AISM (2006) 58: 427–440
DOI 10.1007/s10463-006-0040-1

In Hong Chang · Rahul Mukerjee

Asymptotic results on a general class of empirical statistics: power and confidence interval properties

Received: 10 December 2004 / Revised: 24 May 2005 / Published online: 11 May 2006

Abstract We consider a very general class of empirical statistics that includes (a) empirical discrepancy (ED) statistics, (b) generalized empirical exponential family likelihood statistics, (c) generalized empirical likelihood statistics, (d) empirical statistics arising from Bayesian considerations, and (e) Bartlett-type adjusted versions of ED statistics. With reference to this general class, we investigate higher order asymptotics on power and expected lengths of confidence intervals. For (b)-(e), such results have been hitherto unexplored. Furthermore, our findings help in understanding the presently known results on the subclass (a) from a wider perspective.

Keywords Average power · Bartlett-type adjustment · Confidence interval · Contiguous alternatives · Edgeworth expansion · Empirical likelihood · Minimaxity · Second-order · Third-order

1 Introduction

Ever since its introduction by Owen (1988), empirical likelihood has been of significant interest in the statistics and econometrics literature; see Owen (2001) and Mittelhammer et al. (2000) for accounts of recent developments. Corcoran (1998) introduced a general class of empirical discrepancy (ED) statistics. This class includes the Cressie-Read discrepancy statistics (Baggerly, 1998) and, in particular, the empirical likelihood ratio (ELR) statistic. In a pioneering work, Bravo (2003)

I. H. Chang
Department of Computer Science and Statistics, Chosun University, Gwangju 501–759, South Korea

R. Mukerjee (✉)
Indian Institute of Management Calcutta, Joka, Diamond Harbour Road, Kolkata 700 104, India
E-mail: rmuk1@hotmail.com

reported illuminating results on second order power of Cressie-Read discrepancy statistics under contiguous alternatives. This was followed up by Bravo (2005), Fang and Mukerjee (2005a) and Mukerjee (2005), who worked with the more general class of ED statistics (Corcoran, 1998) and reported results on Bartlett-type adjustments, expected lengths of confidence intervals, and third-order power, respectively.

While the picture for ED statistics is now fairly clear in view of the aforesaid results, there are other important classes of empirical statistics for which higher order asymptotic results on power or confidence interval properties have not yet been explored. Notable among these are statistics arising from generalized empirical exponential family likelihood (Corcoran, 1998, Sect. 4), generalized empirical likelihood (Newey and Smith, 2004), and empirical-type likelihoods that admit a probability matching prior (Fang and Mukerjee, 2005b). Furthermore, the Bartlett-type adjusted versions of ED statistics (Bravo, 2005) are not themselves ED statistics and the issue of a theoretical understanding of their power and confidence interval properties remains open.

In an attempt to fill up this gap to some extent, in the present article we consider a very general class of statistics (see Sect. 2) that covers not only the ED statistics but also the ones mentioned above as subclasses. Higher order asymptotic results on power and expected lengths of confidence intervals are obtained in Sects. 3 and 4, respectively; the implications of the results on second- and third-order power are discussed in Subsects. 3.4 and 3.5. These results are not only more comprehensive than the existing ones on ED statistics, but also help in understanding the findings for the ED subclass from a wider perspective; see Subsect. 3.4. Compared to ED statistics, a new feature with our general class is that the relevant approximate cumulants of fifth and higher orders are not necessarily ignorable. This calls for the development of an indirect technique, akin to that in comparing parametric likelihood-based tests (Mukerjee, 1990a). Another new feature is that the critical value may now involve an additional lower order term that requires careful handling.

In order to give a flavor of the main ideas without making the notation and algebra too heavy, the case of univariate observations is considered here. However, the univariate results can provide useful pointers to what is likely to happen in the multivariate case; see Subsect. 3.5 and Sect. 4. Indeed, at the expense of more excruciating algebra, it should be possible to extend these results to the multivariate case. It is anticipated that this will call for combining the present techniques with those in Mukerjee (2005).

2 A general class of empirical statistics

Let $X_1, \ldots, X_n$ be independent scalar-valued random variables from an unknown common distribution with mean θ. We work under the same conditions as in Bravo (2003); these conditions justify the Edgeworth expansions considered later. Let $\bar{X} = n^{-1}\sum_{i=1}^{n} X_i$, $m_s = n^{-1}\sum_{i=1}^{n}(X_i - \bar{X})^s (s = 2, 3, \ldots)$, $g_3 = m_3/m_2^{3/2}$, $g_4 = m_4/m_2^2$, and $y \equiv y(\theta) = (n/m_2)^{1/2}(\bar{X} - \theta)$.

We consider a very general class of empirical statistics such that any statistic in the class admits a stochastic expansion of the form

$$T(\theta) = \{W(\theta)\}^2 + o_p\left(n^{-1}\right), \tag{1}$$

where

$$W(\theta) = y + n^{-1/2}U(g_3, y) + n^{-1}V(g_3, g_4, y), \tag{2}$$

and

$$U(g_3, y) = \sum\sum c_{\text{sq}} g_3^s y^q,\ V(g_3, g_4, y) = \sum\sum\sum d_{\text{stq}} g_3^s g_4^t y^q, \tag{3}$$

are polynomials in g_3, y and g_3, g_4, y, respectively, the coefficients c_{sq} and d_{stq} therein being $O(1)$ constants. The specific forms of these polynomials depend on the particular statistic under consideration.

