Indian Statistical Institute
Research and Training School

ISEC - 15th Term: July 1959 - April 1960

List of Examinations

Sl.no.	Date	Subject	Examiner/s
		A. Periodical Examination	
1.	14.9.59	General Principles	AM
2.	16.9.59	Probability & Mathematics	DB, RRR
3.	18.9.59	Statistical Analysis	SKM, HK, SK
4.	28.11.59	Statistical Methods & Sample Surveys: Theory	AM, SKM, MNK
5.	27.11.59	Statistical Methods & Sample Surveys: Practical	AM, SKM, PKB, MNK
6.			
		B. Final Examination	
6.	19.2.60	Paper I : Theory	(SKM, AM, (T.Okuno, NSN, (DNG, AB, NSI
7.	19.2.60	Paper II : Practical	(PKB, AM, (T.Okuno, (DNG, AB, NSI
8.	20.2.60	Paper III. Specialization	
		(i) Sample Surveys	NSN, DNG
		(ii) Machine Tabulation	N.C.Dutta
		(iii) Flood Studies	MKBose, AKS
		(iv) Econometrics	AB, NSI
		(v) Demography	SJP, MVR
		(vi) Crop Surveys	KPS
		(vii) Industrial Statistics (Census of Manufacturing Industries)	KOK, SPS
		(viii) Psychometrics	Mrs. HDas
		(ix) Design of Experiments	T.Okuno
		(x) Estimation & Testing of Hypothesis (Statistical Inference)	JR, PKB

INTERNATIONAL STATISTICAL EDUCATION CENTRE, CALCUTTA

Thirteenth Term (July 1959 - April 1960)

Periodical Examination

General Principles

Date: 14 September, 1959 Full marks - 100 Time: 10.10 A.M. 12.40 P.M.

Answer any FIVE questions

1. Explain the following terms:-

   i. statistic ii. statistical data iii. sample

   iv. population v. attribute vi. multivariate data

2. With the help of an example, show how a series of two-attribute data can be condensed into a two-way frequency table, without loss of much information. What types of information are lost in this process of tabulation.

3. Explain with the help of a simple illustration the main difference between the ordinary graph and the semi-logarithmic graph.

4. State i) purpose of Tabulation
   ii) different parts of a Statistical Table
   iii) principles to be followed in the preparation of tables .

5. a) When should you present data in bar-diagrams and when in line-graphs?

   b) Briefly indicate how the following are useful in the study of relationship between two variables

      i) line curves

      ii) scatter diagrams.

6. In not more than 50 lines write a descriptive note explaining how modern statistics evolved and what role statistics plays in governmental, business, scientific and other activities.

Examiner: A. Matthai

INTERNATIONAL STATISTICAL EDUCATION CENTRE-CALCUTTA

Thirteenth Term (July 1959 - April 1960)

Periodical Examination

Probability and Mathematics

Full marks 100

Date: 16 September, 1959 Time: 10.10 A.M.-12.40 P.M.

Figures in the margin indicate full marks.

1. Carefully define each of the following terms:-

   a) The 'mode' of a random variable.

   b) The 'mean' of a random variable.

   c) 'Independent events'.

   d) 'Exclusive events'. (20)

2. a) Suppose a symmetric die is rolled twice. Find the probability distribution and the mode of the total score. (10)

   b) A bag contains 4 balls of which 2 are white and the other 2 are black. If 2 balls are drawn one by one without replacements and

   A = the event that both are white
   B = the event that both are black

   then find the probability of each of the following three events

   i) $\bar{A}$, ii) A+B, iii) $\bar{A}\,\bar{B}$ (15)

3. Differentiate

   (1) $x^2 e^{x^2}$ (2) $x \log x$ (3) $\frac{e^x+1}{x+1}$. (15)

4. State clearly the procedure to be followed in finding the maxima or minima of a function f(x) of a single variable. (15)

5. Let (x_1, y_1), (x_2, y_2), .. (x_n, y_n) be n-pairs of observations of two random quantities X and Y. A statistical procedure for prediction of Y when X is known is the following

   $$Y = \alpha + \beta X$$

   where α and β are the values at which the function

   $$f(u,v) = \sum_{i=1}^{n} (y_i - u - vx_i)^2$$

   is a minimum. Find α and β. (15)

6. State the rule for differential coefficient of product of two functions. Hence find the differential coefficient of $f(x) = e^x(x-1)$. Deduce the value of

   $$\int x e^x \, dx;$$ (10)

Examiners: D. Basu and R. Ranga Rao

INTERNATIONAL STATISTICAL EDUCATION CENTRE, CALCUTTA

Thirteenth Term (July 1959 - April 1960)

Periodical Examination

Statistical Analysis

Full marks - 100

Date: 30 September, 1959 Time: 10.10A.M. - 12.40 P.M.

Figures in the margin indicate full marks

The following list gives the sex (male/female), the blood group (A, B, AB or O) and the ability (+) or inability (-) to taste PTC of 40 individuals investigated in a genetical study.

Summarise the information in a table, with a suitable title, column and row headings and other essential details to make the table self-explanatory.

serial no.	sex	blood group	taste reaction to PTC	serial no.	sex	blood group	taste reaction to PTC
1	male	AB	+	21	female	A	+
2	female	AB	+	22	male	O	+
3	male	B	+	23	male	O	-
4	female	B	+	24	male	O	-
5	male	B	-	25	female	O	-
6	male	A	+	26	male	B	+
7	male	A	-	27	male	B	+
8	female	A	+	28	female	O	+
9	male	A	+	29	female	O	+
10	female	O	+	30	female	A	+
11	female	O	+	31	female	A	+
12	female	O	+	32	male	A	+
13	male	O	+	33	female	O	+
14	female	A	+	34	female	O	+
15	male	A	-	35	female	A	+
16	male	B	-	36	female	O	+
17	male	B	+	37	male	O	+
18	female	A	+	38	male	O	+
19	male	AB	-	39	male	A	-
20	female	O	+	0	female	B	-

(15)

Certain data obtained from a study of a group of 1000 employees in a cotton mill as to their religion, sex and marital status were reported as follows:

525 Hindus ; 312 males

470 married ; 42 Hindu males

147 married Hindus ; 86 married males

25 married Hindu males ;

Examine this for internal consistency. (15)

P.T.O.

3. (a) Find by linear interpolation the median of the distribution of marks obtained by 80 students as given below:-

Marks	Frequency
0- 10	3
10- 20	9
20- 30	15
30- 40	30
40- 50	18
50- 60	5

Draw also the cumulative frequency curve (ogive) of the above distribution and read off the value of the median. (15)

(b) For a certain group of earners their annual income figures (in rupees) in 1957 and in 1958 are available. The median income in 1958 was found to be 15 rupees more than that in 1957.

Mr. X concludes that on the average people were earning more in 1958 than in 1957. Comment. (10)

[Hint: Show by constructing an example that this need not be so. In fact, the median income can be more even though most of the people are earning less.]

4. (a) What are measures of dispersion of a distribution? Why is the standard deviation most commonly used as a measure of dispersion in statistics? (15)

(b) For two samples of sizes 10 and 15 the mean values and the variances were computed as follows:-

	first sample	second sample
size	10	15
mean	67	64
variance	12	15

Find the mean and the variance for the combined sample of size 25. (15)

(c) For a frequency distribution of marks in history of 200 candidates (grouped in intervals 0-5, 5-10, ... etc.) the mean and standard deviation were found to be 40 and 15. Later it was discovered that the score 43 was misread as 53 in obtaining the frequency distribution. Find the corrected mean and standard deviation corresponding to the corrected frequency distribution. (15)

Examiners: S.K. Mitra, R. Krishnamurty and S. Krishnamurty

INTERNATIONAL STATISTICAL EDUCATION CENTRE-CALCUTTA

Thirteenth Term (July 1959 - April 1960)

Periodical Examination

Statistical Methods and Sample Surveys

(Theory)

Full marks - 100

Date: 28 November, 1959 Time: 10.10 A.M.-12.40 P.

1. What are the considerations in support of the product moment coefficient of correlation as a measure of the degree of correlation between two variables x and y.

   or

   Describe methods you know of judging how far a regression function estimated by the least square method, fits the data.

2. Write down the "normal equations" for fitting a second degree polynomial regression to bivariate data.

   or

   x is a random variable having expectation 20 and variance 16 ; nothing further is known about its distribution.

   What can you say about the expectation and variance of $\bar{x}$, the mean based on n independant observations of x in a random sample. State what could be said further about the distribution of $\bar{x}$.

3. Write short notes on -

   (a) level of significance in a test procedure

   (b) power of a test

4. From a normal population with variance 100 two samples of size 100 and 300 were drawn. What is the probability that the variance in the first sample will exceed the variance in the second by more than 10. (Use large sample approximations).

5. Explain briefly what you understand by -

   (i) population, (ii) random sample, (iii) sampling frame,

   (iv) bias of an estimator, (v) variance of an estimator, and

   (vi) mean square error of an estimator.

6. Suppose 2 families are selected with equal probability without replacement from a population of 4 families for estimating the income per person in the population.

family	number of persons	family income
1	4	800
2	2	1000
3	3	600
4	1	500

   Find out the value of bias, variance and mean square error of your estimator.

Examiners: A. Matthai, S.K. Mitra, and M.N. Murthy

INTERNATIONAL STATISTICAL EDUCATION CENTRE, CALCUTTA

Thirteenth Term (July 1959 - April 1960)

Periodical Examination

Statistical Methods & Sample Surveys

(Practical)

Full marks - 100

Date: 27 November, 1959 Time: 2 PM - 5 PM

Figures in the margin indicate full marks

1. Obtain a formula from the following data which will enable one to forecast yield of rice from amount of manure.

Yield of rice in bushels per acre

		150	160	170	180	190	200	210	220
farm-manure in cart-loads per acre	0.5	12	3						
	1.0	7	24	15					
	1.5		13	38	15	3			
	2.0		6	17	27	20	1		
	2.5			5	11	7	12		
	3.0					2	3	5	
	3.5						1	3	
	4.0							1	1
	4.5								2

or

An economist wanted to derive a formula for predicting the national income in money terms (y), in terms of the money supply with the public (x_1) and national income in real terms (x_2). He decided to use a linear relation

$$y = \alpha + \beta_1 x_1 + \beta_2 x_2$$

Given the following totals and uncorrected sums of squares and sums of products, estimate the constants α, β_1, β_2

	Totals	uncorrected s.s. and s.p.		
		x_1	x_2	y
x_1	137.59	2731.1055	11970.6825	13963.9650
x_2	605.57		68795.3900	70218.9200
y	708.10			71882.4900

Note: Data were obtained for 7 sets of observations (30)

2(a) Find as accurately as possible the value of $e^{-2.25}$ making use of the following table only:

x	e^{-x}
2.2	.110803
2.3	.100259
2.4	.090718
2.5	.082085

(8)

(b) The average number of deaths from street accidents per year in a certain locality is 2.25. Find the probability of (i) at least one death, (ii) more than one death in a year from street accidents. (12)

3. A company sells seeds of a certain flowering plant in packets containing a dozen seeds. It is known that 80 percent of these seeds germinate. What is the probability (exact) that at least 10 out of 12 seeds in a packet will germinate? Find the same by any approximate formula that you know and the percentage error in the approximation

or

Examine on the basis of the following data collected from the attendance register of an office on a certain day whether it is justified to assert that there is some association between distance of office from residence and late attendance

punctuality in coming to office	Distance of office from residence	
	within a mile	more than a mile
in time	28	40
late	5	21

(20)

4. (a) Select a sample of 10 stores with equal probability without replacements from the population of 28 stores given below (6)

(b) Estimate the average sales per store and the standard error of the estimate on the basis of the sample. (12)

(c) Estimate also the proportion of stores with sales greater than Rs. 100 and the standard error of the estimate on the basis of the sample (12)

Store	Sales (Rs.)	Store	Sales (Rs.)	Store	Sales (Rs.)
1	252	11	138	21	179
2	405	12	102	72	401
3	234	13	112	23	190
4	150	14	32	24	112
5	103	15	88	25	119
6	287	16	344	26	128
7	240	17	152	27	186
8	278	18	85	28	89
9	65	19	89		
10	207	20	61		

Examiners: Matthai, S. Krishnamurty, P.K. Bhattacharyya and M.N. Murthy

INTERNATIONAL STATISTICAL EDUCATION CENTRE, CALCUTTA.

Thirteenth Term

Final Examination

Paper I

Full Marks : 100

Date : 19 February 1960 Time : 10 A.M. to 1 P.M.

1. Answer any five of the following questions
2. All questions carry equal marks.

STATISTICAL ANALYSIS

1. Horse racing fans often maintain that in a race around a circular track significant advantages accrue to the horses in certain post positions. Any horse's post position is his assigned post in the starting line up. Position 1 is closest to the rail on the inside of the track; position 8 is on the outside, farthest from the rail in an 8-hourse race.

New York Post (August 30, 1955) gives the following race results for the first month of racing in the 1955 season at a particular circular track

WINS ACCRUED ON A CIRCULAR TRACK
BY HORSES FROM EIGHT POST POSITIONS

Post positions	1	2	3	4	5	6	7	8	Total
No. of wins	29	19	18	25	17	10	15	11	144

Let us start with the null hypothesis H_o which states that the suspicion of these horse racing fans is baseless i.e. post position in no way influences the results of the race.

(i) Assuming this null hypothesis to be true and that the horses are assigned at random to the various post positions before the race is run, what is the probability that in a single race the win will be registered at a particular post position?

(ii) Hence compute the number of wins which are expected at various post positions in 144 races, if H_o is true.

(iii) Compare these expected frequencies with the observed frequencies as given in the data by computing the Goodness of Fit Chi-square and on the basis of the computed value suggest whether it would still be reasonable to assume that H_o is true or whether there are positive indications in the opposite direction.

2. The following table gives the correlation coefficients between stature and sitting height of Muslims in 4 districts of Bengal and also the size of the samples from which those coefficients were computed.

district	sample size	correlation coefficient
Barisal	131	0.719
Dacca	337	0.824
Mymensingh	299	0.547
Nadia	170	0.793

Describe a test procedure for testing the homogeneity of these coefficients giving complete details of the computational layout and stating clearly the assumptions under which the procedure is valid. No computations are required.

Examiner : S.K.Mitra

Please turn over

DESIGN OF EXPERIMENTS

3. Show how the following three conceptions: (i) Replication, (ii) Randomization and (iii) Local control, are introduced into the randomized block design, and explain briefly what kinds of role they could play.

4. (a) A split-plot experiment on corn is being conducted to try 3 rates of planting (stands) with 3 levels of fertilizer in irrigated and non-irrigated plots. The design is randomized blocks with 4 replications. The whole plot (the first unit) carries the irrigation treatments. Show one example of the field arrangements corresponding to this purpose (with randomization).

(b) For a six-month course in Statistics, 6 teachers (1,2,3,4,5,6) are available for the tutorial classes. The students are divided into 6 groups (A,B,C,D,E,F). Each group of students is to work under the supervision of each teacher for one month. Prepare a tutorial programme for the six months (I,II, III, IV,V,VI) in a Latin square with randomization.

Examiner : T. Okuno.

SAMPLE SURVEYS

5. (i) What is cluster sampling? (ii) Is cluster sampling more efficient in practice than a simple random sample of elements if the same number of elements is covered in both the cases? Give reasons for your answer.

(iii) For the same fixed cost, would you prefer cluster sampling to simple random sampling (of elements)? Give a suitable example to support your answer.

(iv) To estimate the average number of persons per household (household size) a sample of villages is to be drawn and total number of households and persons in each sampled village is to be counted. Write down a suitable estimator you will use.

6. (a) (i) What is stratified sampling? Why is it used? What is the maximum number of strata you can have?

(ii) If from each stratum, a simple random sample is to be drawn and the ordinary unbiased estimate is to be used, how will you group the units to form strata?

(iii) Will you change the method of stratification if you intend to use a ratio estimate? Illustrate your answer with an example.

(b) Write down the expression in a stratified simple random sample for

(i) the estimate of the mean,
(ii) the variance of the estimate of the mean
(iii) sample sizes in different strata under optimum (Neyman) allocation
(iv) the variance of the estimate of the mean under optimum (Neyman) allocation.

(Proper explanation of the symbols used should be given).

Examiners: N.S. Nanjamma and D.N.Ghosh.

Please turn over

ECONOMETRICS

7. Show how national income is determined by the equality of saving and investment. What is the effect of an increase in investment on national income?

Or,

Give three definitions of national income and demonstrate their equivalence.

8. What do you understand by the term index number of seasonal variation in economic time series? Indicate two important methods of computing seasonal indices. Describe the various steps involved in any one of the methods.

Or,

Define clearly the concept of Engel elasticity. How do you compute it for a given commodity such as sugar? What type of data do you require? In your calculation of such elasticities from household budget data, what adjustments will you make for the size and composition of the household?

Examiners: A.K.Biswas and N.S.Iyengar.

STATISTICAL QUALITY CONTROL

9. Distinguish between the terms 'statistical control' and 'economic control'. State how a control chart may be defined.

Or,

Examine how, when a quality characteristic may have any unknown distribution, the use of three-sigma limits are justified in a control chart. Why not use exact 'probability limits' on control charts at least when the form of the distribution is known?

10. Distinguish between the two situations in each of the following pairs:-

(i) control charts when standards are given and when standards are not given;

(ii) control charts for small samples and for large samples;

(iii) control charts for individuals and for samples of individuals

Examiner : A. Matthai.

oooOooo

INTERNATIONAL STATISTICAL EDUCATION CENTRE, CALCUTTA.

Thirteenth Term

Final Examination

Paper II

Full Marks : 100

Date : 19 February 1960. Time : 2 P.M. - 5 P.M.

1. Answer any five of the following questions.

2. All questions carry equal marks.

STATISTICAL ANALYSIS

1. Samples of different sizes were selected from lots supplied by five different manufacturers A, B, C, D, E and the number of defective articles detected are tabulated below.

Manufacturer	A	B	C	D	E
Sample size	100	200	150	250	150
Number of defectives	20	45	25	48	23

Do you consider lots supplied by the different manufacturers to be of the same quality?

2. In a feeding experiment, a number of animals were paired up according to initial weight. Ten such pairs were taken and one animal in each pair was given ration 1 and the other ration 2, The gains in pound at the end of the feeding period are given below.

Pair	Ration 1	Ration 2	Pair	Ration 1	Ration 2
1	12.3	10.9	6	10.5	9.4
2	11.4	10.2	7	7.2	7.2
3	8.2	8.4	8	8.6	7.5
4	13.5	11.3	9	10.8	9.2
5	10.2	8.3	10	11.0	11.3

Do the results of the experiment provide sufficient evidence to indicate that ration 1 is better than ration 2?

Examiner : P. K. Bhattacharyya.

DESIGN OF EXPERIMENTS

3. The following data come from a randomized block experiment:-

variety	block I	II	III	IV
A	12	14	14	14
B	10	12	10	13
C	8	10	15	19
D	15	16	20	23
E	23	20	25	24

(i) Carry out the analysis of variance.
(ii) Show that variety E differs significantly from A, B, and C.
(iii) Set 95% confidence limits on the difference between varieties A and B.

Please turn over

4. The following data show the treatment totals of average daily gains of swine in 2^3 factorial experiment. The design was randomized blocks with 7 replications.

Lysine %	Protein %	Sex	Treatment totals
0	12	M	8.5
		F	8.6
	14	M	11.0
		F	10.7
0.6	12	M	10.2
		F	9.1
	14	M	10.3
		F	9.6

Compute the mean effects and sums of squares for main effects, two-factor interactions, and three-factor interaction, and then, test these effects against the error mean square of 0.0224.

Examiner : T. Okuno.

SAMPLE SURVEYS

5. The table below gives the area (in acres) under rice in 1936 and 1937 for a set of 18 villages. Select a sample of 6 villages with probability proportional to the area under rice in 1936 and with replacement. Estimate

(i) the total area under rice in 1937
(ii) the variance of the estimate in (i)
(iii) the coefficient of variation of the estimate in (i).

Area under rice in acres

sl. no. of village	1936	1937	sl. no. of village	1936	1937
1	80	50	10	130	135
2	160	150	11	130	145
3	130	110	12	130	105
4	100	110	13	190	180
5	110	100	14	70	60
6	80	80	15	60	80
7	80	105	16	70	60
8	50	30	17	140	100
9	90	85	18	200	140

6. A set of 100 tehsils is grouped to form two strata such that stratum 1 contains 40 tehsils and stratum 2, 60 tehsils. A sample of 8 tehsils has been drawn with equal probability without replacement from stratum 1 and a sample of 10 tehsils from stratum 2 in a similar way. The total population in stratum 1 and 2 are 3,675,000 and 4,155,000 respectively. The table below gives the data on agricultural population and total population in each of the 18 sampled tehsils.

* $\left[\hat{V}_{HH}(\hat{Y}) = \frac{1}{n(n-1)} \left\{ \sum_{i=1}^{n} \frac{y_i^2}{p_i^2} - n\hat{Y}^2 \right\} \right]$

Please turn over

(a) Estimate the total agricultural population using

(i) a separate ratio estimator
(ii) a combined ratio estimator

and using the total population as auxiliary characteristic.

(b) Also estimate the variances of the above estimates.

Stratum 1			Stratum 2		
tehsil no.	Agricultural population (in .000)	Total population (in .000)	Tehsil no.	Agricultural population (in .000)	Total population (in .000)
1	115	135	1	95	110
2	55	80	2	65	80
3	125	185	3	55	70
4	65	95	4	85	125
5	80	110	5	60	70
6	60	95	6	95	135
7	70	85	7	40	70
8	75	90	8	90	100
			9	50	80
			10	90	125

$$\hat{V}\left(\hat{\bar{y}}_{RS}\right) = \sum_{h=1}^{L} \frac{N_h\left(N_h - n_h\right)}{n_h\left(n_h - 1\right)} \left\{ \sum_{i=1}^{n_h} y_{ih}^2 + \hat{R}_h^2 \sum_{i=1}^{n_h} x_{ih}^2 - 2\hat{R}_h \sum_{i=1}^{n_h} y_{ih} x_{ih} \right\}$$

$$\hat{V}\left(\hat{\bar{y}}_{RC}\right) = \sum_{h=1}^{L} \frac{N_h\left(N_h - n_h\right)}{n_h\left(n_h - 1\right)} \left\{ \sum_{i=1}^{n_h} y_{ih}^2 + \hat{R}^2 \sum_{i=1}^{n_h} x_{ih}^2 - 2\hat{R} \sum_{i=1}^{n_h} y_{ih} x_{ih} \right\}$$

Examiners : N.S.Nanjamma and D.N.Ghosh

ECONOMETRICS

7. The numbers (in hundreds) of letters posted in a certain city on each day in a period of five consecutive weeks are given below. Calculate the indices of average fluctuations from day to day within each week.

Number of letters posted in a certain city

week	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
first	18	161	170	154	143	161	76
second	18	165	179	157	168	195	85
third	21	162	169	153	139	185	82
fourth	24	171	182	170	162	179	95
fifth	27	172	196	180	170	202	120

Please turn over

8. The following table gives the distribution of persons by per capita monthly income and percentage consumption of cereals in different groups for rural and urban areas of West Bengal.

per capita income (Rs.)	number of persons (thousands)		consumption of cereals (p.c.)	
	rural	urban	rural	urban
1 - 7	463	61	3.64	0.86
8 - 10	915	38	5.22	0.29
11 - 13	1098	151	6.26	2.08
14 - 16	2320	387	13.23	5.05
17 - 19	1668	475	9.51	5.43
20 - 24	2329	912	13.28	9.47
25 - 29	2855	1200	16.28	15.71
30 - 39	2996	802	17.09	12.61
40 - 49	1366	698	7.79	10.85
50 and above	1526	1022	8.70	37.65

Draw concentration curves of consumption of cereals and study the difference between rural and urban areas, if there is any.

Examiner : N. S. Iyengar.

STATISTICAL QUALITY CONTROL

9. Samples were taken from 8 different consignments of a product in order to find out if any assignable causes of variation existed in a quality characteristic (x). Examine the results given below, for control and write a brief note giving your conclusions.

consignment	average	standard deviation
1	33.5	4.20
2	32.7	4.35
3	33.9	4.38
4	40.6	4.56
5	35.6	4.51
6	33.2	4.39
7	34.5	4.90
8	35.7	4.86

10. Interpret the following control chart constants indicating clearly when they are to be used : c_2, d_3, A_2, D_4, B_4.

Examiner : A. Matthai

ooooOoooo

INTERNATIONAL STATISTICAL EDUCATION CENTRE, CALCUTTA.

Thirteenth Term

Final Examination

Paper III

Special Subject : SAMPLE SURVEYS

Full Marks : 100

Date : 20 February 1960 Time : 10 A.M.- 1 P.M.

Figures in the margin indicate full marks.

1. (i) Prove that the sample mean is an unbiased estimate of the population mean for a simple random sample drawn without replacement.

(ii) Derive an expression for the variance of the sample mean for a simple random sample drawn without replacement. 15

2. A forgetful statistician, in charge of a sample survey, used the formula

$$n_i = n.\frac{N_i S_i^2}{\sum N_i S_i^2}$$

to obtain the allocations for the different strata he had constructed. What was the mistake he committed? What would be the variance of his estimate for the total of the characteristic he wants to measure? Will it be more or less than what he would have obtained if he had allocated proportionally? (It is supposed that samples are to be drawn with equal probability and without replacement from each stratum). 12

3. A sample survey is to be conducted among 1200 persons to estimate their total monthly expenditure. The error that can be tolerated is 10 percent of the true value on either side with a confidence level of 90 percent. The coefficient of variation of the distribution of monthly expenditure is known to be 60 percent. Calculate the sample size required if the sample is drawn with equal probability and without replacement.

probability level	50 percent	80 percent	90 percent	95 percent	99 percent
Normal deviate	0.67	1.28	1.64	1.96	2.58

4. The authorities of a college decided to conduct a sample survey to estimate the no. of smokers among the students. A random sample of 110 Asian students from the 600 Asian students and another random sample of 25 European students from the 100 European students were selected. The selection in both the cases was done with equal probability and without replacement. The table below gives the results obtained in the enquiry. 16

stratum	smokers	non-smokers	total
Asian students	33	77	110
European students	15	10	25

Estimate (i) the total no. of smokers in the college, (ii) the variance of your estimate.

Please turn over

5. The following table gives the population of 36 villages of Bihar State, India. 18

sl. no. of the village	population	sl.no. of the village	population	sl.no. of the village	population	sl.no. of the village	popula-tion
1	823	10	412	19	289	28	281
2	736	11	576	20	274	29	167
3	710	12	507	21	250	30	149
4	616	13	367	22	252	31	134
5	574	14	328	23	259	32	129
6	530	15	328	24	284	33	164
7	551	16	304	25	231	34	95
8	480	17	345	26	315	35	87
9	414	18	293	27	325	36	69

Draw two systematic samples of size 5 and obtain (i) the best unbiased estimate of the total population of the 36 villages and (ii) an unbiased estimate of the variance of the estimate.

6.(a) Mention in brief the important steps to be taken at the central office in planning a survey.

(b) It is proposed to conduct an Agricultural Labour Enquiry all over India concerning the debt and employment position of agricultural labourers. Suggest what type of data you will require to construct your strata.

(c) From the investigator variability and error considerations it is thought that a sample of 10 villages per investigator for each season (a season being a three monthly period to be considered as one subround) will be suitable. The time requirements are as given below

Schedule	Description	Time requirement
0.1	Listing of households	3 days./ village
10.0	Employment and Unemployment	3 hhs./2 days
7.0	Household indebtedness	3 hhs./ day
3.01	Price collection	1 day/village

To estimate the seasonal variation in employment position, Schedule 10.0 is to be canvassed among the same set of households every 6 months and Schedule 3.01 every month from the same village. Construct a work programme assuming that in each subround of 3 months there are 70 working days so that for the whole of the round period of one year the investigator will work for 280 days. 25

Examiners: D. N. Ghosh and G. Parthasarathy.

INTERNATIONAL STATISTICAL EDUCATION CENTRE - CALCUTTA

Thirteenth Term

Final Examination

Paper III

Special Subject : Punched Card Accounting

Full marks 100.

Date: 20 February, 1960 Time: 10 A.M.-1 P.M.

