50 Papers theoretical (Th) practical (Pr)	Broad subject indications of papers	1971 Issue May (pages marked 'M')	1971 Issue Nov. (pages marked 'N')
(1)	(2)	(3)	(4)
	(Th) Official Statistics: Descriptive Statistics	1- 2	1- 2
I	(Th) Probability Theory: Statistical Methods	3- 4	3- 4
II	(Th) Sample Surveys: Design & Analysis of Experiments	5- 6	5- 6
V	(Th) Applied Statistics : (2 theor.½ papers on any two of)	(7-10)	(7- 9)
	Gr.a: Economic Statistics	7	7
50 marks for each group	Gr.b: Statistical Quality Control	8	8
	Gr.c: Statistical Methods in Genetics	n.r.c.	n.r.c.
	Gr.d: Vital Statistics & Demography	9	9
	Gr.e: Educational & Psychological Statistics	10	10
	(Pr) Methods of Numerical Computation; Descriptive Statistics; Official Statistics	(11-14)	(10-13)
I	(Pr) Statistical Methods; Design & Analysis of Experiments; Sample Surveys	15-19	14-17
II	(Pr) Applied Statistics: (2 pract.½ papers on any two of)	(21-26)	(18-23)
50 marks for each group	Gr.a: Economic Statistics	21-22	18-19
	Gr.b: Statistical Quality Control	22-24	19-21
	Gr.c: Statistical Methods in Genetics	24	n.r.c.
	Gr.d: Vital Statistics & Demography	24-25	21-22
	Gr.e: Educational & Psychological Statistics	25-26	22-23
	SUBJECTS OF SPECIALISATION (Papers VIII to X)		
III	(Th) Specialisation group - First Paper (I)	(27-34)	(24-27)
First paper for the subject of specialisation under different groups (theoretical)	Gr.A: Economic Statistics - Econometrics	n.r.c.	n.r.c.
	Gr.B: Techno-commercial Statistics - S.Q.C.	27-28	24-25
	Gr.C: Biometry - Biometric Methods	n.r.c.	n.r.c.
	Gr.D: Design & Analysis of Expts.-Statistical Aspects	n.r.c.	n.r.c.
	Gr.E: Sample Surveys - Theoretical Aspects	29-31	26-27
	Gr.F: Techniques of Computation- Numerical Analysis	n.r.c.	n.r.c.
	Gr.G: Statistical Inference - General Theory	31-32	n.r.c.
	Gr.H: Probability Theory - Basic Probability	32-34	n.r.c.
IV	(Th) Specialisation group - Second Paper (II)	(35-46)	(28-35)
Second paper for the same group subject of specialisation (theoretical)	Gr.A: Economic Statistics- Ind.Econ. & Planning	35-36	n.r.c.
	Gr.B: TCS - Soc. I: Operat. Research (OR)	36-38	28-30
	(Alternative) : Elements of Book-keeping and Accountancy	39-41	32-33
	- Soc.II: Methods in Business(SCB)	38-39	31
	Gr.C: Biometry - Statistical Methods in Genetics	41-42	n.r.c.
*Practical numerical computation work with desk calculating machines)	Gr.D: Design & Analysis of Expts.-Combinatorial Aspects	n.r.c.	n.r.c.
	Gr.E: Sample Surveys- Organisational Aspects	43-44	34-35
	Gr.F: Techniques of Computation - Practical*	45-46	n.r.c.
	Gr.G: Statistical Inference - Special Topics	n.r.c.	n.r.c.
	Gr.H: Probability Theory - Limit Distributions	n.r.c.	n.r.c.
	(Pr) Specialisation group - Third Paper (III)	(47-56)	(36-42)
Third Paper covering for each subject group of specialisation, topics of papers VIII, IX & X Practical) ** Practical numerical computation with Unit Record Machines)	(Practical paper based on topics of special papers I and II)		
	Gr.A: Economic Statistics - Practical Paper	n.r.c.	n.r.c
	Gr.B: TCS - Soc.I : SQC	47-48	36-37
	Soc.II (alt: O.R./E.B.A.	48-49	38
	Soc.III : S.M.B.	49	39
	Gr.C: Biometry - Practical Paper	50-51	n.r.c.
	Gr.D: Design & Analysis of Expts.- Practical Paper	n.r.c.	n.r.c.
	Gr.E: Sample Surveys - Practical Paper	51-53	40-42
	Gr.F: Techniques of Computation - Practical Paper**	53-56	n.r.c.
	Gr.G: Statistical Inference - Practical Paper	n.r.c.	n.r.c.
	Gr.H: Probability Theory - Practical Paper	n.r.c.	n.r.c.

n.r.c. = No registered candidate and hence no paper available

INDIAN STATISTICAL INSTITUTE

Statistician's Diploma Examination - May 1971

Paper I : Official Statistics and Descriptive Statistics (Theoretical)

Time : 4 hours Full marks : 100

Figures in the margin indicate full marks

Group A : Official Statistics

(Attempt any three questions from this group)

1. Describe the broad groups of Statistical systems prevailing in different countries. Discuss the merits and demerits of the Indian Statistical system. (6+10)=16

2. Write short notes on any two of the following :

   i) U.N. Statistical Commission

   ii) Functions of the Office of the Registrar General, Ministry of Home Affairs

   iii) Rôle of the Central Statistical Organisation in dissemination of statistical information (5+9)=16

3. Describe the publications dealing with Housing Statistics. What rôle has NSS played in collecting housing statistics ? (10+6)=16

4. Describe the main aspects of educational statistics on which data are available in India. Mention the publications and the collection system. (8+3)=16

5. Describe the information available in India on agricultural production with respect to both its extent and coverage. (16)

   MEANNESS (2)

Group B : Descriptive Statistics

(Attempt any three questions from this group)

6.(a) The values of the arithmetic mean and the standard deviation of the following frequency distribution of a continuous variable, obtained by using an arbitrary working origin and an arbitrary scale are 135.3 lb. and 9.6 lb., respectively. Determine the actual class-intervals.

u	-4	-3	-2	-1	0	1	2	3	Total
f	2	5	9	13	22	13	8	4	80

(b) For the data on 250 heights, the following values were obtained :

mean = 54 inches

standard deviation = 3 inches

$\beta_1 = 0$,

$\beta_2 = 3$,

It was later found on scrutiny, that the two items 64 in. and 50 in. in the original data had been wrongly written in place of the correct values 62 in. and 52 in., respectively. Determine the correct values of the above four frequency constants. (6+10)=16

Please Turn Over

7. Derive the general differential equation which generates the Pearsonian system of frequency curves.

Specify the conditions under which the different individual types are obtained.

Solve the differential equation to obtain the frequency function of one of these types.

Discuss how you will estimate the parameters involved in that type.

(5+5+3+3)=16

8. "As a measure of the association between two variables, the correlation coefficient has its limitations. But it may be considered a perfect measure in case the variables are distributed in the bivariate normal form". Discuss.

Let $x' = x + \epsilon_x$ and $y' = y + \epsilon_y$, where ϵ_x and ϵ_y are observational errors that are independent of x and y and also independent of each other. If the correlation coefficient between x and y is ρ, and that between x' and y' is ρ', show that

$$|\rho'| < |\rho| \qquad (10+6)=16$$

9. What tests will you use to judge any formula for the calculation of an index number of prices ?

Examine the formulae of Laspeyres, Paasche and Fisher and the weighted geometric average of price relatives, in terms of these tests. (8+8)=16

10. Why is analysis of economic time series considered important? Explain with suitable examples.

Describe the different methods you know to evaluate the seasonal variation of a time series. Examine their relative merits.

(6+10)=16

HEATHERS (2)

INDIAN STATISTICAL INSTITUTE

Statistician's Diploma Examination - May 1971

Paper II: Probability Theory and Statistical Methods (Theoretical)

Time: 4 hours | Full marks: 100

Figures in the margin indicate full marks

Group A : Probability Theory

(Attempt any four questions from this group)

1.(a) If A, B and C denote three events, then show that

$$P(A \cup B \cup C) = P(A)+P(B)+P(C)-P(AB)-P(AC)-P(BC)+P(ABC)$$

(b) Forty-eight bulbs are in stock at a shop and three of them are bad. If a customer picks up five bulbs at random, what is the probability that he picks up two bad bulbs? (5+7)=12

2. If X has a Binocial distribution B(n,p), then

i) find the moment generating function of

$$Y = \frac{X - np}{\sqrt{(npq)}}, \text{ where } q = 1 - p$$

ii) Find the β_1 and β_2 of Y and show that as $n \to \infty$ (p remaining fixed), these converge respectively to 0 and 3. (5+7)=12

3. If X_1 and X_2 have a joint probability density function

$$f(x_1, x_2) = 2 \quad 0 < x_1 < x_2 < 1$$

$$= 0 \quad \text{elsewhere,}$$

then,

i) find the marginal probability density functions of X_1 and X_2 and check whether they are statistically independent.

ii) find the conditional mean of X_1 given $X_2 = x_2$

iii) find the conditional variance of X_2 given $X_1 = x_1$. (3+3+6)=12

4.(a) If X is a random variable such that $E(X) = 3$ and $E(X^2) = 13$, then use Chebychev's inequality to determine a lower bound for the probability $P(-2 < X < 8)$.

(b) $X_1, X_2, \ldots$ are independent random variables, each subject to the following distribution :

$$P(X=x) = (1-\theta).\theta^x; \quad 0 < \theta < 1; \quad x = 0, 1, 2, \ldots$$

Check whether $\frac{1}{n}(X_1 + \ldots + X_n)$ converges in probability, to any limit as $n \to \infty$. (6+6)=12

5. If $x_1, x_2, \ldots, x_n$ is a random sample from a $N(\mu, \sigma^2)$ distribution, then show that $\bar{x}$ and s^2 defined by

$$\bar{x} = \frac{1}{n}\sum_{1}^{n} x_i, \quad s^2 = \frac{1}{n-1}\sum_{1}^{n}(x_i - \bar{x})^2$$

are statistically independent. Obtain the sampling distribution of $u = \sqrt{n}(\bar{x} - \mu)/\sigma$ (12)

WEARNESS (2)

Please Turn Over

Group B : Statistical Methods

(Attempt any three questions from this group)

6.(a) Explain the method of maximum likelihood and minimum chi-square in estimation and discuss their merits and demerits.

(b) Obtain the sufficient statistic for α for the distribution

$$f(x; \alpha) = e^{-(x-\alpha)} \quad \alpha \leq x \leq \infty ;$$

hence find the best unbiased estimate of α . (8+8)=16

7.(a) Explain the concept of a confidence interval as developed by Neyman and discuss with examples how in many situations the percentage points of the normal distribution can be utilised in determining confidence intervals in large samples.

(b) Obtain a 95% confidence interval for the parameter of a Poisson distribution on the basis of a sample of size n. (12+4)=16

8.(a) Give an outline of Neyman's formulation of similar critical regions for testing a composite hypothesis with one degree of freedom.

(b) Derive a best critical region for testing the hypothesis $\mu = \mu_0$ in $N(\mu, \sigma^2)$ when σ^2 is unspecified. (10+6)=16

9.(a) Describe briefly what you mean by non-parametric inference.

(b) Describe a procedure for testing the homogeneity of two independent ungrouped samples. (8+8)=16

10.(a) How do you test for the equality of two regression coefficients, when independent samples of sizes n_1 and n_2 are available from two bivariate normal populations?

(b) For testing independence in a m x n contingency table prove that $\chi^2 \leq (n-1)\,N$ or $(n-1)N$, where N is the total frequency of the table. (8+8)=16

MATHEMATICS

INDIAN STATISTICAL INSTITUTE

Statistician's Diploma Examination - May 1971

Paper III: Sample Surveys and Design & Analysis of Experiments (Theoretical)

Time : 4 hours | Full marks : 100

Figures in the margin indicate full marks

Group A : Sample Surveys

(Attempt any three questions from this group)

1.(a) What are the different sources of errors in a sample survey?

(b) Describe briefly how these errors can be controlled.

(c) State briefly the advantages of sampling over complete enumeration. (4+10+2)=16

2.(a) Explain the procedure of stratified sampling.

(b) Giving reasons, state the conditions under which you would recommend the use of this sampling procedure.

(c) Suggest an unbiased estimate of the population mean $\bar{Y}$ under this procedure and obtain the expressions for its variance under

i) proportional

and ii) Neyman allocation.

Compare the two allocations from the point of view of efficiency of the estimate. (2+4+10)=16

3.(a) Giving an example, explain the concept of cluster sampling.

(b) A population consists of N clusters each containing M units. A simple random sample of n clusters is drawn without replacement. Is the mean of the cluster means an unbiased estimate of the population mean?

(c) Obtain the expression for the variance of the above estimate in terms of intraclass correlation and discuss its efficiency as compared to the mean of a simple random sample of nM units drawn without replacement. (4+2+10)=16

4.(a) What is the ratio method of estimation?

(b) Stating clearly the assumptions made, obtain the approximate expression for the bias in a ratio estimate and derive the condition for the ratio estimate to be more efficient than the sample mean. (2+14)=16

5. Write notes on any two of the following :

i) Double sampling

ii) Circular systematic sampling

iii) Indian National Sample Survey (6)

WITNESS (2)

Please Turn Over

Group B : Design & Analysis of Experiments

(Attempt any three questions from this group)

6.(a) Set up the analysis of variance table of a randomized block experiment with v-treatments in b blocks.

Denoting by t_i (i=1, 2, ...) the treatment effect of different treatments, how would you test the hypothesis that

$c_1t_1 + c_2t_2 + c_3t_3 = 2.4$, where c_1, c_2 and c_3 are known constants such that $c_1 + c_2 + c_3 = 0$?

(b) A randomized block experiment has been carried out in 4 blocks with 5 treatments A, B, C, D and E. The reading for treatment B in block No.2, appears to be missing.

Give the analysis using missing-plot technique. (8+8)=16

7. A factorial experiment has been carried out in randomized blocks with three factors K, P and N. The factors K and P are at two levels each, and the factor N is at three levels. If the experiment be carried out in 4 blocks of 12 plots each, set up the analysis of variance table. Indicate (using suitable notations) how the different sums of squares would be computed. (16)

8. A split plot experiment is planned with two factors - A (at α levels) and B(at two levels in r blocks, each block containing α plots. Set up the analysis of variance table.

Using suitable notations, indicate how the different sums of squares would be computed. (16)

9. Define a Youden square design. With practical illustrations, discuss situations where such a design is useful. Give the details of the intra-block analysis of a Youden square design, pointing out the different computational steps. (6+10)=16

10. Write notes on any two of the following :

i) Analysis of covariance

ii) Square Lattice design

iii) Partial confounding in factorial experiments

NEATNESS

INDIAN STATISTICAL INSTITUTE

Statistician's Diploma Examination - May 1971

Paper IV : Applied Statistics Group Papers (Theoretical)

Time : 4 hours | Full marks : 100

(a) Candidates will be required to answer questions from those two groups of subjects only for which they have registered their option.

(b) Separate answer books are to be used for each of the two groups attempted.

(c) Figures in the margin indicate full marks.

(d) Use of calculating machines is not permitted

GROUP A : ECONOMIC STATISTICS : Half Paper of 50 marks

(Attempt any three questions from this group)

1. Describe the different schemes for explaining the oscillations in a stationary time series. Explain the use of correlograms for discriminating between the above schemes.

(16)

2. (a) A product y is produced by the use of two factors x_1 and x_2 according to the linear homogeneous production function, $y = k\sqrt{x_1 x_2}$ where k is a constant. Given the factor prices (p_1 and p_2), find the optimum usage of the factors for minimum cost for given product y. Show that, at this optimum : $x_1 = c_1 y$ and $x_2 = c_2 y$ i.e., the factors are always combined in fixed proportions (depending on P_1 and P_2).

(b) Indicate the limitations of Leontief's open input-output models.

(8+8) = 16

3. (a) What does a Lorenz curve describe in statistical terms? What do you mean by the coefficient of concentration?

(b) Derive for the income distribution shown in either (i) or (ii) the coefficients of concentration.

i) $dF(x) = \begin{cases} \frac{\alpha}{x_0}\left(\frac{x_0}{x_1}\right)^{\alpha+1} & \text{for } x > x_0 \\ 0 & \text{for } x \leq x_0 \end{cases}$ $\alpha > 0$

ii) $dF(x) = \begin{cases} \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{\log x - \mu}{\sigma}\right)^2\right\} & \text{for } x > 0 \\ 0 & \text{for } x \leq 0 \end{cases}$ (8+8) = 16

4. What do you mean by income elasticity of demand ? How would you estimate this elasticity from family budget data? Enumerate the disadvantages of constant elasticity demand functions.

(16)

5. What is a chain index? Discuss its advantages and disadvantages over a fixed base index number. Examine the important formulae for the calculation of price index numbers in the light of various tests devised for this purpose.

(13)

Neatness : (2)

(17)

Please Turn Over

GROUP B : STATISTICAL QUALITY CONTROL : Final paper of 50 marks

(Attempt any three questions from this group)

1.(a) What is meant by specification limits and control limits? Does a process in statistical control, ensure that all the product will be within specifications? Justify your statement by means of an example.

(b) Define process capability when the process shows a state of control. Discuss in detail the various courses of action open to you under the following situations :

(i) The process capability is greater than the specified tolerance

(ii) The process capability is approximately equal to the specified tolerance

(iii) The process capability is less than the specified tolerance

(c) When only a single specification limit is given, discuss its relationship with the process capability and the actions that can be taken under different settings of the machine.

(4+8+4) = 16

2.(a) A company manufactures four models of radios and receives quite a few complaints from customers. In order to improve the quality of radios, the management decides to use control chart technique for the defects observed at the final testing. The number of radios manufactured daily is not constant.

Suggest a procedure, to set up control chart or charts for controlling the defects, giving clearly the assumptions made and the statistical concepts used.

(b) Discuss the situations under which the control chart for 'fraction defective' is to be preferred to the control chart for 'number defective' and compare their relative advantages or disadvantages.

(c) Describe briefly the procedure for obtaining the OC curve for a 'number defective' chart.

(8+4+4) = 16

3. Obtain an unknown sigma acceptance sampling plan by variables for the case of one-sided (upper) specification limit U, such that under the plan, lots with 100 p_1 and 100 p_2 percent defectives ($p_2 > p_1$), would be accepted with probabilities $1-\alpha$ and β respectively.

(16)

4.(a) Discuss the important features of Military Standard Sampling Plans in details for Attributes.

(b) Bring out the important differences between MIL-STD 105B and MIL-STD 105D.

(10+6) = 16

Footnote : (2)

(X)

GROUP C : STATISTICAL METHODS IN GENETICS : Half paper of 50 marks

No candidate available

GROUP D : VITAL STATISTICS AND DEMOGRAPHY : Half-paper of 50 marks

(Attempt any three questions from this group)

1.(a) Distinguish between census and registration records as sources of demographic data. Define and illustrate a rate and a ratio in connection with vital events.

(b) Indicate the appropriate death rates for the comparison of mortality (i) between two communities, (ii) between two periods of time, (iii) from two causes of death and (iv) in two occupations.

(6+10) = 16

2.(a) Explain differences between the central death rate (m_x), probability of death (q_x) and force of mortality (μ_x) in a life table. Describe inter-relations among them.

(b) Under necessary assumptions (to be clearly stated) obtain a mathematical relation between the net reproduction rate, the rate of natural increase and the length of a female generation.

(8+8) = 16

3. Derive the equation/logistic curve for a population within a limited area, interpreting the parameters that occur in the equation.

Describe one method of estimating these parameters.

Indicate possible generalisations of the logistic law for population prediction.

(7+7+2) = 16

4. Write notes on any two of the following :

(a) infant mortality rate,

(b) Makeham's law of graduation,

(c) morbidity incidence rate.

(8+8) = 16

Footnotes : (2)

Please Turn Over

GROUP B : EDUCATIONAL AND PSYCHOLOGICAL STATISTICS Multi-paper of 3C

(Attempt any three questions from this group)

1.(a) Why is it considered desirable to convert gross scores to some standard scores? Define 'standardized scores' and 'normalized scores' and describe how they are derived.

(b) Explain the concept of percentile scale, and describe a practical method of its computation from raw data.

(16)

2.(a) Define 'reliability' and 'validity' of a test. Give an account of the available methods of assessing the reliability of a test, commenting on their merits and demerits.

How would you assess the validity of a test ?

(b) Describe briefly the rationale and the technique of item analysis. Explain the similarities and differences between (i) item difficulty and item discrimination and (ii) achievement tests and aptitude test.

(16)

3.(a) Give an account of the different classificatory techniques used in personnel selection and vocational guidance programmes.

(b) Define tetrachoric and biserial correlations, indicating their applications in Psychometry.

(16)

4.(a) Explain the purpose of factor analysis. Describe the centroid meth of estimating the factor-loadings of the tests included in a battery State the underlying assumptions.

(b) Give an account of the maximum-likelihood method of estimation of factor-loadings. What are factor scores and how are they estimated

(16)

5.(a) Write a note on the learning models and measurement of learning and forgetting. Indicate the method of estimation of the parameters of a suitable stochastic model for learning.

(b) Discuss the appropriateness of the available methods for measuring the goodness of fit of a stochastic model to a set of experimental data.

(16)

Neatness : (2)

INDIAN STATISTICAL INSTITUTE

Statistician's Diploma Examination - May 1971

Paper V : Methods of Numerical Computation ; Descriptive Statistics and Official Statistics (Practical)

Time : 5 hours Full marks : 100

(i) Figures in the margin indicate full marks

(ii) Use of calculating machines is permitted.

GROUP A : Methods of Numerical Computation.

(Attempt any two questions)

1.(a) The number of members of a society are as given in the following table:

Year	Number of Members	Year	Number of Members
1910	845	1915	772
1911	837	1917	757
1913	836	1918	761
1914	821	1919	793

Interpolate for the missing figures for the years 1912 and 1916.

1.(b) Given the following table:

$f(0) = 23.1234$
$f(6) = 23.7252$
$f(12) = 24.3838$
$f(18) = 25.1320$

Estimate by a suitable method of interpolation the values of $f(7)$, $f(9)$, $f(10)$ and $f(11)$.

$(6+6\frac{1}{2}) = 12\frac{1}{2}$

2. Find the value of

$$\int_0^{1.93} \frac{(1+t^2/10)^{-5.5}}{B(0.5, 5.0)\sqrt{10}}$$

by Weddle's Rule, taking the ordinates correct to four decimal places.

$(12\frac{1}{2})$

3. Solve the equations

$$A\underset{\sim}{x} = \underset{\sim}{b}$$

Where

$$A = \begin{pmatrix} 2.236 & 0.0 & 5.333 \\ 0.35 & 1.25 & 0.45 \\ 1 & 1 & 1 \end{pmatrix}$$

$$\underset{\sim}{x}' = (x_1, x_2, x_3)$$

$$\underset{\sim}{b}' = (3.598, 0.438, 0.807)$$

$(12\frac{1}{2})$

SD-V : GROUP B - Descriptive Statistics

(Attempt any three of the items in question 4)

4. (a) A car hire firm has two cars, which it hires out day by day. The number of demands for a car on each day is distributed as a Poisson variable with mean 1.5. Calculate,

(i) the proportion of days on which neither car is used and

(ii) the proportion of days on which some demand is refused.

