Statistics and Probability Letters 82 (2012) 2008–2016

Contents lists available at SciVerse ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Generalized Cordeiro–Ferrari Bartlett-type adjustment

Yoshihide Kakizawa

Faculty of Economics, Hokkaido University, Nishi 7, Kita 9, Kita-ku, Sapporo 060–0809, Japan

ARTICLE INFO

Article history:
Received 7 April 2012
Received in revised form 18 June 2012
Accepted 20 June 2012
Available online 26 June 2012

MSC:
62E20
62F03
62F05

Keywords:
Asymptotic expansion
Bartlett-type adjustment
Chi-squared approximation
Edgeworth expansion

ABSTRACT

The Bartlett-type adjustment is a higher-order asymptotic method for reducing the errors of the chi-squared approximations to the null distributions of various test statistics, which ensures that the resulting test has size $\alpha + o(N^{-1})$, where $0 < \alpha < 1$ is the significance level and N is the sample size. Recently, Kakizawa (2012) has revisited the Chandra–Mukerjee/Taniguchi adjustments in a unified way, since Chandra and Mukerjee (1991) and Taniguchi (1991b) originally considered the test of the simple null hypothesis, except for Mukerjee (1992). This paper considers a generalization of the adjustment due to Cordeiro and Ferrari (1991).

1. Introduction

The null distributions of the likelihood ratio, Rao's and Wald's test statistics have asymptotic expansions in powers of N^{-1}, where N is the sample size. The Bartlett/Bartlett-type adjustment is then designed to make the chi-squared approximation accurate up to order N^{-1}. Historically, it was first exploited by Bartlett (1937) in his classical test for homogeneity of variances. In general, a simple mean adjustment for the likelihood ratio test statistic $\mathrm{LR}^{(N)}$ through multiplication by a constant of the form $1 + b/N$ implies $P^{(N)}[(1 + b/N)\mathrm{LR}^{(N)} \le x] = \mathrm{pr}[\chi_f^2 \le x] + o(N^{-1})$ under the null hypothesis, where f is the number of restrictions under test; see Lawley (1956) and Hayakawa (1977). This fact became widely known as the Bartlett correctability of $\mathrm{LR}^{(N)}$, where substitution of a suitable consistent estimator for b is allowed. Among the vast literature, we further mention Taniguchi (1988), Bickel and Ghosh (1990), Jensen (1993) and Kakizawa (2011) for the theoretical issues.

Strangely, the test statistics $T^{(N)}$ other than $\mathrm{LR}^{(N)}$ are, in general, not Bartlett correctable. Thus, Chandra and Mukerjee (1991) and Taniguchi (1991b) first proposed the Bartlett-type adjustments on the basis of additional information of score or the maximum likelihood estimator (MLE), respectively. Cordeiro and Ferrari (1991) gave a kth order polynomial in $T^{(N)}$ without a constant term, where $k \in \mathbf{N}$ and the coefficients are determined according to an asymptotic expansion for the null distribution of $T^{(N)}$; see Kakizawa (1996) for the corresponding monotone version. Note that for the case $f = k = 1$, these proposals (Chandra and Mukerjee, 1991; Cordeiro and Ferrari, 1991; Taniguchi, 1991b) are essentially identical to the traditional multiplicative Bartlett adjustment. However, several attempts (Chandra and Mukerjee, 1991; Taniguchi, 1991b) and a unified treatment for the Chandra–Mukerjee/Taniguchi adjustments (Kakizawa, 2010) were restricted to the test of the simple null hypothesis. Extending Mukerjee (1992), who considered Rao's test statistic about a scalar parameter in the presence of a scalar nuisance parameter, Kakizawa (2012) has recently revisited the Bartlett-type adjustments for a class of

E-mail address: kakizawa@econ.hokudai.ac.jp.

0167-7152/$ – see front matter
doi:10.1016/j.spl.2012.06.022

test statistics in the framework of a composite hypothesis about a subvector of the parameters, where both the parameter of interest and the nuisance parameter are multidimensional.

Notice that the literature on the Cordeiro–Ferrari adjustment (CF adjustment for short) with $k = 2, 3$ is very extensive; some related papers during the last two decades are found in Kakizawa (2012). A contribution of this paper is to generalize the CF adjustment in a sense, by noting that it is not a unique form, as pointed out by Kakizawa (2011) for the traditional Bartlett adjustment $(1 + b/N)\mathrm{LR}^{(N)}$. Although, in this paper, we focus on the independent and identically distributed case for notational simplicity, we can construct the Bartlett-type adjustments even in a non-identical or dependent case where some regularity conditions are met for the log-likelihood derivatives according to the situations under consideration. So, the results are applicable in wide generality since they allow both the interest and nuisance parameters to be vector-valued, for which there is no assumption regarding the global parameter orthogonality.

