DOCUMENT RESUME

ED 347 210 TM 018 742

AUTHOR Kim, Seock-Ho; And Others

TITLE An Investigation of Hierarchical Bayes Procedures in Item Response Theory.

PUB DATE Apr 92

NOTE 52p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, April 20-24, 1992). Some pages contain light, broken type.

PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC03 Plus Postage.

DESCRIPTORS Ability; *Bayesian Statistics; Comparative Analysis; *Equations (Mathematics); *Estimation (Mathematics); *Item Response Theory; *Mathematical Models; Probability; Sample Size; Simulation; *Statistical Bias; Statistical Distributions; Test Length

IDENTIFIERS A Posteriori Index; A Priori Tests; *Hierarchical Analysis; Item Parameters; Two Parameter Model

ABSTRACT

Hierarchical Bayes procedures were compared for estimating item and ability parameters in item response theory. Simulated data sets from the two-parameter logistic model were analyzed using three different hierarchical Bayes procedures: (1) the joint Bayesian with known hyperparameters (JB1); (2) the joint Bayesian with information hyperpriors (JB2); and (3) the marginal Bayesian with known hyperparameters (MB). Prior and posterior distributions focusing on one- and two-stage hierarchical priors are presented, and two joint Bayesian methods that consider the specific priors are discussed. The MB procedure yielded consistently smaller root mean square differences than did either the JB1 or JB2 procedure for item and ability estimates. The maximum a posteriori estimation used along with the MB procedure yielded larger biases than did the joint Bayes model estimation in JB1 and JB2. As the sample size and test length increased, the three Bayes procedures yielded essentially the same result. Fifteen tables present study data, and there is a 39-item list of references. (SLD)

An Investigation of Hierarchical Bayes Procedures In Item Response Theory

Seock-Ho Kim, Allan S. Cohen, Frank B. Baker,
Michael J. Subkoviak, and Tom Leonard
University of Wisconsin–Madison

Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA
April, 1992

Abstract

Hierarchical Bayes procedures were compared for estimating item and ability parameters in item response theory. Simulated data sets from the two-parameter logistic model were analyzed using three different hierarchical Bayes procedures: the joint Bayesian with known hyperparameters (JB1), the joint Bayesian with informative hyperpriors (JB2), and the marginal Bayesian with known hyperparameters (MB). MB yielded consistently smaller root mean square differences than either JB1 or JB2 for item and ability estimates. The maximum a posteriori estimation used along with MB yielded larger biases then the joint Bayes modal estimation in JB1 and JB2. As the sample size and test length increased, the three Bayes procedures yielded essentially the same result.

Key words: Bayes estimation, hierarchical prior, item response theory, joint Bayesian estimation, marginal Bayesian estimation.

Introduction

A common situation in item response theory (IRT) is that in which both item and ability (i.e., structural and incidental) parameters have to be estimated simultaneously. When this is the case, Bayesian estimation may be preferable to maximum likelihood estimation. Bayesian methods yield item discrimination parameter estimates which never become infinite; lower asymptote estimates of item characteristic curves which do not have implausible values; and ability estimates which are automatically restricted to a reasonable range (Lord, 1986). Although Bayes procedures have been available for some time, the properties of these techniques have not been studied as thoroughly as those of maximum likelihood methods. The purpose of this study, therefore, was to compare different Bayes procedures for estimation of item and ability parameters in IRT.

Bayesian approaches in IRT can be distinguished on the basis of whether estimation of item parameters is done with or without marginalization over ability parameters. If marginalization is used, the solution is marginal Bayesian estimation; if marginalization is not used, the solution is joint Bayesian estimation.

Swaminathan and Gifford (1982, 1985, 1986) developed the joint Bayesian procedures for the one-, two-, and three-parameter item characteristic curve models. Their methods implement the hierarchical Bayes procedures for the specification of prior beliefs following the approach taken by Lindley (1971) and Lindley and Smith (1972). Evidence presented by Swaminathan and Gifford indicated that joint Bayesian parameter estimates were superior to those obtained via joint maximum likelihood estimation in that they remained in the parameter space, had smaller mean square differences from the underlying values, and were less biased (Gifford & Swaminathan, 1990).

Mislevy (1986) employed the hierarchical Bayesian estimation model of Lindley and Smith (1972) to extend the marginal maximum likelihood approach to a marginal Bayesian solution.

This permitted prior distributions to be posited for item parameters. Supplementary Bayesian procedures can also be used to obtain ability estimates once the marginal Bayesian estimates are obtained for item parameters. Mislevy and Bock (1989) implemented this Bayesian approach in the BILOG computer program. Tsutakawa and Lin (1986) also proposed a marginal Bayesian estimation to compute the posterior mode using the EM algorithm.

Evidence has been subsequently presented which points to the likelihood that marginal modes may provide better approximations than joint modes to posterior means when nuisance (i.e., ability) parameters are present (Mislevy, 1986; O'Hagan, 1976; Tsutakawa & Lin, 1986). As yet, however, no empirical analyses have been reported which test this point.

Bayesian approaches are characterized by incorporation of prior information or beliefs into the estimation of parameters in order to improve the accuracy of those estimates. Specification of priors in Bayesian analysis is a subjective matter. A number of different forms of priors have been studied (e.g., Leonard & Novick, 1985; Lord, 1980;, Mislevy, 1986; Mislevy & Bock, 1989; Swaminathan & Gifford, 1986; Tsutakawa & Lin, 1986). The terminology describing the structure of priors can sometimes be quite confusing. In a classical Bayesian approach, a single prior can be selected for the ordinary parameters. It is possible to recognize some uncertainty in priors. When priors are expressed in terms of family or class of prior, we call the parameters in the class of priors as hyperparameters. Hyperparameters describe the distributional characteristics of the prior information. It is sometimes also convenient to specify prior information on the hyperparameters as well. This second prior is called a hyperprior and contains parameters which are referred to as hyperhyperparameters (Good, 1980, 1983; Lindley, 1971, Lindley & Smith, 1972).

To completely exploit the potential of the Bayesian estimation requires understanding of its mathematical underpinnings, particularly the role of prior distributions in estimating parameters. In the present study, we compared the effectiveness of three hierarchical Bayes procedures for obtaining item and ability estimates: the joint Bayesian estimation with known hyperparameters (JB1), the joint Bayesian estimation with informative hyperpriors (JB2), and the marginal Bayesian estimation with known hyperparameters (MB).

In the following sections, we present a discussion of joint and marginal Bayesian estimation in IRT. Included is a presentation of prior and posterior distributions focusing specifically on one- and two-stage hierarchical priors. Finally, we present a discussion of the two joint Bayesian methods considering the specific priors dealt with in this paper.

Background

The Model

Item characteristic curve models are expressed as mathematical equations of the probability of a correct response to a test item as a function of the ability of the person responding. Consider binary responses to a set of n test items by a set of N examinees. A response of an examinee i to an item j is represented in these models by a random variable U_{ij}, where $i = 1, \ldots, N$ and $j = 1, \ldots, n$. The probability of a correct response to item j is represented by

$$P(U_{ij} = 1|\theta_i, \underset{\sim}{\xi}_j) = P_j(\theta_i), \tag{1}$$

and the probability of an incorrect response is given by

$$P(U_{ij} = 0|\theta_i, \underset{\sim}{\xi}_j) = Q_j(\theta_i), \tag{2}$$

depending on a real-valued ability parameter θ_i, and a real- or vector-valued item parameter $\underset{\sim}{\xi}_j$.

The item characteristic curve of the three-parameter model[1] is given by

$$P_j(\theta_i) = c_j + (1 - c_j)\left[1 + \exp\{-a_j(\theta_i - b_j)\}\right]^{-1}, \tag{3}$$

where a_j is the item discrimination parameter, b_j is the item difficulty parameter, c_j is the lower asymptote of the item characteristic curve for the item j, and θ_i is the ability parameter of the person i.

Likelihood Function

Under typical testing conditions, a sample of N examinees are drawn at random from a population of examinees possessing the underlying ability. No assumption is necessary as to the distribution of the examinees over the ability continuum (Lord & Novick, 1968). For each examinee there is a vector of dichotomously scored item responses of length n denoted by $\mathbf{U}_i = (U_{i1}, \ldots, U_{in})'$. One such vector exists for each of the N examinees. The resulting $N \times n$ matrix of item responses is denoted by $\mathbf{U}$.

Under the local independence assumption, the probability of $\mathbf{U}_i$ given ability θ_i and item parameters $\underline{\xi}$ is

$$p(\mathbf{U}_i|\theta_i, \underline{\xi}) = \prod_{j=1}^{n} P_j(\theta_i)^{U_{ij}} Q_j(\theta_i)^{1-U_{ij}}, \tag{4}$$

where $\underline{\xi} = (\underline{\xi}_1, \ldots, \underline{\xi}_n)'$. If $\underline{\theta}$ is the vector of the N examinee trait scores, $\underline{\theta} = (\theta_1, \ldots, \theta_N)'$, the joint probability of $\mathbf{U}$ given by $\underline{\theta}$ and $\underline{\xi}$ can be written as

$$p(\mathbf{U}|\underline{\theta}, \underline{\xi}) = \prod_{i=1}^{N} \prod_{j=1}^{n} P_j(\theta_i)^{U_{ij}} Q_j(\theta_i)^{1-U_{ij}}. \tag{5}$$

When we make inferences about both ability and item parameters from the observed data $\mathbf{u}$ of the $N \times n$ matrix of item responses, the probability of $\mathbf{u}$ given by $\underline{\theta}$ and $\underline{\xi}$ is

$$p(\mathbf{u}|\underline{\theta}, \underline{\xi}) = \prod_{i=1}^{N} \prod_{j=1}^{n} P_j(\theta_i)^{u_{ij}} Q_j(\theta_i)^{1-u_{ij}} = l(\underline{\theta}, \underline{\xi}). \tag{6}$$

[1]Because of inclusiveness, that is, the one- and two-parameter item characteristic curve models are regarded as the special cases of the three-parameter model, all expressions are developed below only for Birnbaum's three-parameter model (Birnbaum, 1968).

The likelihood, $l(\underline{\theta}, \underline{\xi})$, is a function of the parameters of the n item characteristic curves and the N abilities.

Parameter Estimation in IRT

The four main approaches currently used in IRT for parameter estimation are (a) joint maximum likelihood estimation, (b) joint Bayesian estimation, (c) marginal maximum likelihood estimation, and (d) marginal Bayesian estimation. The following discussion presents a description of the Bayesian procedures as the extensions of the maximum likelihood methods where the priors are posited for the item and ability parameters.

The joint maximum likelihood estimation (Birnbaum, 1968; Lord, 1980; Wingersky, Barton, & Lord, 1982) simultaneously maximizes the likelihood function $l(\underline{\theta}, \underline{\xi})$ in Equation 6.

The joint Bayesian estimation (Swaminathan & Gifford, 1982, 1985, 1986) simultaneously maximizes the posterior distribution

$$\pi(\underline{\theta}, \underline{\xi}|\mathbf{u}) \propto l(\underline{\theta}, \underline{\xi})\pi(\underline{\theta}, \underline{\xi}), \tag{7}$$

where $\propto$ denotes proportionality and $\pi(\underline{\theta}, \underline{\xi})$ is the joint prior density of the parameters $\underline{\theta}$ and $\underline{\xi}$. Equivalently, the posterior distribution of parameters given the matrix of observations $\mathbf{u}$ is written as

$$\pi(\underline{\theta}, \underline{\xi}|\mathbf{u}) = \frac{l(\underline{\theta}, \underline{\xi})\pi(\underline{\theta}, \underline{\xi})}{m(\mathbf{u})}, \tag{8}$$

where $m(\mathbf{u})$ is the marginal probability density function of $\mathbf{u}$ defined as

$$m(\mathbf{u}) = \int_\Theta \int_\Xi l(\underline{\theta}, \underline{\xi})\pi(\underline{\theta}, \underline{\xi})d\underline{\xi}d\underline{\theta}, \tag{9}$$

where Θ and Ξ are the parameter spaces for ability and item parameters, respectively. The posterior density function is a revised expression of the belief one has about the parameters once the data have been collected. It contains all the information necessary for making probability statements regarding the parameters of interest.

