Supporting Documentation for

"Some Statistical Properties of Efficiency Robust Tests with Applications to Genetic Association Studies"

By Gang Zheng, Qizhai Li, and Ao Yuan

We first show, in appendices S1-S3, that the likelihood functions of some common genetic association studies can be written as $L_n(\lambda_1, \lambda_2, \alpha)$ as described in the paper, where λ_1 and λ_2 are the parameters of interest and α is a vector of nuisance parameters. The study designs that we consider here include case-control data, population-based quantitative traits, and case-parents trio data. Then, in appendices S4-S5, we provide technical details of the derivations and the proofs of the main results. Finally, in appendix S6, Table S2 reports conditional probabilities of trio data under the null and alternative hypotheses.

Appendix S1. Case-control data

Retrospective likelihood. Case-control data are often collected retrospectively given the disease status. This design is cost-effective, especially for rare diseases. It is also one of the most common study designs in genetic association studies. Analyzing case-control data retrospectively involve a large number of nuisance parameters.

Let G be the genotype of a di-allelic marker with alleles A and B. Without loss of generality, assume B is the risk allele when the marker is associated with the disease of interest (i.e., the alternative hypothesis H_1). Hence, the risk of getting the disease is higher when the number of allele B increases in the genotype. Denote the three genotypes as $\{G_0, G_1, G_2\}$, with the subscript 0, 1, 2 indicating the number of allele B in the genotype, and $D = 1$ for a case and $D = 0$ for a control. The genotype G_i is coded as g_i, $i = 0, 1, 2$, in the analysis, given by $g_0 = (0, 0)^T$, $g_1 = (1, 0)^T$

and $g_2 = (0,1)^T$. Denote $X \in \{X_1, \ldots, X_J\}$ as a vector of discrete covariates with distinctive values. If a component of X is continuous, we assume its value is obtained from a discrete variable with the support equal to the unique observed values (Gustafson et al., 2002). Let the data for an individual be (g, D, X), where $g \in \{g_0, g_1, g_2\}$ is the coding for the genotype $G \in \{G_0, G_1, G_2\}$, $D \in \{0,1\}$ is the disease indicator, and $X \in \{X_1, \ldots, X_J\}$ is the vector of covariates.

For a retrospective likelihood, denote

$$p_{idX_j} = \mathrm{pr}(g = g_i, X = X_j | D = d), \quad i = 0,1,2;\ d = 0,1;\ j = 1, \ldots, J.$$

Denote an indicator function for the jth covariate X_j as $I_j(X) = 1$ if $X = X_j$ and 0 otherwise, and row vectors $\delta_{0j}(X) = (I_j(X), 0, 0)$, $\delta_{1j}(X) = (0, I_j(X), 0)$ and $\delta_{2j}(X) = (0, 0, I_j(X))$ for $j = 1, \ldots, J$. Using a common parameterization for the case-control data (Seaman and Richardson, 2004), denote

$$p_{i1X_j} = \frac{p_{i0X_j} \exp\left(\alpha_0^T X_j + \beta^T g_i\right)}{\sum_{i=0}^{2} \sum_{j=0}^{J} p_{i0X_j} \exp\left(\alpha_0^T X_j + \beta^T g_i\right)} = \frac{\exp\left(\alpha^T \widetilde{X}_j + \beta^T g_i\right)}{\sum_{i=0}^{2} \sum_{j=0}^{J} \exp\left(\alpha^T \widetilde{X}_j + \beta^T g_i\right)},$$

where $\widetilde{X}_j = (\delta_{i1}(X_j), \ldots, \delta_{iJ}(X_j), X_j^T)^T$, $\beta = (\beta_1, \beta_2)^T$, and

$$\alpha^T = (\log p_{00X_1}, \log p_{10X_1}, \log p_{20X_1}, \ldots, \log p_{00X_J}, \log p_{10X_J}, \log p_{20X_J}, \alpha_0^T).$$

Given the covariate $X = x$, $\lambda_i = \exp(\beta_i) = (p_{i1x} p_{00x})/(p_{i0x} p_{01x})$ is the odds ratio for G_i vs. G_0 $(i = 1,2)$. Under H_0, $\lambda_1 = \lambda_2 = 1$ and α is a vector of nuisance parameters estimable under H_0. Under H_1, due to the dose-effect of the genotype, i.e., an individual with a genotype with more allele B has greater probability of developing the disease, we have $\lambda_2 \geq \lambda_1 \geq 1$ with at least one strict inequality.

Let r_{ij} (and s_{ij}) be the count of cases (and controls) with genotype G_i and covariate X_j for $i = 0,1,2$ and $j = 1, \ldots, J$. Then $\{r_{ij} : i = 0,1,2; j = 1, \ldots, J\}$ and $\{s_{ij} : i = 0,1,2; j = 1, \ldots, J\}$

follow two independent multinomial distributions. The retrospective likelihood function can be written as

$$L_n(\lambda_1, \lambda_2, \alpha) = \prod_{i=0}^{2} \prod_{j=1}^{J} \left(p_{i1X_j}^{r_{ij}} p_{i0X_j}^{s_{ij}} \right), \tag{1}$$

where $n = \sum_{i,j}(r_{ij} + s_{ij})$.

For a special case without covariates, let the genotype count for G_i in r cases and s controls be r_i and s_i, respectively, where $i = 0, 1, 2$ and $n = r + s$. Denote $n_i = r_i + s_i$ $(i = 0, 1, 2)$. Then the likelihood function (1) can be written as

$$L_n(\lambda_1, \lambda_2, \alpha) = \frac{\lambda_1^{r_1} \lambda_2^{r_2} \prod_{i=0}^{2} p_{i0}^{n_i}}{(p_{00} + \lambda_1 p_{10} + \lambda_2 p_{20})^r},$$

where $\lambda_2 \geq \lambda_1 \geq 1$ and $\alpha = (p_{00}, p_{10}, p_{20})^T$. This simplified likelihood function was used to construct a likelihood ratio test for case-control genetic association studies (Wang and Sheffield, 2005).

Prospective likelihood. Using the same notation, for an individual with data (g, D, X), denote the logistic regression model as

$$p_{1iX_j} = \mathrm{pr}(D = 1 | g = g_i, X = X_j) = \frac{\exp(\alpha^T \widetilde{X}_j + \beta^T g_i)}{1 + \exp(\alpha^T \widetilde{X}_j + \beta^T g_i)},$$

where $\widetilde{X}_j = (1, X_j^T)^T$ and $p_{0iX_j} = 1 - p_{1iX_j}$. Denote $\lambda_i = \exp(\beta_i)$ for $i = 1, 2$. Then, given $X = x$, $\lambda_i = (p_{1ix} p_{00x})/(p_{0ix} p_{10x})$ is the odds ratio of G_i vs. G_0 $(i = 1, 2)$. As before, due to the dose-effect of the genotype, $\lambda_2 \geq \lambda_1 \geq 1$. Under H_0, $\lambda_1 = \lambda_2 = 1$ and α contains nuisance parameters estimable under H_0.