Figures in the margin indicate full marks

1. How many holes can you punch in a column of a card ? Why is the corner of the card cut ? 7

2. How many keys are there on a key punch machine ? What is the difference between a key punch and a verifier machine ? 8

3.(a) How many pockets are there on a sorter machine ? 3

(b) If you are to sort a pack of cards to columns 12 to 16, from which column should you start sorting ? 3

(c) Is it possible to select all cards punched '5' from a pack of cards in column '10' on a sorter without disturbing the sequence of the other cards ? If so how ? 7

4.(a) What is the speed of a Reproducer ? Is it possible to Gang punch information from columns 15 to 18 of card to columns 1-4 a set of cards ? 6

(b) Prepare plug-chart for the following jobs :- 18

Reproduce from Cols. 10-14 to Cols. 20 - 24
" 30-35 to " 30 - 35
" 41-42 to " 41 - 42

5.(a) How many printing Bars are there on a 416 type tabulator ? What are the basic steps in plugging for accumulation ? 8

(b) Prepare a plug-chart for the following job on 416 type tabulator ? 25

Card design

1.	C.D.I	Col. 1-4
2.	Zone	" 5
3.	State	" 6
4.	Qty.	" 7-9
5.	Value	" 10-12

You are to prepare a statement showing the totals of value and Qty by state for each zone i.e. zone x state.

6.(a) How many Banks and adding counters are there on a Hollerith Senior Tabulator ? What is the maximum number of adding wheels in a counter ? 5

(b) Is it possible to subtract on it ? What is the adding speed of it (Hollerith Senior Tabulator) ? 5

7. How many feeds are there on a Collator ? Can you print on it ? 5

Examiner: N. C. Dutta.

INTERNATIONAL STATISTICAL EDUCATION CENTRE, CALCUTTA.

Thirteenth Term

Final Examination

Paper III

Full Marks : 100

Special Subject : Flood Estimation

Date: 20 February 1960 Time: 10 A.M.-1 P.M.

1. Define the terms 'flood intensity' and 'flood run-off' stating the the units in which they are generally expressed. What are the major factors affecting floods? State briefly how some of these factors are interconnected and some are constant for a particular catchment.

Discuss the applicability of Dickens' formula

$$Q = 825a^{\frac{3}{4}}$$

(Q = maximum intensity of flood in cubic feet per second and 'a' = area of the catchment in square miles). Mention also the main limitations of such flood formulae.

2. Describe briefly either (a) Probability method or (b) Unit Hydrograph method for estimation of floods.

3. The following table gives the observed daily gauges at two sites A and B of certain river, during the period August 1 - 15, 1955. Gauges are observed at a certain fixed hour on each day. The river flows from A to B.

Date	River gauges (in feet) at A	B
August 1	78.50	22.00
2	78.50	20.20
3	76.60	21.50
4	77.00	20.60
5	76.60	19.70
6	76.60	17.60
7	76.60	17.90
8	76.60	17.30
9	82.40	16.90
10	80.40	24.40
11	79.30	25.50
12	78.50	20.60
13	77.80	19.20
14	78.00	20.60
15	77.80	21.10

Can we estimate the time-lag between the occurrences of the gauges at the two places from the above data? Obtain a linear regression of 'Gauge at B' on the 'Gauge at A' with one day time-lag.

oooOooo

Examiners : N.K.Bose and A.K.Sinha.

INTERNATIONAL STATISTICAL EDUCATION CENTRE - CALCUTTA

Thirteenth Term

Final Examination

Paper III

Special Subject: Econometrics

Full marks 100

Date: 20 February 1960 Time: 10 A.M.- 1 P.M.

1. Develop the procedure of construction of input-output tables for an open economy.

OR

Given the investment fund available for planned expenditure how would you decide to allocate it to producers goods and consumers goods sectors ?

[Suggested time: 30 minutes]

2. Making suitable assumptions, to be stated clearly, show how you would use the Lorenz curve and concentration curves together for computing the Engel elasticities.

OR

How would you determine the likely increase in consumption of cereals per capita by the end of the Third Plan, given that the average income per person would increase by 100 α percent as compared to that in 1955-56 while the inequality of incomes would fall by 100 β percent ? What assumptions would you make use of in solving the problem ?

[Suggested time : 30 minutes]

3. Describe the process of increasing the rate of saving in the most general case and then proceed to suggest some ways in which it could be done in India.

OR

What is the prime facie reason that we should expect the capital coefficient to be lower in a planned economy than in an unplanned economy, when the two economies are at the same technological level.

In this connection, suggest some way of reducing the overall capital coefficient while not relenting on the question of introducing the most modern techniques of production in the heavy industries.

[Suggested time: 30 minutes)

4. Suppose the following problem is given. How would you reformulate it as a problem in linear programming ?

Let the Indian economy be considered as consisting of the four sectors:

S_1 : basic investment goods

S_2 : factory consumer goods

S_3 : household industries (including agriculture) and

S_4 : services.

It has been determined on the basis of some studies that the income-investment ratios in the above sectors are respectively 0.20, 0.35, 1.25 and 0.45. The capital requirements per employed person in the sectors are Rs.20,000, Rs.8,750, Rs.2,500 and Rs.3,750 respectively.

P.T.O.

The Planning Commission desires that the additional employment in the household industries sector and services sector should not together exceed 9 million. Further the Commission has the information that the additional employment that would have to be created during the Plan period will not exceed 11 million. Given that the total resources available for investment during the period is less than Rs.5,600 crores, suggest an optimum investment plan for the Planning Commission.

[Suggested time : 30 minutes]

5. Using the data given below draw (i) Lorenz curve for total expenditure and (ii) the specific concentration curve for the expenditure on clothing. Making suitable assumptions compute the Engel elasticity of demand for clothing items.

Table: Consumer expenditure in rupees per person per month All-India (rural and urban) 1955-56.

	Expenditure classes in rupees per person per month												All classes
	0-8	8-11	11-13	13-15	15-18	18-21	21-24	24-28	28-34	34-43	43-55	55 and above	
average expenditure on clothing.	0.21	0.53	0.79	1.03	1.43	1.92	2.19	2.60	3.86	4.93	4.33	10.24	1.77
total expenditure per capita	6.26	9.41	11.98	13.96	16.49	19.54	22.51	25.79	30.68	37.72	47.24	83.29	18.71
p.c.of population	14.01	17.31	11.74	9.14	11.55	8.21	7.04	5.53	5.17	4.85	2.59	2.86	100

[Suggested time : 1 hour]

Examiners:- A. K. Biswas and N.S.Iyengar.

INTERNATIONAL STATISTICAL EDUCATION CENTRE - CALCUTTA

Thirteenth Term

Final Examination

Paper III

Special Subject : Demography

Full marks 100

Date: 20 February 1960 Time: 10 A.M.-1 P.M.

1. The crude death rates in two groups of male labourers (groups A and B) are 6.21 and 11.39 per 1000. In the following table are given the age specific death rates in the two groups and the age distribution of the standard population comprising of all male labourers. Determine the relative value of death rates in group A and group B after eliminating the effect of age-composition on the crude death rates.

Age	Death rate per 1000		Standard
	Group A	Group B	population
20 - 24	2.65	4.13	1651416
25 - 34	2.68	4.46	3029176
35 - 44	4.04	7.68	2489703
45 - 54	7.49	14.50	2277821
55 - 64	16.45	28.93	1736275

2. The following table gives the enumerated number of children in the census taken at the end of 1950 in a certain area together with the completness of enumeration and survival rates from birth to mid-point of specified age. Calculate the total expected number of births which should have occurred during the five-year period, 1946-50.

age	enumerated number in 1950 (in thousands)		completeness of enumeration		survival rate from birth to mid point of specified age	
	male	female	male	female	male	female
0 year	274	262	.6335	.6623	.88815	.90163
1 year	268	259	.6958	.7327	.73521	.75595
2 year	298	291	.8776	.9440	.67303	.69344
3 year	301	300	.9740	1.0440	.64049	.65867
4 year	307	305	1.0280	1.0581	.62194	.63885

3. From the following table calculate the gross and net reproduction rates assuming that the proportion of female births to all births is 0.48.

age	female population (1951 census)	number of births 1951	survival rate of females from birth to specified age
15 - 19	16545	783	.66769
20 - 24	15080	2333	.62601
25 - 29	13968	3617	.57884
30 - 34	13910	3915	.53235
35 - 39	10269	3179	.48820
40 - 44	9122	2148	.44656

4. Briefly discuss: (i) defacto and dojure enumeration; (ii) incidence and prevalence rates;(iii) ad-hoc and recurrent surveys.

Examiner: M. V. Raman

INTERNATIONAL STATISTICAL EDUCATION CENTRE - CALCUTTA
Thirteenth Term
Final Examination

Paper III
Full marks 100
Special Subject: Crop Surveys

Date:-20 February 1960 Time: 10 A.M.-1 P.M.

Figures in the margin indicate full marks

1. Name the important crops of your country and give their sowing and harvesting time. Give a suitable sampling design and draft a schedule of investigation for a sample survey to estimate the area under the crops and the production of crops. Also indicate the method you would adopt to estimate the area and production of crops. 50

2. Describe briefly the method of crop-cutting experiments conducted by the National Sample Survey (NSS), to estimate the yield rate of crops dealing in particular with the following:

   a) reasons for two concentric circular cuts
   b) the problem of bias due to border plants and the method of minimising it. 30

3. Give some of the various estimation procedures which can be used in large scale crop surveys to estimate area under and production of crops, the design being similar to that of the 15th round National Sample Survey. 20

Examiner: R. P. Saha.

INTERNATIONAL STATISTICAL EDUCATION CENTRE

Thirteen Term

Final Examination

Paper III

February 1960 Special Subject : Census of manufacturing industries.
Full marks : 100 Time 10 A.M. – 1 P.M.

Question No.1 is compulsory . Attempt any two of the following

1. Design a schedule for the purpose of studying wages and employment structure in the manufacturing sector, carefully stating the items of information to be collected.

   Write the instructions (clearly fifining the items) for filling up the schedule. Draw a tabulation program for the report.

2. In order to estimate the 'value added' by the manufacturing sector, a sample survey of factories is to be undertaken.

   Industry-wise list of factories giving the number of workers employed by each of them is available. The total sample size is fixed.

   (a) Give a suitable sampling design, indicating the method of allocation of samples to different industries. Also give the estimation procedure.

   (b) What more information you will like to have in order to improve the allocation procedure and the sampling ? State reasons.

3. (a) Differenciate between Ex-factory price and Market price.

   (b) Give the different positive components that are included in the Ex-factory value of output.

   (c) Explain the terms

      1. Gross value added

      2. Net value added

   (d) The records of a certain factory is given below . Calculate

      1. Ex-factory value of output

      2. Gross value added

      3. Net value added

Records

1. Sale proceeds of the total output	50.00	crores of Rupees
2. Salaries and wages paid to workers engaged in production	10.25	"
3. Value of raw material etc.	15.00	"
4. Discount to pruchasers	0.60	"
5. Taxes on total sales	1.05	"
6. Transport and delivery charges at customers 'destination	2.00	"

7.	Insurance charges for the output	0.05 crores of Rupees
8.	Profits of the factory	8.15 "
9.	Other input items	2.00 "
10.	Depreciation charges	5.25 "

4. (a) What is meant by index of industrial production ?

Give a suitable method of computing one. How will you adjust it for seasonal variation ?

(b) The general indices of industrial production for the years 1957 and 1958 are 136 and 137 respectively.

What more minumum information (if any) you will require in order to compare the production in real terms for the two years. State reasons.

Examiners: K. K. Kar and S. P. Sangal.

TISTICAL EDUCATI(
Thirteenth Term

Final Examination

Paper III

Special Subject: Psychometry and Educational Statistics

Full marks 100

Date: 20 February 1960 Time: 10 A.M. - 1 P.M.

1. Give brief definitions of the following terms:

   (a) limited response answer
   (b) reliability
   (c) standard score
   (d) percentile score
   (e) validity

2. A test is to be chosen for selecting candidates for admission to the University. Several tests have been tried out to see which one would be best for this purpose. Their reliability and validity coefficients are as follows:

Test	Reliability		Validity	
A.	Split-half correlation	= .80	Correlation with high school marks	= .72
B.	Equivalent forms correlation	= .76	Correlation with University final examinations	= .61
C.	Split-half correlation	= .84	Correlation with University final examinations	= .52
D.	Equivalent forms correlation	= .78	Correlation with high School marks	= .68

   Which test should be chosen for selecting students, and why ? Explain briefly.

3. Briefly answer the following:

   (a) Explain what is meant by item analysis
   (b) Indicate the purposes of item analysis
   (c) Describe the two major types of item indices.

4. Suggest some ways in which psychometric concepts and methods may be usefully applied in the educational field in Japan.

Examiner: Rhea S. Das.

INTERNATIONAL STATISTICAL EDUCATION CENTRE, CALCUTTA

Thirteenth Term

Final Examination

Paper III

Full Marks : 100

Special Subject : Design of Experiments

Date 20 February 1960 Time 10 A.M. - 1 P.M.

Q.1. Suppose the following data are obtained from a randomized block experiment

Treatment	Blocks			
	I	II	III	IV
1	4	3	6	7
2	18	19	18	13
3	26	25	24	21
4	38	35	28	31
5	44	43	39	38

(1) Express each observation as a sum of the following 4 items:-

$\hat{\mu}$ = estimate of general mean

$\hat{\alpha}_i$ = estimate of the i-th treatment effect

$\hat{\beta}_j$ = estimate of the j-th block effect

e_{ij} = residual

corresponding to the mathematical model

$x_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$

(2) Calculate the sum of squares for residuals.

Q.2. Here is a Latin square with one missing plot. Treatments are indicated A, B, C and D.

Row	Column 1	2	3	4
1	A 3	B 4	D 5	C 5
2	B 12	A 11	C 12	D 9
3	C 15	D 11	B 13	A 14
4	D 6	C [ ]	A 5	B 7

(i) Estimate the missing value.

(ii) Analyse the variance

Q. 3. State briefly the advantages and disadvantages of the factorial experiment.

Q. 4. Consider the following designs for comparing 4 treatments A,B,C,D with one another in the following manners, each having 12 observations. Show the partition of degrees of freedom and expected values of mean squares in analysis of variance table, corresponding to each case

Please turn over

Order of treatments

	1	2	3	4	5	6	7	8	9	10	11	12
(i)	A	D	C	D	B	B	C	A	B	D	A	C

(completely randomized)

	1	2	3	4	5	6	7	8	9	10	11	12
(ii)	B	C	A	D	A	B	D	C	C	B	D	A

(randomized within each replication)

	1	2	3	4	5	6	7	8	9	10	11	12
(iii)	A	A	A	D	D	D	C	C	C	B	B	B

(each treatment given only once, with 3 readings)

Q. 5. Here is a split-plot experiment for easy computation. Analyse the variance and give a brief comment. The whole plot carries the factor A and to the split-plots the levels of factor B are assigned.

Treatment		Block I	Block II	Block III
A_0	B_0	0	2	1
	B_1	1	3	2
	B_2	2	4	3
A_1	B_0	-4	-2	-3
	B_1	-3	-1	-2
	B_2	-4	-2	-3
A_2	B_0	-1	1	0
	B_1	0	2	1
	B_2	0	2	1
A_3	B_0	-1	1	0
	B_1	0	2	1
	B_2	1	3	2

Q.6. The total yields for 5 spacings obtained from the Latin square experiment are given below

Please turn over

spacings	treatment totals	coefficients for orthogonal polynomials			
		linear	quadratic	cubic	quartic
A : 2"	349	-2	+2	-1	+1
B : 4"	314	-1	-1	+2	-4
C : 6"	262	0	-2	0	+6
D : 8"	191	+1	-1	-2	-4
E : 10"	188	+2	+2	+1	+1
Divisor		10	14	10	70
λ_i		1	1	5/6	35/12

where, the orthogonal polynomials are given as follows, n showing the number of levels of treatment.

$$u_1 = \lambda_1(x - \bar{x})$$
$$u_2 = \lambda_2\left[(x - \bar{x})^2 - \frac{n^2 - 1}{12}\right]$$

(i) Test each regression comparison against the error mean square 1,056, and give a brief comment.

(ii) Fit the orthogonal polynomial of 2 degrees to the treatment means .

oooOooo

Examiner : T. Okuno.

INTERNATIONAL STATISTICAL EDUCATION CENTRE - CALCUTTA
Thirteenth Term
Final Examination
Paper III

Special Subject: Statistical Inference
Full marks : 100

Date: 20 February 1960 Time: 10 A.M.-1 P.M.

1.(a) Define the terms 'unbiased estimate', 'sufficient statistic', and 'efficient estimate'.

(b) $X_1, \ldots, X_{n_2}$ are n independent random samples from a population with variance σ^2. Show that

$$s^2 = \frac{1}{n-1} \sum_{1}^{n} (X_i - \bar{X})^2$$

is an unbiased estimate of σ^2 where $\bar{X} = \sum_{1}^{n} X_i / n$. Can you recommend s as an unbiased estimate of σ ? Give reasons.

2. T_1 and T_2 are two unbiased estimates of θ. $V(T_1) = \sigma_1^2$, $V(T_2) = \sigma_2^2$ and $Cov(T_1, T_2) = 0$. Show that it is always possible to form a linear combination T of T_1 and T_2 which will be an unbiased estimate of θ and will have a variance less than that of T_1 or of T_2.

3.(a) Explain the maximum likelihood method of estimating a parameter θ.

(b) State what you know about the large sample properties of this method.

(c) The following are independent samples from a Poisson population with unknown mean μ :

1, 0, 2, 3, 1.

(i) Write down the likelihood function , and

(ii) find the maximum likelihood estimate of μ, say $\hat{\mu}$.

(iii) Obtain an estimate of the variance of $\hat{\mu}$.

4.(a) Explain the following terms:

(i) Level of significance of a test, (ii) Power of a test.

(b) X has Poisson distribution with unknown mean μ, The hypothesis H H_0 ($\mu = 1$) is to be tested. Consider the following test based on four independent random samples (X_1, X_2, X_3, X_4):

Let $T = X_1 + X_2 + X_3 + X_4$

Reject H_0 if $T \geq 5$

Accept H_0 if $T < 5$.

Find the level of significance of this test and its power when $\mu = 2$.

P.T.O.

5(a) State and prove the Nuyman-Pearson lemma in testing of hypotheses.

(b) $X_1, \ldots, X_{10}$ are independent random samples from a normal population with mean 0 and variance σ^2 (unknown). Find the most powerful test having type one error = .05 for the hypothesis $H_0(\sigma^2 = 2)$ against $H_1(\sigma^2 = 1)$.

6. There are 3 normal populations with means μ_1, μ_2, μ_3 and a common variance σ^2. μ_1, μ_2, μ_3 and σ^2 are all unknown. Independent random samples of size 5 are taken from each of the three populations. Derive the likelihood ratio test criterion for testing the hypothesis $H_0(\mu_1 = \mu_2 = \mu_3)$ and with the help of this criterion give the test procedure having type one error = .01.

Examiners: J. Roy and P.K.Bhattacharya.

Indian Statistical Institute
Research and Training School

Short-Term Course (Junior) : February 1959 - August 1959

LIST OF EXAMINATIONS

Sl.no.	Date	Subject	Examiner/s
		A. Preliminary Examination	
1.	7. 5. 59	Presentation of Data & Economic Statistics	AM, SK
2.	21.5.59	Probability, Numerical Mathematics, Descriptive Statistics	JR, CSR, PKB
3.	18.6.59	Descriptive Statistics, Sample Surveys and Vital Statistics	JR, MNM, PKB
		B. Final Examination	
4.	14.8.59	Paper I	JR, VPG
5.	18.8.59	Paper II	JR, SKM, PKB, SJ
6.	20.8.59	Paper III	JR, MNM, NSN
7.	22.8.59	Paper IV	AM, SK
8.	16.8.59	Paper V (Practicals) Session 1	JR, PKB
9.	16.8.59	Paper V (Practicals) Session 2	GKC, SJ, PKB
10.	23.8.59	Paper VI (Practicals) Session 1	IMC, SK
11.	23.8.59	Paper VI (Practicals) Session 2	AM, MSI, NSN, PKB

Indian Statistical Institute
Research and Training School

Short-Term Course (Junior): February 1960 - ~~March~~ August 1960
LIST OF EXAMINATIONS

Sl.no.	Date	Subject	Examiner/s
		A. Preliminary Examination	
1.	5.4.60	Data Processing; Numerical Mathematics	DB, AM, PKB
2.	17.5.60	Descriptive Statistics; Probability; Economic Statistics	DB, JR, NSI
3.	28.5.60	Economic Statistics, (S_1) Vital Statistics (Practical)	NSI, MVRaman
4.	28.5.60	Descriptive Statistics (S_2)	JS
5.	2.6.60	Algebra and Sample Surveys	SKM, SRR
		b. Final Examination	
6.	29.7.60	Paper I (S_2)	RNG
7.	1.8.60	Paper II	AM, JR, NSI
8.	3.8.60	Paper III	JR, DNG, NSN
9.	5.8.60	Paper IV	AM, NSI
10.	31.7.60	Paper V. Session 1	JS
11.	7.8.60	Paper VI. Session 1	NSI, DNG, NSN
12.	7.8.60	Paper VI. Session 2	JS, AM, MVR

Short-Term Course (Junior) : August 1959 - February 1960

		A. Preliminary Examination	
1.	15.10.59	Probability	DB
2.	15.10.59	Presentation of Data & Numerical Mathematics	AM, PKB
3.	26.11.59	Descriptive Statistics & Economic Statistics	JR, NSI
		B. Final Examination	
4.	8. 2. 60	Paper I	DB
5.	10.2. 60	Paper II	JR, AM, NSI
6.	12.2. 60	Paper III	JR, NSN
7.	13.2. 60	Paper IV	AM, NSI
8.	7. 2. 60	Paper V. Session I	JR
9.	7. 2. 60	Paper V. Session 2	SKM, SRSV, PKB
10.	14.2. 60	Paper VI. Session 1	JR, NSI
11.	14.2. 60	Paper VI. Session 2	NSN, AM, PKB

INDIAN STATISTICAL INSTITUTE

Research and Training School

Short-term Statistician's Course - Junior

First Periodical Examination

DATA-PROCESSING AND ECONOMIC STATISTICS

Date : 7 May 1959 Time : 2½ hours

Group A

1. Give two illustrations each of typical statistical data relating on (1) one discrete variable, (2) two continuous variables and (3) two variables and two attributes.

2. Enumerate, giving short examples, three of the usual methods available for scrutinizing statistical data.

3. Mention with brief description the main types of (1) design, (2) data-sheet and (3) observation methods that are used for data collection.

Group B

4. Explain the factor reversal test and the time reversal test for index numbers. Do you or do you not think that the above tests are reasonable? Substantiate your view.

5. What is meant by the trend of a time series? How will you fit a logistic curve to given time series data?

6. From the following price-data,

commodities	prices		price relatives
	1951	1958	
rice	Rs.18/- per md.	Rs.27/- per md.	150
cloth	Rs.1.70 per yd.	Rs.0.85 per yd.	50

one concludes that as the mean of the price-relatives is 100, the price level in 1958 is the same as in 1951. Comment.

You may assume that the above two commodities are the important ones.

oooOooo

Examiners : A.Matthai and S.Krishnamoorthy

INDIAN STATISTICAL INSTITUTE

Research and Training School

Short-term Statistician's Course (Junior)

Second Periodical Test

PROBABILITY, NUMERICAL MATHEMATICS AND DESCRIPTIVE STATISTICS

Date : 21 May 1959 Full marks 100 Time : 3 hours

Figures in the margin indicate full marks.

1. Two dice with faces marked 1,2,3,4,5,6 are thrown, what is the probability that the sum of the marks is 7? (10)

2. Box number one contains 2 red balls and 3 white; box number two contains 5 red balls and 4 white. One ball is transferred from box one to box two and then one ball is drawn from box two.

(a) What is the probability that this ball is red?

(b) If it is known that the ball finally drawn from box two is red, what is the conditional probability that the ball transferred from box one to box two is red? (10)

3. (a) What is a random variable? Define expectation and variance of a random variable and covariance of two random variables.

(b) X_1 and X_2 are two random variables with $E(X_i) = m_i$, $\text{Var}(X_i) = \sigma_{ii}$ and $\text{Cov}(X_1, X_2) = \sigma_{12}$. Let $Y = X_1 + X_2$ and $Z = X_1 - X_2$. Find the covariance between Y and Z.

(c) $X_1, X_2, \ldots, X_n$ are n independent random variables with $E(X_i) = m$, $\text{Var}(X_i) = \sigma^2$. Find the expectation and the variance of

$$\bar{X} = \frac{1}{n}(X_1 + X_2 + \ldots + X_n). \qquad (10)$$

4. If X is a random variable with $E(X) = \mu$ and $\text{Var}(X) = \sigma^2$, show that for $t > 0$

$$\text{Prob}[\,|X - \mu| < \sigma t\,] \geq 1 - \frac{1}{t^2}$$

holds. (10)

5. The National Income of India in 1951 is Rs.99.7 abja (abja = 1000 million), as estimated by the National Income Committee and the Central Statistical Organisation. The population of India according to the census (1951) is 361 million. Compute as accurately as possible the per capita national income. Indicate its reliability, assuming that the official figures are correct to the given number of significant figures. (10)

6. The population of a town in the last four censuses is given below. Estimate the population for 1961.

1921	1931	1941	1951
82,984	86,686	89,547	93,091

(10)

Please turn over

7. From the following series, estimate the missing index number of cost of living of labourers in a town.

Year	1931	1932	1933	1934	1935
Index	149	145	?	131	141

(10)

8. (a) If the mean and standard deviation of n observations, $x_1, \ldots x_n$ are m and σ respectively, and if $u_i = (x_i - m)/\sigma$, show that $u_1, \ldots u_n$ have zero mean and unit standard deviation.

(b) The following table gives the distribution of lengths of 1000 telephone calls through a certain exchange.

time (in seconds)	number of calls
0 - 99	1
100 -199	28
200 -299	88
300 -399	181
400 -499	248
500 -599	262
600 -699	134
700 -799	42
800 -899	11
900 -999	5

Shifting the origin at the mid-point of the class-interval 400-499 seconds and taking 100 seconds as the unit, we get from the above table (10)

$$\sum f u = 280, \quad \sum f u^2 = 2278.$$

Find the mean and the standard deviation of the distribution.

9. Give an example from practice to which the Poisson distribution is a suitable model. Show that for the Poisson distribution the mean is equal to the variance. (10)

10. (a) The telephone exchange referred to in Q. 8(b) is going to put an upper limit of 10 minutes to the duration of telephone calls and they want to know the percentage of calls which may be cut short when this upper limit is introduced.

Assume that the lengths of telephone calls are normally distributed with the mean and standard deviation as obtained in 8(b), and find the required per centage.

(b) With the above assumption, find also the time limit that should be introduced such that only one call in a hundred is cut short. (10)

oooOooo

Examiners : J. Roy, C.S.Ramakrishnan and P.K.Bhattacharyya.

INDIAN STATISTICAL INSTITUTE
Research and Training School

Short-term Statistician's Training Course (Junior)

Third Periodical Examination
Descriptive Statistics, Sample Survey and Vital Statistics.

Date: 10 June, 1959. Figures on the margin indicate full marks Time: 3 hours

1. For the joint distribution of two variates X and Y, let $E(X) = \lambda$, $E(Y) = \mu$, $E(Y|X = x) = \mu(x)$, $\beta = \text{Cov}(X, Y)/\text{Var}(X)$ and $\alpha = \mu - \beta\lambda$. Show that (30)

$$E(Y - \mu)^2 = E[Y - \mu(X)]^2 + E[\mu(X) - \alpha - \beta X]^2 + E[\alpha + \beta X - \mu]^2$$

and discuss the significance of this result.

OR

For each student completing a certain technical course in the past, the score in the B.Sc. examination (X), the score in an admission test (Y) and the score in the final test on completion of the course (Z) are available. This time there are three hundred applicants for admission to the course from whom only fifty have to be selected. For each applicant the score X in the B.Sc. examination and the score Y in the admission test are known. Describe how you would utilise the information to make a rational selection.

2. Discuss briefly the relative merits and demerits of (10)

(i) complete enumeration survey
(ii) sample enumeration survey.

3. Write short notes on the following explaining clearly the concepts involved: (10)

(i) random sample (ii) unbiased estimator
(iii) standard error.

4. Describe the selection procedures for the following sampling schemes (10)

(i) simple random sampling without replacement
(ii) circular systematic sampling.

5. Suppose in a sample of 100 persons drawn with equal probability with replacement 70 persons were found to have a particular disease. Estimate the population proportion of persons having that disease and estimate its sampling variance. (10)

6. Define:-

(a) the gross and net reproduction rates (15)
(b) the true rate of natural increase.

OR

The age distribution and the age specific death rates at a certain time are available for the whole of a country as well as for a particular city in that country. Why is it not proper to compare the general mortality conditions experienced by the population in that city with the general mortality conditions experienced by the entire population of the country on the basis of crude death rates ? Suggest a suitable method of comparison.

