CHAPTER SEVEN

STATISTICAL ANALYSIS OF THE DATA

Statistical Analysis of the Data

Data collected from the test have little meaning until they have been classified in a systematic way. The first task that confronts us, then, is the organization of our material and this leads naturally to a grouping of the scores into classes.

Analysis of the Total Test Scores

The maximum score that a testee can obtain on this test is 148 and the lowest score that can be obtained is zero. The highest score that is obtained on the test is 114 while the lowest score is 17. The range between the highest and the lowest scores is therefore 98. This range within which the scores are distributed is divided into eleven class intervals each interval being of 10 units. The distribution of the scores is given in Table No.9.

The frequency distribution of the scores obtained by the pupils in each of the sub-tests are also tabulated separately in Table Nos. 10 to 16.

Table :9: Frequency Distribution of the Scores - Whole Test

Scores	Mid-points	f	Cum.f	d	fd	fd^2
111-120	110.5	8	2000	+5	40	200
101-110	105.5	48	1992	+4	192	768
91-100	95.5	132	1944	+3	396	1188
81-90	85.5	252	1812	+2	504	1008
71-80	75.5	384	1560	+1	384	384
61-70	65.5	426	1176	0	0	0
51-60	55.5	390	750	-1	-390	390
41-50	45.5	230	360	-2	-460	920
31-40	35.5	99	130	-3	-297	891
21-30	25.5	27	31	-4	-108	432
11-20	15.5	4	4	-5	- 20	100
		N = 2000			$\Sigma fd = +241$	$\Sigma fd^2 = 6281$

Calculation of the Mean by the 'Assumed Mean' method

$$C = \frac{\Sigma fd}{N} = \frac{+241}{2000}$$

True mean = Assumed mean + Correction factor (Ci)

$$= 65.5 + \frac{241}{2000} \times 10$$

$$= 65.5 + 1.21$$

$$= 66.71$$

$$\text{Mdn} = 1 + \left(\frac{N/2 - F}{f_m} \right) i$$

$$= 60.5 + \left(\frac{1000 - 750}{426} \right) 10$$

$$= 60.5 + \frac{250 \times 10}{426}$$

$$= 60.5 + 5.87$$

$$= 66.37$$

$$\text{Mode} = 3\text{Mdn} - 2\text{Mean}$$

$$= 3 \times 66.37 - 2 \times 66.71$$

$$= 65.69$$

Calculation of the Standard Deviation

$$\sigma = i \times \sqrt{\frac{\Sigma fd^2}{2000} - \left(\frac{\Sigma fd}{N} \right)^2}$$

$$= 10 \sqrt{\frac{6281}{2000} - \left(\frac{241}{2000} \right)^2}$$

$$= 10 \sqrt{3.13}$$

$$= 10 \times 1.769$$

$$= 17.69$$

Table :10: Frequency Distribution of Scores - Sub-Test 1

Sr.No.	Scores	Mid-points	Frequency	Cum. frequency
1	24-26	25	56	2000
2	21-23	22	118	1944
3	18-20	19	359	1826
4	15-17	16	452	1467
5	12-14	13	474	1015
6	9-11	10	329	541
7	6-8	7	165	212
8	3-5	4	40	47
9	0-2	1	7	7
			N = 2000	

Table :11: Frequency Distribution of Scores - Sub-test 2

Sr.No.	Scores	Mid-points	Frequency	Cum. Frequency
1	24-26	25	33	2000
2	21-23	22	126	1967
3	18-20	19	265	1841
4	15-17	16	374	1576
5	12-14	13	424	1202
6	9-11	10	368	778
7	6-8	7	257	410
8	3-5	4	120	153
9	0-2	1	33	33
			N = 2000	

Table :12: Frequency Distribution of Scores – Sub-test 3

Sr.No.	Scores	Mid points	Frequency	Cum. frequency
1	18-20	19	1	2000
2	15-17	16	39	1999
3	12-14	13	289	1960
4	9-11	10	584	1671
5	6-8	7	623	1087
6	3-5	4	387	464
7	0-2	1	77	77

N = 2000

Table :13: Frequency Distribution of Scores Sub-test 4

Sr.No.	Scores	Mid points	Frequency	Cum. frequency
1	18-20	19	15	2000
2	15-17	16	180	1985
3	12-14	13	470	1805
4	9-11	10	700	1335
5	6-8	7	430	635
6	3-5	4	170	205
7	0-2	1	35	35

N = 2000

Table :14: Frequency Distribution of Scores - Sub-test 5

Sr.No.	Scores	Mid point	Frequency	Cum. frequency
1	18-20	19	29	2000
2	15-17	16	189	1971
3	12-14	13	523	1782
4	9-11	10	644	1259
5	6-8	7	430	615
6	3-5	4	145	125
7	0-2	1	40	40

N = 2000

Table :15: Frequency Distribution of Scores- Sub-test 6

Sr.No.	Scores	Mid point	Frequency	Cum. frequency
1	18-20	19	10	2000
2	15-17	16	88	1990
3	12-14	13	432	1902
4	9-11	10	720	1470
5	6-8	7	530	750
6	3-5	4	180	220
7	0-2	1	40	40

N = 2000

Table :16: Frequency Distribution of Scores - Sub-test 7

Sr.No.	Scores	Mid point	Frequency	Cum. Frequency
1	12-14	13	3	2000
2	9-11	10	134	1997
3	6-8	7	783	1863
4	3-5	4	821	1080
5	0-2	1	259	259

N = 2000

RELIABILITY OF MEAN, MEDIAN AND STANDARD DEVIATION

The results obtained above of the parameters mean, median and standard deviation of the frequency distribution of the whole test are from the random sampling. These may deviate from the population parameters. We have tried to arrive at statistics that would approximate the corresponding parameters very closely, by selecting an adequate sampling. As no guarantee can be given for the reliability of the statistics, it is necessary to test the reliability of the same. The use of standard errors and other sampling statistics can be made to estimate how far our obtained statistics may have deviated from their corresponding parameters. The reliability of each of the above statistics is tested by calculating its standard error.

1. Standard error (SE) of the mean

SE_M or σ_M = $\frac{\sigma}{\sqrt{N}}$ where σ = the standard deviation of the total distribution

N = No.of cases in the sample.

$$= 17.69 / \sqrt{2000}$$

$$= 0.3956$$

The true mean lies between 66.71 ± 0.3956 x 2.58 at .01 level i.e. between 65.689 and 67.731. Thus the obtained mean is highly reliable as the true mean lies within the narrow range.

2. Standard error (SE) of the median

$$SE_{Mdn} = \frac{1.253\ \sigma}{\sqrt{N}}$$

$$= \frac{1.253 \times 17.69}{\sqrt{2000}}$$

$$= 0.4956$$

∴ The true median lies between 64.596 and 67.152 at 0.01 level. (i.e.) between 66.37 ± 0.4956 x 2.58. The median obtained is quite reliable.

3. Standard error (SE) of the standard deviation

$$SE_{SD} = \frac{0.71\ \sigma}{\sqrt{N}}$$

$$= \frac{0.71 \times 17.69}{\sqrt{2000}}$$

$$= 0.2808$$

$\therefore$ The true standard deviation lies between 17.69 $\pm$ 0.2808x2.58 at 0.01 level, i.e. between **16.966** and **18.414**. The standard deviation obtained is also highly reliable.

It can be concluded from these results that all the parameters lie within the narrow ranges and hence the results obtained are highly reliable.

Calculation of Skewness of the Distribution

There are two different formulae for the calculation of skewness. Skewness is calculated by using both of these formulae.

(1) $Sk = \frac{3(\text{mean} - \text{median})}{\sigma}$ .... (I)

and (2) $Sk = \frac{P_{90} + P_{10}}{2} - P_{50}$ .... (II)

Calculation of Sk by formula I

The values of mean, median and standard deviation of the distribution are :

Mean = 66.71

Mdn = 66.37

and SD = 17.69

$$Sk = \frac{3(\text{mean} - \text{median})}{\sigma}$$

(a measure of skewness in terms of frequency distribution).

$$= \frac{3(66.71 - 66.37)}{17.69}$$

$$= \frac{3 \times .34}{17.69}$$

$$= +\ 0.0576$$

The Skewness obtained is slightly positive.

Calculation of Sk by Formula II :

$$P_{90} = 89.96$$

$$P_{10} = 43.54$$

$$P_{50} = 66.37$$

$$Sk = \frac{P_{90} + P_{10}}{2} - P_{50}$$

(a measure of skewness in terms of percentiles)

$$= \frac{89.96 + 43.54}{2} - 66.37$$

$$= 66.75 - 66.37$$

$$= +0.38$$

The value of skewness obtained by this formula also indicates a positive value but slightly higher than the previous one. The results obtained by these two formulae differ slightly. According to 'Garrett' the two measures of skewness are computed from different reference values in the distribution and hence are not directly comparable.

Significance of Skewness

For concluding whether the obtained skewness is significant or not, the standard error of skewness should be known. The standard errors for formulae I and II used here, are not very satisfactory and Garrett states that " the measures of skewness as they stand are often sufficient for many problems in psychology and education. " But the author calculated the standard error of skewness by using formula II and from the value obtained it is concluded that the skewness is not at all significant.

$$\sigma_{Sk} = \frac{0.5185D}{\sqrt{N}} \quad \text{where } D = P_{90} - P_{10}$$

$$= \frac{0.5185D}{\sqrt{2000}}$$

$$= \frac{0.5185 \times (89.96 - 43.54)}{44.72}$$

$$= 0.5381$$

The skewness obtained by using the formula II is equal to +0.38

Critical Ratio

$$CR = \frac{+0.38}{0.5381} = +0.7062$$

The CR (Critical Ratio) falls within the limits $\pm$ 2.58 which determine the 0.01 level of significance. Hence it is clear that +0.7062 represents no real deviation of this frequency distribution from normality.

Calculation of Kurtosis of the distribution

The following formula is used for the calculation of the Kurtosis

$$Ku = \frac{Q}{P_{90} - P_{10}} \quad , \quad \text{where } Q = \frac{P_{75} - P_{25}}{2}$$

$$= \frac{(P_{75} - P_{25})/2}{P_{90} - P_{10}}$$

$$= \frac{(78.94 - 54.90) / 2}{89.96 - 43.54}$$

$$= \frac{12.02}{46.42}$$

$$= 0.2628$$

The kurtosis of the frequency distribution is thus equal to 0.2628. The value is slightly less than 0.263 and less by 0.0002. It indicates that the distribution is slightly leptokurtic.

Significance of Kurtosis

To estimate the significance of deviation of Ku thus obtqined from the Ku of the normal curve the SE of Ku is calculated by the formula given below :

$$\sigma_{Ku} = \frac{0.28}{\sqrt{N}}$$

$$\sigma_{Ku} = \frac{0.28}{\sqrt{44.72}} = 0.0062$$

and the CR (Critical Ratio)

$$= \frac{D}{\sigma_{Ku}}$$ where D is the deviation of Ku of the obtained distribution from Ku (0.263) of normal distribution.

$$= -\frac{0.0002}{0.0063}$$

$$= -0.03175$$

The CR (-0.03175) falls well within the ± 1.96 limits which determine the 0.05 level of significance. So it is concluded that the Kurtosis 0.2628 represents no real deviation of the frequency distribution from normality.

GRAPHICAL REPRESENTATION OF THE TEST SCORES

Aid in analysing numerical data may often be obtained from a graphic or pictorial treatment of the frequency distribution. The advertiser has long used graphic methods because these devices catch the eye and hold the attention when the most careful array of statistical evidence fails to attract notice. For this and other reasons the research worker has to utilise the attention-getting power of visual presentation, and at the same time seek to translate numerical facts, often abstract and difficult of interpretation, into more concrete and understandable form. The procedure suggested by 'Garrett' is followed in toto to represent the frequency distribution graphically.

