

INDIAN STATISTICAL INSTITUTE  
 One-year Evening Course (1980-81) : Part I  
 Periodical Examination

Descriptive Statistics - Theory-cum-Practical

Date: 3.9.80

Maximum Marks: 100

Time: 2 hours

Answer any four questions

- 1(a) Define the following terms giving suitable examples:  
 Individual, population, sample, attribute, variate.
- (b) Explain briefly the advantages of the sampling approach over complete enumeration. (15+10)=(25)
2. Define median and mode of a set of observations on a variate, taking care to mention all possible situations.  
 Compute median and mode for the following frequency distribution:

<u>monthly wages</u> (Rs.)	<u>no. of workers</u>
0-100	17
100-200	42
200-300	78
300-400	129
400-500	84
500-600	62
600-700	41
700-800	29
800-900	12
900-1000	6
<u>total</u>	<u>500</u>

(10+15)=(25)

3. The following shows the average marks scored by three groups of students appearing at a certain examination along with the corresponding s.d.'s. Find the pooled values of average and s.d. taking all three groups of students together:

<u>group number</u>	<u>no. of students in the group</u>	<u>average marks</u>	<u>s.d. of marks</u>
1	35	42.7	15.2
2	64	47.2	16.8
3	11	53.1	17.2

(25)

4. You are given  $n$  observations on a variate  $x$  denoted  $x_1, x_2, \dots, x_n$  and you apply the transformation:  

$$u_i = \frac{x_i - A}{B}, \quad i = 1, 2, \dots, n,$$
 where  $A$  and  $B$  are constants.  
 Establish the relationship between mean of  $x$  and mean of  $u$  and also that between the s.d. of  $x$  and s.d. of  $u$ .  
 What happens if  $A = \bar{x}$  and  $B = \text{s.d. of } x$ ? Examine fully. (25)

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

- (a) The ratio chart.  
 (b) Advantages of arithmetic mean over other averages.  
 (c) Moments of a frequency distribution and their uses.

(25)

INDIAN STATISTICAL INSTITUTE  
One-year Evening Course in Statistical Methods  
and Applications: 1980-81

Part I

PERIODICAL EXAMINATIONS

Paper III: Economic Statistics (Index Numbers  
and National Accounts)

Date: 1.10.80

Maximum Marks: 100

Time: 2 hours

Note: Answer all the questions. Marks allotted to each question are given in brackets.

1. Define an Index Number. Mention the well-known index number formulae distinguishing clearly between the 'aggregative' and 'average' type index numbers. Is it possible to interpret the Laspeyre's and the Paasche's formulae as 'average' type indices? Justify your answer. [5+13+8]=[26]
2. What are the 'time reversal' and 'factor reversal' tests? Examine whether the Laspeyre's, Paasche's, Edgeworth-Marshall and Fisher's formulae satisfy these tests. [8+16]=[24]
3. The following table gives the average prices and quantity consumed per person of important food items by the agricultural labourers in a village during December 1974 and December 1975. Calculate the Laspeyre's and Fisher's price index and Paasche's quantity index for this group of items for the period December 1975 using the period December 1974 as base. [50]

Item	December 1974		December 1975	
	average price per kg. (in Rs.)	quantity purchased per person (in gms.)	average price per kg. (in Rs.)	quantity purchased per person (in gms.)
	(1)	(2)	(4)	(5)
1. Rice (medium)	2.55	200	2.08	250
2. Rice (coarse)	2.33	500	1.82	700
3. Wheat (coarse)	1.94	1000	1.46	1000
4. Barley (coarse)	1.90	200	1.36	180
5. Arhar Dal	3.90	100	2.69	110
6. Urd Dal	2.86	75	2.95	60
7. Gram Dal	3.33	115	2.70	95
8. Masur Dal	3.02	100	2.44	75
9. Mung Dal	3.15	150	2.88	125
10. Jowar	1.66	800	1.19	700
11. Bajra	1.76	450	1.12	475
12. Maize	1.59	1000	1.02	1200
13. Khesari Dal	2.31	500	1.68	600
14. Jagi	1.61	2050	1.18	2000

INDIAN STATISTICAL INSTITUTE  
One Year Evening Course: 1980-81  
Economic Statistics  
Part - I  
PERIODICAL EXAMINATION

1980-81: E322

Date: 29.10.1980

Maximum Marks: 100

Time: 2 hours

Note: Answer all questions.

1. Attempt any two of the following:

- (a) What preliminary adjustments are to be made before time series data can be properly analysed ?
- (b) Indicate the factors that give rise to seasonal fluctuations. Why is it important to measure these fluctuations ?
- (c) Describe the method of moving averages for determining trend values. What are the advantages and disadvantages of this method ?

(15 + 15) = [30]

2. Discuss the characteristics of the logistic curve.

[30]

3. Fit a logistic curve to the census populations of England and Wales given below:

Year	Population (in millions)
1861	20.1
1871	22.7
1881	26.0
1891	29.0
1901	32.5
1911	36.1

[40]

INDIAN STATISTICAL INSTITUTE  
One Year Evening Course on Statistical  
Methods and Application

1980-81: E 22

Part - I

Paper II : Descriptive Statistics

PERIODICAL EXAMINATION

Date: 5.11.80

Maximum Marks: 100

Time: 2 hours

Note: Answer any two questions from Group A and  
Q.No.4 from Group B.

Group A

1. Suppose you have fitted a linear regression  $Y = a + bx$  by least squares method using  $n$  observation-pairs  $(x_1, y_1), \dots, (x_n, y_n)$  on the two variates  $x$  and  $y$ . Find the mean and variance of the expected value  $\hat{Y}$ , and also the mean and variance of the residual  $e = y - \hat{Y}$ . Hence explain what is measured by the correlation coefficient  $r$  between the two variates.

[30]

2. Define the correlation ratio  $\eta_{yx}$  and show by any method that  $\eta_{yx}^2 \geq r^2$ , where  $r$  is the correlation coefficient. (You need not prove every result).

What is measured by  $\eta_{yx}$  and by the discrepancy  $\eta_{yx}^2 - r^2$ ? Explain fully.

[30]

3. Show how one can fit a linear equation of the form  $y = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k$  by least squares method using  $n$  sets of observations of the form  $(y_i, x_{1i}, \dots, x_{ki})$ ,  $i = 1, 2, \dots, n$ .

Define the multiple correlation coefficient  $R_{y \cdot x_1 x_2 \dots x_k}$  and give expressions which may be used to compute this coefficient.

[30]

p.t.o.

Group B

4. Either

Compute the correlation coefficient between household income (x) and expenditure on food items (y) from the joint distribution given below:

monthly household income (Rs.)	monthly expenditure on food (Rs.)						Total
	101-200	201-300	301-400	401-500	501-600	601-700	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
201-400	18	2					20
401-600	1	12	4				17
601-800		2	14	3			19
801-1000			5	11	4		20
1001-1200				4	6	3	13
1201-1400					8	2	10
Total	19	16	23	18	18	5	99

[30]

Or

The following results are based on a study of three variates: y = performance score in mathematics,  $x_1$  = intelligence score, and  $x_2$  = average number of hours spent on mathematics per week.