(b) In a precision bombing attack there is a 50% chance that any one bomb will strike the target. Two direct hits are required to destroy the target completely. How many bombs must be dropped to give a 99% chance or better, of completely destroying the target?

(c) (i) For a certain normal distribution the first moment about 10 is 40 and the fourth moment about 50 is 40. What are the mean and variance of the distribution?

(ii) The mean height of soldiers is known to be 68.22 inches with a variance of 10.8 $(\text{in})^2$. How many soldiers in a regiment of 1,000 would you expect to be over 5 feet tall?

(Assume the distribution of heights to be normal).

(d) If skulls are classified as A, B and C according as the length breadth index is under 75, between 75 and 80 and over 80, find approximately (assuming that the distribution is normal) the mean and standard deviation of a series in which A are 50%, B are 20% and C are 4 %.

(3+3+6) = 12

5. For certain large cities in USA, the following values were computed with

X_1 = Crime rate, being the number of known offences per thousand of population,

X_2 = percentage of male inhabitants,

and X_3 = percentage of total inhabitants who are foreign-born.

Means	Standard deviations	Correlation coefficient
$\bar{x}_1 = 10.9$	$\sigma_1 = 7.0$	$r_{12} = +0.43$
$\bar{x}_2 = 49.2$	$\sigma_2 = 1.3$	$r_{13} = -0.31$
$\bar{x}_3 = 10.2$	$\sigma_3 = 4.0$	$r_{23} = +0.25$

(i) Obtain the linear regression equation of x_1 on the other two variables.

(ii) Evaluate also, the multiple correlation coefficient $r_{1.23}$.

(8+4) = 12

6. The total time required to bring an automobile to a stop after noticing danger is composed of the reaction time (time between recognition of danger and application of brakes) plus braking-time (time for stopping after application of brakes). The following table gives the stopping distance D(feet) of an automobile travelling at the speed V (m.p.h.) at the instant the danger is sighted:

Speed V (m.p.h)	20	30	40	50	60	70
Stopping distance D (feet)	54	90	136	208	292	390

(i.e. For the operative part of the question, see next page).

table.

6. To the speed and distance figures, given in the immediately preceding
(i) Fit a least-square parabola of the form $D = a_0 + a_1 V + a_2 V^2$ to the data
(ii) Estimate D when V = 45 m.p.h.

(6+2) = 10

7. EITHER

The following data show the different group indices with their respective weights for the middle class and menial class people of a State in 1952. Construct the cost of living indices for the two classes and compare.

(Base year 1939 = 100)

Group	Group Index: Middle class	Group Index: Menial class	Group weight: Middle class	Group weight: Menial class
Food	303.4	300.2	71.22	71.20
Clothing	503.9	301.2	4.51	2.00
Fuel and light	337.9	336.8	6.50	9.27
House rent	110.9	110.9	8.97	8.03
Miscellaneous	233.4	233.3	8.72	9.37

OR

On the basis of monthly sales (in million rupees) of a certain commodity for a certain number of years, the following calculations were made:-

Trend : $y = 25.74 + 0.15\ t$,

where origin is at January 1965; t = time unit (one month) and y = monthly sales.

Seasonal Indices:

Month	Jan	Feb	March	April	May	June
Seasonal index	79	76	95	93	103	97

Month	July	Aug	Sept	Oct	Nov	Dec
Seasonal index	86	89	103	122	113	133

Estimate the monthly sales for 1965.

(10)

SD V- GROUP C : Official Statistics

(Attempt both the questions - Wherever possible the collected data should be shown in a neat tabular form).

8. From the official publications placed at your disposal, obtain the following information:

(a) Principal targets and achievements in the first three 5-year plans in respect of:

(i) Agricultural production, (ii) Power and (iii) Technical Education in India.

(b) Decennial growth rate of population in India as a whole. Obtain these values for six of the largest States according to 1951 population.

(Figures are to be shown from 1901 onwards) (6+7) = 13

9. Collect data for the latest 5 years with regard to any five of the items (i) to (vii):

(i) Employment in Central Government and State Government establishments in India

(ii) Number of foreign tourists to India

(iii) Total yield of food-grains in India

(iv) Registered medical practitioners in India

(v) Number of recognised educational institutions in India

(vi) Expenditure on rural education in India

(vii) Number of printing presses at work and number of books published in India.

INDIAN STATISTICAL INSTITUTE

Statistician's Diploma Examination - May 1971

Paper VI - Statistical Methods; Design and Analysis of Experiments and Sample surveys.

Time : 6 hours Full marks : 100

(i) Figures in the margin indicate full marks.

(ii) Use of calculating machines is permitted.

GROUP A - Statistical Methods

(Answer any two questions from this group)

1.(a) The following represent independent samples from two normal populations. Assuming the variances to be equal, test whether the mean of the first population is three times the mean of the second.

Sample 1 : 67.2, 73.5, 61.0, 69.3, 71.6, 80.7, 69.0, 63.7, 70.8, 77.1

Sample 2 : 23.1, 21.6, 25.8, 23.2, 27.0, 26.9, 20.0, 22.7.

(b) The correlation coefficients between two characters obtained from 2 samples of sizes 28 and 35 were 0.50 and 0.30, respectively. Is difference between the two coefficients significant at the 5% level?

(12+8) = 20

2.(a) In a textile factory, 100 pieces of cloth were inspected and the number of defects in each of these pieces were recorded. The inspection data are shown in the following frequency table:

No. of defects	Frequency of cloth pieces
0	70
1	18
2	2
3	1
4 or more	0

Fit a Poisson distribution to the data and test the goodness of fit.

(b) Samples of different sizes were selected from lots supplied by four manufacturers, A, B, C and D, and the number of defective articles for each sample is tabulated below

Manufacturer	A	B	C	D
Sample size	120	200	250	400
Number of defectives	20	35	37	43

Do you consider the lots supplied by the different manufacturers to be of the same quality?

(12+8) = 20

3. The following table gives the observed frequencies of different combinations of colour and pollen shape in sweet peas in a certain experiment:

Pollen shape \ Colour	Purple	Red
Long	296 $(2+\theta)/4$	27 $(1-\theta)/4$
Round	19 $(1-\theta)/4$	85 $(\theta)/4$

The theoretical proportions are given in brackets, where $\theta = (1-p)^2$, p being the recombination fraction. Estimate p by the method of maximum likelihood and calculate the standard error of the estimate.

(20)

SD VI - GROUP B: Design and analysis of experiments

(Attempt both the questions from this group)

4. Either:

For a 3^3 design in nine blocks of three plots each, the composition of the blocks is as follows:

(000, 011, 022), (101, 112, 120), (202, 210, 221)

(001, 012, 020), (102, 110, 121), (200, 211, 222)

(002, 010, 021), (100, 111, 122), (201, 212, 220)

Determine the main effects and the interactions which have been confounded.

Or:

Construct the Balanced Incomplete Block design.

$$v = b = 21,\ r = k = 5,\ \lambda = 1$$

10

5. Either:

The following are the yields of 10 treatments in 15 blocks, the layout being a balanced incomplete block design, in blocks of 4 plots each. Perform the intra-block analysis.

Block	Treatments 1	2	3	4	5	6	7	8	9	10
1	9.7	8.7		8.4	8.0	–	–	–	–	–
2	–	9.6	8.0	–	–	8.3	–	–	–	8.0
3	–	9.0	–	7.3	–	8.0	4.3	–	–	–
4	8.3	–	8.7	–	8.0	–	8.0	–	–	–
5	10.0	–	–	7.5	–	–	–	4.2	–	8.5
6	–	9.6	–	–	–	–	5.1	8.0	8.0	–
7	–	9.0	–	–	7.4	–	–	6.4	–	8.8
8	–	–	–	–	9.4	–	6.3	–	8.1	8.0
9	9.3	9.3	8.2	–	–	–	–	–	8.3	–
10	–	–	–	8.7	9.0	8.0	–	–	8.0	–
11	9.7	–	–	–	–	8.7	8.0	–	–	8.0
12	–	–	8.3	8.1	–	–	–	–	8.7	8.1
13	9.0	–	–	–	–	7.4	–	8.4	8.0	–
14	–	–	8.0	8.3	–	–	4.0	8.0	–	–
15	–	–	9.3	–	8.3	8.3	–	3.0	–	–

OR

The following table gives the results of an experiment undertaken to determine the effect of five fertilizers each at two levels, on the yield of potatoes.

Table - Plan and yields in pounds of potatoes

Block I	Block II	Block III	Block IV
(1)	s1d	p1d	d
740	1328	844	1248
sr	spkn	nd	pknd
1134	1232	1104	1100
spk	knd	sk	k
934	1006	1150	734
s1nd	sp1d	spknd	snd
1260	1324	1503	1376
kd	pd	spa	sp1d
898	1000	1212	1353
spd	pk	kn	s1n
1234	962	1000	1376
pnd	n	sd	sp
928	700	1178	1008
p1n	s	p	pr
830	1100	888	964

Determine the confounded interactions and analyse the data.

(20)

SD-VI : GROUP C - Sample Surveys

(Attempt both the questions from this group)

1. From a region containing 1000 villages and having a total cultivable area 935000 bighas, a simple random sample of 64 villages was drawn. The next page table shows the total cultivable area and also the area under wheat, in these villages.

Obtain from the figures shown in the appended table (next page) the ratio and regression estimates of the total area under wheat in the region.

Total cultivable area and area under wheat.

Area under wheat (bighas)	Total cultivable area (bighas)				
	200	700	1200	1700	2200
100	12	6	-	-	-
300	-	10	4	2	1
500	-	4	7	3	1
700	-	-	-	2	1
900	-	-	-	1	2

Obtain the estimates of the variances of the estimates you have calculated, and compare them with the variance of the estimate you would use if for the sample, only information regarding the area under wheat in each village had come to be recorded.

(3)

The following data show the classification by size of all farms in a country. The average acreage under corn per farm ($\bar{Y}_h$) and the square root (S_h) of mean square, are given for each class.

Farm size	N_h	$\bar{Y}_h$	S_h
0-40	304	5.4	0.3
41-80	401	16.3	13.3
81-120	391	21.3	15.1
121-160	334	31.5	19.8
161-200	160	42.1	24.5
201-240	113	50.1	28.0
241-280	148	63.3	35.2

For a stratified sample of size 100 with classes as strata, compute the sample sizes in each class under:

(i) proportional allocation;
(ii) optimum allocation.

For the given population compare the precisions of these two sampling methods, with that of simple random sampling.

15

2. Suppose that there is a pond exactly in the shape of an inverted and truncated cone with smaller diameter 10 metres (at the bottom), larger diameter 50 metres (at the top) and height 4 metres. You are required to locate 15 points at random inside this pond in order to sample the water for bacteriological examination. Please note that a depth of 3 cm from the surface is to be excluded from the sample. You are to indicate the selected points by any convenient system of co-ordinates, correct to the nearest centimetre.

(3)

Draw a random sample of size 16 with replacement from a bivariate population represented by the following bivariate frequency table (next page):

x \ y	0	10	11	12	13	14	15	13	17	18	19	20	21	22	23	24	25
1	0	0	0	0	1	3	7	2	3	0	1	3	7	0	1	2	4
2	1	0	2	3	0	4	30	13	5	10	11	3	2	1	4	1	2
3	10	-	17	3	20	13	17	2	20	25	7	0	10	11	13	7	2
4	4	3	0	9	11	15	10	17	10	15	23	27	10	12	15	4	0
5	6	0	2	11	5	9	21	4	0	3	1	25	17	11	14	2	9
6	5	3	1	10	4	7	20	5	3	0	2	21	11	12	10	2	0
7	4	2	7	13	2	10	17	8	10	4	1	25	7	10	15	1	7
8	2	1	0	10	4	11	21	5	11	0	7	17	3	12	11	2	5
9	1	2	4	13	2	10	14	2	9	7	4	11	7	11	5	1	0
10	5	3	0	11	1	0	12	1	0	2	2	6	0	9	4	0	1
11	6	5	0	9	7	0	11	7	9	2	3	5	4	5	3	0	2
12	11	6	0	0	6	8	23	21	7	4	5	4	2	3	0	0	0
13	2	7	6	10	5	0	10	0	10	9	6	3	1	2	3	2	3
14	9	8	13	6	3	0	0	5	5	8	2	1	0	1	2	3	6
15	17	5	17	7	5	0	3	2	6	7	7	5	4	7	9	10	7
16	0	4	4	0	6	7	5	6	7	0	4	2	3	2	3	4	3
17	11	10	1	4	1	7	10	11	3	6	2	3	2	4	5	7	2
18	4	5	2	0	2	7	6	5	0	2	10	7	0	0	1	2	1
19	3	3	5	11	0	0	3	4	12	7	0	3	0	7	0	5	6
20	2	0	0	2	7	0	0	7	4	5	2	3	4	3	4	3	2
21	5	11	2	11	0	0	4	3	2	4	3	4	2	0	1	2	1
22	7	0	3	9	5	0	5	2	1	3	0	3	1	2	3	4	5
23	0	0	4	7	2	5	2	1	3	1	3	2	3	1	2	1	2
24	0	10	3	3	11	5	3	2	5	2	1	2	0	0	1	1	0
25	1	0	3	2	3	5	4	5	0	2	1	0	2	0	1	0	2

INDIAN STATISTICAL INSTITUTE

Statistician's Diploma Examination - May 1971

Paper VII : Applied Statistics Groups (Practical)

Time : 5 hours Full marks : 140

(a) Candidates will be required to answer questions from these two groups of subjects only, for which they have registered their options.

(b) Separate answer books are to be used for each of these two groups attempted.

(c) Figures in the margin indicate full marks.

(d) Use of calculating machines is permitted.

GROUP A : ECONOMIC STATISTICS Full marks (70 marks)

(Attempt all questions from this group)

1. The table below gives the family budget data for a few samples of four income classes of families of a country, for a certain year.

Items	Yearly income per consumer unit in rupees			
	below 600	600-750	750-1050	above 1050
(1)	(2)	(3)	(4)	(5)
1. Number of households in the sample	136	179	111	22
2. Average number of consumer units per household	2.30	2.57	2.50	2.43
3. Average income per consumer unit (Rs.)	543.1	681.3	881.9	1232.8
4. Average expenditure on food per consumer unit (Rs.)	201.8	231.6	274.4	407.1

Calculate the income elasticity of demand of food, assuming that the demand function has constant elasticity.

(25)

2. From the following table showing the monthly receipts (in suitable units) of State Governments in India, obtain measures of seasonal variations.

Year \ Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1952	23	39	82	17	18	16	20	17	12	22	20	13
1953	25	28	105	20	22	20	23	18	23	29	15	18
1954	32	36	93	21	21	22	29	21	15	27	27	21
1955	32	42	99	24	24	23	29	24	21	32	28	21

(25)

3. From the input-output table given below, ascertain what would have happened if export from the manufacturing sector had been 10 per cent higher than what actually were. (No changes are to be assessed in the demand of final buyers and in the rate of indirect taxes.)

Input-output table of the United Kingdom for 1950
(figures in million pounds sterling)

Purchases (m £) / Sales (m £)	Agri.& mining (1)	Mnfg. inds. (2)	Other inds. (3)	Exports (4)	Govt (5)	Households (6)	Capital (7)	Total output (8)
1. Agriculture and mining	-	810	211	57	12	550	-5	1,041
2. Manufacturing industries	301	-	934	1,030	124	3,510	810	6,001
3. Other industries	199	1,010	-	512	1,044	3,011	752	6,007
4. Imports	64	1,475	250	115	80	747	52	2,702
5. Indirect taxes less subsidies	-11	701	314	-	11	400	21	1,535
6. Income and depreciation	1,035	3,040	6,379	-	-	-	-	11,354
7. Total input	1,041	6,001	6,007	2,713	2,171	6,217	1,630	33,530

(25)

SDVII- GROUP B : STATISTICAL QUALITY CONTROL Half paper (30 marks)

1. Data on roller diameter were obtained on 25 samples each of size four, after a certain operation. The following table gives the observed values

Sample No.	Roller diameter in inch on piece			
	1	2	3	4
1	0.1005	0.1035	0.1007	0.1070
2	0.1070	0.1070	0.1000	0.1007
3	0.1005	0.1055	0.1057	0.1070
4	0.1073	0.1000	0.1070	0.1070
5	0.1005	0.1001	0.1000	0.1000
6	0.1077	0.1070	0.1074	0.1077
7	0.1001	0.1079	0.1076	0.1070
8	0.1001	0.1002	0.1001	0.1007
9	0.1001	0.1002	0.1005	0.1001
10	0.1000	0.1007	0.1002	0.1001
11	0.1005	0.1090	0.1007	0.1001
12	0.1000	0.1007	0.1000	0.1001
13	0.1070	0.1075	0.1072	0.1072
14	0.1000	0.1079	0.1070	0.1000
15	0.1070	0.1071	0.1073	0.1074
16	0.1070	0.1005	0.1000	0.1070
17	0.1000	0.1000	0.1070	0.1079
18	0.1000	0.1000	0.1090	0.1074
19	0.1070	0.1071	0.1073	0.1000
20	0.1000	0.1070	0.1002	0.1002
21	0.1077	0.1071	0.1001	0.1070
22	0.1077	0.1073	0.1071	0.1072
23	0.1070	0.1071	0.1070	0.1073
24	0.1073	0.1073	0.1075	0.1000
25	0.1070	0.1071	0.1077	0.1070

The specification for the roller diameter is $0.1875 \pm \begin{matrix} 0.0040 \\ -0.0010 \end{matrix}$

With reference to the observed value of the roller diameter given in the table on the previous page:

(i) Set up control charts for X and R
(ii) Find the process capability
(iii) Give comments

(15+5+5) = 25

2. A double sampling plan is given by

$$N = 2000$$
$$n_1 = 100 \quad n_2 = 150$$
$$c_1 = 1 \quad c_2 = 4$$

The lots rejected by the plan are 100% inspected and all the defectives found are replaced by good ones.

(a) Draw OC curve for this plan, calculating at least 8 points and stating the approximations used if any.

(b) Draw the ASN curve for the same plan.

(c) Find the average amount of inspection per lot for lots having 3% defectives.

(12+8+5) = 25

3. An organic chemical used in the manufacture of certain dyestuffs is produced as a paste containing 30% - 50% solids. The yield of a batch is estimated by weighing the batch and analysing a sample. Considerable variation was found from batch to batch, and before investigating the reason for the variation it was necessary to investigate the errors of sampling and testing. For this purpose, 9 consecutive batches from normal production were sampled. From each batch three random samples of paste were taken. Each sample was thoroughly mixed and divided into two sub-samples. This gave 54 samples which were coded and sent for analyses in a random order. The test results taken as deviation from 40% are given in the following table:

Batch No	Percent strength of samples of paste					
	Sample I		Sample II		Sample III	
	sub-sample 1	sub-sample 2	sub-sample 1	sub-sample 2	sub-sample 1	sub-sample 2
1	10.3	9.8	10.1	9.5	11.1	9.4
2	5.8	5.4	4.4	4.7	4.7	4.4
3	1.0	1.4	2.7	1.6	3.1	3.3
4	8.7	10.0	8.0	10.4	7.9	8.5
5	8.9	9.4	8.4	6.8	6.5	5.4
6	7.0	6.1	7.4	7.2	5.1	7.5
7	6.3	5.0	4.6	4.0	5.6	4.2
8	4.9	2.3	5.1	3.4	3.3	1.6
9	15.7	15.4	16.3	16.3	15.1	15.0

Analyse the data suitably and estimate the sampling and testing errors. Give a suitable sampling and testing procedure so as to estimate the true strength of a batch within 1% of its true value, with 95% confidence.

(25)

5. An industrial experimenter wishes to compare the effects of 4 types of grids, A, B, C & D on the vacuum of radio tubes. Scaling machines and operators are two factors which may affect vacuum. A Latin square experiment was conducted and following coded data obtained. (next page)

Effects of types of grids on the vacuum of radio tubes.

Operator	Sealing Machines 1	2	3	4
1	A 5.5	C 2.9	D 9.6	B 6.8
2	C 6.2	D 6.5	B 5.2	A 9.2
3	D 8.9	B 2.9	A 5.8	C 6.6
4	B 6.1	A 5.1	C 2.8	D 9.8

Analyse the data and draw suitable conclusions.

SD VII - GROUP C : GENETICAL LINKAGE [illegible]

1. The following table gives the segregation for two factors R-r and T-t in a back cross for Poppy.

Test for evidence of linkage; estimate the recombination fraction and its standard error.

	RT	Rt	rT	rt
Observed frequency	191	37	33	203

(1

2. The following table gives the mean squares and mean products between and within sires for three characters (1) number of eggs (2) egg weight in gm (x_2) and (3) time period of maturity in days (x_3) for poultry. The data pertain to 11 randomly selected daughters of each of 15 cocks of breed, the matings being at random.

	d.f.	Mean square and mean products (x_1^2)	(x_2^2)	(x_3^2)	$(x_1 x_2)$	$(x_1 x_3)$	$(x_2$
Between sires	14	432.6	70.15	2073	40.41	-131.7	50.
Within sires	135	337.3	12.14	1532	1.03	-178.6	1.

Assuming the set of weights, $a_1 : a_2 : a_3 = 7 : 10 : 0$, find the best linear discriminant function for selection. Compare its efficiency with sim selection.

3. The following distribution was observed in a simple random sample fr a population studied for A-B-O blood groups:

Type:	O	A	B	AB
Frequency:	100	70	90	22

Estimate the gene frequencies by maximal likelihood method and test whether the data conform to Bernstein's hypothesis.

SD VII - GROUP D : VITAL STATISTICS AND DEMOGRAPH

1. Determine (a) the net reproduction rate and (b) the nuptial net rep rate for the female population presented in the table below:

Age group	Female Fertility rate	Survival factor	Proportion married wo
15-19	.0156	.925	.118
20-24	.1428	.914	.626
25-29	.1109	.902	.809
30-34	.0327	.891	.916
35-39	.0557	.878	.920
40-44	.0133	.865	.927

Hence determine approximately the length of the female generation an rate of growth of the female population.

2. Calculate the complete expectation of life at ages 11-15 years, given that it is 29.72 years at age 10, when $l_{10} = 71,503$ and 1000 q_x has values as follows:

x (age in years)	1,000 q_x	x (age in years)	1,000 q_x
10	2.01	13	3.93
11	2.52	14	4.68
12	3.20	15	5.41

3. The following table gives the size of the Indian population as recorded at each of the past eight decennial censuses. Determine the equation to the logistic curve that would fit these data, preferably by using Rhodes' method.

Census year	Population (millions)	Census year	Population (millions)
1901	238.3	1941	310.5
1911	252.0	1951	351.0
1921	251.2	1961	439.1
1931	278.0	1971	

(25)

SD VII - GROUP I : EXPERIMENTAL AND PSYCHOLOGICAL STATISTICS (50)

1. A number of students were examined in a subject by three examiners, E_1, E_2 and E_3 independently. The standards of marking of the examiners are reflected in the percentage frequency distribution of scores given below:

Marks	Percentage frequency distribution of E_1	E_2	E_3
0-10	5	10	5
10-30	15	20	25
30-50	50	60	50
50-70	21	6	10
70-90	5	2	0
90-100	1	-	2

Determine the relative ranks of the three students A, B, C who have scored the following marks with the three examiners E_1, E_2, E_3.

