On Tail Behaviour of An ARCH(1) Process With A Mixture of Normal Noise

Zhiqiang Zhang[1,2] Howell Tong[1,3]

[1]The University of Hong Kong, Hong Kong
[2]East China Normal University, Shanghai, China
[3]London School of Economics, U.K.

Abstract. In this paper, we study an ARCH(1) process $X_t = Z_t\sqrt{\beta + \lambda X_{t-1}^2}$, where the noise $\{Z_t\}$ is a sequence of independent and identically distributed random variables each as a mixture of normals with zero mean and unit variance, and $\beta > 0$ and $\lambda > 0$ are two constants. Compared with the ARCH(1) process with standard normal noise, the stationarity region for this model becomes wider though the noise has a heavier tail. When the process is stationary, the tail probability $\mathrm{P}(|X_t| > x)$ shrinks more slowly than the standard normal case for $\lambda \in (0, 1)$ but faster for $\lambda > 1$. Two examples are also given in this paper.

Keywords and phrases: mixture of normal, extremes, ARCH, tail behaviour.

AMS 2000 Subject Classification: Primary: 60G10;
Secondary: 60K05, 62E05.

1. Introduction

The distributional form of returns on financial assets has important implications for theoretical analysis in economics and finance. Earlier theories were based on normal distribution, while the excess kurtosis found in Fama (1965) and others led to the rejection of the normal assumption and adoption of the stable Paretian distribution as a statistical model for asset returns. Since then, there has been a proliferation of distributions to describe the law of asset returns. Distributions such as Student-t, mixtures of normals and the Weibull provide in many instances models with better statistical fit than the stable Paretian law; see Mittnik and Rachev (1993), and other references therein.

On the other hand, aspects of conditional distributions are the focus of the class of autoregressive conditionally heteroscedastic models which were introduced by Engle (1982) and extended by Bollerslev (1986), and are widely used in financial modeling. These of ARCH/GARCH models have attracted substantial attention in the literature. However, it seems that few papers have discussed both the distributions and conditional heteroscedasticity together.

In this paper, we consider the following process

$$X_t = Z_t \sqrt{\beta + \lambda X_{t-1}^2}, \qquad t \in \mathbb{N}, \tag{1.1}$$

where the noise $\{Z_t\}$ is a sequence of independent and identically distributed random variables with a common mixture of normal distributions and with zero mean and unit variance. Here, $\beta > 0$ and $\lambda > 0$ are two constants. We will call model (1.1) an ARCH(1) process driven by a mixture of normals, and denote it by ARCH-MN(1).

When Z_t has a standard normal distribution instead of a mixture of normals, we have the classical ARCH(1) process, which we also call the normal ARCH(1)

and denote by $\{X_t^{(0)}\}$. Thus,

$$X_t^{(0)} = Z_t^{(0)} \sqrt{\beta + \lambda (X_{t-1}^{(0)})^2}, \qquad t \in \mathbb{N}, \tag{1.2}$$

where the noise $\{Z_t^{(0)}\}$ is a sequence of independent and identically distributed standard normal random random variables. This normal ARCH(1) process was studied in Embrechts *et al.* (1997), including the stationarity region, some properties on moments, the tail behaviour and extremes. We summarize some of their results in the following theorem.

Theorem 1.1. (Embrechts *et al.* (1997)) *Let $\{X_t^{(0)}\}$ be a normal ARCH(1) process defined by (1.2).*

(a) When $\lambda \in (0, 2e^{\gamma})$, where $\gamma \approx 0.5772$ is Euler's constant, there exists some random variable $X^{(0)}$ such that $X_t^{(0)} \xrightarrow{\mathrm{d}} X^{(0)}$ ("$\xrightarrow{\mathrm{d}}$" means convergence in distribution). Moreover, if $X_0^{(0)} \stackrel{\mathrm{d}}{=} X^{(0)}$ ("$A \stackrel{\mathrm{d}}{=} B$" stands for that A and B have the same distribution), then the process $\{X_t^{(0)}\}$ is strictly stationary.

