Journal of the Korean Data & Information Science Society
2009, **20**(2), 453–458

한국데이터정보과학회지

Numerical study on Jarque-Bera normality test for innovations of ARMA-GARCH models †

Taewook Lee[1]

Department of Information Statistics, Hankuk University of Foreign Studies
Received 15 January 2009, revised 25 February 2009, accepted 3 March 2009

Abstract

In this paper, we consider Jarque-Bera (JB) normality test for the innovations of ARMA-GARCH models. In financial applications, JB test based on the residuals are routinely used for the normality of ARMA-GARCH innovations without a justification. However, the validity of JB test should be justified in advance of the actual practice (Lee et al., 2009). Through the simulation study, it is found that the validity of JB test depends on the shape of test statistic. Specifically, when the constant term is involved in ARMA model, a certain type of residual based JB test produces severe size distortions.

Keywords: ARMA-GARCH model, Jarque-Bera test, test for normality.

1. Introduction

The autoregressive moving average (ARMA) models for the conditional mean with the generalized autoregressive conditional heteroscedastic (GARCH) models for the conditional variance have gained popularity in the analysis of financial time series data. In actual applications, it is common to put the normal assumption on the innovation random variables of ARMA-GARCH models (Park and Lee, 2007). However, the normality assumption on innovations has been frequently violated in the real data analysis. Therefore, the validity of the normality assumption should be examined before modeling financial time series data with ARMA-GARCH models. For this reason, the normality test for the ARMA-GARCH innovations has been paid much attention and several test procedures are intensively studied in the past decades.

Among the existing normality tests, we focus on the Jarque-Bera (JB) test since it has been well known to have merits of being simple and producing good powers compared to others such as the Kolmogorov-Smirnov and Bickel-Rosenblatt (Bickel and Rosenblatt, 1973) tests. In the literature, two most popular types of JB test are defined as follow:

$$S_T = T(\hat{\tau}^2/6 + (\hat{\kappa}-3)^2/24) \ and \ \tilde{S_T} = T(\tilde{\tau}^2/6 + (\tilde{\kappa}-3)^2/24), \qquad (1.1)$$

† This work was supportted by the Korea Research Foundation Grant funded by the Korean Government (MOEHRD, Basic Research Promotion Fund) (KRF-2008-331-C00059).
[1] Full time Lecturer, Department of Information Statistics, Hankuk University of Foreign Studies, Gyeonggi-do 449-791, Korea. E-mail: twlee@hufs.ac.kr

where

$$\hat{\tau} = \frac{1}{T}\sum_{t=1}^{T} {\epsilon_t}^3, \hat{\kappa} = \frac{1}{T}\sum_{t=1}^{T} {\epsilon_t}^4 \tag{1.2}$$

and

$$\tilde{\tau} = \frac{\frac{1}{T}\sum_{t=1}^{T}(\epsilon_t - \bar{\epsilon})^3}{(\tilde{\sigma}^2)^{3/2}}, \tilde{\kappa} = \frac{\frac{1}{T}\sum_{t=1}^{T}(\epsilon_t - \bar{\epsilon})^4}{(\tilde{\sigma}^2)^2}, \tilde{\sigma}^2 = \frac{1}{T}\sum_{t=1}^{T}(\epsilon_t - \bar{\epsilon})^2, \bar{\epsilon} = \frac{1}{T}\sum_{t=1}^{T}\epsilon_t. \tag{1.3}$$

