Communications in Statistics—Simulation and Computation®, 38: 2171–2183, 2009

ISSN: 0361-0918 print/1532-4141 online
DOI: 10.1080/03610910903259659

Modifications of the Empirical Likelihood Interval Estimation with Improved Coverage Probabilities

ALBERT VEXLER, SHULING LIU, LE KANG, AND ALAN DAVID HUTSON

Department of Biostatistics, The State University of New York at Buffalo, Buffalo, New York, USA

The empirical likelihood (EL) technique has been well addressed in both the theoretical and applied literature in the context of powerful nonparametric statistical methods for testing and interval estimations. A nonparametric version of Wilks theorem (Wilks, 1938) can usually provide an asymptotic evaluation of the Type I error of EL ratio-type tests. In this article, we examine the performance of this asymptotic result when the EL is based on finite samples that are from various distributions. In the context of the Type I error control, we show that the classical EL procedure and the Student's t-test have asymptotically a similar structure. Thus, we conclude that modifications of t-type tests can be adopted to improve the EL ratio test. We propose the application of the Chen (1995) t-test modification to the EL ratio test. We display that the Chen approach leads to a location change of observed data whereas the classical Bartlett method is known to be a scale correction of the data distribution. Finally, we modify the EL ratio test via both the Chen and Bartlett corrections. We support our argument with theoretical proofs as well as a Monte Carlo study. A real data example studies the proposed approach in practice.

Keywords Bartlett Correction; Empirical likelihood; Interval estimation; Likelihood ratio; Modified *t*-test; Nonparametric test; Significance level; Type I error.

Mathematics Subject Classification 62G10; 62G20.

1. Introduction

Likelihood methods are common statistical tools that provide very powerful methods of statistical inference. The parametric likelihood methodology requires the assumption that the distribution function of the individual observations have known forms (e.g., Vexler et al., 2009). To relax the parametric assumptions, Owen (1988) introduced the empirical likelihood (EL) methodology as a nonparametric version of the parametric likelihood technique. In a wide range of applications,

Received March 13, 2009; Accepted August 13, 2009
Address correspondence to Albert Vexler, Department of Biostatistics, The State University of New York, Farber Hall, Room 246, Buffalo, NY 14214, USA; E-mail: avexler@buffalo.edu

the nonparametric EL ratio tests have comparable advantages to those of the parametric likelihood methodologies. EL methods have been well accepted in the literature as very powerful statistical procedures in the distribution-free setting relative to parametric approaches; e.g., see DiCiccio et al. (1991), Hall and Scala (1990), Owen (1988, 1990, 1991, 2001), Qin and Lawless (1994), and Yu et al. (2009). Other results include that of Owen (1990) who proved that the asymptotic distribution function of twice the logarithm of an EL ratio-type test has the form of a chi-square distribution function. This asymptotic result is a nonparametric version of the Wilks' theorem (e.g., Wilks, 1938), which is known in the context of the asymptotic null distribution of the maximum parametric likelihood ratio test statistic. However, it is well known that the distance between the actual null distribution of a test statistic and its asymptotic approximation can be significant. To acquire the Type I error control related to the EL ratio test, the Bartlett correction can be applied; e.g., see Owen (2001). The Bartlett correction is made by multiplying the relevant test thresholds by a scale parameter, values of which can be chosen to improve upon Wilks' approximation.

In practice, a finite sample size strongly restricts the Type I error control that is obtained through Wilks' approximation. In this case, the application of Bartlett's correction can be also difficult. This is due to the fact that this correction is based upon asymptotic principles, which in turn depend functionally on unknown parameters, such as the 3rd and 4th unknown moments of the underlying population, and hence have to be estimated. When the sample size is relatively small, the unknown parameters utilized within Bartlett's correction can be poorly estimated such that the efficiency of the Bartlett's technique deteriorates. It should also be noted that the EL ratio test statistic in certain situations is not Bartlett correctable; e.g., see Chen and Cui (2006) and Lazar and Mykland (1999). Corcoran (1998), as well as Davison and Stafford (1998), remarked that the Bartlett correction applied to EL ratios does not provide significantly effective results in certain practical situations. In this article, we investigate the Bartlett-corrected EL ratio test based on samples with finite sizes.

Data with skewed distributions can also affect the efficiency of the EL method. In this article, utilizing both theoretical and empirical arguments, we show that, in the context of the Type I error control, when the data is non normally distributed and skewed (especially when the sample size is relatively small), the EL ratio test can be corrected in a similar manner to that of a Student's t-test correction. That is, if some correction can improve Student's t-type tests, then it can also be adopted to the EL ratio-type test. Chen (1995) proposed a modification of the t-test that was shown to be more accurate and more powerful than both the Johnson's modified t-test and the Sutton's composite test. The Chen approach is very efficient for small sample sizes. Thus, we propose to apply the Chen (1995) method to modify the EL ratio test. This correction is different from the Bartlett modification that rescales the log EL ratio in order to improve the second-order approximation via a chi-square distribution. Alternatively, we show that Chen correction transforms observations, changing the location of data. To take advantages of both the scale and location corrections, we propose a new modification of the EL ratio test combining the Chen and Bartlett techniques.

