Indian Statistical Institute
Mid-semester Question paper on Reliability II
M.Tech (QR & OR).

Please answer all questions. Marks allotted to each question is given in [ ].

Maximum Marks = 60 Time = 2 hrs. Date : 3 September 2018.

1. Let the prior distribution of the number of undetected bugs be denoted by π (h,t) with the history 'h', which is known and 't' denotes the number of test cases tested. Let $g_k(\pi(h,t))$, k≥1, denotes the maximum expected gain after k test cases are tested. Assume that a reward is obtained for identifying one bug and debugging it immediatelyand let it be 1 unit, whereas the cost of testing each test case is 'c' unit, where c<1. Assume that $g_o(\pi(h,t))=0$. Show that $g_k(\pi(h,t))$ is bounded above by $\mu_{\pi(h,t)}$, the finite mean of the prior distribution π (h,t). [Note that the proofs of all other results that may be needed to prove the given result also need to be proved]

[20]

2. a) Define likelihood ratio ordering of probability distributions.
   b) If $\pi'(h_1,t_1)$ is preferred to $\pi(h_2,t_2)$ in the likelihood ratio sense then show that $\sum_{n\geq k}\pi'_n(h_1,t_1) \geq \sum_{n\geq k}\pi_n(h_2,t_2)$ for all k≥1.
   c) Using (b) above, show that the probability of success under $\pi'(h_1,t_1)$ is greater than or equal to the probability of success under $\pi(h_2,t_2)$.
   d) Show that for any π (h,t), π (h,t) is preferred to π(hf,t+1) in the likelihood ratio sense.
   e) Also if $\pi'(h_1, t_1)$ is preferred to $\pi(h_2, t_2)$ in the likelihood ratio sense then show that $\pi'(h_1s,t_1+1)$ is preferred to $\pi(h_2s,t_2+1)$ in the likelihood ratio sense and $\pi'(h_1f,t_1+1)$ is preferred to $\pi(h_2f,t_2+1)$ in the likelihood ratio sense.

[5+10+5+5+10=35]

3. Explain what is meant by discrete framework in the software reliability field and why is it important? [5]

INDIAN STATISTICAL INSTITUTE

Mid-Semester Examination: 2018-19

Course Name: M. Tech. (QR & OR) II

Subject Name: Advanced Statistical Methods

Date of Examination: 04.09.2018 Maximum Marks: 70 Duration: 2 hours

Note:
1. This paper carries 80 marks. Answer all questions but the maximum you can score is 70.
2. All notations have their usual meanings

---

1. Distinguish between *dependence* and *interdependence* techniques. Give examples. [5]

2. Suppose **y** is $N3(\mu, \Sigma)$, where $\mu = \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix}$, $\Sigma = \begin{bmatrix} 5 & 2 & -6 \\ 2 & 8 & 1 \\ -6 & 1 & 3 \end{bmatrix}$

   Find the

   a) Distribution of $z = 2y_1 - y_2 + 3y_3$.
   b) Joint distribution of $z_1 = y_1 + y_2 + y_3$ and $z_2 = y_1 - y_2 + 2y_3$.
   c) Distribution of y_2.
   d) Joint distribution of y_1 and y_3.
   e) Joint distribution of y_1, y_3, and $\frac{1}{2}(y_1 + y_2)$.
   Justify your answers in (a) to (e)

   [1½ + 3 + ½ + 1 + 3 = 9]

3. Assume y and x are sub vectors, each 2×1, where $\begin{pmatrix} x \\ y \end{pmatrix}$ is $N_4(\mu, \Sigma)$ with

$$\mu = \begin{bmatrix} 1 \\ 3 \\ 5 \\ 2 \end{bmatrix}, \Sigma = \begin{bmatrix} 3 & 0 & 0 & 4 \\ 0 & 7 & -2 & 0 \\ 0 & -2 & 1 & 0 \\ 4 & 0 & 0 & 9 \end{bmatrix}$$

   i) Find $E(\mathbf{y}|\mathbf{x})$
   ii) Find $\mathrm{Cov}(\mathbf{y}|\mathbf{x})$ [3 + 3 = 6]

4. Consider the following problems:

a) In a steel melting shop there are two furnaces. The source of raw material is the same for both the furnaces. Each day a single heat is produced and from each heat of steel a single sample is collected and chemical analysis is carried out. Percentages of four elements are recorded. A team of engineers has hypothesized that two furnaces produce steel of different chemical compositions.

b) Two batches of steel components have arrived. Physical properties (three in number) are measured taking samples from each batch. The manager wants to know if the physical properties are the same for both the batches

Answer the following questions in each of case a) and b) above

i) Suggest the data collection plan that would facilitate testing the stated hypothesis
ii) Write down the hypothesis in each case
iii) Write down the necessary assumptions
iv) Write down the test statistic you propose to use in each case.
v) What distributions do they follow?
vi) What should be the minimum sample size in each case?
vii) What are the test criteria?
viii) Would you prefer univariate tests instead? Justify

[3 + 2 + 3 + 2 + 2 + 3 + 2 + 3 = 20]

5. In a study involving 2 variables four groups with 50 cases in each group were studied.

The sample means for the four groups were $\bar{x}_1 = \begin{pmatrix} 10 \\ 20 \end{pmatrix}$, $\bar{x}_2 = \begin{pmatrix} 15 \\ 20 \end{pmatrix}$, $\bar{x}_3 = \begin{pmatrix} 20 \\ 25 \end{pmatrix}$, $\bar{x}_4 = \begin{pmatrix} 25 \\ 40 \end{pmatrix}$

The sample covariance matrices for each group were,

$$S_1 = \begin{bmatrix} 4 & 2 \\ 2 & 5 \end{bmatrix}, \quad S_2 = \begin{bmatrix} 6 & 2 \\ 2 & 4 \end{bmatrix}, \quad S_3 = \begin{bmatrix} 6 & 3 \\ 3 & 9 \end{bmatrix}, \quad S_4 = \begin{bmatrix} 12 & 2 \\ 2 & 9 \end{bmatrix}$$

The experimenter wished to test the hypothesis □□□□μ□□□□μ□□□□□μ□□□μ□□□about the means of these variables:

Answer the following questions:

(a) Write down the underlying model and assumptions
(b) What is the value of the Wilk's lambda?
(c) What is the critical value of the significance test at the α level of .05?
(d) What is the conclusion regarding the null hypothesis at the α level of .05?
(e) Set up ANOVA for individual variables if required.
(f) What are your overall conclusions?

