Biometrika (2000), **87,** 4, *pp.* 946–953

Printed in Great Britain

Empirical likelihood confidence intervals for local linear smoothers

By SONG XI CHEN

Department of Statistics and Applied Probability, National University of Singapore, Singapore 117543

stacsx@nus.edu.sg

AND YONG SONG QIN

Department of Statistics and Finance, University of Science and Technology of China, Hefei, Anhui 230026, China

ysqin@ustc.edu.cn

Summary

Empirical likelihood is considered in conjunction with the local linear smoother to construct confidence intervals for a nonparametric regression function with bounded support. The coverage error of the empirical likelihood confidence intervals is evaluated and is shown to be of the same order throughout the support of the regression function. This is a significant improvement over confidence intervals based directly on the asymptotic normal distribution of the local linear estimator, which have a larger order of coverage error near the boundary. This improvement is attributable to the natural variance estimator that empirical likelihood implicitly chooses for the local linear smoother.

Some key words: Bartlett correction; Confidence interval; Coverage error; Empirical likelihood; Local linear smoother; Nonparametric regression.

1. Introduction

Let $\{(X_i, Y_i)\}_{i=1}^n$ be n independent pairs of observations where the Y_i are responses associated with design points $X_i \in [0, \infty)$ according to $Y_i = m(X_i) + \varepsilon_i$, where m is an unknown regression function with support $[0, \infty)$, and the ε_i are independent errors each with zero mean and finite variance.

The advantages of local polynomial smoothing in nonparametric estimation of m have been explored by Fan (1993), Ruppert & Wand (1994) and others, and are well summarised in Fan & Gijbels (1996).

In this paper we consider the construction of confidence intervals for $m(x)$ at fixed x based on $\hat{m}_l(x)$, a local linear smoother of $m(x)$. Recently, in an unpublished La Trobe University Technical Report, S. X. Chen and Y. S. Qin studied confidence intervals based on the asymptotic normal distribution of $\hat{m}_l(x)$, and found that the coverage error near the boundary was of a larger order than the coverage error in the interior. This is surprising as $\hat{m}_l(x)$ is well known for its ability to adapt to the boundary. The reason, as shown in the technical report by Chen and Qin, is that the second-order term of $\text{var}\{\hat{m}_l(x)\}$ is of a larger order near the boundary.

In this paper, we construct confidence intervals for $m(x)$ using empirical likelihood in conjunction with the local linear smoother. Empirical likelihood, introduced by Owen (1988, 1990), is a nonparametric approach for constructing confidence intervals. Not only does it produce confidence intervals that have natural shape and respect the range of the parameter, but it also has the advantage of studentising automatically. This means that, in the current context, empirical likelihood will select a variance estimator for $\hat{m}_l(x)$ implicitly. It turns out that the coverage error of the empirical likelihood confidence intervals is the same throughout $[0, \infty)$, as a result of the natural variance estimator implicitly prescribed.

In § 2 we introduce the empirical likelihood confidence intervals for $m(x)$, the coverage accuracy of the confidence intervals is studied in § 3, and some empirical results are reported in § 4.

2. Empirical likelihood confidence intervals

Let f be the density function of the design points X_i, and let

$$V(x) = E[\{Y - m(x)\}^2 \mid X = x] \tag{1}$$

be the conditional variance of Y given $X = x$. Let K be a kernel function which is itself a probability density, and let h be the smoothing bandwidth. We assume the following regularity conditions:

(i) K is a symmetric bounded density function compactly supported on $[-1, 1]$;

(ii) as $n \to \infty$, $h \to 0$ and $nh \to \infty$, and there is a real number $s \geqslant 4$ such that $E|Y|^s < \infty$, $nh^{2s} \to 0$ and $n^{s-2}h^s \to \infty$;

(iii) f, m and V have continuous derivatives up to the second order in a neighbourhood of x, and both $f(x) > 0$ and $V(x) > 0$.

The local linear estimator for $m(x)$ at any $x \in [0, \infty)$ is

$$\hat{m}_l(x) = \frac{\sum_{i=1}^n W_i Y_i}{\sum_{i=1}^n W_i}, \tag{2}$$

where

$$W_i = K_h(x - X_i)\left\{s_{n,2} - \frac{(x - X_i)s_{n,1}}{h}\right\}, \quad s_{n,l} = (nh)^{-1} \sum_{i=1}^n \frac{K_h(x - X_i)(x - X_i)^l}{h^l},$$

for $l = 0$, 1 and 2, and $K_h(.) = K(./h)$.

