Indian Statistical Institute

Mid-Semester Examination: 2010-2011

M.Tech(QR & OR) II

Subject: Reliability II

Maximum Marks: 100 Duration : 3 hours Date: 15 September 2010

Notations used in this paper are the usual notations used in the class.

1. Fill in the following blanks from the given choices :

   (a) The largest class of life distributions among the following is ———————— (*IFR, IFRA, DMRL, NBU, NBUE).*

   (b) If $\pi\prime(h_1, t_1) \overset{LR}{>} \pi(h_2, t_2)$ *then the ratio* $\frac{\pi\prime_n(h_1,t_1)}{\pi_n(h_2,t_2)}$ *is* ———————— *(increasing, decreasing) in* n.

   (c) If F *is IFR then* $\bar{F}(x|t)$ *is* ———————— *(increasing, decreasing) in* t, $x \geq 0$.

   (d) If $\pi\prime(h_1, t_1) \overset{LR}{>} \pi(h_2, t_2)$ *then* $\alpha(\pi\prime(h_1, t_1))$ ——— $(\geq, \leq)\alpha(\pi(h_2, t_2))$.

   (e) If $\pi\prime(h_1, t_1) = (1, 0, 0, \ldots)$ *and* $\pi(h_2, t_2) = (0, 1, 0, \ldots)$ *then* $g(\pi\prime(h_1, t_1))$ ———— $g(\pi(h_2, t_2))(\geq, \leq)$.

   [2 × 5 = 10]

2. If $\pi\prime(h_1, t_1) \overset{LR}{>} \pi(h_2, t_2)$ then show that $g(\pi\prime(h_1, t_1)) \geq g(\pi(h_2, t_2))$.

   [30]

3. Show that a coherent system of iondependent IFRA components itself has an IFRA life distribution.

   [20]

4. Bring out the differences between software and hardware reliabilities.

   [5]

5. State the assumptions in the Jelinski-Moranda model and give your comments on each of the assumptions.

   [10]

P.T.O

6. Define *IFR, DMRL* and *NBUE* classes of life distributions. Show and prove the relationship between these three classes of life distributions.

[13]

7. A new component is to be designed. A stress analysis revealed that the component is subjected to a tensile stress. But there are variations in the load and the tensile stress is found to be normally distributed with mean of 35,000 psi and a standard deviation of 4000 psi. The manufacturing operations create a residual compressive stress that is normally distributed with mean of 10,000 psi and a standard deviation of 1,500 psi. A strength analysis of the component showed that the mean value of the significant strength is 50,000 psi. The variations introduced by various strength factors are not clear at the present time. The engineer wants to know the maximum value of the standard deviation for the strength that will insure that the component reliability does not drop below 0.999.

[12]

INDIAN STATISTICAL INSTITUTE
M. Tech. (QR & OR) 2nd YEAR
Year: 2010
MIDSEMESTER EXAMINATION

Subject: Operations Research-II

Date of Exam: 16.09.2010 **Max. Marks: 100**

Time: 3 hours

Group-A [60 Marks]

Answer any four.

1. State the geometrical aspects for finding the solution of LP and NLP problems.
Suppose A is an $m \times n$ matrix and c is an n vector. Then, exactly one of the following two systems has a solution:

$$System\ 1 \quad Ax < 0 \quad \text{for some } x \in R^n$$
$$System\ 2 \quad A^t y = 0 \text{ and } y \geq 0 \quad \text{for some } y \in R^m.$$

[9 + 6 = 15]

2. Define convex function, pseudo convex function and quasi convex function.
Let $f(x_1, x_2) = 2x_1 + 2x_1^2 - 2x_1x_2 + x_2^2$. Find the Hessian matrix $H(x)$ and show that $H(x) \in$ PD.

[9 + 6 = 15]

3. Define epigraph and sub-gradient of a function.
Let S be a nonempty convex set in R^n and let $f : S \to R$. Then f is convex if and only if *epi* f is a convex set.

[6 + 9 = 15]

4. State the duality theorem.
Let S be a nonempty closed convex set in R^n and $y \notin S$. Then there exist a nonzero vector p and a scalar α such that $p^t y > \alpha$ and $p^t x \leq \alpha$ for each $x \in S$.

[7 + 8 = 15]

5. Suppose that $f : R^n \to R$ is differentiable at $\bar{x}$. Prove that if there is a vector d such that $\nabla f(\bar{x})' d < 0$, then there exists a $\delta > 0$ such that $f(\bar{x} + \lambda d) < f(\bar{x})$ for each $\lambda \in (0, \delta)$.

Suppose that $f : R^n \to R$ is differentiable at $\bar{x}$. Prove that $\nabla f(\bar{x}) = 0$ and $H(\bar{x})$ is positive semi-definite if $\bar{x}$ is a local minimum.

[7 + 8 = 15]

Group-B [40 Marks]

Answer all questions.

