[1992-93 E-113]

INDIAN STATISTICAL INSTITUTE
ONE YEAR PART-TIME COURSE IN STATISTICAL METHODS
AND APPLICATIONS: 1992-93
GRADE 1: DESCRIPTIVE STATISTICS (THEORY)
PART-I FINAL EXAMINATION

Date: 14.1.93 | Maximum Marks: 100 | Time: 2 Hours

Note: Answer any four questions. Each question carries 25 marks.

1.(a) Write short note on any one of the following:

(i) Scrutiny of data, (ii) Measures of skewness.

(b) Discuss the effect of changes of location and scale (i.e., of linear transformation) on the following measures:

(i) Arithmetic mean, (ii) Standard deviation,

(iii) Simple correlation coefficient and

(iv) Regression coefficient in the simple regression.

[9+16=25]

2.(a) Define Range, Mean deviation and standard deviation as measures of dispersion. Compare the merits and demerits of the three measures.

(b) Obtain mean and standard deviation of n terms of an arithmetic series. Hence find the mean and standard deviation of first n natural numbers. What will be the value of the coefficient of skewness g_1 of first n natural numbers? [15+10=25]

3. Write down the density function of normal distribution. Prove that the 2r th central moment of normal distribution is

$$\mu_{2r} = (2r-1)(2r-3)\ldots 3.1.\sigma^{2r},$$

where σ^2 is the variance of the distribution.

Discuss in brief how you will fit a normal distribution from a set of observations with a reasonably large size. [25]

contd. ....2.

4.(a) Define partial correlation coefficient in a trivariate set up. Derive the expression of partial correlation coefficient in terms of the simple correlation coefficients. Give an example to show how simple correlation coefficient can give misleading inference whereas the corresponding value of the partial correlation coefficien can resolve the issue.

(b) Find the coefficient of variation of $|X|$ where

$X \sim N(0, \sigma^2)$. [18+7=25]

5.(a) Prove that for any set of positive values $x_1, x_2, \ldots, x_n$;

the AM $\geq$ GM $\geq$ HM. When do the equalities hold ?

(b) If $x_1 = y_1+y_2$, $x_2 = y_2+y_3$ and $x_3 = y_3+y_1$, where y_1, y_2 and y_3 are uncorrelated variables and each of which has zero mean and unit variance, find the multiple correlation coefficient between x_1 and the two variables x_2 and x_3.

(c) For a trivariate distribution, show that if $r_{12}=0$ then $r_{13}^2 + r_{23}^2 \leq 1$ where r_{12}, r_{13} and r_{23} are the simple correlations with usual meaning of subscripts. [12+8+5=2

INDIAN STATISTICAL INSTITUTE

One Year Part-time Course in Statistical Methods and Applications : 1992-93

PART I

Grade 1 : Descriptive Statistics (Theory and Practical)

Periodical Examination

Date : 8.10.1992 Maximum Marks : 100 Time : 2 Hours.

Note : Use separate answerscript for Group A and B.

GROUP A

(Theory) Max. Marks : 50

Note : Answer all questions. Allotted marks are given in brackets [ ] at the end of each question.

1. Write short notes on any two of the following :
   (a) Non-sampling Error
   (b) Scrutiny of Data
   (c) Quantiles
   (d) Coefficient of Variation. [8+8=16]

2. Give comments on the following sets of values :
   (a) AM = 10 and GM+HM=21
   (b) $Q_1 = 35.5$, $Q_2 = 49.2$ and $Q_3 = 42.7$
   (c) $MD_{\bar{x}} = 7.3$ and $MD_{Me} = 7.1$
   (d) $n = 15$, $\Sigma x_i = 60$ and $\Sigma x_i^2 = 240$
   (e) $n = 20$, $\Sigma x_i = 100$ and $\Sigma x_i^2 = 400$
   (f) Of the 200 students in a class, 48 percent are good in English, 90 percent are good in Mathematics and 70 students are good in both the subjects.

   [2+2+2+3+3+5=17]

3.(a) Define mean, median and mode of a frequency distribution. Give an example where median is a better measure of central tendency than mean.

p.t.o.

3.(b) For a frequency distribution of marks in History of n candidates (grouped in intervals of equal length ℓ as $c_0 - c_1, c_1 - c_2, \ldots, c_{k-1} - c_k$) the mean and s.d. were found to be μ and σ respectively. Later it was discover that the score x^0 in the interval $(c_{r-1} - c_r)$ was misre as x^* in the interval $(c_r - c_{r+1})$ in obtaining the frequency distribution. Find the corrected mean and s.d. corresponding to the corrected frequency distribution.

[6+11 = 17]

GROUP B

(Practical)

Note : This paper contains 50 marks. Answer all the questions.

1. The following table gives the frequency distribution of IQ for 230 six-year old children.

IQ	No. of children
60–69	5
70–79	17
80–89	42
90–99	60
100–109	55
110–119	38
120–129	13

(a) Compute mean, median and mode from the data after necessary modifications of the frequency distribution.

(b) Draw the histogram and less-than type ogive and hence find mode and median from the appropriate diagrams.

(c) Calculate numericallythe first quartile and the ninth decile. Also calculate the number of children with IQ more than 105.

[10+15+10=35]

2. The first three moments about the origin 51 kg. calculated from the data on the weights of 25 college students are given by :

$m_1' = -0.4$ kg.

$\sqrt{m_2'} = 1.2$ kg.

$(m_3')^{1/3} = -0.25$ kg.

Determine the mean, standard deviation and the coefficient of skewness.

[15]

INDIAN STATISTICAL INSTITUTE
ONE YEAR PART-TIME COURSE IN STATISTICAL METHODS
AND APPLICATIONS: 1992-93
GRADE 1: DESCRIPTIVE STATISTICS (PRACTICAL)
PART-I
FINAL EXAMINATION

Date: 15.1.93 Maximum Marks: 100 Time: 2 Hours

1.(a) In a trivariate distribution, the three variates x_1, x_2 and x_3 are measured from their respective means, i.e. $x_1 = X_1 - \bar{X}_1$, $x_2 = X_2 - \bar{X}_2$ and $x_3 = X_3 - \bar{X}_3$.

Given

$$\sigma_1 = .27 = \sigma_3 \quad ; \quad \sigma_2 = 2.4$$

$$r_{12} = 0.28 \quad ; \quad r_{23} = 0.49; \quad r_{13} = 0.51$$

calculate $r_{13.2}$. If $x_4 = x_1 + x_2$ obtain r_{42}, r_{43} and $r_{43.2}$.

(b) Two judges rank a number of competitors in a certain art competition as follows:

	Competitor											
	1	2	3	4	5	6	7	8	9	10	11	12
Judge A	5	1	4	2	7	3	6	8	10	9	11	12
Judge B	10	5	1	2	3	4	7	6	8	11	9	12

Measure the association between the judgement of the two judges by using Kendall's τ.

(c) Two independent random variables X and Y have the following variances

$$\sigma_X^2 = 36 \quad ; \quad \sigma_Y^2 = 16$$

calculate r_{UV} given U = X+Y and V = X-Y.

[(20+6+4)=30]

2. In calculating the moments of a frequency distribution based on 100 observations, the following results were obtained:

Mean = 9, variance = 19, $b_1 = 0.7$ (m_3+ive), $b_2 = 4$.

Later on it was found that one observation 12 was read as 21. Obtain the correct value of its first four central moments. [20]

contd. ...2.

