/INDIAN STATISTICAL INSTITUTE

'One year Evening Course in Statistical Methods and Applications 1977-78

Part II Final Examinations (Supplementary)

Paper IV(i): Inference - Theory

Date: 5.7.78 Maximum Marks: 75 Time: 2 hours

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

1.a) Let $X \sim N(\mu, \sigma^2)$. Let x_1, x_2, x_3 and x_4 be a random sample of size 4 from X. Write down the distributions of each of the following statistics: (You don't have to give proofs)

i) $x_1 + x_2 - x_3$. (ii) $\frac{1}{\sigma^2}\{(x_1 - \mu)^2 + (x_3 - \mu)^2\}$

iii) $\frac{2(x_2 - \mu)^2}{(x_3 - \mu)^2 + (x_4 - \mu)^2}$.

b) Let $X \sim N(\mu, \sigma^2)$ Let $x_1, \ldots, x_n$ be a random sample of size n from X. Show that

$$\sum_{i=1}^{n} x_i \quad \text{and} \quad \sum_{i=1}^{n} (x_i - \bar{x})^2$$

are jointly sufficient for μ and σ^2.

2. Explain clearly the test procedure you use with computational layouts in each of the following situations:

i) $X \sim N(\mu, \sigma^2)$ and $x_1, \ldots, x_n$ is a random sample of size n from X. σ^2 is unknown. You have to test $H_0: \mu = 4$ against $H_1: \mu > 4$.

ii) test for independence in a two way contingency table.

3.a) What is meant by 'confidence interval'? Consider the situation in 2(i). Obtain 95% confidence interval for μ.

b) Let X be a random variable having the density

$$\frac{\theta^k}{\Gamma k} e^{-\theta x} x^{k-1} \qquad \alpha > 0, \; k > 0, \quad 0 < x < \infty$$

Let $x_1, \ldots, x_n$ be a random sample of size n from X. Obtain the minimum variance unbiased estimator of $\frac{1}{\theta}$.

4. Let X be a random variable taking only two values 1 and 0 with probabilities p and 1 - p respectively ($0 < p < 1$) Let $x_1, \ldots, x_n$ be a random sample of size n from X. Obtain the maximum likelihood estimator of p. Show that the maximum likelihood estimator is sufficient for p.

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

i) most powerful tests and Neyman-Pearson lemma.

ii) Point estimation and interval estimation.

iii) Use of sampling distributions in testing statistical hypotheses.

iv) Analysis of variance for one-way classification of data.

INDIAN STATISTICAL INSTITUTE

One-year Evening Course in Statistical Methods and Applications: 1977-78

Part II Final Examinations

CALCUTTA and DELHI

Paper VI(b): Control Chart and Acceptance Sampling Theory and Practical

Date: 2.6.78 Maximum Marks: 75 Time: 2 hours

Note: Qs. 1 and 2 are compulsory. Answer any one question from questions 3 and 4. Marks allotted for each question are given in brackets [ ].

1.a) Explain the term 'quality of a manufactured item and discuss the common ways of specifying quality giving an example in each case. [6]

b) A company sells tea packets of net weight 450 gms., each. Daily five packets are being sampled from the production line and weighed in the laboratory. The sample averages and ranges of the net weight in grams for 20 days are as follows:

Day	Average	Range	Day	Average	Range
1	445.2	9.4	11	452.2	16.8
2	456.8	20.2	12	451.8	28.2
3	454.9	21.2	13	457.9	18.0
4	458.0	30.2	14	449.0	24.8
5	455.0	25.8	15	453.5	17.8
6	450.5	15.2	16	459.0	18.8
7	454.2	20.8	17	447.5	19.0
8	452.9	13.2	18	456.2	18.0
9	454.4	16.0	19	450.2	24.2
10	448.6	15.2	20	453.9	22.2

i) Draw an $\bar{X}$ and R chart neatly on a graph paper and examine if the process is under statistical control. [15]

ii) Estimate process standard deviation and hence obtain process capability. [3]

iii) If the market specifications are 450 gms $\pm$ 20 gms, estimate what percentage of packets being supplied to the market are either heavier than the upper specification limit or lighter than the lower specification limit. [6]

2.a) Define the terms (i) AOQL and (ii) AQL. [6]

b) For the single sampling plan with sample size 20 and acceptance number 2 obtain the probabilities of acceptance for suitable values of incoming lot qualities and draw an OC curve.

(Hint: Use Poisson approximation and refer suitable tables on cumulative Poisson probabilities.) [24]

3.a) Compare 100 % inspection with sampling inspection. [4]

b) Write down procedure for double sampling plan by attributes and mention some advantages of double sampling plan over single sample plan. [6]

c) Write down the steps to read out single sampling LTPD plans from Dodge and Romig's published sampling inspection tables. [5]

---

4. Write short notes on any three of following:-

a) Rational subgroups

b) Specification VS Process capability

c) Fraction defective chart

d) Producer's and Consumer's risk. [5+5+5]=[15]

INDIAN STATISTICAL INSTITUTE

One-year Evening Course in Statistical Methods and Applications: 1977-78

CALCUTTA AND DELHI

Part II: FINAL EXAMINATION

Time: 2 hours Paper VI(a): Design of Experiments (Theory and Practical) Max.Marks: 75

N.B.: Attempt any number of questions.

1. Describe briefly the various methods available for increasing the accuracy of agricultural field experiments. [15]

2. (a) Describe a Latin Square arrangement of order r and indicate situations where such a design would be appropriate.

   (b) Describe the method of analysis of a Latin Square Design with one missing observation. [8+12]

3. (a) What is a factorial experiment? In what respects is it different from a number of single factor experiments?

   (b) Define the terms main effects and interaction effects in relation to a 2^3 experiment. [10+10]

4. Five treatments A, B, C, D and E are compared and four observations for each treatment are made. The experiment is arranged in a randomised block design with four blocks. The data obtained and the design are:

Block 1	56.6(A)	63.3(C)	64.7(E)	60.0(D)	56.0(B)
Block 2	65.5(E)	61.9(D)	57.1(B)	53.6(A)	59.7(C)
Block 3	57.0(C)	54.0(A)	59.9(D)	63.4(E)	55.5(B)
Block 4	59.3(D)	62.8(E)	57.2(C)	55.8(B)	54.0(A)

Perform the formal analysis of variance for this design and make your conclusions. [40]

INDIAN STATISTICAL INSTITUTE

One-year Evening Course in Statistical Methods and Applications: 1977-78

Part II Final Examinations

CALCUTTA and DELHI

Paper V(ii): Sample Surveys: Practical

Date: 29.5.78 Maximum Marks: 75 Time: 2 hours

Note: Attempt as many questions as you can. The maximum you can score is 75. The paper carries 90 marks. Marks allotted for each question are given in brackets [ ].

---

1. The following table gives the data of all the 400 cultivators' holdings in a subdivision of U.P. which are stratified according to size. The number of holdings (N_i), mean area under wheat per holding ($\bar{Y}_i$) and s.d. of area under wheat (S_i) are given below for each stratum:

Stratum No.	Holding size (acres)	N_i	$\bar{Y}_i$	S_i
1	0 - 50	80	16	10.2
2	51 - 100	160	24	18.2
3	101 - 150	90	34	30.5
4	> 150	70	40	35.1

A sample of 40 holdings is to be drawn with SRSWOR in each stratum.

i) Obtain sample sizes under proportional allocation.

ii) Obtain sample sizes under Neyman allocation.

iii) Obtain the relative efficiency of the sampling strategy A over B and C where

A = (Simple strat. sampling with Neyman allocation, $\bar{y}_{st}$)

B = (Simple strat. sampling with proportional allocation, $\bar{y}_{st}$)

C = (SRSWOR, $\bar{y}$)

and $\bar{y}_{st}$ is an usual unbiased estimator for $\bar{Y}$.

[2+6+16]=[24]

---

2. There are 10 villages, in a Police Station, having 20 Plots each. The first stage units are villages and second stage units are plots. A SRSWOR of 3 villages is selected and then 4 plots are selected in each selected village using SRSWOR. The yields of Jute, in suitable units, are obtained as follows:

Plot No.	Village 1	Village 2	Village 3
1	10	5	17
2	20	30	11
3	15	20	20
4	30	8	10

i) Obtain the estimate for the total yield in the Police station.

