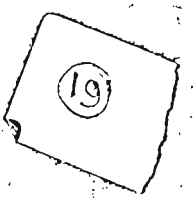
INDIAN STATISTICAL INSTITUTE

QUESTION PAPERS

for

Statistician's Diploma Examinations

May & November 1962



Price Rupee one

INDIAN STATISTICAL INSTITUTE

STATISTICIAN'S DIPLOMA EXAMINATION—MAY 1962

PAPER I: OFFICIAL STATISTICS AND DESCRIPTIVE STATISTICS (THEORETICAL)

Time: 4 hours Full marks: 100

Figures in the margin indicate full marks.

GROUP A

(Answer *any three* questions from this group)

1. Give an account of the various statistical reporting systems existing in India, listing the different types of statistics collected under each. Do you consider that administrative statistics have been neglected in this country? Illustrate your answer by examples. (10+6)

2. Give an account of any recent sample survey in India with which you are familiar, covering the purpose of the survey, sampling design (in detail), nature and quality of data collected. Suggest in what way the survey plan could have been improved? (12+4)

3. Describe the set-up of the statistical offices of the United Nations and its specialized agencies. (8)

Give a list of the standards which they have brought out for the purpose of International comparability of statistics. (4)

Discuss the usefulness of these standards from the point of view of under-developed countries. (4)

4. Discuss critically the present position of national and regional income statistics in the country in the light of improvements which have been effected during the past decade of planning. (12)

What further steps are needed for eliminating the existing deficiencies? (4)

Neatness. (2)

GROUP B

(Answer *any three* questions from this group)

5. Write down the frequency laws of the Binomial, Poisson and Negative Binomial distributions. Give examples of distributions arising in practice, which conform to these laws.

On an average, there is one misprint per page in a book of 2000 pages. Estimate the number of pages in the book that are completely free of misprints. Justify your estimate. (16)

6. Explain the concept of statistical dependence of one characteristic on another. How do the product moment Correlation Coefficient and the Correlation ratio measure this dependence?

Given that the linear regression of Y on X is $Y=5-2X$ and that of X on Y is $X=3-\frac{1}{4}Y$, find out the means of X and Y, and the Coefficient of Correlation between them. (16)

7. Why are time series decomposed into Components? What does the trend in a time series stand for? Describe the different methods available for determining trend in time series. (16)

8. In the distribution of diameters of a large number of spherical balls, the

mean = 50 mm.
median = 48 mm.
second central moment = 1 $(mm)^2$.
third central moment = 2 $(mm)^3$.

The weight w in mg. of a ball of diameter d mm. is given by the formula $w=5d^3$. Find the mean and median weights of the balls. (16)

Neatness. (2)

PAPER II : PROBABILITY THEORY AND STATISTICAL METHODS (THEORETICAL)

Time : 4 hours Full marks : 100

Figures in the margin indicate full marks.

GROUP A

(Answer *any three* questions from this group)

1. (*a*) Describe Kolmogorov's axiomatic approach to the notion of 'probability'. (6)

(*b*) A coin is tossed until for the first time the same result appears twice in succession. To every possible outcome requiring n tosses attribute probability $\frac{1}{2^n}$. Describe the sample space. (3)

Find the probability of the following events :

(*i*) The experiment ends before the fifth toss. (3)

(*ii*) An even number of tosses is required. (4)

2. Define Student's t and obtain its sampling distribution, under the usual assumptions.

Obtain the variance of the t-distribution (for d.f.>4), and the kurtosis coefficient γ_2. Find limiting values of the expressions you obtain when the number of degrees of freedom tends to infinity, and comment on the results. (8+8)

3. Let the probability p_n that a family has exactly n children be αp^n when $n \geq 1$ and $p_0 = 1 - \frac{\alpha p}{1-p}$. If the probability that a child may be a boy is half for any family, show that, for $k \geq 1$, the probability that a family has exactly k boys is $\frac{2\alpha p^k}{(2-p)^{k+1}}$. Given that a family includes at least one boy, what is the probability that there are two or more boys? (12+4)

4. Define (*i*) variance and (*ii*) median of a random variable. When does the variance of a random variable vanish? Let X and Y be independent random variables with finite variance. Then show that the variance of $X+Y$ is equal to the sum of the variances of X and Y. If $X+Y$ takes the constant value C with probability one show that X and Y are constant with probability one. If X has a density function derive the distribution of the median based on a sample of size $2n+1$. (1+1+2+5+7)

5. Define the moment generating function of a random variable. Let X and Y be independent random variables with moment generating functions $\phi(t)$ and $\psi(t)$. Prove that the moment generating functions of $X+Y$, $c\,X$ and $X-a$ are $\phi(t)\,\psi(t)$, $\phi(ct)$ and $e^{-ta}\,\phi(t)$ respectively, where c and a are some constants.

Evaluate the moment generating function of a Poisson random variable with parameter λ. If $X_1, X_2, \ldots, X_n$ are n independent Poisson variates with parameter λ, what is the moment generating function of $\frac{X_1 + \ldots + X_n - n\lambda}{\sqrt{n\lambda}}$. What is the limit of this moment generating function as $n \to \infty$? (1+1+1+1+3+3+6)

Neatness (2)

GROUP B

(Answer *any three* questions from this group)

6. Explain the terms (*i*) biased and unbiased estimators (*ii*) sampling distribution of an estimator (*iii*) standard error of an estimator.

Obtain the standard error of the mean in random samples of any size from a population. Obtain the standard error of the median in large samples from a population.

Hence show that in large samples from a normal population with mean 0 the sample-mean is more efficient than the sample-median as an estimator of 0. (6+2+4+4)

7. Explain the terms (i) errors of the first and second kind (ii) level of significance (iii) most powerful critical region (iv) uniformly most powerful test (v) similar region.

Suppose you are to test the hypothesis that a coin is unbiased, against the alternative that it is more likely to show up a head. Describe a suitable test procedure based on 10 tosses of the coin and at an approximate level of 5%. Determine the power of this test when the probability of the head is $\frac{3}{4}$. (An approximate calculation is only required). (8+8)

8. Explain the following methods of estimation:

(*i*) method of maximum likelihood.

(*ii*) method of moments.

If $f(x, \alpha) = \frac{\alpha^\lambda}{\Gamma(\lambda)} x^{\lambda-1} e^{-\alpha x}$ $(x > 0, \alpha > 0)$

is a density function with an unknown parameter α and λ is a known positive constant show that the method of maximum likelihood and the method of moments give the same estimate $\frac{\lambda}{\bar{x}}$ for α on the basis of n independent observations, $\bar{x}$ being the sample mean. Show also that $\frac{n\lambda-1}{n\bar{x}}$ is an unbiased estimate of α. (8+8)

9. Write down the density function of the χ^2 distribution with n degrees of freedom and explain its usefulness in tests for goodness of fit.

A statistician wants to investigate the dependence between eye colour and hair colour among a group of students. He classified eye colour as Black, Brown and Blue and hair colour as Black, Brown and Golden. How should he collect and present the data? Give the method of analysis of the data with proper justifications. (8+8)

Neatness (2)

PAPER III : SAMPLE SURVEYS AND DESIGN & ANALYSIS OF EXPERIMENTS (THEORETICAL)

Time : 4 hours Full marks : 100

Figures in the margin indicate full marks.

GROUP A

(Answer *any three* questions from this group)

1. (*a*) Define 'random sampling numbers' and mention some of their important uses in statistics. (5)

(*b*) Describe the method adopted for construction of any two well known series of random sampling numbers. (5)

(*c*) Describe the various tests that are usually applied to test for randomness in a series of random sampling numbers. (6)

2. You are required to plan a sample survey to study the problem of 'indebtedness' among the rural agricultural population in India. Suggest a suitable plan touching briefly, but adequately, on each of the following points :

(*a*) frame (*b*) sampling units (*c*) method of sampling (*d*) estimation procedures. (10)

Draft a simple questionnaire that may be used in this connection. (6)

3. (*a*) What are pilot surveys and what are their uses? (6)

(*b*) What are cost and variance functions and how are they obtained from pilot surveys? (10)

4. State how non-sampling errors differ from sampling errors. (4)

What are interpenetrating samples? Explain their use and limitations in the control of non-sampling errors in large-scale sample surveys. (12)

Neatness (2)

GROUP B

(Answer *any three* questions from this group)

5. In a varietal experiment with 11 varieties, a design is constructed using 11 blocks of size 5 as under:

(1, 2, 3, 5, 8),	(2, 3, 4, 6, 9),	(3, 4, 5, 7, 10),	(4, 5, 6, 8, 11),
(5, 6, 7, 9, 1),	(6, 7, 8, 10, 2),	(7, 8, 9, 11, 3),	(8, 9, 10, 1, 4),
(9, 10, 11, 2, 5),	(10, 11, 1, 3, 6,),	(11, 1, 2, 4, 7).	

Identify the type of the design and find out the parametres of the design. Describe the method of analysis you would adopt, using only intra-block information. (16)

6. (*a*) Describe a randomised block experiment and derive its usual analysis under suitable assumptions. (10)

(*b*) Explain fully the part played by 'randomisation' in such experiments. (6)

7. (*a*) Discuss how, through 'analysis of covariance' one utilises information on an auxiliary variate to increase the sensitivity of designed experiments and indicate situations where such a procedure would be effective. (8)

(*b*) Draw up the blank analysis of covariance table (one auxiliary variate) for a randomised block experiments comparing 6 treatments in 4 blocks of 6 plots each, showing the degrees of freedom for the various components. (4)

(*c*) Suppose the treatments are the possible combinations of two levels of Factor A and three levels of Factor B. Write down the improved estimate for main effect of A after accounting for the regression of yield on the auxiliary variate. (4)

8. (*a*) Explain 'confounding', total and partial, in connection with factorial experiments. (6)

(*b*) You are required to plan a 2^5 factorial experiment in blocks of size 8, involving 5 factors A, B, C, D and E so that the second order interactions ABC and BCD are confounded. Write down the combinations appearing in each block separately. (8)

(*c*) Is any other d.f. so confounded? If so, what effect does it represent? (2)

Neatness (2)

PAPER IV : APPLIED STATISTICS (THEORETICAL)

Time : 4 hours | Full marks : 100

(*i*) Answer questions only from the two groups relating to the subjects you have opted for.

(*ii*) Use a separate answer book for each group.

(*iii*) Figures in the margin indicate full marks.

GROUP A—ECONOMIC STATISTICS

(Answer *any three* questions from this group)

1. If it is required to write a statistical note on the trend of prices of gold in India during post-independence years, mention briefly the basic sources of data and show how you would proceed to analyse them.

How would you correct the monthly series of data for seasonality? (6+10)

2. Give a short account of the nature and uses of the Cobb-Douglas production function. Describe how you would proceed to determine such functions for an agricultural crop, such as paddy in Andhra Pradesh. (6+10)

3. Explain clearly the concepts of national income at factor cost, national income at market prices, gross national product, personal income, disposable income and capital formation. (16)

4. *Either,*

Describe the main features of the index number of agricultural production published by the Ministry of Food and Agriculture, Government of India. (16)

Or,

You are required to compile an index number of cost of building construction in Calcutta. Describe the data you will require and the procedure you will adopt for the computation of the index. (16)

Neatness (2)

GROUP B—STATISTICAL QUALITY CONTROL

(Answer *any three* questions from this group)

5. (*a*) Given the values of Range in k samples each of size n, how will you obtain a point estimate of the process standard deviation? State the assumptions, if any, that are made in your method of estimation. (8)

(*b*) Explain how a $(1-\alpha)$ per cent confidence interval can be set up for the process standard deviation (*i*) using sample range (*ii*) using sample standard deviation. (8)

6. (*a*) Explain each of the different types of stipulations that are commonly placed on acceptance sampling plans for their construction. (8)

(*b*) Comment on the following statement :

"The US Military Standard Plans could not have been constructed without reference to the Dodge-Romig Sampling Inspection Tables". (8)

7. (*a*) Explain the procedures to be followed for setting up Control Charts on (*i*) averages and ranges (*ii*) fraction defective and (*iii*) number of defects, with varying sample size. Explain briefly the theoretical basis and practical use of having control limits on these charts. (10)

(*b*) How is the efficiency of a control chart measured? Determine the OC curve of a control chart of averages of samples of size 4, with 3 sigma control limits used for action, assuming that the standard deviation is known and constant. (6)

8. A factory has the option to purchase its raw material from 5 different sources to manufacture a certain product. The factory has 5 machines for the manufacture of the product, each machine being operated by one operator. The factory is interested in conducting an experiment to determine which raw material gives best results.

Suggest a suitable design for the experiment. Using the usual notations, give the estimates of the components of variance. (16)

Neatness (2)

GROUP C—STATISTICAL METHODS IN GENETICS

(Answer *any two* questions from this group)

9. Explain how linkage between two factors can be detected and measured.

Derive the expected frequencies for the four phenotypic classes into which the offsprings of the matings (*i*) Ab/aB × Ab/aB and (*ii*) AB/ab × ab/ab can be classified.

Examine which of the two matings mentioned is more informative for the estimation of the linkage parameter. (24)

10. Determine the expected frequencies of the O, A, B and AB blood groups in a population under 'random mating'.

How do you estimate the gene frequencies from observed frequencies of the O, A, B, and AB classes?

Indicate how to detect whether the population is subject to some amount of inbreedings. (24)

11. Mention what data are needed and how to compute a discriminant function for selection of plants for a given characteristic like the yield.

How is genetic advance due to the use of discriminant function measured. (24)

Neatness (2)

GROUP D—VITAL STATISTICS AND DEMOGRAPHY

(Answer *any three* questions from this group)

12. Define incidence and prevalence rates of sickness. Discuss the respective roles of general health surveys, case finding (diagnostic) surveys, and hospital records as sources of sickness data. Discuss briefly the results of any studies which have been made to assess the effect of (*i*) proxy interviews, and (*ii*) length of recall period on the responses in a general health survey? (16)

13. Define stationary and stable populations. How can one construct Life Tables from vital statistics and census returns? What are the main adjustments necessary in the pirmary data? (16)

14. Comment on the following statements (made in connection with some demographic surveys): (16)

(*i*) "It is .... clear that the survey material is mainly of use not so much for calculation of overall figures or rates as for detailed analysis of varying situations and conditions; that is, differences between regions, classes and strata in society. For the overall figures reliance will continue to be placed, in the main, on such data as that of the census and the birth and death registration statistics."

(*ii*) "For the rural investigation, ... in selecting the villages, an attempt was made to select villages with population not more than 1500.... The investigators were given instruction to select a sample of 15 to 30 households in each selected village so as to cover all sections of the village site."

(*iii*) "Among males in the labour force 86.35 per cent belong to the working age group (16—61), 10.09 per cent are boys (7—15 years), and only 5.35 per cent are old (62 and above). Among females in the labour force 85.73 per cent belong to the working age group, 12.42 per cent are girls of tender age and only 1.63 per cent are old. Participation of females in the labour force thus starts somewhat earlier in life and also terminates earlier, compared to the males."

15. Discuss briefly the Indian Census operations in 1961, pointing out the differences between this and the previous census.

If you are required to forecast the population by age × sex groups in India in the year 1966, how would you proceed? (16)

Neatness (2)

GROUP E—EDUCATIONAL AND PSYCHOLOGICAL STATISTICS

(Answer *any three* questions from this group)

16. Describe the method of transforming raw scores on tests for a group of students into (*i*) standard or z-scores, (*ii*) McCall's T-scores and (*iii*) Percentile scores. Indicate the situations in which one or more of them have useful interpretation.

Why is the present system of comparing and combining the examination scores considered as faulty and how could one improve upon it using percentile scores of students? (10+6)

17. Define biserial and tetrachoric correlations and describe the situations in which they are used. Explain the concept of rank correlation and justify its use with psychological data in preference to product moment correlation. Describe the suitability of Kendall's T-Coefficient as a measure of rank correlation stating its advantages over Spearman's ρ. (5+5+6)

18. Define 'reliability' and 'validity of tests'. Explain reliability in terms of true error and observed variances of test scores. (4)

Obtain the formula for reliability in terms of test length. (4)

How would you match different test items of a battery to get a unique value for the reliability Coefficient. (4)

Why is 'split-half' method of finding reliability not suitable for speed tests? (4)

19. Define parallel tests. (5)

Show that $r_{xl} = \frac{1}{K\sqrt{r_{gh}}} + \left(1 - \frac{1}{K}\right)\sqrt{r_{gh}}$, where the symbols have the usual meaning. (5)

How do you examine whether three or more given tests are parallel? (6)

Neatness (2)

PAPER V: METHODS OF NUMERICAL COMPUTATION, DESCRIPTIVE STATISTICS AND OFFICIAL STATISTICS (PRACTICAL)

Time : 5 hours Full marks : 100

(i) Figures in the margin indicate full marks.
(ii) Use of calculating machines is permitted.

1. (a) Values of $y=f(x)$ are given for six values of x. Find by a suitable method of interpolation the value of y for $x=0.7$ correct to six decimal places. Give reasons for the choice of the formula of interpolation used by you. (12)

x	0.2	0.4	0.6	0.8	1.0	1.2
y	.5792597	.6554217	.7257469	.7881446	.8413447	.8849303

(b) Find by numerical method of differentiation the value of $\frac{d^2y}{dx^2}$ at $x=0.7$ for the function in (a) above. How many significant digits are there in your results? (5)

(c) Solve the following equations in x, y, z:

$$63 = 32x+126y+106z$$
$$251 = 126x+501y+413z$$
$$225 = 106x+413y+368z$$ (8)

2. *Either,*

The frequency distribution of monthly income (in Rs.) of a sample if middle class families in Calcutta are given below. Fit a normal curve to the distribution and test the goodness of fit. (20)

monthly income (Rs.)	number of families	monthly income (Rs.)	number of families
0— 75	15	300— 500	309
75—100	51	500— 750	162
100—150	199	750—1000	66
150—200	240	1000—1500	50
200—300	324	1500—1800	27
		Total	1443

Or,

(*a*) In the following table, X denotes the deviation of mid-point of a class interval from an arbitrary origin expressed in terms of constant width h of class-interval as unit and f the frequency of the class-interval.

X	−3	−2	−1	0	1	2	3	4
f	3	15	45	57	50	36	25	9

Find h and the boundaries of all the class-intervals, if the mean and standard deviation (ignoring Sheppard's correlation) are respectively 40.004 and 7.92. (*N.B.*—Square of standard deviation is equal to the sum of the squares of deviation from the mean divided by the total number of observations). (10)

(*b*) The mean of a normally distributed variable X is 2.08 and 9.68% of the values of X are negative. Find the probability that a value of X taken at random will be greater than 5.0. If the total frequency is 1250, draw the frequency curve by calculating a suitable number of ordinates. (10)

3. The following data relate to average values of land cultivated in acres (X_1), land owned in acres (X_2), and size of family (X_3):

land cultivated in acres (X_1)	land owned in acres (X_2)	size of families (X_3)
3.44	2.33	5.13
4.27	3.30	5.70
5.07	4.14	6.14
5.80	5.28	6.40
7.54	6.36	6.84
7.17	7.28	7.06
8.38	8.26	7.58
8.68	9.26	7.31
9.87	10.12	8.71
10.70	11.35	8.01
11.99	13.34	8.12
12.33	14.26	9.24
15.45	15.64	9.68
13.05	17.24	8.48
15.87	18.86	9.04
17.29	19.57	10.20
16.37	20.59	10.37
19.65	16.24	13.61

Determine the regression equation of X_1 on X_2 and X_3 and calculate the corresponding multiple correlation coefficient. (15)

4. From the following data relating to middle class Bengalees in Calcutta in the income group Rs. 101—150 per month, calculate the cost of living index for 1949 with 1939 as base year. (15)

item	expenditure on the item as % of the total expenditure	unit	price in Rs.		price index (base year 1939)
			1939	1949	
Food					
Rice	15.41	seer	0.14	0.50	
Wheat products	4.53	,,	0.13	0.53	
Chira-muri	0.61	,,	0.14	1.13	
Pulses	2.89	,,	0.18	0.60	
Fish	6.00	,,	0.57	1.99	
Meat	2.08	,,	0.60	2.75	
Egg	0.38	20	0.53	1.85	
Milk	4.80	seer	0.19	0.79	
Butter and Ghee	2.60	,,	1.59	5.25	
Oil	4.72	,,	0.48	2.14	
Potato	3.48	,,	0.12	0.50	
Other vegetables	5.74	,,	0.10	0.35	
Salt	0.42	,,	0.06	0.09	
Spices	1.62	,,	0.60	2.74	
Sugar and Gur	2.30	,,	0.29	0.70	
Tea and Tiffin	4.71	—	0.69	2.76	
Rent and Tax	9.81	—	—	—	110.00
Fuel and Light					
Coal	3.46	Md.	0.38	1.69	
Other fuel	3.18	—	—	—	364.65
Clothing	6.43	—	—	—	472.64
Miscellaneous					
Intoxicants	2.56	—	—	—	250.00
Toilet	1.13	—	—	—	300.00
Education	1.83	—	—	—	150.00
Medicine	1.89	—	—	—	250.00
Other items	7.33	—	—	—	300.0

5. *Either,*

(*a*) Name one publication in which each of the following series is available and also mention the issuing authority and the periodicity, of the publication. (9)

(*i*) Index number of employment in manufacturing by different countries of the world.

(*ii*) Percentage absenteeism in mills of Bombay and Calcutta.

(*iii*) Production, despatch and stock of coal in Bihar.

(*b*) From the publications supplied to you, compile the following data in a neat tabular form with proper headings and spacing for any six consecutive months, with footnotes if necessary. (16)

(*i*) Value of imports of merchandise into India by sea, air and land.

(*ii*) Value and quantum index numbers of imports into and exports from India, of food items by sea, air and land.

(*iii*) Working class index number of consumer prices.

Mention the source from which you have compiled the data.

