Asymptotic Relative Efficiencies of the Score and Robust Tests in Genetic Association Studies

Ao Yuan[1], Ruzong Fan[1], Jinfeng Xu[2], Qizhai Li[3]

[1]Department of Biostatistics, Bioinformatics and Biomathematics, Georgetown University, Washinton DC 20057;

[2]Department of Statistics and Actuarial Science, University of Hong Kong, Hong Kong;

[3]Academy of Mathematics and Systems Science, Chinese Academy of Sciences, China, Beijing 100190, P.R. China.

Abstract

We compare the asymptotic behavior of two commonly used test statistics, the score statistic $Z(\theta)$ and the maximin efficient robust test statistic Z_{MERT} arise in genetic association study, by computing their asymptotic relative efficiencies (AREs) relative to each other. Four commonly used ARE measures, the Pitman ARE, Chernoff ARE, Hodges-Lehmann ARE and the Bahadur ARE are considered. Some modifications of these methods are made to simplify the computations. We found the Chernoff, Hodges-Lehmann and Bahadur AREs are suitable for our setting. Based on our study, the efficiencies of the two test statistic varies for different criterion used, and for different parameter values under the same criterion, which described in the context.

Key words: Asymptotic relative efficiency, genetic association study, maximin efficiency robust test.

1. Introduction

In genetic association studies, several test statistics are often used. It is of meaning to compare their asymptotic performances. Although for likelihood ratio based test statistic for testing hypothesis of simple null versus simple alternative, there is an uniformly most powerful test under some regularity conditions. However, most test statistics are not constructed directly from likelihood ratio, the hypothesis are composite, and there is generally no such optimal test. Therefore, the classical method to compare any two test statistics is to evaluate the asymptotic relative efficiency (ARE) between them.

The ARE is a well studied area, with vast literatures and numerous different definitions. But often the computation of ARE is very difficult in the general case, some of the classical methods for ARE require that the test statistics have some standard forms, such as they have the same asymptotic distribution, or have the forms of i.i.d. summations. However, in practice, such as in genetic association studies, some test statistics do not have these forms. Sitlani and McKnight (2011) studied AREs for the trend test under different models and stratifications. In this communication, we compare the asymptotic behavior of two commonly used test statistics the score statistic $Z(\theta)$ and the maximin efficient robust test statistic Z_{MERT}, arise in case-control genetic association study, as given in Zheng, Li and Yuan (2011), hereafter ZLY, by evaluate their AREs relative to each other. Four commonly used ARE measures, the Pitman ARE, Chernoff ARE, Hodges-Lehmann ARE and the Bahadur ARE are considered. Pitman's ARE does not apply directly. We found the Chernoff, Hodges-Lehmann and the Bahadur AREs are suitable for our setting. Some modifications of these methods are made to simplify the computations.

Existing studies on ARE are mainly focused on two categories. One is to compare efficiencies of estimators of the same parameter; the other is to compare test statistics of the same hypothesis, in which the test statistics may not estimate the same parameter. The latter study can be under the assumption that the test statistics in comparison are asymptotic normality. In this case, the ARE can often be easily computed. There are also methods for compare ARE of different test statistics in general, in which different test statistics of the same hypothesis may have different asymptotic distributions. In this general case, Pitman, Bahadure and Hodges-Lehmann proposed different ways to compute the ARE, and it is often difficult. Although, when the test statistics have the same asymptotic distribution, the ARE can be computed essily. We also give a simple definition of ARE, so that the ARE can be computed in the case of different asymptotic distributions, as long as the asymptotic distributions of the test statistics are known.

In Section 2, we describe the background of the genetic association study problem and a brief review of the classical definitions of ARE. In Section 3 we compare the ARE of the test statistics arose from our genetic association study. We found that he performances, or the efficiencies of the two test statistic varies for different criterion used, and for different parameter values under the same criterion, which described in the context.

2. Background

Denote the log-likelihood function as $l_n(\lambda_1, \lambda_2, \eta) = \sum_{i=1}^{n} \log f(Y_i|\lambda_1, \lambda_2, \eta^T X_i)$, where Y_i is the outcome, $(\lambda_1, \lambda_2) \in \Lambda \subset R^2$ are the parameters of interest, $\eta \in R^m$ is a vector of parameters $(m \geq 0)$ for the covariate $X_i = (x_{i1}, \ldots, x_{im})^T$, and n is the sample size. We test H_0 : $(\lambda_1, \lambda_1) = (1,1)$ against the alternative hypothesis H_1 : $(r_1, r_2) \in \Lambda - \{(1,1)\}$,

(λ_1, λ_2) (λ_1, λ_2)

由 扫描全能王 扫描创建

where Λ has two edges with known slopes θ_0 and θ_1, and the null point $(1, 1)$ is on the boundary of Λ. We assume $-\infty < \theta_0 < \theta_1 < \infty$ and the endpoints θ_0 and θ_1 satisfy some constraints as specified in ZLY. When $\theta_1 = \infty$ which corresponds to a vertical edge, we can switch λ_1 and λ_2 and define new (θ_0, θ_1) so that $-\infty < \theta_0 < \theta_1 < \infty$ is satisfied by the new (θ_0, θ_1). For example, we can write $\lambda_1 = 1 + (\lambda_2 - 1)/\theta_1 = 1 + \theta_1^*(\lambda_2 - 1)$ and $\lambda_1 = 1 + (\lambda_2 - 1)/\theta_0 = 1 + \theta_0^*(\lambda_2 - 1)$, where $-\infty < \theta_1^* < \theta_0^* < \infty$.

