Middle-East Journal of Scientific Research 14 (11): 1502-1507, 2013
ISSN 1990-9233

DOI: 10.5829/idosi.mejsr.2013.14.11.2164

Efficient Bootstrap Forecast Intervals for Return and Volatility Using the Linear Estimator of ARCH Models

[1]*Farhat Iqbal and* [2]*Sohail Chand*

[1]Department of Statistics, University of Balochistan, Quetta-Pakistan
[2]College of Statistical and Actuarial Sciences, University of the Punjab, Lahore-Pakistan

Abstract: In this paper, we construct efficient forecast intervals for autoregressive conditional heteroscedastic (ARCH) models using the bootstrap. Forecast intervals for returns and volatility are constructed using the linear estimator (LE) for ARCH model. An advantage of LE over the widely used quasi maximum likelihood estimator (QMLE) is that its computation is very easy and requires less CPU time which enables us to construct these forecast intervals in quick time. Monte Carlo results show that although both estimators provide good mean coverage, the LE can be considered favourable in terms of its mean lengths close to empirical with low standard errors. The bootstrapped prediction intervals for volatilities capture the asymmetry commonly present in real data sets.

Key words: ARCH model · Linear estimator · Quasi-Maximum likelihood estimator · Bootstrapping

INTRODUCTION

The autoregressive conditional heteroscedastic (ARCH) model was suggested by [1]. In this model, the volatility of the current return of an asset is described as a linear function of the squares of past returns. Quasi-maximum likelihood estimator (QMLE) is frequently used for the estimation of ARCH models. The asymptotic properties of the QMLE under the existence of fourth-order moment on the ARCH process were established by [2]. The QMLE does not admit a closed form expression and numerical optimization methods must be used to obtain the estimates.

The linear estimator (LE) for the parameters of ARCH model was proposed by [3]. The linear estimator has a closed form and is obtained by solving linear equations. Hence, it can be easily implemented and does not require the use of any numerical optimization methods or the choice of initial values of parameters. The linear estimator requires very little computational time for the estimation of the parameters of ARCH model. This advantage enables researchers to perform computer intensive tasks such as volatility forecasting using recursive scheme and bootstrapping volatility models in short time. [4] compared the LE with the QMLE in estimating and forecasting ARCH models and found that the LE provides estimate as accurate as the QMLE and had better predictability.

Predicting the distribution of the future returns has become an increasingly interesting area of research amongst financial practitioners and researchers. Accurate prediction of future volatilities is important for the implementation and evaluation of asset and derivative pricing [5]. Measuring the financial risk such as value-at-risk (VaR) is also very important and an accurate measure of this risk estimate is desisatle. Most of the surveys deal with predicting point forecast of returns, volatilities and VaR; see [6-8] among others for discussion on forecasting.

These studies focus on point forecasts and most importantly ignore parameter uncertainty. The parametric bootstrap prediction intervals was first discussed by [9]. [10] gave non-parametric bootstrap intervals for AR models. [11] proposed a bootstrap method for prediction intervals of future observation in ARMA models with ARCH errors without considering parameters uncertainty. [12] extended bootstrap methods to ARIMA models [13] compared nonparametric and parametric bootstrap with Baillie and Bollerslev (BB) Gaussian asymptotic prediction interval in a Monte Carlo experiment.

Corrersponding Author: Sohail Chand, College of Statistical and Actuarial Sciences, University of the Punjab, Lahore-Pakistan.

In this paper, we use bootstrap to obtain prediction intervals for returns and volatilities. The bootstrap prediction intervals are obtained using both the LE and the QMLE. These prediction intervals along with point estimate will help practitioners to evaluate the forecasting performance of their models. We investigate the difference in bootstrap prediction intervals of both estimators. It is important to mention again that LE can be estimated in quick time and thus developing bootstrapped confidence intervals. Using LE requires very small processing time as compared to the QMLE.

Monte Carlo results showed that our LE bootstrap method generates reliable prediction intervals. We found, that although both the QMLE and LE provide good mean coverage, the LE can be considered superior in terms of its mean lengths close to empirical behaviour with low standard errors. The bootstrapped prediction intervals for volatilities capture the asymmetry commonly present in real data sets.

