

Adjusted empirical likelihood with high-order one-sided coverage precision

JIAHUA CHEN* AND YUKUN LIU^{†,‡}

Constructing confidence intervals with high-order coverage probability precision is more difficult for one-sided intervals than for two-sided intervals. Many existing methods can achieve precision of order n^{-2} for two-sided intervals but only $n^{-1/2}$ for one-sided intervals. Through a creative use of adjusted empirical likelihood, we develop a new procedure that attains coverage precision of order $n^{-3/2}$ for one-sided intervals while retaining order n^{-2} precision for two-sided intervals. We provide detailed comparisons of the asymptotic properties of the new method and those of representative existing methods. Simulation results show that the new method offers many advantages.

AMS 2000 SUBJECT CLASSIFICATIONS: Primary 62G20; secondary 62E20.

KEYWORDS AND PHRASES: Bartlett correction, Confidence limit, Edgeworth expansion, Zero-inflated population.

1. INTRODUCTION

The problem of constructing high-precision confidence intervals based on a random sample for the population mean has perhaps been over-analyzed in statistical literature. For a population with a normal distribution, a confidence interval based on a student- t statistic (or, more precisely, pivot) gives a perfect solution. Even if the data are from a non-normal distribution, the t -interval is still highly satisfactory. Nevertheless, when the population distribution is skewed and/or the parameter has natural boundaries, we look for superior alternatives. For instance, the variance stabilization method is used for the binomial and Poisson distributions to obtain sensible asymmetric intervals. Bartlett correction is used to obtain likelihood intervals with higher-order precision. Resampling methods are used to achieve high-order coverage precision under minimum distributional assumptions (Efron, 1987; Hall, 1988; Loh, 1991). The empirical likelihood or EL (Owen, 1990; Hall and La Scala, 1990) allows us to obtain confidence intervals with a data-driven shape and range-preserving and transformation-respecting

properties. The EL is further found to be Bartlett correctable (DiCiccio, Hall and Romano, 1991) which leads to confidence regions with high-order precision.

In some applications, we are interested in a lower bound on the parameter of interest. For instance, an accounting firm may search for a lower bound on the average overclaim (Kvanli, Shen and Deng, 1998). In such applications, the population contains a large number of zeroes and other observations are positive and heavily skewed. Constructing precise one-sided confidence intervals for such applications is challenging. A two-sided confidence interval may have over-coverage at the lower limit and under-coverage at the upper limit with a satisfactory overall coverage probability. When the population is skewed, a two-sided t -interval has coverage error of order $O(n^{-1})$, whereas the corresponding one-sided t -interval has error of order $O(n^{-1/2})$, where n is the size of the random sample.

Many papers specifically target the precision of one-sided confidence intervals, particularly in the presence of population skewness. Johnson (1978) proposed transforming the t -statistic to improve the precision of the coverage probability. Because this transformation is neither monotone nor invertible, the resulting confidence set is complex. Its asymptotic coverage errors remain $O(n^{-1})$ and $O(n^{-1/2})$ for two-sided and one-sided confidence intervals respectively. Hall (1992) introduced a monotone transformation for constructing both one-sided and two-sided confidence intervals for the population mean. In Hall's interval the two-sided coverage error remains $O(n^{-1})$ and the one-sided coverage error is improved to $O(n^{-1})$. Kvanli, Shen and Deng (1998) proposed an interval based on a parametric mixture model. The application of the plain EL confidence intervals to such populations is investigated in both Chen, Chen and Rao (2003) and Chen and Qin (2003). They found that the one-sided EL intervals are satisfactory without posing a restrictive parametric assumption on the population distribution.

We find that the adjusted empirical likelihood or AEL (Chen, Variyath and Abraham, 2008) provides more satisfactory one-sided confidence intervals with high-order precision. The AEL method adds pseudo-observations to the original data set before the usual EL procedure is applied. It removes a non-existence problem associated with the EL and hence eliminates a blind spot in EL applications. The AEL has the same first-order asymptotic properties as the EL and its numerical computation is much simpler. In addition, Liu and Chen (2010) found the level of adjustment

*Partially supported by NSERC of Canada and by MITACS.

[†]Partially supported by the NNSF of China Grants 11001083 and 11101156, and Innovation Research Grant from ECNU.

[‡]Corresponding author.

at which the AEL achieves the same high-order precision as the Bartlett-corrected EL. The idea is more broadly applicable (Li et al., 2011). However, the improvement in the precision of one-sided confidence intervals is less than that for two-sided intervals.

We report in this paper that an AEL with two fine-tuned pseudo-observations improves the one-sided coverage accuracy to $O(n^{-3/2})$ and the two-sided coverage accuracy to $O(n^{-2})$. This is an improvement on the transformation method of Hall (1992). While high-order Edgeworth expansion (Hall, 1988) or iterative resampling (Loh, 1991) may lead to the same high-order precision, our method is conceptually straightforward and provides another interesting dimension of the EL/AEL methodology.

The rest of the paper is organized as follows. In Section 2, we review several existing methods. Section 3 presents the proposed method and its asymptotic coverage errors. A simulation study is reported in Section 4. All proofs and derivations are postponed to the Appendix.

2. EL CONFIDENCE INTERVALS

Let X_i , $i = 1, 2, \dots, n$ be a set of independent and identically distributed random variables with common distribution F , and x_i be the observed value of X_i . The empirical likelihood (Owen, 1990) of F is then given by

$$L_n(F) = \prod_{i=1}^n p_i$$

with $p_i = P(X_i = x_i)$. Let $\mu = E(X_1)$ and assume its finiteness. The profile empirical likelihood function of μ is defined as

$$L_n(\mu) = \sup \left\{ \prod_{i=1}^n p_i : p_i > 0, \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i(x_i - \mu) = 0 \right\}.$$

Replacing $\sum_{i=1}^n p_i(x_i - \mu) = 0$ by $\sum_{i=1}^n p_i g(x_i; \theta) = 0$ for some $g(x; \theta)$, Qin and Lawless (1994) successfully extended EL for parameters defined through estimating equations.

We focus on the mean of a one-dimensional population. Note that $L_n(\mu)$ attains its maximum value when $\mu = \bar{x}_n$, the sample mean. The empirical log-likelihood ratio function of μ is defined as

$$(1) \quad R_n(\mu) = 2\{\log L_n(\bar{x}) - \log L_n(\mu)\}.$$

If μ_0 is the true value of μ ,

$$\text{PR}\{R_n(\mu_0) \leq x\} = \text{PR}\{\chi_1^2 \leq x\} + O(n^{-1})$$

where χ_1^2 denotes a random variable with the χ_1^2 distribution. Thus, an EL confidence set for μ at level $(1 - 2\alpha)$ can be constructed as

$$(2) \quad I_{\text{EL}} = \{\mu : R_n(\mu) \leq \chi_{1,1-2\alpha}^2\},$$

where $\chi_{1,1-2\alpha}^2$ is the $(1 - 2\alpha)$ th quantile of the χ_1^2 distribution. Clearly the coverage precision of I_{EL} is $O(n^{-1})$ since

$$\text{PR}\{\mu_0 \in I_{\text{EL}}\} = 1 - 2\alpha + O(n^{-1}).$$

In fact, asymptotically I_{EL} always has lower coverage probability than the nominal level. This shortcoming can be corrected by scaling $R_n(\mu)$ down via Bartlett correction. DiCiccio, Hall and Romano (1991) found that

$$(3) \quad \text{PR}\{R_n(\mu_0) \leq (1 + b/n)x\} = \text{PR}\{\chi_1^2 \leq x\} + O(n^{-2})$$

as $n \rightarrow \infty$, where the Bartlett correction factor

$$b = (1/2)\alpha_4/\alpha_2^2 - (1/3)\alpha_3^2/\alpha_2^3 = (1/2)\kappa - (1/3)\gamma^2 + (3/2)$$

with $\alpha_k = E(X - \mu)^k$ for $k = 1, 2, \dots$, where γ and κ are the population skewness and kurtosis. The Bartlett-corrected EL (BEL) confidence interval is given by

$$I_{\text{BEL}} = \{\mu : R_n(\mu) \leq (1 + b/n)\chi_{1,1-2\alpha}^2\}.$$

Its coverage precision is improved to $O(n^{-2})$. The moment relationship implies $b > 0$, so the BEL-intervals always have higher coverage probabilities than the EL-intervals. The Bartlett correction is also applicable to multidimensional μ and for parameters defined through estimating equations. See DiCiccio, Hall and Romano (1991) and Chen and Cui (2006, 2007) for further discussion.

