应用概率统计　第二十三卷
第一期　2007 年 2 月

Chinese Journal of Applied Probability
and Statistics Vol. 23 No. 1 Feb. 2007

Coverage Accuracy of Confidence Intervals for a Conditional Probability Density Function

LEI QINGZHU　　QIN YONGSONG

(*School of Mathematical Sciences, Guangxi Normal University, Guilin, 541004*)

Abstract

Point-wise confidence intervals for a conditional probability density function are considered. The confidence intervals are based on the empirical likelihood. Their coverage accuracy is assessed by developing Edgeworth expansions for the coverage probabilities. It is shown that the empirical likelihood confidence intervals are Bartlett correctable.

Keywords: Confidence interval, empirical likelihood, double kernel estimator, Bartlett correction.

AMS Subject Classification: Primary 62E20.

§ 1. Introduction

Let $(X_1, Y_1), \cdots, (X_n, Y_n)$ be an i.i.d. sample from a population $(X, Y) \in R^p \times R^q$, and $\theta = f(y|x)$ be the conditional probability density function of Y given $X = x \in R^p$. The double kernel estimator[1,2] for $f(y|x)$ at any given $(x, y) \in R^p \times R^q$ is

$$\widehat{f}(y|x) = \frac{a^{-p}b^{-q}\sum K_1\Big(\dfrac{x-X_i}{a}\Big)K_2\Big(\dfrac{y-Y_i}{b}\Big)}{a^{-q}\sum K_1\Big(\dfrac{x-X_i}{a}\Big)}, \tag{1.1}$$

where K_i, $i = 1, 2$ are kernel functions, and $a = a_n$ and b_n are smoothing bandwidths. Strong consistency and asymptotic normality of the estimator are studied in [1, 2].

This paper is concerned with the construction of point-wise empirical likelihood confidence intervals for $f(y|x)$ at any fixed $(x, y) \in R^p \times R^q$ in conjunction with the double kernel estimator. It is aimed at studying coverage accuracy of the confidence intervals based on the empirical likelihood by developing Edgeworth expansions for the coverage probabilities. It is shown that the empirical likelihood confidence interval is Bartlett correctable. Moreover, the empirical likelihood automatically studentizes so that there is no need to estimate variances in the limiting distributions. These results lead to a conclusion that, in the sense of coverage accuracy, empirical likelihood approach suggested in the present paper is particularly competitive with the normal approximation methods implied in [2].

A conditional density provides the most informative summary of the relationship between independent and dependent variables. The conditional density function plays a pivotal role in

Received 2004. 8. 20. Revised 2005. 9. 2.

financial econometrics (see [14]). The estimates and their asymptotic properties of the nonparametric kernel estimates are investigated in [2] and [14], among others. Often confidence intervals for $f(y|x)$ are required, for instance, to test a hypothesis about $f(y|x)$.

The paper is structured as follows. Section 2 describes the concept of empirical likelihood for conditional densities, and presents results about Wilks' theorem which can be used to construct the empirical likelihood confidence intervals. The coverage errors of the empirical likelihood confidence intervals are given in Section 3 by developing Edgeworth expansions, and the results for Bartlett correction are obtained in Section 4. Simulation results are reported in Section 5. Some proofs are given in the Appendix.

§2. Empirical Likelihood Confidence Intervals for $f(y|x)$

Use $g(\cdot,\cdot)$ and $h(\cdot)$ to denote the joint probability density function of (X,Y) and the probability density function of X, respectively. We first introduce some notations and assumptions.

We assume the following regularity conditions:

(i) g and h have continuous partial derivatives up to the r-th order in a neighborhood of (x,y) and x respectively, and $g(x,y)>0,\ h(x)>0$;

(ii) For $i=1,2$, K_i are compactly supported kernels which satisfy

$$\int u_1^{j_1}\cdots u_p^{j_p}K_1(u)\mathrm{d}u=\begin{cases}1, & \sum\limits_{i=1}^{p}j_i=0;\\ 0, & 1\le\sum\limits_{i=1}^{p}j_i\le r-1,\end{cases}$$

$$\int v_1^{j_1}\cdots v_q^{j_q}K_2(v)\mathrm{d}v=\begin{cases}1, & \sum\limits_{i=1}^{q}j_i=0;\\ 0, & 1\le\sum\limits_{i=1}^{q}j_i\le r-1,\end{cases}$$

where $u=(u_1,\cdots,u_p)^\tau$, $v=(v_1,\cdots,v_q)^\tau$, j_s $(s=1,2,\cdots,\max\{p,q\})$ are non-negative integers;

