



Assessing statistical significance in variance components linkage analysis: A theoretical justification



Gengxin Li^a, Yuehua Cui^{b,c,*}

^a Department of Mathematics and Statistics, Wright State University, Dayton, OH 45435, United States

^b Department of Statistics and Probability, Michigan State University, East Lansing, MI 48824, United States

^c School of Economics, Wuhan University of Technology, Wuhan, China

ARTICLE INFO

Article history:

Received 1 January 2016

Received in revised form 12 April 2016

Accepted 7 June 2016

Available online 16 June 2016

Keywords:

Identity-by-descent

Likelihood ratio test

Linkage analysis

Mixture chi-square

Quantitative trait loci

ABSTRACT

Variance components analysis has been a standard means in family-based genetic data analysis. The variance components technique treats genetic effects as random, and tests whether variance components are zero using the likelihood ratio (LR) test. In the literature, the asymptotic distribution of the LR is claimed to follow a mixture chi-square distribution, where the mixture proportions are calculated based on the binomial coefficients, a special case in Self and Liang (1987). This threshold calculation, however, often yields conservative test results as discussed in a number of studies, especially in multi-trait analyses. In this work, we show that the LR statistic asymptotically follows a mixture chi-square distribution where the mixture proportions depend on the estimated Fisher information matrix in both univariate and multivariate trait analyses. We provide a general approximation form for the distribution of the LR under the null hypothesis of no genetic effects. We illustrate our idea with three variance components models in genetic linkage analysis. The performance of the new threshold calculation method is demonstrated via simulation studies, and its application is further illustrated via a real data analysis.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Genes are the functional units responsible for inheriting biological variations in phenotypes from parents. Inheritance of these characteristics of quantitative traits is attributed to multiple genes working in a coordinated manner. The detectable regions of the genome that contain or are closely linked to causal genes are defined as quantitative trait loci (QTL). The association between QTL and closely linked genes contributing to phenotypic variations is termed as genetic linkage. In human linkage analysis, a variance components (VCs) model is a powerful tool for QTL mapping. In a VC analysis, genetic effects are often partitioned as additive, dominance and polygene effects whereby each one is treated as random (Amos, 1994; Xu and Atchley, 1995). Consider two alleles (A and a) at a gene locus, the additive genetic effect measures “the quantitative change in a trait that is associated with substituting one allele (one genotype) with that of another allele” (from Wikipedia) within a population. Specifically, “the additive effect is half of the difference between the mean of all cases that are homozygous for one version of the allele (a/a) compared to the mean of all cases that are homozygous for the other allele (A/A)” (from Wikipedia). The dominance effect is a non-linear genetic effect which measures the effect where one allele masks the contribution of a second allele at the same locus. For example, if the effect of allele A masks the effect

* Corresponding author at: Department of Statistics and Probability, Michigan State University, East Lansing, MI 48824, United States.
E-mail addresses: gengxin.li@wright.edu (G. Li), cui@stt.msu.edu (Y. Cui).

of allele a , then allele A is called a dominant allele and allele a is called a recessive allele (Lynch and Walsh, 1998). In case of over-dominance, the effect of heterozygote genotype Aa is larger than the effect of homozygous genotypes AA and aa . In plants, this is called the hybrid vigor. The polygene accounts for the effect of genes or QTLs not located on the same region as the tested genes (Lynch and Walsh, 1998). Thus, the polygene effect is treated as the background gene effect. In a VC linkage analysis, the interest is to test whether the variance component of a genetic effect (i.e., additive and/or dominance effect) is significantly different from zero. Likelihood ratio (LR) test is often applied for such purpose (Amos, 1994). Due to irregular conditions (i.e., parameter boundary problem), the asymptotic distribution of the LR does not follow a regular chi-square distribution, rather a mixture χ^2 distribution, where the mixture proportions are calculated with standard binomial coefficients, a special case discussed in Self and Liang (1987).

A number of studies have demonstrated the asymptotic distribution of LR under irregular conditions, see for example, Chernoff (1954), Self and Liang (1987) and Shapiro (1988). Chernoff (1954) showed that the limiting distribution of the LR has a mixture chi-square distribution when parameters of interest are on one side of a hyperplane, or in the first quadrant within an \mathbb{R}^2 space. Self and Liang (1987) extended the Chernoff's result to boundary cases. For a multivariate normal distribution, Kudô (1963) showed the geometric nature of the LR with respect to the slipping means and its mixture chi-square distribution. Kudô and Choi (1975) later on extended the argument to the inference of a one-sided test. Afterwards, Self and Liang (1987) and Shapiro (1988) demonstrated that the boundary problem is equivalent to that of a restricted mean under a multivariate Gaussian distribution. Shapiro (1985) provided one simple proof of the large sample distribution of the LR under the boundary condition for any convex cone. The author further developed a unified theory corresponding to the inequality constrained testing in multivariate normal population for general cases (Shapiro, 1988).

In a univariate linkage analysis with a variance components model, the boundary problem occurs with variance components testing. The distribution of LR in case 9 in Self and Liang (1987) has been commonly applied for a threshold determination (e.g., Amos, 1994 and Hanson et al., 2001). This result is based on the assumption that the unknown parameters are independent, leading to a diagonal variance-covariance matrix for unknown parameters. In reality, the above assumption could be easily violated. This matter consequently leads to conservative hypothesis tests (Allison et al., 1999).

In a bivariate linkage analysis, Amos et al. (2001) proposed an approach to approximate the null distribution of the LR. Again, their derivation assumes a diagonal Fisher information matrix. Additionally, they assumed that the genetic correlation between two traits is perfectly correlated either positively ($\rho = 1$) or negatively ($\rho = -1$) in their derivation, which is unrealistic in reality. Recently Morris et al. (2009) defined a constrained likelihood ratio test (CLRT). They applied Geyer's regularity (1994) to show the asymptotic distribution of the CLRT, but with uncertainty on whether the global M-maximizer can be attained. Because of this limitation, a simulation-based method was developed. However, the computational burden limits its applicability.

In this work, we revisit the LR statistic in testing variance components in linkage analysis under three genetic models (Almasy and Blangero, 2010; Xu and Atchley, 1995; Wang and Zeng, 2009; Nagy et al., 2014), and show that it follows a mixture chi-square distribution. According to the distribution, we provide a new calculation of the mixture proportions based on the estimated Fisher information matrix. The simulation results show the improved performance of the approximation based on the new mixture proportion calculation method. We applied the method to a Genetic Analysis Workshop (GAW) 18 dataset in which two genetic traits, systolic blood pressure (SBP) and diastolic blood pressure (DBP) were measured (Almasy et al., 2014). One can analyze each trait separately and estimate genetic (additive and dominance) effect size. Since the two traits are correlated, we may benefit by a joint analysis of multiple traits using a multi-trait VC model. The rest of this paper is organized as follows. Section 2 introduces three classical VC models with both univariate and multivariate trait analysis. The main result is illustrated in Section 3. Section 4 shows the performance of the new approximation via simulation examples. We demonstrate the utility of the method via a real data analysis in Section 5.

2. Motivating models

The variance components model in a genetic linkage study is typically composed of both fixed and random effects, with non-genetic effects treated as fixed and genetic effects treated as random (Goldgar, 1990; Amos, 1994). The total genetic effect is typically decomposed into additive, dominance and polygenic effects and all are treated as random (Amos, 1994). Below we briefly introduce three VC models and investigate the limiting distribution of the LR statistics under these model setups.

2.1. Model I

Assuming that K families are collected and the phenotype for the k th family is denoted by \mathbf{y}_k with n_k offsprings. For example, \mathbf{y}_k can be a vector of SBP or DBP measures for members in the k th family. Under the variance components model mapping framework, the total genetic effect is partitioned into several components expressed as

$$\mathbf{y}_k = \mu \mathbf{1}_{n_k} + \mathbf{a}_k + \mathbf{d}_k + \mathbf{g}_k + \mathbf{e}_k \quad (1)$$

where μ is the overall mean; $\mathbf{a}_k \sim N(\mathbf{0}, \mathbf{\Pi}_k \sigma_a^2)$ is the random additive effect of the major QTL; $\mathbf{d}_k \sim N(\mathbf{0}, \mathbf{\Delta}_k \sigma_d^2)$ is the dominance effect of QTL; $\mathbf{g}_k \sim N(\mathbf{0}, \mathbf{\Phi}_k \sigma_g^2)$ is the polygenic effect that reflects the effects of unlinked QTLs; $\mathbf{e}_k \sim N(\mathbf{0}, \mathbf{I}_k \sigma_e^2)$

is the residual term; $\mathbf{\Pi}_k$ is the matrix of proportion of alleles shared identical-by-descent (IBD) in the pedigree, and $\mathbf{\Delta}_k$ is the matrix of the probability that pedigree shares both alleles IBD, $\mathbf{\Phi}_k$ is a matrix of the expected proportion of alleles shared IBD, and \mathbf{I}_k is an identity matrix (Amos, 1994). When a testing QTL is not on a marker position, the variance–covariance of the phenotype for a pair of siblings y_{ki} and y_{kj} in the k th family can be expressed as

$$\text{cov}(y_{ki}, y_{kj} | \pi_{ij}, \delta_{ij}, \phi_{ij}, \theta) = \begin{cases} \sigma_a^2 + \sigma_d^2 + \sigma_g^2 + \sigma_e^2 & \text{if } i = j \\ b_{ij}(\theta, \pi_{ij})\sigma_a^2 + c_{ij}(\theta, \pi_{ij}, \delta_{ij})\sigma_d^2 + \phi_{ij}\sigma_g^2 & \text{if } i \neq j \end{cases}$$

where π_{ij} is the proportion of marker alleles shared IBD between the i th and j th sibs; δ_{ij} is the probability that a pair of sibs share two alleles IBD; ϕ_{ij} is the kinship coefficient between two sibs; θ is the recombination fraction, which indicates linkage between a trait locus and a marker. When a trait locus is not at the marker, θ takes none zero value. The coefficients of additive and dominance variances between two sibs are expressed as: $b_{ij}(\theta, \pi_{ij}) = \frac{1}{2} + (1 - 2\theta)^2(\pi_{ij} - \frac{1}{2})$ and $c_{ij}(\theta, \pi_{ij}, \delta_{ij}) = 4\theta^2(1 - \theta)^2 + (1 - 2\theta)^2\pi_{ij} + (1 - 2\theta)^4\delta_{ij}$ (Amos et al., 2001). In matrix notation, the phenotypic variance–covariance matrix among individuals in family k can be expressed as

$$\mathbf{\Sigma}_k = \mathbf{\Pi}_k\sigma_a^2 + \mathbf{\Delta}_k\sigma_d^2 + \mathbf{\Phi}_k\sigma_g^2 + \mathbf{I}_k\sigma_e^2.$$

The QTL effect is assessed via testing the following hypotheses,

$$\begin{cases} H_0 : \sigma_a^2 = \sigma_d^2 = 0 \\ H_1 : \sigma_a^2 > 0 \text{ or } \sigma_d^2 > 0. \end{cases} \quad (2)$$

Define $\theta_1 = \sigma_a^2, \theta_2 = \sigma_d^2, \theta_3 = \sigma_g^2, \theta_4 = \sigma_e^2$ and $\theta_5 = \mu$. Let $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5)^T \in \Omega = [0, \infty) \times [0, \infty) \times (0, \infty) \times (0, \infty) \times \mathbb{R}$ be the true parameter space. Under the null of hypothesis (2), the parameter space is reduced to $\boldsymbol{\theta}_0 = (\theta_{10}, \theta_{20}, \theta_{30}, \theta_{40}, \theta_{50})^T = (0, 0, \sigma_g^2, \sigma_e^2, \mu)^T \in \Omega_0 = \{0\} \times \{0\} \times (0, \infty) \times (0, \infty) \times \mathbb{R}$. Thus two parameters under the null are on the boundary of the true parameter space (Ω). In current applications, the LR statistic for the above test has been commonly claimed to be a mixture chi-square distribution, i.e., $\text{LR} \sim \frac{1}{4}\chi_2^2 : \frac{1}{2}\chi_1^2 : \frac{1}{4}\chi_0^2$, a special case (case 9) discussed in Self and Liang (1987). We will illustrate by simulation that this approximation produces conservative results and derive a new mixture proportion calculation method.

