Semiparametric Maximum Likelihood Variance Component Estimation Using Mixture Moment Structure Models

Kristian E. Markon

Department of Psychology, University of Minnesota, Minneapolis, United States of America

Nonnormal phenotypic distributions introduce significant problems in the estimation and selection of genetic models. Here, a semiparametric maximum likelihood approach to analyzing nonnormal phenotypes is described. In this approach, distributions are explicitly modeled together with genetic and environmental effects. Distributional parameters are introduced through mixture constraints, where the distribution of effects are discretized and freely estimated rather than assumed to be normal. Semiparametric maximum likelihood estimation can be used with a variety of genetic models, can be extended to a variety of pedigree structures, and has various advantages over other approaches to modeling nonnormal data.

Nonnormal distributions are frequently encountered in genetic modeling, in the analysis of pathological as well as nonpathological phenotypes. The importance of nonnormal distributions in making inferences about genetic models, moreover, has been well demonstrated: inappropriately assuming that a distribution is normal may lead to biased parameter estimates (van den Oord et al., 2000) and to selection of overly complex models, in terms of relative as well as absolute fit (Allison et al., 1999; Markon & Krueger, 2004; van den Oord et al., 2000). Nonnormality is ubiquitous and greatly impacts model inference, underscoring the importance of methods for modeling nonnormal data.

Although existing methods for modeling nonnormal data have numerous advantages, they also have various disadvantages. Normalizing transformations (e.g., log, Blom), for example, are often preferable in cases where the scale of variables is arbitrary, but do not always eliminate nonnormality, leading to the problem that initially motivates their use. Adjusted fit statistics (Satorra & Bentler, 1994), similarly, are attractive in modeling nonnormality, but do not account for parameter estimation bias, and may be inappropriate when used for purposes other than was assumed in their original derivation (e.g., as a value in the calculation of another statistic). Nonnormal parametric methods (e.g., parametric Poisson models) are appealing, but may be analytically complex, and require specification of an appropriate distribution, which may be difficult. Finally, simulation methods (e.g., bootstrap, Monte Carlo) are powerful, but can be computationally prohibitive in terms of time as well as memory.

Here, a semiparametric maximum likelihood (SPML) approach to modeling of nonnormal phenotypes is described. Although it has proven valuable in other areas of statistical inference, SPML estimation has not been widely applied in genetic modeling. In the approach explored here, nonnormal distributions are explicitly modeled together with genetic and environmental effects. Distributional parameters are introduced through mixture models, where the distribution of an effect is discretized and freely estimated rather than assumed to be normal. SPML estimation can be used with a variety of genetic models, can be extended to a variety of pedigree structures, and has various advantages over other approaches to modeling nonnormal data. The performance of the SPML estimator is illustrated using a small-scale simulation, and an example using the approach is presented.

Background

Density Estimation Using Normal Mixture Models

One particularly attractive approach to modeling an unknown distribution is to represent the distribution using a normal mixture model (Hoff, 2000; Roeder & Wasserman, 1997). Under this approach, a density is represented using a mixture of normal subdistributions (Figure 1). Normal mixture model density estimation is similar in logic to many other nonparametric methods for representing densities, in that the distribution of interest is approximated by a mixture of some type of subdistribution. Histogram density estimates, for example, represent a density using a

Received 4 November, 2005; accepted 30 January, 2006.

Address for correspondence: Kristian E. Markon, Department of Psychology, University of Minnesota, Elliott Hall, 75 E. River Rd, Minneapolis, MN 55455, USA. E-mail: mark0060@tc.umn.edu

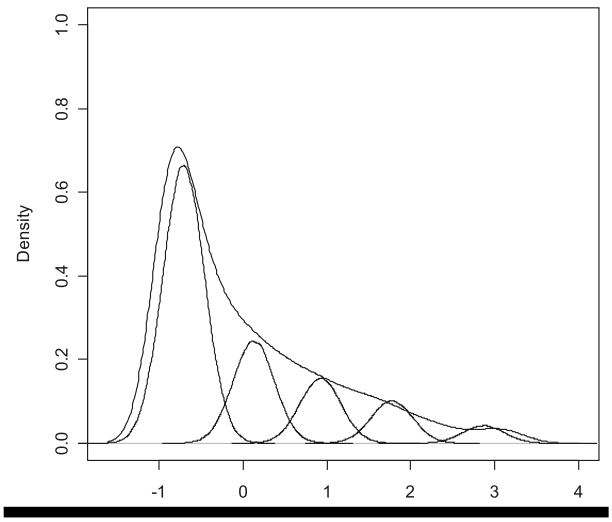

Figure 1

Mixture representation of a hypothetical nonnormal distribution.

mixture of rectangular subdistributions rather than a mixture of normal subdistributions.