Let $p_1, \ldots, p_n$ be nonnegative numbers adding to unity. For any $x \in [0, \infty)$, the empirical likelihood at a candidate value θ of $E\{\hat{m}_l(x)\}$ is

$$L(\theta) = \sup_{\sum p_i W_i (Y_i - \theta) = 0} \prod_{i=1}^n p_i.$$

If we apply the standard method of Lagrange multipliers to find the optimal p_i, the log empirical likelihood ratio is

$$l(\theta) = -2 \log\{L(\theta)n^n\} = 2 \sum \log\{1 + \lambda(\theta)W_i(Y_i - \theta)\},$$

where $\lambda(\theta)$ satisfies

$$\sum_{i=1}^n W_i(Y_i - \theta)\{1 + \lambda(\theta)W_i(Y_i - \theta)\}^{-1} = 0. \tag{3}$$

The empirical likelihood formulated above is for $E\{\hat{m}_l(x)\} = m(x) + \text{bias}$, rather than for $m(x)$. To convert the empirical likelihood to one for $m(x)$, one can either correct explicitly via direct estimation of the bias (Hall, 1991), or one can reduce the bias via undersmoothing (Hall, 1992; Neumann, 1995). Hall (1992) showed that better coverage accuracy is achieved by the undersmoothing method. We consider the latter approach in this paper. The Wilks' theorem given below assures that $l\{m(x)\}$ has a limiting chi-squared distribution under the conditions (i)–(iii) and a condition that implies undersmoothing.

Let

$$U_i = W_i\{Y_i - m(x)\}, \quad \bar{U}_j = (nh)^{-1} \sum_{i=1}^n U_i^j, \quad \alpha_j(x/h) = \int_{-1}^{x/h} u^j K(u)\, du$$

for positive integers j. For nonnegative integers i, j, k, define

$$\bar{\omega}_{ijk} = (nh)^{-1} \sum_{l=1}^{n} \left(\frac{x - X_l}{h}\right)^i K_h^j(x - X_l)\{Y_i - m(x)\}^k, \quad \mu_{ijk} = E(\bar{\omega}_{ijk}).$$

It may be shown that

$$\bar{U}_1 = \bar{\omega}_{210}\bar{\omega}_{011} - \bar{\omega}_{110}\bar{\omega}_{111}, \tag{4}$$

$$\bar{U}_2 = \bar{\omega}_{210}^2\bar{\omega}_{022} - 2\bar{\omega}_{110}\bar{\omega}_{210}\bar{\omega}_{122} + \bar{\omega}_{110}^2\bar{\omega}_{222}, \tag{5}$$

$$\bar{U}_3 = \bar{\omega}_{210}^3\bar{\omega}_{033} - 3\bar{\omega}_{110}\bar{\omega}_{210}^2\bar{\omega}_{133} + 3\bar{\omega}_{110}^2\bar{\omega}_{233} - \bar{\omega}_{110}^3\bar{\omega}_{333}, \tag{6}$$

$$\bar{U}_4 = \bar{\omega}_{210}^4\bar{\omega}_{044} - 4\bar{\omega}_{110}\bar{\omega}_{210}^3\bar{\omega}_{144} + 6\bar{\omega}_{110}^2\bar{\omega}_{210}^2\bar{\omega}_{244} - 4\bar{\omega}_{110}^3\bar{\omega}_{210}\bar{\omega}_{344} + \bar{\omega}_{110}^4\bar{\omega}_{444}. \tag{7}$$

Finally, define $\alpha_j = \int_{-\infty}^{x/h} u^j K(u)\, du$ for $j = 0$, 1 and 2.

