1985-86	E263

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Applications : 1985-86

FINAL EXAMINATION

PART - II

Industrial Statistics (Theory and Practical)

Date: 16.6.1986 Maximum Marks: 100 Time: 2 hrs.

Note: Answer any <u>three</u> questions. All questions carry equal marks.

1.(a) Define acceptance rectification scheme for a lot by lot sampling inspection plan. Define average outgoing quality. Derive the necessary expressions for Average outgoing quality limit for a single sampling plan under this scheme.

(b) Draw the AOQ curve for the sampling plan for lots of size 1000, sample size 100, and acceptance number 2. Using poisson approximation.

From the curve find out approximately the value of the process average at which AOQ takes its maximum value i.e., AOQL and also give the AOQL value.

Now use the Dodge and Romig's tables for arriving at the exact values for the above two parameters.

(18+15) = [33]

2.(a) Find a Dodge and Romig's single sampling fraction defective sampling inspection plan that has an AOQL = .01 and for lot size 7500, minimises the total inspection when the process average is .0075. What quality product has a 0.10 chance of lot acceptance ? Calculate the value of ATI at process average .0075.

(b) Given LTPD = 5% find the Dodge and Romig's double sampling plan that minimises the ATI for lots of size 11,000 at a process average of 2%. What is the AOQL of this plan ?

Now find out the probability of acceptance at process average 2% of the plan.

Contd..... 2/-

Contd... Q.No.2

(c) Give the expression of ASN for a double sampling plan with parameters (n_1, n_2, c_1, c_2) at process average p. Calculate the ASN for the double sampling plan with $n_1 = 50$, $n_2 = 100$, $c_1 = 2$, $c_2 = 4$ at $p = .02$ and at $p = .10$.

(7+12+14) = [33]

3.(a) Define the type A and type B OC curve. Give the expression for both OC functions for a single sampling plan. Also give the approximate formulas (wherever applicable) for calculating the probability of acceptance under type A and type B situation.

(b) Define AQL, LTPD, producer's risk and consumer's risk.

(c) Under poisson approximation the np value which gives 95% probability of acceptance (denoted as $np_{.95}$) and which gives 10% probability of acceptance (denoted as $np_{0.10}$) are given as under for different C values.

C	$np_{.95}$	$np_{.10}$
1	.355	3.89
2	.818	5.32
3	1.366	6.68
4	1.970	7.99

Given an AQL of 1% and LTPD and of 6.5% how do we obtain a single sampling plan for a large lot which gives α = and $\beta = 0.10$. If the LTPD is changed to 4% what would be the change in n and C.

(9+8+16) = [33]

4.(a) Define process capability. Given the upper and lower specification limits of a product as 10 and 30 units where do we set the process if the process cabability is

(i) 15 units and working at lower level is more economic.

(ii) 25 units.

(b) Subgroups of 5 items each are taken from a manufacturing process at regular intervals. A certain quality characteristic is measured and $\overline{X}$ and R values are computed for

Contd..... Q.No.4.(b)

each subgroup. After 25 subgroups $\Sigma \bar{X}$ = 357.50 and ΣR = 8.80. Compute the control limits. All points on both the charts fall within these limits. If specification limits are 14.40 ± 0.40 what conclusions can you draw about the ability of the existing process to produce items within specifications ? What percentage of products, if any, would you expect to find outside the specification limits if the specification limits are tightened to 14.40 ± .35.

(c) A p chart is used to analyse the September recerd for 100% inspection of certain radio transmitting tubes. The total number inspected during the month and the total number of defective was 2196 and 158 respectively. Compute $\bar{p}$, compute individual 3 sigma limits for the following 3 days and state whether the fraction defective fell within control limits for each day.

Date	Number of inspected	Number of defectives
September 14	54	8
15	162	24
16	213	3

(10+13+10) = [33]

:bcc:

1985-86	E263

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Applications: 1985-86

FINAL EXAMINATION

Design of Experiments (Theory and Practical)

Date: 12.6.1986 Maximum Marks: 100 Time: 2 hrs.

Note: Attempt Question No.1 and any TWO from the rest. Marks alloted to a question are indicated in brackets [ ] at the end.

1. An engineer is interested in the effect of cutting speed (A), tool geometry (B), and cutting angle (C) on the life of a machine tool. Two levels of each factor are chosen, and three replicates of a 2^3 factorial design are run. The results are given below. Analyse the data and draw conclusions.

Treatment combination	Replicate I	Replicate II	Replicate III
(1)	22	31	25
a	32	43	29
b	35	34	50
ab	55	47	46
c	44	45	38
ac	40	37	36
bc	60	50	54
abc	39	41	47

[30]

2. Consider the following linear model (Ω):

$$y_1 = \theta_1 + \theta_4 + \theta_5 + e_1 ,$$
$$y_2 = \theta_2 - \theta_4 + \theta_5 + e_2 ,$$
$$y_3 = \theta_3 - 2\theta_5 + e_3 ,$$
$$y_4 = \theta_1 + \theta_2 + 2\theta_5 + e_4 ,$$
$$y_5 = \theta_1 + \theta_3 + \theta_4 - \theta_5 + e_5$$

Contd..... 2/-

Contd..... Q.No.2

Show that $\sum_{i=1}^{5} \lambda_i \theta_i$ is estimable if and only if $\lambda_4 = \lambda_1 - \lambda_2$ and $\lambda_5 = \lambda_1 + \lambda_2 - 2\lambda_3$. Get a least square estimator for $\{\theta_i\}$ under suitably chosen side restrictions. Obtain the B.L.U.E. of $\theta_1 + \theta_2 + \theta_3$ in terms of $\{y_i\}$. Give also an unbiased estimator of σ^2 in terms of $\{y_i\}$.

(10+10+7+8) = [35]

3. Define a randomised block design with an example, and postulate a suitable linear model for its analysis. Identify the estimable functions under the model postulated, and develop the analysis of variance for such designs.

(5+5+10+15) = [35]

4. Define the main effects and interactions of a 2^2 factorial, and show that they represent 3 mutually orthogonal treatment contrasts. Describe Yates' method of sums and differences for computing sum of squares due to the various factorial effects. Indicate briefly the analysis of variance for the data from a 2^n factorial.

(10+10+15) = [35]

:bcc:

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Applications : 1985-86

FINAL EXAMINATION

Part - II

Sample Surveys (Practical)

Date: 10.6.1986 Maximum Marks: 100 Time: 2 hrs.

Note: Attempt as many questions as you can. The paper carries 120 Marks but the maximum you can score is 100 marks. The marks allotted are given at the end of each question.

1. A rural zone consisting of 50 villages is divided into three strata having 20, 12 and 18 villages respectively. Samples, using SRSWOR, were selected independenty from the above three strata and the number of households and the population of selected villages were observed. The data yielded:

Stratum	No. of selected villages (n_h)	No. of households (x_h)	Population (y_h)
1	4	60, 72, 56, 48	600,360,250,600
2	2	55, 45	300,450
3	4	90,120,100,50	900,1000,800,800

(i) Give unbiased estimates for average no. of households and average population per village in the rural zone.

(ii) Estimate the size of household in the rural zone.

(iii) Estimate unbiasedly by the percentage of villages, in the rural zone, with no. of households more than 50.

(iv) Estimate the standard error of the estimate in (iii) above.

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

p.t.o.

2. For a survey on 'Cigarette-smoking' among college going male students, a sample of 5 colleges was drawn, out of 50 colleges under a University, using SRSWOR. Each college has 15 Departments. Samples of 3 Departments were drawn, using SRSWOR, from each of the selected college and no. of students who smoke Cigarattes was found out. The investigation yielded.

Serial No. of Selected college	No. of smokers among male students in the selected departments		
1	50,	150,	60
2	90,	40,	120
3	40,	30,	50
4	120,	90,	60
5	80,	40,	50

(i) Give unbiased estimates (a) for the total no. of smokers (among the male students) in the 50 colleges under consideration; (b) for the average no. of smokers per college; (c) for the average no. of smokers per Department.

(ii) Estimate the standard errors of the above estimates.

(15+15) = [30]

3. The following table gives the total cultivated area (x) and area under what (y) in acres for 10 villages in a Block of Kanpur District (rural) of U.P.

Village No.	1	2	3	4	5	6	7	8	9	10
x	400	650	1000	1500	800	1200	2000	1200	800	1500
y	75	200	300	1000	600	900	1500	700	500	1200

(i) Select 3 villages, using Lahiri's method, with probabilities proportional to total cultivated area and with replacement and give unbiased estimate for the average area under wheat.

Contd..... 3/-

Contd..... Q.No.3

(ii) Give estimate for the standard error of the above estimate.

(10+15) = [25]

4. (i) Select a sample of 5 villages, from the 10 villages in question no.3 above, using SRSWOR.

(ii) Give ratio and regression estimates for the average area under wheat.

(iii) Estimate unbiasedly the percentage of villages having area under wheat as 1000 acres or more.

(iv) Estimate the ratio of average area under wheat to the average area cultivated.

(v) Estimate the standard errors of the ratio and regression estimates in (ii) above.

(3+15+4+3+10) = [35]

:bcc:

1985-86	E253

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Aplications : 1985-86

FINAL EXAMINATION

PART - II

Sample Surveys (Theory)

Date: 6.6.1986 Maximum Marks: 100 Time: 2 hrs.

Note: Attempt as many questions as you can. The paper carries 120 marks. However, the maximum you can score is 100 marks. The marks allotted are specified at the end of each question.

1. (i) What do you mean by stratified sampling ?

(ii) Mention the main problems which arise in stratification.

(iii) Define unbiased estimators for population total Y, mean $\bar{Y}$, and proportion P based on simple stratified sampling.

Derive the expression for variance of any one of the above estimators.

(6+6+13) = [25]

2. (i) What do you mean by PPS sampling ?

(ii) Describe the cumulative size method and Lahiri's method of selecting a PPSWR sample.

(iii) Give an unbiased estimator for population mean based on PPSWR sampling and derive the expression for its variance.

(6+9+10) = [25]

3. (i) Define ratio, regression, difference and product estimators for population mean $\bar{Y}$ when the population mean $\bar{X}$ of an auxiliary character x is known

(ii) Consider the sampling strategies

$T_1 \equiv$ (SRSWOR, sample mean);

$T_2 \equiv$ (SRSWOR, ratio estimator) and

p.t.o.

Contd..... Q.No.3.(ii)

T_3 ■ (SRSWOR, product estimator) for population mean $\bar{Y}$, let ρ be the correlation coefficient between y and x and C_y the coefficient of variation of y. Let the sample size be large.

Mention for each of the following situations, separately, one of the above strategies you would suggest to use giving reasons for the same.

(a) $0 < \rho \leq 0.8$; $0 < C_y \leq 2$; $C_x \geq 4$

(b) $\rho \geq 0.4$; $C_y \geq 1$; $0 < C_x \leq 0.7$

(c) $\rho \leq -0.3$; $0 < C_y \leq 2$; $C_x \geq 1$

(d) $\rho \geq -0.8$; $C_y \geq 1.4$; $C_x \leq 2$

(iii) Give three sampling strategies for estimation of the population ratio $R = Y/X$, where Y and X denote the population totals of characters y and x respectively.