The data were collected for  $n = 20$  students.

$$\Sigma y = 1019$$

$$\Sigma x_1 = 2027$$

$$\Sigma x_2 = 405$$

Corrected sums of squares and products matrix:

$$\begin{matrix} & y & x_1 & x_2 \\ y & \left( \begin{array}{ccc} 4532 & 2692 & 787 \\ & 4372 & 534 \\ & & 348 \end{array} \right) \\ x_1 & & & \\ x_2 & & & \end{matrix}$$

Find the multiple regression equation of y on  $x_1$  and  $x_2$ , and compute the multiple correlation coefficient  $R_{y \cdot x_1 x_2}$ .

[30]

5. Practical Record.

[10]

INDIAN STATISTICAL INSTITUTE  
One Year Evening Course in Statistical  
Methods and Applications

Part - II

Sample Surveys (Theory and Practical)

PERIODICAL EXAMINATION

Date: 9.3.81

Maximum Marks: 100

Time: 2 hours

Note: Do as many questions as you can. The paper carries 110 marks but the maximum you can score is 100 marks. Marks allotted are given at the end of each question.

1. (i) Define simple random sampling and show that
- (a)  $P_{i,r} = 1/N$  for  $i = 1, \dots, N$  and  $r = 1, \dots, n$  in both SRSWR and SRSWOR and that (b) sample mean is an unbiased estimate of the population mean.

(ii) Obtain variance of the sample mean.

(iii) Obtain an unbiased estimate of the variance of sample mean.

(10+5+5) = [20]

2. (i) What do you mean by proportional and Neyman allocations in stratified sampling.

(ii) Show that

$$V_N(\bar{y}_{st}) \leq V_p(\bar{y}_{st}) \leq V_{ran}(\bar{y})$$

(5+15) = [20]

3. (i) What do you mean by 'pps' sampling ?

(ii) Define the general scheme of ups wor and give an unbiased estimate for the population mean.

(iii) Give an unbiased estimate for mean and the expression for its variance in case of upswr.

(5+5+10) = [20]

4. The following data gives the size of landholdings (in acres) and indebtedness (in 00' Rs.) for 15 households.

Sr. No.	1	2	3	4	5	6	7	
Size of landholding	2	10	2.5	3	1	0.5	0	
Indebtedness	30	8	30	25	20	10	12	
Sr. No.	8	9	10	11	12	13	14	15
Size of land holding	1.5	0.5	0	8	10	15	7	2
Indebtedness	30	25	25	12	10	2	9	20

- (i) Estimate the average (per household) indebtedness on the basis of a SRSWOR of size 7. Obtain an unbiased estimate of the variance of estimate.
- (ii) Stratify the households into three strata on the basis of size of land holdings and estimate the average indebtedness on the basis of a sample of 7 households drawn according to (a) proportional allocation (b) Neyman-allocation (based on size of land holdings) Obtain an unbiased estimate of variance in case of proportional allocation.
- (iii) Select a sample of 7 households with probabilities proportional to size of land holdings and estimate the average indebtedness.

$$(10+20+10) = [40]$$

5. A SRSWOR of size  $n$  out of  $N$  units in a population is drawn and two variables  $(y, x)$  are observed. Obtain the expression for the covariance between the sample means  $\bar{y}$  and  $\bar{x}$ .

INDIAN STATISTICAL INSTITUTE  
One Year Evening Course in Statistical Methods  
and Applications

1980-81: E621

Part - II

Design of Experiments (Theory and Practical)

Date: 16.3.81

Maximum Marks: 100

Time: 2 hours

1. Describe the role of randomisation and replication in the planning of an experiment with suitable examples from agriculture and industry.

[25]

2. What is a linear model ?

Write down an appropriate linear model, stating clearly all the assumptions for two way classified data with one observation per cell. Write down the analysis of variance table for this model, giving expressions for the different sums of squares. Describe clearly a method for testing any hypothesis of interest in such a problem.

[25]

3. The observations given below in some coded units correspond to three district treatments.

Treatment	1	:	40,	58,	55,	45,	45,	60,	70
"	2	:	72,	50,	55,	58,	62,	65,	56
"	3	:	40,	33,	65,	68,	42,	43,	52

- (i) Test for the equality of the three treatment means.  
(ii) Test whether the effects of treatment 1 and treatment 3 are equal.

[50]

---



## Sample Surveys (Theory and Practical)

## PERIODICAL EXAMINATION

Date: 23.3.81

Maximum Marks: 100

Time: 2 hours

Note: Do as many questions as you can. The paper carries 110 marks but the maximum score is limited to 100 marks. Marks allotted are given at the end of each question.

- 1.(a) There are  $N$  villages in a district. The average no. of live stock owning households per village (say  $\bar{X}$ ) is known. Give four sampling strategies for estimating the total live stock in the district.
- (b) Obtain Bias and MSE of any one of the sampling strategies.
- (c) Obtain estimates of bias and MSE of the sampling strategy discussed in (b) above.

$$(5+10+10) = [25]$$

- 2.(a) There are 100 tea gardens in India. Give a suitable sampling strategy to estimate ratio of total female workers to total <sup>male</sup> workers for the 100 tea gardens, on the basis of a sample of 10 tea gardens.
- (b) Obtain Bias and MSE of the strategy.

$$(5 + 10) = [15]$$

- 3.(a) Let a poor peasant household be defined as that which owns land upto 5 acres. There are  $N$  households in Block which have been divided into 4 subsets according some consideration. Give a suitable sampling strategy to estimate the percentage of poor peasant households in the block. Also give estimate for the percentage of land owned by all poor pleasant households (to the total land).
- (b) Give Mean and MSE of the sampling strategy.
- (c) Give an estimate of the MSE.

$$(5+10+5) = [20]$$

4. Investigation for 150 factories were carried on. The average sale per factory for the investigation year was found to be Rs.20 millions. A simple random sample of 15 factories yielded the following data:

Sr. No.	1	2	3	4	5	6	7	8
Grossprofit (Rs. millions)	13	11	1	1	2	2	1	2
Sales (Rs. millions)	62	51	5	3	8	6	9	7
Sr. No.	9	10	11	12	13	14	15	
Grossprofit (Rs. millions)	2	2	10	18	4	4	3	
Sales (Rs. millions)	12	11	42	50	47	48	22	

- (a) Give simple, ratio, regression, difference and product estimates for  $\bar{Y}$ , the average Grossprofit per factory for the investigation year.
- (b) Obtain the estimated MSEs of the estimates.
- (c) Obtain the estimated relative efficiencies of the regression estimate over the other estimates.
- (d) Give estimates for the following: (i) Ratio of Grossprofit to Sales, (ii) Percentage of the factories whose Grossprofit is Rs.1 crore or above, (iii) percentage of the Grossprofit for the factories in (ii) to the total grossprofit for all factories.
- (20+10+10+5) = [45]
5. Show that  $s_{xy}$  is unbiased for  $S_{xy}$  in SRSWOR and for  $\sigma_{xy}$  in SRSWR.