7. Define the essential columns of a life table and explain briefly the concept of 'life table population'. (15)

Examiners: J. Roy, M.N.Moorty and P.K. Bhattacharya.

INDIAN STATISTICAL INSTITUTE
Research and Training School

Short-term Statistician's Course (Junior)

Final Examination, August, 1959.

Paper I

Date: 14 August, 1959. Time: 6 P.M.-9 P.M.

1. From a box containing 3 white and 7 black balls, 2 balls are drawn. What is the chance that none of the balls drawn is black when the drawing is (a) with replacement (b) without replacement ?

2. Box X contains 2 white and 3 black balls and box Y contains 4 white and 5 black balls. One ball is taken from box X and put in box Y and then a ball is drawn from box Y. If this ball is white, what is the conditional probability that the ball transferred was also white ?

3. What is a random variable ? If μ is the expectation and σ^2 the variance of a random variable X, show that for any given positive number t

$$\text{Prob}\,(|X - \mu| \geq \sigma t) \leq 1/t^2$$

4. The joint probability distribution of two random variables X and Y is given below:

(a) Find the probability distribution of $Z = X + Y$.

(b) Find also the conditional probability distribution of $Z = X + Y$ given that $W = X^2 + Y^2 = 5$.

Y \ X	0	1	2	3
-2	.05	.10	.02	.08
-1	.15	0	.05	0
0	.08	.07	.05	.10
1	.02	.04	.04	.15

5. Let x be a random variate such that

$$E(x) = \theta$$

Further $X_1, \ldots, X_n$ denote n observed values of x. Then define:

(i) A linear estimate of θ.
(ii) A linear unbiassed estimate of θ.

Prove that of all the linear unbiassed estimates of θ the estimate

$$\frac{X_1 + \ldots + X_n}{n}$$

has the smallest variance.

6. How would you justify the use of an unbiassed estimate with smallest variance?

Examiners: J. Roy and V.P.Godambe.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Training Course (Junior)
Final Examination, August 1959
PAPER II

Date 18 August, 1959. Time: 6 P.M.-9 P.M.

1. EITHER. Write a small note (not exceeding 300 words) on the statistical system in India.

OR. Write the names of at least two official publications, (together with its frequency of publication and the name of the publishing authority) giving statistics relating to each of the following topics: (a) agriculture (b) trade and (c) education in India.

2. Fifteen experiments were performed to determine the boiling point of a certain oil. The observations were in Fahrenheit Scale. The arithmetic mean in Centigrade scale was computed in two different ways, (a) the arithmetic mean of the original observations was computed and was then converted to Centigrade Scale by the formula $C = 5/9 (F - 32)$, where C and F are measurements of the same temperature in Centigrade and Fahrenheit respectively, and (b) each observation was converted to Centigrade Scale and their arithmetic mean was computed. The two methods gave the same result. But the geometric mean of Fahrenheit observations converted to Centigrade was different from the geometric mean of the observations in Centigrade.

Explain why the two methods of computation gave the same result for arithmetic mean but different results for geometric mean ? Can you say whether the medians obtained by the above two methods will be the same ? Give reasons.

3. A manufacturer of Shuttlecocks has found that one percent of his products are defective because of improper distribution of weight. The shuttlecocks are sold in dozen boxes. What theoretical distribution would fit an observed distribution of number of defectives in dozen boxes. What is the probability that in a dozen box (a) no defective (b) more than one defectives ?

4. The following statistical constants were obtained for the joint distribution of the yield of dry bark in ounces (y) and the age in years (x) of Cinchona plants:

Linear regression equation of y on x : $y = 2.5 + 1.4x$

coefficient of correlation $r = 0.6$ correlation ratio $\eta = 0.7$ variance of y : $\sigma^2 = 16.0$.

Write a small note in non-technical language explaining the significance of these constants. What estimate would you give for the yield of dry bark from a 5 year old Cinchona plant and what is the margin of uncertainty of your estimate ?

P.T.O.

5. To select computers for appointment in a statistical institute, the applicants were given two psychological tests. The scores obtained in these tests by computers already working in the institute and their efficiency-ratings are also available. How would you make use of these data scores in these two tests in selecting the best applicants ?

6. Give a large sample test for testing whether the standard deviation of one normal population is C times that of another.

7. What are the conditions under which the chi-square test of goodness of fit (or the frequency chi-square test) can be applied ? Explain how you will determine the degree of freedom of the chi-square statistic, and discuss why a frequency chi-square test is not appropriate in examining the goodness of fit of a logistic curve describing the growth of a population.

Examiners:- J. Roy, S.K.Mitra, P.K.Bhattacharya and S. John.

INDIAN STATISTICAL INSTITUTE
Research and Training School

Short-term Statistician's Training Course (Junior)

Final Examination, August 1959
Paper III

Date: 20 August 1959. Time: 6 P.M. - 9 P.M.

1. Write short explanatory notes on each one of the following three basic principles in design of experiments:

(i) randomisation, (ii) replication, (iii) error control.

2. Why is the use of Latin square designs advocated at times in preference to simple randomized block lay outs ? Draw up a Latin Square lay out for a comparative trial with six treatments. (Describe how you randomised). Give the (blank) analysis of variance table for the above lay out showing the degrees of freedom for the different components.

Suppose in the analysis of the above lay out the treatment effect comes out significant. If you are further interested in the difference between any two particular treatment effect, how would you proceed to test for its significance ?

3. What is a factorial experiment ? Define the terms 'main effects' and 'interactions' as used in this context. What is the usefulness of such experiments ? Draw a (blank) analysis of variance table for a factorial experiment involving 3 factors each at 2 levels in four randomised blocks each of eight plots.

4. Tick the best fitting definition among the 4 given after each statement.

(a) A purposive sampling procedure is one where

(i) there is some purpose behind the sampling procedure adopted;

(ii) the units are purposively divided into a number of groups and samples are drawn at random from each group;

(iii) many purposes are satisfied at the same time;

(iv) the units are selected purposively without any recourse to a random mechanism.

(b) An estimate calculated from a sample of a given size drawn according to a given method is called unbiased, if

(i) it is expected that the value of the estimate obtained from any such sample will be close to the true value, you wish to estimate;

(ii) the average of the different values, the estimate can take weighted by the corresponding probabilities is equal to the true value;

(iii) on drawing samples of larger and larger size, the value of the estimate tends closer and closer to the true value;

(iv) the value of the estimate obtained from any such sample is equal to the true value.

P.T.O.

(c) Proportional allocation in a stratified simple random sample consists in

(i) selecting units in each stratum with probability proportional to the size;

(ii) allocating the number of units to a stratum in proportion to the total number of units in that stratum;

(iii) selecting one stratum with probability proportional to the number of units in that stratum;

(iv) allocating the number of units to a stratum in proportion to the total in the [illegible] of supplementary variable.

(d) A simple random sample of size n drawn without replacement is one where

(i) every unit has an equal chance of occurring in the sample;

(ii) every possible sample of size n (all combinations of size n) has an equal chance of occurrence;

(iii) the chance of any unit coming in the sample is random;

(iv) Every unit in the population has some chance of occurring in the sample.

5. (a) Describe briefly the method of sampling exactly n units systematically when N, the total number of units, is not a multiple of n and give an unbiased estimator of the population total.

(b) How do you make use of the information on a supplementary characteristic available for each unit to get better estimates when adopting a systematic sampling scheme ?

6. (a) State the principles to be followed in stratification and allocation.

(b) What do you consider to be the difference, if any, between stratified sampling and systematic sampling?

Examiners: J. Roy, M.N.Murty and N.S.Nanjamma.

INDIAN STATISTICAL INSTITUTE
Research and Training School

Short-term Statistician's Training Course (Junior)

Final Examination, August 1959.

Paper IV.

Date: 22 August 1959. GROUP A Time: 4 p.m. - 7 p.m.

Answer any two questions.

1. Interpret the control chart constants: d_2, d_3, c_2, D_4, B_1 and A_2.

2. Explain in what different ways could an item quality characteristic, which is a variable, be specified.

   Explain the use of statistical principles, giving example, in determining specification limits.

3. Explain, giving an illustration, the principle of rational subgrouping for control charts.

4. Write notes on : 1. Control chart for individuals.
   2. AOQL.
   3. Acceptance sampling procedures.
   4. Group control charts.

GROUP B

Answer any three questions

5. (a) What is a chain index ? What are its advantages and disadvantages?

   (b) Explain the construction and uses of cost of living index numbers.

6. Describe the main problems that will have to be faced in constructing an index number series. You may consider for this purpose an index number series to be constructed to depict the progress of the Second Five Year Plan. You should also indicate how you will overcome the difficulties.

7. (a) What are the main components of a time series ?

   (b) How will you 'deseasonalize' a given time series ?

8. (a) Comment upon the multiplicative model explaining the resultant of the different components of a time series.

   (b) Given the monthly data relating to the passenger-miles operated by the Indian Railways for the past five years, discuss the procedure of forecasting the demand on Indian Railways for passenger-traffic in September 1959.

Examiners: A. Matthai and S. Krishnamurty.

INDIAN STATISTICAL INSTITUTE
Research and Training School

Short-term Statistician's Course (Junior)

Final Examination, August 1959

Paper V Session 1.

Date: 16 August 1959 Time: 11 A.M.-2 P.M.

1. A frequency distribution of the scores of an unselected group of 150 candidates in an examination are given below :-

Score	Number of candidates
11 - 20	3
21 - 30	16
31 - 40	29
41 - 50	57
51 - 60	26
61 - 70	14
71 - 80	4
81 - 90	1

(a) Find the mean and the standard deviation of the scores.

(b) The pass mark in this examination was 40 and those who passed got one of three grades, A, B,C according to their performance. The candidates scoring 60 or above got grade A, those scoring below 60 but not less than 50 got grade B and all others (who passed) got grade C. Find the percentage of students in the four groups (A, B, C and fail) assuming the scores to be normally distributed with the mean and standard deviation as calculated in (a).

2. The following table gives the distribution of length in cm (x) and weight of dry fibre in gms (y) of 350 jute plants.

Obtain an appropriate formula for predicting y on the basis of x. Find the correlation coefficient and also correlation ratio of y on x.

Length of plant in cms and weight of dry fibre in gms of 350 jute plants.

Weight (in gms) \ Length (in cm)	106.5-118.5	118.5-130.5	130.5-142.5	142.5-154.5	154.5-166.5	166.5-178.5	178.5-190.5	190.5-202.5	202.5-214.5
0.255 - 2.255	15	17	33	31	15	3	1		
2.255 - 4.255	1	1	5	19	60	33	10	2	1
4.255 - 6.255				1	5	23	23	5	1
6.255 - 8.255					1	4	10	14	
8.255 -10.255						1	4	4	1
10.255 -12.255			1					5	

Examiners: J. Roy and P.K.Bhattacharya.

INDIAN STATISTICAL INSTITUTE
Research and Training School

Short-term Statistician's Course (Junior)
Final Examination August 1959.
Paper V. Session 2.

Date: 16 August 1959. Time: 2.30 P.M.-5.30 P.M.

In questions 1, 2, and 3 apply appropriate tests of significance.

1. In 1953 a large firm analysed records for a random sample of cars purchased in 1952 and operated during 1952 by company salesmen. For each car operating expenses - gas, oil, repairs etc., during the months of February, May, July and December, 1952 were ascertained from the salesmen's reports. The total expenses was then divided by the total mileage during these months. The resulting expenses per mile for 12 Fords, 10 Chevrolet and 8 Plymouth are given below:

Operating costs per mile (in cents) for three makes of car.

Chevrolet	Ford	Plymouth
3.93	4.70	2.43
3.45	4.15	2.98
2.00	4.55	3.04
2.28	3.31	4.94
3.49	2.13	3.15
4.25	4.69	2.46
2.38	2.68	3.34
3.02	2.56	2.38
3.26	3.93	
4.08	1.56	
	4.29	
	1.74	
average 3.214	3.341	3.090
standard deviation 0.785	1.192	0.832

(a) Examine, whether there is any difference in the three makes of car, so far as operating costs are concerned.

(b) Do you think the experiment was properly planned? If not, what other observations would you ask for ?

2. When it was found that for a particular breed (Breed 1) of cattle the weight (x, in lbs) of calves at age one was a good predictor of their weight (y, in lbs) at age three, it was decided to investigate the same for two other breeds (breeds 2 and 3). Accordingly data were collected and some statistics were computed, which are presented below:

	Breed 1	Breed 2	Breed 3
Sample size	48	56	31
$\bar{x}$	95.10	84.41	98.49
$\bar{y}$	186.34	152.50	192.32
Var (x)	4916.28	3598.14	5004.31
Cov (x, y)	6012.31	4601.26	5492.67
Var (y)	19648.42	14391.45	20051.42

(a) Examine, for breeds 2 and 3, whether it is possible at all to predict y from x.

(b) Do you think, x is equally effective in predicting y, in all the three breeds ?

3. The relative sensitivity of four explosives A, B, C and D were tested under controlled conditions. Hundred cartridges of explosive A, 65 of B, 85 of C and 40 of D were fired in turn in a standard gas chamber and a record was made of whether each shot resulted in an ignition or not. The results were as follows:

	Ignition	Non-ignition	Total
Explosive A	10	90	100
Explosive B	10	55	65
Explosive C	7	78	85
Explosive D	9	31	40

Is there clear evidence of a difference between the four explosives in liability to cause ignition ?

4. The following values of x and $\log_{10} x$ are given.

x	$\log_{10} x$
1.250	.0969100
1.260	.1003705
1.265	.1020905

Find by linear interpolation the values of $\log_{10} x$ for $x = 1.257$ and $x = 1.262$. Round off your answers such that in each case the error in the last digit retained is less than half a unit.

5. Find by iteration the positive root of the equation

$$x = 3(1 - e^{-x})$$

correct to two places after decimal. [You can start with $x = 3$ as a first approximate.]

Examiners: S.K.Mitra, S.John and P.K.Bhattacharya.

INDIAN STATISTICAL INSTITUTE
Research and Training School

Short-term Statistician's Training Course (Junior)

Final Examination, August 1959

PAPER VI. Session 1.

Date:-23 August 1959 Time: 11 A.M.-2 P.M.

1. A randomised block design was used to compare four treatments in an agricultural experiment. The yields are recorded below.

	treatment			
block	1	2	3	4
1	18.5	16.2	14.1	13.0
2	11.7	12.9	14.4	16.9
3	15.4	15.5	20.3	18.4
4	16.5	12.7	15.7	16.5

(a) Make an analysis of variance of the above data to test for the significance of differences between the treatments.

(b) Suppose the four treatments are composed as follows:

treatment	nitrogen	phosphorus
1	= (10 lbs per plot,	15 lbs per plot)
2	= (25 lbs per plot,	15 lbs per plot)
3	= (10 lbs per plot,	20 lbs per plot)
4	= (25 lbs per plot,	20 lbs per plot)

Sub-divide the sum of squares due to treatments into main effects and interaction sums of squares and test for their significance.

2. (a) Quarterly averages of the index numbers of 'Government and Semi-government Securities' (base 1930=100) for the 4 years 1948-1951 are given in the following table. Obtain the seasonal index by the method of link relatives, assuming a linear trend.

Index numbers of Security prices (1930=100)

	Quarter 1	Quarter 2	Quarter 3	Quarter 4
1948	103.2	102.0	102.3	101.7
1949	101.4	101.5	101.1	101.2
1950	100.9	100.8	100.6	99.8
1951	98.7	98.3	98.2	95.6

(Source: Monthly Abstract of Statistics.)

(b) The table below gives the indices of Seasonal Variation in Absenteeism of Textile Mill workers in various centres in and for the State of Madras. Examine the data and prepare a brief note bringing out the salient features.

Indices of Seasonal Variation.

	Jan.	Feb.	March	April	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
Madras City	97	95	96	101	105	107	101	100	107	100	98	94
Madura	100	111	114	111	111	107	94	93	91	90	90	89
Coimbatore	93	106	112	109	130	113	94	92	90	58	98	87
Tinnevelly	110	118	111	100	110	101	91	87	97	91	95	91
State	98	106	106	109	109	108	94	91	96	93	97	93

Source: Indian Economic Journal, Volume 3, page 84.

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

Business Activity in India

Item	Monthly average in 1948	May 1949	Weights
I. Industrial Production	-	-	46
II. Foreign Trade			6
Exports (Value in Rs.)	1912539	1221671	3
Imports (Value in Rs.)	1040125	2902524	3
III. Shipping Activity			2
Tonnage entered	13376	27848	1
Tonnage cleared	13068	38536	1
IV. Financial Activity			23
Cheek clearance (Rs.crores)	137.8	127.1	18
Note Circulation (Rs.crores)	341.5	315.4	5
V. Inland Trade			
Tons lifted by Railways	5588	7102	23

Examiners: I.M.Chakravarti and S. Krishnamurty.

INDIAN STATISTICAL INSTITUTE
Research and Training School

Short-term Statistician's Training Course (Junior)
Final Examination, August 1959
Paper VI - Division 2.

Date: 23 August 1959 Time: 2.30 p.m.-5.30 p.m.

1. The means and ranges for 6 samples of product, with respect to a dimension, each sample relating to a different shift, are given below. Examine whether the process is in a state of control.

sample no.	no. of items of product in sample	mean dimensions (in inches)	range
1	5	2.3126	0.4216
2	4	2.8625	0.5628
3	5	2.5321	0.3713
4	5	2.4962	0.3596
5	4	2.5716	0.4862
6	5	2.8212	0.3953

Table of control chart constants.

no. of observations in sample	chart for averages: factors for control limits			chart for averages: factors for central line		chart for standard deviations: factors for control limits			
n	A	A_1	A_2	c_2	$1/c_2$	B_1	B_2	B_3	B_4
4	1.500	1.860	0.729	0.7979	1.2533	0	1.808	0	2.266
5	1.342	1.596	0.577	0.8407	1.1894	0	1.756	0	2.089

no. of observations in sample	chart for ranges: factors for central line			chart for ranges: factors for control limits				chart for individual factors for control limits	
n	d_2	$1/d_2$	d_3	D_1	D_2	D_3	D_4	E_1	E_2
4	2.059	0.4857	0.880	0	4.698	0	2.282	3.760	1.457
5	2.326	0.4299	0.864	0	4.918	0	2.115	3.568	1.290

2. The following table contains information on sales for two months for 28 stores in Calcutta. Suppose you know the figures for May 1957 before the survey for May 1958, this information can be used as 'size' for drawing pps. sample to increase the efficiency of the estimate.

   1) Describe briefly the steps involved in selecting a sample of size 5 with probability proportional to size with replacement, by each of the following methods:

      i) Cumulative method
      ii) Lahiri's method

   2) Draw 2 samples of size 5, one by each of the above methods.

   3) Estimate the average sales in May 1958 per store and obtain an estimate of its variance - from each of the two samples.

   4) Comment on the observed difference between the two estimates, and between the two estimates of variance.

P.T.O.

Sales of Grocery stores in Calcutta for May 1957 and May 1958 in Rs.

May 1957	May 1958	May 1957	May 1958
252	333	88	104
405	229	344	393
234	219	152	154
150	189	85	91
103	108	89	102
287	443	61	81
240	264	179	206
278	304	401	485
65	61	190	213
207	152	112	127
138	141	119	142
102	107	128	155
182	203	196	186
32	29	85	105

3. The following table gives the age distribution of women between 12 and 46 and female births (during one year) classified by mother's age as obtained from a sample as well as the probability of survival up to each pivotal age in the female life-table of the population from which the sample was taken.

Age-group	Pivotal age (x)	Female population ('000)	Female Births ('000)	Life-table probability of survival upto age x
12-16	14.5	15068	242	0.5928
17-21	19.5	12759	1605	0.5700
22-26	24.5	13603	1615	0.5455
27-31	29.5	10680	1069	0.5110
32-36	34.5	8569	574	0.4661
37-41	39.5	7486	395	0.4165
42-46	44.5	5900	129	0.3667

(a) Find the specific female fertility rate for each of the above quinquennial age-groups.

(b) In the life-table referred to above, take 1000 as radix and find the l_x values at the pivotal ages.

(c) Find the net reproduction rate, assuming the deaths over each five-yearly age interval to be uniformly distributed over the whole interval and the age specific fertility rate to be constant over each five-yearly age interval.

Examiners: A. Matthai, M.N.Murty, N.S.Nanjamma and P.K.Bhattacharya.

INDIAN STATISTICAL INSTITUTE
Research and Training School

Short-term Statistician's Course (Junior)

Periodical Examination I

Presentation of Data and Numerical Mathematics

Date: 15 October 1959. Time: 7 P.M. to 9 P.M.

1.(a) State the principles to be observed in drawing up Statistical Tables.

(b) With reference to statistical data consisting of two characteristics observed on n individuals, examine what types of information are lost and what types of information are preserved in doing the following:

i) forming a frequency table without class intervals;
ii) forming a frequency table with class intervals;
iii) forming two one-way frequency tables, one for the first characteristic and the other for the second characteristic;
iv) forming a two-way frequency table.

2.(a) State the importance of scrutinising data collected. Giving an illustration in each case, mention four methods available for scrutinising data.

(b) Write a note on each of any four of the following types of graphical presentation of data, bringing out its special features:

i) bar diagram
ii) line curve
iii) scatter diagram
iv) Z-chart
v) component parts chart
vi) semi-logarithmic chart
vii) pictograph.

3. The following experiment was conducted to study the effect of subjecting the seeds to different temperatures on the yield of a certain plant. A large number of seeds were divided into three groups and these groups were subjected to temperatures 60°F, 90°F and 120°F respectively. The results of the experiment are given below.

Temperature (in °F)	Average yield per plant (in gms)
60	15.7
90	22.5
120	18.4

(a) Find by interpolation the average yield per plant from seeds subjected to a temperature of 80°F by making the best use of the above table.

(b) For what temperature in the range covered by this experiment is the average yield a maximum ?

4. The values of the function

$$f(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$$

are given below for $t = 0, .05, .10$. Evaluate

$$\int_0^{0.1} f(t)dt$$

as accurately as possible, using only the given values of $f(t)$.

t	f(t)
.00	.39894
.05	.39844
.10	.39695

Examiners: A. Matthai and P. K. Bhattacharya.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Training Course
(UNION)
Periodical Examination II
Descriptive Statistics and Economic Statistics

Date: 26 November 1959. Time: 6 P.M.-9 P.M.

GROUP A

Answer any three questions

1. The mean (m) and the variance (v) of n = 100 observations were computed as m = 10.00, v = 25.00. It is later discovered that an observation was wrongly recorded as x' = 15.00 while its correct value is x = 5.00.

Find the correct values of the mean and the variance.

2. A sample of 1000 families, each with 5 children, is taken. Write down the expected frequency distribution of the number of male children in the family. Explain in not more than three sentences why you expect such a distribution. What would be the mean and the variance of this distribution ?

3. The average number of deaths per year from street accidents in a certain locality is 2. Find the chance that in a particular year there are 3 or more deaths from street accidents.

4. The quality characteristic of a manufactured product is distributed Normally with mean 20 and standard deviation 5. The lower specification limit of quality is 12; that is, an article is unacceptable if its quality characteristic is 12 or less.

Find the percentage of unacceptable articles.

If changes are made in the production process so that the mean is raised to 22, standard deviation remaining the same; what would now be the percentage of unacceptable goods?

GROUP B

Answer all the three questions

5. How do you eliminate the trend from a time series when it is given

(a) annually for fifty years and
(b) monthly for five years ?

6. Describe briefly the method of compiling a monthly index of industrial production in India. How can a seasonal industry such as the sugar industry be included in such an index ?

7. Graduation by the least square method leads to the same results as by the method of weighted moving average. Illustrate this by fitting a quadratic to five consecutive items in a time series.

Examiners: J. Roy and N. S. Iyengar.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Course, Junior
Final Examination. February, 1960

Paper I

Date: 8 February, 1960. Time: 6 p.m.-9 p.m.

1.(a) State and prove Chebyscheff's Lemma.

(b) Let X be a random variable with mean μ and the 14th central moment μ_{14}. Indicate how you would use Chebyscheff's Lemma to get an upper bound to the probability of the event 'X differs from μ by at least 20'.

2. Suppose we go on tossing an unbiased coin until we get two consecutive heads. Let ν be the number of tosses required for the event to occur.

a) What is the range of the random variable ν ?

b) What is $P(\nu = 5)$?

c) What is $P(\nu > 2)$?

3.(a) State the Law of Large Numbers.

(b) Discuss the bearing of the Law of Large Numbers on the Method of Moments for Estimation.

4.(a) Explain carefully the Maximum Likelihood and the Minimum χ^2 methods of Estimation.

(b) Suppose a coin is tossed n times and let p be the probability of getting a head in a single throw. Show that the Maximum Likelihood Method and the Minimum - χ^2 Method lead to the same estimator for p.

5. Let $X_1, X_2, \ldots X_n$ be n independent observations on a random variable X. Let $EX = \mu$ and $V(X) = \sigma^2$. Let $\bar{X} = \frac{1}{n}(X_1 + \ldots + X_n)$ and $s^2 = \frac{1}{n-1}\sum (X_i - \bar{X})^2$.

a) Prove that s^2 is an unbiased estimator of σ^2 and that s is a negatively biased estimator of σ.

b) Indicate how you would prove that $\frac{100\, s}{\bar{X}}$ is a consistent estimator of the co-efficient of variation of X.

6. Carefully explain each of the following expressions:

a) Cumulative Distribution Function.

b) Sample space.

c) Unbiased Estimator.

d) Statistic.

e) Variance of the sample mean.

Examiner: D. Basu

INDIAN STATISTICAL INSTITUTE
Research and Training School

Short-term Statistician's Course. Junior

Final Examination. February 1960.

Paper II

Date: 10 February 1960 Time: 6 P.M.-9 P.M.

Answer the two groups in seperate books

GROUP A

Attempt any two questions

1. A variable X takes the values x with relative frequencies

$$f_x = (1 - p)p^x \qquad (0 < p < 1)$$

for $x = 0, 1, 2, \ldots \ldots \infty$. Find the mean, median, mode and variance of this distribution.

2. (a) Write down the expression for the pdf of a Normal distribution with mean μ and sd σ.

(b) The frequency distribution of the monthly income of individuals in a certain town is available in about ten classes. It has been suggested that the logarithm of income follows the Normal distribution. Describe how you will examine this point.

(c) The distribution of heights of males is Normal with mean μ and standard deviation σ. The distribution of heights of females is Normal with mean μ' and standard deviation σ'. The ratio of males to females in the population is $p : (1 - p)$. For the combined distribution of heights of males and females find the first four moments.

3. (a) Consider the joint distribution of two random variables X and Y. Show that $D = E(Y - \alpha - \beta X)^2$ is minimised if $\alpha = E(Y) - \beta E(X)$ and $\beta = \text{Cov}(X, Y)/V(X)$ and that the minimum value of D is

$$D_{min} = V(Y) - [\text{Cov}(X, Y)]^2/V(X).$$

(b) Write ten sentences to explain the statistical significance of the above result, introducing the concepts of regression, best linear formula for prediction, correlation coefficient and correlation ratio.

GROUP B

4. Discuss with the help of examples, why scrutiny of data is necessary and how far data can be improved through scrutiny.

OR

What are the essential parts of a statistical report. Mention the important principles to be observed in report-writing.

(Do not spend more than 30 minutes on this question).

5. Describe the role of the Central Statistical Organisation in the current statistical system of India.

OR

Information is required on the items given below:

(i) Overall balance of trade in merchandise and treasure of India.

(ii) India's defence expenditure during the last few years.

(iii) Central Government employment in any recent year.

(iv) Production of cotton yarn and cloth variety-wise and state-wise.

Indicate where they are primarily published and mention, if any, their limitations.

OR

Name two important enquiries collecting statistics on a voluntary basis with an all-India coverage. What types of information are collected in these enquiries and how are they collected?

(Do not spend more than 30 minutes on this question).

Examiners: J. Roy, A. Matthai and N.S.Iyengar.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Course. Junior
Final Examination, February 1960

Paper III

Date: 12 February 1960 Time: 6 P.M.-9 P.M.

Answer the two groups in separate books

GROUP A

1. Explain the roles played by Randomisation, Replication and Error control in planning an experiment.

[Do not write more than fifteen sentences]

2. Describe the randomised block and the Latin Square layouts. In blank tables of analysis of variance, show for each layout the partitioning of the degrees of freedom.

Write not more than five sentences to describe the disadvantages of these two layouts.

3. Define the term 'interaction' as used in connection with a factorial experiment involving two factors, each at two levels.

Given that: main effect of A = 10, main effect of B = 16 and interaction AB = 4 units; complete the following table of average effects for each of the four factorial combinations:

		levels of A	
		a_1	a_2
levels of B	b_1	80	*
	b_2	*	*

(The higher levels of A and B are denoted by a_2 and b_2 respectively).

