

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods
and Applications: 1976-77

PERIODICAL EXAMINATION

Paper IIIa: Basic Mathematics

Date: 8.9.76

Maximum Marks: 50

Time: $1\frac{1}{2}$ hours

Note: Answer all the questions. Marks allotted for each question are given in brackets [].

- 1.a) How many numbers greater than 4,000,000 can be formed from the digits 2, 3, 4, 4, 5, 5, 5 ?
- b) A club of 60 members has four kinds of membership, say I, II, III, IV. There are 20 members of the I type, 16 of II, 14 of III and 10 of IV. In how many ways can you choose a committee of 8 members so that it has two people from each type of membership ?
- a) What is the coefficient of x^n in the expansion of $(x^2 + \frac{1}{x})^{n+2}$ [5+5]=[15]
- 2.a) Define the concept of the limit of a sequence $\{a_n\}$ of real numbers.
Show that $\lim_{n \rightarrow \infty} a_n = 1$ if $a_n = n^{\frac{1}{n}}$, $n = 1, 2, 3, \dots$
- b) Let $f(x)$ be the function defined by:
- $$f(x) = x^2 \sin \frac{1}{x} \quad \text{if } x \neq 0$$
- $$= 0 \quad \text{if } x = 0.$$
- What is the derivative $f'(x)$ of f at x where $x \neq 0$?
(You may assume that $\frac{d(\sin x)}{dx} = \cos x$)
What is $f'(0)$? [10+10]=[20]
- 3.a) Define the rank of an $m \times n$ matrix A ,
- b) Solve the following system of linear equations:
- $$2x_1 + 3x_2 + 4x_3 = 2$$
- $$x_1 + x_2 + 5x_3 = 0$$
- $$2x_1 + 3x_2 + x_3 = 3$$
- [5+10]=[15]

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods
and Applications: 1976-77

Part I: PERIODICAL EXAMINATION

Paper Ia: Compilation and Presentation of Data

Date: 22.9.76

Maximum Marks: 50

Time: 2 hours

Group A: Answer any two questions.

1. Define the following terms, giving suitable illustrations: Individual, sample, population, attribute, discrete variate, continuous variate. [10]
2. Why is it that results based on samples are often more accurate than those based on complete enumeration? Discuss, explaining the sources of non-sampling errors in statistical data. [10]
3. Compare the interview method and the mail questionnaire method of collecting statistical data from a sample of respondents. [10]
4. Mention the main principles to be followed in representing data through bar diagrams. (You need not consider the variants of the bar diagram.) [10]

Group B: Answer any two questions.

5. Present the following information in a suitable tabular form with appropriate headings and footnotes:

Estimates of savings of the household sector of the Indian economy during 1960-61 and 1961-62 were prepared by three agencies. The NCAER, New Delhi, published their estimates in 1966 in their Occasional Paper 16, 'Saving in India during the Plan Period', the CSO (Govt. of India) published their estimates in 1969 in a publication entitled 'Estimates of Saving in India 1960-61 to 1965-66', and the RBI published theirs in the RBI Bulletin of March 1965. All the estimates were at current prices and expressed in Rs. lakhs. The total savings in 1960-61 were estimated as 10825 by the NCAER, as 75100 by the CSO and as 102580 by the RBI, and the corresponding figures for 1961-62 were 97160, 83800 and 92220. Savings was split into two broad components: financial assets and (net) physical assets. The estimates of total financial assets for 1960-61 were 50771 for the NCAER, 45700 for the CSO and 43600 for the RBI, the corresponding figures for 1961-62 being 40806, 49200 and 47600. The estimates of total (net) physical assets were 29381 for 1960-61 and 34573 for 1961-62 according to the CSO, 55054 for 1960-61 and 56354 for 1961-62 according to the NCAER and 58980 for 1960-61 and 44560 for 1961-62 according to the RBI. Financial assets had various forms, one of which was 'currency'. Estimates for 'currency' were 13523 (NCAER), 14505 (CSO) and 16030 (RBI) for the year 1960-61 and 9630 (NCAER), 9416 (CSO) and 9520 (RBI) for the year 1961-62. [15]

6. Give a suitable graphical representation to the following data:

Table: Official estimates of area under jowar and of yield rate of jowar in India during 1957-58 to 1964-65.

Year	area (thousand hectares)	yield rate (kg./hectares)
(1)	(2)	(3)
1957 - 1958	17311	499
1958 - 1959	17960	503
1959 - 1960	17707	484
1960 - 1961	18412	533
1961-- 1962	18249	440
1962 - 1963	18414	529
1963 - 1964	18376	501
1964 - 1965	18056	536

[15]

7. Draw an appropriate diagram to represent the following data:

head of expenditure	average monthly expenditure per household (in Rs.)	
	1970-71	1975-76
(1)	(2)	(3)
food	192	322
fuel and light	28	41
housing	35	49
clothing	34	60
miscellaneous	61	120
total	350	592

[15]

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods
and Applications: 1976-77

Part I: PERIODICAL EXAMINATION

Paper II: Descriptive Statistics - Theory

Date: 13.10.76

Maximum Marks: 75

Time: 2 hours

Note: Answer any four questions. All questions carry equal marks.

- 1.a) Explain with examples the distinction
- i) between an attribute and a variable and
 - ii) between a discrete variable and a continuous variable.
- b) Discuss the different considerations to be kept in view in drawing up a frequency distribution for data on a continuous variable.
- 2.a) What are the different measures of central tendency? Discuss the relative merits and demerits of using these measures.
- b) The mean weight per student in a group of 6 students is 119 lb. The individual weights of 5 of them are 115 lb, 109 lb, 129 lb, 117 lb and 114 lb. What is the weight of the other student of the group?
3. Suppose you have computed the measures a.m., median, s.d. and coefficient of variation from a given data.
- What effect would adding a constant 5.2 to all observations have upon the numerical values of these measures? What would be the effect of adding 5.2 and then multiplying the sums by 8.0? Would it make any difference in the above quantities if you multiply the observations by 8.0 first and then add 5.2?
4. What are skewness and kurtosis? Give some suitable measures for skewness and kurtosis. Given first four raw moments, write down the formulae to find the first four central moments in terms of these raw moments.
5. Define product moment correlation coefficient and state its important properties. Prove that the correlation coefficient lies between -1 and +1. Interpret the cases when it is equal to -1 and +1.
- 6.a) Define conditional and marginal frequency distributions from a bivariate frequency table. What do you mean by the regression of one variable on another?
- b) For 20 Army personnel, the regression of weight of kidneys (y) on weight of heart (x), both measured in oz, is
- $$Y = .399x + 6.93$$
- and the regression of weight of heart on weight of kidneys is
- $$X = 1.212y - 2.461.$$
- Find the correlation coefficient between x and y and their means. Find their s.d.s as well.

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods
and Applications. 1976-77
PERIODICAL EXAMINATION

Paper II. Descriptive Statistics (Practical)

Date: 27.10.76

Maximum Marks.75

Time: 2 hours

Note. Answer all the questions. Marks allotted for each question are given in brackets []. In each question, write down the procedure briefly. Don't use Sheppard's correction.

1. The following table gives the frequency distribution of scores in English obtained by students who passed in the subject (i.e., scored 36 or more) in a School Leaving Certificate examination.