3.(a) If $\log_{10} X$ is normally distributed with mean 4 and variance 4 find,

$$P\,(1.[illegible] < X < 83180000)$$

(b) Fit a [illegible] distribution to the following data with respect to the [illegible] of red blood corpuscels (x) per cell:

x	0	1	2	3	4	5
no.of cells (f)	142	156	69	27	5	1

(c) An owner of a small hotel with five rooms is considering buying T.V. sets to rent to room occupants. He expects that about half of his customers would be willing to rent sets and finally buys three sets. Assuming 100% occupancy at all times,

(i) what fraction of the evenings will there be more request than T.V. sets?

(ii) what is the probability that a customer who requests a T.V. set will receive one?

(iii) if the owner's cost per set per day is C, what rental R must be charged in order to break even in the long run?

[8+12+(3+8+4)=3

4. Practical record books and performance in class. [15

INDIAN STATISTICAL INSTITUTE
ONE YEAR PART TIME COURSE IN STATISTICAL METHODS
AND APPLICATIONS 1992-93
PART I
SUPPLEMENTARY EXAMINATION
GRADE 1: DESCRIPTIVE STATISTICS (THEORY AND PRACTICAL)

Date: 10.3.93 Maximum Marks: 100 Time: 2 hours

GROUP A

Maximum Marks: 50

Note: This part of the question paper contains 50 marks. Answer all questions. Allotted marks are indicated in brackets at the end of each question.

1. Define r^{th} order ''moment about an arbitrary origin'' of a frequency distribution. Derive the expression of r^{th} order central moment in terms of ''moments about an arbitrary origin'' of order r and less. Hence find the central moments up to fourth order in terms of raw moments (about zero). [15]

2. (a) Derive the Poisson distribution as a limiting form of the binomial distribution. Hence, find the expressions for mean and variance of Poisson distribution from those of binomial distribution.

   (b) If $X \sim N(0,\sigma^2)$, derive $E(|X|)$ and hence find the coefficient of variation of $|X|$. [10+10=20]

3. Using LS principle, obtain the multiple regression equation of y on x_1 and x_2 in terms of means, the standard deviations and the inter-correlations of the variables. [15]

contd. 2/-

GROUP B

Maximum Marks:50

1. Find the missing frequencies in the following incomplete frequency distribution. The median and mode of the distribution are 22.5 and 23.08 respectively and the total frequency is 75.

Marks	Frequency	Marks	Frequency
0-5	2	20-25	-
5-10	-	25-30	16
10-15	7	30-35	-
15-20	13	35-40	3

[15]

2. Calculate the correlation coefficient from the following table giving the ages of 100 husbands and their wives in years.

Age of husbands → / Age of wives ↓	20-30	30-40	40-50	50-60	60-70
15-25	5	9	3		
25-35		10	25	2	
35-45		1	12	2	
45-55			4	16	5
55-65				4	2

[20

3.(a) The third decile and the upper quartile of a normal distribution are 56 and 63 respectively. Find the mean and variance of the distribution. Hence find the median and the 37th percentile of the distribution. [15

INDIAN STATISTICAL INSTITUTE
One Year Part-time Course in Statistical Methods and Applications: 1992-93

Part-I Periodical Examination
Grade:2 Probability and Sampling Distributions

Date: 19.11.92 Maximum Marks:100 Time: 2 Hrs.

Note: Attempt all questions. The paper carries 110 marks but your maximum score can not exceed 100 marks. The marks allotted are given at the end of each question.

1. (i) Give axiomatic definition of probability

(ii) State and prove the formula for the probability of an event $E=A\cup B$ which is the union of two events A and B.

(5+10)=[15]

2. (i) State and prove Baye's theorem on conditional probability.

(ii) Of the total production during the first four days of a week 30%, 25% and 35% of the items were produced during the first, second and third day respectively. The percentage of defective items during these days is supposed to be 2%, 3%, 2% and 4% respectively. An item is selected randomly out of all these items

(a) Find the probability that it is of acceptable quality (non-defective)

(b) If the selected item is found to be defective then find the probability that it was produced on the fourth day of the week.

(12+13)=[25]

3. (i) Using the characteristic function find the mean and variance of a binomial distribution.

(ii) Derive the characteristic function of gamma distribution.

(10+10)=[20]

4. The classification of one thousand households was found as follows with some missing entries:

	(i) Poor	(ii) Middle economic status	(iii) Rich	
(i) Scheduled caste and scheduled tribes	150	-	10	-
(ii) Brahmin	20	-	10	100
(iii) Others	-	-	-	650
	600	-	90	1000

p.t.o.

(a) complete the table

(b) obtain the marginal probabilities

(c) if a house-hold is selected randomly and found to be of poor economic status then what is the probability it belongs to Brahmin Category.

(d) what is the probability that a non-Brahmin house-hold is not rich. [20]

5. The joint probability density of (X,Y) is given by

$$f(x,y) = K-x-2y \qquad 1 < x < 2,\ 2 \leq y \leq 5$$
$$= 0 \qquad \text{elsewhere}$$

(i) Find the value of K

(ii) Find mean and variance of Y

(iii) Let $Z=3+5x$. Find mean and standard deviation of Z.

(iv) Find conditional distributions of X. [20

6. The amount of funds (X) received annually from external source for supporting the research projects of an Institute are found to be as follows for the last ten years:

In Rs. lakhs : 15,20,10,12,30,35,18,25,40,10

Assuming the distribution of X as normal with mean 20 and standard deviation 15 find the probabilities that

(a) $P(X < 25)$ (b) $P(X < 30/ X > 15)$ (c) $P(X < 25/ X < 30)$
(d) $P(X < 30/ X < 25)$ [10

INDIAN STATISTICAL INSTITUTE
ONE YEAR PART TIME COURSE IN STATISTICAL METHODS AND APPLICATIONS
PART I
FINAL EXAMINATION
GRADE 2: PROBABILITY AND SAMPLING DISTRIBUTIONS

Date:25.1.93 Maximum Marks:100 Time:$2\frac{1}{2}$ Hours

Note: Attempt All the questions. The paper carries 115 Marks but your maximum score can not exceed 100 marks. The marks allotted are specified at the end of each question.

1. (i) Give the axiomatic definition of probability

(ii) For any events A and B show that

$$P(AB) - P(A)\,P(B) \leq 1/4$$

(iii) If A and B are independent events then show that

(a) A and $\bar{B}$ and (b) $\bar{A}$ and $\bar{B}$ are also independent.

(4+8+8)=[20]

(i) State and prove Chebyshev's inequality.

(ii) State and prove Weak Law of Large Numbers.

(iii) Let $\bar{x}$ denote the mean of a sample of size n from a population whose variance is 10.9. It is required that $|\bar{X} - u| < 5.2$ in at least 90 percent cases. Find the range for sample size n in case (a) distribution of population is unknown (b) population is normal. (10+10+10)=[30]

3. (i) Define characteristic function

(ii) Three companies functioning independently have the profils (x_1, x_2, x_3) distributed normally with means (4.5, 2.6, 5.2) and standard deviations (1.6, 2.8, 7.9), in Rs. crores). Find the probability that the average profit of these companies is at least Rs.5.10 crores. (5+10)=[15]

(i) State Central Limit Theorem for i.i.d. r.vs

(ii) A mechine is producing steel cylinders whose lengths $x_1, x_2, \ldots$ have an unknown distribution with mean 3.5 cms and standard deviation 0.2 cms. To make an instrument thirty such cylinders are chosen randomly and joined end to end. What is the probability that the combined length of these 30 cylinders is less than one meter. (5+15)=[20]

5. Let T denote the time to send a message over a single communication channel of a particular type. Suppose the p.d.f. of T is given by

$$f(t) = \frac{1}{2} e^{-\frac{1}{2}t} \qquad t \geq 0$$

Let X denote the time required to send four independent messages one after the other on the channel.