1. Consider the following project for which three time estimates are given.

Activity	t_o	t_m	t_p
1-2	1	2	3
2-3	3	4	5
2-4	2	3	10
3-5	2	2	2
4-5	3	4	11
5-6	1	3	11

a) Draw the network.

b) Estimate the most expected project completion time and its variance from the given three time estimates, after finding out critical path.

c) Compute the probability that the project will be completed within 12 days and also within 17 days. [Given: $\Phi(-1.18) = 0.119$; $\Phi(0.79) = 0.785$]

[5+10+5=20]

2. Answer the following questions related to Job Sequencing problems.

a. Write down at least three assumptions to formulating a general Job Shop problem.

b. Prove that a schedule which is optimal with respect to L_{max} is also optimal with respect to T_{max}. Comment on its converse result.

c. Suppose, 5 jobs are to be processed, each of which must go through the 2 machines, A and B in the order A - B. Processing times are given in the table below:

Job	M/c. A	M/c. B
1	5	2
2	1	6
3	9	7
4	3	8
5	10	4

By applying the Johnson's rule, determine a sequence for the 5 jobs that will minimize the elapsed time T and then find out elapsed time T. Identify the component of idle time, if there is any.

[3+7+10=20]

INDIAN STATISTICAL INSTITUTE

Mid-Semester Examination: 2010-11 (First Semester)

Course Name: M-TECH (QROR) II Year.

Subject Name: Software Engineering

Date: 17 - 09 - 2010 *Maximum Marks: 50* *Duration: 2½ hours*

Answer any five questions.

1. Suppose a program for solving a problem costs X, and an industrial level software for solving that problem costs 10X. Where do you think this extra 9X cost is spent? Suggest a possible breakdown of this extra cost with your reasons and justifications?

[5+5=10]

2. What is the relationship between a process model, process specification, and process for a project? What are the key outputs in a development project that follows the prototyping model? Write an ETVX specification for this process model?

[3+3+4= 10]

3. What is SRS? What are functional and non-functional requirements in software engineering? The basic goal of the requirement activity is to get an SRS that has some desirable properties. Explain these properties.

[3+4+3=10]

4. Which is the single attribute of software that allows a program to be intellectually manageable and why? We strive for the lowest possible coupling and high cohesion, while designing the software. Why?

[5+5=10]

5. What is the role of software architecture in a software system? What are the different views of architecture? Describe each of them in detail.

[4+6=10]

6. What is the importance of design in the software engineering? If some existing modules are to be re-used in building a new system, will you use a top-down or bottom-up approach and why?

[5+5=10]

7. Differentiate between the following terms: (a) Milestone and deliverable, (b) Requirements Definition and Specification, and (c) a software product and a software process.

[3+4+3=10]

---XXX---

Note: Books, note books, computers, mobiles etc. are not allowed during examination.

INDIAN STATISTICAL INSTITUTE

M.Tech (QR-OR). II Year

Mid-Semestral Examination

Subject: **Industrial Experimentation**

Date: 20.09.2010 Time: 3 hours Maximum Marks:100

[Answer any part of any question. The maximum mark you can score is 100]

1. A field experiment was performed to study the effects of two levels of nitrogen (0 and 230) and two levels of sulphur (0 and 40) and their interaction on the yield of wheat with a randomized block design. The design and data are shown below:

Block 1			Block 2			Block 3		
Nitrogen	Sulphur	%Yield	Nitrogen	Sulphur	%Yield	Nitrogen	Sulphur	%Yield
0	0	0.7	0	40	0.7	230	40	1.5
230	0	0.8	230	0	1.6	0	0	0.4
230	40	1.8	0	0	0.5	0	40	0.3
0	40	0.5	230	40	1.8	230	0	0.7

(a) Briefly explain how you would construct and randomize a design like the one shown.

(b) Analyse the data and offer your comments.

[5+15=20]

2. (a) Six groups of oysters (3 male groups and 3 female groups) were selected from a clean site and six groups of oysters (3 male groups and 3 female groups) were selected from a polluted site in a tidal creek. (Oyster size is a reasonable proxy for gender). Six oysters from each group were sacrificed and morphological measurements were made--the response was an average morphological measurement from each group of a half-dozen oysters. Afterwards, the six groups of oysters from the clean site were transferred to the polluted site and the six groups of oysters from the polluted site were transferred to the clean site. After six weeks, six more oysters from each group were sacrificed and the measurements were repeated. Can you identify the design? Is there any serious confounding in the experiment? Can both gender and treatment (clean/polluted) effects be tested? How would you test them?

2 (b) A physician wants to compare a new drug A with the current standard drug B for a certain chronic illness. For each of the following types of design, briefly state how a design for comparing the two drugs would be constructed and randomized. In doing so, suppose that each of the designs involves 10 patients in total.

(i) a completely randomized design with equal replication,

(ii) a matched-pairs design,

(iii) a cross-over design with two periods.

