The Annals of Statistics
1996, Vol. 24, No. 1, 365–369

EXPONENTIAL EMPIRICAL LIKELIHOOD IS NOT BARTLETT CORRECTABLE

BY BING-YI JING AND ANDREW T. A. WOOD

Hong Kong University of Science and Technology and Australian National University

In a recent paper, DiCiccio, Hall and Romano established that Owen's empirical likelihood is Bartlett correctable. This is an intriguing and perhaps surprising result and is the only nonparametric context in which Bartlett correctability is known to hold. An alternative, closely related nonparametric likelihood, referred to here as exponential empirical likelihood, may be constructed using Efron's method of nonparametric tilting. The purpose of this note is to show that exponential empirical likelihood is not Bartlett correctable.

1. Introduction. Suppose that L is a parametric likelihood function based on a sample of size n, and consider two hypotheses H_0 and H_1 of dimension d and $d + k$, respectively. If L_0 and L_1 denote the maximum of L under H_0 and H_1, then, quite generally, the maximised log-likelihood ratio $W = 2\log(L_1/L_0)$ satisfies

$$P(W > c_{\alpha,k}) = \alpha + O(n^{-1}), \tag{1.1}$$

where $c_{\alpha,k}$ is such that $P(\chi_k^2 > c_{\alpha,k}) = \alpha$. Bartlett correctability entails the following: there exists a scaling factor b, known as a Bartlett adjustment, such that

$$P(W/b > c_{\alpha,k}) = \alpha + O(n^{-2}). \tag{1.2}$$

When such a b exists, it can be taken to be $E(W)/k$ or a quantity which differs from $E(W)/k$ by $O_p(n^{-1/2})$.

Bartlett correctability, indicated in (1.2), is a remarkable property of parametric likelihood. This property of W is shared by few other types of test statistic. For further details of Bartlett correction, see Lawley (1956), Barndorff-Nielsen and Cox (1984), McCullagh and Cox (1986), McCullagh [(1987), Chapter 7] and Bickel and Ghosh (1990). For a derivation of (1.2), which is stronger than the usual statement of the Bartlett correction property, see Barndorff-Nielsen and Hall (1988).

In a recent paper, DiCiccio, Hall and Romano (1988, 1991) showed that Owen's (1988, 1990) empirical likelihood is Bartlett correctable in the "smooth function of means" model. This is an interesting result, and it is perhaps surprising that Bartlett correctability, which is a very delicate property,

Received March 1994; revised April 1995.
AMS 1991 *subject classifications*. 62G20, 62E20.
Key words and phrases. Bartlett adjustment, nonparametric likelihood, tilting.

holds in a nonparametric setting. See also Hall and LaScala (1990) for discussion of (1.1) and (1.2) in the empirical likelihood context.

It will be sufficient here to consider empirical likelihood for a scalar mean. Suppose that $X_1, \dots, X_n$ is a sample of independent and identically distributed random variables, and let $p_1, \dots, p_n$ be a probability vector which assigns probability mass p_i to observation i. Then the empirical log-likelihood ratio W_μ, evaluated at a candidate mean value μ using Lagrange multiplier arguments, is given by

$$(1.3)\qquad W_\mu \equiv \sup_{\sum p_i X_i = \mu} 2 \sum_{i=1}^{n} \log\{(np_i)^{-1}\} = 2 \sum_{i=1}^{n} \log\{1 + t(X_i - \mu)\},$$

where t is the solution to

$$(1.4)\qquad \sum_{i=1}^{n} n^{-1}\{1 + t(X_i - \mu)\}^{-1}(X_i - \mu) = 0.$$

An alternative nonparametric likelihood, referred to here as exponential empirical likelihood, may be constructed using Efron's (1981, 1982) method of nonparametric tilting. In effect, we restrict the p_i to be of the form