Remark 1 As hinted earlier, the above class includes:

(a) ED statistics,
(b) generalized empirical exponential family likelihood statistics (Corcoran, 1998, Sect. 4; see also Schennach, 2005),
(c) generalized empirical likelihood statistics (Newey and Smith, 2004),
(d) statistics arising from empirical-type likelihoods that admit a probability matching prior (Fang and Mukerjee, 2005b), and
(e) Bartlett-type adjusted versions of ED statistics, obtained by Bravo (2005) in the spirit of Cordeiro and Ferrari's (1991) work in the parametric case.

For (a)–(c), it can be shown that

$$U(g_3, y) = cg_3y^2, \tag{4}$$

where c is a $O(1)$ constant. As noted in Corcoran (1998), any statistic in subclass (a) is defined via a constrained optimization. Fang and Mukerjee (2005a) established (4) for these statistics by expressing, with margin of error $o_p(n^{-1})$, the associated Lagrangian multipliers as polynomials in y, g_3 and g_4; note that our y is the same as $(-y)$ in their notation. In the same spirit, (4) can be established for the generalized empirical likelihood statistics in (c) by finding a similar solution to an optimization problem underlying the definition of any of these statistics. With a lengthy algebra, the same approach shows that (4) holds also for the subclass (b), with $c = 1/3$. Following Fang and Mukerjee (2005a), c in (4) equals 1/3 for the ELR statistic as well. Thus all the statistics in the subclass (b) as well as the ELR statistic have the same $U(.)$, though, of course, they have different $V(.)$. For the subclass (d),

$$U(g_3, y) = g_3\left(\frac{1}{3}y^2 - \frac{1}{2}\right). \tag{5}$$

Following Fang and Mukerjee (2005b), if one works with an empirical-type likelihood of the form $L(\theta) \propto \exp\{-T(\theta)/2\}$, where $T(\theta)$ is given by (1–3) with $U(.)$ as in (5), then the posterior quantiles of θ under the flat prior (i.e., the Lebesgue measure) have approximate frequentist validity. Finally, for (e), the Bartlett-type adjusted version of any ED statistic has the same $U(.)$ but possibly different $V(.)$ compared to the original statistic, because the adjustment term is a polynomial in the original ED statistic, the coefficients therein being at most of order $O_p(n^{-1})$.

Remark 2 In contrast with the stochastic expansions considered in Baggerly (1998), Bravo (2003) or Mukerjee (2005), the one in (1) does not involve any unknown population moment other than θ. This helps not only in finding an asymptotic representation for the associated confidence interval (cf. Fang and Mukerjee, 2005a) but also in understanding the power properties in a more transparent manner (see Subsect. 3.5).

3 Power properties

3.1 Preliminaries

Suppose interest lies in the null hypothesis $H_0 : \theta = \theta_0$. Let σ be the unknown population standard deviation under θ_0, and we consider contiguous alternatives of the form $H_n : \theta = \theta_n$, where $\theta_n = \theta_0 + n^{-1/2}\gamma\sigma$ and γ is free from n. As in Bravo (2003), we note that in the present context of nonparametric inference, the population distribution functions under the null and contiguous alternative hypotheses are related in the sense that they are both assumed to belong to the same class of distributions, indexed by the mean θ. Thus the distribution of $X_i - \theta_n$, under θ_n, is the same as that of $X_i - \theta_0$, under θ_0 (the difference with a parametric location model is that the form of the distribution is now unknown). Consequently, defining $Z_i = (X_i - \theta_n)/\sigma$, the standardized moments $\beta_s = E_{\theta_n}(Z_i^s)$ $(s = 3, 4)$ do not depend on γ.

Let z be the upper $(\alpha/2)$-point of a standard normal variate. In view of (1), we consider a critical region of the form

$$|W(\theta_0)| > z + n^{-1/2}a(g_3, z) + n^{-1}b(g_3, g_4, z), \tag{6}$$

where $a(.)$ and $b(.)$ are polynomials in g_3, z and g_3, g_4, z, respectively, the coefficients therein being $O(1)$ constants. These polynomials are to be so chosen that the critical region has size $\alpha + o(n^{-1})$. The right-hand side of (6) is motivated by the fact that the $(1-\alpha)$ th quantile of $|W(\theta_0)|$, under H_0 and with margin of error $o(n^{-1})$, has a similar structure with g_3, g_4 replaced by β_3, β_4. As will be seen later, $a(.)$ can be nonzero for our general class of statistics, a fact that may be contrasted with what happens for ED statistics (Bravo, 2003). Under contiguous alternatives, the power function corresponding to (6) is given by

$$P(\gamma) = P_{\theta_n}(M_1 > z) + P_{\theta_n}(M_2 < -z) + o\left(n^{-1}\right), \tag{7}$$

where

$$M_1 = W(\theta_0) - n^{-1/2}a(g_3, z) - n^{-1}b(g_3, g_4, z), \tag{8}$$

$$M_2 = W(\theta_0) + n^{-1/2}a(g_3, z) + n^{-1}b(g_3, g_4, z). \tag{9}$$

3.2 Approximate characteristic functions

In view of (7), we now consider the approximate characteristic functions (cf) of M_1 and M_2 under θ_n. Some additional notation is required for this purpose. Let $U'(g_3, y)$ and $a'(g_3, z)$ be the first partial derivatives of $U(g_3, y)$ and $a(g_3, z)$ with respect to g_3, e.g., by (3),

