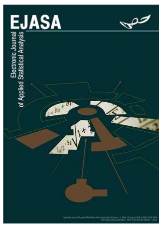

Electronic Journal of Applied Statistical Analysis
EJASA (2012), Electron. J. App. Stat. Anal., Vol. 5, Issue 3, 360 – 366
e-ISSN 2070-5948, DOI 10.1285/i20705948v5n3p360

COMPARISON OF TWO BIAS REDUCTION TECHNIQUES FOR THE RASCH MODEL

Lucio Bertoli Barsotti[*(1)], **Antonio Punzo**[(2)]

[(1)] *Dipartimento di Scienze Aziendali, Economiche e Metodi Quantitativi, University of Bergamo, Italy*
[(2)] *Dipartimento di Impresa, Culture e Società,Università di Catania, Italy*

Received 20 July 2012; Accepted 07 October 2012
Available online 16 November 2012

Abstract: *This study examines the effect of two different techniques of bias reduction in the case of the fixed persons-fixed items formulation of the Rasch model. A first approach can be considered "corrective", because it consists simply in correcting ex-post the joint maximum likelihood estimates by a factor (m-1)/m, were m represents the number of items and/or persons. A second approach, which is an application of a quite general formula for reducing the maximum likelihood estimation bias, can be considered "preventive", because it arises from a modification of the score function. A comparative study of these two techniques was done using simulated data.*

Keywords: *Maximum likelihood estimation, Rasch model, bias, modified score.*

1. Introduction

For the Rasch Model (RM), the Joint Maximum Likelihood (JML) is an estimation procedure in which item and person parameters are estimated simultaneously. One of the major drawbacks of the JML approach is that item parameters cannot be estimated consistently if the number of subjects, n, approaches infinity and the number of items, k, is fixed. More specifically, it is known that the JML estimation of item parameters is biased (both for the case of finite samples, and asymptotically). Indeed the JML estimate of the item difficulty parameters have an approximate bias that is a function of the constant $k/(k-1)$ ([1], p. 244). As a practical solution for reducing this bias, [1] proposed the use of a multiplicative bias correcting factor $(k-1)/k$.
The main purpose of this study is to compare the properties of two possible bias-reducing procedures for the JML estimation of the RM parameters: i) the $(m-1)/m$ bias correction (where m can be either n or k); ii) the procedure of bias reduction suggested by [6] and based on a

[*] E-mail: lucio.bertoli-barsotti@unibg.it

suitable modification of the score function (modified score). The two methods will be simply denoted as C-JML and JML*, respectively, hereafter. While C-JML is "corrective", because the Maximum Likelihood Estimate (MLE) is first calculated and then corrected, JML* is "preventive" in character, because the likelihood function is modified before the MLE is computed. The C-JML approach is actually adopted by the software Winsteps ([8]), while the Firth's approach does not seems to be implemented in any programs for the JML estimation of the RM. It is important to emphasize these approaches to bias reduction apply both to item and person parameters. Note that, since the correction factor is smaller than one, the C-JML estimator reduces the standard error of the JML estimator. Now, it is known that the JML approach overspreads the estimates of item difficulties under the Rasch model, and that the C-JML estimator yields practically unbiased estimates ([11]). Then, one can also deduce that the C-JML always has a higher precision, in terms of mean square error, with respect to the JML estimator. Nevertheless, this estimator suffers of the unbounded nature of the JML estimation: if the JML estimate is infinite, so is the C-JML estimate. In particular, the $(m-1)/m$ bias correction applies only for non-extreme score vectors. Moreover, this approach does not apply to other cases in which the JML estimate does not exist (see [7] for necessary and sufficient conditions for the existence of a JML estimate for the RM).