(8+10+7) = [25]

4. (i) What do you mean by two stage sampling ?

(ii) A rural District is divided into N zones consisting of M villages each. Give a suitable sampling procedure and estimators to estimate

(a) Total milk production in the District.

(b) Average milk production for a zone in the District.

(c) Average milk production per village in the District.

(iii) Derive the expression for variance of one of the estimators in (ii) above.

(6+6+13) = [25]

5. (i) Outline the principal steps to be taken for conducting a large scale sample survey.

(ii) Name the main types of non-sampling errors and indicate how do they arise.

(10+10) = [20]

:bcc:
5586

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Applications : 1985-86

FINAL EXAMINATION

Statistical Inference (Practical)

Date: 4.6.1986 Maximum Marks: 50 Time: 2 hours

Note: Answer as many questions as you can.
Maximum that you can score is 50.
α represents the level of significance.

1. A coin is tossed 18 times and the number of times a head has appeared is 12. Let p represent the probability of occurrence of a head. At $\alpha = 0.05$, will you reject the null hypothesis if $H_0: p = \frac{1}{2}$ and $H_1: p > \frac{1}{2}$. [7]

2. When the first proof of a book containing 250 pages was read the following distribution of misprints was found:

No. of misprints per page	Frequency
0	139
1	76
2	28
3	4
4	2
5	1
≥ 6	0
Total	250

Test the goodness of fit of poisson distribution. [12]

3. Let $Y_1, Y_2, \ldots, Y_n$ be a random sample from $N(\theta,1)$. Construct power curve for $H_0: \theta = 3$ Vs $H_1: \theta < 3$ for $\alpha = 0.05$. [8]

4. Electroencephalograms are records showing fluctuations of electrical activity in the brain. Among the several kinds of brain waves produced, the dominant ones are usually alpha waves.

Contd... 2/-

Contd..... Q.No.4

An experiment was conducted to see whether sensory deprivation over an extended period of time has any effect on the alpha wave pattern. The subjects were 20 inmates in a Canadian prison. They were randomly split into two equal sized groups. Members of one group were placed in solitary confinement. Those in the other group were allowed to remain in their own cells. Seven days later, alpha wave frequencies were measured for all 20 subjects.

Non confined (x_i)	Solitary confinement (y_i)
10.7	9.6
10.7	10.4
10.4	9.7
10.9	10.3
10.5	9.2
10.3	9.3
9.6	9.9
11.1	9.5
11.2	9.0
10.4	10.9

Assume the distribution of X_i's to be normal, Y_i's to be normal and they are independent. Will you be able to test the equality of means of the above two normal distributions ? If so, find cut at $\alpha = 0.05$. Whether the means can be considered to be same. [12]

5. Nine workers were tested to determine the effects of certain exposures on various respiratory functions. One such function, air flow rate is measured by computing the ratio of a person's forced respiratory volume (FEV_1) to his vital capacity (VC). In persons with no lung dysfunction the value of FEV_1/VC is 0.80. It can be assumed that if the exposure does have an effect, it will be to reduce FEV_1/VC. The data are listed below.

Contd... 3/-

Contd..... Q.No.5

Subject	FEV_1/VC
1	0.78
2	0.84
3	0.83
4	0.61
5	0.70
6	0.63
7	0.76
8	0.82
9	0.67

Do the appropriate hypothesis test. Set $\alpha = 0.05$.

[6]

6. Staffs from 15 different colleges participated in a surveillance program to monitor the number of patients experiencing adverse reactions to prescribed medication. The percentages for the fifteen hospitals were 5.8, 5.3, 4.5, 3.9, 4.6, 5.4, 7.9, 8.2, 6.9, 5.7, 4.6, 6.3, 8.4, 4.6 and 7.3. Let μ denote the true average percentage represented by these figures. Construct a 95% confidence interval for μ.

[5]

7. Class assignments. [5]

:bcc:

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods
and Applications : 1985-86

FINAL EXAMINATION

PART - II

Statistical Inference (Theory)

Date: 2.6.1986 Maximum Marks: 100 Time: 2 hours

Note: Solve as much as you can. Maximum possible score is 100.

1.(a) Let $X_1, X_2, \ldots, X_n$ be a random sample from the pdf $f(x,\theta)$. Define (i) an unbiased estimator for θ, (ii) a sufficient statistic for θ.

(b) Let T be an unbiased estimator for the parameter θ. Is T^2 unbiased for θ^2. Justify.

(c) An estimator $T_n = r_n(X_1, X_2, \ldots, X_n)$ for the parameter θ is said to be asymptotically unbiased if $\lim_{n \to \infty} E(T_n) = \theta$.

If $X_1, X_2, \ldots, X_n$ is a random sample from the uniform $(0,\theta)$ distribution then show that $X_{(n)}$, the nth order statistic, is not unbiased for θ. Is it asymptotically unbiased ?

(d) Let $X_1, X_2, \ldots, X_n$ be a random sample from the pdf $f(x,\theta)$ given by $f(x,\theta) = e^{A(x)\alpha(\theta) + B(x) + \beta(\theta)}$.

Find a sufficient statistic for θ.

(5+4+6+8) = [23]

p.t.o

2.(a) Let $X_1, X_2, \ldots, X_n$ be a random sample from the pdf $f(x,\theta)$. Define

(i) the Cramer-Rao Lower Bound (CRLB) for estimating the parameter θ.

(ii) a consistent estimator for θ.

(b) Let $X_1, X_2, \ldots, X_n$ be a random sample from the uniform $(0,\theta)$ distribution. Show that $X_{(n)}$, the nth order statistic, is consistent for θ.

(c) Show that the estimator $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, based on a random sample $X_1, X_2, \ldots, X_n$ from Poisson distribution, is an efficient estimator for the occurrence rate λ, where

$$f(x,\lambda) = \frac{e^{-\lambda}\lambda^x}{x!}\ ;\ x = 0, 1, 2, \ldots .$$

(d) Let $X_1, X_2, \ldots, X_n$ be a random sample from

$$f(x,\theta) = \frac{1}{\theta^2}\, x\, e^{-x/\theta}\ ,\ 0 < x < \infty\ ,\ 0 < \theta < \infty$$

Find the MLE for θ.

(9+9+9+9) = [36]

3.(a) Based on a random sample $X_1, X_2, \ldots, X_n$ from the pdf $f(x,\theta);\ \theta \in \Omega$, suppose we are interested in testing the hypothesis

$$H_0:\ \theta \in \Omega_0 \text{ against } H_1:\ \theta \in \Omega_1$$

then define (i) a randomised test, its size and power (ii) a uniformly most powerful test.

(b) A random sample of size 2 is drawn from a uniform $(0,\theta)$ pdf. We wish to test $H_0:\ \theta = 2$ against $H_1:\ \theta < 2$ by rejecting H_0 when $x_1 + x_2 \leq k$. Find the value of k for which this test of ours will have size .05.

Contd... 3/

Contd... Q.No.3

(c) Let X_1, X_2 be a random sample of size 2 from the pdf

$$f(x,\theta) = \theta\, x^{\theta - 1} \qquad 0 < x < 1$$

Obtain the best critical region of size .1 for testing H_o: $\theta = 1$ against H_1: $\theta = 2$.

(9+10+8) = [27]

4.(a) Explain the principle of the generalised likelihood ratio test (GLRT).

(b) Let X_1, X_2 ..., X_n be a random sample from the $N(\mu, \sigma^2)$ distribution. Obtain the GLRT for testing H_o: $\mu = \mu_o$ against H_1: $\mu \neq \mu_o$, σ^2 unknown. Set α, the size, equal to .05.

(c) Explain the goodness of fit λ^2-test to test the hypothesis that a random sample has come from Poisson pdf with the occurrence rate parameter $\lambda = 6.2$.

(5+10+8) = [23]

:bcc:

1985-86	E263

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Applications : 1985-86

PERIODICAL EXAMINATION

PART - II

Industrial Statistics (Theory and Practical)

Date: 7.5.1986 Maximum Marks: 100 Time: 2 hrs.

Note: Answer Question No.1 and any TWO from the rest.

1. The following data pertains to the daily sample results of observations of certain dimension of Bomb Bases. A sample of size 5 is taken daily from the production process. The means and ranges of 15 samples are as under. The specification limits are given as 0.840 and 0.820.

Over all Heights of Fragmantation Bomb Bases

Day	$\bar{X}$	R
1	0.8324	0.014
2	0.8306	0.008
3	0.8262	0.020
4	0.8326	0.004
5	0.8290	0.013
6	0.8316	0.013
7	0.8336	0.012
8	0.8310	0.020
9	0.8323	0.024
10	0.8336	0.010
11	0.8332	0.018
12	0.8288	0.006
13	0.8310	0.016
14	0.8294	0.023
15	0.8322	0.003

Set up trial control limits for a $\bar{X}$-R chart for future use from the above data. Estimate process capability and compute the minimum unavoidable proportion defectives.

[40]

p.t.o.

2. The maximum temperatures reached by a given heating process show an average of 113.30° Cent.egrade and a standard deviation of 5.6° C. Assume that these maximum temperature variations are random and normally distributed, determine

   (a) What per cent of the maximum temperatures are less than 116.10° ?

   (b) What value is exceeded by 57.78 per cent of the temperature readings ?

   (c) What limits include the middle 50 per cent of the temperature readings ?

   (d) What per cent of maximum temperatures lie between 114.98° and 116.66° ?

[30]

3. The number of defects per group in 20 successive groups of 5 radio sets each were as follows. Use an appropriate control chart to analyse the data w.r.t. the stability of the defect level.

1	77	11	57
2	64	12	40
3	75	13	22
4	93	14	92
5	45	15	89
6	61	16	55
7	49	17	25
8	65	18	54
9	45	19	22
10	77	20	49

[30]

4. Consider setting up a p chart with reference to production of railway-car side frames. Suppose that these frames are in continuous production in a foundry and that a random sam of 50 frames is taken from each day's output and inspected. The result of 20 days of past operations were as follows:

Date	Number rejected	Date	Number rejected
March 27	4	April 6	34
28	9	7	25
29	10	8	18
30	11	9	12
31	13	10	4
April 1	30	11	3
2	26	12	11
3	13	13	8
4	8	14	14
5	23	15	21

Set up a control chart for controlling the fraction defect of railway car side frames.

[30]

1985-86	E262

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Applications : 1985-86

PERIODICAL EXAMINATION

PART - II

Design of Experiments (Theory and Practical)

Date: 30.4.1986 Maximum Marks: 100 Time: 2 hrs.

Note: Answer Question No.1 and any TWO from the rest. Details of marks alloted to a question are indicated in brackets [ ] at the end.

1. An experiment was run to determine whether four specific firing temperatures affect the density of a certain type of brick. The experiment led to the following data:

Temperature	Density				
100	21.8	21.9	21.7	21.6	21.7
125	21.7	21.4	21.5	21.4	
150	21.9	21.8	21.8	21.6	21.5
175	21.9	21.7	21.8	21.4	

Analyse the data and draw conclusions.