Here, S_T can be obtained in the construction of $\tilde{S_T}$ by the replacement of $\bar{\epsilon}$ and $\tilde{\sigma}^2$ with 0 and 1 , respectively. Such a replacement may be acceptable if

$$\bar{\epsilon} \ \rightarrow \ 0 \text{ and } \tilde{\sigma}^2 \ \rightarrow \ 1 \tag{1.4}$$

in probability. Under the assumption that ϵ_t 's are either IID Gaussian random variables (Jarque and Bera, 1980; Bera and Jarque, 1981) or weakly dependent data (cf. Bai and Ng, 2005) with $E(\epsilon_t) = 0$ and $E(\epsilon_t^2) = 1$, the asymptotic results in (1.4) are satisfied and consequently, it can be shown that both S_T and $\tilde{S_T}$ are asymptotically distributed as χ^2 distribution with 2 degrees of freedom (χ^2_2). For concerning the JB normality test for the innovations of heteroskedastic model, we refer to Chen and Kuan (2003), Fiorentini, Sentana and Calzolari (2004), Kulperger and Yu (2005), Lee and Ha (2007) and Lee et al. (2009) and the papers cited in those articles. Note that Chen and Kuan (2003) can be obtained at http://www.econ.sinica.edu.tw/upload/file/03-a003.2008090211040515.pdf.

In financial applications of testing the normality of innovations in ARMA-GARCH models, JB test statistics S_T and $\tilde{S_T}$ based on the residuals are routinely used, since the asymptotic distribution of the test statistics under the null hypothesis is expected to achieve χ^2_2 distribution. However, the limiting distribution of S_T for innovations of some volatility models is not always χ^2_2 distribution. Therefore, the validity of JB test should be justified in advance of the actual practice. For instance, Chen and Kuan (2003) considers S_T based on the residuals for the innovations of AR-ARCH models and showed that S_T is not valid and produces severe size distortions. In order to overcome this defect, they proposed the modified JB test. Recently, some authors reported that unlike $\tilde{S_T}$, residual based test statistic S_T for the normality of GARCH innovations suffers from severe size distortions (Kulperger and Yu, 2005; Lee et al., 2009).

In this paper, our goal is to investigate the validity of two popular types of JB test statistics S_T and $\tilde{S_T}$ based on the residuals for the normality of ARMA-GARCH innovations. This is not a trivial extension, since the residual based tests behave differently, depending on the structure of the time series models (Chen and Kuan, 2003; Lee et al., 2009). Through simulation studies, it is shown that the size of JB test depends on the shape of test statistic. Specifically, when the constant term is involved in ARMA model, residual based test statistic S^r_T (cf. (2.1) below) does produce size distortions, which is not true for $\tilde{S}^r_T$ (cf. (2.1) below).

2. JB test for ARMA-GARCH innovations

Consider the ARMA(m,s)-GARCH(p,q) models:

$$r_t = \phi_0 + \sum_{i=1}^{m} \phi_i r_{t-i} + a_t + \sum_{j=1}^{s} \theta_j a_{t-j} \ , \ a_t = \epsilon_t \sqrt{h_t},$$
$$h_t = \omega + \sum_{k=1}^{q} \alpha_k a_{t-k}^2 + \sum_{l=1}^{p} \beta_l h_{t-l}.$$

where $\phi_i, \theta_j \in R, 1 \le i \le m, 1 \le j \le s, \omega > 0, \alpha_k \ge 0, 1 \le k \le q, \beta_l \ge 0, 1 \le l \le p$ and the innovations ϵ_t are IID random variables with $E(\epsilon_t) = 0$ and $E(\epsilon_t^2) = 1$.

In this section. we consider the problem of testing the following hypotheses:
H_0 : ϵ_t 's are normally distributed. vs.
H_1 : Not H_0 .
To construct JB test statistics, we employ the quasi maximum likelihood estimators $\hat{\theta}_T = (\hat{\phi}, \hat{\theta}, \hat{\omega}, \hat{\alpha}, \hat{\beta})'$ (Francq and Zakoian, 2004) and obtain the residuals as follows:

$$\tilde{\epsilon_t} = \frac{r_t - \tilde{\mu_t}}{\sqrt{\tilde{h}_t}}, \quad t = 1, \dots, T,$$