(b) For $\lambda \in (0, 1)$, $\{X_t^{(0)}\}$ is stationary with finite variance $\mathrm{E}[(X_t^{(0)})^2] = \beta/(1-\lambda)$, and when $1 \le \lambda < 2e^{\gamma} \approx 3.5621$, $\{X_t^{(0)}\}$ is stationary with infinite variance.

(c) When $\{X_t^{(0)}\}$ is strictly stationary, there exist two positive constants $c_0 = c_0(\beta, \lambda) > 0$ and $\kappa_0 = \kappa_0(\lambda) > 0$ such that

$$\mathrm{P}(|X_t^{(0)}| > x) \sim c_0 x^{-2\kappa_0}, \qquad x \to \infty. \tag{1.3}$$

(d) Let $M_n^{(0)} = \max(|X_1^{(0)}|, \cdots, |X_n^{(0)}|)$. When $\{X_t^{(0)}\}$ is strictly stationary, there exists a third positive constant $\theta_0 = \theta_0(\beta, \lambda)$ such that

$$\lim_{n\to\infty} \mathrm{P}(n^{-1/2\kappa_0} M_n^{(0)} \le x) = \exp\{-c_0 \theta_0 x^{-2\kappa_0}\}, \qquad x > 0, \tag{1.4}$$

where c_0 and κ_0 are the same as in (c).

Compared with the normal ARCH(1) process $\{X_t^{(0)}\}$, the ARCH-MN(1) process $\{X_t\}$ has the following properties.

(A) stationarity region. The stationarity region for $\{X_t\}$ is $\lambda \in (0, 2e^{\gamma+\delta})$, where $\delta > 0$ is a positive constant. The stationarity region becomes **wider**. In this case, there also exists some random variable X such that $X_t \xrightarrow{\mathrm{d}} X$, and when $X_0 \stackrel{\mathrm{d}}{=} X$, $\{X_t\}$ is strictly stationary.

(B) Variance. For $\lambda \in (0,1)$, $\{X_t\}$ is stationary with finite variance $\mathrm{E}(X_t^2) = \beta/(1-\lambda)$, and when $1 \leq \lambda < 2e^{\gamma+\delta}$, $\{X_t\}$ is stationary with infinite variance.

(C) Tail behaviour. When $\{X_t\}$ is strictly stationary, there exist two positive constants $c = c(\beta, \lambda) > 0$ and $\kappa = \kappa(\lambda) > 0$ such that

$$\mathrm{P}(|X_t| > x) \sim c x^{-2\kappa}, \qquad x \to \infty. \tag{1.5}$$

When $\lambda \in (0,1)$, $\kappa < \kappa_0$, that's to say, $\{X_t\}$ has **heavier** tail than $\{X_t^{(0)}\}$ in this sense; while when $\lambda > 1$, $\kappa > \kappa_0$, $\{X_t\}$ has **lighter** tail.

(D) Extremes. Let $M_n = \max(|X_1|, \cdots, |X_n|)$. When $\{X_t\}$ is strictly stationary, there exists a positive constant $\theta = \theta(\beta, \lambda)$ such that

$$\lim_{n\to\infty} \mathrm{P}(n^{-1/2\kappa} M_n \leq x) = \exp\{-c\theta x^{-2\kappa}\}, \qquad x > 0, \tag{1.6}$$

where c and κ are the same as in **(C)**.

The positive constant δ in **(A)** depends on the mixing structure. The constants c and κ depend on the mixing structure as well as the parameters β and λ. We will discuss this in detail in section 2. Two examples, $\{Z_t\}$ in the form of a discrete mixture of normals and the other a continuous mixture, are given in section 3. Some numerical results are also given in this section.

2. Properties of ARCH(1) process with mixing normal noise

Mixtures of measures or distributions occur frequently in the theory and applications of probability and statistics. It has drawn a lot of attentions from

both statisticians and economitrians for more than half a century; an early reference is Robbins (1948).

