이학석사학위논문

A Goodness of Fit Test for Threshold GARCH Models

Threshold GARCH 모형에 대한 적합도 검정

2013년 02월

서울대학교 대학원

통계학과

MENG HAN

A Goodness of Fit Test for Threshold GARCH Models

지도교수 이상열

이 논문을 이학석사학위논문으로 제출함

2012년 12월

서울대학교 대학원

통계학과

MENG HAN

MENG HAN의 석사학위논문을 인준함

2013년 2월

위 원 장 ____________________ (인)

부 위 원 장 ____________________ (인)

위 원 ____________________ (인)

A Goodness of Fit Test for Threshold GARCH Models

By

MENG HAN

A Thesis
submitted in fulfillment of the requirement
for the degree of
Master of Science
in Statistics

The Department of Statistics
College of Natural Sciences
Seoul National University

February, 2013

Abstract

In this paper, an entropy based goodness of fit test is applied to conditional heteroscedasticity models such as the threshold GARCH model. (cf. Lee et al. (2011)). The asymptotic distribution of the test statistic is derived under the null hypothesis so that the test can be taken into practice. In this paper, a bootstrap method is applied in the simulation study to deal with the simple vs. simple test and the composite hypothesis case is stated theoretically. In the simulation study, the performance of the test for four pairs of different m, n values and model coefficients are evaluated through Monte Carlo simulations. The result shows an adequate performance of the test. Finally, a real data analysis is conducted by using the stock price of Intel Corporation from January 2009 to November 2012 with 988 observations.

Keywords: Maximum entropy test, threshold GARCH model, bootstrap method, goodness of fit.

Student Number: 2010-24004

Contents

Chapter 1

Introduction

It is well known that the maximum entropy principle is a criterion to select a priori probabilities. The principle of maximum entropy is a prime doctrine in Bayesian probability theory. This theory states that, subject to precisely stated prior data, which must be a proposition that expresses testable information: the probability distribution which represents the current state of knowledge best is the one with largest information-theoretical entropy.

The idea is to let some precisely stated prior data or testable information about a probability distribution function be provided. Consider the set of all trial probability distributions that encode the prior data. Of those, the one that maximizes the information entropy is the proper probability distribution under the given prior data.

The principle was first expounded by E.T. Jaynes in his two papers in 1957 in which he emphasized a natural correspondence between statistical mechanics and information theory. He argued that the entropy of statistical mechanics and the information entropy of information theory is principally identical. Consequently, statistical mechanics should be considered just as a particular application of a general tool of logical inference and information theory. Now, the maximum entropy method is broadly applied to diverse research fields in science and engineering, such as spatial physics, computer vision, natural language processing, and information science.

According to Forte(1984), the Boltzmann-Shannon entropy of X is defined as

$$H(f) = -\int_{-\infty}^{\infty} f(x)\log(f(x))dx,$$

which measures the average amount of uncertainly concerning the value x of a random variable X with probability density $f(x)$. where $f(x)$, $\log(f(x))$ is understood to be zero whenever $f(x)=0$. In the connection with maximum entropy distributions, this form of definition is taken as a standard form. However, there is a related formular of entropy as follows:

$$H_c(p(x) \parallel q(x)) = -\int p(x)\log(\frac{p(x)}{q(x)})dx,$$

where $q(x)$ is proportional to the limiting density of discrete points. The relative entropy is always referred to as the Kullback-Leibler divergence of q from p. The inference principle is to minimum it.

On the other hand, Forte and Hughes(1988) apply the function $-\sum p_i \log(p_i/(x_i - x_{i-1}))$ as an expression of entropy, where $p_i = \Pr[x_{i-1} < X < x_i] = \int_{x_{i-1}}^{x_i} f(x)dx, i = 1,2,\cdots,n-1$ and $a = x_0 < \cdots < x_n = b$ when X is a discrete random variable.

Goodness of fit (GOF) tests measure the degree of agreement between the distribution of an observed data sample and a theoretical statistical distribution. Since then, a large number of GOF methods including the chi-squared test and various empirical distribution function (edf) tests have been developed by many researchers. Also, goodness of fit tests include measures of entropy, divergence, and information, etc. Many of the tests implementing currently such as the likelihood ratio, the chi-squared, the score and Wald tests can be defined in terms of the appropriate measures. Meanwhile, there are some other similar tests include those based on the ϕ-family of Csiszar measures.

Recently, Lee et al. (2011) have proposed a maximum entropy test for iid random variables and demonstrated its advantages over other existing tests. Since this test has merit to outperform several existing GOF tests, we consider applying the maximum entropy test to the threshold GARCH models based on this method in this study. See Lee et al. (2012).

The GOF test in time series models has been an important practical issue. In time series studies, the normality test has been an important issue as the normality of time series guarantees several advantageous properties while non-normal time series does not. So, many researchers have focused on the Gaussian test. On the other hand, a prior information of non-normality is also crucial since a heavy-tailed distribution is ofter required in financial time series modeling. As relevant references, see Lee and Wei (1999) who considered the empirical process GOF test in autoregressive models. And see Bera and Jarque (1981), Bai (2003), and Lee et al. (2010), who dealt with the Gaussian test in GARCH models.

In the next few chapters , we will review the maximum entropy test, introduce the test statistic, study its properties in iid settings, apply this test to threshold GARCH models, perform a simulation study to explore the capabilities of the proposed test statistic. Finally, we conduct a real data analysis as example.

