CHAPTER VI

STATISTICAL ANALYSIS OF THE DATA

6.1 INTRODUCTION

Any psychological test needs a systematic way of construction and its pilot testing before administration of the final version of the test. The steps of IETAT construction and its pilot testing already have been discussed in previous chapters. Administration of the final version of IETAT has also been discussed in the preceding chapter. This chapter presents the statistical analysis of the test data which is required for the standardization. The following steps have been followed for standardization of IETAT.

- The statistical analysis of the data including the computation of mean, median, SD etc. of the whole sample, graphical presentation and statistical discussions of the nature of the distribution of the test scores.
- Estimating the test reliability.
- Estimating the test validity.
- Establishing the test norms.
- Factor analysis of the test data.

In this chapter, the first four steps have been treated thoroughly. The statistical analysis of the data is presented under the heading measures of central tendency and variability of the scores, reliability of mean, median and SD of the whole test, deciding the nature of distribution of test scores including calculation of Skewness and kurtosis, graphical presentation of test scores and chi-square test of the normal distribution. Estimation of test reliability and validity along with established test norms have been also presented in this chapter. The detailed statistical analysis of the data is presented below.

6.2 MEASURES OF CENTRAL TENDENCY AND VARIABILITY OF THE IETAT SCORES

The central tendency of a distribution is an estimate of the center of a distribution value. Mean, median, and mode are three major types of measures of central tendency. Thus, the mean, median, and mode are calculated from the distribution of the scores obtained through the constructed IETAT.

The value of the mean, median, and mode indicate the central value of the distribution, however, the central value alone is not sufficient to fully describe the distribution and therefore we require measuring the spread of actual scores. The extent of spread may vary from one distribution to another and the variability of the actual scores in the distribution can be measured by the range and Standard Deviation [SD]. The range describes the difference in the highest and lowest score obtained. The maximum score that a pre-service teacher can obtain on the constructed IETAT is 50 and the minimum score that can be obtained is 0. The highest score that is obtained on the IETAT is 35 while the lowest score is 7. The range between the obtained highest and lowest scores is, therefore (35 – 7) + 1 =29. This range within which the scores are distributed is divided into six class intervals each interval being of five units. SD is an accurate measure of dispersion. Thus, SD along with the mean, median, mode and are calculated by using the following formulas IV, V, VI and VII from the distribution of the scores as given in the following Table 6.1.

Mean = $\frac{\sum fd}{N}$ (Formula IV)

Median = $L + \frac{\frac{n}{2} - Cfi}{fi} \times i$ (Formula V)

Mode = 3 Median – 2 Mean (Formula VI)

SD (σ) = $i \times \sqrt{\frac{\sum fd^2}{N} - \left(\frac{\sum fd}{N}\right)^2}$ (Formula VII)

Table 6.1

Frequency Distribution, Mean, Median, Mode and SD of the Scores in IETAT

Score Intervals	Midpoint (X)	Frequencies (f)	Cum. Frequencies (Cf)	d	fd	fd^2
31-35	33	17	552	+3	51	153
26-30	28	80	535	+2	160	320
21-25	23	181	455	+1	181	181
16-20	18	182	274	0	0	0
11-15	13	76	92	-1	-76	76
06-10	8	16	16	-2	-32	64
		N=552			$\sum fd = 284$	$\sum fd^2 = 794$

Mean = 20.57

Median = 20.55

Mode = 20.51

SD = 5.45

From the table 6.1 it can be observed that the value of mean, median and mode are found to be 20.57, 20.55 and 20.51 respectively. This indicates that there is no much difference between the mean, median and mode. Hence the distribution is normal and the selected sample is representative of the population.

The frequency distributions of the scores obtained by the pre-service teachers in each of the sections are also tabulated separately in Table 6.2.

Table 6.2

Section wise Frequency Distribution of the Scores

Sr. No.	Scores	Mid-points	Section I		Section II		Section III		Section IV		Section V	
			F	CF	F	CF	F	CF	F	CF	F	CF
1	9-10	9.5	15	552	-	-	13	552	-	-	-	-
2	7-8	7.5	156	537	21	552	141	539	6	552	11	552
3	5-6	5.5	214	381	120	531	202	398	129	546	106	541
4	3-4	3.5	118	167	262	411	135	196	235	417	275	435
5	0-2	1	49	49	149	149	61	61	182	182	160	160
			N=552		**N=552**		**N=552**		**N=552**		**N=552**	
Mean			**5.39**		**3.55**		**5.18**		**3.35**		**3.38**	
Median			**5.15**		**2.25**		**4.77**		**2.31**		**1.5**	
SD			**1.96**		**1.60**		**2.01**		**1.55**		**1.49**	

From the table 6.2 it can be seen that there is difference in the mean performance of pre-service teachers in different sections of IETAT. The highest mean observed in section I (5.39) whereas the lowest mean observed in the section IV (3.35).

6.3 RELIABILITY OF MEAN, MEDIAN AND STANDARD DEVIATION OF WHOLE IETAT

The result obtained above of the parametric mean, median and SD of the frequency distribution of the whole test is from the random sampling. These may deviate from the population parameter. The result presented in the table 7.1 indicates the selected sample is representative of population. 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 errors. The calculation of Standard Error [SE] of mean, median, and SD is done by using the following formulas VIII, XIV and X, and the result is presented below in the following table 6.3.

SE_M or σ_M = $\frac{\sigma}{\sqrt{N}}$ (Formula VIII)

Where,

σ = the standard deviation of the total distribution

N = no. of cases in the sample

SE_{Mdn} = $\frac{1.253\sigma}{\sqrt{N}}$ (Formula IX)

Where,

σ = the standard deviation of the total distribution

N = no. of cases in the sample

SE_{SD} = $\frac{0.71\sigma}{\sqrt{N}}$ (Formula X)

Where,

σ = the standard deviation of the total distribution

N = no. of cases in the sample

Table 6.3

SE and Confidence Interval of Mean, Median and SD

N	**SE of Mean** SE_M **= 0.2320**		**SE of Median** SE_{Mdn} **= 0.2891**		**SE of SD** SE_{SD} **= 0.1638**	
	Confidence Interval		**Confidence Interval**		**Confidence Interval**	
	0.95	**0.99**	**0.95**	**0.99**	**0.95**	**0.99**
	Mean±SE_M×1.96	Mean±SE_M×2.58	Mdn±SE_{Mdn}×1.96	Mdn±SE_{Mdn}×2.58	SD±SE_{SD}×1.96	SD±SE_{SD}×2.58
552	21.02 to 20.11	21.16 to 19.97	21.11 to 19.98	21.29 to 19.80	5.74 to 5.09	5.84 to 4.99

From the table 6.3 it can be observed that the values of SE_M, SE_{Mdn} and SE_{SD} found to be 0.2307, 0.2891 and 0.1638 respectively.

The true mean lies between 20.57 ± 0.2320 × 1.96 at .05 level and 20.57 ± 0.2320 × 2.58 at .01 level. The confidence interval is 21.02 to 20.11 at .05 level whereas 21.16 and

19.97 at .01 level. This result indicates that out of 100, 95 times the population mean will lie between the ranges 21.02 to 20.11 and 99 times between the ranges 21.16 to 19.97. Thus the obtained mean is highly reliable as the true mean lies within the narrow range at both the levels.

The true median lies between 20.57 ± 0.2891 × 1.96 at .05 level and 20.57 ± 0.2891 × 2.58 at .01 level. The confidence interval is 21.11 to 19.98 at .05 level whereas 21.29 and 19.80 at .01 level. This result indicates that out of 100, 95 times the population median will lie between the ranges 21.11 to 19.98 and 99 times between the ranges 21.29 to 19.80. Thus the obtained median is quite reliable as the true median lies within the narrow range at both the levels.

The true SD lies between 5.42 ± 0.1638 × 1.96 at .05 level and 5.42 ± 0.1638 × 2.58 at .01 level. The confidence interval is 5.74 and 5.09 at .05 level whereas 5.84 and 4.99 at .01 level. This result indicates that out of 100, 95 times the population SD will lie between the ranges 5.74 to 5.09 and 99 times between the ranges 5.84 to 4.99. Thus the obtained SD is highly reliable as the true SD lies within the narrow range at both- the levels.

Thus, it can be concluded from these results that all the parameters lie within the narrow ranges and hence the result obtained is highly reliable.

6.4 DECIDING THE NATURE OF DISTRIBUTION OF IETAT SCORES

If the test scores are distributed normally, we assume that the constructed test is satisfactory. While any significant deviation of the distribution to either side, suggests any of the following.

(i) The sample may not be fully representative of the population i.e. it may be biased.

(ii) The sample selected may not be large enough to represent the population.

(iii) Wrong procedures for selecting test-items might have been followed.

(iv) Any other error, likely to affect the distribution adversely, might have been committed in test-construction.

The following three procedures were used to study the distribution of IETAT scores.

(i) Calculation of 'skewness [Sk]' and 'kurtosis [Ku]'

(ii) Graphical representation of test-scores and their interpretation

(iii) Chi-square test of the hypothesis of normal distribution

6.4.1 Calculation of Sk of the Distribution

There are two different formulas for calculating the Sk of the distribution. Sk is calculated by using both the formulas XI and XII given below. For concluding whether the obtained Sk is significant or not, the SE of Sk should be known. Thus the SE of Sk is also calculated along with the Critical Ratio [CR] by using the following formulas XIII and XIV. The calculation of Sk of distribution is presented in table 7.4.

$$\text{Sk} = \frac{3\,(\text{Mean} - \text{Median})}{\text{SD}} \qquad \text{(Formula XI)}$$

(a measure of skewness in terms of frequency distribution)

$$\text{Sk} = \frac{(P_{90} + P_{10})}{2} - P_{50} \qquad \text{(Formula XII)}$$

(a measure of skewness in terms of percentiles)

$$\text{Significance of Sk } (\sigma_{sk}) = \frac{0.5185D}{\sqrt{N}} \qquad \text{(Formula XIII)}$$

(where $D = P_{90} - P_{10} = 28.11 - 13.07 = 15.04$)

$$\text{CR} = \frac{\text{Sk}}{\sigma_{sk}} \qquad \text{(Formula XIV)}$$

Table 6.4

Skewness of the Distribution

	Sk in terms of Frequency Distribution	**Sk in terms of Percentiles**	**Significance**	**Critical Ratio [CR]**
Skewness	+0.01	+0.03	0.3319	0.09

From the table 6.4, it can be seen that the value of Sk obtained in terms of percentile indicates a positive value (+0.03) but slightly higher than the Sk value obtained through in terms of frequency distribution (+0.01). The result obtained in terms of the frequency distribution and percentile differs slightly. According to Garrett, the two measures of Sk computed from different reference values in the distribution and hence are not directly comparable.

Significance of Sk

The SEs for both the formulas used here are not very satisfactory and Garrett states that "the measures of Sk as they stand are often sufficient for many problems in psychology and education." The deviation of our measure of Sk (obtained in terms of percentile) from 0 is +0.03. Thus the value obtained is not at all significant.

Our Sk, therefore, deviates +0.09 σ_k from 0. The CR (0.09) falls well within the limits ± 2.58 which determines the 0.01 level of significance. Hence it is clear that +0.03 represents no real deviation of this frequency distribution from normality.

6.4.2 Calculation of Ku of the Distribution

The calculation of Ku is done by using the formula XV given below. To estimate the significance of deviation of Kurtosis thus obtained from the Ku of the normal curve, the SE of Ku is also calculated by the following formula XVI and the result is presented in table 7.5.

$$Ku = \frac{Q}{P_{90} - P_{10}} \quad \text{(Formula XV)}$$

Where $Q = \frac{P_{75} - P_{25}}{2} = \frac{24.37 - 16.76}{2} = \frac{7.61}{2} = 3.80$

The significance of Ku $(\sigma_{Ku}) = \frac{0.28}{\sqrt{N}}$ (Formula XVI)

$$CR = \frac{D}{\sigma_{Ku}}$$

Where D is the deviation of Ku of the obtained distribution from Ku (0.263) of the normal distribution

Table 6.5

Kurtosis of the Distribution

	Ku	**Significance**	**Critical Ratio [CR]**
Kutosis	0.2527	0.0119	-0.8655

From the table 6.5, it can be observed that the Kurtosis of the frequency distribution is thus equal to 0.2527. The value is slightly less than 0.263 and lower by -0.010. The negative direction of the deviation indicates that the distribution tends slightly towards leptokurtic.