(i) Find the pdf of X

(ii) Obtain mean and standard deviation of X. (5+10)=[15]

6. (i) Define t - statistic

(ii) Define X^2 (Chi-square) distribution and derive its characteristic function. Hence prove its reproductive property. (5+10)=[15

INDIAN STATISTICAL INSTITUTE

One Year Part-time Course in Statistical Methods and Applications : 1992-93

Grade 2 : Probability and Sampling Distribution

Periodical Examination

PART I

Date : 22.12.1992 Maximum Marks : 100 Time : 2 Hours

Do all questions. The paper carries 110 marks but your maximum score can not exceed 100 marks. The marks allotted to each question are specified.

1.(i) Define characteristic function (c.f.)

(ii) Establish relationship between moments and c.f. of a distribution.

(iii) Let $\{X_\alpha\}$, $\alpha = 1,\ldots,n$ be independent random variables such that X_α follows normal distribution with mean μ_α and variance σ_α^2. Obtain the distribution of

$$Y = \sum_{\alpha=1}^{n} W_\alpha X_\alpha, \quad W_\alpha \text{ : constants}$$

What are the mean and variance of Y ?

[5+10+15] = [30]

2.(i) State and prove Chebyshev's inequality.

(ii) Let $\{X_1,\ldots,X_n\}$ be a random sample from a population with mean μ (unknown) and coefficient of variation 0.5. Find the range for the sample size n so that Probability that the relative deviation $\frac{|\bar{x}-\mu|}{\mu}$ of sample mean $\bar{x}$ from μ does not exceed 0.1 is atleast 0.95.

[15+10] = [25]

3.(i) Define the terms (a) Convergence in probability (b) Asymptotically normal (c) Weak Law of Large Number (d) Central Limit Theorem

p.t.o.

3.(ii) State C.L.T. for i.i.d.rvs. (Lindgeberg-Levy case)

(iii) State De'Moivre's theorem for binomial proportions.

[20+5 +5] = [40]

4.(i) In a population 10 percent of the people are supposed to be suffering from a particular disease. Thirty batches of people, each batch consisting of 15 persons, are selected randomly and inspected for the incidence of disease. Find the probability that of the total number of persons inspected atleast fifty will be found to suffer from the disease.

(ii) Show that sample mean converges in probability to population mean.

[10+5]=[15]

INDIAN STATISTICAL INSTITUTE
ONE YEAR PART TIME COURSE IN STATISTICAL METHODS AND APPLICATIONS
PART-I
FINAL EXAMINATION
GRADE 3: TESTS OF STATISTICAL HYPOTHESES (PRACTICAL)

Date:22.1.93 Maximum Marks:100 Time: 2 hours

1. The percentage distribution of household in a certain district A was obtained on the basis of a large scale survey. A random sample of 292 households in another district B was taken and the frequency distribution was obtained. Test whether the structure of household size of B differs significantly from that of A

Family size	Percentage dist.(A)	Observed frequency (B)
1 - 3	15.4	42
4 - 6	42.5	127
7 - 9	30.8	93
10 - 12	9.6	25
13 - 15	1.4	4
16 - 18	0.3	1

[20]

2. With a view to test whether a certain additive will increase the strength of a certain type of synthetic yarn, two random samples, are without the additive and the other with the additive were taken. The results noted are as follows

Type	Sample size (n)	$\bar{X}$	$\Sigma(X_i-\bar{X})^2$
With additive	10	82.5	324
Without Additive	15	76.5	370

Can we conclude that the strength has increased due to the use of the additive? [15]

contd. .....2.

3. Packages made by a standard method for a long time indicate that $\sigma_0 = 11.75$ for the variation of weight. A less expensive and less time consuming new method is tried, it shows that for a sample of size 20, the quantity $\sum_{i=1}^{20}(x_i-\bar{x})^2 = 2656.08$.

Test whether this method results in an increase of the variability of the quality of the packages in term of these weights. [10]

4. A coin is thrown 50 times where head occurs 29 times and Tail 21 times. Test whether the coin is defective so as to show head more often than Tail. [10]

5. Two experiments A and B take repeated measurements. Test whether B's measurements are more accurate than A's

A's measurements (in mm)		B's measurements (in mm)	
12.47	12.44	12.06	12.34
11.90	12.13	12.23	12.46
12.77	11.86	12.46	12.39
11.96	12.25	11.98	
12.78	12.29	12.22	

[15]

6. The weights of ten boys before they are subjected to a change of diet and after a lapse of six months are recorded below.

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

Test regarding the ratio (σ_x/σ_y) [15]

7. Practical Note Book. [15]

1992-93|E133

INDIAN STATISTICAL INSTITUTE
ONE YEAR PART TIME COURSE IN STATISTICAL METHODS AND APPLICATIONS
PART I
FINAL EXAMINATION
GRADE 3: TESTS OF STATISTICAL HYPOTHESES (THEORY)

Date: 21.1.93 Maximum Marks: 100 Time: 2 Hours

Note: Answer all the questions. Marks are indicated along the margin.

1.(a) What is meant by test of (statistical) significance? Explain.

(b) Explain expressions 'Nue hypothesis', 'Alternative hypothesis', 'acceptance region' 'region of rejection'

(c) Discuss what is meant by one and two sided tests. What are the critical values for a one-sided test and a two-sided test at (i) 0.025% (ii) 0.15% level of significance. [15]

2. When is a statistical test said to be unbiased? Define likelihood ratio test.

Consider n Bernoullian trials with probability of success P. Test the hypothesis $H_0: p = 1/2$ against $H_1: p \neq 1/2$. [20]

3. Describe how one may test for association between two attributes in case the sample size is not large. [15]

4. Why is it that, for data of somewhat similar types, we sometimes use Fisher's t-test and at some other times the paired t-test?

Explain the distinction between the two tests with examples. [15]

5.(i) A die is thrown 60 times in which an odd face shows 36 times. Test whether the die shows odd faces more often than even faces. [5]

(ii) A random sample of size 10 from a normal population gives $s^2=90$. Test at the 5% lavel of significance:

$$H_0: \sigma_0^2 (=80) \text{ against } H_1: \sigma^2 \neq \sigma_0^2 (=80)$$ [5]

p.t.o.

(iii) The correlation co-efficient between nasal length and stature for a group of 20 indian adult males was found to be 0.203. Test whether there is any correlation between the characters of the population.

(iv) 60 students are examined in statistics, Physics and Mathematics. The total correlation between the scores obtained as

$$r_{12}=0.64, \; r_{13}=0.75, \; r_{23}=0.82$$

(where x_1, x_2 and x_3 are taken to denote the scores in Statistics, Physics and Mathematics). It is conjectured that the correlation between x_1 and x_2 is due to the influence of x_3 on both x_1 and x_2.

Test whether this conjecture seems valid in the light of the above data. [10]

6. Define and explain the following

(a) type I and type II error and power of test.

