Ann. Inst. Statist. Math.
Vol. 55, No. 4, 689–703 (2003)

ADJUSTED EMPIRICAL LIKELIHOOD METHOD FOR QUANTILES

WANG ZHOU AND BING-YI JING

Department of Mathematics, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China, e-mail: wangzhou@ust.hk; majing@ust.hk

(Received June 21, 2002; revised November 28, 2002)

Abstract. Empirical likelihood (EL) was first applied to quantiles by Chen and Hall (1993, *Ann. Statist.*, **21**, 1166–1181). In this paper, we shall propose an alternative EL approach which is also some kind of the kernel method. It not only eliminates the need to solve nonlinear equations, but also is extremely easy to implement. Confidence intervals derived from the proposed approach are shown, by an nonparametric version of Wilks' theorem, to have the same order of coverage accuracy (order $1/n$) as those of Chen and Hall. Numerical results are presented to compare our method with other methods.

Key words and phrases: Confidence interval, empirical likelihood, quantile, Edgeworth expansion.

1. Introduction

Let $X_1, X_2, \dots, X_n$ be a random sample from the unknown distribution $F(x)$ with density $f(x)$. Given $0 < q < 1$, we define the q-th quantile by $F^{-1}(q) = \inf\{x : F(x) \geq q\}$. In this paper, we will construct the confidence interval for $\theta_0 = F^{-1}(q)$.

Quantile is an important population characteristic. In some instances the quantile approach is feasible and useful when other approaches are out of the question. For example, to estimate the parameter of a Cauchy distribution, with density $f(x) = 1/\pi[1+(x-\mu)^2]$, $-\infty < x < \infty$, the sample mean $\overline{X}$ is not a consistent estimate of the location parameter μ. However, the sample median $\theta_{1/2}$ is $AN(\mu, \pi^2/4n)$ and thus quite well-behaved.

Let $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ be the order statistics of the X_i's. As we know, the common estimator of θ_0 is the sample quantile

$$\hat{\theta}_0 = X_{[nq]:n},$$

where $[nq]$ is the integer part of nq. The disadvantage of the sample quantile is its deficiency. See Falk (1984). In order to improve its efficiency, Sheather and Marron (1990) proposed the kernel quantile estimators

$$KQ_q = \sum_{i=1}^{n} \int_{(i-1)/n}^{i/n} K_h(t-q)dtX_{i:n},$$

where $K_h(\cdot) = h^{-1}K(\cdot/h)$ with K a density function symmetric about 0. The asymptotic normality of KQ_q was established by Yang (1985). Therefore, one could use the kernel quantile estimator KQ_q (properly studentized) to construct a confidence interval for the population quantile θ_0. However, the coverage of the kernel quantile estimator based confidence interval is not very accurate as our simulation results show.

In the nonparametric case empirical likelihood methods are powerful techniques for constructing confidence intervals and tests, notably in enabling the shape of a confidence region determined by the sample data. Since Owen (1988, 1990) introduced empirical likelihood into statistics, many developments have taken place. For a review, see Owen (2001). When applied to quantiles, Owen's method yields the so-called binomial method confidence intervals. However, because of the discreteness of the binomial distribution, the size of coverage error is of order $O(n^{-1/2})$. In order to improve the accuracy, it is natural to use smoothing methods. Chen and Hall (1993) first applied the method of smoothed empirical likelihood to sample quantiles and obtained very accurate results. But their method involves solving a system of nonlinear equations. Adimari (1998) presented a new version of the empirical log-likelihood ratio function for the quantiles which also yielded good results. For other works on application of EL for smoothing problems, we refer to Chen (1996, 1997).

The rest of this paper is arranged as follows. In Section 2, we propose a new version of the empirical likelihood method to quantiles. Based on the asymptotic chisquare distribution of the log-empirical likelihood ratio, we construct confidence intervals for the quantiles. We do some simulations in Section 3 in order to compare all kinds of methods numerically. The proof is deferred to Section 4.

2. Methodology and main results

We know that θ_0 coincides with the M-estimates defined by the equation

$$\int_{-\infty}^{\infty} \phi(x-\theta)dF(x) = 0, \tag{2.1}$$

with

$$\phi(z) = \begin{cases} -1 & \text{if} \quad z \le 0, \\ q/(1-q) & \text{if} \quad z > 0. \end{cases}$$

The empirical likelihood ratio for θ_0 is

$$R(\theta_0) = \sup_{p_1,\ldots,p_n} \prod_{i=1}^{n} (np_i), \tag{2.2}$$

subject to

$$p_i \ge 0, \quad i = 1,\ldots,n, \quad \sum_{i=1}^{n} p_i = 1, \quad \sum_{i=1}^{n} p_i\phi(X_i - \theta_0) = 0. \tag{2.3}$$

From (2.2), we have

$$\log R(\theta_0) = \sup_{p_1,\ldots,p_n} \sum_{i=1}^{n} \log p_i + n\log n, \tag{2.4}$$

where p_i, $i = 1,\ldots,n$ satisfy the constraint (2.3). The method of Lagrange multiplier may be used to maximize $\sum_{i=1}^{n} \log p_i$ subject to the constraint (2.3). Arguing this, we may prove that the maximization point occurs with

$$p_i = \frac{1}{n}\frac{1}{1+\lambda(\theta_0)\phi(X_i-\theta_0)}, \qquad i=1,\ldots,n, \tag{2.5}$$

where $\lambda(\theta_0)$ satisfies the equation

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\phi(X_i-\theta_0)}{1+\lambda(\theta_0)\phi(X_i-\theta_0)} = 0. \tag{2.6}$$