Significance of Ku

The SE of Ku is found to be 0.0119 and the CR (-0.8655) falls well within the ± 1.96 limits which determines the 0.05 level of significance. So it can be concluded that the CR of Ku 0.2527 represents no real deviation of this frequency distribution from normality.

These results were subjected to the graphical representation for further confirmation of conclusions reached earlier.

6.4.3 Graphical Representation of Test Scores

Generally, the tabulated data always convey a picture of a numerical data. However, analyzing numerical data may often be obtained from a graphic or a pictorial treatment of the frequency distribution. Graphic or pictorial device catch the eyes and holds the attention easily than any other form when the most careful array of statistical evidence falls to attract notice. This could be the main reason for using the graphics or pictorial treatment of statistical data so that the researcher can utilize attention seeking the power of visual presentation. The other reason is to seek translation of abstract and difficult numerical data into the concrete and understandable form. Thus it is advisable and convenient to compare the obtained distribution by eye with that normal curve which best fits the data. Garrett also mentions that the direction and extent of asymmetry often strike us more convincingly when seen in a graph than when expressed by measures of

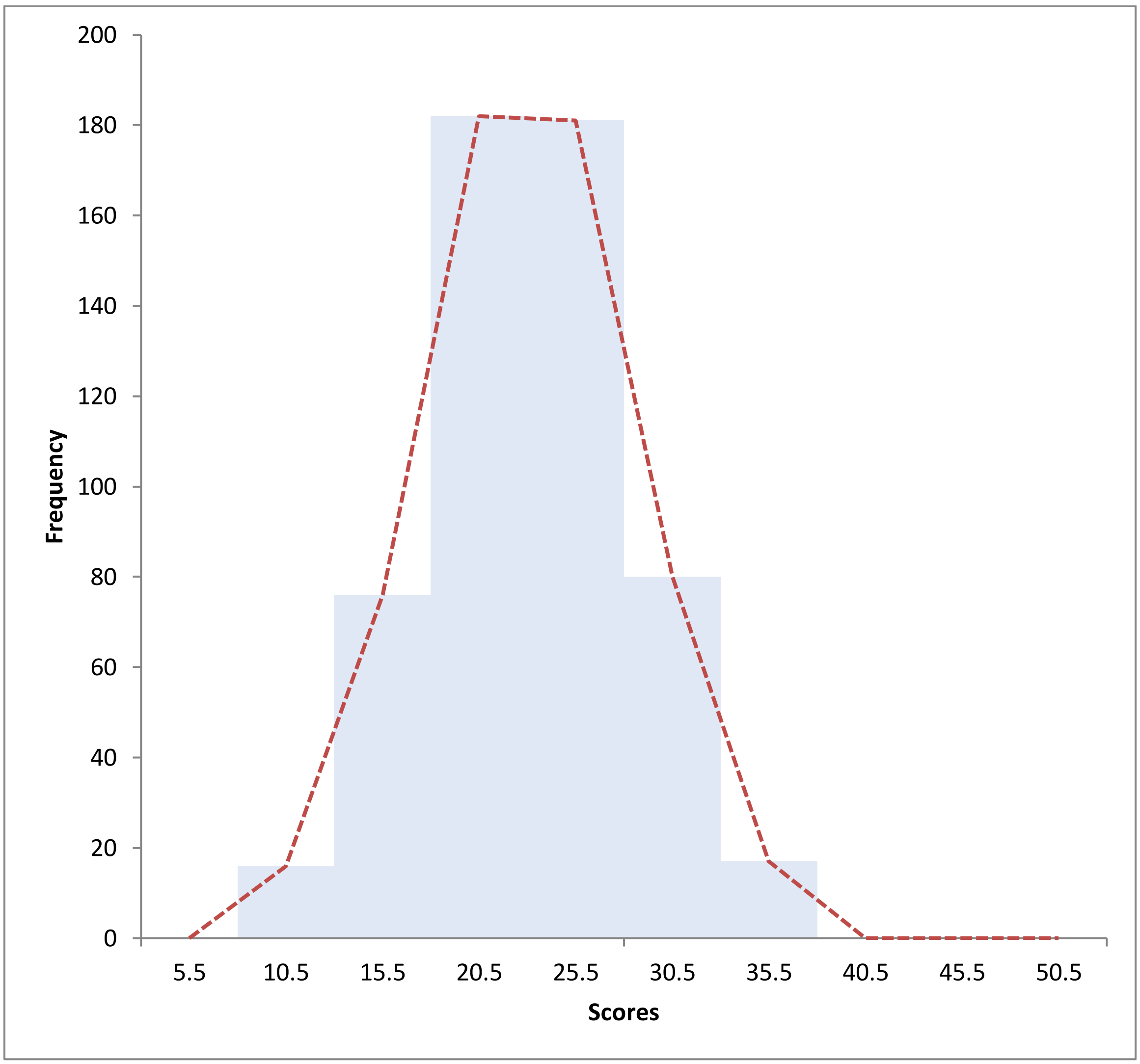
Skewness and kurtosis. Based on the procedure suggested by Garrett (1966), the frequency curves of the whole sample are presented in the graph 6.1 to 6.5.

6.4.3.1 Frequency Polygon

The method given by Garrett in his book for the construction of the frequency polygon was followed. Data used for drawing the frequency polygon are given in table 6.1 on page 83. The polygon is constructed as per the method given by Garrett (1966).

Graph 6.1

Frequency Polygon



Graph 6.1 shows the histogram and frequency polygon on the same axis of a whole sample of 552 pre-service teachers. It can be observed from the graph 6.1 that the distribution towards both the end is normal as the frequency polygon touches the base line at both the ends.

6.4.3.2 The 'Smoothed' Frequency Polygon

In order to iron out chance irregularities, the frequency polygon is 'smoothed' for which a series of 'moving' or 'running' averages are taken from which new or adjusted frequencies are determined. The smoothed frequencies are calculated based on the method given by Garrett (1966) are given in the following table 6.6.

Table 6.6

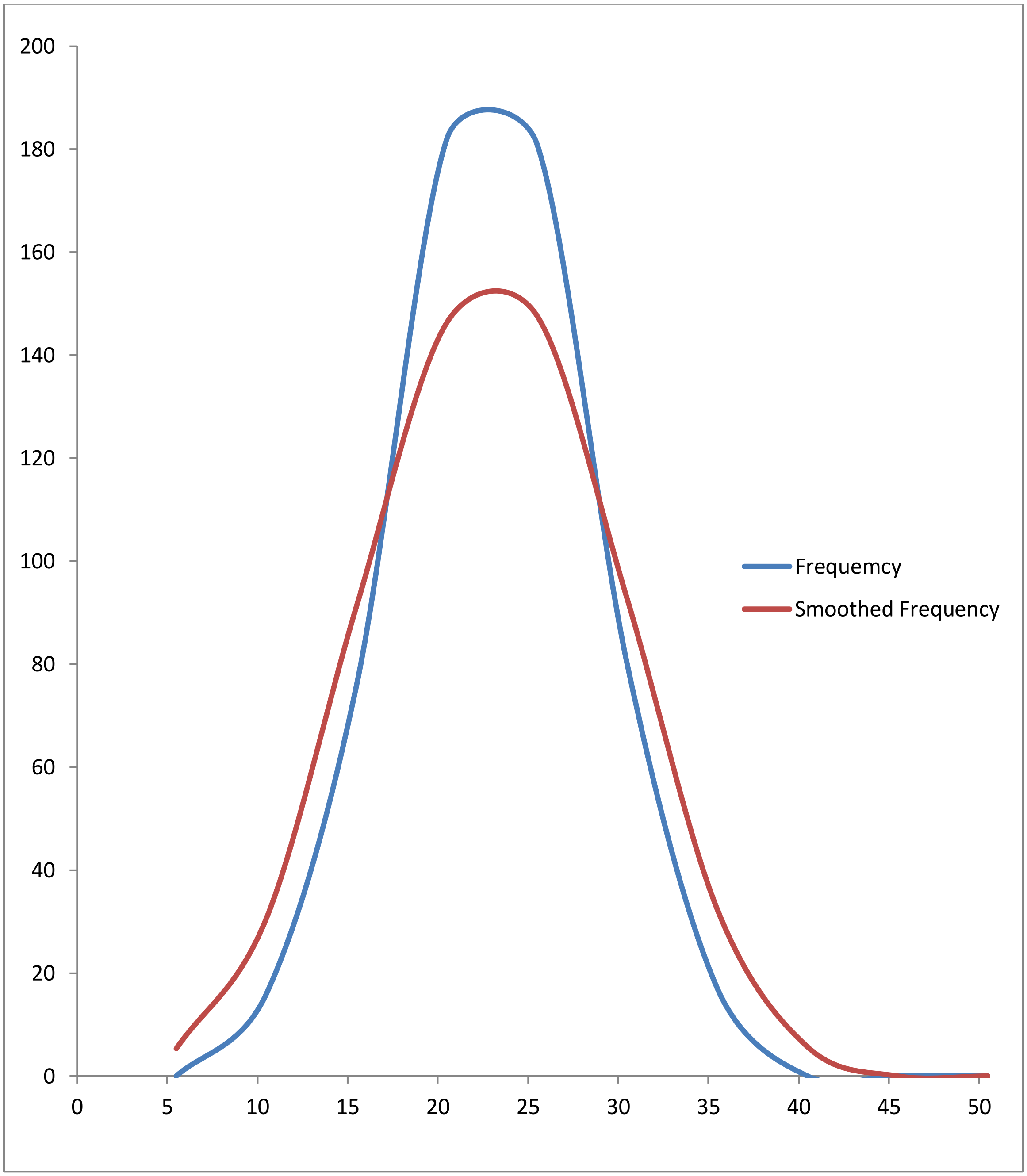
Smoothed Frequencies of the Distribution

Scores	**Original Frequency**	**Smoothed Frequency**
46-50	0	0
41-45	0	0
36-40	0	5.66
31-35	17	32.33
26-30	80	92.66
21-25	181	147.66
16-20	182	146.33
11-15	76	91.33
6-10	16	30.66
1-5	0	5.33
	N=552	

Table 6.6 shows the differences in the original frequencies and smoothed frequencies in the whole sample. After smoothing the frequencies, smoother frequency polygon have been plotted and presented in the following graph 6.2.