Table :17: Showing Smoothed Frequencies of the Distribution

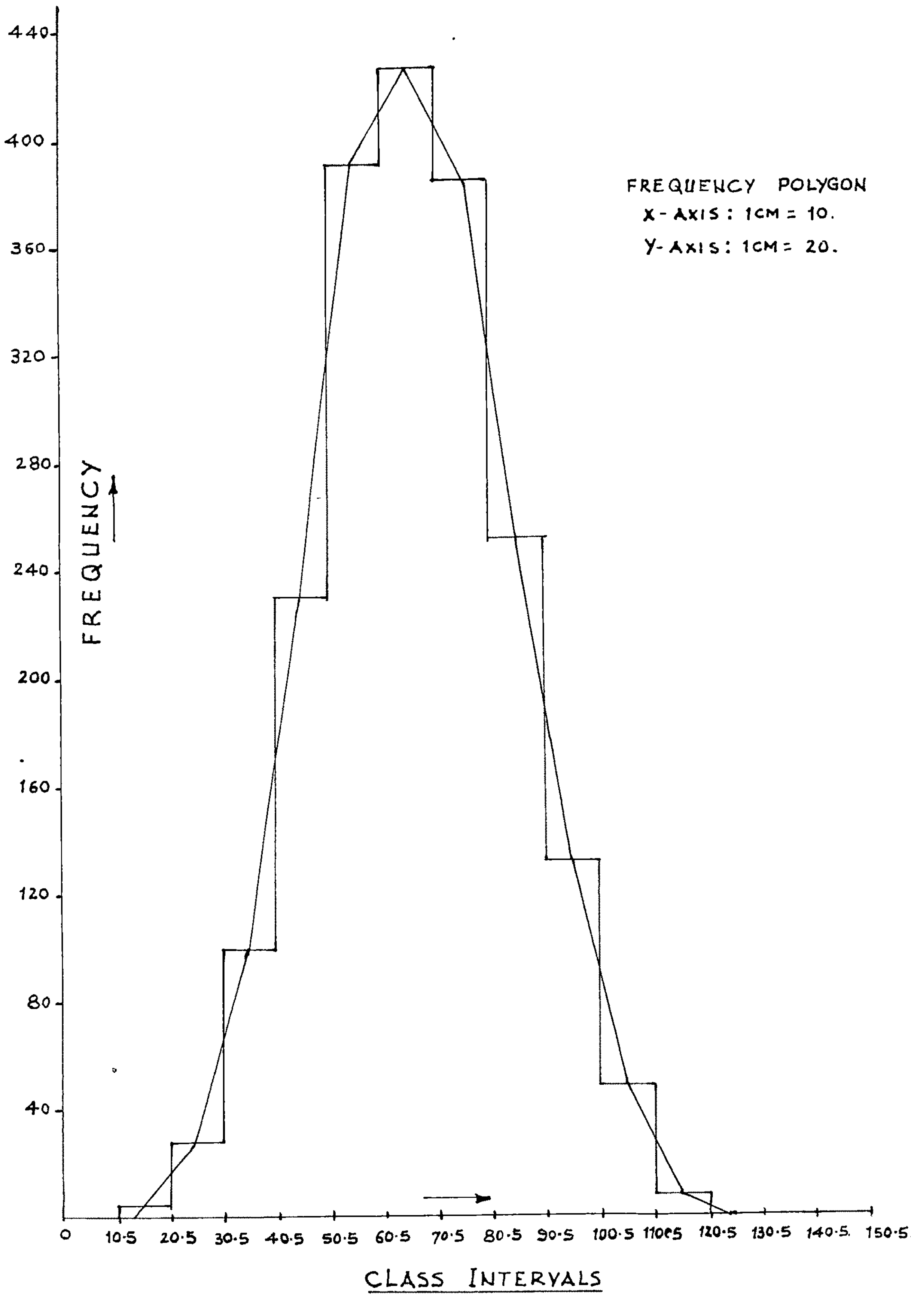
Scores	Original frequency	Smoothed frequency
111-120	8	18.66
101-110	48	62.66
91-100	132	144.00
81-90	252	256.00
71-80	384	354.00
61-70	426	400.0
51-60	390	348.66
41-50	230	239.66
31-40	99	118.66
21-30	27	43.33
11-20	4	10.33
	N = 2000	

The following curves have been drawn of the frequency distribution:

1. Frequency Polygon

Data used for drawing the frequency polygon are given in Table 9 on page 110. The polygon is constructed as per the method given by Garrett.[1]

[1] Garrett, H.E. Op.Cit., 10-13pp.


FREQUENCY POLYGON
X - AXIS : 1CM = 10.
Y - AXIS : 1CM = 20.
FREQUENCY
440
400
360
320
280
240
200
160
120
80
40
0
10·5
20·5
30·5
40·5
50·5
60·5
70·5
80·5
90·5
100·5
110e5
120·5
130·5
140·5
150·5.
CLASS INTERVALS

2. <u>The 'Smoothed' frequency Polygon</u>

In order to iron out chance irregularities the frequency polygon is 'smoothed' as shown in the graph , page 125. In smoothening a series of 'moving' or 'running' averages are taken from which new or adjusted frequencies are determined. The smoothed frequencies calculated are given in Table 17 on page 122. In smoothening the frequency polygon, the method is followed as per the suggestions given by Garrett.[1]

3. <u>Histogram</u>

A second way of representing a frequency distribution graphically is by means of a 'Histogram' or 'Column Diagram'. Even here the method given by Garrett[2] is used for the construction of the histogram. The graph is drawn on Page 126 .

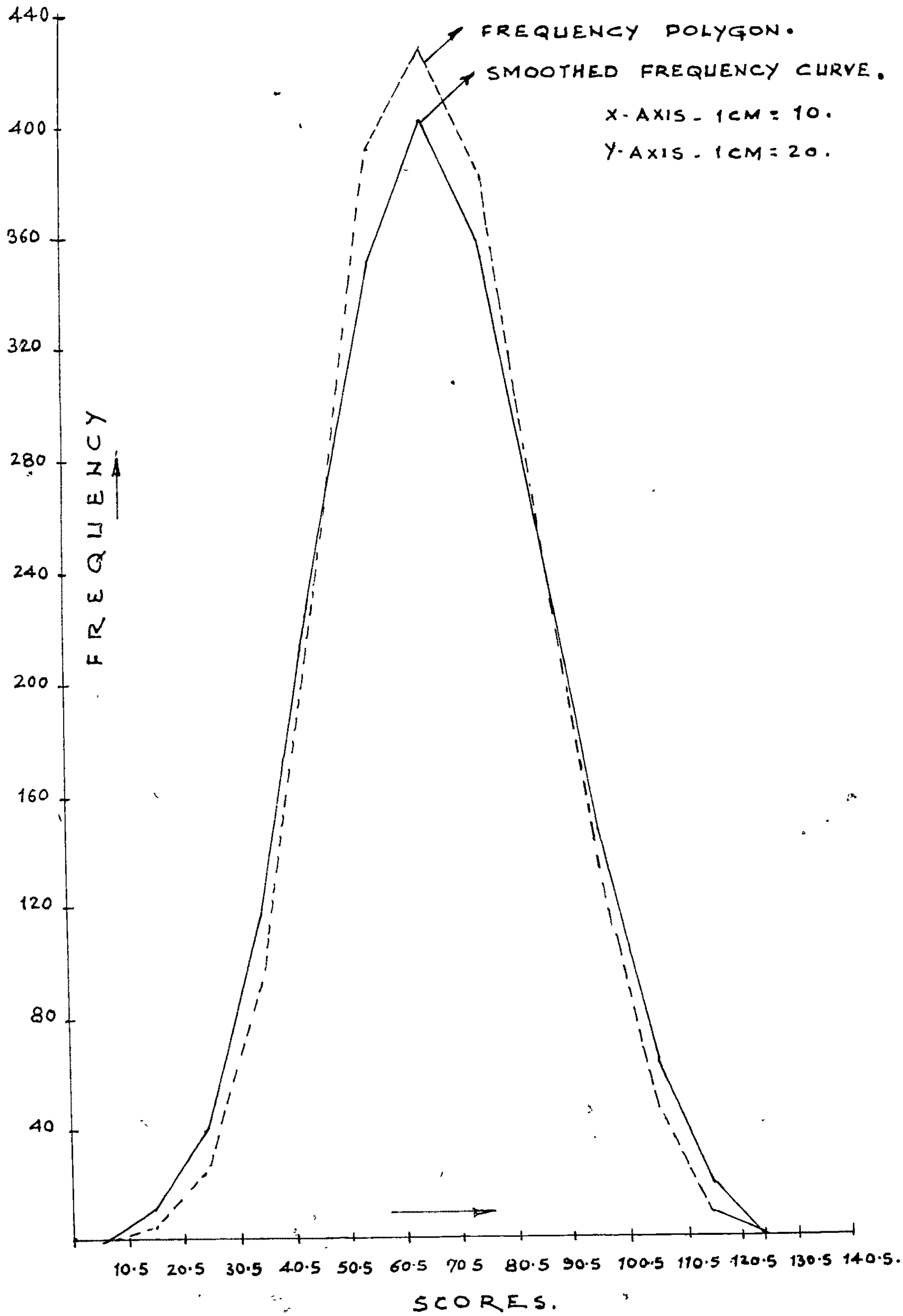
4. <u>Construction of the Cumulative Frequency graph</u>

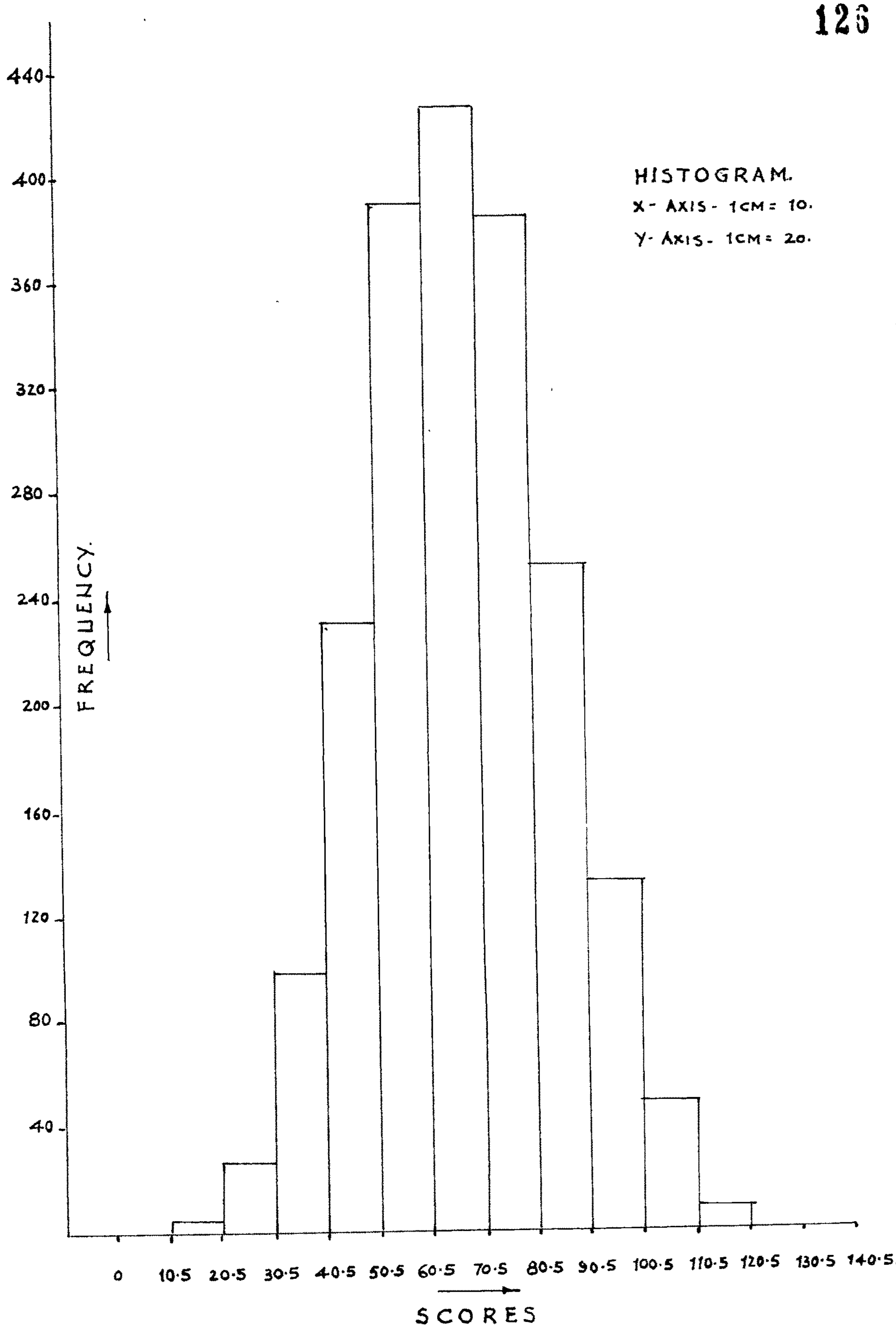
The cumulative frequency graph is another way of representing a frequency distribution by means of a diagram. Before the cumulative frequency graph is plotted, the scores of the distribution must be added serially or cumulated as shown in Table 18. The last cumulative frequency is equal to 2000, the total frequency. In cumulative frequency curve, each cumulative frequency is plotted at the exact upper limit of the interval upon which it falls. The plotted points are joined to give the S-shaped cumulative frequency graph No. , page 127

