STATISTICA, anno LXXX, n. 2, 2020

ON AN ABSOLUTE AUTOREGRESSIVE MODEL AND SKEW SYMMETRIC DISTRIBUTIONS

Dong Li
Center for Statistical Science and Department of Industrial Engineering, Tsinghua University, Beijing 100084, China
Howell Tong [1]
School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu 611731, China
Center for Statistical Science and Department of Industrial Engineering, Tsinghua University, Beijing 100084, China
Department of Statistics, London School of Economics and Political Science, London, UK

1. Introduction

Although skew symmetric distributions have quite a long history, they have attracted serious attention only relatively recently, perhaps not much earlier than the turn of the millennium. Partly due to their tractability and applications in areas such as robust statistics, financial econometrics, and others, where departure from normality is called for, we have been witnessing increasing research activities in the area for nearly twenty years, so much so that there are now a small number of dedicated monographs, e.g., Genton (2004) and Azzalini and Capitanio (2014).

It is well known that a distribution comes alive only when it is related to a random dynamical process, also called a stochastic process or a time series model. Examples abound: normal distribution with Brownian motion, Poisson distribution with renewal process, negative binomial distribution with queuing process, and others. For the skew-normal distribution, the connection was first discovered by Anděl *et al.* (1984) and further studied by Azzalini (1986) and Chan and Tong (1986). Specifically, Chan and Tong (1986) studied the integral equation that underlies the connection systematically by reference to some symmetry groups. The purpose of this paper is to exploit the method developed by Chan and Tong (1986) and show how the stochastic process approach can lead to other skew symmetric distributions, including a correct version of a skew-Cauchy distribution that is different from the one studied in Genton (2004)

[1] Corresponding Author. E-mail: howell.tong@gmail.com

and Azzalini and Capitanio (2014), and some discrete skew symmetric distributions as well as some singular distributions. We briefly study the least squares method for the estimation of the skewness parameter for dependent data.

2. ABSOLUTE AUTOREGRESSIVE MODEL

The following model is a special case of the class of threshold autoregressive models

$$y_t = \alpha|y_{t-1}| + \varepsilon_t, \tag{1}$$

where $\{\varepsilon_t\}$ is a sequence of independent and identically distributed (i.i.d.) random variables and ε_t independent of $\{y_{t-j} : j \geq 1\}$. See Anděl *et al.* (1984) and Example 4.7 in Tong (1990). Let us call it an absolute autoregressive model of order 1.

First, we give a sufficient condition on the existence of a stationary solution to model (1). It is weaker than the one given on page 140 in Tong (1990).

THEOREM 1. *Suppose that* $\{\varepsilon_t\}$ *is i.i.d. with* $\mathbb{E}\max\{1, \ln|\varepsilon_t|\} < \infty$. *If* $|\alpha| < 1$, *then model* (1) *has a unique strictly stationary solution.*

PROOF. Let $\varphi(x) = \alpha|x|$. Using the construction technique in Ling *et al.* (2007), define, for any fix t and any $n \geq 1$,

$$\xi_{n,t} = \varepsilon_t + \varphi(\varphi(\cdots\varphi(\varphi(0) + \varepsilon_{t-n+1}) + \varepsilon_{t-n+2})\cdots) + \varepsilon_{t-1}), \tag{2}$$

where $\xi_{n,t}$ starts from $y_{t-n} = 0$. Then, for $n > m$,

$$\begin{aligned}|\xi_{n,t} - \xi_{m,t}| &\leq |\alpha||\xi_{n-1,t-1} - \xi_{m-1,t-1}| \leq \cdots \\ &\leq |\alpha|^m \sum_{j=0}^{\infty} |\alpha|^j |\varepsilon_{t-m-j}|.\end{aligned}$$

Since $\mathbb{E}\max\{1, \ln|\varepsilon_t|\} < \infty$, by the Kolmogorov three series theorem, we have that $\sum_{j=0}^{\infty} |\alpha|^j |\varepsilon_{t-m-j}| < \infty$ a.s.. Thus, $\{\xi_{n,t} : n \geq 1\}$ is a Cauchy sequence. Therefore, it converges a.s.. Write

$$\xi_t = \lim_{n\to\infty} \xi_{n,t}, \quad \text{a.s..} \tag{3}$$

Note that $\xi_{n,t} = \varphi(\xi_{n-1,t-1}) + \varepsilon_t$. By the continuity of $\varphi(\cdot)$, letting $n \to \infty$, we have

$$\xi_t = \varphi(\xi_{t-1}) + \varepsilon_t = \alpha|\xi_{t-1}| + \varepsilon_t.$$

By (2) and (3), $\{\xi_t\}$ is strictly stationary. Thus, there exits a strictly stationary solution to model (1). The uniqueness is clear due to $|\alpha| < 1$. The proof is complete. □

When ε_t is symmetric and has a density $f(x)$ and $|\alpha|<1$, y_t has a unique stationary probability density, which we denote by h, and it holds that h is given by the following integral equation

$$h(y)=\int_{\mathbb{R}} h(x)f(y-\alpha|x|)dx. \tag{4}$$

When f is a normal density or a Cauchy density, Anděl *et al.* (1984) and Anděl and Barton (1986) used the 'guess-and-check' approach to solve this integral equation. Specifically, they guessed a function for h and checked if the integral on the right hand side of equation (4) integrates to the same guessed function. As we shall see, this approach is not foolproof. In this paper, we adopt the systematic method developed by Chan and Tong (1986).

3. SKEW-SYMMETRIC DISTRIBUTIONS

Under the aforementioned conditions, we have

$$\begin{aligned} h(y) &= \int_{\mathbb{R}} h(x)f(y-\alpha|x|)dx \\ &= \int_0^\infty h(x)f(y-\alpha x)dx + \int_{-\infty}^0 h(x)f(y+\alpha x)dx. \end{aligned} \tag{5}$$

By the symmetry of f, we also have

$$h(-y)=\int_0^\infty h(x)f(y+\alpha x)dx + \int_{-\infty}^0 h(x)f(y-\alpha x)dx.$$

Thus,

$$h(y)+h(-y)=\int_{\mathbb{R}}[h(x)+h(-x)]f(y-\alpha x)dx.$$

Let $\bar{h}(y)=[h(y)+h(-y)]/2$. Then,

$$\bar{h}(y)=\int_{\mathbb{R}} \bar{h}(x)f(y-\alpha x)dx,$$

which is the integral equation for the stationary density of $\{x_t\}$ satisfying the (linear) autoregressive model of order 1, or in short, AR(1) model:

$$x_t=\alpha x_{t-1}+\varepsilon_t, \quad |\alpha|<1. \tag{6}$$

By (5), we have

$$h(y) = \int_0^\infty [h(x) + h(-x)] f(y - \alpha x) dx = 2 \int_0^\infty \bar{h}(x) f(y - \alpha x) dx.$$

Thus, to find the density of y_t, it suffices to find that of model (6), which is *linear*.

Particularly, when $\varepsilon_t \sim N(0,1)$, then it follows that

$$h(y) = \sqrt{\frac{2(1-\alpha^2)}{\pi}} \exp\left\{-\frac{(1-\alpha^2)y^2}{2}\right\} \Phi(\alpha y), \quad y \in \mathbb{R}, \quad |\alpha| < 1. \tag{7}$$

If $\varepsilon_t \sim N(0, 1-\alpha^2)$ with $|\alpha| < 1$, then it follows that

$$h(y) = 2\phi(y)\Phi(\tilde{\alpha} y), \quad \tilde{\alpha} = \frac{\alpha}{\sqrt{1-\alpha^2}} \in \mathbb{R}, \quad y \in \mathbb{R}, \tag{8}$$

where $\phi(y)$ and $\Phi(y)$ are the probability density and distribution of $N(0,1)$, respectively.