In the normal mixture model approach to density estimation, an unknown density is estimated by estimating the location and size of each normal subdistribution in the mixture. Specifically, the mean vector of each normal subdistribution, the variance–covariance matrix of each subpopuation, and the proportion of individuals in each subdistribution, is estimated. Each of these sets of parameters can be assumed to differ across subdistributions, or can be assumed to be the same.

Mixture Moment Structure Models

The moments (e.g., mean, variance) of any subdistribution of a mixture model can be parameterized directly. Alternatively, the moments can be parameterized indirectly through some structural model of interest. In the latter case, the model is a form of moment structure model (MSM; i.e., structural equation model) in which a structural model is assumed to hold for each subdistribution of the mixture. The data is modeled as a mixture of normal subdistributions, each subdistribution described by a structural model. Constraints can be placed on the parameters of the MSM to equate or free parameters across subdistributions of the mixture.

SPML Variance Components Estimation

Mixture moment structure models provide a compelling framework for SPML estimation, as they allow for parameters representing an unknown distribution as well as for parameters representing a causal model. The structural model used for genetic variance components estimation can be reformulated as a mixture moment structure model, with certain parameters representing the distribution of the data, and other parameters representing the causal model of interest. The data is then modeled as a mixture of distributions, with the distributional parameters (i.e., means, proportions) allowed to differ across mixtures, but the causal parameters (i.e., covariances, loadings) constrained to be the same across mixtures.

Variance Components Model

To understand the SPML formulation of a genetic variance components model, first consider the standard parametric normal maximum likelihood (ML) variance components model:

$$\begin{aligned} \Sigma &= A\Sigma_a A' + C\Sigma_c C' + E\Sigma_e E' \\ \mu &= \nu + A\mu_a + C\mu_c + E\mu_e \end{aligned} \qquad [1]$$

Here, Σ represents the observed phenotypic covariance matrix for a set of relatives (e.g., twin pairs), and μ represents the observed phenotypic mean vector for that set of relatives. Σ_a, similarly, represents the expected correlation matrix among additive genetic factors for the set of relatives (e.g., a matrix of ones for pairs of monozygotic [MZ] twins, and a standardized matrix with .5 on the off-diagonal for dizygotic [DZ] twins); Σ_c represents the expected shared environmental correlation matrix (a matrix of ones for relatives raised in the same household), and Σ_e represents the nonshared environmental correlation matrix (generally a unit diagonal matrix). μ_a represents the mean vector for the additive genetic factors, μ_c represents the mean vector for the shared environmental factors, and μ_e represents the mean vector for the nonshared environmental factors. Generally, it is assumed that $\mu_a = \mu_c = \mu_e = 0$, and the phenotypic mean vector is therefore modeled solely by the intercept ν. A represents the set of additive genetic factor loadings, C represents the set of shared environmental factor loadings, and E represents the set of nonshared environmental factor loadings. The proportion of variance due to each etiologic effect is estimated through the magnitude of A, C, and E.

The SPML variance components model assumes that nonnormal data can be represented using a mixture of normal distributions. The parametric normal ML model in Equation 1 is assumed to hold for each subdistribution of the mixture, with some constraints. Distributional parameters are allowed to vary across subdistributions, but the causal parameters are constrained to be equal:

$$\begin{aligned} \Sigma &= A\Sigma_a A' + C\Sigma_c C' + E\Sigma_e E' \\ \mu_i &= \nu + A\mu_{ai} + C\mu_{ci} + E\mu_{ei} \end{aligned} \qquad [2]$$

The observed phenotypic distribution is thus modeled as a mixture of normal subdistributions, each subdistribution i having covariance matrix Σ and mean vector μ_i. Note that the subdistributions vary only in location, not in causal structure: while the subdistribution means μ_i are allowed to differ, the covariance matrix Σ is the same for each subdistribution, as are the loadings A, C, and E. The phenotypic means of

https://doi.org/10.1375/twin.9.3.360 Published online by Cambridge University Press

mixture subdistributions — which represent values of the unknown distribution — are modeled in terms of latent additive genetic, shared environmental, and nonshared environmental means μ_{ai}, μ_{ci}, and μ_{ei}.

