INDIAN STATISTICAL INSTITUTE

Statistician's Diploma Examination - November 1975

Paper I (Theoretical): Official Statistics and Descriptive Statistics

Time: 4 hours Full marks: 100

(a) Figures in the margin indicate full marks.

(b) Use of calculating machines is not permitted.

GROUP A: Official Statistics (50 marks)

(Attempt any three questions from this group.)

1. You are required to prepare a note on the availability of foodgrains per capita in India for a few years. Which statistical sources you will use to obtain the necessary information? Observe that you have to use different sources and indicate the information you propose to pull from these sources. Which would you consider to be more reliable? (16)

2. You are seeking an answer to the question, is investment in India in the recent past adequate for the growth of real national income? Describe the contents of the statistical sources you will use for this purpose and indicate the major gaps in data. (16)

3. Describe the statistical work of any two of the following:

i) an important Central Government statistical office,

ii) a specialised UN agency,

iii) an important State Government statistical office. (8+8)=16

4. (a) Describe a sample survey indicating the design, questionnaire, field organisation, tabulation procedure and the important results obtained.

(b) Write a critical note on estimates based on sample surveys, complete enumeration and official records obtained as a by-product of administration. (10+6)=16

5. (a) Write short notes on available Indian statistics relating to any two of the following:

(i) health services,

(ii) money and banking,

(iii) mining,

(iv) motor vehicles and roads.

(b) Write a brief note on the available data on production and imports of crude petroleum in India. Which countries are the leading producers of crude petroleum and what do you understand by OPEC? Give some idea of the petroleum price hike in the recent past. (8+8)=16

GROUP B: Descriptive Statistics (50 marks)

(Attempt any three questions from this group.)

6. (a) Discuss range as a measure of variability and explain why it is widely used in statistical quality control.

(b) Explain the advantages of a measure of relative dispersion like the coefficient of variation over measures of dispersion like the s.d.

Please turn over

6. (c) (contd.) Consider a set of observations $x_1, x_2, \ldots, x_n$, all of which are positive. Show that if all the deviations $x_i - \bar{x}$ are small in comparison with $\bar{x}$, the geometric mean x_g is approximately given by

$$x_g = \bar{x}\left(1 - \frac{1}{2}\frac{s^2}{\bar{x}^2}\right)$$

Here $\bar{x}$ denotes the arithmetic mean and s the standard deviation. (5+5+6)=16

7. (a) Define Hermite polynomials and prove their orthogonality property.

(b) Give an account of the Gram-Charlier Type A series representation of a continuous type probability density function. Express the first few coefficients of the series in terms of the moments of the distribution. (7+9)=16

8. Define the correlation ratio η_{yx} of y on x.

Prove the following decomposition of the total variability of y in a set of bivariate observations where corresponding to $x = x_i$, y assumes n_i values $y_{i1}, y_{i2}, \ldots, y_{in_i}$ $(i = 1, 2, \ldots, k)$:

$$\sum\sum(y_{ij} - \bar{y})^2 = \sum\sum(y_{ij} - \bar{y}_i)^2 + \sum n_i(\bar{y}_i - Y_i)^2 + \sum n_i(Y_i - \bar{y})^2$$

Here $\bar{y}_i = \sum y_{ij}/n_i$, $\bar{y} = \sum\sum y_{ij}/\sum n_i$ and Y_i is the expected value of y when $x = x_i$ based on the least squares regression of y on x.

Discuss in the light of this relation the cases where $\eta^2_{yx} = 0$, $\eta^2_{yx} = 1$ and $\eta^2_{yx} = r^2$, where r is the correlation coefficient between x and y. What is the significance of the discrepancy $\eta^2_{yx} - r^2$? (3+8+3+2)=16

9. (a) Define the partial correlation coefficient $r_{12.34 \ldots p}$ and express it in terms of the (total) correlation coefficients among the variables $x_1, x_2, \ldots, x_p$.

(b) If all the (total) correlation coefficients among $x_1, x_2, \ldots, x_p$ are equal to ρ, what will be the values of

(i) $r_{12.34 \ldots p}$

and (ii) the multiple correlation coefficient $R_{1.23 \ldots p}$? (10+6)=16

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

(a) Uses of pictorial diagrams and statistical maps for presentation of statistical data.

(b) Cumulants - their properties and uses.

(c) The bivariate normal distribution.

(d) The four components of a time series. (8+8)=16

INDIAN STATISTICAL INSTITUTE

Statistician's Diploma Examination - November 1975

Paper II : Probability Theory and Statistical Methods (Theoretical)

Time : 3 hours Full marks : 100

Figures in the margin indicate full marks

Group A : Probability Theory (50 marks)

(Attempt any four questions from this group)

1. The frequency function (that is, density function) of the random variable X is kx on the interval $(0, 1)$ and 0 outside this interval; k is a constant.

(a) What is the value of the constant k ?

(b) Obtain the density function of the random variable X^2.

(c) Calculate the expectation (mean) of the random variable $(X^2 - X)$.

(d) What is the distribution function of X ? (2+4+2+4) = 12

2(a) X is a random variable taking 2 values only. Prove by direct algebraical calculation that $E(X^2) \geq \{E(X)\}^2$.

(b) X and Y are two random variables such that $P_r\{X = -1, Y = 0\} = P_r\{X = 0, Y = 1\} = P_r\{X = 1, Y = 0\} = \frac{1}{3}$.

Are X and Y independent ? Give reasons. What is the correlation coefficient of X and Y ?

(c) In general, what is the connexion between independence and correlation (of 2 random variables) ?

(d) X and Y are two random variables. Each of them takes only the values 0 and 1. Prove that if they have zero correlation, they are independent. (3+3+2+4) = 12

3. The frequency function of the joint distribution of X and Y is as follows :

$f(x,y) = kxy$ inside the triangle whose vertices are $(0,0)$, $(1,0)$ and $(0,1)$,
$f(x,y) = 0$ outside this triangle.

(a) Determine the value of the constant k.

(b) Determine the conditional expectation $E\{Y \mid X\}$.

(c) Are X and Y independent ? Give reasons.

(d) Calculate the probability $P_r\{(X + 2Y) < 1\}$. (3+3+3+3) = 12

(Please turn over)

4.(a) Prove Chebyshev's inequality $P\{|X-a| \geq k\sigma\} \leq \frac{1}{k^2}$.

(b) Give a specific example where $P\{|X-a| \geq k\sigma\} = \frac{1}{k^2}$.

(c) When do we say that the sequence $\{X_m\}$ of random variables converges in probability to the constant a?

(d) There is an infinite population with mean μ and standard deviation σ (which is finite). Prove that the sample mean $\frac{x_1+x_2+\ldots+x_n}{n}$ converges in probability to a constant.

(e) The frequency function of a bivariate distribution is

$f_{X,Y} = \frac{1}{k}(x^2+4xy+5y^2)$.

What are the means and variance of X and Y ? What is the value of the constant K ? What is the correlation coefficient of X and Y ?
(You need not perform any double integration) (3+2+2+2+3) = 12

5.(a) Show that in the case of samples from a normal population, the sample mean and sample variance are independent.

(b) 5 chairs are arranged in a row. They are going to be occupied by 2 boys and 3 girls. What is the probability that all the 3 girls are together (that is, the 3 girls sit on adjacent chairs, ? Give the answer in its simplest possible form without factorial. (6+6) = 12

Neatness. 2

Group B : Statistical Methods (50 marks)

(Attempt any three questions from this group)

6. Suppose $X_1, X_2, \ldots, X_n$ are independent and identically distributed (i.i.d.) random variables with density

$$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

(a) Let $\bar{X}_n = \frac{1}{n}(X_1+X_2+\ldots+X_n)$,

and $s_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X}_n)^2$, $n \geq 2$

Show that s_n^2 and $\bar{X}_n$ are independent.

(b) Verify that $\frac{(n-1)}{\sigma^2} s_n^2$ has χ^2 distribution with (n-1) degrees of freedom. $(n \geq 2)$

(c) Is the parameter μ/σ estimable ?

(8+2+6) = 16

7. Let X and Y be random variables having joint density

$$f_\theta(x,y) = e^{-\theta x} e^{-y/\theta}, \quad x,y > 0$$

and $\theta > 0$ is unknown.

(a) Is (X,Y) sufficient for θ?

(b) Is $T_1 = (Y/X)^{1/2}$ the maximum likelihood estimate of θ?

(c) Show that $T_2 = (XY)^{1/2}$ has a distribution free of θ, but (T_1, T_2) is sufficient for θ.

[2+5+(6+3) = 16]

8(a) Suppose $X_1, X_2, \ldots, X_n$ are independent and each is distributed uniformly on the interval $(0,\theta)$. Let $T_n = \max(X_1, X_2, \ldots, X_n)$. Is T_n sufficient for θ? Is T_n consistent for θ?

(b) Let X be a random variable with the density as in 7.(c). Is $e^{-\mu+\sigma^2/2}$ estimable?

[(5+5)+6 = 16]

9(a) Define a uniformly minimum variance unbiased estimator (UMVUE) of a parametric function.

(b) Let X have a Poisson distribution truncated at 0, so that

$$P_\theta\{X = x\} = \frac{\theta^x e^{-\theta}}{x!\,(1-e^{-\theta})} \quad \text{for } x = 1,2,\ldots.$$

Show that $T(x) = (-1)^{x+1}$ is the UMVUE of $e^{-\theta}$.

(c) Define admissibility of an estimator with respect to a given loss function.

(d) Show that $T(x)$ in (b) is not admissible for estimating $e^{-\theta}$ with respect to squared-error loss function.

(3+3+3+5) = 15

10(a) Define the most powerful (MP) test of level α for testing a simple null hypothesis against a simple alternative hypothesis.

(b) Define unbiasedness of a test. Show that the most powerful test for a simple null hypothesis against a simple alternative is unbiased.

(c) Suppose $X_1, X_2, \ldots, X_n$ are i.i.d random variables with density

$$f_{\mu,\sigma}(x) = \frac{1}{\sigma} e^{-\left(\frac{x-\mu}{\sigma}\right)}, \quad x > \mu, \; -\infty < \mu < \infty$$

To test H_0 : $\mu = 0, \sigma = 1$ against

H_1 : $\mu = 1, \sigma = 2$,

find the MP test of level $\alpha = 0.05$. [3+(3+3)+7] = 16

INDIAN STATISTICAL INSTITUTE

Statistician's Diploma Examination : November 1975

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

Time : 4 hours Full marks : 100

(a) Figures in the margin indicate full marks
(b) Use of calculating machines is not permitted

Group A : Sample Surveys (50 marks)

(Attempt any three questions from this group)

1. $Y_1, Y_2, \ldots, Y_n$ represent observations on a random sample taken without replacement from a population of size N. Show that of all unbiased estimates of population mean which are linear combinations of $Y_1, Y_2, \ldots, Y_n$ the sample mean has least variance. Derive the expression for the variance of the sample mean. What should be the value of n in order that this variance may be exactly half of the variance of mean of random sample of same size taken with replacement.

(10+4+2) = 16

2(a) A finite population consists of 20 units serially numbered from 1 to 20. If the chance p_i of selecting the ith unit in any one draw is given, describe the operational steps to be adopted to draw a probability sample of 5 units with replacement from this population.

(b) A finite population consists of 3 strata. If the stratum standard deviations are in the ratio 1 : 2 : 3 and Neyman allocation provides the same size of samples from all the strata, what will be the proportional allocation of a sample of size 11 to the 3 strata ?

(c) Describe Lahiri's method for P.P.S sampling with replacement.

(8+4+4) = 16

3. Distinguish between Multi-stage and Multi-phase Sampling.

For two-stage sampling from a population with equal first stage units, the same number of second stage units is chosen randomly from each selected 1st-stage unit. Set up an unbiased estimate of the population total and derive its sampling variance. Also indicate how this sampling variance can be estimated unbiasedly.

(6+10) = 16

4. What is regression method of estimation ? Derive an approximate expression for the bias and variance of regression estimator, stating clearly the assumptions involved. Compare the precision of the regression estimator with that of a ratio estimator.

(3+8+5) = 16

5. Write notes on any two of the following :

(a) Sampling frame
(b) Systematic Sampling
(c) Double Sampling (8+8) = 16

Neatness

(Please turn over)

(NO)

Group B : Design and Analysis of Experiments (50 marks)

(Attempt any three questions from this group

6. Suppose Y_{ij}'s are independent samples from normal populations each with variance σ^2 and $E(Y_{ij}) = \mu_i$, $i = 1, \ldots, k$; $j = 1, \ldots, n_i$.

(a) What will be your testing procedure to decide about the equality of the means, if σ were known ?