Now, $h(y)$ in (8) is the probability density function (pdf) of the so-called skew-normal distribution with parameter $\tilde{\alpha} \in \mathbb{R}$, which is the normal density skewed (via multiplication) by the normal distribution with parameter $\tilde{\alpha}$. This skewing method has been exploited and generalized in recent years to cover distributions beyond the normal and beyond univariate distributions. See, e.g., Azzalini and Capitanio (2014), which also gives a short history of this skewing method to construct skew-normal distributions.

For the AR approach, we first recall that all the finite stationary joint pdfs generated by a stationary nonlinear AR model of general order driven by a white noise process with symmetric distribution are symmetric if and only if the autoregressive function is skew-symmetric at points where the pdf is positive. See, e.g., Tong (1990). It is clear that the absolute AR approach obtains the skew-normal distribution by turning the nonlinear integral equation involving $|x|$ into a linear integral equation without the modulus sign; the latter corresponds to a linear AR model. It is therefore of interest to investigate what this approach gives us in respect of the Cauchy distribution.

When $\varepsilon_t \sim$ Cauchy distribution, with density

$$f(x) = \frac{1}{\pi} \frac{1}{1+x^2}, \quad x \in \mathbb{R},$$

by (6), we have

$$\bar{h}(y) = \frac{1}{\pi} \frac{1-|\alpha|}{1+(1-|\alpha|)^2 y^2}, \quad x \in \mathbb{R}, \quad |\alpha| < 1.$$

Thus,

$$h(y) = \frac{2}{\pi^2 [y^2 + (x-1)^2][y^2 + (x+1)^2]} \left\{ xy \ln\left(\frac{y^2+1}{x^2}\right) \right.$$
$$\left. + x(y^2 + x^2 - 1)\arctan(y) + \frac{|x|\pi}{2}(y^2 + x^2 - 1) + \frac{\pi}{2}(y^2 - x^2 + 1) \right\},$$

where $\varkappa = \alpha/(1-|\alpha|)$ and $|\alpha| < 1$. Although the pdf and the cumulative distribution function of a Cauchy random variable both appear in the above skew-Cauchy distribution, they no longer do so in the simple product form as in the normal case. Effectively, here is an alternative skew-Cauchy distribution to the one studied in the existing literature, e.g., Azzalini and Capitanio (2014). Note also that our skew-Cauchy distribution differs from the one due to Anděl and Barton (1986), which can take negative values; this illustrates the fact that the 'guess-and-check' approach is not always foolproof.

Figure 1 plots the density $h(y)$ for different values of α when $\varepsilon_t \sim N(0,1)$ and when ε_t has a Cauchy distribution, respectively. From the figure, we can see that h is skewed to the right (left) when α is positive (negative), even though ε_t is symmetric.

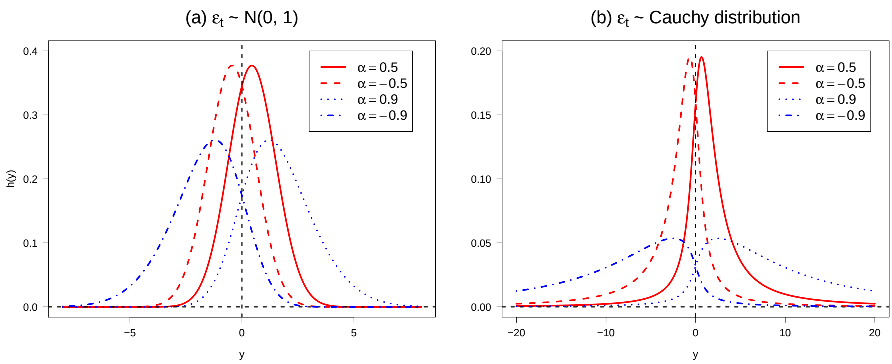

Figure 1 – The density $h(y)$ for different values of α when the noise term ε_t is $N(0,1)$ and the Cauchy distribution, respectively.

For the absolute AR approach, a natural question concerns the extension to distributions beyond the normal and the Cauchy. The fact that we have a skew symmetric distribution in *closed-form* for these two cases has to do with the self-decomposability of the normal distribution and the Cauchy distribution. In general, simple closed-forms are rare. For example, if the noise term ε_t has a t-distribution with odd degrees of freedom, then x_t in model (6) will have a mixture of t-distributions, leading to a more complicated closed-form for the skew counterpart.

4. Numerical technique

For symmetrically (or asymmetrically) distributed ε_t other than the normal and the Cauchy, it is generally difficult to obtain an explicit expression of h. Let $H(\cdot)$ denote the distribution function of y_t. From model (1), we have the following integral equation

$$H(y) = \int_{\mathbb{R}} K(y,x)H(dx), \tag{9}$$

where $K(y,x) = \mathbb{P}(\varepsilon_t \le y - \alpha|x|)$.

Chan and Tong (1986) studied theoretical properties of (9) related to the stationary pdfs of nonlinear autoregressive processes. It is generally difficult to obtain closed form formulae for $H(x)$. Several quite complicated approximate numerical methods are provided in §4.2 of Tong (1990) for some stationary stochastic processes.

Here, we use the numerical technique in Li and Qiu (2020) to solve (9). More specifically, suppose that $-\infty = y_0 < y_1 < y_2 < ... < y_m < y_{m+1} = \infty$ is a partition of $\mathbb{R}$. Then, by the definition of the Riemann-Stieltjes integral, we have the following approximation

$$
\begin{aligned}
H(y_k) &\approx \sum_{j=1}^{m+1} K(y_k, y_j^*)[H(y_j) - H(y_{j-1})] \\
&= \sum_{j=1}^{m} \left[K(y_k, y_j^*) - K(y_k, y_{j+1}^*)\right] H(y_j) + K(y_k, y_{m+1}^*),
\end{aligned}
$$

where $y_j^* \in [y_{j-1}, y_j]$ for $j = 1, ..., m+1$. Here, we take $y_j^* = (y_{j-1} + y_j)/2$. Let $\mathbf{K} = (\varkappa_{ij})_{m\times m}$ and $\mathbf{a} = (a_i)_{m\times 1}$, where

$$
\varkappa_{ij} = K(y_i, y_j^*) - K(y_i, y_{j+1}^*) \quad \text{and} \quad a_i = K(y_i, y_{m+1}^*).
$$

Denote $\mathbf{H} = (H(y_1), ..., H(y_m))'$. Then

$$
\mathbf{H} = \mathbf{K}\mathbf{H} + \mathbf{a},
$$

which yields

$$
\mathbf{H} = (\mathbf{I} - \mathbf{K})^{-1}\mathbf{a}.
$$

Now, if ε_t has a density $f(\cdot)$, then the density $h(\cdot)$ of y_t is numerically

$$
h(y) = \int_{\mathbb{R}} f(y - \alpha|u|) H(du) \approx \sum_{i=1}^{m+1} f(y - \alpha|y_i|)[H(y_i) - H(y_{i-1})].
$$

From the above discussion, we can see that the error of our approximation method is from the approximation of the related Riemann-Stieltjes integral. This error can get arbitrarily smaller by a much finer partition of $\mathbb{R}$.

