

## Empirical Likelihood Confidence Intervals for Linear Regression Coefficients

SONG XI CHEN

*Australian National University, Canberra, Australia*

Nonparametric versions of Wilks' theorem are proved for empirical likelihood estimators of slope and mean parameters for a simple linear regression model. They enable us to construct empirical likelihood confidence intervals for these parameters. The coverage errors of these confidence intervals are of order  $n^{-1}$  and can be reduced to order  $n^{-2}$  by Bartlett correction. © 1994 Academic Press, Inc.

### 1. INTRODUCTION

Empirical likelihood is a nonparametric technique for constructing confidence regions. It has sampling properties similar to those of bootstrap. However, instead of putting equal probability weight  $n^{-1}$  on each data value, empirical likelihood chooses the weights by profiling a multinomial likelihood supported on the sample. The use of empirical likelihood methods to construct confidence regions for  $\beta$ , which is the vector of unknown regression coefficients in a linear regression model, has been studied by Owen [1] and Chen [2]. Owen [1] pioneered this work by proving a nonparametric version of Wilks' theorem for the empirical likelihood ratio of  $\beta$ , which enables us to construct confidence regions for  $\beta$  using  $\chi^2$  tables. The second order properties of empirical likelihood confidence regions were discussed by Chen [2], showing that coverage errors are of order  $n^{-1}$  and Bartlett correction can be employed to reduce the coverage error to order  $n^{-2}$ .

However, it is not enough to just construct confidence regions for  $\beta$ . In practice, statisticians are often confronted with problems of constructing confidence intervals for a particular regression coefficient or certain linear combinations of  $\beta$ .

Received May 11, 1992; revised May 6, 1993.

AMS 1980 subject classification: primary 62E20.

Key words and phrases: Bartlett correction, confidence interval, coverage, Edgeworth expansion, empirical likelihood, simple linear regression model.

In this paper we address the above problem under the simple linear regression model. A simple linear regression model is

$$y_i = a_0 + b_0 x_i + \varepsilon_i, \quad 1 \leq i \leq n, \quad (1.1)$$

where all the variables appearing in (1.1) are scalars. Among them,  $x_i$  and  $y_i$  are the  $i$ th fixed design point and response, respectively, the  $\varepsilon_i$ 's are independent and identically distributed random errors with mean zero and variance  $\sigma^2$ , and  $a_0$  and  $b_0$  are the unknown intercept and slope parameters, respectively.

There are two aims in this paper. First, we show how to construct empirical likelihood confidence intervals for the slope parameter  $b_0$  and means  $y_0 = a_0 + b_0 x_0$  for any fixed  $x_0$ , under model (1.1). Obviously the latter case includes the intercept parameter  $a_0$  when one chooses  $x_0 = 0$ . Second, we study the coverage accuracy and Bartlett correctability of empirical likelihood confidence intervals for these parameters.

Analyses in Sections 3 and 4 show that both empirical likelihood confidence intervals for  $b_0$  and  $y_0$  have coverage errors of order  $n^{-1}$ , and that both confidence intervals are Bartlett correctable. Thus, simple scale adjustments can improve the coverage accuracy of those confidence intervals from order  $n^{-1}$  to order  $n^{-2}$ . A simulation study is presented in Section 5.

## 2. PRELIMINARIES

In this section we introduce some notation and basic formulae which are used throughout this paper. We use  $\hat{a}_0$  and  $\hat{b}_0$  for the least squares estimates of  $a_0$  and  $b_0$ , respectively,  $\mu_j$  for the  $j$ th moment of  $\varepsilon_1$  for  $j = 1, 2$ , and  $\bar{x}$  and  $\bar{y}$  for the means of  $x_i$ 's and  $y_i$ 's, respectively. We define auxiliary variables  $z_i(a, b) = (1, x_i)^T (y_i - a - bx_i)$  for  $i = 1, \dots, n$ , where  $a$  and  $b$  are any candidate values for  $a_0$  and  $b_0$ . Specifically we write  $z_i$  as  $z_i(a_0, b_0)$ . Furthermore, put

$$\begin{aligned} \sigma_x^2 &= n^{-1} \sum (x_i - \bar{x})^2, & m_j &= n^{-1} \sum (x_i - \bar{x})^j, & j &= 3, 4 \\ \hat{\sigma}^2 &= n^{-1} \sum \hat{\varepsilon}_i^2, & \hat{\mu}_j &= n^{-1} \sum \hat{\varepsilon}_i^j, & j &= 3, 4, \\ \tilde{\varepsilon} &= \bar{y} - a_0 - b_0 \bar{x}, & \text{where } \hat{\varepsilon}_i &= y_i - \hat{a} - \hat{b}x_i. \end{aligned}$$

Let

$$V_n = \sigma^2 \begin{pmatrix} 1 & \bar{x} \\ \bar{x} & n^{-1} \sum x_i^2 \end{pmatrix}$$

be the average covariance matrix of auxiliary variables  $z_i$ 's, let  $v_{1n}$  and  $v_{2n}$  be the largest and smallest eigenvalues of  $V_n$ , respectively, and let

$$U_n = \begin{pmatrix} u_1^1 & u_1^2 \\ u_2^1 & u_2^2 \end{pmatrix} = V_n^{-1/2}$$

be the inverse of the square root matrix of  $V_n$ . Moreover we define

$$\begin{aligned} g_{j_1 j_2 \dots j_k}(x_i) &= \prod_{l=1}^k (u_{j_l}^1 + u_{j_l}^2 x_i), \\ \bar{\alpha}^{j_1 j_2 \dots j_k} &= n^{-1} \sum E\{g_{j_1 j_2 \dots j_k}(x_i) e_i^k\}, \\ A^{j_1, j_2, \dots, j_k}(a, b) &= n^{-1} \sum g_{j_1 j_2 \dots j_k}(x_i) (y_i - a - b x_i)^k - \bar{\alpha}^{j_1, j_2, \dots, j_k}. \end{aligned}$$

For simplicity of notation we write

$$A_0^{j_1, j_2, \dots, j_k} = A^{j_1, j_2, \dots, j_k}(a_0, b_0) \quad \text{and} \quad A^{j_1, j_2, \dots, j_k} = A^{j_1, j_2, \dots, j_k}(a, b).$$

We assume the following regularity conditions.

There exist positive constants  $C_1$  and  $C_2$  such that uniformly in  $n$ ,

$$C_1 < v_{pn} \leq v_{1n} < C_2; \quad \text{and} \quad n^{-2} \sum_{j=1}^n E \|z_j\|^4 \rightarrow 0, \quad (2.1)$$

where  $\| \cdot \|$  is the Euclidean norm. For any candidate values  $a$  and  $b$  or  $a_0$  and  $b_0$ ,

$$a = a_0 + O_p(n^{-1/2}) \quad \text{and} \quad b = b_0 + O_p(n^{-1/2}). \quad (2.2)$$

Let  $l(a, b)$  be the empirical log likelihood ratio evaluated at  $(a, b)$ . Write  $p_1, \dots, p_n$  for nonnegative numbers adding to unity. Then, according to the definition of empirical likelihood,

$$l(a, b) = -2 \min_{\sum p_i z_i(a, b) = 0} \sum_{i=1}^n \log(np_i).$$

Using the Lagrange method gives us

$$l(a, b) = 2 \sum \log \{1 + \lambda(1, x_i)^T (y_i - a - b x_i)\},$$

and  $\lambda = (\lambda_1, \lambda_2)$  satisfies

$$\sum \frac{(1, x_i)^T (y_i - a - b x_i)}{1 + \lambda(1, x_i)^T (y_i - a - b x_i)} = 0.$$

Since the analytic solutions for both  $\lambda$  and  $l(a, b)$  are difficult to obtain, we have to resort to expansions. Using (2.4) of Chen [2], under conditions (2.1) and (2.2), we have for  $l(a, b)$  the Taylor expansion

$$\begin{aligned} n^{-1}l(a, b) = & A^j A^j - A^{jk} A^j A^k + \frac{2}{3} \bar{\alpha}^{jkl} A^j A^k A^l + A^{jl} A^{kl} A^j A^k \\ & + \frac{2}{3} A^{jkl} A^j A^k A^l - 2\bar{\alpha}^{jkm} A^{lm} A^j A^k A^l \\ & + (\bar{\alpha}^{jkn} \bar{\alpha}^{lmn} - \frac{1}{2} \bar{\alpha}^{jklm}) A^j A^k A^l A^m + O_p(n^{-5/2}). \end{aligned} \quad (2.3)$$

Here we use the summation convention according to which, if an index occurs more than once in an expression, summation over the index is understood.

