INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Applications : 1987-88

FINAL EXAMINATION

PART - II

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

Date: 4.7.1988

GROUP - A

Point and Interval Estimation
(Practical)

Maximum Marks: 30 Time: 45 Minutes

Note: Attempt all questions. Your maximum score for the group cannot exceed 30 marks.

1. Items produced by a company are packed into boxes containing two dozen items each. A sample of 60 boxes is selected randomly and with replacement and number of defective items is found as

No. of boxes	1	20	23	4	6	6
No. of def. items	5	1	2	0	3	4

Let θ be the probability of an item being defective

(a) Obtain unbiased, consistent, sufficient estimate of θ.

(b) Obtain MLE for θ.

(c) Obtain unbiased consistent and sufficient estimate for $X(\theta)$, the expected number of defective items in a box.

(8+3+4) = [15]

2. Annual earnings of households in a zone are assumed to be distributed normally. Let μ_1 be the mean and σ_1^2 the variance of the distribution at initial stage and μ_2 and σ_2^2 after a development project has been completed. Independent random samples drawn before and after the completion of the project showed the annual earnings as (in Rs 1000).

Contd..... 2/-

Contd..... Q.No.2

Sample 1 : 2.5, 3.1, 1.5, 2.2, 4.7, 2.2, 1.9, 2.5

Sample 2 : 3.1, 3.3, 2.8, 2.9, 4.1, 2.4, 3.1, 3.0, 4.5, 3.6.

(a) Obtain best unbiased estimate of average change $\mu_2 - \mu_1$ and also obtain an unbiased estimate of the variance of the estimate.

(b) Obtain 95% confidence interval for $\mu_2 - \mu_1$.

(c) Obtain MLE for $\mu_2 - \mu_1$.

GROUP - B

Linear Estimation
(Practical)

Maximum Marks: 30 Time: 45 Minutes

Note: Answer all the questions.

1. The following data show the birth-weights (in lb) of babies born, classified according to the age of mother.

Birth - weights

Age of mothers (in years)				
15-20	20-25	25-30	30-35	35 and above
5.1	5.0	5.1	5.0	5.0
5.0	5.3	5.1	4.9	5.1
4.8	5.3	4.9	6.0	5.9
5.2	6.2	4.8	5.8	5.6
5.4	6.5	6.0	5.5	6.0
			5.5	6.2

Test whether the age of mother significantly affects the birth-weight of babies.

[13]

2. Data were collected on the effect of lacquare concentrations at four levels and standing times at three levels on screeen quality as indicated on a 20-point scale. Two observations were made for each of twelve treatment combinations. The resulting data produced the following incomplete ANOVA table.

Screen quality ANOVA

Source	d.f.	S.S.	M.S.	F	$F_{.05}$
Lacquer (L)	-	63.79	-	-	-
Time (T)	-	126.58	-	-	-
L X T interaction	-	38.07	-	-	-
Error	-	-	-		
Total	-	262.95			

(a) Complete the ANOVA table.

(b) Set up a mathematical model for the experiment stating the appropriate assumptions and indicate the hypotheses to be tested in terms of your model.

(c) On the basis of computed F values in the table draw your conclusions.

(5+6+6) = [17]

GROUP - C

Design of Experiments
(Practical)

Maximum Marks: 40 Time: 1 hour

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 blocks of a 2^3 factorial design are run. The results follow:

Contd..... 4/-

Contd..... Q.No.1

Treatment combination	Block		
	I	II	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

Analysis the data and make a report. Also estimate the error-variance σ^2. You should write down the model clearly and explain all the notation.

(35+5) = [40]

:bcc:

1987-88	E253

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Applications : 1987-88

FINAL EXAMINATION

PART - II

Date: 1.7.1988 Maximum Marks: 100 Total Time: $2\frac{1}{2}$ hours

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

GROUP - A

Point and Interval Estimation (Theory)

Maximum Marks: 30 Time: 45 minutes

Note: Attempt all questions. Your maximum score for this group can not exceed 30 marks.

1. (i) Define a sufficient statistic.
   (ii) State Fisher-Neyman factorization theorem for sufficieney.
   (iii) Give maximum likelihood method of estimation.
   (iv) State the properties of maximum likelihood estimates.

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

2. (i) State the Cramer-Rao inequality along with the regularity conditions.
   (ii) Obtain sufficient statistics for (μ, σ^2) on the basis of a random sample from $N(\mu, \sigma^2)$.
   (iii) Obtain MLEs for α and β for a population characterized by

$$f(x) = \frac{1}{\beta - \alpha} \qquad \alpha < x < \beta$$

(7+4+3) = [14]

3. Prove or disprove the following statements:
   (i) MLE is always unbiased.
   (ii) A sufficient statistic is always unbiased.
   (iii) Sample mean is BLUE in the class $d = \sum_{i=1}^{n} a_i$ with $\sum_{1}^{n} a_i = 1$.

Contd..... 2/-

Contd..... Q.No.3

(iv) Sample mean is consistent for the population mean.

(v) Sample mean is UMVU for μ in $N(\mu, \sigma^2)$.

[5]

GROUP - B

Linear Estimation
(Theory)

Maximum Marks: 30 Time: 45 Minutes

Note: Answer any TWO questions.

1. Define a linearly estimable parametric function. State Gauss-Markov linear model. Write down the linear model with all the assumptions for a one-way classified data. Hence describe the method of obtaining best linear unbiased estimators of the parametric functions of this linear model.

(3+3+9) = [15]

2. What is analysis of variance ?

Describe the detail analysis of a two-way classified fixed effects model with one observation per cell. Represent the entire analysis in a tabular form.

(5+10) = [15]

3. What is regression model ? How does it differ from experimental design model ?

Describe how would you apply analysis of variance technique to determine (i) whether there exists a relationship between two variables, (ii) whether the relationship is linear or not.

(5+10) = [15]

P.t.o.

GROUP - C

Design of Experiments (Theory)

Maximum Marks: 40 Time: 1 hour

Note: Answer all the questions.

1. Explain, with reference to CRD, RBD and LSD, the principles of randomisation and local control.

(8+12) = [20]

2. In a 2^3 experiment with the treatments A, B and C, define the main effect of B, interaction AC. How will you compute the sum of squares due to these effects. Explain clearly the notation you are using.

[(3+3) + (2+2)] = [10]

3. Obtain a layout of a LSD with 5 treatments. Explain the different steps clearly.

[10]

:bcc:

1987-88	E273

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Applications : 1987-88

FINAL EXAMINATION

PART - II

Grade 7 : (i) Survey Sampling and Organisation Aspects
(ii) Elements of OR | Practical

Date: 24.6.1988

GROUP - A

Survey Sampling and Organisation Aspects
(Practical)

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

Note: This paper carries 100 marks. The marks for each question are given within square brackets at the end of each question .

1. To estimate the total yield of paddy in a district of West Bengal, a two stage sampling scheme was used. At first stage, a SRSWOR sample of 4 villages, out of 50 villages in the district, was taken and at second stage a SRSWOR sample of 3 plots from each selected village was drawn. The other relevant information is given below:

Sample village	Number of plots	Yield of paddy (in Kg./in sampled plot		
		1	2	3
1	72	24	32	27
2	35	26	13	9
3	86	11	21	21
4	20	32	32	42

Given that the total number of plots in the district is 2535.

Estimate unbiasedly the total yield in the district and also the rse of your estimate.

Assuming that the cost function of the survey is given by $C = 20n + 12mn$, find optimum values of m and n for a fixed total cost Rs.500/-.

(10+25+15) = [50]

p.t.o.

2. The following table gives the number of workers (x) and output (in 000, Rs.) (y) of 15 factories in a certain region. Estimate the total output of all the factories in the region by drawing a PPSWR sample of size 5, size being x, the number of workers and rse of your estimate. Also estimate the gain in precision in using PPSWR over SRSWR on the basis of PPSWR sample so drawn.

Sl.No.	x	y	Sl.No.	x	y
1	20	10	9	15	5
2	25	12	10	75	50
3	10	5	11	35	25
4	30	2	12	80	75
5	50	8	13	20	10
6	22	12	14	25	15
7	80	35	15	15	2
8	60	20			

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

GROUP - B

Elements of OR (Practical)

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

Note: Answer all the questions. The marks allotted to each question are given in [ ] in the margin.

1. Solve the following problem using simplex method

$$\text{Maximise } Z = 3x_1 + 5x_2 + 4x_3$$

$$\text{subject to } 2x_1 + 3x_2 \leq 8$$

$$2x_2 + 5x_3 \leq 10$$

$$3x_1 + 2x_2 + 4x_3 \leq 15$$

$$x_1 \geq 0,\ x_2 \geq 0,\ x_3 \geq 0$$

[40]

2.(a) Five machines are engaged to produce components of a agricultural equipment. From past experience it is known that these machines will operate for an average of 80 hours,

Contd..... 3/-

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

then require an average of 20 hours repair. Each machine will give profit of Rs 150 per productive hour and the cost of servicing the machine is Rs 50 per hour per service man. The cost per service man is incurred irrespective of the fact that whether he is actually working or not. Obtain optimal number of serviceman maximising the profit.

(b) In a university computer centre jobs arrive at the rate of 4 jobs per minute in a Poisson manner. The person who handles the receipt of these jobs is found to spend T minutes per job where distribution of the random variable T is found to be as follows:

T (in minutes)	Probability
0.1	0.20
0.2	0.40
0.3	0.30
0.4	0.10

What is the average time a job takes in the system for the receiving operation ?

(25+15) = [40]

3. The user of a commodity has the option of either manufacturing it or buying it from the market. The set up cost is Rs.30/- per set up. Other details are given below.

Annual demand	=	20,000 units
Price/Unit	=	Rs.2.00
Order Cost	=	Rs.40/- per order
Inventory carrying cost	=	Rs.0.20 per rupee value of inventory per annum
Manufacturing cost/unit	=	Rs.2.50

Which option should the user accept and why ?

[20]

:bcc:

1987-88	E273

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Applications : 1987-88

FINAL EXAMINATION

PART - II

Grade 7 : (i) Survey Sampling and Organisation Aspects (Theory)
(ii) Elements of OR (Theory)

Date: 22.6.1988

GROUP - A

Survey Sampling and Organisation Aspects (Theory)

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

Note: Answer any FOUR questions. Each question carries equal marks. Maximum marks for each subdivision of a question have been indicated in the margin.

1.(a) Write in details the different steps in the excecution of a large scale sample survey.

(b) Suppose you want to make a "Family Budget" Survey for the state of West Bengal. Suggest a suitable sampling plan. Indicate the items of information to be collected and suggest a schedule.

Or

Suggest a sampling design for the survey of manufacturing establishments for the state. Consider the small, medium and large industries separately, after defining these groups of factories suitably. Indicate the items of information to be collected and suggest a schedule.

(8 + (6+4+7)] = [25]

2.(a) How would you estimate the gain in precision in using PPSWR - sampling over SRSWR on the basis of a PPSWR - Sample ? On the basis of a SRSWR ?

Contd..... 2/-

Contd..... Q.No.2

(b) Find a sampling design for which the ratio estimator of population total is unbiased. How would you adjust the mean of the ratio's, $\frac{1}{n}\sum_{i=1}^{n} r_i$, $r_i = \frac{y_i}{x_i}$, in SRSWOR, to obtain an unbiased estimator of population total.

[(6+6) + (6+7)] = [25]

3.(a) In two stage sampling with the selection of one fsu with equal probability, obtain two different estimates of population total and their mean square errors.

(b) Obtain Durbin - Desraj - Rao estimator of variance of a linear estimator of population total in two stage sampling.

(10+15) = [25]

4.(a) Obtain ratio to mean estimator of population total in two stage sampling, its mean square error and an estimate of the mse.

(b) In two stage sampling with simple random sampling at each stage and a constant second stage sampling fraction from a population having equal sized second - stage units, obtain the optimum values of m and n assuming a linear cost function.

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

5. In two stage sampling with fsu's being selected with probability proportional to the number of second stage units in them and with replacement obtain unbiased estimator of population total, its variance and an estimator of variance. When is such a design self-weighted ? For such a sampling design obtain optimum value of m and n, assuming a linear cost function.

(4+6+4+2+9) = [25]

6. Write short notes on

(a) Non-sampling errors and their control in sample surveys.

(b) Inter-penetrating subsampling techniques.

(c) Pilot surveys.

(10+8+7) = [25]

GROUP - B

Elements of OR (Theory)

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

Note: Answer all the questions. The marks allotted to each question are provided in [ ] in the margin.

1.(a) What is a basic feasible solution to a linear programming problem ? What is a degenerate solution ?

(b) Explain the idea behind the two phase method for solving a LP problem.

(c) Derive the necessary formula for carrying out simplex algorithm for a maximisation LP problem. State and prove optimality condition for such a problem.