Figure 2 shows the performance of the numerical technique. Here, $m = 80$ is used to partition $\mathbb{R}$, $\varepsilon_t \sim N(0,1)$ and $\alpha = 0.5$ and 0.9, respectively. The theoretical density is from (7). From the figure, we can see that the numerical approximation is quite accurate.

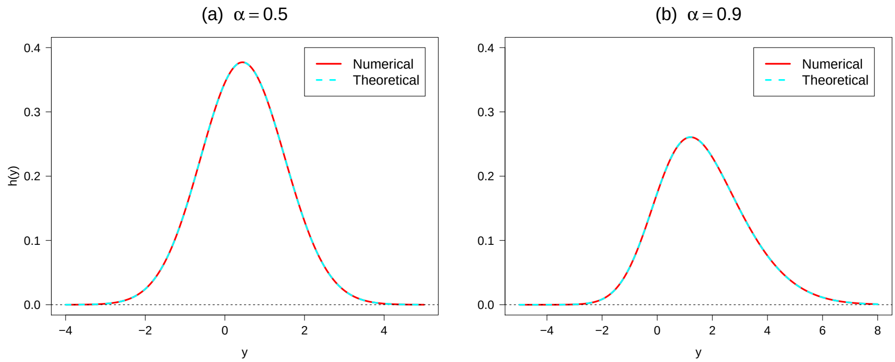

Figure 2 – The density $h(y)$ by the numerical technique and theoretical expression for different values of α when the error ε_t is $N(0,1)$, respectively.

5. The Rademacher case

Theorem 1 also allows a discrete distribution for ε_t. Here, we give an example and derive the distribution of y_t for a special case.

Theorem 2. *Suppose that* $\{\varepsilon_t\}$ *is i.i.d. with* $\mathbb{P}(\varepsilon_t = \pm 1) = 1/2$ *(called the Rademacher random variable). If* $\alpha = 1/2$*, then* y_t *in* (1) *has the following distribution function*

$$H(y) = \begin{cases} 0, & \text{if } y < -1, \\ (y+1)/2, & \text{if } -1 \le y < 0, \\ 1/2, & \text{if } 0 \le y < 1, \\ y/2, & \text{if } 1 \le y < 2, \\ 1, & \text{if } y \ge 2. \end{cases}$$

When $\alpha = -1/2$*, the distribution function of* y_t *is* $1 - H(-y)$*.*

Figure 3 plots the distribution $H(y)$ and its density $h(y)$.

Consider the linear AR(1) process:

$$x_t = 0.5 x_{t-1} + \varepsilon_t,$$

where $\{\varepsilon_t\}$ is i.i.d. with $\mathbb{P}(\varepsilon_t = \pm 1) = 1/2$. Using the characteristic function technique, we have that x_t is a uniform distribution on $[-2,2]$, i.e., $x_t \sim U[-2,2]$.

From Figures 3-4, we can see that the density of y_t is very different from that of x_t.

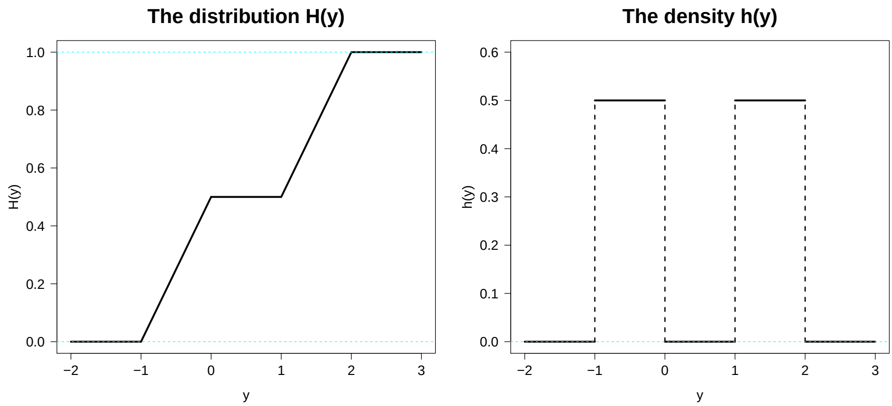

Figure 3 – The distribution function $H(y)$ with its density $h(y)$.

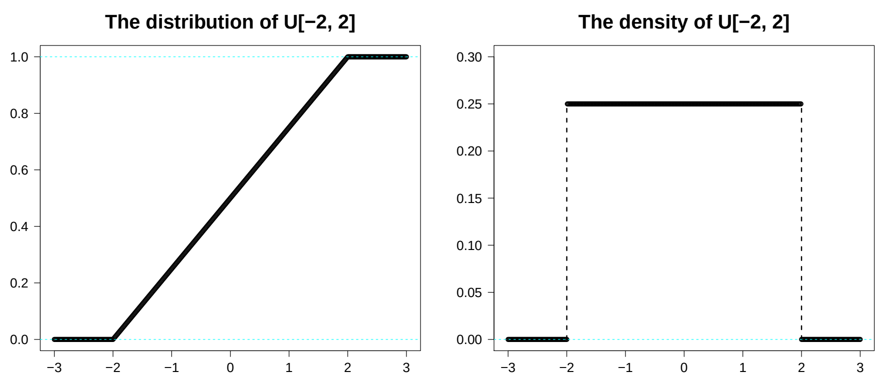

Figure 4 – The distribution function of $U[-2,2]$ with its density.

THEOREM 3. *Suppose that $\{\varepsilon_t\}$ is i.i.d. with $\mathbb{P}(\varepsilon_t = \pm 1) = 1/2$. If $\alpha \in (0, 1/2)$, then the distribution function of y_t in (1) satisfies*

$$H_\alpha(y) = \begin{cases} 0, & \text{if } \frac{y-a}{b} \le 0, \\ \frac{2j-1}{2^{n+1}}, & \text{if } \frac{y-a}{b} \in I_{n,j}, \\ 1, & \text{if } \frac{y-a}{b} \ge 1. \end{cases} \qquad n = 0, 1, 2, \ldots, \quad j = 1, \ldots, 2^n, \tag{10}$$

where

$$a = 2\alpha - \frac{1}{1-\alpha}, \quad b = \frac{2(1-\alpha+\alpha^2)}{1-\alpha}, \quad \rho = \frac{\alpha^2}{1-\alpha+\alpha^2},$$

and

$$I_{n,j} = \Bigg[(2j-1)\alpha^n \rho + (j-1)\ell_n + \sum_{k=0}^{n-1} \left\lfloor \frac{j + 2^{n-k-1} - 1}{2^{n-k}} \right\rfloor \ell_k, \\ (2j-1)\alpha^n \rho + j\ell_n + \sum_{k=0}^{n-1} \left\lfloor \frac{j + 2^{n-k-1} - 1}{2^{n-k}} \right\rfloor \ell_k \Bigg],$$

where $\lfloor \cdot \rfloor$ is the floor function and $\ell_n = (1-2\alpha)\alpha^{n-1}\rho$, with the convention $I_{0,1} = [\rho, 1-\rho]$ and $\ell_0 = 1 - 2\rho$.

When $\alpha \in (-1/2, 0)$, the distribution function of y_t is $1 - H_{-\alpha}(-y)$.

Figure 5 plots the distribution $H_\alpha(y)$ in (10) with $\alpha = 1/3$ and $1/5$, from which $H_\alpha(y)$ looks like the Cantor function.

