A better test for autocorrelation in financial time series

William J. Egan, Ph.D.
wjegan@gmail.com
January 19, 2008

Abstract

This paper tests the accuracy of the commonly used cutoffs for determining the statistical significance of autocorrelations in time series. Monte Carlo simulations with 50,000 replicates were used to generate 95% confidence limits by varying sample size from 21 to 252 using both normally distributed and t-distributed data. The simulations show that the confidence limits derived from the commonly used formulas are biased at sample sizes of less than several hundred and should not be used.

Keywords: confidence limits, statistical significance testing, autocorrelation function, acf, a.c.f., partial autocorrelation function, pacf, p.a.c.f., Monte Carlo simulation, autoregressive, moving average, ARMA, ARIMA, Box-Jenkins, time series, sample size

Introduction

Building a model to forecast the future values of a time series requires that we determine if there are any statistically significant autocorrelations in the data. The classic autoregressive integrated moving average (ARIMA) model of Box and Jenkins, as well as the autoregressive conditional heteroskedasticity (ARCH) family of models, use analysis of autocorrelation to guide the model building process. [1-5] The Box-Jenkins model identification procedure involves tests of the statistical significance of the elements of the autocorrelation function (ACF) and partial autocorrelation function (PACF). These tests are used to determine if autoregressive and/or moving average patterns are present in the time series.

These statistical tests check if the observed autocorrelations exceed theoretical cutoffs. If they do, the autocorrelation at that time lag is deemed significant. Incorrect cutoffs would cause serious problems. Using incorrect cutoffs in the model building process will miss real autocorrelations or lead to the inclusion of false autocorrelations; either case will produce poor models giving misleading predictions. There are two possible causes of incorrect cutoffs which we will investigate in this paper.

First, the usual calculations of the standard errors of the autocorrelation coefficients assume that the autocorrelation coefficients are asymptotically normally distributed. This means that as the sample size increases, the distribution of the autocorrelation coefficients becomes more normal. Therefore, the first question is: how much error is introduced by the assumption of normality when the sample size is small? It is interesting that texts on time series analysis commonly only provide approximate formulas.

Second, the distribution of most financial time series is non-normal with heavy tails. As I have previously shown, the distribution of the S&P 500 stock index is matched extremely well by the t-distribution with location-scale parameters. [6] Thus, the second question is: are the standard cutoffs for testing the statistical significance of autocorrelations correct when applied to heavy-tailed financial data?

Standard calculations

The ACF is calculated using equations 2.1.9 and 2.1.10 from Box, Jenkins, Reinsel (BJR) [1]:

$$r_k = \frac{c_k}{c_0} \quad \text{where } c_k = \frac{1}{N}\sum_{t=1}^{N-k}(z_t - \bar{z})(z_{t+k} - \bar{z}) \text{ and } k = 0, 1, 2, \ldots, K$$

$\bar{z}$ is the sample mean of the time series, N is the number of observations, t is the index of the observation, and k is the lag. BJR recommends that the analyst keep $k < N/4$ and $N > 50$ for meaningful results.

BJR notes that the autocorrelation coefficients tend to be independently and normally distributed with mean 0 and variance $1/N$ for large N. The main formula used to compute approximate 95% confidence limits of for the r_k values is $-1/N \pm 2/N^{½}$, which is often shortened to $\pm 2/N^{½}$. [1,2].

The PACF was calculated by the method of successively fitting linear regressions and retaining the last regression coefficient computed at each step. The asymptotic 95% confidence intervals for PACF values are also commonly computed using the formula $\pm 2/N^{½}$.[1,2] PACF values outside these cutoffs are considered statistically significant.

The algorithms used to compute ACF and PACF were verified against the published results for the batch data in BJR (Tables 2.1, 2.2, 3.1).[1] Results agree and are shown in Table 1.

Answering the questions with Monte Carlo simulations

A straightforward way to answer both questions is to use Monte Carlo simulations. By generating thousands of random time series of varying lengths using both the normal and t-distributions, we can obtain excellent estimates of the best cutoff values to test the statistical significance of autocorrelation coefficients. We also avoid the difficult task of determining an exact formula which combines the effects of data distribution and sample size.

Simulation parameters were as follows. Each simulation generated 50,000 random time series. The ACF and PACF were computed for each of the 50,000 time series. To generate the random normally distributed data, parameters were set to mean = 0 and standard deviation = 1. To generate random data with a t-distribution, I used the previously published parameter set that fit the daily percentage changes of the S&P 500 stock index.[6] The 2.5% and 97.5% points of the ACF and PACF distributions at a given time lag were used to define the Monte Carlo 95% confidence limits. The Monte Carlo confidence limits were then compared with the theoretical confidence limits.

