INDIAN STATISTICAL INSTITUTE

STATISTICIAN'S DIPLOMA EXAMINATION, PART I, 1950.

PAPER I

Time allowed: 4 hours. Full marks 100.

(It is not necessary to answer all the questions in order to secure full marks. Attempt as many as you like remembering that the questions do not carry equal marks and credit will be given according to the quality of the answers.)

1. Under what circumstances the 't'-statistic was introduced in statistical analysis? Enumerate fully the various uses of the 't'-statistic.

2. Explain the logic underlying the technique of Analysis of Variance. Show how this method can be used to test for the null hypothesis regarding the multiple correlation coefficient.

3. What is the Z-transformation of the correlation coefficient and what are its uses? Show how this transformation can be used to test for the equality of a number of observed correlation coefficients.

4. In regression problems what are the logical differences (i) when the independent variables are stochastic (ii) when they are non-stochastic? Do we apply different tests for an observed linear regression coefficient under the above two hypotheses? (Assume the usual normal homoscedastic hypothesis regarding the dependent variable.) Give full explanation of your answer.

5. Define Pearsonian χ^2 for goodness of fit and describe its various uses. Explain fully the reasons for the following rules in its application.

(i) It is necessary to reduce the number of degrees of freedom by one unit for each parameter estimated from the sample.

(ii) In no group shall the expected number be less than 5.

What devices are available when the rule (ii) cannot be satisfied?

6. Explain the differences between the outlooks of the two methods of estimation of a parameter (i) method of point estimation, (ii) method of interval estimation. Briefly describe the various methods of estimation introduced by Fisher and by Neyman.

7. If p denotes the proportion of A's and $q=(1-p)$, the proportion of *not* A's in a population of size N, then

(a) show that the variance of the estimate of p on a sample of n is given by

$$V(p_n)=\frac{N}{N-1}\frac{n}{pq}$$

and (b) obtain an unbiassed estimate of $V(p_n)$ in terms of the sample estimates p_n and $q_n=(1-p_n)$

8.(a) Explain the following terms :(i) contingency table & (ii) correlation table.

(b) Comment on the following

90 per cent of people who smoke die before reaching the age of 60 years. Smoking is therefore bad for health.

(c) The following summary appears in a report on a sample survey covering 1000 fields. Scrutinize the numbers and point out if there be any misprint.

Manured field	510	Fields both irrigated and manured	180
Irrigated fields	490	Fields both manured and growing improved varieties	140
Fields growing improved varieties	427	Fields both irrigated and growing improved varieties	85

9. (a) Obtain the values of the mean, standard deviation, mean difference and the interquartile range for the distribution

$$dF = y_0 e^{-x/\sigma}\, dx \quad 0 \leqslant x \leqslant \infty$$

(b) Obtain the maximum likelihood estimate of the parameter σ of the above curve.

10. Goddard, the captain of the West Indies cricket team, is reported to have observed the rule of calling 'tails' every time the toss was made during the five matches of the last Test series with the Indian Team. What is the probability of his winning the toss in all the five matches?

How will the probability be affected if (a) he had made a rule of tossing a coin privately to decide whether to call 'heads' or 'tails' on each occasion, (b) the factors determining his choice were not predetermined but he called out whatever occurred to him on the spur of the moment?

11.(a) Derive Lagrange's formula for interpolation. State the underlying assumptions.

(b) Use the formula to obtain the value of u_{13}

x	11	17	21	23	31
u_x	14,646	83,526	194,486	270,846	923,526

STATISTICIAN'S DIPLOMA EXAMINATION, PART I, 1950.

PAPER II

Time allowed: 4 hours. Full marks 100.

(It is not necessary to answer all the questions in order to secure full marks. Attempt as many as you like remembering that the questions do not carry equal marks and credit will be given according to the quality of the answers.)

1. Describe the principles underlying the two factor theory of Spearman. Comment on its adequacy in explaining the variations observed in test scores. Do you consider the multiple factor theory as an improvement over the two factor theory?

2. Discuss the role of blood groups in

(i) studying the problem of evolution

(ii) judging the homogeneity of a group of individuals

(iii) medical practice

(iv) cases of disputed parentage.

3.(a) Can you suggest a suitable questionnaire for the 1951 Census with a view to furnish information for the construction of

1. mortality tables
2. net reproduction rates?

(b) What statistics would you compare to assess the relative importance of two places for the establishment of a health centre?

4. (a) What is meant by the design of an experiment?

(b) Explain the significance of factorial experiments.

(c) In a factorial experiment with 3 factors each at 2 levels conducted in 6 randomised blocks of 8 plots it was discovered later that 2 factors used were identical. This

means that the experiment consisted of only two factors each at 2 levels but the combinations with one factor repeated twice. How would you analyse such an experiment?

5. A lady claims that on tasting a cup of tea she can say whether milk or decoction has been added first to the cup. Design a suitable experiment to test the lady explaining the various principles you may introduce.