When $\theta_0 \in [X_{1:n}, X_{n:n}]$, the solution $\lambda(\theta_0)$ of (2.6) is

$$\lambda(\theta_0) = (q - F_n(\theta_0))/q,$$

where F_n is the empirical distribution function given by $F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I\{X_i \le x\}$, with $I\{\cdot\}$ being the indicator function. Hence the log-empirical likelihood ratio is

$$\begin{aligned} l(\theta_0) &= -2\log R(\theta_0) \\ &= 2n\left(F_n(\theta_0)\log\frac{F_n(\theta_0)}{q} + (1-F_n(\theta_0))\log\frac{1-F_n(\theta_0)}{1-q}\right). \end{aligned} \tag{2.7}$$

The above formula also appeared in Adimari (1998). From (2.7) $\log R(\theta_0)$ is a step function with jumps at the observed values. The fact that $\log R(\theta_0)$ can take only a finite number values makes the χ_1^2 approximation not very accurate. It is reasonable to replace $F_n(\cdot)$ by some smoothed version of the empirical distribution.

Let the bandwidth be $h = h_n > 0$. Suppose $h \to 0$ as $n \to \infty$. Choose some Borel measurable function $K(x)$ as the kernel and define $G(t) = \int_{-\infty}^{t} K(x)dx$. So the smoothed empirical distribution is

$$\hat{F}_n(x) = \frac{1}{n}\sum_{i=1}^{n} G\left(\frac{x-X_i}{h}\right). \tag{2.8}$$

We propose to use $\hat{F}_n(\theta_0)$ in (2.7) instead of $F_n(\theta_0)$. Thus the adjusted log-empirical likelihood ratio is

$$\hat{l}(\theta_0) = 2n\left(\hat{F}_n(\theta_0)\log\frac{\hat{F}_n(\theta_0)}{q} + (1-\hat{F}_n(\theta_0))\log\frac{1-\hat{F}_n(\theta_0)}{1-q}\right). \tag{2.9}$$

Clearly, our $\hat{l}(\theta_0)$ has a closed form, while Chen and Hall's version is implicitly determined by a nonlinear equation. In order to state our theorem, we give some regularity conditions.

(i) Let $f(x) = F'(x)$. For some integer $r \ge 2$, f and $f^{(r-1)}$ exist in a neighborhood of θ_0, and are continuous at θ_0. Also $f(\theta_0) > 0$.

(ii) The kernel $K(\cdot)$ is bounded and has a compact support $[a,b]$. For some decomposition, $a = u_0 < u_1 < \cdots < u_m = b$, $K(\cdot)$ is either strictly positive or strictly negative on each interval (u_{j-1}, u_j), where $j = 1,\ldots,m$. Also suppose

$$\int uK(u)G(u)du = 0,$$

$$\int u^j K(u)du = \begin{cases} 1, & j=0, \\ 0, & 1 \le j \le r_1 - 1, \\ C, & j = r_1, \end{cases}$$

where C is some finite constant and r_1 is some integer greater than or equal to r.

(iii) $nh/\log n \to \infty$, nh^2 is bounded as $n \to \infty$.

Let us give some remarks about the conditions. The first requires that the distribution function F be sufficiently smooth in a neighborhood of θ_0. That $f(\theta_0) > 0$ guarantees the asymptotic variance of the sample quantile is of order n^{-1}. Without that assumption, the asymptotic theory is quite different. We refer to Feldman and Tucker (1966). The second condition specifies that $K(\cdot)$ is different from the commonly used kernel in nonparametric density estimation. For example, $K(u) = \{\frac{21-9\sqrt{21}}{8}u^2 + \frac{-3+3\sqrt{21}}{8}\}I(|u| \leq 1)$ satisfies condition (ii) with $r_1 = 2$. Finally, condition (iii) implies that the bandwidth h does not converge to zero too fast or too slowly.

THEOREM 2.1. *Let $\hat{l}(\theta_0)$ be defined by* (2.9). *Assume that conditions* (i)–(iii) *hold. Then we have as $n \to \infty$,*

$$P(\hat{l}(\theta_0) \leq x) - P(\chi_1^2 \leq x) = O(n^{-1})$$

for each fixed x.

We postpone the proof of Theorem 2.1 to Section 4.

By Theorem 2.1, we have

$$\lim_{n\to\infty} P\{\theta_0 \in I_{hc}\} = P(\chi_1^2 \leq c),$$

where $I_{hc} = \{\theta : \hat{l}(\theta) \leq c\}$. If c is chosen to satisfy $P(\chi_1^2 \leq c) = \alpha$, then the coverage probability of the interval I_{hc} will approximate α with error $O(n^{-1})$ as $n \to \infty$. Also note that our result is not uniform in x.

3. Simulation results

A Monte Carlo study was conducted to investigate the coverage accuracy of empirical likelihood confidence interval. We generated 10,000 pseudo random samples of various sizes from standard normal, exponential and chi-square with degree 1 respectively. The kernel function K has been chosen to be $K(u) = \{\frac{21-9\sqrt{21}}{8}u^2+\frac{-3+3\sqrt{21}}{8}\}I(|u| \leq 1)$, which satisfies condition (ii) with $r = 2$. And we employed bandwidths $h = n^{-1/2}, n^{-3/4}, n^{-1}$, where $h = n^{-1/2}$ satisfies condition (iii).

Also shown in Tables 1–6 are smoothed empirical likelihood method (SELM) proposed by Chen and Hall (1993) and Bartlett adjusted SELM (BSELM) respectively. In this case, we use $K(u) = \frac{15}{16}(1-u^2)^2 I(|u| \leq 1)$, and $h = n^{-3/4}$. Adimari's result (1998) is included too for comparison.