Let r_{ij} and s_{ij} be defined as before. Then using the binomial distribution, the likelihood function can be written as

$$L_n(\lambda_1, \lambda_2, \alpha^T) = \prod_{i=0}^{2} \prod_{j=1}^{J} \left(p_{1iX_j}^{r_{ij}} p_{0iX_j}^{s_{ij}} \right). \tag{2}$$

The likelihood in (2) contains less nuisance parameters than the one in (1). It has been shown that inferences of (λ_1, λ_2) using (1) and (2) are equivalent (Prentice and Pyke, 1979; Qin and Zhang, 1997; Seaman and Richardson, 2004).

For a special case without covariate X, (2) can be written as

$$L_n(\lambda_1, \lambda_2, \alpha) = \frac{\lambda_1^{r_1} \lambda_2^{r_2} \exp(r\alpha)}{\{1 + \exp(\alpha)\}^{n_0} \{1 + \lambda_1 \exp(\alpha)\}^{n_1} \{1 + \lambda_2 \exp(\alpha)\}^{n_2}},$$

where α is a scaler parameter. Many common test statistics for case-control genetic association are based on the above simplified likelihood function, including Pearson's test, the trend tests, and robust tests (Sasieni, 1997; Freidlin et al., 2002; Zheng et al., 2012).

Appendix S2. Quantitative trait

For an individual, the quantitative trait is denoted as Y. Genotype and covariates are still denoted as before. Conditional on g and X, the following linear model is often used:

$$Y = \mu + \beta^T g + \alpha_0^T X + \epsilon = \alpha_1^T \widetilde{X} + \beta^T g + \epsilon,$$

where $\widetilde{X}^T = (1, X^T)$, $\alpha_1^T = (\mu, \alpha_0^T)$, $\beta = (\lambda_1, \lambda_2)^T$ is the parameter of interest, and ϵ usually follows $N(0, \sigma^2)$. Denote $\alpha^T = (\mu, \sigma, \alpha_0^T)$, which is the nuisance parameter. The constraints on (λ_1, λ_2) still hold if the increasing number of allele B in the genotype corresponds to the higher trait value under H_1. Then the likelihood function can be written as

$$L_n(\lambda_1, \lambda_2, \alpha) = \prod_{j=1}^{n} \phi \left(\frac{y_j - \alpha_1^T \widetilde{X}_j + \beta^T g_j}{\sigma} \right)$$

where ϕ is the density of $N(0, 1)$, y_j is the trait value of the jth individual with covariate X_j and genotype coding g_j.

Table S1: Conditional probability $\mathrm{pr}(G|\mathrm{MT}, D)$ for observing the genotype of an offspring who has the disease given the mating type (MT). Genotype is denoted as $(G_0, G_1, G_2) = (AA, AB, BB)$.

Parental		Offspring			
				Conditional Probability	
MT	count	G	count	(λ_1, λ_2)	(λ, θ)
$AA \times AB$	n_1	AA	n_{10}	$p_{10} = \frac{1}{1+\lambda_1}$	$p_{10} = \frac{1}{2-\theta+\theta\lambda}$
		AB	n_{11}	$p_{11} = \frac{\lambda_1}{1+\lambda_1}$	$p_{11} = \frac{1-\theta+\theta\lambda}{2-\theta+\theta\lambda}$
$AB \times AB$	n_2	AA	n_{20}	$p_{20} = \frac{1}{1+2\lambda_1+\lambda_2}$	$p_{20} = \frac{1}{3-2\theta+(2\theta+1)\lambda}$
		AB	n_{21}	$p_{21} = \frac{2\lambda_1}{1+2\lambda_1+\lambda_2}$	$p_{21} = \frac{2(1-\theta+\theta\lambda)}{3-2\theta+(2\theta+1)\lambda}$
		BB	n_{22}	$p_{22} = \frac{\lambda_2}{1+2\lambda_1+\lambda_2}$	$p_{22} = \frac{\lambda}{3-2\theta+(2\theta+1)\lambda}$
$AB \times BB$	n_3	AB	n_{31}	$p_{31} = \frac{\lambda_1}{\lambda_1+\lambda_2}$	$p_{31} = \frac{1-\theta+\theta\lambda}{1-\theta+(1+\theta)\lambda}$
		BB	n_{32}	$p_{32} = \frac{\lambda_2}{\lambda_1+\lambda_2}$	$p_{32} = \frac{\lambda}{1-\theta+(1+\theta)\lambda}$

Appendix S3. Case-parents trio data

In the above, we have discussed the likelihood for population-based genetic association studies with a binary or quantitative trait. Here we consider a simple family-based genetic association study using case-parents trio data. An individual with a disease ($D = 1$) is ascertained whose genotype and whose parental genotypes are all obtained. Inference can be based on the conditional probability of genotype $G \in \{G_0, G_1, G_2\}$ of an offspring given the parental mating type (MT) and $D = 1$ (Schaid and Sommer, 1993).

Among six MTs, only three of them, $G_0 \times G_1$, $G_1 \times G_1$ and $G_1 \times G_2$, contribute to the conditional likelihood. Denote the sample sizes of the three informative MTs as n_1, n_2 and n_3, respectively. Assume that (n_1, n_2, n_3) are fixed by the study design. The genotype counts for offspring with the three MTs are denoted as (n_{10}, n_{11}), (n_{20}, n_{21}, n_{22}), and (n_{31}, n_{32}), respectively, where the second subscript is the number of allele B in the offspring genotype (Table S1). Denote $\lambda_i = P(\mathrm{case}|G_i)/P(\mathrm{case}|G_0)$ as the genotype relative risk for $i = 1, 2$. By the dose-effect in the genotype, we have $\lambda_2 \geq \lambda_1 \geq 1$. Under H_0, $\lambda_1 = \lambda_2 = 1$. There is no nuisance parameter α here.

Table S1 can be found in Schaid and Sommer (1993). The conditional likelihood is obtained by the independence among the three MTs. For each MT, the data follow a binomial distribution ($AB \times AB$) or multinomial distribution ($AA \times AB$ and $AB \times BB$). Then the likelihood function can be written as

$$\begin{aligned} L_n(\lambda_1, \lambda_2) &= p_{10}^{n_{10}} p_{11}^{n_{11}} p_{20}^{n_{20}} p_{21}^{n_{21}} p_{22}^{n_{22}} p_{31}^{n_{31}} p_{32}^{n_{32}} \\ &= \frac{\lambda_1^{n_{11}+n_{21}+n_{31}} \lambda_2^{n_{22}+n_{32}}}{(1+\lambda_1)^{n_1}(1+2\lambda_1+\lambda_2)^{n_2}(\lambda_1+\lambda_2)^{n_3}}, \end{aligned}$$

where $n = n_1 + n_2 + n_3$.

