

# Determining the Significance of Estimated Signed and Unsigned Areas Between Two Item Response Functions

Nambury S. Raju  
Illinois Institute of Technology

Asymptotic sampling distributions (means and variances) of estimated signed and unsigned areas between two item response functions (IRFs) are presented for the Rasch model, the two-parameter model, and the three-parameter model with fixed lower asymptotes. In item bias or differential item functioning research, it may be of interest to determine whether the estimated signed and unsigned areas between IRFs calibrated

with two different groups are significantly different from 0. The usefulness of these sampling distributions in this context is discussed and illustrated. More empirical research with the proposed significance tests is necessary. *Index terms: asymptotic mean and variance, differential item functioning, item bias, item response functions, item response theory.*

Raju (1988) presented formulas for computing the exact signed and unsigned areas between two item response functions (IRFs) for the one-, two-, and three-parameter logistic models. He also showed that when the lower asymptotes are not equal for the three-parameter model, the exact area between two IRFs was infinite.

Raju's formulas are functions of item parameters. In practice, only estimates of item parameters are available, hence the areas based on these estimates are subject to sampling fluctuations. Without some notion of the degree of variability to be expected due to sampling, it is impossible to determine whether two IRFs differ only in terms of sampling error. This is an important question for researchers who use the estimated area between two IRFs as a measure of item bias or differential item functioning. Therefore, a major purpose of this paper is to present the asymptotic sampling distributions (means and variances) of Raju's signed and unsigned area estimates for the one- and two-parameter models, and for the three-parameter model when the lower asymptotes are fixed. The other purpose of this paper is to illustrate their applicability in item bias research.

## Background and Notation

Let  $F_1(\theta)$  and  $F_2(\theta)$  represent three-parameter IRFs for the same item for Group 1 and Group 2, respectively. Let

$$F_1 \equiv F_1(\theta) = c_1 + (1 - c_1)P_1 \quad (1)$$

---

APPLIED PSYCHOLOGICAL MEASUREMENT

Vol. 14, No. 2, June 1990, pp. 197-207

© Copyright 1990 Applied Psychological Measurement Inc.

0146-6216/90/020197-11\$1.80

$$F_2 \equiv F_2(\theta) = c_2 + (1 - c_2)P_2 \quad , \quad (2)$$

where

$$P_1 \equiv P_1(\theta) = \frac{\exp[Da_1(\theta - b_1)]}{1 + \exp[Da_1(\theta - b_1)]} \quad , \quad (3)$$

$$P_2 \equiv P_2(\theta) = \frac{\exp[Da_2(\theta - b_2)]}{1 + \exp[Da_2(\theta - b_2)]} \quad , \quad (4)$$

$a_i$ ,  $b_i$ , and  $c_i$  are the three item parameters for the  $i$ th IRF, and  $D$  is a scaling constant, usually set to 1.7 to equate the logistic ogive approximately to the normal ogive (Lord, 1980). Note that Equations 1 and 2 reduce to two-parameter IRFs when  $c_i = 0$  and to one-parameter (Rasch) IRFs when, in addition,  $D$  is absorbed into  $a_i$  and  $a_i = 1$ . Furthermore, define the signed and unsigned areas between  $F_1$  and  $F_2$  with estimated  $\theta$ s and item parameters as

$$\text{Signed Area} = SA_k = \int_{-\infty}^{\infty} (\hat{F}_1 - \hat{F}_2) d\theta \quad (5)$$

$$\text{Unsigned Area} = UA_k = \int_{-\infty}^{\infty} |\hat{F}_1 - \hat{F}_2| d\theta \quad , \quad (6)$$

where  $k$  refers to the item response theory (IRT) model and

$$l = \begin{cases} 0 & (a_i \text{ are equal or unequal, signed index}) \\ 1 & (a_i \text{ are equal, unsigned index}) \\ 2 & (a_i \text{ are unequal, unsigned index}) \end{cases} \quad (7)$$

The following formulas are used in the development of the asymptotic mean and variance of the estimated unsigned (absolute) area between two IRFs. If  $X$  is a continuous random variable with probability density function  $f(x)$ , the expectation of the absolute value of  $X$  can be written as

$$E(|X|) = \int_{-\infty}^{\infty} |x| f(x) dx = -\mu(X) + 2 \int_0^{\infty} x f(x) dx \quad . \quad (8)$$

If  $X$  is normally distributed with  $\mu(X)$  and  $\sigma(X)$ , then the expectation of  $|X|$  can be written, after simplification, as

$$E(|X|) = \mu(|X|) = \mu(X)[1 - 2\Phi(z_0)] + (2/\pi)^{1/2}\sigma(X) \exp(-z_0^2/2) \quad (9)$$

(Johnson & Kotz, 1970), where

$$\Phi(z_0) = \int_{-\infty}^{z_0} g(z) dz \quad , \quad (10)$$

$z_0 = [0 - \mu(X)]/\sigma(X)$ , and  $g(z)$  is a normal probability density function with  $\mu(Z) = 0$  and  $\sigma(Z) = 1$ .