THEOREM 1. *Under conditions* (i)–(iii), *$l\{m(x)\}$ has an asymptotic χ_1^2 distribution if $nh^5 \to 0$, and this condition is also necessary if $m''(x) \neq 0$.*

Proof. Let $\mu^{(j)} = E(\bar{U}_j)$. As $\bar{U}_1$ is a smooth function of the sample means $(\bar{\omega}_{210}, \bar{\omega}_{011}, \bar{\omega}_{110}, \bar{\omega}_{111})$, $\bar{U}_1 - \mu^{(1)}$ is asymptotically normally distributed, by the Central Limit Theorem. As

$$\mu^{(1)} = \frac{1}{2}\{\alpha_2^2(x/h) - \alpha_1(x/h)\alpha_3(x/h)\}m''(x)h^2 + O(h^3)$$

and $\mathrm{var}\{\bar{U}_1\} = (nh)^{-1}\mu^{(2)} + o\{(nh)^{-1}\}$, then $\bar{U}_1 = O_p\{(nh)^{-1/2} + h^2\}$. Using condition (ii), we may show, along the lines of Owen (1990), that

$$\lambda\{m(x)\} = O_p\{(nh)^{-1/2} + h^2\}.$$

As $\sum U_i = \lambda\{m(x)\} \sum U_i^2 + O_p[nh\{(nh)^{-1/2} + h^2\}^2]$, we have

$$\lambda\{m(x)\} = (\bar{U}_2)^{-1}\bar{U}_1 + O_p[\{(nh)^{-1/2} + h^2\}^2].$$

By Taylor expansion, we obtain

$$\begin{aligned} l\{m(x)\} &= 2\lambda\{m(x)\} \sum U_i - \lambda^2\{m(x)\} \sum U_i^2 + O_p[nh\{(nh)^{-1/2} + h^2\}^3] \\ &= [Z + (nh)^{1/2}\{\mu^{(2)}\}^{-1/2}\mu^{(1)}]^2 + o_p[nh\{(nh)^{-1/2} + h^2\}^2], \end{aligned}$$

where $Z = (nh)^{1/2}(\bar{U}_1 - \mu^{(1)})\{\mu^{(2)}\}^{-1/2}$ is asymptotically $N(0, 1)$. Therefore, $l\{m(x)\}$ is asymptotically χ_1^2 if and only if $(nh)^{1/2}\{\mu^{(2)}\}^{-1/2}\mu^{(1)} \to 0$ which, in turn, implies that $nh^5 \to 0$ if $m''(x) \neq 0$. □

Based on the above theorem, an empirical likelihood confidence interval for $m(x)$ with nominal confidence level α is $I_{\alpha,\mathrm{EL}} = \{\theta : l(\theta) \leqslant c_\alpha\}$, where c_α satisfies $\mathrm{pr}(\chi_1^2 \leqslant c_\alpha) = \alpha$.

A conventional approach to constructing confidence intervals for $m(x)$, as considered in Chen and Qin's technical report, is based directly on the asymptotic normal distribution of $\hat{m}_l(x)$. Define

$$b(x/h) = \frac{\alpha_2^2(x/h) - \alpha_1(x/h)\alpha_3(x/h)}{\alpha_0(x/h)\alpha_2(x/h) - \alpha_1^2(x/h)},$$

$$v(x/h) = \frac{\int_{-1}^{x/h} \{\alpha_2(x/h) - u\alpha_1(x/h)\}^2 K^2(u)\, du}{\{\alpha_0(x/h)\alpha_2(x/h) - \alpha_1^2(x/h)\}^2}.$$

Standard results in local linear smoothing (Fan & Gijbels, 1996) show that

$$\left\{\frac{nh}{v(x/h)}\right\}^{1/2} \frac{\hat{m}_l(x) - m(x)}{\{V(x)/f(x)\}^{1/2}} \to N(B_n, 1) \tag{8}$$

in distribution, where $B_n = \frac{1}{2}b(x/h)\{v(x/h)V(x)f^{-1}(x)\}^{1/2}m''(x)(nh)^{1/2}h^2$. By letting $nh^5 \to 0$ such that $B_n \to 0$, we obtain a confidence interval with approximate nominal coverage α:

$$I_{\alpha,\mathrm{asy}} = \hat{m}_l(x) \pm z_\alpha \left\{\frac{v(x/h)\hat{V}(x)}{nh\hat{f}(x)}\right\}^{1/2}, \tag{9}$$

where z_α is the $(1+\alpha)/2$-quantile of $N(0,1)$ and $\hat{V}(x)$ and $\hat{f}(x)$ are consistent nonnegative estimators of $V(x)$ and $f(x)$. In their report, Chen and Qin use

$$\hat{V}(x)=\sum_{i=1}^{n}K_h(x-X_i)\{Y_i-\hat{m}_l(x)\}^2 \Big/ \sum_{i=1}^{n}K_h(x-X_i),$$

$$\hat{f}(x)=(nh)^{-1}\sum_{i=1}^{n}\frac{K_h(x-X_i)}{a_0(x/h)}.$$

Compared with $I_{\alpha,\mathrm{asy}}$, the empirical likelihood confidence interval is simpler as it does not require us to estimate $V(x)$ and $f(x)$ nor to remember to put in $v(x/h)$; the internal optimisation conducted by empirical likelihood carries these out automatically.