Assume θ_0 and θ_1 are known based on the problem of interest and/or scientific knowledge. Given $\lambda_1 = \lambda \geq 1$, λ_2 can be written as $\lambda_2 = 1 - \theta + \theta\lambda$, $\theta \in [\theta_0, \theta_1]$. We treat θ as a nuisance parameter not estimable under $H_0 : \lambda = 1$, but assume η is estimable under H_0. The log-likelihood function becomes $l_n(\lambda, \eta, \theta)$. The score test for $H_0 : \lambda = 1$ is given by

$$Z(\theta) = \frac{\frac{\partial}{\partial\lambda} l_n(\lambda, \eta, \theta)|_{H_0, \widehat{\eta}_n}}{\left\{Var_{H_0, \widehat{\eta}_n}\left(\frac{\partial}{\partial\lambda} l_n(\lambda, \eta, \theta)\right)\right\}^{1/2}},$$

where $\widehat{\eta}_n$ is the MLE of η under H_0. It would be difficult to deal with $l_n(\lambda, \eta, \theta)$ because θ in $Z(\theta)$ is implicitly expressed.

We work with $l_n(\lambda, 1 - \theta + \theta\lambda, \eta)$, where θ is explicitly expressed. We consider $l_n(\lambda, 1 - \theta + \theta\lambda, \eta)$ as a multivariate function with three variables $x_1 = \lambda$, $x_2 = 1 - \theta + \theta\lambda$ and $x_3 = \eta$. Denote $l_{n,u} = \partial l_n / \partial x_u$ for $u = 1, 2, 3$, $l_{n,uv} = \partial^2 l_n / \partial x_u \partial x_v$ for $u = 1, 2$ and $v = 1, 2, 3$, and $l_{n,33} = \partial^2 l_n / \partial x_3 \partial x_3^T$. Assume $l_{n,uv} = l_{n,vu}$ for $u, v = 1, 2$, $l_{n,uv} = l_{n,vu}^T$ for $u = 1, 2$ and $v = 3$.

Suppose we have a family of asymptotically normally distributed tests $T_0 = \{Z(\theta) : \theta \in [a, b]\}$, where $Z(\theta) \xrightarrow{D} N(0, 1)$ under $H_0 : \lambda = 1$ for a given $\theta \in [a, b]$, which determines the data-generating model under $H_1 : \lambda \geq 1$. When $\theta = \theta^{(0)} \in [a, b]$ is the true value, $Z(\theta^{(0)})$ is asymptotically most powerful (optimal). In this case, when $\theta^{(1)} \neq \theta^{(0)}$ is used, the Pitman ARE of $Z(\theta^{(1)})$ relative to $Z(\theta^{(0)})$ is given by (Gastwirth, 1966, 1985)

$$e\left(Z(\theta^{(1)}), Z(\theta^{(0)})\right) = \rho^2_{\theta^{(0)}, \theta^{(1)}},$$

where $\rho_{\theta^{(0)}, \theta^{(1)}}$ is the asymptotic null correlation coefficient between $Z(\theta^{(0)})$ and $Z(\theta^{(1)})$ for $\theta^{(0)}, \theta^{(1)} \in [a, b]$. Let T_1 be a set of all convex linear combinations of T_0. A simple robust test derived under efficiency robust theory (Gastwirth, 1966, 1985; Birnbaum and Laska, 1967) is the maximin efficient robust test (MERT), denoted as Z_{MERT}.When $T_0 = \{Z(\theta_i), Z(\theta_j)\}$, Z_{MERT} is given by

$$Z_{\text{MERT}} = (Z(\theta_i) + Z(\theta_j)) \,/\, \left\{2(1 + \rho_{\theta_i, \theta_j})\right\}^{1/2}.$$

When T_0 has more than two members, generally Z_{MERT} exists and is unique (Gastwirth 1966), but its computation needs quadratic programming methods (Rosen 1960). However, when there is an extrem pair $(Z(\theta_i), Z(\theta_j))$ in T_0, i.e. $\rho_{\theta_i,\theta_j} = \inf_{\theta,\theta' \in [a,b]} \rho_{\theta,\theta'} > 0$, then $Z_{\text{MERT}} = (Z(\theta_i) + Z(\theta_j)) / \{2(1+\rho_{\theta_i,\theta_j})\}^{1/2}$ is MERT for T_0 if and only if (Gastwirth 1970)

$$\rho_{\theta_i,\theta} + \rho_{\theta_j,\theta} \geq 1 + \rho_{\theta_i,\theta_j}, \quad \forall\, \theta \in [a,b].$$

and thus

$$e\left(Z_{\text{MERT}}, Z(\theta^{(0)})\right) = \sup_{Z \in T_1} \inf_{\theta \in [a,b]} e(Z, Z(\theta)).$$

That is, the MERT reaches the maximin ARE due to model uncertainty. The MERT was first derived for linear rank tests for the two-sample problem (Gastwirth, 1966; Birnbaum and Laska, 1967) and later extended to a family of asymptotically normally distributed tests T_0 (Gastwirth, 1985).