If  $\mu(X) = 0$ , Equation 9 can be rewritten as

$$\mu(|X|) = (2/\pi)^{1/2}\sigma(X) \quad . \quad (11)$$

Similarly, the variance of  $|X|$  can be written as

$$\sigma^2(|X|) = E(|X|^2) - [E(|X|)]^2 = E(X^2) - [E(|X|)]^2 = \sigma^2(X) + \mu^2(X) - \mu^2(|X|) \quad . \quad (12)$$

Also, if  $\mu(X) = 0$ , Equation 12 can be rewritten as

$$\sigma^2(|X|) = \sigma^2(X) - \mu^2(|X|) = \sigma^2(X)(1 - 2/\pi) \quad . \quad (13)$$

### One-Parameter (Rasch) Model

According to Raju (1988), the estimated signed and unsigned areas between two IRFs in the Rasch model can be written as

$$SA_{10} = \hat{b}_2 - \hat{b}_1 \quad (14)$$

$$UA_{11} = |\hat{b}_2 - \hat{b}_1| \quad (15)$$

Because  $\hat{b}_i$  converges asymptotically to  $b_i$  (Wright & Stone, 1979), the asymptotic mean and variance of  $SA_{10}$  can be expressed as

$$\mu(SA_{10}) = E(\hat{b}_2) - E(\hat{b}_1) = b_2 - b_1 \quad (16)$$

$$\sigma^2(SA_{10}) = \text{Var}(\hat{b}_2) + \text{Var}(\hat{b}_1) \quad (17)$$

(Note that because  $\hat{b}_1$  and  $\hat{b}_2$  are based on two independent samples, the covariance term is not needed in Equation 17.) In the Rasch model, the asymptotic variance of  $\hat{b}_i$  can be expressed as

$$\text{Var}(\hat{b}_i) = \left[ \sum_{j=1}^{N_i} P_i(\theta_j) Q_i(\theta_j) \right]^{-1} \quad (18)$$

(Wright & Stone, 1979), where  $N_i$  is the number of examinees in group  $i$ ,  $\theta_j$  is the trait level of examinee  $j$ , and  $Q_i = 1 - P_i$ . In practice, Equation 18 is evaluated with estimated item parameters and  $\theta$ s. Letting  $X = \hat{b}_2 - \hat{b}_1$  and assuming that  $X$  is normally distributed, the mean and variance of  $UA_{11}$  can be written, using Equations 9 and 12, as

$$\mu(UA_{11}) = \mu(SA_{10})[1 - 2\Phi(z_0)] + (2/\pi)^{1/2}\sigma(SA_{10}) \exp(-z_0^2/2) \quad (19)$$

where  $z_0 = [0 - \mu(SA_{10})]/\sigma(SA_{10})$ , and

$$\sigma^2(UA_{11}) = \sigma^2(SA_{10}) + \mu^2(SA_{10}) - \mu^2(UA_{11}) \quad (20)$$

Equations 19 and 20 can be further simplified if  $\mu(SA_{10}) = 0$ .

### Two-Parameter Model

According to Raju (1988), the signed and unsigned areas between 2 two-parameter IRFs with estimated item parameters can be expressed as

$$\text{Signed Area } (SA_{20}) = \hat{b}_2 - \hat{b}_1 \quad (21)$$

$$\text{Unsigned Area } (UA_{21}) = |\hat{b}_2 - \hat{b}_1| \text{ when } \hat{a}_1 = \hat{a}_2 \quad (22)$$

$$\text{Unsigned Area } (UA_{22}) = \left| \frac{2(\hat{a}_2 - \hat{a}_1)}{D\hat{a}_1\hat{a}_2} \ln \left\{ 1 + \exp \left[ \frac{D\hat{a}_1\hat{a}_2(\hat{b}_2 - \hat{b}_1)}{\hat{a}_2 - \hat{a}_1} \right] \right\} - (\hat{b}_2 - \hat{b}_1) \right| \text{ when } \hat{a}_1 \neq \hat{a}_2 \quad (23)$$

#### Mean and Variance of $SA_{20}$

Because  $\hat{b}_i$  converges asymptotically to  $b_i$  (Lord, 1980), the asymptotic mean of  $SA_{20}$  can be written as

$$\mu(SA_{20}) = E(SA_{20}) = E(\hat{b}_2) - E(\hat{b}_1) = b_2 - b_1 \quad (24)$$

Because  $\hat{b}_1$  and  $\hat{b}_2$  are obtained from two independent samples, the variance of  $SA_{20}$  can be expressed as

$$\sigma^2(SA_{20}) = \text{Var}(\hat{b}_2 - \hat{b}_1) = \text{Var}(\hat{b}_2) + \text{Var}(\hat{b}_1) \quad (25)$$

In the two-parameter model, the asymptotic variance of  $\hat{b}_i$  can be written as

$$\text{Var}(\hat{b}_i) = \frac{I_{a_i}}{I_{a_i}I_{b_i} - I_{a_i b_i}^2} \quad (26)$$