To achieve our goals mentioned above, sketch theoretical proofs of the new modifications are provided. A Monte Carlo study is used to evaluate the control of Type I error related to the proposed procedures. A real data example is provided in order to illustrate the proposed method. Section 5 contains our concluding remarks.

2. Methods

Assume, for simplicity, we observe independent identically distributed (i.i.d.) random variables $X_1, X_2, \ldots, X_n$ with $E|X_1|^3 < \infty$. We would like to test the hypothesis

$$H_0 : E(X_1) = \theta \quad \text{vs.} \quad H_1 : E(X_1) \neq \theta. \tag{1}$$

Furthermore, before we proceed note that the hypothesis above can be extended easily to the more general case, namely,

$$H_0 : E(g(X)) = \theta \quad \text{vs.} \quad H_1 : E(g(X)) \neq \theta,$$

where g is a given function. To test the hypothesis above defined at (1) nonparametrically, first define the EL function in the form of

$$L(X_1, \ldots, X_n \mid \theta) = \prod_{i=1}^{n} p_i,$$

where $\sum_{i=1}^{n} p_i = 1$. (Here, p_i's, $i = 1, \ldots, n$, have a sense of $\Pr\{X_i = x_i\}$.) The values of p_i's are unknown and should be obtained under the null or the alternative hypothesis. Under the null hypothesis, the EL methodology (e.g., Owen, 2001; Yu et al., 2009) requires one to find the value of the p_i's that maximize the EL that satisfy the empirical constraints $\sum_{i=1}^{n} p_i = 1$ and $\sum_{i=1}^{n} p_i X_i = \theta$. Using Lagrange multipliers, one can show that under the null hypothesis, the maximum EL function has the following form:

$$L(X_1, \ldots, X_n \mid \theta) = \sup_{\substack{0<p_1,p_2,\ldots,p_n<1 \\ \Sigma p_i=1,\ \Sigma p_i X_i=\theta}} \prod_{i=1}^{n} p_i = \prod_{i=1}^{n} \frac{1}{n[1 + \lambda(X_i - \theta)]}, \tag{2}$$

where λ is a root of

$$\sum_{i=1}^{n} \frac{X_i - \theta}{1 + \lambda(X_i - \theta)} = 0.$$

Under the alternative hypothesis, just the empirical constraint $\sum_{i=1}^{n} p_i = 1$ is in effect. Then

$$L(X_1, \ldots, X_n) = \sup_{\substack{0<p_1,p_2,\ldots,p_n<1 \\ \Sigma p_i=1}} \prod_{i=1}^{n} p_i = \prod_{i=1}^{n} \frac{1}{n} = \left(\frac{1}{n}\right)^n. \tag{3}$$

Combining (2) and (3), we obtain the log EL ratio, test statistic for (1) in the form of

$$l(X_1, \ldots, X_n \mid \theta) = -2 \log \frac{L(X_1, \ldots, X_n \mid \theta)}{L(X_1, \ldots, X_n)} = 2 \sum_{i=1}^{n} \log[1 + \lambda(X_i - \theta)]. \tag{4}$$

Owen (1988) proved that $l(X_1, \ldots, X_n \mid \theta)$ has an asymptotic chi-square distribution, under the null hypothesis. Hence, we can test H_0 at the approximated significance level α using the test

$$l(X_1, \ldots, X_n \mid \theta) = 2\sum_{i=1}^{n} \log[1 + \lambda(X_i - \theta)] > C_\alpha,$$

where the test threshold C_α is the $100(1-\alpha)\%$ percentile for a chi-square distribution.

Note that one can show

$$l(X_1, \ldots, X_n \mid \theta) \approx \frac{[\sum_{i=1}^{n}(X_i - \theta)]^2}{\sum_{i=1}^{n}(X_i - \theta)^2} \quad \text{as } n \to \infty, \tag{5}$$

where $[\sum_{i=1}^{n}(X_i - \theta)]^2 / \sum_{i=1}^{n}(X_i - \theta)^2$ has an asymptotic chi-square distribution under the null hypothesis given at (1). The asymptotic proposition (5) and its proof are similar to results that were demonstrated and utilized by Qin and Lawless (1994) in a different context. For details, see the Appendix.