2. (contd.)

ii) Obtain the estimate for the mean yield per plot.

iii) Obtain the estimate for the mean yield per village.

iv) Give an unbiased estimate for the variance of the estimate in (i) above. [5+2+2+15]=[24]

---

3. The following table gives the values for a sample of 4 factories with PPSWR, size being x, from a population of 15 factories.

Sl. No.	1	2	3	4
Number of workers (x)	20	25	10	80
Output in Rs.('000) (y)	1500	1300	950	2500

The population total of x is known to be 500.

i) Estimate unbiasedly the total output Y.

ii) Give an unbiased estimate for the variance of the above estimate. [6+12]=[18]

---

4. A simple random sample of 49 schools was selected from a population of 196 schools and the number of students (x) and expenditures (y), in Rs., on some items were noted for each school. The data yielded

$$\sum_{i=1}^{n} y_i = 6262, \quad \sum_{i=1}^{n} y_i^2 = 152\,7882, \quad \sum_{i=1}^{n} x_i y_i = 1251630,$$

$$\sum_{i=1}^{n} x_i = 5054, \quad \sum_{i=1}^{n} x_i^2 = 1044504.$$

The total number of students in the population is known to be 22919.

a) Estimate the total expenditure Y using

i) Simple random estimator

ii) Ratio estimator.

b) Obtain estimates of the variances of the above estimates. [3+5+16]=[24]

INDIAN STATISTICAL INSTITUTE

One-year Evening Course in Statistical Methods and Applications:1977-78
CALCUTTA AND DELHI

Part II: Final Examinations

Paper VI: Sample Surveys Theory

Date: 26.5.78 Maximum Marks: 75 Time: 2 hours

Note: The paper carries 105 marks. Answer as many as you can. Maximum you can score is 75. Marks allotted for each question are given in brackets[ ].

1. How will you estimate unbiasedly the total number of literates in a city using simple random sampling without replacement. Show that the estimate you suggest is actually unbiased. Obtain the variance of your estimate.

[4+5+12=21]

2. For a fixed total sample size compare the precisions of the usual unbiased estimators for population mean in the cases of

   (a) Simple random sampling with replacement

   (b) Stratified random sampling with replacement with proportional allocation.

   (c) Stratified random sampling with replacement with Neyman's optimum allocation.

[16]

3. Derive Neyman's optimum allocation of given total sample size to different strata. [15]

4. Work out the approximate expressions for bias and variance of the ratio estimate of population total of a variate under study in case of a simple random sampling. State clearly the assumptions involved and their implications.

[7+10+3=20]

5. Describe Lahiri's method of drawing a sample from a finite population with probabilities proportional to sizes of the sampling units. Is this procedure more convenient than 'Cumulative Total' method? Give reasons.

[10+5=15]

6. (a) What methods do you suggest for estimating the variance of the sample mean in case of systematic sampling?

   (b) What are interpenetrating sub-samples? What are their uses?

   (c) Write a short note on 'Census vs sample survey'.

[7+6+5=18]

INDIAN STATISTICAL INSTITUTE

One-year Evening Course in Statistical Methods and Applications: 1977-78

CALCUTTA AND DELHI

Paper IV(b): Inference (Practical)

Date: 24.5.1978 Max.Marks:75 Time: 2 hours

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

1. (a) Two suppliers are competing to receive the order for bulk supply of a certain item to a factory. The factory inspector obtained random samples from both the suppliers and the results are as follows:

Supplier	Sample size	Percent defective
1	104	30.55
2	219	32.05

Can we choose a particular supplier on the basis of the anove data?

(b) Given below are the probabilities and the observed frequencies in four phenotypic classes, AB, Ab, aB and ab, in a genetical experiment. Estimate the parameter θ by the maximum likelihood method and find its standard error.

Frequencies in Phenotypic classes in a Genetical Experiment

Phenotypic class	Probability	Observed frequency
AB	$\frac{1}{4}(2+\theta)$	102
Ab	$\frac{1}{4}(1-\theta)$	25
aB	$\frac{1}{4}(1-\theta)$	28
ab	$\frac{1}{4}\theta$	5

[10+15]

2. (a) Ten young men were put through a strenuous physical programme by the Army. Their weights (lbs) were recorded before and after with the following results.

Man	1	2	3	4	5	6	7	8	9	10
Weight before	127	195	162	170	143	205	168	175	197	136
Weight after	135	200	160	182	147	200	172	186	194	141

Does the programme affect the average weight of young men?

(b) Two machines both produce 1/2 inch bolts. Machine 1 is somewhat more expensive than machine 2. We would like to buy machine 2 unless we can prove that machine 1 makes bolts with lower variance of diameters. Sample bolts collected from the two machines show the following results:

Machine	Sample size	Estimation of variance
1	25	.00028
2	30	.00045

Which machine should be purchased? [10+15]

3. In an experiment on immunization of cattle from tuberculosis the following results were obtained:

	Affected by disease	Unaffected by disease
Inoculated	12	26
Not inoculated	16	6

Examine the effect of vaccine in controlling susceptibility to tuberculosis. [15]

4. (a) The scores (out of 100 marks) of 10 students in the two subjects mathematics (x) and physics (y) are given below:

	(x)	(y)
1	41	40
2	73	69
3	62	65
4	59	70
5	57	68
6	49	53
7	81	72
8	92	83
9	40	42
10	55	43

Does the performance in mathematics depend on the performance in physics? [12]

(b) Three persons suspected to be driving under the influence of liquor were stopped and blood samples taken from each were sent to a laboratory. Five determinations on percentage of alcohol in the blood were made on each sample.

Determination of percentage of alcohol

	Alcohol ( ) in blood sample				
Person	1	2	3	4	5
1	0.08	0.06	0.06	0.07	0.09
2	0.14	0.18	0.16	0.15	0.12
3	0.00	0.02	0.01	0.02	0.03

Are all the three drivers equallyintoxicated? [13]

INDIAN STATISTICAL INSTITUTE

One-year Evening Course in Statistical Methods and Applications: 1977-78

Part II Final Examinations

CALCUTTA and DELHI

Paper IV(i): Methods of Statistical Inference Theory

Date: 23.5.78 Maximum Marks: 75 Time: 2 hours

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

---

1. EITHER

a) State clearly Rao-Blackwell theorem and Cramér-Rao inequality. State (without proof) a necessary and sufficient condition for the equality in Cramér-Rao inequality.

b) Let X have the density

$$\frac{\theta^k}{\lceil k} e^{-\theta x} x^{k-1} \qquad 0 \leq x < \infty$$
$$\theta > 0 \quad \text{and} \quad k > 0.$$

Obtain the function $\psi(\theta)$ which has an unbiased estimator t (based on a random sample of size n from x) the variance of which attains Cramer-Rao lower bound. What is the variance of this estimator ?

c) Consider the statistic r, the observed number of successes in n Bernoullian trials with probability π of success. Show that r is a sufficient statistic for π. [2+2+2+6+3+4]=[19]

OR

d) Explain clearly the concept of sufficiency. State (without proof) the factorisation theorem for sufficient statistics.

e) Let $X_1, X_2, \ldots, X_n$ be i.i.d. r.v.'s each distributed as $N(0, 1)$. Find the distribution of $Y = X_1^2 + X_2^2 + \ldots + X_n^2$. (Give the complete proof.)

---

2. a) Explain clearly the maximum likelihood method of estimating parameters.

b) Let X follow a Poisson distribution with mean λ. Obtain the maximum likelihood estimator of λ based on a random sample of size n from X.

c) Prove or disprove: The maximum likelihood estimator is unbiased. [5+7+7]=[19]

---

3. a) Give an example of a parameter which does not possess an unbiased estimator.

b) Give an example of a parameter which possesses unbiased estimators but does not possess a minimum variance unbiased estimator.

c) Let $X \sim N(\mu, \sigma^2)$. Let $x_1, \ldots, x_n$ be a random sample of size n from X, show that $\bar{x} = n \sum_{i=1}^{n} x_i$ and

3.c) (contd.)

$S^2 = \sum_{i=1}^{n} (x_i - \bar{x})^2$ are jointly sufficient for μ and σ^2

[$5\frac{1}{2}$ + $5\frac{1}{2}$ +8]=[19]

---

4. EITHER

a) State and prove Neyman-Pearson lemma for testing a simple hypothesis against a simple alternative.

b) Let $X \sim N(\mu, 4)$. Obtain the uniformly most powerful test of size α for testing $H_0: \mu = 1$ against $H_1: \mu > 1$ based on a random sample of size n from X.