6. Explain in detail the information that could be obtained in a census of Industrial production; and also the recent efforts made by the Government of India to obtain this information for the important industries through the application of the Industrial Statistics Act. How far is it possible to compile total physical output from the data of two such Censuses? Give a scheme for constructing Index Number of Industrial production.

7. What is meant by the "Balance of payments" of a country in international trade? What are the different items that enter into its computation and how are they estimated?

Point out the gaps in the information for the correct estimation of India's Balance of payments and explain how they can be filled up.

8. It is proposed to check up the accuracy of the area under different crops recorded each year by patwaris in their registers. Prepare the design for a suitable multistage sampling scheme for a province. Point out how to fix the number of units for the different stages and the controls you use to check up the investigator bias etc.

9. Point out the objects of a cost of living Index Number and explain in detail the method of constructing such a series.

Examine some of the existing series of cost of living Index Numbers for important Industrial cities in India and point out their merits and defects. Is it feasible to have an All India series of cost of living Index Numbers?

10. Point out the different methods of estimating the national income of a country and the uses of such estimates. Prepare a scheme for estimating the National Income of India pointing out the deficiencies of available statistical data and how they can be made up?

11. Explain the theory underlying the Shewhart Control Charts for detecting existence of assignable causes for variation in the quality of manufactured articles in any process of production.

Explain in detail the method of constructing a P chart and X & R charts.

STATISTICIAN'S DIPLOMA EXAMINATION, PART I, 1950.

PAPER III (Practical)

Time: 6 hours. Full marks 100.

1. The following table shows the percentage success of bolls from flowers for 5 different strains of the cotton plant.

Blocks / Kinds	I	II	III	IV	V
Control	38.2	37.7	38.9	37.9	38.2
Selection A	43.2	41.0	42.3	41.2	40.2
Selection B	46.5	45.3	45.0	45.6	44.7
Selection C	46.8	47.4	49.3	47.1	46.5
Selection D	49.5	46.6	48.7	49.6	47.6

(i) Compare control with all the selected varieties.

(ii) Arrange the different varieties in order of the percentage of success of bolls from flowers. 30

2. The following table gives the frequency distribution of the number of tillers.

No. of tillers	Observed frequency
0	3
1	11
2	7
3	37
4	144
5	245
6	410
7	421
8	400
9	305
10	145
11	87
12	26
13	12
14	5
15	1
16	0
17	1

Frequency of Constants	
Size of the sample	2250
Mean	7.2249 ± 0.0425
S.D.	2.0160 ± 0.0301
β_1	$.0551 \pm .0263$
β_2	3.3436 ± 0.2003

Fit a suitable curve to the data. Test the goodness of fit by the (P, χ^2) test. From the χ^2 test write down your observations. 30

3. Other things being equal the pressure (p) of a gas is dependent on its volume (v) according to the relation $pv^a = b$ where a and b are constants. In an experiment the following values are obtained. Estimate a and b. Show the expected and observed values of p for different values of (v) on a graph paper.

p(kg./cm²)	0.5	1.0	1.5	2.0	2.5	2.0
v(litres)	1.62	1.00	0.75	0.62	0.52	0.46

20

4. When women were employed in airplane factories in U.S.A., one official stated that the number of accidents involving women would be twice the number involving men. Use the following data from 5 different factories to test the hypothesis that the number of women involved in accidents is twice as large as the number of men involved in accidents.

Number of accidents

Factories	Male workers	Female workers
A	15	22
B	12	19
C	18	28
D	20	32
E	18	26

Assume that the factories were of the same size and employed the same number of men and women. What other hypothesis would it be advisable to test by using the same data? State the hypothesis carefully and draw a conclusion. 20

STATISTICIAN'S DIPLOMA EXAMINATION, PART I, 1950.
PAPER IV (Practical)

Time allowed: 6 hours. Full marks 100.

1. The table below gives a classification of 11,950 male deaths, by age at death and civil condition.

Civil condition	Age at death in years									
	10–14	15–19	20–24	25–34	35–44	45–54	55–64	65–74	75–84	85 & over
Bachelors	130	206	287	578	797	885	939	914	867	391
Married	130	205	278	500	558	492	447	329	241	77
Widowers	—	—	6	74	233	356	397	441	391	137
Divorced	—	—	—	2	6	30	86	141	223	176

Does this indicate any relation between longevity and civil condition? 28

2. Rate of gain in pounds per day (y), initial weight in pounds (x_1) and initial age in days (x_2) recorded for 30 swine in an experimental station, are given in the table below. 40

y	x_1	x_2	y	x_1	x_2
1.61	74	78	1.22	42	83
1.31	75	99	1.39	47	71
1.12	64	80	1.39	42	66
1.35	48	75	1.56	40	67
1.29	52	94	1.36	40	67
1.24	42	91	1.40	62	77
1.29	62	75	1.47	55	71
1.43	43	63	1.37	62	78
1.29	50	62	1.15	43	70
1.26	40	67	1.22	57	95
1.67	80	78	1.48	51	96
1.41	61	83	1.31	41	71
1.73	62	79	1.27	40	63
1.23	47	70	1.22	45	62
1.49	59	86	1.36	39	67

How is the rate of gain dependent on initial age and weight? What would be the correlation coefficient between rate of gain and initial weight, if the ages of these swine were equal? Is this correlation significant?