(b) Neyman and Pearson's theory of testing of hypotheses

[(2)+(8)=10]

[1992-93] E142

INDIAN STATISTICAL INSTITUTE
ONE YEAR PART TIME COURSE IN STATISTICAL METHODS AND APPLICATIONS
PART I
GRADE 4: APPLIED STATISTICS (THEORY AND PRACTICAL)
PERIODICAL II EXAMINATION

Date: 24.12.92 Maximum Marks: 40 Time: 2 Hours

GROUP-A (Theory)

1.(a) Let P_i be the unconditional probability of getting the i^{th} sample at the first stage of randomisation and $P_{j/i}$ be the conditional probability of getting $(i-j)$ th sample at the second stage of randomisation given that the i^{th} sample has been selected at the first stage of randomisation. Let t_{ij} be the estimate based on the $(i-j)^{th}$ sample. Show that

$$E(t) = E_1E_2(t)$$

and $$V(t) = E_1V_2(t) + V_1E_2(t)$$ [5]

(b) Define a two stage sampling procedure and provide an unbiased estimate for population total based on two-stage sampling design.

Also show that variability of the estimate can be decomposed into two components, namely, (1) variation due to first stage units and (2) variation due to second stage units within the first stage unit. [15]

Examine the important formula for the calculation of price Index numbers in the light of the various tests developed for the purpose. [10]

(a) Write short notes on:

(1) Secular trend (2) Seasonal variation discuss (in brief) different methods of measuring trend.

(b) Describe the method of fitting the curve of the type

$$Y_t = K + ab^t$$ [12+8=20

contd. ...2

GROUP B (PRACTICAL)

Note: The questions given in this section carry 50 marks. Attempt as much as you can but the maximum attainable score is 40.

1. The following data relate to a stratified sample taken from a population of 105 factories divided into 3 strata with 45,32 and 28 factories respectively:

Stratum Sl.No.	No. of factories	Type of sampling	Sample size	No. of workers (y)
1	45	SRSWOR	6	75,101,5,78,78,45
2	32	SRSWOR	5	247,238,259,129, 223
3	28	Systematic (two interpenetrating) sub-samples	4	407,326,481,445
			4	335,412,503,348

Estimate the total no. of workers and also the standard error of your estimate. [20

2.(a) Following data relate to the production of pure sulphuric acid in India for the years 1962-66. Fit a quadratic trend to the data and obtain the trend values for the year 1962-66.

Year	:	1962	1963	1964	1965	1966
Production (in 000 tonnes)	:	470	565	630	600	590

(b) Given the following data, test which of the index nos. (among Laspeyre's, Paasche's, Edgeworth-Marshall and Fishers' index nos.) satisfy the time reversal and factor reversal test.

Year	Rice Price	Rice quantity	Wheat Price	Wheat quantity	Jowar Price	Jowar quantity
1927	9.3	100	6.4	11	5.1	5
1934	4.5	90	3.7	40	2.7	3

[(18+12)=30]

contd. ...3

GROUP C

VITAL STATISTICS

Maximum Marks: 20 Time: 2 Hours

1. What do you mean by a life table. Distinguish between cohort and current life tables. Describe the structure of a complete life table. Explain how the different columns of a life table may be computed on the basis of observed age-specific mortality rates. [2]+[2]+[4]+[2]

2. Write short notes on the following :

   (a) Crude death rate [1]

   (b) Standardised death rate [3]

   (c) General fertility rate [2]

   (d) Total fertility rate [2]

   (e) Net reproduction rate [2]

1992-93 | E132

INDIAN STATISTICAL INSTITUTE

One Year Part-time Course in Statistical Methods and Applications : 1992-93

PART I

Grade 3: Tests of Statistical Hypotheses

Periodical Examination

Date : 22.10.1992 Maximum Marks : 100 Time : 2 Hours

1. A purchaser of bricks believes that the quality of the bricks is deteriorating. From past experience, the mean crushing strenth of such bricks is 400 pounds, with a s.d. of 20 pounds A sample of 100 bricks yielded a mean of 395 pounds.

   (a) Set up the null hypothesis.

   (b) Suggest the alternative hypothesis.

   (c) Define the test statistic and test the hypothesis at 5% level of significance.

   (d) In case the s.d. is unknown what would be the test statistic ?

   (e) Test the hypothesis, when an estimate of variance estimated from the sample is found to be 196 units.

[40]

2.(a) A biologist has mixed a spray desired to kill 50% of certain type of insects. If a spraying of 200 such insects killed 120 of them, Would you conclude his mixture was satisfactory.

   (b) In one Section of city, 62 out of 450 tax payers were delinquent with these tax payments, whereas in another Section 40 out of 500 were delinquent. Test to see if the delinquency rate is the same for these two Sections of the city.

[20]

p.t.o.

3. Find t from the table such that

   (i) $P(Z>t) = 0.025$, (ii) $P(Z<t) = 0.025$
   (iii) $P(Z>t) = 0.05$, (iv) $P(Z<t) = 0.05$
   Hence, find t such that
   (a) $P(-t<Z<t) = .90$, (b) $P(-t<Z<t) = .95$,

   where Z is a Standard Normal deviate.

   [20]

4. Defie and explain the following :
   (a) Type I and Type II error
   (b) Critical Region
   (c) Level of significance
   (d) Null hypothesis and Alternative hypothesis.

   [20]

INDIAN STATISTICAL INSTITUTE
One Year Part-time Course in Statistical Methods and Applications : 1992-93
PART I
Grade 4 : Applied Statistics (Theory)
Final Examination

Date : 18.1.1993 Maximum Marks : 100 Time : 2 Hours

GROUP A and B Max.Marks : 70

Sample Survey (Index Number and Time Series)

Answer all the questions. The marks are indicated in the margin.

1.(a) Define stratified sampling. When would you recommend use of stratified sampling design in preference to simple random sampling design ? Explain with the help of some suitable illustration.

(b) A population is divided in K strata with N_i units in stratum i, $\sum_{i=1}^{K} N_i = N$. Random samples of size i are drawn from the stratum i, i=1,2..K and $\bar{x}_i$ is the sample mean of i^{th} stratum. Find the mean and variance of the estimator

$$\bar{x}_{st} = \sum_{i=1}^{K} (N_i/N)\bar{x}_i$$

[15]

2. Explain the linear and circular methods of systematic sampling. When does systematic sampling lead to more precision compared to simple random sampling ? Discuss.

[15]

3. What is meant by 'Seasonal Variation' in time series ? Illustrate your answer with suitable examples. Describe the 'ratio-to-trend method' for measurement of seasonal fluctuation in a time series. When would you recommend this method.

[20]

4. Write short notes on the following :
   (a) Multistage design
   (b) Laspayere's and Paasche's Indices are not measures of true Indices.

[20]

p.t.o.

GROUP C max.Marks : 30

Vital Statistics

1.(a) Define crude death rate. Why is the rate called 'crude'? [4]

(b) Define standardised death rate. In what way are the standardised rates superior ? [4]

(c) Calculate the crude and standardised death rates for the local population from the following data and compare them with crude death rate of the standard population.

Age group	Standard		Local	
	Population	deaths	Population	deaths
0 - 10	600	18	400	16
10 - 20	1000	5	1500	6
20 - 60	3000	24	2400	24
60 - 100	400	20	700	21

[7]

(d) Fill in the blanks in the life table given below :

Age x	l_x	d_x	q_x	L_x	T_x	e_x^0
4	95000	500	?	?	4850350	?
5	?	400	?	?	?	?

[5]

2. Define and explain any four of the following :
   (i) Crude birth rate
   (ii) General fertility rate
   (iii) Age-specific fertility rates
   (iv) Total fertility rate
   (v) Gross reproduction rate
   (vi) Net reproduction rate.