Graph 6.2

Smoothed Frequency Polygon



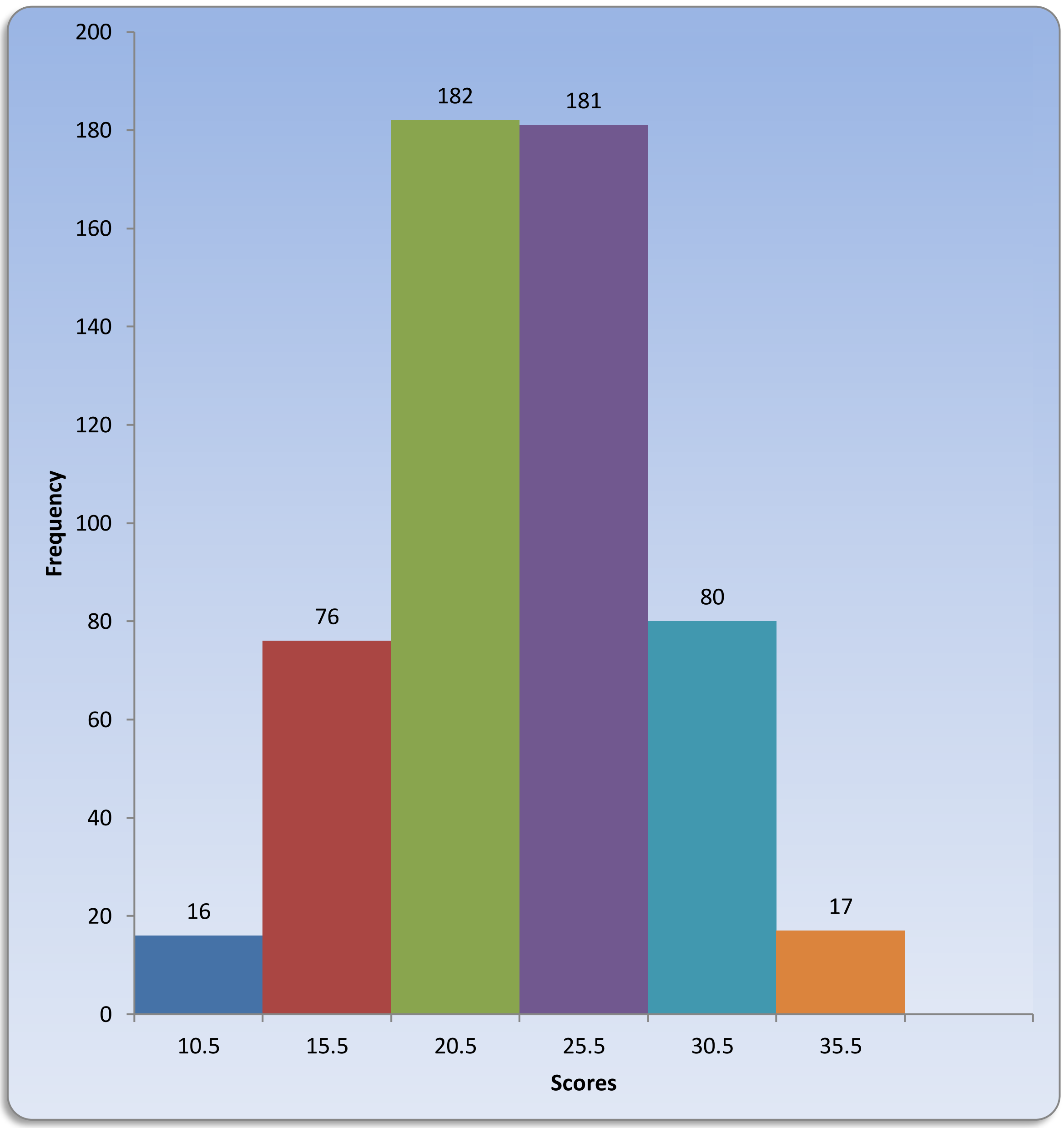
From the graph 6.2 it can be observed that the height of the frequency polygon is more than that of the smoothed frequency polygon. Both the polygons follow the normal distribution as its baseline touches at both the ends of low and high scores.

6.4.3.3 Histogram

A second way of representing a frequency distribution graphically is by means of a 'Histogram' or 'Colum Diagram'. Even here the method given by Garrett (1966) is used for the construction of the histogram. The plotted histogram is given in the following graph 6.3.

Graph 6.3

Histogram

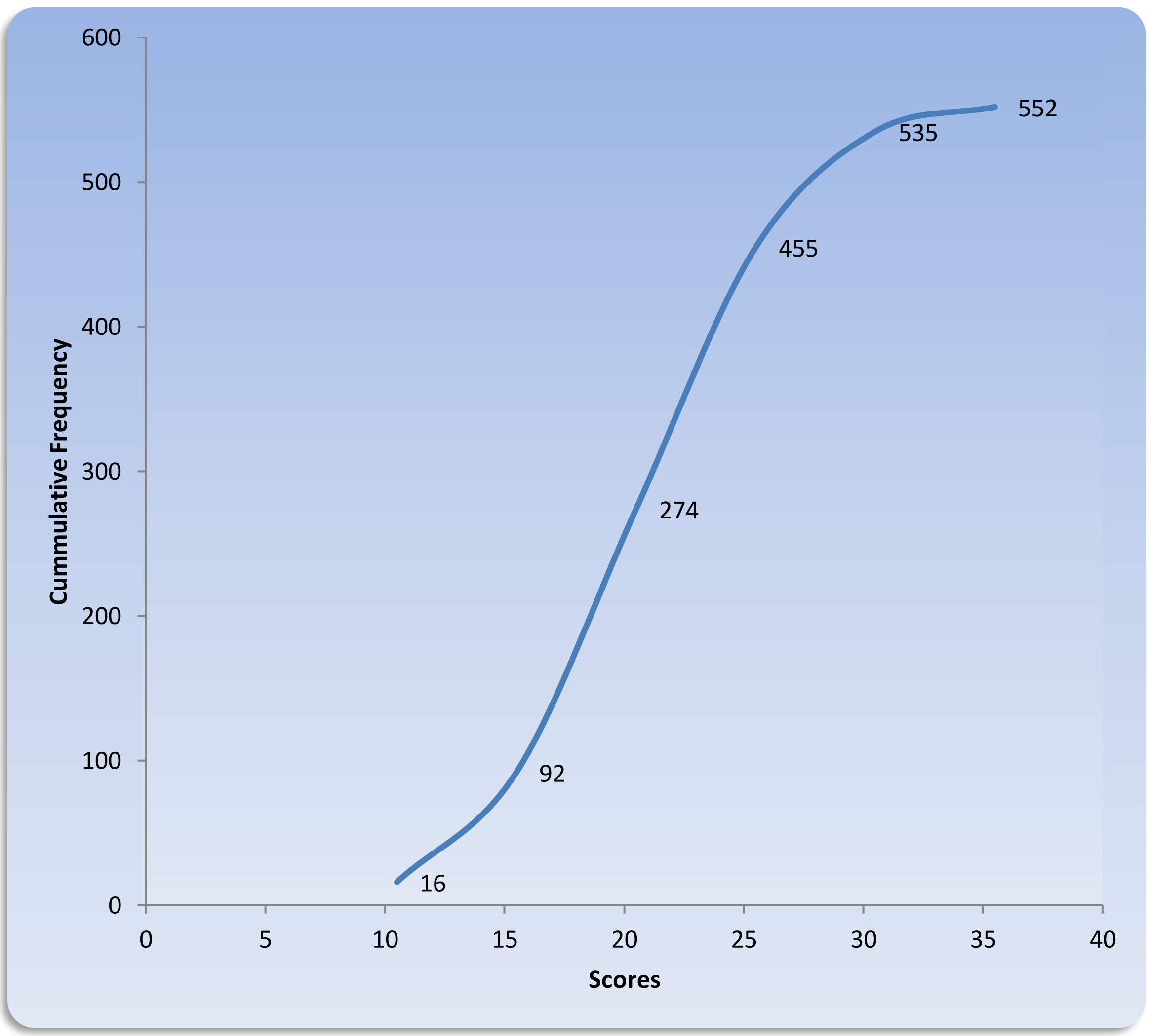


6.4.3.4 Cumulative Frequency Curve

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 6.7. The last cumulative frequency is equal to 552, 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 6.4 given below.

Graph 6.4

Cumulative Frequency Curve



The graph 6.4 shows the cumulative frequency curve for the IETAT scores of 552 pre-service teachers. The plotted points when joined gives the 'S' shaped curve. The obtained 'S' shape depicts the distribution as normal.

6.4.3.5 Cumulative Percentage Curve

The cumulative percentage curve or ogive differs from the cumulative frequency graph in that frequencies are expressed as cumulative percent of N on the Y-axis instead of as cumulative frequencies. The scores and the cumulative percent calculated are given in the following table 6.7.

Table 6.7

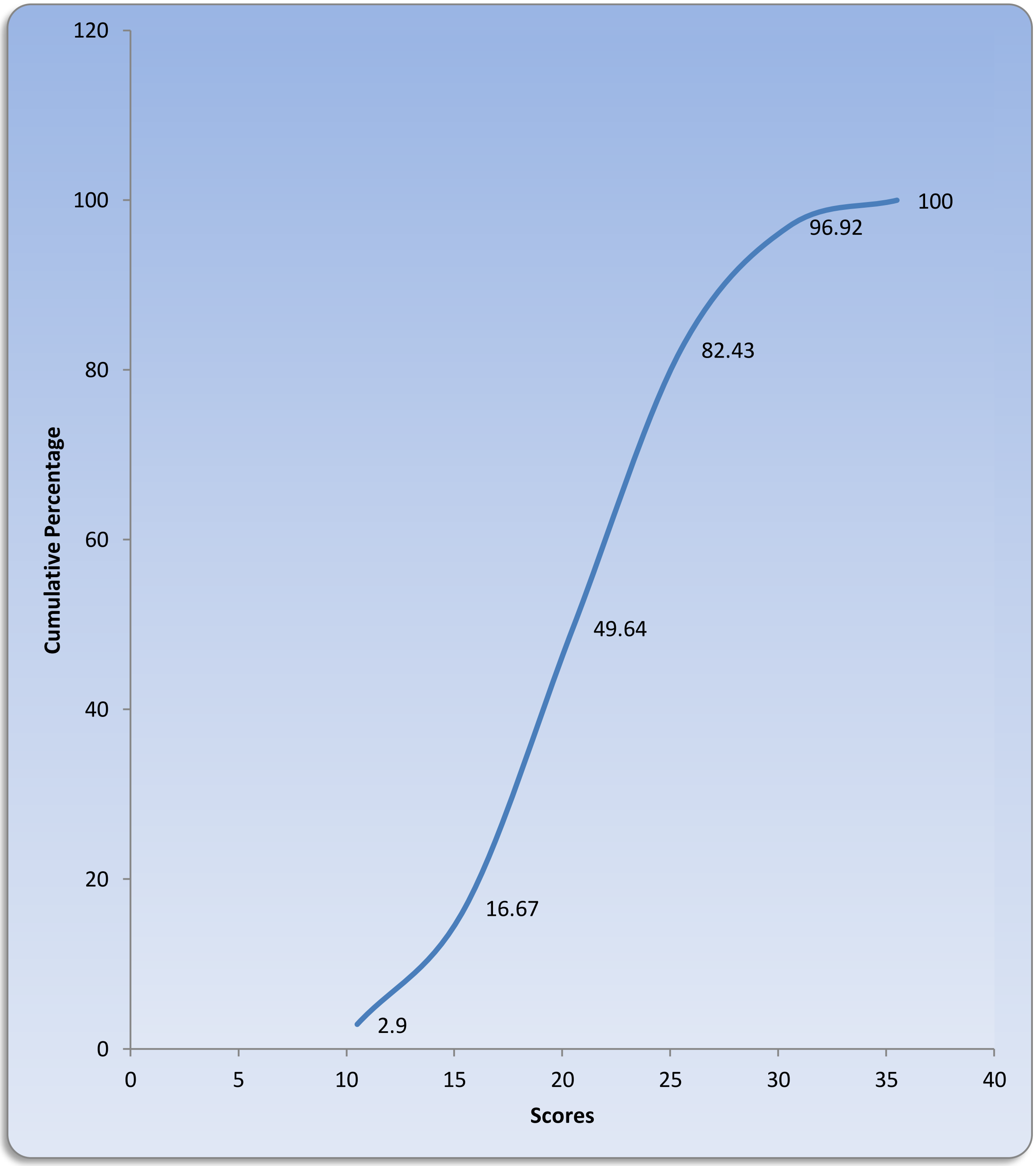
Cumulative Frequencies and Cumulative Percentage of Frequencies

Score Intervals	Exact Score Intervals	Frequency (f)	Percentage (%)	Cumulative Frequency	Cumulative Frequency Percent
31-35	30.5-35.5	17	3.08	552	100
26-30	25.5-30.5	80	14.49	535	96.92
21.25	20.5-25.5	181	32.79	455	82.43
16-20	15.5-20.5	182	32.97	274	49.64
11-15	10.5-15.5	76	13.77	92	16.67
6-10	5.5-10.5	16	2.90	16	2.90
Total		**552**	**100**		

From the table 6.7 it can be observed that 181 pre-service teachers lie in the class interval 21-25 in which mean falls whereas 97 and 274 pre-service teachers lie in the class intervals from 26 to 35 and 6 to 20 i.e. above and below the mean respectively. Further maximum percent (65.79) of frequency lie in the class intervals 21-25 and 16-20. The cumulative percentages are plotted at the exact upper limit of the interval upon which it falls. The graph drawn is given below as graph no 6.5.

Graph 6.5

Cumulative Percentage Curve



The graph 6.5 shows the cumulative percentage curve for the IETAT scores of 552 (100%) pre-service teachers. The plotted points when joined gives the 'S' shaped curve. Thus, the obtained 'S' shape depicts the distribution as normal.

6.4.4 Best Fitting Normal Distribution Curve

The equation of the normal probability curve reads as follows given in formula XVII:

$$Y = \frac{N}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} \quad \text{(Formula XVII)}$$

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

N = Number of cases

σ = Standard deviation of the distribution

π = 3.1416 (the ratio of the circumstance 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_0) should first be calculated. The Y_0 can be determined from the equation of the normal curve.