S4. Derivations of $\sigma(\alpha_0, \theta_i, \theta_j)$ and ρ_{θ_i,θ_j}

The $(m+1)\times(m+1)$ Fisher information matrix of $n = 1$ under H_0 is given by

$$\begin{aligned} I(\alpha^*, \theta) &= \begin{pmatrix} I_{\lambda\lambda} & I_{\lambda\alpha} \\ I_{\alpha\lambda} & I_{\alpha\alpha} \end{pmatrix} \\ &= -\begin{pmatrix} L_{22}(\alpha^*) + 2\theta L_{12}(\alpha^*) + \theta^2 L_{11}(\alpha^*) & L_{23}(\alpha^*) + \theta L_{13}(\alpha^*) \\ L_{32}(\alpha^*) + \theta L_{31}(\alpha^*) & L_{33}(\alpha^*) \end{pmatrix}. \end{aligned}$$

Denote $L_{uv}(\alpha^*)$ as L_{uv} for simplicity. Denote $I^{-1}(\alpha^*, \theta) = \begin{pmatrix} I^{\lambda\lambda} & I^{\lambda\alpha} \\ I^{\alpha\lambda} & I^{\alpha\alpha} \end{pmatrix}$, where

$$\begin{aligned} I^{\lambda\lambda} &= \left(L_{\lambda\lambda} - L_{\lambda\alpha} L_{\alpha\alpha}^{-1} L_{\alpha\lambda}\right)^{-1} \\ &= \left\{\left(L_{13}L_{33}^{-1}L_{31} - L_{11}\right)\theta^2 + 2\left(L_{23}L_{33}^{-1}L_{31} - L_{12}\right)\theta + \left(L_{23}L_{33}^{-1}L_{32} - L_{22}\right)\right\}^{-1}, \\ I^{\lambda\alpha} &= -I^{\lambda\lambda}(L_{23} + \theta L_{13})L_{33}^{-1}, \\ I^{\alpha\alpha} &= L_{33}^{-1}(L_{32} + \theta L_{31})I^{\lambda\lambda}(L_{23} + \theta L_{13})L_{33}^{-1} - L_{33}^{-1}. \end{aligned}$$

Since $\hat{\alpha}_n$ is the MLE of α^* under H_0, it is asymptotically normal with mean α^*, i.e., $\hat{\alpha}_n - \alpha^* = O_P(n^{-1/2})\mathbf{1}$, and satisfies

$$\mathbf{0} = l_{n,3}(1, 1, \hat{\alpha}_n) = l_{n,3}(1, 1, \alpha^*) + l_{n,33}(1, 1, \alpha^*)(\hat{\alpha}_n - \alpha^*) + o_P(||\hat{\alpha}_n - \alpha^*||^2)\mathbf{1}.$$

Since $\{n^{-1}l_{n,33}(1,1,\alpha^*)\}^{-1} \to L_{33}^{-1}(\alpha^*) < \infty$ (a.s.),

$$n^{1/2}(\hat{\alpha}_n - \alpha^*) = -L_{33}^{-1}(\alpha^*)n^{-1/2}l_{n,3}(1,1,\alpha^*) + o_P(1)\mathbf{1}.$$

The score function, evaluated under H_0 and at $\hat{\alpha}_n$, is

$$\begin{aligned} U_n(1,1,\widehat{\alpha}_n,\theta) &= l_{n,2}(1,1,\hat{\alpha}_n) + \theta l_{n,1}(1,1,\hat{\alpha}_n) \\ &= l_{n,2}(1,1,\alpha^*) + l_{n,23}(1,1,\alpha^*)(\hat{\alpha}_n - \alpha^*) \\ &\quad +\theta l_{n,1}(1,1,\alpha^*) + \theta l_{n,13}(1,1,\alpha^*)(\hat{\alpha}_n - \alpha^*) + o_P(||\hat{\alpha}_n - \alpha^*||^2). \end{aligned}$$

Hence

$$\begin{aligned} & n^{-1/2}U_n(1,1,\hat{\alpha}_n,\theta) \\ = \;& n^{-1/2}\left[l_{n,2}(1,1,\alpha^*) + \theta l_{n,1}(1,1,\alpha^*) - \{L_{23}(\alpha^*) + \theta L_{13}(\alpha^*)\}\, L_{33}^{-1}(\alpha^*)l_{n,3}(1,1,\alpha^*)\right] \\ & +o_P(1) \\ = \;& n^{-1/2}\left\{U_n(1,1,\alpha^*,\theta) - I_{\lambda\alpha}I_{\alpha\alpha}^{-1}l_{n,3}(1,1,\alpha^*)\right\} + o_P(1) \\ := \;& n^{-1/2}S_n(\theta,\alpha^*) + o_P(1), \end{aligned} \tag{3}$$

where $S_n(\theta,\alpha^*)$ is a summation of n independent and identically distributed random variables. Let $s(\theta,\alpha) = U_1(1,1,\alpha,\theta) - I_{\lambda\alpha}I_{\alpha\alpha}^{-1}l_{1,3}(1,1,\alpha)$, which is also given in (6) in the paper. Then $E_{H_0}\left(S_n(\theta,\alpha^*)\right) = nE_{H_0}s(\theta,\alpha^*) = 0$ and

$$\text{var}_{H_0}(S_n(\theta,\alpha^*)) = n\text{var}_{H_0}s(\theta,\alpha^*) = n\left(I_{\lambda\lambda} - I_{\lambda\alpha}I_{\alpha\alpha}^{-1}I_{\alpha\lambda}\right). \tag{4}$$

Denote $\sigma(\alpha^*,\theta,\theta) = I_{\lambda\lambda} - I_{\lambda\alpha}I_{\alpha\alpha}^{-1}I_{\alpha\lambda} = (I^{\lambda\lambda})^{-1}$ with a consistent estimator $\sigma(\hat{\alpha}_n,\theta,\theta)$. By the central limit theorem, $n^{-1/2}U_n(1,1,\hat{\alpha}_n,\theta) = n^{-1/2}S_n(\theta,\alpha^*)+o_P(1)$, which converges to $N(0,\sigma(\alpha^*,\theta,\theta))$ in distribution.