3. Coverage accuracy

We now assume three extra conditions:

(iv) $h=o\{n^{-1/5}\}$,

(v) $nh(\log n)^{-1}\to\infty$ and $E|Y|^{15}<\infty$,

(vi) $K(u)$ is continuous differentiable and $\{1, uK(u), u^2K(u)\}$ are linearly independent in $C[-1,1]$.

Condition (iv) implies undersmoothing as required by Theorem 1, and (v) and (vi) are needed to develop Edgeworth expansions for the coverage probability.

Using the standard Taylor expansion for the empirical likelihood ratio (Chen & Hall, 1993), we have

$$\begin{aligned} l(\theta)=nh\{&\bar{U}_2^{-1}\bar{U}_1^2+\tfrac{2}{3}\bar{U}_2^{-3}\bar{U}_3\bar{U}_1^3+(\bar{U}_2^{-5}\bar{U}_3^2-\tfrac{1}{2}\bar{U}_2^{-4}\bar{U}_4)\bar{U}_1^4\\ &+(8\bar{U}_2^{-6}\bar{U}_3\bar{U}_4-8\bar{U}_2^{-7}\bar{U}_3^3-\tfrac{8}{5}\bar{U}_2^{-5}\bar{U}_5)\bar{U}_1^5\}\\ &+nh\sum_{k=5}^{j}R_{1k}\bar{U}_1^{k+1}+O_p[nh\{(nh)^{-1/2}+h^2\}^{j+2}], \end{aligned} \tag{10}$$

where R_{1k} denotes $\bar{U}_2^{-(2k-1)}$ multiplied by a polynomial in $\bar{U}_2,\ldots,\bar{U}_{k+1}$ with constant coefficients. As in Chen & Hall (1993), we may write $l(\theta)=\{(nh)^{1/2}S_j'\}^2$, where

$$\begin{aligned} S_j'&=\bar{U}_2^{-1/2}\{\bar{U}_1+\tfrac{1}{3}\bar{U}_2^{-2}\bar{U}_3\bar{U}_1^2+(\tfrac{4}{9}\bar{U}_2^{-4}\bar{U}_3^2-\tfrac{1}{4}\bar{U}_2^{-3}\bar{U}_4)\bar{U}_1^3\}+\Delta\\ &=S+\Delta \end{aligned} \tag{11}$$

and $\Delta=O_p[\{(n/h)^{-1/2}+h^2\}^4]$. Define $\mu_0^{(j)}$ for $j=1,\ldots,4$, by replacing $\bar{\omega}_{ijk}$ by μ_{ijk} in (4) to (7). Then $\mu^{(j)}=E(\bar{U}_j)=\mu_0^{(j)}+O\{(nh)^{-1}\}$.

Let k_i be the ith cumulant of $(nh)^{1/2}S$. Derivations given in a second technical report by S. X. Chen and Y. S. Qin show that

$$\begin{aligned} k_1&=(nh)^{1/2}(\mu_0^{(2)})^{-1/2}\mu_0^{(1)}-\tfrac{1}{6}(\mu_0^{(2)})^{-3/2}\mu_0^{(3)}(nh)^{-1/2}+O\{(nh)^{1/2}h^4+(nh)^{-1/2}h^2+(nh)^{-3/2}\},\\ k_2&=1+\tfrac{1}{3}(\mu_0^{(2)})^{-2}\mu_0^{(3)}\mu_0^{(1)}+\{-\tfrac{13}{36}(\mu_0^{(2)})^{-3}(\mu_0^{(3)})^2+\tfrac{1}{2}(\mu_0^{(2)})^{-2}\mu_0^{(4)}\}(nh)^{-1}\\ &\quad+\gamma(\mu_0^{(2)})^{-1}h^2+O\{h^4+(nh)^{-1}h^2+(nh)^{-3/2}\},\\ k_3&=O\{(nh)^{-1/2}h^2+(nh)^{-3/2}\},\\ k_4&=O\{h^4+(nh)^{-1}h^2+(nh)^{-2}\},\\ k_l&=O\{(nh)^{-(l-2)/2}\}\quad(l\geqslant 5), \end{aligned} \tag{12}$$

where

$$\begin{aligned} \gamma=h^{-2}(&\mu_{111}^2\mu_{220}+\mu_{011}^2\mu_{420}+2\mu_{210}\mu_{011}\mu_{221}-2\mu_{210}\mu_{111}\mu_{121}-2\mu_{110}\mu_{011}\mu_{321}\\ &+2\mu_{110}\mu_{111}\mu_{221}-2\mu_{011}\mu_{111}\mu_{320}). \end{aligned}$$