[10]

1992-93 | E142

INDIAN STATISTICAL INSTITUTE

One Year Part-time Course in Statistical Methods and Applications : 1992-93

Grade 4 : Applied Statistics (Theory and Practical)

Periodical Examination

Date : 29.10.1992 Maximum Marks : 100 Time : 2 Hours

Note : Use separate answerscripts for Group A and Group B.

GROUP A (Theory)

1. (a) Define and illustrate with examples the following
   (i) Population, (ii) Sampling Unit, (iii) Frame
   (iv) Census (v) Sampling.

   (b) Discuss merits and demerits of sample survey and compute enumeration. [20]

2. For the case of SRSWOR of size n from a finite population
   (a) Prove that sample mean is unbiased for population mean
   (b) Derive the variance of the sample mean
   (c) Obtain an unbiased estimator of the variance of the sample mean. [20]

3. What is stratification ? Discuss the situation when stratification is needed.
   Obtain an unbiased estimator for the population mean, when it is stratified into a number of strata. [10]

GROUP B (Practical)

1. Given below are the data on number of workers(x) and output (y) for 30 factories in a certain region :

x	y (in 000, Rs.)	x	y (in 000, Rs.)	x	y (in 000, Rs.)
51	135	152	512	481	682
68	339	60	328	544	707
87	372	57	290	100	389
53	266	335	632	65	342
65	342	93	376	51	118
107	397	452	685	750	761
70	340	71	342	198	556
253	586	144	509	443	660
125	438	211	563	166	523
352	651	528	730	85	360

p.t.o.

(a) Estimate the total output through :

(i) a SRSWOR of size 10.

(ii) a stratified random sample without replacement of size 10, stratification being done with respect to the no. of workers and with proportional allocation.

(b) Estimate also the standard error of your estimates. Give as far as possible the details of sample selection.

[50]

INDIAN STATISTICAL INSTITUTE

One Year Part-time Course in Statistical Methods and Applications : 1992-93

PART I

Grade 4 : Applied Statistics (Practical)
Final Examination

Date : 19.1.1993 Maximum Marks : 100 Time : 2 Hours

This paper contains 115 marks. Do as much as you can but the maximum attainable score is 100.

1. In order to estimate the total milk production per day in a city, all the 100 khatals in the city were stratified in 4 strata according to the size (no. of milch animals) of the khatal. Then from each of the stratum a SRSWOR of 5 khatals was drawn and the milk production in sampled khatals was noted for that particular day. The relevant data are given below :

Stratum Sl.No.	No. of khatals	Milk production (in 00 litres) in sampled khatal				
		1	2	3	4	5
1	25	2.0	2.5	1.8	2.6	3.0
2	22	6.5	7.6	7.0	8.8	9.0
3	38	8.0	12.0	11.5	16.0	13.5
4	15	15.0	14.0	20.0	25.0	18.0

Estimate the total milk production per day in the city and estimate the standard error of your estimate.
Using the sample variances for each stratum as obtained above re-allocate a sample of 20 khatals to 4 strata in an optimum manner.

[25]

2. It is desired to estimate the sample size needed from a population of size 430 using SRSWOR where by a pilot survey, the coefficient of variation was estimated as 0.5. It is also desired that a relative deviation of the estimates from the parameter of 'd' be allowed with a chance 'α'. Tabulate n for d=0.1, 0.2 and α = 0.05, 0.10.

[15]

p.t.o.

3. From the following data calculate Paasche's quantity index number for the year 1969 with 1951 as base.

Commodity	Quantity		Value
	1951	1969	1969
A	54	250	500
B	93	[illegible]	825
C	18	56	448
D	6	8	56
E	23	47	141

[15]

4. Obtain indices of seasonal variation from the following time series data on consumption of cold drinks, which contain only seasonal and irregular variation.

Quarter / Year	Consumption of cold drinks ('000 bottles)			
	I	II	III	IV
1971	90	90	87	70
1972	85	80	78	75
1973	80	75	75	72
1974	85	82	80	81

Remove the effect of seasonal variation from the original series and plot the series so obtained together with original series. Use a multiplicative model.

[45]

5. Practical record books and performance in class.

[15]

INDIAN STATISTICAL INSTITUTE
ONE YEAR PART TIME COURSE IN STATISTICAL METHODS
AND APPLICATIONS 1992-93
PART I
SUPPLEMENTARY EXAMINATION
GRADE 4: APPLIED STATISTICS (THEORY AND PRACTICAL)

Date: 10.3.93 Maximum Marks: 100 Time: 2 Hours

GROUP A

Maximum Marks: 40

Note: Attempt all questions.

1. In which of the following situations would you prefer a sample survey to a complete enumeration (with the reason)?
   (a) Determination of the average number of match sticks in a box and the proportion that actually light from a carton containing a gross of match boxes.
   (b) A survey of tuberculosis prevalence in an area (population about 50,000)
   (c) A longitudinal fertility survey in England
   (d) Preparation of electoral registers
   (e) A survey of electoral registers. [10]=[5x2]

2. Show that for stratified SRSWOR

$$V_{ran} = V_{prop} + \frac{(N-n)}{n(N-1)}\left[\Sigma N_h(\bar{Y}_h-\bar{Y})^2 - \frac{1}{N}\sum_h (N-N_h)S_h^2\right]$$

where, V_{ran}, V_{prop}, N_h, S_h^2, $\bar{Y}_h$ and $\bar{Y}$ have their usual meanings. [10]

3. Define linear systematic sampling. Show that sample mean defined under linear systematic sampling is unbiased for population mean $\bar{Y}$. Is it always better than simple random sampling? If not, state the condition under which linear systematic sampling will be better than simple random sampling. [5+5+10]=[20]

contd. ...2/-

GROUP B

Maximum Marks: 40

1. Following data give the production of coal in India for a number of years. Fit a quadratic trend to the data:

Year	1959	1960	1961	1962	1963	1964	1965
Production (y): (in million metric tons)	48	52	56	61	66	62	67

Obtain trend values for all the years and plot them along with the original series. [20]

2. Given the following data, obtain all the index nos. commonly used for the year 1934 with 1927 as base year:

	Rice		Wheat		Jowar	
Year	Price	Quantity	Price	Quantity	Price	Quantity
1927	9.3	100	6.4	11	5.1	5
1934	4.5	90	3.7	40	2.7	3

[10]

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.

How would you distribute optimally, a sample of 1000 households between the two strata? Compare your allocation with proportional allocation. [10]

contd. 3/-

GROUP C

Maximum Marks:20

1.(a) Define crude death rate. Why is the rate called 'crude'? [2]

(b) Define standardised death rate. In what way are the standardised rates superior? [4]

(c) Distinguish between complete and abridged life tables. Derive life table mortality rate q_x in terms of age specific death rate m_x. [4]

2.(a) Describe the different measures of fertility. [5]

(b) Define and explain the following [5]

(i) Gross reproduction rate

(ii) Net reproduction rate.

1992-93 E251

INDIAN STATISTICAL INSTITUTE

ONE YEAR PART-TIME COURSE IN STATISTICAL METHODS AND APPLICATIONS- 1992-93

PART II

FINAL EXAMINATION

Grade:5 Estimation, Linear Estimation and Design of Experiments (Practical)

Date: 17.6.93 Maximum Marks:100 Time: 2 Hours

Note: Answer any THREE

1. For a factorial experiment with 3 factors A,B and C each at two levels, the design and yield per plot are given below.