In addition to the means of each subdistribution, the proportions of individuals in each subdistribution, P_i, are also estimated. The unknown distribution is therefore semiparametrically estimated through the locations of each subdistribution — that is, the mean vectors μ_{ai}, μ_{ci}, and μ_{ei} — and the proportion of individuals in each subdistribution, P_i. This corresponds to estimation of the values and probabilities defining the distribution, respectively. In this regard, in contrast to the parametric normal ML model, the latent mean vectors μ_{ai}, μ_{ci}, and μ_{ei} will generally not equal zero.

Standardization of Parameters

Although the covariance matrix and mean vectors are being fit to each subdistribution of the mixture, it is not each subdistribution that is of interest, but rather, the entire distribution that is being estimated semiparametrically. It is important to note the means and covariances within each subdistribution will be different from the means and covariances taken across subdistributions. If parameters are initially standardized to the mean and variances within each subdistribution, it will be necessary to restandardize to the mean and variances of the entire distribution as a whole.

The mean of a latent nonshared distribution k (e.g., the shared environmental distribution c), for example, is given by

$$\mu_k = \sum\nolimits_i P_i \mu_{ki} \qquad [3]$$

This is a standard formula for the mean — that is, the sum of the values multiplied by their probabilities. It is important to note in this regard that the latent means of each subdistribution (e.g., μ_{ki}) here represent values of the unknown density that is being semiparametrically estimated. The variance of the latent nonshared environmental distribution, similarly, is given by

$$\begin{aligned} \sigma_k^2 &= \sum\nolimits_i P_i (\mu_{ki} - \mu_k)^2 \\ \Sigma_k &= \sum\nolimits_i P_i (\mu_{ki} - \mu_k)\,(\mu_{ki} - \mu_k)' \end{aligned} \qquad [4]$$

The means and variances for the additive genetic and shared environmental distributions would be given by similar equations. Using these means and variances, the original parameter estimates can then be restandardized using standard formulae (e.g., Bollen, 1989).

Parameter Constraints

Implicit in Equations 1 to 4 is an assumption that relatives have identical latent values, and that each subdistribution i is defined by a single value of the latent distribution (i.e., for a mean vector $\mu_i = [\mu_{i1}, \mu_{i2}, \mu_{ir} \ldots]$, the assumption is $[\mu_{i1} = \mu_{i2} = \mu_{ir} \ldots]$, where r is an index of the relative). This assumption may be reasonable under certain circumstances, for example, when modeling the latent shared environmental distribution, where relatives' values are often assumed to be perfectly correlated, and each family can be assigned to some subdistribution i that reflects a value along the shared environmental distribution. However, in general, it will be necessary to relax the assumption to allow relatives to have different values of the latent variable.

In the general case, the latent distribution is multivariate, not univariate, and a multivariate mixture subdistribution is needed to represent each section of the latent multivariate space (imagine, for example, that the distribution in Figure 1 is bivariate rather than univariate, and the mixture subdistributions are similarly bivariate). If relatives' latent values are assumed to be perfectly correlated (e.g., a latent shared environmental distribution) this multivariate distribution can be reduced to a univariate distribution. If they are not assumed to be perfectly correlated, however (e.g., a latent genetic distribution), a multivariate distribution must be modeled, and that multivariate distribution is subject to nonlinear constraints of the model (e.g., with DZ twin pairs, the latent bivariate genetic distribution must be estimated subject to a constraint that the implied latent correlation equals .5).

