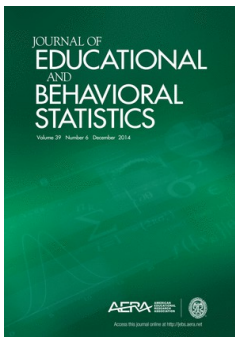

Journal of educational and behavioral statistics, ISSN 1076-9986
Volume 40, number 2 (april 2015)

Abstract
This article proposes a novel approach to detect differential item functioning (DIF) among dichotomously scored items. Unlike standard DIF methods that perform an item-by-item analysis, we propose the "LR lasso DIF method": logistic regression (LR) model is formulated for all item responses. The model contains item-specific intercepts, an effect of the sum score, and item-group interaction (i.e., DIF) effects, with a lasso penalty on all DIF parameters. Optimal penalty parameter selection is investigated through several known information criteria (Akaike information criterion, Bayesian information criterion, and cross validation) as well as through a newly developed alternative. A simulation study was conducted to compare the global performance of the suggested LR lasso DIF method to the LR and Mantel–Haenszel methods (in terms of false alarm and hit rates). It is concluded that for small samples, the LR lasso DIF approach globally outperforms the LR method, and also the Mantel–Haenszel method, especially in the presence of item impact, while it yields similar results with larger samples.

Abstract
When fitting hierarchical regression models, maximum likelihood (ML) estimation has computational (and, for some users, philosophical) advantages compared to full Bayesian inference, but when the number of groups is small, estimates of the covariance matrix (Σ) of group-level varying coefficients are often degenerate. One can do better, even from a purely point estimation perspective, by using a prior distribution or penalty function. In this article, we use Bayes modal estimation to obtain positive definite covariance matrix estimates. We recommend a class of Wishart (not inverse-Wishart) priors for Σ with a default choice of hyperparameters, that is, the degrees of freedom are set equal to the number of varying coefficients plus 2, and the scale matrix is the identity matrix multiplied by a value that is large relative to the scale of the problem. This prior is equivalent to independent gamma priors for the eigenvalues of Σ with shape parameter 1.5 and rate parameter close to 0. It is also equivalent to independent gamma priors for the variances with the same hyperparameters multiplied by a function of the correlation coefficients. With this default prior, the posterior mode for Σ is always strictly positive definite. Furthermore, the resulting uncertainty for the fixed coefficients is less underestimated than under classical ML or restricted maximum likelihood estimation. We also suggest an extension of our method that can be used when stronger prior information is available for some of the variances or correlations.

Abstract
In an earlier paper, we presented methods for estimating achievement gaps when test scores are coarsened into a small

number of ordered categories, preventing fine-grained distinctions between individual scores. We demonstrated that gaps can nonetheless be estimated with minimal bias across a broad range of simulated and real coarsened data scenarios. In this article, we extend this previous work to obtain practical estimates of the imprecision imparted by the coarsening process and of the bias imparted by measurement error. In the first part of this article, we derive standard error estimates and demonstrate that coarsening leads to only very modest increases in standard errors under a wide range of conditions. In the second part of this article, we describe and evaluate a practical method for disattenuating gap estimates to account for bias due to measurement error.

An Evaluation of Empirical Bayes's Estimation of Value-Added Teacher Performance Measures P. 190-222

Cassandra M. Guarino - Michelle Maxfield - Mark D. Reckase - Paul N. Thompson - Jeffrey M. Wooldridge

Abstract

Empirical Bayes's (EB) estimation has become a popular procedure used to calculate teacher value added, often as a way to make imprecise estimates more reliable. In this article, we review the theory of EB estimation and use simulated and real student achievement data to study the ability of EB estimators to properly rank teachers. We compare the performance of EB estimators with that of other widely used value-added estimators under different teacher assignment scenarios. We find that, although EB estimators generally perform well under random assignment (RA) of teachers to classrooms, their performance suffers under nonrandom teacher assignment. Under non-RA, estimators that explicitly (if imperfectly) control for the teacher assignment mechanism perform the best out of all the estimators we examine. We also find that shrinking the estimates, as in EB estimation, does not itself substantially boost performance.