Submitted to the Annals of Statistics
arXiv: arXiv:0000.0000

SUPPLEMENT TO "ORACALLY EFFICIENT ESTIMATION OF AUTOREGRESSIVE ERROR DISTRIBUTION WITH SIMULTANEOUS CONFIDENCE BAND" *

By Jiangyan Wang[1], Rong Liu[2], Fuxia Cheng[3] and Lijian Yang[1,4]

[1]*Soochow University,* [2]*University of Toledo,* [3]*Illinois State University,* [4]*Michigan State University*

1. Proofs of Lemma A.3 - Lemma A.7.

Proof of Lemma A.3. The inequalities in (A.1) can be expressed in terms of a as follows

$$a\left(2+\eta\right)>1, h\leq n^{-a(2+2\eta)-1}, h\geq n^{2a-1}\left(\log n\right)^{2}.$$

By Condition (C4), $n^{-3/8}\ll h=h_n\ll n^{-\{2(1+\beta)\}^{-1}}$ hence we only need to choose an $a>0$, such that

$$a>\left(2+\eta\right)^{-1}, a(2+2\eta)-1>-\left\{2\left(1+\beta\right)\right\}^{-1}, 2a-1<-3/8$$

which simply are

$$a>\left(2+\eta\right)^{-1}, a>\left(1+2\beta\right)\left(1+\eta\right)^{-1}/4\left(1+\beta\right), a<5/16.$$

Note that $\beta\in(1/3,1]$ and by Condition (C5), $\eta>6/5$, so

$$\left(1+2\beta\right)\left(1+\eta\right)^{-1}/4\left(1+\beta\right)<15/88, \left(2+\eta\right)^{-1}<5/16,$$

thus an a that satisfies all the above does exist, the lemma therefore holds.

*Supported in part by NSF award DMS 1007594, Jiangsu Specially-Appointed Professor Program SR10700111, Jiangsu Province Key-Discipline (Statistics) Program ZY107002, National Natural Science Foundation of China award 11371272, Research Fund for the Doctoral Program of Higher Education of China award 20133201110002, and Summer Fellowship of University of Toledo.

Proof of Lemma A.4. We decompose X_{t-r} into truncated part and tail part:

$$X_{t-r} = X^{D_n}_{t-r,1} + X^{D_n}_{t-r,2} + X^{D_n}_{t-r,3},$$

where

$$X^{D_n}_{t-r,1} = X_{t-r} I\left(|X_{t-r}| > D_n\right),$$

$$X^{D_n}_{t-r,2} = \mathrm{E}\left[X_{t-r} I\left\{|X_{t-r}| \le D_n\right\} \mid Z_t\right],$$

$$X^{D_n}_{t-r,3} = X_{t-r} I\left\{|X_{t-r}| \le D_n\right\} - X^{D_n}_{t-r,2}.$$

Based on Lemma A.3

$$\sum\nolimits_{n=1}^{\infty} P\left(|X_{n-r}| > D_n\right) \le \sum\nolimits_{n=1}^{\infty} D_n^{-(2+\eta)} \mathrm{E}\left(|X_{n-r}|^{2+\eta} \mid Z_t\right) < \infty.$$

By Borel-Cantelli Lemma, one has

$$P\left\{\omega \mid \exists N_1(\omega), |X_{n-r}| \le D_n, \forall n > N_1(\omega)\right\} = 1.$$

Since $\{D_n\} = \{n^a\}$ is increasing,

$$P\left\{\omega \mid \exists N(\omega), |X_{t-r}| \le D_n, t = 1, \cdots, n, n > N(\omega)\right\} = 1. \tag{S.1}$$

Let $\xi_t = \xi_t(z) = n^{-1} K_h(z - Z_t) X_{t-r} = \xi_{t,1} + \xi_{t,2} + \xi_{t,3}$ with

$$\xi_{t,l} = \xi_{t,l}(z) = n^{-1} K_h(z - Z_t) X^{D_n}_{t-r,l}, l = 1, 2, 3. \tag{S.2}$$

So from (S.1), with probability 1, one has $\sum_{t=1}^{n} \xi_{t,1} = 0$ for large n.