2.2. Model II

When a QTL has a pleiotropic effect on bivariate traits or when several QTLs are closely linked, a multivariate model considering two traits would be more powerful than a single trait analysis (Jiang and Zeng, 1995; Evans, 2002). To illustrate the idea, we consider a bivariate trait model assuming only additive effect. The VC model for family k can be expressed as

$$(\mathbf{y}_{k1}, \mathbf{y}_{k2}) = (\mu_1 \mathbf{1}_{n_k}, \mu_2 \mathbf{1}_{n_k}) + (\mathbf{a}_{k1}, \mathbf{a}_{k2}) + (\mathbf{g}_{k1}, \mathbf{g}_{k2}) + (\mathbf{e}_{k1}, \mathbf{e}_{k2}), \quad k = 1, \dots, K \quad (3)$$

where \mathbf{y}_{kh} is the h th ($h = 1, 2$) phenotypic vector for the k th family, for example, SBP and DBP traits in the GAW 18 dataset; $\boldsymbol{\mu}_h$ ($h = 1, 2$) is the overall mean for the h th phenotypic trait, $(\mathbf{a}_{k1}, \mathbf{a}_{k2})$ contains the random QTL additive effects for two phenotypic traits; \mathbf{g}_{kh} and \mathbf{e}_{kh} are the random polygene effect and residual for the h th phenotypic trait, respectively. All the random terms are assumed to be normally distributed with mean $\mathbf{0}$. The phenotypic variance–covariance matrix for family k can be expressed as

$$\text{Cov} \begin{pmatrix} \mathbf{y}_{k1} \\ \mathbf{y}_{k2} \end{pmatrix} = \begin{pmatrix} \sigma_{a1}^2 & \sigma_{a12} \\ \sigma_{a12} & \sigma_{a2}^2 \end{pmatrix} \otimes \mathbf{\Pi}_k + \begin{pmatrix} \sigma_{g1}^2 & \sigma_{g12} \\ \sigma_{g12} & \sigma_{g2}^2 \end{pmatrix} \otimes \mathbf{\Phi}_k + \begin{pmatrix} \sigma_{e1}^2 & \sigma_{e12} \\ \sigma_{e12} & \sigma_{e2}^2 \end{pmatrix} \otimes \mathbf{I}_k$$

where \otimes is the Kronecker product; σ_{a12} , σ_{g12} , and σ_{e12} are the covariances of the additive, polygene and residual terms between two traits, respectively. All the other terms are defined similarly as in Model I.

The hypothesis test to detect a QTL under the bivariate model is formulated as

$$\begin{cases} H_0 : \sigma_{a1}^2 = \sigma_{a2}^2 = \sigma_{a12} = 0 \\ H_1 : \sigma_{a1}^2 > 0 \text{ or } \sigma_{a2}^2 > 0. \end{cases} \quad (4)$$

Under the alternative, when either one of the variance terms is zero, the covariance term σ_{a12} is set to zero. Amos et al. (2001) proposed a mixture chi-square approximation to the LR statistics. Again, it gives conservative result as will be revealed in our simulation study.

2.3. Model III

Bivariate variance components linkage analysis model assuming both additive and dominance effects has been applied in real applications (e.g., Nagy et al., 2014). In this model (Model III), a bivariate trait VC model assuming both additive and

dominance effects is defined as

$$(\mathbf{y}_{k_1}, \mathbf{y}_{k_2}) = (\mu_1 \mathbf{1}_{n_k}, \mu_2 \mathbf{1}_{n_k}) + (\mathbf{a}_{k_1}, \mathbf{a}_{k_2}) + (\mathbf{d}_{k_1}, \mathbf{d}_{k_2}) + (\mathbf{g}_{k_1}, \mathbf{g}_{k_2}) + (\mathbf{e}_{k_1}, \mathbf{e}_{k_2}), \quad k = 1, \dots, K \tag{5}$$

where $(\mathbf{d}_{k_1}, \mathbf{d}_{k_2})$ represents the random dominance effects of major gene at the QTL for two phenotypic traits. The variance–covariance matrix between two traits in the k th family is given as,

$$\text{Cov} \begin{pmatrix} \mathbf{y}_{k_1} \\ \mathbf{y}_{k_2} \end{pmatrix} = \begin{pmatrix} \sigma_{a_1}^2 & \sigma_{a_{12}}^2 \\ \sigma_{a_{12}}^2 & \sigma_{a_2}^2 \end{pmatrix} \otimes \mathbf{\Pi}_k + \begin{pmatrix} \sigma_{d_1}^2 & \sigma_{d_{12}}^2 \\ \sigma_{d_{12}}^2 & \sigma_{d_2}^2 \end{pmatrix} \otimes \mathbf{\Delta}_k + \begin{pmatrix} \sigma_{g_1}^2 & \sigma_{g_{12}}^2 \\ \sigma_{g_{12}}^2 & \sigma_{g_2}^2 \end{pmatrix} \otimes \mathbf{\Phi}_k + \begin{pmatrix} \sigma_{e_1}^2 & \sigma_{e_{12}}^2 \\ \sigma_{e_{12}}^2 & \sigma_{e_2}^2 \end{pmatrix} \otimes \mathbf{I}_k.$$

The hypotheses to test major gene effect under this model are formulated as

$$\begin{cases} H_0 : \sigma_{a_1}^2 = \sigma_{a_2}^2 = \sigma_{a_{12}}^2 = \sigma_{d_1}^2 = \sigma_{d_2}^2 = \sigma_{d_{12}}^2 = 0 \\ H_1 : \sigma_{a_1}^2 > 0 \text{ or } \sigma_{a_2}^2 > 0 \text{ or } \sigma_{a_{12}}^2 > 0 \text{ or } \sigma_{d_1}^2 > 0 \text{ or } \sigma_{d_2}^2 > 0. \end{cases} \tag{6}$$

Similarly, under the alternative, when either one of the variance terms (additive or dominance) is zero, the corresponding covariance term is set to zero. In this case, the null distribution of LR and the mixture probabilities corresponding to the χ^2 components are not straightforward to calculate. To explore the geometric properties of the LR in the VC models described above, we propose a general approximation approach in next section.

3. Main results

For a random sample X_1, X_2, \dots, X_n of size n with a common density function $f(x, \boldsymbol{\theta})$, let $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_m)^T \in \Omega$ be the parameter vector of interest, and $\boldsymbol{\theta}_0$ be the corresponding population version. Let $\ell(\boldsymbol{\theta}) = \sum \log f(x, \boldsymbol{\theta})$ be the log-likelihood function.

Condition 1. Following Chernoff (1954), the following conditions are assumed:

- I. For every $\boldsymbol{\theta} \in G$ where G is a closure neighborhood centered at $\boldsymbol{\theta}_0$, the first three derivatives of $\ell(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$ exist for almost all \mathbf{x} ;
- II. For every $\boldsymbol{\theta} \in G$, $|\frac{\partial \ell(\boldsymbol{\theta})}{\partial \theta_i}|$ and $|\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j}|$ are bounded by a finite integrable function $K(\mathbf{x})$, and $|\frac{\partial^3 \ell(\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j \partial \theta_l}| < K(\mathbf{x})$ where $E[K(\mathbf{x})] < \infty$, $(i, j, l = 1, \dots, m)$;
- III. The information matrix $M(M_{ij} = -\frac{1}{n} E(\frac{\partial \ell(\boldsymbol{\theta})}{\partial \theta_i} \frac{\partial \ell(\boldsymbol{\theta})}{\partial \theta_j}))$ is nonsingular for $\boldsymbol{\theta} \in G$, and $\|M\| < \infty$ $(i, j = 1, \dots, m)$.

Proposition 1. Under Condition 1, there exists a vector $\hat{\boldsymbol{\theta}}_\zeta$ in Ω , such that $\hat{\boldsymbol{\theta}}_\zeta \rightarrow \boldsymbol{\theta}_0$ in probability, and $(\hat{\boldsymbol{\theta}}_\zeta - \boldsymbol{\theta}_0) = O_p(n^{-\frac{1}{2}})$.

Denote a local maximum estimator of $\boldsymbol{\theta}_0$ by $\tilde{\boldsymbol{\theta}}_\zeta$. In addition to the regularity conditions on the parameter set defined in Chernoff (1954), additional conditions are required to achieve the asymptotic equivalence of a local estimator.

Condition 2. Consider the following conditions:

- IV. $\hat{\boldsymbol{\theta}}_\zeta$ and $\tilde{\boldsymbol{\theta}}_\zeta$ are \sqrt{n} -consistent optimizers;
- V. The parameter set Ω is a nearly convex set at $\boldsymbol{\theta}_0$;
- VI. Condition vi in Theorem 3.2 in Shapiro (2000) is met.

Proposition 2. If Conditions 1 and 2 stated above are satisfied, $\hat{\boldsymbol{\theta}}_\zeta - \tilde{\boldsymbol{\theta}}_\zeta = o_p(n^{-\frac{1}{2}})$.

It is well known that a cone contains several desirable properties that may simplify the optimization problem. According to the arguments of Chernoff (1954) and Self and Liang (1987), a cone is defined as follows.

Definition 1. The set $\Omega \subset \mathbb{R}^m$ is approximated by a cone C_Ω at $\boldsymbol{\theta}_0$, if

$$\inf_{\mathbf{s} \in C_\Omega} \|\mathbf{s} - \mathbf{t}\| = o(\|\mathbf{t} - \boldsymbol{\theta}_0\|) \quad \text{for all } \mathbf{t} \in \Omega; \quad \inf_{\mathbf{t} \in \Omega} \|\mathbf{s} - \mathbf{t}\| = o(\|\mathbf{s} - \boldsymbol{\theta}_0\|) \quad \text{for all } \mathbf{s} \in C_\Omega.$$

Note that the cone C_Ω is positively homogeneous if $\mathbf{s} \in C_\Omega$, $c(\mathbf{s} - \boldsymbol{\theta}_0) + \boldsymbol{\theta}_0 \in C_\Omega$ when $c \geq 0$. Moreover, $C_\Omega - \boldsymbol{\theta}_0$ with vertex at the origin is acquired by translating the cone C_Ω with vertex at $\boldsymbol{\theta}_0$. Thus, Ω can be approximated by a closed convex cone C_Ω with vertex at $\boldsymbol{\theta}_0$.