[9+10]=[19]

OR

c) What is meant by saying that the mean of a sample differs significantly from a specified value say 100, at the 1% level of significance ?

d) Describe how you will test the hypothesis that the correlation coefficient is equal to zero against the alternative that it is not zero on the basis of a sample $(x_1, y_1), (x_2, y_2) \cdots (x_n, y_n)$ from a Bivariate Normal distribution.

---

5. For each of the following explain clearly with computational layouts how you test the corresponding hypotheses.

a) X and Y are jointly distributed as bivariate normal with mean μ_1 and μ_2, variances σ_1^2 and σ_2^2 and correlation coefficient ρ. Let (x_i, y_i), $i = 1, \cdots, n$ be a random sample of size n. Test $H_0: \sigma_0^2 = \sigma_1^2$ against $H_1: \sigma_0^2 \neq \sigma_1^2$.

b) Let $X_i \sim N(\mu_i, \sigma^2)$, $i = 1, \cdots, k$. Independent random samples of sizes $n_1, \cdots, n_k$ are available on $X_1, \cdots, X_k$ respectively. Test $H_0: \mu_1 = \cdots = \mu_k$ against H_1 : at least two μ_i's are different.

[9+10]=[19]

---

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

a) Confidence intervals

b) χ^2 - test for goodness of fit

c) Level and power function of a test

d) Test of significance for multiple and partial correlation coefficients.

e) Least square estimators. [3 x $6\frac{1}{3}$]=[19]

INDIAN STATISTICAL INSTITUTE
One-year Evening Course in Statistical Methods and Applications: 1977-78

PERIODICAL EXAMINATIONS

(Part II)

Paper VIb: Control Charts and Acceptance Sampling

Date: 26.4.78 Maximum Marks: 75 Time. 2 hours

Note: Qs. 1 and 2 are compulsory. Answer any one question out of questions 3 and 4. Distribution of marks shown in brackets [ ].

1.a) Explain the meaning of a process being in a state of statistical control. [6]

b) The following table gives the data on a measurable characteristic with specification limits 0.7538 and 0.7588 in some units. The sample size is 4.

Sample No.	$\bar{X}$	R	Sample No.	$\bar{X}$	R
1	0.7540	0.0011	9	0.7547	0.0007
2	0.7542	0.0014	10	0.7549	0.0015
3	0.7542	0.0009	11	0.7541	0.0017
4	0.7546	0.0010	12	0.7542	0.0010
5	0.7550	0.0008	13	0.7545	0.0011
6	0.7539	0.0009	14	0.7548	0.0009
7	0.7541	0.0012	15	0.7551	0.0012
8	0.7543	0.0011			

i) Set up an $\bar{X}$-R chart. (Draw the chart on a graph sheet). [15]

ii) Write a brief report on your findings. [3]

iii) Suggest some alternative courses of action which may be taken in the interest of the producer. [6]

2.a) Define the term (i) Acceptance Sampling (ii) AOQ, (iii) Screening. [8]

b) Compute the probabilities of acceptance for a lot of size 1000 having 2 % defective items under the following procedure: [6]

Take a random sample of 10 items and inspect the sample. If the sample contains either 0 or 1 defectives accept the lot otherwise reject the lot.

c) Give examples of industrial quality characteristics following Poisson law. [6]

d) Defects on an electronic component are recorded below:

No. of components inspected	No. of defects in all	No. of components inspected	No. of defects in all
4	17	6	18
7	23	7	28
5	24	7	29
7	27	6	31
7	32	8	39
7	33	6	29

Analyse the data by means of a control chart assuming that the samples were collected in the order of production. [10]

3. Explain how would you proceed to investigate whether a given process is capable to produce items satisfying known specification limits (Assume that the quality characteristic is measurable one). Write down the steps involved. [15]

4. Explain the procedure involved under percentage inspection. Compare sampling inspection plan and percentage inspection pointing out their relative merits and demerits. [15]

INDIAN STATISTICAL INSTITUTE

One-year Evening Course in Statistical Methods and Applications: 1977-78

PERIODICAL EXAMINATIONS

Design and Analysis of Experiments

12.4.78 Maximum Marks: 75 Time: 2 hours

Note: Answer as much as you can. Maximum you can get, however, is 75.

1. Write short notes on

i) randomisation, (ii) local control,
iii) least significant difference, and
iv) Latin square design. [5x4]=[20]

2. For an experiment observations were recorded as Y variables. Explain, how the transformation

$$X = a + bY, \quad b \neq 0$$

affects the analysis with X variables. [10]

3. Following table gives the one-week growth of certain plant at the time of flowering from 4 locations.

Location	Observations
A	15.0, 17.5, 11.5, 18.0, 14.0, 17.5
B	19.0, 21.5, 22.0
C	33.0, 27.0, 35.0
D	32.0, 28.0, 28.0

i) Write an appropriate linear model for analysis of the above data.

ii) Write the ANOVA table.

iii) Give estimates of parameters in the linear model.

iv) Give your findings by testing necessary hypotheses.

(table level = 0.05) [5+10+5+10]=[30]

4. Discuss how a randomised block design experiment without subsamples but with block-treatment interactions differ from a randomised block design experiment with subsamples but without block-treatment interactions. [20]

INDIAN STATISTICAL INSTITUTE

One-year Evening Course in Statistical Methods and Applications: 1977-78

Part II: PERIODICAL EXAMINATIONS

Paper V: Sample Surveys: Theory and Practical

Date: 29.3.78 Maximum Marks: 75 Time. 2 hours

Note: Do as many questions as you can. The paper carries 83 marks. The maximum you can score is 75. Marks allotted for each question are given in brackets [ ].

1. What do you mean by linear systematic sampling and circular systematic sampling ? Give unbiased estimator for population total in both cases. [15]

2. What do you mean by multistage sampling ? Give an unbiased sampling strategy for the populatior mean per ssu. Obtain variance of the sampling strategy. [15]

3. What do you mean by sampling and non-sampling errors ? Name some non-sampling errors and mention in brief how they creep into surveys. Give an unbiased sampling strategy for the population mean in the situation where some population units do not respond at least at the first enquiry. [10]

4. The following table gives number of fruits (y) on 16 guava trees.

Tree No.	1	2	3	4	5	6	7	8	9	10
No. of fruits	200	100	50	150	175	200	250	150	300	150

11	12	13	14	15	16
300	350	400	200	150	300

Give an unbiased estimate of total number of fruits on 16 trees

i) based on a linear systematic sample of 5 trees

ii) based on a circular systematic sample of 6 trees. [7+8]=[15]

5. A District is divided into 10 Divisions of 7 Subdivisions each. A srs wor of 3 Divisions is selected and then srs wor of 4 Subdivisions are selected from each sampled Division. The number of farm labourers is given below for the sampled Subdivisions.

Sampled Division	Number of farm labourers in the sampled subdivisions			
D_1	1200	2000	2500	1500
D_2	1000	1500	1100	1200
D_3	3000	4200	3500	3200

i) Give an unbiased estimate for total farm labourers in the District

ii) Give an unbiased estimate for average number of farm labourers per Subdivision.

iii) Give an unbiased estimate for average number of farm labourers per Division.

iv) Give an unbiased estimate for average number of farm labourers per Subdivision per Division.

v) Give unbiased estimate for the variance of the estimator in (ii) above. [5+3+3+3+9]=[23]

6. Define ratio, regression, difference and product estimators for population mean when SRS WOR is adopted. [5]

INDIAN STATISTICAL INSTITUTE
One-year Evening Course in Statistical Methods and Applications: 1977-78

PART II PERIODICAL EXAMINATIONS

Paper IV: Methods of Statistical Inference

Date: 8.3.78 Maximum Marks: 75 Time: 2 hours

Note: Answer all questions.

1.a) Describe clearly how you test the equality of means of k normal populations with same variance based on independent random samples of sizes $n_1, \ldots, n_k$ respectively.

b) The following data relate to lives in hours of sample lamps taken from four batches of electric lamps. Test whether the batches differ among themselves in average length of life.

Life of electric lamps in hours

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

[10+25]=[35]

---

2. The following table gives the scores of 10 workers working in a match factory in a psychological test X and also their efficiency index, Y. Test whether the correlation coefficient between the score in the psychological test and the efficiency index is significantly different from zero.