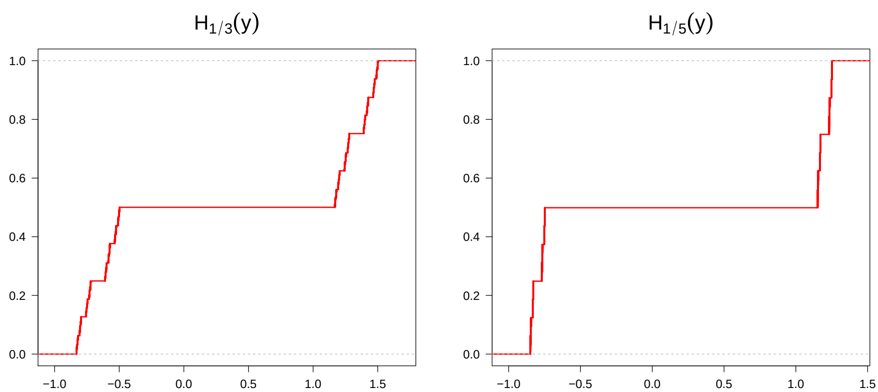

Figure 5 – The distribution $H_\alpha(y)$ with $\alpha = 1/3$ and $1/5$.

Thus, we call $H_\alpha(y)$ in (10) the quasi-Cantor distribution function. Clearly, $H_\alpha(y)$ is a singular distribution when $0 < \alpha < 1/2$. Together with Theorem 2, we can see that H_α changes from a discrete distribution to a singular one to an absolutely continuous one as α varies from 0 to $(0, 1/2)$ to $1/2$. A simulated bifurcation plot is provided in Figure 6, in which the vertical axis represents all possible values of y_t. Here, the sample size 10,000 is used and ε_t is the Rademacher random variable.

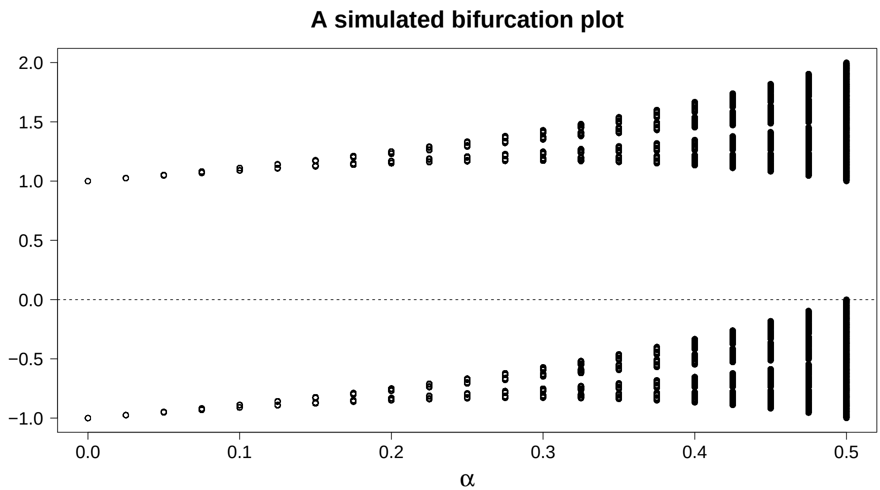

Figure 6 – The vertical axis represents all possible values of y_t in simulations for different α. The sample size is 10,000. ε_t is the Rademacher random variable.

On the other hand, we can in principle find the distribution of y_t when $1/2 < |\alpha| < 1$. However, it is much more complicated and contains two sub-cases, i.e., $1/2 < |\alpha| < 2/3$ and $2/3 \le |\alpha| < 1$. For the sub-case $1/2 < \alpha < 2/3$, a simple calculation yields that

$$H_\alpha(y) = \begin{cases} 0, & \text{if } y < -1, \\ *, & \text{if } -1 \le y < \frac{2\alpha-1}{1-\alpha}, \\ \frac{1}{2}, & \text{if } \frac{2\alpha-1}{1-\alpha} \le y < 1, \\ H_\alpha(y-2) + \frac{1}{2}, & \text{if } 1 \le y \le 1/(1-\alpha), \\ 1, & \text{if } y \ge 1/(1-\alpha). \end{cases}$$

Unfortunately, we here fail to find the closed form of $H_\alpha(y)$ when $-1 \le y < (2\alpha - 1)/(1-\alpha)$. We leave this for future work.

Theorems 2 and 3 give the explicit expressions of $H_\alpha(y)$ for $0 < \alpha \le 1/2$ when ε_t is the Rademacher random variable. Of course, theoretically, ε_t can be also allowed to be any discrete uniform random variable supported by more than two points. However, it is decidedly difficult to derive an explicit expression of $H_\alpha(y)$, even for the simple case with $\alpha = 1/2$ and $\mathbb{P}(\varepsilon_t = \pm 1) = \mathbb{P}(\varepsilon_t = 0) = 1/3$.

6. THE LEAST SQUARES ESTIMATION

In this section, we consider the estimation of α_0 for the model

$$y_t = \alpha_0 |y_{t-1}| + \varepsilon_t,$$

where $\{\varepsilon_t\}$ is i.i.d. with zero mean and finite variance. Note that α_0 controls the skewness of the distribution of y_t.

Assume that a sample $\{y_1, y_2, ..., y_n\}$ of size n is available. Note that they are generally dependent data. The initial value is y_0. Then, the least squares estimation of α_0 is

$$\widehat{\alpha}_n = \frac{\sum_{t=1}^n y_t |y_{t-1}|}{\sum_{t=1}^n |y_{t-1}|^2}.$$

THEOREM 4. *Suppose that $\{\varepsilon_t\}$ is i.i.d. with $\mathbb{E}\varepsilon_t = 0$ and $\sigma^2 = \mathbb{E}\varepsilon_t^2 \in (0, \infty)$.*

(i). *If $|\alpha_0| < 1$ and $\{y_t\}$ is ergodic, then*

$$\sqrt{n}\,(\widehat{\alpha}_n - \alpha_0) \Longrightarrow N\big(0, 1 - \alpha_0^2\big),$$

where '$\Longrightarrow$' stands for weak convergence.

(ii). *If $\alpha_0 = 1$, then*

$$n\,(\widehat{\alpha}_n - 1) \Longrightarrow \frac{[\mathbb{B}(1)]^2 - 1}{2\int_0^1 [\mathbb{B}(s)]^2 ds},$$

where $\mathbb{B}(s)$ is the standard Brownian motion on $[0, \infty)$, and if $\alpha_0 = -1$, then

$$n\,(\widehat{\alpha}_n + 1) \Longrightarrow \frac{1 - [\mathbb{B}(1)]^2}{2\int_0^1 [\mathbb{B}(s)]^2 ds}.$$

(iii). *If $|\alpha_0| > 1$ and $\mathbb{E}y_0^2 < \infty$, then*

$$\frac{\alpha_0^n\,(\widehat{\alpha}_n - \alpha_0)}{\alpha_0^2 - 1} \Longrightarrow \frac{\zeta^*}{\xi^*},$$

where ζ^* and ξ^* are independent random variables, and

$$\zeta^* \stackrel{d}{=} \zeta := \sum_{t=1}^{\infty} \frac{\varepsilon_t}{\alpha_0^t},$$

$$\xi^* \stackrel{d}{=} \xi := \sum_{k=1}^{\infty} \Big(\prod_{j=k}^{\infty} d_j\Big) \frac{\varepsilon_k}{\alpha_0^k} + \Big(\prod_{j=1}^{\infty} d_j\Big) |y_0|$$

with $d_j = I(y_j \geq 0) - I(y_j < 0)$.