Proposition 3. When $\boldsymbol{\theta} = \mathbf{0}$, F is the distribution of the MLE $\hat{\boldsymbol{\theta}}_\zeta$ based on one realization \mathbf{Y} with the population distribution $N(\boldsymbol{\theta}, M^{-1})$ where $\boldsymbol{\theta} \in C_\Omega - \boldsymbol{\theta}_0$. If all previous conditions hold, $n^{\frac{1}{2}}(\hat{\boldsymbol{\theta}}_\zeta - \boldsymbol{\theta}_0)$ converges weakly to F , a multivariate normal distribution with mean $\mathbf{0}$ and covariance matrix M^{-1} .

See the proof of Theorem 2 in Self and Liang (1987) for a proof.

Condition 3. Further assume the following condition:

VII. Let C_{Ω_0} and C_{Ω_1} be two closed convex cones with vertex at θ_0 to approximate Ω_0 and Ω_1 . Then $C_{\Omega_0} - \theta_0$ with vertex at origin is also a closed convex cone by translating C_{Ω_0} at θ_0 .

Theorem 3.1. If Conditions 1–3 stated above hold and when $\theta = \theta_0$, the large sample distribution of the LR statistic is the same as that of testing $\theta \in C_{\Omega_0}$ against $\theta \in C_{\Omega_1}$ based on one realization \mathbf{Y} generated from a population distribution $N(\theta, M^{-1})$. Moreover, the LR statistics is distributed as a mixture chi-square distribution with the tail probability given as

$$\Pr(LR > c) = \sum_{i'=1}^{l(q)} \Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^{i'}) \Pr(\chi_{\text{rank}(T_{v^*0}|\psi_{v^*|\mathbf{Z}}^{i'})}^2 > c)$$

where c is a positive constant; $\Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^{i'})$ is the mixing proportion corresponding to the chi-square component where $\psi_{v|\mathbf{Y}}^{i'}$ is a set of all faces of \mathbf{Y} ; $l(q)$ is the total number of sets of \mathbf{Y} ; $\sum_{i'=1}^{l(q)} \Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^{i'}) = 1$; $\text{rank}(T_{v^*0}|\psi_{v^*|\mathbf{Z}}^{i'})$ denotes the degrees of freedom of the corresponding chi-square distribution; and \mathbf{Z} follows a multivariate normal distribution with mean zero and identity covariance.

The proof is given in the [Appendix](#). Following [Theorem 3.1](#), we now evaluate the distribution of the LR statistic for the three models described in [Section 2](#). For Model I, the distribution of the LR statistic is approximated as,

$$\Pr(LR > c) = \frac{\pi - \cos^{-1} \rho_{12}}{2\pi} P(\chi_2^2 > c) + \frac{1}{2} P(\chi_1^2 > c) \quad (7)$$

where ρ_{12} is the correlation between the estimates of additive and dominance effects.

For Model II, the tail probability of LR is calculated as

$$\Pr(LR > c) = \frac{\pi - \cos^{-1} \rho_{12}}{2\pi} P(\chi_3^2 > c) + \frac{1}{2} P(\chi_1^2 > c) \quad (8)$$

where ρ_{12} is the correlation of additive effects between two traits.

For Model III, the tail probability of LR can be expressed as

$$\Pr(LR > c) = w_0 P(\chi_6^2 > c) + w_1 P(\chi_4^2 > c) + w_2 P(\chi_3^2 > c) + w_2 P(\chi_2^2 > c) + w_3 P(\chi_1^2 > c) \quad (9)$$

where w_0, w_1, w_2, w_2 , and w_3 are the weights. The detailed derivation of these results is rendered in the [Appendix](#) section.

4. Simulation

We designed simulations to evaluate the limiting distribution of the LR. The results of the new approximation in univariate and multivariate analysis were compared with those commonly applied as described in [Self and Liang \(1987\)](#) and [Amos and Andrade \(2001\)](#). In brief, [Self and Liang \(1987\)](#) method was primarily derived from the univariate model, and we directly extended it to the multivariate cases without considering the covariance terms. The method adopted by [Amos and Andrade \(2001\)](#) was developed from Self and Liang's approach plus considering covariance terms in the multivariate trait models. However, both approaches assumed that the unknown parameters were independent.

In accordance with the genetic models described in [Section 2](#), the genotype and phenotype data were simulated from outbred populations. In an outbred population, four different alleles could segregate in a family. The marker genotypes were generated from these four basic alleles with equal frequency. The number of families was assumed to be 20 and 40 with each having 5 offspring. Phenotypes of multiple traits were assumed to follow a multivariate normal distribution. The parameters of Models I, II, III were given in [Table 1](#). The limiting distributions of our method for the three motivating models were given in earlier sections. For the method in [Self and Liang \(1987\)](#), it is given as $\frac{1}{4}\chi_2^2 : \frac{1}{2}\chi_1^2 : \frac{1}{4}\chi_0^2$ for Model I, $\frac{1}{4}\chi_2^2 : \frac{1}{2}\chi_1^2 : \frac{1}{4}\chi_0^2$ for Model II and $\frac{1}{16}\chi_4^2 : \frac{4}{16}\chi_3^2 : \frac{6}{16}\chi_2^2 : \frac{4}{16}\chi_1^2 : \frac{1}{16}\chi_0^2$ for Model III. We denoted this method as SL in later presentation. For the one described in [Amos and Andrade \(2001\)](#), it is given as $\frac{1}{4}\chi_2^2 : \frac{1}{2}\chi_1^2 : \frac{1}{4}\chi_0^2$ for Model I, $\frac{1}{4}\chi_3^2 : \frac{1}{2}\chi_1^2 : \frac{1}{4}\chi_0^2$ for Model II and $\frac{1}{16}\chi_6^2 : \frac{4}{16}\chi_4^2 : \frac{2}{16}\chi_3^2 : \frac{4}{16}\chi_2^2 : \frac{4}{16}\chi_1^2 : \frac{1}{16}\chi_0^2$ for Model III. We referred to this method as Amos in later presentation. Two different simulation designs were considered with one considering 20 families and 5 offspring in each family (denoted as 20×5 design), and another one considering 40 families with 5 offspring in each family (denoted as 40×5 design). For each case, the result of 1000 simulation replications was recorded.

A quantile plot of the empirical error rates vs the nominal levels is given in [Fig. 1](#). Under Model I, the proposed method gives error rates quite close to the nominal levels, while SL underestimates the error rates. Since the threshold determination for SL and Amos was the same, the result for using Amos method was omitted. Under Model II, the new method gives quite accurate approximation compared to the other two methods which largely overestimate the error rates. Under Model III, SL overestimates the error rates and Amos tends to underestimate the error rates. Although our new approximation underestimates the error rates under the 20×5 design, the approximation improves when the number of family increases to 40. Overall, the 40×5 design gives better approximation than the 20×5 design does, which is consistent with the large sample property. In summary, the new approximation method shows improved approximation compared to other two methods under the three analytical models.

Table 1
Parameter setting under the null hypothesis.

Model	Parameters													
	$\sigma_{a_1}^2$	$\sigma_{a_2}^2$	$\sigma_{a_{12}}$	$\sigma_{d_1}^2$	$\sigma_{d_2}^2$	$\sigma_{d_{12}}$	$\sigma_{g_1}^2$	$\sigma_{g_2}^2$	$\sigma_{g_{12}}$	$\sigma_{e_1}^2$	$\sigma_{e_2}^2$	$\sigma_{e_{12}}$	μ_1	μ_2
Model I	0	–	–	0	–	–	2	–	–	2.5	–	–	3	–
Model II	0	0	0	0	0	0	2	2	1.8	2.5	2.5	2.2	4	2
Model III	0	0	0	0	0	0	2	2	1.8	2.5	2.5	2.2	4	2

Table 2
Information of selected families in GAW 18 dataset.

	Families																
	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9	f_{10}	f_{11}	f_{12}	f_{13}	f_{14}	f_{15}	f_{16}	f_{17}
Number of offspring	10	5	5	5	5	8	4	2	4	8	4	5	4	4	1	8	7

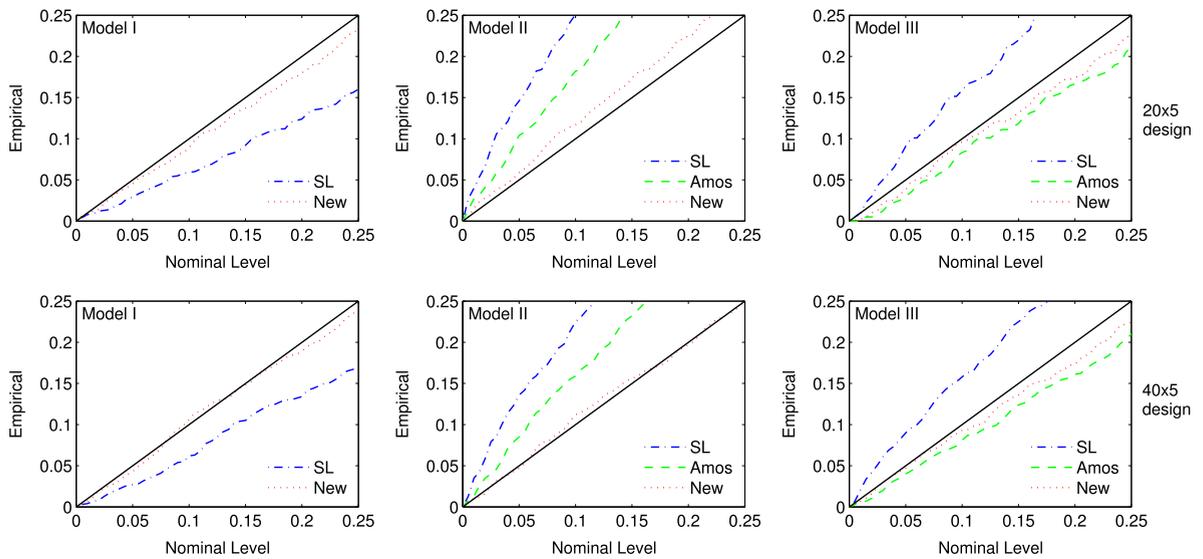


Fig. 1. The quantile plot of the empirical error rates under different models for the three methods. SL refers to the method proposed in [Self and Liang \(1987\)](#). Amos refers to the method described in [Amos and Andrade \(2001\)](#).

5. Real data application

We applied the proposed method to a dataset offered by the Genetic Analysis Workshop (GAW) 18 ([Almasy et al., 2014](#)). The data contain whole-genome single nucleoid polymorphism (SNP) data sequenced from 20 large pedigrees with blood pressure measurements. We selected 17 independent families with various number of offsprings within each family. [Table 2](#) lists these 17 independent families and the number of offspring in each family. The number of offspring in a family varies from 1 to 10, and the average number of offspring per family is around 5. Two quantitative traits, systolic blood pressure (SBP) and diastolic blood pressure (DBP), were measured. The goal is to identify SNP variants that are associated with SBP or DBP, or both. We applied the univariate trait analysis to SBP and DBP separately, and to SBP and DBP jointly for the multi-trait analysis. Individual blood pressure was measured over several time points and we picked time 1 in this analysis. Details about the data generating process can be found at [Almasy et al. \(2014\)](#). We selected a genetic region between 60cM and 100cM on chromosome 3 which contains 21 SNP markers. The aforementioned three variance components models were applied to this data. We compared the results with two other methods, namely SL and Amos, described in the simulation section.