Sl. No.	Score in test X	Efficiency index Y
1	15	47
2	77	67
3	18	38
4	8	30
5	16	26
6	54	44
7	95	62
8	22	40
9	69	67
10	75	72

[25]

---

3. Measurements on the length of hind femur of males of two species of desert locusts were taken and the following computations were made.

Means and Standard deviations of lengths of hind femur of two species of desert locusts.

Species	Sample size	Length of hind femur in mm. mean	Standard deviation
Gregaria	567	24.21	1.15
Kakko Swarm	852	24.92	1.64

INDIAN STATISTICAL INSTITUTE
One-year Evening Course in Statistical Methods and Applications: 1977-78

Part II. PERIODICAL EXAMINATIONS
Paper V-A: Sample Surveys, Theory and Practical

Date: 1 March 1978 Maximum Marks: 75 Time: 2 hours

Note: Do as many questions as you do. The paper carries 80 marks. The maximum you can score is 75. The marks allotted to each question are given in brackets [ ].

1. What do you mean by simple random sampling? Give an unbiased estimate of population total of a character y based on simple random sample wor. Obtain expression for the standard error of the estimate. [20]

2. What do you mean by proportional allocation and Neyman allocation? Give an unbiased estimate for the population mean of a character y based on a simple stratified sample. Obtain the expression for variance under proportional allocation and Neyman allocation. [15]

3. There are N households in Baranagore. The average size of household X is known. Give a sampling strategy to estimate the total number of unemployed in the population using the knowledge of X. Obtain mean square error of the estimator and an estimate of the mean square error. [20]

4. The following table gives the data on census population (x) and cultivated area (y) in acres for a population of 5 villages.

Village	(x)	y
1	110	320
2	175	450
3	180	413
4	250	1015
5	85	375

a) Draw a simple random sample of size 2. Estimate the total cultivated area Y using

i) simple random estimater

and ii) difference estimator.

For difference estimator you may take $\lambda = 3$. Obtain variance of the estimate and estimate of the variance.

b) Draw a sample of 2 villages with PPSWR, size being x. Estimate the total cultivated area Y. Obtain variance of the estimate. Also obtain estimate of the variance.

c) Compare the relative efficiency of the sampling strategy in (b) with other sampling strategies in (a).

[12+10+3]=[25]

INDIAN STATISTICAL INSTITUTE
One-year Evening Course in Statistical Methods and Applications: 1977-78

PART II: PERIODICAL EXAMINATIONS

PAPER IV: Methods of Statistical Inference Theory and Practical

Date: 8.2.78 Maximum Marks: 75 Time: 2 hours

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

1. Explain each of the following terms clearly with the help of suitable examples:

a) composite hypothesis
b) critical region
c) power of a test. [3x4]=[12]

2. Let $X \sim N(0, \sigma^2)$. Let x_1, x_2, x_3 and x_4 be a random sample of size of from X. Write down the distribution of each of the following statistics. (You need not derive.)

a) $\frac{1}{\sigma^2}(x_1^2 + x_2^2 + x_3^2)$

b) $\dfrac{x_4}{[\frac{1}{3}(x_1^2 + x_2^2 + x_3^2)]^{1/2}}$

(c) $\dfrac{x_1^2}{x_1^2 + x_2^2 + x_3^2}$.

(d) $\dfrac{x_1^2 + x_2^2}{x_3^2 + x_4^2}$. [4x4]=[16]

3.a) State and prove Neyman-Pearson lemma for testing a simple hypothesis against a simple alternative.

b) Let X be a random variable taking two values 1 and 0 with probabilities p and (1 - p) respectively ($0 < p < 1$). Let $x_1, \cdots, x_5$ be a random sample of size 5 from X.

i) Obtain the distribution of $\sum_{i=1}^{5} x_i$.

ii) Obtain the most powerful test of size 0.05 for testing $p = 0.2$ against $p = 0.1$. What is the power of this test ? [10+8+5+5]=[28]

4.a) Let $X \sim N(\mu, \sigma^2)$ where σ^2 is unknown. Let $x_1, \cdots, x_n$ be a random sample of size n from X. How do you test $H: \mu = \mu_0$ against $K: \mu \neq \mu_0$?

b) A random sample of 16 bulbs was chosen from a lot and the life of these bulbs (in hours)were observed. They are given below:

308, 320, 276, 340, 282, 263, 272, 312, 318, 276, 283, 287, 307, 302, 288, 270.

Examine whether the average life of a bulb differs significantly from 300 hours. State your assumptions.

[6+13]+[19]

INDIAN STATISTICAL INSTITUTE

One-year Evening Course in Statistical Methods and Applications: 1977-78

Part I Supplementary Examination

CALCUTTA AND DELHI

Paper Ib: Topics in Applied Statistics:

Index Numbers and Time Series Analysis

Date: 25.1.78 Maximum Marks.100 Time:$2\frac{1}{2}$ hours

Note: Answer Groups A and B in separate answerscripts. Marks allotted for each question are given in brackets.

Group A. Index Number

Answer any two questions. Max. Marks. 50

1. Define a 'true' cost of living index. Show that the Paasche's index underestimates the 'true' cost of living index. [5+20]=[25]

2.a) Define a 'chain' index. Suppose a series of 'fixed base' indices, $P_{01}, P_{02}, \ldots, P_{0n}$, is available. Under what condition this series will be independent of the choice of the base period? [5+8]=[13]

b) The table below gives the retail prices of the group 'foodgrains and products' in two centres: Calcutta and Bombay in 1960. The column (2) gives the average family expenditures (of industrial workers) in Calcutta in 1960. Using this column as weight calculate the consumer price index number of 'foodgrains and products' for Bombay, taking Calcutta as base. [12]

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

3.a) Write a note on the consumer price index number. [15]

b) The all commodity wholesale price index (WPI) in India increased by 88.4 per cent in 1971-1972 over 1961-62. The WPI of different groups for 1971-72 (base 1961-62 = 100) are given in the following table except that of food articles:

Group	Weight	WPI for 1971-72 (base 1961-62 = 100)
Liquor and Tobacco	25	194.8
Fuel, Power, Light and Lubricants	61	172.1
Industrial Raw Materials	121	191.0
Chemicals	7	197.0
Machinery and Transport equipment	79	159.0
Manufactures	294	167.1
Food articles	413	

Calculate WPI of food articles for 1971-72 with 1961-62 as base. [10]

p.t.o.

Group B: Time Series Analysis

Max. Marks. 50

Note. Answer either Q.4 or Q.5 and either Q.6 or Q.7.

4.a) Describe briefly the various components that are usually present in an economic time series. Give examples.

b) How would you decide whether an additive model or a multiplicative model would be appropriate for analysing an observed time series which contains a trend, seasonality and irregular movement. [15+9]=[24]

5. Write notes on any three of the following.

a) Moving average method
b) Logistic curve
c) Constant and moving seasonal indices
d) Method of Least Squares in trend analysis

6. The following data relate to the production of pig iron in the U.S.A. during 1855-1895.

Year	Production (000 tonnes)	Year	Production (000 tonnes)
1855	700	1880	3835
1860	821	1885	4045
1865	832	1890	9203
1870	1665	1895	9446
1875	2024		

Fit a Gompertz trend curve to the above data by the method of group sums. Also, estimate the production for 1915. [21+5]=[26]

7. Calculate the seasonal indices for the following quarterly time series data on average weekly earnings by the method of ratio to moving average.

Year	Quarter I	Quarter II	Quarter III	Quarter IV
1951-52	13.27	14.52	13.72	14.59
1952-53	15.14	16.18	14.70	15.77
1953-54	16.03	16.93	15.53	16.55
1954-55	16.79	17.63	16.27	17.81
1955-56	17.94	18.83	17.54	18.00
1956-57	18.86	19.92	18.35	19.53
1957-58	19.47	20.59	18.68	19.93

[26]

INDIAN STATISTICAL INSTITUTE
One-year Evening Course in Statistical Methods
and Applications 1977-78

Part I Supplementary Examination (CALCUTTA and DELHI)

Paper IIa. Descriptive Statistica Practical

Date. 25.1.78 Maximum Marks. 75 Time. 2 hours

Note: Attempt as much as you can. Maximum score can be 75. Marks allotted for each part are mentioned in brackets.