This asymptotic result motivates us to consider the t-test statistic

$$T(X_1, \ldots, X_n \mid \theta) = \frac{\sum_{i=1}^{n}(X_i - \theta)}{\sqrt{\sum_{i=1}^{n}(X_i - \overline{X})^2}} \quad \left(\overline{X} = \sum_{i=1}^{n} X_i / n\right),$$

which is asymptotically equivalent to 2 log EL ratio, i.e., by virtue of (5) one can show that the 2 log EL ratio is approximated by T^2 that in turn converges in distribution to a chi-square distribution. This indicates that the EL ratio test and the Student's t-test have similar properties in the context of the Type I error control. Note that Owen (1990) concluded that large sample sizes strongly improve the accuracy of the Wilks-type approximations of the distribution of null EL ratio test statistic. In the context of the Type I error control, we demonstrate that the classical EL procedure and Student's t-test based on observations from the same distributions have similar properties. That is, when data are close to being normally distributed, the Type I error control related to the tests based on the data is expected to be accurate even when n, the sample size, is small. However, when the sample size is large and the observations are from a lognormal distribution, neither the t-test nor the EL ratio test control the Type I error well. In this case, both tests can be affected strongly by the skewness of data. In Sec. 3, we confirm this conclusion via a Monte Carlo study.

In certain situations, the EL ratio test is Bartlett correctable; e.g., see Owen (2001). The Bartlett correction is a theoretical adjustment for the second-order approximation of a log-likelihood ratio distribution; e.g., see Hall and Scala (1990). However, the Bartlett correction cannot completely solve the problem of the Type I error control when the data is skewed. Moreover, when the Bartlett correction is applied certain unknown parameters need to be estimated. The accuracy of this estimation can affect the overall efficiency of the Bartlett correction. Note that Lazar and Mykland (1999) proved that when nuisance parameters are present the EL ratio statistic is not Bartlett correctable. When the Bartlett correction is in effect, the test statistic is multiplied by a coefficient. This rescaling of the log likelihood ratio is used to reduce the discrepancy between a chi-square distribution and the asymptotic null distribution of the log likelihood ratio. Thus, we refer to this a scale correction. Alternatively, since some modifications of the t-test, different from the Bartlett methodology, can improve the relevant Type I error control, we can consider a similar modification of the EL ratio test following the asymptotic equivalence between the two tests mentioned above. We outline the approach of Chen, showing this method changes the location of the sample instead of providing a scaling correction. In the context of the hypothesis

$$H_0 : E(X) = \theta \quad \text{vs.} \quad H_1 : E(X) > \theta, \tag{6}$$

define $\hat{\beta}_1 = \sum_{i=1}^{n}(X_i - \overline{X})^3/nS^3$, where $S = \left(\sum_{i=1}^{n}(X_i - \overline{X})^2/(n-1)\right)^{0.5}$. Then the Chen modified t-test can be presented as

$$T(X_1, \ldots, X_n \mid \theta) + \Delta(X_1, \ldots, X_n \mid \theta), \tag{7}$$

where

$$\Delta(X_1, \ldots, X_n \mid \theta) = \frac{1}{6\sqrt{n}}\hat{\beta}_1 \left[1 + 2\left\{\frac{\sqrt{n}(\overline{X} - \theta)}{S}\right\}^2\right] + \frac{1}{9n}\hat{\beta}_1^2 \left[\frac{\sqrt{n}(\overline{X} - \theta)}{S} + 2\left\{\frac{\sqrt{n}(\overline{X} - \theta)}{S}\right\}^2\right].$$

This correction was obtained using an Edgeworth expansion applied to the t-test when the distribution of data is positively skewed (Chen, 1995). The test statistic (7) can be represented by the regular t-test statistic based on observations

$$X_i^* = X_i + \frac{\Delta(X_1, \ldots, X_n \mid \theta)}{\sqrt{n}} S, \quad i = 1, \ldots, n. \tag{8}$$

Thus, the Chen approach changes the location of X_i's. Therefore, we can classify the Chen method as a location correction.

Assume we change the location of X_i's and define the EL ratio based on relocated data and derive an optimal change of the location of the X_i's. Towards this end, the EL ratio test is created utilizing the constraints

$$\sum_{i=1}^{n}(X_i + d)p_i = \theta \quad \text{and} \quad \sum_{i=1}^{n} p_i = 1,$$

under the null H_0, where d is a shift of the location. Let us obtain a form of d to correct the EL ratio test statistic following the Chen approach. By virtue of (5), we have the Type I error

$$P_{H_0}\{l(X_1, \ldots, X_n \mid \theta) > C_\alpha\} \cong P_{H_0}\left\{\frac{\left[\sum_{i=1}^{n}(X_i + d - \theta)\right]^2}{\sum_{i=1}^{n}(X_i + d - \theta)^2} > C_\alpha\right\}$$
$$= P_{H_0}\left\{R(d, \theta) > \sqrt{C_\alpha} \quad \text{or} \quad R(d, \theta) < -\sqrt{C_\alpha}\right\}, \tag{9}$$