(b) If σ is unknown, describe a method of testing

$$\mu_1 : \mu_2 : \ldots\ldots : \mu_k = 1 : 2 : \ldots\ldots k.$$

(6+10) = 16

7(a) Discuss the advantages and disadvantages of using a latin square design in practice.

(b) Suppose in a 4x4 latin square experiment, observations on the first treatment are missing in the first row - first column and second row - fourth column. Find the error sum of squares, and give a statistic for testing the equality of treatment effects.

(5+11) = 16

8(a) What is a balanced incomplete block design with the parameters v, b, r, k and λ ? For such a design, prove that $vr = bk$, $\lambda(v-1) = r(k-1)$.

(b) Suppose in a v x v latin square experiment, all the v observations on the first row are not available. Can you suggest a simple method of analysing this data ?

(3+4)+9 = 16

9(a) What is a 2^n factorial experiment ? How do you compute the main effects and interactions in a 2^3 experiment ?

(b) Give a design for a 2^5 factorial in blocks of eight plots so that no main effect or two-factor interactions are confounded.

(3+5+8) = 16

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

(a) Optimum size and shape of plots
(b) Youden square design
(c) Partial confounding
(d) Use of concomittant variables

(8+8) = 16

Contd.... 2

(N7)

INDIAN STATISTICAL INSTITUTE

Statistician's Diploma Examination - November 1975

Paper IV (Theoretical): Applied Statistics Group Papers

Time : 4 hours (for two groups) Full marks : 100

(a) Candidates will be required to answer questions from those two groups of subjects only for which they have already 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 (Half-Paper 50 marks)

(Attempt any three questions from this group.)

1. Explain briefly the problems of :

i) identification,

ii) multi-collinearity and

iii) bias and their consequences in the estimation of statistical demand functions from time series data. (6+5+5)=16

2. Suppose that the behaviour of households in a community can be represented by the following model

$$e_j = e_j^* + \beta_j (E - E^*)$$

where the household variables E and e_j stand for monthly total outlay and expenditure on a commodity j ($j = 1, 2, \ldots, m$), respectively, e_j^*, E^* and β_j ($\beta_j \geq 0, \sum \beta_j = 1$) are constant for all households.

(a) Give economic interpretation of the parameters e_j^*, E^* and β_j

(b) Show that the index of consumer prices in this model varies with household's total outlay.

(c) Discuss the implications of such variation for inter temporal comparisons of consumption disparities. (5+6+5)=16

3. (a) Define a homogeneous production function.

(b) Show that the model of production as formulated by Solow, Minhas, Arrow and Chenery is homogeneous. What is the degree of homogeneity?

(c) Indicate a method of estimating the substitution parameter in such models. (5+5+6)=16

4. (a) Explain the different entries in an Indian inter-industry transactions matrix.

(b) Using an inter industry table, establish the equivalence of the different approaches to the measurement of national income. (8+8)=16

Please turn over

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

i) Measures of income inequality

ii) Engel elasticities

iii) Decomposition of economic time series

iv) Measurement of returns to scale in production

v) Estimates of availability of food in India. (8+8)=16

NEATNESS

GROUP (b): STATISTICAL QUALITY CONTROL (Half-paper 50 marks)

(Attempt any three questions from this group.)

1.(a) Explain the meaning of "a process being in a state of statistical control".

(b) What is meant by process capability? Write down the steps you would take to determine the process capability of a given process. Under what condition(s) are modified control limits used and how are they constructed? Indicate some practical situations where by taking some possible alternative courses of action the use of modified control limits could be avoided. (4+12)=16

2.(a) For certain characteristic the average value was observed to be shifting downward in a linear fashion at a constant rate. The lower and upper specification limits are given. Suggest a control procedure to control the characteristic so that unnecessary resetting of the process could be reduced. Assume that the process variability is under control.

(b) From a process items are selected at random in a sample of size n and are classified into defective and non-defective by using go and no-go gauge, set at the specification limits. Suggest a suitable control chart procedure of

(i) n remains the same from sample to sample

and (ii) n varies from sample to sample. (10+6)=16

3.(a) Devise a sequential sampling procedure under which a lot with 100 p_1 % defective would be accepted with a probability of $1-\alpha$ and a lot with 100 p_2 % defective would be accepted with probability of β.

(b) Derive expressions for sample size n and the constant k, the two parameters of the acceptance sampling plan by variables when AQL, LTPD, consumer's risk and producer's risks are specified and sigma (σ) is not known. Write down the expression for probability of acceptance. (6+10)=16

4.(a) Explain what is meant by completely confounded and partially confounded effect by giving an example from industrial experiments. You should be brief in your answer.

(b) For a 2^5 factorial design with factors A,B,C,D and E, obtain the system of confounding in 8 blocks with 4 plots each such that the effects AD, BE, ABC, BCD, CDE, ACE and ABDE are confounded.

(c) Write short note on the method of steepest ascent as a tool for optimising response. (5+5+6)=16

NEATNESS (2)

GROUP (c): STATISTICAL METHODS IN GENETICS

(Candidates not available)

(NO) Please turn over

GROUP (d): VITAL STATISTICS AND DEMOGRAPHY (Half-paper 50 marks)

(Attempt any three questions from this group.)

1. Describe in details the components method of population projection. (16)

2. (a) Explain the concepts of stationary and stable populations.

(b) How does a quasi-stable population differ from a stable population? Explain how the concept of stable population has been used to estimate the demographic characteristics of a country which does not have reliable census and vital statistics. (6+10)=16

3. (a) Prove $$\eta q_x = \frac{2n \cdot \eta m_x}{2 + n \cdot \eta m_x}$$

the symbols having their usual significance.
State clearly the assumptions involved.

(b) In a population, known to be stationary, the complete expectation of life at birth is 55 years. Determine the level of (1) crude birth rate, (2) crude death rate and (3) crude rate of natural increase of the population.

(c) Explain the difference between complete and curtate expectation of life at age x. (6+6+4)=16

Write notes on any three of the following:

(a) Female age at marriage as a factor affecting fertility.

(b) Age bias in Indian Census.

(c) Net reproduction rate.

(d) Recall-lapse in demographic enquiries.

(e) Vital registration in India. (16)

NEATNESS (2)

GROUP (e): EDUCATIONAL & PSYCHOLOGICAL STATISTICS (Half-paper 50 marks)

(Attempt any three questions from this group.)

(a) Describe the method of construction of an achievement test. How are parallel forms of a test prepared? Also explain what are meant by the norms of a test for a given population.

(b) Explain the method of determining IQ from an intelligence test. Give the relation between mental age, chronological age and IQ. (4+3+3+5+1)=16

(a) Define the terms reliability and validity, as used in test theory.

(b) Test X has a validity coefficient of .65 and a reliability of .75, whereas the validity of test Y is .67 and its reliability .95 Each of these tests is a 50-item test. Which type of item (that in test X or in test Y) would probably show the greater validity for a 200-item test?

(c) Prove that, if a test of n items is a subtest of a test with m items ($n < m$), the correlation r_{nm} is

$$r_{nm} = \left(\frac{\frac{1 - r}{m} + r}{\frac{1 - r}{n} + r} \right)^{\frac{1}{2}}$$

where r is the reliability of a unit test. (5+5+6)=16

3. (a) What are the defects of the usual system of ranking in scholastic tests? Discuss in this connection the role of normalised scores and linear derived scores.

(b) In a certain college examination, student A got the highest marks in English Honours (in a group of 10), but he had the second highest score, together with another student, in Vernacular (in a group of 200).

Determine the percentile ranks of student A in the two subjects and interpret the results. (4+8+4)

4. (a) Compare and contrast the multiple factor analytic model and the true-score decomposition model of the classical test theory.

(b) Critically comment on any two tests of statistical significance in factor analysis. (8+8)

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

(a) Mahalanobis D^2 statistic.

(b) The use of regression analysis in psychometry.

(c) Compound symmetry pattern.

(d) The place of factor analysis in a test development programme.

(e) Non-linear models in education. (5½×3)

STATISTICS (2)

(N11)

INDIAN STATISTICAL INSTITUTE

Statistician's Diploma Examination - November 1975

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

Time: 5 hours Full marks:100

(a) Figures in the margin indicate full marks.

(b) Use of calculating machines is permitted.

GROUP A : Methods of Numerical Computation (26 marks)

(Attempt any two questions from this group.)

1. (a) From the following information on the decennial growth of population in India during 1901-1971, estimate by using finite differences the size of the Indian population in 1960:

Year	1901	1911	1921	1931	1941	1951	1961	1971
Population (in crores)	23.8	25.2	25.1	27.9	31.8	36.1	43.9	54.7

(b) Solve by Cramer's rule the following system of equations:

$$6.0x + 2.5y - 9.3w = 4.7$$
$$3.0x + 4.2y + 2.0z = 17.9$$
$$5.0x + 9.0z - 11.2w = 18.8$$
$$2.3y - 4.0z + 6.9w = 1.5$$

(2) (6+7)=13

2. (a) Given $AB = C$,

where $B = \begin{pmatrix} 63.5 & 154.5 & 10.6 \\ 154.5 & 637.1 & 46.2 \\ 10.6 & 46.2 & 6.5 \end{pmatrix}$

and $C = \begin{pmatrix} 2.9 & 1.0 & 0.6 \\ 4.1 & 0.0 & 8.2 \\ 1.5 & 2.1 & 5.4 \end{pmatrix}$

find A.

(b) Find by any numerical method known to you, the positive root of the equation $2x^x = 3x + 250$, correct to 3 decimal places. (7+6)=13

The ordinates $f(x) = (2\pi)^{-1/2} \exp(-x^2/2)$ corresponding to various values x of a random variable X following the normal law with mean zero and standard deviation unity, are given below:

x	f(x)
0.0	0.3989423
0.1	0.3989223
0.2	0.3988625
0.3	0.3987628
0.4	0.3986233
0.5	0.3984439
0.6	0.3982248

Using the above information, find approximately

(a) the value of x at which the ordinate is 0.3989;

(b) the probability that x takes values smaller than or equal to 0.6.

(6+7)=13

Please turn over

GROUP B: Descriptive Statistics (50 marks)

(Attempt all questions from this group)

4. The number of students, mean marks and s.d. of marks in a mathematical aptitude test for 5 different schools are given below. Compute the mean and s.d. of marks for all the schools combined.

school	number of students	mean marks	s.d. of marks
1	55	43.5	11.6
2	46	40.6	12.5
3	52	56.4	10.2
4	50	49.8	11.4
5	35	58.0	8.3

(2+5)=7

5. The results of a particular examination are shown below in a summary form:

Result	Percentage of candidates
Passed with 1st class Hons.	12.50
Passed with 2nd class Hons.	52.75
Failed	34.75

It is known that a candidate fails if he obtains less than 40 marks (out of 100) and passes with 1st class Honours if he gets at least 60 marks. Find the mean and s.d of the distribution of marks, assuming that it is of the normal type.

Also, find the mean and s.d of marks of the candidates who failed in the examination. (6+6)=12

6. During an investigation in an agricultural farm in Bengal, the length (in cm.) of green jute plant and weight (in gm.) of dry jute fibre were observed for 350 plants. With these data the following bivariate frequency distribution was obtained :

Weight of dry jute fibre (in gm.) Class mark \ Length of green plant (in cm.) Class mark	111.5	127.5	143.5	159.5	175.5	191.5	207.5
1.175	12	25	15	1			
2.775	1	4	33	59	29	3	
4.375	1		4	28	35	14	2
5.975				2	20	18	1
7.575				1	1	14	5
9.175					4	8	2
10.775						3	2
12.375							3

Compute the correlation coefficient and the correlation ratio of weight of dry fibre on length of green plant and comment. (5+5+2)=12

Please turn over

7. For the three variables x_1, x_2 and x_3, we have the following values for the total correlation coefficients :

$$r_{12} = 0.80, \quad r_{13} = -0.40 \text{ and } \quad r_{23} = -0.56.$$

Find the partial correlation coefficient $r_{13.2}$ and the multiple correlation coefficient $r_{1.23}$. Interpret the results. (2+2+2)=6

8. Statistics of goods carried by Indian Railways during 1963-64 to 1969-70 are given below:

Year (April-March)	Goods carried (thousand metric tons)
1963-64	192,262
1964-65	195,062
1965-66	204,150
1966-67	202,697
1967-68	197,795
1968-69	205,197
1969-70	209-088

(a) Fit a linear trend to the yearly totals, and hence obtain the trend line for monthly values.

(b) Estimate the amount of goods carried in April, 1970, assuming the seasonal index for April is 90 . (10+3)=13

GROUP C : Official Statistics (24 marks)

(Attempt both the questions from this group.)

