8-2006

Smoothed Block Empirical Likelihood for Quantiles of Weakly Dependent Processes

Song Xi Chen
Iowa State University, songchen@iastate.edu

Chiu Min Wong
Temasek Junion College

Smoothed Block Empirical Likelihood for Quantiles of Weakly Dependent Processes

Abstract
Inference on quantiles associated with dependent observation is a commonly encountered task in risk management these days. This paper considers employing the empirical likelihood to construct confidence intervals for quantiles of the stationary distribution of a weakly dependent process. To accommodate data dependence and avoid any secondary variance estimation, the empirical likelihood is formulated based on blocks of observations. To reduce the length of the confidence intervals, the weighted empirical distribution is smoothed by a kernel function and a smoothing bandwidth. It shows that a rescaled version of the smoothed block empirical likelihood ratio admits a limiting chi-square distribution with one degree of freedom, which facilitates likelihood ratio confidence intervals for the quantiles.

Keywords
alpha-mixing, empirical likelihood, kernel smoothing, quantile, Value-at-Risk, Peking University

Disciplines
Statistics and Probability

Comments
This preprint was published as Song Xi Chen and Chiu Min Wong, " Smoothed Block Empirical Likelihood for Quantiles of Weakly Dependent Processes", *Statistica Sinica* (2009): 71-81.

Smoothed Block Empirical Likelihood for Quantiles of Weakly Dependent Processes

Song Xi Chen
Department of Statistics
Iowa State University
Ames, Iowa, USA

Chiu Min Wong
Department of Mathematics
Temasek Junior College
Singapore

August 11, 2006

Abstract: Inference on quantiles associated with dependent observation is a commonly encountered task in risk management these days. This paper considers employing the empirical likelihood to construct confidence intervals for quantiles of the stationary distribution of a weakly dependent process. To accommodate data dependence and avoid any secondary variance estimation, the empirical likelihood is formulated based on blocks of observations. To reduce the length of the confidence intervals, the weighted empirical distribution is smoothed by a kernel function and a smoothing bandwidth. It shows that a rescaled version of the smoothed block empirical likelihood ratio admits a limiting chi-square distribution with one degree of freedom, which facilitates likelihood ratio confidence intervals for the quantiles. The practical performance of these confidence intervals is confirmed by a simulation study.

Key Words: α-mixing, Empirical Likelihood, Kernel Smoothing, Quantile, Risk Analysis, Value-at Risk;

1 Introduction

Let $X_1, ..., X_N$ be a sequence of weakly dependent stationary random variables, and F be their common marginal distribution. The interest of this paper is to construct confidence intervals for $\theta_q =: F^{-1}(q) = \inf\{x|F(x) \geq q\}$, the q-th quantile of F for $q \in (0,1)$. In financial risk management, θ_q is called the Value-at-Risk which specifies the level of excessive losses at a confidence level $1-q$. As financial returns are most likely dependent, the proposed confidence intervals for θ_q have direct applications in risk management.

We propose using the empirical likelihood for the construction of confidence intervals for θ_q. Empirical likelihood introduced by Owen (1988, 1990) is a nonparametric method of inference that enables a likelihood type inference in a nonparametric setting. Two striking properties of the empirical likelihood are the Wilks' theorem and Bartlett correction, which mirror those of a parametric likelihood. Qin and Lawless (1994) established the Wilks' theorem for estimating equations based empirical likelihood, and Chen and Cui (2006a, 2006b) showed that this empirical likelihood is Bartlett corretable with or without nuisance parameters. Tsao (2004) studied the effect of the number of constraints on the coverage probability of the empirical likelihood confidence intervals for a mean parameter. Like its parametric counterpart, the empirical likelihood confidence intervals/regions are constructed by contouring the empirical likelihood ratio, which brings two benefits. One is that their shape and orientations are naturally determined by data. Another is that the intervals/regions are obtained without a secondary estimation.

These features of the empirical likelihood confidence intervals are the major motivations for our current proposal for quantiles. Indeed, when considering extreme quantiles in risk analysis, the distribution of the sample quantile estimator can be quite skewed. Therefore, it is more appealing to have confidence intervals which are naturally determined by the data rather than forcing them to be symmetric about an point estimate as the case for intervals based on the asymptotic normality of the sample quantile estimator. The fact that the empirical likelihood intervals are obtained by contouring the likelihood ratio without an secondary variance estimation is particularly advantageous for dependent data. This is because the data dependence leads to a variance which involves covariances at all lags, which makes its estimation much more involved than independent cases.

A key ingredient of our proposal is to smooth the weight empirical distribution function. The purpose of the kernel smoothing is to reduce the length of the confidence intervals, which is clearly demonstrated in our simulation study. Combining the empirical likelihood and the kernel smoothing for confidence intervals of a quantile with independent and identically distribution was proposed in Chen and Hall (1993), which shows that employing a kernel quantile estimator together with the empirical likelihood remarkably reduces the coverage errors from $O(N^{-1/2})$ to $O(N^{-1})$ before the Bartlett correction and to $O(N^{-7/4})$ after Bartlett correction. Further investigations have been carried out by Zhou and Jing (2003a) and (2003b). Quantile estimation using empirical likelihood in the context of survey sampling is considered in Chen and Wu (2004).

The paper is organized as follows. We introduced an kernel smoothed empirical likelihood for a quantile based on blocks of data in Section 2. Section 3 reports the main results of the paper. Results from a simulation study are reported in Section 4. All the technical details are relegated in the appendix.

