

INDIAN STATISTICAL INSTITUTE

STATISTICIAN'S DIPLOMA EXAMINATION, PART I - JULY 1948.

FIRST PAPER

Time allowed 4 hours

Maximum 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 answer).

1.(a) Two dice have their faces marked as follows :

Dice No. 1 - - 2, 4, 4, 6, 6, 6

Dice No. 2 - - 1, 1, 1, 3, 3, 5

What is the most probable sum of throws with the two dice ?

(b) A coin is tossed ($m : n$) times ($m > n$). Show that the probability of at least m consecutive heads is $\frac{n+2}{2^{m+1}}$.

2. An Urn contains $a : \alpha$ red and $b : \beta$ black balls. $a + \beta$ balls are drawn at random from the Urn and it is found that there are α red and β black balls in the sample. Find the probability of drawing the above sample.

Show that the above can be used to furnish an exact test for the absence of association in a fourfold contingency table.

3. (a) Show that the Poisson's Distribution $m^x e^{-m}/x!$ can be obtained as the limit, under certain conditions, of the Binomial Distribution $n C_x p^x (1-p)^{n-x}$.

Obtain the Maximum Likelihood estimate of m , given the following sample values.

x	0	1	2	...	k
frequencies	n_0	n_1	n_2	...	n_k

(b) Obtain in terms of a set of observations (x_1, x_2, \dots, x_n) the maximum likelihood estimates of the parameters a and b occurring in a population of the form $kx/(b-a)$; given that $a < x_1 < x_2 < \dots < x_n < b$.

4. State when the Poisson's Distribution is used in statistical analysis.

During the flying-bomb attack on London frequent assertions were made that the points of impact of the bombs tended to be grouped in clusters. To test this assertion an area was selected comprising of 144 sq. kilometres of South London. This area was divided into 576 squares of $\frac{1}{4}$ sq. kilometres and a count was made of the number of squares containing 0, 1, 2 etc. bomb-hits.

No. of bomb-hit	0	1	2	3	4	5 & over	Total
No. of such squares	229	211	93	35	7	1	576

Do the data substantiate the assertion ?

5. A sample of size n , giving arithmetic mean \bar{x} and standard deviation S has been drawn from the population

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-a)^2}{2\sigma^2}} + \frac{1}{\sqrt{\pi\sigma^2}} e^{-\frac{(x-2a)^2}{\sigma^2}}$$

Use the method of moments to determine a and σ .

Give a short rationale of the method of moments — its advantages and drawbacks specially as compared to other methods of estimation that you may know of.

6. Prove that in any distribution

- (i) the mean deviation about the median is less than the mean deviation about the mean.
- (ii) the mean deviation about the mean is less than the standard deviation about the mean.

(iii) if the skewness be defined as $\frac{\text{Mean-Median}}{\text{Standard deviation}}$, then

$$|\text{skewness}| < 1$$

7. Write a critical account of the various uses of the Chi-square (χ^2) Statistics in tests of significance. Explain in particular under what conditions and in what situations the partition or combination of χ^2 's would be valid and useful, and also how they would be valid.

8. (a) Define Partial and Multiple Correlations and explain the uses of these coefficients. Explain in general terms how you would estimate these co-efficients or test any hypotheses concerning them, indicating in particular how the 'null hypotheses' regarding absence of correlation.

(b) Express the partial correlation coefficient in terms of the partial regression coefficients and the multiple correlation coefficient in terms of the partial correlation coefficients.

9. Write short notes on

- (i) Un-biased estimates
- (ii) Sheppard's correction
- (iii) Intraclass correlation coefficients
- (iv) Uses of the Ogive

10. Show how it is possible to obtain one uniformly most powerful test for the hypotheses $a = a_0$ with respect to alternatives $a > a_0$ and another with respect to alternatives $a < a_0$ in the case of the exponential population :

$$\frac{1}{a} e^{-\frac{x}{a}} dx \quad (a > 0, 0 < x < \infty)$$

11. There is a multinomial probability law relating to 4 cells in which the probabilities are $\frac{1}{2}(1-p)$, $\frac{1}{2}p$, $\frac{1}{2}(1-p)$ and another relating to 10 cells in which the probabilities are

$$\frac{1}{2}(1-p)^2, \frac{1}{2}p(1-p), \frac{1}{2}p(1-p), \frac{1}{2}p^2, \frac{1}{2}p^2, \frac{1}{2}p(1-p), \frac{1}{2}p^2, \frac{1}{2}(1-p)^2, \frac{1}{2}p(1-p), \frac{1}{2}(1-p)^2$$

By actually calculating the informations in both cases with respect to the parameter p , compare the rate of information per unit observation in one case with that in the other.

STATISTICIAN'S DIPLOMA EXAMINATION, PART 1—JULY 1948.