[4 + 4 + 2 + 1 + 1 + 4 + 2 = 18]

6. In a study to assess the effects of Solder-Bath Temperature (SBT) at 3 levels and Conveyor Speed (CS) also at 3 levels on generation of defects in Printed Circuit Boards (PCBs), a set of experiments with all combinations of SBT and CS were run. Three PCBs were soldered at each factor combinations. Two defect types D1 and D2 were observed. Some intermediate computations are furnished below. Notations have their usual meaning

a) Write down the underlying model and the associated assumptions.
b) Write down the hypotheses that are to be tested.
c) Test the hypotheses.
d) Carry out univariate analyses if required.
e) Draw conclusions.

$$E = \begin{bmatrix} 526.67 & -151.33 \\ -151.33 & 4636.67 \end{bmatrix}, \quad H_{SBT} = \begin{bmatrix} 96.52 & 295.37 \\ 295.37 & 2716.96 \end{bmatrix},$$

$$H_{CS} = \begin{bmatrix} 6.741 & -343.52 \\ -343.52 & 26787.19 \end{bmatrix}, \quad H_{SBT \times CS} = \begin{bmatrix} 117.48 & 34.85 \\ 34.85 & 3251.70 \end{bmatrix},$$

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

INDIAN STATISTICAL INSTITUTE

Mid-Semester Examination: 201

Course Name: M. Tech. (QR & OR) II

Subject Name: Advanced Multivariate Analy

Mid Sem Q.P.

Date of Examination: 04.09.2018 Maximum Marks: 70 Duration: 2 hours

Note:
1. This paper carries 80 marks. Answer all questions but the maximum you can score is 70.
2. All notations have their usual meanings

---

1. Distinguish between *dependence* and *interdependence* techniques. Give examples. [5]

2. Suppose $\mathbf{y}$ is $N3(\mu, \Sigma)$, where $\mu = \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix}$, $\Sigma = \begin{bmatrix} 5 & 2 & 0 \\ 2 & 8 & 1 \\ 0 & 1 & 3 \end{bmatrix}$

   Find the

   a) Distribution of $z = 2y_1 - y_2 + 3y_3$.
   b) Joint distribution of $z_1 = y_1 + y_2 + y_3$ and $z_2 = y_1 - y_2 + 2y_3$.
   c) Distribution of y_2.
   d) Joint distribution of y_1 and y_3.
   e) Joint distribution of y_1, y_3, and $\frac{1}{2}(y_1 + y_2)$.

   Justify your answers in (a) to (e)

   [1½ + 3 + ½ + 1 + 3 = 9]

3. Assume y and x are sub vectors, each 2×1, where $\begin{pmatrix} x \\ y \end{pmatrix}$ is $N_4(\mu, \Sigma)$ with

$$\mu = \begin{bmatrix} 1 \\ 3 \\ 5 \\ 2 \end{bmatrix}, \Sigma = \begin{bmatrix} 3 & 0 & 0 & 4 \\ 0 & 7 & -2 & 0 \\ 0 & -2 & 1 & 0 \\ 4 & 0 & 0 & 9 \end{bmatrix}$$

   i) Find $E(\mathbf{y}|\mathbf{x})$
   ii) Find $\mathrm{Cov}(\mathbf{y}|\mathbf{x})$ [3 + 3 = 6]

4. Consider the following problems:

a) In a steel melting shop there are two furnaces. The source of raw material is the same for both the furnaces. Each day a single heat is produced and from each heat of steel a single sample is collected and chemical analysis is carried out. Percentages of four elements are recorded. A team of engineers has hypothesized that two furnaces produce steel of different chemical compositions.

b) Two batches of steel components have arrived. Physical properties (three in number) are measured taking samples from each batch. The manager wants to know if the physical properties are the same for both the batches

Answer the following questions in each of case a) and b) above

i) Suggest the data collection plan that would facilitate testing the stated hypothesis
ii) Write down the hypothesis in each case
iii) Write down the necessary assumptions
iv) Write down the test statistic you propose to use in each case.
v) What distributions do they follow?
vi) What should be the minimum sample size in each case?
vii) What are the test criteria?
viii) Would you prefer univariate tests instead? Justify

[3 + 2 + 3 + 2 + 2 + 3 + 2 + 3 = 20]

5. In a study involving 2 variables four groups with 50 cases in each group were studied.

The sample means for the four groups were $\bar{x}_1 = \begin{pmatrix} 10 \\ 20 \end{pmatrix}$, $\bar{x}_2 = \begin{pmatrix} 15 \\ 20 \end{pmatrix}$, $\bar{x}_3 = \begin{pmatrix} 20 \\ 25 \end{pmatrix}$, $\bar{x}_4 = \begin{pmatrix} 25 \\ 40 \end{pmatrix}$

The sample covariance matrices for each group were,

$$S_1 = \begin{bmatrix} 4 & 2 \\ 2 & 5 \end{bmatrix}, \quad S_2 = \begin{bmatrix} 6 & 2 \\ 2 & 4 \end{bmatrix}, \quad S_3 = \begin{bmatrix} 6 & 3 \\ 3 & 9 \end{bmatrix}, \quad S_4 = \begin{bmatrix} 12 & 2 \\ 2 & 9 \end{bmatrix}$$

The experimenter wished to test the hypothesis $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$ about the means of these variables:

Answer the following questions:

(a) Write down the underlying model and assumptions
(b) What is the value of the Wilk's lambda?
(c) What is the critical value of the significance test at the α level of .05?
(d) What is the conclusion regarding the null hypothesis at the α level of .05?
(e) Set up ANOVA for individual variables if required.
(f) What are your overall conclusions?