[10+15=25]

3. To arrange a 2^6 design in 16 blocks, two blocking schemes are considered. The first scheme has the block generators: $B_1=126$, $B_2=136$, $B_3=346$, $B_4=456$ and the second scheme has the block generators: $B_1=136$, $B_2=1234$, $B_3=3456$, $B_4=12456$. Compare the two schemes in terms of their respective $g_i(b)$ and indicate which blocking scheme is more advantageous.

[15]

4 (a)In a medical experiment, two doses of the standard preparation of gastrin and two of the preparation of unknown potency were required to be administered on rats to test the efficacy of the drug. Obviously it would have been best to give all four doses to each rat (using as many rats as are necessary for the accuracy required) so that the estimate of potency of the unknown preparation is not biased by differences in sensitivity to gastrin between the individual rats. What type of design do you think could have been used with the above arrangement?

4(b) However, Lai (1962) found it impracticable to obtain from each animal responses to more than three of the four treatments. The duration of action of each dose was such that if all four were given to the same animal responses would have had to be so close together that they would interact, or else the assay would have had to be impossibly long. Consequently the fourth dose had to be carried over to a different animal. The treatments were next assigned randomly to the letters. In the present example, A was equivalent to High dose of standard preparation (HS), B was equivalent to Low dose of standard preparation (LS), C was equivalent to High dose of test (unknown) preparation (HT) and D was equivalent to Low dose of test (unknown) preparation (LT). Thus, according to the scheme in the following table, the first rat was given first HS, then LS and finally HT. The second rat was given LT, HS and LS in that order: and so on. Responses were expressed as the mean rate of acid secretion (μ equiv./10 min) (Lai, 1962)

Rats	Order-1	Order-2	Order-3
1	A=2.190	B=0.975	C=1.700
2	D=1.570	A=3.130	B=1.850
3	C=2.570	D=1.680	A=3.000
4	B=1.150	C=2.275	D=0.730

What is the name of the above design? Write the model. Are there any differences among the treatments insofar as the mean rate of acid secretion (by the rats) is concerned?

[5+15=20]

5. Which of the two possible designs (A, B) will you choose in each case and why?

(i) 2^{7-3} experiment with the following two possibilities.

Design A: 2^{7-3} with the defining relation I=1235=1246=12347

Design B: 2^{7-3} with the defining relation I=1235=1246=1347

(ii) 2^{6-2} experiment with the following two possibilities.

Design A: 2^{6-2} with the defining relation I=12345=1236

Design B: 2^{6-2} with the defining relation I=1235=2346

(iii) 2^{9-4} experiment with the following two possibilities.

Design A :2^{9-4} ; design generators: 6=123, 7=124, 8=125, 9=1345

Design B :2^{9-4} ; design generators: 6=123, 7=124, 8=134 , 9=2345

(iv) 2^{8-4} experiment with the following two possibilities.

Design A :2^{8-4} ; design generators: 5=123, 6=124, 7=134, 8=234

Design B :2^{8-4} ; design generators: 5=123, 6=134, 7=145, 8=246

[32]

6. (a)Verify whether the following design is a balanced incomplete block design (BIBD). If so, state why? Write down the parameters of the given BIBD

			Block						
I	II	III	IV	V	VI	VII	VIII	IX	X
0	0	0	0	0	1	1	1	2	2
1	1	2	3	4	2	3	4	3	3
2	3	4	5	5	5	4	5	4	5

(b) Briefly explain how the design in (a) should be randomized.

(c) State advantages and disadvantages of an incomplete block design being *balanced.*

(d) Suppose that six varieties of a crop are to be compared in an experiment. One possibility is to use the design in (a) with ten greenhouses, growing three of the varieties in each greenhouse. Another possibility is to use five larger greenhouses and grow all six varieties in each of them. Assuming that in the first case the error variance is equal to 8, and that in the second case it is equal to 15, which of the two designs would you prefer for estimating the difference of any pair

of varieties and why? What are the names of the two designs and show how would you arrange the experiment?

[5+2+3+10=20]

7. Using an appropriate orthogonal array, give the experimental layouts of the following experiments. Give also the Linear graphs.

(a) One 5- level factor (A), Two 3-level factors (B,C) and one 2-level factor(D). All the main effects are required to be estimated.

(b) Two 2-level factors (A,B), Two 2-level factors C,D such that, C is nested within the low level of A and D is nested within the high level of A, One 2-level factor (E) such that. E is nested within the higher level of B, One 4-level factor (F). The interaction AF is required to be estimated along with the main effects.

[4+6=10]

-x-

INDIAN STATISTICAL INSTITUTE

First-Semester Examination: 2010-11

Course Name: M-TECH (QROR) II Year.

Subject Name: Software Engineering

Date: 03 - 12 - 2010 ***Maximum Marks: 60*** ***Duration: 3 hours***

Answer any five questions.

1. What is UML? Explain the significance of UML. Why do we need to build model before building the actual system? Explain any two relevant metrics for OO development paradigm.

[1+3+4+4=12]

2. Explain software reliability. How is it different from Hardware reliability? Describe four metrics that help in measuring the software reliability.