(Hulin, Drasgow, & Parsons, 1983; Lord, 1980), where

$$I_{a_i} = D^2 \sum_{j=1}^{N_i} (\theta_j - b_i)^2 P_i(\theta_j) Q_i(\theta_j) \quad (27)$$

$$I_{b_i} = D^2 a_i^2 \sum_{j=1}^{N_i} P_i(\theta_j) Q_i(\theta_j) \quad (28)$$

$$I_{a_i b_i} = D^2 a_i \sum_{j=1}^{N_i} (\theta_j - b_i) P_i(\theta_j) Q_i(\theta_j) \quad (29)$$

Even though the mean and variance of  $SA_{20}$  involve true item and  $\theta$  parameters, they are typically evaluated with estimated item and  $\theta$  parameters.

#### Mean and Variance of $UA_{21}$ When $\hat{a}_1 = \hat{a}_2$

Let  $X = \hat{b}_2 - \hat{b}_1$ . Assuming that  $X$  is normally distributed, and using Equations 9 and 12, the mean and variance of  $|X|$ , which is the same as  $UA_{21}$ , can be written as

$$\mu(|X|) = \mu(UA_{21}) = \mu(SA_{20})[1 - 2\Phi(z_0)] + (2/\pi)^{1/2} \sigma(SA_{20}) \exp(-z_0^2/2) \quad (30)$$

where  $z_0 = [0 - \mu(SA_{20})]/\sigma(SA_{20})$ , and

$$\sigma^2(|X|) = \sigma^2(UA_{21}) = \sigma^2(SA_{20}) + \mu^2(SA_{20}) - \mu^2(UA_{21}) \quad (31)$$

because  $X = SA_{20}$  and  $|X| = UA_{21}$ . Equations 30 and 31 can be readily evaluated using the equations for the mean and variance of  $SA_{20}$ .

#### Mean and Variance of $UA_{22}$ When $\hat{a}_1 \neq \hat{a}_2$

First, let

$$H = \frac{2(\hat{a}_2 - \hat{a}_1)}{D\hat{a}_1\hat{a}_2} \ln \left\{ 1 + \exp \left[ \frac{D\hat{a}_1\hat{a}_2(\hat{b}_2 - \hat{b}_1)}{\hat{a}_2 - \hat{a}_1} \right] \right\} - (\hat{b}_2 - \hat{b}_1) \quad (32)$$

where  $|H| = UA_{22}$ . Then, using the Taylor series approximation (or the delta method), the mean and variance of  $H$  can be approximated by

$$E(H) = \mu(H) = \frac{2(a_2 - a_1)}{Da_1a_2} \ln \left\{ 1 + \exp \left[ \frac{Da_1a_2(b_2 - b_1)}{a_2 - a_1} \right] \right\} - (b_2 - b_1) \quad (33)$$

$$\text{Var}(H) = \sigma^2(H) = B_1^2 \text{Var}(\hat{b}_1) + B_2^2 \text{Var}(\hat{b}_2) + A_1^2 \text{Var}(\hat{a}_1) + A_2^2 \text{Var}(\hat{a}_2) + 2B_1A_1 \text{Cov}(\hat{b}_1, \hat{a}_1) + 2B_2A_2 \text{Cov}(\hat{b}_2, \hat{a}_2) \quad (34)$$

(Kendall & Stuart, 1977, pp. 246–247), where, after simplification,

$$B_1 = \frac{\partial H}{\partial b_1} = 1 - \frac{2 \exp(Y)}{1 + \exp(Y)} \quad (35)$$

$$B_2 = \frac{\partial H}{\partial b_2} = -B_1 \quad (36)$$

$$A_1 = \frac{\partial H}{\partial a_1} = \frac{2}{a_1^2} \left\{ \frac{a_1 a_2 (b_2 - b_1)}{a_2 - a_1} \frac{\exp(Y)}{1 + \exp(Y)} - \frac{\ln[1 + \exp(Y)]}{D} \right\} \quad (37)$$

$$A_2 = \frac{\partial H}{\partial a_2} = \frac{-2}{a_2^2} \left\{ \frac{a_1 a_2 (b_2 - b_1)}{a_2 - a_1} \frac{\exp(Y)}{1 + \exp(Y)} - \frac{\ln[1 + \exp(Y)]}{D} \right\} = -\frac{a_1^2}{a_2^2} A_1 \quad (38)$$

and

$$Y = \frac{D a_1 a_2 (b_2 - b_1)}{a_2 - a_1} \quad (39)$$

The variance of  $\hat{b}_i$  needed in Equation 34 is defined in Equation 26. Following Lord (1980) and Hulin et al. (1983), the variance of  $\hat{a}_i$  can be written as

$$\text{Var}(\hat{a}_i) = \frac{I_{b_i}}{I_{a_i} I_{b_i} - I_{a_i b_i}^2} \quad (40)$$

