IES | ISS Exam, 2019

XZH-T-STSS

STATISTICS

Paper – IV

Time Allowed : Three Hours

Maximum Marks : 200

Question Paper Specific Instructions

Please read each of the following instructions carefully before attempting questions :

*There are **FOURTEEN** questions divided under **SEVEN** sections.*

*Candidate has to choose any **TWO** Sections and attempt the questions therein. All the Sections carry equal marks.*

The number of marks carried by a question / part is indicated against it.

Wherever any assumptions are made for answering a question, they must be clearly indicated.

Diagrams / Figures, wherever required, shall be drawn in the space provided for answering the question itself.

Unless otherwise mentioned, symbols and notations have their usual standard meanings.

Chi-Square Distribution Table is given with Q. 10(b).

Attempts of questions shall be counted in sequential order. Unless struck off, attempt of a question shall be counted even if attempted partly.

Any page or portion of the page left blank in the Question-cum-Answer Booklet must be clearly struck off.

*Answers must be written in **ENGLISH** only.*

SECTION A

(Operations Research and Reliability)

Q1. (a) A department head has four subordinates and four tasks to be performed. The subordinates differ in efficiency and the tasks differ in their intrinsic difficulty. His estimate of the time each man would take to perform each task is given in the matrix below :

Task	*Men* E	F	G	H
A	18	26	17	11
B	13	28	14	26
C	38	19	18	15
D	19	26	24	10

How should the tasks be allocated to a man so as to minimize the total man hours ? *10*

(b) There are four villages V_1, V_2, V_3 and V_4 which are affected due to floods. Food items are to be dropped in these villages by aircrafts A_1, A_2 and A_3. Details of trips are given in the following table :

	V_1	V_2	V_3	V_4	Total Number of Trips from A_i
A_1	2	5	9	10	50
A_2	10	8	6	3	60
A_3	8	10	5	3	40
Total Number of Trips from V_j	30	50	60	40	

Solve the transportation problem to drop maximum food. *10*

(c) Solve the following zero-sum game for two persons. Obtain the optimum strategies for both players and value of the game. *10*

Player A

$$\text{Player B} \begin{bmatrix} 2 & 4 & 1 \\ 3 & -1 & 2 \\ -1 & 2 & 3 \end{bmatrix}$$

(d) An aircraft uses rivets at a constant rate of 5000 kg per year. The rivets cost ₹ 20 per kg and the company personnel estimate that it costs ₹ 200 to place an order and the carrying cost of inventory is 10% per year.

(i) How frequently should orders for rivets be placed and what quantities should be ordered for ?

(ii) If the actual costs are ₹ 500 to place an order and 15% for carrying cost, the optimum policy would change. How much will the company be losing per year because of imperfect cost information ? *10*

(e) Consider a system of n independent components with failure rates $\lambda_1, \lambda_2, ..., \lambda_n$ respectively. If we assume that the lifetime of the components follow exponential distribution, what is the probability that the component j is the first component to fail ? *10*

Q2. Answer any ***two*** of the following :

(a) Find the optimal solution to the LP problem :

Maximize $z = -x_1 + 3x_2 - 3x_3$

subject to the constraints

$$3x_1 - x_2 + x_3 \leq 7$$

$$-x_1 + 2x_2 \leq 6$$

$$-4x_1 + 3x_2 + 8x_3 \leq 10$$

and $x_1, x_2, x_3 \geq 0$.

(i) Within what range does the cost of x_1 vary so that the optimality remains unaffected ?

(ii) Within what range does the cost of x_3 vary so that the optimality remains unaffected ? *25*

(b) (i) Describe the problem of replacement of items whose maintenance cost increases with time and value of money also changes with time.

