Bernoulli **8**(2), 2002, 255–274

Estimation of the innovation quantile density function of an AR(p) process based on autoregression quantiles

FAOUZI EL BANTLI* and MARC HALLIN**

I.S.R.O., Université Libre de Bruxelles, B-1050 Bruxelles, Belgium.
E-mail: *faouzi.elbantli@fortisbank.com, **mhallin@ulb.ac.be

In this paper, we propose two types of estimator (one of histogram type, the other a kernel estimate) of the quantile density (or *sparsity*) function $\alpha \mapsto [f(F^{-1}(\alpha))]^{-1}$ associated with the innovation density f of an autoregressive model of order p. Our estimators are based on autoregression quantiles. Contrary to more classical estimators based on estimated residuals, they are autoregression-invariant and scale-equivariant. Their asymptotic behaviour is derived from a uniform Bahadur representation for autoregression quantiles – a result of independent interest. Simulations are carried out to illustrate their performance.

Keywords: autoregression; autoregression quantiles; Bahadur–Kiefer representation; histogram estimator; kernel estimator; quantile density function; sparsity function

1. Introduction

Denote by $\mathbf{Y} := (Y_1, \ldots, Y_n)^{\mathrm{T}}$ an observed realization of length n of some solution of the autoregressive (AR) model of order p,

$$Y_t = \varrho_1 Y_{t-1} + \ldots + \varrho_p Y_{t-p} + \varepsilon_t, \qquad t \in \mathbb{Z}, \tag{1.1}$$

where p is a fixed integer, $\boldsymbol{\varrho} = (\varrho_1, \ldots, \varrho_p)^{\mathrm{T}}$ is a vector of unknown parameters, and $\{\varepsilon_t, t \in \mathbb{Z}\}$ a sequence of independent and identically distributed (i.i.d.) random variables with distribution function F and density f. We do not assume that f is known; we just require that it belongs to the family $\mathcal{F}$ of non-vanishing densities satisfying

$$\int x \mathrm{d}F(x) = 0, \qquad 0 < \sigma_F^2 := \int x^2 \mathrm{d}F(x) < \infty. \tag{1.2}$$

As usual, the autoregressive parameter $\boldsymbol{\varrho}$ is supposed to be such that the polynomial

$$\varrho(z) := 1 - \sum_{i=1}^{p} \varrho_i z^{p-i}, \qquad z \in \mathbb{C}, \tag{1.3}$$

has no root within the unit disc (the usual *causality* assumption). Model (1.1) then admits a stationary solution; however, we do not require $(Y_1, \ldots, Y_n)^{\mathrm{T}}$ to be a realization of this stationary solution, since all solutions of (1.1) are *asymptotically stationary*; see Hallin and

1350–7265

Werker (1998) for a detailed discussion of this issue. Furthermore, we assume that $(Y_{-p+1}, \ldots, Y_0)$ are also observed; if they are not, they can safely be set equal to zero without affecting asymptotic results.

Denote by $Q : \alpha \mapsto Q(\alpha) := \inf\{x : F(x) \geqslant \alpha\}$, $0 < \alpha < 1$, the quantile function of f, and put $q : \alpha \mapsto q(\alpha) := [f(Q(\alpha))]^{-1}$. This function q, or, more precisely, its value $q(\alpha)$ at some $\alpha \in (0, 1)$, plays a role in a variety of inference problems – mainly in the semi-parametric or nonparametric context, as a nuisance. For instance, the construction of confidence intervals for the population quantile of order α involves the asymptotic variance $\sigma_\alpha^2 = \alpha(1-\alpha)q^2(\alpha)$ of the corresponding empirical quantile; see, for instance, Csörgő (1983) or Csörgő and Révész (1984). The same factor appears in the asymptotic covariance matrices of the α-regression and α-autoregression quantiles (Koenker and Bassett 1978; Koul and Saleh 1995). The adaptive procedures and tests based on L_1-regression or L_1-autoregression also involve the estimation of $q(\alpha)$ at some fixed value of α (Koenker 1987; McKean and Schrader 1987). Finally, $q(\alpha)$ appears in the problem of building bounded-length sequential confidence intervals for quantiles (Geertsema 1992). Tukey (1965) proposed the term *sparsity function* for q, which never really caught on; Parzen (1979) reverts to the more classical term *quantile density function*, to which we henceforth also adhere.

A vast literature has been devoted to the estimation of $q(\alpha)$ in the context of location or linear regression models, including works by Siddiqui (1960), Bloch and Gastwirth (1968), Bofinger (1975), Lai *et al*. (1983), Sheather and Maritz (1983), Yang (1985), Falk (1986) and Zelterman (1990) in the one-sample case, and Welsh (1987a; 1987b), Koenker and Bassett (1982), Dodge and Jurečková (1987, 1991, 1992), and Dodge *et al*. (1991) in the regression context.

The most general results in the linear regression set-up have been obtained by Koul *et al*. (1987) (see also Koul 1992), who discuss the consistency of functionals of the type

$$Q(f) := \int f \,\mathrm{d}\phi(F) = \int_0^1 f(F^{-1}(\alpha))\mathrm{d}\phi(\alpha), \tag{1.4}$$

where ϕ is a non-decreasing right-continuous function on (0, 1) with bounded total variation.

The time series context has so far been much less explored. To the best of our knowledge, the only attempt in the direction of serially dependent observations can be found in the monograph by Koul (1992). The methodology described there, and further developed in Koul (1996), would allow for a class of nonlinear regression and autoregression models much larger than the AR(p) models considered here. On the other hand, this methodology only allows for consistency results over bounded intervals of the form $[\alpha, 1-\alpha]$, $0 < \alpha < \frac{1}{2}$, or, equivalently, over bounded functions ϕ in (1.4). Moreover, the resulting estimates do not necessarily enjoy the desirable *invariance* or *equivariance* properties we now describe (for instance, the estimator Q_n^α proposed by Koul (1992) is invariant when based on equivariant estimators of the nuisance parameter – $\boldsymbol{\rho}$ in this context – only).

In view of its role in the estimation of the variance of estimators of location and/or regression, it is highly desirable that an estimator $T_n(\mathbf{Y})$ of $q(\alpha)$ be location- and/or regression-invariant. This was pointed out by Dodge and Jurečková (1995) in the context of regression models with independent observations. The same arguments remain valid in the autoregression context, where an estimator $T_n(\mathbf{Y})$ of $q(\alpha)$ should be scale-equivariant and *autoregression*-invariant. More precisely, defining

$$\mathbf{x}_t^* := (Y_t, \ldots, Y_{t-p+1})^{\mathrm{T}} \quad \text{and} \quad \mathbf{x}_t := (1, \mathbf{x}_t^{*\mathrm{T}})^{\mathrm{T}}, \qquad t = 0, \ldots, n-1,$$

consider the random matrices

$$\mathbf{X}_n^* := (\mathbf{x}_0^*, \ldots, \mathbf{x}_{n-1}^*)^{\mathrm{T}} \quad \text{and} \quad \mathbf{X}_n := (\mathbf{x}_0, \ldots, \mathbf{x}_{n-1})^{\mathrm{T}}$$

($(n \times p)$ and $(n \times (p+1))$), respectively. An estimator $T_n(\mathbf{Y})$ of $q(\alpha)$ should satisfy the following properties:

(i) $T_n(a\mathbf{Y}) = aT_n(\mathbf{Y})$, for all $a > 0$ (scale equivariance);
(ii) $T_n(\mathbf{Y} + \mathbf{X}_n\mathbf{b}) = T_n(\mathbf{Y})$, for all $\mathbf{b} \in \mathbb{R}^{p+1}$ (autoregression invariance).