The marginal maximum likelihood estimation of item parameters (Bock & Aitkin, 1981; Bock & Lieberman, 1970; Harwell, Baker, & Zwarts, 1988) maximizes the marginal likelihood function

$$m(\underline{\xi}) = \prod_{i=1}^{N} \int_{\Theta} l(\theta_i, \underline{\xi}) \pi(\theta_i) d\theta_i, \tag{10}$$

where $\pi(\theta_i)$ denotes a prior distribution of ability and

$$l(\theta_i, \underline{\xi}) = \prod_{j=1}^{n} P_j(\theta_i)^{u_{ij}} Q_j(\theta_i)^{1-u_{ij}} = p(\mathbf{u}_i | \theta_i, \underline{\xi}). \tag{11}$$

Supplementary maximum likelihood estimation and Bayesian estimation procedures can be used to obtain ability parameter estimates.

Bayesian priors on item parameters may also be used in the marginal maximum likelihood estimation to obtain the marginal Bayesian estimation of item parameters (Harwell & Baker, 1991; Mislevy, 1986). The marginal Bayesian estimation maximizes the marginal posterior distribution

$$\pi(\underline{\xi} | \mathbf{u}) \propto m(\underline{\xi}) \pi(\underline{\xi}), \tag{12}$$

where $m(\underline{\xi})$ is the marginal likelihood function and $\pi(\underline{\xi})$ is the prior distribution of item parameters.

Prior and Posterior Distributions

Prior Distribution

A flexible family of prior distributions is available by transforming item parameters to new parameters which may be taken to possess a multivariate normal prior distribution. To this end Leonard and Novick (1985) and Mislevy (1986) recommend the following transformations:

$$\alpha_j = \ln a_j \tag{13}$$

and

$$\gamma_j = \ln\{c_j/(1 - c_j)\}. \tag{14}$$

Since b_j is a difficulty parameter, we also use the following expression:

$$\beta_j = b_j. \tag{15}$$

In order to define the posterior distribution precisely, we first specify the prior belief about the parameters. We assume $\underline{\theta}$ and $\underline{\xi}$ priors which are independently distributed with probability density functions $\pi(\underline{\theta})$ and $\pi(\underline{\xi})$, respectively.

Since we use the three-parameter model,

$$\underline{\xi} = (\alpha_1, \beta_1, \gamma_1, \ldots, \alpha_n, \beta_n, \gamma_n)'. \tag{16}$$

We assume the vector of item parameters possesses a multivariate normal distribution conditional on the respective mean vector $\underline{\mu}_{\xi}$ and covariance matrix $\underline{\Sigma}_{\xi}$. This prior specification is more general than previous suggestions in the literature. The prior distribution of item parameters is

$$\pi(\underline{\xi}|\underline{\eta}) = (2\pi)^{-3n/2} \left|\underline{\Sigma}_{\xi}\right|^{-1/2} \exp\left\{-\frac{1}{2}(\underline{\xi} - \underline{\mu}_{\xi})'\underline{\Sigma}_{\xi}^{-1}(\underline{\xi} - \underline{\mu}_{\xi})\right\}, \tag{17}$$

where the hyperparameter $\underline{\eta} = (\underline{\mu}_{\xi}, \underline{\Sigma}_{\xi})$.

If we assume the vectors of the parameters $\underline{\alpha} = (\alpha_1, \ldots, \alpha_n)'$, $\underline{\beta} = (\beta_1, \ldots, \beta_n)'$, and $\underline{\gamma} = (\gamma_1, \ldots, \gamma_n)'$ to be independent, we can take the vectors $\underline{\alpha}$, $\underline{\beta}$, and $\underline{\gamma}$ to possess independent multivariate normal distributions, conditional on their respective mean vectors $\underline{\mu}_{\alpha}$, $\underline{\mu}_{\beta}$, and $\underline{\mu}_{\gamma}$, and covariance matrices $\underline{\Sigma}_{\alpha}$, $\underline{\Sigma}_{\beta}$, and $\underline{\Sigma}_{\gamma}$ (Leonard & Novick, 1985). The prior distribution of item parameters in this case is

$$\pi(\underline{\xi}|\underline{\eta}) = \pi(\underline{\alpha}|\underline{\eta}_{\alpha})\pi(\underline{\beta}|\underline{\eta}_{\beta})\pi(\underline{\gamma}|\underline{\eta}_{\gamma}), \tag{18}$$

where $\underline{\eta}_{\alpha} = (\underline{\mu}_{\alpha}, \underline{\Sigma}_{\alpha})$, $\underline{\eta}_{\beta} = (\underline{\mu}_{\beta}, \underline{\Sigma}_{\beta})$, $\underline{\eta}_{\gamma} = (\underline{\mu}_{\gamma}, \underline{\Sigma}_{\gamma})$,

$$\pi(\underline{\alpha}|\underline{\eta}_{\alpha}) = (2\pi)^{-n/2} \left|\underline{\Sigma}_{\alpha}\right|^{-1/2} \exp\left\{-\frac{1}{2}(\underline{\alpha} - \underline{\mu}_{\alpha})'\underline{\Sigma}_{\alpha}^{-1}(\underline{\alpha} - \underline{\mu}_{\alpha})\right\}, \tag{19}$$

and $\pi(\underline{\beta}|\underline{\eta}_\beta)$ and $\pi(\underline{\gamma}|\underline{\eta}_\gamma)$ are defined similarly.

If we further assume exchangeability for all three parameters, we may take $\underline{\mu}_\alpha = \mu_\alpha \mathbf{1}$, $\underline{\mu}_\beta = \mu_\beta \mathbf{1}$, $\underline{\mu}_\gamma = \mu_\gamma \mathbf{1}$, $\underline{\Sigma}_\alpha = \sigma^2_\alpha \mathbf{I}_n$, $\underline{\Sigma}_\beta = \sigma^2_\beta \mathbf{I}_n$, and $\underline{\Sigma}_\gamma = \sigma^2_\gamma \mathbf{I}_n$, where μ_α, μ_β, μ_γ, σ^2_α, σ^2_β, and σ^2_γ are scalars, $\mathbf{1}$ is an $n \times 1$ vector of ones, and $\mathbf{I}_n$ is an identity matrix of order n (Leonard & Novick, 1985). The prior distribution of item parameters, assuming exchangeability, is

$$\pi(\underline{\xi}|\underline{\eta}) = \prod_{j=1}^{n} \pi(\alpha_j|\mu_\alpha, \sigma^2_\alpha)\pi(\beta_j|\mu_\beta, \sigma^2_\beta)\pi(\gamma_j|\mu_\gamma, \sigma^2_\gamma), \tag{20}$$

where

$$\pi(\alpha_j|\mu_\alpha, \sigma^2_\alpha) = (2\pi\sigma^2_\alpha)^{-1/2} \exp\left\{-\frac{1}{2\sigma^2_\alpha}(\alpha_j - \mu_\alpha)^2\right\}, \tag{21}$$

and $\pi(\beta_j|\mu_\beta, \sigma^2_\beta)$ and $\pi(\gamma_j|\mu_\gamma, \sigma^2_\gamma)$ are defined similarly. This form of the prior distribution of item parameters is used in the present study for the joint Bayesian estimation as well as for the marginal Bayesian estimation procedures. A hierarchical Bayes approach is developed below in which another stage priors are assigned to the prior parameters, μ_α, μ_β, μ_γ, σ^2_α, σ^2_β, and σ^2_γ.

Hierarchical Approach

We can specify prior distributions for the parameter vectors $\underline{\theta}$ and $\underline{\xi}$ in two stages. This type of prior distribution is a hierarchical prior (Berger, 1985; Good, 1983) also called a multistage prior (Lindley, 1971; Lindley & Smith, 1972). The idea is that one may have structural and subjective prior information at the same time and that it is often convenient to model this in stages.

The structural knowledge that the θ_i are independent and identically distributed leads to the first stage prior description

$$\pi_1(\underline{\theta}) = \prod_{i=1}^{N} \pi_0(\theta_i). \tag{22}$$

The subscript 1 on π_1 is to indicate that this is the first stage. The hierarchical approach then places a second stage subjective prior on π_0. If we use Γ to denote a class of priors, the hierarchical approach is most commonly used when the first stage, Γ , consists of priors of a certain functional form. Thus, if

$$\Gamma_\theta = \{\pi_1(\underline{\theta}|\underline{\tau}) : \pi_1 \text{ is of a given functional form and } \underline{\tau} \in T\}, \tag{23}$$

then the second stage would consist of putting a prior distribution, $\pi_2(\underline{\tau})$, on the hyperparameter $\underline{\tau}$. Such a second stage prior is sometimes called a hyperprior (Berger, 1985; Good, 1983).

The structural assumption of independence of the θ_i, together with the assumption that they have a common normal distribution (i.e., we assume that the information on these parameters is exchangeable), leads to

$$\Gamma_\theta = \left\{\pi_1(\underline{\theta}|\underline{\tau}) : \pi_1(\underline{\theta}|\underline{\tau}) = \prod_{i=1}^{N} \pi_0(\theta_i), \pi_0 \text{ being } \mathcal{N}(\mu_\theta, \sigma_\theta^2), -\infty < \mu_\theta < \infty \text{ and } \sigma_\theta^2 > 0\right\}, \tag{24}$$

where

$$\pi_1(\underline{\theta}|\underline{\tau}) = \prod_{i=1}^{N}(2\pi\sigma_\theta^2)^{-1/2} \exp\left\{-\frac{1}{2\sigma_\theta^2}(\theta_i - \mu_\theta)^2\right\}. \tag{25}$$

Similarly, assumptions that item parameters are independent and identically distributed and that the information on each of the item parameters is exchangeable lead to Γ_α, Γ_β, and Γ_γ with the hyperparameter $\underline{\eta}$. Then the first stage prior distribution of item parameters assuming independence and exchangeability is

$$\pi_1(\underline{\xi}|\underline{\eta}) = \prod_{j=1}^{n} \pi_1(\alpha_j|\mu_\alpha, \sigma_\alpha^2)\pi_1(\beta_j|\mu_\beta, \sigma_\beta^2)\pi_1(\gamma_j|\mu_\gamma, \sigma_\gamma^2). \tag{26}$$

The complete prior for the hierarchical model, assuming independence between ability and item parameters, is

$$\pi(\underline{\theta}, \underline{\tau}, \underline{\xi}, \underline{\eta}) = \pi_1(\underline{\theta}|\underline{\tau})\pi_2(\underline{\tau})\pi_1(\underline{\xi}|\underline{\eta})\pi_2(\underline{\eta}), \tag{27}$$

where $\pi_1(\underline{\theta}|\underline{\tau})$ is the first stage density of $\underline{\theta}$ conditional on $\underline{\tau}$ which takes the second stage density $\pi_2(\underline{\tau})$ and $\pi_1(\underline{\xi}|\underline{\eta})$ is the first stage density of $\underline{\xi}$ conditional on $\underline{\eta}$ which takes the second stage density $\pi_2(\underline{\eta})$.

Second Stage Prior

Noninformative priors are often used at the second stage because of the difficulty in specifying second stage priors (Berger, 1985). Sometimes, it is simply assumed that hyperparameters are known. For example, in the joint Bayesian estimation procedures, identifying restrictions can be incorporated directly into the prior (Swaminathan & Gifford, 1986) because the three-parameter model does not need to be identified. Therefore, we set $\mu_\theta = 0$ and $\sigma_\theta^2 = 1$, so that

$$\pi_1(\underline{\theta}|\underline{\tau}) = (2\pi)^{-N/2} \exp\left(-\frac{1}{2}\sum_{i=1}^{N}\theta_i^2\right). \tag{28}$$

In the above specification, setting $\mu_\theta = 0$ and $\sigma_\theta^2 = 1$ contains the explicit assumption that the hyperparameter $\underline{\tau}$ is known.