(ii) A machine costs ₹ 5000. Operating costs are ₹ 800 per year for the first five years and increasing by ₹ 200 per year in the sixth and subsequent years. Assuming a 10% discount rate of money per year, find the optimum length of time to hold the machine before it is replaced. (Assume that the machine will be eventually sold for scrap at a negligible price). 25

(c) (i) For $M|M|C : (\infty|FIFO)$ queue system, derive the steady state distribution of the queue length.

(ii) A tax consulting centre has four service counters to receive their customers. There is an average arrival rate of ten persons per hour to this centre and each tax advisor spends an irregular amount of time, which have been found to follow exponential distribution. The average service time is 20 minutes. Calculate the average number of customers in the system, average queue length and average waiting time for a customer. 25

(d) (i) Consider an item with failure rate function

$$h(t) = \frac{t}{1+t}.$$

Obtain reliability function, MTTF, conditional survival function : $P(T > x + t \mid T > t)$ and mean residual life function.

(ii) Suppose 100 electronic components are subjected to life testing until all of them failed. The results are given in the following table. Estimate the reliability at the end of each interval and failure rate at the start of each interval. *25*

Time interval in hours		*Number of good components*	
Start	*End*	*Start*	*End*
0	10	100	76
10	20	76	56
20	30	56	38
30	40	38	26
40	50	26	15
50	60	15	7
60	70	7	3
70	80	3	1
80	90	1	0

SECTION B
(Demography and Vital Statistics)

Q3. (a) Explain the meaning of Vital Statistics. Give its usefulness and shortcomings. *10*

(b) What is Sample Registration System (SRS) ? What factors are studied under SRS ? Give its shortcomings. *10*

(c) Derive an algebraic expression relating the probability of a person dying between the age of x and (x + 1), q_x to the force of mortality, μ_x. *10*

(d) Describe how the Abridged life table will be constructed by Greville's method. *10*

(e) Complete the following Life Table (with usual notation) : *10*

Age in Years	l_x	d_x	q_x	p_x	L_x	T_x	e_x^0
11	12000	?	?	?	?	?	?
12	8000	?	?	?	?	?	?
13	6000	?	?	?	?	?	?
14	4000	–	–	–	3000	?	?

Q4. Answer any ***two*** of the following :

(a) (i) Explain any one method of fitting Gompertz growth curve. *15*

(ii) What is Chandrasekaran – Deming formula ? Explain it. *10*

(b) With usual notation, in a life table, e_1 is equal to e_0 and $l_x = (l_1 / l_0)^x \, l_0$, for $0 \leq x < 1$, then show that $p_0 = \exp(-1/e_1)$. *25*

(c) Explain different net migration rates as proposed by Hamilton. *25*

(d) Given the life table of three men A, B and C aged 90, 91 and 92 years respectively as follows :

Age x	90	91	92	93	94	95	96	97	98	99	100
l_x	16090	11490	8012	5448	3607	2320	1447	873	590	98	0

where l_x = Number living at age x. Complete the life table and find the probability that

(i) A, B and C will be alive in two years time (i.e. at the end of two years)

(ii) All will be dead within two years, and

(iii) C will be alive for 6 years' time. *25*

SECTION C

(Survival Analysis and Clinical Trials)

Q5. (a) Consider the survival model characterized by the following piecewise-constant hazard function :

$$\lambda(t) = \begin{cases} \lambda_1, & 0 \le t < \pi_1 \\ \lambda_2, & \pi_1 \le t < \pi_2 \\ \lambda_3, & \pi_2 \le t \end{cases}$$

where π_1 and π_2 are known constants. Based on right censored data (x_i, δ_i), i = 1, ... n, derive the MLEs of λ_1, λ_2 and λ_3 and their standard errors. *10*

(b) The survival times of 11 patients with Acute Myelogenous Leukemia are given below :

$$9, 13, 13^+, 18, 23, 28^+, 31, 34, 45^+, 48, 161^+,$$

where censored data are marked by "+" in the superscript.