We can see that Theorem 4 mimics results for the classical AR(1) model; see, e.g., Anderson (1959), Dickey and Fuller (1979), etc. When $|\alpha_0| > 1$, the limiting distribution of $\widehat{\alpha}_n$ is complicated and closely related to the error and the process itself.

Figure 7 gives the histograms of (a) $\sqrt{n}(\widehat{\alpha}_n - \alpha_0)$ with $\alpha_0 = 0.5$, (b) $n(\widehat{\alpha}_n - 1)$, and (c) $n(\widehat{\alpha}_n + 1)$. Here, the sample size is 500, 2000 replications are used. The error is standard normal.

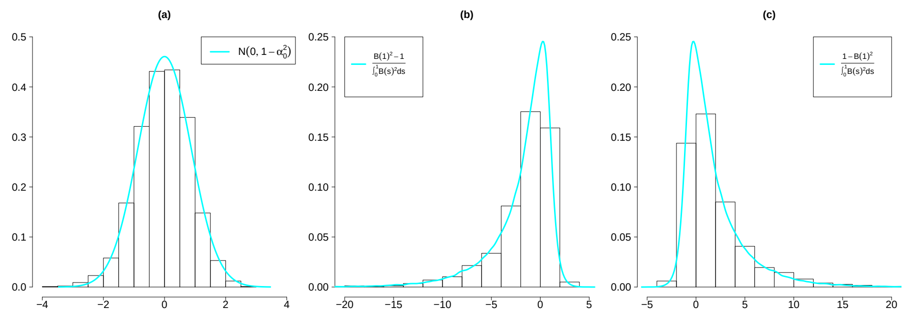

Figure 7 – The histograms of (a) $\sqrt{n}(\widehat{\alpha}_n - \alpha_0)$, (b) $n(\widehat{\alpha}_n - 1)$, and (c) $n(\widehat{\alpha}_n + 1)$.

Acknowledgements

The authors would like to thank the Editor and two anonymous referees for their very helpful comments. Li's research is supported in part by the NSFC (Nos. 11771239, 71973077 and 11571348) and the Tsinghua University Initiative Scientific Research Program (No. 2019Z07L01009). Tong's in part by a research grant from the University of Electronic Science and Technology of China, Chengdu, China.

Appendix

A. Technical Proofs

A.1. *Proof of Theorem 2*

Clearly, by the expression of y_t in (1),

$$2H(y) = \mathbb{P}(|y_{t-1}| \le 2(y+1)) + \mathbb{P}(|y_{t-1}| \le 2(y-1)). \tag{11}$$

(i). When $y < -1$, then $H(y) = 0$. This is trivial.
(ii). When $y \ge 2$, by (i) and (11), we have

$$2H(y) = H(2(y+1)) + H(2(y-1)).$$

Particularly, on letting $y=2$, it follows that

$$2H(2)=H(2(2+1))+H(2(2-1))=H(6)+H(2),$$

i.e., $H(2)=H(6)$. Since H is monotonic, we have $H(y)\equiv H(2)$ for $y\in[2,6]$. Similarly, let $y=6$. Then, $2H(6)=H(14)+H(10)$, i.e., $0=[H(14)-H(6)]+[H(10)-H(6)]\geq 0$ by the monotonicity of H. Thus $H(14)=H(6)$, which implies that $H(y)=H(6)=H(2)$ for $y\in[2,14]$. Repeating this procedure, we can get $H(y)=H(2)$ for $y\geq 2$. Thus, $H(2)=\lim_{y\to\infty}H(y)=1$, and in turn $H(y)=1$ for $y\geq 2$.

(iii). When $0\leq y<1$, by (11) and (i) again, we have

$$2H(y)=H(2(y+1)).$$

Note that $2(y+1)\geq 2$; by (ii), then $2H(y)=1$, i.e., $H(y)=1/2$ for $0\leq y<1$.

(iv). When $1\leq y<2$, by (11) and (i)-(ii), it follows that

$$2H(y)=1+H(2(y-1))-\mathbb{P}(y_{t-1}<-2(y-1)). \tag{12}$$

Particularly, if $1.5\leq y<2$, by (i), then

$$2H(y)=1+H(2(y-1)),$$

which is equivalent to, after the transformation $y\mapsto 1+\frac{y}{2}$, $y\in[1.5,2)$,

$$H(y)=2H\Big(1+\frac{y}{2}\Big)-1,\quad 1\leq y<2. \tag{13}$$

Now, we study (13). After n-step iterations, it follows that

$$\begin{aligned}H(y)&=2^nH\Big(1+\frac{1}{2}+\ldots+\frac{1}{2^{n-1}}+\frac{y}{2^n}\Big)-[2^{n-1}+2^{n-2}+\ldots+2+1]\\&=2^nH\Big(2-\frac{1}{2^{n-1}}+\frac{y}{2^n}\Big)-[2^n-1],\end{aligned}$$

that is, by $H(2)=1$ in (ii),

$$H(y)-1=\frac{H\Big(2-\frac{1}{2^{n-1}}+\frac{y}{2^n}\Big)-H(2)}{\frac{1}{2^n}}=\frac{H\Big(2-(\frac{1}{2^{n-1}}-\frac{y}{2^n})\Big)-H(2)}{\frac{1}{2^{n-1}}-\frac{y}{2^n}}\cdot(2-y).$$

Thus,

$$\frac{H(y)-1}{2-y}=\frac{H\Big(2-(\frac{1}{2^{n-1}}-\frac{y}{2^n})\Big)-H(2)}{\frac{1}{2^{n-1}}-\frac{y}{2^n}},\quad\text{for any } n\geq 1,$$

where the left-hand side does not vary with n for each fixed $y \in [1,2)$, which means the existence of the left-derivative of $H(y)$ at $y = 2$ since $\left(\frac{1}{2^{n-1}} - \frac{y}{2^n}\right) \downarrow 0$ from the right as $n \to \infty$. Write the limit in the right-hand side by $c := H'(2-)$. Thus, $H(y)$ has the linear form

$$H(y) = 1 + 2c - cy \quad \text{for} \quad y \in [1,2). \tag{14}$$

Note that $\lim_{y \uparrow 2} H(y) = \lim_{y \uparrow 2}(1 + 2c - cy) = 1 = H(2)$; then $H(y)$ is continuous at $y = 2$.

By (12), we have

$$2H(1) = 1 + H(0) - \mathbb{P}(y_{t-1} < 0).$$

Note that, by (i) and the continuity of $H(y)$ at $y = 2$,

$$\begin{aligned}\mathbb{P}(y_t < 0) &= \frac{1}{2}\mathbb{P}(|y_{t-1}| < 2) = \frac{1}{2}\mathbb{P}(-2 < y_{t-1} < 2) \\ &= \frac{1}{2}\mathbb{P}(y_{t-1} < 2) = \frac{1}{2}\mathbb{P}(y_{t-1} \le 2) = 1/2.\end{aligned}$$

Therefore, by $H(0) = 1/2$ in (iii), we have

$$2H(1) = 1 + H(0) - \mathbb{P}(y_{t-1} < 0) = 1 + \frac{1}{2} - \frac{1}{2} = 1, \quad \text{i.e.,} \quad H(1) = \frac{1}{2}.$$

Using the closed-form of H in (14), it follows that $H(1) = 1 + 2c - c = 1 + c = 1/2$, i.e., $c = -1/2$, and then $H(y) = y/2$ for $1 \le y < 2$ in (14).