[5]

One Year Evening Course in Statistical  
Methods and Applications: 1980-81

## PERIODICAL EXAMINATION

## Industrial Statistics (Theory and Practical)

Date: 6.4.81

Maximum Marks: 100

Time: 2 hours

Note: Answer all the questions.

Marks allotted to each questions are  
provided in the margin.Answers should be to the point and  
brief.

- 1.(a) Explain what is meant by 'quality' of a product.
- (b) Describe different kinds of quality characteristics which are encountered in quality control and give three examples of quality characteristics in each category.
- (c) Give five practical examples of quality characteristics where Poisson law holds good. State briefly the conditions for applicability of this law.

(5+15+10) = [30]

- 2.(a) Explain what is meant by 'a process being in a state of statistical control'.
- (b) The following measurement data were obtained on certain variable quality characteristic for 10 successive subgroups. The specification limits are LSL = 60 and USL = 80.

Sub-group no.	Measurements			
	1	2	3	4
1	67	70	68	62
2	66	69	63	64
3	66	67	62	66
4	64	68	65	67
5	70	63	65	66
6	68	71	69	66
7	65	67	65	67
8	69	65	68	67
9	66	65	66	65
10	61	63	61	60

Contd..... Q.No.2

- (i) Carry out an  $\bar{X}$ -R chart analysis and comment whether the process is in a state of statistical control (Hint:  $D_4 = 2.282$ ,  $A_2 = 0.729$ ,  $d_2 = 2.059$ ).
- (ii) Estimate process capability and indicate whether process is capable to meet the specification limits.
- (iii) Estimate the percentage of defective items being produced.

$$(10+15+5+10) = [40]$$

3. The number of defects (c) observed on 16 Dial Telephone Assemblies were as follows:

Assembly No.	No. of defects (c)	Assembly No.	No. of defects (c)
1	3	9	3
2	8	10	5
3	6	11	6
4	6	12	5
5	4	13	9
6	2	14	5
7	3	15	4
8	8	16	6

- (i) Compute the control limits  $UCL_c = \bar{c} + 3\sqrt{\bar{c}}$  and  $LCL_c = \bar{c} - 3\sqrt{\bar{c}}$  where symbols have the usual meaning.
- (ii) Examine the data for control and comment.
- (iii) Draw c-chart on a graph sheet.

$$(15+5+10) = [30]$$

One Year Evening Course in Statistical Methods  
and Applications: 1980-81

## Statistical Inference (Theory and Practical)

## PERIODICAL EXAMINATION

Date: 25.4.81

Maximum Marks: 100

Time: 2 hours

**Note:** Answer any four. For all the questions state your assumptions, Hypothesis and conclusions clearly.

1. A company buys electric switches in lots of 100. A lot is to be rejected if it contains more than 10 defectives. The purchase division of the company wants to take decision on each lot after examining 10 switches from the lot. What procedure would you adopt if you were in charge of the purchase ?
2. The mean drying time for a brand of spray paint is known to be 90 seconds. The research division of the company that produces this paint believes that by adding a new chemical to the paint, the mean drying time can be reduced. To test this claim, the paint with the chemical additions is sprayed on 15 surfaces and the mean drying time was recorded. The mean and standard duration computed from these measurements are 86 seconds and 4.5 seconds respectively. Do these data justify the claims of the research division ?
3. A plastic sheet manufacturing company contemplates replacing the machine in use by a new machine. Uncontrollable factors cause fluctuation in the thickness of the plastic sheets. The company would like to purchase the new machine if the new machine reduces the fluctuation. The thickness in millimeters of 10 specimens from each of the machines is given below.

Existing machine : 226, 228, 226, 225, 232, 228, 227,  
229, 225, 230.

New machine : 225, 224, 224, 225, 230, 230, 226,  
228, 225, 232.

Would you recommend replacement of the existing machine by the new machine ?

4. An anthropologist measuring the length and breadth of skull obtains a correlation of  $r = .82$  based on 27 skulls. Assume that the data are taken from a bivariate normal population. Test  $H_0 : \rho = 0$  against  $H_1 : \rho \neq 0$ .
  
  5. In an experiment with flowers of a certain species theory predicts that there should be four flower types in the ratio 9:3:3:1. If an experiment produces the frequencies 120, 50, 40, 10 for these four types, is it compatible with the theory ?
-

INDIAN STATISTICAL INSTITUTE  
One Year Evening Course in Statistical  
Methods and Applications: 1980-81

Part - I (Supplementary)  
Descriptive Statistics (Theory and Practical)

ANNUAL EXAMINATION

Date: 14.1.81

Maximum Marks: 100

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

Group - A

Answer any two questions

1. (a) Define the mode of a set of observations and describe how it is computed in different types of situations.
- (b) Compare arithmetic mean and mode as measures of location.
- (5+10+15) = [30]

2. (a) Three variates  $x$ ,  $y$  and  $z$  are used in a correlational study. The calculations yield the values of three coefficients viz.  $r_{xy}$ ,  $r_{xy.z}$  and  $R_{x.yz}$ .

(i) Define the three coefficients mentioned above as clearly as you can.

(ii) For each coefficient mention the limits within which the coefficient must lie, and state when the coefficient exactly attains each of these limits.

- (b) What do you learn if you are told that  $r_{xy} = -0.5$  ?

(9+12+9) = [30]

3. Define the correlation ratio  $\eta_{yx}$ , and explain its meaning. Discuss the cases where (i)  $\eta_{yx} = 0$  (ii)  $\eta_{yx} = 1$  and (iii)  $\eta_{yx} = |r|$ . (Here  $r$  denotes the correlation coefficient.)

Describe how  $\eta_{yx}$  can be computed from a given set of data. What is the significance of the difference  $\eta_{yx} - |r|$  ?

(8+9+8+5) = [30]

Group - B

Answer any one question

4. The following shows the sums obtained for computation of moments from a frequency distribution of 82 5-room wood houses in a certain region of USA according to their costs ( $x$ ):

$$\Sigma fd = 10, \quad \Sigma fd^2 = 236, \quad \Sigma fd^3 = -90, \quad \Sigma fd^4 = 3032.$$

Contd..... Q.No.4

Here  $d = (x - 6500)/1000$  is the transformed variate, and  $f$  denotes frequency.

Compute the arithmetic mean, s.d. and  $g_1$  and  $g_2$  of the frequency distribution of  $x$ .

What do you learn from the values of  $g_1$  and  $g_2$  ?

(32+8) = [40]

5. Fit a Poisson distribution to the following data:

no. of defects	no. of articles
0	125
1	97
2	38
3	11
4	2
5 or more	2
Total	275

(Compute the expected frequencies. You need not plot any graph, nor test the goodness of fit.)

[40]

-----



INDIAN STATISTICAL INSTITUTE  
One Year Evening Course in Statistical  
Methods and Applications: 1980-81

Part - I (Supplementary)

Paper-III (Economic Statistics)

ANNUAL EXAMINATION

Date: 14.1.81

Maximum Marks: 100

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

Note: Answer any four questions. Marks are given in brackets [ ].