9. From the official publications placed at your disposal, collect data on any three of the following items for latest four consecutive years for which figures are available :

(a) Total revenue receipts and expenditure of the Indian Central Government.

(b) Number of deaths from small-pox in West Bengal and Maharastra.

(c) Total number of persons employed and total ex-factory value of output in the different industries in India taken together.

(d) Number of reported cases of robbery and murder in India.

(e) Total quantity and value of exports from India of undressed skins of goats and kids. (3x4)=12

Either

It has been claimed by a Government source that our increasing success in the agricultural front over the past few years has enabled us to depend less and less on the mercy of other countries in the matter of food. Offer your comments on this claim, so far as only rice is concerned, by collecting relevant data for the latest 6 consecutive years from the official publications supplied. (9+3)=12

Or

You are required to compare the percentages of area under rice and wheat brought under irrigation facilities in the States of West Bengal, Punjab, Uttar Pradesh and Madhya Pradesh. Collect relevant data for any recent year from the official publications supplied and present these in a neat tabular form facilitating the comparison. (12)

Note : i) Full references to publications consulted are to be given along with answers. Any other information considered relevant is also to be given.

ii) The data furnished should be methodically arranged in a neat tabular form.

iii) Arithmetic manipulation of collected data, if necessary, is permissible. In that case, the original data collected are also to be shown side by side.

(P.T.O.)

INDIAN STATISTICAL INSTITUTE

Statistician's Diploma Examination - November 1975

Paper VI : Statistical Methods; Design & Analysis of Experiments and Sample Surveys (P r a c t i c a l)

Time : 5 hours Full marks: 1[illegible]

(a) Figures in the margin indicate full marks
(b) Use of calculating machines is permitted

GROUP A : Statistical Methods (40 marks)

(Attempt any two questions from this group)

1(a) A drug was given to 25 subjects half an hour before bed time while 25 other subjects were kept as controls. The next morning subjects reported the time taken by them to fall asleep. The following table gives the reported times of the two groups :

Time in minutes

Controls : 15, 25, 30, 15, 35, [illegible], 25, [illegible], 35, 35, 42, 25, 35, 2[illegible], 25, 40, 15, 15, 2[illegible], 3[illegible], 25, 1[illegible], 5[illegible], 3[illegible], 45.

Treated with drug : 25, 3[illegible], 40, 4[illegible], 15, 15, 20, 25, 30, 25, 20, 15, 10, 1[illegible], 15, 25, 35, 1[illegible], 10, 15.

Ascertain from the data whether the drug had any quickening effect on sleep (i) assuming normality, and (ii) without making any assumption about the form of parent distributions. (5+5)

(b) The following table gives estimates, together with their standard errors, of the expenditure per person per month on food grains in rural India as obtained in four independent surveys :

surveys	expenditure (estimate ± S.E.)
1	8.35 ± 0.032
2	7.60 ± 0.357
3	7.25 ± 0.721
4	7.43 ± 0.018

Are the estimates in statistical agreement ? If so, obtain the pooled estimate of the expenditure and its standard error. (6+2+2)

(Please turn over)

2(a) The net incomes of 162 and 419 families in two cities in a certain year are given below :

annual net income(Rs.)	no. of families	
	city I	city II
1500 - 3000	-	4
3000 - 4500	10	43
4500 - 6000	23	95
6000 - 7500	40	120
7500 - 9000	32	67
9000 -10500	23	51
10500 -12000	20	17
12000 -13500	4	9
13500 -15000	2	5
15000 -16500	1	3
16500 -18000	2	2
18000 -19500	-	1

On the assumption that the families included are a random sample of middle class families in the corresponding town, test whether the true income distributions for such families in the two towns may be supposed to be identical. (8)

(b) The age in years (x) and the chest-girth (y) were recorded for two groups of school boys consisting of 13 and 18 boys, respectively. On the basis of these data, the following values were obtained :

	Group I	Group II
$\sum_i x_i$	202.7	244.1
$\sum_i y_i$	23.3	53.8
$\sum_i x_i^2$	2742.56	3314.21
$\sum_i y_i^2$	44.77	174.40
$\sum_i x_i y_i$	315.57	729.82

Determine for each group the linear regression equation of y on x. Also, examine if the corresponding population regression equations (assumed linear) may be supposed to be parallel.

3(a) In a group of three independent tests of significance, the probabilities (for the test - statistic taking more extreme values than the observed values) were found as 0.0170, 0.0214 and 0.0352. Would you consider the combination of the tests to be significant ?

(Please turn

3(b. In an experiment on yield of sugar beets (tons/acre), there were two levels of irrigation treatment and three levels of fertilizer treatment, and each combination of treatments was carried out in five replications. The analysis of variance table was as follows :

Variation	S.S.	D.F.	M.S.
Irrigation	12[illegible].0	1	12[illegible].0
Fertilizer	221.7	2	11[illegible].0
Interaction	35.0	2	17.5
Error	1[illegible]6.0	24	4.5
T o t a l :	4[illegible]4.7	29	-

Assuming that it makes sense to regard the irrigation and fertilizer effects as random, estimate the components of variance. (6)

(c. The yield in bushels of a certain type and size of potato plot is found to be normally distributed with a standard deviation of 2.36. It is hoped that the application of a certain fertilizer will increase the yield by at least 7.5 bushel. How large a sample of plots should be used to detect a difference of this amount, using the mean sample yield as a criterion, with a test of size 5% and power 9[illegible]% ? (10)

GROUP B : Design and Analysis of Experiments (3[illegible] marks)

(Attempt all questions from this group)

4. Either
Obtain the layout of a randomised block experiment, using six treatments and a control in two blocks. (12)

Or
Obtain the layout of a suitable 2^4 - experiment in 2^2 blocks in a single replicate. (12)

5. Either
In order to study the variation in moisture content from plant to plant and from leaf to leaf of turnip greens, a 5x5 latin square experiment was conducted. The treatments were the five times of sampling (i, ii, iii, iv, v). The plan and data on moisture content of turnip leaves are presented in the following table :

Table : Plan and moisture content

plants	leaf size (A to E, smallest to largest)				
	A	B	C	D	E
1	(i) 82.37	(iv) 83.15	(v) 84.29	(ii) 84.93	(iii) 85.62
2	(iii) 81.40	(i) 80.77	(iv) 81.40	(v) 83.84	(ii) 82.93
3	(ii) 83.32	(iii) 84.53	(i) 84.50	(iv) 85.00	(v) 85.50
4	(v) 80.92	(ii) 81.00	(iii) 83.20	(i) 83.86	(iv) 83.28
5	(iv) 80.88	(v) 82.16	(ii) 83.83	(iii) 81.98	(i) 82.51

Test all meaningful hypotheses and state your conclusions. Compute the standard error of a treatment mean. (2[illegible])

(Please turn over)

5. (contd.) OR

The following table gives the data on the use of drugs in the treatment of leprosy. On each patient six sites on the body at which leprosy bacilli tend to congregate were selected. The variate x (based on laboratory tests) is a score representing the abundance of leprosy bacilli at those sites before the experiment began. The variate y is a similar score after several months of treatment. Drugs A and D are antibiotics while drug C is an inert drug included as a control. Ten patients were selected for each treatment :

Table : Scores before (x) and after (y) treatment

drugs					
A		D		C	
X	Y	X	Y	X	Y
11	6	6	0	16	13
8	0	6	2	13	10
5	2	7	3	11	18
14	8	8	1	9	5
19	11	18	18	21	23
6	4	8	4	16	12
10	13	19	14	12	5
6	1	8	9	12	16
11	8	5	1	7	1
3	00	15	0	12	20

Analyse the data to find out whether the drugs had any effect after eliminating the dependence of Y on X. Is it worthwhile to perform a covariance analysis in the present example ? (25)

GROUP C : Sample Surveys (30 marks)

(Attempt all questions from this group.

6(a) In a state lottery the tickets bearing the following numbers under three series are to be included in the draw :

series	serial number of ticket
A	000001 - 015008
B	000001 - 002573
C	000101, 000200 - 001352.

By drawing a simple random sample of size 3, determine the series and ticket numbers that will get 1st, 2nd and 3rd prizes in the lottery. A ticket can get only one prize.

(b) 200 cultivators' holdings are stratified as follows according to size :

holding size (acres)	number of holdings
0 - 40	42
41 - 80	48
81 - 120	32
121 - 160	34
161 - 200	18
201 and above	26

Draw a stratified random sample (with proportional allocation, of size 10. (5+5)

(Please turn over)

7. Either

A population is divided into 5 strata in a certain crop survey. The following table gives for the ith stratum (i = 1,2,....,5) the number of villages (N_i), the standard deviation of the variable under study per village (σ_i, in Kg.) and the cost per unit (C_i, in Rs.) :

i	N_i	σ_i (in Kg.)	C_i (in Rs.)
1	3,780	23.5	3.50
2	5,260	18.6	2.75
3	8,270	27.6	2.25
4	4,180	27.2	3.00
5	2,800	16.8	2.50

The total approved cost of the survey for estimating the population mean is Rs.5,000/- and the overhead cost is Rs.550/-.

Obtain (a) the optimum values of n_i's, the sample size of ith stratum and (b) the corresponding variance of the estimator. Find also (c) the relative sample sizes for the different strata, using Neyman's allocation. (8+3+6) = 20

Or

The following preliminary study was conducted to allocate resources in an optimum manner in two-stage sampling.

10 beets were chosen from each of 100 plots in a uniformity trial in order to develop a sampling technique for the determination of the sugar percentage in field experiments on sugar beets. The sugar percentage was obtained separately for each beet. The following mean squares (MS) were obtained :

MS between plots (i.e. primary units) = 2.9254
MS between beets (i.e. sub-units within plots = 2.1374

The average cost was Rs.2.80 per primary unit and Rs.1.30 per sub-unit.

(a) Determine the optimum number of primary units and sub-units for a preassigned total cost of Rs.1050/- including an overhead cost of Rs.50/-.

(b) Obtain an unbiassed estimate of the variance of the sample mean of the two-stage sample scheme obtained in (a).

(c) If the standard error of the sample mean is not to exceed 0.2% what combinations of primary units and sub-units are allowable ? (Assume the total number of sub-units to be the same as that obtained under scheme (a).) (10+5+5) =2.

INDIAN STATISTICAL INSTITUTE

Statistician's Diploma Examination - November 1975

Paper VII (Practical): Applied Statistics Group Papers

Time: 5 hours Full marks: 150

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

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

(Attempt any two questions from this group.)

The following data give the average per capita monthly consumption expenditure at 1952-53 prices of the bottom and top deciles as well as the whole population of rural India over the period from 1952-53 to 1961-62.

Draw suitable sketches to represent this information and comment on the changes in the level and distribution of consumption in India.

Table 1. Average per capita consumption in constant rupees of 1952-53, Rural India, 1952-1962.

Year / Decile	1952-53	1953-54	1954-55	1955-56	1956-57	1957-58	1958-59	1959-60	1960-61	1961-62
Bottom decile	6.35	5.16	4.89	5.45	5.27	5.16	5.56	5.86	5.79	4.97
Top decile	50.13	40.00	37.53	44.13	36.33	40.94	40.59	37.91	38.23	47.97
Whole population	20.57	16.53	15.29	17.78	15.73	16.27	17.23	16.74	16.80	18.30

(25)

Using a random number table, select a series of 20 numbers (corresponding to, say, quarterly data over a period of five years). Partition this series into seasonal, trend, cyclical, and random components. (25)

A crude input-output table for a country for the year 1961 is given below:

(Rs. crores)

Producing sector	Using sector: Agriculture	Manufacturing	Services	Final demand	Total output
Agriculture	1000	500	100	400	2000
Manufacturing	200	1000	800	2000	4000
Services	300	600	150	950	2000
Number of persons employed (million)	100	20	30		

(a) Obtain the Technological Coefficient matrix (A) with the usual Leontief assumptions.

(b) If it is planned to produce gross output of 9000, 3250 and 4000 Rs. crores in the three sectors (agriculture, manufacturing and services) show that the plan is inconsistent.

Please turn over

3. (c) (contd.) If in 1975 the government decided the final demand in the three sectors to be 1000, 2500 and 1500 £ crores, find the gross output in the three sectors.

(d) What will be the total employment in the economy if the production plan in (c) is carried out? (25)

GROUP (b): STATISTICAL QUALITY CONTROL (Half-paper 50 marks)

(Attempt any two questions from this group.)

1. (a) The following table gives the results of inspection for spots on enamel plates of the same size.