Our study is important because many researchers avoid bootstrapping ARCH models with large sample size due to exhaustive computer time QMLE takes for estimation. The LE on the other hand estimates the same model in short time and thus various computer intensive tasks can be applied on ARCH models.

The rest of the paper is organized as follows. In Section 2 we define the linear estimator for ARCH model. In Section 3, bootstrap method for constructing forecast intervals for returns and volatilities are discussed. Results of Monte Carlo simulations are presented in Section 4. Finally, Section 5 concludes the results.

The Linear Estimator for ARCH Model: Consider the following ARCH model where one observes $\{X_t; 1\text{-}p \le t \le T\}$ satisfying

$$X_t = h_t^{1/2}(\hat{a})\varepsilon_t \quad 1 \le t \le T, \tag{2.1}$$

where $\hat{a} = [\beta_1, \beta_2, \ldots, \beta_p]'$ is the unknown parameter to be estimated with $\hat{a}_0 > 0, \beta_j \ge 0; 1 \le j \le p,$

$$h_t(\hat{a}) = \beta_0 + \beta_1 X_{t-1}^2 + \beta_2 X_{t-2}^2 + \ldots \beta_p X_{t-p}^2$$

with $\{\varepsilon_t;\ 1 \le t \le T\}$ are independently and identically distributed (IID) with mean zero and unit variance. It is assumed that $\{\varepsilon_t;\ 1 \le t \le T\}$ are independent of $\{X_t;\ 1\text{-}p \le t \le T\}$. It is also assumed that (2.1) holds, $\{X_t;\ t \ge 1\text{-}p\}$ is a stationary and ergodic process and $E(\varepsilon_t^4) < \infty$.

Let $Y_t = X_t^2$ for $p \le t \le T$, thus

$$Z_{t-1} = [1, Y_{t-1}, \ldots, Y_{t-p}]' = [1, X_{t-1}^2, \ldots, X_{t-p}^2]'$$

and $\eta_t = \varepsilon_t^2 - 1; 1 \le t \le T.$ Then squaring both sides of (2.1) and using the form $h_{t-1}(\hat{a}) = Z'_{t-1}\hat{a}$ we get

$$Y_t = Z'_{t-1}\hat{a} + h_{t-1}(\hat{a})\eta_t \quad 1 \le t \le T \tag{2.2}$$

where $E\{h_{t-1}(\hat{a})\eta_t\} = E\{h_{t-1}(\hat{a})\}E\{\eta_t\} = 0$.

[3] defined a preliminary least squares estimator $\hat{a}_{pr}$ of $\hat{a}$ as the solution of

$$\sum_{t=1}^{T} [Z'_{t-1}\{Y_t - Z'_{t-1}\hat{a}\}] = 0 \tag{2.3}$$

which yields the estimator given by

$$\hat{a}_{pr} = (Z'Z)^{-1} Z'Y$$

where Z is the $T \times (1+p)$ matrix whose t-th row equals Z'_{t-1} and Y is the vector with t-th entry Y_1; $1 \le t \le T$.

An improved estimator $\hat{\hat{a}}$ of $\hat{a}$ can be obtained as follows. Dividing (2.2) by $h_{t-1}(\hat{a})$, we get

$$\frac{Y_t}{Z'_{t-1}\hat{a}} = \left(\frac{Z_{t-1}}{Z'_{t-1}\hat{a}}\right)' \hat{a} + \eta_t$$

Now replacing $Z'_{t-1}\hat{a}$ by $Z'_{t-1}\hat{\hat{a}}$ yields

$$\frac{Y_t}{Z'_{t-1}\hat{\hat{a}}_{pr}} \approx \left(\frac{Z_{t-1}}{Z'_{t-1}\hat{\hat{a}}_{pr}}\right)' \hat{a} + \eta_t$$

Therefore, a linear estimator of $\hat{a}$ is defined as the solution of

$$\sum_{t=1}^{T} \left[\left\{ \frac{Z_{t-1}}{Z'_{t-1}\hat{\hat{a}}_{pr}} \right\} \left\{ \frac{Y_t}{Z'_{t-1}\hat{\hat{a}}_{pr}} - \left(\frac{Z_{t-1}}{Z'_{t-1}\hat{\hat{a}}_{pr}} \right)' \hat{a} \right\} \right] = 0$$

yielding the linear estimator

$$\hat{\hat{a}} = \left[\sum_{t=1}^{T} \left\{ \frac{Z_{t-1}Z'_{t-1}}{\left(Z'_{t-1}\hat{\hat{a}}_{pr}\right)^2} \right\} \right]^{-1} \left[\sum_{t=1}^{T} \left\{ \frac{Z_{t-1}Y_t}{Z'_{t-1}\hat{\hat{a}}_{pr})^2} \right\} \right]$$