---

1. Calculate standard deviation, variance and median for the following data.

Variate values	950-1000	1000-1050	1050-1100	1100-1150	1150-1200	1200-1250	1250-1300
Frequencies	19	20	18	18	12	11	7

What can you say about the skewness of the distribution. [4x4]=[16]

---

2. Fit a Poisson distribution to the following data which gives the number of yeast cells per square for 400 squares.

No. of cells per square	0	1	2	3	4	5	6	7	8	9	10
No. of squares	103	143	98	42	8	4	2	0	0	0	0

[20]

3.a) In a precision bombing attack there is a 50 % chance that any one bomb will strike the target. Two direct hits are required to destroy the target completely. How many bombs must be dropped to give a 99 % chance or better of completely destroying the target ? [10]

b) For a normal distribution whose mean is 10 and variance 9, find the value of the variate in the r.h.s. of its mean such that the probability of the variate from mean to that point is .4485. [10]

---

4. Following are the observations on three variables X, Y and Z

X	1	2	3	5	7
Y	7	8	10	11	11
Z	14	12	8	6	5

Predict Z when X = 6 by least square methods. Calculate the partial correlation coefficient between X and Z when the linear effect of Y is removed. [10+15]=[25]

INDIAN STATISTICAL INSTITUTE
One-year Evening Course in Statistical Methods
and Applications 1977-78

Part I Supplementary Examination

Paper IIa: Descriptive Statistics Theory

Date: 18.1.78 Maximum Marks. 75 Time: 2 hours

Note. Attempt as many questions as you can.
Maximum score can be 75.
Marks allotted for each part are mentioned in brackets.

---

1. Discuss the effect of the changes of location and scale on the following measures -

Mean, Median, Standard Deviation and Mean Deviation. [16]

---

2. Define Poisson distribution. Give two examples where Poisson distribution is a suitable approximation.

If the two independent random variables X_1 and X_2 have Poisson distributions with means m_1 and m_2 respectively, find the distribution of $X_1 + X_2$. [20]

---

3. Calculate mean deviation of the normal distribution having mean 5 and variance 9. [8]

Establish a recurrence relation between the central moments of N(0, 1). [10]

---

4. What is the effect of the changes of origin and scale on regression and correlation coefficients.

Define correlation ratio. What is it a measure of ? [4x4]=[16]

---

5. Obtain expressions of multiple and partial correlation coefficients in terms pairwise correlation coefficients when there are total three random variables under consideration. [10+10]=[20]

INDIAN STATISTICAL INSTITUTE

One-Year Evening Course in Statistical Methods and Applications: 1977-78

CALCUTTA AND DELHI

Part I: Final Examinations

Paper III b: Probability

Date: 16.12.1977 Max.Marks:100 Duration: 2½ hours

Note: Answer as many questions as you can. The paper carries 140 marks, the maximum you can score is 100.

---

1. (a) A deck of cards is dealt into four hands containing 13 cards each. What is the probability that (i) at least one hand has a complete suit of cards, (ii) each hand has an ace.

   (b) Find the probability that while placing r balls at random in n cells, a specified cell contains exactly k balls. [8+5+7=20]

---

2. (a) Express the following statements in terms of two events A_1 and A_2 defined on a sample space S, for k=0,1 and 2:

   (i) at least k of the events A_1 and A_2 occur;

   (ii) at most k of the events A_1 and A_2 occur.

   (b) Let A and B be two events defined on a sample space S. Show that the formula for the probability that exactly one of the two events occur is given by

   $$P[(AB')\cup(BA')] = P(A)+P(B)-2P(AB)$$

   (c) Define independence of (i) two events, (ii) three events.

   (d) Given that the events A and B and C are independent. Prove or disprove that AUB and C are independent.

   [(3+3)+5+5+4=20]

---

3. (a) State and prove 'Bayes' theorem of inverse probability.

   (b) An urn contains g green and r red balls. One ball is drawn at random. It is replaced and c more balls of the same colour are added to the urn, c being a given positive integer. A ball is now drawn at random from the urn. Find the conditional probability that the first ball is red given that the second ball is green.

   [8+12=20]

---

4. (a) Define the probability distribution on a discrete sample space.

   (b) A fair die is rolled once. Let X be thrice the score that appears and Y assumes value 0 or 1 according as the score is odd or even respectively. Derive joint probabilities $p(x,y)$ and marginal probabilities $p(x)$ and $p(y)$.

   (c) Define the moment generating function of a random variable.

   (d) Derive the moment generating function for a poisson random variable.

   [4+(6+3)+(2+5)=20]

---

5. (a) Given the joint probability density function $f(x,y)=xy\ e^{-(x+y)}$, $x > 0$, $y > 0$.

   Calculate $P(x < 1, y < 1)$. Find marginal distributions of x and y.

   (b) If two random variables x and y assumes only two values each, and if $Cov(x,y)=0$, then x and y are independent.

   [(4+6)+10=20]

---

6. Suppose an urn contains pn white and qn black balls, the total number of balls being n, and that a random sample of r balls is taken one by one without replacement. Let

   x_i = 1 if the ith ball drawn is white
   = 0 if the ith ball drawn is black.

   (a) show that $E(x_i) = p$, $Var.(x_i) = pq$.

   (b) show that the covariance between x_j and x_k $(j \neq k)$ is $-pq/(n-1)$

   (c) From (a) and (b), obtain the variance of $S_r=x_1+x_2\ldots+x_r$.

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

---

7. (a) State and Prove Tchebychev's theorem.
   (b) State the weak Law of large numbers.
   (c) State the central limit theorem. Hence or otherwise show that a binomial distribution can be approximated by a normal distribution for large number of trials.

   [7+3+(3+7)=20]

INDIAN STATISTICAL INSTITUTE

One-year Evening Course in Statistical Methods and Applications: 1977-78

CALCUTTA AND DELHI

Part I Final Examinations

Paper IIIa: Basic Mathematics

Date: 14.12.77 Maximum Marks: 50 Time: $1\frac{1}{2}$ hours

Note: Answer as many questions as you can.
The maximum you can score is 50.
Marks allotted for each question are given in brackets [ ].

1.a) A candidate is required to answer 6 out of 10 questions which are divided into 2 groups, each containing 5 questions, and the candidate is not permitted to attempt more than 4 questions from any group. Prove that the candidate can make up the choice in exactly 200 ways. [7]

b) If $(1+x)^n = C_0 + C_1 x + C_2 x^2 + \ldots + C_n x^n$, prove that

$$C_2 + 2C_3 + \ldots + (n-1)C_n = 1 + (n-2)2^{n-1}.$$ [5]

2.a) Prove that $\lim_{n \to \infty} \frac{n!}{n^n} = 0$. [5]

b) Let f be the function defined on the real line as:

$$f(x) = x - [x]$$

where $[x]$ denotes the largest integer less than or equal to x. Draw the graph of the function. Find the continuity points and the discontinuity points of f. Justify your answer. [7]

3.a) Find the derivative with respect to x of

$$\alpha^{\sin x} (\log x)^2$$

where $\alpha > 0$ and $x \geq 2$. [8]

b) Given n fixed positive real numbers $a_1, \ldots, a_n$ determine x so that

$$\sum_{i=1}^{n} (a_i - x)^2 \quad \text{is a minimum.}$$ [6]

---

4. Find the following integrals:

a) $\int_{-\pi}^{\pi} \cos mx \cos nx \, dx$, m, n are integers.

b) $\int_2^3 \frac{1}{1-x^2} \, dx$.

c) $\int x^\alpha \log x \, dx$, α is an integer. [6+6+6]=[18]

5.a) State which of the following relation is true and which is false. Justify your answer in each case.

For any three subsets A, B and C of an universal set we have

i) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

ii) $A - (B - C) = (A - B) - C$. [5+5]=[10]

b) Define the product of two square matrices of the same order. Given an example of two nonzero square matrices of order 3 such that their product is the zero matrix. [6]

INDIAN STATISTICAL INSTITUTE

One-year Evening Course in Statistical Methods and Applications: 1977-78

CALCUTTA AND DELHI

Part I Final Examinations
Paper II(ii): Descriptive Statistics Practical

Date: 12.12.77 Maximum Marks: 75 Time: 2 hours

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

1. The weights in grams of 30 apples picked out at random from a consignment are as follows:

106, 107, 76, 82, 109, 107, 115, 93, 95, 123,
125, 111, 92, 70, 126, 129, 119, 115, 90, 84,
107, 104, 82, 92, 87, 103, 96, 112, 105, 81.