(iii) $a\to0,\ b\to0$ and $na^pb^{3q}\to\infty$ as $n\to\infty$;

(iv) $na^{p+2r}b^q\to0,\ na^pb^{q+2r}\to0$ as $n\to\infty$.

Remark 2.1 We will use Conditions (i) to (iv) to prove a Wilks Theorem for the empirical likelihood ratio statistic (defined below). Let us compare these conditions with those employed in [2] to obtain the asymptotic distribution of the double kernel estimator. In [2], only the situation that $r=2,\ p=q=1$ is considered. In this situation, it can be seen that our conditions are weaker than those in [2] in some degree.

Define $\omega_i(\theta)=K_1((x-X_i)/a)\{b^{-q}K_2((y-Y_i)/b)-\theta\}$ with $\theta=f(y|x)$ and, for positive integers j,

$$\varpi_j=(na^p)^{-1}\sum_{i=1}^{n}\omega_i^j(\theta),\qquad \mu_j=\mathsf{E}(\varpi_j).$$

It can be shown, by Taylor expansion, Conditions (i) and (ii), $a\to0$ and $b\to0$, that

$$\mu_1=O(a^r+b^r),\qquad \mu_2=b^{-q}\Big\{f(x,y)\int K_1^2(u)K_2^2(v)\mathrm{d}u\mathrm{d}v+o(1)\Big\}.\tag{2.1}$$

By Central Limit Theorem,

$$\sqrt{na^p}\mu_2^{-1/2}(\varpi_1-\mu_1)\to_d N(0,1), \qquad \sqrt{na^pb^{3q}}(\varpi_2-\mu_2)=O_p(1). \tag{2.2}$$

In this paper, the empirical likelihood method is used to construct the confidence intervals for θ. The empirical likelihood, introduced by Owen[3,4], is a computer intensive statistical method like the bootstrap. However, instead of using an equal probability weight n^{-1} for all data values, the empirical likelihood chooses the weights, say p_i on the i-th data points (X_i, Y_i), by profiling a multinomial likelihood under a set of constraints. The constraints reflect the characteristics of the quantity of interests. A review of the empirical likelihood is given in Hall and Scala[5].

Let $p_1,\cdots,p_n$ be nonnegative numbers adding to unity. The empirical likelihood at θ, is defined as

$$L(\theta)=\sup_{\sum\limits_{i=1}^{n}p_i\omega_i(\theta)=0}\prod_{i=1}^{n}p_i.$$

After using a Lagrange multiplier to find the optimal p_i, the log empirical likelihood ratio is

$$\ell(\theta)=-2\log\{L(\theta)n^n\}=2\sum_{i=1}^{n}\log\{1+\lambda(\theta)\omega_i(\theta)\},$$

where $\lambda(\theta)$ satisfies

$$\sum_{i=1}^{n}\omega_i(\theta)\{1+\lambda(\theta)\omega_i(\theta)\}^{-1}=0. \tag{2.3}$$

An empirical likelihood confidence interval with nominal coverage of α, denoted as I_α, is

$$I_{\alpha,el}=\{\theta|\ell(\theta)\le c_\alpha\}, \tag{2.4}$$

where c_α is the αth quantile of the χ_1^2 distribution. A special feature of the empirical likelihood confidence interval is that no explicit variance estimator, like the one in [2], is required in its construction as the studentizing is carried out internally via the optimization procedure.