The 'x' at the mean of the normal curve is '0'. When x = 0, $e^{-\frac{x^2}{2\sigma^2}} = 1$.

$$\text{Thus, } Y_0 = \frac{N}{\sigma\sqrt{2\pi}} \quad \text{(Formula XVIII)}$$

The σ in interval unit is used in the equation since the units on the x-axis are in terms of class-intervals, Y_0 is the frequency at the mean point in the score distribution.

In the present test,

N = 552, σ = 1.09 and $\sqrt{2\pi} = 2.51$

$$\text{Therefore, } Y_0 = \frac{552}{1.09 \times 2.51}$$

$$= \frac{552}{2.7359}$$

$$= 201.76$$

Thus, the value of Y_0 is found equal to 201.76 when the 'x' at the mean of the normal curve is '0'. The values of Y, the heights of the ordinates at different σ – distances from the mean, are found out from the statistical table B for the ordinates of the normal probability curve expressed as fractional parts of the mean ordinate, y_0 (Garrett, 1966, 4th

Indian Reprint-2014, p. 459) and the corresponding values of Y when Y_0 = 201.76 are computed. The final values of the ordinates at different σ distances are given in the following table 6.8.

Table 6.8

Normal Curve Ordinates at Mean

N	**Mean**	**SD* (σ)**	**Y_0 ($\sqrt{2\pi} = 2.51$)**	**σ distance from the Mean**	**Value of Y when Y_0 = 1 (Read from Table)**	**Value of Y when Y_0 = 201.76 (obtained from data)**	**Height of the ordinate**
552	20.57	1.09	201.76	± 0.5 σ	0.88250	0.88250 × 201.76	178.05
				± 1 σ	0.60653	0.60653 × 201.76	122.37
				± 1.5 σ	0.32465	0.32465 × 201.76	65.50
				± 2 σ	0.13534	0.13534 × 201.76	27.30
				± 3 σ	0.01111	0.01111 × 201.76	2.24

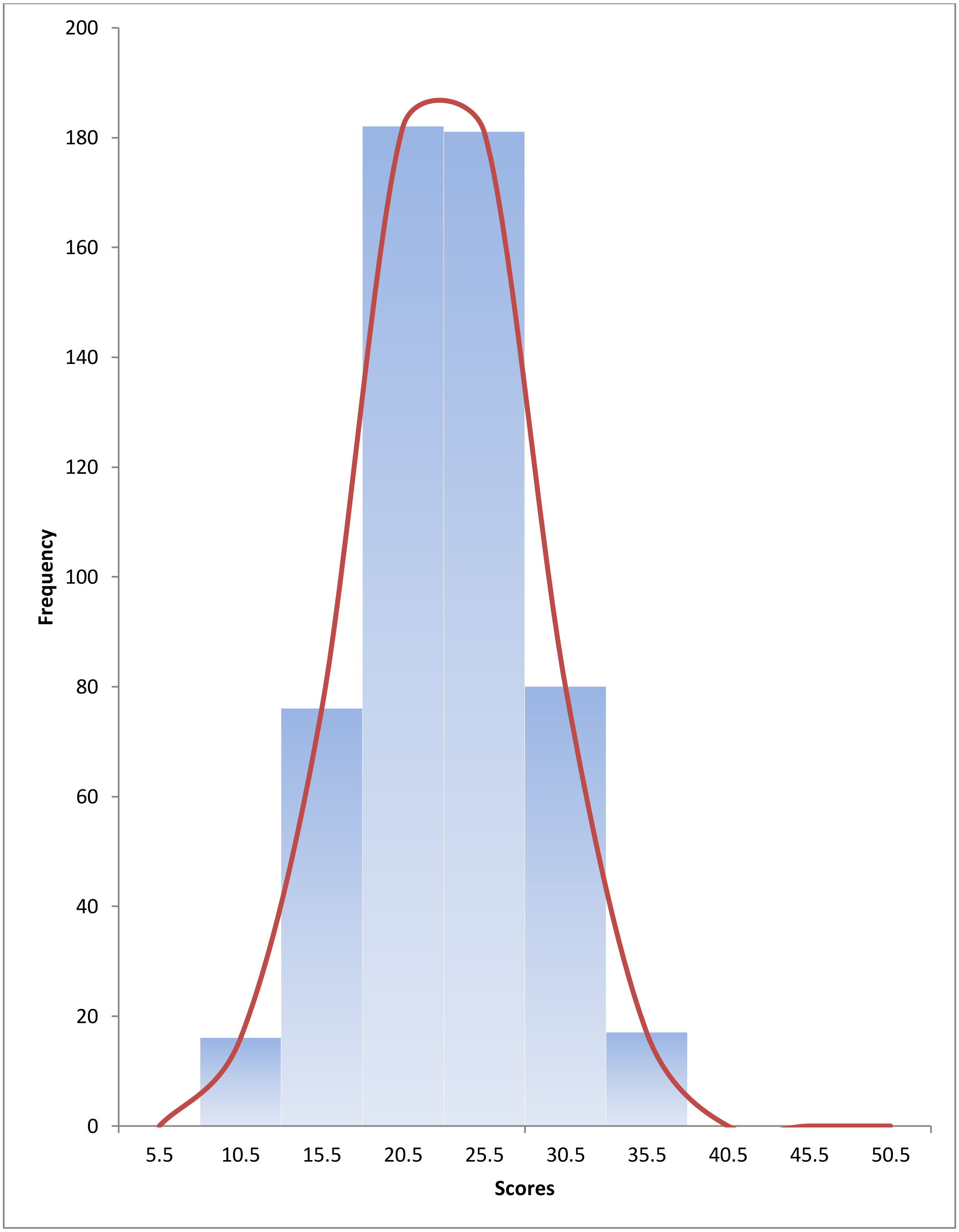
(**in class interval unit*)

The data given in the above table 6.8 were used to super-impose the ideal (best-fitting) normal curve on the obtained histogram given on page 18.

The skewness of the distribution is found to be + 0.03. The value indicates a low degree of positive skewness in the data. The kurtosis of the distribution is 0.2527 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 18, on the whole, fits in with the obtained distribution well enough to warrant our treatment of data as normal.

Graph 6.6

Superimposition of the Best Fitting Curve on the Obtained Histogram for the Whole IETAT



The normal curve ordinates at the mean, ± 1σ, ± 2σ, ± 3σ distances for each of the five sections are calculated and given in the following tables 6.9.

Table 6.9

Normal Curve Ordinates at Mean for Sections I to V of IETAT

Section	N	Mean	SD* (σ)	Y_0 ($\sqrt{2\pi} = 2.51$)	σ distance from the Mean	Value of Y when $Y_0 = 1$ (Read from Table)	Value of Y when $Y_0 = 39.55$ (obtained from data)	Height of the ordinate
I	552	5.39	0.9783	224.39	± 1 σ	0.60653	0.60653 × 224.39	136.09
					± 2 σ	0.13534	0.13534 × 224.39	30.36
					± 3 σ	0.01111	0.01111 × 224.39	2.49
II	552	3.55	0.7990	276.00	± 1 σ	0.60653	0.60653 × 276.00	167.40
					± 2 σ	0.13534	0.13534 × 276.00	37.35
					± 3 σ	0.01111	0.01111 × 276.00	3.06
III	552	5.18	1.005	219.04	± 1 σ	0.60653	0.60653 × 219.04	132.85
					± 2 σ	0.13534	0.13534 × 219.04	29.64
					± 3 σ	0.01111	0.01111 × 219.04	2.43
IV	552	3.35	0.7755	284.54	± 1 σ	0.60653	0.60653 × 284.54	172.58
					± 2 σ	0.13534	0.13534 × 284.54	38.50
					± 3 σ	0.01111	0.01111 × 284.54	3.16
V	552	3.38	0.7471	293.62	± 1 σ	0.60653	0.60653 × 293.62	178.08
					± 2 σ	0.13534	0.13534 × 293.62	39.74
					± 3 σ	0.01111	0.01111 × 293.62	3.26

(**in class interval unit*)

6.4.5 Chi-square Test of the Hypotheses of Normal Distribution

The chi-square test is found to be quite useful in testing hypotheses. It is the sum ratios in which 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 hypotheses to be tested here are:

- The distribution of the scores on the IETAT follows the normal curve.
- If there is any discrepancy between the observed and expected frequencies is insignificant and is due to chance factor(s) only.

The procedure discussed in Biometrica tables for statisticians (Pearson, 1914), is followed thoroughly for calculating chi-square values. The values of $P_{(x)}$ are measured directly from the New Tables of the Probability Integral given by Sheppard (1905). The values of $\Delta P_{(x)}$ are obtained by subtracting the value of $P_{(x)}$ by its upper value. The expected frequencies are calculated by multiplying the value of $\Delta P_{(x)}$ with the total sample size 552.

The values of ‘df’ indicated in the table 6.10 to 6.15 are 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 section also have been calculated with a view to studying the nature and role of each section in the whole test. The scores of the section should also be tested to find out whether they are also distributed normally. The result of the chi-square test for the whole test is presented in table 6.10 whereas section wise result of chi-square test is presented in the tables 6.11 to 6.15.

Table 6.10

Chi-square Test of Normal Distribution for Whole Test

Score Intervals	Exact Score Interval	f_o	X	X–M	$\frac{X-M}{\sigma}$	Area $P_{(x)}$	$\Delta P_{(x)}$	$f_e = N\times\Delta P_{(x)}$	f_o-f_e	$(f_o-f_e)^2$	$\frac{(fo-fe)^2}{f_e}$
31-35	30.5-35.5	17	30.5	9.93	1.83	0.96637	0.03366	19	2	4	0.2105
26-30	25.5-30.5	80	25.5	4.93	0.90	0.81593	0.15044	83	3	9	0.1084
21-25	20.5-25.5	181	20.5	-0.07	-0.01	0.47210	0.34383	190	9	81	0.4263
16-20	15.5-20.5	182	15.5	-5.07	-0.94	0.17361	0.29849	165	17	289	1.7515
11-15	10.5-15.5	76	10.5	-10.07	-1.86	0.03145	0.14216	78	2	4	0.0513
6-10	5.5-10.5	16	5.5	-15.07	-2.78	0.00272	0.02873	16	0	0	0
		N=552									**2.548**
Mean = 20.57 SD = 5.45		Degrees of Freedom (df) = 3					From the table at .01 level = 11.345 at .05 level = 7.815 **Thus the value obtained is less than the table values at both the levels and therefore the value obtained is not significant at both .01 and .05 levels.**				

Table 6.11

Chi-square Test of Normal Distribution for Section I

Score Intervals	**Exact Score Interval**	**f_o**	**X**	**X–M**	$\frac{X-M}{\sigma}$	**Area $P_{(x)}$**	**$\Delta P_{(x)}$**	**$f_e = N \times \Delta P_{(x)}$**	**$f_o - f_e$**	**$(f_o - f_e)^2$**	$\frac{(fo-fe)^2}{f_e}$
9-10	8.5-10.5	15	8.5	3.16	1.53	0.93699	0.06301	35	20	400	11.43
7-8	6.5-8.5	156	6.5	1.16	0.56	0.71226	0.22473	124	32	1024	8.258
5-6	4.5-6.5	214	4.5	-0.84	-0.40	0.34458	0.36768	203	11	121	0.596
3-4	2.5-4.5	118	2.5	-2.84	-1.38	0.08380	0.26078	144	26	676	4.694
0-2	0-2.5	49	0	-5.34	-2.59	0.00480	0.07900	44	5	25	0.568
		N=552									**= 25.546**
Mean = 5.39 SD = 1.96		Degrees of Freedom (df) = 2					From the table at .01 level = 9.210 at .05 level = 5.991 **Thus the value obtained is higher than the table values at both the levels and therefore the value obtained is significant at both .01 and .05 levels.**				