Or,

(*a*) Collect relevant information from the publications supplied and write a note on the changes in the per-capita national income in India since 1948-49. (13)

(*b*) From the official publications supplied collect data on index number of agricultural production in India showing separately figures for food grains, non-food grains and all commodities for as many years as possible and comment on the salient features of the data. (12)

PAPER VI: STATISTICAL METHODS, DESIGN & ANALYSIS OF EXPERIMENTS AND SAMPLE SURVEYS (PRACTICAL)

Time : 5 hours Full marks : 100

(*i*) Figures in the margin indicate full marks.

(*ii*) Use of calculating machines is permitted.

GROUP A

(Answer *any two* questions from this group)

1. The following data give the gains in weight (in suitable unit) of 10 matched pairs of rats. One rat from each pair received its protein from raw peanuts while the other received its protein from roasted peanuts. Test to see whether roasting the peanuts had any effect on their protein value. (20)

raw	61	60	56	63	56	63	59	56	44	61
roasted	55	54	47	59	51	61	57	54	62	58

2. 878 professional scientists were asked about their opinions (Agree/Disagree) on each of the following questions :

x_1 : The development of new ideas is the scientists' greatest source of satisfaction.

x_2 : The scientist will make his maximum contribution to society when he has freedom to work on problems which interest him.

x_3 : The neglect of basic scientific research would be the equivalent of "killing the goose that laid the golden eggs".

The frequency distribution of responses were as follows:

response			frequency
x_1	x_2	x_3	
0	0	0	605
0	0	1	65
0	1	0	70
0	1	1	22
1	0	0	72
1	0	1	18
1	1	0	21
1	1	1	5

0 = Agree

1 = Disagree

Test if there is any association between opinions expressed on these three questions. (20)

3. The following table gives the variances and covariances for scores in two tests based on a sample of 50 examinees. Test if the correlation between the two tests is significantly different from zero. If not, set up a 95% confidence interval for it.

dispersion matrix

tests	A	B
A	25.0704	12.4363
B		28.2021

(20)

4.

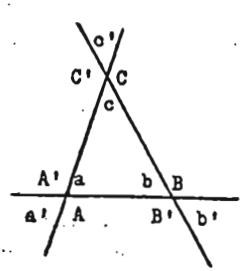


Fig.

Independent measurements on the angles of the triangle above are given below.

$a = 59.1$	$a' = 58.6$	$A = 120.5$	$A' = 122.1$
$b = 60.4$	$b' = 61.3$	$B = 119.8$	$B' = 118.7$
$c = 60.1$	$c' = 59.2$	$C = 120.7$	$C' = 121.5$

Making use of geometrical properties known, estimate by the method of least squares the internal angles of the triangle (Fig.) and also the precision of measurement. (20)

GROUP B

(Answer *any two* questions from this group)

5. Four varieties of paddy were compared in five randomised blocks of four plots each. The yields were observed as follows:

block	variety			
	A	*B*	*C*	*D*
1	56.8	55.0	48.0	52.3
2	54.5	56.5	48.8	55.8
3	53.3	52.3	49.3	50.8
4	53.0	54.0	46.0	54.3
5	56.3	54.3	49.8	55.3

(*a*) Set up the analysis of variance table and demonstrate the significance of the treatment mean square. (10)

(*b*) It is suspected that the variation between treatments is primarily due to the small yield of variety C. Test if this is supported by the above data. (5)

6. An experiment was performed to test four materials *A*, *B*, *C* and *D* on each of four runs of a machine with four positions. A Latin Square design was chosen. The numbers in the table denote the loss of weight of the material and the letters *A* to *D* refer to the four materials. The loss of weight of the material *B* on the fourth run of the machine in the fourth position could not be determined.

run	position in the machine			
	4	2	1	3
2	A (251)	B (241)	D (227)	C (229)
3	D (234)	C (273)	A (274)	B (226)
1	C (235)	D (236)	B (218)	A (268)
4		A (270)	C (230)	D (225)

Analyse the experiment and make critical comments on the basis of statistical tests of significance. (15)

7. The table below gives the plan and yields in lbs per plot of a 3^3 experiment on turmeric laid out in 3 randomised blocks of nine plots each, the treatments being all combinations of

n_0 = no nitrogen	p_0 = no super phosphate	k_0 = no potash
n_1 = 60 lbs N/acre	p_1 = 45 lbs P_2O_5/acre	k_1 = 100 lbs K_2O/acre
n_2 = 120 lbs N/acre	p_2 = 90 lbs P_2O_5/acre	k_2 = 200 lbs K_2O/acre

Plan and yields of raw turmeric in lbs per plot.

Block I	$n_1 p_1 k_2$ 114	$n_0 p_2 k_2$ 90	$n_2 p_1 k_0$ 76	$n_0 p_1 k_1$ 68
	$n_0 p_0 k_0$ 92	$n_1 p_2 k_0$ 105	$n_2 p_2 k_1$ 95	$n_1 p_0 k_1$ 91	$n_2 p_0 k_2$ 81
Block II	$n_1 p_0 k_2$ 79	$n_2 p_0 k_0$ 76	$n_0 p_0 k_1$ 57	$n_2 p_2 k_2$ 97
	$n_1 p_2 k_1$ 88	$n_0 p_1 k_2$ 75	$n_0 p_2 k_0$ 71	$n_1 p_1 k_0$ 76	$n_2 p_1 k_1$ 76
Block III	$n_0 p_2 k_1$ 73	$n_1 p_0 k_0$ 74	$n_2 p_0 k_1$ 66	$n_2 p_1 k_2$ 106
	$n_0 p_0 k_2$ 66	$n_2 p_2 k_0$ 107	$n_1 p_2 k_2$ 88	$n_0 p_1 k_0$ 65	$n_1 p_1 k_1$ 84

Analyse the data and state your conclusions. (15)

GROUP C

(Answer *any two* questions from this group)

8. The following table gives the stratification of all tea estates by registered acreage under tea and the average and the standard deviation of the number of workers on books.

registered acreage under tea	number of tea estates	number of workers on books	
		average	s.d.
0— 50	312	40	2.3
51—100	185	75	4.1
101—200	117	180	8.7
201—400	86	351	15.3
401—600	75	518	22.9
601—800	71	721	29.1
801—1000	68	837	36.8
Over 1000	31	1215	50.3

It is desired to choose a sample of 50 estates to estimate the total number of tea workers in the region. Compare the precision of the following three methods of sample selection:

(a) an unrestricted random sample. (5)

(b) stratified random sample—proportional allocation. (5)

(c) stratified random sample—Neyman's optimum selection. (5)

9. The following table gives the total area under cultivation x, and the yield of paddy y for 30 villages in a tehsil.

village	area in acres (x)	yield of paddy in maunds (y)	village	area in acres (x)	yield of paddy in maunds (y)
1	870	8521	16	1224	11913
2	883	8554	17	1229	11517
3	894	8783	18	1260	11955
4	901	8862	19	1260	11639
5	914	7025	20	1302	12428
6	973	8887	21	1319	12120
7	995	9569	22	1350	12172
8	1031	9598	23	1415	13221
9	1043	10316	24	1461	13040
10	1054	8963	25	1476	12103
11	1089	9502	26	1536	9560
12	1111	10512	27	1582	14804
13	1123	10874	28	1591	14818
14	1177	11560	29	1601	15772
15	1203	11031	30	1650	15023

(a) Draw a sample of size 6 with replacement with probabilities proportional to area. Obtain an unbiased estimate to the total yield of paddy and an unbiased estimate of the variance of the estimator. (7)

(b) Divide the villages into two strata of nearby equal total area, grouping the villages in their serial order. Draw samples of size 3 each, from each stratum with equal probabilities with replacement. Compute an unbiased estimate of the total yield of paddy and an unbiased estimate of the variance of the estimator. (8)

10. In a stratified sample obtained to estimate the crop in a region, it is given that the variance V of the estimate and the total cost W are of the following forms:

$$V = \sum_i \frac{A_i^2 a_i}{n_i x_i g_i}$$

$$W = b + \sum_i n_i (c_i + d_i x_i)$$

Where A_i = geographical area,
b = Rs. 10000.00,
n_i = number of sample units

and x_i = size of sample unit in the ith stratum and other letters denote some constants. The number of strata is 2 and the numerical values of constants are given below.

	A_i (sq. mile)	a_i	g_i	c_i (Rs.)	d_i (Rs.)
1	7000	0.09	0.4	30	3.50
2	9000	0.12	0.5	36	4.00

Find out the optimum values of n_i and x_i for W = Rs. 50,000.00 (15)

PAPER VII : APPLIED STATISTICS (PRACTICAL)

Time : 5 hours Full marks : 100

(i) Answer questions only from the two groups relating to the subjects you have opted for.
(ii) Use a separate answer book for each Group.
(iii) Figures in the margin indicate full marks.
(iv) Use of calculating machines is permitted.

GROUP A—ECONOMIC STATISTICS

(Answer *any two* questions from this group)

1. The following table gives some basic time-series data relating to Indian manufactures during the period 1947–1957. These data include (*a*) number of persons employed, (*b*) total salaries and wages paid, (*c*) total material and fuel consumption and depreciation and (*d*) total of products and by-products manufactured for sale.

Compute the following for each year and interprot your results. You may draw rough graphs for this purpose.

(i) value added by manufacture.
(ii) value added per person employed.
(iii) non-labour payments.
(iv) non-labar payments as percentage of value added. (25)

description	1947	1948	1949	1950	1951	1952	1953	1954	1955	1956	1957
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Number of persons employed (thousand)	1633	1704	1685	1632	1633	1648	1628	1715	1784	1886	1895
Total salaries and wages paid (Rs. lakhs)	13576	16582	17719	17217	18918	20075	20521	21856	23114	25580	26480
Total material and fuel consumed and depreciation (Rs. lakhs)	50139	63631	70338	74408	95965	86894	78852	91463	96865	114555	125688
Total of products and by-products for sale (Rs. lakhs)	74361	95365	97607	102801	130686	118397	111278	128754	140610	161429	172380

Basic source : Summary tables for the years 1947–55 in the Report of the "Census of Indian Manufactures–1955", Reports of the Census of Indian Manufactures 1956 and 1957.

2. The following table gives the monthly statistics of production of sugar in India during 1957—1961. Calculate the indices of seasonal variation by the method of moving average. (25)

TABLE 4. MONTHLY STATISTICS OF PRODUCTION OF SUGAR IN INDIA DURING 1957—1961*

(thousand tons**)

month	year				
	1957	1958	1959	1960	1961
(1)	(2)	(3)	(4)	(5)	(6)
January	361.4	429.8	444.6	498.3	541.3
February	355.7	411.5	411.2	473.5	501.8
March	375.8	377.6	366.5	443.9	505.5
April	277.0	197.6	174.3	265.3	416.9
May	153.1	51.8	32.3	76.5	208.4
June	39.3	3.6	0.2	2.9	43.1
July	9.9	4.3	4.5	5.3	8.3
August	7.6	7.2	9.3	10.1	14.6
September	8.2	10.6	10.3	16.2	17.1
October	17.4	18.0	20.0	39.0	—
November	136.2	89.8	179.3	239.0	—
December	358.0	355.2	430.8	486.9	—

*Figures are in respect of cane sugar only.

**From September 1960 onwards production figures are in Thousand Matric Tonnes (1 Ton = 1.02 Metric Tonnes).

Source : Monthly Statistics of the Production of Selected Industries of India, 1957—1961.

3. Fit a parabola by the method of least squares to the following data on export of gasoline from the United States and plot the two curves for observed and graduated values, on the same sheet of graph paper. (25)

years	exports of gasoline (millions of barrels)
1916	8.5
1917	9.9
1918	13.3
1919	8.9
1920	15.3
1921	12.7
1922	13.8
1923	20.1
1924	28.3
1925	30.6
1926	42.5
1927	44.3
1928	52.9
1929	62.1
1930	65.6

(25)

4. The distribution of Central Government employees by various pay ranges and the percentage distribution of their pay in different groups are given below for 1953 and 1958. Find out the values of Lorenz ratio for the two periods. Find out by linear interpolation the ordinate values of the Lorenz Curves corresponding to the lower 90 per cent and 95 per cent of the employees ranked according to pay. (25)

DISTRIBUTION OF CENTRAL GOVERNMENT EMPLOYEES AND THEIR PAY

pay ranges (Rs.)	number of employees		percentage of pay	
	1953	1958	1953	1958
Below 51	814995	1009260	30.1	28.1
51—100	420839	520291	29.5	27.5
101—150	103218	168369	12.0	14.7
151—200	53258	67841	8.7	8.3
201—250	20243	28564	4.2	4.5
251—300	9741	16350	2.5	3.1
301—350	6814	9958	2.0	2.3
351—500	8698	11904	3.4	3.5
501—750	4719	6644	2.7	2.9
751—1000	2115	3149	1.7	1.9
1001—1500	1143	1793	1.3	1.6
1501 and above	937	1148	1.9	1.6

GROUP B—STATISTICAL QUALITY CONTROL

5. (*a*) Obtain 95% confidence intervals:

(*i*) for the proportion defective in the process, given that in a sample of 200 items 13 defectives were found. (3)

(*ii*) for the average number of defects per yard in the process, given that on 15 pieces of cloth, each five yards long, a total of 32 defects were noticed. (3)

(*iii*) for the process mean, given that in a sample of 25 items the mean was 12.32 units and the range 2.51 units. (3)

(*b*) Obtain the values necessary for drawing the control chart lines:

(*i*) in an $\bar{x}$-chart, given that in 4 samples each of size 5, the means were 4.8, 5.3, 4.7, 4.9 and the ranges were 1.9, 1.3, 2.0, 1.8. (3)

(*ii*) in an s-chart, given as a standard the value 3.2 for the standard deviation, and given that the sample size is either 6 or 5. (3)

(*iii*) in a u-chart (defects per unit), given that the number of units per sub-group is 15 or 20 and that the average number of defects per unit based on 8 sub-groups is 1.4. (3)

(*c*) From the Dodge-Romig Sampling Inspection Tables, select plans to satisfy:

(*i*) lot size: 3400; process average: 1.30 per cent; A O Q L: 2.50 per cent; type of sampling: double. (3)

(*ii*) lot size : 12000; L T P D : 2.5 per cent; process average : 0.7 per cent; type of sampling : single. (3)

(*d*) For each of the plans in (*c*) above, obtain the approximate probabilities of acceptance (OC) at the following values of incoming quality : 0.5, 1.0, 1.5, 2.0, 5.0, 10.0 and 15.0 per cent. (6)

6. *Either,*

Construct a chart for recording the observations of a sequential sampling plan for testing $p_0 = 0.3$ against $p_1 = 0.4$ with $\alpha = 2$ per cent and $\beta = 1$ per cent. Sketch the ASN Curve for this test. (20)

Or,

The following table gives the results of an experiment conducted by a factory to study the effect of three factors (method of processing, temperature and raw material) on the yield of a process. The process was carried out by each of the two methods M_1 and M_2; the main stage was carried out at each of the three temperatures T_1, T_2 and T_3; four batches of raw material B_1, B_2, B_3 and B_4 were used. Two trials were carried out with each set.

Analyse the data on the basis of this experiment. What specific suggestions can you make to increase the yield? (20)

	B_1		B_2		B_3		B_4	
	M_1	M_2	M_1	M_2	M_1	M_2	M_1	M_2
T_1	76.1	80.5	68.6	73.2	67.1	80.2	74.4	77.5
	74.0	79.3	69.4	76.0	71.4	72.9	78.6	75.3
T_2	74.0	80.9	70.0	74.4	77.6	76.8	70.2	75.6
	76.1	80.2	73.0	79.3	75.4	80.2	74.5	80.3
T_3	74.0	78.2	74.8	81.0	73.6	80.9	78.2	85.8
	82.9	83.8	74.9	79.0	72.8	77.8	77.6	79.6

GROUP C—STATISTICAL METHODS IN GENETICS

7. The following data relate to 3 back cross tests for the genes *A* and *B*.

test	frequencies			
	AB	Ab	aB	ab
I	58	72	105	67
II	53	81	75	45
III	68	73	75	40

(*i*) Examine whether the proportions in the four classes are significantly different in the three tests.

(*ii*) Examine the data for Linkage between A and B.

(*iii*) Find the confidence limits to the linkage factor. (30)

8. The frequencies of the blood groups MM, MN and NN in two samples from two populations A and B are as follows:

genotype	probability	frequencies	
		A	B
MM	m^2	36	49
MN	$2mn$	52	42
NN	n^2	16	15

Examine whether the gene frequencies of m and n are the same in both the populations.

(20)

GROUP D—VITAL STATISTICS AND DEMOGRAPHY

(Answer *any two* questions from this group)

9. Fit a logistic or any other suitable curve to the following Indian Census data and from the fitted curve estimate the population in 1971 and 2001 A.D. (25)

Census year	1891	1901	1911	1921	1931	1941	1951	1961
Population (million)	235.9	233.0	250.7	249.0	277.4	316.9	359.2	438.0

10. From the following data, calculate the gross and net reproduction rates.

You can assume a sex-rates of 105.8 male per 100 female births. (25)

age group (years)	number of females	number of live births	proportion surviving from birth to midpoint of age group
(1)	(2)	(3)	(4)
10—14	19561	8	.96928
15—19	17423	1220	.96683
20—24	15767	3002	.96338
25—29	14230	2349	.95915
30—34	12352	1267	.95381
35—39	11259	587	.94658
40—44	8254	120	.93569
45—49	6050	6	.91912

11. From the following data, calculate the working life expectancy at ages 15, 20, 25, 35, 45 and 65 for Indian males. The working life expectancy is defined as

$$e_x^{10} = \frac{\Sigma\left[\left(\sum_x^{x+m} L_x\right) p_x : x+m\right]}{l_x}$$

where l_x = number of persons living at age x,

L_x = number of persons living between ages x and $x+1$,

$p_x : x + m$ = labour force proportion for the age group x to $x+m$.

The summation outside the square brackets extends from the base age x to the end of life. (25)

age (x)	l_x	T_x	$p_x : x+m$
(1)	(2)	(3)	(4)
15	87,138	3,157,529	.5704
20	82,748	2,733,007	.7023
25	78,258	2,330,363	.7550
35	67,968	1,597,394	.7669
45	55,512	978,394	.6527
65	22,458	183,609	.3164

Note that $\sum_{x}^{x+m} L_x = T_x - T_{x+m}$

GROUP E—EDUCATIONAL AND PSYCHOLOGICAL STATISTICS

(Answer *any two* questions from this group)

12. Frequency distributions of marks obtained by two groups of candidates *A* and *B*, in two major vernaculars are given below. Make a comparative study of achievements in Bengali and Hindi. (25)

marks	bengali		hindi	
	I	II	I	II
0— 4	—	—	—	—
5— 9	—	4	—	—
10—14	3	7	—	—
15—19	17	24	—	—
20—24	58	68	3	—
25—29	133	152	—	2
30—34	202	179	3	3
35—39	802	828	20	22
40—44	892	718	27	26
45—49	687	628	19	12
50—54	466	462	11	17
55—59	164	275	5	8
60—64	79	119	8	5
65—69	11	44	2	2
70—74	13	17	—	1
75—79	1	3	—	—

13. The following table gives the reduced correlation matrix of sex tests. Assuming that the number of common factors is 2, evaluate the factor loadings by any suitable method. (25)

$$R = \begin{vmatrix} h_1^2 & 0.72 & 0.75 & 0.49 & 0.42 & 0.28 \\ — & h_2^2 & 0.78 & 0.42 & 0.36 & 0.24 \\ — & — & h_3^2 & 0.35 & 0.30 & 0.20 \\ — & — & — & h_4^2 & 0.42 & 0.28 \\ — & — & — & — & h_5^2 & 0.24 \\ — & — & — & — & — & h_6^2 \end{vmatrix}$$

14. From the Item-score matrix given below for a 10-item test taken by 8 individuals, find—

(a) Item variances and their sum.

(b) Indivdual score means and the variance.

(c) Apply Kudor-Richardson Formula—20 to estimate the Reliability of the test. (25)

MATRIX OF ITEM-SCORES FOR A 10-ITEM TEST TAKEN BY 8 INDIVIDUALS

individuals	items									
	a	*b*	*c*	*d*	*e*	*f*	*g*	*h*	*i*	*j*
1	0	0	0	0	0	0	0	0	0	0
2	1	0	0	1	0	0	0	0	0	0
3	1	1	1	0	1	0	0	0	0	0
4	1	1	1	1	0	0	1	0	0	0
5	1	1	1	0	1	1	0	0	0	0
6	1	1	1	1	0	1	0	1	1	0
7	1	1	1	1	1	1	1	1	0	0
8	1	1	1	1	1	1	1	1	1	1

PAPER VIII : SUBJECTS OF SPECIALISATION—I

Time : 4 hours Full marks : 100

(*i*) Answer questions only from the section on the subject of spcialisation you have opted for.

(*ii*) Figures in the margin indicate full marks.

(a) ECONOMIC STATISTICS—ECONOMETRICS

(Answer *any four* questions from this section)

1. Let E^n be an operator which when applied to a variable displaces it by one time unit, that is $E^n x_t = x_{t+n}$ and $E^{-n} x_t = x_{t-n}$. With this notation, a simple aggregative model of an economy is given by the equations,

$$x_1 = (1 - aE^{-1})\, x_3$$
$$x_2 = (b\, E^{-1} + c\, E^{-2})\, x_3$$
$$x_4 = x_2$$

where x_2, x_3, x_4 respectively stand for asset formation, national income and saving, expressed as deviations from the means.