### 3. EMPIRICAL LIKELIHOOD CONFIDENCE INTERVAL FOR $b_0$

In this section we show how to construct empirical likelihood confidence intervals for the slope parameter  $b_0$  and analyse the coverage properties of these confidence intervals. We first prove a nonparametric version of Wilks' theorem for the empirical log likelihood ratio for  $b_0$  (Theorem 3.1). Then we develop an Edgeworth expansion of the distribution of the empirical log likelihood ratio for  $b_0$  (Theorem 3.2), which is used to show that the coverage errors of the confidence intervals are of order  $n^{-1}$ . Furthermore we demonstrate that the empirical likelihood confidence intervals are Bartlett correctable (Theorem 3.3). This means that simple scale adjustments can reduce the coverage errors from  $O(n^{-1})$  to  $O(n^{-2})$ .

The empirical log likelihood ratio for  $b_0$  may be obtained by minimizing  $l(a, b_0)$  with respect to  $a$ , which is treated as a nuisance parameter in this section, since we are only interested in constructing confidence intervals for  $b_0$ . Let  $\tilde{a}$  be the optimal  $a$  which minimizes  $l(a, b_0)$ . Then

$$l(b_0) = l(\tilde{a}, b_0) = \min_a l(a, b_0).$$

From (2.3), we know that

$$\begin{aligned} n^{-1}l(a, b_0) = & A^j(a, b_0) A^j(a, b_0) - A^{jk}(a, b_0) A^j(a, b_0) A^k(a, b_0) \\ & + \left\{ \frac{2}{3} \bar{\alpha}^{jkl} A^l(a, b_0) + A^{jl}(a, b_0) A^{kl}(a, b_0) \right\} A^j(a, b_0) A^k(a, b_0) \\ & + \left\{ \frac{2}{3} A^{jkl}(a, b_0) - 2\bar{\alpha}^{jkm} A^{lm}(a, b_0) \right\} \\ & \times A^j(a, b_0) A^k(a, b_0) A^l(a, b_0) \\ & + (\bar{\alpha}^{jkn} \bar{\alpha}^{lmn} - \frac{1}{2} \bar{\alpha}^{jklm}) A^j(a, b_0) A^k(a, b_0) A^l(a, b_0) A^m(a, b_0) \\ & + O_p(n^{-5/2}). \end{aligned} \quad (3.1)$$

Consider an expansion of  $\tilde{a} = \hat{a} + a_1 + a_2 + a_3$ , where  $a_j = O_p(n^{-j/2})$ ,  $j = 1, 2, 3$ . We determine  $a_1, a_2, a_3$  successively. Put

$$\begin{aligned}\gamma_j &= n^{-1} \sum g_j(x_i), \quad \gamma_{jk} = n^{-1} \sum g_{jk}(x_i), \\ \gamma_{jk,1}(a, b) &= n^{-1} \sum g_{jk}(x_i)(y_i - a - bx_i), \\ \gamma_{jkl,2}(a, b) &= n^{-1} \sum g_{jkl}(x_i)(y_i - a - bx_i)^2.\end{aligned}$$

Some algebra shows that

$$\begin{aligned}a_1 &= A^j(\hat{a}) \gamma_j / \gamma_j \gamma_j = \bar{x}(\hat{b} - b_0), \\ a_2 &= -(\gamma_i \gamma_i)^{-1} \gamma_j \{A^k(\hat{a}) - \gamma_k a_1\} [A^{jk}(\hat{a}) - \bar{\alpha}^{jkl} \{A^l(\hat{a}) - \gamma_l a_1\}]\end{aligned}$$

and  $a_3 = O_p(n^{-2})$ . In summary we have

$$\tilde{a} = \hat{a} + \bar{x}(\hat{b} - b_0) - (\gamma_i \gamma_i)^{-1} \gamma_j \{A^k(\hat{a}) - \gamma_k a_1\} [A^{jk}(\hat{a}) - \bar{\alpha}^{jkl} \{A^l(\hat{a}) - \gamma_l a_1\}].$$

The above formula suggests using  $\hat{a} + \bar{x}(\hat{b} - b_0)$  as an initial value for  $a$  in numerically searching for  $\tilde{a}$ . In the author's experience, this works well. Now, with  $\tilde{a}$  substituted into (3.1), the empirical likelihood ratio statistic at  $b_0$  is given by

$$\begin{aligned}n^{-1}l(b_0) &= \{A^j(\hat{a}) - \gamma_j a_1\} \{A^l(\hat{a}) - \gamma_l a_1\} \\ &\quad - \{A^{jk}(\hat{a}) - 2\gamma_{jk,1}(\hat{a}) a_1 + \gamma_{jk} a_1^2\} \{A^l(\hat{a}) - \gamma_l a_1\} \\ &\quad \times \{A^k(\hat{a}) - \gamma_k a_1\} + \frac{2}{3} \bar{\alpha}^{jkl} \{A^l(\hat{a}) - \gamma_l a_1\} \\ &\quad \times \{A^k(\hat{a}) - \gamma_k a_1\} \{A^l(\hat{a}) - \gamma_l a_1\} + A^{jl}(\hat{a}) A^{kl}(\hat{a}) \{A^l(\hat{a}) - \gamma_l a_1\} \\ &\quad \times \{A^k(\hat{a}) - \gamma_k a_1\} - \gamma_j \gamma_l a_2^2 + \frac{2}{3} \{A^{jkl}(\hat{a}) - \gamma_{jkl,2}(\hat{a}) a_1\} \\ &\quad \times \{A^l(\hat{a}) - \gamma_l a_1\} \{A^k(\hat{a}) - \gamma_k a_1\} \{A^l(\hat{a}) - \gamma_l a_1\} - 2\bar{\alpha}^{jkm} A^{lm}(\hat{a}) \\ &\quad \times \{A^l(\hat{a}) - \gamma_l a_1\} \{A^k(\hat{a}) - \gamma_k a_1\} \{A^l(\hat{a}) - \gamma_l a_1\} + \bar{\alpha}^{jkn} \bar{\alpha}^{lmn} \\ &\quad \times \{A^l(\hat{a}) - \gamma_l a_1\} \{A^k(\hat{a}) - \gamma_k a_1\} \{A^l(\hat{a}) - \gamma_l a_1\} \{A^m(\hat{a}) - \gamma_m a_1\} \\ &\quad - \frac{1}{2} \bar{\alpha}^{jklm} \{A^l(\hat{a}) - \gamma_l a_1\} \{A^k(\hat{a}) - \gamma_k a_1\} \{A^l(\hat{a}) - \gamma_l a_1\} \\ &\quad \times \{A^m(\hat{a}) - \gamma_m a_1\} + O_p(n^{-5/2}).\end{aligned}$$