In the present study, we use the identical form of prior for each of the item parameters. Detailed examples, therefore, are given below only for the transformed item discrimination parameter. Hyperpriors for μ_α and σ_α^2 can be specified by assuming that μ_α and σ_α^2 are independent, μ_α has a noninformative uniform distribution, and σ_α^2 has an inverse gamma distribution with parameters ν_α and λ_α, $\mathcal{IG}(\nu_\alpha, \lambda_\alpha)$. That is,

$$\pi_2(\underline{\eta}_{\underline{\alpha}}) = \pi_2(\mu_\alpha)\pi_2(\sigma_\alpha^2|\nu_\alpha, \lambda_\alpha) = \frac{1}{\Gamma(\nu_\alpha)\lambda_\alpha^{\nu_\alpha}(\sigma_\alpha^2)^{\nu_\alpha+1}} \exp\left(-\frac{1}{\lambda_\alpha\sigma_\alpha^2}\right), \tag{29}$$

where $\nu_\alpha > 0$ and $\lambda_\alpha > 0$. Since $E(\sigma_\alpha^{-2}) = \nu_\alpha\lambda_\alpha$, we consider $\frac{1}{\nu_\alpha\lambda_\alpha}$ as a prior variance estimate and $2\nu_\alpha$ as a prior sample size for the variance of item discrimination (Leonard, 1972; Novick, 1969). The prior for $\underline{\alpha}$ can be expressed as

$$\pi_1(\underline{\alpha}|\underline{\eta}_{\underline{\alpha}})\pi_2(\underline{\eta}_{\underline{\alpha}}) = \prod_{j=1}^{n} \pi_0(\alpha_j)\pi_2(\underline{\eta}_{\underline{\alpha}}) \tag{30}$$

$$= (2\pi\sigma_\alpha^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma_\alpha^2}\sum_{j=1}^{n}(\alpha_j - \mu_\alpha)^2\right\} \frac{1}{\Gamma(\nu_\alpha)\lambda_\alpha^{\nu_\alpha}(\sigma_\alpha^2)^{\nu_\alpha+1}} \exp\left(-\frac{1}{\lambda_\alpha\sigma_\alpha^2}\right). \tag{31}$$

The above expression depends on the nuisance parameters, μ_α and σ_α^2. These can be integrated out to yield

$$\int_0^\infty \int_{-\infty}^\infty \pi_1(\underline{\alpha}|\underline{\eta}_{\underline{\alpha}})\pi_2(\underline{\eta}_{\underline{\alpha}})d\mu_\alpha d\sigma_\alpha^2 \propto \left\{\frac{2}{\lambda_\alpha} + \sum_{j=1}^n(\alpha_j - \bar{\alpha})^2\right\}^{-(n+2\nu_\alpha-1)/2}. \tag{32}$$

Therefore,

$$\pi(\underline{\alpha}|\nu_\alpha, \lambda_\alpha) \propto \left\{\frac{2}{\lambda_\alpha} + \sum_{j=1}^n(\alpha_j - \bar{\alpha})^2\right\}^{-(n+2\nu_\alpha-1)/2}. \tag{33}$$

Similar prior specifications yield

$$\pi(\underline{\beta}|\nu_\beta, \lambda_\beta) \propto \left\{\frac{2}{\lambda_\beta} + \sum_{j=1}^n(\beta_j - \bar{\beta})^2\right\}^{-(n+2\nu_\beta-1)/2} \tag{34}$$

and

$$\pi(\underline{\gamma}|\nu_\gamma, \lambda_\gamma) \propto \left\{\frac{2}{\lambda_\gamma} + \sum_{j=1}^n(\gamma_j - \bar{\gamma})^2\right\}^{-(n+2\nu_\gamma-1)/2}, \tag{35}$$

where $\bar{\beta} = \frac{1}{n}\sum_{j=1}^n \beta_j$ and $\bar{\gamma} = \frac{1}{n}\sum_{j=1}^n \gamma_j$.

In the context of the hierarchical approach (Goel, 1983; Goel & DeGroot, 1981), we can illustrate the above specification of priors of item parameters as

$$\pi(\underline{\xi}, \underline{\eta}) = \pi_1(\underline{\xi}|\underline{\eta})\pi_2(\underline{\eta}), \tag{36}$$

where $\pi_2(\underline{\eta})$ is viewed as

$$\pi_2(\underline{\eta}) = \pi_{2,1}(\underline{\eta}^{(1)}|\underline{\eta}^{(2)})\pi_{2,2}(\underline{\eta}^{(2)}). \tag{37}$$

It can be seen that $\underline{\eta} = (\underline{\eta}^{(1)}, \underline{\eta}^{(2)})$. We integrate out the nuisance parameter $\underline{\eta}^{(1)}$ explicitly assuming $\underline{\eta}^{(2)}$ is known:

$$\int_{H^{(1)}} \pi(\underline{\xi}, \underline{\eta})d\underline{\eta}^{(1)} = \int_{H^{(1)}} \pi_1(\underline{\xi}|\underline{\eta})\pi_2(\underline{\eta})d\underline{\eta}^{(1)} = \pi(\underline{\xi}|\underline{\eta}^{(2)}). \tag{38}$$

From the assumption of independence of the respective vectors of item parameters,

$$\pi(\underline{\xi}|\underline{\eta}^{(2)}) = \pi(\underline{\alpha}|\underline{\eta}_{\underline{\alpha}}^{(2)})\pi(\underline{\beta}|\underline{\eta}_{\underline{\beta}}^{(2)})\pi(\underline{\gamma}|\underline{\eta}_{\underline{\gamma}}^{(2)}). \tag{39}$$

For the transformed item discrimination parameter, for example,

$$\underline{\eta}_{\underline{\alpha}} = (\underline{\eta}_{\underline{\alpha}}^{(1)}, \underline{\eta}_{\underline{\alpha}}^{(2)}) = (\mu_\alpha, \sigma_\alpha^2; \nu_\alpha, \lambda_\alpha). \tag{40}$$

We can integrate out $\underline{\eta}_{\underline{\alpha}}^{(1)}$ from the prior distribution to yield

$$\int_{H^{(1)}} \pi_1(\underline{\alpha}|\underline{\eta}_{\underline{\alpha}})\pi_2(\underline{\eta}_{\underline{\alpha}})d\underline{\eta}_{\underline{\alpha}}^{(1)} = \pi(\underline{\alpha}|\underline{\eta}_{\underline{\alpha}}^{(2)}) = \pi(\underline{\alpha}|\nu_\alpha, \lambda_\alpha). \tag{41}$$

Posterior Distribution

Bayesian analysis is performed by combining the prior information and the sample information into what is called the posterior distribution; all decisions or inferences are made about the parameter of interest from the posterior distribution. The joint posterior density of $\underline{\theta}$ and $\underline{\xi}$, given observations $\mathbf{u}$, $\underline{\tau}$, and $\underline{\eta}$, is

$$\pi(\underline{\theta}, \underline{\xi}|\mathbf{u}, \underline{\tau}, \underline{\eta}) \propto l(\underline{\theta}, \underline{\xi})\pi(\underline{\theta}|\underline{\tau})\pi(\underline{\xi}|\underline{\eta}). \tag{42}$$

When ignorance (i.e., noninformative) priors are assigned to the hyperparameters, $\underline{\tau}$ and $\underline{\eta}$, the posterior evaluation will be based largely on the data. This will provide Stein-type shrinkage estimates for the item and ability parameters, smoothing each of these toward respective average values (Leonard & Novick, 1985). When the hyperparameters are assumed to be known, the simultaneous maximization of the joint posterior results in JB1.

In JB2, the following joint posterior distribution will be simultaneously maximized to find the joint modal estimates:

$$\pi(\underline{\theta}, \underline{\xi}|\mathbf{u}, \underline{\tau}, \underline{\eta}^{(2)}) \propto l(\underline{\theta}, \underline{\xi})\pi(\underline{\theta}|\underline{\tau})\pi(\underline{\xi}|\underline{\eta}^{(2)}). \tag{43}$$

In the marginal Bayesian estimation context (Harwell & Baker, 1991; Mislevy, 1986), assuming the hyperparameters are known, the examinee parameters $\underline{\theta}$ are integrated over their distribution to obtain the marginal posterior distribution

$$\pi(\underline{\xi}|\mathbf{u}, \underline{\tau}, \underline{\eta}) \propto m(\underline{\xi}|\underline{\tau})\pi(\underline{\xi}|\underline{\eta}). \tag{44}$$

Marginal Bayesian modal estimates of item parameters can be found by maximizing the marginal posterior distribution.

When the two-stage hierarchical priors are employed in the marginal Bayesian estimation of item parameters, assuming the ability hyperparameter $\underline{\tau}$ is known, the marginal posterior distribution can be defined as

$$\pi(\underline{\xi}|\mathbf{u}, \underline{\tau}, \underline{\eta}^{(2)}) \propto m(\underline{\xi}|\underline{\tau})\pi(\underline{\xi}|\underline{\eta}^{(2)}). \quad (45)$$

In the marginal Bayesian estimation procedures, ability parameters are estimated after obtaining the item parameter estimates assuming these are true values. Two Bayes methods are available; Bayes modal estimation and Bayes expected a posteriori (EAP) estimation (Bock & Mislevy, 1982).

Since detailed mathematical derivation can be found for the marginal Bayesian estimation procedures (Harwell & Baker 1991; Mislevy, 1986), in the next section we presents only the two joint Bayesian estimation procedures.

Joint Bayesian Estimation

JB2 Estimation

In order to estimate the item and ability parameters, the log posterior distribution $\ln \pi(\underline{\theta}, \underline{\xi}|\mathbf{u}, \underline{\tau}, \underline{\eta}^{(2)})$ is to be maximized by taking partial derivatives with respect to the parameters and setting them equal to zeros. A procedure such as the Newton-Raphson method is then used to obtain the joint modal estimators.

Since the parameters for all n items and the abilities for all N examinees are unknown, we first to take derivatives of the logarithm of the posterior distribution with respect to these parameters. These are then set equal to zero and the $3n + N$ simultaneous equations solved to obtain the Bayes modal estimates of the unknown parameters. Assuming item and ability parameters are independent, we can obtain joint Bayes modal estimates via Birnbaum's (1968) method.

In Birnbaum's method the item parameter estimation part and the ability parameter estimation part are repeated iteratively until a stable set of item and ability estimates is obtained. In the item parameter estimation part, the Newton-Raphson (Kennedy & Gentle, 1980) equation is

$$\hat{\underline{\xi}}_j^{(s)} = \hat{\underline{\xi}}_j^{(s-1)} - \left\{\mathbf{H}_j^{(s-1)}\right\}^{-1} \mathbf{f}_j^{(s-1)} \tag{46}$$

where s indexes the iteration, $\mathbf{f}_j$ is the gradient vector, $\mathbf{H}_j$ is the Hessian matrix of the log posterior distribution, $F = \ln \pi(\underline{\theta}, \underline{\xi}|\mathbf{u}, \underline{\tau}, \underline{\eta}^{(2)})$. The Newton-Raphson equation of the ability parameter estimation part, for examinee i, is

$$\hat{\theta}_i^{(s)} = \hat{\theta}_i^{(s-1)} - \left(\frac{\partial^2 F}{\partial \theta_i^2}\right)^{-1}_{(s-1)} \left(\frac{\partial F}{\partial \theta_i}\right)_{(s-1)}. \tag{47}$$

We take a partial derivative of the log posterior distribution with respect to each item parameter, for example α_j, and set to zero. The resulting equation becomes

$$\frac{\partial}{\partial \alpha_j} \ln l(\underline{\theta}, \underline{\xi}) + \frac{\partial}{\partial \alpha_j} \ln \pi(\underline{\xi}|\underline{\eta}^{(2)}) = 0. \tag{48}$$

Similarly, when we take a partial derivative of the log of the posterior distribution with regard to an examinee's ability parameter, θ_i, and set to zero, the resulting equation is

$$\frac{\partial}{\partial \theta_i} \ln l(\underline{\theta}, \underline{\xi}) + \frac{\partial}{\partial \theta_i} \ln \pi(\underline{\theta}|\underline{\tau}) = 0. \tag{49}$$

In the subsequent sections, we derive the individual elements which are needed in the Newton-Raphson method for the joint Bayesian-2 estimation procedure.

