Chapter 9
An Alternative Continuization Method: The Continuized Log-Linear Method

Tianyou Wang

9.1 Introduction

von Davier, Holland and Thayer (2004b, pp. 45–47) described a five-step, test-equating framework: (a) presmoothing, (b) estimating score probabilities, (c) continuization, (d) equating, and (e) calculating the standard error of equating. In this framework, the presmoothing step is usually done with log-linear smoothing. Step 2 is to transform smoothed distribution into two marginal distributions for the target population (sometimes called synthetic population). In their framework, Step 3 is done with an adjusted Gaussian kernel procedure.

The advantage of the von Davier et al. (2004b) framework is that it modularizes the equating process so that different designs and methods only affect certain steps. For instance, different data collection designs will result in different design functions in Step 2. For a random-groups design, Step 2 is usually omitted in the traditional description of the equating process, but in this framework an identity design function is used. Likewise, different equating methods only affect Step 4.

The main difference between this framework and previous equating procedures is that it has a continuization step, so that the equating step is based on two continuous distributions rather than two discrete distributions. Denote the random variables for the test scores for test X as *X* and for test Y as *Y*, and the target population cumulative distributions of *X* and *Y* as *F(X)* and *G(Y)*, respectively. Then the equipercentile equating function $\hat{e}_Y(x)$ is given by Equation 9.1:

$$\hat{e}_Y(X) = G^{-1}(F(X)), \tag{9.1}$$

The traditional, percentile rank-based, equating procedure also can be viewed as a uniform-kernel continuization procedure under this framework. However,

---

T. Wang
The University of Iowa, 210B Lindquist Center, Iowa City, IA 52242, USA
e-mail: tianyouwang@yahoo.com

A.A. von Davier (ed.), *Statistical Models for Test Equating, Scaling, and Linking*,
Statistics for Social and Behavioral Sciences,
DOI 10.1007/978-0-387-98138-3_9,

uniform kernel produces piecewise linear culumlative distributions, which may not be the ideal procedure. Wang (2008) proposed an alternative continuization method that directly takes the log-linear function in the presmoothing step and transforms it into a continuous distribution. This method is called the continuized log-linear (CLL) method and is described for different data collection designs in the next two sections.

9.2 The CLL Method for the Equivalent-Groups Design

For the equivalent-groups design, the design function is the identity function. The distributions obtained from Step 2 are the same as those from Step 1. For this design, an alternative continuization procedure that utilizes the polynomial log-linear function obtained in the log-linear smoothing step is presented here, the CLL distribution. The probability density function (PDF) is expressed as

$$f(x) = \frac{1}{D} exp(\mathbf{b}^T \boldsymbol{\beta}), \tag{9.2}$$

where $\mathbf{b}^T = (1, x, x^2, \ldots, x^M)$ is a vector of polynomial terms of test X score x, $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \ldots, \beta_M)^T$ is the vector of parameters, and M is the order (or degree) of the polynomial. Holland and Thayer (1987, 2000) gave detailed descriptions about model parameter estimation and how to select the polynomial degree M. D is a normalizing constant that ensures that $f(x)$ is a PDF.

It is easy to show that all the moments of the CLL distribution are approximately equal to those of the smoothed discrete distribution by the following relationship between *i-th* noncentral moments of the CLL distribution and the smoothed discrete distribution:

$$\frac{\int_l^u x^i exp(\mathbf{b}^\mathbf{T} \boldsymbol{\beta}) dx}{\int_l^u exp(\mathbf{b}^\mathbf{T} \boldsymbol{\beta}) dx} \approx \frac{1}{N} \sum_{x=0}^{J} x^i exp(\mathbf{b}^\mathbf{T} \boldsymbol{\beta}), \tag{9.3}$$

where J is the number of test items, and l and u are the lower and upper limit of integration. In this case, they are set to be -0.5 and $J + 0.5$, respectively, so that the probabilities of the end points of the discrete distribution are allowed to spread out in both directions. N is the sample size. This approximation holds because the right side of the equation is actually an expression for numerical integration of the left side with equally spaced quadrature points. The numerator and denominator of the left side can be separately expressed as

$$\int_l^u x^i exp(\mathbf{b}^\mathbf{T} \boldsymbol{\beta}) dx \approx \sum_{x=0}^{J} x^i exp(\mathbf{b}^\mathbf{T} \boldsymbol{\beta}),$$

and

$$D = \int_{l}^{u} exp(\mathbf{b}^{\mathrm{T}}\boldsymbol{\beta})dx \approx \sum_{x=0}^{J} exp(\beta_0 + \beta_1 x + \cdots + \beta_M x^M) = N. \tag{9.4}$$

This means that the normalizing constant is approximately equal to the sample size, which is known prior to equating. This result significantly simplifies the computation. The above expressions are very similar to the trapezoidal rule (see Thisted, 1988, p. 264; note that the subinterval length equals 1). The range of the continuous distribution is set from -0.5 to $J + 0.5$ so that in the quadrature the function is evaluated at the midpoints of the subintervals rather than at the end points, as in the regular trapezoidal rule. This range is consistent with the range of the percentile-rank method in conventional equipercentile equating (Kolen & Brennan, 2004, pp. 39–46). Because of the smoothness of the log-linear function, the approximation can be quite close when the number of quadrature points (i.e., the score points J) gets large.