[4 + 4 + 2 + 1 + 1 + 4 + 2 = 18]

6. In a study to assess the effects of Solder-Bath Temperature (SBT) at 3 levels and Conveyor Speed (CS) also at 3 levels on generation of defects in Printed Circuit Boards (PCBs), a set of experiments with all combinations of SBT and CS were run. Three PCBs were soldered at each factor combinations. Two defect types D1 and D2 were observed. Some intermediate computations are furnished below. Notations have their usual meaning

a) Write down the underlying model and the associated assumptions.
b) Write down the hypotheses that are to be tested.
c) Test the hypotheses.
d) Carry out univariate analyses if required.
e) Draw conclusions.

$$E = \begin{bmatrix} 526.67 & -151.33 \\ -151.33 & 4636.67 \end{bmatrix}, \quad H_{SBT} = \begin{bmatrix} 96.52 & 295.37 \\ 295.37 & 2716.96 \end{bmatrix},$$

$$H_{CS} = \begin{bmatrix} 6.741 & -343.52 \\ -343.52 & 26787.19 \end{bmatrix}, \quad H_{SBT \times CS} = \begin{bmatrix} 117.48 & 34.85 \\ 34.85 & 3251.70 \end{bmatrix},$$

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

INDIAN STATISTICAL INSTITUTE

Mid Semester Examination: 2018 – 19

Course Name: M Tech (QROR), 2nd Year
Subject Name: Business Analytics

Date: 06.09.2018 Maximum Marks: 60 Duration: 2 Hours

Notes: Answer all the questions.

1. Suppose you are trying to build a model to predict whether a potential buyer would buy a particular type of mobile phone or not. You have decided to use the Naïve Bayes' classifier. You have identified the following explanatory variables – age (3 levels – young, middle aged, old), gender (2 levels – male, female), income (3 levels – high, medium, low), type of job (3 levels – government, corporate, other) and past buying habit (2 levels – traditional, experimental).
    a. How many parameters would you need to estimate assuming that you are using all the explanatory variables? [3]
    b. Explain briefly the assumptions to be made to apply the Naïve Bayes' classification model in this specific context (i.e. which events are assumed to be independent of each other?). [3]
    c. How will you use two-way tables to get an idea about the reasonableness of the Naïve Bayes' assumption? Show the structure of a sample table. [6 + 3 = 9]
    d. Can you use Naïve Bayes' classifier with only retrospective data? Explain. [5]

2. Let $g(x) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p$ be the logit of a multiple logistic regression model, where $X_1, X_2, \ldots X_p$ are the explanatory variables. Let Y be the binary response variable taking values 0 and 1.
    a. What diagrams can you use to check whether the linearity assumption for the logit is reasonable? Explain briefly. [5]
    b. Express $\text{Prob}(Y = 1 / X_1 = x_1, X_2 = x_2, \ldots X_p = x_p)$ in terms of the logit. [2]
    c. Suppose there are 5 explanatory variables (i.e. p = 5), all measured in ratio scale. You have fitted the model and have found that the estimated coefficients for the 5 variables $X_1, X_2, \ldots X_5$ to be – 0.024, – 0.014, 1.004, 0.433, and – 0.049 respectively. Interpret the coefficients taking any one as an example. [3]
    d. Suppose the standard errors of the 5 estimated coefficients given above are 0.0337, 0.0065, 0.4979, 0.3622 and 0.1672 respectively.
        i. Which variables are likely to be significant? Explain. [5]
        ii. Suppose the log likelihood of the fitted model is -111.286. Is it possible to have the likelihood of the null model to be -111.200? Explain. [2]
        iii. Let G_1 and G_2 be the log likelihoods of the fitted model and null models respectively in this case. How will you use the values of G_1 and G_2 to test the global null hypothesis that all 5 coefficients are zero? [3]

3. In 1970s, some legal experts believed that the verdict of death penalty in cases of homicide (murder) in the US depended on the race of the victim (the person murdered) and the defendant (the person accused of the murder). A detailed study covering 326 cases where death penalty was awarded to the defendant was carried out and the following were observed. Out of 151 cases where both defendant and victim were white, 19 were awarded death penalty. Out of 9 cases where defendant was white and victim black, none were awarded death penalty. Out of 63 cases where defendant was black and victim white, 11 were awarded death penalty. Finally, out of 103 cases where both defendant and victim were black, 6 were awarded death penalty.
    a. Present the data in tabular format. [5]
    b. Do you think that the data support the belief that the race of the victim (white or black) and defendant has an impact on the chance of awarding death penalty? Explain, preferably with a visual presentation. [8]
    c. Suppose you were told to fit a logistic regression model to estimate the chance of death penalty being awarded in different situations.
        i. Identify the response and explanatory variables. [2]
        ii. Do you think the model fitting will encounter any problem? Explain. [5]

INDIAN STATISTICAL INSTITUTE

Mid-Semestral Examination : 2018-19

Course name : M. Tech. (QR & OR)-II

Subject Name : Industrial Experimentation

Date: 07/09/2018 Maximum Marks: 100 Duration 2 hours

NOTE: (i) This paper carries 115 marks. Question no. 5 is compulsory. Answer as much as you can from the remaining but the maximum you can score in this remaining portion is 70. The marks are indicated in [ ] on the right margin.

(ii) The symbols and notations have the usual meaning as introduced in your class.

(iii) Give examples wherever feasible.

1. Define the following five terms with suitable example wherever feasible:

   Experiment, Treatment, Interaction effect, Experimental unit, Orthogonal contrasts, Orthogonal data.

   $(4\times4+5\times2) = [26]$

2. Four treatments are being compared in five blocks using a randomized complete block design.

   a) Write down an appropriate fixed effect model for the design when each treatment is replicated once in a block.

   b) Find the $E(MS_{Treatments})$. [You may assume that the treatment and block effects in the model are expressed as deviation from their respective means.]

   c) If an observation is found missing then how you do propose to estimate it for an approximate analysis. What will be the error degrees of freedom for this analysis?

   $(4+10+\overline{4+1}) = [19]$

3. List the outline or steps in designing an experiment. What are orthogonal *Latin* squares and what is a *Graeco-Latin* square design? What is a standard *Latin* square?

   $(5+(5+2)+2) = [14]$

4. An engineer wishes to test the effect of **three** different **assembly methods** (A, B, C) on the assembly time of a mobile phone. Each run requires the services of an operator. Furthermore, each assembly method produces such fatigue that the time required for the last assembly may be greater than the time required for the first assemble, regardless of the method. To account for this variability, the engineer decides to include order of assembly as a block factor. The entire experiment is replicated over **three** different **lines of production**. Different sets of **three operators** are used for different lines but the order of assembly is kept same for all the lines. Analyse the data (given overleaf in Table 1) from this experiment ($\alpha = 0.05$) and draw appropriate conclusions. If assembly methods are different then which method/(s) will you recommend for use and why?

