PSYCHOMETRIKA—VOL. 80, NO. 1, 200–204
MARCH 2015
DOI: 10.1007/S11336-013-9378-5

A NOTE ON WEIGHTED LIKELIHOOD AND JEFFREYS MODAL ESTIMATION OF PROFICIENCY LEVELS IN POLYTOMOUS ITEM RESPONSE MODELS

DAVID MAGIS

UNIVERSITY OF LIÈGE AND KU LEUVEN, BELGIUM

Warm (in Psychometrika, 54, 427–450, 1989) established the equivalence between the so-called Jeffreys modal and the weighted likelihood estimators of proficiency level with some dichotomous item response models. The purpose of this note is to extend this result to polytomous item response models. First, a general condition is derived to ensure the perfect equivalence between these two estimators. Second, it is shown that this condition is fulfilled by two broad classes of polytomous models including, among others, the partial credit, rating scale, graded response, and nominal response models.

Key words: item response theory, polytomous models, weighted likelihood, Jeffreys modal.

1. Introduction

The purpose of this short note is to generalize the equivalence of two proficiency level estimators, namely the weighted likelihood (WL) estimator (Warm, 1989) and the so-called Jeffreys modal (JM) estimator (Magis & Raîche, 2012), to the context of polytomous item response models. The equivalence of these two estimators was established under some dichotomous item response theory (IRT) models (Warm, 1989). The goal is to derive a necessary and sufficient condition on the structure of the polytomous model for this equivalence between estimators to hold. In a second step, this condition is investigated for two broad classes of polytomous models, namely the "difference" models and the "divide-by-total" models (Thissen & Steinberg, 1986).

2. Weighted Likelihood and Jeffreys Modal Estimation

Let n be the number of polytomous items administered to an examinee with proficiency level θ. Each item j ($j = 1, \ldots, n$) has $g_j + 1$ response categories, coded as $0, 1, \ldots, g_j$, and each item has a set of specific parameters, denoted by $\boldsymbol{p}_j$. Let $P_{jk}(\theta) = \Pr(X_j = k|\theta, \boldsymbol{p}_j)$ be the probability that the response X_j of the examinee to item j is the k-th one ($k = 0, 1, \ldots, g_j$), given examinee's proficiency θ and item parameters $\boldsymbol{p}_j$. The items are assumed to be pre-calibrated so that item parameters $\boldsymbol{p}_j$ can be considered as fixed values. Moreover, the response probability function $P_{jk}(\theta)$ is assumed to be several times differentiable with respect to θ.

To define the WL and JM estimators, set τ_{jk} as the random variable being equal to one if $X_j = k$ (that is, if the answer to item j is response category k) and zero otherwise. Then the likelihood $L(\theta)$ and log-likelihood $l(\theta)$ functions take the form

$$L(\theta) = \prod_{j=1}^{n} \prod_{k=0}^{g_j} P_{jk}(\theta)^{\tau_{jk}} \tag{1}$$

Requests for reprints should be sent to David Magis, Department of Education (B32), University of Liège, Boulevard du Rectorat 5, 4000 Liège, Belgium. E-mail: david.magis@ulg.ac.be

© 2013 The Psychometric Society

and

$$l(\theta) = \sum_{j=1}^{n} \sum_{k=0}^{g_j} \tau_{jk} \log P_{jk}(\theta), \tag{2}$$

respectively. The JM estimator is a particular Bayesian maximum a posteriori (MAP) estimator of proficiency, obtained by selecting a non-informative prior, proportional to the squared root of the test information function (Jeffreys, 1946): $f(\theta) \propto \sqrt{I(\theta)}$ with

$$I(\theta) = -E\left(\frac{d^2 l(\theta)}{d\theta^2}\right) = -\sum_{j=1}^{n} \sum_{k=0}^{g_j} P_{jk}(\theta) \frac{d^2 \log P_{jk}(\theta)}{d\theta^2}. \tag{3}$$

Making use of the Jeffreys' prior distribution yields the so-called Jeffreys modal (JM) estimator, obtained by solving

$$\frac{I'(\theta)}{2I(\theta)} + \frac{dl(\theta)}{d\theta} = 0, \tag{4}$$

with $I'(\theta)$ standing for the first derivative of $I(\theta)$ with respect to θ.