	Block 1	Block 2	Block 3
(1)	32	44	24
bc	24	36	20
ac	32	34	32
ab	30	30	36
a	30	46	28
c	32	39	32
abc	36	42	30
b	27	32	26

Analysing the data, test for the main - effects and interactions. [33]

2. Consider the following fixed effects linear model for six observations:

$$\underset{6\times1}{\underline{Y}} = \underset{6\times4}{X}\ \underset{4\times1}{\underline{\theta}} + \underset{6\times1}{\underline{e}}$$

where $E(\underline{e}) = \underline{0}$, $D(\underline{e}) = \sigma^2 I_6$. Further, let

$$X'X = \begin{bmatrix} 3 & -1 & -1 & -1 \\ -1 & 3 & -1 & -1 \\ -1 & -1 & 3 & -1 \\ -1 & -1 & -1 & 3 \end{bmatrix}$$

and $\underline{Y}'X = (1 \quad 4 \quad -3 \quad -2)$.

(a) Show that $\sum_{i=1}^{4} l_i \theta_i$ is estimable if and only if $\sum_{i=1}^{4} l_i = 0$.

p.t.o.

2. (b) Using any non-estimable function as side restriction, solve the normal equations and obtain a set of LSE of θ_i, i=1,2,3,4. Hence find the BLUE of $\theta_1+\theta_2+2\theta_3-4\theta_4$. [33]

3. Suppose the observations on X and Y are given as:

X	Y
59	75
65	70
45	55
52	65
60	60
62	69
70	80
55	65
45	59
49	61

where n = 10 students, and Y = Marks in Maths, X = Marks in Econ. Compute the least squares regression of Y on X (i.e., fit the model $y_i = a+bx_i+e_i$). Furthermore, test for H_0: a=b=0. [33]

4. A field experiment was conducted using a CRD. The details of the experiment are given below.
No. of Treatments : 3
The digestibility co-efficients recorded in the experiment

Treatments		
T_1	T_2	T_3
38.83	36.44	37.53
29.16	34.66	41.75
29.54	44.82	50.77
35.76		30.47

Analyse the data and draw your conclusions. [33]

<u>One Mark is kept for Neatness</u>.

1992-93	E251

INDIAN STATISTICAL INSTITUTE

One Year Part-time Course in Statistical Methods and Applications : 1992-93

PART II

FINAL EXAMINATION

Grade 5: Estimation, Linear Estimation and Design of Experiments (Theory)

Date: 15.6.93 Maximum Marks: 100 Time: 2 hours

GROUP-A

Maximum Marks: 30

Note: Answer any TWO questions. Marks allotted to each question are given within brackets.

1. Discuss, with examples, the notion of minimum - variance unbiased estimation. (15)

2.(a) Find the maximum likelihood estimator of θ based on a random sample $x_1, x_2, \ldots, x_n$ from a population having density function

$$f(x) = \exp[-(x-\theta)], \quad \theta \leq x < \infty.$$

(b) Let $x_1, x_2, \ldots, x_n$ be a random sample from $N(\mu, \sigma^2)$. Find the Rao-Cramer lower bound of the variance for an unbiased estimator of σ^2. $(7\frac{1}{2} \times 2 = 15)$

3. Write short notes on any two of the following

(a) Interval estimation;

(b) Method of moments; and

(c) "Sufficiency" property of an estimator. $(7\frac{1}{2} \times 2 = 15)$

GROUP B

Maximum Marks: 70

Note: Answer Question No.1 and any TWO out of the rest. Marks allotted are indicated in brackets [ ] at the end of the question.

1.(a) Define estimable functions and give two necessary and sufficient conditions for estimability known to you.

(b) State clearly the Gauss-Markoff theorem and explain its importance.

p.t.o.

(c) Consider the following linear model of rank 2 for five observations $y_1, y_2, \ldots, y_5$ involving 3 fixed effects $\theta_1, \theta_2, \theta_3$ and random errors e_i's with mean 0 and variance σ^2:-

$y_1 = \theta_1 - \theta_2 + \theta_3 + e_1$; $y_2 = \theta_1 + 2\theta_3 + e_2$; $y_3 = \theta_2 + \theta_3 + e_3$;

$y_4 = 2\theta_1 - \theta_2 + 3\theta_3 + e_4$; $y_5 = \theta_2 + \theta_2 + e_5$.

(i) Obtain a necessary and sufficient condition for $C_1\theta_1 + C_2\theta_2 + C_3\theta_3$ to be estimable and get its B.L.U.E.

(ii) Obtain an unbiased estimator for σ^2 in terms of $\{y_i\}$.

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

2. Develope the analysis of variance of two-way classified data with m (m > 1) observations per cell. [15]

3. Define a randomised block design with an example. When do you recommend the use of such designs. Describe its analysis of variance. [3+2+10=15]

4.(a) Define the main effects and interactions of a 2^2 factorial experiment, and show that they represent 4-1=3 mutually orthogonal treatment contrasts.

(b) Give the analysis of variance of a 2^2 experiment in r randomised blocks, explaining clearly the method of Yates' Sum and Difference to compute various sums of squares. [5+10=15]

1992-93 | E261

INDIAN STATISTICAL INSTITUTE

One Year Part-time Course in Statistical Methods and Applications : 1992-93

PART II

FINAL EXAMINATION

Grade 6 : Programming (Theory)

Date : 21.6.1993 Maximum Marks : 100 Time : 2 Hours

The questions carry total 110 marks. You can score maximum 100. Marks allotted for each question are shown in [ ]. Answer as much as you can.

1. Write short-notes on the following :

   (a) Secondary storage, (b) C.P.U., (c) Memory

   (d) File, record and field, (e) Operating system. [4x5=20]

2. What is a binary number ? Explain its importance in digital computer. [5]

3. Describe the following MSDOS commands (with examples) :

   (a) COPY (b) FORMAT (c) DELETE (d) DIR (e) MKDIR [2x5=10]

4. Describe the following BASIC statement/functions (with examples):

   (a) WHILE-WEND (b) ON GOSUB (c) OPEN and CLOSE

   (d) MID$ (e) INSTR. [4x5=20]

5. What is a flow-chart ? Draw a flow-chart for FOR-NEXT statement. [5]

6. Let A be an array of N numbers. Given a number F you have to search it in the array A and print the corresponding index. Print index value as 0(zero) if F is not found in the array. Draw a flow-chart and write a BASIC program. [10]

7. Two ascending sequence of M and N entries are stored in two arrays A and B respectively. Use another array C of size M+N to merge A and B into one ascending sequence and output it. Write a BASIC program. [20]

8. A sequential file contains time spent for study (T) and total marks obtained in examination (M) for a group of stndents, Write a BASIC program to compute correlation coefficient (ρ) between T and M. [20]

1992-93 E261

INDIAN STATISTICAL INSTITUTE

One Year Part-time Course in Statistical Methods and Applications : 1992-93

PART II

FINAL EXAMINATION

Grade 6 : Programming (Practical)

Date : 18.6.1993 Maximum Marks : 100 Time : 2 Hours

1. Create a sequential file containing the following information.

Salesman Code	Memo Number	Value of Sales (Rs.)
01	X001	800.00
01	X002	300.00
01	X003	50.50
01	X004	100.00
02	Y102	500.00
02	Y103	67.00
02	Y104	400.00
03	M501	900.00
04	A005	1000.00
04	A006	2000.00
04	A007	500.00

[20]

2. Write a program in BASIC to print the total value of sales, commission and bonus (the computation rule is stated below) of each salesman in the following format, using the file created in question No.1.

Sales Statement for the month .... Year ....

Salesman Code	Memo No.	Amount (Rs.)	Commission (Rs.)	Bonus (Rs.)
..	....	....		
	....	....		
	....	....		
	Total	....	....	....
..	....	....		
	....	....		
	....	....		
	Total	....	....	....