$$p_i(\tilde{t}) = \frac{\exp(\tilde{t} X_i)}{\sum_{j=1}^{n} \exp(\tilde{t} X_j)};$$

and the exponential empirical log-likelihood ratio is then given by

$$(1.5)\qquad \tilde{W}_\mu = 2 \sum_{i=1}^{n} \log\{np_i(\tilde{t})\},$$

where $\tilde{t}$ satisfies

$$(1.6)\qquad \sum_{i=1}^{n} p_i(\tilde{t})(X_i - \mu) = 0.$$

As noted by DiCiccio and Romano (1990), the empirical likelihood (1.3) minimises forward Kullback–Leibler distance, while the exponential empirical likelihood (1.5) minimises backward Kullback–Leibler distance, where in both cases the minimisation is over those discrete distributions with mean μ which are concentrated on the data points $X_1, \dots, X_n$. Exponential empirical likelihood is also discussed in Davison, Hinkley and Worton [(1992), Section 4].

The purpose of this note is to establish that exponential empirical likelihood is not Bartlett correctable. We do this by comparing the relevant expansions for empirical likelihood and exponential empirical likelihood.

Tom DiCiccio and a referee have pointed out that the result given here is contained implicitly in the work of DiCiccio and Romano. More specifically, it follows from the formula below (3.5) in DiCiccio and Romano (1990), combined with the argument outlined in Section 2.3 of DiCiccio, Hall and Romano (1991).

Given a sequence $X, X_1, X_2 \dots$ of independent and identically distributed random variables, we shall write

$$\alpha_r = E(X^r), \qquad A_r = n^{-1} \sum_{i=1}^{n} X_i^r - \alpha_r, \qquad r = 1, 2, \dots .$$

Without loss of generality, it will be assumed that $\alpha_1 = 0$ and $\alpha_2 = 1$. Thus, it will be appropriate to restrict attention to $\mu = 0$ in (1.3)–(1.6).

2. Bartlett correctability.

2.1. *Expansions for empirical likelihood.* In the process of establishing Bartlett correctability of empirical likelihood, DiCiccio, Hall and Romano (1988, 1991) obtained expansions for quantities such as the empirical likelihood ratio W_0. The relevant expansions, specialised to the case of a scalar mean, are summarised below.

First, we give an expansion for t, the solution to (1.4) with μ set to zero. The expansion may be written $t = t_1 + t_2 + t_3 + O_p(n^{-2})$, where

$$\begin{aligned} t_1 &= A_1, \qquad t_2 = \alpha_3 A_1^2 - A_1 A_2, \\ t_3 &= A_1 A_2^2 + A_1^2 A_3 + 2\alpha_3^2 A_1^3 - 3\alpha_3 A_1^2 A_2 - \alpha_4 A_1^3. \end{aligned} \tag{2.1}$$

Observe that t_i is $O_p(n^{-i/2})$ for $i = 1, 2, 3$.

Next, we Taylor-expand $n^{-1}W_0$ as a function of t, ignoring terms which are $O_p(n^{-5/2})$ or smaller, and then substitute for t using (2.1). The resulting expansion for $n^{-1}W_0$, given by DiCiccio, Hall and Romano [(1991), formula 4.1], is

$$\begin{aligned} n^{-1}W_0 = {} & A_1^2 - A_1^2 A_2 + \tfrac{2}{3}\alpha_3 A_1^3 + A_1^2 A_2^2 + \tfrac{2}{3} A_1^3 A_3 + \alpha_3^2 A_1^4 \\ & - 2\alpha_3 A_1^3 A_2 - \tfrac{1}{2}\alpha_4 A_1^4 + O_p(n^{-5/2}). \end{aligned} \tag{2.2}$$

Then we obtain a quantity R which satisfies

$$\begin{aligned} n^{-1}W_0 &= R^2 + O_p(n^{-5/2}), \qquad R = R_1 + R_2 + R_3, \\ R_j &= O_p(n^{-j/2}), \qquad j = 1, 2, 3. \end{aligned} \tag{2.3}$$