SECOND PAPER

Time allowed 4 hours

Minimum 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 answer.)

1. Explain the general method of construction of Index Numbers of the Cost of Living illustrating your remarks with the help of any typical index in current use in this country.

How far can a change in the cost of living be measured over a period in which there have been considerable modifications of diet or other changes in the consumption of necessary articles, and in what way?

2. Given the monthly statistics for the past ten years of sales of an article with an inelastic demand, how would you analyse these data with a view to judging the results of the current year, as separate monthly figures for that year become gradually available?

3. Examine the problem of the statistical derivation of the demand curve for a staple agricultural crop from time series data, giving a critical appraisal of the link-relative, trend-ratio and multiple regression methods that are usually employed in this connexion.

4. Describe briefly the nature and scope of official data relating to agricultural statistics available in India. Comment on their adequacy and reliability, suggesting lines of improvement, if any.

5. Write notes on any two of the following :—

- (a) National income
- (b) Engel's curves
- (c) Pareto's income law

and (d) Balance of Trade.

6. Discuss briefly the nature of the evidence in support of genetical linkage (where it is supposed to exist) and the technique of preparation of autosomal or chromosomal maps. Also indicate very broadly either (i) the common statistical methods of detection of linkage, or, (ii) assuming linkage, the common statistical methods of estimating it.

7. With regard to the coming 1951 census operations in India what suggestions have you got to offer (and why) in respect of (i) the questionnaire, (ii) the actual taking of the census, and (iii) the tabulation?

8. Given that the West Bengal Government in particular and the Indian Union in general is faced with the problem of rehabilitation of refugees from East Bengal, suppose you were asked to plan a quick sample survey to estimate certain items of relevant information. Draw up a scheme of survey giving a rough indication of the (i) questionnaire, (ii) the organization, and (iii) the tabulation.

9. Explain the anatomy of a life table with special reference to the purpose and implication of the concept of a stationary population. Explain further how any distortion of the age distribution of a growing population gets eventually corrected.

10. Write a short essay on the application of statistical methods to the discovery and standardisation of proper tests for vocational selection or guidance.

Or

Write a short essay on factor analysis bringing out clearly its object and how far it can succeed in attainment thereof.

STATISTICIAN'S DIPLOMA EXAMINATION, PART I—JULY 1948.

THIRD PAPER (Practical)

Time allowed 6 hours

Maximum Marks 100.

1. Find by Newton's formula for interpolation, or otherwise, the value of y corresponding to $x=2.4$ from the following table

x	1	3	5	7	9	11
y	6	32	130	348	734	1336

2. Find a Fourier series to fit the following observations which represent twelve equidistant ordinates of a periodic curve.

U_0	U_1	U_2	U_3	U_4	U_5	U_6	U_7	U_8	U_9	U_{10}	U_{11}
3.717	4.045	3.137	2.276	1.791	1.408	1.373	1.643	1.194	0.646	0.566	1.770

3. Fit a normal curve to the distribution given below of depth of sapwood in 1370 telephone poles.

Depth of sapwood in inches	Number of telephone poles	Depth of sapwood in inches	Number of telephone poles
0.85-1.14	2	3.55-3.84	123
1.15-1.44	20	3.85-4.14	62
1.45-1.74	62	4.15-4.44	48
1.75-2.04	106	4.45-4.74	27
2.05-2.34	153	4.75-5.04	14
2.35-2.64	186	5.05-5.34	5
2.65-2.94	193	5.35-5.64	1
2.95-3.24	188		
3.25-3.54	151		
		Total	1370

Test the goodness of fit and represent the distribution graphically. Have you any comments to make?

4. The mean of a certain normal distribution is equal to the standard error of the mean of samples of 100 from that distribution. Find the probability that the mean of a sample of 25 from the distribution will be negative.

5. Represent graphically the relation between n and p given by

$$1 \pm n(n-1) \log_e (1-p) = 0, \quad \text{when } 0.001 \leq p \leq .05$$

(Plot about ten points. Arrange the calculations necessary in a neat tabular form).

6. In 1940, it was decided to publicize the monthly sales (in tons) of a particular brand of coffee, with wide sale over the whole of America, along with the supplementary information of its sale in a particular state "Ohio".

1940-Coffee sales (in tons)

	Ohio	U.S.A.
January	850	6800
February	880	7000
March	900	7100
April	990	7500
May	1000	7500
June	950	7400
July	900	7000
August	900	6900
September	1000	7000
October	1200	7400
November	1350	7800
December	1500	8000

Arrange the information in a suitable graph that would show the relative changes in sales in Ohio and U.S.A.

What further data (if any) would be needed to test the effect of additional local advertising campaign (at Ohio state).