Furthermore, the covariance between  $\hat{a}_i$  and  $\hat{b}_i$  can be written as

$$\text{Cov}(\hat{a}_i, \hat{b}_i) = \frac{-I_{a_i b_i}}{I_{a_i} I_{b_i} - I_{a_i b_i}^2} \quad (41)$$

Several points should be made about the mean and variance of  $H$  given in Equations 33 and 34:

1. It is assumed that  $E(\hat{b}_i) = b_i$  and  $E(\hat{a}_i) = a_i$ . The asymptotic convergence of  $\hat{b}_i$  and  $\hat{a}_i$  to  $b_i$  and  $a_i$ , respectively, is generally accepted in the IRT literature (Hambleton & Swaminathan, 1985; Lord, 1980).
2. In order to compute the mean and variance of  $H$ , as previously noted,  $\hat{b}_i$ ,  $\hat{a}_i$ , and  $\hat{\theta}_i$  need to be substituted for  $b_i$ ,  $a_i$ , and  $\theta_i$ , respectively.
3. Equation 34 for the variance of  $H$  does not contain the covariances between  $\hat{b}_1$  and  $\hat{a}_2$ ,  $\hat{b}_1$  and  $\hat{b}_2$ ,  $\hat{a}_1$  and  $\hat{b}_2$ , and  $\hat{a}_1$  and  $\hat{a}_2$  because the item parameters involved in these pairs are based on two independent samples.

If it is assumed that  $H$  is normally distributed, then the mean and variance of  $UA_{22}$ , which is identical to  $|H|$ , can be written as

$$\mu(UA_{22}) = \mu(|H|) = \mu(H)[1 - 2\Phi(z_0)] + (2/\pi)^{1/2} \sigma(H) \exp(-z_0^2/2) \quad (42)$$

where  $z_0 = [0 - \mu(H)]/\sigma(H)$ , and

$$\sigma^2(UA_{22}) = \sigma^2(H) + \mu^2(H) - \mu^2(|H|) \quad (43)$$

### Three-Parameter Model

Because the area between two IRFs is infinite when the lower asymptotes are not equal for the three-parameter model, only the special case in which  $\hat{c}_1 = \hat{c}_2 = c$  will be considered here. According to Raju (1988), the signed and unsigned areas for this special case are given by

$$SA_{30} = (1 - c)SA_{20} \quad (44)$$

$$UA_{31} = (1 - c)UA_{21} \quad (\hat{a}_1 = \hat{a}_2) \quad (45)$$

$$UA_{32} = (1 - c)UA_{22} \quad (\hat{a}_1 \neq \hat{a}_2) \quad (46)$$

Assuming that  $c$  is a constant, the means and variances associated with the expressions given in Equations 44–46 can be written as

$$\mu(SA_{30}) = (1 - c)\mu(SA_{20}) \quad (47)$$

$$\sigma^2(SA_{30}) = (1 - c)^2\sigma^2(SA_{20}) \quad (48)$$

$$\mu(UA_{31}) = (1 - c)\mu(UA_{21}) \quad (49)$$

$$\sigma^2(UA_{31}) = (1 - c)^2\sigma^2(UA_{21}) \quad (50)$$

$$\mu(UA_{32}) = (1 - c)\mu(UA_{22}) \quad (51)$$

$$\sigma^2(UA_{32}) = (1 - c)^2\sigma^2(UA_{22}) \quad (52)$$

These means and variances can be computed readily using the equations given for the two-parameter model, provided the expressions for  $I$  given in Equations 27–29 are redefined as follows:

$$I_b = \frac{D^2 a_i^2}{(1 - c)^2} \sum_{j=1}^{N_i} \left\{ [F_i(\theta_j) - c]^2 \frac{1 - F_i(\theta_j)}{F_i(\theta_j)} \right\} \quad (53)$$

$$I_a = \frac{D^2}{(1 - c)^2} \sum_{j=1}^{N_i} \left\{ (\theta_j - b_i)^2 [F_i(\theta_j) - c]^2 \frac{1 - F_i(\theta_j)}{F_i(\theta_j)} \right\} \quad (54)$$

$$I_{a,b} = \frac{D^2 a_i}{(1 - c)^2} \sum_{j=1}^{N_i} \left\{ (\theta_j - b_i) [F_i(\theta_j) - c]^2 \frac{1 - F_i(\theta_j)}{F_i(\theta_j)} \right\} \quad (55)$$

(Lord, 1980, p. 191). These quantities are needed for estimating  $\text{Var}(\hat{b}_i)$ ,  $\text{Var}(\hat{a}_i)$ , and  $\text{Cov}(\hat{a}_i, \hat{b}_i)$  using Equations 26, 40, and 41, respectively, when  $c \neq 0$ . Equations 53–55 are typically evaluated with estimated item and  $\theta$  parameters.