Calculate the Kaplan-Meier estimate of S(20) and obtain the standard of log of that estimate. Compare the estimate under the assumption of an exponential distribution. *10*

(c) Consider a family of lifetime distributions which has a hazard function $\lambda(t) = 2t/\beta^2$. Write down the survival function S(t) for this family of models and obtain the MLE $\hat{\beta}$ of β. Also write the expression for the standard error. *10*

(d) A clinical trial studies the survival of lung cancer patients, where the patients were randomly assigned to an experimental treatment or to a placebo. The survival times of patients were recorded.

Based on the data set, a researcher fits the model below :

coxph(formula = surv(time, obs) ~ trt, data = X, method = "exact")

	coef	exp(coef)	se(coef)
trt	−0·8188	0·4410	0·7376

State the model and the assumptions. Test whether the treatment is effective. State the null and the alternative hypotheses and your conclusions at the confidence level of 95%. *10*

(Given that $z_{.975} = 1{\cdot}96$)

(e) Define a competing risk model. Explain a non-parametric inference procedure for such a model. *10*

Q6. Answer any ***two*** of the following :

(a) Assume the hazard rate is at time 0 and linearly increases λ ($\lambda > 0$) as time increases 2 units. Derive the survival function, the probability density function and the median of survival time. Discuss the shape of hazard function and the probability density function. *5+5+5+10=25*

(b) Suppose that the time to death X has an exponential distribution with hazard rate λ and the right censoring time C is exponential with hazard rate θ. Let

$T = \min(X, C)$ and $\delta = 1$ if $X \leq C$; 0 if $X > C$.

Assume that X and C are independent.

(i) Find $P_r(\delta = 1)$.

(ii) Derive the survival function of T.

Let (T_1, δ_1), ... (T_5, δ_5) be 5 randomly taken observations. Assume they follow the above model and it can be shown that δ and T are independent. Construct the likelihood function and the MLE of λ. *5+5+15=25*

(c) The survival time of two treatment groups are recorded as follows :

Treatment A : 3, 5^+, 8, 10^+

Treatment B : 2, 2, 6

(i) Compare the survival function of two treatment groups using two sample log rank tests. First, write out the null and alternative hypotheses. Then calculate the test statistic and the variance of test statistics. Describe the distribution of the test statistic under the null hypothesis. Do not calculate p-value, but simply interpret the conclusions in different situations.

(ii) Assume the two treatments have no difference and so we can combine all observations in one group. Find the MLEs of the parameters when the survival time is assumed to follow the Weibull distribution with shape parameter $\alpha = 2$ i.e. $f(t) = 2\lambda t\, e^{-\lambda t^2}$ $t > 0$. *13+12=25*

(d) Justify the use of Phase III in clinical trials. Discuss briefly the analysis in Phase I – III trials when the outcomes are ordinal. *10+15=25*

SECTION D

(Quality Control)

Q7. (a) Explain the concept of Statistical Quality Control. Give its advantages and limitations. *10*

(b) Distinguish between Charts for Variables and Charts for Attributes. *10*

(c) Explain the following : *10*

(i) $3 - \sigma$ limits

(ii) Natural control limits

(iii) Specification limits

(iv) Modified control limits

(d) A sub-group of 5-items each is taken from a manufacturing process at a regular interval. A certain quality characteristic is measured and $\overline{X}$ and R values computed. After 25 sub-groups it is found that $\Sigma \overline{X} = 357{\cdot}50$ and $\Sigma R = 8{\cdot}8$. If the specification limits are $14{\cdot}4 \neq 0{\cdot}40$; and if the process is in statistical control, what conclusion can you draw about the ability of the process to produce items within specifications ? (For sub-group of 5-items, $d_2 = 2{\cdot}326$) *10*

(e) Explain the situations in which Moving Average Charts and Exponentially Weighted Moving Average Charts are useful. Give limits of these two charts. *10*

Q8. Answer any ***two*** of the following :