Confidence intervals by the normal approximation method, denoted by S.M. in Tables 1–6, can be obtained as follows. From Yang (1985), we know

$$n^{1/2}(KQ_q - \theta_0) \to_L N(0, \sigma^2),$$

where

$$\sigma^2 = \frac{p(1-p)}{f^2(\theta_0)}. \tag{3.1}$$

Table 1.

$q = 0.5$. $(N(0,1))$					
	nominal level	0.8	0.9	0.95	0.99
$n = 20$	$h = n^{-0.5}$	0.789	0.894	0.949	0.986
	$h = n^{-0.75}$	0.773	0.891	0.953	0.986
	$h = n^{-1}$	0.756	0.888	0.955	0.985
	SELM	0.769	0.886	0.948	0.984
	BSELM	0.773	0.888	0.950	0.985
	Adimari	0.808	0.913	0.925	0.989
	S.M.	0.843	0.927	0.967	0.994
$n = 50$	$h = n^{-0.5}$	0.802	0.897	0.951	0.989
	$h = n^{-0.75}$	0.801	0.892	0.946	0.989
	$h = n^{-1}$	0.799	0.887	0.942	0.990
	SELM	0.796	0.892	0.947	0.990
	BSELM	0.798	0.894	0.947	0.990
	Adimari	0.794	0.905	0.950	0.988
	S.M.	0.827	0.921	0.962	0.991
$n = 100$	$h = n^{-0.5}$	0.802	0.901	0.950	0.989
	$h = n^{-0.75}$	0.804	0.903	0.951	0.987
	$h = n^{-1}$	0.807	0.909	0.949	0.987
	SELM	0.801	0.902	0.950	0.988
	BSELM	0.802	0.902	0.950	0.988
	Adimari	0.763	0.886	0.951	0.991
	S.M.	0.819	0.912	0.959	0.992

A consistent estimator of σ^2, $\hat{\sigma}^2$, can be obtained by replacing $f(\theta_0)$ in the above formula by their empirical versions (or smoothed ones when appropriate). Thus, a two-sided confidence interval based on the normal approximations can be taken to be

$$I_{1-\alpha}^{(N)} = (KQ_q - z_{1-\alpha/2}\hat{\sigma}/\sqrt{n}, KQ_q + z_{1-\alpha/2}\hat{\sigma}/\sqrt{n}),$$

where $z_{1-\alpha}$ is the $1-\alpha$ quantile of a standard normal distribution. For simplicity, in our simulation studies conducted here, we employ the true value of σ^2 rather than its consistent estimator $\hat{\sigma}^2$. We did this since it is reasonable to expect that the latter will not usually outperform the former. We also use the biquadratic kernel in the calculation of KQ_q. The bandwidth is selected to be $h = n^{-1/2}$.

We make the following observations from the numerical studies.

(1) Our method is more accurate almost everywhere than SELM. And it is very competitive with BSELM and Admirani's result, even in small samples.

(2) The new method is robust with respect to bandwidth selections.

(3) Compared with SELM, the adjusted empirical method is very efficient in the computation. For example, when $n = 20$, $F = N(0,1)$, the time needed to get the first four numbers $(0.7895, 0.8941, 0.9496, 0.9869)$ is less than 10 seconds when we run C-program in Sun microsystems. But the time needed to get the corresponding four numbers of SELM is more than 4 minutes.

(4) In many cases, we find that BSELM is not better than SELM. Even

Table 2.

	$q = 0.5$. (e^{-x})				
	nominal level	0.8	0.9	0.95	0.99
$n = 20$	$h = n^{-0.5}$	0.792	0.898	0.948	0.987
	$h = n^{-0.75}$	0.778	0.894	0.953	0.986
	$h = n^{-1}$	0.759	0.890	0.957	0.986
	SELM	0.776	0.890	0.950	0.985
	BSELM	0.782	0.893	0.950	0.986
	Adimari	0.813	0.892	0.948	0.986
	S.M.	0.841	0.921	0.955	0.985
$n = 50$	$h = n^{-0.5}$	0.798	0.901	0.952	0.990
	$h = n^{-0.75}$	0.796	0.895	0.949	0.990
	$h = n^{-1}$	0.795	0.889	0.945	0.991
	SELM	0.793	0.895	0.950	0.991
	BSELM	0.794	0.896	0.951	0.991
	S.M.	0.832	0.923	0.962	0.989
	Adimari	0.794	0.911	0.952	0.989
	S.M.	0.832	0.923	0.962	0.989
$n = 100$	$h = n^{-0.5}$	0.803	0.902	0.949	0.988
	$h = n^{-0.75}$	0.807	0.903	0.947	0.988
	$h = n^{-1}$	0.811	0.908	0.943	0.987
	SELM	0.805	0.902	0.948	0.989
	BSELM	0.806	0.903	0.949	0.989
	Adimari	0.810	0.890	0.952	0.989
	S.M.	0.825	0.915	0.959	0.991

in some cases where BSELM is better than SELM, the improvements are very marginal. If we examine the simulation results in Chen and Hall (1993) carefully, we can see that Bartlett adjustment improves the order of accuracy insignificantly. As pointed out by one referee, this is because when h is small the Bartlett factor is small and the entire correction factor is very close to one. It is a quite common phenomenon in smoothing, as observed in Chen and Qin (2000). So Bartlett correction for smoothing problem may not be very effective and its usefulness is mainly a theoretical nature as it indicates there is something delicate going on with the EL which resembles a parametric likelihood.

(5) For small sample size n, the kernel quantile estimator method seems to perform worse than the smoothed empirical likelihood methods. However, the advantage of the latter method gradually disappears when the sample sizes get large, as one might expect.

We also do some simulation results for $q = 0.1$, 0.25. The conclusions are the same.

4. Proof of Theorem 2.1

Let us first introduce a few lemmas, which will be useful in proving the main theorems.

Table 3.