[3+3+6=12]

3. Write notes on the following (any three):
   a. Component based Software engineering
   b. Cleanroom Software Engineering
   c. Reengineering and Reverse-Engineering
   d. Capability Maturity Model (CMM)
   e. Change Management (CM)

[4x3=12]

4. List the scenarios for order processing and billing in a restaurant. Create sequence diagram and object diagram for this problem.

[2+5+5=12]

5. What is structured programming and how does it help improve code quality? What are the major concepts that help to make a program more readable?

[6+6=12]

6. Discuss the difference between black-box testing and structural testing. Suggest how these may be used together in the defect removal process.

[6+6=12]

7. Describe the following OO concepts with examples:
   a. Association and Aggregation
   b. Generalization and Specialization
   c. Inheritance and Polymorphism

[4x3=12]

8. Explain software quality? Define at least eight software quality attributes briefly.

[4+8=12]

---XXX---

Note: Books, note books, computers, mobiles etc. are not allowed during examination.

INDIAN STATISTICAL INSTITUTE

Semestral Examination: (2010 - 2011)

Course Name: M. Tech. (QROR) **Year: 2nd year**

Subject Name: Database Management Systems

Date: December 6, 2010 Maximum Marks: 100 Duration: 3 hrs

Answer as many questions as possible.

a) Describe various constraints that need to be satisfied while updating a relational database. 10

b) Describe various mapping cardinalities, with suitable real life examples, between two entity sets. 10

a) Elaborate in details what a safe tuple relational calculus expression means. 10

b) Consider a binary relation between two schemas: customer (cust_ID, cust_name) and account (cust_ID, account_no, account_balance), the names of the schemas and the attributes being self-explanatory. Write the expressions that result in customer names, and their account numbers and corrsponding account balance of all the customers, in

i) relational algebra

ii) tuple relational calculus (in reduced form), and

iii) domain relational calculus (in reduced form). 2+5+3 = 10

Consider a relation schema $R = (A, B, C, D, E)$ with the set F of functional dependencies $F = \{A \rightarrow BC, CD \rightarrow E, B \rightarrow D, E \rightarrow A\}$. Check whether R is in BCNF or not with respect to F. If not, decompose R into two or more relation schemas so that the resulting schemas are in BCNF. Explain whether this decomposition is dependency preserving or not. 2+6+2 = 10

Consider the proposed rule for functional dependencies: If $\alpha \rightarrow \beta$ and $\gamma \rightarrow \beta$, then $\alpha \rightarrow \gamma$. Prove that this rule is not sound by showing a real life relation that satisfies $\alpha \rightarrow \beta$ and $\gamma \rightarrow \beta$, but does not satisfy $\alpha \rightarrow \gamma$. 4

Define partial dependency. Show that every partial dependency is a transitive dependency. 2+4 = 6

P. T. O.

a) Consider a relational schema: book (ISBN, title, publisher, publisher_address) with the set functional dependencies as {ISBN → title, ISBN → publisher, publisher → publisher_addres Determine the normal form which the schema is in, with necessary justification.

b) Consider a relational schema: order (order _no, product_ID, product_name) with the set functional dependency as {product_ID → product_name}. Determine the normal form which schema is in, with proper justification.

Consider a database of Indian Statistical Institute with the following schemas. The names of schemas and the attributes are self explanatory.

i) employee (roll_number, fname, mint, lname, PAN, bdate, address, gender, basic_ super_PAN, dname), with the primary key as PAN, and unique roll_number;

ii) department (dname, officer_PAN, officerstartdate), with the primary key as dname;

iii) project (pname, pnumber, dname) with the primary key as pnumber, and unique pname;

iv) works_on (roll_number, employee_PAN, pnumber, hours_worked), with the primary ke (employee_PAN, pnumber);

v) dependent (roll_number, employee_PAN, dependent_name, gender, bdate, relationship), the primary key as (employee_PAN, dependent_name).

Consider the query, "Find the last names of the employees born after 1960 who work o project named *Soft Computing*." Assume that "date1 > date2", "date1 = date2" and "date date2" indicate "date1 is after date2", "date1 is identical to date2" and "date1 is before dat respectively.

a) Write the relational algebra expression, in standard form, for the query.

b) Hence write the corresponding expression in SQL.

c) Draw the query tree for the relational algebra expression.

d) Modify the query tree in order to minimize the execution time.

a) Define multivalued dependency with a suitable real life example.

b) Define 4th normal form (4NF).

c) Give an example of a relation schema R and a set of dependencies such that R is in BCNF, but i in 4NF.

Indian Statistical Institute

Semester Examination: 2010-2011

M.Tech(QR & OR) II

Subject: Reliability II

Maximum Marks: 100 Duration : 3 hours Date: 8 December 2010

ote: Answer as many questions as possible. Maximum you can score is only 100.

1. (a) Define availability, limiting average availability. When availability is equal to reliability?

   (b) If limiting availability exists, prove that limiting average availability is equal to limiting availability.

   (c) Define warranty, free replacement warranty and pro rata warranty.