Denote the mean of the truncated part as $X^{D_n}_{t-r,2}$, since

$$\left|X^{D_n}_{t-r,2}\right| = \left|\mathrm{E}\left[X_{t-r} I\left\{|X_{t-r}| > D_n\right\} \mid Z_t\right]\right| \le D_n^{-(1+\eta)} \mathrm{E}\left(|X_{t-r}|^{2+\eta} \mid Z_t\right),$$

according to Lemma A.3, $D_n^{-(1+\eta)} n^{1/2} h^{1/2} \to 0$, thus classic kernel smoothing theory yields

$$\sum\nolimits_{t=1}^{n} \xi_{t,2} = U_p\left(D_n^{-(1+\eta)}\right) = u_p\left(n^{-1/2} h^{-1/2}\right).$$

Lemma A.1 will be applied to $\xi_{t,3}$. Note $\mathrm{E}\left(X^{D_n}_{t-r,3} \mid Z_t\right) = 0$,

$$\mathrm{var}\left(X^{D_n}_{t-r,3} \mid Z_t\right) = \mathrm{E}\left(|X_{t-r}|^2\right) + U_p\left(D_n^{-\eta} + D_n^{-2(1+\eta)}\right).$$

Clearly $\mathrm{E}\,\xi_{t,3}=0$, and

$$\text{(S.3)}\quad \begin{aligned}\mathrm{E}\left|\xi_{t,3}\right|^2 &= n^{-2}\,\mathrm{E}\,K_h^2\left(z-Z_t\right)\mathrm{E}\left(X_{t-r,3}^{D_n}\mid Z_t\right)^2\\ &= \mathrm{E}\left(\left|X_{t-r}\right|^2\right)f\left(z\right)n^{-2}h^{-1}\int K^2\left(v\right)dv\left\{1+o\left(1\right)\right\}\end{aligned}$$

so for n large enough,

$$\text{(S.4)}\quad \mathrm{E}\left|\xi_{t,3}\right|^2\geq n^{-2}h^{-1}c_f\left\|K\right\|_2^2/2, m_2^2=\mathrm{E}\left|\xi_{t,3}\right|^2\leq 2n^{-2}h^{-1}C_f\left\|K\right\|_2^2.$$

For any integer $k\geq 3$,

$$\text{(S.5)}\quad \begin{aligned}\mathrm{E}\left|\xi_{t,3}\right|^k &= n^{-k}\,\mathrm{E}\left|K_h\left(z-Z_t\right)\left\{X_{t-r}I\left\{\left|X_{t-r}\right|\leq D_n\right\}-X_{t-r,2}^{D_n}\right\}\right|^k\\ &\leq n^{-k}\,\mathrm{E}\left|K_h\left(z-Z_t\right)\right|^k\left(2D_n\right)^k\leq n^{-k}\left(2D_nh^{-1}\left\|K\right\|_\infty\right)^k h.\end{aligned}$$

Therefore, (S.4) and (S.5) entail

$$\begin{aligned}\mathrm{E}\left|\xi_{t,3}\right|^k/\,\mathrm{E}\left|\xi_{t,3}\right|^2 &\leq 2n^{-k}\left(2D_nh^{-1}\left\|K\right\|_\infty\right)^k hn^2hc_f^{-1}\left(\left\|K\right\|_2^2\right)^{-1}\\ &\leq k!\left(C_{f,K}D_nn^{-1}h^{-1}\left\|K\right\|_\infty\right)^{k-2}, k\geq 3\end{aligned}$$

in which $C_{f,K}=\left\{8c_f^{-1}\left(\left\|K\right\|_\infty/\left\|K\right\|_2\right)^2\right\}^{1/(k-2)}$.