P.T.O.

Table 1: Data on Assembly Time (coded units)

			Line 1			Line 2			Line 3		
			Operator			Operator			Operator		
			1	2	3	1	2	3	1	2	3
Order of assembly	I		*A* (7)	*B* (8)	*C* (9)	*C* (8)	*B* (4)	*A* (7)	*B* (8)	*A* (7)	*C* (8)
	II		*B* (4)	*C* (5)	*A* (6)	*B* (6)	*A* (3)	*C* (6)	*A* (4)	*C* (8)	*B* (6)
	III		*C* (6)	*A* (3)	*B* (4)	*A* (5)	*C* (8)	*B* (7)	*C* (8)	*B* (4)	*A* (7)

(20+6) = [26]

5. Class Assignments. [30]

----- x x x x x -----

F distribution (5%) Table

$F_{0.05, v1, v2}$

	Degree of freedom for the Numerator (v_1)										
Degree of freedom for the Denominator (v_2)	1	2	3	4	5	6	7	8	10	12	24
2	18.5	19.0	19.2	19.2	9.3	19.3	19.4	19.4	19.4	19.4	19.5
3	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.79	8.74	8.64
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	5.96	5.91	5.77
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.74	4.68	4.53
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.06	4.00	3.84
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.64	3.57	3.41
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.35	3.28	3.12
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.14	3.07	2.90
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	2.98	2.91	2.74
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.85	2.79	2.61
12	4.75	3.88	3.49	3.26	3.11	3.00	2.91	2.85	2.75	2.69	2.51
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.67	2.60	2.42
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.60	2.53	2.35
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.54	2.48	2.29
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.49	2.42	2.24
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.45	2.38	2.19
18	4.41	3.55	3.16	2.93	2.77	3.66	2.58	2.51	2.41	2.34	2.15
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.38	2.31	2.11
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.35	2.28	2.08

PERCENTAGE POINTS OF THE T DISTRIBUTION

		Tail Probabilities							
One Tail		0.10	0.05	0.025	0.01	0.005	0.001	0.0005	
Two Tails		0.20	0.10	0.05	0.02	0.01	0.002	0.001	
	1	3.078	6.314	12.71	31.82	63.66	318.3	637	1
	2	1.886	2.920	4.303	6.965	9.925	22.330	31.6	2
	3	1.638	2.353	3.182	4.541	5.841	10.210	12.92	3
	4	1.533	2.132	2.776	3.747	4.604	7.173	8.610	4
	5	1.476	2.015	2.571	3.365	4.032	5.893	6.869	5
	6	1.440	1.943	2.447	3.143	3.707	5.208	5.959	6
	7	1.415	1.895	2.365	2.998	3.499	4.785	5.408	7
	8	1.397	1.860	2.306	2.896	3.355	4.501	5.041	8
	9	1.383	1.833	2.262	2.821	3.250	4.297	4.781	9
	10	1.372	1.812	2.228	2.764	3.169	4.144	4.587	10
	11	1.363	1.796	2.201	2.718	3.106	4.025	4.437	11
	12	1.356	1.782	2.179	2.681	3.055	3.930	4.318	12
	13	1.350	1.771	2.160	2.650	3.012	3.852	4.221	13
	14	1.345	1.761	2.145	2.624	2.977	3.787	4.140	14
	15	1.341	1.753	2.131	2.602	2.947	3.733	4.073	15
	16	1.337	1.746	2.120	2.583	2.921	3.686	4.015	16
	17	1.333	1.740	2.110	2.567	2.898	3.646	3.965	17
	18	1.330	1.734	2.101	2.552	2.878	3.610	3.922	18
	19	1.328	1.729	2.093	2.539	2.861	3.579	3.883	19
	20	1.325	1.725	2.086	2.528	2.845	3.552	3.850	20
	21	1.323	1.721	2.080	2.518	2.831	3.527	3.819	21
	22	1.321	1.717	2.074	2.508	2.819	3.505	3.792	22
	23	1.319	1.714	2.069	2.500	2.807	3.485	3.768	23
	24	1.318	1.711	2.064	2.492	2.797	3.467	3.745	24
	25	1.316	1.708	2.060	2.485	2.787	3.450	3.725	25
	26	1.315	1.706	2.056	2.479	2.779	3.435	3.707	26
	27	1.314	1.703	2.052	2.473	2.771	3.421	3.690	27
	28	1.313	1.701	2.048	2.467	2.763	3.408	3.674	28
	29	1.311	1.699	2.045	2.462	2.756	3.396	3.659	29
	30	1.310	1.697	2.042	2.457	2.750	3.385	3.646	30
	32	1.309	1.694	2.037	2.449	2.738	3.365	3.622	32
	34	1.307	1.691	2.032	2.441	2.728	3.348	3.601	34
	36	1.306	1.688	2.028	2.434	2.719	3.333	3.582	36
	38	1.304	1.686	2.024	2.429	2.712	3.319	3.566	38
	40	1.303	1.684	2.021	2.423	2.704	3.307	3.551	40
	42	1.302	1.682	2.018	2.418	2.698	3.296	3.538	42
	44	1.301	1.680	2.015	2.414	2.692	3.286	3.526	44
	46	1.300	1.679	2.013	2.410	2.687	3.277	3.515	46
	48	1.299	1.677	2.011	2.407	2.682	3.269	3.505	48
	50	1.299	1.676	2.009	2.403	2.678	3.261	3.496	50
	55	1.297	1.673	2.004	2.396	2.668	3.245	3.476	55
	60	1.296	1.671	2.000	2.390	2.660	3.232	3.460	60
	65	1.295	1.669	1.997	2.385	2.654	3.220	3.447	65
	70	1.294	1.667	1.994	2.381	2.648	3.211	3.435	70
	80	1.292	1.664	1.990	2.374	2.639	3.195	3.416	80
	100	1.290	1.660	1.984	2.364	2.626	3.174	3.390	100
	150	1.287	1.655	1.976	2.351	2.609	3.145	3.357	150
	200	1.286	1.653	1.972	2.345	2.601	3.131	3.340	200
Two Tails		0.20	0.10	0.05	0.02	0.01	0.002	0.001	
One Tail		0.10	0.05	0.025	0.01	0.005	0.001	0.0005	
		Tail Probabilities							

D E G R E E S O F F R E E D O M

INDIAN STATISTICAL INSTITUTE

First Semester Examination: 2018 – 19

Course Name: M Tech (QROR), 2nd Year
Subject Name: Business Analytics

Date: 12 November 2018 **Maximum Marks: 100** **Duration: 3 Hours**

Notes: Answer questions 3 and 5 and any two from 1, 2 and 4.