Are the parameters a, b, c identifiable? Give reasons. (12)

Obtain the fundamental autoregressive equation of the system. (7)

Discuss the pattern of growth of national income given by the autoregressive equation,

$$(1-1.6E^{-1}+0.6\ E^{-2})x_3 = 0. \tag{6}$$

2. Derive Cobb-Douglas production function from the basic assumptions and discuss its properties. What kind of data would you use to estimate the parameters of a Cobb-Douglas production function? (15)

J. R. N. Stone has used a production function,

$$y = y^* - \beta\,(a - a^*)$$

where y and a respectively denote output and real assets per head and the symbols y^*, a^* and β denote the parameters. In particular, y^* denotes the highest attainable output per head which is reached when the actual amount of assets per head, a, is equal to a^*. If Y, A and n respectively stand for total output, total assets and population, show that

$$Y = \frac{\partial Y}{\partial n} \cdot n + \frac{\partial Y}{\partial A} \cdot A \tag{10}$$

3. Describe the methods used for deriving demand curves by making use of time series data, briefly indicating the nature of data utilised for the purpose. (17)

Obtain the price elasticity of demand of wheat from the demand curve,

$$p = \frac{2180}{D^2}$$

where p is price and D is the quantity demanded. (8)

4. Describe how national income can be conceived as an aggregate of either values added or factor shares or final expenditures. (15)

Briefly explain the notion of an input-output table, and show that the revenues of production accounts can be obtained as a product of a matrix multiplier and the vector of final purhases. (10)

5. Explain clearly the procedure of fractile graphical analysis and indicate some of the problems in which the procedure is appropriate. (25)

6. Given cross section data on household consumer expenditure, how would you fit the equations,

$$Y = a\,X^b, \text{ and } \quad Y = a\,\frac{X}{X+b}$$

where Y is expenditure on a particular item of consumption and X is total expenditure? (15)

What do you understand by the terms:

price elasticity, income elasticity, cross elasticity, arc elasticity? (10)

(b) TECHNO-COMMERCIAL STATISTICS—STATISTICAL QUALITY CONTROL

(Answer *any four* questions from this section)

1. Explaining clearly the basic assumptions, describe how the sample range and the sample standard deviation could each be used for assessing process variability?

In what circumstances would you advocate the use of range in place of standard deviation and why?

Indicate the uses of range on control charts and acceptance sampling. (25)

2. (a) Explain with illustrative examples how the control chart for averages and ranges could help interpret the quality of manufacture in relation to specification limits? (10)

(b) Examine the use of median for control of the average level of a measurable characteristic, in the place of the arithmatic mean. Briefly discuss its advantages and disadvantages? (15)

3. What are the circumstances in which you will recommend the use of each of these charts—group control charts, control charts for moving averages and ranges, control charts with sloping limits, and two way control charts. (12)

Sketch the steps in the installation and maintenance in a plant, of any one of these charts, with special reference to the detection of assignable causes of variation. (13)

4. (a) Describe with the help of a suitable chart or diagram the important factors to be considered in establishing a sampling inspection plan for lot acceptance. (15)

(b) Discuss the value of the control chart as a tool in acceptance inspection. (10)

5. (a) Beifly describe the statistical tests available to test for equality of a set of sample estimates of variance and the significance of the largest of a set of sample estimates of variance. Examine their uses in industrial quality control work. (15)

(b) Examine the use of factorial experiments in seeking the optimum operating conditions of an industrial process. Suggest alternative designs for this purppse. (10)

(d) DESIGN AND ANALYSIS OF EXPERIMENTS—STATISTICAL ASPECTS

(Answer *any four* questions from this section)

1. Discuss critically, with examples, how in the design and analysis of experiments the following techniques may be used to increase the precision of experimental results:

(a) grouping of experimental units.

(b) choice of size, shape and structure of experimental units.

(c) correction for concomitant variation. (25)

2. What is a balanced incomplete block design? How would you analyse the results obtained from such an experiment and also extract the information from block totals? Is it always worthwhile to extract inter-block information? (25)

3. A 2^3 factorial experiment was conducted on coconut trees, using single tree plots and blocks of 8 plots, and achieving balanced loss of information on partially confounded three-factor and four-factor interactions in 5 replications, and preserving all other effects.

Explain clearly how you would analyse the results of this experiment. (15)

It was later found that some of the trees had a clockwise spiral of leaves while the others had an anticlockwise spiral. How would you examine if the treatments have differential effects on these two types of trees? (10)

4. In order to estimate the parameters in the formula

$$\eta = \alpha+\beta_1 x_1+\beta_2 x_2+\gamma_{11}x_1^2+\gamma_{12}x_1x_2+\gamma_{22}x_2^2$$

connecting the expected response η with the levels (x_1, x_2) of two quantitative factors, an experiment was conducted and the response y_i was observed at the levels (x_{1i}, x_{2i}) of the two factors, $i = 1, 2, \ldots n$. The levels were chosen to satisfy the following conditions:

$$m(1, 0) = m(0, 1) = 0$$
$$m(2, 0) = m(0, 2) = 1, m(1, 1) = 0$$
$$m(3, 0) = m(0, 3) = m(2, 1) = m(1, 2) = 0$$
$$m(4, 0) = m(0, 4) = 3m(2, 2) = \lambda, m(3, 1) = m(1, 3) = 0$$

where

$$m(l_1, l_2) = \frac{1}{n}\sum_{i=1}^{n}\begin{pmatrix} l_1 & l_2 \\ x_{1i} & x_{2i}\end{pmatrix}$$

Obtain least-square estimates of the parameters.

Assuming that the errors of observation are independent and have equal variances, work out the variances and the covariances of the estimates of the parameters. (25)

5. A teacher wants to carry out an experiment to determine which of the following two methods of memorizing a poem is better: (1) memorizing each line separately or (2) memorizing stanza by stanza. Students of several boys' and girls' schools in the city are available for the experiment.

Suggest a suitable plan for the experiment, explaining clearly what is sought to be achieved by each special feature of the plan. (25)

6. Write critical notes on *any* two of the following:

(a) designs for two-way elimination of heterogeneity.

(b) planning of long-term experiments.

(c) agronomic extension experiments on cultivators' land. ($12\frac{1}{2}\times 2$)

(c) SAMPLE SURVEYS—THEORETICAL ASPECTS

(Answer *any four* questions from this section)

1. (*i*) Suppose d distinct units occur in a sample of n units selected with equal probability with replacement from a population of N units. Show that the estimator $\left(\sum_{i=1}^{d} y_i\right)/d$ where y_i is the value of the characteristic under consideration for the ith distinct unit in the sample, is unbiased for the population mean $\bar{Y}$. (5)

(*ii*) Obtain an unbiased estimator of the variance of this estimator. (10)

(*iii*) Suppose in a list of N factories serially numbered, m factories have gone out of existence and n new factories have been added to the list making the present total number of factories $(N-m+n)$. Give a simple procedure for selecting one factory with equal probability from the $(N-m+n)$ factories avoiding renumbering of the original N factories and show that your procedure achieves equal probability for the new factories. (10)

2. (*i*) Describe 'linear' and 'circular' systematic sampling procedures and discuss briefly their advantages and disadvantages. (7)

(*ii*) Show that (*a*) $I\,T$ and (*b*) $N\,T/n$ are unbiased estimator of the population total in "linear" and "circular" systematic sampling respectively, where I is the sampling interval, N the number of units in the population, n the number of units in the sample, and T the sample total. (8)

(*iii*) Show that systematic sampling will be more efficient than sampling with equal probabilty without replacement if the intra-class correlation coefficient $\rho < -\frac{1}{N-1}$, where N is the number of units in the population. (10)

3. (*i*) Let y_i and p_i $(i=1, 2, \ldots, n)$ be the value of the characteristic under consideration and probability of selection at any draw for the ith unit in a sample of n units selected with varying probabilities with replacement. Obtain an unbiased estimator of the population total Y and derive an unbiased estimator of its variance. (10)

(*ii*) If a sample of n units is selected with varying probabilities without replacement, show that the estimator $\left(\sum_{i=1}^{n} y_i/\pi_i\right)$ where π_i is the probability of inclusion of the ith unit in the sample, is unbiased for the population total Y. Derive an unbiased estimator of the variance of this estimator. (15)

4. (*i*) Describe cluster sampling and discuss briefly its advantages and disadvantages. (5)

(*ii*) Compare the efficiency of selecting n clusters of M units with equal probability without replacement with that of sampling nM units directly with equal probability without replacement. (8)

(*iii*) Given a sample of n clusters of M units selected with equal probability with replacement, obtain an estimator of loss of efficiency in using cluster sampling as compared to sampling nM units directly with equal probability with replacement. (12)

5. (*i*) Discuss briefly the advantages of stratified sampling. (4)

(*ii*) Mention the points one should consider in evolving a suitable stratified sample design for a survey. (6)

(*iii*) Suppose the objective is to estimate the difference in the rates of incidence of a particular disease in a 'model village' and a neighbouring village. Treating the villages as strata and assuming the cost of diagnosis of the disease to be C_1 and C_2 per person respectively in the two villages, determine the optimum allocation of sample size to the two villages when the cost of the survey is fixed at C. Let N_1 and N_2 be the total number of persons and P_1 and P_2 be the rates of incidence of the disease in the two villages. Assume equal probability sampling without replacement in the strata. (15)

6. (*i*) If y and x are unbiased estimators of the population totals Y and X of the characteristic under consideration and the supplementary variate respectively based on any sample design, show that the ratio of the exact bias of the ratio estimator $\frac{y}{x}X$ to its standard error is not greater than the coefficient of variation (relative standard error) of the estimator x. (8)

(*ii*) Derive approximate expressions for the bias and the mean square error of the ratio estimator $\frac{y}{x}X$ stating clearly the assumptions involved. (10)

(*iii*) Let y_s and x_s be unbiased estimates of the population totals Y and X of the characteristic under consideration and the supplementary information respectively based on the sample 's' selected with probability p_s. Show that the ratio estimator $\frac{y_s}{x_s}X$ will be unbiased for the population total Y if the sample 's' were selected with probability proportional to $x_s\, p_s$. Obtain the variance of this estimator. (7)

7. (*i*) Define a self-weighting design and discuss briefly its advantages and disadvantages. (6)

(*ii*) Show that a two-stage design, where n villages are selected with probability proportional to number of households in them, in the first stage, and m households are selected with equal probability without replacement in the second stage from each selected village, is self-weighting. (7)

(*iii*) Derive an unbiased estimator of the variance of the self-weighting estimator obtained in (*ii*). (12)

(*g*) STATISTICAL INFERENCE—GENERAL THEORY

(Answer *any five* questions from this section)

1. (*a*) Define the "consistency" of an estimator (5)

(*b*) Show that if $E(T_n)\to 0$ and $V(T_n)\to 0$ as $n\to\infty$ then T_n is a consistent estimator of θ. (5)

(*c*) Give two examples of consistent estimators which are not unbiased. (5)

(*d*) Examine if the sample mean is a consistent estimator of the parameter μ of the Cauchy distribution (5)

$$f(x, \mu) = \frac{1}{\pi} \quad \frac{1}{1+(x-\mu)^2}, \quad -\infty < x < +\infty$$

2. (*a*) Show that the coefficient of correlation between two unbiased efficient estimators T and T' of a parameter θ is unity. (4)

(*b*) Using the above, show that an unbiased efficient estimate is unique in the sense that if T and T' are two such estimates, the probability of those sample points for which the values of T and T' are different is zero. (4)

(*c*) Obtain the Cramer-Rao expression for the minimum variance which an unbiased estimator of θ can attain. (8)

(*d*) Derive, if possible, a sharper bound for this minimum variance than that obtained in (*c*). (4)

3. (*a*) Define a "sufficient statistic" and a "sufficient set of statistics". (5)

(*b*) Examine if the sample mean is a sufficient statistic for the parameter μ of a normal distribution $N(\mu,\sigma)$ where μ and σ are unknown. (5)

(*c*) State a general form of distributions which admit of a sufficient statistic and examine if $N(0,\sigma)$ has this form with respect to the estimation of σ. (5)

(*d*) If T is an unbiased estimator of θ and T' a sufficient statistic, show that the function $h(T') = E(T|T')$ is also an unbiased estimator and that the variance of $h(T')$ cannot exceed that of T. (5)

4. (*a*) Define a "most powerful" test and show how such a test is constructed in the case of a simple hypothesis against a simple alternative. Give proofs for your assertions. (7)

(*b*) Examine if a most powerful test is necessarily unbiased. (6)

(*c*) Construct a most powerful test for the hypothesis $\sigma = \sigma_0$ for a normal population $N(0,\sigma)$ and examine if the test obtained is also uniformly most powerful. (7)

5. (*a*) Define a "composite hypothesis" and give two examples of such hypothesis. (4)

(*b*) What are "similar regions" in sample space and how are they useful in test construction? (4)

(*c*) State a property which the frequency function of a population should have in order that similar regions may be obtained. How would you proceed to construct most powerful tests from similar regions? (6)

(*d*) Describe the construction of similar regions and a most powerful test for the hypothesis $\mu = \mu_0$ for a normal population $N(\mu,\sigma)$. (6)

6. (*a*) State Neyman-Pearson's lemma. (6)

(*b*) Define "locally most powerful unbiased tests" and show how the above lemma can be used to construct these tests. (7)

(c) Construct a locally most powerful unbiased test for the hypothesis $\sigma = \sigma_0$ for a normal population $N(0, \sigma)$. (7)

7. (a) What is a confidence interval in the Neyman sense? (6)

(b) Explain clearly with the help of an example Neyman's concept of the "shortest" confidence interval. (6)

(c) Show how the above ideas are connected with the construction of uniformly most poweful tests, illustrating your statements with the consideration of the parameter μ of a normal population $N(\mu, 1)$. (8)

8. Write in brief a description of the problem of statistical inference viewed as a decision problem and bring out how estimation of parameters and testing of hypotheses are two special cases of decision problems. (10)

Show how the solution of an appropriate zero-sum two person game leads to the solution of a decision problem. (10)

PAPER IX : SUBJECTS OF SPECIALISATION—II

Time : 4 hours Full marks : 100

(i) Answer questions only from the section on the subject of specialisation you have opted for.

(ii) Figures in the margin indiate full marks.

(a) ECONOMIC STATISTICS—INDIAN ECONOMICS AND ECONOMICS OF PLANNING

GROUP A

(Answer *any four* questions from this group)

1. Briefly discuss the merits and demerits of direct and indirect taxes in the Indian context. On what economic grounds would you justify taxes on luxuries? (12½)

2. What is the function and performance of the Industrial Finance Corporation? Why is the Indian banking system unequal to the task of financing industries? (12½)

3. *Either,*

What are the main steps taken by the Reserve Bank of India to extend facilities for bank credit to Indian agriculturists? (12½)

Or,

What are the basic advantages of an extensive bill market? Enumerate the steps taken by the Reserve Bank of India to promote an active bill market. (12½)

4. Give the current official estimates of the percentage distribution of national income over agricultural and non-agricultural sectors of the Indian economy and the percentage distribution of the working force over these sectors.

What picture do you get if you compare these with those of some of the developed economies? Bring out the significance of the change from one type of sectoral distribution of working force to another. (12½)

5. What is the bearing of the Indian transport system on the current coal crisis? Suggest some major lines on which a solution of the crisis has to be sought. (12½)

6. Give a critical appraisal of the present system of subsidies to coastal shipping. (12½)

7. Write notes on *any two* of the following: (6¼+6¼)

(a) The place of engineering industries in economic growth.

(b) The progress of engineering industry in India.

(c) The importance of Indian jute industry and the difficulties facing the industry.

8. *Either,*

Trace the general progress achieved in imposing ceiling on land holdings in different states of the Indian Union. (12½)

Or,

Report the progress of consolidation of holdings in India. Discuss how far the two reforms of fixation of ceiling and consolidation are contradictory. (12½)

GROUP B

(Answer *any three* questions from this group)

9. Elaborate Mahalanobis' two-sector growth model to show that a demand function is implicit in the model. What is the marginal propensity to consume, according to that demand function? (16⅔)

10. State Harrod's equations for natural rate of growth and warranted rate of growth. Carefully explain the terms and bring out the oscillatory character of growth in this model. (16⅔)

11. What do you understand by the term transaction matrix for a given economy? Explain in what manner this is useful for national planning? (16⅔)

12. What are the estimated indices or rates of growth of investment in the public and the private sectors of industry in India over the decade 1950-60? Broadly indicate the nature of the problem of deflation involved in such an estimate and give a model procedure for deflation. (16⅔)

13. (*a*) State the estimated magnitudes of unemployment at the end of the Second and the Third Five Year Plans. Give reasons for the trend and suggest remedies that are consistent with sound economics. (10)

(*b*) Briefly discuss the financial resources for the Third Five Year Plan (6⅔)

(b) TECHNO-COMMERCIAL STATISTICS

(i) Answer Group A and either Group B_1 or Goup B_2 whichever you have opted for.

(ii) Use a separate answer book for each group.

GROUP A: STATISTICAL METHODS IN BUSINESS

(Answer *both* the questions from this group)

1. (a) What is Market Research, its scope and purpose? List the various activities of marketing research personnel. What are the various stages of a market research enquiry, and what points should generally be covered in the final report? (6)

(b) What are the advantages and disadvantages of conducting a Market Survey by mailed questionnaires as compared to personal interview method? (6)

(c) Briefly explain the Hansen-Hurwitz method of dealing with non-response and its use in market research. (6)

2. *Either,*

(a) What is Business forecasting? What is the type of data and what are the methods used for forecasting? (6)

(b) A company uses the following formula:

Forecast Sales = Disposable income (weighted for industry) × Industry trend × company share of market.

Explain the various terms in the formula and how the data for these are obtained. (6)

Or,

(a) What is Job Evaluation? What is its central purpose? Outline the main steps in the Job Evaluation approach. What are the factors to be taken into account in evaluating a Job? (5)

(b) Give a brief summary description of each of the four basic methos of evaluation. Explain their comparative advantages and disadvantages. (4)

(c) What is merit rating? How does it basically differ from Job Evaluation? (3)

GROUP B_1: OPERATIONS RESEARCH

(Answer question 7 and *any two* from the rest in this group)

3. A contractor has a firm order to supply goods at a uniform rate R per unit time. He starts a production run every t time units, where t is fixed. His set up cost each time he starts a run is C_3. Production time is negligible. The cost of holding unit amount of inventory for unit time is C_1.

(a) Assume that shortages are not permitted. How frequently should a production run be made and what quantity of goods should be produced? (10)

(*b*) Suppose shortages are permitted, and let C_2 be the unit cost of shortage per unit time. How frequently should a production run be made and what should be the planned initial inventory level at the time each batch is completed? (15)

4. (*a*) A shop keeper has to decide how much quantity of bread he should stock every week. The quantity of bread demanded in any week is assumed to be a continuous random variable with a given probability density function $f(x)$. Let A be the unit cost of purchasing bread, B unit sale price, C refund on stale bread per unit quantity, and finally D is the unit penalty cost.

Find the optimum quantity of bread to be stocked. (15)

(*b*) If $A=8$, $B=20$, $C=2$, $D=5$ and the distributions of demand is rectangular between 1000 and 2000, show that the optimum quantity is 1740. (10)

5. (*a*) Explain clearly what is meant by the dual of a given linear programme and the mutual relationship of the two problems. (8)

(*b*) Write down the dual of the following linear programming problem. Maximise $3x_1+5x_2+4x_3$ subject to the restrictions x_1, x_2, x_3 non-negative and $2x_1+3x_2 \leqslant 8$, $2x_2+5x_3 \leqslant 10$ and $3x_1+2x_2+4x_3 \leqslant 15$. (8)

(*c*) The optional simplex tableau of the above problem with x_4, x_5, x_6 as slack variables is given below.

	x_1	x_2	x_3	x_4	x_5	x_6	right hand side
x_1	1	0	0	$\frac{-2}{41}$	$\frac{-12}{41}$	$\frac{15}{41}$	$\frac{89}{41}$
x_2	0	1	0	$\frac{15}{41}$	$\frac{8}{41}$	$\frac{-10}{41}$	$\frac{50}{41}$
x_3	0	0	1	$\frac{-2}{41}$	$\frac{-12}{41}$	$\frac{15}{41}$	$\frac{89}{41}$
c_j-z_j	–	–	–	$\frac{-45}{41}$	$\frac{-24}{41}$	$\frac{-11}{41}$	$Z=\frac{765}{41}$

From this tableau, find the optimum value of the dual variables and verify the duality theorem. (9)

6. (*a*) Write down (without proof) the equations for the steady state probabilities of a queuing system with a single server, when both the arrival and service time distributions are negative exponential, the queue discipline is first come first served and the mean arrival rate λ is less than the mean service rate μ.

Also write down the average time an arrival spends in the system, both in the queue and the counter. (12)

(*b*) The arrival distribution of ships in a harbour is Poisson with mean rate λ ships per week and service time distribution is exponential with mean rate of μ

unloading per week. The queue discipline is : first come first serviced. The ships have to wait till the dock is free. There is only one unloading dock in the harbour, and the cost per unloading operations is proportional to the speed of service, that is, it is equal to $D\mu$ where D is the cost per operation per unit service rate of operating the dock. The average cost of having one ship idle in the port (the cost of crew, overhead etc.) is proportional to the average time spent in port, that is, it is equal to $C\,W$ where C is the cost of ship delay per week, W is the mean time including service spent in the port, expressed in weeks.

Given λ, C, D, determine the value of μ so that the total average cost per ship unloaded is minimum. From this value of μ show that we may include as follows.

(*i*) If the crew cost C is small compared to dock cost D, the optimum service rate need not be much larger than the arrival rate i.e. one should keep dock utilization high at the expense of ship delay. (6)

(*ii*) But if the crew cost C is larger compared to dock cost D, the optimum value of μ would be considerably larger than λ i.e. one could afford to have one's efficient dock idle most of the time in order to reduce costly ship delays. (7)

7. Supposing that you are being interviewed by the Manager of a commercial firm for a job in its operational research department, explain, in non-technical language, the scope and purpose of O.R. and its usefulness to the firm. Give some examples of application of O.R. in industry. (20)

GROUP B_2 : ELEMENTS OF BOOK-KEEPING AND ACCOUNTANCY

(Attempt *all* questions from this group)

8. From the following balances of S. K. Bose, a sole trader, prepare Trading and Profit & Loss account and Balance Sheet, as at 31st December 1961 :

	Rs.		Rs.
Capital Account	20,500	Rent Received	300
Carriage Inwards	750	Rent and Insurance	9,955
Cash in hand	55	Purchase returns	2,000
Creditors-Trade	15,000	Plant and Machinery	10,000
Cash at Bank	20,645	Power	4,500
Stock on 1 January 1961	16,000	Purchases	85,500
Sales	1,65,050	Bad debts	600
Salaries and wages	17,200	Bad Debt Reserve	
Sales Returns	300	(as of 1 January 1961)	300
Discount Received	900	Advertising (Development)	4,000
Debtors	7,300	Goodwill	2,500
Discount allowed	2,500	Travellers samples	1,350
Creditors-(Expenses)	3,400	Travellers Commission	1,445
Drawings	2,500	Travellers salaries	4,650
General Charges	4,300	Wages-Manufacturing	11,500

In preparing the accounts the following are to be taken into consideration:

(*i*) The closing stock was Rs. 11,500/- but there has been a loss by fire on 20 December 1961 to the extent of Rs. 10,000/- covered by insurance.