$$U'(g_3, y) = \sum\sum c_{\text{sq}} s g_3^{s-1} y^q. \tag{10}$$

Also, note that the square of $U(g_3, y)$ is again a polynomial in g_3, y, and hence write

$$\{U(g_3, y)\}^2 = \sum\sum c^*_{\text{sq}} g_3^s y^q, \tag{11}$$

the coefficients c^*_{sq} being $O(1)$ constants. Let $K_j(\gamma)$ be the jth raw moment, about zero, of the univariate normal distribution with mean γ and variance unity. Define

$$h_{01} = \frac{1}{6}\left(\gamma^2 - 1\right), \quad h_{11} = \frac{1}{6}\gamma, \quad h_{21} = -\frac{1}{3}, \tag{12}$$

$$h_{02} = -\frac{1}{72}\left(\gamma^5 - 10\gamma^3 + 15\gamma\right)\beta_3^2 - \frac{1}{24}\left(\gamma^3 - 3\gamma\right)(\beta_4 - 3), \tag{13}$$

$$\begin{aligned}
h_{12} &= -\frac{1}{72}\left(\gamma^4 - 15\gamma^2 + 12\right)\beta_3^2 - \frac{1}{8}\gamma^2(\beta_4 - 3) + \frac{1}{4}\beta_4, \\
h_{22} &= \frac{1}{72}\left(5\gamma^3 - 33\gamma\right)\beta_3^2 + \gamma\left(\frac{1}{4}\beta_4 - 1\right), \\
h_{32} &= \frac{1}{72}\left(\gamma^2 + 8\right)\beta_3^2 - \frac{1}{12}\beta_4 + \frac{1}{2}, h_{42} = -\frac{1}{9}\gamma\beta_3^2, h_{52} = \frac{1}{18}\beta_3^2, \\
h_{03} &= -\frac{1}{6}\gamma^3,\ h_{13} = \frac{1}{2}, h_{23} = \frac{1}{2}\gamma,\ h_{33} = -\frac{1}{3}.
\end{aligned} \tag{14}$$

Let $\xi = (-1)^{1/2}\tau$, where τ is an auxiliary variate. Then, an argument sketched in the appendix shows that the approximate cf of M_1, under θ_n, is given by

$$E_{\theta_n}\{\exp(M_1\xi)\} = \left\{1 + \xi\left(n^{-1/2}\Delta_1 + n^{-1}\Delta_2\right)\right\}\exp\left(\gamma\xi + \frac{1}{2}\xi^2\right) + o\left(n^{-1}\right), \tag{15}$$

where

$$\Delta_1 = \beta_3 \sum_{j=0}^{2} h_{j1} K_j(\gamma + \xi) + \sum\sum c_{\text{sq}} \beta_3^s K_q(\gamma + \xi) - a(\beta_3, z), \tag{16}$$

$$\Delta_2 = \sum_{j=0}^{5} h_{j2} K_j (\gamma + \xi) + \sum\sum\sum d_{\mathrm{stq}} \beta_3^s \beta_4^t K_q (\gamma + \xi) - b(\beta_3, \beta_4, z)$$

$$+\beta_3 \sum_{j=0}^{3} h_{j3} \left\{ \sum\sum c_{\mathrm{sq}} \beta_3^s K_{q+j} (\gamma + \xi) - a(\beta_3, z) K_j (\gamma + \xi) \right\}$$

$$+ \left(\beta_4 - 3 - \frac{3}{2} \beta_3^2 \right) \left\{ \sum\sum c_{\mathrm{sq}} s \beta_3^{s-1} \left(K_{q+1} (\gamma + \xi) \right. \right.$$

$$\left. \left. - K_q (\gamma + \xi) \gamma \right) - a' (\beta_3, z) \xi \right\}$$

$$+ \frac{1}{2} \xi \sum\sum c_{\mathrm{sq}}^* \beta_3^s K_q (\gamma + \xi) - a(\beta_3, z) \xi \sum\sum c_{\mathrm{sq}} \beta_3^s K_q (\gamma + \xi)$$

$$+ \frac{1}{2} \xi \{ a(\beta_3, z) \}^2, \tag{17}$$

and the c_{sq}, d_{stq} and c_{sq}^* are as in (3) and (11). Because of (8) and (9), the approximate cf of M_2, under θ_n, can be obtained from (15–17) replacing $a(.)$ and $b(.)$ by their negatives.

Remark 3 The presence of terms like $K_q(\gamma + \xi)$ and $K_{q+j}(\gamma + \xi)$ in (16) and (17) shows that the fifth and higher order approximate cumulants of M_1 and M_2 may not be ignorable even when one works with margin of error $o(n^{-1})$. This is in contrast with what happens for the subclass of ED statistics.