GROUP B

4. Enumerate the advantages and disadvantages of sample surveys as compared to complete censuses, explaining the reasons for your points in brief.

5. (a) The number of persons and total expenditure on food for each of the 4 families is given below. If a sample of 2 families is drawn with equal probability without replacement, devise a suitable estimator to estimate the per person expenditure on food for this population. Calculate (i) the bias, (ii) the sampling variance and (iii) the mean square error of your estimator:

Family serial number	No. of persons in the family	Total exp. on food in Rs.
1	5	80
2	1	40
3	3	60
4	4	100

(b) Write short notes on (i) sampling unit, (ii) sampling frame, (iii) random number tables.

P.T.O

6. Prove that for a sample of n units drawn without replacement the sample mean is an unbiased estimate of the population mean. Derive an unbiased estimate of the variance of this estimate.

7. Is stratified sampling a case of purposive sampling ? - Explain. What is meant by optimum allocation. For stratified simple random sampling with replacement, derive the formula for the optimum allocation when the total cost is fixed and the cost of enumeration of a unit varies from stratum to stratum. In most cases the within stratum variances for the characteristic under consideration are not known. What then is the practical utility of this allocation ?

Examiners: J. Roy and N.S.Nanjamma.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Course. Junior.
Final Examination, February, 1960.
Paper IV

Date: 13 February, 1960. Time: 3 P.M.-6 P.M.

Answer the two groups in separate books

GROUP A

Answer any three questions

1. Explain how 'process capability' and 'consumer requirement' are concerned with framing specifications.

2. Distinguish between the terms 'statistical control' and 'economic control' State how a control chart may be defined or described.

3. Examine how, when a quality characteristic may have any unknown distribution, the use of three-sigma limits are justified in a control chart. Why not use exact 'probability limits' on control charts at least when the form of the distribution is known?

4. Distinguish between the two situations in each of the following pairs:
   i) control charts when standards are given and when standards are not given;
   ii) control charts for small samples and for large samples;
   iii) control charts for individuals and for samples of individuals.

5. State the problem in lot acceptance sampling and indicate how acceptance sampling plans are generally constructed.

GROUP B

Answer any two questions

6. What is meant by the trend of a time series ? Give an account of the different methods of obtaining the trend of a given time series.

7. Compare the Fixed-base method with the chain-base method in calculating an index number. What are the considerations in favour of the latter method ?

8. Explain what is meant by the Index-number of seasonal variation in economic time series. Describe the Ratios to trend method of constructing such indices.

Examiners: A. Matthai and N.S.Iyengar.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Course Junior
Final Examination. February, 1960

Paper V. Session 1.

Date: 7 February 1960. Time: 11 a.m.-2 p.m.

1. The following table gives the frequency distribution of the annual incomes of the members of a certain profession (based on a sample):

annual income in rupees	frequency
below 4000	163
4001 - 5000	435
5001 - 8000	600
8001 - 10000	198
10001 - 15000	176
15001 - 25000	86
25001 - 40000	38
above 40000	15

It is proposed to introduce a professional tax on the incomes at the following slab rates:

annual income in rupees	rate of tax (percent)
below 4200	0
4201 - 6000	2
6001 - 8400	5
8401 - 12000	8
12001 - 18000	10
18001 - 24000	12
above 24000	15

(a) The total number of members of the profession is estimated as 25000. Estimate the amount that would be realised from the tax.

(b) Also examine to what extent the variation in the incomes is diminished because of the imposition of the tax.

[Only approximate answers are required; but you should clearly explain what assumptions you are making in your computations. Do not spend more than 45 minutes on this question]

2.(a) A purchaser uses the following rule: From each lot containing a very large number of articles, he takes a sample of 5 articles and accepts the lot only if there is no defective articles among these five. If the proportion of defectives in a lot is 0.15 what is the chance that he will accept the lot ?

(b) The average number of misprints per page in a book of 1000 pages is 1.5. In how many pages would you expect the number of misprints to be 3 or more ?

(c) The distribution of scores in a certain examination was found to be Normal with mean 51.2 and standard deviation 7.4. Find the percentage of students scoring not less than 45 and not exceeding 60.

[Do not spend more than 45 minutes on this question].

p.t.o.

3a. The following table gives the distribution of 100 casts of steel by the percentage of iron in the form of pig iron (x) and the lime consumption in hundredweight per cast (y).

Obtain the linear regression equation of y on x. Compute the coefficient of correlation and the correlation ratio of y on x.

Estimate the lime consumption for a cast for which the percentage of iron in the form of pig iron is 35.

Frequency distribution of percentage of pig iron and lime consumption

lime consumption	percentage of pig-iron						
in cwt. per cast.	20-24	25-29	30-34	35-39	40-44	45-49	50-54
100 - 124				1			
125 - 149		2	1	7	1		
150 - 174	1	1	6	6	1	2	2
175 - 199			6	12	3	6	5
200 - 224				3	8	11	3
225 - 249			1	1	2	3	1
250 - 274						1	1
275 - 299				1			
300 - 324						1	

Examiner: J. Roy.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Course. Junior.
Final Examination. February, 1960.

Paper V. Session 2.

Date: 7 February 1960. Time: 2.30 p.m.-5.30 p.m.

1. In the 400 digits occurring in a page of certain random number tables the frequency of integers 0, 1, ..., 9 were observed as follows.

integer	0	1	2	3	4
frequency	47	41	38	35	48
integer	5	6	7	8	9
frequency	41	42	30	46	32

(a) Do you consider this to be sufficient evidence against randomness? (Hint: If the integers were really drawn at random what would be their expected frequencies ? Test whether the observed frequencies differ significantly from the expected frequencies).

(b) It was suspected that the method by which these numbers were generated would lead to a preponderence of even integers. Can the observed excess in the frequency of even integers over odd be explained by chance? Apply suitable test.

2. An eye specialist wanted to determine whether eye defects are hereditary among males. For that he examined pairs (father and son) and classified the eye defects in 3 groups: normal, slight and serious. The following 3 x 3 table gives the distribution of the 278 pairs classified according to the degree of seriousness of eye defects in father and in son. What type of distribution is expected if eye defects are not hereditary ? Analyse the data and apply suitable test to determine if there is any indication as to the hereditariness of these defects.

		eye defect in father			
		normal	slight	serious	total
eye defect in son	normal	84	24	9	117
	slight	31	61	18	110
	serious	11	18	22	51
	total	126	103	49	278

3. A factory manufactures bullets with a diameter specification of 0.2 inches. However although the mean diameter of bullets manufactured was 0.2 inches the variance was as high as 0.0008 (inches)^2. The production manager of the factory wanted to increase the efficiency of the production process by reducing the variance in diameter. He tried a new process and from a sample of size 100 estimated the variance as 0.0004 (inches)^2. Test if the apparent improvement can be due to random causes or due to actual improvement in the quality. (State the assumptions involved in your test).

4.(a) Twenty samples were taken from a container of a particular brand of Vanaspati. Four analysts were given five different samples each and asked to determine the melting point. The results are given below.

Examine by suitable test if the analysts are consistent in their determinations of the melting point.

Melting point of Vanaspati in degrees Fahrenheit

analyst	individual determinations of melting points				
1	93.60	94.64	96.30	93.62	93.51
2	96.44,	96.53	98.38	97.00	97.63
3	92.57	94.01	92.49	93.29	90.87
4	95.55	95.90	94.25	95.80	96.21

(b) The analysts 1 and 2 learnt their techniques from two different laboratories. Do they agree in their determinations ? Apply suitable test.

5. X is a random variable with

$$P[X > a] = (1 + \frac{a}{2})e^{-a/2}, \qquad a \geq 0.$$

Find as accurately as possible the upper 5 percent point of the distribution of X with the help of the following table:-

a	$(1 + a)e^{-a}$
4.6	.05629
4.7	.05184
4.8	.04773

6. Find the values of the function

$$f(x) = 2x - \log_{10} x - 7$$

at some points of x lying between 3 and 4 and with the help of these find an approximation to the root of the equation $f(x) = 0$ lying in this interval. Taking this as a first approximate, apply any suitable method of iteration to arrive at an accuracy of 4 places after decimal.

Examiners: S.K.Mitra, S.R.Srinivasa Varadhan and P.K.Bhattacharya.

Indian Statistical Institute
Research and Training School
Short-term Statistician's Course. Junior
Final Examination. February, 1960

Paper VI. Session 1.

Date: 14 February, 1960 Time: 11 A.M.- 2 P.M.

1. Prepare a randomised layout for a factorial experiment involving three factors each at two levels in four blocks each of eight plots.

Draw up the blank analysis of variance table and show the degrees of freedom for each component.

2. The following table gives the results of an experiment with four treatments A, B, C, D in a 4 x 4 Latin square. Carry out the analysis of variance and arrange the treatments in increasing order of yield.

B 154	A 144	D 137	C 132
D 137	C 176	B 177	A 129
C 138	D 139	A 121	B 171
A 98	B 173	C 133	D 128

3. Compute the general index number of wholesale prices for the week ended 12th December 1959 from the following table:

Table 1. Index number of wholesale prices for sub-groups of commodities.

	group and sub-group	weight	(Base: 1952-53 = 100) index number for week ended 12.12.1959.
I	food articles	(504)	...
	(a) cereals	382	100.9
	(b) pulses	84	91.2
	(c) fruits and vegetables	45	120.8
	(d) milk and ghee	167	110.4
	(e) edible oils	93	128.3
	(f) fish, egg and meat	34	112.8
	(g) sugar and gur	95	150.9
	(h) others	100	187.9
II	liquor and tobacco	(21)	101.4
III	fuel, power, light and lubricants	(30)	116.5
IV	industrial raw materials	(155)	...
	(a) fibres	393	119.6
	(b) oil seeds	388	130.4
	(c) minerals	14	95.2
	(d) others	205	130.4
V	manufactures:	(290)	112.6
	A. Intermediate products	141	118.2
	B. Finished products	859	111.7
	(a) Textiles	506	108.3
	(i) cotton	313	116.3
	(ii) jute	126	93.9
	(iii) woollen	12	104.2
	(iv) silk and rayon	55	96.3
	(b) metal products	41	144.6
	(c) chemicals	70	106.9
	(d) oil cakes	33	145.2
	(e) machinery and transport equipment	106	106.0
	(f) others	103	113.7

P.T.O.

4. Using the data below calculate the indices of seasonal variation of quarterly revenue receipts by the method of moving averages:

Table 2. Quarterly revenue receipts in a treasury.

year	1st quarter Jan-March	II quarter April-June	III quarter July-Sept.	IV quarter Oct.-Dec.
1950	80.8	75.9	102.0	93.5
1951	122.6	103.5	138.1	119.9
1952	136.3	96.8	118.1	113.5
1953	101.0	92.2	111.0	96.0
1954	107.4	85.0	117.4	109.3

Note: The figures are in units of Rs.10^7

Examiners: J. Roy and N.S.Iyergar.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Course. Junior
Final Examination. February, 1960.

Paper VI, Session 2.

Date: 14 February 1960. Time: 2.30 p.m.-5.30 p.m.

1. A survey is to be conducted to estimate the total income of a group of 100 persons within Rs.1000/- of the true value with a 19 in 20 chance. For this purpose find the sample size required when the sample is drawn with equal probability without replacement. the population standard deviation for income is known to be Rs.25.

Probability level	50	80	90	95	99
Normal deviate	0.67	1.28	1.64	1.96	2.58

2. To study the reaction of employees in a factory to the question, 'Are you willing to work in the night shift ?' a very quick sample survey was undertaken by selecting 180 out of 900 regular employees and 80 out of 600 casual employees with equal probability with replacement. The following table gives the distribution of the employees by the type of answers given.

Stratum	Answers		Total
	Yes	No	
Regular employees	80	100	180
Casual employees	60	20	80

Estimate the number of employees (regular and casual combined) in the factory who are (i) willing to work in the night shift, (ii) not willing to work in the night shift.

Estimate the coefficients of variation of the estimates obtained. Which of the estimates do you think is more reliable ? Can you explain why?

3. The following table contains information on sales for two months for 28 stores of Ontario. Suppose before the survey you know the figures of sales for May 1958 (Col. 2) and you are to estimate the total sales of those 28 stores for May 1959 through a sample survey.

Sales of Grocery stores in Ontario for
May 1958 and May 1959

sl. no.	May 1958	May 1959	sl. no.	May 1958	May 1959
1	252	353	15	88	104
2	405	229	16	344	393
3	234	219	17	152	154
4	150	189	18	85	91
5	103	108	19	89	102
6	287	443	20	61	81
7	240	264	21	179	206
8	278	304	22	401	485
9	65	61	23	190	213
10	207	152	24	112	127
11	138	144	25	119	142
12	102	107	26	128	155
13	182	203	27	186	186
14	32	29	28	89	105

p.t.o.

Answer either A or B

(A) (i) Draw a simple systematic sample of 5 stores, profitably using May, 1958 sales.

(ii) Obtain an unbiased estimate of the total sales in May,1959.

(iii) Can you estimate the standard error of your estimate using only the sample drawn ? Obtain an estimate of the standard error of your estimate by a suitable method.

(B) (i) Draw a sample of 5 stores with probability proportional to the sales in May,1958, with replacement.

(ii) Obtain an estimate of the total sales in May, 1959.

(iii) Obtain an estimate of the standard error of your estimate.

4. One specimen of a chemical in liquid form was collected from each of 10 consecutively produced bulks, and inspected for a quality characteristic (x) and the results are given below.

bulk No.	1	2	3	4	5	6	7	8	9	10
value of x:	12.3	13.8	12.7	14.6	13.5	12.6	14.0	14.9	12.3	15.6

Set-up suitable control charts and interpret.

OR

Samples were taken from 8 different consignments of a product in order to find out if any assignable causes of variation existed in a quality characteristic (x). Examine the results given below, for control and write a brief note giving your conclusions.

consignment	average	standard deviation
1	33.5	4.20
2	32.7	4.35
3	33.9	4.38
4	40.6	4.56
5	35.6	4.51
6	33.2	4.39
7	34.5	4.90
8	35.7	4.86

(Do not spend more than 45 minutes on this question).

5. For the life table of a population, the following are given

x	20	21	22	23	24
q_x (per thousand)	10.49	10.76	11.09	11.40	11.67

$l_{20} = 82742$, $T_{20} = 2,733,007$

Complete all columns of the life-table for x = 20 to 24.

(Do not spend more than 45 minutes on this question).

Examiners: N.S.Nanjamma, A. Matthai and P.K.Bhattacharya.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Training Course, Junior.
Periodical Examination I
Probability, Data Processing and Numerical Analysis.

Date: 5 April 1960. Time: 6 P.M.-9 P.M.

1. Define very carefully the following terms:

(a) Sample space, (b) The favourable cases for an event
(c) Mutually exclusive events, (d) The range of a random variable
(e) The modal value of a random variable.

2. Find the probability distribution of each of the following three random variables (no proof required). In each of the three cases the random experiment is that of rolling a symmetric die twice.

(a) X is the absolute difference between the two scores.
(b) Y is the product of the two scores.
(c) Z is the larger of the two scores if they are unequal and is the common score if they are equal.

3.(a) State the different principles to be observed in the framing of a good data-sheet. Give illustrations.

(b) Prepare a blank layout of a statistical table and indicate its different parts.

(c) Why scrutiny of data ? Mention the different methods you know for scrutinising data. Give examples whenever possible.

4. Write down the expression for the polynomial function $f(x)$ passing through the points (x_0, y_0), (x_1, y_1), ..., (x_n, y_n), and explain briefly how it comes useful in inverse interpolation.

5. Give an example where it is very difficult to evaluate a definite integral analytically but numerical integration provides a rough and ready answer.

Derive Simpson's One-third rule for numerical integration.

Examiners: D. Basu, A. Matthai and P.K.Bhattacharya.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Course, Junior.
Periodical Examination
Probability, Descriptive Statistics and Economic Statistics

Date: 17 May, 1960 Time: 6 P.M.-9 P.M.

Figures in the margin indicate full marks.
Answer group A and group C in one answer book and group B in a separate answer book.

Group A

Do not spend more than 1 hour on this group.
Answer all questions.

1. Assuming that $E(X+Y) = E(X) + E(Y)$ for all X and Y, prove that $V(X-Y) = V(X) + V(Y) - 2\,\mathrm{Cov}\,(X, Y)$. [8]

2. How often must a symmetric die be rolled in order to make the probability of getting at least one six equal to at least $\frac{1}{2}$? Give reasons. [8]

3. Given that the random variable X takes the four values -1, 0, 1 and 2 with equal probabilities, find out

   (i) $E(X)$; (ii) the probability distribution and the mode of the random variable X^2. [8]

4. Given that the random variable X^2 takes only the two values 0 and 1 with equal probabilities, can you find the distribution of X ? If not, why not ? If you are given the further information that the mode of X is zero, can you find the probability distribution of X ? What if you are told that $E(X) = 0$? [8]

Group B

Do not spend more than 1 hour 15 minutes on this group
Attempt as many questions as you like. The maximum that you can score on this group is 40.

5. What class-intervals will you take for preparing a frequency distribution in each of these cases ?

   (i) heights of adult Bengali males recorded to the nearest inch [min. ht: 56", max. ht: 74", no. of individuals: 500]

   (ii) monthly expenditure per person in households in India [min. exp. Re 0.83, max. exp. Rs 5256.70, no. of households: 5000]

   (iii) age returns suspected to be rounded off to multiples of two or five in many cases. [2+2+2]

6. Let $x_1, x_2, \ldots, x_N$ be N values of a variate. Write down the expressions for the arithmetic mean A, the geometric mean G and the harmonic mean H of these values. For $N = 2$, prove that $A \geq G \geq H$ and that if any two of them are equal, all the three of them must be equal.

   State without proof if the above inequalities hold generally for any N. [2+5+1]

7. What is meant by the cumulative distribution function (cdf) of a variate ?

A population consists of five students, and the characteristic noted for each is his score in a certain test. The scores are 75, 50, 60, 75 and 80. Find the cdf. [2+1]

8. What is a median ? Find the median for the distribution of scores in question 7. [2+2]

9. Write down the expression for the variance of N values, $x_1, x_2, \ldots, x_N$.
Show that $\sum_{i=1}^{N} (x_i - c)^2$ is minimum when $c = \frac{1}{N} \sum_{i=1}^{N} x_i$. [2+3]

10. If the probability of success in a single trial is θ, write down (without proof) the probability of getting x successes in n independent trials. Write down the mean and the variance of the number of successes (proof not required). [2+1]

11. If the average number of misprints per page in a book is μ, and if the book has in all N pages, how many pages would you expect to be free of misprints ? Write two sentences to justify your answer. [2+2]

12. Write down the expression for the frequency density function of a Normal distribution with mean μ and standard deviation σ.

Let $P(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}t^2} dt$. The life in hours of an electric bulb is Normally distributed with mean 1000 and standard deviation 80. If you buy a large number of bulbs, what fraction would you expect to live for not less than 900 hours but not more than 1080 hours ? Write answer in terms of the P(x) function - numerical value is not required. [3+5]

13. Show that for a Normal distribution $\beta_1 = 0$, $\beta_2 = 3$. [8]

14. For a bivariate distribution of (X, Y), write down the expressions for

(i) the least-squares linear regression coefficient of Y on X

(ii) ρ : the coefficient of correlation of X and Y

(iii) η : the correlation ratio of Y on X . [3+3+3]

15. Prove that as defined in question 14.

$$1 \geq \eta^2 \geq \rho^2 \geq 0$$ [10]

P.T.O.

Group C

Do not spend more than 45 minutes on this group.
Answer any two.

16. (a) Explain the construction and uses of cost of living index numbers.

(b) Describe the procedure adopted in India for constructing weekly index numbers of wholesale prices. [14]

17. What do you understand by the seasonal index of variation in economic time series ? How would you determine such indices from quarterly time series data ? [14]

18. Write short notes on :-

(a) chain index numbers

(b) variate difference method

(c) factor reversal test. [14]

Examiners: D. Basu, J. Roy and N. S. Iyengar.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Training Course, Junior.
Periodical Examination

Descriptive Statistics - Practical

Date: 28 May 1960. Time:- 3 P.M.-5 P.M.

Neatness in computation carry some marks.

1. The data in table I gives the frequency distribution of the diameter of 534 wedding rings.

   (a) Find the expected probabilities in each class interval assuming the distribution to be normal.

   (b) If a jeweller wants to manufacture 10,000 wedding rings how should be distribute them in these class intervals.

   (c) Find the actual third decile of the distribution and the theoretical third decile as computed from the normality assumption.

2. The following table gives the bivariate frequency table of the ages of the husband and wife among 100 couples. (Refer to Table II).

   (a) Compute the correlation coefficient between the ages of the husband and the wife.

   (b) Obtain the linear prediction formula for predicting the age of the wife on the basis of the age of the husband.

Examiner: J. Sethuraman.

P.T.O.

Table I.

Diameter in mms.	frequency
0.65 - 0.85	39
0.85 - 1.05	75
1.05 - 1.25	105
1.25 - 1.45	116
1.45 - 1.65	98
1.65 - 1.85	66
1.85 - 2.05	35
	534

mean = $\bar{x}$ = 1.2861 variance = 0.1073 s.d. = 0.3276

Table II. X : age of the husband

Y : age of the wife

X	Y	15–25	25–35	35–45	45–55	55–65	f_u
	u \ v	-2	-1	0	1	2	
20–30	-2	5					5
30–40	-1	9	10	1			20
40–50	0	3	25	12	4		44
50–60	1		2	2	16	4	24
60–70	2				5	2	7
	f_v	17	37	15	25	6	100

You may make use of the following computations:

$$\Sigma u = 8 \quad \Sigma v = -34 \quad \Sigma u^2 = 92 \quad \Sigma v^2 = 154$$

So that all you now need is Σuv, which is all that you have to obtain from the bivariate table.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Training Course, Junior.
Periodical Examination

Economic Statistics and Vital Statistics, Practical.

Date: 28 May, 1960. Time:- 5.15 P.M.-8.15 P.M.

Figures in the margin indicate full marks.

1. The following table gives index numbers of wholesale prices by groups and sub-groups of commodities for the week ended March 19, 1960, with March 1953 = 100. Fill up the gaps where necessary and compute the overall index of wholesale prices for the given week. Write a short summary based on your calculations.

	groups and sub-groups	weight	index for week ended 19.3.60.
I	food articles	(504)	....
	(a) cereals	382	103.1
	(b) pulses	84	89.5
	(c) fruits and vegetables	45	127.0*
	(d) milk and ghee	167	113.5
	(e) edible oils	93	136.2
	(f) fish, eggs and meat	34	115.9
	(g) sugar and gur	95	140.1
	(h) others	100	156.4
II	Liquor and tobacco	(21)	96.1
	(a) tobacco	941	93.8
III	fuel, power, light and lubricants	(30)	117.7
IV	industrial raw materials	(155)	.....
	(a) fibre	393	126.5
	(b) oil seeds	388	141.2
	(c) minerals	14	98.1*
	(d) others	205	131.8*
V	manufactures	(290)	.....
	A. intermediate products	141	121.1
	B. furnished products	859	.....
	(a) textiles	506	115.0
	(i) cotton	313	124.4*
	(ii) jute	126	97.9*
	(iii) woollen	12	104.2*
	(iv) silk and rayan	55	102.9*
	(b) metal products	41	148.2
	(c) chemicals	70	105.2
	(d) oil cakes	33	141.4
	(e) machinery and transport equipment	106	108.5
	(f) others	103	113.8
	All commodities	(1000)	.....

* : estimated

N.B. Figures in brackets indicate the weights assigned in combining the five groups into a single all commodities index. Weights attached to various sub-groups within groups are shown without brackets. (10)

2. The following table gives monthly average spot prices of gold in the Bombay Bullion Market during January 1956 - February 1960. Construct suitable indices for the monthly variation in prices of gold.

Table: Spot prices of gold in the Bombay Bullion Market (Rs. per tola).

year	Jan.	Feb.	March	April	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
1956	96	98	104	104	104	103	101	104	103	104	105	104
1957	106	108	106	106	108	106	107	108	108	107	109	108
1958	110	113	112	113	112	109	106	107	108	109	113	114
1959	117	118	120	122	122	121	120	121	122	124	125	125
1960	130	131										

Source: RBI Bulletins and CSO Monthly Abstracts of statistics. (15)

3. In the three departments of a large factory the number of days lost by sickness per worker per annum were: Department A, 9.03; Department B, 8.68; and Department C, 9.42. The following table gives the age-distribution of the workers in the above three departments and the number of days lost on account of sickness in a year. Calculate a standardized rate for each department using the total population of the factory as the standard. Examine the pattern exhibited by the crude and standardized rates.

Age in years	Department A		Department B		Department C	
	no. of workers	days lost	no. of workers	days lost	no. of workers	days lost
20 -	100	650	350	2450	50	300
30 -	220	1540	400	3400	70	455
40 -	420	3570	100	1000	120	840
50 -	300	3000	50	600	300	2850
60 -	140	1890	50	800	250	3000
total	1180	10650	950	8250	790	7445

(15)

4. A total of 567990 children were born in a particular year in a country which had its female population in the age group 15-44 years distributed as shown in table below. The proportional distribution of births by age of mother as obtained from a sample study is also given in column (3). Assuming that the ratio of female babies to all births is 48 %, calculate the gross reproduction rate.

age	female population ('000)	proportion of births
(1)	(2)	(3)
15 - 19	789	.2305
20 - 24	765	.3355
25 - 29	717	.2474
30 - 34	642	.0978
35 - 39	540	.0635
40 - 44	435	.0253
all ages	3888	1.0000

(10)

Examiners: N. S. Iyenger and Shd M.V.Raman

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Training Course - Junior
Final Examination, July - August, 1960
Paper I. Session 2.

Date: 29 July 1960. Time: 8 P.M. - 9 P.M.

1. Discuss very briefly the reasons for requiring that an estimate be unbiased and of minimum variance.

State without proof the Cramér-Rao inequality, and give an example illustrating its use.

2. Define the term 'random sampling without replacement' in the context of sampling from a finite population. Show that, for such a sample, the sample mean is an unbiased estimate of the population mean. State without proof the formula for the variance of this estimate.

3. Describe the method of maximum likelihood.

Suppose that in a random sample of 900 households in a certain large city, it was found that 100 households possessed at least one musical instrument. What is the maximum likelihood estimate of the fraction of such households in the city ? What is the (estimated) standard deviation of this estimate ?

Examiner: R. R. Bahadur.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Training Course - Junior
Final Examination, July-August, 1960.

Paper II.

Date: 1 August 1960. Time: 6 P.M.-9 P.M.

Answer questions 1, 10 and any six of the rest.
You are advised to devote 30 minutes each for questions 1 and 10 and 20 minutes each for the other six questions that you answer.
Answers must be brief and to the point.

1. EITHER

(a) Discuss the importance of scrutinising statistical data.

(b) State the important principles to be observed in the drawing up of statistical tables.

OR

(c) Enumerate the different parts of a typical statistical report. Briefly indicate the requirements of a good report.

(d) Explain briefly (i) questionnaire, (ii) schedule, (iii) mail enquiry and (iv) census.

2. Suppose you are given the frequency distribution of the incomes of households in Calcutta. Explain clearly how you will compute from this table:

(a) the total income of all households in Calcutta

(b) the total income of the poorest 10 percent, 20 percent, ..., 90 percent of the households in Calcutta. (You may use interpolation, if necessary).

3. A population consists of N individuals with characteristic values $X_1, X_2, \ldots, X_N$. Write down the expressions for

(a) the arithmetic mean

(b) the variance

(c) the third and the fourth moments

(d) β_1 and β_2 coefficients

for this distribution.

4. The arithmetic mean of the scores of 100 students in an examination was 48 and the standard deviation was 4. Show that the number of students who scored between 40 and 56 could not have been less than 75.

5. For a certain frequency distribution, a computer obtained for the beta-coefficients, $\beta_1 = 1.8$, $\beta_2 = 2.5$. The supervising statistician said that his computations were wrong. What do you say ? Give reasons.

6. Explaining the symbols used, write down the probability laws for the Binomial, Poisson and Normal distributions.

P.T.O.

7. The linear regression equation of weight W in Kg on height H in cm for adult Bengalee males was found to be

$$W = \lambda + \mu H$$

Convert this for use when weight is measured in lbs and height h is measured in inches. [1 Kg = 2.2 lbs, 1 inch = 2.5 cm, approximately].

8. Explain the terms correlation ratio η_{xy}, and correlation coefficient ρ. Without going into algebraic details, show that

$$1 \geq \eta \geq \rho \geq 0$$

When can we get equalities in the above relations ?