Form the grouped frequency table by dividing the variate range into intervals of equal width. Draw the frequency curve and comment on its skewness. Obtain the median and the two quartiles. [20]

2.a) Assuming that 30 % of the people in a certain city are suffering from a certain stomach disease and assuming that 60 investigators each took 12 individuals to see whether they have the disease, how many investigators do you expect to report that 4 or less were affected by the disease? [7]

b) In a certain factory turning out small-sized screws, there is a chance 1/25 for any screw to be defective. The screws are supplied in packets of 100. The manufacturer announces that if a packet contains more than 10 defectives, he will refund the price of the packet. Out of 200 packets, for how many packets on the average do you think that he will have to refund the price? [8]

3. The following is the frequency distribution of the right hand grip of 400 European males.

Right hand grip (in lb.)			Frequency
29.5	-	39.5	2
39.5	-	49.5	3
49.5	-	59.5	16
59.5	-	69.5	60
69.5	-	79.5	106
79.5	-	89.5	110
89.5	-	99.5	75
99.5	-	109.5	21
109.5	-	119.5	6
119.5	-	129.5	1

Find the expected frequencies for the above classes, assuming that the population distribution of right hand grip is approximately normal. [20]

4. The following marks have been obtained by a class of students in statistics.

Paper I (x):	45	55	56	58	60	65	70
Paper II (y):	56	50	48	60	62	64	67
Paper III (z):	35	36	40	42	43	46	46

(a) Find the equation of the line of regression of y on x. [10]

(b) Calculate the multiple correlation coefficient between z and x, y. [10]

5. The following data give the distribution of monthly income of 610 families selected randomly from a certain city. Draw the concentration curve for the data and comment on the income distribution of the city on its basis.

Monthly income (in Rs.)	Frequency
Less than 100	60
100 - 150	83
150 - 200	124
200 - 250	209
250 - 300	73
300 - 350	31
350 - 400	15
400 - 450	9
450 - 500	4
500 and more	2

[15]

INDIAN STATISTICAL INSTITUTE

One-year Evening Course in Statistical Methods and Applications: 1977-78

Calcutta and Delhi

Part I Examinations

Paper IIa: Descriptive Statistics Theory

Date: 9.12.1977 Marks: 75 Duration: 2 hours

Note: Answer as many questions as you can. Maximum you can score is 75 out of 97.

---

1. Discuss briefly the different considerations, to be kept in view, in drawing up a frequency distribution for data on continuous variable. [6]

---

2. Show that (a) the mean square deviation is minimum when measured from the mean and (b) the mean absolute deviation is minimum when measured from the median. [5+7]

---

3. (a) Determine the mode of binomial distribution.

(b) Show that for the binomial distribution B(n,p)

$$\mu_{r+1} = p(1-p)\left[n\, r\, \mu_{r-1} + \frac{d\,\mu_r}{dp}\right] \qquad [5+10]$$

---

4. Define the poisson distribution with parameter λ. Derive the mean and Variance for this distribution. Give two examples of situations where this model is appropriate. [10]

---

5. Show that all odd order central moments for the normal distribution vanish, where as the central moments of even order are given by

$$\mu_{2k} = 1,3,5.....(2k-1)\,\sigma^{2k} \quad k=0,1,2. \qquad [10]$$

---

6. (a) Define the concept of correlation coefficient 'r', using an example. Write down the formula you would use for evaluating r^2 from data given on two variables x and y.

(b) How would you interpret the following results:

(1) $r^2=0$, (2) $r^2=1$ (3) $\eta^2_{yx}=0$ (4) $\eta^2_{yx}=1$

(c) What is the significance of the discrepancy $\eta^2_{yx} - r^2$.

[5+5+2]

7. (a) Define the term regression. Given a set of observations on two variables x and y, how would you determine the regression lines for x and y?

(b) For 20 army personnel, the regression of weight of brains (y) on weight of heart (x) both measured in grams is $Y=0.4\ X + 6.9$, and the regression of weight of heart on the weight of brains $X = 1.2\ Y - 2.5$.

Find the correlation coefficient between the two variables and also their mean values.

[7+5]

---

8. Define clearly the coefficient of multiple correlation between the variable x_1 and the two variables x_2 and x_3.

Show that $R^2_{1(23)} = \dfrac{r^2_{12} + r^2_{13} - 2r_{12}\ r_{23}\ r_{31}}{1 - r^2_{23}}$ [5+5]

---

9. Define the partial correlation coefficient $r_{12.3}$ and express it in terms of the (total) correlation coefficients among the variables x_1 x_2 and x_3.

Prove the identity $b_{12.3}\ b_{23.1}\ b_{31.2} = r_{12.3}\ r_{23.1}\ r_{31.2}$

[5+5]

INDIAN STATISTICAL INSTITUTE

One-year Evening Course in Statistical Methods and Applications 1977-78

Part I Final Examinations

CALCUTTA and DELHI

Paper Ib: Topics in Applied Statistics:

Index Numbers and Time Series Analysis

Date: 7.12.77 Maximum Marks: 100 Time: $2\frac{1}{2}$ hours

Note: Answer Groups A and B in separate answerscripts.
Marks allotted for each question are given in brackets [ ].

Group A: Index Numbers

Maximum Marks: 50

Note: Answer any two questions.

1.a) Examine the following statement: "Laspeyres' index is inferior to Fisher's Ideal index, but it is popular compared to either Fisher's Index or Paasche's index". [8]

b) What are the different types of errors involved in the construction of an index number? [9]

c) The table below gives Laspeyres' price index numbers for three periods 1, 2 and 3. (For example, the figure in the second row and first column gives Laspeyres' price index for period 1 with period 3 = 100):

base period	current period 1	2	3
2	97.7	100.0	119.2
3	82.7	84.6	100.0
4	79.8	81.7	96.5

using Paasche-type link indices, calculate the chain index for period 4, taking period 1 as the base period. [8]

---

2. What is an index of industrial production? How is this index constructed? [25]

3.a) Write a note on factor reversal test, indicating its importance. [9]

b) From the following data relating to middle class Bengalees in Calcutta in the income bracket 101 - 150 rupees per month, compute a cost of living index for January 1949 with 1939 as base. [16]

	Percentage of total expenditure spent on the item	Unit	Price in Rs. 1939	January 1949	Price index for Jan.'49 (1939 = 100)
FOOD					
food grains	23.44				357.7
fish, egg, meat	8.55				361.0
milk etc.	7.40				383.5
salt, spices, vegetables etc.	11.26				366.8
oil	4.72	seer	0.48	2.14	
sugar and gur	2.30	,,	0.29	0.70	
tea and tiffin	4.71	,,	0.69	2.76	
RENT and TAX	9.81	...	....	....	110.0
FUEL and LIGHT					
coal	3.46	md.	0.38	1.69	
other fuel	3.18	...	....	....	364.7
CLOTHING	6.43	...	....	....	472.6
MISCELLANEOUS	14.74	...	....	....	272.6

p.t.o.

Group B: Time Series Analysis

Maximum Marks: 50

Note: Answer either Q.4 or Q.5 and either Q.6 or Q.7.

4.a) Describe briefly the various components that are usually present in an economic time series. Give examples.

b) Explain the implications of the additive and the multiplicative models for decomposition of an observed time series. [14+10]=[24]

5. Write notes on any three of the following:

a) correlogram
b) cyclical fluctuations
c) constant and moving seasonal indices,
d) growth curves. [8 x 3]=[24]

6. The following data give the population figures of the U.S.A. recorded through decennial consuses.

Year	Population (millions)	Year	Population (millions)
1860	31.40	1910	91.97
1870	38.60	1920	105.71
1880	50.16	1930	122.78
1890	62.95	1940	131.40
1900	76.00		

Fit a logistic trend curve to the above data and estimate the population expected in 1978. [21+5]=[26]

7. The operating season for a fruit processing unit is from May to September each year. The following data relate to the monthly production (000 kg.) from 1968 to 1970.