7. Results of tests on worn wire rope of nominal diameter $1/2$ to $5/8$ inch are given below.

Tensile strength in thousands of pounds per sq. inch.	Number of broken wires in worst lay	Length of worn surface
174	8	0.14
185	0	0.00
188	8	0.12
190	14	0.11
179	0	0.00
183	0	0.11
191	0	0.09
177	0	0.00
183	0	0.12
186	0	0.13
180	0	0.19
184	3	0.15
175	5	0.13
175	0	0.00
166	2	0.16
170	14	0.15
180	0	0.00
181	0	0.12
201	12	0.15
172	0	0.00
184	5	0.16
145	11	0.15
172	0	0.00
133	8	0.11
157	21	0.14
175	0	0.00

Predict the tensile strength when number of broken wires in worst lay is 7 and length of worn surface is 0.08

STATISTICIAN'S DIPLOMA EXAMINATION, Part I—JULY 1948.

FOURTH PAPER (Practical)

Time allowed 6 hours.

Maximum marks 100.

1. The following data give the results of field tests on the corrosion of non-bituminous pipe coatings for underground use.

Soil type	Lead coated steel pipe	Bare Steel pipe
A	27.3	41.4
B	18.4	18.9
C	11.9	21.7
D	28.7	9.8
E	11.3	16.8
F	14.8	9.0
G	20.8	19.3
H	21.6	11.1
I	17.0	32.1
J	7.8	7.4
K	18.6	68.3
L	14.7	20.7
M	19.0	34.4
N	65.3	76.2

Do these types of pipe differ significantly in their resistance to corrosion.?

2. In a public preference survey, the following table classifies the people interviewed according to their opinion regarding interaste marriage.

Opinion	Age in years			
	19-25	26-35	35-55	over 55
Unconditional support	76	125	96	10
Conditional support	69	117	126	17
Condone	14	27	35	4
Oppose unconditionally	60	168	210	46

Examine in what way the nature of opinion changes with age.

3. The distribution of a group of families according to total monthly expense and percentage of expense on food is given below.

Percentage of expense on food	Total monthly expenses in rupees									
	50-60	70-80	90-100	110-120	130-140	150-160	170-180	190-200	210-220	
20-23		1				2				
24-27		1				3				
28-31				1		4	1	3		
32-35	1			3	2	4				1
36-39		5		2	5	8	8	2		
40-43	2	5		11	8	6	3			
44-47	3	9		6	13	8	2	1		
48-51	4	12		4	2	6	3	1		
52-55	4	10		5	2	1		1		
56-60	1	5								
	1	1								

Find the coefficient of correlation between total expense and percentage of expense on food and comment on its implication. Indicate also how one may try to test its significance and what might be the theoretical danger of such a test.

4. The following table gives the value of sales in a commercial firm in thousands of rupees (a) test whether there are statistically significant seasonal fluctuations (b) construct a series exhibiting the residual variation after removing seasonal fluctuations and linear trend.

	1931	1932	1933	1934	1935	1936
January	432	421	384	501	467	512
February	487	494	423	552	540	574
March	542	495	421	590	564	652
April	555	524	459	610	620	664
May	537	504	488	598	582	645
June	554	496	490	584	578	714
July	501	442	484	524	577	642
August	519	478	531	547	583	655
September	537	500	548	576	622	690
October	560	507	566	625	646	719
November	566	494	566	627	695	744
December	1101	918	1123	1218	1353	1451

5. " In 1943-44 production of cotton cloth, factory sugar, and steel ingots reached the levels of 4871 million—, 1.27 million tons and 1.37 million tons respectively. In 1947-48 according to an estimate the output of mill cloth, jute goods, factory sugar, paper, cement, sulphuric acid, coal and steel ingots amounted to 3,800 million—, 1.05 million tons, 0.95 million tons, 1.25 million—, 1.35 million tons, 0.60 million cwts, and 0.90 million tons.

In 1941-42 peak production were noticeable under jute goods 1.26 million tons, paper 1.87 million cwts, sulphuric acid 0.86 million cwts, cement and coal productions in the year were 2.22 and 26.5 million tons respectively.

Increase in the number of days lost through strikes and lock outs in the 1st quarter of 1947 approximated to 5 million working days against 12.3 million in 1946, 4.05 million in 1945 and only 2.34 million in 1943.

In agricultural production the output of cereals in India in 5 years ending 1946-47 averaged 46.8 million tons per annum as against an output of 51.8 million tons in the corresponding period upto 1944-45."

The above are disjointed extracts from an article in a paper—Fill up the gaps in units marked—.

Re-arrange the information using suitable tabular form, headings etc.

Comment on industrial production using the relevant figures only, and not exceeding 20 lines.

COMPUTER'S CERTIFICATE EXAMINATION, MAY 1948.

PART 1A : SECTION 1.