where

$$R(d, \theta) = \frac{\sum_{i=1}^{n}(X_i + d - \theta)}{\sqrt{\sum_{i=1}^{n}(X_i + d - \theta)^2}}$$
$$= \frac{\sqrt{n}(\overline{X} - \theta)}{S}\left(\frac{S}{\sqrt{\sum_{i=1}^{n}(X_i + d - \theta)^2/n}}\right) + \frac{d_n}{\sqrt{\sum_{i=1}^{n}(X_i + d - \theta)^2}}.$$

Using the Taylor expansion with d around 0, we obtain that

$$R(d, \theta) = T(X_1, \ldots, X_n \mid \theta) + \frac{d_n}{\sqrt{\sum_{i=1}^{n}(X_i - \theta)^2}} + o(d^2) \quad \text{as } d \to 0.$$

Here, the corresponding remained term has an order of d^2 and can be ignored.

Since $R(d, \theta) > \sqrt{C_\alpha}$ and $R(d, \theta) \approx T(X_1, \ldots, X_n \mid \theta) + d_n/\sqrt{\sum_{i=1}^{n}(X_i - \theta)^2}$, then by the Chen methodology we must set up

$$d : \frac{d_n^{0.5}}{\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \theta)^2\right)^{0.5}} = \Delta(X_1, \ldots, X_n \mid \theta)$$

as in (7) in order to correct the Type I error. Thus, we conclude that the optimal change in the data location given the rule $R(d, \theta) > \sqrt{C_\alpha}$ is

$$d = \frac{1}{\sqrt{n}}\Delta(X_1, \ldots, X_n \mid \theta)S.$$

By virtue of (9), and in order to approximate the decision rule $\{l(X \mid \theta) > C_\alpha\}$, we should also consider $\{R(d, \theta) < -\sqrt{C_\alpha}\}$. In a similar manner to the analysis mentioned above, for a negatively skewed population, we should set up

$$d = \frac{1}{\sqrt{n}}\Delta(-X_1, \ldots, -X_n \mid \theta)S.$$

Here, when we considered $\{R(d, \theta) < -\sqrt{C_\alpha}\}$, to preserve the statement of (6), we utilized observations $-X_1, \ldots, -X_n$ keeping the sign '>' of the alternative hypothesis of (6). That explains the form of $d = n^{-0.5}\Delta(-X_1, \ldots, -X_n \mid \theta)S$, comparing with the location change $n^{-0.5}\Delta(X_1, \ldots, X_n \mid \theta)S$ given the rule $\{R(d, \theta) > \sqrt{C_\alpha}\}$.

Thus, we propose the following modification of the EL ratio test:

1. Calculate values of the statistics

$$l_1(\theta) = l\left(X_1 + \frac{\Delta(X_1, \ldots, X_n \mid \theta)}{\sqrt{n}}S, \ldots, X_n + \frac{\Delta(X_1, \ldots, X_n \mid \theta)}{\sqrt{n}}S \,\middle|\, \theta\right) \quad \text{and}$$
$$l_2(\theta) = l\left(X_1 + \frac{\Delta(-X_1, \ldots, -X_n \mid \theta)}{\sqrt{n}}S, \ldots, X_n + \frac{\Delta(-X_1, \ldots, -X_n \mid \theta)}{\sqrt{n}} \,\middle|\, \theta\right).$$

2. Reject the null hypothesis of (1) if

$$(l_1(\theta))^{0.5} \geq Z\left(1 - \frac{\alpha}{2}\right) \quad \text{or} \quad (l_2(\theta))^{0.5} \geq Z\left(1 - \frac{\alpha}{2}\right),$$

where $Z(1-\alpha)$ is the $100(1-\alpha)$ percentile of the standard normal distribution. Note that, if we can presuppose the underlying distribution function of observations is positively or negatively skewed, we then can define the test statistic

$$\begin{cases} l(\theta) = l\left(X_1 + \dfrac{\Delta(X_1, \ldots, X_n \mid \theta)}{\sqrt{n}} S, \ldots, X_n + \dfrac{\Delta(X_1, \ldots, X_n \mid \theta)}{\sqrt{n}} S \,\middle|\, \theta\right) \\ \quad \text{for positively skewed data;} \\ l(\theta) = l\left(X_i + \dfrac{\Delta(-X_1, \ldots, -X_n \mid \theta)}{\sqrt{n}} S, \ldots, X_n + \dfrac{\Delta(-X_1, \ldots, -X_n \mid \theta)}{\sqrt{n}} S \,\middle|\, \theta\right) \\ \quad \text{for negatively skewed data,} \end{cases}$$

respectively, and reject the null hypothesis if $l(\theta) \geq \chi^2(1, 1-\alpha)$, where $\chi^2(1, 1-\alpha)$ is the $100(1-\alpha)\%$ percentile of the chi-square distribution with the degree of freedom of 1.

