Research Journal of Recent Sciences ISSN 2277-2502
Vol. **2(5),** 39-43, May **(2013)** *Res.J.Recent Sci.*

Testing Goodness-of-Fit in Autoregressive Fractionally Integrated Moving-Average Models with Conditional Hetroscedastic Errors of Unknown form

Ali Amjad [1], Salahuddin[2] and Alamgir[2]
[1]Department of Statistics, Islamia College University Peshawar, PAKISTAN
[2]Department of Statistics, University of Peshawar, Peshawar, PAKISTAN

Available online at: www.isca.in
Received 8th December 2012, revised 18th January 2013, accepted 11th February 2013

Abstract

This paper considers testing goodness-of-fit in Autoregressive fractionally integrated moving-average models with conditional hetroscedasticity. We extend the applicability of Hong's and power transformed Hong's test statistics as goodness-of-fit tests in ARFIMA-GARCH models, where the structural form of GARCH model is unknown. Simulation study is performed to assess the size and power performance of both tests.

Key Words: Conditional hetroscedasticity, ARFIMA, GARCH, Goodness-of-Fit Tests.

Introduction

It is a nontrivial task to find an appropriate or a parsimonious model in regression and time series data analysis. Residuals analysis is commonly used as model diagnostics in time series model building. The adequacy of the fitted time series model is commonly tested by checking the assumption of white noise residuals. If the appropriate model has been chosen, there will be zero autocorrelation in the residuals series. Let e_t be the series of the residuals from the fitted model, then in hypothesis testing settings we can state our null and alternative hypothesis as

$H_0 : \rho_e(j) = 0$ for all $j \neq 0$ versus $H_1 : \rho_e(j) \neq 0$ for some $j \neq 0$.

In frequency domain approach the above hypothesis can be stated as

$H_0 : f_e(v) = 1/2\pi$, versus $H_1 : f_e(v) \neq 1/2\pi$ for some $v \in (-\pi, \pi)$,

where $f_e(\theta) = (2\pi)^{-1} \sum_{k \in z} \rho_e(k) e^{-ik\theta}$ is the normalized spectral density function of e_t. Rejecting the above null hypothesis implies the inadequacy of the fitted model. Several tests have been developed to test the hypothesis of zero autocorrelation. Box and Pierce[1] have developed a portmanteau test to test the adequacy of the fitted time series model. The test statistic is given as:

$$Q_n = n \sum_{j=1}^{m} \hat{\rho}_e^2(h) \quad (1)$$

where $\hat{\rho}_e(h)$ is the autocorrelation of e_t at lag h and m is assumed to be fixed. They showed that for large n (sample size), the statistic Q_n has chi-square distribution with m degrees of freedom assuming that e_t series is independently and identically distributed. If e_t are the residuals from a fitted time series model, then Q_n is distributed as χ^2 with $m - p$ degrees of freedom, where p is the number of parameters in the model. Davis et al.[2] showed that the distribution of Q_n can deviate from chi-square and the true significance level is likely to be lower than the predicted significance level. A modified version of Box and pierce[1] test statistic was proposed by Ljung and Box[3] , which has the following form:

$$Q_n = n(n+2) \sum_{j=1}^{m} \hat{\rho}_e^2(h) / n - j \quad (2)$$

They preformed a comparative study of their test with the test of Box and Pierce[1] and showed that their test has substantially improved approximation to chi-square distribution. For various choices of m, Ljung[4] examined the properties of Box and Pierce[1] test statistic. They suggested a modified version of Box and Pierce test statistic that allowed the use of various values of m. Their simulation studies showed that the modified test is more powerful under various innovations distributions. Hong[5] introduced three classes of consistent one sided tests for testing serial correlation of the residuals of the linear dynamic model that include both lagged dependent and independent variables. Under the null hypothesis of zero autocorrelation, they showed that the standardized form of all these test statistics is asymptotically $N(0,1)$. To improve asymptotic normality of Hong's tests, Chen and Deo[6] introduced power transformed Hong's test. They examined the performance of Hong's and power transformed Hong's test statistics as goodness-of-fit tests for different time series models with identically independent errors.

