DOCUMENT RESUME

ED 046 981 | TM 000 342

AUTHOR Lindsay, Carl A.; Prichard, Mark A.
TITLE An Analytical Procedure for the Equipercentile Method of Equating Tests.
PUB DATE Feb 71
NOTE 12p; Paper presented at the Annual Meeting of the American Educational Research Association, New York, New York, February 1971

EDRS PRICE EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS Comparative Statistics, *Computer Programs, Data Analysis, *Equated Scores, *Goodness of Fit, *Statistical Analysis, Statistics, *Test Interpretation, Test Results
IDENTIFIERS *Equipercentile Equating, FORTRAN IV

ABSTRACT

Prior use of the equipercentile method of test equating was based on a graphic procedure which is tedious, subject to smoothing errors, and non-analytical. Recognition of the equipercentile method as a curve-fitting procedure for two cumulative percentage distributions leads to a proposed analytical solution to the problem through use of linear estimates for successive "missing" score points. A complete equipercentile procedure which uses the proposed method and provides linear and quadratic functions for goodness-of-fit and extrapolation is discussed and illustrated with data from a test equating project. A FORTRAN IV program for the complete procedure is available. (Author)

AN ANALYTICAL PROCEDURE FOR THE EQUIPERCENTILE METHOD OF EQUATING TESTS

CARL A. LINDSAY

and

MARK A. PRICHARD

THE PENNSYLVANIA STATE UNIVERSITY

Paper Presented at the 1971 Annual American Educational Research Association, New York City, February 4 - 7, 1971

100 Word Abstract for

An Analytical Procedure for the
Equipercentile Method of Equating Tests

Carl A. Lindsay and Mark A. Prichard
The Pennsylvania State University

Prior use of the equipercentile method of test equating was based on a graphic procedure which is tedious, subject to smoothing errors, and non-analytical. Recognition of the equipercentile method as a curve-fitting procedure for two cumulative percentage distributions leads to a proposed analytical solution to the problem through use of linear estimates for successive "missing" score points. A complete equipercentile procedure which uses the proposed method and provides linear and quadratic functions for goodness-of-fit and extrapolation is discussed and illustrated with data from a test equating project. A FORTRAN IV program for the complete procedure is available.

An Analytical Procedure for the

Equipercentile Method of Equating Tests

Carl A. Lindsay and Mark A. Prichard
The Pennsylvania State University

The purpose of this paper is neither to defend nor decry the practice of equating tests, nor to discuss the various methods extant and their applicability to different situations (see Flanagan, 1950; Angoff, in press; also see Marks and Lindsay, 1970 for a theoretical and empirical treatment of test equating). Rather we shall restrict our attention to a discussion of an analytical procedure for the equipercentile method of equating tests.

According to Lord (1955), tests X and Y can be considered equivalent for a given group when the score scales on the two tests are so adjusted that both tests have the same frequency distribution of true scores in the given group. When both tests are equally reliable, or unreliable, then it makes little practical difference whether estimated true scores or obtained scores are equated. If we ignore the issue of whether or not we are dealing with obtained scores or estimated true scores, the equipercentile test equating method defines two scores on test X and Y as equivalent if their corresponding percentile ranks or relative cumulative frequency distributions in any given group are equal. The equipercentile method is thus seen as a curvilinear or area transformation problem rather than a linear or linear regression problem. Possible differences in the shapes of the two distributions are taken into account with the equipercentile method.

A graphic procedure, discussed by Flanagan (1950) and Angoff (in press), has previously been used for finding equivalent raw score points with the equipercentile method. First, relative cumulative frequencies are developed throughout the obtained score range for each test. The relative cumulative frequency distributions for tests are then plotted on arithmetic probability paper and a smooth line drawn en the points to obtain percentile ranks. Next, raw scores for both tests, associated with selected percentile ranks, are plotted on regualr graph paper. Finally, a smoothed line, drawn between the plotted points, is used to record the conversion of scores from test X to test Y.

It is easily seen that this procedure is tedious, subject to smoothing errors, and non-replicable, i.e., it does not use an analytical solution for finding equated score points. Although there have been two analytical methods proposed for smoothing obtained score distributions (Cureton & Tukey, 1951; Keats & Lord, 1962), no analytical solution for the equipercentile method has been proposed previously which both smooths the obtained distributions and develops equivalent scores. In addition, since the graphic procedure is done by hand, it is not convenient to develop functional equations for predicting equated scores on one test from the other and for extrapolation beyond the obtained data points.