Score	Frequency
36 - 40	441
41 - 45	397
46 - 50	333
51 - 55	186
56 - 60	107
61 - 65	38
66 - 70	14
71 - 75	4
76 - 80	2

- a) Draw the two ogives and hence find out approximate values of the three quartiles. [3+6]=[9]
- b) Calculate approximate values of the 5th and 95th percentiles. [3]
- c) Calculate a measure of dispersion as well as a measure of skewness. [6]
2. The following table gives the frequency distribution of height for 177 Indian adult males. Compute the kurtosis of this distribution.

Height	Frequency
144.6 - 149.5	1
149.6 - 154.5	3
154.6 - 159.5	24
159.6 - 164.5	58
164.6 - 169.5	60
169.6 - 174.5	27
174.6 - 179.5	2
179.6 - 184.5	2

- 3.a) In a batch of 10 children, the I.Q. of a dull boy is 36 below the average I.Q. of the other children. Show that the s.d. of the I.Q. for all the children cannot be less than 10.8. If this s.d. is actually 11.4, determine what the s.d. will be when the dull boy is left out. [10]
- b) Examine the following results of a piece of computation for internal consistency. [5]
- $$n = 100, \quad \sum x = -114, \quad \sum x^2 = 103.$$
4. Neatness [2]

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods
and Applications, 1976-77

PERIODICAL EXAMINATION

Paper Ib.: Index Numbers and Time Series

Date: 10.11.76 Maximum Marks: 100 Time: 2 $\frac{1}{2}$ hours

Note: Answer Q.5 and any three questions out of the rest. Marks allotted for each question are given in brackets [].

- 1.a) What are time reversal and circular tests of consistency of index number formula? Mention one formula which satisfies the former test but not the latter. [4+4+4]=[12]
- b) Average monthly expenditure per household in 1951 on food articles in the rural areas of North India and the prices of these articles in 1951 and in 1954 are given below:

commodity	expenditure per household in 1951 (Rs.)	average price per seer (Rs.)	
		1951	1954
food-grains	53.27	0.40	0.45
pulses	7.11	0.43	0.38
edible oils	4.07	2.12	1.50
vegetables	2.19	0.44	0.42
milk	4.80	0.75	0.62
meat and fish	1.41	1.50	1.62
fruits	0.34	1.60	1.50
salt	0.56	0.07	0.07
spices	2.23	2.40	2.40
sugar	4.46	0.70	0.75

Calculate a suitable index showing the change in food prices in this area in 1954 compared to 1951. [13]

- 2.a) Suppose, you are given national product at current prices in each of two years - 1960 and 1975 - and a general price index for 1975 with 1960 as base. Can you obtain an index of real national product in 1975 with 1960 as base? [10]
- b) The general consumer price index for the middle class people of Calcutta increased by 266 per cent in 1957 over 1937. The table below gives the consumer price indices (1939=100) of all groups of items except clothing along with their weight.

Group	Weight	Consumer price index in 1957 (base 1939=100)
Food	61	411.8
Fuel and lighting	7	388.0
Housing	9	116.9
Miscellaneous	19	284.5
Clothing	4	?
	<u>100</u>	

Calculate the consumer price index for clothing group in 1957 (with 1939 = 100) [15]

3. Is the following assertion correct: 'the Laspeyres' index never exceeds the Paasche's index'? Give reasons for your answer. [25]

- 4.a) Are there any considerations involved in the choice of a base period for comparing prices over time? [8]
- b) What is a 'quantity index'? Write Laspeyres' quantity index number formula. [5+2]=[7]
- c) Do you agree with the view that Fisher's ideal index number formula is superior to Laspeyres' formula? Give reasons for your answer. [10]
5. What do you mean by secular trend of a time series? Describe when and how you would fit a second degree trend equation to a time series data. [5+20]=[25]

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods
and Applications: 1976-77

PART I PERIODICAL EXAMINATION

Paper III(b): Probability

Date: 24.11.76

Maximum Marks: 100

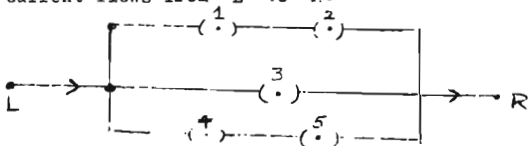
Time: $2\frac{1}{2}$ hours

Note: This paper carries 135 marks. Answer as many questions as you can. Maximum you can score is 100. Marks allotted for each question are given in brackets [].

- 1.a) Give the axiomatic definition of probability in finite sample spaces.
- b) Prove that for any two events A and B, we have that

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$
 [8+7]=[15]
2. Two six-sided dice, one red and the other green, are thrown and the numbers on their upper faces are observed. Write down the sample space. Find the probability, assuming that all out comes are equally likely, of:
 - a) throwing a double,
 - b) $r + g = 10$,
 - c) $g \geq r + 3$.

[5+5+6+4]=[20]
3. Define the term 'complete independence of three events'. Give an example of three events which are pairwise independent but not complete independent. Justify your answer. [5+10]=[15]
4. In the following Figure assume that the probability of each switch being closed is p and that each is open or closed independently of any other. Find the probability that current flows from L to R.



- 5.a) Define the terms 'a priori probability, a posteriori probability'. State Bayes' theorem.
- b) In a factory, machine A produces 30% of the output, machine B produces 25%, and machine C produces the remaining 45%. One percent of the output of machine A is defective, as is 1.2% of B's output, and 2% of C's. An item drawn at random from a day's output is defective. What is the probability that it was produced by A? by B? by C? Clearly indicate the a priori probabilities, a posteriori probabilities in a separate table. [25]
6. Let the probability function of a random variable X be as follows:

Values, X	2025	2050	2075
Probability, f(x)	0.3	0.2	0.5

Find the variance of X.

[10]

7. Find the mean and the variance of the binomial distribution $B(n, p)$. [5+8]=[13]
8. Give an example of two random variables X, Y for which $E(XY) \neq E(X) \cdot E(Y)$. Justify your answer. [4]
9. The joint probability distribution of X, Y is given in the following table. Prove that X and Y are independent. Find all the conditional probability distributions. [10 +5]=[15]

X \ Y	1	2	3	4
0	$\frac{1}{24}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{24}$
1	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{12}$
2	$\frac{1}{24}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{24}$

[10 +5]=[15]

INDIAN STATISTICAL INSTITUTE
RESEARCH AND TRAINING SCHOOL

One-year Evening Course in Statistical Methods and Applications: 1976-77

Part I Final Examination
Paper Ia: Collection and Presentation of Data
(Theory and Practical)

Time: 1.30 hrs.

Date: 3.1.1977

Max. Marks: 50

N.B.: Answer question 1 (which is compulsory) and any of the remaining questions. The maximum you can score is 50 marks.

- 1 (i) The net domestic product at factor cost for India amounted to 13335 crores of rupees in 1960-61: of which the contribution of the primary sector (i.e. agriculture, forestry, etc) was 6965 crores of rupees, the contribution of the secondary sector (i.e. manufacturing construction etc.) was 2549 crores of rupees, and the contribution of the tertiary sector was 3821 crores of rupees.

The corresponding figures for 1973-74 were 20179, 8589, 4671 and 6921 crores of rupees.

Source: National Accounts Statistics, 1960-61 to 1973-74, C.S.O., 1976.

Prepare a blank tabular layout (with a suitable title) to present the above information.

- 1(ii) The following table gives the age distribution for the male and female population in India during 1971.

Represent the data in a suitable graphical form and comment.

Age groups (years)	Population (in millions)	
	Male	Female
0 - 14	119	111
15 - 29	67	64
30 - 49	64	58
above 50	35	31

Source: Census of India, 1971.