Likelihood

Taking logarithms, the log likelihood function is

$$\ln l(\underline{\theta}, \underline{\xi}) = \sum_{i=1}^{N} \sum_{j=1}^{n} \left[u_{ij} \ln \{P_j(\theta_i)\} + (1 - u_{ij}) \ln \{Q_j(\theta_i)\}\right]. \tag{50}$$

First, we need partial derivatives of $P_j(\theta_i)$ with respect to each item parameter. The partial derivatives of $P_j(\theta_i)$ with respect to α_j, β_j, and γ_j are

$$\frac{\partial}{\partial \alpha_j} P_j(\theta_i) = \exp(\alpha_j)\{1 - \Psi(\gamma_j)\}(\theta_i - \beta_j)P_j^*(\theta_i)Q_j^*(\theta_i), \tag{51}$$

$$\frac{\partial}{\partial \beta_j} P_j(\theta_i) = -\exp(\alpha_j)\{1 - \Psi(\gamma_j)\}P_j^*(\theta_i)Q_j^*(\theta_i), \tag{52}$$

and

$$\frac{\partial}{\partial \gamma_j} P_j(\theta_i) = \Psi(\gamma_j)\{1 - \Psi(\gamma_j)\}Q_j^*(\theta_i), \tag{53}$$

where $P_j^*(\theta_i) = [1 + \exp\{-\exp(\alpha_j)(\theta_i - \beta_j)\}]^{-1}$ and $Q_j^*(\theta_i) = 1 - P_j^*(\theta_i)$. Using these expressions and the relationship

$$\frac{Q_j^*(\theta_i)}{Q_j(\theta_i)} = \frac{1}{1 - \Psi(\gamma_j)}, \tag{54}$$

the derivatives of the log likelihood with respect to the item parameters are

$$\frac{\partial}{\partial \alpha_j} \ln l(\underline{\theta}, \underline{\xi}) = \exp(\alpha_j)\{1 - \Psi(\gamma_j)\}\sum_{i=1}^{N}(\theta_i - \beta_j)w_{ij}\{u_{ij} - P_j(\theta_i)\}, \tag{55}$$

$$\frac{\partial}{\partial \beta_j} \ln l(\underline{\theta}, \underline{\xi}) = -\exp(\alpha_j)\{1 - \Psi(\gamma_j)\}\sum_{i=1}^{N} w_{ij}\{u_{ij} - P_j(\theta_i)\}, \tag{56}$$

and

$$\frac{\partial}{\partial \gamma_j} \ln l(\underline{\theta}, \underline{\xi}) = \Psi(\gamma_j)\sum_{i=1}^{N}\{P_j(\theta_i)\}^{-1}\{u_{ij} - P_j(\theta_i)\}, \tag{57}$$

where

$$w_{ij} = \frac{P_j^*(\theta_i)Q_j^*(\theta_i)}{P_j(\theta_i)Q_j(\theta_i)}. \tag{58}$$

The partial derivative of $P_j(\theta_i)$ with respect to θ_i is

$$\frac{\partial}{\partial \theta_i} P_j(\theta_i) = \exp(\alpha_j)\{1 - \Psi(\gamma_j)\}P_j^*(\theta_i)Q_j^*(\theta_i) \tag{59}$$

and hence the derivative of the log likelihood with respect to the ability parameter is

$$\frac{\partial}{\partial \theta_i} \ln l(\underline{\theta}, \underline{\xi}) = \sum_{j=1}^{n} \exp(\alpha_j)\{1 - \Psi(\gamma_j)\}w_{ij}\{u_{ij} - P_j(\theta_i)\}. \tag{60}$$

Second Derivatives of the Likelihood

The Newton-Raphson procedures require the second derivatives of the log posterior distribution with respect to each parameter. Following standard practice (Finney, 1971; Rao, 1973), the expectations of the second derivatives of the log likelihood for respective item parameters are

$$E\left\{\frac{\partial^2}{\partial\alpha_j^2}\ln l(\underline{\theta},\underline{\xi})\right\} = -\exp(2\alpha_j)\{1-\Psi(\gamma_j)\}^2\sum_{i=1}^{N}(\theta_i-\beta_j)^2 w_{ij}P_j^*(\theta_i)Q_j^*(\theta_i), \tag{61}$$

$$E\left\{\frac{\partial^2}{\partial\beta_j^2}\ln l(\underline{\theta},\underline{\xi})\right\} = -\exp(2\alpha_j)\{1-\Psi(\gamma_j)\}^2\sum_{i=1}^{N} w_{ij}P_j^*(\theta_i)Q_j^*(\theta_i), \tag{62}$$

$$E\left\{\frac{\partial^2}{\partial\gamma_j^2}\ln l(\underline{\theta},\underline{\xi})\right\} = -\{\Psi(\gamma_j)\}^2\{1-\Psi(\gamma_j)\}\sum_{i=1}^{N}\{P_j(\theta_i)\}^{-1}Q_j^*(\theta_i), \tag{63}$$

$$E\left\{\frac{\partial^2}{\partial\alpha_j\partial\beta_j}\ln l(\underline{\theta},\underline{\xi})\right\} = \exp(2\alpha_j)\{1-\Psi(\gamma_j)\}^2\sum_{i=1}^{N}(\theta_i-\beta_j)w_{ij}P_j^*(\theta_i)Q_j^*(\theta_i), \tag{64}$$

$$E\left\{\frac{\partial^2}{\partial\alpha_j\partial\gamma_j}\ln l(\underline{\theta},\underline{\xi})\right\} = -\exp(\alpha_j)\Psi(\gamma_j)\{1-\Psi(\gamma_j)\}^2\sum_{i=1}^{N}(\theta_i-\beta_j)w_{ij}Q_j^*(\theta_i), \tag{65}$$

and

$$E\left\{\frac{\partial^2}{\partial\beta_j\partial\gamma_j}\ln l(\underline{\theta},\underline{\xi})\right\} = \exp(\alpha_j)\Psi(\gamma_j)\{1-\Psi(\gamma_j)\}^2\sum_{i=1}^{N} w_{ij}Q_j^*(\theta_i). \tag{66}$$

The expectation of the second derivative of the log likelihood with respect to the ability parameter is

$$E\left\{\frac{\partial^2}{\partial\theta_j^2}\ln l(\underline{\theta},\underline{\xi})\right\} = -\sum_{j=1}^{n}\exp(2\alpha_j)\{1-\Psi(\gamma_j)\}^2 w_{ij}P_j^*(\theta_i)Q_j^*(\theta_i). \tag{67}$$

Derivatives of Priors

The logarithm of the prior of the item parameters is

$$\ln\pi(\underline{\xi}|\underline{\eta}^{(2)}) = \ln\pi(\underline{\alpha}|\nu_\alpha,\lambda_\alpha) + \ln\pi(\underline{\beta}|\nu_\beta,\lambda_\beta) + \ln\pi(\underline{\gamma}|\nu_\gamma,\lambda_\gamma), \tag{68}$$

where

$$\ln\pi(\underline{\alpha}|\nu_\alpha,\lambda_\alpha) \propto -\left(\frac{n+2\nu_\alpha-1}{2}\right)\ln\left\{\frac{2}{\lambda_\alpha}+\sum_{j=1}^{n}(\alpha_j-\bar{\alpha})^2\right\}, \tag{69}$$

$$\ln \pi(\underline{\beta}|\nu_\beta, \lambda_\beta) \propto -\left(\frac{n+2\nu_\beta-1}{2}\right) \ln \left\{\frac{2}{\lambda_\beta} + \sum_{j=1}^{n}(\beta_j - \bar{\beta})^2\right\}, \tag{70}$$

and

$$\ln \pi(\underline{\gamma}|\nu_\gamma, \lambda_\gamma) \propto -\left(\frac{n+2\nu_\gamma-1}{2}\right) \ln \left\{\frac{2}{\lambda_\gamma} + \sum_{j=1}^{n}(\gamma_j - \bar{\gamma})^2\right\}. \tag{71}$$

The partial derivative of the log prior of the item parameters with respect to α_j is

$$\frac{\partial}{\partial \alpha_j} \ln \pi(\underline{\alpha}|\nu_\alpha, \lambda_\alpha) \propto -\frac{1}{s_\alpha^2}(\alpha_j - \bar{\alpha}), \tag{72}$$

and similarly,

$$\frac{\partial}{\partial \beta_j} \ln \pi(\underline{\beta}|\nu_\beta, \lambda_\beta) \propto -\frac{1}{s_\beta^2}(\beta_j - \bar{\beta}), \tag{73}$$

and

$$\frac{\partial}{\partial \gamma_j} \ln \pi(\underline{\gamma}|\nu_\gamma, \lambda_\gamma) \propto -\frac{1}{s_\gamma^2}(\gamma_j - \bar{\gamma}), \tag{74}$$

where

$$s_\alpha^2 = \frac{\frac{2}{\lambda_\alpha} + \sum_{j=1}^{n}(\alpha_j - \bar{\alpha})^2}{n + 2\nu_\alpha - 1}, \tag{75}$$

$$s_\beta^2 = \frac{\frac{2}{\lambda_\beta} + \sum_{j=1}^{n}(\beta_j - \bar{\beta})^2}{n + 2\nu_\beta - 1}, \tag{76}$$

and

$$s_\gamma^2 = \frac{\frac{2}{\lambda_\gamma} + \sum_{j=1}^{n}(\gamma_j - \bar{\gamma})^2}{n + 2\nu_\gamma - 1}. \tag{77}$$

Second Derivatives of Priors

The second derivatives of the log prior of the item parameters are

$$\frac{\partial^2}{\partial \alpha_j^2} \ln \pi(\underline{\alpha}|\nu_\alpha, \lambda_\alpha) \propto -\frac{\left(1-\frac{1}{n}\right) s_\alpha^2 - 2(\alpha_j - \bar{\alpha})^2/(n+2\nu_\alpha - 1)}{s_\alpha^4}, \tag{78}$$

$$\frac{\partial^2}{\partial \beta_j^2} \ln \pi(\underline{\beta}|\nu_\beta, \lambda_\beta) \propto -\frac{\left(1-\frac{1}{n}\right) s_\beta^2 - 2(\beta_j - \bar{\beta})^2/(n+2\nu_\beta - 1)}{s_\beta^4}, \tag{79}$$

and

$$\frac{\partial^2}{\partial \gamma_j^2} \ln \pi(\underline{\gamma}|\nu_\gamma, \lambda_\gamma) \propto -\frac{\left(1-\frac{1}{n}\right) s_\gamma^2 - 2(\gamma_j - \bar{\gamma})^2/(n+2\nu_\gamma - 1)}{s_\gamma^4}. \tag{80}$$

Since

$$\ln \pi(\underline{\theta}|\underline{\tau}) \propto -\frac{1}{2}\sum_{i=1}^{n}\theta_i^2, \tag{81}$$

the partial derivative of the log prior distribution of ability parameters with respect to θ_i is

$$\frac{\partial}{\partial \theta_i}\ln \pi(\underline{\theta}|\underline{\tau}) \propto -\theta_i \tag{82}$$

and the second derivative is

$$\frac{\partial^2}{\partial \theta_i^2}\ln \pi(\underline{\theta}|\underline{\tau}) \propto -1. \tag{83}$$

Initial Values for the Newton-Raphson Method

The Newton-Raphson method typically requires close approximations to the solution as starting points. Initial values for these starting points may be obtained from the following equations (Baker, 1987; Swaminathan & Gifford, 1986):

$$\alpha_j^{(0)} = \ln a_j^{(0)} = \ln\left(\frac{1.702 r_{b_j}}{\sqrt{1 - r_{b_j}^2}}\right), \tag{84}$$

$$\beta_j^{(0)} = \frac{z_j}{r_{b_j}}, \tag{85}$$

$$\gamma_j^{(0)} = \ln\left(\frac{c_j^{(0)}}{1 - c_j^{(0)}}\right) = \ln\left(\frac{1}{m_j - 1}\right), \tag{86}$$

and

$$\theta_i^{(0)} = \ln\left(\frac{\sum_{j=1}^{n} u_{ij} + \frac{1}{2}}{n - \sum_{j=1}^{n} u_{ij} + \frac{1}{2}}\right), \tag{87}$$

where r_{b_j} is the biserial correlation of the item j and the item-excluded total score, z_j is the normal deviate $z_j = \Phi^{-1}(1 - p_j)$, Φ denotes the standard normal cumulative density function, p_j is the classical item difficulty (i.e., $p_j = \sum_{i=1}^{N} u_{ij}/N$), and m_j is the number of options in multiple choice item j.

JB1 Estimation

The difference between the two joint Bayesian estimation procedures lies in the form of the prior distributions. Since JB1 also requires the Newton-Raphson method, we need partial and second derivatives of the log likelihood and log prior distributions. When we take a partial derivative of the log posterior distribution with respect to an item parameter, say α_j, and set to zero, we obtain

$$\frac{\partial}{\partial \alpha_j} \ln l(\underline{\theta}, \underline{\xi}) + \frac{\partial}{\partial \alpha_j} \ln \pi(\underline{\xi} | \underline{\eta}) = 0. \tag{88}$$

Since the partial derivative of log likelihood function is the same as one used in JB2 estimation, we dispense with description of the likelihood part and present the elements for the item priors.