3. The following table gives the per capita national income (in dollars) of U.S.A. from farm and non-farm activities during 1910-1941. 32

year	farm income	non-farm income	year	farm income	non-farm income	year	farm income	non-farm income
1910	139	482	1921	110	718	1931	114	605
1911	122	468	1922	153	716	1932	74	442
1912	135	484	1923	180	820	1933	94	417
1913	136	522	1924	180	788	1934	111	487
1914	140	483	1925	223	810	1935	158	540
1915	135	502	1926	216	856	1936	170	626
1916	155	580	1927	209	818	1937	197	671
1917	258	630	1928	222	829	1938	164	621
1918	304	670	1929	223	870	1939	171	658
1919	319	763	1930	170	761	1940	179	717
1920	265	876				1941	254	826

Does this indicate that the importance of the farm sector in contributing to the total national income is gradually dwindling, while that of the other sector is rising?

INDIAN STATISTICAL INSTITUTE

STATISTICIAN'S DIPLOMA EXAMINATION, PART II, 1950.

PAPER I

Time allowed: 4 hours. Full marks 100.

(Attempt any four questions.)

1. (a) Find, to four decimal places, the real root of the equation $x^3+x-1=0$ by whichever of the following methods you consider most suitable, stating briefly the reasons of your choice:

(i) divided differences applied inversely

(ii) successive approximation

(iii) elimination of third differences.

(b) For what value of x is the following tabulated function a minimum?

x	y
0.2	0.9182
0.3	0.8975
0.4	0.8873
0.5	0.8862
0.6	0.8935
0.7	0.9086

2.(a) Find the sum to n terms of the series

1, 4, 13, 30, 59,.....

(i) assuming fourth differences constant,

and (ii) assuming third differences increase in geometrical progression.

Compare the two answers where $n=6$.

(b) In the following table examples are given for certain combination of ages of the annual premiums charged by a British life office for an assurance of £100 payable on the first death of two persons.

Ages	Annual premium £. s. d.
50 & 50	5 1 1
53 & 53	5 17 4
56 & 56	6 17 8
50 & 53	5 9 10
50 & 56	5 19 10
53 & 56	6 6 11

Making the fullest use of the available data, calculate to the nearest penny the annual premium for such an assurance in the case of a pair of persons aged 51 and 52 respectively.

3.(a) Explain briefly the objects of numerical integration and the circumstances in which they are used.

(b) Prove that

$$\int_0^6 u_x\,dx = \frac{3}{10}\left[(u_0+u_6) + 5(u_1+u_3)+(u_2+u_4)+6u_5\right]$$

approximately. Also show that the formula is true to fifth differences and ascertain the sixth difference error.

(c) Establish Newton-Coles formula of integration.

4.(a) An urn contains a white balls and b black balls. Two players, taking turns, draw out the balls, one at a time, and whoever extracts the first white balls wins the game. By considering the chances of winning the game, find the sum of the series

$$1+\frac{b}{a+b-1}+\frac{b(b-1)}{(a+b-1)(a+b-2)}+\ldots \text{ to } (b+1) \text{ terms.}$$

(b) Show how the value of π can be determined with sufficient accuracy with the help of a simple experiment in geometrical probability.

5. Explain clearly the distinction between stochastic convergence and mathematical convergence to a limit.

Establish Tchebycheff's inequality for a stochastic variable having an arbitrary distribution with finite variance. Hence show that for such a population the sample mean converges stochastically to the population mean.

6. A bi-variate universe consists of discrete variables x and y, in which x takes the values $x_1, x_2, \ldots x_n$ and y takes the values $y_1, y_2, \ldots y_m$. If $\eta(x_i)$ denotes the *a priori* probability that x takes the value x_i and $P(y_j|x_i)$ the conditional probability that y takes the value y_j when it is known that x has taken the value x_i, show that the conditional probability $P(x_i|y_j)$ is given by

$$\frac{\eta(x_i)P(y_j|x_i)}{\sum_{1}^{n}\eta(x_i)P(y_j|x_i)}$$

An urn contains five balls, all black, all white, or of both colours. Three balls are drawn together at random, of which two are found to be black and one white. To find the probability that the urn contains three black and two white balls,

(i) Show that the problem has no unique solution.

(ii) State the assumption known as "Bayes' postulate" which would permit a definite solution of the problem. Find the solution under this assumption.

(iii) Write a critical note on Bayes' postulate.

STATISTICIAN'S DIPLOMA EXAMINATION, PART II, 1950.

PAPER II

Time allowed: 4 hours. Full marks 100.