Classical estimators of $q(\alpha)$ – see the references above – unfortunately do not satisfy (i) and (ii). In the linear regression model, this led Dodge and Jurečková (1995) to construct estimators based on *regression quantiles*, which do possess these two essential properties. The objective of the present paper is to propose similar estimators, based on *auto*regression quantiles, for the quantile density of the non-observable innovation of the AR(p) model (1.1).

Two estimators based on autoregression quantiles – the first one of histogram type, the second one a kernel estimator – are proposed in Section 2. In order to study their asymptotic behaviour, we first obtain, in Section 3, a uniform Bahadur–Kiefer representation for *autoregression quantiles*. This asymptotic representation, of independent theoretical interest, is used in Section 4, along with the strong approximation results for quantile processes obtained by Csörgő and Révész (1978) and Csörgő (1983), in the derivation of the asymptotic distribution of our estimators. Finally, the finite-sample performance of the proposed estimators is investigated, in Section 5, by means of numerical simulations.

2. Two estimators of the quantile density function, based on autoregression quantiles

Following Koul and Saleh (1995), we define the α-autoregression quantile as a solution

$$\hat{\boldsymbol{\varrho}}_n(\alpha) := (\hat{\varrho}_n^0(\alpha), \hat{\boldsymbol{\varrho}}_n^1(\alpha)), \qquad \hat{\varrho}_n^0(\alpha) \in \mathbb{R}, \hat{\boldsymbol{\varrho}}_n^1(\alpha) \in \mathbb{R}^p$$

of the minimization problem

$$\sum_{t=1}^{n} h_\alpha(Y_t - r_0 - \mathbf{x}_{t-1}^{*\mathrm{T}}\mathbf{r}_1) := \min, \tag{2.1}$$

where the minimum is taken with respect to $\mathbf{r} = (r_0, \mathbf{r}_1^{\mathrm{T}})^{\mathrm{T}} \in \mathbb{R}^{p+1}$, and

$$h_\alpha(u) := |u|\{\alpha I[u > 0] + (1-\alpha)I[u \leqslant 0]\}, \qquad u \in \mathbb{R}, \alpha \in (0, 1).$$

The finite-sample and asymptotic properties of autoregression quantiles and the related autoregression rank score processes are studied in Koul and Saleh (1995) and Hallin and Jurečková (1999). Letting $\varepsilon_{t;\alpha} := \varepsilon_t - F^{-1}(\alpha)$, $\psi_\alpha := \alpha - I[u \leqslant 0]$, $0 < \alpha < 1$, $u \in \mathbb{R}$, and

$$\alpha_n := n^{-1}(\log n)^2(\log\log n)^2, \tag{2.2}$$

it has been shown – see Theorem 3.1 or Hallin and Jurečková (1999) – that, under the assumptions (F1) and (F2) given below in Section 3,

$$n^{1/2}\sigma_\alpha^{-1}(\hat{\boldsymbol{\varrho}}_n(\alpha) - \boldsymbol{\varrho}(\alpha)) = n^{-1/2}q(\alpha)(\boldsymbol{\Sigma}_n)^{-1}\sum_{t=1}^{n}\mathbf{x}_{t-1}^{*}\psi_\alpha(\varepsilon_{t;\alpha}) + \mathbf{R}_n(\alpha), \tag{2.3}$$

where $\boldsymbol{\varrho}(\alpha) := (0, \boldsymbol{\varrho}^{\mathrm{T}})^{\mathrm{T}} + F^{-1}(\alpha)\mathbf{e}_1$, $\mathbf{e}_1 := (1, 0, \ldots, 0)^{\mathrm{T}}$, $\boldsymbol{\Sigma}_n := n^{-1}\sum_{t=1}^{n}\mathbf{x}_{t-1}\mathbf{x}_{t-1}^{\mathrm{T}}$ and

$$\sup_{\alpha_n \leqslant \alpha \leqslant 1-\alpha_n} \mathbf{R}_n(\alpha) = o_{\mathrm{P}}(1), \tag{2.4}$$

as $n \to \infty$.

The results obtained by Koul and Saleh (1995) require less stringent assumptions, but consistency then only holds over fixed-length intervals, of the form $\alpha_0 \leqslant \alpha \leqslant 1 - \alpha_0$, $\alpha_0 < \frac{1}{2}$, while the uniformity of (2.4) plays an essential role in the context of this paper.

We propose two estimators of $q(\alpha)$. Our first is of histogram type,

$$H_n(\alpha) := \frac{\hat{\varrho}_n^0(\alpha + h_n) - \hat{\varrho}_n^0(\alpha - h_n)}{2h_n}, \tag{2.5}$$

where $h_n = o(n^{-1/3})$ and $nh_n \to \infty$ as $n \to \infty$; this estimator is in the spirit of Siddiqui (1960). Our second, $\kappa_n(\alpha)$, is of kernel type, in the spirit of Falk (1986),

$$\kappa_n(\alpha) := \int_0^1 \hat{\varrho}_n^0(u)h_n^{-2}k\left(\frac{\alpha - u}{h_n}\right)\mathrm{d}u, \tag{2.6}$$

where the following condition holds:

(K1) $h_n = o(n^{-1/3})$ and $nh_n^2 \to \infty$ as $n \to \infty$, and $k : \mathbb{R} \to \mathbb{R}$ is a compactly supported continuous kernel function such that $\int k(x)\mathrm{d}x = 0$ and $\int xk(x)\mathrm{d}x = -1$.

In order to derive the limiting distributions of the proposed estimators $H_n(\alpha)$ and $\kappa_n(\alpha)$, however, we need the exact rate at which the approximation error $\mathbf{R}_n$ in (2.4) goes to zero. We therefore establish a uniform Bahadur–Kiefer representation for autoregression quantiles.

3. A uniform Bahadur–Kiefer representation for autoregression quantiles

In this section, we establish a uniform Bahadur representation for $\hat{\boldsymbol{\varrho}}_n$. This representation, which is of independent interest, will be used in Section 4, in the study of the asymptotic behaviour of $H_n(\alpha)$ and $\kappa_n(\alpha)$. This Bahadur representation, however, requires some regularity assumptions on the innovation density f, which we borrow from Hallin and Jurečková (1999):

(F1) $f \in \mathcal{F}$ is absolutely continuous, with derivative f' almost everywhere and finite Fisher information for location $\mathcal{I}(f) := \int (f'(x)/f(x))^2 f(x)\mathrm{d}x < \infty$; moreover, there exists $K_f \geqslant 0$ such that, for all $|x| > K_f$, f has two bounded derivatives, f' and f'', respectively;

(F2) f is monotonically decreasing to 0 as $x \to \pm\infty$ and, for some $b = b_f > 0$, $r = r_f \geqslant 1$,

$$\lim_{x\to-\infty} \frac{-\log\, F(x)}{b|x|^r} = \lim_{x\to\infty} \frac{-\log(1 - F(x))}{b|x|^r} = 1.$$

We then have the following result.

Theorem 3.1 *Assume that* $f \in \mathcal{F}$ *satisfies* (F1) *and* (F2). *Then*

$$n^{1/2}(\hat{\varrho}_n^0(\alpha) - F^{-1}(\alpha)) = n^{-1/2} q(\alpha) \sum_{t=1}^{n} \psi_\alpha(\varepsilon_{t;\alpha}) + O_{\mathrm{P}}(n^{-1/4}(\log n)^{1/r}(\log\log n)^{1/4})$$

$$n^{1/2}(\hat{\boldsymbol{\varrho}}_n^1(\alpha) - \boldsymbol{\varrho}) = n^{-1/2} q(\alpha)(\Sigma^*)^{-1} \sum_{t=1}^{n} \mathbf{x}_{t-1}^* \psi_\alpha(\varepsilon_{t;\alpha}) + O_{\mathrm{P}}(n^{-1/4}(\log n)^{1/r}(\log\log n)^{1/4}),$$

as $n \to \infty$, *uniformly in* $\alpha \in [\alpha_n, 1 - \alpha_n]$, *where* $\boldsymbol{\Sigma}^*$ *denotes the* $p \times p$ *autocovariance matrix in the distribution of the stationary solution of* (2.1) *and* α_n *is given in* (2.2).