$q = 0.75$. $(N(0,1))$					
	nominal level	0.8	0.9	0.95	0.99
$n = 20$	$h = n^{-0.5}$	0.797	0.885	0.947	0.988
	$h = n^{-0.75}$	0.802	0.881	0.948	0.991
	$h = n^{-1}$	0.804	0.873	0.954	0.991
	SELM	0.788	0.877	0.953	0.994
	BSELM	0.793	0.888	0.955	0.994
	Adimari	0.791	0.875	0.958	0.989
	S.M.	0.844	0.931	0.967	0.993
$n = 50$	$h = n^{-0.5}$	0.800	0.901	0.949	0.988
	$h = n^{-0.75}$	0.806	0.902	0.949	0.987
	$h = n^{-1}$	0.809	0.901	0.950	0.987
	SELM	0.802	0.902	0.950	0.987
	BSELM	0.804	0.903	0.951	0.987
	Adimari	0.771	0.901	0.941	0.989
	S.M.	0.834	0.922	0.961	0.992
$n = 100$	$h = n^{-0.5}$	0.798	0.898	0.948	0.989
	$h = n^{-0.75}$	0.798	0.898	0.949	0.989
	$h = n^{-1}$	0.796	0.896	0.950	0.989
	SELM	0.793	0.895	0.949	0.989
	BSELM	0.794	0.896	0.950	0.989
	Adimari	0.815	0.891	0.943	0.989
	S.M.	0.825	0.914	0.960	0.993

LEMMA 4.1. *Under conditions* (i)–(iii), *we have* $\forall \epsilon > 0$,

$$\sup_{|t|>\epsilon} \left| \int_{-\infty}^{\infty} \exp\{itG(u)\} d_u F(\theta_0 - hu) \right| \leq 1 - C(\epsilon)h$$

for all sufficiently small h, *where* $C(\epsilon) > 0$ *and* d_u *means the integrand is a function of* u.

PROOF. Observe that

$$\begin{aligned} I(t) &:= \int_{-\infty}^{\infty} \exp\{itG(u)\} d_u F(\theta_0 - hu) \\ &= [1 - F(\theta_0 - ha)]e^{it} + F(\theta_0 - hb) + \int_a^b \exp\{itG(u)\} d_u F(\theta_0 - hu) \\ &= [1 - F(\theta_0 - ha)]\cos t + F(\theta_0 - hb) + i[1 - F(\theta_0 - ha)]\sin t \\ &\quad - hf(\theta_0) \int_a^b \exp\{itG(u)\} du \\ &\quad - h \int_a^b \exp\{itG(u)\}[f(\theta_0 - hu) - f(\theta_0)] du. \end{aligned} \tag{4.1}$$

Table 4.

$q = 0.75$. (e^{-x})					
	nominal level	0.8	0.9	0.95	0.99
$n = 20$	$h = n^{-0.5}$	0.801	0.885	0.950	0.990
	$h = n^{-0.75}$	0.803	0.880	0.951	0.991
	$h = n^{-1}$	0.803	0.874	0.957	0.993
	SELM	0.786	0.874	0.952	0.994
	BSELM	0.789	0.883	0.954	0.994
	Adimari	0.804	0.892	0.962	0.992
	S.M.	0.842	0.916	0.954	0.985
$n = 50$	$h = n^{-0.5}$	0.804	0.902	0.950	0.990
	$h = n^{-0.75}$	0.809	0.902	0.951	0.989
	$h = n^{-1}$	0.812	0.903	0.952	0.989
	SELM	0.807	0.903	0.953	0.990
	BSELM	0.808	0.904	0.954	0.990
	Adimari	0.807	0.903	0.946	0.991
	S.M.	0.830	0.922	0.961	0.990
$n = 100$	$h = n^{-0.5}$	0.806	0.902	0.951	0.989
	$h = n^{-0.75}$	0.806	0.901	0.951	0.989
	$h = n^{-1}$	0.805	0.898	0.952	0.989
	SELM	0.803	0.898	0.952	0.989
	BSELM	0.803	0.900	0.952	0.989
	Adimari	0.806	0.896	0.948	0.989
	S.M.	0.828	0.917	0.959	0.991

We shall show shortly that the conditions imposed on K imply that for each $\epsilon > 0$, there exists $\epsilon' > 0$ such that

$$\sup_{|t|>\epsilon} \left| (b-a)^{-1} \int_a^b \exp\{itG(u)\} du \right| \le 1 - 3\epsilon'. \tag{4.2}$$

Note that for h sufficiently small,

$$\int_a^b |f(\theta_0 - hu) - f(\theta_0)| du \le (b-a) f(\theta_0)\epsilon', \tag{4.3}$$

$$\begin{aligned} &|[1 - F(\theta_0 - ha)]\cos t + F(\theta_0 - hb) + i[1 - F(\theta_0 - ha)]\sin t| \\ &\quad \le [1 - F(\theta_0 - ha)] + F(\theta_0 - hb) \\ &\quad = 1 - h\int_a^b f(\theta_0 - hu) du \\ &\quad \le 1 - h(b-a) f(\theta_0)(1 - \epsilon'). \end{aligned} \tag{4.4}$$

Combining (4.1)–(4.4) gives

$$\begin{aligned} |I(t)| &\le 1 - h(b-a)f(\theta_0)(1-\epsilon') + h(b-a)f(\theta_0)(1-3\epsilon') + h(b-a)f(\theta_0)\epsilon' \\ &= 1 - h(b-a)f(\theta_0)\epsilon'. \end{aligned}$$

Table 5.