Plate no.	Number of spots	Plate no.	Number of spots
1	8	13	10
2	7	14	5
3	9	15	28
4	11	16	24
5	12	17	35
6	8	18	10
7	10	19	23
8	18	20	11
9	10	21	13
10	10	22	16
11	18	23	14
12	19	24	13

(i) Examine whether the process is under statistical control and set standards for the average number of spots per plate.

(ii) Obtain the upper control limit for the number of spots per plate which would be exceeded only once in 100 when the process is stable at the above standard.

(b) Tool wear causes a steady trend in the average length of a manufactured item at the time of machining. The following values of $\bar{X}$ and R were obtained from sub-groups of size 4 from the machine at half-hour intervals.

Sub-group no.	$\bar{X}$	R	Sub-group no.	$\bar{X}$	R
1	8.220	.004	9	8.251	.005
2	8.228	.006	10	8.252	[illegible]
3	8.223	.005	11	8.256	.002
4	8.231	.009	12	8.256	.004
5	8.236	.011	13	8.263	.007
6	8.241	.008	14	8.262	.011
7	8.244	.010	15	8.270	.009
8	8.246	.006	16	8.256	.002

Estimate the standard deviation of the process and the equation of a least squares trend line fitted to $\bar{X}$ values. (10+15)=25

Please turn over

2. (a) A supplier inspects his own products under a Dodge and Romig 2% AOQL plan, screening all rejected lots before shipment. The lot size is 1000, and the assumed process average is 1%. Single sampling is used. The purchaser inspects the same lots under a Dodge and Romig 5% LTPD plan, returning all the rejected lots to the supplier. The plan used is taken from the process average column headed '0.51 - 1.00'. Single sampling is used.

(i) What is the probability that a 2% defective lot will pass both the supplier's and purchaser's inspection?

(ii) What is the probability that a 2% defective lot will pass the supplier's inspection and be rejected by the purchaser's?

(b) A lot contains a large number of items. Devise a single sampling attribute classification plan for it to meet the stipulations:

AQL = .01 α = .05

LTPD = .09 β = .10

where α and β are the producer's and consumer's risks respectively. What would be the probability of acceptance for a lot containing 5% defective items under this plan. (12+13)=25

The yields in a 2^3 factorial experiment in three randomised blocks on the effects of three fertiliser Nitrogen (N), Phosphorous (P) and Potassium (K) on certain response are tabulated below:

Treatment	Block I	Block II	Block III
(1)	32	43	27
n	41	19	47
p	26	36	28
k	29	39	64
np	61	52	30
pk	51	35	63
nk	26	66	48
npk	38	48	34

i) Estimate the various main effects and interactions.

ii) Carry out an analysis of variance (ANOVA) test and comment.

iii) Point out the drawback of the above design and suggest your modification if any to improve its efficiency. (8+12+5)=25

Please turn over

GROUP (c) : STATISTICAL METHOD IN GENETICS (Half-paper 50 marks)

(Attempt any two questions from this group

1. The joint segregation of the two factors A, a and B, b has been studied and the frequencies are given in the table below. Test (i) whether there is any linkage between A and B. If so, estimate the linkage ratio and test whether the different families agree in giving the same result. Calculate the standard error of the pooled estimate of linkage.

Cross : (Aa Bb x aa bb)

Phenotype :	AB	Ab	aB	ab	total
Family :					
I	191	37	36	203	467
II	380	62	61	398	901
III	90	17	16	100	223
total :	661	116	113	701	1591

(25)

2. The analysis of the data pertaining to the number of eggs laid (x_1), egg weight (x_2) and age at maturity (x_3) of 218 Brown Leghorn pullets which were progenies of different sires, gave the following results :

	D.F.	Mean Squares		
		SS (x_1)	SS (x_2)	SS (x_3)
Between sires	14	463	70	2673
Within sires	201	337	12	1582
		Mean Products		
		sp(x_1x_2)	sp(x_1x_3)	sp(x_2x_3)
Between sires	14	48	- 432	59
Within sires	201	2	- 179	2

Find the best discriminat function and the corresponding ANOVA. (25)

3. Blood specimens were collected for 427 people and after testing according to blood-group antigens, the following classifications were made for different phenotypes in two systems of tests :

A - B - O blood group		M - N blood group	
Phenotype	Number	Phenotype	Number
O	122	MM	211
A	97	MN	162
B	163	NN	54
AB	45		
	427		427

Determine the gene-frequencies of A-B-O and M-N blood groups by using the maximum-likelihood and gene-counting methods; obtain standard errors of the estimates and comment on the blood group behaviour of the community. (25)

Please turn over

GROUP (d): VITAL STATISTICS AND DEMOGRAPHY (Half-paper. 50 marks)

(Attempt any three questions from this group.)

1. Below is given an extract of a life table for a male population (the radix of the life table being $l_0 = 1000$) :

Age group x to x+5	Life table population in the age group x to x+5 ${}_5L_x$	Age group x to x+5	Life table population in the age group x to x+5 ${}_5L_x$
0 - 5	4852	40 - 45	4508
5 - 10	4804	45 - 50	4400
10 - 15	4786	50 - 55	4249
15 - 20	4765	55 - 60	4034
20 - 25	4733	60 - 65	3727
25 - 30	4692	65 - 70	3300
30 - 35	4647	70 +	6512*
35 - 40	4586		

* Denotes life table population aged 70 and above.

Given r (the true rate of natural increase for the stable population) = .0025, compute the stable age-distribution for the population by quinquennial age groups. (16)

2. From the following data, calculate :

(a) crude birth rate, assuming that women in ages 15-44 form 20 percent of total population;

(b) gross reproduction rate, assuming the sex-ratio at birth to be 107 males : 100 females.

Age group	Number of women	Annual no. of births
15 - 19	255,927	29,074
20 - 24	254,851	63,850
25 - 29	232,731	54,766
30 - 34	217,017	42,644
35 - 39	172,273	25,858
40 - 44	170,932	12,004

(7+9)=16

3. (a) Calculate the conventional infant mortality rates from the following data :

Year	Live births	Infant deaths among births of: Preceding year	This year
1967	12502	-	-
1968	12412	135	345
1969	12375	129	340
1970	11282	120	320

Also calculate the more refined infant mortality rates.

3. (b) (contd.) From the following values of life table functions, calculate complete expectation of life at ages 20, 25 and 30 :

Age group (x to x+5)	$_5q_x$
20 - 25	.0212
25 - 30	.0284
30 - 35	.0352
35 - 40	.0501

$l_{20} = 20{,}000$

$\mathring{e}_{35} = 24.8$

(10+6)=16

4. The following table shows the population of a certain country as recorded at successive censuses :

Census year	Population (million)
1880	50.2
1890	62.9
1900	76.0
1911	92.0
1921	105.7
1931	122.8
1941	131.7
1951	150.7
1961	179.3

Fit a logistic curve to the data and comment on the goodness of fit.
(13+3)=16

HEATNESS (2)

GROUP (c): EDUCATIONAL AND PSYCHOLOGICAL STATISTICS (Half-paper 50 marks)

(Attempt any three questions from this group.)

1. (a) Give the error of measurement, standard deviation of true scores, correlation between observed and error scores, and the index of reliability for each of the following tests:

Test	Number of items	Mean	Stand. devn.	Reliability
A	50	100.0	15.0	0.91
B	100	211.6	25.7	0.84
C	80	57.4	11.3	0.78
D	700	361.9	76.5	0.87
E	200	127.4	21.9	0.76

(b) Assume a normal distribution of error scores. Using the above data give the true score limits (approximately 0.3 per cent level) for persons making each of the following scores:

(i) A score of 115 on Test A;

(ii) A score of 500 on test D; (11+5)=16

Please turn over

2. (a) In a study of 4 educational tests, A, B, C and D, the following results were obtained :

Test	Mean	Standard deviation	Number of items	Reliability
A	18.4	4.2	30	0.72
B	28.9	9.8	60	0.96
C	37.2	8.1	50	0.90
D	39.2	11.5	75	0.92

Suppose that Test A was administered to a new group by a different investigator. The mean and standard deviation of scores obtained by this new sample on Test A turned out to be 25.3 and 8.4 respectively. What reliability would you expect the test to have for this new group?

(b) It is reported that test B has been administered to a new group and the reliability coefficient is only 0.90. What would account satisfactorily for this lowered reliability without indicating any faults of test administration or scoring?

(c) Test C is administered to a new group with the following results:

Mean= 31.9, stand. devn.=12.7, reliability = .96.

Are these results in reasonable agreement with those reported in the above table for Test C?

(d) A teacher wished to use Test D for sectioning a class and finds a mean score of 45.3 and a standard deviation of 3.9. What comment would you make on this proposal ? (4+4+4+4)=16

3. Using the data presented in the table below, answer the questions appearing under, (a) - (e):

Test	Mean	Standard deviation	No. of items	Reliability	Validity*
A	16.5	4.4	30	0.72	0.68
B	12.6	3.5	20	0.77	0.50
C	53.2	10.7	100	0.88	0.68
D	32.3	8.0	50	0.91	0.71
E	66.3	17.2	120	0.95	0.75

* Refers to the product moment correlation between School grade average (criterion) and the test in question. It is to be noted that the criterion itself has a reliability of 0.70.

(a) If Test A is lengthened to a 100 -item test, what would you expect the new mean, stand. devn., reliability and validity to be? (Assume that the criterion has not been altered).

(b) Which of the above five tests is, in its present form, best for use in predicting school grade average?

(c) If Test B is lengthened to increase its reliability to 0.90, how many new items will be needed? What will the new validity be, assuming that the criterion test remains unchanged?

(d) If the reliability of the criterion is 0.70 as assumed, what is the correlation between true criterion scores and true scores for each of the above five tests?

(e) Give the true variance, and error variance for Test C. Estimate the true and error variances for Test C if it is increased to 300 items. (3+3+3+4+3)=16

Please turn over

4. (a) Extract the Spearman general factor-("g") from the following correlation matrix:

$$\underset{\sim}{R} = \begin{bmatrix} - & .48 & .40 & .40 & .30 \\ .48 & - & .40 & .40 & .25 \\ .40 & .40 & - & .34 & .28 \\ .40 & .40 & .34 & - & .20 \\ .30 & .25 & .28 & .20 & - \end{bmatrix}$$

(Please note that the elements in the principal diagonal are unity.)

(b) Calculate the obtained communality for each of the four tests from the following factor-loading matrix extracted by the technique of principal axes:

Test No.	FACTOR I	FACTOR II
One	.9	-.2
Two	.3	.9
Three	.8	-.9
Four	.9	.2

Also, calculate the percentage of total variance and percentage of common variance accounted for by each of the factors.

(8+8)=16

NEATNESS (2)

INDIAN STATISTICAL INSTITUTE

Statistician's Diploma Examination : November 1978

Paper VIII : Subjects of First paper of Specialisation (Theoretical)

Time : 4 hours Full marks : 1[illegible]

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

(b) Figures in the margin indicate full marks

GROUP A : ECONOMIC STATISTICS

Econometrics - Special Paper I

(Attempt any five questions from this group)

1(a) State clearly the assumptions of the classical linear regression model. Derive the least squares estimators of the parameters in this model and state their properties.

(4+4+4) = 12

(b) Show that if the disturbances are normally distributed, the least squares estimators of the coefficients in the classical linear regression model are the same as the maximum likelihood estimators

8

2. Specify the generalised linear regression model and derive (without using any transformation) the best linear unbiased estimator of the regression coefficients of the model. Also obtain the variance - covariance matrix of the generalised least squares (GLS) estimator. Interpret the GLS estimator as the classical least squares estimator in a transformed model.

(12+3+5) = 20

3. Describe the three-stage least squares (3SLS) method of estimation of a simultaneous equation model. Under what conditions will the 3SLS estimators be identical with the 2SLS estimators ? Justify your answer.

(12+8) = 20

4. State Paretos' Law of Income Distribution and mention the important properties of the distribution. Derive the expression for the Lorenz curve and the Lorenz ratio for the Pareto distribution

(8+12) = 20

How is an empirical input-output table constructed ? Discuss the various conceptual and practical problems faced in preparing such a table.

(8+12) = 20

6 Mention the commonly used algebraic forms of the Engel curve and discuss their properties. For any one of these forms, describe how you would estimate the parameter of the curve from grouped family budget data.

(12+8) = 20

State the important properties of the Cobb - Douglas production function. Describe how you would estimate the parameters of the Cobb - Douglas production function in a simultaneous equations set-up using cross-section data on firms.

(5+15) = 20

Please turn over

8. Consider a consumer with a given income (M) to be allocated on two goods with given prices (P_1 and P_2). Suppose his utility function is given by :

$$u = q_1^{\alpha} q_2^{\beta}, \alpha, \beta > 0,$$

where q_i is the quantity of the i-th goods consumed by him.