The $Z(\theta)$ statistic has the following property (ZLY): Let $\theta \in [\theta_i, \theta_j] \subseteq [\theta_0, \theta_1]$. Then

$$Z(\theta) = \sum_{l=i,j} W_l(\theta) Z(\theta_l),$$

where $W_i(\theta) = \{\sigma(\theta_i)/\sigma(\theta)\}\{(\theta_j - \theta)/(\theta_j - \theta_i)\}$ and $W_j(\theta) = \{\sigma(\theta_j)/\sigma(\theta)\}\{(\theta - \theta_i)/(\theta_j - \theta_i)\}$.

Let $\widehat{\eta}_{0,n}$ be the MLE of η under H_0, and $(\widehat{\eta}_{1,n}, \hat{\lambda}_n)$ be that of (η, λ) under H_1. For given θ, the χ^2 likelihood ratio test statistic is $T(\theta) = 2[l_n(\hat{\lambda}_n, 1 - \theta + \theta\hat{\lambda}_n, \widehat{\eta}_{1,n}) - l_n(1, 1, \widehat{\eta}_{0,n})]$. 么?
By the Wilk's theorem, under H_0,

$$T(\theta) \xrightarrow{D} \chi_1^2,$$

the chisquared distribution with one degree of freedom. The likelihood ratio test is also widely used in genetic association studies, its properties, including its ARE is well studied in the literature, so we will not investigate it here.

Let the MLE $\widehat{\eta}_n$ satisfy $\partial l_n / \partial \eta|_{H_0, \widehat{\eta}_n} = l_{n,3}(1, 1, \widehat{\eta}_n) = \mathbf{0}$. Here $\mathbf{0}$ presents a vector of 0's. Let η_0 be the true value (unknown) of η under either H_0 or H_1. Since η_0 is unknown, we define the score function as

$$U_n(1,1,\widehat{\eta}_n) = \frac{\partial l_n}{\partial \lambda}|_{H_0, \widehat{\eta}_n} = l_{n,1}(1,1,\widehat{\eta}_n) + \theta l_{n,2}(1,1,\widehat{\eta}_n)$$

and the test statistic for H_0 as

$$z(\theta) = \frac{U_n(1,1,\widehat{\eta}_n)}{\{Var_{H_0}(U_n(1,1,\widehat{\eta}_n))\}^{1/2}} \sim \frac{n^{-1/2} U_n(1,1,\widehat{\eta}_n)}{\{(1,\mathbf{0}) I^{-1}(\eta_0)(1,\mathbf{0})^T\}^{-1/2}}$$

$$= \frac{n^{-1/2}\{l_{n,1}(1,1,\widehat{\eta}_n) + \theta l_{n,2}(1,1,\widehat{\eta}_n)\}}{(A_{\eta_0}\theta^2 + 2B_{\eta_0}\theta + C_{\eta_0})^{1/2}},$$

是 η_0 还是 $\widehat{\eta}_n$

where "~" means asymptotically equivalent, $A_\eta = L_{23}(\eta) L_{33}^{-1}(\eta) L_{32} - L_{22}(\eta)$, $B_\eta = L_{13}(\eta) L_{33}^{-1}(\eta) L_{31}(\eta) - L_{12}(\eta)$ and $C_\eta = L_{13}(\eta) L_{33}^{-1}(\eta) L_{31}(\theta) - L_{11}(\eta)$.

忘记定义L，下面是否正确　　是 $L_{31}(\eta)$ 么？

$L_{uv}(\eta) = E_{H_0}(l_{1,uv}(1,1,\eta))$

计算时 $= \frac{1}{n} l_{n,uv}(1,1,\eta)$

Denote $W = \begin{pmatrix} 1 & 0 \\ W_0(\theta) & W_1(\theta) \\ 0 & 1 \end{pmatrix}$. For a vector $v = (v_1, v_2, v_3)^T$, denote $||v|| = \sum_{i=1}^{3} v_i$.