Bootstrap Prediction Intervals: Consider an ARCH(p) model where one observe $\{X_t;\ 1-p \le t \le T\}$ satisfying

$$X_t = h_t^{1/2}(\hat{a})\varepsilon_t \text{ and } h_t(\hat{a}) = \beta_0 + \beta_1 X_{t-1}^2 + \beta_2 X_{t-2}^2 + \dots \beta_p X_{t-p}^2$$

Our aim is to estimate the distribution of s-steps ahead returns X_{T+s} and volatilities h_{T+s}. The bootstrap method is described in the following steps:

Fitting the ARCH Model: Fit an ARCH model to the given data set and estimate the parameters of the model $\hat{a} = [\beta_1, \beta_2, ..., \beta_p]$. We use both the QMLE and LE for the estimation of the parameter vector. Let the estimated parameter vector be $\hat{a} = [\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_p]'$ and the estimated residuals are computed as

$$\hat{\varepsilon}_t = X_t \big/ \hat{h}_t^{1/2}$$

where $\hat{h}_t = \hat{h}_t(\hat{a}) = \hat{\beta}_0 + \hat{\beta}_1 X_{t-1}^2 + \hat{\beta}_2 X_{t-2}^2 + \dots \hat{\beta}_p X_{t-p}^2$.

Bootstrapping: Use the fitted model to generate bootstrap draws of the parameter. First, we generate ε_t^* , random draws with replacement from $\hat{F}_t$, where $\hat{F}_t$ is the empirical distribution function of the centered residuals $\left(\hat{\varepsilon}_t - \sum_{t=1}^{T} \hat{\varepsilon}_t / T\right)$. Then the following replicates are generated:

$$\hat{h}_t^* = \hat{\beta}_0 + \hat{\beta}_1 X_{t-1}^{*2} + \hat{\beta}_2 X_{t-2}^{*2} + \dots \hat{\beta}_p X_{t-p}^{*2}$$

$$X_t^* = \hat{h}_t^{*1/2} \varepsilon_t^* \text{ for } t = 1, 2, \cdots, T.$$

The parameters of this generated series are estimated and the estimated parameters of this bootstrap series $\hat{a}^* = [\hat{\beta}_1^*, \hat{\beta}_2^*, \cdots, \hat{\beta}_p^*]'$ are used to obtain future values.

Future Realisations: Generate future realisations of returns and volatilities. We want to estimate the distribution of future returns X_{T+1} and future volatilities h_{T+1} for $s>0$, where s is the forecast step. In order to get these future realisations, we need $X_{T+1-i}^* = X_{T+1-i}\ (1 \le i \le p)$ and ε_{T+s}^* (random draws with replacement from $\hat{F}_t$). Using the above, the future realisations of returns are generated recursively as:

$$\hat{h}_{T+s}^* = \hat{\beta}_0{}^* + \hat{\beta}_1{}^* X_{t-1}^{*2} + \hat{\beta}_2{}^* X_{t-2}^{*2} + \dots \hat{\beta}_p^* X_{t-p}^{*2}$$

$$X_{T+s}^* = \hat{h}_{T+s}^{*1/2} \varepsilon_t^* \text{ for } s = 1, 2, \cdots$$

Prediction Intervals: Once the set of B bootstrap future values, $(X_{T+s}^{*(1)}, \dots\dots X_{T+s}^{*(B)})$ are obtained, the prediction intervals are defined as quantiles of the bootstrapped cumulative distribution function (CDF) of X_{T+s}^*. More specifically, we define the bootstrapped CDF of X_{T+s}^* by $G_X^*(l) = \Pr\{X_{T+s}^* \le l\}$ and its Monte Carlo estimate by $G_X^*(l) - \#\{X_{T+s}^{*(b)} \le l\}$, where #(.) counts the number of cases when the condition within brackets are satisfied and b = 1, 2, …, B. The for a given ϕ, a 100(1-ϕ)% prediction interval for X_{T+s}^* is given by

$$[L_{X,B}^*(X), U_{X,B}^*(X)] = \left[Q_{X,B}^*\left(\frac{\phi}{2}\right), Q_{X,B}^*\left(1 - \frac{\phi}{2}\right)\right]$$

where $Q_{X,B}^* = G_{X,B}^{*-1}$.

