Journal of the Korean
Statistical Society
Vol. 25, No. 2, 1996

On Bootstrapping Bartlett Adjusted Empirical Likelihood Ratio Statistic in Regression Analysis †

Woochul Kim, Duk-Hyun Ko [1] and Keewon Lee [2]

Abstract

The bootstrap calibration method for empirical likelihood is considered to make a confidence region for the regression coefficients. Asymptotic properties are studied regarding the coverage probability. Small sample simulation results reveal that the bootstrap calibration works quite well.

Key Words : Empirical likelihood; Bootstrap method; Bartlett correction; Regression model.

† This research was supported by the Ministry of Education, 1995 during the first author's visit to Hallim University, Chunchon.

[1] Department of Computer Science and Statistics, Seoul National University, Seoul, 151-742, Korea.

[2] Department of Statistics, Hallim University, Chunchon, 200-702, Korea

1. INTRODUCTION

Likelihood based inference is the most popular method in parametric model. In nonparametric case, Owen (1988, 1990) suggested the use of empirical likelihood based inference as an alternative to the bootstrap method. An important feature of empirical likelihood is that it uses only the data to determine the shape and orientation of a confidence region, it preserves the range of the parameter of interest and it is invariant to the transformation of parameter.

After Owen (1988), many authors discovered that the empirical likelihood has similar properties to those of parametric likelihood. A major advantage of empirical likelihood is that it admits a Bartlett correction. So, a simple empirical correction of the mean of empirical log likelihood ratio statistic reduces the coverage error to $O(n^{-2})$. DiCiccio, Hall and Romano (1991) showed that Bartlett correction works in the case of empirical likelihood for the smooth function of the mean.

Owen (1991) extended the idea of empirical likelihood to linear model, and showed that in this case empirical log likelihood ratio statistic has asymptotic χ^2 distribution. Chen (1993) showed that, in linear regression and correlation model, empirical likelihood ratio statistic admits Bartlett correction.

In this paper, we use the bootstrap method to estimate the quantile of the distribution of empirical log likelihood ratio statistic and its Bartlett corrected version in regression model. It is shown that the bootstrap based empirical likelihood ratio confidence region has coverage error of order $O(n^{-2})$, and that bootstrapping the Bartlett corrected version reduces the coverage error to $O(n^{-3})$.

Section 2 reviews the empirical likelihood method in regression model and the Bartlett adjustment. Section 3 deals with the bootstrap calibration of the empirical likelihood, and provides the theoretical results. In Section 4, results of a simulation study are provided to shed the insight into the small sample performance of the bootstrap calibration.

2. EMPIRICAL LIKELIHOOD AND BARTLETT ADJUSTMENT

2.1 Empirical Likelihood

Consider the linear regression model,

$$y_i = x_i\beta + \epsilon_i, \quad 1 \leq i \leq n.$$

where β is a $p \times 1$ vector of unknown parameter, x_i is a $1 \times p$ design point and y_i is a response. The error ϵ_i's are independent random variables with mean 0 and variance $\sigma^2(x_i)$.

Adopting Owen (1988)'s notations, we define a working variable u_i and a variance matrix V_n as follows:

$$\begin{aligned} u_i &\equiv x_i^t(y_i - x_i\beta) \\ V_n &\equiv n^{-1}\sum_{i=1}^{n} Cov(u_i) = n^{-1}\sum_{i=1}^{n} x_i^t x_i \sigma^2(x_i) \end{aligned}$$

Then, the empirical log likelihood ratio statistic evaluated at the true parameter β is defined by

$$\ell(\beta) = 2\sum_{i=1}^{n} \log(1 + \lambda^t u_i) \tag{2.1}$$

where λ is a $p \times 1$ vector satisfying

$$n^{-1}\sum_{i=1}^{n}(1 + \lambda^t u_i)^{-1} u_i = 0.$$

In terms of standardized variables $w_i \equiv V_n^{-1/2} u_i$, the equation (2.1) becomes

$$\ell(\beta) = 2\sum_{i=1}^{n} \log(1 + t^t w_i) \tag{2.2}$$

where t satisfies

$$n^{-1}\sum_{i=1}^{n}(1 + t^t w_i)^{-1} w_i = 0. \tag{2.3}$$

The following theorem in Owen (1991) shows that the empirical log likelihood ratio statistic (2.1) has asymptotic chi-squared distribution and the empirical likelihood based confidence region of β has coverage error of order $O(n^{-1})$.