Let $f(y|\lambda, 1-\theta+\theta\lambda, \eta)$ be the true density of the data y. The null model is $f(\cdot|1,1,\eta)$ and the alternative model is $f(\cdot|\lambda, 1-\theta+\theta\lambda, \eta)$. The following notation is also used under H_1. For fixed (λ, θ), let

$$\eta_\theta \triangleq \alpha(\lambda,\theta) = \arg\sup_\eta \int f(x|\lambda, 1-\theta+\theta\lambda, \eta_0) \log f(x|1,1,\eta) dx.$$

Under H_1, the empirical version of η_θ is just $\hat{\eta}_n$. We denote the Fisher information and its inverse in the blocked forms as

$$I(\eta_0) = \begin{pmatrix} I_{\lambda\lambda} & I_{\lambda\eta} \\ I_{\eta\lambda} & I_{\eta\eta} \end{pmatrix}, \quad \text{and} \quad I^{-1}(\eta_0) = \begin{pmatrix} I^{\lambda\lambda} & I^{\lambda\eta} \\ I^{\eta\lambda} & I^{\eta\eta} \end{pmatrix}.$$

Let $\sigma^2(\theta) = \{(1,0)I^{-1}(\alpha_0)(1,0)^T\}^{-1} = (I^{\lambda\lambda})^{-1}$,

$$s(\theta,\eta) = l_{1,1}(1,1,\eta) + \theta l_{1,2}(1,1,\eta) - (L_{13}^T(\eta) + \theta L_{23}^T(\eta))L_{33}^{-1}(\eta) l_{1,3}(1,1,\eta),$$

$\mu(\lambda,\theta) = E_{H_1,\eta_0}(s(\theta,\eta_\theta))$, $u(\lambda,\theta) \triangleq s(\theta,\eta_\theta) - \mu(\lambda,\theta)$, $\tau^2(\lambda,\theta) = E_{H_1,\eta_0}(u(\lambda,\theta))^2$, $\tilde{\sigma}^2(\lambda,\theta) = \tau^2(\lambda,\theta)/\sigma^2(\theta)$, and $\tilde{\Omega} = (\tilde{\omega}_{ij})_{2\times 2}$ with $\tilde{\omega}_{11} = \tilde{\sigma}^2(\lambda,\theta_0)$, $\tilde{\omega}_{22} = \tilde{\sigma}^2(\lambda,\theta_1)$, and $\tilde{\omega}_{12} = E_{H_1,\alpha_0}\{u(\lambda,\theta_0)u(\lambda,\theta_1)\} / \{\sigma(\theta_0)\sigma(\theta_1)\}$. The empirical version of $\tilde{\Omega}$ is given by $\tilde{\Omega}_n$, where η_θ in $\tilde{\Omega}$ is replaced by $\hat{\eta}_n$.

计算时 $\mu(\lambda,\theta) = \frac{1}{n}[l_{n,1}(1,1,\eta) + \theta l_{n,2}(1,1,\eta) - (L_{13}^T(\eta) + \theta L_{23}^T(\eta))L_{33}^{-1}(\eta) l_{n,3}(1,1,\eta)]$ 对么?

参数λ体现在哪?

Below we give a brief review of the notions of ARE for test statistics in the general case, more detailed account can be found in Serfling (1980) and Nikitin (2011).

The calculation of the existing of versions of ARE is generally not easy, as in the examples (Serfling, 1980; Nikitin, 1995; van der Varrt, 1998). We only point out, that the Pitman ARE is based on the central limit theorem for test statistics, that the Bahadur ARE requires the large deviation asymptotics of test statistics under the null-hypothesis, while the Hodges-Lehmann ARE is connected with large deviation asymptotics under the alternative. Each type of ARE has its own advantage and dis-advantage, and the different notions of ARE are not always give consistent conclusion.

If the condition of asymptotic normality (or common asymptotic distribution) fails, considerable difficulties arise when calculating the Pitman ARE as it may not at all exist or may depend on α and β. Usually one considers limiting Pitman ARE as $\alpha \to 0$. Wieand (1976) established the correspondence between this kind of ARE and the limiting approximate Bahadur efficiency which is easy to calculate.

$$\mu^{(1)}(\lambda,\theta) = \frac{1}{n}[L_{n,11}(1,1,\eta) + \theta L_{n,21}(1,1,\eta) - (L_{13}^T(\eta) + \theta L_{23}^T(\eta))L_{33}^{-1}(\eta) L_{n,31}(1,1,\eta)]$$

$$= L_{11}(\eta) + \theta L_{21}(\eta) - (L_{13}^T(\eta) + \theta_{23}^T(\eta))L_{33}^{-1}(\eta) L_{31}(\eta)$$ 对么?

Let $I(\lambda_0)$ be the Fisher information at λ_0. Under some additional conditions, Rao (1963) proved that

$$\left(\frac{\sigma(\lambda_0)}{\mu^{(1)}(\lambda_0)}\right)^2 \geq I^{-1}(\lambda_0).$$

Any test statistic S_n achieves the equality in the above is called *Pitman efficient.*

Under suitable conditions, Pitman ARE can be expressed in terms of correlation coefficient between the two test statistics in their standardized form, as given below.

(P5). $(S_{1,n} - \mu_{1,n}(\lambda))/\sigma_{1,n}(\lambda)$ and $(S_{2,n} - \mu_{2,n}(\lambda))/\sigma_{2,n}(\lambda)$ are asymptotic joint normal uniformly in a neighborhood of λ_0.

