Classification of GARCH Time Series: A Simulation Study

Thomas Kalantzis and Demetrios Papanastassiou

University of Macedonia,
Dept. of Applied Informatics,
156 Egnatia Street,
54006 Thessaloniki, Greece
(e-mail: tkalant@uom.gr, papanast@uom.gr)

Abstract. We examine a discrimination rule for time series data generated by a GARCH(1,1) process that classifies a sample into a group in terms of its unconditional variance. A simulation study indicates that our rule is more efficient than a benchmark rule in all cases, except from a narrow range of alternatives lying on the right side of the null.
Keywords: Local heteroscedasticity, Discrimination rule, Likelihood ratio.

1 Introduction

In analyzing high frequency financial time series data, the common practice is to examine the first differences of the logged observations, known as returns. Contrary to the raw prices, returns are considered to be more amenable to statistical manipulations. Under some fundamental economic hypotheses, they form a sample of uncorrelated second order stationary series.

However, if we look at a typical returns plot of reasonable length, we shall observe clusters of different variation, which, at a first sight, may cast some doubt on the issue of the conventional equal variance perception. The main characteristics of this idiosyncratic regular local heteroscedasticity are captured by the widely used GARCH models introduced by [Bollerslev, 1986]. For a returns series, the variance is of practical interest, since it is widely considered as a measure of the risk involved on investing on the particular stock, [Tsay, 2002].

If we want to classify such a series in terms of its variance into one of two groups, in principle we can treat it as an independent sample from identically distributed observations and apply the usual discriminant function, see [Johnson and Wichern, 1992]. However, because of the presence of the local heteroscedasticity, and the fact that independence and normality are challenged both on empirical and theoretical basis, we were motivated to seek discrimination rules which take these facts into account.

In this paper we introduce a likelihood ratio type discrimination rule to classify a GARCH(1,1) process into two categories. Since it is the unconditional long term variance which is mainly of interest, the test concentrates on this aspect. Methods and theory for discriminating processes on an overall basis, mainly of the linear type, are reviewed by [Taniguchi and Kakizawa, 2000] and [Shumway and Stoffer, 2000].

In the sequel, in Section 2 we discuss the discrimination rule for an appropriately parameterized GARCH(1,1) model. In Section 3 we present the results of a simulation study comparing our approach against the benchmark rule of independent and identically distributed, iid, observations. The final section summarizes our conclusions and suggestions.

2 The GARCH(1,1) Discrimination Rule

Let $Y_t, t = 1, 2, ..., n$, be a set of normally distributed iid observations. Suppose one samples from either of two groups, $G_1 : N(0, \sigma_1^2)$ or $G_2 : N(0, \sigma_2^2)$. The conventional likelihood ratio based rule, see [Johnson and Wichern, 1992], states that

classify the sample as belonging to G_1 when $ln\frac{L_1}{L_2} \geq 0$,

while

classify the sample as belonging to G_2 when $ln\frac{L_1}{L_2} < 0$, (1)

where L_j is the sample likelihood value, supposing it comes from G_j, $j = 1, 2$. More precisely, the discriminant function is

$$\ln\frac{L_1}{L_2} = -\frac{T}{2}\ln\frac{\sigma_1^2}{\sigma_2^2} - \frac{1}{2}\sum_{t=1}^{t=T} y_t^2\left(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2}\right). \tag{2}$$

On the other hand, suppose our data are generated by a stationary GARCH(1,1) process, [Bollerslev, 1986],

$$\begin{aligned} Y_t &= u_t, \\ u_t &= \varepsilon_t h_t^{1/2}, \varepsilon_t \overset{iid}{\text{sim}} N(0,1), \\ h_t &= a_0 + a_1 u_{t-1}^2 + b_1 h_{t-1}, \end{aligned} \tag{3}$$

ε_t independent of h_t, and $a_0 > 0$, $a_1, b_1 \geq 0$, are constant parameters. It is easy to see that the unconditional variance of Y_t is

$$\sigma^2 = E(Y_t^2) = \frac{a_0}{1 - a_1 - b_1},$$

and that, although Y_t are uncorrelated, they are not independent and normally distributed, see [Hamilton, 1994], amongst many others. Since σ^2 is the parameter of our prime interest, we reparameterize the model in terms of σ^2, writing

$$h_t = \sigma^2(1 - a_1 - b_1) + a_1 u_{t-1}^2 + b_1 h_{t-1}. \tag{4}$$

Group G_j, $j = 1, 2$, is described as the set of all possible GARCH(1,1) models that have the same variance σ_j^2, $j = 1, 2$. We are interested to allocate a sample $Y_t, t = 1, 2, ..., T$, to one of the G_1 and G_2 groups in terms only of its variance σ^2. The a_1 and b_1 parameters, parameterizing the dynamic behavior of the conditional variance, are a sort of nuisance parameters.