Furthermore, Samejima (1998) developed the polytomous version of the weighted likelihood (WL) estimator (Warm, 1989). This estimator is obtained by solving

$$\frac{J(\theta)}{2I(\theta)} + \frac{dl(\theta)}{d\theta} = 0, \tag{5}$$

where

$$J(\theta) = \sum_{j=1}^{n} \sum_{k=0}^{g_j} \frac{P'_{jk}(\theta) P''_{jk}(\theta)}{P_{jk}(\theta)}, \tag{6}$$

and $P'_{jk}(\theta)$ and $P''_{jk}(\theta)$ stand respectively for the first and second derivatives of $P_{jk}(\theta)$ with respect to θ. The closeness of Equations (4) and (5) is the starting point of further comparisons between the WL and the JM estimators of ability.

To compare the WL and the JM estimators with polytomous IRT models, set

$$f_{\mathrm{WL}}(\theta) = \frac{J(\theta)}{2I(\theta)} + \frac{dl(\theta)}{d\theta} \quad \text{and} \quad f_{\mathrm{JM}}(\theta) = \frac{I'(\theta)}{2I(\theta)} + \frac{dl(\theta)}{d\theta}, \tag{7}$$

and let us establish the conditions on the response probabilities $P_{jk}(\theta)$ that ensure $f_{\mathrm{WL}}(\theta) = f_{\mathrm{JM}}(\theta)$ for any θ. This equality is fulfilled whenever the difference $I'(\theta) - J(\theta)$ is consistently equal to zero. To compute this difference, note first

$$\frac{d^2 \log P_{jk}(\theta)}{d\theta^2} = \frac{d}{d\theta}\left(\frac{P'_{jk}(\theta)}{P_{jk}(\theta)}\right) = \frac{P''_{jk}(\theta) P_{jk}(\theta) - P'_{jk}(\theta)^2}{P_{jk}(\theta)^2}, \tag{8}$$

so that

$$I(\theta) = \sum_{j=1}^{n} \sum_{k=0}^{g_j} \left(\frac{P'_{jk}(\theta)^2}{P_{jk}(\theta)} - P''_{jk}(\theta)\right). \tag{9}$$

Moreover, since $\sum_{k=0}^{g_j} P_{jk}(\theta) = 1$ for any θ, the function $h(\theta) = \sum_{k=0}^{g_j} P_{jk}(\theta)$ is constant with respect to θ, and each derivative of $h(\theta)$ with respect to θ is consistently equal to zero. Hence,

$$I(\theta) = \sum_{j=1}^{n} \sum_{k=0}^{g_j} \frac{P'_{jk}(\theta)^2}{P_{jk}(\theta)}, \tag{10}$$

and

$$I'(\theta)=\sum_{j=1}^{n}\sum_{k=0}^{g_j}P'_{jk}(\theta)\left(\frac{2P''_{jk}(\theta)}{P_{jk}(\theta)}-\frac{P'_{jk}(\theta)^2}{P_{jk}(\theta)^2}\right). \tag{11}$$

Eventually,

$$I'(\theta)-J(\theta)=\sum_{j=1}^{n}\sum_{k=0}^{g_j}P'_{jk}(\theta)\left(\frac{P''_{jk}(\theta)}{P_{jk}(\theta)}-\frac{P'_{jk}(\theta)^2}{P_{jk}(\theta)^2}\right)=\sum_{j=1}^{n}\sum_{k=0}^{g_j}P'_{jk}(\theta)\frac{d}{d\theta}\left(\frac{P'_{jk}(\theta)}{P_{jk}(\theta)}\right), \tag{12}$$

according to (8). In sum, the WL and JM estimators are completely identical if and only if

$$\sum_{j=1}^{n}\sum_{k=0}^{g_j}P'_{jk}(\theta)\frac{d}{d\theta}\left(\frac{P'_{jk}(\theta)}{P_{jk}(\theta)}\right)=0, \tag{13}$$

for any θ. Note that a sufficient (but not necessary) condition to ensure (14) is to verify that

$$\sum_{k=0}^{g_j}P'_{jk}(\theta)\frac{d}{d\theta}\left(\frac{P'_{jk}(\theta)}{P_{jk}(\theta)}\right)=0, \tag{14}$$

for any θ and any item. The purpose of the next sections is to demonstrate that relationship (13) holds for two classes of polytomous item response models, the "difference" models and the "divide-by-total" models (Thissen & Steinberg, 1986).

3. Difference Models

Difference models can be used to model ordered category polytomous items. Such models are defined by a cumulative probability function, say $P^*_{jk}(\theta)=\Pr(X_j\geq k)$, describing the probability that the item response is one of the categories $k,k+1,\ldots,g_j$. The response category probabilities $P_{jk}(\theta)$ are then obtained as the successive differences $P_{jk}(\theta)=P^*_{jk}(\theta)-P^*_{j,k+1}(\theta)$ with the convention that $P^*_{j0}(\theta)=1$ and $P^*_{jk}(\theta)=0$ for any $k>g_j$.

