## 8·1. INTRODUCTION

In this chapter we shall be concerned with statistical methods applicable in the field of Psychology and Education. A new discipline called 'Psychometry' has been developed as a branch of psychology which deals with the measurement of psychological traits or the mental abilities like intelligence, aptitude, opinion, interest, personality or scholastic achievement, etc. When individuals are ranked (or arranged) in an ordinal series according to their scholastic achievement, then the problem relates to the education. As such, the educational statistics may also be considered as a part of psychometry when individuals are arranged in a series with respect to some attribute (or trait) and these ranks will give us the serial position of the object in the group. In fact, the psychological and educational traits or characteristics are rather abstract in nature and they can be measured only with some approximation.

For the measurement of psychological and educational characteristics (which are rather abstract in nature as compared with physical or biological characteristics) and consequently for the scaling of psychological and educational data, various devices, many of them based upon the use of normal probability curve, have been used. Here the most practical consideration is that the scales for different tests should be comparable. Although arbitrary depending upon the choice of the investigator, the scale units should be equal, meaningful and stable and should provide comparability of the means, dispersions and form of the distribution. Although the zero point of psychological scale is arbitrary, the distances from arbitrary zero are additive. In other words, *psychological scale is an interval scale and not a ratio scale* since there is no absolute zero point. In the following Section (§ 8·2 to § 8·2·4), we shall discuss briefly some of the commonly employed scaling procedures.

## 8·2. SOME SCALING PROCEDURES

In this Section, we shall discuss some of the common scaling procedures used in psychology and education :

(*i*) Scaling individual test items in terms of difficulty.

(*ii*) Scaling of scores on a test : $Z$ (or $\sigma$) score, standard scores, Normalised scores, $T$-scores, and percentile scores.

(*iii*) Scaling of rankings in terms of normal probability curve.

(*iv*) Scaling of ratings in terms of normal probability curve.

In the scaling procedures developed for psychological and educational research it is often assumed that the variable trait ($X$) is normally distributed (over population of elements whose trait we want to measure). The origin and unit of measurement of the scale may be chosen arbitrarily, but they should remain fixed throughout the use of the scale over a group of subjects.

**8·2·1. Scaling Individual Test Items in Terms of Difficulty.** In this case a number of problems or test items, say $n$ all designed to test the same psychological trait or test, are administered to a large group of individuals who are selected at random out of those for whom the final test is intended and we are interested in arranging these items in order of difficulty, say, from very simple to very difficult. For this the set of problems is given to a group of individuals for solving them and for each problem the proportion of those who could solve it is obtained. Thus the proportion $p_i$ of the individuals solving the $i$th problem ($i = 1, 2, \ldots, n$) is given by :

$$p_i = \frac{\text{Number of individuals answering } i\text{th problem correctly}}{\text{Number of individuals taken in the group}}$$

and thus items can be arranged in order of '*percentage difficulty*'. Of course, the larger the percentage of the people passing a particular test item, the lower it is in order of difficulty. Thus, for example, an item answered successfully by 80% of the individuals is obviously much easier as compared to a problem solved correctly by only 45%. But comparison of percentage difficulty is only a crude method, since these percentages do not successfully reflect the differences in difficulty.

In the construction of the difficulty scale we assume that the ability or the trait ($X$) being measured is distributed normally about some mean $\mu$ and standard deviation $\sigma$. Without loss of generality we can assume $\mu = 0$. Under the assumption of the normality of the trait ($X$) the heterogeneity (or variability) of the group provides a better difficulty scale, known as $\sigma$-scale. In fact, the *difficulty value of an item is usually defined as the minimum ability to answer this item correctly under the assumption that the ability is distributed normally* $N(0, \sigma^2)$. If $p_i$ is the proportion of the individuals answering $i$th item successfully then its *difficulty value is given by* $\sigma > z_i$, where $z_i$ is determined from the following relation :

$$P(Z > z_i) = \frac{1}{\sqrt{2\pi}} \int_{z_i}^{\infty} e^{-t^2/2}\, dt = p_i\,, \ (i = 1, 2, \ldots, n) \qquad \ldots (8\cdot1)$$

where $Z \sim N(0, 1)$. $z_i$'s are also known as $\sigma$-distances from the mean.

For given $p_i$'s, the values of $z_i$'s can be read from the table of areas under standard normal probability curve.

It will be seen that some of the $z_i$'s computed from (8·1) will be negative while others will be positive. It is somewhat difficult to compare a negative value with a positive value. To overcome this difficulty, we shift the origin in $\sigma$-distances from mean ($z_i$'s) or in difficulty values ($\sigma z_i$) to some suitable constant $\theta$ corresponding to zero point of the difficulty scale. $\theta$ is usually taken as $-3\sigma$ but is also sometimes taken as the value corresponding to the %age of individuals failing all items or passing one or other item (*level of minimum difficulty*). For example, suppose that 4% of the entire group fail to answer correctly a single problem. Then this percentage will be represented by the left tail area of 0·04 under normal curve. Thus we have (from *Fig. 8·1*).

$$P(0 < Z < z_1) = 0\cdot46$$

$\Rightarrow \quad z_1 = 1\cdot75$ (from Normal Tables)

Hence $\quad z_1' = -z_1 = -1\cdot75$

$\therefore \quad \theta = -1\cdot75\sigma$

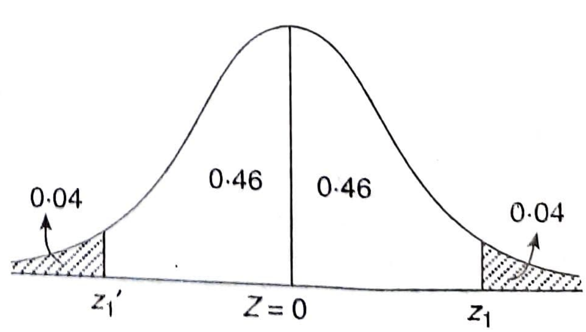


*Fig. 8·1*

This gives the level of minimum difficulty and we may shift the origin in the difficulty values $\sigma z_i$ to the point $-1\cdot75\sigma$ which implies that we shall add $1\cdot75\sigma$ to each difficulty value. If we are dealing with $\sigma$-distances from mean, *viz.*, $z_i$'s then we obtain $z_i + 1\cdot75$, ($i = 1, 2, \ldots, n$) which gives the $\sigma$-distances from the level of minimum or zero difficulty.