[8+8]

2 Write brief notes on any three of the following:

1. Population
2. Sample
3. Inductive Inference
4. Characteristic
5. Non-sampling errors

[9]

3 Give a brief account of the NSS highlighting its salient features.

OR

Describe briefly the current position of the statistical system in India.

[10]

4 State only (1) the name of the publication and (2) the authority which publishes it, where you can find data on three of the following: (Do not mention the same source more than once).

1. National income of India;
2. National income of different countries of the world;
3. agricultural situation in India;
4. statistics relating to planning in India;
5. statistics relating to Indian manufacturing industries.

[6]

5 Distinguish between primary and secondary data.

What are the different methods of collecting primary data?

[10]

6 Write a short note explaining the usefulness of statistics in the current economic situation in India.

[10]

Neatness

[5]

INDIAN STATISTICAL INSTITUTE

One-year Evening Course in Statistical Methods
and Applications. 1976-77.

Part V: FINAL EXAMINATIONS
DELHI AND CALCUTTA

Paper Ib: Topics in Applied Statistics: Index Numbers
and Time Series - Theory and Practical.

Date: 5.1.77

Maximum Marks: 100

Time: $2\frac{1}{2}$ hours

Note. Answer Groups A and B in separate answer-
scripts.

Group A. Index Numbers

Note: Answer any two questions. Marks allotted for
each question are given in brackets [].

- 1.a) Distinguish between an 'aggregative' and an 'average type'
index number formula. Can you interpret Laspeyres' formula
in terms of either of the above two approaches? [5+5+5]=[15]
- b) The following table gives index numbers of wholesale prices
of groups/sub-groups of items in India for the last week of
January, 1974 (base 1961-62 = 100) as well as their respec-
tive weights.

Group/sub-group	Weight in the group	Weight in the general index	Index for last week of January 1974 (1961-62 = 100)
1. Food articles	...	413	...
(1.1) Food grains	358	...	314.4
(1.2) Fruits and vegetables	56	...	300.5
(1.3) Milk and milk products	148	...	307.6
(1.4) Edible oils	131	...	376.5
(1.5) Fish, eggs and meat	48	...	464.1
(1.6) Others	259	...	252.3
2. Industrial raw materials	...	121	319.5
3. Liquor and tobacco	...	25	252.7
4. Fuel, power, light and lubricants	...	61	237.1
5. Machinery and transport equipments	...	79	195.8
6. Other manufactures (including chemicals)	...	301	228.0

Calculate a general index of wholesale prices for the last
week of January, 1974 (1961-62 = 100). [10]

- 2.a) What is meant by the term 'consumer price index'? If you
are asked to construct consumer price index numbers for
industrial workers in Calcutta, how would you select vari-
ous commodities including their varieties and obtain their
weights? [5+10]=[15]
- b) The table below presents values (in million pounds) of
exports of manufactures by U.K. in two years 1938 and
1949.

2.(contd.)

Items	1938	1949	
	value at current prices	value at current prices	value at 1938 average prices
Pottery, glass etc.	9.6	48.8	20.4
Metal manufactures	54.0	190.2	88.2
Electrical apparatus	13.6	79.1	31.4
Machinery	57.9	278.7	119.9
Textile manufactures	101.2	359.5	100.2
Chemicals	22.3	86.1	35.4
Vehicles	44.5	313.5	144.8
Other manufactures	62.2	202.7	85.7

What type of quantity and price index numbers of U.K. exports can be constructed from such data? Compute these for 1949 with 1938 as the base. [5+5]=[10]

- 3.a) Define a chain index. Suppose a series of fixed base indices $I_{01}, I_{02}, \dots, I_{0k}$ is available. Under what condition this series would be independent of the choice of the base period? [3+7]=[10]
- b) The index of Business Activity is constructed by taking the weighted average of activity relatives in the different business sectors of the economy. The sectors, their activity levels in 1948 and May 1949 as also the weights for the different sectors are shown in the table below. The index of industrial production in May 1949 with 1948 as base is 117.6. Calculate the index of Business Activity in India in May 1949 (1948 = 100).

Sector	Business Activity in India		Weight
	Monthly average in 1948	May 1949	
I Industrial production	-	-	46
II Foreign trade, exports and imports (Rs. lakhs)	32	41	6
III Financial Activity, Cheque clearance and note circulation (Rs. crores)	479.3	442.5	23
IV Transport: Cargo carried by ships and railways (thousand tons).	32	73	25

Group B. Time Series

Note: Answer Q.6 and any one of the rest. Marks allotted for each question are given in brackets [].

4. State the characteristics of the modified exponential curve and describe, in detail, how you will fit the curve to some empirical data. [2]

5. Write short notes on any two of the following

- i) Changing seasonal pattern.
- ii) Logistic curve.
- iii) Cyclical fluctuations and methods of measuring it.

[20]

6. The following table gives the import of raw jute in Calcutta for a number of years. Assuming the seasonal pattern to be constant, compute monthly seasonal indices by the ratio-to-moving average method. Explain clearly what the seasonal indices really indicate ?

Import of raw jute (000 tons) into Calcutta.

Year \ Month	1955	1956	1957	1958
January	103.4	130.2	131.8	132.2
February	105.6	121.8	101.0	96.5
March	89.5	115.1	85.8	86.0
April	69.2	85.3	67.3	74.4
May	55.6	65.2	75.4	77.8
June	48.7	53.6	63.7	54.3
July	42.4	55.7	59.9	57.0
August	47.8	74.5	53.1	63.6
September	87.3	79.4	75.5	73.0
October	105.9	110.0	109.9	122.0
November	143.9	138.6	141.4	152.7
December	138.9	135.7	138.0	158.0

[30]

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods
and Applications. 1976-77

Part I FINAL EXAMINATIONS
DELHI AND CALCUTTA

Paper II(i): Descriptive Statistics - Theory

Date: 7.1.77

Maximum Marks: 75

Time: 2 hours

Note: Attempt any three questions.
All questions carry equal marks.

-
- 1.a) Let \bar{x}_1, \bar{x}_2 and s_1^2, s_2^2 be the means and variances respectively based on two independent sets of observations of sizes n_1 and n_2 . Find the expressions for the combined mean and combined variance.
- b) Find the mean deviation about mean and standard deviation of the series $a, a+d, a+2d, \dots, a+nd$ and prove that the latter is greater than the former.
- c) What are moments? Show that some measures of central tendency, dispersion, skewness and kurtosis can be described in terms of moments.
-
- 2.a) Define probability mass function (p.m.f.) and probability density function (p.d.f.) of a random variable. State under what conditions a function can be treated as p.m.f. or p.d.f.
- b) Starting with a suitable probability model how will you develop Binomial probability distribution? Show that under certain limiting conditions it reduces to Poisson probability distribution.
- c) Define uniform probability distribution and obtain the mean, s.d., skewness and kurtosis of this distribution.
-
3. Define the normal probability distribution and show that one of the parameters represents the mean, median and mode of the distribution.
Explain how to show that Binomial and Poisson distributions tend to normal distribution.
-
- 4.a) Carefully examine each of the following statements and say whether the statement is true or false adding a few lines to substantiate your conclusion.
If the correlation coefficient between random variables X and Y be negative, then it means
- correlation coefficient between $-X$ and $-Y$ is positive,
 - $E(XY) < E(X)E(Y)$,
 - regression coefficient of X on Y and that of Y on X are both negative,
 - the regression of Y on X is linear and the line is downward sloping.
- b) Find the most likely price in Bombay corresponding to the price of Rs.70 in Calcutta from the following data:

Average price in Calcutta = 65

Average price in Bombay = 67

s.d. at Calcutta = 2.5

s.d. at Bombay = 3.5

Correlation coefficient is +.8 between the two prices of the commodity in the two towns.