As $\mu_{ij1} = O(h)$, whether x is in the interior or in the boundary, $\gamma = C + O(h)$, where C is a constant which depends on f, m, V_j and K.

By setting up the generating function of $(nh)^{1/2}S$ from the above cumulants and inverting it, we obtain the following Edgeworth expansion for the coverage probability of $I_{\alpha,\mathrm{EL}}$:

$$\begin{aligned}\mathrm{pr}\{m(x) \in I_{\alpha,\mathrm{EL}}\} \\ &= \alpha - c_\alpha^{1/2}\phi(c_\alpha^{1/2})[nh(\mu_0^{(2)})^{-1}(\mu_0^{(1)})^2 + \{-\tfrac{1}{3}(\mu_0^{(2)})^{-3}(\mu_0^{(3)})^2 + \tfrac{1}{2}(\mu_0^{(2)})^{-2}\mu_0^{(4)}\}(nh)^{-1} + \gamma(\mu_0^{(2)})^{-1}h^2] \\ &\quad + O\{nh^7 + h^4 + (nh)^{-1}h^2 + (nh)^{-3/2}\}, \qquad (13)\end{aligned}$$

where Φ and ϕ are the standard normal distribution and density functions, respectively.

Remark 1. The validity of the above Edgeworth expansion in (13) can be established using the arguments given in Chen and Qin's first technical report for $I_{\alpha,\mathrm{asy}}$.

Remark 2. The above Edgeworth expansion is valid for all $x \in [0, \infty)$ regardless of whether it is on the boundary or not. The coverage error is $O\{nh^5 + h^2 + (nh)^{-1}\}$ throughout the support of the curve. This is a substantial improvement over $I_{\alpha,\mathrm{asy}}$ whose coverage error, as shown again in Chen and Qin's first report, is $O\{nh^5 + h + (nh)^{-1}\}$ on the boundary and $O\{nh^5 + h^2 + (nh)^{-1}\}$ in the interior. The larger coverage error near the boundary for $I_{\alpha,\mathrm{asy}}$ occurs because the second-order term of the variance is $O(n^{-1})$ rather than $O(hn^{-1})$ as in the interior.

Remark 3. It then follows from (13) that

$$\mathrm{pr}\{m(x) \in I_{\alpha,\mathrm{EL}}\} = \alpha + b_1 nh^5 + b_2 h^2 + b_3(nh)^{-1} + O\{nh^7 + h^3 + (nh)^{-1}h^2 + (nh)^{-3/2}\}, \qquad (14)$$

where

$$\begin{aligned} b_1 &= -\tfrac{1}{4}c_\alpha^{1/2}\phi(c_\alpha^{1/2})(\mu_0^{(2)})^{-1}(\alpha_2^2 - \alpha_1\alpha_3)^2 f^4(x)\{m''(x)\}^2, \\ b_2 &= -c_\alpha^{1/2}\phi(c_\alpha^{1/2})(\mu_0^{(2)})^{-1}\gamma, \\ b_3 &= -c_\alpha^{1/2}\phi(c_\alpha^{1/2})\{-\tfrac{1}{3}(\mu_0^{(2)})^{-3}(\mu_0^{(3)})^2 + \tfrac{1}{2}(\mu_0^{(2)})^{-2}\mu_0^{(4)}\}. \end{aligned}$$

The optimal h that minimises the leading coverage error term is

$$h^* = \left\{\frac{-b_2 + (b_2^2 + 5b_1 b_3)^{1/2}}{5b_1}\right\}^{1/3} n^{-1/3},$$

provided $b_2^2 + 5b_1 b_3 > 0$. The optimal coverage error is $O(n^{-2/3})$.