Firth's approach is defined in a rather general framework, that is for both exponential and nonexponential models. In particular, for the case of an exponential family in its canonical parameterization (that is the case of the RM) the method consists in maximizing a modified log-likelihood function $l^* = l + A$, where l is the log-likelihood function and A is one-half of the logarithm of the square root of the determinant of the Fisher information matrix. Still other bias reducing techniques are possible; for example, one could consider other types of adjustements of the score function (for example by taking observed instead of expected information function). Alternatively, one could rely on a suitable modification of a profile likelihood function; the interested reader is referred to [3] for an up-to-date review concerning techniques of bias reduction based on modified likelihood, or modified profile likelihood. However, an investigation along these lines goes beyond the scope of the present study. Interestingly, [12] introduced a special case of Firth's bias reduction method, by applying a similar formula to the 3-parameter logistic model (which includes the RM as a special case), but his approach is only devoted to the problem of the estimation of the ability parameter, under the assumption that the item parameters are *known*. This estimate is defined as the Weighted Likelihood Estimate (WLE) in [12]. The WLE is currently adopted, as a default, by the software RUMM 2020 ([2]) to obtain person parameter estimates. In a first step, this software uses the (Pairwise) Conditional Maximum Likelihood (CML) estimation method for obtaining the item parameters. Then, in a second step, the MLE (or WLE) approach is used to estimate ability parameters, treating the previously estimated item parameters as if they were the true quantities.

This paper focuses on the JML estimation approach for the RM; special attention is devoted to the problem of the bias of item parameters (for the simple reason that for person parameters this problem is less likely to occur). More specifically, we will explore comparatively the effect of the MLE bias reducing formula proposed by Firth in the special case of the JML estimation of the RM, with respect to that given by the C-JML estimator.

2. Materials and Methods

According to the RM for dichotomous responses 0/1, the logit of the probability of a 1-response is $\ln\left(P\left(X_{vi}=1\right)/P\left(X_{vi}=0\right)\right)=\theta_v-\beta_i$, where $X_{vi}=1$ denotes a 1-response of person *v* ($v=1,...,n$) to item *i* ($i=1,...,k$) and where the parameters θ_v and β_i represent, respectively, the ability of person *v* and the difficulty of item *i*. The JML is not simply an estimation method for the RM but it should be considered a model formulation, also known as fixed persons-fixed items RM ([5]). In a fixed persons-fixed items RM, item and person parameters are estimated simultaneously by maximizing the log-likelihood function $l=\sum_{vi}\ln p_{vi}$, where $p_{vi}=P(X_{vi}=x_{vi})$, with x_{vi} taking values on the set $\{0,1\}$. In particular, let $P_{vi}=P\left(X_{vi}=1\right)$, $Q_{vi}=1-P_{vi}$ and $U_{vi}=P_{vi}Q_{vi}$. It is easy to see that $\partial l/\partial\theta_v=\sum_i\left\{x_{vi}-exp\left(\theta_v-\beta_i\right)/\left[1+exp\left(\theta_v-\beta_i\right)\right]\right\}$, and that $\partial l/\partial\beta_i=\sum_v\left\{-x_{vi}+exp\left(\theta_v-\beta_i\right)/\left[1+exp\left(\theta_v-\beta_i\right)\right]\right\}$. Then $\partial^2 l/\partial\theta_v^2=-\sum_i U_{vi}$; $\partial^2 l/\partial\theta_v\partial\theta_w=0$, for every $v\neq w$; $\partial^2 l/\partial\beta_i^2=-\sum_v U_{vi}$; $\partial^2 l/\partial\beta_i\partial\beta_j=0$, for every $i\neq j$; $\partial^2 l/\partial\theta_v\partial\beta_i=U_{vi}$.

Now, since only $k+1$ different test scores are possible for persons – that is $x_{v\bullet}=0,1,...,k$ - only $k+1$ different theta estimates are possible. Then, by considering one identifiability constraint, the dimension of the canonical parameter is $(k+1)+k-1=2k$ (at most, depending on the dataset). Let f_t be the number of persons having test score *t*, $t=0,1,...,k$. There is no loss of generality by considering $\left(\eta_1,\eta_2,...,\eta_{2k}\right)$ as the canonical parameter (simply by renaming person and item parameters), where $\eta_t=\lambda_t$, $t=1,...,k$, and $\eta_{k+i}=\beta_i$, $i=1,...,k$, where λ_t represents the parameter of the persons with test score *t*, and where θ_0 (for example) is taken equal to zero as an identifiability constraint. The elements of the information matrix $\Im$ are defined by the relationships $-E\,\partial^2 l/\partial\eta_t^2=\sum_i f_t U_{ti}$, $-E\,\partial^2 l/\partial\eta_t\partial\eta_{k+i}=-f_t U_{ti}$, $-E\,\partial^2 l/\partial\eta_{k+i}^2=\sum_t f_t U_{ti}$, and zero otherwise (for simplicity we assume that there are no ties among the item total scores $x_{\bullet 1},...,x_{\bullet k}$; otherwise a similar grouping of items may also be considered). The JML* estimator is obtained by maximizing the modified log-likelihood function $l^*=l+A$, where $A=2^{-1}\ln\left|\Im\right|$.