(a) A food company puts mango juice into cans and advertise as containing 200 mL of the juice. After filling, 20 samples of 4 cans each were collected by a random method at an interval of 60 minutes and quantity of juice in those cans were measured. The excess of 200 mL in each can was presented below. Construct $\overline{X}$-chart to control the volume of mango juice and comment on whether the production is under control or not. (Given $A_2 = 0{\cdot}73$ for n = 4) *25*

Can	Sample Number																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	15	10	8	12	18	20	15	13	9	6	5	3	6	12	15	18	13	10	5	6
2	12	8	15	17	13	16	19	23	8	10	12	15	18	9	15	17	16	20	15	14
3	13	8	17	11	15	14	23	14	18	24	20	18	12	15	6	8	5	8	10	12
4	20	14	10	12	4	20	17	16	5	20	15	18	10	18	16	15	4	10	12	14

(b) The following data is on defective electronic machine parts from a standard brand of manufacturer. Sample sizes and number of defectives in each sample of 12 independent samples are given below. Using suitable control limits, test whether the process is statistically controlled or not. *25*

Sample No.	1	2	3	4	5	6	7	8	9	10	11	12
Sample Size	115	220	210	220	220	255	440	365	255	300	280	330
No. of Defectives	15	18	23	22	18	15	44	47	13	33	42	46

(c) (i) What is control chart ? Explain the basic principles underlying the control charts. *10*

(ii) Give the criterion for detecting lack of control in Shewhart's control chart for mean and range. *5*

(iii) What are the advantages and limitations of Sampling Inspection technique ? *10*

(d) (i) Explain the three different important control charts for Attributes. *15*

(ii) Explain (a) ASN curve, (b) AOQ curve, (c) ATI curve, and (d) OC curve. *10*

SECTION E
(Multivariate Analysis)

Q9. (a) Let $X = (X_1 \; X_2 \; X_3)' \sim N_3(\mu, I)$ with $\mu = (2 \; -3 \; 1)'$. Then obtain the distribution of (i) $3X_1 - 2X_2 + 5X_3$ and (ii) $(X_1 - X_2 \quad X_2 - X_3)'$. *10*

(b) Let $X \sim N_p(0, \Sigma)$ with

$$\Sigma = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 5 & 3 \\ 1 & 3 & 4 \end{pmatrix}.$$

(i) Derive the condition under which $X'AX$ and $X'BX$ are independent, where A and B are any real symmetric matrices of order $p \times p$.

(ii) Are $Y_1 = 2X_1^2 + X_2^2 + 2X_3^2 + 2X_1X_3 + 2X_2X_3$

and $Y_2 = X_1^2 + 2X_2^2 + 3X_3^2 + 2X_1X_2 + 5X_1X_3 - 2X_2X_3$

independent ? Justify your answer. *10*

(c) If X is a p-variate normal distribution with mean vector μ and dispersion matrix Σ, derive an expression for the expected value of $X'AX$, for any non-zero real symmetric matrix A of real numbers. *10*

(d) Define Hotelling's T^2 statistic. Show that it is invariant under linear transformation. *10*

(e) Briefly describe the classification problem. Obtain Bayes' classification rule. *10*

Q10. Answer any ***two*** of the following :

(a) Let $Z = \begin{pmatrix} X \\ Y \end{pmatrix} \sim N_4(\mu, \Sigma)$, in which X and Y are sub-vectors, each of order 2×1, with

$$\mu = \begin{pmatrix} 3 \\ 1 \\ 2 \\ -1 \end{pmatrix} \text{ and } \Sigma = \begin{pmatrix} 3 & 0 & 5 & -2 \\ 0 & 3 & -1 & 4 \\ 5 & -1 & 1 & 2 \\ -2 & 4 & 2 & 1 \end{pmatrix}.$$

