Guest Editorial: Advances in Statistical Models and Methods

Chandra A. Reynolds

Department of Psychology, University of California at Riverside, Riverside, California, United States of America

A variety of recent methods and models are presented in this special issue devoted to advances in behavioral genetic methodology, ranging from tutorials to the introduction of new models and procedures that invite further exploration given their promising advantages. The articles included in this special issue can be organized under one or more of the following emergent themes: (a) alternatives to maximum likelihood (ML) estimation, (b) advantages of mixed effects software for fitting biometrical models, (c) testing for latent heterogeneity, and (d) violation of standard assumptions.

Alternatives to Maximum Likelihood Estimation

Alternatives to customary ML estimation in behavioral genetic modeling have been sought in cases where distributional assumptions are not met or in the evaluation of complex high dimensional models. Two articles in this special issue address the benefits of Bayesian inference and semiparamteric ML estimation, respectively, in biometrical model fitting.

Bayesian inference and Markov-chain Monte Carlo (MCMC) methods for high complexity models. The authors van den Berg, Beem and Boomsma illustrate the advantages of Bayesian MCMC methods as a flexible and useful alternative to ML estimation in the case of complex computationally laden models (e.g., multithreshold phenotypes, longitudinally measured phenotypes, modeling G × E interaction or GE correlation). Accessible examples using the freely available BUGS software are illustrated with the application of bivariate ACE and ADE models to repeated measures data on hormone levels in young twins.

Semiparamteric maximum-likelihood and nonnormality. Markon introduces a semiparametric maximum likelihood (SPML) approach to the modeling of genetic and environmental influences on nonnormal phenotypes. The biometrical variance decomposition model is reframed in terms of a mixture moment structural model and is estimated alongside parameters of the nonnormal distribution. Simulation results and an applied example are presented which suggest some advantages of SPML over ML estimation, including evidence of greater accuracy in model fit.

Using Mixed Effects Software to Fit Models to Genetically Informative Data

Three articles in this issue make use of mixed effects procedures in widely available statistical packages to fit biometrical models (Beem and Boomsma; McArdle; Wang et al.). As highlighted in the individual articles by McArdle and by Beem and Boomsma in this issue, the benefits that these statistical packages afford include ease in data management and input, graphing capabilities and examining the tenability of model assumptions using built-in options. Nonetheless, tradeoffs may include restricted choices in modeling the biometric covariance structures in comparison to dedicated structural equation modeling (SEM) software (see Beem and Boomsma, this issue).

Longitudinal model fitting. Biometrical latent growth models (LGM) have been presented in prior work using specialized SEM software (e.g., McArdle & Hamagami, 2003). Such approaches are now augmented with the ability to use common statistical packages to fit LGM with twins or extended kinships in order to decompose growth parameter variance into genetic and environmental components. In this issue, McArdle provides a tutorial on fitting biometrical LGM using a variety of software, including the PROC MIXED procedure in the SAS package. Simulated and empirical data are examined in detail with accessible SAS scripts.

Haplotype association. Wang and colleagues make use of specialized multilevel modeling software (MlwiN) to consider the association of measured SNPs and inferred haplotypes with phenotypes measured longitudinally in multiethnic youth. The authors outline the incorporation of the haplotype trend regression technique (HTR; Zaykin et al., 2002) into a nonlinear latent growth model in order to test the association of inferred haplotypes with growth trajectories of blood pressure and ventricular mass.

Combined association-linkage. Combined linkage-association analyses have been typically performed using

Address for correspondence: Chandra A. Reynolds, Department of Psychology, University of California at Riverside, Riverside, California 92521, USA. E-mail:chandra.reynolds@ucr.edu

https://doi.org/10.1375/twin.9.3.311 Published online by Cambridge University Press

specialized software (e.g., QTDT, Mx). Beem and Boomsma provide a detailed tutorial on the implementation of a family-based quantitative transmission disequilibrium test (QTDT) using the mixed effects capabilities of widely used standard statistical packages, including SPSS.