   (d) Comment on preventive and corrective maintenances.

   [3+5+5+4]=[17]

2. (a) Prove that $A_{av} = lim_{t\to\infty}\frac{E(U(t))}{t} = (1 + \sum_{j=1}^{n}\frac{v_j}{\mu_j})^{-1}$, symbols have usual

   (b) Consider a system of 4 components (as given in the following table)in series under the assumptions that no 2 components can fail together and if one failed component is repaired others remain in suspended animation. Find:

      i. long run fraction of time the system is operating.

      ii. long run fraction of time the system is down due to failure of each of 4 components.

      iii. long run average number of failures per hour of each of 4 components.

      iv. long run average of system up times.

      v. long run average of system down times.

Mean times to failure and repair of components of the system

Components	Mean Life(hrs.)	Mean Repair time(hrs.)
1	100	0.2
2	200	0.4
3	2000	2.0
4	20,000	40.0

[8+10]=[18]

3. (a) Write down the assumptions that are generally used in developing ALT models.

   (b) Silicon carbide is used as an optional material for manufacturing semiconductor devices, especially for those devices operating under high temperatures and high electric fields conditions.An extensive accelerated life experiment is conducted by subjecting 6H-SiC metal-oxide-silicon capacitors to temperatures of $140, 240,$ and $300°C$. The recorded data follow exponential distribution at all the temperatures. The mean failure times at different temperatures are $8234, 4425$ and 1100 respectively. Determine the mean time to failure of the capacitors at $30°C$.

[5+10]=[15]

4. Suppose X and Y denote respectively the strength and stress variable.

   (a) Consider the case where a component's stress and strength both follow Weibull distribution. Derive an expression for reliability of the component.

   (b) Derive an expression for reliability of the component whose stress and strength both follow Gamma distribution with scale parameters not equal to 1.

[10+ 10]= [20]

5. (a) What is the difference between the Cox model and the extended Cox model?

   (b) Why the Cox model is called a semiparametric model? Also why the Cox likelihood is called a "'partial"' likelihood?

   (c) In a 10-year follow-up study conducted in the US, involving persons 60 years or older, for evaluating the relationship of social support to mortality status. A Cox proportional hazards model was used to describe the relationship of a measure of social network to time until death. The social network index

was denoted as SNI and took on integer values between 0 (poor social network) to 5 (excellent social network). Variables to be considered for control in the analysis as either potential confounders or potential effect modifiers were AGE(treated continuously), RACE(0,1), and SEX(0,1).

i. State an initial PH model that can be used to assess the relationship of interest, which considers the potential confounding and interaction effects of the AGE,RACE, and SEX (assume no higher than two factor products involving SNI with AGE,RACE, and SEX).

ii. For the model developed above, give an expression for the hazard ratio that compares a person with SNI =4 to a person with SNI= 2 and the same values of the covariates being controlled.

iii. Describe how would you test for interaction using the above model. In particular, state the null hypothesis, the general form of the test statistic, with its distribution and degrees of freedom under the null hypothesis.

[3+2+[3+5+6]] = [19]

6. (a) Describe and formulate Effort Minimization Algorithm along with the assumptions. Give a solution to the optimization problem. Is this solution unique?

(b) A system consists of four subsystems that must function if the system has to function properly. The system reliability goal is 0.950. All the four subsystems have identical reliability improvement effort functions. The estimated subsystem reliabilities at the present time are 0.75, 0.85, 0.90, 0.95. What reliability goal should be apportioned to the subsystems so as to minimize the total effort spent on the system improvement?

[(6+3+1)+10]= [20]

INDIAN STATISTICAL INSTITUTE

M.Tech (QR & OR) 2nd YEAR

Academic Session: 2010

SEMESTER EXAMINATION

Subject: Operations Research-II

Date of Exam: 10-12-2010 **Time: 10:30-13:30** **Maximum marks: 100**

Use separate answer scripts for Group-A and Group-B.

Group – A [50 marks]

Answer question (5) and any three from (1) to (4).

1. a) Let $f: R^n \to R$, $g_i: R^n \to R$ for $i=1,\ldots,m$. Consider the problem to minimize $f(x)$ subject to $g_i(x) \le 0$ for $i=1,\ldots,m$. Let $\bar{x}$ be a feasible solution and suppose that f and g_i are differentiable at $\bar{x}$. If $\bar{x}$ locally solves the problem, then prove that there exist scalars u_i, $i=1,2,\ldots,m$ such that

$$\nabla f(\bar{x}) + \sum_{i=1}^{n} u_i \nabla g_i(\bar{x}) = 0$$
$$u_i g_i(\bar{x}) = 0 \quad \text{for } i = 1, \ldots, m$$
$$0 \neq u \geq 0$$

b) State the primal feasibility, dual feasibility and complementary slackness condition of the problem in (a).

[6+6=12]

2. a) Characterize the stationary point of a nonlinear programming problem in connection with the cone of feasible direction and the cone of descent direction.

b) State the KKT sufficient conditions of optimality for a nonlinear programming problem.