Denote $\rho(\lambda)$ the asymptotic correlation coefficient between them under λ, and Φ and ϕ be the distribution and density function of $N(0,1)$. The following result is true.

(van Eden, 1963). Assume that $S_{1,n}$ and $S_{2,n}$ satisfy (P1)-(P5) in their standardized form with $H = \Phi$, $k = 1$ and $d(n) = n^{1/2}$, and that $\rho(\lambda_n) \to \rho(\lambda_0) := \rho$ as $\lambda_n \to \lambda_0$. Then
(i) For $0 \leq \lambda \leq 1$, tests of the form $S_{\gamma,n} = (1-\gamma)S_{1,n} + \gamma S_{2,n}$ satisfy (P1)-(P5), and the "best" $S_{\gamma,n}$ which maximizes $e_P(\{S_{\lambda,n}\},\{S_{1,n}\})$ is the one with

$$\gamma = \frac{c_1 - \rho c_2}{(1-\rho)(c_1+c_2)} = \frac{e_P^{1/2}(\{S_{2,n}\},\{S_{1,n}\}) - \rho}{(1-\rho)[1 + e_P^{1/2}(\{S_{2,n}\},\{S_{1,n}\})]}$$

and

$$e_P(\{S_{\gamma,n}\},\{S_{1,n}\}) = 1 + \frac{[e_P^{1/2}(\{S_{2,n}\},\{S_{1,n}\})]^2}{1-\rho^2}.$$

(ii) If $S_{1,n}$ is the best test satisfying (P1)-(P5), then

$$e_P(\{S_{2,n}\},\{S_{1,n}\}) = \rho^2. \tag{4}$$

In the typical case, S_n is an i.i.d. summation (upto scale), then $\mu_n(\lambda) = n\mu(\lambda)$, $\sigma_n(\lambda) = \sqrt{n}\sigma(\lambda)$, $d(n) = \sqrt{n}$, $k = 1$, $c = \sigma(\lambda_0)/\mu'(\lambda_0)$ and $\lambda_n = \lambda_0 + n^{-1/2}\sigma(\lambda_0)/\mu'(\lambda_0)[F^{-1}(1-\alpha) - F^{-1}(\beta)]$.

Note $e_P(\{S_{1,n}\},\{S_{2,n}\})$ does not depend on (α,β), thus if $e_P(\{S_{1,n}\},\{S_{2,n}\}) < 1$ or $c_1 < c_2$, then $\{S_{1,n}\}$ is better than $\{S_{2,n}\}$ for all (α,β).

是不是 $e_p(\{S_{2,n}\},\{S_{1,n}\})<1$

Pitman ARE given by (3) or (4) are easy to use. However, they require the two comparing test statistics have the same asymptotic distribution (after standardization), (4) require further that they are jointly asymptotic normal. In practice, these conditions some times

(ii)

$$\tilde{Q}_{Z(\theta^{(0)})} = 2\left(1 - \Phi\left(\frac{\mu^{(1)}(\lambda_0, \theta^{(0)})}{2\sigma(\theta^{(0)})}\right)\right),$$

$$\tilde{Q}_{Z_{MERT}} = 2\left(1 - \Phi\left(\left[\frac{\mu^{(1)}(\lambda_0, \theta_i)}{\sigma(\theta_i)} + \frac{\mu^{(1)}(\lambda_0, \theta_j)}{\sigma(\theta_j)}\right] / \sqrt{8(1+\rho_{\theta_i,\theta_j})}\right)\right).$$

这里的 λ_0，在计算时 $\lambda_0 = 1$ ？

Hodges-Lehmann ARE. Consider testing the null hypothesis be $H_0 : \lambda \in \Lambda_0$ vs $H_1 : \lambda \in \Lambda_1$, given a level α test statistic S_n with critical value $t_n(\alpha) : \alpha_n := \sup_{\lambda\in\Lambda_0} P_\lambda(S_n \geq t_n(\alpha)) \to \alpha$. For $\lambda \in \Lambda_1$, the type II error at λ is $\beta_n(\lambda) = P_\lambda(S_n \leq t_n(\alpha))$. Typically, $\beta_n(\lambda)$ tends to zero at exponential rate, the faster the better S_n is. Hodges and Lehmann (1956) proposed

$$d(\lambda) = \lim_n -2n^{-1} \log \beta_n(\lambda)$$

as a measure of the performance of S_n and it called the Hodges-Lehmann index of the statistic S_n. For two test statistics $S_{1,n}$ and $S_{2,n}$ for the same H_0 vs H_1, with $d_1(\lambda)$ and $d_2(\lambda)$, the Hodges-Lehmann ARE of $\{S_{1,n}\}$ relative to $\{S_{2,n}\}$ at $\lambda \in \Lambda_1$ is defined as

$$e_{HL}(\{S_{1,n}\}, \{S_{2,n}\}) = \frac{d_1(\lambda)}{d_2(\lambda)}.$$

For probability density functions f and g, let $K(f,g) = \int f(x)\log[f(x)/g(x)]dx$ be the Kullback-Leibler divergence between f and g. For any test statistic $S_n(X_1, ..., X_n)$ based on $X_1, ..., X_n$ i.i.d. density $f(\cdot|\lambda)$, the Hodges-Lehmann index has the following property

$$\lim_n (1 - \beta_n(\lambda)) \geq -\inf\{K(f(\cdot|\lambda_0), f(\cdot|\lambda)) : \lambda_0 \in \Lambda_0\},$$

and any test statistic acheive the equality in the above is said to be *Hodges-Lehmann efficient.*

Compared to the Pitman and Chernoff ARE, the Hodges-Lehmman ARE does not require the comparing test statistic have the same asymptotic distribution, nor they have the form of i.i.d. summations, so it has wilder application scope.