[Fig. 2](#) displays the plot of $-\log_{10}(p\text{-values})$ against the 21 SNPs in the analyzed region (60–100cM), with the horizontal axis representing the actual position of the SNPs (cM) analyzed with the three models. The two figures on the top row show the results analyzed with Model I where each blood pressure trait is analyzed separately. The bottom left and right plots show the results analyzed with Models II and III respectively. Recall that Models II and III are bivariate trait analysis where SBP and DBP are jointly analyzed. For the univariate analysis fitted with Model I, we observed consistently smaller p -values with our method compared to the SL method. This is consistent with the simulation result in which SL underestimates the error rate (see the left two plots in [Fig. 1](#)). For the bivariate trait analysis fitted with Model II, the p -values calculated with our method are larger than the other two. This pattern is consistent with the simulation results in which both SL and Amos

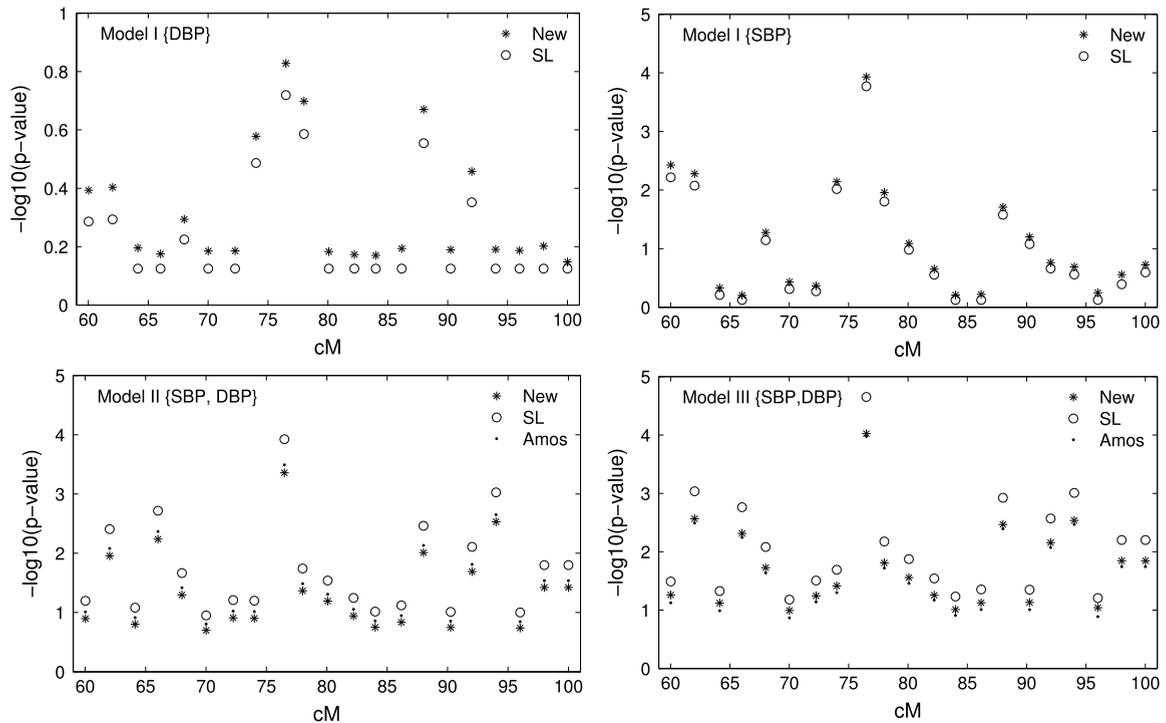


Fig. 2. The plot of $-\log_{10}(p\text{-value})$ under the three models. New refers to the proposed method; SL refers to the method proposed in [Self and Liang \(1987\)](#); and Amos refers to the method described in [Amos and Andrade \(2001\)](#). The two plots on the top row show the univariate results analyzed with Model I. The two plots at the bottom show the multi-trait results analyzed with Models II and III respectively.

Table 3

List of p -values for SNP rs704530 (located at 76.49cM) calculated with different methods under Models I, II and III.

P-value	Model I		Model II	Model III
	SBP	DBP	{SBP, DBP}	{SBP, DBP}
SL	1.7E-04	0.1909	1.2E-04	2.2E-05
Amos	–	–	3.2E-04	1.01E-04
New	1.2E-04	0.1486	4.4E-04	9.5E-05

overestimate the error rate (see the middle two plots in [Fig. 1](#)), hence tend to give larger false positives. For Model III (bottom right plot in [Fig. 2](#)), the p -values obtained with our method are between the ones calculated with SL and Amos method. Again, this result is consistent with the simulation results shown in the right two plots in [Fig. 1](#) in which SL overestimates the error rate and Amos underestimates the error rate. Although the p -values obtained with the three methods do not show large difference, the results do confirm the benefit of applying the newly proposed threshold determination method.

The univariate analysis results indicate that there is one significance SNP marker at 76.49cM (rs704530) which shows significant association with SBP, but not with DBP. The p -value for this SNP marker is $1.2E-04$ with our method and is $1.7E-04$ with SL. [Table 3](#) lists the p -values for this SNP under the three models. It is worth mentioning that our new method gives very consistent p -values for this SNP, which makes biological sense since we may not get benefit by the joint analysis as this SNP is not associated with the DBP trait at all. However, SL method gives smaller p -values in Model III compared to the p -values obtained with Models I and II. This is due to the nature of the method itself which overestimates the error rate as shown in the simulation study ([Fig. 1](#)).

6. Discussion

In a family-based genetic analysis, it has been proven that the variance components model is an efficient and powerful tool for QTL mapping. In addition, many studies have shown that multiple trait analysis provides higher power and mapping precision in QTL detection compared to a univariate trait analysis, especially when traits are highly correlated ([Jiang and Zeng, 1995](#); [Evans, 2002](#)). Although likelihood ratio test has been broadly applied in assessing the significance of different variance components, there is a lack of rigorous investigation on the limiting distribution of the LR statistic under the three motivating genetic models ([Amos and Andrade, 2001](#)). As shown by the simulation studies, inappropriate application of threshold determination could result in poor estimation of error rates. This is further confirmed by a real data analysis.

In this paper, we focused on three VC genetic models covering univariate and multivariate trait analysis. Likelihood ratio test for variance components models has been investigated extensively in the literature (e.g., Chernoff, 1954; Self and Liang, 1987; Shapiro, 1985, 1988). Because the parameters of interests under the null lie on the boundary of the parameter space, the regular chi-square approximation fails due to the violation of the regularity conditions. We reevaluated the asymptotic distribution of the LR statistic under the three model setups and investigated the geometric nature of the null distribution of the LR statistic. We further proposed an analytical method to calculate the tail probability of the LR statistic. The new threshold determination method provides better approximation to the distribution of the LR statistic under the three evaluated models compared to their counterparts.

Although association analysis has been broadly applied to identify genetic variants associated with a disease trait, the role of linkage analysis cannot be undermined (Ott et al., 2015). Our simulation study indicates that SL method underestimates the error rate in the univariate analysis considering both additive and dominance effects. This is primarily due to the fact that SL ignores the correlation between the additive and dominance effects while determining the threshold. The new method involving the correlation calculates the mixing proportions corresponding to the relevant chi-square distribution, thus yields better approximation. When jointly analyzing SBP and DBP under Models II and III, SL approximation is limited by ignoring the covariance of multiple traits which greatly influences the degrees of freedom of the mixture chi-square distribution. This matter leads to higher false positive rates as shown in simulation studies (see Fig. 1) than the methods Amos and New do. The difference between Amos and New in Model II is that the new method estimates the correlation of two traits to adjust the mixing proportions of the chi-square distribution, but Amos applies a fixed mixing proportion, hence yields higher false positive results than the new method does. For Model III, both Amos and New incorporate the covariance of two traits to adjust the degrees of freedom of the mixture chi-square distribution. The discrepancy between the two lies in the calculation of mixing proportions. Again, Amos fixes the mixing components, while our new method estimates the correlation of the two effects to calculate the mixing proportions, hence is more data adaptive. The simulation studies demonstrate the performance of the three methods under different model setup. The real data analysis further confirms the performance and is consistent with the simulation results.

The three motivating models represent the most important models in variance components linkage analysis. It is worth noting that the commonly applied mixture chi-square approximation is still valid under the univariate trait analysis assuming only the additive effect component. When extended to multiple traits analysis, it is limited to adopt. Our extensive simulation studies demonstrate the utility of the proposed threshold calculation method. We expect the new method can reduce false positives or false negatives in determining a linkage signal, hence reduce the cost of unnecessary investigations in a lab condition due to false results.

Acknowledgments

The authors greatly appreciate the reviewer's insightful and constructive comments that greatly improved the presentation of the manuscript. This work was partially supported by National Science Foundation (NSF) grant DMS-1209112. The Genetic Analysis Workshops are supported by NIH grant R01 GM031575. The GAW18 whole genome sequence data were provided by the T2D-GENES Consortium, which is supported by NIH grants U01 DK085524, U01 DK085584, U01 DK085501, U01 DK085526, and U01 DK085545. The other genetic and phenotypic data for GAW18 were provided by the San Antonio Family Heart Study and San Antonio Family Diabetes/Gallbladder study, which are supported by NIH grants P01 HL045222, R01 DK047482, and R01 DK053889. Andrew R. Wood is supported by European Research Council grant SZ-245 50371-GLUCOSEGENES-FP7-IDEAS-ERC.

Appendix

Proof of Proposition 1. Following the arguments in Lehmann and Casella (1998), it is possible to define a sequence of points $\hat{\theta}_\zeta$ in the closed set G about θ_0 to locally maximize $\ell(\theta)$. Then following lemma 1 in Chernoff (1954), the \sqrt{n} -consistency of $\hat{\theta}_\zeta$ can be proved.

Proof of Proposition 2. See the proof of Theorem 3.2 in Shapiro (2000). In brief, two key steps are involved. First, the parameter set is nearly convex at θ_0 . Comparing with convexity, near convexity is a loose condition. In particular, near convexity can be achieved by some smooth constraints in real application. When the fitted function is monotonically nondecreasing and twice continuously differentiable on a given interval, the parameter set is nearly convex at θ_0 under the Mangasarian–Fromovitz constraints. Next, define the Lipschitz continuous functions $F_n(\hat{\theta}_\zeta)$ and $F_n(\tilde{\theta}_\zeta)$ as $-\frac{1}{n}\ell(\theta)$ in terms of $\hat{\theta}_\zeta$ and $\tilde{\theta}_\zeta$, respectively. When $F_n(\hat{\theta}_\zeta)$ and $F_n(\tilde{\theta}_\zeta)$ satisfy conditions 3.8 and 3.9 of Theorem 3.2 in Shapiro (2000), the asymptotic equivalence of $\hat{\theta}_\zeta$ and $\tilde{\theta}_\zeta$ can be achieved by the property of the near convexity (condition A in Shapiro, 2000).

Proof of Theorem 3.1. First note that LR can be approximated as the difference of two quadratic forms with respect to Ω_0 and Ω_1 (see Theorem 1 in Chernoff, 1954). Following the \sqrt{n} -consistency of the optimizer and the property of the approximating cones, the large sample distribution of LR is the same as that of testing $\theta \in C_{\Omega_0}$ against $\theta \in C_{\Omega_1}$. Next, we prove that LR asymptotically follows a mixture chi-square distribution.

Table A.1
List of possible subsets of $\psi_{v|\mathbf{Y}}$ without covariance terms.