(*ii*) Depreciate Plant and Machinery by 10% and Travellers samples by 33⅓%.

(*iii*) Increase the Bad debts Reserve by Rs. 1,000/-.

(*iv*) Write 50% off Advertising Development Account.

(*v*) Annual premium on Insurance expiring 1 March 1962 was Rs. 600/-.

(*vi*) A commission of one per cent on the gross profit is to be provided for works manager and a commission of 5% of net profit (after charging the works manager's commission) is to be credited to General Manager. (20)

9. From the following entries in the Bank Columns of the Cash Book of M/s Roy & Co. and the corresponding Pass Book you are required to prepare a Bank Reconcilation Statement as at 31 December 1961. (8)

BANKS COLUMNS OF CASH BOOK

1961		Rs.	1961		Rs.
Dec. 1	To balance b/d	1,400	Dec. 5	By Pandey & Sons	200
3	,, Son & Co	245	18	,, Garlick & Co.	300
7	,, Rajaram & Sons	150	20	,, Islam Bros.	425
12	,, Smith Bros.	375	24	,, Charles & Co.	320
17	,, Chatterjee & Co.	400	30	,, Rent of premises	875
20	,, Interest on Investment	530	31	,, Office slaries	1,750
25	,, Isaac & Sons	1,470	31	,, Workshop wages	2,250
30	,, Black & Co.	2,400	31	,, Legal Charges	250
31	,, Balaram & Co.	1,450	31	,, Drawings	600
31	,, Interest on Deposit	350	31	,, Interest on Loan	500
31	,, Yusuf Bros.	1,800	31	,, Bank charges	50

M/S Roy & CO. IN ACCOUNT WITH THE BANK Ltd.

1962		Rs.	1962		Rs.
Jany 1	To Balance (over Draft)	3,025	Jany 3	By Issac & Sons	1,470
8	,, Garlick & Co.	300	13	,, Mitra Bros	740
14	,, Perry & Co.	575	19	,, Yusuf Bros	2,400
23	,, Islam Bros	425	28	,, Sultan Bros	900
30	,, Murray Bros	700	31	,, Black & Co.	1,800
31	,, M. Chowdhury	180	31	,, Talukdar & Bros	1,800
31	,, Charles & Co.	320	31	,, Balaram & Co.	1,450

10. Roy, for the mutual accommodation of himself and Bose, draws upon the latter a bill at three months date for Rs. 8,000/- dated 1st January 1961. The bill is discounted by Roy at 5% and half the proceeds are remitted to Bose.

Bose at the same time draws a bill at three months date on Roy for Rs. 4,000/-. After securing Roy's acceptance, the bill is discounted at 6% by Bose who remits half the proceeds to Roy. Bose becomes bankrupt on 31st March and 25 nP in the rupee is received on 15th May as first and final dividend from his estate.

Write up the journal entries in the books of Roy and the ledger entries (Roy's account) in the books of Bose. (16)

11. *Either,*

(*a*) Write short notes on the following :

(*i*) Suspense account (2)

(*ii*) Analytical Petty Cash Book (give proforma example) (2)

(*iii*) Subsidiary Journal (2)

(*iv*) Imprest system of Petty Cash (2)

(*b*) 'Balance Sheet and Profit and Loss Account are dependent on each other' —Explain. (4)

Or,

(*a*) Is trial balance agreement a complete guarantee as to the accuracy of the books of accounts? (5)

(*b*) What steps would you take to discover errors in cases of disagreement? (5)

(*c*) Can closing stock be included in trial balance? (2)

12. *Either,*

(*a*) Explain :

(*i*) Fictious Assets and Wasting Assets. (3)

(*ii*) Capital and Revenue Expenditure. (3)

(*iii*) Deferred Revenue Expenditure. (3)

(*b*) How should stock-in-trade be valued and adjusted at the end of each financial year? (5)

Or,

(*a*) Distinguish between Trading, Manufacturing and Working account. (3)

(*b*) How should the following items be treated in finanl accounts?

(*i*) Raw materials (*ii*) Finished products (*iii*) Non-productive wages (6)

(*c*) Mention the points that would call for enquiry when the percentage of gross profit is more than justifiable. (5)

(*d*) DESIGN AND ANALYSIS OF EXPERIMENTS—COMBINATORIAL ASPECTS

(Answer *any four* questions from this section)

1. Define a finite field and a finite projective Geometry and show how you will construct a finite field with four elements. (12)

Show how you will construct $PO(2, s)$ with the help of $GF(s)$ where s is a prime power and in particular write down the "points" and "lines" of $PO(2, 3)$. (13)

2. Define a set of mutually orthogonal latin squares of a given order. (4)

If $N(v)$ denotes the maximum number of mutually orthogonal latin squares of order v, show that

$$N(v) \leqslant v-1 \tag{6}$$

Show that any set of $v-2$ mutually orthogonal latin squares of order v can be completed to a full set of $v-1$ mutually orthogonal latin squares. (9)

Give an example of a latin square of order 4 which does not have an orthogonal mate. (6)

3. Define the notion of an orthogonal array with v levels, k constraints, strength d and index λ. (5)

Establish the relation between an orthogonal array indicated above with $d=2$ and $\lambda = 1$ and a set of mutually orthogonal latin squares of order v. (10)

Given a balanced incomplete block design with

$$V = b = 4\lambda - 1, \quad r = k = 2\lambda, \quad \lambda = \lambda$$

show how you will construct an orthogonal array with 2 symbols, $4\lambda-1$ constraints, strength 2 and index λ. (10)

4. Define a balanced incomplete block design and obtain the relations amongst the parameters. (5)

Using matrix methods show that a necessary condition for existence of a symmetric b.i.b.d. is that $r-\lambda$ must be a perfect square when v is even. Also derive the inequality $b \geqslant V$ for any b.i.b.d. (8)

Explain the first fundamental theorem of differences to construct a b.i.b.d and use it to construct a design with parameters

$$V = b = 11, \quad r = k = 5, \quad \lambda = 2 \tag{12}$$

5. Define a partially balanced incomplete block design with two associate classes and establish the relation amongst the parameters. (6)

Obtain the parameters of a group-divisible design and show that these parameters uniquely define a group divisible association scheme. (5)

Obtain the classification of G.D. designs with the help of the characteristic roots of NN' where N is the incidence matrix of the design. (8)

Show that there is a one-to-one correspondence between a singular G.D. design and a certain b.i.b. design. (6)

6. Explain the principle of generalised Interaction with reference to a 2^m symmetric factorial experiment. (10)

A 2^4 experiment is to be carried out in blocks of 4 plots each without confounding main effects. Give a suitable design in 4 replications partially confounding every interaction in some replication or other. Write down what effects are confounded in each replication. (15)

7. Explain clearly what is meant by a rotatable design. (9)

Give a method of constructing rotatable design of the second order in three dimensions. (16)

— (c) SAMPLE SURVEYS—ORGANISATIONAL ASPECTS

(Answer *any four* questions from this section)

1. Write an essay on planning data collection in a countrywide sample survey in India in the socio-economic field. Indicate the special features to be incorporated if the survey is intended to be a continuing one. (25)

2. What should be the duties and functions of a field supervisor placed in charge of an organisation set up to collect data on area and outturn of principal crops grown in a State. Mention some of the control charts that he may maintain, indicating their particular uses. (Illustrate on the basis of any State you know best). (25)

3. On what practical considerations will you base the design of a sample survey to make a reliable estimate of the rate of growth of human population in the different States of India? How would these considerations be different if the objective of the survey were to gain insight into the economic condition of small scale (below factory level) industries? (25)

4. Draw up a schedule, for inquiry into the state of employment, under employment in the rural sector of India. Give your own definitions and concepts with justification and elaborate a set of instructions to field workers. (25)

5. What should be the concurrent and post survey arrangement to ensure maximum accuracy of the data collected in a survey of the type mentioned in Question 4 above, before the data are forwarded to the processing department. (25)

6. Write notes on *any three* of the following, in respect of sample surveys:

(*i*) legal provision (*ii*) publicity arrangements (*iii*) budget preparation (*iv*) survey calendar (*v*) post-enumeration checks (*vi*) call-back. (25)

PAPER X: SUBJECTS OF SPECIALISATION—III

Time: 5 hours Full marks: 100

(*i*) Answer questions only from the section on the subject of specialisation you have opted for.

(*ii*) Figures in the margin indicate full marks.

(*iii*) Use of calculating machines is permitted.

(a) ECONOMIC STATISTICS (PRACTICAL)

(Attempt *any five* questions from this section)

1. The following table gives the per capita consumer expenditure in rupees per month on all items and on foodgrains by certain class ranges of per capita consumer

expenditure. The figures relate to rural India and were collected during the 7th round of the National Sample Survey, November 1953 to March 1954.

Compute Engel elasticity of food expenditure. (20)

per capita expenditure class (Rs.)	per capita monthly expenditure (Rs.)		percentage distribution of persons
	on all items	on food grains	
0— 8	6.20	3.76	15.47
8—11	9.63	5.03	17.80
11—13	11.92	6.43	12.04
13—15	13.96	6.90	10.31
15—18	16.21	7.58	10.83
18—21	19.06	7.99	8.75
21—24	22.06	8.90	6.94
24—28	26.06	8.35	5.77
28—34	30.53	8.57	4.82
34—43	36.89	8.98	3.73
43—55	48.98	11.69	1.07
55 and above	89.05	10.68	1.57
all classes	17.22	6.73	100.00

2. In the Mahalanobis two-sector model given by

$$y_t = y_0 \left[1 - \alpha_0 \frac{\lambda_i \beta_i + \lambda_c \beta_c}{\lambda_i \beta_i} \left\{ \left(1 + \lambda_i \beta_i \right)^t - 1 \right\} \right]$$

where $\lambda_0 = 1000$, $\alpha_0 = 6$ per cent $\beta_i = 0.25$ and $\beta_c = 0.50$, find out from which year, $\lambda_i = 0.5$, would give a larger value of national income than $\lambda_i = 0.1$. (15)

In how many years will national income be (*i*) double and (*ii*) three times, at the annual geometric rate of growth of 6 per cent per year? (5)

3. The following table gives the frequency distribution of income in Britain during 1911-12 and covers persons with higher income.

Fit a Pareto curve to the data, work out the expected frequencies, and obtain an estimate of concentration ratio graphically or otherwise. (15+5)

income in thousands of pound sterling	number of persons in thousands
5—10	7546
10—15	1890
15—20	790
20—25	424
25—35	411
35—45	190
45—55	103
55—65	70
65—75	50
75—100	118
over 100	85

4. Use simplex method to maximise

$$y = 6x_1 + 5.5x_2 + 9x_3 + 8x_4$$

subject to the restrictions,

$$x_1 + x_2 + x_3 + x_4 \leqslant 15$$

$$7x_1 + 5x_2 + 3x_3 + 2x_4 \leqslant 120$$

$$3x_1 + 5x_2 + 10x_3 + 15x_4 \leqslant 100$$

as well as the conditions

$$x_1 \geqslant 0, x_2 \geqslant 0, x_3 \geqslant 0 \text{ and } x_4 \geqslant 0. \qquad (20)$$

5. The following table gives the inter-industrial flows in India in 1949-50. The figures are in Rs. crores (market prices).

	intermediate flows			final products	total products
	agriculture	industry	rest		
intermediate flows					
agriculture	939	942	126	3845	5852
industry	216	452	195	2838	3701
rest	50	676	238	2805	3769
value added	4647	1631	3210		
total product	5852	3701	3769		

If the final products in agriculture, industry and rest respectively go up by 5 per cent 10 per cent and 6 per cent what will be the levels of total products in these sectors? If the total products go up, respectively, by these percentages, obtain the final product vector. (16+4)

6. The following table gives the price and production of oats in the USA expressed as ratios to the respective trends and denoted respectively by Y and X:

year	Y	X	year	Y	X
1881	1.30	0.929	1894	1.07	0.886
1882	1.05	1.036	1895	0.75	1.070
1883	0.90	1.156	1896	0.76	0.983
1884	0.85	1.128	1897	0.96	0.969
1885	0.84	1.165	1898	0.95	1.005
1886	0.77	1.108	1899	0.83	1.074
1887	0.96	1.124	1900	0.86	1.033
1888	0.79	1.151	1901	1.37	0.857
1889	0.81	1.188	1902	1.03	1.131
1890	1.48	0.798	1903	1.14	0.911
1891	1.10	1.088	1904	0.86	1.033
1892	1.09	0.943	1905	0.86	1.000
1893	1.16	0.882			

It has been found that a curve, $Y = \frac{1}{a+bX}$ approximately represents the relations between the two variables. Estimate a and b. (20)

(b) TECHNO-COMMERCIAL STATISTICS (PRACTICAL)

(i) Answer Groups A & C and either Group B_1 or Group B_2 whichever you have opted for.

(ii) Use a separate answer book for each group.

GROUP A : STATISTICAL QUALITY CONTROL

(Answer *any two* questions from this group)

1. The dimention of a component has been under control with an overall average of 0.834" and an average range in samples of size 5 of 0.006". The specification limits for the product are 0.820" and 0.840".

(a) what is the standard deviation of item quality? (2)

(b) what is the 1% significance level value of the sample range? (3)

(c) what is the percentage of "off limits" produced by the process? (4)

(d) what would be the per cent defective produced when the process average shifts to 0.837"? (4)

(e) what is the probability that a sample observation would catch such a shift on the control chart? (4)

(f) how many samples of size 5 would have to be taken to have a probability of 0.95 of catching the shift by at least one of the samples? (4)

(g) what should be the sample size so that there would be an even chance that a sample observation would detect the change? (4)

2. (a) A state trading corporation has specified that the gross weight of cement bags should be between 50 ± 5 kg and that the average weight of 20 randomly chosen bags should not be less than 50 kg. What should be the mean and the maximum standard deviation of the filling process so that the specification requirements could be met most economically? (7)

(b) A company which is developing a new product makes 25 items on a trial basis. 4 of these were found to be defective. Some alterations in the design were then made and n items produced of which *none* were found to be defective. How large should n have been so that the new design could be considered an improvement at 1% level of significance? (13)

(c) Weights of empty tins were found to be distributed with mean = 95 gm and s.d = 4 gm. The weights of filled tins (gross weight) were distributed with mean = 601 gm and s.d = 16 gm. Estimate the mean and standard deviation of the nett weight of the contents. State assumptions made, if any. (5)

3. The following data refer to volume, in units of 0.01 c.c. filled in "2 c.c." ampoules by 6 fillers. 8 ampoules filled by each filler have been tested.

f_1	f_2	f_3	f_4	f_5	f_6
226	218	224	211	227	196
199	205	236	188	252	215
247	231	245	226	229	180
227	235	215	229	244	211
254	195	274	203	226	212
230	220	230	223	219	204
226	218	213	225	203	189
212	229	210	222	216	186

(*a*) Test for differences among the fillers in respect of variability as well as the average level. (15)

(*b*) Estimate the overall variability obtained in the process and the likely improvements when assignable causes, if any, are eliminated. State assumptions made, if any. (10)

GROUP B_1 : OPERATIONS RESEARCH

(Answer *both the questions* from this group)

4. *Either,*

Obtain a good first feasible basic solution of the following transportation problem by Matrix minimum method (i.e. fill the cell with the smallest cost, and then the cell with the next lower cost and so on). Test whether this solution is optimal. Otherwise obtain an improved basic feasible solution and its corresponding cost. (20)

origins	destinations					supplies
	1	2	3	4	5	
1	7	10	7	4	7	10
2	6	2	6	6	4	12
3	6	5	9	11	3	4
4	5	7	10	1	1	18
requirements	8	8	12	6	10	44

Or,

A firm which buys and sells rice has storage space for 4000 tons. The buying and selling prices vary from season to season and are given below.

season	summer	autumn	winter	spring
Cost price Rs./ton	1250	1100	1200	1150
Sales price Rs./ton	1300	1200	1275	1250

During any season, the firm can sell any amount up to total stock at the beginning of the season. Also, it can order any amount during the middle of any season at prices shown above, to be delivered at the beginning of the next season. If in the beginning of summer, the firm has a stock of 1000 tons how much should it plan to purchase and sell for the next four seasons to maximise its profit. (20)

5. The cost per year of running a machine whose purchase price was Rs. 6000, and its resale values are given below.

At what age is a replacement due? (10)

	year						
	1	2	3	4	5	6	7
running cost	1000	1200	1400	1800	2300	2800	3400
resale price	3000	1500	750	375	200	200	200

GROUP B_2: ELEMENTS OF BOOK-KEEPING AND ACCOUNTANCY

(Answer *all* questions from this group)

6. How would you rectify the following errors discovered while preparing the trial balance: (9)

(*a*) An item of Rs. 750 paid for the purchase of Gas Engine has been debited to Cash Purchase account.

(*b*) A credit purchase of Rs. 125 form X & Co. has been omitted to be passed through the invoice book but the payment made subsequently stands debited to their account.

(*c*) A credit sale of Rs. 75 to A. Sen duly entered in the Sales Book has been credited to his account.

(*d*) Rs. 85 paid for the purchase of a sporting gun for one of the partners had been charged to Cash Purchases account.

(*e*) An addition in the Returns Inwards book had been cast 10 nP short.

(*f*) A cheque for Rs. 100 received from a Sundry Debtor was duly recorded in Cash Book and the Party's account. But the subsequent dishonour of the cheque was not taken into account.

7. The Sundry Debtors at 31st December, 1960 were Rs. 40,000 and you are required to make a 5% Reserve for Doubtful debts and also a 5% Reserve for Discounts. The actual bad debts during the year 1961 amounted to Rs. 1,600 and the discount allowed was Rs. 1,700. The debtors at the close of the year 1961 were Rs. 50,000. Reserve for Doubtful debts and Reserve for Discounts are required to be maintained @% as in the previous year. Show how the transactions would be shown in the Profit & Loss account and Balance Sheet for the years ended 1960 and 1961. (9)

8. From the following Trading & Profit & Loss account and the Balance Sheet, prepare a Trial Balance of S. K. Bose as at 31st December 1961. (12)

TRADING ACCOUNT
for the year ended 31.12. 1961.

		Rs.			Rs.
To Opening stock		21,300	By Sales	1,40,000	
„ Purchases	84,000		*less* Returns	5,000	1,35,000
Less returns	4,000	80,000	„ value of stock destroyed by fire		10,000
„ Freight & Duty		10,000	„ closing stock		27,000
„ Gross Profit c/d to P & L account		61,000			
		1,72,300			1,72,300

PROFIT & LOSS ACCOUNT
for the year ended 31.12.1961

To Salaries	18,500	By Gross profit b/d	61,000
„ Loss on Stock destroyed by fire	4,000	„ Profit of textile department	10,000
„ Interest on loan of D. Das	1,350	„ Reserve for discount on creditors	360
„ Insurance charges	800	„ Reserve for discount on Debtors	38
„ Trading Expenses	800		
„ Reserve for Debts	700		
„ Sundry Expenses	600		
„ Printing and Stationery	500		
„ Depreciation	11,450		
„ Net profit transferred to Capital account	32,698		
	71,398		71,398

BALANCE SHEET
as at 31.12.1961

Capital Account			Land & Building		
Balance on 1.1.1961	1,00,000		*Less* Depreciation		58,800
add profit for the year	32,698		Lease hold premises *Less* Depreciation		24,000
	1,32,698		Plant & Machinery *Less* Depreciation		16,000
Less Drawings	14,000	1,18,698	Furniture & Fixture *Less* Depreciation		4,750
			Stock-in-Trade		
Loan from D. Das			General goods	27,300	
add interest		31,350	Textile goods	8,000	35,300
Sundry Creditors	18,000				
Less Reserve for discount	360	17,640	*Sundry Debtors*	18,000	
			less Reserve for Bad debts	900	
				17,100	
			less Reserve for discount	342	16,758
			Insurance Company		6,000
			Prepaid Insurance		200
			Cash in hand		1,280
			Cash at Bank		4,600
		1,67,688			1,67,688

You are informed that the old Reserve for Doubtful Debts was Rs. 1,000.

Prepare the Trial Balance *prior* to the following adjustments which have already been made while preparing the Final Accounts :

(1) The stock on 31.12.1961 was Rs. 27,300.

(2) On 23rd December, 1961 a fire broke out and destroyed stock of the value of Rs. 10,000. The Insurance Company admitted the claim for loss of stock to the value of Rs. 6,000 only and paid the amount on 10th January, 1962.

(3) Of the Sundry Debtors, Rs. 400 are bad and should be written off. The Reserve for Doubtful Debts should be maintained at 5% on Debtors (Net) and the Reserve for Discount on Debtors at 2%. Also a reserve for discount on creditors is maintained at 2%.

(4) Goods of the value of Rs. 6,000 had been received on 27th December 1961, but the purchase invoice was omitted to be recorded in the Purchase Book.

(5) Mr. Bose had utilised goods worth Rs. 2,000 for his personal use, but no record was made of it.

(6) Depreciation is to be provided as under :

Land & Building	@ 2%
Plant & Machinery	@20%
Furniture & Fixture	@ 5%
Leasehold Premises	@20%

(7) Prepaid Insurance was Rs. 200.

(8) Interest on Loan was to be allowed at the rate of 4½% per annum.