So Cramér's condition is fulfilled with $C_{f,K}D_nn^{-1}h^{-1}\left\|K\right\|_\infty$ as the constant c. Applying Lemma A.1 with $k=3$, let $\varepsilon_n=C_\varepsilon n^{-3/2}h^{-1/2}\log n$ for some $C_\varepsilon>0$, and $q=\left[c_q\left(\log n\right)^{-1}n\right]$, for some $c_q>0$, then $q^{-1}n\sim c_q^{-1}\log n$. According to (S.5), one has

$$m_3=\max_{1\leq t\leq n}\left\|\xi_{t,3}\right\|_3\leq\left\{\mathrm{E}\left|\xi_{t,3}\right|^3\right\}^{1/3}\leq\left(nh\right)^{-1}2D_n\left\|K\right\|_\infty h^{1/3},$$

thus $m_3^{6/7}\leq\left(Cn^{a-3}h^{-2}\right)^{2/7}=Cn^{2a/7-6/7}h^{-4/7}$. While by (S.4), one has

$$2\left\{1+\varepsilon_n^2/\left(25m_2^2+5c\varepsilon_n\right)\right\}\sim\left(C_\varepsilon\log n\right)^2\left(25C_f\left\|K\right\|_2^2n\right)^{-1},$$

so $a_1\leq 2c_q^{-1}\log n$ for large n, and

$$\begin{aligned}\frac{q\varepsilon_n^2}{25m_2^2+5c\varepsilon_n} &\geq \frac{c_qC_\varepsilon^2\log n}{50C_f\left\|K\right\|_2^2+5D_n\left\|K\right\|_\infty C_\varepsilon C_{f,K}n^{-1/2}h^{-1/2}\log n}\\ &\geq \frac{c_qC_\varepsilon^2\log n}{100C_f\left\|K\right\|_2^2}\geq\left(9+C_2\right)\log n\end{aligned}$$

for large C_ε and large n. Lastly

$$a_2(3) \le 11n \left\{ 1 + 5C_{f,K} n^{(4a+9)/14} h^{-1/14} \left(C_\varepsilon \log n\right)^{-1} \right\} \le Cn^3.$$

Notice that by taking sufficiently small c_q, for $k = 3$,

$$\alpha\left([n/\left(q+1\right)]\right)^{2k/(2k+1)} \le \alpha\left([\log n/2c_q]\right)^{6/7} \le C_\rho^{6/7} \rho^{6/7[\log n/2c_q]} \le Cn^{-11-C_2}.$$

Putting all together, for some constants $C_\varepsilon, C_q > 0$ and large n,

$$\begin{aligned} P\left\{\left|\sum_{t=1}^n \xi_{t,3}\right| > n\varepsilon_n\right\} &\le 2c_q^{-1}\left(\log n\right)\exp\left\{-\left(9+C_2\right)\log n\right\} + Cn^3 n^{-11-C_2} \\ &\le Cn^{-8-C_2}. \end{aligned}$$

According to Lemma A.1, one obtains

$$P\left\{\left|\sum_{t=1}^n \xi_{t,3}\right| > C\left(nh\right)^{-1/2}\log n\right\} \le n^{-8-C_2}.$$

To bound the truncated sum uniformly for all $z \in [-a_n, a_n]$, we discrete by equally spaced $-a_n = z_0 < z_1 < ... < z_N = a_n$ with $N = n^{4+[C_2]+1}$

$$\begin{aligned} & P\left\{\max_{j=0}^{N}\left|\sum_{t=1}^n \xi_{t,3}\left(z_j\right)\right| > C\left(nh\right)^{-1/2}\log n\right\} \\ \le\ & \sum_{j=0}^N P\left\{\left|\sum_{t=1}^n \xi_{t,3}\left(z_j\right)\right| > C\left(nh\right)^{-1/2}\log n\right\} \le n^{-3}. \end{aligned}$$