The proposed CLL continuization seems to have several advantages over kernel continuization. First, CLL continuization is simpler and more direct. Second, it is smoother and is guaranteed to be without the small bumpiness in the kernel continuization. Third, it preserves all the moments of the discrete distribution to the precision of equally spaced numerical integration with $J + 1$ quadrature points. The next section illustrates these points with two data sets, one from von Davier et al. (2004b) and the other from Kolen and Brennan (2004).

9.3 The CLL Method for Other Designs

The CLL approach for the equivalent-groups design can be extended to other designs, such as the single-group design, the single-group counterbalanced design, and the nonequivalent groups with anchor test (NEAT) design. Typically, these designs require a bivariate log-linear smoothing procedure in Step 1 of the test equating framework described earlier in this paper. With the Gaussian kernel continuization method, Step 2 is the step that applies the design functions, and Step 3 is the continuization step. With the CLL continuization method, because the continuization step must directly utilize the log-linear function from Step 1, continuization must be carried out immediately after Step 1. So, the design function must be applied after the continuization step and must be applied on continuous distribution functions rather than on discrete distributions, as in the kernel method. Another difference in the design function is that with the kernel method, the design functions are applied to discrete distributions and are thus in matrix form (see von Davier et al., 2004b, Ch. 2 for a detailed description). However, with the CLL method, the design function is a conceptual term that encapsulates the procedures (usually expressed as a set of equations) that transform the continuized distributions into two marginal distributions for X and Y in the target population. The following subsections describe the procedures for various equating designs. For the equivalent-groups design described in the precious section, the design function is an identity function, which means that no such procedure is needed.

9.3.1 For the Single-Group, Counterbalanced Design

For the single-group design, both test X and test Y are administered to the same group of examinees. For the counterbalenced design, the whole group takes both test X and test Y; however, approximately half of the group takes test X first and then test Y, whereas the other half takes test Y first and then test X. The first half group will be labeled as Group 1 and the second half as Group 2. The single-group design can be viewed as a special case of the counterbalanced design where there is only Group 1.

The log-linear functions are taken directly from Step 1 (treating them as continuous functions) and normalized to be PDFs. For Group 1, the PDF can be expressed as Equation 9.5:

$$f_1(x,y) = \frac{1}{D_1} exp(\mathbf{b}^T \boldsymbol{\beta}), \tag{9.5}$$

where $\mathbf{b}^T = (1, x, x^2, \ldots, x^{M_X}, y, y^2, \ldots, y^{M_Y}, xy, x^2y, xy^2, \ldots, x^{C_X}y^{C_Y})$ is a vector of polynomial terms of x and y, $\beta = (\beta_{00}, \beta_{01}, \beta_{02}, \ldots, \beta_{0M_X}, \beta_{10}, \beta_{20}, \ldots, \beta_{M_Y0}, \beta_{11}, \beta_{12}, \beta_{21}, \ldots, \beta_{C_XC_Y})^T$ is a vector of parameters, M_X and M_Y are the orders of marginal polynomial terms for X and Y, C_X and C_Y are the orders of the cross-product terms for X and Y, and D_1 is a normalizing constant that ensures that $f_1(x, y)$ is a PDF. Again, it can be shown that the normalizing constant approximates the sample size.

The joint PDF of Group 2, $f_2(x, y)$, can be found in a similar fashion. Given the weights of X and Y for Group 1, w_X and w_Y, the combined marginal distributions of X and Y can be expressed as follows:

$$f(x) = w_X \int_{l_Y}^{u_Y} f_1(x,y)dy + (1 - w_X) \int_{l_Y}^{u_Y} f_2(x,y)dy, \tag{9.6}$$

$$f(y) = w_y \int_{l_X}^{u_X} f_1(x,y)dx + (1 - w_y) \int_{l_X}^{u_X} f_2(x,y)dx. \tag{9.7}$$

Numerical integration is used in carrying out the necessary integrations. The rest of the equating procedure is the same as for the equivalent-groups design.

9.3.2 For the NEAT Design

For the NEAT design, Group 1 from Population 1 takes test X plus the anchor set V, and Group 2 from Population 2 takes test Y plus the anchor set V. The continuous bivariate PDFs $f_1(x,v)$ for X and V, $f_2(y,v)$ for Y and V can be obtained in a similar fashion as described in the previous section for the counterbalanced design. The NEAT design has essentially two equating methods: the frequency estimation (also called poststratification) and the chained equipercentile equating method. The frequency estimation method is based on the assumption that the conditional distributions of test scores conditioning on an anchor test score remain invariant across populations, which can be expressed as follows:

$$f_2(x|v) = f_1(x|v) = f_1(x,v)/f_1(v), \tag{9.8}$$

$$f_1(y|v) = f_2(y|v) = f_2(y,v)/f_2(v), \tag{9.9}$$

The marginal distributions can be found by the following expressions:

$$f_1(x) = \int_{l_V}^{u_V} f_1(x,v)dv, \tag{9.10}$$

$$f_2(y) = \int_{l_V}^{u_V} f_2(y,v)dv, \tag{9.11}$$

$$f_1(v) = \int_{l_X}^{u_X} f_1(x,v)dx, \tag{9.12}$$

$$f_2(v) = \int_{l_Y}^{u_Y} f_2(y,v)dy, \tag{9.13}$$

With this assumption and given the weight of Population 1 in the target population, w_1, the marginal distributions of X and Y for the target population are

$$f_T(x) = w_1 f_1(x) + (1 - w_1)\int_{l_V}^{u_V} f_1(x|v) f_2(v)dv, \tag{9.14}$$

$$f_T(y) = w_1 \int_{l_V}^{u_V} f_2(y|v) f_1(v)dv + (1 - w_1) f_2(y), \tag{9.15}$$

The rest of the equating procedure is the same as in the equivalent-groups design.

The chained equipercentile equating method first equates X to V using $f_1(x)$ and $f_1(v)$, and then equates the V equivalent X scores to Y using $f_2(v)$ and $f_2(y)$. Given all the continuous marginal distributions in Equations 9.10–9.13, Equation 9.1 must be applied twice to accomplish the chain equipercentile equating procedure.

9.4 Standard Error of Equating for CLL under the Equivalent-Groups Design

von Davier et al. (2004b) derived this general expression for the asymptotic standard error of equating (SEE):

$$SEE_Y(x) = \left\| \hat{J}_{eY} \hat{J}_{DF} C \right\|. \tag{9.16}$$

This expression is decomposed into three parts, each relating to a different stage of the equating process. $\hat{J}_{eY}$ is related to continuization (Step 3) and equating (Step 4). $\hat{J}_{DF}$ is related to the estimation of score probabilities (Step 2). C is related to presmoothing (Step 1). Because the CLL method uses the log-linear function directly in the continuization step, the cumulative distribution functions of test X and test Y depend on the the estimated parameter vectors $\hat{\boldsymbol{\beta}}_X$ and $\hat{\boldsymbol{\beta}}_Y$ of the log-linear models rather than on the estimated score probabilities $\hat{r}$ and $\hat{s}$ in von Davier et al. (2004b). Let F denote the cumulative distribution functions of X and G denote the cumulative distribution functions of Y. The equating function from X to Y can be expressed as

$$e_Y(x) = e_Y(x; \boldsymbol{\beta}_\mathbf{X}, \boldsymbol{\beta}_\mathbf{Y}) = G^{-1}(F(x; \boldsymbol{\beta}_\mathbf{X}); \boldsymbol{\beta}_\mathbf{Y}), \tag{9.17}$$

where

$$F(x; \beta_X) = \frac{\int_l^x exp(\mathbf{b}_\mathbf{X}^T \boldsymbol{\beta}_\mathbf{X}) dt}{\int_l^u exp(\mathbf{b}_\mathbf{X}^T \boldsymbol{\beta}_\mathbf{X}) dt}, \tag{9.18}$$

and

$$G(y; \beta_Y) = \frac{\int_l^y exp(\mathbf{b}_\mathbf{Y}^T \boldsymbol{\beta}_\mathbf{Y}) dt}{\int_l^u exp(\mathbf{b}_\mathbf{Y}^T \boldsymbol{\beta}_\mathbf{Y}) dt}. \tag{9.19}$$

Using the δ-method and following a similar approach as in Holland, King, and Thayer (1989), the square of the SEE can be expressed as

$$\boldsymbol{\sigma}_Y^2(x) = (\boldsymbol{\partial} e_Y)^T \Sigma (\boldsymbol{\partial} e_Y) \tag{9.20}$$

where

$$\Sigma = \begin{bmatrix} \Sigma_{\hat{\beta}_X} & \Sigma_{\hat{\beta}_X \hat{\beta}_Y} \\ \Sigma_{\hat{\beta}_X \hat{\beta}_Y} & \Sigma_{\hat{\beta}_Y} \end{bmatrix}, \tag{9.21}$$

and

$$(\partial e_Y) = \begin{bmatrix} \dfrac{\partial e_Y}{\partial \boldsymbol{\beta}_X} \\ \dfrac{\partial e_Y}{\partial \boldsymbol{\beta}_Y} \end{bmatrix}. \tag{9.22}$$