3.3 Power under contiguous alternatives

Equations (15–17) yield an Edgeworth expansion for M_1, under θ_n, in the usual manner. Even though (16) and (17) look formidable, the fact that (Mukerjee, 1990a)

$$K_j \left(\gamma - \frac{\mathrm{d}}{\mathrm{d}u} \right) \phi(u - \gamma) = u^j \phi(u - \gamma),$$

where $\phi(.)$ is the standard univariate normal density, helps in expressing the Edgeworth expansion in a neat and readily integrable form. The same happens for M_2. These Edgeworth expansions, together with (7) and (12–14) yield an expansion for $P(\gamma)$ with margin of error $o(n^{-1})$. In view of (16), (17) and their counterparts for M_2, the expansion for $P(\gamma)$ involves the polynomials $a(.)$ and $b(.)$ that appear in the critical value in (6). One can check that the size condition $P(0) = \alpha + o(n^{-1})$ holds provided

$$a(g_3, z) = \frac{1}{2} \{ U(g_3, z) - U(g_3, -z) \}, \tag{18}$$

$$b\left(g_3,g_4,z\right)=\frac{1}{18}g_3^2z^5+\left(\frac{1}{9}g_3^2-\frac{1}{12}g_4+\frac{1}{2}\right)z^3+\left(\frac{1}{4}g_4-\frac{1}{6}g_3^2\right)z$$
$$+\frac{1}{2}\left\{V\left(g_3,g_4,z\right)-V\left(g_3,g_4,-z\right)\right\}+\frac{1}{8}z\left\{U\left(g_3,z\right)\right.$$
$$\left.+U\left(g_3,-z\right)\right\}^2+\frac{1}{2}z\left(g_4-3-\frac{3}{2}g_3^2\right)\left\{U'\left(g_3,z\right)\right.$$
$$\left.+U'\left(g_3,-z\right)\right\}+\left\{U\left(g_3,z\right)+U\left(g_3,-z\right)\right\}$$
$$\times\left\{g_3\left(\frac{1}{4}z-\frac{1}{6}z^3\right)-\frac{1}{4}\tilde{U}\left(g_3,z\right)+\frac{1}{4}\tilde{U}\left(g_3,-z\right)\right\},\quad (19)$$

where $\tilde{U}(g_3,y)$ is the first partial derivative of $U(g_3,y)$ with respect to y. With $a(.)$ and $b(.)$ as in (18) and (19), the expansion for the power function $P(\gamma)$, under contiguous alternatives, finally simplifies to

$$P\left(\gamma\right)=P^{(0)}\left(\gamma\right)+n^{-1/2}P^{(1)}\left(\gamma\right)+n^{-1}P^{(2)}\left(\gamma\right)+o\left(n^{-1}\right),\quad (20)$$

where

$$P^{(0)}\left(\gamma\right)=2-\Phi\left(z-\gamma\right)-\Phi\left(z+\gamma\right),$$

$$P^{(1)}\left(\gamma\right)=\phi\left(z-\gamma\right)\left\{\frac{1}{6}\gamma\beta_3z+C^{(1)}\left(\gamma\right)\right\}+\phi\left(z+\gamma\right)\left\{\frac{1}{6}\gamma\beta_3z-C^{(1)}\left(\gamma\right)\right\},\quad (21)$$

$$P^{(2)}\left(\gamma\right)=\phi\left(z-\gamma\right)\left\{C^{(2)}\left(\gamma\right)+C\right\}+\phi\left(z+\gamma\right)\left\{C^{(2)}\left(-\gamma\right)-C\right\},\quad (22)$$

$\Phi(.)$ is the standard univariate normal distribution function, and

$$C^{(1)}\left(\gamma\right)=\frac{1}{6}\beta_3\left(\gamma^2-1-2z^2\right)+\frac{1}{2}\left\{U\left(\beta_3,z\right)+U\left(\beta_3,-z\right)\right\},\quad (23)$$

$$C^{(2)}\left(\gamma\right)=\frac{1}{4}\beta_3\left\{U\left(\beta_3,z\right)+U\left(\beta_3,-z\right)\right\}\left(z^2\gamma-\frac{1}{3}\gamma^3\right)$$
$$-\frac{1}{8}\gamma\left\{U\left(\beta_3,z\right)+U\left(\beta_3,-z\right)\right\}^2$$
$$-\frac{1}{2}\gamma\left(\beta_4-3-\frac{3}{2}\beta_3^2\right)\left\{U'\left(\beta_3,z\right)+U'\left(\beta_3,-z\right)\right\}$$
$$-\frac{1}{72}\left(\gamma^5-10\gamma^3+15\gamma\right)\beta_3^2$$
$$-\frac{1}{24}\left(\gamma^3-3\gamma\right)\left(\beta_4-3\right)-z\left\{\frac{1}{72}\left(\gamma^4-15\gamma^2\right)\beta_3^2+\frac{1}{8}\gamma^2\left(\beta_4-3\right)\right\}$$
$$+z^2\left\{\frac{1}{72}\left(5\gamma^3-33\gamma\right)\beta_3^2+\gamma\left(\frac{1}{4}\beta_4-1\right)\right\}+\frac{1}{72}z^3\beta_3^2\gamma^2-\frac{1}{9}z^4\beta_3^2\gamma,\quad (24)$$

$$C = \frac{1}{2}\{V(\beta_3, \beta_4, z) + V(\beta_3, \beta_4, -z)\}$$
$$-\frac{1}{4}\{U(\beta_3, z) + U(\beta_3, -z)\}\left\{\tilde{U}(\beta_3, z) + \tilde{U}(\beta_3, -z)\right\}. \qquad (25)$$

Remark 4 In our general class, $U(g_3, y)$ can involve odd powers of y, and hence there is no guarantee that $a(g_3, z)$, shown in (18), will vanish. This is unlike what happens for ED statistics for which $U(g_3, y)$ is an even polynomial in y (see (4)), so that $a(g_3, z)$ necessarily vanishes.