For the purpose easy analysis, we next express  $l(b_0)$  in terms of powers of  $(\hat{b} - b_0)$ . Let us define  $\eta_j = \sigma_x^2 u_j^2$ , where  $u_j^2$  is the  $(j, 2)$  element in

the matrix  $U_n$ . Using the facts that  $\eta_j \eta_j = \sigma_x^2 / \sigma^2$  and  $\{A^j(\hat{a}) - \gamma_j a_1\} \{A^j(\hat{a}) - \gamma_j a_1\} = (\hat{b} - b_0)^2 \sigma_x^2 / \sigma^2$ , it may be shown that

$$\begin{aligned} n^{-1}l(b_0) &= \frac{\sigma_x^2}{\sigma^2} (\hat{b} - b_0)^2 - \eta_j \eta_k A_0^{jk} (\hat{b} - b_0)^2 + \frac{2}{3} \bar{\alpha}^{jkl} \eta_j \eta_k \eta_l (\hat{b} - b_0)^3 \\ &\quad + \eta_j \eta_k (\gamma_{jk,1} \bar{e} - \gamma_{jk} \bar{e}^2 + A_0^{jl} A_0^{kl}) (\hat{b} - b_0)^2 \\ &\quad - (\gamma_l \gamma_l)^{-1} \gamma_j \gamma_m \eta_k \eta_n \{A_0^{jk} A_0^{mn} - 2\bar{\alpha}^{jkl} \eta_l A_0^{mn} (\hat{b} - b_0) \\ &\quad + \bar{\alpha}^{jkl} \bar{\alpha}^{mnp} \eta_l \eta_p (\hat{b} - b_0)^2\} (\hat{b} - b_0)^2 \\ &\quad + \frac{2}{3} \eta_j \eta_k \eta_l (A_0^{jkl} - 3\gamma_{jkl,2} \bar{e} - 2\bar{\alpha}^{jklm} A_0^{lm}) (\hat{b} - b_0)^3 \\ &\quad + (\bar{\alpha}^{jkn} \bar{\alpha}^{lmn} - \frac{1}{2} \bar{\alpha}^{jklm}) \eta_j \eta_k \eta_l \eta_m (\hat{b} - b_0)^4 \\ &\quad + O_p(n^{-5/2}). \end{aligned} \quad (3.2)$$

The following nonparametric version Wilk's theorem is a direct consequence of expansion (3.2).

**THEOREM 3.1 (Wilks' theorem).** *Assume conditions (2.1). Then,*

$$P\{l(b_0) < c\} = P(\chi_1^2 < c) + o(1), \quad \text{as } n \rightarrow \infty.$$

*Proof.* Since  $\text{Var}(\hat{b} - b_0) = n^{-1} \sigma^2 / \sigma_x^2$ , by the Central Limit Theorem, we know that  $n^{1/2}(\hat{b} - b_0) \sigma_x / \sigma$  has asymptotically a standard normal distribution. Thus from (3.2),

$$l(b_0) = \frac{n\sigma_x^2}{\sigma^2} (\hat{b} - b_0)^2 + O_p(n^{-1/2}) = \chi_1^2 + o_p(1).$$

Hence the theorem is proved. ■

From Theorem 3.1 an empirical likelihood confidence interval for  $b_0$  with nominal coverage level  $\alpha$  can be constructed as follows. First find from  $\chi_1^2$  tables the value  $c_\alpha$  such that  $P(\chi_1^2 < c_\alpha) = \alpha$ . Then  $I_\alpha = \{b_0 | l(b_0) < c_\alpha\}$  is the  $\alpha$  level confidence interval for  $b_0$ . Theorem 3.1 ensures that  $I_\alpha$  has correct asymptotic coverage.

In the remainder of this section we investigate coverage accuracy of  $I_\alpha$ . To do this, we decompose  $l(b_0)$  from (3.2) as

$$l(b_0) = nR_b^2 + O_p(n^{-5/2}), \quad (3.3)$$

where  $R_b = R_{b1} + R_{b2} + R_{b3}$  and  $R_{bj} = O_p(n^{-j/2})$  for  $j = 1, 2, 3$ . Put

$$C_1 = -\frac{1}{2} \sigma^2 \bar{\alpha}^{jkl} \bar{\alpha}^{mnp} \eta_k \eta_l \eta_n \eta_p \left( \gamma_j \gamma_m + \frac{1}{9\sigma_x^2} \eta_j \eta_m \right) \\ + \eta_j \eta_k \eta_l \eta_m \left( \frac{1}{2} \bar{\alpha}^{ikn} \bar{\alpha}^{lmn} - \frac{1}{4} \bar{\alpha}^{iklm} \right).$$

Comparing (3.2) with (3.3) yields

$$R_{b1} = \frac{\sigma_x}{\sigma} (\hat{b} - b_0), \\ \frac{\sigma_x}{\sigma} R_{b2} = -\frac{1}{2} \eta_j \eta_k A_0^{jk} (\hat{b} - b_0) + \frac{1}{3} \bar{\alpha}^{jkl} \eta_j \eta_k \eta_l (\hat{b} - b_0)^2, \\ \frac{\sigma_x}{\sigma} R_{b3} = \eta_j \eta_k (A_0^{jk} \bar{\varepsilon} - \frac{1}{2} \gamma_{jk} \bar{\varepsilon}^2 + \frac{1}{2} A_0^{jl} A_0^{kl}) (\hat{b} - b_0) + C_1 (\hat{b} - b_0)^3 \\ - \left( \frac{\sigma^2}{2} \gamma_k \gamma_n \eta_j \eta_m + \frac{\sigma^2}{8\sigma_x^2} \eta_j \eta_k \eta_m \eta_n \right) A_0^{jk} A_0^{mn} (\hat{b} - b_0) \\ + \frac{1}{3} \eta_j \eta_k \eta_l A_0^{jkl} (\hat{b} - b_0)^2 \\ + \left\{ \sigma^2 \bar{\alpha}^{jkl} \eta_k \eta_m \eta_l \left( \gamma_j \gamma_n + \frac{1}{6\sigma_x^2} \eta_j \eta_n \right) - \bar{\alpha}^{jkm} \eta_j \eta_k \eta_n \right\} \\ \times A_0^{mn} (\hat{b} - b_0)^2. \quad (3.4)$$