Proposition 3. Under conditions of Theorem 4 in Zheng et al. (2010), with $\mu_{MERT}(\lambda) := [\mu(\lambda,\theta_i)/\sigma(\theta_i) + \mu(\lambda,\theta_j)/\sigma(\theta_j)]/\sqrt{2(1+\rho_{\theta_i,\theta_j})}$, and $\tilde{\zeta}$ given in (2), for $\lambda > 1$, we have

$$d_{Z(\theta)}(\lambda) = \frac{\mu^2(\lambda,\theta)}{\sigma^2(\theta)}, \qquad d_{Z_{MERT}}(\lambda) = \mu^2_{MERT}(\lambda).$$

For the chisquared test T, under H_1 its asymptotic distribution is a non-central chisquared distribution, with no-closed form. So its Hodges-Lehmann ARE is not directly available.

这里的

Bahadur efficiency of likelihood ratio test has been studied by a number of researchers for some special distribution families. Arcones (2005, Theorem 3.3) proved that, under some regularity conditions, the likelihood ratio statistic is Bahadur efficient. Let $f(\cdot|\lambda,\theta,\eta)$ be the density function of the data, under his conditions of Theorem 3.3, for each fixed $\lambda > 1$ and θ, we have

$$c_T = -2\inf\inf\{K(f(\cdot|\lambda,\theta,\eta), f(\cdot|\eta,\lambda_0)) : \eta\}.$$

Like the Hodges-Lehmman ARE, Bahadur ARE does not require the comparing test statistic have the same asymptotic distribution, nor they have the form of i.i.d. summations, so it has wide application scope.

For computation easiness, we consider a local version of Bahadur ARE. Consider testing $H_0 : \lambda = \lambda_0$ vs the local alternative $H_n : \lambda = \lambda_0 + n^{-1/2}$. Let F_0 be the asymptotic distribution function of S_n under H_0, we define

$$\tilde{c} = \lim_n [1 - F_0(S_n|H_n)].$$

Typically, $0 < \tilde{c} < 1$. The smaller $\tilde{c}$, the better S_n is. For two test statistics $S_{i,n}$ $(i = 1,2)$ for the same hypothesis with $G_{i,n}$ and $\tilde{c}_i$, we define the local Bahadur ARE of $S_{1,n}$ relative to $S_{2,n}$ as

$$\tilde{e}_B(\{S_{1,n}\},\{S_{2,n}\}) = \frac{\tilde{c}_2}{\tilde{c}_1}.$$

Proposition 4. *i) with $\mu_{MERT}(\lambda)$ given in Proposition 3, we have*

$$c_{Z(\theta)}(\lambda) = \mu^2(\lambda,\theta)/\sigma^2(\theta), \qquad c_{Z_{MERT}}(\lambda) = \mu^2_{MERT}(\lambda).$$

ii) Under conditions of Theorem 4 in ZLY, with $\mu^{(1)}_{MERT}(\lambda)$ be the derivative of $\mu_{MERT}(\lambda)$, we have

$$\tilde{c}_{Z(\theta)} = 1 - \Phi(\mu^{(1)}\underline{(1,\theta_0)/\sigma(\theta_0)}), \quad \tilde{c}_{Z_{MERT}} = 1 - \Phi(\mu^{(1)}_{MERT}(1)).$$

这里的 θ_0 对应真实的θ，也就是之前的 $\theta^{(0)}$ 么?

Appendix

Derivation of $\tilde{c}_i$: From (P3), we have $c_i = \lim_n d(n)\sigma_{i,n}(\lambda_0)/\mu_{i,n}(\lambda_0)$. Also, as in the proof in Serfling (1980, p.317-318), $P_{\lambda_0}(S_{i,n} > u_{\alpha,i,n}) \to \alpha$ and $\beta_{i,n}(\lambda_n) := P_{\lambda_n}(S_{i,n} \le u_{\alpha,i,n}) \to \beta$ if and only if

$$\frac{(\lambda_n - \lambda_0)^k}{k!}\frac{d(n)}{c_i} \to F_i^{-1}(1-\alpha) - F_i^{-1}(\beta) \quad \text{or} \quad \frac{(\lambda_n - \lambda_0)^k}{k!}\frac{d(n)}{\tilde{c}_i} \to 1.$$