- 1.(a) Write a note on time reversal and factor reversal tests. Which of the index number formulas satisfy either of these tests ?
- (b) Show that Laspeyre's formula can be regarded as a weighted average type of formula.
- (20+5) = [25]
- 2.(a) (i) Define national income.  
(ii) Explain the concept of value added.  
(iii) Describe two methods for estimating national income.
- (b) Starting with Net National Product at factor cost indicate what other items you will require for estimating
- (1) Gross National Product at market prices.  
(2) Personal Income
- and state how you will arrive at the estimates with the help of these other items.
- [2+4+12+7) = [25]
- 3.(a) Describe a method of fitting for the logistic curve.
- (b) Show that the logistic curve has a point of inflexion.
- (20+5) = [25]
4. Write briefly on any two of the following:
- (a) Agricultural Statistics in India.  
(b) Industrial Statistics in India.  
(c) Population Statistics in India.  
(d) National Sample Survey Organisation.

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

p.t.o.

5. The table below gives the losses due to fire in different quarters of the year. Calculate the indices of quarterly variation.

Year	Losses (in million dollars)			
	Jan.-March	April-June	July-Sept.	Oct.-December
1949	108	161	149	169
1950	190	178	149	173
1951	209	178	161	183

[25]

INDIAN STATISTICAL INSTITUTE  
One Year Evening Course in Statistical Methods  
and Applications

Part - II .

Statistical Inference Theory and Practical

Date: 25.2.81

Maximum Marks: 100

Time: 2 hours

1. Let  $p$  be the probability of head of a coin. Suppose  $p$  is unknown and we want to estimate  $p$  on the basis of their independent tosses of the coin.

(a) Define an estimate of  $p$ .

(b) The following table gives four functions. Which of these are estimates of  $p$  ?

Outcome/	$T_1$	$T_2$	$T_3$	$T_4$
H H H	1	1	1	1
H H T	$2/3$	$2/3p$	$2/3$	$2/3$
H T H	$1/3$	$1/3p$	$2/3$	$2/3$
H T T	$1/3$	$1/3p$	$1/3$	$3/4$
T H H	$2/3$	$2/3p$	$2/3$	$2/3$
T H T	$2/3$	$2/3p$	$1/3$	$1/3$
T T H	$1/3$	$1/3p$	$1/3$	$1/4$
T T T	0	0	0	0

- (c) Which of  $T_1, T_2, T_3, T_4$  are unbiased estimates of  $p$  ?  
Which among  $T_1, T_2, T_3, T_4$  would you recommend as a good estimate of  $p$  ? Justify your answer.

[25]

2. As in the previous question, suppose you want to estimate  $p$ . Suppose instead of tossing the coin thrice you perform the following experiment. 'Toss the coin till you get 10 heads'. Then the no. of heads required to obtain 10 heads is a random variable, whose distribution depends on  $p$ . What is the minimum variance unbiased estimate of  $p$  in this experiment ? Prove the unbiasedness and minimum variance of the estimate.

[25]

3.(a) Suppose you have  $n$  observations from uniform  $(0, \theta)$ .  
Obtain the minimum variance unbiased estimate of  $\theta$ .  
Justify your claims.

(b) Suppose you have one observation from uniform  $(\theta, \theta + 1)$ .  
If you know further that  $\theta$  takes only even integer values  
i.e.  $\theta$  can be one of  $2, 4, 6, \dots$ ; then how will you  
estimate  $\theta$  on the basis of your observation ?

(c) Suppose you have two observations in the previous can what  
will then be your estimate of  $\theta$  ?

[25]

4. Draw a sample of size 10 from  $N(4, 1)$  and compute.

(a) The following estimates of  $\theta$

$$T_1 = \frac{1}{n} \sum x_i, \quad T_2 = \frac{\max(x_i) - \min(x_i)}{2},$$

$$T_3 = \text{Median}(x_i).$$

(b) The following estimates of  $\sigma^2$

$$T_1 = \frac{1}{n} \sum (x_i - 4)^2$$

$$T_2 = \frac{1}{n-1} \sum (x_i - 4)^2$$

$$T_3 = \frac{1}{n-1} \sum (x_i - \bar{x})^2$$

(c) Comment on the accuracy of your estimates.

[25]

---

INDIAN STATISTICAL INSTITUTE  
One Year Evening Course in Statistical Methods  
and Applications: 1980-81

Part - II

Inference Theory

FINAL EXAMINATION

26.5.81

Maximum Marks: 100

Time: 2 hours

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

1. Suppose  $X_1, X_2, \dots, X_n$  are independent observations from a population distributed as  $f_\theta$  where  $\theta$  is an unknown real parameter.
  - (a) Define an unbiased estimate of  $\theta$ .
  - (b) When  $f_\theta$  is  $N(\theta, 1)$  show that  $\bar{x} = \frac{\sum X_i}{n}$  is a minimum variance unbiased estimate of  $\theta$ .
  
2. Define a sufficient statistic. State (without proof) the factorisation theorem for sufficient statistic. Show that the sample mean is sufficient for  $\theta$  in the case of a sample from  $N(\theta, 1)$ .
  
3. (a) What is a Maximum Likelihood estimate.  
 (b) Find out the maximum likelihood estimates for
  - (i)  $\lambda$  of Poisson  $\lambda$ .
  - (ii)  $\sigma^2$  of  $N(0, \sigma^2)$ .
  
4. (a) Explain the method of least squares for estimating unknown parameters of a distribution.  
 (b) Suppose  $Y_1, Y_2, \dots, Y_n$  are independent Normally distributed random variables such that  $E(Y_i) = \alpha + \beta x_i$  and  $\text{Var}(Y_i) = \sigma^2$   $i = 1, \dots, n$ ; where  $x_1, x_2, \dots, x_n$  are known and  $\alpha, \beta, \sigma^2$  are unknown. Obtain the least square estimates of  $\alpha$  and  $\beta$ .  
 (c) Give an example of a situation where the above model is applicable.

5. Suppose you have 100 observations. Describe how you would test that these came from a Normal population.

6.(a) Let  $\bar{x}$  and  $s^2$  denote the sample mean and sample variance respectively for a sample of size  $n$  ( $\geq 2$ ) from  $N(\mu, \sigma^2)$ . Show that  $\bar{x}$  and  $s^2$  are independent.

(b) Let  $\bar{x}_1, s_1^2$  denote the sample mean and sample variance respectively for a sample of size  $n_1$  ( $\geq 2$ ) from  $N(\mu_1, \sigma^2)$ .

Let  $\bar{x}_2, s_2^2$  denote similar statistics for a sample of size  $n_2$  ( $\geq 2$ ) from  $N(\mu_2, \sigma^2)$ . Let  $T = \frac{[(\bar{x}_1 - \bar{x}_2) - \mu_1 - \mu_2]}{[\sigma^2 (\frac{1}{n_1} + \frac{1}{n_2})]^{1/2}}$

where  $\sigma^2 = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}$ . Show that  $T$  has a

$t$  - distribution with  $n_1 + n_2 - 2$  degrees of freedom.

---

INDIAN STATISTICAL INSTITUTE  
CALCUTTA AND DELHI

One-year Evening Course in Statistical Methods and Applications: 1980-81

Part-II: Annual Examinations

Statistical Inference - Practical

Date: 28.5.1981

Max. Marks: 75

Time: 2 hours

Note: Answer any three questions.