Full marks : 100

Time allowed : 2½ hours

(No machine allowed)

1. Evaluate

(a) $\frac{2ax^3}{(a-x)^2} - \frac{6\sqrt{ax}}{a^2\sqrt{2a+4x}} - \frac{20x^2}{64a}$ when $a = 4$ and $x = 2$

(b) $a^2 + b^2 + c^2 - 3abc$

when $a = 3$, $b = 4$, $c = 5$ and $a = 3/2$, $b = 5/2$, $c = 7/2$

(c) $3.1347 + 7.032 + .07 + 1.345 + .0079$

(d) $\frac{6.501 - 3.07 + 2.124}{5.5}$

2. Three boys agree to start together and run until all come together again, round a circular court 15 yds. in circumference. One runs at the rate of 6, the second 7 and the third 8 miles an hour. In how many seconds will the race end.

3. Copy out the following table correcting any mistakes you can detect.

Province	Population		Increase or Decrease in population (1921-1931) in per cent
	1931	1921	
Madras	46740107	42318985	+ 15.0
Bombay	17992053	16012342	+ 11.8
Sind	3887070	3279377	+ 18.5
Bengal	50114002	46702307	+ 9.2
U. P.	48408763	45375009	+ 6.7
Punjab	23580852	20685478	+ 10.0
Bihar	32371434	29026545	+ 11.0
Orissa	5306142	4908873	+ 8.0
C. P.	15507723	13912700	+ 2.8
Assam	8622251	7459128	+ 15.6
N. W. F. P.	2425076	2251340	+ 7.7
Coorg	163327	16383	- 5.0
Delhi	636240	488452	+ 3.2

Exhibit by means of any suitable diagram the increase or decrease in population (in per cent).

4. Complete the following calculations.

m	n	m+n	m-n	m ²	n ²	(m+n) ²	(m-n) ²	4mn.
14	23							
47	34							
9	12							
0	6							
18	19							
64	57							
81	78							
13	21							
54	48							
9	14							
28	33							
10	23							

Present the following in tabular form.

In 1940-41 the total production in India (in thousand tons) of the principal oil seeds were as follows : groundnuts 3702 ; linseed 434, rape and mustard 1103, castor 105, sesamum 433. Next year the production of each of the first three items fell by 36 per cent and the remaining items fell only by 10 per cent each. In 1942-43 there was an increase compared to the preceding year of 8 per cent in groundnuts, 12 per cent in linseed, 1 per cent in rape and mustard, 50 per cent in castor, and 10 per cent in sesamum. In the next year the figures were respectively 3823, 395, 955, 140 and 447 for groundnuts, linseed, rape and mustard, castor and sesamum.

COMPUTER'S CERTIFICATE EXAMINATION, MAY 1948.

PART 1A: SECTION 2.

Full marks : 100

Time allowed : 3 hours

(No machine allowed. No credit will be given for incomplete or scrappy answer).

1. The following table gives the average consumption in seers per head per month of the various food items, as in the year 1941, in the Jagadlal area and the price rates prevalent in 1941 and in 1945 in rupees per seer is given in the next two columns. With the year 1941 as base find the Index of price in 1945 for each item, as also the weighted Index of Food price.

Items	Consumption per head per annum in seers	Price rates in rupees per seer	
		1941	1945
1. Cereals	217	0.18	0.49
2. Pulses	35	0.15	0.54
3. Potato	32	0.06	0.32
4. Salt & Spices	8	0.10	0.19
5. Sugar	7	0.27	0.50
6. Ghee & Butter	2	1.52	3.20
7. Milk & Products	22	0.19	0.57
8. Vegetable oil	7	0.43	1.31
9. Fish	7	0.30	1.13
10. Meat	6	0.40	1.08

2. The following table gives the distribution of 200 factory workers paid on an output basis, according to their earnings in a particular month. Find the average earning per worker per month. If an incentive payment of extra 75% on all earnings in excess of Rs. 45/- is declared, what is the expected earning per worker per month on the average, assuming that the workers do not increase their output but retain their old habits. Find out the ratio of expected average to the present one.

Earnings per month	Number of workers
Rs. 31—Rs. 35	18
Rs. 36—Rs. 40	50
Rs. 41—Rs. 45	34
Rs. 46—Rs. 50	32
Rs. 51—Rs. 55	25
Rs. 56—Rs. 60	16
Rs. 61—Rs. 65	11
Rs. 66—Rs. 70	5
Rs. 71—Rs. 75	2
Rs. 76—Rs. 80	1
Total	200

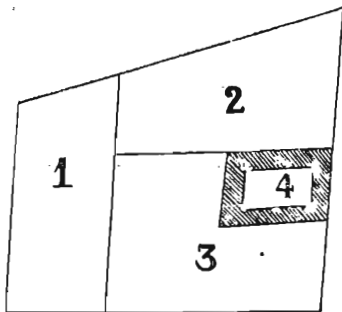
Or.