Hence, we showed that the Chen approach can be also adapted to be applied to the EL ratio test.

In accordance with the previously mentioned remarks, we can consider the Bartlett and Chen techniques as two different test-modifications. The Bartlett correction can reduce the slope caused by the difference between the chi-square distribution and the log EL ratio distribution and the Chen approach corrects the bias caused by the skewness of a null distribution of observations. Thus, we conclude to combine these modifications and to propose the next decision rule. We suggest rejecting the null hypothesis at (1) if

$$(l_1(\theta))^{0.5} \geq Z\left(1 - \frac{\alpha}{2}\right)\left(1 + \frac{a}{n}\right)^{-0.5} \quad \text{or} \quad (l_2(\theta))^{0.5} \geq Z\left(1 - \frac{\alpha}{2}\right)\left(1 + \frac{a}{n}\right)^{-0.5},$$

where a is an estimator of

$$\frac{1}{2}\frac{E((X-EX)^4)}{(E((X-EX)^2))^2} - \frac{1}{3}\frac{(E((X-EX)^3))^2}{(E((X-EX)^2))^3}.$$

Note that, by virtue of the relation between the testing and interval estimation, we can obtain the confidence interval estimator of θ in the form of

$$\mathrm{CI}_{1-\alpha} = \{\theta(\mathbf{X}) \mid TS(\theta) \leq C_{1-\alpha}\},$$

where $TS(\theta)$ presents a test statistic for the hypothesis (1) and $C_{1-\alpha}$ is the corresponding threshold, assuming that the nominal coverage probability is $1-\alpha$. Therefore, the previoulsy considered issue of the Type I error control is also relevant to a problem to improve the accuracy of the interval estimation.

3. Monte Carlo Study

In this section, we provide a Monte Carlo study to compare the following five tests: (1) the standard EL ratio test; (2) the proposed Chen-corrected EL ratio test; (3) the Bartlett-corrected EL ratio test; (4) the proposed Chen–Bartlett-corrected EL ratio test; and (5) the Student's t-test. Towards this end, the null distribution functions were chosen to be in Normal and Lognormal forms, since the asymptotic result mentioned in Sec. 2 showed that we can expect a good approximation to the test-statistic distribution when data are from a normally distributed population, whereas the approximation can be expected to be worse in the case with lognormally distributed data. We also generated data following a chi-square distribution to investigate the proposed method. Random variables were drawn from the null distributions with different parameters. At each baseline distribution and each set of parameters, we repeated 10,000 times data generations. The Monte Carlo Type I errors are presented in Tables 1–3.

As is shown in Table 1, the t-test demonstrates the actual Type I errors that are closer to the expected 0.05 than those of the considered tests, under the normal distribution assumption. The standard EL ratio method has well controlled Type I errors when the sample sizes are greater than 25. The Bartlett and Chen corrections are assumed to reduce a bias caused by the difference between the chi-square distribution and log EL ratio asymptotic distribution. In this case, this estimation is not necessary; however, the proposed Chen–Bartlett correction is readily available.

Table 2 displays situations when data follow a lognormal distribution. The notation $LN(a, b)$ corresponds to a lognormal distribution where a and b are the mean and standard deviation of the logarithm. In accordance with the analysis mentioned in Sec. 2, we can expect corrections of the test statistics are needed in this case. Here, the Monte Carlo study confirms this conclusion. The Chen–Bartlett-corrected EL ratio test is the best of all the five test statistics. Especially when the data are strongly skewed, e.g., in the cases with $LN(1, 2)$, $LN(2, 2)$, $n = 25$, the Monte Carlo Type I errors for the EL ratio test and the t-test are greater than 0.3 (we expected the significance level at 0.05). When the combination of Chen

Table 1
The Monte Carlo Type I Errors of the tests based on data from the null normal distributions: $X \sim \text{Normal}(0, \sigma^2)$and $H_0 : EX = 0$ (The expected Type I error is 0.05)