[30]

2. Explain the meanings of the terms:

(a) plot, (b) treatment, (c) experimental error
Illustrate with examples the three basic principles of experimental designs. Define a latin square design and give an example showing its use.

(3x4) + (3x6) + 5 = [35]

3. Describe a general linear model (say, Ω) for fixed effects. Give the mathematical and statistical definitions of estimable functions and show that they are equivalent. Define a least square estimator (LSE) for the parameter vector $\underline{\beta}$ under Ω, give the normal equations for $\hat{\underline{\beta}}$, and show that these equations always form a consistent system.

5 + 10 + (5+7+8) = [35]

p.t.o.

4.(a) State the Gauss - Markoff theorem, and give an unbiased estimator for σ^2, the common error variance under Ω.

(b) Let y_1, y_2, y_3, y_4 and y_5 be five observations for which we assume the following linear model:

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

$y_3 = \theta_2 + \theta_3 + e_3$; $\quad y_4 = 2\theta_1 - \theta_2 + 3\theta_3 + e_4$;

$y_5 = \theta_2 + \theta_3 + e_5$;

where θ_1, θ_2, θ_3 are fixed but unknown effects, and e_i's are random errors assumed to be i.i.d. with mean zero and variance σ^2. Show that $\lambda_1 \theta_1 + \lambda_2 \theta_2 + \lambda_3 \theta_3$ is estimable if and only if $2\lambda_1 + \lambda_2 = \lambda_3$. Write down the normal equations for $\hat{\theta}_1$, $\hat{\theta}_2$ and $\hat{\theta}_3$, and obtain a LSE for them subject to suitably chosen side restrictions. Give the BLUE of an estimable function in terms of the observations.

(5+3) + 15 + (7+5) = [35]

:bcc:

1985-86 | E252

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Applications : 1985-86

PERIODICAL EXAMINATION

PART - II

Sample Surveys (Theory and Practical)

Date: 23.4.1986 Maximum Marks: 100 Time: 2 hrs

Note: Do as many questions as you can. The paper carries 110 marks but maximum score can not exceed 100 marks. Marks allotted are given at the end of each question.

1. (i) What do you mean by sampling and non-sampling errors ?

(ii) Name the main types of non-sampling errors and indicate how they creep into sample surveys or census.

(iii) Give a sampling strategy to estimate the population mean unbiasedly incase of presence of non-response.

(iv) Out of 50 factories in a town 10 were selected randomly and letters were sent to them to find the no. of staff working on daily basis. The responses were (from six only): 200, 100, 50, 100, 150, 250. A sub-sample of 2 factories from remaining 4 (the non-responding factories) was selected randomly and efforts yielded the responses as 300 and 500 workers.

Estimate the average no. of workers working on daily basis in the above 50 factories.

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

2. (i) There are N universities in India. The no. of affliated colleges to these universities are M_1, ..., M_N respectively. Give a suitable sampling procedure and estimates for (a) total no. of students in the above N universities, coming from the Low Income Group (LIG).

(b) Average no. of students from LIG for each university.

(c) Average no. of students from LIG for a college.

LIG may be defined as the families which have income of Rs.200/- or less per month per family member.

Contd..... 2/-

Contd..... Q.No.2

(ii) A Block consisting of 50 villages is divided suitably into 10 clusters each consisting of 5 villages. A sample of 2 clusters was selected using SRSWOR. Then from each selected cluster, samples of 3 villages were selected using SRSWOR and no. of Cattles recorded. The investigation yielded.

Selected clusters	No. of Cattles for selected villages		
1	500,	1000,	300
2	700,	250,	600

(a) What is the total no. of Cattles in the block ?

(b) Estimate the average no. of Cattles per village and per cluster in the block.

(c) Estimate the standard errors of the estimates in (a) and (b) above.

(10+20) = [30]

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

(ii) A Paddy producing zone, in a rural Block, consists of 6 villages. In a particular year each village had 5 plots of paddy-crop. The following data gives yield of paddy (in '000 kgs) for various plots in the above six villages :

Village no.	1	2	3	4	5	6
Yield of Paddy in various plots	45	60	10	20	22	20
	15	50	20	15	28	18
	10	35	25	10	15	16
	60	40	15	20	18	32
	25	20	25	35	30	40

(a) Estimate the total yield and average yield per plot taking a sample of 2 villages using SRSWOR.

(b) Estimate the standard error of estimates.

(5+20) = [25]

4\. (i) Describe the procedures of linear systematic sampling and circular systematic sampling and give unbiased estimates for population mean.

(ii) Refer to Question no.3.(ii) above. Suppose the population consists of the 30 plots.

(a) Draw a linear systematic sample of 5 plots.

(b) Draw a circular systematic sample of 7 plots.

(c) Estimate the total paddy-yield of the 30 plots basing on the samples in (a) and (b) above.

(10+15) = [25]

:bcc:

1985-86 | E242

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Applications : 1985-86

PERIODICAL EXAMINATION

PART - II

Statistical Inference (Tehory and Practical)

Date: 9.4.1986 Maximum Marks: 100 Time: 2 hrs.

Note: Solve as much as you can. The maximum possible score is 100.

1.(a) What do you understand by 'testing of hypotheses' ? Give an example of a situation where you have to test a hypothesis.

(b) Let $X_1, X_2, \ldots, X_n$ be a random sample from $f(x, \theta)$; $\theta \epsilon \Omega$. Consider the problem of testing

$$H_o : \theta \epsilon \ _o \text{ against } H_1 : \theta \epsilon \Omega_1.$$

Define a nonrandomised test, its size and its power function. Explain the 'errors' a decision rule might lead to in testing the above hypotheses. Can one minimise the probabilities of these errors simultaneously? If not what is the standard method of controlling these probabilities.

(3+3+6+4+2+2) = [20]

2.(a) Let $X_1, X_2, \ldots, X_n$ be a random sample from $f(x, \theta)$ where $\theta \epsilon \{\theta_o, \theta_1\}$. Consider the problem of testing $H_o : \theta = \theta_o$ against $H_1 : \theta = \theta_1$. State and prove the Neyman-Pearson lemma.

(b) Let $X_1, X_2, \ldots, X_n$ be a random sample from $N(\theta, \sigma^2)$. If $\sigma^2 = 1$ then test $H_o : \theta = 2$ against $H_1 : \theta = 4$. Set α, the probability of type I error = .05.

(c) If σ^2 were unknown in the above problem could you still have used Neyman-Pearson lemma to test the same hypothesis? Explain.

(3+8+6+3) = [20]

p.t.o.

3.(a) A coin, with probability of head p, is flipped five times independently and the outcomes are noted. Find out the most powerful critical region for testing $H_o : p = \frac{1}{3}$ against $H_1 : p = \frac{2}{3}$. Set $\alpha = 11/3^5$.

(b) For the same value of α also find out the most powerful critical region for testing $H_o : p = \frac{2}{3}$ against $H_1 : p = \frac{1}{3}$.

(c) Let Z be a standard normal variate and Y have t-distribution with 9 degrees of freedom. Find the values z_1, z_2 and y_1, y_2 such that

Prob $(|Z| > z_1) = 0.01$; Prob $(Z < z_2) = 0.1$

Prob $(|Y-2| > y_2) = 0.05$ and Prob $(Y-4 > y_2-7) = 0.05$.

(6+6+8) = [20]

4.(a) Explain in detail the principle of generalised likelihood ratio test (GLRT).

(b) Let $X_1, X_2, \ldots, X_n$ be a random sample from $N(\mu, \sigma^2)$. Obtain the GLRT for testing $H_o: \sigma^2 = \sigma_o^2$ against $H_1 : \sigma^2 \neq \sigma_o^2$; μ unknown. Set $\alpha = .05$.

(c) Let $X_1, X_2, \ldots X_n$ be a random sample from $N(\mu_1, \sigma^2)$ and $Y_1, Y_2, \ldots, Y_m$ be another random sample, independent of the first, from $N(\mu_2, \sigma^2)$. Obtain the GLRT for testing $H_o : \mu_1 = \mu_2$ against $H_1 : \mu_1 \neq \mu_2$, σ^2 unknown. Set $\alpha = .05$.

(8+10+12) = [30]

5.(a) Define (i) monotone likelihood ratio property, (ii) uniformly most powerful (UMP) test.

(b) Give an UMP test for a family of distributions

$f(x, \theta)$; $\theta \in \Omega$, Ω an interval, that possesses a monotone likelihood ratio property, for testing $H_o: \theta \leq \theta_o$ against $H_1: \theta > \theta_o$.

(c) What is a contingency table. Explain the χ^2-test for testin the independence of two characters in a two-way contingency tables.

(4+4+2+6) = [16]

:bcc:

1985-86 E251

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Applications: 1985-86

PERIODICAL EXAMINATION

PART - II

Sample Surveys (Theory and Practical)

Date: 2.4.1986 Maximum Marks: 100 Time: 2 hrs.

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

1. (i) What do you mean by stratified sampling ?

(ii) Let SSS; simple stratified sampling. Consider the three sampling strategies for population mean $\bar{Y}$ as

T_1 = (SRSWOR, sample mean)

T_2 = (SSS, usual unbiased estimator $\bar{y}st)_P$

T_3 = $(SSS, \bar{y}st)_N$

where T_2 and T_3 are based on proportional and Neyman allocations respectively. Then if the ferms of order $1/N$ are negligible, then show that

$$V(T_3) \leq V(T_2) \leq V(T_1)$$

(iii) A district consisting of 12 Tehsils is divided suitably into three strata. The variate y observed is no. of villages having no electricity. A stratified sample of over all 5 tehsils is to be selected. Suppose the information available is as follows:

Stratum No.	No. of Tehsils N_h	$\bar{Y}_h$	S_h
1	5	6	10
2	3	13	21
3	4	22	24

(a) Obtain sample sizes for different strata under proportional allocation and Neyman allocation.

Contd..... 2/-

Contd..... Q.No.1

(b) Obtain standard error for the unbiased estimator of Y, the total no. of villages in the district having no electricity.

(c) Obtain standard error of the usual unbiased estimator of Y based on SRSWOR of 5 tehsils from the 12 Tehsils in the district.

(d) Obtain relative efficiency of the strategy T_3 over T_1 and T_2 defined in (ii) above.

(5+10+20) = [35]

2. (i) There are N villages in a Block. It is proposed to estimate $\bar{Y}$ the average no. of house holds per village in the Block based on a sample of n villages. The average population of a village in the block, $\bar{X}$, is known from the previous census records.

Define sampling strategies for $\bar{Y}$, based on ratio, regression and difference estimators.

(ii) Derive the conditions under which one sampling strategy is preferable over the other among the strategies:

T_1 ≡ (SRSWOR, sample mean);

T_2 ≡ (SRSWOR, ratio estimator);

T_3 ≡ (SRSWOR, product estimator);

T_4 ≡ (SRSWOR, regression estimator);

T_5 ≡ (SRSWOR, difference estimator)

You may assume that sample size is large.