The class of difference models to be considered in this paper has the following structural form:

$$P^*_{jk}(\theta)=\frac{e_{jk}(\theta)}{1+e_{jk}(\theta)}, \tag{15}$$

with the constraint that

$$e'_{jk}(\theta)=\frac{de_{jk}(\theta)}{d\theta}=\alpha_j e_{jk}(\theta) \tag{16}$$

and α_j is constant with respect to θ. This class of models includes the graded response model (GRM; Samejima, 1969) and the modified graded response model (M-GRM; Muraki, 1990), which can easily be assessed by, e.g., using the notations of Embretson and Reise (2000).

First, for any $j=1,\ldots,n$ and any $k=0,\ldots,g_j$,

$$P^{*\prime}_{jk}(\theta)=\frac{dP^*_{jk}(\theta)}{d\theta}=\frac{\alpha_j e_{jk}(\theta)}{[1+e_{jk}(\theta)]^2}=\alpha_j P^*_{jk}(\theta)\big[1-P^*_{jk}(\theta)\big], \tag{17}$$

so that

$$P'_{jk}(\theta)=P^{*\prime}_{jk}(\theta)-P^{*\prime}_{j,k+1}(\theta)=\alpha_j\big[P^*_{jk}(\theta)-P^*_{j,k+1}(\theta)\big]\big[1-P^*_{jk}(\theta)-P^*_{j,k+1}(\theta)\big], \tag{18}$$

and by definition of $P_{jk}(\theta)$,

$$\frac{P'_{jk}(\theta)}{P_{jk}(\theta)} = \alpha_j \left[1 - P^*_{jk}(\theta) - P^*_{j,k+1}(\theta)\right]. \tag{19}$$

It follows that

$$\frac{d}{d\theta}\left(\frac{P'_{jk}(\theta)}{P_{jk}(\theta)}\right) = -\alpha_j \left[P^{*\prime}_{jk}(\theta) + P^{*\prime}_{j,k+1}(\theta)\right], \tag{20}$$

and thus

$$P'_{jk}(\theta)\frac{d}{d\theta}\left(\frac{P'_{jk}(\theta)}{P_{jk}(\theta)}\right) = -\alpha_j \left[P^{*\prime}_{jk}(\theta)^2 - P^{*\prime}_{j,k+1}(\theta)^2\right] \tag{21}$$

using (20) and the left equality in (18). This yields eventually

$$\sum_{k=0}^{g_j} P'_{jk}(\theta)\frac{d}{d\theta}\left(\frac{P'_{jk}(\theta)}{P_{jk}(\theta)}\right) = -\alpha_j \left[P^{*\prime}_{j0}(\theta)^2 - P^{*\prime}_{j,g_j+1}(\theta)^2\right] = 0, \tag{22}$$

since $P^{*\prime}_{j0}(\theta) = P^{*\prime}_{j,g_j+1}(\theta) = 0$ because $P^*_{j0}(\theta)$ and $P^*_{j,g_j+1}(\theta)$ are constant with respect to θ. In sum, any model defined by (15) and (16) satisfies the relationship (13), and hence establishes the equivalence between the JM and WL estimators.

4. Divide-by-Total Models

The class of "divide-by-total" models is defined by

$$P_{jk}(\theta) = \frac{\Gamma_{jk}(\theta)}{\sum_{r=0}^{g_j}\Gamma_{jr}(\theta)} = \frac{\Gamma_{jk}(\theta)}{\Gamma_j(\theta)}, \tag{23}$$

with the constraint that

$$\Gamma'_{jk}(\theta) = \frac{d\Gamma_{jk}(\theta)}{d\theta} = h_{jk}\Gamma_{jk}(\theta) \tag{24}$$

and h_{jk} does not depend on θ. This class includes four well-known polytomous item response models: the partial credit model (PCM; Masters, 1982), the generalized partial credit model (GPCM; Muraki, 1992), the rating scale model (RSM; Andrich, 1978), and the nominal response model (NRM; Bock, 1972). Using the notations of Embretson and Reise (2000), it is straightforward to determine, for each model, the values of $\Gamma_{jk}(\theta)$ and h_{jk} functions that ensure the definitions (23) and (24).