	$q = 0.9$. $(N(0,1))$				
	nominal level	0.8	0.9	0.95	0.99
$n = 20$	$h = n^{-0.5}$	0.777	0.851	0.886	0.908
	$h = n^{-0.75}$	0.769	0.842	0.873	0.893
	$h = n^{-1}$	0.760	0.838	0.868	0.883
	SELM	0.872	0.954	0.995	1.000
	BSELM	0.880	0.957	0.997	1.000
	Adimari	0.888	0.956	0.983	0.996
	S.M.	0.834	0.929	0.972	0.997
$n = 50$	$h = n^{-0.5}$	0.787	0.889	0.940	0.988
	$h = n^{-0.75}$	0.780	0.894	0.938	0.989
	$h = n^{-1}$	0.771	0.896	0.937	0.989
	SELM	0.778	0.905	0.941	0.994
	BSELM	0.783	0.908	0.944	0.994
	Adimari	0.798	0.895	0.940	0.991
	S.M.	0.855	0.941	0.975	0.995
$n = 100$	$h = n^{-0.5}$	0.792	0.897	0.949	0.988
	$h = n^{-0.75}$	0.793	0.900	0.951	0.987
	$h = n^{-1}$	0.800	0.900	0.953	0.987
	SELM	0.796	0.896	0.950	0.987
	BSELM	0.799	0.898	0.950	0.987
	Adimari	0.821	0.902	0.934	0.992
	S.M.	0.831	0.922	0.961	0.993

Finally we check (4.2). Observe that

$$\int_a^b \exp\{itG(u)\}du = \sum_{j=1}^{m} \int_{G(u_{j-1})}^{G(u_j)} \exp\{itu\}dG^{-1}(u),$$

where $G^{-1}(u)$ is the inverse function of G, which is uniquely determined in each subinterval (u_{j-1}, u_j).

Define $L(u) = dG^{-1}(u)/du = 1/K(G^{-1}(u))$. So $| \int_{G(u_{j-1})}^{G(u_j)} |L(u)||du| = u_j - u_{j-1}$. This implies that L is integrable on the interval $I_j = (a_j, b_j)$ whose endpoints are $G(u_j)$ and $G(u_{j-1})$. Hence (4.2) follows from Riemann-Lebesgue's lemma.

Define $w_j = w_j(X_j) = G(\frac{\theta_0 - X_j}{h}) - q$, $Y_j = w_j - Ew_j$, $j = 1, \ldots, n$, $\sigma^2 = \operatorname{Var} Y_1 = q(1-q) + O(h)$, $\rho_r = E|Y_1|^r$, $\chi_v = v$-th cumulant of Y_1, $\eta_r = E|\frac{Y_1}{\sigma}|^r$.

LEMMA 4.2. *Let $s \geq 3$, there exist two positive constants $C_1(s)$, $C_2(s)$ depending only on s, such that for all t in R satisfying*

$$|t| \leq C_1(s) \frac{n^{1/2}}{\eta_s^{1/(s-2)}},$$

Table 6.

	$q = 0.9.\ (e^{-x})$				
	nominal level	0.8	0.9	0.95	0.99
$n = 20$	$h = n^{-0.5}$	0.776	0.853	0.883	0.903
	$h = n^{-0.75}$	0.765	0.847	0.876	0.892
	$h = n^{-1}$	0.757	0.845	0.875	0.886
	SELM	0.863	0.959	0.997	1.000
	BSELM	0.870	0.962	0.998	1.000
	Adimari	0.860	0.958	0.976	0.996
	S.M.	0.864	0.960	0.990	0.999
$n = 50$	$h = n^{-0.5}$	0.795	0.903	0.949	0.992
	$h = n^{-0.75}$	0.784	0.908	0.949	0.993
	$h = n^{-1}$	0.776	0.912	0.948	0.993
	SELM	0.783	0.916	0.951	0.996
	BSELM	0.786	0.917	0.953	0.997
	Adimari	0.791	0.910	0.951	0.994
	S.M.	0.836	0.919	0.961	0.991
$n = 100$	$h = n^{-0.5}$	0.793	0.898	0.951	0.989
	$h = n^{-0.75}$	0.799	0.900	0.954	0.988
	$h = n^{-1}$	0.803	0.901	0.954	0.988
	SELM	0.799	0.896	0.952	0.988
	BSELM	0.800	0.897	0.953	0.988
	Adimari	0.802	0.902	0.955	0.988
	S.M.	0.818	0.910	0.949	0.986

one has for all α, $0 \le \alpha \le s$,

$$\left| D^{\alpha} \left[\prod_{j=1}^{n} E\left(\exp\left\{ \frac{it}{\sigma\sqrt{n}} Y_j \right\} \right) - e^{-t^2/2} \sum_{r=0}^{s-3} n^{-r/2} \widetilde{P}_r(i\sigma^{-1}t; \{\chi_v\}) \right] \right|$$
$$\le \frac{C_2(s)}{n^{(s-2)/2}} \eta_s [|t|^{s-\alpha} + |t|^{3(s-2)+\alpha}] e^{-t^2/4},$$

where $\widetilde{P}_1(z; \{\chi_v\}) = \frac{\chi_3}{3!} z^3$, $\widetilde{P}_2(z; \{\chi_v\}) = \frac{\chi_4}{4!} z^4 + \frac{\chi_3^2}{2 \times (3!)^2} z^6$, $\widetilde{P}_3(z; \{\chi_v\}) = \frac{\chi_5}{5!} z^5 + \frac{\chi_3 \chi_4}{3!4!} z^7 + \frac{\chi_3^3}{(3!)^4} z^9$. *For the general definition of* $\widetilde{P}_r$, *see Bhattacharya and Rao* (1976).

This is Theorem 9.10 of Bhattacharya and Rao (1976). And we follow the conventions of Bhattacharya and Rao (1976). In the following, ϕ_{0,σ^2} will denote the density function of the normal random variable with mean 0 and variance σ^2. $\Phi(x)$ and $\phi(x)$ mean the standard normal distribution and density respectively.