So Borel-Cantelli Lemma implies that

$$\max_{j=0,\ldots,N}\left|\sum_{t=1}^n \xi_{t,3}\left(z_j\right)\right| = O_{a.s.}\left(n^{-1/2}h^{-1/2}\log n\right).$$

Notice that

$$\begin{aligned} & \sup_{z\in[-a_n,a_n]}\left|\sum_{t=1}^n \xi_{t,3}\left(z\right)\right| \\ & \le \max_{j=0,\ldots,N}\left|\sum_{t=1}^n \xi_{t,3}\left(z_j\right)\right| + \max_{j=0,\ldots,N-1}\sup_{z\in[z_j,z_{j+1}]}\left|\sum_{t=1}^n \xi_{t,3}\left(z_j\right) - \sum_{t=1}^n \xi_{t,3}\left(z\right)\right| \\ & = O_{a.s.}\left(n^{-1/2}h^{-1/2}\log n\right) + O_p\left(n^{-(4+[C_2]+1)}4a_n h^{-2} D_n\right) \\ & = O_{a.s.}\left(n^{-1/2}h^{-1/2}\log n\right). \end{aligned}$$

Applying all the three parts to $\sup_{z\epsilon[-a_n,a_n]}\left|n^{-1}\sum_{t=1}^n K_h\left(z-Z_t\right)X_{t-r}\right|$, equation (A.2) is proved.

Proof of Lemma A.5. One can check that for any $1 \le r, s \le p$,

$$\mathrm{E}\left|X_{t-r}X_{t-s}\right|^{2+\eta} \le \sqrt{\mathrm{E}\left|X_{t-r}\right|^{2(2+\eta)}\mathrm{E}\left|X_{t-s}\right|^{2(2+\eta)}}.$$

By Condition (C5), one has $\mathrm{E}\left|X_{t-r}X_{t-s}\right|^{2+\eta} < \infty$. The proof of Lemma A.4 can be adopted to prove that

$$\begin{aligned} &\sup_{z\epsilon[-a_n,a_n]}\left|n^{-1}\sum_{t=1}^{n}K_h'\left(z-Z_t\right)\left\{X_{t-r}X_{t-s}-\mathrm{E}\,X_{t-r}X_{t-s}\right\}\right| \\ =\ & O_{a.s.}\left(n^{-1/2}h^{-1/2}\log n\right). \end{aligned} \tag{S.6}$$

So equation (A.3) holds. Note that

$$\begin{aligned} &\mathrm{E}\left|X_{t-r}X_{t-s}X_{t-v}\right|^{2+\eta} \le \mathrm{E}\left\{1/3(\left|X_{t-r}\right|+\left|X_{t-s}\right|+\left|X_{t-v}\right|)\right\}^{3(2+\eta)} \\ \le\ & \mathrm{E}\left\{1/3\left(\left|X_{t-r}\right|^{3(2+\eta)}+\left|X_{t-s}\right|^{3(2+\eta)}+\left|X_{t-v}\right|^{3(2+\eta)}\right)\right\} = \mathrm{E}\left|X_t\right|^{3(2+\eta)} < \infty. \end{aligned}$$

Similar with the proof of (A.3), one has

$$\begin{aligned} &\sup_{z\epsilon[-a_n,a_n]}\left|n^{-1}\sum_{t=1}^{n}K_h''\left(z-Z_t\right)\left\{X_{t-r}X_{t-s}X_{t-v}-\mathrm{E}\,X_{t-r}X_{t-s}X_{t-v}\right\}\right| \\ =\ & O_{a.s.}\left(n^{-1/2}h^{-1/2}\log n\right). \end{aligned}$$

Thus equation (A.4) can be obtained. Lemma A.5 follows consequently.

