

A Numerical Study on CUSUM Test for Volatility Shifts Against Long-Range Dependence

Youngsun Lee^a · Taewook Lee^{a,1}

^aDepartment of statistics, Hankuk University of Foreign Studies

(Received December 31, 2013; Revised February 26, 2014; Accepted March 14, 2014)

Abstract

Persistence is one of the typical characteristics appearing in the volatility of financial time series. According to the recent researches, the volatility persistence may be due to either volatility shifts or long-range dependence. In this paper, we consider residual-based CUSUM tests to distinguish volatility persistence, long-range dependence and volatility shifts in GARCH models. It is observed that this test procedure achieve reasonable powers without a size distortion. Moreover, we employ AIC and BIC criteria to estimate the change points and the number of change points in volatility. We demonstrate the superiority of residual-based CUSUM tests on various Monte Carlo simulations and empirical data analysis.

Keywords: Volatility shifts, Long-range dependence, Persistence, CUSUM test, GARCH model.

1. 서론

금융시계열 자료에서 나타나는 변동성은 등분산성이 아니라 조건부 이분산 모형으로 설명되는 것이 일반적이다. 이러한 금융시계열 자료의 변동성을 설명하기 위해 Engle (1982)의 ARCH(AutoRegressive Conditional Heteroscedasticity) 모형과 Bollerslev (1986)의 GARCH(generalized ARCH) 모형이 주로 사용된다. GARCH 모형을 실제 자료에 적용하게 되면, 계수들의 합이 1에 가깝게 나오는 지속성(persistence) 현상이 나타나게 되는데 이를 설명하기 위해서 Engle과 Bollerslev (1986)는 GARCH 모형을 포함하는 IGARCH(integrated GARCH) 모형을 발표했다. 그러나 IGARCH 모형을 이용한 금융 자료의 모형 적합이나 장기적인 예측은 많은 연구자들에게 큰 관심을 받지 못하였다. 그 이유는 IGARCH 모형이 약정상성을 만족하지 못하여 장기 예측 시에 분산이 폭발하는 현상이 나타나는 문제가 있기 때문이다. 또한 금융시계열 자료에서 비조건부 분산의 변화가 있는 변동성 변화모형을 적합 시켜야 하는 자료에 대해서 변동성의 변화를 무시하고 GARCH 모형을 적합 시키게 되면 그 결과가 IGARCH 모형으로 나오게 된다. 이에 대한 자세한 내용은 Mikosch와 Stărică (2004)와 Hillebrand (2005)를 참고하기 바란다. 장기역성(long-range dependence)은 시계열 자료의 공분산 함수가 시차에 따라 서서히 감소하여 과거의 영향이 미래에도 상당히 미치게 되는 현상으로서 $\sum_{h=-\infty}^{\infty} |\gamma_X(h)| =$

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (NRF-2013R1A1A2006796).

¹Corresponding author: Department of statistics, Hankuk University of Foreign Studies, Yongin, Gyeonggi 449-791, Korea. E-mail: twlee@hufs.ac.kr

∞ 을 만족한다. 여기서 $\gamma_X(h)$ 는 시계열 자료 $\{X_t\}$ 에서 h -시차의 공분산을 의미한다. 특히, 장기역성일 때는 서서히 감소하는 공분산 함수가 다음의 식 (1.1)의 형태를 따르게 된다.

$$\gamma_X(h) = \text{Cov}(X_0, X_h) \sim Ch^{2d-1}, \quad \text{as } h \rightarrow \infty. \quad (1.1)$$

단, C 는 상수이며 d 는 0과 1/2 사이에 속하는 장기역성 계수이다. 장기역성과 대조되는 현상으로는 단기역성(short-range dependence)이 있는데, 이는 시계열 자료의 공분산 함수가 빠르게 감소하여 과거가 미래에 미치는 영향이 급격히 감소하는 현상으로 $\sum_{h=-\infty}^{\infty} |\gamma_X(h)| < \infty$ 을 만족한다. 한편, Ding 등 (1993)과 Baillie 등 (1996)에서 장기역성이 금융시계열 자료의 변동성을 적절히 설명하는 모형이라 나와 있지만, 변동성 변화가 있는 금융시계열 자료를 장기역성 모형으로 잘못 적합하는 경우가 종종 발생한다는 사실이 발표되고 있다. 이에 대한 자세한 내용은 Klemesš (1974), Teverovsky 등 (1999), Diebold와 Inoue (2001), Mikosch와 Stáricá (2004)가 자세히 설명하고 있다. 뿐만 아니라 장기역성이 존재하는 자료에 대해서 IGARCH 모형으로 적합되는 문제가 발생하기도 한다. 따라서 금융시계열 자료의 변동성을 분석하기 이전에 IGARCH 모형, 변동성 변화 모형 및 장기역성 모형 중에서 어떤 모형이 자료 분석에 적절한지 결정하는 문제는 많은 연구자로부터 점차 큰 주목을 받고 있다. 이를 위해서 다음의 통계적 가설 검정을 시행하고자 한다.

$$H_0 : \text{VS-R} \quad \text{vs.} \quad H_1 : \text{LRD}.$$

단, 대립가설에서의 LRD는 장기역성 모형을 나타내며, 귀무가설에서의 R 은 알려져 있지 않은 값으로 변동성 변화의 횟수이며 VS-R은 모수 변화에 의해 비조건부 분산이 변하는 변동성 변화가 R 번 일어나는 GARCH 모형을 나타낸다. 참고로 많은 경우에 금융시계열 자료의 변동성의 자기상관성은 시차가 커짐에 따라 서서히 감소하는 경향이 나타나며, 이러한 자료는 IGARCH를 포함하는 변동성 변화가 있는 GARCH 모형 또는 GARCH 모형의 변형으로서 장기역성을 나타내는 FIGARCH 모형 등으로 설명하는 것이 일반적이다. 그러나 서로 성질이 크게 다른 두 가지 모형으로 자료를 설명하는 것은 적절하지 않다고 볼 수 있다. 따라서 본 논문에서는 변동성의 자기상관성이 서서히 감소하는 금융시계열 자료에 대하여 변동성 변화가 있는 GARCH 모형과 장기역성 모형 중 어느 모형이 자료를 생성하는 모형인지를 구분하는 통계적 검정법을 제안하고자 한다. 본 논문에서는 귀무가설이 기각되면 장기역성 모형이 자료 분석을 위한 모형으로 결정된다. 그렇지 않은 경우에는 $R = 0$ 이면 IGARCH 모형을 포함하는 변화가 없는 GARCH 모형이 선택됨을 의미하고 $R \geq 1$ 인 경우에는 변동성 변화가 존재하는 GARCH 모형이 자료 적합에 적절하다고 최종 결정한다. 이에 대한 자세한 내용은 본문에서 다루도록 하겠다. 위의 가설에서 변동성 변화가 일어난 시점을 모르기 때문에 변화점에 대한 추정이 필요하다. 이에 본 논문에서는 AIC와 BIC를 이용하는 통계적 기법을 도입하여 변화점과 변화점의 개수를 추정된 후 모형 선택을 위한 가설검정에 사용할 것이다. 본 논문의 구성은 다음과 같다. 우선 2장에서는 변화가 없는 변동성 모형과 장기역성 모형을 구분하는 검정 방법을 소개하고, 3장에서는 변화가 한 번 이상 있는 변동성 모형과 장기역성 모형을 구분하는 검정을 소개한다. 4장에서는 변동성 변화의 횟수와 시점을 추정하기 위한 방법으로 AIC와 BIC를 이용한 방법을 소개한다. 5장과 6장에서는 본 논문에서 제안한 검정법이 모형 선택을 위해 적합하지 확인하기 위해 몬테카를로 모의실험을 시행하고 다양한 금융 자료를 분석한다.