***Example 8·1.*** *Five problems are solved by 15%, 34%, 50%, 62% and 80% respectively of a large unselected group. If the zero point of ability in this test is taken to be at $-3\sigma$, what is the $\sigma$-value of each problem as measured from this point ? Compare the difference in difficulty between A and B with the difference in difficulty between D and E.*

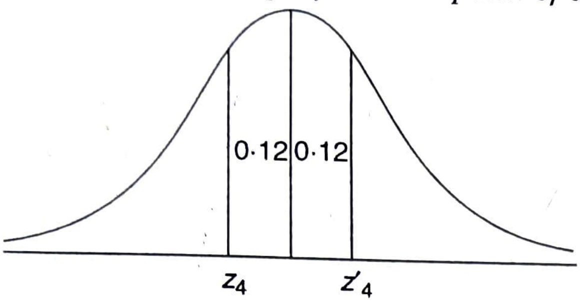


**Fig. 8·2**

**Solution.** In usual notations, it is given :

$p_1 = 0{\cdot}15$, $p_2 = 0{\cdot}34$, $p_3 = 0{\cdot}50$, $p_4 = 0{\cdot}62$ and $p_5 = 0{\cdot}80$.

The $\sigma$-values corresponding to these $p$'s, on using (8·1) are obtained as follows :

$$P(Z > z_1) = 0{\cdot}15 \quad \Rightarrow \quad P(0 < Z < z_1) = 0{\cdot}35$$

$\Rightarrow$ $z_1 = 1{\cdot}04$ (From normal probability tables)

Similarly $P(Z > z_2) = 0{\cdot}34 \quad \Rightarrow \quad P(0 < Z < z_2) = 0{\cdot}16$ *i.e.,* $z_2 = 0{\cdot}42$

$P(Z > z_3) = 0{\cdot}50 \quad \Rightarrow \quad p(0 < Z < z_3) = 0$ *i.e.,* $z_3 = 0$

$P(Z > z_4) = 0{\cdot}62$

$\Rightarrow$ $P(0 < Z < z_4') = 0{\cdot}12$ (By symmetry) [See *Fig. 8·2*]

$\Rightarrow$ $z_4' = 0{\cdot}31$ (from normal tables)

$\therefore$ $z_4 = -z_4' = -0{\cdot}31$

Also $P(Z > z_5) = 0{\cdot}80$

$\Rightarrow P(0 < Z < z_5') = 0{\cdot}30$ (From symmetry)

$\Rightarrow$ $z_5' = 0{\cdot}84$ (From normal tales)

$\therefore$ $z_5 = -z_5' = -0{\cdot}84$ (See *Fig. 8·3*)

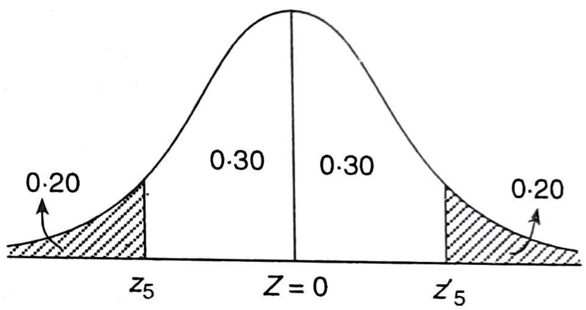


**Fig. 8·3**

The required $\sigma$-values can now be obtained as given in the Table 8·1.

TABLE 8·1

| *Problem* | *A* | *B* | *C* | *D* | *E* |
|---|---|---|---|---|---|
| $\sigma$-distances from mean | 1·04 | 0·42 | 0 | −0·31 | −0·84 |
| $\sigma$-distances from arbitrary zero = −3 | 4·04 | 3·42 | 3 | 2·69 | 2·16 |

The difference in difficulty between *A* and *B* $= 1{\cdot}04 - 0{\cdot}42 = 0{\cdot}62$

Difference in difficulty between *D* and *E* $= -0{\cdot}31 + 0{\cdot}84 = 0{\cdot}53.$

Thus
$$\frac{d_{A-B}}{d_{D-E}} = \frac{0{\cdot}62}{0{\cdot}53} \simeq 1{\cdot}2$$

$\Rightarrow$ The difficulty of *A* relative to *B* is 1·2 times greater than the difficulty of *D* relative to *E*.

***Example 8·2.*** *Given a test question solved by 10% of a large unselected group, a second question solved by 20% of the same group and a third question solved by 30%, determine the relative difficulty of the questions assuming the capacity measured by the test questions to be distributed normally. Given that, with*

$$f(a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-x^2/2}\, dx, \qquad \ldots (*)$$

$f(0{\cdot}52) = 0{\cdot}7$, $f(0{\cdot}84) = 0{\cdot}8$ *and* $f(1{\cdot}28) = 0{\cdot}9$

**Solution.** From (*), we see that $f(a)$ is nothing but the area to the left of the ordinate at $Z = a$, where $Z \sim N(0, 1)$. In other words :

$$f(a) = P(Z \le a) = 1 - P(Z > a) \quad \ldots (**)$$

The σ-values are obtained as follows :

$$P(Z > z_1) = 0{\cdot}10,\ P(Z > z_2) = 0{\cdot}20,\ P(Z > z_3) = 0{\cdot}30$$

On using (**), we get $f(z_1) = 0{\cdot}90$, $f(z_2) = 0{\cdot}80$ and $f(z_3) = 0{\cdot}70$

Hence on using (*), we obtain $z_1 = 1{\cdot}28$, $z_2 = 0{\cdot}84$ and $z_3 = 0{\cdot}52$

The results may be summarised in the *Table 8·2*

TABLE 8·2

| *Problem* | *Passed by* | *Difficulty values* | *Differences* |
|---|---|---|---|
| 1 | 10% | $1{\cdot}28\sigma = d_1$ | |
| | | | $d_1 - d_2 = 0{\cdot}44\sigma$ |
| 2 | 20% | $0{\cdot}84\sigma = d_2$ | |
| | | | $d_2 - d_3 = 0{\cdot}32\sigma$ |
| 3 | 30% | $0{\cdot}52\sigma = d_3$ | |

If we consider the percentages only it appears that problem 2 is as difficult from problem 3 as problem 1 is from problem 2. But the difficulty scale shows that the differences in difficulty between problems 2 and 3 is 0·32 σ which is roughly 3/4th of the difference in difficulty between questions 1 and 2.