In cases where the relatives' latent values are assumed to be perfectly uncorrelated — as is typically the case with the nonshared environmental distribution — parameterization can be simplified somewhat. In that case, each relative can take on any value of the latent distribution, but the possible latent values for each relative are constrained to be the same, as are the probabilities of each value. Assume that there are R relatives in each pedigree, for example, and each relative can take on M possible values θ_1, θ_2, θ_m, ... θ_M of the latent distribution, each value having probability $P(\theta_1)$, $P(\theta_2)$, $P(\theta_m)$, ... $P(\theta_M)$. Then the total number of subdistributions I in the mixture is $I = M^R$, a subdistribution for each possible combination of relatives' latent values. Because the latent values are assumed to be perfectly uncorrelated, however, the probability of each subdistribution i, P_i, is obtained simply by multiplying the marginal probabilities of the latent values represented in subdistribution i. For example, the probability of being in a subdistribution i where the mean vector for two relatives is given by $\mu_i = (\theta_1, \theta_2)$, is given by $P_i = P(\theta_1)\ P(\theta_2)$. In this way, the multivariate latent distribution is fully specified in terms of its marginal distributions, which are constrained to be equal.

When the relatives' latent values are assumed to be correlated at some specific value, as is typically the case with the additive genetic distribution, parameterization of a latent distribution can be more complicated. In this case, nonlinear constraints must be placed on the possible values of the latent distribution and the probabilities of those values in order that the estimated latent distribution has an implied correlation equal to what is assumed a priori. For example, if the two additive genetic values for two relatives are assumed to be correlated .5, constraints must be placed on the

https://doi.org/10.1375/twin.9.3.360 Published online by Cambridge University Press

Table 1

Results of Monte Carlo Simulation

			Bias				Variance		
	Fit	ν	*A*	*C*	*E*	ν	*A*	*C*	*E*
Normal									
250	ML	.993	−.004	−.034	−.051	.000	.0027	.0415	.0437	.0018
	SPML	.007	−.004	−.054	−.054	−.058	.0027	.0573	.0477	.0099
500	ML	.990	−.001	−.027	−.021	.003	.0015	.0195	.0226	.0010
	SPML	.010	−.001	−.039	−.019	−.076	.0015	.0271	.0235	.0098
1000	ML	.985	.000	−.013	−.009	.001	.0008	.0091	.0107	.0005
	SPML	.015	.000	−.015	−.011	−.118	.0008	.0121	.0122	.0136
2000	ML	.978	−.002	−.007	−.003	.000	.0004	.0045	.0053	.0003
	SPML	.022	−.002	−.006	−.003	−.159	.0004	.0043	.0051	.0170
Nonnormal									
250	ML	.000	−.001	−.035	−.042	−.002	.0017	.0307	.0365	.0012
	SPML	1.00	−.001	−.017	−.031	−.075	.0017	.0141	.0189	.0048
500	ML	.000	−.003	−.019	−.018	.002	.0008	.0161	.0168	.0006
	SPML	1.00	−.002	−.005	−.014	−.040	.0008	.0048	.0072	.0018
1000	ML	.000	.000	−.004	−.016	.001	.0004	.0067	.0089	.0003
	SPML	1.00	.000	−.007	−.004	−.037	.0004	.0020	.0028	.0007
2000	ML	.000	.000	.001	−.010	.001	.0002	.0031	.0040	.0001
	SPML	1.00	.000	−.011	−.001	−.039	.0002	.0012	.0016	.0006

Note: Values in the table are the bias and variance of parameter estimates in each condition using each estimation method, as well as the proportion of replication samples in which each estimation method resulted in better fit according to the Bayesian Information Criterion (BIC). Conditions are described in detail in the text. *ML* indicates parametric normal ML estimation; *SPML*, semiparametric ML estimation. ν indicates intercept parameter; *A*, additive genetic path; *C*, shared environmental path; *E*, nonshared environmental path.

values θ_1, θ_2, θ_m, ... θ_M of the latent distribution, and on the probabilities $P(\theta_1)$, $P(\theta_2)$, $P(\theta_m)$, ... $P(\theta_M)$, such that the implied correlation between relatives' values is equal to .5.