It turns out that we may choose

$$\begin{aligned} R_1 &= A_1, \qquad R_2 = \tfrac{1}{3}\alpha_3 A_1^2 - \tfrac{1}{2} A_1 A_2, \\ R_3 &= \tfrac{3}{8} A_1 A_2^2 + \tfrac{4}{9}\alpha_3^2 A_1^3 - \tfrac{5}{6}\alpha_3 A_1^2 A_2 + \tfrac{1}{3} A_1^2 A_3 - \tfrac{1}{4}\alpha_4 A_1^3. \end{aligned} \tag{2.4}$$

2.2. *Expansions for exponential empirical likelihood.* The corresponding expansions for exponential empirical likelihood in the case of a scalar mean are now given. Formally, the derivations are very similar.

The expansion for $\tilde{t}$, the solution to $\sum \exp(\tilde{t}X_i)X_i = 0$ obtained from (1.6), has the form $\tilde{t} = \tilde{t}_1 + \tilde{t}_2 + \tilde{t}_3 + O_p(n^{-2})$, where

$$\tilde{t}_1 = -A_1, \qquad \tilde{t}_2 = A_1 A_2 - \tfrac{1}{2}\alpha_3 A_1^2$$
$$\tilde{t}_3 = \tfrac{3}{2}\alpha_3 A_1^2 A_2 + \tfrac{1}{6}\alpha_4 A_1^3 - A_1 A_2^2 - \tfrac{1}{2}\alpha_3^2 A_1^3 - \tfrac{1}{2}A_1^2 A_3.$$

The exponential empirical log-likelihood ratio has an expansion

$$n^{-1}\tilde{W}_0 = A_1^2 - A_1^2 A_2 + \tfrac{2}{3}\alpha_3 A_1^3 + A_1^2 A_2^2 + \tfrac{2}{3}A_1^3 A_3 + \tfrac{3}{4}\alpha_3^2 A_1^4 - 2\alpha_3 A_1^3 A_2$$
$$- \tfrac{1}{4}\alpha_4 A_1^4 + O_p(n^{-5/2}),$$

and the quantity $\tilde{R} = \tilde{R}_1 + \tilde{R}_2 + \tilde{R}_3$, corresponding to $R = R_1 + R_2 + R_3$ in (2.4), is given by

$$\tilde{R}_1 = R_1, \qquad \tilde{R}_2 = R_2, \qquad \tilde{R}_3 = R_3 + \tfrac{1}{8}(\alpha_4 - \alpha_3^2)A_1^3. \tag{2.5}$$

2.3. *Cumulants of R and $\tilde{R}$.* Given that mild smoothness conditions are satisfied, Bartlett correctability of empirical likelihood is equivalent to the third and fourth cumulants of R satisfying the following:

$$\kappa_3(R) = O(n^{-3}), \qquad \kappa_4(R) = O(n^{-4}) \tag{2.6}$$

[see, e.g., DiCiccio, Hall and Romano (1988, 1991) and McCullagh (1987), pages 214–215]. The checking of (2.6) in the general case of a smooth function of means involves some very heavy calculation [see DiCiccio, Hall and Romano (1988, 1991)].

Using the formulae connecting moments and cumulants [e.g., McCullagh (1987), page 34] and taking into account the order of magnitude of the moments of R_1, R_2 and R_3, it is found that

$$\kappa_3(R) = E(R_1^3) + 3E(R_1^2 R_2) - 3E(R_1^2)E(R_2) + O(n^{-3}) \tag{2.7}$$

and

$$\begin{aligned}\kappa_4(R) &= 4E(R_1^3 R_3) - 12E(R_1^2)E(R_1 R_3)\\ &\quad + (\text{terms involving moments of } R_1, R_2 \text{ only}) + O(n^{-4}).\end{aligned} \tag{2.8}$$