Before we develop an Edgeworth expansion for  $l(b_0)$  we introduce some notations. From (3.4) we see that there exists a smooth function  $H$  such that  $R_b = H(\bar{U})$ , where  $\bar{U} = (\hat{b} - b_0, \bar{\varepsilon}, A_0^{11}, A_0^{12}, A_0^{22}, A_0^{111}, A_0^{112}, A_0^{122}, A_0^{222})$ . Let

$$B_1 = \begin{pmatrix} V_{n1}^{-1/2} \otimes V_{n1}^{-1/2} \\ V_{n1}^{-1/2} \otimes V_{n2}^{-1/2} \\ V_{n2}^{-1/2} \otimes V_{n2}^{-1/2} \end{pmatrix}$$

and

$$B_2 = \begin{pmatrix} V_{n1}^{-1/2} \otimes V_{n1}^{-1/2} \otimes V_{n1}^{-1/2} \\ V_{n1}^{-1/2} \otimes V_{n1}^{-1/2} \otimes V_{n2}^{-1/2} \\ V_{n1}^{-1/2} \otimes V_{n2}^{-1/2} \otimes V_{n2}^{-1/2} \\ V_{n2}^{-1/2} \otimes V_{n2}^{-1/2} \otimes V_{n2}^{-1/2} \end{pmatrix}$$

be  $3 \times 4$  and  $4 \times 8$  matrices, respectively, where  $\otimes$  is the Kronecker product of matrices and  $V_{nj}^{-1/2}$  is the  $j$ th row of  $V_n^{-1/2}$ ,  $j = 1, 2$ . From the definition

of  $A_0^{jk}$  and  $A_0^{kjl}$ ,  $\bar{U}$  can be expressed as  $\bar{U} = n^{-1} \sum U_i$ , where  $U_i$  is a vector of nine dimensions having from

$$U_i = [\sigma_x^{-2}(x_i - \bar{x}) \varepsilon_i, \varepsilon_i, \{(1x_i) \otimes (1x_i)\} B_1^T \varepsilon_i^2, \{(1x_i) \otimes (1x_i) \otimes (1x_i)\} B_2^T \varepsilon_i^3].$$

Put  $T_n = n^{-1} \sum \text{cov}(U_i)$  as the average covariance matrix of  $U_i$ 's and  $g_1$  as the density function of  $\chi_1^2$  distribution. Then, we have the following theorem.

**THEOREM 3.2.** Assume that

- (i) there exists positive constants  $C_1, C_2$  such that uniformly in  $n$ ,  $C_1 \leq v_{2n} \leq v_{1n} \leq C_2$ ; (ii)  $|x_i|$ 's for  $1 \leq i \leq n$  are uniformly bounded; (iii)  $E|\varepsilon_i|^{15} < \infty$ ; (iv) for every positive  $\tau$ ,  $\lim_{n \rightarrow \infty} \int_{\|\varepsilon_1\| > \tau n^{1/2}} \|\varepsilon_1\|^{15} = 0$ ; (v) the smallest eigenvalue of  $T_n$  is bounded away from zero; (vi) the characteristic function  $h$  of  $U_1$  satisfies  $\limsup_{|t| \rightarrow \infty} |h(t)| < 1$ . (3.5)

Then  $P\{l(b_0) < c_x\} = \alpha - (1 + \frac{1}{2}t_1 - \frac{1}{3}t_2)n^{-1}c_x g_1(c_x) + O(n^{-2})$ , where

$$t_1 = \frac{\mu_4}{\sigma^4 \sigma_x^4} m_4, \quad t_2 = \frac{\mu_3^2}{\sigma^6 \sigma_x^6} m_3^2, \quad m_j = n^{-1} \sum (x_i - \bar{x})^j, \quad \text{for } j = 3, 4.$$

*Proof.* Let  $k_{bj}$  be the  $j$ th cumulant of  $n^{1/2}R_b$ . Calculations show that

$$\begin{aligned} k_{b1} &= -\frac{1}{6} t_2^{1/2} n^{-1/2} + O(n^{-3/2}), \\ k_{b2} &= 1 + (1 + \frac{1}{2}t_1 - \frac{13}{36}t_2)n^{-1} + O(n^{-2}), \\ k_{bj} &= O(n^{-3/2}), \quad j \geq 3. \end{aligned}$$

A formal Edgeworth expansion for the distribution function of  $R_b$  can be constructed as

$$P(n^{1/2}R_b < x) = \int_{-\infty}^x \Psi(v) \phi(v) dv + O(n^{-3/2}), \quad (3.6)$$

where  $\Psi(v) = 1 + \frac{1}{6}t_2^{1/2}vn^{-1/2} + \frac{1}{2}(1 + \frac{1}{2}t_1 - \frac{1}{3}t_2)(v^2 - 1)n^{-1}$ . Accepting that expansion (3.6) may be justified, we establish an Edgeworth expansion for  $l(b_0)$  as

$$\begin{aligned} P\{l(b_0) < c\} &= P(-c^{1/2} < n^{1/2}R_b < c^{1/2}) + O(n^{-3/2}) \\ &= \int_{-c^{1/2}}^{c^{1/2}} \Psi(v) \phi(v) dv + O(n^{-3/2}) \\ &= \alpha - (1 + \frac{1}{2}t_1 - \frac{1}{3}t_2)n^{-1}cg_1(c) + O(n^{-3/2}), \end{aligned}$$

where  $g_1$  is the density function of  $\chi_1^2$  distribution. By the evenness and oddness of the polynomials in the above Edgeworth expansion, it can be shown that the  $O(n^{-3/2})$  term is actually  $O(n^{-2})$ . This leads us to Theorem 3.2.

It remains to check that expansion (3.6) is valid. Remember that  $R_b = H(\bar{U})$ , where  $H$  is a sufficient smooth function and  $\bar{U}$  is the mean of independent but not identically distributed random variable  $U_i$ 's. For this case, Bhattacharya and Rao [3, Theorem 20.2] have developed a valid Edgeworth expansion. It may be shown that conditions (3.5) imply the conditions of Theorem 20.2 of Bhattacharya and Rao [3]. Thus, a valid Edgeworth expansion for  $\bar{U}$  can be obtained. Consequently, the Edgeworth expansion of  $\bar{U}$  may be transformed by smooth function  $H$  to yield another valid Edgeworth expansion (3.6) for  $R_b$ , by using the results given by Skovgaard [4]. Therefore the theorem can be established. ■

Theorem 3.2 states that the empirical likelihood confidence interval  $I_x$  has coverage error at order of  $n^{-1}$ . By looking at the coefficient of the  $n^{-1}$  term in the Edgeworth expansion for the distribution function of  $l(b_0)$ , we see that the coverage error is dominated by a combination of four factors: the moments of  $\varepsilon_i$ , the "moments" of the fixed design points, the nominal coverage level, and the sample size  $n$ . We should note that the conditions listed in (3.5) are just sufficient conditions for deriving the Edgeworth expansion given in Theorem 3.2.

Based on the expression for  $R_{bj}$ ,  $j = 1, 2, 3$  in (3.4), we may show that

$$\begin{aligned} E\{l(b_0)\} &= n\{E(R_{b1})^2 + 2E(R_{b1}R_{b2}) + E(R_{b2})^2 + 2E(R_{b1}R_{b3})\} + O(n^{-2}) \\ &= 1 + (1 + \frac{1}{2}t_1 - \frac{1}{3}t_2)n^{-1} + O(n^{-2}). \end{aligned}$$

We see that the difference between the means of  $l(b_0)$  and limiting distribution  $\chi_1^2$  is of order  $n^{-1}$ . Next we show that Bartlett correction can reduce the coverage errors of empirical likelihood confidence intervals to order  $n^{-2}$ . Let  $\rho_b = 1 + \frac{1}{2}t_1 - \frac{1}{3}t_2$  be the Bartlett correction for  $l(b_0)$ . We have the following theorem about the Bartlett correctability of confidence interval  $I_x$ :

**THEOREM 3.3.** *Assume condition (3.5). Then,*

$$P\{l(b_0) < c_\alpha(1 + \rho_b n^{-1})\} = \alpha + O(n^{-2}).$$

*Proof.* The method of proof is identical to that of Theorem 2.3 of Chen [2].