Give suitable explanations for arriving at figures, where necessary. (12)

GROUP C : STATISTICAL METHODS IN BUSINESS

(Answer *all questions* from this group)

9. A departmental store wishes to ascertain the average amount outstanding on its customer accounts. Past experience shows that the sums outstanding are as shown in the following table :

Stratum	Amount outstnaifng (Rs.)	relative frequency	Variance of amount outstanding
1	Less than 5.0	.20	1
2	Between 5.0 and 20.0	.40	6
3	Between 20.0 and 50.0	.25	20
4	Between 50.0 and 100.0	.10	75
5	Between 100 and 250	.03	650
6	above 250	.02	2500

Show how to use stratified sampling to minimise the variance of the estimate of the mean amount outstanding based on a sample of 5000 accounts. (8)

10. A bureau of business research polled a sample of 500 men engaged in different fields of business in a city and obtained the following data :

	bankers	manufac-turers	merchants	farmers
increased activity	40	55	80	75
decreased activity	15	30	40	50
no appreciable change	20	25	30	40

Examine whether there were any differences in attitude towards the prospects for overall business activity among the different fields of business. (7)

11. From the monthly sales record of a departmental store for the years 1951 to 1955, a statistician obtained the following trend equation :

$$Y = 25.74 + 0.455x$$

Origin: January 1951. x:unit 1 month.

y: average monthly sales in thousands of rupees.

He also obtained the following monthly seasonal indices. (correct to 2 decimal places)

January	February	March	April	May	June
.79	.76	.94	.97	1.05	.97
July	August	September	October	November	December
.85	.89	1.02	1.21	1.16	1.31

Forecast the monthly sales for the year 1957. (5)

(*d*) DESIGN AND ANALYSIS OF EXPERIMENTS (PRACTICAL)

(Answer *any three* questions from this section)

1. The following table gives the yields in lbs of green tea leaf in a 3^3 factorial experiment using 2 replicates, each in 3 blocks of 9 plots each in which the same effects were confounded in both the replicates. The levels of the three factors are given below :

factor	level	
N	0	no nitrogen
	1	40 lbs "
	2	80 lbs "
K	0	no K_2O
	1	20 lbs "
	2	40 lbs "
P	0	no P_2O_5
	1	20 lbs "
	2	40 lbs "

The plan and yields are given below. Analyse the data fully and give your recommendations regarding the most effective combination of factors. (33)

plot number	levels of			yield	
	N	P	K	replicate 1	replicate 2
				block 1	block 4
1	0	0	0	454	386
2	0	2	1	359	408
3	0	1	2	525	561
4	1	1	0	419	201
5	1	0	1	408	500
6	1	2	2	390	597
7	2	2	0	608	381
8	2	1	1	500	687
9	2	0	2	510	505
				block 2	block 5
10	0	1	0	327	200
11	0	0	1	354	308
12	0	2	2	335	430
13	1	2	0	547	335
14	1	1	1	423	370
15	1	0	2	367	560
16	2	0	0	498	333
17	2	2	1	391	488
18	2	1	2	465	629
				block 3	block 6
19	0	2	0	276	192
20	0	1	1	451	288
21	0	0	2	263	257
22	1	0	0	370	228
23	1	2	1	291	281
24	1	1	2	373	283
25	2	1	0	401	274
26	2	0	1	362	449
27	2	2	2	620	758

2. An incomplete block design involving 10 treatments in 5 blocks of 4 plots each can be obtained by considering the points and lines of the diagram below respectively as treatments and blocks.

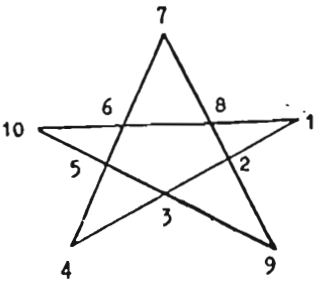


Fig.

(a) Write down complete instructions, giving blank tabular layouts wherever possible, for all the computations necessary in carying out an intra-block analysis of the results of such an experiment. (15)

(b) Work out the efficiency factor of the design. (10)

(c) Find the dual of the above design and identify it. (8)

3. 13 objects numbered (1), (2), ... (13), were weighed in combinations of 4 objects at a time on a spring balance. The results are given below:

serial number of weighing	objects weighed				total weight (gms)
1	(3)	(6)	(9)	(11)	28.7
2	(3)	(4)	(8)	(12)	27.5
3	(10)	(11)	(12)	(13)	45.6
4	(2)	(5)	(8)	(11)	25.8
5	(7)	(8)	(9)	(10)	34.4
6	(4)	(5)	(6)	(10)	25.0
7	(1)	(5)	(9)	(12)	26.9
8	(3)	(5)	(7)	(13)	28.3
9	(1)	(2)	(3)	(10)	16.4
10	(2)	(4)	(9)	(13)	27.4
11	(1)	(4)	(7)	(11)	23.2
12	(1)	(6)	(8)	(13)	28.0
13	(2)	(6)	(7)	(12)	26.9

Estimate the weights of the different objects and also the standard errors of of the estimates. (25)

Would it not have been better to weigh the objects individually? (8)

4. The following table gives the plan and the yields of an experiment arranged in a 5×5 Latin Square with two plots missing.

B 257	A 279	E 230	C 287	•
D 245	E 245	A 283	B 264	C 260
E 182	C 280	B 258	D 246	A 250
A 203	D 227	•	E 193	B 259
C 231	B 266	D 271	A 334	E 338

Analyse the data fully and work out the variances of the pairwise comparisons among A, B, C, D and E. (33)

Neatness (1)

(e) SAMPLE SURVEYS (PRACTICAL)

(Answer question 1 and *any two* of the rest from this section)

1. A survey is to be conducted to study the incidence (number of households indebted) and extent (average amount of indebtedness per household) of indebtedness among the scheduled caste population in a state and to compare this with the incidence and extent of indebtedness in the rest of the population.

(a) Plan a survey describing clearly and in brief—

(i) the sampling units you will use

(ii) the sampling frames required and how you will procure them

(iii) an outline of the sampling procedure

(iv) pilot studies, if you require any. (20)

(b) Give a tabulation programme describing only the important tables that should be compiled. (10)

(c) Draw up a suitable household-schedule for the survey giving—

(i) the identification details of the household

(ii) the relevant items to be collected (10)

(d) List down the different items in planning the survey, collecting the information and processing the data, which should be taken into consideration when drawing up the budget for the survey. (10)

2. A sample of 20 households was selected from a tehsil using a two-stage sampling method, villages being selected in the first stage with probability proportional to their 1951 census population, with replacement and an equal number of households being selected in the second stage, within each sampled village, with equal probability with replacement. The data collected is given in the table below.

village number	p_i	number of households	household number	earnings in Rs.
(1)	(2)	(3)	(4)	(5)
1	15.80	287	1	122
			2	150
			3	108
			4	140
			5	95
2	21.00	257	1	125
			2	122
			3	110
			4	134
			5	215
3	43.69	68	1	300
			2	115
			3	67
			4	110
			5	81
4	27.73	218	1	180
			2	174
			3	125
			4	120
			5	115

p_i = probability of selection of the ith village

(a) Estimate the average earnings per household in the tehsil and the coefficient of variation (relative standard error) of your estimate. (10)

(b) Express the variance of the estimate in the form $V = \frac{A}{n} + \frac{B}{mn}$ where n is the number of sample villages and m the number of households to be sampled in each village. If the cost function is of the form $C = c_1 n + c_2 nm$ where C = Rs. 1000-, c_1 = Rs. 800/- and c_2 = Rs. 5/- get the optimum values of n and m which minimize the variance of the estimate for the fixed cost of Rs. 1000/-, using the estimates for A & B from the data given. [Hint: Estimate A, using the estimates for V and B] (15)

3. It is proposed to estimate the total population of 64 cities in a large country by drawing a stratified single-stage sample of cities with equal probability without replacement in each stratum. The cities are grouped into two strata and the table below gives the stratum-wise mean and variance of population figures obtained from a previous census. The cost of surveying a city in the first stratum is Rs. 2500 and in the second stratum it is Rs. 400/-.

(a) For a total cost of Rs. 30,000/- find the optimum allocation to the strata, and the minimum variance of the estimate. (15)

(b) If instead, if it is desired that the estimate must lie within 5% of the true value with a confidence level fo 95% find the minimum cost of the survey. (10)

stratum number (i)	total number of cities (N_i)	$\bar{Y}_i$ in 10^3	S_i^2 in 10^6
(1)	(2)	(3)	(4)
1	16	630	53824
2	48	108	5625

4. The following table gives the geographical area and area cultivated under rice in 60 villages of a tehsil.

(*a*) Select a sample fo 10 villages with probability proportional to geographical area, with replacement, and estimate the total area under rice in the tehsil. Also, estimate the standard error of this estimate. (15)

(*b*) Using the sample you have drawn estimate the gain or loss in efficiency in selecting a sample of 10 villages as suggested in (*a*) as compared to selecting a sample of 10 villages with equal probability without replacement.

[Hint: An estimate of Y is obtained by $(\hat{Y})^2 - \hat{V}(\hat{Y})$ where Y stands for the total area under rice in the tehsil]

Table showing village-wise geographical area and area under rice in a tehsil

village number	geographical area (in acres)	area under rice (in acres)	village number	geographical area (in acres)	area under rice (in acres)
(1)	(2)	(3)	(1)	(2)	(3)
1	151	81	31	716	305
2	542	292	32	230	127
3	562	275	33	349	245
4	410	118	34	379	270
5	249	44	35	135	64
6	121	56	36	873	445
7	106	33	37	248	69
8	397	147	38	1034	401
9	453	194	39	503	164
10	710	282	40	206	46
11	176	65	41	682	166
12	730	288	42	653	129
13	327	115	43	166	44
14	280	161	44	445	116
15	287	179	45	1495	164
16	404	273	46	501	96
17	124	58	47	1473	373
18	244	58	48	1117	261
19	370	98	49	389	79
20	541	155	50	716	191
21	153	41	51	274	51
22	658	230	52	750	100
23	805	239	53	251	56
24	435	158	54	691	83
25	418	130	55	147	13
26	235	114	56	864	146
27	364	170	57	808	242
28	515	272	58	924	340
29	586	418	59	601	189
30	246	47	60	396	99

Note: Give the reference to the random number tables used for selection, and the pages, column and row of the first random number consulted.

STATISTICIAN'S DIPLOMA EXAMINATION, NOVEMBER 1962

PAPER I : OFFICIAL STATISTICS AND DESCRIPTIVE STATISTICS (THEORETICAL)

Time : 4 hours | Full marks : 100

Figures in the margin indicate full marks.

GROUP A

(Answer *any three* questions from this group)

1. Discuss in detail the role of the Indian National Sample Survey in providng statistical data for purposes of planning in India. (10)

What are the limitations, if any, of these data and what improvements would you suggest for elimininting them? (6)

2. Discuss critically the present position in regard to the statistics of production of crops in India in the light of the improvements that have been made during the past two decades. (12)

What further improvements are necessary for eliminating the present deficiencies? (4)

3. What are the types of data currently included in the official publications of Government of India dealing with foreign trade? Also, name the publications. (10)

Describe briefly, in the context of planning, the limitations, if any, of the publications you name and the statistics you describe. (6)

4. How is the decennial population census organized in India? What were the items included in the questionnaire at the 1961 census and what changes, if any, would you suggest in the same? (12)

What suggestions have you to offer to improve the reliability of the data collected under the census? (4)

Neatness (2)

GROUP B

(Answer *all* questions from this group)

5. Define Binomial and Poisson distributions and indicate by examples their utility. (5)

Find the first four moments of a Binomial distribution and obtain the values of the Beta coefficients of sknowness and kurtosis. (10)

6. Explain clearly the meanings of total, partial and multiple correlation coefficients. (5)

If X_1, X_2, X_3 are three variables for which a sample of n observations (x_1, x_2, x_3) are available, express the value of $R_{1 \cdot 23}$ in terms of r_{12}, r_{13}, r_{23}. (5)

Establish the formula. (5)

$$r_{12\cdot 3} = \frac{r_{12} - r_{13}\, r_{23}}{\sqrt{(1-r_{13}^2)}\sqrt{(1-r_{23}^2)}}$$

Is it possible to obtain the values $r_{12}=0.06$, $r_{13}=0.8$, $r_{23}=-0.5$ for any set of three variables X_1, X_2 and X_3? (5)

7. Explain clearly the method of Least Squares in fitting a straight line to bivariate data. (10)

If $(x_1, y_1), \ldots, (x_n, y_n)$ are n pairs of observations, find the angle between the two lines of regression, one of y on x and the other of x on y. (5)

PAPER II : PROBABILITY THEORY AND STATISTICAL METHODS (THEORETICAL)

Time : 4 hours Full marks : 100

Figures in the margin indicate full marks.

GROUP A

(Answer *any three* questions from this group)

1. (*a*) Show that for any integer m with $1 \leqslant m \leqslant N$, the probability $P_{[m]}$ that exactly m among the N events $A_1, A_2, \ldots, A_N$ occur simultaneously is

$$P_{[m]} = S_m - \binom{m+1}{m} S_{m+1} + \binom{m+2}{m} S_{m+2} + \ldots + \binom{N}{m} S_N$$

where S_i is the sum of $\binom{N}{i}$ probabilities each corresponding to the simultaneous occurrence of some i of the N events. (9)

(*b*) A player tosses a coin and scores one point if head turns up and two points if tail turns up. If P_n is the probability of obtaining exactly n points, find a recurrence relation for P_n and hence its value. (7)

2. State and prove Tchebychev's inequality and show that it cannot be improved in general. (6)

Use the above inequality to prove the weak law of large numbers, assuming the existence of the second moment. (4)

Also prove this law without assuming the existence of the second moment. (6)

3. Prove the statistical independence of the sample mean and variance for a random sample from a univariate normal population. (4)

Also find their distributions. (8)

Find the first two moments about the origin for the distribution of the sample variance. (4)

4. Define the correlation coefficient ρ between two random variables X and Y. Show that the independence of X and Y implies that $\rho = 0$, but the converse is not true. (5)

If X and Y assume only two values each and $\rho = 0$, then prove that X and Y are independent. (8)

If $\rho = 1$, what can you say about the relation between X and Y? (3)

5. (*a*) If X is a discrete random variable having the Poisson distribution

$$g(x) = \frac{m^x e^{-m}}{x!}, \quad x = 0, 1, 2 \ldots$$

and Y is another discrete variable having the conditional distribution

$$f(Y|x) = \binom{x}{y} p^y (1-p)^{x-y}, \quad 0 \leqslant y \leqslant x,$$

Show that the marginal distribution of Y is a Poisson distribution with parameter mp. (10)

(*b*) What transformation will change the random variable X to Y so that Y has the uniform distribution over the interval [0, 1] when the probability density function of X is

$$f(x) = \frac{x-1}{2}, \quad 1 \leqslant x \leqslant 3. \tag{6}$$

Neatness (2)

GROUP B

(Answer *any three* questions from this group)

6. Give a brief description of the theory of confidence intervals as developed by Neyman.

Apply the theory to determine a 95% confidence interval, based on a sample of size n, for the mean of a normal population whose standard deviation is also unknown. Give an outline of the proof. (16)

7. Write a critical note on the method of maximum likelihood in the theory of estimation.

A person is required to shoot at a target till he scores two successes. Suppose he needed n trials. How do you estimate p, the probability of success in a single trial, assuming that the trials are all independent and p remains the same from trial to trial. (16)

8. Dfine the likelihood ratio criterion enunciated by Neyman and Pearson in connection with the testing of statistical hypotheses.

Two independent samples each of size n are drawn from two normal populations. Obtain the likelihood ratio criteria to test the following hypotheses:

(*i*) equality of variances, and

(*ii*) equality of means when the variances are known to be equal *a priori*. (16)

9. Enumerate the various statistical hypotheses which you can test by means of the t-table and indicate in each case the computational procedure by which you can construct the appropriate statistics. Mention in each case the appropriate degrees of freedom under which you would consult the t-table. (16)

Neatness (2)

PAPER III : SAMPLE SURVEYS AND DESIGN & ANALYSIS OF EXPERIMENTS (THEORETICAL)

Time : 4 hours | Full marak : 100

Figures in the margin indicate full marks.

GROUP A

(Answer *any three* questions from this group)

1. A simple random sample of size n is drawn from a population consisting of N units :

(*a*) Define simple random sampling. (2)

(*b*) Is the sample mean a consistent estimate of the population mean? (2)

(*c*) Is the sample mean an unbiased estimate of the population mean? (2)

(*d*) Give the variance of the sample mean and also unbiased estimate of this variance. (2)

(*e*) What is the sample size required to estimate the population mean with a given standard error? (2)

If n_i of the units in the sample are found to be of type A.

(*f*) Give an unbiased estimate of the proportion of units in the population which are of type A. (2)

(*g*) Give the variance of the estimate in (*f*) and also the unbiased estimate of this variance. (2)

(*h*) If the sample size is sufficiently large, write down 95% confidence limits for the unknown proportion of units in the population which are of type A. (2)

2. (*a*) Discuss the merits and demerits of the ratio and regression methods of estimation. (8)

(*b*) Obtain conditions under which the regression method of estimation will give higher precision than the ratio method estimation. (8)

3. A district consists of N villages. To estimate the area under wheat in the district, two different methods are suggested :

(*i*) Select a sample of n villages at random.

(*ii*) From a frequency distribution of the villages into t classes according to the agricultural area of the village. Let the i-th class contain N_i villages $i = 1, 2, \ldots, t$. From the i-th class select a random sample of n_i villages with

$$\sum_{i=1}^{t} n_i = n.$$

(*a*) Give an unbiased estimate of the total area under wheat in the district for each of the two methods. (5)

(*b*) What are the main factors governing the choice of the sample size to be drawn from different classes under method (*ii*)? (5)

(*c*) If n_i is proportional to $N_i (i=1, 2, \ldots, t)$, how do the two methods compare with each other? (5)

4. (*a*) What is systematic sampling? Give the circumstances under which it is to be preferred to simple random sampling. (8)

(*b*) Explain how you will estimate the variance of a systematic sample with a random start. (8)

Neatness (2)

GROUP B

(Answer *any three* questions from this group)

5. (*a*) Describe a Latin Square Design and its analysis, explaining the various assumptions involved.

(*b*) Explain fully the part played by randomisation in such a design.

(*c*) Is there any restriction on the number of treatments that could be accommodated in such a design? Why? (16)

6. In an experiment with 13 treatments the following design in blocks of four has been used :

(2, 7, 6, 9), (8, 13, 12, 2), (11, 3, 2, 5), (6, 11, 10, 13),
(10, 2, 1, 4), (12, 4, 3, 6), (1, 6, 5, 8), (5, 10, 9, 12).
(7, 12, 11, 1), (4, 9, 8, 11), (3, 8, 7, 10), (9, 1, 13, 3), (13, 5, 4, 6).

Identify the type of design and find out its parameters. Describe the method of analysis you would adopt.

Calculate the efficiency factor of the design. (16)

7. One plot is missing in a randomised block experiment. Describe how you would analyse the data including the analysis of variance and tests of significance of differences between pairs of treatment effects.

Also give an estimate of the *loss of information* caused by the missing plot. (16)

8. (*a*) Explain main effects and interactions in the case of a 3^3 factorial experiment.

(*b*) Write down a design for this experiment in blocks of 9 plots partially confounding the highest order interaction. (16)

Neatness (2)

PAPER IV : APPLIED STATISTICS (THEORETICAL)

Time : 4 hours Full marks : 100

(*i*) Answer questions only from the two groups relating to the subjects you have opted for.
(*ii*) Use a separate answer book for each group.
(*iii*) Figures in the margin indicate full marks.

GROUP A—ECONOMIC STATISTICS

(Answer *any three* questions from this group)

1. Write a short essay, on index numbers covering their role in economic analysis and methodology of their construction. (16)

2. (*a*) Describe the components of a time-series and state briefly how you would estimate them. (12)

(*b*) The following regression-equation was found to fit the gross agricultural output series in India (1900-01 to 1950-51) satisfactorily. Give a brief interpretation of the result. All the regression-coefficients were found to be significantly different from zero.

$$Y_t = 4410 + 11.7\,(t - \bar{t}) - 0.576\,(t - \bar{t})^2$$

where Y_t is the repeated moving average of X_t, the agricultural output at time, t, given by

$$Yt = \frac{X_{t+3} + 2X_{t+2} + 2X_{t+1} + 2X_t + 2X_{t-1} + 2X_{t-2} + X_{t-3}}{12} \qquad (4)$$

3. Write a critical note on the recent work done on the estimation of either the demand functions or the production functions on Indian data, covering the data available, functions fitted, methodological short comings or otherwise and interpretation. (16)

4. What are Pareto's and Gibrat's laws of income-distribution? Are they descriptive or explanatory? How far are they adequate for the analysis of income distribution? How are the parameters and concentration-ratios, connected with them, estimated? (16)

5. Write a critical note on the method of estimation of national income in India. (16)

Neatness (2)

GROUP B—STATISTICAL QUALITY CONTROL

(Answer *any three* question from this group)

6. (*a*) Distinguish between specification limits, control limits and confidence limits. (8)

(*b*) What do you understand by a 'rational subgroup'? Discuss its importance in the setting up of a control chart giving an example for each of $\bar{x} - R$, p and c charts. (8)

7. (*a*) The standard deviation 's' of a sample of size n for a characteristic 'x' is obtained as

$$s = \sqrt{\frac{\Sigma(\overline{x - \bar{x}})^2}{n}}$$

Derive its expectation when the parent population is Normal and discuss its usefulness in quality control. (10)

(*b*) State the addition theorem of k independent variances and describe how it can be used to set up realistic tolerances of mating parts. (6)

8. (*a*) What is an acceptance sampling plan? Discuss the role of OC curve in the choice of a suitable plan. (6)

(*b*) Briefly describe the salient points of Mil Std 105A or Mil Std 105B tables. State the procedure to be adopted in the choice of a single sampling plan from the tables for normal inspection and the basis for tightened and reduced inspection while operating these plans. (10)

9. The flow chart of the operations in the manufacture of a ceramics product is as follows:

Mixing → Moulding → Drying →
Glazing → Firing → Sorting inspection

The sorting operation revealed an incidence of 20% cracks in the product and it is desired to carry out a preliminary production experiment to locate the operation or operations that need more intensive investigations.

Clay for moulding is mixed in batches and moulded into proper shape by *three* moulders. There are *two* driers and *one* kiln for firing. Glazing is not expected to contribute to these defects. Each moulder produces aoubt 75 pieces in an hour.