(v). When $-1 \le y < 0$, we first consider the case that $-0.5 \le y < 0$. By (11) and (iv),

$$2H(y) = H(2(y+1)) = \frac{1}{2} \cdot 2(y+1) = y + 1.$$

Thus, $H(y) = (y+1)/2$ for $-0.5 \le y < 0$.

For $-1 \le y < -0.5$, by (11) and (iii) again, we have

$$2H(y) = \frac{1}{2} - \mathbb{P}(y_t < -2(y+1)).$$

Note that, for $-1 \le y < -0.5$, it follows that $-2 < -4(y+1) \le 0$ and $0 < 2-4(y+1) \le 2$.

By the continuity of $H(y)$ obtained above when $y \geq -0.5$, we have

$$\begin{aligned}
\mathbb{P}(y_t < -2(y+1)) &= \frac{1}{2}\mathbb{P}(1+0.5|y_{t-1}| < -2(y+1)) \\
&\quad + \frac{1}{2}\mathbb{P}(-1+0.5|y_{t-1}| < -2(y+1)) \\
&= \frac{1}{2}\mathbb{P}(|y_{t-1}| < -4(y+1)-2) + \frac{1}{2}\mathbb{P}(|y_{t-1}| < 2-4(y+1)) \\
&= \frac{1}{2}\mathbb{P}(|y_{t-1}| < 2-4(y+1)) \\
&= \frac{1}{2}\mathbb{P}(4(y+1)-2 < y_{t-1} < 2-4(y+1)) \\
&= \frac{1}{2}[H(2-4(y+1)) - H(4(y+1)-2)].
\end{aligned}$$

Therefore, for $-1 \leq y < -0.5$, we have

$$4H(y) = 1 + H(4(y+1)-2) - H(2-4(y+1)).$$

Particularly, when $-0.75 \leq y < -0.5$, i.e., $-1 \leq 4(y+1)-2 < 0$ and $0 < 2-4(y+1) \leq 1$, by (iii), we have

$$4H(y) = \frac{1}{2} + H(4(y+1)-2).$$

Let $z = 4(y+1)-2$, i.e., $y = (z+2)/4-1$, then

$$H(z) = 4H\left(\frac{z}{4} - \frac{1}{2}\right) - \frac{1}{2}, \quad -1 \leq z < 0.$$

Similar to (iv), we can get $H(y) = (y+1)/2$ for $-1 \leq y < 0$.

In sum, from (i)-(v), we can get the distribution $H(y)$. The proof is complete.

A.2. A Lemma

To prove Theorem 3, we first give a lemma.

LEMMA 5. *Suppose that* $\{\varepsilon_t\}$ *is i.i.d. with* $\mathbb{P}(\varepsilon_t = \pm 1) = 1/2$. *If* $0 < \alpha < 1$, *then the distribution* $H_\alpha(y)$ *of* y_t *in* (1) *satisfies*

$$H_\alpha(y) = \begin{cases} 0, & \text{if } y < -1, \\ 1, & \text{if } y \geq 1/(1-\alpha). \end{cases}$$

PROOF. By the expression of y_t in (1), it follows that

$$2H_\alpha(y)=\mathbb{P}\Big(|y_{t-1}|\le\frac{y+1}{\alpha}\Big)+\mathbb{P}\Big(|y_{t-1}|\le\frac{y-1}{\alpha}\Big). \tag{15}$$

Clearly, when $y<-1$, we have $H_\alpha(y)=0$.

When $y\ge 1/(1-\alpha)$, by a simple calculation, we then have

$$2H_\alpha(y)=H_\alpha\Big(\frac{y+1}{\alpha}\Big)+H_\alpha\Big(\frac{y-1}{\alpha}\Big).$$

Note that $(y+1)/\alpha>(y-1)/\alpha\ge y$ if $y\ge 1/(1-\alpha)$. By the monotonicity of H_α, we have

$$0=\Big[H_\alpha\Big(\frac{y+1}{\alpha}\Big)-H_\alpha(y)\Big]+\Big[H_\alpha\Big(\frac{y-1}{\alpha}\Big)-H_\alpha(y)\Big]\ge 0.$$

Thus,

$$H_\alpha(y)=H_\alpha\Big(\frac{y+1}{\alpha}\Big),\quad \text{for}\quad y\ge\frac{1}{1-\alpha}.$$

After n-iteration, it follows that

$$H_\alpha(y)=H_\alpha\Big(\frac{y}{\alpha^{2^n}}+\sum_{k=1}^{2^n}\frac{1}{\alpha^k}\Big),\quad \text{for}\quad y\ge\frac{1}{1-\alpha}.$$

Therefore,

$$H_\alpha(y)=\lim_{n\to\infty}H_\alpha\Big(\frac{y}{\alpha^{2^n}}+\sum_{k=1}^{2^n}\frac{1}{\alpha^k}\Big)=H_\alpha(\infty)=1,\quad \text{for}\quad y\ge\frac{1}{1-\alpha}.$$

The proof is complete.

A.3. *Proof of Theorem 3*

The proof will be completed by the following four steps.

First: By (15) and Lemma 5 with $\alpha^{-1}>(1-\alpha)^{-1}$ for $0<\alpha<1/2$, it follows that

$$2H_\alpha(0)=\mathbb{P}(|y_{t-1}|\le 1/\alpha)=H_\alpha(1/\alpha)=1,$$

i.e., $H_\alpha(0)=1/2$. Note that H_α is continuous at $y=0$, since

$$\begin{aligned}\mathbb{P}(y_t=0)&=\mathbb{P}(\varepsilon_t+\alpha|y_{t-1}|=0)=\frac{1}{2}\mathbb{P}(|y_{t-1}|=1/\alpha)\\&=\frac{1}{2}\big[\mathbb{P}(y_t=1/\alpha)+\mathbb{P}(y_{t-1}=-1/\alpha)\big]=0\end{aligned}$$

by Lemma 5 and the fact that $\alpha^{-1} > (1-\alpha)^{-1}$ and $-\alpha^{-1} < -2$ when $0 < \alpha < 1/2$. Thus,

$$\begin{aligned}2H_\alpha(1) &= \mathbb{P}(|y_{t-1}| \le 2/\alpha) + \mathbb{P}(|y_{t-1}| \le 0)\\ &= \mathbb{P}(-2/\alpha \le y_{t-1} \le 2/\alpha) + \mathbb{P}(y_{t-1} = 0)\\ &= H_\alpha(2/\alpha) = 1\end{aligned}$$

by Lemma 5. Then $H_\alpha(1) = 1/2$. By the monotonicity of H_α, we have

$$H_\alpha(y) = 1/2, \quad \text{for } 0 \le y \le 1. \tag{16}$$

Further, noting that

$$\frac{2\alpha - 1}{1-\alpha} \le y \le 0 \quad \text{implies} \quad 1 < \frac{1}{1-\alpha} \le \frac{y+1}{\alpha} \le \frac{1}{\alpha},$$

we have, by (15) and Lemma 5 again,

$$2H_\alpha(y) = \mathbb{P}\Big(|y_{t-1}| \le \frac{y+1}{\alpha}\Big) = \mathbb{P}\Big(-\frac{y+1}{\alpha} \le y_t \le \frac{y+1}{\alpha}\Big) = H_\alpha\Big(\frac{y+1}{\alpha}\Big) = 1.$$