Sample size n	σ	ELR	Chen	Bartlett	Chen & Bartlett	T-test
25	0.5	0.0634	0.0659	0.0563	0.0605	0.0511
	1.0	0.0636	0.0669	0.0581	0.0614	0.0514
	1.5	0.0642	0.0674	0.0573	0.0608	0.0498
	2.0	0.0646	0.0652	0.0581	0.0602	0.0500
35	0.5	0.0580	0.0614	0.0529	0.0573	0.0501
	1.0	0.0595	0.0602	0.0557	0.0565	0.0511
	1.5	0.0599	0.0617	0.0551	0.0573	0.0511
	2.0	0.0583	0.0593	0.0544	0.0555	0.0505
50	0.5	0.0535	0.0549	0.0501	0.0520	0.0490
	1.0	0.0551	0.0562	0.0507	0.0524	0.0497
	1.5	0.0598	0.0612	0.0557	0.0567	0.0542
	2.0	0.0582	0.0583	0.0547	0.0547	0.0549
100	0.5	0.0495	0.0492	0.0482	0.0479	0.0476
	1.0	0.0517	0.0524	0.0497	0.0500	0.0490
	1.5	0.0523	0.0525	0.0511	0.0506	0.0514
	2.0	0.0564	0.0561	0.0546	0.0548	0.0543

Table 2
The Monte Carlo Type I errors corresponded to null lognormal distributions: $X \sim \text{Lognormal}(\mu, \sigma^2) - \exp(\mu + 0.5\sigma^2)$ and $H_0 : EX = 0$ (The expected Type I error is 0.05)

Sample size n	(μ, σ)	ELR	Chen	Bartlett	Chen & Bartlett	T-test
25	(0, 1.0)	0.1229	0.1118	0.1104	0.1035	0.1264
	(0, 1.5)	0.2006	0.1463	0.1858	0.1386	0.2278
	(0, 2.0)	0.3221	0.2064	0.3028	0.1991	0.3641
	(1, 1.0)	0.1213	0.1096	0.1145	0.1005	0.1302
	(1, 1.5)	0.2024	0.1283	0.1853	0.1185	0.2253
	(1, 2.0)	0.3221	0.1671	0.2890	0.1719	0.3574
	(2, 2.0)	0.3197	0.1736	0.2869	0.1641	0.3573
35	(0, 1.0)	0.1061	0.1051	0.0966	0.0965	0.1108
	(0, 1.5)	0.1756	0.1226	0.1626	0.1139	0.1997
	(0, 2.0)	0.2852	0.1839	0.2632	0.1748	0.3329
	(1, 1.0)	0.1085	0.0938	0.0987	0.0850	0.1163
	(1, 1.5)	0.1760	0.1127	0.1611	0.1039	0.2011
	(1, 2.0)	0.2851	0.1566	0.2675	0.1499	0.3337
	(2, 2.0)	0.2855	0.1533	0.2680	0.1438	0.3333
50	(0, 1.0)	0.0851	0.0830	0.0771	0.0695	0.0933
	(0, 1.5)	0.1561	0.1186	0.1450	0.1111	0.1853
	(0, 2.0)	0.2534	0.1574	0.2335	0.1506	0.3074
	(1, 1.0)	0.0950	0.0911	0.0867	0.0831	0.1034
	(1, 1.5)	0.1537	0.1067	0.1420	0.0987	0.1836
	(1, 2.0)	0.2571	0.1381	0.2406	0.1289	0.3042
	(2, 2.0)	0.2554	0.1258	0.2381	0.1164	0.3077
100	(0, 1.0)	0.0688	0.0821	0.0713	0.0676	0.0750
	(0, 1.5)	0.1222	0.0970	0.1137	0.0899	0.1514
	(0, 2.0)	0.2040	0.1151	0.1899	0.1076	0.2550
	(1, 1.0)	0.0809	0.0842	0.0750	0.0780	0.0877
	(1, 1.5)	0.1265	0.1007	0.1181	0.0930	0.1567
	(1, 2.0)	0.2030	0.1139	0.1896	0.1064	0.2541
	(2, 2.0)	0.2042	0.1071	0.1895	0.1005	0.2582

Table 3
The Monte Carlo Type I errors corresponded to null chi-square distributions: $X \sim \chi^2(1) - 1$ and $H_0 : EX = 0$ (The expected Type I error is 0.05)

Sample size n	ELR	Chen	Bartlett	Chen & Bartlett	T-test
25	0.0986	0.0907	0.0911	0.0844	0.0953
35	0.0840	0.0777	0.0788	0.0721	0.0892
50	0.0702	0.0689	0.0699	0.0590	0.0773
100	0.0560	0.0657	0.0667	0.0502	0.0610
300	0.0515	0.0558	0.0568	0.0500	0.0550

and Bartlett corrections were applied, the Type I error was about two times better approximated than that of the two other tests. Therefore, we can conclude that the proposed method can significantly improve the Type I error control and the accuracy to cover probabilities of the interval estimation.

Table 3 is related to the cases where data follow a chi-square distribution. From Table 3, the proposed Chen–Bartlett-corrected EL ratio test is shown to have the best ability to control the Type I error especially when the sample sizes are relatively small. With the increasing of the sample size, all the tests tend to perform well.