(8+12+12+8) = [40]

2.(a) Explain clearly inventory control problem. What is Economic Order Quantity (EOQ) ? Give its derivation.

(b) What is safety stock ? Derive a formula to determine safety stock when lead time is fixed but demand during the dead time follows normal distribution.

(c) Explain A,B,C analysis and give a procedure to obtain ordering table for the C class of items when the annual demand is constant and is equal D units, cost of ordering is Rs. A per order, unit cost is Rs. C and inventory carrying cost is "a" .

(10+10+15) = [35]

3.(a) Explain the machine interference problem ? Out line the procedure to obtain optimal number of service counters when the relevant costs are provided and finite queueing tables are made available.

(b) Explain the meaning of steady state probabilities. Under steady state derive expression for average queue length for the queueing model $M|M|C : (\infty \mid FIFO)$.

(10 + 15) = [25]

:bcc:

1987-88	E263

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Applications : 1987-88

PART - II

FINAL EXAMINATION

Grade 6 : (i) Fundamentals of Sampling Theory
(ii) Elements of SQC] Practical

Date: 20.6.1988

GROUP - A

(1) Fundamentals of Sampling Theory (Practical)

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

Note: This paper carries 100 marks. Marks for each question are given at the end of the question within square Brackets.

1. An experienced farmer makes an eye estimate of the weight of peaches x_i on each tree in an orchard of N = 200 trees. He finds a total weight of X = 11,600 units. The peaches are picked and weighed on a simple random sample of 10 trees, with the following results.

		Tree Number										
		1	2	3	4	5	6	7	8	9	10	Total
Actual weight	y_i	61	42	50	58	67	45	39	57	71	53	543
Estimated weight	x_i	59	47	52	60	67	48	44	58	76	58	569

Estimate the total actual weight Y using.

(a) Regression method

(b) Ratio method

(c) Simple mean estimate

Estimate the Mean Square Error/variances of the estimates obtained in (a), (b) and (c) above.

Compare the above estimates in light of your results obtained above. [50]

p.t.o.

2. The following data relate to the population of 112 villages, divided into two strata. In first stratum, a PPSWR sample of size 3 has been drawn, the sizes being the area under cultivation (x) for a particular crop. In second stratum, a SRSWOR sample of size 8 has been drawn.

Stratum No.	N_h	n_h		Yield (y) and area under cultivation (x) for sampled villages
1	32	3	y :	248 512 364
			x :	52 104 74
2	80	8	y :	74 86 68 74 39 46 64 78

For first stratum $X_1 = \Sigma X_{1i} = 2114$.

(a) Estimate the average yield for the population.

(b) Estimate the sampling Error of your estimate.

(c) If for a future survey, you wish to allocate a total sample of size 20 optimally in the two strata, what sample sizes would you recommended.

[50]

GROUP - B

(ii) Elements of SQC (Practical)

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

Note: (a) Attempt all questions.
(b) Figures in the margin indicates full marks.
(c) Practical Assignments are allowed.

1. The follow' _ data on a process at final inspection are available:

Sample number	Sample size	Number defective	Sample number	Sample size	Number defective
1	200	6	11	100	1
2	250	8	12	100	0
3	250	9	13	100	1
4	250	7	14	200	4
5	200	3	15	200	5
6	200	4	16	200	3
7	150	2	17	200	10
8	150	1	18	200	4
9	150	0	19	250	7
10	150	2	20	250	6

Contd...., 3/-

Contd..... Q.No.1

(a) Set up an appropriate control chart for the data and find the trial control limits.

(b) Design a control chart for future production control with an appropriate sample size such that the lower control limit is positive.

(12+8) = [20]

2. The data below represent the number of nonconformities per 1000 meters of telephone cables.

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

(a) Set up a control chart for the data and obtain the trial control limits. Would you conclude that the process is in statistical control ? What control procedure would you recommend for future production ?

(b) Suppose a new inspection unit is defined as 2500 meters of cable.

(i) What are the center line and control limits for a control chart based on the total number of defects in the new inspection unit ?

(ii) What are the center line and control limits for a control chart for defects per unit used to monitor future production ?

(10+5+5) = [20]

3. A sample of 5 units is taken from a process every half hour. It is known that the process standard deviation is in control with $\sigma = 2.0$. The $\bar{x}$ values for the last 20 samples are

Contd..... 4/-

Contd..... Q.No.3

Sample number	$\bar{x}$	Sample number	$\bar{x}$	Sample number	$\bar{x}$
1	41.5	8	40.2	15	41.8
2	42.7	9	41.4	16	40.7
3	40.5	10	43.9	17	42.8
4	39.8	11	40.6	18	43.4
5	41.6	12	39.4	19	42.0
6	44.7	13	38.6	20	41.9
7	39.6	14	42.5		

(Control Charts need not be drawn).

The specifications on the product are 40 ± 5

(a) Set up a modified control chart on this process. Use 3σ limits on the chart and assume that the largest fraction defective that is tolerable is 0.1%.

(b) Reconstruct the chart in part (a) using 2σ limits. Is there any difference in the analysis of the data ?

(c) Control Charts for $\bar{X}$ and R are maintained on the tensile strength of a metal fastener. After 30 samples of size n = 7 are analysed we find that,

$$\sum_{i=1}^{30} \bar{x}_i = 12,870 \quad \text{and} \quad \sum_{i=1}^{30} R_i = .1350$$

(i) Compute control limits on R-Chart.

(ii) Assuming that R-Chart exhibits control, estimates the parameters μ and σ.

(iii) If the process output is Normally distributed and if the specifications are 440 ± 40, can the process meet the specifications ? Estimate the fraction conforming.

(iv) If the variance remains constant where should the mean be located to minimise the fraction non-conforming ?

(5+5+2+2+4+2) = [20]

4. MIL - STD 105 D is being used to inspect incoming lots of size N = 5000, general inspection level II and an AQL = 0.65%. is being used.

   (a) Find the Normal, Tightened and Reduced inspection plans for single and Double sampling.

   (b) Draw the OC curves for the Normal and Tightened inspection for single sampling plans on the same graph.

   (c) While using Normal inspection for lots of size N = 750, at an inspection level II for MIL - STD 105 D, the following are the number of defectives observed in the recommended sample size for the single sampling plan from 10 consecutive lots,

   0, 1, 0, 0, 1, 1, 2, 0, 1, 2

   Does the above result justify switching over to reduced inspection for (i) AQL = 1.0% and (ii) AQL = 2.5% ?

   (6+10+4) = [20]

5. Practical Assignments. [20]

:bcc:

1987-88	E263

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Applications : 1987-88

FINAL EXAMINATION

PART - II

Grade 6 : (i) Fundamentals of Sampling Theory } Theory
(ii) Elements of SQC

Date: 17.6.1988 Maximum Marks: 100 Time : $2\frac{1}{2}$ hours

GROUP - A

Fundamentals of Sampling Theory
(Theory) Max.Marks: 100

Note: Answer any FOUR questions

1.(a) Explain how you would select a random sample of size 3 from a population of size 112 without replacement. [5]

(b) For simple random sampling without replacement write down the unbiased estimator for the population total and an unbiased estimator for the variance of the above estimator.
(2+3) = [5]

(c) Explain Lahiri's method and derive the probability of selection of a unit that you would get using this method.
(4+6) = [10]

(d) Explain what you understand by 'inclusion probability of a unit' and 'inclusion probability of a pair of units' in a sample.
(2+3) = [5]

2.(a) What are the advantages of stratified sampling ? [5]

(b) Derive Neyman's optimum allocation of a total sample size n to strata when simple random sampling without replacement design is used in all strata. Explain how you would use the allocation in actual practice.
(8+4) = [12]

Contd..... 2/-

Contd..... Q.No.2

(c) If the optimum allocations are n_i and the actual allocation applied in practice turns out to be n'_i, obtain an expression for the relative loss due to deviation from the optimum in terms of the allocation. Comment on a bound for this relative loss.

(6+2) = [8]

3.(a) How do you estimate the population total using the ratio method of estimation ? [5]

(b) Derive the bias and mean squared error of the above estimator. (5+7) = [12]

(c) How does this estimator compare with the unbiased estimator which does not use any auxiliary information ? [8]

4.(a) What is Horvitz-Thompson (HT) estimator for a population total ? [4]

(b) Derive the variance of the HT estimator. [7]

(c) Write down the variance estimators due to (i) Horvitz and Thompson and (ii) Yates and Grundy and comment on their relative performances.

(6+6+2) = [14]

5. Write notes on any <u>three</u> :

(a) Systematic sampling

(b) Determination of sample size

(c) Regression estimators

(d) Questionnaires and schedules

(e) Interpenetrating network of subsamples. [25]

GROUP - B

Elements of SQC (Theory)

Max. Marks: 100

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

1.(a) Give the theoretical distribution on which the central line and control limits of the number defective charts are based.

(b) How would you interprete points above UCL and points below LCL in a fraction defective chart ?

(c) State whether the following statements are necessarily true:

(i) When a manufacturing process is under statistical control, all products manufactured satisfy design specifications.

(ii) If the process capability of a manufacturing process is less than the design tolerence, all products manufactured will satisfy design specifications.

(iii) The range chart with usual sample size 4 or 5 is not particularly effective in detecting small shifts in process standard deviation.

(iv) The standard deviation method of estimating process capability is more efficient than the range method for subgroups of size 'n' greater than 10.

(v) Modified control limits in an $\bar{X}$-chart are the limits of its chance variation.

(vi) A double sampling plan gives more protection against rejecting a good lot compared to the corresponding single sampling plan.

(d) Write short notes on any two

(i) Control Chart for individuals

(ii) Group Control Chart

(iii) Modified Control Chart.

(2+2+6+10) = [20]

p.t.o.

2.(a) "In a random assembly stricter control is desirable on components of assembly with wider tolerances than those with narrower tolerances" — Justify.

(b) An assembly of two parts is formed by fitting a shaft into a bearing. It is known that the inside diameters of bearings are normally distributed with a mean of 2.010 cm. and standard deviation of 0.02 cm., and that the outside diameters of the shafts are normally distributed with a mean of 2.004 cm. and standard deviation 0.001 cm. Determine the distribution of clearance between the parts if random assembly is carried out. What is the probability that the clearance is positive ?

(c) A machine fixes caps on bottles by applying an average torque of 8 units with a standard deviation of 1 unit. The bottles can stand a torque of 10 units on an average with a standard deviation of 2 units. Assuming normality and independence of distributions, estimate the percentage of breakage.

(5+10+5) = [20]

3.(a) Explain what is meant by process capability of a manufacturing process.

(b) Assuming that a process average can be adjusted at the desired level, derive an expression for the optimum setting of the process mean to minimise the total cost of rejection and rework due to undersize and oversize when the process capability is not adequate to meet the design tolerance.

(c) The specification limits for a quality characteristic are 41.0 ± 5.0. The process standard deviation as estimated from the control chart data maintained on the process is 2.50. If an item exceeds the upper specification limit, it can be reworked at an additional cost of Rs.2.50 per piece, and if it is below the lower specification limit it is scrapped incurring a loss of Rs.4.50 per piece. At what average the process should be controlled for future production to minimise the total cost of rework and rejection.

(5+10+5) = [20]

4.(a) Define the terms,

(i) AQL, (ii) LTPD, (iii) Producers Risk and (iv) Consumer's Risk.

(b) Give exact and approximate expressions for Type B OC function for a single sampling lot by lot inspection.

(c) Discuss briefly the following aspects of Military standard 105 D :

(i) Inspection level,
(ii) Relation between lot size and sample size,
(iii) Relation between sample size, acceptance number, and AQL
(iv) Producer's Risk at AQL.

(8+4+8) = [2]

5.(a) Derive an expression for item by item sequential sampling plan using the cumulative number of defective items as the criterion for taking a decision at each stage of sampling.

(b) Construct an item by item sequential sampling plan with $p_1 = 0.02$, $p_2 = 0.08$, $\alpha = 0.005$, $\beta = 0.10$.

(c) An inspector tests 40 units without finding a single defective item. Under the given plan would he have decided to accept the lot before reaching the 40th unit, or would he have to continue further.

[given that, log 2 = 0.301030 ; log 3 = 0.477121
log 9.2 = 0.963788 ; log 9.5 = 0.977724
log 9.8 = 0.991226]

(10+4+6) = [20]

6.(a) Describe how you will design a single sampling plan when AQL, LTPD and the associated risks are specified.

(b) Derive an expression of ASN for a doubling sampling plan by attribute for complete inspection of the second sample.

(c) State the conditions for switching over to Reduced inspection from Normal inspection, and from Reduced inspection to Normal inspection.