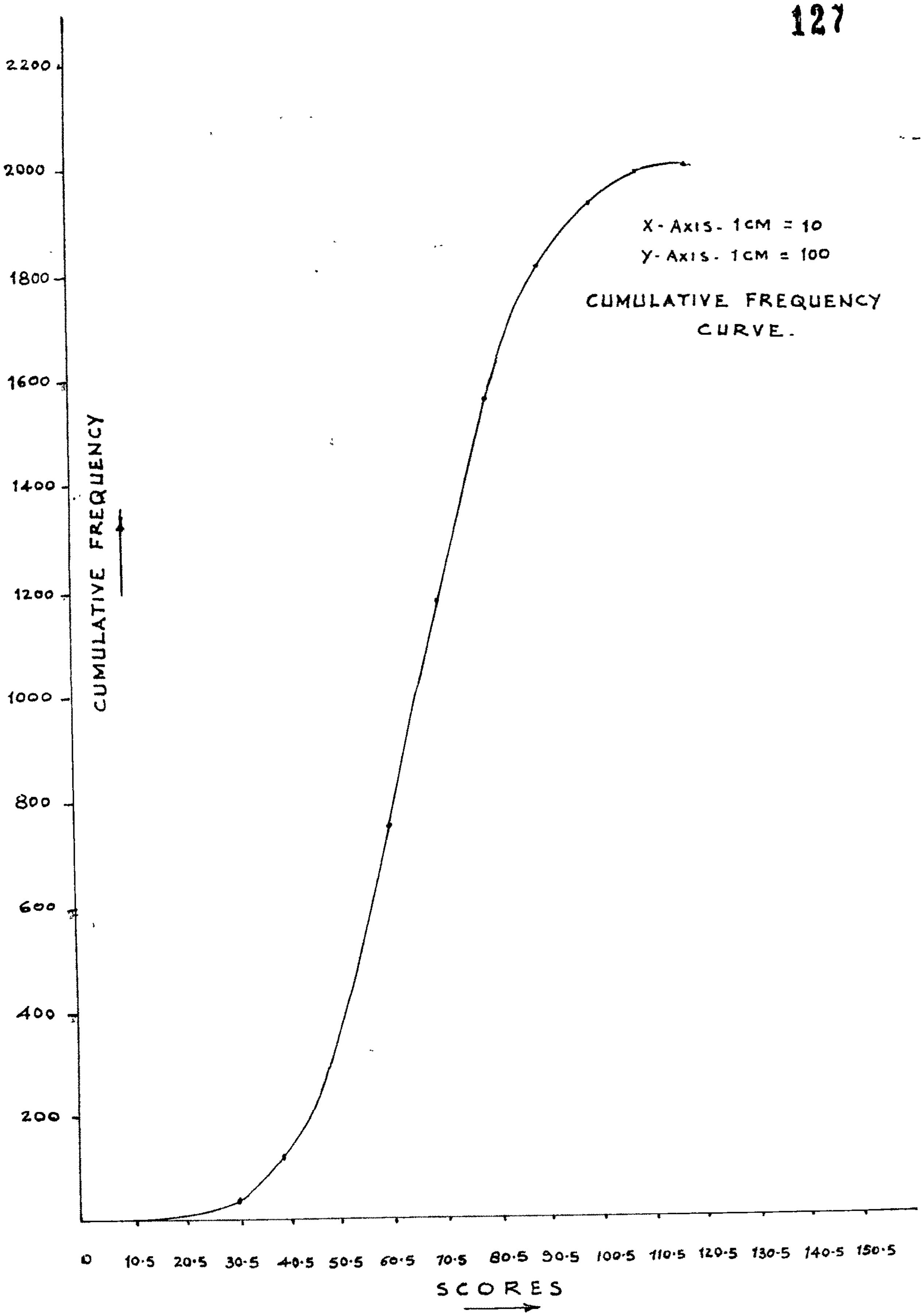
---

[1] Ibid., p.10

[2] Ibid., p.15.


FREQUENCY POLYGON.
SMOOTHED FREQUENCY CURVE.
X-AXIS - 1CM = 10.
Y-AXIS - 1CM = 20.
440
400
360
320
280
240
200
160
120
80
40
FREQUENCY
10·5 20·5 30·5 40·5 50·5 60·5 70·5 80·5 90·5 100·5 110·5 120·5 130·5 140·5.
SCORES.


HISTOGRAM.
X - AXIS - 1CM = 10.
Y - AXIS - 1CM = 20.
FREQUENCY
440
400
360
320
280
240
200
160
120
80
40
0
10·5
20·5
30·5
40·5
50·5
60·5
70·5
80·5
90·5
100·5
110·5
120·5
130·5
140·5.
SCORES

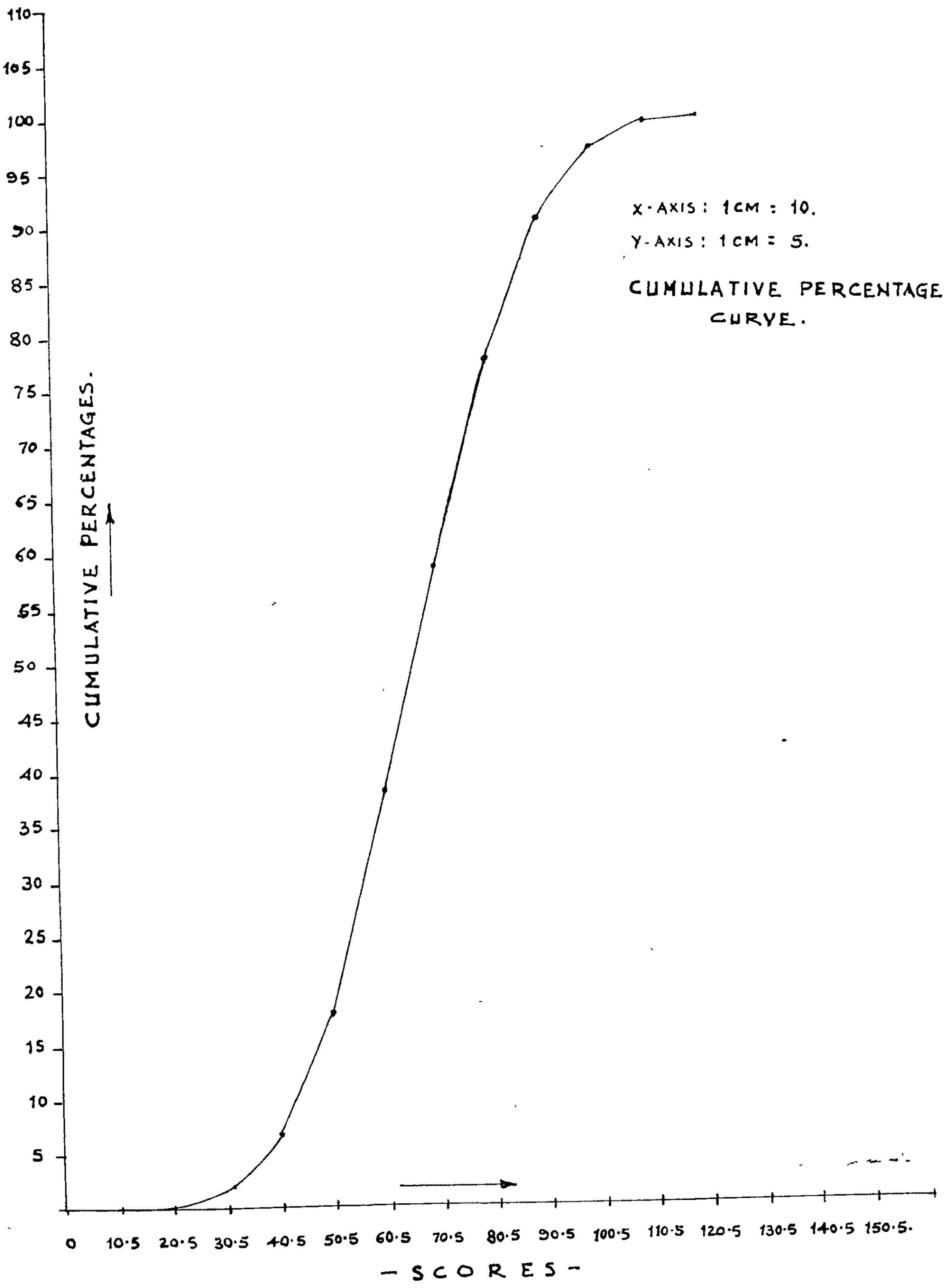

2200
2000
1800
1600
1400
1200
1000
800
600
400
200
CUMULATIVE FREQUENCY
X - AXIS - 1 CM = 10
Y - AXIS - 1 CM = 100
CUMULATIVE FREQUENCY CURVE.
0
10·5
20·5
30·5
40·5
50·5
60·5
70·5
80·5
90·5
100·5
110·5
120·5
130·5
140·5
150·5
SCORES

5. Cumulative Percentage Curve or Ogive

The cumulative percentage curve or Ogive differs from the cumulative frequency graph in that frequencies are expressed as cumulative percents of N on the Y-axis instead of as cumulative frequencies. The scores and the cumulative percents calculated are given in Table 18 below. The cumulative percents are plotted at the exact upper limit of the interval upon which it falls. The graph drawn is given on page 129.

Table :18: Cumulative and Cumulative Percent Frequencies

Scores	Frequency f	Cum.f.	Cum.Percent f.
111-120	8	2000	100.00
101-110	48	1992	99.60
91-100	132	1944	97.20
81-90	252	1812	90.60
71-80	384	1560	78.00
61-70	426	1176	58.80
51-60	390	750	37.50
41-50	230	360	18.00
31-40	99	130	6.50
21-30	27	31	1.55
11-20	4	4	0.20
	N = 2000		


X-AXIS : 1CM : 10.
Y-AXIS : 1CM = 5.
CUMULATIVE PERCENTAGE CURVE.
CUMULATIVE PERCENTAGES.
110
105
100
95
90
85
80
75
70
65
60
55
50
45
40
35
30
25
20
15
10
5
0
10·5
20·5
30·5
40·5
50·5
60·5
70·5
80·5
90·5
100·5
110·5
120·5
130·5
140·5
150·5.
— SCORES —

6. The Best fitting Normal Distribution Curve

The equation of the normal probability curve reads as follows :

$$Y = \frac{N}{\sigma\sqrt{2\pi}} \quad e^{-\frac{x^2}{2\sigma^2}}$$

where

x == scores (expressed as deviations from the mean) laid off along the base line or x-axis.