It follows from (2.5)–(2.8) that $\kappa_3(\tilde{R}) = O(n^{-3})$ and

$$\begin{aligned}\kappa_4(\tilde{R}) &= \frac{1}{2}(\alpha_4 - \alpha_3^2)\{E(A_1^6) - 3E(A_1^2)E(A_1^4)\} + O(n^{-4})\\ &= \frac{1}{2n^3}(\alpha_4 - \alpha_3^2)(15\sigma^6 - 3 \times 3\sigma^6) + O(n^{-4})\\ &= 3n^{-3}(\alpha_4 - \alpha_3^2) + O(n^{-4}),\end{aligned} \tag{2.9}$$

where $\sigma^2 \equiv \operatorname{Var}(X) = 1$ by assumption. Since it is necessary for $\kappa_4(\tilde{R})$ to be $O(n^{-4})$ for Bartlett correctability to hold, and since $\alpha_4 > \alpha_3^2$ for all continuous distributions with zero mean and unit variance, we may conclude from (2.9) that exponential empirical likelihood is not Bartlett correctable.

REFERENCES

BARNDORFF-NIELSEN, O. E. and COX, D. R. (1984). Bartlett adjustments to the likelihood ratio statistic and the distribution of the maximum likelihood estimator. *J. Roy. Statist. Soc. Ser. B* **46** 483–495.

BARNDORFF-NIELSEN, O. E. and HALL, P. (1988). On the level error after Bartlett adjustment of the likelihood ratio statistic. *Biometrika* **75** 374–378.

BICKEL, P. J. and GHOSH, J. K. (1990). A decomposition for the likelihood ratio statistic and the Bartlett correction—a Bayesian argument. *Ann. Statist.* **18** 1070–1090.

DAVISON, A. C., HINKLEY, D. V. and WORTON, B. J. (1992). Bootstrap likelihoods. *Biometrika* **79** 113–130.

DICICCIO, T., HALL, P. and ROMANO, J. (1988). Bartlett adjustment for empirical likelihood. Technical Report 298, Dept. Statistics, Stanford Univ.

DICICCIO, T., HALL, P. and ROMANO, J. (1991). Empirical likelihood is Bartlett correctable. *Ann. Statist.* **19** 1053–1061.

DICICCIO, T. and ROMANO, J. (1990). Nonparametric confidence limits by resampling methods and least favorable families. *Internat. Statist. Rev.* **58** 59–76.

EFRON, B. (1981). Nonparametric standard errors and confidence intervals (with discussion). *Canad. J. Statist.* **9** 139–172.

EFRON, B. (1982). *The Jackknife, the Bootstrap and Other Resampling Plans*. SIAM, Philadelphia.

HALL, P. and LASCALA, B. (1990). Methodology and algorithms of empirical likelihood. *Internat. Statist. Rev.* **58** 109–127.

LAWLEY, D. N. (1956). A general method for approximating to the distribution of likelihood-ratio criteria. *Biometrika* **43** 295–303.

MCCULLAGH, P. (1987). *Tensor Methods in Statistics*. Chapman and Hall, London.

MCCULLAGH, P. and COX, D. R. (1986). Invariants and likelihood ratio statistics. *Ann. Statist.* **14** 1419–1430.

OWEN, A. (1988). Empirical likelihood ratio confidence intervals for a single functional. *Biometrika* **75** 237–249.

OWEN, A. (1990). Empirical likelihood ratio confidence regions. *Ann. Statist.* **18** 90–120.

DEPARTMENT OF MATHEMATICS
HONG KONG UNIVERSITY
OF SCIENCE AND TECHNOLOGY
CLEAR WATER BAY, KOWLOON
HONG KONG

CENTRE FOR MATHEMATICS AND ITS APPLICATIONS
AUSTRALIAN NATIONAL UNIVERSITY
CANBERRA, ACT 0200
AUSTRALIA