[5+7=12]

3. a) Formulate quadratic programming as a linear complementarity problem LCP (q, M). State the formulation in case of linear programming problem.

b) State the Lemke's algorithm mentioning the termination criteria.

[6+6=12]

4. a) Define positive semidefinite matrix and copositive matrix.

Consider the matrix $A = \begin{bmatrix} 4 & 1 & 2 & 5 \\ 2 & 1 & 4 & 1 \\ 4 & 1 & 0 & -1 \\ 1 & 8 & 2 & 1 \end{bmatrix}$

Is the matrix A copositive-plus?

b) Suggest a method to solve a linear fractional programming problem as linear programming problem.

[6+6=12]

5. State an additive type model for dynamic programming problem.

The sales manager for a publisher of college text books has four sale persons to assign to three different regions of the country. Determine how many sales persons should be assigned to the respective regions in order to maximize sale using dynamic programming problem. The following table gives the estimated increase in sales in each region:

Salesman	Estimated increase in sales (in appropriate unit) for three regions		
	1	2	3
1	35	21	28
2	48	42	41
3	70	56	63
4	89	70	75

[4+10=14]

Group-B [60 Marks]

Answer Q.1, Q.2 and any two questions. Maximum you can score 50.

1. Under which conditions an n jobs, 3 machines problem is solvable? Solve the following 2 jobs, 4 machines problem.

	Machines			
Jobs	A	B	C	D
1	2	4	5	1
2	2	5	3	6

Technological ordering of Job-1 is A-B-C-D and of Job-2 is D-B-A-C.

i) Derive the waiting time of each job on each of the four machines.
ii) Derive the simultaneous processing times of both the jobs independently on the machines.
iii) Derive the time spent on Job-1 and Job-2 only when there is no simultaneous processing time
iv) Estimate the minimum makespan schedule.

[15]

2. Define the matrix – Unimodular and Totally Unimodular. What is the necessary condition for a matrix to be totally unimodular? Prove that, if a matrix A is totally unimodular, then every basic solution of the LP: $Ax = b$, $x \geq 0$, where A and b are assumed to be integer, is integer.

[4+2+9=15]

3. Solve the following integer programming problem by branch and bound method. Draw the tree diagram and mention the reasons for fathoming a live node in the tree.

Maximize $Z = 4x_1 + 3x_2$

Subject to

$5x_1 + 3x_2 \geq 30,$

$x_1 \leq 4,\ x_2 \leq 6,$

$x_1 \geq 0,\ x_2 \geq 0,$ integer

[15]

4. Define separation and branching of Branch-and-Bound technique. Explain the B-B algorithm for calculation of upper and lower bounds for an IP problem

[3+3+9=15]

5. Derive Gomory's cut for an ILP problem. Generate the initial two cuts for the following ILP problem.

$$\text{Maximize } 7x_1 + 9x_2$$
$$\text{Subject to}$$
$$-x_1 + 3x_2 \leq 6$$
$$7x_1 + x_2 \leq 35$$
$$x_1, x_2 \geq 0, \text{ integer}$$

given below the optimal tableau of the corresponding LP.

Variable	x_1	x_2	x_3	x_4	Soln. (X_B)
x_2	0	1	7/22	1/22	7/2
x_1	1	0	-1/22	3/22	9/2

[15]

INDIAN STATISTICAL INSTITUTE
M. Tech. (QR OR); II Year (2010-2011)
Semester: III
Subject: Advance Statistical Methods

Date: 13.12.2010 Maximum Marks: 100 Duration: 3½ hours

Note: Answer Group A and Group B in separate Answer sheets.

Group: A – 60 Marks but the maximum you can score is 50.

1) Perspiration from 20 females was observed. Three components were measured. They were X_1: Sweat rate, X_2: Sodium content and X_3: Potassium content. The Summary of the data are given below:

$$\bar{X} = \begin{bmatrix} 4.640 \\ 45.400 \\ 9.965 \end{bmatrix} \quad \text{and} \quad S^{-1} = \begin{bmatrix} 0.586 & -0.022 & 0.258 \\ -0.022 & 0.006 & -0.002 \\ 0.258 & -0.002 & 0.402 \end{bmatrix}$$

The scientist wants to test the hypothesis if the mean vector $\mu = (4, 50, 10)'$

i) Write down the test statistic you propose to use. What distributions does it follow?

ii) Write down the necessary assumptions.

iii) Test the hypothesis?

iv) Would you prefer univariate tests instead? Justify your answer.

$(2 + 2 + 6 + 2) = [12]$

2) (a) Write down the multiple linear regression model in matrix form

(b) State the underlying assumptions

(c) Write down the normal equations

(a) Show that $\hat{\beta}$ is unbiased for β

(d) Derive the expression for $V(\hat{\beta})$

The symbols used above have their usual meaning

[2+2+2+3+1 = 10]

P.T.O.