Y = The height of the curve above the x-axis that is the frequency of a given x-value.

The other tems of the equation are constants.

N = Number of cases

σ = Standard deviation of the distribution

π = 3.1416 (the ratio of the circumference of a circle to its diameter).

e = 2.7183 (base of the Napierian system of logarithms).

The best fitting curve is to be superimposed on the obtained histogram. To plot a normal curve over this histogram, the height of the maximum ordinate (y_o) should first be calculated. This can be determined from the equation of the normal curve as shown below :

The 'x' at the mean of the normal curve is '0'

When $x = 0$, $e^{-x^2/2\sigma^2} = 1$

$$\therefore Y_o = N / \sigma\sqrt{2\pi}$$

In the present test,

$$N = 2000 , \quad \sigma = 1.769 \quad \text{and} \quad \sqrt{2\pi} = 2.51$$

$$\therefore Y_o = \frac{2000}{1.769 \times 2.51}$$

$$= 450.4$$

So the value of Y_o = 450.4

The values of Y , the heights of the ordinates at different σ - distances from the mean are found out from the statistical table[1] and the corresponding values of Y when Y_o = 450.4 are computed. The final values of the ordinates at different σ-distances are given in Table 19.

Table :19: Showing Normal curve Ordinates at Mean

Y_o = 450.4

Mean = 66.71 ; σ = 1.769

σ distance from the Mean	Value of Y when y_o=1 (Read from the table)	Value of Y When Y_o=450.4 (obtained from the data)	Height of the ordinate
$\pm$ 0.5σ	0.88250	0.88250 x 450.4	397.5
$\pm$ 1 σ	0.60653	0.60653 x 450.4	273.2
$\pm$ 1.5σ	0.32465	0.32465 x 450.4	146.2
$\pm$ 2σ	0.13534	0.13534 x 450.4	60.94
$\pm$ 3σ	0.01111	0.01111 x 450.4	5.003

[1] Garrett, Op.Cit. p.447

The data given in Table 19 are used to super-impose the ideal (best-fitting) normal curve on the obtained histogram. The curve drawn is given on Page 133

The skewness of the distribution is found to be +0.38. The value indicates a low degree of positive skewness in the data. The kurtosis of the distribution is 0.2628 and the distribution is slightly leptokurtic. The divergence indicated is not at all significant of a 'real' discrepancy between the data and that of the normal distribution. The normal curve given on Page 133 on the whole fits in with the obtained distribution well enough to warrant our treatment of data as normal.

The normal curve ordinates at mean, $\pm 1\sigma$, $\pm 2\sigma$, $\pm 3\sigma$ distances for each of the seven sub-tests are calculated and given in Tables from 20 to 26.

Table :20: Showing Normal Curve Ordinates at Mean For Sub-Test 1

N = 2000, Mean = 14.46, σ = 1.48 (class interval units)

Y_o = 539.1

σ-distance from the mean	Value of y, when Y_o=1 (Read from the table)	Value of Y_o when Y_o=539.1 obtained from the data	Height of the ordinate
$\pm 1\sigma$	.60653	.60653 x 539.1	327.1
$\pm 2\sigma$	.13534	.13534 x 539.1	72.95
$\pm 3\sigma$	.01111	.01111 x 539.1	5.99

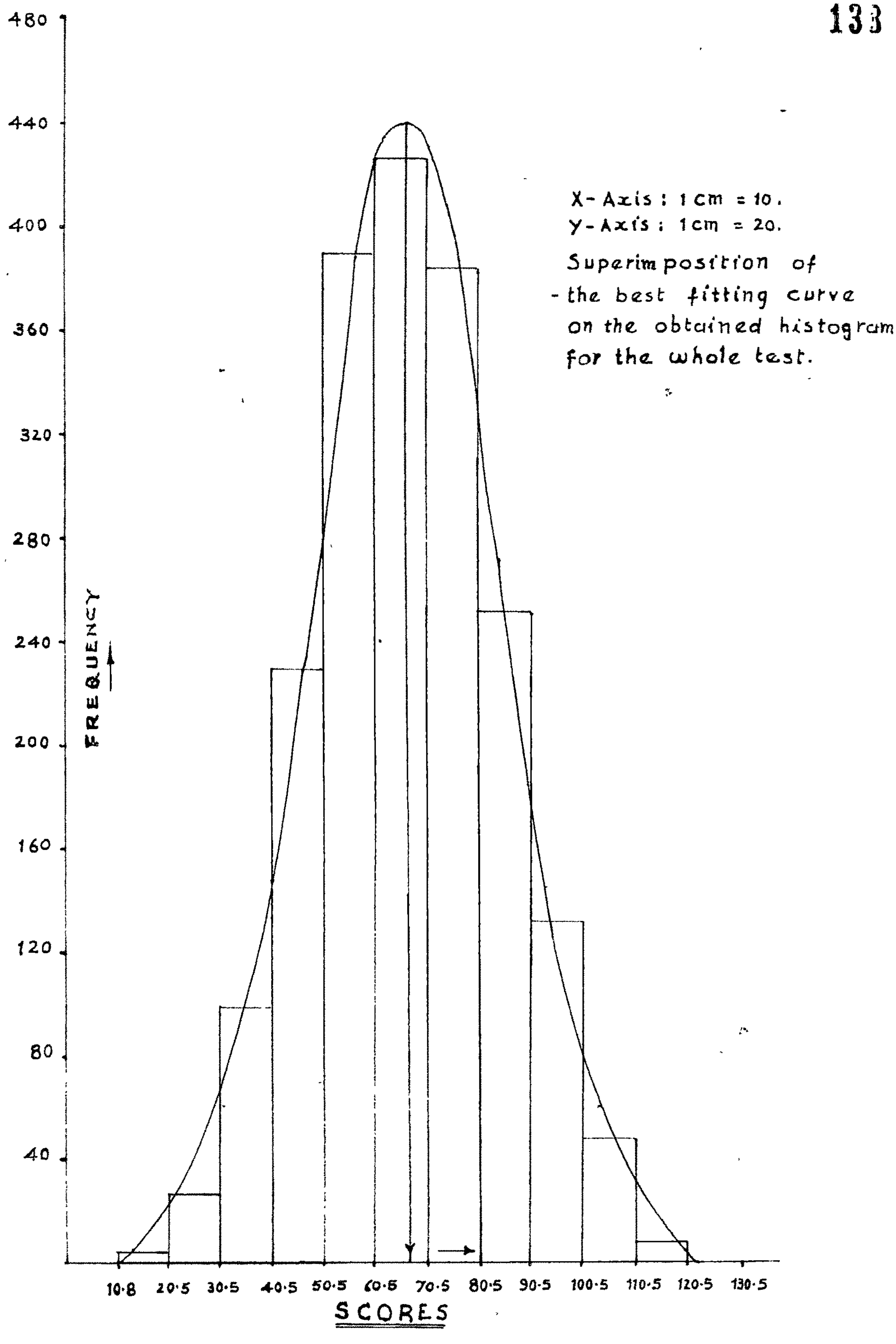

X-Axis : 1 cm = 10.
Y-Axis : 1 cm = 20.
Superimposition of the best fitting curve on the obtained histogram for the whole test.
FREQUENCY
SCORES
480
440
400
360
320
280
240
200
160
120
80
40
10·8
20·5
30·5
40·5
50·5
60·5
70·5
80·5
90·5
100·5
110·5
120·5
130·5

Table :21: Showing Normal Curve Ordinates at Mean For Sub-test II

N = 2000 ; Mean = 13.06 ; σ = 1.75 (interval units)

Y_o = 455.9

σ distance from the Mean	Value of y, when Y_o=1 (Read from tables)	Value of Y when Y_o=455.9 obtained from the data	Height of the ordinate
± 1 σ	.60653	.60653 x 455.9	276.6
± 2 σ	.13534	.13534 x 455.9	61.69
± 3 σ	.01111	.01111 x 455.9	5.065

Table :22: Showing Normal Curve Ordinates at Mean Sub-test III

N = 2000 ; Mean = 8.113 ; σ = 1.123(Interval Units)

Y_o = 710.4

σ distance from the Mean	Value of y when Y_o=1 (Read from tables)	Value of y when Y_o=710.4 obtained from the data	Height of the ordinate
± 1 σ	.60653	.60653 x 710.4	430.9
± 2 σ	.13534	.13534 x 710.4	96.12
± 3 σ	.01111	.01111 x 710.4	7.89

Table :23: Showing Normal Curve Ordinates at Mean
Sub-test IV

N = 2000 ; Mean = 10 ; σ= 1.1173 (Interval units)
Y_o = 680.1

σ-distance from the Mean	Value of Y when Y_o=1 (Read from tables)	Value of Y when Y_o=680.1 obtained from the ordinate	Height of the ordinate
$\pm$ 1σ	.60653	.60653 x 680.1	412.6
$\pm$ 2σ	.13534	.13534 x 680.1	92.02
$\pm$ 3σ	.01111	.01111 x 680.1	7.556

Table :24: Showing Normal Curve Ordinates at Mean
Sub-test V

N = 2000 ; Mean = 10.22 ; σ = 1.203 (Interval Units)
Y_o = 663.1

σ-distance from the Mean	Value of Y when Y_o=1 (Read from tables)	Value of Y when Y_o=663.1 obtained from the ordinates	Height of the ordinate
$\pm$ 1 σ	.60653	.60653 x 663.1	402.3
$\pm$ 2 σ	.13534	.13534 x 663.1	89.73
$\pm$ 3 σ	.01111	.01111 x 663.1	7.367

Table :25: Showing Normal Curve Ordinates at Mean
Sub-test VI

N = 2000 ; Mean = 9.4 ; σ = 1.03 (Interval Units)
Y_o = 774.7

σ-distance from the Mean	Value of Y when Y_o=1 (Read from tables)	Value of Y when Y_o=774.7 obtained from the data	Height of the ordinate
$\pm 1\sigma$	.60653	.60653 x 774.7	470.1
$\pm 2\sigma$	.13534	.13534 x 774.7	104.8
$\pm 3\sigma$	.01111	.01111 x 774.7	8.610

Table :26: Showing Normal Curve Ordinates at Mean
Sub-test VII

N = 2000 ; Mean = 5.202 ; σ = .803 (Interval Units)

Y_o = 993.6

σ-distance from the Mean	Value of Y when Y_o=1 (Read from tables)	Value of Y when Y_o=774.7 obtained from the data	Height of the ordinate
$\pm 1\sigma$	.60653	.60653 x 993.6	602.7
$\pm 2\sigma$	.13534	.13534 x 993.6	134.5
$\pm 3\sigma$	.01111	.01111 x 993.6	11.04

CHI-SQUARE TEST OF THE HYPOTHESIS OF NORMAL DISTRIBUTION

The chi-square test is found to be quite useful in testing some hypothesis. It is the sum ratios. Each ratio is between a squared (discrepancy or) difference and an expected frequency. The discrepancy is between an obtained frequency and a frequency expected on the basis of the hypothesis we are testing.

The hypothesis to be tested here is :

1. The distribution of the scores on the aptitude test follows the normal curve;
2. If there is any discrepancy between the observed and the expected frequencies it is insignificant and is due to chance factor/factors only.

The procedure discussed in Biometrika tables for statisticians, is followed thoroughly for calculating chi-square values.

The value of 'df' indicated in the Tables (27-34) is the number of class intervals minus 3. One degree of freedom has been lost in computing the mean, a second in computing the standard deviation and a third for N, the size of the sample, along with the statistics for total test scores, the statistics for scores on each sub-test also have been calculated with a view to studying the nature and role of each sub-test in the whole aptitude test battery. The sub-test scores should also be tested to find out whether they are also distributed normally.