Below is a graphed sketch of a land-holding, comprising 4 plots, one small division on the graph representing 8 feet on the ground. The plot numbered 1, is a 'Sali' land, valued at Rs. 210 per cottah; plot 2 is 'Soona', valued at Rs. 175/- per cottah, plot 3 is a 'Bastoo' land, valued at Rs. 625/- a cottah, and plot 4 is a Tank valued at Rs. 190 per cottah.

Find: (a) Total value of the land

(b) What should be the overall flat rate per cottah, in case plot 4 is to be reserved as a public tank, and the whole value is to be realised by selling the remaining three plots.

N.B. One standard 'bigha' measures $120' \times 120'$ and contains 20 'cottahs'.



3. (a) Solve: $3x + \frac{24}{r}$

(b) With the help of tables find the value of the following expression correct upto 4 places of decimal.

$$\left[\frac{113.7091 \times (1.030918)^{3.5}}{1.11111} \right]^{2.5}$$

4. In an examination of 2357 candidates, the following was the distribution of the marks scored.

Scores	Number of candidates	Scores	Number of candidates	Scores	Number of candidates
0-5	12	31-35	205	61-65	190
6-10	21	36-40	234	66-70	132
11-15	41	41-45	264	71-75	85
16-20	67	46-50	262	76-80	68
21-25	103	51-55	263	81-85	37
26-30	118	56-60	232	86-90	16
				91-95	7

- (a) Calculate the standard error of the Mean Score.
 (b) Draw a smooth frequency curve by hand and
 (c) Find out, within what range of the Mean Score 67.24% of the candidates lie?

COMPUTER'S CERTIFICATE EXAMINATION, MAY 1948.

PART 1B: SECTION 1.

Full marks: 100

Time allowed: 3 hours.

(Machines & Tables allowed. No credit will be given for incomplete or scrappy answer.)

1. The frequency distribution of hourly earnings of 6201 factory workers is given below:

Hourly earnings (in cents)	frequency	Hourly earnings (in cents)	frequency
27.5-32.5	120	72.5-	915
32.5-	152	77.5-	745
37.5-	170	82.5-	530
42.5-	214	87.5-	259
47.5-	410	92.5-	152
52.5-	429	97.5-	107
57.5-	568	102.5-	50
62.5-	650	107.5-	25
67.5-	795		
			6201

Calculate the median, quartile points, and the points above and below of which ten per cent of cases lie.

2. The following table gives the number of points obtained in an efficiency test performed on 100 workers of different ages.

Calculate from the table the correlation coefficient of age and number of points.

Age (years)	Points	Age (years)	Points	Age (years)	Points	Age (years)	Points
67	51	34	79	36	78	48	62
49	77	31	78	61	55	34	63
43	67	45	78	38	76	32	69
27	79	61	61	29	69	47	60
36	76	59	59	46	81	30	80
32	53	30	88	33	77	45	90
49	81	48	63	11	81	53	95
36	83	37	77	47	78	47	74
30	90	46	93	50	63	28	85
35	86	35	79	38	82	36	72
46	73	43	69	34	81	62	61
69	54	50	73	28	91	44	80
43	60	60	72	39	67	47	66
42	72	49	58	44	60	37	92
39	72	42	76	42	78	31	73
34	84	24	82	38	68	59	73
41	86	43	88	50	72	40	70
46	70	48	65	31	74	37	67
26	81	59	69	53	81	36	71
61	58	46	77	41	76	29	88
27	83	58	61	41	85	50	82
55	55	60	77	34	82	30	84
42	74	51	53	43	83	43	75
27	91	33	75	56	79	43	77
43	78	37	86	42	85	25	88

3. Construct a two way frequency table from the data given in Q. 2 with class intervals 5 for each of the variables.

(a) Denoting Age by x and Number of points by y calculate the mean values of y 's corresponding to the mid point of each of the class intervals of x . Fit by method of least squares, a straight line to these values of x and y .

(b) Similarly calculate the mean values of x 's corresponding to the mid point of each of the class intervals of y . Fit by method of least squares, a straight line to these values of y and x .

4. The following table gives the number of females in the district of Hooghly of different marital conditions at different age groups.

Age group	Number of females		
	unmarried	married	widowed
0-5	81946	149	1195
5-10	81589	2393	765
10-15	37573	23449	1239
15-20	5431	56772	2806
20-25	1052	57446	6854
25-50	517	47636	10012
30-35	407	38277	14951
35-40	263	26672	16600
40-45	263	18348	16041
45-50	210	9444	16596
50-55	186	4598	15739
55-60	103	2142	13951
60-65	156	1218	10189
65-70	54	664	6276
70 & over	156	618	8828

Represent the data in a suitable diagram

COMPUTER'S CERTIFICATE EXAMINATION, MAY 1948.

PART IB SECTION 2.

