Journal of Multivariate Analysis 114 (2013) 359–377

Contents lists available at SciVerse ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

# Asymptotic properties of the Bayes and pseudo Bayes estimators of ability in item response theory

Haruhiko Ogasawara

*Department of Information and Management Science, Otaru University of Commerce, 3-5-21, Midori, Otaru 047-8501, Japan*

ARTICLE INFO

*Article history:*
Received 23 August 2011
Available online 27 August 2012

*AMS subject classifications:*
primary 62F12
secondary 62F15

*Keywords:*
IRT
Asymptotic cumulants
Bayes modal
Mean square error
Jeffreys prior
Bias
Higher-order asymptotic variance
Model misspecification

ABSTRACT

Asymptotic cumulants of the Bayes and pseudo Bayes estimators of ability in item response theory are obtained up to the fourth order with the higher-order asymptotic variance under possible model misspecification. Typical estimators are treated as special cases of the (pseudo) Bayes estimator with the general weight. The asymptotic cumulants of the estimators after studentization are also derived. From the comparison of the mean square errors, the Bayes modal estimator with the standard normal prior is recommended for point estimation. For interval estimation, however, the maximum likelihood estimator is appropriate considering its small bias after studentization.



## 1. Introduction

For an ability test developed using item response theory (IRT), estimation of the proficiencies of examinees separately from the calibration for items is a main purpose of the test. When the parameters of items have been well estimated in an item bank using a large sample size, the item parameters may be seen as population ones though they are actually estimates under possible model misspecification (p.m.m.). With the assumption of known item parameters, the ability of an examinee, seen as an unknown fixed parameter, is estimated by various ways. Maximum likelihood (ML) estimation is a standard method [24], [4, Section 20.3], where it is well known that the asymptotic variance of the ML estimator (MLE) is given by the reciprocal of the test information or the Fisher information. Lord [25] gave the asymptotic biases of the MLE of an ability and its monotone transformations under correct model specification (c.m.s.). Under p.m.m., Ogasawara [29] derived the asymptotic cumulants up to the fourth order and the higher-order asymptotic variances of the MLEs of an ability and its transformations with and without studentization.

In addition to ML estimation, Bayesian estimation has been used, where the maximum a posteriori (MAP) estimator [34, Chapter 2], [6] and the expected a posteriori (EAP) estimator [2,3,6] are typical ones. The MAP estimator is also called the Bayes modal estimator (BME). Lord [26] gave the asymptotic bias of the BME with the standard normal prior under c.m.s. One of the advantages of the BME is that finite values of the estimators are available in the cases of no correct/incorrect responses while the corresponding MLEs are infinite.

The weighted likelihood estimator (WLE) given by Warm [36] looks like a BME, where the weight corresponding to the first derivative of the logarithm of a prior with respect to the ability parameter is introduced such that the asymptotic bias

*E-mail address:* hogasa@res.otaru-uc.ac.jp.

0047-259X/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2012.08.013

disappears. Since it is difficult to derive the prior explicitly, the WLE may be seen as a pseudo BME. The properties of the WLE and its applications were investigated using simulations by Kim and Nicewander [20] and Hoijtink and Boomsma [17]. Penfield and Bergeron [33] used the WLE in the generalized partial credit model. Another BME is given by using the Jeffreys [18], [19, Section 3.10] uninformative prior written as JME in this paper. It is known that the WLE becomes equal to the JME in the case of the 2-parameter logistic model (2PLM).

Further, in tailored testing a sequential method of ability estimation is available [31,32,35]. However, the method is not dealt with in this paper. That is, it is supposed that estimation of ability is performed based on a fixed number of items.

The MLE, BME, WLE and JME can be seen as special cases of the estimator using the general weight, given as a function of an ability, where the weight corresponds, in form, to the log-prior derivative with respect to the ability parameter. In the case of the MLE, the weight is defined as null. The main purpose of this paper is to derive the asymptotic properties such as the asymptotic cumulants up to the fourth order with the higher-order asymptotic variance and the mean square error for the pseudo Bayes estimator using the general weight. We also show their applications to the four typical estimators mentioned above. The asymptotic properties will be given for the estimators with and without studentization. It is known that the so-called sandwich estimator of the asymptotic variance is robust against p.m.m. in the case of independent and identically distributed (i.i.d.) cases (e.g., [28]). Unfortunately, for estimation of ability, the items in an ability test are generally unidentical and do not give the advantage of the sandwich estimator (note that generally only one observation per examinee is available for an item). So, studentization is performed using the estimated information with the assumption of c.m.s. However, the asymptotic properties of the estimators will be given under p.m.m. as well as under c.m.s.