1. Data relating to heart weight (in gms.) of 12 male and 15 female cats have been summarized as follows:

<u>sex</u>	<u>mean</u>	<u>s.d.</u>
male	11.0	9.3
female	8.6	4.1

Examine separately (i) if there is a significant difference in the variance of males and females (ii) if there is significant difference in the mean of males and females, assuming that the true variances of the male and female populations are same.

[10+15]

2. (a) Mean and s.d. of 10 readings on the length of a given rod is 22.1 and 2.5 inches, respectively. Examine if the length of the rod is significantly larger than 21 inches.
- (b) Twenty samples were taken from a container of a particular brand of vanaspati. Four analysts were given five different samples each and asked to determine the melting point. The results are given below:

<u>Analysist:</u>	<u>individual determination of melting points</u>				
1	93.60,	94.64,	96.30,	93.62,	93.51
2	96.44,	96.53,	98.38,	97.00,	97.63
3	92.57,	94.01,	92.49,	93.29,	90.87
4	95.55,	95.90,	94.25,	95.80,	96.21

Examine if the analysts are consistent in their determination of the melting point.

[8+17]

3. (a) The co-efficient of correlation between the scores in two halves of a psychological test applied on a group of 20 students was 0.71. Examine if the correlation co-efficient is significantly different from zero.
- (b) The following data give the weight (in gms) of anterior muscles of both hind legs of 10 normal rabbits:

<u>rabbit no.</u>	1	2	3	4	5	6	7	8	9	10
<u>left leg:</u>	5.0	4.8	4.3	5.1	4.1	4.0	7.1	5.9	5.3	5.3
<u>right leg:</u>	4.9	5.0	4.3	5.3	4.1	4.0	6.9	6.3	5.2	5.5

Is there any difference between the two legs in respect of the average weight of the anterior muscle.

[10+15]

4. (a) The data given below represent the frequencies of offsprings in four classes:

<u>Class :</u>	1	2	3	4
<u>frequency:</u>	102	17	36	5

Examine if the observed frequencies in the four classes are in agreement with the expected ratios 9 : 3 : 3 : 1.

- (c) Inter-penetrating sub-samples.  
(d) Techniques for the assessment and control of non-sampling errors.

4. (b) The following data relate to the degree of speech defects and physical defects of a number of handicapped school children:

<u>Physical defect</u>	<u>Speech defect</u>		
	<u>mild</u>	<u>intermediate</u>	<u>serious</u>
<u>mild</u>	17	10	4
<u>intermediate</u>	21	32	50
<u>serious</u>	12	20	45

Examine if there is any association between the degree of the two types of defects.

OR

In an experiment designed to determine the relationship between the dose of a compost fertiliser ( $x$ ) and the yield of a crop ( $y$ ), the data were summarized as follows:

<u>no. of plots</u>	<u>means</u>		<u>corrected sum of squares and products</u>		
<u>n</u>	<u><math>\bar{x}</math></u>	<u><math>\bar{y}</math></u>	<u><math>S_x^2</math></u>	<u><math>S_y^2</math></u>	<u><math>S_{xy}</math></u>
15	10.8	122.7	70.6	98.5	68.3

- (i) Estimate the slope of the regression line of  $y$  on  $x$ .  
(ii) Examine if the true slope exceeds 1.5.

[10+15]



One-year Evening Course in Statistical Methods and Applications:1980-81

ANNUAL EXAMINATION

PART-II

Sample Survey(Theory)

Date:30.5.1981

Max.Marks:100

Time: 2 hours

Note: Attempt as much as you can. Maximum you can score is 100.

- Describe the linear systematic sampling procedure for selecting a sample of  $n$  units from a population of  $N$  units. Show that usually sample mean is not an unbiased estimator of the population mean. Give an unbiased estimator.
  - Describe circular systematic sampling and gave an unbiased estimate of the population mean.
  - How can one obtain an unbiased estimator of the variance of an estimator based on the systematic sampling? [15+5+5]
- Give unbiased estimates and prove their unbiasedness in the following cases:
  - population mean  $\bar{Y}$ , based on a stratified simple random sample of size  $n$  ( $= \sum_{s=1}^k n_s$ ) sample units are selected from the  $s$ th ( $s=1, \dots, k$ ) stratum of population with size  $N_s$ .
  - proportion of population units belonging to a category  $C$ , based on a simple random sample of  $n$  clusters from a population of  $N$  clusters, where each cluster is of the same size  $M$ .
  - population total  $Y$ , based on a two-stage sample of  $n$  villages and  $n_i$  households from the  $i$ -th selected village with SRS used at both the stages, assuming that the population consists of  $N$  villages and  $i$ -th village has  $M_i$  households. [5+5+10]
- Define proportional allocation and Neyman's allocation in stratified sampling. Prove that
 
$$V_{opt} \leq V_{prop} \leq V_{ran}$$
 where  $V_{opt}$ ,  $V_{prop}$  and  $V_{ran}$  stand for variances of the estimator of the population total based on stratified Neyman's optimum allocation srs, stratified proportional allocation srs and unstratified srs respectively. [25]
- Define ratio method of estimation for estimating the population mean.
  - Obtain mean and mean square error of the ratio estimator.
  - $NM$  population units are grouped at random to form  $N$  clusters of  $M$  units each. Compare the efficiency of sampling a clusters with simple random sample of  $nM$  units. [5+10+8]
- Write notes on any two of the following:
  - Complete enumeration versus sample survey,
  - Methods of selection of sample with probability proportional to size.
  - Inter-penetrating sub-samples.
  - Techniques for the assessment and control of non-sampling errors..

INDIAN STATISTICAL INSTITUTE  
 One Year Evening Course in Statistical Methods  
 and Applications: 1980-81  
 Part - II  
 Sample Survey - Practical

FINAL EXAMINATION

1.6.81                      Maximum Marks: 100                      Time: 2 hours

Note: Do as much as you can. Your maximum score will, however, not exceed 100 marks.

1. Three thousand factories in a city were divided into strata according to the number of workers employed. The following table gives no. of factories and mean and S.D. of output for each stratum.

Stratum No.	No. of workers	No. of factories	Output per factory (in '000 Rs.)	Standard deviation (in '000 Rs)
1	1- 49	1800	100	80
2	50- 99	500	250	200
3	100-249	350	500	600
4	250-999	250	1550	1600
5	1000 and above	100	2050	2500

A stratified total sample of 350 factories is to be drawn using SRSWOR in each stratum.

- (i) Obtain the sample sizes for each stratum in case of (a) proportional allocation (b) Neyman allocation.
- (ii) Obtain variance for the unbiased estimate for gross output in case of both the allocations.
- (iii) If without doing any stratification, a sample of 350 factories is drawn, using SRSWOR, from the population of 3000 factories, find variance of an unbiased estimate for gross output.

Contd..... Q.1-

(iv) Obtain the relative efficiency of Neyman allocation over (a) proportional allocation and (b) SRSWOR.

$$(10+20+15+5) = [50]$$

2. A simple random sample without replacement (SRSWOR) of 15 factories out of 100 factories yielded the following data for the investigation year.

