A RAPID AND RELIABLE METHOD FOR ESTIMATING PERCENTILE SCORES OF EXAMINATION MARKS AND PSYCHOMETRIC TEST SCORES.

RHEA S. DAS
Indian Statistical Institute
Calcutta.

Percentile scores are one type of derived scores obtained by a non-linear transformation of the original test scores. A percentile is defined as a score value below which fall a given percent of frequencies. The popularity and value of percentile scores lie in their ease of interpretation and the information which they provide of the relative standing of individuals in a group. While the traditional 'percent marks' of the universities are also easily interpreted, they do not indicate the position of individual students relative to the group as a whole. Arbitrary standards are set to indicate first, second and third class marks, but the percentage of students in the several classes varies considerably from year to year. The standing of a student, relative to his group is not indicated by his 'percent marks', as for example '60' might refer to the 90th percentile one year and to the 80th percentile another year. In the case of aptitude and achievement tests, raw scores neither reveal comparative standing of students nor are they readily interpretable as the range of possible scores varies widely from test to test. If marks or test scores are altered in some way to indicate the position of individuals relative to the group, their practical value is enhanced for such purposes as promotion and failure, awards, counselling, vocational guidance and reports of progress to parents. Transforming the marks or scores to percentile scores is one method of accomplishing this purpose.

For the computation of percentile scores three methods are generally described in standard texts (1, 4, 5, 6): first, direct computation of the one hundred percentile points; second, computation of the percentiles for the mid-points of class intervals and arithmetic interpolation of the remaining percentiles; and third, plotting the percentiles corresponding to the tops of class intervals and drawing a smoothed curve from which the percentile scores are read off. The educator or psychologist may select one of these methods in terms of the following criteria: economy of time, reliability or repeatability, and freedom from chance fluctuations in the original score distributions. If the first method mentioned is examined in terms of these criteria, it will be found that it is reliable but not economical of time nor free from chance fluctuations. The second method is reliable and relatively free from chance fluctuations, but is still relatively time consuming. It is also interesting that the standard texts do not detail the appropriate method of interpolation which may not be familiar to most users of percentile scores. Finally, while the third method is economical of time and relatively free from chance fluctuation, it suffers from relatively lower reliability. The difficulties associated with these methods suggest that a rapid method should be developed which is reliable and relatively free from chance fluctuations. The method outlined below was developed in an attempt to satisfy these criteria.

As Adkins (1) has shown, the relationship between test score and the cumulative percentage receiving that score or less is sigmoidal or S-shaped. When the relationship is plotted with cumulative percentage as ordinate and total score as abscissa, it is the familiar ogive, or cumulative percentage curve, of statistics. This sigmoidal relationship does not permit an arithmetic transformation of the original scores to percentile scores. If a transformation could be applied to the relationship making it linear, arithmetical operations could be employed to estimate percentile scores. A transformation which reduces the sigmoidal response curve to a straight line on the basis of the normal integral has been developed and refined for problems concerning the toxicity of insecticides and fungicides (3). This transformation, known as the 'probit transformation', was first enunciated by the psycho-physicist Fechner (2), but it has not been widely applied in the analysis of psychological data in recent years. In the paragraphs below the use of the probit transformation to obtain the estimates of percentile scores is outlined.

Probits may be defined as transformed scores with a mean, K, of 5 and standard deviation, S, of 1. They are related to the original scores by the formula

$$Y = K + S\left(\frac{X - \bar{X}}{\sigma}\right) \qquad [1]$$

where

Y = probit
K = derived mean, 5
S = derived standard deviation, 1
X = any original score
$\bar{X}$ = mean of the original scores
σ = standard deviation of the original scores

This formula is the same as that given in standard texts (e.g., 1,6) to compute standard or Z scores. The relationship between probit scores and percentages has been tabulated to obtain probits from percentages (3). If the relationship is tabulated in the reverse direction, to read percentages from probits, percentile scores can be rapidly estimated from scores transformed so that their mean is 5 and their standard deviation equals 1. The detailed steps are given below, with some simplified formulae to facilitate computation:

1. For a normal distribution, equation [1] gives the maximum likelihood estimate of the probit values.

a. Compute $\overline{X}$ and σ of the original test scores, following a method outlined in standard texts (e.g., 4,5). For example, the sum of scores $(\Sigma X)^2$, the sum of the squared scores ΣX^2, and number of scores N, may be used to compute $\overline{X}$ and σ.

$$\overline{X} = \frac{\Sigma X}{N} \quad \ldots\ldots \quad \ldots\ldots \quad [2]$$

$$= \frac{1}{N}\sqrt{N\,\Sigma X^2 - (\Sigma X)^2} \quad \ldots\ldots \quad [3]$$

b. Obtain the constants A and B where

$$A = \frac{1}{\quad} \quad \ldots\ldots \quad \ldots\ldots \quad [4]$$

$$B = 5 - \frac{\overline{X}}{\sigma} \quad \ldots\ldots \quad \ldots\ldots \quad [5]$$

c. Prepare a table listing in the first column the original scores X in descending order. In the second column give the product A(X) for each score. Add B to the product A(X) for each score, and enter this sum in the third column. The third column then gives Y by the formula $Y = A(X) + B$....[6]

which is a simplified form of formula [1] (6, p. 269).

