ACT Research Report Series 91-5

Method of Moments Estimates for the Four-Parameter Beta Compound Binomial Model and the Calculation of Classification Consistency Indexes

Bradley A. Hanson

September 1991

ACT

For additional copies write:
ACT Research Report Series
P.O. Box 168
Iowa City, Iowa 52243

Method of Moments Estimates for the Four-Parameter Beta Compound Binomial Model and the Calculation of Classification Consistency Indexes

Bradley A. Hanson

American College Testing

September, 1991

Abstract

This paper presents a detailed derivation of method of moments estimates for the four-parameter beta compound binomial strong true score model. A procedure is presented to deal with the case in which the usual method of moments estimates do not exist or result in invalid parameter estimates. The results presented regarding the method of moments estimates are used to derive formulas for computing classification consistency indices under the four-parameter beta compound binomial model.

Acknowledgement. The author thanks Robert L. Brennan for carefully reading an earlier version of this paper and offering helpful suggestions and comments.

The first part of this paper discusses estimation of the four-parameter beta compound binomial model by the method of moments. Most of the material in the first part of this paper is a restatement of material presented in Lord (1964, 1965) with some added details. The second part of this paper uses the results presented in the first part to obtain formulas for computing the classification consistency indexes described by Hanson and Brennan (1990).

The purpose of this paper is to provide a detailed description of the procedures used in computer programs written by the author which compute estimates for the four-parameter beta compound binomial model and indexes of classification consistency based on the four-parameter beta compound binomial model.

Estimation of the Four-parameter Beta Compound Binomial Model

It is assumed that the test score to be modeled is the sum of K dichotomously scored test items (this score is referred to as the raw or observed test score). The probability that the raw score random variable X in the population of interest equals $i\,(i = 0, \dots, K)$, under the four-parameter beta compound binomial model is given by

$$\Pr(X = i) = \int_l^u \Pr(X = i \mid \tau, k)\, g(\tau \mid \alpha, \beta, l, u)\, d\tau\,, \tag{1}$$

where τ is the proportion correct true score. (For simplicity of notation the dependence of the marginal test score distribution $[\Pr(X = i)]$ on the parameters k, α, β, l, and u is not denoted.) The true score distribution $[g(\tau \mid \alpha, \beta, l, u)]$ is assumed to belong to the four-parameter beta family of distributions. The four-parameter beta distribution is a generalization of the usual beta distribution that in addition to the two shape parameters ($\alpha > 0$ and $\beta > 0$) has parameters for the lower (l) and upper (u) limits of the distribution ($0 \le l < u \le 1$). The four-parameter beta density function [defined on the interval (l, u)] is

$$g(\tau \mid \alpha, \beta, l, u) = \frac{(-l + \tau)^{\alpha - 1}\,(u - \tau)^{\beta - 1}}{(u - l)^{\alpha + \beta - 1} B(\alpha, \beta)}\,, \tag{2}$$

where $B(\alpha, \beta)$ is the beta function, which is related to the gamma function $[\Gamma(x)]$ by

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\,\Gamma(\beta)}{\Gamma(\alpha + \beta)}\,.$$

The conditional error distribution $[\Pr(X = i \mid \tau, k)]$ is assumed to be Lord's (1965) two term approximation to the compound binomial distribution. The probability density function of the two-term approximation to the compound binomial distribution is (Lord, 1965, Equation 5)

$$\Pr(X = i \mid \tau, k) = p(i \mid K, \tau) \\ - k\tau(1 - \tau)\big[p(i \mid K - 2, \tau) - 2p(i - 1 \mid K - 2, \tau) + p(i - 2 \mid K - 2, \tau)\big]\,, \quad (3)$$

where $p(i \mid n, \tau) \equiv \Pr(Z = i \mid n, \tau)$ where the distribution of Z is binomial with parameters n (the number of binomial events) and τ (the binomial probability).

The two-term approximation to the compound binomial distribution given by Equation 3 involves the parameter k (note that this is distinct from upper case K which is used in this paper to donate the number of items on the test), in addition to the binomial parameter. When $k = 0$ the conditional distribution of X given τ is binomial. Lord (1965) gives a method of estimating k by setting the theoretical value of the average error variance (or reliability) under the two term approximation of the compound binomial distribution (which is a function of k) equal to an estimate of the average error variance (or reliability). The number-correct true score variance, assuming the conditional error distribution is Lord's two-term approximation to the compound binomial distribution, is (Lord, 1965, Equation 39)

$$\frac{K^2\,\sigma_x^2 - (K - 2k)\,\mu_x\,(K - \mu_x)}{K(K - 1) + 2k}\,, \quad (4)$$

where μ_x and σ_x^2 are the raw score mean and variance. Subtracting Equation 4 from σ_x^2 gives the average error variance (denoted σ_e^2) under the two term approximation of the compound binomial distribution

$$\sigma_e^2 = \frac{\sigma_x^2\big[K(K - 1) + 2k - K^2\big] + (K - 2k)\mu_x(K - \mu_x)}{K(K - 1) + 2k}\,. \quad (5)$$

Solving Equation 5 for k gives

$$k = \frac{K\big[(K-1)(\sigma_x^2 - \sigma_e^2) - K\sigma_x^2 + \mu_x(K-\mu_x)\big]}{2\big[\mu_x(K-\mu_x) - (\sigma_x^2 - \sigma_e^2)\big]} . \tag{6}$$

Given an estimate of the average error variance, σ_e^2, and estimates of μ_x and σ_x^2 in Equation 6, an estimate of k can be calculated. The value of k as a function of the reliability (ρ), raw score mean, and raw score variance is given by substituting $\sigma_x^2(1-\rho)$ for σ_e^2 in Equation 6.

The conditional error distribution given by Equation 3 may not be a probability distribution since for some values of τ and i it is possible that $\Pr(X = i \mid \tau, k) < 0$. Lord (1965) states that for usual values of K and k negative probabilities are typically negligible for $.01 \leq \tau \leq .99$ so that for practical purposes it is appropriate to treat $\Pr(X = i \mid \tau, k)$ given in Equation 3 as a probability distribution.

After a value of k has been determined there are two steps involved in the estimation of the observed score distribution under the four-parameter beta compound binomial model using the method of moments. First, the parameters of the four-parameter beta true score distribution are estimated, and second these parameter estimates are used to produce the estimated observed score distribution.