Set	Subset	Count
$\psi_{v \mathbf{Y}} = \{\mathbf{Y}; \mathbf{Y} \in \mathbb{R}^q, \mathbf{g}(\mathbf{Y}) \in v\}$	$\psi_{v \mathbf{Y}}^1 = \{\mathbf{Y}; Y_1 > 0, Y_2 > 0, \dots, Y_q > 0, \mathbf{g}(\mathbf{Y}) \in v\}$	$\binom{q}{0}$
	$\psi_{v \mathbf{Y}}^2 = \{\mathbf{Y}; Y_1 \leq 0, Y_2 > 0, \dots, Y_q > 0, \mathbf{g}(\mathbf{Y}) \in v\}$	$\binom{q}{1}$
	$\psi_{v \mathbf{Y}}^3 = \{\mathbf{Y}; Y_1 > 0, Y_2 \leq 0, \dots, Y_q > 0, \mathbf{g}(\mathbf{Y}) \in v\}$	
	...	
	$\psi_{v \mathbf{Y}}^{q+1} = \{\mathbf{Y}; Y_1 > 0, Y_2 > 0, \dots, Y_q \leq 0, \mathbf{g}(\mathbf{Y}) \in v\}$	
	...	
	$\psi_{v \mathbf{Y}}^{2^q} = \{\mathbf{Y}; Y_1 \leq 0, Y_2 \leq 0, \dots, Y_q \leq 0, \mathbf{g}(\mathbf{Y}) \in v\}$	$\binom{q}{q}$

Following Theorem 1 in Chernoff (1954), the asymptotic distribution of LR is equivalent to the following quadratic approximation

$$LR = \inf_{\theta \in C_{\Omega_0}} (\mathbf{Y} - \theta)'M(\mathbf{Y} - \theta) - \inf_{\theta \in C_{\Omega_1}} (\mathbf{Y} - \theta)'M(\mathbf{Y} - \theta) \tag{A.1}$$

where $\mathbf{Y} \sim N(\theta, M^{-1})$. Subtracting θ_0 from \mathbf{Y} and θ , we get an equivalent form of (A.1) as

$$LR = \inf_{\theta \in C_{\Omega_0} - \theta_0} (\mathbf{Y} - \theta)'M(\mathbf{Y} - \theta) - \inf_{\theta \in C_{\Omega_1} - \theta_0} (\mathbf{Y} - \theta)'M(\mathbf{Y} - \theta) \tag{A.2}$$

with $\mathbf{Y} \sim N(\mathbf{0}, M^{-1})$ and M is the Fisher information matrix.

Let $C^\vee = (C_{\Omega_1} - \theta_0) \cap (C_{\Omega_0} - \theta_0)^\perp$, where $(C_{\Omega_0} - \theta_0)^\perp$ is the orthogonal complement of $(C_{\Omega_0} - \theta_0)$. Following the Pythagoras theorem (Shapiro, 1988), the statistic LR in (A.2) can be expressed as

$$LR = \inf_{\theta \in C^\vee} (\mathbf{Y} - \theta)'M(\mathbf{Y} - \theta). \tag{A.3}$$

It can be seen that C^\vee is a closed polyhedral convex cone with dimension q ($q \leq m$) since C^\vee is the intersection of convex cones. Thus a polar cone $C^{\vee 0}$ can be defined as $C^{\vee 0} = \{\gamma \in \mathbb{R}^q; \gamma'\theta \leq 0, \forall \theta \in C^\vee\}$, and $(C^{\vee 0})^0 = C^\vee$ by the basic property of the polar cone (Stoer and Witzgall, 1970, Chapter 2).

Let $\mathbb{F}(C^\vee)$ represent the set of all faces of C^\vee . Following Shapiro (1985), we can select a face $v \in \mathbb{F}(C^\vee)$ corresponding to a polar face $v^0 \in \mathbb{F}(C^{\vee 0})$, and the linear spaces produced by v and v^0 are orthogonal to each other. For one face v (or v^0), we can find a projection T_v (or T_{v^0}) (a symmetric idempotent matrix giving projection onto the spaces by v (or v^0)) and $T_v = I - T_{v^0}$ since they are orthogonal. Then $T_v\mathbf{Y}$ (or $T_{v^0}\mathbf{Y}$) is a projection of a random vector \mathbf{Y} onto C^\vee (or $C^{\vee 0}$). For a given \mathbf{Y} , let $\mathbf{g}(\mathbf{Y}) = (g_1(\mathbf{Y}), g_2(\mathbf{Y}), \dots, g_q(\mathbf{Y}))^T$ be the minimizer to achieve the infimum in (A.3). Define $\psi_{v|\mathbf{Y}} = \{\mathbf{Y} \in \mathbb{R}^q : \mathbf{g}(\mathbf{Y}) \in v\}$ so that $\mathbf{g}(\mathbf{Y}) \in v$ if and only if $T_v\mathbf{Y} \in C^\vee$ and $T_{v^0}\mathbf{Y} \in C^{\vee 0}$. By Shapiro (1985), $\psi_{v|\mathbf{Y}}$ can also be defined by the inequalities as $\psi_{v|\mathbf{Y}} = \{\mathbf{Y} \in \mathbb{R}^q : e'T_v\mathbf{Y} \leq 0, e \in C^{\vee 0}, f'T_{v^0}\mathbf{Y} \leq 0, f \in C^\vee\}$. Thus, $\mathbf{g}(\mathbf{Y}) = T_v\mathbf{Y} \in C^\vee$, for all $\mathbf{Y} \in \psi_{v|\mathbf{Y}}$. Consequently, the likelihood ratio statistic in (A.3) can be expressed as,

$$LR = (\mathbf{Y} - \mathbf{g}(\mathbf{Y}))'M(\mathbf{Y} - \mathbf{g}(\mathbf{Y})) \quad \text{for all } \mathbf{Y} \in \psi_{v|\mathbf{Y}}. \tag{A.4}$$

Note that the set $\psi_{v|\mathbf{Y}}$ is composed of several almost disjoint sets $\psi_{v|\mathbf{Y}}^{i'}$, $i' = 1, \dots, l(q)$. The total number of these disjoint subsets $l(q)$ is calculated by the general form of binomial theorem, i.e., $l(q) = 2^{q-o}$, where q is the number of parameters in C^\vee and o is the number of covariance terms in C^\vee . Moreover, all these subsets are classified into $q - o + 1$ categories. To display these subsets, we start from the simple case that no covariance term is in $\psi_{v|\mathbf{Y}}$ ($o = 0$). The subsets of $\psi_{v|\mathbf{Y}}$ are given in Table A.1.

The case when covariance terms are involved in C^\vee is calculated differently. Take a simple case as an example and let $C^\vee = \{\theta; \theta_1 > 0, \theta_2 > 0, \theta_3 \in \mathbb{R}\}$, where θ_1 is a variance term for trait one, θ_2 is the variance term for trait 2, θ_3 is the covariance between the two traits. Because of the definition of covariance, θ_3 occurs only when $\theta_1 > 0$ and $\theta_2 > 0$, so θ_3 is represented by $\theta_3 I(\theta_1 > 0, \theta_2 > 0)$. Corresponding to this constraint, the set $\psi_{v|\mathbf{Y}}$ is denoted as $\psi_{v|\mathbf{Y}} = \{\mathbf{Y}; Y_i \in \mathbb{R}, i \in q^\vee \setminus o^\vee, Y_j I(Y_{j-2} > 0, Y_{j-1} > 0) \in \mathbb{R}, j \in o^\vee, \mathbf{g}(\mathbf{Y}) \in v\}$, where the set q^\vee is defined as $q^\vee = \{1, 2, \dots, q\}$ and o^\vee is denoted as a subset of q^\vee , and is shown as $o^\vee = \{3, 6, \dots, q\} = \{3k^\vee, k^\vee = 1, 2, \dots, \frac{q}{3}\}$. Thus, the whole subsets under this constrained condition can be shown as in Table A.2.

The general form of the whole number of subsets is $l(q) = 2^{q-o}$, and these subsets consist of $q - o + 1$ groups. Therefore,

$$\psi_{v|\mathbf{Y}} = \bigcup_{i'=1}^{l(q)} \psi_{v|\mathbf{Y}}^{i'}$$

Considering a linear transformation on \mathbf{Y} and θ , a new closed convex cone C^* is defined as $C^* = \{\theta^*; \theta^* = \Lambda^{\frac{1}{2}}P'\theta, \theta \in C^\vee\}$, where $PAP' = M$. Let \mathbf{Z} ($\mathbf{Z} = \Lambda^{\frac{1}{2}}P'\mathbf{Y}$, a linear transformed random vector) which is distributed as a multivariate

Table A.2

The whole possible subsets of $\psi_{v|Y}$ with covariance terms.

Subset	Count
$\psi_{v Y}^1 = \{Y; Y_1 > 0, Y_2 > 0, Y_3 \in \mathbb{R}, Y_4 > 0, Y_5 > 0, Y_6 \in \mathbb{R}, \dots, Y_{q-2} > 0, Y_{q-1} > 0, Y_q \in \mathbb{R}, g(Y) \in v\}$	$\binom{q-0}{0}$
$\psi_{v Y}^2 = \{Y; Y_1 \leq 0, Y_2 > 0, Y_4 > 0, Y_5 > 0, Y_6 \in \mathbb{R}, \dots, Y_{q-2} > 0, Y_{q-1} > 0, Y_q \in \mathbb{R}, g(Y) \in v\}$	$\binom{q-0}{1}$
$\psi_{v Y}^3 = \{Y; Y_1 > 0, Y_2 \leq 0, Y_4 > 0, Y_5 > 0, Y_6 \in \mathbb{R}, \dots, Y_{q-2} > 0, Y_{q-1} > 0, Y_q \in \mathbb{R}, g(Y) \in v\}$	
$\psi_{v Y}^4 = \{Y; Y_1 > 0, Y_2 > 0, Y_3 \in \mathbb{R}, Y_4 \leq 0, Y_5 > 0, \dots, Y_{q-2} > 0, Y_{q-1} > 0, Y_q \in \mathbb{R}, g(Y) \in v\}$	
...	
$\psi_{v Y}^{q-0+1} = \{Y; Y_1 > 0, Y_2 > 0, Y_3 \in \mathbb{R}, Y_4 > 0, Y_5 > 0, Y_6 \in \mathbb{R}, \dots, Y_{q-2} > 0, Y_{q-1} \leq 0, g(Y) \in v\}$	
...	
...	
$\psi_{v Y}^{2^q-0} = \{Y; Y_1 \leq 0, Y_2 \leq 0, Y_4 \leq 0, Y_5 \leq 0, \dots, Y_{q-2} \leq 0, Y_{q-1} \leq 0, g(Y) \in v\}$	$\binom{q-0}{q-0}$

normal distribution with mean zero and identity covariance. In terms of this random vector Z the likelihood ratio LR in (A.3) is evaluated equivalently as:

$$LR = \inf_{\theta^* \in C^*} \|Z - \theta^*\|^2. \tag{A.5}$$

In the same way, C^* is a closed convex cone and C^{*0} is denoted as the polar cone of C^* with $(C^{*0})^0 = C^*$. So there is a face $v^* \in \mathbb{F}(C^*)$ (or $v^{*0} \in \mathbb{F}(C^{*0})$) such that a symmetric idempotent matrix T_{v^*} (or $T_{v^{*0}}$) giving projection onto the space created by v^* (or v^{*0}) is defined. The linear transformation of Y to Z guarantees that there also exists a minimizer denoted by $d(Z)$ for (A.5), in which $d(Z) = T_{v^*}Z \in C^*, \forall Z \in \psi_{v^*|Z}$, where $\psi_{v^*|Z}$ can be defined by a linear transformation from $\psi_{v|Y}$.