Theorem 1. (Owen (1991)) Under certain regularity conditions, the empirical log likelihood ratio statistic (2.1) has asymptotic chi-squared distribution with the error of order $O(n^{-1})$, i.e., as $n \to \infty$,

$$P(\ell(\beta) < c) \;=\; P(\chi_p^2 < c) + O(n^{-1})$$

where χ_p^2 denotes the chi-squared random variable with p degrees of freedom.

2.2 Bartlett Adjustment

Theorem 2.1 shows that the coverage error of empirical log likelihood regions for β in regression model is the same as that of confidence regions constructed by bootstrap method. But, it is well known that the coverage error of empirical log likelihood is partly due to the fact that the mean of $\ell(\beta)$ does not agree with the mean of χ_p^2.

In this regard, it can be shown that

$$E\ell(\beta) = p(1 + n^{-1}b) + O(n^{-2})$$

where the exact form of b can be found in Chen (1993).

The Bartlett corrected region of $\ell(\beta)$ is defined by

$$\ell_{BC}(\beta) = \ell(\beta) \cdot (1 + n^{-1}\hat{b})^{-1}, \tag{2.4}$$

where $\hat{b}$ denotes b or $\sqrt{n}$-consistent estimate of b. The following theorem shows that it has the coverage error of order $O(n^{-2})$.

Theorem 2. (Chen 1993) Under certain regularity condition, the Bartlett corrected empirical log likelihood ratio statistic has asymptotic chi-squared distribution with the error of order $O(n^{-2})$, i.e., as $n \to \infty$,

$$P\{\ell_{BC}(\beta) < c\} = P\{\chi_p^2 < c\} + O(n^{-2}).$$

3. BOOTSTRAP CALIBRATION OF EMPIRICAL LIKELIHOOD

Bootstrap is a general method to estimate the distribution of a statistic using Monte Carlo method. So, we can use the bootstrap method to estimate the distribution of empirical likelihood ratio statistic rather than to use the chi-squared approximation. In this section, we show that the empirical likelihood confidence region based on bootstrap quantile reduces the size of coverage error by the factor of $O(n^{-1})$.

Let $\mathcal{X}$ denote the observed sample, i.e., $\mathcal{X} \equiv \{(x_1, y_1), \cdots, (x_n, y_n)\}$ and $\mathcal{X}^* \equiv \{(x_1, y_1^*), \cdots, (x_n, y_n^*)\}$ denote the bootstrap sample with

$$y_i^* = x_i\hat{\beta} + \hat{\epsilon}_i^*, \quad i = 1, \cdots, n$$

where $\hat{\beta}$ denotes the ordinary least square estimate of β and $\hat{\epsilon}_i^*$'s are random sample from $\hat{\epsilon}_i = y_i - x_i\hat{\beta}$. Also, let $\ell^*(\hat{\beta})$ denote the bootstrap version of $\ell(\beta)$.

To approximate the distribution of $\ell(\beta)$ and $\ell^*(\hat{\beta})$, consider the formal expansion of $\ell(\beta)$. To this end, we need the following notations,

$$\begin{aligned} \bar{\alpha}^{j_1\cdots j_k} &= n^{-1}\sum_{i=1}^{n} E(w_i^{j_1}\cdots w_i^{j_k}) \\ Z^{j_1\cdots j_k} &= \sqrt{n}(n^{-1}\sum_{i=1}^{n} w_i^{j_1}\cdots w_i^{j_k} - \bar{\alpha}^{j_1\cdots j_k}) \end{aligned}$$