Students	Marks given by E_1	E_2	E_3
A	25	62	73
B	48	51	35
C	73	25	50

2. The following information on a test is available:

Item No.	Difficulty value	Item No.	Difficulty value
1	.70	11	.60
2	.60	12	.44
3	.65	13	.52
4	.75	14	.40
5	.40	15	.32
6	.42	16	.30
7	.35	17	.21
8	.80	18	.23
9	.74	19	.25
10	.36	20	.70

The standard deviation of the total score on the test is 4.5. Find the test reliability by using (i) the difficulty values of items and standard deviations of the total score, and (ii) mean and sd. of the total scores.

(P.T.O.)

3. Twelve subjects were independently ranked by three judges in respect a trait. The results are given below:

Judges	Subjects											
	1	2	3	4	5	6	7	8	9	10	11	
I	6	1	4	2	7	10	12	11	9	8	3	
II	7	3	5	1	5	12	8	9	10	11	2	
III	4	1	6	2	5	10	12	8	11	9	3	

Test whether the rankings by the three judges can be regarded as concu

In case they are concordant, combine the three rankings in a suitable and assign final ranks to the subjects.

4. The following table gives the matrix of correlations for a battery of four tests:

Variable	x_1	x_2	x_3	x_4
x_1	1.00	.50	.35	.37
x_2	.50	1.00	.47	.48
x_3	.35	.47	1.00	.33
x_4	.37	.48	.33	1.00

Calculate the loadings of the first two factors using a suitable factor analysis method.

5. A test has 40 items ($N = 40$). Each item has five alternative answers ($n = 5$). When the test was administered to 30 students, each student attempted and answered all the items. The number of right answers (r) the number of wrong answers (w) equals the number of items (N). The following table gives the values of R for the 30 students.

Sl.No.	R	Sl.No.	R	Sl.No.	R
1	10	11	12	21	12
2	14	12	8	22	6
3	12	13	33	23	10
4	10	14	17	24	7
5	9	15	20	25	15
6	13	16	14	26	27
7	6	17	17	27	21
8	10	18	9	28	8
9	10	19	22	29	10
10	21	20	11	30	14

For each student, complete his score corrected for chance guessing.

Write brief notes on the assumptions and interpretations of corrected scores.

INDIAN STATISTICAL INSTITUTE

Statistician's Diploma Examination - May 1971

Paper VIII : Subjects of Specialisation - I

Time : 4 hours Full marks : 100

(a) Candidates will be required to answer questions from that group only for which they have registered their options.

(b) Figures in the margin indicate full marks.

GROUP A : ECONOMIC STATISTICS

Special Paper I - Econometrics

(No candidate available)

GROUP B : TECHNO-CHEMICAL STATISTICS

(Special Paper I - Statistical Quality Control)

Attempt any five questions from this group

1. (a) What is the Average Run Length (ARL) of a Control Chart ? In a Control Chart for a sample statistic, (using only an upper control limit) the decision rule adopted is to take action, if λ successive points fall beyond the Control limit. If P is the probability of a point violating the control limit, show that

$$ARL = \frac{(1-p^{\lambda})}{p^{\lambda}(1-p)}$$

(b) For $n = 4$, $\sigma = 0.5$, $\lambda = 2$ and $ARL = 200$, Calculate the upper control limit for the range chart (12+8)=20

2. (a) Describe the operating procedure for the one-sided decision interval CUSUM Scheme. Briefly indicate how to choose the appropriate reference value and the decision interval for a normal variate.

(b) Prove the equivalence of the V-mask procedure and two one-sided decision interval criteria. (12+8)=20

3. (a) What are tolerance limits ? Explain the difference between the tolerance limits and confidence limits, by citing practical examples from any industry.

(b) Explain, how you would construct tolerance limits such that it can be said with 100β percent Confidence that they will include atleast 100γ percent of the population assuming, either a normal parent population with known mean 'm' or a parent population with an unknown form. (5+12)=20

4. (a) Compare and contrast the salient aspects of single, double, and multiple sampling inspection plans for attribute inspection.

(b) Derive an expression for the chance of acceptance of a lot of submitted product containing a proportion of defectives p, under a double sampling acceptance/rectification scheme.

(c) Comment on : "the long term process average outgoing percent defective, may be rarely expected to exceed one-half of the AOQL value associated with the inspection plan in use".

(3+9+8)=20

5. (a) Write a short note on the salient features of the MIL-STD-114 Sampling procedures for inspection by variables for percent defective.

(b) What are the conditions under which the mixed variables-attribute inspection plans are to be used? State the procedures of acceptance. (15+5)=20

6. (a) What is meant by a Rotatable Design? State the necessary and sufficient conditions for the rotatability of a second order design.

(b) If n_c points of the cube of type $(\pm 1, \pm 1, \ldots \pm 1)$ are used along with $n_s = 2k$ points of the star of type $(\pm\alpha, 0, 0 \ldots 0)$, $(0, \pm\alpha, 0, \ldots 0), \ldots (0, 0, \ldots \pm\alpha)$, to form a central composite design with k factors.

Prove that $\alpha = \sqrt[4]{n_c}$ for rotatability.

(c) Derive the conditions for orthogonal blocking for second order rotatable designs. Show that the condition for orthogonal blocking for the central composite design is given by,

$$\frac{\sqrt{n_c}}{2} = \frac{n_s + n_{co}}{n_s + n_{so}}$$

where n_{co} is the number of central points added with the cube in the first block and n_{so} is the number of central points added with the star in the second block

(3+3+4) = 20

7. (a) What is narrow limit gauging? Examine the advantages of and the problems involved in the use of narrow limit gauging for control of a measurable characteristic.

(b) Examine the following statement.

'The MIL-STD sampling plans and procedures are more appropriate when the supply is regular and continuous, whereas the Dodge-Romig plans are more suitable when the supply is occasional or sporadic'.

(10+10) = 20

8. Write brief notes on the following :

a) Role of standardization and variety reduction in Quality Control.

b) The importance of Vendor development for assurance of incoming quality.

(10+10) = 20

SD Paper VIII : STATISTICS
Group C : Special Paper I - Biometric Methods
(No candidate available)

Please turn over

(220)

ED VIII-Gr D: Design and Analysis of Experiments
Special Paper I - Statistical Aspects
(No candidate available)

ED VIII GROUP D : DESIGN AND ANALYSIS OF EXPERIMENTS

(Special Paper I - Theoretical Aspects)

Attempt any five questions from this group

(a) To draw a simple random sample (S.R.S) of size 'n' from a population of size 'N' without replacement (W.O.R) two methods are available : the direct method in which we first draw one unit from the whole population (giving equal probabilities to all the units) and then omitting the selected unit from the population, draw one more unit with equal probabilities from the remaining N-1 units, and then omitting again the two selected units, draw one more unit from the remaining N-2 units; the process of drawing has to be continued like this until a sample of size 'n' is obtained.

The indirect method is to go on drawing a S.R.S with replacement (W.R) until n distinct units are obtained and then to take these n distinct units (i.e. ignoring the repetitions) as the sample.

Prove that the two methods can be considered as equivalent, hence point out the advantages of the indirect method.

(b) In sampling for estimation of a rare attribute in a population, Haldane's method is to draw a simple random sample without replacement, until a preassigned number n of units possessing the rare attribute, are observed. If n is the size of the sample thus drawn,

(i) prove that (ignoring finite population correction)

$$\text{Prob}\,(n = k) = \binom{k-1}{n-1} P^n (1-p)^{k-n}$$

for $k \geq n$ (where p is the population proportion of the rare attribute)

(ii) Hence or otherwise, prove that

$$\hat{P} = \frac{n-1}{n-1}$$

is an unbiased estimate of P.

(c) Describe (i) the cumulative total method and (ii) Lahiri's method for drawing a 'pps' sample from a population and discuss their relative advantages and disadvantages.

(8+2+8+3) = 20

(a) In stratified sampling involving two strata, it is proposed to take a probability-proportional-to-$\bar{x}$ (pps) sample with replacement from the first stratum and a S.R.S. without replacement from the second stratum. If the total sample is to be of size 'n', derive the optimum allocation of sample size to the two strata, assuming that the cost per sampling unit is c_i for the i-th stratum (i=1,2), and making use of the usual estimators.

(b) On the basis of the sample drawn as per plan in (a) above, estimate the gain accrued by this method of sampling over that of an unstratified S.R.S. without replacement of size n, from the whole population.

[Assume cost to be proportional to n]

(12+8) = 20

Please turn over

3. (a) Describe the regression method of estimation and define the regression estimator $\hat{Y}_{reg}$ of the population total Y.

(b) Prove that $\hat{Y}_{reg}$ is not unbiased for Y, and derive an approximate expression for its bias.

(c) Derive an approximate expression for the mean-square error (m.s.e. of $\hat{Y}_{reg}$.

(d) Indicate how you will estimate the bias and m.s.e. of $\hat{Y}_{reg}$, on the basis of the sample.

(2+5+6+7) = 20

4. For a two stage sampling method the p.s.u's are selected with probabilities proportional to y (say) and the second stage units are selected with probabilities proportional to another auxiliary variable z (say). In both the cases the sampling is with replacement.

(i) Give an unbiased estimate $\hat{Y}$ of Y, the population total of the main variable y and prove its unbiasedness.

(ii) Derive $V(\hat{Y})$, the variance of $\hat{Y}$.

(iii) Give an unbiased estimate $\hat{V}(\hat{Y})$ of $V(\hat{Y})$ and prove its unbiasedness.

(5+7+8) = 20

5. For any sampling method let $\mathcal{S}$ be the collection of all possible samples and 's' a typical random sample. Let P_s denote the probability of drawing the sample 's', so that $\sum_{s \in \mathcal{S}} P_s = 1$.

For the i-th unit U_i of the population, the inclusion probability π_i is defined to be the probability that a random sample contains the i-th unit (irrespective of whether it occurs only once or more in the sample). Similarly, for the i-th and j-th units U_i and U_j of the population, the joint inclusion probability π_{ij} is defined to be the probability that a random sample contains both the i-th and the j-th units. In the above notation, then

$$\pi_i = \sum_{s \supset i} P_s \quad \text{and} \quad \pi_{ij} = \sum_{s \supset i,j} P_s$$

where the sums are over all samples that contain the i-th unit in the case of π_i, and the i-th and j-th units in the case of π_{ij}.

Assume that all the π_i's and π_{ij}'s are strictly positive.

With these preliminaries,

(a) prove that the well-known Horvitz-Thompson estimator given by

$$\hat{Y}_{HT} = \sum_{i \in s} \frac{Y_i}{\pi_i}$$

, is unbiased for the population to

(where the sum is over all distinct units U_i that belong to the sample i.e. repetitions are ignored)

(b) Derive the variance of $\hat{Y}_{HT}$ in the form $\sum_{i=1}^{N} a_{ii} Y_i^2 + \sum\sum_{i \neq j} a_{ij} Y_i Y_j$.

(30)

5(c) Contd. Prove that

$$\sum_{i\in s}\frac{Y_i^2}{\pi_i^2}(1-\pi_i)+\sum\sum_{\substack{i\neq j\\ \in s}}\frac{Y_iY_j}{\pi_i\pi_j}\,\frac{\pi_{ij}-\pi_i\pi_j}{\pi_{ij}}$$

is an unbiased estimator of the variance derived in (b) above.

[As in (a), the sums are over distinct U_i's belonging to the sample s and distinct pairs (U_i, U_j) belonging to s, respectively.]

(5+7+8) = 20

Bringing out the relative merits and demerits, and using suitable mathematical and heuristic arguments, make critical comparisons between the following:

(i) multistage sampling versus unistage sampling;

(ii) pps sampling versus Ratio and Regression methods;

(iii) simple random sampling versus linear and circular systematic sampling;

(iv) sampling with replacement versus sampling without replacement.

(5 X 4)=20

Write short notes on any five of the following:

(i) Sampling from incomplete and outdated frames.

(ii) Nonresponse.

(iii) Nonsampling errors.

(iv) Two-phase sampling.

(v) Multipurpose surveys.

(vi) Interpenetrating subsamples.

(5 X 3) = 20

GROUP F : THEORY OF COMPUTING

(Special Paper I-Numerical Analysis

(No candidate available)

GROUP G : STATISTICAL METHODS

(Special Paper I - General Theory)

Attempt any five questions from this group.

(a) Show that if a population has a finite mean, it is unbiasedly and consistently estimated by the sample-mean.

(b) Give one example of each of :-

(i) An estimator which is biased, but consistent.

(ii) An estimator which is unbiased but not consistent.

(c) For the distribution with p.d.f.

$$f(x;\theta)=\frac{1}{\theta}e^{-\frac{x}{\theta}}\qquad (0\le x<\infty;\ \theta>0),$$

examine whether an unbiased estimator of the population standard deviation exists, and if so find the same. (6+3+3+8)=20

Please turn over

2(a). Define (i) a complete family of probability distributions,
(ii) a complete statistic.

(b) Show that any function of a complete sufficient statistic is the minimum variance unbiased estimator of its expected value.

(c) Obtain the minimum variance unbiased estimator of the standard deviation of a normal population with known mean μ.

(2+7+3) = 20

3(a). Explain the technique of the method of minimum chi-square and comment on its relationship to the method of maximum likelihood.

(b) State and prove the relationship between a maximum likelihood estimate and a sufficient estimator.

(c) Show that if an unbiased estimator of a parameter θ attains the Cramér-Rao lower bound, it is also the maximum likelihood estimator of θ.

Does this property hold good for the minimum variance unbiased estimator of a parameter?

(3+4+3) = 20

4(a). Show that if a uniformly most powerful test (U.M.P.) exists, it is unbiased.

(b) Define a locally most powerful test and show how to obtain such a test.

(c) Obtain a locally most powerful test for testing $H_0 : \theta = \theta_0$ for the parameter θ with p.d.f. distribution:

$$f(x;\theta) = \theta e^{-\theta x}, \quad \text{where} \quad (0 < x < \infty ; \theta > 0)$$

(7+5+3) = 20

5(a) Apply the likelihood ratio criterion to obtain a test for equality of variances in two normal populations on the basis of independent samples.

(b) Obtain a confidence interval with confidence coefficient α for the range of a rectangular population on the basis of a random sample.

(10+10) = 20

6(a) Explain the terms (i) risk function, (ii) convexity, (iii) randomized decision function, and (iv) Bayes decision function.

(b) Explain how a statistical decision problem can be viewed as a zero-sum two person game.

(c) Show how the Neyman-Pearson theory of testing of hypotheses can be viewed as a special case of the decision theory.

(6+6+8) = 20

SD VIII : GROUP B – PROBABILITY THEORY

(Special Paper I – Basic Probability)

SECTION I

Attempt all questions from this section.

1. Either

(a) Define outer measure and measurability with respect to an outer measure.

(b) If the sets A and B are measurable with respect to a given outer measure, show that A ∪ B is also measurable.

(c) Show that every set of outer measure zero is measurable. (6+6+8)=20

Give the statement of the dominated convergence theorem for integrals.

Why does the dominated convergence theorem fail to hold in the following examples :

(i) $\Omega = [0,1]$, $\mathcal{B}$ the Borel σ-field of $[0,1]$, P the Lebesgue measure on $\mathcal{B}$, and

$$f_n(\omega) = \begin{cases} n & \text{if } \omega \in (0, \frac{1}{n}] \\ 0, & \text{otherwise} \end{cases}$$

(ii) Ω is the real line, $\mathcal{B}$ the Borel σ-field on it, P the Lebesgue measure on Ω, and

$$f_n(\omega) = \begin{cases} \frac{1}{n} & \text{if } \omega \in (-n,n) \\ 0, & \text{otherwise.} \end{cases}$$

(6+7+7)=20

$(X,\mathcal{A})$ and $(Y,\mathcal{B})$ are measurable spaces and $(X \times Y, \mathcal{A} \times \mathcal{B})$ denotes the product space.

If $h(x,y)$ is a function on $X \times Y$, define the x-section and y-section of $h(x,y)$.

Show that if $h(x,y)$ is a measurable function, then every x-section of $h(x,y)$ is measurable.

If $E \in \mathcal{A} \times \mathcal{B}$, $\alpha > 0$ and β are real numbers, show that the set

$$\{ (x,y) \mid (x, \alpha y + \beta) \in E \}$$

is a measurable subset of $X \times Y$.

(4+3+6) = 13

Define signed measure.

If μ is a signed measure on $(X,\mathcal{A})$ show that there exist two disjoint sets A,B such that

(i) $A \cup B = X$

(ii) $A \cap C$, $B \cap C \in \mathcal{A}$, whenever $C \in \mathcal{A}$

(iii) $\mu(A \cap C) \geq 0$, $\mu(B \cap C) \leq 0$ for all $C \in \mathcal{A}$.

(5+10)=15

SECTION VIII Group C : Probability Theory (contd.) :

SECTION III (Attempt all questions from this section)

If f is a finite valued random variable defined on a probability space $(\Omega, \mathcal{B}, P)$ and F is its distribution function. Show that F is a monotonic non-decreasing function with $\lim_{x \to \infty} F(x) = 1$.

$f, f_1, f_2, \ldots$ are finite valued random variables defined on the probability space $(\Omega, \mathcal{B}, P)$ with distribution functions $F, F_1, F_2, \ldots$ respectively. Show that if $f_n \to f$ in probability, then $F_n \to F$ in the sense of convergence of distribution functions.

(5+15) = 20

(2 33)

Please turn over

4. (Contd.)

Let F be a continuous, monotonic non-decreasing function on the real line. The set function μ is defined for half-open intervals $(a,b]$ by

$$\mu\{(a,b]\} = F(b) - F(a)$$

Show that the set function μ can be extended to a measure on the Borel σ-field of the real line.

(20)

5(a). If P is a probability measure on the measurable space $(\Omega,\mathcal{A})$ show that

$$\lim_n P(E_n) = P(E)$$

where $E_n \in \mathcal{A}$ and $E = \lim_n E_n$.

(b) $\{E_n\}$ is a sequence of events such that $\sum_{n=1}^{\infty} P(E_n)$ is finite.

Find the probability that an infinite number of the events E_n happen.

(6+6) =12

6(a). Define independent random variables.

(b) If f,g are independent random variables, show that the random variables

$$af + b \; ; \; cg + d$$

are also independent. $(a,b,c,d$ are real numbers with $a, c \neq 0)$.

(c) If ϕ, ψ are continuous functions on $(-\infty, \infty)$, and f,g are independent random variables, show that $\phi(f)$ and $\psi(g)$ are independent random variables.

(4+4+7)=15

INDIAN STATISTICAL INSTITUTE

Statistician's Diploma Examination - May 1971

Paper IX - Subjects of second paper of Specialisation (Theoretical)

Time : 4 hours Full marks : 100

(a) Candidates are required to answer questions from that group only for which they have registered their options.

(b) Figures in the margin indicate full marks.

GROUP A : ECONOMIC STATISTICS - Special Paper II

(Indian Economics & Economics of Planning)

Section I - Indian Economics

(Attempt any three questions from this Group)

1. Discuss the main characteristics of the occupational distribution of the working population of India. Have there been any important changes in this distribution since independence? (16)

2. Critically discuss the major problems of defining and measuring 'disguised' unemployment in Indian agriculture. (16)

3. How would you try to assess the degree of incidence of poverty in India? What are the reasons for believing that there are now more people below the poverty line than there were, say, in the beginning of the 1950's? (16)

4. Suggest suitable measures for the solution of long term 'balance of payments' difficulties in India and give reasons supporting the suggested measures. (16)

5. Do you think that zonal restriction on the movement of foodgrains and compulsory procurement of foodgrains have deepened our food problems? Give reasons for your answer. (16)

6. Discuss whether the shortage of industrial finance has hampered industrial growth in India and critically review the working of the new institutions for providing industrial finance in the light of your answer. (16)

Section II - Economics of Planning

(Attempt any three questions from this Group)

7. Discuss the usefulness and limitations of the concept of incremental capital-output ratio as a planning tool. (16)

8. What are the principles that should govern the choice of production techniques in a poor country with a large volume of surplus labour? (16)

9. Discuss the dual role of "foreign aid" (including loans) in bridging the savings-investment and the export-import gaps in an underdeveloped economy. Is "more aid to end all aid" a good slogan?

Give reasons for your answer. (16)

10. Discuss the uses of the von Neumann model of growth for the purposes of planning. (16)

11. Discuss the manner in which a slow rate of growth of agriculture could affect the rate of growth of industry. Illustrate your answer with examples drawn from Indian experience. (16)

12. Write short notes on any two of the following:

(1) The Golden Rule of accumulation.

(2) The open static Leontief model

(3) The shadow price of labour in an underdeveloped economy.

(4) The optimum rate of saving.

(8+8) = 16

Neatness (Sections I and II) (4)

SD-II - GROUP B: TECHNO COMMERCIAL STATISTICS - Special Paper II

Section I - OPERATIONS RESEARCH (70 marks)

Section II - STATISTICAL METHODS IN BUSINESS (30 marks)

Section I (Alternative): ELEMENTS OF BOOK-KEEPING AND ACCOUNTANCY (70 marks)

Section I: OPERATIONS RESEARCH

(i) Use a separate answer book for this section.

(ii) Attempt question 1 and any other two questions from this section.

1(a) Explain the meaning of following terms used in linear programming: Slack variables, artificial variables, feasible solution and basic feasible solution.

(b) If the constraints of a linear programming problem are consistent, show there exists a basic feasible solution.

(c) Establish the correspondence between basic feasible solutions and extreme points of the convex set of feasible solutions.

(d) Consider the linear programming problem Max.z= cx subject to $Ax = b, x \geq 0$. If $x_B = B^{-1}b$, is a basic feasible solution such that $C_B B^{-1}P_j \leq C_j$ for all non-basic variables, prove that it is an optimal feasible solution.

(Notation: B is the basis matrix, CB is the m component row vector of cost coefficients of the basic variables and P_j is the jth column of A).

(4+6+6+10) = 30

2(a) Explain the solution procedure for solving the assignment problem with the help of the following example:

There are 5 jobs to be assigned one each to 5 machines and the associated cost matrix is as follows: (see next page)

Machine Cost Matrix

M/C	1	2	3	5	5
A	11	17	8	16	20
B	9	7	12	6	15
C	13	13	15	12	16
D	21	24	17	28	26
E	14	10	12	11	15

(b) Give a mathematical formulation of the transportation problem for material, available at m sources and required at n destinations, assuming the total available material is just sufficient to satisfy the total requirements. Discuss the theoretical basis for any one of the algorithms for solving a transportation problem.

(10+10) = 20

3(a) Derive expressions for economic lot size and minimum cost for the following inventory situation:

(i) Demand is uniform at a rate of r per unit time.

(ii) Replenishment is not instantaneous but takes place at a uniform rate p per unit time (p > r)

(iii) Cost of carrying inventory is C_1 rupees per unit per unit time.

(iv) Cost of production set up is C_2 rupees per set up.

(v) No shortage is permitted.

(b) The uniform annual demands for two bulky items are 90 tons and 160 tons respectively. The carrying costs are Rs. 250/- and Rs.200/- per ton per year and the set up costs are Rs. 50/- and Rs.40/- per production run respectively. No shortages are allowed. Space considerations restrict the average amount in inventory of both items to 4000 ft^3. A ton of the first item occupies 1000 ft^3 and a ton of the second item 500 ft^3. Find the optimal lot size of each item.