The elements of Σ are further obtained by the following equations:

$$\boldsymbol{\Sigma}_{\hat{\beta}_X} = (B_X^T \boldsymbol{\Sigma}_{mX} B_X)^{-1}, \tag{9.23}$$

where B_X is the design matrix for X in the log-linear model (see Holland & Thayer, 1987) and

$$\boldsymbol{\Sigma}_{mX} = N(D_{p_X} - p_X p_X^T), \tag{9.24}$$

where p_X is the vector of probabilities in the multinomial categories for Form X and D_{p_X} is a diagonal matrix made from p_X. $\Sigma_{\hat{\beta}_Y}$ can be obtained in a similar fashion. Because X and Y are two independent groups, it follows that the model parameter estimates for the two groups are also independent, which is expressed as

$$\boldsymbol{\Sigma}_{\hat{\beta}_X \hat{\beta}_Y} = \mathbf{0}. \tag{9.25}$$

The elements of (∂e_Y) can be obtained from Equations 9.26 and 9.27:

$$\frac{\partial e_Y}{\partial \boldsymbol{\beta}_{Xi}} = \frac{1}{\frac{\partial G(y;\boldsymbol{\beta}_Y)}{\partial y}} \Big|_{y=e_Y(x)} \frac{\partial F(x;\boldsymbol{\beta}_X)}{\partial \boldsymbol{\beta}_{Xi}} \tag{9.26}$$

$$\frac{\partial e_Y}{\partial \boldsymbol{\beta}_{Yi}} = -\frac{1}{\frac{\partial G(y;\boldsymbol{\beta}_Y)}{\partial y}} \Big|_{y=e_Y(x)} \frac{\partial G(y;\boldsymbol{\beta}_Y)}{\partial \boldsymbol{\beta}_{Yi}} \Big|_{y=e_Y(x)}. \tag{9.27}$$

Given Equations 9.18 and 9.19, the derivatives in Equations 9.26 and 9.27 can be derived straightforwardly. Their expressions can be quite messy and thus are omitted here.

The general expression of SEE in Equation 9.20 applies to all designs. However, for designs other than the equivalent-groups design, calculating expression in Equation 9.22 could be quite complicated, depending on the specific design and equating method, and is beyond the scope of this chapter.

9.5 Illustration With Real Test Data

9.5.1 Comparison of the Continuization Procedures

Because the CLL method performs the continuization step before applying the design function, and the kernel method applies the design function before the continuization step, the two continuization procedures only can be compared directly under the equivalent-groups design where the design function is the

Table 9.1 The Moments and Differences in Moments for the 20-Item Data Set (With Kernel Moments Computed Based on Formula)

Test X	Raw dist.	Log-linear	Kernel			CLL
			0.33	0.622	1.0	
Moments						
Mean	10.8183	10.8183	10.8183	10.8183	10.8183	10.8283
SD	3.8059	3.8059	3.8059	3.8059	3.8059	3.7909
Skewness	0.0026	−0.0649	−0.0641	−0.0623	−0.0587	−0.0502
Kurtosis	2.5322	2.6990	2.6588	2.5604	2.3616	2.6723
Difference in moments with the log-linear discrete distribution						
Mean	-	-	0.0000	0.0000	0.0000	0.0100
SD	-	-	0.0000	0.0000	0.0000	−0.0150
Skewness	-	-	0.0007	0.0025	0.0062	0.0147
Kurtosis	-	-	−0.0401	−0.1386	−0.3373	−0.0267

Note. CLL = continuzed log-linear

identity function and thus can be skipped. This section compares the CLL and kernel continuization methods using two real data sets in terms of the smoothness of the continuous distribution and preservation of moments.

The first data set is taken from von Davier et al. (2004b, Table 9.1) and is a 20-item test data set. Only test form X data are used here. First, the log-linear model is fitted with degree 2 to the raw frequency data. Then the kernel continuization is implemented with three different bandwidth parameter values: $h = 0.33$, $h = 0.622$, and $h = 1.0$. The h value of 0.622 represents the optimal h that minimizes the combined penalty function for this data set. The other two h values are somewhat arbitrary, but with one somewhat smaller than the optimal value and the other somewhat larger than the optimal value.

The CLL distribution is plotted against the kernel distribution in Figure 9.1. The upper part shows that the kernel distributions are very close to the CLL distribution. In fact, the three lines almost coincide with each other, except with $h = 1$ making the kernel distribution depart slightly from the CLL distribution, especially at the ends of the score scale. As discussed previously, this departure reflects a distortion of the shape of the discrete distribution.

The lower part of Figure 9.1 plots the differences between the kernel distributions and the CLL distribution. It can be seen that with h =.622 the kernel distribution still has some bumps,although they are too small to be seen in the upper part of Figure 9.1.(Note that the vertical scales for the upper and lower part of Figure 9.1 are very different.)

The moments for different continuizations for this data set are in Table 9.1. Note that log-linear smoothing with degree 2 maintains the first two moments of the raw score distribution. The moments for the kernel distributions were computed based on the theoretical results in von Davier et al. (2004b), namely, that the first two moments of kernel distribution are the same as the log-linear discrete distribution, but the skewness and kurtosis differ by a factor of $(a_X)^3$ and $(a_X)^4$, respectively. The moments for CLL were empirically computed using numerical integration. For the