The AEL was proposed by Chen, Variyath and Abraham (2008) to overcome the obstacle in EL applications caused by $\sum_{i=1}^n p_i g(x_i; \theta) = 0$ having no solution in non-negative p_1, \dots, p_n . Simplistically, they added a pseudo-observation $g_{n+1} = -a\bar{g}_n$ for some $a > 0$ where $\bar{g}_n = n^{-1} \sum_{i=1}^n g(x_i; \theta)$ to the data set before computing R_n defined similarly to (1). Choosing a low level of adjustment a retains all the first-order asymptotic properties of EL and enables a smoother numerical implementation.

Interestingly, Liu and Chen (2010) found that setting $a = b/2$ improves the precision of the AEL two-sided confidence intervals to $O(n^{-2})$. Hence, AEL and BEL attain the same precision. In this paper, we apply AEL to obtain high-order one-sided intervals.

3. HIGH-ORDER ONE-SIDED AEL INTERVALS

The likelihood ratio function $R_n(\mu)$ in (1) is a convex function of μ . When μ approaches the maximum or minimum observation, $R_n(\mu)$ goes to infinity. Hence, I_{EL} defined in (2) is a bounded two-sided confidence interval. A one-sided EL-interval can be easily constructed. Let $\hat{\mu}_1 < \hat{\mu}_2$ be two solutions to $R_n(\mu) = \chi_{1,1-2\alpha}^2$. Then two one-sided $(1 - \alpha)$ -intervals for μ are $[\hat{\mu}_1, \infty)$ and $(-\infty, \hat{\mu}_2]$. It can be easily verified that

$$\text{PR}(\hat{\mu}_1 \leq \mu_0) = 1 - \alpha + O(n^{-1/2}).$$

The Bartlett correction given by Diccicio, Hall and Romano (1991) does not improve the coverage precision of the above one-sided interval.

Let $x_{n+1} = \mu - a_1(\bar{x} - \mu)$ and

$$L_n(\mu; a_1) = \sup \left\{ \prod_{i=1}^{n+1} p_i : p_i > 0, \sum_{i=1}^{n+1} p_i = 1, \sum_{i=1}^{n+1} p_i(x_i - \mu) = 0 \right\}.$$

Define $R_n(\mu; a_1) = 2\{\log L_n(\bar{x}; a_1) - \log L_n(\mu; a_1)\}$ and let $\tilde{\mu}_1 < \tilde{\mu}_2$ be the solutions to $R_n(\mu; a_1) = \chi_{1,1-2\alpha}^2$. Liu and Chen (2010) showed that when $a_1 = b/2$ the two-sided interval $[\tilde{\mu}_1, \tilde{\mu}_2]$ contains μ_0 with probability $1 - 2\alpha + O(n^{-2})$.

The AEL given by Liu and Chen (2010) mainly inflates the coverage probability of the EL confidence region to correct the under-coverage problem. To further improve the precision of the one-sided interval, we must also address the imbalance caused by the skewness of the population. For this purpose, we add another pseudo-observation as follows:

Let x_{n+1} be the same as before and add a second pseudo-observation $x_{n+2} - \mu = a_2$. This choice adjusts the center of the data to provide a more precise one-sided confidence interval. Let the new AEL profile likelihood function of μ be

$$L_n(\mu; \mathbf{a}) = \sup \left\{ \prod_{i=1}^{n+2} p_i : p_i > 0, \sum_{i=1}^{n+2} p_i = 1, \sum_{i=1}^{n+2} p_i(x_i - \mu) = 0 \right\}$$

where $\mathbf{a} = (a_1, a_2)$, $x_{n+1} = \mu - a_1(\bar{x} - \mu)$, and $x_{n+2} = \mu + a_2$. The constrained maximization problem can be easily solved by the Lagrange multiplier method. We find

$$\log L_n(\mu; \mathbf{a}) = - \sum_{i=1}^{n+2} \log\{1 + \lambda(x_i - \mu)\} - (n+2) \log(n+2)$$

in which λ is the solution to

$$\sum_{i=1}^{n+2} \frac{x_i - \mu}{1 + \lambda(x_i - \mu)} = 0.$$

We hence obtain the profile likelihood ratio function of μ as

$$(4) \quad R(\mu; \mathbf{a}) = 2 \sum_{i=1}^{n+2} \log\{1 + \lambda(x_i - \mu)\}.$$

A choice of $a_1 > 0$ ensures the constraint always has a solution. We further tune a_1 and a_2 to obtain high-order one-sided AEL confidence intervals.

Theorem 3.1. *Let x_1, \dots, x_n be a random sample from population F . Assume that F satisfies Cramér's condition and has a finite 18th moment. Denote the true value of its mean μ_0 . Let the two pseudo-observations be defined as $x_{n+1} - \mu = -a_1(\bar{x} - \mu)$ and $x_{n+2} - \mu = a_2$ with*

$$a_1 = (1/2)b; \quad a_2 = (1/6)\alpha_3/\alpha_2 = (1/6)\gamma\alpha_2^{1/2}$$

where b is the Bartlett correction factor, and $\gamma = \alpha_3/\alpha_2^{3/2}$ is the population skewness. Let the likelihood ratio function be defined as in (4). Then

$$\Pr\{R_n(\mu_0; \mathbf{a}) \leq \chi_{1,1-2\alpha}^2\} = 1 - 2\alpha + O(n^{-2}).$$

In addition, let $\hat{\mu}_L < \hat{\mu}_U$ be the two solutions to $R_n(\mu; \mathbf{a}) = \chi_{1,1-2\alpha}^2$. Then

$$\begin{aligned} \Pr(\mu_0 \geq \hat{\mu}_L) &= 1 - \alpha + O(n^{-3/2}); \\ \Pr(\mu_0 \leq \hat{\mu}_U) &= 1 - \alpha + O(n^{-3/2}); \\ \Pr(\hat{\mu}_L \leq \mu_0 \leq \hat{\mu}_U) &= 1 - 2\alpha + O(n^{-2}). \end{aligned}$$

Under Cramér's condition and the finite moment assumption, the distribution of the smooth functions of the sample mean has an Edgeworth expansion. Cramér's condition requires $\limsup_{|t| \rightarrow \infty} |E \exp(itX)| < 1$. This condition is satisfied by all continuous distributions.

Interestingly, the recommended level of adjustment for inflating the coverage probability is again half of the Bartlett correction factor $b/2$. Further, the recommended shift of the center is related to the population skewness and coincides with that of Hall (1992). Note that the multiplication factor $\alpha_2^{1/2}$ in a_2 makes the adjustment scale-invariant. Compared to the Bartlett correction given by DiCiccio, Hall and Romano (1991) and the AEL of Liu and Chen (2010), the new adjustment attains the same precision for two-sided intervals and improves the precision of one-sided intervals.

The above theorem cannot be directly applied because the level of adjustment must be estimated. As usual, the estimation error is of the order of $O_p(n^{-1/2})$. In contrast to the results of Diccicio, Hall and Romano (1991) and Liu and Chen (2010), when we estimate a_2 this error has a non-negligible effect on the precision of the coverage probability. The sample version of the above theorem must therefore be revised.

Let $\hat{\alpha}_k = n^{-1} \sum_{i=1}^n (x_i - \bar{x})^k$ be the moment estimator of α_k , and $\hat{\gamma} = \hat{\alpha}_3/\hat{\alpha}_2^{3/2}$ be the moment estimator of the population skewness. The result is as follows:

Theorem 3.2. *Assume the same conditions as in Theorem 3.1. Let*

$$\hat{a}_1 = \hat{b} - (1/12)\hat{\alpha}_2^{-2}\hat{\alpha}_4, \quad \hat{a}_2 = (1/6)\hat{\gamma}(\hat{\alpha}_2)^{1/2} + (1/2)(\bar{x} - \mu),$$

and $\hat{\mathbf{a}} = (\hat{a}_1, \hat{a}_2)$. Define the likelihood ratio function as in (4). Then

$$\Pr\{R_n(\mu_0; \hat{\mathbf{a}}) \leq \chi_{1,1-2\alpha}^2\} = 1 - 2\alpha + O(n^{-2}).$$

Let $\hat{\mu}_L$ and $\hat{\mu}_U$ be two solutions to $R_n(\mu; \hat{\mathbf{a}}) = \chi_{1,1-2\alpha}^2$. Then

$$\begin{aligned} \Pr(\mu_0 \geq \hat{\mu}_L) &= 1 - \alpha + O(n^{-3/2}); \\ \Pr(\mu_0 \leq \hat{\mu}_U) &= 1 - \alpha + O(n^{-3/2}); \\ \Pr(\hat{\mu}_L \leq \mu_0 \leq \hat{\mu}_U) &= 1 - 2\alpha + O(n^{-2}). \end{aligned}$$

In Theorem 3.2, we replace a_2 by its moment estimator with an additional $(1/2)(\bar{x} - \mu)$. This arrangement allows a simple and nonnegative \hat{a}_1 . There is flexibility in Theorem 3.2 when the corresponding AEL is implemented. In general, any estimators that differ from \hat{a}_1 by $O(n^{-1/2})$ and differ from \hat{a}_2 by $O(n^{-1})$ do not alter the conclusion of Theorem 3.2. Through some simulation experiments, we found that this helps to make the estimation of a_1 and a_2 robust. More details will be provided in the simulation section.

