Heavy Tail Behavior and Parameters Estimation of GARCH (1, 1) Process

Hailong Chen[1, 2] and Chunli Liu[2]

[1]*College of Computer Science and Technology, Harbin Engineering University, Harbin China*
[2]*School of Computer Science and Technology, Harbin University of Science and Technology, Harbin China*
hrbustchl@163.com

Abstract

In practice, Financial Time Series have serious volatility cluster, that is large volatility tend to be concentrated in a certain period of time, and small volatility tend to be concentrated in another period of time. While GARCH models can well describe the dynamic changes of the volatility of financial time series, and capture the cluster and heteroscedasticity phenomena. At the beginning of this paper, the definitions and basic theories of GARCH(1,1) models are discussed. Secondly, show the heavy tail behavior of GARCH(1,1) process with α *-stable residuals* $\{\varepsilon_t\}_{t\in Z}$ *,* $\alpha \in (0,2]$ *and* $\{\varepsilon_t\}_{t\in Z}$ *errors. In fact, both these processes have heavy-tailed properties, but generally the tail of GARCH(1,1) process is heavier than the tail of* $\{\varepsilon_t\}_{t\in Z}$ *errors. And then the modification of maximum likelihood function has been constructed as the theoretical basis of this study, make use of Holder inequality and Jensen's inequality to estimate parameters of GARCH(1,1) model with residuals having regularly varying distributions with index* $\alpha > 0$*. Finally, the consistency and asymptotic normality of the estimates constructed are further proved.*

Keywords: *GARCH process; errors; regularly varying distribution; stable distribution; heavy-tailed distribution; parameter estimation*

1. Introduction

In the analysis of financial data, the best-known and most often used processes are Autoregressive Conditionally Heteroskedastic (ARCH) and its extensions generalizations – Generalized Autoregressive Conditionally Heteroskedastic (GARCH) processes, see F. Engle [1], T. Bollerslev [2].

It is known that, provided the errors distribution has finite fourth moment, quasi-maximum likelihood estimators for ARCH/GARCH processes are asymptotically normally distributed with the standard rate $\sqrt{n}$. When the errors have heavy tail probability, parameters estimation of GARCH process has investigated by T. Mikosch and D. Straumann in [3]. They showed the consistency and asymptotic distribution of quasimaximum likelihood estimation. In paper [4], we consider a modification of quasi–maximum likelihood estimator for GARCH (1,1) process with the errors, whose squares have regularly varying tail with index $\alpha, \alpha > 0$. We showed that, this estimator is unbiased and asymptotically normal with standard convergence rate of $n^{1/2}$ regardless of whether the errors are heavy–tailed. In the case of GARCH(1,1) process, this

estimator is applied for a larger class of GARCH (1,1) processes with heavy–tailed errors than P. Hall and Q. Yao in [5] and D. Straumann in [6].

This paper introduces the basic definitions and related theories of GARCH(1,1) models as theoretical basis for this study; Secondly, show the heavy tail behavior of GARCH(1,1) process with α-stable residuals $\{\varepsilon_t\}_{t\in Z}$, $\alpha \in (0,2]$ and $\{\varepsilon_t\}_{t\in Z}$ errors. And then the modification of maximum likelihood function has been constructed, in order to assess with the index for $\alpha > 0$ change regularly distributed disturbance GARCH (1,1) model parameters, and finally prove that the structure further assess the credibility and asymptotically normal distribution.

2. GARCH (1,1) Models

Let's consider GARCH (1,1) model of the following type

$$y_t = \sigma_t \varepsilon_t, \ \ t \in Z, \tag{1}$$

where

$$\sigma_t^2 = \omega_0(1-\beta_0) + \alpha_0 y_{t-1}^2 + \beta_0 \sigma_{t-1}^2, \tag{2}$$

$$\omega_0 > 0, \ \ \alpha_0 > 0, \ \ 0 \le \beta_0 < 1, \tag{3}$$

$\{\varepsilon_t\}$ are independent identically distributed regularly varying random variables with index $\alpha > 0$.

Let $\theta_0 = (\omega_0, \alpha_0, \beta_0)$ be unknown parameters of the model, the parametric space

$$\Theta = \{\theta = (\omega, \alpha, \beta) : 0 < \omega_1 \le \omega \le \omega_2, 0 < \alpha_1 \le \alpha \le \alpha_2, 0 < \beta_1 \le \beta \le \beta_2 < 1, E\ln(\beta + \alpha\varepsilon_0^2) < 0\},$$

θ_0 be internal point of space Θ, then the parametric model under review can be expressed as $y_t = \sigma_t(\theta)\varepsilon_t, \ \ t \in Z,$ moreover following the performance of

$$\sigma_t^2(\theta) = \omega + \alpha \sum_{k=0}^{\infty} \beta^k y_{t-k-1}^2, \ t \in Z. \tag{4}$$