(Answer any four questions.)

1. Deduce the Binomial probability distribution. Prove that it tends to the Poisson and Normal laws under certain conditions, stating those conditions explicitly.

2. Starting with the Pearson's differential equation, deduce the system of frequency curves and show how they can be represented in the $\beta_1-\beta_2$ plane.

3. For the bivariate normal surface, $ke^{-\frac{1}{2}(ax^2+2hxy+by^2)}$ interpret the constants a, b, c and k in terms of well-known statistical constants of a bivariate distribution. Show that both the regressions are linear and the array distribution are normal and homoscedastic.

4. Discuss the use of the 'method of moments' in graduating frequency data in relation to the 'method of least squares' and the 'method of maximum likelihood'.

5. Prove the following properties of the multivariate normal distribution:—

(a) the marginal distribution of each variable is normal;

(b) the array-distribution of each variable is normal with a constant variance but a mean which is a linear function of the other variables.

(c) the joint distribution of any two variables for given values of all other variables is bi-variate normal with a correlation independent of the variables whose values are fixed.

STATISTICIAN'S DIPLOMA EXAMINATION, PART II, 1950

PAPER III

Time allowed: 4 hours. Full marks 100.

(Answer any three questions.)

1.(a) What is the difference between an unbiased statistic and a consistent statistic?

Show that $S^2 = \Sigma(x-\bar{x})^2/(n-1)$ is an unbiased estimate of σ^2 and that S is a biased estimate of σ irrespective of the form of the population. Calculate the extent of the bias in S when the population follows the Normal law.

(b) Derive the exact distribution function of $S^2 = \Sigma(x-\bar{x})^2/(n-1)$ for samples of n from a Normal population and show that the standard deviation of S^2 is $\sigma^2\sqrt{2/(n-1)}$.

2.(a) Prove the theorem:

If $\chi_1^2, \chi_2^2, \ldots \chi_k^2$ possess independent χ^2 distributions with $\nu_1, \nu_2, \ldots, \nu_k$ degrees of freedom respectively then $\chi_1^2 + \chi_2^2 + \ldots + \chi_k^2$ will possess a χ^2 distribution with $\nu_1 + \nu_2 + \ldots + \nu_k$ degrees of freedom.

(b) Find the distribution of Pearsonian χ^2 for testing goodness of fit.

3. Discuss the uses of D^2-statistic in problems of discrimination and classification. In what way does this statistic differ from Karl Pearson's C^2-statistic?

4. (a) Write notes on
   (i) discriminant functions
   (ii) various uses of the χ^2- statistic.

   (b) Explain
   (i) Small differences can be detected provided samples are large.
   (ii) A significantly small χ^2 also disproves the null hypothesis.

INDIAN STATISTICAL INSTITUTE

COMPUTER'S CERTIFICATE EXAMINATION, 1950.

PART 1A: SECTION 1.

Time allowed: 3 hours. Full marks 100.

(Use separate answer books for the two groups. No machine allowed.)

GROUP A. (40 marks)

1. Complete the table below

X	X^2	$(X-5)^2$	Y	Y^2	$(Y-6)^2$	XY	$(X-5)(Y-5)$
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
3			5				
7			6				
6			9				
5			3				
4			8				
2			3				
1			4				
8			2				
9			11				
5			9				
Total							

Also verify that

(a) total of col. (3) = total of col. (2) − 250

(b) total of col. (6) = total of col. (5) − 360

(c) total of col. (8) = total of col. (7) − 300

Or

Evaluate:

(i) $x^3 - 10.9x^2 + 9.4x - 102$, when $x = 5$, 8, 9, and 10.

(ii) If $a = 0.00731$, find the values of a^2 and a^3, correct to five decimal places.

Show that within this degree of accuracy $1 + a + a^2$ and $1 + a + a^2 + a^3$ yield the same result.

(iii) An *Are* is the area of a square on a side of 10 metres; and 1 metre = 1.0936.... yards. Express *Are* in terms of square yards correct to the nearest hundredth.

2. Find the sum of squares of all integers from 1 to 20 and also the sum of all even integers from 1 to 20, and find their difference.

Group B. (60 marks)

1. The following informations were collected, as a result of investigations into the general health of school students on a sample of 24.

Serial No.	Age in years	Sex	Home district	Height in inches	Weight in lbs.
1	13	Male	Burdwan	60	118
2	9	Male	Calcutta	50	80
3	11	Female	Birbhum	64	110
4	12	Male	Burdwan	54	96
5	8	Male	Nadia	40	75
6	17	Male	Burdwan	66	120
7	11	Male	Burdwan	59	111
8	10	Female	Calcutta	56	98
9	10	Male	Birbhum	49	99
10	6	Male	Birbhum	38	80
11	7	Male	Nadia	37	67
12	9	Female	Calcutta	44	80
13	12	Female	Birbhum	56	100
14	12	Male	Calcutta	63	104
15	13	Male	Birbhum	60	96
16	14	Male	Calcutta	70	113
17	10	Male	Burdwan	53	94
18	11	Female	Birbhum	55	101
19	7	Female	Burdwan	36	65
20	7	Male	Nadia	32	50
21	8	Male	Calcutta	40	79
22	13	Female	Nadia	59	93
23	11	Male	Calcutta	58	102
24	10	Male	Calcutta	47	86

Find the following:

(1) Average height and average weight of Male and Female students separately.