Proof of Lemma A.6. Define $S_n = \sum_{t=1}^{n}\left|X_{t-r}X_{t-s}X_{t-v}X_{t-w}\right|$, by Condition (C5),

$$\begin{aligned} &\mathrm{E}\left|X_{t-r}X_{t-s}X_{t-v}X_{t-w}\right| \le \mathrm{E}\left\{1/4(\left|X_{t-r}\right|+\left|X_{t-s}\right|+\left|X_{t-v}\right|+\left|X_{t-w}\right|)\right\}^4 \\ \le\ & \mathrm{E}\left\{1/4\left(\left|X_{t-r}\right|^4+\left|X_{t-s}\right|^4+\left|X_{t-v}\right|^4+\left|X_{t-w}\right|^4\right)\right\} = \mathrm{E}\left|X_t\right|^4 < \infty. \end{aligned}$$

Thus the conditions of Proposition 2.8 in [1] are fulfilled, i.e., as $n \to \infty$, $S_n/n \overset{a.s.}{\to} \mathrm{E}\left(\left|X_{t-r}X_{t-s}X_{t-v}X_{t-w}\right|\right)$, Lemma A.6 is proved.

Proof of Lemma A.7. By definition of M_n, notice that $\left|X_{t-r}X_{t-s}X_{t-v}\right| \le \left|X_{t-r}\right|\left|X_{t-s}\right|\left|X_{t-v}\right|$, one has

$$\begin{aligned} &\sup_{z\epsilon(-\infty,-a_n)}\left|n^{-1}\sum_{t=1}^{n}K_h''\left(z-Z_t\right)X_{t-r}X_{t-s}X_{t-v}\right| \\ \le\ & n^{-1}\sum_{t=1}^{n}I\left\{Z_t\in(-\infty,-a_n+h]\right\}h^{-1}\left\|K''\right\|_\infty M_n^3. \end{aligned} \tag{S.7}$$

Denote $A = n^{-1}\sum_{t=1}^{n}I\left\{Z_t\in(-\infty,-a_n+h]\right\}$, which can be written as

$$A = P\left(Z_t \le -a_n+h\right) + \sum_{t=1}^{n}\xi_{tn} \tag{S.8}$$

where $\xi_{tn} = n^{-1}\left[I\left\{Z_t \le -a_n + h\right\} - P\left(Z_t \le -a_n + h\right)\right]$.

The first part of A

$$P\left\{Z_t \in \left(-\infty, -a_n + h\right]\right\} \le \mathrm{E}\left|Z_t\right|^{6+3\eta}\left(a_n - h\right)^{-6-3\eta} \le \mathrm{E}\left|Z_t\right|^{6+3\eta} n^{-\delta(6+3\eta)}$$

by the selection of $a_n = h + n^{\delta}$.

Consider the second part of A, $\mathrm{E}\,\xi_{tn} = 0$,

$$\mathrm{E}\,\xi_{tn}^2 = P\left(Z_t \le -a_n + h\right)\left\{1 - P\left(Z_t \le -a_n + h\right)\right\}/n^2 \le \mathrm{E}\left|Z_t\right|^{6+3\eta} n^{-\delta(6+3\eta)-2}.$$

Therefore $\sum_{t=1}^{n} \xi_{tn} = O_p\left(n^{-\{\delta(6+3\eta)+1\}/2}\right)$, and $A = O_p\left(n^{-\{\delta(6+3\eta)+1\}/2}\right)$. Using this bound on A, (S.7) and $M_n = Op\left(\log\left(n\right)^{1/\lambda}\right)$, one obtains that

$$\begin{aligned} & \sup_{z\epsilon(-\infty,-a_n)}\left|\frac{1}{n}\sum_{t=1}^{n} K_h''\left(z - Z_t\right) X_{t-r}X_{t-s}X_{t-v}\right| \\ = \;& O_p\left(n^{-\{\delta(6+3\eta)+1\}/2} n^{3\gamma}/h\right) = O_p\left(n^{-\{\delta(6+3\eta)+1\}/2+3\gamma+3/8}\right) = O_p\left(n^{-1}\right) \end{aligned}$$

as the choice of $\delta > \left(7/4 + 6\gamma\right)\left(6 + 3\eta\right)^{-1}$ entails that $-\left\{\delta\left(6 + 3\eta\right) + 1\right\}/2 + 3\gamma + 3/8 < -1$.