To evaluate the coverage of $I_{\alpha,el}$, we notice from (2.3) that

$$\sum_{i=1}^{n}\omega_i(\theta)-\lambda(\theta)\sum_{i=1}^{n}\frac{\omega_i^2(\theta)}{1+\lambda(\theta)\omega_i(\theta)}=0. \tag{2.5}$$

Put $Z_n=\max_i|\omega_i(\theta)|=O(b^{-q})$, a.s.. Then

$$\frac{|\lambda(\theta)|}{1+|\lambda(\theta)|Z_n}\varpi_2=\varpi_1.$$

It is clear that $Z_n=O(b^{-q})$ almost surely, $\varpi_1=O_p\{(na^pb^q)^{-1/2}+a^r+b^r\}$ and $\varpi_2=\mu_2+O_p\{(na^pb^{3q})^{-1/2}\}$ from (2.1) and (2.2). Then combining with Conditions (iii) and (iv)

$$\lambda(\theta)=O_p[b^q\{(na^pb^q)^{-1/2}+a^r+b^r\}].$$

From (2.5), we have $\lambda(\theta)=(\varpi_2)^{-1}\varpi_1+O_p[\{(na^pb^q)^{-1/2}+a^r+b^r\}^2]$. Thus,

$$\begin{aligned}\ell(\theta)&=na^p\mu_2^{-1}\varpi_1^2+O_p[na^p\{(na^pb^q)^{-1/2}+a^r+b^r\}^3]\\&=\{(na^p)^{1/2}(\varpi_1-\mu_1)\mu_2^{-1/2}+(na^p)^{1/2}\mu_2^{-1/2}\mu_1\}^2+O_p[na^p\{(na^pb^q)^{-1/2}+a^r+b^r\}^3].\end{aligned}$$

Now Conditions (iii), (iv) and (2.2) imply that $\ell(\theta)$ is asymptotically χ_1^2.

In summary, we have the following Wilks' Theorem for the empirical likelihood statistic:

Theorem 2.1 Under Conditions (i) to (iv),

$$\ell(\theta) \to_d \chi_1^2.$$

From this result, we can see that the empirical likelihood confidence interval $I_{\alpha,el}$ has asymptotic coverage probability α.

§ 3. Coverage Accuracy

In studying the coverage accuracy of the confidence intervals, we assume an extra condition

(v) $na^p/\log n \to \infty$, $nb^t \to 0$ for some $t > 0$.

The derivation deferred until the Appendix shows that the coverage probability of $I_{\alpha,el}$ admits the following Edgeworth expansion:

Theorem 3.1 Under Conditions (i) to (v),

$$\begin{aligned}\mathsf{P}\{\theta \in I_{\alpha,el}\} &= \alpha - \Big\{na^p\mu_1^2\mu_2^{-1} + \Big(\frac{1}{2}\mu_2^{-2}\mu_4 - \frac{1}{3}\mu_2^{-3}\mu_3^2\Big)(na^p)^{-1}\Big\}c_\alpha\phi(c_\alpha)\\ &\quad +O\{(na^p)(a^{3r}+b^{3r}) + (na^pb^q)^{-2}\},\end{aligned} \tag{3.1}$$

where $\phi(\cdot)$ is the the probability density function of the standard normal random variable.

Now, we investigate the optimal bandwidth choice. Put

$$\kappa_{l1} = \int_{R^p} u_l^r K_1(u_1,\cdots,u_p)\mathrm{d}u, \quad 1\le l\le p, \qquad \kappa_{s2} = \int_{R^q} v_s^r K_2(v_1,\cdots,v_q)\mathrm{d}v, \quad 1\le s\le q.$$

It can be shown that

$$\begin{aligned}\mu_1 &= (-1)^r(r!)^{-1}\Big\{a^r\sum_{l=1}^p\kappa_{l1}\frac{\partial}{\partial u_l} + b^r\sum_{s=1}^q\kappa_{s2}\frac{\partial}{\partial v_s}\Big\}g(u,v)|_{u=x,v=y}\\ &\quad -(-1)^r(r!)^{-1}a^r\sum_{l=1}^p\kappa_{l1}\frac{\partial}{\partial u_l}h(u)|_{u=x} + o(a^r+b^r)\\ &\widehat{=} \xi_1a^r + \xi_2b^r + o(a^r+b^r),\\ \mu_2 &= b^{-q}\Big\{g(x,y)\int K_1^2(u)K_2^2(v)\mathrm{d}u\mathrm{d}v + o(1)\Big\} \widehat{=} \xi_3b^{-q} + o(b^{-q}),\\ \mu_3 &= b^{-2q}\Big\{g(x,y)\int K_1^3(u)K_2^3(v)\mathrm{d}u\mathrm{d}v + o(1)\Big\} \widehat{=} \xi_4b^{-2q} + o(b^{-2q}),\\ \mu_4 &= b^{-3q}\Big\{g(x,y)\int K_1^4(u)K_2^4(v)\mathrm{d}u\mathrm{d}v + o(1)\Big\} \widehat{=} \xi_5b^{-3q} + o(b^{-3q}).\end{aligned}$$