Derivatives of Priors

The term $\frac{\partial}{\partial \alpha_j} \ln \pi(\underline{\xi} | \underline{\eta})$ represents the contribution of the item priors. The partial derivatives of $\ln \pi(\underline{\xi} | \underline{\eta})$ with respect to α_j, β_j, and γ_j are

$$\frac{\partial}{\partial \alpha_j} \ln \pi(\underline{\xi} | \underline{\eta}) = -\frac{1}{\sigma_\alpha^2}(\alpha_j - \mu_\alpha), \tag{89}$$

$$\frac{\partial}{\partial \beta_j} \ln \pi(\underline{\xi} | \underline{\eta}) = -\frac{1}{\sigma_\beta^2}(\beta_j - \mu_\beta), \tag{90}$$

and

$$\frac{\partial}{\partial \gamma_j} \ln \pi(\underline{\xi} | \underline{\eta}) = -\frac{1}{\sigma_\gamma^2}(\gamma_j - \mu_\gamma). \tag{91}$$

Second Derivatives of Prior

The second derivatives of the priors for the item parameters are

$$\frac{\partial^2}{\partial \alpha_j^2} \ln \pi(\underline{\xi} | \underline{\eta}) = -\frac{1}{\sigma_\alpha^2}, \tag{92}$$

$$\frac{\partial^2}{\partial\beta_j^2}\ln\pi(\underline{\xi}|\underline{\eta}) = -\frac{1}{\sigma_\beta^2}, \tag{93}$$

and

$$\frac{\partial^2}{\partial\gamma_j^2}\ln\pi(\underline{\xi}|\underline{\eta}) = -\frac{1}{\sigma_\gamma^2}. \tag{94}$$

Empirical Study

In this section we present an empirical comparison of the three Bayesian methods. Data were simulated under the following conditions: (1) number of examinees ($N = 100, 300$), (2) number of items ($n = 15, 45$), (3) estimation (JB1, JB2, MB), and (4) prior condition (prior-α_L, prior-α_T, prior-$\alpha\beta_T$). The sample sizes and the test lengths were selected to emulate the situation in which estimation procedures and priors might have some impact upon item and ability parameter estimates. The sample size and test length, were completely crossed to yield four situations.

Three Bayesian estimation procedures were used: JB1 is the joint Bayes modal estimation procedure with known hyperparameters; JB2 is the joint Bayes modal estimation procedure with informative hyperpriors; and MB is the marginal Bayes modal estimation of item parameters with known hyperparameters and the EAP estimation of ability parameters.

Each estimation procedure had the three prior conditions: prior-α_L, prior-α_T, and prior-$\alpha\beta_T$. The prior-α_L condition used a loose prior for the transformed item discrimination; the prior-α_T condition used a tight prior for the transformed item discrimination; and the prior-$\alpha\beta_T$ condition used tight priors for both the transformed item discrimination and the item difficulty. The exact specification of the prior condition is presented in a subsequent section on the item and ability parameter estimation.

Data Generation

Using the two-parameter logistic model,

$$P_j(\theta_i) = [1 + \exp\{-a_j(\theta_i - b_j)\}]^{-1}, \tag{95}$$

dichotomous item response vectors were generated via the computer program GENIRV (Baker, 1982). Based on the usual ranges of item parameters for the two-parameter logistic model, the underlying item discrimination parameters were assumed to be normally distributed with mean 1.046 and variance 0.103, $a_j \sim \mathcal{N}(1.046, 0.103)$; that is, $\alpha_j \sim \mathcal{N}(0.0, 0.09)$. The underlying item difficulty parameters are distributed normally with mean 0.0 and variance 1.0, $b_j \sim \mathcal{N}(0, 1)$.

For data generation purposes, an approximation based on histograms was adopted. Item discrimination and item difficulty parameters for the 15-item test were set to have three different values respectively. For the 45-item test, each of the item parameters was set to have five different values. Item parameters used to generate the data sets are given in Table 1 and Table 2 for the 15-item test and for the 45-item test, respectively.

Insert Tables 1 and 2 about here

The underlying ability parameters were matched to the item difficulty distribution. Hence, a normal distribution with mean 0.0 and variance 1.0, $\theta_i \sim \mathcal{N}(0, 1)$, was used to specify the underlying ability parameters. Table 3 shows the ability groups and the number of examinees in each ability group for samples of 100 and 300.

Insert Table 3 about here

For each of the factors of sample size and test length, four replications of the simulated data were generated. Since the two factors were completely crossed, a total of 16 GENIRV runs was needed to obtain the data sets for the empirical comparison.

Item and Ability Parameter Estimation

Each of the generated data sets was analyzed via the computer program BILOG (Mislevy & Bock, 1989) for the marginal Bayesian estimation and via the computer program

JBAYES, specifically developed for this study to provide the joint Bayesian estimates. In each estimation procedure, three prior conditions, prior-α_L, prior-α_T, and prior-$\alpha\beta_T$, were employed. Hence, for example, the generated item response data set for the first replication of sample size 100 and test length 15 was analyzed by nine computer runs (three estimation procedures with three prior conditions).

The default options of the computer program BILOG (Mislevy & Bock, 1989) provide the marginal Bayesian modal estimates of item parameters and the expected a posteriori estimates of ability parameters for the two-parameter model. In the prior-α_L condition for MB, a lognormal prior with mean 0.0 and variance 0.25 was used, that is, $\ln a_j \sim \mathcal{N}(0, 0.25)$. This is, in fact, the default prior specification in BILOG for the two-parameter model. In the prior-α_T condition, a lognormal distribution with mean 0.0 and variance 0.09, $\ln a_j \sim \mathcal{N}(0, 0.09)$, was used. For the prior-$\alpha\beta_T$ condition, the same prior in the prior-α_T condition along with a normal prior was used for the item difficulty with mean 0.0 and variance 1.0, $\beta_j \sim \mathcal{N}(0, 1)$.

For JB1 estimation via JBAYES, $\alpha_j \sim \mathcal{N}(0, 0.25)$ was used for the prior-α_L condition. For the prior-α_T condition, $\alpha_j \sim \mathcal{N}(0, 0.09)$ was used. The prior-$\alpha\beta_T$, used $\alpha_j \sim \mathcal{N}(0, 0.09)$ and $\beta_j \sim \mathcal{N}(0, 1)$. For JB2 estimation, the mean hyperparameter was assumed to have a noninformative uniform distribution and the variance hyperparameter was set to have an inverse gamma distribution. In the prior-α_L condition, the inverse gamma distribution with $\nu_\alpha = 4$ and $\lambda_\alpha = 1$ was used for the variance hyperparameter of the transformed item discrimination parameters: $\sigma_\alpha^2 \sim \mathcal{IG}(4, 1)$. The inverse gamma distribution with parameters $\nu_\alpha = 11$ and $\lambda_\alpha = 1$ was used in the prior-α_T condition: $\sigma_\alpha^2 \sim \mathcal{IG}(11, 1)$. Two inverse gamma distributions with parameters $\nu_\alpha = 11$ and $\lambda_\alpha = 1$, and $\nu_\beta = 4$ and $\lambda_\beta = 0.25$ for the variance hyperparameters of the transformed item discrimination and of the item difficulty, respectively, were adopted for the prior-$\alpha\beta_T$ condition: $\sigma_\alpha^2 \sim \mathcal{IG}(11, 1)$ and $\sigma_\beta^2 \sim \mathcal{IG}(4, 0.25)$.

When the mean hyperparameter is assumed to have a fixed value, μ, then the specification of the variance hyperparameter by the inverse gamma distribution with parameters ν and λ, $\mathcal{IG}(\nu,\lambda)$, yields the parameter of interest which is distributed as a t with mean μ, variance $\frac{1}{\nu\lambda}$, and degrees of freedom 2ν, that is, $\mathcal{T}(2\nu,\mu,\frac{1}{\nu\lambda})$ (Berger, 1985). Therefore, for the transformed item discrimination, assuming the mean hyperparameter μ_α has a fixed value, specification of the hyperparameter of variance by the inverse gamma with $\nu_\alpha = 4$ and $\lambda_\alpha = 1$ yields a transformed item discrimination parameter which is distributed as a t with mean μ_α, variance $\frac{1}{\nu_\alpha\lambda_\alpha} = 0.25$, and degrees of freedom $2\nu_\alpha = 8$, that is, $\alpha_j \sim \mathcal{T}(8,\mu_\alpha,0.25)$. Similarly, the specification $\sigma_\alpha^2 \sim \mathcal{IG}(11,1)$ implies $\alpha_j \sim \mathcal{T}(22,\mu_\alpha,0.09)$; and the specification $\sigma_\beta^2 \sim \mathcal{IG}(4,0.25)$ yields $\beta_j \sim \mathcal{T}(8,\mu_\beta,1)$. In the above illustration, because we assumed a noninformative prior for the mean hyperparameter, the specifications used in JB2 will not produce the same specifications of item hyperparameters used in the marginal Bayesian and the joint Bayesian-1 procedures. These specifications are $\sigma_\alpha^2 \sim \mathcal{IG}(4,1)$, $\sigma_\alpha^2 \sim \mathcal{IG}(11,1)$, and $\sigma_\beta^2 \sim \mathcal{IG}(4,0.25)$ and are similar to their counterparts in the MB and JB1 estimation procedures.

The EAP estimation was used in MB for the ability estimation via BILOG. Bayes modal estimation was employed in the ability estimation for both joint Bayesian procedures via JBAYES. All three Bayesian estimation procedures used a standard normal distribution as the prior for the ability parameters.

Metric Transformation

In parameter recovery studies, such as the present one. comparisons between two or more sets of estimates and the underlying parameters require that the item and ability estimates obtained from different calibration runs and their parameters be placed on a common metric (Baker & Al-Karni, 1991; Yen, 1987). Parameter estimation procedures under IRT yield metrics which are unique up to a linear transformation. To link both sets of estimates and parameters, it is necessary to determine the slope and intercept of the equating coefficients required for the transformation. The estimates of the item and ability parameters for each of the estimation procedures were placed on the scale of the true parameters using test characteristic curve method by Stocking and Lord (1983) as implemented in the computer program EQUATE (Baker, 1990).

Criteria

The empirical comparisons in this study involved three criteria: root mean square differences (RMSD), correlation, and bias. RMSD is the square root of the average of the squared differences between estimated and true values. For item discrimination, for example, RMSD is defined as

$$\sqrt{\frac{1}{n}\sum_{j=1}^{n}(\hat{a}_j - a_j)^2}. \tag{96}$$

The bias B_a of a point estimator $\hat{a}$ is given by $B_a = E(\hat{a}) - a$; the bias for item difficulty is given by $B_b = E(\hat{b}) - b$; and the bias for the ability estimator is defined by $B_\theta = E(\hat{\theta}) - \theta$ (Mendenhall, Scheaffer, & Wackerly, 1981). For the 15-item test, B_a (or B_b) was obtained with regard to the three different underlying parameters across the four replications. For the 45-item test, B_a (or B_b) was calculated with regard to the five different underlying parameters across the four replications. The bias B_θ was obtained for the 11 ability levels over the four replications.

Results

RMSD and Correlation Results

RMSD and Correlation Results for Item Discrimination. RMSDs of item discriminations for each data set are reported in Table 4. As sample size increased, RMSDs decreased; marginal RMSD means were 0.24924 and 0.20646 for sample sizes 100 and 300, respectively.

Insert Tables 4 and 5 about here

MB yielded smaller RMSDs than either of the two joint Bayesian procedures. For the two joint Bayesian procedures, JB1 yielded larger RMSDs. Increasing the number of items reduced the size of RMSDs, particularly for JB1 and JB2. For the 15-item test, MB yielded smaller RMSD values although all three estimation methods produced nearly the same values for the 45-item test. RMSDs for the third replication of the sample size 100 and 15-item test were slightly smaller than for the other cases and RMSDs for the fourth replication of the sample size 100 and 15-item test were slightly larger than for the other cases. These differences were probably due to sampling fluctuations in the data generation procedures used in this study. The effect of this probable sampling fluctuation could also be seen for the respective correlations in Table 5.

When the loose prior was used in JB1 and JB2, it yielded comparatively larger values of RMSD than did either of the tight prior conditions. This was particularly the case for the short 15-item test.

The correlations between true and estimated values of item discriminations are given in Table 5. For each data set, the three Bayesian estimation procedures yielded practically the same correlations. Generally, the larger the sample sizes the higher correlations. Also, increasing the number of items tended to produce slightly higher correlations. For the three prior condition used, there seemed no definitive tendency observed in the correlations.

RMSD and Correlation Results for Item Difficulty. Table 6 contains RMSDs for item difficulty. The pattern of results was nearly the same as that for item discrimination. An increase in sample size appeared to be associated with a decrease in the size of RMSDs. For JB1 and JB2, increasing the number of items appeared to slightly decrease RMSDs. The values of RMSD from MB were nearly the same regardless of the test size. MB consistently yielded the smallest RMSDs.

Prior-$\alpha\beta_T$ condition yielded a relatively smaller RMSDs than did either the prior-α_L or prior-α_T conditions. MB consistently yielded smaller RMSDs than JB1 and JB2 regardless the prior condition employed.