Sr. No.	Gross profit (Rs. millions)	Sales (Rs. millions)
1	14	60
2	12	50
3	2	5
4	2	4
5	3	8
6	3	9
7	2	6
8	3	10
9	3	12
10	3	13
11	10	40
12	20	50
13	5	20
14	5	25
15	4	18

The total sales for all the 100 factories were found to be Rs.2500 millions.

- (i) (a) Give an estimate for  $\bar{Y}$ , the average gross profit per factory, in case information on sales data is not utilized.  
(b) Give ratio estimate for  $\bar{Y}$  using sales as auxiliary variable.

Contd..... Q.No.2

- (ii) Obtain the estimates of mean square errors of the two estimates in (i) above. Obtain relative efficiency of the ratio estimate over the other estimate.
- (iii) Obtain estimates for the following:
  - (a) ratio of average gross profit to average sales (for all the 100 factories).
  - (b) percentage of the factories (out of all 100 factories) whose gross-profits are Rs.5 million and above.
- (iv) Obtain the estimated mean square error of the estimate in (iii) (b) above.

$$(10+25+10+5) = [50]$$

3. Consider the data of Q.No.2 above as population data.

- (i) Draw a sample of 5 factories (out of the 15) with probabilities proportional to sales and give an unbiased estimate for the total gross profit for the 15 factories.
- (ii) Obtain an unbiased estimate of the variance of the estimate in (i) above.
- (iii) Draw a linear and circular systematic samples of 5 factories each and give unbiased estimates for the total gross profit (for all the 15 factories) based on the above samples separately.

$$(10+10+10) = [30]$$

---

INDIAN STATISTICAL INSTITUTE  
CALCUTTA AND DELHI

One-year Evening Course in Statistical Methods and Applications:1980-81

ANNUAL EXAMINATION

PART II

Design and Analysis of Experiments  
(Theory and Practical)

Date: 3.6.1981

Max.Marks:100

Time: 2 hours

Note: Attempt question 4 and as many as you like from the remaining three. The maximum number of marks you can score is 100.

- (a) Describe the role of linear Models in the design and analysis of Experiments. [13]

(b) State the basic principles of Experimental design. What are their roles and purposes? Discuss to what extent these principles are met in a Randomized Block Design. [25]
- State clearly the situations under which a factorial experiment is best applicable. Considering a  $2^2$  factorial design explain two methods of obtaining factorial effect totals. [9+16]
- Give the lay out and analysis of a Randomised Block Design with 'a' treatments and 'b' blocks. State clearly the assumptions behind the model. [20+5]
- An experimenter is studying the effect of five different formulations A,B,C,D,E of an explosive mixture used in the manufacture of dynamite on the observed explosive force. Each formulation is mixed from a batch of raw material that is only large enough for five formulations to be tested. Further-more the formulations are prepared by five operators. Data are given in the following table in the form of a Latin square design. Analyse the data and give your comments. [40]

Batches of material	Operators				
	1	2	3	4	5
1	24(A)	20(B)	19(C)	24(D)	21(E)
2	17(B)	24(C)	30(D)	27(E)	36(A)
3	18(C)	38(D)	26(E)	27(A)	21(B)
4	26(D)	31(E)	26(A)	23(B)	22(C)
5	22(E)	30(A)	20(B)	29(C)	31(D)

.....

## INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods  
and Applications: 1980-81

## Part - II

Paper VI(b): Industrial Statistics (Theory and Practical)

## FINAL EXAMINATION

5.6.81

Maximum Marks: 100

Time: 2 hours

Note: Question No.2 is compulsory. Answer any other two questions from the remaining questions. Marks allotted to each questions are provided in the margin. Answers should be to the point and brief.

1. Explain the following :

- (a) Quality of a product                      (c) Indifference Quality Level (I Q L)  
(b) Statistical Control                      (d) A Q L and  
(e) Consumer's Risk

(6 x 5) = [30]

2. A sugar mill sells sugar cubes in packets of net weight equal to 450 gms. Daily five packets are being sampled from the production line and weighed in the laboratory. The averages and ranges of the net weight in gms. for twenty-four days are as follows:

Day	Average	Range	Day	Average	Range
1	437.2	19.0	13	452.9	13.2
2	446.6	21.2	14	454.4	16.0
3	449.6	32.8	15	448.6	15.2
4	459.5	13.0	16	454.2	16.8
5	446.6	24.7	17	451.8	28.2
6	445.2	9.4	18	457.9	18.0
7	456.8	20.2	19	449.0	24.8
8	454.9	21.2	20	453.5	17.8
9	458.0	30.2	21	459.0	18.8
10	455.0	25.8	22	447.5	19.0
11	450.5	15.2	23	456.2	18.0
12	454.2	20.8	24	453.9	22.0

Contd.... Q.No.1

- (a) Carry out an  $\bar{X}$  - R chart analysis.
- (b) Draw an  $\bar{X}$  - R chart and examine whether the process is under control.
- (c) Obtain estimate of  $\mu$  and standard deviation under statistical control and hence compute the proportion of under-weight packets being produced by the mill.

$$(15+15+10) = [40]$$

3.(a) Write a brief note on hundred percent inspection as compared to sampling inspection. Describe procedure under double sampling plan by attributes.

(b) Write down a note on

Either

AOQL plan of Dodge and Romig describing its main features including the type of protection it provides.

Or

Indian Standard Plan by attributes (IS 2500 Part (1)) describing its main features.

$$(10+8+12) = [30]$$

4. For the sampling plan  $N = 3000$ ,  $n = 100$  and  $c = 1$

- (a) Draw the O.C curve.
- (b) Read the value of IQL approximately.

$$(25+5) = [30]$$

5. The following data give the results of inspection of enamel plates of a standard size for spots.

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

Contd..... Q. No.5

- (a) Analyse the data by means of a suitable control chart and examine if the process is under statistical control.
  
- (b) On the basis of results obtained in (a) obtain the probability that a plate selected at random from the process will have number of spots more than 8.

(22+8) = [30]

---



Part - I

PROBABILITY

Date: 3.12.80

Maximum Marks: 100

Time: 3 hours

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) Assume that peoples birthdays are equally likely to occur among the 365 days of the year (ignoring leap years). What is the probability that no two people in a group of  $n$  people will have a common birthday ?

- (b)  $n$  balls are distributed at random into  $r$  boxes. What is the probability that exactly  $K$  balls are in the first  $r_1$  boxes ?

$$(7+8) = [15]$$

- 2.(a) Show that if  $A$  is a subset of  $B$ , then  $P(A) \leq P(B)$ .

- (b) For any two events  $A$  and  $B$ , show that

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$(6+7) = [13]$$

- 3.(a) State Baye's rule for computation of conditional probability.

- (b) (i) An insurance company believes that people can be divided into two classes -- those that are accident prone and those that are not. Their statistics show that an accident-prone person will have an accident at some time within a fixed 1-year period with probability 0.4, whereas this probability decreases to 0.2 for a non-accident-prone person. If we assume that 30 per cent of the population is accident-prone, what is the probability that a new policy holder will have an accident within a year of purchasing his policy ?