[1]Abridged from Karl Pearson, Tables for Statisticians and Biometricians Part I, London: Cambridge University Press, 1924, pp. 2-6

Table :27: Chi-square test of the Normal Distribution Hypothesis applied to the Frequency Distribution : Whole Test

Scores	Scores	Frequency f_o	x - M	$\frac{x-M}{\sigma}$	Area P(x)	$\Delta P(x)$	$f_e = N \times \Delta P(x)$	$f_o - f_e$	$(f_o - f_e)^2$	$\frac{(f_o-f_e)^2}{f_e}$
111-120	110.5-120.5	8	43.79	2.42	0.99224	.00776	16	8	64	4.0000
101-110	100.5-110.5	48	33.79	1.91	0.97193	.02031	41	7	49	1.1860
91-100	90.5-100.5	132	23.79	1.34	0.90988	.06205	124	8	64	0.5162
81-90	80.5-90.5	252	13.79	0.78	0.78230	.12758	255	3	9	0.0353
71-80	70.5-80.5	384	3.79	0.21	0.58317	.19913	398	14	196	0.4970
61-70	60.5-70.5	426	-6.21	-0.35	0.36317	.22000	440	14	196	0.4454
51-60	50.5-60.5	390	-16.21	-0.92	0.17879	.18438	369	21	441	1.1950
41-50	40.5-50.5	230	-26.21	-1.48	0.06944	.10935	219	11	121	0.5526
31-40	30.5-40.5	99	-36.21	-2.05	0.02018	.04926	99	0	0	0
21-30	20.5-30.5	27	-46.21	-2.61	0.00453	.01565	31	4	16	0.5160
11-20	10.5-20.5	4	-56.21	-3.18	0.00074	.00379	8	4	16	2.0000
	N=	2000								χ^2=10.9435

Mean = 66.71

σ = 17.69

Degrees of Freedom (df) = 8

From the χ^2 Table

At 0.01 level χ^2 = 20.09

0.05 level χ^2 = 15.507

χ^2 value obtained is not significant at .01 and .05 levels

Table :28: Chi-square test of the Normal Distribution Hypothesis applied to the Frequency Distribution : Sub-test I

Scores	Scores	f_o	x-M	$\frac{x-M}{\sigma}$	Area P(x)	$\Delta P(x)$	$f_e = N \times \Delta P(x)$	(f_o-f_e)	$(f_o-f_e)^2$	$\frac{(f_o-f_e)^2}{f_o}$
24-26	23.5-26.5	56	9.09	2.05	0.97982	0.02018	40	16	256	6.398
21-23	20.5-23.5	118	6.09	1.37	0.91466	0.06516	130	12	144	1.108
18-20	17.5-20.5	359	3.09	0.70	0.75804	0.15662	313	46	2116	6.761
15-17	14.5-17.5	452	0.09	0.02	0.50798	0.25006	500	48	2304	4.068
12-14	11.5-14.5	474	-2.91	-0.66	0.25463	0.25335	507	33	1089	2.15
9-11	8.5-11.5	329	-5.91	-1.33	0.09176	0.06287	326	3	9	0.0276
6-8	5.5-8.5	165	-8.91	-2.01	0.02222	0.06954	139	26	676	4.863
3-5	2.5-5.5	40	-11.91	-2.645	0.00415	0.01807	36	4	16	0.4444
0-2	0 -2.5	7	-14.91	-3.12	0.00090	0.00325	6	1	1	0.1660
	N =	2000							χ^2 =	25.99

Mean = 14.41

σ = 4.43

df = 6

χ^2 obtained is slightly significant at both the levels.

From the χ^2 table

At 0.01 level, $\chi^2 = 16.812$

0.05 level $\chi^2 = 12.592$

Table :29: Chi-square test of the Normal Distribution Hypothesis applied to a Frequency Distribution : Sub-test II

Scores	Scores	f_o	x-M	$\frac{x-M}{\sigma}$	Area P(x)	$\Delta P(x)$	$f_e = N \times \Delta P(x)$	(f_o-f_e)	$(f_o-f_e)^2$	$\frac{(f_o-f_e)^2}{f_e}$
24-26	23.5-26.5	33	10.44	1.972	0.97558	0.02442	49	16	256	5.224
21-23	20.5-23.5	126	7.44	1.420	0.92220	0.05338	107	19	361	3.374
18-20	17.5-20.5	265	4.44	0.8474	0.80234	0.11986	240	25	625	2.604
15-17	14.5-17.5	374	1.44	0.2749	0.61026	0.19208	384	10	100	0.260
12-14	11.5-14.5	424	-1.56	-0.2977	0.38209	0.22817	456	32	1024	2.245
9-11	8.5-11.5	368	-4.56	-0.8704	0.19215	0.18994	380	12	144	0.379
6-8	5.5-8.5	257	-7.56	-1.443	0.07493	0.11722	234	23	529	2.261
3-5	2.5-5.5	120	-10.56	-2.015	0.02169	0.05324	107	13	169	1.580
0-2	0-2.5	33	-13.56	-2.587	0.00480	0.01689	34	1	1	0.029
		N = 2000								χ^2 = 17.956

Mean = 13.06

σ = 5.24

df = 6

χ^2 obtained is slightly significant at both the levels.

From the χ^2 table

At 0.01 level χ^2 = 16.812

0.05 level χ^2 = 12.592

Table :30: Chi-square test of the Normal Distribution Hypothesis applied to a frequency distribution : Sub-test III

Scores	Scores	f_o	x - M	$\frac{x-M}{\sigma}$	Area P(x)	$\Delta P(x)$	$f_e = N \times \Delta P(x)$	$f_e - f_e$	$(f_o - f_e)^2$	$\frac{(f_o - f_e)^2}{f_e}$
18-20	17.5-20.5	1	9.39	2.79	0.00736	0.00264	5	4	16	3.200
15-17	14.5-17.5	39	6.39	1.90	0.97128	0.2608	52	13	169	3.250
12-14	11.5-14.5	289	3.39	1.01	0.84375	0.12753	255	34	1156	4.517
9-11	8.5-11.5	584	0.39	0.12	0.54776	0.29599	592	8	64	0.108
6-8	5.5-8.5	623	-2.61	0.78	0.21770	0.33006	660	37	1369	2.074
3-5	2.5-5.5	387	-5.61	-1.67	0.04746	0.17024	341	46	2116	6.204
0-2	0-2.5	77	-8.61	-2.56	0.00523	0.04233	85	8	64	0.753
		N = 2000								χ^2 = 20.106

Mean = 8.113

σ = 3.37

df = 4

χ^2 value obtained is slightly significant at both the levels.

From the χ^2 table

At 0.01 level χ^2 = 13.277

0.05 level χ^2 = 9.488

Table :31: Chi-square test of the Normal Distribution Hypothesis applied to a Frequency Distribution : Sub-test IV

Scores	Scores	f_o	X - M	$\frac{X - M}{\sigma}$	Area P(x)	$\Delta P(x)$	$f_e = N \times \Delta P(x)$	$f_o - f_e$	$(f_o - f_e)^2$	$\frac{(f_o - f_e)^2}{f_e}$
18-20	17.5-20.5	15	7.5	2.131	0.98341	0.01659	33	18	324	9.900
15-17	14.5-17.5	180	4.5	1.28	0.89973	0.08368	167	13	169	1.012
12-14	11.5-14.5	470	1.5	0.426	0.66640	0.23333	467	4	16	0.034
9-11	8.5-11.5	700	-1.5	-0.426	0.33360	0.33280	666	1156	1156	1.740
6-8	5.5-8.5	430	-4.5	-1.28	0.10027	0.23333	467	37	1369	2.910
3-5	2.5-5.5	170	-7.5	-2.131	0.01659	0.08368	167	3	9	0.054
0-2	0-2.5	35	-10.5	-2.984	0.00144	0.01515	30	5	25	0.830
	N =	2000							χ^2 =	16.48

Mean = 10   df = 4

σ = 3.52

χ^2 value obtained is slightly significant at both the levels

From the χ^2 table

At 0.01 level χ^2 = 13.277

0.05 level χ^2 = 9.485

Table :32: Chi-square test of the Normal Distribution Hypothesis applied to the Frequency Distribution : Sub-test V

Scores	Scores	f_o	x-M	$\frac{x-M}{\sigma}$	Area P(x)	$\Delta P(x)$	$f_e = N \times \Delta P(x)$	$f_o - f_e$	$(f_o - f_e)^2$	$\frac{(f_o - f_e)^2}{f_e}$
18-20	17.5-20.5	29	7.28	2.031	0.97882	0.02118	42	13	169	4.024
15-17	14.4-17.5	189	4.28	1.185	0.88298	0.09584	192	3	9	0.048
12-14	11.5-14.5	523	1.28	0.3546	0.64058	0.24240	485	38	1444	2.98
9-11	8.5-11.5	644	-1.72	-0.4764	0.31561	0.32497	650	6	36	0.055
6-8	5.5-8.5	430	-4.72	-1.307	0.09510	0.22051	441	11	121	0.2745
3-5	2.5-5.5	145	-7.72	-2.138	0.01618	0.7892	158	13	169	1.069
0-2	0-2.5	40	-10.72	-2.970	0.00149	0.01469	29	11	121	4.17
	N =	2000							χ^2 =	12.621

Mean = 10.22

σ = 3.61

df = 4

χ^2 value obtained is slightly significant at .05 level and not significant at 0.01 level

From the χ^2 table

At 0.01 level χ^2 = 13.279

0.05 level χ^2 = 9.488

Table :33: Chi-square test of the Normal distribution Hypothesis applied to a Frequency Distribution : Sub-test VI

Scores	Scores	f_o	x-M	$\frac{x-M}{\sigma}$	Area P(x)	$\triangle P(x)$	$f_e = N \times \triangle P(x)$	$f_e - f_e$	$(f_o-f_e)^2$	$\frac{(f_o-f_e)^2}{f_e}$
18-20	17.5-20.5	10	8.1	2.613	0.99547	0.00453	9	1	1	0.111
15-17	14.5-17.5	88	5.1	1.645	0.95053	0.04494	90	2	4	0.044
12-14	11.5-14.5	432	2.1	0.6774	0.75175	0.20878	418	14	196	0.469
9-11	8.5-11.5	720	-0.9	-0.2902	0.38591	0.36584	732	12	144	0.197
6-8	5.5-8.5	530	-3.9	-1.258	0.10383	0.28208	564	34	1156	2.049
3-5	2.5-5.5	180	-6.9	-2.225	0.01287	0.09096	182	2	4	0.221
0-2	0-2.5	40	-9.9	-3.193	0.00071	0.01216	24	16	256	10.670
	N =	2000							χ^2 =	13.761

Mean = 9.4

σ = 3.1

df = 4

The χ^2 value obtained is slightly significant at .05 and at 0.01 levels.

From the χ^2 table

At 0.01 level χ^2 = 13.279

At 0.05 level χ^2 = 9.488

Table :34: Chi-square test of the Normal Distribution Hypothesis applied to a Frequency Distribution : Sub-test VII

Scores	Scores	f_o	x-M	$\frac{x-M}{\sigma}$	Area P(x)	$\triangle P(x)$	$f_e = N \times \triangle P(x)$	$f_o - f_e$	$(f_o - f_e)^2$	$\frac{(f_o - f_e)^2}{f_e}$
12-14	11.5-14.5	3	6.3	2.614	0.99547	0.00453	9	6	36	4.00
9-11	8.5-11.5	134	3.3	1.370	0.91466	0.08081	162	28	784	4.84
6-8	5.5-8.5	783	0.3	0.1245	0.55172	0.36294	726	57	3249	4.475
3-5	2.5-5.5	821	-2.7	-1.120	0.13136	0.42036	841	20	400	0.476
0-2	0-2.5	259	-5.7	-2.365	0.00889	0.12247	250	9	81	0.32
	N =	2000							χ^2 =	14.11

Mean = 5.2015

σ = 2.41

df = 2

The χ^2 value obtained is found to be slightly significant at both the levels.