Firstly, we describe the mixing structure considered in this paper. Let $\mathcal{F} = \{\phi(x,y), y \in H\}$ be a family of probability density functions, where H is some index set, and for each $y \in H$,

$$\phi(x,y) = (\sqrt{2\pi}\sigma(y))^{-1} \exp\{-x^2/2\sigma^2(y)\}, \qquad x \in \mathbb{R}$$

is the probability density function of a normal distribution with zero mean and variance $\sigma^2(y) > 0$. Moreover, let $G(\cdot)$ be a probability measure (or, distribution) on H. We define a new probability density function $f(x)$ as follows:

$$f(x) = \int_{y \in H} \phi(x,y) dG(y), \qquad x \in \mathbb{R}. \tag{2.1}$$

Then, $f(x)$ is a mixing normal probability density function, or, a mixture of normal distributions. Obviously, such a probability density function $f(x)$ has zero mean, and the corresponding variance is

$$\sigma_f^2 = \int_{x \in \mathbb{R}} x^2 f(x) dx = \int_{y \in H} \sigma^2(y) dG(y).$$

Furthermore, for convenience of comparison, we assume that $f(x)$ has unit variance and finite moments of all orders, that is,

$$\int_{y \in H} \sigma^2(y) dG(y) = 1, \tag{2.2}$$

$$\int_{y \in H} \sigma^{2k}(y) dG(y) < \infty, \qquad k \in \mathbb{N}. \tag{2.3}$$

Throughout this section, we assume that $\{X_t\}$ is an ARCH-MN(1) process defined by (1.1), where the noise $\{Z_t\}$ is a sequence of independent and identically distributed random variables having a common probability density function $f(x)$ defined by (2.1).

We have the following Lemma.

Lemma 2.1. *Let Z be a mixture of normal random variables which has a probability density function $f(x)$ defined by (2.1). Assume that condition (2.2) holds. Then,*

$$\mathrm{E}[\ln(\lambda Z^2)] < 0 \tag{2.4}$$

for any $\lambda \in (0, 2e^{\gamma+\delta})$, where $\gamma \approx 0.5772$ is Euler's constant, and

$$\delta = -\int_{y\in H} \ln \sigma^2(y) dG(y) \geq 0 \tag{2.5}$$

is a constant. $\delta = 0$ if and only if G is a trivial (singular) measure.

Proof. Notice that for the standard normal random variable $Z^{(0)}$,

$$\mathrm{E}[\ln(Z^{(0)})^2] = -\gamma - \ln 2 < 0,$$

where γ is Euler's constant. And if $\widetilde{Z} \sim N(0, \sigma^2)$, then,

$$\mathrm{E}(\ln Z^2) = \ln \sigma^2 - \gamma - \ln 2.$$

So, by the definition of $f(x)$, we have

$$\begin{aligned}
\mathrm{E}[\ln(\lambda Z^2)] &= \ln \lambda + \int_{x\in\mathbb{R}} \ln x^2 f(x) dx \\
&= \ln \lambda + \int_{y\in H} \int_{x\in\mathbb{R}} \ln x^2 \phi(x, y) dx dG(y) \\
&= \ln \lambda + \int_{y\in H} [\ln \sigma^2(y) - \gamma - \ln 2] dG(y) \\
&= \ln \lambda - \gamma - \ln 2 + \int_{y\in H} \ln \sigma^2(y) dG(y) \\
&= \ln \lambda - \gamma - \ln 2 - \delta,
\end{aligned}$$

where, by Jensen's inequality and equation (2.2),

$$\delta = -\int_{y\in H} \ln \sigma^2(y) dG(y) \geq -\ln\Big[\int_{y\in H} \sigma^2(y) dG(y)\Big] = 0,$$

and the equation holds if and only if G is a trivial measure. Therefore, $\mathrm{E}[\ln(\lambda Z^2)] < 0$ for any $\lambda \in (0, 2e^{\gamma+\delta})$. The proof is completed. ■

Now, we have the following theorem on the stationarity of the ARCH-MN(1) model (1.1). The probability measure $G(\cdot)$ is assumed to be nontrivial.

