

INDIAN STATISTICAL INSTITUTE

Computers Cert.; Diploma & Field Survey
Question Papers.

1951

COMPUTER'S CERTIFICATE EXAMINATION, 1951

PART IA: SECTION II

Time: 3 hours

Full marks 100

- (1) Answers to the different groups are to be given in separate books.
- (2) Use of Calculating machines is not permitted.
- (3) Figures in the margin indicate full marks.

GROUP A

1. (a) Solve for x :
 $432x^2 + 1283x + 951 = 0$.
- (b) Evaluate:
 $\Sigma x^2 + \Sigma y^2 - \Sigma z^2$

for values of $x = 524$ to 529 , $y = 68$ to 80 , $z = 35$ to 44 with the help of appropriate tables. 20

2. (i) Using the values given below of weights and indices, calculate the weighted mean index of (a) Food and (b) Fuels.

(ii) Using the values of indices obtained under (i) and also the values given below of weights and indices calculate the cost of living index by computing the weighted average of the separate indices Food, Fuel, etc.

Food items	Weight	Index	Fuel items	Weight	Index
Rice	33	383.3	Kerosene	27	258.2
Ata	11	402.5	Firewood	35	310.0
Dal	8	615.9	Coal	25	502.1
Ghee	5	823.8	Matches	13	320.0
Oil	5	611.0		100	41376
Salt	1	251.1			
Spices	3	991.6	"Clothings" index	= 882.0	547
Sugar	2	605.4	"Rent" index	= 315.6	
Refreshments	7	353.3	"Miscellaneous" index	= 441.8	
Milk	7	415.0			
Fish	9	607.1			
Vegetables	9	711.4			

100

Weights of different groups

Groups	Weights
Food	69
Fuels	6
Clothings	8
Rent	5
Miscellaneous	12
	100

-75 365.3

3. The following table shows the distribution of income and expenditure of a number of families. Find the mean and standard deviation of the income distribution and the mean and standard deviation of the expenditure distribution of all the families taken together. 20

Expenditure (Rs.) Income (Rs.)	30-35	35-40	40-45	45-50	50-55	55-60	60-65	65-70	70-75	75-80
30-35	4	3	—	—	—	1	—	—	—	—
35-40	—	6	8	1	1	2	—	1	—	—
40-45	1	—	14	10	4	3	—	3	—	1
45-50	—	2	4	24	18	10	9	6	—	1
50-55	—	—	—	6	40	10	22	6	1	2
55-60	—	—	1	—	36	48	12	7	—	4
60-65	—	—	1	—	—	6	27	20	15	—
65-70	—	—	—	—	2	12	10	18	11	6
70-75	—	—	—	—	—	—	4	3	11	5
75-80	—	—	—	—	1	—	2	—	6	8

GROUP B

- 4.(a) Evaluate Y for the values, $x=1$ to $x=10$ from the relation:

$$y = 28.78 \sqrt{x}$$

- (b) with the help of mathematical Tables or otherwise, find the value of

$$\frac{99.78 \times 0.465 \times (1.0003)^{100}}{1000.3 \times 0.00395712}$$

5. The following is a statement compiled from the census returns of a certain state:-

Census year	Population in lakhs			
	Hindus	Muslims	Others	Total
1871	27.0	19.5	3.5	50.0
1881	28.0	20.4	4.0	52.4
1891	29.6	21.6	4.4	55.6
1901	31.8	23.1	4.6	59.5
1911	35.4	24.9	4.7	65.0
1921	36.8	27.1	4.8	68.7
1931	38.2	30.1	4.8	73.1
1941	39.4	33.2	4.9	77.5

- (a) Represent graphically the growth of population by decades and the relation with by communities.

- (b) Make estimates of the population by communities for the years 1922 and 1937.

COMPUTER'S CERTIFICATE EXAMINATION 1051.

PART 1B: SECTION I (First set)

Time: 3 hours

Full marks 100

Answers to the different groups are to be given in separate books.

Use of Calculating machines is permitted.

Figures in the margin indicate full marks.

GROUP A

1. The table below gives the index of prices of farm products in U.S.A. during the period 1910-1948.

Year	Index	Year	Index	Year	Index
1910	103	1923	142	1936	114
1911	95	1924	143	1937	122
1912	99 →	1925	156	1938	97
1913	102	1926	146	1939	95
1914	102	1927	141	1940	100
1915	99	1928	149	1941	123
1916	119	1929	148	1942	158
1917	178	1930	125	1943	192
1918	206	1931	87	1944	196
1919	218	1932	65	1945	206
1920	212	1933	70	1946	234
1921	124	1934	90	1947	275
1922	131	1935	109	1948	285

Compute a five-year moving average, and draw a graph showing the data and the moving averages. 25

2. The following table shows the distribution of 203 families according to total yearly expense and the percentage of yearly expense for food.

		Total yearly expense in dollars								
		500-	700-	900-	1100-	1300-	1500-	1700-	1900-	2100-
Percentage of ex- pense for food.	56-	1	1							
	52-	1	5		1					
	48-	4	10	5	2	1		1		
	44-	4	12	4	2	6	3	1		
	40-	3	9	6	13	8	2	1		
	36-	2	5	11	8	6	3	2		
	32-	1	5	2	5	8	8	2		
	28-	1		3	2	4	4			1
	24-		1		1		3	1	3	
	20-		1				2	2		

Calculate the correlation coefficient between total yearly expense and percentage of expense for food. 25

GROUP B

3. The following data show the weight in grams of fixed length of wool thread from 50 "ends" being spun on two bobbins, 25 ends to each bobbin. The interest lies in two factors, the variation between the bobbins and the variation in the 25 ends on the same bobbin, according to their position.

(a) Set out the appropriate analysis of variance for between bobbins, between ends and the residual.

TABLE			
End Number	Bobbin 1	Number 2	Totals
1	7.50	7.23	14.73
2	7.52	7.81	15.33
3	7.70	7.94	15.64
4	7.93	7.94	15.87
5	7.78	7.89	15.67
6	7.73	8.23	15.96
7	8.07	8.27	16.34
8	8.01	8.54	16.55
9	8.22	8.24	16.46
10	8.24	8.35	16.59
11	8.17	8.29	16.46
12	8.09	8.54	16.63
13	8.11	8.45	16.56
14	7.96	8.43	16.39
15	8.09	8.47	16.56
16	8.04	8.33	16.37
17	7.78	8.47	16.25
18	8.11	8.63	16.74
19	8.17	8.31	16.48
20	8.12	8.31	16.43
21	8.13	8.10	16.23
22	8.01	8.01	16.02
23	8.17	7.92	16.09
24	8.05	8.27	16.32
25	7.91	7.92	15.83
	199.61	204.89	404.60

(b) Derive the equation of linear regression of weight of wool thread spun on the 'End Number' taking into account the totals of the two bobbins for the first 20 ends only.

(c) Set out the analysis of variance so as to show what part of the variation between the first 20 ends is explained by the linear regression.

COMPUTER'S CERTIFICATE EXAMINATION, 1951.

PART 1B: SECTION I (Second set)

Time: 3 hours

Full marks 100

- (a) Answers to the different groups are to be given in separate books.
 (b) Use of Calculating machines is permitted.
 (c) Figures in the margin indicate full marks.

GROUP A

1. The frequency distribution of wages of 526 employees is given below.

Yearly wages in dollars	No. of employees
400-599	3
600-799	10
800-999	19
1000-1199	39
1200-1399	82
1400-1599	103
1600-1799	121
1800-1999	101
2000-2199	33
2200-2399	12
2400-2599	1
2600-2799	2

Calculate mean, standard deviation, and β_1 , β_2 for the distribution, using suitable checks in the calculation. Calculate also median and the quartiles,

2. The prices of agricultural commodities for 1946-47 and for the month of December in 1950, are given below along with the value of output of these commodities in 1946-47.

Commodities	Unit	Prices		Value of output in million rupees
		1946-47 Rs. a. p.	Dec., 1950 Rs. a. p.	
Rice	md.	13 12 0	13 12 0	8364
Wheat	md.	9 11 0	9 11 0	2207
Jowar	md.	6 0 6	8 0 0	876
Bajra	md.	6 13 0	8 0 0	373
Maize	md.	7 12 0	13 0 0	295
Barley	md.	7 2 0	7 2 0	484
Groundnut	531 lbs.	148 0 0	171 0 0	850
Rape seed	650 lbs.	107 0 0	174 0 0	553
Linseed	md.	22 0 0	25 2 0	102
Castor seed	cwt.	34 12 0	28 8 0	14
Cotton, raw	784 lbs.	466 0 0	433 0 0	701
Jute, raw	400 lbs.	87 0 0	200 0 0	556
Tea	lb.	1 3 4	1 13 11	534
Coffee	28 lbs.	34 9 0	42 0 0	40

Calculate the weighted Index number of prices of these commodities for December 1950, taking 1946-47 as base.

20

GROUP B

3. From the records of a school medical examination the data for the table given below are found.

y	x							Total
	7	8	9	10	11	12	13	
38								1
39								4
40								15
41								24
42								31
43								41
44								36
45								52
46								49
47								54
48								56
49								46
50								55
51								37
52								34
53								25
54								12
55								9
56								2
57								4
58								3
Total	40	80	108	103	109	96	54	
	1671	3460	4859	4873	5358	4849	2818	

x = age in years, y = height in inches.

(a) Derive the equation of linear regression of height on age and analyse the total variation in heights so as to exhibit that part of it which is explicable as due to age and further that part of the latter which is accounted for by the linear regression of height on ages.

(b) On scrutiny of the original material, it appeared that the height, 51 inches, of a child aged 12 years was misread as 41 inches while preparing the table given above. Make the correction, recalculate the regression equation and set down the revised analysis of variance.

COMPUTER'S CERTIFICATE EXAMINATION, 1951.

PART 1B: SECTION II (First set)

Time: 3 hours

Full marks 100

- (a) Answers to the different groups are to be given in separate books.
 (b) Use of Calculating machines is permitted.
 (c) Figures in the margin indicate full marks.

GROUP A

1. The frequency distribution of the weights (in grams) of a number of iron bars is shown below:—

Weight	Frequency
67.8-69.0	6
69.0-70.2	30
70.2-71.4	42
71.4-72.6	60
72.6-73.8	94
73.8-75.0	120
75.0-76.2	102
76.2-77.4	80
77.4-78.6	54
78.6-79.8	14
79.8-81.0	12

It is known that the mean and the standard deviation of this distribution are 75.0 grams and 2.55 grams respectively.

(a) Assuming the above to be a random sample from a normal population, find out the expected frequencies corresponding to the given observed frequencies and the goodness of fit (of the assumed normal distribution).

(b) Find out the values of the ordinates of the hypothetical normal curve at following values of the variate:—70.8, 73.2, 75.6 and 78.0 grams.

2. The following table shows the frequencies of a number of people having different eye-colours and hair-colours.

Attribute	Hair colour				Total
	Fair	Brown	Black	Red	
Eye-colour					
Blue	1768	807	189	47	2811
Grey	946	1387	746	53	3132
Brown	115	438	288	16	857
Total	2829	2632	1223	116	6800

Assuming that the above is a random sample, test whether there is any association between eye-colour and hair-colour in the population.

3. From a very large population, 1630 persons were chosen at random and their monthly wages recorded; the mean wage was found to be Rs. 350 and 10 annas, while the standard deviation was Rs. 12 and 4 annas. Do you think that the mean wage for the population as a whole may be Rs. 351 and 14 annas? 10

GROUP B

4. (a) Find the value of

$$0.2x^{0.13} + x^{\frac{1}{3}} + \log_e x,$$

when $x = 2.31$.

(b) The following table gives $I(u,p)$ for different values of u and p . Find the value of $I(u,p)$ when $u = 6.23$ and $p = 3.81$ by linear interpolation.

p	u		
	6.2	6.3	6.4
3.7	.9981130	.9983959	.9986374
3.8	.9980781	.9983675	.9986145
3.9	.9980418	.9983380	.9985905

20

5. The following table gives the military expenditure (y_1) and the total expenditure (y_2) of the Government of India during the years 1923-24 to 1931-32. Calculate the percentage of military expenditure to total expenditure (y_1) and draw the graphs of y_1 , y_2 , and y_1/y_2 against different years on the same graph paper.

Year	1923 —24	1924 —25	1925 —26	1926 —27	1927 —28	1928 —29	1929 —30	1930 —31	1931 —32
Military Expenditure	610	600	600	610	560	580	590	583	559
Total Expenditure	2960	2100	2160	2220	2190	2220	2300	2304	2201

20

6. For any five of the following items of information write the name of at least one publication from which you can get the information.

(i) Gross amount of import duty collected at all the ports of India in the month of January, 1931.

(ii) Harvest price of winter rice in India for the 5 years 1938-39 to 1942-43.

(iii) Number of trade disputes in India resulting in stoppage of work, number of worker involved and number of man days lost during December, 1950.

(iv) Average percentage absenteeism in December 1950 in the Mills of Bombay, Ahmedabad and Sholapur.

(v) Earnings from carrying rice (in husk) by class I Railways in Indian Union during the year 1949-50.

(vi) Number of Ration cards registered upto end of June 1947 in Calcutta industrial areas.

(vii) Number of males having production of raw materials as the principal means of livelihood in Burdwan District in 1941.

COMPUTER'S CERTIFICATE EXAMINATION, 1951.

PART IB: SECTION II (Second set)

Time: 3 hours

Full marks 100

- (a) Answers to the different groups are to be given in separate books.
 (b) Use of Calculating machines is permitted.
 (c) Figures in the margin indicate full marks.

GROUP A

1. The frequency distribution of the heights (in inches) of a number of students is shown below:—

Height	Frequency
54.0-55.2	6
55.2-56.4	30
56.4-57.6	42
57.6-58.8	66
58.8-60.0	94
60.0-61.2	120
61.2-62.4	102
62.4-63.6	60
63.6-64.8	54
64.8-66.0	14
66.0-67.2	12

It is known that the mean and the standard deviation of this distribution are 60.6 inches and 2.55 inches respectively.

(a) Assuming the above to be a random sample from a normal population, work out the expected frequencies corresponding to the given observed frequencies, and test the goodness of fit (cf the assumed normal distribution).