From (3), for any given $\theta_i, \theta_j \in (-\infty, \infty)$,

$$\mathrm{cov}_{H_0}\left(n^{-1/2}U_n(1,1,\hat{\alpha}_n,\theta_i), n^{-1/2}U_n(1,1,\hat{\alpha}_n,\theta_j)\right) = \mathrm{cov}_{H_0}\left(s(\theta_i,\alpha^*), s(\theta_j,\alpha^*)\right) + o(1). \quad (5)$$

Denote $I_{\lambda\lambda}^{ij} = \mathrm{cov}_{H_0}\left(U_1(1,1,\alpha^*,\theta_i), U_1(1,1,\alpha^*,\theta_j)\right) = -\left\{L_{22} + L_{12}(\theta_i+\theta_j) + L_{11}\theta_i\theta_j\right\}$ and $L_{\lambda\alpha}^i = \mathrm{cov}_{H_0}\left(U_1(1,1,\alpha^*,\theta_i), l_{1,3}(1,1,\alpha^*)\right) = -\left(L_{23} + \theta_i L_{13}\right)$. Then

$$\mathrm{cov}_{H_0}\left(s(\theta_i,\alpha^*), s(\theta_j,\alpha^*)\right) = I_{\lambda\lambda}^{ij} - I_{\lambda\alpha}^{i} I_{\alpha\alpha}^{-1} I_{\alpha\lambda}^{j} = \sigma(\alpha^*, \theta_i, \theta_j). \quad (6)$$

Under the regularity conditions, the derivatives of $A(\alpha)$, $B(\alpha)$ and $C(\alpha)$ with respective to α, evaluated at α^*, are all bounded. It follows that $\{\sigma(\hat{\alpha}_n,\theta,\theta)\}^{-1/2} = \{\sigma(\alpha^*,\theta,\theta)\}^{-1/2} + o_P(1)$. Thus, from (5),

$$\mathrm{cor}_{H_0}(z(\theta_i), z(\theta_j)) = \mathrm{cov}_{H_0}\left(\frac{n^{-1/2}U_n(1,1,\hat{\alpha}_n,\theta_i)}{\{\sigma(\alpha^*,\theta_i,\theta_i)\}^{1/2}}, \frac{n^{-1/2}U_n(1,1,\hat{\alpha}_n,\theta_j)}{\{\sigma(\alpha^*,\theta_j,\theta_j)\}^{1/2}}\right) + o(1),$$

where the right hand side converges to ρ_{θ_i,θ_j} as $n \to \infty$.

Appendix S5. Proofs

Proof of theorem 1. Denote $l_{n,2}(1,1,\widehat{\alpha}_n)$ as $\widehat{l}_{n,2}$, $l_{n,1}(1,1,\widehat{\alpha}_n)$ as $\widehat{l}_{n,1}$, and $\sigma(\widehat{\alpha}_n,\theta,\theta)$ as $\widehat{\sigma}(\theta,\theta)$, Then, for any $\theta \in (\theta_i,\theta_j) \subset \Theta$, from (2) and (4) in the paper,

$$\begin{aligned}
z(\theta) &= \frac{n^{-1/2}(\widehat{l}_{n,2} + \theta\widehat{l}_{n,1})}{\widehat{\sigma}^{1/2}(\theta,\theta)} \\
&= \frac{n^{-1/2}\widehat{l}_{n,2}}{\widehat{\sigma}^{1/2}(\theta_i,\theta_i)} \frac{\widehat{\sigma}^{1/2}(\theta_i,\theta_i)}{\widehat{\sigma}^{1/2}(\theta,\theta)} \frac{\theta_j-\theta}{\theta_j-\theta_i} + \frac{n^{-1/2}\theta_i\widehat{l}_{n,1}}{\widehat{\sigma}^{1/2}(\theta_i,\theta_i)} \frac{\widehat{\sigma}^{1/2}(\theta_i,\theta_i)}{\widehat{\sigma}^{1/2}(\theta,\theta)} \frac{\theta_j-\theta}{\theta_j-\theta_i} \\
&\quad + \frac{n^{-1/2}\widehat{l}_{n,2}}{\widehat{\sigma}^{1/2}(\theta_j,\theta_j)} \frac{\widehat{\sigma}^{1/2}(\theta_j,\theta_j)}{\widehat{\sigma}^{1/2}(\theta,\theta)} \frac{\theta-\theta_i}{\theta_j-\theta_i} + \frac{n^{-1/2}\theta_j\widehat{l}_{n,1}}{\widehat{\sigma}^{1/2}(\theta_j,\theta_j)} \frac{\widehat{\sigma}^{1/2}(\theta_j,\theta_j)}{\widehat{\sigma}^{1/2}(\theta,\theta)} \frac{\theta-\theta_i}{\theta_j-\theta_i} \\
&= z(\theta_i)\frac{\widehat{\sigma}^{1/2}(\theta_i,\theta_i)}{\widehat{\sigma}^{1/2}(\theta,\theta)} \frac{\theta_j-\theta}{\theta_j-\theta_i} + z(\theta_j)\frac{\widehat{\sigma}^{1/2}(\theta_j,\theta_j)}{\widehat{\sigma}^{1/2}(\theta,\theta)} \frac{\theta-\theta_i}{\theta_j-\theta_i} \\
&= w_i(\theta)z(\theta_i) + w_j(\theta)z(\theta_j),
\end{aligned}$$

where $w_i(\theta)$ and $w_j(\theta)$ are given in theorem 1. This proves part (a).

Under H_0, $\widehat{\alpha}_n \to \alpha^*$ (a.s.). Hence, $w_l(\theta)$ $(l = i, j)$ is a consistent estimator of $W_l(\theta)$, where $\widehat{\alpha}_n$ in $w_l(\theta)$ becomes α^* in $W_l(\theta)$. Using (4) in the paper and the expressions of $W_i(\theta)$ and $W_j(\theta)$, under H_0, we have

$$\rho_{\theta_i,\theta_j} W_j(\theta) + W_i(\theta) = \frac{\sigma(\alpha^*, \theta_i, \theta_j)(\theta - \theta_i) + \sigma(\alpha^*, \theta_i, \theta_i)(\theta_j - \theta)}{\sigma^{1/2}(\alpha^*, \theta, \theta)\sigma^{1/2}(\alpha^*, \theta_i, \theta_i)(\theta_j - \theta_i)},$$

where the numerator is

$$\begin{aligned} & \sigma(\alpha^*, \theta_i, \theta_j)(\theta - \theta_i) + \sigma(\alpha^*, \theta_i, \theta_i)(\theta_j - \theta) \\ = \; & \{A(\alpha^*)\theta_i\theta_j + B(\alpha^*)(\theta_i + \theta_j) + C(\alpha^*)\}(\theta - \theta_i) \\ & + \{A(\alpha^*)\theta_i^2 + 2B(\alpha^*)\theta_i + C(\alpha^*)\}(\theta_j - \theta) \\ = \; & \{A(\alpha^*)\theta\theta_i + B(\alpha^*)(\theta + \theta_i) + C(\alpha^*)\}(\theta_j - \theta_i) = \sigma(\alpha^*, \theta_i, \theta)(\theta_j - \theta_i). \end{aligned}$$

Hence,

$$\rho_{\theta_i,\theta_j} W_j(\theta) + W_i(\theta) = \rho_{\theta_i,\theta}. \tag{7}$$

Likewise, we can prove

$$\rho_{\theta_i,\theta_j} W_i(\theta) + W_j(\theta) = \rho_{\theta_j,\theta}. \tag{8}$$

Expressions for $W_i(\theta)$ and $W_j(\theta)$ in (b) can be obtained directly by solving equations (7) and (8).

