Statistical Science
2009, Vol. 24, No. 3, 270–272
DOI: 10.1214/09-STS277A
Main article DOI: 10.1214/09-STS277

Discussion of Likelihood Inference for Models with Unobservables: Another View

Thomas A. Louis

1. INTRODUCTION

Lee and Nelder identify important issues and provide excellent advice and warnings associated with inferences and interpretations for models with unobserved, latent variables (random effects). Their discussion of prediction versus estimation goals is insightful and I have some sympathy with their call for use of comprehensive probability models. They provide a clear explanation of their h-likelihood approach and a spirited promotion of it. Unfortunately, the value and impact of the their advice are compromised by their singular focus on promoting h-likelihood. Their claim that it is an almost universally preferred approach is, to put it mildly, a stretch. The h-likelihood approach by no means "trumps" all competitors and has its own deficits. Over promotion makes the article more of an opinion-piece than a scientific comparison of approaches.

2. POINT/COUNTERPOINT

I identify and discuss principal points of (partial) agreement and of disagreement. Statements by Lee and Nelder are in *italics*; my responses and comments are in Roman.

2.1 Modeling Strategies

Lee and Nelder write, "*However, we believe that such a choice is inappropriate because the choice of an estimation method for a particular parameterization (marginal parameter) should not pre-empt the process of model selection.*" I agree. Estimation methods are a means to an end and usually not, themselves, the end (in methods research they can be the goal). Of course, the estimation method might influence model choice in that an inefficient method may miss important covariates and an inappropriate method may lead to bias. Sometimes the means/ends distinction gets blurred. For example, several years ago someone wrote to let me know that he thought the EM estimate was the absolute best; far better than the MLE!

Unified Probability Models are absolutely necessary: I do take issue with this claim. One should not discount the effectiveness of analyses and algorithms that are not fully probability-based or comprehensive. These have and will continue to play an important role. While a unified approach with marginals, conditionals, etc., all generated by a joint distribution is without question the ideal, often it is not attainable. Data limitations, limitations in scientific understanding and computing constraints can thwart use of this holy grail. Even attainment can be illusory because the unified model may not be correct and may mislead. So, while I favor the unified approach, I'm very comfortable with an approach that validly and effectively addresses a specific goal.

"*. . . so that care is necessary in making inferences about unobservables.*" Absolutely! Extreme care and caution are most definitely needed. Inferences on latent effects are always model-based to some degree, and some assumptions cannot be verified empirically. For example, models using the standard Poisson distribution as baseline rather than the more general negative binomial will "identify" unaccounted (extra-Poisson) variation and allocate it to a latent effect. If a negative binomial model is used, much of this variation will be absorbed into the baseline model. Both approaches can produce similar predictions of observable quantities, but will produce very different inference for latent effects. All modeling approaches need to deal with such issues, and the h-likelihood is not a panacea. In contrast, use of latent variable models and hierarchical models to generalize the mean and association structure of models for observeds is quite safe. Therefore, I agree with Lee and Nelder that focus on the prediction space rather than the parameter space avoids mis- or over-interpretation of parameter estimates.

2.2 H-likelihood and Competitors

"*. . . that when applied appropriately h-likelihood methods are both valid and efficient in such settings.*" It is most surely the case that in some settings, with

Thomas A. Louis is Professor, Department of Biostatistics, Johns Hopkins Bloombergh School of Public Health, Baltimore, Maryland, USA (e-mail: tlouis@jhsph.edu).