Simulation

In order to evaluate the performance of SPML variance components estimation, a small-scale Monte Carlo simulation was conducted. The goal of the simulation was to compare SPML variance components estimation to parametric normal ML estimation in cases where the population is nonnormal, to determine whether SPML improves the accuracy of parameter estimates.

Methods

Data were simulated from two population conditions: one in which the nonshared environmental (E) distribution was assumed to be normal, that is, $\sim N(0,1)$, and another in which the E distribution was assumed to be nonnormal, that is, $\sim \Gamma(.25,.50)$. In all conditions, the additive genetic (G) and shared environmental (C) distributions were assumed to be normal, that is, $\sim N(0,1)$. The source of nonnormality was restricted to E in the current simulations because it was relatively tractable computationally, and represents an important case where there is no a priori hypothesis about the nature of the nonnormality. In all conditions the proportion of A, C, and E variance was assumed to be .3, .2, and .5, respectively.

Samples of simulated MZ and DZ twin data were generated, each sample comprising a total of 250, 500, 1000, or 2000 twin pairs, with equal numbers of MZ and DZ pairs. There were 500 Monte Carlo replications in each of the eight conditions created by crossing population distribution with sample size. In each sample, ACE models were fit using both estimation methods, parametric normal ML estimation and SPML estimation. SPML estimation was performed using five mixture subdistributions to represent the E distribution.

Results

Table 1 presents the bias and variance of the parametric ML and SPML parameter estimates for each condition. The bias and variance of SPML estimates was comparable to that of ML estimates, especially under conditions in which the population distribution was nonnormal. Among nonnormal population conditions, the bias and variance of SPML E estimates was comparable to but slightly larger than that of ML estimates. The bias and variance of A and C estimates, in contrast, was generally comparable to or slightly smaller than that of ML estimates. Among normal population conditions, the bias and variance of SPML estimates was comparable to that of ML estimates, with the exception of the SPML E estimates, which were notably larger in terms of bias and variance than

https://doi.org/10.1375/twin.9.3.360 Published online by Cambridge University Press

Table 2

ML and SPML Variance Components Estimates for the MPQ Alienation Scale

	ML	*SPML*
ln(L)	–3241.51	–2462.74
k	4	12
AIC	3255.18	2503.77
BIC	3245.51	2474.74
ν	.002	.002
A	.516	.533
C	.248	.000
E	.820	.846

Note: *ML* indicates parametric normal maximum likelihood estimation; *SPML*, semiparametric maximum likelihood estimation. *ln(L)* indicates log-likelihood; *k*, number of parameters; *AIC*, Aikaike's Information Criterion; *BIC*, Bayesian Information Criterion; ν, intercept parameter; *A*, additive genetic path; *C*, shared environmental path; *E*, nonshared environmental path.

the ML estimates. Inspection of individual replications suggests that the larger bias and variance of SPML E estimates under normal population conditions was due to overfitting a single normal distribution with multiple mixture subdistributions.

Table 1 also presents the per cent of samples in which ML or SPML models fit better according to the Bayesian Information Criterion (BIC). As is evident in the table, with close to perfect accuracy in nearly all conditions, the normal ML models fit better in samples from normal populations, and the SPML models fit better in samples from nonnormal populations.

Example

To illustrate differences between ML and SPML estimation in a specific dataset, ML and SPML approaches were used to estimate A, C, and E components of variance in the Alienation scale of the Multidimensional Personality Questionnaire (MPQ; Tellegen, 2000). The Alienation scale measures a general tendency to feel socially alienated, victimized, or mistreated; some of its items have somewhat quasi-paranoid content, reflecting the belief that others intend harm of some sort. Alienation scores are typically highly skewed in representative samples, providing a good example of a phenotype that poses challenges to variance components estimation due to violations of the normality assumption.

Data were obtained on 933 twin pairs (381 MZ and 552 DZ) from the Minnesota Twin Registry (Krueger & Johnson, 2002). Data were corrected for the linear and quadratic effects of age and sex, and were modeled using a univariate ACE variance components model (Equation 2). The SPML model assumed that nonnormality was due to nonnormality in the E distribution, which was modeled using five mixture subdistributions, the mean of the third distribution being set to zero for identification purposes. This model was also fit using normal ML estimation in Mplus (Muthén & Muthén, 2004; an example script is available upon request).