Similarly, we can define the bootstrap prediction intervals for volatilities. For future volatilities $\left(\hat{h}_{T+s}^{*(1)}, \dots\dots, \hat{h}_{T+s}^{*(B)}\right)$, the prediction intervals are defined as quantiles of the bootstrap CDF of $\hat{h}_{T+s}^*$. The bootstrap CDF of $\hat{h}_{T+s}^*$ is given by $G_h^*(l) = \Pr\{\hat{h}_{T+s}^* \le l\}$ and its Monte Carlo estimate by $G_h^*(l) = \#\{\hat{h}_{T+s}^{*(b)} \le l\}$. Then, a 100(1-$\phi$)% prediction interval for $\hat{h}_{T+s}^*$ is given by

$$[L_{X,B}^*(h), U_{h,B}^*(h)] = \left[Q_{h,B}^*\left(\frac{\phi}{2}\right), Q_{h,B}^*\left(1 - \frac{\phi}{2}\right)\right],$$

where $Q_{X,B}^* = G_{X,B}^{*-1}$.

Monte Carlo Simulation: Two different studies are conducted to develop bootstrap prediction intervals for returns and volatilities and to compare the results of the QMLE and LE. The model simulated is an ARCH(2) model

$$X_t = h_t^{\frac{1}{2}} \varepsilon_t,$$

$$h_t = 0.1 + 0.4 X_{t-1}^2 + 0.2 X_{t-2}^2.$$

For this study errors are generated from the standard normal and student-t distribution with 3 degrees of freedom. The motivation behind these values for parameters is not to use high persistent values. In future we aim to user different combination of parameter values and other fat-tailed distributions. The ARCH(2) model is simulated with sample size T=500 and parameters are estimated using the QMLE and LE. For this model, under particular sample size and error distribution, R=1000 future values of X_{T+S} and X_{T+S} are generated with true parameter values, where the forecast steps considered are s = 1,2 and 20. Using the bootstrap method with B=999, a 100(1-ϕ)% prediction intervals for returns denoted by (L_X^*, U_X^*) and for volatilities denoted by (L_h^*, U_h^*) are obtained.

The conditional coverage and length for returns are computed as $\overline{1}-\hat{\phi}_X^* = \#\{L_X^* \le X_{T+s}^r \le U_X^*\}/R,$ for r = 1,2,....,R. Choices of nominal coverage considered are 80%, 95% and 99%, though only results for 99% prediction intervals are presented as this interval could be of more interest in risk management. The length is defined as $LEN_X = U_X^* - L_X^*$. Similarly the conditional coverage and length for volatilities are obtained. The coverage of the left and right tail of the distribution of returns and volatilities are also obtained. The average and the standard deviation for coverage and length and the average proportion of observation lying out of the left and right quantiles are computed based on K=100 Monte Carlo replicates.

For empirical lengths, R=10,000 independent replication each of size 500 are generated. The root mean squared error (RMSE) for both estimators at each step length is calculated, where RMSE for returns is defined as

$$RMSE_X = \sqrt{\frac{1}{R}\sum_{i=1}^{R}(\hat{\phi}_{i,X}^* - \phi)^2}.$$

Table 1 reports the mean coverage and the corresponding standard errors together with the mean length with its corresponding standard errors and the mean coverage on the left and right tails and the RMSE when ARCH(2) model is generated with standard normal and student-t distribution with 3 df for predicting intervals for returns for s = 1,2 and 20 steps ahead. It can be seen that the mean coverage and their corresponding standard errors for both estimators are close to each other and provide good match to the empirical coverage with the QMLE having slightly high probability. By examining the results of mean length we found that the lengths for LE are close to empirical length and their standard errors are below than those of QMLE. These findings become more prominent in the case of student-t distribution. The mean lengths of QMLE for all step lengths are found greater