LEMMA 4.3. *Let $Z_1, \ldots, Z_n$ be a sequence of standard normal random variables independent of $X_1, X_2, \ldots, X_n$, and let q_n denote the probability density of $\sqrt{n}\bar{Y} + n^{-c}\sqrt{n}\bar{Z}$, where $\bar{Y} = \frac{1}{n}\sum_{j=1}^{n} Y_j$ and $\bar{Z} = \frac{1}{n}\sum_{j=1}^{n} Z_j$. Then for each pair of positive integers (α, s),*

$\exists c_0(\alpha, s) > 0$ *such that for all* $c \geq c_0(\alpha, s)$,

$$\sup_{x \in R}(1+|x|^\alpha)\left|q_n(x) - \sum_{j=0}^{s-3} n^{-j/2}P_j(-\phi_{0,\sigma^2};\{\chi_v\})(x)\right| = O(n^{-(s-2)/2}),$$

as $n \to \infty$, *where* $P_j(-\phi_{0,\sigma^2};\{\chi_v\}) = \widetilde{P}_j(-D;\{\chi_v\})\phi_{0,\sigma^2}$.

PROOF. Let $\hat{Q}_n(t) = Ee^{it(\sqrt{n}\bar{Y}+n^{-c}\sqrt{n}\bar{Z})}$. For each n, $D_\alpha\hat{Q}_n$, α-th derivative of $\hat{Q}_n$, is integrable for $0 \leq |\alpha| \leq s$. Writing for $|\alpha| \leq s$,

$$h_n(x) = x^\alpha\left[q_n(x) - \sum_{j=0}^{s-3} n^{-j/2}P_j(-\phi_{0,\widetilde{\sigma}^2};\{\widetilde{\chi}_v\})(x)\right],$$

$$\widetilde{h}_n(t) = (-i)^\alpha D^\alpha\left[\hat{Q}_n(t) - \sum_{j=0}^{s-3} n^{-j/2}\widetilde{P}_j(it;\{\widetilde{\chi}_v\})\exp\left(-\frac{\widetilde{\sigma}^2}{2}t^2\right)\right],$$

where $P_j(-\phi_{0,\widetilde{\sigma}^2};\{\widetilde{\chi}_v\}) = \widetilde{P}_j(-D;\{\widetilde{\chi}_v\})\phi_{0,\widetilde{\sigma}^2}$, $\widetilde{\sigma}^2 = \mathrm{Var}(Y_1+n^{-c}Z_1) = \sigma^2+n^{-2c}$, $\widetilde{\chi}_v = v$-th cumulant of $Y_1+n^{-c}Z_1$, one has (by the Fourier inversion theorem)

$$h_n(x) = (2\pi)^{-1}\int e^{-itx}\widetilde{h}_n(t)dt. \tag{4.5}$$

Define $\widetilde{\eta}_s = E|\widetilde{\sigma}^{-1}(Y_1+n^{-c}Z_1)|^s$. By Lemma 4.2

$$|\widetilde{h}_n(t)| \leq \frac{C_2(s)}{n^{(s-2)/2}}\widetilde{\eta}_s[(\widetilde{\sigma}^{-1}t)^{s-\alpha} + (\widetilde{\sigma}^{-1}t)^{3(s-2)+|\alpha|}]e^{-(1/4)\widetilde{\sigma}^{-1}t^2} \tag{4.6}$$

for all t satisfying

$$|t| \leq \widetilde{\sigma}^{-1}C_1(s)n^{1/2}\widetilde{\eta}_s^{-1/(s-2)} =: a_n n^{1/2}. \tag{4.7}$$

Since $\widetilde{\sigma}^2 = \sigma^2 + n^{-2c} = q(1-q) + n^{-2c} + O(h)$, $\widetilde{\eta}_s = \widetilde{\sigma}^{-s}E|Y_1+n^{-c}Z_1|^s$, we have that if $c > 0$, a_n is bounded away from 0 as $n \to \infty$. Hence we can suppose if $n \geq n_0$ and $c > 0$, $a_n \geq a_0 > 0$, where a_0 is some constant. In view of (4.5), (4.6) and (4.7), it is enough to prove

$$\int_{(|t|>a_0n^{1/2})} |D^\alpha\hat{Q}_n(t)|dt = O(n^{-(s-2)/2}), \tag{4.8}$$

$$\int_{(|t|>a_0n^{1/2})}\left|D^\alpha\left[\sum_{j=0}^{s-3} n^{-j/2}\widetilde{P}_j(it;\{\widetilde{\chi}_v\})e^{-(1/2)\widetilde{\sigma}^2t^2}\right]\right|dt = O(n^{-(s-2)/2}). \tag{4.9}$$

(4.9) is true because of the presence of the exponential term and $\widetilde{\sigma}^2 \to q(1-q)$ as $n \to \infty$.

Simple calculations show that for $|t| > a_0n^{1/2}$,

$$\begin{aligned} &|D^\alpha\hat{Q}_n(t)| \\ &\quad \leq C_3(s)(\widetilde{\rho}_\alpha + \widetilde{\rho}_{\alpha-1} + \cdots + \widetilde{\rho}_1 + 1)n^{\alpha/2}\delta^{n-\alpha-1}|Ee^{it/\sqrt{n}(Y_1+n^{-c}Z_1)}|, \end{aligned} \tag{4.10}$$

where $\widetilde{\rho}_r = E|Y_1 + n^{-c}Z_1|^r$, $r = 1, \ldots, \alpha$, $\delta = \sup_{|t|>a_0 n^{1/2}} |Ee^{it/\sqrt{n}(Y_1+n^{-c}Z_1)}|$. By Lemma 4.1,

$$\delta \leq \sup_{|t|>a_0} |Ee^{itG((\theta_0-X_1)/h)}| \leq 1 - C(a_0)h.$$

Also we have

$$\begin{aligned}
&\int_{|t|>a_0 n^{1/2}} |Ee^{it/\sqrt{n}(Y_1+n^{-c}Z_1)}|dt \\
&\quad \leq \int_{|t|>a_0 n^{1/2}} |Ee^{it/\sqrt{n}n^{-c}Z_1}|dt \\
&\quad \leq \int_{|t|>a_0 n^{1/2}} e^{-(t^2/2)n^{-1-2c}} dt \\
&\quad \leq \sqrt{2\pi} n^{1/2+c}.
\end{aligned}$$