3. Simulation study

A simulation study was carried out to determine the bias and the precision of the three estimation methods: JML, JML* and C-JML. The entire simulation was performed in the R computing environment ([9]). Maximization of the likelihood is carried out by using the `nlm` function, which adopts a Newton-type algorithm, in the R-package `stats`. The R functions necessary to obtain the estimates, according to the considered methods, are available from the website `http://www.economia.unict.it/punzo`.

3.1 Design

We considered 6 test lengths k ranging from 5 to 30, in increments of 5, and 3 sample sizes, n =100, 400, 1000. For each couple (n,k), R=100 replications were taken into account, leading to a total of 1800 datasets. In detail, for each couple (n,k) and for each replication:

- n values of θ were generated from a standard normal distribution;
- k values of β were drawn from a standard normal distribution and *a posteriori* centered to have zero mean;
- a $n\times k$ dataset was generated from a RM;
- JML, JML* and C-JML estimates were derived for the dataset at hand, with an identifiability constraint of zero mean on the β parameters. This identifiability constraint was chosen in order to “tune” true and estimated item parameters.

3.2 Error indices

For each couple (n,k), the overall precision of the obtained item parameter estimates was assessed by (an estimated) root mean square error $\mathrm{RMSE}=\sqrt{R^{-1}\sum_{r=1}^{R}k^{-1}\sum_{j=1}^{k}\left(\hat{\beta}_{jr}-\beta_{jr}\right)^2}$, where $\hat{\beta}_{jr}$ is the estimated parameter of item j in replication r, and β_{jr} is the corresponding true parameter value. In order to evaluate the bias note that, due to the *a posteriori* centering of the β_{jr} and due to the way the $\hat{\beta}_{jr}$ are constrained, the average bias $(Rk)^{-1}\sum_{r,j}\left(\hat{\beta}_{jr}-\beta_{jr}\right)$ masks the overall magnitude of the bias, because in this formula negative and positive biases cancel each other out. The following formula (see [10]) was adopted to evaluate the bias $\Delta_I=\left(S_I\right)^{-1}\sum_{(j,r)\in\Omega_I}\hat{\beta}_{jr}/\beta_{jr}$, where $\Omega_I=\left\{(j,r),\, j=1,...,k,\, r=1,...,R: \left|\beta_{jr}\right|\in I\right\}$, $S_I=\mathrm{card}\left(\Omega_I\right)$ and I represents the interval of interest. Naturally, values of Δ_I near to one indicate low bias. The intervals I of interest were [0.5,1), [1,1.5), [1.5,2), $[2,\infty)$, and $[0.5,\infty)$, with the latter summarizing the information arising from the previous ones.

3.3 Results

Table 1 shows the comparative results in terms of the bias index Δ_I. As expected, regardless from the estimation method, the precision of the item estimates increases when k and/or n increase too. Also, although JML* works better than JML, C-JML clearly outperforms both of them. In this ranking among methods, it is also important to note as JML* performs more similar to JML than C-JML. Tables 2 and 3 illustrate the results in terms of RMSE for the values k=5, 10, 15 and k=20, 25, 30, respectively. These tables, beside to confirm the above considerations, also confirm that the JML underestimates the difficulties of easy items but overestimates those of difficult items. In other words, it overspread the difficulty estimates.

4. Discussion

This study compares the effect of two different techniques of bias reduction of the JML estimates of item parameters for a RM. The interest in this estimation method is simply due to the popularity of the software Winsteps, which appears to be the most widely used Rasch analysis program ([4]), and that effectively adopts the JML estimation method, as well as the C-JML bias-correction approach. Our simulations confirm that the bias (computed by the index Δ_I) is quite near to the conjectured factor *k/*(*k*-1), especially for large samples.