Show that the space restriction increases costs by about Rs.75/- per year.

(10+10) = 20

4(a) For a single server queue with Poisson arrivals, with mean arrival rate λ and exponential service time with mean service rate μ ($\lambda < \mu$), obtain the following results:

(i) P_n = Probability that there are n units in the system

$$= 1 - \frac{\lambda}{\mu} \left(\frac{\lambda}{\mu}\right)^n \qquad n \geq 0$$

(ii) E(n) = Expected number of units in the system

$$= \frac{\lambda}{\mu - \lambda}$$

(iii) E(w/w > 0) = Average waiting time of an arrival who waits

$$= \frac{1}{\mu - \lambda}$$

(b) Arrivals at a telephone booth are considered to be following Poisson law of distribution with an average time of 10 minutes between one arrival and the next. The length of a phone call is assumed to be distributed exponentially with mean 3 minutes.

(i) What is the probability that a person arriving at the booth will have to wait?

(ii) What is the average length of the queue?

(iii) The telephone department will install a second booth when convinced that an arrival would expect to have to wait at least 3 minutes for the phone. By how much, must the flow of arrivals be increased, in order to justify a second booth?

(4+3+3 + 3+3+4) = 20

5(a) (i) What features characterise Dynamic Programming Problems?

(ii) State Bellman's Principle of optimality.

(b) A computer contains 10,000 resistors. When each resistor fails, it is replaced. The cost of replacing a resistor individually is Re.1/- only. If all the 10,000 resistors are replaced at the same time, the cost per resistor would be reduced to 35 paise. The percent surviving S(t) at the end of month t and P(t) the probability of failure during the month t are:

t	S(t)	P(t)
0	100	-
1	97	.03
2	90	.07
3	70	.20
4	30	.40
5	15	.15
6	0	.15

What is the optimum replacement plan?

(4+2+14) = 20

SUBJECT - GROUP B : SECTION II - STATISTICAL METHODS IN BUSINESS

(i) Use separate answer book for this section.

(ii) Attempt any two questions from this section.

1(a) Describe briefly the various methods of sales forecasting.

(b) How are seasonal variations estimated? (10+5) = 15

2(a) The life of electric light bulbs is known to be a normally distributed random variable, with unknown mean μ and standard deviation 200 hours. The value of a lot of 1000 bulbs is (1000)(1/5000) μ rupees. A random sample of n bulbs is to be drawn by a prospective buyer and 1000 (1/5000)$\bar{x}$ rupees paid to the manufacturer. How large n should be so that the probability is 0.95 that the buyer does not overpay or underpay the manufacturer by more than Rs.20/-?

(b) A large company is faced with the problem of verification of its huge inventory. It has 54000 items in stock valued at Rs.300 lakhs. With two additional lines of manufacture, it was expected that the number of items to be carried in stock would increase by 10-15% and the corresponding increase in value of some 30%. Previous experience with verification of physical inventory had shown that on an average 28% of the total number of items checked had differences between book records and actual stock. The company wants a method of verification and estimation of inventory without an increase in the existing manpower while still satisfying the requirements of accuracy of inventory counts.

Can you discuss a suitable statistical technique for the inventory verification work? Would you consider additional information necessary? If so, give broad indications. (7+8) = 15

3(a) What is payment by results? Give a brief account of the various systems of payment by results.

(b) A sample poll was taken in a constituency to predict the results of election. Interviews with 1000 randomly selected registered voters showed the following preference:

For candidate A : 532
For candidate B : 468

On the basis of the above, the prediction of A's election was published. Election day came and B won with 51.1% of the votes Comment on the situation.

(10+5) = 15

SDII - GROUP B : SECTION I (Alternative) - ELEMENTS OF BOOK KEEPING AND ACCOUNTANCY

(i) Use separate answer book for this section.

(ii) Attempt question 5 and any other three from this group.

1. How many different classes of accounts are usually to be found in a Ledger? State the characteristic features of each such class of account and give three examples for each class of account.

(5+5+5) = 15

2(a) Describe briefly what is meant by each Book.

(b) Enter the following transactions in a three column Cash Book of Messrs. General Trading Co.

April 1971 - 1 Cash in hand Rs. 237; balance at Bank Rs.6,594

4 Received from K. Sen Rs. 590, allowed him discount Rs.10/-.

6 Paid salaries for March 1971 by cash, Rs.200/-. Cash sales Rs.134/-.

7 Paid B.K. Bose by cheque Rs.300/-, cash purchase Rs.60/-.

10 Paid Ahmed by cheque Rs.594; discount received Rs.26.

14 Cash sales Rs.112; paid cartage and coolie fare Rs.6/-.

18 Withdrew from Bank for office use Rs.200; paid rent in cash Rs.50/-.

19 Cash sales Rs.212; received from G.C. Dhar Rs.194/- by cheque; discount Rs.6/-.

22 Deposited into the Bank Rs.1000/-. Purchased a Motor Car for Rs.5,300 and drew a cheque for the amount.

28 Cash sales Rs.299/-.

30 Bank notifies that Robert and Co's cheque has been dishonoured

(3+12) = 15

3(a) Distinguish between Capital and Revenue Expenditure.

(b) 'Chhayabani' Ltd., constructed a cinema house and had incurred the following expenditure during the first year ending 31st March 1971.

i) Second-hand furniture worth Rs.9000 was purchased; repairing of the furniture cost Rs.1000. The furniture was installed by own workmen - wages for this being Rs.200.

ii) Expenses in connection with obtaining a licence for running the cinema were Rs.20,000. During the course of the year the cinema was fined Rs.100 for contravening rules. Renewal fee, Rs.200, for next year has also been paid.

iii) Fire insurance, Rs.1000, was paid on 1st October 1970 for one year.

iv) Temporary huts were constructed costing Rs.1,200. These were necessary for the construction of the cinema house and came to be demolished when the house was complete.

Point out how you would classify the above items.

(5+10) = 15

4. A draws on B a bill of exchange for 3 months for Rs.1000 which B accepts on 1st January 1970. A endorses the bill in favour of C. Before maturity B approaches A with the request that the bill be renewed for a further period of 3 months at six per cent per annum interest. A pays the sum to C on the due date and agrees to the proposal of B.

Pass journal entries in the books of A, assuming that the second bill is duly met.

15

5. S. Sen Gupta

Trial Balance as at 31st December 1970.

			Dr Rs.	Cr Rs.
1.	S. Sen Gupta: Capital	..		4,250
2.	S. Sen Gupta: Drawings	..	710	
3.	Plant and Machinery	..	950	
4.	Stock, 1st January 1970	..	1,460	
5.	Purchases	..	10,362	
6.	Purchases Returns	..		291
7.	Sales	..		11,906
8.	Sales Returns	..	210	
9.	General Expenses	..	440	
10.	Rent	..	120	
11.	Rates	..	200	
12.	Aprenticeship Premium	..		80
13.	Bank overdraft	..		240
14.	Bad Debts	..	172	
15.	Sundry Debtors	..	3,200	
16.	Sundry Creditors	..		1,000
17.	Cash in hand	..	48	
18.	Bad Debts Reserve	..		105
			17,872	17,872

You are required to prepare Trading and Profit and Loss Account, for the year ended 31st December 1970, and Balance Sheet as at that date, having regard to the following:

i) Write off depreciation on plant and machinery at 10 per cent per annum.

ii) The Bad Debts Reserve is to be increased to 5 per cent on Sundry Debtors.

iii) Sen Gupta has drawn out goods for his own consumption amounting to Rs.40, of which there is no record in the books.

iv) Rs.100 Plant and Machinery purchased on 31st December 1970 has been inadvertently included in Purchases.

v) Rent accrued is Rs.40/-

vi) Rates of Rs.80 are paid in advance.

vii) Stock-in-hand at 31st December 1970 is valued at Rs.1,730.

(7+10+8) = 25

SD-IX - GROUP C : BIOLOGY (Special Paper II)

Section I : Statistical Methods in Genetics

(Attempt any two questions from this Group)

1. Explain in brief the terms 'linkage' and 'recombination fraction'. For what value of the recombination fraction does one conclude that there is no linkage?

Can you observe linkage from data on families where none of the parents is doubly heterozygous? Explain.

Estimate the recombination fraction and obtain its standard error, when the data are collected from offspring of a back-cross mating between a double heterozygote in the coupling phase and a double recessive one

(i.e., $\frac{ab}{AB} \times \frac{ab}{ab}$ mating).

(3+3+2+2 + 15) = 25

2. What is a population said to be in genotypic equilibrium with respect to an autosomal locus, with two alleles?

Show that whatever be the initial genotypic proportions, under random mating, the population reaches equilibrium in a single generation.

In a random mating population, how do you estimate the recessive gene frequencies from data on mother-child combination?

(5+10+10) = 25

3. Describe briefly the genetic theory explaining the genotype-phenotype relationship in the ABO blood group system.

In a family where the blood groups of the parents are A and AB respectively, what are the possible blood groups for a child of these parents?

Show that for a random mating population the maximum likelihood estimates of A, B and O gene frequencies (denoted by p, q and r respectively), are given by the solution of

$$\frac{\overline{A} + \overline{AB}}{P} = 2\,(\overline{O}) + \frac{\overline{A}}{2 + P} + \frac{2\,(\overline{B})}{2 + Q}$$

$$\frac{\overline{B} + \overline{AB}}{Q} = 2\,(\overline{O}) + \frac{2(\overline{A})}{2 + P} + \frac{\overline{B}}{2 + Q},$$

Where, $\overline{O}$, $\overline{A}$, $\overline{B}$ and $\overline{AB}$ denote the observed phenotypic proportions and $P = p/r$ and $Q = q/r$.

(8+2+15) = 25

SD-IX - Group B - Section II : Bio-Assays

(Attempt any two questions from this group)

4. Explain the terms direct assay and indirect assay. What are the major disadvantages of direct assays? Clearly indicate how by suitably designing an indirect assay, one can remove these difficulties.

(3+3+9+10) = 25

5. Indicate a method to estimate the median effective dose and obtain its standard error on the basis of quantal data, when the tolerance distribution is assumed to be normal.

Explain in brief the terminologies underlined and used in this question.

(16+3+3+3) = 25

6. Write short notes on any two of the following methods of analysing data on quantal responses:

(i) The method of characteristic curves.

(ii) Estimation of equivalent doses.

(iii) Estimation by least squares.

(iv) The method of extreme effective doses

(12½+12½) = 25

SD-IX: GROUP D - DESIGN & ANALYSIS OF EXPERIMENTS - SPECIAL PAPER II (COMBINATIONAL ASPECTS)

No candidate available

SD-IX: GROUP E - SAMPLE SURVEYS - (Special Paper II)
(Organisational aspects)

(Answer any five questions from this group)

1.(a) Discuss the importance of concepts and definitions in conducting a sample survey. Illustrate by means of two practical examples, how lack of precise definitions and concepts can completely abort the efforts undergone in conducting a survey.

(b) Indicate the situations in which the following methods of inquiry may be considered most suitable:

(i) physical observation; (ii) interview; (iii) mailed questionnaire.

For each method, illustrate your answer with the help of two practical examples.

(4+1 + 4+4+4) = 20

2.(a) What are the nature of biases, which you would be on the lookout for and guard against, while interviewing a candidate for the post of an investigator.

(b) Discuss the relative merits and demerits of the following methods of recruitment of field investigators:

(i) Interview and consideration of academic records.

(ii) No interview but only a written test with a medical certificate of fitness for touring duties.

(c) Prepare a note bringing out relative advantages and disadvantages of recruiting Inspectors through:

(i) Direct recruitment from the open market.

(ii) Promotion from amongst field investigators.

(6 + 4+4 + 3+3) = 20

3.(a) The following are the estimates of expenditure for completing various stages of work of a survey by employing two alternative types of workers, (i) wholetime workers and (ii) part-time workers:

Stages of work	Expenditure on the basis of wholetime workers (Rupees)	part-time workers (Rupees)
(i) Preliminary work and collection of primary data	15,000	30,000
(ii) Processing and analysis of data	5,000	20,000
(iii) Supervision and report writing	5,000	15,000

Indicate clearly, how would you organise the work to obtain the best results, if the total cost of the survey is not to exceed Rs.50,000. Give reasons for your choice?

(b) Make out what you would consider a strong case for a higher pay scale to be given to the investigators having minimum academic qualifications equivalent to those in other cadres, but enjoying a lower pay scale. (12+8) = 20

4. Mention the steps you would take to reduce errors in final estimates of a sample survey, arising out of the following causes:

(i) deliberate attempts on the part of informants to give biased information.

(ii) non-response of sample households.

(iii) mistakes due to possible malfunctioning of ITM equipment.

(iv) careless and inefficient scrutiny on the part of the scrutinisers

(v) defective verification of punched cards.

(vi) careless and inefficient checking of the computation sheets and tables.

(vii) tendency of the investigators to fill in schedules without actually contacting the sample households.

(viii) inconsistency of figures supplied by respondents.

$(2\frac{1}{2} \times 8) = 20$

5.(a) Give the layout of the main blocks of a schedule to be designed for undertaking a sample check of the Population Census completed in April 1971. Also draft for the use of the investigators a set of instructions for filling up these blocks.

(b) Discuss the relative advantages and disadvantages of writing a report on a sample survey by adopting the following alternative methods:

(i) Dividing the report into two parts, viz.,

Part I : Description of the survey and discussion of the results.

Part II: Statistical Tables.

(ii) Arranging the report in such a manner that the description and analytical results are interspersed with short and handy tables.

$(7+7 + 3+3) = 20$

6. A mass of tabulated data collected in connection with investment and indebtedness in the rural areas of the country is handed over to you for preparing a draft report. Give a detailed synopsis of the report, assuming that all the information needed by you for the report is available in the tables (20)

7. Write short notes on any four of the following:

(i) Mechanical versus Manual Tabulation

(ii) Centralised versus decentralised training of field investigators.

(iii) Weeding of filled in schedules and punched cards.

(iv) Testing of schedules and instructions.

(v) Importance and type of propaganda needed for successful completion of the survey.

(vi) Scrutiny according to a given programme as against intelligent scrutiny without any given programme.

(vii) Single versus repetitive surveys.

$(5 \times 4) = 20$

SC:X GROUP F : TECHNIQUES OF COMPUTATION

Special Paper II - Practical

(Calculation with Desk Calculators)

(Attempt any five questions from this group)

Time : 3 hours Full marks : 100

1.(a) The values and the absolute errors of x_1, x_2 and x_3 are

$$x_1 = 22.3925 \pm 0.00025 ;$$

$$x_2 = 5.3123 \times 10^5 \pm 5.0 \text{ and}$$

$$x_3 = 0.000103321 \pm 0.5 \times 10^{-9}$$

Find the relative errors in each of these. Let y be defined as

$$y = x_1 \cdot \left(\frac{x_2}{x_3}\right)^{1/2} .$$

Compute the value of y and its error.

(b) How many terms are to be used to find the value of $\log_e 2$ correct upto 5 places of decimal using the following expression :-

$$\frac{1}{2}\log_e x = \left(\frac{x-1}{x+1}\right) + \frac{1}{3}\left(\frac{x-1}{x+1}\right)^3 + \frac{1}{5}\left(\frac{x-1}{x+1}\right)^5 + \ldots\ldots$$

(5+5+2+10)=20

Tabulate the function $f(h,k)$ for $h = -1.0(0.5)1.0$, $k = -1.0(0.5)1.0$ where $f(h,k)$ is defined by

$$f(h,k) = \begin{cases} \dfrac{|h|+|k|}{2} & \text{if } h \text{ and } k \text{ are either both positive or both negative (i.e. } h \geq 0 \text{ and } k \geq 0 \text{ or } h \leq 0 \text{ and } k \leq 0) \\ \dfrac{h^2+k^2}{2(|h|+|k|)} , & \text{otherwise} \end{cases}$$

where $|\ldots|$ denotes the absolute value.

From this table find the value of $f(0.1, -0.1)$ using any interpolation formula, and find its error.

(10+7+3) = 20

3. Find the value of the integral I, by using any mechanical quadrature formula, where I is given by

$$I = \int_0^1 \frac{1}{1+x^2}\, dx$$

Compare your result with the actual value. (15+5) = 20

4.(a). Let $T_n(x)$ denote Chebychev's polynomial of degree n. Write explicitly $T_n(x)$ for $n = 1(1)10$.

(b) The function $f(x) = x^{1/3}$, is to be approximated by a polynomial $p(x)$ of degree 3 in such a way that ϵ defined by

$$\epsilon = \max_{-1 \leq x \leq 1} |f(x)-p(x)| \text{ is minimum.}$$

Find $p(x)$.

[Hints: Write $f(x) = \sum_{r=0}^{3} a_r T_r(x)$ and obtain the coefficients a_0, a_1, a_2 and a_3 using the orthogonality property of Chebychev polynomials.]

(10+10) = 20

5. Solve numerically the differential equation $\frac{dy}{dx} = y+1$, with the initial condition $y(0) = 0$. (20)

6. Determine to five places of decimal, a real root of the following equations :

$$x = \sin(x+y)$$
$$y = \cos(x-y) \qquad (20)$$

7. Solve the following set of equations :

$$0.437\, x_1 + 5.137\, x_2 + 3.141\, x_3 + 2.063\, x_4 = 29.912$$
$$5.137\, x_1 + 6.421\, x_2 + 2.617\, x_3 + 2.003\, x_4 = 25.053$$
$$3.141\, x_1 + 2.617\, x_2 + 4.120\, x_3 + 1.623\, x_4 = 10.557$$
$$2.063\, x_1 + 2.003\, x_2 + 1.623\, x_3 + 3.416\, x_4 = 12.690 \qquad (20)$$

SD Paper I - Group V : Statistical Inference
Special Paper II - Special Topics
(No candidate available)

SD Paper I - Group VI : Probability Theory
Special Paper II - Limit Distributions
(No candidate available)

INDIAN STATISTICAL INSTITUTE

Statistician's Diploma Examination - May 1971

Paper X - Subjects of Specialization - III (Practical)

Time : 5 hours Full marks : 100

SD-X : GROUP A: ECONOMIC STATISTICS -
Special Paper III - Practical
(No candidate available)

SD-X : GROUP B - TECHNO-COMMERCIAL STATISTICS.

Section I - Statistical Quality Control : 50 marks

(Attempt question 1 and any other two from this Section)

1. The following table provides the summarised data on a measurable characteristic with specification limits 0.7528 to 0.7571, the sample size being 4.

Sample	$\bar{X}$	R	Sample	$\bar{X}$	R
1	0.7540	.0011	9	0.7517	.0007
2	0.7542	.0014	10	0.7549	.0015
3	0.7542	.0009	11	0.7542	.0010
4	0.7546	.0010	12	0.7541	.0017
5	0.7550	.0008	13	0.7545	.0011
6	0.7539	.0009	14	0.7548	.0009
7	0.7541	.0012	15	0.7551	.0012
8	0.7543	.0011	-		

(a) Examine the data carefully and give a control scheme which in your opinion may be considered as most advantageous.

(b) Suppose it is decided to use samples of size 5 for control purposes; obtain the control limits for average and range charts. (9+5) = 14

2.(a) An inspector with 80 per cent efficiency (i.e. for him the probability of classifying a defective as a non-defective is 0.20) uses the sampling plan n = 100, c = 1. Plot the effective OC curve.

(b) Construct a single sampling plan for attributes, given the following data:

AQL = 0.05; Producer's risk = 0.05

LTPD = 0.20; Consumer's risk = 0.10

(12+6) = 18

3.(a) For a lot of size N = 2000, the following single sampling plan is being used for major defects, n = 120, c = 1. It is desired to use for minor defects, a single sampling plan having the same sample size and an AOQL as close to 5% defective as possible, but not greater than 5%.

What single sampling plan should be used and what is the actual AOQL of the plan chosen?

(b) The tolerance specified for the outer diameter of a shell is $14'' \pm 0.067''$. The oversized shells are reworked, while the undersized ones are scrapped. The cost of reworking 5 shells is equal to the loss incurred by scrapping one shell. The process standard deviation was found to be 0.028 inches.

Obtain the average level by minimising the total expected cost of rework and scrap.

(over)

(c) Find the equations of acceptance and rejection lines and plot the five point OC curve for the sequential sampling plan with,

AQL = 0.05, Producer's risk = 0.05

LTPD = 0.10, Consumer's risk = 0.15

(6+6+6) = 18

4. In a chemical experiment 12 combinations of 2 factors were tried. The table below gives the response y and the corresponding levels of two factors, x_1 and x_2 (suitably scaled).

factor (x_1)	factor (x_2)	response (y)	factor (x_1)	factor (x_2)	response (y)
-1	-1	15.67	0	$-\sqrt{2}$	25.39
1	-1	10.17	0	$\sqrt{2}$	18.18
-1	1	16.50	0	0	22.22
1	1	5.80	0	0	19.49
$-\sqrt{2}$	0	17.65	0	0	22.76
$\sqrt{2}$	0	0.20	0	0	24.27

i) Fit a second degree response function of y on x_1 and x_2.

ii) Test for the lack of the fitted response.

(12+6) = 18

SD-IV : GROUP B: SECTION II : OPERATIONS RESEARCH - 50 marks

(Attempt question 5 and any other two from this section)

5. Find an optimal product mix for a factory producing paper towels of three different sizes, using the data given below

Department	Time (in minutes) required per unit of size A	size B	size C	Capacity (minutes)
Cutting ..	10.7	5.0	2.0	2705
Folding ..	5.4	10.0	4.0	2210
Packaging ..	0.7	1.0	2.0	415
Profit/unit ...	Rs.10	Rs.15	Rs.20	

(20)

6. A company proposes to install a large boiler. A certain automatic monitoring unit is crucial for the operation of the whole system. At the time of the original order, spares for this unit can be purchased for Rs.2000 each. The probability distribution of demand of this spare (unit) during the life time of the installation is known to be,

Demand	Probability
0	0.35
1	0.25
2	0.20
3	0.15
4	0.05

If a spare (unit) is needed and not available, the cost to the company in down time and replacement, will be Rs.15,000. Unused spares have no salvage value.

What is the optimal number of spares to order?

(10)

7. An oil company is constructing a service station on a high way. Traffic analysis indicates that customer arrivals would approximate a Poisson process, with a mean of 30 automobiles per hour. Previous studies show that one pump could service a mean of 16 automobiles per hour, with the service-time distribution approximating the negative exponential.

i) If 4 pumps are installed, what is the probability that an arrival would have to wait in a line?

ii) Determine the average number of automobiles in the system, average waiting time and average time spent by a customer in the system.

iii) For what percentage of time would a pump be idle on an average?

(Consider only the steady-state situation that would be attained in the long run)

(3+4+3) = 10

SD-X : GROUP B: Section III - STATISTICAL METHODS IN BUSINESS

(Attempt any two questions from this Section) 20 marks

8. Analyse the following matrix of transitional probabilities and determine the long run equilibrium market share for each firm. Give your reasons.

From	To X	Y	Z
Firm X	1.0	0	0
Firm Y	0.10	0.75	0.15
Firm Z	0.10	0.05	0.85

Given that the initial market share is (0, 0.5, 0.5) for X, Y and Z, obtain the market share for Y, after two transitions.