4. ASYMPTOTIC COMPARISON OF ONE-SIDED CONFIDENCE INTERVALS

There are many methods for constructing one-sided confidence intervals. In this section, we make asymptotic comparisons of the new method with a number of existing methods. For each method, we give an asymptotic expansion of its upper confidence limit. Because of symmetry, the asymptotic expansion of the lower confidence limit is the same.

The most commonly used one- or two-sided intervals are based on a t -statistic. The upper confidence limit of the one-sided t -interval is given by

$$U_T = \bar{x} + n^{-1/2} s_n t_{n-1, 1-\alpha}$$

where \bar{x} and s_n^2 are the sample mean and variance, respectively. Its coverage probability is exactly $1 - \alpha$ when the random sample is from a normal distribution. When the population distribution is not normal, this upper limit can be expanded, under some moment conditions, as

$$U_T = \bar{x} + n^{-1/2} s_n z_{1-\alpha} + O_p(n^{-3/2}).$$

For notational simplicity, we take $s_n^2 = n^{-1} \sum_{i=1}^n (x_i - \bar{x})^2$. This minor deviation from the traditional sample variance does not invalidate the above expansion.

The upper limit of the H-interval (Hall, 1992) is obtained from a transformed t -statistic. It has the following expression and asymptotic expansion:

$$\begin{aligned} U_H &= \bar{x} - 3\hat{\gamma}^{-1} s_n \{1 + \hat{\gamma}(n^{-1/2} z_\alpha - n^{-1} \hat{\gamma}/6)\}^{1/3} - 1 \\ &= \bar{x} + n^{-1/2} s_n z_{1-\alpha} + n^{-1} s_n \hat{\gamma} (1 + 2z_{1-\alpha}^2)/6 \\ &\quad + O_p(n^{-3/2}). \end{aligned}$$

The interval of Johnson (1978) is insightful and likely motivated the above improved version. Because it is asymptotically inferior, its asymptotic expansion is not presented here.

The upper confidence limit of the one-sided EL-interval is given by

$$\begin{aligned} U_{EL} &= \bar{x} + n^{-1/2} s_n z_{1-\alpha} + n^{-1} s_n z_{1-\alpha}^2 \gamma / 3 \\ &\quad + n^{-3/2} s_n z_{1-\alpha}^3 (9 + 9\kappa - 8\gamma^2) / 36 + O_p(n^{-2}). \end{aligned}$$

We have expanded it to order n^{-2} to demonstrate the Bartlett correction method that follows.

The BEL inflates the confidence limits by replacing $z_{1-\alpha}^2$ with $(1 + b/n)z_{1-\alpha}^2$. It affects only the $n^{-3/2}$ term:

$$\begin{aligned} U_{BEL} &= U_{EL} + n^{-3/2} s_n z_{1-\alpha} b / 2 + O_p(n^{-2}) \\ &= \bar{x} + n^{-1/2} s_n z_{1-\alpha} + n^{-1} s_n z_{1-\alpha}^2 \gamma / 3 \\ &\quad + n^{-3/2} s_n z_{1-\alpha}^3 (9 + 18\kappa - 7\gamma^2) / 18 + O_p(n^{-2}). \end{aligned}$$

The AEL of Liu and Chen (2010) has the same expansion as that of U_{BEL} to order $O_p(n^{-2})$. The AEL-interval of Theorem 3.2 has an upper confidence limit

$$\begin{aligned} U_{AEL} &= \bar{x} + n^{-1/2} s_n z_{1-\alpha} + n^{-1} s_n (1 + 2z_{1-\alpha}^2) \hat{\gamma} / 6 \\ &\quad + n^{-3/2} s_n z_{1-\alpha} \{ (9\kappa - 8\gamma^2 + 9) z_{1-\alpha}^2 / 48 \\ &\quad + (30\kappa - 23\gamma^2 + 54) / 72 \} + O_p(n^{-2}). \end{aligned}$$

The derivations of the three EL-based upper limits will be given in the Appendix.

The coverage probabilities of the above one-sided confidence intervals are as follows. For simplicity, we keep only the terms whose order is relevant for demonstrating the high-order precision. We do not give explicit forms of $q_1(\cdot)$ and $q_2(\cdot)$, which are odd and even polynomial functions respectively.

$$\begin{aligned} \text{PR}\{\mu \leq U_T\} &= 1 - \alpha + (1/6)n^{-1/2} \gamma (2z_{1-\alpha}^2 + 1) \phi(z_{1-\alpha}) \\ &\quad + O(n^{-1}); \\ \text{PR}\{\mu \leq U_{EL}\} &= 1 - \alpha + (1/6)n^{-1/2} \gamma \phi(z_{1-\alpha}) + O(n^{-1}); \\ \text{PR}\{\mu \leq U_{BEL}\} &= 1 - \alpha + (1/6)n^{-1/2} \gamma \phi(z_{1-\alpha}) + O(n^{-3/2}); \\ \text{PR}\{\mu \leq U_H\} &= 1 - \alpha + n^{-1} q_1(z_{1-\alpha}) \phi(z_{1-\alpha}) + O(n^{-3/2}); \\ \text{PR}\{\mu \leq U_{AEL}\} &= 1 - \alpha + n^{-3/2} q_2(z_{1-\alpha}) \phi(z_{1-\alpha}) + O(n^{-2}). \end{aligned}$$

Note that U_T , U_{EL} , and U_{BEL} have the same order of precision. Because $\phi(\cdot)$ is an even function, when U_{EL} and U_{BEL} are combined with their corresponding lower confidence limits, say L_{EL} and L_{BEL} , over-coverage at one end is balanced out by under-coverage at the other end. Hence, their two-sided intervals have higher-order precision.

The one-sided H-interval has straightforward high-order precision $O(n^{-1})$. The new one-sided AEL-interval is a clear asymptotic winner with coverage precision $O(n^{-3/2})$. This order is further improved to $O(n^{-2})$ if a two-sided interval is constructed.

5. SIMULATION BASED ON ARTIFICIAL AND REAL POPULATIONS

We now investigate the performances of the new one-sided AEL-interval and the existing methods reviewed earlier. We report the results for five intervals denoted T, H, EL, BEL, and AEL. This AEL is not that given by Liu and Chen (2010). We first simulate random samples from four

artificial distributions: a standard normal, a chi-squared distribution with 4 degrees of freedom, an exponential distribution with mean 1, and a standard log normal distribution. Later we will also use a zero-inflated exponential distribution and a real data set. The first four populations are in order of increasing skewness. We use the sample sizes $n = 20$ and $n = 40$ and generate 10,000 random samples from each population. We count the proportions of $\mu < L$ and $\mu > U$ where L and U stand for the lower and upper confidence limits. The target values are 5% corresponding to the one-sided 95% confidence interval. With 10,000 random samples, the simulation error is about 0.5%. The results for a target coverage of 90% are similar.

The direct computation of \hat{a}_1 and \hat{a}_2 in Theorem 3.2 results in unstable outcomes for the AEL interval. Let $x_{(i)}$ be the i th order statistic, $i = 1, 2, \dots, n$. In the simulation, we replaced $x_{(n)}$ and $x_{(1)}$ by $x_{(n-1)}$ and $x_{(2)}$ when computing \hat{a}_1 and \hat{a}_2 . We also replaced \hat{a}_1 by $\hat{a}_1/(1 + 2\hat{a}_1/n)$. In addition, we replaced $\bar{x} - \mu$ by

$$(\bar{x} - \mu)/\{1 + 0.2(\bar{x} - \mu)^2/\hat{\alpha}_2\}$$

to limit the impact when μ is clearly far from μ_0 . Our asymptotic results are based on the assumption that $\bar{x} - \mu = O_p(n^{-1/2})$. Hence, the theoretical result works best when $\bar{x} - \mu$ is not exceptionally large. At the same time, the above modifications do not invalidate Theorem 3.2.

We present the simulation results in two figures, each containing twelve plots corresponding to two sample sizes, and one row for each population distribution. In each row, the first plot gives the non-coverage probabilities of the lower and upper limits. A one-sided confidence interval is accurate if its point is close to the horizontal 5% line. The second plot provides box-plots of the lower confidence limits. Provided the coverage probabilities are comparable, the larger and more concentrated the lower limit, the better the interval. The third plot provides box-plots of the upper confidence limits. Given similar coverage, those that are smaller and more concentrated are better.