where $\tilde{\mu_t}$'s and $\tilde{h_t}$'s are defined recursively by using

$$\tilde{\mu_t} = \hat{\phi_0} + \sum_{i=1}^{m} \hat{\phi_i} r_{t-i} + \sum_{j=1}^{s} \hat{\theta_j} \tilde{a}_{t-j}, \tilde{h}_t = \hat{\omega} + \sum_{k=1}^{q} \hat{\alpha_k} \tilde{a}_{t-k}^2 + \sum_{l=1}^{p} \hat{\beta_l} \tilde{h}_{t-l}, \tilde{a}_t = r_t - \tilde{\mu_t}$$

and the initial random variables for $r_0, \cdots, r_{1-(q-s)-m}$, $\tilde{a}_0, \cdots, \tilde{a}_{1-\max(s,q)}$, $\tilde{h}_0, \cdots, \tilde{h}_{1-p}$ are chosen to be fixed (Francq and Zakoian, 2004). Using those residuals and taking the form of JB test statistics S_T and $\tilde{S_T}$ in (1.1), we define

$$S_T^r = T(\hat{\tau}_r^2/6 + (\hat{\kappa}_r - 3)^2/24) \text{ and } \tilde{S_T^r} = T(\tilde{\tau}_r^2/6 + (\tilde{\kappa}_r - 3)^2/24), \tag{2.1}$$

where $\hat{\tau}_r$, $\hat{\kappa}_r$, $\tilde{\tau}_r$, and $\tilde{\kappa}_r$ are obtained by replacing true innovations ϵ_t in (1.2) and (1.3) with residuals $\tilde{\epsilon_t}$. This is simply because ϵ_t is not observable as in many other times series analysis. Under H_0 , it is expected that the limiting distributions of S_T^r and $\tilde{S_T^r}$ are χ_2^2 distribution, which is identical to those of S_T and $\tilde{S_T}$ based on true innovations. Then, we reject H_0 if $S_T > C_\alpha$ or $\tilde{S_T} > C_\alpha$, where the critical value C_α is the $(1-\alpha)$ -quantile point of χ_2^2 distribution.

3. Simulation study

In this section, we evaluate the validity and performance of S_T^r and $\tilde{S_T^r}$ through a simulation study. The empirical sizes and powers are calculated at the nominal level 0.05 in both cases.

We consider ARMA(1,1)-GARCH(1,1) models as follows:

$$r_t = \phi_0 + \phi_1 r_{t-1} + a_t + \theta a_{t-1}, a_t = \epsilon_t \sqrt{h_t}$$
$$h_t = \omega + \alpha a_{t-1}^2 + \beta h_{t-1}$$

Table 3.1 Empirical sizes of $\tilde{S_T^r}$ and S_T^r with $\phi_0 = 0.0$

		$\phi_1 = 0.2$				$\phi_1 = -0.2$			
		$\theta = 0.2$		$\theta = -0.2$		$\theta = 0.2$		$\theta = -0.2$	
		$\omega = 0.1$	$\omega = 0.1$	$\omega = 0.1$	$\omega = 0.1$	$\omega = 0.1$	$\omega = 0.1$	$\omega = 0.1$	$\omega = 0.1$
		$\alpha = 0.3$	$\alpha = 0.1$	$\alpha = 0.3$	$\alpha = 0.1$	$\alpha = 0.3$	$\alpha = 0.1$	$\alpha = 0.3$	$\alpha = 0.1$
	n	$\beta = 0.3$	$\beta = 0.8$	$\beta = 0.3$	$\beta = 0.8$	$\beta = 0.3$	$\beta = 0.8$	$\beta = 0.3$	$\beta = 0.8$
	500	0.041	0.044	0.049	0.044	0.036	0.042	0.040	0.047
$\tilde{S_T^r}$	1000	0.040	0.043	0.045	0.046	0.041	0.051	0.047	0.051
	2000	0.044	0.056	0.039	0.045	0.052	0.048	0.054	0.050
	500	0.162	0.161	0.159	0.162	0.165	0.164	0.162	0.179
S_T^r	1000	0.152	0.170	0.158	0.172	0.173	0.177	0.165	0.171
	2000	0.144	0.172	0.172	0.181	0.191	0.173	0.175	0.190