Therefore, the dominant coverage error term in (3.1) becomes

$$\Big\{na^pb^q(\xi_1a^r+\xi_2b^r)^2\xi_3^{-1} + \Big(\frac{1}{2}\xi_3^{-2}\xi_5 - \frac{1}{3}\xi_3^{-3}\xi_4^2\Big)(na^qb^q)^{-1}\Big\}c_\alpha\phi(c_\alpha). \tag{3.2}$$

If we take $a = b = h$, then the dominant coverage error term is

$$\Big\{(\xi_1+\xi_2)^2\xi_3^{-1}nh^{p+q+2r} + \Big(\frac{1}{2}\xi_3^{-2}\xi_5 - \frac{1}{3}\xi_3^{-3}\xi_4^2\Big)(nh^{p+q})^{-1}\Big\}c_\alpha\phi(c_\alpha). \tag{3.3}$$

In this case, The optimal h that minimizes (3.3) is

$$h^{\star}_{el} = \left\{ \frac{(p+q)\Big(\dfrac{1}{2}\xi_3^{-1}\xi_5 - \dfrac{1}{3}\xi_3^{-2}\xi_4^2\Big)}{(p+q+2r)(\xi_1+\xi_2)^2} \right\}^{1/[2(p+q+r)]} n^{-1/(p+q+r)}. \tag{3.4}$$

Choosing $a = b = h = O(n^{-1/(p+q+r)})$, then the optimal coverage probability of the empirical likelihood confidence interval $I_{\alpha,el}$ is $\mathsf{P}\{\theta \in I_{\alpha,el}\} = \alpha - O(n^{-r/(p+q+r)})$.

§ 4. The Bartlett Correction

The results in the last section show that the optimal coverage error of $I_{\alpha,el}$ is at the order of $O(n^{-r/(p+q+r)})$. What we are going to show in this section is that the coverage error of the empirical likelihood confidence interval can be reduced by Bartlett correction.

The Bartlett correction is a novel and elegant property of classical parametric likelihood. A simple adjustment in the mean of the likelihood ratio statistic will improve the coverage accuracy of the likelihood ratio based confidence intervals by one order of magnitude. It has been shown by DiCiccio, Hall and Romano[6], Chen[7,8,9] and Chen and Hall[10] that the empirical likelihood possesses the Bartlett property for wide range of situations. Thus far the only known case where the empirical likelihood does not admit the property is that found by Jing and Wood[11] by restricting the distributions within the exponential family.

We will show that in the current situation the empirical likelihood admits the Bartlett property. It may be shown that

$$\mathsf{E}\{\ell(\theta_0)\} = 1 + (na^p)^{-1}\beta + o\{(na^p)(a^{2r}+b^{2r}) + (na^p b^q)^{-1}\},$$

where θ_0 is the true value of the parameter and

$$\beta = \mu_2^{-1}(na^p\mu_1)^2 + \frac{1}{2}\mu_2^{-2}\mu_4 - \frac{1}{3}\mu_2^{-3}\mu_3^2. \tag{4.1}$$

Notice that β appears in the leading coverage error term in (3.1). Based on (3.1) and choosing $a = b = h = O(n^{-1/(p+q+r)})$, we have

$$\begin{aligned}
& \mathsf{P}[\ell(\theta) \le c_\alpha\{1+\beta(nh^p)^{-1}\}] \\
= {} & \mathsf{P}[\chi_1^2 \le c_\alpha\{1+\beta(nh^p)^{-1}\}] \\
& -(nh^p)^{-1}\beta c_\alpha^{1/2}\{1+\beta(nh^p)^{-1}\}^{1/2}\phi[c_\alpha^{-1/2}\{1+\beta(nh^p)^{-1}\}^{1/2}] + O\{(nh^p)h^{3r} + (nh^{p+q})^{-2}\} \\
= {} & \mathsf{P}(\chi_1^2 \le c_\alpha) + O\{(nh^p)h^{3r} + (nh^{p+q})^{-2}\} = \alpha + O(n^{-2r/(p+q+r)}).
\end{aligned} \tag{4.2}$$

Therefore, the empirical likelihood is Bartlett correctable in the current case.