(6+6+8) = [20]

:bcc:

INDIAN STATISTICAL INSTITUTE 1987-88 E272
One Year Evening Course in Statistical Methods and Applications : 1987-88

PERIODICAL EXAMINATION

PART - II

Grade 7 : Survey Sampling and Organisation Aspects and Elements of OR

Date: 9.5.1988 Time: 2 hrs.

Note: Answer the TWO groups in separate answerscripts.

GROUP - A

Survey Sampling and Organisation Aspects

Maximum Marks: 100

Note: Answer any FOUR questions. Each question carries equal marks. Maximum marks for each subdivision of a question have been indicated in the margin.

1. A linear unbiased estimator of population total based on a two stage sampling design is given by

$$\hat{Y} = \sum_{i \in s} b_{is} \hat{Y}_i$$

When $\hat{Y}_i$ is an estimate of ith fsu - total. Find an expression for variance $V(\hat{Y})$ and $v(\hat{Y})$ an unbiased estimator for $V(\hat{Y})$.

Hence or otherwise find variance and unbiased variance estimator of $Y_{HT(a)}$, a two-stage Horvitz - Thompson estimator of population total.

(16+9) = [25]

2.(a) In two stage sampling with selection of a single first stage unit with equal probability, obtain two different estimators of population total and obtain their MSE's (variances). Also obtain the variance of the estimator when the single first stage unit is selected with probability proportioned to the size of the first stage unit.

(b) Compare the proformances of the above estimators with respect to the following aftificial population.

Contd..... 2/-

Contd..... Q.No.2.(b)

Unit	y_{ij}	M_i
1	0, 1	2
2	1, 2, 2, 3	4
3	3, 3, 4, 4, 5, 5	6

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

3.(a) In two stage sampling with units of unequal first stage units and sampling at each stage with SRSWOR, consider the following estimator of population the

$$\hat{Y} = \frac{\sum_1^n M_i \bar{y}_i}{\sum_1^n M_i} \cdot M, \qquad M = \sum_1^n M_i$$

Obtain MSE of $\hat{Y}$ and its estimator. When is the estimator Y self-weighting ? If M is not known how would you estimate it ?

(b) The following data give the age of scientists selected randomly from each of randomly selected 5 pages of the "American Men of Science". Obtain $\hat{\bar{Y}}$, the average age of a scientist as listed in the above book and an estimator of its MSE.

Unit No.	M_i	Ages	
1	15	47	30
2	19	38	51
3	19	43	45
4	16	55	41
5	16	59	45

[(5+5+3+2) + 10] = [25]

4.(a) In two stage sampling where the first stage units are selected with equal probability and second stage units by SRSWOR obtain ratio estimator of population total Y by using the auxiliary information as supplied by another variable on the units in the population. Obtain its MSE and its estimator.

(b) How would you find the ratio estimator of population total when the fsu's are selected with prob. p_i and with replacement. Indicate expresions of its MSE and its estimator.

(3+6+6+5+5) = [25]

5.(a) In two stage sampling from a population with fsu's of equal sizes, when the units are selected with equal probability and without replacement at each stage with a constant second stage sampling fraction obtain the optimum values of m and n assuming a linear cost function.

(b) For a survey of the above type one has the following:

(i) Cost of enumerating a psu = Rs.100/-.

(ii) Cost of enumerating a ssu = Rs. 12/-.

(iii) $\frac{1}{N(M-1)} \sum_i \sum_j (Y_{ij} - Y_i)^2 = 351.8.$

(iv) $\frac{1}{M-1} \sum_{i=1}^{N} (\bar{Y}_i - \bar{\bar{Y}})^2 = 498.3$

Obtain the optimum value of m.

(c) Considering two-stage sampling from a population with unequal sizes of fsu's and selection at each stage with SRSWOR and, a constant second-stage sampling fraction and the estimater $\hat{\bar{Y}}$ as in Q.3(a), obtain optimum value of m assuming a linear cost fn, that incorporates the constant cost of listing a ssu, a constant cost of enumeration of fsu and a constant cost of enumeration of ssu.

GROUP B

Elements of OR (Theory and Practical)

Maximum Marks : 100

Note: Answer all the questions. The answer should be brief and to the point. The marks allotted to each questions are provided in [ ] in the margin.

1.(a) Consider the queueing model M/M/1 : (∞|FIFO). Derive the following:

(i) waiting time distribution

(ii) average waiting time of a customer

(iii) average waiting time of an arrival who has to wait.

Contd..... 4/-

Contd..... Q.No.1

(b) A transport unit runs a fleet of 20 cars to meet the demand of various departments of an Institute. The demand for one of the cars for field work is treated separately. Number of requisitions for this car arrive with mean rate of 2.5 per day but only 3 requisitions on an average can be taken up for service each day. Unfulfilled demand can be taken up the next day. Analyse the waiting line problem and comment in respect of average number of cases which wait for service. Suggest changes for improving the service explaining criterion to measure the efficiency of the system.

(15+10+10+10) = [45]

2.(a) Describe the problem of machine interference using the terminologies of M/M/C ; (N/FIFO) queueing system. Give an illustration either from a transport organisation or a printing press which ever you have visited recently.

(b) At a workshop there are 12 machines in operation. From past experience it is known that these machines will operate for an average of 50 hours, then require an average of 30 hours repair. Each machine will yield a profit of Rs.120 per productive hour and the cost of servicing the machine is Rs.60/- per hour per service man. It is given that the cost per serviceman continues whether are not he is actually working. Treat this as machine interference problem and obtain optimal number of serviceman maximising the net profit.

(c) Obtain optimal number of servicemen maximising the net profit in the problem stated in (b) under following changed situation.

A preventive maintenance costing Rs.120/- per hour for the total of twelve machines is introduced resulting in reducing the service time from 30 hours to 20 hours which increasing the production time from 50 hours to 80 hours.

(5+20+20) = [45]

Practical assignments. [10]

:bcc:

1987-88	E262

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Applications : 1987-88

PERIODICAL EXAMINATION

PART - II

Grade 6 : Fundamentals of Sampling Theory
and
Elements of SQC

Date: 25.4.1988 Time: 2 hrs.

GROUP A

Fundamentals of Sampling Theory

Maximum Marks: 100

Note: Answer ALL questions.

1.(a) What are the advantages of stratified sampling ? [8]

(b) A population of size 112 is divided into two strata of sizes 80 and 32. The table below gives the details of the sampling procedure:

Stratum No.	Stratum size	Sample size	Sampling Design	Y - values
1	80	8	SRSWOR	148,164,214,198,173,201, 186,194
2	32	4	PPSWR	y_i: 456, 349, 601, 218 x_i: 98, 64, 127, 43

From the first stratum a sample of size 8 is chosen by Simple Random Sampling Without Replacement (SRSWOR) Design and the y-values of a study variable are given in Col.(5). From the second stratum a probability proportional to size (PPS) sample is chosen with replacement (WR) where the size is the auxiliary information on the variable x. It is also known that the total size of x_i values in stratum 2 is 2919.

(i) Estimate the population mean. [20]

(ii) Estimate the sampling error of your estimate obtained in (i) above. [32]

p.t.o.

2.(a) Derive Neyman's optimum allocation of a total sample size n to strata when simple random sampling without replacement design is used in all strata. [10]

(b) How do you obtain this optimum allocation in practice ? [5]

(c) Calculate the variance of the estimate of the population mean under this allocation and obtain an expression for the gain due to this allocation compared to proportional allocation. (10+10+5) = [25]

GROUP B

Elements of SQC (Theory and Practical)

Maximum Marks: 100

Note: Question no.1 is compulsory and any TWO from the remaining questions to be attempted.

1. Draw the Type B OC curves for the following single sampling plans in the same background (same coordinate axes) using Poisson approximation.

(i) n = 50 ; c = 1
(ii) n = 100 ; c = 1
(iii) n = 100 ; c = 2
(iv) n = 200 ; c = 2

Comment on the behaviour of the respective plans in discreminating bad lots from good ones. (10 x 4) = [40]

2. A company has the following plan. It takes a sample equal to 10% of the lot being inspected. If 1% or less of the sample is defective, the lot is accepted. If more than 1% of the sample is defective, the lot is rejected.

(a) If a lot contains 4000 items of which 80 are defective what is the probability of its being accepted under the plan ?

Contd..... 3/-

Contd..... Q.No.2

(b) If a lot contains 1000 items of which 20 are defective what is the probability of its being accepted under the plan ?

(c) If the lot contains 10,000 items of which 200 are defective what is the probability of its being accepted.

(d) If 0.02 is the lot tolerence percent defective, what can you say about the protection given by the above plan ?

$(7\frac{1}{2} \times 4) = [30]$

3.(a) A control chart for fraction nonconforming is to be established using a centerline $p' = 0.10$. What sample size is required if we wish to detect a shift in the process fraction nonconforming to 0.16 with a probability of 0.50 ?

(b) A certain product is manufactured to a specification of 120 ± 5 for one of the quality characteristics. At present the estimated process average is 120.0 and the standard deviation is 1.50.

(i) Compute the 3-sigma control limits for an $\bar{X}$ and R chart based on a subgroup size of 4.

(ii) What percentage of the product does not meet the specification under the present condition ?

(iii) What percentage of the product will not meet the specifications if the process average suddenly shifts to 121.0, assuming normally distributed product characteristics.

(10+8+6+6) = [30]

4.(a) Control charts for $\bar{X}$ and R are maintained on the tensile strength in pounds of a certain yarn. The subgroup size is 5. The values of $\bar{X}$ and R are computed for each subgroup. After 25 subgroups, $\Sigma \bar{X} = 514.8$ and $\Sigma R = 120.0$. Compute the values of 3-sigma limits for the $\bar{X}$ and R charts, and estimate the value of σ' on the assumption that the process is in statistical control.

Contd..... 4/-

Contd..... Q.No.4

(b) There is a single specification limit (lower limit) L = 15 lbs. for the tensile strength of yarn in problem (a). Make the necessary calculations based on the control chart data given in (a) to determine whether or not you would expect this product to meet the specification. If not, approximately what percentage would you expect to find below the lower specification limit. Assume that the normal distribution is applicable.

(c) A fraction nonconforming chart is designed with $p' = 0.08$ and a sample size of 100. Draw the OC curve of the control chart and comment on its efficiency.

(10+5+15) = [30]

:bcc:

1987-88	E252

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods and Applications : 1987-88

PERIODICAL EXAMINATION

PART - II

GRADE - 5 : Estimation, Linear Estimation and Design of Experiments.

Date: 18.4.1988 Maximum Marks: 100 Time: 2 hours

Note: Answer the THREE groups in separate answerscripts.

GROUP - A

Point and Interval Estimation

Max. Marks: 30

Note: Do all questions. Your maximum score can not exceed 30 marks.

1. Show that mean $\bar{x}$ of a random sample drawn from a normal population $N(\mu, \sigma^2)$ is unbiased, consistent, sufficient, MLE, hits C-R bound and is UMVUE for μ.

[15]

2. Define log-normal distribution and obtain MLEs for its mean and variance.

[15]

3. Give examples for the following:

   (i) An estimator may be unbiased but not consistent and vice-versa.

   (ii) An MLE may not be unbiased.

   (iii) A sufficient statistic need not be unbiased.

   (iv) An MLE need not be consistent.

   (v) Unbiased estimator may not always exist.

[5]

p.t.o.

GROUP - B

Linear Estimation

Max. Marks: 30

Note: Answer all the questions.

1.(a) Distinguish between regression model and experimental design model with reference to general Gauss-Markov linear model.

(b) What is reparamatrization of a linear model ? Explain why this technique is useful in the analysis of a linear model.

(5+5) = [10]

2. To study the effect of drug dosage in milligrams on a person's reaction time in milliseconds, 15 patients were randomly assigned one of five dosage of a drug — 0.5, 1.0, 1.5, 2.0, 2.5 mg — with three patients assigned to each level of drug dosage. The reaction times were recorded. An analysis of variance of these data yields the following table:

Source	d.f.	S.S.	M.S.	F	Level of significance 5%
Linear	-	192.53	-	-	-
Deviation from linear	-	-	-	-	-
Between dosages	-	-	-	-	-
Error	-	-	2.00	-	-
Total	14	249.73	-	-	-

(a) Complete the ANOVA table.

(b) Test whether there is a significant effect of dosage on reaction time.

(c) If (b) is found to be true, test whether the regression of reaction time on dosage is linear or not.

(10+5+5) = [20]

GROUP - C

Design of Experiments

Max. Marks: 40

Note: Answer all the questions.

1. Explain the terms

   treatment, experimental error in the context of designs of experiment.

   (7+8) = [15]

2. Describe briefly a randomised block design. How would you analyse the data obtained from such a design ?

   (10+15) = [25]

:bcc:

1987-88	E271

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods and Applications : 1987-88

PERIODICAL EXAMINATION

PART - II

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

(b) Elements of OR (Theory and Practical)

Date: 4.4.1988

GROUP - A

Survey Sampling and Organisation Aspects (Theory and Practical)

Maximum Marks: 100 Time: 2 hrs.

