Statistics and Probability Letters 86 (2014) 38–43

Contents lists available at ScienceDirect

Statistics and Probability Letters

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

On Bartlett correctability of empirical likelihood in generalized power divergence family☆

Lorenzo Camponovo[a], Taisuke Otsu[b,*]

[a] University of St.Gallen, Bodanstrasse 6, 9000 St.Gallen, Switzerland

[b] London School of Economics, Houghton Street, London WC2A 2AE, UK

ARTICLE INFO

Article history:
Received 3 September 2013
Received in revised form 6 December 2013
Accepted 6 December 2013
Available online 25 December 2013

Keywords:
Empirical likelihood
Bartlett correction
Power divergence family

ABSTRACT

Baggerly (1998) showed that empirical likelihood is the only member in the Cressie–Read power divergence family to be Bartlett correctable. This paper strengthens Baggerly's result by showing that in a generalized class of the power divergence family, which includes the Cressie–Read family and other nonparametric likelihood such as Schennach's (2005, 2007) exponentially tilted empirical likelihood, empirical likelihood is still the only member to be Bartlett correctable.

1. Introduction

Since Owen (1988), empirical likelihood has been used as a device to construct nonparametric likelihood for numerous statistical problems and models as surveyed by Owen (2001). In spite of its nonparametric construction based on observed data points, empirical likelihood shares similar properties to parametric likelihood. For example, the empirical likelihood ratio statistic obeys the chi-squared limiting distribution, the so-called Wilks phenomenon. Another distinguishing feature of empirical likelihood is that it admits Bartlett correction, a second-order refinement based on a mean adjustment. This point was first made by DiCiccio et al. (1991) and extended to other contexts, such as quantiles (Chen and Hall, 1993), time series models (Kitamura, 1997; Monti, 1997), local linear smoothers Chen and Qin (2001), among others. Also Bartlett correctability has been studied for other constructions of nonparametric likelihood. Jing and Wood (1996) showed that exponential tilting (or empirical entropy) likelihood is not Bartlett correctable. Corcoran (1998) constructed some Bartlett correctable nonparametric likelihood based on a Taylor expansion of empirical likelihood. Baggerly (1998) strengthened Jing and Wood's (1996) result by showing that empirical likelihood is the only member in the Cressie and Read (1984) power divergence family to be Bartlett correctable.

The Cressie–Read type nonparametric likelihood is computed by choosing a tuning constant to define both the shape of the criterion function and the form of weights allocated to data points. Schennach (2005, 2007) suggested to choose different tuning constants for the shape of the criterion and the form of weights, and proposed a more general class of nonparametric likelihood. In particular, Schennach (2005) showed that exponentially tilted empirical likelihood (where the criterion is log-likelihood but the weights are computed by exponential tilting) can emerge as a valid likelihood function for Bayesian inference by a limiting argument. Also Schennach (2007) argued that when generalized estimating equations are misspecified, the point estimator based exponentially tilted empirical likelihood shows some robustness compared

☆ The authors would like to thank Elvezio Ronchetti for helpful comments.

* Corresponding author. Tel.: +44 2079557509.
E-mail addresses: lorenzo.camponovo@unisg.ch (L. Camponovo), t.otsu@lse.ac.uk (T. Otsu).

0167-7152/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.spl.2013.12.008

to the one based on empirical likelihood. Given this background, it is of interest to extend Baggerly's (1998) analysis to accommodate such new likelihood constructions and to study their Bartlett correctability.

In this paper, we confirm that in a generalized class of the power divergence family containing two tuning constants, empirical likelihood is still the only member to be Bartlett correctable. This result not only includes Baggerly's (1998) result as a special case, but also implies that Schennach's (2005, 2007) exponentially tilted empirical likelihood is not Bartlett correctable. Technically we follow a conventional approach based on the Edgeworth expansion (DiCiccio et al., 1991). We focus on characterizing the third and fourth order joint cumulants of the signed root of the test statistic based on the generalized power divergence family, and show that those cumulants vanish at sufficiently fast rates only when we employ the empirical likelihood statistic. We illustrate the theoretical findings by a small simulation study, which indicates that the empirical likelihood statistic with Bartlett correction has better coverage properties than other statistics.

2. Generalized power divergence family

We begin by introducing the generalized power divergence statistic. Consider a scalar random variable X drawn from an unknown distribution F_0 with mean μ_0. Following Owen (1988), the log-empirical likelihood ratio statistic for the mean is written as

$$\ell_{\mathrm{EL}}(\mu_0) = -2 \max_{p_1,\ldots,p_n} \sum_{i=1}^{n} \log(np_i), \quad \text{subject to } \sum_{i=1}^{n} p_i = 1, \ \sum_{i=1}^{n} p_i X_i = \mu_0.$$

It is known that under suitable regularity conditions the statistic $\ell_{\mathrm{EL}}(\mu_0)$ converges in distribution to the χ_1^2 distribution (Owen, 1988) and admits Bartlett correction, which yields a confidence interval with coverage error of order $O\left(n^{-2}\right)$ (DiCiccio et al., 1991).