So the integral in (4.8) is dominated by

$$\begin{aligned}
&C_4(s)(\widetilde{\rho}_\alpha + \widetilde{\rho}_{\alpha-1} + \cdots + \widetilde{\rho}_1 + 1)n^{\alpha/2+1/2+c}(1 - C(a_0)h)^n \\
&\quad \leq C_4(s)(\widetilde{\rho}_\alpha + \widetilde{\rho}_{\alpha-1} + \cdots + \widetilde{\rho}_1 + 1)n^{\alpha/2+1/2+c} \exp(-C(a_0)nh) \\
&\quad = O(n^{-a_1})
\end{aligned}$$

for all $a_1 > 0$ since $nh/\log n \to \infty$. This completes the proof of (4.8). So we have

$$\sup_x |h_n(x)| = O(n^{-(s-2)/2}).$$

Since $\widetilde{\chi}_v = \chi_v + \chi'_v$, where χ'_v is the v-th cumulant of $n^{-c}Z_1$, we have

$$\begin{aligned}
&\widetilde{\chi}_v = \chi_v, \qquad v \geq 3, \\
&\widetilde{\sigma}^2 = \sigma^2 + n^{-2c}.
\end{aligned}$$

Thus if $c \geq \frac{s-2}{2}$, we have

$$\sup_x \left| (1 + x^\alpha) \left(q_n(x) - \sum_{j=0}^{s-3} n^{-j/2} P_j(-\phi_{0,\sigma^2}; \{\chi_v\})(x) \right) \right| = O(n^{-(s-2)/2}).$$

This completes the proof.

LEMMA 4.4. *Assume the conditions of Lemma 4.3. Let $\mathcal{B}$ be the class of Borel sets $B \subset R$ that satisfy*

$$\sup_{B\in\mathcal{B}} \int_{(\partial B)^\epsilon} e^{-(1/2)x^2} dx = O(\epsilon)$$

as $\epsilon \to 0$, then for each integer $s \geq 4$,

$$\sup_{B\in\mathcal{B}} \left| P(\sqrt{n}\bar{Y} \in B) - \int_B \sum_{j=0}^{s-3} n^{-j/2} p_j(-\phi_{0,\sigma^2}; \{\chi_v\})(x)dx \right| = O(n^{-(s-2)/2}).$$

PROOF. Taking $\alpha = 2$ in Lemma 4.3, we deduce that for all $c \geq c_0(2, s)$,

$$\sup_{B\in\mathcal{B}_0}\left|P(\sqrt{n}\bar{Y} + \sqrt{n}n^{-c}\bar{Z} \in B) - \int_B \sum_{j=0}^{s-3} n^{-j/2} P_j(-\phi_{0,\sigma^2}; \{\chi_v\})(x)dx\right| = O(n^{-(s-2)/2}), \tag{4.11}$$

where $\mathcal{B}_0$ denotes Borel σ-field. Put $\delta = \delta(n) = n^{-c/2}$. Now

$$\begin{aligned}
&|P(\sqrt{n}\bar{Y} \in B) - P(\sqrt{n}\bar{Y} + \sqrt{n}n^{-c}\bar{Z} \in B)| \\
&\quad \leq |P(\sqrt{n}\bar{Y} \in B, |\sqrt{n}n^{-c}\bar{Z}| \leq \delta) - P(\sqrt{n}\bar{Y} + \sqrt{n}n^{-c}\bar{Z} \in B, |\sqrt{n}n^{-c}\bar{Z}| \leq \delta)| \\
&\qquad + |P(\sqrt{n}\bar{Y} \in B, |\sqrt{n}n^{-c}\bar{Z}| > \delta) - P(\sqrt{n}\bar{Y} + \sqrt{n}n^{-c}\bar{Z} \in B, |\sqrt{n}n^{-c}\bar{Z}| > \delta)| \\
&\quad \leq P(\sqrt{n}\bar{Y} + \sqrt{n}n^{-c}\bar{Z} \in (\partial B)^{\delta}) + P(|\sqrt{n}\bar{Z}| > n^{c/2}).
\end{aligned}$$

We may deduce from (4.11) that if $c \geq s - 2$,

$$\sup_{B\in\mathcal{B}} P(\sqrt{n}\bar{Y} + \sqrt{n}n^{-c}\bar{Z} \in (\partial B)^{\delta}) = O(n^{-(s-2)/2}),$$

and of course, $P(|\sqrt{n}\bar{Z}| > n^{c/2}) = O(n^{-c})$ for all $c > 0$. Therefore

$$\sup_{B\in\mathcal{B}} |P(\sqrt{n}\bar{Y} \in B) - P(\sqrt{n}\bar{Y} + \sqrt{n}n^{-c}\bar{Z} \in B)| = O(n^{-(s-2)/2}).$$

Now, (4.11) implies the conclusion.