Full marks : 100

Time allowed : 3 hours.

(Credit will be given for neatness and precise answer).

1. The distribution of 86 scores in a memory test is given below.

Scores	Frequency
9-10	0
11-13	4
14-16	3
17-19	8
20-22	13
23-25	9
26-28	17
29-31	14
32-34	8
35-37	5
38-40	4
41-43	1
44-46	0

The mean of the above distribution is 26.1 and its standard deviation is 6.45.

Find the expected frequencies in the above class intervals for a normal distribution with the same mean and S.D. as above.

Are the discrepancies between the calculated and observed frequencies statistically significant?

2. (a) Name the publications in which the following information may be obtained
- Price of rice in Calcutta in August, 1939
 - Production of cotton in the district of Surat (India) in 1937
 - No. of municipal towns in the district of Midnapore, Bengal
 - Production of Pig Iron in India in May, 1936
 - No. of Buffaloes in the district of Gaya (Bihar)
- (b) Draw a pie chart to represent the production of wheat in India as shown in the following table :

Provinces	Production
Punjab	3,526,000 tons
U.P.	2,980,000 ..
C.P.	1,587,000 ..
Bombay	864,000 ..
Rest of India	1,793,000 ..
Total	10,750,000 tons

3. (a) The number of bacteria in a given culture 't' hours after they were first observed was found to be as follows :

t	N
0	125
1	209
2	340
3	651
4	924
5	1525
6	1512

If the above data can be graduated by $N = 125 (1.649)^t$, calculate the theoretical values of N at each stage as shown above.

(b) Draw two curves to show the agreement between the actual and calculated values of the number of bacteria.

4. (a) A sample of 900 small rods was found to have a mean length of 3.4 cms. Could it be reasonably regarded as a sample from a large population with mean = 3.25 cms. and S.D. = 2.61 cms.?

(b) The mean height of 6400 Englishmen was found to be 67.85 inches with S.D. 2.56 inches and the mean height of 1600 Australians was found to be 68.55 inches with S.D. 2.52. Do the data indicate that Australians are on an average taller than Englishmen?

COMPUTER'S CERTIFICATE EXAMINATION, MAY 1948.

PART 1B : SECTION I. Supplementary

Full marks : 100

Time allowed : 3 hours.

(Machines and Tables allowed. No credit will be given for incomplete and scrappy answer.)

1. The weights in pounds of 330 working class children are given by the following distribution.

Weight in pounds	Frequency
50-55	2
55-	3
60-	11
65-	34
70-	84
75-	65
80-	48
85-	33
90-	30
95-	11
100-	4
105-	3
110-	1
115-120	1
	330

(a) Calculate the mean, standard deviation and β_1, β_2 .

(b) Represent the distribution in a suitable diagram, and from this diagram determine the percentages of cases falling between mean \pm S.D. and Mean \pm 2 S.D.

2. Prices of a manufactured good (X) and the corresponding prices of the raw material used in its manufacture (Y) are given in the table below. Construct a two way table with class intervals 50 for X and 10 for Y.

Calculate from the table the coefficient of correlation between X and Y.

X - 334	390	351	302	516	372	510	429	551	323	421	566
Y - 09	78	89	66	102	76	166	85	98	76	80	121
X - 562	303	291	414	388	342	351	426	646	313	290	412
Y - 145	50	43	80	83	87	80	85	127	75	69	71
X - 310	502	341	332	444	358	292	467	307	442	371	300
Y - 74	91	71	66	62	73	47	89	73	70	71	75
X - 419	402	302	261	333	308	433	264	440	288	302	492
Y - 80	88	67	55	61	60	72	52	72	59	76	124
X - 251	344	440	394	432	401	344	373	372	448	475	302
Y - 52	79	88	73	78	88	71	76	88	89	64	60
X - 374	320	354	278	280	584	296	424	424	414	267	500
Y - 72	72	78	55	69	129	67	85	63	88	55	61
X - 567	314	272	298	332	344	320	301	200	315	378	316
Y - 78	74	84	91	90	62	64	66	68	70	69	78
X - 324	424	302	411	302	268	285	316	434	282	261	354
Y - 75	86	64	77	68	70	85	75	81	66	51	70
X - 402	384	341	203	347	511	423	254	364	444	257	453
Y - 87	87	79	63	75	104	88	55	71	83	57	73

3. The following table shows what the consumption per capita of a certain article has been corresponding to the prices of that article ruling at different periods.