## Applications

### Significance Test for the Signed Areas

Assuming that the observed  $SA$ s are normally distributed, the mean and variance equations given above can be used in assessing the significance of an observed area between two IRFs. To establish whether an observed signed area is significantly different from 0, compute

$$Z = \frac{SA - 0}{\sigma(SA)} \quad (56)$$

and then determine if  $-z < Z < z$ , where  $z$  is the tabled value from the unit normal distribution (e.g.,  $z = 1.96$  for a two-tailed test when  $\alpha = .05$ ). If the observed  $Z$  is outside the limits, then the null hypothesis that  $SA = 0$  is erroneously rejected with  $p = .05$  when the null hypothesis is, in fact, true. Because IRT analyses typically involve large samples (500 or more) and the standard deviation of  $SA$  is sample-size dependent, it may be desirable to use a value of  $z$  (e.g.,  $\pm 3$ ) corresponding to a very small value of  $\alpha$ . The  $SA$  in Equation 56 is written without subscripts to indicate that this general expression is valid for signed areas in all three models.

### Significance Test for the Unsigned Areas

Equation 56 is not useful for establishing the significance of observed unsigned areas because the assumption of normality will not be valid for unsigned areas. Because the distribution of  $H$  in Equation

32 is assumed to be normal in the development of the mean and variance formulas for  $|H|$ , the distribution of  $|H|$  is, by definition, a (top) *half-normal* distribution. The same reasoning is also valid for  $|\hat{b}_2 - \hat{b}_1|$ .

One solution to the problem of non-normality with unsigned areas is to consider the statistical significance of an observed  $H$  rather than  $|H|$ . That is, let

$$Z = \frac{H-0}{\sigma(H)} \tag{57}$$

and then determine if  $-z < Z < z$ , where  $z$  is the tabled value from the unit normal distribution. The rationale for this recommendation is that the distribution of  $|H|$  is a (top) half-normal distribution and, therefore, an observed  $|H|$  will fall above  $z\sigma(H)$  only  $\alpha$  percent of the time under the null hypothesis that  $H$  is normally distributed with  $E(H) = 0$ . This recommendation is equally suitable for the unsigned area given by  $|\hat{b}_2 - \hat{b}_1|$ . It may be desirable to use a value of  $z$  corresponding to a very small  $\alpha$  level, because, as previously noted, most IRT analyses typically involve large samples.

### An Example

To illustrate the procedures described above, two subsamples were selected from the Fall standardization sample of the SRA Survey of Basic Skills (SBS; Science Research Associates, 1985). Examinees for both samples consisted of students in Grades 4 through 6 who took the vocabulary test of Level 34, Form P of the SBS. Sample 1 consisted of 1,000 White examinees and Sample 2 consisted of 1,000 Black examinees. The vocabulary test has 40 multiple-choice items with four responses per item. The test was analyzed with the LOGIST program (Wingersky, Barton, & Lord, 1983), using the three-parameter model with the  $c$  parameter fixed at .20.

The number-correct score summary for the White and Black examinees is given in Table 1. The mean score of 26.98 for the White group is more than seven points greater than the mean for the Black group. The standard deviations and KR-20 estimates of reliability are about the same for both groups.

The item parameter estimates from the LOGIST run are presented in Table 2 for both groups. Also given in this table are the signed and unsigned areas ( $SA$  and  $UA$ ) and  $H$ , the signed portion of  $UA$ . Following the procedures described in Cook and Eignor (1983), the item parameter estimates ( $as$  and  $bs$  only) from the Black group were linearly transformed so that the transformed item parameter estimates were on the same scale as the item parameter estimates for the White group. This transformation was necessary because the item parameters were separately calibrated for the White and Black groups. The transformed  $a$  and  $b$  parameters for the Black group were used in the subsequent computations of the signed and unsigned areas and their sampling distributions. The additive and multiplicative constants associated with the linear transformation were  $-.93$  and  $1.13$ , respectively. Note that there is no need to transform the  $c$  parameters (Cook & Eignor, 1983).

Table 1  
 Number-Correct Score Mean and Standard  
 Deviation for the White and Black Examinees

Statistic	White Group	Black Group
Mean	26.98	19.44
SD	8.47	8.38
KR-20	.91	.90
N	1,000	1,000

Table 2  
 Item Parameter Estimates for the White Group, Original and Transformed Parameter Estimates for the Black Group, and Area Measures