(2) Average height and average weight of the school students separately for the Age-groups (i) 6 years–9 years (ii) 10 years-12 years (iii) over 13 years.

(3) Total number and their percentage to total of the students who have

(i) a height above 50 inches

(ii) an weight above 100 lbs

(iii) both, i.e. those with a height above 50 inches and at the same time having a weight above 100 lbs.

2. The following table was copied from a computing sheet, in which the total yield of paddy, rate of paddy in maunds per acre and the square of these yield rates, were computed for nine plots, together with the relevant totals.

At the time of copying certain figures were omitted. A few mistakes were also made either at the stage of computing or at the stage of copying.

	Ser. No. of family	Area of plot in acres	Total yield of paddy in mds.	Yield rate in mds. per acre	Squares of Col.(4)
	(1)	(2)	(3)	(4)	(5)
	1	3	42	14.0	196.00
	2	5			116.64
	3	2	23	11.5	132.25
	4	2	25	12.5	144.25
	5	4	56	14.0	196.00
Sub-total		15			785.14
	6	4	34	8.5	70.25
	7	2			
	8	2	15	7.5	56.25
	9	4	24	6.0	36.00
Sub-total		12			
Grand total		28		11.0	

Scrutinize the above carefully and supply the missing figures, including subtotals and totals after correcting the obvious mistakes.

3. Complete the following working sheet retaining figures correct to 4 decimal places.

Ser. No.	x	$1/x$	$\sqrt{x}$	x	$\log_{10}x$	$\sqrt{x}$	x	x^2	$\sqrt{x}$
	(1.1)	(1.2)	(1.3)	(2.1)	(2.2)	(2.3)	(3.1)	(3.2)	(3.3)
1	1.0			0.40			0.160		
2	1.1			.44			0.176		
3	1.2			.48			0.192		
4	1.3			.52			0.208		
5	1.4			.56			0.224		
6	1.5			.60			0.240		
7	1.6			.64			0.256		
8	1.7			.68			0.272		
9	1.8			.72			0.288		
10	1.9			.76			0.304		
11	0.20			.080			.020		
Total									

Or

In a small town with a total population of 6250, 30% is reported to be of age 12 or under, 50% of age above 12 and below 50 and the rest above 50.

Assuming that annual per capita rate of consumption for children below 12 years is 2 maunds, adults above 12 and below 50 consume 4 maunds and old men above 50 consume 3.5 maunds of rice per head per year, estimate

(a) the total requirement of rice for the town in a year

(b) total annual requirement for the town, if people are forced to reduce the rate of consumption by 6¼ % in the case of children, 10% in the case of adults and 12½% in the case of old persons,

COMPUTER'S CERTIFICATE EXAMINATION, 1950.

PART 1A: SECTION 2.

GROUP A (50 marks)

(Use separate answer books for the two groups. No. machine allowed.)
Attempt two out of the three questions of this group.

1. (*a*)

12	13	14	15	16	17	18	19	20
4718	5600	6700	7962	9404	11038	12876	14930	17212

Plot the ten pairs of values of x, y given in the table above on a piece of graph paper and join the plotted points by smooth (free hand) curve. Read from the graphs the value of y corresponding to $x=14.5$ and 16.5.

(*b*) Draw the figure bounded by the following four lines:

(i) line at $x=12$, parallel to the y-axis, (ii) line at to the $x=16$, parallel x-axis, (iii) line at $y=4000$, parallel to the x-axis, and (iv) the (x, y) curve drawn in question 1(*a*) above. Find out the area of this figure.

2. (a) Find out the required values of y in the following three cases by simple interpolation:

(i) when $x=-1.5$, $y=-2.39$; and when $x=-2.8$, $y=-4.90$
Find the value of y for $x=-2.3$.

(ii) when $x=-1.0$, $y=-3.36$; and when $x=3.7$, $y=2.07$.

(iii) when $x=-3.8$, $y=-2.98$; when $x=4.2$, $y=0.17$.

(*b*) work out the value of 3.48925×4.15682 correct to 4 decimal places, by the method of contracted multiplication, or otherwise.

3. (a) Solve the equation $3.89x^2+24.5x=2.89$.

(*b*) Using the mathematical tables for squares, cubes, etc, work out the square, cube, fourth power, reciprocal and the square root of the reciprocal for the following numbers:—

3.13, 45200, 254, 0.000218 and -0.000218.