Insert Tables 6 and 7 about here

For each data set, the three estimation procedures yielded nearly the same correlations between estimates and parameters (see Table 7). Generally, the larger sample sizes yielded higher correlations. Increasing the number of items tended to produce slightly higher correlations. There seemed to be no definitive trends in the correlations among the three prior conditions.

RMSD and Correlation Results for Ability. The RMSD results between ability estimates and the underlying parameters are reported in Table 8. As expected, RMSD values were much smaller for the 45-item test than for the 15-item test. Smaller values were consistently obtained for MB than for either of the two joint Bayesian procedures. The differences between MB and either of the two joint Bayesian procedures were particularly noticeable with the short test. As the number of items increased, the differences in RMSDs among the three estimation procedures appeared to decrease.

Insert Tables 8 and 9 about here

Prior conditions did not have an apparent impact on the size of RMSD values for ability. This might be expected as the prior conditions used were manipulated only with respect to item parameters.

The correlations between the ability estimates and the true values are reported in Table 9. The correlations were nearly identical across the three estimation procedures for each data set. The 45-item test yielded higher correlations than the 15-item test. The prior conditions did not seem to affect the correlations between the ability estimates and the underlying parameters. As was the case with RMSD results, the prior used in the context of the item parameter estimation had minimal effect when estimating ability parameters.

Bias Results

Bias Results for Item Discrimination. The bias results for item discrimination, presented in Table 10, appear to reflect influence by a number of factors. Each bias statistic was obtained by combining all four replications together; that is, the numbers of items used to obtain bias values were 16, 28, and 16, for a = 0.66, 1.00, and 1.51, respectively, for the 15-item test. For the 45-item test, 16, 36, 76, 36, and 16 items were used for a = 0.57, 0.76, 1.00, 1.32, and 1.77, respectively.

Insert Table 10 about here

For each test length, increasing the sample size resulted in a decrease in bias values. In general, positive bias values were observed for the smaller item discrimination parameters (i.e., a = 0.66 for the 15-item test, and a = 0.57 and 0.76 for the 45-item test) due to the regression toward the mean of the prior distribution. Negative values of bias were obtained for the relatively larger item discrimination parameters (i.e., a = 1.51 for the 15-item test, and a = 1.32 and 1.77 for the 45-item test). This shrinkage effect can be observed for all data sets except when the loose prior on item discrimination (prior-α_{L}) was used for the 15-item test. When a large sample size was used with 45-item test, all three estimation procedures yielded similar results.

For the three different levels of item discrimination, both JB1 and JB2 produced more positive bias for the 15-item test than did MB. The two tight prior conditions, prior-α_{T} and prior-$\alpha\beta_{\mathrm{T}}$, yielded similar pattern of bias for all data sets.

Bias Results for Item Difficulty. The bias results for item difficulty are reported in Table 11. The pattern of results was somewhat different from that for item discrimination. For the 15-item test, the two joint Bayesian methods yielded negative bias values for the easy items (b = −1.38) and positive bias values for the difficult items (b = 1.38). When both priors on item difficulty and item discrimination were used, the same pattern was observed. Even though the test size and sample size increased, the same pattern was observed for three methods of estimation. MB yielded the smallest bias for all item difficulty levels in all data sets.

Insert Table 11 about here

Bias Results for Ability. The bias results for ability from the 100-examinee-15-item data set are presented in Table 12. Those for the 100-examinee-45-item, 300-examinee-15-item, and 300-examinee-45-item data sets are presented in Tables 13, 14, and 15, respectively. It can be seen from these tables that shrinkage was more evident when a small number of items was used. The prior conditions employed in item parameter estimation did not produce any difference among the bias results. The expected a posteriori estimation of ability employed in MB yielded consistently larger sizes of bias than the Bayes modal method used in the two joint Bayesian methods. JB1 and JB2 yielded nearly the same pattern of bias for all data sets. JB2 yielded relatively smaller values of bias, however, then the other two methods. It should be noted that the bias values for the different ability levels were obtained by combining the four replications.

Insert Tables 12, 13, 14, and 15 about here

Discussion

Maximum likelihood approaches in IRT suffer from a number of problems, an important one being the possibility that unreasonable values will be obtained for parameter estimates, particularly for item discrimination and pseudo-guessing. In addition, these approaches perform poorly when estimating item and ability parameters for unusual response patterns such as all correct or all incorrect answers. These problems have led to interest in the development of Bayesian approaches for estimation of item and ability parameters. In the present study, we used a recovery study approach to compare parameter estimates obtained via a marginal Bayesian algorithm, MB, and two joint Bayesian algorithms, JB1 and JB2.

Analysis of item parameter recovery results indicated that MB yielded parameter estimates which were generally better than those obtained from JB1 or JB2. RMSD and Bias results for item discrimination and difficulty were smaller for MB estimates. JB1 and JB2 estimates were similar although JB2 results were slightly better. These differences were primarily evident in the small sample and short test conditions. This superiority was likely due to the fact that MB permits item parameters to be estimated without the concurrent need to estimate ability. Differences due to sample size are interesting if only for the fact that the two sample sizes simulated in the present study, 100 and 300 examinees, were both relatively small. In reality, all three Bayesian methods performed well, yielding item parameter estimates which were not markedly different from the underlying values. Failure of the joint Bayesian methods to provide estimates as accurate as MB under these conditions should not be viewed as something that indicates a serious deficiencies for the joint Bayesian methods. Rather, what these results suggest is that marginalized Bayesian solutions are relatively powerful under the somewhat extreme conditions simulated in the present study.

The EAP ability estimates obtained via MB were more accurate in terms of RMSD than those from either of the two joint methods. The bias values for EAP estimates, however, were larger than Bayes model estimates of ability for JB1 and JB2. This is well-known result and demonstrates the impact of the use of the posterior mean in the EAP estimation rather than the posterior mode (Bock & Mislevy, 1982).

The effectiveness of the marginalization in MB may depend in part on the accuracy of the ability hyperparameters. Seong (1991) has shown that item parameter estimates from the marginalized distribution are sensitive to misspecification of the ability distributions. In this study we generated the ability had standard normal distribution. Consequently, the marginalization of the posterior distribution was performed under an optimal situation.

Both the shape and the variance of the prior distribution play a part in the estimation of parameters. The more informative the prior (i.e., the smaller the variance), the more the parameter estimate tends to be pulled toward the mean of the prior. The tight prior conditions used in the present study, prior-α_T and prior-$\alpha\beta_T$, yielded better item parameter estimates than did the loose prior, prior-α_L. The use of tight priors seems appropriate when there is strong a priori information about the parameters. In the MB context, the misspecification of prior information has not been found to be a serious problem except when the mean of the underlying item discrimination parameters was quite smaller than the mean of the prior (Al-Karni, 1990).

Incorrect specification of the prior may result in more serious consequences for JB1 and MB than for JB2. This condition was not tested in the present study because priors were relatively well-matched to the generated data sets.

Several issues remain to be studies in the present context. In particular, little has been done on the shrinkage effect except for Al-Karni (1990) and Gifford and Swaminathan (1990). Neither are the effects of priors well-known with respect to the robustness of two-stage hierarchical models. This kind of research is particularly valuable for small samples and short tests. Marginal Bayesian estimation was arguably the more desirable algorithm in the present study. Even so, it remains to be seen whether incorporation of a two-stage hierarchical procedure might improve marginal Bayes modal estimates.

References

Al-Karni, A. (1990). *The impact of test structure on Bayesian test analysis*. Unpublished doctoral dissertation, University of Wisconsin–Madison.

Baker, F. B. (1982). *GENIRV: A program to generate item response vectors*. Unpublished manuscript, University of Wisconsin–Madison, Laboratory of Experimental Design, Madison.

Baker, F. B. (1987a). Methodology Review: Item parameter estimation under the one-, two-, and three-parameter logistic models. *Applied Psychological Measurement, 11*, 111-141.

Baker, F. B. (1990). *EQUATE* (computer program). University of Wisconsin–Madison, Department of Educational Psychology, Laboratory of Experimental Design, Madison.

Baker, F. B., & Al-Karni, A. (1991). A comparison of two procedures for computing IRT equating coefficients. *Journal of Educational Measurement*, *28*, 147-162.

Berger, J. O. (1985). *Statistical decision theory and Bayesian analysis* (2nd ed.) New York: Springer-Verlag.

Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F. M. Lord & M. R. Novick, *Statistical theories of mental test scores* (pp. 395-479). Reading, MA: Addison-Wesley.

Bock, R. D., & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. *Psychometrika*, *46*, 443-459.

Bock, R. D., & Lieberman, M. (1970). Fitting a response model for n dichotomously scored items. *Psychometrika*, *35*, 179-197.

Bock, R. D., & Mislevy, R. J. (1982). Adaptive EAP estimation of ability in a microcomputer environment. *Applied Psychological Measurement*, *4*, 431-444.

Finney, D. J. (1971). *Probit analysis* (3rd ed.). Cambridge, England: Cambridge University Press.

Gifford, J. A., & Swaminathan, H. (1990). Bias and the effect of priors in Bayesian estimation of parameters of item response models. *Applied Psychological Measurement*, *14*, 33-43.

Goel, P. K. (1983). Information measures and Bayesian hierarchical models. *Journal of the American Statistical Association*, *78*, 408-410.

Goel, P. K., & DeGroot, M. H. (1981). Information about hyperparameters in hierarchical models. *Journal of the American Statistical Association*, *76*, 140-147.

Good, I. J. (1983). The robustness of a hierarchical model for multinomials and contingency tables. In G. E. P. Box, T. Leonard, & C. F. Wu (Eds.), *Scientific inference, data analysis, and robustness* (pp. 191-211). New York: Academic Press.

Harwell, M. R., & Baker, F. B. (1991). The use of prior distributions in marginalized Bayesian item parameter estimation: A didactic. *Applied Psychological Measurement*, *15*, 375-389.

Harwell, M. R., Baker, F. B., & Zwarts, M. (1988). Item parameter estimation via marginal maximum likelihood and EM algorithm: A didactic. *Journal of Educational Statistics*, *13*, 243-271.

Kennedy, W. J., Jr., & Gentle, J. E. (1980). *Statistical computing*. New York: Marcel Dekker.

Leonard, T. (1972). *Bayesian methods for multinominal data* (ACT Technical Bulletin No. 4). Iowa City, IA: The American College Testing Program.

Leonard, T., & Novick, M. R. (1985). *Bayesian inference and diagnostics for the three parameter logistic model* (ONR Technical Report 85-5). Iowa City, IA: The University of Iowa, CADA Research Group. (ERIC Document Reproduction Service No. ED 261 068)

Lindley, D. V. (1971). The estimation of many parameters. In V. P. Godambe & D. A. Sprott (Eds.), *Foundations of statistical inference* (pp. 435-455). Toronto: Holt, Rinehart & Winston of Canada.

Lindley, D. V., & Smith, A. F. M. (1972). Bayes estimates for the linear model (with discussion). *Journal of the Royal Statistical Society, Series B*, *34*, 1-41.

Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Lord, F. M. (1986). Maximum likelihood and Bayesian parameter estimation in item response theory. *Journal of Educational Measurement*, *23*, 157-162.

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.

Mendenhall, W., Scheaffer, R. L., & Wackerly, D. D. (1981). *Mathematical statistics with application*. Boston, MA: Duxbury Press.

Mislevy, R. J. (1986). Bayes modal estimation in item response models. *Psychometrika*, *51*, 177-195.

Mislevy, R. J., & Bock, R. D. (1989). *PC-BILOG 3: Item analysis and test scoring with binary logistic models*. Mooresville, IN: Scientific Software.

Novick, M. R. (1969). Multiparameter Bayesian indifference procedures (with discussion). *Journal of the Royal Statistical Society, Series B*, *31*, 29-64.

O'Hagan, A. (1976). On posterior joint and marginal modes. *Biometrika*, *63*, 329-333.

Rao, C. R. (1973). *Linear statistical inference and its applications* (2nd ed.). New York: John Wiley & Sons.

Seong, T.-J. (1990). Sensitivity of marginal maximum likelihood estimation of item and ability parmaters to the characteristics of the prior ability distributions. *Applied Psychological Measurement*, *14*, 299-311.

Stocking, M. L., & Lord, F. M. (1983). Developing a common metric in item response theory. *Applied Psychological Measurement*, *7*, 201-210.

Swaminathan, H., & Gifford, J. A. (1982). Bayesian estimation in the Rasch model. *Journal of Educational Statistics*, 7, 175-191.

Swaminathan, H., & Gifford, J. A. (1985). Bayesian estimation in the two-parameter logistic model. *Psychometrika*, *50*, 349-364.