Table 6.12

Chi-square Test of Normal Distribution for Section II

Score Intervals	Exact Score Interval	f_o	X	X–M	$\frac{X-M}{\sigma}$	Area $P_{(x)}$	$\Delta P_{(x)}$	$f_e = N \times \Delta P_{(x)}$	$f_o - f_e$	$(f_o - f_e)^2$	$\frac{(fo-fe)^2}{f_e}$
7-8	6.5-8.5	21	6.5	3.09	1.74	0.95907	0.06093	34	13	169	4.970
5-6	4.5-6.5	120	4.5	1.09	0.61	0.72906	0.23001	127	7	49	0.3858
3-4	2.5-4.5	262	2.5	-0.91	-0.51	0.30503	0.42403	234	28	784	3.350
0-2	0-2.5	149	0	-3.41	-1.92	0.02743	0.27760	153	4	16	0.104
		N=552									**= 8.8098**
Mean = 3.55 SD = 1.60		Degrees of Freedom (df) = 1					From the table at .01 level = 6.635 at .05 level = 3.841 **Thus the value obtained is higher than the table values at both the levels and therefore the value obtained is significant at both .01 and .05 levels.**				

Table 6.13

Chi-square Test of Normal Distribution for Section III

Score Intervals	Exact Score Interval	f_o	X	X–M	$\frac{X - M}{\sigma}$	Area $P_{(x)}$	$\Delta P_{(x)}$	$f_e = N \times \Delta P_{(x)}$	$f_o - f_e$	$(f_o - f_e)^2$	$\frac{(fo - fe)^2}{f_e}$
9-10	8.5-10.5	13	8.5	3.4	1.62	0.94738	0.05262	29	16	256	8.828
7-8	6.5-8.5	141	6.5	1.4	0.67	0.74857	0.19881	110	31	961	8.736
5-6	4.5-6.5	202	4.5	-0.6	-0.29	0.38591	0.36266	200	2	4	0.02
3-4	2.5-4.5	135	2.5	-2.6	-1.24	0.10749	0.27842	154	19	361	2.344
0-2	0-2.5	61	0	-5.10	-2.43	0.00755	0.09994	55	6	36	0.655
		N=552									**= 20.538**
Mean = 5.18 SD = 2.01		Degrees of Freedom (df) = 2					From the table at .01 level = 9.210 at .05 level = 5.991 **Thus the value obtained is higher than the table values at both the levels and therefore the value obtained is significant at both .01 and .05 levels.**				

Table 6.14

Chi-square Test of Normal Distribution for Section IV

Score Intervals	Exact Score Interval	f_o	X	X–M	$\frac{X-M}{\sigma}$	Area $P_{(x)}$	$\Delta P_{(x)}$	$f_e = N \times \Delta P_{(x)}$	$f_o - f_e$	$(f_o - f_e)^2$	$\frac{(fo - fe)^2}{f_e}$
7-8	6.5-8.5	6	6.5	3.33	1.89	0.97062	0.02938	16	10	100	6.25
5-6	4.5-6.5	129	4.5	1.33	0.76	0.77637	0.19425	107	22	484	4.523
3-4	2.5-4.5	235	2.5	-0.67	-0.38	0.35198	0.42439	234	1	1	0.004
0-2	0-2.5	182	0	-3.17	-1.80	0.03594	0.31604	174	8	64	0.368
		N=552									**= 11.145**
Mean = 3.35 SD = 1.55		Degrees of Freedom (df) = 1					From the table at .01 level = 6.635 at .05 level = 3.841 **Thus the value obtained is higher than the table values at both the levels and therefore the value obtained is significant at both .01 and .05 levels.**				

Table 6.15

Chi-square Test of Normal Distribution for Section V

Score Intervals	Exact Score Interval	f_o	X	X–M	$\frac{X-M}{\sigma}$	Area $P_{(x)}$	$\Delta P_{(x)}$	$f_e = N \times \Delta P_{(x)}$	f_o-f_e	$(f_o-f_e)^2$	$\frac{(fo-fe)^2}{f_e}$
7-8	6.5-8.5	11	6.5	3.49	2.07	0.98077	0.01923	11	0	0	0
5-6	4.5-6.5	106	4.5	1.49	0.88	0.81057	0.17020	94	12	144	1.532
3-4	2.5-4.5	275	2.5	-0.51	-0.30	0.38209	0.42848	237	38	1444	6.093
0-2	0-2.5	160	0	-3.01	-1.78	0.03754	0.34455	190	30	900	4.737
		N=552									**= 12.362**
Mean = 3.38 SD = 1.49		Degrees of Freedom (df) = 1					From the table at .01 level = 6.635 at .05 level = 3.841 **Thus the value obtained is higher than the table values at both the levels and therefore the value obtained is significant at both .01 and .05 levels.**				

From the table 6.10, it can be observed that a perusal of the chi-square value shows that the value is not significant at .05 and .01 levels for the whole test. Thus the hypothesis, 'Distribution of the scores on the IETAT follows normal curve' is retained and it can be said that the distribution of the scores on the IETAT follows the normal curve. This result is also supported by the graphical representation of the result and superimposition of the best fitting curve on the obtained histogram given on page 99. However, from the table 6.11 to 6.15 it can be seen that section wise obtained chi-square values is significant at .05 and .01 levels for the five sections of IETAT.

Thus based on the above result, it can be concluded that the distribution of scores in the IETAT followed the normal curve since the chi-square value obtained for the whole test is not significant at .05 and .01 levels.

The comparison of the results obtained from the study of the nature of curves by three different techniques is presented in the following table 6.16.

Table 6.16

Normal Curve Ordinates at Mean for IETAT

Sr. No.	**Technique**	**Result**
1	Nature of curve revealed through measures of divergence	Normal distribution
2	Nature of curve revealed through superimposition of an ideal curve	Normal distribution
3	Nature of curve revealed through chi-square test	Normal distribution
Conclusion		**Normal distribution**

Based on the above results shown in the table 6.16, it can, thus, be concluded safely that the distribution of the total scores, though, not perfectly normal, is much sensibly normal.

6.5 RELIABILITY OF THE IETAT

Determining of the reliability of the test is most essential characteristics of a good measuring instrument and therefore it is necessary to check the reliability of the constructed test. Any measuring device must fulfill the conditions of reliability and validity. Reliability of the measuring device mentions consistency of scores when re-examined on the same sample at different times (Anastasi & Urbina, 1997). There are different methods for establishing the reliability of a test such as split-half method, rational equivalence method, Hoyt's technique, alternate or parallel forms and test-retest method. The reliability of the present test is estimated by using the following three methods.

(i) Split-half method
(ii) K-R formula 20
(iii) Hoyt's 'Analysis-of-variance' technique

The 'test-retest' method is not applied as it was not feasible due to four-month internship of sampled pre-service teachers. The 'parallel-forms' method also is not applied as it is not feasible to construct parallel forms of the test. The application of each method to estimate the reliability of the IETAT is presented below.

6.5.1 Application of Split Half Method

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'.

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 items 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 occasion, hence variation introduced by differences between the two testing situations are eliminated. Therefore, 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 IETAT. A small sample of 110 (20% of total sample size i.e. 552) respondents was selected for the purpose of applying 'split-half' method to estimate the reliability of whole test. The testis 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 'Spearman-Brown Prophecy' formula XIX given below.

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

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

The scatter-gram of scores used in split half method is given in the following table 6.17.

Table 6.17

Scattergram of Scores used in Split Half Method

	Scores	**Even Item Scores**							**fy**
		1-3	**4-6**	**7-9**	**10-12**	**13-15**	**16-18**	**19-21**	
Odd Item Scores	**19-21**	--	--	--	--	--	--	--	--
	16-18	--	--	--	--	--	4	2	6
	13-15	--	--	--	--	19	1	--	20
	10-12	--	--	--	36	6	--	--	42
	7-9	--	6	27	1	--	--	--	34
	4-6	2	6	--	--	--	--	--	**8**
	1-3	--	--	--	--	--	--	--	**--**
fx		**2**	**12**	**27**	**37**	**25**	**5**	**2**	**110**

Product Moment r = 0.943 P.E.r. = 0.001

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

$$= \frac{2 \times 0.943}{1 + (2-1)\, 0.943}$$

$$= \frac{1.886}{1 + 1 \times 0.943}$$

$$= \frac{1.886}{1.943}$$

$$= 0.97$$

$$\text{P.E.'r'} = 0.6745 \times \frac{1 - r^2}{\sqrt{N}}$$

$$= 0.6745 \times \frac{1 - (0.97)^2}{\sqrt{552}}$$

$$= 0.6745 \times \frac{1 - 0.9409}{23.49}$$

$$= 0.6745 \times \frac{0.0591}{23.49}$$

$$= 0.6745 \times 0.0025$$

$$= 0.0001$$

The reliability coefficient based on a test half as long as the entire test is 0.943. The reliability of the entire test calculated by the Spearman-Brown Prophecy formula is found to be 0.97. Since the reliability coefficient of the IETAT is considerably high, the test may be considered to be highly reliable.

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 – Z function and find the standard error of Z function. The formula XX given below is for the SE of Z (σ_Z).

$\sigma_Z = \frac{1}{\sqrt{N-3}}$ where N = 552 (Formula XX)

Standard error of Z,

$$
\begin{aligned}
SE_Z &= \sigma_Z \\
&= \frac{1}{\sqrt{552-3}} \\
&= \frac{1}{\sqrt{549}} \\
&= \frac{1}{23.430} \\
&= 0.0427
\end{aligned}
$$

From the table of conversion of Pearson r into a corresponding Fisher's z coefficient, we read that an r of 0.994 corresponds to a Z of 2.65.

The value of Z of 2.65 for corresponding r of 0.97 ranges between 2.72 and 2.57 (i.e. 2.65 $\pm$ 1.96 $\times$ 0.04) converting these values of Z's back into r's we get a confidence interval from 0.990 to 0.985, 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. (i) its sampling distribution is approximately normal and (ii) it's SE depends only upon the size of the sample N and is independent of the size of r.

6.5.2 Application of the Kuder-Richardson Method [‘Whole Test’ – ‘Total Sample’]

Dissatisfied with the Split-half Method, Kuder and Richardson developed a new procedure 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 XXI used here is given below.

$r_{11} = \frac{n}{n-1} \times \frac{\sigma_t^2 - \sum pq}{\sigma_t^2}$ (Formula XXI)

Where,

r_{11} = Reliability coefficient of the whole test

n = Number of items in the test

σ_t = Standard deviation of the test scores

p = Proportion of the group answering a test item correctly

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

To apply this method also, the original sample of 552 respondents was used. Thus in the above formula XX, n is equal to 552 and σ_t is equal to 5.45. The proportion of the group answering a test item correctly is found out for each of the 50 test items. From the values of ‘p’, the corresponding values of ‘q’ are calculated. The values of ‘pq’ for each of the items are calculated and the sum of all ‘pq’ values is found to be 10.72. Item wise calculated values of ‘pq’ are given in the following table 6.18.

Table 6.18

‘pq’ Values of 50 Test Items (n=552)

Item No.	M’ (Correct Responses)	n - M’ (Incorrect Responses)	P $\left(\frac{M'}{n}\right)$	q $\left(\frac{n-M'}{n}\right)$	‘pq’
1	366	186	0.6630	0.3369	0.2233
2	333	219	0.6632	0.3967	0.2630
3	391	161	0.7083	0.2916	0.2065
4	174	378	0.3152	0.6847	0.2158
5	408	144	0.7391	0.2608	0.1927
6	316	236	0.5724	0.4275	0.2447
7	183	369	0.3315	0.6884	0.2282
8	452	100	0.8188	0.1811	0.1482
9	134	418	0.2427	0.7572	0.1837
10	223	329	0.4039	0.5960	0.2407
11	224	328	0.4057	0.5942	0.2410
12	171	381	0.3097	0.6902	0.2137
13	185	367	0.3351	0.6648	0.2227
14	321	231	0.5815	0.4184	0.2432
15	207	345	0.3750	0.6250	0.2343
16	73	479	0.1322	0.8677	0.1147
17	195	357	0.3532	0.6467	0.2284
18	193	359	0.3496	0.6503	0.2273
19	136	416	0.2463	0.7536	0.1856
20	224	328	0.4057	0.5942	0.2410
21	281	271	0.5090	0.4909	0.2498
22	288	264	0.5217	0.4782	0.2494
23	451	101	0.8170	0.1829	0.1494
24	332	220	0.6014	0.3985	0.2396
25	252	300	0.4565	0.5434	0.2480

Item No.	**M’ (Correct Responses)**	**n - M’ (Incorrect Responses)**	**P** $\left(\frac{M'}{n}\right)$	**q** $\left(\frac{n-M'}{n}\right)$	**‘pq’**
26	243	309	0.4402	0.5597	0.2463
27	222	330	0.4021	0.5978	0.2403
28	157	395	0.2844	0.7155	0.2034
29	250	302	0.4528	0.5471	0.2477
30	358	194	0.6485	0.3514	0.2278
31	170	382	0.3079	0.6920	0.2130
32	106	446	0.1920	0.8079	0.1551
33	158	394	0.2862	0.7137	0.2042
34	148	404	0.2681	0.7318	0.1916
35	248	304	0.4492	0.5507	0.2473
36	242	310	0.4384	0.5615	0.2461
37	245	307	0.4438	0.5561	0.2467
38	246	306	0.4456	0.5543	0.2469
39	126	426	0.2282	0.7717	0.1761
40	98	454	0.1775	0.5224	0.0927
41	241	311	0.4365	0.5634	0.2459
42	269	283	0.4873	0.5126	0.2497
43	119	433	0.2155	0.7844	0.1690
44	283	269	0.5126	0.4873	0.2497
45	178	374	0.3224	0.6775	0.2184
46	212	340	0.3840	0.6159	0.2365
47	136	416	0.2463	0.7536	0.1856
48	106	446	0.1920	0.8079	0.1551
49	174	376	0.3152	0.6811	0.2146
50	127	425	0.2300	0.7699	0.1770
					$\sum pq = 10.7216$

The reliability coefficient is calculated using the K-R formula XX.