Note that the set $\psi_{v^*|Z}$ is also a polyhedral convex cone by its definition and satisfies the conditions of Lemma 3.1 (Shapiro, 1985), and $T_{v^{*0}}$ is an symmetric idempotent matrix corresponding to face v^{*0} , then the likelihood ratio statistic in (A.5) can be written as

$$LR = \|Z - d(Z)\|^2 = \|Z - T_{v^*}Z\|^2 = Z'(I - T_{v^*})Z = Z'T_{v^{*0}}Z, \text{ for all } Z \in \psi_{v^*|Z}. \tag{A.6}$$

It is clear that the minimum value of LR obtained for $Y \in \psi_{v|Y}$ in (A.4) is equivalent to the infimum value of LR obtained for $Z \in \psi_{v^*|Z}$ in (A.6).

Note that the set $\psi_{v^*|Z}$ is also made up of several almost disjoint sets, i.e., $\psi_{v^*|Z} = \bigcup_{i'=1}^{l(q)} \psi_{v^{i'}|Z}$. Condition on $Z \in \psi_{v^{i'}|Z}$, LR follows a chi-square distribution with $\text{rank}(T_{v^{i'0}}) = \text{rank}(I - T_{v^{i'*}})$ degrees of freedom. By Bayes' theorem, the distribution of LR in (A.6) is derived to be a mixture chi-square distribution.

Given a positive number $c > 0$ and a random vector Y , the tail probability of LR under the null hypothesis can be evaluated as,

$$\begin{aligned} \Pr(LR > c) &= \Pr((Y - g(Y))'M(Y - g(Y)) > c, Y \in \psi_{v|Y}) \\ &= \Pr\left((Y - g(Y))'M(Y - g(Y)) > c, Y \in \bigcup_{i'=1}^{l(q)} \psi_{v^{i'}|Y}\right). \end{aligned} \tag{A.7}$$

Applying the union rule for these almost disjoint sets, the representation of (A.7) is changed to

$$\begin{aligned} \Pr(LR > c) &= \Pr\left(\bigcup_{i'=1}^{l(q)} \{(Y - g(Y))'M(Y - g(Y)) > c, Y \in \psi_{v^{i'}|Y}\}\right) \\ &= \sum_{i'=1}^{l(q)} \Pr((Y - g(Y))'M(Y - g(Y)) > c, Y \in \psi_{v^{i'}|Y}) \\ &= \sum_{i'=1}^{l(q)} \Pr(Y \in \psi_{v^{i'}|Y}) \Pr((Y - g(Y))'M(Y - g(Y)) > c | Y \in \psi_{v^{i'}|Y}). \end{aligned} \tag{A.8}$$

According to the resemblance between the result in (A.4) and the comment in (A.6), the representation of the probability can be further expressed as

$$\begin{aligned} \Pr(LR > c) &= \sum_{i'=1}^{l(q)} \Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^{i'}) \Pr(\mathbf{Z}'T_{v^*0}\mathbf{Z} > c | \mathbf{Z} \in \psi_{v^*|\mathbf{Z}}^{i'}) \\ &= \sum_{i'=1}^{l(q)} \Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^{i'}) \Pr(\chi_{\text{rank}(T_{v^*0}|\psi_{v^*|\mathbf{Z}}^{i'})}^2 > c) \end{aligned} \quad (\text{A.9})$$

where $\Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^{i'})$ is the mixing proportion corresponding to the chi-square components with $\sum_{i'=1}^{l(q)} \Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^{i'}) = 1$, and $\text{rank}(T_{v^*0}|\psi_{v^*|\mathbf{Z}}^{i'})$ denotes the degrees of freedom of the corresponding χ^2 distribution. This completes the proof.

The distribution of LR under Model I: The parameters of this model are given as $\boldsymbol{\theta} = \{\theta_1, \theta_2, \theta_3, \theta_4, \theta_5\} = \{\sigma_a^2, \sigma_d^2, \sigma_g^2, \sigma_e^2, \mu\}$, and the approximating cone under the null hypothesis is defined as $C_{\Omega_0} = \{\boldsymbol{\theta}; \theta_1 = 0, \theta_2 = 0, \theta_3 > 0, \theta_4 > 0, \theta_5 \in \mathbb{R}\}$ against $C_{\Omega_1} = \{\boldsymbol{\theta}; \theta_1 > 0, \theta_2 > 0, \theta_3 > 0, \theta_4 > 0, \theta_5 \in \mathbb{R}\}$ under the alternative. The number of parameters to be tested is 2. Thus the set of \mathbf{Y} consists of $l(q) = 4$ almost disjoint sets with 3 categories:

- (i) $\psi_{v|\mathbf{Y}}^1 = \{\mathbf{Y}; Y_1 > 0, Y_2 > 0, g(\mathbf{y}) \in v\}$;
- (ii) $\psi_{v|\mathbf{Y}}^2 = \{\mathbf{Y}; Y_1 > 0, Y_2 \leq 0, g(\mathbf{y}) \in v\}$, $\psi_{v|\mathbf{Y}}^3 = \{\mathbf{Y}; Y_1 \leq 0, Y_2 > 0, g(\mathbf{y}) \in v\}$;
- (iii) $\psi_{v|\mathbf{Y}}^4 = \{\mathbf{Y}; Y_1 \leq 0, Y_2 \leq 0, g(\mathbf{y}) \in v\}$.

When $\mathbf{Y} \in \psi_{v|\mathbf{Y}}^1$, $LR = Z_1^2 + Z_2^2 \sim \chi_2^2$ where $Z_i \stackrel{i.i.d.}{\sim} N(0, 1)$ for $i = 1, 2$, and the corresponding mixture proportion is estimated by $\Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^1)$. For \mathbf{Y} in category (ii) (i.e., $\mathbf{Y} \in \psi_{v|\mathbf{Y}}^2$ or $\mathbf{Y} \in \psi_{v|\mathbf{Y}}^3$), $LR \sim \chi_1^2$ with the corresponding mixing proportion calculated by $\Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^2) + \Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^3)$. For the last category, $LR \sim \chi_0^2$ for $\mathbf{Y} \in \psi_{v|\mathbf{Y}}^4$, and the mixing proportion is $\Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^4)$. The calculation of the mixing proportion follows Plackett (1954) or Kendall (1954). Specifically, $\Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^1) = \frac{\pi - \cos^{-1} \rho_{12}}{2\pi}$, $\Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^2) + \Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^3) = \frac{1}{2}$, and $\Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^4) = \frac{\cos^{-1} \rho_{12}}{2\pi}$, where ρ_{12} is the correlation between the estimators of additive and dominance effects. Therefore, the distribution is approximated as

$$\Pr(LR > c) = \frac{\pi - \cos^{-1} \rho_{12}}{2\pi} P(\chi_2^2 > c) + \frac{1}{2} P(\chi_1^2 > c). \quad (\text{A.10})$$

The distribution of LR under Model II: For Model II in (3), only the random additive effects (a_{k_1}, a_{k_2}) of QTL for two traits are considered. Denote the parameters of additive major gene effects as $\sigma_{a_1}^2, \sigma_{a_2}^2, \sigma_{a_{12}}$ where $\sigma_{a_{12}}$ is the covariance term between the two traits. Similarly, two covariance terms are denoted for the polygene and random residual terms. All parameters in this model are defined as: $\boldsymbol{\theta} = \{\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6, \theta_7, \theta_8, \theta_9, \theta_{10}, \theta_{11}\} = \{\sigma_{a_1}^2, \sigma_{a_2}^2, \sigma_{a_{12}}, \sigma_{g_1}^2, \sigma_{g_2}^2, \sigma_{g_{12}}, \sigma_{e_1}^2, \sigma_{e_2}^2, \sigma_{e_{12}}, \mu_1, \mu_2\}$. The approximating parameter cone under the null hypothesis is defined as $C_{\Omega_0} = \{\boldsymbol{\theta}; \theta_1 = 0, \theta_2 = 0, \theta_3 = 0, \theta_4 > 0, \theta_5 > 0, \theta_6 \in \mathbb{R}, \theta_7 > 0, \theta_8 > 0, \theta_9 \in \mathbb{R}, \theta_{10} \in \mathbb{R}, \theta_{11} \in \mathbb{R}\}$. Similarly, the cone under the alternative hypothesis is denoted as $C_{\Omega_1} = \{\boldsymbol{\theta}; \theta_1 > 0, \theta_2 > 0, \theta_3 \in \mathbb{R}, \theta_4 > 0, \theta_5 > 0, \theta_6 \in \mathbb{R}, \theta_7 > 0, \theta_8 > 0, \theta_9 \in \mathbb{R}, \theta_{10} \in \mathbb{R}, \theta_{11} \in \mathbb{R}\}$. Corresponding to the hypothesis test, the number of tested parameters is 3 ($q = 3$) and there is 1 testing covariance term, then the set $\psi_{v|\mathbf{Y}}$ has 4 almost disjoint subsets and all these subsets are classified into 3 groups.

- (i) $\psi_{v|\mathbf{Y}}^1 = \{\mathbf{Y}; Y_1 > 0, Y_2 > 0, Y_3 \in \mathbb{R}, g(\mathbf{y}) \in v\}$;
- (ii) $\psi_{v|\mathbf{Y}}^2 = \{\mathbf{Y}; Y_1 > 0, Y_2 \leq 0, g(\mathbf{y}) \in v\}$, $\psi_{v|\mathbf{Y}}^3 = \{\mathbf{Y}; Y_1 \leq 0, Y_2 > 0, g(\mathbf{y}) \in v\}$;
- (iii) $\psi_{v|\mathbf{Y}}^4 = \{\mathbf{Y}; Y_1 \leq 0, Y_2 \leq 0, g(\mathbf{y}) \in v\}$.