(iii) A simple random sample of 5 factories from 50 factories yielded the sales (x) and profits)y) as (In Rs Lakhs) (x,y); (120,30); (50,10); (40,10); (100,20); (80,20). The average sale per factory is known to be Rs.80 lakhs. Estimate the average profit per factory using sample mean and ratio, regression estimators.

(5+15+15) = [35]

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

Contd..... Q.No.3

(ii) Following table gives the number of persons (x) and male labour force (y) in six blocks of a small town:

Block No.	x	y
1	1650	600
2	400	100
3	750	300
4	1000	350
5	500	150
6	2500	1200

(a) Draw ppswr samples of 2 Blocks using (i) cumulative size method, (ii) Lahiries method.

(b) Estimate the total male labour force in the town based on the samples in (a) above.

(c) Estimate the standard error of the estimates in (b) above.

(5+20) = [25]

4. (i) Give a sampling strategy to estimate P, the percentage of land holdings below 1.0 hectare in a Block of 50 villages.

(ii) Give a sampling strategy to estimate R, the ratio of total male working force to female working force in a Block of 40 villages.

(iii) A simple random sample of 5 small hotels/restaurants was selected and total no. of workers (x) and no. of child-workers were found to be

Hotel No.	1	2	3	4	5
x	10	20	10	15	5
y	2	5	3	4	3

What is the percentage of child-workers ?

(5+5+5) = [15]

:bcc:

1985-86	E241

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Applications: 1985-86

PERIODICAL EXAMINATION

PART - II

Statistical Inference (Theory and Practical)

Date: 19.3.1986 Maximum Marks: 100 Time: 2 hrs.

Note: Solve as much as you can. The maximum possible score is 100.

1.(a) What is meant by inductive logic and deductive logic ? Give examples.

(b) Define a random sample. Give a real life example.

(c) Describe briefly the estimation set up that you are familiar with. Define an estimator. When would you call it unbiased ?

(d) Let $X_1, X_2, \ldots, X_n$ be a random sample from $N(\mu, \sigma^2)$. Show that $\sum_{i=1}^{n} (X_i - \overline{X})^2/(n-1)$ is unbiased for σ^2, obtain its variance, where $n\overline{X} = \sum_{i=1}^{n} X_i$.

(6+4+6+8) = [24]

2.(a) Let $X_1, X_2, \ldots, X_n$ be a random sample from $f(x, \theta)$, $\theta \in \Omega$. Define sufficient statistic.

(b) Let $X_1, X_2, \ldots, X_n$ be a random sample from a Pareto distribution,

$$f(x, \theta) = \theta/(1+x)^{\theta+1}, \quad 0 < x < \infty, \quad 0 < \theta < \infty.$$

Deduce a sufficient statistic for θ.

(c) Let X_1, X_2, X_3 be three Bernoulli trials with parameter p. Let $T = X_1 + 2X_2 + 3X_3$. Show that T is not sufficient for p.

(d) Let $\{T_n, n \geq 1\}$ be a sequence of estimators for θ. When do you call it consistent.

Contd..... 2/-

Contd..... Q.No.2

(e) A sequence $\{ T_n, n \geq 1 \}$ of estimators for θ is said to be squared-error consistent if $\lim_{n \to \infty} E(T_n - \theta)^2 = 0 \quad \forall \theta \in$
Show that if T_n is squared-error consistent then it is consistent as well. (Hint: For any r.v. X and a constant $a > 0$, $P[X^2 < a] > 1 - \frac{E(X^2)}{a}$).

(f) Let $X_1, X_2, \ldots, X_n$ be a random sample from $N(0, \sigma^2)$.
Show that $\frac{1}{n} \sum_{i=1}^{n} X_i^2$ is consistent for σ^2.

(3+7+6+2+7+5) = [30]

3.(a) What do you understand by Cramer-Rao lower bound ?

(b) Find the Cramer-Rao lower bound for estimating the parameter θ in a Cauchy distribution,

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

(c) Define efficient estimator and a best estimator. What is the difference between them ?

(3+12+6) = [21

4.(a) Let $X_1, X_2, \ldots, X_n$ be a random sample from $f(x,\theta)$. Defin likelihood function. What is the rationale behind the maximum likelihood estimator ?

(b) A commuter's trip home consists of first riding a subway to a bus stop and then taking a bus home. The bus she would like to catch arrives uniformly over the interval (θ_1, θ_2). She would like estimates for both θ_1 and θ_2 so she has some idea of when she should be at the stop (θ_1) and when she is probably too late and will have to wait for the next bus (θ_2). Over an eight-day period, she makes certain to be at the stop early enough to reco the arrival times of the bus; 5:15, 5:21, 5:14, 5:23, 5:29, 5:17, 5:15 and 5:18. She consults you, being a statistician, for estimating θ_1 and θ_2 by maximum likeli- hood method of estimation. How would you carry out your job. Also obtain maximum likelihood estimator for the mean of the arrival distribution.

Contd..... 3/-

Contd.... Q.No.4

(c) Use the method of moments to estimate θ in the one-parameter beta distribution,

$$f(x,\theta) = \frac{\Gamma(\theta+1)}{\Gamma\theta\,\Gamma 1} x^{\theta-1} \quad ; \; 0 < x < 1 \; ; \; \theta > 0$$

(d) Let $X_1, X_2, \ldots, X_n$ be a random sample from the uniform distribution defined over $(0, \theta)$. Find the formula for a 90% confidence interval for θ based on the statistic $T = (n+1)X_{(1)}$, where $X_{(1)}$ is the smallest order statistic.

(5+10+5+8) = [28]

5.(a) Let $X_1, X_2 \ldots, X_m$ be a random sample from $N(0,1)$ and $Y_1, Y_2, \ldots, Y_n$ be another independent random sample also from $N(0,1)$. Let $U = \sum_{i=1}^{m} X_i^2$, $V = \sum_{i=1}^{n} Y_i^2$. Define $Z = \dfrac{U/m}{V/n}$. Obtain the probability density function of Z.

(b) If Z is distributed as $F(m,n)$ then $1/Z$ is distributed as $F(n,m)$. Prove.

(7+7) = [14]

:bcc:

1985-86 | E111S

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Applications : 1985-86

SUPPLEMENTARY EXAMINATION

Probability

Date: 26.2.1986 Maximum Marks: 100 Time: 2 hrs.

Note: The paper is set for 107 marks. You may attempt any part of any question. The maximum you can score is 100. The marks alloted for each question are given in brackets [ ].

1.(a) When two random variables X and Y are said to be independent? Let X and Y be two independent random variables each geometrically distributed with parameter p. Find the distribution of X+Y.

(b) Let X be uniformly distributed on 0, 1, ..., 49. Calculate $P(2.6 < X < 12.2)$.

(c) A die is rolled until a 6 appears. How many rolls are required so that the probability of getting 6 is at least 1/2 ?

(2+6+5+5) = [18]

2.(a) Let X and Y be two random variables having finite expectations. Show that

(i) if c is a constant and $P(X = c) = 1$, then $EX = c$.

(ii) if c is a constant, then cX has finite expectation and $E(cX) = cEX$.

(b) Give an example to show that the expectation of a product of two random variables is not the product of their expectations.

(c) Let X be a random variable having a finite second moment. Find the value of λ that minimizes $E(X-\lambda)^2$.

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

p.t.o.

3.(a) Let X be a random variable such that for some constant M, $P(|X| \leq M) = 1$. Prove that X has finite expectation and $|EX| \leq M$.

(b) Suppose X and Y are two independent random variables such that $EX^4 = 2$, $EY^2 = 1$, $EX^2 = 1$, and $EY = 0$. Compute $\mathrm{Var}(X^2 Y)$.

(c) Let N be a positive integer and let f be the function defined by

$$f(x) = \begin{cases} \dfrac{2x}{N(N+1)}, & x = 1, 2, \ldots, N. \\ 0, & \text{elsewhere} \end{cases}$$

Show that f is a discrete density and find its mean.

(5+5+5) = [15

4.(a) Let $g(x) = x(1-x)$, $0 \leq x \leq 1$, and $g(x) = 0$ elsewhere. Normalize g to make it a density.

(b) When a random variable is said to be a symmetric random variable ? Let X be a random variable that has a density. Show that f has a symmetric density if and only if X is a symmetric random variable.

(c) Let X be a random variable having the normal density $n(0, \sigma^2)$. Find the density of the random variable $Y = X^2$.

(d) Let X have the normal density $n(\mu, \sigma^2)$. Find $a > 0$ such that $P(\mu - a \leq X \leq \mu + a) = .9$. It is given that $\Phi(1.645) = .95$.

(5+2+5+5+5) = [2

5.(a) Let X and Y be independent random variables such that X has the gamma density $\Gamma(\alpha_1, \lambda)$ and Y has the gamma densit $\Gamma(\alpha_2, \lambda)$. Show that $X+Y$ has the gamma density $\Gamma(\alpha_1 + \alpha_2,$

(b) Let $f(x, y) = c(y-x)^\alpha$, $0 \leq x < y \leq 1$ and $f(x, y) = 0$ elsewhere. For what values of α can c be chosen to make f a density function ?

(10+5) = [

6. Define expectation of a continuous random variable X having a density f. Let X be uniformly distributed on (a , b). Find EX.

(2+5) = [7]

7. State (true or false except in (iv) and (vi) where you give the statement only) :

(i) Poisson distribution can arise as a limit of binomial distribution.

(ii) The random variable X has finite expectation if and only if $E\,|X| < \infty$.

(iii) $Cov\,(X , Y) = 0$ whenever X and Y are independent random variables.

(iv) Give the statement of Chebyshev's inequality.

(v) For a distribution function $F(.)$ of a random variable $F(x+) = F(x)$ for all x .

(vi) State Central limit theorem.

(vii) The conclusion of Central limit theorem holds even if no moments of X_1 exist beyond the second.

(viii) For exponentially distributed random variable X, $P(X > a+b) \neq P(X > a)\,P(X > b)$, $a \geq 0$ and $b \geq 0$.

(ix) The correlation coefficient ρ is always between -1 and 1.

(x) The expectation of a sum of two random variables is the sum of their expectations.

[10]

:bcc:

1985-86 | E121S

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Applications: 1985-86

SUPPLEMENTARY EXAMINATION

Descriptive Statistics (Theory and Practical)

Date: 26.2.1986 Maximum Marks: 100 Time: 2 hrs.

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

1.(a) Discuss the effect of the changes of location and scale on the following measures:

Mean, Median, Standard Deviation and Mean Deviation

(b) Show that :

(i) the mean squared deviation is minimum when measured from the mean and

(ii) the mean absolute deviation is minimum when measured from the median.

(10+5+5) = [20]

2. Answer any three questions from the following:

(a) Show that

$$(\Sigma fi)(\Sigma fi(x_i - \bar{x})^2) = \sum_i \Big(\sum_{j>i} f_i f_j (x_i - x_j)^2 \Big)$$

(b) Show that the measure of skewness given by

$$\frac{3(\bar{x} - M_e)}{s}$$

must lie between -3 and 3.