The proof for $z\epsilon\left(a_n, \infty\right)$ is similar. Hence (A.13) holds.

If one replaces M_n^3 with M_n and M_n^2, and correspondingly $\left\|K''\right\|_\infty$ with $\left\|K\right\|_\infty$ and $\left\|K'\right\|_\infty$ on the right hand side of (S.7), the proof of (A.13) can be adopted to prove (A.11) and (A.12), Lemma A.7 is proved.

2. Global errors of $\hat{F}$ and $\hat{F}_n$. The following Table 4 compares global errors of the smooth estimator $\hat{F}$ defined in (1.3) and the nonsmooth $\hat{F}_n$ defined in (1.4) of the main text, based on the same set of residuals $\left\{\hat{Z}_t\right\}_{t=1}^{n}$. Clearly the smooth estimator has much smaller errors than the nonsmooth one in all scenarios.

Comparing SCB coverage frequencies for $\hat{F}$ and $\hat{F}_n$ is infeasible as such construction does not yet exist for $\hat{F}_n$.

References.

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

Jiangyan Wang
Center for Advanced Statistics and Econometrics Research
Soochow University
Suzhou 215006, CHINA
E-mail: wangjiangyan2007@126.com

Rong Liu
Department of Mathematics and Statistics
University of Toledo
Toledo, OH 43606, USA
E-mail: rong.liu@utoledo.edu

ϕ	n	$\overline{D}_n(\hat{F_n})$	$\overline{D}_n(\hat{F})/\overline{D}_n(\hat{F_n})$	$\overline{\text{MISE}}(\hat{F_n})$	$\overline{\text{MISE}}(\hat{F})/\overline{\text{MISE}}(\hat{F_n})$
-0.8	50	0.2836	0.3022	0.0066	0.4236
	100	0.2701	0.2404	0.0034	0.4427
	500	0.2541	0.1206	0.0007	0.4560
	1000	0.2458	0.0928	0.0003	0.4711
-0.2	50	0.2875	0.3009	0.0067	0.4260
	100	0.2701	0.2392	0.0034	0.4402
	500	0.2536	0.1212	0.0007	0.4582
	1000	0.2453	0.0930	0.0003	0.4848
0.2	50	0.2894	0.3037	0.0068	0.4474
	100	0.2696	0.2404	0.0034	0.4488
	500	0.2534	0.1213	0.0007	0.4578
	1000	0.2451	0.0930	0.0003	0.4902
0.8	50	0.2830	0.3336	0.0070	0.6512
	100	0.2713	0.2536	0.0035	0.5556
	500	0.2536	0.1227	0.0007	0.4832
	1000	0.2456	0.0935	0.0003	0.4976

TABLE 4. *Comparing $\hat{F}$ and $\hat{F_n}$: standard normal distribution errors in $AR(1)$.*

FUXIA CHENG
DEPARTMENT OF MATHEMATICS
ILLINOIS STATE UNIVERSITY
NORMAL, IL 61790, USA
E-MAIL: fcheng@ilstu.edu

LIJIAN YANG
CENTER FOR ADVANCED STATISTICS AND ECONOMETRICS RESEARCH
SOOCHOW UNIVERSITY
SUZHOU 215006, CHINA
AND
DEPARTMENT OF STATISTICS AND PROBABILITY
MICHIGAN STATE UNIVERSITY
EAST LANSING, MI 48824, USA
E-MAIL: yanglijian@suda.edu.cn