$W_i(\theta) + W_j(\theta) > 1$ is equivalent to

$$\sigma^{1/2}(\alpha^*, \theta_i, \theta_i)(\theta_j - \theta)/(\theta_j - \theta_i) + \sigma^{1/2}(\alpha^*, \theta_j, \theta_j)(\theta - \theta_i)/(\theta_j - \theta_i) > \sigma^{1/2}(\alpha^*, \theta, \theta).$$

Since $\theta_i(\theta_j - \theta)/(\theta_j - \theta_i) + \theta_j(\theta - \theta_i)/(\theta_j - \theta_i) = \theta$, it suffices to show $\sigma^{1/2}(\alpha^*, \theta, \theta) = \{A(\alpha^*)\theta^2 + 2B(\alpha^*)\theta + C(\alpha^*)\}^{1/2}$ is a strictly convex function for $\theta \in (\theta_i, \theta_j)$. We show $\partial^2\sigma^{1/2}(\alpha^*, \theta, \theta)/\partial\theta^2 > 0$, which is equivalent to $A(\alpha^*)C(\alpha^*) > \{B(\alpha^*)\}^2$. First, $A(\alpha^*) \geq 0$. If not, let $A(\alpha^*) < 0$. Given that $A(\alpha^*)$, $B(\alpha^*)$ and $C(\alpha^*)$ are all bounded under the regularity conditions, when θ is sufficiently large,

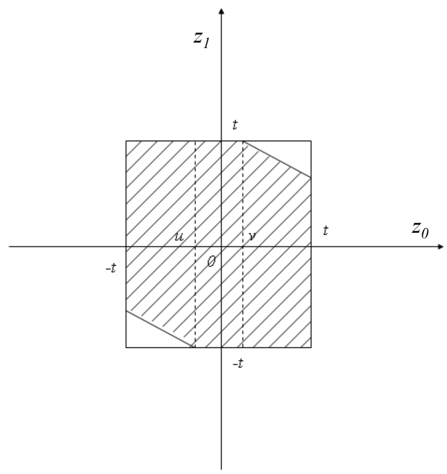
$A(\alpha^*)\theta^2+2B(\alpha^*)\theta+C(\alpha^*) < 0$ due to $A(\alpha^*) < 0$, which is a contradiction to $\sigma(\alpha^*, \theta, \theta) = A(\alpha^*)\theta^2 + 2B(\alpha^*)\theta+C(\alpha^*) > 0$. Since $A(\alpha^*) \geq 0$ and $A(\alpha^*)\theta^2+2B(\alpha^*)\theta+C(\alpha^*) > 0$ for any θ, it follows that $\{B(\alpha^*)\}^2 - A(\alpha^*)C(\alpha^*) < 0$. Thus, $W_i(\theta) + W_j(\theta) > 1$. $W_i^2(\theta) + W_j^2(\theta) + 2W_i(\theta)W_j(\theta)\rho_{\theta_i,\theta_j} = 1$ directly follows from $\text{var}_{H_0}(z(\theta)) = \text{var}_{H_0}(W_i(\theta)z(\theta_i) + W_j(\theta)z(\theta_j)) + o_P(1)$. This completes the proof of (b).

Proof of theorem 2. Let $\theta \in (\theta_i, \theta_j)$. From theorem 1, we have $W_i(\theta) + W_j(\theta) > 1$. Thus, from (7) and (8), $\rho_{\theta_i,\theta} + \rho_{\theta,\theta_j} = \{W_i(\theta) + W_j(\theta)\}(1 + \rho_{\theta_i,\theta_j}) > (1 + \rho_{\theta_i,\theta_j})$. Then, from (7), (8), and $W_i^2(\theta) + W_j^2(\theta) + 2W_i(\theta)W_j(\theta)\rho_{\theta_i,\theta_j} = 1$, we have $\rho^2_{\theta,\theta_i} = (W_i + W_j\rho_{\theta_i,\theta_j})^2 = 1 - W_j^2(1 - \rho^2_{\theta_i,\theta_j})$ and $\rho^2_{\theta,\theta_j} = 1 - W_i^2(1 - \rho^2_{\theta_i,\theta_j})$. Thus, $\rho^2_{\theta,\theta_i} + \rho^2_{\theta,\theta_j} = 2 - (1 - \rho^2_{\theta_i,\theta_j})(W_i^2 + W_j^2) = 2 - (1 - \rho_{\theta_i,\theta_j})(1 + \rho_{\theta_i,\theta_j})(W_i^2 + W_j^2) < 1 + \rho_{\theta_i,\theta_j}$ because $(1 + \rho_{\theta_i,\theta_j})(W_i^2 + W_j^2) > W_i^2 + W_j^2 + 2W_iW_j\rho_{\theta_i,\theta_j} = 1$. This proves (a). Then, from $\rho_{\theta_i,\theta} + \rho_{\theta,\theta_j} \geq 1 + \rho_{\theta_i,\theta_j}$, we have $\rho_{\theta_i,\theta} > \rho_{\theta_i,\theta_j}$ and $\rho_{\theta,\theta_j} > \rho_{\theta_i,\theta_j}$ for any $\theta \in (\theta_i, \theta_j)$. If there exists $(s,t) \in (\theta_i, \theta_j)$ such that $\rho_{s,t} < \rho_{\theta_i,\theta_j}$, $\rho_{\theta_i,s}+\rho_{\theta_i,\theta_j} > \rho_{\theta_i,s}+\rho_{s,t} \geq 1+\rho_{\theta_i,t}$, which implies $\rho_{\theta_i,t} < \rho_{\theta_i,\theta_j}$ for $t \in (\theta_i, \theta_j)$, a contradiction. This proves (b).

Proof of theorem 3. From the results of theorems 1 and 2, the extreme pair for the family of any linear combination of $z(\theta_s)$ and $z(\theta_t)$ with $\theta_s, \theta_t \in [\theta_i, \theta_j]$ is $(z(\theta_i), z(\theta_j))$. Moreover, $\rho_{\theta_i,\theta} + \rho_{\theta,\theta_j} \geq 1+\rho_{\theta_i,\theta_j}$ for any $\theta \in [\theta_i, \theta_j]$. Thus, the MERT for the family is the MERT for the extreme pair, which is given in (5) in the paper. Let $\theta^* \in [\theta_i, \theta_j]$ be the true value of θ under H_1. The ARE of the MERT is $\lim_{n\to\infty} \text{cor}^2_{H_0}(z_{\text{MERT}}, z(\theta^*)) = (\rho_{\theta^*,\theta_i} + \rho_{\theta^*,\theta_j})^2/\{2(1 + \rho_{\theta_i,\theta_j})\} \geq (1 + \rho_{\theta_i,\theta_j})^2/\{2(1 + \rho_{\theta_i,\theta_j})\} = (1 + \rho_{\theta_i,\theta_j})/2$. This proves (a).