The computation of four AREs

Yuan Xue
xueyuan115@mails.ucas.ac.cn

July 29, 2018

1 Introduction

1.1 Notation

likelihood function

$$L_n(\lambda_1, \lambda_2, \eta) = \frac{\lambda_1^{r_1} \lambda_2^{r2} exp(r\eta)}{\{1 + exp(\eta)\}^{n_0}\{1 + \lambda_1 exp(\eta)\}^{n_1}\{1 + \lambda_2 exp(\eta)\}^{n_2}} \tag{1}$$

log-likelihood function $l_n(\lambda, \eta, \theta)$
we work with $l_n(\lambda, 1 - \theta - +\theta\lambda, \eta)$, where θ is explicitly epressed.

$$\begin{aligned}
x_1 =& \lambda \\
x_2 =& 1 - \theta - \theta\lambda \\
x_3 =& \eta
\end{aligned}$$

Denote

$$\begin{aligned}
l_{n,\mu} =& \partial l_n / \partial x_\mu \quad \text{for} \quad \mu = 1, 2, 3 \\
l_{n,\mu\nu} =& \partial^2 l_n / \partial x_\mu \partial x_\nu \quad \text{for} \quad \mu = 1, 2, \nu = 1, 2, 3 \\
l_{n,33} =& \partial^2 l_n / \partial x_3 \partial x_3 \\
l_{n,\mu\nu} =& l_{n,\nu\mu} \quad \text{for} \quad \mu, \nu = 1, 2 \\
l_{n,\mu\nu} =& l^T_{n,\nu\mu} \quad \text{for} \quad \mu = 1, 2, \nu = 3 \\
L_{\mu\nu}(\eta) =& E_{H_0}(l_{1,\mu\nu}(1, 1, \eta)) \quad \text{for} \quad \mu = 1, 2, 3, \nu = 1, 2, 3
\end{aligned}$$

$$s(\theta, \eta) = l_{1,1}(1, 1, \eta) + \theta l_{1,2}(1, 1, \eta) - (L^T_{13}(\eta) + \theta L^T_{23}(\eta))L^{-1}_{33}(\eta) l_{1,3}(1, 1, \eta)$$

1.2 Algorithm

- Input: Y, G, $\theta^{(0)}$, θ_i, θ_j
- Output: $e_P(Z_{MERT}, Z_{\theta^{(0)}})$, $\tilde{e}_C(Z_{MERT}, Z_{\theta^{(0)}})$, $e_{HL}(Z_{MERT}, Z_{\theta^{(0)}})$, $e_B(Z_{MERT}, Z_{\theta^{(0)}})$

step 1 Estimate $\hat{\eta}$. where $\hat{\eta}$ satisfy $\partial l_n / \partial \eta|_{H_0, \hat{\eta}_n} = l_{n,3}(1, 1, \hat{\eta}_n) = 0$.

step 2 Compute $l_n(1, 1, \hat{\eta})$; $l_{n,\mu}(1, 1, \hat{\eta})$, for $\mu = 1, 2, 3$; $l_{n,\mu\nu}(1, 1, \hat{\eta})$ for $\mu = 1, 2, 3$, $\nu = 1, 2, 3$.

step 3 Compute $L_{\mu\nu}(\hat{\eta}) = E_{H_0}(l_{1,\mu\nu}(1, 1, \hat{\eta})) = \frac{1}{n} l_{n,\mu\nu}(1, 1, \hat{\eta})$

step 4 Compute $\sigma(\theta^{(0)})$, $\sigma(\theta_i)$, $\sigma(\theta_j), \sigma(\theta^{(0)},\theta_i), \sigma(\theta^{(0)},\theta_j), \sigma(\theta_i,\theta_j)$, where

$$\sigma(\theta_i,\theta_j) = A_{\hat\eta}\theta_i\theta_j + B_{\hat\eta}(\theta_i+\theta_j) + C_{\hat\eta}$$

$$\begin{aligned} A_\eta =& L_{23}(\eta)L_{33}^{-1}(\eta)L_{32}(\eta) - L_{22}(\eta) \\ B_\eta =& L_{13}(\eta)L_{33}^{-1}(\eta)L_{31}(\eta) - L_{12}(\eta) \\ C_\eta =& L_{13}(\eta)L_{33}^{-1}(\eta)L_{31}(\eta) - L_{11}(\eta) \end{aligned}$$

step 5 Compute $\mu(\lambda,\theta^{(0)}), \mu(\lambda,\theta_i), \mu(\lambda,\theta_j)$,

$$\mu(\lambda,\theta) = E_{H_1,\eta_0}(s(\theta,\eta_\theta)) = \frac{1}{n}(l_{n,11}(n,1,\hat\eta) + \theta l_{n,21}(n,1,\hat\eta) - (L_{13}^T(\hat\eta) + \theta L_{23}^T(\hat\eta))L_{33}^{-1}(\hat\eta)l_{n,31}(1,1,\hat\eta))$$

step 6 Compute $\mu^{(1)}(\lambda,\theta^{(0)}), \mu^{(1)}(\lambda,\theta_i), \mu^{(1)}(\lambda,\theta_j)$,