- (ii) Suppose a new policy holder has an accident within a year of purchasing his policy. What is the probability that he is accident-prone ?

$$(5 + 7+8) = [20]$$

4. Suppose that an airplane engine will fail, when in flight, with probability  $1-p$  independently from engine to engine. Also suppose that the airplane will make a successful flight if at least 50 percent of its engines remain operative. For what values of  $p$  is a 4-engine plane preferable to a 2-engine plane ?

$$[10]$$

5.(a) State Chebyshev's Inequality.

- (b) It is known that the number of items produced in a factory during a week is a random variable with mean 50 and variance 25. What can be said about the probability that the production in a week will be between 40 and 60.

$$(5+7) = [12]$$

6. The life time in hours of a certain kind of radio tube is a random variable (X) having a probability density function given by

$$f(x) = \begin{cases} 0 & , x \leq 100 \\ \frac{100}{x^2} & , x > 100 \end{cases}$$

What is the probability that exactly 2 of 5 such tubes in a radio set will have to be replaced within the first 150 hours of operation? Assume that the events  $E_i$ ,  $i = 1, 2, 3, 4, 5$  that the  $i$ th such tube will have to be replaced within this time, are independent.

$$[12]$$

- 7.(a) The joint density of X and Y is given by

$$f(x,y) = \begin{cases} e^{-(x+y)} & , 0 < x < \infty, 0 < y < \infty \\ 0 & \text{otherwise} \end{cases}$$

Find the density function of the random variable  $X|Y$ .

- (b) A man and a woman decide to meet at a certain location. If each person independently arrives at a time uniformly distributed between 1 p.m. and 2 p.m. find the probability that the first to arrive has to wait longer than 10 minutes.

$$(8+12) = [20]$$

8. The random variables X and Y are said to have a bivariate normal distribution if their joint density function is given by

$$f(x,y) = \frac{1}{2\pi \sigma_x \sigma_y \sqrt{1-\rho^2}} \exp \left[ -\frac{1}{2(1-\rho^2)} \left\{ \frac{(x-\mu_x)^2}{\sigma_x^2} + \frac{(y-\mu_y)^2}{\sigma_y^2} - 2\rho \frac{(x-\mu_x)(y-\mu_y)}{\sigma_x \sigma_y} \right\} \right]$$

- (a) Show that X and Y are both normal random variables with respective parameters  $\mu_x, \sigma_x^2$  and  $\mu_y, \sigma_y^2$ .
- (b) Show that X and Y are independent when  $\rho = 0$ .

$$(10+10) = [20]$$

9. Write short notes on:

- (i) Central Limit Theorem  
 (ii) Probability Generating Function

Part - I  
Descriptive Statistics (Theory)

THEORETICAL EXAMINATION

Date: 5.12.80

Maximum Marks: 100

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

Note: Answer any four questions.

- 1.(a) Define median of a set of observations  $x_1, x_2, \dots, x_n$ .

Show that the sum  $\sum_{i=1}^n |x_i - A|$  is minimum when A is the median.

- (b) Explain fully why the arithmetic mean is preferred to the median in most applications. Can you mention some situations where the median is preferable to the arithmetic mean?

(10+15) = [25]

2. Define the moments of the frequency distribution of a variate x and show how they help in measurement of dispersion, skewness and kurtosis.

State what you know about the properties of the measures mentioned by you.

[25]

3. Define the concentration curve of the size distribution of a non-negative variate. Define also the concentration coefficient.

Discuss fully the cases where the concentration curve takes (i) its highest possible position and (ii) its lowest possible position.

Describe how the concentration coefficient can be computed from a frequency distribution of the variate.

(7+8+10) = [25]

4. Either

- (a) A regression equation of the form  $Y = a + bx$  has been fitted by least squares method to n observation-pairs  $(x_1, y_1), \dots, (x_n, y_n)$ . Let  $Y_i = a + bx_i$ ,  $i = 1, 2, \dots, n$ , and  $e_i = y_i - Y_i$ .

Find the means  $\bar{e}$  and  $\bar{Y}$ , and also show that

$$V(Y) = r^2 V(y) \text{ and } V(e) = (1-r^2) V(y).$$

Contd.....Q.No.4. Either (a)

Hence interpret the value of  $|r|$  or  $r^2$ , mentioning particularly the cases where  $r \rightarrow \pm 1$ .

- (b) What is the message conveyed by the sign of  $r$  ?  
(c) Show that  $-1 \leq r \leq 1$ . (Can you give two proofs of this ?)

(15+5+5) = [25]

Or

- (a) Define the correlation ratio  $\eta_{yx}$  and show that  $\eta_{yx} \leq |r|$  where  $r$  is the correlation coefficient. What is the interpretation of the difference  $\eta_{yx} - |r|$  ?  
(b) What is the major limitation of  $r$  as a measure of correlation between two variates ?  
(c) What can one learn from a scatter diagram showing the observations  $(x_1, y_1), \dots, (x_n, y_n)$  of a bivariate study ?

(12+6+7) = [25]

5. Suppose you are given observations on 3 variates denoted  $y, x_1$  and  $x_2$  for each of  $n$  individuals. Show how you can estimate a regression equation of the form

$$y = b_0 + b_1 x_1 + b_2 x_2$$

by least squares method.

Define the multiple correlation coefficient  $R_{y, x_1 x_2}$  and discuss its usefulness, mentioning necessary formulae. How do you compute it in practice ? What is the range of values that this coefficient can assume ? When does it take (i) the highest possible value and (ii) the lowest possible value ?

(12+5+8) = [25]

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

- (a) Nonsampling errors in statistical enquiries.  
(b) Different methods of collecting data.  
(c) Fitting a normal distribution to an empirical frequency distribution.  
(d) Concept and measurement of partial correlation.

(2 x 12 $\frac{1}{2}$ ) = [25]

INDIAN STATISTICAL INSTITUTE  
CALCUTTA AND DELHI

One-year Evening Course in Statistical Methods and Applications:1980-81

Final Part I Examinations

Descriptive Statistics(Practical)

Date: 8.12.1980

Max.Marks: 100

Time: 2½ hours

Note: Answer any three questions.

1. (a) The following table gives the frequency distribution of monthly wages of workers in a factory. Compute any two of the following (i) the standard deviation (ii) Quartile deviation, and (iii) the coefficient of variation.

<u>Monthly wages(Rs.)</u>	<u>No. of workers</u>
162.5 - 167.5	2
167.5 - 172.5	22
172.5 - 177.5	19
177.5 - 182.5	14
182.5 - 187.5	3
187.5 - 192.5	4
192.5 - 197.5	6
197.5 - 202.5	1
202.5 - 207.5	1

- (b) Certain data obtained from a study of a group of 1000 employees in a cotton mill were reported as follows:  
525 Hindus, 470 married, 147 married Hindus, 25 married Hindu males, 312 males, 42 Hindu males, 86 married males.  
Examine these frequencies for internal consistency. [20+13]

Fit a normal distribution to the following frequency distribution.