Remark 4. The Edgeworth expansion for the coverage probability can be extended to confidence intervals based on a local quadratic estimator of $m(x)$ provided that m has a continuous fourth derivative and the rest of conditions (i)–(vi) are true. The coverage error for these intervals can be shown to be $O\{nh^9 + h^2 + (nh)^{-1}\}$. The main difference is that the contribution from the bias is nh^9, because of the local quadratic smoothing, which is a smaller order than nh^5. Provided $nh^7 \to 0$, the coverage is $O\{h^2 + (nh)^{-1}\}$.

We now show why the empirical likelihood confidence interval has a constant order of coverage error throughout the support of the curve. The leading term in the expansion (10) of $l(\theta)$ is

$$l_1\{m(x)\} = (nh)\bar{U}_1^2/\bar{U}_2.$$

It may be shown that $\mathrm{var}\{\bar{U}_1/(\bar{U}_2)^{1/2}\} = (nh)^{-1} + O\{h^4 + (nh)^{-1}h^2 + (nh)^{-2}\}$ for any x. As

$$l_1\{m(x)\} = \left[\frac{(nh)^{1/2}\{\hat{m}_l(x) - m(x)\}}{\bar{U}_2/(s_0 s_2 - s_1^2)^2}\right]^2,$$

empirical likelihood implicitly uses

$$(nh)^{-1}(s_0 s_2 - s_1^2)^{-1}\bar{U}_2 = \{(s_0 s_2 - s_1^2)\}^{-2}(nh)^{-2}\sum K_h^2(x - X_i)\left(s_{n,2} - \frac{x - X_i}{h}s_{n,1}\right)^2\{Y_i - m(x)\}^2$$

as an 'estimator' of the variance of $\hat{m}_l(x)$. It is not really an estimator as it involves $m(x)$. However, it will be if we replace $m(x)$ by $\hat{m}_l(x)$ and denote it by $\widehat{\text{var}}_{\text{EL}}\{\hat{m}_l(x)\}$.

A confidence interval which compromises between $I_{\alpha,\text{asy}}$ and $I_{\alpha,\text{EL}}$ is

$$I_{\alpha,\text{EL1}} = \hat{m}_l(x) \pm z_\alpha[\widehat{\text{var}}_{\text{EL}}\{\hat{m}_l(x)\}]^{1/2}.$$

It may be shown in a manner similar to that used in the study of $I_{\alpha,\text{EL}}$ that the coverage error of $I_{\alpha,\text{EL1}}$ is $O\{nh^5 + h^2 + (nh)^{-1}\}$ throughout.

Bartlett corrections have been established for empirical likelihood for almost all the situations considered so far; see for example DiCiccio, Hall & Romano (1992), Chen (1996) and Chen & Hall (1993). The fact that k_3 and k_4 are of smaller order than $nh^5 + h^2 + (nh)^{-1}$ establishes the foundation for $l(\theta)$ being Bartlett correctable. To appreciate this, note that, based on the results of the first two cumulants given in (12),

$$E\{l(\theta)\} = k_1^2 + k_2 = 1 + \beta + O\{h^4 + (nh)^{-1}h^2 + (nh)^{-3/2}\},$$

where

$$\beta = nh(\mu_0^{(2)})^{-1}(\mu_0^{(1)})^2 + \gamma h^2(\mu_0^{(2)})^{-1} + \{-\tfrac{1}{3}(\mu_0^{(2)})^{-3}(\mu_0^{(3)})^2 + \tfrac{1}{2}(\mu_0^{(2)})^{-2}\mu_0^{(4)}\}(nh)^{-1}.$$

According to (13),

$$\begin{aligned}\text{pr}\{l(\theta) \leqslant c_\alpha(1+\beta)\} &= \text{pr}\{\chi_1^2 \leqslant c_\alpha(1+\beta)\} - c_\alpha^{1/2}(1+\beta)^{1/2}\phi\{c_\alpha^{1/2}(1+\beta)^{1/2}\}\beta \\ &\quad + O\{nh^7 + h^4 + (nh)^{-1}h^2 + (nh)^{-3/2}\} \\ &= \alpha + O\{n^2h^{10} + nh^7 + h^4 + (nh)^{-1}h^2 + (nh)^{-3/2}\}. \qquad (15)\end{aligned}$$

The simple adjustment in the mean therefore leads to more accurate approximation of the distribution of $l(\theta)$ by the χ_1^2 distribution. However, to conduct the adjustment in practice, β has to be estimated. An estimator for β can be found by replacing $\mu_0^{(j)}$, μ_{ijk} and $m(x)$ by $\bar{U}_j$, $\bar{\omega}_{ijk}$ and $\hat{m}_l(x)$ respectively in the expression for β.