1. A company sales its products to other companies on credit, i.e. the company sends the item first and collects the money later. The company has 200 customers and it is noted that the number of days taken by the customers to make the payment varies. It is further noted that the company sales at least 100 times to each customer and variation in the number of days to make payment for each sale. Thus the actual number of days required for payment may be denoted by y_{ij}, where i = 1, 2, ...200 give the customer number and j = 1, 2, ...give the number of the sales deals. The company finds that some customers take much longer to make payment compared to others and decided to group the customers on the basis of their actual payment time (measured in number of days).
    a. Suppose you are asked to group the customers with respect to their average payment time, i.e. average number of days to make payment.
        i. Suggest a method to carry out the grouping.
        ii. Write the model.
        iii. What assumptions did you make? [3 + 3 + 3 = 9]
    b. It is known that the customers do not vary much with respect to their average or the lower percentiles of the payment time. The main variation occurs in the higher percentiles, typically from the 80th percentile onward. Accordingly, it was decided to cluster the customers using hierarchical clustering using the variate (80th, 90th, 95th) percentile for the number of days to make payment for each customer. Further, it was observed that the higher percentiles of the number of days to make payment are large (80th percentile over 50 days) for all customers.
        i. The distance between two customers may be defined as the Euclidian distance or city block distance. Which one would you choose and why? [4]
        ii. Suppose you have been asked to choose between single linkage, complete linkage and average linkage. Which linkage method will you choose? Explain. Define the three different linkage types. [5 + 3 X 3 = 14]
        iii. Which diagram may be used to visually present the payment time for the different customers in the identified clusters? Explain briefly. [3]
2. A software maintenance company receives maintenance requests over mail. They want to allocate the maintenance tasks automatically to one of the four different expert groups numbered 1, 2, 3 and 4. The allocation is made using 5 different keywords K_1, K_2,... K_{10}. Let p_{ij} = Prob(Request contains keyword K_j / Request belongs to expert group i).
    a. In this context
        i. Identify the response and explanatory variables and their measurement scales. [2 + 2 = 4]
        ii. Formulate this problem as a problem of classification clearly stating the probability that need to be estimated. [5]
    b. Show how you can use the Naïve Bayes' classification methodology to allocate a maintenance request to a group. Which probabilities would you need to know apart from p_{ij}'s to carry out this classification? [12]
    c. Suppose you have data on 10000 request mails with the values of K_1, K_2,... K_{10} and the actual expert group to which the requests were allocated. Present this data in a tabular manner and show how this table may be used to devise a classification rule. [3 + 6 = 9]
3. Answer the following
    a. Explain briefly the concept of cost complexity pruning in the context of a decision tree. What happens when the tuning parameter is zero or very large? [5 + 3 = 8]
    b. What are the formulations for ridge regression and lasso? [6 + 6 = 12]

P.T.O

4. Suppose the credit rating of an individual is a continuous variable varying between 0 – 10. An individual with a high credit rating is very likely to pay back his / her loan on time. Suppose a bank is trying to use the credit rating score to determine whether an individual should be given a loan or not. The bank has decided to use Linear Discriminant Analysis (LDA) to build this model.
    a. Identify the response and explanatory variables and their scales of measurement? [4]
    b. What assumptions are required to be able to use the LDA effectively? What graphical techniques may be used to judge the validity of these assumptions? Explain. [3 + 5 = 8]
    c. Suppose data on the result of loan repayment and credit rating of a large number of individuals are available with the bank. Explain how you will fit the Linear Discriminant Analysis model to this data. [8]
    d. It is known that about 5% of the loans go bad (i.e. the person taking the loan does not pay back). You wish to use logistic regression to estimate the probability that a person would default given his / her credit rating. You choose 100 bad loans and 100 good loans and find the credit rating of the persons availing these loans
        i. Do you think taking equal number of good and bad loans is reasonable even when you are aware that only 5% of the loans go bad? Explain. [5]
        ii. Can you use this sample to estimate the probability of default? Explain. [5]
5. Explain the following briefly
    a. Express MSE in terms of bias and variance. [3]
    b. Suppose a binary classifier is being used to detect fraudulant transactions on credit card. You were told that the rate of true positive is 100% (i.e. the probability of testing positive given that the transaction is fraudulant is 1) and hence the classifier must be very good. In fact it was concluded that whenever the classifier shows a positive result, the transaction must be fraudulant. Do you agree? Explain. [10]
    c. If you need to use ridge regression or lasso for the purpose of feature selection, which one will you choose? Explain. [7]

INDIAN STATISTICAL INSTITUTE

First Semester EXAMINATION: 2018

Course Name: M. TECH (QROR) II

Subject Name: OR II

Date: 16 /11/2018 **Maximum Marks: 100** **Duration: 3 hours**

Answer 1. and any four from 2. to 6.

1. a) Define linear complementarity problem LCP (q, M).
   b) Formulate linear programming problem and quadratic programming problem as linear complementarity problem. State the Lemke's algorithm.

[2+2+2+4=10]

2. a) 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 $\lambda \in (0, \delta)$.

   b) 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 semidefinite if $\bar{x}$ is a local minimum.