**Example 8·3.** *Given three test items 1, 2 and 3 are passed by 50%, 40% and 30% respectively of a large group, on the assumption of normality of the distribution, what percentage of this group must pass test item 4 in order for this to be as much more difficult than 3, as 2 is more difficult than 1 ?*

**Solution.** Proceeding exactly similarly as in *Example 8·2*, we shall get

TABLE 8·3

| *Test item* | *Passed by* | *Difficulty value* |
|---|---|---|
| 1 | 50% | $0{\cdot}00\sigma = d_1$ |
| 2 | 40% | $0{\cdot}25\sigma = d_2$ |
| 3 | 30% | $0{\cdot}52\sigma = d_3$ |
| 4 | | $d_4$ (say) |

We are given that $d_1 - d_2 = d_3 - d_4$

$\Rightarrow \quad 0{\cdot}52\sigma - d_4 = 0{\cdot}25\sigma$

$\Rightarrow \quad d_4 = 0{\cdot}77\sigma$

Hence, the difficulty value of item 4 should be 0·77σ and consequently its σ-value is 0·77. Therefore, the percentage $p$ of individuals passing item 4 is given by :

$$p = \frac{1}{\sqrt{2\pi}} \int_{0{\cdot}77}^{\infty} e^{-t^2/2}\, dt; \quad P(Z \ge 0{\cdot}77);\ Z \sim N(0, 1)$$

$$= 0{\cdot}5 - P(0 \le Z \le 0{\cdot}77) = 0{\cdot}5 - 0{\cdot}2794 = 0{\cdot}2206 \text{ (from Normal Probability Tables)}$$

**Example 8·4.** *Four items are to be constructed so that they are equally spaced on the difficulty scale. If the easiest item is passed by 80% of the group and the most difficult item by 20%, find approximately the percentage of individuals in the group passing the other two items.*

***Solution.*** Let the difficulty values of the items 1, 2, 3 and 4 be $\sigma z_1$, $\sigma z_2$, $\sigma z_3$ and $\sigma z_4$ respectively. Then we are given :

$$P(Z > z_1) = 0{\cdot}80$$
$$\Rightarrow \quad P(Z < z_1) = 0{\cdot}20 \qquad [Fig.\ 8{\cdot}4]$$
$$\text{and} \quad P(Z > z_4) = 0{\cdot}20$$
$$\Rightarrow \quad P(0 < Z < z_4) = 0{\cdot}3 \qquad [Fig.\ 8{\cdot}4]$$
$$\Rightarrow \quad z_4 = 0{\cdot}84$$

[From Normal Probability Tables.]

Also, since $P(Z > z_4) = P(Z < z_1) = 0{\cdot}20$, [*Fig. 8·4*]

It is obvious (by symmetry) that :

$$z_1 = -z_4 = -0{\cdot}84$$

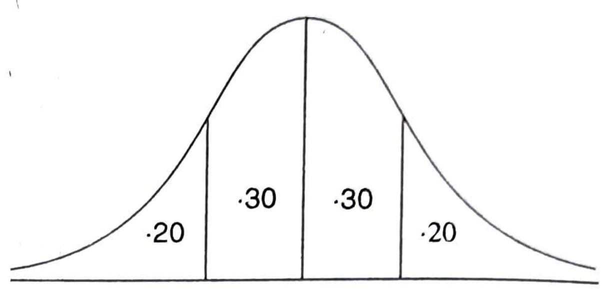


***Fig. 8·4***

Further, since the four items are equally spaced on the difficulty scale, we have

$$\sigma(z_4 - z_3) = \sigma(z_3 - z_2) = \sigma(z_2 - z_1) \quad \Rightarrow \quad z_4 - z_3 = z_3 - z_2 = z_2 - z_1$$

*i.e.*, the range between $z_1 = -0{\cdot}84$ and $z_4 = 0{\cdot}84$ is to be equally divided into three parts so that

$$z_4 - z_3 = z_3 - z_2 = z_2 - z_1 = \frac{\text{Range}}{3} = \frac{0{\cdot}84 - (-0{\cdot}84)}{3} = 0{\cdot}56$$

$$\therefore \quad z_3 = z_4 - 0{\cdot}56 = 0{\cdot}84 - 0{\cdot}56 = 0{\cdot}28$$
$$z_2 = z_1 + 0{\cdot}56 = -0{\cdot}84 + 0{\cdot}56 = -0{\cdot}28$$

We have for find : $P(Z > z_2)$ and $P(Z > z_3)$.

$$P(Z > z_3) = P(Z > 0{\cdot}28)$$
$$= 0{\cdot}5 - P(0 < Z < 0{\cdot}28) \qquad [Fig.\ 8{\cdot}5]$$
$$= 0{\cdot}50 - 0{\cdot}11 = 0{\cdot}39$$

[From Normal Probability Tables)

$\therefore$ 39% of the individuals in the group passed the item 3.

$$P(Z > z_2) = P(Z > -0{\cdot}28)$$
$$= 0{\cdot}50 + P(0 < Z < 0{\cdot}28) \text{ (By symmetry)}$$
$$= 0{\cdot}50 + 0{\cdot}11 = 0{\cdot}61$$

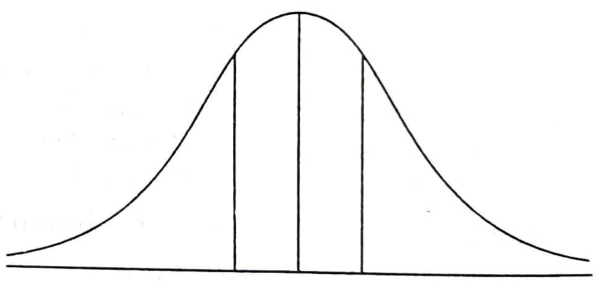


***Fig. 8·5***

$\therefore$ 61% of the individuals in the group passed the item 2.

**8·2·2. Scaling of Scores on a Test.** Suppose a number of candidates are given five different tests, say, in English, Statistics, Physics, Psychology and Scholastic aptitude. The usual system of judging the ability of an individual consists in adding the raw scores of each individual in the five tests to get his grand total or composite score and ranking them on the basis of the grand totals; an individual with the highest total securing the first position, and so on. The question arises : Are we justified in making comparisons on the basis of the sums of raw scores ? The answer is 'No', since the same raw scores $x$(say) in different tests, *e.g.*, English and Statistics may require different degrees of ability and hence may not be equivalent and therefore cannot be compared meaningfully. In order to make valid comparisons between the raw scores, we need a common scale which is obtained under some assumption regarding the distribution of the trait being measured by the test. The *standard scores and T-scores* furnish such common scale. These scores are used to combine and compare scores originally expressed in different units and are greatly used in aptitude and achievement tests.