The likelihood based rule will remain as in (1), but the likelihood ratio in (2) is modified to take into account the special form of our heteroscedastic data. Noting that the conditional distribution of Y_t given h_t is a normal $N(0, h_t)$, the decomposition of the likelihood function of a time series process yields the discriminant function

$$\ln \frac{L_1}{L_2} = -\frac{1}{2}\sum_{t=1}^{T}\left(\ln \frac{h_{1t}}{h_{2t}}\right) - \frac{1}{2}\sum_{t=1}^{T} y_t^2 \left(\frac{1}{h_{1t}} - \frac{1}{h_{2t}}\right). \tag{5}$$

3 Simulation Study

We carried out a simulation study to assess the GARCH discriminant function in (5) against (2), which we consider as a sort of benchmark rule. The experimental data come from the GARCH(1,1) model in (3) with its conditional variance reparameterized as in (4). Examining real daily or weekly series of stock or exchange rate returns, we calculated their free variance to be of the order of $5 \cdot 10^{-5}$. We considered that as a typical variance value of real life data, and in our experiments we set the variance of group G_1 equal to this value, that is $\sigma_1^2 = 5 \cdot 10^{-5}$. The remaining parameters a_1 and b_1 take a range of values within what is considered as typical in the relative literature. We mention that the condition for (3) to be stationary is $a_1 + b_1 < 1$. Usually, in real series applications, the sum of a_1 and b_1 lies close to 1, and a_1 is always smaller than b_1. When $a_1 + b_1 = 1$ the model is still stationary, but with infinite variance and therefore makes no sense for our study.

Models	0	1	2	3	4	5
Parameters						
a_1	0.00	0.10	0.10	0.40	0.40	0.40
b_1	0.00	0.50	0.80	0.50	0.55	0.58
Sum	0.00	0.60	0.90	0.90	0.95	0.98

Table 1. Models tested in the simulation

The criterion to assess our findings was the error rate $P(2|1)$, that is the probability to allocate a sample to G_2 when it truly comes from G_1. These probabilities are reported in the corresponding tables and are calculated by repeating the same experiment 300 times. Factors we felt that might be of influence in the efficiency of the discrimination rules were the sample size

T, the magnitude of the alternative variance in G_2, and the combination of the a_1 and b_1 values. The simulation study was designed to take into consideration all these factors. In Table 1 we present only a few selected combinations of a_1 and b_1 values from those examined, declared as models 0 to 5. Practically, Model 0 is an iid series.

σ_2^2	1	2	3	4	4.5	4.9	5.1	5.5	6	7	8	9
series length												
GARCH rule												
300	.000	.053	.313	.500	.550	.567	.417	.393	.350	.300	.277	.250
1000	.000	.007	.127	.400	.513	.560	.417	.383	.337	.280	.227	.193
benchmark rule												
300	.000	.000	.003	.107	.317	.487	.420	.240	.107	.020	.003	.000
1000	.000	.053	.313	.013	.143	.460	.397	.160	.010	.000	.000	.000

Table 2. Error rates for Model 5. Alternative variance values should be multiplied by 10^{-5}. The true variance of the series is $5 \cdot 10^{-5}$

Before presenting our results, we clarify the computational flow of our procedure. Once we had in hand a series from G_1, we maximized L_1 with respect to a_1 and b_1 considering σ^2 known and equal to $\sigma_1^2 = 5 \cdot 10^{-5}$. Next, we maximized L_2 for a_1 and b_1 setting now $\sigma^2 = \sigma_2^2$, one of the alternatives. A conjugate gradient routine was written to maximize the loglikelihoods, after transforming a_1 and b_1 so that the restrictions, $a_1, b_1 \geq 0$ and $a_1 + b_1 < 1$ were fulfilled. If maximization of both L_1 and L_2 was terminated successfully, then (5) was calculated and the series was classified into G_1 or G_2 accordingly.