Chapter 2

The Maximum Entropy Test

2.1 The Maximum Entropy Test for a Simple Hypothesis

In this chapter, we firstly start to review the maximum entropy test for iid random variables. For this, the readers are referred to as Lee et al. (2011). Consider $Y_i, i=1,2,\cdots,n$ to be a random sample from a distribution with an unknown distribution function expressed as F. Let's think about this following test of fit:

$$H_0 : F = F_0 \leftrightarrow H_1 : F \neq F_0.$$

For continuous distribution, the following generalization of the Jaynes (1957,1963) maximum entropy expression is proposed:

$$S^{\omega}_{\max}(F) = -\sum_{i=1}^{m} \omega_i (F(s_i) - F(s_{i-1})) \log\left(\frac{F(s_i) - F(s_{i-1})}{s_i - s_{i-1}}\right).$$

where ω's are appropriate weights with $0 \le \omega_i \le 1$ and $\sum_{i=1}^{m} \omega_i = 1$, m is the number of disjoint intervals for partitioning the data range. Moreover, $-\infty < a \le s_1 \le s_2 \le \cdots \le s_m \le b < \infty$ are preassigned partition points. For a properly chosen constant c, the null hypothesis will be rejected if $\max_{\omega} \left| S^{\omega}(F_n) - S^{\omega}(F_o) \right| \ge c$, where F_n is the empirical distribution based on the data sample, namely,

$$F_n(x) = n^{-1} \sum_{i=1}^{n} I(Y_i \le x).$$

The proposed test statistic is certified by the fact that it is closely related to the entropy measure $H(f)$ given above. Indeed, the unweighted form of the entropy expression can be seen as:

$$S_{\max}(F) \approx -\sum_{i-1}^{m} f(s_i)(s_i - s_{i-1})\log(f(s_i))$$

$$\rightarrow -\int_{-\infty}^{\infty} f(x)\log(f(x))dx = -E_F(\log f(x)) \equiv H(f),$$

where f is the probability density function with $\int f dx = F_0$ under the null hypothesis.

If F_0 is the uniform distribution in [0,1], then $S^{\omega}(F_0) = 0$. This goodness of fit test can be reduced to a uniform test by using the probability integral transform $F_0(Y_i) = U_i$, and a uniformity test can be implemented to the U_i's. We can always use this transform of data aim to do GOF test since F_0 is not determined without estimation of parameters.

Using the transformed random variables U_i 's, the test we concerned about can be reduced as:

$$H_0' : U_i \sim U[0,1] \leftrightarrow H_1' : \text{not } H_0'.$$

The theorem below provided by Lee et al.(2011) shows an asymptotic distribution for the test statistic mentioned above.

Theorem Let $Y_1, Y_2, \cdots, Y_n$ be a continuous random sample with a cumulative distribution function F. Under null hypothesis $H_0 : F = F_0$, as $n \rightarrow \infty$, we have

$$T_n := \sqrt{n}\, sup \left| S_{\max}^{\omega}(F_n) \right| \xrightarrow{d} sup \left| \sum_{i=1}^{m} \omega_i (B(s_i) - B(s_{i-1})) \right|,$$

where F_n is the sample distribution based on $U_i = F_0(Y_i)$, $B(s)$ is a Brownian bridge on [0,1]. The supremum is taken among the space of bounded weights $\omega_i : [0,1] \rightarrow [0,1]$ with $\sum_{i=1}^{m} \omega_i = 1$, and $0 = s_0 \le s_1 \le \cdots \le s_m = 1$.

This test turned out to be well performed among various situations according to the simulation results of the study by Lee et al.(2011). However, the role of weights' selection can not be ignored before the implementation of the uniform transformation since the performance of the test can vary with the choice of m even in the simple vs. simple hypothesis test. But in our case, since the data will be transformed into uniform random variables and $s_i = i/m$ will be used to make a uniform spacing of the unit interval. The supremum over all weights will be used to deal with any possibilities pictured by specific alternatives. Markedly, this approach can moderate difficulty in choosing an optimal weight. The issue of computing the supremum will be introduced later. Conventionally, neither too small nor too large, m should be chosen to be much less than the sample size n so that m/n can be close to 0.

2.2 The Maximum Entropy Test for Composite Hypothesis

Besides test for simple hypothesis, a composite test is also widely utilized. It is normal to test the hypothesis that the unknown distribution F belongs to a parametric family $F_{\theta_{\theta \subset \Theta}}$, where Θ is an open set in R^k. In this case, the test problem becomes like this:

$$H_0 : F \in F_0(x;\theta); \theta \in \Theta \leftrightarrow H_1 : \text{not } H_0, \tag{2.2.1}$$

where F_0 ,θ_0 are denoted to be a continuous underlying distribution under the null hypothesis and the true parameter, respectively. Below is the summary of Lee (2013).

Now, suppose that $Y_1, \cdots, Y_n$ are observed sample and $\widehat{\theta_n}$ is an estimator of θ_0 with $n^{1/2}(\widehat{\theta_n} - \theta_0) = Op(1)$ under the null hypothesis.