The proof of this theorem relies on the following lemmas.

Lemma 3.1 *Under* (F1) *and* (F2),

$$\left\| \sum_{t=1}^{n} \mathbf{x}_{t-1}[I(\varepsilon_t \leqslant \mathbf{x}_{t-1}^{\mathrm{T}}(\hat{\boldsymbol{\varrho}}_n(\alpha) - \boldsymbol{\varrho}(\alpha)) + F^{-1}(\alpha)) - \alpha] \right\| = O_{\mathrm{P}}((\log n)^{1/r}(\log\log n)^{1/4}),$$

as $n \to \infty$, *uniformly in* $[\alpha_n, 1 - \alpha_n]$.

Proof. Denote by $\mathcal{H}_\alpha$ the subset of $p+1$ values of t for which $Y_t = \mathbf{x}_{t-1}^{\mathrm{T}} \hat{\boldsymbol{\varrho}}_n(\alpha)$, and by $\mathbf{Y}_{\mathcal{H}_\alpha}$ the matrix with rows $\{\mathbf{x}_{t-1}^{\mathrm{T}},\, t \in \mathcal{H}_\alpha\}$. Optimality of $\hat{\boldsymbol{\varrho}}_n(\alpha)$ implies (Koenker and Bassett 1978, Theorem 3.3) that

$$\sum_{t=1}^{n} \mathbf{x}_{t-1}[I(Y_t \leqslant \mathbf{x}_{t-1}^{\mathrm{T}} \hat{\boldsymbol{\varrho}}_n(\alpha)) - \alpha] - (1 - \alpha) \sum_{t \in \mathcal{H}_\alpha} \mathbf{x}_{t-1} = \mathbf{Y}_{\mathcal{H}_\alpha} \mathbf{u},$$

for all $\alpha \in [\alpha_n, 1 - \alpha_n]$, where $\mathbf{u} = (u_1, \ldots, u_{p+1})^{\mathrm{T}}$ is a vector in $\mathbb{R}^{p+1}$ such that $-\alpha \leqslant u_j \leqslant 1 - \alpha$, $j = 1, \ldots, p+1$. From Hallin and Jurečková (1999, Section 5.1), we have

$$r_n := \max_{1 \leqslant t \leqslant n} \|\mathbf{x}_{t-1}\| = O_{\mathrm{P}}((\log n)^{1/r}(\log\log n)^{1/4}) \tag{3.1}$$

as $n \to \infty$. The lemma thus follows from the fact that

$$\|\mathbf{Y}_{\mathcal{H}_\alpha}\mathbf{u}\| \leqslant (p+1)[\mathrm{tr}(\mathbf{Y}^{\mathrm{T}}_{\mathcal{H}_\alpha}\mathbf{Y}_{\mathcal{H}_\alpha})]^{1/2} \leqslant (p+1)^2 \max_{1 \leqslant t \leqslant n} \|\mathbf{x}_{t-1}\|$$

and

$$\left\| \sum_{t \in \mathcal{H}_\alpha} \mathbf{x}_{t-1} \right\| \leqslant (p+1) \max_{1 \leqslant t \leqslant n} \|\mathbf{x}_{t-1}\|.$$

□

Now, for all $\boldsymbol{\delta} \in \mathbb{R}^{p+1}$ and $\alpha \in [a_n, 1-\alpha_n]$, define

$$Q_t(\boldsymbol{\delta}, \alpha) := \mathbf{y}_{t-1}[I[\varepsilon_t \leqslant \mathbf{x}^{\mathrm{T}}_{t-1}\boldsymbol{\delta} + F^{-1}(\alpha)] - I[\varepsilon_t \leqslant F^{-1}(\alpha)]], \qquad 1 \leqslant t \leqslant n. \tag{3.2}$$

Denote by $\mathcal{A}_t$ the σ-algebra generated by $\{-Y_{p+1}, \ldots, Y_0; \varepsilon_s | s \leqslant t\}$. For each t, ε_t is independent of $\mathcal{A}_{t-1}$, and the sequence

$$\sum_{t=1}^{n} \left[Q_t^{(j)}(\boldsymbol{\delta}, \alpha) - \mathrm{E}\left(Q_t^{(j)}(\boldsymbol{\delta}, a) | \mathcal{A}_{t-1} \right) \right] := \sum_{t=1}^{n} \xi^{(j)}_{n;t}, \qquad 1 \leqslant j \leqslant p+1, \tag{3.3}$$

forms a square-integrable martingale with respect to the filtration $\{\mathcal{A}_t\}$.

Lemma 3.2 *Under* (F1) *and* (F2), *for any* $\lambda > 0$ *and any* $K > \lambda^2/2$,

$$\mathrm{P}\left[\left| \max_{1 \leqslant j \leqslant p+1} \sum_{t=1}^{n} \xi^{(j)}_{n;t} \right| \geqslant \lambda B_n \right] \leqslant n^{-K},$$

as $n \to \infty$, *where* $B_n := Cn^{1/4}(\log n)^{2/r}(\log\log n)^{1/2}$.

Proof. We have, by (3.1),

$$\mathrm{P}\left[\left| \sum_{t=1}^{n} \xi^{(j)}_{n;t} \right| \geqslant \lambda B_n \right] \leqslant \mathrm{P}\left[\left| \sum_{t=1}^{n} \xi^{(j)}_{n;t} \right| \geqslant \lambda B_n,\ r_n \leqslant K(\log n)^{1/r}(\log\log n)^{1/4} \right] + \frac{1}{2} n^{-K}.$$

Again by (3.1), we have that, for $n \geqslant n_0$, $-\tau_n^2 \leqslant \xi^{(j)}_{n;t} \leqslant \tau_n$, where $\tau_n := K(\log n)^{1/r}$ $(\log\log n)^{1/4}$. Hence, for $n \geqslant n_0$, $0 \leqslant (\xi^{(j)}_{n;t} + \tau_n^2)/(\tau_n(1+\tau_n)) \leqslant 1$, and thus, by Hoeffding's inequality, for any $\lambda > 0$ and for $n \geqslant n_0$,

$$\mathrm{P}\left[\left|\sum_{t=1}^{n}\xi_{n;t}^{(j)}\right| \geqslant \lambda B_n,\ r_n \leqslant K(\log n)^{1/r}(\log\log n)^{1/4}\right]$$

$$\leqslant \mathrm{P}\left[\left|\sum_{t=1}^{n}\left[\frac{\xi_{n;t}^{(j)}+\tau_n^2}{\tau_n(1+\tau_n)} - \mathrm{E}\left(\frac{\xi_{n;t}^{(j)}+\tau_n^2}{\tau_n(1+\tau_n)}\right)\right]\right| \geqslant \frac{\lambda B_n}{\tau_n(1+\tau_n)}\right]$$

$$\leqslant \exp\left(-\frac{\lambda^2}{2}\cdot\frac{B_n^2}{\tau^4}\right) \leqslant \frac{1}{2}n^{-\lambda^2/2} \leqslant \frac{1}{2}n^{-K},$$

for $K > \lambda^2/2$; the lemma follows. □

Lemma 3.3 *Under* (F1) *and* (F2),

$$\sup_{(\boldsymbol{\delta},\alpha)\in\boldsymbol{\Delta}_n\times[\alpha_n,1-\alpha_n]}\left\|\sum_{t=1}^{n}\xi_{n;t}(\boldsymbol{\delta},\alpha)\right\| = O_{\mathrm{P}}(B_n)$$

as $n\to\infty$, *where* $\boldsymbol{\Delta}_n := \{\boldsymbol{\delta}\in\mathbb{R}^{p+1} \mid \|\boldsymbol{\delta}\| \leqslant C(n^{-1/2}\log n)^{(1/r)-2}(\log\log n)^{-1/2}\}$.