Theorem 2.1. **(Stationarity)** *Let $\{X_t\}$ be the ARCH-MN(1) process defined by (1.1) and the noise $\{Z_t\}$ be a sequence of independent and identically distributed random variables with a common distribution $f(x)$ defined by (2.1). Assume that conditions (2.2) and (2.3) hold. Then, when $\lambda \in (0, 2e^{\gamma+\delta})$, where γ is Euler's constant and δ is a positive constant defined by (2.5), there exists some random variable X such that*

$$X_t \stackrel{\mathrm{d}}{\to} X, \qquad t \to \infty. \tag{2.6}$$

Moreover, if we choose $X_0 \stackrel{\mathrm{d}}{=} X$ and be independent of $\{Z_t\}$, then $\{X_t\}$ is strictly stationary.

Proof. Notice that the squared process $\{X_t^2\}$ satisfies the following stochastic recurrence equation,

$$X_t^2 = A_t + B_t X_{t-1}^2 = \beta Z_t^2 + \lambda Z_t^2 X_{t-1}^2, \qquad t \in \mathbb{N}, \tag{2.7}$$

where $(A_t, B_t) = (\beta Z_t^2, \lambda Z_t^2)$ is a sequence of independent and identically distributed random vectors. Furthermore, by Lemma 2.1, when $\lambda \in (0, 2e^{\gamma+\delta})$,

$$\mathrm{E}[\ln^+(\beta Z_t^2)] \leq \ln^+ \beta + \mathrm{E}(Z_t^2) < \infty \qquad \text{and} \qquad \mathrm{E}[\ln(\lambda Z_t^2)] < 0. \tag{2.8}$$

Therefore, by Proposition 8.4.3 in Embrechts *et al.* (1997), the proof is completed. ■

Remark 2.1. It's well known that the mixture of normal distribution has a kurtosis greater than 3, in which sense, it has a heavier (fatter) tail than the normal distribution. At first sight, it might be felt that the ARCH(1) process with a noise having heavier tail would diverge more readily than that with a lighter tail noise, thereby leading to a narrower stationarity region. However, comparing Theorem 2.1 with Theorem 1.1, we can see that the ARCH-MN(1)

process $\{X_t\}$ defined by (1.1) has a wider stationarity region than the normal ARCH(1) process $\{X_t^{(0)}\}$.

Before considering the tail behaviour of $\{X_t\}$, we give another lemma on some properties of the mixture of normal distribution.

Lemma 2.2. *Let Z be a mixture of normal random variable which has a probability density function $f(x)$ defined by (2.1). Assume that conditions (2.2) and (2.3) hold. For $\lambda \in (0, 2e^{\gamma+\delta})$, where the constants γ and δ are defined as in Lemma 2.1, define*

$$h(u) = \mathrm{E}(\lambda Z^2)^u, \qquad u \geq 0.$$

Then,

$$h(u) = \frac{(2\lambda)^u}{\sqrt{\pi}}\Gamma\Big(u+\frac{1}{2}\Big)\int_{y\in H}(\sigma^2(y))^u dG(y). \tag{2.9}$$

The function h is strictly convex in u, and there exists a unique solution $\kappa = \kappa(\lambda)$ to the equation $h(u) = 1$. Moreover,

$$\kappa(\lambda)\begin{cases} >1, & \text{if } \lambda \in (0,1), \\ =1, & \text{if } \lambda = 1, \\ <1, & \text{if } \lambda \in (1, 2e^{\gamma+\delta}), \end{cases} \tag{2.10}$$

and

$$\mathrm{E}[(\lambda Z^2)^\kappa \ln(\lambda Z^2)] > 0. \tag{2.11}$$

Proof. By symmetry of the mixture of normal density $f(x)$ and integration by parts, we have

$$\begin{aligned} h(u) &= \mathrm{E}(\lambda Z^2)^u = \lambda^u \mathrm{E}[(Z^2)^u] \\ &= \frac{\lambda^u}{\sqrt{2\pi}}\int_{y\in H}\int_{x\in\mathbb{R}}(x^2)^u\sigma^{-1}(y)\exp\{-x^2/2\sigma^2(y)\}dx\,dG(y) \\ &= \frac{2\lambda^u}{\sqrt{2\pi}}\int_{y\in H}\int_0^\infty(\sigma^2(y))^u(2z)^{u-1/2}e^{-z}dz\,dG(y) \\ &= \frac{(2\lambda)^u}{\sqrt{\pi}}\Gamma\Big(u+\frac{1}{2}\Big)\int_{y\in H}(\sigma^2(y))^u dG(y), \end{aligned}$$

which giving (2.9), By (2.3), $h(u)$ is finite.