The estimator of covariance term is only observed in $\psi_{v|\mathbf{Y}}^1$, and it will vanish automatically when $Y_1 \leq 0$ or $Y_2 \leq 0$. The LR statistic can be expressed as

$$LR = \begin{cases} Z_1^2 + Z_2^2 + Z_3^2 \sim \chi_3^2 & \text{with mixing prop. } \Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^1) \\ Z_1^2 \sim \chi_1^2 & \text{with mixing prop. } \Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^2) \\ Z_2^2 \sim \chi_1^2 & \text{with mixing prop. } \Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^3) \\ 0 \sim \chi_0^2 & \text{with mixing prop. } \Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^4) \end{cases}$$

For $\mathbf{Y} \in \psi_{v|\mathbf{Y}}^1$, $LR \sim \chi_3^2$, and the corresponding mixture proportion is calculated as $\Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^1) = \frac{\pi - \cos^{-1} \rho_{12}}{2\pi}$. For $\mathbf{Y} \in \psi_{v|\mathbf{Y}}^2$ or $\mathbf{Y} \in \psi_{v|\mathbf{Y}}^3$, $LR \sim \chi_1^2$ with mixing probability $\Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^2) + \Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^3) = \frac{1}{2}$. For $\mathbf{Y} \in \psi_{v|\mathbf{Y}}^4$, $LR \sim \chi_0^2$, and the mixing probability is evaluated as $\Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^4) = \frac{\cos^{-1} \rho_{12}}{2\pi}$, where ρ_{12} is the correlation of additive effects between two traits. Thus, the tail probability of LR under Model II is given as

$$\Pr(LR > c) = \frac{\pi - \cos^{-1} \rho_{12}}{2\pi} P(\chi_3^2 > c) + \frac{1}{2} P(\chi_1^2 > c). \quad (\text{A.11})$$

The distribution of LR under Model III: In Model III, random additive effects (a_{k_1}, a_{k_2}) and dominance effects (d_{k_1}, d_{k_2}) are considered. The parameters under this model are denoted as: $\theta = \{\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6, \theta_7, \theta_8, \theta_9, \theta_{10}, \theta_{11}, \theta_{12}, \theta_{13}, \theta_{14}\} = \{\sigma_{a_1}^2, \sigma_{a_2}^2, \sigma_{a_{12}}^2, \sigma_{d_1}^2, \sigma_{d_2}^2, \sigma_{d_{12}}^2, \sigma_{g_1}^2, \sigma_{g_2}^2, \sigma_{g_{12}}^2, \sigma_{e_1}^2, \sigma_{e_2}^2, \sigma_{e_{12}}^2, \mu_1, \mu_2\}$. The approximating cone under the null hypothesis is denoted as $C_{\Omega_0} = \{\theta; \theta_1 = 0, \theta_2 = 0, \theta_3 = 0, \theta_4 = 0, \theta_5 = 0, \theta_6 = 0, \theta_7 > 0, \theta_8 > 0, \theta_9 \in \mathbb{R}, \theta_{10} > 0, \theta_{11} > 0, \theta_{12} \in \mathbb{R}, \theta_{13} \in \mathbb{R}, \theta_{14} \in \mathbb{R}\}$, and the cone under the alternative hypothesis is denoted as $C_{\Omega_1} = \{\theta, \theta_1 > 0, \theta_2 > 0, \theta_3 \in \mathbb{R}, \theta_4 > 0, \theta_5 > 0, \theta_6 \in \mathbb{R}, \theta_7 > 0, \theta_8 > 0, \theta_9 \in \mathbb{R}, \theta_{10} > 0, \theta_{11} > 0, \theta_{12} \in \mathbb{R}, \theta_{13} \in \mathbb{R}, \theta_{14} \in \mathbb{R}\}$. The number of testing parameters in Model III is 6 ($q = 6$) and the number of testing covariance terms is 2 ($o = 2$). Then the set $\psi_{v|Y}$ can be partitioned into 16 almost disjoint subsets that comprise 5 categories, i.e.,

- (i) $\psi_{v|Y}^1 = \{Y; Y_1 > 0, Y_2 > 0, Y_3 \in \mathbb{R}, Y_4 > 0, Y_5 > 0, Y_6 \in \mathbb{R}, g(\mathbf{y}) \in v\}$;
- (ii) $\psi_{v|Y}^2 = \{Y; Y_1 > 0, Y_2 \leq 0, Y_4 > 0, Y_5 > 0, Y_6 \in \mathbb{R}, g(\mathbf{y}) \in v\}$,
 $\psi_{v|Y}^3 = \{Y; Y_1 \leq 0, Y_2 > 0, Y_4 > 0, Y_5 > 0, Y_6 \in \mathbb{R}, g(\mathbf{y}) \in v\}$;
- (iii) $\psi_{v|Y}^4 = \{Y; Y_1 > 0, Y_2 > 0, Y_3 \in \mathbb{R}, Y_4 > 0, Y_5 \leq 0, g(\mathbf{y}) \in v\}$,
 $\psi_{v|Y}^5 = \{Y; Y_1 > 0, Y_2 > 0, Y_3 \in \mathbb{R}, Y_4 \leq 0, Y_5 > 0, g(\mathbf{y}) \in v\}$;
- (iv) $\psi_{v|Y}^6 = \{Y; Y_1 \leq 0, Y_2 \leq 0, Y_4 > 0, Y_5 > 0, Y_6 \in \mathbb{R}, g(\mathbf{y}) \in v\}$;
- (v) $\psi_{v|Y}^7 = \{Y; Y_1 > 0, Y_2 > 0, Y_3 \in \mathbb{R}, Y_4 \leq 0, Y_5 \leq 0, g(\mathbf{y}) \in v\}$;
- (vi) $\psi_{v|Y}^8 = \{Y; Y_1 \leq 0, Y_2 > 0, Y_4 \leq 0, Y_5 > 0, g(\mathbf{y}) \in v\}$;
- (vii) $\psi_{v|Y}^9 = \{Y; Y_1 \leq 0, Y_2 > 0, Y_4 > 0, Y_5 \leq 0, g(\mathbf{y}) \in v\}$;
- (viii) $\psi_{v|Y}^{10} = \{Y; Y_1 > 0, Y_2 \leq 0, Y_4 \leq 0, Y_5 > 0, g(\mathbf{y}) \in v\}$;
- (ix) $\psi_{v|Y}^{11} = \{Y; Y_1 > 0, Y_2 \leq 0, Y_4 > 0, Y_5 \leq 0, g(\mathbf{y}) \in v\}$;
- (x) $\psi_{v|Y}^{12} = \{Y; Y_1 \leq 0, Y_2 \leq 0, Y_4 \leq 0, Y_5 > 0, g(\mathbf{y}) \in v\}$;
- (xi) $\psi_{v|Y}^{13} = \{Y; Y_1 \leq 0, Y_2 \leq 0, Y_4 > 0, Y_5 \leq 0, g(\mathbf{y}) \in v\}$;
- (xii) $\psi_{v|Y}^{14} = \{Y; Y_1 \leq 0, Y_2 > 0, Y_4 \leq 0, Y_5 \leq 0, g(\mathbf{y}) \in v\}$;
- (xiii) $\psi_{v|Y}^{15} = \{Y; Y_1 > 0, Y_2 \leq 0, Y_4 \leq 0, Y_5 \leq 0, g(\mathbf{y}) \in v\}$;
- (xiv) $\psi_{v|Y}^{16} = \{Y; Y_1 \leq 0, Y_2 \leq 0, Y_4 \leq 0, Y_5 \leq 0, g(\mathbf{y}) \in v\}$;

Based on these 16 almost disjoint subsets, the limiting distribution of LR can be expressed as

$$LR = \begin{cases} Z_1^2 + Z_2^2 + Z_3^2 + Z_4^2 + Z_5^2 + Z_6^2 \sim \chi_6^2 & \text{with mixing prop. } \Pr(\mathbf{Y} \in \psi_{v|Y}^1) \\ Z_1^2 + Z_4^2 + Z_5^2 + Z_6^2 \sim \chi_4^2 & \text{with mixing prop. } \Pr(\mathbf{Y} \in \psi_{v|Y}^2) \\ Z_2^2 + Z_4^2 + Z_5^2 + Z_6^2 \sim \chi_4^2 & \text{with mixing prop. } \Pr(\mathbf{Y} \in \psi_{v|Y}^3) \\ Z_1^2 + Z_2^2 + Z_3^2 + Z_4^2 \sim \chi_4^2 & \text{with mixing prop. } \Pr(\mathbf{Y} \in \psi_{v|Y}^4) \\ Z_1^2 + Z_2^2 + Z_3^2 + Z_5^2 \sim \chi_4^2 & \text{with mixing prop. } \Pr(\mathbf{Y} \in \psi_{v|Y}^5) \\ Z_4^2 + Z_5^2 + Z_6^2 \sim \chi_3^2 & \text{with mixing prop. } \Pr(\mathbf{Y} \in \psi_{v|Y}^6) \\ Z_1^2 + Z_2^2 + Z_3^2 \sim \chi_3^2 & \text{with mixing prop. } \Pr(\mathbf{Y} \in \psi_{v|Y}^7) \\ Z_2^2 + Z_5^2 \sim \chi_2^2 & \text{with mixing prop. } \Pr(\mathbf{Y} \in \psi_{v|Y}^8) \\ Z_2^2 + Z_4^2 \sim \chi_2^2 & \text{with mixing prop. } \Pr(\mathbf{Y} \in \psi_{v|Y}^9) \\ Z_1^2 + Z_5^2 \sim \chi_2^2 & \text{with mixing prop. } \Pr(\mathbf{Y} \in \psi_{v|Y}^{10}) \\ Z_1^2 + Z_4^2 \sim \chi_2^2 & \text{with mixing prop. } \Pr(\mathbf{Y} \in \psi_{v|Y}^{11}) \\ Z_5^2 \sim \chi_1^2 & \text{with mixing prop. } \Pr(\mathbf{Y} \in \psi_{v|Y}^{12}) \\ Z_4^2 \sim \chi_1^2 & \text{with mixing prop. } \Pr(\mathbf{Y} \in \psi_{v|Y}^{13}) \\ Z_2^2 \sim \chi_1^2 & \text{with mixing prop. } \Pr(\mathbf{Y} \in \psi_{v|Y}^{14}) \\ Z_1^2 \sim \chi_1^2 & \text{with mixing prop. } \Pr(\mathbf{Y} \in \psi_{v|Y}^{15}) \\ 0 \sim \chi_0^2 & \text{with mixing prop. } \Pr(\mathbf{Y} \in \psi_{v|Y}^{16}) \end{cases}.$$

Therefore, the tail probability of LR under Model III is given as

$$\begin{aligned} \Pr(LR > c) &= \Pr(\mathbf{Y} \in \psi_{v|Y}^1)P(\chi_6^2 > c) + \sum_{i=2}^5 \Pr(\mathbf{Y} \in \psi_{v|Y}^i)P(\chi_4^2 > c) \\ &+ \sum_{i=6}^7 \Pr(\mathbf{Y} \in \psi_{v|Y}^i)P(\chi_3^2 > c) + \sum_{i=8}^{11} \Pr(\mathbf{Y} \in \psi_{v|Y}^i)P(\chi_2^2 > c) \\ &+ \sum_{i=12}^{15} \Pr(\mathbf{Y} \in \psi_{v|Y}^i)P(\chi_1^2 > c). \end{aligned} \tag{A.12}$$

The mixture proportions can be calculated following the results of Kudô (1963, p.415) and Shapiro (1985, p.141). Specifically,

(1) The mixing proportion corresponding to the chi-square distribution with 4 df is $w_1 = \sum_{i'=2}^5 \Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^{i'})$, and this probability can be estimated as

$$\begin{aligned} w_1 &= \sum_{i'=2}^5 \Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^{i'}) = \sum_{\substack{a=1,b=2,c=4,d=5,e=6 \\ a=2,b=1,c=4,d=5,e=6 \\ a=4,b=5,c=1,d=2,e=3 \\ a=5,b=4,c=1,d=2,e=3}} \Pr(Y_a < 0, Y_b > 0, Y_c > 0, Y_d > 0, Y_e \in \mathbb{R}) \\ &= \sum_{\substack{a=1,b=2,c=4,d=5 \\ a=2,b=1,c=4,d=5 \\ a=4,b=5,c=1,d=2 \\ a=5,b=4,c=1,d=2}} \Pr(Y_a < 0, Y_b > 0, Y_c > 0, Y_d > 0) \\ &= \frac{1}{8\pi} \left(8\pi - \sum_{\substack{a>b;a\neq c,b\neq c \\ a,b,c=1,2,4,5}} \cos^{-1} \rho_{ab|c} \right) \end{aligned}$$

where $\rho_{ab|c}$ is calculated as $\rho_{ab|c} = \frac{\rho_{ab} - \rho_{ac} \rho_{bc}}{\sqrt{\rho_{ac}^2} \sqrt{\rho_{bc}^2}}$.