Estimation of True Score Distribution

Under the assumption that the conditional error distribution is given by the two term approximation to the compound binomial distribution, Lord (1965) produces a formula which gives the non-central moments (moments about zero) of the proportion correct true score distribution in terms of the factorial moments of the observed score distribution. The first moment of the proportion correct true score distribution can be written as

$$\mu_{1\tau} = \frac{1}{K}\mu_{[1]x} , \tag{7}$$

where $\mu_{i\tau}$ is the i-th central moment of the proportion correct true score distribution and $\mu_{[i]x}$ is the i-th factorial moment of the number correct observed score distribution (the

i-th factorial moment of a random variable X is defined as $E[X(X+1)\dots(X+i-1)]$). The i-th ($i \geq 2$) non-central moment of the proportion correct true score distribution ($\mu'_{i\tau}$) is (Lord, 1965, Equation 37)

$$\mu'_{i\tau} = \frac{\frac{\mu_{[i]x}}{(K-2)^{[i-2]}} + ki^{[2]}\mu'_{(i-1)\tau}}{K(K-1) + ki^{[2]}}\,, \tag{8}$$

where for integers j and l, $j^{[l]} = j(j-1)\dots(j-l+1)$. Given estimates of the first four factorial moments of the observed score distribution (these can be obtained from the central or non-central observed score moments, see Kendall & Stuart, 1977, page 66), estimates of the mean and second through fourth non-central moments of the proportion correct true score distribution can be obtained from Equations 7 and 8.

The estimates of the first four true score moments are used to produce method of moments estimates of the parameters of the four-parameter beta true score distribution. The four parameters to be estimated are the two shape parameters (α and β), the lower limit of the distribution (l) and the upper limit of the distribution (u). If the proportion correct true score distribution is a member of the four-parameter beta family then its mean, variance, skewness and kurtosis are given by (Johnson & Kotz, 1970, pages 40-44)

$$\begin{aligned}
\mu'_{1\tau} &= l + (u-l)\left(\frac{\alpha}{\alpha+\beta}\right) \\
\mu_{2\tau} &= (u-l)^2 \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \\
\gamma_3 \equiv \frac{\mu_{3\tau}}{\left(\sqrt{\mu_{2\tau}}\right)^3} &= 2(\beta-\alpha)\frac{\sqrt{\alpha+\beta+1}}{(\alpha+\beta+2)\sqrt{\alpha\beta}} \\
\gamma_4 \equiv \frac{\mu_{4\tau}}{{\mu_{2\tau}}^2} &= 3(\alpha+\beta+1)\frac{2(\alpha+\beta)^2 + \alpha\beta(\alpha+\beta-6)}{\alpha\beta(\alpha+\beta+2)(\alpha+\beta+3)}\,,
\end{aligned} \tag{9}$$

where $\mu_{i\tau}$ is the i-th central moment (moment about the mean) of the proportion correct true score distribution. The central moments of Equation 9 can be written in terms of the non-central moments of Equation 8 (see Kendall & Stuart, 1977, page 58).

Let

$$r = \frac{6(\gamma_4 - \gamma_3^2 - 1)}{(6 + 3\gamma_3^2 - 2\gamma_4)}\,. \tag{10}$$

Substituting the values for γ_3 and γ_4 given in Equation 9 into Equation 10 and simplifying gives $r = \alpha + \beta$. Substituting r for $\alpha + \beta$ in the expression for γ_4 given in Equation 9 and solving for $\alpha\beta$ gives

$$\alpha\beta = \frac{6\,r^2\,(r+1)}{(r+2)\,(r+3)\,\gamma_4 - 3\,(r-6)\,(r+1)}\,. \tag{11}$$

Solving for α using the expressions $\alpha + \beta = r$ and Equation 11 gives the two solutions

$$\alpha = \frac{r}{2}\left(1 \pm \sqrt{1 - \frac{24\,(r+1)}{(r+2)\,(r+3)\,\gamma_4 - 3\,(r-6)\,(r+1)}}\right)\quad . \tag{12}$$

If the solutions given in Equation 12 are real then one of these solutions will be the value of α and the other solution will be the value of β. If the skewness (γ_3) is positive then the larger solution will be the value of β, otherwise the larger solution will be the value of α.

Substituting estimated moments for population moments in Equation 12 gives method of moments estimates of the parameters α and β. Solving the expressions for $\mu'_{1\tau}$ and $\mu_{2\tau}$ in Equation 9 for l and u gives

$$\begin{aligned} l &= \mu'_{1\tau} - \frac{\alpha\sqrt{\mu_{2\tau}\,(\alpha+\beta+1)}}{\sqrt{\alpha\beta}} \\ u &= \mu'_{1\tau} + \frac{\beta\sqrt{\mu_{2\tau}\,(\alpha+\beta+1)}}{\sqrt{\alpha\beta}}\,. \end{aligned} \tag{13}$$

Substituting estimates of $\mu'_{1\tau}$, $\mu_{2\tau}$ and the method of moments estimates of α and β from Equation 12 in Equation 13 gives method of moments estimates of l and u.

The method of moments estimates of the four parameters do not exist when the solutions for α and β given in Equation 12 are not real. Even if solutions exist the estimates of some parameters may not be valid (i.e., $u > 1$, $l < 0$, $\alpha < 0$, $\beta < 0$). When the method of moments solution using the first four moments does not exist or one or more of the estimated parameters is invalid it is suggested an estimation procedure be used in which the first three moments are fit to determine three of the four parameters (α, β and u) and the remaining parameter (l) is chosen such that the kurtosis of the fitted distribution is as near as possible to kurtosis as calculated directly from the observed score moments (using Equation 8). To implement this procedure, method of moments estimates of α, β and u (given a specified value of l) based on the first three moments of the proportion correct true score distribution are needed.

To derive method of moments estimates of α, β and u (given a specified value of l) the following definitions are made

$$\xi_i \equiv \frac{\alpha + i - 1}{\alpha + \beta + i - 1}, \quad i = 1, 2, 3. \tag{14}$$

Using Equation 14, ξ_3 can be written as a function of ξ_1 and ξ_2:

$$\xi_3 = \frac{\xi_1 - 2\,\xi_2 + \xi_1\,\xi_2}{2\,\xi_1 - \xi_2 - 1}. \tag{15}$$

The first two non-central moments of the proportion correct true score distribution can be written as

$$\begin{aligned} \mu'_{1\tau} &= l + (u - l)\,\xi_1 \\ \mu'_{2\tau} &= l^2 + 2\,l\,(u - l)\,\xi_1 + (u - l)^2\,\xi_1\,\xi_2. \end{aligned} \tag{16}$$

Solving Equation 16 for ξ_1 and ξ_2 gives

$$\begin{aligned} \xi_1 &= \frac{\mu'_{1\tau} - l}{u - l} \\ \xi_2 &= \frac{\mu'_{2\tau} - 2\,l\,(\mu'_{1\tau} - l) - l^2}{(\mu'_{1\tau} - l)(u - l)}. \end{aligned} \tag{17}$$