2. In order to convert the Y scores to percentiles, Table 1 has been prepared.

a. The probit or Y values, in 0.1 steps, are given in the first column of Table 1. The remaining columns give Y in 0.01 steps. The corresponding percentile values are given in the body of the table.

Table 1: Transformation of probits, or Y scores with mean of 5 and standard deviation of 1, to percentiles

Probits	0	.01	.02	.03	.04	.05	.06	.07	.08	.09
2.6	0.82	0.84	0.87	0.89	0.91	0.94	0.96	0.99	1.02	1.04
2.7	1.07	1.10	1.13	1.16	1.19	1.22	1.25	1.29	1.32	1.36
2.8	1.39	1.43	1.46	1 50	1.54	1.58	1.62	1.66	1.70	1.74
2.9	1.79	1.83	1.88	1.92	1.97	2.02	2.07	2.12	2.17	2.22
3.0	2.28	2.33	2.39	2.44	2.50	2.56	2.62	2.68	2.74	2.81
3.1	2.87	2.94	3.01	3.07	3.14	3.22	3.29	3.36	3.44	3.51
3.2	3.59	3.67	3.75	3.84	3.92	4.01	4.09	4.18	4.27	4.36
3.3	4.46	4.55	4.65	4.75	4.85	4.95	5.05	5.16	5.26	5.37
3.4	5.48	5.59	5.71	5.82	5.94	6.06	6.18	6.30	6.43	6.55
3.5	6.68	6.81	6.94	7.08	7.21	7.35	7.49	7.64	7.78	7.93
3.6	8.08	8.23	8.38	8.53	8.69	8.85	9.01	9.18	9.34	9.51
3.7	9.68	9.85	10.03	10.20	10.38	10.56	10.75	10.93	11.12	11.31
3.8	11 51	11.70	11.90	12.10	12.30	12.51	12.71	12.92	13.14	13.35
3.9	13.57	13.79	14.01	14.23	14.46	14.69	14.92	15.15	15.39	15.62
4.0	15.87	16.11	16.35	16.60	16.85	17.11	17.35	17.62	17.88	18.14
4.1	18.41	18.67	18.94	19.22	19.49	19.77	20.05	20.33	20.61	20.90
4.2	21.19	21.48	21.77	22.06	22.36	22.66	22.96	23.27	23.58	23.89
4.3	24.20	24.51	24.83	25.14	25.46	25.78	26.11	26.43	26.76	27.09
4.4	27.43	27.76	28.10	28.43	28.77	29.12	29.46	29.81	30.15	30.50
4.5	30.85	31.21	31.56	31.92	32.28	32.64	33.00	33.36	33.72	34.09
4.6	34.46	34.83	35.20	35.57	35.94	36.32	36.69	37.07	37.45	37.83
4.7	38.21	38.59	38.87	39.36	39.74	40.13	40,50	40.90	41.29	41.68
4.8	42.07	42.47	42.86	43.25	43.64	44.04	44.43	44.83	45.22	45.62
4.9	46.02	46.41	46.81	47.21	47.61	48.01	48.40	48.80	49.20	49.60

(Table 1 continued)

Probits	0	.01	.02	.03	.04	.05	.06	.07	.08	.09
5.0	50.00	50.40	50.80	51.20	51.60	51.99	52.39	52.79	53.19	53.59
5.1	53.98	54.38	54.78	55.17	55.57	55.96	56.36	56.75	57.14	57.53
5.2	57.93	58.32	58.71	59.10	59.48	59.87	60.26	60.64	61.03	61.41
5.3	61.79	62.17	62.55	62 93	63.31	63.68	64.06	64.43	64.80	65.17
5.4	65.54	65.91	66.28	66.64	67.00	67.36	67.72	68.08	68.44	68.79
5.5	69.15	69.50	69.85	70.19	70.54	70.88	71.23	71.57	71.90	72.24
5.6	72.57	72.91	73.24	73.57	73.89	74.22	74.54	74.86	75.17	75.49
5.7	75.80	76.11	76.42	76.73	77.04	77.34	77.64	77.94	78.23	78.52
5.8	78.81	79.10	79.39	79.67	79.95	80.23	80.51	80.78	81.06	81.33
5.9	81.59	81.86	82.12	82.38	82.64	82.89	83.15	83.40	83.65	83.89
6.0	84.13	84.38	84.61	84.85	85.08	85.31	85.54	85.77	85.99	86.21
6.1	86.43	86.65	86.86	87.08	87.29	87.49	87.70	87.90	88.10	88.30
6.2	88.49	88.69	88.88	89.07	89,25	89.44	89.62	89.80	89.97	90.15
6.3	90.32	90.49	90.66	90.82	90.99	91.15	91.31	91.47	91.62	91.77
6.4	91.92	92.07	92.22	92.36	92.51	92.65	92.79	92.92	93.06	93.19
6.5	93.32	93.45	93.57	93.70	93.82	93.94	94.06	94.18	94.29	94.41
6.6	94.52	94.63	94.74	94.84	95.95	95.05	95.15	95.25	95.35	95.45
6.7	95.54	95.64	95.73	95.82	95.91	95.99	96.08	96 16	96.25	96.33
6.8	96.41	96.49	96.56	96.64	96.71	96.78	96.86	96.93	96.99	97.06
6.9	97.13	97.19	97.26	97.32	97.38	97.44	97.50	97.56	97.61	97.67
7.0	97.72	97.78	97.83	97.88	97.93	97.98	98.03	98.08	98.12	98.17
7.1	98.21	98.26	98.30	98.34	98.38	98.42	98.46	98.50	98.54	98.57
7.2	98.61	98.64	98.68	98.71	98.75	98.78	98.81	98.84	98.87	98.90
7.3	98.93	98.96	98.98	99.01	99.04	99.06	99.09	99.11	99.13	99.16