Theorem 3.3 implies that a simple Bartlett correction can increase the coverage accuracy of empirical likelihood confidence intervals for  $b_0$  from  $O(n^{-1})$  to  $O(n^{-2})$ . However,  $\rho_b$  is usually unknown because of

unknown  $\mu_3$  and  $\mu_4$ , the third and fourth moments of  $\varepsilon_1$ , in  $t_1$  and  $t_2$ . An  $n^{1/2}$ -consistent estimate of  $\rho_b$ , denoted by  $\hat{\rho}_b$ , can be obtained by defining  $\hat{\rho}_b = 1 + \frac{1}{2} \hat{t}_1 - \frac{1}{3} \hat{t}_2$ , where  $\hat{t}_1$  and  $\hat{t}_2$  are obtained by replacing  $\mu_3$  and  $\mu_4$  in  $t_1$  and  $t_2$  by  $\hat{\mu}_3$  and  $\hat{\mu}_4$ , respectively, where  $\hat{\mu}_3$  and  $\hat{\mu}_4$  are the moment estimators of  $\mu_3$  and  $\mu_4$ . We may get the same order of accuracy by replacing  $\rho_b$  with  $\hat{\rho}_b$  in Theorem 2.3, under moderate conditions such as: the joint distribution of components of the  $l(b_0)$  and  $\hat{\rho}_b$  admits multivariate Edgeworth expansions.

#### 4. EMPIRICAL LIKELIHOOD CONFIDENCE INTERVAL FOR MEANS

In this section we construct empirical likelihood confidence intervals for the mean value  $y_0 = E(y|x=x_0) = a_0 + b_0 x_0$ , for any fixed  $x_0$ . Since  $y_0 = a_0$  when  $x_0 = 0$ , we may confine our attention to constructing empirical likelihood confidence intervals for a general  $y_0$ . The empirical log likelihood ratio for  $y_0$ , denoted as  $l(y_0)$ , may be obtained by minimizing  $l(a, b)$  given in (2.3), under the constraint of  $a + b x_0 = y_0$ , that is,

$$l(y_0) = l(\tilde{a}, \tilde{b}) = \min_{a + b x_0 = y_0} l(a, b).$$

Suppose  $\tilde{a}$  and  $\tilde{b}$  have expansions  $\tilde{a} = \hat{a} + a_1 + a_2 + a_3$  and  $\tilde{b} = \hat{b} + b_1 + b_2 + b_3$ , where  $a_j, b_j = O_p(n^{-j/2})$ ,  $j = 1, 2, 3$ . Note that we use notations  $\tilde{a}$  and  $a_j$  again here, but with meanings different from those in Section 3. In the following,  $a_j, b_j, j = 1, 2, 3$ , are determined successively. Put

$$\begin{aligned} \beta_j &= n^{-1} \sum g_j(x_i) x_i, & \beta_{jk} &= n^{-1} \sum g_{jk}(x_i) x_i, \\ \beta_{jkl} &= n^{-1} \sum g_{jkl}(x_i) x_i, & \beta_{jk2} &= n^{-1} \sum g_{jk}(x_i) x_i^2, \\ \beta_{jk,1}(a, b) &= n^{-1} \sum g_{jk}(x_i)(y_i - a - b x_i), \\ \beta_{jkl,2}(a, b) &= n^{-1} \sum g_{jkl}(x_i)(y_i - a - b x_i)^2. \end{aligned}$$

Let  $W_0 = \hat{a} + \hat{b} x_0 - y_0 = O_p(n^{-1/2})$  and  $t^j = A^j(\hat{a}, \hat{b}) - \gamma_j a_1 - \beta_j b_1$ . After some algebra we may show that

$$\begin{aligned} a_1 &= -\frac{\sigma_x^2 + \bar{x}(\bar{x} - x_0)}{\sigma_x^2 + (\bar{x} - x_0)^2} W_0, & b_1 &= \frac{(\bar{x} - x_0)}{\sigma_x^2 + (\bar{x} - x_0)^2} W_0, \\ b_2 &= \{(\beta_p - \gamma_p x_0)(\beta_p - \gamma_p x_0)\}^{-1} \\ &\quad \times \{\bar{x}^{jk} t^j t^k (\beta_l - \gamma_l x_0) - A^{jk}(a, \hat{b}) t^j (\beta_k - \gamma_k x_0)\}, \\ a_2 &= -b_2 x_0, & a_3 &= O_p(n^{-2}), & b_3 &= O_p(n^{-2}). \end{aligned}$$

Now substitute  $\tilde{a}, \tilde{b}$  into the formula for  $n^{-1}l(\tilde{a}, \tilde{b})$ , obtaining

$$\begin{aligned}
 n^{-1}l(y_0) = & t^j t^j - \{A^{jk}(\hat{a}, \hat{b}) - 2\gamma_{jk,1}a_1 - 2\beta_{jk,1}b_1 \\
 & + \gamma_{jk}a_1^2 + 2\beta_{jk}a_1b_1 + \beta_{jk2}b_1^2\} t^j t^k \\
 & + \frac{2}{3} \tilde{\alpha}^{jkl} t^j t^k t^l - (\beta_j - \gamma_j x_0)(\beta_j - \gamma_j x_0) b_2^2 \\
 & + A^{jl}(\hat{a}, \hat{b}) A^{kl}(\hat{a}, \hat{b}) t^j t^k \\
 & + \frac{2}{3} \{A^{jkl}(\hat{a}, \hat{b}) - 3\gamma_{jkl,2}a_1 - 3\beta_{jkl,2}b_1\} t^j t^k t^l \\
 & - 2\tilde{\alpha}^{jkm} A^{lm}(\hat{a}, \hat{b}) t^j t^k t^l + (\tilde{\alpha}^{jkn} \tilde{\alpha}^{lmn} - \frac{1}{2} \tilde{\alpha}^{jklm}) t^j t^k t^l t^m \\
 & + O_p(n^{-5/2}). \tag{4.1}
 \end{aligned}$$

Define  $\alpha^2(x_0) = \sigma_x^2 / \{\sigma_x^2 + (\bar{x} - x_0)^2\}$  and  $\xi^j = u_j^1 + u_j^2 x_0, j = 1, 2$ . Then we have  $t^j = \alpha^2(x_0) \xi^j W_0$ ,  $t^j t^j = \alpha^2(x_0) \sigma^{-2} W_0^2$ , and  $(\beta_j - \gamma_j x_0)(\beta_j - \gamma_j x_0) = \sigma_x^2 \sigma^{-2} \alpha^{-2}(x_0)$ . Substituting these formulae into (4.1), we obtain

$$\begin{aligned}
 n^{-1}l(y_0) = & \alpha^2(x_0) \sigma^{-2} W_0^2 - \alpha^4(x_0) \xi^j \xi^k A_0^{jk} W_0^2 \\
 & + \frac{2}{3} \alpha^6(x_0) \tilde{\alpha}^{jkl} \xi^j \xi^k \xi^l W_0^3 - \sigma_x^{-2} \sigma^2 \alpha^6(x_0) \\
 & \times \{\tilde{\alpha}^{jkl} \alpha^2(x_0) \xi^j \xi^k (\beta_l - \gamma_l x_0) W_0^2 - \xi^j (\beta_k - \gamma_k x_0) A_0^{jk} W_0\}^2 \\
 & + 2\alpha^4(x_0) \xi^j \xi^k (\beta_{jk,1} - \gamma_{jk,1} x_0)(b_1 + \hat{b} - b_0) W_0^2 \\
 & - \alpha^4(x_0) \xi^j \xi^k \{(\beta_{jk2} - 2\beta_{jk} x_0 + \gamma_{jk} x_0^2) \\
 & \times (b_1 + \hat{b} - b_0)^2 - A_0^{jl} A_0^{kl}\} W_0^2 \\
 & + \alpha^6(x_0) \xi^j \xi^k \xi^l \{\frac{2}{3} A_0^{jkl} - 2(\beta_{jkl} - \gamma_{jkl} \sigma^2 x_0) \\
 & \times (b_1 + \hat{b} - b_0) - 2\tilde{\alpha}^{jkm} A_0^{lm}\} W_0^3 \\
 & - 2\alpha^6(x_0) \xi^j \xi^k \xi^l W_0^3 + \alpha^8(x_0) \\
 & \times (\tilde{\alpha}^{jkn} \tilde{\alpha}^{lmn} - \frac{1}{2} \tilde{\alpha}^{jklm}) \xi^j \xi^k \xi^l \xi^m W_0^4 \\
 & + O_p(n^{-5/2}). \tag{4.2}
 \end{aligned}$$

The following nonparametric version of Wilks' theorem is a conclusion of (4.2).