Substituting the expressions for ξ_1 and ξ_2 given in Equation 17 into the expression for ξ_3 given in Equation 15 and simplifying yields

$$(u - l)\,\xi_3 = \frac{-l^3 + 3\mu'_{1\tau}l^2 - (\mu'_{2\tau} + 2\,[\mu'_{1\tau}]^2)\,l + \mu'_{1\tau}\mu'_{2\tau} + (u - l)\,([\mu'_{1\tau}]^2 - 2\mu'_{2\tau} + 2\mu'_{1\tau}l - l^2)}{l^2 - 2\mu'_{1\tau}l + 2(\mu'_{1\tau})^2 - \mu'_{2\tau} + (u - l)\,(l - \mu'_{1\tau})}. \tag{18}$$

The third non-central moment of the proportion correct true score distribution is

$$\mu'_{3\tau} = l^3 + 3l^2(u - l)\,\xi_1 + 3l(u - l)^2\,\xi_1\xi_2 + (u - l)^3\,\xi_1\xi_2\xi_3. \tag{19}$$

Substituting the expressions for ξ_1 and ξ_2 from Equation 17 into Equation 19 and simplifying produces

$$\mu'_{3\tau} = l^3 + 3l^2(\mu'_{1\tau} - l) + 3l\left[\mu'_{2\tau} - l^2 - 2l(\mu'_{1\tau} - l)\right] + \left[\mu'_{2\tau} - l^2 - 2l(\mu'_{1\tau} - l)\right](u - l)\xi_3 \,. \quad (20)$$

The only expression involving u in Equation 20 is $(u - l)\xi_3$. Substituting the right hand side of Equation 18 for $(u - l)\xi_3$ in Equation 20, solving for u and simplifying gives

$$u = \frac{l\left([\mu'_{1\tau}]^2\mu'_{2\tau} - 2[\mu'_{2\tau}]^2 + \mu'_{1\tau}\mu'_{3\tau}\right) + \mu'_{1\tau}(\mu'_{2\tau})^2 - 2(\mu'_{1\tau})^2\mu'_{3\tau} + \mu'_{2\tau}\mu'_{3\tau}}{l\left(2[\mu'_{1\tau}]^3 - 3\mu'_{1\tau}\mu'_{2\tau} + \mu'_{3\tau}\right) + 2(\mu'_{2\tau})^2 - (\mu'_{1\tau})^2\mu'_{2\tau} - \mu'_{1\tau}\mu'_{3\tau}} \,. \quad (21)$$

Substituting estimated moments in Equation 21 in place of the population moments gives a method of moments estimate of u (l is assumed to be specified).

Let τ^* be a transformation of the proportion correct true score given by

$$\tau^* = \frac{\tau - l}{u - l} \,. \quad (22)$$

Since τ has a four-parameter beta distribution, τ^* has a two-parameter beta distribution with parameter α and β. The parameters α and β are given in terms of the non-central moments of the distribution of τ^* as

$$\begin{aligned} \alpha &= \frac{\mu'_{1\tau^*}(\mu'_{1\tau^*} - \mu'_{2\tau^*})}{\mu'_{2\tau^*} - (\mu'_{1\tau^*})^2} \\ \beta &= \frac{(1 - \mu'_{1\tau^*})(\mu'_{1\tau^*} - \mu'_{2\tau^*})}{\mu'_{2\tau^*} - (\mu'_{1\tau^*})^2} \,. \end{aligned} \quad (23)$$

Given values of u and l, the non-central moments of the distribution of τ^* can be obtained from the non-central moments of the distribution of τ. Substituting estimated moments of the distribution of τ^* in Equation 23 in place of the population moments gives method of moments estimates of α and β.

Given a value of l, Equations 21 and 23 can be used to obtain method of moments estimates of u, α, and β. If the method of moments estimates of l, u, α and β given by Equations 12 and 13 do not exist or are not valid, then a solution is selected such that the first three moments are fit (which determines u, α, and β) and l is chosen such that the fitted true score kurtosis is as close as possible to the true score kurtosis as calculated directly from the observed score moments (using Equation 8).

The expressions for α and β given in Equation 23 can be written in terms of l and the first three non-central moments of the proportion correct true score distribution ($\mu'_{1\tau}$, $\mu'_{2\tau}$ and $\mu'_{3\tau}$). These values of α and β can be used in the expression for the kurtosis given in Equation 9 to compute the fitted true score kurtosis as a function of l. The function of l given by the squared difference of fitted true score kurtosis (a function of l) and the true score kurtosis calculated directly from the observed score moments (using Equation 8) is to be minimized. The minimization is constrained to be over those values of l for which valid method of moments estimates of u, α and β using Equations 21 and 23 exist.

A method of finding the value of l that solves this constrained optimization problem is to first compute the two solutions in which l and u are on the boundary of the parameter space (the solution for which $l = 0$ and the solution for which $u = 1$). Either or both of these solutions may not be a valid solution. The solution for which $l = 0$ can be computed using Equations 21 and 23 . The solution for $u = 1$ can be computed by solving Equation 21 for l giving

$$l = \frac{u\left([\mu'_{1\tau}]^2\mu'_{2\tau} - 2[\mu'_{2\tau}]^2 + \mu'_{1\tau}\mu'_{3\tau}\right) + \mu'_{1\tau}(\mu'_{2\tau})^2 - 2(\mu'_{1\tau})^2\mu'_{3\tau} + \mu'_{2\tau}\mu'_{3\tau}}{u\left(2[\mu'_{1\tau}]^3 - 3\mu'_{1\tau}\mu'_{2\tau} + \mu'_{3\tau}\right) + 2(\mu'_{2\tau})^2 - (\mu'_{1\tau})^2\mu'_{2\tau} - \mu'_{1\tau}\mu'_{3\tau}} \,. \tag{24}$$

Substituting $u = 1$ in Equation 24 gives a method of moments estimate of l. This estimate of l along with $u = 1$ are used in Equation 23 to give estimates of α and β. For each of these two solutions (the solution for which $l = 0$ and the solution for which $u = 1$) the fitted kurtosis is calculated (assuming the solutions are valid). The solution with the smallest squared difference in the fitted kurtosis and the kurtosis calculated using Equation 8 (this squared difference will be referred to as the squared kurtosis difference) is used as the initial solution. A grid search (Thisted, 1988, page 200) is then used for values of $l > 0$ to search for a solution with a smaller squared kurtosis difference than the initial solution. For almost all the situations that have been encountered in practice either the solution for which $l = 0$ or the solution for which $u = 1$ produces the smallest squared kurtosis difference. There are examples, though, in which a solution with $l > 0$ and $u < 1$ can give a smaller squared kurtosis difference than either the solution for which $l = 0$ or the solution for which $u = 1$.