2. 변화가 없는 변동성과 장기역성을 구분하는 검정

2.1. CUSUM 통계량

이 장에서는 변동성 변화가 없는 GARCH 모형과 장기역성 모형을 구분하는 검정방법을 소개한다. 일

반적으로 금융시계열 자료의 변동성 분석을 위해 다음과 같이 정의하는 로그수익률을 사용한다.

$$r_t = \log P_t - \log P_{t-1}. \quad (2.1)$$

단, P_t 는 t 시점에서의 주식의 가격 또는 주가지수를 나타낸다. 식 (2.1)에서 정의한 로그수익률 r_t 가 다음의 GARCH(p, q)모형을 만족한다고 하자.

$$r_t = \sigma_t(\theta)\epsilon_t, \quad \sigma_t^2(\theta) = \omega + \sum_{i=1}^p \alpha_i r_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2(\theta), \quad (2.2)$$

여기서 ϵ_t 는 서로 독립이고 동일한 $N(0, 1)$ 을 따른다. 또한 $\omega > 0$, 모든 i, j 에 대해서 $\alpha_i \geq 0, \beta_j \geq 0$ 을 만족해야 한다. 식 (2.2)의 GARCH 모형은 단기억성의 대표적인 모형으로 변화가 없는 변동성 모형이다. 우선 로그수익률 r_t 에 대한 관심 있는 통계적 가설은 다음과 같다.

$$H_0 : \text{VS-0} \quad \text{vs.} \quad H_1 : \text{LRD}.$$

위의 가설에서 귀무가설을 기각하지 못할 경우에는 로그수익률 r_t 가 IGARCH 모형을 포함하는 변동성 변화가 없는 GARCH 모형을 따르는 것이고, 그렇지 않은 경우에는 장기억성 모형을 따른다. 가설 검정에서 사용하게 될 통계량은 CUSUM 검정통계량으로, 변화점을 알아내기 위한 검정방법 중에서 가장 흔히 쓰는 방법이다. 로그수익률 r_t 의 변동성 변화에 대한 검정에서 사용하게 될 CUSUM 통계량은 다음과 같이 정의된다.

$$T_n = \frac{1}{\sqrt{ns_n}} \max_{1 \leq k \leq n} \left| \sum_{t=1}^k r_t^2 - \frac{k}{n} \sum_{t=1}^n r_t^2 \right|, \quad (2.3)$$

여기서 s_n^2 은 long-run 분산 $\sigma^2 = \sum_{j=-\infty}^{\infty} \gamma_j$ 의 추정량으로, γ_j 는 r_t^2 의 j -시차의 공분산을 뜻한다. CUSUM 검정통계량의 기초가 되는 내용과 극한 분포에 대해서는 Pooter와 Dijk (2004)과 Berkes 등 (2006)를 참고하기 바란다. 로그수익률의 제곱 값인 r_t^2 은 자기상관성이 존재하므로 조건부 이분산의 존재성을 만족하여 Andrews (1991)에 나와 있는 다음의 HAC(Heteroscedasticity and Autocorrelation Consistent) 추정량을 사용한다.

$$\hat{s}_n^2 = \hat{\gamma}_0 + 2 \sum_{j=1}^l w_{j,l} \hat{\gamma}_j, \quad \text{where } w_{j,l} = 1 - \frac{j}{l+1}, \quad l = 15 \times \log_{10}(n). \quad (2.4)$$

식 (2.4)에서 정의된 HAC 추정량은 Bartlett 커널 추정량으로 조율모수(tuning parameter)인 $w_{j,l}$ 값에 따라 검정력이 달라진다. 그러므로 적절한 $w_{j,l}$ 값을 이용하여 추정량을 계산해야 하는 어려움이 있다. 이와 같이 HAC 추정량으로 로그수익률에 대한 검정을 실시하면 size 왜곡현상과 검정력이 낮게 나오는 현상이 발생하게 된다. 이러한 현상들을 개선하기 위해서 다음 절에서는 GARCH 모형의 잔차를 이용한 통계량을 정의해 보고자 한다.

2.2. GARCH 모형의 잔차에 대한 CUSUM 통계량

로그수익률 자료를 이용한 CUSUM 검정통계량의 단점을 보완하기 위해 식 (2.2)의 GARCH 모형의 오차항을 이용한 CUSUM 검정통계량을 다음과 같이 정의한다.

$$\hat{T}_n = \frac{1}{\sqrt{n\tau}} \max_{1 \leq k \leq n} \left| \sum_{t=1}^k \epsilon_t^2 - \frac{k}{n} \sum_{t=1}^n \epsilon_t^2 \right|, \quad (2.5)$$

여기서 GARCH 모형에서의 오차인 $\{\epsilon_t\}$ 는 실제로 관측되지 않는 값이므로 GARCH 모형 적합 후에 얻을 수 있는 잔차를 이용한다. 우선 식 (2.2)의 GARCH 모형에서 모수 θ 의 참값을 θ_0 라고 하고, θ_0 의 추정량으로서 다음을 만족하는 Gaussian QMLE를 $\hat{\theta} = (\hat{\omega}, \hat{\alpha}_1, \dots, \hat{\alpha}_p, \hat{\beta}_1, \dots, \hat{\beta}_q)'$ 이라 하자.

$$\sqrt{n} \left\| \hat{\theta} - \theta_0 \right\| = O_P(1). \quad (2.6)$$

식 (2.6)에 대한 자세한 내용은 Berkes 등 (2003)과 Francq와 Zakoian (2004) 논문을 참고하기 바란다. 식 (2.2)를 이용하여 GARCH 모형의 잔차를 다음과 같이 정의한다.

$$\tilde{\epsilon}_t = \frac{r_t}{\tilde{\sigma}_t(\hat{\theta})}, \quad t = 1, \dots, n. \quad (2.7)$$

위 식 (2.7)에서 분모의 $\tilde{\sigma}_t^2(\hat{\theta})$ 는 다음의 식 (2.8)을 이용하여 순차적으로 계산되는 값이다.

$$\tilde{\sigma}_t^2(\hat{\theta}) = \hat{\omega} + \sum_{i=1}^p \hat{\alpha}_i r_{t-i}^2 + \sum_{j=1}^q \hat{\beta}_j \sigma_{t-j}^2(\theta). \quad (2.8)$$

식 (2.8)를 계산하기 위해서 초기값인 r_0^2, \dots, r_{1-p}^2 과 $\sigma_0^2, \dots, \sigma_{1-q}^2$ 가 필요하다. Francq와 Zakoian (2004)와 마찬가지로 본 논문에서는 $\sigma_0^2, \dots, \sigma_{1-q}^2$ 의 초기값을 GARCH 모형의 비조건부 분산의 추정값으로 하겠다. 식 (2.7)에서 정의된 잔차를 이용한 CUSUM 검정통계량에서 사용되는 long-run 분산 τ^2 추정량은 다음의 두 가지 방법으로 계산할 수 있다. 첫 번째는 HAC 추정량인 Bartlett 커널 추정량을 이용하여 τ^2 의 추정량을 계산하는 것으로 식 (2.9)와 같다.

$$\hat{\tau}_B^2 = \hat{\gamma}_0 + 2 \sum_{j=1}^l w_{j,l} \hat{\gamma}_j, \quad \text{where } w_{j,l} = 1 - \frac{j}{l+1}, \quad l = 15 \times \log_{10}(n). \quad (2.9)$$

두 번째로는 $\{\epsilon_t\}$ 의 기본가정인 서로 독립이라는 점을 이용하여 다음과 같이 계산할 수 있다.

$$\hat{\tau}_V^2 = \frac{1}{n} \sum_{t=1}^n \left(\tilde{\epsilon}_t^2 - \frac{1}{n} \sum_{t=1}^n \tilde{\epsilon}_t^2 \right)^2 = \frac{1}{n} \sum_{t=1}^n \tilde{\epsilon}_t^4 - \left(\frac{1}{n} \sum_{t=1}^n \tilde{\epsilon}_t^2 \right)^2. \quad (2.10)$$

식 (2.10)의 $\hat{\tau}_V^2$ 는 $\tau^2 = \text{Var}(\tilde{\epsilon}_t^2)$ 의 적률추정량이다. GARCH 모형의 잔차 $\{\tilde{\epsilon}_t\}$ 를 이용한 CUSUM 검정통계량은 τ 의 추정량 식 (2.9)와 (2.10)에 따라서 다음의 식 (2.11)과 (2.12)로 정의된다.

$$\tilde{T}_{n,B} = \frac{1}{\sqrt{n\hat{\tau}_B}} \max_{1 \leq k \leq n} \left| \sum_{t=1}^k \tilde{\epsilon}_t^2 - \frac{k}{n} \sum_{t=1}^n \tilde{\epsilon}_t^2 \right|, \quad (2.11)$$

$$\tilde{T}_{n,V} = \frac{1}{\sqrt{n\hat{\tau}_V}} \max_{1 \leq k \leq n} \left| \sum_{t=1}^k \tilde{\epsilon}_t^2 - \frac{k}{n} \sum_{t=1}^n \tilde{\epsilon}_t^2 \right|. \quad (2.12)$$

Lee 등 (2014), Kulperger와 Yu (2005)와 Berkes 등 (2006)에 근거하여 위의 두 검정통계량 $\tilde{T}_{n,B}$ 와 $\tilde{T}_{n,V}$ 의 극한 분포는 $\sup_{0 \leq u \leq 1} |B^o(u)|$ 임을 알 수 있으며, 이에 대한 성능은 5장에서 모의실험을 통하여 비교하고자 한다. 단, B^o 는 표준화된 Brownian bridge를 의미한다.