The results given in this paper will be useful when we choose good point estimators among the typical ones used in practice by e.g., seeing the amounts of their mean square errors, which will be algebraically given and numerically illustrated. Estimators are also used for interval estimation of the population value of an ability, where the asymptotic distributions of the estimators after studentization are required. For interval estimation more accurate than the usual Wald confidence interval (CI), the asymptotic cumulants in addition to the usual asymptotic variance are needed and will be given with numerical illustration.

This paper is organized as follows. In Section 2, the definition of the (pseudo) Bayes estimator using the general weight is given with their special cases. The population value of ability under model misspecification (m.m.) is also defined. Section 3 gives the asymptotic cumulants of the ability estimator before studentization while in Section 4 the corresponding results for the studentized estimator are given with their advantages summarized as theorems. In Section 5, numerical examples are shown using artificial and real data, where the results implied by the theorems are numerically illustrated. Finally, Section 6 gives some conclusive comparisons for typical ability estimators.

## 2. The pseudo Bayes estimator with the general weight

Let $U_k$ be the dichotomous variable for the $k$th item ($k = 1, \ldots, n$), where $U_k = 1$ denotes a correct response, $U_k = 0$ an incorrect one by an examinee with ability $\theta$, and $n$ is the number of items. The probability of $U_k = 1$ for the examinee by an IRT model is generally written as $P_k = P_k(\theta) = \Pr(U_k = 1|\theta)$ with $Q_k \equiv 1 - P_k$ ($k = 1, \ldots, n$), which is usually a monotone function of $\theta$ and is assumed to be differentiable a required number of times. For instance, the familiar 3-parameter logistic model (3PLM) is

$$P_k = \Pr(U_k = 1|\theta, a_k, b_k, c_k) = c_k + \frac{1 - c_k}{1 + \exp\{-Da_k(\theta - b_k)\}} \quad (k = 1, \ldots, n), \tag{2.1}$$

where $D = 1.7$, and the item parameters $a_k$, $b_k$ and $c_k$ ($k = 1, \ldots, n$) are assumed to be known. We also suppose that the model may be misspecified. The corresponding true probability of $U_k = 1$ is denoted by

$$P_{Tk} \equiv E_T(U_k|\theta) \quad \text{with } Q_{Tk} \equiv 1 - P_{Tk} \ (k = 1, \ldots, n), \tag{2.2}$$

where $E_T(\cdot)$ indicates that the expectation is taken using the true probability function for $U_k$. That is $P_k \neq P_{Tk}$ for at least one item under m.m.

Let $\mathbf{u} = (U_1, \ldots, U_n)'$; then with the assumption of local independence, the likelihood of $\theta$ is written as $L = L(\theta|\mathbf{u}) = \prod_{k=1}^{n} P_k^{U_k} Q_k^{1-U_k}$. Let $p(\theta)$ be the prior density for $\theta$. Then, the estimator $\hat{\theta}_{\mathrm{GW}}$ of ability using the general weight $g(\theta) \equiv \partial \log p(\theta)/\partial\theta$ is defined as the $\theta$ that maximizes

$$\begin{aligned} \bar{l}_{\mathrm{GW}} &\equiv n^{-1}\{\log L + \log p(\theta)\} \equiv \bar{l} + n^{-1}\log p(\theta) \\ &= n^{-1}\sum_{k=1}^{n}\{U_k \log P_k + (1 - U_k)\log Q_k\} + n^{-1}\log p(\theta), \end{aligned} \tag{2.3}$$

where $\bar{l}$ is the mean log-likelihood. Note that even if $p(\theta)$ is not available in an explicit form, $g(\theta)$ can still be defined. In this case $\hat{\theta}_{\mathrm{GW}}$ is defined using the modified score in pseudo Bayesian estimation. The four special cases are summarized as follows:

$$\begin{aligned} &\hat{\theta}_{\mathrm{GW}} = \hat{\theta}_{\mathrm{ML}} \quad \text{when } g(\theta) = 0, \qquad \hat{\theta}_{\mathrm{GW}} = \hat{\theta}_{\mathrm{BM}} \quad \text{when } g(\theta) = -\theta, \\ &\hat{\theta}_{\mathrm{GW}} = \hat{\theta}_{\mathrm{WL}} \quad \text{when } g(\theta) = \bar{j}/(2\bar{i}), \quad \text{and} \quad \hat{\theta}_{\mathrm{GW}} = \hat{\theta}_{\mathrm{JM}} \quad \text{when } g(\theta) = \bar{i}'/(2\bar{i}), \end{aligned} \tag{2.4}$$