Let $I_{\alpha,bcel} = [\theta | \ell(\theta) \le c_\alpha\{1+\beta(nh^p)^{-1}\}]$ be the Bartlett corrected empirical likelihood confidence interval. From (4.2), we see that $I_{\alpha,bcel}$ has coverage errors of $n^{-2r/(p+q+r)}$ if h is chosen to be $O(n^{-1/(p+q+r)})$. In practice, the Bartlett factor β has to be estimated which in turn requires estimation of μ_j for $j = 1, 2, 3$ and 4. Estimators for μ_j, $j \ge 2$, can be defined as

$$\widehat{\mu}_j = (na^p)^{-1} \sum_{i=1}^n K_1^j\Big(\frac{x - X_i}{a}\Big)\Big\{b^{-q}K_2\Big(\frac{y-Y_i}{b}\Big) - \widehat{f}(y|x)\Big\}^j.$$

§5. Simulations

We conducted a small simulation study on the finite sample performance of normal approximation and empirical likelihood based confidence intervals on $f(y|x)$, with (X_i, Y_i)'s generated by the process

$$X_i \sim N(0,1), \qquad Y_i \Big| X_i \sim N\Big(\beta_1 X_i, \frac{1+\beta_2 X_i^2}{1+\beta_2}\Big).$$

We generated $1,000$ random samples of data $\{X_i, Y_i, i=1,\cdots,n\}$ for $n=60,\ 120$ and 180 from the above model. For nominal confidence level $1-\alpha=0.95$, using the simulated samples, we evaluated the coverage probability (CP) and the average length of the interval (AL) of the normal approximation based (NA), empirical likelihood based (EL) and Bartlett correction based (BC) intervals.

Throughout we use two fourth-order kernel (i.e. $r=4$),

$$K_1(u) = K_2(u) = \begin{cases} \dfrac{105}{64}(1-3u^2)(1-u^2)^2, & \text{if } |u| \le 1, \\ 0, & \text{otherwise.} \end{cases}$$

We take $h = c \times n^{-1/6}$ for $c=0.5,\ 1$ and 1.5. Table 1 reports the simulation results for $f(y|x)$ at $(x,y)=(0,0)$. It is seen that the coverage probability (CP) for NA, EL and BC increases with n, approaching the nominal level 0.95; The performance of EL and BC confidence interval's is significantly better than that of NA.

Table 1 Confidence interval coverage probability (CP) and average length (AL) for $f(y|x)$ at $(x,y)=(0,0)$

(β_1,β_2)	n	CI	CP ($c=0.5,\ 1,\ 1.5$)	AL ($c=0.5,\ 1,\ 1.5$)
$(\beta_1,\beta_2)=(1,1)$	60	NA	0.885, 0.905, 0.901	2.26, 2.31, 2.27
		EL	0.906, 0.911, 0.914	2.27, 2.28, 2.10
		BC	0.923, 0.917, 0.918	1.95, 1.89, 2.00
	120	NA	0.891, 0.912, 0.915	2.21, 2.27, 2.31
		EL	0.927, 0.934, 0.938	1.87, 1.90, 1.95
		BC	0.930, 0.937, 0.928	1.65, 1.69, 1.50
	180	NA	0.931, 0.941, 0.935	2.15, 1.98, 1.83
		EL	0.949, 0.938, 0.944	1.90, 1.87, 1.56
		BC	0.946, 0.941, 0.954	1.05, 1.23, 2.10
$(\beta_1,\beta_2)=(1,2)$	60	NA	0.905, 0.886, 0.890	2.13, 2.01, 2.11
		EL	0.931, 0.956, 0.934	2.12, 1.87, 1.32
		BC	0.930, 0.948, 0.933	1.93, 2.10, 1.75
	120	NA	0.935, 0.942, 0.950	1.31, 2.32, 2.28
		EL	0.947, 0.940, 0.951	1.16, 0.91, 1.36
		BC	0.944, 0.948, 0.961	1.01, 0.86, 0.95
	180	NA	0.946, 0.945, 0.952	2.17, 2.20, 1.57
		EL	0.949, 0.953, 0.948	1.21, 1.53, 1.12
		BC	0.948, 0.947, 0.934	0.85, 0.73, 0.64