Estimation of Observed Score Distribution

This section discusses estimation of the observed score distribution under the four-parameter beta compound binomial model assuming a value of k has been determined and estimates of α, β, l, and u have been calculated. The derivations presented in this section closely follow those presented in Lord (1964) with some added details. These derivations and the resulting formula for computing the observed score probabilities are not presented in Lord (1965).

Case 1: $k = 0$. The case is first considered in which the conditional error distribution $[\Pr(X = i \mid \tau, k)]$ in Equation 3 is binomial ($k = 0$). In this case the observed score distribution is

$$\Pr(X = i) = \int_l^u \binom{K}{i} \tau^i (1-\tau)^{K-i} \left[\frac{(-l+\tau)^{\alpha-1}(u-\tau)^{\beta-1}}{(u-l)^{\alpha+\beta-1} B(\alpha,\beta)} \right] d\tau \,. \qquad (25)$$

Let τ^* be defined as in Equation 22 so that $\tau = (u-l)\tau^* + l$ and $d\tau = (u-l)d\tau^*$. Making the change of variable from τ to τ^* in Equation 25 gives

$$\Pr(X = i) = \frac{\binom{K}{i}}{B(\alpha,\beta)} \int_0^1 \left[(u-l)\tau^* + l\right]^i \left[1 - (u-l)\tau^* - l\right]^{K-i} \cdot \left[\frac{\left[(u-l)\tau^*\right]^{\alpha-1} \left[(u-l)(1-\tau^*)\right]^{\beta-1}}{(u-l)^{\alpha+\beta-1}} \right] (u-l)\, d\tau^* \,. \qquad (26)$$

Simplifying the expression of the true score density times $u - l$ in Equation 26 and substituting $(1-u) + (u-l)(1-\tau^*)$ for $1 - (u-l)\tau^* - l$ gives

$$\Pr(X = i) = \frac{\binom{K}{i}}{B(\alpha,\beta)} \int_0^1 \left[(u-l)\tau^* + l\right]^i \left[(1-u) + (u-l)(1-\tau^*)\right]^{K-i} \cdot (\tau^*)^{\alpha-1} (1-\tau^*)^{\beta-1}\, d\tau^* \,. \quad (27)$$

Using the binomial theorem gives the following two equations

$$\left[(u-l)\,\tau^* + l\right]^i = \sum_{r=0}^{i} \binom{i}{r} \left[(u-l)\,\tau^*\right]^r l^{i-r}\,, \tag{28}$$

and

$$\left[(1-u) + (u-l)\,(1-\tau^*)\right]^{K-i} = \sum_{s=0}^{K-i} \binom{K-i}{s} \left[(u-l)\,(1-\tau^*)\right]^s (1-u)^{K-i-s}\,. \tag{29}$$

Substituting Equations 28 and 29 into Equation 27 gives

$$\begin{aligned}
\Pr(X=i) &= \frac{\binom{K}{i}}{B(\alpha,\beta)} \int_0^1 \sum_{r=0}^{i} \binom{i}{r} \left[(u-l)\,\tau^*\right]^r l^{i-r} \\
&\quad \cdot \sum_{s=0}^{K-i} \binom{K-i}{s} \left[(u-l)\,(1-\tau^*)\right]^s (1-u)^{K-i-s} \left[(\tau^*)^{\alpha-1}\,(1-\tau^*)^{\beta-1}\right] d\tau^* \\
&= \frac{\binom{K}{i}}{B(\alpha,\beta)} \sum_{r=0}^{i} \sum_{s=0}^{K-i} \binom{i}{r} \binom{K-i}{s} l^{i-r} (1-u)^{K-i-s} (u-l)^{s+r} \\
&\quad \cdot \int_0^1 (\tau^*)^{r+\alpha-1}\,(1-\tau^*)^{s+\beta-1}\, d\tau^*\,.
\end{aligned} \tag{30}$$

Using the definition of the Beta function, the integral in Equation 30 can be written as

$$\int_0^1 (\tau^*)^{r+\alpha-1}\,(1-\tau^*)^{s+\beta-1}\, d\tau^* = B(\alpha+r, \beta+s) = \frac{\Gamma(\alpha+r)\Gamma(\beta+s)}{\Gamma(\alpha+\beta+r+s)}\,.$$

Consequently, Equation 30 can be written as

$$\begin{aligned}
\Pr(X=i) &= \sum_{r=0}^{i} \sum_{s=0}^{K-i} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\,\Gamma(\beta)}\, \frac{\Gamma(\alpha+r)\Gamma(\beta+s)}{\Gamma(\alpha+\beta+r+s)} \\
&\quad \cdot \binom{K}{i} \binom{i}{r} \binom{K-i}{s} l^{i-r} (1-u)^{K-i-s} (u-l)^{s+r}\,.
\end{aligned} \tag{31}$$

The gamma functions on the right side of Equation 31 can be written as

$$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\,\Gamma(\beta)}\, \frac{\Gamma(\alpha+r)\,\Gamma(\beta+s)}{\Gamma(\alpha+\beta+r+s)} = \frac{(\alpha)_r\,(\beta)_s}{(\alpha+\beta)_{r+s}}\,, \tag{32}$$

where $(x)_j = x(x+1)\ldots(x+j-1)$ and $(x)_0 \equiv 1$. Substituting Equation 32 in Equation 31 gives

$$\Pr(X=i) = \binom{K}{i} \sum_{r=0}^{i} \sum_{s=0}^{K-i} \binom{i}{r} \binom{K-i}{s} \frac{(\alpha)_r\,(\beta)_s}{(\alpha+\beta)_{r+s}} l^{i-r} (1-u)^{K-i-s} (u-l)^{s+r}\,. \tag{33}$$

Let $r' = i - r$ and $s' = K - i - s$. Changing the index of the first summation in Equation 33 from r to r' and the index of the second summation from s to s' produces

$$\begin{aligned}
\Pr(X = i) &= \binom{K}{i} \sum_{r'=0}^{i} \sum_{s'=0}^{K-i} \binom{i}{r'} \binom{K-i}{s'} \frac{(\alpha)_{i-r'} (\beta)_{K-i-s'}}{(\alpha+\beta)_{K-s'-r'}} l^{r'} (1-u)^{s'} (u-l)^{K-r'-s'} \\
&= (u-l)^K \sum_{r'=0}^{i} \sum_{s'=0}^{K-i} \binom{K}{i} \binom{i}{r'} \binom{K-i}{s'} \\
&\qquad \cdot \frac{(\alpha)_{i-r'} (\beta)_{K-i-s'}}{(\alpha+\beta)_{K-s'-r'}} \left[\frac{l}{u-l}\right]^{r'} \left[\frac{1-u}{u-l}\right]^{s'} \qquad (34)
\end{aligned}$$