(*a*) ***Z* (*or* σ) Scores.** Deviations from the mean expressed in terms of the standard deviation σ are called *σ-scores* or *Z-scores* or *reduced scores*. For example, if the distribution of raw scores ($X$) in a test has a mean μ and a *s.d.* σ, *i.e.*, if $E(X) = \mu$ and $\sigma_X = \sigma$, then σ-score or $Z$-score corresponding to the raw score $X$ is given by :

$$Z = \frac{X - \mu}{\sigma} \qquad \ldots (8{\cdot}2)$$

We have

$$E(Z) = E\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma}[E(X) - \mu] = 0$$

and

$$\text{Var}(Z) = \text{Var}\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma^2}\text{Var}(X - \mu) \qquad [\because \text{Var}(ax) = a^2 \text{Var}(X)]$$

$$= \frac{1}{\sigma^2} \cdot \text{Var}(X) \qquad [\because \text{Variance is independent of change of origin}]$$

$$= 1$$

Hence, *the mean of a set of σ scores is always zero and its standard deviation is unity.* Accordingly a *σ-scale* is one that has a mean of zero and a standard deviation of 1. In the construction of σ-scale, the assumption is that the distributions of the trait under considerations differ only in mean and s.d. Although theoretically σ-scores will do for making valid comparisons between raw scores, from practical point of view they are less convenient to use than others due to following shortcomings :

(*i*) Since, in general, about 50% of raw scores will lie below mean μ, approximately half of the σ-scores will be negative in sign.

(*ii*) Another disadvantage is the very large unit *viz.*, one standard deviation, making σ-scores small decimal fractions which are somewhat awkward to deal with in computation.

Both these objections can be overcome by adding to σ-scores a constant μ′ (say) so that all of them become positive and multiplying them by another constant σ′ (say) preferably by 10, to make the unit smaller and range in total units greater. This amounts to transforming the σ-scores to a new scale with mean μ′ and *s.d.* σ′.

(*b*) ***Standard Scores.*** The σ-scores transformed to the new mean μ′ and *s.d.* σ′ are called *standard scores*. Thus the standard score $X'$ (with mean μ′ and *s.d.* σ′) corresponding to the raw score $X$ with mean μ and s.d. σ is given by the relation :

$$\frac{X' - \mu'}{\sigma'} = \frac{X - \mu}{\sigma} \Rightarrow X' = \mu' + \sigma'\left(\frac{X-\mu}{\sigma}\right), \text{ i.e., } X = \mu' + \sigma' Z, \qquad \ldots (8{\cdot}3)$$

where $Z = \dfrac{X - \mu}{\sigma}$ is the σ-score corresponding to $X$.

A convenient form of (8·3) for practical purposes is as follows :

$$X' = (\sigma'/\sigma)X - [(\sigma'/\sigma)\mu - \mu'] \qquad \ldots (8{\cdot}3a)$$

All the quantities on the right-hand side are given. Hence, the relation for the conversion of raw score $X$ to standard score $X'$ is a straight line given in (8·3*a*).

***Remarks*** **1.** In order to facilitate computational arithmetic, the new mean μ′ is usually taken as some round figure such as 50, 100, 200, etc. and the new s.d. σ′ as 10, 20, 50, etc.

**2.** σ-scores or standard scores are obtained on the assumption that the actual means, dispersions and form of the distribution are same of all the trials under consideration. We should beware of these limitations and draw our conclusions with some reservation in line with these limitations.

**3.** It has already been pointed out that the raw scores obtained by the same individual in different tests cannot be compared directly due to differences in test units. However σ-scores or standard scores provide a method for comparisons provided the distributions of the raw scores are of the same form. Fortunately, the distribution of most of raw scores is approximately normal.

***Example 8·5.*** *(a) The fifth grade norms for a reading examination are Mean = 60, s.d. = 10, for an arithmetic examination Mean = 26, s.d. = 4. Ram scores 55 on the reading test and 24 on the arithmetic test. Compute his σ-scores. In which test is he better ?*

*(b) Compare his standard scores in a distribution with mean 100 and s.d. 20.*

***Solution.*** (a) Using (8·2), *Ram's* σ-score :

In Reading $= \dfrac{55-60}{10} = -0.5 = z_1$ (say) ; and in Arithmetic $= \dfrac{24-26}{4} = -0.5 = z_2$ (say)

Since Ram's σ-scores are same in reading and arithmetic tests, he is equally good in both.

(b) On using (8·3), *Ram's Standard Score* :

For Reading $= \mu' + \sigma' z_1 = 100 + 20(-0.5) = 90$ ; For Arithemetic $= \mu' + \sigma' z_2 = 90 \quad (\because z_1 = z_2)$

***Example 8·6.*** *(a) A test is administered on 400 pupils. It gave mean 60 and standard deviation 12. Complete the following table of equivalent raw scores.*

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| *Raw Score* | : | 84 | 78 | 72 | 66 | 60 | 54 | 48 | 42 | 36 |
| *σ-Score* | : | — | — | 1 | — | 0 | — | — | — | — |
| *Standard Score* | : | — | — | — | — | — | 45 | — | — | — |

*(b) Convert the ten scores 1, 2, ..., 10 into standard scores with mean 50 and standard deviation 10.*

***Solution.*** (a) Let $X$ denote the raw score. Then we are given : $E(X) = \mu = 60$ , $\sigma_X = \sigma = 12$

σ-scores ($Z$) and standard scores ($X'$) are given by :

$$Z = \frac{X-\mu}{\sigma} = \frac{X-60}{12} \quad \text{and} \quad X' = 50 + 10Z$$

TABLE 8·2 : COMPUTATION OF σ-SCORES AND STANDARD SCORES

| Raw Score ($X$) | 84 | 78 | 72 | 66 | 60 | 54 | 48 | 42 | 36 |
|---|---|---|---|---|---|---|---|---|---|
| σ-Score ($Z$) | 2 | 1·5 | 1 | 0·5 | 0 | –0·5 | –1 | –1·5 | –2 |
| Standard Score ($X'$) | 70 | 65 | 60 | 55 | 50 | 45 | 40 | 35 | 30 |

(b) The mean and variance of the first $n$ natural numbers is given by :

$$\mu = \frac{n+1}{2}, \quad \sigma^2 = \frac{n^2-1}{12}$$

Hence, the mean and variance of the first 10 natural numbers is given by :

$$\mu = \frac{10+1}{2} = \frac{11}{2} = 5.5 \quad \text{and}$$

$$\sigma^2 = \frac{100-1}{12} = 8.25 \quad \Rightarrow \quad \sigma = 2.87$$

The standard scores are obtained in Table 8·5.