The Bartlett-corrected confidence interval for $m(x)$ is

$$I_{\alpha,\text{BCEL}} = \{\theta \mid l(\theta) \leqslant c_\alpha(1+\hat{\beta})\}.$$

Its coverage error will be slightly larger than the remainder term of (15) because β has to be estimated.

4. Empirical results

We report some simulation results designed to investigate the performance of the three empirical likelihood confidence intervals $I_{\alpha,\text{EL}}$, $I_{\alpha,\text{BCEL}}$ and $I_{\alpha,\text{EL1}}$. The asymptotic normal distribution based interval $I_{\alpha,\text{asy}}$ is also included for comparison.

The model considered in the simulation was given by

$$Y_i = 0{\cdot}3\exp\{-4{\cdot}0(4{\cdot}0X_i - 1{\cdot}0)^2\} + 0{\cdot}7\exp\{-16{\cdot}0(4{\cdot}0X_i - 3{\cdot}0)^2\} + \sigma(X_i)\varepsilon_i$$

for $i = 1, \ldots, n$, where X_i and ε_i were, respectively, Un[0, 1] and standard normal random variables, and $\sigma(x) = 0{\cdot}15\exp(-x)$. The sample size n ranges from 200 to 1000, and x was chosen to be 0 on the boundary and 0·5 in the interior. For each sample size considered and each sample simulated, 30 equally spaced levels of bandwidth h centred around $0{\cdot}6n^{-1/3}$ were used to construct the four intervals with nominal 95% level of confidence. This was designed to obtain information about the sensitivity of the coverages of the confidence intervals and their lengths compared to h.

The results of the simulation can be well represented by those for $n = 200$, 600 and 1000 shown in Fig. 1 for $x = 0$, which contains the average coverage errors and lengths of the confidence intervals based on 1000 replications. We also chose $x = 0{\cdot}5$ in the simulation as an example of an interior point. At $x = 0{\cdot}5$, the distinctions in coverage and length among $I_{\alpha,\text{asy}}$, $I_{\alpha,\text{EL}}$ and $I_{\alpha,\text{BCEL}}$ were small, and are therefore not presented in a figure here.