In the current study, we consider model diagnostic checking of *ARFIMA* models when its innovations are conditionally hetroscedastic of unknown form.

The Model

Long memory processes have been widely used in the analysis of time series data. Nile river data is an outstanding example which exhibits long memory behaviour[7]. Other examples are the Ethernet traffic time series studied by Leland et.al.[8] and foreign exchange rate returns studied by Goodhart and Hare[9]. The common feature of these time series is that the decay of the autocorrelation function is like a power function rather than exponential as in the case of short memory time series. The spectral density of such processes behaves just like a power function and diverges as the frequency goes to zero.

Autoregressive fractionally integrated moving average process (*ARFIMA*(p, d, q)) is a well known class of long memory time series. These models take into account the hyperbolic decay of autocorrelation function. *ARFIMA*(p, d, q) were independently introduced by Granger and Joyeux[10] and Hosking[11]. This model is a generalization of the ARIMA(p, d, q) model, where d is taken to be an integer.

It is defined as

$$\phi(B)X_t = \theta(B)(1-B)^{-d} e_t, \tag{3}$$

where $\phi(B) = \sum_{i=0}^{p} \phi_i B^i$ and $\theta(B) = \sum_{j=0}^{q} \theta_j B^j$, $\theta_0 = 1$, $\phi_0 = 1$ (and B is the backward shift operator), are the autoregressive and moving-average operators respectively; $\phi(B)$ and $\theta(B)$ have no common roots, $(1-B)^{-d}$ is fractionally differencing operator defined by the binomial expansion

$$(1-B)^{-d} = \sum \frac{\Gamma(j+d)}{\Gamma(j+1)} B^j, \; j = 0,1,2,..............., \tag{4}$$

for $d < 0.5$, $d \neq 0, -1, -2, ...$ and e_t is a white noise sequence with finite variance. If $d > 0$, the series exhibit long memory. *ARFIMA* models have proven useful tools in the analysis of long range dependence processes. Autoregressive fractionally integrated moving average (*ARFIMA*) models with *GARCH* errors have been widely used in time series data analysis. Baillie et al.[12] used *ARFIMA-GARCH* models to analyze the inflation of ten different countries. To model daily data on the Swiss 1-month Euromarket interest rate during the period 1986–1989, Hauser and Kunst[13] used fractionally integrated models with ARCH errors. Other applications of fractionally integrated models with conditionally hetroscedastic errors can be found in Hauser and Kunst[13], Lien and Tse[14], Eleck and Markus[15] and Koopman et al[16]. A two stage model building strategy is generally used to fit an *ARFIMA-GARCH* model. In the first step an *ARFIMA* model is fitted to the given series and then a *GARCH* model to the residuals of the *ARFIMA* model. So, it is important to select a correct *ARFIMA* model in the first stage. The misspecification of *ARFIMA* model in the first stage will lead to misspecification of the *GARCH* model in the second stage[17]. The tests developed by Chen and Deo[18], Delgado et al.[19], Delgado and Velasco[20] and Hidalgo and Kreiss[21] all work for long memory time series models. However, they assumed Gaussian or linear processes with conditionally homoscedastic noise processes. Ling and Li[22] and Li and Li[23] have studied BP type tests for model diagnostics of *ARFIMA-GARCH* models but assuming that the parametric form of *GARCH* model is known. In the present work, we considered model diagnosis of *ARFIMA* models with *GARCH* errors of unknown form. We investigate the performance of Hong;s statistic as a goodness of fit test for *ARFIMA-GARCH* models through simulation study. We also examine the performance of power transformed Hong's statistic of Chen and Deo[6] in the above settings.