- 5.a) Define multiple and partial correlation coefficients in case of three variables x_1, x_2, x_3 .

Show that multiple correlation coefficient is always non-negative.

- b) If the relation $ax_1 + bx_2 + cx_3 = 0$ holds for all sets of values of x_1, x_2 and x_3 , what must be the partial correlation coefficient?
-

INDIAN STATISTICAL INSTITUTE
RESEARCH AND TRAINING SCHOOL

One-year Evening Course in Statistical Methods and Applications: 1976-77

Part I Annual Examination

PAPER II(ii) DESCRIPTIVE STATISTICS PRACTICAL

Time: 2 hrs.

Date: 10.1.1977

Max. Marks: 75

N.B.: Attempt as many questions as you like.

1 (a) The following Table gives the I.O. for 300 six-year old children.

I.O.	Number of children
130-149	25
110-129	112
90-109	135
70- 89	22
50- 69	6

Compute the third central moment and hence a measure of skewness for this distribution.

[12]

(b) In the recently concluded Ranji Trophy Matches, two batsmen A and B from a certain competing team were at top of their form. Both of them batted in 8 innings each and the number of runs scored by each has been as shown below.

Innings	1	2	3	4	5	6	7	8	All innings
Runs scorel	53	40	86	62	71	58	49	80	508
by	113	27	39	32	199	0	150	46	506

Which of the two batsmen is more consistent?

[2]

- 2 When the first proof of 200 pages of an encyclopedia of 5000 pages was read, the distribution of printing mistakes was found to be as shown below:

Frequency Distribution of number of misprints in a page

No. of misprints	Frequency
0	112
1	63
2	20
3	3
4	2

- (a) Fit a Poisson Law to the frequency distribution of printing mistakes.
- (b) Estimate the total cost of correcting the first proof of the whole encyclopedia if the cost of detection and correction is 20 paise per page for pages containing less than 3 mistakes and 30 paise per page for pages containing more than 2 mistakes.

[10+5]

- 3 Under suitable conditions a gunner from a bomber tries to hit a square target of side 250 ft. with the centre of the target just below the plane. The errors of the gunner along axes parallel to the sides of the target are supposed to be independent and normally distributed with the centre of the target as mean and 400 ft. standard deviation. Find the probability (a) that the target will be hit at the first attempt and (b) that the target will be hit at least once in ten attempts.

[15]

- 4 During an investigation in an agricultural farm, the length - X (in cm) - of green jute plant and the weight - Y (in gm.) of dry jute fibre were observed for 100 plants. With these data the following bivariate frequency table was obtained.

		length of green jute plants (cm)				
		101-120	121-140	141-160	161-180	181-200
Weight	2 cm and less	0	5	6	2	0
of	- 4 cm	2	10	4	1	1
dry	- 6 gm	5	9	11	3	2
jute	- 8 gm	3	2	4	11	5
fibre	-10 gm	0	1	3	2	8
(gm)						

From the above table obtain:

- (a) the regression line of Y on X .
- (b) the correlation coefficient between X and Y .

[10+10]

The following means, standard deviations and correlations are found for:

X_1 = seed hay crop in cwt. per acre

X_2 = spring rainfall

X_3 = accumulated temperature above 42°F in spring
in a certain district during 20 years

$\bar{X}_1 = 28.02$ $s_1 = 4.42$ $r_{12} = 0.80$

$\bar{X}_2 = 4.91$ $s_2 = 1.10$ $r_{13} = -0.40$

$\bar{X}_3 = 594.00$ $s_3 = 85.00$ $r_{23} = -0.56$

Obtain (i) the regression equation for hay crop on spring rainfall and accumulated temperature; and (ii) the partial correlation coefficient between spring rainfall and accumulated temperature.

[20]

INDIAN STATISTICAL INSTITUTE
RESEARCH AND TRAINING SCHOOL

One-year Evening Course in Statistical Methods and Applications: 1976-77

Part I Annual Examination

BASIC MATHEMATICS

Time: 1.30 hrs.

Date: 12.1.1977

Max. Marks: 50

N.B.: Attempt any number of questions

1 (a) Derive an expression for the number of permutations of n things taken r at a time.

(b) Find the term independent of x in the expansion of

$$\left(3x^3 - \frac{7}{x^2} \right)^{10}$$

(c) How many distinguishable permutations of the letters of the word ALLAHABAD are there?

[6+6+4]

2 (a) What do you understand by convergence and divergence of series? Under what condition is a geometric series convergent?

(b) Test for the convergence of the series

$$\frac{1}{2} + \frac{3}{2^2} + \frac{5}{2^3} + \dots$$

[3+5]

3 (a) Multiply the following matrices:

$$\begin{pmatrix} 2 & 3 \\ 4 & 7 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 5 & 0 & 7 \\ -2 & 3 & 11 \end{pmatrix}$$

(b) Find the inverse of the following matrix:

$$\begin{pmatrix} 1 & 5 & 6 \\ 2 & 3 & 1 \\ 0 & 1 & 2 \end{pmatrix}$$

[4+10]

4 (a) Define continuity and differentiability of a function at a point.

(b) Find the derivatives of (i) $(\cos x)^{\sin x}$ (ii) $\frac{ax + b}{cx + d}$

[4+8]

5 Evaluate the following Integrals:

i) $\int \frac{\sin x}{1 + \cos x} dx$

ii) $\int e^x \sin x dx$

iii) $\int x (5x^2 - 3)^7 dx$

[4+4+4]

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods
and Applications: 1976-77

Part I. FINAL EXAMINATIONS

DELHI AND CALCUTTA

Paper IIIb. Probability

Date: 14.1.1977

Maximum Marks: 100

Time: 2½ hrs.

Note: The paper carries 140 marks. Answer as many questions as you can. Maximum you can score is 100. Marks allotted for each question are given in brackets [].

- 1.a) State the axiomatic definition of probability, in finite sample spaces. [5]
- b) If A and B are any two events, then prove that

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$
 [8]
- c) Two six-sided dice, one red and the other green, are thrown and the numbers which appear on their upper faces are noted. Assuming that all the outcomes are equally likely, find the probability of the following events, where r, g are numbers on the red and green die respectively.
- i) $r \geq g + 3,$
 - ii) throwing a double,
 - iii) r and g have the same parity. [4+3+5]=[12]

- 2.a) Define the terms 'a priori, a posteriori probabilities'. State Bayes' theorem. [10]
- b) In a factory, machine A produces 30% of the output, machine B produces 25% and machine C produces the remaining 45%. One percent of the output of machine A is defective, as is 1.2% of B's output, and 2% of C's. An item drawn at random from a day's output is defective. What is the probability that it was produced by A? by B? by C? Summarize, in a separate table, the a priori and a posteriori probabilities. [15]

- 3.a) Define the terms marginal, conditional probability distributions associated with a two-dimensional random variable (X, Y). [7]
- b) Let (X, Y) be a 2-dimensional random variable with the joint probability distribution given in Table 1. Find the marginal, conditional probability distributions of (X, Y).