Find (i) $E(X \mid Y)$, and (ii) Dispersion matrix of X given $Y = y$. *25*

(b) Explain how you will test $H_0 : \mu = \mu^{(0)}$ in $N_p(\mu, \Sigma)$ when Σ is known. Height and weight of five male college students are given. Assuming the sample is from a bivariate normal with covariance matrix $\Sigma = \begin{pmatrix} 60 & 110 \\ 110 & 265 \end{pmatrix}$, test whether the population mean vector is $(166 \;\; 68)'$. *25*

Height :	170	172	168	159	169
Weight :	73	65	62	63	70

(c) What are principal components ? Obtain their relation with eigenvalues and eigenvectors of the dispersion matrix. The following are measurements on the test scores (X, Y) of 6 candidates for two subject examinations :

(50, 55), (62, 92), (80, 97), (65, 83), (64, 95), (73, 93)

Determine the first principal components for the test scores, by using Hotelling's iterative procedure. *25*

χ^2-DISTRIBUTION : VALUES OF $\chi^2_{\alpha \cdot \nu}$

ν \ α	0·995	0·99	0·975	0·95	0·05	0·025	0·01	0·005
1	0·000	0·000	0·001	0·004	3·841	5·024	6·635	7·879
2	0·010	0·020	0·051	0·103	5·991	7·378	9·210	10·597
3	0·072	0·115	0·216	0·352	7·815	9·348	11·345	12·838
4	0·207	0·297	0·484	0·711	9·488	11·143	13·277	14·860
5	0·412	0·554	0·831	1·145	11·070	12·832	15·086	16·750
6	0·676	0·872	1·237	1·635	12·592	14·449	16·812	18·548
7	0·989	1·239	1·690	2·167	14·067	16·013	18·475	20·278
8	1·344	1·646	2·180	2·733	15·507	17·535	20·090	21·955
9	1·735	2·088	2·700	3·325	16·919	19·023	21·666	23·589
10	2·156	2·558	3·247	3·940	18·307	20·483	23·209	25·188
11	2·603	3·053	3·816	4·575	19·675	21·920	24·725	26·757
12	3·074	3·571	4·404	5·226	21·026	23·337	26·217	28·300
13	3·565	4·107	5·009	5·892	22·362	24·736	27·688	29·819
14	4·075	4·660	5·629	6·571	23·685	26·119	29·141	31·319
15	4·601	5·229	6·262	7·261	24·996	27·488	30·578	32·801
16	5·142	5·812	6·908	7·962	26·296	28·845	32·000	34·267
17	5·697	6·408	7·564	8·672	27·587	30·191	33·409	35·718
18	6·265	7·015	8·231	9·390	28·869	31·526	34·805	37·156
19	6·844	7·633	8·907	10·117	30·144	32·852	36·191	38·582
20	7·434	8·260	9·591	10·851	31·410	34·170	37·566	39·997
21	8·034	8·897	10·283	11·591	32·671	35·479	38·932	41·401
22	8·643	9·542	10·982	12·338	33·924	36·781	40·289	42·796
23	9·260	10·196	11·688	13·091	35·172	38·076	41·638	44·181
24	9·886	10·856	12·401	13·848	36·415	39·364	42·980	45·558
25	10·520	11·524	13·120	14·611	37·652	40·646	44·314	46·928
26	11·160	12·198	13·844	15·379	38·885	41·923	45·642	48·290
27	11·808	12·879	14·573	16·151	40·113	43·194	46·963	49·645
28	12·461	13·565	15·308	16·928	41·337	44·461	48·278	50·993
29	13·121	14·256	16·047	17·708	42·557	45·722	49·588	52·336
30	13·787	14·953	16·791	18·493	43·773	46·979	50·892	53·672
40	20·706	22·164	24·433	26·509	55·759	59·342	63·691	66·766
50	27·991	29·707	32·357	34·764	67·505	71·420	76·154	79·490
60	35·535	37·485	40·482	43·188	79·082	83·298	88·379	91·952
70	43·275	45·442	48·758	51·739	90·531	95·023	100·425	104·215
80	51·172	53·540	57·153	60·391	101·879	106·629	112·329	116·321
90	59·196	61·754	65·647	69·126	113·145	118·136	124·116	128·299
100	67·328	70·065	74·222	77·929	124·342	129·561	135·807	140·169

For larger values of ν, the variable $\sqrt{2\chi^2} - \sqrt{2\nu - 1}$ may be used as a standard normal variable.