(c) Show that the combined variance of two groups of observation can be calculated from the relation

$$\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2} + \frac{n_1 n_2}{(n_1 + n_2)^2} (\bar{x}_1 - \bar{x}_2)^2$$

(d) The life time of a particular model of a stereo cartridge is normally distributed with mean $\mu = 1000$ hours and a standard deviation $\sigma = 100$ hours. Find the probability that one of these cartridges will last :

Contd..... 2/-

Contd..... Q.No.2.(d)

(i) between 700 and 1200 hours,
(ii) 930 hours or less,
(iii) between 750 and 950 hours,
(iv) More than 870 hours.

(e) In a certain factory turning out razor blades, there is a small chance $\frac{1}{500}$ for any blade to be defective. The blades are supplied in packets of 10. Use the poisson distribution to calculate the appropriate no. of packets containing no defective, one defective and two defective blades respectively in a consignment of 10,000 packets.

$[\exp(-0.02) = 0.9802]$

(10 x 3) = [30]

3. The following constants were obtained from measurements on Serum Cholesterol in mg/100 c.c. (Y), Weight in Kg (X_1) and Systolic Blood Pressure (X_2) of 11 normal males between the ages of 14 and 24 years:

$\Sigma X_1 = 639$	$\Sigma X_2 = 1306$	$\Sigma Y = 1821$
$\Sigma X_1^2 = 37184$	$\Sigma X_2^2 = 155410$	$\Sigma Y^2 = 302557$
$\Sigma YX_1 = 105650$	$\Sigma YX_2 = 216624$	$\Sigma X_1X_2 = 75900$

(a) Find the regression of Y on X_1 and X_2 and hence find the coefficient of determination of Y on X_1 and X_2.

(b) Find the coefficient of determination of Y on X_1 and X_2 using the simple correlation coefficients.

(c) Obtain all the partial correlation coefficients.

(d) Find also the expected amount of serum cholesterol in a person of weight 60 kg. and systolic blood pressure 125.

(25+10+12+3) = [50]

:bcc:

1985-86	E134

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Applications : 1985-86

FINAL EXAMINATION

PART - I

Economics Statistics (c and d)

Date: 18.12.1985 Maximum Marks: 100 Time: 2 hrs.

Note: Answer Q.No.4 and any TWO out of the rest. Marks allotted to different questions are shown in brackets [ ].

1. The following data relate to the quarterly revenue expenditures (in lakhs of rupees) of the Govt. of India, during the years 1954-55 to 1957-58 :

year \ quarter	I	II	III	IV
1954-55	3867	4404	5726	9816
1955-56	4669	5327	5811	10350
1956-57	4693	5640	5957	12961
1957-58	5518	6887	7782	19718

A quadratic trend curve was fitted to the annual figures for the above data and the estimated trend equation is :

$$T_t = 27728.8 + 2837.2\,t + 389.9\,t^2$$

where the unit of t is two quaters and the origin of t is at the end of the fourth quarter of 1955-56.

(a) Compute trend values for all quarters of the years, 1954-55 to 1957-58.

(b) Compute seasonal indices by the ratio-to-trend method.

(16+17) = [33]

2. (a) The following table gives the data on the wholesale price index numbers of cereal and manufactures (with base 1970-71 = 100) for each of the years 1971-1984 :

Contd..... 2/-

Contd..... Q.No.2.(a)

Year	Price Index of cereals (P_c)	Price index of manufactures (P_m)
1971	100.0	110.9
1972	111.4	119.0
1973	128.8	133.8
1974	178.1	163.1
1975	185.7	173.1
1976	151.6	171.7
1977	161.2	179.4
1978	158.1	177.9
1979	167.0	203.9
1980	189.5	249.9
1981	213.1	269.5
1982	229.4	269.6
1983	259.2	287.9
1984	246.9	313.3

The relative price of cereals is defined to be the ratio of P_c to P_m.

(i) Draw a diagram showing the movements in the relative price of cereals over the years 1971-84.

(ii) Which components of a time series are present in the diagram in (i) ? Comment on the behaviour of these components.

(b) What is a "modified exponential" trend curve ? Describe briefly a method of estimating the parameters of this curve.

(21+12) = [33]

3.(a) Fit an exponential trend to the following data on the annual Australian exports of merchandise over the years 1965-66 to 1971-72 :

Year	1965-66	1966-67	1967-68	1968-69	1969-70	1970-71	1971-72
Exports (in crores of Australian dollars)	263	293	294	324	399	424	477

Contd..... 3/-

Contd..... Q.No.3

(b) Critically examine the method of moving average as a method of computing secular trend values.

(22+11) = [33]

4. Write short notes on <u>any two</u> of the following items :

(a) Some of the major points of enquiry in the 1971 population census;

(b) employment statistics in India;

(c) government statistical organisations (at the centre) in India.

(17+17) = [34]

:bcc:

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Applications: 1985-86

FINAL EXAMINATION

PART - I

Economics Statistics (a and b)

Date: 16.12.1985 Maximum Marks: 100 Time: 2 hrs.

Note: This question paper carries 125 marks. You can answer any part of any question. But the maximum that you can score is 100.

1. What is 'homogeneity error' ? [6]

2. Describe factor reversal test and F-error. (6+4) = [10]

3. The values of cost of living index in Sweden from 1945 to 1952 with 1948 = 100 were as follows:

   93, 93, 96, 100, 102, 103, 119 and 128.

   Obtain the values of this index by shifting the base to 1951. [10]

4. 'Weighing is only one of many problems in constructing index numbers'. Discuss this statement with special reference to index numbers you are familiar with. [13]

5. The table below relates to the weekly pay (before tax and other deductions) of the manual wage earners on a company's payroll.

Category of employees	1950		1960	
	Number	Total pay (£)	Number	Total pay (£)
Men aged 21 and over	350	2500	300	4200
Women aged 21 and over	400	1600	1200	8000
Youth and boys	150	450	100	560
Girls	100	250	400	1540

Construct an index of weekly earnings for 1960 based on 1950, showing the rise in earnings for all employees in one figure. [12]

p.t.o.

6. The following table gives figures of imports of board from Finland by the United Kingdom. Construct a suitable quantity index to show the change in the quantity of board entering the United Kingdom from Finland in 1957 as compared to 1956.

Type	1956 Quantity (000 cwt)	1956 Value(£ 000)	1957 Quantity (000 cwt)	1957 Value(£ 000)
Machine glazed board	90	296	143	549
Folding bore board	71	262	5	20
Kraft board	184	567	225	649
Wood pulp board	94	254	81	227
Other board	39	108	104	264

[13]

7. Distinguish between

(a) gross domestic product and gross national product,

(b) personal income and personal disposable income.

(8 x 2) = [16]

8. The average weekly wages and the consumer price indexes for all manufacturing industries in the U.S.A. for the seven months from November 1955 to May 1956 were as follows:

Average weekly wages ($)	79.52	79.71	78.58	78.17	78.78	78.99	78.0[illegible]
Consumer price index	115.0	114.7	114.6	114.6	114.7	114.9	115.4

Use the figures for consumer price index to deflate the wages and then find the percentage change in real wages over the seven months.

(7+8) = [15]

9. Calculate gross national product at market prices, the net national product at factor cost, net total capital formation and personal disposable income for the Indian economy for the year 1973-74 from the following figures (in crores of rupees).

Item	1973-74
Imports of goods and services	2950
Gross fixed capital formation	8887
Govt. final consumption expenditure	5057
Exports of goods and services	2658
Change in stocks	1849
Private final consumption expenditure	43685
Depreciation of fixed capital	3206
Indirect taxes	5876
Subsidies	717
Net factor income from abroad	- 323
Income from domestic product accruing to the govt.	654
Current transfers to households	1100
National debt interest	502
Business savings	410
Corporate tax	583
Direct taxes paid by households	1220

(5 x 4) = [20]

10. Practical Records [10]

:bcc:

1985-86 | E113

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods
and Applications : 1985-86

FINAL EXAMINATION

PART - I

Probability

Date: 13.12.1985 Maximum Marks: 100 Time: $2\frac{1}{2}$ hrs.

Note: The paper is set for 114 marks. You may attempt any part of any question. The maximum you can score is 90. Ten marks are alloted for assignments. The marks for each question are given in brackets [ ].

1.(a) Prove that if the random variables X and Y have moments of order r, then X + Y also has a moment of order r.

(b) A die is rolled until an ace appears. What is the probability that at most ten rolls are needed ?

(5+5) = [10]

2.(a) State and prove Chebyshev's Inequality.

(b) Give an example to show that in general the bound given by Chebyshev's Inequality cannot be improved upon.

(c) State Weak Law of Large Numbers.

(2+5+5+3) = [15]

3.(a) Let φ be a differentiable strictly increasing or strictly decreasing function on an interval I, and let $\varphi(I)$ denote the range of φ and φ^{-1} the inverse function to φ. Let X be a continuous random variable having density f such that $f(x) = 0$ for $x \notin I$. Show that $Y = \varphi(X)$ has density g given by

$$g(y) = 0 \quad \text{for} \quad y \notin \varphi(I) \quad \text{and}$$

$$g(y) = f(\varphi^{-1}(y))\left|\frac{d}{dy}\varphi^{-1}(y)\right|, \quad y \in \varphi(I).$$

(b) Let X be a random variable having an exponential density with parameter λ. Find the density of $Y = X^{1/\beta}$, where $\beta \neq 0$.

(10+5) = [15]

p.t.o.

4.(a) Let $g(x) = e^{-x^2/2}$, $-\infty < x < \infty$. Normalize g to make it a density. What this density usually is called ? Is it a symmetric density ?

(b) If a random variable Y is distributed as $n(\mu, \sigma^2)$, then show that the random variable $a + bY$, $b \neq 0$, is distributed as $n(a + b\mu, b^2 \sigma^2)$.

(5+1+1+5) = [12]

5.(a) Let X be a random variable having the normal density $n(0, \sigma^2)$. Find the density of the random variable $Y = X^2$.

(b) Define gamma densities in general.

(c) Let X have the normal density with parameters μ and $\sigma^2 = .50$. Find the constant c such that $P(|X-\mu| \leq c) = .95$.

(6+3+5) = [14]

6.(a) Determine c so that the function

$$f(x, y) = c\, e^{-(x^2-xy+y^2)/2}, \quad -\infty < x, y < \infty$$

becomes a joint density function of two random variables X and Y. Hence show that X and Y are not independent.

(b) Let X and Y be continuous random variables having joint density f given by $f(x,y) = \lambda^2 e^{-\lambda y}$, $0 \leq x \leq y$, and $f(x,y) = 0$ elsewhere. Find the marginal densities of X and Y. Find the joint distribution function of X and Y.

(8+2+3+3+4) = [20]

7.(a) Describe a continuous random variable X having a given density for which expectation does not exist as a finite number.

(b) Let X have the gamma density $\Gamma(\alpha, \lambda)$. Find the variance of X.

(5+5) = [10]

8.(a) State Central Limit Theorem.