(b) Find out the values of the ordinates of the hypothetical normal curve at the following values of the variate:—57.0, 59.4, 61.8 and 64.2 inches. 25

2. One thousand subjects of English, French, German, Italian and Spanish nationality (supposed to be chosen at random) were asked to name their preferences among the music of those five nationalities. The results are shown in the form of a frequency table below.

(E=English, F=French, G=German, I=Italian and S=Spanish)

Nationality of subject	Nationality of music preferred					Totals
	E	F	G	I	S	
E	32	16	75	47	30	200
F	10	67	42	41	40	200
G	12	23	107	36	22	200
I	16	20	44	76	44	200
S	8	63	30	43	66	200
Totals	78	170	298	243	202	1000

Test the association between the nationality of the subject and the nationality of the music preferred. 15

3. From a collection 1,00,000 stone-blocks, 900 blocks were chosen at random and their weights recorded; the mean weight was found to be 16 maunds and 3 seers, while their standard deviation was 4 mls. 2 seers. What are the limits within which the mean weight of all the 1,00,000 blocks is likely to be? 10

GROUP B

4. (a) Evaluate:

$$\frac{5.0^{0.12} + (2.31)^{\frac{1}{2}}}{1 + \log_e 5.6}$$

(b) The following table gives z for different values of x and y . Find z when $x=.14$ and $y=.56$ by linear interpolation.

x	y		
	.4	.5	.6
.1	.3183	.2910	.2630
.2	.3076	.2830	.2573
.3	.2943	.2728	.2498

20

5. The amount of sugar content (y) in an aqueous solution at different time intervals (x) after addition of acid to it is given below. The relation between the sugar content (y) and the time interval (x) is given by:

$$y = 289.5e^{-0.0109x}$$

Represent graphically the observed and expected values of y .

X	0	45	90	150	210
Y	131	114	90	81	67

20

6. For any five of the following items of information write the name of at least one publication from which you can get the information.

- Production of yarn and woven goods in Bombay city in the year 1949.
- Number of persons between the ages 21-25 unemployed for less than one year in the district of Birbhum in 1941.
- Number of deaths due to cholera in Baranagar during the years 1939 to 1943.
- Number and earning from passengers carried by all Railways in Indian Union during the year 1949-50.
- Units of Electrical Energy generated and sold in the Indian dominion during December, 1950.
- Area irrigated from Government canals in different districts of Bihar during the five years 1938-39 to 1942-43.
- Total value of principal articles of merchandise imported by sea and air from foreign countries into India in January 1951. 10

COMPUTER'S CERTIFICATE EXAMINATION, 1951.

PART 1C: SECTION I

Time: 4 hours

Full marks 100

- (a) Answers to the different groups are to be given in separate books.
 (b) Use of Calculating machines is permitted.
 (c) Figures in the margin indicate full marks.

GROUP A

1. The following data relate to the Schopper and Riegler freeness tests on paper pulp samples taken, at hourly intervals during the beating process.

Test on Paper Pulp samples

Hours of beating	1	2	3	4	5	6	7	8	9	10	11	12	13
Schopper and Riegler Freeness: Degrees	17	21	22	27	36	49	56	64	80	86	88	92	94

Determine graphically the appropriate curve to be fitted to the data and find precisely the values of the constants of the equation to the curve.

2. Analyse the following data with a view to grading the different treatments in proper order.

Yield of Napier grass in Srs. per line

Blocks Treatments	1	2	3	4	5	6
Control	88	63	35	23	27	32
Bono & Dung	89	86	44	48	46	61
Dung	84	90	56	65	63	63
Niciphos	93	94	68	75	77	83
Bono and Rape	108	83	48	87	83	88
Bono and Castor	132	112	84	88	96	68

20

GROUP B

3. The following table shows the crop year prices of wheat at Minneapolis and Liverpool and the production of wheat in the United States for different crop years:

Crop year	Crop-year prices, cents per bushel		Production 0,000,000 bushels in United States	Crop year	Crop-year prices, cents per bushel		Production 0,000,000 bushels in United States
	At Minneapolis	At Liverpool			At Minneapolis	At Liverpool	
	x_1	x_2	x_3		x_1	x_2	x_3
1891-92	84	114	68	1903-04	90	90	66
1892-93	66	85	61	1904-05	111	97	56
1893-94	59	73	51	1905-06	84	98	71
1894-95	81	70	54	1906-07	84	94	74
1895-96	57	78	54	1907-08	107	110	63
1896-97	71	89	52	1908-09	116	122	64
1897-98	94	117	61	1909-10	108	117	68
1898-99	69	85	77	1910-11	103	107	63
1899-1900	68	87	66	1911-12	108	114	62
1900-01	73	86	60	1912-13	86	112	73
1901-02	72	88	76	1913-14	88	106	75
1902-03	76	89	69				

Calculate: $r_{11,3}$ and test its significance.

25

4. Evaluate $y = x^{\frac{1}{2}} e^{-x}$ for integral values of x from zero to ten, and determine graphically the value of x for which y is maximum. 10

5. (a) Find the present worth of Rs. 2521 due five years hence at three per cent compound interest.

(b) Calculate the approximate values of Γn for integral values of n from eleven to fifteen with the help of the following formula:

$$\Gamma n = \sqrt{\frac{2}{\pi}} \pi n^n e^{-n} \quad 15$$

6. From statistical tables find:

(i) Ordinate of a normal curve $y = \frac{1}{\sqrt{0.7}} e^{-\frac{x^2}{0.7}}$ for $x = -0.23$ and $x = 0.31$, and

find the area under this curve between these two values of x .

(ii) Upper and lower 5% values of χ^2 for 21 d.f.

(iii) Upper 5% value of $z = \frac{1}{2} \log_e F$ for $n_1 = 8$ and $n_2 = 17$. 10

COMPUTER'S CERTIFICATE EXAMINATION, 1951.

PART IC: SECTION II

Time : 4 hours

Full marks 100

- (a) Answers to the different groups are to be given in separate books.
 (b) Use of Calculating machines is permitted.
 (c) Figures in the margin indicate full marks.

GROUP A

1. Below is given the frequency distribution of 1,000 tests of fibre-strength of a variety of Surat cotton:

Strength-class (grams)	Frequency
0.0-0.9	38
1.0-1.9	165
2.0-2.9	188
3.0-3.9	159
4.0-4.9	137
5.0-5.9	114
6.0-6.9	81
7.0-7.9	48
8.0-8.9	29
9.0-9.9	19
10.0-10.9	15
11.0-11.9	6
12.0-12.9	0
13.0-13.9	11
Total	1,000

Ascertain the appropriate type of Pearsonian curve to be fitted to data. 15

2. (a) Interpolate the missing value in the following table

n = degrees of freedom	One per cent values of t -statistic
3	5.841
4	4.604
5	4.032
6	—
7	3.499
8	3.355
9	3.250

(b) The following table gives certain values of a function $f(n, \theta)$

θ n	0°	15°	30°	45°	60°	75°	90°
8	2.306	2.293	2.261	2.225	2.196	2.176	2.170
12	2.179	2.175	2.167	2.163	2.164	2.168	2.176

Find, by interpolation, the value of $f(n, \theta)$ for (i) $n=11$, and $\theta = 49^\circ 47'$, (ii) $n=12$ and $\theta = 37^\circ 30'$.

3. Evaluate the determinant

$$R = \begin{vmatrix} 1 & 0.70 & 0.60 \\ 0.70 & 1 & 0.40 \\ 0.60 & 0.40 & 1 \end{vmatrix}$$

Find also the co-factors of the three unit elements and in each case evaluate the expression

$$\sqrt{1 - \frac{R}{\text{Co-factor}}}$$

GROUP B

4. Prepare a layout for a manurial experiment with 6 blocks and 5 treatment besides a control.

An area of ground was sown with peas and divided into twentyfive equal size plots. The plots received five different fertilizer treatments:—M, N, C, P, and K. The actual layout is shown below. The figures give yields in pounds per plot.

N	P	K	C	M
172	164	154	160	178
C	K	M	P	N
156	152	160	161	171
M	N	C	K	P
175	167	163	158	160
P	C	N	M	K
157	158	170	183	162
K	M	P	N	C
153	173	154	166	159

Apply the technique of analysis of variance to test if all the treatments can be regarded as equally effective.

5. The following figures show the net earnings (in units of 100,000 rupees) of Public Utility Concerns over a period of five years:

	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec
1926	669	618	607	595	549	557	492	499	569	609	658	703
1927	744	669	654	649	612	592	539	536	619	653	702	759
1928	790	743	728	689	677	675	623	618	682	737	814	910
1929	920	860	850	830	825	790	710	730	800	830	920	1000
1930	920	900	880	895	860	830	706	715	808	841	883	890

Calculate the index of seasonal variations in the net earnings.

20

6. For any five of the following items of information relating to India name at least one publication from which you will obtain the information:

- Weekly wholesale prices of oilseeds
- Monthly exports of pig iron
- Monthly production of sulphate of ammonia
- Monthly index number of wholesale prices of metal products
- Number of industrial disputes (Annual)
- Monthly collection of sales tax in Calcutta
- Passengers killed (Annual)

10

STATISTICIAN'S DIPLOMA EXAMINATION, 1951.
S PAPER I: THEORETICAL STATISTICS (GENERAL)

Time: 4 hours

Full marks 100

- Answers to the different groups are to be given in separate books.
- Figures in the margin indicate full marks.
- Use of Calculating machines is not permitted.

GROUP A

Attempt question No. 1 and any two of the rest from Group A.

1. Comment carefully on the truth or falsity of the following statements:

- A large number of stone chips (C_1) of fairly uniform size (as indicated by the coefficient of variation) are mixed with a smaller number of chips (C_2) which are not so uniform. Obviously the uniformity in the size of the chips in the resulting mixture cannot be less than that of C_2 .
- It is reported that Yule's coefficient of association as well as the coefficient of colligation (in respect of the two attributes under study) are nearly equal to -1 . We may therefore safely conclude that the attributes occur together rather rarely.
- If a map with necessary details is available a simple way of selecting houses with equal probability is to perform the following operation for the requisite number of times. Throw a random point on the map and then select the house nearest to this point.
- A method of choosing a sample of villages (i.e. localities with a population of 5,000 or less) with probability proportional to size of population, without recourse to the method of 'cumulative totals', is to fix from a complete list of localities (i.e. villages and towns) showing population of each unit—a village at random and then to include it in the sample only if its population is not less than a randomly chosen number between 1 and 5,000. The whole operation is to be repeated until the requisite number of villages is included in the sample.

20

2. Explain clearly how the utility of the usual interpolation formulae depends upon the closeness with which the variable under consideration can be represented locally by a polynomial of fairly low degree. Illustrate your answer by means of the following problem in which attempts are to be made to estimate the value of P for (a) $x^2 = 0.5$, ($n' = 7$) and (b) $x^2 = 55$, ($n' = 29$) by interpolation from Table XII in "Tables for Statisticians and Biometricians, Part II". The relevant portions of the Table reproduced here for ready reference.

x^2	P for $n' = 7$	x^2	P for $n' = 7$	x^2	P for $n' = 29$
1	0.985612	6	0.423190	30	0.363218
2	0.910609	7	0.320847	40	0.066128
3	0.808847	8	0.238103	50	0.006467
4	0.676676	9	0.173578	60	0.000407
5	0.543813	10	0.124652	70	0.000019

16

3. An urn contains 2 white and 3 black balls, and a second urn contains 3 white, 4 black and 2 red balls. Two balls chosen at random from the first urn are transferred to the second, and then two balls similarly chosen from the second urn are returned to the first. Find the chance that the final and initial distributions of the balls according to colour in either urn are identical.

4. Three points are taken at random on a line of length l . Find the mean distance of the intermediate point from the middle of the line. Assume uniform distribution of probability.

5. State and prove the Limit Theorem concerning the convergence of the binomial to the normal distribution.

GROUP B

Answer an three questions.

6. Show that the sum of squares of n independent normal variables with mean zero and variance unity is the χ^2 -distribution with n d.f.

$$\text{Const. } e^{-\frac{1}{2} \chi^2} (\chi^2)^{\frac{n-2}{2}} d\chi^2$$

Also show that if Y and Z are two independent variables with χ^2 -distributions with m and n d.f. respectively then $Y+Z$ has a χ^2 -distribution with $(m+n)$ d.f.

7. Write a note on the mathematical set up of the randomised block and Latin square experiments explaining the assumptions clearly and the method of analysis.

8. Let M be the median for a sample of size n from a normal distribution with mean m and variance σ^2 . Calculate the large sample standard error of M and compare this with the standard error of the sample mean \bar{x} .

9. Write a note on the transformations used for the stabilisation of variance, when a set of random variables have

(i) binomial distribution with different values of p , where p is the probability of success.

(ii) Poisson distribution with different mean values m .

Discuss the utility of these methods in analysis of variance.

STATISTICIAN'S DIPLOMA EXAMINATION, 1951.

PAPER 2: APPLIED STATISTICS (General)

Time: 4 hours

Full marks 100

- (a) Answers to the different groups are to be given in separate books.
 (b) All questions carry equal marks.
 (c) Use of Calculating machines is not permitted.

Attempt any three questions from each group.

GROUP A

1. Discuss the principles of randomization and of local control with reference to agricultural field experiments and point out their salient features. How will you choose between randomized block and Latin square lay-outs for trying out a given number of treatments? Prepare a lay-out using Latin square for an experiment consisting of 5 treatments.

2. Discuss the advantages of a 'Factorial experiment'.

Explain the basis of the practice of confounding applicable to factorial experiments. Illustrate your answer from a $2 \times 2 \times 2$ experiment in which (i) the three-factor interaction is confounded with blocks, (ii) the two-factor interactions are partially confounded with blocks.

3. Give a critical account of the statistical methods that are usually adopted for studying the growth of populations, commenting on the adequacy or otherwise of the Indian Census data for this purpose.

4. Describe the two usual methods of standardization of crude mortality rates. Explain with reasons which method you would use for comparing the mortality experience of (a) industrial insurance companies, and (b) a small tuberculosis hospital.

5. For one and two traits (supposed for simplicity to depend respectively on one or two genes) write a short critical note on the nature of (a) autosomal, and (b) chromosomal inheritance, contrasting the two types; and for two traits contrasting further under each type (a) the case where the genes are located in the same autosome or chromosome with (b) the case where the genes are located in different autosomal or chromosomal pairs.