Computation rule for calculating Commission and Bonus :

The commission and bonus of a salesman is calculated on his total value of sales. The commission is calculated as 8% of his total value of sales. A salesman fails to get the bonus if his total value of sales is <Rs. 1000. He gets 2% of total value of sales as bonus if it is >=Rs.1000 and <Rs.3000 otherwise he gets 3% of total value of sales as bonus.

[40]

3. Assignments. [40]

1992-93 E261

INDIAN STATISTICAL INSTITUTE
ONE YEAR PART-TIME COURSE IN STATISTICAL METHODS
AND APPLICATIONS: 1992-93
PART II
FINAL EXAMINATION
GRADE 6: FUNDAMENTALS OF SAMPLING THEORY (PRACTICAL)

Date: 18.6.93 Maximum Marks: 100 Time: 2 Hours

Note: This paper carries 115 marks. Attempt all the questions but the maximum attainable score is 100.

1. Given below are the data on number of workers (x) and value added (y) (in million's of Rs.) by 30 medium sized factories in a certain region. Draw a PPSWR sample of size 5 with probabilities proportional to the no. of workers (x) and estimate the average value added per factory and the relative standard error of your estimate:

Sl. No.	x	y	Sl. No.	x	y	Sl. No.	x	y
1	95	47.1	13	57	30.2	25	63	41.2
2	79	45.2	14	132	70.7	26	83	51.7
3	30	15.6	15	47	31.2	27	124	70.2
4	45	23.2	16	43	11.2	28	31	16.9
5	28	11.7	17	116	61.7	29	96	51.3
6	142	98.6	18	65	21.9	30	60	15.5
7	125	51.7	19	103	15.7			
8	81	42.2	20	52	38.4			
9	43	27.3	21	67	32.2			
10	53	31.8	22	22	31.3			
11	148	67.2	23	75	36.5			
12	89	50.9	24	69	38.9			

[30]

2. Stratify the population of factories given in Problem (1) into two strata with respect to auxiliary variable (x). Allocate a sample size 10 to the strata optimally and draw samples through SRSWOR independently in each stratum.

Obtain the ratio and regression estimates (combined or separate type) of total value added with the estimates of their relative standard errors. [40]

3. A sampler proposes to take a stratified randam sample. He expects that his field costs will be of the form $\Sigma\, C_h n_h$. His advance estimates of relevant quantities for the two strata are as given below:

Stratum	W_h	S_h	C_h
1	0.4	10	\$ 4
2	0.6	20	\$ 9

p.t.o.

3. (a) Find the values of n_1/n and n_2/n that minimize the total field cost for a given value of $V(\bar{y}st)$.

(b) Find the sample size required under this allocation to make $V(\bar{y}st) = 1$. Ignore the fpc.

(c) How much will the total field cost be ? [18]

4. In a study of the possible use of sampling to cut down the work in taking inventary in a stock room a count is made of the value of the articles on each of the 36 shelves in the room. The estimate of total value made from a sample selected according to SRSWOR is to be correct within 200 ฿ apart from a 1 in 20 chance. Estimate the sample size. Given $\sum_{i=1}^{36} y_i = 218$,

$\sum_{i=1}^{36} y_i^2 = 131,682$. [12]

5. Practical record books. [15]

1992-93 E261

INDIAN STATISTICAL INSTITUTE
ONE YEAR PART-TIME COURSE IN STATISTICAL METHODS
AND APPLICATIONS: 1992-93
PART II
FINAL EXAMINATION
GRADE:6 FUNDAMENTALS OF SAMPLING THEORY (THEORY)

Date: 21.6.93 Maximum Marks: 100 Time: 2 Hours.

1. Under simple random sampling, show that
   (a) sample proportion $p = a/n$ is an unbiased estimator of the population proportion $P = A/N$, a, A, N have usual meanings.
   (b) An unbiased estimate of the variance of p is

$$v(p) = s_p^2 = \frac{(N-n)}{(n-1)N} \cdot pq.$$ [10]

2. Under what considerations and how the size of a sample is estimated?

   In the population of 676 petition sheets, how a large must the sample be if the total number of signatures is to be estimated with a margin of error 1000, apart from a 1 in 20 chance? Given that $S^2 = \Sigma (y_i - \bar{Y})^2/(N-1) = 229.0$. [10]+[5]=15

3. Define regression method of estimation in estimating population total. Discuss the relative performance of ratio and regression estimators. [20]

4. A simple random sample of n-clusters each containing M elements is drawn from the N-clusters in the population. Obtain the unbiased estimator for the population mean.

   Show that $V(\bar{y}) \simeq \frac{1-f}{nM} S^2 [1+(M-1)\rho]$

   f = sampling fraction, $S^2 = \Sigma(y_i - \bar{Y})^2/(N-1)$,

   ρ = Intracluster Correlation. [20]

5. Define inclusing probabilities of different orders. Show that Horvitz-Thompson estimator is unbiased for population total.

   Obtain the estimator of the variance of the estimator (Yates-Grundy Expression). [5]+[5]+[10]=[20]

p.t.o.

6. Discuss the following:

(i) Use of auxiliary information

(ii) Systematic sampling (linear)

(iii) On the choice of the triput $(P, \hat{Y}, V(\hat{Y}))$ [15]

1992-93 E271

INDIAN STATISTICAL INSTITUTE
ONE YEAR PART-TIME COURSE IN STATISTICAL METHODS AND APPLICATIONS: 1992-93
PART II
FINAL EXAMINATION

Grade 7: Survey Sampling and Organisation Aspects (Pract.)

Date: 25.6.93 Maximum Marks: 100 Time: 2 Hours.

1. Out of 24 villages in an area, two linear systematic samples of 4 villages each were selected. The total area under wheat in each of these sample villages is given in the following Table. Estimate the total area under wheat in the area with the standard error

Area under wheat of the two linear systematic samples of 4 villages each from 24 villages

Linear Systematic Sample	Sample village 1	2	3	4
1	427	326	481	445
2	335	412	503	348

[15]

2. In an area there are 349 schools with a total of 27215 students. Eight schools were first selected with probability proportional to the number of students and in each selected school, 50 students were selected at random.

The following Table 2 gives for each school, the number of students with trachoma and the number with multiple scars. Estimate (a) the proportion of students with trachoma and (b) among those with trachoma the proportion with multiple scars with the corresponding standard errors of the estimators.

Table 2. Number of students with trachoma and number with multiple scars among samples of 50 students each selected at random from each of the eight schools selected with probability proportional to the total number of students

School No. i	Number of students with trachoma: Total number	Number with multiple scars
1	40	3
2	31	0
3	47	16
4	41	8
5	43	8
6	36	5
7	39	2
8	48	13
Total	325	55
Row Sum of Squares	13421	591
Row sum of Products	2426	

[20]

p.t.o.

3. The following Table 1 gives the analysis of variance of a samp of $n = 40$ from $N = 1000$ f.s.u.'s and $m_0 = 5$ sample s.s.u.'s in each selected fsu($M_0 = 50$), sampling at both stages being sim random

(a) Estimate unbiasedly the variance of the mean and its components.

(b) Given the total Cost of a survey C = Rs.6000, the Cost of including an f.s.u. in the sample C_1 = Rs.50 and the Cost of getting the data from an s.s.u. within each selected f.s.u., c_2 = Rs.20, obtain the optimum sample numbers of f.s.u.'s and s.s.u.'s within selected f.s.u..