Let us first examine Fig. 1 where $n = 20$. The T-interval is the clear winner under normality, as expected. However, the other four methods are acceptable in all respects. When we move down so that the population becomes more and more skewed, the coverage of the T-interval deteriorates. The other four methods also get worse for the upper confidence limit but to a lesser degree, and they maintain comparable and accurate coverage for the lower confidence limit. In terms of the coverage probability, the H-interval is the best and the EL-interval is the worst. However, the H-interval has an unstable upper limit. The second row of plots shows that all the intervals except the T-interval have nearly the same average lower limits. A closer examination of the lower coverage probability reveals that the AEL-interval has an edge over the BEL-interval. Note that the scale change has partially masked the true difference.

The results for $n = 40$ are given in Fig. 2. It is clear that all the methods produce more efficient intervals and more precise coverage probabilities than for $n = 20$. Yet the relative performances are the same. In conclusion, while the new AEL is not a star among the five methods investigated, it is a good overall performer. It is a good choice for applications where the population is apparently skewed.

Now we move to a zero-inflated population and provide a real-data example. We generated data from a mixture model that contains 15% nonzero observations from an exponential with mean 5. Note that 15% is the probability that each observation is nonzero; a sample with 100 observations does not necessarily have exactly 15 nonzero observations. Such populations are common in accounting (Kvanli, Shen, and Deng, 1998). The challenge is to give an accurate lower confidence limit. We choose a typical sample size, $n = 100$, in this simulation. See the plots in the first row in Fig. 3. It is simple to conclude that the AEL-interval is a good method. The H-interval has the most accurate coverage probabilities, but its upper confidence limits are unstable. A stabilization measure may help but such a measure could reduce the coverage precision.

Our previous examples were all based on artificially generated populations. Next we conduct a simulation based on a real data set from Chen, Chen and Rao (2003). The data from the Canadian Labour Force Survey of 2000 were provided by Statistics Canada through the TriUniversity Data Resources website. The address is <http://tdr.uoguelph.ca>, but access is restricted. We only took a random sample containing 10% of the data from the province of Ontario. We use the number of extra hours worked as the response variable y . Among the 17,415 sampled observations, there are 3,677 nonzero values. The proportion of nonzero observations is 21%. The mean of the entire data set is 19.46. We wish to examine the performance of the AEL-interval, particularly in comparison with that of the EL advocated by Chen, Chen and Rao (2003).

We constructed 95% one-sided confidence intervals based on 10,000 samples (without replacement) of size 100. Because the sample fraction is small, we ignore the finite-population correction factor. The outcomes are summarized via three plots in the second row of Fig. 3. We note that the H-intervals are impressively accurate. For the more important lower limit, AEL is more accurate. AEL has a tighter average lower limit (12.71) than BEL (12.60) while their standard deviations are 3.76 and 3.75 respectively. The corresponding figures for EL are 12.62 and 3.74. Thus, the AEL-interval has a better lower limit. The one-sided H-, EL-, and AEL-intervals have 4.5%, 4.5%, and 4.8% non-coverage, all of which are close to the 5% target.

These results again show that the new AEL is not a star among the five methods investigated, but it is a good overall performer and a good choice in applications where the population is apparently skewed or zero-inflated.

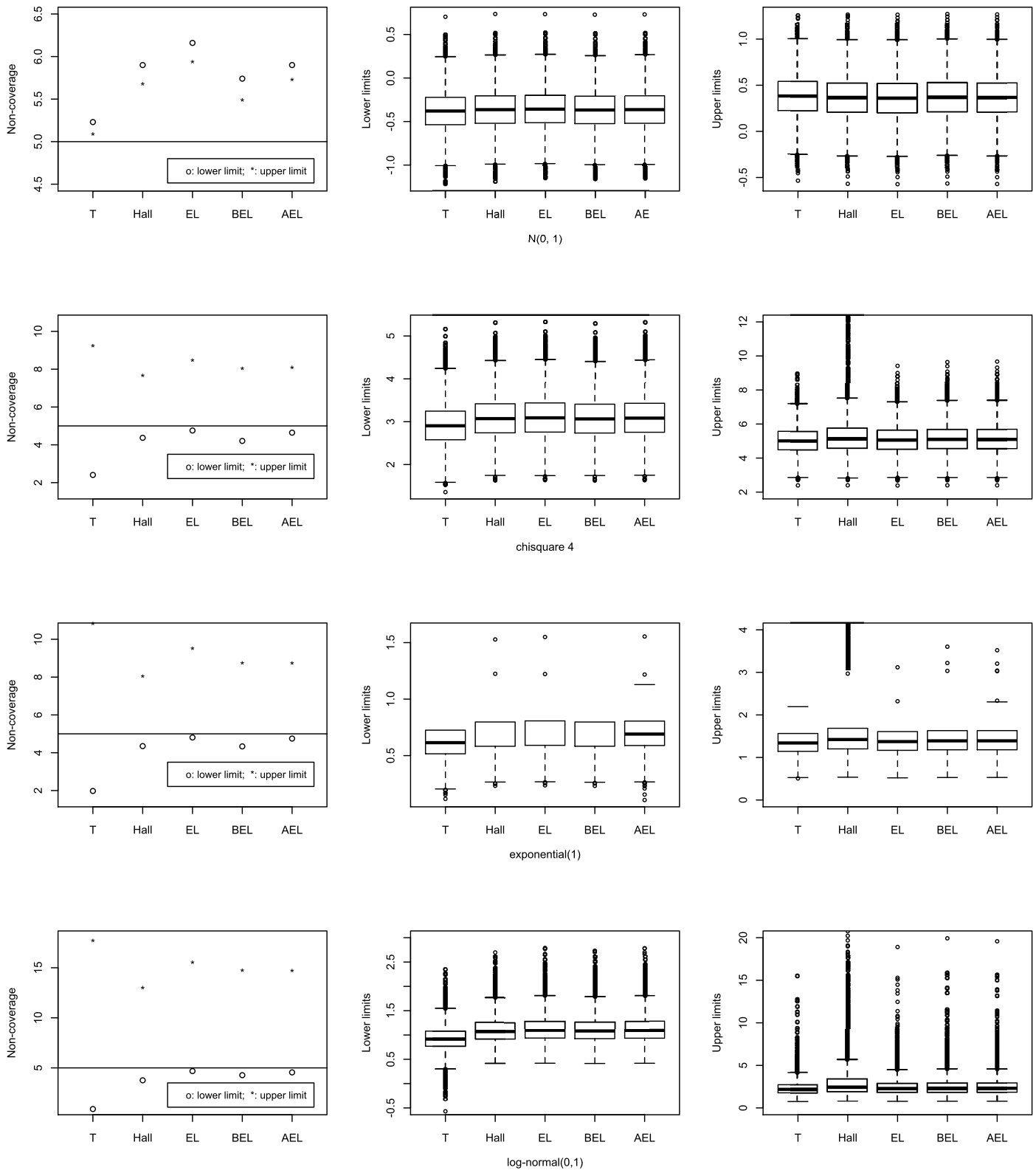


Figure 1. Simulation results for one-sided intervals when $n = 20$.