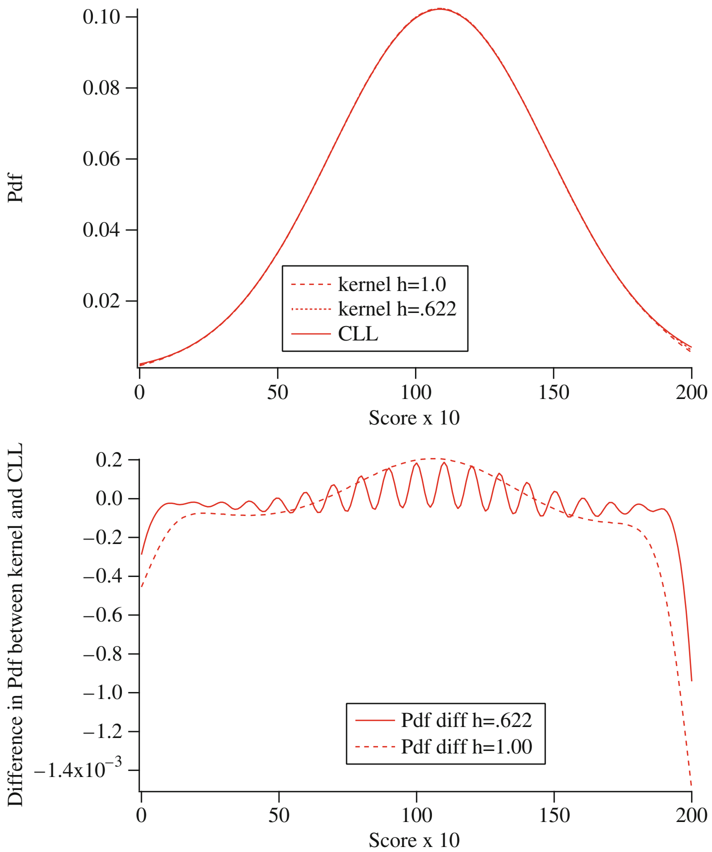

Fig. 9.1 Comparisons of kernel continuization and continuized log-linear (CLL) for the 20-item data set. Pdf = probability density function

kernel method, the case of $h = .33$ can be ignored, since it produced unacceptably large bumps. All CLL moments approximate those of the log-linear distribution reasonably well, whereas the kernel methods have bigger differences in kurtosis. The kernel continuization did not distort the skewness of the distribution, even when a large h was specified, because the skewness of the discrete distribution was very small.

The same analyses were repeated for the 40-item ACT mathematics data in Kolen and Brennan (2004). A log-linear model with a degree of 6 was fitted to the raw frequency. The same kernel and CLL procedures were applied as for the first illustrative example. Three h parameter values were used for this data set: 0.33, 0.597, and 1.0. The value 0.597 represents the optimal h that minimizes the combined penalty function. (It turns out that in both data sets, the second penalty function PEN_2 does not have any effect on the combined penalty because there is no

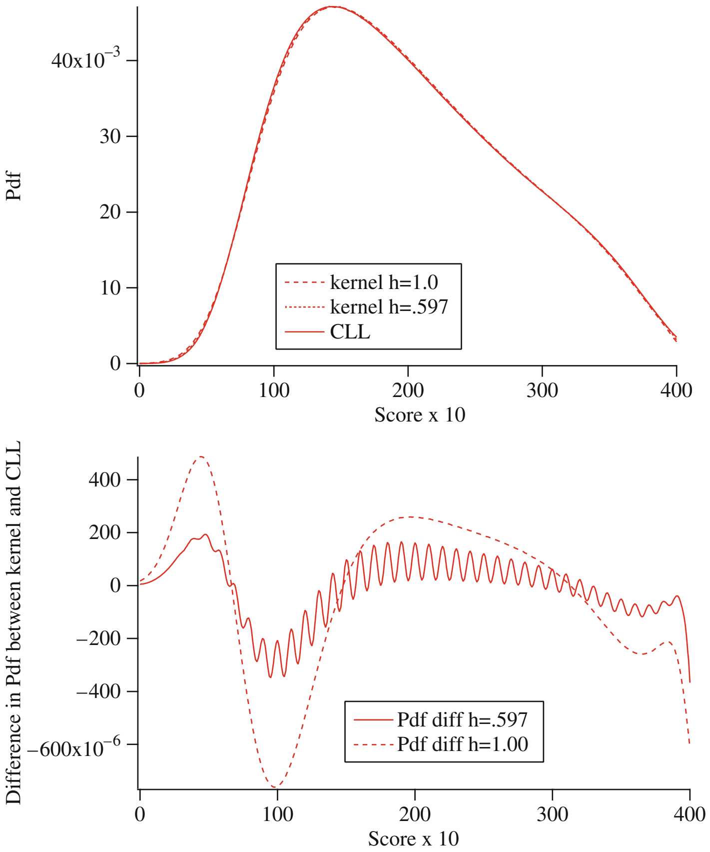

Fig. 9.2 Comparisons of kernel continuization and continuzed log-linear (CLL) for the 40-item data set. Pdf = probability density function

U-shaped distribution around any score point.) Results are plotted in Figure 9.2. These plots show similar patterns of comparisons to those for the 20-item data set in the first example.