Proof. We shall apply the chaining argument of Pollard (1984), in the same spirit as in Lemma A.2 of Koenker and Portnoy (1987). First, let $\mathcal{S}_n$ denote the set of centres of spheres of radius $n^{-5/2}$ covering $\boldsymbol{\Delta}_n\times[\alpha_n, 1-\alpha_n]$. Clearly, the number of elements in $\mathcal{S}_n$ is $O(n^{5(p+1)/2})$. Hence, letting $\lambda = \sqrt{5(p+1)}$, Lemma 3.2 implies

$$\mathrm{P}\left[\left\|\sup_{(\boldsymbol{\delta},\alpha)\in\mathcal{S}_n}\sum_{t=1}^{n}\xi_{n;t}(\boldsymbol{\delta},\alpha)\right\| \geqslant 2p\sqrt{p+1}B_n\right] \leqslant Cn^{5(p+1)/2}n^{-K} \to 0 \tag{3.4}$$

as $n\to\infty$. Now, consider two elements $(\boldsymbol{\delta}_1, \alpha_1)$ and $(\boldsymbol{\delta}_2, \alpha_2)$ in the same sphere. Define

$$\begin{aligned}
D_n &:= \left\|\sum_{t=1}^{n}[Q_t(\boldsymbol{\delta}_1,\alpha_1) - Q_t(\boldsymbol{\delta}_2,\alpha_2)]\right\| \\
&= \left\|\sum_{t=1}^{n}\mathbf{x}_{t-1}[I[\varepsilon_t \leqslant \mathbf{x}_{t-1}^{\mathrm{T}}\boldsymbol{\delta}_1 + F^{-1}(\alpha_1)] - I[\varepsilon_t \leqslant \mathbf{x}_{t-1}^{\mathrm{T}}\boldsymbol{\delta}_2 + F^{-1}(\alpha_2)]]\right\| \\
&\leqslant \max_{1\leqslant t\leqslant n}\|\mathbf{x}_{t-1}\|\sum_{t=1}^{n}|I[\varepsilon_t \leqslant \mathbf{x}_{t-1}^{\mathrm{T}}\boldsymbol{\delta}_1 + F^{-1}(\alpha_1)] - I[\varepsilon_t \leqslant \mathbf{x}_{t-1}^{\mathrm{T}}\boldsymbol{\delta}_2 + F^{-1}(\alpha_2)]| \\
&= \max_{1\leqslant t\leqslant n}\|\mathbf{x}_{t-1}\|S_n, \text{ say.}
\end{aligned} \tag{3.5}$$

By (F1), $|F^{-1}(\alpha_1) - F^{-1}(\alpha_2)| \leqslant Cn^{-5/2}$, so that

$$\begin{aligned}
|\mathbf{x}_{t-1}^{\mathrm{T}}\boldsymbol{\delta}_1 + F^{-1}(\alpha_1) - \mathbf{x}_{t-1}^{\mathrm{T}}\boldsymbol{\delta}_2 - F^{-1}(\alpha_2)| &\leqslant Cn^{-5/2} + Cn^{-5/2}(\log n)^{1/r}(\log\log n)^{1/4} \\
&= O_{\mathrm{P}}(n^{-5/2}(\log n)^{1/r}(\log\log n)^{1/4}).
\end{aligned}$$

For $t_1 \neq t_2$, we have

$$\mathrm{P}[|\varepsilon_{t_1} - \varepsilon_{t_2}| \leqslant Cn^{-5/2}(\log n)^{1/r}(\log\log n)^{1/4}]$$
$$= P[\varepsilon_{t_1} \in [\varepsilon_{t_2} - Cn^{-5/2}(\log n)^{1/r}(\log\log n)^{1/4},\ \varepsilon_{t_2} - Cn^{-5/2}(\log n)^{1/r}(\log\log n)^{1/4}]]$$
$$\leqslant Cn^{-5/2}(\log n)^{1/r}(\log\log n)^{1/4}.$$

Therefore,

$$\mathrm{P}\left[\min_{t_1\neq t_2} |\varepsilon_{t_1} - \varepsilon_{t_2}| \leqslant Cn^{-5/2}(\log n)^{1/r}(\log\log n)^{1/4}\right] \leqslant Cn(n-1)n^{-5/2}(\log n)^{1/r}(\log\log n)^{1/4} = o(1).$$

This means that, with probability arbitrarily close to one, $S_n \neq 0$. Hence, still with probability arbitrarily close to one,

$$D_n \leqslant 2 \max_{1\leqslant t\leqslant n} \|\mathbf{x}_{t-1}\| = O_{\mathrm{P}}((\log n)^{1/r}(\log\log n)^{1/4}).$$

Now we have

$$\left\|\left[\sum_{t=1}^{n} \mathrm{E}(Q_t(\boldsymbol{\delta}_1, \alpha_1)|\mathcal{A}_{t-1}) - \mathrm{E}(Q_t(\boldsymbol{\delta}_2, \alpha_2)|\mathcal{A}_{t-1})\right]\right\|$$
$$\leqslant \sum_{t=1}^{n} \|\mathbf{x}_{t-1}\|\, |\mathbf{x}_{t-1}^{\mathrm{T}}\boldsymbol{\delta}_1| f(F^{-1}(\alpha_1^*)) + \sum_{t=1}^{n} \|\mathbf{x}_{t-1}\|\, |\mathbf{x}_{t-1}^{\mathrm{T}}\boldsymbol{\delta}_2| f(F^{-1}(\alpha_2^*))$$
$$\leqslant \sum_{t=1}^{n} \|\mathrm{x}_{t-1}\|^2 \|\boldsymbol{\delta}_2\|\, |f(F^{-1}(\alpha_1^*)) - f(F^{-1}(\alpha_2^*))| + \sum_{t=1}^{n} \|\mathbf{x}_{t-1}\|^2 f(F^{-1}(\alpha_1^*))\|\boldsymbol{\delta}_1 - \boldsymbol{\delta}_2\|$$
$$= O_{\mathrm{P}}(n^{-3/2}(\log n)^{(3/r)-2}).$$

Thus, with probability arbitrarily close to one,

$$\left\|\sum_{t=1}^{n} \xi_{n;t}(\boldsymbol{\delta}_1, \alpha_1) - \xi_{n;t}(\boldsymbol{\delta}_2, \alpha_2)\right\| = O((\log n)^{1/r}(\log\log n)^{1/4}),$$

uniformly in $(\boldsymbol{\delta}_1, \boldsymbol{\delta}_2) \in \boldsymbol{\Delta}_n^2$ and $(\alpha_1, \alpha_2) \in [\alpha_n, 1-\alpha_n]^2$, where $\|\boldsymbol{\delta}_1 - \boldsymbol{\delta}_2\| \leqslant n^{-5/2}$ and $|\alpha_1 - \alpha_2| \leqslant n^{-5/2}$. The lemma thus follows from (3.4). □

Proof of Theorem 3.1. A Taylor expansion yields, for any $\boldsymbol{\delta} \in \boldsymbol{\Delta}_n$,

$$\sum_{t=1}^{n} \mathrm{E}(Q_t(\boldsymbol{\delta}, \alpha)|\mathcal{A}_{t-1}) = \sum_{t=1}^{n} \mathbf{x}_{t-1}[F(\mathbf{x}_{t-1}^{\mathrm{T}}\boldsymbol{\delta} + F^{-1}(\alpha)) - F^{-1}(\alpha)]$$
$$= \sum_{t=1}^{n} \mathbf{x}_{t-1}\mathbf{x}_{t-1}^{\mathrm{T}}\boldsymbol{\delta} f(F^{-1}(\alpha)) + \sum_{t=1}^{n} \mathbf{x}_{t-1}(\mathbf{x}_{t-1}^{\mathrm{T}}\boldsymbol{\delta})^2 f'(F^{-1}(\alpha^*))$$
$$= nf(F^{-1}(\alpha))\boldsymbol{\Sigma}_n\boldsymbol{\delta} + \sum_{t=1}^{n} \|\mathbf{x}_{t-1}\|^3 \cdot O_{\mathrm{P}}(n^{-1}(\log n)^{(2/r)-4}(\log\log n)^{-1}).$$