Now, one has first

$$\Gamma'_j(\theta) = \frac{d\Gamma_j(\theta)}{d\theta} = \sum_{r=0}^{g_j} h_{jr}\Gamma_{jr}(\theta) \quad \text{and} \quad \frac{\Gamma'_j(\theta)}{\Gamma_j(\theta)} = \sum_{r=0}^{g_j} h_{jr}P_{jr}(\theta), \tag{25}$$

by definition of $P_{jr}(\theta)$. It follows that

$$P'_{jk}(\theta) = \frac{\Gamma'_{jk}(\theta)}{\Gamma_j(\theta)} - \frac{\Gamma_{jk}(\theta)\Gamma'_j(\theta)}{\Gamma_j(\theta)^2} = P_{jk}(\theta)\left[h_{jk} - \sum_{r=0}^{g_j} h_{jr}P_{jr}(\theta)\right], \tag{26}$$

and hence

$$\frac{d}{d\theta}\left(\frac{P'_{jk}(\theta)}{P_{jk}(\theta)}\right) = \frac{d}{d\theta}\left(h_{jk} - \sum_{r=0}^{g_j} h_{jr}P_{jr}(\theta)\right) = -\sum_{r=0}^{g_j} h_{jr}P'_{jr}(\theta), \tag{27}$$

which does not depend on the particular category index k. In sum,

$$\sum_{k=0}^{g_j} P'_{jk}(\theta)\frac{d}{d\theta}\left(\frac{P'_{jk}(\theta)}{P_{jk}(\theta)}\right) = -\sum_{r=0}^{g_j} h_{jr}\,P'_{jr}(\theta)\sum_{k=0}^{g_j} P'_{jk}(\theta) = 0, \tag{28}$$

by definition of the aforementioned function $h(\theta)$. In sum, the equivalence between the JM and WL estimators also holds for the "divide-by-total" models defined by relationships (23) and (24).

5. Concluding Remarks

This note studied the relationships between the WL and JM estimators of proficiency under polytomous IRT models. Their equivalence was established for two broad classes of polytomous models that encompass most-known polytomous IRT models. Furthermore, fixing $g_j = 1$ for all items leads back to the results and equations displayed in Magis and Raîche (2012) and Warm (1989) in the dichotomous framework.

Acknowledgements

The author wishes to thank Roger E. Millsap, Editor, and three anonymous reviewers for their helpful comments. This research was funded by a post-doctoral research grant "Chargé de recherches" of the National Funds for Scientific Research (FNRS, Belgium), the IAP Research Network P7/06 of the Belgian State (Belgian Science Policy), and the Research Funds of the KU Leuven, Belgium.

References

Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika*, *43*, 561–573. doi:10.1007/BF02293814.

Bock, R.D. (1972). Estimating item parameters and latent ability when responses are scored in two or more nominal categories. *Psychometrika*, *37*, 29–51. doi:10.1007/BF02291411.

Embretson, S.E., & Reise, S.P. (2000). *Item response theory for psychologists*. Mahwah: Lawrence Erlbaum Associates.

Jeffreys, H. (1946). An invariant form for the prior probability in estimation problems. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, *186*, 453–461.

Magis, D., & Raîche, G. (2012). On the relationships between Jeffreys modal and weighted likelihood estimation of ability under logistic IRT models. *Psychometrika*, *77*, 163–169. doi:10.1007/s11336-011-9233-5.

Masters, G.N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, *47*, 149–174. doi:10.1007/BF02296272.

Muraki, E. (1990). Fitting a polytomous item response model to Likert-type data. *Applied Psychological Measurement*, *14*, 59–71. doi:10.1177/014662169001400106.

Muraki, E. (1992). A generalized partial credit model: application of an EM algorithm. *Applied Psychological Measurement*, *16*, 159–176. doi:10.1177/014662169201600206.

Samejima, F. (1969). *Estimation of latent ability using a response pattern of graded scores*. *Psychometrika monograph* (Vol. 17).

Samejima, F. (1998, April). *Expansion of Warm's weighted likelihood estimator of ability for the three-parameter logistic model to general discrete responses*. Paper presented at the annual meeting of the National Council on Measurement in Education, San Diego, CA.

Thissen, D., & Steinberg, L. (1986). A taxonomy of item response models. *Psychometrika*, *51*, 567–577. doi:10.1007/BF02295596.

Warm, T.A. (1989). Weighted likelihood estimation of ability in item response models. *Psychometrika*, *54*, 427–450. doi:10.1007/BF02294627.

Manuscript Received: 10 DEC 2012
Published Online Date: 27 NOV 2013