) In a study to assess the effects of Solder-Bath Temperature (SBT) and Wave Height (WH) on generation of defects in PCBs a 3^2 design was run. Three PCBs were soldered at each factor combination. Two types of defect were observed. Some intermediate computations are furnished below. Notations have their usual meaning.

i) Write down the underlying model and the associated assumptions.

ii) Write down the hypotheses that are to be tested.

iii) Test the hypotheses.

iv) Carry out univariate analyses.

v) Draw conclusions.

Error Sums of Squares and Cross Products

Variable	Defect 1	Defect 2
Defect 1	525.3333	438.33
Defect 2	438.33	11710.67

Between Effects Sums of Squares and Cross Products Matrices

	Variable	Defect 1	Defect 2
SBT	Defect 1	246.2963	900.19
	Defect 2	900.19	5222.74
WH	Defect 1	2.0741	26.96
	Defect 2	26.96	55330.07
SBT*WH	Defect 1	49.4815	-89.52
	Defect 2	-89.52	9385.93

[4 + 3 + 6 + 6 + 3 = 22]

4) An experiment was conducted to study the relationship of yield of a synthetic analogue to jojoba oil (y) in terms of

x_1: reaction temperature,

x_2: Initial amount of catalyst and

x_3: Pressure.

Following table gives the data.

Expt No.	x_1	x_2	x_3	y
1.	-1	-1	-1	17
2.	1	-1	-1	44
3.	-1	1	-1	19
4.	1	1	-1	46
5.	-1	-1	1	7
6.	1	-1	1	55
7.	-1	1	1	15
8.	1	1	1	41
9.	0	0	0	29
10.	0	0	0	28
11.	0	0	0	30
12.	0	0	0	27
13.	0	0	0	28

a) Fit a multiple linear regression model of y on x_1, x_2 and x_3 and test for lack of fit.

b) Is it possible to get the contribution of each of the regressors unconditionally on the others? Explain your answers?

[12 + 4 = 16]

GROUP-B

Full marks: 50

1. Assignments [20]

2. Write True or False [1X5=5]

   i) Single linkage method & complete linkage method lead to the same cluster.
   ii) Prior Knowledge about the distribution of the variables is a prerequisite for clustering.
   iii) Principal components do not always lead to meaningful interpretation
   iv) One can go back to the original variable in the case of principal component regression.
   v) A discriminant function always minimizes the distance between pairs of observations.

3. Obtain the Fisher's discriminant function for several populations [7]

4. i) Discuss the differences between Hierarchical clustering and Partitional algorithms [5+5]
   ii) What are Scree plots and Dendograms?

5. What is a factor score? Show how such scores can be estimated. [8]

INDIAN STATISTICAL INSTITUTE

Mid Semestral Examination: (2010 - 2011)

Course Name: M.Tech (QROR) **Year: 2nd year**

Subject Name: Database Management Systems

Date: September 14, 2010 Maximum Marks: 50 Duration: 2 hrs

Answer as many questions as possible.

1. Describe in details the advantages of using the database approach over the file processing approach. 10

2. Define the following terms with suitable examples.
   a) Superkey, candidate key, primary key, foreign key, partial key 5
   b) DDL, DML, data dictionary 5
   c) Composite attribute, multivalued attribute, derived attribute 5

3. Consider a part of bank database. There are the following entity sets:
   a) BANK is described by the attributes code, name and address.
   b) Each bank has several branches. Thus BRANCH entity set is described by the attributes address and branch-number.
   c) Each branch has customers. CUSTOMER entity set is described by customer-id, name, address, PAN and phone-number.
   d) A customer may have a loan account and/or a deposit account. LOAN-AC is described by the attributes loan-number, amount and loan-type (i.e., personal, car, house-building, loan-against-property). DEPOSIT-AC is described by account-number, balance-amount and type (i.e., savings, current, fixed-deposit)

   Design an E-R diagram illustrating the relationships among these entity sets. Show the keys and the mapping cardinalities properly along with your assumptions. 30

INDIAN STATISTICAL INSTITUTE

M.Tech (QR & OR), 2010-2011

II Year, Semester-III Examination

Subject: **Industrial Experimentation**

Date: 15.12.2010 Time: 3 hours Maximum Marks:100

[Answer any part of any question. The maximum mark you can score is 100

1. A 3-component mixture experiment is subject to the following constraints for the mixture components that define the two regions of experimentation.

(i) $0.15 \le x_1$, $0.25 \le x_2$, $0.10 \le x_3$ and $x_1 + x_2 + x_3 = 1$

(ii) $x_1 \le 0.5$, $x_2 \le 0.6$, $x_3 \le 0.6$ and $x_1 + x_2 + x_3 = 1$

Give two sets of reasonable experimental runs for fitting second order mixture models for the response characteristic.

[20]

2. The region of experimentation of a 4-component mixture system is defined by the following constraints for the four mixture components.

$0.3 \le x_1 \le 0.6$, $0.1 \le x_2 \le 0.5$, $0.4 \le x_2 \le 0.7$, $0.05 \le x_2 \le 0.2$ and $x_1 + x_2 + x_3 + x_4 = 1$

Apply XVERT algorithm to identify a set of candidate points that would be considered as the selected design by the experimenter for fitting a quadratic model to the data on response characteristic.