<u>Scores in English</u>	<u>No. of students</u>
14.5 - 19.5	9
19.5 - 24.5	11
24.5 - 29.5	10
29.5 - 34.5	44
34.5 - 39.5	45
39.5 - 44.5	54
44.5 - 49.5	37
49.5 - 54.5	26
54.5 - 59.5	8
59.5 - 64.5	5
64.5 - 69.5	1
	<u>Total 250</u>

Compute the expected frequencies, but you need not draw the fitted curve, nor apply the  $\chi^2$ -test for goodness of fit.

[33]

3. An experimental study was conducted for building up a formula for predicting the yield of dry bark in ounces ( $X_1$ ) of a plant from its height in inches ( $X_2$ ) and its girth in inches ( $X_3$ ) at a level six inches above the ground. Measurements on  $X_1, X_2$  and  $X_3$  were taken on each of 32 plants. The following means and corrected sums of squares and products were computed:

$$\text{Means: } \bar{X}_1 = 21.68 \quad \bar{X}_2 = 166.40 \quad \bar{X}_3 = 3.14$$

Corrected sums of squares and products:

$$\begin{aligned} s_{11} &= 6498.00 & s_{12} &= 9483.84 & s_{13} &= 321.26 \\ s_{22} &= 102767.64 & s_{23} &= 599.58 & s_{33} &= 33.95 \end{aligned}$$

- Find (i) the multiple correlation between  $X_1$  and  $X_2, X_3$ ,  
(ii) the multiple regression equation of  $X_1$  on  $X_2$  and  $X_3$ ,  
(iii) the partial correlation coefficient of  $X_1$  and  $X_2$  eliminating the effect of  $X_3$ .

[11x3=33]

4. (a) The means of two samples of sizes 50 and 100 respectively are 54.4 and 50.3 and the corresponding standard deviations are 8 and 7. Obtain the mean and the standard deviation of the sample of size 150 obtained by combining the two samples.  
(b) In a partially destroyed laboratory record of a correlational study the following results only are legible:

$$\text{Variance of } x = 9$$

$$\text{Regression equations: } 8x - 10y + 66 = 0 \text{ and } 40x - 18y = 214$$

What were (a) the mean values of  $x$  and  $y$ ,

(b) the standard deviation of  $y$ , and

(c) the coefficient of correlation between  $x$  and  $y$ ?

- (c) Following are the results based on heights and weights of 1000 men students:  $\bar{y} = 68$  inches,  $\bar{x} = 150$  lbs,  
 $s_y = 1.5$  inches,  $s_x = 20$  lbs,  $r = 0.6$

Aggarwal is five feet and six inches tall and Rao's weight is 120 lbs. Estimate Rao's height and Aggarwal's weight.

[11x3=33]

INDIAN STATISTICAL INSTITUTE  
 Statistical Methods and Applications: 1980-81  
 Part - I

1980-81: E35

Economics Statistics (Index Numbers, National Income Accounts)

EXAMINATION

Date: 10.12.80

Maximum Marks: 100

Time: 2 hours

Note: Answer question number 5 and any one from the rest. Marks allotted for each question are given in brackets [ ].

- 1.(a) Describe the index number problem with special reference to a price index number. Discuss how one should choose the items and the base year in constructing a price index number. (5+10) = [15]
- (b) What is the 'circular test'? Is there any advantage in using a formula which satisfies the circular test? Explain your answer. (5+5) = [10]
- (c) The following data relate to the group indices (I) of cost of living and the corresponding weights (w) for the non-manual class in a centre in India for the year 1970-73:

Group	1970		1971		1972		1973	
	I	w	I	w	I	w	I	w
Food	100	60	122	65	150	60	160	61
Clothing	100	10	111	8	126	9	140	11
Housing	100	15	115	13	135	12	145	14
Fuel and Light	100	7	114	6	130	8	142	5
Misc.	100	8	116	8	139	11	153	9

Using Laspeyre's formula calculate the cost of living index for 1973 with 1970 as base by the (i) the chain-base method and (ii) fixed-base method and compare the two indices.

[25]

P.T.O.

2.(a) What is meant by a consumer price index number ? How does it differ from a cost of living index number ?

(5+5) = [10]

(b) Describe the method of constructing the consumer's price index number for Industrial worker's in India explaining clearly the procedure of obtaining the weights.

[40]

3.(a) Define National Income. Describe briefly the different methods of measuring National Income.

(4+8) = [12]

(b) Given data on National Income at current prices over a number of years, what problem will you face in comparing them ? Discuss how you will deal with it.

(4+9) = [13]

(c) The following data relate to the Indian economy for the year 1977-78:

(i) Net domestic product at factor cost	=	75058	Crores
(ii) Personal Income	=	74294	"
(iii) Indirect tax less subsidies	=	9108	"
(iv) Gross domestic Capital formation	=	19429	"
(v) Private Income	=	75850	"
(vi) Net factor Income from abroad	=	- 264	"
(vii) Corporation Tax	=	1221	"
(viii) Gross National Product at factor cost	=	79776	"

Find out from the above data

Net National Product at factor cost ;

Corporate Saving ;

Depreciation of fixed Capital ;

Gross National Product at market prices ;

and Net Domestic Capital formation.

[25]



INDIAN STATISTICAL INSTITUTE

One-year Evening Course in Statistical Methods and Applications:  
1980-81

Part - I Final Examinations

DELHI AND CALCUTTA

Economics Statistics

(Time Series and Official Statistics)

Date: 12.12.1980

Maximum Marks:100

Time: 2 hours

Note: From group A Answer Question 1 and any two from the rest. From group B answer both the questions.

1. Compute the seasonal indices for the following data relating to the revenue expenditure, government of India.

Revenue Expenditure (Rs. crores)

<u>Year</u>	<u>April-June</u>	<u>July-Sept.</u>	<u>Oct.-Dec.</u>	<u>Jan.-March</u>
1971-72	78	62	56	100
1972-73	84	64	61	81
1973-74	92	70	63	72
1974-75	100	81	83	96

If the total expenditure in the first quarter of an year is Rs.185 crores, determine how much would be expenditure during remaining three quarters of the year. [27+6]

OR

- (a) Fit a straight line trend by the method of least squares to the following data, which relates to the yearly sales of a company.

<u>Year</u>	<u>Sales(000 Rs.)</u>
1970	112
1971	120
1972	124
1973	134
1974	125
1975	139
1976	143
1977	138
1978	149

- (b) Obtain the trend value for the year 1980.

- (c) Obtain the monthly trend equation and the trend value for January 1975. [21+4+7]

2. What purpose is served by analysing Time Series? Describe the different components of a time series. [22]
3. Describe briefly the methods known for determining the trend in a time series. [22]
4. (a) How do you propose to make adjustments for price variation and population variation in Time Series data in order to secure comparability?  
(b) A sequence has M numbers. Prove that in a moving average of order N there will be M-N+1 numbers.  
(c) Describe the method of fitting an exponential curve. [7+5+10]

Group B: Official Statistics

5. Write short notes on any two of the following:  
(i) Availability of industrial statistics in India  
(ii) Statistical System in India  
(iii) National Sample Survey Organisation. [9+9]
6. Write the names of 5 important publications of the United Nations. [5]