3. The Nature of Heavy Tails

Both stable distributions and GARCH(1,1) process have heavy-tailed properties, We will introduce below:

3.1. The Nature of Heavy Tails of Stable Distributions

Let us consider independent, identically distributed α- stable random variables $\{\varepsilon_t\}_{t\in Z}$, $\alpha \in (0,2]$, having characteristic function:

$$\Psi_{\varepsilon_1}(\mathrm{u}) = i\mu u - \sigma^\alpha |u|^\alpha + i\sigma^\alpha t\omega(u,\alpha,\beta), \ u \in R,$$

where $\beta \in [-1,1], \sigma > 0, \mu \in R,$

$$\omega(u,\alpha,\beta) = \begin{cases} |u|^{\alpha-1} \beta tg(\alpha\pi/2), & \alpha \ne 1, \\ -2\beta \ln|u|/\pi, & \alpha = 1. \end{cases}$$

Then let us denote $\varepsilon_1 \sim S_\alpha(\sigma,\beta,\mu)$.

In the paper [7], we know, that if

$$\varepsilon_1 \sim S_\alpha(\sigma,\beta,\mu),\ 0<\alpha\le 2,$$

then

$$P(\varepsilon_1 < x) \sim \begin{cases} |x|^\alpha \dfrac{1-\beta}{2}\sigma^\alpha \dfrac{1-\alpha}{\Gamma(2-\alpha)\cos(\pi\alpha/2)}, & \alpha\ne 1, \\ |x|^\alpha \dfrac{1-\beta}{2}\sigma^\alpha \dfrac{2}{\pi}, & \alpha = 1 \end{cases} \quad \text{when } x\to -\infty \tag{5}$$

and

$$P(\varepsilon_1 > x) \sim \begin{cases} x^{-\alpha} \dfrac{1+\beta}{2}\sigma^\alpha \dfrac{1-\alpha}{\Gamma(2-\alpha)\cos(\pi\alpha/2)}, & \alpha\ne 1, \\ x^{-\alpha} \dfrac{1+\beta}{2}\sigma^\alpha \dfrac{2}{\pi}, & \alpha = 1 \end{cases} \quad \text{when } x\to +\infty \tag{6}$$

where $f(x)\sim g(x)$ with $x\to\infty$ means, that

$$\lim_{x\to\infty}[f(x)/g(x)] = const.$$

In this case, it is said that α-stable distributions have heavy-tailed properties with α index, and α index is referred to as the heavy tail index (Please see [8, 9]).

3.2. The Nature of Heavy Tails of GARCH(1,1) Process

Let us consider a stationary process (1), $\varepsilon_t \sim S_\alpha(1,0,0)$, $\alpha\in(0,2]$. In paper [10] it has been proven, that if formula:

$$EA_1^{k/2} = 1,\ k\in R, \tag{7}$$

has only one positive root $k=k_0$, then the process (1) is a regularly varying process with index k_0 (Please see [11]), i.e. there exists a positive constant C_0, such that:

$$P(\sigma > x) \sim C_0 x^{-k_0} \quad \text{when } x\to\infty \tag{8}$$

and

$$P(|y|>x) \sim E|\varepsilon_0|^{k_0}\ P(\sigma>x) \text{ when } x\to\infty, \tag{9}$$

Let us prove, that in case α-stable residuals, $\alpha\in(0,2]$, GARCH(1,1) process is a regularly varying process.

Theorem *GARCH(1,1) process defined by relation (1), is a regularly varying process with a certain index k_0, on condition that $\{\varepsilon_t\}_{t\in Z}$ α - stable random variables, $\alpha\in(0,2]$.*

Proof. Indeed, if the $\{\varepsilon_t\}_{t\in Z}$ residuals are α-stable random variables, then $\{A_t = \beta_0 + \alpha_0\varepsilon_{t-1}^2\}_{t\in Z}$ – regularly varying random variables with index $\alpha/2\le 1$. Therefore, there always exists constant $h_0>0$, such that $EA_1^h<\infty$ for all $h<h_0$ and $EA_1^{h_0}=\infty$. On the other hand, out of $E(\ln A_1)<0$ condition, it follows, that $EA_1<1$. Using the continuity of $f(k)=EA_1^{k/2}$ function, we conclude, that formula

$$EA_1^{k/2} = 1, k \in R,$$

always has only one positive root $k = k_0$. Therefore, on condition, that $\{\varepsilon_t\}_{t\in Z}$ residuals – α -stable random variables, relations (8) and (9) are satisfied, *i.e.*, $\{y_t\}_{t\in Z}$ is a regularly varying process with index k_0.