2 Block Empirical Likelihood for Quantiles

let $\mathcal{F}_k^l$ be the σ-algebra of events generated by $\{X_t, k \le t \le l\}$ for $l \ge k$. The α-mixing coefficient introduced by Rosenblatt (1956) is

$$\alpha(k) = \sup_{\mathcal{A} \in \mathcal{F}_1^i, \mathcal{B} \in \mathcal{F}_{i+k}^\infty} |P(AB) - P(A)P(B)|.$$

The series is said to be α-mixing if $\lim_{k\to\infty} \alpha(k) = 0$. The dependence described by α-mixing is the weakest, as it is implied by other types of mixing; see Doukhan (1994) for a comprehensive discussion on mixing.

Let $F_n(x) = n^{-1} \sum_{i=1}^n I(X_i \le x)$ be the empirical distribution of the weakly dependent data $\{X_i\}_{i=1}^N$, where $I(\cdot)$ is the indicator function. We first smooth the empirical distribution with a kernel K and a smoothing bandwidth h. And then convert it to obtain kernel quantile estimator which is smoother than the conventional sample quantile estimator.

Let K be a r-th order kernel which satisfies

$$\int u^j K(u)du = \begin{cases} 1, & \text{if } j = 0; \\ 0, & \text{if } 1 \leq j \leq r-1; \\ \kappa, & \text{if } j = r. \end{cases} \tag{2.1}$$

for some integer $r \geq 2$ and some $\kappa \neq 0$. Furthermore let $G_h(x) = \int_{-\infty}^{x/h} K(y)dy$ where h is the smoothing bandwidth such that $h \to 0$ as $N \to \infty$.

A kernel estimator of $F(x)$ is $\hat{F}_{n,h}(x) = n^{-1}\sum_{i=1}^n G_h(x - X_i)$ and the kernel quantile estimator $\hat{\theta}_{q,h}$ is the solution of

$$\hat{F}_{n,h}(x) = q.$$

Kernel estimators have been applied to estimation and testing for time series data; see Robinson (1989), Hjellvik and Tjøstheim (1995), Hjellvik, Chen and Tjøstheim (2004) and the book of Fan and Yao (2003).

Chen and Tang (2005) studied the statistical properties of $\hat{\theta}_{n,h}$ and variance estimation. Unlike estimation of a regression or a probability density function for weakly dependent observations, the data dependence contributes to the leading order variance of the kernel quantile estimator. In particular, for each $h > 0$ let $\gamma_h(k) = Cov\{G_h(\theta_q - X_1), G_h(\theta_q - X_{k+1})\}$. The leading variance term of $\hat{\theta}_{n,h}$ can be written as

$$\sigma_{n,h}^2 = q(1-q) + 2\sum_{k=1}^{n-1}(1 - k/n)\gamma_h(k) \tag{2.2}$$

which indicates clearly the first order effect of dependence. Chen and Tang (2005) proposed estimating the variance via a kernel estimation of the spectral density of a derived sequence. The variance estimator together with the asymptotic normality of $\hat{\theta}_{q,h}$ can be used to obtain confidence intervals for θ_q.

There are two limitations with confidence intervals based on the asymptotic normality. One is that the intervals are always symmetric. However, for extreme quantiles commonly used in risk analysis, the finite sample distribution of the quantile estimator can be quite skewed. Therefore, it is more appealing to have asymmetric confidence intervals to reflect the skewness of the underlying distribution. Another limitation is that a secondary variance estimation is required for (2.2). In Chen and Tang (2005), the estimation is via estimating the spectral density function. For spectral density estimation, see Brockwell and Davis (1991).

The proposed empirical likelihood intervals for θ_q are not only asymmetric but also free of any secondary variance. The latter is due to empirical likelihood's ability to standardise internally via its built-in algorithm.

Let $\{p_i\}_{i=1}^N$ be probability weights adding to one. A weighted kernel estimator for the distribution function F is

$$\hat{F}_{p,h}(x) = \sum_{i=1}^{N} p_i G_h(x - X_i). \tag{2.3}$$

If the observations were independent and identically distributed, we could formulate the empirical likelihood for the quantile θ_q by

$$L_h(\theta_q) = \sup \prod_{i=1}^{N} p_i$$

subject to $\hat{F}_{p,h}(\theta_q) = q$ and $\sum_{i=1}^N p_i = 1$. This is the formulation of Chen and Hall (1993). However, as pointed out by Kitamura (1997) in the context of estimating equations, the above single point based empirical likelihood would ignore data dependency and cause the empirical likelihood ratio to lose its limiting chi-square distribution. The latter has been a major attraction of the likelihood ratio statistics.

In this paper we extend Kitamura (1997) by introducing a smoothing bandwidth into the estimating equation which makes the estimating equation dependent of the sample size. In order to capture data dependence, we employ the blocking technique which was first applied for the bootstrap method (Carlstein, 1986; Künsch, 1989) and applied to empirical likelihood by Kitamura (1997). The data blocking divides the entire sample into a sequence of data blocks. The block length is taken to be sufficiently large so that the data dependence can be captured. At the same time, the weakly dependence allows us to treat the blocks as independent if the gap between successive blocks becomes large, although this gap will be much smaller than the block length.

Let M be a positive integer representing the block length and L be the gap between the beginnings of two consecutive blocks, and Q be the total number of blocks so that

$Q = [(N-M)/L] + 1$. Assumptions on M and L will be specified in Condition C4 in the next section.

For $i = 1, \ldots, Q$ define $g_h(X_i, \theta_q) = G_h(\theta_q - X_i) - q$ and $T_i(\theta_q) = \frac{1}{M}\sum_{j=1}^{M} g_h(X_{(i-1)L+j}, \theta_q)$ be the i-th block average.