The simulations

The first two simulations compared the ACF at lag = 1 for both distributions. Figure 1 shows the Monte Carlo 95% confidence limits for the normally distributed data (blue) vs. the standard approximation for the theoretical confidence limits ($\pm\ 2/N^{½}$ in black). While the lower bounds match quite well, the upper bound of the simulation is lower than the theoretical limits, especially for smaller sample sizes. Figure 2 shows the same plot as Figure 1 with the Monte Carlo 95% confidence limits from second simulation using data generated with the t-distribution plotted in red. These bounds are slightly narrower.

The effect of the “full” approximation ($-1/N \pm 2/N^{½}$) for the 95% theoretical confidence limits is important. Figure 3 shows how the $-1/N$ term shifts the theoretical confidence limits and makes them more symmetric when compared to the center of the confidence limits from the simulations.

Figure 4 to 6 show the same plots for the PACF. Similar effects are present.

The simulations also show that using a constant cutoff for statistical significance of autocorrelations at various time lags for the same N is incorrect for smaller N. Tables 2 and 3 show the Monte Carlo vs. theoretical 95% confidence limits for the ACF lags 1 to 15 for N = 63 and N = 252. Figure 7 plots the Monte Carlo confidence limits for the ACF and PACF at N=63 with the $\pm\ 2/N^{½}$ limits.

At the smaller sample size, the Monte Carlo 95% confidence limits are narrower by approximately 0.09 at lag 15 for the ACF. The effect is much smaller (0.017) at lag 15 for N = 252. (The values 63 and 252 for N were chosen because they are the number of trading days in a quarter and a year, respectively.)

Why do the confidence limits narrow as the lag is increased for the ACF at small sample sizes? At first glance, this seems backwards. Larger lags decrease the sample size, and confidence limits widen as sample size decreases. The equation used to compute r_k (shown above) provides the explanation. The mean of the time series ($\bar{z}$) and N are not adjusted as k increases. Consequently, as the lag k increases, the denominator c_o is constant and only the numerator decreases.

Other the other hand, Monte Carlo confidence limits for the PACF widen because regressions at higher lags require removal of data and thus reduce sample size. The alternating stair pattern in the PACF confidence limits 1) slowly disappears as sample size increase and 2) is present for PACF values calculated using both linear regression and the Durbin recursive algorithm (data not shown). This effect is likely due to the discrete jumps that occur when lagging data for the regression/correlation calculation when the sample size is small.

Conclusions

The answered given by the Monte Carlo simulations to our two questions revealed several problems with the usual tests for autocorrelation:

1. The commonly used approximation $\pm 2/N^{½}$ used to define 95% confidence limits is incorrect for small sample sizes, but only for the upper limit. For N = 50, the approximation gives $\pm$ 0.283 while the simulation using the t-distribution gives 95% confidence limits of -0.283 to + 0.239. This effect would cause some real positive autocorrelations to be missed. As Figure 2 shows, the effect persists out to N = 252.

2. The “full” approximation $-1/N \pm 2/N^{½}$ shifts the theoretical bounds up. For N = 50, the “full” approximation gives -0.303 to +0.263 for the 95% confidence limits while the simulation using the t-distribution gives 95% confidence limits of -0.283 to + 0.239.

3. There is a consistent, albeit small difference between Monte Carlo 95% confidence limits generated from normally distributed data vs. heavy-tailed data.

4. As Table 2 and Figure 7 show, even when the sample size is moderate (N=63), the Monte Carlo 95% confidence limits for the ACF shrink considerably as the lag is increased. In contrast, the cutoffs given by the usual approximations are constant. This effect is quite small at larger sample size (N = 252); see Table

3. For the PACF, the Monte Carlo confidence limits widen as lag increases (Figure 7).

It is recommended that Monte Carlo simulation be used to generate the cutoffs for testing the statistical significance of autocorrelations for time series if the sample size is less than several hundred.

Acknowledgements

I would like to thank Dr. Marie Egan for her helpful review of this paper.