3.4 Second-order power

The first-order term $P^{(0)}(\gamma)$ in (20) is the same for all statistics under consideration. Turning to the second-order term $P^{(1)}(\gamma)$, from (21) and (23), we note that $P^{(1)}(\gamma)$ is an odd function of γ. Furthermore,

$$P^{(1)}(\gamma) = \gamma\phi(z)\,z\left\{U(\beta_3, z) + U(\beta_3, -z) - \frac{2}{3}\beta_3 z^2\right\} + O\left(\gamma^3\right),$$

so that $P^{(1)}(\gamma) = O(\gamma^3)$ provided the polynomial $U(.)$ satisfies

$$\frac{1}{2}\{U(\beta_3, z) + U(\beta_3, -z)\} = \frac{1}{3}\beta_3 z^2. \qquad (26)$$

By (21) and (23), all possible $U(.)$ satisfying (26) have the same $P^{(1)}(\gamma)$. Thus, as in Bravo (2003), the condition (26) characterizes second-order local maximinity in our class, in the sense of maximizing the minimum of $P^{(1)}(\gamma)$ and $P^{(1)}(-\gamma)$ for sufficiently small $|\gamma|$; note that this refers to points equidistant from the null hypothetical value. Following Remark 1, the ELR statistic as well as all generalized empirical exponential family likelihood statistics (Corcoran, 1998, Sect. 4) satisfy (26). *Thus, the second-order local maximinity of the ELR statistic, that was so far known in the Cressie-Reid and ED subclasses (Bravo, 2003), now stands extended to our general class.*

There are two key ingredients in the phenomenon just noted: (i) $P^{(1)}(\gamma) = O(\gamma^3)$ for the ELR statistic, (ii)$P^{(1)}(\gamma)$ is an odd function for every statistic in our class, even though unlike with ED statistics, $U(g_3, y)$ can involve odd powers of y. While (i) is known from Bravo (2003), the explicit formulae (21) and (23) establish (ii). Note that fact (ii) is nontrivial – for example, had there been a statistic in our class satisfying, say $P^{(1)}(\gamma) = \gamma^2\Omega + o(\gamma^2)$ for some Ω (> 0) free from γ, then notwithstanding (i), the second-order local maximinity of the ELR statistic would not hold. Fact (ii) rules out such possibilities.

Comparing (5) with (26), it also follows that empirical statistics in the subclass (d) of Remark 1 will be dominated by the ELR statistic with regard to second-order local maximinity.

3.5 Third-order average power and its implications

The fact (ii) in the previous subsection implies that the arithmetic mean of $P^{(1)}(\gamma)$ and $P^{(1)}(-\gamma)$ is zero for every statistic in our class. Thus, under the criterion of average local power over equidistant points from the null hypothetical value, one needs to consider the third-order term $P^{(2)}(\gamma)$ in (20). Let $\bar{P}^{(2)}(\gamma)$ be the arithmetic mean of $P^{(2)}(\gamma)$ and $P^{(2)}(-\gamma)$. From (22) and (25), it is readily seen that $\bar{P}^{(2)}(\gamma)$ is free from C and hence from $V(.)$, i.e., *statistics with the same* $U(.)$ *will have the same* $\bar{P}^{(2)}(\gamma)$ *in addition to the same* $P^{(1)}(\gamma)$. The same phenomenon is expected to persist if one attempts to extend the present results to the multivariate case. This is analogous to what happens with parametric likelihood-based tests (Bickel et al., 1981; Mukerjee, 1990b). However, even for ED statistics, this was so far hidden because a different kind of stochastic expansion, involving higher order population moments, was used; see Remark 2.

Remark 5 The fact noted in the last paragraph has major implications with regard to the important subclasses, (a), (b), (c) and (e), of statistics considered in Remark 1.

(i) As noted in Remark 1, all generalized empirical exponential family likelihood statistics in the subclass (b) have the same $U(.)$ as the ELR statistic. Hence they all have the same $P^{(1)}(\gamma)$ and $\bar{P}^{(2)}(\gamma)$ as the ELR statistic, which is a member of the subclass (a) of ED statistics.
(ii) It was seen in Remark 1 that all generalized empirical likelihood statistics in the subclass (c) have $U(.)$ as in (4). Now, following Fang and Mukerjee (2005a), any c in (4) is attainable within the subclass (a) of ED statistics (in fact, within the subclass of Cressie-Reid statistics). In other words, given any statistic in the subclass (c), there is one in the subclass (a) with the same $U(.)$and hence the same $P^{(1)}(\gamma)$ and $\bar{P}^{(2)}(\gamma)$.
(iii) The Bartlett-type adjusted version of any ED statistic, as obtained by Bravo (2005) in the spirit of Cordeiro et al. (1991) (vide the subclass (e)), has the same $U(.)$ and hence the same $P^{(1)}(\gamma)$ and $\bar{P}^{(2)}(\gamma)$ as the original statistic.
(iv) As a result of (i)–(iii) above, given any member of the subclasses (b), (c) or (e), there is a member of the subclass (a) with the same $P^{(1)}(\gamma)$ and $\bar{P}^{(2)}(\gamma)$. Therefore, once the subclass (a) of ED statistics is considered, no further gain in $P^{(1)}(\gamma)$ or $\bar{P}^{(2)}(\gamma)$ is possible via the subclasses (b), (c) or (e), so that these latter subclasses may be left out of consideration in so far as higher order power is concerned.