TABLE 8·5 : COMPUTATION OF STANDARD SCORES

| *Raw Scores* ($X$) | $X-\mu$ | $Z = \dfrac{X-\mu}{\sigma}$ | *Standard Score* $= 50 + 10Z$ |
|---|---|---|---|
| 1 | –4·5 | –1·57 | 34·3 |
| 2 | –3·5 | –1·25 | 37·5 |
| 3 | –2·5 | –0·87 | 41·3 |
| 4 | –1·5 | –0·52 | 44·8 |
| 5 | –0·5 | –0·17 | 48·3 |
| 6 | 0·5 | 0·17 | 51·7 |
| 7 | 1·5 | 0·52 | 55·2 |
| 8 | 2·5 | 0·87 | 58·7 |
| 9 | 3·5 | 1·25 | 62·5 |
| 10 | 4·5 | 1·57 | 65·7 |

***Example 8·7.*** *A number of students were examined in a subject by three examiners $E_1$, $E_2$ and $E_3$ independently. The standards of marking of the examiners are reflected in the percentage frequency distribution of scores given in adjoining Table 8·6.*

TABLE 8·6

| *Marks* | *Percentage frequency distribution of* $E_1$ | $E_2$ | $E_3$ |
|---|---|---|---|
| 0—10 | 5 | 10 | 5 |
| 10—30 | 15 | 20 | 25 |
| 30—50 | 50 | 60 | 50 |
| 50—70 | 24 | 8 | 10 |
| 70—90 | 5 | 2 | 8 |
| 90—100 | 1 | — | 2 |

*Determine the relative ranks of the three students A, B and C who have scored the marks with the three examiners $E_1$, $E_2$ and $E_3$ as given in adjoining Table 8·6(A).*

TABLE 8·6(A)

| *Student* | *Marks given by examiner* $E_1$ | $E_2$ | $E_3$ |
|---|---|---|---|
| A | 25 | 62 | 73 |
| B | 48 | 51 | 35 |
| C | 78 | 25 | 50 |

**Solution.** We first calculate mean and standard deviation separately for the distribution of marks by the examiners $E_1$, $E_2$ and $E_3$.

Mean of the distribution of marks by $E_1$ is given by $\bar{x}_1 = \frac{1}{N} \Sigma xf_1$.

In the same way we can calculate the means $\bar{x}_2$ and $\bar{x}_3$ of the distribution of marks by $E_2$ and $E_3$ respectively.

Standard deviation of the distribution of marks by examiner $E_1$ is given by :

$$\sigma_1^2 = \left[\frac{1}{N} \Sigma f_1 x^2 - \bar{x}_1^2\right], N = \Sigma f_1.$$

Similarly, we can obtain $\sigma_2^2$ and $\sigma_3^2$ for the distribution of marks by examiners $E_2$ and $E_3$ respectively. Next we calculate the σ-scores by using the formula :

$$\sigma\text{-score} = \frac{X - \bar{x}}{\sigma}$$

and then student getting the highest aggregate of σ-scores by the examiners $E_1$, $E_2$ and $E_3$ is given first rank and so on.

TABLE 8·7 : CALCUALTIONS FOR MEAN AND S.D.

| *Marks* | $x$ | $f_1$ | $f_2$ | $f_3$ | $f_1x$ | $f_2x$ | $f_3x$ | $f_1x^2$ | $f_2x^2$ | $f_3x^2$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 0—10 | 5 | 5 | 10 | 5 | 25 | 50 | 25 | 125 | 250 | 125 |
| 10—30 | 20 | 15 | 20 | 25 | 300 | 400 | 500 | 6,000 | 8,000 | 10,000 |
| 30—50 | 40 | 50 | 60 | 50 | 2,000 | 2,400 | 2,000 | 80,000 | 96,000 | 80,000 |
| 50—70 | 60 | 24 | 8 | 10 | 1,440 | 480 | 600 | 86,400 | 28,800 | 36,000 |
| 70—90 | 80 | 5 | 2 | 8 | 400 | 160 | 640 | 32,000 | 12,800 | 51,200 |
| 90—100 | 95 | 1 | 0 | 2 | 95 | 0 | 190 | 9,025 | 0 | 18,050 |
| Total | | 100 | 100 | 100 | 4,260 | 3,490 | 3,955 | 2,13,550 | 1,45,850 | 1,95,375 |

$$\bar{x}_1 = \frac{4{,}260}{100} = 42·6\ ;\ \bar{x}_2 = \frac{3{,}490}{100} = 34·9\ ;\ \bar{x}_3 = \frac{3{,}955}{100} = 39·55$$

$$\sigma_1^2 = \left[\left(\frac{2{,}13{,}550}{100}\right) - \left(\frac{4{,}260}{100}\right)^2\right] = 2{,}135·50 - 1{,}814·76 = 320·74 \quad \Rightarrow \quad \sigma_1 = 17·9$$

$$\sigma_2^2 = \left[\frac{1{,}45{,}850}{100} - \left(\frac{3{,}490}{100}\right)^2\right] = 1{,}458·50 - 1{,}218·01 = 240·49 \quad \Rightarrow \quad \sigma_2 = 15·5$$

$$\sigma_3^2 = \left[\frac{1{,}95{,}375}{100} - \left(\frac{3{,}955}{100}\right)^2\right] = 389·55 \quad \Rightarrow \quad \sigma_3 = 19·6$$

TABLE 8·8

| | Examiner | | |
|---|---|---|---|
| | $E_1$ | $E_2$ | $E_3$ |
| σ-Score for *A* | $\frac{25 - 42·6}{17·9} = -·983$ | $\frac{62 - 34·9}{15·5} = 1·748$ | $\frac{73 - 39·55}{19·6} = 1·706$ |
| σ-Score for *B* | $\frac{48 - 42·6}{17·9} = ·302$ | $\frac{51 - 34·9}{15·5} = 1·038$ | $\frac{35 - 39·55}{19·6} = -0·232$ |
| σ-Score for *C* | $\frac{78 - 42·6}{17·9} = 1·983$ | $\frac{25 - 34·9}{15·5} = -·639$ | $\frac{50 - 39·55}{19·6} = 0·533$ |

The combined σ-scores for the students *A*, *B* and *C* as given by the three examiners and their relative ranks are obtained in *Table 8·8*. The student with the highest σ-score is given the rank 1 and the student with the lowest σ-score is given the rank 3.

TABLE 8·6 : COMPUTATION OF COMBINED σ-SCORE AND RANKS

| *Student* | *Combined σ-score* | *Rank* |
|---|---|---|
| A | $-·983 + 1·748 + 1·706 = 2·471$ | I |
| B | $·302 + 1·038 - ·232 = 1·108$ | III |
| C | $1·983 - ·639 + ·533 = 1·877$ | II |

**(c) Normalised Scores.** Here the scaling procedure is based on the *assumption that the trait under consideration* ($X$) is normally distributed with mean $\mu_x$ and *S.D.* $\sigma_x$ and the raw scores are converted into a system of normalised scores by transforming them into the equivalent points of a normal distribution.