Explain how you will plan the collection and analysis of the data. (16)

Neatness (2)

GROUP C—STATISTICAL METHODS IN GENETICS

(Answer *any two* question from this group)

10. The O—A—B blood groups of a number of children born to mothers of 0 group are obtained. Obtain the expected frequencies of O, A, B and AB groups in children assuming that no choice was exercised by the 0-group mothers in selecting husbands of particular blood groups. The O, A, B gene frequencies for the population of husbands may be taken as r, p, q. (25)

11. Derive the expected frequencies of the four phenotypic classes for the offsprings of the mating $Ab/aB \times Ab/aB$, introducinga linkage parameter. How are the expectations altered if the gene A is lethal in homozygons condition, i.e., offsprings receiving A from the mother as well as the father are not born? Explain how linkage is estimated in such a situation. (25)

12. Explain the purpose of inbreedings. Determine the matrix of transition probabilities of different types of mating under *selfing* considering only one locus with two alleles A, a. (25)

GROUP D—DEMOGRAPHY AND VITAL STATISTICS

(Answer question number 13 and *any two* of the rest from this group)

13. (*a*) Enumerate the principal sources of demographic data, with brief accounts of these sources. (7)

(*b*) "Defects in original (demographic) data fall into two groups—the failure to include all the intended people in the statistics, and errors in the information recorded about the people who *were* included."

(*i*) Examplify the above statement. (3)

(*ii*) Describe the important methods of detecting defects of the above two types. (5)

(*iii*) Describe a method of correcting for "heaping" in age statistics. (3)

14. (*a*) What types of data would you need to directly calculate m_x, the age-specific death rates, and q_x, the probability of death within a year at age x? (4)

(*b*) Given only census data, how would you calculate q_x? (4)

(*c*) Explain the different columns of an abridged life table and describe the successive steps to compute the values in these columns. (8)

15. (*a*) If P_1 and P_2 are the populations at an initial period and a later period respectively, and B, D and M are the number of births, deaths and net migration in the intervening period, then

$$P_2 = P_1 + B - D + M.$$

What are the difficulties in estimating P_2 from P_1 using vital statistics in the above identity? (3)

(*b*) Define a "rate of natural increase". How is this rate calculated? (5)

(*c*) Describe briefly a method for projecting—

(*i*) the total population, and (4)

(*ii*) the age-and sex-wise population. (4)

16. (*a*) Write a brief discussion on the relativity of the concepts of "normal health" and "sickness". (4)

(*b*) What are the different categories of information which you would collect under "morbidity statistics"? (4)

(*c*) Write an account of the various sources of morbidity statistics. (4)

(*d*) Suggest some methods for collecting information on general sickness for the whole population of a country. (4)

GROUP E—EDUCATIONAL AND PSYCHOLOGICAL STATISTICS

(Answer *any two* questions from this group)

17. Describe briefly how you would construct and validate a battery of objective tests which is to be used to select from amongst a large number of applicants a limited number for admission to a post graduate course in statistics.

In an objective test, there are N items of equal importance, and for each item m alternative answers are provided. The candidates are required to choose the best answer (from the m alternatives) for every item. What score would you award to a candidate who has correctly answered n_1 items, wrongly answered n_2 items and left n_3 items unanswered, $n_1+n_2+n_3=N$, so as to correct for the effects of guessing? (10+6)

18. What is "Intelligence Quotient"? How is it measured? What are its uses? What are the drawbacks? (3+7+3+3)

19. N subjects are independently ranked by m judges in respect of the same trait. How would you examine if the judges are concordant? Describe the appropriate statistical procedure, explaining carefully its basis.

If the judges are concordant, how would you assign scores to the subjects in respect of this particular trait? (10+6)

20. Describe briefly Thurstone's multiple factor theory, explaining the significance of the model and the procedure for estimating the factor loadings by the centroid method. (16)

Neatness (2)

PAPER V : METHODS OF NUMERICAL COMPUTATION, DESCRIPTIVE STATISTICS AND OFFICIAL STATISTICS (PRACTICAL)

Time : 5 hours Full marks : 100

(*i*) Figures in the margin indicate full marks.

(*ii*) Use of calculating machines is permitted.

GROUP A

(Answer *any two* questions from this group)

1. Compute the inverse of the following matrix :

$$\begin{vmatrix} 3 & -2 & 0 & -1 \\ 0 & 2 & 2 & 1 \\ 1 & -2 & -3 & -2 \\ 0 & 1 & 2 & 1 \end{vmatrix} \quad (12\frac{1}{2})$$

2. Given $\sin 45° = 0.7071$, $\sin 50° = 0.7660$, $\sin 55° = 0.8192$, $\sin 60° = 0.8660$, find $\sin 52°$ by using any method of interpolation. $(12\frac{1}{2})$

3. By means of a suitable quadrature formula, find the value of π from the relation

$$\frac{\pi}{4} = \int_0^1 \frac{dx}{1+x^2} \text{ (use intervals of 0.1 for } x) \quad (12\frac{1}{2})$$

GROUP B

(Answer question number 4 and *any two* of the rest from this group)

4. *Either,*

The following table gives the female population between the ages 15-49 of Puerto Rico in 1950, in 5-yearly age-groups.

age-group	women enumerated in census, 1960
15—19	101,760
20—24	93,207
25—29	73,857
30—34	57,394
35—39	57,598
40—44	45,161
45—49	34,833

(a) Fit a second degree polynomial to the population as a function of age, assuming that the mid-point of an age-group represents the age-group. (13)

(b) Under the same assumption plot the given data on a graph and show the fitted polynomial on the same graph. (7)

Or,

The following table shows the weight, stature and 'frontal breadth' of 10 persons.

serial (number)	weight (lb.)	stature (cm.)	frontal breadth (mm.)
1	116	165	105
2	89	160	102
3	102	168	100
4	108	158	105
5	88	159	103
6	104	163	100
7	118	172	104
8	100	155	104
9	116	162	110
10	102	157	105

(a) Compute the least square linear regression of weight on stature and frontal breadth. (13)

(b) Estimate the weights of the persons with serial number 1, 3 and 10 from the measuremets of their stature and frontal breadth. (7)

5. The following table shows the frequencies of the different values of a variable x:

x	0	1	2	3	4	5	6	7
f	305	365	210	80	28	9	2	1

(a) Calculate the expected frequencies, assuming a Poisson distribution for x. (10)

(b) Calculate the variance of the given distribution. What should be the value of this variance under the assumption made in (a)? (5)

6. The following is the distribution of the size of an item:

size of item in ft.	frequency
1—10	9
11—20	13
21—30	80
31—40	231
41—50	110
51—60	45
61—70	12

(*a*) Calculate the values of β_1 and β_2 coefficients of this distribution and suggest if a normal distribution would fit the data. (10)

(*b*) Assuming a normal distribution, calculate the expected frequencies in the classes 1-10, 31-40 and 61-70. (5)

7. The table below gives the wholesale prices and quantities produced of a number of commodities in India in 1951 and 1952.

commodity	price (Rs. per md.)		quantity (000 tons)	
	1951	1952	1951	1952
Rice	16.87	17.50	20,960	22,537
Jowar	10.09	11.03	5,981	7,243
Bajra	10.07	13.33	2,309	3,142
mazie	21.75	15.07	2,043	2,825
Ragi	9.45	9.30	1,291	1,315
Wheat	18.60	23.67	6,085	7,382
Barley	20.66	17.29	2,330	2,882
Gram	24.09	19.01	3,334	4,142

Calculate Laspeyre's Paasche's and Fisher's Ideal index numbers of prices in 1952 with 1951 as base. (15)

GRPUP C

(Answer *any one* question from this group)

8. (*a*) Name the publications in respect of *any three* of the following series in which they are available. Give the name of the authority issuing them and also mention their frequency:

(*i*) Statistics of production of crops for different countries of the world (3)
(*ii*) Production, despatches, stocks and distribution of coal in India. (3)
(*iii*) Statistics of infant mortality in India. (3)
(*iv*) Statistics of production of raw silk in India. (3)
(*v*) Statistics relating to area and outturn of forest produce in India. (3)

(*b*) Collect the data in respect of (*i*) and *any two* of the rest in the following, from the publications supplied to you and present them in neat tabular forms. Mention the source from which you have compiled the data.

(*i*) Statistics of land utilization for India for any five consecutive years. (6)
(*ii*) Statistics of installed capacity production and distribution of electricity in India for any four consecutive months. (5)
(*iii*) Number of livestock and poultry in India for any one year. (5)
(*iv*) Statistics of production of ores and minerals in India for any one year. (5)
(*v*) Statistics of railway wagons loaded in India for any six consecutive months. (5)

9. (*a*) From the publications supplied collect figures of area irrigated source-wise in the country for the latest five consecutive years for which these are available and discuss the trends in these areas. Give the source from which you have compiled the data. (13)

(*b*) Collect relevant data regarding production, arrivals, consumption and export of jute for India for the latest twelve months for which these are available and write a critical note on the salient features of these data. Give the source of your data. (12)

PAPER VI : STATISTICAL METHODS, DESIGN & ANALYSIS OF EXPERIMENTS AND SAMPLE SURVEYS (PRACTICAL)

Time : 5 hours Full marks : 100

(*i*) Figures in the margin indicate full marks.
(*ii*) Use of calculating machines is permitted

GROUP A

(Answer both the questions from this group)

1. Answer *any three* of the following :

(*a*) Estimates of the percentage of unemployed persons in a certain city, obtained from two independent random sample surveys are given below. Is the difference in the two estimates due to fluctuations of sampling? (8)

survey	sample size	per cent unemployed
1	2345	19.678
2	1789	20.123

(*b*) The coefficient of correlation between the weight of a green jute plant and the weight of the dry fibre extracted from the plant was found to be 0.72 in a sample of 40 plants. Set up 95% confidence limits for the value of the correlation coefficient in the population. (8)

(*c*) The following table gives certain statistics relating to the lactation yields (x) of 200 dairy cows classified according to manner of milking and to whether they have calved before or not. In each cell the first figure given is the number (n) of cows, the second is the sum (Σx) of the yields of these cows during the lactation period and the third the sum of squares (Σx^2) of the yields (uncorrected).

Is there evidence that yield is lowered by machine milking? If so, is the decrease same for first calvers and others? Do first calvers yield less milk than other cows? (8)

manner of milking	first calvers	others
hand milked	$n = 50$ $\Sigma x = 7156$ $\Sigma x^2 = 1120864$	$n = 50$ $\Sigma x = 8357$ $\Sigma x^2 = 1475865$
machine milked	$n = 50$ $\Sigma x = 6296$ $\Sigma x^2 = 1387003$	$n = 50$ $\Sigma x = 6783$ $\Sigma x = 1000048$

(*d*) The following table gives the result of an experiment to compare the effect of a newly discovered medicine on a certain disease with that of the prevailing treatment (control).

Do the data confirm the superiority of the new medicine?

treatment	cases treated		
	total	cured	not cured
control	8	3	5
new medicine	8	5	3

(8,

2. Answer *any one* of the following :

(*a*) A, B, C, D are four points in order on a straight line. Lengths of different segments were measured to the thousandth of an inch and are given below. These are naturally affected by errors of measurement, which can be assumed to be of the nature of independent random variables, each with zero mean and a common unknown variance.

Estimate the length of the line AD and find the standard error of the estimate. (10+6)

segment	length (in inches)
AB	1.234
AC	3.568
AD	7.024
BC	2.345
BD	5.812
CD	3.456

(*b*) In a certain genetical experiment, the relative frequencies of four different types of pea plants are expected to be $\frac{1}{4}(2+\theta)$, $\frac{1}{4}(1-\theta)$, $\frac{1}{4}(1-\theta)$ and $\frac{1}{4}\theta$ respectively where $0<\theta<1$ is an unknown parameter. The frequencies observed were 102, 21, 32, and 5 respectively.

(*i*) Estimate the parameter θ by the method of maximum likelihood and find the standard error of the estimate. (6+4)

(*ii*) Examine the goodness of fit of the distribution. (6)

GROUP B

(Answer *any two* questions from this group)

3. The lay-out plan and yield in lbs. of a manurial experiment on sugarcane crop involving 5 manures A, B, C, D, E and conducted using a Latin square design are given below:

E 805	B 843	C 749	D 1029	A 807
B 1011	D 1016	E 749	A 913	C 740
A 898	C 905	D 674	E 667	B 563
C 670	E 1100	A 770	B 812	D 1195
D 769	A 727	B 723	C 796	E 740

(*a*) Perform the analysis of variance of the above yield data.

(*b*) Given that the size of each plot is 1/40 acre, present in a tabular form the mean yield of cane in tons per acre for each manure and its standard error.

(*c*) Write a brief report on the results of the analysis. (15)

4. The lay-out plan and yield of each plot of a $2\times2\times2$ factorial experiment involving the factors *N*, *P* and *K* [the control being denoted by (1)] are given below:

block								
I	*nk* 159	(1) 179	*pk* 135	*np* 130	*p* 153	*npk* 202	*n* 153	*k* 182
II	*npk* 155	*np* 129	*k* 151	(1) 159	*n* 191	*nk* 138	*p* 188	*pk* 210
III	*p* 154	*nk* 77	(1) 92	*npk* 128	*k* 143	*np* 119	*n* 115	*pk* 179
IV	*n* 113	*pk* 127	*npk* 185	(1) 148	*p* 136	*nk* 107	*k* 182	*np* 212

Analyse the data and state your conclusions. (15)

5. The following table gives data on number of beets per plot and yield in tons per acre in a manurial experiment using six randomised blocks of 4 plots each. The upper figure in each cell gives the number of beets and the lower figures the yield.

treatment	block					
	1	2	3	4	5	6
(1)	183 2.45	176 2.25	291 4.38	254 4.35	225 3.42	249 3.27
p	356 6.71	300 5.44	301 4.92	271 5.23	288 6.74	258 4.74
k	224 3.22	258 4.14	244 2.32	217 4.42	192 3.28	230 4.00
pk	329 6.34	283 5.44	308 5.22	326 8.00	318 6.96	318 6.96

Construct the table of analysis of variance and covariance and test whether the treatments had significant effects on (*i*) either of the two varietes, (*ii*) on yield adjusted for number per plot. Explain whether you are justified in doing (*ii*).

Write also a brief report containing interpretation of the results. (15)

GROUP C

(Answer *any two* questions from this group)

6. The following table gives the population in hundreds of all the 58 villages in a tehsil.

serial number of village	population	serial number of village	population	serial number of village	population
1	16	21	47	41	68
2	26	22	39	42	77
3	38	23	53	43	11
4	76	24	9	44	86
5	51	25	18	45	79
6	76	26	54	46	56
7	81	27	39	47	72
8	85	28	62	48	29
9	28	29	70	49	24
10	68	30	59	50	82
11	100	31	32	51	97
12	75	32	17	52	43
13	82	33	23	53	45
14	12	34	61	54	37
15	20	35	87	55	81
16	52	36	29	56	27
17	15	37	93	57	74
18	21	38	35	58	68
19	63	39	68		
20	58	40	43		

(*a*) If the total population of the tehsil is estimated from a simple sample of 6 villages drawn with replacement, what would be the sampling variance of the estimate?

(*b*) If the tehsil is divided into two strata in the following way:

Stratum 1: villages with population not exceeding 50 (hundred)

Stratum 2: all other villages.

and from each stratum a simple sample of 3 villages is drawn with replacement, and the total population of the tehsil is estimated from the samples drawn this way, what would be the variance of the estimate? (15)

7. To estimate the total number of beetles in a square plot of land 250 feet, the plot was divided into 625 square cells, each 10 feet by 10 feet. A simple random sample of 10 cells were selected, and for each selected cell two independent estimates of the number of beetles were obtained by actual enumeration of two randomly located circles of equal area within the cell. The results are given below:

cells	1	2	3	4	5	6	7	8	9	10
estimates	61	38	25	20	71	95	32	39	24	0
	39	51	32	45	56	79	26	60	29	5

Estimate the total number of beetles in the plot and find the standard error of the estimate. (15)

8. In a two stage sampling with replacement scheme the variance of the sample mean $V = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_1 n_2}$ where n_1 is the number of first stage units and n_2 the number of second stage units per first stage unit. The cost of carrying out such a survey is $T = a + b\,n_1 + c\,n_1\,n_2$

The values of the constants σ_1, σ_2, a, b, c as estimated from a pilot survey, are as follows:

$\sigma_1 = 16.5$, $\sigma_2 = 28.7$

$a =$ Rs. 1527.00 $b =$ Rs. 6.40 $c =$ Rs. 1.24

A preliminary estimate of the population mean is $\mu = 150$.

Determine the optimum values of n_1, n_2 (in the sense of minimising the cost) so that the estimate of the population mean will have a standard error not exceeding 5% of its value. Find also the corresponding cost. (15)

PAPER VII: APPLIED STATISTICS (PRACTICAL)

Time: 5 hours Full marks: 100

(*i*) Answer questions only from the two groups relating to the subjects you have opted for.

(*ii*) Use a separate answer book for each group.

(*iii*) Figures in the margin indicate full marks.

(*iv*) Use of calculating machines is permitted.

GROUP A—ECONOMIC STATISTICS

(Answer *any two* questions from this group)

1. From the data given below calculate net national product at market prices, private income and consumer-expenditure on goods and services for India for the years 1950-51 to 1957-58. Also work out an estimate of the propensity to consume of the private-sector. Estimate the rates of growth of net national product at market prices and of private income. Interpret the results. (25)

	Item	in current prices in Rs. Abjas			
		1950-51	1951-52	1952-53	1953-54
1	Net domestic product at factor cost	99.5	99.9	98.3	104.8
2	Earned Income from abroad	—0.2	—0.2	—0.1	0.0
3	Indirect taxes including miscellaneous fees	5.4	6.3	5.6	5.8
4	Subsidies	0.4	0.4	0.4	0.2
5	Income from domestic product accruing to Government	0.7	0.8	0.6	0.5
6	National debt interest	0.4	0.4	0.4	0.4
7	Transfer payments	0.6	0.6	0.7	0.8
8	Depreciation	4.2	4.5	4.7	5.0
9	Gross capital formation including business inventories	10.4	10.5	11.1	11.2
10	Government current expenditure on goods and services	5.6	5.9	6.0	6.4
11	Net private donations from abroad	0.4	0.4	0.5	0.4

	Item	in current prices in Rs. Abjas			
		1954-55	1955-56	1956-57	1957-58
1	Net domestic product at factor cost	96.1	99.8	113.0	114.0
2	Earned Income from abroad	0.0	0.0	0.1	—0.1
3	Indirect taxes including miscellenous fees	6.3	7.0	8.0	9.4
4	Subsidies	0.1	0.2	0.3	0.5
5	Income from domestic product accruing to Government	0.6	0.7	0.9	0.9
6	National debt interest	0.4	0.5	0.5	0.6
7	Transfer payments	0.9	1.3	1.4	1.5
8	Depreciation	5.1	5.7	5.5	6.4
9	Gross capital formation including business inventories	12.2	14.7	19.1	19.7
10	Government current expenditure on goods and services	6.7	7.3	7.9	9.2
11	Net private donations from abroad	0.3	0.4	0.5	0.5

2. The following table gives indices of all food-grains production, total agricultural production, cotton-textiles production and total industrial production for India for the years 1950-51, 1951-52 to 1960-61 (base year 1949-1950 = 100).

Fit the trends $X = ab^t$ to the four series and estimate the rates of growth and compare. Which series is more adequately explained by the trends and why? Project the four indices for 1961-62 to 1964-65 and comment. (25)

years	all foodgrains index	all agricultural production index	cotton textiles index	all industrial production index
1950–51	90.5	95.6	—	—
1951–52	91.1	97.5	109.3	112.2
1952–53	101.1	102.0	116.6	113.4
1953–54	119.1	114.3	123.7	117.3
1954–55	115.0	117.0	126.7	129.1
1955–56	115.3	116.8	127.9	138.6
1956–57	120.8	124.0	134.7	151.0
1957–58	107.9	114.6	127.3	152.5
1958–59	130.1	132.3	123.4	157.2
1959–60	124.3	127.2	127.5	175.0
1960–61	131.6	135.0	133.1	194.3

3. The following table gives index number of wholesale prices of food articles for India from January 1947 to January 1951. (Base year ending August 1939 = 100). Calculate the indices of monthly variation by the moving average method and comment. (25)

month	indices of wholesale prices of food articles in (years)				
	1947	1948	1949	1950	1951
January	290.6	347.7	385.3	379.1	413.5
February	290.1	348.5	383.8	395.3	
March	286.2	347.1	376.5	396.2	
April	279.9	348.8	373.8	399.8	
May	278.6	357.6	377.0	401.2	
June	287.3	377.0	381.6	402.8	
July	291.4	390.7	395.9	422.8	
August	297.4	397.7	410.6	426.5	
September	296.2	396.6	403.1	430.4	
October	295.2	393.1	406.8	427.3	
November	294.8	394.1	405.1	424.3	
December	321.6	397.5	374.1	423.9	

4. The following table gives the estimated number of households distributed by size-level of household ownership holding in South India (N.S.S. Reference-period July 1953-June 1954 major crop season).

Fit a Pareto curve when (*i*) the first class (landless class) is included and (*ii*) it is not included; comment. Hence or otherwise obtain an estimate of the concentration-ratio in both the cases. (25)

area of household ownership (acres)	number of households (000)	area of household ownership (acres)	number of households (000)
.00	3801	10.00–14.99	468
.01– .04	1484	15.00–19.99	250
.05– .09	332	20.00–24.99	113
.10– .49	860	25.00–29.99	64
.50– .99	1031	30.00–39.99	71
1.00–1.49	779	40.00–49.99	31
1.50–2.49	1161	50.00–74.99	50
2.50–4.99	1519	75.00–99.99	7
5.00–7.49	962	100.00–249.99	20
7.50–9.99	565	25.00–	4

GROUP B—STATISTICAL QUALITY CONTROL

(Answer *any two* questions from this group)

5. 15 spinning frames running the same quality of yarn were each observed 20 times at random intervals for ends down (number of spindles not producing yarn due to an unmended break). A set of 5 frames are fed from one finisher. There are 3 finishers, designated A, B, C. The total number of ends down in the 20 rounds are given below for each of the 15 frames.

finisher	frame	total ends down in 20 rounds
A	1	52
	2	65
	3	73
	4	84
	5	47
B	6	83
	7	78
	8	75
	9	68
	10	92
C	11	55
	12	59
	13	64
	14	59
	15	50

It is desired to locate the frames and the finishers that may need corrective action. Make a suitable control chart analysis of the data and give your findings. (25)

6. (*a*) The average range of 20 samples each of size 5 for a measurable quality characteristic was found to be 4.0. Obtain—

(*i*) an estimate of the process standard deviation.