Comment In case, when $\{\varepsilon_t\}_{t\in Z}$ errors have standard normal distribution ($\alpha = 2$) and $\alpha_0 + \beta_0 = 1$ (IGARCH(1,1) process), formula (7) has only one root $k = 2$, *i.e.*, IGARCH (1,1) process is a regularly varying process with index 2:

$$P(\sigma > x) \sim C_0 x^{-2},$$

$$P(|y| > x) \sim C_0 x^{-2}$$

when $x \to \infty$.

Thus, IGARCH (1,1) process with $\{\varepsilon_t\}_{t\in Z}$ residuals, having standard normal distribution, has heavy-tailed properties with index $k = 2$.

Out of the conclusion, that $\{\varepsilon_t\}_{t\in Z}$ – α -stable random variables, we obtain that $\{A_t = \beta_0 + \alpha_0 \varepsilon_{t-1}^2\}_{t\in Z}$ are regularly varying random variables with index $\alpha / 2$, therefore, $E(A_t)^k < \infty$ if and only if $k < \alpha / 2$. Thus, formula (7) has only one root $k = k_0 < \alpha$. Or if stated differently, the tail of GARCH(1,1) process is heavier than the tail of $\{\varepsilon_t\}_{t\in Z}$ errors. Thus, an extreme behavior of GARCH (1,1) process is conditional on an extreme behavior of the errors. (see Figure 1 and Figure 2).

However, the heavy tail behavior of GARCH (1,1) process is less different from the heavy tail behavior of errors, when heavy tail index converges to 2, especially, in case of IGARCH(1,1) process. Indeed, $f(k) = EA_1^{k/2} = E(\beta_0 + \alpha_0 \varepsilon_0^2)^{k/2} = 1$ is a continuous convex function, therefore, when $\alpha \to 2$, the root of formula (7) converges to the root of $E(\beta_0 + \alpha_0 \xi^2)^{k/2} = 1$ formula, where $\xi \sim N(0,1)$. In case of IGARCH (1,1) process with residuals, having standard normal distribution, $E(\beta_0 + \alpha_0 \xi^2)^{k/2} = 1$ has only one root 2 *i.e.*, IGARCH (1,1) process index k_0 converges to 2 and coincides with heavy tail index of $\{\varepsilon_t\}_{t\in Z}$ residuals when $\alpha \to 2$. Thus, an extreme behavior of GARCH (1,1) process is similar to an extreme behavior of $\{\varepsilon_t\}_{t\in Z}$ residuals when $\alpha \to 2$. (see Figure 3 and Figure 4).

Figure 1. 700 Realizations of $\{\varepsilon_t\}_{t\in Z}$ Residual Process with $\alpha = 1,5$ and GARCH(1,1) Process with $\omega_0 = 0,2,\ \alpha_0 = 0,2,\ \beta_0 = 0,4$ Parameters

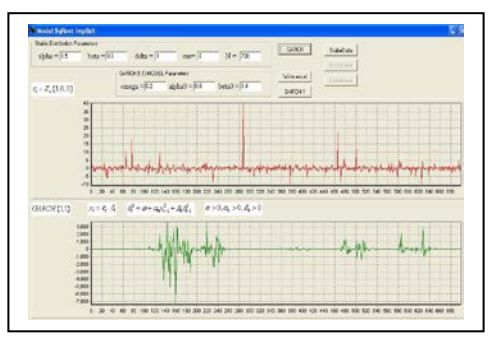

Figure 2. 700 Realizations of $\{\varepsilon_t\}_{t\in Z}$ Residual Process with $\alpha = 1,5$ and IGARCH(1,1) Process with $\omega_0 = 0,2,\ \alpha_0 = 0,6,\ \beta_0 = 0,4$ Parameters

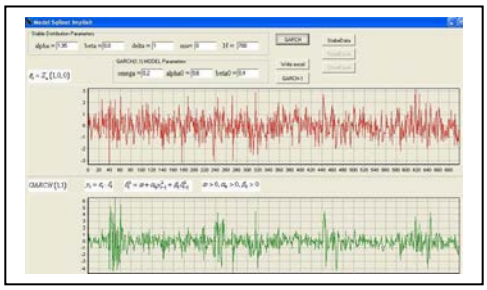

Figure 3. 700 Realizations of $\{\varepsilon_t\}_{t\in Z}$ Residual Process with $\alpha = 1{,}95$ and GARCH(1,1) Process with $\omega_0 = 0{,}2,\ \alpha_0 = 0{,}2,\ \beta_0 = 0{,}4$ Parameters

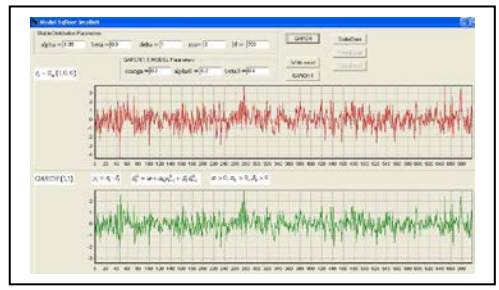

Figure 4. 700 Realizations of $\{\varepsilon_t\}_{t\in Z}$ Residual Process with $\alpha = 1{,}95$ and IGARCH(1,1) Process with $\omega_0 = 0{,}2,\ \alpha_0 = 0{,}6,\ \beta_0 = 0{,}4$ Parameters