An explicit formula for $\bar{P}^{(2)}(\gamma)$ helps for comparisons that are not covered by Remark 5 above. From (22) and (24), it can be seen that $\bar{P}^{(2)}(\gamma) = \gamma^2\phi(z)(R_0 + R_1) + O(\gamma^4)$, where

$$R_0 = z^3\left(\frac{1}{2}\beta_4 - 2 - \frac{8}{9}\beta_3^2\right) - \frac{2}{9}z^5\beta_3^2$$

is the same for all statistics under consideration, and

$$\begin{aligned}R_1 = &\frac{1}{2}z^3\beta_3\left\{U(\beta_3, z) + U(\beta_3, -z)\right\}\\ &-z\left(\beta_4 - 3 - \frac{3}{2}\beta_3^2\right)\left\{U'(\beta_3, z) + U'(\beta_3, -z)\right\}\\ &-\frac{1}{4}z\left\{U(\beta_3, z) + U(\beta_3, -z)\right\}^2 \end{aligned} \quad (27)$$

depends on the particular statistic. If one specializes to ED statistics, then (27) is in agreement with the findings in Mukerjee (2005).

Indeed, (27) involves the unknown β_3 and β_4, a feature shared by other results in this general area – e.g., those on the mean squared error of point estimators without bias correction (Newey and Smith, 2004, p. 235). However, in the spirit of Newey and Smith (2004, p. 236), one can use (27) for making meaningful comparisons over various possibilities for β_3 and β_4. As an illustration, let R_{11} be the R_1 for the ELR statistic and R_{12} be the R_1 for statistics arising from Bayesian considerations (vide subclass (d) of Remark 1). From (5), note that the latter statistics have the same $U(.)$ and hence the same $\bar{P}^{(2)}(\gamma)$ and R_1. By (27) and Remark 1, it can be seen that

$$R_{11} - R_{12} = \beta_3^2\left(\frac{1}{6}z^3 + \frac{7}{4}z\right) - (\beta_4 - 3)\,z.$$

Thus $R_{11} - R_{12}$ is positive and hence the ELR statistic is superior to statistics arising from Bayesian considerations if either (a) $\beta_4 < 3$, or (b) $\beta_4 = 3$ and $\beta_3 \neq 0$, or (c) $\beta_4 > 3$ and $\beta_4 - 3 < \beta_3^2\left(\frac{1}{6}z^2 + \frac{7}{4}\right)$.

Alternatively, one can consider expected R_1 under suitable prior specification for β_3 and β_4 and optimize over sufficiently large subclasses of our general class. One such subclass consists of statistics for which

$$U(g_3, y) = g_3 B(y)\,, \quad (28)$$

where $B(.)$ is a polynomial with coefficients given by $O(1)$ constants. Observe that all the statistics listed in Remark 1 satisfy (28). Under (28), the formula (27) for R_1 simplifies to

$$R_1 = \beta_3^2 z^3 \bar{B}(z) - 2\left(\beta_4 - 3 - \frac{3}{2}\beta_3^2\right) z\bar{B}(z) - \beta_3^2 z\left\{\bar{B}(z)\right\}^2,$$

where $\bar{B}(z)$ is the arithmetic mean of $B(z)$ and $B(-z)$. Hence if $E_1(> 0)$ and E_2 be the expectations of β_3^2 and $\beta_4 - 3 - \frac{3}{2}\beta_3^2$ under a prior, then one needs to choose $B(.)$ so as to maximize

$$E_1 z^3 \bar{B}(z) - 2E_2 z\bar{B}(z) - E_1 z\left\{\bar{B}(z)\right\}^2,$$

and this is achieved when

$$\bar{B}(z) = \frac{1}{2}z^2 - \left(\frac{E_2}{E_1}\right). \quad (29)$$

No ED statistic, including the ELR statistic, satisfies (29) unless $E_2 = 0$. However, any statistic with

$$B(y) = \frac{1}{2}y^2 - \left(\frac{E_2}{E_1}\right) \tag{30}$$

satisfies (29) for every z. *Thus we get an example of a situation where going beyond the ED subclass can help*. In particular, one may assign a uniform prior on (β_3^2, β_4) over

$$\left\{\left(\beta_3^2, \beta_4\right) : 0 \le \beta_3^2 \le 1.8, \beta_3^2 + 1 < \beta_4 < \frac{15}{8}\beta_3^2 + \frac{9}{2}\right\}, \tag{31}$$

which is the range for the Pearsonian system of distributions; see Pearson and Hartley (1958, p. 210). Then $E_1 = 234/245$, $E_2 = -607/1960$, and (30) reduces to $B(y) = (1/2)y^2 + (607/1872)$.