2. Preliminaries

2.1. Notation

Let $P_\theta^{(N)}$ denote the $\boldsymbol{\theta}$-distribution of $\mathbf{X}_1, \ldots, \mathbf{X}_N$, which are independent and identically distributed random vectors, taking values of $\mathbf{R}^{d_X}$, according to a density $f(\mathbf{x}, \boldsymbol{\theta})$, $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, where $\boldsymbol{\Theta}$ is an open convex subset of $\mathbf{R}^p$. Throughout this paper, we assume the same regularity conditions as in Kakizawa (2012). The parameter vector $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_p)'$ is composed of two parts, a parameter of interest $\boldsymbol{\theta}_{(1)} = (\theta_1, \ldots, \theta_{p_1})'$ and a nuisance parameter $\boldsymbol{\theta}_{(2)} = (\theta_{p_1+1}, \ldots, \theta_{p_1+p_2})'$ with $p = p_1 + p_2$, where $\boldsymbol{\theta} = (\boldsymbol{\theta}'_{(1)}, \boldsymbol{\theta}'_{(2)})'$ and $\boldsymbol{\Theta} = \boldsymbol{\Theta}_{(1)} \times \boldsymbol{\Theta}_{(2)}$. The log-likelihood is denoted by $\mathcal{L}^{(N)}(\boldsymbol{\theta}) = \sum_{i=1}^N \log f(\mathbf{X}_i, \boldsymbol{\theta})$. We want to test a composite hypothesis $H_0 : \boldsymbol{\theta}_{(1)} = \boldsymbol{\theta}_{(1)0}$ against $H_1 : \boldsymbol{\theta}_{(1)} \neq \boldsymbol{\theta}_{(1)0}$, where $\boldsymbol{\theta}_{(1)0} \in \boldsymbol{\Theta}_{(1)}$ is specified while $\boldsymbol{\theta}_{(2)} \in \boldsymbol{\Theta}_{(2)}$ remains unspecified. Let $\widehat{\boldsymbol{\theta}}_{\mathrm{ML}}^{(N)} \in \boldsymbol{\Theta}$ be the MLE of $\boldsymbol{\theta}$, and let $\widetilde{\boldsymbol{\theta}}_{(2)\mathrm{ML}}^{(N)} \in \boldsymbol{\Theta}_{(2)}$ be the MLE of $\boldsymbol{\theta}_{(2)}$ under the constraint $\boldsymbol{\theta}_{(1)} = \boldsymbol{\theta}_{(1)0}$, where we write

$$\widetilde{\boldsymbol{\theta}}_{\mathrm{ML}}^{(N)} = \begin{pmatrix} \boldsymbol{\theta}_{(1)0} \\ \widetilde{\boldsymbol{\theta}}_{(2)\mathrm{ML}}^{(N)} \end{pmatrix}.$$

As usual, the Rth partial derivative of the log density $\log f(\mathbf{x}, \boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$ is denoted by

$$\ell_{j_1 \cdots j_R}(\mathbf{x}, \boldsymbol{\theta}) = \frac{\partial}{\partial \theta_{j_1}} \cdots \frac{\partial}{\partial \theta_{j_R}} \log f(\mathbf{x}, \boldsymbol{\theta}) \quad (R \in \mathbf{N};\ j_1, \ldots, j_R \in \{1, \ldots, p\}).$$

We introduce $I_R = j_1 \cdots j_R$ for notational simplicity and denote the cumulants of the $\ell_{I_R}(\mathbf{X}, \boldsymbol{\theta})$'s by

$$\nu_{I_{R_1}, \ldots, I_{R_v}}(\boldsymbol{\theta}) = \mathrm{cum}_\theta[\ell_{I_{R_1}}(\mathbf{X}, \boldsymbol{\theta}), \ldots, \ell_{I_{R_v}}(\mathbf{X}, \boldsymbol{\theta})],$$

where descending order $R_1 \geq \cdots \geq R_v \geq 1$ on the size $R_i = |I_{R_i}|$ is assumed, since $\nu_{I_{R_1}, \ldots, I_{R_v}}(\boldsymbol{\theta})$ is symmetric under permutation of $\{I_{R_1}, \ldots, I_{R_v}\}$. We assume that