The ARE of $z(\theta)$ and $z(\theta_i)$ (or $z(\theta_j)$) is ρ^2_{θ,θ_i} (or ρ^2_{θ,θ_j}). Since $1 + \rho_{\theta_i,\theta_j} \geq \rho^2_{\theta_i,\theta} + \rho^2_{\theta_j,\theta}$ for any $\theta \in [\theta_i, \theta_j]$, $(1 + \rho_{\theta_i,\theta_j})/2 \geq \min(\rho^2_{\theta_i,\theta}, \rho^2_{\theta_j,\theta}) \geq \min_{\theta^*\in[\theta_i,\theta_j]} \rho^2_{\theta,\theta^*}$, the minimum ARE of $z(\theta)$, $\theta \in [\theta_i, \theta_j]$ when the true model $\theta^* \in [\theta_i, \theta_j]$ is unknown. This proves (b).

Figure S1. Region R_z is given in the shaded area in the (z_0, z_1) coordinate system (Zheng et al., 2012, p.168).



Proof of theorem 4. We first consider H_0. Then we generalize the results to H_1, although the results under H_0 can be obtained from those under H_1 as a special case. By theorem 1(a), $z(\theta) = W_0(\theta)z(\theta_0) + W_1(\theta)z(\theta_1) + o_p(1)$ for $\theta \in (0, 1)$. We have $(z(\theta_0), z(\theta_1))^T \to N(\mathbf{0}, \Omega)$ in distribution. Use $\leftrightarrow$ to denote two events having the same probability. Then

$$\{\text{MAX3} \leq t\} \leftrightarrow \{z(\theta_0) \in [-t, t], z(\theta_1) \in [-t, t], W_0(\theta)z(\theta_0) + W_1(\theta)z(\theta_1) \in [-t, t]\}.$$

Denote $R_z = \{(z_0, z_1) : |z_0| \leq t, |z_1| \leq t, |W_0(\theta)z_0 + W_1(\theta)z_1| \leq t\}$. The density of the asymptotic distribution of $(z(\theta_0), z(\theta_1))$ under H_0 is given by

$$h_0(z_0, z_1) = h(z_0, z_1; \rho_{\theta_0,\theta_1}) = \frac{1}{\left(1 - \rho^2_{\theta_0,\theta_1}\right)^{1/2}} \phi(z_0) \phi\left(\frac{z_1 - \rho_{\theta_0,\theta_1} z_0}{\left(1 - \rho^2_{\theta_0,\theta_1}\right)^{1/2}}\right).$$

A figure of R_z can be found in Zheng et al. (2012), which is also given here in Figure S1, where $v = t(1 - W_1(\theta))/W_0(\theta)$ and $u = -v$. Since $W_0(\theta) + W_1(\theta) > 1$ (theorem 1) and (7) and (8) above,

$0 < v < t$. Then

$$\mathrm{pr}_{H_0}\left(\mathrm{MAX3} \le t\right) = \int\int_{(z_0,z_1)\in R_z} h(z_0, z_1; \rho_{\theta_0,\theta_1}) dz_0 dz_1. \tag{9}$$

Zang et al. (2010) expressed the right hand side of (9) in terms of single integrations under the special setting. See also Zheng et al. (2012, p.168). Their derivations were based on the results of theorems 1 and 2 under the special setting. Since we generalize those results, their derivations can be extended accordingly. Denote A_1 as the double integration over the whole rectangular and A_2 and A_3 as the double integrations over the two small triangles. That is,

$$\begin{aligned}
A_1 &= \int_{-t}^{t} d\Phi(z_0) \int_{-t}^{t} d\Phi\left(\frac{z_1 - \rho_{\theta_0,\theta_1} z_0}{\sqrt{1-\rho^2_{\theta_0,\theta_1}}}\right) \\
&= \int_{-t}^{t} \phi(z_0) \left\{ \Phi\left(\frac{t - \rho_{\theta_0,\theta_1} z_0}{\sqrt{1-\rho^2_{\theta_0,\theta_1}}}\right) - \Phi\left(\frac{-t - \rho_{\theta_0,\theta_1} z_0}{\sqrt{1-\rho^2_{\theta_0,\theta_1}}}\right) \right\} dz_0 \\
&= 2\int_{0}^{t} \phi(z_0) \left\{ \Phi\left(\frac{t - \rho_{\theta_0,\theta_1} z_0}{\sqrt{1-\rho^2_{\theta_0,\theta_1}}}\right) - \Phi\left(\frac{-t - \rho_{\theta_0,\theta_1} z_0}{\sqrt{1-\rho^2_{\theta_0,\theta_1}}}\right) \right\} dz_0 = 2I_1(t); \\
A_2 &= \int_{-t}^{-v} d\Phi(z_0) \int_{-t}^{-\frac{t+W_0(\theta)z_0}{W_1(\theta)}} d\Phi\left(\frac{z_1 - \rho_{\theta_0,\theta_1} z_0}{\sqrt{1-\rho^2_{\theta_0,\theta_1}}}\right) \\
&= \int_{-t}^{-v} \phi(z_0) \left\{ \Phi\left(-\frac{t + \rho_{\theta_0,\theta} z_0}{W_1(\theta)\sqrt{1-\rho^2_{\theta_0,\theta_1}}}\right) - \Phi\left(-\frac{t + \rho_{\theta_0,\theta_1} z_0}{\sqrt{1-\rho^2_{\theta_0,\theta_1}}}\right) \right\} dz_0 \\
&= \int_{-t}^{-v} \phi(z_0) \left\{ \Phi\left(\frac{t + \rho_{\theta_0,\theta_1} z_0}{\sqrt{1-\rho^2_{\theta_0,\theta_1}}}\right) - \Phi\left(\frac{t + \rho_{\theta_0,\theta} z_0}{W_1(\theta)\sqrt{1-\rho^2_{\theta_0,\theta_1}}}\right) \right\} dz_0 \\
&= \int_{v}^{t} \phi(z_0) \left\{ \Phi\left(\frac{t - \rho_{\theta_0,\theta_1} z_0}{\sqrt{1-\rho^2_{\theta_0,\theta_1}}}\right) - \Phi\left(\frac{t - \rho_{\theta_0,\theta} z_0}{W_1(\theta)\sqrt{1-\rho^2_{\theta_0,\theta_1}}}\right) \right\} dz_0; \\
A_3 &= \int_{v}^{t} d\Phi(z_0) \int_{\frac{t-W_0(\theta)z_0}{W_1(\theta)}}^{t} d\Phi\left(\frac{z_1 - \rho_{\theta_0,\theta_1} z_0}{\sqrt{1-\rho^2_{\theta_0,\theta_1}}}\right) \\
&= \int_{v}^{t} \phi(z_0) \left\{ \Phi\left(\frac{t - \rho_{\theta_0,\theta_1} z_0}{\sqrt{1-\rho^2_{\theta_0,\theta_1}}}\right) - \Phi\left(\frac{t - \rho_{\theta_0,\theta} z_0}{W_1(\theta)\sqrt{1-\rho^2_{\theta_0,\theta_1}}}\right) \right\} dz_0 = A_2 = I_2(t).
\end{aligned}$$

Note that in A_2 and A_3 we apply (7) and (8). Hence,

$$\mathrm{pr}_{H_0}(\mathrm{MAX3} \le t) = A_1 - A_2 - A_3 = 2\{I_1(t) - I_2(t)\}.$$