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

$$= \frac{50}{50-1} \times \frac{(5.42)^2 - 10.72}{(5.42)^2}$$

$$= \frac{50}{49} \times \frac{29.38 - 10.72}{29.38}$$

$$= \frac{50}{49} \times \frac{18.66}{29.38}$$

$$= \frac{50 \times 18.66}{49 \times 29.38}$$

$$= \frac{933}{1439.62}$$

$$= 0.648$$

The reliability coefficient of the present IETAT as measured by K-R formula is 0.648 which is lower than the result obtained by the ‘Split half method’.

6.5.3 Application of Hoyt’s Method [‘Whole Test’ – ‘Total Sample’]

To apply this method, the total sample of 552 was used. All the test-scripts were arranged according to their rank order. The ‘student-item’ chart was constructed. A specimen of the same is given below in table 6.19.

Table 6.19

A Specimen of ‘Student-Item’ Chart

Respondent	**Items**						**Each Respondent’s Total Score ‘t’**	**t^2**
	1	**2**	**3**	**4**	**….**	**50 (n)**		
1							t_1	t_1
2							t_2	t_2
3							t_3	t_3
:							:	:
:							:	:
552 (k)							t_{552}	t^2_{552}
Numbers of correct responses on each item: ‘p’	p_1	p_2	p_3	p_4	….	p_{50}	$\sum_{i=1}^{n} pi = \sum_{i=1}^{n} ti$	
p^2	$p_1{}^2$	$p_2{}^2$	$p_3{}^2$	$p_4{}^2$	….	$p^2{}_{50}$		
	$\sum t^2 = 249465$				$\sum p^2 = 3001837$			

In the table 6.19,

n = number of items

k = number of individuals

p = number of correct responses on each item

t = each respondent's total score

$\sum p =$ sum of all 50 p's

$\sum t =$ sum of all 552 t's

$\sum p^2$ = sum of all p^2's

$\sum t^2$ = sum of all t^2's

After preparing the 'Student-Item' chart, the 'analysis of variance' table was constructed and necessary calculations were made to fill the blanks in the table. The duly completed table is shown below in table 6.20.

Table 6.20

An Analysis-of-Variance Table with the Formulas for Computing the Values under Each Heading and the Actual Values under Each Heading

Source of Variance	Degrees of Freedom (df)	Sum of Squares	Mean of Squares	
Between individuals	k – 1 = 551	$\frac{1}{n}\sum t^2 - \frac{(\sum t)^2}{nk}$ = 301	$\frac{\text{Sum of squares}}{\text{df}}$ = 0.5453	(a)
Between items	n – 1 = 49	$\frac{1}{k}\sum p^2 - \frac{(\sum t)^2}{nk}$ = 750	$\frac{\text{Sum of squares}}{\text{df}}$ = 15.31	(b)
Residual	(n-1) (k-1) = 26999	Total – (between individuals + between items) = 5636	$\frac{\text{Sum of squares}}{\text{df}}$ = 0.2087	(c)
Total	nk – 1 = 27599	$\frac{(\sum t)(nk - \sum t)}{nk}$ = 6687	--	

The details of above calculation given in the table 6.20, is presented below.

Degrees of freedom between individuals = k – 1 = 552 – 1 = 551

Degrees of freedom between items = n – 1 = 50 – 1 = 49

Residual = (n-1) (k-1) = (50 – 1) (552 – 1) = (49) (551) = 26999

Total degree of freedom = nk – 1 = (552 × 50) – 1 = 27600 – 1 = 27599

Sum of square between individuals = $\frac{1}{n}\sum t^2 - \frac{(\sum t)^2}{nk}$

$= \frac{1}{50} \times 249465 - \frac{(11375)^2}{27600}$

$= 4989 - 4688$

$= 301$

Sum of square between items = $\frac{1}{k}\sum p^2 - \frac{(\sum t)^2}{nk}$

$= \frac{1}{552} \times 3001837 - \frac{(11375)^2}{27600}$

$= 5438 - 4688$

$= 750$

Total sum of squares = $\frac{(\sum t)(nk - \sum t)}{nk}$

$= \frac{(11375)\,(27600 - 11375)}{27600}$

$= \frac{(11375)\,(16225)}{27600}$

$= 6687$

Residual sum of squares = total – (between items + between individuals)

= 6687 – (301 + 750)

= 6687 – 1051

= 5636

Mean of squares for individuals = $\frac{301}{551}$ = 0.5453

Mean of squares for items = $\frac{750}{49}$ = 15.31

Mean of squares for residual = $\frac{5636}{26999}$ = 0.2087

The reliability coefficient of the IETAT was obtained by the following formula:

$r_{tt} = \frac{a - c}{a}$

$= \frac{0.5453 - 0.2087}{0.5453}$

$= \frac{0.3366}{0.5753}$

$= 0.617$

The reliability coefficient obtained by this method is 0.617 which is less than that of obtained by the ‘split-half’ method.

The results obtained by the use of different methods are summarized in the following table 6.21.

Table 6.21

Reliability Coefficient of IETAT obtained through Different Methods

Sr. No.	Method Used	Reliability Coefficient obtained	P.E.r
1	Split-half method	0.97	0.0003
2	Kuder-Richardson method	0.65	-
3	Hoyt's Method	0.62	-

From the table 6.21, it can be seen that the reliability coefficient obtained through the application of 'split-half' method gives higher value (0.97). This might be due to the tendency of 'split-half' method to give the high value of reliability coefficient. The K-R formula 20 (0.65) and Hoyt's method (0.62) gives little identical results. Thus the reliability of the present IETAT can be fixed at 0.75 and the value showed that the constructed IETAT is reliable at the satisfactory level.

6.5.4 Reliability in terms of True Scores and Measurement Errors

- ***Reliability co-efficient as a measure of true variance***

The variance of the obtained scores can be divided into variance of the true scores and the variance of chance errors.

The reliability of the present test is fixed at .75 therefore, 75 percent of the variance of test scores is true variance and 25 percent error variance.

- ***Estimating true scores by way of the regression equation and the reliability coefficient***

The regression equation (Formula XXII) estimates true score is given below:

$X_\infty = r_{tt} X_1 + (1 - r_{tt}) M_1$ (Formula XXII)

Where, X_∞ = estimated true score on the test

X_1= obtained score on the test

M_1= mean of test distribution (20.57)

r_{tt} = reliability coefficient of the test (0.75)

$$X_\infty = 0.6 \times 35 + (1 - 0.75)\, 20.57$$
$$= 28.7 + (0.25)\, 20.57$$
$$= 21 + 5.14$$
$$= 26.14$$

The standard error of an estimated true score is given by the following formula XXII.

$SE_\infty = SE_{\infty 1} = \sigma_1 \sqrt{r_{tt} - r^2{}_{tt}}$ (Formula XXII)

Where, σ_1= 5.45 and r_{tt}= 0.75

$$SE_{\infty 1} = 5.45\sqrt{0.75 - (0.75)^2}$$
$$= 5.45\sqrt{0.75 - 0.56}$$
$$= 5.45\sqrt{0.19}$$
$$= 5.45 \times 0.4359$$
$$= 2.38$$

The 0.95 confidence interval is $X_\infty \pm 1.96 \times 2.38$ i.e. $\pi_\infty \pm 4$

- ***The index of reliability***

The correlation between a set of obtained scores and their corresponding true counterparts was found by measuring the index of reliability. For this, the following formula XXIII was employed.

$r_{1\infty} = \sqrt{r_{tt}}$ (Formula XXIII)

Where, $r_{1\infty}$ = the correlation between the obtained and true scores

r_{tt} = the reliability coefficient of the test

$$r_{1\infty} = \sqrt{0.75}$$
$$= 0.866$$

Thus, 0.866 is the maximum correlation which the test is capable of yielding in the present form.

The reliability of the present test is estimated by applying three different methods. It is found to be 0.75 and the value is seen to be satisfactory as far 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. Hence the validity of the measuring instrument is also an important characteristic.

6.6 VALIDITY OF THE IETAT

Validity refers to the truthfulness of the test and is always its most important characteristics. No matter what other merits the test may possess, if it lacks validity, it is not worth its use. Thorndike and Hagen (1958) classified two main types of validity rationale and experimental or statistical. Rational validity includes content and concept or construct validity while experimental or statistical validity includes congruent, concurrent and predictive validity.

The purpose of the present test is to measure inclusive education teaching aptitude and it itself suggest the predictive validity. The predictive validity of the test is the main aspect of this IETAT. Besides this, the content validity is also obtained. The congruent validity was not possible to be established as no aptitude test of the same nature is available in this country. Concurrent validity was also not possible as it concerned with the contemporary criterion while the present IETAT is designed to forecast a future performance in inclusive education teaching future. The validity estimated for the present test is discussed thoroughly as below.

6.6.1 Content Validity

Content validity has been decided on SMEs’ rating on the items constructed and the validity index of each item. The procedure of the Lawshe’s method for determining content validity of items is already described in chapter III of this report. SMEs' rating for each item of the final version of IETAT is given in chapter V.

The Content Validity Ratio [CVR] was calculated for each item based on the formula given by Lawshe (1975). The average of the CVR across all items on the test was found to be 0.64

6.6.2 Construct Validity

Construction of items and its selection in the test has been described thoroughly in chapter III and IV of this report. For validity index, items showing .20 and more than that have been selected while items below .20 validity index have been rejected. All the items in all the sections were included after a careful scrutiny. Only the valid items were selected while the items that were not found to be valid were summarily rejected.

The criterion related validity is again of two types: Convergent validity and predictive validity. Since there is no other aptitude test of this type available in this country, convergent validity of the present IETAT is out of question and thus, the predictive validity is determined for the present test.

6.6.3 Predictive Validity

The predictive validity of the constructed test is determined experimentally by finding the correlation between the test and some independent criterion. The coefficient of validity of a test is the coefficient of correlation between test scores and criterion score. The criterion against which the present IETAT validated is the examination marks of the pre-service teachers in the subject of Creating Inclusive School. The criterion was selected after a careful scrutiny and all the SMEs had agreed that the criterion is satisfactorily reliable and valid. However, it should be noted here that no criterion is a perfect one and it is extremely difficult to fix up criteria to judge success in teaching. So the present criterion also cannot be a perfect one, but the investigator is confident that it is reliable and valid to the extent it is possible to reach in the present circumstances and limited researches in the field.