§ 6. Appendix: Derivations

Derivation of (3.1) Note that

$$\lambda(\theta)=O_p[b^q\{(na^pb^q)^{-1/2}+a^r+b^r\}].$$

Similar to (A1.3) in Chen and Hall[10], we obtain the following Taylor expansion of $\ell(\theta)$.

$$\begin{aligned}\ell(\theta) \;=\;& (na^p)\Big\{\overline{\omega}_2^{-1}\overline{\omega}_1^2+\frac{2}{3}\overline{\omega}_2^{-3}\overline{\omega}_3\overline{\omega}_1^3+\Big(\overline{\omega}_2^{-5}\overline{\omega}_3^2-\frac{1}{2}\overline{\omega}_2^{-4}\overline{\omega}_4\Big)\overline{\omega}_1^4\\ &+\Big(8\overline{\omega}_2^{-6}\overline{\omega}_3\overline{\omega}_4-8\overline{\omega}_2^{-7}\overline{\omega}_3^3-\frac{8}{5}\overline{\omega}_2^{-5}\overline{\omega}_5\Big)\overline{\omega}_1^5\Big\}\\ &+na^p\sum_{k=5}^{j}R_{1k}\overline{\omega}_1^{k+1}+O_p[na^p\{(na^pb^q)^{-1/2}+a^r+b^r\}^{j+1}],\end{aligned}$$

where R_{1k} denotes $\overline{\omega}_2^{-(2k-1)}$ multiplied by a polynomial in $\overline{\omega}_2,\cdots,\overline{\omega}_{k+1}$, with constant coefficients.

As in Chen and Hall[10], we may write

$$\ell(\theta)=\{(na^p)^{1/2}S_j'\}^2,$$

where

$$\begin{aligned}S_j' \;=\;& \overline{\omega}_2^{-1/2}\Big\{\overline{\omega}_1+\frac{1}{3}\overline{\omega}_2^{-2}\overline{\omega}_3\overline{\omega}_1^2+\Big(\frac{4}{9}\overline{\omega}_2^{-4}\overline{\omega}_3^2-\frac{1}{4}\overline{\omega}_2^{-3}\overline{\omega}_4\Big)\overline{\omega}_1^3\\ &+\Big(-\frac{112}{27}\overline{\omega}_2^{-6}\overline{\omega}_3^3+\frac{49}{12}\overline{\omega}_2^{-5}\overline{\omega}_3\overline{\omega}_4-\frac{4}{5}\overline{\omega}_2^{-4}\overline{\omega}_5\Big)\overline{\omega}_1^4+\sum_{k=5}^{j}T_k\overline{\omega}_1^k\Big\}+U_j\\ =\;& S_j+U_j,\end{aligned}$$

where $U_j=O_p[\{(na^pb^q)^{-1/2}+a^r+b^r\}^{j+1}]$, and T_k denotes $\overline{\omega}_2^{-2(k-1)}$ multiplied by a polynomial in $\overline{\omega}_2,\cdots,\overline{\omega}_k$ with constant coefficients. Noting that $na^t\to 0$, $nb^t\to 0$ for some $t>0$, a little additional analysis shows that by choosing j sufficiently large we may ensure that,

$$\mathsf{P}\{|U_j|>(na^p)^{-5/2}\}=O\{(na^p)^{-2}\}. \tag{6.1}$$

Observe that S_j is a function of $\overline{\omega}_1,\cdots,\overline{\omega}_j$. Denote that function by s_j. Put $\mu_k=\mathsf{E}(\overline{\omega}_k)$, $\mu=(\mu_1,\cdots,\mu_j)^T$, $u=(u_1,\cdots,u_j)^T$, $V_k=\overline{\omega}_k-\mu_k$, $V=(V_1,\cdots,V_j)^T$,