Note: Answer any FOUR questions. Each question carries equal marks. Maximum marks allotted for each subdivision of a question has been indicated in the margin.

1.(a) What are the different steps in planning and organising a large scale sample survey.

(b) The presence or absence of each of two characteristics is to be measured on each unit in a simple random sample from a large population. If P_1, P_2 are the percentages of units in the population that possess characteristics 1 and 2, a client wishes to estimate (P_1-P_2) with a standard error not exceeding 2 percentage points. What sample size do you suggest if the client thinks that both P_1 and P_2 should lie between 40% and 60% and that the characteristics are independently distributed on the units ?

(15+10) = [25]

2. In a two stage sampling design the first stage units (fsu) are selected with probability proportional to size of the fsu's and the second stage units with simple random sampling without replacement. Find an unbiased estimator of population total and derive its variance.

Contd..... 2/-

Contd..... Q.No.2

Also find an unbiased estimator of this variance. When is this design said to be self-weighted ?

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

3.(a) In a sample survey using a two stage design in which 5 villages, which are first stage units, are selected with probabilities proportional to their last census populations out of 80 villages and the households, which are second stage units, are selected by circular systematic sampling procedures, the following data are obtained on agricultural area possessed by each sampled household. Obtain an unbiased estimate of total agricultural area possessed by all the households in the population together with an estimate of its variance.

village sl. no.	prob. of selection (P_i)	Total no. of house-holds	area of agricultural land based possessed by each household (acres)				
3	.01	64	1.2	1.3	0.7	1.8	
5	.03	82	2.7	3.2	1.9	0.4	3.2
37	.11	96	1.8	2.1	0.7	3.1	1.5
41	.02	88	2.1	1.6	1.3	1.8	0.1
73	.07	125	1.4	1.6	2.3	0.9	1.8 2.1

(b) What is an interpenetrating subsampling technique. Discuss its usefulness at estimation stage in a sample survey.

(16+9) = [25]

4.(a) In a two stage sampling n first stage units (fsu's) are selected by simple random sampling without replacement (srswor) out of N fsu's each containing M second stage units (ssu's) and m ssu's are selected from each selected fsu by srswor. Compare the efficiency of this procedure with respect to a cluster sampling procedure in which k clusters (fsu's) are selected by srswor, $K = \frac{nm}{M}$.

(b) A linear estimator of population total based on a two stage design is

$$\hat{Y} = \sum_{i \in s} b_{is} \hat{Y}_i$$

Contd..... 3/-

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

When b_{is} is the coefficient attached to unit i in the first stage sample s and $\hat{Y}_i$ is an estimator of ith fsu total. Find conditions required for unbiasedness of $\hat{Y}$. Find an expression for $V(\hat{Y})$.

(16+9) = [25]

5. In a survey of milk production in Calcutta all the Cowsheds are divided into 4 strata according to the number of milch animals in the shed and simple random samples of cowsheds are drawn from each stratum. The following table gives the production of milk in a specific week by the sampled Cowsheds. Find an estimate of the total quentity of milk produced in the week and an estimate of its variance. Also on the basis of the stratified sample thus selected estimate the gain in stratification over simple random sampling.

Stratum	No. of Cowsheds	Milk produced by the sampled cowsheds ('00 litres)					
1	182	3.5	2.9	7.2	3.6	2.1	
2	237	4.2	6.1	8.3	7.9	5.8	6.3
3	148	9.2	11.5	16.9	12.3	16.9	
4	92	23.2	27.9	32.3			
	829						

[25]

p.t.o.

GROUP - B

Elements of OR (Theory and Practical)

Maximum Marks: 100 Time: 2 hours

Note: Answer all the questions. The answer should be brief and to the point. The marks allotted to each questions are provided in [ ] in the margin.

1. Explain clearly the meaning of a basic feasible optimal solution to a linear programming problem (LPP) using usual symbols and notations. State and prove optimality condition for a basic feasible solution to LPP with maximisation of a linear function of basic variables.

(5+10) = [15]

2. A fire work firm is allowed to manufacture 1000 kg. of explosives altogether per month for two varieties of explosive E_1 and E_2. But raw material availability restricts the manufacture of type E_1 and E_2 explosives to a maximum of 600 kg. and 800 kg. respectively. The firm makes profit at the rate of Rs.2000/- and Rs.1500/- per metric tones of explosives E_1 and E_2 respectively

   (a) Formulate the problem as linear programming problem maximising the profit and sketch the problem on a suitable graph sheet indicating the feasible region and hence obtain the optimal solution.

   (b) Suppose monthly ceiling on the manufacture of explosives is raised to 1500 kg.. Will the optimal solution in (a) change ? If so, find out the new optimal solution and indicate the total monthly profit of the firm.

(20+5) = [25]

3. Solve the following linear programming problem using simplex method :

$$\text{Maximise } z = 5x_1 + 7x_2$$

$$\text{Subject to } 2x_1 + x_2 \leq 5$$

$$3x_1 + 4x_2 \leq 10$$

$$x_1, x_2 \geq 0$$

[30]

p.t.o.

4.(a) Maximise $z = c_1x_1 + c_2x_2 + c_3x_3$

Subject to $a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \leq b_1$

$x_1, x_3 \; x_3 \geq 0$

the symbols have usual meaning and $c_j < 0$ for $j = 1, 2, 3$. Write down an optimal solution stating your assumption.

(b) Explain the notation M/M/C : (∞ | F | FO) in waiting line models. Under steady state using balance equations obtain expression for the probability that there are n customers in the system under the model M/M/1 : (∞ | F | FO).

(c) The service time of an enquiry clerk and arrival distribution of the calls to the clerk through a telephone operator follow M/M/1 : (∞ | F | FO) model. If enquiry clerk is busy the operator asks the caller to wait and when the clerk becomes free the operator transfers the longest waiting call to him.

Calls arrive at the rate of 12 per hour and the enquiry clerk can service a call in four minutes, on the average obtain (i) average number of calls waiting to be attended, (ii) average waiting time before the call is attended by the enquiry clerk.

(5+15+10) = [30]

:bcc:
29388

1987-88	E261

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Applications : 1987-88

PERIODICAL EXAMINATION

PART - II

Grade 6: (a) Fundamentals of Sampling Theory (Th. and Pr.)
(b) Elements of SQC (Theory and Practical)

Date: 28.3.1988

GROUP - A

Fundamentals of Sampling Theory
(Theory and Practical)

Maximum Marks: 100 Time: 2 hours

Note: Answer ALL questions.

1.(a) Write down (no proof required) an unbiased estimator of the population total of a study variate based on a sample selected using simple random sampling without replacement (srswor) scheme. Also give expressions for the variance of the above estimator and an unbiased estimator of this variance.

(3+5+5) = [13]

(b) From a population of 248 households 25 households were selected by srswor scheme of which 8 possess a colour T.V. Estimate the proportion of colour T.V.'s in the population and estimate its sampling error.

(6+10) = [16]

2.(a) Explain with a simple illustration what you understand by 'Lahiri's method' of selection. Deduce that the method gives a probability of selection proportional to size for each unit.

(8+10) = [18]

(b) From a population of 78 factories 6 factories were selected with probability proportional to the number of workers in the factory. Data on the output is given in the table below for these selected factories. It is also know that the total number of workers in the population is 5862.

Contd..... 2/-

Contd..... Q.No.2.(b)

Sl. No.	No. of workers	Output
1	124	1420
2	68	710
3	34	692
4	110	1206
5	46	510
6	71	787

(i) Estimate the average output per factory in this population. [12]

(ii) Give an unbiased estimate of the sampling error of your estimate in (i) above. [18]

3.(a) What are the differences between linear and circular systematic sampling ? [12]

(b) Explain how you would estimate the gain due to ppswr sampling over srswr sampling of the same size. [11]

GROUP - B

Elements of SQC (Theory and Practical)

Maximum Marks: 100 Time: 2 hours

Note: Question number 1 is compulsory. Answer any THREE from the remaining questions.
Figures in the margin indicate full marks.

1.(a) Discuss the concepts of chance and assignable causes of variation and the part they play in Statistical Quality Control.

(b) What is meant by the statement that a process is in a state of statistical control.

Contd..... 3/-

Contd..... Q.No.1

(c) The following data were obtained for a 10 day period to initiate $\bar{X}$ and R control charts for a quality characteristics that required a substantial amount of rework.

The specification for this characteristic is given as 171 ± 11. This subgroup size was 5, and all the figures apply to products made on a single machine by a single operator.

(i) Draw a suitable control chart and obtain the trial control limits. Draw your own conclusions about the process from the analysis of the data.

(ii) If the $\bar{X}$ - R charts are to be continued for future production, suggest suitable limits.

Sub-group No.	Average $\bar{x}$	Range R	Sub-group No.	Average $\bar{x}$	Range R
1	177.6	23	11	179.8	9
2	176.6	8	12	176.4	8
3	178.4	22	13	178.4	7
4	176.6	12	14	178.2	4
5	177.0	7	15	180.6	6
6	179.4	8	16	179.6	6
7	178.6	15	17	177.8	10
8	179.6	6	18	178.4	9
9	178.8	7	19	181.6	7
10	178.2	12	20	177.6	10

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

2. A manufacturer wishes to setup a control chart at the final inspection for a gas water heater. Defects in workmanship and visual quality features are checked in this inspection. For the last 22 working days 176 water heaters were inspected and a total of 924 defects reported.

(a) What type of control chart would you recommend here and how would you use it ?

(b) Using two water heaters as inspection unit, calculate the center line and control limits that are consistent with the past 22 days inspection data.

Contd..... 4/-

Contd..... Q.No.2

(c) What is the probability of Type I error for the control chart in part (b) ?

(6+6+8) = [20]

3. The number of defects found on final inspection of a cassette deck is shown below.

Deck No.	No. of defects	Deck No.	No. of defects	Deck No.	No. of defects
1	0	7	1	13	2
2	1	8	3	14	5
3	1	9	2	15	1
4	0	10	1	16	2
5	2	11	0	17	1
6	1	12	3	18	1

(a) Can you conclude that the process is in Statistical Control ? What center line and control limits would you recommend for controlling future production ?

(b) What are the center line and control limits for a control chart for future production based on total number of defects in an inspection unit of four cassette decks ?

(c) Draw an OC curve for the control scheme in part (a)

(5+5+10) = [20]

4. Data of an optical company are given below, each article, 39 per box, inspected being either good or broken.

(a) Make an appropriate control chart for the data and comment.

(b) Why do you think the sample size of n = 39 was used ?

(c) Using a Poisson approximation and $n\bar{p}$ as c', estimate the probability of a box having <u>less</u> than 36 good articles.

Contd..... 5/-

Contd..... Q.No.4

Sample No.	No. broken	Sample No.	No. broken	Sample No.	No. broken
1	2	16	1	31	6
2	1	17	0	32	1
3	1	18	1	33	2
4	0	19	0	34	2
5	2	20	1	35	2
6	3	21	0	36	1
7	2	22	2	37	1
8	1	23	0	38	0
9	1	24	1	39	2
10	2	25	1	40	2
11	1	26	0	41	2
12	1	27	1	42	5
13	1	28	0	43	2
14	2	29	7		
15	1	30	10		

(10+5+5) = [20]

5. (a) A control chart indicates that the current process fraction defective is 0.02. If 50 items are inspected each day, what is the probability of detecting a shift in the fraction defective to 0.04 on the first day after the shift ? By the end of the third day following the shift ?

(b) A company purchases a small metal bracket in containers of 5000 each. Ten containers have arrived at the unloading facility, and 250 brackets are selected at random from each container. The fraction non-conforming in each sample are,

0, 0, 0.004, 0.008, 0, 0.020, 0.004, 0, 0, and 0.008

Does the data from this shipment indicate statistical control ?

(10+10) = [20]

:bcc:

1987-88 | E253

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods and Applications : 1987-88

PERIODICAL EXAMINATION

PART - II

Grade 5 : Estimation, Linear Estimation

Date: 21.3.1988 Maximum Marks: 100 Time: 2 hrs.

Note: Answer the TWO groups in separate answerscripts.

GROUP - A

Point and Interval Estimation. Max. Marks: 50

Note: Do as many questions as you can. Your maximum score for Group A can not exceed 50 marks.

1. (i) State the problems of point estimation and interval estimation.

(ii) Name the desirable properties of a good estimator.

(5+5) = [10]

2. Let $X_1, \ldots, X_n$ be i.i.d. r.v.s. with finite mean μ and variance σ^2. Show that

(a) The class of estimators $D = \{ d(x) \}$, $d(x) = \sum_{i=1}^{n} a_i x_i$ is unbiased for μ iff $\sum_{i=1}^{n} a_i = 1$.