Item No.	White Group			Black Group (Original)			Black Group (Transformed)			Signed Area	Unsigned Area	H
	a	b	c	a	b	c	a	b	c			
1	.50	-2.67	.20	.69	-.65	.20	.61	-1.67	.20	.80	.80	.80
2	.67	-1.37	.20	.89	-.65	.20	.78	-1.67	.20	-.24	.26	.26
3	.91	-1.59	.20	.86	-.65	.20	.76	-1.67	.20	-.06	.15	-.15
4	.58	-1.51	.20	.45	-.66	.20	.40	-1.68	.20	-.14	.53	-.53
5	1.05	-.67	.20	1.24	.30	.20	1.09	-.59	.20	.06	.06	.06
6	.79	-.79	.20	.72	.06	.20	.63	-.86	.20	-.06	.21	-.21
7	1.04	-.25	.20	.76	.49	.20	.67	-.38	.20	-.10	.36	-.36
8	.97	-.29	.20	1.08	.88	.20	.95	.07	.20	.29	.29	-.29
9	.57	-1.00	.20	1.22	.10	.20	1.07	-.82	.20	.15	.55	.55
10	1.31	-.11	.20	1.73	.68	.20	1.52	-.16	.20	-.04	.08	.08
11	1.04	-.57	.20	.82	.68	.20	.72	-.16	.20	.33	.40	-.40
12	.79	.82	.20	.81	.88	.20	.71	.07	.20	-.60	.60	-.60
13	1.61	.39	.20	2.00	1.23	.20	1.76	.47	.20	.06	.06	.06
14	1.10	.77	.20	1.06	1.36	.20	.93	.61	.20	-.13	.15	-.15
15	.73	.44	.20	1.11	1.04	.20	.98	.25	.20	-.15	.26	.26
16	1.68	.78	.20	2.00	1.63	.20	1.76	.92	.20	.11	.11	.11
17	1.03	-.45	.20	1.05	.43	.20	.92	-.44	.20	.01	.07	-.07
18	.47	-.94	.20	.90	-.06	.20	.79	-1.00	.20	-.05	.57	.57
19	.91	-.47	.20	.59	.28	.20	.52	-.61	.20	-.12	.55	-.55
20	1.17	-.20	.20	.99	.43	.20	.87	-.44	.20	-.19	.25	-.25
21	1.11	.02	.20	.86	.85	.20	.76	.03	.20	.01	.27	-.27
22	.51	-.74	.20	.72	.10	.20	.63	-.82	.20	-.06	.26	.26
23	.53	-.89	.20	.91	-.02	.20	.80	-.95	.20	-.05	.42	.42
24	.91	-.71	.20	1.33	.32	.20	1.17	-.57	.20	.11	.19	.19
25	1.28	-.19	.20	1.76	.77	.20	1.55	-.06	.20	.11	.13	.13
26	.73	-.02	.20	.65	.79	.20	.57	-.03	.20	-.01	.25	-.25
27	.57	.43	.20	.64	1.48	.20	.56	.75	.20	.26	.26	-.26
28	1.02	-1.92	.20	1.27	-.97	.20	1.12	-2.03	.20	-.09	.10	.10
29	1.22	-1.15	.20	1.26	-.63	.20	1.11	-1.65	.20	-.40	.40	-.40
30	1.29	-.51	.20	1.37	.11	.20	1.21	-.81	.20	-.24	.24	-.24
31	1.00	.29	.20	1.26	1.22	.20	1.11	.45	.20	.13	.14	.14
32	1.17	-1.33	.20	1.08	-.41	.20	.95	-1.40	.20	-.05	.14	-.14
33	.91	-.56	.20	.98	.43	.20	.86	-.44	.20	.09	.10	-.10
34	.97	-.73	.20	.86	.05	.20	.76	-.88	.20	-.12	.21	-.21
35	1.22	-.08	.20	1.26	.85	.20	1.11	.03	.20	.09	.10	-.10
36	1.63	-.10	.20	1.71	.99	.20	1.51	.19	.20	.23	.23	-.23
37	1.07	-.07	.20	.87	.55	.20	.77	-.31	.20	-.19	.29	-.29
38	1.35	-.07	.20	1.55	.73	.20	1.36	-.10	.20	-.03	.03	.03
39	.86	.47	.20	1.50	1.36	.20	1.32	.61	.20	.11	.28	.28
40	1.59	.27	.20	1.54	1.25	.20	1.36	.49	.20	.17	.18	-.18

The standard errors of estimated parameters, signed and unsigned areas, and  $H$  are presented in Table 3. In this table, the standard error for the unsigned area is for the case when the null hypothesis of  $\mu(H) = 0$  was true.

The signed areas vary between  $-.60$  and  $.80$ . Similarly, the unsigned areas vary between  $.03$  and  $.80$ . Items with  $Z$  scores greater than  $+3.0$  or less than  $-3.0$  are identified with an asterisk in Table 3. These items seem to indicate significantly different performance between White and Black examinees.



Table 3  
 Standard Errors (SE) and Covariances [C(a,b)] of Estimated Item Parameters,  
 Signed and Unsigned Areas (SA and UA), H Values and MH  $\chi^2$  Values