Thus, combining this with (16), we have

$$H_\alpha(y) = \frac{1}{2}, \quad \text{for } \frac{2\alpha-1}{1-\alpha} \le y \le 1. \tag{17}$$

Similarly, noting that

$$1 \le y \le 1 + \alpha \cdot \frac{1-2\alpha}{1-\alpha} \quad \text{implies} \quad 0 \le \frac{y-1}{\alpha} \le \frac{1-2\alpha}{1-\alpha} \le 1,$$

we have, by (15), (16) and the stationarity of y_t,

$$\begin{aligned}2H_\alpha(y) &= 1 + H_\alpha\Big(\frac{y-1}{\alpha}\Big) - \mathbb{P}\Big(y_t < -\frac{y-1}{\alpha}\Big)\\ &= 1 + \frac{1}{2} - \mathbb{P}\Big(y_t < -\frac{y-1}{\alpha}\Big).\end{aligned}$$

Note that, by the expression of y_t in (1) and Lemma 5,

$$2\mathbb{P}\Big(y_t = -\frac{y-1}{\alpha}\Big) = \mathbb{P}\Big(y_t = \frac{1}{\alpha}\Big(1 - \frac{y-1}{\alpha}\Big)\Big) = 0$$

since

$$\frac{1}{1-\alpha} \le \frac{1}{\alpha}\Big(1 - \frac{y-1}{\alpha}\Big) \le \frac{1}{\alpha}.$$

Thus, by (17),

$$2H_\alpha(y) = \frac{3}{2} - H_\alpha\Big(-\frac{y-1}{\alpha}\Big) = \frac{3}{2} - \frac{1}{2} = 1, \quad \text{i.e.,} \quad H_\alpha(y) = \frac{1}{2},$$

since

$$\frac{2\alpha-1}{1-\alpha} \le -\frac{y-1}{\alpha} \le 0.$$

On combining this with (17), it follows that

$$H_\alpha(y) = \frac{1}{2}, \quad \text{for} \quad \frac{2\alpha-1}{1-\alpha} \le y \le 1 + \alpha \cdot \frac{1-2\alpha}{1-\alpha}. \tag{18}$$

Second: By the monotonicity of $H_\alpha(\cdot)$ and (15), it follows that

$$2H_\alpha(y) \le 2H_\alpha(a) = \mathbb{P}\Big(y_t \le \frac{1-2\alpha}{1-\alpha}\Big) - \mathbb{P}\Big(y_t < \frac{2\alpha-1}{1-\alpha}\Big), \quad y \le a.$$

Note that, by the expression of y_t in (1) and Lemma 5 again,

$$2\mathbb{P}\Big(y_t = \frac{2\alpha-1}{1-\alpha}\Big) = \mathbb{P}\Big(|y_{t-1}| = \frac{1}{1-\alpha}\Big) = 0.$$

Thus, by (16) and (18), for $y \le a$,

$$2H_\alpha(y) \le 2H_\alpha(a) = \mathbb{P}\Big(y_t \le \frac{1-2\alpha}{1-\alpha}\Big) - \mathbb{P}\Big(y_t \le \frac{2\alpha-1}{1-\alpha}\Big) = \frac{1}{2} - \frac{1}{2} = 0. \tag{19}$$

Therefore, by combining this with (18), (19) and Lemma 5, $H_\alpha(y)$ satisfies

$$H_\alpha(y) = \begin{cases} 0, & \text{if } \frac{y-a}{b} \le 0, \\ 1/2, & \text{if } \rho \le \frac{y-a}{b} \le 1-\rho, \\ 1, & \text{if } \frac{y-a}{b} \ge 1. \end{cases}$$

Third: If $\alpha\rho \le (y-a)/b \le (1-\alpha)\rho$, then,

$$0 < \frac{1-2\alpha+2\alpha^2}{1-\alpha} \le \frac{y+1}{\alpha} \le \frac{1-2\alpha^2}{1-\alpha},$$
$$\frac{2\alpha^3+3\alpha-2\alpha^2-2}{\alpha(1-\alpha)} \le \frac{y-1}{\alpha} \le \frac{3\alpha-2\alpha^3-2}{\alpha(1-\alpha)} < 0.$$

By (15), it follows that

$$2H_\alpha(y) = \mathbb{P}\Big(y_t \le \frac{y+1}{\alpha}\Big) - \mathbb{P}\Big(y_t < -\frac{y+1}{\alpha}\Big).$$

By the stationarity of y_t and

$$\frac{2\alpha-1}{1-\alpha} \le \frac{1}{\alpha}\Big(1-\frac{y+1}{\alpha}\Big) \le \frac{1-2\alpha}{1-\alpha} < 1,$$

we have

$$\begin{aligned}2\mathbb{P}\Big(y_t = -\frac{y+1}{\alpha}\Big) &= \mathbb{P}\Big(|y_t| = \frac{1}{\alpha}\Big(1-\frac{y+1}{\alpha}\Big)\Big)\\ &\le \mathbb{P}\Big(y_t = \frac{1}{\alpha}\Big(1-\frac{y+1}{\alpha}\Big)\Big) + \mathbb{P}\Big(y_t = -\frac{1}{\alpha}\Big(1-\frac{y+1}{\alpha}\Big)\Big)\\ &= 0+0=0\end{aligned}$$

by (18). Thus, by (18) and (19),

$$\begin{aligned}2H_\alpha(y) &= \mathbb{P}\Big(y_t \le \frac{y+1}{\alpha}\Big) - \mathbb{P}\Big(y_t \le -\frac{y+1}{\alpha}\Big)\\ &= H_\alpha\Big(\frac{y+1}{\alpha}\Big) - H_\alpha\Big(-\frac{y+1}{\alpha}\Big)\\ &= \frac{1}{2} - 0,\end{aligned}$$

i.e.,

$$H_\alpha(y) = \frac{1}{2^2}, \quad \text{for} \quad \alpha\rho \le (y-a)/b \le (1-\alpha)\rho.$$

Similarly,

$$H_\alpha(y) = \frac{3}{2^2}, \quad \text{for} \quad 1-(1-\alpha)\rho \le (y-a)/b \le 1-\alpha\rho.$$

Fourth: We construct a sequence of closed intervals, for each fixed $n \ge 1$ and $j = 1,\dots,2^n$,

$$\begin{aligned}I_{n,j} = \Bigg[&(2j-1)\alpha^n\rho + (j-1)\ell_n + \sum_{k=0}^{n-1}\Big\lfloor \frac{j+2^{n-k-1}-1}{2^{n-k}}\Big\rfloor \ell_k,\\ &(2j-1)\alpha^n\rho + j\ell_n + \sum_{k=0}^{n-1}\Big\lfloor \frac{j+2^{n-k-1}-1}{2^{n-k}}\Big\rfloor \ell_k\Bigg],\end{aligned}$$

where $\lfloor\cdot\rfloor$ is the floor function, $\ell_n = (1-2\alpha)\alpha^{n-1}\rho$. We adopt the convention $I_{0,1} = [\rho, 1-\rho]$ and $\ell_0 = 1-2\rho$.

Similar to the *Third* step, by a tedious calculation, we can prove that

$$H_\alpha(y) = \frac{2j-1}{2^{n+1}}, \quad \text{if} \quad \frac{y-a}{b} \in I_{n,j}, \quad n \ge 0, \quad j = 1,\dots,2^n.$$

Note that the length of the set $\bigcup_{n=0}^{\infty}\bigcup_{j=1}^{2^n} I_{n,j}$ is $(1-2\rho)+\sum_{n=1}^{\infty}2^n(1-2\alpha)\alpha^{n-1}\rho=1$. Thus, $H_\alpha(y)$ in (10) can determine the distribution of y_t. The proof is complete.