Next, we consider H_1. It is known (Huber, 1967; Pfanzagl, 1969) that $\hat{\alpha}_n$ converges to α_θ (a.s.) under H_1. As before, we have $\mathbf{0} = l_{n,3}(1,1,\hat{\alpha}_n) = l_{n,3}(1,1,\alpha_\theta) + l_{n,33}(1,1,\alpha_\theta)(\hat{\alpha}_n - \alpha_\theta) + o_P(||\hat{\alpha}_n - \alpha_\theta||^2)\mathbf{1}$, and $n^{1/2}(\hat{\alpha}_n - \alpha_\theta) = -\left\{n^{-1} l_{n,33}(1,1,\alpha_\theta)\right\}^{-1} n^{-1/2} l_{n,3}(1,1,\alpha_\theta) + o_P(1)\mathbf{1}$. By the definition of α_θ, $l_{n,3}(1,1,\alpha_\theta)$ is a summation of independent and identically distributed random variables with mean

$$E_{H_1,\alpha^*}\left(l_{1,3}(1,1,\alpha_\theta)\right) = \frac{\partial}{\partial \alpha}\left\{\int f(x \mid \lambda, 1-\theta+\theta\lambda, \alpha^*) \log f(x \mid 1,1,\alpha) dx\right\} |_{\alpha=\alpha_\theta} = 0,$$

and $-n^{-1} l_{n,33}(1,1,\alpha_\theta) \to -E(l_{1,33}(1,1,\alpha_\theta)) < \infty$ (a.s.), which is positive definite. So we still have that $n^{1/2}(\hat{\alpha}_n - \alpha_\theta)$ is asymptotical normal with mean zero, or $\hat{\alpha}_n - \alpha_\theta = O_P(n^{-1/2})\mathbf{1}$.

However, $\mu(\lambda,\theta) = E_{H_1,\alpha^*}\left(s(\theta,\alpha_\theta)\right) \neq 0$. Similarly to the steps in (a), we have

$$\begin{aligned}
& n^{-1/2}\left\{U_n(1,1,\hat{\alpha}_n,\theta) - n\mu(\lambda,\theta)\right\} = n^{-1/2}\left\{S_n(\theta,\alpha_\theta) - n\mu(\lambda,\theta)\right\} + o_P(1) \\
= \; & n^{-1/2}\left\{U_n(1,1,\alpha_\theta,\theta) - (L_{23}(\alpha_\theta) + \theta L_{13}(\alpha_\theta)) L_{33}^{-1}(\alpha_\theta) l_{n,3}(1,1,\alpha_\theta) - n\mu(\lambda,\theta)\right\} \\
& + o_P(1) \to N\left(0, \tau^2(\lambda,\theta)\right),
\end{aligned}$$

in distribution. $\sigma(\hat{\alpha}_n,\theta,\theta) \to \sigma(\alpha_\theta,\theta,\theta)$ in probability under H_1 because $\hat{\alpha}_n \to \alpha_\theta$ (a.s.). Thus, under H_1,

$$\begin{aligned}
z(\theta) - \frac{n^{1/2}\mu(\lambda,\theta)}{\sigma^{1/2}(\alpha_\theta,\theta,\theta)} &= \frac{n^{-1/2}\left\{U_n(1,1,\hat{\alpha}_n,\theta)\frac{\sigma^{1/2}(\alpha_\theta,\theta,\theta)}{\sigma^{1/2}(\hat{\alpha}_n,\theta,\theta)} - n\mu(\lambda,\theta)\right\}}{\sigma^{1/2}(\alpha_\theta,\theta,\theta)} \\
&\to N\left(0, \tilde{\sigma}(\lambda,\theta,\theta)\right),
\end{aligned}$$

in distribution, and

$$\left(z(\theta_0) - n^{1/2}\eta_0, z(\theta_1) - n^{1/2}\eta_1\right)^T \to N(\mathbf{0}, \tilde{\Omega})$$

in distribution, where η_i, $i=0,1$, are given in theorem 4 and $\tilde{\Omega}=(\tilde{\omega}_{ij})_{2\times 2}$ is defined right before theorem 4.

The following derivations are similar to those under H_0. Under H_1,

$$\mathrm{pr}_{H_1}(\mathrm{MAX3}\le t)=\mathrm{pr}_{H_1}(z(\theta_0)\in[-t,t],z(\theta_1)\in[-t,t],\tilde{W}_0(\theta)z(\theta_0)+\tilde{W}_1(\theta)z(\theta_1)\in[-t,t]).$$

If we replace α^* with α_θ, then the proofs of theorems 1 and 2 still hold. Hence, R_z given in Figure S1 can still be used under H_1 although the distribution of $(z(\theta_0),z(\theta_1))^T$ is different under H_1. When n is sufficiently large, the density of $(z(\theta_0),z(\theta_1))^T$ under H_1 can be written as

$$h_1(z_0,z_1)=\frac{1}{\sqrt{\tilde{\omega}_{11}}}\phi\left(\frac{z_0-\sqrt{n}\eta_0}{\sqrt{\tilde{\omega}_{11}}}\right)\frac{1}{\sqrt{\tilde{\omega}_{22}\left(1-\tilde{\rho}^2_{\theta_0,\theta_1}\right)}}\phi\left(\frac{\frac{z_1-\sqrt{n}\eta_1}{\sqrt{\tilde{\omega}_{22}}}-\tilde{\rho}_{\theta_0,\theta_1}\frac{z_0-\sqrt{n}\eta_0}{\sqrt{\tilde{\omega}_{11}}}}{\sqrt{1-\tilde{\rho}^2_{\theta_0,\theta_1}}}\right).$$

Hence, the probability over the whole rectangular is given by

$$\begin{aligned}
&\tilde{I}_{1n}(t)\\
=\ &\int_{-t}^{t}\left\{\Phi\left(\frac{\frac{t-\sqrt{n}\eta_1}{\sqrt{\tilde{\omega}_{22}}}-\tilde{\rho}_{\theta_0,\theta_1}\frac{z_0-\sqrt{n}\eta_0}{\sqrt{\tilde{\omega}_{11}}}}{\sqrt{1-\tilde{\rho}^2_{\theta_0,\theta_1}}}\right)-\Phi\left(\frac{\frac{-t-\sqrt{n}\eta_1}{\sqrt{\tilde{\omega}_{22}}}-\tilde{\rho}_{\theta_0,\theta_1}\frac{z_0-\sqrt{n}\eta_0}{\sqrt{\tilde{\omega}_{11}}}}{\sqrt{1-\tilde{\rho}^2_{\theta_0,\theta_1}}}\right)\right\}d\Phi\left(\frac{z_0-\sqrt{n}\eta_0}{\sqrt{\tilde{\omega}_{11}}}\right)\\
=\ &\int_{-\tilde{t}^{(n)}_{0+}}^{\tilde{t}^{(n)}_{0-}}\left\{\Phi\left(\frac{\tilde{t}^{(n)}_{1-}-\tilde{\rho}_{\theta_0,\theta_1}\tilde{z}_0}{\sqrt{1-\tilde{\rho}^2_{\theta_0,\theta_1}}}\right)-\Phi\left(\frac{-\tilde{t}^{(n)}_{1+}-\tilde{\rho}_{\theta_0,\theta_1}\tilde{z}_0}{\sqrt{1-\tilde{\rho}^2_{\theta_0,\theta_1}}}\right)\right\}d\Phi\left(\tilde{z}_0\right),
\end{aligned}$$