Table 1

	X	0	1	2
Y	0	0.1	0.2	0.2
1	0.24	0.08	0.08	
2	0.06	0.12	0.12	

[5+8]=[13]

- 4.a) Stating all the conditions necessary, prove that a binomial distribution can be approximated by means of a poisson distribution. [15]
- b) In a lot of 50 items, 5 are defective. If a random sample of 5 items is selected without replacement, what is the probability that the sample is composed of
- i) all nondefective,
 - ii) 3 nondefective and 2 defective items. [6 + 4] = 10.
-
- 5.a) Define the terms independence of two random variables'. [4]
- b) If X, Y are independent random variables, then prove that $E(X \cdot Y) = E(X) \cdot E(Y)$. [12]
- c) Give an example of two random variables, X, Y for which $E(X \cdot Y) \neq E(X) \cdot E(Y)$. Justify your answer. [9]
- 6.a) State the Chebyshev's inequality. [5]
- b) Suppose that $\text{Var}(X) = 0$. Then prove that $P(X = \mu) = 1$, where $\mu = E(X)$. [5]
- c) Deduce the weak law of large numbers from Chebyshev's inequality. [10]

INDIAN STATISTICAL INSTITUTE

One-year Evening Course in Statistical Methods and Applications:
1976-77 -

Part I: Supplementary Examination

Paper II (ii) Descriptive Statistics - Practical

Time: 2 hrs. Date of Examination: 16.2.1977 Max. Marks: 75

1. Construct a histogram for the following distribution of scores and on the same figure draw the frequency polygon.

Scores	Frequency	Scores	Frequency
20 - 29	3	60-69	79
30 - 39	4	70-79	118
40 - 49	13	80-89	91
50 - 59	39	90-99	22

OR

- Compute (i) Median (ii) 30th centile and
(iii) standard deviation for the above
distribution of scores.

/20/

2. If the number of marks scored in a certain examination is normal random variable with mean 55 and standard deviation 21; what is the likely number of candidates obtaining (i) less than 34 marks, (ii) more than 40 marks but less than 60 marks and (iii) more than 80 marks; out of a total number of 575 candidates appearing in the examination?

/15/

3. Obtain the equation of the regression lines of (i) Y on X and (ii) X on Y for the following data. Draw the two lines on a graph and read the values of the means of Y and X from the graph.

X	2	2	4	4	4	6	6	6
Y	3	6	2	4	8	5	7	10
X	8	8	8	10	19	19	12	12
Y	5	8	10	8	12	5	9	11

/15 + 15/

4	1
5	1
Total	200

4. Presentation and Neatness

 / 19 /

- a) Fit a Poisson law to the frequency distribution of printing mistakes.
- b) Estimate the total cost of correcting the first proof of the encyclopedia by using the information given in the following table.

Table 3

Cost of Detection and Correction of Misprints.

No. of misprints on a page	Cost of detection and correction per page (dollars)
0	.10
1	.16
2	.23
3	.29
4	.34
5	.36

$$[15+10]=25]$$

4. The following table gives the means and corrected sum of products matrix for three variates. x_1 = crude oil vapour pressure, x_2 = gasoline end point, x_0 = gasoline yield percentage based on 32 samples of crude oil.

- a) Build up a linear formula for estimating the percentage of gasoline yield from crude oil, x_0 on the basis of other characteristics x_1 and x_2 ,
- b) Compute the multiple correlation coefficient $R_{0.12}$ of x_0 or x_1 and x_2 .
- c) Find the partial correlation coefficient $r_{01.2}$.

$$[12\frac{1}{2} + 12\frac{1}{2}] = 25]$$

Table 4

Corrected sum of Products and Measure.

	x_1	x_2	x_0	Means
x_1	212.769	-1,688.140	331.456	4.18125
x_2		33,466.500	-3,931.050	241.500
x_0			3,564.070	19.6594

INDIAN STATISTICAL INSTITUTE
One year Evening Course in Statistical Methods
and Applications, 1976-77
PERIODICAL EXAMINATION

Part II

Paper IV: Methods of Statistical Inference:
Theory and Practical

Date: 2 March 1977

Maximum Marks. 75

Time: 2 hours

Note: Answer any three questions.

1. Define clearly and precisely the following concepts.
 - a) Simple hypothesis and composite hypothesis,
 - b) Critical region,
 - c) Probabilities of Type I and Type II errors,
 - d) Level of significance of a test,
 - e) Power of a test.

 - 2.a) In order to test the hypothesis H_0 that the proportion p of defectives produced by a machine is a given value p_0 against the alternative H_1 that it is greater than p_0 a sample of n observations is taken. Derive the uniformly most powerful test for H_0 against H_1 at level of significance α .

State clearly the theorem that gives you the uniformly most powerful test above.
 - b) Obtain the test for $H_0: p = p_0$ against $H_1: p \neq p_0$ based on a sample of size n and of level of significance α .

 - 3.a) A fertilizer mixing machine is set to give 10 pounds of nitrate for every 100 pound bag with variance $\sigma^2 = 6.25$. 9 bags are examined and the nitrate content is found out to be 9, 12, 11, 10, 11, 9, 13, 6, 9 pounds. Test the hypothesis H_0 that the mean nitrate content $\mu = 10$ pounds against the alternative $H_1: \mu \neq 10$. (Take $\alpha = .01$)
 - b) Calculate the power of the test for the values, $\mu = 9, 9.5, 10.5, 11$.

 4. The mean systolic blood pressure of adults is known to be 120. In order to test the variance in the systolic blood pressure a sample of 10 individuals is taken and their systolic blood pressure readings were
116, 124, 110, 114, 100, 130, 125, 130, 90, 110.
Use the data to test whether the variance can be regarded as 25 against the alternative (a) that it is greater than 25, (b) that it is not equal to 25.
(Take $\alpha = .05$ in both cases.)
-

INDIAN STATISTICAL INSTITUTE
One year Evening Course in Statistical Methods
and Applications: 1976-77

PERIODICAL EXAMINATION
Part II

Paper V: Sample Surveys. (Theory and
Practical) ✓

Date: 23.3.77

Maximum Marks: 75

Time: 2 hours

Group A

Attempt any two questions from this Group.
The questions carry equal marks.

- 1.a) Describe briefly the advantages of carrying out a sample survey in preference to a complete enumeration survey. In which situations is complete enumeration survey preferred?
- b) Explain briefly the following terms with examples, whenever possible:
- i) probability sampling; (ii) Sample design;
 - iii) sampling frame; (iv) sampling and non-sampling errors. [20]
- 2.a) Let y_i be the i th sample observation ($i = 1, 2, \dots, n$) in a srsWOR from a population of N units.
- i) Find the variance of y_i , and the covariance between y_i and y_j ($i \neq j$).
 - ii) Hence, or otherwise, obtain an expression for the variance of the sample mean.
- b) i) State (without proof) the expression for the variance of the estimator of the number of population units which belong to a specified class.
- ii) If $N = 1000$ and $n = 100$, show that the above variance cannot exceed 2253, whatever be the unknown number of units belonging to the specified class. [20]
- 3.a) Describe briefly the procedure in linear systematic sampling. What are its disadvantages? How can you modify this procedure to remove these disadvantages?
- b) State and prove a result to substantiate the truth in the statement:
- 'The variance of the systematic sample mean can be reduced by arranging (or numbering) the population units in such a way that the units in each systematic sample are as heterogeneous as possible with respect to the characteristic under study'. [20]

Group D

Answer all the questions.