(b) Suppose the length of life of a certain kind of light bulb, after it is installed, is exponentially distributed with a mean length of 10 days. As soon as one light bulb burns out, a similar one is installed in its place.

Contd... 3/-

Contd..... Q.No.8.(b)

Find the probability that more than 50 bulbs will be required during a one-year period. It is given that ϕ (1.91) = .9719.

(3+5) = [8]

9. State (true or false):

(i) Let $X_1, \ldots, X_n$ be independent random variables having a common distribution. When the X_i have finite mean, the Weak Law of Large Numbers holds.

(ii) Chebyshev bounds one really poor in the case of binomial distribution.

(iii) The distribution function F(,) of a random variable is a non-increasing function.

(iv) For a continuous random variable X, $P(X = x) = 0$ for all x, where $-\infty < x < \infty$.

(v) Let X be a random variable that has a density f. Then f has a symmetric density if X is a symmetric random variable.

(vi) Let X be a random variable such that

$$P(X > a+b) = P(X > a)\, P(X > b), \quad a \geq 0,\ b \geq 0.$$

Then either $P(X > 0) = 0$ or X is exponentially distributed.

(vii) Exponential densities are special cases of gamma densities.

(viii) Var X = 0 iff X is a constant.

(ix) The variance of a sum of two random variables is the sum of the variances.

(x) It is possible for dependent random variables to be uncorrelated.

[10]

:bcc:

1985-86	E123

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Applications : 1985-86

FINAL EXAMINATION

PART - I

Descriptive Statistics (Theory)

Date: 11.12.1985 Maximum Marks: 100 Time: 2 hrs.

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

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

   (a) Mean, Median and Mode as measures of location of a distribution.

   (b) Sampling and Non-sampling errors.

   (c) Standard normal variate.

   (d) Diagrammatic representation of statistical data.

(8 x 2) = [16]

2. (a) Suppose you have been given n observation pairs $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ on the two variates x and y. Estimate the constants a and b in the regression equation of y on x,

$$y = a + bx$$

   using least squares method. Find the variances of the expected value $\hat{y}$ and of the residual $e = y - \hat{y}$. Hence explain what is measured by the correlation coefficient r between the two variables.

   (b) Write down the normal equations for the linear regression of the form $y = b_0 + b_1x_1 + b_2x_2 + \ldots + b_kx_k$. Give expression for the coefficient of determination in terms of the estimated regression coefficients and $s_{yy}, s_{yx_1}, \ldots, s_{yx_k}$.

   (c) The simple correlation coefficients between x_1, x_2 and x_3 were found to be

$$r_{12} = 0.4, \quad r_{13} = 0.8 \text{ and } r_{23} = 0.7,$$

Contd..... 2/-

Contd... Q.No.2.(c)

where x_1 = scores in mathematics, x_2 = no. of visits to films and x_3 = study hours per day. Comment on the relation between x_1 and x_2.

(d) If $x = u + v$, $y = v + w$, $z = w + u$ where u, v, w are uncorrelated variables and each of which has zero mean and unit variance, find the multiple correlation coefficient between x and two variables y and z. What is the multiple correlation coefficient between y and two variables z and x ?

(20 + 6 + 7 + 9) = [42]

3. Answer any three from the following:

(a) A person travels n equal distances with velocities $v_1, v_2, \ldots, v_n$. Show that his average velocity can not exceed $(v_1 + v_2 + \ldots + v_n)/n$. When will it be equal to this value ?

(b) Let S and R be, respectively, the standard deviation and the range of a set of n values of x. Show that

$$\frac{R^2}{2n} \leq S^2 \leq \frac{R^2}{4}.$$

When do the equalities hold ?

(c) The three central moments m_2, m_3 and m_4 of a variable were found to be 5, 25 and 50 respectively. Are the figures consistent ? Assuming that the figures for m_2 and m_4 are correct, find the range of m_3 for which the figures become consistent.

(d) Show that for binomial distribution

$$\mu_{r+1} = pq(rm\,\mu_{r-1} + d\mu_r/dp),$$

hence find μ_2, μ_3 and μ_4 of this distribution.

(e) For a certain normal distribution the first moment about 10 is 40 and that the fourth moment about 50 is 48. What is the probability that if a random sample is drawn from this distribution it will not exceed 51.45 ?

(10 x 3) = [30]

4. Are the following sets of figures consistent ? If not give reasons for each case.

(a) $r_{12} = 0.6$, $r_{23} = 0.8$, $r_{31} = -0.5$,

where r_{ij} is the simple correlation coefficient between the variables x_i and x_j.

(b) $m'_1 = 1$, $m'_2 = 1$ and $m'_3 = 2$,

where m'_i is the ith raw moment about zero of a variable x.

(c) The mean and standard deviation of binomial distribution are 5 and 3 respectively.

[4 x 3) = [12]

:bcc:

1985-86 | E124

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Applications : 1985-86

FINAL EXAMINATION

PART - I

Descriptive Statistics (Practical)

Date: 9.12.1985 Maximum Marks: 100 Time: $2\frac{1}{2}$ hr[illegible]

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

1.(a) During one year, the rate of milk prices per quart to bre[illegible] prices per loaf was 2.50, whereas during the next year, [illegible] rate was 2.00. (i) Find the arithmatic mean of these ratios of bread prices to milk prices for the two year period, (ii) Find the arithmatic mean of these ratios fo[illegible] the two year period, (iii) Discuss the advisability of using the arithmatic mean for averaging ratios, (iv) Disc[illegible] the suitability of geometric mean for averaging ratios.

[10]

(b) The bacterial count in a certain culture increased from 1000 to 4000 in three days. What was the average percentage increase per day. [5]

(c) The following Table shows the I.Q.'s of 480 School children at a certain elementary school. Find the mean and standard deviation. Also determine the percentage of students' I.Q.'s which fall within (i) $\bar{X} \pm s$, (ii) $\bar{X} \pm 2s$, (iii) $\bar{X} \pm 3s$

Class marks (X)	70	74	78	82	86	90	94	98	102	106	110	114	118	122	126
Frequency (F)	4	9	16	28	45	66	85	72	54	38	27	18	11	5	[illegible]

[10]

2.(a) The mean weight of 500 male students at a certain Coll[illegible] is 151 lb and the standard deviation is 15 lb. Assumi[illegible] the weights are normally distributed, find how many students weigh (i) between 120 and 155 lb, (ii) more th[illegible] 185 lb.

[10]

Contd..... Q.No.2

(b) Fit a normal curve to the following data:

Heights of 100 male students

Heights (inches)	Number of students
60 - 62	5
63 - 65	18
66 - 68	42
69 - 71	27
72 - 74	8

and test the goodness of fit. [15]

3. The following Table shows the respective heights X and Y of a sample of 12 fathers and their oldest sons:

Height (X) of father (inches)	65	63	67	64	68	62	70	66	68	67	69	71
Height (Y) of son (inches)	68	66	68	65	69	66	68	65	71	67	68	70

(i) Construct a scatter diagram [5]

(ii) Find the least square regression of Y on X [5]

(iii) Find the least square regression of X on Y

(iv) Compute the standard error of estimate, $S_{Y.X}$ and construct two lines prallel to the regression line of Y on X and having vertical distance $S_{Y.X}$ from it. Also determine the percentage of data points falling between these two lines. [10]

4.(a) In a partially destroyed laboratory record of an analysis of correlation data, the following results are only legible:

Variance of $x = 9$

Regression Equations: $8x - 10y + 66 = 0$;

$40x - 18y = 214$

What were (i) the mean values of x and y, (ii) the standard deviation of y, and (iii) the coefficient of correlation between x and y. $[12\frac{1}{2}]$

Contd..... 3/-

Contd..... Q.No.4

(b) A Computer while Calculating the Correlation Coefficient between two variables x and y from 25 pairs of observations obtained the following constants.

$$n = 25,\ \Sigma x = 125,\ \Sigma x^2 = 650,\ \Sigma y^2 = 460,\ \Sigma xy = 508.$$

It was however later discovered at the time of checking that he had copied down two pairs

x	y
6	14
8	6

While the corrected value were

x	y
8	12
6	8

Obtain the correct value of correlation coefficient.

[$12\frac{1}{2}$]

5. The following Table shows the weights X_1 to the nearest pound, heights X_2 to the nearest inch and ages X_3 to the nearest year of 12 boys:

Weight (X_1)	64	71	53	67	55	58	77	57	56	51	76	68
Height (X_2)	57	59	49	62	51	50	55	48	52	42	61	57
Age (X_3)	8	10	6	11	8	7	10	9	10	6	12	9

Calculate (i) $R_{2.13}$ and $R_{3.12}$ and compare with the value of $R_{1.23}$.

[25]

:bcc:

1985-86	E132

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Applications : 1985-86

PERIODICAL EXAMINATION

PART - I

Economics Statistics (Theory and Practical)(c+d)

Date: 6 November 1985 Maximum Marks: 100 Time: 2 hrs.

Note: Answer Question No.1 and any TWO out of the rest. Marks allotted to different questions are shown in brackets [ ].

1. The following time series relates to the quantity index of world exports y_t, (base 1958 = 100) for the period from 1955 to 1970 :

Year	1955	1956	1957	1958	1959	1960	1961	1962
Index (y_t)	89	95	101	100	108	120	126	135

Year	1963	1964	1965	1966	1967	1968	1969	1970
Index (y_t)	144	158	170	183	193	217	241	262

(a) Plot these data on a diagram and comment on the nature of trend displayed by these points.

[20]

(b) Compute trend values by the moving average method, taking a 3 - year period as the period of the moving average. Plot the trend values on the diagram.

(15+5) = [20]

2. What are the different components of a time series. Discuss these components with suitable illustrations.

[30]

3. Write short notes on the following methods of computing trend values of a time series :

(i) free-hand curve fitting
(ii) moving average

(8+22) = [30]

p.t.o.

4. What is meant by the least squares method of estimation ? Illustrate this method in the context of fitting a quadratic trend curve to a time series variable.

[30]

:-bcc:

1985-86	E122

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods
and Applications : 1985-86

PERIODICAL EXAMINATION

PART - I

Descriptive Statistics (Theory and Practical)

Date: 30.10.1985 Maximum Marks: 100 Time: 2 hrs.

Note: Attempt all questions as indicated.

1.(a) Write down the density function of the normal distribution. Show that for a normal distribution,

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

(ii) $E \mid X - \mu \mid = \sqrt{2/\pi}\, \sigma$

(iii) $E(e^{tx}) = e^{t\mu + \frac{1}{2} t^2 \sigma^2}$

[20]

(b) (i) If X is normally distributed with zero mean and variance unit. Find the expectation and variance of X^2.

[10]

(ii) If X is normally distributed with mean 11 and s.d. 1.5, find the number x_o such that $P(X > x_o) = 0.3$.

[5]

(iii) Let X be a chance variate having a normal distribution with mean 30 and s.d. 5. Find the probabilities,

$P(26 \leq X \leq 40)$, $P(X > 28)$ and $P(X > 33)$

[5]

(c) Derive the poisson distribution as a limiting form of the binomial distribution. Give examples of data for which the poisson distribution is expected to give a good fit.