From the χ^2 table

At 0.01 level χ^2 = 9.210

0.05 level χ^2 = 5.991

A perusal of the Chi-square values shows that the value is not at all significant at .05 and .01 levels for the whole test while the values are slightly significant at both the levels in the sub-tests except the fifth one. In the case of sub-test V the chi-square value is not significant at .01 level but slightly significant at .05 level.

It can be concluded that the distribution of scores in the aptitude test followed the normal curve since the Chi-square value obtained for the whole test is not significant at .05 and .01 levels.

STUDY OF THE PERFORMANCES OF BOYS AND GIRLS

To have a comparative study of the performance of boys and girls in the science aptitude test, a representative sample of 400 from each of the groups of boys andg girls is selected. The Mean, Median and Standard Deviation of the two frequency distributions are calculated. The data pertaining to the two distributions of boys and girls are given in Tables 35 and 36.

Table :35: Data grouped for the Calculation of Mean, Median and S.D. of the Distribution - Sample of 400 Boys

Scores	Mid-point	f	Cum.f.	d	fd	fd^2	Standard Scores, M'=50 & σ' = 10
110-119	114.5	1	400	+ 4	+ 4	16	74.07
100-109	104.5	18	399	+ 3	+54	162	68.19
90-99	94.5	58	381	+ 2	+ 116	232	62.29
80-89	84.5	70	323	+ 1	+70	70	56.40
70-79	74.5	102	253	0	0	0	50.51
60-69	64.5	74	151	- 1	-74	74	44.61
50-59	54.5	45	77	- 2	-90	180	38.71
40-49	44.5	16	32	- 3	-48	144	32.81
30-39	34.5	13	16	- 4	-52	208	26.91
20-29	24.5	3	3	- 5	-15	75	21.01
		N = 400				$\Sigma fd = -35$	$\Sigma fd^2 = 1161$

Mean = 73.63 Mdn. = 74.3 S.D. = 17.03

Note: Mid points of class intervals of the frequency distribution are expressed in standard scores with M' = 50 and σ' = 10.

Table :36: Data grouped for the Calculation of Mean, Median and S.D. of the Distribution - Sample 400 girls.

Scores	Mid point	f	Cum.f	d	fd	fd^2	Standard Scores M'=50 & σ'=10
110-119	114.5	6	400	+ 4	24	96	74.30
100-109	14.5	26	394	+ 3	78	234	68.30
90-99	94.5	41	368	+ 2	82	164	62.30
80-89	84.5	71	327	+ 1	71	71	56.30
70-79	74.5	103	256	0	0	0	50.30
60-69	64.5	80	153	- 1	- 80	80	43.70
50-59	54.5	39	73	- 2	- 78	156	37.70
40-49	44.5	22	34	- 3	- 66	198	31.70
30-39	34.5	8	12	- 4	- 32	128	25.70
20-29	24.5	4	4	- 5	- 20	100	19.70
	N =400			Σ fd =	- 21	Σfd^2=12.27	

Mean = 73.98 Mdn. = 74.06 S.D. = 17.36

Note: Mid points of class intervals of the frequency distribution are expressed in standard scores with M' = 50 and σ' = 10.

The results(mean, median and standard deviation values) indicate that there is no significant difference between boys and girls in their performance in the present 'Aptitude test'.

RELIABILITY OF THE TEST

The determining of the reliability of the test is the most essential characterisation of a good measuring instrument. The most common method used for determining the reliability of a test is the 'Split half method'. On the application of this method the items are divided into equivalent parts or tests by placing the correct odd items in one part and the correct even items in the other. If the items of the test have been well scaled in difficulty two equivalent parts can be tested. These two parts are now treated as two forms of the same test and the coefficient of correlation computed between them. We thus have a reliability coefficient based on a test half as long as the original. From the half test reliability, the self correlation of the whole test is estimated by the 'Spearman-Brown Prophecy Formula'[1]

$$R = \frac{nr}{1+(n-1)\ r}$$ where

r stands for obtained correlation

n for the number of parts of a test (in split half method, it will be two)

R for the Reliability of the whole test

---

[1] P.E.Vernon. Measurement of Abilities(London: University of London Press Ltd.,1953),p.145.

The split half method is employed when it is not possible to construct an alternate form of the test.

Objection has been raised about the split half method on the ground that a test can be divided into two parts in a variety of ways so that the reliability coefficient is not a unique value. This criticism is strictly true only when the items are of equal difficulty. When items are placed in order of difficulty from least to most difficult as in this test, the split into odd and even gives a determination of the reliability coefficient, which is quite dependable.

Again its main advantage is that all the data for determining the test reliability are obtained on one occassion, hence variation introduced by differences between the two testing situations are eliminated. Hence the split half method is regarded as the best of the methods for determining the test reliability.

This method is followed for determining the reliability of the present test. A small sample of 400 testees out of the total sample of 2000 is selected for the purpose of applying 'Split half'method to estimate the reliability of the whole test. The testees selected is based on the odd and even method. Scores secured by the pupils for the odd and even items are found out and tabulated. On the basis of the data a scatter diagram is prepared and the coefficient of correlation is computed. Then the reliability of the whole test is determined with the help of the 'Spearman-Brown Prophecy' formula. The scattergram of scores used in Split half method is given in Table 37.

Table :37: Scattergram of Scores used in Split Half Method

'Even-items' Scores

	Scores	15-19	20-24	25-29	30-34	35-39	40-44	45-49	f_y
ODD ITEMS SCORES	50-54							21	21
	45-49						13	16	29
	40-44				15	12	18	3	48
	35-39			16	18	46	5		85
	30-34		2	15	45	25	5		92
	25-29	3	6	33	5	5	20		72
	20-24	14	27						41
	15-19	12							12
	f_x	29	35	64	83	88	61	40	400

Product moment $r = 0.783$ P.E.r $= \pm 0.0761$

The reliability coefficient based on a test half as long as the entire test is 0.783. The reliability of the entire test is calculated by the Spearman-Brown Prophecy formula.

$$R = \frac{nr}{1 + (n-1)r}$$

$$= \frac{2 \times 0.783}{1 + 1 \times 0.783}$$

$$= \frac{1.56}{1.78}$$

$$= 0.876$$

Since the reliability coefficient of the aptitude test is considerably high, the test may be considered to be highly reliable.

The P.E. of the 'r' (0.876) is given by :

P.E.'r' = $0.6745 \times \frac{1 - r^2}{\sqrt{N}}$ Where the reliability coefficient is denoted by 'r'.

$$= 0.6745 \times \frac{1 - (0.876)^2}{\sqrt{400}}$$

$$= 0.0761$$

The significance of the obtained reliability coefficient is also determined. A good method of testing the significance of the coefficient, when the value is high is to convert it into R.A.Fisher's[1]- Z function and find the standard error of Z function. The formula for the standard error of Z, σ_z is

$$\sigma_z = \frac{1}{\sqrt{N - 3}}$$

Where N = 400

[1] Fisher, R.A. Statistical Methods for Research Workers. (London: Oliver & Boyd, 1941), pp.190-203 cited by H.E.Garrett 'Statistics in Psychology & Education', pp.199-and 448.

From the Tables[1], we read that an r of 0.88 corresponds to a Z of 1.38.

Standard error of Z,

$$SE_Z = \sigma_Z$$

$$= \frac{1}{\sqrt{400-3}}$$

$$= \frac{1}{\sqrt{397}}$$

$$= \frac{1}{19.925}$$

$$= 0.0500$$

The value of Z of 1.38 for corresponding r of 0.88 ranges between 1.48 and 1.28 (i.e. 1.38 $\pm$ 1.96 x 0.05) converting these values of Z's back into r's we get a confidence interval from 0.857 to 0.902, since the range within which the true r lies, is narrow we arrive at the conclusion that r obtained in the test for reliability is considerably significant.

The conversion of r into Fisher's Z function and the determination of SE of Z is necessitated by its two main advantages over r viz. (1) its sampling distribution is approximately normal and (2) its SE depends only upon the

---

[1] Ibid.

size of the sample N and is independent of the size of r.

Application of the Kuder-Richardson Method

Dissatisfied with the Split half Method, Kuder and Richardson developed a new proceedure based on item statistics to estimate the reliability of the test. They split a test into 'n' parts of one item each.

The formula provides an estimate of the internal consistency of the test and thus of the dependability of test scores.

The K-R formula used here is given below :

$$r_{11} = \frac{n}{n-1} \times \frac{\sigma_t^2 - \Sigma pq}{\sigma_t^2}$$

Where

r_{11} = Reliability coefficient of the whole test

n = Number of items in the test

σ_t = The standard deviation of the test scores

p = The proportion of the group answering a test item correctly

q = (1 - p) = the proportion of the group answering a test item incorrectly.

To apply this method, a sample of 400 testees is used. The standard deviation of the test is equal to 17.69. The proportion of the group answering a test item correctly is found out for each of the 148 test items. From the values of 'p', the corresponding values of 'q' are calculated. In the table given on the next page, the values of 'pq' for each of the items are shown. The sum of all 'pq' values is found to be 32.23.

The reliability coefficient is calculated using the K-R formula

$$r_{11} = \frac{n}{n-1} \times \frac{\sigma_t^2 - \Sigma pq}{\sigma_t^2}$$

$$= \frac{148}{147} \times \frac{(17.69)^2 - 32.23}{(17.69)^2}$$

$$= \frac{148 \times 280.67}{147 \times 312.9}$$

$$= 0.9014$$

The reliability coefficient of the present aptitude test as measured by K - R formula method is 0.9014 which is slightly higher than the result obtained by the 'Split half method.'

Sr.No.	Method used	Reliability coefficient obtained	P.E.r
1	Split half method	0.876	± 0.0761
2	Kuder-Richardson method	0.9014	-

The reliability coefficient obtained is fixed at 0.89 and the value showed that the test is highly reliable.

Table :38: Showing 'pq' Values of 148 Test Items

Item No.	'pq'	Item No.	'pq'
1	0.1476	23	0.1924
2	0.1476	24	0.2176
3	0.1344	25	0.2016
4	0.2464	26	0.2500
5	0.2500	27	0.1600
6	0.2016	28	0.2484
7	0.2400	29	0.2304
8	0.2100	30	0.2400
9	0.2496	31	0.2464
10	0.2304	32	0.2464
11	0.2304	33	0.2100
12	0.2016	34	0.2244
13	0.1476	35	0.2496
14	0.1924	36	0.2400
15	0.1476	37	0.2304
16	0.2356	38	0.2176
17	0.2176	39	0.1824
18	0.2464	40	0.2304
19	0.2464	41	0.2304
20	0.2484	42	0.1824
21	0.2484	43	0.2464
22	0.2100	44	0.2464

Table :38: (Contd.)

Item No.	'pq'	Item No	'pq'
45	0.1344	68	0.2464
46	0.1924	69	0.2400
47	0.1824	70	0.1476
48	0.2100	71	0.2244
49	0.2464	72	0.2176
50	0.2304	73	0.2244
51	0.1476	74	0.2496
52	0.1204	75	0.2400
53	0.2304	76	0.2436
54	0.2176	77	0.2304
55	0.2016	78	0.2356
56	0.2464	79	0.2244
57	0.2464	80	0.2400
58	0.2244	81	0.2304
59	0.2304	82	0.1716
60	0.2436	83	0.2400
61	0.2176	84	0.2284
62	0.2356	85	0.2400
63	0.2100	86	0.2244
64	0.2100	87	0.2484
65	0.2400	88	0.2176
66	0.2464	89	0.2244
67	0.2484	90	0.2484

Table :38: (Contd.)