Note that, $\bar{\alpha}^j = 0$ and $\bar{\alpha}^{jk} = \delta^{jk}$ where δ^{jk} denotes Cronecker's delta. Then, in the equation (2.3), we may assume the following formal expansion of t;

$$t \equiv \sum_{r=1}^{\infty} n^{-r/2} t_r, \text{ with } t_r \equiv O_P(1)$$

which implies an expression

$$\begin{aligned} t^{j_1}\cdots t^{j_k} &= \sum_{r_1=1}^{\infty}\cdots\sum_{r_k=1}^{\infty} n^{-(r_1+\cdots+r_k)/2} t_{r_1}^{j_1}\cdots t_{r_k}^{j_k} \\ &= \sum_{r=k}^{\infty} n^{-r/2} \sum_{r_1+\cdots+r_k=r} t_{r_1}^{j_1}\cdots t_{r_k}^{j_k}. \end{aligned}$$

Thus, putting this into the equation (2.3), we have

$$\begin{aligned} t_1^p &= Z^p \\ t_r^p &= \sum_{k=1}^{r-1}(-1)^k\Big\{\sum_{s_1+\cdots+s_{k+1}=r} t_{s_1}^{j_1}\cdots t_{s_{k+1}}^{j_{k+1}}\bar{\alpha}^{j_1\cdots j_{k+1}p} \\ &\qquad - \sum_{r_1+\cdots+r_k=r-1} t_{r_1}^{j_1}\cdots t_{r_k}^{j_k} Z^{j_1\cdots j_k p},\Big\} \quad (r \geq 2) \end{aligned}$$

From Taylor series expansion of the equation (2.2) and the expansion of t, we obtain the following expansion of $\ell(\beta)$;

$$\ell(\beta) = \sum_{k=2}^{5} 2(-1)^k(1-k^{-1})n^{-\frac{1}{2}(k-2)}\xi^{j_1}\cdots\xi^{j_k}(\bar{\alpha}^{j_1\cdots j_k} + n^{-1/2}Z^{j_1\cdots j_k}) + O_P(n^{-2}) \tag{3.1}$$

where $\xi^j \equiv \sqrt{n}t^j$. Note that $n^{-j/2}$ term in $\ell(\beta)$ is the sum of the products of odd/even number of random variables for odd/even j.

It follows from (3.1) that $\ell(\beta)$ admits an expression

$$\ell(\beta) = R^t R + O_P(n^{-2})$$

where $R = R_0 + \sum_{j=1}^{3} n^{-j/2} R_j$, and R_j's are random vectors whose components are sums of product of odd/even number of average type random variables for even/odd j. The detailed formula of R can be found in Chen (1993).

For this R, the following Edgeworth expansion holds (see Chen (1993), for example).

$$\sup_{B \in \mathcal{B}} \mid P(R \in B) - \int_B \sum_{j=1}^{4} n^{-(j-1)/2} \pi_j(v)\phi(v)dv \mid = O(n^{-2}), \tag{3.2}$$

where π_j's are certain even/odd polynomials for odd/even j . Here $\mathcal{B}$ is any class of Borel sets satisfying

$$\sup_{B \in \mathcal{B}} \int_{(\partial B)^\epsilon} \phi(v)dv = O(\epsilon), \quad \epsilon \to 0$$

with (∂B) and $(\partial B)^\epsilon$ denote the boundary of B and its ϵ neighborhood, respectively.

In fact, the following conditions guarantee the validity of Edgeworth expansion.

(C1) There exist positive constants C_1 and C_2 such that $C_1 < v_{nn} \leq v_{1n} < C_2$ uniformly in n, where v_{1n} and v_{nn} denote the maximum and minimum eigenvalues of the covariance matrix V_n, respectively.

(C2) For sufficiently large positive integer l,

$$\sup_n n^{-1} \sum_{i=1}^{n} E||u_i||^l < \infty$$

(C3) For every positive $\epsilon > 0$ and l in (C2),

$$\lim_{n \to \infty} n^{-1} \sum_{j=1}^{n} \int_{||u_j|| > \epsilon n^{1/2}} ||u_j||^l = 0.$$

(C4) For every positive b, the joint characteristic function ϕ_n of $u_1, \cdots, u_n$ satisfies Cramer's condition

$$\limsup_{n\to\infty} \sup_{||t||>b} |\phi_n(t)| < 1.$$

Theorem 3. Under the conditions (C1) through (C4), the following approximations hold;

$$\begin{aligned} P\{\ell(\beta) < c\} &= P\{\chi_p^2 < c\} + n^{-1}q_1(c)g(c) + O(n^{-2}) \\ P\{\ell^*(\hat{\beta}) < c|\mathcal{X}\} &= P\{\chi_p^2 < c\} + n^{-1}\hat{q}_1(c)g(c) + O_P(n^{-2}) \end{aligned}$$

where g denotes the density function of χ_p^2 and q_1 is a linear odd polynomial and $\hat{q}_1$ is its sample version.