The terms in Equation 34 involving combinations can be written as

$$\begin{aligned}
\binom{K}{i}\binom{i}{r'}\binom{K-i}{s'} &= \frac{K!}{i!\,(K-i)!} \frac{i!}{r'!\,(i-r')!} \frac{(K-i)!}{s'!\,(K-i-s')!} \\
&= \frac{K!}{r'!\,s'!} \frac{1}{(i-r')!\,(K-i-s')!} \\
&= \frac{(K-r')!\,(K-s')!}{K!} \frac{K!}{r'!\,(K-r')!} \frac{K!}{s'!\,(K-s')!} \frac{1}{(i-r')!\,(K-i-s')!} \\
&= \frac{(K-r')!\,(K-s')!}{K!} \binom{K}{r'} \binom{K}{s'} \frac{1}{(i-r')!\,(K-i-s')!} \qquad (35)
\end{aligned}$$

Substituting Equation 35 into 34 and rearranging the terms gives

$$\begin{aligned}
\Pr(X = i) = (u-l)^K \sum_{r'=0}^{i} \sum_{s'=0}^{K-i} &\left[\frac{(\alpha)_{i-r'}}{(i-r')!}\right] \left[\binom{K}{r'} \left(\frac{l}{u-l}\right)^{r'}\right] \\
&\cdot \left[\frac{(K-r')!\,(K-s')!}{K!\,(\alpha+\beta)_{K-r'-s'}}\right] \left[\binom{K}{s'} \left(\frac{1-u}{u-l}\right)^{s'}\right] \left[\frac{(\beta)_{K-i-s'}}{(K-i-s')!}\right] \qquad (36)
\end{aligned}$$

Lord (1964) suggests using Equation 36 to calculate the observed score distribution when $k = 0$. Equation 36 can be calculated as the diagonal terms of a matrix product of five $(K+1) \times (K+1)$ matrices $(M_1, M_2, \ldots, M_5)$ with the matrices based on the bracketed terms in Equation 36. The lower triangular matrix M_1 is given by

$$M_1 = \begin{pmatrix}
\frac{(\alpha)_0}{0!} & 0 & 0 & \cdots & 0 \\
\frac{(\alpha)_1}{1!} & \frac{(\alpha)_0}{0!} & 0 & \cdots & 0 \\
\frac{(\alpha)_2}{2!} & \frac{(\alpha)_1}{1!} & \frac{(\alpha)_0}{0!} & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\frac{(\alpha)_K}{K!} & \frac{(\alpha)_{K-1}}{(K-1)!} & \frac{(\alpha)_{K-2}}{(K-2)!} & \cdots & \frac{(\alpha)_0}{0!}
\end{pmatrix}.$$

The diagonal matrix M_2 is given by

$$M_2 = \begin{pmatrix} \binom{K}{0}\left(\frac{l}{u-l}\right)^0 & 0 & 0 & \cdots & 0 \\ 0 & \binom{K}{1}\left(\frac{l}{u-l}\right)^1 & 0 & \cdots & 0 \\ 0 & 0 & \binom{K}{2}\left(\frac{l}{u-l}\right)^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \binom{K}{K}\left(\frac{l}{u-l}\right)^K \end{pmatrix}.$$

The matrix M_3 is given by

$$M_3 = \begin{pmatrix} \frac{K!K!}{K!(\alpha+\beta)_K} & \frac{K!(K-1)!}{K!(\alpha+\beta)_{K-1}} & \frac{K!(K-2)!}{K!(\alpha+\beta)_{K-2}} & \cdots & \frac{K!0!}{K!(\alpha+\beta)_0} \\ \frac{(K-1)!K!}{K!(\alpha+\beta)_{K-1}} & \frac{(K-1)!(K-1)!}{K!(\alpha+\beta)_{K-2}} & \frac{(K-1)!(K-2)!}{K!(\alpha+\beta)_{K-3}} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{2!K!}{K!(\alpha+\beta)_2} & \frac{2!(K-1)!}{K!(\alpha+\beta)_1} & \frac{2!(K-2)!}{K!(\alpha+\beta)_0} & \cdots & 0 \\ \frac{1!K!}{K!(\alpha+\beta)_1} & \frac{1!(K-1)!}{K!(\alpha+\beta)_0} & 0 & \cdots & 0 \\ \frac{0!K!}{K!(\alpha+\beta)_0} & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

The diagonal matrix M_4 is given by

$$M_4 = \begin{pmatrix} \binom{K}{0}\left(\frac{1-u}{u-l}\right)^0 & 0 & 0 & \cdots & 0 \\ 0 & \binom{K}{1}\left(\frac{1-u}{u-l}\right)^1 & 0 & \cdots & 0 \\ 0 & 0 & \binom{K}{2}\left(\frac{1-u}{u-l}\right)^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \binom{K}{K}\left(\frac{1-u}{u-l}\right)^K \end{pmatrix}.$$

The matrix M_5 is given by

$$M_5 = \begin{pmatrix} \frac{(\beta)_K}{K!} & \frac{(\beta)_{K-1}}{(K-1)!} & \frac{(\beta)_{K-2}}{(K-2)!} & \cdots & \frac{(\beta)_0}{0!} \\ \frac{(\beta)_{K-1}}{(K-1)!} & \frac{(\beta)_{K-2}}{(K-2)!} & \frac{(\beta)_{K-3}}{(K-3)!} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{(\beta)_2}{2!} & \frac{(\beta)_1}{1!} & \frac{(\beta)_0}{0!} & \cdots & 0 \\ \frac{(\beta)_1}{1!} & \frac{(\beta)_0}{0!} & 0 & \cdots & 0 \\ \frac{(\beta)_0}{0!} & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

The elements of each of the matrices M_i can be computed by first calculating an easily computed initial element and computing the other elements as a simple factor times an adjacent element. In addition, not all elements of the matrices need to be separately calculated. For example, for matrix M_1 only the first column needs to be computed because the numbers in the other columns are a subset of the numbers in the first column. The fitted probability for raw score i is the $(i+1)$-th diagonal element of the matrix product $M_1 M_2 M_3 M_4 M_5$ times $(u-l)^K$.