(b) The best choice of weights $a_1, \ldots, a_n$ is $a_i = 1/n$ for all i.

(c) Hence or otherwise show that $\bar{X} = \frac{1}{n} \sum_{1}^{n} X_i$ is BLUE for μ.

[10]

3.(a) Define a sufficient statistic.

(b) State Fisher-Neyman factorization theorem for sufficieney.

(c) Obtain sufficient statistics for (μ, σ^2) on the basis of a random sample from $N(\mu, \sigma^2)$.

(5+5+5) = [15]

p.t.o.

4. (i) State Maximum likelihood method of estimation.

   (ii) List down the properties of MLES.

   (7+3) = [10]

5. (i) State Cramer-Rao inequality for variance of an estimator.

   (ii) Let P and λ be the parameters of point binomial and Poisson distribution respectively and μ the mean of a normal population. Obtain unbiased estimators for p^2, λ^2 and μ^2 based on the random samples $X_1, \ldots, X_n$ from corresponding populations.

   (5+10) = [15]

GROUP - B

Linear Estimation Max. Marks: 50

Note: Answer all the questions.

1. Either,

   (a) Define a parametric function. Under what condition a parametric function is said to be linearly estimable ? State Ganss-Markov theorem of linear estimation.

   (b) What is linear model ? State clearly Ganss-Markov linear model.

   Describe the method of obtaining the best linear unbiased estimators of the parametric functions of the linear model in the contest of one-way classified data.

   Or, (10+15) = [25]

   What do you mean by two-way classified fixed effects linear model ?

   Describe the detail analysis of this model when there are more than one but equal number of observations per cell.

   (5+20) = [25]

2. Below are given the yields (in gm.) per plot for three varieties of seed cotton.

Contd..... 3/-

Contd..... Q.No.2

Variety 1	Variety 2	Variety 3
77	109	46
70	106	70
63	137	71
84	79	65
95	134	61
81	78	40
88	126	47
101	98	73

(a) Write down the appropriate linear model stating the assumptions.

(b) Test at 5% level, if the varieties differ significantly among themselves.

(c) If the result of (b) is affirmative, determine which varieties differ.

(3+16+6) = [25]

:bcc:

1987-88 | E123S

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Applications : 1987-88

PART - I

SUPPLEMENTARY EXAMINATION

Probability and Sampling Distributions

Date: 7.3.1988 Maximum Marks: 100 Time: 2 hours

Note: You may answer any part of any question.

1. If n dice are thrown at a time, what is the probability of having each of the points 1, 2, ..., 6 at least once?

[20]

2. Let the density function $f(x,y)$ of X and Y be given by

$$f(x,y) = \begin{cases} 3x & \text{if } 0 \le y \le x \le 1; \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find the marginal densities of X and Y. (7+7) = [14]

(b) Find the marginal distribution functions of X and Y. (5+5) = [10]

(c) Find the conditional densities. (5+5) = [10]

(d) Are X and Y independent ? Supply reasons. [3]

(e) Find $E(X \mid Y = y)$. [3]

3.(a) When is a random variable X said to follow a gamma distribution ? [5]

(b) Find the mean and variance of X. (7+8) = [15]

(c) Find the mean and variance of the Chi-square distribution with n degrees of freedom. (2+3) = [5]

p.t.o.

4. Let X, Y be two independent random variable each following the uniform distribution over the interval (0,1). Find the density function and the distribution function of (a) XY (b) $|X-Y|$.

[(13+7) + (13+7)] = [40]

5. If X follows the Poisson distribution with parameter λ, show that

$$\mu_{r+1} = \lambda(r\mu_{r-1} + \frac{d\mu_r}{d\lambda}), \quad r \geq 1$$

where μ_r is the r-th central moment of X. [15]

6. Find the mode of a binomial distribution. [10]

:bcc:

1987-88	E144

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods and Applications : 1987-88

PART - I

FINAL EXAMINATION

Applied Statistics (Practical)

Date: 20.1.1988 Maximum Marks: 100 Time: 2 hours

Note: This paper carries 110 marks. Do as much as you can. Marks for each question are given under the questions. Marks for lack of neatness in calculation will be deducted suitably.

Group - A

(Sample Surveys)

1. Following data relate to the distribution of factories by number of workers, average output and standard deviation:

Stratum number	Size classes (number of workers)	No. of factories	Output per factory (in'000 Rs)	Standard deviation in'000 Rs)
1	1 - 49	18260	100	80
2	50 - 99	4315	250	200
3	100 - 249	2233	500	600
4	250 - 999	1057	1760	1900
5	1000 and above	567	2250	2500

A sample of 3000 factories is to be selected with SRSWOR within each stratum. Compare the efficiency of optimum allocation with proportional allocation and unstratified SRSWOR.

[50]

Group - C

(Index Numbers)

2. Compute Laspeyre's , Paasche's, Edgeworth-Marshall's and Fisher's ideal index numbers for the following data for the year 1969-70 taking 1968-69 as base:

Contd..... 2/-

Contd..... Q.No.2

Wholesale Crop-prices (Unit: Rs. per Quintal) in 1968-69 and 1969-70

Year	Rice	Wheat	Jowar	Barley	Maize	Gram
1968-69	119.00	82.56	56.00	55.62	60.58	83.42
1969-70	111.67	95.42	56.00	61.40	55.84	101.31

Crop-production (Unit: Thousand Metric tons)

Year	Rice	Wheat	Jowar	Barley	Maize	Gram
1968-69	39761	18651	9804	2424	5701	4309
1969-70	40430	20093	9721	2716	5674	5546

[35]

Group - D

(Time Series Analysis)

3. Following data give the production of Coal in India for a number of years. Fit a Quadratic trend to the data:

Year (t)	1959	1960	1961	1962	1963	1964	1965
Production (y) (000 metric tons)	47800	52593	56065	61370	65956	62440	67162

[25]

:bcc:

1987-88	E133

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods and Applications : 1987-88

PART - I

FINAL EXAMINATION

Applied Statistics

Date: 18.1.1988 Maximum Marks: 100 Time: 2 hours

Note: Answer the THREE GROUPS in separate answerscript.

GROUP - A (Sample Surveys)

Maxi. Marks: 30

Note: Do as many questions as you can from this group. The paper carries 40 marks but the maximum you can score in the group is 30 marks. The marks allotted are given at the end of each question.

1. (i) Define proportional and Neyman allocations in simple stratified sampling.

(ii) Give an unbiased estimator for population mean based on stratified sampling.

(5+5) = [10]

2. In case of SRSWOR, where sample of size n is selected from a population of N units, show that

(i) sample mean $\bar{y} = \frac{1}{n}\sum_{1}^{n} y_i$ is unbiased for population mean.

(ii) variance of $\bar{y}$ is given by $V(\bar{y}) = \frac{N-n}{nN} S^2$

where $S^2 = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \bar{Y})^2$, $\bar{Y}$ = population mean.

(5+10) = [15]

3. (i) Define two stage sampling and give an unbiased estimator for population total.

(ii) Define systematic sampling and give an unbiased estimator of population mean.

(iii) What do you mean by non-sampling errors. Name some important non-sampling errors.

(5+5+5) = [15]

p.t.o.

GROUP - B (Vital Statistics)

Max. Marks: 20

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

1. Write short notes on the standardisation of death rates. Why the mortality conditions of two places cannot usually be compared on the basis of crude death rates ?

[10]

2. Discuss briefly the different measures of fertility and reproduction.

[10]

3. The following table gives the number of births that occured to a population in a particular year classified by age of mother together with the female population in each age group of the child bearing period. The mid year population for the same year was 2285000.

Calculate:

(a) crude birth rate
(b) general fertility rate
(c) total fertility rate
(d) age specific fertility rates

Age	Female population (00)	Number of births
15 - 19	848	2343
20 - 24	700	14,541
25 - 29	727	16,736
30 - 34	759	10,218
35 - 39	751	5134
40 - 44	716	1422
45 - 49	660	93
Total	5161	50487

[10]

GROUP - C (Index Number and Time Series)

Max. Marks: 50

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

1.(a) Prove that Laspeyre's index is greater than Paasche's index.

(b) Discuss the different steps for constructing a wholesale price index number for India.

(13+12) = [25]

2.(a) Discuss how will you determine trend in a time series by fitting a polynomial of suitable degree.

(b) Write short notes on:

(i) Exponential trend
(ii) Modified exponential trend
(iii) Gompertz curve
(iv) Logistic curve

(13+12) = [25]

(a) Explain what do you mean by method of moving averages. Discuss how this method can be used for determining seasonal indices in a time series.

(b) Discuss how will you obtain measures of seasonal fluctuation after fitting a suitable trend to a time series. State the assumptions clearly.

(13+12) = [25]

:bcc:

1987-88	E133

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods and Applications : 1987-88

PART - I

FINAL EXAMINATION

Tests of Statistical Hypothesis

Date: 15.1.1988 Maximum Marks: 100 Time: $2\frac{1}{2}$ hrs.

Note: Do all the questions. The paper carries 120 marks but your maximum score can not exceed 100 marks. Marks allotted are given at the end of each question.

1. (i) State N-P Lemma

   (ii) Using N-P lemma, obtain Most-Powerful test for testing $H_0 : \mu = \mu_0$ against $H_1 : \mu = \mu_1$ in case of Normal populations $N(\mu, \sigma^2)$, σ being known.

   (5+10) = [15]

2. (i) Define unbiased test.

   (ii) Show that a UMP test is always unbiased.

   (iii) Define UMPU test

   (iv) A random sample from a normal population yielded the observations as

   -10.5, 8.5, 3.2, 1.5, -5.5, -2.2, 10.1, 3.1, 6.5, 7.5

   (a) Can you say that population mean is 1.5 ?

   (b) Obtain 95% confidence interval for mean.

   (5+8+5+12) = [30]

3. A drug manufacturing company invented a new drug for treatment of patients suffering from high blood pressure. After treatment of patients for a specified time the fall in the B.P. for a random sample of patients was recorded as :

   20, 10, 5, -5, 15, 10, 9, 10, 18, 14

   Can you say that the drug is effective ?

   [10]

p.t.o.

4. A random sample of 25 factories producing same item was selected and profit during the last year (x_1); production (x_2) and sales (x_3) were observed. The statistical enquiry yielded the following summary of data (in suitable units)

$\bar{x}_1 = 40$; $\bar{x}_2 = 70$; $\bar{x}_3 = 90$

$s^2_{x_1} = 3$; $s^2_{x_2} = 6$; $s^2_{x_3} = 7$

$r_{12} = 0.4$; $r_{23} = 0.5$; $r_{13} = 0.6$

(i) Test the significance of the population multiple correlation coefficient $R_{1(2,3)}$.

(ii) Test the significance of partial correlation coefficient $\rho_{12.3}$.

(iii) Is it justified to say that correlation between x_1 and x_2 in the population is $\rho_{12} = 0.6$?

(iv) Test whether the regression coefficient of x_1 on x_2 is $\beta_{12} = 0.4$.

(v) Obtain 95% confidence interval for β_{12}.

[25]

5. Three candidates were contesting for an assembly seat from a constituency which was partly urban and partly rural. The opinion of 500 randomly selected persons in favour of different candidates was found as follows:

Ca	Candidate A	Candidate B	Candidate C
Rural	150	65	85
Urban	55	75	70

Is there any association between the nature of constituency and candidature ?

[15]

6. A development scheme was implemented in two areas A and B. After one year two separate samples of 30 and 50 households from A and B respectively were selected and increase in their annual income was found. The enquiry gave following summary of data :

Contd..... 3/-

Contd..... Q. No.6

	A	B
sample size :	$n_1 = 30$;	$n_2 = 50$
mean :	$\bar{x}_1 = 500$;	$\bar{x}_2 = 450$
s.d. :	$s_1 = 11.5$;	$s_2 = 9.0$

(a) Does the statistical enquiry suggest that

(i) scheme is equally beneficial for both areas

or (ii) scheme is more beneficial for A than for B

or (iii) scheme is more beneficial for B than for A

(b) Obtain 95% confidence interval for $\mu_1 - \mu_2$ where μ_1 and μ_2 are average increase in incomes for A and B respectively.

(15+5) = [20]

7. Practical record note Books. [5]

:bcc:

1987-88	E114

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Applications : 1987-88

PART - I

FINAL EXAMINATION

Descriptive Statistics (Practical)

Date: 13.1.1988 Maximum Marks: 100 Time: 2 hours

Note: Attempt all the questions. The marks allotted for each question are given within curly brackets under the questions.

1.(a) The following table shows the distribution of 100 families by their expenditure per week :

Expenditure per week (in Rs.)	0-10	10-20	20-30	30-40	40-50
Number of families	14	?	27	?	15

The mode of the distribution is known to be Rs.25/-. Calculate the missing frequencies and then the median of the distribution.