Notice that $h(0) = 1$ for all λ. Furthermore, h has derivatives of all orders. In particular,

$$h'(u) = \mathrm{E}[(\lambda Z^2)^u \ln(\lambda Z^2)], \tag{2.12}$$

$$h''(u) = \mathrm{E}[(\lambda Z^2)^u (\ln(\lambda Z^2))^2] > 0. \tag{2.13}$$

Equation (2.13) implies that h is strictly convex on $\mathbb{R}$. By Lemma 2.1, $h'(0) = \mathrm{E}[\ln(\lambda Z^2)] < 0$ for $\lambda \in (0, 2e^{\gamma+\delta})$. Moreover, it's obvious that $h(u) \to \infty$ as $u \to \infty$; see Figure 2.1. Therefore, there exists a unique $\kappa > 0$ such that $h(\kappa) = 1$, and $h'(\kappa) > 0$, which, together with (2.12), gives (2.11).

Finally, since $h(1) = \lambda$, (2.10) follows by a monotonicity argument. The proof is thus completed. ■

Now, we are well prepared to give the following theorem on the tail behaviour of $\{X_t\}$.

Theorem 2.2. (Tail behaviour) *Let $\{X_t\}$ be an ARCH-MN(1) process defined by (1.1) with parameters $\beta > 0$ and $\lambda \in (0, 2e^{\gamma+\delta})$, where constants γ and δ are defined as in Lemma 2.1, and let the noise $\{Z_t\}$ be a sequence of independent and identically distributed random variables with a common distribution $f(x)$ defined by (2.1). Assume that conditions (2.2) and (2.3) hold. Then,*

$$\mathrm{P}(|X_t| > x) \sim c x^{-2\kappa}, \qquad x \to \infty, \tag{2.14}$$

where

$$c = \frac{\mathrm{E}[((\beta + \lambda X^2)^\kappa - (\lambda X^2)^\kappa)(Z^2)^\kappa]}{\kappa \mathrm{E}[(\lambda Z^2)^\kappa \ln(\lambda Z^2)]} \in (0, \infty) \tag{2.15}$$

for a mixture of normal random variable Z with density $f(x)$, and be independent of $X \stackrel{\mathrm{d}}{=} X_0$.

Proof. The proof is complicated but similar to the proofs of Lemma 8.4.13, Lemma 8.4.14 and Theorem 8.4.12 in Embrechts *et al.* (1997), and is thus omitted. ■

Finally, we proceed to compare the tail behaviour of the ARCH-MN(1) process (1.1) with that of the normal ARCH(1) process (1.2). That is, we compare the κ in Theorem 2.1 with the κ_0 in Theorem 1.1. Such a κ or κ_0 may be called the **tail probability exponent**.

Theorem 2.3. **(Comparison)** *(i) When* $\lambda \in (0, 1)$*, both processes are stationary and the tail probability exponents* κ *and* κ_0 *exist. Moreover,* $1 < \kappa < \kappa_0$.

(ii) When $\lambda \in (1, 2e^{\gamma})$*, both processes are stationary and the tail probability exponents* κ *and* κ_0 *exist. Moreover,* $0 < \kappa_0 < \kappa < 1$.