Table 1: Prediction intervals for returns of ARCH(2) model with nominal coverage of 99%

T = 500 B = 999	Mean coverage	S.D of coverage	Mean length	S.D of length	Mean coverage below/above	RMSE
			1-step ahead			
Empirical	0.9900		2.8933		0.50%/0.50%	
QMLE	0.9869	0.0071	3.0967	0.4891	0.57%/0.74%	0.0107
LE	0.9852	0.0074	2.9622	0.4284	0.63%/0.86%	0.0123
			10-steps ahead			
Empirical	0.9900		3.0550		0.50%/0.50%	
QMLE	0.9828	0.0096	3.0816	0.4992	0.98%/0.75%	0.0155
LE	0.9816	0.0080	2.9936	0.4833	1.02%/0.82%	0.0156
			20-steps ahead			
Empirical	0.9900		3.0416		0.50%/0.50%	
QMLE	0.9919	0.0051	3.0938	0.4647	0.28%/0.54%	0.0059
LE	0.9914	0.0047	3.0161	0.4280	0.28%/0.58%	0.0060
			Student-t(3) Distribution			
			1-step ahead			
Empirical	0.9900		2.7157		0.50%/0.50%	
QMLE	0.9907	0.0066	3.3364	0.9949	0.57%/0.36%	0.0078
LE	0.9889	0.0057	2.9192	0.5496	0.66%/0.46%	0.0084
			10-steps ahead			
Empirical	0.9900		2.8917		0.50%/0.50%	
QMLE	0.9909	0.0053	3.3347	0.8500	0.39%/0.52%	0.0067
LE	0.9874	0.0067	2.9554	0.6913	0.64%/0.62%	0.0071
			20-steps ahead			
Empirical	0.9900		3.1312		0.50%/0.50%	
QMLE	0.9861	0.0074	3.3538	0.9669	0.60%/0.79%	0.0116
LE	0.9836	0.0065	3.0282	0.6815	0.77%/0.87%	0.0121

Table 2: Prediction intervals for volatilities of ARCH(2) model with nominal coverage of 99%

T = 500 B = 999	Mean coverage	S.D of coverage	Mean length	S.D of length	Mean coverage below/above	RMSE
			Standard Normal Distribution			
			1-step ahead			
Empirical	0.9900		1.7390		0.50%/0.50%	
QMLE	0.9857	0.0211	2.0875	1.2881	0.48%/0.94%	0.0290
LE	0.9730	0.0448	1.6939	0.8186	1.45%/1.24%	0.0497
			10-steps ahead			
Empirical	0.9900		1.7957		0.50%/0.50%	
QMLE	0.9846	0.0276	2.0545	1.0999	0.65%/0.89%	0.0294
LE	0.9735	0.0433	1.8277	1.2007	1.57%/1.09%	0.0482
			20-steps ahead			
Empirical	0.9900		1.9711		0.50%/0.50%	
QMLE	0.9891	0.0268	2.2128	1.1523	0.74%/0.35%	0.0273
LE	0.9777	0.0441	1.8338	1.1777	1.66%/0.57%	0.0472
			Student-t(3) Distribution			
			1-step ahead			
Empirical	0.9900		1.7759		0.50%/0.50%	
QMLE	0.9887	0.0150	3.4645	4.2050	0.00%/1.13%	0.0162
LE	0.9748	0.0373	1.3863	0.8745	0.51%/2.01%	0.0422
			10-steps ahead			
Empirical	0.9900		1.8344		0.50%/0.50%	
QMLE	0.9947	0.0108	3.5826	4.1048	0.00%/0.53%	0.0108
LE	0.9835	0.0300	1.4372	1.1011	0.37%/1.28%	0.0320
			20-steps ahead			
Empirical	0.9900		2.3055		0.50%/0.50%	
QMLE	0.9894	0.0127	3.5116	4.1143	0.00%/1.06%	0.0138
LE	0.9743	0.0410	1.4704	1.4169	0.55%/2.01%	0.0457

than both the empirical lengths and LE. This shows that prediction intervals of QMLE are on average larger than the mean length of LE and this may be one of the reasons of high coverage probabilities of QMLE. The mean coverage on the left and right tails of both estimators shows similar results. The root mean squared errors of QMLE are found slightly smaller than the LE.