The most definite of our conclusions is that both rules perform better as alternative variance σ_2^2 takes values further away from σ_1^2. Also, the $P(2/1)$ error rate improves with the sample size, and this can be seen for the case of Model 3 in Table 2. Since the general pattern of $P(2/1)$ is the same for either $T = 300$ or $T = 1000$, for reasons of space economy, we report more detailed results in Table 3 only for $T = 1000$.

Concerning the effect of the sum $\alpha_1 + \beta_1$, the error rate for both rules increases as $\alpha_1 + \beta_1$ approaches unity. For models with the same sum, the rate is worse for larger α_1, see for instance Model 2 versus Model 3 in Table 3. This can be explained by the fact that larger α_1 implies wider local variance bursts.

Regarding the relative performance of the GARCH rule against the benchmark rule, which is of the main interest in our study, there is not a clear pattern for the whole range of alternatives. The GARCH rule is always better than the benchmark for σ_2^2 smaller than the true $\sigma_1^2 = 5 \cdot 10^{-5}$. For a range of alternatives from $5.1 \cdot 10^{-5}$ to approximately $10 \cdot 10^{-5}$, the benchmark rule outperforms the GARCH rule. This can be seen graphically in Fig.1 for the

σ_2^2:	1	2	3	4	4.5	4.9	5.1	5.5	6	7	8	9
GARCH rule												
model 0	.000	.000	.003	.020	.133	.447	.413	.170	.133	.000	.000	.000
model 1	.000	.000	.000	.023	.190	.473	.410	.210	.070	.003	.000	.000
model 2	.010	.000	.000	.117	.316	.483	.447	.310	.197	.073	.027	.000
model 3	.000	.007	.127	.400	.513	.560	.417	.383	.337	.280	.227	.193
model 4	.000	.090	.363	.500	.570	.563	.420	.387	.350	.280	.260	.240
model 5	.090	.487	.663	.740	.783	.787	.193	.177	.167	.143	.110	.080
benchmark rule												
model 0	.000	.000	.000	.013	.143	.460	.397	.160	.010	.000	.000	.000
model 1	.000	.000	.000	.030	.197	.480	.410	.207	.057	.003	.000	.000
model 2	.003	.000	.000	.137	.350	.500	.423	.297	.163	.037	.033	.000
model 3	.000	.033	.290	.051	.610	.677	.303	.257	.227	.197	.167	.133
model 4	.010	.307	.557	.710	.737	.747	.233	.213	.200	.180	.163	.133
model 5	.370	.690	.790	.827	.840	.853	.140	.133	.117	.103	.087	.083

Table 3. Error rates for models 0 to 5. Alternative variance values should be multiplied by 10^{-5}. The true variance of the series is $5 \cdot 10^{-5}$.

case of Model 3. The superiority of the benchmark rule grows larger as the sum $\alpha_1 + \beta_1$ approaches unity.

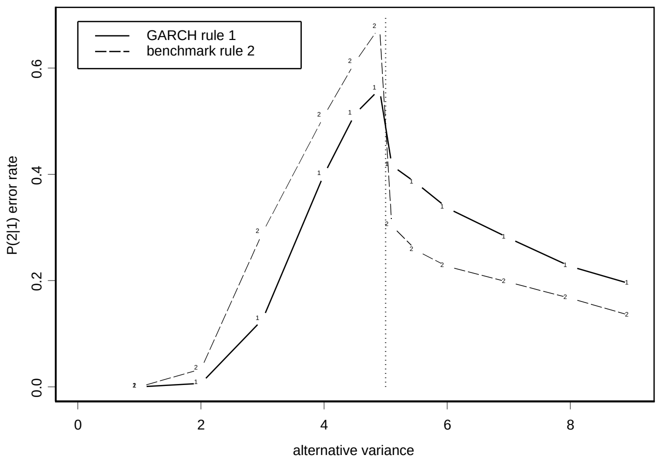

Fig. 1. Error rates for Model 3. Variances should be multiplied by 10^{-5}.