Let $p_1, \cdots, p_Q$ be empirical likelihood weights allocated to the Q blocks respectively. The block empirical likelihood for θ_q is

$$L_h(\theta_q) = \sup \prod_{i=1}^{Q} p_i \tag{2.4}$$

subject to $\sum_{i=1}^{Q} p_i = 1$ and $\sum_{i=1}^{Q} p_i T_i(\theta_q) = 0$. From the standard algorithm of empirical likelihood, the optimal p_i that maximize the profile likelihood (2.4) is

$$p_i = \frac{1}{Q[1 + \lambda(\theta_q) T_i(\theta_q)]} \tag{2.5}$$

where $\lambda(\theta_q)$ is a Lagrange multiplier satisfying

$$\sum_{i=1}^{Q} \frac{T_i(\theta_q)}{1 + \lambda(\theta_q) T_i(\theta_q)} = 0. \tag{2.6}$$

Since $L_h(\theta_q)$ attains its maximum at $p_i = Q^{-1}$ for all $i \in \{1, \ldots, Q\}$, we can define

$$\ell_h(\theta_q) = -2 \log\{L_h(\theta_q)/Q^{-Q}\} \tag{2.7}$$

to be the log empirical log-likelihood ratio for θ_q. From (2.5),

$$\ell_h(\theta_q) = 2 \sum_{i=1}^{Q} log[1 + \lambda(\theta_q) T_i(\theta_q)], \tag{2.8}$$

where $\lambda(\theta_q)$ is the solution of (2.6).

If we choose $h = 0$ in the above formulation, then $g_h(X_i, \theta_q) = G_h(\theta_q - X_i) - q$ is degenerated to $I(X_i \leq \theta_q) - q$, namely the estimating equation is free of N. Then the results in Kitamura (1997) are applicable to this unsmoothed empirical likelihood formulation for the quantile.

3 Main Results

We assume the following conditions in our investigation:

C1: $\{X_i\}_{i=1}^N$ is a strictly stationary α-mixing sequence. The mixing coefficient $\alpha(k)$ satisfy $\sum_{k=1}^{\infty} k\alpha^{1/p}(k) < \infty$ for some $p > 1$. Let $\phi(t)$ be the spectral density function of $\{I(X_k < \theta_q)\}_{k=1}^N$. We assume $\phi(0) > 0$.

C2: Let K be a bounded and compactly supported rth order kernel satisfying (2.1); and the smoothing bandwidth h satisfies $Nh^{2r} \to 0$ but $Nh \to \infty$ as $N \to \infty$.

C3: The distribution function F of X_i is absolutely continuous with a density f which has continuous $(r-1)$-th derivatives in a neighbourhood of θ_q and $f(\theta_q) > 0$.

C4: The block length satisfies $M \to \infty$ and $M = o(N^{1/2})$ as $N \to \infty$ and the gap L between the starting points of two adjacent blocks satisfies $kL \le M$ and $(k+1)L > M$ for some $k > 0$.

We first establish the order of magnitude for the Lagrange multiplier $\lambda(\theta_q)$, which is a key result in establishing stochastic expansions for the empirical likelihood ratio $l_h(\theta_q)$.

Theorem 1 Under Conditions C1-C4, $\lambda(\theta_q) = O_p\{M(N^{-1/2} + h^r)\}$.

The next theorem shows that a scaled version of the empirical likelihood ratio converges to the χ_1^2 distribution.

Theorem 2 Under Conditions C1-C4 and as $N \to \infty$

$$\frac{N}{MQ} l_h(\theta_q) \xrightarrow{d} \chi_1^2.$$

Theorem 2 readily leads to an empirical likelihood confidence interval for θ_q at $1-\alpha$ level of confidence

$$I_{\alpha,h} = \{\theta_q | \frac{N}{MQ} \ell_h(\theta_q) \le c_\alpha\}.$$

where c_α is the upper α-quantile of χ_1^2 such that $P(\chi_1^2 > c_\alpha) = \alpha$. Theorem 2 ensures that I_α will attain the nominal coverage level $1-\alpha$ asymptotically. A major attraction of the proposed confidence interval is in its avoiding any secondary estimation of the variance of the kernel quantile estimator $\hat{\theta}_{q,h}$ given by (2.2).

If we choose not to carry out the kernel smoothing in the empirical likelihood formulation, which effectively assigns $h = 0$ as discussed at the end of last section, then Theorem 2 is still valid as a special case of Kitamura (1997). Let $l_0(\theta_q)$ be the unsmoothed empirical likelihood ratio. Then, a $1-\alpha$ confidence unsmoothed empirical likelihood confidence interval for θ_q is

$$I_{\alpha,0} = \{\theta_q | \frac{N}{MQ} \ell_0(\theta_q) \leq c_\alpha\}.$$

We expect that the smoothed confidence intervals $I_{\alpha,h}$ have shorter length than $I_{\alpha,0}$. This is based on a fact (Chen and Tang, 2005) that the kernel estimator for θ_q reduces the variance of the unsmoothed sample quantile estimator at the second order. This is indeed confirmed by the simulation study reported in the next section.

4 Simulation Results

We report results from a simulation study that is designed to evaluate the performance of the empirical likelihood confidence intervals for the quantile θ_q. For comparison purpose, we carry out simulation for both the kernel smoothed intervals $I_{\alpha,h}$ and unsmoothed intervals $I_{\alpha,0}$. We are interested in the lengths and coverage levels of the confidence intervals.

We considered two time series models in the simulation. One is an AR(1) model

$$X_t = 0.5X_{t-1} + \varepsilon_t$$

and an AR(2) model

$$X_t = \tfrac{5}{6}X_{t-1} - \tfrac{1}{6}X_{t-2} + \varepsilon_t.$$

In both models, ε_t are independent and identically distributed $N(0, 1)$ random variables. Clearly, both models are strictly stationary and α-mixing.