Let $p$ be the proportion of individuals getting scores below a score $x$. Then

$$p = P(X \le x) = P\left[Z \le \frac{x - \mu_x}{\sigma_x} = \xi \text{ (say)}\right], \text{ where } Z \sim N(0, 1)$$

The number $\xi$ given by $P(Z \le \xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\xi} e^{-u^2/2}\, du = p \quad \Rightarrow \quad \Phi(\xi) = p$ ... (8·4)

where $\Phi(.)$ is the distribution function of a standard normal variate, is called the *normalised score* corresponding to $x$. Like σ-scores, normalised scores have their mean zero and variance unity and in all probability they almost lie in the range $-3$ to 3. Equivalent normalised scores represent the same level of the talent or achievement.

For practical convenience, normalised scores are transformed to new scale with mean $\mu$ (say), and standard deviation $\sigma$ (say), by the relation :

$$\frac{\eta - \mu}{\sigma} = \xi \quad \Rightarrow \quad \eta = \mu + \sigma\xi, \qquad \text{... (8·5)}$$

where $\mu$ and $\sigma$ are pre-assigned. $\eta$'s are called *normalised standard scores*.

**(d) T-Scores.** In particular if we take $\mu = 50$ and $\sigma = 10$ in (8·5), we get *T*-scores. Thus *T*-scores are normalised standard scores converted into a distribution with mean 50 and standard deviation 10 and are given by :

$$T = 50 + 10\xi \qquad \ldots (8\cdot 6)$$

*T*-scores were devised by McCall William A. and are named so in memory of the psychologists Terms and Thorndyke.

***Calculation of T-scores for a given Frequency Distribution.*** The procedure for obtaining *T*-scores for any given frequency distribution is outlined in the following steps :

(*i*) Arrange the test scores in *descending order* of magnitude (as is customary with most psychological educational data).

(*ii*) Obtain the cumulative frequency (*c.f.*) starting from the bottom of the distribution.

(*iii*) Obtain the *c.f.* below the mid-value of each class interval under the assumption that the frequencies are uniformly distributed over the class intervals, *i.e.*, find $\left[c.f.) - \frac{1}{2}f\right]$.

(*iv*) Express these cumulative frequencies as percentages or proportions '*p*' of the total frequency *N*.

(*v*) Obtain the normalised scores ξ given by :

$$P(Z \le \xi) = \int_{-\infty}^{\xi} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du = p \ ; \ Z \sim N(0, 1)$$

(*vi*) Finally, *T*-scores are obtained from normalised scores ξ's on using 8·6.

Tables for obtaining *T*-scores directly from the percentages or proportion '*p*' exist (*c.f.* Table *G* in "Statistics in Psychology and Education" by Henry E. Garret).

The steps outlined above can be elegantly displayed in the following *Table 8·10* :

TABLE 8·10 : COMPUTATION OF T-SCORES

| *Class Interval* | *f* | *c.f.* | *c.f. below mid-value of each class* | *Col. (4) as proportion 'p' of N* | | |
|---|---|---|---|---|---|---|
| (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| $x_1 - x_2$ | $f_1$ | $N$ | $N - \frac{1}{2}f_1 = A_1$ (say) | $A_1/N$ | | |
| $x_2 - x_3$ | $f_2$ | $\sum_{i=2}^{n} f_i$ | $\sum_{i=2}^{n} f_i - \frac{1}{2}f_2 = A_2$ | $A_2/N$ | $p = \int_{-\infty}^{\xi} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du$ | $T = 50 + 10\xi$ |
| $x_3 - x_4$ | $f_3$ | $\sum_{i=3}^{n} f_i$ | $\sum_{i=3}^{n} f_i - \frac{1}{2}f_3 = A_3$ | $A_3/N$ | | |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | | |
| $x_{n-1} - x_n$ | $f_{n-1}$ | $f_n + f_{n-1}$ | $f_n + \frac{1}{2}f_{n-1} = A_{n-1}$ | $A_{n-1}/N$ | | |
| $x_n - x_{n+1}$ | $f_n$ | $f_n$ | $\frac{1}{2}f_n = A_n$ | $A_n/N$ | | |

***Example 8·8.*** *Find the T-scores corresponding to the test scores X for the following frequency distribution :*

| $x$ : | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| $y$ : | 5 | 10 | 20 | 5 | 4 | 4 | 2 |

***Solution.*** For the computation of the $T$-scores, we complete the Table on Page 8·11 (as explained in *Table 8·10*)

TABLE 8·11 : CALCULATIONS FOR T-SCORES

| $x$ | $f$ | c.f. | *c.f. below mid-value of each score* | $p = \frac{\text{col. (4)}}{N}$ | $\xi$ $P(Z \le \xi) = p$ | $T = 50 + 10\xi$ |
|---|---|---|---|---|---|---|
| (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| 7 | 2 | 50 | $50 - 1 = 49$ | 0·98 | $2·55 = \xi_1$ | 75 |
| 6 | 4 | 48 | $48 - 2 = 46$ | 0·92 | $1·45 = \xi_2$ | 64 |
| 5 | 4 | 44 | $44 - 2 = 42$ | 0·84 | $0·995 = \xi_3$ | 60 |
| 4 | 5 | 40 | $40 - 2·5 = 37·5$ | 0·75 | $0·675 = \xi_4$ | 57 |
| 3 | 20 | 35 | $35 - 10 = 25$ | 0.50 | $0 = \xi_5$ | 50 |
| 2 | 10 | 15 | $15 - 5 = 10$ | 0·20 | $-0·840 = \xi_6$ | 42 |
| 1 | 5 | 5 | $\frac{1}{2}f = (5/2) = 2·5$ | 0·05 | $-1·645 = \xi_7$ | 34 |

Normalised scores $\xi_i$'s are obtained as follows :

If $Z \sim N(0, 1)$, then from normal probability tables, we get

$$P(Z \le \xi_1) = 0·98 \quad \Rightarrow \quad \xi_1 = 2·55 \quad ; \quad P(Z \le \xi_2) = 0·92 \quad \Rightarrow \quad \xi_2 = 1·45, \text{ and so on.}$$

***Important Remark.*** Here although we are dealing with a discrete distribution, we convert the *c.f.* in Col. (3) to *c.f.* in Col. (4) since *each score is regarded as an interval and not a point on the scale.* For example, the score 6 may be regarded as the mid-point of the class interval 5·5 to 6·5.