9. Discuss how the scores (x) in the B.Sc., examination and (y) in the special psychological test of candidates seeking admission to the Short-term Statistician's Course could be used to select the most suitable students for the course. [You may make use of the scores (x), (y) and the scores (z) in the final statistics examination of students who had previously gone through this course].

10. EITHER

Mention some of the special features of the 1951 population census as compared to the previous census.

OR

Describe the multi-purpose character of the Indian National Sample Survey.

OR

Write a note on the available statistics regarding the Indian manufacturing sector.

Examiners: A. Matthai, J. Roy and M. S. Iyengar.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Training Course: Junior
Final Examination, July-Aug. '60.

Paper VI

Date 3 August 1960 Time: 6 P.M.-9 P.M.

Answers should be brief and to the point

GROUP A

Answer any three

1. Examine the role of randomisation, replication and error control in planning an experiment.

2.(a) Describe a Latin square layout and explain how the fundamental principles of randomisation, replication and error control are made use of in this design.

(b) Draw up a blank analysis of variance table for this case and show the degrees of freedom for each of the components.

3.(a) What are the advantages of a factorial experiment over a one-factor-at-a-time type of experiment ?

(b) Two factors A and B were tried, each at two levels 0 and 1 and the average effects μ_{ij} for the factorial combination $A_i B_j$ (i, j = 0, 1) were found. From these, the main effects and interactions were found as:

$$A = 10, \quad B = 8, \quad AB = -4$$

units. Given that the lower level of the factors is indicated by the symbol 0 and that $\mu_{00} = 20$ fill up the follow table of average effects

	A_0	A_1
B_0	20	?
B_1	?	?

4. Design a simple experiment to find out which of two teachers of statistics is more successful.

GROUP B

Answer all questions

5. Comment in detail on the following statements:

(i) Stratified sampling is a type of purposive sampling.
(ii) The main objection to non-random selection in a sample survey is not that the sample may furnish less precise results, but that there is no scientific means to ascertain its precision.
(iii) Sample survey experts always advocate retaining of periodical censuses.

6. For a simple random sample of size n drawn without replacement, $\bar{x}$ is an unbiased estimate of $\bar{X}$,

$$V(\bar{x}) = \frac{N-n}{Nn} S^2 \quad \text{and} \quad \hat{V}(\bar{x}) = \frac{N-n}{Nn} s^2, \qquad \text{in the}$$

usual notation. Using these results, for a simple random sample of n units drawn without replacement, find an unbiased estimate of the proportion of units in the population in a specified class. Find also the variance of the estimate and an unbiased estimate of the variance in terms of the proportions in the population and sample respectively. Explain clearly the notations you use.

7. The task of a statistician does not end with stratification, - he has to allocate wisely.

(i) Explain the dangers of unwise allocation.

(ii) Has proportional allocation any special advantage over all other allocations ?

(iii) When does optimum allocation become equivalent to proportional allocation ?

Examiners: J. Roy, D. N. Ghosh and N. S. Nanjamma.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Training Course : Junior
Final Examination, July-August 1960
Paper IV

Date: 5 August 1960. Time: 6 P.M.-9 P.M.

Answer all questions

1. Explain the terms: Statistical Control and Economic Control.

2. Distinguish between control charts using given standards and those using no standards.

3. Justify the use of 3-sigma limits as control limits. Which are the situations in which you may not use 3-sigma control limits ?

4. Explain with the help of an example the meaning of rational sub-grouping.

OR

Write notes on: (i) the percentage defective, (ii) the u-chart
(iii) mean-range, (iv) the control chart constant D 4.

5. Discuss the general problem of business forecasting.
Briefly describe some of the methods commonly employed and consider their validity.

OR

What do you understand by the term business cycles ? Describe the Harvard methods of constructing economic barometers.

6. Explain briefly the method of compiling an index number of production. How can a seasonal industry like the sugar industry be included in such an index even though it is monthly ?

OR

Stuvel defines the following price index (Econometrica, 1957)

$$I = \frac{L_p - L_q}{2} + \sqrt{\left(\frac{L_p - L_q}{2}\right)^2 + \frac{V_1}{V_0}}$$

where L_p and L_q are the Laspeyre-type price and quantity indeces, V_0 and V_1 being the values at the base period and current period respectively. Examine whether the above Index satisfies time reversal and factor reversal tests.

Examiners: A. Matthai and N. S. Iyengar.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Training Course - Junior
Final Examination, July-August, 1960

Paper V. Session 1.

Date: 31 July, 1960. Time: 11 A.M.-2 P.M.

1. EITHER.

The following gives the distribution of deaths per day of aged women (aged over 90 years) from the Times obituary over a period of three consecutive years

x = deaths per day	0	1	2	3	4	5	6	7	8	Total
f = number of days	222	262	339	151	79	32	6	4	1	1096

$\Sigma xf = 1937$, mean = 1.7673

(a) Assuming that the number of deaths per day follows a Poisson distribution, calculate

(i) the theoretical probability of 4 deaths in a day,

(ii) the expected number of days with 3 deaths in the next year of 365 days.

OR

The scores in English of students follows a normal distribution with mean 45 and standard deviation 6.

(i) If 30 is the pass mark what is the proportion of students that will pass?

(ii) What should be the pass mark if 70 percent of the students should pass?

(iii) What is the fourth central moment of this distribution?

2. The following results are available from a sample of size 50 on (X, Y)

$\Sigma x_i = 200$; $\Sigma y_i = 400$; $\Sigma x_i^2 = 1600$; $\Sigma y_i^2 = 3650$; $\Sigma x_i y_i = 1900$;

(i) Calculate the correlation coefficient between x and y.

(ii) Calculate the best linear predictor of Y based on X and predict Y when X = 4.

(iii) What interpretation do you give to this predicted value ?

3. Compute (i) the multiple correlation coefficient of X_0 on X_1 and X_2 and (ii) the best linear predictor of X_0 on the basis of X_1 and X_2 from the following data

	corrected sum of squares			means
X_1	54.76			21.7
X_2	3.64	18.23		6.3
X_0	23.86	1.79	15.12	10.6

Examiner: J. Sethuraman

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Training Course - Junior
Final Examination, July-August, 1960

Paper V. Session on 2

Date: 31 July 1960. GROUP Time: 2.30 P.M.-5.30 P.M.

Answer all questions. Suggested time: 2 hr. 15 min.

1. The standard deviation of stature (in inches) of two groups of boys, in of ages 11-13 and the other of ages 14-16, estimated from two samples of sizes 10 and 15 respectively, are 1.15 and 2.3. Does this prove that in respect of stature, adolescent boys (ages 14-16) are more heterogeneous?

2. A potential buyer of electric bulbs (40 W, 220 V) bought 100 each bulbs of each of 2 different brands. On testing the bulbs, he found that brand A had a mean life of 1282 hours with a standard deviation of 80 hours whereas, brand B had a mean life of 1200 hours with a standard deviation of 94 hours. Can the buyer be quite certain that the 2 brands do differ in quality ?

3. The coefficient of correlation between two characters estimated by two investigators from samples of sizes 20 and 25 were 0.318 and 0.253 respectively. Would you regard the difference as due to fluctuation of sampling?

4. My friend Cheatercock and myself contributed equal amounts towards the cost of cultivating a land in a village near Calcutta. Our land almost rectangular in shape (160 yds. x 40 yds) was divided by ails and bunds into 16 square plots (20 yds x 20 yds each) in 8 rows and 2 columns as shown in the plan below.

When the crop was almost ready for harvest, my friend, who was hard-pressed for money requested for his share of the crop. Since I didn't think it wise to harvest the whole crop immediately, I said that he can have the yields from eight of the sixteen plots. To be fair, I insisted that the choice of 8 plots he would immediately harvest should be done in an unbiased manner. A statistician to whom the problem was referred suggested the following random selection which he claimed secures an equitable distribution of the total yield. First, the eight plots should consist of one from each row. Second, to select a plot from each row an unbiased coin has to be tossed (independently for each row) and if head appears the left hand side plot is chosen and otherwise the right hand side plot.

I asked Cheatercock to follow this procedure and select 8 plots and I could not be actually present when the selection was made. He agreed to supply me the figures for individual yields of the plots. A few days later I harvested the rest of the plots. The following table gives the relevant information:-

The plan and yields of the 16 plots (units unspecified)

(The plots selected by Cheatercock are marked as C)

Row No.	Plot No.	Plot yield	Plot No.	Plot yield
1	(1)	10 (C)	(2)	11
2	(3)	7	(4)	12 (C)
3	(5)	9	(6)	8 (C)
4	(7)	10 (C)	(8)	6
5	(9)	10	(10)	10 (C)
6	(11)	7 (C)	(12)	5
7	(13)	13 (C)	(14)	9
8	(15)	14 (C)	(16)	10

(a) Show that whatever may be the yields of the individual plots the expected difference in the total yields between 8 plots chosen in the suggested manner and the rest is zero. (Recall expectation of a variable is $\sum p_i x_i$ where x_i denote the possible values of x and p_i the probability attached to it).

(b) Compute the probability of obtaining a total yield of 8 plots higher than what Cheatercock actually got under the suggested random selection.

(c) On the basis of this computed probability do you think I have reason to suspect that Cheatercock's selection was not strictly according to the agreement. Did he deliberately choose plots which he thought would give him a higher yield ?

(d) When Cheatercock was cornered by me, he declared that he did not try to choose one plot from each row as suggested but that he numbered the plots from 1 to 16 as indicated in the plan and chose 8 Nos. from 1 to 16 giving equal chance to all possible selections of 8 and harvested the corresponding plots. Compute the probability that by following the procedure he indicated, one gets 8 plots with every row represented by one plot. On the basis of this computed probability do you think he is telling another lie to cover up his previous lie ?

GROUP D

Answer any two. Suggested time: 45 min.

5. Find the values of $Z(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}$ at $t = 0, 0.5, 1.0, 1.5, 2.0$.

Hence evaluate $\int_{-2}^{2} Z(t)\,dt$ applying a suitable formula for quadrature.

6. Making use of the following table of $P(x) = \int_{-\infty}^{x} Z(t)\,dt$ find the value of x for which $P(x) = .90$ applying Lagrange's formula for inverse interpolation

x	P(x)
1.27	.89796
1.28	.89973
1.29	.90147

7. Find the root of the equation

$$x^3 - 2x - 5 = 0$$

lying between 2 and 3 correct to 5 places of decimals using any suitable method.

Examiners: S.K.Mitra, S.Das Gupta and P.K.Bhattacharya.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Training Course. Junior
Final Examination, July-August 1960

Paper VI. Session 1.

Date: 7 August 1960. Time: 11 A.M. - 2 P.M.

Answer all questions

1. The following table gives the exports of two varieties of cotton yarns and five varieties of cotton manufactures from the United Kingdom during 1938 and 1943 by quantities and values :

	quantities		values	
	1938	1943	1938	1943
Cotton yarn	(mn lbs.)	(mn lbs.)	£(000)	£(000)
X	170.00	16.37	8397	2817
Y	12.93	2.82	1278	531
Cotton manufactures	(1000 sq.yds)	(1000 sq. yds)	£(000)	£(000)
A	235.3	69.32	3841	2854
B	421.6	83.15	7776	3319
C	281.9	102.90	7152	5805
D	368.8	90.86	10628	6381
E	78.58	28.08	2564	1922

Calculate a quantum index $Q_{1938,\,1943}$ for the British export for 1943 using 1938 as base. Show also that $Q_{1938,\,1943} \neq \frac{1}{Q_{1943,\,1938}}$ by computing $Q_{1943,\,1938}$.

2. The following table gives the estimated personal expenditure on Fuel and Light in the United Kingdom.

		Quarterly £ (million		
		1945	1946	1947
Quarter	1	78	84	92
	2	62	64	70
	3	56	61	63
	4	71	82	85

Assuming a linear trend, calculate the seasonal index.

Forecast the expenditures for the first three quarters of 1948.

3. For the population of 5 persons, given in the table, it is proposed to estimate the <u>total expenditure</u> of the population from a simple random sample of size 2 drawn without replacement. If information on income can also be collected from each individual at the time of survey it is proposed to use the biased estimate $Y_R = (\bar{y}/\bar{x})\,X$ where $\bar{y}$ and $\bar{x}$ are the sample means of expenditure and income respectively and X is the total income for the population. Calculate the bias and mean square error of this estimate by enumerating all possible samples of size 2.

Person no.	1	2	3	4	5
Expenditure (in Rs.)	7	9	14	10	8
Income (Rs.)	9	12	19	15	10

P.T.O.

4. The following table gives a list of 51 urban blocks of Kerala State divided into two strata. Draw one sample of 3 blocks from stratum 1 and another sample of 6 blocks from stratum 2 and estimate the total number of unemployed persons in this population from the samples you have drawn. Also estimate the variance of your estimate.

Stratum 1

Block No.	No. of unemployed persons	Block No.	No. of unemployed persons
1	12	9	7
2	10	10	12
3	15	11	12
4	5	12	6
5	19	13	11
6	11	14	11
7	13	15	12
8	11	16	11

The page and random number column used are to be given.

Stratum 2

Block No.	No. of unemployed persons	Block No.	No. of unemployed persons
1	11	19	10
2	2	20	2
3	13	21	15
4	11	22	15
5	9	23	8
6	9	24	11
7	10	25	11
8	11	26	10
9	7	27	7
10	10	28	10
11	14	29	11
12	12	30	10
13	14	31	11
14	7	32	11
15	10	33	6
16	14	34	1
17	10	35	12
18	13		

5. A sample survey is to be conducted among 1200 factories to estimate their total monthly production. The error that can be tolerated is 10 % of the true value on either side with a confidence level of 90 %. The coefficient of variation of the distribution of monthly production is known to be approximately 60 %. Calculate the sample size required if the sample is to be drawn with equal probability without replacement.

Prob. level	90 %	95 %	99 %
Normal deviate	1.64	1.96	2.58

Examiners: N.S.Iyengar, D. N. Ghosh and N.S.Nanjamma.

Use of practical records is not allowed.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Training Course, Junior
Final Examination July- Aug.'60.
Paper VI. Session 2.

Date 7 August 1960. Time: 2.30 P.M.-5.30 P.M.

Use of practical records is not allowed

1. EITHER

In order to compare the errors of measurements in shoot heights among 6 samplers, the following experiment was conducted 6 areas were selected and the 6 samplers were allocated in 6 different orders to the areas so that no two samplers work in the same region at the same time. The results of such a latin square arrangement are given below.

Samplers' errors in shoot heights (cm.)

Order	Areas 1	2	3	4	5	6	Totals
1	F + 3.5	B + 4.2	A + 6.7	D + 6.6	C + 4.1	E + 3.8	+ 28.9
2	B + 8.9	F + 1.9	D + 5.8	A + 4.5	E + 2.4	C + 5.8	+ 29.3
3	C + 9.6	E + 3.7	F - 2.7	B + 3.7	D + 6.0	A + 7.0	+ 27.3
4	D + 10.5	C + 10.2	B + 4.6	E + 3.7	A + 5.1	F + 3.8	+ 37.9
5	E + 3.1	A + 7.2	C + 4.0	F - 3.3	B + 3.5	D + 5.0	+ 19.5
6	A + 5.9	D + 7.6	E - 0.7	C + 3.0	F + 4.0	B + 8.6	+ 28.4
Total	+ 41.5	+ 34.8	+ 17.7	+ 18.2	+ 25.1	+ 34.0	+ 171.3

$$\Sigma y_{ij}^2 = 1144.73$$

Analyse the data for the equality in capacity of the samplers. If they are not equal, rank them according to their ability.

OR

Three factors, Vitamin (V), Protein (P) and another nutritive substance (K) were used each at two levels, to form eight treatment combinations. All the treatments were given to the rabbits in each litter were randomised before the treatments were given. The following table gives the increases in weights.

P.T.O.

Litter	Increases in weights (gms) of individual rabbits.								
	(O)	(P)	(K)	(VK)	(V)	(PK)	(VPK)	(VP)	Total
1	31	75	49	110	90	60	53	60	528
	(V)	(O)	(VK)	(VPK)	(P)	(PK)	(VP)	(K)	
2	110	50	109	57	90	102	81	63	662
	(P)	(K)	(O)	(PK)	(VPK)	(V)	(VK)	(VP)	
3	60	42	45	87	60	105	110	53	562
4	(VP)	(VPK)	(P)	(V)	(K)	(VK)	(O)	(PK)	
	65	59	95	94	50	103	32	75	571

$$\Sigma y_{ij} = 2325, \quad \Sigma y_{ij}^2 = 187405.$$

Analyse the results in detail and write a short report.

(Suggested time: 1 hour)

2. Obtain the values for the upper and lower control limits for the following, given that from each of 10 batches a sample of size 4 was taken and that the overall mean was 14.2, the average of standard deviations was 2.1, and the mean-range was 4.3.

(i) $\bar{x}$ - chart using range
(ii) s - chart
(iii) R - chart

(Suggested time: 30 minutes)

3. Examine the following data for (i) statistical control and (ii) control with respect to a standard value of 2 end-breaks per yard.

Piece No.	No. of yds.	No. of end-breaks
1	10	32
2	15	31
3	10	27
4	10	23
5	10	24
6	12	21
7	15	18
8	15	32

(Suggested time: 30 minutes)

P.T.O.

4. The following table gives for a certain district the age distribution of population and deaths from all causes in 1940 and 1950.

(a) Calculate the crude and age-specific death rates for the two years. Mention the factors which might account for the time change in the crude death rate in general.

(b) Calculate for each calendar year a total death rate, adjusted for age differences in the populations.

Age (years)	1940		1950	
	Population estimated July 1	Deaths	Population estimated July 1	Deaths
Under 1	65080	4426	52610	2484
1 - 4	284810	1428	229087	480
5 - 14	778218	1217	658729	539
15 - 24	712233	1631	755223	979
25 - 34	646098	2318	669885	1412
35 - 44	631898	3531	615807	2678
45 - 54	500021	5513	564329	5625
55 - 64	356864	8502	402077	9180
65 - 74	196588	10351	259122	13435
75 and over	77804	10416	109852	14344
Total	4249614	49333	4316721	51156

(Suggested time: 1 hour)

Examiners: J. Sethuraman, A. Matthai and M. V. Raman.

Indian Statistical Institute
Research and Training School

Short-Term Course (Senior) : February 1959 - August 1959

LIST OF EXAMINATIONS

Sl.no.	Date	Subject	Examiner/s
		A. Periodical Examination	
1.	7. 5. 59	Algebra and Estimation	VPG, SKM
2.	24.6. 59	Sample Surveys	SRR, MNM
3.	29.7. 59	S.Q.C.	AM
4.	25.8.59	Paper I	
		B. Final Examination	
4.	25.8.59	Paper I	SKM
5.	27.8.59	Paper II	RRB, DB, VPG
6.	12.9.59	Paper III (Practical) Session 1	SKM
7.	12.9.59	Paper III (Practical) Session 2	SKM
8.	9.9.59	S.Q.C. (Paper IV)	AM, SQC Unit
9.	10.9.59	Paper V. Sample Surveys	SRR
10.	29.8.59	Paper VI. S.Q.C. Session 1	SQC Unit
11.	29.8.59	Paper VI. S.Q.C. Session 2	AM
12.	23.8.59	Paper VII. (Practical) Session 1	SRR, MNM
13.	23.8.59	Paper VII. (Practical) Session 2	SRR, MNM

Short-Term Course (Senior) : September 1959 - February 1960

		A. Preliminary Examination	
1.	17.12.59	Economic Statistics	AKB, MMukherjee, NSI
2.	24.12.59	Algebra and Estimation	SKM
3.	31.12.59	S.Q.C.	AM
4.	16.2.60	Paper I	SKM, JR
5.	P		
		B. Final Examination	
4.	16.2.60	Paper I	SKM, JR
5.	18.2.60	Paper II	RRB, SRR
6.	27.2.60	Paper III (Practical) Session 1	SKM, SJ
7.	27.2.60	Paper III (Practical) Session 2	JR, SRR
8.	23.2.60	Paper IV (S.Q.C.)	AM, SQC Unit
9.	25.2.60	Paper V (Econometrics)	AKB, MMukherjee, NSI
10.	14.2.60	Paper VI (S.Q.C.) Session 1	AM
11.	14.2.60	Paper VI (S.Q.C.) Session 2	SQC Unit
12.	20.2.60	Paper VII (Econometrics) Session 1	NSI, AKB
13.	20.2.60	Paper VII (Econometrics) Session 2	AKB, MMukherjee
14.	20.2.60	Paper VII (Sample Surveys) Session 1	SRR
15.	20.2.60	Paper VII (Sample Surveys) Session 2	SRR

Indian Statistical Institute
Research and Training School

Short-Term Course (Senior) : March 1960 - August 1960
LIST OF EXAMINATIONS

Sl.no.	Date	Subject	Examiner/s
		A. Periodical Examination	
1.	24.5.60	Planning & Econometric Statistics	AKB
2.	26.5.60	S.Q.C.	SQC Unit
3.	2.6.60	Algebra and Sample Surveys	SKR, SRR
4.	~~30.6.60~~	~~Paper-I~~	~~SKR, JS, KRS~~
		B. Final Examination	
4.	10.8.60	Paper I	SKR, JS, KRS
5.	12.8.60	Paper II	DB, JS, SRR
6.	7.8.60	Paper III. Session 1	SKR, JS
7.	7.8.60	Paper III. Session 2	KRS, SRR
8.	17.8.60	Paper IV (SQC)	AM, SQC Unit
9.	13.8.60	Paper V (SQC) Session 1	AM
10.	13.8.60	Paper V (SQC) Session 2	SQC Unit
11.	19.8.60	Paper VI (Econometrics)	AKB, NSI
12.	20.8.60	Paper VII. (Econometrics) Session 1	NSI
13.	20.8.60	Paper VII, (Econometrics) Session 2	AKB

INDIAN STATISTICAL INSTITUTE.
Research and Training School

Short-term Statistician's Course - Senior

First Periodical Examination

ALGEBRA AND ESTIMATION

Date : 7 May 1959. Time : 3 hours.

Group A

1. What is meant by a subspace of a vector space V? Consider the vector space V_4 of 4 tuples (x_1, x_2, x_3, x_4).

(i) Pick out from this collection V_4 all vectors whose first co-ordinate $x_1 = 1$ and form a subcollection S_1. Is S_1 a subspace? If not, why not?

(ii) Pick out from V_4 all vectors whose first two co-ordinates x_1 and x_2 satisfy the equation.

$$2x_1 + 3x_2 = 0$$

and form a subcollection S_2. Give an example of a vector which belongs to S_2. Is S_2 a subspace? If so, what is its dimension? Determine its dimension by first obtaining a basis of this subspace. You are required to find a basis which contains (3,-2, 1, 0) and (3,-2, 0,1) as members.

(iii) Also obtain a normal orthogonal basis of this subspace S_2.

(iv) Express both (3, -2, 1,0) and (3,-2.0,1) as linear combinations of the vectors in the normal orthogonal basis which you obtain in (iii).

(v) The equation in (ii) can also be interpreted as implying that the vector (x_1, x_2, x_3, x_4) is orthogonal to some fixed vector (a_1, a_2, a_3, a_4). Obtain numerical values for a_1, a_2, a_3, and a_4.

2. Let A and B be the matrices given below

$$A = \begin{bmatrix} 1 & 3 & 7 \\ 2 & 3 & 2 \\ 6 & 1 & 5 \end{bmatrix}$$

$$B = \begin{bmatrix} 2 & 9 \\ 8 & 6 \\ 3 & 0 \end{bmatrix}$$

(i) Evaluate AB, (ii) What is the rank of A? (iii) Consider the equations

$$x_1 + 3x_2 + 7x_3 = 1$$

$$2x_1 + 3x_2 + 2x_3 = 5$$

$$6x_1 + x_2 + 5x_3 = 2$$

Using the matrix A given above and defining two other matrices suitably, can you express these equations in a compact form? You may use the definition of product of matrices given in your class.

Group B

3. For a distribution of a random variate, which is completely specified excepting for one unknown parameter, define the following terms:-

1. An unbiased estimate of the parameter
2. An unbiased estimate with minimum variance
3. An unbiased estimate having variance equal to the Cramer-Rao lower bound

Please turn over

4. Explain the role of the maximum likelihood method of estimation in (3) above.

5. Under what situation would you use an unbiased estimate with minimum variance? Give the justification.

oooOooo

Examiners: S.K.Mitra and V.P.Godambe

INDIAN STATISTICAL INSTITUTE
Research and Training School

Short-term Statistician's Training Course - Senior (March-Sept.59)

Mid-term class test.

Sample Surveys.

Date: 24.6.59. Time: 3 hours

N.B. (i) Figures in the margin indicate full marks.

(ii) Only Brief and Precise answers are expected.

1. What are the stages of work involved in planning a survey ? (List the points only). [12]

2. (i) What is a sampling frame ? [3]
(ii) What are the likely defects in a sampling frame ? [6]
(iii) How do these defects in a sampling frame affect the estimate from a sample ? [3]

3. Explain briefly any four of the following: [12]
(i) Simple random sampling with replacement.
(ii) Stratified sampling.
(iii) Multistage sampling.
(iv) Double sampling.
(v) Cluster sampling.
(vi) Systematic sampling.
(vii) Sampling with preassigned unequal probabilities.
(viii) Interpenitrating sampling.

4. (i) What is the difference between cluster sampling and systematic sampling ? [7]
(ii) In a sample survey for estimating the total of a character x in a population, the sampling frame is sometimes rearranged according to increasing order of an auxiliary character, y (known for all units in the population). Under what condition is this advantageous ? [5]

5. Either

Under what conditions is systematic sampling better than simple random sampling ? (Derive the necessary condition).

Or

Explain how you would estimate the gain due to stratification from a sample survey data obtained from a stratified sample (simple random sample with replacement). [14]

6. A village consists of 15 households. The list of 15 households along with the number of persons and value of cereals consumed by the households in March 1955 is given.

(i) Draw two samples of 5 households each with probability proportional to number of persons in households systematically. [10]

[Note: If you cannot draw the sample systematically draw samples according to PPS random sampling scheme. But, the marks will be reduced by 50 percent].

P.T.O.

(ii) Estimate the total value of cereals consumed by the households in the village unbiasedly from both the samples separately and also show that the estimator used is unbiased. [16]

(iii) Build up an unbiased estimate of total value of cereals consumed in the village by suitably combining the estimates from the two samples mentioned in (ii). [4]

(iv) From the two sample estimates obtained in (ii), work out the estimate of variance of the combined estimate mentioned in (iii). [8]

Note: All steps should be shown and explained clearly. Computations should be done in neat tabular form as far as possible.

Complete list of 15 households in the village.

serial number of households	number of persons	value of cereals consumed (Rs.)	serial number of households	number of persons	value of cereals consumed (Rs.)
1	5	43.75	9	5	45.75
2	5	36.25	10	4	33.00
3	4	43.25	11	3	21.88
4	3	21.56	12	10	86.75
5	3	32.50	13	7	63.75
6	9	106.25	14	5	44.50
7	5	31.50	15	6	55.75
8	5	43.25			

Examiner:- S. R. Rao

INDIAN STATISTICAL INSTITUTE
Research and Training School

Short-term Statistician's Course - Senior
Periodical Examination
Statistical Quality Control

Date: 29 July 1959. Time: 6 P.M. - 8.30 P.M.

1. Describe in detail any typical AQL plan, paying attention to principles of construction of plan, presentation of tables and method of using of the plan.

2. State a single sampling acceptance plan for variables for a one sided specification limit and for unknown lot standard deviation. Derive the elements in the plan.

3. Explain any four of the terms: (1) AOQL, (2) Reduced and Tightened Inspection, (3) Inspection levels, (4) Narrow limit gauging (5) Snap reading, (6) Components of variance.

4. The standard deviation of errors in measurement of a characteristic is determined to be 5 units and the standard deviation of the measured values of the characteristic 30 units. Obtain the value of the true standard deviation of the characteristic.

Examiner: A. Matthai.

INDIAN STATISTICAL INSTITUTE
Research and Training School

Short-term Statistician's Course (Senior)

Final Examination, August-September 1959

Paper I

25 August 1959 — 6 P.M.- 9 P.M.

1. Let A be the matrix

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 0 & 2 \end{pmatrix}$$

Compute (a) $\Sigma = AA'$ and (b) $(AA')^{-1}$.

2. Solve the following linear equations:

$$x_1 + 2x_2 + 2x_3 = 2$$
$$2x_1 + 5x_2 + 4x_3 = 4$$
$$2x_1 + 4x_2 + 5x_3 = 3$$

3. Compute $\underline{a}\,\Sigma^{-1}\,\underline{a}'$, where $\underline{a}$ is the vector (2, 4, 3).

4. State reasons why no vector other than the null vector can be orthogonal to (a) each row of A (b) each row of Σ.

5. (a) What is meant by a multivariate normal population? State at least two properties of such a distribution.