6. By crossing a doubly heterozygous individual (Aa Bb) with a doubly recessive one (aa bb) you have certain numbers in the different possible classes of offspring. Enumerate these classes and then discuss a suitable statistical technique by which, from the observed data, you can test independently the three different hypotheses

- (a) The character A is segregating in the 1:1 ratio.
- (b) The character B is segregating in the 1:1 ratio.
- (c) There is no linkage between A and B.

GROUP B

7. What does a Wholesale price index number attempt to measure? Prepare a scheme for constructing such an index for India and comment on the existing index numbers of this type in India, in the light of your proposed index.

8. What is meant by business cycle? Describe a method for obtaining the cyclic component of a series showing the number of persons employed in manufacturing industries, by months, for the period 1920-1950.

9. If the cyclic component of the Index of Industrial production is given for the same period and region, how will you find the relation between these two characteristics.

10. What do you mean by agricultural statistics of a country? In the light of your definition, examine the present agricultural statistics of India with regard to its exhaustiveness and suggest suitable methods for collecting information on those items which are found wanting.

11. What are cost and variance functions? How are they obtained? Indicate how you will use these functions for planning a sample survey for obtaining area cultivated under the major crop of your State.

12. Why is intelligence testing important in education? How does it differ from achievement testing? State the essential requisites of intelligence tests and describe in brief how you will proceed to construct a set of such tests to be given to a batch of school children.

STATISTICIAN'S DIPLOMA EXAMINATION, 1951.

S

PAPER 3: STATISTICAL INFERENCE.

Time: 4 hours

Full marks 100

- (a) Answers to the different groups are to be given in separate books.
- (b) All questions carry equal marks.
- (c) Use of Calculating machines is not permitted.

Attempt any two questions from each group.

GROUP A

1. Carefully compare and contrast the theory of "Fiducial Probability" of R.A. Fisher and the theory of "Confidence Interval" of J. Neyman. Try also to give a short historical sketch of each of the theories.

2. Develop the theory of "Testing of Hypothesis" of J. Neyman and E.S. Pearson against a single alternative Hypothesis. Define the "Power of a Test" and investigate how a "Most Powerful Test" can be obtained. Illustrate the above with the help of an example. Write also a short historical note on the development of the theory.

3. Fully discuss the problem of K samples from normal populations and the various hypotheses regarding the population means and standard deviations. Sketch the History of L_1 -test for testing the equality of the sample S.D.'s.

- 4. Write notes on any four of the following:
 - (i) Shortest Confidence Interval.
 - (ii) Joint Fiducial Probability Distribution.
 - (iii) Uniformly Most Powerful Test.
 - (iv) Errors of the First and Second kind.
 - (v) Simple and Composite Hypotheses.
 - (vi) Fisher-Behrens Test.

GROUP B

5. On arriving in a certain town I found the first car I met bearing the number 2674. Assuming that the register numbers of cars were serially assigned what can I say about the total number of cars in the town?

On meeting my friend in that town he told me that his car bears the number 3123. On the basis of these two observations how can I better my estimate.

I was told that a certain number of cars bearing the numbers serially from 1, 2, ... etc was damaged and made inoperative. If I can collect n random observations of the existing register numbers what estimation procedure would you suggest for estimating the effective number of cars.

- 6.(a) What is Bayo's theorem? What are the difficulties encountered in its practical use?
- (b) Show that minimum variance unbiased estimates are explicit functions of sufficient statistic.
- (c) Write a critical note on the criterion of minimum variance in estimation.
- 7.(a) What are the assumptions underlying the test of a null hypothesis concerning the multiple correlation?

(b) What is the effect of introducing a number of concomitant variables which have zero correlation with the variable to be predicted?

(c) Assuming that all variables are equally correlated find the limiting value of the multiple correlation of one variable on the others as the number of variables become large.

(d) How do you test on the basis of measurement of lengths of the right hand and left hand and right leg and left leg whether there is equal growth of the limbs on the right and left sides. Assume that 100 individuals have been measured.

STATISTICIAN'S DIPLOMA EXAMINATION, 1951.

S
PAPER 6: PRACTICAL

Time: 6 hours , Full marks 100

(a) Figures in the margin indicate full marks.

(b) Use of Calculating machines is permitted.

1. Given the following data:—

Values of x	310	320	330	340	350	360
Values of $\log x$	2.4914	2.5052	2.5185	2.5315	2.5441	2.5503

With the help of the above data find the value of $\log 3375$ 10

2. The following data relate to lives in hours of four batches of electric lamps. Test whether the batches differ among themselves in average length of life.

Batch 1:	1600,	1610,	1650,	1680,	1700,	1720,	1800.
Batch 2:	1580,	1640,	1640,	1700,	1750,		
Batch 3:	1460,	1550,	1600,	1620,	1640,	1660,	1740, 1820.
Batch 4:	1510,	1520,	1530,	1570,	1600,	1680.	

20

3. As a result of sampling studies, the distributions of expenditures of middle class families in Calcutta in the years 1939 and 1945 were found as follows:

Expenditure level (in Rs.)	1939	1945
0-50	150	60
51-100	229	348
101-150	204	361
151-200	179	272
201-250	119	192
251-300	94	134
301-400	93	155
401-600	73	128
601-700	63	56

Investigate whether there is any change in the pattern of expenditure of the middle class families. 25

4. An experiment has been conducted to verify the law of falling bodies under gravity expressed by $s = \frac{1}{2}gt^2$, where s is the distance fallen through during a period of time t and g is gravitational constant (32.2ft./sec²). The following results have been obtained.

t (seconds)	1	2	3	4	5
s (feet)	15	70	140	250	380

Using the above data estimate the constant g and test whether your estimated value differs significantly from the theoretical value as stated above.

5. The following table shows the records of 20 farms with varying incomes (x_1), dairy cows (x_2) and crop acreages (x_3). Calculate $R_{1,2}$, $r_{12,3}$ and test their significance.

Income (\$ per year) X_1	No. of dairy cows X_2	Crop acreage (acres) X_3
960	18	60
830	0	220
1260	14	180
610	6	80
590	1	120
900	9	100
820	6	170
880	12	110
860	7	160
760	2	230
1020	17	70
1080	15	120
960	7	240
700	0	160
800	12	90
1130	16	110
780	2	220
740	6	110
980	12	160
800	15	80

35

STATISTICIAN'S DIPLOMA EXAMINATION, 1951.

Time: 6 hours

PAPER 7: PRACTICAL

Full marks 100

- (a) Attempt question no. 1, and any four questions out of the rest.
 (b) Figures in the margin indicate full marks.
 (c) Use of Calculating machines is permitted.

1. *Either*

x	l_x	d_x	p_x	L_x	T_x
25	779727	3116	0.09600	778169
26	..	3190	0.09589	775016	31138339
27	773421	3288	771777	30363323
28	770133	3393	0.09559	768434	29591546
29	766735	3510	0.09542	..	28823112
30	763225	..	0.09522	761400	28058132

A portion of a life table is given above, the usual notations have been used.

- (i) Calculate the missing values.
 (ii) A company employing 500 men, all of whom enter at age 25 has reached a stationary condition. All the workers are transferred to other offices at age 30. Assuming that in the first year of service 25 per cent of those attaining age 26 are dismissed for inefficiency and 15 per cent resign at age 27 and there is no further resignations or dismissals, calculate with the help of the above table the number of (a) annual dismissals and (b) annual transfers.

Or

From the data given below, calculate the following:

- (i) The annual trend on the assumption it is linear.
 (ii) The seasonal index by centred 12-monthly moving average.
 & (iii) The cyclical irregular movements (per cent) from the deseasonalised data.

Year	Jan.	Feb.	March	April	May	June	July	Aug.	Sept	Oct.	Nov.	Dec.
1925	1780	1785	1805	1890	1930	1740	1415	1260	1460	1640	1715	1625
1926	1470	1590	1730	1955	1995	1840	1535	1410	1745	2225	2220	2215
1927	1885	2070	2300	2665	2565	2370	1935	1680	2050	2585	2610	2496
1928	2055	2300	2625	2915	2690	2545	1815	1635	2020	2500	2605	2555
1929	1900	2260	2540	2655	2570	2130	1860	1690	2235	2755	2920	2665

2. Prepare a control chart for the average of samples given below, the expected proportion to lie outside the inner limit being 10 per 100 and that to lie outside the outer limit being 1 per 100.

Sample number					
1	2	3	4	5	6
7.40	7.52	7.48	7.49	7.36	7.32
7.46	7.60	7.44	7.43	7.42	7.42
7.39	7.40	7.46	7.45	7.38	7.40
7.48	7.44	7.43	7.44	7.46	7.36
7.41	7.40	7.48	7.34	7.40	7.38
7.46	7.54	7.47	7.40	7.51	7.45
7.43	7.47	7.49	7.48	7.31	7.40
7.40	7.48	7.38	7.46	7.48	7.34

19

3. From the data given below calculate the weighted aggregate chain index-number of the Building material prices for the year 1945, with counterbalancing change in weights. (Base year is 1930)

Commodity	Weight	Price per unit					
		1930		1944		1945	
		Rs.	as.	Rs.	as.	Rs.	as.
Iron	76	29-	8	45-12	46-	4	
Steel	22	45-	0	72-	8	76- 0	
Gravel	276	9-	4	22-	0	22- 8	
Brick	27	13-12		36-	0	34- 4	
Cement	47	17-	8	20-	4	19-12	
Sand	128	22-	4	30-12		25- 0	
Timber	{	Burmah Teak	24	64-12	260-	0	..
		Assam Teak		..	184-12	206-	0

19

4. In a factorial experiment with N, P and K, the design and yield per plot of which is given below, one interaction has been confounded. Find out which interaction has been confounded. Test the significance of the effects of N and NP: Is the treatment combination NP better than the Control?

Control	PK	NK	NP	I(a)	Control	NP	PK	NK	III(a)
26	26	32	30		32	32	36	40	
N	K	NPK	P	I(b)	P	NPK	N	K	III(b)
32	30	36	28		30	46	40	36	
P	N	NPK	K	II(a)	NK	control	NP	PK	IV(a)
32	42	46	38		36	24	28	20	
NP	NK	control	PK	II(b)	NPK	N	K	P	IV(b)
34	44	30	34		32	30	28	26	

19

5. Five problems A, B, C, D and E have been solved by 16%, 30%, 46%, 60% and 75% respectively of a large unselected number. If the zero point of ability in the test is taken to be -3σ , what is the σ -value of each problem as measured from this point? What percentage of this group will be able to solve the problems F and G respectively, where the problem F is as much more difficult than the problem B as A more difficulty than C and the problem G is as much easier than E as the problem D easier than B.

(Assume normality of distribution)

6. The following table has been extracted from the First Report of the National Income Committee:—

Net Domestic Product of the Indian Union by character of Enterprise: 1948-49

Items	Net output (Rs. 100 crores)	Items	Net output (Rs. 100 crores)	Items	Net output (Rs. 100 crores)
(1)	(2)	(1)	(2)	(1)	(2)
<i>Small enterprise (largely household)</i>		<i>Large enterprise</i>		<i>Unclassified items</i>	
1. Agriculture (other than plantations etc.)	40.0	7. Agriculture (plantations etc.)	0.7	15. Other commerce & transport	
2. Fishery	0.2	8. Forestry	0.6	16. Government services (administrative)	
3. small enterprise & hand trades	8.6	9. Mining	0.6	17. House property	
4. Professions & liberal arts	3.2	10. Factory establishments	5.8	18. Total of unclassified items	
5. Domestic service	1.5	11. Railways	2.0	19. Net domestic product	
		12. Communications	0.3		
		13. Organised banking & insurance	0.5		
6. Total of small enterprise	53.5	14. Total of larger enterprise	10.5		

Complete the table by showing the relative importance of the different sectors of the national economy.

Write a brief report based on the above table, bringing out the salient features of the Indian economy with particular reference to the industrial and agricultural development in India.

STATISTICIAN'S DIPLOMA EXAMINATION, 1951.

PAPERS 4 & 5 (SPECIAL SUBJECTS): ECONOMIC STATISTICS

Time: 4 hours

Full marks 100

- (a) Attempt any four questions.
 (b) All questions carry equal marks.

1. What statistical data are available in India for studying the course of inflation? State their sources and explain how you would make use of them.

2. A general index number of wholesale prices covering raw materials, semi-manufactured articles and manufactured articles can seldom be of any real use. Discuss this with special, but not exclusive, reference to the Indian wholesale price index number (Year ended August, 1939 = 100).

3. (a) Explain the difference between an index number of business activity and an index number of production.

(b) If p_1, p_2 are the prices and q_1, q_2 are the corresponding quantities in two years, then the index number of quantities for the second year with reference to the first year may be written in either of the following forms:

$$J = \frac{\sum p_1 q_2}{\sum p_1 q_1}; J' = \frac{\sum p_2 q_1}{\sum p_2 q_2}$$

In what circumstances is $J > J'$?

Compute J and J' and illustrate your answer from the figures below:

Commodity	First Year		Second Year	
	Quantity	Value	Quantity	Value
A	100	Rs. 300	200	Rs. 300
B	150	Rs. 300	150	Rs. 300
C	200	Rs. 200	100	Rs. 150

4. Consider what information is necessary for measuring either "income elasticity of demand" or "elasticity of substitution" in the simplest cases.

5. What special difficulties arise in applying the least square method to time series data? How can they be overcome?

6. Answer any three of the following:—

(a) Why are the usual tests of significance not valid in the analysis of most economic data?

(b) Why do we sometimes get nonsense correlations in time series?

(c) What is a "Gallup poll"? Why are its forecasts not always correct?

(d) What is "the link relative method" as used in analysing time series? To what extent is it valid?

7. Comment on the following table of net outputs per engaged person (Table No. 4.1 in the First Report of the National Income Committee, April, 1951):—

Items	Net output per engaged person
Agriculture	Rs. 500
Mining and Factory Establishments	Rs. 1,900
Small Enterprises	Rs. 600
Railways and Communications	Rs. 2,100
Banking, Insurance etc.	Rs. 1,600
Professions and Liberal Arts	Rs. 700
Government Services (Administration)	Rs. 1,300
Net Domestic Product (At factor cost)	Rs. 730

8. Explain the debit and credit items that make up India's Balance of Payment. Why do figures for merchandise trade given in the Reserve Bank statement for this balance differ from those published by the Director General, Commercial Intelligence and Statistics?