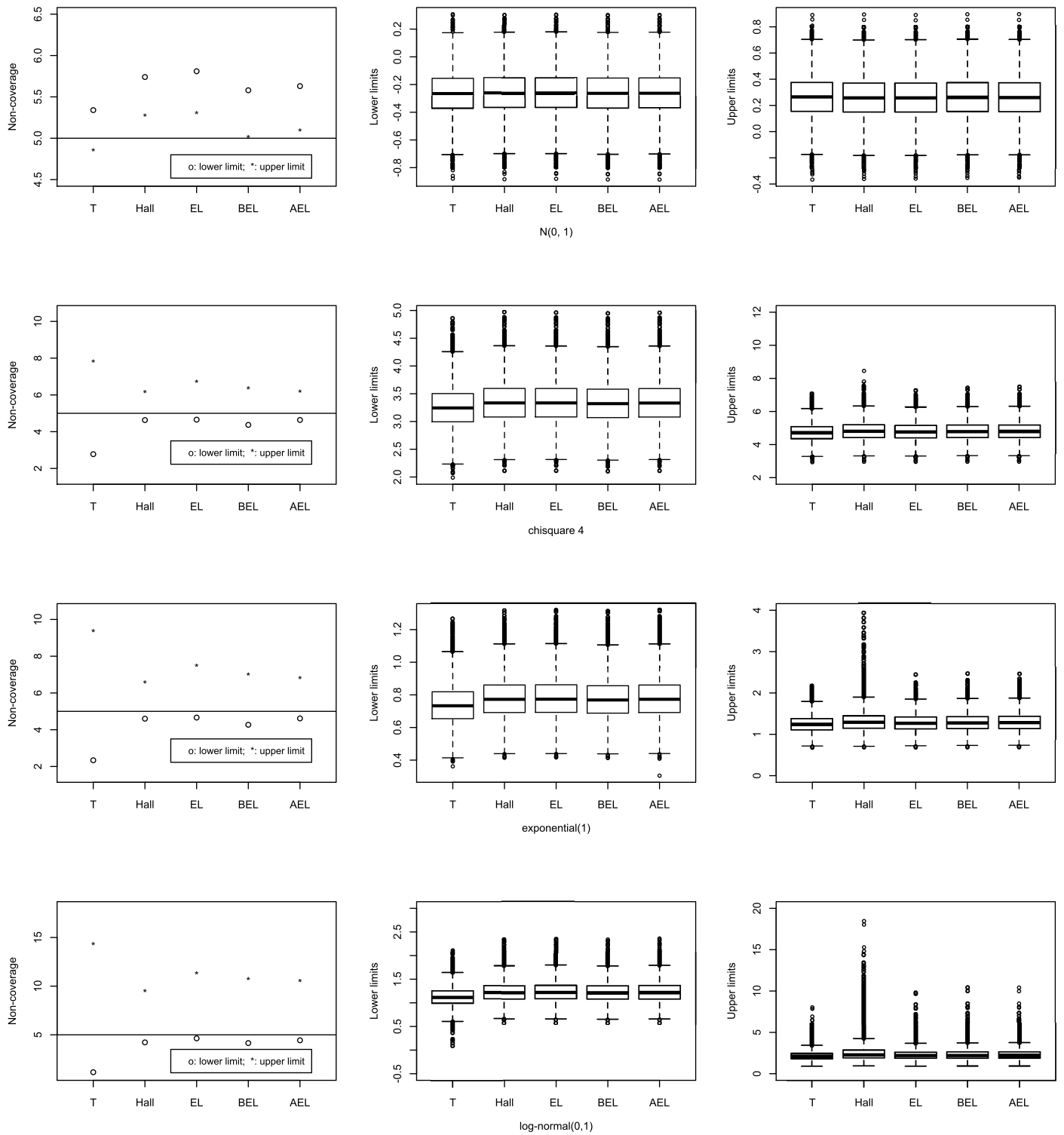


Figure 2. Simulation results for one-sided intervals when $n = 40$.

APPENDIX A. PROOFS

Denote the true value of the population mean μ_0 . Let $\alpha_k = E(X_i - \mu_0)^k$ and $A_k = n^{-1} \sum_{i=1}^n (X_i - \mu_0)^k - \alpha_k$ for $k \geq 2$. Without loss of generality, we assume that $\alpha_2 = 1$ in

the following derivations. Under appropriate finite-moment assumptions, $A_k = O_p(n^{-1/2})$. The moment assumptions are given in Theorem 3.1 so the above order assessments are subsequently assumed.

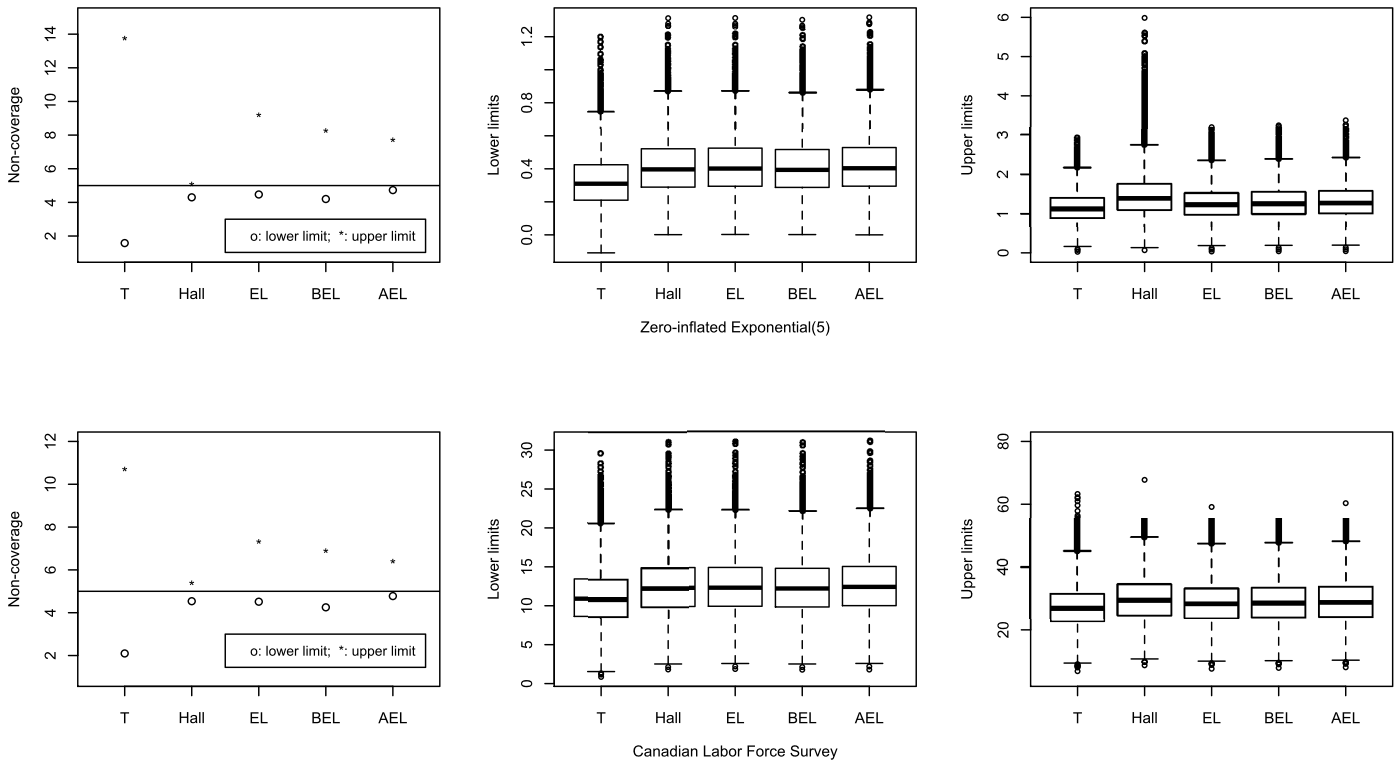


Figure 3. One-sided interval for zero-inflated population.

A.1 Expanding $R(\mu_0; \mathbf{a})$

We first give an asymptotic expansion of $R(\mu_0; \mathbf{a})$ when \mathbf{a} contains two non-random and sample-size independent constants through its connection with $R(\mu_0; \mathbf{0})$. We can easily verify that $R(\mu_0; \mathbf{0})$ equals the usual empirical log-likelihood ratio. Hence, according to DiCiccio, Hall and Romano (1991),

$$R(\mu_0; \mathbf{0}) = n\{T_1 + T_2 + T_3 + O_p(n^{-2})\}^2$$

where

$$\begin{aligned} T_1 &= A_1, \\ T_2 &= -(1/2)A_1A_2 + (1/3)\alpha_3A_1^2, \\ T_3 &= (3/8)A_1A_2^2 - (5/6)\alpha_3A_1^2A_2 + (1/3)A_1^2A_3 \\ &\quad + \{(4/9)\alpha_3^2 - (1/4)\alpha_4\}A_1^3. \end{aligned}$$

We may regard $\sqrt{n}\{T_1+T_2+T_3\}$ as a signed root of $R(\mu_0, \mathbf{0})$.

Given this expansion, we need only investigate the difference $R(\mu_0; \mathbf{a}) - R(\mu_0; \mathbf{0})$ to find an expansion for $R(\mu_0; \mathbf{a})$. The outcome is given in the following lemma:

Lemma 1. *Under the conditions of Theorem 3.1, we have*

$$\begin{aligned} R(\mu_0; \mathbf{a}) - R(\mu_0; \mathbf{0}) &= 2a_2A_1 + (2a_1 + 2\alpha_3a_2 - a_2^2)A_1^2 \\ &\quad - 2a_2A_1A_2 + a_2^2n^{-1} + O_p(n^{-3/2}). \end{aligned}$$

Proof. The constrained optimization problem of the EL can be solved by the Lagrange multiplier method, as discussed previously. The same technique provides a tool for the asymptotic expansion. Let

$$\begin{aligned} \phi_0(\lambda) &= \sum_{i=1}^n \log\{1 + \lambda(x_i - \mu_0)\}, \\ \phi_a(\lambda) &= \sum_{i=1}^{n+2} \log\{1 + \lambda(x_i - \mu_0)\} \end{aligned}$$

with x_{n+1} and x_{n+2} being the pseudo-observations defined in Theorem 3.1. Let λ_0 and λ_a be stationary points of $\phi_0(\lambda)$ and $\phi_a(\lambda)$ respectively. They are Lagrange multipliers for the constrained maximization. It is well known (Owen, 2001) that $\lambda_0 = O_p(n^{-1/2})$ and we claim without proof that $\lambda_a = O_p(n^{-1/2})$. The importance of the Lagrange multipliers is further illustrated by the fact that they satisfy $R(\mu_0; \mathbf{0}) = 2\phi_0(\lambda_0)$ and $R(\mu_0; \mathbf{a}) = 2\phi_a(\lambda_a)$.