[12]

3. A thickening agent is added to a liquid soap product to control the viscosity. The agent usually makes up 10% of the soap by weight. There are three different salt compounds that are blended to make up the thickener. The product formulators have decided to investigate the combination of these ingredients taking recourse to mixture experiment. Let x_1^*, x_2^*, x_3^* represent the actual

P. T. O.

proportions of the three salt compounds with $x_1^* + x_2^* + x_3^* = 0.1$. It is found that, if $x_i = \frac{x_i^*}{0.1}$, i=1,2,3 then the design points can be expressed as proportions $0 \le x_i \le 1$ with the standard mixture constraint $x_1 + x_2 + x_3 = 1$. The experimenter decided to use an augmented simplex-lattice design to study the viscosity of the thickener. The design and the viscosity data are shown in the following table. The centroid point was replicated three times.

Design Points	Component Proportions			Viscosity
	x_1	x_2	x_2	
1	1	0	0	410
2	0	1	0	880
3	0	0	1	1445
4	1/2	1/2	0	683
5	1/2	0	1/2	465
6	0	1/2	1/2	456
7	1/3	1/3	1/3	465
8	1/3	1/3	1/3	392
9	1/3	1/3	1/3	428

(i) Can you fit an adequate quadratic model for viscosity of the thickening agent as function of the three salt compounds of the soap mixture ?

(ii) Suppose that a desirable value for viscosity of this thickening agent is 900 centipoise. Is there a formulation that would produce this value of viscosity ?

[10+5=15]

4.(a) Prove that the model matrix of a second order central composite design (CCD) without centre points is singular if $\alpha = \sqrt{k}$ where k is the number of design variables.

(b) For a CCD with k=2, $I = x_1 x_2$ (x_1, x_2 are design variables) for the factorial part, can you choose a proper α value in order to make the design rotatable?

(c) A first order orthogonal design with all points at ±1 extremes remains rotatable for a first order model if all the design points are duplicated the same number of times. Do you agree? Justify.

(d) An experimenter has used a first order response surface design that was a quarter fraction of a 2^5 design with the defining relations: $x_1x_2x_3=I$ and $x_3x_4x_5=I$. The design factors are all quantitative. Is the design orthogonal for estimating main effects? Is the design variance-optimal?

(e) If 4 centre points are added to the above design, will the resulting design remain variance-optimal for estimating main effects ?

(f) Can you consider the above first order variance optimal design a first order rotatable design or, a D-optimal design? Explain.

(g) How a stationary ridge system is different from a rising or, falling ridge system? Explain the conditions of getting these two types of ridge systems in response surface. Which canonical form of the second order response surface model is more helpful to describe the presence of ridge system?

[4+3+3+5+3+2+5=25]

5. (a) Give a 'face centre cube' design in three variables. Examine whether the design is orthogonal and/or, rotatable. How does this design compare with a Box-Behnken design in three variables w. r. t. orthogonality, rotatability and any other criteria you think appropriate?

(b) Do you think that the following design blocks orthogonally its two blocks as shown below?

Block	x_1	x_2
I	-1	-1
	-1	1
	1	-1
	1	1
	0	0
	0	0

	0	0
	0	0
II	-1	0
	-1	0
	1	0
	1	0
	0	-1
	0	-1
	0	1
	0	1

[12+3=15]

6. Consider the following first order model

$\hat{y} = 15.96 + 1.02x_1 + 3.4x_2 - 2.4x_3$ where, x_i is a coded variable such that $-1 \le x_i \le 1$ and the natural variables are:

Pressure: 24-30 Kg/sq. cm (x_1), Time: 1-2 minutes (x_2) and Caustic Soda Amount 1-2 kg. (x_3).

(i) If the step size in terms of time is 0.5 minutes, find the coordinates of 10 points on the path of steepest ascent.

(ii) Find a unit vector that defines the path of ascent.

[9+3=12]

7. Set up a minimum-bias rotatable design in terms of independent mixture related variables for the three mixture components (x_1, x_2, x_3) having the following constraints with a view to fitting an adequate first order model.

$0.3 \le x_1 \le 0.6$, $0.1 \le x_2 \le 0.5$, $0.4 \le x_2 \le 0.7$; $x_1 + x_2 + x_3 = 1$

[25]

P. T. O.

8.(a) Suppose a crossed array of the type 2_{III}^{6-3} X 2_{III}^{3-1} is proposed to be used for an experiment. However, you wish to consider as many two-factor interactions among the control variables as possible. Your main objective is to fit an appropriate response surface model involving both control and noise variables. Suggest an alternative design.

(b) Consider a situation with three control variables and three noise variables. In order to construct an appropriate response model, the design should contain linear main effects, two-factor interactions among the control variables and two-factor interactions between control and noise variables. What is the minimum number of valid design points required?

[8+7=15]