The impetus for the present research was provided by an actual test equating problem requiring the use of the equipercentile method. Faced with the problem of cross-equating six subtests each from two nationally-used achievement tests given to over 3,000 fifth graders, the authors developed an analytical solution to the equipercentile method. A FORTRAN IV program was written which carries out the equipercentile equating procedure and provides linear and quadratic functions for goodness-of-fit and extrapolation as well as estimates of the error involved in predicting one test from the other.

Method

From an analytical point of view, the graphic procedure for the equipercentile method involves two steps: (a) interpolation and (b) extrapolation based on a curve-fitting procedure. The proposed method uses a linear rule to interpolate the two obtained distributions, then develops functional equations from the interpolated distributions for matching and extrapolation.

Interpolation Procedure

Given two cumulative percentage (CP) test score distributions X and Y, each score and its associated CP can be represented on an ordered pair, (X_i, P_j) (Y_k, P_k) where X_i refers to a given raw (scaled, grade equivalent, etc.) score and P_k to its associated CP. The first task involved with the equipercentile method is to find pairs of raw scores that cut off equal proportions of the two distributions. It is seen, then, that the interpolation problem of the equipercentile method involves the estimation of "missing" raw score points in one distribution for a given CP in the other. The proposed solution to the smoothing problem assumes that the best estimate of a "missing" score point on one distribution lies on a straight line connecting two adjacent CP's and their associated score points. It is derived below:

Given: (X_1, P_1) and (X_2, P_2) from distribution X

and

$(X_n$ and $P_n)$ from distribution Y, where $P_n \neq P_1, P_2$.

The problem is to find the X_n on distribution X associated with the P_n from distribution Y.

We begin with the following identities:

$$\left(\frac{X_n - X_1}{X_2 - X_1}\right) = \left(\frac{P_n - P_1}{P_2 - P_1}\right) \qquad (1)$$

and

$$X_n - X_1 = \left(\frac{P_n - P_1}{P_2 - P_1}\right) (X_2 - X_1) \qquad (2)$$

In other words, the linear rule requires a proportionate increase or decrease on score points in terms of their associated CP's.

Solving for X_n, the "missing" raw score point, we have the computing formula:

$$X_n = \left(\frac{P_n - P_1}{P_2 - P_1}\right) (X_2 - X_1) + X_1 \qquad (3)$$

The method, of course, is general and we may solve for Y_n by appropriate substitution. In practice the method finds successive "missing" points on X and Y by doing "flip-flop" estimates from one to the other.

An illustration of the linear rule for smoothing or interpolation is provided by the data in Table 1.

The left-hand side of Table 1 shows the original CP distributions for the tests. The right-hand side shows the smoothed distributions for the two tests obtained by applying the linear rule. Successive, "flip-flop" estimates of raw scores for "missing" CP's in one distribution are shown in parentheses. It is seen that this method yields two distributions with an equal number of raw score points in each. However, it is also evident that the smoothing procedure does not necessarily yield "equated" scores for all possible score points in the two distributions. So far, a score of 20 on X is equated with a score of 22 on Y, a score of 27 on X is equated with a score of 23.5 on Y, etc. But what about a score of 2 on X or a score of 22 on X? This problem is dealt with in the following section.

Curve-Fitting Procedure

The next step involves the development of functional equations, by the least squares criterion, for fitting the two interpolated distributions. The two smoothed distributions are taken as the best estimate of the complete distributions of scores if an infinite number of observations were available, and are subsequently referred to as actual distributions.

The curve-fitting procedure involves the development of first, second, or higher degree polynomials, using the interpolated score points. Standard least square numerical solutions are used to obtain the constants for the polynomials or prediction equations. The accuracy with which a given polynomial reproduces the actual trend of

Table 1

Original Cumulative Percentage Distribution and Smoothed Cumulative Percentage Distributions for Two Tests Obtained by the Linear Rule (Dummy Data)

Original				Smoothed			
Text X		Test Y		Test X		Test Y	
Score	CP	Score	CP	Score	CP	Score	CP
(X_i	P_i)	(Y_k	P_k)	(X_i	P_i)	(Y_k	P_k)
20	1	22	1	20	1	22	1
27	3	25	5	27	3	(23.5)	(3)
30	8	31	8	(28.2)	(5)	25	5
35	20	35	25	30	8	31	8
40	40	37	40	35	20	(33.8)	(20)
41	59	43	57	(36.3)	(25)	35	25
52	74	44	72	40	40	37	40
53	88	45	88	(40.9)	(57)	43	57
62	96	50	93	41	59	(43.2)	(59)
70	100	55	100	(50.5)	(72)	44	72
				52	74	(44.2)	(74)
				53	88	45	88
				(58.6)	(93)	50	93
				62	96	(52.1)	(96)
				70	100	55	100

Note: Values in parentheses were obtained by applying the linear rule to obtain score points corresponding to "missing" CP points.

the bivariate distribution is assessed by developing an estimate of the average error.

$$E = \sqrt{\frac{(Y - \hat{Y})^2}{K}} \qquad (4)$$

Where E = average error

Y = actual value of Y

$\hat{Y}$ = predicted value of Y (from X)

K = number of data points.

Results

The FORTRAN IV Program

A FORTRAN IV program has been written to carry out the complete analytical equipercentile equating method described above. It is currently available from the authors as a subroutine of the statistical package for the Social Sciences (Nie, Bent, and Hull, 1970). However, it can be modified to stand alone.

The program also has another convenient feature which is illustrated in the next section of this paper. It produces card or tape output containing the actual smoothed distributions data points and data points generated by the linear and quadratic prediction equations. These points serve as input to a CALCOMP plotter which produces a graphic representation of the goodness-of-fit.

Illustrative Results

Shown in Figures 1 and 2 are two CALCOMP plotter graphs based on actual data from a recently completed test equating project (Lindsay, 1969a; 1969b). The FORTRAN IV program briefly described above was written to carry out the equipercentile equating for this project.

As is known, the main object of test equating is to provide an estimate of an individual's score on, say, test Y, given his score on test X. The functional ...tions shown in Figures 1 and 2 were obtained by finding the best fit line (linear,

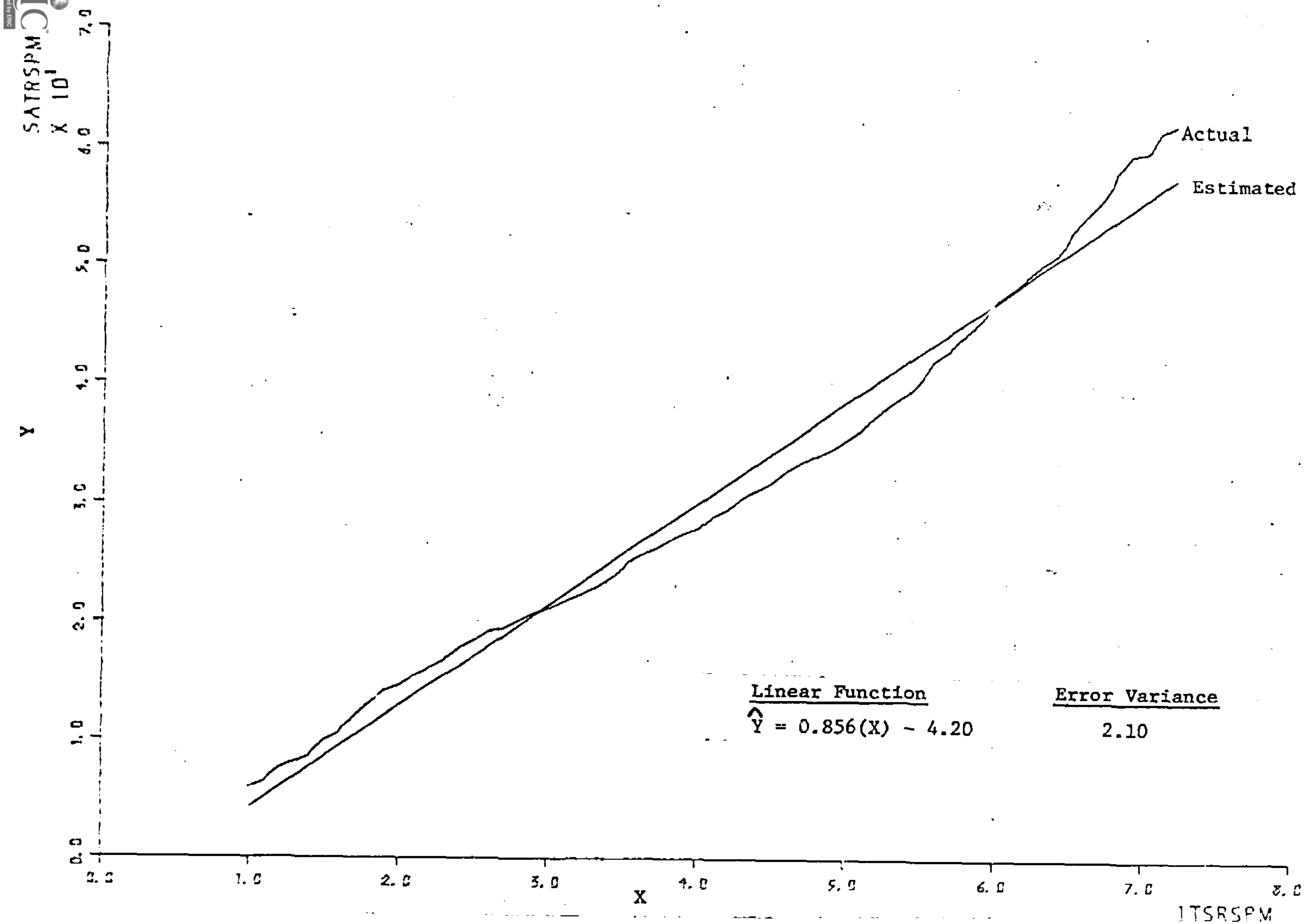

Fig. 1. Plot of data points generated by the linear function for estimating test Y scores from test X scores compared to the actual smoothed distribution obtained by the linear rule.

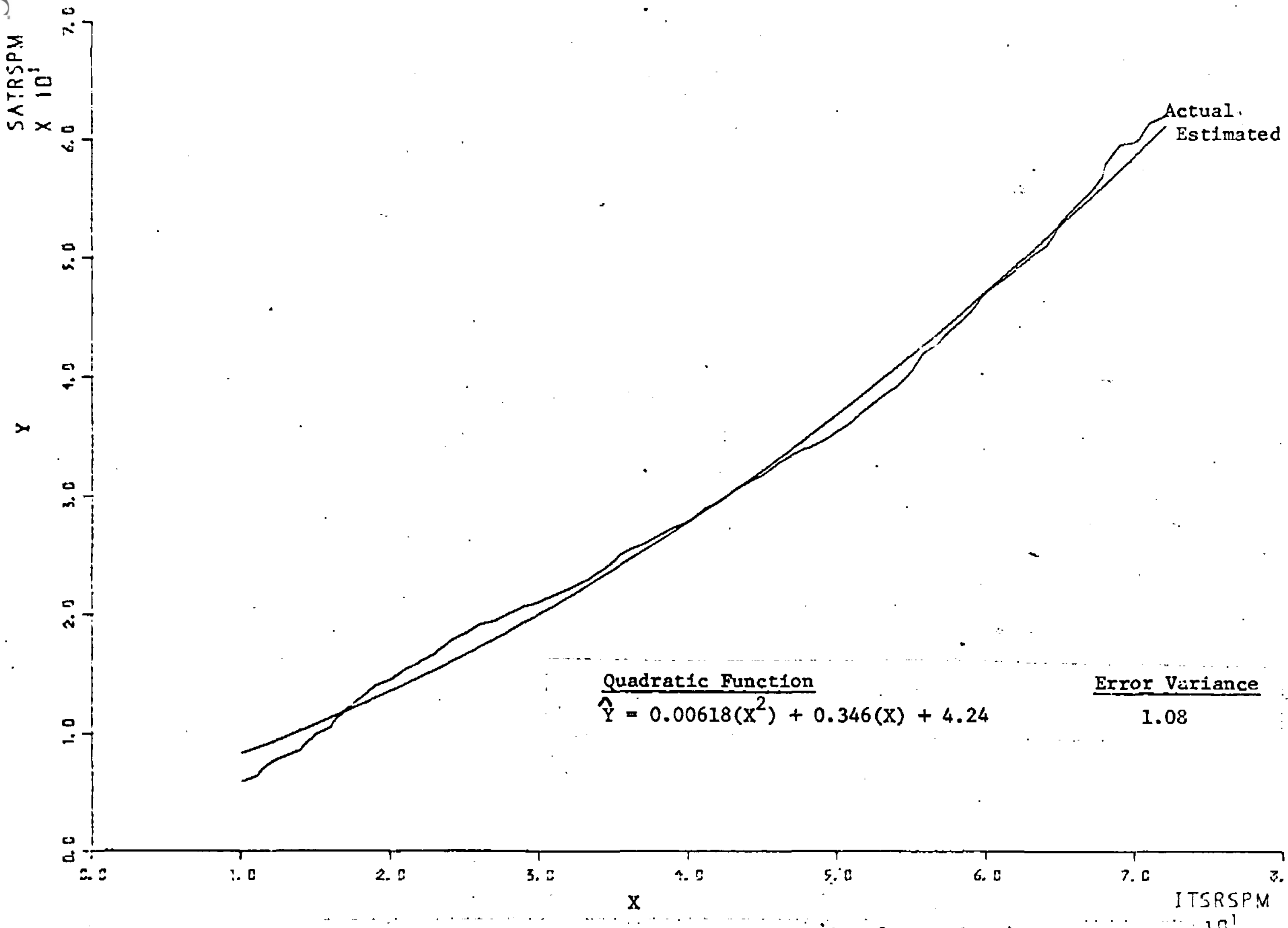

Fig. 2. Plot of data points generated by the quadratic function for estimating test Y scores from test X scores compared to the actual smoothed distribution obtained by the quadratic rule.

Figure 1 and quadratic, Figure 2) for predicting test Y from test X. The two figures then show graphically how well the estimated distribution, generated by a given equation, fits the actual distribution. It is easily seen that the quadratic equation provides the best fit to the actual bivariate distribution. This is borne out by an error variance of 1.08 for the quadratic equation as compared to one of 2.10 for the linear equation.

Conclusions

It is recognized that the proposed analytical solution is appropriate only in those instances that the graphic one is. It does not solve the problems associated with a small number of subjects, highly skewed distributions, etc. In fact the solution should be viewed only as an analogue to the graphic method. It is concluded that the proposed method has merit because the results are verifiable, it is fast and inexpensive if done on a computer, it eliminates tedious hand-smoothing, and puts the equipercentile method on the same analytical basis as the linear method of test equating.

References

Angoff, W.H., Can useful general-purpose equivalency tables be prepared for different college admission tests? In A. Anastasi (Ed.), Testing problems in perspective. Washington, D.C.: American Council on Education, 1966.

Cureton, E.E., and Tukey, J.W., Smoothing frequency distributions, equating tests, and preparing norms. American Psychologist, 1951, 8, 404 (Abstract).

Flanagan, J.C., Units, scores and norms. In E.F. Lindquist (Ed.), Educational Measurement. Washington, D.C.: American Council on Education, 1951.

Keats, J.A., and Lord, F.M., A theoretical distribution for mental test scores. Psychometrika, 1962, 27, 59-72.

Lord, F.M., Equating test scores - a maximum likelihood solution. Psychometrika, 1955, 20, 193-200.

Lindsay, C.A., The development of comparable scores for selected subtests of the Iowa Test of Basic Skills and the Stanford Achievement Test: The Spring 1969, Bureau of Research Test Equating Project. Unpublished Student Affairs Research Report, University Park, Pa.; The Pennsylvania State University, October, 1969a.

Lindsay, C.A., The development and comparison of a set of linear and quadratic equations for the cross-equating of selected Stanford Achievement Test and the Iowa Test of Basic Skills Raw Score Subtests. Unpublished Student Affairs Research Report, University Park, Pa.; The Pennsylvania State University, December, 1969b.

Marks, F., and Lindsay, C.A., Test Equating: some theoretical and empirical considerations. Students Affairs Research Report 70-12, University Park, Pa.; The Pennsylvania State University, August, 1970.

Nie, N.H., Bent, D.H., and Hull, C.H., Statistical package for the social sciences. New York: McGraw-Hill, 1970.