(a) Derive the first-order conditions for the maximisation of the consumer's utility.

(b) Will the second-order conditions for utility maximisation be satisfied ?

(c) Derive his demand curve for the two-goods.

(d) Calculate the income and price elasticities of demand for the two goods.

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

(a) Use of dummy variables for seasonal adjustment.

(b) Durbin - Watson Statistic and its uses.

(c) Pooling of time-series and cross-section data in demand analysis.

(d) Elasticity of substitution and its properties. (10+10)=20

GROUP B : TECHNO-COMMERCIAL STATISTICS

Statistical Quality Control - Special Paper I

(Attempt any five questions from this group)

1. (a) Explain how you would proceed to condense mass of data on a measured characteristic in the form of frequency distribution and draw histogram.

(b) Write down the formula for finding out the mean and standard deviation for the grouped data.

(c) Explain how histograms of measurable characteristics can be made use of to infer about production process.

(8+4+8) = 20

2. (a) Why do you prefer to plot the mean in a control chart instead of individual values while controlling process level ?

(b) Indicate with suitable examples in industry the situations in which the following charts are made use of for control purposes. Also write down the expressions for the control limits in each case.

i) $\bar{X}$-R chart

ii) Chart with modified control limits

iii) Control chart on defects per unit.

(c) Explain briefly without deriving any expressions the cusum chart procedure for controlling the mean of a measured characteristic.

(4+12+4) = 20

Please turn over

3.(a) Distinguish between 'specification limits' and 'control limits'.

(b) Let L and U be the lower and upper specification limits for a product characteristic $\bar{x}$ which is distributed as normal with s.d. σ. C_1 is the cost of rejection when $\bar{X} < L$ and C_2 is the cost of rejection when $\bar{X} > U$. Determine the level at which the process should be maintained so that the total cost of rejection is minimised.

(c) Discuss the statistical aspects of tolerances with suitable industrial examples.

(4+12+8) = 20

4.(a) Derive parameters in a single sampling plan by variables given the lower specification limit on quality of a product when the standard deviation of the lot is not known.

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

(10+10) = 20

5.(a) Describe the main features of any of the published attribute sampling inspection plans which give average quality protection for the consumer.

(b) For a Dodge-Romig single sampling inspection plan with lot quality protection in terms of AOQL, show that

$$AOQL = y\left(\frac{1}{n} - \frac{1}{N}\right)$$

where N = lot size, n = sample size, c = the acceptance number and $y = x \sum_{r=0}^{c} e^{-x} \frac{x^r}{r}$; $x = np'$, p' being the equality level corresponding to AOQL.

(10+10) = 20

6.(a) Explain the use of fractional factorial designs in industrial experimentation.

(b) Write down the treatment combinations for the principal block for $\frac{1}{4}$ replicate of a 2^5 factorial design assuming that the highest order interaction is non-significant.

(c) In an experiment with n factors, each at two levels, what is the minimum number of treatment combinations that you must try in order to estimate the main effects, assuming all interactions to be absent ? Indicate the design and its analysis. (6+8+6)=20

Write brief notes on the following :

(a) Quality costs.

(b) Advantages of factorial experiments over single factor experiments

(c) Work sampling

(8+8+4) = 20

Group C : Biometric Methods

Special Paper I

(No candidate available)

Please turn over

GROUP D : DESIGN AND ANALYSIS OF EXPERIMENTS

Statistical Aspects - Special Paper I

(Attempt any five questions from this group)

1(a) Describe the method of analysis of covariance suitable for two-way classified data.

(b) Discuss in which situations this technique is appropriate.

(c) Find an expression for the estimate of the difference between the average effects of two pecified levels of one of the ways of classification,

(3+4+8) = 15

2. There are five makes of a calculating machine. It is intended to compare the makes by obtaining the time required for certain jobs after eliminating the effects of investigators and also the nature of jobs. For this purpose 5 different jobs are given. Assuming that at the most 5 machines of each make are available give a suitable design and indicate its method of analysis.

(12+1) = 13

3(a) Define partially balanced incomplete block designs.

(b) Indicate in what situation such designs are preferable to P.B.I.B designs.

(c) Show that the square lattice designs are P.B.I.B designs.

(d) Write and prove 4 parametic relation in two-associate P.B.I.B designs.

(5+5+5+5) = 20

4. A B.I.B design with parameters $v = b = 7$, $r = k = 3, \lambda = 1$ was first suggested for an experiment. But as the error degrees of freedom for the design was considered inadequate, the design was repeated once more so as to have in all 14 blocks of three plots each.

(a) Obtain the design and outline the method of its analysis.

(b) Obtain an expression for the variance of the estimate of the difference between two treatments.

(8+8+4) = 20

5(a) Explain the concepts of (i) Factorial experiments (ii) Main effects (iii) Two factor interaction (iv) Three factor interaction and (v) Confounding.

(b) Describe a method of analysis of factorial experiments with each factor at two levels.

(8+12) = 20

Please turn over

3(a) Obtain one replication of a confounded design for 3 factors each at three levels using blocks of 9 plots each and saving all main effects and two factor interactions.

(b) Outline the method of analysis of the design.

(c) Split the sum of squares due to main effect of each factor into linear and quadratic components.

(7+7+6) = 20

7(a) Define Switchover designs with illustration.

(b) In which situation are these designs suitable ?

(c) Outline the method of analysis of a Switchover design with three factors after eliminating residual effect.

(4+4+12) = 20

8. Write notes on any two of the following :

(a) Recovery of interblock information

(b) Youden Square.

(c) Weighing Design.

(4+4+12) = 20

GROUP B : SAMPLE SURVEYS

Theoretical Aspects : Special Paper I

(Attempt any five questions from this group)

(a) A random sample of size n is drawn without replacement from a population containing N units with variate values x_i (i = 1,2...., N). Derive the sampling error of sample mean $\bar{x}_n$.

(b) Consider another scheme of sampling where the N-th unit is always selected and n-1 units are selected without replacement from the remaining N-1 units. Define a suitable estimator $\bar{x}'$ of population mean based on x_N and $\bar{x}_{n-1}$. Under what conditions will $\bar{x}'$ be more precise than $\bar{x}_n$ of (a) above.

(12+8) = 20

(a) A population consists of L strata with Ni (i = 1,2,...,L) units in the i-th stratum. If a sample of n_i units (i=1,2,...,L) be drawn at random with replacement from the i-th stratum, define an estimator of population mean $\bar{y}$. Also derive the expression of the sampling error of the estimator.

(b) If c_i be the cost of sampling one unit from the i-th stratum of the population, derive the expression for optimum allocation of n sample units over the L strata so that for a fixed cost the variance of estimator of population mean is minimised.

(12+13) = 20

Please turn over

3. A sample of n units has been drawn at random with replacement from a population of N units with variate values (x_i, y_i) $(i = 1,2,\ldots, N)$. Define an estimator $\hat{R}$ of mean of y per unit of x based on the sample. Derive the sampling error of $\hat{R}$. Also derive (as shown by Hartly & Ross or otherwise) an expression for the upper bound of the bias in $\hat{R}$ in relation to $\sqrt{v(\hat{R})}$ in terms of C.V.(x) and n.

(2+12+6) = 2

4(a) A district consists of N villages and the i-th village contains M_i households. A scheme of sampling consists in selecting at random with replacement n_1 $(< N)$ villages and selecting n_2 households at random within each selected village for collecting data on the study variate y. Define an estimator of total of y for the district. Find the sampling error of the estimator.

(b) If C_1 be the cost of going to a first stage unit(village) (including preparation of frame for selection of second stage units viz. households) and C_2 be the cost of collecting data on y from one selected second stage unit (household), find what should be optimum allocation of n_1 and n_2 for a fixed cost so that the variance of the estimator is minimised.

(4+10+6) = 20

5. Define what is meant by self-weighting design. A state consists of K - strata, the i-th stratum consisting of Ni first stage unit. From the ith stratum n_i $(< Ni)$ first stage units are selected with probability proportional to size with replacement. Define how the within village sampling fraction should be determined so that design becomes self-weighting at state level.

(4+16) = 20

6. A population consists of nK units. The units are randomly grouped into n groups of K units each. From each group one unit is selected with PPS for estimating character y. Define an estimator for population total Y. Also derive an expression for the sampling error of the estimator.

(8+12) = 20

7. Write notes on any two of the following :

(a) Circular systematic sample

(b) Horvitz - Thompson estimator

(c) Separate and combined ratio estimates

(d) Cluster sampling

(e) Pilot survey.

(10+10) = 20

Please turn over

GROUP F : TECHNIQUES OF COMPUTATION

Numerical Analysis - Special Paper I

(Attempt any five questions from this group)

1(a) Discuss briefly the sources of error in a computed numerical result. Show that if the first significant figure of a number is K and the number is correct to n significant figures, then the relative error is less than $\frac{1}{K} \cdot 10^{-n+1}$

(b) Consider the expressions

$$u = (a-b)/c \quad \text{and} \quad v = a/c - b/c.$$

Assume that a,b and c are all positive and have no inherent error and that a ≈ b. Show that the relative round-off error in v can be much greater than that in u. Illustrate with a = 3.41, b = 3.38, c = 0.75 using two digit floating point arithmetic.

(4+8)+(8+2) = 20

2. Define the general problem of interpolation. Derive Newton's forward interpolation formula with the error term. What is inverse interpolation ? How can one use inverse interpolation to evaluate a root of a numerical equation? (3+6+6+2+3)=20

3. What is the basic difference between Newton - Cote's and Gauss' quadrature formulae ? State the advantages and disadvantages of Gaussian formulae. Explain the method of determining the abscissas and weights in a Gaussian formula. Illustrate your answer by determining abscissas and weights in the formula

$$\int_0^1 \sqrt{x}\, f(x)\, dx \simeq a_1 f(x_1) + a_2 f(x_2).$$

How would you estimate the error in the above formula ?

(2+2+3+5+3) = 20

4. Describe an iterative method to compute a real root of the equation $f(x) = 0$ and derive the condition for its convergence. Generalize the method to solve two simultaneous equations

$$f(x,y) = 0 \ , \quad g(x,y) = 0.$$

Derive an iteration formula to compute $\frac{1}{a}$ without performing the division operation.

(5+5+5+5) = 20

5. Describe Taylor's series method and Euler's method for the solution of an ordinary first order differential equation of the form $y' = f(x,y)$ with the initial condition $y(x_0) = y_0$. How will you compare the efficiency of the two methods ?

(8+8+4) = 20

Please turn over

6. Describe, giving relevant derivations, a method to compute all the latent roots and latent vectors of a real square matrix. 2

7. Describe Gauss' elimination method to solve a system of linear equations. In what way does Jordan's method differ from Gauss' method ? Compare the two methods with respect to the number of basic arithmetic operations required.

(9+1+7) = 20

Group G : Statistical Inference

General Theory - Special Paper I

and

Group H : Probability Theory

Basic Probability - Special Paper I

(No candidate available)

INDIAN STATISTICAL INSTITUTE

Statistician's Diploma Examination - November 1975

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 and Economics of Planning

Section I : Indian Economics (50 marks)

(Attempt any three questions from this group.)

1. (a) Indicate the orders of magnitude of the percentage shares of agriculture in the

(i) national income,

and (ii) labour force, of India

(b) Can you give a rough estimate of the ratio of average productivity of labour in agriculture to the average productivity of labour over the entire economy in India?

(c) Briefly describe the trends in the shares referred in (a) above over the last two decades or so.

(d) What is the significance (theoretical or empirical) of these trends for the growth of the economy? (2+6+6)=16

What are the principal "sectors of origin of national income" distinguished in the official national income estimation in India? Briefly describe the procedures adopted for calculating their separate contributions to national income. Comment on the adequacy of these procedures. (4+8+4)=16

(a) Give a brief account of the major crops in India. Where are they grown? During what seasons?

(b) Which crops and regions have contributed prominently to the growth of agricultural output in India over the past decade or so? Why?

(c) Has there been any shift in the seasonality pattern of total agricultural production over the same period? (6+8+2)=16

India is said to have experienced her worst inflation after the 2nd World War during mid-1972 to end-1974, and this inflation has been variously characterised as (i) a stagflation, (ii) a part of world-wide inflation, and (iii) a failure of monetary policy.

(a) Give a short description of the nature and extent of this inflation.

(b) Do you agree with the view that the inflation really started in mid-1972 and ended in end-1974? Give arguments in support of your answer.

(c) Comment on the validity of each of the three characterisations (i), (ii) and (iii) given above (separately, or jointly, as you choose). (4+2+10)=16

Please turn over

5. Write short-notes on any two of the following :

(a) Road versus water transport as policy options in the present Indian context (where the alternatives exist).