**THEOREM 4.1** (Wilks' theorem). *Assume conditions (2.1) and (2.2). Then,*

$$P\{l(y_0) < c\} = P(\chi_1^2 < c) + o(1), \quad n \rightarrow \infty.$$

*Proof.* From (4.2), we know that

$$\begin{aligned}
 l(y_0) = & n\alpha(x_0)^2 \sigma^{-2} \hat{W}_0^2 + O_p(n^{-1/2}) \\
 = & n\sigma^{-2} \frac{\sigma_x^2}{\sigma_x^2 + (\bar{x} - x_0)^2} \hat{W}_0^2 + O_p(n^{-1/2}).
 \end{aligned}$$

Thus the theorem is proved by the fact that  $W_0$  is asymptotically normal with mean zero and variance  $n^{-1}\sigma^2\sigma_x^{-2}\{\sigma_x^2 + (\bar{x} - x_0)^2\}$ . ■

An empirical likelihood confidence interval for  $y_0$  with asymptotic coverage level  $\alpha$  can be constructed as  $J_x = \{y_0 | l(y_0) < c_x\}$  such that  $P(\chi_1^2 < c_x) = \alpha$ . In the rest of this section we investigate the second order properties of  $J_x$ .

A signed root decomposition of  $l(y_0)$  can be obtained from (4.2) as

$$l(y_0) = nR_{y_0}^2 + O_p(n^{-5/2}),$$

where  $R_{y_0} = R_{y_01} + R_{y_02} + R_{y_03}$ , and  $R_{y_0j} = O_p(n^{-j/2})$  for  $j = 1, 2, 3$ . A little algebra shows that

$$\begin{aligned} R_{y_01} &= \alpha(x_0) \sigma^{-1} W_0, \\ R_{y_02} &= \alpha^3(x_0) \sigma \xi^j \xi^k \left\{ -\frac{1}{2} A_0^{jk} W_0 + \frac{1}{3} \alpha^2(x_0) \bar{\alpha}^{jkl} \xi^l W_0^2 \right\}, \\ R_{y_01} R_{y_03} &= \alpha^4(x_0) \xi^j \xi^k (\beta_{jk,1} - \gamma_{jk,1} x_0) (b_1 + \hat{b} - b_0) W_0^2 + C_2 W_0^4 \\ &\quad - \frac{1}{2} \alpha^4(x_0) \xi^j \xi^k (\beta_{jk,2} - 2\beta_{jk} x_0 + \gamma_{jk} x_0^2) (b_1 + \hat{b} - b_0)^2 W_0^2 \\ &\quad - \alpha^6(x_0) \sigma^2 \xi^j \xi^m \left\{ \frac{1}{2} \sigma_x^{-2} (\beta_k - \gamma_l x_0) (\beta_n - \gamma_n x_0) + \frac{1}{8} \xi^k \xi^n \right\} \\ &\quad \times A_0^{jk} A_0^{mn} W_0^2 + \alpha^8(x_0) \sigma^2 \\ &\quad \times \left\{ \sigma_x^{-2} \bar{\alpha}^{jkl} \xi^j \xi^k \xi^m (\beta_l - \gamma_l x_0) (\beta_n - \gamma_n x_0) \right. \\ &\quad \left. + \frac{1}{6} \bar{\alpha}^{jkl} \xi^j \xi^k \xi^l \xi^m \xi^n - \alpha^{-2}(x_0) \sigma^{-2} \bar{\alpha}^{jkm} \xi^j \xi^k \xi^n \right\} A_0^{mn} W_0^3 \\ &\quad + \frac{1}{2} \alpha^4(x_0) \xi^j \xi^k A_0^{jl} A_0^{kl} W_0^2 + \frac{1}{3} \alpha^6(x_0) \xi^j \xi^k \xi^l A_0^{jkl} W_0^3 \\ &\quad - \alpha^6(x_0) \sigma^2 \xi^j \xi^k \xi^l (\beta_{jkl,2} - \gamma_{jkl,2} x_0) (b_1 + \hat{b} - b_0) W_0^3, \end{aligned}$$

where

$$\begin{aligned} C_2 &= -\frac{1}{2} \alpha^{10}(x_0) \sigma^2 \sigma_x^{-2} \left\{ \bar{\alpha}^{jkl} \xi^j \xi^k (\beta_l - \gamma_l x_0) \right\}^2 \\ &\quad - \frac{1}{18} \alpha^{10}(x_0) \sigma^2 \left\{ \bar{\alpha}^{jkl} \xi^j \xi^k \xi^l \right\}^2 \\ &\quad + \alpha^8(x_0) \left( \frac{1}{2} \bar{\alpha}^{jkn} \bar{\alpha}^{lmn} - \frac{1}{4} \bar{\alpha}^{jklm} \right) \xi^j \xi^k \xi^l \xi^m. \end{aligned}$$

In order to assess the coverage accuracy of the confidence interval  $J_x$  we establish an Edgeworth expansion for the distribution of  $l(y_0)$ . To this end, we note from expressions for  $R_{y_0j}$   $j = 1, 2, 3$ , that there is a smooth function  $Q_1$  such that  $R_{y_0} = Q_1(\bar{S})$ , where

$$\bar{S} = (W_0, b_1 + \hat{b} - b_0, A_0^{11}, A_0^{12}, A_0^{22}, A_0^{111}, A_0^{112}, A_0^{122}, A_0^{222}).$$

Since  $b_1 = (\bar{x} - x_0) \{ \bar{x} - x_0 \}^{-1} W_0$  and  $W_0 = \bar{\varepsilon} + (\bar{x} - x_0)(\hat{b} - b_0)$ , there

exists another smooth function  $Q_2$  such that  $\bar{S} = Q_2(\bar{U})$ , where  $\bar{U}$  was defined in Section 3. So, putting  $Q = Q_1 Q_2$ , we have  $R_{y_0} = Q(\bar{U})$ . Define

$$s_1 = \alpha^4(x_0) \sigma^{-4} \mu_4 q_1, \quad s_2 = \alpha^6(x_0) \sigma^{-6} \mu_3^2 q_2^2, \quad s_3 = \alpha^4(x_0) q_3,$$

where

$$\begin{aligned} q_1 &= 1 + 6 \frac{(\bar{x} - x_0)^2}{\sigma_x^2} - 4 \frac{(\bar{x} - x_0)^3}{\sigma_x^6} m_3 + \frac{(\bar{x} - x_0)^4}{\sigma_x^8} m_4, \\ q_2 &= 1 + 3 \frac{(\bar{x} - x_0)^2}{\sigma_x^2} - \frac{(\bar{x} - x_0)^3}{\sigma_x^6} m_3, \\ q_3 &= 1 - 3 \frac{(\bar{x} - x_0)^2}{\sigma_x^2} + \frac{(\bar{x} - x_0)^2}{\sigma_x^6} m_4 + \frac{(\bar{x} - x_0)^4}{\sigma_x^4} \\ &\quad + 2 \left\{ \frac{(\bar{x} - x_0)^3}{\sigma_x^6} - \frac{(\bar{x} - x_0)}{\sigma_x^4} \right\} m_3. \end{aligned}$$

Now the coverage accuracy of confidence interval  $J_x$  is discussed in the following theorem.