(b) In a distribution γ_1 and γ_2 values are found to be 0.7018 and -2.3457 respectively. If μ_3 be 39.75, find the fourth central moment of the distribution.

(c) A computer, while calculating correlation coefficient between two variates x and y from 25 pairs of observations, obtained the following results :

$n = 25$, $\Sigma x = 125$, $\Sigma x^2 = 650$, $\Sigma y = 100$, $\Sigma y^2 = 460$,

$\Sigma xy = 508$

It was later found that two pairs were wrongly copied down as

x	y
6	14
8	6

instead of

x	y
8	12
6	8

. Obtain the correct value of correlation coefficient.

(10+7+8) = [25]

p.t.o.

2. In the first proof of a book containing 250 pages, 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

Fit a poisson distribution and hence find the probability of getting atmost 3 misprints per page.

(15+5) = [20]

3. The following gives the sums, sums of squares and sums of products obtained in a multiple correlation analysis of 19 observations on y, x_1 and x_2 :

$\Sigma y = 105.5$, $\Sigma x_1 = 210.3$, $\Sigma x_2 = 200.1$, $\Sigma y^2 = 601.25$,
$\Sigma x_1^2 = 2357.17$, $\Sigma x_2^2 = 2121.17$, $\Sigma y x_2 = 1121.47$,
$\Sigma y x_1 = 1184.22$ and $\Sigma x_1 x_2 = 2230.81$.

(a) Obtain the multiple correlation coefficient of y on x_1 and x_2.

(b) Fit the regression plane of y on x_1 and x_2.

(c) Obtain the partial correlation coefficient between y and x_2 keeping x_1 fixed.

(10+15+5) = [30]

4.(a) Obtain the rank correlation coefficient for the following data :

X :	68	64	75	50	64	80	75	40	55	64
Y :	62	58	68	45	81	60	68	48	50	70

(b) Fit an equation of the form $y = ab^x$ to the following data :

x :	2	3	4	5	6
y :	144	172.8	207.4	248.8	298.6

(15+10) = [25]

:bcc:

1987-88	E113

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Applications : 1987-88

PART - I

FINAL EXAMINATION

Descriptive Statistics (Theory)

Date: 11.1.1988 Maximum Marks: 100 Time: 2 hours

Note: Answer any FOUR questions

1.(a) A variable takes values a, ar, ar^2, ..., ar^{n-1}, each with frequency unity. If A, G and H represent the arithmetic mean, geometric mean and harmonic mean of these values respectively, prove that $AH = G^2$.

(b) Suppose there are two values X_1 and X_2 with frequencies f_1 and f_2 respectively. Show that the variance is given by

$$S^2 = \frac{f_1 f_2}{(f_1 + f_2)^2} (X_1 - X_2)^2$$

(c) If $X_1 < \text{---} < X_n$ then show that

$$\frac{(X_n - X_1)^2}{2 \sum_1^n X_i^2} < 1.$$

2.(a) What are moments ? Let m_r' denote the rth order moment about origin 'A' and m_r the rth order central moment.

(i) Express m_r in terms of m_r''s.

(ii) Show that the least value of m_2' is m_2.

(iii) Establish the inequality $b_2 \geq b_1$, where notations have their usual meanings.

(b) Give a short account of the frequency curves that may be encountered in practice.

p.t.o.

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

Two regression equations :

$$8X - 10Y + 66 = 0$$

$$40X - 18Y = 214$$

$Var(X) = 9$.

(i) Determine which one is the regression of X on Y and which one is the regression of Y on X.

(ii) Find the mean value of X and Y and the s.d. of Y.

(iii) Find the correlation coefficient between X and Y.

(b) Distinguish between product correlation and correlation ratio.

Establish the inequality,

$$0 \leq r^2 \leq \eta^2 \leq 1$$

where notations have their usual meanings.

4.(a) Develop Spearman's rank correlation coefficient, assuming that there is no tie. Interpret this coefficient when it attains +1 and -1.

(b) What is contingency table ? Define Yule's measure of association (Q) and the coefficient of colligation (Y) in a 2 x 2 contingency table.

Hence prove that $Q = \dfrac{2Y}{1+Y^2}$.

5.(a) Define multiple correlation and partial correlation. Develop the formula $R_{1.23} = 1 - \dfrac{|R|}{R_{11}}$, where notations have their usual meanings. Show that the multiple correlation coefficient is always greater than all total and partial correlation coefficients.

(b) If the relation $aX_1 + bX_2 + cX_3 = 0$ holds for all sets of values of X_1, X_2, X_3, what must be the partial correlation coefficients ?

6. Write notes on :

   (a) Concentration curve.

   (b) Method of least squares in fitting a curve.

   (c) Diagramatic representation of numerical data.

:bcc:

1987-88	E123

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Applications : 1987-88

PART - I

FINAL EXAMINATION

Probability and Sampling Distributions

Date: 8.1.1988 Maximum Marks: 100 Time: $2\frac{1}{2}$ hours

Note: Question No.1 is compulsory. You may answer any part of the remaining questions.

1.(a) Two dice are thrown n times. Find the probability that each of the six combinations (1 , 1), (2 , 2), ..., (6 , 6) appears at least once.

[13]

(b) Let X, Y, Z be three independent gamma variables with parameters (α , p), (α , q) and (α , r) respectively. Find the distributions of

(i) $X + Y$ (ii) $\frac{X}{X + Y}$ (iii) $\frac{X}{X + Y + Z}$

(5+5+3) = [13]

(c) Let X follow the binomial distribution with parameters n and p, $0 < p < 1$. Find $E(|X-np|)$. Give details.

[15]

(d) Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed random variables. Let

$$P(X_1 = 0) = 1 - P(X_1 = 1) = 1 - p, \quad 0 < p < 1.$$

Let $Y = \min(X_1, \ldots, X_n)$ and $Z = \max(X_1, \ldots, X_n)$. Find $P(Y = 1)$ and $P(Z = 1)$.

(4+5) [9]

2.(a) Let F(x) be the distribution function of a random variable X. Show that F(x) is a right continuous function. Give details.

[10]

Contd..... 2/-

Contd..... Q.No.2

(b) State and prove Boole's inequality. [10]

3.(a) Let X be as in Question no.1.(c). Show that

$$\mu_{r+1} = pq(n r\mu_{r-1} + \frac{d\mu_r}{dp}), \quad r \geq 1$$

where $q = 1 - p$ and μ_r is the r-th central moment of X. [10]

(b) Find the mode of a Poisson variable. [7]

4.(a) What is the meaning of

$$X_n \xrightarrow{p} c$$

where $\{X_n\}$ is a sequence of random variables and c is a real number. [2]

(b) State and prove Chebyshev's WLLN. (5+10) = [15]

(c) Let $\{\lambda_n\}$ be a sequence of pairwise independent random variables such that

$$|X_n| \leq M \quad \text{for each} \quad n \geq 1$$

where M is a real number free from n. Show that $\{X_n\}$ obeys WLLN. [3]

5.(a) What is the meaning of the statement that X_n is $AN(\mu_n ; \sigma_n^2)$, $\sigma_n^2 > 0$. [3]

(b) State the Lindeberg-Levy central limit theorem. [5]

(c) Let $\{X_n\}$ be as in Question no.1.(d). Find

$$\lim_{n \to \infty} P(S_n \leq np)$$

where $S_n = X_1 + \ldots + X_n$. [8]

Contd..... 3/-

6.(a) State Chebyshev's inequality; prove it for the discrete case. Give details.

(2+8) = [10]

(b) Let X be a random variable such that

$$E(X) = m, \quad E(|X - m|^r) = k.$$

Show that

$$E(|X - m| \geq 1) \leq k$$

[5]

(c) Show that $E(E(X|Y)) = E(X)$ under the assumptions that the expectations are finite.

[5]

:bcc:

1987-88 | E122

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods and Applications : 1987-88

Probability and Sampling Distributions

PART - I

PERIODICAL EXAMINATION

21.12.1987 Maximum Marks: 100 Time: 2 hours

Note: Question No.4 is compulsory. You may answer any part of the remaining questions.

1. Let X and Y be two independent random variables each following the uniform distribution over the interval $(0, 1)$. Find the density function and the distribution function of
(a) $X+Y$; (b) $\frac{X}{Y}$.

(13+7) + (13+7) = [40]

2. Let the density function of X and Y be given by

$$f(x, y) = \begin{cases} 2 & \text{if } 0 \leq y \leq x \leq 1 ; \\ 0 & \text{otherwise} \end{cases}$$

(a) Find the marginal densities of X and Y. (7+7) = [14]

(b) Find the marginal distribution functions of X and Y. (5+5) = [10]

(c) Find the joint distribution function of X and Y [7]

(d) Find the conditional densities. (5+5) = [10]

(e) Are X and Y are independent ? Supply reasons. [3]

3. (a) When is a random variable said to follow a normal distribution ? [3]

Contd..... 2/-

Contd..... Q.No.3

(b) Evaluate the mean and variance of the normal distribution.

(5+5, = [10]

(c) Find the expected value of absolute deviation about the mean of the normal distribution. [5]

4. Let X and Y be two independent random variables each following the uniform distribution over the interval (-2, 3). Find the density function of X+Y.

[20]

:bcc:

1987-88 | E132

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods and Applications, 1987-88

PART - I

PERIODICAL EXAMINATION

Tests of Statistical Hypotheses

Date: 14.12.1987 Maximum Marks: 100 Time: 2 hrs.

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

1. (i) Define an unbiased test.

(ii) Prove that if a test is M.P. it is unbiased as well.

(iii) Define an UMPU test.

(3+7+5) = [15]

2. A buyer wants to purchase a large consignment of apples. To be convinced of the quality he selects 600 apples randomly and finds 65 to be below quality.

(a) Can you say that not more than 10% apples are below quality in the consignment ?

(b) Obtain 95% and 99% confidence intervals for the percentage of apples below quality in the consignment.

(c) Obtain almost certain limits for the percentage of below quality apples in the consignment.

(10+5+5) = [20]

3. Marks obtained in a statistics course by 6 male students in an examination were found to be (out of 100)

75, 30, 50, 60, 45, 52.

While the 8 female students got the marks as

90, 25, 60, 52, 44, 57, 70, 65

Contd..... 2/-

Contd..... Q.No.3

(a) Can you say that scores by male and female students are significantly different ?

(b) Can you say that variability in the above two sets is of same order ?

(c) Obtain 95% confidence interval for difference in the mean scores.

(15+5+5) = [25]

4. Marks obtained by 7 students, in English, before and after they received instruction were found to be as follows :

Student	1	2	3	4	5	6	7
% Marks before instruction	30	50	60	55	40	45	52
% Marks after instruction	45	55	70	52	45	46	57

(a) Can you say that there has been significant improvement after instruction ?

(b) Is the correlation between two sets of marks significant ?

(15+15) = [30]

5. (i) Define Fisher ' Z transformation for correlation.

(ii) In question no.4 above can you say that correlation coefficient between two sets of marks is 0.5 ?

(iii) Test the hypotheses that the regression coefficient of marks after instruction over before instruction is 0.95.

(3+7+10) = [20]

:bcc:

1987-88 | E142

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Applications : 1987-88

PART - I

PERIODICAL EXAMINATION

Date: 30.11.1987 Maximum Marks: 100 Time: 2 hours

Note: Answer the TWO GROUPS in separate answerscripts.

GROUP - A Time: 1 hour
(Vital Statistics) Max. Marks : 50

Note: Attempt all questions.

1.(a) Explain the following terms:

(i) Crude death rate.

(ii) age-specific death rates.

(iii) infant mortality rate.

What is the necessity of standardisation ? Explain briefly the two methods of standardisation of death rates (direct and indirect method).

(b) The following table gives the age specific death rates for 'A', the age distribution of 'A' and the 'Standard Population'.

Calculate

(i) Crude death rate for 'A'

(ii) Standardised death rate for 'A' by direct method of standardisation.

Contd..... 2/-

Contd..... Q.No.1.(b)

Age in years	Specific death rate for 'A' (per 1000)	Percentage age distribution	
		'A'	Standard Population
0 - 20	50	20	25
20 - 40	10	40	30
40 - 60	20	30	30
60 +	60	10	15

(17+8) = [25]

2.(a) Define the following rates:

(i) Crude birth rate.
(ii) general fertility rate.
(iii) gross reproduction rate.
(iv) net reproduction rate.