(b) The ratio of direct to indirect tax revenue in India: its behaviour over time and significance for income distribution.

(c) The recent debate over the nationalisation of wholesale trade in foodgrains in India.

(d) The pattern of credit allocation from commercial banks in India before and after the nationalisation of fourteen major private commercial banks. (8+8)=16

Section II : Economics of Planning (50 marks)

(Attempt any three questions from this section)

6. (a) Derive the Harrod-Domar formula for the rate of growth of national income.

(b) What is its use for planning purposes?

(c) Derive the policy implications (for savings) of the following types of targets of planning in terms of the formula in (a) above or some extension of it:

(i) the (immediate) achievement of a stipulated rate of growth of total consumption;

(ii) the (immediate) achievement of a stipulated rate of growth of per capita income;

(iii) the acceleration of growth to approach a stipulated rate of growth of total income, given the initial levels of both income and savings. (3+3+2+2+6)=16

7. What is the usefulness of input-output (IO) analysis in development planning? Critically review any attempt at the application of IO analysis to Indian planning. (6+10)=16

8. Give a theoretical analysis of the draft 5th five-year plan of India indicating the major points of departure of its approach to the problems sought to be tackled by planning from those of the previous plans. (16)

9. What are, roughly, the nature and quantum of "foreign aid" inflow into the Indian economy? How have these changed over the successive five-year plans? Discuss in this connection the prospects of India's achieving self-reliance by 1980. (6+6+4)=16

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

i) The Mahalanobis Model of growth

ii) Shadow prices in linear programming and their possible role in resource allocation.

iii) The role of land reform in development planning, with special reference to Indian planning.

iv) Employment policy in the overall context of a planning strategy in India. (8+8)=16

NEATNESS (4)

Please turn over

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 (70 marks)

(a) Use a separate answer-book for this Section.

(b) Attempt any four questions from this Section.

1. A manufacturing concern produces four different types of products A, B, C, D, each of which is required to be processed in two different types of machines P and Q. Machine type P has three different varieties P_1, P_2, P_3, while machine type Q has two, namely Q_1 and Q_2.

Although product D can be processed in any of the three varieties of machine type P and any of the two varieties of machine type Q, the other products, namely A, B, C have certain restrictions on the use of machine varieties. While product A does not have any restriction on machine type P it must have to be processed only in machine variety Q_2. Product B, on the other hand, cannot be processed in machine variety P_1. Finally, machine variety P_3 is not available for product C.

Data regarding the following items of information are tabulated below:

i) Processing time (in hours) for one unit of each product in each variety of machine.

ii) Machine availability (in hours) for one month.

iii) Operating cost (in Rs.) at full capacity.

iv) Material cost (in Rs.)/per unit.

v) Price (in Rs.) per unit.

Tabular presentation of basic data

Machine variety	Product processing time in hours per unit of product A	B	C	D	Machine availability for one month (hours)	Operating cost at full capacity (Rs.)
P_1	10	-	8	7	690	200
P_2	8	5	5	3	700	210
P_3	4	11	-	10	650	150
Q_1	-	10	7	3	710	250
Q_2	5	4	2	9	600	120
Material cost per unit (Rs.)	0.60	1.10	0.90	0.80		
Price per unit (Rs.)	2.00	3.00	2.50	3.25		

Formulate the problem in a linear programming model in which the objective is to maximise the profit. (17)

Please turn over

2. (a) A manager operates a single-channel service facility in which customers arrive randomly in a Poisson stream at an average rate of λ. The service times are exponentially distributed and the average rate of service is μ. The manager can vary μ. Suppose that he wishes to choose μ so that only a fraction θ of the time will the total queue length equal or exceed n members.

Express the value of μ that achieves this condition explicitly in terms of λ, θ, and n.

(b) An Oil Refinery receives crude oil at an average rate of one tanker per day. The unloading facilities, which operate 24 hours per day, can handle only one tanker at a time, but can unload tankers at an average rate of two per day. Under the usual assumptions of Poisson arrival and exponential service time distribution, determine

(i) average number of tankers in the system,

(ii) average time spent by a tanker in the system,

(iii) average waiting time of a tanker in the queue,

(iv) percentage of time in which exactly two tankers are in the system. (17)

3. The food distributing authority intends to have an investigation made regarding the daily supply of rice to a region for which the pattern of daily consumption (in tons) is as follows:

Daily consumption (tons)	Frequency (number of days for which consumption is in given range)
500 – 509	6
510 – 519	11
520 – 529	20
530 – 539	30
540 – 549	18
550 – 559	10
560 – 569	5

The storage facility of the region is restricted to a capacity of 300 tons, and the initial stock is 250 tons.

Simulate the daily consumption for 10 days and determine the optimal daily supply from the following three levels; explaining clearly the optimality criterion used by you.

(i) 500 tons, (ii) 530 tons, (iii) 550 tons.

A part of the Random Number Table is reproduced below :

4652	3319	8431	2150	2352	2472
0043	3488	9031	7617	1220	4129
7148	1943	4390	1749	2030	2327
7353	6007	9410	9179	2722	8445
0641	1489	0928	0385	8488	0422

(17)

Please turn over

4. Three factories F_1, F_2 and F_3 manufacture a given product P. Each of these factories has a certain production capacity: Factory F_1 can turn out 100 units of product P; factory F_2, 200 units and factory F_3, 50 units.

These factories supply five wholesale houses H_1, H_2, H_3, H_4 and H_5. Each wholesale house has a definite demand for product P. For H_1, it is 145; H_2, H_4 and H_5 need 45 each, and H_3 has a demand of 70.

Both the production capacities and demands refer to the same period of time, for instance, a year. The elements of the following cost matrix give the cost (in hundred rupees) of transporting one unit from a factory to a wholesale house.

Cost Matrix

(Elements in hundred rupees)

Factory	Wholesale house				
	H_1	H_2	H_3	H_4	H_5
F_1	2	5	3	4	8
F_2	1	9	4	5	10
F_3	5	6	1	7	4

Formulate the least cost transportation problem and indicate how it is solved. (17)

5. A company has contracted to manufacture and deliver a unique order of 18 units over the next 90 days. The contract calls for delivery of 5 units at the end of 30 days, 5 units at the end of 60 days and 8 units at the end of 90 days. Because of limited production facilities and the company's decision to work over time rather than expand its facilities over this 90-day period, these units will be produced at increasing costs according to the following relationship, where u_i is the cost in rupees of manufacturing x_i units in the ith 30 - day period (i=1, 2, 3):

$$u_i = 5{,}000 + 1{,}000\, x_i\, (x_i - 1).$$

If any excess or parts of units are manufactured and on hand at the end of any 30-day period, these may be carried over to the next period at a holding cost (interest, storage, security protection, special care and testing, etc.) of Rs.1,000 for each unit carried over from one 30-day period to the next. Holding costs are proportional to partial units carried over. Thus, letting v_i be the number of units carried over into the ith period (i.e. it is the starting inventory for the ith period), the holding cost h_i is, in rupees,

$$h_i = 1{,}000\, v_i.$$

Determine the optimal scheduling sequence to minimize the sum of the manufacturing and holding costs. There is no starting inventory. Carried over inventories need not be integral numbers of units. (17)

6. Write notes on any three of the following:

(i) Assignment technique
(ii) Economic lot size in Inventory Control
(iii) Replacement problems
(iv) Warehousing problem in dynamic programming. (17/3x3)=17

NEATNESS (2)

Please turn over

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

(a) Use separate answer book for this section.

(b) Attempt Question No.1 and any other three from this sec

1. Prepare Trading and Profit & Loss Account and Balance Sheet as on 31st March, 1975 from the following balances:

		Rs.
M. Marigold's Capital Account	...	1,19,400
M. Marigold's Drawings Account	...	10,550
Sundry Creditors	...	59,630
12% Loan Account (Credit)	...	20,000
Cash in hand	...	3,030
Cash at Bank	...	18,970
Sundry Debtors (including Badri Das for Dishonoured Bill of Rs.1,000)	...	62,000
Bills Receivable	...	9,500
Provision for Doubtful Debts	...	2,500
Fixtures & Fittings	...	8,970
Plant & Machinery	...	28,800
Stock, 1st April 1974	...	89,680
Purchases	...	2,56,590
Manufacturing Wages	...	40,970
Sales	...	3,56,430
Returns Inwards	...	2,780
Salaries	...	11,000
Rent and Taxes	...	5,620
Interest & Discount (Debit)	...	5,870
Travelling Expenses	...	1,880
Repairs & Renewals	...	3,370
Insurance (including Premium of Rs.300 per annum paid up to 30.9.1975)	...	400
Bad Debts	...	3,620
Commission Received	...	5,640

Adjustments:

(a) Stock in hand on 31st March, 1975 was Rs.1,28,960.

(b) Create a provision of 5% on Sundry Debtors.

(c) Depreciate Plant and Machinery by 5% and Fixtures & Fittings by 10% per annum.

(d) Interest on Loan for the last two months is not paid.

(8+10+7)=25

2. The following errors were discovered in the books of S. Selvam on December 31, 1974. The difference in the Trial Balance had been entered in a Suspense Account and on correction of the errors the Suspense Account was eliminated:

i) The total of Purchases Day Book had been undercast by Rs.1,500.

ii) The Discount column on the debit side of the Cash Book has been posted to the Discount Column on the credit side of the Cash Book, Rs.210.

iii) Rs.760 for repairs to Motor Van had been taken to the Motor Van Account.

iv) A cheque received from B Bannister Rs.390 had been debited in the Cash Book, but the double entry had not been completed.

v) The Returns Outwards Book had been overcast by Rs.500.

Show by means of Journal entries how these errors would be corrected in the books of S. Selvam. (15)

3. What do you mean by Imprest System of Petty Cash ? Give the ruling of a Petty Cash Book showing therein five distinct heads of petty expenses. (15)

4. What is a Balance Sheet? Discuss its importance. What is meant by 'marshalling' of assets and liabilities in a Balance Sheet? (15)

5. Explain the term 'Bill of Exchange'. Discuss the general accounting procedure for dealing with Bill transactions in a large business concern. (15)

6. Write notes on any three of the following :

(a) Principles of double entry.

(b) Capital and Revenue Expenditure.

(c) Journal Proper.

(d) Trial Balance.

(e) Valuation of Assets. (15)

Section II : STATISTICAL METHODS IN BUSINESS (30 marks)

(a) Use separate answer book for this section.

(b) Attempt any two questions from this section.

1. Describe the various forecasting techniques in Marketing Research. (15)

2. Why is it necessary to use statistical sampling for audit purposes? Illustrate the application of sampling techniques in auditing. (7+8)=15

3. Write notes on any two of the following :

i) Job Evaluation

ii) Planning a Market Survey

iii) Incentive Schemes (7½x2)=15

GROUP C : BIOMETRY
Special Paper II: Biometric Methods.

GROUP D : DESIGN & ANALYSIS OF EXPERIMENTS
Special Paper II: Combinatorial Aspects.

GROUP E : SAMPLE SURVEYS
Special Paper II: Organisational Aspects.

(No candidate available)

Please turn over

GROUP F: TECHNIQUES OF COMPUTATION - PRACTICAL

Special Paper II : Numerical Computation
(Practical with Desk Calculators)

Time: 5 hours Full marks: 100

Attempt any four questions from this group.

1. (a) A result of R = 29.25 was obtained in determining the gas constant for air. Knowing that the relative error of this is 0.1%, find the limits within which R lies.

(b) How many digits are to be taken in computing $\sqrt{20}$ so that error does not exceed 0.1%.

(c) An approximate number a = 24253 has a relative error of 1%. How many correct digits has it ?

(d) Find the difference $u = \sqrt{2.01} - \sqrt{2}$ correct to three significant digits.

(e) If all coefficients in the following definition

$$f(x) = \frac{5.03241x + 0.11095}{0.75995x + 0.014915}$$

are rounded numbers. To how many significant digits can f(x) be determined, where x is known only to rounded 3.26?

(f) Show that the number $(2.46)^{1/64}$ is known within less than one unit in the place of its fifth significant digit if 2.46 is known to be correctly rounded to three digits.

(2+3+3+2+10+5)=25

2. (a) Using Newton-Raphson's method, compute the negative root of the equation

$$x^4 - 3x^2 + 75x - 10000 = 0$$

correct to five significant digits.