***Uses of T-scores.*** The $T$-scale and $T$-scores overcome the objection raised against standard scores; they have a convenient unit and cover a wide range of talent. $T$-scores from different tests are readily comparable since they refer to a standard scale with mean $\mu = 50$ and s.d. $\sigma = 10$ based upon normal probability curve. The underlying assumption in the construction of the $T$-scale is the *normality of the trait being considered,* an assumption which is quite reasonable and feasible since the parent distribution of most mental abilities is more or less normal.

***Comparison of T-scores and Standard Scores.*** Although both the standard scores and $T$-scores are used for the comparability of raw scores of an individual in different tests, they should not be confused with each other since the basic assumptions underlying these two measures are entirely different. In the construction of standard scores, the basic assumption is that *the standard scores have the same form of distribution as the raw scores* and thus they are merely original scores expressed in $\sigma$-units. This type of conversion is justified when the transformation is linear as, for example, in the conversion of inches to centimeters or kilograms to pounds. On the other hand, $T$-scores are based upon the fundamental assumption that "*the parent distribution of the trait being considered is normal*". Thus, *w.r.t.* the original scores, they represent equivalent scores in a normal distribution. It is, therefore, quite clear that standard scores and $T$-scores are not interchangeable. However, if the distribution of raw scores is strictly normal, the standard scores correspond exactly to $T$-scores.

*(e) **Percentile Scores.*** Percentile score of an individual is the percentage (to the nearest integer) of the frequency lying below his raw score $x$ (say), assuming that score is a continuous variable. For any frequency distribution, the method of computing the percentile scores corresponding to any class interval (or mid-value of the class) has incidentally been explained in column (5) of the *Table 8·10* (on *page 8·11*) while calculating the $T$-scores. The percentile scores for different classes are thus given by :

$$P = 100 \times \text{cumulative frequency below mid-value of the class}/N$$
$$= 100 \times \text{column (5) of } \textit{Table 8·10} \text{ on } \textit{page 8·11}.$$

***Advantages and Disadvantages of Percentile Scores.*** The most practical advantage of percentile scores is the ease with which they are understood and calculated. If an individual is subject to several tests then his corresponding percentile scores give his relative achievement in different tests. Percentile scores may be combined to give his final test score.

The basic assumption in the construction of the percentile scale is that the distribution of the trait under consideration is rectangular or uniform such that the percentile differences are equal throughout the scale. By a scale with equal percentile units we mean a scale in which the difference between the percentile scores of 20 and 30 (say) is the same as the difference between percentile scores of 60 and 70. From practical point of view the underlying assumption for percentile scaling is wrong since most of the psychological and educational data follow approximately the normal distribution and the distribution of raw scores is rarely, if ever, rectangular in form. Consequently, in practice, the percentile scores cannot be regarded as representing equal increments of achievement and the percentile scale does not progress by equal units. Hence percentile scale should be used with reservations, if any, keeping in mind its limitations.

**8·2·3. Scaling of Rankings in Terms of Normal Probability Curve.** Suppose $N$ individuals are ranked by a judge in order of merit of a particular trait, say, 'social responsibility' Under the assumption that there is no tie, *i.e.*, no two individuals are bracketed together the frequency of each rank is 1 and the corresponding frequency distribution (in ascending order of ranks) can be written as :

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Individual | : | $x_1$ | $x_2$ | $x_3$ | ... $x_R$ | ... $x_N$ |
| Rank | : | 1 | 2 | 3 | ... $R$ | ... $N$ |

where $x_i$'s take values from 1 to $N$.

Proceeding exactly as in *Table 8·10* the conversion *Table 8·12* for obtaining the percentile scores corresponding to different values of $R$ may be summarised as follows in *Tale 8·12* : [under the assumption of the normality of the trait concerned which indirectly means that a rank $R$ of an individual represents the interval from $\left[\left(R - \frac{1}{2}\right) \text{ to } \left(R + \frac{1}{2}\right)\right]$.

TABLE 8·12 : COMPUTATIONS FOR PERCENTILE SCORES

| *Rank* | *f* | *c.f.* | *c.f. below value of R* | *Cumulative proportion below R* |
|---|---|---|---|---|
| 1 | 1 | $N$ | $N - 1 + 0·5$ | ⋮ |
| 2 | 1 | $N - 1$ | $N - 2 + 0·5$ | ⋮ |
| 3 | 1 | $N - 2$ | $N - 3 + 0·5$ | ⋮ |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| $R - 1$ | 1 | $N - R + 2$ | $N - r + 1·5$ | ⋮ |

| | | | | |
|---|---|---|---|---|
| $R$ | 1 | $N-R+1$ | $N-R+0{\cdot}5$ | $\frac{N-R+0{\cdot}5}{N}=1-\frac{R-0{\cdot}5}{N}=p_r$ |
| $R+1$ | 1 | $N-R$ | $\vdots$ | $\vdots$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $N-1$ | 1 | 2 | 1·5 | $1{\cdot}5/N$ |
| $N$ | 1 | 1 | 0·5 | $0{\cdot}5/N$ |

Percentile score corresponding to rank $R$ of an individual among $N$ individuals is given by :

$$100\, p_R = 100\left(1-\frac{R-0{\cdot}5}{N}\right) \qquad \ldots(8{\cdot}7)$$

This formula enables us to convert any set of ranks into scores if we are justified in assuming normality in the trait for which ranks are given. This method is specially useful when we are dealing with qualitative characteristics which cannot be measured quantitatively, *e.g.*, attributes like beauty, honesty, personality, athletic ability. For example, suppose $N$ individuals are ranked by three different judges $A, B, C$, (say) in order of merit *w.r.t.* certain characteristic, say, beauty. The problem is "*how to combine these ranks by three judges to get the final rankings ?*" The solution consists in obtaining percentile scores by using (8·7) corresponding to the ranks of $N$ individuals for each of the judges $A$, $B$ and $C$. These transmuted ranks may be combined and averaged for giving them final ranks.

***Remarks* 1.** The scores corresponding to the values of $p_R$ on a scale of 10 units are given by :

$$10p_R = 10\left(1-\frac{R-0{\cdot}5}{N}\right) \qquad \ldots(8{\cdot}7a)$$

**2.** The scale values corresponding to $p_R$'s are obtained by finding normalised scores $\xi$'s corresponding to $p_R$ by the relation :

$$p_R = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\xi} e^{-t^2/2}\, dt = \Phi(\xi) \qquad \ldots(8{\cdot}8)$$

where $\Phi(.)$ is the distribution function of a standard normal variate. Since $T$-scores are given by :

$$T = 50 + 10\xi \quad \Rightarrow \quad \xi = \frac{T-50}{10},$$

from (8·8), we get

$$\Phi\left(\frac{T-50}{10}\right) = p_R \qquad \ldots(8{\cdot}9)$$

This equation provides us a method of obtaining $T$-scores from the percentile scores or $p_R$.