Let $U_i = F_0(Y_i;\theta_0)$ and $\widehat{U_i} = F_0(Y_i;\widehat{\theta_n})$, and

$$F_n(s) = \frac{1}{n}\sum_{i=1}^{n} I(U_i \leq s), \quad \widehat{F_n(s)} = \frac{1}{n}\sum_{i=1}^{n} I(\widehat{U_i} \leq s), \text{ where } s \in [0,1]. \quad (2.2.2)$$

Moreover, define that

$$V_n(s) = n^{1/2}(F_n(s) - s) \text{ and } \widehat{V_n(s)} = n^{1/2}(\widehat{F_n(s)} - s). \quad (2.2.3)$$

In order to put the maximum entropy test into practice, it is natural to replace F_n with $\widehat{F_n}$ for the test statistic. However, since the difference of parameter estimation can cause different behavior between V_n and $\widehat{V_n}$, which also can lead to a different limiting distribution of T_n. Thus, in order to get a similar result like the theorem mentioned above, we have to eliminate the estimation influence by considering the K-transformation introduced by Khmaladze (1981), which transforms the estimated empirical process to a more tractable process.

The K-transformation has been actuated the fact that

$$W(s) = V(s) + \int_0^s \frac{V(u)}{1-u} du, \quad (2.2.4)$$

where $W(s)$ and $V(s)$, $s \in [0,1]$, denote a standard Brownian motion and a Brownian bridge, respectively. This formular can be deduced by the contention

$$V(s) = (1-s)\int_0^s \frac{dW(u)}{1-u}. \tag{2.2.5}$$

And more general expression can be put as follows. Let a vector $g(r) = (r, g_1(r), \cdots, g_p(r))'$ be a real-valued function on [0,1] with $\int_0^1 g^{2_j}(s)ds < \infty, j = 1,2,\cdots,p,$ so that $C(s) = \int_s^1 \dot{g}(x)\dot{g}(x)'dx$ is invertible for all $0 \le s < 1$. Then, according to Khmaladze (1981),

$$W(s) = V(s) - \int_0^s [\dot{g}(u)' C^{-1}(u) \int_u^1 \dot{g}(\tau) dV(\tau)] du \tag{2.2.6}$$

forms a standard Brownian motion. When $g(r) = r$, the previous result can be obtained. What's more, if $V_n(s), 0 \le s \le 1$, is a sequence of stochastic processes such that $V_n \Rightarrow V$ in D[0,1]. Then

$$W_n(s) = V_n(s) - \int_0^s [\dot{g}(u)' C^{-1}(u) \int_u^1 \dot{g}(\tau) dV_n(\tau)] du \Rightarrow W(s) \tag{2.2.7}$$

This result can be easily extended to the stochastic process $\widehat{V_n}$ in D[0,1] with

$$\widehat{\upsilon_n}(r) = V_n(r) + C_n(r) + \Delta_n(r) \tag{2.2.8}$$

with $V_n \Rightarrow V, C_n(r) = -\overline{g}(r)'a_n$ and $sup_{0 \le r < 1}|\Delta_n(r)| = op(1)$ where $\overline{g}: [0,1] \rightarrow R^p$ is a real-valued function with $\int_0^1 \| \dot{\overline{g(r)}} \|^2 dr < \infty$ and $\overline{g}(0) = 0$, while a_n is a sequence of random variables with $a_n = Op(1)$. Let $g(r) = (r, \overline{g}')'$ and assume that $C(s)$ is invertible for every $0 < s \le 1$. Then, since we have $C_n(r) = -g(r)'b_n$ and $b_n = (0, a_n')'$,

$$\int_0^s [\dot{g}(u)' C^{-1}(u) \int_u^1 \dot{g}(\tau) dC_n(\tau)] du = -\int_0^s [\dot{g}(\tau)\dot{g}(\tau)' b_n d\tau] du \tag{2.2.9}$$
$$= g(s)'b_n = C_n(s).$$

Considering (2.2.7), we have

$$\widehat{W}_n(s) = \widehat{V}_n(s) - \int_0^s [\dot{g}(u)' C^{-1}(u) \int_u^1 \dot{g}(\tau) d\widehat{V}_n(\tau)] d\tau \Rightarrow W. \qquad (2.2.10)$$

In fact, if we let $h(s) = \frac{\partial F_{\theta_0}}{\partial \theta}(F_{\theta_0}^{-1}(s))$ be set, we can get the conclusion under the null hypothesis as

$$\widehat{V}_n(s) = V_n(s) - h(s)' n^{1/2}(\widehat{\theta}_n - \theta_0) + op(1). \qquad (2.2.11)$$

Using the asymptotic result on the K-transformation, we use another expression $\widehat{\mathscr{F}}_n$ instead of $\widehat{F}_n$ and defined as

$$\widehat{\mathscr{F}}_n(s) = \widehat{F}_n(s) - \int_0^s \dot{g}(u)' C^{-1}(u) \int_u^1 \dot{g}(v) d\frac{\widehat{V}_n(v)}{n^{1/2}} du, \qquad (2.2.12)$$

where for each s, $C(s) = \int_s^1 \dot{g}(u)\dot{g}(u)' du$ with $g(s) = (s, h(s)')'$ is a nonsingular $(d+1) \times (d+1)$ matrix. Though (2.2.10), we can see that

$$\widehat{W}_n(s) = n^{1/2}(\widehat{\mathscr{F}}_n(s) - s) \qquad (2.2.13)$$

converges to a standard Brownian motion weakly under the null hypothesis. This result can be used to obtain the theorem below.