[6 + 14 = 20]

3. a) Let S be a nonempty open set in R^n and $f : R^n \to R$, $g_i : R^n \to R$ for $i=1, \ldots m$. Consider the problem to minimize $f(x)$ subject to $x \in X$ and $g_i(x) \leq 0$ 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 show that there exist scalars u_0 and u_i such that

$$u_0 \nabla f(\bar{x}) + \sum_{i=1}^{m} u_i \nabla g_i(\bar{x}) = 0$$
$$u_i g_i(\bar{x}) = 0 \quad \text{for } i = 1, \ldots, m$$
$$u_0, u_i \geq 0 \quad \text{for } i = 1, \ldots, m$$
$$(u_0, u) \neq (0, 0)$$

   b) State under what condition the above statement can be written as

P.T.O.

$$\nabla f(\bar{x}) + \sum_{i=1}^{n} \bar{u}_i \nabla g_i(\bar{x}) = 0$$

$$\bar{u}_i g_i(\bar{x}) = 0$$

$\bar{u}_i \geq 0$ with at least one $\bar{u}_i > 0$.

[12+ 8 = 20]

4. a) State a method to solve a linear fractional programming problem as linear programming problem.
   b) How do you solve a separable nonlinear programming problem as linear programming problem?

[9 + 11 = 20]

5. a) Let S be a nonempty open convex set in R^n and $f: S \to R$ be differentiable on S. Prove that f is convex if and only if

$$[\nabla f(x_2) - \nabla f(x_1)]^t (x_2 - x_1) \geq 0 \text{ for each } x_1, x_2 \in S.$$

   b) State the definition of quasiconvex and pseudoconvex functions.

[14+6=20]

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

[8+12= 20]

7. Assignment [10]

INDIAN STATISTICAL INSTITUTE

First Semestral Examination: 2018-19

Programme: M. Tech. (QR & OR) II

Course: Advanced Multivariate Analysis

Date of Examination: 19.11.2018 Maximum Marks: 100 Duration: 3½ hours

Note:
1. This paper carries115 marks.
   Answer all questions but the maximum you can score is 100.
2. All notations have their usual meanings.
3. Give to the point answer.

---

1) Write *Agree* or *Disagree* and briefly Justify

i) Single, complete and linkage methods would not always lead to the same cluster solution.

ii) Principal component analysis is a dependence technique.

iii) One can go back to the original variables in the case of principal component regression.

iv) Discriminant function analysis is an interdependence technique.

v) A model developed by multiple linear regression method represents the underlying causal model.

vi) A multiple linear regression model cannot be used if it is hypothesised by the scientist that the response is related to the dependent variables as $y = \beta_0 x_1^{\beta_1} x_2^{\beta_2}$.

vii) One may carry out ANOVA for each variable instead of a MANOVA.

viii) Factor analysis is a dependence technique.

ix) Cluster Analysis is an interdependence technique.

x) Principal Component Regression is an interdependence technique.

[2 x 10 =20]

2) i) What do you understand by multicollinearity? How does it affect estimates of regression coefficients?

ii) What is Variance Inflation Factor? How does it help to detect the multicollinearity?

iii) What are the methods for dealing with multicollinearity?

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

3) a) What are the purposes of Principal Component Analysis?

b) Six haematological variables X_1, X_2, X_3, X_4, X_5, and X_6 were measured on 51 workers. The summary of the data is given below:

Let S and R be the covariance matrix and correlation matrix respectively. The diagonal elements of S are 0.69, 5.4, 2006682.4, 90.3, 56.4 and 18.1. Eigenvalue of S and R are as follows:

Eigenvalues

2006760	2.42
65	1.4
18	1.03
7	0.92
3	0.2
0	0.03

First 3 eigenvectors of S and R are :

S			R		
e1	**e2**	**e3**	**e1**	**e2**	**e3**
0.00016	0.005	-0.0136	0.424	-0.561	-0.150
0.00051	0.017	0.0787	0.426	-0.528	0.087
0.99998	- 0.001	-0.0002	0.563	0.387	-0.051
0.00529	0.698	0.0174	0.454	0.267	0.166
0.00322	-0.716	0.0195	0.303	0.425	-0.296
0.00020	0.025	0.9965	0.073	0.069	0.293

(i) How many principal components you should retain, separately for S and R ?

(ii) Does the large variance of X_3 affect the pattern of the components of S?

(iii) Should we carry out the PCA with S? Justify your answer.

(iv) Interpret the components of either S or R.

[4+(3+2+3+3) = 15]

4) (a) Consider two multivariate normal populations: $N_p(\mu_1, \Sigma_1)$ *and* $N_p(\mu_2, \Sigma_2)$. Derive the minimum Total Probability of Misclassification (TPM) rule to classify an observation x_o, assuming equal prior probabilities. Write down the expression of *Actual Error Rate* for the classification rule.

(b) Various aspects of economic cycles were measured for certain consumer goods and producer goods. The variables are x_1 = length of cycle, x_2 = percentage of rising prices, x_3 = cyclical amplitude and x_4 = rate of change. There were $n_1 = 9$ observations on consumers' goods and $n_2 = 10$ observations on producers' good. The following summary statistic were obtained:

$$\bar{x}_1 = \begin{pmatrix} 48.61 \\ 52.67 \\ 0.056 \\ 0.922 \end{pmatrix}, \quad \bar{x}_1 = \begin{pmatrix} 90.30 \\ 50.50 \\ 17.40 \\ 1.07 \end{pmatrix}, \quad S_{pooled}^{-1} = \begin{pmatrix} 0.00225 & 0.00227 & -0.00592 & 0.03793 \\ 0.00227 & 0.01723 & -0.01059 & 0.13159 \\ -0.00592 & -0.01059 & 0.08180 & -0.35849 \\ 0.03793 & 0.13159 & -0.35849 & 7.10212 \end{pmatrix}$$

$T^2(x_1, x_2, x_3, x_4) = 18.46248$, $T^2(x_2, x_3, x_4) = 11.64578$

$T^2(x_1, x_3, x_4) = 18.43656$ $T^2(x_1, x_2, x_4) = 15.06727$,

$T^2(x_1, x_2, x_3) = 18.45909$.

i) Calculate the partial F-statistic corresponding to each variable, and find out the variables having significant contribution in group separation. Rank the variables in respect of relative contribution of group separation.

ii) Derive the linear discriminant function which maximally separates the groups. Which variable has highest contribution in group separation? Calculate the value of Mahalanobis distance.

iii) Consider an observation $x_0 = (66.5, 48, 15, 1.0)'$. Classify this observation as consumers' goods or producers' goods using the classification rule obtained in (a).