(2) With respect to the chi-square distribution with 3 df, the mixing proportion is given as $w_2^1 = \sum_{i'=6}^7 \Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^{i'})$, and

$$\begin{aligned} w_2^1 &= \sum_{i'=6}^7 \Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^{i'}) = \Pr(Y_1 < 0, Y_2 < 0, Y_4 > 0, Y_5 > 0, Y_6 \in \mathbb{R}) \\ &\quad + \Pr(Y_1 > 0, Y_2 > 0, Y_3 \in \mathbb{R}, Y_4 < 0, Y_5 < 0) \\ &= \Pr(Y_1 < 0, Y_2 < 0, Y_4 > 0, Y_5 > 0) + \Pr(Y_1 > 0, Y_2 > 0, Y_4 < 0, Y_5 < 0) \\ &= \frac{1}{4\pi^2} [\cos^{-1} \rho_{12} (\pi - \cos^{-1} \rho_{45|12}) + \cos^{-1} \rho_{45} (\pi - \cos^{-1} \rho_{12|45})] \end{aligned}$$

where $\rho_{cd|ab}$ is estimated by $\rho_{cd|ab} = \frac{\rho_{ab} - \frac{\rho_{ac} \rho_{bc} + \rho_{ad} \rho_{bd} - \rho_{ac} \rho_{bd} \rho_{cd} - \rho_{ad} \rho_{bc} \rho_{cd}}{1 - \rho_{cd}^2}}{\sqrt{\frac{1 - \rho_{cd}^2 - \rho_{ac}^2 - \rho_{ad}^2 + 2\rho_{ac}^2 \rho_{ad}^2 \rho_{cd}^2}{1 - \rho_{cd}^2}} \sqrt{\frac{1 - \rho_{cd}^2 - \rho_{bc}^2 - \rho_{bd}^2 + 2\rho_{bc}^2 \rho_{bd}^2 \rho_{cd}^2}{1 - \rho_{cd}^2}}}$.

(3) $w_2^2 = \sum_{i'=8}^{11} \Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^{i'})$ is the mixing probability for the chi-square component with 2 df. The mixing proportion is evaluated as

$$\begin{aligned} w_2^2 &= \sum_{i'=8}^{11} \Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^{i'}) = \sum_{a,b=1,2;c,d=4,5;a\neq b;c\neq d} \Pr(Y_a < 0, Y_b > 0, Y_c < 0, Y_d > 0) \\ &= \frac{1}{4\pi^2} [\cos^{-1} \rho_{14} (\pi - \cos^{-1} \rho_{25|14}) + \cos^{-1} \rho_{15} (\pi - \cos^{-1} \rho_{24|15}) \\ &\quad + \cos^{-1} \rho_{24} (\pi - \cos^{-1} \rho_{15|24}) + \cos^{-1} \rho_{25} (\pi - \cos^{-1} \rho_{14|25})] \end{aligned}$$

where $\rho_{cd|ab}$ is defined as in (2).

(4) Following the comment by Shapiro (1985), the mixing proportions are assigned equally on the even and odd places. Thus the mixing probability $w_3 = \sum_{i'=12}^{15} \Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^{i'})$ for the chi-square component with 1 df is calculated as

$$\begin{aligned} w_3 &= \sum_{i'=12}^{15} \Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^{i'}) = \sum_{\substack{a=1,b=2,c=4,d=5 \\ a=2,b=1,c=4,d=5 \\ a=4,b=5,c=1,d=2 \\ a=5,b=4,c=1,d=2}} \Pr(Y_a > 0, Y_b < 0, Y_c < 0, Y_d < 0) \\ &= \frac{1}{2} - \sum_{i'=2}^5 \Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^{i'}) \\ &= \frac{1}{8\pi} \left(\sum_{\substack{a>b;a\neq c,b\neq c \\ a,b,c=1,2,4,5}} \cos^{-1} \rho_{ab|c} - 4\pi \right) \end{aligned}$$

where there is a resemblance about the calculation of $\rho_{ab|c}$ between $w_3 = \sum_{i'=12}^{15} \Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^{i'})$ and $w_1 = \sum_{i'=2}^5 \Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^{i'})$.

(5) Due to the equivalence of weights on odd and even places (Kendall, 1954), we have $w_0 + w_4 = \Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^1) + \Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^{16}) = \frac{1}{2} - \sum_{i'=6}^{11} \Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^{i'}) = \frac{1}{2} - w_2^1 - w_2^2$. Suppose the two mixing proportions $w_0 = \Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^1)$ for the

chi-square component with 6 df and $w_4 = \Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^{16})$ for the chi-square component with 0 df equally share the probability $\frac{1}{2} - \sum_{i'=6}^{11} \Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^{i'})$. Then, the mixing proportion is approximated as

$$\begin{aligned} w_0 = w_4 = \Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^1) = \Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^{16}) &= \frac{1}{4} - \frac{1}{2} \sum_{i'=6}^{11} \Pr(\mathbf{Y} \in \psi_{v|\mathbf{Y}}^{i'}) \\ &= \frac{1}{4} - \frac{1}{8\pi^2} [\cos^{-1} \rho_{12}(\pi - \cos^{-1} \rho_{45|12}) \\ &\quad + \cos^{-1} \rho_{45}(\pi - \cos^{-1} \rho_{12|45}) \\ &\quad + \cos^{-1} \rho_{14}(\pi - \cos^{-1} \rho_{25|14}) \\ &\quad + \cos^{-1} \rho_{15}(\pi - \cos^{-1} \rho_{24|15}) \\ &\quad + \cos^{-1} \rho_{24}(\pi - \cos^{-1} \rho_{15|24}) \\ &\quad + \cos^{-1} \rho_{25}(\pi - \cos^{-1} \rho_{14|25})]. \end{aligned}$$

Back to our original motivating models, the limiting distribution of the LR is given as $\frac{\pi - \cos^{-1} \rho_{12}}{2\pi} \chi_2^2 : \frac{1}{2} \chi_1^2 : \frac{\cos^{-1} \rho_{12}}{2\pi} \chi_0^2$ for Model I, $\frac{\pi - \cos^{-1} \rho_{12}}{2\pi} \chi_3^2 : \frac{1}{2} \chi_1^2 : \frac{\cos^{-1} \rho_{12}}{2\pi} \chi_0^2$ for Model II and $w_0 \chi_6^2 : w_1 \chi_4^2 : w_2 \chi_3^2 : w_2 \chi_2^2 : w_3 \chi_1^2 : w_4 \chi_0^2$ for Model III.

References

- 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. *Am. J. Hum. Genet.* 65, 531–544.
- Almasy, L., Blangero, J., 2010. Variance component methods for analysis of complex phenotypes. *Cold Spring Harb Protoc.* 2010 (5), pdb.top77. <http://dx.doi.org/10.1101/pdb.top77>.
- Almasy, L.A., Dyer, T.D., Peralta, J.M., Jun, G., Wood, A.R., Fuchsberger, C., Almeida, M.A., Kent, J.W., Fowler, S., Blackwell, T.W., Puppala, S., Kumar, S., Curran, J.E., Lehman, D., Abecasis, G., Duggirala, R., Blangero, J., The T2D-GENES Consortium., 2014. Data for genetic analysis workshop 18: human whole genome sequence, blood pressure, and simulated phenotypes in extended pedigrees, *BMC Proc.*, Vol. 8, no. (Suppl 2), p. S2.
- Amos, C.I., 1994. Robust variance-components approach for assessing genetic linkage in pedigrees. *Am. J. Hum. Genet.* 54, 535–543.
- Amos, C., Andrade, M., 2001. Genetic linkage methods for quantitative traits. *Stat. Methods Med. Res.* 10, 3–25.
- Amos, C.I., de Andrade, M., Zhu, K.D., 2001. Comparison of multivariate tests for genetic linkage. *Hum. Hered.* 51, 133–144.
- Chernoff, H., 1954. On the distribution of the likelihood ratio. *Ann. Math. Stat.* 25, 573–578.
- Evans, D.M., 2002. The power of multivariate quantitative-trait loci linkage analysis is influenced by the correlation between the variables. *Am. J. Hum. Genet.* 70, 1599–1602.
- Geyer, C.J., 1994. On the asymptotics of constrained M-estimation. *Ann. Statist.* 22, 1993–2010.
- Goldgar, D.E., 1990. Multipoint analysis of human quantitative genetic variation. *Am. J. Hum. Genet.* 47, 957–967.
- Hanson, R.L., Kobes, S., Lindsay, R.S., Knowler, W.C., 2001. Assessment of parent-of-origin effects in linkage analysis of quantitative traits. *Am. J. Hum. Genet.* 68 (4), 951–962.
- Jiang, C., Zeng, Z.-B., 1995. Multiple trait analysis of genetic mapping for quantitative trait loci. *Genetics* 140, 1111–1127.
- Kendall, M.G., 1954. Note on the bias in the estimation of autocorrelation. *Biometrika* 41, 403–404.
- Kudô, A., 1963. A multivariate analogue of the one-sided test. *Biometrika* 50, 403–418.
- Kudô, A., Choi, J.R., 1975. A generalized multivariate analogue of the one sided test. *Mem. Fac. Sci. Kyushu Univ.* 29, 303–328.
- Lehmann, E.L., Casella, G., 1998. *Theory of Point Estimation*. Springer, New York.
- Lynch, M., Walsh, B., 1998. *Genetics and Analysis of Quantitative Traits*. Sinauer Associates, Inc.
- Morris, N.J., Elston, R., Stein, C.M., 2009. Calculating asymptotic significance levels of the constrained likelihood ratio test with application to multivariate genetic linkage analysis. *Stat. Appl. Genet. Mol. Biol.* 8, 39.
- Nagy, . et al., 2014. Estimation of additive and dominance variance for litter size components in rabbits. *Czech. J. Anim. Sci.* 59, 182–189.
- Ott, J., Wang, J., Leal, S.M., 2015. Genetic linkage analysis in the age of whole-genome sequencing. *Nat. Rev. Genet.* 16, 275–284.
- Plackett, R.L., 1954. A reduction formula for normal multivariate integrals. *Biometrika* 41, 351–360.
- Self, S.G., Liang, K.Y., 1987. Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *J. Amer. Statist. Assoc.* 82, 605–610.
- Shapiro, A., 1985. Asymptotic distribution of test statistics in the analysis of moment structures under inequality constraints. *Biometrika* 72, 133–144.
- Shapiro, A., 1988. Towards a unified theory of inequality constrained testing in multivariate analysis. *Internat. Statist. Rev.* 56, 49–62.
- Shapiro, A., 2000. On the asymptotics of constrained local M-estimators. *Ann. Statist.* 3 (2000), 948–960.
- Stoer, J., Witzgall, C., 1970. *Convexity and Optimization in Finite Dimensions, I*. Springer-Verlag, Berlin.
- Wang, T., Zeng, Z.B., 2009. Contribution of genetic effects to genetic variance components with epistasis and linkage disequilibrium. *BMC Genet.* 10, 52.
- Xu, S., Atchley, W.R., 1995. A random model approach to interval mapping of quantitative trait loci. *Genetics* 141 (3), 1189–1197.