The Test Statistics

In a seminal paper Hong[5] introduced several test statistics that are generalization of the Box and pierce[1] test statistic. These tests are based on the distance between the kernel based spectral density estimator and the spectral density of the noise under the null hypothesis. The standardized form of the Hong's test statistics with quadratic distance is given by

$$H_n = \left(n\sum_{j=1}^{n-1} k^2(j/p_n)\hat{\rho}_e^2(j) - C_n(k) \right) / (2D_n(k)^{1/2}) \tag{5}$$

where $C_n(k) = \sum_{j=1}^{n-1} (1-j/n)k^2(j/p_n)$,

$D_n(k) = \sum_{j=1}^{n-2} (1-j/n)(1-(j+1)/n)k^4(j/p_n)$, $k(.)$ is the kernel function which is non-negative and symmetric and p_n is the bandwidth that depends on the sample size. Under the assumption of *i.i.d* errors of the model, when $p_n \to \infty$ and $p_n = o(n)$, Hong[5] showed that the asymptotic null distribution of H_n is standard normal. Hong and Lee[24] extended the above result relaxing the assumption of *i.i.d* errors and established the results assuming the conditional heteroscedastic errors of unknown form.

Simulation results of Chen and Deo[6] found that for small samples the distribution of Hong's test is right skewed, which results to the size distortion of the test. To deal with this problem, Chen and Deo[6] introduced a power transformed version of Hong's test statistics. The idea behind this transformation is to induce normality.

They showed that the appropriate power β to be used such that H_n^{β} become approximately normal is given by

$$\beta = 1 - \frac{2}{3}\frac{(\sum_{j=1}^{p_n} k_j^2)\sum_{j=1}^{p_n} k_j^6}{(\sum_{j=1}^{p_n} k_j^4)^2}. \qquad (6)$$

Monte Carlo study of Chen and Deo[18] showed that for the above choice of β, the distribution of H_n^{β} could be well approximated by normal distribution. In our Monte Carlo simulations the above value of β is used.

Monte Carlo Evidence

In this section, we investigate, through simulations, the finite sample performance of the Hong's and power transformed Hong's test statistics as goodness-of-fit tests for *ARFIMA (p, d, q)* models with dependent errors. We use two sample sizes $n = 100$ and $n = 300$. The error distribution is taken to be standard normal. We use the following four kernels for both tests to examine the effect of different kernels.

Daniel (DAN): $k(w) = \sin(\pi w)/w, \quad w \in (-\infty, \infty)$

Parzen(PAR):

$$k(w) = \begin{cases} 1 - 6(\pi w/6)^2 + 6|\pi w/6|^3 & |w| \le 3/\pi \\ 2(1 - |\pi w/6)^3| & 3/\pi \le |w \le 6/\pi| \\ 0 & otherwise \end{cases}$$

QS:

$$k(w) = (9/5w^2)\{\sin(\sqrt{5/3}\pi w)/\sqrt{5/3}\pi w - \cos(\sqrt{5/3}\pi w)\}$$
$$w \in (-\infty, \infty)$$

Bartlett(BAR): $k(w) = \begin{cases} 1 - |z| & |z| \le 1 \\ 0 & otherwise \end{cases}$

To investigate the effect of p_n, we use three different rates: $p_n = [\ln(n)]$, $p_n = [3n^{0.2}]$ and $p_n = [3n^{0.3}]$. For $n = 100$ these rates deliver p_n = 5, 8, 12 and for $n = 300$ these rates make $p_n = 8, 10, 17$.

To examine the size performance of Hong's and power transformed Hong's test statistics we consider the following models.

M1: $ARFIMA(0,0.4,0)$ - GARCH((0.05, 0.1),0.85)

M2: ARFIMA(0.5,0.4,0) - GARCH((0.05, 0.1),0.85)

For power performance of both tests the following models are used.

M3: ARFIMA(0,.2, 0.4,0) - GARCH((0.05, 0.1),0.85)
alternative fitting model as ARFIMA(0, d ,0)

M4: ARFIMA(0,0.4,0.2) - GARCH((0.05, 0.1),0.85)
alternative fitting model as ARFIMA(0, d ,0)

M5: ARFIMA(0.5,0.4,0.2) - GARCH((0.05, 0.1),0.85)
alternative fitting model as ARFIMA(1, d, 0) .