$$d_{k_1,\cdots,k_m}=\Big(\prod_{l=1}^{m}\frac{\partial}{\partial u_{k_l}}\Big)s_j(u_1,\cdots,u_j)|_{u=\mu},$$

$$p(u)=s_j(\mu)+\sum_{m=1}^{6}(m!)^{-1}\sum_{k_1,\cdots,k_m\in\{1,\cdots,j\}}d_{k_1,\cdots,k_m}u_{k_1}\cdots u_{k_m}.$$

Let $k_1,k_2,\cdots$ be the cumulants of $(nh^d)^{1/2}p(V)$, and $k^{i_1,\cdots,i_p}$ denote the p-th order multivariate cumulants of $V=(V_1,\cdots,V_j)^T$. Our next step is to calculate $k_1,k_2,\cdots$.

It is clear, from (2.1), that

$$d_1=\mu_2^{-1/2}+\frac{2}{3}\mu_2^{-5/2}\mu_3\mu_1+O(a^{2r}+b^{2r}),\qquad d_2=-\frac{1}{2}\mu_2^{-3/2}\mu_1+O(a^{2r}+b^{2r}),$$
$$d_l=O(a^{2r}+b^{2r}),\qquad \text{for }\ l\ge 3,$$

$$d_{11} = \frac{2}{3}\mu_2^{-5/2}\mu_3 + O(a^{2r}+b^{2r}), \qquad d_{12} = -\frac{1}{2}\mu_2^{-3/2} + O(a^r+b^r),$$
$$d_{lm} = O(a^r+b^r), \qquad \text{for all other second derivatives},$$
$$d_{111} = \frac{8}{3}\mu_2^{-9/2}\mu_3^2 - \frac{3}{2}\mu_2^{-7/2}\mu_4 + O(a^r+b^r), \qquad d_{112} = -\frac{5}{3}\mu_2^{-7/2}\mu_3 + O(a^r+b^r),$$
$$d_{113} = \frac{2}{3}\mu_2^{-5/2} + O(a^r+b^r), \qquad d_{122} = \frac{3}{4}\mu_2^{-5/2} + O(a^r+b^r),$$
$$d_{ijk} = O(a^r+b^r), \qquad \text{for all other third derivatives}.$$

Noting that $\mu_j = a_1^{-p}\mathsf{E}\omega_1^j(\theta)$, we can obtain that

$$k^{11} = (na^p)^{-1}(\mu_2 - a^p\mu_1^2), \qquad k^{12} = (na^p)^{-1}(\mu_3 - a^p\mu_1\mu_2),$$
$$k^{13} = (na^p)^{-1}(\mu_4 - a^p\mu_1\mu_3), \qquad k^{22} = (na^p)^{-1}(\mu_4 - a^p\mu_2^2),$$
$$k^{111} = (na^p)^{-2}(\mu_3 - 3a^p\mu_1\mu_2 + 2a^{2p}\mu_1^3),$$
$$k^{112} = (na^p)^{-2}(\mu_4 - 2a^p\mu_1\mu_3 + 2a^{2p}\mu_1^2\mu_2 - a^p\mu_2^2),$$
$$k^{122} = (na^p)^{-2}(\mu_5 - 2a^p\mu_2\mu_3 - a^p\mu_1\mu_4 + 2a^{2p}\mu_1\mu_2^2),$$
$$k^{1111} = (na^p)^{-3}(\mu_4 - 3a^p\mu_2^2) + O\{(na^p)^{-3}a^p\mu_1\},$$
$$k^{1112} = (na^p)^{-3}(\mu_5 - 4a^p\mu_2\mu_3) + O\{(na^p)^{-3}a^p\mu_1\}.$$