(d) (i) Let $X^{(1)}, X^{(2)}, \ldots, X^{(N)}$ be a random sample from $N_p(\mu, I)$ with

$$\overline{X} = \frac{1}{N} \sum_{j=1}^{N} X^{(j)}$$

Derive the distribution of

$$\sum_{j=1}^{N} (X^{(j)} - \overline{X})(X^{(j)} - \overline{X})'.$$

(ii) Let $V = \begin{pmatrix} X & Y \\ Y & Z \end{pmatrix}$ follow Wishart distribution $W_2(n, I)$. Obtain the distribution of $U = \frac{1}{5}X + \frac{4}{5}(Y + Z)$. *25*

SECTION F

(Design and Analysis of Experiments)

Q11. (a) Assuming 'p' as a prime number, consider

$$L_j = \begin{bmatrix} 0 & 1 & 2 & \dots & p-1 \\ j & j+1 & j+2 & \dots & j+p-1 \\ 2j & 2j+1 & 2j+2 & \dots & 2j+p-1 \\ \dots & \dots & & \dots & \\ (p-1)j & (p-1)j+1 & & \dots & (p-1)j+p-1 \end{bmatrix}$$

and prove that L_j is a Latin Square. *10*

(b) (i) The following is a replicate of Z^5 experiment in a block size of Z^3 :

Block I	1	bc	de	bcde	abd	acd	abc	ace
Block II	d	bcd	e	bce	ab	ac	abde	acde
Block III	b	c	bde	cde	ad	abcd	ae	abce
Block IV	a	abc	ade	abcde	bd	cd	be	ce

Identify the treatment combinations (including generalized interaction) confounded in the 4 blocks above. *4*

(ii) Five factors, each at 2 levels are required to be tested and to control heterogeneity, it is desired that each block should contain 8 plots only. Obtain a balanced arrangement in 5 replications so that the main effect and first order interactions remain unconfounded while $1/5^{th}$ of information on each of 2^{nd} and 3^{rd} order interactions are lost using the following scheme of confounding : *6*

Replicate – I	ABC	BDE	ACDE
Replicate – II	BCD	ADE	ABCE
Replicate – III	ACD	BCE	ABDE
Replicate – IV	ABD	ACE	BCDE
Replicate – V	ABE	CDE	ABCD

(c) Write down Yates' algorithm for analysing 2^n factorial experiments. *10*

(d) Mentioning the utility of Duncan's Multiple Range (DMR) Test, list out the steps in DMR test. State the guidelines for inference using the results of DMR test. *10*

(e) Discuss Analysis of Covariance *5+5*

(i) For RBD

(ii) With two ancillary variates.

Q12. Answer any ***two*** of the following :

(a) Analyse the following results of LSD and draw your observations taking $F_{4,12}(0{\cdot}025) = 4{\cdot}12$; $t_{0{\cdot}005,12} = 3{\cdot}055$. *25*

The objective of the experiment is to study the consumption of petrol by 5 different makes (A, B, C, D, E) of cars using LSD. Here,

Treatments	:	Makes of cars – A, B, C, D, E
Drivers	:	D_1, D_2, D_3, D_4, D_5 taken as rows
Speeds	:	25 kmph, 35 kmph, 50 kmph, 60 kmph, 70 kmph taken as columns.