The moments of various distributions for this data set are in Table 9.2. The CLL moments are slightly closer to the discrete distribution moments than the kernel moments, although both methods produce very close moments. The CLL method preserves moments better in this case because the number of score points is larger and the approximation in Equation 9.3 is more accurate when the number of score points is larger.

Overall, these two illustrations confirm that the CLL continuization method has certain advantages over the kernel method with respect to simplicity, a smoother continuous distribution, and preserving moments better when the number of score points is relative large and the discrete distributions are highly skewed.

Table 9.2 The Moments and Differences in Moments for the 40-Item Data Set (With Kernel Moments Computed Based on Formula)

Test X	Raw dist.	Log-linear	Kernel			CLL
			0.33	0.622	1.0	
Moments						
Mean	19.8524	19.8524	19.8524	19.8524	19.8524	19.8512
SD	8.2116	8.2116	8.2116	8.2116	8.2116	8.2105
Skewness	0.3753	0.3753	0.3744	0.3722	0.3671	0.3751
Kurtosis	2.3024	2.3024	2.2950	2.2783	2.2356	2.3023
Difference in moments with the log-linear discrete distribution						
Mean	-	-	0.0000	0.0000	0.0000	−0.0012
SD	-	-	0.0000	0.0000	0.0000	−0.0012
Skewness	-	-	−0.0009	−0.0031	−0.0082	−0.0001
Kurtosis	-	-	−0.0074	−0.0241	−0.0668	−0.0001

Note. CLL = continuzed log-linear

9.5.2 Comparisons of Equating Functions

The 40-item test data sets are also used to compare the equating functions under the equivalent-groups design based on three methods: (a) the traditional equipercentile equating method based on percentile ranks, (b) the kernel method, and (c) the CLL method. The optimal *h* parameters were used to compute the kernel continuous distributions. The traditional equipercentile method is also applied to the unsmoothed raw frequency data as a baseline for comparison. The results for the 40-item data set are in Table 9.3. The equating functions and their differences are plotted in Figure 9.3. The results showed that the equating functions based on these three methods were quite similar. Except at the end points of the score scale, the differences were within 0.1.

Another set of real test data with a pairs of test forms was taken from Kolen and Brennan (2004, p. 147) to compare the CLL method with the kernel method under the NEAT design. The test had 36 items with a 12-item internal anchor test. The sample size was 1,655 for the X group and 1,638 for the Y group. A bivariate log-linear smoothing procedure was used for the smoothing step. The frequency estimation method was used for computing the equating function. The frequency estimation method under the NEAT design requires a rather complicated design function. Three continuization and equating methods are computed and compared: (a) the traditional equipercentile equating method based on percentile ranks, (b) the kernel method, and (c) the CLL method. The results are in Table 9.4. The equating functions and their differences are plotted in Figure 9.4. The results showed that the CLL method produces equating results similar to the kernel method but slightly different from the traditional log-linear equipercentile method.

9.5.3 Comparison of SEE Estimates

The SEEs for the CLL method were computed for the 20-item data set using Equation 9.25 and are contained in Table 9.5. The SEEs for the kernel method were also computed and are presented in Table 9.5, which shows that the SEEs for the two methods were very similar.

9.6 Summary

Wang (2008) proposed an alternative continuization method for the test equating framework constructed by von Davier et al. (2004b). With this new continuization method, there are two major differences between the proposed CLL method and the kernel method: (a) The proposed CLL method directly uses the function from the

Table 9.3 The Equating Functions for the 40-Item Data Set Under an Equivalent-Groups Design

Score	Raw equating	Log-linear	kernel (.597)	CLL equating
0	0.0000	−0.4384	−0.7031	−0.4199
1	0.9796	0.1239	0.0537	0.1406
2	1.6462	0.9293	0.9143	0.9664
3	2.2856	1.8264	1.8069	1.8473
4	2.8932	2.7410	2.7072	2.7369
5	3.6205	3.6573	3.6082	3.6300
6	4.4997	4.5710	4.5112	4.5266
7	5.5148	5.4725	5.4191	5.4291
8	6.3124	6.3577	6.3355	6.3411
9	7.2242	7.2731	7.2648	7.2668
10	8.1607	8.2143	8.2119	8.2111
11	9.1827	9.1819	9.1819	9.1792
12	10.1859	10.1790	10.1798	10.1762
13	11.2513	11.2092	11.2101	11.2067
14	12.3896	12.2750	12.2761	12.2734
15	13.3929	13.3764	13.3784	13.3770
16	14.5240	14.5111	14.5147	14.5146
17	15.7169	15.6784	15.6790	15.6801
18	16.8234	16.8638	16.8623	16.8647
19	18.0092	18.0566	18.0541	18.0580
20	19.1647	19.2469	19.2449	19.2497
21	20.3676	20.4262	20.4263	20.4312
22	21.4556	21.5911	21.5916	21.5961
23	22.6871	22.7368	22.7365	22.7404
24	23.9157	23.8595	23.8588	23.8623
25	25.0292	24.9594	24.9586	24.9616
26	26.1612	26.0374	26.0369	26.0394
27	27.2633	27.0954	27.0955	27.0973
28	28.1801	28.1357	28.1364	28.1375
29	29.1424	29.1606	29.1621	29.1625
30	30.1305	30.1729	30.1750	30.1746
31	31.1297	31.1749	31.1777	31.1765
32	32.1357	32.1691	32.1726	32.1705
33	33.0781	33.1576	33.1618	33.1588
34	34.0172	34.1424	34.1470	34.1434
35	35.1016	35.1250	35.1300	35.1257
36	36.2426	36.1064	36.1118	36.1068
37	37.1248	37.0873	37.0929	37.0873
38	38.1321	38.0676	38.0729	38.0670
39	39.0807	39.0462	39.0514	39.0448
40	39.9006	40.0202	40.0256	40.0177