4. A Glucose Data Example

In this section, we use the proposed method to construct the confidence interval estimators based on data from a study that evaluates biomarkers related to atherosclerotic coronary heart disease. We consider the biomarker "glucose" that can quantify different phases of oxidative stress and antioxidant status process of an individual (e.g., Schisterman et al., 2001). A population-based sample of randomly selected residents of Erie and Niagara counties, 35–79 years of age, was the focus of this investigation (Schisterman et al., 2001). The sampling frame for adults between the ages of 35 and 65 was from the New York State Department of Motor Vehicle drivers' license rolls, while the elderly sample (age 65–79) was randomly chosen from the Health Care Financing Administration database. A cohort of 917 men and women were selected for the analyses to present a population without myocardial infarction. Participants provided a 12-h fasting food specimen for biochemical analysis at baseline, and a number of parameters were examined from fresh blood samples.

We illustrate the proposed method based on the biomarker glucose that has values with a right-skewed distribution. The strategy was that a sample with size N was randomly selected from the glucose data to estimate the confidence intervals at the level 95% of the mean, say $[a, b]$; the rest of the data was used to calculate the average of the glucose levels, say c. The value of 917-N was chosen to be large, and hence we believe the computed average c is very close to the theoretical mean of the glucose levels. We repeated this strategy 10,000 times calculating the frequencies of the event $c \in [a, b]$. This bootstrap-type test shows that the Chen–Bartlett and Chen corrected EL ratio tests provides the coverage probabilities that are closer to the target 0.95 than those of the EL ratio test and the Bartlett corrected EL ratio test and the t-test. Table 4 presents these results for the different sample sizes of

Table 4
The coverage probabilities related to the glucose biomaker (The expected coverage probability is 0.95)

Sample size n	ELR	Chen	Bartlett	Chen & Bartlett	T-test
15	0.7696	0.8439	0.7854	0.8448	0.7903
25	0.8332	0.8997	0.8455	0.8977	0.8314
35	0.8684	0.9289	0.8798	0.9403	0.8551
50	0.8896	0.9387	0.8996	0.9445	0.8775
100	0.9120	0.9403	0.9188	0.9490	0.9025

Table 5
The confidence interval estimators at the level 95% of the mean of the glucose biomarker measured from 917 patients without myocardial infarction

ELR	(105.3226, 109.7409)
Chen	(105.4954, 128.1496)
Bartlett	(105.3161, 109.7502)
Chen&Bartlett	(105.4897, 128.1475)
T-test	(105.1445, 109.5367)

$N = 15, 25, 35, 50$, and 100. It can be noticed that when the sample size is above 50, the actual coverage probabilities is very close to the nominal one.

Thus, we propose to use the Chen–Bartlett and Chen corrected EL ratio statistics to estimate the confidence interval for the mean of glucose measurements. The investigated test statistics provided the confidence intervals at the level 95% of the mean of the glucose biomarker measured from 917 patients without myocardial infarction (the healthy population). The results are shown in Table 5. The ELR, Bartlett corrected ELR, and T-test show approximately similar results, the use of which can be misleading, taking into account the outputs of Table 4 and that glucose measurements have a right skewed distribution.

5. Conclusions

The EL methodology has been well addressed in the literature as a nonparametric version of its powerful parametric likelihood counterparts. An important property is that the asymptotic distribution of $2\log$ EL ratio test statistics follow a chi-square distribution. However, this approximation can be biased especially when the sample size is small. In this article, we illustrated the skewness of the data distribution can also affect the efficiency of the EL ratio tests. With respect to improving the Type I error control of the tests and coverage probability of the interval estimation the EL technique and the Bartlett correction of the EL ratio tests are well accepted in the literature. We proved that the EL ratio test can be corrected in a similar manner to a classic t-test's correction. Thus, the Chen (1995) approach, which was proposed to correct the skewness of the t-test distribution, was adapted to be applied to the EL ratio test. The Bartlett correction was shown to be based on scaling modified data, whereas the Chen method was presented via changing the location. We noted that the Bartlett technique requires certain additional conditions, e.g., no nuisance parameters related to the statement of problem, Cramér's conditions, which are unnecessary when the Chen correction is assumed to be applied.

In this article, we also proposed to combine both the Bartlett and Chen techniques as the new modification of the EL ratio test in order to improve the accuracy of the interval estimations.

Our arguments are confirmed via both the theoretical consideration and the Monte Carlo study. The Bartlett–Chen-corrected EL ratio test was shown to perform much better than the standard EL ratio test and also the classic Student's t-test when data were from skewed distribution. To demonstrate that the proposed method can be utilized in practice, we applied the proposed method to construct the confidence intervals for the glucose data that is from a coronary heart disease study.