Corollary A.1 in Hallin and Jurečková (1999) implies that $\|\mathbf{x}_t\|$ has finite moments of all orders, so that $\sum_{t=1}^{n}\|\mathbf{x}_{t-1}\|^3 = O_{\mathrm{P}}(n)$. Hence,

$$\sum_{t=1}^{n} \mathrm{E}(Q_t(\boldsymbol{\delta}, \alpha)|\mathcal{A}_{t-1}) = nf(F^{-1}(\alpha))\boldsymbol{\Sigma}_n\boldsymbol{\delta} + O_{\mathrm{P}}((\log n)^{(2/r)-4}(\log\log n)^{-1}). \qquad (3.6)$$

From Theorem 3.1 of Hallin and Jurečková (1999), we have

$$\sup_{\alpha_n \leqslant \alpha \leqslant 1-\alpha_n} \sigma_\alpha^{-1}\|\hat{\boldsymbol{\varrho}}_n(\alpha) - \boldsymbol{\varrho}(\alpha)\| = O_{\mathrm{P}}(n^{-1/2}(\log\log n)^{1/2}).$$

where $\sigma_\alpha := (\alpha(1-\alpha))^{1/2}/f(F^{-1}(\alpha))$; Lemma A.1 of Hallin and Jurečková (1999) implies that

$$\sigma_\alpha(\alpha(1-\alpha))^{1/2}(-\log(\alpha(1-\alpha)))^{1-1/r} \to (rb^{1/r})^{-1}$$

as $\alpha \to 0, 1$. Consequently,

$$\sup_{\alpha_n \leqslant \alpha \leqslant 1-\alpha_n} \sigma_\alpha^{-1}\|\hat{\boldsymbol{\varrho}}_n(\alpha) - \boldsymbol{\varrho}(\alpha)\| = O_{\mathrm{P}}(n^{-1/2}(\log n)^{(1/r)-2}(\log\log n)^{-1/2}). \qquad (3.7)$$

Now define $\hat{\boldsymbol{\delta}}_n := \hat{\boldsymbol{\varrho}}_n(\alpha) - \boldsymbol{\varrho}(\alpha)$; it follows from (3.7) that $\hat{\boldsymbol{\delta}}_n \in \boldsymbol{\Delta}_n$. Applying Lemmas 3.1 and 3.3, we have, in view of (3.6),

$$\begin{aligned}\sum_{t=1}^{n} \mathbf{x}_{t-1}[\alpha - I[\varepsilon_t \leqslant F^{-1}(\alpha)]] = nf(F^{-1}(\alpha))\boldsymbol{\Sigma}_n\hat{\boldsymbol{\delta}}_n &+ O_{\mathrm{P}}((\log n)^{1/r}(\log n \log n)^{1/4}) \\ &+ O_{\mathrm{P}}((\log n)^{(2/r)-4}(\log\log n)^{-1}) \\ &+ O_{\mathrm{P}}(n^{1/4}(\log n)^{2/r}(\log\log n)^{1/2}).\end{aligned}$$

The result then follows using the fact that

$$\boldsymbol{\Sigma}_n = \begin{pmatrix} 1 & 0 \\ 0 & \boldsymbol{\Sigma}^* \end{pmatrix} + o_{\mathrm{P}}(1),$$

as $n \to \infty$, which is an immediate consequence of Corollary A.1 of Hallin and Jurečková (1999). □

4. Asymptotics of the estimators of $q(\alpha)$

The following theorems characterize the asymptotic behaviour of the estimators defined in Section 2. We start with the consistency and asymptotic normality of the histogram-type estimator $H_n(\alpha)$, in the same spirit as Theorems 4.1.3 and 4.1.4 in Csörgő (1983); it can be verified that the range of consistency $[\alpha_n, 1-\alpha_n]$, with α_n given in (2.2), is the same as that appearing in these theorems.

Theorem 4.1. *Under* (F1) *and* (F2)*, we have*

$$\sup_{\alpha_n \leqslant \alpha \leqslant \alpha_n} |H_n(\alpha) - q(\alpha)| = O_{\mathrm{P}}((nh_n)^{-1/2}), \tag{4.1}$$

as $n \to \infty$*, uniformly for* $\alpha \in [\alpha_n, 1-\alpha_n]$.

Proof. Denote by F_n the empirical distribution function corresponding to the non-observable errors $(\varepsilon_1, \ldots, \varepsilon_n)$. Theorem 3.1, the assumptions on h_n in (2.5), and Theorem 4.1.3 in Csörgő (1983) then imply that, as $n \to \infty$,

$$\begin{aligned} H_n(\alpha) = (2h_n)^{-1}[Q(\alpha+h_n) - Q(\alpha-h_n)] + (2h_n)^{-1}\{q(\alpha+h_n)[\alpha+h_n - F_n(Q(\alpha+h_n))] \\ - q(\alpha-h_n)[\alpha - h_n - F_n(Q(\alpha-h_n))]\} + O_{\mathrm{P}}(n^{-1/4}h_n^{-1}(\log n)^{1/r}(\log\log n)^{1/4}) \\ = q(\alpha) + O_{\mathrm{P}}(n^{-1/4}h_n^{-1}(\log n)^{1/r}(\log\log n)^{1/4}) + O_{\mathrm{P}}((nh_n)^{-1/2}) + O(h_n^2) \\ = q(\alpha) + O_{\mathrm{P}}((nh_n)^{-1/2}), \end{aligned}$$

uniformly for $\alpha \in [\alpha_n, 1-\alpha_n]$, as required. □

Theorem 4.2 *Under* (F1) *and* (F2)*, there exists a sequence of Brownian bridges* $\{B_n(u), 0 \leqslant u \leqslant 1\}$ *such that*

$$\sup_{\alpha_n \leqslant \alpha \leqslant 1-\alpha_n} |(2nh_n)^{1/2}(H_n(\alpha) - q(\alpha)) - q(\alpha)(2h_n)^{-1/2}(B_n(\alpha+h_n) - B_n(\alpha-h_n))|$$

$$= O_{\mathrm{P}}((nh_n)^{-1/2}\log n) \tag{4.2}$$

as $n \to \infty$.

Proof. From part (i) of Lemma A.1 of Hallin and Jurečková (1999), and using l'Hôpital's rule, we obtain

$$\sup_{0<\alpha<1} \alpha(1-\alpha)\frac{|f'(F^{-1}(\alpha))|}{f^2(F^{-1}(\alpha))} \leqslant (rb^{1/r})^{-1}.$$

Therefore, one can apply the theory of strong approximations of quantile processes (see Csörgő and Révész 1978; and Theorem 3.2.4 of Csörgő 1983). This implies the existence of a $(\varepsilon_1, \ldots, \varepsilon_n)$-measurable Brownian bridge $B_n(\cdot)$ such that

$$\begin{aligned}(2nh_n)^{1/2}(H_n(\alpha) - q(\alpha)) &= (2h_n)^{-1/2}[q(\alpha + h_n)B_n(\alpha + h_n) - q(\alpha - h_n)B_n(\alpha - h_n)] \\ &\quad + O_{\mathrm{P}}((nh_n)^{-1/2}\log n) \\ &= q(\alpha)(2h_n)^{-1/2}[B_n(\alpha + h_n) - B_n(\alpha - h_n)] + O_{\mathrm{P}}((nh_n^3)^{1/2}) \\ &\quad + O_{\mathrm{P}}((nh_n^2)^{-1/4}) + O(h_n^{1/2}) + O_{\mathrm{P}}((nh_n)^{-1/2}\log n) \\ &= (2h_n)^{-1/2}[q(\alpha + h_n)B_n(\alpha + h_n) - q(\alpha - h_n)B_n(\alpha - h_n)] \\ &\quad + O_{\mathrm{P}}((nh_n)^{-1/2}\log n),\end{aligned}$$

uniformly in $\alpha \in [\alpha_n, 1 - \alpha_n]$, as $n \to \infty$. □

The asymptotic normality of $(nh_n)^{1/2}(H_n(\alpha) - q(\alpha))$ follows as an immediate corollary:

Corollary 4.1 *Under* (F1) *and* (F2), *we have*

$$(nh_n)^{1/2}(H_n(\alpha) - q(\alpha)) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \tfrac{1}{2}q^2(\alpha)\right), \tag{4.3}$$

as $n \to \infty$.