(b) Given the following table, Calculate :

(i) gross reproduction rate (GRR)
(ii) net reproduction rate (NRR)

Age of mother	Age specific fertility rate (female births)	Survival factor
15 - 19	.0108	.969
20 - 24	.0662	.967
25 - 29	.0675	.963
30 - 34	.0413	.958
35 - 39	.0216	.952
40 - 44	.0063	.942
45 - 49	.0004	.928

(16+9) = [25]

GROUP - B (Index Number)

Time: 1 hour
Max. Marks: 50

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

1. Describe the different problems faced in constructing index numbers.

[25]

2.(a) Show that the factor reversal test and time reversal test are not satisfied by Laspeyre's and Paasche's index numbers. Further show that these tests are satisfied by Fisher's ideal index number.

(6+6) = [12]

(b) What is a Chain index ? Discuss its advantages and disadvantages over a fixed-base index number.

(6+7) = [13]

3. The following data relate to the wholesale prices of cerelals at selected centres in India during two different weeks and the corresponding weights:

Item	Weight	Price (Rs. per maund)	
		Week ending 17-11-56	Week ending 21-12-57
Rice	224	20.50	17.50
Wheat	106	18.50	17.40
Jowar	19	16.25	10.50
Bazra	10	15.50	12.44
Barley	10	13.00	11.25
Maize	9	13.00	11.06
Ragi	4	10.12	12.75

How do the wholesale prices of cerelals in India during the week ending 21-12-57 compare with those in the week ending 17-11-56 ?

[25]

:bcc:

1987-88 | E112

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Applications : 1987-88

PART - I

PERIODICAL EXAMINATION

Descriptive Statistics (Theory and Practical)

Date: 16.11.1987 Maximum Marks: 100 Time: 2 hours

Note: Answer the TWO groups in separate answer books.

GROUP - A

Note: Answer any TWO questions

1.(a) A population of values of certain variable is symmetrically distributed about the constant K.

State whether the following results are correct or not:

(i) The mean is K.

(ii) The mean coincides with the mode.

(iii) $\frac{1}{2}$ (third quartile + first quartile) is not equal to K.

(iv) Skewness of the distribution is non-negative.

(v) Odd order central moments are zero.

(vi) The distribution cannot be bimodal.

(b) Let $\bar{X}_1$, $\bar{X}_2$ and S_1^2, S_2^2 be the means and variances respectively, based on two independent sets of observations of sizes n_1 and n_2. Find the expressions for the combined mean and combined variance.

9 + (6+10) = [25]

2.(a) Let r be the simple correlation coefficient obtained from n pairs of observations $(X_1, Y_1), \ldots, (X_n, Y_n)$ on two correlated variables.

Prove that $-1 \leq r \leq 1$.

Interprete the values of $r = \pm 1$.

Contd..... 2/-

Contd..... Q.No.2

(b) Show that the square of the simple correlation coefficient r defined in (a) can be expressed as a ratio of variance due to linear regression of y on x to the total variance of y.

(c) Verify which of the following statements is correct and which one is false.

Suppose the correlation coefficient between X and Y is negative. Then it means

(i) the regression coefficient of Y on X and that of X on Y are both negative.

(ii) the arithmetic mean of the products X_iY_i, i = 1...n, is greater than the product of the arithmetic mean of $X_1, \ldots, X_n$ and that of $Y_1, \ldots, Y_n$.

Add a few lines to substantiate your conclusion.

(10+7+8) = [25]

3. Write short notes on

(a) Fitting of a theoretical distribution to a given data.

(b) The curve of concentration.

(c) Simple correlation and regression.

(8+8+9) = [25]

GROUP - B

Note: Answer all the questions

4. The following data relate to the number of accidents per day in a certain year on B.T. Road:

Number of accidents	Number of days
0	133
1	135
2	67
3	23
4	6
5	1
Total	365

Contd..... 3/-

Contd..... Q.No.4

(a) Fit a Poisson distribution to the above data.

(b) Assuming that the average number of accidents is one per day, find the expected number of days with at least two accidents each day during the year.

(12+3) = [15]

5. The following table shows the frequency distribution of 157 plants according to yield in OZ.

Yield inOZ	Number of plants
3.5 - 7.5	8
7.5 - 11.5	21
11.5 - 15.5	45
15.5 - 19.5	39
19.5 - 23.5	35
23.5 - 27.5	9
Total =	157

Compute the first four central moments of this distribution and hence find β_1 and β_2 coefficients.

[20]

6. The lifetime of a particular model of a Stereo Cartridge is normally distributed with mean 1000 hours and a standard deviation 100 hours.

Find the probability that one of these cartridges will last.

(a) between 850 hrs. and 1050 hrs.

(b) 950 hours or less

In a sample of 1500 Cartridges, how many are expected to last

(c) more than 1200 hours ?

(d) between 850 and 1000 hours ?

(4+2+4+5) = [15]

:bcc:

1987-88	E131

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods and Applications : 1987-88

PART : I

PERIODICAL EXAMINATION

Tests of Statistical Hypotheses

Date: 9.11.1987 Maximum Marks: 100 Time: 2 hours

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

1. (i) State the problems of Point estimation, Interval estimation and Testing of Hypotheses.

(ii) Define the following terms in testing hypotheses :

(a) Two-types of errors

(b) Size of a test and level of significance

(c) Power of a test

(d) Most powerful test

(10+20) = [30]

2. (i) State N-P Lemma.

(ii) Derive Most powerful test for testing $H_0 : \mu = \mu_0$ versus $H_1 : \mu = \mu_1$; incase of normal distribution $N(\mu, \sigma^2)$ where σ is known.

(iii) A random sample of 10 individuals yielded their heights (in cms.) as:

110, 150, 170, 105, 180, 120, 115, 160, 140, 130.

Assuming that the hights follow a normal distribution with mean μ cm. and standard deviation 28.5 cm., test the following hypotheses :

(a) $H_0 : \mu = 150$ against $H_1 : \mu = 140$

(b) $H_0 : \mu = 150$ $H_1 : \mu = 160$

(c) $H_0 : \mu = 145$ $H_1 : \mu > 145$

(d) $H_0 : \mu = 145$ $H_1 : \mu < 145$

(e) $H_0 : \mu = 150$ $H_1 : \mu \neq 150$

Contd..... 2/-

Contd..... Q.No.2.(iii)

Obtain the power of the test in case of (a) and (b).

(15+15+25) = [55]

3. (i) Define the terms: Confidence interval and Confidence-coefficient.

(ii) Assume that profit percentages of industrial units follow a normal distribution with mean μ and s.d. σ.

Twenty industrial units were selected randomly and their profit-percentages were found. The statistical enquiry yielded the following data :

5.0, 10.5, 25.0, 20.5, 30.0, 25.5, 15.5, 20.0, 25.0, 18.5

10.5, 12.5, 40.0, 22.5, 45.5, 50.0, 15.5, 18.5, 20.5, 22.5.

Obtain 95% and 99% confidence intervals for the mean profit percentage μ in case

(a) $\sigma = 12.5$ (b) σ unknown.

(5+20) = [25]

:bcc:

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Applications : 1987-88

PERIODICAL EXAMINATION

Sample Surveys (Theory and Practical)

Date: 19.10.1987 Maximum Marks: 100 Time: 2 hrs.

Note: You must attempt Question No.6 and any FOUR from the remaining questions.

1.(a) Define a population and an elementary unit. Is it always possible to get the exact value of the parameter.? If not why ? Suppose it is not always possible to enumerate completely, what are you going to propose ? Discuss in a few lines about its merits and demerits.

(2+3+5) = [10]

(b) What is non-sampling error ? Discuss the different sources of non-sampling error mentioning the methods by which non-sampling errors can be controlled.

(2+8) = [10]

2.(a) What do you mean by a sampling design and inclusion probability of first and second orders ? Suppose in estimating the population total, the proposed estimator is

$$\hat{Y}_{H-T} = \sum_{i \in s} y_i/\pi_i$$

Is it unbiased ? Calculate the variance of the proposed estimator.

[10]

(b) Hence or otherwise, show that sample mean $\bar{y}$ based on simple random sampling without replacement is unbiased for $\bar{Y}$, the population mean and also find that

$$V(\bar{y}) = \frac{(N-n)}{Nn} S^2$$

where $$S^2 = \frac{1}{N-1} \sum_{i=1}^{N} (y_i - \bar{Y})^2 .$$

[5]

Contd..... 2/-

Contd..... Q. No.2

(c) Suppose the proposed estimates in Q.No.2.(a) is

$$T_2 = \lambda \hat{Y}_{H-T}$$, when λ is a real constant.

Will it be unbiased for Y ? If not what is the amount of bias ?

[5]

3.(a) Discuss the situations when stratification is made. Build up an estimator for population mean and also calculate the variance of the proposed estimator when sampling from each stratum has been made by simple random sampling without replacement.

[10]

(b) With two strata, a sampler would like to have $n_1 = n_2$ for administrative convenience instead of using the values given by Neyman allocation. If $V(\bar{y}_{st})$, $V_{op}(\bar{y}_{st})$ denote the variances given by $n_1 = n_2$ and the Neyman allocation, respectively. Show that the fractional increase in variance

$$[V(\bar{y}_{st}) - V_{opt}(\bar{y}_{st})]/V_{opt}(\bar{y}_{st}) = \left(\frac{r-1}{r+1}\right)^2$$

where, $r = \frac{n_1}{n_2}$. [10]

4.(a) Let P_i is the unconditional probability of getting the ith sample at the first stage of randomization and $P_{j|i}$ is the conditional probability of getting ijth sample at the second stage of randomization given that the ith sample has been selected at the first stage of randomization. Let t_{ij} be the estimate based on the ijth sample. show that

$$E(t) = E_1 E_2(t)$$

and

$$V(t) = E_1 V_2(t) + V_1 E_2(t).$$

[5]

Contd..... 3/-

Contd..... Q.No.4

(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.

(5+10) = [15]

5.(a) Describe the procedure of linear systematic sampling. Show that the sample mean under linear systematic sampling is unbiased for population mean.

[5]

Under what conditions systematic sampling will be more better than simple random sampling with and without replacement.

[10]

Can variance of the estimator be estimated under systematic sampling ? If not why ?

[5]

6. A hypothetical population comprises of 5 units with values of a character y as 5, 3, 8, 2, and 1.

By enumerating all samples of fixed size 2 from the population with SRSWOR, prove the following :

(i) $E(\bar{y}) = \bar{Y}$

(ii) $E(s^2) = S^2$

and (iii) $V(\bar{y}) = \frac{N-n}{Nn} S^2$

where $N = 5$ and $n = 2$.

[20]

:bcc:

1987-88 E121

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods and Applications : 1987-88

PERIODICAL EXAMINATION

PART - I

Probability and Sampling Distributions

Date: 12.10.1987 Maximum Marks: 100 Time: 2 hours

Note: You may answer any part of any question. Special credit will be given to answers which are precise and to the point. State clearly the results you are using.

Question No.6 is compulsory.

1.(a) If $A \subset B$, show that $P(A) \leq P(B)$.

(b) Let $\{A_n\}$ be a sequence of events.

(i) If $A_1 \subset A_2 \subset A_3 \subset \ldots$, show that

$$\lim_{n \to \infty} P(A_n) = P(\bigcup_1^\infty A_n)$$

(ii) If $A_1 \supset A_2 \supset A_3 \supset \ldots$, show that

$$\lim_{n \to \infty} P(A_n) = P(\bigcap_1^\infty A_n).$$

(10+5) = [15]

(c) If $P(A) = \frac{1}{3}$, $P(A \cup B) = \frac{1}{2}$ and $P(A \cap B) = \frac{1}{4}$, find $P(B)$. [5]

2. A die is rolled 12 times. Compute the probability of getting two sixes. [10]

3.(a) Let $A_1, \ldots, A_n$ be n events. Write down the expression of $P(\bigcup_{m=1}^n A_m)$ in terms of $P(A_1), \ldots, P(A_n)$. [5]

(b) n distinct balls marked 1, ..., n are distributed at random to n distinct boxes also marked 1, ..., n.

Contd..... 2/-

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

(i) Find the probability that there are no matches.

(ii) Find the probability that there are exactly r matches, $1 \leq r < n$.

Give details of your arguments.

(20+10) = [30]

4.(a) Define carefully a random variable, a discrete random variable, an absolutely continuous random variable.

(5+5+5) = [15]

(b) What is a distribution function of a random variable ? If F is a distribution function, show that $\lim_{n \to \infty} F(n) = 1$; give details.

(3+12) = [15]

5. When is a random variable said to follow a binomial distribution ? Find its variance; give details.

(3+17) = [20]

6. Let X be a random variable such that $P(X = \frac{1}{3}) = \frac{1}{3}$, $P(X = \frac{1}{2}) = \frac{1}{2}$ and $P(X = 1) = \frac{1}{6}$. Determine the distribution function of X.

[10]

:bcc:

1986-87 | 82633

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods and Applications : 1986-87

PART - II

SUPPLEMENTARY EXAMINATION

Date: 12.10.1987 Maximum Marks: 100 Time: 2 hours

Note: Answer the two groups in seperate answer books.