Two levels of quantiles were considered: the 5% and 50% (median) quantiles. The former is a level commonly used in risk assessment. The sample sizes considered were $N = 300$ and 500. The block length M was 12 for $N = 300$ and 16 for $N = 500$. We set the gap between two successive blocks L to be half of M in all cases. We employed the second order ($r = 2$) Epanechnikov kernel throughout the simulation. Three bandwidths were used for the kernel smoothed interval: $h_1 = 1.50N^{-1/4}$, $h_2 = N^{-1/4}$ and $h_3 = 0.50N^{-1/4}$.

The confidence intervals for the 5% and 50% quantiles with confidence level 0.95 and 0.99 are reported in Tables 1 for the AR(1) model and in Table 2 for the AR(2) model. We observe from Tables 1 and 2 that both smoothed and unsmoothed confidence intervals had similar and both satisfactory coverage in all the cases considered in the simulation. The observed empirical coverage was not sensitive to the choice of the smoothing bandwidth h. From our discussion in the previous section, we have anticipated the kernel smoothed confidence intervals to be substantially shorter than their unsmoothed counterpart and this has turned out to be the case. Our discovery clearly exhibited the usefulness of kernel smoothing in the context of interval estimation for dependent observations.

5 Appendix: Technical Details

For each $h > 0$, let $\gamma_h(\ell) = Cov\{g_h(X_1, \theta_q), g_h(X_{\ell+1}, \theta_q)\}$ and

$$\phi_h(t) = (2\pi)^{-1} \sum_{\ell=-\infty}^{\infty} \gamma_h(\ell) exp(-i\ell t) \quad \text{for} \quad t \in [-\pi, \pi]$$

be the spectral density function of $\{g_h(X_k, \theta_q)\}_{k=1}^N$. We note from C2 that $\phi_h(t)$ exists for each given h and $t \in [-\pi, \pi]$. Similarly, we define $\phi(t)$ to be the spectral density function of $\{I(X_k < \theta_q)\}_{k=1}^N$. From Lemma 1 of Chen and Tang (2005), $\phi_h(0) = \phi(0) + o(1)$ as $N \to \infty$.

Let $T_i(\theta_q) = M^{-1} \sum_{j=1}^M g_h(X_{(i-1)L+j}, \theta_q)$ and $\tilde{T}_\beta(\theta) = Q^{-1} \sum_{i=1}^Q [T_i(\theta_q)]^\beta$, where β is any positive integer. In particular, take $S_Q = \tilde{T}_2(\theta_q)$. We need lemmas in the proof of Theorems 1 and 2.

Lemma A.1 Under Conditions C1-C4, $\tilde{T}_\beta(\theta_q) = O_p((M^{-1/2} + h^r)^\beta)$ for any integer $\beta \geq 2$.

Proof: It can be shown that Condition C1 implies that $\{T_i^\beta(\theta_q)\}_{i=1}^Q$ is strictly stationary. Construct a new sequence $\{\tau_i^\beta(\theta_q)\}_{i=1}^Q$ based on the blocks with mean zero.

$$E[\tilde{T}_\beta(\theta_q)] = E[T_1^\beta(\theta_q)]$$

$$
\begin{aligned}
&= E[\tau_1(\theta_q) + c_0 h^r + o(h^r)]^\beta \\
&= E[\sum_{y=0}^{\beta} \binom{\beta}{y} \{\tau_1(\theta_q)\}^y \{c_o h^r + o(h^r)\}^{\beta-y}]. \qquad \text{(A.1)}
\end{aligned}
$$

Note that $E[\tau_i^\beta(\theta_q)] = O(M^{-\beta/2})$ by applying the moment inequality in Yokoyama (1980). Hence for any $\beta \geq 2$, (A.1) leads to

$$
E[\tilde{T}_\beta(\theta_q)] = O[(h^r)^\beta + (h^r)^{\beta-1}M^{-1/2} + ... + (M^{-1/2})^\beta] = O((M^{-1/2} + h^r)^\beta) \qquad \text{(A.2)}
$$

Next write $u(i) = Cov(T_1^\beta(\theta_q), T_{i+1}^\beta(\theta_q))$. If we take k as the greatest integer satisfying $kL \leq M$, then $Var(\tilde{T}_\beta(\theta_q))$ can be rewritten as

$$
\begin{aligned}
QVar(\tilde{T}_\beta(\theta_q)) &= u(0) + 2\sum_{i=1}^{k+1}(1 - \frac{i}{Q})u(i) + 2\sum_{i=k+2}^{Q-1}(1 - \frac{i}{Q})u(i) \\
&\leq u(0) + 2\sum_{i=1}^{k+1}|u(i)| + 2\sum_{i=k+2}^{Q-1}(|u(i)| \qquad \text{(A.3)}
\end{aligned}
$$

We shall show that the three terms on the right hand side of (A.3) denoted by A_1, A_2 and A_3 respectively will all converge to zero. From (A.2),

$$
u(0) = E(T_1^{2\beta}(\theta_q)) = O((M^{-1/2} + h^r)^\beta).
$$

By applying the definition of covariance and then the Cauchy-Schwartz inequality on A_2, we can get

$$
\begin{aligned}
A_2 &\leq 2\sum_{i=1}^{k+1}\{|E[T_1^\beta(\theta_q)T_{i+1}^\beta(\theta_q)]| + |E^2[T_1^\beta(\theta_q)]|\} \\
&\leq 2\sum_{i=1}^{k+1}\{|E[T_1^{2\beta}(\theta_q)]E[T_{i+1}^{2\beta}(\theta_q)]|^{1/2} + |E^2[T_1^\beta(\theta_q)]|\} \\
&= O((M^{-1/2} + h^r)^\beta).
\end{aligned}
$$