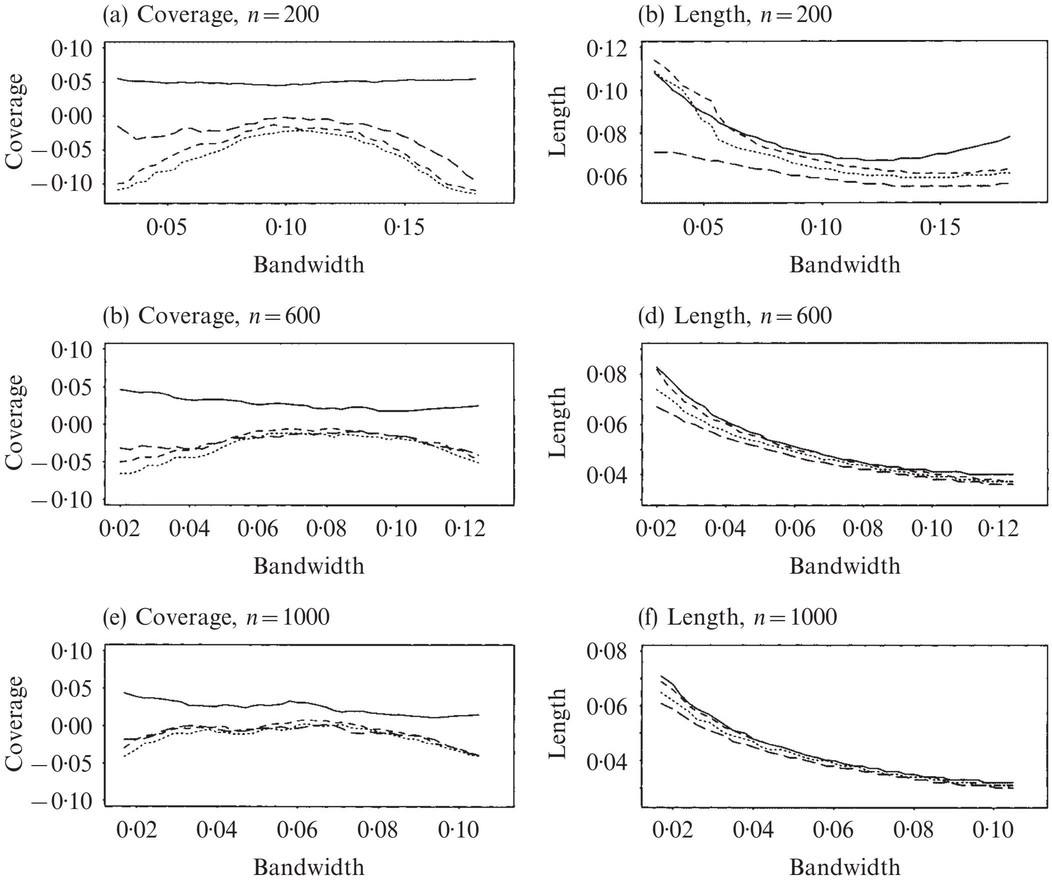

Fig. 1. Coverage and length of $I_{\alpha,\mathrm{asy}}$ (solid lines), $I_{\alpha,\mathrm{EL}}$ (dotted lines), $I_{\alpha,\mathrm{BCEL}}$ (short dashed lines) and $I_{\alpha,\mathrm{EL1}}$ (long dashed lines) at $x = 0$, plotted against bandwidth.

Figure 1 shows gaps in coverage error between $I_{\alpha,\mathrm{asy}}$ and the three likelihood based confidence intervals: $I_{\alpha,\mathrm{asy}}$ had excessive positive coverage error and $I_{\alpha,\mathrm{EL}}$ showed the greatest under-coverage. The interval $I_{\alpha,\mathrm{asy}}$ was the longest in most of the situations. The coverage of $I_{\alpha,\mathrm{EL}}$ was improved by the Bartlett correction which widened $I_{\alpha,\mathrm{EL}}$. It is interesting to see that the asymptotic normal interval using the variance prescribed by empirical likelihood had better coverage and was shorter than its parent $I_{\alpha,\mathrm{EL}}$. We also note that the coverage of $I_{\alpha,\mathrm{EL}}$ was improved by increasing the sample size, while the improvement in coverage of $I_{\alpha,\mathrm{asy}}$ seemed to be less so as n increased. This was probably a consequence of the theoretical result that $I_{\alpha,\mathrm{asy}}$ had a larger order of coverage error near the boundary. The coverage of the three empirical likelihood based intervals was more sensitive to h than $I_{\alpha,\mathrm{asy}}$; however, this sensitivity decreased as n increased.

Acknowledgement

The research was carried out while S. X. Chen was employed by La Trobe University. The authors would like to thank a referee, the associate editor and the editor for comments that led to a better presentation of the paper.

References

Chen, S. X. (1996). Empirical likelihood for nonparametric density estimation. *Biometrika* **83**, 329–41.
Chen, S. X. & Hall, P. (1993). Smoothed empirical likelihood confidence intervals for quantiles. *Ann. Statist.* **21**, 1166–81.

DiCiccio, T. J., Hall, P. & Romano, J. P. (1991). Bartlett adjustment for empirical likelihood. *Ann. Statist.* **19**, 1053–61.

Fan, J. (1993). Local linear regression smoothers and their minimax efficiencies. *Ann. Statist.* **21**, 196–216.

Fan, J. & Gijbels, I. (1996). *Local Polynomial Modelling and its Applications*. London: Chapman and Hall.

Hall, P. (1991). Edgeworth expansions for nonparametric density estimators, with applications. *Statistics* **22**, 215–32.

Hall, P. (1992). On bootstrap confidence intervals in nonparametric regression. *Ann. Statist.* **20**, 695–711.

Owen, A. (1988). Empirical likelihood ratio confidence intervals for a single functional. *Biometrika* **75**, 237–49.

Owen, A. (1990). Empirical likelihood ratio confidence regions. *Ann. Statist.* **18**, 90–120.

Neumann, M. H. (1995). Automatic bandwidth choice and confidence intervals in nonparametric regression. *Ann. Statist.* **23**, 1937–59.

Ruppert, D. & Wand, M. P. (1994). Multivariate locally weighted least squares regression. *Ann. Statist.* **22**, 1346–70.

[*Received October* 1999. *Revised June* 2000]