(b) For the system

$$2x^2 - xy - 5x + 1 = 0$$

$$x + 3\log_{10}x - y^2 = 0$$

find the positive roots to four significant digits by iterative method, given approximate values of the roots x = 3.5 and y = 2.2

(10+15)=25

3. (a) The values x and $y = 0^x$, correct upto three places of decimal are given below:

x	3.50	3.55	3.60	3.65	3.70
y	33.115	34.813	36.598	38.475	40.447

Form the difference table and hence calculate the value of y at x = 3.54. Compare the result thus obtained with the true value.

(b) For the function y = f(x) we have the table

x	0.05	0.15	0.20	0.25	0.35	0.40	0.50	0.55
y	0.9512	0.8607	0.8187	0.7788	0.7047	0.6703	0.6065	0.5769

Find f(0.45). (10+15)=25

Please turn over

4. (a) Evaluate $I = \int_0^1 \frac{dx}{1+x}$ correct upto 5 places of decimal.

Compare the result so obtained with the true value of the integral.

(b) Compare the accuracy of different quadrature formulae with three ordinates for the integral

$$I = \int_{-1}^{1} \sqrt{2+x}\ dx = 2\sqrt{3} - \frac{2}{3} \qquad (13+12)=25$$

5. Obtain by Taylor's Series five consecutive starting values for the numerical solution of $\frac{dy}{dx} = 2x - y$ with $x_0 = 1$, $y_0 = 3$. Check the values and add three more lines to the table. Compare your result with those obtained from the exact analytical solution

$$y = 2x + 3e^{1-x} - 2. \qquad (25)$$

6. Using square root method compute the inverse of the matrix

$$A = \begin{pmatrix} 30 & 31 & -17 & 9 \\ 31 & 39 & -3 & 23 \\ -17 & -3 & 46 & 24 \\ 9 & 23 & 24 & 30 \end{pmatrix}$$

and the determinant of A. (20+5)=25

7. Find the first two numerically largest eigen values and the corresponding eigen vectors of the matrix

$$\begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix}$$

(25)

GROUP G : STATISTICAL INFERENCE
Special Paper : Special Topics

GROUP H : PROBABILITY
Special Paper : Limit Distributions

(No candidate available)

INDIAN STATISTICAL INSTITUTE

Statistician's Diploma Examination : November 1975

Paper X : Subjects of Third Paper of Specialisation(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.

GROUP A : ECONOMIC STATISTICS

Special Paper III - Practical

(Attempt any three questions from this group)

1. Below are shown the corn yields in bushels per acre in 8 experimented plots each receiving a fixed amount of phosphorus(240 lbs/acre). Four levels of nitrogen were applied, and there were two plots receiving each "treatment".

amount of nitrogen (lbs/acre) (N) :	80	160	240
yields(bushels/acre) (Y) :	38.2,35.0	67.2,67.8	127.6,125.8	121.1,114.2

Fit a production function of the form $y = a+b\sqrt{N+c}N$ to the data and find the value of R^2 for the equation.

What is the level of nitrogen giving maximum rate of yield per acre in the light of the fitted regression equation ? (32)

2. The following shows the estimates of net national product and aggregate consumption in the US during 1933-48, expressed at 1929 prices :

year	net national product(10^9)	consumption(10^9)	year	net national product(10^9)	consumption(10^9)
1933	58.2	55.0	1941	113.3	97.9
1934	64.4	63.3	1942	147.8	98.2
1935	75.4	73.1	1943	125.2	98.3
1936	85.0	80.5	1944	147.1	102.2
1937	92.7	84.4	1945	168.8	109.1
1938	85.4	83.0	1946	131.4	122.3
1939	92.3	87.0	1947	130.9	124.9
1940	101.2	91.7	1948	134.7	127.5

Fit a consumption function of the form

$$C_t = \alpha+\beta Y_t+\gamma C_{t-1}$$

to the data and estimate short-and long-term marginal propensities to consume. (32)

Please turn over

3. The following figures relate to the US economy during 1922-35 :

year	retail price of pork(cents/lbs)	pork consumption(lb/person)	disposable personal income ($/person)
1922	26.8	69.7	541
1923	25.3	74.2	616
1924	25.3	74.0	610
1925	31.1	66.8	636
1926	33.3	64.1	651
1927	31.2	67.7	645
1928	29.5	70.9	653
1929	30.3	69.6	682
1930	29.1	67.0	604
1931	23.7	68.4	515
1932	15.6	70.7	390
1933	13.9	69.6	364
1934	18.8	63.1	411
1935	27.4	48.4	459

Ignoring movements in the general price level, estimate income and price elasticities of demand for pork by fitting the constant elasticity demand function to the above data. Test whether the price elasticity is numerically less than unity. (32)

4. $y_{1t} = \beta_{12} y_{2t} + \gamma_{11} x_{1t} + u_{1t}$ is one equation in a two-equation econometric model which contains two other exogenous variables x_{2t} and x_{3t}. The following is the matrix of sums of squares and products of the observations :

	y_1	y_2	x_1	x_2	x_3
y_1	99	202	32	19	45
y_2		895	5	48	164
x_1			102	2	4
x_2				48	6
x_3					42

Find the 2SLS estimates of β_{12} and γ_{11} and also estimate their standard errors. (32)

Please turn over

A farmer wants to choose an optimum combination of the following five activities :

activity	unit of output
Wethers	1. lb of wool
Ewes	1 ewes and lambs
Peas	1[illegible] bushels
Swedes	1[illegible] sugar bags
Vealers	1 cow and calf

The following shows the profits etc. per unit level of each of these activities :

activity no.:	1	2	3	4	5
profit/unit (£, :	17.417	37.[illegible]	58.[illegible]	33.000	19.[illegible]
land(acres)/unit:	2.778	5.[illegible]	1.000	1.000	1[illegible]
capital(£/unit) :	38.867	14.250	67.[illegible]0	42.000	35.000
labour hours/unit:					
season 1	0.356	[illegible].640	2.000	10.000	[illegible].080
2	0.356	1.840	2.500	0	0.080
3	0.711	1.280	0.500	0	0.080
4	0.444	0.800	4.000	0	0.080

Find the levels of these five activities for which total profit is maximum subject to the following constraints :

(a) the farmer has 250 acres of land and £ 2000 of capital;

(b) labour availability cannot exceed 130 hours in any of the four seasons. (32)

Neatness (4)

GROUP B : TECHNO-COMMERCIAL STATISTICS

Section 1 : Statistical Quality Control

(a) Use separate answer book for this section
(b) Attempt all questions from this section

1. In a factory, during 1971, a study was undertaken to analyse the causes of rejection and it was found that machining defects formed the major cause of rejection. It was therefore decided to control machining defects by control chart techniques.

The following data were collected.

Please turn over

1. (contd.)

Sample Size : 25

date	sample no.	no. of defectives	date	sample no.	no. of defectives
15/11	1	3	17/11	9	1
	2	2		10	2
	3	0		11	1
	4	0		12	3
16/11	5	1	18/11	13	3
	6	4		14	1
	7	0		15	0
	8	1		16	2

During the course of study some actions were suggested in order to decrease the number of rejections. The following data were collected after these suggestions had been implemented.

Sample Size : 25

date	sample no.	no. of defectives	date	sample no.	no. of defectives
13/12	1	2	15/12	9	1
	2	0		10	0
	3	1		11	0
	4	1		12	2
14/12	5	3	16/12	13	2
	6	0		14	0
	7	2		15	1
	8	0		16	0

Analyse the above sets of data by means of control chart technique and estimate the process levels before and after the actions were taken. Comment on the results. (29)

2. In an experiment to study the effects of glass type (a,b and c) and phosphor type (A,B,C and D) on the brightness of a TV tube screen, the following data on the current necessary to produce a certain brightness was obtained. Each combination of glass-type and phosphor type was replicated thrice. Analyse the effect of the two factors and their interaction.

Glass type \ Phosphor type	A	B	C	D
a	28	30	28	19
	29	32	27	23
	26	34	30	18
b	24	24	23	13
	23	26	21	19
	25	23	23	15
c	31	29	26	30
	35	28	23	30
	32	24	25	27

(20)

Please turn over

3. Attempt either (a, b) or (c)

Either

(a) What Dodge & Romig LTPD sampling plans would you use for the following conditions ?

	sampling plan single/double	LTPD % def.	lot size	process average % def.
i)	single	2.	1500	0.3
ii)	double	1.0	10000	0.2

(b) What Dodge and Romig AOQL sampling plans would you use for the following conditions ?

	sampling plan single/double	AOQL %	lot size	process average % def.
i)	double	5.0	450	3.0
ii)	single	1.5	200	0.10

(5+5) = 10

Or

(c) Given LTPD = 10.0%. Find a single sampling plan that has a probability of acceptance of 0.10 when lot quality is 0.10 and lot size is 5000 and minimizes inspection when the process average is 0.005. What is the AOQL of the plan ? What quality lots have a 0.95 chance of being accepted ? (10)

GROUP B : Section II : Operations Research

(a) Use separate answer book for this section
(b) Attempt any two questions from this section

1. The Knox - Lix Company has the option of using one or more of four different types of production processes. The input restrictions and profitability data are given below.

	process 1	process 2	process 3	process 4	availability
Labour	1	1	1	1	15
Material Y	7	5	3	2	120
Material Z	3	5	10	15	100
Unit Profit	4	5	9	11	

Obtain the amounts to be produced under each process to maximize the total profit (15)

Please turn over

2. Arrivals at a telephone booth are considered to be poisson, with an average time of 20 minutes between one arrival and the next. The length of a phone calls is assumed to be distributed exponentially with mean 6 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 queues that form from time to time?

(iii) The telephone company will install a second booth when convinced that an arrival would expect to have to wait at least 2 minutes for the phone. By how much will the flow of arrivals have to increase in order to justify a second booth? (5x3)=15

3. (a) A new automatic machine is to be purchased by a plant. It is required to decide, on the spares that are to be ordered with the machine, optimally.

The following information is available.

(i) Cost of a spare R.200/-

(ii) Cost of down time due to non-availability of spare = R.1000/-

(iii) Demand for spares arises as given below:

No. required	Probability
0	0.80
1	0.10
2	0.06
3	0.02
4	0.01
5	0.01
6 or more	0.00

(b) Suppose the cost of down time due to non-availability of spare changes from R.1000 to R.1500 will your decision made in (a) be changed? Give reason. (9+6)=15

Section II (Alternative) : Elements of Book-keeping and Accountancy - Practical (No candidate available)

GROUP B: Section III: Statistical Methods in Business (20 marks)

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

(b) Attempt both the questions from this section.

1. Three news papers namely, New India, Old Times and Freedom are published in a certain city, and a survey shows that of the adult population 20% read New India, 16% Old Times and 14% read Freedom, 8% read both New India and Old Times, 5% read both New India and Freedom, 4% read both Old Times and Freedom and 2% read all the three. If an adult is chosen at random find the probability that

(i) he/she reads at least two if it is known that he reads at least one paper.

(ii) he/she reads exactly one of these papers. (10)

Please turn over

2. A cigarette company uses two different methods to test the moisture content of tobacco processed. Samples of tobacco are drawn from the process and its moisture content tested by both the methods. The following data are available for analysis.

Sample no.	Method A	Method B	Sample no.	Method A	Method B
(1)	(2)	(3)	(1)	(2)	(3)
1	9.9	8.7	7	10.8	10.2
2	9.8	9.4	8	11.9	10.3
3	9.9	9.0	9	10.0	9.8
4	8.8	9.0	10	9.7	8.8
5	8.4	8.0	11	10.1	8.9
6	8.6	8.2	12	10.0	9.1

Can we conclude that method A gives a higher value than Method B in general? Give reason. (10)

GROUP C : BIOMETRIC METHODS

Practical - Special Paper III

(No candidate available)

GROUP D : DESIGN AND ANALYSIS OF EXPERIMENTS

Practical - Special Paper III

(Attempt all questions from this group)

1. An experiment was carried out to compare 9 strains of paddy, which are denoted by (i, j), i = 1, 2, 3; j = 1, 2, 3. The design adopted was an incomplete block design, in two replications, each consisting of three blocks with three plots per block. The table below gives the plan and the yield of grain in lb. per plot.

Replication 1	Block 1	(3,2) 11.3	(3,3) 7.8	(3,1) 7.5
	Block 2	(2,2) 12.0	(2,3) 9.3	(2,1) 8.3
	Block 3	(1,2) 8.8	(1,3) 9.6	(1,1) 6.5
Replication 2	Block 1	(2,3) 6.3	(3,3) 8.5	(1,3) 8.0
	Block 2	(2,2) 13.0	(3,2) 14.0	(1,2) 13.3
	Block 3	(2,1) 11.3	(3,1) 8.9	(1,1) 10.3

(a) Analyse the data after recovering inter-block information

(b) Order the strains of paddy.