Table 2 presents the results of the ML and SPML variance components estimation methods. Substantial improvements in fit with SPML relative to ML estimation are evident in values of the log-likelihood, AIC (Akaike, 1973), and BIC (Schwarz, 1978), with the log-likelihood being much greater, and AIC and BIC being much lower, using SPML estimation. Parameter estimates were similar, indicating that the two methods converge on generally similar assessments of the proportion of A, C, and E variance. However, the estimates of A and E variance were slightly larger, and the estimate of C variance smaller, using SPML relative to parametric normal ML estimation.

Discussion

In SPML estimation, a nonnormal distribution is modeled simultaneously with effects of interest. In the approach described here, nonnormal distributions are parameterized using mixture models, where the distribution corresponding to an effect is discretized and freely estimated rather than assumed to be normal. In contrast to many other approaches to nonnormality which rely on transformations of distributions or fit statistics, SPML estimation models nonnormality directly, without making distributional assumptions about the data.

Current results suggest that, in terms of bias and efficiency of parameter estimates, SPML estimation is comparable to normal ML estimation in the analysis of normal and nonnormal data. An important exception to this is in the analysis of normal data, where use of SPML may lead to overfitting and less accurate estimates than parametric normal ML estimation. This conclusion is intuitively reasonable, in that analysis of normal data using normal models seems preferable.

In terms of model selection, simulation results suggest that in nonnormal samples SPML estimation may provide a more accurate assessment of model fit than normal ML estimation. The current simulations, however, were limited to comparisons between normal and semiparametric models that were identical in terms of the structural components of variance modeled. Expanded simulations will be necessary to determine whether or not SPML estimation affords more accurate assessments of the relative fit of models differing in structural complexity (e.g., AE versus ACE models, or independent versus common pathway models).

The small-scale simulation results presented here are generally consistent with previous studies in suggesting that under conditions of nonnormality different estimation methods produce similar parameter estimates, but differ in assessments of model fit (Lei & Lomax, 2005). Although more extensive simulations and analysis are necessary, results of the present simulation are consistent with those of previous studies, and suggest that SPML estimation

https://doi.org/10.1375/twin.9.3.360 Published online by Cambridge University Press

provides a viable alternative to classical normal ML estimation in parsing components of genetic and environmental variance.

The general topic of SPML model selection and inference is important and merits future research. Information-theoretic methods provide a particularly compelling approach to SPML model selection, as they successfully discriminate between mixture models is other contexts (McLachlan & Peel, 2000). The performance of information-theoretic statistics in selecting semiparametric variance component models is not well understood, but theoretical arguments (Barron & Cover, 1991; Vereshchagin & Vitanyi, 2004) suggest that they would likely exhibit similar patterns of performance across a range of settings. Another approach is to develop SPML fit statistics based on asymptotic sampling distribution theory. The distribution of likelihood-based statistics is increasingly well characterized for mixture models (Lo et al., 2001), and may provide a framework for semiparametric inference as well. One particularly important question is how novel theoretically derived indices of fit perform relative to existing nonparametric methods for evaluating model fit, especially randomization-based methods (e.g., bootstrap, permutation).

Another particularly important issue meriting further research is the number of mixture subdistributions to use when modeling effect distributions of interest. A nonnormal distribution can be approximated by an arbitrary number of values, and it is important to determine how many values are adequate to model the distribution. The approach implicitly adopted here is to treat the number of subdistributions as being arbitrarily large, as is done in quadrature density estimation. Another possibility is to treat the number of subdistributions as a model selection problem, and select the number of subdistributions based on a model fit statistic. Yet another possibility, delineated by Skrondal and Rabe-Hesketh (2004), is to treat the number of subdistributions as an optimization problem, and incorporate it in the estimation algorithm as an estimable parameter.

Finally, it is important to note that other methods for SPML modeling exist; use of mixture models is only one, albeit convenient, method of SPML estimation. Diao and Lin (2005), for example, have described a SPML variance components model for linkage analysis in which the data distribution is modeled indirectly through a transformation. Their approach is similar to other transformation-based approaches to modeling nonnormal data, except that the transformation is unspecified and estimated along with structural parameters of the model. Another form of SPML estimation is fully nonparametric maximum likelihood (NPML) estimation, where the distributional parameters are saturated with regard to available data, and the distribution is estimated nonparametrically subject to constraints implied by the structural model (Hoff, 2000; Owen, 2001). Developing and integrating these different forms of SPML estimation represents an important direction for future research.