Next, we analyse the performance of both LE and QMLE prediction intervals for future volatilities. Using same DGP as in the previous case, we develop 99% bootstrapped prediction intervals for s = 1,2 and 20 steps ahead volatilities. The results when errors are generated form Gaussian and student-t distribution with 3 df are tabulated in Table 2. The mean coverage for QMLE is found greater than LE with low standard errors. The mean lengths of LE are close to empirical lengths where as that of QMLE are larger in size with large standard errors. Again this feature can be seen in the case of Students-t(3). The results of the average coverage on the left and right tails reveal that the shape of the volatility is asymmetric which is often observes in real data sets. The RMSEs of QMLE are found smaller than the LE and this can be due to the wider lengths of QMLE.

We conclude this section by highlighting our contributions and findings. We defined bootstrap prediction intervals for returns and volatilities for ARCH models. We showed that our method is easy to apply when the LE is used for the estimation of ARCH models. Results of our simulations indicated that the proposed bootstrap method is appropriate for predicting interval forecasts. We found that LE provides better prediction intervals than the QMLE in most of the cases.

We write our own MATLAB and Fortran code and checked the CPU time (in seconds) taken by both the LE and the QMLE for estimating an ARCH(3) model. The experiment was performed on Intel Core 2 Duo CPU running at 2 Ghz with 2 GB of random access memory (RAM). The sample size used was T = 1,000 and the experiment was repeated K = 10,000 times. The Linear Estimator took 215.55 sec whereas the QMLE took 902.43 sec for estimating the same data sets. This clearly reveals the advantage of using the LE for estimating the parameters of ARCH models as the LE takes around one-fourth of the time than the QMLE and also is not only efficient but estimates the parameters as accurately as the QMLE. This difference becomes very significant when resampling methods are used on large data sets.

CONCLUSION

Prediction intervals for returns and volatilities are developed using a simple bootstrap method. Monte Carlo results showed that although both estimators provide good mean coverage, the LE can be considered favourable in terms of its mean lengths close to the empirical with low standard errors. Furthermore, the LE takes little CPU time on computer intensive tasks as compared to the QMLE.

REFERENCES

1. Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of UK inflation, Econometrica, 50(4): 987 - 1008.
2. Weiss, A.A., 1986. Asymptotic Theory for ARCH models: Estimation and testing. Econometric Theory, 2: 107-131.
3. Bose, A. and K. Mukherjee, 2003. Estimating the ARCH parameters by solving linear equations, Journal of Time Series Analysis, 24(2): 127-136.
4. Iqbal, F. and Y.Z. Jafri, 2010. The accuracy of the linear estimator for ARCH model. Science International (Lahore), 22(1): 113-118.
5. Pascual, L., J. Romo and E. Ruiz, 2006. Bootstrap prediction for returns and volatilities in GARCH models. Computational Statistics & Data Analysis, 50(9): 2293 - 2312.
6. Baillie, R.T. and T. Bollerslev, 1992. Prediction in dynamic models with time-dependent conditional variances. Journal of Econometrics, 52(1-2): 91-113.
7. Andersen, T.G., T. Bollerslev, F.X. Diebold and P. Labys, 2001. The distribution of realized exchange rate volatility. Journal of the American Statistical Association, 96(153): 42-55.
8. Tsay, R.S., 2005. Analysis of Financial Time Series. 2nd Edition. Wiley. New York.
9. Beran, R., 1990. Calibrating prediction regions. Journal of the American Statistical Association 85(411): 715-723.
10. Thombs, L.A. and W.R. Schucany, 1990. Bootstrap prediction intervals for autoregression. Journal of the American Statistical Association, 85(410): 486 - 492.
11. Miguel, J.A. and P. Olave, 1999. Bootstrapping forecast intervals in ARCH models. Test, 8(2): 345 - 364.
12. Pascual, L., J. Romo and E. Ruiz, 2004. Bootstrap predictive inference for ARIMA processes. Journal of Time Series Analysis, 25(4): 449 - 465.
13. Reevs, J.J., 2005. Bootstrap prediction intervals for ARCH models. International Journal of Forecasting, 21(2):L 237-248.