**THEOREM 4.2.** Assume condition (3.5). Then,

$$P\{l(y_0) < c_x\} = \alpha - \left( \frac{s_1}{2} - \frac{s_2}{3} + s_3 \right) n^{-1} c_x g_1(c_x) + O(n^{-3/2}). \quad (4.3)$$

*Proof.* Let  $k_{y_0j}$ ,  $j = 1, 2, \dots$ , denote the  $j$ th cumulants of  $n^{1/2} R_{y_0}$ . Calculations show that

$$\begin{aligned} k_{y_01} &= -\frac{1}{6} s_2^{1/2} n^{-1/2} + O(n^{-3/2}), \\ k_{y_02} &= 1 + \left( \frac{1}{2} s_1 - \frac{13}{36} s_2 + s_3 \right) n^{-1} + O(n^{-2}), \\ k_{y_0j} &= O(n^{-3/2}), \quad j \geq 3. \end{aligned} \quad (4.4)$$

A formal Edgeworth expansion for the distribution of  $n^{1/2} R_{y_0}$  can be set up from (4.4) as

$$P(n^{1/2} R_b < x) = \int_{-\infty}^x \Pi(v) \phi(v) dv + O(n^{-3/2}), \quad (4.5)$$

where  $\Pi(v) = 1 + \frac{1}{6} s_2^{1/2} v n^{-1/2} + \frac{1}{2} \left( \frac{1}{2} s_1 - \frac{1}{3} s_2 + s_3 \right) (v^2 - 1) n^{-1}$ . The validity of expansion (4.5) can be demonstrated in the same way that validity for expansion (3.6) in the proof of Theorem 3.2 was demonstrated. And the theorem can be obtained from expansion (4.5) in the same way that we derived Theorem 3.2. ■

Theorem 4.2 states that the coverage errors of empirical likelihood confidence intervals for  $y_0 = a + bx_0$  are of order of  $n^{-1}$ , provided that the  $x_0$  is fixed and independent of sample size  $n$ . From the  $n^{-1}$  order term in (4.3) and the definitions of  $s_1$ ,  $s_2$ , and  $s_3$ , we see that the coverage error is dominated by the combination of the following five factors: the moments of  $\varepsilon_i$ , the “moments” of the fixed design points, the nominal coverage level, the sample size  $n$ , and the size of  $(\bar{x} - x_0)/\sigma_x$ —the standard distance between  $x_0$  and the centre,  $\bar{x}$ , of the design points.

In analogy with the Bartlett correction for the slope parameter  $b_0$  developed in Theorem 3.3, we do the same thing here for  $y_0$ . Calculations reveal that

$$E\{l(y_0)\} = n\{E(R_{y_0})^2 + O(n^{-2})\} = 1 + \left(\frac{s_1}{2} - \frac{s_2}{3} + s_3\right)n^{-1} + O(n^{-2}).$$

Put  $\rho_{y_0} = (s_1/2 - s_2/3 + s_3)$ , the Bartlett correction for  $l(y_0)$ . The Bartlett correction property for the empirical likelihood confidence intervals for  $y_0$  is proved by the following theorem.

**THEOREM 4.3.** *Assume conditions (3.5). For any  $\alpha > 0$  and fixed  $x_0$ ,*

$$P\{l(y_0) < c_\alpha(1 + \rho_{y_0}n^{-1})\} = \alpha + O(n^{-2}).$$

Thus a simple scale adjustment can increase the coverage accuracy of empirical likelihood confidence intervals for  $y_0$  from  $O(n^{-1})$  to  $O(n^{-2})$ . In practice,  $\rho_{y_0}$  is usually unknown, because  $\mu_3$  and  $\mu_4$  are unknown. However, an  $n^{1/2}$ -consistent estimate  $\hat{\rho}_{y_0}$  of  $\rho_{y_0}$  can be obtained by replacing  $\sigma^2$ ,  $\mu_3$ , and  $\mu_4$  with  $\hat{\sigma}^2$ ,  $\hat{\mu}_3$ , and  $\hat{\mu}_4$ , respectively, in  $s_1$  and  $s_2$ , that is,  $\hat{\rho}_{y_0} = (\hat{s}_1/2 - \hat{s}_2/3 + s_3)$ , where  $\hat{s}_1 = \alpha^4(x_0) \hat{\sigma}^{-4} \hat{\mu}_4 q_1$  and  $\hat{s}_2 = \alpha^6(x_0) \hat{\sigma}^{-6} \hat{\mu}_3^2 q_2^2$ . It may be shown that we may obtain the same order of accuracy by replacing  $\rho_{y_0}$  with  $\hat{\rho}_{y_0}$  in Theorem 3.3.

## 5. SIMULATION RESULTS

This section describes simulation experiments carried out to examine the coverage properties of the empirical likelihood confidence intervals for  $b_0$  and  $y_0$  proposed in the previous sections. The following simple linear regression model was treated:

$$y_i = 1 + x_i + \varepsilon_i, \quad i = 1, \dots, n.$$

The data set  $x_i$  was the one which has been displayed in Chen [2]. We chose sample sizes  $n = 15, 30, 50$  and nominal coverage level  $\alpha = 0.90, 0.95$ .

We assigned two error patterns for  $\varepsilon_i$ . One was  $\varepsilon_i = N(0, 1)$  and another was  $\varepsilon_i = E(1.00) - 1.00$ , where  $N(0, 1)$  and  $E(1.00)$  were random variables with the standard normal distribution and the exponential distribution with unit mean, respectively. The normal and exponential random variables were generated by the routines of Press *et al.* [5].

For each combination of  $n$ ,  $\alpha$ , and  $\varepsilon_i$  we display in Table I the coverages of the uncorrected confidence intervals and two Bartlett corrected confidence intervals based on 10,000 simulations. One of the corrected confidence intervals used the theoretical Bartlett correction; another used the empirical Bartlett correction. Standard errors are given for each of the simulated coverages. To empirically justify the expansions developed in Theorems 3.2 and 4.2, we also calculated theoretical coverages up to second order in Edgeworth expansions for  $l(b_0)$  and  $l(y_0)$ . Since the coverages can be obtained without simulation, they are called “predicted coverages”.