4. In a simple random sample of 50 households drawn with replacement from a total of 250 households in a village, only 8 were found to possess transistor radios. These households had respectively 3,5,3,4,7,4,4 and 5 members. Estimate the total number of households that possess transistor radios and the total number of persons in these households along with their standard errors. [20]
5. Draw a circular systematic sample of 5 households from a village which has 24 households and describe your procedure clearly. [15]

INDIAN STATISTICAL INSTITUTE

One year Evening Course in Statistical Methods
and Applications: 1976-77 ✓

PERIODICAL EXAMINATION
Part II

Paper IV: Methods of Statistical Inference:
Theory and Practical

Date: 30.3.77

Maximum Marks: 75

Time: 2 hours

Note: Attempt any four questions. All questions
carry equal marks.

1. Talcum power is packed into tins by a machine. A random sample of 11 tins is drawn, and their contents are found to weigh (in lb.) as follows.
0.44, 0.51, 0.49, 0.52, 0.45, 0.48, 0.46, 0.45, 0.47, 0.45
and 0.47
Test of the average packing can be taken to be 0.50 lb.
 2. The height of six randomly chosen sailors are in inches: 65, 63, 68, 69, 71, 72.
Those of 10 randomly chosen soldiers are: 61, 62, 65, 66, 69, 69, 70, 71, 72 and 73.
Do these data suggest that sailors are on the average taller than soldiers?
 3. Out of 10 randomly chosen boys of ^acertain school 3 were found using spectacles. Can we conclude from these data that at least 25% boys of the school use spectacles.
 - 4.a) In a certain area the average rainfall during the Monsoon season June-September was required to plan for providing the irrigation facilities. Data on several years were obtained from the meteorological department. Based on the data devise a test to verify that the average rainfall in any given season is not more than a specified value.
(State your steps clearly and precisely).
 - b) In order to introduce new security methods, the record of the number of major thefts in a railway stockyard per week was taken over the last 100 weeks and it was observed that on the average there were 1.5 major thefts per week. However the statistician reported that the average number was more than 2, basing his computations on the assumption that the thefts follow a Poisson distribution. Analyse the statistician's report and give your comment whether you would agree with the statistician.
 - 5.a) In order to see whether a longer lapse of time between the last day of the class and the time of final examination makes a difference in the performance of the students, a class of N students is divided into two groups of m and n students. The first batch of n students was given the examination immediately and the second batch one month later. It is observed that there is more variation in the scores of the second batch. Suggest a procedure to verify whether the variation is less in the first batch of students than in the second batch. to be
 - b) Two different additives are compared to see which one is better for improving the durability of concrete. A sample of m observations from the first mix and a sample of n observations from the second are taken. Based on the data it is proposed to make the statement that the variance of the first mix is twice that of the second mix, suggest a procedure to test this statement. (State your alternative and assumptions clearly).
-

~~INDIAN STATISTICAL INSTITUTE~~

One year Evening Course in Statistical Methods
and Applications: 1976-77

PERIODICAL EXAMINATION ✓

Part II

Paper V: Sample Surveys (Theory and Practical)

Date: 20.4.77

Maximum Marks: 75

Time: 2 hours

Note: Answer all questions.

Marks allotted for each question are given
in brackets [].

- 1.a) Explain the purpose of stratification in sample surveys.
- b) What are the factors that are taken into consideration in allocating the sample size to different strata ?
- c) Obtain the expression for gain in precision (reduction of variance) due to stratification with proportional allocation, over a simple random sample (of the same size) without replacement from the unstratified population. Give the interpretation of the expression when the strata means are all equal. [8+4+(6+2)]=[20]
2. Describe Lahiri's method for drawing a 'pps' sample from a population and prove that this method ensures inclusion of the units in the sample with the preassigned sizes. [8+7]=[15]
3. The households in a town are to be sampled in order to estimate the average amount of assets per household that are readily convertible into cash. The households are stratified into a high-rent and a low-rent stratum. A house in the high rent stratum is thought to have about nine times as much assets as one in the low-rent stratum and S_h is expected to be proportional to the square root of the stratum mean.

There are 4000 households in the high-rent stratum and 20,000 in the low-rent stratum.

- a) How would you distribute a sample of 1000 households between the two strata ?
- b) If the object is to estimate the difference between assets per household in the two strata, how should the sample be distributed? [20]
4. From 35 farms with a total geographical area of 5759, a sample of 5 farms drawn with probability proportional to the area of the farms gave the data shown below:

<u>Farm No.</u>	3	18	28	34	35
Total area	52	110	300	410	430
Area under a crop	10	24	59	72	103
Proportion of area under the crop	0.1923	0.2182	0.1967	0.1756	0.2395

Estimate the average and the total area under the crop with respective standard errors.

[20]

INDIAN STATISTICAL INSTITUTE
One year Evening Course in Statistical Methods
and Application 1976-77 ✓

PERIODICAL EXAMINATION
Part II
Paper VIa: Design and Analysis of Experiments
Theory and Practical

Date: 4.5.77

Maximum Marks. 75

Time: 2 hours

Note: Answer all questions.

Marks allotted for each question are given in brackets [].

1. Describe the following terms:

(a) Treatment	(c) Randomization	
(b) Plot	(d) Replication	[4 x 5] = [20]

2. For a Randomized Block Design experiment with t treatments in b blocks of t plots each, discuss

(a) the lay-out	
(b) the model and the estimation of parameters	
(c) the analysis of variance table.	[5+12+8] = [25]

3. Following table gives the nitrogen content in milligrams of red clover plants inoculated with five different trifolii strains and a control in a random order.

Control	Trifolii strains				
17.5	19.4	17.7	17.0	20.7	14.3
19.4	32.6	24.8	19.4	21.0	14.4
19.1	27.0	27.9	9.1	20.5	11.8
16.9	32.1	25.2	11.9	18.8	11.6
20.8	33.0	24.3	15.8	18.6	14.2

- (a) Write the analysis of variance table and the model.
- (b) Give estimates of treatment effects and the error.
- (c) Test the equality of treatment effects.

[20+6+4] = [30]

INDIAN STATISTICAL INSTITUTE
 One year Evening Course in Statistical Methods
 and Applications: 1976-77

PERIODICAL EXAMINATION
 Part II

Paper VIb: Control Charts and Acceptance Sampling
 Theory and Practical

Date: 18-5-77

Maximum Marks. 75

Time: 2 hours

- Note: 1) Answer all questions
 2) The marks allotted to each question are provided in [].