(iii) When $\lambda \in [2e^{\gamma}, 2e^{\gamma+\delta})$*, process (1.1) is stationary and the tail probability exponent* κ *exists, and* $0 < \kappa < 1$*; process (1.2) is non-stationary.*

Proof. Define $h_0(u) = \mathrm{E}[\lambda(Z^{(0)})^2]^u$, where $Z^{(0)}$ is a standard normal random variable. The unique positive solution for $h_0(u) = 1$ is denoted by κ_0. By Lemma 8.4.6 of Embrechts *et al.* (1997),

$$h_0(u) = \frac{(2\lambda)^u}{\sqrt{\pi}} \Gamma\Big(u + \frac{1}{2}\Big). \tag{2.16}$$

Comparing $h(u)$ with $h_0(u)$, we have

$$h(u) = h_0(u) \int_{y \in H} (\sigma^2(y))^u dG(y).$$

By Jensen's inequality and condition (2.2), the multiplier is

$$\int_{y \in H} (\sigma^2(y))^u dG(y) > \Big(\int_{y \in H} \sigma^2(y) dG(y)\Big)^u = 1, \qquad \text{when } u > 1,$$
$$\int_{y \in H} (\sigma^2(y))^u dG(y) < \Big(\int_{y \in H} \sigma^2(y) dG(y)\Big)^u = 1, \qquad \text{when } u < 1.$$

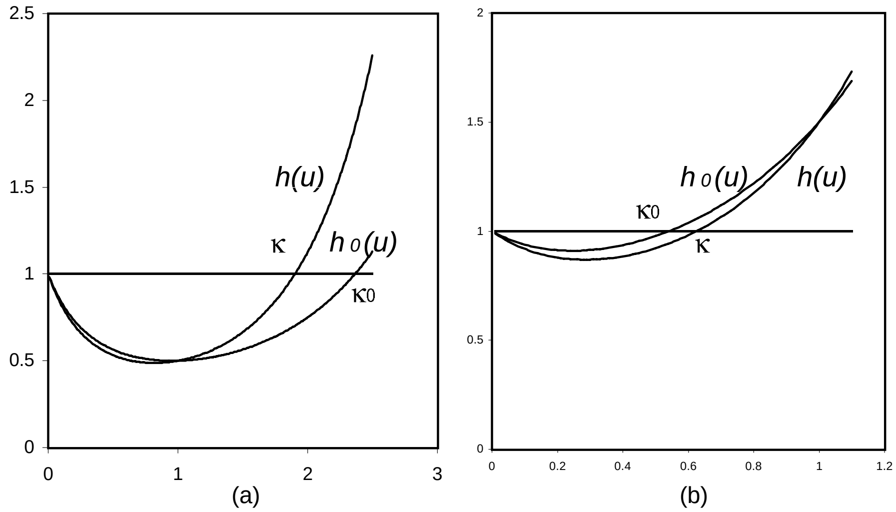

Figure 2.1: *The tail probability exponents* κ *and* κ_0 *for different parameter values* $\lambda = 0.5$ *(a) and* $\lambda = 1.5$ *(b).*

Therefore, $h(u) < h_0(u)$ when $u < 1$ and $h(u) > h_0(u)$ when $u > 1$, see Figure 2.1. Notice that $\kappa(\lambda) > 1$ for $\lambda < 1$ and $\kappa(\lambda) < 1$ for $\lambda > 1$, then, the proof can be completed by a monotonicity argument. ■

Remark 2.2. We add the univariance condition (2.2) to the mixture of normal distribution $f(x)$ just for convenience of comparison. Actually, if this condition fails to hold, that is, $\sigma_f^2 \neq 1$, we may define

$$\widetilde{Z}_t = \sigma_f^{-1} Z_t, \qquad \widetilde{\beta} = \sigma_f^2 \beta, \qquad \widetilde{\lambda} = \sigma_f^2 \lambda.$$

and the model (1.1) can be rewrite as

$$X_t = \widetilde{Z}_t \sqrt{\widetilde{\beta} + \widetilde{\lambda} X_{t-1}^2}, \qquad t \in \mathbb{N}, \tag{2.17}$$

where $\{\widetilde{Z}_t\}$ has zero mean and unit variance. Since $\lambda Z^2 = \widetilde{\lambda} \widetilde{Z}^2$, all results in this section still hold for model (2.17) with δ changed by $\widetilde{\delta} = \delta + \ln \sigma_f^2$.