(10)

9. Business cycles in U.S.A. and England have had the following duration as measured to nearest year. The observations below are given in chronological order during the period 1796 to 1925.

U.S.A.:- 6, 6, 5, 3, 7, 3, 3, 5, 4, 3, 6, 12, 6, 4,
3, 5, 5, 4, 9, 5, 3, 2, 3, 4, 3, 4, 2, 3, 5, 2, 3

England:- 4, 6, 4, 3, 5, 4, 6, 4, 2, 6, 10, 7, 4, 8, 8, 9,
8, 10, 7, 6, 5, 2

i) Calculate the average duration of business cycle in each country separately.

ii) Test whether the duration of business cycle is on an average more in England than in U.S.A.

(10)

10. The following are the profits of a concern during the years 1955 to 1961:

Years:	'55	'56	'57	'58	'59	'60	'61
Profit (thousand rupees)	75	70	72	65	50	54	41

Fit an appropriate trend line to the data. Hence, predict the profit of the concern in the year 1963 and obtain the standard error of your estimate.

Comment on the data and your results. - (10)

NOTE: Scheme of marking followed for Group B (Special Paper III - Practical) : Sec.I:S.Q.C. - 50 marks; Sec.II: O.R./(Alt.) E.B.A. 30 marks; Sec.III: S.M.D. - 20 marks.

SD-X - GROUP B - Section II (Alternative):
Elements of Book Keeping & Accountancy (Practical)
(No candidate available) Spl. Paper III

SD-X : GROUP C : BIOMETRIC METHODS (Special Paper III - Practical)

(Attempt any three questions from this group)

1. Four measurements Stature (ST), Head Length (HL), Head Breadth (HB) and Nose Breadth (NB) are made on each member of two very large random samples, taken from the African tribes Batutsi and Swahili. The mean values for these measurements for the two tribes are given separately below:

	ST	HL	HB	NB
Batutsi tribe	1723.03	195.21	142.88	39.58
Swahili tribe	1619.10	191.80	137.22	44.08

The inverse of the pooled variance - covariance matrix of the measurements is:

$$\begin{vmatrix} .000327 & -.000720 & -.000357 & -.000340 \\ & .020043 & -.000722 & -.012250 \\ & & -.021510 & -.000579 \\ & & & .126965 \end{vmatrix}$$

i) Find the best linear discriminant function for distinguishing between the two groups.

ii) Using a cut off point which makes the two probabilities of misclassification equal, determine the common value of this probability, assuming that all observations are from multivariate normal distributions.

iii) An African, known to belong to one of the two tribes, has the following anthropometric measurements:

ST	HL	HB	NB
1765	190	144	34

Is he more likely to be a Batutsi or a Swahili?

(32)

2.(a) In a sample consisting of 14,345 foxes, 12 were found to be black, 678 intermediate and 13,655 red.

Obtain the frequencies of the black and the red genes in the fox population.

Are the observed numbers consistent with those expected from the Hardy-Weinberg Law?

(b) If there are two brachydactylous individuals per thousand in a large random mating population and it is assumed that homozygous brachydactyly is lethal, what would be the mutation rate from normal to brachydactyly gene in order to balance loss of the latter gene through selection?

(16+16) = 32

3. The following segregations are obtained on backcrossing Groute Locusts (Acridium aranosum), heterozygous for the two factors W and Ny, to the double recessive:

	W Ny	W ny	w Ny	w ny	Total
Repulsion	30	70	2	24	126
Coupling	519	119	12	349	999

(For operative part of this question, see next page)

Q. 3 (cont'd)

Test for linkage between the two factors after testing for single factor segregations. Estimate the recombination fraction between the two factors and obtain its standard error.

(32)

4.(a) The data obtained in a randomised micro-biological assay of pantothenic acid content of plant tissues are given below. The tubes of standard contained .02, .04, .06 and .08 ml. of calcium pantothenate and those of unknowns .05, .10, .15 and .20 ml. of plant extract. At each dose two tubes were tested for the pantothenic acid content. Find out the pantothenic acid content in one millilitre of plant extract along with the fiducial limits assuming a linear relationship between dosage and response:

Dose	Titre (in milli-litre) Standard (1)	Standard (2)	Unknown (plant tissues) (1)	Unknown (plant tissues) (2)
1	2.4	2.7	2.8	2.9
2	3.6	3.8	4.1	3.7
3	7.7	5.0	5.5	5.5
4	6.1	6.3	6.4	6.7
0	1.3	1.7	-	-

4.(b) In an assay with petroleum ether extract by direct spraying on gram weevils, the following mortalities were observed out of 50 exposed at each concentration:

Concentration	Mortality
.0625	41
.03125	32
.01562	20
Control	0

Calculate LD 50 and its fiducial limits. (16+16) = 32

Neatness = (4)

SD-X: GROUP D: DESIGN & ANALYSIS OF EXPERIMENTS

Special Paper III - Practical

(No candidate available

SD-X - GROUP E : SAMPLE SURVEYS - Special Paper III - Practical

(Attempt any three questions from this Group)

1. The following table shows the values of two characters x and y for a sample of 10 units selected by simple random sampling without replacement:

Sample values of x and y

Sr. No. (1)	x (2)	y (3)	Sr. No. (1)	x (2)	y (3)
1	56	46	6	58	52
2	105	88	7	86	67
3	92	60	8	141	110
4	71	55	9	103	78
5	95	77	10	74	60

Population total of x = 16,000.

(For operative part of this question, see next page)

(p.t.o.)

Q.1 (contd)

It is required to estimate the population total of y by using the following estimators:

(a) usual unbiased estimator

(b) usual ratio estimator

(c) unbiased ratio-type estimator

(d) almost unbiased ratio estimator - by splitting the sample into two random groups.

Calculate the required estimates. Also find the efficiency of estimators (b), (c) and (d) as compared to (a).

(12+20) = 32

2. Select a sample of 2 units with probability proportional to size at each draw and without replacement from the population of 5 units given below. Estimate the population total on the basis of your sample and also estimate the sampling variance by Yates and Grundy formula.

Population of 5 units

Unit No.	Size X	Value Y
(1)	(2)	(3)
1	4	24
2	10	50
3	2	8
4	5	25
5	9	33

(8+10+14) = 32

3. The table below gives the number of villages and area under wheat for 30 circles of a district of Uttar Pradesh. It is required to estimate the total area under wheat by drawing a sample of 5 circles and surveying all the villages of the sampled circles.

Circle No.(i)	No. of villages (Mi)	Area in acres (Yi)	Circle No. (i)	No. of villages (Mi)	Area in acres (Yi)
(1)	(2)	(3)	(1)	(2)	(3)
1	6	1510	16	3	723
2	5	1030	17	4	1260
3	4	1692	18	8	2400
4	5	870	19	2	440
5	4	498	20	4	812
6	2	376	21	3	609
7	4	760	22	4	1210
8	2	350	23	4	920
9	5	995	24	3	741
10	1	34	25	5	2100
11	3	627	26	2	300
12	4	749	27	2	130
13	3	534	28	3	927
14	1	186	29	1	210
15	1	82	30	2	408

Calculate the sampling variance and the bias (if any) of the estimate of total area under wheat for each of the following procedures of sampling and estimation:

(a) simple random sampling without replacement,

(i) estimate based on the mean of the circle means

(ii) estimate based on the mean of the circle totals.

(b) sampling with probability proportional to the circle size (i.e. number of villages in a circle) and with replacement estimate being based on the mean of the circle means. (10+10+12) = 32

(372)

4. Draw up a small schedule for collecting data on births, deaths and fertility from a sample of families selected for a demographic enquiry.

Also draw up a scrutiny programme for checking the filled in schedules before processing.

(20+12) = 32

5. For collecting data on household consumer expenditure, a large-scale sample survey was conducted in a state of India. The survey design was a stratified two-stage one with districts as strata, villages as first-stage units and households as second-stage units. The filled-in schedules received at the processing centre are required to be processed and tabulated manually. Different stages of work will be as follows:

i) receipt of schedules in the records section,

ii) scrutiny of schedules in the scrutiny section

iii) primary tabulation of data including calculation of multipliers, weighting of sample values by multipliers etc. in the computation section,

iv) checking of tabulation in the computation section,

v) preparation of final tables in the computation section.

Assuming that each section has got one or more supervisors and several workers under each supervisor, draw up the required forms and proformas for:

(a) recording receipt of schedules from the field

(b) allotment of work to individual workers and units (under supervisors)

(c) progress of work in the units

(d) working sheets for calculation of multipliers and weighted estimates of State totals.

(4+10+8+10) = 32

Neatness .. .. .. (4)

Special
SD-X : GROUP F - TECHNIQUES OF COMPUTATION (Paper III: Practical)
based on punched card equipment - Unit Record Machines).

(Attempt to answer the entire question at all its stages)

For the operations proposed to be carried out by you, you will be supplied one deck of cards designated by 78F2 in column 1 to 4. The card design is given below:

Sl. No.	Card Design description	cols read.	cols allotted	Remarks
A. Identification Particulars				
1.	Identification	4	1 - 4	78F2
2.	Serial No.	3	5 - 7	
3.	Stratum	1	8	
4.	Sub-stratum	1	9	
5.	Sub-sample	1	10	
6.	Sub-round	1	11	
7.	Sample h.h. No.	3	12 - 14	
8.	Party	1	15	
9.	Sl. No. in bl. 3.1	2	16	
10.	Sl. No. in bl. 3.1 h.h	2	17	

Sl. No.	description	cols read	cols allotted	Remarks
B. Pregnancy particulars :				
11.	Have you ever been pregnant?	1	18	
12.	Did you have twins?	1	19	
13.	How many twins	1	20	
14.	Age at consumation of marriage	2	21-22	
15.	Marital status	1	23	
16.	Age in completed years	2	24-25	
17.	Year of effective marriage	2	26-27	
18.	Are you currently pregnant?	1	28	
19.	No. of pregnancies (last 3L.No.)	2	29-30	
20.	Order of pregnancy	2	31-32	
21.	Outcome code and date	5	33-37	
22.	Birth certificate	1	38	
23.	Born alive or not	1	39	
24.	Present age of child	2	40-41	
25.	If dead, age at death	2	42-43	
26.	Sex	1	44	
27.	Interval between successive pregnancy termination	3	45-47	

The various stages of operations required to be carried out by you are indicated below:

Stage 1

(i) Note down the total volume of cards handled, in the answer sheet provided.

(ii) Take the 78F2 cards and sort them by interval between successive pregnancy terminations and form group codes as under, and punch them in column '79' of the deck.

Interval range	Code
Upto 11 month	1
12-17 "	2
18-23 "	3
24-29 "	4
30-35 "	5
36-41 "	6
42-48 "	7
49-59 "	8
60 and above	9
No record	0

Note down the number of cards replaced (if any) during sorting and gang-punching and hand over the damaged cards (if any) to the examiner after the replacement.

(iii) Wire the control panel of a reproducing punch and carry out operations under Stage 1(ii)

(2+10+8) = 21

Stage 2

(i) Consider column 8 of the cards and transfer the weights given below in column 77-78 of the 78F2 cards after Stage 1.

(Next page)

The weights proposed are :

Stratum	Weight
1	12
2	08
3	05
4	00

(ii) Multiply the Pregnancy Interval by the weight transferred in (i). Punch the product 'PI' in columns 72-73.

(iii) Prepare a panel chart to carry out the operations in (ii) above.

(10+10+10) = 30

Stage 3

(i) Sort by Sex (punched in Column 41) & Pregnancy interval group codes (punched in Column 71).

(ii) Tabulate the sorted cards in (i) in the form given below with sub-totals by pregnancy interval class code and also the grand total.

Proforma of table for sorted cards

Sex	Pregnancy-interval class code	number of cards	Summation of 'PI' products
1	0	-	-
1	1	-	-
:	:	:	:
Sub-total →			
2	0	-	-
2	1	-	-
:	:	:	:
Sub-total →			

Grand total

(iii) Prepare a Panel Chart to carry out the operations in (ii) above.

(5+20+15)=40

Stage 4

(i) Write table heading etc. on the machine print and hand over the tabulated sheet, cards and wired panel charts in respect of stages 1,2 and 3 to the examiner.

(ii) Put your signature on the performance sheet provided. Produce it at the end of every stage for counter signature. Hand over this record also to the examiner at the conclusion of the examination.

(10)

(*) For performance record, see Appendix Sheet P.10.

GROUP G : STATISTICAL METHODS
(Special Paper III - Practical)
No candidate available

GROUP H : PROBABILITY THEORY
(Special Paper III - Practical)
No candidate available

(155)

SD-X: GROUP F : NUMERICAL COMPUTATION WITH UNIT RECORD MACHINES

(Paper III : Practical)

Performance Sheet : May 1971

Time of start .....

Stage 1

Number of cards sorted for 78% .....

Number of damaged cards replaced ..... (If any)

Help required to be rendered during operations Yes/No

I Supervisor's counter signature — Time of completion ..... (at the end of first stage)

Stage 2

Number of damaged cards replaced (If any) .....

Assistance given during calculating punch 'wiring'. Yes/No

Assistance given during operating calculating punch. Yes/No

II Supervisor's counter-signature — Time of completion..... (at the end of second stage)

Stage 3

Number of sheets produced .....

Assistance given during panel wiring... Yes/No

Assistance given during operation .... Yes/No

III Supervisor's counter signature — Time of completion..... (at the end of third stage)

Stage 4

Assistance needed in writing Table heading ..... Yes/No

IV Supervisor's counter signature — Time of completion..... (at the end of fourth stage)

NOTE: If at any of the stages of operation, assistance has to be rendered to the candidate, a note of the assistance given against the stage number has to be recorded as below:

Stage No.	Nature of assistance given
1	
2	
3	
4	

Signature of Supervisor/s

Signature of candidate (examinee)

Statistician's Diploma Examination - November 1971

Paper I : Official Statistics and Descriptive Statistics (Theoretical)

Time : 4 hours Full marks : 100

(a) Figures in the margin indicate full marks.

GROUP A

(Attempt any three questions from this group)

1. What do you understand by statistical standards? How are they developed at the international level? Describe the efforts which have been made to develop statistical standards in India in the field of either labour statistics or trade statistics. (3+5+8) = 16

2. Describe the design, scope and coverage of the "Annual Survey of Industries". What are the lacunae in the industrial statistics of India and how can these be removed? (10+6) = 16

3. Indicate in detail the type of data that are thrown up by the Indian Population Census and name the publications which contain these statistics. How have these data helped us in formulating our five-year plans ? (10+6) = 16

4. Which publications present statistics on either foreign trade or education in India? Describe in detail the scope and coverage of any three of these publications and indicate their short-comings. (4+12) = 16

5. Define briefly any eight of the following terms :

a) bank advances
b) balance of payments
c) capital formation
d) coastal trade
e) dead weight tonnage
f) harvest prices
g) index of industrial production
h) international migration
i) labour force
j) literacy
k) morbidity
l) national income
m) national classification of industries
n) passenger miles
o) unemployment
p) value added by manufacture

(2x8) = 16

GROUP B

(Attempt any three questions from this group)

6. (a) Let μ'_r denote the r^{th} moment about zero and μ_r the r^{th} moment about the mean of a random variable; $r = 0, 1, 2, \ldots$. Obtain the formula

$$\mu_r = \mu'_r - \binom{r}{1}\mu'_{r-1}\mu'_1 + \binom{r}{2}\mu'_{r-2}\mu'^2_1 - \ldots + (-1)^{r-1}\binom{r}{r-1}\mu'^{r-1}_1\mu'_1 + (-1)^r\mu'^r_1$$

Hence, derive expressions for μ_2, μ_3 and μ_4 in terms of moments about zero (3+2) = 5

(b) Show that $\beta_2 \geq \beta_1 + 1$ (5)

(c) X is a random variable with

$$P(X=x) = \binom{x-1}{r-1} p^r q^{x-r}, \quad x = r, r+1, \ldots; \quad 0<p<1, \; p+q=1$$

Show that the moment generating function of X about zero

$$i.e. \; E(e^{tX}) = (pe^t)^r (1-qe^t)^{-r}.$$

Hence, derive expressions for the mean and the variance of X. (4+2) = 6

Please turn over

7.(a) Starting with the general differential equation

$$\frac{dy}{dx} = \frac{y(x-a)}{b_0 + b_1 x + b_2 x^2}$$

for the unimodal Pearsonian system of frequency distribution, show that a measure of skewness for such distributions is $\frac{\sqrt{\beta_1}(\beta_2+3)}{2(5\beta_2-6\beta_1-9)}$, where the measure of skewness is defined as (mean-mode)/standard deviation. (8)

(b) Obtain as particular cases (i) the normal, (ii) the type II, and (iii) the type III Pearsonian curves from the general differential equation given in (a). (2+3+3)=8

8.(a) Define the correlation ratio of a variable Y on another variable X. Show that it cannot be smaller than the square of the correlation coefficient between X and Y. How can you use this as a measure of divergence of the regression of Y on X from linearity? (2+4+2)= 8

(b) Define (i) Spearman's rank correlation coefficient, and (ii) Kendall's tau (τ). Derive the expression for Spearman's rank correlation coefficient in terms of rank differences when there are no ties in either of the two rankings. ($1\frac{1}{2}$+$1\frac{1}{2}$+5)= 8

9.(a) Describe the chain-base method in the construction of index numbers indicating the conditions where this method is suitable. (4+2)= 6

(b) Describe the basic components of a time-series and explain how you would measure seasonal variations. What additional steps would you take for the purpose, when the seasonal variation is moving? (2+6+2)=10

10. Comment on the sources of bias in any six from the following :

a) To study voting pattern, a sample of voters is selected by choosing every hundredth man from a list of registered voters and including both him and his wife in the sample.

b) To find the average number of apples per kilogram in a barrel of apples, a sample is drawn by taking a handful from the top.

c) To study the proportion of sand in a mixture of sand and saw-dust, a sample is taken by scooping up a quantity from the bottom.

d) To study the attitudes of readers of a newspaper towards an event of topical interest, readers are sampled by printing in the paper an invitation to them to send their observations on the event.

e) In a household expenditure survey, a random selection of households is made and the requisite information is collected during a year. For repeating the inquiry year after year, the same households are retained in the sample.

f) The size of families in a town is investigated by a house-to-house enquiry in the (i) mornings only, (ii) afternoons only, ignoring those houses at which there is no reply.

g) The mean income per family in a certain town is estimated from a sample chosen from the telephone directory.

h) The average yield of wheat in a village is estimated by taking a simple random sample of wheat growing fields, and taking a sample of harvested produce from each such selected field and calculating a simple average of the yield figures.

($2\frac{2}{3}$ x 6)= 16

NEATNESS (:+2) (:)

INDIAN STATISTICAL INSTITUTE

Statistician's Diploma Examination - November 1971

Paper II : Probability Theory and Statistical Methods (Theoretical)

Time : 4 hours | Full marks : 100

Figures in the margin indicate full marks.

GROUP A

(Attempt any four questions from this group)

1.(a) Let $\mathcal{X} = \{x : 0 < x < 10\}$ denote a sample space and let $P(A_1) = 3/8$, $P(A_2) = 7/8$ be the probabilities of the sets A_1 and A_2 defined by

$$A_1 = \{x : 0 < x \leq 2\}, \quad A_2 = \{x : 0 < x \leq 4\}.$$

Find $P(A_3)$, where $A_3 = \{x : 0 < x < 2\} \cup \{x : 4 < x < 10\}$

(b) Three groups of children contain respectively 3 girls and 1 boy; 2 girls and 2 boys; 1 girl and 3 boys. One child is selected at random from each group. Find the probability that selected children consist of 2 boys and 1 girl. (5+7) = 12

2.(a) Establish the convergence of Binomial distribution to the Poisson distribution, stating the assumptions involved.

(b) If $X_1, X_2, \ldots, X_n$ are independent random variables, each taking values 1 and 0 with probabilities θ and $1 - \theta$ respectively, then find the probability generating function of their sum $S_n = X_1 + \ldots + X_n$. (6+6) = 12

3. Determine K so that

$$f(x_1, x_2) = K \cdot x_1^2 \cdot x_2^3, \quad 0 < x_1 < x_2 < 1$$
$$= 0 \text{ elsewhere,}$$

may be the joint density function of two random variables X_1 and X_2.

Also, determine the variance of X_2 and the conditional variance of X_1 given $X_2 = x_2$, where $0 < x_2 < 1$. (3+3+6) = 12

4.(a) Let X be a random variable taking values -1, 0 and 1 with probabilities $\frac{1}{8}$, $\frac{6}{8}$ and $\frac{1}{8}$ respectively. Show that $P(|X| \geq 1) = \frac{1}{4}$ which is also the upper bound provided by Chebychev's inequality.

(b) Let X be a random variable such that $P(X \leq 0) = 0$ and let $\mu = E(X)$ exist.

Show that $P(X \geq 2\mu) \leq \frac{1}{2}$. (2+4+6) = 12

5. What is meant by the 'Sampling distribution' and the 'Standard error' of a Statistic ?

If $x_1, \ldots, x_n$ are independent observations from a $\mathcal{N}(\theta, 1)$ distribution, then obtain the sampling distribution of $u = \sqrt{n}(\bar{x} - \theta)/s$, where $\bar{x} = \sum_1^n x_i/n$ and $s^2 = \sum_1^n (x_i - \bar{x})^2/(n-1)$. (1+1+10) = 12

Please turn over

(3)

GROUP B

(Attempt any three questions from this group)

6.(a) Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed random variables. Under θ, let $P_\theta \left\{ X_i = 1 \right\} = \theta$ and $P_\theta \left\{ X_i = 0 \right\} = 1 - \theta$, $0 < \theta < 1$.

Find the Cramer-Rao lower bound to the variance of unbiased estimators of θ and hence find the efficient unbiased estimator of θ.

(b) Let $X_1, \ldots, X_n$ be independently distributed with a common 'Cauchy' density

$$f_\theta(x) = \frac{1}{\pi} \left\{ \frac{1}{1 + (x - \theta)^2} \right\}.$$

Find the Cramer-Rao lower bound to the variance of an unbiased estimator of θ.

Does there exist an unbiased estimator, which attains this lower bound? (8+8) = 16

7. Explain the method of maximum likelihood for point estimation. Derive the maximum likelihood estimators of the parameters of a bivariate normal population (with all the parameters unknown) given a random sample of size n :

$$(x_1, y_1), \ldots, (x_n, y_n).$$

How will the estimates be modified if the observations $y_{n+1}, \ldots, y_n$ $(n < n)$ be not available ? (5+7+1) = 16

8. With reference to the problem of testing a simple hypothesis involving a single real parameter, define a uniformly most powerful unbiased test.

Given a random sample of size n from the population $N(m, 1)$, derive the uniformly most powerful unbiased test for the hypothesis $m = 0$ against $m \neq 0$, and write down the power function of this test. (4+12) = 16

9.(a) Samples of various sizes are available from each of $K_1 + K_2$ normal populations, with unknown means and variances. The first K_1 of these populations are known to have the same variance equal to σ_1^2 (say) and the remaining K_2 are also known to have the same variance equal to σ_2^2 (say).

Derive the likelihood ratio test for the hypothesis $\sigma_1^2 = \sigma_2^2$

(b) Indicate how one can compute the power function of the suggested test procedure. (10+6) = 16

10. Write short notes on the following :

(a) Simple and Composite Hypothesis.