Group B (50 marks)

1. The following data relate to weights of contents of cans (size No. 2½) of tomatoes.

wts. of cans	frequency
25.5	1
25.0	2
24.5	6
24.0	9
23.5	14
23.0	14
22.5	31
22.0	34
21.5	33
21.0	41
20.5	26
20.0	24
19.5	12
19.0	7
18.5	2
18.0	0
17.5	2
17.0	0
16.5	1
16.0	0
15.5	1

Calculate mean, standard deviation and standard error of mean weight of contents of tomato cans.

2. The table below gives the Annual indices of Industrial Production during the year 1939 to 1946.

	Indian Cotton	Jute	Steel	Pig-iron	Cement	Sugar	Electricity	Coal	Misc.
Weights	40	17	8	7	3	10	5	6	5
1939-40	96.8	115.3	108.0	109.5	103.0	101.1	105.0	102.0	109.9
1940-41	106.7	100.0	120.7	116.7	103.0	171.1	114.2	105.7	126.4
1941-42	125.9	115.3	137.5	120.2	132.1	117.1	137.5	107.7	129.1
1942-43	128.8	112.6	131.5	107.1	129.8	160.8	140.8	103.7	120.1
1943-44	138.7	96.4	137.8	100.6	152.6	165.9	155.0	91.6	129.2
1944-45	131.0	100.0	127.7	77.4	121.4	100.0	165.8	98.0	130.9
1945-46	132.3	100.0	131.2	83.9	128.0	125.2	167.5	107.7	106.3

Evaluate a weighted General Index of Industrial Production for all the years.

COMPUTER'S CERTIFICATE EXAMINATION, 1950.

PART 1B: SECTION 1

GROUP A (50 marks)

Time allowed: 3 hours.

(Use separate answer books for the two groups.)

1. Draw up in detail with proper attention to headings, spacing, double lines etc. and showing all sub-totals, a neat blank table in which could be shown the comparative statements between the years 1948 and 1949 of the number of persons engaged in six industries distinguishing males and females and among the latter single, married and widowed.

2. The following table gives two series of measurements of maximum head-length and maximum head-breadth in m.m. respectively of 104 artisan class of U.P.

maximum head length in m.m.	maximum head-breadth in m.m.	maximum head-length in m.m.	maximum head-breadth in m.m.	maximum head-length in m.m.	maximum head-breadth in m.m.	maximum head-length in m.m.	maximum head breadth in m.m.
183	131	193	143	184	139	187	139
185	136	195	135	185	136	190	139
188	135	193	144	189	149	189	138
196	139	193	136	182	136	205	144
173	135	184	135	188	142	192	139
195	146	187	143	185	139	188	137
187	137	183	141	189	128	185	147
183	136	194	137	182	141	198	142
192	142	175	139	189	135	185	139
184	139	195	137	186	139	190	139
183	140	188	134	182	136	185	131
184	137	184	135	190	139	186	140
184	135	186	132	197	134	193	139
190	138	183	146	194	144	175	139
175	140	192	134	182	131	185	129
177	137	195	138	189	138	191	147
192	137	197	141	195	133	184	137
197	142	192	137	183	135	185	134
191	136	184	133	186	135	201	141
171	140	187	133	190	133	191	137
191	139	184	146	187	140	179	136
190	129	201	140	188	139	187	137
192	147	185	138	179	131	188	142
190	137	180	135	174	137	201	141
187	139	185	133	189	141	170	128
184	141	191	139	185	128	196	139

Prepare a two-way frequency table and calculate the correlation coefficient between the two characters.

Group B (50 marks)

1.(a) A weight moving under a certain force was observed at the end of each second (t) to have passed on a distance (s) from the beginning. The results were as follows:

t	1	2	3	4	5	6	7	8
s	0.31	1.12	2.20	3.82	6.36	9.02	12.00	16.10

Fit the straight line of closest fit by the method of least squares.

(b) With the same data fit by the method of least squares, a second order parabola of closest fit.

2. The following data relate to yields (in suitable units) of 6 varieties over 36 randomized plots each variety being replicated 6 times.

Varieties						
A	248	217	227	210	218	215
B	245	217	240	210	205	219
C	238	228	205	191	224	211
D	254	223	189	180	216	209
E	249	221	226	242	240	218
F	240	212	194	211	202	215

Set up the Analysis of Variance Table.

COMPUTER'S CERTIFICATE EXAMINATION, 1950.

Part 1B: Section 2.

Group A (50 marks)

Time allowed: 3 hours.

(Use separate answer books for the two groups.)

1.(a) Find the value of

$$\left[1+\frac{(1.2)^{5.4}}{(0.06)^{5.8}}\right]^{1.8}$$

(b) The following is an extract of a two-way table showing the value of Z, which is a function of x and y for certain values of x and y.

y	x	
	0.4	0.5
0.3	0.2943	0.2728
0.4	0.2784	0.2602

Find the value of Z when $x=0.467$ $y=0.320$ by linear interpolation.