Denoting price as X and consumption as Y fit a straight line by the method of least squares to the data, of the form

$$(i) Y = a + bx$$

$$(ii) X = a' + b'y$$

Price (annas)	Consumption per capita (lbs)
6.0	27.0
6.3	27.1
6.5	21.2
7.0	18.6
7.2	18.6
7.6	18.0
8.0	19.3
8.1	17.1
8.4	16.6
8.6	11.3
10.0	7.9
10.5	3.7

4. In order to chalk out a plan of rehabilitating the refugee families from the Western Pakistan, the following information is needed

1. Name of the Head of the family
 2. Present address
 3. Home district
 4. Number of persons in the family, children, adults, males and females.
 5. Religion and caste
 6. Number of former earners (who are still living) in different occupations; agriculture, business, service, misc.
 7. Yearly income from different occupations
 8. Total landed property owned by the family
 9. Which province in India is preferred by the family for rehabilitation.
 10. Number of persons killed or missing in the family, children, adult, male, female.
 11. Number of earners killed or missing and their annual income.
- Prepare a suitable form in which this information may be recorded.

COMPUTER'S CERTIFICATE EXAMINATION, MAY 1948.

PART 1B : SECTION 2. SUPPLEMENTARY

Full marks : 100

Time allowed : 3 hours.

(Credit will be given for neatness and precise answer)

1. Marks obtained by 179 examinees are represented by the following frequency distribution :

Scores	frequency
34-37	1
38-41	0
42-45	3
46-49	7
50-53	20
54-57	32
58-61	39
62-65	21
66-69	34
70-73	5
74-77	8
78-81	8
82-85	1

The mean of the above distribution is 61.1 and the standard deviation is 8.4.

Fit a normal curve to the above data. Is the theoretical curve (obtained) a good fit to the observed data? Support your statement.

2. (a) The following table represents the sale of shares in the N. Y. stock exchange during different years :

1921	14,300,000 shares
1923	19,770,000 ..
1925	37,090,000 ..
1927	48,080,000 ..
1929	93,750,000 ..

If the number of shares sold is graduated by $(13.956,000)(1.178)^x$ when x represents the number of years that elapsed after 1910, calculate the number of shares sold in 1922, 1924 and 1928.

(b) Draw graphs showing the actual and calculated values for the number of shares sold between 1921 and 1929.

3. (a) Draw bar diagrams to represent the following facts :

India imported 223,142 tons of sugar in 1935, 201,158 tons in 1936, 23,075 tons in 1937, 14,389 tons in 1938 and 35,679 tons in 1939. Of the total quantities the share of Java was 176,146 tons in 1935, 150,416 tons in 1936, 15,022 tons in 1937, 10,293 tons in 1938 and 24,510 tons in 1939 and the rest came from various other countries.

(b) Name the publications from which the following information may be obtained:

- Total population of Rangpur town
- Exports of Jute from Bengal ports
- Imports of hides into Calcutta from other parts of India during 1938
- Total number of labourers employed in the Jute Mills of Bihar
- Acreege under Paddy in the district of Burdwan in 1941.

4. (a) The mean lengths of two samples of beams of 1000 and 2000 are 67.5 inches and 68.0 inches respectively while the standard deviations are 15.5 and 17.4 inches respectively. Can the samples be regarded as drawn from two (normal) populations with the same mean?

(b) A distribution of 150 people in normal condition gave an average pulse rate of 79.88 ± 0.15 beats per minute but after being administered a certain drug they showed an average pulse rate of 81.12 ± 0.20 beats per minute. Is it probable that the increase in pulse rate is due to the drug or is the result due to a variation in the sampling?

COMPUTER'S CERTIFICATE EXAMINATION, MAY 1948.

PART IC : SECTION I.

Full marks : 100

Time allowed : 3 hours

(Machines and Tables allowed. No credit will be given for incomplete or scrappy answer.)

1 (i) Fit the cubic $Y = a_0 + a_1x + a_2x^2 + a_3x^3$ to the following data. Calculate the expected values and draw the fitted curve through the points corresponding to the observed values.

x	1	2	3	4	5	6	7
y	10.16	12.00	13.90	15.01	17.93	20.07	22.71
x	8	9	10	11	12	13	
y	26.97	29.00	32.53	36.07	37.89	39.95	

(ii) Solve graphically the following equation

$$a_0 + a_1x + a_2x^2 + a_3x^3 = 3.333x + 1.667$$

where a_0, a_1, a_2, a_3 have the same values as in the fitted curve under (i).

2. The following table gives the yields of three varieties of paddy, each variety being treated with four kinds of manure. The varieties are V_1, V_2, V_3 and Manures are M_1, M_2, M_3, M_4 . The experiment was laid out in a randomized block design. Test the significance of interaction between Variety and Manure.

Blocks	V_1				V_2				V_3			
	M_1	M_2	M_3	M_4	M_1	M_2	M_3	M_4	M_1	M_2	M_3	M_4
I	96	187	222	109	146	262	246	277	115	208	220	430
II	70	163	125	97	133	181	263	293	143	220	341	371
III	77	143	134	138	164	224	104	260	117	234	268	484
IV	80	179	173	113	140	248	190	325	120	253	297	460

3. Assuming the correctness of the first two columns check the following calculation correcting mistakes whenever found. Thence calculate the first four corrected moments.