Proof. The proof of the first part can be found in Chen (1993). To this end, it should be emphasized that the parity property of π_j's in (3.2) can be derived not from the smooth function model, but from the form of the statistic R. Note that R_j is the sum of the products of odd/even numbers of random variables for even/odd j. From this fact, and Theorem 2.1 in Hall (1992), the r^{th} cumulant of R can be expanded as a power series of n^{-1} with the leading order being $n^{-(r-2)/2}$.

To prove the second part, we need stringent moment conditions given in (C2) and (C3). Let M_r be the sample moment of order r, i.e.,

$$M_r = n^{-1}\sum_{i=1}^{n}\prod_j (u_i^j)^{l_j}$$

where l_j's are nonnegative integers satisfying $\sum l_j = r, (r = 1, 2, 3, 4)$. Then we need to choose sufficiently large l so that

$$P(|M_r - E(M_r)| > \epsilon) = O(n^{-\lambda})$$

for some $\lambda > 1$.

With such l, applying Lemma 5.5 of Hall (1992), we can obtain the following approximation for R^*, the bootstrap analogue of R;

$$\sup_{B\in\mathcal{B}} |\ P(R^* \in B|\mathcal{X}) - \int_B \sum_{j=1}^{4} n^{-(j-1)/2}\hat{\pi}_j(v)\phi(v)dv\ |= O_P(n^{-2}), \tag{3.3}$$

where $\hat{\pi}_j$'s denotes the sample analogue of π_j. Now, the expansion of the distribution of $\ell^*(\hat{\beta})$ has no $n^{-j/2}$ term for odd j due to the parity of $\hat{\pi}_j$'s and the Edgeworth expansion (3.3). Q.E.D.

Let y_α denote the α quantile of the empirical log likelihood ratio statistic and $\hat{y}_\alpha$ denote its bootstrap estimator, i.e.,

$$P\{\ell(\beta) < y_\alpha\} = \alpha, \quad P\{\ell^*(\hat{\beta}) < \hat{y}_\alpha | \mathcal{X}\} = \alpha.$$

Corollary 1. Under the conditions (C1) through (C4), the following approximation hold;

$$P\{\ell(\beta) < \hat{y}_\alpha\} = \alpha + O(n^{-2}).$$

Proof. From Theorem 3, we can obtain the following expansions of y_α and $\hat{y}_\alpha$;

$$\begin{aligned} y_\alpha &= \chi_p^2(\alpha) - n^{-1} q_1(\chi_p^2(\alpha))/g(\chi_p^2(\alpha)) + O(n^{-2}) \\ \hat{y}_\alpha &= \chi_p^2(\alpha) - n^{-1} \hat{q}_1(\chi_p^2(\alpha))/g(\chi_p^2(\alpha)) + O_P(n^{-2}) \end{aligned}$$

where $\chi_p^2(\alpha)$ denotes the α quantile of χ_p^2. Thus we have

$$\hat{y}_\alpha - y_\alpha = O_P(n^{-3/2}).$$

Therefore,

$$P\{\ell(\beta) < \hat{y}_\alpha\} = P\{\ell(\beta) \cdot (1 + n^{-3/2}\zeta_1)^{-1} < y_\alpha\} + O(n^{-2})$$

where ζ is a random variable of order $O_P(1)$. Now it should be noted that $\ell(\beta) \cdot (1 + n^{-3/2}\zeta_1)^{-1}$ has the same property as $\ell(\beta)$,i.e., its $n^{-j/2}$ term is the sum of the products of odd/even number of random variables for odd/even j. Hence the argument in the proof of Theorem 3 can be employed to show that the expansion of the distribution of $\ell(\beta) \cdot (1 + n^{-3/2}\zeta_1))$ as a power series of $n^{-1/2}$ has no $n^{-3/2}$ term, i.e.,