Case 2: $k > 0$. When $k > 0$, substituting the conditional error distribution given in Equation 3 into Equation 1 gives the observed score distribution as

$$\Pr(X = i) = p_0(i) - \int_l^u k\,\tau(1-\tau)\big[p(i \mid K-2,\tau) - 2p(i-1 \mid K-2,\tau) \\ + p(i-2 \mid K-2,\tau)\big]\,g(\tau \mid \alpha,\beta,l,u)\,d\tau\,, \quad (37)$$

where $p_0(i)$ is the probability given in Equation 25 of raw score i when the conditional error distribution is binomial, $p(i \mid n,\tau)$ is a binomial probability as defined in Equation 3, and $g(\tau \mid \alpha,\beta,l,u)$ is the four-parameter beta true score distribution given by Equation 2. The terms $\tau(1-\tau)\,p(j \mid K-2,\tau)$ in Equation 37 can be written as

$$\begin{aligned} \tau(1-\tau)\,p(j \mid K-2,\tau) &= \binom{K-2}{j}\tau^{j+1}(1-\tau)^{K-1-j} \\ &= \frac{(K-2)!}{j!(K-2-j)!}\,\tau^{j+1}(1-\tau)^{K-1-j} \\ &= \frac{(j+1)(K-1-j)}{K(K-1)}\,\frac{K!}{(j+1)!(K-1-j)!}\,\tau^{j+1}(1-\tau)^{K-1-j} \\ &= \frac{(j+1)(K-1-j)}{K(K-1)}\,p(j+1 \mid K,\tau) \end{aligned} \quad (38)$$

Substituting the result in Equation 38 into 37 produces

$$\Pr(X = i) = p_0(i) - \frac{k}{K(K-1)}\int_l^u \big[(i+1)\,(K-1-i)\,p(i+1 \mid K,\tau) \\ -2i(K-i)\,p(i \mid K,\tau) + (i-1)\,(K+1-i)\,p(i-1 \mid K,\tau)\big]\,g(\tau \mid \alpha,\beta,l,u)\,d\tau. \quad (39)$$

Define $p_0^*(j)$ as

$$p_0^*(j) \equiv j(K-j)\,p_0(j)\,. \quad (40)$$

Using Equation 40, Equation 39 can be written as

$$\Pr(X = i) = p_0(i) - \frac{k}{K(K-1)}\big[p_0^*(i-1) - 2p_0^*(i) + p_0^*(i+1)\big]\,. \quad (41)$$

The procedure for calculating the observed score distribution for a value of $k > 0$, say k_0, is to first calculate the observed score distribution assuming $k = 0$ using Equation 36 (producing $p_0(i), i = 0, \ldots, K$) and then to use Equation 41 to calculate the observed score distribution for $k = k_0$.

When $k > 0$ some of the probabilities computed using Equation 41 may be negative. When negative values are computed they are usually very small in magnitude.

Classification Consistency Indexes

In this section the results presented previously will be used to obtain formulas for computing the classification consistency indexes described by Hanson and Brennan (1990) assuming the four-parameter beta compound binomial model holds for the test score in question. The two types of classification consistency indexes discussed by Hanson and Brennan (1990) are considered separately. First, the calculation of the probability of a consistent decision (coefficient p) and coefficient κ is discussed. Next, the calculation of the false negative and false positive error rates is discussed.

Probability of a Consistent Decision and Coefficient Kappa

The probability of a consistent decision (coefficient p) is defined in terms of the bivariate distribution of scores on two independent administrations of a test. Assuming the test score random variables on two independent administrations of a test (denoted X_1 and X_2) follow the four-parameter beta compound binomial model the probability of a randomly chosen examinee obtaining raw scores i and j on the two test administrations is (using the notation of Equation 1)

$$\Pr(X_1 = i, X_2 = j) = \int_l^u \Pr(X_1 = i \mid \tau, k) \Pr(X_2 = j \mid \tau, k) g(\tau \mid \alpha, \beta, l, u)\, d\tau \,. \qquad (42)$$

The classification consistency index p is defined as

$$p \equiv \sum_{i=0}^{x_0-1} \sum_{j=0}^{x_0-1} \Pr(X_1 = i, X_2 = j) + \sum_{i=x_0}^{K} \sum_{j=x_0}^{K} \Pr(X_1 = i, X_2 = j)\,, \qquad (43)$$

where x_0 is the raw score cut point.

Coefficient κ is defined as

$$\kappa \equiv \frac{p - p_c}{1 - p_c}, \tag{44}$$

where p_c is the probability of consistent classification by chance given by

$$p_c \equiv \sum_{i=0}^{x_0-1} \sum_{j=0}^{x_0-1} \Pr(X_1 = i) \Pr(X_2 = j) + \sum_{i=x_0}^{K} \sum_{j=x_0}^{K} \Pr(X_1 = i) \Pr(X_2 = j) .$$

The bivariate distribution of scores on two independent administrations of a test given by Equation 42 can be computed based on the conditional error and true score distributions estimated from a single test administration. Thus, the first step in computing the bivariate distribution given in Equation 42 is to estimate the parameters of the four-parameter beta true score distribution using data from a single test administration as previously described. After the parameters of the true score distribution have been estimated, the bivariate distribution given in Equation 42 can be computed.

<u>Case 1: $k = 0$.</u> For computing the integral in Equation 42 the case is first considered in which the conditional error distributions $[Pr(X_1 = i \mid \tau, k)$ and $Pr(X_2 = j \mid \tau, k)]$ are binomial ($k = 0$). In this case Equation 42 can be written as

$$\begin{aligned}
\Pr(X_1 = i, X_2 = j) &= \int_l^u \Pr(X_1 = i \mid \tau, k = 0) \Pr(X_2 = j \mid \tau, k = 0)\, g(\tau \mid \alpha, \beta, l, u)\, d\tau \\
&= \int_l^u \binom{K}{i} \tau^i (1 - \tau)^{K-i} \binom{K}{j} \tau^j (1 - \tau)^{K-j} g(\tau \mid \alpha, \beta, l, u)\, d\tau \\
&= \int_l^u \binom{K}{i} \binom{K}{j} \tau^{i+j} (1 - \tau)^{2K-i-j} g(\tau \mid \alpha, \beta, l, u)\, d\tau \\
&= \frac{\binom{K}{i}\binom{K}{j}}{\binom{2K}{i+j}} \int_l^u \binom{2K}{i+j} \tau^{i+j} (1 - \tau)^{2K-i-j} g(\tau \mid \alpha, \beta, l, u)\, d\tau
\end{aligned} \tag{45}$$

The integral on the last line of Equation 45 is the probability of a test score of $i+j$ on a test with $2K$ items where the test score distribution is assumed to follow the four-parameter beta binomial model with the same true score distribution as the test scores X_1 and X_2. This integral can be calculated using Equation 36 (using the estimated parameters of the true score distribution computed from a single test administration).