Use the bound derived by Roussas and Ioannides (1987, p109) on A_3 to obtain

$$
A_3 \leq \sum_{i=k+2}^{Q-1} 10\{E[T_1^{\beta q*}(\theta_q)]\}^{1/q*}\{E[T_{i+1}^{\beta r*}(\theta_q)]\}^{1/r*}\alpha_T(i)^{1/p*},
$$

where $p*, q*$ and $r* > 1$ and $\frac{1}{p*}+\frac{1}{q*}+\frac{1}{r*} = 1$. Since $\{E[T_1^{\beta q*}(\theta_q)]\}^{1/q*} = \{E[T_{i+1}^{\beta r*}(\theta_q)]\}^{1/r*} = O((M^{-1/2}+h^r)^{\beta/2})$, and $\alpha_T(i) \leq \alpha_(iL-M)$ for $i \geq k+2$ from Politis and Romano (1992), we can conclude that

$$\begin{aligned} A_3 &\leq C_1(M^{-1/2}+h^r)^\beta \sum_{i=k+2}^{Q-1} (\alpha(iL-M))^{1/p*} \\ &\leq C_1(M^{-1/2}+h^r)^\beta \sum_{i=1}^{\infty} \alpha(i)^{1/p*} = O((M^{-1/2}+h^r)^\beta). \end{aligned}$$

Therefore $Var(\tilde{T}_\beta(\theta_q)) = O((M^{-1/2}+h^r)^\beta Q^{-1})$. Lastly a straightforward usage of Cheybeshev's inequality concludes the proof.

Lemma A.2 Under Conditions C1-C4, $MS_Q \xrightarrow{p} 2\pi\phi(0) > 0$.

Proof: Although Lemma A.1 tells us that $E[MS_Q] = O((1+M^{1/2}h^r)^2)$, but here we will attempt to show that its limit is $2\pi\phi(0)$.

$$\begin{aligned} E(MS_Q) &= \frac{1}{M} Var\{\sum_{i=1}^{M} g_h(X_i,\theta_q)\} = \gamma_h(0) + 2\sum_{i=1}^{M-1}(1-\frac{i}{M})\gamma_h(i) \\ &= \gamma_h(0) + 2\sum_{i=1}^{\infty}\gamma_h(i) - 2\sum_{i=M}^{\infty}\gamma_h(i) - \frac{2}{M}\sum_{i=1}^{M-1} i\gamma_h(i). \end{aligned}$$

Clearly the first two terms in the last step is $2\pi\phi(0)$. For the remaining two terms, we can use Davydov's inequality to show that

$$\begin{aligned} |\frac{2}{M}\sum_{i=1}^{M-1} i\gamma_h(i) + 2\sum_{i=M}^{\infty}\gamma_h(i)| &\leq \frac{2}{M}\sum_{i=1}^{\infty} i|\gamma_h(i)| \\ &\leq \frac{2}{M}\sum_{i=1}^{\infty} i2p(2\alpha_g(i))^{1/p}\{E[g_h(X_i,\theta_q)]\}^{1-1/p} \\ &\leq \frac{C_1}{M}\sum_{i=1}^{\infty} i\alpha(i)^{1/p}\{c_o h^r + o(h^r)\}^{1-1/p} = o(M^{-1}h^r). \end{aligned}$$

Hence $E(MS_Q) = 2\pi\phi(0) + o(1)$. Observe also that $Var(MS_Q) = O((1+M^{1/2}h^r)^2 Q^{-1})$ from Lemma A.1. Our result then follows from Cheybeshev's inequality.

Lemma A.3. Under Conditions C1-C4, $\tilde{T}_1(\theta_q) = O_p(N^{-1/2}+h^r)$.

Proof:

$$\begin{aligned} Var\{\tilde{T}_1(\theta_q)\} &= Var[Q^{-1}\sum_{i=1}^{Q}\tau_i(\theta_q) + c_0h^r + o(h^r)] \\ &= Var[Q^{-1}\sum_{i=1}^{Q}\tau_i(\theta_q)] \\ &= Q^{-1}E[\tau_1^2(\theta_q)] = O\{(MQ)^{-1}\} = O(N^{-1}). \end{aligned}$$

Since $E[\tilde{T}_1(\theta_q)] = c_oh^r + o(h^r)$, it can be easily deduced that

$$\tilde{T}_1(\theta_q) = O_p\{N^{-1/2} + h^r\} = O_p(N^{-1/2})$$

which follows from Condition C2.

Lemma A.4. Take $\xi_Q = Q^{-1}\sum_{i=1}^{Q}\tau_i^2(\theta_q)$. Under Conditions C1-C4,

$$M\xi_Q \xrightarrow{p} 2\pi\phi(0) > 0.$$

Proof:

$$\begin{aligned} E[M\xi_Q] &= \frac{M}{Q}\sum_{i=1}^{Q}\{E[T_i(\theta_q)] - (c_0h^r + o(h^r))\}^2 \\ &= E(MS_Q) - 2ME[\tilde{T}_1(\theta_q)][c_0h^r + o(h^r)] + M[c_0h^r + o(h^r)]^2] \\ &= E(MS_Q) + O(Mh^{2r}) = 2\pi\phi(0) + o(1). \end{aligned}$$

It can also be shown that $V(M\xi_Q) = O(Q^{-1})$ by a proof similar to Lemma A.1. Hence from Cheybeshev's result, the statement is proved.