4. The Proof of Parameters Estimation Theory

Lemma 1. *Whenever* $i = \overline{1,n}$ *, we will have the following inequality*

$$\frac{\sigma_t^2}{\sigma_{t-i}^2} \le \prod_{j=1}^{i} (1 + \beta_0 + \alpha_0 \varepsilon_{t-j}^2). \tag{10}$$

Proof. We will obtain the proof using mathematical induction method. For $i = 1,$ using (2), (3) we will obtain the following

$$\frac{\sigma_t^2}{\sigma_{t-1}^2} = \frac{\omega_0(1-\beta_0) + \alpha_0 y_{t-1}^2 + \beta_0 \sigma_{t-1}^2}{\sigma_{t-1}^2} =$$

$$= \frac{\omega_0(1-\beta_0)}{\sigma_{t-1}^2} + \alpha_0 \varepsilon_{t-1}^2 + \beta_0 \le$$

$$\le 1 + \beta_0 + \alpha_0 \varepsilon_{t-1}^2. \tag{11}$$

Let's assume that lemma is correct where $i = n-1,$ then

$$\frac{\sigma_t^2}{\sigma_{t-n+1}^2} \le \prod_{j=1}^{n-1} (1 + \beta_0 + \alpha_0 \varepsilon_{t-j}^2). \tag{12}$$

Let's prove that lemma is correct for $i = n$. Using (2), (3) we will obtain the following

$$\frac{\sigma_t^2}{\sigma_{t-n}^2} = \frac{\sigma_t^2}{\sigma_{t-n+1}^2} \times \frac{\sigma_{t-n+1}^2}{\sigma_{t-n}^2} =$$

$$= \frac{\sigma_t^2}{\sigma_{t-n+1}^2} \times \frac{\omega_0(1-\beta_0) + \alpha_0 y_{t-n}^2 + \beta_0 \sigma_{t-n}^2}{\sigma_{t-n}^2} =$$

$$= \frac{\sigma_t^2}{\sigma_{t-n+1}^2} \times \left[\frac{\omega_0(1-\beta_0)}{\sigma_{t-n}^2} + \alpha_0 \varepsilon_{t-n}^2 \frac{\sigma_{t-n}^2}{\sigma_{t-n}^2} + \beta_0 \frac{\sigma_{t-n}^2}{\sigma_{t-n}^2} \right] =$$

$$= \frac{\sigma_t^2}{\sigma_{t-n+1}^2} \times \left[\frac{\omega_0(1-\beta_0)}{\sigma_{t-n}^2} + \alpha_0 \varepsilon_{t-n}^2 + \beta_0 \right] \le$$

$$\leq \frac{\sigma_t^2}{\sigma_{t-n+1}^2} \times (1+\beta_0+\alpha_0\varepsilon_{t-n}^2).$$

From (12) we will obtain

$$\frac{\sigma_t^2}{\sigma_{t-n}^2} \leq \prod_{j=1}^{n}(1+\beta_0+\alpha_0\varepsilon_{t-j}^2).$$

The lemma is proved.

Lemma 2. *Whenever any natural number* $M>0$, *we will have inequality*

$$\frac{\sigma_t^2(\theta)}{\sigma_t^2} \geq \frac{\alpha}{\prod_{j=1}^{M+1}(1+\beta_0+\alpha_0\varepsilon_{t-j}^2)} \sum_{k=0}^{M}\beta^k\varepsilon_{t-k-1}^2.$$

Proof. It follows from (4) that

$$\frac{\sigma_t^2}{\sigma_t^2(\theta)} = \frac{\sigma_t^2}{\omega+\alpha\sum_{k=0}^{\infty}\beta^k y_{t-k-1}^2} \leq \frac{\sigma_t^2}{\alpha\sum_{k=0}^{M}\beta^k y_{t-k-1}^2}, \tag{13}$$

whenever any natural number $M>0$.

Using lemma 1 and (13) we will obtain the following

$$\frac{\sigma_t^2(\theta)}{\sigma_t^2} \geq \frac{\alpha\sum_{k=0}^{M}\beta^k y_{t-k-1}^2}{\sigma_t^2} = \alpha\sum_{k=0}^{M}\beta^k\varepsilon_{t-k-1}^2\frac{\sigma_{t-k-1}^2}{\sigma_t^2} \geq$$

$$\geq \alpha\sum_{k=0}^{M}\beta^k\varepsilon_{t-k-1}^2\frac{1}{\prod_{j=1}^{k+1}(1+\beta_0+\alpha_0\varepsilon_{t-j}^2)} \geq$$

$$\geq \frac{\alpha}{\prod_{j=1}^{M+1}(1+\beta_0+\alpha_0\varepsilon_{t-j}^2)}\sum_{k=0}^{M}\beta^k\varepsilon_{t-k-1}^2.$$

Lemma is proved.