1. Control charts for \bar{X} and R are maintained on the tensile strength in pounds on a certain yarn. The subgroup size is 5. The values of \bar{X} and R are computed for each subgroup. After 25 subgroups, $\Sigma X = 514.8$ and $\Sigma R = 120.0$. Compute the values of 3-sigma limits for the \bar{X} and R charts, on the assumption that the process is in statistical control. [12]
- 2.a) Explain the terms, specifications and process capability.
 b) Explain how an examination of various relationships of process capability and specification limits can be useful for economical gain in industry. [8+12]=[20]
3. Write short notes on any three of the following.
 a) Chance cause variation
 b) Statistical control
 c) Rational subgroup
 d) Interpretation of \bar{X} , R chart. [18]
4. The following data were obtained over a 20-day period to initiate \bar{X} and R control charts for a quality characteristic of a certain manufactured product that had required a substantial amount of rework. The subgroup size was 5.

subgroup number	\bar{X}	R	subgroup number	\bar{X}	R
1	177.6	13	11	179.8	9
2	176.6	8	12	176.4	8
3	178.4	22	13	178.4	7
4	176.6	12	14	178.2	4
5	177.0	7	15	180.6	6
6	179.4	8	16	179.6	6
7	178.6	15	17	177.8	10
8	179.6	6	18	178.4	9
9	178.8	7	19	181.6	7
10	178.2	12	20	177.6	10

- a) Examine the data for homogeneity.
- b) The specification for the above quality characteristic is given as 178 ± 10 . Determine the control limits for the \bar{X} , R charts to be used in future. [18+7]=[25]

INDIAN STATISTICAL INSTITUTE

One-year Evening Course in Statistical Methods and Applications:
1976-77

Part II Final Examinations
Delhi and Calcutta

Paper IV (a): Methods of Statistical Inference: Theory

Date: 27.6.77

Maximum Marks: 75

Time: 2 hours

Note: You may attempt any part of any question. Any score over 75 will be counted as 75.

1. If x_1, x_2, \dots, x_n is a sample of n independent observations drawn from a normal distribution with mean and variance σ^2 , and $s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{(n-1)}$ where $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$, then prove that

- (i) s^2 is an unbiased and a consistent estimator of σ^2 .
(ii) s^2 is a sufficient statistic for σ^2 .

$$\underline{15+7+3} = \underline{25}$$

2. (a) x_1, x_2, \dots, x_n are independent random variables with common frequency function $f(x; \theta)$ where θ is unknown. Explain the method of maximum likelihood for estimating θ on the basis of observations on x_1, x_2, \dots, x_n .
(b) Find the joint maximum likelihood estimators of μ and σ^2 for a normal distribution $N(x; \mu, \sigma^2)$. Drive moment estimators also.

$$\underline{15+10+7} = \underline{32}$$

3. Let X_1, X_2, \dots, X_n ($n \geq 5$) be n independent and identically distributed normal variables, each with mean m and variance σ^2 . What are the probability distributions of the following random variable?

- (i) $Y_1 = \frac{1}{n} \sum_{i=1}^n X_i$
(ii) $Y_2 = \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - m)^2$
(iii) $Y_3 = \frac{1}{\sigma} \sqrt{\frac{n}{\sum_{i=2}^n (X_i - m)^2 / (n-1)}}$
(iv) $Y_4 = \frac{1}{4} \sum_{i=1}^4 (X_i - m)^2 / \left[\frac{1}{n-4} \sum_{i=5}^n (X_i - m)^2 \right]$

$$\underline{2+3+4+5} = \underline{14}$$

4. Describe briefly the procedure of one-way analysis of variance for testing the hypothesis $H_0 : \mu_1 = \mu_2 = \dots = \mu_k$, on the basis of sample of size $n = n_1 + n_2 + \dots + n_k$, where n_i is the size of sample drawn from the i th normal population with mean μ_i and variance σ^2 .

/15/

4. x_1, x_2, \dots, x_n are independent observations from a normal population with mean μ and variance σ^2 . Describe how you will test the hypothesis $H_0 : \mu = \mu_0$ against the alternative $H_1 : \mu \neq \mu_0$, assuming (i) σ^2 is known, and (ii) σ^2 is not known.

/7+7/ = /14/

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

- (a) Two kinds of errors of a test and power of a test.
- (b) Properties of maximum likelihood estimations.
- (c) Chi-square Test for goodness of fit.
- (d) Test of significance for multiple and partial correlation coefficients.
- (e) Cramer - Rao in-equality.

/8+8+8/ = /24/

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods
and Applications: 1976-77

Calcutta and Delhi

PART II FINAL EXAMINATIONS

Paper IV: Methods of Statistical Inference.
Practical

Date. 29.6.77

Maximum Marks: 75

Time: 2 hours

Note. Attempt as much as you can. Maximum score
can be 75. Marks allotted for each question
are given in brackets [].

1.a) The means of two large samples of 1000 and 2000 mem-
bers are 67.5 and 68.0 inches. Can the samples be regarded
as drawn from the same population with standard deviation
2.5 inches? [8]

b) Average of 150 observations from population with standard
deviation 4.3 is 16.2. Find 99% confidence limits
for population mean. [7]

2.a) Mean and variance of 10 readings on the length of a given
rod is 20 inches and 7.3 inches². Does the result contradict
the assumption that the length of rod is 19.9 inches?
Find 90% confidence interval for the length of rod, [7+4]=[11]

b) Two samples of sizes 9 and 8 have the sum of squares of
deviations from their respective means equal to 160 inches²
and 90 inches². Do you agree that the samples are drawn from
the same normal population? [10]

3. The weight of ten boys before they are subjected to a
change of diet and after a lapse of six months are recorded
below:-

Sl.No.	1	2	3	4	5	6	7	8	9	10
Before	109	112	98	114	102	97	88	101	89	91
After	115	120	99	117	105	98	91	99	93	89

Test whether there has been any significant gain in weight
as a result of the change of diet. [15]

4. A random sample of size 140 was drawn from a certain popu-
lation. The frequency distribution is given below. Fit
a normal distribution and test the goodness of fit.

Class interval	Frequency
-5 and below	4
-5 - 0	8
0 - 5	20
5 - 10	35
10 - 15	43
15 - 20	14
20 - 25	10
25 and above	6

[20]

5. The following data relate to lives in hours of sample lamps taken from batches of electric lamps. Test whether the batch differ amongst themselves in average length of life.

<u>Batch</u>	<u>Sample size</u>	<u>Life of individual lamps</u>
A	7	1600, 1610, 1650, 1680, 1700, 1720, 1800
B	5	1580, 1640, 1690, 1700, 1750
C	8	1500, 1550, 1600, 1620, 1640, 1660, 1740, 1820
D	6	1510, 1520, 1530, 1570, 1600, 1680

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods
and Applications: 1976-77

Calcutta and Delhi

PART II FINAL EXAMINATIONS

Paper VI: Sample Surveys Theory

Date: 1-7-77

Maximum Marks: 75

Time: 2 hours

Note: Attempt Questions 1, 2 and any two from the rest. Explain clearly the symbols which you may use in your answers. Marks allotted for each question are given in brackets [].

1. Answer any three of the following: -

- a) What are the different sources of errors in a sample survey? Explain, very briefly, how these errors can be controlled.
- b) State, giving reasons, the conditions under which the use of stratified sampling may be recommended in a large-scale survey.
- c) What is a two-stage design and when should it be used?
- d) How and when, the need for selecting units with unequal probabilities from a population, is justified? [3 x 7] = [21]

2. Explain clearly an operational procedure for selecting 10 units with EQUAL PROBABILITY and without replacement from a population of 490 units, with the help of a table of random numbers. [YOU ARE NOT REQUIRED TO DRAW AN ACTUAL SAMPLE] [14]

3. A population of N (known) units has N_1 (unknown) units belonging to a specified class C . Obtain an unbiased estimator of N_1 as well as an unbiased estimator of its variance, under SRSWOR. [8+12] = [20]

4. Consider a population consisting of ' k ' strata and the following four sampling procedures for the estimation of 'population total':

- i) Unstratified SRS with replacement.
 - ii) Stratified SRS with replacement, and with equal allocation.
 - iii) Stratified SRS with replacement, and with proportional allocation.
 - iv) Stratified SRS with replacement, and with optimum allocation.
- a) State (without proof) the customary unbiased estimator of the population total for each of the above mentioned sampling procedures, total sample size being the same for all the designs.

b) Obtain also the variance of the estimator in each case. [4 x (2+3)] = [20]

- 5.a) Consider a two-stage sampling procedure used for a population which consists of 'N' first-stage units (fsu), each of which contains 'M' second-stage units (ssu). Let a sample of 'n' fsu's be selected and then from each selected fsu, let 'm' ssu's be selected, let $y_{ij} = km$.

