Journal of Korean
Data & Information Science Society
2004, Vol. 15, No. 2, pp. 441~447

A Robust Heteroscadastic Test for ARCH Models[1)]

Sahmyeong Kim[2)]

Abstract

Li and Mak (1994) developed a test statistic for detecting the non-linearity and the heteroscedasticity of the time series data. But it is well known that the test statistic may be very sensitive in case of heavy-tailed distributions of the errors. Jiang et al.(2001) suggested the robust method for ARCH models but the calculation procedures for the estimation are very complicated. We suggested the robust method based on Huber's function and our method works quite well rater than the Li and Mak(1994). Also our method is relatively easy to calculate the test statistic.

Keywords : ARCH모형, 로버스트 통계량, 이분산성 검정통계량

1. 서 론

대부분 시계열 자료의 분석 방법은 등분산성을 가정한 모형들이 많이 활용 되었으나, 주가, 이자율, 환율 등과 같은 재정 자료는 시간의 추이에 따라 변동성(volatility)이 매우 심한 특징을 가지고 있다. 변동성이 강한 자료에서 먼저 관찰할 수 있는 점은 자료의 급락과 급상승이 자주 반복된다는 사실이다. 변동성은 통계학의 관점에서 볼 때 분산에 해당되는 개념임으로 강한 변동성이 나타나는 이유는 미래값의 분산(변화가능의 폭)이 현재의 상황에 전적으로 의존하면서 결정되기 때문이다. 만약 현재의 상황에 별 관계가 없다면 변동성의 급락 또는 급상승과 같은 효과를 기대할 수는 없을 것이다. 그래서 자료의 현시점에서 변동폭, 즉 분산이 최근의 자료에 영향을 받는다는 것에 착안하여 Engle(1982)이 오차의 분산이 자기회귀적으로 변하는 조건부 이분산 자기회귀모형(autoregressive conditional heteroskedastic model : ARCH model)을 제시하였다.

통계적인 모형수립은 전통적으로 모형식별, 모형적합, 모형진단의 반복적인 과정을 통해 수립된 모형으로 예측을 하는 것이다. 특히 모형 진단 단계에서 유용한 분석법

---

1) 이 논문은 2003년도 중앙대학교 학술연구비 지원에 의한 것임.
2) 서울시 동작구 흑석동 221번지 중앙대학교 통계학과 부교수
E-mail : sahm@cau.ac.kr

은 잔차(residual)를 이용한 분석법이다.

Box와 Pierce(1970)에 의해 포트맨토우 검정 (portmanteau test)이 제안되었으며, n이 작을 때 Q^2 통계량의 값이 너무 보수적이 되어 실제로는 모형이 잘 적합 되지 않았는데도 모형이 잘 적합 된 것으로, 즉 귀무가설을 기각하지 못하는 것으로 판정하는 경향이 있음이 알려져 있다.

Ljung과 Box(1978)는 이를 보완하여 자기회귀이동평균모형(autoregressive moving average model : ARMA)이 옳다는 귀무가설 하에 근사 $\chi^2(K-p-q)$분포를 따르는 검정통계량을 제안하였다. Ljung과 Box 통계량은 일반적으로 로버스트(robust)한 것으로 알려져 있다. 여기서 p와 q는 $ARMA$모형의 모수의 차수이다. 한편, McLeod(1979)는 자기회귀이동평균모형에서 제곱잔차자기상관(squared residual autocorrelation)을 기초로 하여 모형 타당성을 위한 포트맨토우 검정을 유도하였고 McLeod와 Li(1983)는 제곱잔차자기상관의 점근분포가 동질성의 귀무가설 하에서 $ARMA$ 모수들에 의존하지 않는다는 것을 보였으며 결과적으로 보통의 표준오차(standard error) $1/\sqrt{n}$을 모형 점검에 사용할 수 있게 되었다.

조건부이분산성 시계열 모형의 모형화에 대해 많은 사람들이 연구해 왔지만, 실제적으로 그런 모형들에서 유도된 제곱잔차자기회귀의 점근분포에 대한 결과는 없었다. Li와 Mak(1994)은 조건부이분산성 과정에 대해 제곱잔차자기상관을 고려하여 점근분포를 유도 하였고, 그 결과 조건부 이분산성 시계열모형의 타당성을 확인하는데 유용하다. Jiang등(2001)이 L_1-추정의 방법을 제안하였지만 이는 추정방법이 복잡하다는 단점이 있어 본 논문에서는 최소제곱추정법으로 모수를 추정하여 Li와 Mak(1994)이 제안한 Q통계량의 자기상관을 보정해 $ARCH$모형을 진단하는 방법에 대해 논의 하려 한다.

2. 이분산 검정통계량

$ARCH$모형에는 $h_t = \alpha_0 + \alpha_1 \epsilon_{t-1}^2$에 ϵ에 대한 제곱의 항이 있는 반면에 포트맨토우 검정은 $(\epsilon_t - \bar{\epsilon}_t)^2$여기에 제곱항이 없으므로 이분산성에 대한 검정을 할 수 없다.

다음의 모형을 고려해 보자.

$$\begin{aligned} y_t &= \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \sqrt{h_t} e_t, \\ \epsilon_t &= \sqrt{h_t} e_t,\ e_t \sim iid\ (0, \sigma^2)\text{이고}, \\ h_t &= \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \cdots + \alpha_p \epsilon_{t-p}^2\text{이다.} \end{aligned} \tag{1}$$

잔차에 제곱을 하여 $ARCH$ 모형에 대한 검정을 수행할 수 있는 검정통계량은 Li(1994)이 제시한 검정통계량으로서 다음과 같다.

$$r_k = \frac{\sum_{t=k+1}^{n} (\hat{\epsilon}_t^2/\hat{h}_t - \bar{\epsilon})(\hat{\epsilon}_{t-k}^2/\hat{h}_{t-k}^2 - \bar{\epsilon})}{\sum_{t=1}^{n} (\hat{\epsilon}_t^2/\hat{h}_t - \bar{\epsilon})^2}, \quad k = 1, 2, \cdots,$$

$$\bar{\epsilon} = n^{-1}\sum_{t=1}^{n}\hat{\epsilon}_t^2/\hat{h}_t \tag{2}$$

n은 표본크기이며, 모형이 적당하다면 $\bar{\epsilon}$은 1에 확률적으로 수렴한다. r_k을 아래와 같이 바꿔 쓸 수 있다.

$$r_k = \frac{\sum_{t=k+1}^{n} (\hat{\epsilon}_t^2/\hat{h}_t - 1)(\hat{\epsilon}_{t-k}^2/\hat{h}_{t-k}^2 - 1)}{\sum_{t=1}^{n}(\hat{\epsilon}_t^2/\hat{h}_t - 1)^2} \tag{3}$$

여기서 $n^{-1}\sum_{t=1}^{n}(\hat{\epsilon}_t^2/\hat{h}_t - 1)^2$은 어떤 상수로 수렴하게 된다. 따라서 아래의 점근분포만 고려하면 된다.

$$\hat{C}_k = \frac{1}{n}\sum\left(\frac{\hat{\epsilon}_t^2}{\hat{h}_t} - 1\right)\left(\frac{\hat{\epsilon}_{t-k}^2}{\hat{h}_{t-k}} - 1\right) \tag{4}$$

다음의 정리는 Li와 Mak의 통계량의 점근적 성질을 나타낸다.
정리2.1 주어진 조건에 의하여 다음과 같은 등식이 성립한다.

$$(1)\ \sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N(0, V) \tag{5}$$

$$(2)\ Q(M) = n\hat{r}^T \hat{V}^{-1}\hat{r} \sim \chi^2(M)$$

여기서 $\boldsymbol{X} = (X_1, \cdots, X_M)'$, $X_k = -\frac{1}{n}\sum \frac{1}{h_t}\frac{\partial h_t}{\partial \theta}\left(\frac{\epsilon_{t-k}^2}{h_{t-k}} - 1\right)$이고, $G = \boldsymbol{X}'\boldsymbol{X}$이고, $V_1 = I - \frac{1}{4}XG^{-1}X^T$이다.

증명: Li와 Mak(1994)을 참조.

3. 로버스트 이분산 검정통계량

Jiang등(2001)이 주장하기를 Li와 Mak의 검정통계량은 과대추정(over-estimate)하는 경향이 있다고 하였다. 즉 식(2)의 r_k값이 커져서 귀무가설을 기각하는 경향이 있다고 하였다. 그래서 우리는 로버스트 검정통계량을 제시하여 Li와 Mak의 통계량을 개선하고자 한다. 제시된 로버스트 제곱잔차자기상관함수는 아래와 같다.

$$r_k^* = \frac{\sum_{t=k+1}^{n} \psi(\frac{\hat{\epsilon}_t^2}{\hat{h}_t} - 1)\psi(\frac{\hat{\epsilon}_{t-k}^2}{\hat{h}_{t-k}} - 1)}{\sum_{t=1}^{n} \psi^2(\frac{\hat{\epsilon}_t^2}{\hat{h}_t} - 1)} \tag{6}$$

여기서

$$\psi(x) = \begin{cases} x & |x| \le m \\ m & x > m \\ -m & x < -m \end{cases} \quad , \ m > 0,$$

식 (6)에서 분모$n^{-1}\sum\psi^2(\hat{\epsilon}_t^2/\hat{h}_t - 1)$ 은 상수로 수렴한다. 따라서 아래와 같이 로버스트 제곱잔차자기상관함수를 정의할 수 있다.

정의 : $\hat{r_k} = \frac{1}{C}\sum\psi(\frac{\hat{\epsilon}_t^2}{\hat{h}_t} - 1)\psi(\frac{\hat{\epsilon}_{t-k}^2}{\hat{h}_{t-k}} - 1),\ k = 1, \cdots, M,$ (7)

$\frac{1}{n}\sum\psi^2\left(\frac{\hat{\epsilon}_t^2}{\hat{h}_t} - 1\right)$ 이며, 만일 $\psi(x) = x$ 라면 Li와 Mak의 통계량과 같다는 점이며, Li와 Mak 통계량보다 로버스트 할 수 있다는 것이다.

로버스트 검정통계량 $\hat{r}_k$에 대한 점근적 분포는 다음의 정리에서 구할 수 있다.

정리3.1. 다음의 조건을 고려해보자.

(1) $\psi(x)$는 미분가능하고 제한된(bounded) 함수이며, $E\psi(x) = 0$이다.

(2) $\{Y_t\}$는 정상성을 가지는 자료이며 $EY_t^2 < \infty$이다.

위의 조건하에서 다음의 결과를 얻을 수 있다.

(1) $\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N(0, V_2)$ (8)

(2) $Q^*(M) = n\hat{r}^T \hat{V}_2^{-1}\hat{r} \sim \chi^2(M)$

여기서 $V_2 = I - \frac{1}{4}YH^{-1}Y^T$이고, $H = (\boldsymbol{Y'Y})$이며, $\hat{r} = (r_1, \cdots, r_M)$,

$Y_t = -\frac{1}{n}\sum\psi\left(\frac{1}{h_t}\frac{\partial h_t}{\partial\theta}\right)\psi\left(\frac{\epsilon_{t-k}^2}{h_{t-k}} - 1\right)$이며, $\boldsymbol{Y} = (Y_1, \cdots, Y_M)'$이다.

4. 모의실험

우리가 고려한 $AR(1) - ARCH(1)$모형은 아래와 같다.

$$y_t = \phi y_{t-1} + \sqrt{h_t}e_t \qquad \{e_t\} \sim IID\, N(0,1) \tag{9}$$
$$h_t = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 \qquad \alpha_0 > 0,\ \alpha_1 \geq 0$$

먼저 초기치를 $\phi = 0.6,\ \alpha_0 = 0.8,\ \alpha_1 = 0.4$를 주고, 최소제곱법에 의한 모수추정값을 다음과 같이 구하겠다.

[표 1] AR(1)-ARCH(1)모형의 모수추정 결과

	$\hat{\phi}$	$\widehat{\alpha_0}$	$\widehat{\alpha_1}$
Normal	0.6001686	0.8202159	0.3849366

모의실험 과정을 간편하게 수행하기 위해 $ARCH$모형만 고려하였으며, α_0, α_1의 값을 초기값으로 하여 변화시켜가면서 Li와 Mak의 Q통계량과 본 논문의 로버스트 Q통계량을 비교 하였다. 결과는 각 Case별로 100번씩 반복한 평균을 기재한 것이다. Normal은 $N(0,1)$에서 100%추출한 것이고, Mixed Normal은 10%를 $N(0,5^2)$에서 추출하고, 나머지 90%를 $N(0,1)$에서 추출하였다.

아래의 결과는 $\alpha_0 = 0.1, \alpha_1 = 0.1$의 초기치를 부여하여 실행한 결과이다.

[표 2] 초기치가 $\alpha_0 = 0.1,\ \alpha_1 = 0.1$일 때 Case별 Normal과 Mixed Normal의 Q통계량 비교

	Normal		Mixed Normal	
Case	Q(Li&Mak)	Q(로버스트)	Q(Li&Mak)	Q(로버스트)
100	7.0765398	7.452158	5.7803127	7.5604142
400	7.7896262	8.7693789	10.456276	8.1428579
800	7.9025787	10.618723	13.858426	8.441286
1000	7.6569866	10.814108	17.061577	9.070238

[표 2]는 오차의 분포가 Normal인 경우 Li와 Mak의 Q계량과 로버스트 Q통계량 모두 모형이 적합[3]하다고 판단하였다. 그러나 오차의 분포가 Mixed Normal인 경우, Li와 Mak의 Q통계량은 모형이 적합하지 않다고 판단한 반면 로버스트 Q통계량은 적합하다고 판단한다. 즉 오차의 분포가 정규분포보다 두터운 꼬리를 가진 경우 Li와 Mak의 Q통계량은 잘못된 결과가 도출될 수 있음을 보여준다.

3) $\chi^2(8, \alpha = 0.1) = 13.362$

[표 3] 초기치가 $\alpha_0 = 0.2$, $\alpha_1 = 0.2$일 때 Case별 Normal과 Mixed Normal의 Q통계량 비교

	Normal		Mixed Normal	
Case	Q(Li와Mak)	Q(로버스트)	Q(Li와Mak)	Q(로버스트)
100	6.3231469	7.4522757	5.1009236	7.4717936
400	7.8229707	9.0307063	9.228747	8.276476
800	7.3457874	10.615813	13.02607	7.4245983
1000	7.6733098	10.507971	14.44110	7.8341732

[표 4] 초기치가 $\alpha_0 = 0.3$, $\alpha_1 = 0.3$일 때 Case별 Normal과 Mixed Normal의 Q통계량 비교

	Normal		Mixed Normal	
Case	Q(Li와Mak)	Q(로버스트)	Q(Li와Mak)	Q(로버스트)
100	7.1012013	7.2274237	6.4083678	7.4948819
400	8.0598439	9.598843	11.742365	8.9399277
800	8.2351751	11.216512	14.587066	9.0762078
1000	7.1366265	11.050918	20.797345	9.1598296

[표 5] 초기치가 $\alpha_0 = 0.4$, $\alpha_1 = 0.4$일 때 Case별 Normal과 Mixed Normal의 Q통계량 비교

	Normal		Mixed Normal	
Case	Q(Li와Mak)	Q(로버스트)	Q(Li와Mak)	Q(로버스트)
100	6.3180509	6.9454149	5.1176275	7.4389183
400	7.6611936	10.753203	9.3750309	7.6148595
800	7.6001048	11.77611	11.824656	7.7449014
1000	7.6465857	12.366293	13.466356	8.4602769

각 α_0, α_1 값에 따라 설명되지 않는 오차를 Mixed Normal로 한 검정에서 Li와 Mak의 Q통계량이 과대 추정될 수 있음을 보여준다.

5. 결 론

분산이 최근의 자료에 영향을 받는다는 것에 착안하여 Engle(1982)이 오차의 분산이 자기회귀적으로 변하는 조건부이분산자기회귀모형을 제시한 이래로 Li와 Mak(1994)이 조건부이분산성에 대해 제곱잔차자기상관을 고려하여 점근분포를 유도

하였고, 그 결과 조건부 이분산성 시계열모형의 타당성을 확인하는데 유용하게 사용되었다. 또한 Jiang등(2001)이 L_1-추정 방법을 사용한 $ARCH$모형의 로버스트 모형화에 대해 제안하였지만, 추정방법이 복잡하다는 단점이 있어 본 논문에서는 최소제곱추정법으로 모수를 추정하여 Li와 Mak(1994)이 제안한 Q통계량의 자기상관을 보정해 $ARCH$모형 진단 방법에 대한 논의 결과 정규성으로부터 멀어질수록 Li와 Mak(1994)이 제안한 Q통계량이 잘못된 결과를 도출한 반면 본 논문의 로버스트 검정통계량은 설명되지 않는 오차의 분포가 정규성에서 멀어지더라도 로버스트한 결과를 보여 주었다.

향후 과제로 다른 로버스트 함수를 사용하여 비교해 보는 것도 의미 있을 것이며, 더 나아가 $ARCH$ 모형을 일반화한 $GARCH$ 모형으로의 확장에 대한 논의도 고려해 볼만할 것이다.

참고문헌

1. Bollerslev, T. (1986). Generalized autoregressive conditional heterocedasticity. *Journal of Econometrics,* 31, 307-327.
2. Box, G. E. P. and Pierce, D. A. (1970). Distribution of the residual autocorrelation in autoregressive integrated moving average time series models. *J. Am. Stat. Assoc.* 65, 1509-26.
3. Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity and the Variance of U.K.Inflation. *Econometrica*, 46, 1287-1294.
4. Huber, P. J. (1964). Robust estimation of a location Parameter, *The Annals of Mathematical Statistics,* 35, 73-101.
5. Jiang, J., Zhao, Q. & Hui, Y, V. (2001). Robust Modelling of ARCH Models. *J. Forecast*, 20, 111-133.
6. Li, W. K. & Mak, T. K. (1994). On the squared residual autocorrelations in non-linear time series with conditional heteroscedasticity. *Journal of Time Series Analysis,* 15, 627-636.
7. Ljung, G. M. and G. E. O. Box. (1978). On a measure of lack of fit in time series models. *Journal of Applied Probability*, 25, 553-64.
8. Tong, H. (1990). Nonlinear time series. *Oxford University Press, Oxford.*
9. Weiss, A. A. (1984). Asymptotic theory for ARCH models : estimation and testing. *Econometric Theory,* 2, 107-131.
10. Wong, H. and Li, W. K. (2002). Detecting and Diagnostic Checking ultivariate Conditional Heteroscedastic Time Series Models. *Analysis of Institute of Statistical Mathematics,* 54, 45-59.

[2004년 2월 접수, 2004년 5월 채택]