GROUP - A Max. Marks: 50

Design of Experiments (Theory and Practical)

Note: Answer any two questions.

1. An experiment was set up to test 7 varieties of corn with 7 different spacings in four replicates. Suggest 3 possible experimental designs with a break-up of the degrees of freedom for each designs. Discuss briefly the applicability of each experimental design listed.

[25]

2. The following data relate the body weights of calves of 8 weeks of age. There were three levels of feeding given to a random sample of five calves each. The completely randomized design was used. Obtain the analysis of variance table and analyse the data.

Level of feeding		
Subnormal	Normal	Supernormal
118	142	162
122	129	173
121	134	168
126	132	182
109	135	172

[25]

3.(a) Define main effect and two factor interaction in 2^2 factorial experiment. Give an analysis of such an experiments.

Contd..... 2/-

Contd..... Q.No.3

(b) Compute main effect and interaction from the following data :

treatment	yield
(1)	25
a	40
b	60
ab	70

Write down the analysis of variance table.

[25]

GROUP - B Max. Marks: 50

Industrial Statistics (Theory and Practical)

Note: Attempt Q.No.1 (compulsory) and any TWO from the remaining.

1.(a) Give a brief review of the areas of application of statistical methods in controlling the quality of manufactured product.

(b) When would you use an np-chart in preference over a p - chart ?

(c) What are the assumptions about the underlying distribution for different values of $n\bar{p}$ of a set of observations collected from a process ?

(d) Give a brief account of the Dodge Romig System of lot by lot Sampling Inspection by attribute.

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

2. The specifications for a certain brand of hair oil packed in bottles of 400 ml. is given as 400 ± 5 ml. 25 samples of size 4 each collected from the filling line gives an average net content ($\bar{X}$) of 395 ml. and average range ($\bar{R}$) of 4.05. Both $\bar{X}$ and R charts are in a state of statistical control.

(i) Is the process capability adequate to meet the specifications ?

Contd..... 3/-

Contd..... Q. No.2

(ii) What percentage of the filled bottles are expected to meet the specifications under the existing conditions ?

(iii) If it is decided to continue the control chart for future control with a sample size of 4 bottles find the control limits and the central line.

(4+8+3) = [15]

3.(a) The following data gives the number of missing rivets noted at aircraft final inspection :

Serial No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
No. of missing rivets	8	16	14	19	11	15	8	11	21	12	23	16	9	25	10

draw a suitable control chart for the data and give your comment.

(b) Draw the oc-curve of the control chart.

(7+8) = [15]

4.(a) Describe the procedure of a single sampling plan by attribute and indicate its advantages over 100% inspection.

(b) Explain the terms, (i) AOQ, (ii) OC-curve of a sampling plan.

(c) Find an appropriate Dodge and Romig single sampling plan for the following :

(i) Lot size = 2000; LTPD = 0.5%; process Average = 0.1%.

(ii) Lot size = 2500; AOQL = 1.0%; process Average = 0.2%.

(5+6+4) = [15]

1987-88 | E111

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Applications : 1987-88

PART - I

PERIODICAL EXAMINATION

Descriptive Statistics (Theory and Practical)

Date: 28.9.1987 Maximum Marks: 100 Time: 2 hours

Answer the two groups in seperate answer books.

GROUP A

Note: Answer any TWO questions.

1.(a) In each of the following data identify the characteristic being studied and state whether it is qualitative or quantitative :

(i) The types of metal that are used to make automobiles.

(ii) The amount of anesthesia required for each of the three surgical patients.

(iii) The planting distances for certain types of garden vegetables.

(iv) The nationality of each of twenty immigrants to India.

(b) Let G_X denote the geometric mean of $X_1, \ldots, X_N$, G_Y denote that of $Y_1, \ldots, Y_N$ and G that of the combined observations.

Show that $G^2 = G_X G_Y$.

(c) Given $X > 0$, $Y > 0$, let A represent the arithmetic mean of the reciprocals of X and Y and B represent the reciprocal of their arithmetic mean. Prove that $A \geq B$.

(8+8+9) = [25]

2.(a) It is given that the arithmetic mean of 12 observations is 65, that of first 3 observations is 45 and that of last three observations is 60. Find the arithmetic mean of the remaining observations.

(b) Prove that the standard deviation is always less than the root mean square deviation (RMSD). When are they equal ?

Contd..... 2/-

Contd..... Q.No.2

(c) Suppose a set of observations consists of only two numbers X_1 and X_2 with equal frequencies. Show that the variance of the set of observations is $\frac{1}{4}(X_1 - X_2)^2$.

(8+8+9) = [25]

3.(a) For a given set of observations $X_1, \ldots, X_n$, show that the mean deviation about mean is given by

$$(MD)_{\bar{X}} = \frac{2}{n} \sum_{X_i > \bar{X}} (X_i - \bar{X}).$$

(b) A veterinarian was interested in determining how many animals were treated in the clinic each day. A random sample of 20 days' records produced the following results:

15, 17, 24, 16, 18, 15, 19, 22, 25, 21
18, 17, 20, 20, 20, 18, 10, 16, 12, 21.

The arithmetic mean and s.d. of the above data had been calculated and found to be 18.2 and 3.69 respectively.

Suppose the veterinarian had seen two additional animals on each of 20 days. Using this information and the a.m. and s.d. obtained above, compute the a.m. and s.d. of the new data. Give your comments on the results so obtained.

(c) Define raw moments and central moments on the basis of a random sample $X_1, \ldots, X_n$.

Express the fourth order central moment in terms of raw moments.

Define b_1 and b_2 coefficients and show that $b_2 \geq b_1$.

(8+7+10) = [25]

Group B

Answer all the questions.

4. Following is the distribution of marks obtained by 160 candidates in a selection test for a certain job:

Contd..... 3/-

Contd..... Q.No.4

Marks (out of 100)	Students	Marks (out of 100)	Students
0-10	3	50-60	20
10-20	18	60-70	6
20-30	30	70-80	3
30-40	45	80-90	0
40-50	35	90-100	0

(a) What is the marks obtained by maximum number of candidates ?

(b) Find the coefficient of variation of the distribution.

(c) It was decided by the selection committee to call only 50 candidates for the interview. Find out the minimum marks necessary for a candidate to be called for the interview.

(10 x 3) = [30]

5. A distribution consists of two components. The first component has 100 items with mean 15 and s.d. 3. If the distribution has 250 items with mean 15.6 and s.d. $\sqrt{13.44}$, find the mean and s.d. of the second component.

[10]

6. The mean and s.d. of 20 items is found to be 10 and 2 respectively. At the time of checking it was found that one item 8 was incorrect. Calculate the mean and s.d. if :

(a) the wrong item is omitted.

(b) it is replaced by 12.

[10]

:bcc:

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Applications : 1986-87

PART - II

SUPPLEMENTARY EXAMINATION

Statistical Inference (Theory and Practical)

Date: 28.9.1987 Maximum Marks: 100 Time: 2 hours

Note: You may answer any part of any question.

1. Let $X_1, \ldots, X_n$ be i.i.d., each following the $N(\mu, \sigma^2)$ distribution where μ and σ^2 are unknown, $\sigma^2 > 0$. Show that $(\sum_{i=1}^{n} X_i, \sum_{i=1}^{n} X_i^2)$ is sufficient for (μ, σ^2). What will be the sufficient statistic if it is given that $\mu = \sigma^2$; justify your answer.

(20+20) = [40]

2. Let $X_1, \ldots, X_n$ be i.i.d., each following the uniform distribution over $(2, 2+\theta)$, $\theta > 0$. Find the maximum likelihood estimator of θ. [20]

3. State and prove the Rao - Blackwell theorem.

(5+15) = [20]

4. 20 samples were taken from a container of a particular brand of hydrogenated Negetable oil. Each of four analysts was given five different samples and asked to determine the melting point. The results are given in the following table:

Analyst	Individual determinations of melting points (in Farenhite)				
1	93.60	94.64	96.30	93.62	93.51
2	96.44	96.53	98.38	97.00	97.53
3	92.57	94.01	92.49	93.29	90.87
4	95.55	95.90	94.25	95.80	96.21

Examine whether the analysts are consistent in their determination of the melting point. [35]

5. Write a short note on the method of moments. [15]

:bcc:

1986-87 E/5[illegible]3

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods and Applications : 1986-87

PART - II

SUPPLEMENTARY EXAMINATION

Sample Survey (Theory and Practical)

Date: 28.9.1987 Maximum Marks: 100 Time: 2 hours

Note: Answer any FOUR questions.

1.(a) Prove that under certain assumptions to be stated by you :

$$V_{ran} \geq V_{prop} \geq V_{opt}$$

Can it happen that

$$V_{ran} < V_{prop} \quad ?$$

Explain.

(b) For a socio-economic survey all the villages in a region including the uninhabitated ones were grouped into 4 strata on the basis of their altitude above sea-level and population density and samples were selected from each stratum by SRSWOR. The data on number of households in each of the sampled villages are given below.

Stratum Sl. No.	Total no. of villages	No. of households in sample villages
1	114	10, 21, 24, 0, 10
2	92	12, 32, 15, 21
3	105	31, 27, 25, 32, 14
4	74	15, 8, 9

(i) Estimate the total no. of households unbiasedly.

(ii) Compare the efficiency of the present allocation with that of the optimum allocation keeping the total sample size fixed.

(10+15) = [25]

p.t.o.

2.(a) A survey is to be made of the prevalance of the common diseases in a large population. For any disease that affects at least 1% of the individuals in the population it is desired to estimate the total number of cases with a coefficient of variation of not more than 20%. What size of simple random sample is needed, assuming that the presence of the disease can be recognized without mistakes ?

(b) Values of a variable y for a population of 40 units are:

Unit	y	Unit	y	Unit	y	Unit	y
1	0	11	10	21	22	31	39
2	1	12	11	22	25	32	43
3	2	13	13	23	29	33	46
4	1	14	12	24	30	34	50
5	4	15	12	25	32	35	53
6	5	16	15	26	35	36	52
7	7	17	14	27	33	37	57
8	7	18	17	28	38	38	59
9	9	19	20	29	40	39	63
10	8	20	23	30	41	40	62

(i) Draw 5 independent systematic sample of 4 units each.

(ii) Obtain a single combined estimate of Y based on all these 5 samples.

Making use of 20 sample observations obtained in (i) estimate the relative efficiency of systematic sampling to that of srswor for estimating the total value.

(10+15) = [25]

3.(a) Obtain the expression of variance of the ratio estimator of population mean and find out its estimate. Is this estimator unbiased ?

(b) The following table gives the present population (x) and the previous census population (y) of 15 villages in an area. Draw a SRSWOR of 5 villages and estimate the present population by using the previous census population as an auxiliary variable. Find out an estimator of the MSE of the estimate.

Contd..... 3/-

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

Village	y	x	Village	y	x
1	14	15	9	120	71
2	51	14	10	10	11
3	72	53	11	0	5
4	66	38	12	27	16
5	111	70	13	35	15
6	0	4	14	72	35
7	29	15	15	8	9
8	48	13			

(13+12) = [25]

4.(a) Obtain an unbiased estimate of population total under PPSWR scheme of sampling and an unbiased estimator of its variance.

(b) Draw a sample of 5 factories from the given population by Lahiri's method using the no. of workers as size of a factory and obtain an estimate of the total value added of all the 15 factories. Find an estimate of the rse of the estimate.

Factory	Value added	No. of workers	Factory	Value added	No. of workers
1	410	31	8	35	21
2	21	12	9	106	72
3	39	23	10	29	18
4	124	73	11	172	84
5	1089	519	12	2358	1059
6	23	15	13	351	123
7	71	19	14	61	34
			15	31	27

(13+12) = [25]

5. Write short notes on any Three

(a) Interpenetrating subsampling.

(b) Comparison between SRSWOR and SRSWR.

(c) Cluster sampling.

(d) Stages in conducting large scale sample Surveys.

[25]

6.(a) Explain the concept of two-stage sampling. Obtain an unbiased estimate of population total using SRSWOR on both the stages. Obtain an unbiased estimate of its variance when SRSWR is used at the first stage of sampling.

(b) To determine the yield rate of paddy in a district of West Bengal, 6 villages were selected at random out of 55 villages and 3 plots were selected by SRSWOR in each selected village, the yields in suitable units were obtained as follows :

Village / Sample Plots	Yield and Plots						Total
	1	2	3	4	5	6	
1	24	11	26	26	32	27	146
2	57	21	13	13	32	28	164
3	24	21	36	9	42	18	150
Total	105	53	75	48	106	73	460
Total number of plots	18	19	15	17	14	16	

(i) Estimate the total yield in all the 55 villages.

(ii) Obtain the estimate of the SE of estimate of mean yield.

(15+10) = [25]

---

bcc: