Statistics for Social Science and Public Policy

Advisors:

S.E. Fienberg D. Lievesley J. Rolph

Springer

New York

Berlin

Heidelberg

Hong Kong

London

Milan

Paris

Tokyo

Alina A. von Davier
Paul W. Holland
Dorothy T. Thayer

The Kernel Method of Test Equating

With 63 Illustrations

Springer

Alina A. von Davier
Paul W. Holland
Dorothy T. Thayer
Educational Testing Service
Rosedale Road 12-T
Princeton, NJ 08541
USA

avondavier@ets.org
pholland@ets.org
dthayer@ets.org

Advisors:

Stephen E. Fienberg
Department of Statistics
Carnegie Mellon University
Pittsburgh, PA 15213
USA

Denise Lievesley
Institute for Statistics
Room H.113
UNESCO
7 Place de Fontenoy SP
75352 Paris 07
France

John Rolph
Department of Information and
Operations Management
Graduate School of Business
University of Southern California
Los Angeles, CA 90089
USA

Library of Congress Cataloging-in-Publication Data

Davies, Alina von.

The Kernel method of test equating / Alina von Davier, Paul Holland, Dorothy Thayer.

p. cm. — (Statistics for social science and public policy)

Includes bibliographical references and index.

ISBN 0-387-01985-5 (alk. paper)

1. Examinations—Scoring. 2. Examinations—Interpretation. 3. Examinations—Design and Construction. 4. Educational tests and measurements—Standards. I. Holland, Paul W.

II. Thayer, Dorothy T. III. Title. IV. Series.

LB3060.77.D38 2003

371.26'01'3—dc21

2003050494

ISBN 0-387-01985-5

Printed on acid-free paper.

© 2004 Springer-Verlag New York, Inc.

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer-Verlag New York, Inc., 175 Fifth Avenue, New York, NY 10010, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden. The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

Printed in the United States of America.

9 8 7 6 5 4 3 2 1

SPIN 10925911

Typesetting: Pages created by the authors using a Springer $\text{\LaTeX}$ macro package.

www.springer-ny.com

Springer-Verlag New York Berlin Heidelberg

A member of BertelsmannSpringer Science+Business Media GmbH

To Matthias and Thomas
—A.A.v.D.

To John, Dawn, and Frank
—P.W.H.

To Richard
—D.T.T.

Preface

The antecedents of this book begin with the Program Statistics Research Project at Educational Testing Service (ETS). This was an initiative started in 1978 by Donald Rubin (in the Research Division) and Robert Solomon (the Executive Vice President of ETS at that time). Its purpose was to focus the statistical research interests of the newly formed Research Statistics Group on problems relevant to the work of the testing programs at ETS. Paul Holland was given the responsibility for test equating research while Rubin primarily worked on problems related to predictive validity (Rubin, 1980). These two topics are reoccurring themes for the testing programs. Of course, research on test equating has had a long history at ETS; the most notable early work was that of Lord (1950, 1955a) and Wilks (1961). In addition, Ledyard Tucker was renowned for his many theoretical and practical developments in this area.

The work on test equating of the Program Statistics Research Project led to a conference on test equating in 1980 reported in Holland and Rubin (1982). A new equating method, Section Pre-Equating (Holland and Thayer, 1981), was invented as a direct consequence of the basic equating research of the project and the need for a technical response to the New York test disclosure legislation of the early 1980's (McAllister, 1993).

Kernel Equating (KE) was developed in a series of ETS technical reports as a natural next step in this program of research on test equating. The starting point was the development of useful probability models for fitting the score distributions that arise in test equating (Holland and Thayer, 1987). The log-linear models for score distributions discussed there were uniquely able to fit the real data of many practical testing situations. Prior

to that work there were only a few models available for fitting score distributions and these did not deal with many of the complexities of actual test data. The next step was to develop a new equating method that could fully exploit the log-linear models for score distributions. This was quick to follow. The result was Kernel Equating. KE was developed during the late 1980's and documented in two ETS technical reports, (Holland and Thayer, 1989; Holland et al., 1989). Liou and Cheng (1995) also contributed to the development of KE by connecting it to problems of estimation when there is missing data.

After the initial flurry of this half-decade of development, further work on Kernel Equating stalled. While it remained a specialized method of some theoretical interest, during this period it was used in only one practical circumstance of which we are aware, reported in Dorans (2002). Other research and practical issues, such as DIF (Holland and Wainer, 1993), took attention away from further developments of KE at ETS. Work was done by others that showed the utility of log-linear smoothing of score distributions in test equating applications (Livingston, 1993a; Hanson, 1996). However, except for a brief evaluation study (Livingston, 1993b) nothing was done to widen the applicability of KE beyond the initial two designs discussed in the earlier technical reports—the Equivalent Groups and the Non-Equivalent groups with Anchor Test Designs.

In the summer of 2000, Holland returned to ETS after a seven-year period at the UC Berkeley Graduate School of Education and resumed his long-standing research collaboration with Dorothy Thayer. In the spring of 2001, he met Alina von Davier, a mathematical statistician who was looking for interesting research questions. Their conversations led her to join the research staff at ETS in the summer of 2001 and to the subsequent decision for the three authors to write this book. It was realized early on that to be useful, the book would need to include ways to apply KE to all of the usual equating designs. This required a good deal of new research that was carried out as the book was written and supported by the ETS Research Allocation.

Not only were we able to apply KE to all of the usual designs, but our research also led us to new ideas including the Standard Error of Equating Difference (SEED), a measure of the accuracy of the estimated difference between two equating functions.

By the winter of 2002, we had a good working draft of *The Kernel Method of Test Equating*, and began to try it out on those of our colleagues in R&D at ETS who have operational responsibilities for equating tests on a regular basis. We ran a workshop on KE at ETS and used it to improve the material in our initial draft of the book. This book is the final result.

However, we believe that serious further research on Kernel Equating is not only possible, but is a program likely to bear fruit. Throughout this book we indicate topics that can further illuminate and improve upon what we present here. We hope that researchers and graduate students in

statistics, psychometrics, and educational measurement, looking for useful research topics, will consider some of the possibilities we indicate here, and push the work reported in this book in new directions.

It is important to remember that test equating is about 100 years old, and was used as early as the Army Alpha tests of the early part of the 20th century. By now some think that modern test theory has removed the need to equate tests, but we think this is a dream yet unrealized, and perhaps unrealizable. The equivalence of scores from different tests, no matter how carefully constructed to be parallel, is always an empirical question. The methods and models of test equating have been, for nearly a century, the way that data and theory have come to grips with this hoped-for equivalence. We do not expect it to be any different in the near future, or, perhaps, ever.

How this book is organized. The book is divided into two parts. In the first we develop the theory of Kernel Equating, and in the second, we apply it to real examples of the most important equating designs. Chapter 1 is an introduction in which we give a little history of equating, some general discussion about test linking and test equating in particular, and develop the basic notation that we use throughout the rest of the book. In addition, we review the linear and equipercentile equating functions and the standard error of equating (SEE). Chapter 2 reviews the four basic test-equating designs as well as the application of both the chain and the post-stratification approaches to the Non-Equivalent groups with Anchor Test Design. Our development of the equating designs is aimed at our later discussion of them within the context of KE. Chapters 3, 4, and 5 all deal with the mathematical basis of KE. These are technical chapters that spell out, in detail, the theory of KE. Chapter 6 compares KE to both the linear method and percentile rank method of equipercentile equating and addresses what we believe are its advantages over both.

Part II of the book begins with Chapter 7. Each of the five chapters in Part II addresses a different equating design and shows how to apply KE to it. In these chapters we use real data to show how KE can be applied in each situation. We have attempted to give enough information in each of these application chapters that practitioners can see how the new tools that KE provides work, as well as how the actual use of KE can proceed in practice. Furthermore, to aid in the use of this book as a reference for KE, we tried to make each of the chapters in Part II as similar as we could. The result is a certain amount of repetition in Chapters 7—11. We also make continual references to the appropriate parts of Part I in Chapters 7—11, as well.

Our application of KE to the Counterbalanced Design gives a new approach to equating using this design. Our treatment of both Chain and Post-Stratification Equating in the Non-Equivalent groups with Anchor Test Design is, to our knowledge, the first to give parallel developments

of these two approaches from a common point of view. Finally, our development of the standard error of equating (SEE) for all of the common equating designs, including Chain Equating for the Non-Equivalent groups with Anchor Test Design, is probably the most comprehensive comparison of the SEE's since the early work of Lord (1950, 1955a).

Acknowledgements. We are indebted to ETS and to many ETSers for their assistance and encouragement in the production of this book: Drew Gitomer for financial support; John Mazzeo, Neil Dorans, Skip Livingston, Krishna Tateneni, Wen-Ling Yang, Cathy Wendler, and Miao-Hsiang Lin for encouragement, review, and valuable suggestions; Kristen Huff and Miriam Feigenbaum for their support in the process of gathering the data; and Diane Rein, Elizabeth Brophy, and Martha Thompson for their assistance in the production of the book.

This work is a collaboration in every respect and the order of the authorship is alphabetical.

Princeton, NJ

Alina A. von Davier

September, 2003

Paul W. Holland

Dorothy T. Thayer

Contents

Preface	vii
List of Notation	xv
1 Introduction and Notation	1
1.1 Introduction	1
1.2 The Notation Used in This Book	5
1.3 The Linear Equating Function	8
1.4 The Equipercntile Equating Function	9
1.5 The Relationship Between $\text{Lin}_Y(x)$ and $\text{Equi}_Y(x)$	11
1.6 Data Collection Designs	13
1.7 Sample Estimates	14
1.8 A Summary of the New Material in This Book	15
Part I—The Kernel Method of Test Equating: Theory	17
2 Data Collection Designs	19
2.1 The Equivalent-Groups Design (EG)	21
2.2 The Single-Group Design (SG)	22
2.3 The Counterbalanced Design (CB)	27
2.4 Non-Equivalent groups with Anchor Test Design (NEAT)	32
2.4.1 Chain Equating (CE)	35
2.4.2 Post-Stratification Equating (PSE)	39
2.4.3 Internal Anchor Tests and Structural Zeros	41

2.4.4	The EG Design with an Anchor Test	43
2.5	Random versus Spiraled Samples	43
2.6	Summary	44
3	Kernel Equating: Overview, Pre-smoothing, and Estimation of r and s	45
3.1	The Five Steps of Kernel Equating: Overview	45
3.2	Pre-smoothing Using Log-Linear Models	47
3.2.1	Estimating a Univariate Score Distribution	49
3.3	Estimation of the Score Probabilities	52
4	Kernel Equating: Continuization and Equating	55
4.1	Continuization	55
4.1.1	Gaussian Kernel Smoothing	56
4.1.2	Choice of the Bandwidth	61
4.2	Equating	64
5	Kernel Equating: The SEE and the SEED	67
5.1	Introduction	68
5.2	The δ -Method Divides the Problem in Three	69
5.3	The SEE and the SEED for Kernel Equating	73
5.3.1	Computing $\mathbf{J}_{e_Y}$ for Kernel Equating	73
5.3.2	The SEE for Specific Equating Designs	76
5.3.3	The SEED	79
5.4	The SEE and SEED for Chain Equating	81
6	Kernel Equating versus Other Equating Methods	87
6.1	KE versus Linear Equating	88
6.2	KE versus the Percentile Rank Method	90
6.2.1	The Percentile Rank Method	90
6.2.2	Some Facts About PRM	91
6.2.3	Distributional Characteristics of PRM	92
6.2.4	The Issue of the Finite Range	92
6.3	Viewing PRM from the KE Perspective	93
6.4	Advantages of KE over PRM	94
Part II—The Kernel Method of Test Equating: Applications		97
7	The Equivalent-Groups Design	99
7.1	Pre-smoothing	101
7.2	Estimation of the Score Probabilities	103
7.3	Continuization	104
7.4	Equating	106
7.5	Standard Error of Equating	108
7.6	Deciding Between $\hat{e}_Y(x)$ and $\widehat{\text{Lin}}_Y(x)$	110

8	The Single-Group Design	113
8.1	Pre-smoothing	116
8.2	Estimation of the Score Probabilities	121
8.3	Continuization	123
8.4	Equating	124
8.5	Standard Error of Equating	126
8.6	Deciding Between $\hat{e}_Y(x)$ and $\widehat{\text{Lin}}_Y(x)$	128
9	The Counterbalanced Design	131
9.1	Pre-smoothing	133
9.2	Estimation of the Score Probabilities	139
9.3	Continuization	141
9.4	Equating	143
9.5	Standard Error of Equating	145
9.6	Deciding Between $\hat{e}_{Y1}(x)$ and $\hat{e}_{Y\frac{1}{2}}(x)$	148
9.7	Diagnosis of the Equating Process	149
9.8	Deciding Between $\hat{e}_{Y\frac{1}{2}}(x)$ and $\widehat{\text{Lin}}_{Y\frac{1}{2}}(x)$	150
9.9	Appendix: The Data Used in This Chapter	152
10	The NEAT Design: Chain Equating	155
10.1	Pre-smoothing	159
10.2	Estimation of the Score Probabilities	167
10.3	Continuization	168
10.4	Equating	170
10.5	Standard Error of Equating	174
10.6	Deciding Between $\hat{e}_{Y(CE)}(x)$ and $\widehat{\text{Lin}}_Y(x)$	175
11	The NEAT Design: Post-Stratification Equating	179
11.1	Estimation of the Score Probabilities	181
11.2	Continuization	183
11.3	Equating	184
11.4	Standard Error of Equating	187
11.5	The Choice of the Target Population	188
11.6	Deciding Between $\hat{e}_{\frac{1}{2}Y}(x)$ and $\widehat{\text{Lin}}_{\frac{1}{2}Y}(x)$	190
11.7	Comparing the KE Functions for PSE and CE	192
11.8	CE versus PSE: Which One to Choose?	194
	Appendix	197
A	The δ-Method	197
B	Bivariate Smoothing	199
B.1	Assessing the Fit of the Log-Linear Models	201
B.1.1	Covariance Matrix of the Parameters	201
C	Other Univariate Moments	203

D Review of the Use of Matrices in This Book	205
Bibliography	216
Author Index	223
Subject Index	225

List of Notation

The list of notation is ordered, approximately, by first appearance in the book.

TABLE 1:

Symbol	Explanation	First appearance
$X, Y :$	The names of two tests to be equated.	1.2
$A :$	The name of the anchor test in the Non-Equivalent groups with Anchor Test (NEAT) Design.	1.2
$\mathbf{X}, \mathbf{Y} :$	The scores from X and Y , respectively, regarded as random variables.	1.2
$\mathbf{A} :$	The score on the anchor test, $\mathbf{A}$, regarded as a random variable.	1.2
$J, K, L :$	Number of possible $\mathbf{X}$ -, $\mathbf{Y}$ -, or $\mathbf{A}$ -scores, respectively.	1.2
$x_j :$	A possible score value for $\mathbf{X}$, $j = 1$ to J .	1.2
$y_k :$	A possible score value for $\mathbf{Y}$, $k = 1$ to K .	1.2
$a_l :$	A possible score value for $\mathbf{A}$, $l = 1$ to L .	1.2

continued on the next page

TABLE 1: *continued*

Symbol	Explanation	First appearance
$T :$	The target population of examinees on which the equating of $\mathbf{X}$ and $\mathbf{Y}$ takes place.	1.2
$\sum_j :$	Sum over the full range of j .	1.2
$G^{-1}(u) :$	Inverse function of G . The solution to $u = G(y)$, for a specified u .	1.4
$r_j :$	$\text{Prob}\{\mathbf{X} = x_j T\}$ = the score probability of $\mathbf{X} = x_j$ over the target population, T .	1.2
$s_k :$	$\text{Prob}\{\mathbf{Y} = y_k T\}$ = the score probability of $\mathbf{Y} = y_k$ over the target population, T .	1.2
$t_l :$	$\text{Prob}\{\mathbf{A} = a_l T\}$ = the score probability of $\mathbf{A} = a_l$ over the target population, T .	1.2
$F(x) :$	$\text{Prob}\{\mathbf{X} \leq x T\}$ = cumulative distribution function (cdf) of $\mathbf{X}$ over T .	1.2
$G(y) :$	$\text{Prob}\{\mathbf{Y} \leq y T\}$ = cdf of $\mathbf{Y}$ over T .	1.2
$H(a) :$	$\text{Prob}\{\mathbf{A} \leq a T\}$ = cdf of $\mathbf{A}$ over T . Subscripts on F , G , and H indicate that the cdf's refer to other populations, i.e., P or Q .	1.2
$\mu_X, \sigma_X :$	The mean and standard deviation of $\mathbf{X}$ over the target population, T .	1.2
$\mu_Y, \sigma_Y :$	The mean and standard deviation of $\mathbf{Y}$ over the target population, T .	1.2
$\text{Lin}_Y :$	The linear equating function for equating $\mathbf{X}$ to $\mathbf{Y}$ over T .	1.3
$\text{Equi}_Y :$	Any version of the equipercentile equating function for equating $\mathbf{X}$ to $\mathbf{Y}$ over T .	1.4
$e_Y :$	Usually, the Kernel Equating function for equating $\mathbf{X}$ to $\mathbf{Y}$ over T . It is used in Chapter 1 to refer to any equating function that equates $\mathbf{X}$ to $\mathbf{Y}$ over T .	1.7

continued on the next page

TABLE 1: *continued*

Symbol	Explanation	First appearance
KE :	The Kernel Method of Equating.	1.1
EG :	The Equivalent-Groups Design.	1.6
SG :	The Single-Group Design.	1.6
CB :	The Counterbalanced Design.	1.6
NEAT :	The Non-Equivalent groups with Anchor Test Design.	1.6
SEE_Y :	The Standard Error of Equating for an equating function that equates $\mathbf{X}$ to $\mathbf{Y}$ on T .	1.7
P and Q :	Populations of examinees.	2
$r_{Pj}, s_{Qk},$ t_{Pl}, t_{Ql} :	Score probabilities for $\mathbf{X}$, $\mathbf{Y}$, and $\mathbf{A}$ over the populations P and Q (the NEAT Design).	2.4.1
$\mathbf{r}$:	The (column) vector of score probabilities, r_j .	2.1
$\mathbf{s}$:	The (column) vector of score probabilities, s_k .	2.1
$\mathbf{t}$:	The (column) vector of score probabilities, t_l .	2.4.1
$\mathbf{r}_P, \mathbf{s}_Q,$ $\mathbf{t}_P, \mathbf{t}_Q$:	The (column) vectors of score probabilities for $\mathbf{X}$, $\mathbf{Y}$, and $\mathbf{A}$ over the populations P and Q (the NEAT Design).	2.4.1
DF :	Design Function.	2
df :	Degrees of freedom.	9
$\mathbb{R}^J$:	Euclidean J -dimensional space.	2.1
Ω_J :	The set of all J -dimensional score probability vectors, $\mathbf{r}$.	2.1
p_{jk} :	$\text{Prob}\{\mathbf{X} = x_j, \mathbf{Y} = y_k T\}$ = the bivariate score probability of $\mathbf{X} = x_j$ and $\mathbf{Y} = y_k$ over the target population, T .	2.2

continued on the next page

TABLE 1: *continued*

Symbol	Explanation	First appearance
$\mathbf{P}$:	For the SG Design, the J by K matrix of bivariate score probabilities for $(\mathbf{X}, \mathbf{Y})$. For the NEAT Design, the J by L matrix of bivariate score probabilities, p_{jl} , for $(\mathbf{X}, \mathbf{A})$.	2.2
$\mathbf{p}_k$:	The k th column vector of $\mathbf{P}$.	2.2, 2.4
$v(\mathbf{P})$:	The vectorization of the matrix, $\mathbf{P}$. In the SG Design, $v(\mathbf{P})$ is a $-dimensional column vector.$	2.2
q_{kl} :	$\text{Prob}\{\mathbf{Y} = y_k, \mathbf{A} = a_l Q\}$ = the bivariate score probability of $\mathbf{Y} = y_k$ and $\mathbf{A} = a_l$ over the population, Q .	2.4
$\mathbf{Q}$:	For the NEAT Design, the K by L matrix of bivariate score probabilities for $(\mathbf{Y}, \mathbf{A})$ on Q .	2.4
$\mathbf{q}_l$:	The l th column of $\mathbf{Q}$.	2.4
$v(\mathbf{Q})$:	The vectorization of $\mathbf{Q}$. In the NEAT Design, $v(\mathbf{Q})$ is a KL -dimensional column vector.	2.4
$\mathbf{I}_J$:	The J by J identity matrix.	2.1
$\mathbf{0}$:	A matrix or vector of all zeros of appropriate dimension.	2.1
$\mathbf{M}$:	A matrix of zeros and ones used to compute the column vector of the row sums of a bivariate score probability matrix, $\mathbf{P}$, from its vectorized version, $v(\mathbf{P})$. $\mathbf{M}$ is defined in equations (2.9) and (2.11). $\mathbf{M}$ with subscripts denotes versions of $\mathbf{M}$ that arise in various equating designs throughout this book.	2.2
$\mathbf{N}$:	A matrix of zeros and ones used to compute the column vector of the column sums of a bivariate score probability matrix, $\mathbf{P}$, from its vectorized version, $v(\mathbf{P})$. $\mathbf{N}$ is defined in equation (2.10) and (2.12). $\mathbf{N}$ with subscripts denotes versions of $\mathbf{N}$ that arise in various equating designs throughout this book.	2.2

continued on the next page

TABLE 1: *continued*

Symbol	Explanation	First appearance
$\mathbf{1}_J :$	A column vector of all ones.	2.2
$\mathbf{X}_1 :$	In the CB Design, the $\mathbf{X}$ -score that is obtain when examinees are tested first by X and second by Y . Likewise, $\mathbf{X}_2$, $\mathbf{Y}_1$, and $\mathbf{Y}_2$ are defined analogously.	2.3
$\mathbf{P}_{(12)} :$	The version of the bivariate score probability matrix that arises in the CB design when X is taken first and Y second. $P_{(21)}$ is defined analogously.	2.3
$p_{(12)jk} :$	The entries in the matrix, $\mathbf{P}_{(12)}$. $p_{(21)jk}$ are the entries in $\mathbf{P}_{(21)}$.	2.3
$w_X, w_Y :$	In the CB Design, the weight given to the data that are not subject to order effects.	2.3
$\mathbf{r}_1, \mathbf{s}_1, \mathbf{r}_2, \mathbf{s}_2 :$	In the CB Design, the score probability vectors for $\mathbf{X}_1$, $\mathbf{Y}_1$, $\mathbf{X}_2$, and $\mathbf{Y}_2$, respectively.	2.3
$w :$	In the NEAT Design, the weight given to population P in the definition of the Target Population, $T = wP + (1 - w)Q$.	2.4
$e_{Y(CE)} :$	Chain Equating function equating $\mathbf{X}$ to $\mathbf{Y}$.	2.4.1
$e_{Y(PSE)} :$	Post-Stratification Equating function equating $\mathbf{X}$ to $\mathbf{Y}$.	2.4.2
$\hat{e}_{Yh_Xh_Y} :$	KE function equating $\mathbf{X}$ to $\mathbf{Y}$ using bandwidths, h_X and h_Y , in the continuization step.	3.1
$\hat{F}_{h_X} :$	The KE continuized version of $F(x)$ using a bandwidth of h_X . Defined in (4.5).	4.1
$\hat{G}_{h_Y} :$	The KE continuized version of $G(y)$ using a bandwidth of h_Y . Defined in (4.8).	4.1
$\mathbf{u} :$	The vector used to specify a log-linear model for score probabilities, whose dimension equals the number of score values. In our examples, $\mathbf{u} = \mathbf{0}$.	3.2.1

continued on the next page

TABLE 1: *continued*

Symbol	Explanation	First appearance
$\mathbf{B}$:	The $\mathbf{B}$ -matrix used to specify a log-linear model for score probabilities. Row dimension of $\mathbf{B}$ = the number of parameters in the model and the Column dimension of $\mathbf{B}$ = the number of score values.	3.2.1
$I_S(x_j)$:	The subset indicator variable, which has the value 1 if $x_j \in S$, and 0 otherwise.	3.2.1
$\Sigma_{\hat{\mathbf{r}}, \hat{\mathbf{s}}}$:	The cross-covariance matrix of the estimated score probability vectors, $\hat{\mathbf{r}}$ and $\hat{\mathbf{s}}$.	3.2.1
$\Sigma_{\hat{\mathbf{r}}}$:	The covariance matrix of the estimated score probability vector, $\hat{\mathbf{r}}$.	3.2.1
$\mathbf{C}_r$:	The $\mathbf{C}$ -matrix factor of $\Sigma_{\hat{\mathbf{r}}}$. $\mathbf{C}$ -matrices occur throughout this book, with and without subscripts, to make them appropriate to specific equating designs.	3.2.1
$\mathbf{D}_r$:	The diagonal matrix with the vector $\mathbf{r}$ along its main diagonal.	3.2.1
$\sqrt{\mathbf{r}}$:	The vector whose entries are the (positive) square roots of the vector $\mathbf{r}$.	3.2.1
a_X :	The value defined in equation (4.4), that is part of the definition of the KE continuization process using Gaussian Kernel smoothing.	4.1.1
h_X, h_Y :	The two bandwidths that are used to define the KE continuizations of $F(x)$ and $G(y)$. They are positive numbers. Large values of the bandwidths lead to linear equating, while smaller values give more “equipercentile-like” equating functions.	4.1.1
$\mathcal{N}(\mu, \sigma^2)$:	The Normal or Gaussian distribution with mean μ and variance, σ^2 .	4.1.1
Φ :	The Normal or Gaussian cdf.	4.1.1

continued on the next page

TABLE 1: *continued*

Symbol	Explanation	First appearance
$\text{PEN}_1 :$	A penalty function used to chose an “optimal” bandwidth in the continuization process. Defined in equation (4.27).	4.1.2
$\text{PEN}_2 :$	A penalty function used to chose an “optimal” bandwidth in the continuization process. Defined in equation (4.29).	4.1.2
$\text{PRE}(p) :$	The “Percent Relative Error” in the p th moments. A tool to compare the distribution of $\mathbf{Y}$ with the equated values, $e_Y(\mathbf{X})$.	4.2
$\mathbf{R}, \mathbf{S} :$	A general notation to denote the pre-smoothed data that arises in any equating design. Values of $\mathbf{R}$ and $\mathbf{S}$ are specified in Table 5.1 for each equating design in this book. The related notation for the number of score values, parameter estimates and $\mathbf{C}$ -matrices are given in Table 5.2.	5.2
$\mathbf{J}_{e_Y} :$	The Jacobian matrix of the KE function with respect to the score probabilities, $\mathbf{r}$ and $\mathbf{s}$. Defined in equations (5.9) and (5.19).	5.2
$\mathbf{J}_{DF} :$	The Jacobian matrix of the Design Function with respect to the score probabilities of the pre-smoothed data in each equating design. Defined in Tables 5.3 and 5.4 for the designs used in this book.	5.2
$\ \mathbf{v}\ :$	The Euclidean length (norm) of the vector, $\mathbf{v}$. The square-root of the sum of squares of its coordinates.	5.2
$\frac{\partial F}{\partial \mathbf{r}} :$	The row vector of partial derivatives of the KE continuized cdf, $F(x; \mathbf{r})$ with respect to each coordinate of $\mathbf{r}$. Analogously for $\frac{\partial G}{\partial \mathbf{s}}$.	5.3.1
$F', G' :$	The density functions of the continuized cdf's F and G .	5.3.1
$\mathbf{U}, \mathbf{V} :$	Matrices, with and without various subscripts, that are all similar and defined in the discussion of Tables 5.5 and 5.6.	5.3.2

continued on the next page

TABLE 1: *continued*

Symbol	Explanation	First appearance
SE-vector:	A vector, of the general form, $\mathbf{J}_{e_Y} \mathbf{J}_{DF} \mathbf{C}$, whose length is the value of $\text{SEE}_Y(x)$. An important computational tool for KE.	5.2
SEED_Y :	The Standard Error of Equating Difference for the difference between two equating functions that both equate $\mathbf{X}$ to $\mathbf{Y}$ on T .	5.3.3
PRM_Y :	The percentile rank method of equipercntile equating $\mathbf{X}$ to $\mathbf{Y}$.	6.2
N, M :	The samples sizes for various designs. Subscripts are added as necessary to deal with individual equating designs.	7
LR :	Likelihood ratio Chi-square statistic.	9.1