[12+ (6 + 6 + 6) = 30]

5) i) What are the similarities and dissimilarities between Principal Component Analysis and Factor Analysis?

ii) Write down the orthogonal factor model and the associated assumptions. Explain the notations used.

iii) Show that the assumptions of the factor models and communality remain unchanged under orthogonal transformation.

iv) Consider the three quality characteristics X_1, X_2 *and* X_3. Based on a random sample of size n = 30, the following correlation matrix was obtained.

$$\boldsymbol{R} = \begin{bmatrix} 1.0 & 0.63 & 0.45 \\ 0.63 & 1.0 & 0.35 \\ 0.45 & 0.35 & 1.0 \end{bmatrix}.$$

Eigenvalues and corresponding eigenvectors of R are:

$\lambda_1 = 1.96$, $e_1 =$ $(0.625, \quad 0.593, \quad 0.507)'$

$\lambda_2 = 0.68$, $e_2 =$ $(-0.219, \quad -0.491, \quad 0.843)'$

$\lambda_3 = 0.36$, $e_3 =$ $(-0.749, \quad -0.638, \quad -0.177)'$

How many factors may be extracted? Calculate the Loading matrix.

[4 + 4 + 6 + 6 = 20]

6) (a) Differentiate between Partitional and Hierarchical clustering.

(b) The distance matrix between pairs of five items is given below.

$$\begin{bmatrix} 0 & 2 & 12 & 1 & 6 \\ 2 & 0 & 9 & 7 & 8 \\ 12 & 9 & 0 & 4 & 5 \\ 1 & 7 & 4 & 0 & 3 \\ 6 & 8 & 5 & 3 & 0 \end{bmatrix}$$

Cluster the five items using any-linkage method. Draw the dendrogram and interpret.

[4+(8 + 3) = 15]

INDIAN STATISTICAL INSTITUTE

Semester Examination: 2018-19

(First Semester)

Course Name: M.Tech (QR&OR) II year

Subject Name : Reliability II

Date: 22 November 2018 Maximum Marks: 100 Duration: 3 hours

1. a) Define the following classes of life distributions.

   i) DFR
   ii) DFRA
   iii) NWU
   iv) NWUE

   b) Show and prove the relationships between the above four classes of life distributions.

   [2x4+ 4x3=20]

2. a) A system consists of six subsystems that must function if the system has to function properly. The system reliability goal is 0.990. All the six subsystems have identical reliability improvement effort functions. The estimated subsystem reliabilities at the present time are 0.70, 0.75, 0.80, 0.85, 0.90, 0.95. Determine what reliability goal should be apportioned to the subsystems so as to minimize the total effort spent on the system improvement?
   b) If we apply the ARINC apportionment technique with the system mission time as 20 hours and the present reliability of the subsystems calculated for 50 hours, determine the apportioned reliabilities of the subsystems in order to achieve the system reliability goal of 0.990.

   [15+15=30]

3. Suppose the stress acting on a component is uniformly distributed over an interval $[S_{min}, S_{max}]$. The strength of the component has a two-parameter Weibull distribution with parameters S_0 (shape parameter) and θ (scale parameter). Derive an expression for the reliability of the component. Let

P.T.O

S_{min}=10, S_{max}=30
S_0 =20, θ=30.

Find out the reliability of the component.

[15+5=20]

4. a) Define likelihood ratio ordering of probability distributions.
   b) Suppose $\pi(h,t)$ denote the prior probability distribution of undetected number of bugs in a software with history h and t test cases having tested already. Suppose the likelihood ratio ordering of two probability distributions be denoted by π^* (h, t) $>^{LR}$ $\pi(h,t)$ indicating π^* (h, t) as a preferable prior. Show that if $\pi^*(h_1,t_1) >^{LR} \pi(h_2,t_2)$ then

   $\sum_{n\geq k} \pi_n^*(h_1,t_1) \geq \sum_{n\geq k} \pi_n(h_2,t_2)$ for all k≥1, where π_n(h,t) is the probability of having n bugs in the software with 'h' as the history and 't' numbers of test cases already tested. [5+15=20]

5. a) Write down clearly the steps involved in planning for an accelerated life test (ALT).
   b) Classify the different models that are applicable for ALT.

[7+3=10]

INDIAN STATISTICAL INSTITUTE

Semester Examination: 2018-19

(First Semester)

Course Name: M.Tech (QR&OR) II year

Subject Name : Reliability II

Date: 29.01.19 Maximum Marks: 70 Duration: 3 hours

Back Paper

1. a) Define decreasing failure rate (DFR) distribution and decreasing failure rate average (DFRA) distribution in the context of life distribution.

   b) Show that F is IFRA if and only if $\bar{F}(\alpha t) \geq \bar{F}^{\alpha}(t)$, for all $0<\alpha<1$ and $t \geq 0$

   c) Prove that IFRA distributions possess closure property.

   [4+4+12=20]

2. a) Describe the problem of optimizing reliability allocation to each of the subsystems of a given system using effort minimization algorithm.

   b) Describe specifically the assumptions to be made on the effort function for transforming the reliability level of the subsystem from R_i to R_i^* ($R_i^* > R_i$).

   c) State the solution to the above problem (proof not required).

   d) A system consists of six subsystems that must function if the system has to function properly. The system reliability goal is 0.998. The present reliability of the subsystems are 0.70, 0.75, 0.80, 0.85,0.90 and 0.95 calculated for 50 hours, The system mission time is 20 hours. Use ARINC methodology for determining target reliability to be achieved by the subsystems in order to achieve a system reliability target of 0.998 and find out target reliabilities for the subsystems.

   [5+4+2+14 = 25]

P.T.O.

3. Let the prior distribution $\pi = (\pi_0, \pi_1, \pi_2, \ldots)$ of the undetected number of faults be given as {0.01, 0.02, 0.03, 0.04, 0.2, 0.3, 0.2, 0.1, 0.05, 0.05, 0, 0, 0, ...}. Consider that the conditional probability of success given that there were i bugs in the software is denoted by $p_i = \frac{\pi_i}{20}$. Find out the two-step-posterior probability $\pi_2(\mathcal{SS},2)$.