STATISTICIAN'S DIPLOMA EXAMINATION, 1951.

PAPERS 4 & 5 (SPECIAL SUBJECT): PROBIT ANALYSIS.

Time: 4 hours

Full marks 100

Answer question 1 and any two of the rest.

1. Write an essay on the validity and meaning of results of bio-assays stating clearly the assumptions regarding the aspects

(i) chemical or biochemical about the nature of the test preparation whose potency has to be assayed and that of the standard preparation with which it is to be compared,

(ii) biological or pharmacological about the response evoked in the experimental animals by the stimulus of the doses administered,

(iii) statistical about the computational processes, and

(iv) philosophical—if not metaphysical—about the relation between the theoretical abstractions of pure mathematics and the realities of bio-assay.

2. An experiment to study the relative potencies of two drugs was designed to yield the following information.

First drug			Second drug	
Dosage	no. exposed	no. killed	no. exposed	no. killed
d_1	n_1	n_{1k}	n_1'	n_{1k}'
d_2	n_2	n_{2k}	n_2'	n_{2k}'
..
..

Explain the method of comparison (a) without making any assumption about the dose-response curve and (b) only assuming the response curve to be linear and (c) the tolerance distribution is normal.

3. Write explanatory notes on the following:

(a) Adjustment for natural mortality in the estimation of the parameters of the tolerance distribution.

(b) Types of joint actions of drugs, independent, similar and synergistic, the definition and suitable tests.

4. Explain briefly the Karber, the extreme effective dose, Reed—Muench and the moving average methods of estimating the median effective dose and examine in particular the bias and efficiency of the estimate of the mean of a normal tolerance distribution by the Karber method.

STATISTICIAN'S DIPLOMA EXAMINATION, 1951.

PAPERS 4 & 5 (SPECIAL SUBJECT): VITAL STATISTICS AND POPULATION STUDIES.

Time: 4 hours

Full marks 100

- (a) Attempt any 3 questions but all of them should not belong to the same group.
 (b) All questions carry equal marks.

GROUP A

1.(a) Reproduce with your own comments the substance of the observations and recommendations made by the Population Data Committee of India on the collection and publication of Indian Vital Statistics.

(b) Explain how you will proceed to obtain a dimensional idea of the extent of under-registration in births and deaths in any Indian region on the basis of existing material in possession of Government and other public bodies.

(c) Give a methodological account of sample surveys in the determination of the extent of under-registration of births and deaths.

2.(a) Describe various methods of comparing different mortality experience pointing out their respective advantages and limitations. Which mortality index are particularly suitable for use when a large number of sets of rates has to be compared as for example in the study of occupational mortality?

(b) How will you compare the mortalities of two groups of population by a standardization process when the following data are available:-

(i) deaths by age of the two populations combined, but not for each of the component populations;

(ii) total number of deaths of each component population, not distributed by age; and (iii) the component populations distributed by age. Is it necessarily true that the standardized death-rate of the combined population always falls between the standardized rates of its components, as it would do in the case of crude rates?

(c) Discuss the theoretical and practical difficulties in the measurement of mortality of population classified by (i) marital status, and (ii) occupation.

3.(a) Describe the nature of misstatement in the age returns of the Indian Census. Which in your opinion is likely to be most effective way of minimizing such misstatement: asking (i) age last birthday, (ii) age next birthday, (iii) age nearest birthday or (iv) age in years and months? Do you expect substantial differences in the outcomes of these different methods?

(b) Explain the systems of grouping recommended by the actuaries associated with the successive censuses so as to eliminate to some extent such misstatement of age. Discuss in particular the reasons which led Vaidyanathan to favour the '2-6' method of grouping. What method would you follow to derive the age distribution in standard quinquennial age groups, 0-4, 5-9 etc, from that obtained by the '2-6' method of grouping?

(c) Give an outline of the method followed in the construction of life tables in the 1931 Census.

GROUP B

4.(a) Show that a population with an arbitrary age distribution subject to a constant system of age-specific fertility and mortality rates ultimately tends towards a stable age distribution. Briefly indicate how the rapidity of the approach to the stable form may be studied.

(b) Show that the fundamental integral equation of Lotka's system, viz.

$$\int_0^{\infty} e^{-rx} l_x b_x dx = 1,$$

can be expressed in the form

$$rK_1 - \frac{r^2}{2!} K_2 + \frac{r^3}{3!} K_3 - \dots = \log_e R_0,$$

where the symbols have their usual meaning.

5. Give a considered account of the criticisms which has recently been levelled against the use of gross and net reproduction rates as measures of replacement in any period of marked change. Outline any methods which in your opinion yield much more information on the problem of replacement.

6 Explain the method followed by the Royal Commission on Population of Great Britain in tackling their problem of population projection. State explicitly the assumptions made in course of this investigation and the reasons put forward by the Commission in support of these assumptions.

STATISTICIAN'S DIPLOMA EXAMINATION, 1951.

PAPERS 4&5 (SPECIAL SUBJECT): ANTHROPOMETRY.

Time: 4 hours

Full marks 100

(a) Attempt any five questions.

(b) All questions carry equal marks.

1. Trace the growth of 'anthropometry' as a science. What measures have been taken to standardise anthropometric measurements?

2. Compare the advantages and disadvantages of the C.R.L. and D² as competent tools for anthropological taxonomy.

3. 'If the characters on which measurements are available are insufficiently numerous or inadequately distributed over the entire human body to ensure approximately stable limiting values, D² as a means of classifying or clustering human groups loses much of its uniqueness.' Comment.

4. 'Caloric requirement is not the function of body weight, but of the surface Area of the body.' How would you estimate the 'Surface Area' of the body?

5. State the conditions under which a regression formula constructed from a series of observations from one population can be used for a prediction in another population.

Describe how the stature of prehistoric men can be predicted on the basis of their fossil remains.

6. Apply the method of maximum likelihood to estimate the lengths of right and left femora when n_1 skeletons provide the length of the right femora alone, n_2 the left femora alone and n provide both the measurements.

7. Suppose that records of age and measurements of some characters like stature, chest girth etc are available for a number of individuals. How do you determine the age at which these measurement attain stability?

8. How do you test whether an estimated standard deviation s_1 exceeds another independent estimate s_2 ?

It is said that women on the whole are less variable than men. Describe briefly the observations you need and the tests you employ to establish this.

STATISTICIAN'S DIPLOMA EXAMINATION, 1951.

PAPERS 4 & 5 (SPECIAL SUBJECT): PSYCHOLOGY AND EDUCATION

Time: 4 hours

Full marks 100

(a) Attempt any four questions.

(b) All questions carry equal marks.

1. Discuss in details the suitability of using the coefficient r as a measure of rank correlation.

Obtain the relation

$$U(n+1, \Sigma) = u(n, \Sigma-n) + u(n, \Sigma-n+2) \\ + u(n, \Sigma-n+4) + \dots + u(n, \Sigma+n-2) + U(n, \Sigma+n)$$

where Σ = actual score for any given ranking

$u(n, \Sigma)$ = number of values of Σ in the aggregate of $n!$ possible values obtained by correlating this with all possible rankings.

and hence show how the above relation can be utilised in finding out the distribution of τ .

As $n \rightarrow \infty$

prove that the distribution of τ tends to normality.

2. Starting from the fundamental equation of the multiple factor theory, explain the concept of the reliability of tests.

Let

$$C_1 = \sum_{i=1}^n b_i x_i,$$

where b_i are regression coefficients derived from a matrix by a principal axis method and x_1, x_2, \dots, x_n are the initial variables expressed as deviations from their respective means.

If x_2 and x_3 represent equivalent halves of x_1 , x_4 and x_5 represent equivalent halves of x_1 , etc. obtain an algebraic expression for the reliability coefficient.

3. In the multiple factor theory show that the estimation of factors depends on the solution of a matrix equation

$$F F' = R$$

where F = reduced factor matrix

R = reduced correlation matrix.

Show further that the interpretation of factors depends on the rotation of the axes of reference.

4. In experimental psychology unitary traits have been classified into surface traits and source traits. The former are revealed by a cluster of positively inter-correlating elements, in which each correlates with every other member of the group while the latter may be regarded as the basic independent influences behind the clusters.

Discuss the different statistical procedures which are adopted in the different situations cited above.

5. If Z_0 is an occupation, Z_1 battery of tests, the estimate of a candidate's occupational ability

$$Z_0 = r_0' R^{-1} Z$$

Where r_0' are the correlations of the occupation with the tests and R is the matrix of correlations.

Explain the above with the help of the following example.

Let us suppose we have an "occupation" Z_0 and four tests Z_1, Z_2, Z_3 and Z_4 . The tables of loadings of the factors and the three sets of regression coefficients are as follows:

	f_1	f_2	f_3		f_1	f_2	f_3	f_4
Z_0	.55	.45	.60	f_1	.300	.095	.532	.095
Z_1	.60	.52	.21	f_2	.353	.581	-.352	-.153
Z_2	.52	.60	—	f_3	.121	-.148	-.206	.747
Z_3	.74	—	—					
Z_4	.37	—	.71					

If the set of scores Z_1, Z_2, Z_3 and Z_4 for a man be .2, -.4, .7 and .6, estimate his occupational ability. Without using factors at all we arrive at the following equation.

$$\hat{Z}_0 = .300Z_1 + .222Z_2 + .018Z_3 + .431Z_4$$

With the help of the above illustration, discuss in details "what is the use of estimating a man's factor at all".

STATISTICIAN'S DIPLOMA EXAMINATION, 1951.

PAPERS 1&5 (SPECIAL SUBJECT): MATHEMATICAL THEORY OF SAMPLING DISTRIBUTIONS

Time: 4 hours

Full marks 100

- (a) Attempt any four questions.
 (b) All questions carry equal marks.

1. Show that for a given distribution function $F(x)$ there always exists a characteristic function $\phi_x(t)$. Is the converse always true?

Given that x is a discrete variable which can take only positive integral values and that

$$\phi_x(t) = \text{Exp.} [-a \{1 - e^{-b(1-e^{it})}\}], a > 0, b > 0,$$

then show that x follows Neyman's Contagious Distribution

$$p(x) = \frac{e^{-a} b^x}{x!} \sum_{r=0}^{\infty} \frac{a^r r^x e^{-rb}}{r!}$$

Find the distribution function corresponding to the characteristic function

$$\phi_x(t) = e^{\sqrt{1-it} - 1}, x \text{ being a continuous variable.}$$

2. Suppose x_1, x_2, \dots, x_n are n values of a stochastic variable following a probability law $p(x)$, explain how you would obtain the conditional probability law of x function of x 's subject to a linear constraint on them.

If the x 's are normal with mean zero and standard deviation unity, obtain, using a suitable variate transformation, the distribution of

$\sum_{i=1}^n x_i^2$ when the x 's are subject to $K (< n)$ constraints given by

$$\sum_{i=1}^n C_{ij} x_i = P_j, j=1, 2, \dots, K.$$

Prove that, in a sample drawn from an uncorrelated bivariate normal population (x, y) , the sum of squares of deviations of y from the line of regression of y on x is distributed independently of the sum of squares due to regression and that they are distributed respectively as $\sigma_y^2 \chi^2$ with $(n-2)$ and 1 degrees of freedom. Derive the distribution of the coefficient of regression b_{yx}

3. Random samples of n observations are drawn from a distribution $f(x)$ with

finite range (a, b) and with probability integral $F(x) = \int_a^x f(x) dx$.

Let the observations be arranged in ascending order of magnitude, thus x_1, x_2, \dots, x_n . Write down expressions for the frequency functions of (1) x_m and (2) the range $w = x_n - x_1$.

If the distribution of x is rectangular, obtain the sampling distributions of (1) the median and (2) the range.

4. Let (x_1, x_2, \dots, x_n) be a sample from a normal population with mean zero. Obtain geometrically the distribution of the ratio of the sample mean to the sample standard deviation. Hence derive the distribution of

$$t = \frac{\bar{x} \sqrt{n}}{s},$$

$$\text{where } s^2 = \frac{1}{(n-1)} \sum_{i=1}^n (x_i - \bar{x})^2,$$

If the population mean is $\mu \neq 0$, write down the distribution function of $(\bar{x} \sqrt{n})/\sigma$.

To test whether the means of two populations are equal, the following criterion is suggested

$$u = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\left\{ \frac{\Sigma_1 + \Sigma_2}{n_1 + n_2 - 1} \left(\frac{1}{n_1} + \frac{1}{n_2} \right) \right\}}}$$

where \bar{x}_1 and Σ_1 are the mean and the sum of squares about the mean in a sample of n_1 from the first population and \bar{x}_2 , Σ_2 and n_2 refer similarly to a sample from the second population. State under what conditions is it appropriate to refer u to the t -distribution.

If these conditions are not satisfied, propose an alternative criterion to test that the means are equal.

5. Two variables x, y are distributed normally with means ξ, ζ , and variances σ_x^2, σ_y^2 . Suppose (x_1, x_2, \dots, x_n) and (y_1, y_2, \dots, y_n) are two sets of observed values of x and y . Given that x and y are independent, obtain the distribution of the ratio of the sample variances $\frac{s_x^2}{s_y^2}$ under the hypothesis H_0 ($\sigma_x^2 = \sigma_y^2$) and derive the test for H_1 .

(Assume the distributions of s_x^2 and s_y^2).

If, on the other hand, x and y are known to be correlated, find the likelihood ratio λ to test H_0 . By writing $X^2 = \frac{1}{2}(x_1 + y_1)^2 + \frac{1}{2}(x_1 - y_1)^2$ Show that λ can be put into the form:

$$(1 - r^2)^{\frac{n}{2}}$$

r being the sample correlation coefficient of X, Y . Hence derive the distribution of $\lambda^{1/n}$

6. N observations fall into K classes, the frequency in the i th class being n_i . Suppose m_i is the expected frequency under a simple hypothesis. (Here $\Sigma n_i = \Sigma m_i = N$). Obtain the distribution of the statistic

$$\phi^2 = \sum \frac{(n_i - m_i)^2}{m_i}$$

stating clearly the assumptions made.