Table 1. Values of Δ_I for the three considered estimators, JML, JML*, and C-JML.

k	*I*	*k/(k*-1)	*n*=100			*n*=400			*n*=1000		
			JML	JML*	C-JML	JML	JML*	C-JML	JML	JML*	C-JML
5	[0.5,1)	1.25	1.36	1.32	1.09	1.29	1.28	1.03	1.28	1.28	1.02
	[1,1.5)	1.25	1.37	1.33	1.10	1.31	1.30	1.05	1.31	1.30	1.04
	[1.5,2)	1.25	1.38	1.34	1.11	1.33	1.32	1.06	1.36	1.35	1.08
	[2, ∞)	1.25	1.27	1.22	1.01	1.40	1.39	1.12	1.37	1.36	1.09
	[0.5, ∞)	*1.25*	*1.36*	*1.32*	*1.09*	*1.31*	*1.30*	*1.04*	*1.30*	*1.30*	*1.04*
10	[0.5,1)	1.11	1.15	1.12	1.03	1.12	1.12	1.01	1.12	1.12	1.01
	[1,1.5)	1.11	1.15	1.13	1.04	1.13	1.13	1.02	1.13	1.12	1.01
	[1.5,2)	1.11	1.18	1.15	1.06	1.13	1.13	1.02	1.13	1.13	1.02
	[2, ∞)	1.11	1.18	1.14	1.06	1.12	1.11	1.00	1.14	1.14	1.03
	[0.5, ∞)	*1.11*	*1.15*	*1.13*	*1.04*	*1.13*	*1.12*	*1.01*	*1.13*	*1.12*	*1.01*
15	[0.5,1)	1.07	1.08	1.06	1.01	1.08	1.08	1.01	1.08	1.08	1.01
	[1,1.5)	1.07	1.10	1.07	1.02	1.08	1.08	1.01	1.08	1.08	1.01
	[1.5,2)	1.07	1.08	1.05	1.01	1.08	1.07	1.01	1.08	1.08	1.01
	[2, ∞)	1.07	1.07	1.03	1.00	1.09	1.08	1.02	1.09	1.09	1.02
	[0.5, ∞)	*1.07*	*1.08*	*1.06*	*1.01*	*1.08*	*1.08*	*1.01*	*1.08*	*1.08*	*1.01*
20	[0.5,1)	1.05	1.08	1.06	1.03	1.07	1.07	1.02	1.06	1.05	1.00
	[1,1.5)	1.05	1.06	1.04	1.01	1.06	1.06	1.01	1.06	1.06	1.01
	[1.5,2)	1.05	1.10	1.08	1.05	1.05	1.05	1.00	1.07	1.06	1.01
	[2, ∞)	1.05	1.07	1.03	1.01	1.05	1.04	1.00	1.06	1.05	1.00
	[0.5, ∞)	*1.05*	*1.08*	*1.06*	*1.02*	*1.07*	*1.06*	*1.01*	*1.06*	*1.06*	*1.01*
25	[0.5,1)	1.04	1.06	1.04	1.02	1.04	1.04	1.00	1.05	1.05	1.01
	[1,1.5)	1.04	1.06	1.04	1.02	1.05	1.04	1.01	1.04	1.04	1.00
	[1.5,2)	1.04	1.08	1.05	1.04	1.04	1.03	1.00	1.05	1.05	1.01
	[2, ∞)	1.04	1.07	1.03	1.02	1.05	1.05	1.01	1.05	1.05	1.01
	[0.5, ∞)	*1.04*	*1.06*	*1.04*	*1.02*	*1.05*	*1.04*	*1.00*	*1.05*	*1.05*	*1.01*
30	[0.5,1)	1.03	1.04	1.02	1.00	1.03	1.03	1.00	1.04	1.04	1.00
	[1,1.5)	1.03	1.06	1.04	1.02	1.04	1.03	1.01	1.04	1.04	1.00
	[1.5,2)	1.03	1.06	1.04	1.03	1.04	1.03	1.00	1.04	1.04	1.01
	[2, ∞)	1.03	1.08	1.05	1.04	1.05	1.04	1.01	1.04	1.03	1.00
	[0.5, ∞)	*1.03*	*1.05*	*1.03*	*1.02*	*1.04*	*1.03*	*1.00*	*1.04*	*1.04*	*1.00*