Theorem 1. *Whenever* λ, $0<\lambda<\alpha$, *we will have*

$$E\left\{\sup_{\theta\in\Theta}\frac{\sigma_t^2}{\sigma_t^2(\theta)}\right\}^{\lambda} < \infty. \tag{14}$$

Proof. Using Holder inequality and lemma 2, whenever p', $\lambda<p'<\alpha$, we will have

$$E\left\{\sup_{\theta\in\Theta}\frac{\sigma_t^2}{\sigma_t^2(\theta)}\right\}^{\lambda} \leq \frac{1}{\alpha^{\lambda}}E\left[\left(\prod_{j=1}^{M+1}(1+\beta_0+\alpha_0\varepsilon_{t-j}^2)\right)^{\lambda}\left(\frac{1}{\sum_{k=0}^{M}\beta^k\varepsilon_{t-k-1}^2}\right)^{\lambda}\right] \leq$$

$$\leq \frac{1}{\alpha^{\lambda}}\left\{E\left(\prod_{j=1}^{M+1}(1+\beta_0+\alpha_0\varepsilon_{t-j}^2)\right)^{p'}\right\}^{\lambda/p'} \times \left\{E\left(\sum_{k=0}^{M}\beta^k\varepsilon_{t-k-1}^2\right)^{-\lambda p'/(p'-\lambda)}\right\}^{(p'-\lambda)/p'}. \tag{15}$$

From the independence of random variables $\{\varepsilon_t\}$, whenever any natural number $M > 0$, we will obtain

$$\left\{E\left(\prod_{j=1}^{M+1}(1+\beta_0+\alpha_0\varepsilon_{t-j}^2)\right)^{p'}\right\}^{\lambda/p'} = \left(\prod_{j=1}^{M+1}E(1+\beta_0+\alpha_0\varepsilon_{t-j}^2)^{p'}\right)^{\lambda/p'} < \infty. \tag{16}$$

Based on the above lemmas since $\{\varepsilon_t\}$ constitutes regularly varying random variables with index α, $\{\varepsilon_t^2\}$ have regularly varying distribution with index $\alpha/2$. Then $\sum_{k=0}^{M}\beta^k\varepsilon_{t-k-1}^2$ constitutes regularly varying random variable with index $\alpha/2$. From lemma 3.1 [12] we have

$$E\left(\sum_{k=0}^{M}\beta^k\varepsilon_{t-k-1}^2\right)^{-\lambda p'/(p'-\lambda)} < \infty. \tag{17}$$

Using (15)-(17) we will obtain (14).
The theorem is proved.
Let us consider the modified likelihood function in the following form

$$L_n(\theta) = \frac{1}{n}\sum_{t=1}^{n} l_t(\theta),$$

$$l_t(\theta) = -\left[\ln\sigma_t^2(\theta) + \frac{1}{p}\left(\frac{y_t^2}{\sigma_t^2(\theta)}\right)^p\right], \tag{18}$$

p – a certain constant, $0 < p < \alpha$.

Theorem 2. *Let* $L(\theta) = El_0(\theta)$, *where* $l_0(\theta) = -\left[\ln\sigma_0^2(\theta) + \frac{1}{p}\left(\frac{y_0^2}{\sigma_0^2(\theta)}\right)^p\right]$, *then with any* p, $0 < p < \alpha$, *we have* $L(\theta) < \infty$.

Proof. Let us prove that $L(\theta) < \infty$ when all $\theta \in \Theta$. Indeed, we will obtain from theorem 1 the following

$$E\left(\frac{y_0^2}{\sigma_0^2(\theta)}\right)^p = E\left(\varepsilon_t^2\right)^p E\left(\frac{\sigma_0^2}{\sigma_0^2(\theta)}\right)^p < \infty, \tag{19}$$

where $t = 0, 1, 2, \ldots$.

Therefore, it follows from regularly varying random variables $\{\varepsilon_t\}$ and paper [13] that there exists constant ρ, $\rho > 0$, such that

$$E[\sigma_0^2(\theta)]^\rho < \infty.$$

It follows from here that using Jensen inequality we will have

$$E\ln\sigma_t^2(\theta) < \infty. \tag{20}$$

From (19) and (20) we will obtain the required.
The theorem is proved.

Let $\hat{\theta}_n$ be M estimator of θ_0 parameter of (1)-(3) model with likelihood function (18), *i.e.*,

$$\hat{\theta}_n = \arg\max_{\theta\in\Theta} L_n(\theta). \tag{21}$$

The consistency of $\hat{\theta}_n$ estimator is proved in the following theorem.

Theorem 3. *Let us assume that,* $0 < p < \alpha$, $E|\varepsilon_0^2|^p = 1$, *then*

$$\hat{\theta}_n \xrightarrow{\text{a.s.}} \theta_0 \ \ with \ \ n \to \infty.$$

The theorem is being proved in the same way as in paper [4].

Theorem 4. *If* $\{\varepsilon_t\}$ *are independent identically distributed regularly varying random variables with index* $\alpha > 0$, *then for* $0 < p < \alpha$, *we will have*

$$\sup_{\theta\in\Theta}\{|L_n(\theta) - L(\theta)|\} \xrightarrow{\text{a.s.}} 0 \ \ with \ \ n \to \infty,$$

$$\sup_{\theta\in\Theta}\{|L_n'(\theta) - L'(\theta)|\} \xrightarrow{\text{a.s.}} 0 \ \ with \ \ n \to \infty,$$

$$\sup_{\theta\in\Theta}\{|L_n''(\theta) - L''(\theta)|\} \xrightarrow{\text{a.s.}} 0 \ \ with \ \ n \to \infty.$$

Proof. We know that process (18) is measurable function of strictly stationary and ergodic process $\{\varepsilon_t\}$. That is why $\{l_t(\theta)\}_{t\in Z}$ is also an ergodic process with a finite expectation $L(\theta) = El_0(\theta) < \infty$, $\theta \in \Theta$. Using ergodic theorem we obtain that for all $\theta \in \Theta$

$$L_n(\theta) \xrightarrow{\text{a.s.}} L(\theta) \ \ with \ \ n \to \infty.$$

Using theorem 3.5.8 [14] we come to the conclusion that $\left\{\sup_{u,v\in\Theta}\left\{\frac{1}{|u-v|}|l_t(u) - l_t(v)|\right\}\right\}_{t\in Z}$ is stationary, ergodic process with first-order stopping time. Using the ergodic theorem we will obtain

$$\sup_{u,v\in\Theta}\left\{\frac{1}{|u-v|}|L_n(u) - L_n(v)|\right\} \le \frac{1}{n}\sum_{t=1}^{n}\left[\sup_{u,v\in\Theta}\left\{\frac{1}{|u-v|}|l_t(u) - l_t(v)|\right\}\right] \overset{a.s.}{=} O(1).$$

From this it follows that function $\{L_n(\theta)\}$ is uniformly continuous with a probability of 1 and converge to $L(\theta)$ with $n \to \infty$. Out of compactness of parametric space Θ we have that $\{L_n(\theta)\}$ uniformly converge to $L(\theta)$ a.s. (almost surely).

Out of definition $L_n(\theta)$ in (18) we have that

$$L_n'(\theta) = \frac{1}{n}\sum_{t=1}^{n}\left[\left(\frac{y_t^2}{\sigma_t^2(\theta)}\right)^p - 1\right]\frac{(\sigma_t^2(\theta))'}{\sigma_t^2(\theta)}, \tag{22}$$

$$L_n''(\theta) = \frac{1}{n}\sum_{t=1}^{n}\left\{\left[\left(\frac{y_t^2}{\sigma_t^2(\theta)}\right)^p - 1\right]\frac{(\sigma_t^2(\theta))''}{\sigma_t^2(\theta)} - \left[1 + (p-1)\left(\frac{y_t^2}{\sigma_t^2(\theta)}\right)^p\right]\left(\frac{(\sigma_t^2(\theta))'}{\sigma_t^2(\theta)}\right)^T\frac{(\sigma_t^2(\theta))'}{\sigma_t^2(\theta)}\right\}. \tag{23}$$

Using lemmas 1, 2, as in case with proving theorem 1 [4], we will obtain that

$$E\left\{\sup_{\theta\in\Theta}\left|\frac{\left(\sigma_t^2(\theta)\right)'}{\sigma_t^2(\theta)}\right|\right\}^{\lambda}<\infty \text{ and } E\left\{\sup_{\theta\in\Theta}\left|\frac{\left(\sigma_t^2(\theta)\right)''}{\sigma_t^2(\theta)}\right|\right\}^{\lambda}<\infty \text{ when all } \theta\in\Theta.$$

From (22) and (23) we will obtain that $E\left\{\sup_{\theta\in\Theta}\left|L'(\theta)\right|\right\}<\infty$ and $E\left\{\sup_{\theta\in\Theta}\left|L''(\theta)\right|\right\}<\infty$ when all $\theta\in\Theta$. Using ergodic theorem we will have that $L_n'(\theta)\xrightarrow{\text{a.s}}L'(\theta)$ and $L_n''(\theta)\xrightarrow{\text{a.s.}}L''(\theta)$ when $n\to\infty$ for all $\theta\in\Theta$.

In a similar manner as in case with proving lemma 2 [4], we will obtain that

$$\sup_{\theta\in\Theta}|L_n'(\theta)-L'(\theta)|\xrightarrow{\text{a.s.}}0 \text{ and } \sup_{\theta\in\Theta}\{|L_n''(\theta)-L''(\theta)|\}\xrightarrow{\text{a.s.}}0 \ \textit{when}\ n\to\infty.$$

The theorem is proved.

From (21) we obtain that $L_n(\theta)$ achieves the maximum in point $\hat{\theta}_n$, that is why for n that are large enough $\hat{\theta}_n$ is the internal point of compact space Θ. Then there exists $n_0\in N$, such that $L_n'(\hat{\theta}_n)=0$ with any $n\geq n_0$. It follows that

$$L_n'(\hat{\theta}_n)-L_n'(\theta_0)=-L_n'(\theta_0).$$

Using mean value theorem we have

$$(\hat{\theta}_n-\theta_0)L_n''(\eta)=L_n'(\hat{\theta}_n)-L_n'(\theta_0)=-L_n'(\theta_0), \tag{24}$$

where $\eta\in\Theta$, so that

$$|\eta-\hat{\theta}_n|\leq|\hat{\theta}_n-\theta_0|,\ |\eta-\theta_0|\leq|\hat{\theta}_n-\theta_0|.$$

Using the same method as with proving theorem 1 and (3), we will obtain

$$E\,\|l_0''(\theta)\|<\infty \text{ when all } \theta\in\Theta.$$

From the continuity of the function $L_n''(.)$ and theorem [4], we will obtain

$$L_n''(\eta)=L_n''(\theta_0)+o(1). \tag{25}$$

I. Berkes and L. Horváth in their work proved that $E\left[\left(\frac{(\sigma_0^2(\theta_0))'}{\sigma_0^2}\right)^T\frac{(\sigma_0^2(\theta_0))'}{\sigma_0^2}\right]$ is a non-singular matrix.

We know that $\sigma_0^2(\theta_0)=\sigma_0^2$, then under conditions of theorem [4], using (24) and independence of random variables ε_t and $\frac{\sigma_t^2}{\sigma_t^2(\theta)}$, we will have the following

$$L_n''(\theta_0)=El_0'(\theta_0)=-E\left[1+(p-1)\left|\varepsilon_0^2\right|^p\right]$$

$$E\left[\left(\frac{(\sigma_0^2(\theta_0))'}{\sigma_0^2}\right)^T \frac{(\sigma_0^2(\theta_0))'}{\sigma_0^2}\right].$$

Using similar method as in case with random process $\{l_t'(\theta)\}_{t\in Z}$, we will obtain that $\{l_t''(\theta_0)\}_{t\in Z}$ is strictly stationary and ergodic process. It follows that

$$L_n''(\theta_0) \xrightarrow{\text{a.s.}} L''(\theta_0), \text{ when } n\to\infty, \tag{26}$$

Using (22), we will have

$$L_n'(\theta_0) = \frac{1}{n}\sum_{t=1}^{n}\left(1-|\varepsilon_t^2|^p\right)\frac{(\sigma_t^2(\theta_0))'}{\sigma_t^2}. \tag{27}$$

From (24)-(27) we will obtain the following

$$n^{1/2}(\hat{\theta}_n-\theta_0)(B+o(1)) = n^{-1/2}\sum_{t=1}^{n}\left(1-|\varepsilon_t^2|^p\right)\frac{(\sigma_t^2(\theta_0))'}{\sigma_t^2}, \tag{28}$$

Theorem 5. *Let's assume that conditions of theorem from paper [4] are satisfied, then for* $p<\alpha/4$ *we will have*

$$n^{1/2}(\hat{\theta}_n-\theta_0) \xrightarrow{d} N(0, B^{-1}\Sigma B^{-1}) \text{ when } n\to\infty, \tag{29}$$

where

$$\Sigma = (E(\varepsilon_0^4)^p - 1)A,$$

$$A = E\left[\left(\frac{(\sigma_0^2(\theta_0))'}{\sigma_0^2}\right)^T \frac{(\sigma_0^2(\theta_0))'}{\sigma_0^2}\right],\ B = El_0''(\theta_0).$$

Proof. Indeed, when $p<\alpha/4$, in view of the condition of $E|\varepsilon_t^2|^p=1$ and independence of random variables ε_t and $\frac{(\sigma_t^2(\theta_0))'}{\sigma_t^2}$ we have $E|l_t'(\theta_0)|^2<\infty$, and $El_t'(\theta_0)=0$ when any $t\in Z$. It follows from $\left\{l_t'(\theta_0)=\left(1-|\varepsilon_t^2|^p\right)\frac{(\sigma_t^2(\theta_0))'}{\sigma_t^2}\right\}_{t\in Z}$ being stationary martingale difference process, that $\{l_t'(\theta_0)\}_{t\in Z}$ is ergodic process. If n is large enough, there exists $(B+o(1))^{-1}$ which is equal to $(B^{-1}+o(1))$. On the other hand, matrix $B^{-1}\Sigma B^{-1}$ is a covariance matrix $\left(1-|\varepsilon_t^2|^p\right)\frac{(\sigma_t^2(\theta_0))'}{\sigma_t^2}B^{-1}$. It follows from here that, using (28) and Cramer-Wold theorem [15] we will obtain (29).

The theorem is proved.

5. Conclusions

This paper mainly studies parameters estimation and heavy tail behavior of GARCH (1,1) with the errors having regularly varying distribution. Firstly, we introduce the basic definitions and related theories of GARCH (1,1) models as theoretical basis for this study.

Secondly, show the heavy tail behavior of GARCH(1,1) process with α-stable residuals $\{\varepsilon_t\}_{t\in Z}$, $\alpha \in (0,2]$ and $\{\varepsilon_t\}_{t\in Z}$ errors. And then assess with the index for $\alpha > 0$ change regularly distributed disturbance GARCH (1,1) model parameters and modify maximum likelihood function. Finally, prove the credibility and asymptotically normal distribution of the estimation with related theories.

The related research of this paper will help us analyze deeply some practical problems shown spike and fat-tailed features in the various fields, and make it easier to realize how to correctly describe the degree of heavy-tailed distribution, that is, accurately estimate the tail index of heavy-tailed. Meanwhile, this paper presents parameters estimation and heavy tail behavior of GARCH (1,1) models to provide a reference basis for the study of many practical issues of other financial time series data of the "volatility". For example, GARCH models are applied to the futures, foreign exchange, interest rate and other economic and financial fields.

Acknowledgements

This work was financially supported by Doctoral Fund of Ministry of Education of China (Grant No. 20122303120005), The 51th class General Financial Grant from the China Postdoctoral Science Foundation (Grant No. 2012M510926), Heilongjiang Postdoctoral Financial assistance (Grant No. LBH-Z12069), Science and Technology research project of Heilongjiang Province Management of Education (Grant No. 11541068).

References

[1] R. F. Engle, “Autogressive conditional heteroscedasticity with estimates of the variance of U.K. inflation”, Econometrica, vol. 50, **(1982)**, pp. 987-1007.
[2] T. Bollerslev, “Generalized autoregressive conditional heteroskedasticity”, Journal of Econometrics, vol. 31, **(1986)**, pp. 307-327.
[3] T. Mikosch and D. Straumann, “Whittle estimation in a heavy–tailed GARCH (1,1) model”, Stochastic processes and their applications, vol. 100, **(2002)**, pp. 187-222.
[4] C. Hailong, Y. Chengji and C. Deyun, “Parameters estimator of GARCH(1, 1) process with heavy tailed errors”, Proceedings-2012 7th International Forum on Strategic Technology, IFOST, **(2012)**, pp. 1-5.
[5] P. Hall and Q. Yao, “Inference in ARCH and GARCH models with heavy–tailed errors”, Econometrics, vol. 71, no. 1, **(2003)**, pp. 285-317.
[6] D. Straumann, “Estimation in conditionally heteroscedastic time series models”, Lecture notes in statistics 181. Berlin Heidelberg: Springer, **(2005)**.
[7] R. Weron, “On the Chambers-Mallows-Stuck method for simulating skewed stable random variables”, Statistics & Probability Letters, vol. 28, **(1996)**, pp. 165-171.
[8] J. Feller, “An introduction to probability theory and its applications”, Wiley, New York, vol. 2, **(1966)**, pp. 439-448.
[9] V. V. Uchaikin and V. M. Zolotarev, “Change and stability: Stable distributions and their applications”, Series modern probability and statistics, Utrecht, CSP, vol. 570 P, **(1999)**, pp. 143-150.
[10] P. Bougerol and N. Picard, “Strict stationarity of generalized autoregressive processes”, Ann. Probab., vol. 20, **(1992)**, pp. 1714-1730.
[11] T. Mikosch and C. Stărică, “Limit theory for the sample autocorrelations and extremes of a GARCH(1,1) process”, Ann. Statis., vol. 28, **(2000)**, pp. 1427-1451.
[12] A. H. Jessen and T. Mikosch, “Regularly varying functions”, Publications de L’Institut Mathématique, Tome, vol. 80, no. 94, **(2006)**, pp. 171-192.
[13] I. Berkes and L. Horváth, “GARCH processes: Structure and estimation”, Bernoulli, vol. 9, **(2003)**, pp. 201-227.
[14] W. F. Stout, “Almost sure convergence”, New York: Academic press, **(1974)**.
[15] P. Billingsley, “Convergence of probability measure”, Wiley, New York, **(1968)**.