Proof of Theorem 1: By following the standard procedure in empirical likelihood, for instance that outlined in Owen (1990),

$$\begin{aligned} 0 &= |g(\lambda(\theta_q))| \\ &\geq \frac{|\lambda(\theta_q)|}{Q}\sum_{i=1}^{Q}\frac{T_i^2(\theta_q)}{1+\lambda(\theta_q)T_i(\theta_q)} - \frac{1}{Q}|\sum_{i=1}^{Q}T_i(\theta_q)| \\ &\geq \frac{|\lambda(\theta_q)|}{1+|\lambda(\theta_q)|max_{1\leq i\leq Q}|T_i(\theta_q)|}MS_Q - |M\tilde{T}_1(\theta_q)|. \end{aligned}$$

From Lemma A.2, $MS_Q = 2\pi\phi(0) + o_p(1)$. Recall also Lemma A.3, that $|M\tilde{T}_1(\theta_q)| = O_p(MN^{-1/2})$. This means that

$$\frac{|\lambda(\theta_q)|}{1 + |\lambda(\theta_q)| \max_{1\le i\le Q} |T_i(\theta_q)|} = O_p(MN^{-1/2}).$$

As $MN^{-1/2} = o(1)$, $h^r = o(N^{1/2}M^{-1})$ where $N^{1/2}M^{-1} \to \infty$ as $N \to \infty$. From a result in Künsch (1989), we see that

$$\begin{aligned}\max_{1\le i\le Q} |T_i(\theta_q)| &= \max_{1\le i\le Q} |\tau_i(\theta_q) + c_o h^r + o(h^r)| \\ &= o(N^{1/2}M^{-1}) + O(h^r) = o(N^{1/2}M^{-1}).\end{aligned}$$

Then we conclude that $|\lambda(\theta_q)| = O_p(MN^{-1/2})$ because

$$1 + |\lambda(\theta_q)| \max_{1\le i\le Q} |T_i(\theta_q)| = 1 + o_p(1).$$

Proof of Theorem 2: We first develop an expansion for $\lambda(\theta_q)$. Note that

$$\begin{aligned}0 &= g(\lambda(\theta_q)) = Q^{-1}\sum_{i=1}^{Q} \frac{T_i(\theta_q)}{1+\lambda(\theta_q)T_i(\theta_q)} \\ &= Q^{-1}\sum_{i=1}^{Q} T_i(\theta_q)[1 - \lambda(\theta_q)T_i(\theta_q) + \frac{\lambda^2(\theta_q)T_i^2(\theta_q)}{1+\lambda(\theta_q)T_i(\theta_q)}] \\ &= \tilde{T}_1(\theta_q) - \lambda(\theta_q)S_Q + Q^{-1}\sum_{i=1}^{Q} T_i(\theta_q)[\frac{\lambda^2(\theta_q)T_i^2(\theta_q)}{1+\lambda(\theta_q)T_i(\theta_q)}]. \qquad \text{(A.4)}\end{aligned}$$

To obtain an explicit expression for $\lambda(\theta_q)$, it is necessary to get the order of $\sum_{i=1}^{Q} |T_i(\theta_q)|^3$, and this is

$$\begin{aligned}Q^{-1}\sum_{i=1}^{Q} |T_i(\theta_q)|^3 &\le \max_{1\le i\le Q} |T_i(\theta_q)| S_Q \\ &= o(N^{1/2}M^{-1}) O_p\{(M^{-1/2} + h^r)^2\} \\ &= o_p\{N^{1/2}M^{-2}(1 + M^{1/2}h^r)^2\} = o_p(N^{1/2}M^{-2})\end{aligned}$$

as $Nh^{2r} \to 0$ implies $Mh^{2r} \to 0$ and $M^2N^{-1/2} = o(N^{1/2})$. The last term in (A.4) becomes

$$Q^{-1}\sum_{i=1}^{Q} T_i(\theta_q)[\frac{\lambda^2(\theta_q)T_i^2(\theta_q)}{1+\lambda(\theta_q)T_i(\theta_q)}] \le Q^{-1}\sum_{i=1}^{Q} |T_i(\theta_q)|^3 |\lambda(\theta_q)|^2 (1 + \lambda(\theta_q)| \max_{1\le i\le Q} T_i(\theta_q)|)^{-1}$$

$$
\begin{aligned}
&= Q^{-1}\sum_{i=1}^{Q}|T_i(\theta_q)|^3|\lambda(\theta_q)|^2(1+o_p(1))^{-1}\\
&= o_p(N^{1/2}M^{-2})O_p(MN^{-1/2})^2 = o_p(N^{-1/2})
\end{aligned}
$$

Now we have an expansion on $\lambda(\theta_q)$ based on (A.4):

$$\lambda(\theta_q) = \frac{\tilde{T}_1(\theta_q)}{S_Q} + \beta,$$

where $\beta = o_p(MN^{-1/2})$. It can be shown that $\beta = o_p\left(\frac{\tilde{T}_1(\theta_q)}{S_Q}\right)$.

The adjusted empirical log-likelihood ratio

$$
\begin{aligned}
& \frac{2N}{MQ}\sum_{i=1}^{Q} log(1+\lambda(\theta_q)T_i(\theta_q)) \\
= & \frac{2N}{M}\sum_{j=1}^{\infty}\frac{(-1)^{j+1}\lambda^j(\theta_q)\tilde{T}_j(\theta_q)}{j} \\
= & \frac{2N}{M}\tilde{T}_1(\theta_q)\lambda(\theta_q) - \frac{N}{M}S_Q\lambda^2(\theta_q) + \frac{2N}{MQ}\sum_{i=1}^{Q}\eta_i \\
= & \frac{2N}{M}\tilde{T}_1(\theta_q)(\frac{\tilde{T}_1(\theta_q)}{S_Q}+\beta) - \frac{N}{M}S_Q(\frac{\tilde{T}_1(\theta_q)}{S_Q}+\beta)^2 + \frac{2N}{MQ}\sum_{i=1}^{Q}\eta_i \\
= & \frac{N}{M}\frac{\tilde{T}_1^2(\theta_q)}{S_Q} - \frac{N}{M}S_Q\beta^2 + \frac{2N}{MQ}\sum_{i=1}^{Q}\eta_i, \qquad \text{(A.5)}
\end{aligned}
$$