In what follows, we link λ_0 and λ_a and other quantities through asymptotic expansions. Being a stationary point, λ_a solves the following equation:

$$\begin{aligned} (5) \quad 0 &= \phi'_a(\lambda_a) \\ &= \phi'_a(\mathbf{0}) + \phi''_a(\mathbf{0})\lambda_a + (1/2)\phi'''_a(\mathbf{0})\lambda_a^2 + O_p(n^{-1/2}). \end{aligned}$$

The derivatives of $\phi_a(\cdot)$ have the following expansions:

$$(6) \quad \begin{aligned} \phi'_a(\mathbf{0}) &= nA_1 + a_2 + O_p(n^{-1/2}), \\ \phi''_a(\mathbf{0}) &= -n(1 + A_2) + O_p(1), \\ \phi'''_a(\mathbf{0}) &= 2n\alpha_3 + O_p(n^{1/2}). \end{aligned}$$

Substituting (6) into (5), we get

$$(7) \quad A_1 - (1 + A_2)\lambda_a + a_2n^{-1} + \alpha_3\lambda_a^2 = O_p(n^{-3/2}).$$

Because $A_k = O_p(n^{-1/2})$ and $\lambda_a = O_p(n^{-1/2})$, the above equation becomes

$$A_1 - \lambda_a = O_p(n^{-1}).$$

This leads to $\lambda_a = A_1 + \epsilon_n$ with $\epsilon_n = O_p(n^{-1})$. Substituting $\lambda_a = A_1 + \epsilon_n$ back into (7), we find

$$\epsilon_n = a_2n^{-1} + \alpha_3A_1^2 - A_1A_2 + O_p(n^{-3/2}).$$

Consequently, we find

$$\lambda_a = A_1 - A_1A_2 + \alpha_3A_1^2 + a_2n^{-1} + O_p(n^{-3/2}).$$

Setting $\mathbf{a} = \mathbf{0}$ in the above expansion, we also get

$$\lambda_0 = A_1 - A_1A_2 + \alpha_3A_1^2 + O_p(n^{-3/2}).$$

Using the above expansions for $\lambda_0 - \lambda_a$ and noting that $\phi'_a(\lambda_a) = 0$, we get

$$\begin{aligned} 2\phi_a(\lambda_0) &= 2\phi_a(\lambda_a) + 2(\lambda_0 - \lambda_a)\phi'_a(\lambda_a) + (\lambda_a - \lambda_0)^2\phi''_a(\lambda_a) \\ &\quad + O_p(n^{-3/2}) \\ &= 2\phi_a(\lambda_a) + a_2^2n^{-2}\phi''_a(\lambda_a) + O_p(n^{-3/2}). \end{aligned}$$

Computing $\phi''(\lambda_a)$ with the expansion of λ_a , we find $\phi''_a(\lambda_a) = -n + O_p(n^{1/2})$. Hence, we have

$$(8) \quad 2\phi_a(\lambda_0) = 2\phi_a(\lambda_a) - a_2^2n^{-1} + O_p(n^{-3/2}).$$

Linking $\phi_a(\cdot)$ with $R(\mu_0; \mathbf{a})$, and using the relationship between $\phi_a(\cdot)$ and $\phi_0(\cdot)$, we find

$$\begin{aligned} R(\mu_0; \mathbf{a}) &= 2\phi_a(\lambda_a) \\ &= 2\phi_a(\lambda_0) + a_2^2n^{-1} + O_p(n^{-3/2}) \\ &= 2\phi_0(\lambda_0) + 2\log\{1 + \lambda_0(x_{n+1} - \mu_0)\} \\ &\quad + 2\log\{1 + \lambda_0(x_{n+2} - \mu_0)\} \\ &\quad + a_2^2n^{-1} + O_p(n^{-3/2}) \\ &= R(\mu_0; \mathbf{0}) + 2\log\{1 + \lambda_0(x_{n+1} - \mu_0)\} \\ &\quad + 2\log\{1 + \lambda_0(x_{n+2} - \mu_0)\} \\ &\quad + a_2^2n^{-1} + O_p(n^{-3/2}). \end{aligned}$$

Finally, by expanding the logarithm terms given the expansion of λ_0 , we get the conclusion of this lemma. \square

A.2 Expanding the signed root of $R(\mu_0; \mathbf{a})$

While the results of Theorems 3.1 and 3.2 are stated in terms of the likelihood ratio function, their proofs are built on Edgeworth expansions of their corresponding signed roots. As a preliminary step, we first find a signed root of $R(\mu_0; \mathbf{a})$ through its relationship with $R(\mu_0; \mathbf{0})$ obtained in the last subsection.

Suppose $R(\mu_0; \mathbf{a})$ has the following decomposition

$$R(\mu_0; \mathbf{a}) = n\{T_1^* + T_2^* + T_3^* + O_p(n^{-2})\}^2,$$

such that $T_k^* = O_p(n^{-k/2})$ for $k = 1, 2, 3$. Matching the expansion of $R(\mu_0; \mathbf{a})$ order by order, it is easy to verify that the following T_k^* provide a solution to the signed root:

$$\begin{aligned} T_1^* &= T_1, \\ T_2^* &= T_2 + a_2n^{-1}, \\ T_3^* &= T_3 + \{[-a_1 + (2/3)\alpha_3a_2 - (1/2)a_2^2]A_1 \\ &\quad - (1/2)a_2A_2\}n^{-1}. \end{aligned}$$

The Edgeworth expansion of $\sqrt{n}T^* = \sqrt{n}\{T_1^* + T_2^* + T_3^*\}$ is determined by its cumulants. Let κ_k^* denote the k -th cumulant of T^* . When $a_2 = (1/6)\gamma$ that is not random, some lengthy but simple algebra leads to $\kappa_1^* = O(n^{-2})$, $\kappa_3^* = O(n^{-3})$, $\kappa_4^* = O(n^{-4})$, and

$$\kappa_2^* = n^{-1} + \{(1/2)\alpha_4 - (1/3)\alpha_3^2 - 2a_1\}n^{-2} + O(n^{-3}).$$

Thus, by setting $a_1 = (1/2)b = (1/4)\alpha_4 - (1/6)\alpha_3^2$, which is half of the Bartlett correction factor, we get a simple second cumulant value $\kappa_2^* = n^{-1} + O(n^{-3})$. This leads to the Edgeworth expansion needed in the proofs of the theorems in the next subsection.

A.3 Edgeworth expansion and the proofs

The neat cumulants of $\sqrt{n}T^*$ with \mathbf{a} specified in Theorem 3.1 translate to a simple Edgeworth expansion as follows:

$$\Pr(\sqrt{n}T^* < y) = \int_{-\infty}^y \{1 + n^{-3/2}p(t)\}\phi(t)dt + O(n^{-2}),$$

for some odd-degree polynomial $p(t)$ (up to degree 9), where $\phi(t)$ is the density function of the standard normal distribution. We do not give details on how to obtain the Edgeworth expansion from the cumulants of T^* . In general, if two such statistics have the same first k cumulants, then they have the same Edgeworth expansion up to order $n^{-(k-2)/2}$; see Bhattacharya and Ghosh (1978). Clearly, $\sqrt{n}T^*$ has the same first four cumulants as the standard normal, which leads to the conclusion of Theorem 3.1. When $\alpha_2 \neq 1$, the argument follows with the data x_i 's replaced by $\alpha_2^{-1/2}x_i$'s. This completes the proof of Theorem 3.1.