(b) Do you have reasons to suspect that (2,4,3) could not have been drawn in random sampling from a trivariate normal population with mean values zero and variance covariance matrix Σ (as given in 1(a)).

6. Explain the following terms : (a) estimable parametric functions, (b) best linear unbiased estimate, (c) zero function in the context of linear estimation theory.

7. Stating clearly the problem (assumptions and the hypothesis) derive the analysis of variance for oneway classification from appropriate theorems (which you are required to state) in the theory of least squares.

oooOooo

Examiner : S.K.Mitra.

INDIAN STATISTICAL INSTITUTE
Research and Training School

Short-term Statistician's Training Course (Senior)

Final Examination, August-Sept.1959
Paper II

Date: 27 August 1959 Time: 6 P.M.-9 P.M.

GROUP A

Answer any two questions

1. (a) Let X be a Poisson variable with mean μ and let Y take the two values 0 and 1 with equal probabilities. If X and Y be independent then what are the prob. distributions of

(i) X + Y and (ii) XY ?

(b) If X and Y are defined as before then what should be μ in order that X + Y and X - Y are incorrelated ? Can they be independent ? Give reasons.

2. (a) Let $X_1, X_2, \ldots X_n$ be independent normal variables with means μ and s.d.'s σ. Briefly sketch how the sampling distribution of $\frac{\bar{X} - \mu}{s/\sqrt{n}}$ is obtained.

(b) If the X's be defined as above then what is the distribution of $(X_1 - X_2)^2 + (X_3 - X_4)^2$?

3. (a) Describe the Beta and the Gamma distributions. Find the probability distribution of the ratio of two independent Gamma variables with the same scale parameter.

(b) If X and Y be independent standard normal variables then what is the distribution of $\frac{X^2}{X^2 + Y^2}$?

GROUP B

Answer both questions

4. $X_1, \ldots, X_n$ are n independent observations on a random variate x, having the frequency distribution $p(x \mid \theta)$, which is completely specified but for the unknown value of the parameter θ.
Define:
(i) An unbiased estimate of θ,
(ii) A uniformly minimum variance, unbiased estimate of θ.

5. With the above notation, write the expression for the Crammer-Rao lower bound for the variance of an unbiased estimate of θ. Give an illustrative example, showing that an unbiased estimate may have uniformly minimum variance, but its variance may not attain the Crammer-Rao lower bound.

P.T.O.

GROUP C

Answer both questions

6. State and prove the lemma of Neyman and Pearson concerning testing a simple hypothesis against a simple alternative.

Show that if $s = (x_1, x_2, \ldots, x_n)$ is a sample of n independent observations from a normal population with unknown mean μ and known variance σ_0^2, and it is required to test $\mu = 0$ against $\mu > 0$, then there exists a uniformly most powerful test of size α. Give an explicit description of this test.

7. Define the term 'unbiased test'. State (without proof) sufficient conditions for the existence of a uniformly most powerful unbiased size α test.

Given the outcome of n independent Bernoulli trials each with success probability θ, describe how to construct a u.m.p. unbiased level α test of $\theta = 1/4$ against $\theta \neq 1/4$.

Examiners: R.R.Bahadur, D. Basu and V.P.Godambe.

INDIAN STATISTICAL INSTITUTE
Research and Training School

Short-term Statistician's Training Course (Senior)

Final Examination, August-September 1959

Paper III Session I.

Date: 12 September 1959 Time: 3 P.M. - 5.30 P.M.

Figures in the margin indicate full marks

1. The results of a public opinion survey in Calcutta revealed from amongst Hindu males, classified in 4 age groups the following frequencies for the different types of opinions on the question of widow marriage.

Opinions of Hindu males on widow marriage.

Opinion	Age-Groups				
	19 - 25	26 - 35	36 - 55	over 55	Total
unconditional support	49	79	69	9	206
conditional support	149	296	297	43	785
opposition	7	9	14	2	32
unconditional opposition	18	59	95	23	195
Total	223	443	475	77	1218

(a) Examine whether the age factor influences the type of opinion expressed.
(b) A sociologist who was looking through this data commented that it was needless to have 4 classes for opinion and that just two classes 'support' and 'oppose' would have shown the same amount of discrimination. Do you agree ?

[Hint: Compute the χ^2 for heterogeneity with 2 classes only and its difference from the χ^2 obtained in (a). Test the significance of this difference. (How many d.f ?)]. (30)

2. The following gives the weights of the anterior muscles of both hind legs of 10 normal rabbits.

Rabbit No.	Muscle weights (grams) Left w_1	Right w_2	Rabbit No.	Muscle weights (grams) Left w_1	Right w_2
1	5.0	4.9	6	4.0	4.0
2	4.8	5.0	7	7.1	6.9
3	4.3	4.3	8	5.9	6.3
4	5.1	5.3	9	5.3	5.2
5	4.1	4.1	10	5.3	5.5

(a) Examine if there is any difference between the two legs in respect of : (i) the average weight of the anterior muscle (ii) the variance in weight of the anterior muscle.
(b) Why is the variance ratio test not appropriate for a(ii) ?

[Hint for a(ii) Compute the correlation coefficient between $w_1 + w_2$ and $w_1 - w_2$ and test if this is significantly different from zero. How is it a test for the equality of variances]. (20)

Examiners: S. K. Mitra.

INDIAN STATISTICAL INSTITUTE
Research and Training School

Short-term Statisticians Training Course (Senior)

Final Examination, August - September, 1959.

Paper III, Section 2.

Date: 12 September 1959. Time: 5.45 P.M.-8.15 P.M.

Figures in the margin indicate full marks

3. The following table gives the means and the corrected sum of products (SP) matrix for four variates

X_1 = crude oil gravity
X_2 = crude oil vapour pressure
X_3 = crude oil 10 percent point ASTM.
X_0 = gasoline yield content

based on 32 samples of crude oil.

	SP matrix				Mean
	X_1	X_2	X_3	X_0	
X_1	984	284	-4592	461	39.25
X_2		213	-2763	334	4.18
X_3			33467	-3931	241.50
X_0				3564	19.66

(a) Obtain by the method of least squares the multiple regression equation

$$X_0 = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$$

for predicting X_0 on the basis of the other three measurements and also the value of the multiple correlation coefficient $R_{0.123}$.

(b) Test if $R_{0.123}$ is significantly different from zero.

(c) Examine if it is necessary to include X_3 in the predicting equation, when it has already been decided that X_1 and X_2 should be included.

(Hint: Test if the computed value of β_3 is significantly different from zero). (35)

4. Let y_1, y_2, y_3 and y_4 be 4 observations on random variables Y_1, Y_2, Y_3 and Y_4 respectively which are such that

$$E(Y_1) = 2\theta_1 + \theta_2 + \theta_3$$
$$E(Y_2) = \theta_1 + 2\theta_2 + 2\theta_3$$
$$E(Y_3) = \theta_1 + \theta_2 + \theta_3$$
$$E(Y_4) = 2\theta_1 + 2\theta_2 + 2\theta_3$$

and $V(Y_i) = \sigma^2$ $(i = 1, 2, 3, 4)$.

(a) Can θ_1 be estimated from y_1, y_2, y_3, y_4 ? Give reasons.

(b) Give an example of a non estimable parametric function, showing clearly that it cannot be estimated.

(c) With numerical values for the y's which were actually observed the normal equations were written down. Solving these the following particular solution was obtained

$$\hat{\theta}_1 = 2.5, \quad \hat{\theta}_2 = 1.0, \quad \hat{\theta}_3 = 0$$

with the corresponding pseudo variance covariance matrix as

$$\frac{\sigma^2}{19} \begin{bmatrix} 10 & 9 & 0 \\ 9 & 10 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

The residual sum of squares was computed as

$$R_0^2 = 1.20$$

Test the hypotheses

$$\text{(i)} \quad \theta_1 + \theta_2 + \theta_3 = 3.0$$

$$\text{(ii)} \quad \theta_1 - \theta_2 - \theta_3 = 1.0 \; . \qquad (15)$$

Examiner : S. K. Mitra.

Indian Statistical Institute
Research and Training School

Short-term Statistician's Course (Senior)

Final Examination - August-September, 1959

Paper IV

Date: 9 September, 1959 Time: 6 P.M. - 9 P.M.

Answer any four questions.

1. (a) Explain the meaning of 'Capability of a machine' or 'process'. How does the knowledge of capability help in planning production?

   A control chart maintained on a machine showed a perfect state of control. The estimates of $\bar{X}$ and $\bar{R}$ obtained from the chart are .9995'' and .001'' respectively. The specification on the characteristic is 1.000'' $\pm$ .002''. Comment (value of $d_2 = 2.326$).

   (b) Write a brief note on Quality Control on short run jobs.

2. Explain the use of the addition theorem of a number of independent variances in setting tolerances for products going into an assembly.

   Two mating components forming an assembly have their tolerances specified as $\pm$ .001'' and $\pm$ .005'' respectively.

   If the tolerance for the assembly is $\pm$ .006'' what will be your recommendation to the production department?

   A Capping machine gives a torque of 7 units with a standard deviation of 1 unit. If the breaking strength of the Caps has mean of 11 units with a s.d. of 2 units what proportion of Caps are expected to break at the Capping operation?

3. In a rolling mill, the following factors were considered to be important in an investigation made into the causes of hot rolling cracking defects. Each factor is decided to be examined at two levels. Only interaction of factors C and D given below is considered important. Prepare a suitable layout for the experimental design with minimum of experiments.

Factors	Levels
A. Composition of cast	1, 2
B. Holding time in hours	1, 2
C. Level of metal in mould (inches)	1, 2
D. Cast temperature (^{o}C)	1, 2
E. Pouring time in minutes	1, 2

4. It is required to instal control over the machine utilisation in a machine-shop. How would you plan collection of information? Explain with the help of suitable forms the relevant data to be collected and the type of analysis that you would make. What precautions or any measures would you take to ensure validity of the data?

5. State Dodge's sampling plan for continuous production and derive the values of the components in the AOQ formula. State how plans are provided for practical use.

6. Derive the Sequential Probability Ratio Test criteria from the acceptance sampling point of view. Illustrate the test in the Binomial case for testing $p = p_1$ against $p = p_2$.

7. Write notes on: i) OC of control charts, ii) narrow limit gauging, iii) group control charts, iv) control charts for individuals, v) inspection levels.

8. Compare Attribute inspection with measurement inspection.
State and derive the elements in a single sampling variables plan for use when lot standard deviation is not known and when item quality is specified by a lower one-sided limit.

.....

Examiners: A. Matthai and S.Q.C. Unit

Indian Statistical Institute
Research and Training School
Short-term Statistician's Training Course (Senior)
Final Examination - August-September, 1959

Paper V

Full marks : 100

Date: 10 September 1959 Time: 6 P.M. - 9 P.M.

All questions carry equal marks

1. What are sampling and non-sampling errors? What precautions do you suggest to control the non-sampling errors?

2. (a) What are pilot surveys and how are they useful?

   (b) What are the different methods of data collection? What are their advantages and disadvantages?

3. State whether the following statements are true or false and give reasons for your answers.

   (i) Systematic sampling is always more efficient than simple random sampling.

   (ii) Cluster sampling is always less efficient than simple random sampling.

   (iii) The greater the number of strata, the less will be the variance of the estimator.

   (iv) Two stage sampling is always less efficient than uni-stage sampling.

   (v) Des Raj's estimator for a pps sample drawn without replacement is more efficient than the usual estimator for a pps sample drawn with replacement.

   (vi) Stratified proportional allocation sampling with equal probability with replacement is more efficient than unstratified srs with replacement.

   (vii) Des Raj's estimator in case of pps sampling without replacement cannot be improved upon.

   (viii) It is impossible to obtain the bias in a ratio estimator.

   (ix) In srs with replacement, the sample mean is the minimum variance unbiased estimator.

   (x) Lahiri's method of selecting a sample with pps with replacement is more economical than the cumulative total method.

4. Either

   For estimating the area under paddy in Madras State, a stratified sampling design is suggested where the sampling units (plots) are to be selected from each stratum with probability proportional to area and with replacement.

   (i) Obtain an unbiased estimator of area under paddy and an unbiased estimator of variance for your estimator.

   (ii) Under what conditions, if any, is the allocation of the sample plots to the strata proportional to $c\sqrt{p_s q_s}$ the optimum? (g is the geographical area, p_s is the proportion of area under paddy, $q = 1-p$ and the subscript 's' stands for the sth stratum.)

P.T.O

4. Or

Suppose there are 2 strata and N_1 and N_2 are the number of units in them. Let x_{ij} be the size of u_{ij}, the jth unit in the ith stratum. One unit (say u_{i_1}) is selected with probability proportional to its size from the whole population of N_1+N_2 units. One unit from the remaining (N_i-1) units in the ith stratum and 2 units from the other stratum are selected with equal probability without replacement.

(i) Find the probability of selecting the sample

$$u_{1i}, u_{ij}, u_{2k}, u_{21}$$

(ii) Show that the estimator of the population total y

$$\hat{y} = \frac{N_1\bar{y}_1 + N_2\bar{y}_2}{N_1\bar{x}_1 + N_2\bar{x}_2} X$$

is unbiased. (Where X is the sum of the sizes of all the units and $\bar{y}_1$ and $\bar{x}_1$ are the sample means in ith stratum for the character in question and the size respectively.)

(iii) Obtain an unbiased estimator of the variance of $\hat{y}$.

5. (i) What is a self-weighting design?

(ii) What are its advantages and disadvantages?

(iii) Indicate how to make a two stage sampling design where the first stage units are to be selected with pps with replacement and the second stage units are to be selected systematically self-weighting such that the expected number of second stage units to be selected per sample first stage unit is a given number (say n).

Examiner: S. Raja Rao

INDIAN STATISTICAL INSTITUTE
Research and Training School

Short-term Statistician's Training Course (Senior)

Final Examination, August-Sep.1959

Paper VI. Session 1.

Date: 29 August 1959 Time: 3.30 P.M.-6 P.M.

1. The following data are obtained from an experiment conducted to test the alkali resistance number of wrapping paper for packing a detergent. Four Chemists (C) from two laboratories (T) tested samples from each of 6 randomly selected sheets (S) from a lot. The results were recorded as deviations from the standard. It is desirable that the resistance number should not exceed the standard value of 10. The data are given below:

Table

		C_1	C_2	C_3	C_4	Total
S_1	T_1	-2	3	2	2	5
	T_2	2	2	1	4	9
S_2	T_1	-2	2	1	-1	0
	T_2	4	1	-2	-2	-7
S_3	T_1	0	1	1	-2	0
	T_2	0	1	0	4	5
S_4	T_1	0	2	0	-1	1
	T_2	1	0	0	2	3
S_5	T_1	2	1	-2	5	6
	T_2	3	1	3	4	11
S_6	T_1	4	1	4	-4	5
	T_2	4	1	-1	5	9

Analyse the data and estimate the variation in the alkali resistance number of the lot after eliminating the various errors that may be introduced in testing.

2. The following data were obtained on an automatic lathe producing an outside diameter without any change whatsoever in the setting. B indicates change of raw material (bar). Thus samples 1 to 3 are from the same raw material whereas : 4 and 5 are from another and so on.

Estimate the contribution of different factors to the overall variance. If it is known that it is not possible to secure better raw material and an overall average of 40 units is satisfactory prepare a suitable scheme for control of this dimension.

P.T.O.

	sample no.	observations				
		1	2	3	4	5
	1	42	42	42	43	44
	2	40	41	39	39	42
	3	40	39	40	42	41
B	4	43	42	43	42	42
	5	45	45	45	44	45
B	6	40	41	41	42	40
	7	39	40	40	40	39
	8	39	40	39	41	39
B	9	39	39	40	39	41
	10	40	40	40	42	40
B	11	44	45	43	44	44
	12	44	46	46	44	44
	13	43	41	44	40	41
B	14	41	39	41	37	38
	15	38	38	36	37	35
B	16	40	42	39	40	40
B	17	42	42	43	44	41
	18	41	43	41	41	42
B	19	42	41	39	41	39
	20	38	39	40	40	40

Examiner : SQC Unit.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Training Course (Senior)
Final Examination, August-Sept.1959
Paper VI. Session 2

Date: 29 August 1959 Time: 6.15 P.M.-8.45 P.M.

1. Set up a Tabular Scheme for sequential inspection for number of defects so that lots containing an average of 2 defects per item will be accepted 95 percent of the times and those containing an average of 5 defects only 10 percent of the times.

   Plot the 5 point OC and ASN of the plan.

2. Using Probability Tables choose the elements in the single sampling attributes plan which will accept 95 percent of the times lots containing 1 percent defectives and 10 percent of the times those with 4 percent defectives.

3. Set up 95 percent confidence intervals for the following:

   i) lot standard deviation, given in a sample of 10 observations the range was 6.253.

   ii) lot mean, given the sample mean based on 12 observations was 3.253 and the sample range 0.623.

   iii) lot percentage defective, given that in a sample of 300 items 21 were defectives.

Examiner: A. Matthai.

INDIAN STATISTICAL INSTITUTE
Research and Training School

Short-term Statistician's Training Course (Senior)

Final Examination August-Sept. 1959

Paper VII. Session 1.

Date: 23 August 1959. Full marks 50 Time: 11 A.M.-2 P.M.

Note: Figures in the margin indicate full marks.

1. In order to estimate the average age of the scientists, 20 pages were selected at random from the volume 'American men of Sciences'. The number of biographies M_i per page varies in general from about 14 to 22. On each selected page, 2 biographies were chosen at random and the age of the scientist was recorded. The data recorded is given below.

Total number of pages = $N = 2823$

$$\sum_{i=1}^{N} M_i = 50000$$

unit number	M_i	Ages		unit number	M_i	Ages	
		y_{i1}	y_{i2}			y_{i1}	y_{i2}
1	15	47	30	11	20	66	59
2	19	38	51	12	20	37	49
3	19	43	45	13	19	60	53
4	16	50	44	14	18	45	83
5	16	55	40	15	18	30	37
6	19	39	36	16	16	60	35
7	18	40	40	17	15	60	44
8	18	44	46	18	19	36	33
9	18	40	30	19	19	60	38
10	18	46	59	20	19	50	30

(i) Estimate the average age of scientists and also its sampling error.

(ii) Assuming the cost (in rupees) of the survey (T) is represented by

$$T = n + 2nm$$

where n = number of pages selected and m = number of biographies selected per page,

find the optimum values of n and m that may be used for the subsequent surveys in order that the average age of the scientists may be estimated with standard error of 5 years at minimum cost. What is the minimum cost ? [15]

Note: You may estimate the necessary components of variance function from the sample data given above.

The following table shows the geographical area and cultivated area for 40 sample villages in a region of India divided into 4 strata. The sample was obtained by selecting systematically two independent sub-samples of 6 villages each with equal probability within each stratum.

Calculate the total cultivated area for the state using the Geographical area as the auxiliary variable and by a suitable ratio estimator. Estimate the variance of this estimate. [20]

P.T.O.

Geographical area and cultivated area of sample villages in a state.

sub-sample	sample village no.	geographical area (sq. miles) (x)	cultivated area (acres) (y)	sub-sample	sample village no	geographical area (sq. miles) (x)	cultivated area (acres) (y)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Stratum 1; total villages = 2034 geographical area = 2955 sq.miles				Stratum 2; total villages = 1250 geographical area = 2900 sq.miles			
1	1	1.24	520	1	1	0.80	77
	2	5.31	1[illegible]		2	1.30	740
	3	5.01	1230		3	1.38	1380
	4	0.32	55		4	2 [illegible]	600
	5	1.25	475		5	2.50	605
	6	1.40	233		6	0.17	79
2	1	3.20	1160	2	1	4.30	9
	2	0.60	175		2	5.52	1410
	3	2.86	1070		3	2.45	1510
	4	1.36	24		4	4.82	925
	5	1.02	400		5	5.90	817
	6	1.01	90		6	0.07	20
Stratum 3; total villages = 1264 geographical area = 2430 sq.miles				Stratum 4; total villages = 1250 geographical area = 5142 sq.miles			
1	1	2.05	1015	1	1	2.14	380
	2	0.60	165		2	0.12	8
	3	3.90	617		3	0.40	62
	4	0.30	176		4	3.72	2008
	5	0.72	426		5	1.27	52
	6	1.33	622		6	0.57	109
2	1	0.70	317	2	1	0.35	106
	2	1.11	34		2	0.57	90
	3	1.58	750		3	0.36	60
	4	4.10	1629		4	0.10	24
	5	4.33	1556		5	0.55	192
	6	1.24	537		6	0.21	3165

3. Prepare a first year's budget estimate for the cost of field work on crop cutting experiments for 10 Rabi crops of a State. The following points should be taken into consideration in preparing the budget.

(i) One investigator can take 2 cuts per day;

(ii) There are 240 police stations (administrative units) in the State;

(iii) In each police station, 30 sample cuts should be taken for each of the ten crops;

(iv) Labour charges will have to be paid at the rate of Re.1/- per cut;

(v) There should be provision for supervisory staff, printing of forms and stationery etc.

(vi) You may assume appropriate [illegible]tional estimates of pay allowances etc., that has to be paid to the staff. Mention these figures clearly.

(vii) Mention other assumptions which you may make, if any. [15]

Examiners: S. Raja Rao and M. N. Murty.

INDIAN STATISTICAL INSTITUTE
Research and Training School

Short-term Statistician's Training Course (Senior)

Final Examination: August-Sept. 1959
Paper VII. Session 2.

Date: 23 August 1959. Full marks 50 Time: 2.30 P.M.-5.30 P.M.

Note: Figures in the margin indicate full marks.

1. Enclosed is a list of sub-divisions in West Bengal along with their population figures for 1951 (table A):

Point out the defects in this frame. [10]

2. In a sample survey, the investigators are asked to visit a number of sample villages, take a sample of households according to a rule of procedure and carry out some Socio-economic Investigations. Write a self contained instruction for drawing the sample of households in a sample village making use of the following information.

(i) The sample villages are the villages as defined for the purpose of 1951 census.

(ii) A household is defined as "A group of persons normally living together and ordinarily taking food from the same kitchen".

(iii) The households of the villages are to be arranged in such a manner that all households of sizes less than or equal to 4 are to be given the sampling serial number first and the others should follow later.

(iv) A sample of 8 households are to be selected in a systematic manner from this list after rearrangement. A random number table will be supplied to the Investigator. You should also explain the action to be taken when the total number of households is less than 8.

(v) The enclosed schedule (0.1) has been prepared for this purpose. [20]

3. A sample survey of land holdings was conducted in India and some characteristics of operational holdings are given in table (1) enclosed. Present the data in a suitable graph and write your technical comments and inference from the table. Also give your suggestions, if any, for improving the presentation of table (with reasons).

The following basic concepts are given to facilitate easy understanding of the tables.

(i) An operational holding consists of all land put to agricultural as well as to non-agricultural uses under the management of one district technical and economic unit. An operational holding may be managed either by an individual unit or jointly by more than one unit.

(ii) The estimates for each size class are based on the total number of sample villages and total number of households (the 1st and the 2nd stage units in a two stage sampling design). For the households which do not belong to a particular size class, the value of the character estimated is taken as zero. [20]

Examiners: S. Raja Rao and M.N.Murty.

Table (1) Cumulative percentage of estimated number of operational holdings and of area operated by size level of operational holdings.

area of operational holdings (acres)	sub-sample 1		sub-sample 2		both samples combined	
	cumulative p.c. of operational holdings	cumulative p.c. of area operated	cumulative p.c. of operational holdings	cumulative p.c. of area operated	cumulative p.c. of operational holdings	cumulative p.c. of Area operated
(1)	(2)	(3)	(4)	(5)	(6)	(7)
0.00(1)	6.02	-	13.51	-	10.96	-
0.01-0.99	41.19	1.20	42.58	1.20	42.08	1.20
- 2.49	55.24	5.57	56.61	5.57	56.15	5.58
- 4.99	70.79	15.94	71.39	15.40	71.23	15.60
-7.49	79.97	26.21	79.93	25.06	79.99	25.48
-9.99	85.86	35.07	85.29	35.65	85.42	34.16
-14.99	90.97	47.24	90.97	46.36	90.99	46.64
-19.99	94.20	57.21	94.03	56.08	94.11	56.55
-24.99	95.90	64.44	95.69	62.86	95.78	63.38
-29.99	97.05	70.22	96.85	68.66	96.93	69.17
-49.99	98.88	83.07	98.38	82.85	98.90	82.92
50 and above	100.00	100.00	100.00	100.00	100.00	100.00
actual estimates	63876000	348151000 acres	60916000	330342000 acres	617800000	335711000 acres
average size of operational holdings	5.45 (5.80)		5.42 (6.26)		5.43 (6.10)	
number of sample villages	1410		3021		4431	
number of sample households	34366		51354		75720	

(1) includes operational holdings of size less than 0.005 acres.

2. Figures in brackets are obtained by omitting holding of size less than 0.005 acres.

Table A: List of Tehsil in West Bengal

serial no.	district	sub-divisions	population in 1951 (0000)
1	Burdwan	Sadar	80
2		Kalna	30
3		Katwa	31
4		Katwa	31
5		Asansol	76
6	Bankura	Sadar	97
7		Vishnupur	
8	Midnapore	Sadar	106
9		Contai	74
10		Tamluk	79
11		Ghatal	31
12		Jhargram	46
13	Hooghly	Sadar	45
14		Serampur	73
15		Arambag	37
16	Howrah	Sadar	93
17		Uluberia	
18	Birbhum	Sadar	64
19		Rampurhat	43
20	24-Parganas	Sadar	151
21		Basirhat	71
22		Basirhat	71
23		Barasat	31
24		Bangaon	21
25		Barrackpore	88
26		Diamond Harbour	90
27	Nadia	Sadar	70
28		Ranaghat	44
29	Murshidabad	Sadar	54
30		Lalbagh	39
31		Jangipur	43
32		Kandi	
33	Malda	Sadar	93
34	West Dinajpur	Balurghat	33
35		Raiganj	39
36		Balurghat	33
37	Jalpaiguri	Sadar	54
38		Alipur Duars	37
39	Darjeeling	Sadar	17
40		Siliguri	12
41	Coooh Bihar	Sadar	17
42		Tufanganj	10
43		Dinhata	16
44		Mathabhanga	15
45		Mekliganj	9

Schedule 0.1 : Listing of households

[1] identification of sample village

1. serial number of sample village .....
2. zone-state-natural division .....
3. stratum number .....
4. state .....
5. district .....
6. tehsil .....
7. village .....

[2] certain particulars of sampling

1. number of households in village .....
2. sampling interval .. ..
3. random start .....
4. random pair .....
5. random column .....
6. random row .. ..

[3] list of households and record of selection

house number	household sl.no.	name of head of household	name of father/mother/husband of head of household	relation to head (a)	household size	sampling serial number for household size		order of selection
						1-4	above 4	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)

code number: (a) relation to head : father-1; mother-2; husband-3.

INDIAN STATISTICAL INSTITUTE
Research and Training School

Short-term Statistician's Training Course (Senior)

Periodical Examination I.
Economic Statistics

Date: 17 December 1959 Time: 6 P.M.-9 P.M.

1.(a) Construct a simple numerical example to show how national income can be conceived as an aggregate of (i) income payments (ii) values added and (iii) final expenditures.

(b) Construct a simple production account of an enterprise and a consumption account for household with arbitrary figures.

What is meant by an articulated system of accounts ?

2. Distinguish between the concepts (i) net national income and gross national income (ii) national income at market prices and national income at factor cost (iii) national income and domestic product and (iv) national income and personal income before tax.

or

Describe briefly the method of estimation of national income in India, giving details about any two industrial sectors.

3. Give a brief account of the procedure of construction of inter-industry tables with explicit reference to the two methods of taking account of foreign trade.

or

Construct a simple inter-industry example to illustrate the method of obtaining imports and domestic production requirements, given capacity limitations and targets of consumption in different sectors of the economy.

4. 'Increase in consumption in the long run is attainable at the cost of consumption at present or in the immediate future by means of larger investment in the producers' goods sector'. Explain the statement on the basis of an appropriate planning model.

or

Show how the rate of increase in gross output can be expressed as the product of the proportion of gross output invested and the overall output-outlay ratio for the economy.

5. Explain what is meant by price elasticity and income elasticity of demand. Give an account of the data needed and the methods used for estimating these elasticities. Describe some of their uses.

or

Describe how income elasticity of demand can be estimated from family budget data. Mention its important uses.

Examiners:- A.K.Biswas, M. Mukherjee, N.S.Iyengar.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Training Course (Senior)
Periodical Examination II.

Algebra and Estimation

Date: 24 December 1959. Time: 6 P.M.-9 P.M.