[15]

2. Answer any four of the following questions:

(i) The incidence of occupational disease in an industry is such that the workers have a 20% chance of suffering from it. What is the probability that out of 6 workers, 4 or more will catch the disease ?

[5]

Contd... 2/-

Contd..... Q.No.2

(ii) Let S and R be respectively the s.d. and the range of a set of n values of x. Show that

$$R^2/2n \leq S^2 \leq R^2/4$$ [5]

(iii) Suppose that variable x takes positive values and that deviation $(x_i - \bar{x})$ are small compared to $\bar{x}$. Show that in such cases,

$$x_g \simeq \bar{x}(1 - \tfrac{1}{2}\, s^2/\bar{x}^2) \quad \text{and} \quad x_h \simeq \bar{x}(1 - \frac{s^2}{\bar{x}^2})$$

[5]

(iv) A random variable X has the probability density function

$$f(X = x) = K.\ \text{Sin}^{2n}x\ \text{Cos}\,x\ , \ -\pi/2 \leq x \leq \pi/2$$

Determine K and calculate the probability that

(a) $X \leq \theta$ and $|X| \geq \theta$ for any given θ and

(b) $-\pi/6 \leq X \leq \pi/6$ [5]

(v) The probability density function of the random variable X is, for X = x, proportional to $1/(1+x^4)$ defined in the doubly infinite range $(-\infty < x < \infty)$. Find the proportionality factor and identify the form of the distribution. Also, show that the variance of X is unity. [5]

3. When the first proof of a book containing 250 pages was read, the following distribution of misprints was found:

Number of misprints per page	Frequency
0	139
1	76
2	28
3	4
4	2
5	1
	250

Fit a poisson distribution to the above data and also test the goodness of fit.

[25]

:bcc:

1985-86	E131

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Applications : 1985-86

PERIODICAL EXAMINATION

PART - I

Economics Statistics (Theory and Practical)(a+b)

Date: 9.10.1985 Maximum Marks: 100 Time: 2 hrs.

Note: This question paper carries 125 marks. You can answer any part of any question; but the maximum that you can score is 100.

1. What are 'time reversal' and 'circular' tests ? Find two index number formulae which satisfy the former test but not the latter one.

(9+6) = [15]

2. Distinguish between an 'aggregative' and an 'average-type' index number formula. Can you interpret Laspeyres' formula in terms of either of the above two approaches ?

(9+6) = [15]

3.(a) Define a chain-base price index number and discuss its usefulness.

(b) The following table gives prices per kg. in Rs. (p_t) and quantities in thousand kgs. (q_t) of different varieties of apples during the period 1959-63:

Year	Variety A		Variety B		Variety C	
	p_t	q_t	p_t	q_t	p_t	q_t
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1959	4.4	30.5	7.1	17.1	7.7	13.5
1960	4.3	31.6	7.5	13.6	9.5	9.0
1961	4.5	35.0	7.8	15.2	10.2	7.6
1962	5.9	30.0	8.6	12.4	9.2	12.6
1963	6.2	26.8	7.8	15.8	7.7	13.5

p.t.o.

Contd..... Q.No.3.(b)

(i) Calculate Marshall-Edgeworth quantity index number for the year 1963 with 1960 as the base.

(ii) Calculate Fisher's price index number for the year 1962 with 1961 as the base.

(iii) Calculate chain-base price index number for the year 1962 with 1959 as the base using Laspeyres'-type link indices.

(8+7+10) = [25]

4.(a) Write a short note on consumer price index number. [14]

(b) In 1971-72 the all-commodity index number of wholesale prices in India was 188.4 (base 1961-62 = 100). Percentage increases in wholesale prices of all groups except food articles in this year over 1961-62 are given below with their corresponding weights in the all-commodity index. What was the percentage increase in wholesale prices of food articles over this decade ?

Group	Percentage increase in prices in 1971-72 over 1961-62	Weight
(1)	(2)	(3)
Liquor and tobacco	94.8	25
Fuel, power, light and lubricants	72.1	61
Industrial raw materials	91.0	121
Chemicals	97.0	7
Machinery and transport equipments	59.0	79
Manufactures	67.1	294
Food articles	?	413

[16]

p.t.o

5.(a) The index of Business Activity is constructed by taking the weighted average of activity relatives in the different business sectors of the economy. The sectors, their activity levels in 1948 and in May 1949 as also the weights of the different sectors are shown in the table below. The index of industrial production in May 1949 with 1948 as base is 117.6. Calculate the index of Business Activity in India in May 1949 (1948 = 100).

Item	Business Activity in India		Weight
	Monthly average in 1948	May 1949	
(1)	(2)	(3)	(4)
Industrial production	-	-	46
Foreign trade	-	-	6
Export (value Rs.)	1512539	1221671	3
Import (value Rs.)	1640115	2902524	3
Shipping activity	-	-	2
Tonnage entered	13376	27848	1
Tonnage cleared	13068	38536	1
Financial activity	-	-	23
Cheque clearance (Rs. crores)	137.8	127.1	18
Note circulation (Rs. crores)	341.5	315.4	5
Tons lifted by Railways	5588	7102	23

[17]

(b) The table below gives the retail prices of the group 'foodgrains and products' in two centres : Calcutta and Bombay in 1960. Column 2 of this table gives the average family expenditure (of industrial workers) in Calcutta in 1960. Using these informations, calculate the consumer

Contd..... 4/-

Contd..... Q.No.5.(b)

price index number of 'foodgrains and products' for Bombay taking Calcutta as base.

Item	Average family expenditure in Calcutta in 1960 (Rs.)	Unit	Price (Rs.) in	
			Calcutta	Bombay
(1)	(2)	(3)	(4)	(5)
Rice	17.55	kg.	0.74	0.70
Wheat	4.21	kg.	0.40	0.41
Bread	0.21	500 gms.	0.40	0.48
Arhar dal	0.92	kg.	0.75	0.78
Gram dal	0.27	kg.	0.58	0.60
Moog dal	0.77	kg.	0.78	0.90
Musur dal	0.79	kg.	0.73	0.73

[13]

:bcc:
71085

1985-86 | E112

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistics and Applications : 1985-86

PERIODICAL EXAMINATION

PART - I

Probability

Date: 16.10.1985 Maximum Marks: 100 Time: 2 hrs.

Note: The paper is set for 109 marks. You may attempt any part of any question. The maximum you can score is 100. The marks alloted for each question are given in brackets [ ].

1. Define probability generating function ϕ_X for a random variable X. Hence calculate $\phi_X(t)$ for X having binomial distribution and Poisson distribution.

   Let X and Y be two independent, non-negative integer-valued random variables. Show that $\phi_{X+Y}(t) = \phi_X(t)\ \phi_Y(t)$.

   Let $X_1, \ldots, X_r$ be independent random variables. If X_i, $1 \leq i \leq r$ has the negative binomial distribution with parameters α_i and p, then show that $X_1 + \ldots + X_r$ has the negative binomial distribution with parameters $\alpha_1 + \ldots + \alpha_r$ and p.

   (2+5+5+5+8) = [25]

2. Define expectation of a discrete random variable. Calculate expectation of a random variable X having geometric distribution with parameter p.

   Give an example of a density that does not have finite expectation.

   (2+5+5) = [12]

3. Let X and Y be two random variables having finite expectation.

   (i) Show that X+Y has finite expectation and $E(X+Y) = EX + EY$.

   (ii) Suppose $P(X \geq Y) = 1$. Then $EX \geq EY$.

   Is it always true that $E(XY) = (EX)(EY)$? Give reasons for your answer.

   (5+5+5) = [15]

p.t.o.

4. Let X and Y be two random variables having moments of order r. Show that X+Y also has a moment of order r.

   Define variance of a random variable X. Show that Var X = 0 if and only if X is a constant.

   Let X be a random variable having a finite second moment. Find the value of λ that minimizes $E(X - \lambda)^2$.

   It is given that $EX = \phi_X'(1)$ and $\text{Var } X = \phi_X''(1) + \phi_X'(1) - [\phi_X'(1)]^2$, where $\phi_X(t)$ is the probability generating function of X. Find the mean and variance of X having Poisson distribution.

   (5+2+5+5+5) = [22]

5. Define covariance of two random variables X and Y.

   Is it always true that var(x + Y) = var X + var Y ? Give reasons for your answer.

   Give the statement of Schwarz inequality.

   Define correlation coefficient of two random variables X and Y. Using Schwarz inequality show that correlation coefficient is always between -1 and 1.

   (2+5+3+2+3) = [15]

6. Let X and Y be independent random variables each geometrically distributed with parameter p. Find the distribution of min(X , Y).

   Suppose a box contains 7 red balls and 3 blue balls. If 5 balls are selected at random without replacement, determine the density function of the number of red balls that will be obtained.

   If 10 percent of the balls in a certain box are red and if 20 balls are selected from the box at random with replacement, what is the probability that more than 3 red balls will be obtained ?

   Suppose that X and Y have a discrete joint distribution for which the joint density is defined as follows:

$$f(x,y) = \begin{cases} \frac{1}{30}(x+y) & \text{for } x = 0, 1, 2 \text{ and } Y = 0, 1, 2, 3 \\ 0 & \text{elsewhere} \end{cases}$$

   Are X and Y independent ?

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

:bcc:

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Applications: 1985-86

PERIODICAL EXAMINATION

PART - I

Descriptive Statistics (Theory and Practical)

Date: 25.9.1985 Maximum Marks: 100 Time: 2 hrs.

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

1. Choose the correct answer among the four alternatives from each of the following:

(a) The average depth of a river is 4 feet. The average height of the members of a family is 4.5 feet. The family can safely cross the river if

(i) the height of each member of the family is more than 4 feet.

(ii) the breadth of the river is not more than 20 feet.

(iii) the maximum depth of the river is not more than 4.5 feet.

(iv) none of these.

(b) A population of values is symmetrically distributed about the constant K. Then

(i) the mean coincides with mode.
(ii) the median coincides with mode.
(iii) $(Q_1 + Q_3)/2 = K$.
(iv) none of these.

(c) The average speed of a man travelling r_1 miles at $1/r_1$ m.p.h. r_2 miles at $1/r_2$ m.p.h., ..., r_k miles at $1/r_k$ m.p.h. is x where x is

(i) the arithmetic mean of $r_1, r_2, \ldots, r_k$.

(ii) the harmonic mean of $r_1, r_2, \ldots, r_k$.

(iii) the arithmetic mean of $r_1, r_2, \ldots, r_k$ taking weights as $1/r_1, 1/r_2, \ldots, 1/r_k$ respectively.

(iv) none of these.

p.t.o.

Contd..... Q.No.1

(d) For a distribution of 10 observations on a variable x, the mean and the variance are 10.5 and 2.25 respectively. However, on scrutinising the data it is found that one observation which should be correctly read as 10, had been wrongly recorded as 15, then

(i) either the mean or the variance or both have been wrongly calculated.

(ii) all the observations are equal.

(iii) the correct variance will be greater than 2.25.

(iv) none of these.