$$P\{\ell(\beta) < \hat{y}_\alpha\} = \alpha + O(n^{-2}). \qquad \text{Q.E.D.}$$

In fact, a result similar to Corollary 1 can be obtained for the Bartlett corrected version $\ell_{BC}(\beta)$. To be precise, under certain conditions, we have

$$P\{\ell_{BC}(\beta) < \hat{y}_\alpha^{BC}\} = \alpha + O(n^{-3}). \tag{3.4}$$

where y_α^{BC} denotes a α quantile of $\ell_{BC}(\beta)$ and $\hat{y}_\alpha^{BC}$ is its bootstrap estimator.

Assuming that the Bartlett correction factor $\hat{b}$ in (2.4) has the following expansion;

$$\hat{b} = b + n^{-\frac{1}{2}} c_{1i}\eta^i + n^{-1} c_{2ij}\eta^i\eta^j + n^{-\frac{3}{2}} c_{3ijk}\eta^i\eta^j\eta^k + O_P(n^{-2}) \tag{3.5}$$

Then, from (3.1) and (3.5),

$$\begin{aligned} \ell_{BC}(\beta) \;\equiv\; & \sum_{j=1}^{p}(\gamma^j)^2 + \sum_{k=3}^{7} n^{-\frac{1}{2}(k-2)} d_{1,j_1\cdots j_k}\gamma^{j_1}\cdots\gamma^{j_k} + \sum_{k=2}^{5} n^{-\frac{1}{2}k} d_{2,j_1\cdots j_k}\gamma^{j_1}\cdots\gamma^{j_k} \\ + & \sum_{k=2}^{3} n^{-\frac{1}{2}(k+2)} d_{3,j_1\cdots j_k}\gamma^{j_1}\cdots\gamma^{j_k} + O_P(n^{-3}) \end{aligned}$$

where γ is a vector consisting of ξ's, η's and Z's, and γ^j's are ordered so that $\gamma^j = \xi^j (j = 1, \cdots, p)$.

Applying the argument in Theorem 3.1 and adopting the result of DiCiccio, Hall and Romano (1991), we can obtain the following approximation;

$$P\{\ell_{BC}(\beta) \le c\} = P\{\chi_p^2 \le c\} + n^{-2}Q(c) + O(n^{-3})$$

From this, we can obtain

$$\hat{y}_\alpha^{BC} = y_\alpha^{BC}\{1 + n^{-5/2}\zeta_2\} + O_P(n^{-3})$$

where ζ_2 is a random variable of order $O_P(1)$. Hence, we have

$$P\{\ell_{BC}(\beta) < \hat{y}_\alpha^{BC}\} = P\{\ell_{BC}(\beta)\cdot(1 + n^{-5/2}\zeta_2)^{-1} < y_\alpha^{BC}\} + O(n^{-3}).$$

By applying the argument in the proof of Corollary 1, we can show that the expansion of the distribution of $\ell_{BC}(\beta)\cdot(1 + n^{-5/2}\zeta_2)^{-1}$ as a power series of $n^{-1/2}$ has no $n^{-5/2}$ term. Hence, the result (3.4) follows.

4. RESULTS OF A SIMULATION STUDY

To get the insight into the small sample performance of various confidence regions, we have run a simulation study. In this study, we consider the simple linear regression model

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad i = 1, \cdots, n$$

with the following three types of error distribution;

(M1) Homogeneous normal error: $\epsilon_i \sim \mathcal{N}(0,1)$,

(M2) Variance contaminated error: $\epsilon_i \sim 0.9\mathcal{N}(0,1) + 0.1\mathcal{N}(0,4)$,

(M3) Heteroscedastic normal error: $\epsilon_i \sim (0.5x_i)^{1/2}\mathcal{N}(0,1)$,

where (M1) and (M3) have been considered in Chen (1993).