4 Expected lengths of confidence intervals

In view of (6), a confidence set for θ, with coverage probability $1 - \alpha + o(n^{-1})$, is given by

$$S = \left\{\theta : |W(\theta)| \le z + n^{-1/2}a(g_3, z) + n^{-1}b(g_3, g_4, z)\right\},$$

where $a(.)$ and $b(.)$ are as in (18) and (19). Since $y = (n/m_2)^{1/2}(\bar{X} - \theta)$, by (2) and following Mukerjee and Reid (1999) who considered parametric likelihood based inference, S can be further approximated by the interval

$$I = \left\{\theta : J_1 \le (n/m_2)^{1/2}\left(\bar{X} - \theta\right) \le J_2\right\},$$

with

$$\begin{aligned} J_1 = &-z - n^{-1/2}\{a(g_3, z) + U(g_3, -z)\} \\ &-n^{-1}\left[b(g_3, g_4, z) + V(g_3, g_4, -z) - \{a(g_3, z) + U(g_3, -z)\}\tilde{U}(g_3, -z)\right], \end{aligned}$$

$$\begin{aligned} J_2 = &z + n^{-1/2}\{a(g_3, z) - U(g_3, z)\} \\ &+n^{-1}\left[b(g_3, g_4, z) - V(g_3, g_4, z) - \{a(g_3, z) - U(g_3, z)\}\tilde{U}(g_3, z)\right], \end{aligned}$$

in the sense that the probability for the symmetric difference of S and I to include θ is $o(n^{-1})$. Denote the length of I by L. Then using (18), (19) and expansions analogous to (39) in the appendix, we get

$$E\left(n^{1/2}L\right) = 2z\sigma + n^{-1}(G_0 + G_1)\sigma + o\left(n^{-1}\right). \tag{32}$$

Here

$$G_0 = \frac{1}{9}z^5\beta_3^2 + z^3\left(\frac{2}{9}\beta_3^2 - \frac{1}{6}\beta_4 + 1\right) + z\left\{\frac{1}{4}(\beta_4 - 3) - \frac{1}{3}\beta_3^2\right\} \tag{33}$$

is the same for all statistics under consideration,

$$G_1 = \left(\frac{1}{2}z - \frac{1}{3}z^3\right)\beta_3\left\{U(\beta_3, z) + U(\beta_3, -z)\right\}$$
$$+z\left(\beta_4 - 3 - \frac{3}{2}\beta_3^2\right)\left\{U'(\beta_3, z) + U'(\beta_3, -z)\right\}$$
$$+\frac{1}{4}z\left\{U(\beta_3, z) + U(\beta_3, -z)\right\}^2 \quad (34)$$

depends on the particular statistic, and $\sigma^2, \beta_3, \beta_4$ are now interpreted as $\sigma^2 = E_\theta(X_i - \theta)^2$, $\beta_s = E_\theta\left\{(X_i - \theta)/\sigma\right\}^s$ $(s = 3, 4)$.

If one specializes to ED statistics then (32–34) are in agreement with the findings in Fang and Mukerjee (2005a). Observe that G_0 and G_1 do not involve $V(.)$. This phenomenon, which is expected to persist in any multivariate extension of our results, has the same implications as in Remark 5. Unlike the results on power, the above formula for expected length is non-local. However, the expression for G_1 is similar to that for R_1 in (27) and can be used in the same way as in Subsect. 3.5.

Appendix

On the approximate characteristic function of M_1: By (2) and (8),

$$M_1 = y_0 + n^{-1/2}U_1 + n^{-1}V_1,$$

where $y_0 = y(\theta_0) = (n/m_2)^{1/2}(\bar{X} - \theta_0)$, and

$$U_1 = U(g_3, y_0) - a(g_3, z), \quad V_1 = V(g_3, g_4, y_0) - b(g_3, g_4, z), \quad (35)$$

Hence

$$E_{\theta_n}\left\{\exp(M_1\xi)\right\} = E_{\theta_n}\left[\exp(y_0\xi)\left\{1 + n^{-1/2}U_1\xi + n^{-1}\left(V_1\xi + \frac{1}{2}U_1^2\xi^2\right)\right\}\right]$$
$$+o\left(n^{-1}\right). \quad (36)$$

First consider the leading term on the right-hand side of (36). It can be checked that

$$y_0 = (A_1 + \gamma)\left\{1 - \frac{1}{2}n^{-1/2}A_2 + n^{-1}\left(\frac{1}{2}A_1^2 + \frac{3}{8}A_2^2\right)\right\} + o_p\left(n^{-1}\right), \quad (37)$$

under θ_n. Here

$$A_1 = n^{-1/2}\sum_{i=1}^{n} Z_i,\ A_2 = n^{-1/2}\sum_{i=1}^{n}\left(Z_i^2 - 1\right),\ A_3 = n^{-1/2}\sum_{i=1}^{n}\left(Z_i^3 - \beta_3\right).$$

Hence one can calculate the first four approximate cumulants of y_0 under θ_n, check that the approximate cumulants of still higher orders are $o(n^{-1})$, and thus, after a considerable algebra, show that

$$E_{\theta_n}\left\{\exp(y_0\xi)\right\} = \left[1 + n^{-1/2}\beta_3\xi\sum_{j=0}^{2}h_{j1}K_j(\gamma + \xi) + n^{-1}\xi\sum_{j=0}^{5}h_{j2}K_j(\gamma + \xi)\right]$$
$$\times \exp\left(\gamma\xi + \frac{1}{2}\xi^2\right) + o\left(n^{-1}\right). \quad (38)$$