**8·2·4. Scaling of Ratings in Terms of Normal Probability Curve.** Let us suppose that $N$ individuals have been rated by different judges *w.r.t.* some trait say, honesty and the corresponding frequency distributions of the ratings of the judges are given or known. The problem is :

'Can we assign weights or numerical scores to these ratings so as to make them comparable from judge to judge ? The answer is 'yes' provided we are justified in assuming (*i*) the normality of the trait being considered, and (*ii*) different judges are equally competent.

Let us suppose that the distribution of the trait, say $X$, is $N(0, 1)$. Suppose that the individuals with trait values in the interval $(x_1 - x_2)$ are given a rating 'A' by a judge. The *scale value* (or *σ-value*) corresponding to this rating 'A' is defined to be the average trait value of all these individuals and is accordingly given by the formula (due to Likert)

$$\text{Scale value} = \frac{\int_{x_1}^{x_2} u\,\phi(u)du}{\int_{x_1}^{x_2} \phi(u)du} = \frac{\int_{x_1}^{x_2} u\,\phi(u)du}{\Phi(x_2) - \Phi(x_1)}$$

where $\phi(.)$ and $\Phi(.)$ are respectively the *p.d.f.* and distribution function of a standard normal variate.

$$\text{Scale value} = \frac{\left| -\frac{1}{\sqrt{2\pi}} e^{-u^2/2} \right|_{x_1}^{x_2}}{\Phi(x_2) - \Phi(x_1)} = \frac{\phi(x_1) - \phi(x_3)}{\Phi(x_2) - \Phi(x_1)} \qquad \ldots (8{\cdot}10)$$

For a given distribution of ratings the values on the right-hand side of (8·10) can be easily obtained by using normal probability tables for areas and ordinates. It may be pointed out that the denominator in (8·10) gives the proportion of individuals placed in the rating 'A'. The numerical score for rating 'A' is now obtained by shifting the origin in the scale value ($\sigma$-value) to $-$ 3·0 as an arbitrary origin, multiplying each $\sigma$-value so abtained by 10 and rounding them to the nearest integer.

***Example 8·9.** Letter grades A, B, C, D, E (A being the highest and E the lowest) are assigned by two teachers X and Y to the students of class for 'Honesty'. The Table 8·10 gives the distribution of the proportion of individuals in each rating.*

| → *Honesty* / ↓ *Teacher* | *A* | *B* | *C* | *D* | *E* |
|---|---|---|---|---|---|
| *X* | *0·10* | *0·15* | *0·50* | *0·20* | *0·05* |
| *Y* | *0·20* | *0·40* | *0·20* | *0·10* | *0·10* |

*Find the numerical scores corresponding to each grade and each teacher.*

**Solution.** First of all, we shall obtain the scale values ($\sigma$-scores) corresponding to these percentages. Since $A, B, C, D, E$ are the ratings in the descending order, we have for teacher $X$, on using, normal probability tables : [See *Fig. 8·6*.]

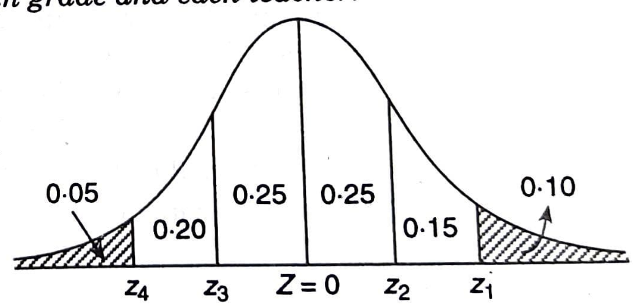


**Fig. 8·6**

$A : P(Z > z_1) = 0{\cdot}10 \quad \Rightarrow \quad z_1 = 1{\cdot}28$

$B : P(Z > z_2) = 0{\cdot}25 \quad \Rightarrow \quad z_2 = 0{\cdot}675$

$C : P(Z > z_3) = 0{\cdot}75 \quad \Rightarrow \quad z_3 = -z_3 = -\,0{\cdot}675$

$D : P(Z > z_4) = 0{\cdot}95 \quad \Rightarrow \quad z_4 = -\,1{\cdot}645$

$$\sigma\text{-Score for 'A'} = \frac{\phi(z_1) - \phi(\infty)}{\Phi(\infty) - \Phi(z_1)} = \frac{\phi(1{\cdot}28)}{0{\cdot}10} = \frac{0{\cdot}1758}{0{\cdot}10} = 1{\cdot}76$$

$$\sigma\text{-Score for 'B'} = \frac{\phi(z_2) - \phi(z_1)}{\Phi(z_1) - \Phi(z_2)} = \frac{\phi(0{\cdot}675) - \phi(1{\cdot}28)}{0{\cdot}15} = \frac{0{\cdot}3176 - 0{\cdot}1758}{0{\cdot}15} = 0{\cdot}95$$

$$\sigma\text{-Score for 'C'} = \frac{\phi(z_3) - \phi(z_2)}{\Phi(z_2) - \Phi(z_3)} = \frac{\phi(-\,0{\cdot}675) - \phi(0{\cdot}675)}{0{\cdot}50} = 0$$

$$\sigma\text{-Score for 'D'} = \frac{\phi(z_4) - \phi(z_3)}{\Phi(z_3) - \Phi(z_4)} = \frac{\phi(-\,1{\cdot}645) - \phi(-\,0{\cdot}675)}{0{\cdot}20} = \frac{0{\cdot}1031 - 0{\cdot}3176}{0{\cdot}20} = -1{\cdot}07$$

$$\sigma\text{-Score for 'E'} = \frac{\phi(-\infty) - \phi(z_4)}{\Phi(z_4) - \Phi(-\infty)} = \frac{0 - 0{\cdot}1031}{0{\cdot}05} = -2{\cdot}06$$

Proceeding similarly, we shall get the scores corresponding to ratings by teacher $Y$ as

| $A$ | $B$ | $C$ | $D$ | $E$ |
|---|---|---|---|---|
| 1·40 | 0·27 | –0·53 | –1·04 | –1·76 |

Shifting the origin to – 3·0 and multiplying by 10, we may summarise the numerical scores (rounded to two digits) as follows :

| *Teacher* | $A$ | $B$ | $C$ | $D$ | $E$ |
|---|---|---|---|---|---|
| $X$ | 48 | 40 | 30 | 19 | 9 |
| $Y$ | 44 | 33 | 25 | 20 | 12 |