Swaminathan, H., & Gifford, J. A. (1986). Bayesian estimation in the three-parameter logistic model. *Psychometrika*, *51*, 581-601.

Tsutakawa, R. K., & Lin, H. Y. (1986). Bayesian estimation of item response curves. *Psychometrika*, *51*, 251-267.

Wingersky, M. S., Barton, M. A., & Lord, F. M. (1982). *LOGIST user's guide*. Princeton, NJ: Educational Testing Service.

Yen, W. M. (1987). A comparison of the efficiency and accuracy of BILOG and LOGIST. *Psychometrika*, *52*, 275-291.

Table 1: Item Discrimination and Item Difficulty Parameters for 15-Item Test

Item	Discrimination[a]	Difficulty
1	0.66 (-0.41)	-1.38
2	0.66 (-0.41)	0.00
3	0.66 (-0.41)	0.00
4	0.66 (-0.41)	1.38
5	1.00 (0.00)	-1.38
6	1.00 (0.00)	-1.38
7	1.00 (0.00)	0.00
8	1.00 (0.00)	0.00
9	1.00 (0.00)	0.00
10	1.00 (0.00)	1.38
11	1.00 (0.00)	1.38
12	1.51 (0.41)	-1.38
13	1.51 (0.41)	0.00
14	1.51 (0.41)	0.00
15	1.51 (0.41)	1.38

[a]Parentheses contain the transformed item discrimination.

Table 2: Item Discrimination and Item Difficulty Parameters for 45-Item Test

Item	Discrimination[a]	Difficulty
1	0.57 (-0.57)	-0.95
2–3	0.57 (-0.57)	0.00
4	0.57 (-0.57)	0.95
5	0.76 (-0.28)	-1.90
6–7	0.76 (-0.28)	-0.95
8–10	0.76 (-0.28)	0.00
11–12	0.76 (-0.28)	0.95
13	0.76 (-0.28)	1.90
14–15	1.00 (0.00)	-1.90
16–18	1.00 (0.00)	-0.95
19–27	1.00 (0.00)	0.00
28–30	1.00 (0.00)	0.95
31–32	1.00 (0.00)	1.90
33	1.32 (0.28)	-1.90
34–35	1.32 (0.28)	-0.95
36–38	1.32 (0.28)	0.00
39–40	1.32 (0.28)	0.95
41	1.32 (0.28)	1.90
42	1.77 (0.57)	-0.95
43–44	1.77 (0.57)	0.00
45	1.77 (0.57)	0.95

[a]Parentheses contain the transformed item discrimination.

Table 3: Number of Examinees at Each of the 11 Ability Levels

	Number of Examinees	
θ Level	$N = 100$	$N = 300$
-2.5	1	4
-2.0	3	8
-1.5	7	20
-1.0	12	36
-0.5	17	52
0.0	20	60
0.5	17	52
1.0	12	36
1.5	7	20
2.0	3	8
2.5	1	4

Table 4: Root Mean Square Differences of Item Discrimination

N	n	r^a	Joint Bayesian-1 α_L	Joint Bayesian-1 α_T	Joint Bayesian-1 $\alpha\beta_T$	Joint Bayesian-2 α_L	Joint Bayesian-2 α_T	Joint Bayesian-2 $\alpha\beta_T$	Marginal Bayesian α_L	Marginal Bayesian α_T	Marginal Bayesian $\alpha\beta_T$
100	15	1	0.315	0.254	0.251	0.263	0.250	0.253	0.272	0.251	0.258
		2	0.282	0.231	0.211	0.245	0.229	0.214	0.253	0.208	0.225
		3	0.349	0.217	0.194	0.285	0.225	0.215	0.181	0.158	0.175
		4	0.332	0.337	0.295	0.301	0.294	0.291	0.313	0.290	0.296
	45	1	0.270	0.233	0.241	0.234	0.240	0.239	0.261	0.227	0.233
		2	0.241	0.233	0.233	0.239	0.249	0.250	0.252	0.240	0.249
		3	0.313	0.264	0.261	0.261	0.264	0.263	0.299	0.259	0.266
		4	0.225	0.209	0.206	0.215	0.228	0.228	0.206	0.197	0.204
300	15	1	0.204	0.195	0.188	0.199	0.199	0.195	0.152	0.160	0.167
		2	0.329	0.184	0.174	0.195	0.179	0.173	0.178	0.169	0.176
		3	0.595	0.288	0.277	0.533	0.291	0.281	0.277	0.211	0.209
		4	0.755	0.231	0.228	0.260	0.229	0.228	0.212	0.191	0.189
	45	1	0.155	0.137	0.134	0.138	0.136	0.137	0.151	0.132	0.134
		2	0.203	0.189	0.183	0.188	0.180	0.181	0.199	0.182	0.182
		3	0.166	0.152	0.151	0.152	0.153	0.153	0.164	0.151	0.153
		4	0.206	0.179	0.172	0.178	0.171	0.169	0.208	0.171	0.174

[a]Number of Examinees (N), Number of Items (n), and Replication (r).

Table 5: Correlations Between Estimates and Parameters for Item Discrimination

N	n	r[a]	Joint Bayesian-1 α_L	Joint Bayesian-1 α_T	Joint Bayesian-1 $\alpha\beta_T$	Joint Bayesian-2 α_L	Joint Bayesian-2 α_T	Joint Bayesian-2 $\alpha\beta_T$	Marginal Bayesian α_L	Marginal Bayesian α_T	Marginal Bayesian $\alpha\beta_T$
100	15	1	0.615	0.612	0.614	0.616	0.610	0.592	0.618	0.621	0.612
		2	0.770	0.745	0.772	0.771	0.740	0.774	0.748	0.748	0.703
		3	0.809	0.826	0.852	0.817	0.828	0.840	0.879	0.893	0.862
		4	0.383	0.362	0.385	0.389	0.372	0.388	0.423	0.423	0.400
	45	1	0.695	0.677	0.691	0.677	0.687	0.698	0.700	0.695	0.683
		2	0.674	0.678	0.685	0.679	0.676	0.683	0.665	0.659	0.641
		3	0.526	0.559	0.564	0.559	0.569	0.572	0.566	0.577	0.562
		4	0.742	0.752	0.765	0.757	0.773	0.771	0.783	0.796	0.796
300	15	1	0.869	0.865	0.874	0.870	0.865	0.872	0.878	0.872	0.857
		2	0.865	0.860	0.870	0.860	0.869	0.874	0.846	0.845	0.831
		3	0.688	0.761	0.761	0.701	0.760	0.761	0.766	0.776	0.780
		4	0.574	0.767	0.766	0.758	0.765	0.750	0.784	0.798	0.804
	45	1	0.906	0.905	0.909	0.906	0.906	0.908	0.906	0.908	0.906
		2	0.821	0.813	0.819	0.815	0.822	0.820	0.817	0.820	0.819
		3	0.879	0.878	0.880	0.878	0.880	0.882	0.877	0.878	0.876
		4	0.843	0.843	0.848	0.845	0.845	0.848	0.850	0.854	0.850

[a]Number of Examinees (N), Number of Items (n), and Replication (r).

Table 6: Root Mean Square Differences of Item Difficulty

N	n	r[a]	Joint Bayesian-1			Joint Bayesian-2			Marginal Bayesian		
			α_L	α_T	$\alpha\beta_T$	α_L	α_T	$\alpha\beta_T$	α_L	α_T	$\alpha\beta_T$
100	15	1	0.374	0.360	0.344	0.374	0.375	0.355	0.325	0.310	0.308
		2	0.379	0.361	0.327	0.381	0.389	0.359	0.259	0.263	0.248
		3	0.481	0.499	0.462	0.498	0.522	0.496	0.382	0.402	0.388
		4	0.346	0.337	0.310	0.344	0.355	0.333	0.292	0.290	0.286
	45	1	0.370	0.346	0.313	0.349	0.342	0.319	0.335	0.312	0.294
		2	0.314	0.306	0.298	0.316	0.319	0.308	0.304	0.301	0.299
		3	0.314	0.303	0.251	0.308	0.310	0.272	0.269	0.260	0.246
		4	0.330	0.308	0.274	0.314	0.315	0.289	0.282	0.276	0.272
300	15	1	0.347	0.334	0.320	0.345	0.343	0.333	0.167	0.170	0.165
		2	0.330	0.301	0.283	0.316	0.304	0.292	0.172	0.174	0.174
		3	0.344	0.329	0.295	0.343	0.330	0.305	0.222	0.188	0.186
		4	0.213	0.203	0.192	0.222	0.211	0.198	0.133	0.121	0.120
	45	1	0.157	0.189	0.174	0.192	0.193	0.180	0.158	0.153	0.152
		2	0.226	0.209	0.197	0.213	0.208	0.198	0.208	0.188	0.184
		3	0.262	0.255	0.236	0.257	0.257	0.242	0.232	0.228	0.228
		4	0.227	0.215	0.194	0.219	0.220	0.203	0.189	0.174	0.171

[a]Number of Examinees (N), Number of Items (n), and Replication (r).

Table 7: Correlations Between Estimates and Parameters for Item Difficulty

	Jcint Bayesian-1			Joint Bayesian-2			Marginal Bayesian		
N n r^a	α_L	α_T	$\alpha\beta_T$	α_L	α_T	$\alpha\beta_T$	α_L	α_T	$\alpha\beta_T$
100 15 1	0.946	0.950	0.951	0.950	0.951	0.953	0.946	0.951	0.952
2	0.976	0.975	0.976	0.977	0.974	0.976	0.976	0.975	0.973
3	0.937	0.935	0.937	0.937	0.934	0.936	0.939	0.936	0.935
4	0.955	0.957	0.961	0.960	0.960	0.961	0.957	0.959	0.959
45 1	0.948	0.955	0.958	0.958	0.961	0.962	0.949	0.957	0.959
2	0.953	0.955	0.956	0.955	0.955	0.955	0.953	0.955	0.955
3	0.970	0.972	0.976	0.972	0.972	0.975	0.970	0.972	0.973
4	0.956	0.960	0.964	0.961	0.961	0.963	0.961	0.963	0.963
300 15 1	0.993	0.993	0.993	0.993	0.992	0.992	0.993	0.993	0.992
2	0.985	0.987	0.988	0.987	0.987	0.988	0.988	0.987	0.986
3	0.968	0.976	0.979	0.970	0.976	0.979	0.975	0.982	0.983
4	0.989	0.993	0.993	0.992	0.993	0.994	0.992	0.994	0.994
45 1	0.988	0.988	0.989	0.988	0.988	0.989	0.988	0.989	0.989
2	0.980	0.983	0.984	0.983	0.984	0.984	0.979	0.983	0.983
3	0.972	0.974	0.976	0.974	0.974	0.976	0.973	0.974	0.974
4	0.985	0.986	0.987	0.986	0.986	0.987	0.984	0.986	0.986

[a]Number of Examinees (N), Number of Items (n), and Replication (r).

Table 8: Root Mean Square Differences of Ability

	Joint Bayesian-1			Joint Bayesian-2			Marginal Bayesian		
N n r[a]	α_L	α_T	$\alpha\beta_T$	α_L	α_T	$\alpha\beta_T$	α_L	α_T	$\alpha\beta_T$
100 15 1	0.501	0.492	0.501	0.514	0.512	0.520	0.456	0.455	0.454
2	0.534	0.522	0.532	0.542	0.533	0.547	0.497	0.491	0.491
3	0.570	0.554	0.565	0.572	0.565	0.572	0.537	0.536	0.536
4	0.609	0.582	0.589	0.616	0.603	0.610	0.546	0.538	0.538
45 1	0.316	0.327	0.334	0.334	0.336	0.344	0.315	0.319	0.320
2	0.352	0.350	0.360	0.362	0.360	0.376	0.342	0.341	0.341
3	0.310	0.308	0.314	0.310	0.311	0.317	0.314	0.310	0.309
4	0.298	0.302	0.308	0.309	0.308	0.317	0.286	0.291	0.292
300 15 1	0.552	0.549	0.554	0.556	0.555	0.558	0.521	0.517	0.517
2	0.552	0.551	0.556	0.557	0.559	0.562	0.521	0.522	0.523
3	0.565	0.557	0.560	0.566	0.560	0.563	0.539	0.537	0.538
4	0.582	0.548	0.553	0.557	0.560	0.568	0.498	0.498	0.498
45 1	0.325	0.323	0.325	0.325	0.325	0.328	0.320	0.318	0.318
2	0.337	0.339	0.344	0.344	0.345	0.349	0.326	0.328	0.328
3	0.304	0.305	0.308	0.306	0.306	0.309	0.302	0.301	0.301
4	0.339	0.339	0.341	0.342	0.342	0.345	0.330	0.329	0.329

[a]Number of Examinees (N), Number of Items (n), and Replication (r).