The following broad conclusions may be drawn from the results summarized in Table I. First, the differences between the uncorrected coverages and their corresponding “predicted coverages” converge to zero as  $n$  increases. This gives empirical justification for Theorems 3.2 and 4.2. Second, substantial improvements on coverage accuracy have been made by implementing Bartlett corrections. This can be observed by looking at

TABLE I  
Estimated True coverages, from 10,000 Simulations, of  $\alpha$ -level Empirical Likelihood  
Confidence Regions for  $b_0$  and  $y_0$ 's

| $\varepsilon_i$                         |                    | $N(0, 1)$    |               | $E(1.00) - 1.00$ |              |
|-----------------------------------------|--------------------|--------------|---------------|------------------|--------------|
| $n$                                     | $\alpha$           | 0.90         | 0.95          | 0.90             | 0.95         |
| (1) Coverages for slope parameter $b_0$ |                    |              |               |                  |              |
| 15                                      | predic.            | 0.840        | 0.909         | 0.750            | 0.849        |
|                                         | uncorr.            | 0.803 (0.40) | 0.860 (0.35)  | 0.789 (0.41)     | 0.858 (0.35) |
|                                         | $\rho_{b_0}$       | 0.859 (0.35) | 0.911 (0.28)  | 0.904 (0.30)     | 0.950 (0.22) |
|                                         | $\hat{\rho}_{b_0}$ | 0.853 (0.35) | 0.906 (0.29)  | 0.856 (0.35)     | 0.916 (0.28) |
| 30                                      | predic.            | 0.878        | 0.935         | 0.845            | 0.913        |
|                                         | uncorr.            | 0.862 (0.35) | 0.919 (0.27)  | 0.840 (0.37)     | 0.902 (0.30) |
|                                         | $\rho_{b_0}$       | 0.884 (0.32) | 0.935 (0.25)  | 0.880 (0.31)     | 0.931 (0.24) |
|                                         | $\hat{\rho}_{b_0}$ | 0.883 (0.32) | 0.934 (0.25)  | 0.871 (0.34)     | 0.928 (0.26) |
| 50                                      | predic.            | 0.888        | 0.939         | 0.870            | 0.930        |
|                                         | uncorr.            | 0.882 (0.32) | 0.9386 (0.24) | 0.860 (0.35)     | 0.926 (0.26) |
|                                         | $\rho_{b_0}$       | 0.896 (0.31) | 0.948 (0.22)  | 0.887 (0.32)     | 0.944 (0.23) |
|                                         | $\hat{\rho}_{b_0}$ | 0.896 (0.31) | 0.948 (0.22)  | 0.880 (0.33)     | 0.938 (0.24) |

(Table continued)

TABLE 1-Continued

| $\varepsilon_t$                                          |                    | $N(0, 1)$    |              | $E(1.00) - 1.00$ |              |
|----------------------------------------------------------|--------------------|--------------|--------------|------------------|--------------|
| $n$                                                      | $\alpha$           | 0.90         | 0.95         | 0.90             | 0.95         |
| (2) Coverages for intercept parameter $a_0$              |                    |              |              |                  |              |
| 15                                                       | predic.            | 0.858        | 0.921        | 0.802            | 0.884        |
|                                                          | uncorr.            | 0.822 (0.38) | 0.884 (0.32) | 0.805 (0.40)     | 0.868 (0.34) |
|                                                          | $\rho_{y_0}$       | 0.861 (0.35) | 0.918 (0.27) | 0.883 (0.32)     | 0.927 (0.26) |
|                                                          | $\hat{\rho}_{y_0}$ | 0.857 (0.35) | 0.915 (0.28) | 0.848 (0.36)     | 0.900 (0.30) |
| 30                                                       | predic.            | 0.880        | 0.937        | 0.865            | 0.921        |
|                                                          | uncorr.            | 0.864 (0.34) | 0.922 (0.27) | 0.840 (0.37)     | 0.901 (0.30) |
|                                                          | $\rho_{y_0}$       | 0.888 (0.32) | 0.937 (0.24) | 0.874 (0.33)     | 0.933 (0.25) |
|                                                          | $\hat{\rho}_{y_0}$ | 0.884 (0.32) | 0.936 (0.24) | 0.863 (0.34)     | 0.922 (0.27) |
| 50                                                       | predic.            | 0.887        | 0.941        | 0.871            | 0.931        |
|                                                          | uncorr.            | 0.883 (0.32) | 0.933 (0.25) | 0.860 (0.35)     | 0.920 (0.27) |
|                                                          | $\rho_{y_0}$       | 0.894 (0.31) | 0.942 (0.23) | 0.884 (0.32)     | 0.942 (0.23) |
|                                                          | $\hat{\rho}_{y_0}$ | 0.894 (0.31) | 0.942 (0.23) | 0.877 (0.33)     | 0.933 (0.25) |
| (3) Coverages for mean parameter $y_0$ with $x_0 = 5.00$ |                    |              |              |                  |              |
| 15                                                       | predic.            | 0.865        | 0.926        | 0.840            | 0.909        |
|                                                          | uncorr.            | 0.837 (0.37) | 0.899 (0.30) | 0.815 (0.39)     | 0.869 (0.34) |
|                                                          | $\rho_{y_0}$       | 0.871 (0.34) | 0.924 (0.27) | 0.868 (0.34)     | 0.908 (0.29) |
|                                                          | $\hat{\rho}_{y_0}$ | 0.867 (0.34) | 0.922 (0.27) | 0.846 (0.36)     | 0.893 (0.31) |
| 30                                                       | predic.            | 0.885        | 0.940        | 0.875            | 0.934        |
|                                                          | uncorr.            | 0.882 (0.32) | 0.936 (0.25) | 0.861 (0.35)     | 0.922 (0.27) |
|                                                          | $\rho_{y_0}$       | 0.897 (0.30) | 0.946 (0.23) | 0.884 (0.32)     | 0.938 (0.24) |
|                                                          | $\hat{\rho}_{y_0}$ | 0.897 (0.30) | 0.946 (0.23) | 0.876 (0.33)     | 0.932 (0.25) |
| 50                                                       | predic.            | 0.889        | 0.943        | 0.879            | 0.936        |
|                                                          | uncorr.            | 0.887 (0.32) | 0.937 (0.24) | 0.871 (0.33)     | 0.923 (0.27) |
|                                                          | $\rho_{y_0}$       | 0.898 (0.30) | 0.945 (0.23) | 0.891 (0.31)     | 0.939 (0.25) |
|                                                          | $\hat{\rho}_{y_0}$ | 0.897 (0.30) | 0.944 (0.23) | 0.884 (0.32)     | 0.933 (0.25) |

Note. Rows headed "predic.," "uncorr.," " $b_0$ " or " $y_0$ " and " $\hat{b}_0$ " or " $\hat{y}_0$ " give the predicted, uncorrected, and Bartlett-corrected coverages, respectively. The figures in parentheses are  $10^2$  times the standard errors associated with the simulated coverages.

both the standard errors and the absolute errors. Third, the empirical Bartlett correction performs similarly to its theoretical Bartlett correction counterpart, except for small sample skewed case.

#### ACKNOWLEDGMENTS

I express my gratitude to Professor Peter Hall for this encouragement and guidance during preparation of this paper. I also thank Dr. Andrew Wood for beneficial discussions.

REFERENCES

[1] OWEN, A. (1991). Empirical likelihood for linear models. *Ann. Statist.* **19** 1725–1747.

[2] CHEN, S. X. (1993). “On the Accuracy of the Empirical Likelihood Confidence Regions for Linear Regression Model.” *Ann. Inst. Statist. Math.* **45** 621–637.

[3] BHATTACHARYA, R. N., AND RAO, R. R. (1976). *Normal Approximation and Asymptotic Expansions*. Wiley, New York.

[4] SKOVGAARD, IB. M. (1981). Transformation of an Edgeworth expansion by a sequence of smooth functions. *Scand. J. Statist.* **8** 207–217.

[5] PRESS, W. H., FLANNERY, B. F., TEUKOLSKY, S. A., AND VETTERLING, W. T. (1989). *Numerical Recipes: The Art of Scientific Computing*. Cambridge Univ. Press, Cambridge.