1.(a) Examine the four vectors

$\alpha_1 = (2, 1, 0, 5)$ $\alpha_2 = (3, 6, 2, 1)$

$\alpha_3 = (1, -10, -1, 3)$ $\alpha_4 = (5, 1, 3, 0)$

for linear dependence

(b) What is the rank of the vector space $V(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ generated by $\alpha_1, \alpha_2, \alpha_3$ and α_4.

(c) Obtain a Normal Orthogonal Basis of this vector space $V(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$.

(d) Examine if the vector $\beta = (10, 0, 4, 9)$ belongs to $V(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ or not.

(e) If not express β as $\beta_0 + \beta_1$ such that β_0 is in $V(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ and β_1 is orthogonal to $V(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$.

2. Consider the matrix

$$A = \begin{bmatrix} 2 & 3 & 1 \\ 3 & 6 & 2 \\ 1 & 2 & 5 \end{bmatrix}$$

(a) Compute A^{-1} and also solve the equations

$$\underline{x} \; A = \underline{C}$$

where $\underline{C} = (2, 1, 6)$

(b) Compute $|A|$ (the determinant of A)

(c) Compute $\underline{C} \, A^{-1} \, \underline{C}'$

(d) Obtain a transformation $\underline{y} = \underline{x} \, B$ under which the quadratic form $\underline{x} \, A^{-1} \, \underline{x}'$ is reduced to sum of squares of the y's (that is $\underline{y} \, \underline{y}'$)

(e) If A is the variance covariance matrix of random variables X_1, X_2 and X_3 and $\underline{Y} = \underline{X} \, B$ (same B as in 2.(d)) what would be the variance covariance covariance matrix of Y_1, Y_2 and Y_3?

P.T.O.

3.(a) Explain the terms (i) unbiased estimate (ii) uniformly minimum variance unbiased estimate, as they occur in the theory of estimation.

(b) If U is the uniformly minimum variance unbiased estimate of a parametric function as computed from certain sample observations and Z is another function of the sample observations with identically zero expectation and finite variance, show that U is necessarily uncorrelated with Z. State the converse of this proposition. Is the converse also true ?

Examiner: S. K. Mitra.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Training Class (Senior)
Periodical Examination III
Statistical Quality Control.

Date: 31 December, 1959. Time: 6 p.m. - 9 p.m.

Answer any two questions from Group A and any one from Group B.

GROUP A

1. Classify lot acceptance sampling plans by type of stipulation. Give a brief account of each type of stipulation.

2. Mention the different steps in Dodge's plan for sampling from a continuous production stream. Derive the terms in the AOQ of the plan.

3. Derive the Sequential Probability Ratio Test from the acceptance sampling point of view. Show how the problem of testing difference between two proportions can be reduced to a sequential test for a single binomial proportion.

4. Write notes on:(i) Confidence intervals; (ii) tolerance limits; (iii) Control charts for individuals; (iv) ASN.

GROUP B

1.(a) The means and ranges in 5 batches of items inspected are given below:

Obtain values for the Central line and Control limits for setting up control charts : (i) for the mean; (ii) for the range.

batch	no. of items	mean	range
1	5	3.28	1.1
2	4	3.33	1.3
3	4	3.61	1.7
4	5	3.40	1.2
5	5	2.98	1.8

(b) By plotting at least six suitable points draw the OC curve of the single sampling attributes plan having for its elements $n = 30$ and $c = 1$ (Tables of Binomial or other probability need not be used).

2. Set up a tabular procedure for a sequential defects-per unit plan by which lots having an average number of 1 defect per unit will be accepted 95% of the times and lots with an average of 2.5 defects per unit will be accepted only 10% of the times.

Plot the 5-point OC and ASN of the plan.

Examiner: A. Matthai.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Course. Senior
Final Examination. February, 1960
Paper I

Date: 16 February, 1960. Time: 6 p.m. - 9 p.m.

Answer the two groups in separate books

GROUP A

1. Under the usual linear estimation set up (which you are required to state) define the terms: estimable linear parametric functions and best linear unbiased estimate (BLUE)

(a) Show that the best linear estimate is necessarily uncorrelated with any linear function of the observations which has identically a zero expectation and that the converse of this result is also true.

(b) Hence or otherwise show that any solution of the normal equations when substituted for the unknown parameters in a linear parametric function L always leads uniquely to the BLUE for L if L is estimable.

(c) Derive the variance of this BLUE and show how it can be estimated from the observations.

OR

$x_1, x_2, \ldots x_m$ and $y_1, y_2, \ldots y_n$ are two samples from two normal populations $N(\mu_x, \sigma_x)$ and $N(\mu_y, \sigma_y)$. Using the theory of least squares derive a test of the hypothesis $\mu_x - \mu_y = 2$ when it is known that

$$\sigma_x^2 = 9\,\sigma_y^2$$

2. Consider the following equations

$$\begin{aligned} x_1 + 2x_2 + x_3 + 2x_4 &= 0 \\ x_1 \quad\quad + x_3 \quad\quad &= 0 \\ x_1 + x_2 + x_3 + x_4 &= 0 \end{aligned}$$

(a) Show that the collection C of all solutions (x_1, x_2, x_3, x_4) of these equations is a subspace of the vector space $V_4(R)$ of all 4-tuples.

(b) Is (-1, 5, 1, -5) a member of C ?

(c) What is the dimension of C ?

(d) Suggest any 4-tuple (a_1, a_2, a_3, a_4) which is orthogonal to every vector C.

(e) Obtain a basis of C and a normal orthogonal basis.

OR

(a) If Σ is the variance covariance matrix of random variables (X_1, X_2, X_3) and B is a square matrix (3 x 3), what would be the variance covariance matrix Λ of (Y_1, Y_2, Y_3) when $\underline{Y}' = B\,\underline{X}'$?

(b) Show that if both Σ and Λ are non singular

$$\underline{X}\,\Sigma^{-1}\,\underline{X}' = \underline{Y}\,\Lambda^{-1}\,\underline{Y}' \quad .$$

(You may use the fact the inverse of a product of non singular matrices is a product of their inverses taken in the reverse order).

p.t.o.

(c) Let

$$\Sigma = \begin{bmatrix} 2 & 0 & 6 \\ 0 & 2 & 4 \\ 6 & 4 & 20 \end{bmatrix}$$

Obtain a matrix B, such that the corresponding Y's, defined by $\underline{Y}' = B\underline{X}'$, are mutually uncorrelated and have individually unit variance (that is, $\Lambda = I$).

(d) If the distribution of $\underline{X}$ is multivariate normal and the X's have zero mean values what would be the distribution (i) of $\underline{Y}$ as obtained in (c) (ii) of $\underline{Y}\,\underline{Y}'$ (iii) of $\underline{X}\,\Sigma^{-1}\,\underline{X}'$.

GROUP B

3. Describe a Balanced Incomplete Block Design and the randomisation procedure used in such a design.

Explain how the intra-block analysis is carried out for this type of designs.

OR

What is meant by confounding in a factorial experiment ? Draw up a balanced partially confounded design for a factorial experiment involving three factors each at two levels, in blocks of four plots in four replications.

Examiners: S. K. Mitra & J. Roy

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Course. Senior
Final Examination. February, 1960

Paper II

Date: 18 February, 1960 Time: 6 P.M.-9 P.M.

Answer the two groups separate books

GROUP A

(Suggested time : 2 hours)

1. Find the probability distribution of $Y = X_1 + X_2$ if X_1 and X_2 are independent and identically distributed variable in the following cases: (a) X_1 is uniformly distributed over the interval (0, 1); (b) X_1 is normally distributed with mean zero and variance 1.

2. Let $X_1, X_2, \ldots, X_n$ be a sample of independent observations from a normal population with mean μ and variance σ^2. Let $\bar{X}_n = \sum_1^n X_i/n$ and $S^2 = \sum_1^n (X_i - \bar{X})^2$. Show that $\bar{X}_n$ and S^2 are independent random variables.

State without proof the density functions of sampling distributions of (i) $T = \sqrt{n}\,(\bar{X}_n - \mu)/\sqrt{S^2/n-1}$ and (ii) $U = S^2/\sigma^2$.

3. Define the terms 'unbiased estimate', and 'sufficient statistic', and 'efficient estimate'.

Show that corresponding to any given unbiased estimate there exists an unbiased estimate which is a function only of the sufficient statistic and such that the variance of the latter estimate never exceeds that of the given estimate. Explain the utility of this result by means of an example.

4. Let X be a variable with a uniform distribution over the interval $(0, \theta)$ where θ is an unknown parameter, $0 < \theta < \infty$, i.e. the probability density of X is $f(x, \theta) = 1/\theta$ if $0 < x < \theta$ and $f(x,\theta)=0$ otherwise. Let $S = (X_1, X_2, \ldots, X_n)$ be n independent observations on X.

Given a sample S, (a) write down the likelihood function, and (b) find the maximum likelihood estimate of θ, say $\hat{\theta}(S)$. (c) Find the probability distribution of $\hat{\theta}$.

5. State and prove the Cramér-Rao inequality concerning the variance of an unbiased estimate.

Give examples where the lower bound to the variance is (a) attained, and (b) not attained. What can be said in cases where the bound is not attained ?

6. State and prove the Neyman-Pearson lemma in the theory of testing hypotheses.

Let $S = (X_1, X_2, \ldots, X_n)$ be a sample form a normal population with unknown mean μ and known variance σ^2. It is desired to test $H : \mu = 0$ against the alternative that $\mu > 0$. It is required that the probability of an error of type one be ε, where ε is a given constant. Show that there exists a u.m.p. test, and describe this test in convenient form. Is the test unbiased ?

P.T.O.

GROUP B

(Suggested time : 1 hour)

7. Explain any two of the following sampling schemes and mention their advantages and disadvantages.

(i) Sampling of units with varying probabilities.
(ii) Cluster sampling.
(iii) Multistage sampling.

8. Derive the condition under which sampling clusters of ultimate units (cluster sampling) will be more efficient than simple random sampling of ultimate units.

OR

What are non-sampling errors ? How do they affect the estimates from sample surveys ? What are the possible sources of such errors ?

Examiners: R.R. Bahadur and S.Raja Rao.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Course. Senior.
Final Examination, February, 1960
Paper III.Session 1.

Date: 27 February 1960 Time: 3 P.M.-5.30 P.M.

1. The density function of a certain bivariate population is given to be

$$K.\ e^{-(x^2 + 5y^2 - 4xy + 2x + 15y)} \qquad (-\infty < x, y < \infty)$$

Obtain the following:

(i) V(x); (ii) V(y); (iii) Cov (x, y);

(iv) The regression line of y on x.

2. Data regarding the heights of 35 individuals at the ages of $5(X_1)$, $10(X_2)$ and $25(X_3)$ are summarized below:

$$\bar{x}_1 = 42.4",\ \bar{x}_2 = 51.7",\ \bar{x}_3 = 66.6" .$$

Estimated variances and covariances (in sq. inches)

X_1	X_2	X_3
20.9	8.4	11.1
	43.6	5.8
		20.2

Examine, in the light of these data, the hypothesis that on the average people grow by 10" between the ages 5 and 10 and by 15" between the ages 10 and 25.

OR

What is meant by an error function (zero function)? Let y_1, y_2, y_3 be three independent random variables having the same variance σ^2, but expectation given by

$$E(y_1) = t_1 - 2t_2 + 3t_3$$

$$E(y_2) = 3t_2 - 2t_3$$

$$E(y_3) = 2t_1 - t_2 + 4t_3$$

where t_1, t_2 and t_3 are three unknown parameters.

(a) How many linearly independent error functions are there in this model.

(b) Obtain an error function under this model.

(c) Without obtaining the best linear unbiassed estimator of $3t_1 + 5t_3$ how will you conclude that $\frac{1}{3}(y_2 + 3y_3)$ is not the best linear unbiassed estimator of $3t_1 + 5t_3$?

(d) Calculate the efficiency of $\frac{1}{3}(y_2 + 3y_3)$ as an estimator of $3t_1 + 5t_3$ relative to the BLUE ?

P.T.O.

3. In an experiment exploring the possibility of estimating the gasolene yield from certain crude oil properties, the following observations (suitably coded) were made on 32 samples of crude oil.

X_1 : crude oil gravity
X_2 : gasolene end pt.
X_3 : crude oil pour pressure
X_4 : crude oil 10 p.c. point ASTM
Y : gasolene yield percent

The following figures computed from the sample are available.

	Total	Sum of products (uncorrected)					
		X_1	X_2	X_3	X_4	Y	
X_1	12.560	5.[illegible]33	4.1319	5.5357	2.9873	2.5154	
X_2	10.627		3.6800	4.2746	2.5999	2.2542	$\Sigma(Y-\bar{Y})^2=0.3564$
X_3	13.380			7.?222	2.9550	2.9649	
X_4	7.728				1.9100	1.4800	
Y	6.291						

(a) Obtain the best linear function of (i) X_1 and X_2 (ii) X_1, X_2, X_3, X_4 for predicting Y.

(b) Test whether the second prediction formula is a real improvement over the first (i.e. Test whether the drop in the residual sum of squares could not be attributed to chance.)

(c) Test whether the following observations on a sample of oil taken a few months later are in agreement with the predicting equation (i)

$$x_1 = 0.400, \qquad x_2 = 0.365, \qquad y = 0.182$$

Examiners: S.K.Mitra and S.John.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Course. Senior
Final Examination, February, 1960
Paper III. Session 2.

Date: 27 February 1960 Time: 6 P.M.-8.30 P.M.

Figures in the margin indicate full marks

1. The layout and the yields in a varietal trial involving 7 varieties in a Balanced Incomplete Block Design are given below.

Analyse the data and write a report.

A	C	G	D
60	72	78	51
B	F	C	G
26	35	67	83
F	B	A	D
39	24	53	41
E	G	F	A
18	80	36	59
C	F	D	E
65	39	44	18
B	E	A	C
30	20	58	74
D	G	E	B
48	72	16	30

[25]

2. Explain the operational procedure of drawing a sample of n units with replacement with probability proportional to a given measure of size by Lahiri's method from a population of N units. [7]

3. You are given a list of 30 villages and their population. Draw a sample of 5 villages with replacement with probability proportional to population either by Lahiri's method or by cumulative method. Give suitable reference to random numbers used.

sl.no.of village	popula-tion	sl.no.of village	popula-tion	sl.no.of village	popula-tion
1	165	11	199	21	70
2	226	12	1199	22	87
3	55	13	876	23	261
4	75	14	240	24	3700
5	370	15	1152	25	556
6	79	16	200	26	131
7	350	17	108	27	113
8	1604	18	261	28	127
9	554	19	182	29	146
10	111	20	556	30	19

[8]

4. The yield of green weight of Jute plants on the corresponding weights of dry fibre obtained from five random sample cuts of 100 sq.ft. each (drawn with replacement) are given below.

green weight (tolas) x	86	37	57	43	81
dry weight (tolas) y	57	24	31	22	46

The total green weight of jute plants from the entire region under study was 2149 tolas. Estimate the yield of dry fibre from the entire region and an estimate of its error by ratio method of estimation. [10]

Examiners: J. Roy and S.Raja Rao.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Course. Senior.
Final Examination. February, 1960

Paper IV.

Date: 23 February, 1960 Time: 6 P.M.-9 P.M.

Answer the two groups in separate books

GROUP A

Attempt any two

1. State Dodge's sampling plan for continuous production.

Derive the terms in the AOQ of the plan. Indicate how the plans are presented for use in industry.

2. Enumerate the different important acceptance sampling plans available in published form briefly indicating, the type of inspection, the type of stipulations behind the plan, the situation in which the plan is mostly applicable, in each case.

3. Describe the circumstances in which control charts for individuals are set up. State also how such charts are set up under the two situations (i) standards given, (ii) standards not given.

GROUP B

Attempt any two

4. Explain the uses of range in Quality Control work.

The weight of a cricket ball is required to lie within 5.50 ± 0.25 ozs. What is the maximum value of the mean range in samples of size of 4 with which it would be possible to ensure that at least 99 p.c. of the balls would conform to specifications - state the assumptions.

Following data gives the volume of fuel in c.c. delivered through three types of nozzles in a given time.

Nozzle type		
A	B	C
96.6	96.6	97.0
97.2	96.4	96.0
96.4	97.0	95.0
97.4	96.2	95.8
97.8	96.8	97.0

Examine by the use of ranges whether the nozzle types are significantly different.

5. Explain the use of the addition theorem of variances in setting tolerances for products going into an assembly.

It is known that pieces of 10 yds. of sliver are distributed with a mean of 100 and standard deviation of 2 in some units. Assuming that weights of consecutive pieces of 10 yds. length of sliver are distributed independently of each other derive the range of variation for pieces of sliver of length 100 yds.

Two mating components form an assembly which has tolerances of ± .012". Determine the tolerances within which the components are to be made assuming equal tolerances for both.

P.T.O.

6. It is known the efficiency obtained on a loom depends on the following factors:

A	Batch of jute
B	Oil content of yarn
C	Twist of yarn
D	Sizing medium used
E	Speed of the loom.

Only interaction of factors A and B is considered important. It is desired to conduct an experiment with 3 levels on each of these factors. Prepare a suitable layout for the experimental design.

Indicate your reasoning for the choice of the design suggested and discuss the analyses you will carry out to examine whether the experiment has been carried out under controlled conditions.

Examiners: A. Matthai and S.Q.C. Unit.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Course. Senior.
Final Examination. February, 1960
Paper V.

Date: 25 February, 1960 Time: 6 p.m.-9 p.m.

1.(a) Indicate the adjustments necessary to obtain personal income after tax from net national output at factor cost. Is personal income tax a part of the net contribution to national income ?

(b) Distinguish between net domestic product and net national product.

OR

Discuss different methods of estimation of national income and indicate briefly the procedure followed in India to obtain the contribution to national income by various sectors.

2. State clearly the Pareto law of income distribution and obtain the mean and the coefficient of variation of the distribution. Derive the equation of the Lorenz curve and hence obtain an expression for the area of concentration.

OR

In deriving relationships between expenditures on various items and income (total expenditure) from family budgets, what procedure will you adopt to eliminate the effect of the variation in family composition ? Describe a method for estimating adult equivalent scale from family budget data.

3. Show that in a two-country model, a rise in investment in one country increases the national income of both the countries and makes the balance of trade unfavourable to the country which initiates the increase in investment expenditure.

OR

Set up appropriate national income model to examine the comparative efficacy of the following types of fiscal policy in improving national income:

(i) increase in government expenditure matched by an increase in taxes;

(ii) reduction in taxes, government expenditure remaining constant.

4. Reconcile the following statements:

(i) saving and investment are equal by definition;

(ii) saving is brought into equality with investment by appropriate changes in national income.

Construct an arithmetic example to explain the nature of the two propositions.

OR

What is the relation between increase in investment and marginal propensity to save ? Does this relation apply to underdeveloped countries ? Explain.

5. Develop a suitable procedure for comparing the employment potential of different sectors of an economy.

OR

You are required to estimate the consumption expenditure on different items in India in 1960, assuming income will rise by 5 p.c. during the period, 1955-1960. What type of data do you require for this purpose ? Explain how you are going to obtain the estimates if the required data are supplied. Indicate some of the limitations of your estimates.

Examiners: A.K.Biswas, M.Mukherjee and N.S.Iyengar.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Course. Senior
Final Examination. February, 1960
Paper VI. Session 1.

Date: 14 February, 1960. Time: 11 a.m.-2 p.m.

1. Set a tabular procedure for sequential attribute inspection by which lots having 3 percent defectives will be accepted only 10 percent of the times and lots with 1 percent defectives will be accepted 95 percent of the times.

Plot the OC and ASN Curves for the plan.

2. (a) Sketch the procedure, giving necessary formulae, for setting up a control chart for range when standards are not given and when samples which are small are all of not the same size.

(b) Find out the values of the elements n and k of a single sampling variables plan (s.d. known) which will accepts lots having 2 percent defectives with probability 0.95 and lots having 5 percent defectives with probability 0.10.

Examiner: A. Mathai

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Course. Senior.
Final Examination. February 1960
Paper VI. Session 2.

Date: 14 February, 1960. Time: 2.30 p.m.-5.30 p.m.

1. From past experience it is known that the percentage yield in a chemical reaction depends on three factors: T, the temperature of the bath; P, the phosphorous content of the raw material and C, the percentage impurities in the catalyst. A 3^3 experiment was conducted with T at three levels (130°C, 135°C, 140°C), P at three levels (1 %, 2%, 3%) and C at three levels (2%, 4%, 6%). The experimental data are given below:

		P_1	P_2	P_3
C_1	T_1	79	78	75
	T_2	84	82	80
	T_3	79	76	75
C_2	T_1	78	74	77
	T_2	88	80	83
	T_3	79	75	77
C_3	T_1	78	77	75
	T_2	84	82	83
	T_3	78	77	76

Test whether the three levels of temperature give significantly different yields. Also determine the optimum level of temperature.

Assuming that this optimum level is adopted by the factory give a control procedure. In giving the control procedure you should note that it is not possible to change the phosphorous content or the impurities in catalyst.

2. The following data pertains to the weight of certain randomly chosen types of fan belts observed after the skiving operation. Data on each type of belt consist of several samples of size 3 drawn at regular intervals. The analysis is summarised below.

type of belt	average weight	no. of samples	within samples d.f.	within samples S.S.	between samples d.f.	between samples S.S.
V - 234	325	35	70	5318.00	34	8639.39
V - 445	297	17	34	2255.33	16	7720.08
V -1001	262	15	30	2686.00	14	11672.90
V - 18	193	43	86	3601.33	42	5437.66
V - 110	184	16	32	1201.33	15	2263.17

Classify the fan belts into homogeneous groups. Suggest a control scheme for one of the groups.

Examiner: S.C.C. Unit.

INDIAN STATISTICAL INSTITUTE
Research and Training School

Short-term Statistician's Course. Senior

Final Examination. February, 1960

Paper VII. Session 1.

Special Subject: Econometrics.

Date: 20 February, 1960 Time: 3 P.M. - 5.30 P.M.

Figures in the margin indicate full marks

1. Estimate the constants of the Pareto curve which fit the data below and discuss to what extent the Pareto law is verified. Also find the concentration ratio.

Number (n) of income recipients earning more than Rs x during 1948-49.

Income (Rs x)	Number (n)
15,000	9,100
10,000	21,700
5,000	62,500
1,500	176,000

(15)

2. Fit a Tornquist demand curve $T = \alpha x/(x + \beta)$ for cereals from the data given below. Estimate the expenditure elasticity of cereal consumption for the classes (0, 15) and (15,24) and also for all groups together.

Consumer Expenditure in rupees per month per person
all-India (1955-56)

Expenditure classes in Rs per person per month

item	0-8	8-11	11-13	13-15	15-18	18-21	21-24	24-28	28-34	34-43	43-55	55 and over
per capita exp. on cereals (Rs)	3.49	4.69	5.89	6.35	6.93	7.56	8.26	8.42	8.71	9.33	10.08	13.98
per capita total exp.(Rs)	6.26	9.41	11.98	13.96	16.49	19.54	22.51	25.79	30.68	37.72	47.24	83.29
percentage of persons	14.01	18.31	11.74	9.14	11.55	8.21	7.04	5.53	5.17	4.85	2.59	2.86

(25)

3. Calculate the inflationary gap from the following data:

Income paid out	-	Rs 10000 crores per annum
indirect taxes	-	Rs 700 " "
Government expenditure	-	Rs 1500 " "
Private investment	-	Rs 5000 " "
Consumption expenditure	=	90 per cent of disposable income

If taxation is used as a weapon for curbing the inflationary gap, obtain the amount of tax-increase that will be required for the purpose. (10)

Examiners: N.S.Iyengar and A. Biswas.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Course. Senior.
Final Examination, February, 1960
Paper VII. Session 2.

Special Subject: Econometrics

Date: 20 February, 1960 Time: 6 P.M.-8.30 P.M.

Figures in the margin indicate full marks

1. The table below gives some data on wage rates, prices and population in India:

year	Wage rates (Rs/year)		Price index	Population
	agricultural workers	skilled workers	Base 1948-49	(million)
(1)	(2)	(3)	(4)	(5)
1875	62	162	14	188
1880	72	181	16	189
1890	76	207	17	205
1900	83	203	21	210

F.J. Atkinson gives the following data on national income in India for 1875:

	no. of adult males (million)	national income (Rs.crores)
agricultural class ...	36	313
other classes ...	21	262
total ...	57	575

Obtain average income per adult male for both classes, carry these forward by index numbers of appropriate wage rates; obtain estimates of adult male for other years by using the population estimates and assuming statistics of 1875 proportions and prepare estimates of national income and per capita income at current and at 1948-49 prices for all the years. (20)

2. The following table supplies the elasticity (i.e. percentage change in per capita expenditure on item expressed as a ratio to percentage change in per capita total consumption expenditure) for a number of items in rural India.

Items of expenditure	Elasticity
1. foodgrains	0.52
2. edible oil	1.02
3. milk and milk products	1.99
4. meat, egg and fish	1.14
5. sugar	1.55
6. clothing	1.66
7. education	1.82
8. medicine	1.53
9. conveyance	1.24
10. footwear	1.13

Assuming that the rural population increases at the rate of 2 percent per annum, obtain estimates for the percentage increase in demand for the ten items for (i) 3 p.c. increase in per capita consumption expenditure over a period of 5 years and (ii) 10 p.c. increase in per capita consumption expenditure over a period of 10 years. (20)

P.T.O.

3. The following table supplies capital and labour requirements per unit of net value added in four sectors of the Indian economy :

Parameters of capital and labour requirements

sector	capital (Rs.millions)	labour (man-years)
1	5.00	250
2	2.86	327
3	0.80	320
4	2.22	593

Out of a total investment fund of Rs.56,000 million, Rs.18,500 million is allocated to sector 1 (producers goods industries) on the basis of considerations of long-run growth. The planning authority fixes the employment target at 11 million man-years, and wants to maximize national income (which is given here by the sum of the net value added in the four sectors).

Present the planning problem cited above in the framework of linear programming. (10)

Examiners: A.K.Biswas and M.Mukherjee.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Course. Senior
Final Examination. February, 1960

Paper VII. Session 1.

Special Subject: Sample Surveys

Date: 20 February, 1960 Time: 3 P.M.-5.30 P.M.

Figures in the margin indicate full marks

1. The following table shows the number of inhabitants (in 000's) for the years 1920 and 1930 in each of a random sample of 20 cities from a population of 187 cities.

sl. no.	population (000) in		sl. no.	population (000) in	
	1920 (x_i)	1930 (y_i)		1920 (x_i)	1930 (y_i)
1	76	80	11	121	113
2	67	67	12	44	58
3	381	464	13	64	63
4	37	63	14	56	142
5	61	69	15	40	64
6	93	104	16	136	139
7	78	106	17	46	53
8	60	57	18	243	291
9	2	50	19	30	111
10	179	260	20	256	288

Estimate the total number of inhabitants in the 187 cities in 1930 and the standard error of this estimate by the following methods:

i) simple random sampling
ii) ratio method of estimation
iii) regression method of estimation (33)

The total number of inhabitants in all the 187 cities in 1920 was 22000.

2. The following table shows the data derived from a stratified random sample of tyre dealers.

stratum (h)	tyre dealers in stratum N_h	number of sample tyre dealers n_h	average number of tyres in hand per dealer $\bar{y}_h$	variance s_h^2
1	199	30	4.1	34.8
2	33	6	13.0	92.2
3	10	3	25.0	174.2
4	6	2	38.2	320.4

Estimate the gain in precision due to stratified sampling compared to that of simple random sampling. (17)

Examiner: S. Raja Rao.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statiatician's Course. Senior
Final Examination. February, 1960

Paper VII. Session 2.

Special Subject: Sample Surveys

Date: 20 February, 1960 Time: 6 P.M.-8.30 P.M.

1. It is proposed to carry out a sample survey to study the extent of indebtedness of scheduled cast and scheduled tribes in India.

(a) What sampling plan do you propose ?

(b) Draw up a suitable schedule for enquiry.

(c) Define the important terms that you have used in the schedule and also state all the important concepts which an investigator should keep in mind while collecting data.

OR

It is proposed to carry out a sample survey of landholdings in 1960. The schedules proposed to be used and some notes on these schedules are supplied to you. The purpose is to study the distribution of holdings of different category and their utilisation. Draw up the important tables with suitable tabular and columnar headings that you wish to present as end products of the survey.

Examiner: S. Raja Rao.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Training Course, Senior
Periodical Examination

Planning and Economic Statistics

Date: 24 May 1960. Time: 6 P.M.-9 P.M.

Answer any four

1. Give three definitions of national income and demonstrate their equivalence.

2. 'Whereas investment in building, say, a textile mill has a once for all effect in enlarging final (consumers') output when the mill goes into production, investment in a steel plant, by increasing the annual rate at which new textile mills etc., can be built in the future, has an effect on the future growth-rate of final output'. - Develop fully the ideas contained in this statement.