We now turn to the asymptotic behaviour of the kernel-type estimator $\kappa_n(\alpha)$ defined in (2.6).

Theorem 4.3 *Under* (F1), (F2) *and* (K1),

$$\sup_{\alpha_n \leqslant \alpha \leqslant \alpha_n} |\kappa_n(\alpha) - q(\alpha)| = O_{\mathrm{P}}((nh_n)^{-1/2}), \tag{4.4}$$

as $n \to \infty$.

Proof. It follows from Theorem 3.1, the definition (2.6) of κ_n, and Theorem 4.1.4 in Csörgő (1983) that

$$\begin{aligned}\kappa_n(\alpha) &= h_n^{-2}\int_0^1 Q(u)k\left(\frac{\alpha-u}{h_n}\right)\mathrm{d}u + h_n^{-2}\int_0^1 q(u)[u-F_n(Q(u))]k\left(\frac{\alpha-u}{h_n}\right)\mathrm{d}u\\ &\quad + O_{\mathrm{P}}(n^{-1/4}h_n^{-1}(\log n)^{1/r}(\log\log n)^{1/4})\\ &= q(\alpha) + h_n^{-1}\int_a^b q(\alpha-h_n)[\alpha-h_nx-F_n(Q(\alpha-h_nx))]\mathrm{d}K(x)\\ &\quad + O_{\mathrm{P}}(n^{-1/4}h_n^{-1}(\log n)^{1/r}(\log\log n)^{1/4}) + O(h_n^2)\\ &= q(\alpha) + O_{\mathrm{P}}((nh_n)^{-1/2}) + O_{\mathrm{P}}(n^{-1/4}h_n^{-1}(\log n)^{1/r}(\log\log n)^{1/4}) + O(h_n^2)\\ &= q(\alpha) + O_{\mathrm{P}}((nh_n)^{-1/2}),\end{aligned} \tag{4.5}$$

uniformly for $\alpha \in [a_n, 1-\alpha_n]$, as $n \to \infty$. □

Theorem 4.4 *Under* (F1)*,* (F2) *and* (K1)*,*

$$(nh_n)^{1/2}(\kappa_n(\alpha) - q(\alpha)) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0,\, q^2(\alpha)\int K^2(x)\mathrm{d}x\right) \tag{4.6}$$

as $n \to \infty$*, where*

$$K(x) := \int_{-\infty}^{x} k(y)\mathrm{d}y. \tag{4.7}$$

Proof. Asymptotic normality readily follows from (4.5) and classical central limit theorems. □

Clearly, the kernel estimator $\kappa_n(\alpha)$ beats (asymptotically) the histogram estimator $H_n(\alpha)$ if and only if

$$\int K^2(x)dx < \tfrac{1}{2}.$$

If we choose K as the Epanechnikov (1969) kernel, that is, $K(x) := \int_{-\infty}^{x} k(y)\mathrm{d}y$ with

$$k(x) := -\frac{3}{2a^3}x, \qquad -a \leqslant x \leqslant a, \tag{4.8}$$

then

$$\int K^2(x)dx = \frac{3}{5a} < \frac{1}{2} \qquad \text{if } a > \frac{6}{5}. \tag{4.9}$$

Note that the choice of the bandwidth h_n for computing the estimators $H_n(\alpha)$ and $\kappa_n(\alpha)$ is crucial. Bofinger (1975) showed that

$$h_n = n^{-1/5}\left(4.5\left(\frac{q(\alpha)}{q''(\alpha)}\right)^2\right)^{1/5} \tag{4.10}$$

is optimal (in the minimum mean square error sense) under mild regularity conditions on F. Note that $q(\alpha)/q''(\alpha)$ is location- and scale-invariant, so that (4.10) is only influenced by the shape of the distribution function F. Of course, $q(\alpha)$ and $q''(\alpha)$ are unknown in practice; Sheather and Maritz (1983) accordingly propose a preliminary estimation of q and q'', showing that, under appropriate regularity assumptions,

$$\frac{q(\alpha)}{q''(\alpha)} = \frac{f^2}{2(f'/f)^2 + [(f'/f)^2 - f''/f]}.$$

Fortunately, this quantity is not very sensitive to F, and little is lost if the bandwidth h_n is chosen as if the underlying distribution were Gaussian. In the Gaussian case, (4.10) yields

$$h_n = n^{-1/5}\left(4.5\frac{\phi^4(\Phi^{-1}(\alpha))}{(\Phi^{-1}(\alpha))^2+1}\right)^{1/5}, \tag{4.11}$$

where ϕ and Φ denote the density and distribution function of the standard normal distribution, respectively. A bandwidth choice suggested by Hall and Sheather (1988), based on Edgeworth expansions for studentized quantiles, is

$$h_n = n^{-1/3} z_{\alpha/2}^{2/3}\left(1.5\frac{q(\alpha)}{q''(\alpha)}\right)^{1/3}, \tag{4.12}$$

where $z_\alpha := \Phi^{-1}(1-\alpha)$. Note that, for symmetric F, $h_n(\alpha) = h_n(1-\alpha)$.

5. Simulation results

Considering model (1.1), with $p = 1$, we generated $N = 1000$ series of length $n = 100$, 150, 200, and 500 from the AR(1) model

$$Y_t = \theta Y_{t-1} + \varepsilon_t, \qquad t \in \mathbb{Z}, \tag{5.1}$$

with initial value $Y_0 = 0$, $\theta = 0.5$, and i.i.d. innovations ε_t, under standard normal, standard Laplace and standard Cauchy densities. The estimators H_n and κ_n were evaluated at two distinct values, $\alpha = 0.25$ and $\alpha = 0.50$, for two choices of the bandwidth, referred to as Bofinger (equation (4.11)) and Hall and Sheather (4.12), respectively. The estimator κ_n was computed on the basis of the Epanechnikov kernel (4.8), with $a = 4/3$. The classical estimator of Siddiqui,

$$S_n := \frac{F_n^{-1}(\alpha + h_n) - F_n^{-1}(\alpha - h_n)}{2h_n},$$

where F_n denotes the residual empirical distribution function resulting from some preliminary estimation, was also computed. We report two versions of this estimate, S_n^{LS} and S_n^{AQ}, based on least-squares and autoregression quantile residuals, respectively.