where $\sum_{i=1}^{Q}\eta_i$ consists of the third and higher order terms in the Taylor expansion of $2\sum_{i=1}^{Q} log\{1+\lambda(\theta_q)T_i(\theta_q)\}$ respect to $\lambda(\theta_q)T_i(\theta_q)$, each having coefficient of the form $\lambda^r(\theta_q)\sum_{i=1}^{Q}T_i^r(\theta_q)$ for $r \geq 3$. In particular, there exists a positive constant D such that

$$P(|\eta_i| \leq D\lambda^3(\theta_q)\sum_{i=1}^{Q}T_i^3(\theta_q)) \to 1.$$

We note from Lemma A.2 and Lemma A.4 that

$$MS_Q = M\xi_Q + o_p(1) = M\xi_Q[1+o_p(1)].$$

Now the first term in (A.5)

$$\frac{N}{M}\frac{\tilde{T}_1^2(\theta_q)}{S_Q} = \frac{N}{M}\frac{\bar{\tau}^2(\theta_q)}{\xi_Q(1+o_p(1))} + \frac{N\{c_0h^r+o(h^r)\}\{2\bar{\tau}(\theta_q)+c_0h^r+o(h^r)\}}{2\pi\phi(0)+o_p(1)}$$

$$
\begin{aligned}
&= \frac{N}{M}\frac{\bar{\tau}^2(\theta_q)}{\xi_Q} + O_p\{Nh^r(N^{-1/2}+h^r)\} \\
&= \frac{N}{M}\frac{\bar{\tau}^2(\theta_q)}{\xi_Q} + o_p(1).
\end{aligned}
$$

We note that the central limit theorem for α-mixing sequences (Bosq, 1998) implies that

$$
\sqrt{\frac{N}{M}}\frac{\bar{\tau}(\theta_q)}{\sqrt{\xi_Q}} \xrightarrow{L} N(0,1).
$$

Hence

$$
\frac{N}{M}\frac{\bar{\tau}^2(\theta_q)}{\xi_Q} \xrightarrow{L} \chi_1^2.
$$

Furthermore,

$$
\begin{aligned}
\frac{N}{M}S_Q\beta^2 &= O(NM^{-1})O_p\{(M^{-1/2}+h^r)^2\}o_p(M^2N^{-1}) \\
&= o_p(1+M^{1/2}h^r)^2 = o_p(1).
\end{aligned}
$$

Finally as $MQ \geq N$

$$
\begin{aligned}
\frac{N}{MQ}|2\sum_{i=1}^{Q}\eta_i| &\leq \frac{2N}{MQ}D|\lambda(\theta_q)|^3\sum_{i=1}^{Q}|T_i(\theta_q)|^3 \\
&\leq 2D|\lambda(\theta_q)|^3\sum_{i=1}^{Q}|T_i(\theta_q)|^3 \\
&= O_p(M^3N^{-3/2})o_p(N^{1/2}M^{-2}Q) \\
&= o_p(MQN^{-1}) = o_p(1).
\end{aligned}
$$

This concludes the proof.

References:

Bosq, D. (1998). Nonparametric Statistics For Stochastic Processes: Estimation and Prediction, Springer-Verlag: New York, 1998.

Brockwell P. J. and Davies, R. A. (1991). Time Series: Theory and Methods, Springer-Verlag: New York, 1991.

Carlstein, E. (1986). The Use of Subseries Values for Estimating the Variance of a General Statistic from a Stationary Sequence, *The Annals of Statistics* **14**, p1171-1179.

Chen, J. and Wu, Changbao. (2002), Estimation of distribution function and quantiles using the model-calibrated pseudo empirical likelihood method. *Statist. Sinica* 12, 1223–1239.

Chen, S. X. and Cui, H. J. (2006a), On the Bartlett properties of empirical likelihood with moment restrictions. *Journal of Econometrics*, to appear.

Chen, S. X. and Cui, H. J. (2006b), On Bartlett Correction of Empirical Likelihood in the Presence of Nuisance Parameters. *Biometrika* 93, 215-220.

Chen, S. X. and Hall P. (1993). Smoothed Empirical Likelihood Confidence Intervals for Quantiles, *The Annals of Statistics* **21**, p1166-1181.

Chen, S. X. and Tang, C. Y. (2005). Nonparametric Inference of Value at Risk for dependent Financial Returns, *Journal of Financial Econometrics*, 3, 227-255.

Doukhan P. (1994). Mixing: Properties and Examples, Springer-Verlag: New-York, 1994.

Fan, J., and Q. Yao (2003) *Nonlinear Time Series. Nonparametric and Parametric Methods*, Springer-Verlag, New York.

Freedman D. A. (1984). On Bootstrapping Two-Stage Least Squares Estimates in Stationary Linear Models, *The Annals of Statistics* **12**, p827-842.

Hjellvik, V., Chen, R. and Tjøstheim, D. (2004). Nonparametric estimation and testing in panels of intercorrelated time series. *Journal of Time Series Analysis* **25**, 831–872.

Hjellvik, V. and Tjøstheim, D. (1995). Nonparametric tests of linearity for time series. *Biometrika* **82**, , 351–368.

Kitamura Y. (1997). Empirical Likelihood Methods with Weakly Dependent Processes, *The Annals of Statistics* **25**, p2084-2102.