2. The frequency distribution of the diameter (in m.m. to the nearest whole number) of ears of Maize is given below.

Dia (in m.m.)	frequency	Dia (in m.m.)	Frequency	Dia (in m.m.)	Frequency
x	f	x	f	x	f
38	4	42	25	46	30
39	7	43	48	47	29
40	16	44	46	48	17
41	24	45	43	49	8
				50	3
					300

The mean and standard deviation of this distribution are 44.157 m.m. and 2.448 m.m. respectively. Fit a normal curve to this distribution and test the goodness of fit.

3. The following table shows the classification of families according to the performance of parents (x) and the performance of first born sons (y) at a university.

	x				
y	Poor	Average	Good	Very good	Total
Poor	9	8	2	1	20
Average	40	342	108	4	500
Good	13	292	365	30	700
Very good	2	8	45	25	80
Total	70	650	520	60	1300

Is there any significant association between the performance of parents and permance of sons?

GROUP B. (50 marks)

1.(a) The following values were calculated during the course of a sample survey.

District	Number of sample units	Average yield rate in mds. per acre (Rice crop)	Standard deviation
Bankura	236	11.13	4.15
W. Dinajpur	214	8.12	3.66

Are the average yield rates for the two districts significantly different from one another?

(b) The estimated area under rice crop in a country, based on 18523 sample units, was 8374.5 (000 acres) and the standard error of this estimate was 47.8 (000 acres). If the cost of collecting information for a single sample unit is 5 annas what will be the total cost required to collect information for the necessary number of sample units in a sample in order to get an estimate which differs from the true value by less than 98.1 (000 acres) in 90 out of 100 samples?

2. The following gives the premium income in lakhs of Rupees earned by Indian Insurance companies for each of the years 1919-1924.

year (t)	premium income (y)	year (t)	premium income (y)
1919	128	1922	174
1920	148	1923	186
1921	160	1924	204

If this can be graduated by $y=208(1.12)^{t-1926}$ find the expected values and draw graphs of the observed and expected values. Is there a good agreement between observed and expected values?

3. Name the publications in which you can get the following informations.

(i) Number and Amount of foreign Money Order paid in India during the year 1939-40.

(ii) Index Number of wholesale prices in India during week ending 22nd April, 1950.

(iii) Exports of raw cotton from India in the month of November, 1949.

(iv) Yield per acre of cleansed rice in Bihar during the year 1938-39.

(v) Number of married females between the ages 0-5 in Hooghly District of Bengal in the year 1941.

(vi) Working class cost of living index number for Bombay city in the month of January, 1950.

COMPUTER'S CERTIFICATE EXAMINATION, 1950.

PART 1C: SECTION 1.

GROUP A (50 marks)

Time allowed: 4 hours

(Use separate answer books for the two groups)

1. Figures for the yield of wheat, height of shoots at ear emergence and number of plants at tillering are given below.

Yield of wheat	Height of shoots (at ear emergence)	Average number of plants at tillering
x	x	x
19.0	25.6	14.9
22.2	25.4	13.5
35.3	30.8	4.6
32.8	33.0	14.4
25.3	28.5	12.8
35.8	28.0	7.5
32.4	25.4	7.2
32.2	28.3	9.5
43.7	35.3	6.8
35.7	32.4	9.7
28.3	25.9	9.2
35.2	24.2	7.5
26.2	27.9	18.5
34.7	34.4	22.2
40.0	32.5	10.0
29.6	27.5	17.6
20.6	23.7	14.4
47.2	32.8	7.9

Test the significance of the partial correlation coefficients $r_{13\cdot2}$ and $r_{12\cdot3}$

2. From statistical tables, find

(i) the ordinate of the Normal Curve

$y=0.3989\,e^{-u^2/2}$ corresponding to
$x=0.856$.

(ii) the area under the same Normal Curve between $x=-1.2$ and $x=2.5$

(iii) the 5% value of χ^2 for d.f. $=20$

(iv) the 2.5% value of F, $n_1=8, n=14$

(v) the lower 5% value of F, $n_1=8, n_1=14$.

3. Given

$$1 + i = \left(\frac{1+j}{p}\right)^p$$

Find j corresponding to (i) $p = 6$, $i = .01$ (ii) $p = 52$, $i = .0125$, each correct to 6 places of decimals.

4.(a) Expand and evaluate the Binomial

$(1 + .03)^6$

(b) Find the common logarithm corresponding to the Naperian logarithm 2.19722.