(b) Level of significance and power of a test.

(c) Shortest and shortest unbiased confidence intervals.

(5+5+6) = 16

INDIAN STATISTICAL INSTITUTE

Statistician's Diploma Examination - November 1971

Paper III : Sample Surveys and Design & Analysis of Experiments (Theoretical)

Time : 4 hours | Full marks : 100

Figures in the margin indicate full marks

GROUP A

(Attempt any three questions from this group)

1. (a) Define :

i) Sampling frame

ii) Simple random sampling without replacement.

iii) Simple random sampling with replacement.

(b) A population of N units contains N_1 units belonging to class A and $N_2 (= N - N_1)$ units belonging to class B. Obtain an unbiased estimator of N_1, and an unbiased estimator of its variance, under simple random sampling without replacement.

(6+10) = 16

2. (a) A population is divided into K strata. Obtain the value of n_i, the size of the sample to be selected from the ith stratum, so that the standard unbiased estimator of the population mean, has maximum precision for a fixed cost.

The cost of sampling a unit is constant in different strata.

(b) Obtain the expression for gain in precision due to stratification under the allocation of the sample as obtained in (a) above. What guidelines does it provide for stratifying a population?

(4+12) = 16

3. (a) With an illustration explain the procedure of systematic sampling.

(b) Is the sample mean always an unbiased estimator of the population mean, under this procedure of sampling?

(c) Obtain the expression for the efficiency of systematic sampling relative to simple random sampling without replacement, and interpret the result.

(3+3+10) = 16.

4. (a) Describe the ratio and regression methods of estimation.

(b) Obtain the approximate expression for the variances of the ratio and regression estimators. When are the two estimators equally efficient?

(4+12) = 16

5. Write notes on any two of the following :

i) Probability - proportional-to-size sampling.

ii) Choice of sampling unit.

iii) Interpenetrating sub-samples. (8+8) = 16.

Please turn over

GROUP B

(Attempt any three questions from this group)

6. (a) Why do you need a design at all? Discuss the requirements of a good experiment.

(b) Set out the important assumptions underlying the analysis and interpretation of data obtained from the standard experimental designs.

(10+6)= 16

7. Explain the principles of covariance analysis. Describe the different computational steps in the covariance analysis of a randomised block experiment with 'b' blocks and 't' treatments, there being only one concomitant variable. Derive the average variance for the difference of two adjusted treatment means.

(5+7+6)= 18

8. What is an orthogonal design? Is BIBD an orthogonal design ? How does a Youden Square differ from a BIBD? When is a BIBD called resolvable? Show that for a resolvable BIBD with parameters b, v, r, k and λ the following inequality must hold :

$$b \geq v + r - 1 .$$

(5+2+2+2+7)= 18

9. (a) A 3^3 experiment with factors A,B and C is conducted in blocks of 9 plots and 4 replicates, confounding ABC, AB^2, AB^2C, AB^2C^2 in the four replicates respectively.

Find the composition of the blocks in each replicate and describe the method of analysis.

(b) Distinguish between confounding and fractional replication and discuss when they are appropriate.

(12+6)= 18

10. Write notes on any three of the following :

i) Efficiency of a design.

ii) Lattice designs.

iii) Construction of 5 x 5 orthogonal Latin Squares.

iv) Connected design.

(15)

[illegible] (A and B) (3)

INDIAN STATISTICAL INSTITUTE

Statistician's Diploma Examination - November 1971

Paper IV : Applied Statistics (Theoretical)

Time : 4 hours | Full marks : 100 (For two groups combined)

(a) Candidates will be required to answer questions from these two groups of subjects only for which they have registered their options.

(b) Separate answer books are to be used for each of the two groups attempted.

(c) Figures in the margin indicate full marks.

(d) Use of calculating machines is not permitted.

GROUP A : ECONOMIC STATISTICS (50 marks) - Half-paper

(Attempt any three questions from this group)

1. Describe the different components of a time series. Establish the effect of eliminating trend by the method of moving averages on the other components of a time series. (6+10)= 16

2. (a) Discuss the steps you would take in constructing a cost of living index number for a certain class of people in a city.

(b) Show that the factor reversal test and the time reversal test are not satisfied by Laspeyre's and Paasche's index numbers. Further show that both these tests are satisfied by Fisher's Ideal index numbers. (6+8)= 13

3. What do you mean by income-elasticity of demand? How would you estimate this elasticity from family budget data? What adjustments would you make for variation in the size of the family? (5+8+4)= 13

4. (a) Define the concentration curve and coefficient of concentration of a distribution.

(b) Write down Pareto's law of income distribution and derive its coefficient of concentration

(c) Write a note on the salient features of any other well-known law of income distribution. (4+6+3)= 13

5. (a) Considering only two distinct inputs, viz labour (L) and capital (C) the Cobb-Douglas production function relates output X to inputs as follows :

$$X = A L^{\alpha} C^{\beta}$$

where A, α and β are positive constants.

From a time-series data of output and the two inputs how would you estimate the parameters?

(b) If L and C are intercorrelated, how would you obtain more precise estimates if outside information (say from a sample survey data) regarding α were available? (10+3)= 13

MATHS (2)

Please turn over

GROUP D : STATISTICAL QUALITY CONTROL (50 marks) - half-hour

(Attempt any three questions from this group)

1. (a) Explain clearly the distinction between control limits and specification limits.

(b) A Company using an $\bar{X}$ and R chart is having two sided specification limits for the characteristic.

i) How will you obtain process capability?

ii) Briefly discuss the comparison of process capability with specified tolerance range assuming that the process is in a state of statistical control.

(c) A Company is maintaining an $\bar{X}$ and R chart on a measurable characteristic, on a routine basis. The range chart is showing a fairly good state of control. The control chart for $\bar{X}$ is showing too frequent violations of the control limits on both the sides.

Interpret such a situation and give a procedure for revising the control limits if you think it necessary. (4+6+4)= 14

2. (a) Given AQL = p_0, LTPD = p_1 and the producer's and consumer's risks, α and β respectively, explain the detailed procedure of devising a single sampling plan by attributes for acceptance of lots and satisfying the stipulated requirements.

State clearly the assumptions that you may make.

(b) Obtain the acceptance criterion for lots, under item by item sequential sampling inspection, for values of AQL, LTPD, producer's risk and consumer's risk as given in (a).

(c) Discuss briefly the relative advantages and disadvantages of single, double, multiple and sequential sampling plans. (5+3+5)= 13

3. (a) Explain the difference between 3σ limits and probability limits of a control chart.

(b) A machine is producing on an average, 2% defective items. It has been decided to instal a p-chart for the process using a sample size of 30. If it is desired to detect a change of 1% in the percentage of defectives on either side with a probability of 0.1 what would be the control limits?

State the assumptions that you make.

(c) Suppose you are using the control limits for p as obtained in (b) and half-hourly samples of size 30 are being taken. If the process shifts by 2% on either side of the standard (i.e. 2% ± 2%) it is desired to detect this shift within two hours on an average after the shift has occurred.

Would you consider the present sample size to be adequate for the purpose? If not, what should be the sample size? State the assumptions you may make. (4+3+6)= 13

4. Write short notes on any three of the following:

(a) Modified control charts

(b) Group control charts

(c) Lot-plot technique

(d) Dodge-Romig's plans

(e) Dodge's continuous sampling plan (15)

(c)

GROUP D : STATISTICAL TABLES & FORMULAE

(to candidate available)

Please turn over

GROUP D : VITAL STATISTICS AND DEMOGRAPHY (50 marks)

(Attempt any three questions from this group).

1. (a) Distinguished between (i) 'de facto' and 'de jure' populations and (ii) householder and enumerator methods of census-taking.

(b) Briefly describe the procedure of constructing an abridged life table, given total number of persons and number of deaths for each five year age group. (3+3+10) = 16

2. (a) Examine the merits and demerits of net reproduction rate as a measure of the prospect of population growth.

(b) Derive a relation connecting the true rate of natural increase and net reproduction rate. State clearly the assumptions you make.

(c) Show that in a stationary population, the average age at death is the reciprocal of the crude birth rate. (3+3+4) = 16

3. How do you differentiate population estimates from population projections? Give an outline of the component method. Compare the component method with methods using mathematical curves for population projection. (4+3+9) = 16

4. Write notes on any two of the following :

(a) Force of mortality

(b) Cohort life table versus current life table

(c) Comparative mortality index. (8+8) = 16

STATISTICS (2)

GROUP E : ECONOMIC AND PSYCHOLOGICAL STATISTICS (50 marks) - Half paper

(Attempt any three questions from this group)

1. Explain briefly the similarities and differences :

(a) between an achievement test and an aptitude test,

(b) between reliability and validity,

(c) between item difficulty and item discrimination measure, and

(d) between factor analysis and regression analysis. (16)

2. Describe the Kuder-Richardson method of assessing the reliability of a test. Give a brief account of the

(i) test-retest,

(ii) parallel form, and

(iii) split-half method of assessing the reliability,

commenting on their merits and demerits. (16)

3. Define product moment correlation and rank correlation and describe the situations in which each is applied. Distinguish between the concepts of correlation and regression, and describe how the multiple regression technique can be used in selection and guidance work.

4. Describe Spearman's two-factor theory and Thurstone's multiple factor theory of factor analysis. Indicate the methods of obtaining factor loadings in these cases. Give an account of the method of principal components as applied to the problem of estimation of factor loadings. (16)

5. 'Learning data are well adapted to mathematical treatment' - Elucidate the above statement and indicate the role of mathematical model in the measurement of learning. Assuming a mathematical model in a learning situation, derive the maximum likelihood estimators of the parameters. (16)

STATISTICS (2)

INDIAN STATISTICAL INSTITUTE

Statistician's Diploma Examination - November 1971

Paper V : Methods of Numerical Computation, Descriptive Statistics and Official Statistics (Practical)

Time : 5 hours Full marks : 100

i) Figures in the margin indicate full marks.

ii) Use of calculating machines is permitted.

GROUP A : (25 marks)

(Attempt any two questions from this group)

1. (a) Given that $f(0) = 23.1234$

$f(6) = 23.7252$

$f(12) = 24.0833$

$f(18) = 26.1330$

Complete the series from $f(6)$ to $f(12)$, by sub-tabulation of intervals.

(b) Given the following table values, obtain the value of log 3375 by making use of a suitable central difference formula.

x	:	310	320	330	340	350	360
log x	:	2.4914	2.5052	2.5185	2.5315	2.5441	2.5563

$(6+6\frac{1}{2}) = 12\frac{1}{2}$

2. (a) Following are the values of an empirical function $f(x)$ for certain values of x :

x	:	0.4	0.5	0.6	0.7	0.8
$f(x)$	:	1.5333	1.7071	2.0412	1.9352	1.8321

Evaluate $\frac{df(x)}{dx}$ and $\frac{d^2 f(x)}{dx}$ at the point x = 0.65. Also find the value of x for which the function $f(x)$ is a maximum.

(b) Find the real root of the equation : $x^2 + 4 \sin x = 0$ correct to four places of decimals by any method that you consider suitable. $(7+5\frac{1}{2}) = 12\frac{1}{2}$

3. Find the inverse of the following symmetric matrix :

1.000	0.213	0.266	0.307
	1.000	0.628	0.553
		1.000	0.489
			1.000

$(12\frac{1}{2})$

Please turn over

GROUP B (50 marks)

(Attempt any four questions from this group)

4. The data given below relate to the marks secured in English by a group of 60 high school students both males and females belonging to different religions. The following abbreviations have been used.

Sex : m = males, f = females

Religion : H = Hindu, M = Muslim, C = Christian and R = Other religions.

(Hm) - 32	(Mf) - 43	(Cm) - 20	(Rm) - 18	(Rm) - 64
(Hm) - 58	(Hm) - 21	(Cf) - 27	(Rf) - 56	(Mf) - 13
(Mf) - 33	(Rm) - 50	(Cm) - 32	(Mf) - 12	(Rm) - 15
(Rm) - 44	(Mf) - 45	(Cm) - 41	(Mf) - 57	(Cf) - 19
(Mf) - 68	(Mf) - 30	(Mf) - 60	(Hm) - 58	(Hm) - 55
(Cf) - 23	(Mf) - 34	(Cf) - 20	(Hm) - 25	(Hm) - 46
(Hm) - 47	(Cm) - 51	(Mf) - 39	(Rm) - 63	(Hm) - 55
(Mf) - 33	(Mf) - 36	(Rm) - 69	(Cf) - 43	(Hm) - 66
(Mf) - 26	(Hm) - 40	(Rm) - 25	(Cm) - 53	(Mf) - 23
(Hm) - 21	(Cf) - 22	(Cm) - 48	(Mf) - 27	(Hm) - 27
(Mf) - 41	(Cm) - 52	(Rm) - 62	(Rm) - 24	(Hm) - 39
(Rm) - 74	(Hm) - 15	(Mf) - 38	(Hm) - 36	(Rm) - 55

Present the information in the form of table giving a suitable heading. Comment on the performance of religion and sex group. (12)

5. (a) The mean weight of 500 male students at a certain college is 60 kg. and the standard deviation is 7 kg. Assuming that the weights are normally distributed, find how many students are expected to weigh (i) between 54 kg. and 70 kg. and (ii) more than 84 kg.

(b) In a bag there are 20 black beans and the remainder are white beans. Determine the probability that a single sample of 8 beans will contain 6 or more white beans. (6+6) = 12

6. In a random sample consisting of 54 women, the intake of two nutrients (protein and fat) was determined together with the age and the concentration of cholesterol in blood of these women.

	Age	Protein	Fat
Protein	-0.1865	-	
Fat	-0.5290	+ 0.5794	
Cholesterol	+0.4737	- 0.4210	- 0.3125

Obtain the correlation between age and cholesterol, eliminating the effect of the intake of protein and fat. (12)

Please turn over

7. The table below gives the net earnings by Government Railways for fifteen years (1951-1965):

Net earnings by Government Railways

year	net earnings (billion Rs.)	year	net earnings (billion Rs.)	year	net earnings (billion Rs.)
1951	632	1956	761	1961	1026
1952	621	1957	713	1962	1178
1953	605	1958	676	1963	1260
1954	624	1959	634	1964	1345
1955	671	1960	872	1965	1419

i) Obtain the trend by fitting a second-degree curve.

ii) Estimate the earnings during 1970 and 1971. (3+4)= 12

8. The table below gives the monthly index numbers of wholesale prices (with 1961-62 = 100 as the base) during the five years from 1966 to 1970.

Monthly index number of wholesale prices
(general index for all commodities)

month	year				
	1966	1967	1968	1969	1970
January	135	155	156	163	174
February	135	159	163	162	174
March	137	159	166	165	176
April	141	160	163	165	176
May	145	164	164	167	179
June	149	168	163	172	181
July	151	173	163	174	181
August	151	171	170	174	182
September	153	172	174	174	183
October	149	173	171	171	183
November	151	172	168	168	181
December	153	167	162	170	181

Find monthly seasonal indices by the moving average method. Hence deseasonalize the original data. (3+9)= 12

9. Compute an index number to show the average variation in the production of the following group of commodities :

Commodity	Units produced ('000)			Weight
	q_0	q_1	q_2	w_0
A	65	75	82	1.5
B	28	32	40	2.0
C	150	127	115	1.2
D	83	85	72	1.0

q_0 = quantity produced in year y0 (base year)

q_1 = quantity produced in year y1

q_2 = quantity produced in year y2

w_0 = number proportional to value at cost in year y0 = weight

(12)

HEADINGS (2)

Please turn over

GROUP C (25 marks)

(Attempt both the questions from this group)

From the official publications placed at your disposal :

10. Obtain from the 1961 Indian Census data,

a) the total population of India

b) the working population of India shown classified under various economic categories

c) the total non-working population

The figures are to be quoted in terms of millions only.

Tabulate the collected data neatly and compare in terms of the total working population, the distribution of male and female workers under the various specified economic categories. Any comments ? (13)

11. Present neatly the information in respect of any three of the following for the latest two available consecutive years :

I) Index numbers of :

a) agricultural production of India under

i) total foodgrains, ii) total oilseeds, iii) total fibres

b) industrial profits in India under

i) mining and quarrying, ii) other industries,

iii) private limited companies

II) 'Installed capacity' and 'production' of

a) vegetable oil products, b) cigarettes, c) cycle tyres.

III) Quantity and value of exports under

a) coffee (unroasted), b) leather footwear,

c) manganese ore and concentrates

IV) Figures of

a) net national product (i.e. national income)

b) per capita net product at

i) current prices

ii) 1960-61 prices

V) Figures of average per capita annual earnings

a) for factory workers for the States of Maharashtra and West Bengal

b) for factory employees of

i) chemicals and chemical products

ii) electricity, gas and steam (3x4) = 12

N.B. Note that full references of publications consulted, are required to be noted alongside answers.

INDIAN STATISTICAL INSTITUTE

Statistician's Diploma Examination - November, 1971.

Paper VI : Statistical Methods : Design & Analysis of Experiments and Sample Surveys (Practical)

Time : 5 hours Full Marks : 100

(a) Figures in the margin indicate full marks.
(b) Use of calculating machines is permitted.

GROUP A.

(Attempt two questions from this group)

1.(a) The additional hours of sleep gained after using each of two drugs by the same group of 12 patients are given below:

Patient	Additional hours of sleep		Patient	Additional hours of sleep	
	Drug 1	Drug 2		Drug 1	Drug 2
1.	2.1	3.6	7.	- 1.3	- 0.4
2.	0.2	4.7	8.	- 0.3	1.9
3.	0.9	1.8	9.	- 1.7	2.0
4.	3.8	5.5	10.	0.7	1.7
5.	3.5	4.6	11.	- 0.8	2.1
6.	- 0.2	- 0.3	12.	1.3	1.1

Test (i) whether the second drug gives, on the average at least an hour more of sleep than the first drug, (ii) whether the Std. devns. of the two series are significantly different.

(b) Estimates of the error variance in three comparable experiments together with their degrees of freedom are given below: Examine whether the variances could be regarded as equal, except for fluctuations of sampling.

Comparable Experiment.	Degrees of freedom	Error variance (estimate).
1.	6	3203
2.	8	8370
3.	11	2696

(12+8) = 20.

2 (a) The distributions in four blood-group classes O, A, B and AB of 140 Christians who are army cadets and 295 other Christians are given below. Can the two samples be regarded as coming from the same population ?

Blood-group frequencies.

Samples.	Blood Group			
	O	A	B	AB
Army Cadets	56	60	18	6
Others.	120	122	42	11

(b) To compare the average number of available shaves from three popular brands of safety razor blades, an experiment was conducted in which two blades of each brand were tried by each of four individuals. The number of shaves obtained from these blades are given below. Analyse the data.

Individuals	Number of shaves from blades.		
	B_1	B_2	B_3
1.	8, 2	7, 3	9, 6.
2.	4, 9	5, 3	11, 9.
3.	6, 3	8, 7	12, 10.
4.	7, 5.	10, 3	8, 9.

(10+10) = 20.

Please turn over

3. A series of values of a random variable are given below:

5.105, 7.283, 6.060, 5.059, 5.627, 6.029
7.672, 6.218, 7.417, 7.672, 5.095, 4.849, 6.133.

It is known that the distribution is normal with mean 6. The standard deviation is also known to be either 1 or 2.5. Starting with the first five observations as a first sample, proceed sequentially to test the hypothesis H_0 that = 1, against the alternative that = 2.5. Take = 0.05 and = 0.01.

Also draw the OC curve of the test.

GROUP B

(Attempt both questions from this group)

4. EITHER

Sixteen treatment combinations representing four factors each at two levels, are arranged in four blocks of four plots each. The intra-blocks of four plots each. The intra-block subgroup is (1), abc, abd, cd. Determine the confounded effects and interactions.

OR

A, B and C stand for three factors, each at three levels. The 3^3 treatment combinations are replicated once in 3 blocks of 9 plots such that AB^2 is confounded. Give the lay-out. (10)

5. EITHER

The following table gives the plan and yields of a 2^4 field experiment on Beans. The yields are given in lbs and the manurial factors as follows:

Dung (d): 0, 1 ton per acre.
Nitrochalk (n): 0, 0.4 cwt. of N per acre.
Super Phosphate (p): 0, 0.6 Cwt of P_2O_5 per acre.
Muriate of Potash (K)0, 1.0 cwt of K_2O per acre.

Analyse the data and interpret your results.

	BLOCK I				BLOCK II			
Rep.I	p	k	d	npk	dp	nk	dk	pk
	46	56	44	37	50	44	43	51
	dnk	dnp	dpk	n	dnpk	(1)	dn	np
	42	49	55	41	41	43	57	49

	BLOCK III				BLOCK IV			
Rep.II	npk	d	n	dkn	nk	dp	(1)	np
	42	41	38	34	44	53	58	40
	n	dnp	k	dpk	nk	dk	dnpk	dn
	43	53	51	45	56	52	54	42

OR

A nutrition experiment on the growth of pigs was carried out with 4 treatments. From each of 4 lots, 4 pigs were randomly chosen, and, for each of the 16 pigs the initial weight in lbs. (x) were recorded. For each lot, the 4 treatments were then allotted randomly to the selected pigs and the experiment carried out.

Please turn over.

After the experiment the rate of gain in weight in lbs. per day [illegible] was noted for each pig. The figures came out as follows. Analyse the data to see if, apart from variations in initial weights, the treatments differ significantly among themselves.

TREATMENTS.

	I		II		III		IV	
Lots	x	y	x	y	x	y	x	y
1.	81	1.40	75	1.31	80	1.67	62	1.91
2.	54	1.29	64	1.12	47	1.73	62	1.90
3.	45	1.53	52	1.29	47	1.39	57	1.42
4.	57	1.34	50	1.30	59	1.49	33	1.52

(20)

GROUP C

(Attempt both the questions from this group)

6. EITHER

Drew a random sample of size 10 from a normal population with mean 3.50 and S.D. 2.12.

OR

In the following table, the first figure relates to the village No. and the second figure relates to the total area in acres under cultivation for 20 villages in a tehsil. Select 5 villages with replacement and probabilities proportional to area.

Village Serial Number.	Total Cultivable Area	Village Serial Number.	Total Cultivable Area.
1.	870	11.	1089
2.	885	12.	1111
3.	894	13.	1122
4.	901	14.	1177
5.	914	15.	1203
6.	973	16.	1224
7.	995	17.	1239
8.	1031	18.	1269
9.	1043	19.	1302
10.	1084	20.	1350

7. Raw wool contains varying amount of grease, dirt and foreign material and its quality is measured by the "Clean Content" - the percentage which the weight of the clean wool bears to the original weight of the raw wool. To estimate the clean content an electrical core boring machine is used, which takes cores of about ½ lb. from a bale, which is then subjected to laboratory analysis.

In an experiment, 6 bales were drawn from a lot and in each 4 cores were taken and clean content measured. Estimate the average clean content from the data given below and standard error of the estimate. Would it have been better to take 12 bales and 2 cores from each ? Give your views.

BALES

Core	1	2	3	4	5	6
1.	52.3	57.0	51.9	54.6	59.9	57.9
2.	58.2	59.7	60.1	57.5	57.9	59.7
3.	62.9	58.2	59.7	51.3	60.3	59.6
4.	50.5	57.4	55.6	57.5	57.5	58.1

(16)

Please turn over.