LEMMA 4.5. *Assume the conditions of Lemma* 4.3. *Then*

$$\begin{aligned}
P(\sqrt{n}\bar{Y}/\sigma \leq y) &= \Phi(y) - \frac{\chi_3}{6\sqrt{n}}(y^2 - 1)\phi(y) \\
&\quad - \left(\frac{\chi_4}{24n}y(y^2 - 3) + \frac{\chi_3^2}{72n}y(y^4 - 10y^2 + 15)\right)\phi(y) + O(n^{-3/2}).
\end{aligned}$$

PROOF. Letting $s = 5$ in Lemma 4.4 gives the result.

PROOF OF THEOREM 2.1. Bernstein inequality implies

$$\begin{aligned}
&P(|\hat{F}_n(\theta_0) - E\hat{F}_n(\theta_0)| \geq n^{-1/2+\epsilon}) \\
&\quad \leq 2\exp\left(-\frac{nn^{-1+2\epsilon}}{2\operatorname{Var} G\left(\dfrac{\theta_0 - X_1}{h}\right) + \dfrac{2}{3}mn^{-1/2+\epsilon}}\right),
\end{aligned}$$

where m satisfies $P(|G(\frac{\theta_0 - X_1}{h}) - EG(\frac{\theta_0 - X_1}{h})| \leq m) = 1$ and ϵ is some arbitrarily small positive number. Under conditions (i) and (ii),

$$\begin{aligned}
E\hat{F}_n(\theta_0) &= q + O(h^2) = q + O(n^{-1}), \\
\operatorname{Var} G\left(\frac{\theta_0 - X_1}{h}\right) &= q(1 - q) + O(h).
\end{aligned}$$

Hence after neglecting a set with probability $O(n^{-2})$, we have

$$\left|\frac{\hat{F}_n(\theta_0)}{q} - 1\right| \le 2n^{-1/2+\epsilon}.$$

Write $z = \frac{\hat{F}_n(\theta_0)}{q}$. Taylor expansion gives

$$\begin{aligned}\hat{l}(\theta_0) &= n\frac{q}{1-q}\left((1-z) + \frac{1}{2}\frac{-2q+q^2}{3(1-q)}(1-z)^2\right)^2 + O(n^{-1})\\ &= \left(R + \frac{1}{\sqrt{n}}\sqrt{\frac{q}{1-q}}\frac{2-q}{3}R^2\right)^2 + O(n^{-1}),\end{aligned}$$

where $R := \sqrt{n}\frac{\hat{F}_n(\theta_0)-q}{\sqrt{q(1-q)}} = \sqrt{n}\frac{z-1}{\sqrt{(1-q)/q}}$. Thus

$$P(\hat{l}(\theta_0) \le x) = P\left(\left(R + \frac{1}{\sqrt{n}}\sqrt{\frac{q}{1-q}}\frac{2-q}{3}R^2\right)^2 + O(n^{-1}) \le x\right) + O(n^{-2}). \tag{4.12}$$

By the condition $\int_{-\infty}^{\infty} uK(u)G(u)du = 0$, we have

$$\operatorname{Var} G\left(\frac{\theta_0 - X_1}{h}\right) = q(1-q) + O(h^2) = q(1-q) + O(n^{-1}).$$

Applying Lemma 4.5, after tedious calculations we have from (4.12)

$$\begin{aligned}P(\hat{l}(\theta_0) \le x) &= P\left(-\frac{\sqrt{x}}{\frac{1}{2} + \sqrt{\frac{1}{4} - \frac{a\sqrt{x}}{\sqrt{n}}}} \le R \le \frac{\sqrt{x}}{\sqrt{\frac{1}{4} + \frac{a\sqrt{x}}{\sqrt{n}}} + \frac{1}{2}}\right) + O(n^{-1})\\ &= \Phi(\sqrt{x}) - \Phi(-\sqrt{x}) + O(n^{-1})\\ &= P(\chi_1^2 \le x) + O(n^{-1}),\end{aligned}$$

where $a = \sqrt{\frac{q}{1-q}\frac{2-q}{3}}$.

Acknowledgements

The authors would like to thank the referees for their valuable suggestions and remarks.

References

Adimari, G. (1998). An empirical likelihood statistic for quantiles, *J. Statist. Comput. Simulation*, 60(1), 85–95.

Bhattacharya, R. N. and Rao, R. R. (1976). *Normal Approximation and Asymptotic Expansions*, Wiley, New York.

Chen, S. (1996). Empirical likelihood confidence intervals for nonparametric density estimation, *Biometrika*, **83**(2), 329–341.

Chen, S. (1997). Empirical likelihood-based kernel density estimation, *Austral. J. Statist.*, **39(1)**, 47–56.

Chen, S. and Hall, P. (1993). Smoothed empirical likelihood confidence intervals for quantiles, *Ann. Statist.*, **21**(3), 1166–1181.

Chen, S. and Qin, Y. (2000). Empirical likelihood confidence intervals for local linear smoothers, *Biometrika*, **87**(4), 946–953.

Falk, M. (1984). Relative deficiency of kernel type estimators of quantiles, *Ann. Statist.*, **12(1)**, 261–268.

Feldman, D. and Tucker, H. G. (1966). Estimation of non-unique quantiles, *Ann. Math. Statist.*, **37**(2), 451–457.

Owen, A. B. (1988). Empirical likelihood ratio confidence intervals for a single functional, *Biometrika*, **75**(2), 237–249.

Owen, A. B. (1990). Empirical likelihood ratio confidence regions, *Ann. Statist.*, **18**(1), 90–120.

Owen, A. B. (2001). *Empirical Likelihood*, Chapman and Hall, London.

Sheather, S. J. and Marron, J. S. (1990). Kernel quantile estimators, *J. Amer. Statist. Assoc.*, **85**(410), 410–416.

Yang, S. S. (1985). A smooth nonparametric estimator of a quantile function, *J. Amer. Statist. Assoc.*, **80**(392), 1004–1011.