For estimating predictive validity, 20% (110 pre-service teachers) sample is selected for determining the validity of the test. The raw scores on total test obtained by 110 pre-service teachers were converted into the standard scores. The raw test scores and raw criterion scores are converted into standard scores with the help of the formula XXIV given below. The raw scores indicated in the table 7.18 and 7.19 are expressed in standard scores in a distribution where M = 23.90, $\sigma = 4.53$, M' = 50 and σ' = 10.

$$X' = \frac{\delta'}{\delta}(X - M) + M' \text{ (Formula XXIV)}$$

Where,

X = a score in the original distribution

X' = a standard score on the new distribution

M = mean of the raw score

M' = mean of the standard score distribution

δ = SD of raw score

δ' = SD of standard score

Table 6.22

Raw Scores of the Final Test and their Corresponding Standard Scores

Raw Scores	Standard Scores	Raw Scores	Standard Scores	Raw Scores	Standard Scores	Raw Scores	Standard Scores	Raw Scores	Standard Scores
1	13.99	11	32.39	21	50.79	31	69.19	41	87.59
2	15.83	12	34.23	22	52.63	32	71.03	42	89.43
3	17.67	13	36.07	23	54.47	33	72.87	43	91.27
4	19.51	14	37.91	24	56.31	34	74.71	44	93.11
5	21.35	15	39.75	25	58.15	35	76.55	45	94.95
6	23.19	16	41.59	26	59.99	36	78.39	46	96.79
7	25.03	17	43.43	27	61.83	37	80.23	47	98.63
8	26.87	18	45.27	28	63.67	38	82.07	48	100.47
9	28.71	19	47.11	29	65.51	39	83.91	49	102.31
10	30.55	20	48.95	30	67.35	40	85.75	50	104.15

The examination scores of the subject of Creating Inclusive School taken as criterion scores and converted into standard scores with the help of the formula XXIV given above. The raw scores are expressed in standard scores in a distribution where M = 39.36, $\sigma = 3.56$, M' = 50 and σ' = 10.

Table 6.23

Raw Criterion Scores (External Examination Marks in Inclusive Education Subject) and their Corresponding Standard Scores

Raw Scores	Standard Scores	Raw Scores	Standard Scores
26	12.59	36	40.59
27	15.39	37	43.39
28	18.19	38	46.19
29	20.99	39	48.99
30	23.79	40	51.79
31	26.59	41	54.59
32	29.39	42	57.39
33	32.18	43	60.19
34	34.99	44	62.99
35	37.79	45	65.79

The two sets of scores were arranged in the form of a scatter diagram and the product-moment coefficient of correlation was calculated. The scatter diagram pertaining to the standard test scores and criterion inclusive education examination scores is given in the table 6.24.

Table 6.24

Scatter Diagram of Standard Test Score and Standard Criterion Scores (External Examination Marks in Inclusive Education Subject)

Scores		**Standard Criterion Scores**								**fy**
		11-15	**16-20**	**21-25**	**26-30**	**31-35**	**36-40**	**41-45**	**46-50**	
Standard Test Scores	46-50	--	--	--	--	--	23	--	--	**23**
	41-45	--	--	--	--	--	12	--	--	**12**
	36-40	--	--	--	--	1	--	--	--	**1**
	31-35	--	--	--	--	6	--	--	--	**6**
	26-30	--	--	--	--	2	--	--	--	**2**
	21-25	--	--	--	2	--	--	8	38	**48**
	16-50	--	--	--	--	3	10	--	--	**13**
	11-15	--	--	--	5	--	--	--	--	**5**
Fx		**0**	**0**	**0**	**7**	**12**	**45**	**8**	**38**	**110**

Product Moment r = 0.5195 P.E.r. = 0.1871

The product-moment coefficient of correlation ‘r’ was calculated according to the usual procedure from the scatter diagram shown in table 6.19. The value obtained in this case is found to be 0.5195. The probable error of this ‘r’ also was calculated. It was found to be 0.01871. The predictive validity of the present test is much satisfactory and it can be said that the test is a good predictor of inclusive education teaching aptitude. Even then, obviously, other things being equal, the higher the correlation, the better.

According to Bingham (1937), the validity of prediction can be increased in the following ways.

(i) By improving the adequacy and reliability of the measures of success used as criteria.

(ii) By choosing those traits for testing which are most indicative of aptitude for achieving the success specified.

(iii) By choosing the tests which most accurately measure those traits.

(iv) By ascertaining the proper weight to be assigned to the scores in each factor, in order most closely to predict the criterion.

While constructing the test, utmost care was taken in selecting factors indicative of teaching aptitude for inclusive education in constructing test items and finally in securing adequate and reliable criterion. However, the investigator feels that adding more items and assigning proper regression weights to test score could increase the test validity.

Gulliksen (1950) suggested that multiplying the length of a test by K increases the validity. For this, he gave the following formula.

$$R_{k1} = \frac{r_{cx}\sqrt{K}}{\sqrt{1+(K-1)r_{tt}}} \qquad \text{(Formula XXV)}$$

Where,

R_{k1} = the augmented validity coefficient

r_{cx} = validity coefficient of the test

r_{tt} = reliability coefficient of the test

K = the number of times the test is increased in length

If this formula XXV is applied and we increase the IETAT length two or three times then the following result obtained.

$$R_{k1} = \frac{0.51\sqrt{2}}{\sqrt{1+(2-1)0.60}} = \frac{0.51 \times 1.41}{\sqrt{1+0.60}} = \frac{0.7191}{\sqrt{1.60}} = \frac{0.7191}{1.2649} = 0.57$$

$$R_{k1} = \frac{0.51\sqrt{3}}{\sqrt{1+(3-1)0.60}} = \frac{0.51 \times 1.73}{\sqrt{1+1.2}} = \frac{0.8823}{\sqrt{2.2}} = \frac{0.8823}{1.48} = 0.59$$

Thus, application of the Gulliksen (1950) formula shows that the increase in the test length will increase the test validity only to some extent but not substantially.

The validity of the present IETAT is, therefore, finally accepted to be 0.519 ± 0.18.

6.7 NORMS OF THE IETAT

The most difficult phase of aptitude testing is the interpretation of results. After the test has been carefully administered and painstakingly scored, the findings must be appraised and translated into information helpful to the individuals 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. For the present IETAT results, the standard-score norms, T-score norms, and percentile norms are established.

6.7.1 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 researchers. The raw scores obtained on the present IETAT are converted into the standard scores with the help of the formula VI given on the page no. 123 in a distribution of M' = 100 and σ' = 20 as well as in a distribution of M' = 50 and σ' = 10. The shift from raw to standard score requires a linear transformation. This transformation does not change the shape of the distribution in any way. The standard test scores obtained are given in the table 6.25 with their corresponding raw scores.

Table 6.25

Raw Scores of the Final Test along with their Corresponding Standard-Scores and T-Scores

Raw Scores	Standard Scores		T-Scores	Raw Scores	Standard Scores		T-Scores
	M'=100, σ'=20	M'=50, σ'=10			M'=100, σ'=20	M'=50, σ'=10	
1	28.18	14.18	--	26	119.93	59.94	61
2	31.85	16.01	--	27	123.60	61.77	63
3	35.52	17.85	--	28	127.27	63.60	64
4	39.19	19.68	--	29	130.94	65.43	66
5	42.86	21.51	--	30	134.61	67.26	68
6	46.53	23.34	--	31	138.28	69.10	71
7	50.20	25.17	12	32	141.95	70.92	75
8	53.87	27	24	33	145.62	72.75	76
9	57.54	28.83	28	34	149.29	74.58	76
10	61.21	30.66	30	35	152.96	76.41	81
11	64.88	32.49	32	36	156.63	78.24	--
12	68.55	34.32	33	37	160.30	80.10	--
13	72.22	36.15	35	38	163.97	81.90	--
14	75.89	37.98	37	39	167.64	83.73	--
15	79.56	39.81	39	40	171.31	85.56	--
16	83.23	41.64	41	41	174.98	87.39	--
17	86.90	43.47	43	42	178.65	89.22	--
18	90.57	45.30	45	43	182.32	91.05	--
19	94.24	47.13	47	44	185.99	92.88	--
20	97.91	48.96	49	45	189.66	94.71	--
21	101.58	50.79	51	46	193.33	96.54	--
22	105.25	52.62	53	47	197	98.37	--
23	108.92	54.45	55	48	200.67	100.20	--
24	112.59	56.28	57	49	204.34	102.02	--
25	116.26	58.11	59	50	208	103.86	--

6.7.2 The T-Score Norm

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 'normalized' σ scores by transforming them directly into equivalent points in a normal distribution. Normalized standard scores are generally called T-scores. T-scaling was devised by McCall (1949). T-scores are normalized standard scores converted into a distribution with a mean of 50 and σ of 10. The procedure suggested by Garrett (1966) is followed in the calculation of the T-scores. The model table for calculation of T-scores is given below.

Test Score	**F**	**Cf**	**Cf below Score + $\frac{1}{2}$ of Given Score**	**Column IV Values in %**	**T-Score**
I	**II**	**III**	**IV**	**V**	**VI**

The T-scores are given in the table 6.26 with their corresponding raw-scores and standard-scores.

Table 6.26

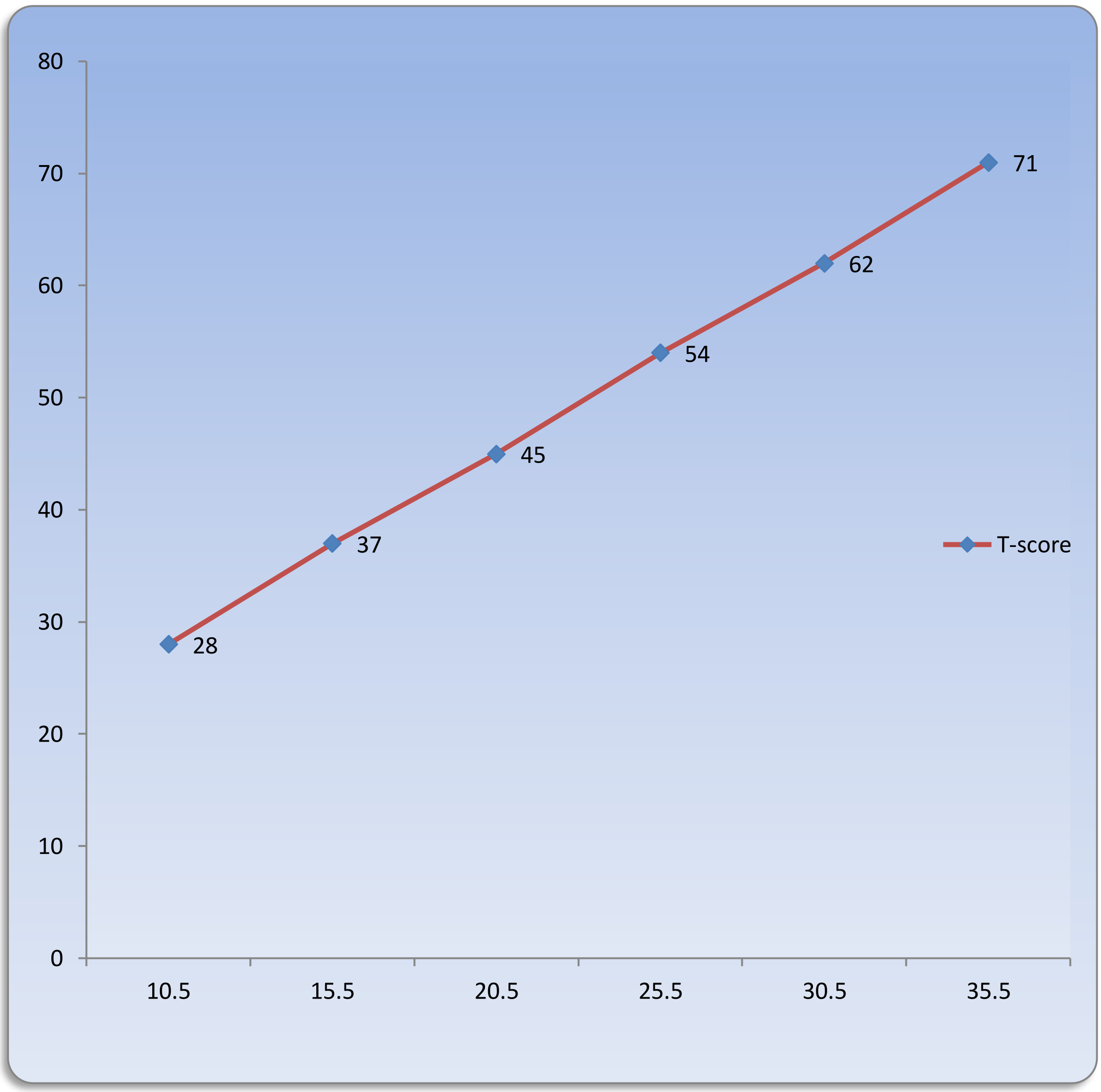
T-Score Values for the Distribution

Score Interval	**f**	**Cf**	**Cf below Score + $\frac{1}{2}$ of Given Score**	**Column IV Values in %**	**T-Score**
I	**II**	**III**	**IV**	**V**	**VI**
31-35	17	552	543.5	98.46	71
26-30	80	535	495	89.67	62
21-25	181	455	364.5	66.03	54
16-20	182	274	183	33.15	45
11-15	76	92	54	9.782	37
6-10	16	16	8	1.449	28
	N = 552				

A graph is also 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. For any integral raw score points, the corresponding T-score points could be found out from the graph 6.7.