If, however, the hypothesis is a composite one with $r < (K-1)$ unspecified parameters in the population law and m_i 's are the expected frequencies estimated from the sample, explain how the distribution of ϕ^2 is modified.

In two infinite populations the proportions of individuals possessing a certain character A are p_1 and p_2 . Two samples of size, m, n are drawn from the two populations and the observations are shown in a 2×2 table.

	With A	Without A	Total
Sample I	a	c	m
Sample II	b	d	n
Total	r	s	o

Write down the probability of obtaining the doublet (a, b) and derive the conditional distribution $p(a/r)$. Using the normal approximation to this conditional distribution of a , show that the test of the hypothesis $p_1 = p_2$ is practically equivalent to the usual χ^2 -test.

7. Show that a partial correlation coefficient of the s th order in samples of size n is distributed as a total correlation coefficient in samples of size $(n-s)$.

Prove that the distribution of the sample multiple correlation coefficient R when the population value is zero, is in the form of the Beta-distribution.

8. Obtain Wishart's Distribution of second order moments in samples of size n from a normal K -variate population.

If a_{ij} is a second order moment, find the distribution of the roots l of the determinantal equation.

$$\begin{vmatrix} a_{11}-l & a_{12} & \dots & a_{1k} \\ a_{21} & a_{22}-l & \dots & a_{2k} \\ \dots & \dots & \dots & \dots \\ a_{k1} & a_{k2} & \dots & a_{kk}-l \end{vmatrix} = 0$$

STATISTICIAN'S DIPLOMA EXAMINATION, 1951.

PAPER 4 & 5 (SPECIAL SUBJECT): STATISTICAL QUALITY CONTROL

Time. 4 hours

Full marks 100

- (a) Attempt any four questions.
 (b) All questions carry equal marks.

1. Explain the use of control chart for variables in industry. Give statistical reasons why it works. What is the difference between "specification limits" and "control limits"? What action should be taken if specification limits are tighter than control limits?

Explain the theory of runs. How would you modify it in the case of control chart moving averages?

2. Write down the distribution of "Range". For the normal case, derive the expression for the mean and variance of the range. Explain the use of range in control charts.

3. What are the advantages of the sequential test procedure over other sampling schemes? Work out a sequential test procedure for two alternate hypotheses H_1 and H_2 , such that under H_i ($i=1,2$), X is normally distributed with mean μ_i ($i=1,2$) and unit standard deviation. Show that the procedure reduces to counting algebraic sum of scores.

4. 1000 lots with 50 articles in each lot are submitted for inspection. The process is on an average 4% defective. The acceptance procedure adopted is as follows:

$$N=50, n=5, C=0.$$

What is the quality of (1) the accepted lots and (2) the rejected lots after the observed defective pieces have been removed. Give your comments on the above acceptance procedure.

5. If σ is the population standard deviation and \bar{x} the expected value of sample s.d. = $\left\{ \frac{1}{n} \sum (x_i - \bar{x})^2 \right\}^{\frac{1}{2}}$ for samples of size n , obtain the expression $\bar{\sigma} = C_2 \sigma$

$$C_2 = \sqrt{\frac{2}{n}} \frac{\left(\frac{n-2}{2}\right)!}{\left(\frac{n-3}{2}\right)!}$$

Explain the importance of the factor c_2 in control charts.

How will you modify the control chart as a basis for action when you know that the machine tool wear causes a known trend?

6. Write a short note on Dodge and Roming sampling inspection schemes and Simon charts. How do these inspection schemes help in maintaining statistical quality control?

STATISTICIAN'S DIPLOMA EXAMINATION, 1951.

PAPERS 4 & 5 (SPECIAL SUBJECT): SAMPLE SURVEYS (Applied).

Time: 4 hours

Full marks 100

- (a) Attempt any five questions only.
 (b) All questions carry equal marks.

1. Discuss how will you plan a pilot survey for collecting relevant information regarding cost and variance factors, which information will be used later in planning a large-scale sampling survey for estimating the yield-rate of wheat in a province.

2. Trace the history of different sampling methods (both subjective and objective) used for estimating crop-acreage and yield in India.

3. Describe the design, field-organization of the checking procedures and scrupulous attention to detail at all stages. Moreover, a reliable checking system can only be devised by careful study of the types of error that are likely to remain undetected"

4. Describe and discuss the "Master Sample of Agriculture" of the United States of America.

5. Describe the uses of Hollerith or Powers-Samas cards (punched cards) in the analysis of data accumulated in the course of a large-scale survey.

6. Discuss the following statement:

"The attainment of a high standard of accuracy in computational work is extremely difficult, and demands most careful organization of the checking procedures and scrupulous attention to detail at all stages. Moreover, a reliable checking system can only be devised by careful study of the types of error that are likely to remain undetected"

—F. Yates: "Sampling Methods for Censuses and Surveys".

7. Discuss the following statement:—

"In any survey there are many sources of error. A perfect survey is a myth. Some surveys are of course, better than others, but even the best surveys contain imperfections. It must not be supposed however, that all surveys are worthless because 'all have errors'. —W. E. Deming: "Some Theory of Sampling".

STATISTICIAN'S DIPLOMA EXAMINATION, 1951.

PAPERS 4 & 5 (SPECIAL SUBJECT): SAMPLE SURVEYS. (THEORY)

Time: 4 hours

Full marks 100

- (a) Attempt any five questions.
 (b) All questions carry equal marks.

1. Find the variance of the mean of a sample of n observations drawn from a population of N individuals.

If the cost function is given by $c = \sum_{i=1}^L c_i n_i$

where C_i is the cost of making an observation in the i th stratum and n_i is the number of observations included in the sample from the i th stratum containing N_i observations and L denotes the number of strata, find the optimum allocation which minimises the variance of the mean for a given cost. Deduce optimum allocation due to Neyman.

Find the loss in efficiency when the stratification is resorted to in estimation only after selecting a random sample from all the strata put together.

2. Find the variance of the mean of a systematic sample of size n out of a population of size kn . Compare its efficiency with regard to unrestricted random sampling and stratified random sampling for a certain class of populations.

How will you modify the above if the population consists of $M = km$ clusters and m clusters are systematically selected and each selected cluster is completely enumerated.

Indicate how you will find the estimate of the population variance based on a single systematic sample assuming that there is a polynomial trend in the history of individuals.

3. What do you understand by interpenetrating checks (I.P.C.). Give the different ways with their suitability or otherwise in which they can be arranged. What are the advantages claimed for I.P.C. Examine critically its use in the conduct of sample surveys.

4. What is a master sample? Give details of the master sample developed for agricultural surveys in U.S.A. and the national sample proposed by Mahalanobis in India. What are the advantages or the drawbacks of such a sample?

5. Develop the formulae for the variance of the ratio and regression estimates. Compare the merits of the two.

Explain how you will use the regression technique to improve the estimate on the second occasion when the information on the earlier occasion is also given, the sample on the two occasions are only partially matched. Obtain also the variance of the revised estimate.

6. Develop the analysis of variance for a multi-stage sampling scheme for a crop cutting survey in which t_i tehsils are selected out of the i th district, n_{ij} villages out of the j th tehsil in the i th district in the sample, n_{ijk} fields out of the k th village in the j th tehsil of the i th district in the sample. There are L districts in the province. Assuming unrestricted random sampling at each stage as well as an infinite population give the expectations of different mean squares and an expression for the variance of the mean of the sample. What modifications will be necessary in the case of a finite population?

7. What are the considerations involved in the choice of a size of a sampling unit. Find the variance of the mean of n clusters of M unit each out of a population of N clusters in term of the inter-correlation coefficient ρ .

In case the cost function is given by $c = c_1 n M + c_2 \sqrt{n}$, where c_1 is the cost of enumerating a unit and c_2 is the cost of travelling unit distance between clusters, discuss the optimum allocation of M and n for a fixed cost.

8. Write short notes on:—

- (1) Multi-phase sampling
- (2) Sampling vs. Complete enumeration.
- (3) Line and area sampling.

STATISTICIAN'S DIPLOMA EXAMINATION, 1931.

PAPERS 4 & 5 (SPECIAL SUBJECT): Design of Experiments

Time: 4 hours

Full marks 100

- (a) Attempt any five questions.
- (b) All questions carry equal marks.

1. A design is formed in the following manner:

To each block of a balanced incomplete block design, with parameters v, b, r, k, λ a control treatment (the same for each block) is added, so that each block has now $(K+1)$ plots. Show that the ratio of the sampling variance of the comparison between the control and one of the original treatments to the sampling variance of the comparison between two of the original treatments is

$$\frac{\lambda + r}{2r}$$

Comment on the case $\lambda = r$

2. What do you understand by the term "confounding" in factorial designs? What are the main considerations involved in resorting to confounding in factorial designs, and what are its advantages and disadvantages? What are the practical considerations underlying the choice of degrees of freedom to be confounded?

'a', 'b', 'c', 'd' and 'e' are five factors at two levels each. In a confounded design with these factors, if (1) $ad, ae, bc, de, abcd, abce, bcde$ are the treatment combinations constituting the principal block, what are the contents of the other blocks? What effects would be confounded in this design? Is a better design possible with blocks of the same size; and if so, what will be the degrees of freedom confounded in one such design?

3. What are the practical situations under which it becomes necessary to resort to fractional replication in a factorial design? Why is the device limited in its usefulness and what precautions have to be taken to ensure that an experiment with fractional replication does not become valueless from the practical point of view?

In the half-replicate of a 2^6 factorial design (with the six factors 'a', 'b', 'c', 'd', 'e' and 'f' each at two levels) arranged in blocks of eight units, the contents of the principal block are (1), $ab, acde, bcde, ef, abef, acdf, bcdef$, the defining contrast being $A B C D E F$. Write down the contents of the other three blocks and state what interactions are confounded in this design. Also, give the structure of analysis of variance in this case.

4. A field experiment has to be conducted with three levels of irrigation (i_1, i_2, i_3), three cultivations (c_1, c_2, c_3), three levels of nitrogen (n_1, n_2, n_3) and three levels of super phosphato (p_1, p_2, p_3). Suggest at least three different types of designs for the layout of this experiment and give the structure of analysis of variance of each design.

Also discuss the relative merits of the different designs, stating clearly under what circumstances each would be an eminently suitable design to adopt.

5.(a) Prove that for a balanced incomplete block design with parameters r, k, λ , the inequality $b > r + r - k$ holds. Is this inequality more stringent than the Fisherian inequality $b > r$?

Also establish that for $b = r, k = r$, we have $b > K + 1 > \lambda + 2$

(b) Twenty one experimental plants have each five leaves growing serially along the stem. Show how to allocate 21 treatments to the leaves, so that each pair of treatments occurs once on the same plant, and each treatment occurs once on a first leaf, and once on a leaf in the other ordinal positions. Give the design and state what special name is given to it.

6. Give brief critical comments on each of the following statements (which are imaginary extracts either from statistical text books or from reports on field experiments), and suggest any corrections or alterations or qualifications you think desirable:-

(i) "If the produce from a single plot in a randomized block experiment is lost by accident, it may be estimated from the formula

$$x = \frac{rB + tT - G}{(r-1)(t-1)} . "$$

(ii) "In a randomized block experiment, the adoption of randomization increases its precision."

(iii) "In a 4×3 factorial experiment in four replications, for comparing the efficacy of four qualities of nitrogen (s, m, e, u) in combination with the three quantities (0,1,2), the analysis of variance is

	D.F
Blocks	3
Treatments	11
Error	33
	47

(iv) "If the analysis of variance is to be applied to the results of a randomized block experiment, the yields of all the plots must be assumed to follow a normal distribution".

(v) "If in a 2^n factorial experiment, the treatment differences do not come out to be statistically significant, the degree of freedom for treatments should be split up into their unitary components for ascertaining which of these contrasts are statistically significant."

7. "The Latin square provides more opportunity than randomized blocks for the reduction of errors by skilful planning". Discuss, carefully bringing out the advantages of the Latin square design in comparison with the randomized block design.

Under what practical situations is the Latin square more advantageous than randomized blocks even when the plots form a continuous line? Give the layout plan for such a design when the number of treatments is five, rows being compact blocks of land. What do columns specify in this case?

STATISTICIAN'S DIPLOMA EXAMINATION, 1951.
PAPERS 4 & 5 (SPECIAL SUBJECT) ACTUARIAL STATISTICS.

Time: 4 hours

Full marks 100

- (a) Attempt any five questions.
(b) All questions carry equal marks.
(c) Use of Calculating machines is permitted.

1.(i) A fund is accumulating at interest in such a way that the force of interest at any time is proportionate to the square of the amount of the fund at that time. If the fund has increased from 1.0 to 1.2 in the last 5 years, find the amount of increase in the fund in the next 5 years.

(ii) Show that the present value of an annuity for n years under which the payment at the end of t years is equal to t^2 is given by

$$(1+i) \frac{d a_n}{d i} + (1+i)^2 \frac{d^2 a_n}{d i^2}$$

2.(i) Show from first principles that among m lives aged x , the most probable number of deaths within an year is the greatest integer in $(m+1)q_x$, and the expected number is $m q_x$.

(ii) In a stationary community, for which the annual death rate is 20.9 per 1000, 60% of the population is under age 38, and the average at death of those lives who die under that age is 11 years. Calculate

- (a) the percentage of the population who attain age 38 each year;
(b) the annual death rate among lives over age 38; and
(c) the expectation of life at birth.

3. Reduce to their simplest terms the following expressions:

(i) $[v \dot{a}_{x:n} - a_{x:n}] + [\dot{a}_{x:n} - a_{x:n}]$;

(ii) $[P_x(1+i)(P_{x+1}+1) - P_{x+1}] \div [P_x + d]$;

(iii) $\frac{P_x}{a_x} - \frac{C_x}{\Delta V_{x+1}}$.

4. State briefly the theoretical basis of formulae for summation graduation. Find the missing part of the following 17-term formula and evaluate its smoothing coefficient:

$$\frac{[5][7]}{35} \left\{ \frac{3}{5} [3] + 1 \right\}$$

Also analyze and criticize the following formula

$$\frac{[5][13]}{65} (-u_5 + u_1 + u_6 + u_{-1} - u_{-2})$$

5.(i) Write down the expression which will state the probability that at least 2 out of 4 lives now aged x will be living t years before the death of a life now aged $(x-t)$

(ii) Write down an expression in integral form for the combined probability that in the case of 4 lives aged 25, 20, 20, and 20, the life aged 25 will be the first to die, and at least 5 years will separate the first and the second deaths. Reduce, if possible, your expression to its simplest form.