Case 2: $k > 0$. When $k > 0$, substituting the conditional error distribution given in Equation 3 into Equation 42 gives the bivariate distribution of scores on two independent administrations of a test as

$$\begin{aligned}
\Pr(X_1 = i\,, X_2 = j) = \int_l^u & \Big\{ p(i \mid K, \tau) \\
& - k\,\tau(1-\tau)\big[p(i \mid K-2, \tau) - 2p(i-1 \mid K-2, \tau) + p(i-2 \mid K-2, \tau)\big]\Big\} \\
& \cdot \Big\{ p(j \mid K, \tau) - k\,\tau(1-\tau)\big[p(j \mid K-2, \tau) - 2p(j-1 \mid K-2, \tau) + p(j-2 \mid K-2, \tau)\big]\Big\} \\
& \cdot g(\tau \mid \alpha, \beta, l, u)\, d\tau\,, \qquad (46)
\end{aligned}$$

where $p(i \mid n, \tau)$ is a binomial probability as defined in Equation 3. Using the equality

$$\tau(1-\tau)\, p(j \mid K-2, \tau) = \frac{(j+1)(K-1-j)}{K(K-1)}\, p(j+1 \mid K, \tau)$$

from Equation 38 and expanding the terms in Equation 46 gives

$$\begin{aligned}
\Pr(X_1 = i\,, X_2 = j) = \int_l^u & \Big\{ p(i \mid K, \tau)\, p(j \mid K, \tau) \\
& + \frac{k}{K(K-1)} \Big[-(i+1)\,(K-i-1)\, p(i+1 \mid K, \tau)\, p(j \mid K, \tau) \\
& \quad -(i-1)\,(K-i+1)\, p(i-1 \mid K, \tau)\, p(j \mid K, \tau) \\
& \quad -(j+1)\,(K-j-1)\, p(i \mid K, \tau)\, p(j+1 \mid K, \tau) \\
& \quad -(j-1)\,(K-j+1)\, p(i \mid K, \tau)\, p(j-1 \mid K, \tau) \\
& \quad +2\left[(K-i)\,i + (K-j)\,j\right] p(i \mid K, \tau)\, p(j \mid K, \tau) \Big] \\
& + \frac{k^2}{K^2(K-1)^2} \Big[(i+1)\,(j+1)\,(K-i-1)\,(K-j-1)\, p(i+1 \mid K, \tau)\, p(j+1 \mid K, \tau) \\
& \quad +(i-1)\,(j+1)\,(K-i+1)\,(K-j-1)\, p(i-1 \mid K, \tau)\, p(j+1 \mid K, \tau) \\
& \quad +(i+1)\,(j-1)\,(K-i-1)\,(K-j+1)\, p(i+1 \mid K, \tau)\, p(j-1 \mid K, \tau) \\
& \quad +(i-1)\,(j-1)\,(K-i+1)\,(K-j+1)\, p(i-1 \mid K, \tau)\, p(j-1 \mid K, \tau) \\
& \quad -2\,i\,(j+1)\,(K-i)\,(K-j-1)\, p(i \mid K, \tau)\, p(j+1 \mid K, \tau) \\
& \quad -2\,i\,(j-1)\,(K-i)\,(K-j+1)\, p(i \mid K, \tau)\, p(j-1 \mid K, \tau)
\end{aligned}$$

$$-2\,(i+1)\,j\,(K-i-1)\,(K-j)\,p(i+1 \mid K,\tau)\,p(j \mid K,\tau)$$
$$-2\,(i-1)\,j\,(K-i+1)\,(K-j)\,p(i-1 \mid K,\tau)\,p(j \mid K,\tau)$$
$$+4\,i\,j\,(K-i)\,(K-j)\,p(i \mid K,\tau)\,p(j \mid K,\tau)\Big]\Big\}\, g(\tau \mid \alpha,\beta,l,u)\,d\tau\,, \quad (47)$$

Using Equation 45, the integral in Equation 47 can be evaluated to give

$$\begin{aligned}
\Pr(X_1 = i\,, X_2 = j) &= p_0(i,j) \\
&+\frac{k}{K(K-1)}\Bigg\{\frac{k}{K(K-1)}\left[(i-1)\,(j-1)\,(K-i+1)\,(K-j+1)\right]p_0(i-1,j-1) \\
&\qquad +(i-1)\,(K-i+1)\left[\frac{-2k}{(K-1)K}\,j\,(K-j)-1\right]p_0(i-1,j) \\
&\quad +\frac{k}{K(K-1)}\left[(i-1)\,(j+1)\,(K-i+1)\,(K-j-1)\right]p_0(i-1,j+1) \\
&\qquad +(j-1)\,(K-j+1)\left[\frac{-2k}{(K-1)K}\,i\,(K-i)-1\right]p_0(i,j-1) \\
&\quad +\Big\{2\left[(K-i)\,i+(K-j)\,j\right]+\frac{k}{K(K-1)}\left[4ij\,(K-i)\,(K-j)\right]\Big\}\,p_0(i,j) \\
&\qquad +(j+1)\,(K-j-1)\left[\frac{-2k}{(K-1)K}\,i\,(K-i)-1\right]p_0(i,j+1) \\
&\quad +\frac{k}{K(K-1)}\left[(i+1)\,(j-1)\,(K-i-1)\,(K-j+1)\right]p_0(i+1,j-1) \\
&\qquad +(i+1)\,(K-i-1)\left[\frac{-2k}{(K-1)K}\,j\,(K-j)-1\right]p_0(i+1,j) \\
&\qquad\quad +\frac{k}{K(K-1)}(i+1)\,(j+1)\,(K-i-1)\,(K-j-1)\,p_0(i+1,j+1)\Bigg\}\,, \quad (48)
\end{aligned}$$

where $p_0(i,j)$ is the probability of scores i and j on two independent administrations of the test when the conditional error distribution is binomial (given by Equation 45).

When $k > 0$ Equation 45 is used to first calculate the values of $p_0(i,j)$, $i,j = 0,\ldots,K$ which are used in Equation 48 to calculate the bivariate distribution of scores on two independent administrations of the test from which coefficient p and coefficient κ can be computed.