[10]

4. A manufacturer of Bourdon tubes (used as a part of pressure sensors in avionics) wishes to determine its MTTE. The manufacturer defines the failure as a leak in the tube. The tubes are manufactured from 18 Ni (250) maraging steel and operate with dry 99.9% nitrogen or hydraulic fluid as the internal working agent. Tubes fail as a result of hydrogen embrittlement arising from the pitting corrosion attack. Because of the criticality of these tubes, the manufacturer decides to conduct ALT by subjecting them to different levels of pressures and determining the time for a leak to occur. The units are continuously examined using an ultrasound method for detecting leaks, indicating failure of the tube. Units are subjected to three stress levels of gas pressures and the times for tubes to show leak are recorded. It is given that the time to leak in hours for each stress level follow Weibull distribution with the same shape parameter, but with different scale parameters.

The mean and 50th percentile at each stress level is given as follows:

Pressure (psi)	100	120	140
Mean	9276.1	4767.01	835.3
50th percentile	9050	4681	821

Determine the mean lives and the reliability functions for design pressures of 80 and 90 psi.

[15]

INDIAN STATISTICAL INSTITUTE

Back Paper Examination : 2018-19

Course name : M. Tech. (QR & OR)-II

Subject Name : Industrial Experimentation

Date: 30 / 01 / 19 Maximum Marks: 100 Duration 3 hours

NOTE: (i) This paper carries 100 marks. Answer all the questions. The marks are indicated in [ ] on the right margin.

(ii) The symbols and notations have the usual meaning as introduced in your class.

(iii) Assume factors are fixed and normality assumptions are satisfied, unless stated otherwise

1. Define a contrast. When are two contrasts said to be orthogonal? When is a set of two-way classified variable data said to be orthogonal? Discuss the Scheffe's procedure for constructing simultaneous confidence intervals for all possible contrasts. What is a random factor? What is meant by a components of variance model? When are factors said to be crossed?

[2+2+5+6+2+4+3 = 24]

2. What is a Latin square design? How are the basic principles of experimentation implemented for such a design?

[3+4 = 7]

3. Given the following three choices of generators for a 2_{IV}^{7-2} design, which one would you prefer and why?

Design 1: $F = ABC$, $G = BCD$,

Design 2: $F = ABC$, $G = ADE$,

Design 3: $F = ABCD$, $G = ABDE$

[5]

4. (a) Explain the concept of AB and AB^2 components of AB interaction where A and B are two three-level factors.

(b) Illustrate the use of such components of interaction in deriving a 3^{4-2} fractional factorial design. What is the alias structure of your design?

[5+(6+4) = 15]

5. a) What is a nested factor? What is a three-stage nested design?

b) A foundry wishes to investigate the hardness of two different formulations (A) of a metal alloy. To this end, an experiment is conducted in which three heats (B) of each alloy formulation are prepared, two ingots (C) are selected at random from each heat

for testing and two hardness measurements (y) are made on each ingot. Using the data that followed, an analysis of variance is done based on the restricted form of the mixed model and the ANOVA table is given below.

Table 1: ANOVA on Hardness of Alloy

Source	Df	Sum_Sq	Mean_Sq	F_value	Pr(>F)
A	1	315.4	315.38	2.9099	0.337550
B	2	6129.1	3064.54	49.0013	0.019999
C	1	651.0	651.04	3.6481	0.080321
AB	2	324.8	162.38	0.2420	0.805127
AC	1	108.4	108.38	0.6073	0.450900
BC	2	125.1	62.54	0.3504	0.711350
ABC	2	1341.8	670.88	3.7593	0.054001
Error	12	2141.5	178.46		
Total	23	11137.1			

Comment on correctness of the analysis with justification(s). Write the appropriate model for the experiment. Rectify the analysis of variance table, if inaccurate, in line with your proposed model. Write the expected mean squares for different effects and test for significance of these effects. Estimate the appropriate variance components of your proposed model.

[(2+3)+(3+3+5+7+4) = 28]

6. What is response surface methodology? What are the designs used for fitting response surfaces called? What is meant by method of steepest ascent? What is usually taken as the path of steepest ascent? Construct a simplex design in three variables. For the first-order model $\hat{y} = 6.5 + 1.5x_1 - 0.8x_2 = 2.0x_3$, find three points along the path of steepest ascent.

[3+1+3+3+6+5 = 21]

Table 2: F distribution (5%) Table

$F_{0.05,\, v1,\, v2}$

/	df_1=1	2	3	4	5	6	7	8	9	10	12	24
df_2=2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.45
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.64
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.77
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.53
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.84
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.41
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.12
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	2.90
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.74
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.61
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.51
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.42
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.35
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.29
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.24
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.19
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.15
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.11
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.08
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.05
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.03
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.01
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	1.98
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	1.96
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	1.89
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.52

INDIAN STATISTICAL INSTITUTE

Second Mid-Semester Examination (2018 – 2019)

Course Name	:	**M. Tech (QR & OR)**
Subject	:	**Industrial Engineering and Management**
Date	:	**25/02/2019 at 10:30 AM**
Maximum Marks	:	**30**
Duration	:	**90 minutes**

Question Paper

Answer *any three* questions

1. (a) Risk is associated with crashing of projects. Comment and justify.

   (b) Assume the following project data given in Table. It is required to crash the project duration from its original duration to a final duration of 110 days. Assume daily indirect cost of Rs 100. Show all necessary steps and calculations.

Activity	Preceded by	Normal		Crash	
		Duration (Day)	Cost (Rs.)	Duration (Day)	Cost (Rs.)
A	---	120	12000	100	14000
B	----	20	1800	15	2800
C	B	40	16000	30	22000
D	C	30	1400	20	2000
E	D,F	50	3600	40	4800
F	B	60	13500	45	18000

[2+8]

2. (a) Why do you standardize the procedures?

   (b) What are the seven wastes in lean system? [3+7]

3. (a) Prove the required number of observations to predict true time within ± 5% precision and 95% confidence level as

$$N = \left(\frac{40\sqrt{N\sum x^2 - (\sum x)^2}}{\sum x}\right)^2$$

(Symbols have their usual meanings)

   (b) How do you improve the inherent safety of a plant?

[5+5]

4. Write short notes on ***any two***: **[2X5=10]**

   (a) Time study and wage incentive

   (b) Total ineffective time

   (c) Function of a manager

   (d) Steps of finding preferred methods

   (e) Delphi method