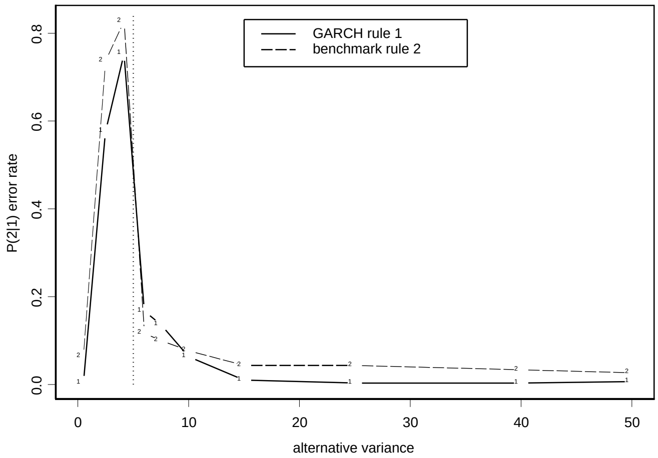

Fig. 2. Error rates for Model 5. Variances should be multiplied by 10^{-5}.

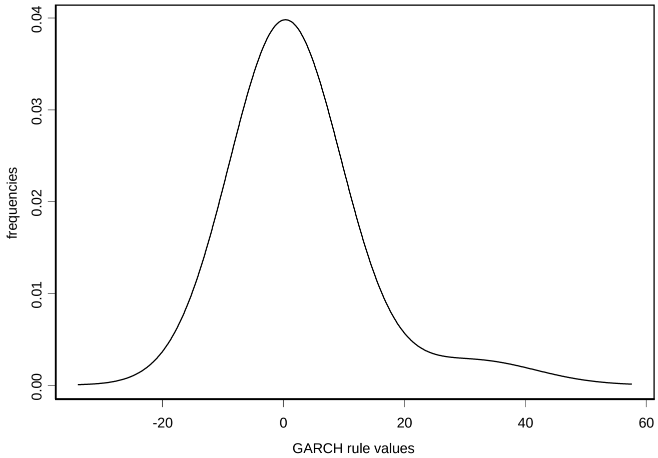

Fig. 3. Smoothed frequency curve from 2000 replications from the GARCH rule, Model 5, $\sigma_2^2 = 7 \cdot 10^{-5}$.

Moving farther to the right of σ_1^2, the pattern is reversing. Fig.2 illustrates the case for Model 5. Note that error rates in this interval are not reported in Table 3. We can not explain this behavior. Finally, Fig.3 gives a smoothed plot of the distribution of (5) for Model 5 with $\sigma_2^2 = 7 \cdot 10^{-5}$. The simulated distribution was plotted from 2000 replications.

4 Conclusions

We conducted an empirical study classifying GARCH(1,1) time series data on the basis of their unconditional variance. The procedure may prove useful to classify financial returns data into different risk groups.

The GARCH rule is better than the benchmark rule, except from a small range of alternatives starting from the null σ_1^2 and going approximately up to $\sigma_2^2 = 2\sigma_1^2$. This is a point that deserves further investigation, and a proper derivation of the distribution of (5) may shed some light.

Rule (5) generalizes easily for higher order GARCH models, although for most applications a simple GARCH(1,1) suffices. Experience with real data could allow us to cross examine rule (5) with other risk classifying criteria, such as the β coefficient value provided by financial econometric theory.

References

[Bollerslev, 1986]T. Bollerslev. Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, pages 307–327, 1986.

[Hamilton, 1994]J.D. Hamilton. *Time Series Analysis.* Princeton University Press, Chichester, 1994.

[Johnson and Wichern, 1992]R.A. Johnson and D.W. Wichern. *Applied Multivariate Statistical Analysis.* Prentice-Hall International, New York, 1992.

[Shumway and Stoffer, 2000]R.H. Shumway and D.S. Stoffer. *Time Series Analysis and Its Applications.* Springer-Verlag, New York, 2000.

[Taniguchi and Kakizawa, 2000]M. Taniguchi and Y. Kakizawa. *Asymptotic Theory of Statistical Inference for Time Series.* Springer-Verlag, New York, 2000.

[Tsay, 2002]R.S. Tsay. *Analysis of Financial Time Series.* Wiley-Interscience, 2002.