(*ii*) 2.5% and 97.5% limits for the sample ranges. (10)

(b) Samples of size 5 are drawn from 3 machines on different occasions and the average ranges of the samples computed for each machine are given below.

Test whether the average ranges are significantly different. (15)

machine	number of samples	average range
1	16	0.536
2	12	0.921
3	15	0.626

7. (a) Draw the OC curve of the single sampling plan with

$$n = 100 \quad \text{and} \quad c = 4$$

Draw also the curve for the same plan if a good item may be misclassified with a probability of 0.01. (15)

(b) The minimum requirement of the gross weight (contents + tin) of a filled tin is 580 gm.

It is desired to instal a suitable sampling scheme with the following properties.

Specify a single sampling variable plan for the purpose. (10)

AQL	= 2.5%	Producer's risk	= 0.05
LTPD	= 8.0%	Consumer's risk	= 0.10

GROUP D—VITAL STATISTICS AND DEMOGRAPHY

(Answer *any two* questions from this group)

8. The following table shows the fertility of Australian women in 1932-34, in 5-year age-groups.

age-group	number of females (30 June, 1953)	yearly total births (1932-34)	number of females in life-table population (1932-34)
15—19	303,619	7,075	4,681.00
20—24	286,612	28,448	4,640.23
25—29	256,509	31,045	4,586.00
30—34	237,663	22,866	4,522.70
35—39	237,498	14,339	4,447.63
40—44	226,473	5,665	4,358.48
45—49	199,389	521	4,246.75

(a) Calculate the gross and net reproduction rates, it being known that the yearly total number of female births during the period in question was 53,898.

(b) The annual rate of growth was found to be $r = -0.00159$. Find the initial stable population figures in the age-groups 15-19, 20-24, 25-29 (25)

9. The following table gives the population of the U.S.A. at each decennial census, from 1860 to 1950.

year	population (millions)
1860	31.443
1870	38.558
1880	50.156
1890	62.948
1900	75.995
1910	91.972
1920	105.711
1930	122.775
1940	131.669
1950	150.697

Fit a simple logistic curve to the above data and find the maximum value which the population would be expected to reach. (25)

10. The following table gives the mortality rates ${}_1q_0$ and ${}_nq_x$ $(x=1, 5, ..; n=5)$ for females in Chile, 1940.

age	${}_nq_x$	age	${}_nq_x$
0	.18848	40	.05558
1	.10276	45	.06222
5	.01688	50	.07818
10	.01626	55	.10498
15	.03309	60	.14179
20	.04352	65	.20558
25	.04580	70	.28596
30	.04721	75 और	1.0
35	.05063		

Calculate e_x, the expectation of life at each age x of the above table. (25)

GROUP E—EDUCATIONAL AND PSYCHOLOGICAL STATISTICS

(Answer *any two* questions from this group)

11. A test consisting of 5 items was applied on a large number of subjects and their scores (Y) on a reliable criterion were also recorded. The following table gives the values of—

p_i = proportion of subjects answering the i-th item correctly.

p_{ij} = proportion of subjects answering both the i-th and the j-th items correctly.

μ_i = average criterion score for those subjects who have answered the i-th item correctly.

i \ j	values of p_{ij} 2	3	4	5	p_i	μ_i
1	0.68	0.56	0.40	0.50	0.75	12.4
2	—	0.52	0.35	0.48	0.70	14.7
3	—	—	0.28	0.40	0.60	17.9
4	—	—	—	0.20	0.45	19.2
5	—	—	—	—	0.55	16.8

The average criterion score is $\mu = 16.5$ and the standard deviation of the criterion score is $\sigma = 2.7$.

Let X denote the total number of items correctly answered by a subject. Calculate the mean and the variance of X and the covariance of X and Y.

Hence obtain the correlation coefficient between X and Y.

If you are required to cut down the length of the test to only 4 items, which one of the items will you omit and why? (25)

12. The following table gives the intercorrelations of scores for a battery of 5 tests administered to group of persons.

On the assumption that there exist only two common factors, estimate the factor loadings by Thurstone's centroid method.

Do you think the complex can be explained more or less adequately by postulating the existence of only a single factor instead of two? (25)

tests	1	2	3	4	5
1	1.000	0.915	0.830	0.918	0.892
2	—	1.000	0.909	0.951	0.939
3	—	—	1.000	0.970	0.922
4	—	—	—	1.000	0.949
5	—	—	—	—	1.000

13. Letter grades A, B, C, D and E were assigned to a group of 85 trainees by three teachers in a dictation test—the students receiving the grade A being the best and those receiving the grade E the worst group, B. C. D being the intermediate groups in that order. The table below shows the number of students placed in each grade by the different teachers.

grade	teacher 1	teacher 2	teacher 3
A	25	15	10
B	21	28	17
C	32	24	27
D	6	12	22
E	1	6	9
	85	85	85

Rank students $S_1, S_2, \ldots, S_5$ in order of merit, who got the following grades from the teachers:

student	teacher 1	teacher 2	teacher
S_1	B	A	D
S_2	A	D	B
S_3	D	B	A
S_4	E	D	B
S_5	A	C	E

(25)

PAPER VIII : SUBJECTS OF SPECIALISATION—I

Time : 4 hours | Full marks : 100

(*i*) Answer questions only from the section on the subject of specialisation you have opted for.

(*ii*) Figures in the margin indicate full marks.

(*b*) TECHNO-COMMERCIAL STATISTICS—STATISTICAL QUALITY CONTROL

(Answer *any four* questions from this section)

1. What is a continuous sampling plan? Describe Dodge's continuous sampling plan and derive the quantities which go into the formula for AOQ of the plan. (25)

2. Derive the sequential probability ratio test from the acceptance sampling point of view.

Describe the sequential procedure for testing whether the proportion of defectives produced by a given process is equal to that produced by a different process. (25)

3. (*a*) Enumerate the different types of charts available to control the mean of a measurable characteristic. Indicate the circumstances under which you would prefer to use each of these. (15)

(*b*) Describe briefly the procedure you would adopt to control the number of defects per unit of a product where the defects have to be classified according to their seriousness. (10)

4. (*a*) Describe the method of steepest ascent and local exploration for determining the optimum operating conditions of a production process. (15)

(*b*) Explain the terms Rating, Normal time and Standard time as applied to time study. Give one or two situations when work sampling can be more profitably employed than actual time study. (10)

5. Write short notes on :

(*i*) Statistical tolerances.

(*ii*) Fractional factorial designs.

(*iii*) Lot plot techniques. (25)

(*d*) DESIGN AND ANALYSIS OF EXPERIMENTS—STATISTICAL ASPECTS

(Answer *any five* questions from this section)

1. It is desired to examine whether the average difference between morning and evening stature of a population is zero or not. If two anthropologists are available, how do you design an experiment to obtain the desired information taking into account that differences may exist in the measurements taken by the two anthropologists and also interaction between investigators and times of measurement. Write down the analysis of variance appropriate to analyse the data collected according to your design. (20)

2. A *pani maharaj* (water diviner) claims that he can detect the existence of water between 10 and 20 feet below the point where he stands. Design an experiment to test his claim. Explain clearly the various principles of design of experiments you would employ in working out a suitable design. (20)

3. An experiment has to be conducted to compare the yields of two varieties of rice K and D in combination with four treatments t_1, t_2, t_3 and t_4 defined as follows:

t_1 = Potash at 15 lbs + Phosphoric acid at 12 lbs
t_2 = Potash at 15 lbs
t_3 = Phosphoric acid at 12 lbs
t_4 = No nutrient.

Suggest a suitable design for the above experiment in eight blocks of four plots each. Indicate how you would analyse the experimental data. (20)

4. Describe how uniformity trial data are used to determine the shape and size of experimental plots. (20)

5. What is meant by partial confounding in a factorial experiment? Give a balanced scheme of partial confounding in a 2^3 factorial experiment in block of 8 plots. Explain the computation of partially confounded interactions. (20)

6. What are the combinatorial requirements of a balanced incomplete block design? Show that in such a design the number of blocks cannot be less than the number of varieties.

Show how the analysis of such a design is altered if it is known after the experiment is laid out that two of the varieties used in the experiment are not really distinct, but the same variety called by two different names. (20)

(*c*) SAMPLE SURVEYS—THEORETICAL ASPECTS

(Answer *any four* questions from this section)

1. A finite population is divided into N clusters each containing M elements. The parameter to be estimated is the population mean μ. A sample of n clusters is selected at random and from each selected cluster m elements are chosen at random. Obtain an unbiased linear estimate of μ, the variance of the estimate and an unbiased estimate of the variance. (Clusters and elements were selected with replacement). (15)

Obtain the optimum sampling and sub-sampling fractions when C_1 is the cost of sampling a primary unit, C_2 is the cost of sampling a secondary unit and $C = C_1 n + C_2 mn$ is the total fixed budget available for the survey. (10)

2. State and prove the principal conditions which must be satisfied for a ratio estimate to be the best unbiased linear estimate of the population total. Show that this estimate is consistent. Under what conditions is its efficiency higher than that of an estimate based on sample average. (25)

3. In a two-stage sampling design, consider the following selection procedures (selection being done with replacement):

(*i*) Select n primary units with probabilities proportional to their size M_i (the number of elements in the ith unit) and m elements from each primary unit selected.

(*ii*) Select *n* primary units with equal probability and a constant proportion of elements in each primary unit selected such that the overall sampling fraction is the same as in (*i*)

Describe the optimum procedures to be followed in each case for estimating the population total of a character. Derive the expression for the estimate proposed in each case. Comment on the relative usefulness of the two selection procedures suggested. (25)

4. (*a*) Discuss with suitable illustrations the importance of "cost" and "variance" functions in planning a large-scale sample survey. (12)

(*b*) What are non-sampling errors? How do you distinguish between sampling and non-sampling errors? What methods would you suggest to assess and to control non-sampling errors ? (13)

5. (*a*) In a finite population, an increase in sample size operates in two different ways to reduce the sampling error. What are these two ways and which is ordinarily the more important? (9)

(*b*) Suppose there are 500 rocks of varying sizes and weights on a table before you. If you wanted to estimate the mean weight of the rocks on the basis of five rocks, how would you select the five? Why ? If you wanted to estimate the mean on the basis of 50 rocks, would you change your method of selection? Why? (16)

PAPER IX : SUBJECTS OF SPECIALISATION—II

Time : 4 hours Full marks : 100

(*i*) Answer questions only form the section on the subject of specialisation you have opted for.

(*ii*) Figures in the margin indicate full marks.

(*b*) TECHNO-COMMERCIAL STATISTICS

(*i*) Answer Group A and either Group B_1 or Group B_2 whichever you have opted for.

(*ii*) Use a seperate answer book for each group.

GROUP A : STATISTICAL METHODS IN BUSINESS

(Attempt *any two* questions from this group)

1. Write a note on how and which sampling techniques are useful in auditing. (15)

2. Develop (*i*) 'the basic exponential model' for forecasting (*ii*) the exponential model with ratio seasonals and (*iii*) the exponential model with ratio seasonals and trend. (12)

How would you estimate the weights and other parameters involved in the model? (3)

3. (*a*) What are the different steps in personnel selection? Where do aptitude and achievement tests, work-measurement, job-analysis come in personnel selection? (9)

(*b*) Describe the points system used for job-evaluation. What are the pre-requisities for the system? Explain also how job-evaluation is necessary for payment by results. (6)

GROUP B_1 : OPERATIONS RESEARCH

(Answer *any four* questions from this group)

4. Explain the concept, scope and tools of operations research as applicable to business and industry. (17)

5. Describe the linear programming technique covering the formulation, assumptions and some simple methods of solution. Illustrate how it is a tool for taking quantitative management decisions in business and industry. How is the solution of the dual programming problem relevant to management decisions? (17)

6. Dynamic programming is supposed to characterise the joint application of two ideas: (1) recursive optimality and (2) summarisation of a "state variable". Explain the two ideas in detail and explain also how they are useful in inventory-control problems, using a hypothetical example to illustrate the argument. (17)

7. (*a*) Work out answers to the two questions (*i*) and (*ii*) below in the following problem: (12)

A shipping company has a single unloading bearth with ships arriving in a Poisson fashion at an average rate of three per day. The unloading time-distribution for a ship with n unloading crews is found to be exponential with average unloading time $\frac{1}{2n}$ days. The company has a large labour supply without regular working hours, and to avoid long waiting lines the company has a policy of using as many unloading crews on a ship as there are ships waiting in line or being unloaded.

(*i*) Under these conditions what will be the average number of unloading crews working at any time?

(*ii*) What is the probability that more than 4 crews will be needed?

(*b*) If C is the purchase price of a machine, R_n its running cost in year n, i the annual interest rate, then work out the year in which the machine is to be replaced if the owner is assumed to borrow his expenditure at the same interest rate i and he repays it making fixed annual payments throughout the life of the machine. (5)

8. Derive Harris's economic lot size formula specifying clearly the conditions under which it is derived. Are these conditions realistic? How are these conditions modified to present a more realistic analysis? (17)

9. Write an explanatory note on the role of the Monte Carlo method in operations research illustrating it if possible by a hypothetical example in operations research. (17)

Neatness (2)

10. *Either,*

Tabulate the following items and state opposite each (*i*) class of account, (*ii*) whether it usually has a debit or credit balance, (*iii*) whether the balance appears in the Trading Account, the Profit and Loss account or Balance Sheet.

(*a*) Prepaid Rent.
(*b*) Discount Received.
(*c*) Carriage Inwards.
(*d*) Salaries and Wages.
(*e*) Interest on Bank Overdraft.
(*f*) Bank Overdraft. ($6 \times 2 = 12$)

Or,

(*a*) What are the advantages of Imprest system of Petty cash? (4)

(*b*) Give the form of a Petty Cash Book on Imprest system, enter therein Ten imaginary figures over a fortnight taking imprest to be Rs. 100 and show at the end of each succeeding week recoupment of Imprest. (8)

11. The following facts relate to the business of Mr. Roy on a particular date :

Balance as per Cash Book (Over drawn)	..	Rs. 80
Unpresented Cheques	..	Rs. 144
Uncredited Cheques	..	Rs. 26

The following further particulars are given :

(*i*) A cheque for Rs. 20 paid to Mr. Peck has been entered erroneously in the cash column.

(*ii*) Bank commission of Rs. 8 has not been accounted for, and

(*iii*) It is revealed that the debit side of the Cash Book (Bank column) has been undercast by Rs. 50.

What is the balance as per Pass Book of Mr. Roy as of that date? (8)

12. *Either,*

(*a*) Write explanatory notes on the following : ($1\frac{1}{2} \times 4 = 6$)
(*i*) Liquid Assets, (*ii*) Documentary bill, (*iii*) Watered capital, and (*iv*) Errors of principle.

(*b*) How will you proceed to detect the difference if the Trial Balance does not agree? (6)

Or,

(*a*) Clearly distinguish between 'Capital Expenditure' and 'Revenue Expenditure'. (6)

(*b*) State how the following items will be treated in the Books of Accounts :
(*i*) Construction of a students Common Room Rs. 20,000
(*ii*) Repairs to Benches and Tables Rs. 150
(*iii*) White-washing the buildings Rs. 1500, and
(*iv*) Pulling out an old structure Rs. 750 and Re-building it Rs. 7500.
($1\frac{1}{2} \times 4 = 6$)

13. On January 3, *Q* purchased from *P* goods to the value of Rs. 6000 and settled the account by means of three bills of Exchange for Rs. 2000 each, due respectively in two, three and four months. A week later *P* discounted the first of the three bills with his Bank, receiving Rs. 1990, the others he held. The first two bills were paid at maturity. The third *Q* was unable to meet and it was retired by arrangements. *Q* paying Rs. 500 in cash and giving *P* a fresh bill for three months to cover the balance with interest at the rate of 5% per annum. This bill was duly met at maturity.

Journalise the above transactions in *P*'s books of Accounts. (14)

14. The following balances stood in the books of K. C. Joseph as at 31.12.61. You are required to prepare a Trading and Profit and Loss account for the year ended 31.12.61 and a Balance Sheet as at that date, after making the necessary adjustments :

	Rs.
Mr. Joseph's Capital Account	80,000/-
Mr. Joseph's Drawing Account	6,000/-
Plant and Machinery ..	20,000/-
Plant and Machinery (additions on 1.7.61)	5,000/-
Stock on 1.1.61 ...	15,000/-
Purchases	82,000/-
Returns Inwards ..	2,000/-
Sundry Debtors ..	20,600/-
Furniture and Fixtures ..	5,000/-
Frieght and Duty ..	2,000/-
Carriage Outwards ..	500/-
Rent, Rates and Taxes ..	4,600/-
Printing and Stationery ..	800/-
Trade Expenses ..	400/-
Sundry Creditors ..	10,000/-
Sales ..	120,000/-
Returns Outwards ..	1,000/-
Postage and Telegrams ..	800/-
Reserve for Doubtful Debts ..	400/-
Discounts Received ..	800/-
Rent of premises sub-let up to 30.6.62 ..	1,200/-
Insurance charges ..	700/-
Salaries and Wages ..	21,300/-
Cash in Hnad ..	6,200/-
Cash at Bank ..	20,500/-

ADJUSTMENTS

(*i*) Stock on 31.12.61 was valued at Rs. 14,600/-.
(*ii*) Write off Rs. 600/- as Bad debts.
(*iii*) The Reserve for Doubtful Debts is to be maintained at 5% on S/Debtors.
(*iv*) Create Reserves for Discount on Debtors and Discount on Creditors at 2%.

(v) Depreciate Furniture and Fixture and Plant and Machinery at the rate of 5% and 20% per annum respectively.

(vi) Insurance Prepaid was Rs. 100/-

(vii) A fire occurred on 25.12.61.in the godown and stock of the value of Rs. 5,000/- was destroyed. The Insurance Company admitted the claim in full.

(Make suitable assumptions, if necessary) (24)

(d) DESIGN & ANALYSIS OF EXPERIMENTS—COMBINATORIAL ASPECTS

(Attempt *any four* questions from this section)

1. Define projective and Euclidean geometries of dimension m over a finite field $GF(p^n)$. (15)

Obtain the expressions for the numbers of points and lines in $PG(2, p^n)$ and $EG(2, p^n)$ and show that $EG(2, p^n)$ can be embedded in $PG(2, 2^n)$. (10)

2. Define a balanced incomplete block design (BIBD) and obtain the relations between $(v, b, r, k$ and $\lambda)$. (3)

Let $N=(n_{ij})$ denote the incidence matrix of a BIBD where $n_{ij}=1$ or -1 according as treatment i occurs in block j or not. Show that the rows of N are orthogonal if and only if $b=4(r-\lambda)$. (4)

Let M and N denote incidence matrices for two BIBD's with parameters $(v_1, b_1, r_1, k_1, \lambda_1)$ and $(v_2, b_2, r_2, k_2, \lambda_2)$ respectively with $b_i=4(r_i-\lambda_i)$, $i=1, 2$. Then show that $[M \times N = m_{ij}(N)]$ also corresponds to a BIBD with (v, b, r, k, λ) where $b=4.(r-\lambda)$. (15)

Hence obtain a solution of $v=b=16$, $r=k=6$, $\lambda=2$. (3)

3. Define a partially balanced incomplete block design (PBIBD). (4)

Show that the dual of a BIBD with $\lambda=1$ is a PBIBD with two associate classes. (12)

If in a PBIBD with two associate classes and parameters $v, b, r, k, \lambda_1, \lambda_2, n_1, n_2, p^i_{jk}$ we have $p^1_{11}=p^2_{11}=\lambda$, then show that we can construct a BIBD with $v=b$, $r=k=n_1$, λ. (9)

4. Define an orthogonal array $[\lambda\ S^2, k, S, 2]$ in S symbols, k constraints, strength 2 and index λ and prove that k satisfies.

$$k \leq \frac{\lambda S^2-1}{S-1}$$

Prove a one-to-one correspondence between $[\lambda\ S^2, k, S, 2]$ and the group divisible design with k groups of s treatments. $v=kS$, $b=\lambda S^2$, $r=\lambda S$, k, $\lambda_1=0$, $\lambda_2=1$. (25)

5. How many factors each at 3 level can be accomodated in blocks of 9 plots without confounding any main effects or 2-factor interactions? Obtain such an

arrangement using the maximum number of factors in two replications partially confounding sum of the 3-factor interactions. (25)

6. What is a rotatable design?

Give a method of constructing a rotatable design of the second order in three dimensions.

(25)

(c) SAMPLE SURVEYS—ORGANIZATIONAL ASPECTS

(Answer *any four* questions from this section)

1. On what practical considerations will you base the design of a sample survey for estimating the cost of production of an important crop in a State? Give the procedure of collection of data which would minimize non-sampling errors and provide estimates with maximum precision and minimise cost. What items of information would you like to include and what alternate results would you like to present? (25)

2. Draw up a schedule of inquiry into the volume of rural savings in India with a view to ascertaining the resources available in the rural area for the third five year plan. Give your own definitions and concepts, with justification, and draw up a set of instructions to field workers. (25)

3. You are required to study the changes in the economic conditions of the people of your State during the period of the second five year plan. What sampling unit will you use and what type of frame will you choose? What type of information will you collect and what method or methods of inquiry will you adopt? (25)

4. Discuss the role of a pilot survey in taking decisions on the appropriate method of inquiry to be adopted in a large-scale sample survey, preparing the schedules and instructions and constructing cost and variance functions. (25)

5. In an integrated scheme of field survey, it is proposed to cover within one year the following type of inquiries:

(a) Estimation of area and average yield of principal crops.