Item No.	'pq'	Item No.	'pq'
91	0.2304	114	0.2400
92	0.2176	115	0.2496
93	0.2400	116	0.2356
94	0.2244	117	0.2496
95	0.1476	118	0.2464
96	0.2100	119	0.2400
97	0.2244	120	0.2244
98	0.2244	121	0.0564
99	0.2484	122	0.2176
100	0.2356	123	0.1924
101	0.2176	124	0.2100
102	0.2100	125	0.2244
103	0.2400	126	0.2244
104	0.2016	127	0.2400
105	0.2016	128	0.2100
106	0.2100	129	0.2304
107	0.2356	130	0.2304
108	0.2100	131	0.2464
109	0.2244	132	0.2304
110	0.2400	133	0.2176
111	0.2100	134	0.2244
112	0.2304	135	0.2496
113	0.2244	136	0.2400

Table :38: (Contd.)

Item No.	'pq'	Item No.	'pq'
137	0.2356	143	0.2100
138	0.2244	144	0.2244
139	0.2016	145	0.1824
140	0.2100	146	0.2176
141	0.2100	147	0.2356
142	0.2400	148	0.1600

THE ESTIMATION OF TEST VALIDITY

The reliability of the present test is estimated by applying two different methods. It is found to be 0.89 and the value is seen to be quite satisfactory as faf as the test is concerned.

But the test constructor is not to be satisfied merely with the reliability of the test. He has to know more about the test viz. whether the test measures what it purports to measure. Unless he is sure about this, he cannot recommend its use for any definite purpose.

Validity of a Test

The validity of a test depends on the efficiency with which it measures what it attempts to measure.

Ross[1] while defining validity says :

" One kind of validity concerns the degree to which the test or other measuring instrument measures what it claims to. In a word, validity means truthfulness."

According to Gulliksen[2], " the validity of a test is the correlation of the test with some criterion."

Validity thus refers to the truthfulness of the test and is always its most important characteristic. No matter what other merits the test may possess, if it lacks validity, it is not worth its use.

The validity of a test is determined experimentally by finding the correlation between the test and some independent criterion.

1. The criterion against which the present test is validated is the annual examination marks of the pupils in Science, of the preceeding year. A sample of 400 testees is selected for determining the validity of the test.
2. Secondly, the opinion of the science teacher in the class is also taken for finding the validity of the test. The teacher's estimation of the pupils is taken

---

[1] Ross,C.C. Measurement in Today's Schools. (New York:Prentice-Hall,Inc.,1955),p.107.

[2] Gulliksen,Harold. Theory of Mental Tests. (New York: John Wiley & Sons,Inc.,1950),p.88.

on a seven point scale and the correlation coeffcient is calculated between teacher's estimation of the pupils and the test scores.

The test scores and the criterion scores indicated in Tables 39 and 40 are expressed in standard scores. The raw test scores and the raw criterion scores are converted into the standard scores with the help of the formula given on Page 79 The raw scores here are expressed in standard scores in a distribution where $M' = 50$ and $\sigma' = 10$.

The annual examination marks in science of the preceeding year are taken as criterion scores and correlated with the standard test scores and the value is found to be 0.76. The scatter diagram pertaining to the standard test scores and criterion scores is given in Table 41. Similarly the correlation coefficient between the test score and the teacher's estimation on a seven point scale is also calculated by the Product moment method. The value is found to be 0.72, and the related scatter diagram is given in Table 42. The values obtained in both the cases are found to be fairly high and it testifies the validity of the test.

Table :39: Raw Scores of the Final Test and their Corresponding Standard Scores

$M = 70.15$, $\sigma = 17.65$ $M' = 50$, $\sigma' = 10$

Raw Test Scores	Standard Scores	Raw Test Scores	Standard Scores
1	8.6	21	20.6
2	9.2	22	21.2
3	9.8	23	21.8
4	10.4	24	22.4
5	11.0	25	23.0
6	11.6	26	23.6
7	12.2	27	24.2
8	12.8	28	24.8
9	13.4	29	25.4
10	14.0	30	26.0
11	14.6	31	26.6
12	15.2	32	27.2
13	15.8	33	27.8
14	16.4	34	28.4
15	17.0	35	29.0
16	17.6	36	29.6
17	18.2	37	30.2
18	18.8	38	30.8
19	19.4	39	31.4
20	20.0	40	32.0

Table :39: Contd.

Raw Test Scores	Standard Scores	Raw Test Scores	Standard Scores
41	32.6	63	45.8
42	33.2	64	46.4
43	33.8	65	47.0
44	34.4	66	47.6
45	35.0	67	48.2
46	35.6	68	48.8
47	36.2	69	49.4
48	36.8	70	50.0
49	37.4	71	50.6
50	38.0	72	51.2
51	38.6	73	51.8
52	39.2	74	52.4
53	39.8	75	53.0
54	40.4	76	53.6
55	41.0	77	54.2
56	41.6	78	54.8
57	42.2	79	55.4
58	42.8	80	56.0
59	43.4	81	56.6
60	44.0	82	57.2
61	44.6	83	57.8
62	45.2	84	58.4

Table :39: Contd.

Raw Test Scores	Standard Scores	Raw Test Scores	Standard Scores
85	59.0	107	72.2
86	59.6	108	72.8
87	60.2	109	73.4
88	60.8	110	74.0
89	61.4	111	74.6
90	62.0	112	75.2
91	62.6	113	75.8
92	63.2	114	76.4
93	68.8	115	77.0
94	64.4	116	77.6
95	65.0	117	78.2
96	65.6	118	78.8
97	66.2	119	79.4
98	66.8	120	80.0
99	67.4	121	80.6
100	68.0	122	81.2
101	68.6	123	81.8
102	69.2	124	82.4
103	69.8	125	83.0
104	70.4	126	83.6
105	71.0	127	84.2
106	71.6	128	84.8

Table :39: Contd.

Raw Test Scores	Standard Scores	Raw Test Scores	Standard Scores
129	85.4	139	91.4
130	86.0	140	92.0
131	86.6	141	92.6
132	87.2	142	93.2
133	87.8	143	93.8
134	88.4	144	94.4
135	89.0	145	95.0
136	89.6	146	95.6
137	90.2	147	96.2
138	90.8	148	96.8

(It may please be noted that ' 148 ' is the maximum attainable score in the present test).

Table :40: Raw Criterion Scores and their corresponding Standard Scores

M = 51.75 σ = 12.8   M'=50 ; σ'=10

Raw Criterion Scores	Standard Scores	Raw Criterion Scores	Standard Scores
20	24.4	25	28.4
21	25.2	26	29.2
22	26.0	27	30.0
23	26.8	28	30.8
24	27.6	29	31.6

Table :40: Contd.

Raw Criterion Scores	Standard Scores	Raw Criterion Scores	Standard Scores
30	32.4	56	53.2
31	33.2	57	54.0
32	34.0	58	54.8
33	34.8	59	55.6
34	35.6	60	56.4
35	36.4	61	57.2
36	37.2	62	58.0
37	38.0	63	58.8
38	38.8	64	59.6
39	39.6	65	60.4
40	40.4	66	61.2
41	41.2	67	62.0
42	42.0	68	62.8
43	42.8	69	63.6
44	43.6	70	64.4
45	44.4	71	65.2
46	45.2	72	66.0
47	46.0	73	66.8
48	46.8	74	67.6
49	47.6	75	68.4
50	48.4	76	69.2
51	49.2	77	70.0
52	50.0	78	70.8
53	50.8	79	71.6
54	51.6	80	72.4
55	52.4	-	-

Table :41: Scatter Diagram Between Standard Test Scores and Standard Criterion Scores

		Standard Criterion Scores									
	Scores	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	Total
Standard Test Scores	70-74								1	2	6	9
	65-69						1	3	13	6	2	25
	60-64					5	5	10	14	2		36
	55-59				4	12	14	16	6	5		57
	50-54			1	15	25	20	25	5	5		96
	45-49			4	20	16	15					70
	40-44		1	4	21	28	6					60
	35-39		1	4	15							20
	30-34		1	11	6							18
	25-29		5	4								9
	Total		8	28	81	86	61	63	44	21	8	N=400

Product moment r = 0.76   P.E.r = ± 0.021

Table :42: Scatter Diagram Between Standard Test Scores and Teacher's Estimation

	Scores	$\bar{C}$	C	$\overset{+}{C}$	$\bar{B}$	B	$\overset{+}{B}$	A	Total
STANDARD TEST SCORES	70–74						1	8	9
	65–69					2	12	12	26
	60–64				1	18	20	2	41
	55–59				8	26	29	7	70
	50–54			2	35	30	36	2	105
	45–49			7	18	23	10	2	60
	40–44		2	6	27	13			48
	35–39		1	4	14				19
	30–34		3	10	3				16
	25–29		4	2					6
	Total		10	31	106	112	108	33	N =400

Product moment r = 0.72 P.E.r = ± 0.024

NORMS OF THE TEST

The most difficult phase of aptitude testing is interpretation of results. After the tests have been carefully administered and painstakingly scored, the findings must be appraised and translated into information helpful to the individual tested. A yardstick is therefore required to measure the magnitude of the deviation of a person's score from the general population average or from the average of his group. A norm is a standard of reference, so a table of norms serves as our yardstick.

Flanagan[1] defines test norms as " Estimates of some characteristic of a distribution of test scores for a specified population." Norms describe the actual performance of specified groups of individuals.

The terms 'norms' and 'standards' are frequently used interchangeably and the confusion arises over the fact that norms are used with standard tests and that a part of the process of standardisation is the derivation of norms. It is therefore necessary at the outset to distinguish clearly between a 'norm' and a 'standard'.

Flanagan[2] also emphasizes this distinction. He says, " Standards on the other hand are desirable or desired levels of attainment preferably expressed in terms of outcomes of instructions.

According to Greene, Jorgenson and Gerberich[3] " the term standard, when used to refer to a level of pupil's achievement implies an ultimate goal to be achieved, while norms are the levels of achievement which typical pupils actually attain."

Standards are formulated arbitrarily to suit one's requirement. Norms are derived from test results. The first ones are subjective while the second ones are objective. In the present case, the following norms are established for the test

[1]Lindquist,E.F., Educational Measurement,American Council on Education,Washington D.C.,1955,p.698

[2]Ibid.,p.698

[3]Greene,H.A.,Jorgensen,A.N., and Gerberich,J.R.,Measurement and Evaluation in the Secondary School",Longmans,Green & Co.,New York, 1955,p.102.

results :

i) Grade Norms
ii) Standard Score Norms
iii) Percentile Norms
iv) T-Score Norms.

Grade Norms

A grade norm may be defined as the mean or median achievement of pupils in a school grade on a given standardised test or it may be defined as the average status of pupils in a given grade with regard to a single factor.

The present test is administered for pupils of grade IX and as such the mean (66.71) and median (66.37) worked out for the distribution are the norms established for grade IX.

Standard Score Norms

A standard score is expressed as a deviation of a score from the arithmetic average of the normative group in which the standard deviation of the normative group is used as the unit of measurement.

Such scores simplify interpretation and increase comparability. The standard score is used most frequently by psychologists and research workers. The raw scores obtained on the test are converted into the standard scores with the help of the formula given on Page 79 in a distribution of $M' = 50$ and $\sigma' = 10$. The standard test scores obtained are given in Table 39.

Percentile Norms

A percentile norm may be defined as a point on a scale of measurement determined by the percentage of individuals in a given population that lies below this point. Percentile norms are widely used in achievement test of various subjects for high school children, in interest inventories, personality inventories and rating scales.

These norms are especially useful in dealing with educational achievement examination when we wish to evaluate and compare the achievement of given students in a number of subject matter tests.