Table 1. Estimated Coverage

model	level	method	Sample Size		
			$n=30$	$n=40$	$n=50$
M1	0.90	EL	0.8368	0.8606	0.8682
		EL_{BC}	0.8612	0.8794	0.8834
		EL^*	0.8932	0.8922	0.8946
		EL^*_{BC}	0.8978	0.8964	0.8982
	0.95	EL	0.9010	0.9156	0.9274
		EL_{BC}	0.9198	0.9302	0.9364
		EL^*	0.9472	0.9442	0.9476
		EL^*_{BC}	0.9510	0.9454	0.9494
M2	0.90	EL	0.8414	0.8598	0.8675
		EL_{BC}	0.8662	0.8850	0.8864
		EL^*	0.8938	0.8978	0.8963
		EL^*_{BC}	0.8978	0.8996	0.8997
	0.95	EL	0.8992	0.9170	0.9249
		EL_{BC}	0.9168	0.9302	0.9367
		EL^*	0.9458	0.9458	0.9456
		EL^*_{BC}	0.9476	0.9466	0.9485
M3	0.90	EL	0.8304	0.8660	0.8704
		EL_{BC}	0.8530	0.8844	0.8826
		EL^*	0.8900	0.9058	0.8952
		EL^*_{BC}	0.8928	0.9068	0.8966
	0.95	EL	0.8964	0.9220	0.9260
		EL_{BC}	0.9122	0.9332	0.9364
		EL^*	0.9434	0.9480	0.9480
		EL^*_{BC}	0.9462	0.9490	0.9490

Throughout the simulation, the design point x_i's are uniform grid on$[1, 10]$, and the sample size $n = 30, 40$ and 50 are considered. For nominal levels $\alpha = 0.90$ and 0.95, the empirical coverages based on 5,000 simulations are computed for various confidence regions for $\beta = (\beta_0, \beta_1)$. The results are given in Table 1.

The confidence regions under this study are the uncorrected empirical likelihood confidence region (EL), Bartlett corrected region (EL_{BC}), bootstrap calibrated empirical likelihood region (EL^*) and bootstrap calibrated Bartlett corrected region (EL^*_{BC}). The bootstrap re-sampling size is 200, and the appropriate Bartlett correction factors in Chen (1993) (iid case factor for M1 and M2, and non-iid case factor for M3) are used.

Overall, the simulation results reveal that the bootstrap calibration is more than satisfactory. It can be observed from Table 1 that, even for homogeneous normal error, the uncorrected empirical likelihood confidence region have coverage probabilities significantly smaller than the nominal level. The Bartlett correction improves coverage at least 1.4% throughout the cases studied. But its empirical coverage probabilities are still below the nominal level as can be observed from Table 1.

On the other hand, as can be observed from Table 1, the bootstrap calibration improves the coverage probabilities quite well. In fact, the bootstrap calibrated empirical likelihood regions (EL^*) have the coverage probabilities all within 1% of the nominal level.

The bootstrap calibration of the Bartlett corrected region (EL^*_{BC}) performs the best. Its improvement over the EL^*, however, is not very significant compared to the improvement of EL^* over EL_{BC}. This simulation result tells us that the bootstrap calibration as well as Bartlett correction should be carried out in practice whenever one employs the empirical likelihood confidence region.

Finally, it should be remarked that throughout the simulation study the ordinary least-square estimate of $\beta = (\beta_0, \beta_1)$ has been used to construct the confidence regions even in the case of heteroscedastic errors. Some limited simulations with the weighted least square estimate were run to show results similar to those in Table 1.

REFERENCE

(1) Chen, S. X. (1993). On the accuracy of empirical likelihood confidence region for linear model. *Annals of the Institute of Mathematical Statistics.*

(2) DiCiccio, T.J., Hall, P. and Romano, J.P. (1991). Empirical likelihood is Bartlett- correctable. *Annal of Statistics*, 1053-1061.

(3) Hall, P. (1992). *The Bootstrap and Edgeworth Expansion.* Springer-Verlag, NewYork.

(4) Owen, A. (1988). Empirical likelihood ratio confidence intervals for a single functional. *Biometrika*, 237-249.

(5) Owen, A. (1990). Empirical likelihood ratio confidence regions. *Annal of Statistics*, 90-120.

(6) Owen, A. (1991). Empirical likelihood for linear models. *Annal of Statistics*, 1725-1747.