SPML estimation provides a compelling method for modeling nonnormal data that complements other methods in facilitating inference about genetic models. Additional research is needed to clarify how SPML estimation compares to existing methods for nonparametric analysis, as well as to parametric nonnormal models that specify more precisely how nonnormal data is created. As a hybrid of fully nonparametric and parametric approaches, SPML estimation relies on assumptions that are more restricted than those of nonparametric methods, but more relaxed than those of parametric approaches. It will be important to determine the impact of these assumptions, and to clarify situations in which SPML estimation is more or less preferable to other approaches to modeling nonnormal genetic data.

References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov & F. Csaki (Eds.), *Proceedings of the Second International Symposium on Information Theory* (pp. 267–281). Budapest, Hungary: Akademiai Kiado.

Allison, D. B., Neale, M. C., Zannolli, R., Schork, N. J., Amos, C. I., & Blangero, J. (1999). Testing the robustness of the likelihood-ratio test in a variance-component quantitative-trait loci-mapping procedure. *American Journal of Human Genetics, 65,* 531–544.

Barron, A. R., & Cover, T. M. (1991). Minimum complexity density estimation. *IEEE Transactions on Information Theory, 37,* 1034–1054.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.

Diao, G., & Lin, D. Y. (2005). A powerful and robust method for mapping quantitative trait loci in general pedigrees. *American Journal of Human Genetics, 77,* 97–111.

Hoff, P. D. (2000). Constrained nonparametric maximum likelihood via mixtures. *Journal of Computational and Graphical Statistics, 9,* 633–641.

Krueger, R. F., & Johnson, W. (2002). The Minnesota Twin Registry: Current status and future directions. *Twin Research, 5,* 488–492.

Lei, M., & Lomax, R. G. (2005). The effect of varying degrees of nonnormality in structural equation modeling. *Structural Equation Modeling, 12,* 1–27.

Lo, Y., Mendell, N. R., & Rubin, D. (2001). Testing the number of components in a normal mixture. *Biometrika, 88,* 767–778.

Markon, K. E., & Krueger, R. F. (2004). An empirical comparison of information-theoretic selection criteria for multivariate behavior genetic models. *Behavior Genetics, 34,* 593–610.

https://doi.org/10.1375/twin.9.3.360 Published online by Cambridge University Press

McLachlan, G., & Peel, D. (2000). *Finite mixture models*. New York: Wiley.

Muthén, B., & Muthén, L. (2004). *Mplus user's guide* (3rd ed.). Los Angeles, CA: Muthén & Muthén.

Owen, A. B. (2001). *Empirical likelihood*. Boca Raton, FL: Chapman & Hall/CRC.

Roeder, K., & Wasserman, L. (1997). Practical Bayesian density estimation using mixtures of normals. *Journal of the American Statistical Association, 92,* 894–902.

Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye & C. C. Clogg (Eds.), *Latent variables analysis: Applications for developmental research* (pp. 399–419). Thousand Oaks, CA: Sage.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics, 6,* 461–464.

Skrondal, A., & Rabe-Hesketh, S. (2004). *Generalized latent variable modeling*. Boca Raton, FL: Chapman & Hall.

Tellegen, A. (2000). *Manual of the Multidimensional Personality Questionnaire*. Minneapolis, MN: University of Minnesota Press.

van den Oord, E. J. C. G., Simonoff, E., Eaves, L. J., Pickles, A., Silberg, J., & Maes, H. (2000). An evaluation of different approaches for behavior genetic analyses with psychiatric symptom scores. *Behavior Genetics, 30,* 1–18.

Vereshchagin, N. K., & Vitanyi, P. M. B. (2004). Kolmogorov's structure functions and model selection. *IEEE Transactions on Information Theory, 50,* 3265–3290.

https://doi.org/10.1375/twin.9.3.360 Published online by Cambridge University Press