Künsch H. R. (1989). The Jackknife and the Bootstrap for General Stationary Observations, *The Annals of Statistics* **17**, p1217-1241.

Owen A. (1988). Empirical Likelihood Ratio Confidence Intervals for a Single Functional, *Biometrika* **75**, p237-249.

Owen A. (1990). Empirical Likelihood Ratio Confidence Regions, *The Annals of Statistics* **18**, p90-120.

Owen A. (2001). Empirical Likelihood, Chapman and Hall: Boca Raton, FL, 2002.

Politis D. N. and Romano J. P. (1992). A General Resampling Scheme for Triangular Arrays of α-mixing Random Variables with Applications to the Problem of Spectral Density Estimation, *The Annals of Statistics* **20**, p1985-2007.

Qin, J. and Lawless, J. (1994), Empirical likelihood and general estimation equations. *Annals of Statistics* 22, 300-325.

Robinson, P. (1989). Hypothesis testing in semiparametric and nonparametric models for econometric time series. *Review of Economic Studies*, **56**, 511–534.

Rosenblatt, M. (1956). "A central limit theorem and the α-mixing condition." *Proceedings of the National Academy of Sciences. U.S.A.*, **42**, 43-47.

Roussas G. G. and Ioannides D. (1980). Moment Inequalities for Mixing Sequences of Random Variables, *Stochastic Anal. Appl.* **5**, p61-120.

Tsao, Min (2004), Bounds on coverage probabilities of the empirical likelihood ratio confidence regions. *Ann. Statist.* 32, 1215–1221.

Yokoyama R. (1980). Moment Bounds for Stationary Mixing Sequences, *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **52**, p45-57.

Zhou, W. and Jing, B.-Y. (2003a). Smoothed empirical likelihood confidence intervals for the difference of quantiles. *Statist. Sinica* **13**, 83–95.

Zhou, W. and Jing, B.-Y. (2003b). Adjusted empirical likelihood method for quantiles. *Ann. Inst. Statist. Math.* **55** , 689–703.

Table 1. Average coverage levels and length (in parentheses) of the empirical likelihood confidence intervals for quantiles for the $AR(1)$ process.

(a) $N = 300, M = 12, l = 6$ for $AR(1)$ process.

q	Coverage	Unsmoothed	$h_1 = 1.50N^{-1/4}$	$h_2 = N^{-1/4}$	$h_3 = 0.50N^{-1/4}$
	0.90	0.923	0.889	0.888	0.889
		(0.516)	(0.417)	(0.430)	(0.437)
0.05	0.95	0.963	0.940	0.942	0.942
		(0.622)	(0.502)	(0.518)	(0.526)
	0.99	0.992	0.987	0.987	0.987
		(0.820)	(0.672)	(0.694)	(0.706)
	0.90	0.913	0.906	0.905	0.903
		(0.282)	(0.256)	(0.261)	(0.264)
0.50	0.95	0.952	0.948	0.947	0.947
		(0.337)	(0.308)	(0.314)	(0.317)
	0.99	0.985	0.982	0.982	0.982
		(0.450)	(0.413)	(0.421)	(0.426)

(b) $N = 500, M = 16, l = 8$ for $AR(1)$ process.

q	Coverage	Unsmoothed	$h_1 = 1.50N^{-1/4}$	$h_2 = N^{-1/4}$	$h_3 = 0.50N^{-1/4}$
	0.90	0.918	0.896	0.895	0.896
		(0.387)	(0.328)	(0.337)	(0.342)
0.05	0.95	0.963	0.944	0.947	0.948
		(0.461)	(0.394)	(0.404)	(0.411)
	0.99	0.994	0.988	0.990	0.990
		(0.614)	(0.527)	(0.541)	(0.549)
	0.90	0.910	0.904	0.903	0.904
		(0.217)	(0.200)	(0.204)	(0.205)
0.50	0.95	0.952	0.948	0.949	0.949
		(0.259)	(0.240)	(0.244)	(0.246)
	0.99	0.989	0.986	0.986	0.987
		(0.345)	(0.321)	(0.327)	(0.330)

Table 2. Average coverage levels and length (in parentheses) of the empirical likelihood confidence intervals for quantiles for the $AR(2)$ process.

(a) $N = 300, M = 12, l = 6$ for $AR(1)$ process.

q	Coverage	Unsmoothed	$h_1 = 1.50N^{-1/4}$	$h_2 = N^{-1/4}$	$h_3 = 0.50N^{-1/4}$
0.05	0.95	0.963	0.941	0.942	0.942
		(0.781)	(0.642)	(0.658)	(0.667)
	0.99	0.992	0.987	0.988	0.987
		(1.029)	(0.860)	(0.883)	(0.895)
0.50	0.95	0.952	0.948	0.947	0.947
		(0.423)	(0.391)	(0.397)	(0.401)
	0.99	0.985	0.982	0.982	0.983
		(0.565)	(0.525)	(0.533)	(0.538)

(b) $N = 500, M = 16, l = 8$ for $AR(1)$ process.

q	Coverage	Unsmoothed	$h_1 = 1.50N^{-1/4}$	$h_2 = N^{-1/4}$	$h_3 = 0.50N^{-1/4}$
0.05	0.95	0.963	0.946	0.947	0.948
		(0.579)	(0.502)	(0.514)	(0.520)
	0.99	0.994	0.989	0.990	0.990
		(0.768)	(0.672)	(0.687)	(0.695)
0.50	0.95	0.952	0.949	0.949	0.949
		(0.325)	(0.305)	(0.309)	(0.311)
	0.99	0.989	0.987	0.987	0.987
		(0.434)	(0.408)	(0.414)	(0.417)