Note. CLL = continuzed log-linear.

log-linear smoothing step and makes it into a PDF, and (b) the continuization step occurs before the design function is applied. The illustration with real test data shows that with a relatively long test length, the CLL method produces smoother continuous score distributions and preserves the moments better than the kernel method. The equating results from the CLL method are quite similar to the kernel

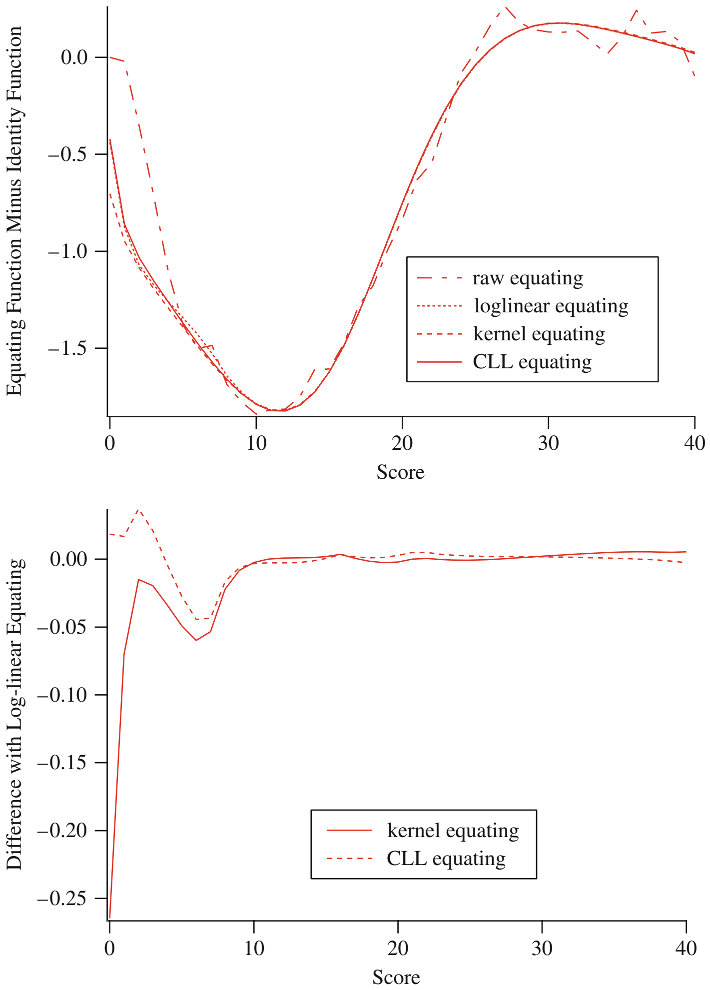

Fig. 9.3 Comparisons of equating functions for the 40-item data set under an equivalent-groups design. CLL = continuzed log-linear

method results under both the equivalent-groups design and the NEAT design. The similarity of the equating results make it difficult to make any recommendation about which method is the best choice under real testing situations. The comparisons are not comprehensive and lack objective criteria to evaluate the equating errors. A more thorough simulation study is needed to compare the kernel and CLL methods in order to make some practical recommendations.

A few differences between the CLL method and the kernel method merit discussion. First, because the CLL method requires that the continuization step occur before the design function is applied, the design function is applied to

Table 9.4 The Frequency-Estimation Equating Functions for the 36-Item Data Set Under a NEAT Design

Score	Log-linear frequency estimation	Kernel frequency estimation	CLL frequency estimation
0	−0.0129	0.0313	0.0059
1	1.0242	1.0827	1.0723
2	2.0988	2.1552	2.1357
3	3.1986	3.2375	3.2256
4	4.3091	4.3239	4.3132
5	5.4200	5.4122	5.4031
6	6.5194	6.5007	6.4929
7	7.5971	7.5880	7.5828
8	8.6759	8.6729	8.6688
9	9.7542	9.7546	9.7515
10	10.8305	10.8322	10.8303
11	11.9035	11.9053	11.9048
12	12.9721	12.9733	12.9738
13	14.0353	14.0357	14.0370
14	15.0924	15.0923	15.0941
15	16.1426	16.1423	16.1447
16	17.1854	17.1854	17.1881
17	18.2200	18.2209	18.2238
18	19.2460	19.2483	19.2513
19	20.2627	20.2668	20.2698
20	21.2694	21.2756	21.2783
21	22.2653	22.2740	22.2770
22	23.2495	23.2612	23.2636
23	24.2209	24.2362	24.2386
24	25.1784	25.1982	25.2004
25	26.1207	26.1464	26.1480
26	27.0466	27.0804	27.0813
27	27.9550	27.9997	27.9998
28	28.8454	28.9046	28.9050
29	29.7179	29.7964	29.7949
30	30.5739	30.6769	30.6760
31	31.4295	31.5496	31.5461
32	32.2939	32.4192	32.4163
33	33.1700	33.2920	33.2908
34	34.0732	34.1764	34.1807
35	35.0130	35.0836	35.1035
36	35.9983	36.0326	36.0800