Next consider the second term on the right-hand side of (36). Analogously to (37), it can be seen that

$$g_3 = \beta_3 + n^{-1/2}\left(A_3 - 3A_1 - \frac{3}{2}\beta_3 A_2\right) + o_p\left(n^{-1/2}\right), \qquad (39)$$

under θ_n, so that by (35),

$$\begin{aligned} E_{\theta_n}\{U_1 \exp(y_0\xi)\} = E_{\theta_n}&\left[\{U(\beta_3, y_0) - a(\beta_3, z)\}\exp(y_0\xi)\right] \\ &+ n^{-1/2}E_{\theta_n}\left[\left(A_3 - 3A_1 - \frac{3}{2}\beta_3 A_2\right)\left\{U'(\beta_3, y_0) - a'(\beta_3, z)\right\}\right. \\ &\left.\times \exp(y_0\xi)\right] + o\left(n^{-1/2}\right). \qquad (40) \end{aligned}$$

The first term on the right-hand side of (40) can be handled via an Edgeworth expansion for y_0 obtained from (38). Under appropriate moment assumptions (Bravo, 2003), the second term can be handled by a conditioning argument, similar to that in Mukerjee (2005), noting that $y_0 = A_1 + \gamma + o_p(1)$ [see (37)] and that, up to the first order of approximation, the limiting distributions of $(A_1, A_2)^T$ and $(A_1, A_3)^T$, under θ_n, are bivariate normal with appropriate parameters – e.g., recalling (10), this involves steps like

$$\begin{aligned} E_{\theta_n}\left\{A_3 U'(\beta_3, y_0)\exp(y_0\xi)\right\} = \beta_4 &\sum\sum c_{sq} s \beta_3^{s-1} \\ &\times \left\{K_{q+1}(\gamma+\xi) - K_q(\gamma+\xi)\gamma\right\} \\ &\times \exp\left(\gamma\xi + \frac{1}{2}\xi^2\right) + o(1), \end{aligned}$$

and so on. Similar considerations apply to the remaining terms on the right-hand side of (36).

If one substitutes (38) and the calculations indicated in the last paragraph in (36), then upon simplification, it follows that the approximate cf of M_1 is as in (15–17).

Acknowledgements We thank the referees for very constructive suggestions. This work was supported by research funds, starting 2003, from Chosun University.

References

Baggerly, K. A. (1998). Empirical likelihood as a goodness-of-fit measure. *Biometrika 85*, 535–547.

Bickel, P. J., Chibisov, D. M., van Zwet, W. R. (1981), On the efficiency of first and second order. *International Statistical Review 49*, 169–175.

Bravo, F. (2003). Second-order power comparisons for a class of nonparametric likelihood-based tests. *Biometrika 90*, 881–890.

Bravo, F. (2005). Bartlett-type adjustments for empirical discrepancy test statistics *Journal of Statistical Planning and Inference* (in press).

Corcoran, S. A. (1998). Bartlett adjustment of empirical discrepancy statistics. *Biometrika 85*, 967–972.

Cordeiro, G. M., Ferrari, S. L. P. (1991). A modified score test statistic having chi-squared distribution to order n^{-1}. *Biometrika 78*, 573–582.
Fang, K. T., Mukerjee, R. (2005a). Expected lengths of confidence intervals based on empirical discrepancy statistics. *Biomerika 92*, 499–503.
Fang, K. T., Mukerjee, R. (2005b). Empirical-type likelihoods allowing posterior credible sets with frequentist validity: Higher order asymptotics. *Biometrika* (in press).
Mittelhammer, R., Judge, G., Miller, D. (2000). *Econometric foundations*, London: Cambridge University Press.
Mukerjee, R. (1990a). Comparison of tests in the multiparameter case I. Second order power. *Journal of Multivariate Analysis 33*, 17–30.
Mukerjee, R. (1990b). Comparison of tests in the multiparameter case II. A third order optimality property of Rao's test. *Journal of Multivariate Analysis 33*, 31–48.
Mukerjee, R. (2005). Higher order power properties of empirical discrepancy statistics. In: J. Fan G. Li (Eds.), *Contemporary multivariate analysis and experimental designs* (in honor of Professor K.T. Fang.) Singapore: World Scientific. pp 75–85.
Mukerjee, R., Reid, N. (1999). On confidence intervals associated with the usual and adjusted likelihoods. *Journal of the Royal Statistical Society, Series B 61*, 945–953.
Newey, W. K., Smith, R. J. (2004). Higher order properties of GMM and generalized empirical likelihood estimators. *Econometrica 72*, 219–255.
Schennach, S. M. (2005). Bayesian exponentially tilted empirical likelihood. *Biometrika 92*, 31–46.
Owen, A. B. (1988). Empirical likelihood ratio confidence intervals for a single functional. *Biometrika 75*, 237–249.
Owen, A. B. (2001). *Empirical likelihood*. London: Chapman and Hall.
Pearson, E. S., Hartley, H. O. (1958). *Biometrika tables for statisticians*, vol. I. London: Cambridge University Press.