The following formula[1] is used for calculating the percentiles. The method of calculating the percentiles is essentially as the one employed in finding the median.

$$P_p = l + \left(\frac{P_N - F}{f_p}\right) \times i$$

Where

P_p = Percentage of the distribution wanted e.g. 10%, 33% etc.

l = Exact lower limit of the class interval upon which P_p lies

P_N = Part of N to be counted off in order to reach P_p

F = Sum of all scores upon intervals below l

f_p = Number of scores within the interval upon which P_p falls.

i = Length of the Class interval

[1]Garrett, H.E. Statistics in Psychology and Education. (Bombay: Applied Pacific Private Ltd.,), p.65.

The percentiles calculated with the help of the formula given on the previous page, are shown in Table 43.

Table :43: Percentile Norms

Percentile	Score	Percentile	Score
P_1	26.43	P_{22}	52.55
P_2	31.41	P_{23}	53.06
P_3	33.43	P_{24}	53.58
P_4	35.45	P_{25}	54.90
P_5	37.47	P_{26}	54.60
P_6	39.49	P_{27}	55.12
P_7	40.93	P_{28}	55.63
P_8	41.80	P_{29}	56.14
P_9	42.67	P_{30}	56.65
P_{10}	43.54	P_{31}	57.17
P_{11}	44.42	P_{32}	57.68
P_{12}	45.28	P_{33}	58.19
P_{13}	46.15	P_{34}	58.70
P_{14}	47.12	P_{35}	59.22
P_{15}	47.89	P_{36}	59.73
P_{16}	48.76	P_{37}	60.24
P_{17}	49.63	P_{38}	60.74
P_{18}	50.50	P_{39}	61.20
P_{19}	50.87	P_{40}	61.67
P_{20}	51.53	P_{41}	62.14
P_{21}	52.04	P_{42}	62.61

Table :43: Contd.

Percentile	Score	Percentile	Score
P_{43}	63.08	P_{69}	75.81
P_{44}	63.55	P_{70}	76.33
P_{45}	64.02	P_{71}	76.85
P_{46}	64.49	P_{72}	77.38
P_{47}	64.96	P_{73}	77.90
P_{48}	65.43	P_{74}	78.41
P_{49}	65.90	P_{75}	78.94
P_{50}	66.37	P_{76}	79.46
P_{51}	66.84	P_{77}	79.98
P_{52}	67.31	P_{78}	80.50
P_{53}	67.78	P_{79}	81.29
P_{54}	68.25	P_{80}	82.08
P_{55}	68.72	P_{81}	82.87
P_{56}	69.14	P_{82}	83.66
P_{57}	69.66	P_{83}	84.44
P_{58}	70.13	P_{84}	85.23
P_{59}	70.60	P_{85}	86.03
P_{60}	71.13	P_{86}	86.81
P_{61}	71.65	P_{87}	87.60
P_{62}	72.17	P_{88}	88.39
P_{63}	72.69	P_{89}	89.18
P_{64}	73.21	P_{90}	89.96
P_{65}	73.73	P_{91}	91.11
P_{66}	74.25	P_{92}	92.62
P_{67}	74.77	P_{93}	94.14
P_{68}	75.29	P_{94}	95.65
-	-	P_{95}	97.17

Table :43: Contd.

Percentile	Score	Percentile	Score
P_{96}	98.68	P_{99}	108.00
P_{97}	100.20	P_{100}	120.50
P_{98}	103.83	-	-

Percentile Ranks

The percentile ranks corresponding to the raw scores obtained are also calculated. The procedure given in Garrett for computing percentile ranks is followed. The percentile rank corresponding to each raw score is given in Table 44.

The distinction between percentile and percentile rank is that in calculating percentiles one starts with a certain percent of N say 15% or 62%. Then one counts into the distribution the given percent and the point reached is the required percentile e.g. P_{15} or P_{62}. The procedure followed in computing percentile ranks is the reverse of the process. Here we begin with an individual score and determine the percentage of scores which lies below it. If this percentage is 62 say, the score has a percentile rank of PR on a scale of 100.

Table :44: Percentile Ranks

Raw Score	Percentile Rank	Raw Score	Percentile Rank
11	0.0100	33	2.7875
12	0.0300	34	3.2825
13	0.0500	35	3.7775
14	0.0700	36	4.2725
15	0.0900	37	4.7675
16	0.1100	38	5.2625
17	0.1300	39	5.7575
18	0.1500	40	6.2525
19	0.1700	41	7.0750
20	0.1900	42	8.2250
21	0.2675	43	9.3750
22	0.4025	44	10.5250
23	0.5375	45	11.6750
24	0.6725	46	12.8250
25	0.8075	47	13.9750
26	0.9425	48	15.1250
27	1.0775	49	16.2750
28	1.2125	50	17.4250
29	1.3475	51	18.9750
30	1.4825	52	20.9250
31	1.7975	53	22.8750
32	2.2925	54	24.8250

Table :44: Contd.

Raw Score	Percentile Rank	Raw Score	Percentile Rank
55	26.7750	77	71.28
56	28.7250	78	73.20
57	30.6750	79	75.12
58	32.6250	80	77.04
59	34.5750	81	78.63
60	36.5250	82	79.89
61	38.5650	83	81.15
62	40.6950	84	82.41
63	42.8250	85	83.67
64	44.9500	86	84.93
65	47.0350	87	86.19
66	49.2150	88	87.45
67	51.3450	89	88.71
68	53.4750	90	89.97
69	55.6050	91	90.93
70	57.7350	92	91.59
71	59.7600	93	92.25
72	61.6800	94	92.91
73	63.6000	95	93.57
74	65.5200	96	94.23
75	67.4400	97	94.89
76	69.3600	98	95.55

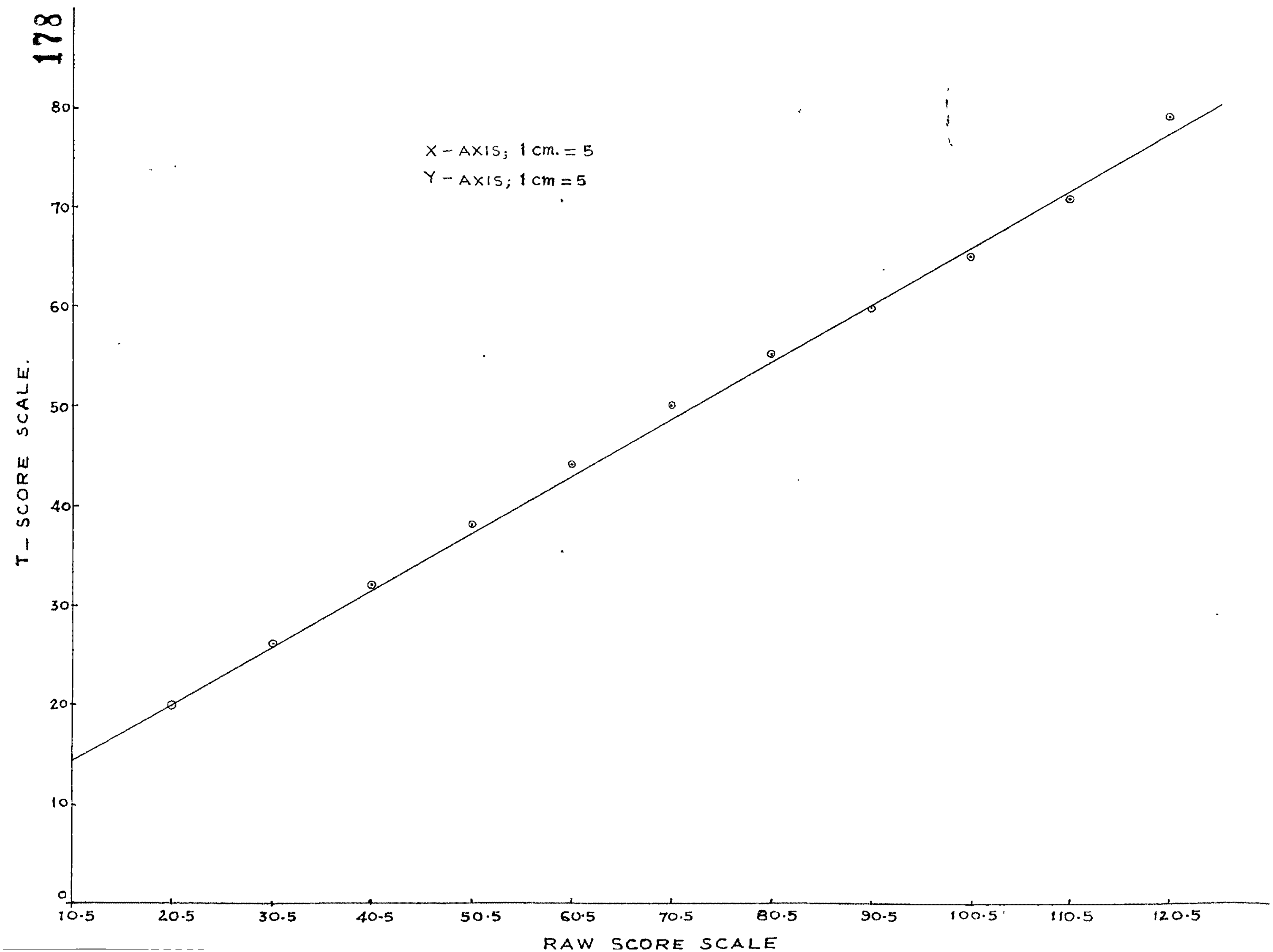
Table :44: Contd.

Raw Scores	Percentile Rank	Raw Score	Percentile Rank
99	96.21	110	99.48
100	96.87	111	99.62
101	97.32	112	99.66
102	97.56	113	99.70
103	97.80	114	99.74
104	98.04	115	99.78
105	98.28	116	99.82
106	98.52	117	99.86
107	98.76	118	99.90
108	99.00	119	99.94
109	99.24	120	99.98

The T-Score Norms

The well known T-scale overcomes the objections raised against standard scores and adds besides an advantage peculiar to itself. It adopts as its unit one tenth of a standard deviation, so that an ordinary distribution with a range of 5 to 6 σ on its base line yields 50 to 60 integral T-scale scores. In addition T-scale goes beyond any ordinary distribution, extending over a spread of 10 standard deviations or 100 units in all.

The obtained scores of the frequency distribution are converted into a system of 'normalised' σ scores by transforming


X - AXIS; 1 cm. = 5
Y - AXIS; 1 cm = 5
T_ SCORE SCALE.
80
70
60
50
40
30
20
10
0
10·5
20·5
30·5
40·5
50·5
60·5
70·5
80·5
90·5
100·5
110·5
120·5
RAW SCORE SCALE

them directly into equivalent points in a normal distribution. Normalised standard scores are generally called T-scores. T-scaling was devised by McCall. T-scores are normalised standard scores converted into a distribution with a mean of 50 and σ of 10. The procedure suggested by Garrett is followed in the calculation of the T-scores. The calculated T-scores are given in Table 45 and a graph is drawn showing the relation between the upper limits of the class intervals and T-scores. If the distribution of the scores is normal, the points should fall rather close to a straight line. From the graph on page 178 it is seen that the points fall on a straight line and it shows the distribution is normal.

For any integral raw score points the corresponding T-score points could be found out from the graph.

Table :45: Showing the T-score values for the distribution.

Scores	f	Cum.f	Cum.f below score + 1/2 on given score	Col.4 in %'s	T-Scores
1	2	3	4	5	6
111-120	8	2000	1996	99.80	79
101-110	48	1992	1968	98.40	71
91-100	132	1944	1878	93.90	65
81-90	252	1812	1686	84.30	60
71-80	384	1560	1368	68.40	55
61-70	426	1176	963	48.15	50
51-60	390	750	555	27.75	44
41-50	230	360	245	12.25	38
31-40	99	130	80.5	4.025	32
21-30	27	31	17.5	0.875	26
11-20	4	4	2	0.1	20
	N=2000				