Table 3.2 Empirical sizes of $\tilde{S_T^r}$ and S_T^r with $\phi_0 = 0.2$

		$\phi_1 = 0.2$				$\phi_1 = -0.2$			
		$\theta = 0.2$		$\theta = -0.2$		$\theta = 0.2$		$\theta = -0.2$	
		$\omega = 0.1$	$\omega = 0.1$	$\omega = 0.1$	$\omega = 0.1$	$\omega = 0.1$	$\omega = 0.1$	$\omega = 0.1$	$\omega = 0.1$
		$\alpha = 0.3$	$\alpha = 0.1$	$\alpha = 0.3$	$\alpha = 0.1$	$\alpha = 0.3$	$\alpha = 0.1$	$\alpha = 0.3$	$\alpha = 0.1$
	n	$\beta = 0.3$	$\beta = 0.8$	$\beta = 0.3$	$\beta = 0.8$	$\beta = 0.3$	$\beta = 0.8$	$\beta = 0.3$	$\beta = 0.8$
	500	0.044	0.038	0.041	0.064	0.037	0.047	0.057	0.048
$\tilde{S_T^r}$	1000	0.045	0.054	0.045	0.058	0.052	0.046	0.058	0.054
	2000	0.045	0.049	0.064	0.047	0.039	0.059	0.052	0.047
	500	0.056	0.039	0.041	0.068	0.046	0.052	0.061	0.046
S_T^r	1000	0.055	0.055	0.056	0.064	0.060	0.055	0.068	0.052
	2000	0.057	0.052	0.071	0.054	0.046	0.062	0.056	0.049

For obtaining empirical sizes and powers, sets of 500, 1000 and 2000 observations are generated from ARMA(1,1)-GARCH(1,1) models with $\phi_0 = 0.0, 0.2$, $\phi_1 = 0.2, -0.2$, $\theta = 0.2, -0.2$ $(\omega, \alpha, \beta) = (0.1, 0.3, 0.3)$, $(0.1, 0.1, 0.8)$. In this simulation, 1000 initial observations are discarded to remove initialization effects. In order to observe the power, we consider two alternative hypothesis under which the error distribution is assumed to be either more skewed or heavy-tailed than the normal distribution, viz.,

(a) The t -distribution with 10 degrees of freedom;

(b) The skewed t -distribution with 10 degrees of freedom and shape parameter 1;

The skewness of the distribution increases as the shape parameter increases (Gupta, 2003). Further, the mean and variance are set to be 0 and 1, respectively for all the cases.

The figures in Tables 1-6 indicate the proportion of the number of rejections of the null hypothesis H_0 out of 1000 repetitions. From Table 1, we can observe that

${S_T}^r$ produces severe size distortion when $\phi_0 = 0.0$: for example, the empirical sizes of ${S_T}^r$ are mostly between 0.15 and 0.20 much larger than the nominal level 0.05. This can be considered the strong evidence against the validity of ${S_T}^r$. However, it can be seen that the empirical sizes of $\tilde{S_T}^r$ still remain very close to 0.05. On the contrary, the figures in Table

Table 3.3 Empirical powers of $\tilde{S_T^r}$ and S_T^r for (a) with $\phi_0 = 0.0$