6. A policy effected on a life aged x provides the following benefits:

(A) in the event of death occurring within n years of the outset,

- (i) Rs.500 at the end of the year of death, and
- (ii) Re. 100 per annum payable for the balance of the n -year period, the first payment being made at the end of the year of death, and the last at the end of the n -year period, together with
- (iii) Rs.500 at the end of the n -year period.

(B) in the event of the life surviving the first n years, Rs.1000 payable at the end of the year of death.

Write down the expression for the annual level premium payable throughout life.

Obtain also expressions for the net values of the policy after t years by (a) the prospective method, and (b) the retrospective method (t being less than n), and show that the two values will be equal if the policy be valued on the same mortality and interest bases as those on which the premium has been calculated.

7.(i) Establish the result.

$$I_{x+t}(V_x + P_x)(1+i) = I_{x+t+1} \times {}_{t+1}V_x + d_{x+t}$$

Explain the meaning of the above equation in simple words.

(ii) Calculate the values of nV_x and nV_{x+n} given that

$$2(1 - A_{x:n}) = 3(A_{x:n} - A_{x+n}) = 4(A_{x+n} - A_x)$$

STATISTICIAN'S DIPLOMA EXAMINATION, 1951.

PAPER 8: ECONOMIC STATISTICS (Practical)

Time: 3 hours

Full marks 50

- (a) Attempt any three questions.
- (b) All questions carry equal marks.
- (c) Use of Calculating machines is permitted.

1. Estimated age distributions of the average number of wage-earners on colliery books in the United Kingdom.

Age group	Number of wage-earners (thousands)		
	1931	1937	1945
14 and 15	31	28	12
16 and 17	51	43	22
18 and 19	51	37	50
20 — 25	129	107	99
26 — 30	112	98	52
31 — 35	99	101	76
36 — 40	84	90	89
41 — 45	75	73	86
46 — 50	89	84	65
51 — 55	59	54	57
56 — 60	48	43	48
61 — 64	24	24	29
65 and over	18	16	24
Total	850	778	709

What changes have taken place in the age composition of miners? Compute any averages and ratios which you think would help to a better understanding of this table.

2. The following data represent percentages of 840 women with x live births. Fit a Poisson distribution and discuss its adequacy.

x (live births)	0	1	2	3	4	5	Total
f (percentages)	44.2	36.9	15.0	2.8	0.9	0.2	100.0

3. The following table shows the number of individual incomes earned in a particular industry in different ranges of income:—

Range of incomes (x) Rs.	Number of incomes
150-500	13,175,000
500-1000	652,000
1000-2000	137,500
2000 and over	35,500

Assume that this distribution of incomes $f(x)$ is linked with the normal distribution

$$N(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$$

by the relationship

$$\int_{-\infty}^z N(z) dz = \int_{150}^x f(x) dx;$$

where

$$z = a \log (x-150) + b.$$

Obtain estimates for a and b from the data and find the number of incomes between Rs. 250 and Rs. 800

z	$N(z)$	z	$N(z)$
1.56	0.9406	0.50	0.6915
1.57	0.9418	0.55	0.7088
2.24	0.9875	0.60	0.7257
2.25	0.9878	0.65	0.7422
2.81	0.9975	0.70	0.7580

4. The table below gives monthly income and average expenditure on broad groups of 641 families of workers in Madras in 1935-36.

Compare the expenditure pattern of different classes of families and calculate the income elasticity of demand of the food group at an income level of Rs. 40.

Monthly income Rs.	No. of families	Monthly expenditure in rupees on					Total
		Food	Housing	Clothing	Fuel & lighting	Miscellaneous	
Below 20	47	10.40	2.98	0.87	1.70	2.28	18.23
20-30	167	14.75	3.10	1.21	2.04	5.09	26.19
30-40	198	18.58	3.66	1.50	2.34	7.74	33.82
40-50	118	22.07	4.07	1.89	2.73	12.43	43.79
50-60	69	25.09	5.33	2.31	2.93	15.90	52.16
60-70	20	30.11	6.20	2.91	3.49	18.08	61.69
Over 70	22	39.71	9.02	3.09	4.69	23.42	81.73

STATISTICIAN'S DIPLOMA EXAMINATION, 1951.

PAPER 8: PROBIT ANALYSIS (Practical)

Time: 3 hours

Full marks 50

(Tables to be supplied: Fisher and Yates)

1. The following data give x =dose, n =the number of animals used and r =the number of animals killed for two drugs at different dosages.

Drug A			Drug B		
x	n	r	x	n	r
1.01	50	44	1.70	48	48
.89	49	42	1.61	50	47
.71	46	24	1.48	49	47
.58	48	16	1.31	48	34
.41	50	6	1.00	48	18
			.71	49	16

Examine,

(a) Whether there is linear relationship between the probit of kill and log concentration in the case of drug A,

(b) whether the regression lines are parallel for both the drugs (regression of probit of kill on log concentration), and test whether LD 50 is significantly different in both cases.

STATISTICIAN'S DIPLOMA EXAMINATION, 1951.

PAPER 8: VITAL STATISTICS & POPULATION STUDIES (Practical)

Time: 3 hours

Full marks 50

(a) Use of Calculating Machines is permitted.

Assuming the applicability of the mortality experience as given in the Life table of Madras (1931 Census) to the 1941 population of Mysore estimate the number of female survivors in different quinquennial age-groups in 1946.

Assuming further the continuance of the "marriage habits" and the "family-building habits" (as far as those concepts can be elaborated consistent with the usability of Civil condition and Fertility Tables in the 1941 Census Report) as depicted in the 1941 Census Tables, indicate how to project the average annual number of births for the next three quinquennial periods (1941-46, 1946-51, 1951-56). Estimate the figure for the period 1941-46 only. State clearly any other assumptions which you have to make. (In view of the limited time at your disposal you are free to make any simplifying assumptions which may be necessary to save time in your computation).

You are provided with

- (1) 1931 Census Report of India containing the actuarial report
- (2) 1941 Mysore Census Report (Tables)

STATISTICIAN'S DIPLOMA EXAMINATION, 1931.

PAPER 8: EDUCATIONAL STATISTICS (Practical)

Time: 3 hours

Full marks 50.

Use of Calculating machines is permitted.

1. The following table gives Correlation Coefficients for a battery of ten tests

Test	1	2	3	4	5	6	7	8	9	10
1	1.00	.00	.00	.35	.61	.35	.61	.00	.00	.40
2	.00	1.00	.00	.01	.35	.00	.00	.35	.61	.40
3	.00	.00	1.00	.00	.00	.61	.35	.61	.35	.40
4	.35	.61	.00	1.00	.61	.18	.30	.30	.53	.55
5	.61	.35	.00	.61	1.00	.30	.53	.18	.30	.55
6	.35	.00	.61	.18	.30	1.00	.61	.53	.30	.55
7	.61	.00	.35	.30	.53	.61	1.00	.30	.18	.55
8	.00	.35	.61	.30	.18	.53	.30	1.00	.61	.55
9	.00	.61	.35	.53	.30	.30	.18	.61	1.00	.55
10	.40	.40	.40	.55	.55	.55	.55	.55	.55	1.00

- (i) Determine the number of common factors in the above set
 (ii) Estimate the common factor loadings for the ten tests.

Explain how these can be used for the interpretation of factors.

2. Suppose that two teachers have ranked a group of individuals on a five point scale. The result is given below

Ranking	T ₁	T ₂
1	10%	20%
2	15%	40%
3	50%	20%
4	20%	10%
5	5%	10%

Discuss how we can assign "weights" so as to make the rankings of two teachers comparable?

STATISTICIAN'S DIPLOMA EXAMINATION, 1931.

PAPER 8: SAMPLING DISTRIBUTIONS (Practical)

Time: 3 hours

Full marks 50

- (a) Attempt any two questions
 (b) All questions carry equal marks
 (c) Use of Calculating machines is permitted.

1. If a variate is distributed as a χ^2 with n degrees of freedom, it is known that

$$x = \frac{\chi^2 - n}{\sqrt{2n}}$$

is asymptotically normal $N(0,1)$. Hence the distribution function of χ^2 can be well represented by a Gram-Charlier Type A series. Thus, to $O(n^{-1/2})$,

$$\int_0^{\chi^2} p(\chi^2) d\chi^2 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\chi^2} e^{-\frac{1}{2}x^2} dx - \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\chi^2} \left[\frac{1}{6} \frac{k_3}{k_1^{3/2}} H_3 + \left(\frac{1}{24} \frac{k_2}{k_1^2} H_2 + \frac{1}{72} \frac{k_2^3}{k_1^3} H_1 \right) + \left(\frac{1}{120} \frac{k_3}{k_1^{5/2}} H_4 + \frac{1}{144} \frac{k_2 k_2}{k_1^{7/2}} H_3 + \frac{1}{1296} \frac{k_2^3}{k_1^{9/2}} H_2 \right) \right]$$

where k_1, k_2, \dots are the cumulants of χ^2 and H, H_1, \dots are the Hermite polynomials.

Using this expression, calculate the value of the distribution function for $n=8$ and $\chi^2=12$. Compare this with the true value.

$$[H_1 = x^2 - 1; H_2 = x^2 - 3x; H_3 = x^3 - 6x^2 + 3;]$$

$$H_4 = x^4 - 10x^3 + 15x^2; H_5 = x^5 - 15x^4 + 45x^3 - 15;$$

$$H_6 = x^6 - 28x^5 + 210x^4 - 420x^3 + 105]$$

2. If a sample of size n is drawn from each of k normal populations and it is desired to test the hypothesis that the variances are equal consider the statistic

$$L_1 = \frac{\left(\prod_{i=1}^k s_i^2 \right)^{1/k}}{\frac{1}{k} \sum_{i=1}^k s_i^2}$$

in which s_i^2 is the variance in the i th sample. The q th moment of L_1 is given by

$$\mu_q = k^q \frac{\Gamma\left(\frac{k(n-1)}{2}\right)}{\Gamma\left(\frac{k(n-1)}{2} + q\right)} \left\{ \frac{\Gamma\left(\frac{n-1}{2} + \frac{q}{k}\right)}{\Gamma\left(\frac{n-1}{2}\right)} \right\}^k$$

Fit a Type I curve to the distribution of L_1 for the case $n=5, k=5$, and obtain the approximate value of the 5% point of L_1 . (Use Thompson's percentage point tables of the Beta Distribution).

In five samples of five observations each, the means and variances are given below.

Mean	s_i^2
8.4	2.57
8.2	9.41
7.9	14.33
8.1	4.50
7.8	8.85

Calculate L_1 and test for the hypothesis that the variances are equal.

3. Two samples of sizes 20 and 15 are drawn from populations with variances σ_1^2 and σ_2^2 respectively. Write down the expression for the power function of the test for $H_0 (\sigma_1^2 = \sigma_2^2)$ against the alternatives $H_1 (\sigma_1^2 > \sigma_2^2)$, the significance level being taken at 0.05. Compute the values of power for

$$\frac{\sigma_1^2}{\sigma_2^2} = 2, 3, 5 \text{ and } 7 \text{ and construct the power curve,}$$

Use this to find how large $\frac{\sigma_1^2}{\sigma_2^2}$ should be before the odds are 4 to 1 in favour of establishing that $\sigma_1 > \sigma_2$ using the 5% level.

STATISTICIAN'S DIPLOMA EXAMINATION, 1951.

PAPER 8: STATISTICAL QUALITY CONTROL (Practical)

Time: 3 hours

Full marks 50

- (a) Answer all questions.
 (b) All questions carry equal marks.
 (c) Use of Calculating machines and of Control chart tables is permitted.

1. A manufacturer wished to control the quality of a type of electrical apparatus with respect to two adjustment characteristics at a level such that the fraction defective, $p' = .0020$. Following table gives observed values of "number of defectives" for this item found in samples drawn from successive lots. Plot control chart. Is the process under control at the specified value of p' ? If the value of p' were not specified would the process be under control?

Lot	Sample size n	Number of defectives	Fraction defective p
1	1550	5	0.0032
2	2000	2	0.0010
3	1550	0	0.0
4	780	3	0.0038
5	260	0	0.0
6	2000	15	0.0075
7	1550	7	0.0045
8	950	2	0.0021
9	950	5	0.0053
10	950	2	0.0021
11	35	0	0
12	330	3	0.0091
13	200	0	0
14	600	4	0.0067
15	1300	8	0.0062
16	780	4	0.0051
Total	15785	60	

2. In a manufacturing plant it was known that a trend existed as the result of tool wear. The dimension for the characteristic inspected was specified as 0.644 ± 0.004 in. Following table gives values of \bar{X} and R for subgroups of five components measured every half hour. Plot control charts.

Subgroup number	\bar{X}	R
1	0.6417	0.0011
2	0.6418	0.0016
3	0.6424	0.0010
4	0.6431	0.0015
5	0.6433	0.0009
6	0.6437	0.0010
7	0.6433	0.0014
8	0.6436	0.0004
9	0.6441	0.0006
10	0.6444	0.0011
11	0.6456	0.0009
12	0.6457	0.0007
13	0.6454	0.0009
Total	8.3681	0.0131

What precautions should be taken in the collection of data for control charts when trends are known to exist.

3. For two alternate hypotheses given by

$$H_1: p_1 = 0.005$$

$$H_2: p_2 = 0.05$$

and

$$\alpha \text{ (producer's risk)} = 0.05$$

$$\beta \text{ (consumer's risk)} = 0.10$$

determine the equations of the acceptance and rejection lines.

Samples are taken in groups of 20. Write down the acceptance and rejection numbers for $N = 20, 40, 60, 80, 100, 120$.

4. A consumer receives batches containing 10,000 components each. Generally, 9 out of 10 batches are free from defects and 1 out of 10 batches contains 20% defectives. If 20% defective batch is accepted the loss incurred by putting the defective components on production line is Rs. 3000 for 10 batches. If the inspection line is half anna per item, find the number of items inspected per batch to minimise the total of inspection costs and consumer costs. It is assumed that acceptance number is 0.



STATISTICIAN'S DIPLOMA EXAMINATION, 1951.