The results for *M1 – M5* have been shown in table 1 to 5. These results report the percentage rejection rates at nominal levels of 5% and 10% based on 5000 replications. For small sample size *n=100* size distortions occur for both tests but come close to the nominal size for *n=300*. The power transformed test is more undersized as compared to Hong's test statistics. The size is better for *M2* compared to *M1*. There is no significant effect of different kernels on the size of both tests. The size becomes better as we increase the bandwidth p_n. This is true for both sample sizes, tests and different kernels.

Both tests have good power performance for different sample sizes but the power increases as we increase the sample size from 100 to 300. Different kernels have no significant effect on the power of both tests.

Table-1
Rejection rate under the *ARFIMA*(0.2,0.4,0)-*GARCH*((0.05, 0.1),0.85) alternative, fitting model *ARFIMA*(0,d,0)

	n=100						n=300					
	p=5		p=8		p=12		p=8		p=10		p=17	
	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
H BAR	33.12	40.56	29.78	37.56	27.64	35.04	57.12	64.22	55.36	62.72	51.74	59.34
TUK	33.76	41.26	29.26	36.58	26.40	33.60	56.34	63.62	54.26	61.42	49.72	57.54
QS	31.08	38.66	27.16	34.38	25.62	32.10	53.70	61.14	51.84	59.42	47.54	55.32
DAN	32.04	39.58	27.96	35.18	25.80	32.62	54.62	61.94	52.66	60.10	48.22	55.90
H^B BAR	26.56	40.36	24.26	36.60	22.94	34.06	51.10	63.18	49.74	61.70	46.92	58.12
TUK	27.38	41.48	23.68	35.88	22.02	32.28	50.58	62.86	48.62	60.46	44.60	55.98
QS	25.08	38.08	22.30	33.20	21.32	30.88	48.28	59.94	46.32	58.04	42.76	53.92
DAN	26.28	38.68	22.86	34.16	21.80	31.34	49.32	60.66	47.56	58.90	43.70	54.54

Table-2
Rejection rate under the *ARFIMA*(0,0.4,0)-*GARCH*((0.05, 0.1),0.85)

	n=100						n=300					
	p=5		p=8		p=12		p=8		p=10		p=17	
	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
H BAR	6.84	9.42	7.10	10.12	7.16	10.76	3.84	9.02	4.54	9.90	5.66	10.10
TUK	6.78	9.44	7.12	10.10	7.26	10.84	3.96	9.04	4.76	9.60	5.94	9.96
QS	6.84	10.08	7.28	10.48	7.40	11.10	4.10	9.12	5.24	9.70	6.10	10.10
DAN	6.76	9.92	7.16	10.14	7.44	11.16	4.44	9.20	5.34	9.74	6.52	10.20
H^B BAR	4.68	9.34	5.26	9.68	5.26	10.12	3.30	6.70	3.66	7.46	4.80	9.22
TUK	4.58	9.50	5.28	9.78	5.28	10.18	3.38	6.64	3.76	7.66	5.16	9.40
QS	4.92	9.72	5.34	9.92	5.56	10.50	3.86	7.82	4.60	8.56	5.96	10.24
DAN	4.94	9.62	5.34	9.80	5.56	10.58	3.94	7.40	4.40	8.14	5.80	10.20

Table-3
Rejection rate under the *ARFIMA* (0,0.4,0.2)-*GARCH*((0.05, 0.1), 0.85) alternative fitting model *ARFIMA*(0,d,0)

	n=100						n=300					
	p=5		p=8		p=12		p=8		p=10		p=17	
	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
H BAR	33.88	41.58	30.84	37.94	28.12	35.20	63.10	70.12	61.22	68.68	56.80	65.06
TUK	34.40	42.10	30.18	37.38	26.86	33.96	62.48	69.34	60.00	67.46	54.78	62.52
QS	32.00	39.50	28.12	34.94	24.80	32.30	59.52	66.84	57.16	65.02	52.26	60.70
DAN	32.54	21.00	29.02	21.48	25.44	20.66	60.28	68.00	58.04	65.86	52.62	61.38
H^B BAR	27.12	40.34	24.56	35.90	22.94	32.64	56.48	69.08	55.14	67.64	51.58	63.74
TUK	27.76	42.38	24.22	36.48	21.72	32.80	56.02	68.70	54.00	66.58	49.12	61.24
QS	25.74	38.84	22.22	34.02	21.18	30.80	53.42	65.88	51.76	63.72	46.72	59.44
DAN	26.60	39.52	23.28	34.72	21.36	31.28	55.10	66.90	53.16	64.66	47.90	60.18