OR .

In a survey for the production of milk from private cow sheds in a city, the cow sheds were classified into three strata according to the number of milch animals in them. A simple random sample of cowsheds was chosen from each stratum. The following data are given.

Estimate the total production of milk.

Also, estimate the gain in precision due to stratification.

Number of cow sheds in stratum	Number of cow sheds selected.	Milk yields (in litres) of selected sheds during the period.
I (one or two milch animals per shed.)	2461	19	0, 27, 1053, 0, 7, 0, 0, 0, 0, 81, 418, 0, 132, 337, 0, 0, 0, 0, 0.
II (Three to six milch animals per shed.)	2330	27	960, 56, 0, 0, 0, 0, 513, 170, 913, 1627, 661, 0, 511, 0, 1361, 232, 0, 991, 477, 905, 864, 2422, 2050, 303, 0, 0, 565.
III (More than six milch animals per shed.)	520	6	1535, 1744, 1821, 2496, 7974, 10233.

(20)

INDIAN STATISTICAL INSTITUTE

Statistician's Diploma Examination - November, 1971.

Paper VII : Applied Statistics (Practical)

Time : 6 hours Full marks : 100

(a) Candidates will be required to answer questions from those two groups of subjects only, for which they have registered their options.

(b) Separate answer books are to be used for each of these two groups attempted.

(c) Figures in the margin indicate full marks.

(d) Use of calculating machines is permitted.

GROUP A : ECONOMIC STATISTICS Half Paper 50 Marks.

(Attempt any two questions from this group)

1. The following table gives the monthly expenditure on clothing and monthly total household expenditure of the civilian staff employed in Defence headquarters.
Derive the Engel curve for clothing, assuming its form to be a straight line in double logarithmic scale.
Obtain also the income elasticity of demand for clothing.

Family Group	Number of households	Average family size	Average monthly household expenditure (in Rs.) on clothing	total
(1)	(2)	(3)	(4)	(5)
1	439	4.5	16.6	174.4
2	361	5.0	17.1	198.3
3	129	5.1	22.6	252.9
4	784	5.4	24.7	297.4
5	192	5.4	26.2	342.2
6	49	5.2	27.6	337.5
7	48	4.7	33.9	468.2
8	40	6.6	47.7	570.1
9	73	6.0	47.2	869.2
10	45	6.4	71.1	1253.3

(25)

2. Fit a Cobb-Douglas production function of the form, $P = A L^{\alpha} C^{\beta}$ to the following time-series data relating to Indian Iron & Steel industry.

time points	Indices of net output P	employment L	capital C
(1)	(2)	(3)	(4)
1.	82.5	40.7	92.7
2.	87.3	55.3	108.0
3.	78.4	65.0	97.6
4.	93.2	71.5	100.6
5.	92.5	81.1	112.3
6.	95.9	94.1	105.0
7.	99.9	100.6	99.0
8.	100.3	98.1	103.0
9.	120.4	100.2	117.7
10.	120.2	105.0	131.0

Please turn over

(25)

3.(a) A regression equation of the form $Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \epsilon$

Where Y - log consumption of pork per head.
X_1 - log income per head.
X_2 - log relative price of pork.
β_1 - the income elasticity coefficient
β_2 - the price elasticity of demand for pork.

is proposed to be fitted to the data on consumption of pork in U.K. during the years 1920-1938.
The analysis of a set of family budget data yields the value

$\hat{\beta}_1 = 0.575793$ as the estimate of β_1 with an estimated variance of 0.0559764. From the annual figures of aggregate consumption income, etc. the following values of squares and products of deviations are obtained.

$\Sigma y^2 = 0.0352895$. $\Sigma x_1 y = -0.0025041$

$\Sigma x_1^2 = 0.0086619$ $\Sigma x_2 y = -0.0235779$

$\Sigma x_2^2 = 0.0410816$ $\Sigma x_1 x_2 = 0.09714$

How would you estimate β_2 utilising the a priori information regarding β_1. Comment the precision of the estimate with that of the least square estimate.

3.(b) Fit a quadratic trend to the data given below :

Production of Coal in India

Year	Production (000 metric tons)	Year	Production (000 metric tons)
(1)	(2)	(1)	(2)
1950	32,826	1954	37,472
1951	34,301	1955	38,839
1952	36,887	1956	33,910
1953	36,557		

$(12^1 + 12^4) = 25.$

PAPER VII - GROUP B : STATISTICAL QUALITY CONTROL

(Attempt any two questions from this group)

1.(a) Data from 30 heats were collected for the rejections that can be seen at the 'knock-off' stage for castings. The number of bad castings obtained out of 100 inspected for each heat, is given below:

0	14	7	4	0	5
2	1	11	7	4	7
0	6	4	4	4	1
5	6	7	7	0	0
8	2	3	10	3	5

Examine the data and set-up the standard for control chart on 'number-defective'.

Please turn over

1.(a) The control chart had been used over a long time and the rejecti were found stable at 4 percent defective. The foreman decided use a sample of size 25. Find the control limits for the revise chart. At a later date, the number of defectives observed in tw samples was as follows :

0	0	2	3	0
2	1	1	1	0
0	1	2	1	1
2	1	1	3	3

The Quality Control in-charge decided on the basis of these data, that the process has to be checked. Do you agree ? If so, give reasons. (10+25)= 2

2. The outside diameter of a head plug after knurling is known to be dependent on the value before knurling. To study the actual relationship, measurements were taken on 30 randomly chosen items over a period of three days, before and after knurling. The relevant measurements in coded units are given below :

Sample no.	Diameter	
	Before knurling	After knurling
1	520	559
2	462	503
3	508	550
4	420	465
5	480	505
6	470	500
7	505	550
8	492	572
9	492	523
10	490	490
11	578	593
12	493	545
13	490	529
14	506	545
15	500	530
16	490	495
17	510	592
18	572	555
19	543	544
20	522	515
21	515	499
22	555	552
23	509	530
24	494	550
25	523	570

Sample no.	Diameter	
	Before knurling	After knurling
26	513	495
27	499	535
28	481	499
29	509	534
30	495	502
31	500	570
32	500	525
33	490	570
34	490	470
35	509	549
36	495	548
37	535	507
38	502	500
39	490	529
40	529	500
41	459	515
42	465	540
43	450	500
44	510	509
45	453	537
46	450	490
47	500	502
48	475	570
49	529	503
50	472	573

If the diameter after knurling is to be maintained within 550 ± 50 units, within what limits should the value be held before knurling? Analyse the data and establish the required limits.

State clearly the assumptions involved in arriving at your answer. (25)

[Knurling is a process of bead or ridge to help better grip.]

Please turn over

3. (a) Under a quality control plan it was decided that a lot with 4 defects per unit would be accepted in 10 percent of cases and a lot with 10 defects per unit would be accepted in 95 percent of cases.

Prepare a 'defect per unit' sampling scheme where the inspection unit is 1000 square metres of cloth.

(b) Draw OC curve for the following single sampling plan n = 150, C = 3 for lots of size 5000, and obtain the AOQL.

State any assumptions that you make. (12+13)= 25

PAPER VII - GROUP C : STATISTICAL METHODS IN GENETICS

(No candidate available)

PAPER VII - GROUP D : VITAL STATISTICS AND DEMOGRAPHY

(Attempt any two questions from this group)

1. A part of a life table is given here with most of the entries missing. On the basis of the available figures, supply the missing ones and complete the table.

Age (x)	l_x	d_x	$1000q_x$	L_x	T_x	$\mathring{e}_x$
10	93,102		0.82			
11			0.93			
12			0.72			
13			0.80			
14			0.85			
15			1.05			
16			1.12			
17			1.23			
18			1.32			
19			1.10		4,012,443	

Hence determine the probabilities

a) that a child of age 10 will live at least 5 years more.

b) that two children aged 10 and 11 will each live at least 5 years more.

c) that of two children aged 10 and 11, at least one will die within 8 years.

2. Decide, giving reasons, which of the two formulae - Gompertz's and Makeham's, you would consider more suitable for the graduation of the following data. Compute the constants of graduation for the formula you choose.

Age (x)	l_x	Age (x)	l_x
30	89035	45	77913
35	86127	50	72703
40	82277	55	66566

Find also the fitted values of l_x and thereby obtain by interpolation or otherwise, the value of $\sum_{t=1}^{5} l_{x+t}$ (10+15)

Please turn over

3. The following table may be taken to be characteristic of the Indian population in respect of females :

age-group	age-specific marital fertility rate	proportion married	life-table survivors ($l_x=1,000$)
15 - 19	.193	.551	2304
20 - 24	.303	.830	2262
25 - 29	.227	.947	2247
30 - 34	.190	.947	2213
35 - 39	.121	.927	2147
40 - 44	.067	.736	1950

Calculate the net reproduction rate, ignoring illegitimate fertility, and assuming that the sex-ratio at birth is 106 male births to 100 female births. (25)

PART VII - GROUP B : EXPECTATION AND STOCHASTIC STATISTICS

(Attempt all questions from this group)

1. Out of 20 students, 10 have been awarded a prize. Receipt and non-receipt of the prize are indicated in the following table by the values 1 and 0, respectively, of a variable x ; the scores on an intelligence test taken by the same group of students are shown side by side as values of another variable y.

sl. no.	x	y	sl. no.	x	y	sl. no.	x	y	sl. no.	x	y
1	1	73	6	1	78	11	0	64	16	0	49
2	1	78	7	1	63	12	0	57	17	0	62
3	1	64	8	1	64	13	0	61	18	0	71
4	1	63	9	1	72	14	0	55	19	0	61
5	1	72	10	1	72	15	0	60	20	0	46

Compute a suitable correlation coefficient between x and y, and write a note on the result. (10)

2. The following table gives the product moment correlations of the two predictors x_1 and x_2, and the criterion y, along with their means and standard deviations.

	x_1	x_2	y	mean	standard deviation
x_1	-	.206	.473	34.1	9.16
x_2	.206	-	.160	29.2	8.130
y	.473	.160	-	61.2	8.76

Calculate the regression coefficients and the constant term for the multiple regression equation of y on x_1 and x_2.

Obtain also the multiple correlation coefficient of y on x_1 and x_2, and test for its significance, assuming the sample size to be 80. (25)

Please turn over

3. Using the intercorrelation matrix below, compute the factor loadings of the first two factors by the 'centroid' method. Each of the communalities should be estimated by the maximum value of the correlation in the corresponding column of the matrix.

	X_1	X_2	X_3	X_4	X_5	X_6
X_1	-	.33	.00	.00	.27	.00
X_2	.33	-	.32	.33	.21	.00
X_3	.00	.32	-	.72	.27	.21
X_4	.00	.33	.72	-	.00	.00
X_5	.27	.21	.27	.00	-	.72
X_6	.00	.00	.21	.00	.72	-

(20)

INDIAN STATISTICAL INSTITUTE

Statistician's Diploma Examination - November 1971

Paper VIII : Subjects of Specialisation - (I) [Grs. A to E]

Time : 4 hours Full marks : 100

(a) Candidates will be required to answer questions from that group only for which they have registered their options.

(b) Figures in the margin indicate full marks.

GROUP A : ECONOMIC STATISTICS - Special Paper I
(No candidate available)

Paper VIII - GROUP B : TECHNO-ECONOMICAL STATISTICS - Paper I (B.C)

(Attempt any five questions from this group)

1. (a) What is the theoretical basis for control charts for measurement data on quality characteristics? Derive expressions for the 3σ control limits on the chart for averages and ranges when no standards are given.

(b) What is meant by process capability? How does its evaluation and its comparison with design tolerance (specification limits) help in economic control of the process? (12+8)= 20

2. (a) It is desired to devise an acceptance sampling plan by variables for sigma-unknown and one sided specification limit L (lower limit) case . The acceptance criterion is as follows.

Draw a sample of size n from the lot and if $x_1, x_2, \cdots x_n$ denote the measurements on the n items then :

Accept the lot if $\bar{x} - k s \geq L$ and

Reject the lot if $\bar{x} - k s < L$

where $\bar{x} = \sum_{i=1}^{n} x_i/n$ and $s = \sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2/(n-1)}$.

Derive an expression for n and k when the following information are given :

i) Acceptable quality level $(AQL) = p_1$; ii) Producer's risk = α

iii) Lot tolerance percent defective $(LTPD) = p_2$; iv) Consumer's risk = β

(b) Write down the O.C. function for the plan in (a) and present a detailed computational procedure to sketch such an O.C. curve.

(15+5)= 20

Please turn over

2. (a) Dodge's CSP-1 scheme for continuous sampling plan for inspection by attributes calls for 100 percent inspection at the start of the plan. As soon as 'i' consecutive units of product are found to be free of defect, 100 percent inspection is discontinued and only a fraction (f) of the units are inspected. These individual sample units are to be selected one at a time at random from the flow of product. If a sample unit is found defective, reversion is immediately made to 100 percent inspection and this cycle is repeated. Any defective unit found during inspection is either rectified or replaced by a good unit.

Obtain with reference to the above scheme, expressions for the following :

i) Average number of pieces inspected in a 100 percent sampling sequence following the finding of a defective.

ii) Average number of pieces passed under the sampling procedure before a defective is found.

iii) The average fraction of total produced units inspected in the long run.

iv) The average fraction of units passed under the sampling procedure.

(b) When is a continuous sampling plan recommended ? (10+10)= 20

3. (a) Write a short note on the salient features of the MIL-STD-105D sampling procedures for inspection by attributes.

(b) What is curtailed inspection? Obtain the expression for the AOQ curve, for a double sampling plan (attribute inspection) with curtailed inspection. (14+6)= 20

5. (a) Explain the basic concepts of fractional factorial designs with factors at two levels by taking the case of 2^3 design.

(b) Write down the treatment combinations for the principal block for 1/2 replicate of a 2^5 factorial design using the highest order interaction as the defining contrast.

(c) Outline the method of steepest ascent in response surface studies. (5+5+10)= 20

6. (a) Discuss the basic concepts in Life testing and Reliability with reference to :-

i) failure rate analysis,

ii) measurement and prediction of reliability,

iii) methods adopted in life-testing for lot-by-lot acceptance of a product.

(b) Write a short note on the lot-plot method. (12+8)= 20

7. Write brief notes on the following :

(a) Total quality control approach in industry

(b) Mark sampling versus continuous time study techniques. (10+10)= 20

PAPER VIII - GROUP A : STATISTICAL METHODS

- GROUP B : DESIGN AND ANALYSIS OF EXPERIMENTS

(No candidate available)

Please turn over

PAPER VIII - [illegible] STATISTICS

Special Paper I - Theoretical Aspects.

(Attempt any five questions from this group)

1. From a population of size N, a sample is drawn by simple random sampling with replacement (s.r.s.w.r.) until a specified number n of distinct units is obtained. Let m be the sample size thus obtained.

(a) Find E(m).

(b) Let $\bar{y}$ denote the sample mean of all the m units and $\bar{y}^*$ denote the mean over the n distinct units obtained.

i) Are $\bar{y}$ and $\bar{y}^*$ unbiased for the population mean ?

ii) Find $V(\bar{y})$ and $V(\bar{y}^*)$ and hence compare the relative efficiencies of $\bar{y}$ and $\bar{y}^*$.

2. For sampling with probabilities proportional to size and with replacement (p.p.s.w.r.).

(a) prove that $\hat{Y} = \frac{1}{n} \sum_{i=1}^{n} \frac{y_i}{p_i}$

is unbiased for Y, the population total, where y_i is the i^{th} value of the i^{th} selected unit and p_i is the probability (preassigned and proportional to size) attached to the i^{th} selected unit.

(b) Derive the variance of $\hat{Y}$.

(c) Consider the case n = 2 and let the sampling be with probability proportional to size and without replacement (p.p.s.w.o.r.) With the same notations as above, let

$$t_1 = \frac{y_1}{p_1}$$

and $$t_2 = y_1 + y_2 \frac{(1-p_1)}{p_2}.$$

Show that t_1 and t_2 are uncorrelated. Compare the relative efficiencies of $\bar{t}$ given by $\bar{t} = \frac{t_1+t_2}{2}$ and $\hat{Y}$ given by (a) above, ignoring the cost considerations.

Consider a population consisting of k strata and the following four sampling methods:

i) unstratified s.r.s.w.r.

ii) stratified s.r.s.w.r. with equal allocation

iii) stratified s.r.s.w.r. with proportional allocation

iv) stratified s.r.s.w.r. with optimum allocation.

(a) Give the customary unbiased estimators of the population total in each case.

(b) Compare the four methods with each other regarding their efficiency.

Please turn over

Consider a two-stage sampling procedure in which the population consists of N primary units each of which contains M secondary units. Let a sample of n primaries be selected and from each selected primary unit let m secondary units be selected. Assume that sampling at both stages is by s.r.s.w.o.r.

(a) Give an unbiased estimator $\hat{\bar{Y}}$ of the population mean $\bar{Y}$.

(b) Derive the variance of $\hat{\bar{Y}}$.

(c) Let a population consist of two strata of which the first stratum has the structure given above and the second stratum has N' units. Let sampling from first stratum be as given for (a) and (b) above and sampling from second stratum be unistage s.r.s.w.o.r. of size n'

i) Give an unbiased estimator of mean $\bar{Y}$ in this case

ii) For given values of N, M, n and n' and for the cost function

$$C = c_0 + n c_1 + n' c_2 + n m c_3$$

where all the constants C, c_0, c_1 c_2, and c_3 are known; derive the optimum value of m.

5. Consider a population consisting of two strata. Let a s.r.s.w.o.r. of size n_1 be drawn from the first stratum and a p.p.s.w.r. sample of size n_2 be drawn from the second stratum.

(a) Give a suitable unbiased estimator $\hat{Y}$ of the population total Y and prove its unbiasedness. [Standard procedure should be adopted].

(b) Derive $V(\hat{Y})$ and a non-negative unbiased estimator of $V(\hat{Y})$.

6. (a) A population has N clusters each of size M units. A s.r.s.w.o.r. of n clusters is drawn.

i) Give an unbiased estimator $\hat{Y}$ of the population total.

ii) Derive the variance $V(\hat{Y})$ of $\hat{Y}$ in terms of the intra-cluster correlation coefficient.

(b) For the Madhow circular systematic sampling

i) Give an unbiased estimator $\hat{\bar{Y}}$ of the population mean.

ii) Derive the variance of $\hat{\bar{Y}}$.

iii) Give an unbiased estimator of the variance obtained in (ii) above.

7. Write short notes on any four of the following :

(a) Multistage and multiphase sampling.

(b) Combined ratio estimator versus separate ratio estimator in stratified sampling.

(c) Sampling on successive occasions.

(d) Self-weighting designs.

(e) Balanced systematic sampling.

(f) Integration of surveys.

PAPER VIII - [illegible] : [illegible]

[illegible] : [illegible]

[illegible] : [illegible]

(No candidate available)

INDIAN STATISTICAL INSTITUTE

Statistical Service Examination - October, 1[illegible].

Paper I : Subject Specification - II (Theoretical).

Time : 4 hours. Full Marks : 100

(a) Candidates are required to answer questions from that group only for which they have registered their options.

(b) Figures in the margin indicate full marks.

SP.IX: GROUP A - ECONOMIC STATISTICS (Special Paper).

(No Candidate Available).

SD Paper IX - GROUP B : TECHNO-COMMERCIAL STATISTICS (Special Paper - II).

Section I : OPERATIONS RESEARCH (70 marks)

Section I : (Alternative): ELEMENTS OF BOOK-KEEPING AND ACCOUNTANCY. (70 marks)

Section II: STATISTICAL METHODS IN BUSINESS (30 marks)

Section I : OPERATIONS RESEARCH.

(i) Use a separate answer book for this section.

(ii) Attempt question 1 and any other three questions from this section.

Q.1 a) What is meant by a basic feasible solution to a Linear Programming problem ? In the case of a maximization problem if an optimal basic solution is non-degenerate and $Z_j - C_j \geq 0$ for every vector not in the basis, show that the optimal solution is unique.

b) Consider the set of equations.

$$2x_1 - 3x_2 + 4x_3 + 6x_4 = 28.$$
$$x_1 + 2x_2 + 3x_3 - 3x_4 + 5x_5 = 12.$$

Note that $x_1 = 2$, $x_2 = 1$, $x_3 = 3$, $x_4 = 2$, $x_5 = 1$ is a feasible solution Reduce this feasible solution to two different basic feasible solutions.

c) Can a vector which is inserted at one iteration in the simplex method be removed immediately at the next iteration ? When can it occur and when is it impossible ?

d) When do you use artificial variables in solving a Linear Programming Problem ? What are the disadvantages of using Charnes' M Method ?

e) How do you detect inconsistency and redundancy in the constraints of the problem ?

(5 x 5) = 25.

Q.2 Suppose that an order of q units placed on January 1st is received on February 1st. The stock level on January 1st, after receipt of delivery is s units. The stock level on February 1st, after receipt of delivery of order placed on January 1st is $s + q - x$ (shortages are not lost). x is the monthly demand. C_1 is the cost of ending inventory per unit, and C_2 is the cost per unit short. The one month demand density function is f (x) and is independent of demand in other months.

a) If y is the demand in February, show that the total expected cost (TEC) in February is expression given in details in the next page.

Please turn over

(Cont)

SD (M1)-EE (c)

Q.2 a) (Contd.)

$$TEC = \int_0^S \left[c_1 \int_{y=0}^{S-x} (S-x-y) f(y)\,dy + c_2 \int_{S-x}^{\infty} (y+x-S) f(y)\,dy \right] f(x)\,dx$$

$$+ c_2 \int_{x=S}^{\infty} \int_{y=0}^{\infty} (y+x-S) f(y)\,dy \; f(x)\,dx \quad \text{where } S = s + c .$$

b) Show that the optimal value of S_y is given by

$$\int_0^{S_0} F(S_0 - x) f(x)\,dx = \frac{c_2}{c_1 + c_2} \quad \text{where } F(x) = \int_0^x f(t)\,dt$$

c) Show that if $f(x) = b e^{-bx}$, the optimal value of S is a solution of

$$e^{-bS} (1 + bS) = \frac{c_1}{c_1 + c_2}$$ (5+5+5) = 15

Q.3 An air freight company picks up and delivers freight where customers require. Company has two types of air craft X and Y with equal loading capacities but different operating costs. These are shown below :

Type of aircraft	Cost per mile (money units).	
	Empty	Loaded.
X	1.00	2.00
Y	1.50	3.00

The present location of the four air crafts which the company is having as follows:-

Location	Type of air craft.
I	X
K	Y
L	Y
M	X

Four customers of the company located at A, B, C and D want to transport nearly the same size of load to their final destinations. The final destinations are at a distance of 600, 300, 1000 and 500 miles from the loading point A, B, C and D respectively.

Distances in miles between location of the aircraft and the Loading points are as follows:

Present Location of aircraft.	Loading Points.			
	A	B	C	D.
I	300	200	400	100
K	300	100	500	200
L	400	100	100	500
M	200	200	400	200

Determine the allocation which minimises the total cost of transportation.