Table 1. Analysis of variance

Source of variation	Degrees of freedom	Mean Square
Between f.s.u.'s	39	634
Between s.s.u.'s within f.s.u.'s	160	409

[20]+[20]=[40]

4. Practical Note Book: [25]

INDIAN STATISTICAL INSTITUTE
ONE YEAR PART-TIME COURSE IN STATISTICAL METHODS AND APPLICATIONS: 1992-93
PART II
FINAL EXAMINATION
GRADE: 7 NUMERICAL METHODS (PRACTICAL)

Date: 25.6.93 Maximum Marks: 100 Time: 2 Hours

Note: This paper carries 135 marks. Attempt all the questions but the maximum attainable score is 100.

1.(a) Verify that Weddle's formula gives the more accurate value than Simpson's and Trapezoidal formulae by evaluating $\int_{-3}^{3} x^4 \, d_x$ with seven equidistant points and finding the relative error in each of the cases.

(b) Evaluate $\sum_{m=0}^{\infty} \frac{1}{(10+m)^2}$ by Euler - Maclaurin's formula.

[20+15]=35

2. Using suitable approximation formula

(a) Find the missing term in the following table:

x	0	1	2	3	4
y	1	3	9	-	81

(b) Find the tenth term of the following series:

0, 0, 2, 6, 12, 20

(c) Find the function whose first difference is $9x^2+11x+5$.

[(7+6+7)=20]

3. Find the inverse of the following matrix by Crout's method. Also find the value of its determinant.

$$\begin{bmatrix} 12 & 20 & 30 \\ 25 & 15 & 21 \\ 30 & 28 & 18 \end{bmatrix}$$

[25]

p.t.o.

4.(a) By the method of successive approximation find the value of x for $f(x) = 16.00$, given $f(0) = 16.35$, $f(5) = 14.88$, $f(10) = 13.59$, and $f(15) = 12.46$.

(b) Evaluate the following binary expression:

1100101 - 1111001 + 101101 x 1111 [(15+5)=20]

5.(a) Find by Newton-Raphson method a real root of the polynomial $x^3 + 1.2x^2 - 4x - 4.8 = 0$.

(b) Find an iterative formula to evaluate $\sqrt{29}$ and evaluate its value correct to four decimal places. [(13+7)=20

6. Practical record Books and performance in class. [15]

INDIAN STATISTICAL INSTITUTE

One Year Part-time Course in Statistical Methods and Applications : 1992-93

PART II
FINAL EXAMINATION

Grade:7 : Survey Sampling and Organisation Aspects (Theory)

Date : 23.6.1993 Maximum Marks : 100 Time : 2 Hours

1.(a) Discuss the reasons for multi-stage sampling.

(b) Out of the N f.s.u's, n are selected and out of the M_i s.s.u's in the i^{th} selected first-stage unit $(i = 1,2,..,n)$, m_i are selected. Sampling at the first stage is ppswr, sizes being the number of s.s.u's and at the second stage is SRSWR. Let y_{ij} be the value of the study variable in the j^{th} selected s.s.u. $(j=1,2,..,m_i)$ in the i^{th} selected fsu $(i=1,2,..,n)$.

Obtain the estimate of the term Y with the standard error of the estimator.

[20]

2. Consider a two stage design with $n(\geq 2)$ of s.s.u's being selected with replacement out of the N total f.s.u's with varying probabilities p_i, say, and a fixed number $m_0(\geq 2)$ of s.s.u's being selected out of the M_i total s.s.u's in the i^{th} selected f.s.u with equal probabilities. Taking a simple cost function

$$C = nc_1 + nm_0c_2.$$

Determine the optimum sizes of m_0 and n.

[20]

3. Obtain a ratio type estimator with the standard errors for population total Y of a variable y in a stratified multistage sampling. Consider the case of (i) Combined and (ii) Separate estimators.

When the two estimators likely to behave equally ? Multistage means two stage and consider simple random sampling at both the stages.

$[12\frac{1}{2} + 12\frac{1}{2} + 5] = [30]$

4. Discuss the following :

(a) Sampling errors and Sampling biases

(b) Errors common to both censuses and samples

(c) Internal consistency checks

(d) Control of Non-Sampling errors.

[5+10+5+10] = [30]

INDIAN STATISTICAL INSTITUTE
ONE YEAR PART-TIME COURSE IN STATISTICAL METHODS
AND APPLICATION: 1992-93
PART II
FINAL EXAMINATION
GRADE: 7 NUMERICAL METHODS (THEORY)

Date: 23.6.93 Maximum Marks: 100 Time: 2 Hours

Note: Answer any four questions. Each question carries 25 marks.

1.(a) ~~Describe~~ Derive Newton's general formula of interpolation.

(b) For equispaced arguments derive the relation between simple difference and divided difference of r^{th} order. Hence prove that Newton's forward formula is a special case of Newton's general formula. [9+16=25]

2.(a) Describe with proof the Birge - Vieta method to find a root of a polynomial function.

(b) Prove that Newton - Raphson method is quadratically convergent when $f'(x) \neq 0$ and is linearly convergent when $f'(x) = 0$. [15+10=25]

3.(a) Describe Cholesky's method to find inverse of a matrix. When can you apply Cholesky's method and why? How will you solve a system of linear equations by this method?

(b) What is the condition on 'a' and 'b' so that the following matrix is singular?

$$\begin{pmatrix} 2 & 1 & a \\ 1 & 2 & 2 \\ 2 & b & 1 \end{pmatrix}$$

Is it always possible to make the matrix singular by manipulating 'b' when 'a' is given? [19+6=25]

4.(a) Assuming $f(x)$ to be of (i) first, (ii) second, (iii) third degree polynomial in x, express

$$\int_0^6 f(x)\,dx$$

in terms of $f(0), f(1), \ldots, f(6)$.

(b) Find x_1 and x_2 such that the following approximate formula becomes exact to the highest possible order:

$$\int_{-1}^{1} f(x)\,dx \simeq \frac{1}{3}\left\{f(-1) + 2f(x_1) + 3f(x_2)\right\}$$ [20+5=25]

p.t.o.

5.(a) $2x^3+4x^2-2x-5 = 0$ has a root near $x = 1$. Find at least two rearrangements (for iteration) that will converge to this root beginning with $x_0 = 1$.

(b) Find the following sum:

$$\sum_{x=1}^{25} (x^5-11x^4+41x^3-61x^2+31x-21)$$

(c) Find the polynomial with lowest degree so that

$f(0)=24$, $f(1)=0$, $f(2)=0$, $f(3)=0$, $f(4)=0$ and $f(5)=144$.

[10+10+5=25]

6. Find the sum of each of the following series:

(a) $$\frac{1}{x+1} - \frac{n}{(x+1)(x+2)} + \frac{n(n-1)}{(x+1)(x+2)(x+3)} - \cdots + (-1)^n \frac{n!}{(x+1)(x+2)\ldots(x+n+1)}.$$

[Hint: use $\Delta^r x^{(-1)}$ for $r = 1,2,\ldots, n$]

(b) $x^3+ (x+1)^3+ (x+2)^3+\ldots+ (x+n-1)^3$

(c) $$\frac{1}{a+1} + \frac{2!}{(a+1)(a+2)} + \frac{3!}{(a+1)(a+2)(a+3)} + \ldots + \frac{n!}{(a+1)(a+2)\ldots(a+n)}$$

[Hint: $\Delta(x+a) \binom{x+a}{a}^{-1} = \binom{x+a}{a}^{-1} (1-a)$ when differenciating w.r.t. a]

[8+9+8=25]