ξ	f	$f\xi$	$f(\xi+1)$	$f\xi^2$	$f(\xi+1)^2$	$f\xi^3$	$f(\xi+1)^3$	$f\xi^4$	$f(\xi+1)^4$
-10	2	-20	-18	200	162	-2000	-1458	20000	13122
-9	4	-36	-32	324	256	-2916	-2048	26244	16384
-8	14	-112	-98	896	686	-7168	-4802	57344	33614
-7	41	-287	-246	309	1476	-1463	-8856	59992	53136
-6	83	-498	-415	2998	2075	-17928	-10375	107568	61875
-5	169	-846	-676	4225	2704	-21125	-10816	105625	43264
-4	394	-1576	-1182	6304	3546	-25216	-10638	100864	31914
-3	669	-2007	-1338	8021	2676	-18063	-5352	54189	10704
-2	990	-1980	-990	3960	890	-7920	-990	15840	990
-1	1223	-1223		1223		-1223		1223	
0	1329		1329		1329		1329		1329
1	1230	1230	2460	1230	4920	1230	9840	1230	19680
2	1063	2126	3189	4252	9567	8504	28701	17008	86103
3	646	1938	2584	5814	10336	17442	41344	52326	165376
4	392	1568	1960	6272	9800	25088	49000	106352	245000
5	202	1010	1212	5050	7272	25250	43632	126250	261792
6	79	474	553	2844	3871	17064	27697	102384	189679
7	32	224	256	1568	2948	10976	16384	76832	131072
8	16	128	144	1042	1296	8336	11664	66688	104976
9	5	46	50	405	500	3645	5000	32805	50000
10	2	20	22	200	242	2000	2662	20000	29282
	8585	8763	13759	64910	65752	107245	236653	1142061	1539292

COMPUTER'S CERTIFICATE EXAMINATION, MAY 1948.

PART 1C: SECTION 2.

Full marks: 100

Time allowed: 2 hours

(Machine & Tables allowed. No credit will be given for incomplete or scrappy answer.)

1. Evaluate correct to three places of decimals.

$1/2.273$	$\sqrt{22.61}$	$\log 0.97$
$1/1.035$	$\sqrt{54.98}$	$\log 5.87$
$1/0.683$	$\sqrt{124.10}$	$\log 24.37$

2. The value of a function corresponding to the values 1, 2, 3, 4, 5 of the argument are 0.198669, 0.237702, 0.276355, 0.314566, 0.352274 respectively. Calculate the value of the function when the argument has the value 1.25.

3. Analyse the results of the following Randomized Block Experiment.

Block I			Block II		
E	F	B	C	B	F
26.4	35.1	40.8	45.1	41.2	32.9
D	A	C	E	A	D
26.4	38.0	37.9	27.9	38.7	24.3
F	B	D	E	D	A
32.7	40.0	25.2	25.0	29.8	48.7
A	C	E	C	B	F
37.5	32.9	27.2	47.0	49.7	36.5
Block III			Block IV		

Are all the treatments equally effective ?

Is the treatment F better than D ?

Show your knowledge of the use of a table of random numbers by assigning six treatments to the several plots in a lay-out of the above type.

4. The following monthly data represent bank debits deflated for price changes (5 ciphers omitted) in a certain country. Calculate an index of seasonal variation.

Month	1923	1924	1925	1924	1927	1928	1929	1930
January	1756	1606	1877	1783	1754	1572	1647	1833
February	1434	1445	1494	1481	1509	1432	1433	1543
March	1962	1745	1855	1855	1897	1749	1724	1785
April	1734	1616	1762	1742	1712	1600	1699	1801
May	1728	1678	1609	1664	1654	1683	1681	1792
June	1871	1552	1626	1709	1719	1757	1646	1737
July	1651	1604	1685	1719	1674	1618	1721	1701
August	1552	1451	1540	1536	1508	1593	1694	1584
September	1545	1563	1537	1587	1641	1584	1629	1615
October	1766	1757	1821	1767	1745	1762	1864	1758
November	1545	1471	1550	1591	1568	1535	1735	1608
December	1662	1547	1705	1819	1651	1669	1725	1650

5. You are asked to calculate on some old materials

(1) β_1 and β_2 ,

(2) $r_{1,2}$.

On going through the working sheets you find the following :

(1) Moments about an arbitrary origin,

$$\mu_1' = -1.0047, \mu_2' = 2.5279, \mu_3' = -10.2768, \mu_4' = 25.7387$$

(2) The total correlation coefficients

$$r_{1,2} = 0.9387, r_{1,3} = 0.7046, r_{2,3} = 0.3053$$

Are these figures of any help to you ?