Baggerly (1998) adapted the Cressie and Read (1984) power divergence family for goodness-of-fit to the present context and considered the test statistic in the form of

$$\ell_{\gamma}(\mu_0) = \min_{p_1,\ldots,p_n} \frac{2}{\gamma(\gamma+1)} \sum_{i=1}^{n} \left\{(np_i)^{\gamma+1} - 1\right\}, \quad \text{subject to } \sum_{i=1}^{n} p_i = 1, \ \sum_{i=1}^{n} p_i X_i = \mu_0, \tag{1}$$

if $\gamma \neq -1, 0$, otherwise $\ell_{-1}(\mu_0) = \min_{p_1,\ldots,p_n} -2\sum_{i=1}^{n} \log(np_i)$ and $\ell_0(\mu_0) = \min_{p_1,\ldots,p_n} 2n\sum_{i=1}^{n} p_i \log(np_i)$. Here $\gamma \in \mathbb{R}$ is a user-specified tuning constant. The empirical likelihood ratio statistic $\ell_{\mathrm{EL}}(\mu_0)$ corresponds to the case of $\gamma = -1$. The case of $\gamma = 0$ is often called the exponential tilting or empirical entropy statistic. Other popular choices for γ include the Neyman modified χ^2 ($\gamma = 1$), Hellinger or Freeman–Tukey ($\gamma = -\frac{1}{2}$), and Pearson's χ^2 ($\gamma = -2$). Baggerly (1998) showed that the power divergence statistic $\ell_{\gamma}(\mu_0)$ converges in distribution to the χ_1^2 distribution for any γ, and that $\ell_{\gamma}(\mu_0)$ is Bartlett correctable only for the case of $\gamma = -1$, the empirical likelihood ratio statistic. As Baggerly (1998) argued, a key insight of (lack of) Bartlett correctability is that the third and fourth order cumulants of the signed root of $\ell_{\gamma}(\mu_0)$ do not vanish at sufficiently fast rates when $\gamma \neq -1$.

From different perspectives, Schennach (2005, 2007) introduced the exponentially tilted empirical likelihood statistic

$$\ell_{\mathrm{ETEL}}(\mu_0) = -2 \sum_{i=1}^{n} \log\left(np_{\mathrm{ET},i}\right),$$

i.e., the criterion function is defined by $\ell_{\gamma}(\mu_0)$ with $\gamma = -1$, where $p_{\mathrm{ET},1}, \ldots, p_{\mathrm{ET},n}$ solve the minimization problem of $\ell_{\gamma}(\mu_0)$ with $\gamma = 0$,

$$\min_{p_1,\ldots,p_n} \sum_{i=1}^{n} p_i \log(np_i), \quad \text{subject to } \sum_{i=1}^{n} p_i = 1, \ \sum_{i=1}^{n} p_i X_i = \mu_0.$$

Schennach (2007) considered generalized estimating equations and studied asymptotic properties of a point estimator based on this statistic. Also Schennach (2005) argued that the function $\ell_{\mathrm{ETEL}}(\mu)$ can be interpreted as a valid likelihood function for Bayesian inference. It should be noted that the statistic $\ell_{\mathrm{ETEL}}(\mu_0)$ does not belong to the power divergence family (1). Therefore, Bartlett correctability of the statistic $\ell_{\mathrm{ETEL}}(\mu_0)$ is an open question.

In order to address this issue, we generalize the power divergence statistic as follows:

$$\ell_{\gamma,\phi}(\mu_0) = \frac{2}{\gamma(\gamma+1)} \sum_{i=1}^{n} \left\{\left(np_{\phi,i}\right)^{\gamma+1} - 1\right\}, \tag{2}$$

if $\gamma \neq -1, 0$, otherwise $\ell_{-1,\phi}(\mu_0) = \min_{p_1,\ldots,p_n} -2\sum_{i=1}^{n} \log\left(np_{\phi,i}\right)$ and $\ell_{0,\phi}(\mu_0) = \min_{p_1,\ldots,p_n} 2n\sum_{i=1}^{n} p_{\phi,i} \log\left(np_{\phi,i}\right)$, where $p_{\phi,1}, \ldots, p_{\phi,n}$ solve

$$\min_{p_1,\ldots,p_n} \ell_{\phi}(\mu_0), \quad \text{subject to } \sum_{i=1}^{n} p_i = 1, \ \sum_{i=1}^{n} p_i X_i = \mu_0. \tag{3}$$