Prepared by Rhea S. Das, Indian Statistical Institute, Calcutta-35, India.

b. Add a fourth column to the list prepared in Step 1c. Locate the percentiles corresponding the Y scores in the third column. Enter the percentiles in the fourth column.

In order to illustrate these steps a numerical example is given below.

Table 2: Original scores of 30 students, with sums of scores and squares.

Serial Number	Score X	Serial Number	Score X	Serial Number	Score X
1	29	11	22	21	14
2	29	12	21	22	14
3	28	13	20	23	14
4	28	14	18	24	14
5	27	15	18	25	13
6	26	16	17	26	13
7	25	17	17	27	13
8	24	18	17	28	8
9	24	19	16	29	6
10	23	20	14	30	5

$\Sigma X = 557$ $\quad \Sigma X^2 = 11669$

$(\Sigma X)^2 = 310249$ $\quad N = 30$

Table 2 gives the original scores, their sums and sums of squares.

Step 1.a. $\bar{X} = \frac{557}{30} = 18.567$

$$= \frac{1}{30}\sqrt{(30)\,(11,669)-(557)^2}=6.652$$

Step 1.b. $A = \frac{1}{6.652} = 0.150$

$$B = 5 - \frac{18.567}{6.652} = 2.209$$

Step 1.c. See Table 3. The first column lists the scores, X; the second column gives A(X); and the third column gives $Y = A(X) + B$.

Step 2.b. See Table 3. The fourth column gives the percentile scores corresponding to the Y scores in the third column.

Table 3: A computational example for estimation of percentile scores using probits

X (1)	A(X) (2)	A(X) + B = Y (3)	Percentile Estimate— (4)
29	4.35	6.559	94.06
28	4.20	6.409	92.07
27	4.05	6.259	89.62
26	3.90	6.109	86.65
25	3.75	5.959	83.15
24	3.60	5.809	79.10
23	3.45	5.659	74.54
22	3.30	5.509	69.50
21	3.15	5.359	64.06
20	3.00	5.209	58.32
18	2.70	4.909	46.41
17	2.55	4.759	40.50
16	2.40	4.609	34.83
14	2.10	4.309	24.51
13	1.95	4.159	20.05
8	1.20	3.409	5.59
6	0.90	3.109	2.94
5	0.75	2.959	2.07

A method for the estimation of percentile scores, for use with examinations and psychometric tests, has been outlined and illustrated. Unlike the customary methods of computing the percentile scores, it is based on the assumption that the original set of scores are drawn from a normal population. The resulting scores may not be, therefore, identical with percentile scores obtained by any of the 3 methods mentioned earlier, and are therefore termed percentile estimates. It is interesting, however, that the scores obtained by these three methods also differ from each other. In absolute value the percentile estimates do not differ significantly from scores computed by arithmetic methods. Empirical studies have shown that they are perfectly correlated, within limits of rounding error, with percentile scores computed by standard arithmetic methods. They may be interpreted in the same way as percentile scores, and have the same advantages of ease in interpretation and of providing information concerning the comparative standing of individuals in the group. The proposed method of computation is both rapid and reliable. In addition, the preliminary transformation of scores in terms of mean and standard deviation also reduces the effects of chance fluctuation on the final percentile estimates.

It is hoped that the method and table presented here will be of practical value in facilitating computation of percentile scores and therefore make possible their increased use in schools, colleges, vocational guidance bureaus, and other educational institutions.

REFERENCES

1. ADKINS, D. C., *Construction and analysis of achievement tests.* Washington: U. S. Government Printing Office, 1947.
2. FECHNER, G. T., *Elemente der Psychophysik.* Leipzig: Breitkopf und Hartel, 1860.
3. FINNEY, D. J., *Probit Analysis.* Cambridge: University Press, 1947.
4. GARRETT, H. E., *Statistics in psychology and education.* 3rd. Ed., New York: Longmans, Green, 1947.
5. GUILFORD, J. P., *Fundamental statistics in psychology and education.* McGraw-Hill, 1950.
6. GULLIKSEN, H., *Theory of mental tests* New York: John Wiley, 1950.