PAPER 8: SAMPLE SURVEYS, APPLIED (Practical)

Time: 3 hours

Full marks 50

- (a) Figures in the margin indicate full marks.
(b) Use of Calculating machines is permitted

1. The formula for the sampling variance, V , of the estimate in the "double sampling procedure is approximately given by

$$V = V_0 \left\{ \frac{\rho^2}{N} + \frac{(1-\rho^2)}{n} \right\}$$

where V_0 is the variance of y , the variate under enquiry, ρ is the correlation coefficient between y and x (x being the auxiliary variate), and n and N are respectively the sizes of the two samples in the first of which both x and y are measured and in the second only x . The cost, W , of such a survey is approximately given by

$$W = a + bn + cN,$$

where $a = \text{Rs. } 470$, $b = \text{Rs. } 24$ and $c = \text{Rs. } 2$ as. 10 only. It is known that $V_0 = 291.8$ (mils/acre)², and the total expenditure sanctioned for the survey is Rs. 10,000 only.

Work out the optimum values of n and N and the corresponding V for the following eight values of $\rho = 0.0, 0.5, 0.3, 0.1, -0.1, -0.3, -0.5, -0.9$. Also draw the (free-hand) (ρ, V) graph. 20

2. Using "random sampling numbers" select two random samples, each of size 5, from the following two populations respectively:— 10

(a) a normal population of mean $= -10.28$ degrees (centigrade) and standard deviation $= 2.87$ degrees (centigrade), and (b) a rectangular population with a constant probability density between the limits -10.7 degrees (centigrade) and $+9.8$ degrees (centigrade). 10

3. A circular region is determined in terms of the rectangular cartesian co-ordinates by the relation $x^2 + y^2 \leq 169$ sq. metres. With the help of "random sampling numbers" locate ten points randomly inside the region. (It will do if the positions of the selected points are written down correct to decimeters).

4. Draw up a questionnaire form for a family budget enquiry in a middle-class area in a big city, say, Calcutta or Bombay. (Special credit will be given for neatness, proper captions, spacing, etc. The form has to be drawn up in English only). 10

STATISTICIAN'S DIPLOMA EXAMINATION, 1951.

PAPER 8: SAMPLE SURVEYS—THEORY (Practical)

Time: 3 hours

Full marks 50

- (a) Attempt any two questions only.
- (b) All questions carry equal marks.
- (c) Use of Calculating machines is permitted.

1. Table 1 gives the number of villages and the area under wheat in each of the 80 patwari circles in Hapur Tahsil of the Meerut district and table 2 gives the analysis of variance on a village basis. It is required to estimate the total area under wheat in the tahsil using a patwari circle as the unit of sampling. We will assume that a sample of 20 patwari circles is to be selected. Calculate the sampling variance of the estimate of the total area under wheat for each of the following procedures of sampling and estimation:

- (a) equal probability, mean per cluster estimate
- (b) equal probability, total per cluster estimate.
- (c) probability proportional to the size of the circle, mean per cluster estimate.

Also calculate the variance for an equivalent sample with village as unit of sampling and compare the relative efficiencies of the three methods. Finite population correction should not be ignored.

TABLE 2.

Analysis of variance of areas under wheat in village of the Tahsil Hapur (bighas).

Source of variation	d.f.	S.S.	M.S.
Between circles	88	10924593	124123
Within circles between villages	210	9587900	45657
Total	298	20512592	68834

TABLE NO. 1. SHOWING THE NUMBER OF VILLAGES AND THE AREA UNDER WHEAT
IN PATWARI CIRCLE OF HAPUR.

Sl.no. of Circle	Number of villages	Area under Wheat (in bighas)	Sl. no. Circle	Number of villages	Area under Wheat (in bighas)
1	6	1502	46	2	667
2	5	1003	47	3	941
3	4	1691	48	1	710
4	5	271	49	1	387
5	4	458	50	10	3516
6	2	736	51	8	2002
7	4	1224	52	9	3622
8	2	996	53	2	1400
9	5	475	54	2	1584
10	1	34	55	3	830
11	3	1027	56	8	167
12	4	1393	57	3	622
13	3	692	58	2	691
14	1	524	59	5	273
15	1	602	60	2	781
16	3	1522	61	2	1101
17	4	2087	62	2	789
18	8	2474	63	2	601
19	2	461	64	3	628
20	4	846	65	4	1141
21	3	1036	66	1	1208
22	4	948	67	5	1633
23	4	1412	68	4	902
24	3	438	69	3	1280
25	5	2111	70	5	1299
26	2	977	71	7	1947
27	3	814	72	3	741
28	1	319	73	2	674
29	2	583	74	7	2554
30	4	1150	75	4	669
31	3	670	76	1	1137
32	2	499	77	2	652
33	4	714	78	1	51
34	4	1081	79	1	1265
35	1	389	80	8	1423
36	7	2675	81	2	794
37	3	868	82	1	1604
38	2	1412	83	3	1021
39	2	445	84	2	1764
40	5	706	85	6	2668
41	2	642	86	1	1076
42	4	2050	87	1	348
43	6	2530	88	4	1224
44	1	247	89	4	1490
45	2	421			
				299	98078

2. The following table gives the measured volumes of timber on 25 systematically located plots of $\frac{1}{10}$ acre, and eye estimates in the stands in which they occurred. The total area under timber is 5124 acres and the total volume of timber, from eye estimates of all these stands, is 6,110,000 Cu. ft. i.e. 1192 cubic feet per acre. Obtain the regression and the ratio estimates of the total volume of the timber on the basis of the data and compare their relative efficiencies.

TABLE. MEASURED VOLUME, y , ON 25 SAMPLE PLOTS, AND EYE ESTIMATE, x , OF CORRESPONDING STANDS (CUBIC FEET PER $\frac{1}{10}$ ACRE)

y	x	y	x	y	x
170	102	255	208	125	65
47	14	135	110	100	194
64	57	140	110	287	167
91	70	154	110	201	268
126	95	110	110	169	152
146	92	112	128	182	152
87	110	153	79	74	148
195	208	216	177	24	207
				255	167
				3684	3302

3. (a) The following tables give information about the different strata forming the population

strata	N_i	σ_i	$\sqrt{c_i}$	where c_i represents
I	2525	325	4	the cost per unit in the
II	3150	275	3	i th stratum.
III	4975	275	2	
IV	1050	250	1	

(i) Find an optimum allocation corresponding to a fixed cost, the cost func-

tion being given by $\sum_{i=1}^L c_i n_i$

(ii) If $c=1000$, find the relative efficiency of an unrestricted random sample of the whole population where size is determined by (i) when $c=1000$.

(b) If the response rate in a market survey is 25% and the variance within the non-response group is $\frac{2}{3}$ in the whole population, supposing it costs Rs.2 to contact a unit Rs.4 to enumerate and process information on that unit and Rs.8 to enumerate and process information on the unit in the non-response group. Determine the optimum values of n and μ , the proportion of units enumerated in the non-response group so that the mean of the population may be measured with variance equal to $3\sigma^2/100$

4. The following tables gives the analysis of variance over a number of years for the wheat yield per field (mbs. per acre).

	1944		1945		1946		1947		1948	
	d.f.	M.S.	d.f.	M.S.	d.f.	M.S.	d.f.	M.S.	d.f.	M.S.
Between tahsils [*]	4	66.5	6	318.4	4	79.4	5	82.3	4	206.8
Within tahsils between villages	11	38.9	12	27.1	7	62.2	19	52.7	14	65.3
Within villages between fields	-	-	15	22.8	8	31.2	11	24.2	8	12.1
Within field between plots	16	5.33	40	6.20	22	11.39	39	6.60	28	9.80

(i) Test whether the within field variation is the same from year to year.

(ii) Pool the analysis of variance for the five years and obtain the estimates of the different components of variation.

(iii) Find the variance of the unweighted mean of all the fields if 10 villages per tahsil and 3 fields per selected village with the same amount of sampling per field, on the average, also the same variance components as given by the pooled analysis of variance.

STATISTICIAN'S DIPLOMA EXAMINATION 1951

PAPER 8: DESIGN OF EXPERIMENTS (Practical)

Time: 3 hours

Full marks 50

- (a) Attempt any two questions.
- (b) All questions carry equal marks.
- (c) Use of Calculating machines is permitted.
- (d) Using of Fisher and Yates' Tables is allowed.

1. Below are given the plan and yields in lbs. per plot of a 6×6 Latin square experiment on the wheat crop carried out for testing the effect of sulphate of ammonia applied at five different times (1,2,3,4,5), 0 denoting control (or no sulphate of ammonia). One of the observations is missing. Analyse the data.

5	2	3	1	0	4
28.8	27.0	20.5	24.7	22.3	24.8
0	4	5	3	1	2
31.8	35.7	33.4	28.9	33.6	32.2
2	1	4	5	3	0
37.1	34.4	—	38.0	33.9	28.3
1	5	0	2	4	3
33.0	36.5	27.1	38.2	41.6	30.4
4	3	2	0	5	1
38.6	34.4	41.8	34.4	41.7	36.0
3	0	1	4	2	5
39.4	29.6	36.5	40.2	38.6	34.0

2. In a varietal trial with thirteen varieties (0), (1), (2), . . . , (12) of paddy, arranged in a balanced incomplete block design, the following is the lay-out plan with the yields of paddy in lbs. per plot.

Blocks	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII
(0)20	(1)31	(2)54	(3)44	(4)36	(5)58	(6)66	(7)34	(8)50	(9)54	(10)64	(11)73	(12)71	
(1)53	(2)49	(3)50	(4)38	(5)47	(6)70	(7)53	(8)36	(9)49	(10)36	(11)56	(12)43	(0)69	
(3)39	(4)51	(5)28	(6)22	(7)54	(8)45	(9)43	(10)42	(11)24	(12)38	(0)66	(1)42	(2)57	
(0)37	(10)48	(11)63	(12)29	(0)61	(1)74	(2)32	(3)58	(4)47	(5)28	(6)44	(7)55	(8)64	

Obtain the design got by omitting the first block and all the varieties which occur in the first block. Carry out an "intra-block" analysis of variance and calculate the standard error for comparing two varieties. What is the efficiency of the present lay-out in comparison with randomized blocks without recovery of interblock information?

3. The annexed table shows the plan and yields in lbs. per plot of a 2^2 field experiment on wheat, the treatments being all combinations of two levels of dung (0, d), two

levels of ammonium sulphate (0, n) two levels of superphosphate (0, p) and two levels of potash (0, h).

Block I				Block III			
p	h	d	np ^h	np ^h	d	p	dn ^h
45	55	53	38	43	42	39	34
dn ^h	dnp	dph	n	n	dnp	h	dph
41	48	55	42	47	52	50	44

Block II				Block IV			
dp	nh	dh	ph	nh	dp	(1)	np
50	44	43	51	43	52	57	39
dn ^{ph}	(1)	dn	np	ph	dh	dn ^{ph}	dn
44	53	41	50	56	52	54	42

Analyse the data and discuss the results of analysis.

4. In a randomized block experiment with 9 replications carried out for testing the difference between the yields of two varieties V_1 , V_2 of Jowar, the variety V_1 being replicated twice and the variety V_2 thrice within each block, the following are the yields per plot in lbs.

Blocks	V_1	V_2
I	10.8	13.0
	13.0	16.0
		14.5
II	14.1	14.0
	17.4	17.7
		17.6
III	9.9	11.2
	11.7	14.1
		12.6
IV	11.2	12.7
	11.0	14.1
		12.2
V	10.5	18.5
	13.5	14.5
		16.2
VI	12.3	13.7
	11.0	16.5
		13.3
VII	12.6	12.0
	10.7	14.0
		13.6
VIII	9.7	10.0
	9.5	12.2
		11.0
IX	14.4	10.5
	15.1	17.0
		16.0

Perform the analysis of variance and interpret the results of analysis.

STATISTICIAN'S DIPLOMA EXAMINATION, 1951.

PAPER 8: ACTUARIAL STATISTICS (PRACTICAL)

Time: 3 hours

Full marks 50.

- (a) Attempt any four questions.
(b) All questions carry equal marks.
(c) Use of Calculating Machines and of interest tables will be permitted.

1. A loan of Rs. 1,00,000 contracted on January 1, 1942 is to be repaid by ten equal annual payments commencing 30th June 1951, followed by an additional and final payment of Rs. 5,000 on 30th September 1960. All payments are to include both principal and interest which latter should work out to an effective rate of 5 per cent per annum throughout the period. Draw up a schedule showing to the nearest rupee each payment and the division of each payment into principal and interest.

2. In connection with a Resettlement programme, a Government made on August 15, 1947, a large number of small loans amounting in all to Rs. 1,00,00,000. The loans are repayable over 20 years by level annuities of principal and interest combined calculated at 4 per cent per annum, the first payment being made on August 15, 1948.

In order to raise the needed funds, the Government issues on August 15, 1947 a loan of Rs. one Crore, carrying interest of 3 per cent per annum payable half yearly. On 15th August each year there will be allotted as service to this loan a sum equal to the capital portion of the receipts due from the small borrowers' annuities on the same date, plus a level additional sum of Rs. X per annum. The balance of this service after paying interest on the outstanding bonds will be applied to the redemption of capital by annual drawings at par. What should the value of X be in order that the loan may be exactly redeemed by the drawing due on August 15, 1967?

What price per cent could a purchaser of a substantial portion of the loan at issue offer in order to secure a net yield of 2 per cent per annum convertible half-yearly after allowing for income tax at 4 annas in the Rupee?

3. 5 years ago, a 25-year capital, redemption assurance policy for Rs. 10,000 was issued subject to level annual premiums payable throughout. The policy holder now desires to shorten the term by 5 years towards which the assurance society suggests two alternative methods:

(i) the policy to run in its present form till 20 years from the outset when it shall be surrendered; the balance of Rs. 10,000 then required to be provided by a new policy which shall mature at that time by additional annual premiums commencing at the present time.

(ii) the policy to be surrendered now and the surrender value applied as a single premium for a new policy maturing in 15 years from now; the balance of Rs. 10,000 then required to be provided by a new policy which shall mature at that time by additional annual premiums commencing at the present time.

Calculate the future annual premiums payable by the policyholder under each method assuming.