False Positive and False Negative Error Rates

The false positive error rate is given by

$$\sum_{i=x_0}^{K} \int_{l}^{\tau_0} \Pr(X = i \mid \tau, k)\, g(\tau \mid \alpha, \beta, l, u)\, d\tau \; , \tag{49}$$

where τ_0 is the true score cutoff. The false negative error rate is given by

$$\sum_{i=0}^{x_0-1} \int_{\tau_0}^{u} \Pr(X = i \mid \tau, k)\, g(\tau \mid \alpha, \beta, l, u)\, d\tau \; , \tag{50}$$

Case 1: $k = 0$. The case is first considered in which the conditional error distribution is binomial ($k = 0$). Based on a series of steps analogous to those which produced Equation 30 from Equation 25 each term in the sum given in Equation 49 can be written as

$$\int_{l}^{\tau_0} \Pr(X = i \mid \tau, k=0)\, g(\tau \mid \alpha, \beta, l, u)\, d\tau$$
$$= \frac{\binom{K}{i}}{B(\alpha,\beta)} \sum_{r=0}^{i} \sum_{s=0}^{K-i} \binom{i}{r} \binom{K-i}{s} l^{i-r} (1-u)^{K-i-s} (u-l)^{s+r}$$
$$\cdot \int_{0}^{\tau_0^*} (\tau^*)^{r+\alpha-1} (1-\tau^*)^{s+\beta-1}\, d\tau^* , \tag{51}$$

where τ^* is defined as in Equation 22 and $\tau_0^* = (\tau_0 - l)/(u - l)$. Using the definition of the incomplete beta function $[I_x(a,b)]$

$$I_x(a,b) \equiv \frac{1}{B(a,b)} \int_0^x t^{a-1} (1-t)^{b-1}\, dt \, ,$$

Equation 51 can be written as (analogous to Equation 31)

$$\int_{l}^{\tau_0} \Pr(X = i \mid \tau, k=0)\, g(\tau \mid \alpha, \beta, l, u)\, d\tau = \sum_{r=0}^{i} \sum_{s=0}^{K-i} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\,\Gamma(\beta)} \frac{\Gamma(\alpha+r)\Gamma(\beta+s)}{\Gamma(\alpha+\beta+r+s)}$$
$$\cdot I_{\tau_0^*}(\alpha+r, \beta+s) \binom{K}{i} \binom{i}{r} \binom{K-i}{s} l^{i-r} (1-u)^{K-i-s} (u-l)^{s+r} . \tag{52}$$

In a manner analogous to that by which Equation 36 was obtained from Equation 31, Equation 52 can be rewritten as

$$\int_{l}^{\tau_0} \Pr(X = i \mid \tau, k=0)\, g(\tau \mid \alpha, \beta, l, u)\, d\tau = (u-l)^K \sum_{r'=0}^{i} \sum_{s'=0}^{K-i} \left[\frac{(\alpha)_{i-r'}}{(i-r')!} \right.$$

$$\cdot \left[\binom{K}{r'} \left(\frac{l}{u-l} \right)^{r'} \right] \left[\frac{(K-r')!\,(K-s')!}{K!\,(\alpha+\beta)_{K-r'-s'}} \right] I_{\tau_0^*}(\alpha+i-r', \beta+K-i-s')$$

$$\cdot \left[\binom{K}{s'} \left(\frac{1-u}{u-l} \right)^{s'} \right] \left[\frac{(\beta)_{K-i-s'}}{(K-i-s')!} \right] \quad (53)$$

Equation 53 can be used to calculate the terms in the sum of Equation 49 to give the false positive rate when $k = 0$.

Case 2: $k > 0$. When $k > 0$ an adjustment formula analogous to Equation 41 can be used. If z_i is the value of the integral in Equation 53 when $k = 0$ then the value of the integral when $k > 0$ is

$$z_i - \frac{k}{K(K-1)} \left[(i-1)(K-i+1)\, z_{i-1} - 2i(K-i)\, z_i + (i+1)(K-i-1)\, z_{i+1} \right]. \quad (54)$$

The false negative error rate can be calculated using the values given by Equation 53 (or when $k > 0$ Equation 54) since

$$\int_l^u \Pr(X = i \mid \tau, k)\, g(\tau \mid \alpha, \beta, l, u)\, d\tau =$$
$$\int_l^{\tau_0} \Pr(X = i \mid \tau, k)\, g(\tau \mid \alpha, \beta, l, u)\, d\tau + \int_{\tau_0}^u \Pr(X = i \mid \tau, k)\, g(\tau \mid \alpha, \beta, l, u)\, d\tau\,,$$

so that

$$\int_{\tau_0}^u \Pr(X = i \mid \tau, k)\, g(\tau \mid \alpha, \beta, l, u)\, d\tau = 1 - \int_l^{\tau_0} \Pr(X = i \mid \tau, k)\, g(\tau \mid \alpha, \beta, l, u)\, d\tau\,. \quad (55)$$

Values of the false positive and false negative error rates are computed based on the values from Equation 53 (or Equation 54 when $k > 0$) and 55, respectively.

Computer Programs

A set of functions written in the C language implementing the procedures presented in this paper for producing method of moments estimates of the true score distribution and observed score distribution under the four-parameter beta compound binomial model and computing classification consistency indexes are available from the author. These functions can be compiled on any computer with an ANSI C compiler.

Macintosh programs which use these routines are also available from the author. One program calculates the method of moments estimates of the four-parameter beta compound binomial model, displays the results in a text window and produces graphic displays of the fitted distributions. This program also computes fitted raw score distributions based on a log-linear model (Holland and Thayer, 1987). Another Macintosh program computes and displays the classification consistency indexes described for three models: the beta binomial, the four-parameter beta binomial, and the four-parameter beta compound binomial. This program also displays the fitted models graphically.

References

Hanson, B. A. & Brennan, R. L. (1990). An investigation of classification consistency indexes estimated under alternative strong true score models. *Journal of Educational Measurement, 27*, 345-359.

Holland, P. W. & Thayer, D. T. (1987). *Notes on the use of log-linear models for fitting discrete probability distributions.* Educational Testing Service Research Report 87-31. Princeton, NJ: Educational Testing Service.

Johnson N. L., & Kotz, S. (1970). *Continuous Univariate Distributions-2.* New York: Houghton Mifflin.

Kendall, M., & Stuart, A. (1977). *The advanced theory of statistics* (4th ed., Vol. 1). New York: MacMillan.

Lord, F. M. (1964). *A strong true-score theory, with applications.* Educational Testing Service Research Bulletin 64-19. Princeton, NJ: Educational Testing Service.

Lord, F. M. (1965). A strong true score theory, with applications. *Psychometrika, 30*, 239-270.

Thisted, R. A. (1988). *Elements of statistical computing: Numerical computation.* New York: Chapman and Hall.