According to the results given by James and Mayne[12], it follows that

$$k_1 = n^{1/2}s_j(\mu)a^{p/2} - \frac{1}{6}\mu_2^{-3/2}\mu_3 a_1^{-p/2}n^{-1/2} + O\{(na^p)^{-1/2}(a^r+b^r) + (na^pb^q)^{-3/2}\},$$
$$k_2 = \sigma^2 + \Big(\frac{1}{2}\mu_2^{-2}\mu_4 - \frac{13}{36}\mu_2^{-3}\mu_3^2\Big)(na^p)^{-1} + O\{(na^p)^{-1}(a^r+b^r) + (na^pb^q)^{-2}\},$$

where

$$\sigma^2 = 1 + \frac{1}{3}\mu_2^{-2}\mu_3\mu_1 + O(a^{2r}+b^{2r}),$$
$$k_3 = O\{(na^p)^{-1/2}(a^r+b^r) + (na^p)^{-5/2}\}, \qquad k_4 = O\{(na^p)^{-1}(a^r+b^r) + (na^pb^q)^{-2}\},$$
$$k_l = O\{(na^p)^{-(l-2)/2}\} \qquad \text{for } l \ge 5.$$

Thus we could develop a formal Edgeworth expansion for the distribution of $(na^p)^{1/2}p(V)$:

$$\begin{aligned} &\mathsf{P}\{n^{1/2}a^{p/2}p(V) \le t\} \\ = \;& \Phi(t) - \frac{1}{12}(na^p)^{-1}\{6\mu_2^{-1}(na^p\mu_1)^2 + 3\mu_2^{-2}\mu_4 - 2\mu_2^{-3}\mu_3^2\}t\phi(t) \\ &+ (\text{even polynomial in } t)\phi(t) + O\{(na^p)(a^{3r}+b^{3r}) + (na^pb^q)^{-2}\}, \end{aligned}$$

which, in turn as in [10], gives the Edgeworth expansion for $\ell(\theta)$ in (3.1).

References

[1] Zhao, L.C. and Liu, Z.J., Strong consistency of the kernel estimators of a conditional density function, *Acta. Math. Sinica (New Ser.)*, **4**(1985), 314–318.

[2] Xue, L.G. and Guo, L.Z., Asymptotic distributions of dpuble kernel estimators for conditional densities (in Chinese), *J. Engineering Math.*, **6**(1989), 59–65.

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

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

[5] Hall, P. and La Scala, B., Methodology and algorithms of empirical likelihood, *Internat. Statist. Rev.*, **58**(1990), 109–127.

[6] DiCiccio, T.J., Hall, P. and Romano, J.P., Bartlett adjustment for empirical likelihood, *Ann. Statist.*, **19**(1991), 1053–1061.

[7] Chen, S.X., On the coverage accuracy of empirical likelihood confidence regions for linear regression model, *Ann. Inst. Statist. Math.*, **45**(1993), 621–637.

[8] Chen, S.X., Empirical likelihood confidence intervals for linear regression coefficients, *J. Mult. Anal.*, **49**(1994), 24–40.

[9] Chen, S.X., Empirical likelihood for nonparametric density estimation, *Biometrika*, **83**(1996), 329–341.

[10] Chen, S.X. and Hall, P., Smoothed empirical likelihood confidence intervals for quantiles, *Ann. Statist.*, **21**(1993), 1166–1181.

[11] Jing, B-Y. and Wood, A.T.A., Exponential empirical likelihood is not Bartlett correctable, *Ann. Statist.*, **24**(1996), 365–369.

[12] James, G.S. and Mayne, A.J., Cumulants of functions of random variables, *Sankhya, Ser. A*, **24**(1962), 47–54.

[13] Hall, P., *The bootstrap and Edgeworth Expansions*, New York: Springer-Verlag, 1992.

[14] Fan, J. and Yim, T.H., A crossvalidation method for estimating conditional densities, *Biometrika*, **91**(2004), 819–834.

条件密度置信区间的覆盖精度

雷庆祝 秦永松

(广西师范大学数学科学学院，桂林， 541004)

本文用经验似然方法讨论了条件密度的置信区间的构造．通过对覆盖概率的 Edgeworth 展开得到了经验似然置信区间的覆盖精度，同时证明了条件密度的经验似然置信区间的 Bartlett 可修正性．

关 键 词： 置信区间，经验似然，双重核估计， Bartlett 可修正性.

学 科 分 类 号： O212.