Item No.	White Group			Black Group			SA	SA		UA	SE	H	H		MH $\chi^2$
	a	b	C(a,b)	a	b	C(a,b)		SE	Z				SE	Z	
1	.06	.28	-.02	.06	.11	.00	.80	.24	3.29*	.80	.14	.80	.23	3.51*	24.52*
2	.07	.14	-.01	.07	.09	.00	-.24	.13	-1.86	.26	.09	.26	.15	1.76	5.35
3	.09	.12	-.01	.07	.09	.00	-.06	.12	-.50	.15	.05	-.15	.08	-1.83	.64
4	.06	.16	-.01	.05	.15	.00	-.14	.17	-.80	.53	.13	-.53	.21	-2.55	1.99
5	.09	.07	.00	.10	.06	.00	.06	.08	.79	.06	.04	.06	.07	.86	4.25
6	.08	.09	.00	.06	.09	.00	-.06	.10	-.59	.21	.07	-.21	.12	-1.71	.42
7	.10	.06	.00	.07	.09	.00	-.10	.09	-1.15	.36	.07	-.36	.12	-3.11*	.16
8	.09	.07	.00	.11	.08	.00	.29	.08	3.52*	.29	.05	-.29	.08	-3.52*	14.60*
9	.06	.13	-.01	.10	.06	.00	.15	.11	1.31	.55	.07	.55	.12	4.58*	.33
10	.12	.05	.00	.16	.05	.00	-.04	.06	-.66	.08	.04	.08	.07	1.20	.13
11	.09	.07	.00	.08	.09	.00	.33	.09	3.63*	.40	.06	-.40	.10	-4.17*	19.57*
12	.10	.08	.00	.08	.10	.00	-.60	.10	-5.79*	.60	.06	-.60	.10	-5.79*	.05
13	.16	.04	.00	.27	.06	.00	.06	.06	1.01	.06	.03	.06	.06	1.04	3.19
14	.13	.06	.00	.13	.11	.01	-.13	.10	-1.30	.15	.07	-.15	.12	-1.22	2.95
15	.08	.08	.00	.12	.08	.00	-.15	.09	-1.66	.26	.07	.26	.11	2.35	.22
16	.19	.05	.00	.36	.09	.01	.11	.08	1.40	.11	.05	.11	.08	1.40	13.66
17	.09	.07	.00	.09	.07	.00	.01	.08	.13	.07	.05	-.07	.09	-.77	.68
18	.06	.15	-.01	.07	.07	.00	-.05	.13	-.38	.57	.11	.57	.19	3.06*	.59
19	.08	.07	.00	.06	.10	.00	-.12	.10	-1.18	.55	.09	-.55	.15	-3.56*	.01
20	.11	.06	.00	.09	.07	.00	-.19	.07	-2.58	.25	.04	-.25	.07	-3.42*	.49
21	.10	.06	.00	.09	.09	.00	.01	.09	.12	.27	.07	-.27	.11	-2.41	.02
22	.06	.13	.00	.06	.09	.00	-.06	.12	-.49	.26	.11	.26	.19	1.39	.41
23	.06	.13	-.01	.08	.07	.00	-.05	.12	-.42	.42	.10	.42	.16	2.59	.61
24	.08	.08	.00	.11	.06	.00	.11	.08	1.37	.19	.04	.19	.06	2.93	1.40
25	.12	.05	.00	.17	.05	.00	.11	.06	1.78	.13	.03	.13	.05	2.52	2.76
26	.08	.08	.00	.07	.11	.00	-.01	.11	-.09	.25	.10	-.25	.17	-1.50	.02
27	.07	.09	.00	.09	.18	.01	.26	.16	1.59	.26	.10	-.26	.16	-1.59	4.08
28	.11	.13	-.01	.11	.08	-.01	-.09	.12	-.72	.10	.08	.10	.13	.76	.05
29	.11	.08	-.01	.10	.07	.00	-.40	.08	-4.71*	.40	.05	-.40	.08	-4.71*	11.64*
30	.11	.06	.00	.11	.06	.00	-.24	.07	-3.64*	.24	.04	-.24	.07	-3.64*	1.57
31	.10	.06	.00	.15	.08	.00	.13	.08	1.58	.14	.05	.14	.08	1.65	2.20
32	.11	.09	-.01	.09	.07	.00	-.05	.09	-.54	.14	.04	-.14	.07	-2.14	1.97
33	.08	.08	.00	.08	.07	.00	.09	.08	1.06	.10	.05	-.10	.09	-1.15	1.53
34	.09	.08	.00	.07	.08	.00	-.12	.09	-1.37	.21	.05	-.21	.08	-2.53	.11
35	.11	.05	.00	.12	.07	.00	.09	.07	1.28	.10	.04	-.10	.07	-1.44	.76
36	.15	.05	.00	.19	.06	.00	.23	.06	3.82*	.23	.04	-.23	.06	-3.82*	6.56
37	.10	.06	.00	.08	.08	.00	-.19	.08	-2.38	.29	.06	-.29	.09	-3.13*	1.31
38	.12	.05	.00	.14	.06	.00	-.03	.06	-.48	.03	.04	.03	.06	.48	.05
39	.09	.07	.00	.20	.08	.01	.11	.08	1.30	.28	.07	.28	.11	2.46	5.16
40	.15	.04	.00	.20	.07	.00	.17	.07	2.49	.18	.04	-.18	.06	-2.81	6.44

The signed area measure identified 7 biased items, whereas the unsigned area measure identified 13 biased items. Because the major purpose of this example is simply to illustrate the computational aspects of the new area measures, no attempt is made to provide reasons for the magnitude and degree of bias observed for the vocabulary items. Finally, note that the covariances between the estimates of *a* and *b* parameters are 0 or close to 0 for each item in both groups.