Assume also that selection of units at first stage is done by SRS with replacement, second stage by systematic sampling. Suggest an unbiased estimator of the population mean. [8

- b) Describe the cumulative total method and Lehiri's method for drawing a PPS sample. Describe situations in which each method is advantageous over the other. [12

INDIAN STATISTICAL INSTITUTE

One-year Evening Course in Statistical Methods and Applications:
1976-77

Part II Final Examinations
Delhi and Calcutta

Paper V (b) : Sample Surveys (Practical)

Date : 4.7.77

Maximum Marks : 75

Time : 2 hours

Note : This paper in all carries 95 marks. You may answer as many questions as you like but the maximum you can score is 75.

=====

Q.1 From a list of 3042 names and addresses, a simple random sample of 200 names showed on investigation 38 wrong addresses. Estimate the total number of addresses needing correction in the list and find the standard error of this estimate.

/15/

=====

Q. 2 The N_i 's and σ_i 's are given for 5 strata into which a population is divided. Obtain the optimal values of n_i 's (sample size of the i th strata) and the corresponding variance of the estimator if the population mean is to be estimated, assuming SRSWR sampling scheme within strata. N_i and σ_i are the i th stratum standard deviation

i	N_i	σ_i
1	3780	28.5
2	5260	18.6
3	8200	27.6
4	4160	27.2
5	2980	16.8

24380

/25/

=====

Q. 3 From 35 farms with a total geographical area of 5759 hectares, a sample of 5 farms drawn with probability proportional to the total area of the farms gave the data shown in the following table. Estimate the total area under the crop and its standard error.

Table : Area under the crop and the total area of 5 sample farms selected with pps to the total area.

(in hectares)

Farm No.	1	2	3	4	5
Total area	52	110	300	410	430
area under a crop	10	24	59	72	103

/25/

- Q. 4 The number of inhabitants (in thousands) for each of a Srs without replacement of 20 cities, drawn from a group of 196 large cities, for the census year. (X_i) and the current year (Y_i) is as given below. The total population of 196 cities in the census year is also known to be 22919.

X_i	Y_i	X_i	Y_i
76	80	60	57
138	143	507	630
67	67	172	260
29	50	121	113
381	364	77	89
120	115	64	63
61	69	64	77
78	106	56	142
66	86	40	60
130	145	40	64

Obtain estimate of the total population in the current year by ratio method of estimation. Estimate its bias and mean square error.

/30/

INDIAN STATISTICAL INSTITUTE

One year Evening Course in Statistical Methods
and Applications: 1976-77

PART II FINAL EXAMINATIONS: Calcutta and Delhi

Paper VIa: Design and Analysis of Experiments
(Theory and Practical)

Date: 6.7.77

Maximum Marks: 75

Time: 2 hours

Note. Answer as much as you can. However, maximum
you can score is 75. Level of significance is
to be taken as 5%.

1. Write short notes on

- i) Randomisation
ii) Latin Square Design

iii) Factorial Experiment.

[6 x 3] = [18]

2. The following table gives 8 observations on the products
of three manufacturing plants.

Plants		
A	B	C
6.0	6.0	7.0
7.0	6.5	8.0
7.0	6.0	

- i) Compute the Analysis of Variance table.
ii) Describe the model and give the estimator of the
parameters in the model.
iii) Test the equality of plant effects of A and B,
and A and C.

[12+7+6] = [25]

3. EITHER

Data is available from a Randomised Block Design experiment
with subsamples. There are 4 treatments, 5 blocks and 3
subsamples (observations) per plot.

Describe the analysis of variance with an appropriate model.

[25]

OR

The following table gives the yields of four strains of
wheat in five randomized blocks (pounds per plot) with
two missing values

Strain	BLOCK					Observed total
	1	2	3	4	5	
A	32.3	34.0	34.3	35.0	36.5	172.1
B	33.3	-	36.3	36.8	34.5	140.9
C	30.8	34.3	35.3	32.3	35.3	168.5
D	-	26.0	29.8	28.0	20.8	112.6
Observed total	96.4	94.3	135.7	132.1	135.6	594.1

3. (contd).

- i) Estimate the missing values.
 - ii) Set up the ANOVA.
 - iii) What is the standard error for comparing treatment A and B? $[10+10+5]=[25]$
- =====

4. The following data are the results of an experiment to study the influence of time of bleeding on plasma phospholipid in lambs, with or without diethylstilbestrol treatment.

Diethylstilbestrol treatment

Time of bleeding	With		Without	
	AM	PM	AM	PM
	17.53	32.00	8.53	39.14
	21.07	23.80	20.53	26.20
	20.80	28.87	12.53	31.33
	17.33	25.06	14.00	45.80
	20.07	29.33	10.80	40.20

- i) Estimate the factorial effects of time of bleeding, diethylstilbestrol treatment and their interaction.
 - ii) Error sum of square for the above data is 379.92. Would you say that the influence of time of bleeding is significant? $[10+7]=[17]$
-

INDIAN STATISTICAL INSTITUTE

One-year Evening Course in Statistical Methods and Applications:
1976-77

Part II Final Examinations
Delhi and Calcutta

Paper VI (b): Control Charts and Acceptance sampling
(Theory and Practical)

Date: 8 July 1977 Maximum Marks: 75 Time: 2 hours

- 1(a) Point out the correct one with argument. A process is said to be in a state of statistical control if
- i) It produces all the items within the specification limits
 - ii) All the points on the control chart (s) are within the control limits.
 - iii) The points on the control charts are evenly distributed around the central line
 - iv) None of the above three is necessarily correct.

(b) Briefly explain the following:

Principle of rational sub-grouping plays an important role in operating a control chart.

6.6/

2. (a) Distinguish between chance and assignable causes of variation.
- (b) The length of a machined component is distributed normally with $\mu = 10.5$ and $\sigma = 0.2$. The component less than 10 or more than 11 are to be scrapped. Find the proportion of scrapped items. Unit of measurement is m.m.
- (c) Fifty samples of waste extraction from a carding machine are collected and measured in gms. The sample size is 5. The \bar{X} and \bar{R} are found to be 15.16 and 5.755 respectively. Find suitable control limits.

4.6.5/

=====

3. Write short notes on any two of the following

- a) Statistical aspects of tolerances
- b) Specification Vs capabilities
- c) Control chart for fraction defective and number defective
- d) Process capability

/6.6/

=====

4. Obtain an attribute single sampling plan to satisfy

$$\Delta QL = 0.02$$

$$LTPD = 0.08$$

$$\alpha = 0.05$$

$$\beta = 0.10$$

/12/

=====

5. Draw O.C. and A.O.Q. curves for single Sampling Acceptance/Rectification plan given that

$$N = 10,000, \quad n = 200, \quad c = 0$$

given $\Delta QL = 0.005$, what is producer's risk? What is ΔOQL of the plan?

/12/

=====

6. Write short note on any four of the following:

- a) AOQL
- b) Consumer's risk
- c) Producer's risk
- d) LTPD
- e) A.T.I.

/3.3.3.3/

=====