Table 1. Mean bias and mean square error of the histogram-type estimator H_n, for various series lengths n, under standard normal, Laplace, and Cauchy innovation densities f, and $\alpha = 0.25, 0.50$

		Bofinger				Hall and Sheather			
n	f	$\alpha = 0.25$		$\alpha = 0.50$		$\alpha = 0.25$		$\alpha = 0.50$	
100	Normal	0.347	(0.416)	0.176	(0.130)	0.254	(0.394)	0.099	(0.138)
	Laplace	0.831	(1.914)	0.411	(0.475)	0.562	(1.484)	0.241	(0.406)
	Cauchy	4.745	(42.069)	1.069	(1.627)	3.088	(20.268)	0.653	(0.844)
150	Normal	0.273	(0.287)	0.171	(0.099)	0.168	(0.263)	0.089	(0.102)
	Laplace	0.723	(1.306)	0.332	(0.298)	0.478	(1.019)	0.175	(0.293)
	Cauchy	3.595	(21.538)	0.846	(1.006)	2.113	(9.778)	0.472	(0.494)
200	Normal	0.250	(0.220)	0.135	(0.078)	0.147	(0.215)	0.064	(0.081)
	Laplace	0.588	(0.902)	0.315	(0.251)	0.338	(0.729)	0.175	(0.234)
	Cauchy	2.858	(13.052)	0.710	(0.718)	1.516	(5.435)	0.377	(0.348)
500	Normal	0.170	(0.102)	0.097	(0.038)	0.080	(0.115)	0.036	(0.043)
	Laplace	−0.990	(1.157)	−1.479	(2.219)	−1.195	(1.644)	−1.685	(2.880)
	Cauchy	1.650	(4.834)	0.438	(0.263)	1.348	(2.815)	0.186	(0.122)

Table 2. Mean bias and mean square error of the kernel-type estimator κ_n, for various series lengths n, under standard normal, Laplace, and Cauchy innovation densities f, and $\alpha = 0.25, 0.50$

		Bofinger				Hall and Sheather			
n	f	$\alpha = 0.25$		$\alpha = 0.50$		$\alpha = 0.25$		$\alpha = 0.50$	
100	Normal	0.421	(0.420)	0.213	(0.122)	0.269	(0.346)	0.127	(0.112)
	Laplace	1.014	(1.933)	0.474	(0.453)	0.641	(1.358)	0.278	(0.356)
	Cauchy	7.723	(102.011)	1.273	(2.082)	4.092	(31.131)	0.722	(0.884)
150	Normal	0.329	(0.283)	0.180	(0.085)	0.198	(0.246)	0.095	(0.080)
	Laplace	0.798	(1.272)	0.382	(0.312)	0.466	(0.929)	0.203	(0.257)
	Cauchy	4.797	(33.509)	0.992	(1.246)	2.528	(11.989)	0.521	(0.496)
200	Normal	0.300	(0.224)	0.157	(0.067)	0.172	(0.195)	0.073	(0.067)
	Laplace	0.648	(0.870)	0.351	(0.245)	0.346	(0.640)	0.181	(0.198)
	Cauchy	3.580	(18.668)	0.85	(0.861)	1.757	(6.421)	0.407	(0.337)
500	Normal	0.191	(0.106)	0.105	(0.034)	0.092	(0.107)	0.034	(0.041)
	Laplace	−0.905	(0.994)	−1.453	(2.143)	−1.140	(1.521)	−1.876	(3.559)
	Cauchy	2.092	(6.007)	0.509	(0.327)	0.911	(1.951)	0.199	(0.121)

Table 3. Mean bias and mean square error of the Siddiqui-type estimator S_n^{LS}, for various series lengths n, under standard normal, Laplace, and Cauchy innovation densities f, and $\alpha = 0.25, 0.50$

		Bofinger				Hall and Sheather			
n	f	$\alpha = 0.25$		$\alpha = 0.50$		$\alpha = 0.25$		$\alpha = 0.50$	
100	Normal	0.294	(0.359)	0.156	(0.113)	0.181	(0.318)	0.090	(0.116)
	Laplace	−0.649	(1.166)	−1.197	(1.833)	−0.881	(1.521)	−1.397	(2.122)
	Cauchy	5.103	(35.657)	1.303	(2.712)	3.438	(27.533)	0.874	(1.601)
150	Normal	0.236	(0.254)	0.126	(0.081)	0.132	(0.247)	0.061	(0.091)
	Laplace	−0.779	(1.091)	−1.277	(1.736)	−1.006	(1.713)	−1.485	(2.316)
	Cauchy	3.537	(22.892)	1.046	(1.649)	2.193	(11.695)	0.645	(0.899)
200	Normal	0.207	(0.197)	0.116	(0.069)	0.106	(0.192)	0.061	(0.073)
	Laplace	−0.854	(1.166)	−1.331	(1.855)	−1.088	(1.639)	−1.542	(2.475)
	Cauchy	3.022	(14.707)	0.844	(1.000)	1.815	(12.151)	0.512	(0.533)
500	Normal	0.154	(0.100)	0.097	(0.036)	0.063	(0.110)	0.043	(0.048)
	Laplace	−1.009	(1.197)	−1.488	(2.248)	−1.205	(1.658)	−1.692	(2.905)
	Cauchy	1.741	(4.949)	0.524	(0.154)	0.852	(2.219)	0.266	(0.193)

Table 4. Mean bias and mean square error of the Siddiqui-type estimator S_n^{AQ}, for various series lengths n, under standard normal, Laplace, and Cauchy innovation densities f, and $\alpha = 0.25, 0.50$

		Bofinger				Hall and Sheather			
n	f	$\alpha = 0.25$		$\alpha = 0.50$		$\alpha = 0.25$		$\alpha = 0.50$	
100	Normal	0.222	(0.308)	0.132	(0.103)	0.102	(0.312)	0.046	(0.110)
	Laplace	−0.783	(1.322)	−1.261	(1.741)	−1.051	(1.800)	−1.466	(2.310)
	Cauchy	0.945	(1.339)	2.401	(9.882)	0.551	(0.683)	1.139	(4.110)
150	Normal	0.195	(0.245)	0.116	(0.077)	0.072	(0.230)	0.033	(0.089)
	Laplace	−0.843	(1.239)	−1.303	(1.803)	−1.114	(1.748)	−1.523	(2.434)
	Cauchy	0.756	(0.843)	3.933	(17.263)	0.398	(0.420)	1.799	(8.068)
200	Normal	0.171	(0.176)	0.102	(0.064)	0.050	(0.191)	0.020	(0.077)
	Laplace	−0.903	(1.221)	−1.367	(1.951)	−1.142	(1.732)	−1.594	(2.634)
	Cauchy	0.635	(0.590)	3.687	(27.432)	0.307	(0.278)	2.297	(13.890)
500	Normal	0.137	(0.091)	0.086	(0.036)	0.035	(0.105)	0.017	(0.046)
	Laplace	−1.009	(1.206)	−1.486	(2.242)	−1.198	(1.668)	−1.702	(2.941)
	Cauchy	0.439	(0.274)	1.586	(3.920)	0.162	(0.122)	0.653	(1.547)

The numerical results on the asymptotic performance of our estimators are summarized in Tables 1–4, where we report the biases along with mean square errors (in brackets). The tables show that, for the three innovation densities considered, the four estimators H_n, κ_n, S_n^{LS} and S_n^{AQ} perform better with the Hall and Sheather bandwidth than with Bofinger's. The kernel estimator κ_n performs better, and seems to converge more rapidly to the true values than the histogram estimator H_n and the Siddiqui estimators S_n^{LS} and S_n^{AQ}. Note that convergence appears to be slower under Cauchy innovations, irrespective of the method adopted.

Two other simulation studies were conducted in order to explore the resistance of the various estimators in the presence of *innovation* and *additive* outliers, respectively. Still from model (5.1), with initial value $Y_0 = 0$, standard normal innovation density, and $\theta = 0.5$, $N = 1000$ series of length $n = 200$ were generated. Ten per cent of the innovation values, however, were (randomly) contaminated with a $\mathcal{N}(0, 3)$ distribution (innovation outliers). The same model with i.i.d. standard normal innovations was then contaminated with additive outliers: a quantity 10 was added randomly, with probability 0.1, to Y_{t-1} on the right-hand side of (5.1). Results are reported in Table 5.

Clearly, our two estimators H_n and κ_n perform better, from the point of view of robustness, than S_n^{AQ} which, in turn, seems to be more robust than S_n^{LS}. In all cases, the Hall and Sheather bandwidth choice yields better results than Bofinger's.