(c) It was discovered later that the 9 strains of paddy were actually obtained from a factorial structure, the numbers i and j being the levels of two factors A and B respectively. Split up the treatment sum of squares into components due to the main effects and the interactions and examine if the data indicate the presence of interactions. (16+8+8)=32

(1151) Please turn over

2. An experiment in $\frac{1}{8}$th replicate of a 2^8 factorial experiment with factors A, B, C, D, E, F, G, and H - each at two levels and arranged in 4 blocks of 8 plots each, was carried out to study the effect of these factors on the yield of paddy crop. In the table below are given the layout plan and yield of paddy (in lbs per plot), the treatments within a block being assigned at random to the 8 plots in the block:

Layout plan and yield of paddy

Block I		Block II		Block III		Block IV	
Treatment	Yield	Treatment	Yield	Treatment	Yield	Treatment	Yield
adg	26.99	aeg	31.34	bdefh	26.22	(1)	25.72
cdef	32.81	abcdf	20.85	acfh	29.90	bcefh	36.98
aceh	33.26	cdfgh	32.29	adefg	30.32	abgh	27.77
bdh	33.13	ef	31.02	egh	32.95	abcde	32.24
abf	31.30	adeh	30.45	cd	31.73	adfh	34.29
abcdefgh	35.43	beh	38.13	abcdgh	29.81	bdfg	31.73
fgh	34.15	bdeg	30.45	abe	28.46	acefg	33.58
bceg	35.57	abfgh	33.64	befg	28.91	cdegh	31.60

(a) Identify the defining contrasts and the confounded interactions.

(b) Analyse the data and interpret the results of your analysis. (8+24)=32

3. A nutritional experiment on the growth of pigs was carried out with 4 treatments. From each of the four lots, four pigs were randomly selected and for each pig the initial age in days (x_1) and the initial weight in lbs (x_2) were recorded. For each lot 4 treatments were then assigned randomly to the selected pigs and the experiment carried out. After the experiment, the rate of gain in weight in lbs. per day (y) was noted for each pig. The figures come out as follows. Analyse the data to see if, apart from variations in initial age and weights, the treatments differ significantly among themselves.

	Treatments											
	1			2			3			4		
lots	x_1	x_2	y	x_1	x_2	y	x_1	x_2	y	x_1	x_2	y
I	78	61	1.40	99	75	1.31	78	80	1.71	77	62	1.40
II	80	54	1.28	80	64	1.12	70	47	1.23	78	62	1.37
III	75	45	1.55	75	52	1.29	71	47	1.41	95	57	1.22
IV	83	57	1.34	62	50	1.40	85	59	1.49	67	39	1.36

(32)

Please turn over

GROUP - I : SAMPLE SURVEYS

Practical - Special Paper III

(Attempt Question No.1 and any two of the rest)

1. A stratified sample of 125 households is to be drawn with simple random sampling without replacement (SRS WOR) in each stratum using the data given in the following table as population-values, where all the 1000 households in a village are stratified according to the size of land holdings (Z) in acres. The table gives the stratum means $\bar{X}_i$ and $\bar{Y}_i$ and standard deviations $S_i(x)$ and $S_i(y)$ for annual income (x) and annual expenditure (y) in hundred. The correlation coefficients ρ_i between x and y and the number of households N_i are also given for each stratum. Let $R_i = \bar{Y}_i/\bar{X}_i$ and $R = \bar{Y}/\bar{X}$ where $\bar{Y}$ and $\bar{X}$ are population means of y and x respectively.

Table

Size of land holdings (in acres)	$\bar{X}_i$	$\bar{Y}_i$	$S_i(x)$	$S_i(y)$	ρ_i	N_i
$Z = 0$	6.1	9.0	130	140	0.1	250
$0 < Z \leq 5$	11.2	12.1	160	200	0.3	260
$5 < Z \leq 15$	25.2	19.5	500	600	0.6	360
$Z > 15$	250.5	100.2	800	1000	0.4	130

i) Obtain the sample sizes for strata in case of (a) proportional allocation, (b) allocation proportional to $N_i d_i$ where $d_i^2 = S_i^2(y) - 2R_i \rho_i S_i(x) S_i(y) + R_i^2 S_i^2(x)$; (c) allocation proportional to $N_i q_i$ where q_i is d_i^2 in (b) above with R_i replaced by R.

ii) Obtain the mean square error of separate ratio estimator for $\bar{Y}$ under the allocations (a) and (b)

iii) Obtain the mean square error of the combined ratio estimator for $\bar{Y}$ under the allocations (a) and (c).

iv) Obtain the relative efficiency of the combined ratio estimators in (iii) over that of separate ratio estimator in (ii) above. (40).

2. Following table gives the annual expenditure (y) on stationary for four departments of two universities. The Universities are such that total annual budget allocations (x) in the departments are same for both the universities.

Table

Departments	x (in Rs.'00)	y (in R.'00) University A	y (in R.'00) University B
1	150	14	20
2	100	10	14
3	80	9	11
4	50	6	11

Please turn over

2. (contd.) Consider the following sampling strategies (combination of sampling and estimation procedure) to estimate the total of Y (for all the four departments under consideration) respectively for A and B.

$S_1 \equiv$ (SRS WOR, usual unbiased estimator)

$S_2 \equiv$ (ppx WR, usual unbiased estimator)

$S_3 \equiv$ (SRS WOR, ratio estimator)

i) Give estimates under above sampling strategies based on samples of size 2 (departments) in each case, by drawing samples using random sampling numbers provided.

ii) Identify the sampling strategy (from among under consideration) which is the best for population A and that which is best for B.

iii) Suggest one, out of these three strategies, which may work well for both the universities A and B. (30

3. To estimate the total value of Puja purchases and also the average per household made by the residents of a locality, a two stage sample survey was conducted by considering the houses as first stage units and households as second stage units. A simple random sample of 12 houses was selected without replacement from 180 houses of the locality under consideration and the value of Puja purchases was obtained from a sample of 1/3 of these households. These households were also selected with equal probabilities and without replacement. The data are given in the following table

serial number of sample houses	number of households observed	Puja purchasings (in Rs.) for the sampled households
1	3	50, 40, 15
2	5	500, 300, 700, 300, 200
3	6	100, 200, 150, 50, 40, 20
4	4	40, 80, 100, 50
5	2	400, 800
6	3	700, 400, 200
7	2	300, 200
8	5	40, 60, 100, 125, 60
9	4	90, 50, 80, 100
10	3	40, 60, 70
11	2	100, 200
12	3	400, 200, 150

(a) Obtain estimates for (i) the total value of Puja purchases made by the locality and also (ii) the average per house and (iii) per household.

(b) How will you modify your estimate in (iii) above, if it is known that the average number of households per house in the locality is 10.

(c) Obtain the variance of the estimate in (b) above. (30)

4. Work out a scheme for conducting a Live-stock survey in a State, giving outlines of (i) sampling design, (ii) information to be collected, (iii) tables to be prepared, (iv) organisational structure, (v) time estimate and (vi) cost-estimate. (30)

Please turn over

GROUP F : TECHNIQUES OF COMPUTATION

(Special Paper III - Practical
based on UR Machines)

(Attempt all questions from this group.)

In the operations to be carried out you have to deal with two decks of cards (1) Household card (Master), (2) Consumption cards (Detail). There is one household card (Master) corresponding to a group of consumption cards (Detail). The information punched on these cards and the columns for each of the items of information, is given on the card design given below:

Card Designs

Household card (Master)

Srl. no.	Item	No. of cols.	Card cols.	Remarks
(1)	(2)	(3)	(4)	(5)
1.	Card design index	4	1 - 4	Punch 1111
2.	State Code	2	5 - 6	
3.	Sub-sample	1	7	
4.	Srl.no. of Block	4	8 - 11	
5.	Sample h.h. no.	1	12	
6.	Household size	2	13 - 14	

Consumption card (Detail)

1.	Card design index	4	1 - 4	Punch 2222
2.	State code	2	5 - 6	
3.	Sub-sample	1	7	
4.	Srl. no. of block	4	8 - 11	
5.	Sample h.h. no.	1	12	
6.	Household size	2	13 - 14	To be transferred from Master card 1111
7.	Item code	2	15 - 16	
8.	Quantity consumed	5	17 - 21	Kg-gr
9.	Value of consumption	5	22 - 26	Rs - Paise

STEP 1

You are given a deck of Household consumption cards (CDI, 2222). Sort in State Code X Srl.no. of Block X Sample h.h. no. Mark the sorted cards as 01 on the top and bottom cards of the deck. Hand it over to the examiner.

Write down the card count on the answer sheet. (10)

STEP 2

You are supplied with a deck of Household consumption cards (CDI,2222) and Household card (Master) (CDI, 1111), both arranged in State Code X Srl. no. of Block X Sample h.h. no.

Please turn over

STEP 2 (Contd.)

Match and merge on a card collator, the two decks into one deck taking into consideration the sorting order. Master card should follow the Household detail cards belonging to the same household. Select unmatched cards from both the decks. Mark merged cards as 02 and selected cards 03 and 04 respectively for Master cards and Detail cards. Hand over these cards to the examiner.

On the answer sheet write down the CDI State Code, Srl. no. of Sample Block and Sample h.h. no. of the selected cards and also the count of merged cards.

Prepare wiring chart for the Collator and hand it over to the examiner. (30)

STEP 3

You are supplied with a merged deck of cards. Each Master card succeeds Detail cards belonging to the corresponding household.

Make use of the Calculating Punch to obtain for each household the (1) total expenditure and (2) the per capita expenditure and punch the per capita expenditure on the Master Cards on columns 27 to 31.

After the end of the above operation, separate Master cards from Detail cards without disturbing the order of both the sets.

Mark the Master cards as 05 and Detail cards as 06 and hand over to the examiner.

On the answer sheet write down the card count of both the sets. Prepare a wiring chart and hand over to the examiner. (20)

STEP 4

You are given a deck of household cards (Master) arranged in ascending order of h.h. no. code. With the help of an Accounting machine, prepare a statement showing the following information in a suitable format and also give total value of consumption of all the households.

1. State Code
2. Sub-sample
3. Srl.no. of Block
4. Sample h.h. no.
6. Household size
7. Per-capita expenditure
8. Value of consumption. (25)

Viva voce (15)

Please turn over

GROUP G : STATISTICAL INFERENCE

Special Paper III - Practical

(Attempt all questions from this group)

1. A sample of size 20 was drawn from a population with an elementary probability law

$$dF(x) = \frac{1}{\theta} e^{-x/\theta} dx, \quad \theta > 0, \quad 0 < x < \infty$$

and the sample mean came out as 12.6. Find the maximum likelihood estimate of θ and variance of its estimate.

How would you modify the estimate if it is known, during drawing the sample, that observations exceeding 60 were rejected? (10)

2. Of 131 earthquakes that occured in a country over 20 years, 27 occurred in Spring, 27 in Summer, 18 in Autumn and 29 in Winter. Test whether occurance of earthquakes in the country may be considered to vary from season to season. (20)

3. Following table gives the number of trials to re-learning required by 8 post operatives (E-rats) and number of trials to re-learning by 15 normal rats (C-rats)

E-rats : 20, 55, 29, 27, 75, 56, 31, 45

C-rats : 23, 8, 24, 15, 6, 21, 9, 16, 22, 18, 14, 28, 13, 30, 11.

Apply suitable non-parametric test for the hypothesis of the homogeneity of the two populations. (15)

4. Reaction times to visual stimuli were obtained from 20 young normal men under three conditions A, B and C of stimulus display. The mean reaction times in hundredths of seconds were 21.05, 21.65 and 28.95 respectively. The sample covariance matrix was

$$S = \begin{bmatrix} 2.2605 & 2.1763 & 1.6342 \\ & 2.6605 & 1.8237 \\ & & 2.4710 \end{bmatrix},$$

Test whether the three stimulus condition effects are equal. If the hypothesis is rejected, determine which condition effects are significantly different. (20)

5. Following are the S.S. and S.P. matrices obtained from samples of sizes 50 each from the populations

$$\begin{pmatrix} 13.0552 & 4.1740 & 8.9620 & 2.7332 \\ & 4.8250 & 4.0500 & 2.0190 \\ & & 10.8200 & 3.5820 \\ & & & 1.9162 \end{pmatrix},$$

$$\begin{pmatrix} 6.0882 & 4.8616 & .8014 & .5062 \\ & 7.0408 & .5732 & .4556 \\ & & 1.4778 & .2974 \\ & & & .5482 \end{pmatrix}$$

Test the hypothesis of equality of two dispersion matrices at 5% level. (25)

GROUP H : PROBABILITY THEORY

SPECIAL TOPICS - Special Paper III

(No candidate available)