GROUP B. (50 marks)

1. Fiducial limits of p are obtained from the formula:—

$$p = \frac{t}{n} + \frac{2}{n} \pm \sqrt{\frac{1}{n^2} + \frac{1}{n}\left(\frac{t}{n}\right)\left(1 - \frac{t}{n}\right)} \Big/ \left(1 + \frac{4}{n}\right)$$

Scrutinize the following table, correct errors, if there are any, fill up the missing columns and calculate the values of p

Sl. No.	n	t	$\frac{1}{n}$	$\frac{2}{n}$	$1 + \frac{4}{n}$	$\frac{1}{n_2}$
1	7981	7134	.0001253	.0002507	1.0005012	.0000000
2	504	425	.0019841	.0039683	1.0079365	.0000039
3	401	375	.0024938	.0049875	1.0099751	.0000062

Sl. No.	$\frac{t}{n}$	$1 - \frac{t}{n}$	$\frac{t}{n}\left(1 - \frac{t}{n}\right)$	$\frac{1}{n}\left(\frac{t}{n}\right)\left(1 - \frac{t}{n}\right)$	$\frac{1}{n_2} + \frac{1}{n}\left(\frac{t}{n}\right)\left(1 - \frac{t}{n}\right)$
1	.893873	.106127	.0918641	.0000119	.0000119
2	.843265	.156745	.1321767	.0002623	.0002662
3	.935162	.064838	.0606340	.0001512	.0001574

Sl. No.	$\sqrt{\frac{1}{n^2} + \frac{1}{n}\left(\frac{t}{n}\right)\left(1 - \frac{t}{n}\right)}$
1	.001091
2	.051503
3	.039674

2. Solve graphically :—

$f(x) = 2.3x^2 + 3.2x + 0.0 = 0$

and read from the graph the value of x for which $f(x)$ is least.

3. Analyse the variance and test whether the observer difference and the machine difference are significant:—

	Observer 1	Observer 2	Observer 3
Readings on machine No. 1	253	261	254
	261	257	263
	232	244	246
	244	249	248
	251	247	243
Readings on machine No. 2	271	267	261
	260	270	270
	258	262	254
	262	260	259
	274	270	271

COMPUTER'S CERTIFICATE EXAMINATION, 1950.

PART 1C: Section 2

GROUP A (50 marks)

Time allowed: 4 hours.

(Use separate answer books for the two groups.)

1. Evaluate the following determinant and all its cofactors:

$$\begin{vmatrix} 12.1 & 3.0 & 6.7 \\ 13.5 & 9.8 & 8.5 \\ 6.3 & 1.2 & 0.9 \end{vmatrix}$$

2. Find the values of $f(x)$ by interpolation from the table given below when (i) $x=3.25$ and (ii) $x=8.75$. In case (i) use the first five values in the table while for (ii) use the last five values.

x	$f(x)$	x	$f(x)$
3	0.0523360	8	0.1391731
4	0.0697565	9	0.1564345
5	0.0871557	10	0.1736482
6	0.1045285	11	0.1908090
7	0.1218693	12	0.2079117

3. Prepare a price index for the data given below with 1933 as the base year. A, B, C etc. are different commodities whose prices are noted. The weights are given at the bottom of the table.

Year	Prices						
	A	B	C	D	E	F	G
1933	4	3	5	8	6	9	7
1934	5	4	6	9	7	9	8
1935	5	3	5	9	8	10	8
1936	4	4	6	12	9	10	10
1937	5	4	7	14	8	10	10
1938	5	5	8	13	9	11	12
1939	6	5	9	14	9	12	12
1940	6	5	9	15	10	13	13
Weights	(3)	(4)	(2)	(6)	(4)	(3)	(3)

GROUP B (50 marks)

1. The table below gives the frequency distribution of 1088 teak trees according to size of diameter in inches. Calculate the mean and standard deviation and obtain the equation of the Normal Curve to fit this distribution.

Diameter in inches	Frequency	Diameter in inches	Frequency
4.5– 7.5	8	22.5–25.5	213
7.5–10.5	28	25.5–28.5	145
10.5–13.5	50	28.5–31.5	76
13.5–16.5	120	31.5–34.5	36
16.5–19.5	181	34.5–37.5	16
19.5–22.5	217		
		Total	1088

Draw a histogram of the distribution and on it draw the normal curve after calculating the ordinates at suitable points.

2.(a) Tabulate the index numbers of the cost of living in India, U.K. and U.S.A. for each of the ten years 1937 to 1946 and plot them in a graph (source of data:— Monthly Bulletin of Statistics issued by the Statistical Office of the United Nations).

(b) From an appropriate statistical publication of the Government of India obtain the figures of value of total imports of merchandise and treasure into India for each official year from 1941-42 to 1945-46 and the corresponding figures for total exports from India.

3. With a view to determining the effect of manuring on the production of leaves of a camphor-yielding variety of *tulsi* plant, seven manurial treatment A, B, C, D, E, F and G of which C is control (no manure) were tried in replicated plots of 1/20th acre size each, in a randomized block lay-out as per diagram below:

Blocks

I	II	III	IV
D	E	C	B
17	22	15	13
F	F	E	E
21	19	15	30
C	A	A	F
20	18	14	29
G	C	F	G
36	14	15	18
E	G	B	D
15	43	19	13
B	B	D	C
13	18	27	11
A	D	G	A
12	14	24	21

The yield of leaves in lbs. at zero moisture for each plot is entered in the diagram. Perform analysis of variance on the data and calculate the mean yield per acre for each treatment.