Replacing a_2 in Theorem 3.1 by some root- n consistent estimator \hat{a}_2 makes $R(\mu_0, \hat{\mathbf{a}})$ different from $R(\mu_0, \mathbf{a})$ by a

quantity of order $O_p(n^{-1})$, which is too large to be ignored in the expansion. Replacing a_1 by a root-n consistent estimator has an effect only on the $O_p(n^{-3/2})$ term. We prove Theorem 3.2 by quantifying the difference between $R(\mu_0; (a_1, \hat{a}_2))$ and $R(\mu_0; (a_1, a_2))$ such that $a_2 = (1/6)\gamma\alpha_2^{1/2}$, with its replacement $\hat{a}_2 = (1/6)\hat{\alpha}_2^{-1}\hat{\alpha}_3 + (1/2)(\bar{x} - \mu)$, and $a_1 = O_p(1)$ to be determined. Keeping the convention that $\alpha_2 = 1$, it is easy to show that

$$\hat{a}_2 - a_2 = (1/6)(A_3 - \alpha_3 A_2) + O_p(n^{-1}).$$

It is straightforward to see that for any fixed a_1 , we have

$$\begin{aligned} R(\mu_0; (a_1, \hat{a}_2)) &= R(\mu_0; (a_1, a_2)) + 2(\hat{a}_2 - a_2)A_1 + O_p(n^{-3/2}) \\ &= R(\mu_0; (a_1, a_2)) + (1/3)(A_3 - \alpha_3 A_2)A_1 \\ &\quad + O_p(n^{-3/2}). \end{aligned}$$

As noticed earlier, the first two terms in the signed root of $R(\mu_0; \mathbf{a})$ do not depend on the value of a_1 provided it is $O_p(1)$. Because of this, the first two terms of the signed root of $R(\mu_0; \hat{\mathbf{a}})$ are the same as T_1^* and T_2^* . The third term is given by

$$\tilde{T}_3 = T_3^* + (1/6)(A_3 - \alpha_3 A_2)n^{-1}.$$

Some lengthy but simple algebra shows that the corresponding cumulants have order assessments $\tilde{\kappa}_1 = O(n^{-2})$, $\tilde{\kappa}_3 = O(n^{-3})$, $\tilde{\kappa}_4 = O(n^{-4})$, and

$$\tilde{\kappa}_2 = n^{-1} + \{(5/6)\alpha_4 - (2/3)\alpha_3^2 - 2a_1\}n^{-2} + O(n^{-3}).$$

Setting

$$a_1 = (5/12)\alpha_2^{-2}\alpha_4 - (1/3)\alpha_2^{-3}\alpha_3^2 = b - (1/12)\alpha_2^{-2}\alpha_4,$$

where we have restored α_2 for comprehension, we get $\tilde{\kappa}_2 = n^{-1} + O(n^{-3})$. The cumulant results lead to the conclusion of Theorem 3.2.

APPENDIX B. EXPANSIONS OF THE UPPER CONFIDENCE LIMITS

The upper confidence limits of EL intervals are obtained by finding the larger solution in μ to

$$R(\mu; \mathbf{a}) = \xi^2$$

where ξ^2 is some upper quantile of the χ_1^2 distribution. When $\mathbf{a} = 0$, we get the upper limits of the usual EL intervals. Thus, deriving a generic asymptotic expansion of $R(\mu; \mathbf{a})$ is the key step. The expansion of $R(\mu_0; \mathbf{a})$ is no longer useful.

We claim without proof that any of the upper confidence limits $\mu = \mu_0 + O_p(n^{-1/2})$ and that $\mathbf{a} = O_p(1)$. For μ and \mathbf{a} in this range, we further claim that the solution to

$$(9) \quad \sum_{i=1}^{n+2} \frac{x_i - \mu}{1 + \lambda(x_i - \mu)} = 0$$

satisfies $\lambda = O_p(n^{-1/2})$. These claims imply that $\bar{x} - \mu = O_p(n^{-1/2})$ because μ is in an $O_p(n^{-1/2})$ neighborhood of μ_0 .

Expanding (9) with respect to λ at $\lambda = 0$, we have

$$(10) \quad \sum_{i=1}^n \{(x_i - \mu) - \lambda(x_i - \mu)^2 + \lambda^2(x_i - \mu)^3\} + a_2 + O_p(n^{-1/2}) = 0.$$

Note that the $(n+1)$ th term in (9) is $O_p(n^{-1/2})$ and hence it has been counted in the above equation. It is easy to see that

$$\begin{aligned} \sum_{i=1}^n (x_i - \mu) &= n(\bar{x} - \mu) = nA_1; \\ \sum_{i=1}^n (x_i - \mu)^2 &= n\{\hat{\alpha}_2 + (\bar{x} - \mu)^2\} = n(\hat{\alpha}_2 + A_1^2); \\ \sum_{i=1}^n (x_i - \mu)^3 &= n\{\hat{\alpha}_3 + 3\hat{\alpha}_2 A_1\} + O_p(n^{-1/2}). \end{aligned}$$

Note that we borrowed an earlier notation $A_1 = \bar{x} - \mu$ but we do not require $\mu = \mu_0$ here. Substituting these into expansion (10), dividing both sides by n and keeping the leading terms only, we get

$$(11) \quad A_1 - \lambda\hat{\alpha}_2 + \lambda^2\hat{\alpha}_3 + a_2n^{-1} + O_p(n^{-3/2}) = 0.$$

A generic expansion for λ is then found as

$$(12) \quad \lambda = \hat{\alpha}_2^{-1}A_1 + \hat{\alpha}_2^{-3}\hat{\alpha}_3A_1^2 + \hat{\alpha}_2^{-1}a_2n^{-1} + O_p(n^{-3/2}).$$

Note that this expansion is not dependent on the size of a_1 up to order $n^{-3/2}$.

With the expansion for λ , we further obtain

$$\begin{aligned} R(\mu; \mathbf{a}) &= \sum_{i=1}^n \{2\lambda(x_i - \mu) - \lambda^2(x_i - \mu)^2 + (2/3)\lambda^3(x_i - \mu)^3 \\ &\quad - (1/2)\lambda^4(x_i - \mu)^4\} + \{2\lambda(x_{n+1} - \mu) \\ &\quad + 2\lambda(x_{n+2} - \mu) - \lambda^2(x_{n+2} - \mu)^2\} + O_p(n^{-3/2}) \\ &= n\{2\lambda A_1 - \lambda^2(\hat{\alpha}_2 + A_1^2) + (2/3)\lambda^3(\hat{\alpha}_3 + 3\hat{\alpha}_2 A_1) \\ &\quad - (1/2)\lambda^4\hat{\alpha}_4\} + \{-2\lambda a_1 A_1 + 2\lambda a_2 - \lambda^2 a_2^2\} \\ &\quad + O_p(n^{-3/2}). \end{aligned}$$

We now substitute the expansion of λ given by (12), and retain terms up to order n^{-1} , to obtain

$$\begin{aligned} R(\mu; \mathbf{a}) &= n\hat{\alpha}_2^{-1}A_1^2 + \{2\hat{\alpha}_2^{-1}a_2A_1 + (2/3)n\hat{\alpha}_2^{-3}\hat{\alpha}_3A_1^3\} \\ &\quad + (2\hat{\alpha}_2^{-3}\hat{\alpha}_3a_2 - 2\hat{\alpha}_2^{-1}a_1 - \hat{\alpha}_2^{-2}a_2^2)A_1^2 \\ &\quad + n\{\hat{\alpha}_2^{-5}\hat{\alpha}_3^2 + \hat{\alpha}_2^{-2} - (1/2)\hat{\alpha}_2^{-4}\hat{\alpha}_4\}A_1^4 \\ &\quad + \hat{\alpha}_2^{-1}a_2^2n^{-1} + O_p(n^{-3/2}) \end{aligned}$$

where the first line contains $O_p(1)$ and $O_p(n^{-1/2})$ terms, and the second and third lines contain $O_p(n^{-1})$ terms and the high-order residual $O_p(n^{-3/2})$.

The remaining task is to solve $R(\mu; \mathbf{a}) = \xi^2$ for μ based on the above expansion. This task is equivalent to finding a solution in A_1 . When only $O_p(1)$ terms are included in the equation, we quickly find that

$$A_1 = -n^{-1/2}\hat{\alpha}_2^{1/2}\xi + O_p(n^{-1}).$$

Recall that $A_1 = \bar{x} - \mu$ and that $\hat{\alpha}_2 = s_n^2$, the sample variance. Setting $\xi = z_{1-\alpha}$, this solution becomes

$$\mu_U = \bar{x} + n^{-1/2}z_{1-\alpha}s_n + O_p(n^{-1}),$$

which is the upper confidence limit based on the normality assumption.

We can refine the above solution to obtain higher-order terms. Setting

$$A_1 = -n^{-1/2}\hat{\alpha}_2^{1/2}\xi + \epsilon_1$$

we obtain an equation for ϵ_1 :

$$n\hat{\alpha}_2^{-1}A_1^2 + \{2\hat{\alpha}_2^{-1}a_2A_1 + (2/3)n\hat{\alpha}_2^{-3}\hat{\alpha}_3A_1^3\} = \xi^2 + O_p(n^{-1}).$$

The solution for ϵ_1 is

$$\epsilon_1 = -\{a_2 + (1/3)\hat{\alpha}_2^{-1}\hat{\alpha}_3\xi^2\}n^{-1} + \epsilon_2$$

with $\epsilon_2 = O_p(n^{-3/2})$. Repeating this process for ϵ_2 , we find

$$\begin{aligned} \epsilon_2 = & \left[-\{a_1\hat{\alpha}_2^{1/2} + (1/2)\hat{\alpha}_2^{-1/2}a_2^2\}\xi \right. \\ & \left. + \{(1/2)\hat{\alpha}_2^{1/2} + (2/9)\hat{\alpha}_2^{-5/2}\hat{\alpha}_3^2 - (1/4)\hat{\alpha}_2^{-3/2}\hat{\alpha}_4\}\xi^3 \right] \\ & \times n^{-3/2} + O_p(n^{-2}). \end{aligned}$$