where $\tilde{t}^{(n)}_{0\pm}=(t\pm\sqrt{n}\eta_0)/\sqrt{\tilde{\omega}_{11}}$ and $\tilde{t}^{(n)}_{1\pm}=(t\pm\sqrt{n}\eta_1)/\sqrt{\tilde{\omega}_{22}}$ are given in theorem 4. Note that, the probability A_1 in the proof of (a) is a special case of $\tilde{I}_1$ if H_0 is true, under which, $\tilde{\rho}_{\theta_0,\theta_1}=\rho_{\theta_0,\theta_1}$, $\tilde{\omega}_{ii}=1$ $(i=1,2)$, and $\eta_0=\eta_1=0$. So $\tilde{t}^{(n)}_{0\pm}=\tilde{t}^{(n)}_{1\pm}=t$.

Denote $\tilde{v}^{(n)}_{0\pm}=(v\pm\sqrt{n}\eta_0)/\sqrt{\tilde{\omega}_{11}}$, $\tilde{\tilde{t}}^{(n)}_{1\pm}=\left[t\pm\sqrt{n}\left\{\tilde{W}_0(\theta)\eta_0+\tilde{W}_1(\theta)\eta_1\right\}\right]/\sqrt{\tilde{\omega}_{22}}$ and $\tilde{\tilde{\rho}}_{\theta_0,\theta}=\tilde{W}_0(\theta)\sqrt{\tilde{\omega}_{11}}/\sqrt{\tilde{\omega}_{22}}+\tilde{\rho}_{\theta_0,\theta_1}\tilde{W}_1(\theta)$. Then, like the derivation for $\tilde{I}_1$, the probabilities over the two small triangles are given by

$$\tilde{I}_{2n}(t)=\int_{-\tilde{t}^{(n)}_{0+}}^{-\tilde{v}^{(n)}_{0+}}\left\{\Phi\left(\frac{-\tilde{\tilde{t}}^{(n)}_{1+}-\tilde{\tilde{\rho}}_{\theta_0,\theta_1}\tilde{z}_0}{\tilde{W}_1(\theta)\sqrt{1-\tilde{\rho}^2_{\theta_0,\theta_1}}}\right)-\Phi\left(\frac{-\tilde{t}^{(n)}_{1+}-\tilde{\rho}_{\theta_0,\theta_1}\tilde{z}_0}{\sqrt{1-\tilde{\rho}^2_{\theta_0,\theta_1}}}\right)\right\}d\Phi\left(\tilde{z}_0\right)$$

and

$$\tilde{I}_{3n}(t) = \int_{\tilde{v}_{0-}^{(n)}}^{\tilde{t}_{0-}^{(n)}} \left\{ \Phi\left(\frac{\tilde{t}_{1-}^{(n)} - \tilde{\rho}_{\theta_0,\theta_1}\tilde{z}_0}{\sqrt{1-\tilde{\rho}^2_{\theta_0,\theta_1}}} \right) - \Phi\left(\frac{\tilde{\tilde{t}}_{1-}^{(n)} - \tilde{\tilde{\rho}}_{\theta_0,\theta_1}\tilde{z}_0}{\tilde{W}_1(\theta)\sqrt{1-\tilde{\rho}^2_{\theta_0,\theta_1}}} \right) \right\} d\Phi\left(\tilde{z}_0\right).$$

Under H_0, $\tilde{I}_{2n}(t)$ and $\tilde{I}_{3n}(t)$ becomes A_2 and A_3, respectively. Finally, we have

$$\mathrm{pr}_{H_1}\left(\mathrm{MAX3} \le t\right) \approx \tilde{I}_{1n}(t) - \tilde{I}_{2n}(t) - \tilde{I}_{3n}(t).$$

Appendix S6. Conditional probabilities of the trio design

Table S2: Conditional probabilities for the case-parents trio design under H_0 and H_1. Under H_1, the true model is $\theta^* = 0, 0.5, 1$.

		H_1		
MT	H_0	$\theta^* = 0$	$\theta^* = 0.5$	$\theta^* = 1$
$I : AA \times AB$	$p_{10} = \frac{1}{2}$	$p_{10} = \frac{1}{2}$	$p_{10} = \frac{2}{3+\lambda}$	$p_{10} = \frac{1}{1+\lambda}$
	$p_{11} = \frac{1}{2}$	$p_{11} = \frac{1}{2}$	$p_{11} = \frac{1+\lambda}{3+\lambda}$	$p_{11} = \frac{\lambda}{1+\lambda}$
$II : AB \times AB$	$p_{20} = \frac{1}{4}$	$p_{20} = \frac{1}{3+\lambda}$	$p_{20} = \frac{1}{2(1+\lambda)}$	$p_{20} = \frac{1}{1+3\lambda}$
	$p_{21} = \frac{1}{2}$	$p_{21} = \frac{2}{3+\lambda}$	$p_{21} = \frac{1+\lambda}{2(1+\lambda)}$	$p_{21} = \frac{2\lambda}{1+3\lambda}$
	$p_{22} = \frac{1}{4}$	$p_{22} = \frac{\lambda}{3+\lambda}$	$p_{22} = \frac{\lambda}{2(1+\lambda)}$	$p_{22} = \frac{\lambda}{1+3\lambda}$
$III : AB \times BB$	$p_{31} = \frac{1}{2}$	$p_{31} = \frac{1}{1+\lambda}$	$p_{31} = \frac{1+\lambda}{1+3\lambda}$	$p_{31} = \frac{1}{2}$
	$p_{32} = \frac{1}{2}$	$p_{32} = \frac{\lambda}{1+\lambda}$	$p_{32} = \frac{2\lambda}{1+3\lambda}$	$p_{32} = \frac{1}{2}$

Additional references

Gustafson, P., Le, N.D. & Vallee, M. (2002). A Bayesian approach to case-control studies with errors in covariables. *Biostatistics* **3**, 229–243.

Prentice, R. & Pyke, R. (1979). Logistic disease incidence models and case-control studies. *Biometrika* **66**, 403–411.

Qin, J. & Zhang, B. (1997). A goodness-of-fit test for logistic regression models based on case-control data. *Biometrika* **84**, 609–618.

Seaman, S.R. & Richardson, S. (2004). Equivalence of prospective and retrospective models in the Bayesian analysis of case-control studies. *Biometrika*, **91**, 15–25.

Wang, K. & Sheffield, V. (2005). A constrained-likelihood approach to marker-trait association studies. *Am. J. Hum. Genet.* **77**, 768–780