When the vocabulary test was analyzed with the Mantel-Haenszel (MH) technique (Holland & Thayer, 1988; Raju, Bode, & Larsen, 1989), some of the same items were also identified as biased. The MH  $\chi^2$  values are listed in the last column of Table 3. Those items with  $\chi^2$  values greater than 9 are identified with an asterisk. Of the five items identified as biased with the MH technique, four items were also

identified as biased with the signed and unsigned area measures. In the current dataset, the unsigned area measure identified more items as biased than did the signed area measure and the MH technique. More empirical studies are obviously needed to adequately compare and contrast the signed and unsigned area measures with the currently available differential item functioning or item bias techniques.

### Conclusions

The asymptotic sampling distribution formulas and the associated significance tests described here are only appropriate for the area estimates obtained with the Rasch model, two-parameter model, and three-parameter model with fixed  $c_s$ . Therefore, these procedures are not appropriate for the three-parameter model with variable  $c_s$ . Even though this limits their usefulness, there is likely sufficient interest in the one- and two-parameter models (Hulin et al., 1983; Wright & Stone, 1979) to justify a careful study of the methods proposed here.

Note also that likelihood ratio tests (Thissen, Steinberg, & Wainer, 1988) and  $\chi^2$  tests (Lord, 1980) can be used to assess whether item parameter estimates from two different samples are significantly different from each other. These tests can be used for inferring (indirectly) whether an observed area between two IRFs is significantly different from 0. Alternatively, the procedures presented here can be used to assess directly whether the observed signed and unsigned areas are significantly different from 0. Regardless of a particular investigator's decision for a given study, there is certainly a need for monte carlo and empirical studies to assess the degree of robustness and uniformity of item bias results obtained with the likelihood ratio,  $\chi^2$ , and area procedures.

Asymptotic formulas are known to provide poor estimates if the sample sizes are small. Although this is probably not a very serious restriction in the IRT context where the sample sizes are usually quite large, there is still a need for non-asymptotic mean and variance formulas for the signed and unsigned areas. Also, the mean and variance formulas for the unsigned areas are based on the assumption that  $H$  (as well as  $\hat{b}_2 - \hat{b}_1$ ) is normally distributed. Although this does not appear to be a very restrictive assumption, it would be desirable to have mean and variance formulas for the unsigned areas that do not require the normality assumption. Future research in this area should resolve these problems.

### References

- Cook, L. L., & Eignor, D. R. (1983). Practical considerations regarding the use of item response theory to equate tests. In R. K. Hambleton (Ed.), *Applications of item response theory*. Vancouver: Educational Research Institute of British Columbia.
- Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: Principles and applications*. Boston: Kluwer-Nijhoff.
- Holland, P. W., & Thayer, D. T. (1988). Differential item performance and Mantel-Haenszel procedure. In H. Wainer & H. I. Braun (Eds.), *Test validity*. Hillsdale NJ: Erlbaum.
- Hulin, C. L., Drasgow, F., & Parsons, C. K. (1983). *Item response theory: Applications to psychological measurement*. Homewood IL: Dow Jones-Irwin.
- Johnson, N. L., & Kotz, S. (1970). *Continuous univariate distributions—2*. New York: Wiley.
- Kendall, M., & Stuart, A. (1977). *The advanced theory of statistics* (Vol. 1, 4th ed.). New York: Macmillan.
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale NJ: Erlbaum.
- Raju, N. S. (1988). The area between two item characteristic curves. *Psychometrika*, 53, 495–502.
- Raju, N. S., Bode, R. K., & Larsen, V. S. (1989). An empirical assessment of the Mantel-Haenszel statistic for studying differential item performance. *Applied Measurement in Education*, 2(1), 1–13.
- Science Research Associates, Inc. (1985). *Survey of basic skills: Technical information*. Chicago: Author.
- Thissen, D., Steinberg, L., & Wainer, H. (1988). Use of item response theory in the study of group differences in trace lines. In H. Wainer & H. I. Braun (Eds.), *Test validity*. Hillsdale NJ: Erlbaum.
- Wingersky, M. S., Barton, M. A., & Lord, F. M. (1983). *LOGIST: A computer program for estimating examinee ability and item characteristic curve parameters*. LOG-

*IST 5, version 1.* Princeton NJ: Educational Testing Service.  
Wright, B. D., & Stone, M. H. (1979). *Best test design*. Chicago: MESA Press.

*anonymous reviewers for their many valuable comments on an earlier version of this manuscript.*

### **Acknowledgments**

*The author thanks Jeffrey A. Slinde, the Editor, and two*

### **Author's Address**

Send requests for reprints or further information to Nambury S. Raju, Department of Psychology, Illinois Institute of Technology, Chicago IL 60616, U.S.A.