Consequently, we obtain a generic expansion for the upper confidence limit:

$$\begin{aligned} \mu_U = & \bar{x} + n^{-1/2}\hat{\alpha}_2^{1/2}\xi + n^{-1}\{a_2 + (1/3)\hat{\alpha}_2^{-1}\hat{\alpha}_3\xi^2\} \\ & + n^{-3/2}\left[\{a_1\hat{\alpha}_2^{1/2} + (1/2)\hat{\alpha}_2^{-1/2}a_2^2\}\xi \right. \\ & \left. + \{-(2/9)\hat{\alpha}_2^{-5/2}\hat{\alpha}_3^2 - (1/2)\hat{\alpha}_2^{1/2} + (1/4)\hat{\alpha}_2^{-3/2}\hat{\alpha}_4\}\xi^3\right] \\ & + O_p(n^{-2}). \end{aligned}$$

Let $\hat{\gamma} = \hat{\alpha}_3/\hat{\alpha}_2^{3/2}$ and $\hat{\kappa} = \hat{\alpha}_4/\hat{\alpha}_2^2 - 3$. The upper limit becomes

$$\begin{aligned} \mu_U = & \bar{x} + n^{-1/2}\hat{\alpha}_2^{1/2}\xi + n^{-1}\{a_2 + (1/3)\hat{\alpha}_2^{1/2}\hat{\gamma}\xi^2\} \\ & + n^{-3/2}\left[\{a_1\hat{\alpha}_2^{1/2} + (1/2)\hat{\alpha}_2^{-1/2}a_2^2\}\xi \right. \\ & \left. + \{(1/4)(1 + \hat{\kappa})\hat{\alpha}_2^{1/2} - (2/9)\hat{\gamma}^2\hat{\alpha}_2^{1/2}\}\xi^3\right] \\ & + O_p(n^{-2}). \end{aligned}$$

Setting $a_1 = a_2 = 0$ and $\xi = z_{1-\alpha}$, the above generic expansion results in the U_{EL} obtained via ordinary EL. The effect of the direct Bartlett correction is to replace $\xi = z_{1-\alpha}$ by

$\xi^2 = (1 + b/n)z_{1-\alpha}^2$. Asymptotically, it affects only the order $n^{-3/2}$ term:

$$U_{BEL} = U_{EL} + n^{-3/2}s_n z_{1-\alpha}(b/2).$$

Recall again that $s_n = \hat{\alpha}_2^{1/2}$.

The upper limit U_{AEL} is the solution to the same equation when $a_1 = \hat{b}/2 = (1/12)\hat{\alpha}_2^{-2}\hat{\alpha}_4$, and $a_2 = (1/6)\hat{\gamma}\hat{\alpha}_2^{1/2} + (1/2)A_1$. Because a_2 depends on A_1 , the previous derivation must be slightly revised. With these specific a_1 and a_2 , we find that

$$\begin{aligned} R(\mu; \mathbf{a}) = & nA_1^2 + (1/3)\alpha_3A_1(1 + 2nA_1^2) \\ & + \{(1/2) - (35/36)\alpha_3^2 - (1/6)\alpha_4\}A_1^2 \\ & + n\{1 + \alpha_3^2 - (1/2)\alpha_4\}A_1^4 + O_p(n^{-3/2}). \end{aligned}$$

Setting

$$A_1 = -n^{-1/2}\hat{\alpha}_2^{1/2}\xi - (1/6)n^{-1}(1 + 2\xi^2)\hat{\alpha}_2^{-1}\hat{\alpha}_3 + \epsilon_2$$

and solving $R(\mu; \mathbf{a}) = \xi^2 + O_p(n^{-3/2})$ for A_1 , we get

$$\begin{aligned} A_1 = & -n^{-1/2}\hat{\alpha}_2^{1/2}\xi - (1/6)n^{-1}\hat{\alpha}_2^{1/2}(1 + 2\xi^2)\hat{\alpha}_2^{-3/2}\hat{\alpha}_3 \\ & - n^{-3/2}\hat{\alpha}_2^{1/2}\left[\{(1/4) - (47/72)\hat{\alpha}_2^{-3}\alpha_3^2 \right. \\ & \left. - (1/12)\hat{\alpha}_2^{-2}\alpha_4\}\xi + \{(1/2) - (5/18)\hat{\alpha}_2^{-3}\alpha_3^2 \right. \\ & \left. - (1/4)\hat{\alpha}_2^{-2}\alpha_4\}\xi^2\right]. \end{aligned}$$

Replacing A_1 by $\bar{x} - \mu$ leads to the expansion of U_{AEL} given in Section 4, after $\hat{\alpha}_2^{-3/2}\hat{\alpha}_3$ is denoted as $\hat{\gamma}$ and so on.

ACKNOWLEDGEMENTS

The authors thank the editor, an associate editor and an anonymous referee for their thorough reviews and insightful comments.

Received 1 June 2011

REFERENCES

- [1] BHATTACHARYA, R. N. and GHOSH, J. K. (1978). On the validity of the Edgeworth expansion. *The Annals of Statistics* **6** 431–451. [MR0471142](#)
- [2] CHEN, J., CHEN, S. and RAO, J. N. K. (2003). Empirical likelihood confidence intervals for the mean of a population containing many zero values. *The Canadian Journal of Statistics* **31** 53–68. [MR1985504](#)
- [3] CHEN, J., VARIYATH, A. M. and ABRAHAM, B. (2008). Adjusted empirical likelihood and its properties. *Journal of Computational and Graphical Statistics* **17** 426–443. [MR2439967](#)
- [4] CHEN, S. X. and CUI, H. J. (2006). On Bartlett correction of empirical likelihood in the presence of nuisance parameters. *Biometrika* **93** 215–220. [MR2277752](#)
- [5] CHEN, S. X. and CUI, H. J. (2007). On the second-order properties of empirical likelihood with moment restrictions. *Journal of Econometrics* **141** 492–516. [MR2413478](#)

- [6] CHEN, S. X. and QIN, J. (2003). Empirical likelihood-based confidence intervals for data with possible zero observations. *Statistics and Probability Letters* **65** 29–37. [MR2012622](#)
- [7] DiCICCO, T. J., HALL, P. and ROMANO, J. P. (1991). Empirical likelihood is Bartlett-correctable. *The Annals of Statistics* **19** 1053–1061. [MR1105861](#)
- [8] EFRON, B. (1987). Better bootstrap confidence intervals. With comments and a rejoinder by the author. *Journal of the American Statistical Association* **82** 171–200. [MR0883345](#)
- [9] HALL, P. (1988). Theoretical comparison of bootstrap confidence intervals. With a discussion and a reply by the author. *The Annals of Statistics* **16** 927–985. [MR0959185](#)
- [10] HALL, P. (1992). On the removal of skewness by transformation. *Journal of the Royal Statistical Society Series B* **54** 221–228. [MR1157721](#)
- [11] HALL, P. and LA SCALA, B. (1990). Methodology and algorithms of empirical likelihood. *International Statistical Review* **58** 109–127.
- [12] JOHNSON, N. J. (1978). Modified t tests and confidence intervals for asymmetrical populations. *Journal of the American Statistical Association* **73** 536–544. [MR0514154](#)
- [13] KVANLI, A. H., SHEN, Y. K. and DENG, L. Y. (1998). Construction of confidence intervals for the mean of a population containing many zero values. *Journal of Business & Economic Statistics* **16** 362–368.
- [14] LI, X., CHEN, J., WU, Y. and TU, D. (2011). Constructing non-parametric likelihood confidence regions with high-order precisions. *Statistica Sinica* **21** 1767–1783.
- [15] LIU, Y. and CHEN, J. (2010). Adjusted empirical likelihood with high-order precision. *The Annals of Statistics* **38** 1341–1362. [MR2662345](#)
- [16] LOH, W. Y. (1991). Bootstrap calibration for confidence interval construction and selection. *Statistica Sinica* **1** 477–491. [MR1130129](#)
- [17] OWEN, A. B. (1990). Empirical likelihood ratio confidence regions. *The Annals of Statistics* **18** 90–120. [MR1130129](#)
- [18] OWEN, A. B. (2001). *Empirical Likelihood*. Chapman and Hall/CRC, New York.
- [19] QIN, J. and LAWLESS, J. (1994). Empirical likelihood and general equations. *The Annals of Statistics* **22** 300–325. [MR1272085](#)

Jiahua Chen
 Department of Statistics
 University of British Columbia
 Vancouver, BC, Canada V6T 1Z2
 E-mail address: jhchen@stat.ubc.ca

Yukun Liu
 Department of Statistics and Actuarial Sciences
 School of Finance and Statistics
 East China Normal University
 Shanghai 200241, China
 E-mail address: ykliu@sfs.ecnu.edu.cn