3. 변동성 변화와 장기역성을 구분하는 검정

3.1. CUSUM 통계량

2장에서 변동성 변화가 없는 경우에 대한 가설검정을 고려해보았다면, 3장에서는 변동성 변화가 있는 모형과 장기역성 모형을 구분하는 검정에 대해 소개하겠다. 변동성의 변화 횟수는 R 번으로 알려져 있

고, 변화가 일어난 변화점은 알려져 있지 않은 경우 관심 있는 가설은 아래와 같다.

$$H_0 : \text{VS-R} \quad \text{vs.} \quad H_1 : \text{LRD},$$

여기서 R 은 1 이상의 값을 가지며, 귀무가설을 기각하지 못할 경우에는 변동성 변화 모형을 따르고 기각할 경우에는 장기역성 모형을 따르게 된다. 본 논문에서는 변동성 변화가 한번 있는 $R = 1$ 인 경우만을 고려하고자 하며, R 이 2 이상인 경우는 한 번 변화가 있는 방법을 확장하면 되므로 자세한 기술은 생략하고자 한다. 먼저, 귀무가설 하에서의 모형은 알려져 있지 않은 k^* 에서 한 번의 변화가 일어난 경우이다. 동일한 $\{\epsilon_t\}$ 을 갖는 두 수익률자료 $\{r_{1,t}\}_{t=1}^n$ 과 $\{r_{2,t}\}_{t=1}^n$ 는 다음에서 정의되는 GARCH(p, q) 모형을 따르게 된다.

$$\begin{aligned} r_{1,t} &= \sigma_{1,t}(\theta_1)\epsilon_t, & \sigma_{1,t}^2(\theta_1) &= \omega_1 + \sum_{i=1}^p \alpha_{1,i}r_{1,t-i}^2 + \sum_{j=1}^q \beta_{1,j}\sigma_{1,t-j}^2(\theta_1), \\ r_{2,t} &= \sigma_{2,t}(\theta_2)\epsilon_t, & \sigma_{2,t}^2(\theta_2) &= \omega_2 + \sum_{i=1}^p \alpha_{2,i}r_{2,t-i}^2 + \sum_{j=1}^q \beta_{2,j}\sigma_{2,t-j}^2(\theta_2). \end{aligned} \tag{3.1}$$

식 (3.1)에서 ϵ_t 는 서로 독립이며 동일한 $N(0, 1)$ 을 따른다. 또한 $l = 1, 2$ 에 대해서 $\omega_l > 0$, 모든 i, j 에 대해서 $\alpha_{l,i} \geq 0, \beta_{l,j} \geq 0$ 을 만족해야 한다. 정상성을 만족하는 조건 하에서 얻은 $\{r_{1,t}\}$ 과 $\{r_{2,t}\}$ 는 변화가 한 번 있는 로그수익률 자료로 다음과 같이 정의된다.

$$r_t = \begin{cases} r_{1,t}, & \text{if } 1 \leq t \leq k^*, \\ r_{2,t}, & \text{if } k^* < t \leq n, \end{cases} \tag{3.2}$$

여기서 k^* 은 알려지지 않은 변동성 변화점이다. 이제 변동성 변화점 k^* 의 추정량을 \hat{k} 이라 할 때, 변화점을 기준으로 두 개의 CUSUM 통계량이 다음의 식 (3.3)과 같이 정의된다. 참고로 추정량 \hat{k} 에 대한 자세한 내용은 4장에서 다루도록 하겠다.

$$\begin{aligned} T_{n,1} &= \frac{1}{s_{n,1}} \hat{k}^{-\frac{1}{2}} \max_{1 \leq k \leq \hat{k}} \left| \sum_{t=1}^k r_t^2 - \frac{k}{\hat{k}} \sum_{t=1}^{\hat{k}} r_t^2 \right|, \\ T_{n,2} &= \frac{1}{s_{n,2}} (n - \hat{k})^{-\frac{1}{2}} \max_{\hat{k} < k \leq n} \left| \sum_{t=k}^n r_t^2 - \frac{k - \hat{k}}{n - \hat{k}} \sum_{t=\hat{k}+1}^n r_t^2 \right|, \end{aligned} \tag{3.3}$$

여기서 $T_{n,1}$ 은 변화점 추정량 \hat{k} 이전의 자료에 대한 CUSUM 통계량이고, $T_{n,2}$ 는 변화점 추정량 \hat{k} 이후의 자료에 대한 CUSUM 통계량으로 다음의 식 (3.4)를 만족하는 M_n 이 최종 CUSUM 통계량 값이 된다.

$$M_n = \max\{T_{n,1}, T_{n,2}\}. \tag{3.4}$$

CUSUM 검정통계량 M_n 의 기초가 되는 내용과 극한 분포에 대해서는 Berkes 등 (2006)를 참고하기 바란다. 식 (3.3)에서 $s_{n,1}^2$ 과 $s_{n,2}^2$ 은 long-run 분산으로 변동성 변화가 없는 모형에 대한 검정과 동일하게 로그수익률 자료를 조건부 이분산 의존성으로 가정하여 HAC 추정량인 Bartlett 커널 추정량을 이용해 다음의 식 (3.5)로 구한다.

$$\begin{aligned} \hat{s}_{n,1}^2 &= \hat{\gamma}_{0,1} + 2 \sum_{j=1}^{l_1} w_{j,l_1} \hat{\gamma}_{j,1}, & \text{where } w_{j,l_1} &= 1 - \frac{j}{l_1 + 1}, \quad l_1 = 15 \times \log_{10}(\hat{k}), \\ \hat{s}_{n,2}^2 &= \hat{\gamma}_{0,2} + 2 \sum_{j=1}^{l_2} w_{j,l_2} \hat{\gamma}_{j,2}, & \text{where } w_{j,l_2} &= 1 - \frac{j}{l_2 + 1}, \quad l_2 = 15 \times \log_{10}(n - \hat{k}), \end{aligned} \tag{3.5}$$

여기서 $\gamma_{j,1}$ 은 변화점 추정량 \hat{k} 이전 자료의 j -시차의 공분산을 뜻하고, $\gamma_{j,2}$ 은 변화점 추정량 \hat{k} 이후 자료의 j -시차의 공분산을 뜻한다. 식 (3.5)의 추정량도 로그수익률 자료에 대한 추정량으로 size 왜곡현상과 검정력이 낮게 나오는 현상이 발생하게 되므로 잔차를 이용하여 검정을 실시해 보고자 한다.

3.2. GARCH 모형의 잔차에 대한 CUSUM 통계량

로그수익률 자료에 대한 CUSUM 통계량의 단점을 극복하기 위해서 식 (3.1)의 GARCH 모형 오차인 $\{\epsilon_t\}$ 를 이용하여 CUSUM 통계량을 다음과 같이 정의한다.

$$\begin{aligned}\hat{T}_{n,1} &= \frac{1}{\tau_1} \hat{k}^{-\frac{1}{2}} \max_{1 \leq k \leq \hat{k}} \left| \sum_{t=1}^k \epsilon_t^2 - \frac{k}{\hat{k}} \sum_{t=1}^{\hat{k}} \epsilon_t^2 \right|, \\ \hat{T}_{n,2} &= \frac{1}{\tau_2} (n - \hat{k})^{-\frac{1}{2}} \max_{\hat{k} < k \leq n} \left| \sum_{t=\hat{k}+1}^k \epsilon_t^2 - \frac{k - \hat{k}}{n - \hat{k}} \sum_{t=\hat{k}+1}^n \epsilon_t^2 \right|,\end{aligned}\quad (3.6)$$

여기서 $\{\epsilon_t\}$ 는 GARCH 모형의 오차로, 실제로 관측되지 않는 값이므로 GARCH 모형 적합 후에 얻을 수 있는 잔차를 이용하여 CUSUM 통계량을 정의해야 한다. 우선 식 (3.1)의 GARCH 모형에서 변화점 이전의 모수의 참값을 $\theta_{0,1}$ 이라 하고 변화점 이후의 모수의 참값을 $\theta_{0,2}$ 라 할 때, $\theta_{0,1}$ 과 $\theta_{0,2}$ 의 추정량으로서 Gaussian QMLEs를 변화점 이전의 $\hat{\theta}_1 = (\hat{\omega}_1, \hat{\alpha}_{1,1}, \dots, \hat{\alpha}_{1,p}, \hat{\beta}_{1,1}, \dots, \hat{\beta}_{1,q})'$ 과 변화점 이후의 $\hat{\theta}_2 = (\hat{\omega}_2, \hat{\alpha}_{2,1}, \dots, \hat{\alpha}_{2,p}, \hat{\beta}_{2,1}, \dots, \hat{\beta}_{2,q})'$ 라 하자. 이에 대한 자세한 내용은 Lee 등 (2014)를 참고하기 바란다. 식 (3.1)을 이용한 GARCH 모형의 잔차는 다음과 같이 정의한다.

$$\tilde{\epsilon}_t = \begin{cases} \frac{r_t}{\tilde{\sigma}_{1,t}(\hat{\theta}_1)}, & t = 1, \dots, \hat{k}, \\ \frac{r_t}{\tilde{\sigma}_{2,t}(\hat{\theta}_2)}, & t = \hat{k} + 1, \dots, n.\end{cases}\quad (3.7)$$

식 (3.7)에서 분모의 $\tilde{\sigma}_{1,t}^2(\hat{\theta}_1)$ 와 $\tilde{\sigma}_{2,t}^2(\hat{\theta}_2)$ 는 다음의 식 (3.8)에서 순차적으로 계산되는 값이다.

$$\begin{aligned}\tilde{\sigma}_{1,t}^2(\hat{\theta}_1) &= \hat{\omega}_1 + \sum_{i=1}^p \hat{\alpha}_{1,i} \tilde{r}_{1,t-i}^2 + \sum_{j=1}^q \hat{\beta}_{1,j} \tilde{\sigma}_{1,t-j}^2(\hat{\theta}_1), \\ \tilde{\sigma}_{2,t}^2(\hat{\theta}_2) &= \hat{\omega}_2 + \sum_{i=1}^p \hat{\alpha}_{2,i} \tilde{r}_{2,t-i}^2 + \sum_{j=1}^q \hat{\beta}_{2,j} \tilde{\sigma}_{2,t-j}^2(\hat{\theta}_2).\end{aligned}\quad (3.8)$$

식 (3.8)을 계산하기 위한 $\tilde{\sigma}_{1,0}^2, \dots, \tilde{\sigma}_{1,1-q}^2$ 과 $\tilde{\sigma}_{2,0}^2, \dots, \tilde{\sigma}_{2,1-q}^2$ 의 초기값은 변화점을 기준으로 이전과 이후에 대한 GARCH 모형의 비조건부 분산의 추정값으로 하겠다. 식 (3.7)에서 정의된 잔차를 이용한 CUSUM 검정통계량에서 사용되는 long-run 분산 τ_1^2 과 τ_2^2 의 추정량을 구하는 방법은 2장과 동일하게 두 가지 방법이 있다. 첫 번째는 조건부 이분산 의존성을 가정한 HAC 추정량인 Bartlett 커널 추정량을 이용하여 다음의 식으로 추정할 수 있다.

$$\begin{aligned}\hat{\tau}_{B,1}^2 &= \hat{\gamma}_{0,1} + 2 \sum_{j=1}^{l_1} w_{j,l_1} \hat{\gamma}_{j,1}, \quad \text{where } w_{j,l_1} = 1 - \frac{j}{l_1 + 1}, \quad l_1 = 15 \times \log_{10}(\hat{k}), \\ \hat{\tau}_{B,2}^2 &= \hat{\gamma}_{0,2} + 2 \sum_{j=1}^{l_2} w_{j,l_2} \hat{\gamma}_{j,2}, \quad \text{where } w_{j,l_2} = 1 - \frac{j}{l_2 + 1}, \quad l_2 = 15 \times \log_{10}(n - \hat{k}),\end{aligned}\quad (3.9)$$

여기서는 로그수익률 자료 대신 GARCH 모형에서의 잔차인 $\tilde{\epsilon}_t^2$ 을 이용하며, $\hat{\tau}_{B,1}$ 는 변화점 추정량 \hat{k} 이전에 대한 추정량이고 $\hat{\tau}_{B,2}$ 는 변화점 추정량 \hat{k} 이후의 추정량이다. 두 번째로는 $\{\epsilon_t\}$ 의 기본가정인 서로 독립이라는 점을 이용하여 식 (3.10)으로 계산할 수도 있다.

$$\hat{\tau}_{V,1}^2 = \frac{1}{\hat{k}} \sum_{t=1}^{\hat{k}} \tilde{\epsilon}_t^4 - \left(\frac{1}{\hat{k}} \sum_{t=1}^{\hat{k}} \tilde{\epsilon}_t^2 \right)^2, \quad \hat{\tau}_{V,2}^2 = \frac{1}{n-\hat{k}} \sum_{t=\hat{k}+1}^n \tilde{\epsilon}_t^4 - \left(\frac{1}{n-\hat{k}} \sum_{t=\hat{k}+1}^n \tilde{\epsilon}_t^2 \right)^2, \quad (3.10)$$

여기서 $\hat{\tau}_{V,1}$ 은 변화점 추정량 \hat{k} 이전에 대한 추정량이고, $\hat{\tau}_{V,2}$ 는 변화점 추정량 \hat{k} 이후에 대한 추정량이다. GARCH 모형의 잔차 $\{\tilde{\epsilon}_t\}$ 에 대한 CUSUM 검정통계량은 τ_1 과 τ_2 의 추정량에 따라서 식 (3.11)과 (3.13)로 정의된다.

$$M_{n,B} = \max \left\{ \tilde{T}_{n,B,1}, \tilde{T}_{n,B,2} \right\}. \quad (3.11)$$

위의 식 (3.11)은 Bartlett 커널 추정량을 이용한 방법으로 잔차를 조건부 이분산 의존성으로 가정했을 경우의 CUSUM 검정통계량이다. 여기서 쓰이는 $\tilde{T}_{n,B,1}$ 과 $\tilde{T}_{n,B,2}$ 는 다음의 식 (3.12)에서 계산된 통계량을 이용한다.

$$\begin{aligned} \tilde{T}_{n,B,1} &= \frac{1}{\hat{\tau}_{B,1}} \hat{k}^{-\frac{1}{2}} \max_{1 \leq k \leq \hat{k}} \left| \sum_{t=1}^k \tilde{\epsilon}_t^2 - \frac{k}{\hat{k}} \sum_{t=1}^{\hat{k}} \tilde{\epsilon}_t^2 \right|, \\ \tilde{T}_{n,B,2} &= \frac{1}{\hat{\tau}_{B,2}} (n-\hat{k})^{-\frac{1}{2}} \max_{\hat{k} < k \leq n} \left| \sum_{t=\hat{k}+1}^k \tilde{\epsilon}_t^2 - \frac{k-\hat{k}}{n-\hat{k}} \sum_{t=\hat{k}+1}^n \tilde{\epsilon}_t^2 \right|. \end{aligned} \quad (3.12)$$

식 (3.13)은 적률추정량을 이용하여 long-run 분산을 추정했을 경우의 CUSUM 검정통계량이다.

$$M_{n,V} = \max \left\{ \tilde{T}_{n,V,1}, \tilde{T}_{n,V,2} \right\}. \quad (3.13)$$

위의 식에서 쓰이는 $\tilde{T}_{n,V,1}$ 과 $\tilde{T}_{n,V,2}$ 는 다음의 식 (3.14)에서 계산된 통계량을 이용한다.

$$\begin{aligned} \tilde{T}_{n,V,1} &= \frac{1}{\hat{\tau}_{V,1}} \hat{k}^{-\frac{1}{2}} \max_{1 \leq k \leq \hat{k}} \left| \sum_{t=1}^k \tilde{\epsilon}_t^2 - \frac{k}{\hat{k}} \sum_{t=1}^{\hat{k}} \tilde{\epsilon}_t^2 \right|, \\ \tilde{T}_{n,V,2} &= \frac{1}{\hat{\tau}_{V,2}} (n-\hat{k})^{-\frac{1}{2}} \max_{\hat{k} < k \leq n} \left| \sum_{t=\hat{k}+1}^k \tilde{\epsilon}_t^2 - \frac{k-\hat{k}}{n-\hat{k}} \sum_{t=\hat{k}+1}^n \tilde{\epsilon}_t^2 \right|. \end{aligned} \quad (3.14)$$

식 (3.11)과 (3.13)의 검정통계량 $M_{n,B}$ 와 $M_{n,V}$ 의 극한 분포는 엄밀한 증명이 추가적으로 필요하지만 Lee 등 (2014)과 Berkes 등 (2006)에 근거하여 $\max\{\sup_{0 \leq u \leq 1} |B_1^o(u)|, \sup_{0 \leq u \leq 1} |B_2^o(u)|\}$ 으로 유추할 수 있으며 이에 대한 성능은 5장에서 모의실험을 통하여 비교하고자 한다. 단, B_1^o 과 B_2^o 는 표준화된 Brownian bridge이며 서로 독립이다.

4. 변화점과 변화점 개수 추정

3장에서 변동성의 변화 횟수 R 이 1인 경우, 알려지지 않은 변화점 k^* 의 추정량인 \hat{k} 을 이용하여 변동성 변화와 장기역성을 구분하는 CUSUM 검정통계량을 정의하였다. 이러한 CUSUM 검정통계량의 검정력을 높이기 위해서는 변화점과 변화점 개수를 정확히 추정하는 것이 매우 중요한 과제인데, 본 논문에서

서는 AIC와 BIC를 이용하는 모수적인 방법을 이용하여 변화점을 추정하려고 한다. 먼저 변화점을 찾기 이전에 변화의 개수의 상한으로 R^* 를 설정한다. 변동성 변화가 R 번 있는 경우, $0 \leq R \leq R^*$ 에 속하는 R 에 대해서 식 (4.1)과 (4.2)를 이용하여 GARCH(p, q) 모형에 대한 AIC와 BIC를 다음과 같이 계산한다.

$$\begin{aligned} \text{AIC}(R, k_1^*, \dots, k_R^*) &= -2 \sum_{i=1}^{R+1} l_{k_{i-1}^*}^{k_i^*}(\hat{\theta}_i)/n + \frac{2}{n}R + (R+1)(1+p+q), \\ \text{BIC}(R, k_1^*, \dots, k_R^*) &= -2 \sum_{i=1}^{R+1} l_{k_{i-1}^*}^{k_i^*}(\hat{\theta}_i)/n + \frac{\log(n)}{n}R + (R+1)(1+p+q), \end{aligned} \quad (4.1)$$

여기서 $1 = k_0^* \leq k_1^* < k_2^* < \dots < k_R^* < k_{R+1}^* = n$ 이고, $l_{k_{i-1}^*}^{k_i^*}(\hat{\theta}_i)$ 는 로그수익률의 제곱값 r_t^2 이 $k_{i-1}^* < k_i^* \leq k_{i+1}^*$ 에 속하는 $k_i^* - k_{i-1}^*$ 개의 자료에 대한 GARCH(p, q) 모형 적합 후의 log-likelihood 값이다. 최소의 AIC를 갖는 추정된 변화점의 개수는 \hat{R}_A 이고 이 때의 변화점 추정량은 $(\hat{k}_{1,A}, \dots, \hat{k}_{\hat{R}_A,A})$ 이다. 마찬가지로 최소의 BIC를 갖는 변화점 개수 및 변화점 추정량은 각각 \hat{R}_B 와 $(\hat{k}_{1,B}, \dots, \hat{k}_{\hat{R}_B,B})$ 이다.

$$\begin{aligned} (\hat{R}_A, \hat{k}_{1,A}, \dots, \hat{k}_{\hat{R}_A,A}) &= \operatorname{argmin} \text{AIC}(R, k_1^*, \dots, k_R^*), \\ (\hat{R}_B, \hat{k}_{1,B}, \dots, \hat{k}_{\hat{R}_B,B}) &= \operatorname{argmin} \text{BIC}(R, k_1^*, \dots, k_R^*). \end{aligned} \quad (4.2)$$

식 (4.3)에 의해서 추정된 변화점의 개수와 변화점을 이용하여 2, 3장에서 정의된 CUSUM 검정통계량 중 적절한 방법으로 검정을 실시한다.

5. 모의실험

이 장에서는 몬테카를로 모의실험을 통해서 앞서 소개된 검정방법에 대한 평가를 실시한다. 서론에서 정의된 통계적 가설 $H_0 : \text{VS-R vs. } H_1 : \text{LRD}$ 에서 귀무가설에 해당하는 변동성 변화모형은 대표적인 단기역성 모형인 GARCH(1,1) 모형을 이용하고, 대립가설에 해당하는 장기역성 모형은 다음의 식 (5.1)로 정의되는 FIGARCH(1, d , 1) 모형을 이용하였다.

$$r_t = \sigma_t \epsilon_t, \quad (1 - \beta L)\sigma_t^2 = \omega + ((1 - \beta L) - (1 - \phi L)(1 - L)^d)r_t^2 \quad (5.1)$$

식 (5.1)에서 L 은 후진연산자이고, ϵ_t 는 서로 독립이고 동일한 $N(0, 1)$ 을 따른다. 모형의 정상성 조건에 의해서 $0 < d < 1$ 일 때, $\omega > 0$, $0 \leq d \leq 1 - 2\phi$ 그리고 $0 \leq \beta < \phi + d$ 을 만족해야 한다. 모든 모의실험에서는 자료의 개수 n 이 1000, 2000 그리고 5000개인 경우에 대하여 1000번을 반복하였으며 5% 유의수준을 기초로 한다. Lee 등 (2014), Kulperger와 Yu (2005)와 Berkes 등 (2006)에 의하여 유의수준 5%에서 검정통계량 T_n , $\tilde{T}_{n,B}$ 및 $\tilde{T}_{n,V}$ 의 기각치는 1.36, M_n , $M_{n,B}$ 및 $M_{n,V}$ 의 기각치는 1.48이다. Table 5.1은 변동성 변화가 없는 GARCH(1,1) 모형에 대한 CUSUM 검정통계량 T_n , $\tilde{T}_{n,B}$ 및 $\tilde{T}_{n,V}$ 의 size 결과이다. 오차항의 분포는 IID $N(0, 1)$ 과 자유도가 5인 T-분포를 고려하였다. 전반적으로 n 이 증가할수록 검정통계량 \tilde{T}_n 의 size가 유의수준 0.05에 가까운 값이 나오는 것을 확인할 수 있으며, 이는 $N(0, 1)$ 뿐만 아니라 두터운 꼬리를 설명하는 $t(5)$ 의 오차항에서도 동일하게 나타난다. Table 5.2는 FIGARCH(1, d , 1) 모형을 이용하여 변화가 없는 변동성 모형에 대한 CUSUM 검정통계량 T_n , $\tilde{T}_{n,B}$ 및 $\tilde{T}_{n,V}$ 의 검정력을 평가한 결과이다. 로그수익률 자료를 이용한 결과보다 GARCH(1,1) 모형을 적합한 후의 잔차에 대한 검정력이 더 좋은 것을 확인할 수 있으며, 그 중에서도 오차항 ϵ_t 의 특성인 서로 독립임을 이용한 추정량 \hat{r}_V 에 대한 결과가 더 좋은 것을 확인할 수 있다.

Table 5.1. Empirical size of CUSUM test for no volatility shifts in GARCH(1, 1) models

| | (ω, α, β) | $N(0, 1)$ | | | $t(5)$ | | |
|-------------------|---------------------------|-------------|-------------|-------------|-------------|-------------|-------------|
| | | $n = 1,000$ | $n = 2,000$ | $n = 5,000$ | $n = 1,000$ | $n = 2,000$ | $n = 5,000$ |
| T_n | (.1, .1, .8) | 0.041 | 0.052 | 0.071 | 0.020 | 0.032 | 0.027 |
| | (.1, .1, .6) | 0.020 | 0.040 | 0.044 | 0.021 | 0.019 | 0.023 |
| | (.1, .2, .6) | 0.037 | 0.037 | 0.045 | 0.023 | 0.026 | 0.037 |
| | (.2, .1, .8) | 0.034 | 0.050 | 0.057 | 0.016 | 0.028 | 0.039 |
| $\tilde{T}_{n,B}$ | (.1, .1, .8) | 0.028 | 0.042 | 0.048 | 0.019 | 0.021 | 0.036 |
| | (.1, .1, .6) | 0.027 | 0.038 | 0.042 | 0.021 | 0.026 | 0.034 |
| | (.1, .2, .6) | 0.028 | 0.036 | 0.037 | 0.014 | 0.026 | 0.027 |
| | (.2, .1, .8) | 0.043 | 0.037 | 0.041 | 0.017 | 0.029 | 0.023 |
| $\tilde{T}_{n,V}$ | (.1, .1, .8) | 0.030 | 0.043 | 0.049 | 0.016 | 0.030 | 0.031 |
| | (.1, .1, .6) | 0.033 | 0.049 | 0.039 | 0.024 | 0.020 | 0.033 |
| | (.1, .2, .6) | 0.045 | 0.039 | 0.041 | 0.033 | 0.034 | 0.024 |
| | (.2, .1, .8) | 0.033 | 0.042 | 0.039 | 0.016 | 0.036 | 0.037 |

Table 5.2. Empirical power of CUSUM test for H_0 : VS-0 vs. H_1 : LRD (FIGARCH(1, d , 1))

| | (ω, β, ϕ, d) | $n = 1,000$ | $n = 2,000$ | $n = 5,000$ |
|-------------------|----------------------------|-------------|-------------|-------------|
| T_n | (.3, .1, .2, .2) | 0.121 | 0.224 | 0.379 |
| | (.3, .1, .2, .4) | 0.183 | 0.351 | 0.526 |
| | (.6, .2, .2, .2) | 0.119 | 0.251 | 0.417 |
| | (.6, .2, .2, .4) | 0.216 | 0.401 | 0.568 |
| $\tilde{T}_{n,B}$ | (.3, .1, .2, .2) | 0.169 | 0.281 | 0.506 |
| | (.3, .1, .2, .4) | 0.287 | 0.521 | 0.759 |
| | (.6, .2, .2, .2) | 0.151 | 0.253 | 0.442 |
| | (.6, .2, .2, .4) | 0.247 | 0.520 | 0.739 |
| $\tilde{T}_{n,V}$ | (.3, .1, .2, .2) | 0.313 | 0.474 | 0.673 |
| | (.3, .1, .2, .4) | 0.517 | 0.707 | 0.911 |
| | (.6, .2, .2, .2) | 0.225 | 0.379 | 0.580 |
| | (.6, .2, .2, .4) | 0.406 | 0.618 | 0.860 |

Table 5.3. Estimated number of volatility shifts for GARCH(1, 1) models with no volatility shifts

| (ω, α, β) | 변화점의 개수 | $n = 1,000$ | | $n = 2,000$ | |
|---------------------------|------------|-------------|------|-------------|------|
| | | AIC | BIC | AIC | BIC |
| (.1, .1, .8) | 0 | 529 | 1000 | 414 | 1000 |
| | 1 | 471 | 0 | 586 | 0 |
| (.1, .1, .6) | 0 | 663 | 1000 | 547 | 1000 |
| | 1 | 337 | 0 | 453 | 0 |
| (.1, .1, .4) | 0 | 683 | 1000 | 605 | 1000 |
| | 1 | 317 | 0 | 395 | 0 |
| (.1, .2, .6) | 0 | 553 | 1000 | 504 | 1000 |
| | 1 | 467 | 0 | 596 | 0 |
| (.2, .1, .8) | 0 | 518 | 1000 | 393 | 999 |
| | 1 | 482 | 0 | 607 | 1 |

다음의 Table 5.3과 Table 5.4는 AIC와 BIC 방법으로 변동성 변화점 개수를 얼마나 잘 추정하는지 확인해 본 결과를 나타낸다. 먼저 Table 5.3은 변동성 변화가 없는 모형을 생성하여 변동성 변화가 없거나

Table 5.4. Estimated number of volatility shifts for GARCH(1, 1) models with a volatility shift

| $(\omega_1, \alpha_1, \beta_1) \rightarrow (\omega_2, \alpha_2, \beta_2)$ | 변화점의 개수 | $n = 1,000$ | | $n = 2,000$ | |
|---|------------|-------------|-----|-------------|-----|
| | | AIC | BIC | AIC | BIC |
| $(.1, .1, .8) \rightarrow (.1, .1, .6)$ | 0 | 0 | 354 | 0 | 12 |
| | 1 | 640 | 646 | 550 | 988 |
| | 2 | 360 | 0 | 450 | 0 |
| $(.1, .1, .8) \rightarrow (.1, .1, .4)$ | 0 | 0 | 67 | 82 | 200 |
| | 1 | 647 | 933 | 527 | 800 |
| | 2 | 353 | 0 | 391 | 0 |
| $(.1, .1, .8) \rightarrow (.1, .2, .6)$ | 0 | 0 | 602 | 0 | 43 |
| | 1 | 568 | 398 | 495 | 957 |
| | 2 | 432 | 0 | 505 | 0 |
| $(.1, .1, .8) \rightarrow (.2, .1, .8)$ | 0 | 3 | 917 | 0 | 471 |
| | 1 | 570 | 83 | 484 | 529 |
| | 2 | 427 | 0 | 516 | 0 |

Table 5.5. Empirical size of CUSUM test for a volatility shift in GARCH(1, 1) models

| | $(\omega_1, \alpha_1, \beta_1) \rightarrow (\omega_2, \alpha_2, \beta_2)$ | $n = 1,000$ | $n = 2,000$ | $n = 5,000$ |
|-----------|---|-------------|-------------|-------------|
| M_n | $(.1, .1, .8) \rightarrow (.1, .1, .6)$ | 0.007 | 0.020 | 0.037 |
| | $(.1, .1, .8) \rightarrow (.1, .1, .4)$ | 0.010 | 0.030 | 0.043 |
| | $(.1, .1, .8) \rightarrow (.1, .2, .6)$ | 0.010 | 0.022 | 0.041 |
| | $(.1, .1, .8) \rightarrow (.2, .1, .8)$ | 0.026 | 0.035 | 0.045 |
| $M_{n,B}$ | $(.1, .1, .8) \rightarrow (.1, .1, .6)$ | 0.007 | 0.019 | 0.029 |
| | $(.1, .1, .8) \rightarrow (.1, .1, .4)$ | 0.011 | 0.028 | 0.041 |
| | $(.1, .1, .8) \rightarrow (.1, .2, .6)$ | 0.060 | 0.015 | 0.035 |
| | $(.1, .1, .8) \rightarrow (.2, .1, .8)$ | 0.002 | 0.019 | 0.028 |
| $M_{n,V}$ | $(.1, .1, .8) \rightarrow (.1, .1, .6)$ | 0.016 | 0.025 | 0.035 |
| | $(.1, .1, .8) \rightarrow (.1, .1, .4)$ | 0.025 | 0.025 | 0.026 |
| | $(.1, .1, .8) \rightarrow (.1, .2, .6)$ | 0.012 | 0.031 | 0.035 |
| | $(.1, .1, .8) \rightarrow (.2, .1, .8)$ | 0.007 | 0.020 | 0.037 |

또는 한 번 있는 모형 중 어떤 모형을 선택하는지를 확인해 본 결과인데, AIC에 비하여 BIC는 완벽하게 변동성 변화가 없는 모형을 찾아냄을 확인할 수 있다. Table 5.4는 변동성 변화가 한 번 있는 모형에 대한 결과인데, AIC보다는 BIC가 변동성 변화가 한 번 있다는 것을 잘 찾아냄을 확인할 수 있다. 이 결과들에 비추어 보면 AIC는 BIC보다 변화점의 개수를 과대추정하는 경향이 있는 것으로 확인된다. 그러므로 본 논문에서는 이후의 변화점 탐지방법에 대해서 AIC보다는 BIC를 이용하고자 한다.

다음 Table 5.5와 Table 5.6은 변동성 변화가 한 번 있는 모형과 장기역성 모형을 구분하기 위한 CUSUM 검정통계량 M_n , $M_{n,B}$ 및 $M_{n,V}$ 의 size와 검정력에 대한 결과이다. Table 5.5는 자료의 중간, 즉 $k^* = n/2$ 인 위치에서 변동성 변화가 한 번 있는 GARCH(1, 1) 모형의 size 결과이다. Table 5.1의 변동성 변화가 없는 경우와 마찬가지로 n 이 증가할수록 0.05에 가까운 값이 나오는 것을 확인할 수 있다. Table 5.6은 Table 5.5와 동일한 자료에 대하여 FIGARCH(1, d , 1) 모형을 이용한 변동성 변화가 있는 모형의 검정력을 평가한 결과이다. 로그수익률 자료를 이용한 결과보다 모형을 적합한 후의 잔차에 대한 검정력이 더 좋은 것을 확인할 수 있다. 특히, 적률추정량을 이용한 long-run 분산의 추정량 $\hat{\nu}$ 인 경우가 더 좋은 것을 확인할 수 있다.

Table 5.6. Empirical power of CUSUM test for $H_0 : VS-1$ vs. $H_1 : LRD$ (FIGARCH(1, d , 1))

| | (ω, β, ϕ, d) | $n = 1,000$ | $n = 2,000$ | $n = 5,000$ |
|-----------|----------------------------|-------------|-------------|-------------|
| M_n | (.3, .1, .2, .2) | 0.051 | 0.146 | 0.308 |
| | (.3, .1, .2, .4) | 0.102 | 0.231 | 0.408 |
| | (.6, .2, .2, .2) | 0.057 | 0.147 | 0.320 |
| | (.6, .2, .2, .4) | 0.112 | 0.275 | 0.480 |
| $M_{n,B}$ | (.3, .1, .2, .2) | 0.102 | 0.200 | 0.403 |
| | (.3, .1, .2, .4) | 0.238 | 0.475 | 0.698 |
| | (.6, .2, .2, .2) | 0.083 | 0.188 | 0.353 |
| | (.6, .2, .2, .4) | 0.220 | 0.458 | 0.686 |
| $M_{n,V}$ | (.3, .1, .2, .2) | 0.225 | 0.387 | 0.593 |
| | (.3, .1, .2, .4) | 0.420 | 0.617 | 0.841 |
| | (.6, .2, .2, .2) | 0.146 | 0.246 | 0.476 |
| | (.6, .2, .2, .4) | 0.325 | 0.542 | 0.784 |

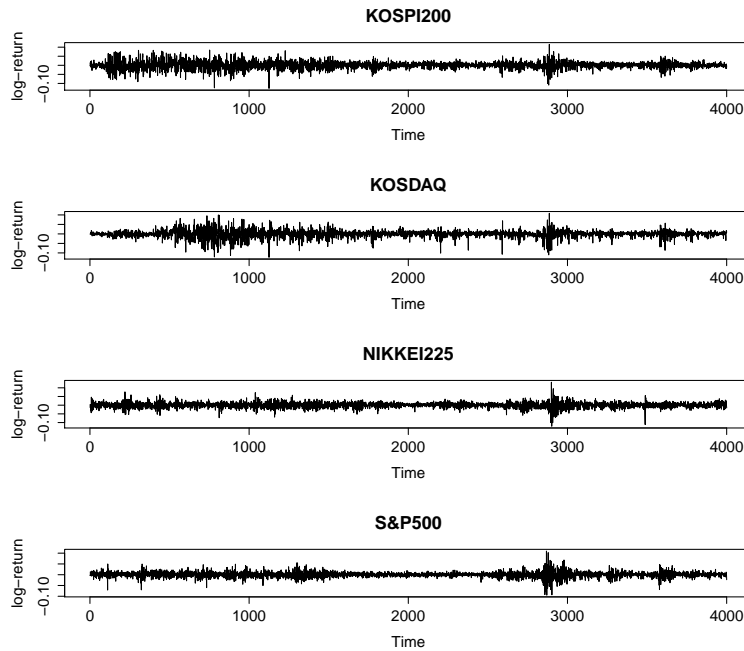


Figure 6.1. Time series plots of log returns of stock indices: KOSPI200, KOSDAQ, NIKKEI225, S&P500

6. 실증 자료 분석

이 장에서는 앞에서 제안한 검정법을 활용하여 국내외 주요 주가지수의 변동성에 변화가 있는지 또는 장기억성의 성질이 있는지 살펴보기로 한다. 분석에서 사용하게 될 자료는 한국, 일본, 미국의 주요 주가지수인 KOSPI200, KOSDAQ, NIKKEI225, S&P500의 로그수익률이며, 2013년 4월 15일까지 최근 4000개를 이용한다. Figure 6.1은 4개의 주가지수에 대한 로그수익률 그래프이다. 4개의 수익률이 비슷한 시점에서 변동성이 갑자기 급증했다가 다시 돌아오는 것을 확인할 수 있는데, 이는 2008년 금융위기 때의 변화를 나타내는 것이다.

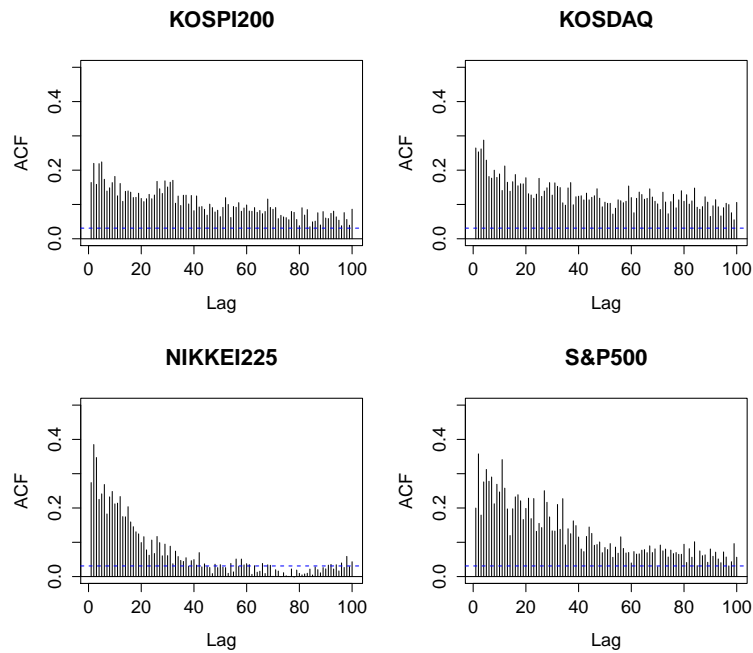


Figure 6.2. ACF plots of squared log returns of stock indices: KOSPI200, KOSDAQ, NIKKEI225, S&P500

Table 6.1. Estimates coefficients of GARCH(1, 1) models for log returns of stock indices: KOSPI200, KOSDAQ, NIKKEI225, S&P500

| 주가지수 | $\hat{\omega}$ | $\hat{\alpha}$ | $\hat{\beta}$ | $\hat{\alpha} + \hat{\beta}$ |
|-----------|----------------|----------------|---------------|------------------------------|
| KOSPI200 | 1.555e-06 | 0.0726 | 0.9258 | 0.9984 |
| KOSDAQ | 6.016e-06 | 0.1676 | 0.8299 | 0.9975 |
| NIKKEI225 | 4.786e-06 | 0.1000 | 0.8818 | 0.9818 |
| S&P500 | 1.756e-06 | 0.0894 | 0.9009 | 0.9903 |

Figure 6.1을 살펴보면 KOSPI200, KOSDAQ의 경우 n 이 약 1,500일 때 변동성 변화가 있는 것처럼 보이고, S&P500의 경우는 n 이 1,500 그리고 2,500일 때 변동성 변화가 있는 것처럼 보이며, NIKKEI225의 경우는 변동성 변화가 크게 나타나 보이지는 않는다. Figure 6.2는 각 주가지수의 로그 수익률 제곱값에 대한 ACF를 그려본 결과이다. Figure 6.2에서는 NIKKEI225를 제외한 나머지 주가지수의 로그수익률 제곱값에 대한 ACF가 모두 천천히 감소하므로 장기역성 모형을 따르는 것처럼 보인다. Table 6.1은 각각의 수익률 자료에 GARCH(1, 1)을 적합 시킨 후의 추정량이다. 4가지의 주가지수의 계수 추정량의 합인 $\hat{\alpha} + \hat{\beta}$ 의 값을 보면 모두 1에 가까운 값을 가지므로 강한 지속성을 나타내는 IGARCH 모형처럼 보이는 것을 확인할 수 있다. 4가지 주가지수에 대해 살펴본 결과, 정확하게 어떠한 모형을 따른다고 결론을 내릴 수 없기 때문에 본론에서 정의한 검정법을 이용하여 각각 어떤 모형을 만족하는지 알아보았다.

정확한 모형을 알아보기 위해서 변화점을 탐지하기 위해 변화가 없는 모형, 한 번 있는 모형 및 두 번 있는 모형에 대한 각각의 BIC 값을 계산해 본 결과가 Table 6.2이다. KOSPI200과 NIKKEI225는 $R = 0$ 일 때 BIC 값이 가장 작은 값을 가졌고, KOSDAQ과 S&P500은 $R = 2$ 일 때 BIC 값이 가장 작은 값을 가졌다. 위의 결과를 이용하여 KOSPI200과 NIKKEI225는 변동성 변화가 없는 모형과 장기역성

Table 6.2. Estimated number of volatility shifts based on BIC criterion

| 변화점의 개수 | 0 | 1 | 2 |
|-----------|---------|---------|---------|
| KOSPI200 | -5.3089 | -5.3072 | -5.3080 |
| KOSDAQ | -5.4630 | -5.4714 | -5.4759 |
| NIKKEI225 | -5.6917 | -5.6901 | -5.6872 |
| S&P500 | -6.1993 | -6.2003 | -6.2034 |

Table 6.3. Results of CUSUM test for $H_0 : VS-R$ vs. $H_1 : LRD$

| 주가지수 | $\tilde{T}_{n,V}, M_{n,V}$ | p -value | 선택된 모형 |
|-----------|----------------------------|------------|--------|
| KOSPI200 | 1.2448 | 0.0902 | IGARCH |
| KOSDAQ | 11.8597 | <0.0001 | 장기역성 |
| NIKKEI225 | 1.0751 | 0.1980 | IGARCH |
| S&P500 | 9.6866 | <0.0001 | 장기역성 |

을 구분하는 가설 $H_0 : VS-0$ vs. $H_1 : LRD$ 에 대한 검정을 실시하였고, KOSDAQ과 S&P500은 변동성 변화가 두 번 있는 모형과 장기역성을 구분하는 가설 $H_0 : VS-2$ vs. $H_1 : LRD$ 에 대한 검정을 실시하였다. 이를 시행한 결과를 Table 6.3에 나타내었다. 최종적으로 KOSPI200과 NIKKEI225는 IGARCH 모형이, KOSDAQ과 S&P500은 장기역성 모형을 만족하는 자료임을 확인할 수 있다.

7. 결론

본 논문에서는 변동성 변화모형과 장기역성 모형을 구분하기 위해 GARCH 모형 적합 후 잔차를 이용한 CUSUM 검정통계량을 제안하였다. CUSUM 검정통계량을 정의하기 위해 필요한 long-run 분산의 추정량으로 의존성을 고려하는 경우 Bartlett 커널 추정량, 독립임을 가정한 경우 적률추정량을 이용하였다. 다양한 모의실험 결과에 의하면 long-run 분산의 적률추정량과 GARCH 모형의 잔차를 이용한 CUSUM 검정통계량이 다른 방법에 비해서 size 왜곡현상 없이 우수한 검정력을 나타내고 있음을 확인하였다. 가설검정에 앞서 귀무가설하의 모형인 VS-R에서 변동성 변화 횟수를 추정하기 위하여 AIC와 BIC를 이용한 모수적인 방법을 사용하였다. 모의실험 결과 과대추정의 성질이 있는 AIC보다는 BIC를 사용하여 변화점의 개수를 추정하는 것이 바람직하다는 사실을 발견하였다. 본 논문에서 제안된 CUSUM 검정방법을 KOSPI200, KOSDAQ, NIKKEI225 그리고 S&P500에 적용시켜본 결과, KOSPI200과 NIKKEI225의 경우에는 IGARCH 모형이 선택되었고, KOSDAQ과 S&P500의 경우는 변동성 변화 모형보다는 장기역성 모형이 선택되었다.

본 논문에서 제안된 검정통계량의 극한분포는 기존 논문의 결과를 토대로 얻었으나, 몇몇 통계량은 엄밀한 증명이 추가적으로 필요하다. 다음으로 모의실험의 결과에서 size가 유의수준 5%와 약간의 차이가 있고 장기역성 모형인 경우의 검정력이 높지 않음을 확인할 수 있어 실증분석 결과의 신뢰성에 대해 논란의 소지가 있을 것으로 보인다. 이는 Berkes 등 (2006), Baek과 Pipiras (2012) 등의 논문에도 잘 나타나 있다. 따라서 CUSUM 검정통계량이 아닌 장기역성 모수 추정량 등을 이용한 검정통계량을 고려해 볼 필요가 있다. 끝으로 AIC 또는 BIC를 이용한 방법 이외에 변동성 변화를 탐지하는 효율적인 방법 개발과 이에 대한 극한 이론에 대한 추가적인 연구도 진행되어야 할 것이다.

References

- Andrews, D. (1991). Heteroscedasticity and Autocorrelation Consistent Covariance Matrix Estimation, *Econometrica*, **59**, 817–858.

- Baek, C. and Pipiras, V. (2012). Statistical tests for a single change in mean against long-range dependence, *Journal of Time Series Analysis*, **33**, 131–151.
- Baillie, R. T., Bollerslev, T. and Mikkelsen, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, **74**, 3–30.
- Berkes, I., Horvath, L. and Kokoszka, P. (2003). GARCH processes: structure and estimation, *Bernoulli*, **9**, 201–227.
- Berkes, I., Horvath, L., Kokoszka, P. and Shao, Q.-M. (2006). On discriminating between long-range dependence and changes in mean. *Annals of Statistics* **34**, 1140–1165.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, **31**, 307–327.
- Diebold, F. X. and Inoue, A. (2001). Long memory and regime switching, *Journal of Econometrics*, **105**, 131–159.
- Ding, Z., Granger, C. W. J. and Engle, R. F. (1993). A long memory property of stock market returns and a new model, *Journal of Empirical Finance*, **1**, 83–106.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica*, **50**, 987–1007.
- Engle, R. F. and Bollerslev, T. (1986). Modelling the persistence of conditional variances, *Econometric Reviews*, **5**, 1–50.
- Franco, C. and Zakoian, J. (2004). Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes, *Bernoulli*, **10**, 605–637.
- Hillebrand, E. (2005). Neglecting parameter changes in GARCH models, *Journal of Econometrics*, **129**, 121–138.
- Klemeš, V. (1974). The Hurst phenomenon: a puzzle?, *Water Resources Research*, **10**, 675–688.
- Kulperger, R. and Yu, H. (2005). High moment partial sum processes of residuals in GARCH models and their applications, *The Annals of Statistics*, **33**, 2395–2422.
- Lee, T., Kim, M. and Beak, C. (2014). Tests for volatility shifts in GARCH models against long-range dependence, submitted for publication.
- Mikosch, T. and Stărică, C. (2004). Nonstationarities in financial time series, the long-range dependence, and the IGARCH effects, *The Review of Economics and Statistics*, **86**, 378–390.
- Pooter, M. D. and Dijk, D. V. (2004). Testing for changes in volatility in heteroskedastic time series—a further examination, Econometric Institute Report EI 2004-38, Erasmus University Rotterdam, Econometric Institute.
- Teverovsky, V., Taqqu, M. S. and Willinger, W. (1999). A critical look at Lo’s modified R/S statistic, *Journal of Statistical Planning and Inference*, **80**, 211–227.

변동성 변화와 장기억성을 구분하는 CUSUM 검정통계량에 대한 실증분석

이영선^a · 이태욱^{a,1}

^a한국외국어대학교 통계학과

(2013년 12월 31일 접수, 2014년 2월 26일 수정, 2014년 3월 14일 채택)

요약

금융시계열 자료의 변동성에 나타나는 대표적인 현상 중에 지속성(persistence)이 있는데, 이를 설명하기 위하여 IGARCH 모형이 주로 사용된다. 최근에 변동성의 지속성은 변동성 변화와 장기억성에 기인한다는 사실이 많은 연구 결과에서 발표되고 있을 뿐만 아니라 장기억성은 변동성 변화로, 변동성 변화는 장기억성으로 보이게 되는 현상이 빈번히 나타난다. 따라서 본 논문에서는 변동성의 지속성, 장기억성 및 변동성 변화를 구분하는 통계적인 방법론을 고려하였다. 이를 위해 GARCH 모형 잔차를 기반으로 하는 CUSUM 통계량을 도입하여, size 왜곡(distortion) 현상을 해결할 뿐만 아니라 우수한 검정력을 얻을 수 있음을 입증하였다. 한편 변동성 변화가 존재하는 경우 변화점 추정이 중요해 지는데, 이를 위해 GARCH 모형을 기반으로 한 AIC 방법과 BIC 방법을 비교하였다. 다양한 모의 실험과 실증자료를 분석하여 우리가 제안하는 잔차 기반의 CUSUM 통계량의 우수성을 입증하였다.

주요용어: 변동성 변화, 장기억성, 지속성, CUSUM 검정, GARCH 모형.

이 논문은 2013년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (NRF-2013R1A1A2006796).

¹교신저자: (449-791) 경기도 용인시 모현면 산 89, 한국외국어대학교 통계학과. E-mail: twlee@hufs.ac.kr