Moreover, simulations confirm that the bias reduction techniques, JML* and C-JML, tend to effectively reduce the bias, without an inflating effect on the standard error of the estimators (with respect to the baseline level of the JML estimator). This may be due to a shrinkage effect on the estimates of both these approaches. In particular, by this comparison, it clearly appears that the C-JML method, based on correcting the JML estimate with the bias-correction factor (*k*-1)/*k*, outperforms the method suggested by Firth, based on a modification of the score function and leading to the JML* estimator. Indeed, in general, JML* shows little differences with respect to the bias of the JML estimator. Besides, an advantage of JML* over C-JML is that the former avoid the issue of the infinite estimate in the case of extreme patterns. But it should also be noted that the existence of JML* estimate for *ill-conditioned* datasets ([7]) has not yet been demonstrated.

Table 2. RMSE for the three considered estimators, JML, JML*, and C-JML, and for low values of *k*.

	k=5			*k*=10			*k*=15		
	JML	JML*	C-JML	JML	JML*	C-JML	JML	JML*	C-JML
n=100	0.41	0.38	0.24	0.29	0.28	0.24	0.27	0.26	0.24
n=400	0.30	0.29	0.12	0.17	0.17	0.11	0.15	0.14	0.12
n=1000	0.26	0.26	0.08	0.14	0.14	0.07	0.11	0.11	0.07

Table 3. RMSE for the three considered estimators, JML, JML*, and C-JML, and for higher values of *k*.

	k=20			*k*=25			*k*=30		
	JML	JML*	C-JML	JML	JML*	C-JML	JML	JML*	C-JML
n=100	0.26	0.25	0.24	0.26	0.25	0.24	0.25	0.24	0.23
n=400	0.13	0.13	0.11	0.13	0.13	0.12	0.13	0.12	0.12
n=1000	0.10	0.10	0.07	0.09	0.09	0.07	0.09	0.09	0.08

References

[1]. Andersen, E.B. (1980). *Discrete statistical models with social sciences applications*. Amsterdam: North-Holland.

[2]. Andrich D., Lyne A., Sheridan B., Lou G. (2003). *RUMM 2020* [Computer Software], Perth, Australia: RUMM Laboratory.

[3]. Bartolucci, F., Bellio, R., Salvan, A., Sartori, N. (2012). Modified Profile Likelihood for Panel Data Models. Available at SSRN repository: http://ssrn.com/abstract=2000666.

[4]. Cohen, J., Chan, T., Jiang, T., Seburn, M. (2008). Consistent estimation of Rasch item parameters and their standard errors under complex sample designs. *Applied Psychological Measurement*, 32, 289-310.

[5]. De Boeck, P. (2008). Random item IRT models. *Psychometrika*, 73, 533-559.

[6]. Firth, D. (1993). Bias reduction of maximum likelihood estimates. *Biometrika*, 80, 27-38.

[7]. Fischer, G.H. (1981). On the existence and uniqueness of maximum-likelihood estimates in the Rasch model. *Psychometrika*, 46, 59-77.

[8]. Linacre, J.M. (2009). *WINSTEPS®. Rasch measurement computer program*. [Computer Software]. Beaverton, Oregon: Winsteps.com.

[9]. R Development Core Team (2011). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

[10]. van den Wollenberg, A. L., Wierda, F. W., Jansen, P. G. W. (1988). Consistency of Rasch model parameter estimation: A simulation study. *Applied Psychological Measurement*, 12, 307-313.

[11]. Wang, W.C., Chen, C.T. (2005). Item parameter recovery, standard error estimates, and fit statistics of the Winsteps program for the family of Rasch models. *Educational and Psychological Measurement*, 65, 376-404.

[12]. Warm, T. A. (1989). Weighted likelihood estimation of in item response theory. *Psychometrika*, 54, 427-450.