A.4. *Proof of Theorem 4*

(i). Note that

$$\sqrt{n}(\widehat{\alpha}_n-\alpha_0)=\frac{n^{-1/2}\sum_{t=1}^n \varepsilon_t|y_{t-1}|}{n^{-1}\sum_{t=1}^n |y_{t-1}|^2}.$$

When $|\alpha_0|<1$, $\{y_t\}$ is strictly stationary and $\mathbb{E}y_t^2=\sigma^2/(1-\alpha_0^2)$. With the ergodicity of $\{y_t\}$, we have

$$\frac{1}{n}\sum_{t=1}^n |y_{t-1}|^2\to\frac{\sigma^2}{1-\alpha_0^2},\quad \text{a.s..}$$

By the martingale central limit theorem in Brown (1971), it follows that

$$\frac{1}{\sqrt{n}}\sum_{t=1}^n \varepsilon_t|y_{t-1}|\Longrightarrow N(0,\sigma^2\mathbb{E}y_t^2).$$

Thus, $\sqrt{n}(\widehat{\alpha}_n-\alpha_0)\Longrightarrow N(0,1-\alpha_0^2)$.

(ii). First, consider the case $\alpha_0=1$. By Theorem 3.1 in Liu *et al.* (2011), we have

$$\frac{y_{[ns]}}{\sqrt{n}}\Longrightarrow\sigma|\mathbb{B}(s)|\quad \text{on } \mathbb{D}[0,1],$$

where $\mathbb{D}[0,1]$ is the Skorokhod space. Note that

$$\varepsilon_t|y_{t-1}|=\frac{(y_t^2-y_{t-1}^2)-\varepsilon_t^2}{2},$$

then

$$2(\widehat{\alpha}_n-1)=\frac{(y_n^2-y_0^2)-\sum_{t=1}^n\varepsilon_t^2}{\sum_{t=1}^n y_{t-1}^2}.$$

By the continuous mapping theorem,

$$n(\widehat{\alpha}_n-1)\Longrightarrow\frac{[\mathbb{B}(1)]^2-1}{2\int_0^1[\mathbb{B}(s)]^2ds}.$$

The proof with $\alpha_0=-1$ is similar and thus omitted.

(iii). The proof is similar to that of Theorem 2.2 in Liu *et al.* (2011). If $|\alpha_0| > 1$, then

$$\frac{y_n}{\alpha_0^n} = \sum_{k=1}^{n}\Big(\prod_{j=k}^{n-1} d_j\Big)\frac{\varepsilon_k}{\alpha_0^k} + \Big(\prod_{j=1}^{n-1} d_j\Big)|y_0| \to \xi, \quad \text{a.s..}$$

Thus,

$$\frac{\alpha_0^2 - 1}{\alpha_0^{2n}} \sum_{t=1}^{n} y_{t-1}^2 \to \xi^2, \quad \text{a.s..}$$

By the same argument as in Theorem 2.2 of Liu *et al.* (2011), we have

$$\frac{\alpha_0^n(\widehat{\alpha}_n - \alpha_0)}{\alpha_0^2 - 1} - \frac{\alpha_0^{-n}\sum_{t=1}^{n}\alpha_0^{t-1}\varepsilon_t}{\xi} \xrightarrow{p} 0,$$

with the convention $\prod_{k=n}^{n-1} \equiv 1$. Note that

$$\frac{y_{\lfloor n/2 \rfloor}}{\alpha_0^{\lfloor n/2 \rfloor}} \xrightarrow{\text{a.s.}} \xi \quad \text{and} \quad \alpha_0^{-n}\left\{\sum_{t=1}^{n}\alpha_0^{t-1}\varepsilon_t - \sum_{t=\lfloor n/2 \rfloor+1}^{n}\alpha_0^{t-1}\varepsilon_t\right\} \xrightarrow{p} 0.$$

By the independence between $\sum_{t=\lfloor n/2 \rfloor+1}^{n}\alpha_0^{t-1}\varepsilon_t$ and $y_{\lfloor n/2 \rfloor}$, we can see that

$$\left(\alpha_0^{-n}\sum_{t=\lfloor n/2 \rfloor+1}^{n}\alpha_0^{t-1}\varepsilon_t,\ \frac{y_{\lfloor n/2 \rfloor}}{\alpha_0^{\lfloor n/2 \rfloor}}\right) \Longrightarrow (\zeta^*, \xi^*).$$

Therefore,

$$\frac{\alpha_0^n(\widehat{\alpha}_n - \alpha_0)}{\alpha_0^2 - 1} \Longrightarrow \frac{\zeta^*}{\xi^*}.$$

The proof is complete.

References

T. Anderson (1959). *On asymptotic distributions of estimates of parameters of stochastic difference equations*. The Annals of Mathematical Statistics, 30, no. 3, pp. 676–687.

J. Anděl, T. Barton (1986). *A note on threshold AR(1) model with Cauchy innovations*. Journal of Time Series Analysis, 7, no. 1, pp. 1–5.

J. Anděl, I. Netuka, K. Zvára (1984). *On threshold autoregressive processes*. Kybernetika, 20, no. 2, pp. 89–106.

A. AZZALINI (1986). *Further results on a class of distributions which includes the normal ones*. Statistica, 46, no. 2, pp. 199–208.

A. AZZALINI, A. CAPITANIO (2014). *The Skew-Normal and Related Families*. Cambridge University Press, Cambridge.

B. M. BROWN (1971). *Martingale central limit theorems*. The Annals of Mathematical Statistics, 42, no. 1, pp. 59–66.

K. S. CHAN, H. TONG (1986). *A note on certain integral equations associated with non-linear time series analysis*. Probability Theory and Related Fields, 73, no. 1, pp. 153–158.

D. A. DICKEY, W. A. FULLER (1979). *Distribution of the estimators for autoregressive time series with a unit root*. Journal of the American Statistical Association, 74, no. 366, pp. 421–431.

M. G. GENTON (2004). *Skew-Elliptical Distributions and Their Applications: A Journey Beyond Normality*. Chapman & Hall/CRC, Boca Raton.

D. LI, J. M. QIU (2020). *The marginal density of a TMA(1) process*. Journal of Time Series Analysis, 41, no. 3, pp. 476–484.

S. LING, H. TONG, D. LI (2007). *Ergodicity and invertibility of threshold moving-average models*. Bernoulli, 13, no. 1, pp. 161–168.

W. LIU, S. LING, Q. M. SHAO (2011). *On non-stationary threshold autoregressive models*. Bernoulli, 17, no. 3, pp. 969–986.

H. TONG (1990). *Non-linear Time Series: A Dynamical System Approach*. Oxford University Press, New York.

SUMMARY

By exploiting the connection between a popular construction of a well-known skew-normal distribution and an absolute autoregressive process, we show how the stochastic process approach can lead to other skew symmetric distributions, including a skew-Cauchy distribution and some singular distributions. In so doing, we also correct an erroneous skew-Cauchy-distribution in the literature. We discuss the estimation, for dependent data, of the key parameter relating to the skewness.

Keywords: Absolute autoregressive process; Estimation of skewness parameter; Singular distributions; Skew-normal distribution; Skew-Cauchy distribution.