		$\phi_1 = 0.2$				$\phi_1 = -0.2$			
		$\theta = 0.2$		$\theta = -0.2$		$\theta = 0.2$		$\theta = -0.2$	
		$\omega = 0.1$	$\omega = 0.1$	$\omega = 0.1$	$\omega = 0.1$	$\omega = 0.1$	$\omega = 0.1$	$\omega = 0.1$	$\omega = 0.1$
		$\alpha = 0.3$	$\alpha = 0.1$	$\alpha = 0.3$	$\alpha = 0.1$	$\alpha = 0.3$	$\alpha = 0.1$	$\alpha = 0.3$	$\alpha = 0.1$
	n	$\beta = 0.3$	$\beta = 0.8$	$\beta = 0.3$	$\beta = 0.8$	$\beta = 0.3$	$\beta = 0.8$	$\beta = 0.3$	$\beta = 0.8$
	500	0.709	0.716	0.715	0.689	0.716	0.688	0.685	0.700
$\tilde{S_T^r}$	1000	0.937	0.925	0.931	0.927	0.919	0.921	0.925	0.929
	2000	0.995	0.996	0.997	0.995	0.998	0.998	0.998	0.996
	500	0.773	0.733	0.769	0.732	0.762	0.728	0.737	0.728
S_T^r	1000	0.940	0.931	0.945	0.927	0.934	0.938	0.938	0.941
	2000	0.996	0.997	0.998	0.996	0.999	0.998	0.998	0.999

Table 3.4 Empirical powers of $\tilde{S_T^r}$ and S_T^r for (a) with $\phi_0 = 0.2$

		$\phi_1 = 0.2$				$\phi_1 = -0.2$			
		$\theta = 0.2$		$\theta = -0.2$		$\theta = 0.2$		$\theta = -0.2$	
		$\omega = 0.1$	$\omega = 0.1$	$\omega = 0.1$	$\omega = 0.1$	$\omega = 0.1$	$\omega = 0.1$	$\omega = 0.1$	$\omega = 0.1$
		$\alpha = 0.3$	$\alpha = 0.1$	$\alpha = 0.3$	$\alpha = 0.1$	$\alpha = 0.3$	$\alpha = 0.1$	$\alpha = 0.3$	$\alpha = 0.1$
	n	$\beta = 0.3$	$\beta = 0.8$	$\beta = 0.3$	$\beta = 0.8$	$\beta = 0.3$	$\beta = 0.8$	$\beta = 0.3$	$\beta = 0.8$
	500	0.701	0.709	0.689	0.683	0.710	0.719	0.704	0.704
$\tilde{S_T^r}$	1000	0.922	0.933	0.928	0.936	0.927	0.941	0.923	0.918
	2000	0.994	0.997	0.998	0.995	0.999	0.997	0.998	0.998
	500	0.688	0.658	0.688	0.667	0.703	0.687	0.700	0.670
S_T^r	1000	0.916	0.925	0.923	0.924	0.926	0.927	0.917	0.903
	2000	0.994	0.996	0.998	0.995	0.998	0.997	0.998	0.996

Table 3.5 Empirical powers of $\tilde{S_T^r}$ and S_T^r for (b) when $\phi_0 = 0.0$

		$\phi_1 = 0.2$				$\phi_1 = -0.2$			
		$\theta = 0.2$		$\theta = -0.2$		$\theta = 0.2$		$\theta = -0.2$	
		$\omega = 0.1$	$\omega = 0.1$	$\omega = 0.1$	$\omega = 0.1$	$\omega = 0.1$	$\omega = 0.1$	$\omega = 0.1$	$\omega = 0.1$
		$\alpha = 0.3$	$\alpha = 0.1$	$\alpha = 0.3$	$\alpha = 0.1$	$\alpha = 0.3$	$\alpha = 0.1$	$\alpha = 0.3$	$\alpha = 0.1$
	n	$\beta = 0.3$	$\beta = 0.8$	$\beta = 0.3$	$\beta = 0.8$	$\beta = 0.3$	$\beta = 0.8$	$\beta = 0.3$	$\beta = 0.8$
	500	0.907	0.899	0.916	0.902	0.905	0.904	0.929	0.901
$\tilde{S_T}$	1000	0.993	0.997	0.993	0.996	0.998	0.998	0.999	0.993
	2000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	500	0.897	0.870	0.892	0.874	0.877	0.879	0.895	0.870
S_T	1000	0.985	0.991	0.983	0.986	0.989	0.985	0.994	0.987
	2000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