Graph 6.7

Corresponding T-Scores of Raw Scores



From the graph 6.7, it can be observed that the points fall on a straight line and it shows the distribution is normal.

6.7.3 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 tests, interest inventories, personality inventories and rating scales. In the present test, it is essential to find out percentile norms of IETAT scores in terms to decide the nature of the distribution of IETAT scores.

The formula given in the Statistics in Psychology and Education (Garrett, 1966) is used for calculating the percentiles. The percentiles calculated with the help of the formula XXVI given below and the percentiles calculated are given in the following table 6.27.

$P_P = l + \left(\frac{P_N - F}{f_P}\right) \times i$ (Formula XXVI)

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

Table 6.27

Percentile Norms

Percentile	Score	Percentile	Score	Percentile	Score	Percentile	Score	Percentile	Score
P_1	7.23	P_{11}	13.44	P_{21}	16.16	P_{31}	17.67	P_{41}	19.19
P_2	8.95	P_{12}	13.80	P_{22}	16.30	P_{32}	17.83	P_{42}	19.34
P_3	10.54	P_{13}	14.16	P_{23}	16.46	P_{33}	17.98	P_{43}	19.49
P_4	10.90	P_{14}	14.53	P_{24}	16.61	P_{34}	18.13	P_{44}	19.64
P_5	11.26	P_{15}	14.89	P_{25}	16.76	P_{35}	18.28	P_{45}	19.79
P_6	11.63	P_{16}	15.26	P_{26}	16.91	P_{36}	18.43	P_{46}	19.94
P_7	11.98	P_{17}	15.55	P_{27}	17.06	P_{37}	18.58	P_{47}	20.10
P_8	12.35	P_{18}	15.70	P_{28}	17.21	P_{38}	18.74	P_{48}	20.25
P_9	12.71	P_{19}	15.85	P_{29}	17.37	P_{39}	18.89	P_{49}	20.40
P_{10}	13.07	P_{20}	16.00	P_{30}	17.52	P_{40}	19.03	P_{50}	20.56

Percentile	Score	Percentile	Score	Percentile	Score	Percentile	Score	Percentile	Score
P_{51}	20.70	P_{61}	22.23	P_{71}	23.76	P_{81}	25.28	P_{91}	28.46
P_{52}	20.86	P_{62}	22.39	P_{72}	23.90	P_{82}	25.43	P_{92}	28.80
P_{53}	21.01	P_{63}	22.54	P_{73}	24.06	P_{83}	25.70	P_{93}	29.15
P_{54}	21.17	P_{64}	22.69	P_{74}	24.21	P_{84}	26.04	P_{94}	29.49
P_{55}	21.32	P_{65}	22.84	P_{75}	24.37	P_{85}	26.39	P_{95}	29.84
P_{56}	21.47	P_{66}	23.00	P_{76}	24.52	P_{86}	26.73	P_{96}	30.18
P_{57}	21.62	P_{67}	23.15	P_{77}	24.67	P_{87}	27.07	P_{97}	30.63
P_{58}	21.78	P_{68}	23.30	P_{78}	24.82	P_{88}	27.42	P_{98}	32.25
P_{59}	21.93	P_{69}	23.45	P_{79}	24.98	P_{89}	27.77	P_{99}	33.88
P_{60}	22.08	P_{70}	23.60	P_{80}	25.13	P_{90}	28.11	P_{100}	35.50

6.7.3.1 Percentile Ranks [PR]

The PR corresponding to the raw scores obtained are also calculated. The distinction between percentile and PR is that in calculating percentiles we start with a certain percent of N e.g. 15% or 62%. Then we count into the distribution. The given percent and the point reached is the required percentile i.e. P_{15} or P62. The procedure followed in computing PR is the reverse of the process. Here we begin with an individual score and determine the percentage of scores which lies below it. If the score is 15, the score has a PR on a scale of 100. The procedure given by Garrett for computing PR is followed. The formula XXVII used for calculating PR. The PR corresponding to each raw score is given in the table 6.28.

$$PR_x = \left(\frac{C_f + \left(\frac{x-ll}{w}\right) \times fi}{n} \right) \times 100 \qquad \text{(Formula XXVII)}$$

Where,

PR_x = Percentile rank of score x

ll = Lower limit of the interval containing the score x

n = Total number of cases

cf = Cumulative frequency of scores below the interval containing score x

f_i = Frequency of scores in the interval containing x

w = Width of the class interval

Table 6.28

Percentile Ranks

Raw Score	Percentile Rank	Raw Score	Percentile Rank
1	--	26	83.80
2	--	27	86.70
3	--	28	89.60
4	--	29	92.49
5	--	30	95.39
6	--	31	97.14
7	0.87	32	97.76
8	1.45	33	98.37
9	2.03	34	98.99
10	2.60	35	99.69
11	4.27	36	--
12	7.02	37	--
13	9.77	38	--
14	12.53	39	--
15	15.28	40	--
16	19.94	41	--
17	26.53	42	--
18	33.12	43	--
19	39.11	44	--
20	46.30	45	--
21	52.87	46	--
22	59.22	47	--
23	65.97	48	--
24	72.53	49	--
25	79.07	50	--

Percentile and percentile ranks can also be read directly from a cumulative percentage curve also called an ogive. Ogive is S-shaped curve and can be easily compared than ordinary cumulative curve because of their common height. The nature of the curve also testifies the normality of distribution of test scores. The ogive and cumulative frequency graph are drawn from the data presented in the table 6.1 on page no. 94. The ogive and cumulative frequency graph 6.4 and 6.5 are shown on the page nos. 94 and 96 respectively.

6.7.4 Letter Grades

The respondents can also be assigned letter grades in accordance with the raw scores they obtain on the present IETAT. The result presented above indicates the test-scores distribution as normal. The respondents were grouped into six grades viz. A, B. C. D, E and F. For this, the baseline of the normal curve was divided into six equal parts, (5б ÷ 6 = 0.83б) each part being equal to 0.83б unit. Then the limits of raw as well as standard scores for different letter grades were marked off. These limits are shown in the table 6.29 and the graph 6.8.

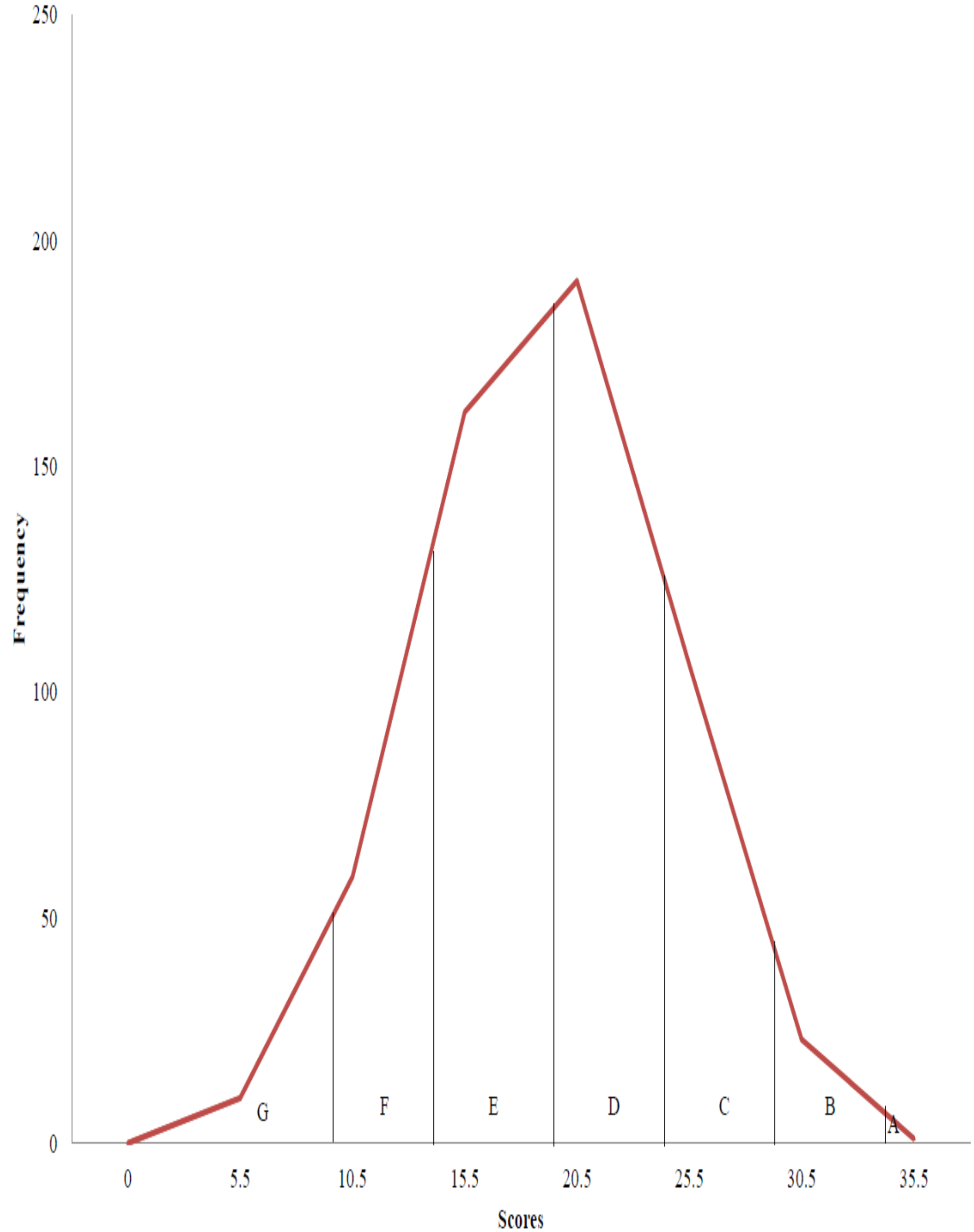
Table 6.29

Assigning Letter Grade to IETAT Scores

Letter Grade	**Limits in terms of б-units**	**Limits in terms of raw scores**	**Limits in terms of standard scores (M'=50, σ'=10)**
A	M + 2.56 б and above	35 and above	76 and above
B	Between M + 1.75 б and M + 2.56 б	Between 30 and 34	Between 67 and 75
C	Between M + 0.78 б and M + 1.75 б	Between 25 and 29	Between 58 and 65
D	Between M - 0.28 б and M + 0.78 б	Between 20 and 24	Between 48 to 57
E	Between M – 1.29 б and M – 0.28 б	Between 15 and 19	Between 41 to 47
F	Between M – 2.15 б and M – 1.29 б	Between 10 and 14	Between 30 to 40
G	Between M – 2.15 б and below	9 and below	29 and below

Graph 6.8

Schematic Representation of the Grade Assignment



6.7.5 Using the Test Norms

The norms established through the present study may be applied in comparing the performance of other samples from the population of pre-service teachers. The pre-service teachers can easily be placed in particular grades and extent of their future success as the inclusive school teachers can be judged. Thus the norms established by this study will be helpful to a reasonable extent in screening out the teachers who really possess aptitude for teaching in inclusive schools.

The test norms will be mainly useful in selecting prospective teachers for pre-service level training and they are most likely to be successful as teachers in inclusive education after the training. As there is a great dearth of trained teachers in inclusive education, the inclusive schools are required to employ untrained teachers and such school authorities may use the present IETAT in appointing the teachers who have teaching aptitude for inclusive education.

The test can also be helpful for vocational guidance purpose. If the test is applied to fresh graduate or post-graduate candidate and if it is found out that s/he possesses a good deal of teaching aptitude towards inclusive education, s/he may be advised to take up the teaching profession by joining either in general or special teacher education courses.