$$\begin{aligned} \mu^{(1)}(\lambda,\theta) =& \frac{1}{n}(l_{n,11}(n,1,\hat\eta) + \theta l_{n,21}(n,1,\hat\eta) - (L_{13}^T(\hat\eta) + \theta L_{23}^T(\hat\eta))L_{33}^{-1}(\hat\eta)l_{n,31}(1,1,\eta)) \\ =& L_{11}(\hat\eta) + \theta L_{21}(\hat\eta) - (L_{13}^T(\hat\eta) + \theta L_{23}^T(\hat\eta))L_{33}^{-1}(\hat\eta)L_{31}(\hat\eta)) \end{aligned}$$

step 7 Compute $e_P(Z_{MERT}, Z(\theta^{(0)}))$.

$$e_P(Z_{MERT}, Z(\theta^{(0)})) = \frac{(\rho_{\theta_i,\theta^{(0)}} + \rho_{\theta_j,\theta^{(0)}})^2}{2(1+\rho_{\theta_i,\theta_j})}$$

$$\rho_{\theta_i,\theta_j} = \frac{\sigma(\hat\eta,\theta_i,\theta_j)}{\sigma(\hat\eta,\theta_i,\theta_i)\sigma(\hat\eta,\theta_i,\theta_j)^{1/2}}$$

step 8 Compute

$$\tilde e_C(Z_{MERT}, Z_{\theta^{(0)}}) = \frac{\tilde Q_{Z_{MERT}}}{\tilde Q_{Z_{\theta^{(0)}}}}$$

$$\tilde Q_{Z_{\theta^{(0)}}} = 2\left(1 - \Phi\left(\frac{\mu^{(1)}(\lambda_0,\theta^{(0)})}{2\sigma(\theta^{(0)})}\right)\right)$$

$$\tilde Q_{Z_{MERT}} = 2\left(1 - \Phi\left(\left[\frac{\mu^{(1)}(\lambda_0,\theta_i)}{2\sigma(\theta_i)} + \frac{\mu^{(1)}(\lambda_0,\theta_j)}{2\sigma(\theta_j)}\right] / \sqrt{8(1+\rho_{\theta_i,\theta_j})}\right)\right)$$

$$\lambda_0 = 1$$

step 9 Compute

$$e_{HL}(Z_{MERT}, Z_\theta) = \frac{d_{Z_{MERT}}(\lambda)}{d_{Z_\theta(\lambda)}}$$

$$d_{Z_\theta}(\lambda) = \frac{\mu^2(\lambda,\theta)}{\sigma^2(\theta)}$$

$$d_{Z_{MERT}}(\lambda) = \mu^2_{MERT}(\lambda)$$

$$\mu_{MERT}(\lambda) = [\mu(\lambda,\theta_i)/\sigma(\theta_i) + \mu(\lambda,\theta_j)/\sigma(\theta_j)]/\sqrt{2(1+\rho_{\theta_i,\theta_j})}$$

step 10 Compute

$$e_B(Z_{MERT}, Z_\theta) = e_{HL}(Z_{MERT}, Z_\theta)$$

$$\tilde e_B(Z_{MERT}, Z_\theta) = \frac{\tilde c_{Z_{MERT}}}{\tilde c_{Z_\theta}}$$

$$\tilde c_{Z_{MERT}} = 1 - \Phi(\mu^{(1)}_{MERT}(1))$$

$$\tilde c_{Z_\theta} = 1 - \Phi(\mu^{(1)}(1,\theta^{(0)})/\sigma(\theta^{(0)}))$$

$$\mu^{(1)}_{MERT}(\lambda)) = [\mu^{(1)}(\lambda,\theta_i)/\sigma(\theta_i) + \mu^{(1)}(\lambda,\theta_j)/\sigma(\theta_j)]/\sqrt{2(1+\rho_{\theta_i,\theta_j})}$$

Table 1: The Four AREs of Z_{MERT} and $Z_{\theta^{(0)}}$

		$\lambda = 1.1$				$\lambda = 1.3$				$\lambda = 1.5$			
MAF	$\theta^{(0)}$	e_P	e_C	e_{HL}	e_B	e_P	e_C	e_{HL}	e_B	e_P	e_C	e_{HL}	e_B
0.15	0												
	1/4												
	1/2												
	1												
0.30	0												
	1/4												
	1/2												
	1												
0.45	0												
	1/4												
	1/2												
	1												

1.3 simulation

Let p be the minor allele frequency (MAF) of the marker of interest in the population. we consider case-control data with $r = 500$ cases and $s = 500$ controls. and $\lambda \in \{1.1, 1.2, 1.3\}$ and $p \in \{0.15, 0.30, 0.45\}$, and the true $\theta^{(0)} \in \{0, 1/4, 1/2, 1\}$. . we generate $Nrep = 1000$ datasets. and we compute the mean and variance of the four AREs to Z_{MERT} and $Z_{\theta^{(0)}}$