Year	May	June	July	August	September
1968	75	86	102	105	90
1969	83	89	110	115	92
1970	84	95	113	118	89

Find out the seasonal indices for this data by ratio-to-moving average method and calculate the deseasonalized figures of production for the August 1970. [22+4]=[26]

INDIAN STATISTICAL INSTITUTE

One-Year Evening Course in Statistical Methods and Applications: 1977-78

CALCUTTA AND DELHI

Final Examination Paper 1A

(Collection and Presentation of Data)

Date:5.12.1977 Max.Marks:50 Time: 1½

Note: Answer all the questions

---

1. The index number of wholesale prices registered a very sharp increase during the period 1970-71 to 1976-77. With 1970-71 as 100, all commodity index rose to 139.7 in 1973-74 and to 174.9 in 1974-75. It showed a slight fall in 1975-76 with a figure of 173.0 but again rose to 176.6 in 1976-77. Among different groups, the maximum increase took place in 'minerals' group which increased to 225.4 in 1973-74 and to 423.5 in 1974-75, reflecting the sharp increase in prices of crude oil in the two years. In the subsequent years the price index showed relatively modest growth with the figures 440.4 in 1975-76 and 449.4 in 1976-77. The 'non-food articles' groups registered marginal increase in the wholesale prices for which the index rose to 163.7 in 1974-75 and to 167.4 in 1976-77. In fact it declined to 139.8 in 1975-76 which was even lower than the 1973-74 level of 146.6. The groups 'food articles' and 'manufactured products' showed different trends. The price index for the former group increased from 136.6 in 1973-74 to 172.1 in 1974-75 but declined to 163.6 in 1975-76 and to 155.3 in 1976-77. However the later group showed continuous increase from 139.5 in 1973-74 to 168.8 in 1974-75, to 171.2 in 1975-76 and 175.2 in 1976-77. The price index for 'fuel and light' group also showed continuous increase with figures 130.6 in 1973-74, 198.3 in 1974-75, 219.2 in 1975-76 and 230.8 in 1976-77.
(Source: Revised index of wholesale prices, March 1977) Ministry of Industry.

(a) Present the above information in a suitable table. [10]

(b) Present the above data as tabulated in (a) in a suitable diagramatic form. [10]

---

2. Describe briefly different methods that are employed in the collection of statistical data. [7]

OR

Give an account of different modes of diagramatic representation of statistical data. [7]

---

3. Explain the difference between any three of the following:

(a) parameter and statistic
(b) variable and attribute
(c) sampling and non-sampling errors
(d) statistical population and sample

[3x3=9]

---

4. (a) Name the important government organisations that are responsible for collecting any four of the following statistics:

(a) National income
(b) Population and demography
(c) Industrial production
(d) Agriculture production
(e) Foreign trade
(f) Banking and finance

[1x4=4]

(b) Mention the name of one important publication for each of the ones you choose in Q.4(a) and describe very briefly the nature of the data contained therein (Do not give the same publication more than once).

[2½x4=10]

INDIAN STATISTICAL INSTITUTE

One-year Evening Course in Statistical Methods and Applications: 1977-78

Part I PERIODICAL EXAMINATION

Paper II: Descriptive Statistics Theory and Practical

Date: 9.11.77 Maximum Marks: 75 Time: 2 hours

Note: Attempt as many questions as you can. Maximum you can score is 75. Marks allotted for each question are given in brackets [ ].

1. Explain the concept of concentration curve. What does the coefficient of concentration measure? [5+3]=[8]

2.a) Six dice are thrown 729 times. How many times do you expect at least three dice to show a five or six? [7]

b) Obtain an expression for the mode of the binomial distribution. [7]

3.a) Establish a recurrence relation among the moments of poisson distribution. [7]

b) Calculate mean deviation of poisson distribution with mean one. [7]

4.a) Obtain first four moments of $N(\mu, \sigma^2)$ and hence show that $\mu_4 / \mu_2^2 = 3$ where μ_i stands for i-th moment of $N(\mu, \sigma^2)$ about mean.

b) For a normal distribution with mean 2 and standard deviation 3, find a point in right side of the mean such that the probability that corresponding random variable lies between mean and that point is 0.4280. [9]

5.a) What is the correlation coefficient a measure of? Obtain the correlation coefficient between two random variables X and Y if they are related as

$$3X + 5Y = 9.$$ [3+5]=[8]

b) Following are the observations on two variables X and Y.

X :	78	89	97	69
Y :	125	137	156	112

Predict Y when X = 83 using least square method. [10]

6. An experiment consisting of 100 Bernoulli trials is repeated 50 times. Fit a suitable distribution to the obtained data, given below, assuming that the probability of success in a trial is very small.

No. of success	0	2	2	3	4	5	6 and above
Frequency	14	18	11	5	1	1	0

[15]

INDIAN STATISTICAL INSTITUTE

One-year Evening Course in Statistical Methods and Applications: 1977-78

PART I PERIODICAL EXAMINATIONS

Paper 1b : Topics in Applied Statistics

Date: 2.11.77 Maximum Marks: 100 Time: $2\frac{1}{2}$ hours

Note: Answer any two questions from Group A and all questions from Group B. Marks allotted for each question are given in brackets [ ]. Answer Groups A and B in separate answerscripts.

Group A: Index Numbers

Maximum Marks: 50

1. Define a 'true' cost of living index. Show that the Laspeyres' index overestimates the true cost of living index. [25]

2.a) Suppose you are asked to construct a cost of living index number for industrial workers in Asansol. How would you select various commodities and obtain their weights? [14]

b) Construct an index of production in food manufacturing industries in India for 1958, (with 1956 = 100), using the statistics given below and taking as weights the values added by different industries in 1956. [11]

Industry	Value added (Rs. crores) in 1956	Unit	Monthly average production 1956	1958
Wheat flour	2	thousand tons	44.8	68.3
Biscuits	1	tons	1235	1475
Sugar	33	thousand tons	155	167
Confectionery	0.1	tons	790	863
Vanaspati	4	thousand tons	21	25
Tea	55	million lbs.	56	60
Coffee	2	hundred tons	29	37
Salt	5	lakh maunds	74	94

3.a) What is a 'chain index'? Compare it with a 'fixed-base index', noting their relative merits and demerits. [5+9]=[14]

b) The following table gives the cost of living index (CLI) numbers for the groups and subgroups of items with their respective weights for the middle class people of Calcutta in 1956 (with 1939 = 100):

Group/ subgroup	Subgroup weight	Group weight	CLI for 1956 (1939 = 100)
1. Food		67	415.3
2. Fuel and Light		8	
2.1 Kerosene	27		252.8
2.2 Firewood	20		410.0
2.3 Coal	40		402.1
2.4 Matches	13		320.0
3. Clothing		10	882.9
4. Rent		5	316.5
5. Miscellaneous		10	441.6

Mr. X was getting a monthly income of Rs.250/- in 1939 and Rs.445/- in 1956. State how much he ought to have received as extra allowance in 1956 to maintain his pre-war standard of living. [11]

P.t.o.

Group B: Time Series Analysis

Maximum Marks: 50

4.a) Discuss briefly the alternative methods of trend analysis of a given observed time series.

b) Suppose an estimated trend equation is given by $T_t = 0.58 + 0.738t$ where $t = 0$ for January 1965 and unit of measurement of t is one year. Estimate the trend value for June 1972. [12+5]=[17]

5. Suppose Y_t, $t = 1,2,\ldots, T$ is an observed time series. What form of trend equation would you choose in each of the following cases?

i) First differences of logarithms of Y_t, $\Delta \log Y_t$, a are approximately constant.

ii) Logarithms of first differences of Y_t, $\log \Delta Y_t$ are approximately linear in t and Y_t tends to a finite maximum as t becomes large.

iii) The graph of rate growth, $\Delta Y_t / Y_t$, when plotted against t resembles a modified exponential curve.

iv) The graph of ~~rate of~~ growth, $\Delta Y_t / Y_t$ ~~against~~ Y_t is approximately linear. [4 x 4]=[16]

3. Give a brief account of the method of ratio-to-moving average for calculating seasonal indices. [17]

INDIAN STATISTICAL INSTITUTE

One-year Evening Course in Statistical Methods and Applications: 1977-78

Part I: PERIODICAL EXAMINATION

Paper IIIb: Probability

Date: 5.10.77 Maximum Marks: 100 Time: $2\frac{1}{2}$ hours

TEST BOOKLET

Instructions

Answer all the questions. The answer to each question must be written down within the space provided below the question. Rough work must be done in the blank pages on the left side of the booklet. Marks allotted to each question are indicated in brackets against the question.

1. Write down the sample space of the following random experiments.

a) A coin is tossed thrice and the sequence of heads and tails that results is observed.

[5]

b) A die is rolled. If the outcome is odd then a coin is tossed and if the outcome is even the die is rolled again and the outcomes at both stages are observed.

[3]

2.a) Consider the sample space of 1(a). Let A denote the event 'exactly one tail occurs' and let B denote the event 'at least one head and at least one tail occur'. Write down the events A and B.