$$\begin{aligned} &\nu_{j_1}(\boldsymbol{\theta}) = 0, \qquad \nu_{j_1 j_2}(\boldsymbol{\theta}) + \nu_{j_1, j_2}(\boldsymbol{\theta}) = 0, \qquad \nu_{j_1 j_2 j_3}(\boldsymbol{\theta}) + \langle 3 \rangle \nu_{j_1 j_2, j_3}(\boldsymbol{\theta}) + \nu_{j_1, j_2, j_3}(\boldsymbol{\theta}) = 0, \\ &\nu_{j_1 j_2 j_3 j_4}(\boldsymbol{\theta}) + \langle 4 \rangle \nu_{j_1 j_2 j_3, j_4}(\boldsymbol{\theta}) + \langle 3 \rangle \nu_{j_1 j_2, j_3 j_4}(\boldsymbol{\theta}) + \langle 6 \rangle \nu_{j_1 j_2, j_3, j_4}(\boldsymbol{\theta}) + \nu_{j_1, j_2, j_3, j_4}(\boldsymbol{\theta}) = 0 \end{aligned} \tag{1}$$

for all $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, where $\langle n \rangle$ before a term with indices is a sum of n similar terms obtained by index permutation. We make use of the Bartlett identities (1) to eliminate $\nu_{j_1 j_2}(\boldsymbol{\theta})$, $\nu_{j_1 j_2 j_3}(\boldsymbol{\theta})$ and $\nu_{j_1 j_2 j_3 j_4}(\boldsymbol{\theta})$ in subsequent calculations. Also, according to the partition $\boldsymbol{\theta} = (\boldsymbol{\theta}'_{(1)}, \boldsymbol{\theta}'_{(2)})'$, we stack the element

$$Z_j^{(N)}(\boldsymbol{\theta}) = \frac{1}{N^{1/2}} \sum_{i=1}^N \ell_j(\mathbf{X}_i, \boldsymbol{\theta}) \quad \text{and} \quad \nu_{j,k}(\boldsymbol{\theta}) = -\nu_{jk}(\boldsymbol{\theta})$$

as follows:

$$[Z_j^{(N)}(\boldsymbol{\theta})]_{j=1,\ldots,p} = \begin{pmatrix} \mathbf{Z}_{(1)}^{(N)}(\boldsymbol{\theta}) \\ \mathbf{Z}_{(2)}^{(N)}(\boldsymbol{\theta}) \end{pmatrix}, \qquad [\nu_{j,k}(\boldsymbol{\theta})]_{j,k \in \{1,\ldots,p\}} = \begin{pmatrix} \boldsymbol{\nu}_{(11)}(\boldsymbol{\theta}) & \boldsymbol{\nu}_{(12)}(\boldsymbol{\theta}) \\ \boldsymbol{\nu}_{(21)}(\boldsymbol{\theta}) & \boldsymbol{\nu}_{(22)}(\boldsymbol{\theta}) \end{pmatrix}.$$

They are the $p \times 1$ score vector $\mathbf{Z}^{(N)}(\boldsymbol{\theta})$ and the $p \times p$ Fisher information matrix $\boldsymbol{\nu}(\boldsymbol{\theta}) = \mathrm{var}_\theta^{(N)}[\mathbf{Z}^{(N)}(\boldsymbol{\theta})]$, respectively. It should be emphasized that we have no assumption regarding the so-called global parameter orthogonality, i.e., $\boldsymbol{\nu}_{(12)}(\boldsymbol{\theta}) \equiv \mathbf{O}_{p_1,p_2}$, where $\mathbf{O}_{p_1,p_2}$ is the $p_1 \times p_2$ zero matrix. Further, we write

$$Z_{j_1 \cdots j_R}^{(N)}(\boldsymbol{\theta}) = \frac{1}{N^{1/2}} \sum_{i=1}^N \{\ell_{j_1 \cdots j_R}(\mathbf{X}_i, \boldsymbol{\theta}) - \nu_{j_1 \cdots j_R}(\boldsymbol{\theta})\} \quad (R = 2, 3, \ldots).$$

Unless otherwise stated, we use the letters $\{j, k\}$ as indices of $\boldsymbol{\theta}$ that run from 1 to p, the letters $\{a, b\}$ as indices of $\boldsymbol{\theta}_{(1)}$ that run from 1 to p_1 and the letters $\{r, s\}$ as indices of $\boldsymbol{\theta}_{(2)}$ that run from $p_1 + 1$ to p. We adopt the Einstein summation