Table 9: Correlations Between Estimates and Parameters for Ability

N	n	r[a]	Joint Bayesian-1 α_L	Joint Bayesian-1 α_T	Joint Bayesian-1 $\alpha\beta_T$	Joint Bayesian-2 α_L	Joint Bayesian-2 α_T	Joint Bayesian-2 $\alpha\beta_T$	Marginal Bayesian α_L	Marginal Bayesian α_T	Marginal Bayesian $\alpha\beta_T$
100	15	1	0.890	0.893	0.893	0.893	0.894	0.893	0.892	0.893	0.894
		2	0.868	0.871	0.871	0.870	0.872	0.872	0.870	0.874	0.873
		3	0.847	0.849	0.849	0.848	0.848	0.849	0.848	0.849	0.848
		4	0.849	0.852	0.853	0.853	0.853	0.854	0.851	0.853	0.853
	45	1	0.951	0.948	0.948	0.948	0.947	0.947	0.950	0.949	0.949
		2	0.944	0.945	0.944	0.944	0.944	0.944	0.944	0.944	0.944
		3	0.951	0.952	0.952	0.952	0.951	0.951	0.950	0.951	0.951
		4	0.958	0.957	0.958	0.957	0.956	0.957	0.959	0.958	0.958
300	15	1	0.854	0.855	0.855	0.855	0.855	0.855	0.857	0.857	0.857
		2	0.856	0.856	0.857	0.856	0.857	0.857	0.855	0.855	0.854
		3	0.837	0.844	0.845	0.838	0.844	0.845	0.845	0.845	0.845
		4	0.858	0.873	0.874	0.973	0.874	0.873	0.874	0.873	0.873
	45	1	0.948	0.948	0.949	0.948	0.948	0.949	0.948	0.949	0.949
		2	0.947	0.947	0.947	0.947	0.947	0.947	0.947	0.946	0.946
		3	0.954	0.954	0.954	0.954	0.954	0.954	0.954	0.954	0.954
		4	0.944	0.944	0.944	0.944	0.944	0.945	0.945	0.945	0.945

[a]Number of Examinees (N), Number of Items (n), and Replication (r).

Table 10: Bias Results for Item Discrimination

N	n	a[a]	Joint Bayesian-1 α_L	Joint Bayesian-1 α_T	Joint Bayesian-1 $\alpha\beta_T$	Joint Bayesian-2 α_L	Joint Bayesian-2 α_T	Joint Bayesian-2 $\alpha\beta_T$	Marginal Bayesian α_L	Marginal Bayesian α_T	Marginal Bayesian $\alpha\beta_T$
100	15	0.66	0.27	0.30	0.28	0.30	0.33	0.32	0.19	0.23	0.22
		1.00	0.06	0.04	0.03	0.05	0.06	0.05	-0.02	-0.03	-0.05
		1.51	0.08	-0.13	-0.16	-0.03	-0.18	-0.19	-0.07	-0.27	-0.30
	45	0.57	0.20	0.26	0.24	0.29	0.32	0.31	0.16	0.23	0.22
		0.76	0.11	0.14	0.14	0.17	0.19	0.18	0.08	0.12	0.10
		1.00	0.07	0.04	0.02	0.04	0.03	0.02	0.03	0.01	-0.01
		1.32	-0.04	-0.13	-0.13	-0.16	-0.20	-0.19	-0.04	-0.15	-0.19
		1.77	-0.25	-0.33	-0.37	-0.38	-0.54	-0.57	-0.25	-0.44	-0.46
300	15	0.66	0.16	0.20	0.19	0.19	0.21	0.21	0.08	0.12	0.12
		1.00	0.07	0.07	0.06	0.07	0.07	0.07	-0.02	-0.02	-0.03
		1.51	0.33	0.03	0.02	0.16	0.02	0.00	-0.02	-0.15	-0.16
	45	0.57	0.06	0.12	0.11	0.12	0.15	0.14	0.02	0.08	0.08
		0.76	0.11	0.13	0.13	0.13	0.14	0.14	0.07	0.09	0.08
		1.00	0.03	0.02	0.02	0.02	0.02	0.02	0.01	-0.01	-0.01
		1.32	-0.01	-0.07	-0.07	-0.06	-0.09	-0.09	-0.04	-0.10	-0.11
		1.77	0.09	-0.05	-0.09	-0.06	-0.12	-0.16	0.07	-0.09	-0.09

[a]Number of Examinees (N), Number of Items (n), and Discrimination (a).

Table 11: Bias Results for Item Difficulty

N n b[a]	Joint Bayesian-1 α_L	Joint Bayesian-1 α_T	Joint Bayesian-1 $\alpha\beta_T$	Joint Bayesian-2 α_L	Joint Bayesian-2 α_T	Joint Bayesian-2 $\alpha\beta_T$	Marginal Bayesian α_L	Marginal Bayesian α_T	Marginal Bayesian $\alpha\beta_T$
100 15 -1.38	-0.19	-0.18	-0.14	-0.22	-0.24	-0.20	0.03	0.00	0.05
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1.38	0.19	0.20	0.15	0.23	0.23	0.20	-0.01	0.00	-0.05
45 -1.90	-0.10	-0.11	0.02	-0.16	-0.17	-0.08	0.05	0.00	0.06
-0.95	-0.13	-0.11	-0.09	-0.12	-0.12	-0.11	-0.06	-0.05	-0.01
0.00	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.01
0.95	0.11	0.10	0.08	0.12	0.11	0.11	0.05	0.05	0.00
1.90	-0.01	0.01	-0.10	0.06	0.07	-0.02	-0.15	-0.10	-0.17
300 15 -1.38	-0.28	-0.28	-0.26	-0.29	-0.29	-0.28	-0.01	-0.01	0.00
0.00	0.00	0.00	0.01	0.02	0.01	0.01	0.01	0.01	0.01
1.38	0.28	0.28	0.26	0.29	0.29	0.27	0.04	0.03	0.01
45 -1.90	-0.09	-0.11	-0.04	-0.12	-0.13	-0.07	0.07	0.04	0.06
-0.95	-0.15	-0.14	-0.14	-0.15	-0.15	-0.14	-0.07	-0.06	-0.04
0.00	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
0.95	0.04	0.04	0.04	0.04	0.05	0.05	-0.05	-0.04	-0.05
1.90	0.17	0.16	0.09	0.17	0.18	0.12	0.00	0.01	0.00

[a]Number of Examinees (N), Number of Items (n), and Difficulty (b).

48

Table 12: Bias Results for Ability from 100-Examinee-15-Item Data Set

	Joint Bayesian-1			Joint Bayesian-2			Marginal Bayesian		
θ Level	α_L	α_T	$\alpha\beta_T$	α_L	α_T	$\alpha\beta_T$	α_L	α_T	$\alpha\beta_T$
-2.5	0.17	0.24	0.18	0.09	0.12	0.07	0.55	0.55	0.56
-2.0	0.21	0.24	0.19	0.16	0.15	0.12	0.48	0.47	0.48
-1.5	0.12	0.15	0.12	0.08	0.10	0.07	0.32	0.33	0.34
-1.0	-0.12	-0.09	-0.12	-0.14	-0.14	-0.15	0.04	0.04	0.05
-0.5	-0.01	0.00	-0.01	-0.02	-0.02	-0.03	0.05	0.05	0.05
0.0	-0.05	-0.05	-0.05	-0.05	-0.06	-0.06	-0.04	-0.05	-0.04
0.5	-0.03	-0.04	-0.03	-0.02	-0.03	-0.02	-0.09	-0.10	-0.10
1.0	-0.04	-0.05	-0.02	-0.01	-0.01	0.01	-0.15	-0.15	-0.15
1.5	-0.21	-0.21	-0.18	-0.16	-0.16	-0.13	-0.37	-0.36	-0.36
2.0	-0.27	-0.31	-0.26	-0.21	-0.24	-0.20	-0.49	-0.51	-0.52
2.5	-0.70	-0.71	-0.66	-0.62	-0.62	-0.58	-0.94	-0.91	-0.91

Table 13: Bias Results for Ability from 100-Examinee-45-Item Data Set

	Joint Bayesian-1			Joint Bayesian-2			Marginal Bayesian		
θ Level	α_L	α_T	$\alpha\beta_T$	α_L	α_T	$\alpha\beta_T$	α_L	α_T	$\alpha\beta_T$
-2.5	0.15	0.15	0.09	0.10	0.11	0.03	0.28	0.27	0.26
-2.0	0.08	0.09	0.03	0.05	0.06	0.00	0.20	0.18	0.18
-1.5	0.05	0.03	-0.01	0.00	0.01	-0.04	0.13	0.11	0.11
-1.0	0.00	-0.01	-0.04	-0.03	-0.02	-0.05	0.05	0.04	0.04
-0.5	-0.02	-0.02	-0.04	-0.03	-0.03	-0.05	0.01	0.00	0.00
0.0	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03
0.5	0.02	0.02	0.04	0.03	0.03	0.04	0.00	0.01	0.00
1.0	-0.01	0.00	0.03	0.03	0.02	0.05	-0.05	-0.03	-0.04
1.5	-0.10	-0.09	-0.05	-0.06	-0.07	-0.03	-0.16	-0.15	-0.15
2.0	-0.10	-0.11	-0.05	-0.07	-0.08	-0.02	-0.19	-0.19	-0.19
2.5	-0.27	-0.24	-0.18	-0.18	-0.18	-0.10	-0.35	-0.33	-0.32

5·)

Table 14: Bias Results for Ability from 300-Examinee-15-Item Data Set

	Joint Bayesian-1			Joint Bayesian-2			Marginal Bayesian		
θ Level	α_L	α_T	$\alpha\beta_T$	α_L	α_T	$\alpha\beta_T$	α_L	α_T	$\alpha\beta_T$
-2.5	0.65	0.57	0.55	0.57	0.53	0.51	0.86	0.85	0.86
-2.0	0.37	0.33	0.31	0.33	0.30	0.29	0.59	0.59	0.60
-1.5	0.10	0.09	0.07	0.08	0.06	0.05	0.31	0.31	0.32
-1.0	0.02	0.03	0.01	0.02	0.01	0.00	0.18	0.17	0.18
-0.5	-0.01	0.01	0.00	0.01	0.00	0.00	0.09	0.09	0.09
0.0	0.02	0.04	0.04	0.05	0.04	0.04	0.04	0.04	0.04
0.5	-0.02	-0.02	-0.01	0.00	-0.01	0.00	-0.10	-0.09	-0.09
1.0	-0.10	-0.08	-0.07	-0.07	-0.07	-0.06	-0.23	-0.22	-0.22
1.5	-0.15	-0.13	-0.11	-0.12	-0.10	-0.09	-0.34	-0.33	-0.33
2.0	-0.28	-0.25	-0.23	-0.24	-0.22	-0.20	-0.53	-0.52	-0.52
2.5	-0.51	-0.47	-0.44	-0.47	-0.43	-0.41	-0.80	-0.79	-0.80

Table 15: Bias Results for Ability from 300-Examinee-45-Item Data Set

	Joint Bayesian-1			Joint Bayesian-2			Marginal Bayesian		
θ Level	α_L	α_T	$\alpha\beta_T$	α_L	α_T	$\alpha\beta_T$	α_L	α_T	$\alpha\beta_T$
-2.5	0.10	0.08	0.05	0.06	0.05	0.02	0.26	0.23	0.23
-2.0	0.11	0.11	0.08	0.09	0.08	0.06	0.23	0.22	0.22
-1.5	0.08	0.07	0.05	0.06	0.05	0.03	0.17	0.15	0.15
-1.0	0.06	0.05	0.03	0.04	0.03	0.02	0.11	0.10	0.10
-0.5	0.01	0.01	0.00	0.01	0.01	0.00	0.04	0.04	0.04
0.0	0.02	0.03	0.03	0.03	0.03	0.03	0.02	0.03	0.02
0.5	0.00	0.00	0.01	0.01	0.01	0.02	-0.03	-0.02	-0.03
1.0	-0.01	0.00	0.02	0.01	0.01	0.03	-0.07	-0.05	-0.06
1.5	0.03	0.04	0.06	0.06	0.06	0.08	-0.07	-0.04	-0.05
2.0	-0.18	-0.17	-0.14	-0.15	-0.15	-0.13	-0.30	-0.28	-0.28
2.5	-0.29	-0.29	-0.26	-0.26	-0.26	-0.23	-0.44	-0.42	-0.42

5?