Remark 2.3. The comparison result (1.6) in **(D)** on extremes is actually an analogue to (1.4) in (d), the normal ARCH case, and is thus omitted in this paper. One can also see Hann *et al.* (1989) for a detailed proof.

3. Examples

In this section, we give some examples of the ARCH-MN(1) process (1.1) in which the noise are dicrete, and respectively, continuous mixtures of normals.

3.1. Examples for discrete mixture

Firstly, we consider the ARCH(1) process (1.1) with a discrete mixture of normal noise. For simplicity, we let $f(x)$ be a mixture of two different normal distributions. The mixing structure, including the index set H, the probability measure $G(y)$ on H and corresponding $\sigma^2(y)$ (or, $\sigma(y)$), is defined as follows.

$$\begin{cases} H = \{1,2\}, \\ dG(\{1\}) = p = 1 - dG(\{2\}), \qquad 0 < p < 1, \\ 0 < \sigma(1) < 1 < \sigma(2). \end{cases} \tag{3.1}$$

In this setting, condition (2.3) automatically holds. In order that condition (2.2) holds, we choose

$$p = \frac{\sigma^2(2) - 1}{\sigma^2(2) - \sigma^2(1)}.$$

For such a mixture of normals distribution, we can immediately find that the constant δ in (2.5) is

$$\delta = -\frac{(\sigma^2(2) - 1)\ln\sigma^2(1) + (1 - \sigma^2(1))\ln\sigma^2(2)}{\sigma^2(2) - \sigma^2(1)}, \tag{3.2}$$

and the fundction $h(u)$ is

$$h(u) = \frac{(2\lambda)^u}{\sqrt{\pi}}\Gamma\Big(u + \frac{1}{2}\Big)\frac{(\sigma^2(2) - 1)\sigma^{2u}(1) + (1 - \sigma^2(1))\sigma^{2u}(2)}{\sigma^2(2) - \sigma^2(1)}. \tag{3.3}$$

For either a standard normal noise or a mixture of normal noise, the equation $h(u) = 1$ cannot be solved explicitly, but numerical solutions can be found as shown in Table 3.1 for several different cases. In Table 3.1, when $\sigma^2(1) = \sigma^2(2) = 1$, the noise is standard normal; we include the normal case into this table for the purpose of comparison. The symbol “NS” in the table means that

the process is non-stationary and therefore the tail probability exponent κ does not exist. The stationarity regions are also given in the table in the form of the upper bound $2e^{\gamma+\delta}$.

$\sigma^2(1)$	1	0.5	0.25	0.1
$\sigma^2(2)$	1	2	4	10
$2e^{\gamma+\delta}$	3.562145	4.487951	8.183531	23.436172
λ	$\kappa(\lambda)$	$\kappa(\lambda)$	$\kappa(\lambda)$	$\kappa(\lambda)$
0.1	13.247	7.516	4.506	2.825
0.3	4.181	2.913	2.084	1.623
0.5	2.366	1.903	1.526	1.309
0.7	1.587	1.415	1.247	1.147
0.9	1.153	1.113	1.069	1.041
1.0	1.000	1.000	1.000	1.000
1.5	0.542	0.627	0.758	0.855
2.0	0.311	0.414	0.604	0.761
2.5	0.170	0.275	0.491	0.692
3.0	0.076	0.176	0.404	0.637
3.5	0.008	0.103	0.333	0.591
4.0	NS	0.045	0.274	0.551
5.0	NS	NS	0.180	0.485
6.0	NS	NS	0.109	0.430
7.0	NS	NS	0.053	0.383
8.0	NS	NS	0.008	0.342
10.0	NS	NS	NS	0.271
12.0	NS	NS	NS	0.213
14.0	NS	NS	NS	0.163
16.0	NS	NS	NS	0.120
18.0	NS	NS	NS	0.082
20.0	NS	NS	NS	0.049
23.0	NS	NS	NS	0.006

Table 3.1 *Tail probability exponents for some ARCH(1) processes with discrete mixture of normal noises.*

We can clearly see from Table 3.1 that the ARCH-MN(1) process (1.1) has a wider stationarity region than the normal ARCH(1) process (1.2). The larger is the variance ratio $\sigma^2(2)/\sigma^2(1)$, the wider is the stationarity region. The tail probability shrinks faster for $\lambda > 1$ but more slowly for $\lambda \in (0, 1)$.