Theorem Let $Y_1, Y_2, \cdots, Y_n$ be a random sample from a continuous cumulative distribution with the distribution function expressed as F. Under$H_0 : F \in F_0(x;\theta)$, (2.2.11) holds and according to (2.2.13) we can have the result with representing a standard Brownian motion as W:

$$\widehat{T}_n := \sqrt{n}\, sup_{\{\omega \in W\}} \left| S^{\omega}_{\max}(\widehat{\mathscr{F}}_n) \right| \xrightarrow{d} sup \left| \sum_{i=1}^{m} \omega_i (W(s_i) - W(s_{i-1})) \right|,$$

where $\omega_i > 0, \sum_{i=1}^{m} \omega_i = 1$ and $0 = s_0 \le s_1 \le \cdots \le s_m = 1$.

We have already known that $\log(x) = 0$ if $x \le 0$. There is a possibility that $\widehat{\mathscr{F}}_n(s_{i+1}) < \widehat{\mathscr{F}}_n(s_i)$, probability of which can be negligible asymptotically.

In order to imply this test into practice, let us consider $\omega_i^{(l)}, l=1,2,\cdots,L$, independent and identically distributed random variables from $U[0,1]$, which are also independent from $U_i \sim U[0,1]$, where L is a fixed positive integer. Then ,if we let

$$\omega_{li} = \frac{\omega_i^{(l)}}{\omega_1^{(l)} + \cdots + \omega_m^{(l)}},$$

as $L \to \infty$,

$$\max_{1 \le l \le L} \left| \sum_{i=1}^{m} \omega_{li} (W(s_i) - W(s_{i-1})) \right| \xrightarrow{d} sup \left| \sum_{i=1}^{m} \omega_i (W(s_i) - W(s_{i-1})) \right|.$$

Subsequently, by taking $s_i = i/m$ for convenience, we can use this as the maximum entropy test statistic

$$\widehat{T}_n = n^{1/2} \max_{1 \le l \le L} \left| \sum_{i=1}^{m} \omega_{li} \left(\widehat{\mathscr{F}}_n\left(\frac{i}{m}\right) - \widehat{\mathscr{F}}_n\left(\frac{i-1}{m}\right)\right) \log m \left(\widehat{\mathscr{F}}_n\left(\frac{i}{m}\right) - \widehat{\mathscr{F}}_n\left(\frac{i-1}{m}\right)\right) \right|$$

$$\overset{d}{\approx} sup \left| \sum_{i=1}^{m} \omega_i \left(W\left(\frac{i}{m}\right) - W\left(\frac{i-1}{m}\right)\right) \right|.$$

The random variable in the above line has the same distribution as that of $m^{1/2} sup \left| \sum_{i=1}^{m} \omega_i Z_i \right|$ with $Z_i \sim N(0,1)\, i.i.d.$, the distribution of which can be approximated by $m^{1/2} \max_l \left| \sum_{i=1}^{m} \omega_{li} Z_i \right|$.

Chapter 3
The Threshold GARCH Model

The threshold GARCH (or TGARCH) model is a volatility model commonly used to handle leverage effects; see Glosten, Jagannathan and Runkle (1993) and Zakoian (1994). A threshold GARCH(m,s) model assumes the form

$$X_t = \sigma_t \epsilon_t$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{s} (\alpha_i + \gamma_i N_{t-i}) X_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2,$$

where N_{t-i} is an indicator for negative X_{t-i}, that is

$$N_{t-i} = \begin{cases} 1 \text{ if } X_{t-i} < 0 \\ 0 \text{ if } X_{t-i} \geq 0 \end{cases},$$

and $\alpha_i, \gamma_i, \beta_i$ are nonnegative parameters satisfying conditions similar to those to GARCH models. From the model, it is obvious that a positive X_{t-i} contributes $\alpha_i X_{t-i}^2$ to σ_t^2, whereas a negative X_{t-i} has a larger impact $(\alpha_i + \gamma_i) X_{t-i}^2$ with $\gamma_i > 0$. The model uses 0 as its threshold to separate the impacts of past shocks. This model is also called the GJR model because Glosten et al. (1993) proposed essentially the same model. The main difference between the GJR model and the threshold model is that GJR model is on conditional standard deviation and the threshold model is on conditional variance. In this paper, we mainly use TGARCH(1,1) model to put the maximum entropy into practice. There are several ways to express threshold GARCH models. Consider the model stated this way:

$$X_t = \sigma_t \epsilon_t$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 (X_{t-1}^+)^2 + \alpha_2 (X_{t-1}^-)^2 + \beta \sigma_{t-1}^2,$$

where $X_{t-1}^+ = \begin{cases} X_{t-1} & \text{if } X_{t-1} > 0 \\ 0 & \text{if } X_{t-1} \le 0 \end{cases}$ and $X_{t-1}^- = \begin{cases} X_{t-1} & \text{if } X_{t-1} \le 0 \\ 0 & \text{if } X_{t-1} > 0 \end{cases}$.

And for all, ϵ_t are iid random variables with $E\epsilon_1 = 1$ and $E\epsilon_1^2 = 1$, $E\epsilon_1^4 < \infty$. What's more $\theta = (\alpha_0, \alpha_i, \beta)'$ with $\alpha_0 > 0, \alpha_i \ge 0, \beta \ge 0$ and $\sum_{i=1}^{2} \alpha_2 + \beta < 1$ are in a compact subset of R^4. The latter constraint on $\alpha_i + \beta$ implies that the unconditional variance of X_t is finite, whereas its conditional variance σ_t^2 may evolve over time.

Similar to GARCH models, ϵ_t is often assumed to follow a standard normal or standardized Students' t-distribution of generalized error distribution.

Here, we want to test:

$$H_0 : \epsilon_t \sim F_0 \quad vs. \quad H_1 : \text{not } H_0,$$

where our main motivation is to test whether F_0 denotes a standard normal density most of the time.

According to the conditions of the model, it is obvious that if the result of $\alpha_i + \beta$ nearly equals 1, the inference is incorrect unless the sample size is considerably large. In that case, GOF test usually performs size distortions. In order to conquer this problem, here recommend the bootstrap method. This method is a non-parameter Monte Carlo method. It estimates the population's distribution and inferences by resampling the observed data. The procedure in Stute et al. (1993) and Lee et al. (2012) is as follows:

1. Based on the observed data $X_1, X_2, \cdots, X_n$, we can obtain the Quasi Maximum Likelihood Estimation $\widehat{\theta_n}$ as model parameters.
2. Generate $\epsilon_1^*, \cdots, \epsilon_n^*$ that follow $F_n(\bullet)$ and construct $X_1^*, \cdots, X_n^*$ obtained through $X_t^* = \sigma_t \epsilon_t^*$ with $\widehat{\theta_n}$ taking the place of θ by letting the initial X_0^* to be 0. Then, calculate the test statistic value $\widehat{T_n}$ based on these X_n^*'s.
3. Repeating the above procedure, say after B times, we can get the value of $\widehat{T_n}$ based on the original observations and the $100(1-\alpha)\%$ percentile.
4. Reject H_0 if $\widehat{T_n}$ obtained is larger than the $100(1-\alpha)\%$ percentile.

Chapter 4

Simulation

In this chapter, for the sake of identifying how the maximum entropy test performs, we propose a simulation study. The empirical sizes of our study are calculated as the number of rejections of the null hypothesis out of 1000 repetitions with the nominal levels of 0.01, 0.05 and 0.10. Using the bootstrap method mentioned at the end of chapter 3, we evaluate the test with B=500 and use threshold GARCH(1,1) models' coefficients $\theta=$(0.1, 0.1, 0.1, 0.1), (0.2, 0.2, 0.2, 0.2), (0.1, 0.3, 0.4, 0.1) and (0.05, 0.1, 0.7, 0.1) for (n, m)=(300, 5), (400, 7), (500, 8) and (1000, 10). Table 4-1 shows the results. The table shows the empirical sizes are close to the nominal level 0.05 and no size distortions appears. Since it is said H_0 should be rejected if $\widehat{T_n}$ obtained is larger than the $100(1-\alpha)\%$ percentile, we can also obtain whether the null hypothesis is rejected due to the calculation results at certain levels. The test is considered to perform adequately.

<Table 4-1 Simulation Result>

$\widehat{\theta}_n$	$n=$, $m=$	$\widehat{T}_n$	$100(1-\alpha)\%$			Size
			0.01	0.05	0.10	
(0.1,0.1,0.1,0.1)	(1000, 10)	5.1526	5.6342	5.1553	4.9063	0.052
	(500, 8)	4.4273	4.4885	4.0996	3.9515	0.046
	(400, 7)	4.8853	5.3318	4.9368	4.6984	0.064
	(300,5)	3.7062	3.7017	3.5800	3.4856	0.041
(0.2,0.2,0.2,0.2)	(1000, 10)	2.6516	2.7105	2.3909	2.2729	0.047
	(500, 8)	2.2607	2.5154	2.2778	2.1576	0.074
	(400, 7)	2.1094	2.3967	2.1849	2.0497	0.086
	(300, 5)	2.4896	2.6840	2.5688	2.3891	0.078
(0.1,0.3,0.4,0.1)	(1000, 10)	1.8266	2.0226	1.8884	1.7933	0.053
	(500, 8)	1.9839	2.0553	1.9247	1.8194	0.048
	(400, 7)	1.9949	2.2677	1.9966	1.9112	0.052
	(300, 5)	1.9637	1.9716	1.8975	1.8623	0.044
(0.05,0.1,0.7,0.1)	(1000, 10)	1.8097	1.9883	1.7944	1.7105	0.049
	(500, 8)	2.1144	2.2286	1.9845	1.9162	0.043
	(400, 7)	1.8228	1.9335	1.7842	1.6799	0.046
	(300, 5)	2.349	2.6941	2.4160	2.3203	0.074

Chapter 5

Real Data Analysis

In this chapter, we apply the maximum entropy test to the stock price of Intel Corporation from January 2009 to November 2012 with 988 observations. Figure 5-1, Figure 5-2 and Figure 5-3 display the daily prices, log return series, the acf and pacf of the log return series, respectively. It is considered that there is some volatility clustering phenomenon. Also, the Ljung-Box statistics $Q(12) = 16.4984(0.1695)$ for X_t and $Q(12) = 242.2057(< 2.2e-16)$ for X_t^2, which in the parenthesis is their p values. This result confirms no serial correlations in the data and conditional heteroscedasticity is satisfied.

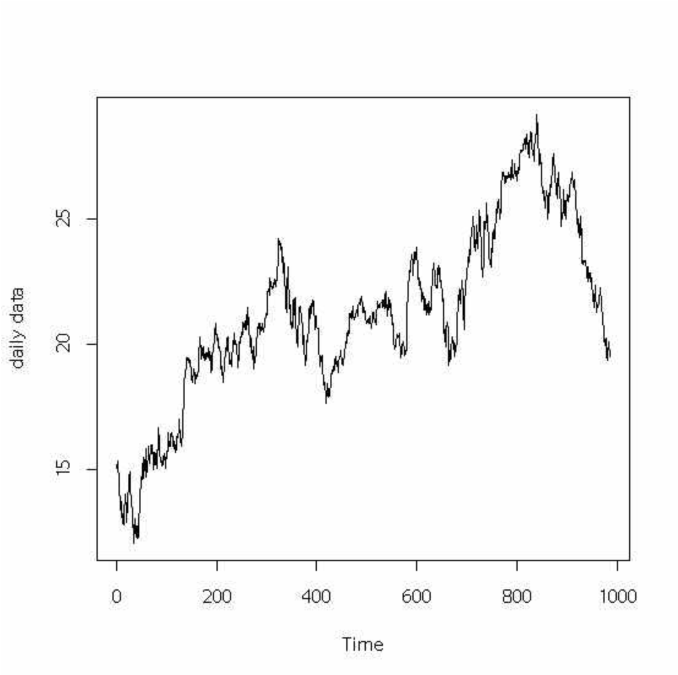

Figure 5-1 Daily Price

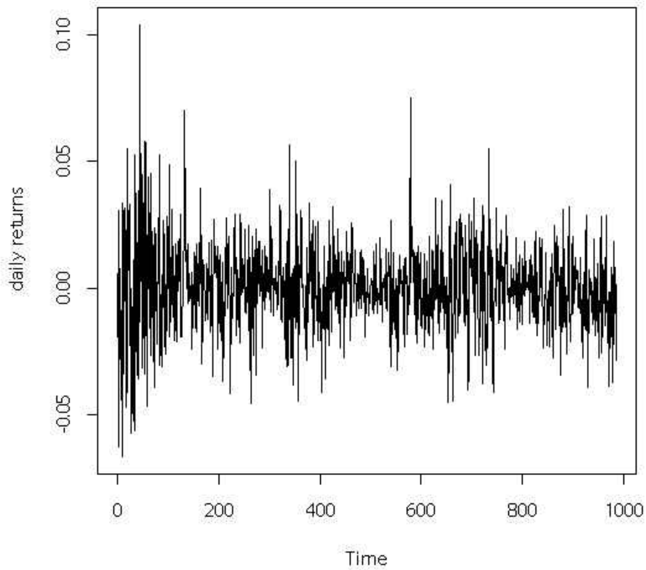

Figure 5-2 log return series

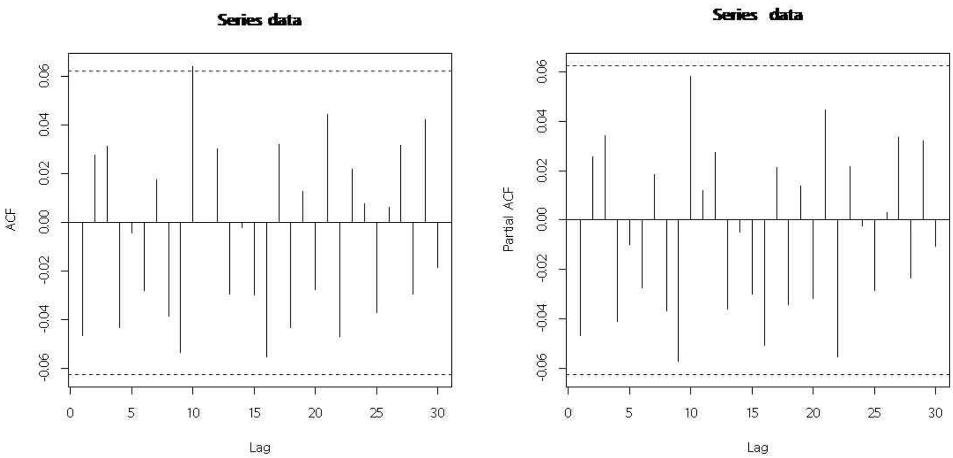

Figure 5-3 The acf and pacf of the daily log return

After fitting the threshold GARCH(1,1) model, the estimated parameters are obtained as follows:

mu=0.000133(0.000491)

omega=0.000004(0.000002)

alpha1=0.007512(0.011704)

beta1=0.938682(0.013444)

gamma1=0.082110(0.020781),

which in the parenthesis is their standard errors. Figure 5-4 shows the Q-Q plot of standardized residuals.

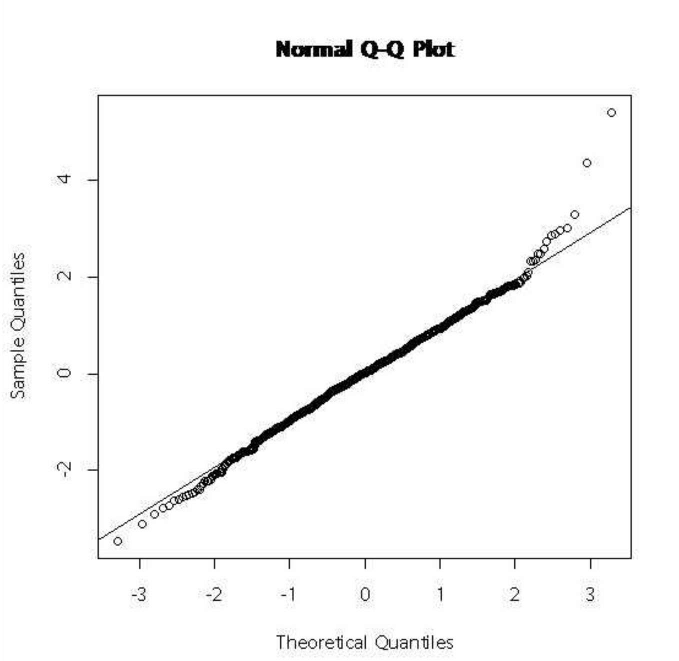

Figure 5-4 Q-Q Plots of the standardized residuals.

The model is confirmed as fellows:

$$X_t = 0.000133 + \sigma_t \epsilon_t$$

$$\sigma_t^2 = 0.000004 + (0.007512 + 0.082110 N_{t-1}) X_{t-1}^2 + 0.938682 \sigma_{t-1}^2,$$

where N_{t-1} is an indicator for negative X_{t-1}, that is

$$N_{t-1} = \begin{cases} 1 \text{ if } X_{t-1} < 0 \\ 0 \text{ if } X_{t-1} \geq 0 \end{cases}.$$

Since it is recommended that m takes the value of $n^{1/3}$, in this case, to test $H_0 : \epsilon_t \sim N(0,1)\ vs.\ H_1 : not\, H_0$, the bootstrap method with $B = 500, m = 10$ is implemented. The outcome is $\widehat{T_n}$ equals 0.6625435 with the $100(1-\alpha)\%$ values being 0.7685466, 0.7120861 and 0.6811656 for α=0.01, 0.05 and 0.10, respectively. So, surprisingly, after fitting the threshold GARCH model, it turns that the null hypothesis cannot be rejected, which stands that $\epsilon_t \sim N(0,1)$. Figure 5-5 shows the plots of the empirical density with the theoretical density of the standard normal distribution.

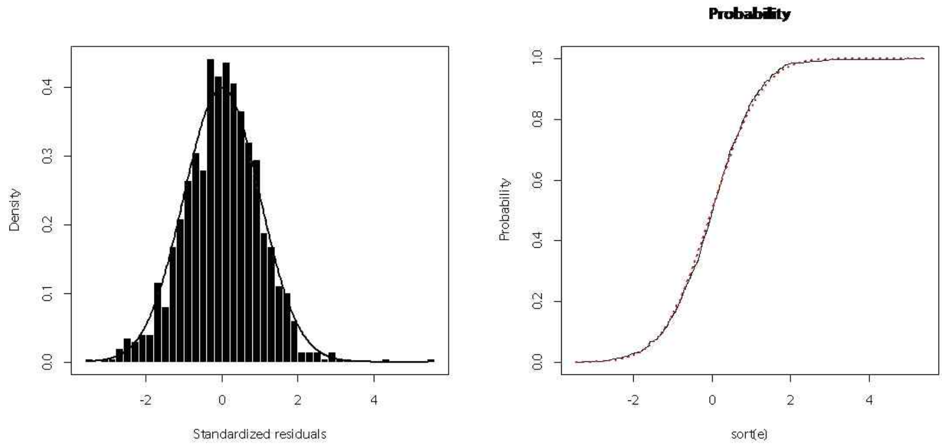

Figure 5-5

Also, we want to take Students' t-distribution and generalized error distribution (scale parameter $\lambda = 1$) as the null hypothesis into consideration. On the one hand, it turns out that $\widehat{T_n}$ equals 0.6445397 and $100(1-\alpha)\%$ values are 0.7583778, 0.6901229 and 0.6597554 for α=0.01, 0.05 and 0.10 under the Students' t-distribution assumption, respectively. On the other hand, $\widehat{T_n}$ is 0.6533133 and $100(1-\alpha)\%$ are 0.7471982, 0.6910033 and 0.6656059 for $\alpha=$0.01, 0.05 and 0.10 under the generalized error distribution assumption, respectively. Figure 5-6 and Figure 5-7 show the plots of the empirical densities with the theoretical densities of the Students' t-distribution and the generalized error distribution. Among these distributions, one can compare the AIC or BIC values to select the best fitted one.

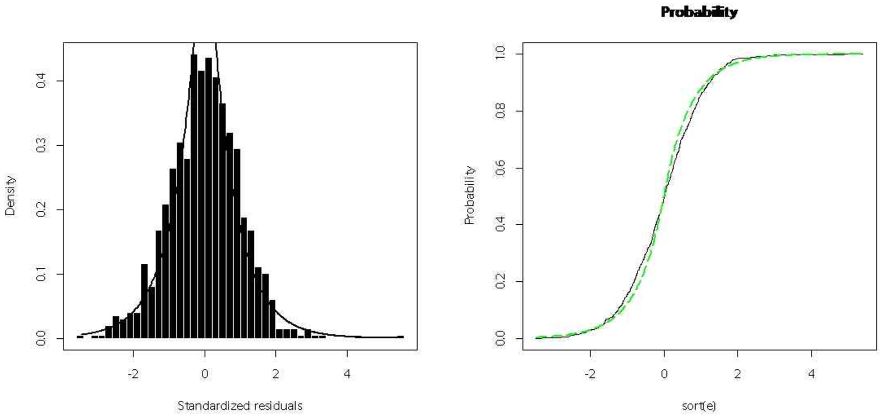

Figure 5-6 for Students' t-distribution

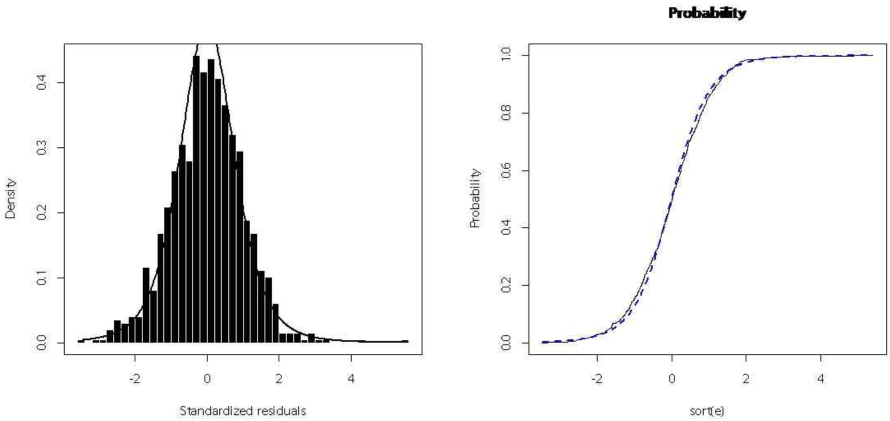

Figure 5-7 for generalized error distribution

Chapter 6

Concluding Remarks

The maximum entropy test is applied to the threshold GARCH model in this paper. This test is proved to perform adequately in the autoregressive models and GARCH models according to the study of Lee et al. While in these papers, it seems that this test performs better with a relatively larger sample size. So a bootstrap test is used to deal with the small sample size and the difficulty of choosing m. The simulation result shows that this method performs adequately and a real data analysis is conducted. In this paper, only simple hypothesis has been put into practice and the composite situation is exposited theoretically. For more information of taking the composite hypothesis case into practice, further study is suggested.

References

Bai, J. (2003). Testing parametric conditional distributions fo dynamic models. *Rev. Econ. Stat*. 85, 531-549.

Bera, A., Jarque, C. (1981). Efficient tests for normality, heteroskedasticity, and serial independence of regression residuals: Monte Carlo evidence. *Econom. Lett*. 7, 313-318.

D' Agostino, R, B. and Stephens, M. A. (1986). *Goodness-of-Fit Techniques*. Marcel Dekker, New York.

Escaciano, J. C. (2010). Asymptotic distribution-free diagnostic tests for heteroskedastic time series models. *Econometric Theory*. 26. 744-773.

E. V. Khmaladze (1993). Goodness of fit problem and scanning innovation martingales. *The Annales of Statistics*, 21, 798-829.

Forte, B and Hughes, W (1988). The maximum entropy principle: a tool to define new entropies. *Reports on mathematical Physics*, 26, 227-235.

Francq, C. and Zakoian, J. M. (2004). Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli*, 10, 605-637.

Karatzas, I., Shreve, S. E. (1991). *Brownian Motion and Stochastic Calculus.* second ed. Springer, New York.

Khmaladze, E. A. (1981). Martingale approach in the theory of goodness-of-fit tests. *Theory Probab. Appl*. **26**, 240-257.

Lee, S. (2013). A maximum entropy type test of fit: composite hypothesis case. Accepted by Computational Statistics and Data Analysis.

Lee. S., Lee, J. and Park, S. (2012). Maximum entropy test for GARCH models. Submitted for publication.

Lee, S., Lee, T., Park, S. (2010). A note on the Jarque-Bera normality test for GARCH innovations. *J. Korean Statist. Soc*. **39**, 93-102.

Lee, S. , Vonta and Karagrigoriou, A. (2011). A maximum entropy type test of fit: Composite hypothesis case. *Computational Statist. Data Anal.* **55**, 2635-2643.

Lee, S., Wei, C. Z. (1999). On residual empirical processes of stochastic regression models with applications to time series. *Ann, Statist.* **27**, 237-261.

Ruey S. Tsay (2010). *Analysis of Financial Time Series*. third ed. John Wiley & Sons, INC., Hoboken.

국문초록

이 논문에서는, 최대 엔트로피 검정을 조건부 이분산성 threshold GARCH 모형에 적용하였다. 이론적으로 귀무가설 아래에서 검정 통계량의 근사 분포를 계산해 낼 수 있기 때문에 이 검정 방법은 실제로 응용되기가 용이하다. 본 논문에서는 부트스트랩 방법으로 모의 실험을 통해 단순 가설 검정을 시도해보고 복합 가설 검정에서의 검정결과를 살펴보았다. 시뮬레이션 연구에서는 4개의 서로 다른 m, n 값과 다양한 모수에서 적절한 결과를 얻을 수 있었다. 논문의 마지막 부분에서는 2009년 1월부터 2012년 11월까지의 Intel Corporation 주가의 988개의 관측 값을 사용해 실제자료 분석을 시도하였다.

키워드: 최대 엔트로피 검정, threshold GARCH 모델, 부트스트랩 방법, 적합도 검정.

학번: 2010-24004