(e) In a certain distribution, the first four raw moments are 1, 3, 7 and 29 respectively, then

(i) $g_1 > 0$ and $g_2 < 0$
(ii) $g_1 = 0$ and $g_2 > 0$
(iii) $g_1 = 0$ and $g_2 < 0$
(iv) none of these.

[Marks will be deducted for wrong choice]

(5 x 4) = [20]

2. Write a few sentences with examples to distinguish between

(a) a variable and an attribute.
(b) sampling errors and non-sampling errors.
(c) a parameter and a statistic.

(3 x 5) = [15]

3. What do you understand by dispersion ? What are the different measures of dispersion ? Explain their uses.

(2+5+3) = [10]

4. Answer any two from (a), (b), (c) and (d).

(a) The frequencies of values 0, 1, 2, ..., n of a variable are given by

$$q^n,\ {}^nc_1\, q^{n-1}\, p,\ \ {}^nc_2\, q^{n-2}\, p^2,\ \ldots,\ p^n$$

where $p+q = 1$. Show that the mean is np.

(b) Show that the mean deviation about A (MD_A) may be obtained by the formula

$$nMD_A = S_2 - S_1 + A\,(n_1 - n_2),$$

Contd..... 3/-

contd..... Q.No.4.(b)

where S_1 is the sum of the values that are less than A and n_1 is the number of such values; while S_2 is the sum of the values that are greater than A and n_2 is the number of such values.

(c) Prove that for any distribution

$$b_2 - b_1 - 1 \geq 0.$$

Discuss the case where $b_2 - b_1 - 1 = 0$.

(d) If in a series of measurements we obtain m_1 values of magnitude x_1, m_2 of magnitude x_2 and so on, and if $\bar{x}$ is the mean value of all the measurements, prove that the variance is

$$\frac{\Sigma\, m_i (K-x_i)^2}{\Sigma\, m_i} - d^2$$

where $\bar{x}$ = K+d, and K is any constant.

(2 x 10) = [20

Particulars relating to the wage-distributions of two manufacturing firms are given below:

	Firm A	Firm B
Mean	277	285
Median	271	262
Mode	260	251
Q_1	262	258
Q_3	276	290
s.d.	32	39

compare the two distributions. [13]

Determine the mean, the median, the mode, the standard deviation and the coefficient of variation for the following distribution of monthly income for 580 middle-class people:

Monthly income (Rs.)	Frequency
50 - 100	53
100 - 150	81
150 - 200	114
200 - 250	195
250 - 300	63
300 - 350	32
350 - 400	20
400 - 450	11
450 - 500	8
500 - 550	3
Total	580

(5+4+4+7+2) = [22]

DCC:

1984-85	E263S

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Applications : 1984-85

SUPPLEMENTARY EXAMINATION

Design of Experiments (Theory and Practical)
and
Industrial Statistics (Theory and Practical)

PART - II FINAL

Date: 13.9.1985 Maximum Marks: 100 Time: 2 hrs.

Note: Answer any three questions. All questions carry equal marks. 2 marks for neatness.

Group A

Design of Experiments (Theory and Practical)

1. Define completely randomised design and randomised block design. When are they used ? What are their advantages and disadvantages ?

(4+4+8) = [16]

2. Define a latin square design. Explain with an example how one randomises a latin square of order 4.

(4+12) = [16]

3. You are asked to compare 4 soap solutions regarding their effectiveness in washing cloth. 20 pieces of cloth of equal size are given. Suggest a plan for the experiment, when
   (a) the pieces are taken from the same material.
   (b) they are taken from 5 different materials.

(8+8) = [16]

4. Write short notes on any two of the following:
   (a) Experimental error and error sum of squares.
   (b) Least square estimation.
   (c) Main effects and interaction.

p.t.o.

Group B

Industrial Statistics (Theory and Practical)

Note: Answer Q.No.1 and any one from the remaining.
Numbers in bracket [ ] indicate full marks.

1. The following data shows the number of defectives observed in samples of 200 each for a period of one month for 24 days:

Sample No.	No. of defectives	Sample No.	No. of defectives
1	6	13	2
2	6	14	4
3	6	15	7
4	5	16	1
5	0	17	3
6	0	18	1
7	6	19	4
8	14	20	0
9	4	21	4
10	0	22	15
11	1	23	4
12	8	24	1

(a) Construct an np-Chart for controlling production in the next month.

(b) Draw the OC-curve for this np-Chart.

(20+15) = [35]

2. (a) What is process capability of a process ? How do you estimate the process capability of a process for which an $\bar{X}$-R Chart is being maintained ?

(b) $\bar{X}$-R Charts are being maintained in controlling the weight of paper (gms/cm^2) on a paper making machine. Both the $\bar{X}$ and R charts show a state of statistical control for a Sample of Size 2. The average range was observed to be 5.8 and the average of the sample averages ($\bar{\bar{X}}$) was 295 gs Examine whether the process capability is adequate to meet the specifications of 300 $\pm$ 15 gsm. Is the out going product meeting the specifications under the existing conditions ?

(5+10) = [15]

Contd..... 3/

3.(a) Explain the use of OC Curve in acceptance sampling.

(b) Compute the probability of acceptance P_a for the following single sampling plans:

(i)	$n = 50$,	$c = 2$,	$p = 0.03$
(ii)	$n = 175$,	$c = 11$,	$p = 0.08$
(iii)	$n = 15$,	$c = 0$,	$p = .004$

(c) Draw a general shape of an OC curve which goes through two points, ($p = 0.012$, $P_a = 0.94$) and ($p = 0.07$, $P_a = 0.05$). State the exact meaning of the P_a values on the OC curve at these two points.

(4+6+5) = [15]

:bcc:

1985-86 E111

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Applications: 1985-86

PERIODICAL EXAMINATION

PART - I

Probability

Date: 11.9.1985 Maximum Marks: 100 Time: 2 hrs.

Note: The paper is set for 109 marks. You may attempt any part of any question. The maximum you can score is 100. The marks alloted for each question are given in brackets [ ].

1. Explain each of the following terms with a suitable example :

   (a) Random experiment
   (b) Sample space
   (c) Event
   (d) Probability space

   (4 x 2) = [8]

2. Prove that

   (i) if events A and B are such that $A \subset B$, then $P(A) \leq P(B)$.

   (ii) for any event A, $P(A') = 1 - P(A)$, where A' denotes the complement of A.

   (4 + 4) = [8]

3. (a) When the two events A and B are said to be independent ?

   (b) Suppose a machine produces bolts, 10% of which are defective. Find the probability that a box of 3 bolts contains at most one defective bolt.

   (2 + 5) = [7]

4. (a) Consider a random permutation of n distinct objects. We say a match occurs at the ith position if the ith object is in the ith position. Find the probability that there are no matches. Hence find the probability that there are exactly r matches.

   (b) From a deck of 52 playing Cards, 13 Cards are drawn at random without replacement. Find the prob. of each of the following:

   A = the ace of spades occurs in the sample.

   B = at least one of the four aces occurs in the sample.

   (10 + 10) = [20]

p.t.o.

5. (a) Define a discrete real-valued random variable and the discrete density function of a random variable.

(b) Show that

$$f(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x}, & x = 0, 1, 2, \ldots, n \\ 0 & , \text{ elsewhere} \end{cases}$$

is a discrete density function.

(c) Let N be a positive integer and let

$$f(x) = \begin{cases} k\, e^x , & x = 1, 2, \ldots, N \\ 0 , & \text{elsewhere} \end{cases}$$

Find the value of k such that f is a discrete density.

(4 + 8 + 8) = [20]

6. (a) When a real-valued function is called a discrete density function ? Show that any discrete density function f is the density function of some random variable X.

(b) Define the distribution function of a random variable.

(c) Let F(x) be the distribution function of a random variable X. Show that

(i) F(x) is a non-decreasing function of x.

(ii) For any two real numbers a, b, $a < b$, $P(a < X \leq b) = F(b) - F(a)$.

(8 + 2 + 10) = [20]

7. (a) Suppose a random variable X has a Poisson distribution for which the parameter $\lambda = 2$. Find $P(X \leq 6)$.

(b) Show that there does not exist any number c such that the following function would be a discrete density function:

$$f(x) = \begin{cases} \frac{c}{x} & , \text{for } x = 1, 2, \ldots \\ 0 & , \text{elsewhere} \end{cases}$$

(5 + 5) = [10]

8. (a) Define a discrete r-dimensional random vector.

   (b) Let X_1, X_2, ..., X_r be r discrete random variables having densities f_1, f_2, ..., f_r respectively. When these random variables are said to be mutually independent ?

   (c) Let X and Y be two random variables having the joint density given by the following table:

X \ Y	-1	0	2	6
-2	1/9	1/27	1/27	1/9
1	2/9	0	1/9	1/9
3	0	0	1/9	4/27

   Find the distribution of X and compute the probability of event Y is even. (3 + 3 + 10) = [16]

;bcc:

1984-85 | E253S

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods and Applications : 1984-85

SUPPLEMENTARY EXAMINATION

Sample Surveys (Theory and Practical)
PART - II FINAL

Date: 11.9.1985 Maximum Marks: 100 Time: 2 hrs.

Note: Answer as much as you can. The maximum possible score is 100.

1. What is a PPSWR sampling scheme ? Explain Lahiri's method. Show that it, infact, gives rise to PPS sampling scheme. How do you improve Lahiri's method ? How would you estimate the population mean based on a sample drawn by Lahiri's method ? Show that your estimator is unbiased. Obtain its variance and estimate it.

(2+3+4+2+3+4+4) = [22]

2. What is a ratio estimator ? When should one go for ratio estimator ? Estimate the population total using ratio method of estimation. Is your estimator unbiased ? If not, obtain its bias. Obtain its MSE. Estimate the bias as well as the MSE.

(2+3+2+4+4+4+4) = [23]

3. Describe the situation in which you can employ a two stage sampling design to estimate the population parameters. Based on a two stage sampling design that uses SRSWOR at both the stages how would you estimate the population total ? Is your estimator unbiased ? Obtain its MSE and estimate it.

(3+2+4+6+6) = [21]

4. A population of 112 villages has been divided into 3 strata. A sample of 5 villages selected with SRSWOR from the first stratum, a PPSWR sample of 3 villages from the second stratum and two independent linear systematic samples of 4 villages each from the third stratum are given.

Let x : cultivated area, y : area under wheat, N_h: no. of villages in the hth stratum, $1 \leq h \leq 3$.

The relevant information is given below

$N_1 = 51$, $N_2 = 36$, $N_3 = 25$.

X_2 = Total cultivated area in stratum 2 = 21860 acres.

Contd..... 2/.

Contd...... Q.No.4

Stratum 1	y : 75, 101, 5, 78, 78	
Stratum 2	x : 729, 617, 569	
	y : 247, 238, 223	
Stratum 3	sample 1	y : 427, 326, 461, 445
	sample 2	y : 335, 412, 503, 348, 398

(a) Estimate the total area under wheat in each stratum separately and also in all the three strata taken together.

(b) Obtain estimates of the variances of the above estim...

(5 x 8) = [40

:bcc: