

붓스트랩 방법을 이용한 일반화 자기회귀 조건부 이분산모형에서의 조건부 분산 예측

김희영^a, 박만식^{1,a}

^a고려대학교 의학통계학교실

요약

일반적으로 일반화 자기회귀 조건부 이분산(GARCH)모형 하에서, 우도함수에 기반한 자료의 예측구간의 추정치는 오차항의 분포에 민감하게 반응하고 더욱이 조건부 분산의 경우 구간 추정이 현실적으로 쉽게 풀리지 않는 문제이다. 이를 해결하기 위해 붓스트랩방법(bootstrap method)이 적용될 수 있음을 최근 연구들을 통해 밝혀졌다. 본 논문에서는 GARCH모형 하에서 자료와 변동성(조건부 분산)의 예측구간 추정을 위해 최근 소개된 Pascual 등 (2006)의 논문을 토대로 붓스트랩 방법을 정리하였다. 실제 사례분석을 위해 국내 주가수익률자료를 이용하였다.

주요용어: 붓스트랩, 일반화 조건부 이분산 모형, 조건부 이분산.

1. 서론

주가, 환율 등 금융시계열자료에서 수익율($= (P_{t+1} - P_t)/P_t$, 여기서 P_t 는 t 시점의 지수)은 Rydberg (2000)가 요약한 바와 같이 다음과 같은 전형적인 특성들을 나타낸다. 첫째, 확률분포가 정규분포에 비하여 꼬리가 두껍다. 즉, 첨도(kurtosis)가 정규분포의 그것보다 크고, 이를 초과첨도(excess kurtosis, leptokurtosis)라 한다. 둘째, 큰 폭으로 움직이는 기간이 지속되다가 또 다른 기간에는 안정국면으로 접어드는 양상을 보이고, 셋째, 확률분포의 왜도(skewness)가 0이 아니라 음이나 양의 값을 갖는다. 대체적으로 오른쪽으로 꼬리가 긴 분포(positively skewed distribution)를 갖는다. 넷째, 장기간 의존성(long-range dependency)를 보인다.

Engle (1982)의 자기회귀 조건부 이분산(autoregressive conditional heteroscedastic: ARCH)모형과 이후 Bollerslev (1986)과 Taylor (1986)의 일반화 자기회귀 조건부 이분산(generalized autoregressive conditional heteroscedastic: GARCH)모형은 위에 열거한 첫째, 둘째 그리고 넷째 특징을 모형화를 통하여 설명하였다. 이 중에서 둘째 성질을 변동성 집중(volatility clustering)현상이라고 하는데, 큰 변동은 큰 변동끼리 작은 변동은 작은 변동끼리 몰리는 행태를 띤다. 금융시계열자료에서는 변동성이 시간에 따라 변하는 것은 일반적이며, 변동성은 위험(risk)을 재는 지표이다. 따라서 변동성에 대한 추정과 예측이 오랫동안 경제분야에서는 중요한 논제였다. 그리고, 미래시점에서의 예측은 점추정(point estimation)보다는 구간추정(interval estimation)을 통해서 훨씬 많은 정보를 얻을 수 있다.

수익율의 예측구간과 변동성의 예측구간은 GARCH 모형에서 오차항의 분포에 따라 달라진다. 현재 GARCH모형에서는 정규분포 이외에 여러 종류의 분포를 사용하고 있는데, 그 이유는 다음과 같다. 시계열이 GARCH과정을 따를 때 비조건부 분포(unconditional distribution)의 첨도는 조건부 분포의 첨도보다 더 커지게 된다. 즉, 오차항을 정규분포로 가정하는 경우 그 GARCH과정의 비조건부

¹ 교신저자: (136-701) 서울시 성북구 안암동 5가1 126-1 고려대학교 의과대학 의학통계학교실 및 의과학연구원 (유전체 및 단백질체 독성연구소), 연구교수. E-mail: man.sik.park@gmail.com

분포의 첨도는 정규분포보다 더 커지게 되므로, 금융시계열의 초과첨도를 일정부분은 설명하는데 도움이 되었다. 그러나, 실제로 금융시계열의 비정규성을 완전히 설명하는 데는 한계가 있다. 따라서, GARCH모형에서는 오차항의 분포로 여러가지의 분포들이 이용되고 있으며, 최근에는 초과첨도 뿐만 아니라 비대칭성도 표현할 수 있는 보다 다양한 분포들을 도입하고 있다.

수익율과 변동성의 예측구간(prediction interval)을 추정하기 위해서는 수익율의 예측오차의 분포와 변동성의 예측오차의 분포를 알아야 한다. 그러나, 이를 이론적으로 규명한 연구는 오차항이 정규분포를 따르는 GARCH(1,1)인 경우에도 이루어지지 않았다. 최근 Pascual 등 (2006)의 연구는 GARCH모형에서 오차항의 특정 분포를 가정하지 않고 수익율과 변동성을 붓스트랩방법(bootstrap method)을 적용하여 예측하였다. 본 논문에서는 Pascual 등 (2006)의 연구를 국내 주가자료에 적용해 보고자 한다.

본 논문의 구성은 다음과 같다. 우선 2절에서는 GARCH모형을 요약하고, 3절에서는 GARCH모형에 붓스트랩방법을 적용하여 수익율과 변동성의 예측구간을 연구한 논문들을 간단히 소개한다. 4절에서는 2005년 1월3일부터 2007년 12월31일까지의 기간동안 국내 2개 기업(삼성엔지니어링, 국민은행)의 수익율과 변동성을 예측하고자 한다. 마지막으로 5절에서는 결론을 제시하고자 한다.

2. GARCH모형

GARCH모형은 Bollerslev (1986)가 Engle (1982)의 ARCH모형을 일반화시킨 것이다. 다음의 식을 만족할 때 $\{X_t\}$ 는 GARCH(1, 1)모형을 따른다고 한다.

$$X_t = \sigma_t \epsilon_t, \quad \sigma_t^2 = \omega + \alpha X_{t-1}^2 + \beta \sigma_{t-1}^2. \quad (2.1)$$

여기서, $\omega > 0$, $\alpha \geq 0$, $\beta \geq 0$, $\{\epsilon_t\} \sim i.i.d(0, 1)$ 이고, ϵ_t 와 $\{X_{t-k}, k \geq 1\}$ 는 독립이고, σ_t^2 를 조건부 분산(conditional variance, conditional volatility)이라고 한다. 식 (2.1)의 확률과정이 공분산 정상성(covariance stationarity)을 만족하기 위한 필요충분조건은 $\alpha + \beta < 1$ 이다.

2.1. σ_{T+m}^2 의 예측

GARCH(1, 1)모형에서 T 시점까지의 자료 $\{x_1, \dots, x_T\}$ 가 관측되었을 때, X_{T+m} 과 σ_{T+m}^2 의 예측에 초점을 맞추어 주요한 성질을 열거하면 다음과 같다. 식 (2.1)에서 σ_t^2 는 $\{x_1, \dots, x_{t-1}\}$ 가 관측되었을 때 X_t 의 조건부 분산이 된다. 즉,

$$\begin{aligned} \text{Var}_{t-1}(X_t) &= E_{t-1}(X_t^2) = \sigma_t^2 = \frac{\omega}{1-\beta} + \alpha \sum_{j=0}^{\infty} \beta^j X_{t-1-j}^2 \\ &= \frac{\omega}{1-\alpha-\beta} + \alpha \sum_{j=0}^{\infty} \beta^j \left(X_{t-j-1}^2 - \frac{\omega}{1-\alpha-\beta} \right). \end{aligned}$$

그리고, GARCH(1,1)모형이 $\alpha + \beta < 1$ 조건을 만족하는 공분산 정상성인 경우, 무조건부 분산(marginal variance)은 $\text{Var}(X_t) = \omega/(1-\alpha-\beta)$ 이다. 또한, m 시차 후의 조건부 분산은 다음과 같이 표현된다.

$$\begin{aligned} \text{Var}_T(X_{T+m}) &= E_T(X_{T+m}^2) = E_T(\sigma_{T+m}^2) \\ &= \omega \sum_{i=1}^{m-1} (\alpha + \beta)^i + (\alpha + \beta)^{m-1} \text{Var}_T(X_{T+1}) \end{aligned} \quad (2.2)$$

$$= \omega \sum_{i=1}^{m-1} (\alpha + \beta)^i + (\alpha + \beta)^{m-1} (\omega + \alpha x_T^2 + \beta \sigma_T^2).$$

따라서, T 시점 까지의 자료, $\{x_1, \dots, x_T\}$ 가 관측되었을때 m 시차 후의 조건부 분산의 점추정값은 식 (2.2)를 사용한다. 그리고, 보다 많은 정보를 주는 예측구간을 원한다면, 예측오차 $v_{T,m} \equiv \sigma_{T+m}^2 - E_T(\sigma_{T+m}^2)$ 의 분포를 구해야 한다. 그러나, 현재까지의 연구는 Baillie와 Bollerslev (1992)가 예측오차 $v_{T,m}$ 의 조건부 평균제곱오차(conditional mean square error)를 구했고, 예측오차 $v_{T,m}$ 의 분포는 유도되지 않았다. $v_{T,m}$ 의 조건부 평균제곱오차는 다음과 같다.

$$E_T(v_{T,m}^2) = \text{Var}_T(\sigma_{T+m}^2) = (\kappa_2 - 1)\alpha^2 \sum_{i=1}^{m-1} (\alpha + \beta)^{2(i-1)} E_T(\sigma_{T+m-i}^4). \quad (2.3)$$

식 (2.3)에 제시된 κ_2 와 $E_T(\sigma_{T+m-i}^4)$ 의 정확한 표현식은 Baillie와 Bollerslev (1992)를 참조하기 바란다.

3. GARCH모형 하에서의 붓스트랩방법

이 절에서는 최근 소개된, GARCH모형 하에서 붓스트랩방법에 대해 연구한 논문들을 소개하고자 한다.

3.1. Miguel와 Olave (1999)의 연구

시계열자료, $\{x_t\}$ 가 다음의 ARMA(p, q)모형을 따르고, 오차항, $\{\epsilon_t\}$ 는 ARCH(r)과정이라고 하자. 즉,

$$\begin{aligned} (1 - \phi_1 B - \dots - \phi_p B^p) x_t &= (1 - \theta_1 B_1 - \dots - \theta_q B^q) \epsilon_t, \\ \epsilon_t &= \sigma_t \varepsilon_t, \\ \sigma_t^2 &= \omega + \sum_{i=1}^r \alpha_i \epsilon_{t-i}^2. \end{aligned} \quad (3.1)$$

여기서, $\{\varepsilon_t\}$ 는 평균이 0, 분산이 1인 *i.i.d.*한 확률분포가 F 를 따르고, $E(\varepsilon_t^4) < \infty$ 라고 가정하자. 관측된 자료가 $\{x_{1-p}, x_{2-p}, \dots, x_T\}$ 로 주어진 경우, Miguel과 Olave (1999)는 다음의 붓스트랩 방법을 제안하였다.

M.1 식 (3.1)에 제시된 모수에 대한 유사 최대우도추정량(Quasi maximum-likelihood estimator), $\widehat{\theta} = (\widehat{\phi}_1, \dots, \widehat{\phi}_p, \widehat{\theta}_1, \dots, \widehat{\theta}_q, \widehat{\omega}, \widehat{\alpha}_1, \dots, \widehat{\alpha}_r)$ 을 구한다.

M.2 다음의 표준화 잔차(standardized residual)를 계산한다.

$$\widehat{\varepsilon}_t = \frac{\widehat{\epsilon}_t}{\widehat{\sigma}_t} = \frac{1}{\widehat{\sigma}_t} \sum_{i=0}^{t-1} \widehat{\beta}_i \left(x_{t-i} - \sum_{n=1}^p \widehat{\phi}_n x_{t-i-n} \right), \quad t = 1, \dots, T.$$

여기서, $\widehat{\beta}_k$ 는 다음의 식으로부터 유도된다.

$$\sum_{k=0}^{\infty} \widehat{\beta}_k z^k = \left(1 + \sum_{i=1}^q \widehat{\theta}_i z^i \right)^{-1}.$$

그리고, 추정된 조건부 분산은 $\widehat{\sigma}_t^2 = \widehat{\omega} + \sum_{k=1}^r \widehat{\alpha}_k \widehat{\varepsilon}_{t-k}^2$ 이다.

M.3 다음의 분포 \widehat{F} 를 갖는 중심화 잔차를 생성한다.

$$\widehat{F}_T(x) = \frac{1}{T} \sum_{i=1}^T 1_{[\widehat{\varepsilon}_i - \bar{\varepsilon}, \infty)}(x).$$

여기서, $\bar{\varepsilon} = 1/T \sum_{i=1}^T \widehat{\varepsilon}_i$ 이다.

M.4 중심화잔차를 \widehat{F} 로부터 임의추출하여 ε_{T+j}^* , $j = 1, \dots, s$ 를 생성한다.

M.5 붓스트랩 표본 $\{x^*\}$ 과 조건부 분산 $\{\sigma^{*2}\}$ 를 다음과 같이 생성한다.

$$x_{T+j}^* = \sum_{i=1}^p \widehat{\phi}_i x_{T+j-i}^* + \sum_{i=1}^q \widehat{\theta}_i \varepsilon_{T+j-i}^* \sigma_{T+j-i}^{*2},$$

$$\sigma_{T+j}^{*2} = \widehat{\omega} + \sum_{i=1}^r \widehat{\alpha}_i \varepsilon_{T+j-i}^{*2} \sigma_{T+j-i}^{*2}, \quad j = 1, \dots, s.$$

여기서, $j \leq 0$ 에 대하여 $x_{T+j}^* = x_{T+j}$, $\varepsilon_{T+j}^* = \widehat{\varepsilon}_{T+j}$, $\sigma_{T+j}^{*2} = \widehat{\sigma}_{T+j}^2$ 이다.

M.6 위의 M.4와 M.5를 B 번 반복한다.

3.2. Pascual 등 (2006)의 연구

시계열자료, $\{x_t\}$ 가 아래의 GARCH(1,1)모형을 따른다고 하자.

$$x_t = \sigma_t \varepsilon_t,$$

$$\sigma_t^2 = \omega + \alpha x_{t-1}^2 + \beta \sigma_{t-1}^2, \quad t = 1, \dots, T. \quad (3.2)$$

여기서, $\alpha + \beta < 1$ 이다. $\{x_1, \dots, x_T\}$ 가 관측되었을 때 Pascual 등 (2006)은 다음의 붓스트랩방법을 제시하였다.

P.1 (3.2)에 제시된 모수에 대한 유사 최대우도추정량, $\widehat{\boldsymbol{\theta}} = (\widehat{\omega}, \widehat{\alpha}, \widehat{\beta})$ 을 구한다.

P.2 조건부 분산을 다음과 같이 추정한다.

$$\widehat{\sigma}_t^2 = \widehat{\omega} + \widehat{\alpha} x_{t-1}^2 + \widehat{\beta} \widehat{\sigma}_{t-1}^2, \quad t = 2, \dots, T.$$

여기서, $\widehat{\sigma}_1^2 = \widehat{\omega} / (1 - \widehat{\alpha} - \widehat{\beta})$ 로 정의하고 이는 $\{x_t\}$ 의 무조건부 분산의 추정값이다.

P.3 $t = 1, \dots, T$ 에 대하여 $\{\widehat{\varepsilon}_t \equiv x_t / \widehat{\sigma}_t\}$ 를 구한다.

P.4 $t = 1, \dots, T$ 시점에서의 붓스트랩 표본, $\{x_t^*\}$ 와 $\{\sigma_t^{*2}\}$ 을 다음과 같이 생성한다.

$$\widehat{\sigma}_t^{*2} = \widehat{\omega} + \widehat{\alpha} x_{t-1}^{*2} + \widehat{\beta} \widehat{\sigma}_{t-1}^{*2}, \quad x_t^* = \varepsilon_t^* \widehat{\sigma}_t^*.$$

여기서, $\widehat{\sigma}_1^{*2} = \widehat{\sigma}_1^2$ 이고, ε_t^* 은 P.3의 잔차를 중심화(centering)한 후 임의추출하여 얻어진다.

P.5 P.4의 붓스트랩 표본 $\{x_t^*\}$ 에 GARCH(1,1)모형을 적합하여 $\widehat{\boldsymbol{\theta}}^* = (\widehat{\omega}^*, \widehat{\alpha}^*, \widehat{\beta}^*)$ 을 구한다.

P.6 P.4의 $\hat{\theta}^* = (\hat{\omega}^*, \hat{\alpha}^*, \hat{\beta}^*)$ 를 이용하여 T 시점 이후의 붓스트랩 표본을 생성한다.

$$\begin{aligned}\hat{\sigma}_{T+j}^{*2} &= \hat{\omega}^* + \hat{\alpha}^* x_{T+j-1}^{*2} + \hat{\beta}^* \hat{\sigma}_{T-j-1}^{*2}, \\ x_{T+j}^* &= \varepsilon_{T+j}^* \hat{\sigma}_{T+j}^*, \quad j = 1, \dots, s.\end{aligned}$$

여기서, ε_{T+k}^* 은 P.3의 잔차를 중심화한 후 임의추출하여 얻어진다. 그리고, T 시점의 x_T^* 은 x_T 로, $\hat{\sigma}_T^{*2}$ 는 아래와 같이 구하게 된다.

$$\hat{\sigma}_T^{*2} = \frac{\hat{\omega}^*}{1 - \hat{\alpha}^* - \hat{\beta}^*} + \hat{\alpha}^* \sum_{j=0}^{T-2} \hat{\beta}^{*j} \left(x_{T-j-1}^2 - \frac{\hat{\omega}^*}{1 - \hat{\alpha}^* - \hat{\beta}^*} \right).$$

P.7 P.4-P.6을 B 번 반복하여 $(x_{T+k}^{*(1)}, \dots, x_{T+k}^{*(B)})$ 과 $(\sigma_{T+k}^{*(1)}, \dots, \sigma_{T+k}^{*(B)})$ 을 생성한다.

자료, 혹은 조건부 분산에 대한 예측에 있어서, 각 붓스트랩 표본으로부터 얻은 추정값들을 이용하는 것이 일반적이나 원자료로부터 얻은 추정값을 이용하기도 한다. 즉, P.5에서 $\hat{\theta}$ 대신에 $\hat{\theta}^*$ 을 사용하기도 한다. 이는 Miguel와 Olave (1999)의 내용과 유사하다. 본 논문에서는 이 두가지 접근 방법을 각각 완전 붓스트랩(full bootstrap: FB)방법과 조건부 붓스트랩(conditional bootstrap: CB)방법이라고 부를 것이다.

4. 실제 사료분석

이 절에서는 3절에서 소개된 방법, 특히, Pascual 등 (2006)에 소개된 GARCH모형 하에서의 자료의 예측 및 조건부 분산의 예측구간 추정에 초점을 맞추어 국내 주가자료에 적용하고자 한다. 우선 실제 자료는 2005년 1월 3일부터 2007년 12월 31일까지의 총 795일 동안 측정된 국내 2개 기업(삼성엔지니어링, 국민은행)의 주식 수익률이다 (박만식 등, 2008 참조). 먼저 GARCH모형의 적합 이전에 자기상관성을 제거하기 위해 자기회귀모형(AR(7))을 적합한 후 각 시계열의 잔차항들을 얻었다. 그런 다음 이 잔차항의 계열에서 $\{x_t; t = 1, \dots, 775\}$ 는 모형 적합을 위해, 나머지 최근 20개 시점의 자료, $\{x_t; t = 776, \dots, 795\}$ 는 예측값과의 비교를 위해 사용되었다. 따라서 이 논문에서는 $\{x_t; t = 1, \dots, 775\}$ 를 원자료로 간주하겠다. 원자료의 모형은 GARCH(1,1)으로 가정하여 붓스트랩 방법을 적용하였다. GARCH모형의 적합은 SAS(version 9.1)의 SAS/ETS("Model procedure")를 사용하였고 공분산 정상성을 위한 조건인 $\alpha + \beta < 1$ 을 제약식으로 두었다. Model procedure에서의 모수 추정방법은 Full information maximum-likelihood estimation(FIML)을 사용하였다. 붓스트랩 방법의 실행은 SAS/IML을 이용하였고 붓스트랩의 반복횟수는 $B = 999$ 로 정하였다.

우선, 자기회귀모형(AR(7))의 적합 후 얻은 모형적합에 사용되는 자료, $\{x_t; t = 1, \dots, 775\}$ 의 시도 표는 그림 1에 나타나 있다. 전체적으로 자료의 변동성이 일정하지 않고, 특히 삼성엔지니어링의 변동 폭이 국민은행에 비해 크다는 것을 알 수 있다. 그리고 두 시계열자료 간의 유사한 변동성 또한 쉽게 발견할 수 있다. 예를 들어, $200 \leq t \leq 230$ 과 $400 \leq t \leq 430$ 에서 두 시계열의 변동성이 다른 시점과 비교하여 상당히 증가함을 알 수 있다. 더욱이, 큰 변동은 큰 변동끼리 작은 변동은 작은 변동끼리 물리는 행태, 즉 변동성 집중현상이 나타나는 것으로 판단하여 GARCH모형을 고려하였다.

각 계열의 원자료, $\{x_t; t = 1, \dots, 775\}$ 과 붓스트랩에 의한 모수 추정의 결과는 표 1에 나타나 있다. 표 1에서 알 수 있듯이, 붓스트랩에 의해 추정된 모수는 ω 에 대해서는 다소 과대추정(overestimated)된 경향을, β 에 대해서는 과소평가(underestimated)된 경향을 보이고 있으며, α 에 대해서는 거의 유사한 결과를 보이고 있다. 하지만, 추정된 모수들의 중위수(Median)의 관점에서는 앞에서 언급한 평균보다

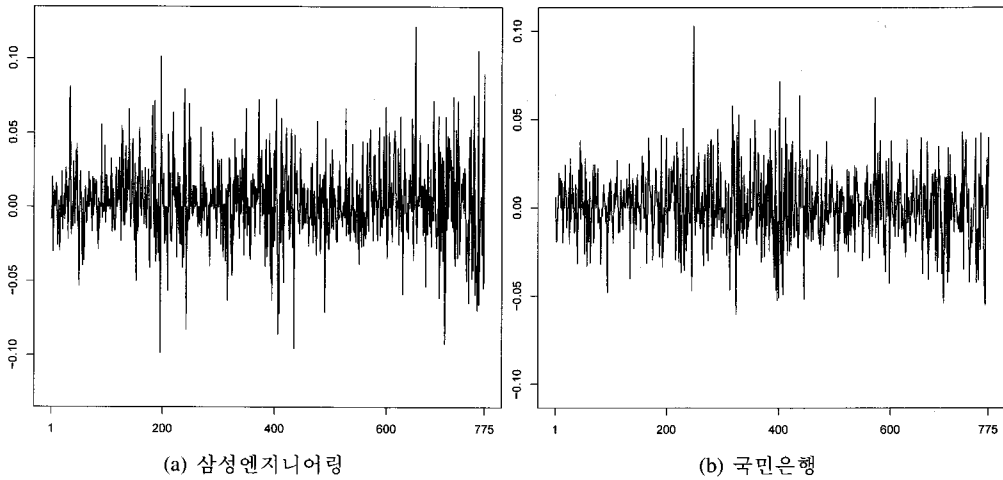


그림 1: AR(7)을 제거한 수익률의 시도표

표 1: 원자료와 붓스트랩에 의한 모수 추정($a^b \equiv a \times 10^{-b}$)

Series	Parameter	FIML		Bootstrap			
		Estimate	Std.Err.	Estimate [†]	Estimate [‡]	Std.Err. [#]	skewness
삼성엔지니어링	ω	6.90 ⁵	3.60 ⁵	8.66 ⁵	7.24 ⁵	5.92 ⁵	3.626
	α	0.107	0.035	0.109	0.107	0.036	0.326
	β	0.805	0.075	0.777	0.795	0.101	-2.668
국민은행	ω	1.10 ⁵	7.90 ⁶	1.88 ⁵	1.39 ⁵	1.90 ⁵	5.306
	α	0.062	0.024	0.064	0.062	0.024	0.554
	β	0.910	0.040	0.864	0.899	0.063	-4.398

†: $\sum_{b=1}^B \widehat{a}^{(b)}/B = \bar{a}^*$; ‡: Median $\{\widehat{a}^{(b)}\}$; #: $\sqrt{\sum_{b=1}^B (\widehat{a}^{(b)} - \bar{a}^*)^2}/B$.

는 훨씬 FIML방법의 결과와 유사함을 알 수 있다. 이는 붓스트랩 방법 하에서 얻은 $B = 999$ 개의 추정값들의 분포가 비대칭성(skewness)을 보이는 데 기인한다. 그리고 α 를 제외한 모수들에 대해 붓스트랩 방법 하에서 얻은 추정값의 표준오차가 FIML방법에서의 그것보다 큰 것으로 나타났다.

표1의 결과를 토대로 최근 20시점에서의 자료, $\{x_t; t = 776, \dots, 795\}$ 과 조건부 분산, $\{\widehat{\sigma}_t^2; t = 776, \dots, 795\}$ 에 대한 실제값이 95% 예측구간에 포함되는 지를 알아보았다. 이 때 미래 시점에 대한 $\widehat{\sigma}_t^2$ 은 다음과 같이 얻어진다.

$$\widehat{\sigma}_t^2 = \widehat{\omega} + \widehat{\alpha}x_{t-1}^2 + \widehat{\beta}\widehat{\sigma}_{t-1}^2, \quad t = 776, \dots, 795.$$

우선, 3절에서 언급한 바와 같이 각 붓스트랩 표본마다의 추정값을 이용하는 경우(FB)와 각 붓스트랩 표본으로부터 예측값들을 산출하는 데 원자료의 모수 추정값을 사용하는 경우(CB)를 같이 고려하였다. 자료값의 예측에 있어서 FIML방법 하에서 도출되는 예측구간의 추정 역시 비교대상에 포함시켰다.

자료의 예측(그림 2(a), (b))에 있어서 모든 계열에 대해서 붓스트랩방법(FB, CB) 하에서의 예측구간이 FIML방법 하에서의 구간보다 넓게 나타났고 FB방법과 CB방법의 예측구간은 거의 일치하였다. FIML방법으로부터 얻은 예측구간은 모든 시점($t = 776, \dots, 795$)에 대해서 붓스트랩 방법들로부터 얻은 구간들에 속한다. 실제값과 각 방법에 따른 예측구간을 비교해 보면, 삼성엔지니어링의 최근 20시

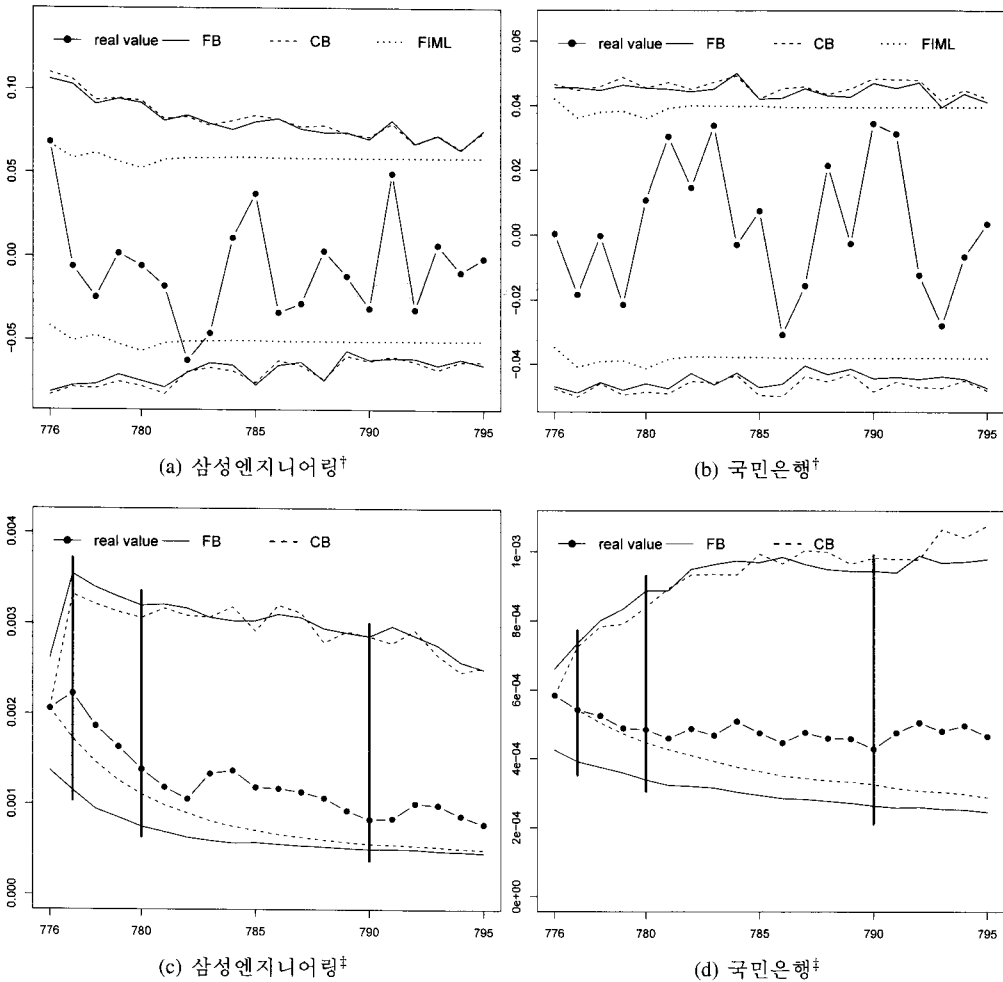


그림 2: $AR(7)$ 을 제거한 수익률의 최근 20 시점($t = 776, \dots, 795$)에서의 예측결과. †: $\{x_t\}$ 와 96% 예측구간; ‡: $\{\hat{\sigma}_t^2\}$ 와 95% 예측구간.

점의 실제값들 중 90% 정도가 FIML방법 하에서 얻게 되는 95% 예측구간에 포함된 반면, 국민은행의 경우, 모든 실제값이 모든 예측구간 안에 놓여져 있음을 알 수 있다.

조건부 분산의 예측 문제는 금융통계학의 주요한 이슈들 중 하나로 현실적인 문제에 있어서 점추정 보다는 구간추정에 더 많은 관심을 보이고 있다. 하지만, FIML방법 하에서 미래시점에 대한 조건부 분산의 예측구간을 구하기가 쉽지 않아서 하나의 대안으로서 붓스트랩방법이 소개되어졌다. 그림 2(c), (d)는 각 계열에서의 조건부 분산의 예측 결과로써, 각 붓스트랩 표본에서 얻게 되는 모수추정값들을 이용하는 FB방법이 CB방법에 비해 조금 넓은 구간을 형성하고 있고, 예측 시점이 멀어질수록 분산의 예측 범위가 증가하다가 점차 일정하게 수렴하는 경향을 보이고 있다. CB방법에 의한 최초 예측시점의 예측값은 모두 같은 값을 가지게 된다. 이는 원자료의 추정값을 사용하는 데 기인한다.

이제, 전체 20개의 미래시점들 중 특정 시점들(그림 2(c), (d)의 세로줄)에 국한시켜서 붓스트랩방법에 의한 조건부 분산의 예측값들을 살펴보자. 그림 3과 4는 특정 시점에서 각 붓스트랩방법에 의해

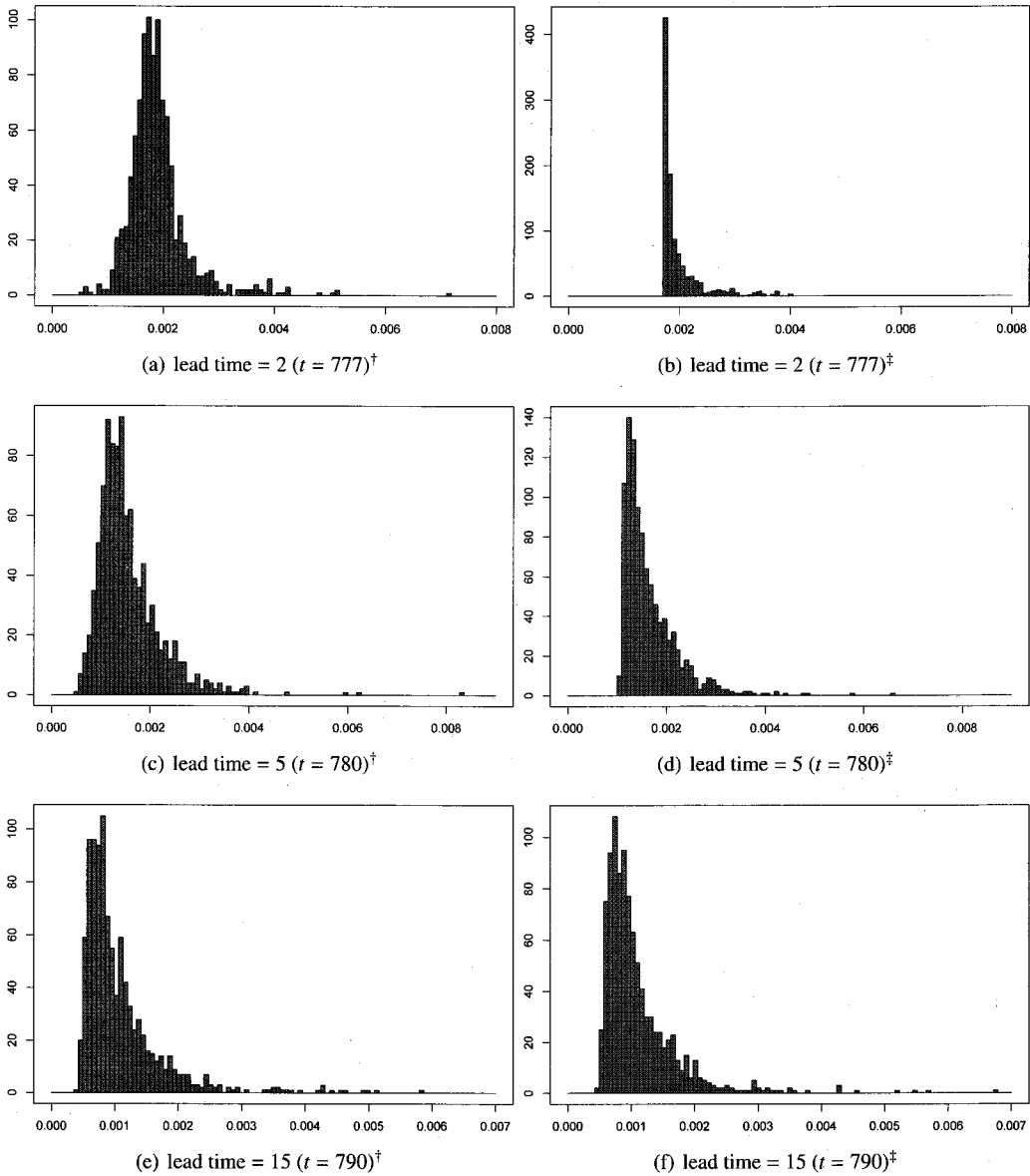


그림 3: 삼성엔지니어링의 특정한 미래 시점($t = 777, 780, 790$)에서의 붓스트랩방법에 의한 조건부 분산.
 †: FB방법, ‡: CB방법.

예측된 조건부 분산의 분포이다. $t = 777$ 시점에서의 CB방법에 의해 예측된 조건부 분산은 그 분포가 FB방법에 의한 것과 비교하여 중심보다 작은 부분이 절단(truncation)된 형태를 취하고 있다. 이는 모든 붓스트랩에 의한 예측과정에서 원자료으로부터 얻은 추정량, $\hat{\theta}$ 에 너무 많이 의존하기 때문이다. 하지만 예측시점이 멀어질수록($t = 780, 790$), $\hat{\theta}$ 의 영향력은 줄어들게 되어 CB방법이 FB방법과 거의 유사한 분포를 띄게 된다.

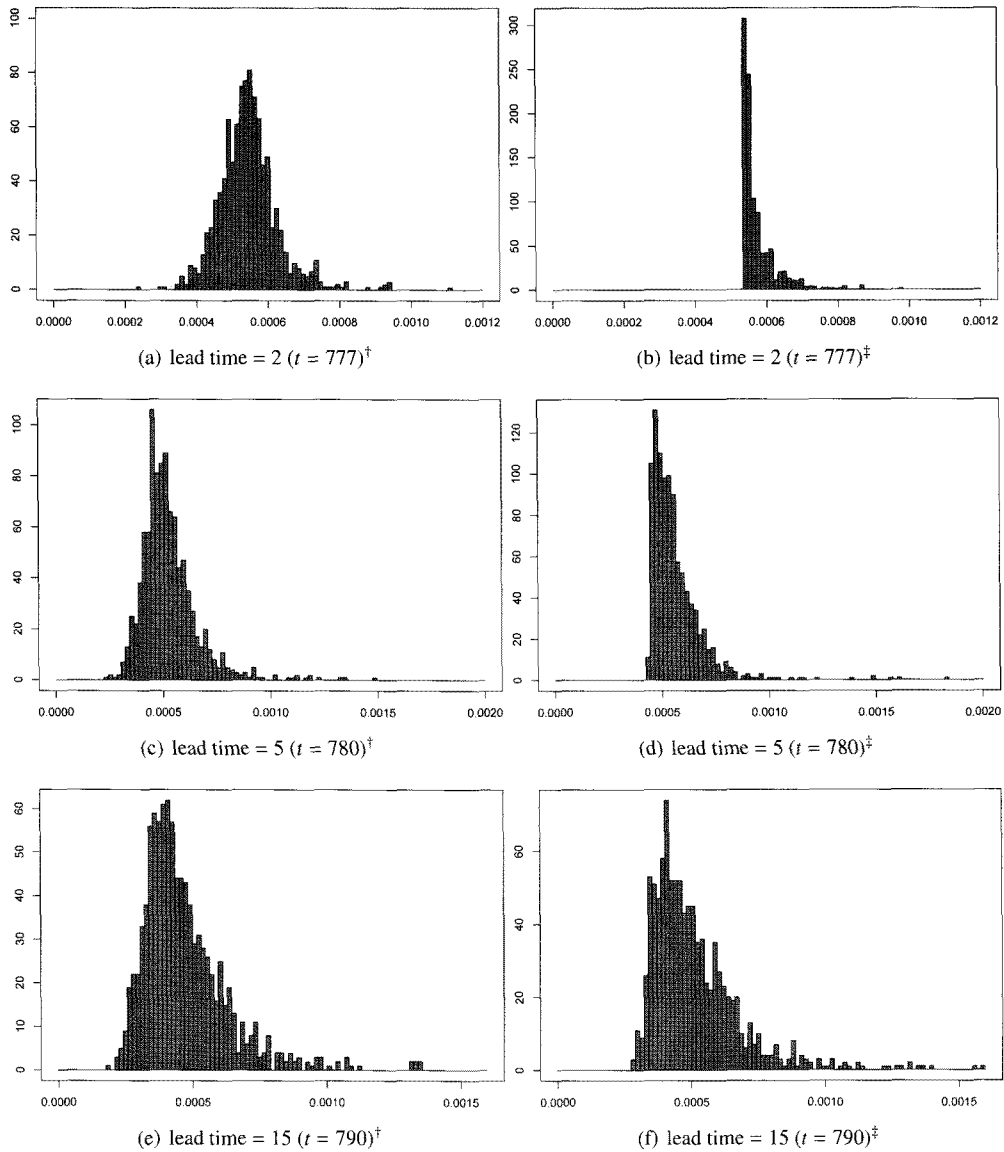


그림 4: 국민은행의 특정한 미래 시점($t = 777, 780, 790$)에서의 붓스트랩방법에 의한 조건부 분산. †: FB방법; ‡: CB방법.

5. 결론 및 토의

주가지수 수익률의 시계열은 주식시장에 내재하는 불확실성으로 인해 변동성이 시간에 따라 변하는 경우가 일반적이고, 이를 모형화하기 위한 접근 방법들 중 일반화 자기회귀 조건부 이분산(GARCH)모형이 있다. 본 논문에서는 GARCH모형 하에서 붓스트랩방법을 통한 계열과 변동성의 예측에 대해 알아보았다. 계열의 예측에 있어서, 우도함수의 기반한 예측구간과 붓스트랩에 의한 예측구간을 비교하였다. 변동성은 위험을 측정하는 지표이므로, 변동성의 측정과 미래시점의 변동성

의 예측은 위험관리 측면에서 중요성이 크다고 할 수 있다. 이를 위한 방법의 일환으로 붓스트랩방법이 Pascual 등 (2006)에 의해 제안되었고 실제 사례분석을 통해 변동성에 대한 구간 예측이 가능하다는 것을 확인하였다. 이 연구에서는 고려하지 않았으나 자기회귀(autoregressive)모형의 모수 역시 붓스트랩 과정에 포함시키는 경우를 고려할 수 있을 것으로 생각된다.

참고 문헌

- 박만식, 김나영, 김희영 (2008). Clustering Korean stock return data based on GARCH model, <한국통계학회논문집>, **15**, 925-937.
- Baillie, R. T. and Bollerslev, T. (1992). Prediction in dynamic models with time dependent conditional variances, *Journal of Econometrics*, **52**, 91-113.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity, *Journal of Econometrics*, **31**, 307-327.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica*, **50**, 987-1007.
- Miguel, J. A. and Olave, P. (1999). Bootstrapping forecast intervals in ARCH models, *Test*, **8**, 345-364.
- Pascual, L., Romo, J. and Ruiz, E. (2006). Bootstrap prediction for returns and volatilities in GARCH models, *Computational Statistics & Data Analysis*, **50**, 2293-2312.
- Rydberg, T. H. (2000). Realistic statistical modelling of financial data, *International Statistical Review*, **68**, 233-258.
- Taylor, S. J. (1986). *Modelling Financial Time Series*, John Wiley & Sons, New York.

Prediction of Conditional Variance under GARCH Model Based on Bootstrap Methods.

Hee-Young Kim^a, Man Sik Park^{1,a}

^aDept. of Biostatistics, Korea Univ.

Abstract

In terms of generalized autoregressive conditional heteroscedastic(GARCH) model, estimation of prediction interval based on likelihood is quite sensitive to distribution of error. Moreover, it is not an easy job to construct prediction interval for conditional variance. Recent studies show that the bootstrap method can be one of the alternatives for solving the problems. In this paper, we introduced the bootstrap approach proposed by Pascual *et al.* (2006). We employed it to Korean stock price data set.

Keywords: Bootstrap, GARCH, conditional variance.

¹ Corresponding author: Research professor, Department of Biostatistics & Department of Preventive Medicine, Medical Research Center for Environmental Toxico-Genomics and Proteomics, College of Medicine, Korea University, 126-1 Anam-Dong, Sungbuk-Gu, Seoul 136-705, Korea. E-mail: man.sik.park@gmail.com