Table-4
Rejection rate under the *ARFIMA* (0.5,0.4,0)-*GARCH*((0.05, 0.1),0.85) alternative fitting model *ARFIMA*(1,d,0)

	n=100						n=300					
	p=5		p=8		p=12		p=8		p=10		p=17	
	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
H BAR	2.36	3.82	3.38	5.38	4.84	6.62	4.74	7.20	4.98	8.34	5.20	10.42
TUK	2.06	3.48	3.22	5.32	4.70	6.76	4.70	6.98	4.88	8.24	5.10	10.64
QS	2.78	4.48	4.30	6.18	5.84	7.64	4.10	8.50	4.94	10.08	5.20	10.94
DAN	2.62	4.32	3.96	6.98	5.52	8.78	4.76	8.26	5.52	9.54	5.84	10.62
H^B BAR	1.54	5.20	2.40	5.94	3.58	7.36	3.36	6.84	4.16	7.78	5.18	9.68
TUK	1.28	3.48	2.28	5.08	3.56	6.32	3.22	6.62	4.18	7.80	5.30	10.22
QS	1.88	4.36	3.12	5.98	4.32	7.30	3.52	7.12	4.28	8.58	5.32	10.26
DAN	1.76	4.18	2.98	5.68	4.16	6.96	3.22	7.82	4.32	8.96	5.04	10.40

Conclusion

We applied the Hong's and power transformed Hong test of Chen and Deo[6] for goodness-of-fit of autoregressive fractionally integrated moving average models with conditionally hetroscedastic errors of unknown form. Our simulation study reveals that for large sample size *(n = 300)* both the tests have good size and power performance, when applied to different long memory models with conditionally hetroscedastic errors but for small sample *(n = 100)* both tests are undersized. The power transformed test is more undersized compared to Hong's test. This size distortion occurs due to the fact that the mean and variance of these test statistics are based on the asymptotic theory and could be misleading in small samples as reported by Chen and Deo[6]. The above results show that some size correction devices are needed in the above test statistics for *ARFIMA* models with dependent errors of unknown form.

Table-5
Rejection rate under the *ARFIMA*(0.5,0.4,0.2)-*GARCH*((0.05, 0.1),0.85) alternative fitting model *ARFIMA(1,d,0)*

	n=100						n=300					
	p=5		p=8		p=12		p=8		p=10		p=17	
	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
H BAR	15.52	20.72	16.58	22.00	16.68	22.28	46.00	53.10	45.66	52.82	42.98	50.90
TUK	14.72	20.04	16.50	21.88	16.74	22.32	46.22	53.20	45.56	52.94	42.06	50.26
QS	16.24	21.14	16.84	22.42	16.46	21.78	45.42	52.88	43.96	51.84	40.82	48.56
DAN	16.08	20.78	16.82	22.24	16.60	22.06	45.88	53.14	44.96	52.08	41.32	48.58
H^B BAR	10.72	20.68	12.44	21.32	13.32	21.42	39.94	52.28	39.24	51.98	38.14	49.58
TUK	10.16	20.22	12.34	21.36	13.24	21.42	39.94	52.42	39.76	51.80	37.52	48.90
QS	11.64	20.92	13.08	21.50	13.84	20.90	39.44	51.82	38.44	50.58	35.86	47.20
DAN	11.76	20.40	13.06	21.34	13.80	21.20	40.30	51.88	39.30	50.92	36.56	47.58