- (a) that single and annual premiums are based on 3 per cent per annum effective, and are loaded 5% of the net premium, and
(b) that the office guarantees surrender values of 50% of the first office premium and 95% of subsequent office premiums all accumulated 2 per cent per annum effective.

Explain also why the premiums under the two alternatives should differ.

4. An Office is asked to quote for 100 one year temporary assurances of Rs. 1,000 each on lives aged 50, the death rate at that age of life being 9 per thousand. Assuming 3 per cent interest and assuming also that claims are payable at the end of the contract period, what is the minimum aggregate premium the office must charge if its risk of a loss is not to be more than 25% of its chance of profit? If on the contrary, 36,500 persons had assured each for Rs. 1,000 for a term of One day only, what, under the same conditions, would the premium be?

5. Calculate the net annual premium for a special assurance on a life aged 30 under which Rs. 2,000 will be payable on survival to the end of 20 years. In the event of death occurring within the 20 years, there will be payable at the end of the year of death the sum of Rs. 1,000 plus the net premium reserve which would be held at that time if the assured were still alive. The premium and reserves are to be calculated at 3% interest by a special mortality table in which

$$100q_{30} = (1.03)^x$$

6. What is meant by "Graduation with reference to a standard table"?

In the following table: E_x is the number exposed to risk of death, and θ_x denotes the number of deaths between age x and $x+1$. It has been decided to graduate the mortality ratios, q_x by means of a formula of the type

$q_x = aq'_{x-1} + b$, where q'_{x-1} of the Standard Table are as set forth in the appropriate column of the table below, Obtain the most appropriate values of a and b and compare the graduated and ungraduated rates of mortality at the ages 25, 35, 45, and 55.

age x	E_x	θ_x	$10^4 q'_{x-1}$	age x	E_x	θ_x	$10^4 q'_{x-1}$
25	1000	5	572	41	1800	14	885
26	1100	4	608	42	1800	14	910
27	1200	5	643	43	1700	10	937
28	1300	6	668	44	1600	13	969
29	1400	6	691	45	1600	12	1001
30	1500	8	707	46	1500	13	1038
31	1500	6	720	47	1500	14	1081
32	1500	8	732	48	1400	12	1122
33	1600	10	746	49	1300	11	1172
34	1600	7	759	50	1200	13	1224
35	1600	7	771	51	1100	8	1281
36	1700	12	787	52	1000	11	1345
37	1700	5	803	53	900	9	1415
38	1700	14	821	54	800	9	1490
39	1800	12	839	55	800	8	1572
40	1800	12	862				

FIELD SURVEY EXAMINATION, 1951

PART I(A)

Time: 3 hours

Full marks 100

Figures in the margin indicate full marks

1.(a) Mention the names of the subdivisions of any district with which you may be familiar and give also the number of police stations, unions and villages of that district

(b) Mention the sources from which you can obtain:

(i) List of households in a village.

(ii) Area of each plot shown on the cadastral survey map of a village.

15

2. Indicate broadly the route you will follow to reach the following places from Calcutta:

- (1) Balurghat (2) Bankura (3) Cooch Behar (4) Maldah (5) Suri

Or

Name the important rivers in your state and name the subdivisions touched by them. Mention the name of any important towns that may lie on these rivers. 15

3. Answer the following

(i) How do you calculate the 3rd class-fare by rail, from Howrah to Asansol, a distance of 132 miles?

(ii) What is the maximum quantity of free luggage a passenger by train is allowed to carry on his (a) Second class (b) Inter class (c) Third class ticket. Does it vary from Mail to other Passenger train

(iii) Mention any three commodities which the railway would not allow you to carry as personal luggage.

(iv) What is the charge for sending Rs. 27/- by money order. How would it differ if it is sent by Telegraph money order.

Or

Supply the following information with respect to your state

(i) total population

(ii) number of females per thousand of the population

(iii) density of the population

(iv) total area of the State

(v) percentage of forest area (reserved forests only) to total area

(Mention must be made of the year to which the figures in your answer refer) 15

4. Name the principal crops of your state, and mention their total acreage and yield for any recent year. Mention the seasons during which they are sown and harvested.

(The year must be mentioned in your answer) 15

5.(a) The incomes quoted below all exceed Rs.250

Rs. 270; 290; 285; 315; 420

Calculate the average of these excesses over Rs. 250. What again would be the average of these excesses over Rs. 300?

(b) Add 25% of 100 to 100; subtract also 25% of 100 from Rs. 125. Would you expect the second calculation to give Rs. 100/-? If not what percentage of 125 should you subtract from Rs. 125/- to get Rs. 100 only. 10

6. Evaluate

$$(a) \frac{.1 \times .1 \times .1}{.2 \times .2 \times .2} + \frac{.01 \times .01 \times .01}{.02 \times .02 \times .02}$$

$$(b) \frac{.1 \times .1 \times .1}{.2 \times .2 \times .2} + \frac{.01 \times .01 \times .01}{.02 \times .02 \times .02}$$

$$(c) (74.5 - .007) \times 0.35$$

$$(d) 5 \text{ annas of } 3.68 \text{ acres} + 11 \text{ annas of } 6.40 \text{ acres} + 16 \text{ annas of } 1.00 \text{ acres.}$$

$$(e) 2\sqrt{A^2 + B^2}$$

$$\text{when } A = 2.4$$

$$B = 1.8$$

Or

(a) A, B and C are three points. If the distances AB and BC are given, what further information would you require to find out the distance between A and C?

(b) ABCD is a field area, such that the portions ABC and BCD form two triangular plots with all their sides equal. If AB = 20ft., calculate the shortest distance from A to D

7. Answer any four of the following:

- (i) How much should 100 pieces of new rupee coins weigh in lbs.
- (ii) How many lbs go to make one maund?
- (iii) How many standard bighas make one square mile.
- (iv) How many square cubits make one cottah?
- (v) How many cubits make 1 metre?

12

N. B.—3 marks for neatness.

FIELD SURVEY EXAMINATION, 1951

PART I (B)

Time: 3 hours

Full marks 100

All questions carry equal marks

Answer any 5 questions

1. Write down the names of all the subdivisional head-quarters of a your State which are not directly served by Railway lines. Also mention the names of the Railway stations nearest to these subdivisional head-quarters. State how you will reach these subdivisional head-quarters from their nearest Railway stations.

Or

Give the boundaries of the following states

(i) Assam (ii) Madhya Bharat (iii) Madhya Pradesh (iv) PEPHU and (v) Uttar Pradesh.

Write down also the names of the principal rivers that flow through these States.

2. When is Jute generally sown in different parts of West Bengal? When is the Jute crop harvested? What are the normal seed requirements for Jute? What is the average yield of Jute per acre? Name the 5 most important Jute growing districts.

(N. B. Candidates not familiar with the growing of Jute in West Bengal are required to supply similar information about the most important cash crop of his state of domicile).

3. An Investigator draws a monthly salary of Rs. 100 plus Dearness Allowance Rs. 35/- and House allowance Rs. 10. He had borrowed Rs. 240 at 5 per cent simple interest from the employees cooperative fund on 1st January 1950, payable in 24 equal monthly instalments and deductible from his pay each month. He was granted extra ordinary leave for six days in June 1950 without pay. Calculate his net dues for the month of June.

4. For a middle class Bengali family consisting of 3 adult members (ages 52, 41 and 33 years) and 3 children (ages 10 yrs, 7 yrs, and a baby of 6 months only) and a servant of 24 years, estimate the monthly consumption of the following food articles: (a) Rice (cleaned), (b) Pulses, (c) Common Salt. Indicate how you arrive at the estimates.

5. Give the final forecast figures for any 3 of the following crops for your state, during the last 3 years.

- (a) Wheat (b) Aus Paddy (c) Jute (d) Sugarcane

Indicate the source of your information and mention the years for which you quote the figures.

6. What are the 5 major occupations of the people of your State? Indicate the relative importance of the different occupational groups in terms of percentages to the total population.

7. You are provided with details of information about a family which have been entered on a printed Economic Enquiry form. Scrutinize the entries appearing below carefully and offer your comments on items which appear to your mind as either inconsistent or incorrect.

I. SAMPLE DETAILS

District—Howrah P.S. Bagnan J.L. No. 36
 Name of the Head of family—Nityananda Chakravarty
 Religion—Hindu, Caste—Kayastha

II. DETAILS ABOUT THE FAMILY MEMBERS

Serial No.	Relationship with Head of the family	Age	Sex	Occupation	Monthly income from occupation
1	Self	45	M	Pleader	Rs. 250/-
2	Wife	50	F	—	—
3	Son	33	M	Teacher	Rs. 120/-
4	Wife	30	F	—	—
5	Son	12	F	—	—

III. ANNUAL INCOME AND EXPENDITURE OF THE FAMILY

Source of income	Annual income Rs.	Items of Expenditure	Expenditure in Rupees Monthly	Total
2. Agriculture	50	1. Food	138	1,658
2. Professional	3700	2. Clothing	100	1,200
3. Others	30	3. House, Light etc.	83	996
		4. Others	150	1,800
Total	450	Total	471	6,652

FIELD SURVEY EXAMINATION, 1951

I(C)

Time: 3 hours

Full marks 100

Answer any five questions

All question carry equal marks

1. Answer any five of the following:

(a) Name any 3 countries with their capitals—along with their approximate population.

(b) Name 3 important commodities of Export and 3 important commodities of Import for India.

(c) Mention 3 big Development Schemes undertaken by Government. Indicate the location of these schemes.

(d) Give the latest figures that you know of the total number of (i) Bulls, (ii) Cows, (iii) Buffaloes (male and female separately) of age above 3 years in your state. Mention the original source of your information.

(e) What are the various types of rural labour and their employment. Mention their modes of payment and the normal wage rates they earn in your district.

(f) Mention the most busy months for a cultivator in your district. What are their major seasonal crops?

(g) What are the methods of irrigation employed in your District and what are the implements used?

(h) What is the population of India: (i) in rural areas and (ii) in urban areas, according to the latest census report in your knowledge.

2. Either (a) Demarcate carefully on the sketch map of India provided with state boundaries, any one of the following

- (i) the areas growing Paddy or Wheat, Jute or Cotton
- (ii) the areas of high, medium and low rainfall.

Or Draw a sketch map of your State and indicate therein any one of the following :

- (i) areas growing the most important crops
- (ii) areas of high, medium and low rainfall
- (iii) areas under forests and hills.

(Sketch Map of India showing State boundaries to be supplied)

3. You have been supplied with a list of random plots from which you are required to make a number of sample cuts, and also to enumerate the proportion of area under a given crop.

(a) State what procedure you would follow in the following cases, when on reaching a plot given in the list, you notice that

either (i) the entire plot, which had been sown with the crop, has been completely eaten up by cattle,

Or (ii) the entire plot has already been harvested by the owner

Or (iii) the condition of the crop appears to be extremely poor Or (iv) the crop under investigation is grown as a mixed crop.

The adjoining plot however contains a normal crop which is readily available for harvesting.

(b) Can you indicate the importance of correctly making these decisions?

4. A cultivator family in the first month of Bhadoi sowings has incurred a cash expenditure of Rs. 135. The family consists of 2 adults and 3 children and has cultivated 3.3 bighas of land during the season. Assuming that the family exhibits a normal expenditure pattern, how would you expect the expenditure of Rs. 135/- to be distributed under the following heads:

- | | | |
|-----------------------|--------------------------|-----------------|
| (i) Cereal foods | (v) Oil and ghee | (ix) Wages paid |
| (ii) Pulses | (vi) Other food articles | (x) |
| (iii) Fish; meat; egg | (vii) Clothing | (xi) |
| (iv) Vegetables | (viii) Fuel and lighting | (xii) |

Clearly specify the expense items which you would like to include under headings (x), (xi) and (xii). Have you reasons to feel that any of these expenditure groups are going to be heavier or lower than the monthly average for the year, and if so for what reasons?

5. A district has an area of 3000 square miles and includes 20 police stations and 4000 villages. A sample survey of its area under the standing crops has to be completed within about 3 months.

200 villages from all over the district chosen at random and 4 groups of 10 adjoining plots in each of them are to be surveyed for the purpose.

- (i) Calculate the approximate staff requirements for this survey. Due allowance must be made for inspection work.
- (ii) Draw up a suitable field programme.
- (iii) Estimate the total cost of field work inclusive of expenses to be incurred for field work at the District Head Quarters.

6. For an all-India survey, camp offices have been opened in several important and far flung towns. A Messenger from the Head office at Calcutta is to be sent by train to hand over urgent records to camp offices at (1) Bangalore, (2) Delhi, (3) Lucknow, (4) Nagpur and (5) Patna (Bihar)

Draw up a tour programme for the messenger to follow so that he may avoid unnecessary delay in journey.

Give an estimate of the approximate number of days that will be required by the messenger to cover the whole journey, allowing for a day's halt at each of the camp offices.

7. The following are the sale returns on a particular day in 30 fish stalls chosen at random out of 580 such stalls in a large City.

Stall no.	Day's sale (in mds)	Stall no.	Day's sale (in mds)	Stall no.	Day's sale (in mds)
(1)	0.5	(11)	2.1	(21)	3.0
(2)	3.8	(12)	1.8	(22)	2.1
(3)	1.9	(13)	5.0	(23)	1.9
(4)	2.4	(14)	1.2	(24)	3.2
(5)	2.0	(15)	1.8	(25)	4.0
(6)	4.2	(16)	2.0	(26)	no report
(7)	0.2	(17)	3.4	(27)	2.2
(8)	1.4	(18)	1.1	(28)	2.3
(9)	3.6	(19)	2.2	(29)	0.9
(10)	0.0	(20)	1.8	(30)	0.0

(i) Calculate the average quantity of fish sold per stall on the day, and estimate the total quantity of fish sold in the whole town on the same day.

(ii) Find out the least and the highest quantities sold by a stall in the above samples. Express these as percentages of the average sale per stall.

FIELD SURVEY EXAMINATION, 1951

PART 2(A)

Time: 3 hours

Full marks 100

All questions carry equal marks

Four questions are to be attempted.

1. A family budget enquiry is required to be made on a sample of 500 middle-class families in Calcutta in order to study their expenditure pattern. Describe how you will organise the work.

Draft a suitable skeleton questionnaire indicating therein the broad expenditure groups only.

What will be the major considerations in the selection of your staff for such an enquiry.