(15)

Please turn over

Q.4 a) An equipment costs "a" and the maintenance costs are "C1 C2 C3 etc ($C_{n+1} > C_n$) during the first year, the second year and the third year etc. Find the optimum replacement policy which minimizes the total of all future discounted costs. You can assume that the above expenditures are incurred at the beginning of each year and that the resale value is zero.

b) The following preventive replacement policy in being followed in case of certain equipment. "Replace the equipment if it fails or reaches an age of T units of time, whichever is earlier". If C_F and C_P ($C_F > C_P$) are the costs of failure replacement and preventive replacement, prove that the average cost per unit time is given by

$$(C_F - C_P)\,\frac{F(T) + \frac{C_P}{C_F - C_P}}{I(T)}$$

Where F (t) is cumulative distribution function of life and

$$I(T) = \int_0^T 1 - F(t)\,dt.$$

Derive the conditions for optimum replacement policy.

(8 + 7 = 15).

Q.5 An operator has been allotted N identical machines for servicing. Whenever a machine stops, the operator has to service the machine before it can be restarted. The machine running time between two successive stoppages and the service time both follow negative exponential distributions with averages $1/\lambda$ and $1/\mu$ respectively. Write down the balance equations for the steady state probabilities.

Hence prove that $P_n = \frac{N!}{(N-n+1)!}\left(\frac{\lambda}{\mu}\right) P_{n-1}$ for n=1 to N

and $P_n = \frac{N!}{(N-n)!}\left(\frac{\lambda}{\mu}\right)^n P_0$

Where P_n is the probability that n machines are not running

i) What is the percentage time the operator will be busy ?

ii) Show that percentage interference delay is given by

$$ICC = \frac{1 + \left(\frac{\lambda}{\mu}\right)(N)}{N\left(\frac{\lambda}{\mu}\right)}, \quad \text{Percent operator utilization}$$

(15)

Q.6 A sample of 100 arrivals of customers at a supermarket are recording to the following distribution.

Time between arrivals, (minutes)	Frequency	Time between arrivals (minutes)	Frequency.
0.5	2	3.5	1
1.0	6	3.0	10
1.5	10	4.0	7
2.0	25	4.5	4
2.5	20	5.0	2
			100

A study of the time required to service the customers by adding up the bill, receiving payment, making change, etc yielded the following distribution.

Service time (minutes)	Frequency	Service time (minutes)	Frequency.
0.5	12	2.5	7
1.0	21	3.0	5
1.5	36		
2.0	19	Total :	100

a) Convert the distributions to Cumulative Probability distributions.

b) Using a simulated sample of 20, estimate the average percent customer waiting time and the average percent idle time of the server. (15)

Please turn over

S.D. II Group B - Section II : STATISTICAL METHODS IN BUSINESS

(i) Use separate answer book for this section.

(ii) Attempt any two questions from this section. (50 Marks)

Time

Q.1 (a) Define and discuss the difference between Consumer Product and Industrial Product, from the point of view of market research.

(b) Describe the Wildebolf advertising model which is used to find out the money to be invested in advertising. Discuss the inefficiency of the primitive percentage model of advertising.

(3+7) = 17.

OR

Let P( ) be the probability of that demand is exactly for d spares. The merchant has to spend c money units to purchase a spare part and a spare part which is not sold involves a loss of U money units. If x denotes the number of spare parts available and W is the loss function when x spares are purchased and d spare options number of spare parts x to be purchased which would minimise the total expected loss is determined by

$$P(d \leq x_0) \geq \frac{U-C}{U} > P(d \leq x_0 - 1) \qquad (15)$$

Q.2 (a) What components of demand are involved in forecast ?

(b) Why do moving averages discount random effects ?

(c) Compare the results of simple moving average forecasts and exponential moving average forecasts

(1+4+7) = 15

Q.3 Write short notes on any three :

(a) Productivity measurement.

(b) Sampling techniques in accountancy and auditing.

(c) Job evaluation and merit rating.

(d) Bonus and incentive schemes.

(5+5+5) = 15.

IX GROUP 5 : Section I (Alternative): SUBJECTS OF BOOK-KEEPING AND ACCOUNTANCY

(a) Use separate answer book for this section
(b) Attempt question 1 and any other three from this group.

1. A & B are partners of the firm A, B & Co. They share profits in the ratio of 2:1. The Trial Balance of A, B & Co as on 31/3/1971 is as below:-

Trial Balance as on 31/3/1971

	₹.			₹.
Building	1,50,000	Capital Account		
Plant & Machinery	50,000	A: 50,000		
Opening Stock:		B: 30,000		80,000
Raw Materials.	30,000			
Finished Goods.	40,000	Sundry Creditors.		25,000
Sundry Debtors.	50,000	Sales.		2,84,500
Factory Expenses.	7,500	Discount.		2,500
Factory Rent & Taxes.	5,500	Commission.		10,000
Wages	20,000	Bank Overdraft.		50,000
Salary of Works Manager.	12,000			
Purchase of Raw Materials	70,000			
Office Expenses.	7,500			
Office Rent.	4,000			
Discount.	3,000			
Insurance.	1,500			
Bad Debt.	2,000			
Cash at Bank.	8,000			
	₹. 4,52,000			₹. 4,52,000

The following additional information is to be taken into consideration.

a) Closing Stock - Raw Materials - ₹.30,000
Finished Goods - ₹.50,000

b) Outstanding Liabilities to be provided for -
Wages - ₹.5,000
Office Salaries - ₹.6000
Office Rent - ₹.4000.

c) Depreciation to be provided for -
Building @ 2½%
Plant & Machinery @ 5%.

d) Interest @ 6% p.a. is payable on the Bank Overdraft and has to be provided for

[NOTE: The overdraft had been existing at the commencement of the year and remained unchanged all through the year.]

You are required to prepare the Trading and Profit & Loss Account for the year ended 31/3/1971 and the Balance Sheet as on that date.

(25)

2. Distinguish with appropriate examples and explanations between:-

i) Floating and Fixed Assets.
ii) Provision and Reserve in Company Accounts.
iii) Capital Reserve and Reserve Capital.

(5+5+5) = 15.

3. What do you mean by the term 'Average Due Date'? How is the starting point generally determined for ascertaining the 'Average Due Date'?

Below are the due dates of payments for certain purchases:-

Purchase Date	5th March, 1971 - ₹.3000	Payment due on	8th April, 1971.
" "	15th April, 1971 - ₹.2000	" -	13th May, 1971
" "	10th May, 1971 - ₹.2750	" -	13th June, 1971
" "	5th June, 1971 - ₹.4000	" -	8th July, 1971.

Work out the 'Average Due Date' on which a single bill is to be drawn for payment of the total amount due.

(4+2+9) = 15.

Please turn over

Q.4 On December 31, 1970, the bank column of the cash book of 'A' shows a debit balance of ₹.922. On examination of the cash book and the bank statement it is found that -

a) Cheques amounting to ₹.1260 issued before December 31 and entered in the cash book, were not presented for payment till after that date.

b) Cheques amounting to ₹.509/- entered in the cash book as sent to the bank on December 31, were entered into the bank statement after that date.

c) A Cheque from a debtor for ₹.146 had been dishonoured prior to December 31, but no record appeared in the Cash Book.

d) A dividend warrant for ₹.76/- was paid direct to the bank and nothing appeared in the cash book.

e) Bank interest and charges amounting to ₹.84 were not entered in the cash book, but appeared in the Bank statement.

f) There was no entry in the cash book of a payment of ₹.20/- paid by bankers order in November, 1970.

g) Bank's postal charges of ₹.2/- for a cheque book received by 'A' was entered in the cash book twice.

h) A cheque for ₹.54/- drawn by 'B' had been charged to 'A's account in error in 1970.

Make appropriate adjustments in the cash book to bring down the correct balance.

Hence prepare a statement reconciling the corrected cash book balance with the balance as per the bank statement.

(9+6) = 15.

Q.5 What is a Negotiable Instrument ?

'A' Purchased goods of the value of ₹.72000/- from 'B' on 1.9.70. He sold them to 'C' for ₹.90,000/- on the same day and agreed to accept a bill of exchange of equal value from him payable on 1.3.1971 plus interest at 10% p.a

'A' negotiated this bill to his bank on 30/9/70 at a discount of 6% p.a.

Show by journal entries, how all these transactions should be recorded in the books of 'A'.

(5+10) = 15.

Q.6 Write short notes on:-

I) Contingent Liability.

ii) Special Crossing of a Cheque.

iii) Suspense Account.

iv) Goodwill.

(4+5+3+3)
= 15.

SD.IX SPECIAL PAPER II

Group C: Biometric Methods in Genetics and Bio-assays.

Group D: Design and Analysis of Experiments.

(No Candidates available)

SCHEME : [illegible] - Special Paper II
(Organisational Aspects)

(a) (Answer any five questions from this group)
(b) Figures in the margin indicate full marks.

1. a) Centralised versus decentralised recruitment of field staff.

b) Technical versus non-technical administrative personnel for managing administrative matters of organisations dealing with sample surveys.

c) Fixed versus running Travelling allowances to investigators.

d) Centralised versus decentralised processing of data.

e) Printing versus cyclostyling of reports.

Compare briefly the merits and demerits of any four of the above. (5x4) = 20

2. a) A consumer expenditure survey is proposed to be undertaken in an industrial town, for working out the weighting diagram for the construction of consumer price index numbers of the industrial workers of the town. Explain in the context of this proposed survey, the role of the following techniques:-

i) Tenement versus Pay roll sampling.

ii) Collection of data from different sample households as against the same set of sample households in different months of the year.

iii) Interpenetrating sub-samples. (4+4+4) = 12

b) State the circumstances under which you would prefer the use of:-

i) Hand tabulation to Mechanical tabulation;

ii) Questionnaire to Schedule. (4+4) = 8

3. A sample survey for estimating the size and distribution of the unemployed persons in the country is proposed to be undertaken. Indicate briefly:

a) a)The sampling design for the survey.

b) Lay out of the main blocks of the schedule.

c) Tabulation programme for the proposed survey.

d) Synopsis of the report of the survey. (4+4+4+8) = 20

4. Write short notes on any four of the following:-

i) Statutory versus Voluntary collection of data.

ii) Role of concepts and definitions in a sample survey.

iii) Error due to differences between investigators.

iv) Time-programme of a survey.

v) Doubling of mechanical equipments versus decision to work on two shifts.

vi) Essential and desirable qualifications for investigators. (4x5) = 20

Please turn over

5. a) Describe in detail the various components of the budget of a large scale sample survey.

b) Draw up an organisation chart for the field staff which may be employed in a large-scale sample survey.

c) What are the essential points which should be borne in mind while organising the machinery for the publicity work relating to a sample survey.

(10+5+5) = 20

6. a) Discuss carefully the salient points that should be taken into account in drafting the instructions for filling up the schedule of a sample survey. Illustrate your answer with the help of examples.

b) How do you propose to build up controls for ensuring correctness of tabulation undertaken mechanically ?

c) What is the importance of a 'flow chart' in a sample survey ?

(9+6+5) = 20

7. a) How should the work in a machine room be organised so that maximum utilisation of a Tabulator could be achieved ?

b) Discuss the utility of maintaining log books in a mechanical tabulation unit ? How are these used ?

c) What are the points to be kept in mind for checking up the final tables of a statistical survey ?

d) What is the role of post-machine tabulation scrutiny ?

(5+5+5+5) 20

SD.IX SPECIAL PAPER II

GROUP : F Techniques of Computation. (Practical)

GROUP : G Statistical Inference.

GROUP : H Probability Theory.

(No Candidates available)

INDIAN STATISTICAL INSTITUTE

Statistician's Diploma Examination - November 1971

Paper X - Subjects of Specialisation - (Special Paper III)

(Practical)

Time : 5 hours Full marks : 100

i) Candidates are required to answer questions from that group only for which they have registered their options.

ii) Figures in the margin indicate full marks.

iii) Use of calculating machine is permitted.

Paper X - GROUP A : ECONOMIC STATISTICS
Special Paper III - Practical
(No candidate available)

...

Paper X - GROUP B : TECHNO-COMMERCIAL STATISTICS (Special Paper III - Practical)

Section I : Statistical Quality Control (50)

Section II : Operations Research/Elements of Book-keeping and Accountancy (30)

Section III : Statistical Methods in Business (20)

Gr. B : Section I : Statistical Quality Control (50 marks)

(Attempt any two questions from this Section)

1. The following are the means and ranges of 20 samples. The samples are sub-groups of size 5 each, drawn at regular intervals from the existing production process. The data pertain to the diameter (in inches) of a manufactured product. The specifications are $0.445'' \pm 0.015''$.

Sample No.	$\bar{X}$	R	Sample No.	$\bar{X}$	R
1	0.4402	0.015	11	0.4360	0.011
2	0.4390	0.018	12	0.4402	0.007
3	0.4448	0.018	13	0.4332	0.008
4	0.4432	0.006	14	0.4358	0.017
5	0.4428	0.008	15	0.4314	0.010
6	0.4392	0.010	16	0.4362	0.015
7	0.4358	0.019	17	0.4390	0.019
8	0.4440	0.011	18	0.4350	0.008
9	0.4366	0.010	19	0.4378	0.011
10	0.4368	0.011	20	0.4384	0.009

(a) Establish standards and determine 3 sigma control limits for $\bar{X}$; R charts.

Calculate the revised control limits when the subgroup size is changed from 5 to 4.

Please turn over

(b) Compare the process capability with the specifications and comment.

(c) If the process average shifts to 0.4315", without changing the variability, what is the probability that you would detect such a shift on your control chart with limits calculated in (a) (Sample size = 5) :

i) On the basis of the first sample drawn following the shift?

ii) At most, in three samples following the shift?

Find the corrected number of samples that have to be taken in order to detect this shift on the control chart.

(15 + 3 + 7) = 25.

2 (a) Draw the OC and average amount of inspection curves for the following double sampling acceptance/rectification plan, the lot size being 1,200.

$n_1 = 40$ $c_1 = 1$

$n_2 = 80$ $c_2 = 3$

(b) Consider the sampling plan by variables with one sided specification limit given by the parameters, n = 16, k = 2.116. σ (Sigma) is known to be 4.5 units.

i) For the above plan, for what lot quality (given by the fraction defective in the lot) would the probability of acceptance be equal to 0.95?

ii) What would be the probability of acceptance of a 10 per cent defective lot?

(16 + 4 + 5) = 25

3 In order to compare the hardness of alloys, three furnaces (F) and four levels of moulds (M) were tried in a S-lit-plot Design with 3 replications. The hardness readings as well as the layout of the design after randomisation are shown below. Analyse the data and offer your comments.

Replication I			Replication II			Replication III		
	M_3	M_2		M_2	M_0		M_1	M_2
	156	118		104	70		108	126
F_3	M_1	M_0	F_3	M_1	M_3	F_2	M_3	M_0
	140	105		89	117		149	70
	M_0	M_1		M_3	M_0		M_3	M_1
	111	130		112	74		144	121
F_1	M_3	M_2	F_1	M_1	M_2	F_3	M_2	M_0
	174	157		89	81		121	96
	M_0	M_1		M_1	M_0		M_0	M_3
	117	114		103	84		81	100
F_2	M_2	M_3	F_2	M_2	M_3	F_1	M_1	M_2
	161	141		132	133		91	97

Gr. B : Section II : Operations Research : (30 marks)

(Attempt any two questions from this Section)

4. A company has four territories open, and four salesmen available for assignment. The territories are not equally rich in their sales potential; it is estimated that a typical salesman operating in each territory, will bring in the following annual sales :

Territory	I	:	Rs. 4,80,000
"	II	:	" 4,00,000
"	III	:	" 3,20,000
"	IV	:	" 2,40,000

The four salesmen are also known to differ in their ability. It is estimated that, working under the same conditions, the yearly sales for the four salesmen A, B, C and D will be in the proportions of 7:5:5:4 respectively. Find the optimal assignment of the salesmen to the territories so as to maximise the expected total sale.

(15)

5. The following table gives the sales, rate of production, inventory and setup costs for 5 items (bolts) manufactured by a Company using the same equipment. Obtain the optimum number of cycles in a year assuming that there are 300 production days available in a year.

Product	Sales per year (quantity)	Production per day	Inventory cost per year per unit (Rs)	Setup Cost (Rs)
1	5,00,000	10,000	0.020	8.00
2	13,000	1,000	0.005	10.00
3	2,00,000	10,000	0.008	17.00
4	1,80,000	2,000	0.022	5.00
5	92,000	1,000	0.015	6.80

(15)

6. Solve the following Linear Programming problem by the Simplex method :

$$\text{Max. } Z = 3x_1 + 4x_2 + x_3 + 7x_4$$

Subject to the following constraints:

$$8x_1 + 3x_2 + 4x_3 + x_4 \leq 7$$

$$2x_1 + 6x_2 + x_3 + 5x_4 \leq 3$$

$$x_1 + 4x_2 + 5x_3 + 2x_4 \leq 8$$

$$\text{all } x_j \geq 0$$

(15)

Gr. B : Section III : Statistical methods in Business (20 marks)

(Attempt question 7 and another from this Section)

7. The seasonal indices of the sales of garments of a particular type in a certain shop are given below :

Quarter			Seasonal Index
I	(Jan - March)	...	97
II	(Apr - June)	...	85
III	(July - Sept)	...	83
IV	(Oct. - Dec.)	...	135

If the total sales in the first quarter of a year be worth Rs.15,000, determine how much worth of garments of this type should be kept in stock by the shop-owner to meet the likely demand for each of the other three quarters of the year.

(6)

8. A manufacturer sells an article at a fixed price of Rs. 1/-. If the weight of the article is less than 8 ounces, he is unable to sell it and it represents a total loss. The weight of the article is normally distributed with mean 'w' ounces and standard deviation one ounce only. The cost per article is Rs. C, where

$$C = 0.05\, w + 0.30$$

Determine the mean weight 'w' which will maximise the manufacturers expected profit.

(14)

9. A manufacturer of dolls, wishes to plan a special consignment for the Christmas Season. On each doll sold he would make a net profit of Rs. 3, whilst the dolls left unsold in the season would be disposed of in the January sale at a net loss of Rs. 2 per doll. On the other hand, if there is a loss of customer in the Season due to failure in the matter of supply, the manufacturer expects a loss to the concern because of a goodwill factor (which may be considered to be Rs. 4 per doll that cannot be supplied). If the demand for dolls in the Christmas Season is denoted by x, then $\log_{10} x$ can be assumed to follow normal distribution with mean 4 and standard deviation 0.25.

Find the optimum number of dolls that the manufacturer should stock in order to maximise the expected profit.

(14)

SD X : PAPER III (Practical)

Gr. C : Biometric Methods

Gr. D : Design & Analysis of Experiments

(No candidates available)

Please turn over

S.D. X : Crops X : Small Farms

Paper III : Practical.

(All questions carry equal marks)
(Attempt any four questions including at least one question from each section).

Section I

1. Given below are three populations U_1, U_2 and U_3, each consisting of four units. For each unit of each of the populations, the values y_i and x_i of variables y and x respectively, are also given.

Sl.no. of unit i.	U_1		U_2		U_3	
	y_1	x_1	y_1	x_1	y_1	x_1
1	-5	30	-10	10	-400	15
2	-8	30	- 8	20	-300	25
3	7	20	14	30	299	40
4	4	20	8	40	399	20

(a) From each of the above three populations draw four samples, each of size 2, one each for the following four methods of sampling:

(i) simple random sampling with replacement (s.r.s.w.r.)

(ii) Simple random sampling without replacement (s.r.s.w.o.r)

(iii) probability proportional to x-measure, with replacement (p.p.x.w.r)

and (iv) probability proportional to x-measure, without replacement (p.p.x.w.o.

(b) Using suitable estimation procedures, calculate unbiased estimates of the population total of the y variable for each of the populations and from each of the samples drawn (four estimates are thus obtained for each of the three totals).

(c) Examine the closeness of the various estimates to the known total of the y variable in each case.

2. Obtain the variances of the four estimators considered by you in question no.1 above, for the three populations and discuss their relative efficiencies.

3. To estimate in a particular year, the total area $Y = \sum_{i}^{N} y_i$ under wheat, of a sub-division, n villages were drawn at random out of total of N villages, with probabilities proportional to x_i, their total areas under cultivation in a previous year, and without replacement. However, while estimating the population total and estimating its variance, the fact that the sampling was with varying probabilities was ignored, and instead the n villages were assumed to be drawn by simple random sampling without replacement. The estimator of the population total Y was taken as the sample mean of the y's multiplied by the population size N.

Please turn over.

(XXX)

Calculate the bias if any, in the estimator and also the error in the estimate of variance of the estimator, with the help of the following data for n = 2.

Derive the necessary formula to estimate the error in variance from the sample.

Sl.number of village.	area under wheat (100 acres) in 1936	total cultivated area (100 acres) in 1931.
(1)	(y_i)	(x_i)
1	3	4
2	2	6
3	3	12
4	4	8

4. It is proposed to draw a sample of n clusters, of M units each, from a population of N clusters and a sub-sample of m units from each sampled cluster using s.r.s.w.r. at both the stages, for estimating the mean per unit of a specified characteristic Y. Assuming the cost function to be of the form $C = C_0 + C_1 n + C_2 nm$, determine the optimum values of n and m in terms of C, C_0, C_1 and C_2.

Find these optimum values of n and m when C = 1000, C_0 = 300, C_1 = 9 and C_2 = 1 (in rupees) using the following analysis of variance table.

ANOVA for the character Y

Source of variation (1)	degrees of freedom. (2)	sum of squares. (3)	mean squares. (4) = (3)/(2)
Between clusters	89	$20 \sum_{i=1}^{90} (\bar{y}_i - \bar{\bar{y}})^2$	190.9
Within clusters.	1710	$\sum_{i=1}^{90} \sum_{j=1}^{20} (y_{ij} - \bar{y}_i)^2$	49.5
Total	1799	$\sum_{i=1}^{90} \sum_{j=1}^{20} (y_{ij} - \bar{\bar{y}})^2$	58.0

NOTE: N, the total number of clusters = 90.

M, the number of units per each cluster = 20.

Section - II

Draw a small sched le for sample checks on demographic data collected in April 1971 census in any given State of India, after suggesting a suitable design for the same. Also draw up a scrutiny programme for checking the filled-in schedules before processing them.

Please turn over

6.(i) A sample survey is to be conducted for assessing poll-prospects of different parties in the coming assembly elections in a big city of India. The survey design is a stratified three-stage design with municipality wards grouped into suitable number of strata, with parts in each selected municipality ward in a stratum as second-stage units and with households in each selected part as third-stage units. The municipal wards in a stratum are sampled with probability proportional to their 1971 populations and with replacement. A suitable scheme of sampling is adopted for subsequent stages.

Draw up a small schedule for this survey. Such schedules, when filled-in, are to be processed at the processing centre and tabulated manually.

(i) Assuming that there are different sections, each with one or more supervisors and several workers under each supervisor, draw up the required forms and proformafor :

(a) recording receipt of schedules from the field,

(b) allotment of work to individual workers and in different sections,

(c) progress of work in the units

(d) working sheets for calculationof suitable multipliers for estimating required totals.

(ii)Will there be a suitable non-negative (i.e. always positive) estimator of variance of the estimator, for different totals ? Give your considered remarks ?

GROUP F : TECHNIQUES OF COMPUTATION
Specialisation Paper III (Practical)

GROUP G : STATISTICAL INFERENCE
Specialisation Paper III (Practical)

GROUP H : PROBABILITY THEORY
Specialisation Paper III (Practical)

(No candidates available)