The proposed methodology can also be applied to different EL-type procedures.

Appendix

A.1. *Sketch Proof of the Asymptotic Chi-Square Distribution of* $l(X_1, \ldots, X_n \mid \theta)$

We have

$$l(X_1, \ldots, X_n \mid \theta) = -2 \log [L(X_1, \ldots, X_n \mid \theta) n^n] = 2 \sum \log[1 + \lambda(X_i - \theta)],$$

where λ is a root of the equation

$$g(\lambda) = \sum_{i=1}^{n} \frac{X_i - \theta}{1 + \lambda(X_i - \theta)} = 0. \tag{A1}$$

Applying the Taylor expansion, we can expand $g(\lambda)$ with λ around 0. Hence, by (A1), we conclude that

$$\lambda \approx n \cdot \frac{\sum_{i=1}^{n} (X_i - \theta)}{\sum_{i=1}^{n} (X_i - \theta)^2} \quad \text{as } n \to \infty. \tag{A2}$$

Since $l(X_1, \ldots, X_n \mid \theta) = 2 \sum \log [1 + \lambda(X_i - \theta)]$, now consider $l(X_1, \ldots, X_n \mid \theta)$ with Taylor series where λ is close to 0 and after plugging (A2) to the statistic $2 \sum \log [1 + \lambda(X_i - \theta)]$ we obtain

$$l(X_1, \ldots, X_n \mid \theta) \approx \frac{[\sum_{i=1}^{n} (X_i - \theta)]^2}{\sum_{i=1}^{n} (X_i - \theta)^2}. \tag{A3}$$

Note that the test statistic $T(X_1, \ldots, X_n \mid \theta)^2$ has an asymptotic chi-square distribution as follows

$$T(X_1, \ldots, X_n \mid \theta) = \frac{\sum_{i=1}^{n} (X_i - \theta)}{\sqrt{\sum_{i=1}^{n} (X_i - \overline{X})^2}} \xrightarrow[n \to \infty]{p} N(0, 1)$$

and hence $l(X_1, \ldots, X_n \mid \theta) \approx T^2(X_1, \ldots, X_n \mid \theta) \xrightarrow[n \to \infty]{p} \chi^2(1)$.

This completes the brief proof.

Acknowledgment

The authors are grateful to the Editor and referee for their helpful comments that clearly improved this article.

References

Chen, L. (1995). Testing the mean of skewed distributions. *Journal of the American Statistical Association* 90:767–772.

Chen, S. X., Cui, H. (2006). On Bartlett correction of empirical likelihood in the presence of nuisance parameters. *Biometrika* 93:215–220.

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

Davison, A. C., Stafford, J. E. (1998). The score function and a comparison of various adjustments of the profile likelihood. *The Canadian Journal of Statistics* 26:139–148.

DiCiccio, T., Hall, P., Romano, J. (1991). Empirical likelihood is Bartlett-correctable. *The Annals of Statistics* 19:1053–1061.

Hall, P., Scala, L. B. (1990). Methodology and algorithms of empirical likelihood. *International Statistical Review/Revue Internationale de Statistique* 58:109–127.

Lazar, N. A., Mykland, A. P. (1999). Empirical likelihood in the presence of nuisance parameters. *Biometrika* 86:203–211.

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

Owen, A. B. (1990). Empirical likelihood confidence regions. *The Annals of Statistics* 18:90–120.

Owen, A. B. (1991). Empirical likelihood for linear models. *The Annals of Statistics* 19:1725–1747.

Owen, A. B. (2001). *Empirical Likelihood*. New York: Chapman and Hall.

Qin, J., Lawless, J. (1994). Empirical likelihood and general estimating equations. *The Annals of Statistics* 22:300–325.

Schisterman, E. F., Faraggi, D., Browne, R., Freudenheim, J., Dorn, J., Muti, P., Armstrong, D., Reiser, B., Trevisan, M. (2001). TBARS and cardiovascular disease in a population-based sample. *Journal of Cardiovascular Risk* 8:1–7.

Vexler, A., Wu, C., Yu, K. F. (2009). Optimal hypothesis testing: From semi to fully Bayes factors. *Metrika* In press. DOI 10.1007/s00184-008-0205-4.

Wilks, S. S. (1938). The large-sample distribution of the likelihood ratio for testing composite hypotheses. *Annals of Mathematical Statistics* 9:60–62.

Yu, J., Vexler, A., Tian, L. (2009). Analyzing incomplete data subject to a threshold using empirical likelihood methods: An application to a pneumonia risk study in an ICU setting. *Biometrics* In press. DOI 10.1111/j.1541-0420.2009.01128.x