3. How do you analyse the effect of an increase in the final demand for a product on other sectors of the economy ? What are the basic assumptions in your analysis ?

4(i) Develop two methods of dealing with imports in an inter-industry table.

(ii) State the procedure of analysing trade and transport margins in the construction of an inter-industry table.

5. If you take a certain rate of increase in per capita income as your objective, how would you proceed to obtain investment requirements ?

Examiner: A. K. Biswas.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Training Course, Senior.
Periodical Examination

Statistical Quality Control

Date: 26 May 1960 Time: 6 P.M.-8.30 P.M.

1. A process is producing 6 percent defective articles.

   (a) Obtain the chance of obtaining 3 defectives in a sample of size 25.

   (b) Obtain 'i' such that the probability of a sample (n = 25) containing i or more defectives is less than 0.05 under controlled conditions.

   (c) If the process deteriorates to 10 percent it is desired to be detected by 90 percent of the samples. Also if there is no change in the process level it is desired that the process should not be stopped in more than 5 percent of the cases. Suggest a suitable control scheme.

2.(a) The range in a sample of size 5 is 0.006". Obtain an estimate of and the 95 percent confidence limits of the standard deviation.

   (b) The average range in 10 samples each of size 5 is 0.002" on machine 1 and 0.003" in 15 samples of size 5 on machine 2. Are the variabilities of the two machines different ?

   (c) 5 samples of size 5 each taken on the same machine at different times gave the following

Average max.	:	30.6 units
min.	:	19.4 units
Range Average	:	5.6 units

   Are the sample averages homogeneous ?

3. How can χ^2 test be used to obtain achievable standards ? Indicate the difference, if any, of analysing attribute data by χ^2 test and control charts.

   How can we overcome the disadvantages of a small sample in an attribute control chart?

Examiner: SQC. Unit.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Training Course, Senior.
Periodical Examination

Algebra and Sample Surveys

Date: 2 June 1960 Time: 6 P.M.-9 P.M.

Answer the two groups in separate answer books.

Group A

1. (a) What is meant by a 'vector space' ? Give an example of a vector space.

   (b) Define 'linear independence' of vectors. Give examples of 'linearly independent' and 'linearly dependent' set of vectors in the vector space which you have considered in (a).

   (c) Define a 'basis' and a 'normal orthogonal basis' of a vector space and discuss their usefulness in vector algebra.

   (d) What is meant by the 'dimension' of a vector space ? Show that in a n-dimensional vector space any n+1 vectors are necessarily dependent.

2. Consider the matrix

$$A = \begin{bmatrix} 2 & 2 & -1 & 2 \\ 3 & 1 & 2 & 0 \\ 0 & -4 & 7 & -6 \end{bmatrix}$$

   (a) Compute AA'.
   (b) Evaluate the rank of AA'.
   (c) Examine if the equations $x\,AA' = (-7, 8, 43)$ are consistent. If so, determine a particular solution of these equations.

Group B

Answer any two of the following.

3. Write notes on the following sampling schemes explaining the operational procedure of drawing the sample, advantages and disadvantages in large scale sample surveys and the method of estimation of the population total of a character and the estimate of its variance.

   (i) Sampling of units with probability proportional to a given measure of size.

   (ii) Cluster sampling.

   (iii) Multistage sampling.

4. (a) 'Cluster samples can be regarded as a particular case of systematic sampling'. - Explain.

   (b) Derive a condition under which a systematic sample is more efficient than simple random sample for any given sample size.

5. What are the stages of work in succession, involved in executing a large scale sample survey ?

Examiners: S.K.Mitra and S.Raja Rao.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Training Course, Senior.
Final Examination, July-Aug.'60

Paper I

Date: 10 August 1960. Time: 6.00 P.M.-9.00 P.M.

N.B. Figures in the margin indicate full marks.

1. Consider the equations

$$\begin{aligned} x_1 + 2x_2 - 7x_3 &= 0 \\ 2x_1 + x_3 - 3x_4 &= 0 \\ 4x_2 - 15x_3 - 6x_4 &= 0 \end{aligned}$$

i) Show that the collection C of all solutions of these equation forms a vector space. You should clearly define vector addition and the product of a vector by a real number and demonstrate that the properties required of a vector space are satisfied in this case. [5]

ii) Is (1, 1, 0, -1) a member of C? [3]

iii) What is the dimension of C ? [3]

iv) Suggest a basis of C. [4]

v) Consider the matrix

$$A = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ 0 & 0 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix}$$

If $\underline{x} = (x_1, x_2, x_3, x_4)$ and $\underline{y} = (y_1, y_2, y_3, y_4)$.
Show that $x A y'$ can be taken to be a valid definition of inner product between x and y. [5]

vi) Accepting this definition of inner product

(a) find a vector y which is orthogonal to (6, -3, 0, -4) [5]

(b) find a normal orthogonal basis of C. [4]

2. Let $X = (X_1, \ldots X_k)$ be a random variable with variance covariance matrix A. Let $Y = (Y_1 \ldots Y_k) = XT$ where T is a k x k matrix.

i) Show that the variance covariance matrix of Y is $T'AT = B$ [5]

ii) Show that $X A^{-1} X' = Y B^{-1} Y'$. [5]

iii) If X have a normal distribution with mean zero and variance covariance matrix A, show that

$X A^{-1} X'$ is distributed as a χ^2 with k d.f. [5]

[Hint: Use the fact that there exists a matrix T with the property $T'AT = I$, and use (ii)].

3. i) State the general problem of linear estimation. [5]

ii) What is a BLUE for a parametric function? [3]

iii) What is a zero function ? [2]

iv) Show that the necessary and sufficient condition that an estimate is a BLUE is that it is uncorrelated with all zero functions. [6]

v) Show that there cannot be two (different) estimates which are BLUE for the same parametric function. [4]

4. EITHER

i) Explain what is meant by a Simple Lattice design. You are given 16 treatments $t_1, t_2, \ldots t_{16}$. Prepare a simple lattice lay-out so as to enable you to carry out the combined intra the inter-block analysis. Give full details of your randomisation procedure. [25]

ii) Obtain the estimates of treatment effects using the general equations. Prepare a blank lay-out for your computational procedure. Give a blank lay-out for the ANOVA table. [15]

OR

i) Explain the general motivation for incomplete block designs and the special importance of the balanced incomplete block designs. Explain briefly what is meant by recovery of inter-block information, showing its advantages and objections to it, if any. [20]

ii) An incomplete block design in which v treatments are applied to b blocks of k plots each such that, each treatment occurs r times in the design, and any two blocks have μ treatments in common is called a Linked block design.

Using the general equations for getting the intra-block estimates of treatment effects, given in the lecture notes, obtain the solutions for treatment effects in the case of a Linked block design.

How will you compute treatment sum of squares adjusted for blocks ? [20]

Hint: Use the P-method of analysis.

Examiners: S.K.Mitra, J. Sethuraman and K.R.Shah.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Training Course, Senior
Final Examination, August 1960
Paper II

12 August 1960 Time: 6 P.M.-9 P.M.

N.B. (1) Figures in the margin indicate full marks.
(2) Answers for the two groups should be written in two separate answer books.

GROUP A

(Suggested time 2 hours)

Sampling Distributions, Estimation and Testing of Hypothesis.

N.B. Answer any five questions. Of questions 1, 2, and 3 at least one should be answered. Questions 7 and 8 are compulsory.

1. Consider the following procedure for testing the null hypothesis that a given coin is unbiased.

Procedure: Toss the coin n times. Reject the null hypothesis if the n tosses are all heads or all tails.

Briefly answer the following questions:-

(a) What is the probability of the first kind of error ?

(b) What is the power function ?

(c) How large should n be in order that the probability of the first kind of error (for the above test procedure) is not more than .05 ?

(d) Is the test uniformly unbiased? Give reasons. [12]

2. An urn contains a very large number of balls, each ball being green, yellow or red. It is known that the relative proportions of green, yellow and red balls are either (Hypothesis H_0) .8, .1 and .1 respectively or (Hypothesis H_1) .1, .5 and .4. In order to determine the composition of the urn, the following experiment is proposed.

Experiment: Draw one ball. If this ball is green or yellow, no more balls are drawn, but if red, one more ball is drawn.

a) Let X be the sample (outcome of the experiment) carefully list the sample space (i.e. the set of possibilities for X).

b) Carefully tabulate the probability distribution of X under both H_0 and H_1.

c) What is the most powerful test of the null hypothesis (in the present case) for which the probability of the first kind of error is .01 ? Briefly give reasons. [12]

3. Let $X_1, X_2, \ldots X_n$ be n independent observations on a random variable X whose frequency function is of the form $a(\theta)\, H(x) e^{b(\theta)T(x)}$, θ being the unknown real valued parameter.

a) Write down the likelihood function for θ.

b) How would you prove that $b(\theta)$ is a monotonic function.

c) Obtain the uniformly most powerful test for the null hypothesis $H(\theta = \theta_0)$ against one sided alternatives.

d) How would you prove that no uniformly most powerful test can exist against both sided alternatives ?

[12]

P.T.O.

4. Let X and Y be two independent random variables each having the uniform distribution over the unit interval. Find the probability distribution of XY.

5. Let $X_1, X_2, \ldots X_n$ be n independent normal variables with equal standard deviation σ (the means may differ). If $Y_1, Y_2, \ldots Y_n$ are obtained from the X_i's by means of an orthogonal linear transformation, then prove that the Y_i's are also independent normal variables.

6. State and prove a result which says that under certain conditions the ratio of two chance variables is a beta variable.

7. State clearly Rao-Blackwell theorem in the theory of estimation. [6

Argue out clearly, how this theorem can be used in practice to obtain minimum variance unbiased estimates. [10

Illustrate your argument by means of an example.

8. State and prove the Cramer-Rao inequality in the theory of unbiased estimation. Under what conditions will the equality sign of the Cramer-Rao inequality be attained. [8

Consider a sample $s = (x_1, \ldots x_n)$ of n independent observations from the normal distribution $N(\theta, 1)$. Find the Cramer-Rao lower bounds for the variance of unbiased estimation of (i) θ and (ii) θ^2. State whether these lower bounds are attained. [

GROUP B

(Suggested time 1 hour)

Sample Surveys

1. EITHER

A simple random sample of n units from a population of N units are drawn without replacement. Two characters, x and y, are measured for each of the n sample units.

$X = \sum_{i=1}^{N} x_i$ is known. [20

(i) How do you estimate the total of character y for all the N units ($= \sum_{i=1}^{N} y_i$) by the ratio method of estimation ?

(ii) Is the estimate biased ? If so, derive the expression for the bias in the ratio estimate, mentioning clearly the assumptions involved (if any).

(iii) If the ratio estimate is generally biased, what is the condition for the ratio estimate to be unbiased ?

(iv) Derive an approximate expression for the variance of the ratio estimate.

P.T.O.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Training Course, Senior.
Final Examination, July-Aug. 1960.
Paper III, Session I

Date: 7 August 1960. Time: 11.00 A.M.-2.00 P.M.

Use of practical records is allowed.

1. The following were computed from data relating to 200 American women and 160 Dutch women.

X_0 = Weight (kg.) X_1 = Stature (cm.)
X_2 = Back length (cm.) X_3 = Back width (cm.)

Results from the data relating to 200 American women

variance covariance matrix

X_1	X_2	X_3	X_0
39.690	7.685	1.127	15.711
	5.570	1.430	6.269
		9.610	24.668
			138.768

Means 160.43 38.02 34.01 60.55

Results from the data relating to 160 Dutch women

variance covariance matrix

X_1	X_2	X_3	X_0
42.250	7.047	1.950	16.055
	5.664	1.618	8.795
		8.352	20.327
			119.028

means 162.05 36.59 36.40 66.80

EITHER

Test whether American women and Dutch women are similar as far as the mean values of the four characteristics X_1, X_2, X_3 and X_0 are concerned.

State clearly the assumptions in your test.

OR

Considering only the data relating to 200 American women

(a) Compute the multiple correlation coefficient of X_0 on X_1, X_2 and X_3.

(b) Test the efficacy of linear prediction of X_0 on the basis of X_1, X_2 and X_3.

(c) Test whether the mean of the characteristics X_1, X_2 X_3 and X_0 are 162, 39, 35 and 60 respectively.

Variance covariance matrix

= (sum of squares and products matrix)/degrees of freedom

P.T.O.

2. EITHER

Consider the following observational equations

$$E(y_1) = 2\theta_1 + 3\theta_2 + 5\theta_3$$
$$E(y_2) = \theta_2 - \theta_3$$
$$E(y_3) = 4\theta_1 + 5\theta_2 + 11\theta_3$$
$$E(y_4) = \theta_1 + \theta_2 + 3\theta_3$$

i) Are all linear functions of θ_1, θ_2, θ_3 estimable?

ii) Give as many linearly-independent linearly estimable functions of θ_1, θ_2, θ_3 as you can.

iii) Suggest a linear function of $\theta_1, \theta_2, \theta_3$ which is not estimable.

iv) Is the linear function $\theta_1 + 5\theta_2 - \theta_3$ estimable ?

v) If so, give the best linear unbiased estimate.

vi) What is the variance of this estimate, under the assumption that all the y's are independent and have variance σ^2 ?

vii) On taking observations we get, $y_1 = 5$, $y_2 = 7$, $y_3 = 9$, $y_4 = 2$.

Test whether $\theta_1 + 9\theta_2 - \theta_3 = 0$

viii) Test simultaneously the hypotheses

$$\theta_1 + 5\theta_2 - \theta_3 = 0$$
$$\theta_1 - 4\theta_3 = 0$$

OR

The following is a record of replicated measurements of profile distances of three subjects (S_1, S_2, S_3) taken by two observers (O_1, O_2).

observer	Measurements of profile distance subject		
	S_1	S_2	S_3
O_1	9.88	9.22	9.60
	9.84	9.26	9.66
		9.23	
O_2	10.08	9.10	9.82
	9.98	9.24	9.54
	10.02		9.60

Let μ_{ij} be the expected value of the measuremant taken by the ith ($i = 1, 2$) observer on the jth subject ($j = 1, 2, 3$) and write

$$\mu_{i\cdot} = \frac{1}{3}\sum_{j=1}^{3} \mu_{ij} \; , \; \mu_{\cdot j} = \frac{1}{2}\sum_{i=1}^{2} \mu_{ij}$$

$$\mu_{\cdot\cdot} = \frac{1}{6}\sum_{i,j} \mu_{ij}$$

The main effect of the ith observer is defined as $\alpha_i = \mu_{i\cdot} - \mu_{\cdot\cdot}$

Similarly the main effect of the jth subject is defined by $\beta_j = \mu_{\cdot j} - \mu_{\cdot\cdot}$

The interaction between the ith observer and the jth subject is defined by

$$\gamma_{ij} = \mu_{ij} - \mu_{i\cdot} - \mu_{\cdot j} + \mu_{\cdot\cdot}$$

(1) Estimate α_i, β_j and γ_{11} and their standard errors.

(2) Test if $\gamma_{11} = 0$.

(3) We say there is no main effect of observers if $\alpha_i = 0$, $\alpha_2 (= -\alpha_1) = 0$. Similarly the hypothesis of no interaction between observers and subjects implies.

$$\gamma_{ij} = 0 \quad \text{for all } i, j$$

which is actually equivalent to

$$H_0 \left\{ \gamma_{11} = 0, \quad \gamma_{12} = 0 \right\}$$

since $\gamma_{13} = -\gamma_{11} - \gamma_{12}$

$$\gamma_{21} = -\gamma_{11}, \quad \gamma_{22} = -\gamma_{12}$$

$$\gamma_{23} = -\gamma_{13}.$$

Analyse the data to examine whether there are significant observer differences, subject differences and observer X subject interaction.

Examiners: S. K. Mitra and J. Sethuraman.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Training Course, Senior
Final Examination, July-Aug. 1960.
Paper III Session II

Date: 7 August 1960. Time: 2.30 P.M.-5.30 P.M.

N.B. 1. Figures in the margin indicate full marks.
2. Answer all questions.
3. Use of practical records is allowed.

DESIGN OF EXPERIMENTS

1. Explain the general motivation for incomplete block designs and the special importance of balanced incomplete block designs. Explain briefly what is meant by recovery of inter-block information, showing it's advantages and the objections to it, if any. (13)

2. An incomplete block design, in which v treatments are applied to b blocks of k plots each such that, each treatment occurs r times in the design, and any two blocks have λ treatments in common, is called a Linked block design.

Using the general equations for getting the intra-block estimates of treatment effects, given in the lecture notes, obtain the solutions for treatment effects in the case of a Linked block design. (13)

Hint: Use the P-method of analysis.

3. In an experiment to compare four seed treatments on three manure varieties, split-plot technique was adopted. The manure varieties denoted by a_0, a_1, a_2 were allotted at random to the whole plots independently in each replicate. The four seed treatments denoted by b_0, b_1, b_2, b_3 were allotted at random to the four sub-plots independently in each whole plot.

The following table gives the number of plants germinating from 100 seeds for each of four seed treatments with each of three manure varieties in three replicates.

Analyse the data.

Repl. I			Repl. II			Repl. III		
a_0b_1 12	a_2b_3 10	a_1b_1 26	a_1b_2 5	a_0b_3 12	a_2b_1 16	a_2b_2 29	a_0b_1 13	a_1b_1 18
a_0b_2 13	a_2b_0 51	a_1b_0 77	a_1b_0 47	a_0b_0 63	a_2b_2 30	a_2b_1 14	a_0b_3 7	a_1b_0 66
a_0b_0 66	a_2b_1 8	a_1b_2 27	a_1b_1 11	a_0b_2 13	a_2b_0 81	a_2b_3 10	a_0b_0 70	a_1b_2 11
a_0b_3 6	a_2b_2 20	a_1b_3 15	a_1b_3 4	a_0b_1 10	a_2b_3 14	a_2b_0 63	a_0b_2 11	a_1b_3 15
97	89	145	67	98	141	116	101	110

(24)

Correction term = 25813.778

Total S.S. = 19652.222

P.T.O.

SAMPLE SURVEYS

3. You are given a list of 32 factories engaged in producing machineries used in Industries and the average number of persons employed per day are also given (please see statement in the reverse page).

(i) Draw a sample of 5 factories with probability proportional to the average number employed per day, with replacement. Explain clearly the steps that you follow.

(ii) You will notice in the statement that the average number employed are not given for two factories. State clearly how best you have dealt with these two factories for the purpose of drawing the sample.

4. A Sample of 20 factories was drawn with probability proportional to the total number employed (x) with replacement from a population of 325 factories. The number employed (x) and the number observed (y) on a day in each of these 20 factories are given below.

(i) Estimate the percentage of workers absent in the population of 325 factories.

(ii) Estimate the percentage error of this estimate.

x	y	x	y	x	y	x	y
148	31	103	18	91	28	96	45
86	22	53	4	125	18	52	16
64	12	42	25	51	19	82	10
81	9	67	27	65	16	85	35
73	13	33	8	43	12	142	15

List of factories producing machineries used in Industry (in greater Calcutta).

sl.	factory	number	average number of persons employed	sl.	factory	number	average number of persons employed
1	23b	2402	19	17	29b	2532	359
2		2404	930	18		2533	91
3		2415	74	19		2534	
4		2435	45	20		2535	104
5		2436	20	21		2537	107
6		2440	187	22		2539	126
7		2451		23		2540	85
8		2462	61	24		2541	716
9		2464	44	25		2547	207
10		2466	225	26		2555	293
11		2467	813	27		2556	46
12		2474	275	28	57	2625	2809
13	29b	2505	53	29	62	2630	578
14		2515	1501	30		2631	617
15		2528	29	31		2632	523
16		2529	178	32		2633	146

(30)

Examiners: K.R.Shah and S. Raja Rao.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Training Course, Senior
Final Examination, August 1960
Paper 4 Theory SQC.

Date: 17 August 1960 Time: 6 P.M.-9 P.M.

GROUP A

1. Write brief notes on the following: (i) control by gauging (ii) OC of a control chart (iii) cost of inspection (iv) estimation of process average.

2. State Dodge's Continuous Sampling Plan. What are the assumptions in the plan? Explain how the plan will be presented for practical use.

3. State a single sampling acceptance procedure for variables inspection for the situation when lot standard deviation is not known and mean-range will be used.

Derive the elements in the plan.

GROUP B

1. Explain the terms loss in utilisation and loss in efficiency as applied to a production process. Under what conditions are these two equal ?

Briefly describe some of the techniques that are useful in estimating the loss in utilisation.

In a workshop consisting of 32 drilling machines, 15 lathes and 16 heavy machines, describe how you would organize the collection of data for estimating the loss in utilisation. Indicate also the statistical analysis that you would do and the manner in which you would present the results in order to enable management to take action.

2. Explain how process capability studies are helpful in planning production in a job-shop taking many short-run jobs.

3. Enumerate some of the points to be taken into account in designing an industrial experiment and interpreting the results.

Examiners: A. Matthai and S.Q.C. Unit.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Training Course, Senior
Final Examination, August, 1960

Date: 13 August 1960. Paper V Session 1 Time: 3 P.M.-5.30 P.M.

Practical S.Q.C.

1. Given that in a sample of 20 items, the mean was 12.5, the standard deviation 1.2, the range 3.3 and the mean-range based on sub-samples of 5 items 2.7, (all in the same units) obtain interval estimates for the following:-

   i) the process mean (using s.d)
   ii) the process mean (using range or mean-range)
   iii) the process standard deviation (using s.d)
   iv) the process standard deviation (using range).

2. Plot by taking at least 7 points the OC curve of a single sampling variables acceptance plan (s.d. unknown) having elements n = 60 and k = 1.362.

3. Explain(with an example but without deviation or actual calculations) how a sequential inspection procedure will be carried out by the Tabular procedure, for defects per unit inspection.

Examiner: A. Matthai.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statisticians Training Course, Senior
Final Examination, August 1960

Paper V Session 2

Practical S.Q.C.

Date : 13 August 1960. Time: 6 P.M.-8.30 P.M.

1. Data collected from a drawing frame in a cotton mill at the third pass is given below. 4 sample pieces of 6 lbs. length of sliver were selected 5 times in a day for 5 days and their weights are given in some units.

Date	Period	Weights of sample pieces				Total	
		1	2	3	4		
4.7.60	7-30	12	10	12	16	50	6
	10-30	10	11	22	16	59	12
	12-30	30	20	30	20	100	10
	3-45	15	22	16	22	75	7
	6-30	18	20	10	18	66	10
6.7.60	7-30	18	8	16	18	60	10
	10-30	21	20	19	11	71	10
	12-30	13	23	13	30	79	17
	3-45	10	8	10	14	42	6
	6-30	9	14	8	15	46	7
7.7.60	7-30	0	0	6	4	10	6
	10-30	10	10	9	7	36	3
	12-30	15	15	15	15	60	0
	3-45	16	9	19	12	56	10
	6-30	22	20	15	15	72	7
8.7.60	7-30	4	16	13	16	49	12
	10-30	20	24	20	20	84	4
	12-30	10	10	12	13	45	3
	3-45	15	15	12	10	52	5
	6-30	10	10	10	10	40	0
9.7.60	7-30	14	19	2	11	46	17
	10-30	10	13	10	7	40	6
	12-30	10	12	10	9	41	3
	3-45	13	12	12	10	47	3
	6-30	15	15	15	11	56	4

Actual sum of squares of all observations is 22302.

(1) Analyse the data and describe briefly the working of the process.

(2) What is the variability in pieces of 6 yds. length under the present working conditions ?

If it is possible to control the process average strictly, what will be the process capability ?

(3) The specification limits on this characteristic are 115 ± 25. Describe in detail how control charts should be set up by assuming that the sampling procedure is to select samples of size 4 several times in a day.

P.T.O.

2. It is required to compare the machine capability of two Herbert Automats which are used to do the turning operation for outside diameter of fan motor shaft. 5 samples of size 5 were selected from each machine production and the outside diameter measured correct to a thou. The deviation from some origin are given below:

Automat 1						Automat 2					
Sample						Sample					
1	8	5	8	0	7	1	6	8	6	9	6
2	5	7	5	8	9	2	8	8	7	5	8
3	5	8	7	8	0	3	8	0	7	5	0
4	8	9	7	5	2	4	5	9	6	6	0
5	6	7	5	0	5	5	9	8	5	8	7

Compare the machine capabilities.

Examiner: SQC Unit.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statisticians' Training Course, Senior
Final Examination, August 1960
Date: 19.8.60 Paper VI. Time: 6 P.M.- 9 P.M.

All questions carry equal marks.
Answer any five out of the following.

1. Analyse the effect of the following types of fiscal policy on consumption and national income:

   (i) balanced budget method

   (ii) tax reduction method.

2. Describe the Tornquist's system of demand curves. In what way do these curves differ from the double logarithmic forms of demand curves ? How would you proceed to estimate the constants of the Tornquist system ?

3. Examine the following propositions:

   (i) an increase in the propensity to save leads to a reduction in national income,

   (ii) planning for economic development implies a rise in the rate of saving.

   Are the two statements contradictory? Explain.

4. What do you understand by a concentration map used in the study of consumers' behaviour ? How would you use such a map in deriving the Engel elasticity of a given commodity ? What are the basic assumptions and how are they tested in actual practice ?

5. Outline briefly the procedure of construction of an inter-industry table distinguishing domestic inputs from imported inputs.

   Assuming that imported inputs are fixed per unit of output in different branches of the economy, analyse the effect of an increase in the final bill of goods on import requirements.

6. Explain briefly any two the following propositions:

   (i) saving and investment are always equal,

   (ii) national income is the aggregate of net value added in different sectors of the economy,

   (iii) if government expenditure is increased by one million rupees, national income will increase by some multiple of one million.

Examiners: A. K. Biswas and N.S.Iyenger.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Training Course, Senior.
Final Examination, August, 1960.
Paper VII, Session 1.

Date: 20 August, 1960. Time: 3 P.M.-5.30 P.M.

1. Using the information contained in table 1 fit a 'quality equation' of the form

$$\log p = \alpha_0 + \alpha_1 \log E$$

where p is the average price obtained as the ratio of the value of consumption to the quantity of consumption, E being the total expenditure per capita. Obtain also 95 percent confidence limits for your estimates and interpret your results.

Table 1. Monthly consumption per capita of cereals in rural India by classes of total expenditure per capita per month.

sl. no.	monthly per capita expenditure classes (Rs.)	population %	total expenditure per capita per month (Rs.)	monthly per capita consumption of cereals: quantity (seers)	value (Rs.)
1	0 - 8	15.57	6.27	12.66	3.39
2	8 - 11	18.07	9.41	14.69	4.66
3	11 - 13	12.23	11.97	17.35	6.03
4	13 - 15	9.38	13.96	18.83	6.38
5	15 - 18	11.35	16.52	24.78	7.09
6	18 - 21	7.92	19.59	28.26	7.79
7	21 - 24	7.03	22.36	22.85	8.60
8	24 - 28	5.12	25.73	23.58	8.91
9	28 - 34	4.66	30.54	24.96	9.31
10	34 - 43	4.35	37.64	25.47	10.18
11	43 - 55	2.08	46.30	31.82	11.46
12	55 and above	2.24	81.05	46.66	14.31

2. Using the data of table 1 estimate the constants of the following demand equation

$$Q = \frac{\alpha E}{E + \beta}, \qquad E > 0$$

Obtain the average Engel elasticities for the expenditure groups Rs. 0 - 15 and Rs. 24 - 34.

Examiner : N. S. Iyengar

INDIAN STATISTICAL INSTITUTE
Research and Training School
Short-term Statistician's Training Course, Senior.
Final Examination, August, 1960.

Paper VII, Session 2.

Date: 20 August 1960. Time: 6 P.M. - 8.30 P.M.

Let the Indian Economy be considered as consisting of the following 4 sectors: S_1 : Basic Investment goods, S_2 : Factory Consumer goods, S_3 : Household Industries (including agriculture) and S_4 : Services. It has been determined on the basis of some studies that the income-investment ratios (income per rupee worth of investment) in the above 4 sectors are respectively 0.20; 0.35; 1.25 and 0.45. The capital requirements per employed person in the sectors are Rs.20,000: Rs.8,750; Rs.2,500 and Rs.3,750 respectively. The Planning Commission desires that the additional employment in the Household Industries Sector and Services Sector should not together exceed 9 million. Further the Commission has the information that the additional employment that would have to be created during the Plan period will not exceed 11 million. Given that the total resources available for investment during the period is less than Rs.5,600 crores, suggest an optimum investment plan for the Planning Commission, so that it maximises national income. (Use the simplex technique of linear programming).

Examiner: A. K. Biswas.