Latent Heterogeneity

Heterogeneity in a statistical sense implies moderation or interaction (Altman & Mathews, 1996). Behavioral genetic models may directly test heterogeneity of genetic and environmental effects in observable groups or classes (e.g., sex-limitation models), or relax the assumption of independence of genetic and environmental effects to determine whether particular genotypes may be differentially sensitive to environmental factors, for example. Three articles in this issue examine latent heterogeneity of sorts through the use of mixture models: (a) underlying heterogeneity in measurement models, or (b) testing the presence of genotype–environment interactions.

Factor mixture analysis. Muthen, Asaparouhov and Rebollo present factor mixture analysis (FMA) that in a single step performs latent class and factor analysis in order to address questions of underlying diversity in latent measurement models. The authors illustrate how FMA can be extended to biometrical analysis (denoted factor mixture heritability analysis) where the relative influence of genetic and environmental factors on class affiliation and differences in resulting biometrical components of variance can be examined. The FMA model is demonstrated using the MPLUS software with an analysis of alcohol abuse and dependence data measured at one point in time though the authors also discuss extensions to longitudinal data, that is, latent transition models and quantitative latent growth models.

Gene–environment interaction. The article by van der Sluis and colleagues examines the power to detect G × E interaction, when they feature as unspecified sources, using marginal maximum likelihood (MML). The MML method is applied to simulated monozygotic (MZ) data using the Mx program and is compared to the classical Jinks and Fulker (1970) approach that regresses MZ pair differences on MZ pair sums. A second article by Gilespie and Neale addresses G × E interaction in terms of latent heterogeneity in components of variance, described as a special case of factor mixture models (cf. FMA described above). In this case the software program Mx is used to fit a combined latent class — latent biometrical factor model to simulated data for both MZ and dizygotic (DZ) twin types.

Violation of Standard Assumptions

The remaining articles in the issue examine the tenability of standard assumptions of phenotypic or biometrical models.

Equal environments assumption. The tenability of the equal environments assumption (EEA) underlying twin models has been long debated in behavioral research and tested for using a variety of methods. Derks, Dolan and Boomsma present an extension to multivariate biometrical models for correlated phenotypes to examine whether the EEA holds. Though certain conditions must be present to adequately test for EEA, the current approach affords greater flexibility in that measured environmental factors do not have to be included to detect possible EEA violations.

Familial dependency in phenotypic factor analysis. Phenotypic factor analysis using ML estimation lead to biases in standard errors and overall model fit if independence of observations is violated, as in family designs. While this is a well-known circumstance, the extent of bias under various patterns of genetic and environmental influence has not been addressed before now. Rebollo and colleagues examine the degree of bias in extended kinships, that is, twins, parents and spouses, assuming various underlying AE and ACE structures and address a software package option in MPLUS that may ameliorate bias. The authors report that a robust maximum likelihood estimation procedure that adjusts for clustering reduces bias to inconsequential levels.

Notes and Acknowledgments

This special issue grew out of well-received paper and plenary sessions on advances in behavioral genetic methodology from the 2005 annual meeting of the Behavior Genetics Association in Hollywood, California, with half of the articles in this issue directly stemming from the meeting. I thank Nick Martin for the invitation and opportunity to compile this special issue highlighting recent advances. Heartfelt appreciation also goes to the several anonymous reviewers for their time and dedication to providing helpful comments, sometimes on short notice, as well as to the commitment and efforts of the authors themselves in meeting deadlines and responding to reviews. I hope that this special issue will encourage further empirical, methodical and theoretical work.

References

Altman, D. G., & Matthews, J. N. (1996). Statistics Notes: Interaction 1: Heterogeneity of effects. *British Medical Journal, 313*, 486.

Jinks, J. L., & Fulker, D. W. (1970). Comparison of the biometrical genetical, MAVA, and classical approaches to the analysis of human behavior. *Psychological Bulletin, 73*, 311–349.

McArdle, J. J., & Hamagami, F. (2003). Structural equation models for evaluating dynamic concepts within longitudinal twin analyses. *Behavior Genetics, 33*, 137–159.

Zaykin, D., Westfall, P., Young, S., Karnoub, M., Wagner, M., & Ehm, M. (2002). Testing association of statistically inferred haplotypes with discrete and continuous traits in samples of unrelated individuals. *Human Heredity, 53*, 79–91.