References

1. Box G.E.P and Pierce D.A., Distribution of the residuals autocorrelations in autoregressive-integrated moving average time series models, *JASA*, **65(332)**, 1509-25 **(1970)**
2. Davies N., Triggs C.M. and Newbold P. Significance levels of the Box-Pierce portmanteau statistics in finite samples, *Biometrika*, **64(3)**, 517-22 **(1977)**
3. Ljung G.M. and Box GEP, On a measure of lack of fit in time series models, *Biometrica*, **65(2)**, 297-303 **(1978)**
4. Ljung G.M., Diagnostic testing of univariate time series models, *Biometrica*, **73(3)**, 725-30 **(1986)**
5. Hong Y., Consistent testing for serial correlation of unknown form, *Econometrica*, 837-864 **(1996)**
6. Chen W. and Deo R.S., Power transformation to induce normality and their applications, *Journal of Royal Statistical Society Series B Statistical Methodology,* **66**, 117-130 **(2004b)**
7. Hurst H., Long-term storage capacity of reservoirs, *Transactions of the American Society of Civil Engineers,* **116**, 770-799 **(1951)**
8. Leland W., Taqqu M., Willinger W. and Wilson D., On the self-similar nature of Ethernet trafic (extended version), *IEEE/ACM Trans*. Network, **2(1),** 1-15 **(1994)**
9. Goodhart C.A.E., and M. O'Hara., High Frequency Data in Financial Markets: Issues and Applications, *Journal of Empirical Finance,* **4**, 73-114 **(1997)**
10. Granger C.W.J. and Joyeux R., An Introduction to Long-Memory Time Series Models and Fractional Differencing, *Journal of Time Series Analysis*, **1**, 15-29 **(1980)**
11. Hosking J.R.M., Fractional differencing, *Biometrika,* **68**, 165–176 **(1981)**
12. Baillie R.T., Chung C.F. and Tieslau M.A., Analysing inflation by the fractionally integrated *ARFIMA-GARCH* model, *Journal of Applied Econometrics*, **11**, 23-40 **(1996)**
13. Hauser M.A. and Kunst R.M., Fractionally integrated models with ARCH errors: With an application to the Swiss 1-month euromarket interest rate, *Review of Quantitative Finance and Accounting,* **10,** 95–113 **(1998)**
14. Lien D. and Tse Y.K., Forecasting the Nikkei spot index with fractional cointegration, *Journal of Forecasting,* **18**, 259–273 **(1999)**
15. Elek P. and Markus L., A long range dependent model with nonlinear innovations for simulating daily river flows, *Natural Hazards and Earth System Sciences,* **4**, 277–283 **(2004)**
16. Koopman S.J., Oohs M. and Carnero M.A., Periodic seasonal Reg-*ARFIMA-GARCH* models for daily electricity spot prices, *Journal of the American Statistical Association,* **102,** 16–27 **(2007)**
17. Lumsdaine R.L. and Ng S., Testing for ARCH in the presence of a possibly misspecified conditional mean, *Journal of Econometrics,* **93**, 257–279 **(1999)**
18. Chen W. and Deo R.S., A generalized portmanteau goodness of fit test for time series models, *Econometric Theory,* **20**, 619-654 **(2004a)**
19. Delgado and Velasco, Distribution-free tests for time series model specification, *Journal of Econometrics*, **155,** 128-137 **(2007)**
20. Delgado M.A., Hidalgo J. and Velasco C., Distribution free goodness-of-fit tests for linear processes, *Annals of Statistics,* **33**, 2568–2609 **(2005)**
21. Hidalgo J. and Kreiss J.P., Bootstrap specification tests for linear covariance stationary processes, *Journal of Econometrics,* **133**, 807–839 **(2006)**
22. Ling S. and Li W.K., On fractionally integrated autoregressive moving-average time series models with conditional heteroscedasticity, *Journal of the American Statistical Association,* **92,** 1184–1194 **(1997)**
23. Li G. and Li W.K., Least absolute deviation estimation for fractionally integrated autoregressive moving average time series models with conditional heteroscedasticity, *Biometrika,* **95**, 399–414 **(2008)**
24. Hong Y. and Lee Y.J., Consistent Testing for Serial Correlation of Unknown Form Under General Conditional Heteroskedasticity, Technical Report, *Departments of Economics and Statistical Sciences, Cornell University*, **(2003)**