Note. CLL = continuized log-linear method; NEAT = nonequivalent groups with anchor test.

continuous distributions. This makes the expression of the design function easier to describe and program than with the kernel method. For example, for the frequency estimation method under the NEAT design, the kernel method applies a complicated set of matrix and vector operations in order to estimate the marginal distributions for the target population. For the CLL method, the design function is expressed nicely in Equations 9.8–9.15.

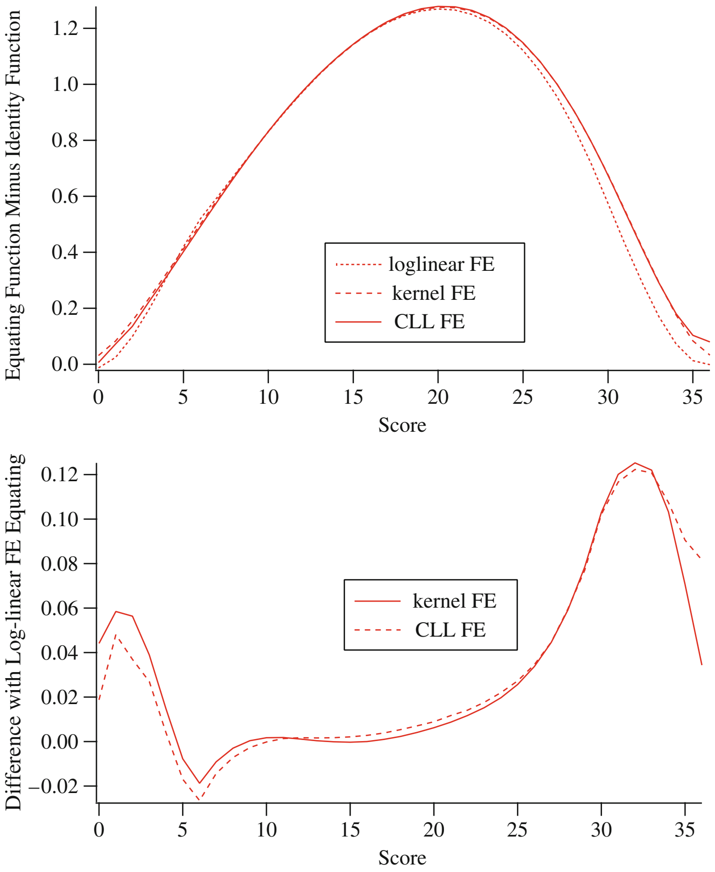

Fig. 9.4 Comparisons of equating functions for the 36-item data set under a nonequivalent groups with anchor test (NEAT) design. CLL = continuzed log-linear; FE = frequency estimation

Seond, the kernel method appears to have closed mathematical form in the continuization and equating steps, whereas the CLL method requires numerical integration. A closer look shows that computing the normal cumulative distribution functions in the kernel method also requires numerical integration or some approximation algorithm. Therefore, computationally speaking, both methods require some kind of numerical method for computation, although the CLL method requires more frequent use of numerical integration.

Finally, the kernel method requires a bandwidth parameter h for the Gaussian kernel. Having this parameter presents advantages and disadvantages. The advantage is that users can manipulate this parameter to achieve some goal. For example, when h is set very large, the kernel method becomes a linear equating method. The disadvantage is that it is rather arbitrary. Although von Davier et. al (2004b) proposed a penalty function to compute the optimal bandwidth, this penalty

Table 9.5 The Standard Errors of Equating (SEEs) for the 20-Item Data Set

Score	Kernel SEE	CLL SEE
0	0.2200	0.2100
1	0.2895	0.2933
2	0.2875	0.2904
3	0.2664	0.2682
4	0.2410	0.2418
5	0.2170	0.2169
6	0.1967	0.1963
7	0.1812	0.1808
8	0.1708	0.1705
9	0.1646	0.1646
10	0.1619	0.1622
11	0.1621	0.1627
12	0.1653	0.1661
13	0.1721	0.1731
14	0.1827	0.1839
15	0.1951	0.1969
16	0.2038	0.2064
17	0.1990	0.2028
18	0.1700	0.1747
19	0.1186	0.1170
20	0.0703	0.0396

function itself is also arbitrary in some sense. The CLL method, on the other hand, does not have such a parameter and thus saves a step in the computation.

The software used to computed the procedures described in this paper is available from the author upon request (tianyouwang@yahoo.com).