Table 5. Mean bias and mean square error of the estimators H_n, κ_n, S_n^{LS} and S_n^{AQ}, in the presence of innovation and additive outliers, respectively, for $\alpha = 0.25$ and $\alpha = 0.50$

	Bofinger				Hall and Sheather			
	$\alpha = 0.25$		$\alpha = 0.50$		$\alpha = 0.25$		$\alpha = 0.50$	
Innovation outliers								
H_n	0.579	(0.612)	0.254	(0.150)	0.443	(0.466)	0.269	(0.156)
κ_n	0.663	(0.573)	0.277	(0.167)	0.465	(0.453)	0.270	(0.144)
S_n^{LS}	0.629	(0.626)	0.379	(0.178)	0.495	(0.501)	0.254	(0.156)
X_n^{AQ}	0.587	(0.564)	0.323	(0.167)	0.436	(0.439)	0.227	(0.156)
Additive outliers								
H_n	0.770	(0.941)	0.776	(0.903)	0.656	(0.772)	0.753	(0.733)
κ_n	0.820	(1.033)	0.923	(0.987)	0.769	(0.875)	0.764	(0.729)
S_n^{LS}	0.942	(1.164)	0.935	(1.021)	0.813	(0.987)	0.913	(0.897)
S_n^{AQ}	0.870	(1.122)	0.887	(0.903)	0.686	(0.772)	0.743	(0.697)

Acknowledgement

The research for this paper was supported by the Fonds d'Encouragement à la Recherche de l'Université Libre de Bruxelles. The second named author was also supported by an ARC contract of the Communauté Française de Belgique. The authors are most grateful to two anonymous referees whose constructive remarks led to a reinforcement of Theorems 4.1, 4.2 and 4.3.

References

Bloch, D.A. and Gastwirth, J.L. (1968) On a simple estimate of the reciprocal of the density function. *Ann. Math. Statist.*, **39**, 1083–1085.

Bofinger, E. (1975) Estimation of a density function using the order statistics. *Austral. J. Statist.*, **17**, 1–7.

Csörgő, M. (1983) *Quantile Processes with Statistical Applications*, CBMS-NSF Reg. Conf. Ser. in Appl. Math. 42. Philadelphia: Society for Industrial and Applied Mathematics.

Csörgő, M. and Révész, P. (1978) Strong approximations of the quantile process. *Ann. Statist.*, **6**, 882–894.

Csörgő, M. and Révész, P. (1984) Two approaches to constructing simultaneous confidence bounds for quantiles. *Probab. Math. Statist.*, **4**, 221–236.

Dodge, Y. and Jurečková, J. (1987) Adaptive combination of least squares and of least absolute deviation estimators. In Y. Dodge (ed.), *Statistical Data Analysis Based on L_1-norm and Related Methods*, pp. 275–284. Amsterdam: Elsevier.

Dodge, Y. and Jurečková, J. (1991) Flexible L-estimation in the linear model. *Comput. Statist. Data Anal.*, **12**, 211–220.

Dodge, Y. and Jurečková, J. (1992) A class of estimators based on adaptive convex combinations of two estimation procedures. In Y. Dodge (ed.), *L_1-Statistical Analysis and Related Methods*, pp. 31–44. Amsterdam: Elsevier.

Dodge, Y. and Jurečková, J. (1995) Estimation of quantile density function based on regression quantiles. *Statist. Probab. Lett.*, **23**, 73–78.

Dodge, Y., Antoch, J. and Jurečková, J. (1991) Adaptive combination of least squares and least absolute deviation estimators: Computational aspects. *Comput. Statist. Data Anal.*, **12**, 87–99.

Epanechnikov, V.A. (1969) Nonparametric estimation of a multivariate probability density. *Theory Probab. Appl.*, **14**, 153–158.

Falk, M. (1986) On the estimation of the quantile density function. *Statist. Probab. Lett.*, **4**, 69–73.

Geertsema, J.C. (1992) A comparison of nonparametric sequential procedures for estimating a quantile. In P.K. Sen and I.A. Salama (eds), *Order Statistics and Nonparametrics: Theory and Applications*, pp. 101–113. Amsterdam: Elsevier.

Hall, P. and Sheather, S.J. (1988) On the distribution of studentized quantiles. *J. Roy. Statist. Soc. Ser. B*, **50**, 381–391.

Hallin, M. and Jurečková, J. (1999) Optimal tests for autoregressive models based on autoregression rank scores. *Ann. Statist.*, **27**, 1385–1414.

Hallin, M. and Werker, B.J.M. (1998) Optimal testing for semi-parametric AR models: from Lagrange multipliers to autoregression rank scores and adaptive tests. In S. Ghosh (ed.), *Asymptotics, Nonparametrics and Time Series*, pp. 295–350. New York: Marcel Dekker.

Koenker, R. (1987) A comparison of asymptotic testing methods for L_1-regression. In Y. Dodge (ed.), *L_1-Statistical Analysis and Related Methods*, pp. 287–295. Amsterdam: Elsevier.

Koenker, R. and Bassett, G. (1978) Regression quantiles. *Econometrica*, **46**, 33–50.

Koenker, R. and Bassett, G. (1982) Tests of linear hypotheses and L_1-estimation. *Econometrica*, **50**, 1577–1582.

Koenker, R. and Portnoy, S. (1987) L-estimation for linear models. *J. Amer. Statist. Assoc.*, **82**, 851–857.

Koul, H.K. (1992) *Weighted Empiricals and Linear Models*, IMS Lecture Notes Monogr. Ser. 21. Hayward, CA: Institute of Mathematical Statistics.

Koul, H.K. (1996) Asymptotics of some estimators and sequential empiricals in non-linear time series. *Ann. Statist.*, **24**, 380–404.

Koul, H.K. and Saleh, A.K.Md.E. (1995) Autoregression quantiles and related rank scores processes. *Ann. Statist.*, **23**, 670–689.

Koul, H.K., Sievers, C. and McKean, J. (1987) An estimator of the scale parameter for the rank analysis of linear models under general score functions. *Scand. J. Statist.*, **14**, 131–143.

Lai, T.L., Robbins, H. and Yu, K.F. (1983) Adaptive choice of mean or median in estimating the center of a symmetric distribution. *Proc. Nat. Acad. Sci. USA*, **80**, 5803–5806.

McKean, J. and Schrader, R.M. (1987) Least absolute errors analysis of variance. In Y. Dodge (ed.), *L_1-Statistical Analysis and Related Methods*, pp. 297–305. Amsterdam: Elsevier.

Parzen, M.I. (1979) Nonparametric statistical data modeling. *J. Amer. Statist. Assoc.*, **7**, 105–131.

Pollard, D. (1984) *Convergence of Stochastic Processes*. New York: Springer-Verlag.

Sheather, S.J. and Maritz, J.S. (1983) An estimate of the asymptotic standard error of the sample mean. *Austral. J. Statist.*, **25**, 109–122.

Siddiqui, M.M. (1960) Distribution of quantiles in samples from a bivariate population. *J. Res. Nat. Bureau Standards Sect. B*, **64**, 145–150.

Tukey, J. (1965) Which part of the sample contains the information? *Proc. Nat. Acad. Sci. USA*, **53**, 127–134.

Welsh, A.H. (1987a) One-step L-estimators for the linear model. *Ann. Statist.*, **15**, 626–641. Correction: *Ann. Statist.* (1988) **16**, 481.

Welsh, A.H. (1987b) Kernel estimates of the sparsity function. In Y. Dodge (ed.), *Statistical Data Analysis based on L_1-norm and Related Methods*, pp. 369–378. Amsterdam: Elsevier.

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

Zeltermann, D. (1990) Smooth nonparametric estimation of the quantile function. *J. Statist. Plann. Inference*, **26**, 339–352.

Received December 2000 and revised October 2001