(b) Size of agricultural holdihngs.

(c) Number of livestock and agricultural implements.

Indicate, with proper justification, the choice of staff requirements (nature, qualification, composition, etc.) necessary for an overall efficient execution of the scheme.

State clearly the details of the scheme of inspection proposed to be followed. (25)

6. Write short notes on *any three* of the following:

(a) Schedules and questionnaires. (8½)

(b) Follow-up of non-response cases. (8½)

(c) Pre-tabulation scrutiny. (8½)

(d) Use of punch-card equipment. (8½)

(e) Training of field staff. (8½)

PAPER X: SUBJECTS OF SPECIALISATION—III

Time: 5 hours Full marks: 100

(*i*) Answer questions only from the section on the subject of specialisation you have opted for.
(*ii*) Figures in the margin indicate full marks.
(*iii*) Use of calculating machines is permitted.

(*a*) ECONOMIC STATISTICS (PRACTICAL)

(Attempt *any four* questions)

1. The following table gives the per capita annual consumption of cereals (y), per capita disposable personal income (x) and wholesale price-indices for all commodities (p) and for cereals (p_1) for India. Take the consumer-behaviour relationship as

$$y = ax^b \left(\frac{p_1}{p}\right)^c$$

and work out the income and price-elasticities for cereal consumption. Are they significantly different from zero? (25)

years	per capita annual consumption cereals (lb. per annum)	per capita disposable personal income (constant prices) (Rs.)	wholesale* price indices all commodities base year 1939=100	wholesale* price indices cereals base year 1939=100
1950-51	273.5	256.3	400.7	471.2
1951-52	274.3	258.3	439.3	483.3
1952-53	262.2	253.0	386.9	450.2
1953-54	285.7	262.1	393.9	450.6
1954-55	298.7	253.7	386.7	403.8
1955-56	303.8	269.0	355.4	354.7
1956-57	306.5	272.2	403.1	472.2
1957-58	296.6	257.7	341.2	538.4

*(the whole-sale price-indices are for the calendar year but may be assumed to correspond to the years given at the olft).

2. The following table gives data on public investment under Central Government for the years 1898-99 to 1928-29. Fit a parabola of the second degree to the above data and comment. (25)

year	public investment under central government (in crores of rupees)	year	public investment under Central Government (in crores of rupees)
1898–99	14.53	1914–15	35.32
1899–1900	14.15	1915–16	39.85
1900–01	16.60	1916–17	45.91
1901–02	17.45	1917–18	48.17
1903–04	20.35	1918–19	46.85
1904–05	21.42	1919–20	40.90
1905–06	23.51	1920–21	36.33
1906–07	29.97	1921–22	43.80
1907–08	29.28	1922–23	51.80
1908–09	34.23	1923–24	42.82
1909–10	33.00	1924–25	50.95
1910–11	27.39	1925–26	36.15
1911–12	32.28	1926–27	22.62
1912–13	31.37	1927–28	13.22
1913–14	35.37	1928–29	12.36

3. Quality-ratios (q_1/q_2) and price-ratios (p_1/p_2) of motor-car imported into India from U.K. and U.S.A. are given below. Work out the least squares estimate of the elasticity of substitution defined by $\frac{d \log (q_1/q_2)}{d \log (p_1/p_2)}$ and assumed to be constant. Comment. (25)

year	U.K./U.S.A. ratio of quantity indices q_1/q_2	U.K./U.S.A. ratio of price-indices p_1/p_2
1921–22	0.9849	2.2620
1922–23	0.3239	2.3072
1923–24	0.3507	1.8621
1924–25	0.5415	1.5304
1925–26	0.5415	1.5304
1926–27	0.5790	1.4825
1927–28	0.6317	1.4302
1928–29	0.5973	1.2797
1929–30	0.3594	1.2497
1930–31	0.3906	1.2065
1931–32	0.5660	1.2506
1932–33	0.6468	1.1806
1933–34	3.2961	0.9521
1934–35	2.4005	1.2176
1935–36	1.1342	1.1684
1936–37	1.7530	0.9791
1937–38	1.6376	0.9866
1938–39	1.3164	1.0328

4. The following table gives the distribution of the households and area owned by size-level of household ownership holding. (N.S.S. July 1953—June 1954 major-crop season Central India).

Draw the Lorenz diagram and work out the concentration-ratio. (25)

Estimated number of households and area owned by size-level of household ownership-holding.

size of holding	number of households (000)	area owned (000 acres)	size of holding	number of households (000)	area owned (000 acres)
.00	2208	—	10.00–14.99	919	11142
.01– .04	1108	18	15–19.49	437	7607
.05– .09	154	10	20–24.99	306	6872
.10– .49	277	60	25–29.99	258	7030
.50– .99	224	147	30–39.99	226	7694
1.00–1.49	254	314	40–49.99	142	6210
1.50–2.49	530	1021	50–74.99	172	9999
2.50–4.99	1115	4011	75–99.99	14	1227
5.00–7.49	828	5068	100–249.99	34	4774
7.50–9.99	625	5435	250–	5	2015

5. The following table gives the wholesale price-indices of rice for India (base year 1952-53 = 100) for the four months March, June, September and December for the years 1957-58 to 1960-61. Remove the seasonal variation by the moving average method, fit a linear trend to the moving average series and work out the first order serial correlation for the residual series after the linear trend is removed.

Is the serial correlation significantly different from zero?

Calculate also the Von-Neumann ratio for testing the time-dependence of the residuals. Comment. (25)

Wholesale price-indices for rice

month \ year	1957–58	1958–59	1959–60	1960–61
March	97	100	92	108
June	017	111	102	114
September	108	118	111	114
December	102	98	101	102

(b) TECHNO-COMMERCIAL STATISTICS (PRACTICAL)

Use a separate answer book for each group

GROUP A: STATISTICAL QUALITY CONTROL

(Answer *any two* questions from this group)

1. *(a)* The observed mean and range, in a sample of 10 items, of a measurable characteristic are 21.5 and 3.8 respectively. Obtain the 95% confidence interval for the process mean.

(b) The average number of visual defects in a specified size of cloth has been found from past records to be 2.8. Determine the lower and upper control limits so that either of these could be reached or exceeded by a sample value with a probability of 0.05 or less. What is the chance of these limits being exceeded when the average number of defects increases to 3.6?

(c) An electrical assembly has been giving considerable rejections at the final inspection. The assembly consists of a number of oporations which two operations (A and B) are considered important. There are four assemblers (A_1, A_2, A_3, A_4) at A and three assemblers (B_1 B_2 B_3) at B. The following study was carried out at this assembly.

480 components were chosen at random from a common source and divided into 12 samples of 40 each. Each of the 12 samples were worked by one of the 12 combinations of assemblers. The other operations have been kept the same for all the 12 samples.

The following table records the rejections obtained for the 12 combinations at the final assembly.

	A_1	A_2	A_3	A_4
B_1	0	1	7	3
B_2	2	14	8	5
B_3	1	4	2	6

Make suitable analysis of the data and comment upon the results. (25)

2. (a) It is desired to control the mean weight of a cigarette at 1.20 gm. A 10% increase in the weight is to be detected with a probability of 0.05. On the other hand the probability of wrongly deciding that the mean weight has increased should not be more than 0.01. Determine the sample size and control limits given that standard deviation of the process is 0.04 gm.

(b) It is known that the final weight (y) of crockery items are dependent, among other factors, on their green weight (x). Paired observations of x and y were taken on 40 items and a summary of the results, in coded units, is given below.

$n = 40$ $\quad \Sigma x = 2537$ $\quad \Sigma y = 2468$

$\Sigma x^2 = 161589$ $\quad \Sigma xy = 157151$ $\quad \Sigma y^2 = 152888$

If x is controlled between 60 and 65 units, obtain the limits within which y may be expected to lie. (25)

3. The following data refers to an investigation carried out to determine the errors involved in the estimate of yield of an organic chemical.

Each of 8 batches were sampled twice at different intervals of time and on each sample the analysis was replicated giving in all $8 \times 2 \times 2 = 32$ readings.

Determine the errors due to analytical errors, sampling and batches respectively.

batch	sample A		sample A	
	analysis		analysis	
	1	2	1	2
1	60.5	60.0	60.3	59.7
2	56.0	55.6	54.6	54.9
3	51.2	51.6	52.9	51.8
4	58.9	60.2	58.2	60.6
5	59.1	59.6	58.6	57.0
6	57.2	56.3	57.6	57.4
7	56.5	55.2	54.8	54.2
8	55.1	52.5	55.3	53.6

GROUP B : OPERATIONS RESEARCH

(Answer question 4 or 5 and question 6 from this group)

4. The following table gives the costs of shipping a ton of steel from the steel-mills at Allentown, Birmingham and Chicago to four warehouses 1, 2, 3 and 4. The total requirements of steel at the warehouses are given in the last column and the total stocks at each mill are given in the last row. The total cost of transportation is to be minimized. Work out a basic feasible solution. See if it is optimum, if it is not, then obtain an improved basic solution. See if it is optimum, if it is not, then obtain a still better basic solution. Is the last solution optimum? For working out the improved solutions, use the implicit prices or any other method. Work out the cost of transportation for all the three solutions. (20)

mill / warehouse	cost in dollars			total requirements (000 tons)
	Allentown	Birmingham	Chicago	
1	73	95	71	12
2	90	89	82	21
3	97	93	100	32
4	76	94	91	11
total capacity (000 tons)	36	20	20	76

5. A manufacturer wants to know what is the optimum stock level of a certain part used in filling orders which come at a relatively constant rate but not in a constant size. Delivery of these parts to him is virtually immediate. He regularly places these orders at the begining of each month. Study of demand reveals that the probabilities shown in the table below are associated with various requirements per month. Finally C_1, the cost of holding a unit in stock for one month, is \$ 1.00 and C_2, the cost of a unit shortage per month, is \$ 20.00. Work out the optimum stock level and the total expected cost at that level and the two neighbouring levels.

number of units required per month (r)	probability of occurrence $P(r)$
0	0.1
1	0.2
2	0.2
3	0.3
4	0.1
>5	0.0

Note : $C(S) = C_1 \sum_{r=0}^{S} P(r)\,(S-\frac{r}{2}) + C_1 \sum_{r=S+1}^{\infty} P(r)\frac{S^2}{2r} + C_2 \sum_{r=S+1}^{\infty} P(r)\frac{(r-S)^2}{2r}$

and the minimum of this occurs when

$$[P(r < (S-1) + (S-\tfrac{1}{2}) \sum_{S}^{\infty} \frac{P(r)}{r}] < \frac{C_2}{C_1+C_2}$$

and $$\frac{C_2}{C_1+C_2} < [P(r < S) + (S+\tfrac{1}{2}) \sum_{S+1}^{\infty} \frac{P(r)}{r}].$$ (20)

6. The following tables give two patterns of the time of arrival of customers, the time at which services begin and the time at which services end. The cost of idle time at the counter is \$1 per minute and the cost of waiting time is also \$1 per minute (the customers are supposed to be paid so much for having had to wait). Work out the idle times, waiting times and queue-lengths for the two patterns and determine for which pattern the extra cost for idle and waiting times is greater. (10)

time of arrival		time at which service begins		time at which service ends	
hour	minute	hour	minute	hour	minute
		1st pattern			
8	07	8	07	8	13
8	14	8	14	8	20
8	25	8	25	8	31
8	39	8	39	8	45
8	43	8	45	8	51
8	56	8	56	9	02
		2nd pattern			
8	07	8	07	8	10
8	14	8	10	8	25
8	25	8	25	8	34
8	39	8	39	8	48
8	43	8	48	8	57
8	56	8	57	9	06

GROUP C: STATISTICAL METHODS IN BUSINESS

(Answer *any two* questions from this group)

7. Fit the log normal distribution to the following data on forecast errors. Comment. (10)

ratio of actual sales of cooking utensils to forecast sales (x)	frequency during first 24 months (f)
0—0.599	1
0.600—0.799	6
0.800—0.999	6
1.000—1.199	6
1.200—1.399	3
2.400—1.999	2
2.000—	0
total	24

[*Note*: If $\log x$ is distributed normally, x is said to be distributed log-normally]

8. The following table gives the preference percentages for a number of brands of toilet soap based on a survey of 534 housewives.

Test if the preference percentage for the brand A is significantly larger than that for each of the brands B, C, D and E.

brand	percentage of those who preferred	total number surveyed
A	11.2	534
B	10.6	534
C	7.5	534
D	6.4	534
E	4.5	534

9. The following table shows the data on sales available to the forecaster in May 1957. Produce a set of trend and seasonally adjusted moving average forecasts for the forthcoming 12 months starting with May. There is an average increase of 3.01% per year as estimated from annual data.

	monthly orders received		
month	1955	1956	1957
January	339	445	205
February	257	176	181
March	271	216	347
April	196	247	255
May	308	203	
June	441	439	
July	552	550	
August	190	186	
September	208	212	
October	107	152	
November	238	175	
December	326	420	

The following simple procedure can be followed.:

S_i = seasonal sales index for the i-th month

$$= \frac{\text{orders received for the } i\text{-th months in 1955 and 1956}}{\text{orders received in 1955 and 1956}}$$

A = previous twelve months orders received, in May 1957

A' = A, corrected for trend

$$= A\left(1 + \frac{8}{12} \times \frac{\text{annual percentage increase}}{100}\right)$$

Sales forecasts for i-th month in 1957 and 1958 $= A' \times S_i$ (10)

(*d*) DESIGN AND ANALYSIS OF EXPERIMENTS (PRACTICAL)

(Answer *any three* questions from this section)

1. The following table gives the yields (lbs./plot) of ear corn (Y) and number of plants (X) of 6 corn varieties, A, B, C, D, E, F in a 6×6 Latin Square.

Analyse the data and write a report comparing the yields of the different varieties, after correcting for the variation in the number of plants. (33½)

	A	B	C	D	E	F
X	18	16	18	14	15	17
Y	8.6	7.5	6.7	6.5	8.2	4.7
	B	C	D	E	F	A
X	16	15	16	19	21	14
Y	7.6	8.2	8.3	4.6	5.2	6.5
	C	E	A	F	B	D
X	15	16	16	19	15	17
Y	5.7	9.6	7.1	4.8	7.5	7.5
	D	F	B	A	C	E
X	18	18	20	18	22	17
Y	8.3	5.3	10.0	6.8	7.8	6.6
	E	D	F	B	A	C
X	15	18	19	16	16	15
Y	8.3	7.9	5.4	9.3	4.8	8.0
	F	A	E	C	D	B
X	18	20	19	17	17	15
Y	5.7	8.7	8.1	6.1	7.2	7.4

2. The following table gives the results of an experiment involving 4 treatments in 4 blocks, each of three plots, with a balanced layout. One of the 12 yields was lost in an accident and this is indicated by an asterisk(*).

Analyse the data fully working out the standard error of every pairwise comparison, and write a brief report on your findings. (33½)

yield in an incomplete block experiment

block	treatments			
	1	2	3	4
1	238	196	254	—
2	*	213	—	312
3	279	—	334	421
4	—	308	367	412

'—' not tried, '*' lost.

3. The following table gives the plan and the yields of an experiment on turmeric with three levels each of nitrogen (ammonium sulphate), phosphorus (super phosphate) and potash (sulphate of potash). The doses used were—

factors	levels		
	(0)	(1)	(2)
(*N*) Nitrogen	nil	60 lbs. *N*/acre	120 lbs. *N*/acre
(*P*) Phosphorus	nil	45 lbs. P_2O_5/acre	90 lbs. P_2O_5/acre
(*K*) Potash	nil	100 lbs. K_2O/acre	200 lbs. K_2O/acre

Plan and yield of 3^3 factorial experiment in 3 blocks of 9 plots each (Treatment combination in bracket, yield of raw turmeric in lbs./plot of 1/96 acre)

block	(treatment) and Yield								
1	(112) 114	(022) 90	(210) 76	(011) 68	(120) 105	(221) 95	(101) 91	(202) 81	(000) 92
2	(102) 79	(200) 76	(001) 57	(222) 97	(121) 88	(012) 75	(020) 71	(110) 76	(211) 76
3	(021) 73	(100) 74	(201) 66	(212) 100	(002) 66	(220) 107	(122) 88	(010) 65	(111) 84

Analyse the data. (33½)

4. (*a*) Give an optimum design for finding the weights of 8 objects by 8 weighings on a common balance with no zero-error.

(*b*) Consider the daul of the design in which every pair of treatments from amongst t treatments is separately tried just once on a block of two plots. Draw up detailed computational instructions for analysing fully the results of an experiment with this kind of design. Work out the efficiency factor of the design when $t=5$. (33½)

(*c*) SAMPLE SURVEYS (PRACTICAL)

(Answer *all* questions from this section)

1. You are required to conduct a survey on the pattern and extent of savings of households in which some members have purchased National Savings Certificates.

(*a*) Plan a survey describing clearly and in brief—

(*i*) the sampling units you will use.

(*ii*) the sampling frames required and how you will procure them.

(*iii*) an outline of the sampling procedure.

(*b*) Draw up a suitable household-schedule for the survey giving—

(*i*) the identification details of the household.

(*ii*) the relevant items to be collected.

(*c*) List down the different items in planning the survey, collecting the information and processing the data, which should be taken into account when drawing up the budget for the survey. (25)

2. The following table has been obtained from a complete census of all the 340 villages of a sub-division. The villages were stratified by size in 4 strata; N_i denotes the number of villages in a stratum $\bar{y}_{N_i}$ the stratum mean of area under wheat, S_{w_i} the standard deviation of area under wheat and S_{a_i} the standard deviation of agricultural area.

stratum number	size of village in bighas	N_i	$\bar{y}_{N_i}$	S_{w_i}	S_{a_i}
1	0 — 500	63	112.1	56.3	129.6
2	501 — 1500	199	276.7	116.4	267.0
3	1501 — 2500	53	558.1	186.0	276.1
4	>2500	25	960.1	361.3	982.2

Calculate the sampling variance of the estimated area under wheat obtained from a sample of 34 villages.—

(*a*) if the villages were selected by simple random sampling without stratification;

(*b*) if the villages were selected by simple random sampling within each stratum, and allocated in proportion to (*i*) N_i : (*ii*) $N_i S_{w_i}$: (*iii*) $N_i S a_i$

Tabulate these sampling variances and the relative efficiency of the different sampling methods with regard to—

(*i*) unstratified simple random sampling.
(*ii*) sampling with proportional allocation. (25)

3. A sample of 6 factories is selected from a population of 100 factories with probability proportional to the number of workers and with replacement, to estimate the total output. Using the data given below for the sampled factories, estimate the gain or loss in efficiency in using the value of fixed capital as the 'size' at the selection stage, instead of the number of workers. (25)

	factory number					
	1	2	3	4	5	6
number of workers	30	50	15	80	50	15
fixed capital (in 000 Rs.)	23	38	8	39	37	20
total output (in 000 Rs.)	50	90	20	100	100	50

Number of workers in all the 100 factories: 2×10^3
Total fixed capital of all the 100 factories: 2×10^7

4. The following table shows the results of a one-dimensional systematic sampling. The variate is area of green land in a given unit. Estimate roughly the standard error of the sample mean by the method of successive differences and also by assuming the sample to be a simple random one (ignoring the *fpc*), and comment on the results. (25)

serial number	1	2	3	4	5	6	7	8	9	10
area of green land	0.5	3.1	2.8	2.7	0.1	2.8	2.6	2.3	3.5	2.4
serial number	11	12	13	14	15	16	17	18	19	20
area of green land	3.8	4.1	4.9	6.0	5.4	2.3	2.9	2.1	6.3	8.2

(*g*) STATISTICAL INFERENCE (PRACTICAL)

(Answer *all* questions from this section)

1. In a genetical experiment on linseed, the observed frequencies for petal and stigma colours are given in the following table, along with the probabilities of the cells in terms of a parameter θ.

stigma colour	petal colour	
	lilac	deep lilac
white	357 $\left(\frac{2+\theta}{4}\right)$	33 $\left(\frac{1-\theta}{4}\right)$
purple	37 $\left(\frac{1-\theta}{4}\right)$	94 $\left(\frac{\theta}{4}\right)$

Estimate θ by the method of maximum likelihood. Also estimate the variance of this estimate of θ. (30)

2. A sample of size 25 is drawn from a normal population whose standard deviation is 3 and whose mean μ is unknown. It is desired to test the hypothesis $\mu=0$, at a 5% level of significance. Draw the power curves of—

(*i*) a left tail end test,

(*ii*) a right tail end test,

(*iii*) a two-sided equal tail ends test,

and comment on the relative merits of the above three tests with aspect to different classes of alternative hypotheses. (20)

3. The following data give heights in millimeters of two groups of persons.

group A				group B			
serial number	height in mms.	serial number	heigt in mms.	serial number	height in mms.	serial number	height in mms.
1	1658	9	1662	1	1676	11	1600
2	1617	10	1712	2	1585	12	1679
3	1628	11	1644	3	1584	13	1625
4	1657	12	1724	4	1674	14	1691
5	1666	13	1696	5	1597	15	1702
6	1630	14	1567	6	1588	16	1610
7	1650	15	1547	7	1689	17	1634
8	1677	16	1608	8	1607	18	1679
				9	1614	19	1675
				10	1690		

Use the follwing tests to examine whether the two groups can be considered to have come from the same population:

(*i*) Wald-Wolfowitz test.

(*ii*) Wilcoxon's test.

(*iii*) Dixon's test. (20)

4. Two groups of army recruits were given three tests A, B and C. The sample size, mean scores and the pooled dispersion matrix are given below:

group	sample size	mean scores		
		A	B	C
I	114	2.9298	1.1667	0.7281
II	33	3.0303	1.2424	0.5455

Dispersion matrix (within groups)

test	A	B	C
A	2.300851	0.251578	0.474169
B	—	0.607466	0.035774
C	—	—	0.595094

(a) Find the coefficients of the best linear function of scores which would discriminate between the two groups.

(b) How would you classify an individual into either of the two groups on the basis of his scores in the three tests?

(c) What would be the error of this method of classification? (30)