A =

B =

[6]

b) Consider the sample space of 1(b). Write down the events.

C = {Even number in the first throw of the die.}
D = {The score in the first throw of the die ≥ 4}

C =

D =

[8]

3.a) Toss a fair coin three times and observe the number of heads. Write down the resulting probability space.

[6]

b) Let 2 items be drawn from a lot containing 12 items of which 4 are defective. Let

A = {both items are defective}
B = {both items are nondefective}
C = {at least one item is defective}

Find (i) P(A) (ii) P(B) and (iii) P(C).

i) P(A) =

ii) P(B) =

iii) P(C) =

3.c) Let $S = \{a_1, a_2, a_3, a_4\}$.
Let P be a function defined on S by $P(a_1) = 1/2$, $P(a_2) = 1/3$, $P(a_3) = 1/4$ and $P(a_4) = 1/5$. Is P a probability function on S?

[4]

4.a) Let B be an event with $P(B) > 0$. Let A be any event. Define $P(A \mid B)$, the conditional probability of A given B.

[3]

b) Show that

$$P(A^C \mid B) = 1 - P(A \mid B).$$

[7]

c) A box contains three coins, two of them fair and one two-headed. A coin is selected at random and is tossed twice. If both tosses result in heads then what is the probability that the coin is two-headed.

[10]

5. Let a fair coin be tossed twice.

Let A = {first toss is a head}
B = {second toss is a head}
C = {exactly one toss is a head}

Are the events A, B and C independent ?

6. A fair die is rolled. Let X denote twice the number appearing and let Y be equal to 1 if an odd number appears and 3 if an even number appears. [10]

i) Find the joint distribution of (X, Y) and the marginal distributions of X and Y.

[10]

6. ii) Find the distribution of $X + Y$.

[5]

iii) Verify that $E(X+Y) = E(X) + E(Y)$.

[5]

iv) Find cov (X, Y).

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods and Applications: 1977-78

Part I: PERIODICAL EXAMINATIONS

Paper II: Descriptive Statistics Theory and Practical

Date: 28.9.77 Maximum Marks: 75 Time: 2 hours

Note: Attempt four questions including Q.1 and Q.5 which are compulsory. Marks allotted for each question are given in brackets [ ]

1. The mean and standard deviation of a set of 25 sample observations were computed as 56 inches and 2 inches respectively. It was detected later that the set included an observation which was wrongly recorded as 64 inches. Exclude this incorrect observation from the set, and compute the mean and S.D. for the remaining 24 observations. [5+10]=[15]

2.a) Describe, briefly, the essential features of the various diagrams which are used for graphical representations of frequency distributions and cumulative frequency distributions. Draw a neat sketch of each diagram.

b) Show that the 'ogives' of a continuous distribution intersect each other at a point whose abscissa is the median value of the distribution. [16+4]=[20]

3.a) State the desirable qualities of a good 'average' of a distribution.

b) Derive the formula for estimating the sample mode from the frequency table of a continuous variable.

c) What can be said about the relative magnitudes of the mean, median and mode of a distribution? [7+7+6]=[20]

4.a) What is 'dispersion' of a distribution? Describe its various measures. (The expression for each measure in terms of the sample values should be stated.)

b) Prove that the 'mean deviation' of a distribution can never exceed its standard deviation. [14+6]=[20]

5. EITHER

The distribution of breaking strength (measured to the nearest thousand pounds per square inch) of 80 test pieces of a certain alloy, is given in the following table:

breaking strength (thousand lb./sq. in.)	44.5-46.5	46.5-48.5	48.5-50.5	50.5-52.5	52.5-54.5
number of pieces	3	24	27	21	5

Compute the (i) mean, (ii) mean deviation, (iii) S.D.,

(iv) percentile rank of a breaking strength of 49.25 thousand lb./sq.inch, and

(v) the proportion of test pieces whose breaking strengths lie within the limits

Mean $\pm$ 2 x (S.D.) [5 x 4]=[20]

p.t.o.

5. OR

The table below gives the frequency distribution of diastolic blood pressure of 250 male adults. The readings were taken to the nearest millimetre, and the class mark of each class is given.

Two class frequencies are, however, missing. The mean diastolic pressure of the entire sample is known to be 80.22 mm.

Blood pressure (class mark)	60	65	70	75	80	85	90	95
frequency	4	5	31	?	114	?	25	2

Calculate the

i) range of the sample.

ii) missing class frequencies.

iii) mean deviation of the distribution.

[4+10+6]=[20

INDIAN STATISTICAL INSTITUTE
One year Evening Course in Statistical Methods and Applications: 1977-78

Part I: PERIODICAL EXAMINATION

Paper Ia: Collection and Presentation of Data

Date: 14.9.77 Maximum Marks: 50 Time: $1\frac{1}{2}$ hours

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

---

1. EITHER

Give an account of the different steps in a statistical investigation. [20]

OR

Write notes on any two of the following: [10 x 2]=[20]

i) National Sample Survey

ii) Annual Survey of Industries

iii) Sampling and non-sampling errors.

---

2. The following shows the distribution of net domestic product at factor cost (at 1960-61 prices) of India in 1974-1975 by industry of origin:

industry	net domestic product (Rs. crores)
1. agriculture, forestry, fishing, mining	8359
2. manufacturing, construction etc.	4748
3. transport, communication and trade	3354
4. finance and real estate	854
5. public administration and defence	1592
6. other services	1374
Total	20281

Draw a suitable diagram to bring out the relative importance of the different industries. [15]

---

3. The official wholesale price index number (base: year ended August 1939 = 100) remained almost stationary at 418.9 during the week ended 20th October, 1956 compared with 418.8 for the previous week. The index was down by 0.2 per cent when compared with corresponding week of last month but was higher by 16.3 per cent than that of a year ago The index for food articles advanced further by 0.5 per cent compared with 393.7 for the previous week, by 3 units compared to that for the corresponding week last month and by 23 % compared to last year. The index for industrial raw materials increased by 0.5 per cent to 505.6 compared to last week due to rise in prices of raw jute, groundnut and rapeseeds. The corresponding index last year was 407.9 whereas the index for the corresponding week of last month was the same as in the current week. The index for semi-manufactures declined by 0.5 per cent as compared with 404.2 for the previous week. The index during corresponding week of last month was 407.7 and that for the last year was 332.2. Lower prices of Rayon and Silk brought down the index for manufactures by 0.3 per cent compared with 390.0 for the last week and was less by 0.8 per cent compared to last month, but was 5 per cent more than for last year. The index number for miscellaneous items was 581.9 compared to 418.8 and 419.8 for the corresponding week of last month and last year, respectively.

Present the above data in a suitable tabular form. [15]

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods and Applications: 1977-78

Part I: PERIODICAL EXAMINATION

Paper IIIa: Basic Mathematics

Date: 31.8.77 Maximum Marks: 50 Time: $1\frac{1}{2}$ hours

Note: The paper carries a total of 65 marks. Answer as many questions as you can. The maximum you can score is 50.

1.a) Find the number of distinct words that can be formed by permuting the letters of the word MISSISSIPPI.

b) If $(1 + x)^n = C_0 + C_1 x + \cdots + C_n x^n$, where n is a positive integer, prove that

$$C_0^2 + C_1^2 + \cdots + C_n^2 = \frac{2n!}{(n!)^2} .$$ [7+8]=[15]

2.a) Let p be a fixed positive real number.

Prove that $\lim_{n \to \infty} \sqrt[n]{p} = 1$.

b) Find the set of continuity points and the set of discontinuity points of the function f defined for every real number x as

$$f(x) = \begin{cases} 0 & \text{if } x^2 > 1 \\ 1 & \text{if } x^2 < 1 \\ 1/2 & \text{if } x^2 = 1 \end{cases}$$ [7+8]=[15]

3.a) Find the derivative of x^x where $x > 0$, with respect to x.

b) Show that the function

$$y = e^{-\alpha x}(a \cos x + b \sin x)$$

satisfies the equation

$$y'' + 2\alpha y' + (\alpha^2 + 1)y = 0.$$ [7+8]=[15]

---

4. Evaluate the following integrals:

i) $\int_{-\pi}^{\pi} \sin mx \cos nx \, dx$ where m, n are positive integers.

ii) $\int e^{ax} \cos bx \, dx$.

iii) $\int \frac{1}{1 + \sin x} \, dx$. [6+7+7]=[20]