2 indicate that the abovementioned phenomenon disappears when $\phi_0 = 0.2$. In this case, both ${S_T}^r$ and $\tilde{S_T}^r$ can be safely applicable. Tables 3-6 show that both $\tilde{S_T}^r$ and ${S_T}^r$ produce good powers, and the power in all cases increases to 1 as n increases. Furthermore, it can be seen that the empirical sized and the powers are not affected by the parameter values for all the cases. From our findings, it is strongly suggested that the usage of $\tilde{S_T}^r$ should be preferred to test the normality for ARMA-GARCH innovations when the constant term is involved in ARMA model. Since our conclusion is based only on the simulation study, it is required to provide a theoretical justification. We leave this as a future task.

Table 3.6 Empirical powers of $\tilde{S}_T^r$ and S_T^r for (b) when $\phi_0 = 0.2$

		$\phi_1 = 0.2$				$\phi_1 = -0.2$			
		$\theta = 0.2$		$\theta = -0.2$		$\theta = 0.2$		$\theta = -0.2$	
		$\omega = 0.1$	$\omega = 0.1$	$\omega = 0.1$	$\omega = 0.1$	$\omega = 0.1$	$\omega = 0.1$	$\omega = 0.1$	$\omega = 0.1$
		$\alpha = 0.3$	$\alpha = 0.1$	$\alpha = 0.3$	$\alpha = 0.1$	$\alpha = 0.3$	$\alpha = 0.1$	$\alpha = 0.3$	$\alpha = 0.1$
	n	$\beta = 0.3$	$\beta = 0.8$	$\beta = 0.3$	$\beta = 0.8$	$\beta = 0.3$	$\beta = 0.8$	$\beta = 0.3$	$\beta = 0.8$
	500	0.910	0.913	0.897	0.916	0.908	0.895	0.921	0.903
$\tilde{S}_T$	1000	0.993	0.992	0.993	0.991	0.998	0.996	0.995	0.992
	2000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	500	0.906	0.904	0.896	0.908	0.912	0.896	0.924	0.892
S_T	1000	0.993	0.992	0.992	0.991	0.997	0.993	0.994	0.989
	2000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

References

Bai, J. and Ng, S. (2005). Tests for skewness, kurtosis, and normality for time series data. *Journal of Business and Economic Statistics*, **23**, 49-60.

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

Bickel, P. and Rosenblatt, M. (1973). On some global measures of the deviations of density function estimates. *Annals of Statistics*, **1**, 1071-1095.

Fiorentini, G., Sentana, E. and Calzolari, G. (2004). On the validity of the Jarque-Bera normality test in conditionally heteroskedastic dynamic regression models. *Economics Letters*, **83**, 307-312.

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

Gupta, A. K. (2003). Multivariate skew t -distribution. *Statistics*, **37**, 359-363.

Jarque, C. and Bera, A. (1980). Efficient tests for normality, heteroskedasticity, and serial independence of regression residuals. *Econometric Letters*, **12**, 255-259.

Kulperger, R. and Yu, H. (2005). High moment partial sum processes of residuals in GARCH models and their applications. *Annals of Statistics*, **33**, 2395-2422.

Lee, S., Park, S. and Lee, T. (2009). A note on the Jarque-Bera normality test for GARCH innovations. *Submitted for publications*.

Lee, T. and Ha, J. (2007). Testing the domestic financial data for the normality of the innovation based on the GARCH(1,1) model. *Journal of the Korean Data & Information Science Society*, **18**, 809-815.

Park, S. and Lee, S. (2007). Modelling KOSPI200 data based on GARCH(1,1) parameter change test. *Journal of Korean Data & Information Science Society*, **18**, 11-16.