The results obtained are,

SS due to Speeds	=	80·73
SS due to Drivers	=	1·36
SS due to Makes	=	65·42
Total SS	=	153·94

Mean no. of km/gallon for different makes are

A : 18·42, B : 16·98, C : 16·84, D : 15·18, E : 19·98.

(b) The following are 3 replications each consisting of a complete block of a 2^r-experiment :

Rep-1

Block-1	a ab
Block-2	1 b

Rep-2

Block-1	b ab
Block-2	1 a

Rep-3

Block-1	1 ab
Block-2	a b

Identify the treatment effects confounded in each of the replicates and outline its ANOVA. *25*

(c) A 6 × 6 LSD was used to compare 6 different types of legume. Yields (in 10 gm units) are given in the following table. Two figures are missing. Analyse the data to find if there is any difference among the legumes. *25*

B 220	F 98	D 149	A 92	E 282	C 160
A 74	E 238	B *	C 228	F 48	D 168
D 188	C 279	F 118	E 278	B 176	A *
E 295	B 222	A 64	D 104	C 213	F 163
C 187	D 90	E 242	F 96	A 66	B 188
F 90	A 124	C 195	B 109	D 79	E 211

(d) Distinguish between split and strip plot designs and outline the analysis of a split plot design. *25*

SECTION G
(Computing with C and R)

Q13. (a) Define a function to evaluate

$$\text{SUM} = \sin x + \frac{1}{3}\sin 3x + \frac{1}{5}\sin 5x + \ldots + \frac{1}{(2n-1)}\sin(2n-1)x$$

and develop a main program in C to print SUM for n = 20 and for all values of x = O(0·1)3. *10*

(b) Write a C-program to find the root(s) of the equation $e^x - 3x = 0$ using Newton-Raphson procedure. *10*

(c) Assuming that a r.v. follows Weibull (b : scale, c : shape) distribution, whose density function is given by

$$\frac{c}{b}\, x^{c-1} \exp(-x^c/b),\ x > 0,\ b, c > 0.$$

Write a C-program to find the estimates of b and c, applying method of moments. *10*

(d) Write an R-code to verify whether the given matrix is skew-symmetric or not. *10*

(e) Given the data (X_i, Y_i), i = 1(1)n, write an R-program to fit the model $Y_i = a + bX_i + c_i$ based on least-squares method. *10*

Q14. Answer any ***two*** of the following :

(a) Given $A_{m \times n}$, write a C-program to find $B = (A'A)^{-1}$ and to verify whether B is an idempotent matrix or not. *25*

(b) Write a C-program to evaluate $\int_a^b f(x)\,dx$ using Weddle's rule and main program to cover the area under

$$\frac{2}{\sigma} x \exp(-x^2/\sigma) \text{ for } x \in [t, \infty[$$

Interpret the area computed in the language of Statistics. *25*

(c) Create the structure to store employees in the following format :

(1) Employee's Number

(2) Employee's Name

(3) Employee's Pay

(4) Date of Joining

to implement the following decision rule :

Decision rule : To increase the pay as per the following guidelines :

$$\text{Pay} = \begin{cases} \leq ₹\ 20{,}000/- & \text{increase pay by 15\%} \\ \leq ₹\ 50{,}000/- & \text{but} > ₹\ 20{,}000/- \quad \text{increase pay by 10\%} \\ > ₹\ 50{,}000/- & \text{no increment} \end{cases}$$

Design a C-program to execute the above decision rule for 2,000 employees of an organization. Also, for short-listing the selections of managerial candidates for internal promotions, write a module in the above program to list out the candidates who have more than 20 years of service. *25*

(d) In usual notations, considering the following linear model, write an R-program to find the estimates of the parameters *25*

$$Y_1 = 2\beta_1 + 3\beta_2 + \varepsilon_1$$

$$Y_2 = 3\beta_1 + 4\beta_2 + \varepsilon_2$$

$$Y_3 = 4\beta_1 + 5\beta_2 + \varepsilon_3$$