3.2. Examples for continuous mixture

In this subsection, we consider two different cases of mixture.

The first one we want to study is a uniform mixture in which we let the variances of the normal distributions for the mixture be uniformally distributed in some interval $H = (0, b)$, where $b > 0$ is a constant. In other words, we let $\sigma^2(y) = y$ for $y \in (0, b)$, and the mixture measure $G(\cdot)$ has a density $g(y) = b^{-1}$ for $y \in (0, b)$. In order that condition (2.2) holds, we choose $b = 2$, and condition (2.3) automatically holds in this case.

We can readily find that

$$\delta = -\frac{1}{2}\int_0^2 \ln y dy = 1 - \ln 2 > 0.$$

So, the stationarity region is $\lambda \in (0, 2e^{\gamma+\delta}) \approx (0, 4.8415)$. Moreover, the function $h(u)$ is

$$\begin{aligned} h(u) &= h_0(u)\int_0^2 \frac{y^u}{2}dy \\ &= \frac{2^u}{u+1}h_0(u) = \frac{(4\lambda)^u}{(u+1)\sqrt{\pi}}\Gamma\Big(u+\frac{1}{2}\Big). \end{aligned}$$

The numerical solution for the tail probability exponent κ can be found in Table 3.2 for several different parameter values.

λ	0.5	1.0	1.5	2.0	3.0	4.0	8.0
normal case	2.366	1.000	0.542	0.311	0.076	NS	NS
uniform mixture	2.000	1.000	0.623	0.417	0.192	0.069	NS
normal mixture	1.566	1.000	0.748	0.597	0.418	0.310	0.104

Table 3.2 *Tail probability exponents for some ARCH(1) processes with continuous mixture of normal noises.*

Another mixture case is the following normal mixture. The mixing structure is defined as

$$H = \mathbb{R}, \qquad \sigma^2(y) = y^2, \qquad \text{and} \qquad G(\cdot) \sim N(0,1).$$

In this case, the constant δ is

$$\delta = -\int_{y\in\mathbb{R}} \ln(y^2)dG(y) = -\mathrm{E}[\ln(Z^{(0)})^2] = \gamma + \ln 2 > 0,$$

where $Z^{(0)} \sim N(0,1)$ and γ is Euler's constant. So, the stationarity region is $\lambda \in (0, 2e^{\gamma+\delta}) \approx (0, 12.6889)$. Moreover, the function $h(u)$ is

$$h(u) = h_0(u) \int_{y \in \mathbb{R}} (y^2)^u dG(y) = h_0(u) \mathrm{E}[(Z^{(0)})^2]^u = \frac{(4\lambda)^u}{\pi} \Gamma^2\Big(u + \frac{1}{2}\Big).$$

Some numerical results are also included in Table 3.2.

REFERENCES

Bollerslev, T. (1986) Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, **31**, 307–327.

Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997) *Modelling Extremal Events.* Springer-Verlag Berlin Heidelberg.

Engle, R.F. (1982) Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, **50**, 987–1007.

Fama, E. (1965) The behavior of stock market prices. *Journal of Business*, **38**, 34–105.

Hann, L. de, Resnick, S.I., Rootzźen, H. and Vries, C.G. de (1989) Extremal behaviour of solutions to a stochastic difference equation with applications to ARCH processes. *Stoch. Proc. Appl.*, **32**, 213–224.

Mittnik, S. and Rachev, S.T. (1993) Modeling asset returns with alternative stable distributions. *Econometric Reviews*, **12**, 261–330.

Robbins, H. (1948) Mixture of distributions. *Ann. Math. Stat.*, **19**, 360–369.

Address:
Zhiqiang Zhang
Department of Statistics and Actuarial Science
The University of Hong Kong
Pokfulam, Hong Kong