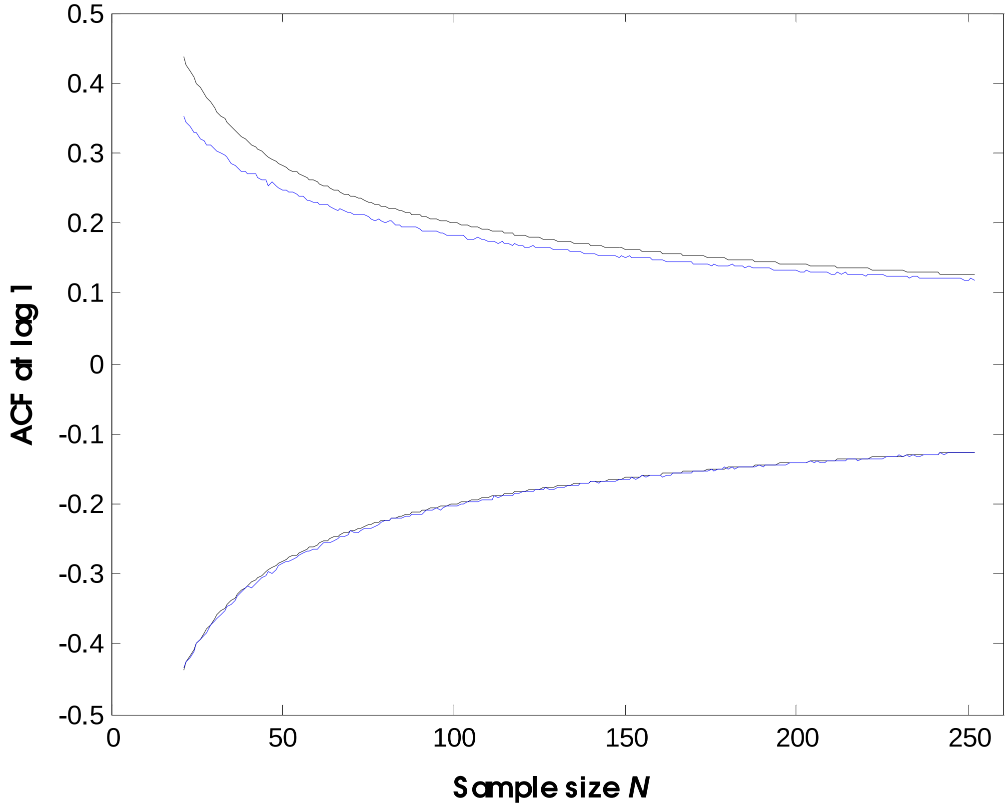

Figure 1. Plot of autocorrelation function vs. sample size for lag = 1. Black lines are the theoretical 95% confidence limits ($\pm\ 2/N^{½}$) and the blue lines are the Monte Carlo 95% confidence limits from the simulation using normally distributed data.

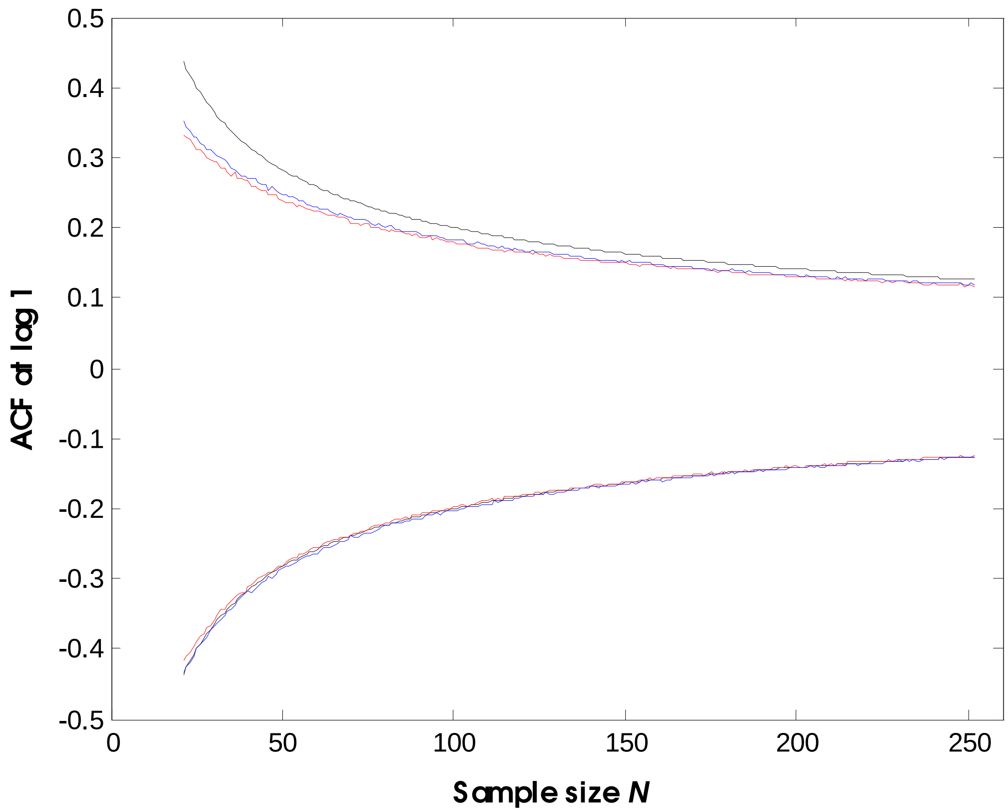

Figure 2. Same as Figure 1 with added red lines representing the Monte Carlo confidence limits generated using data from the t-distribution.

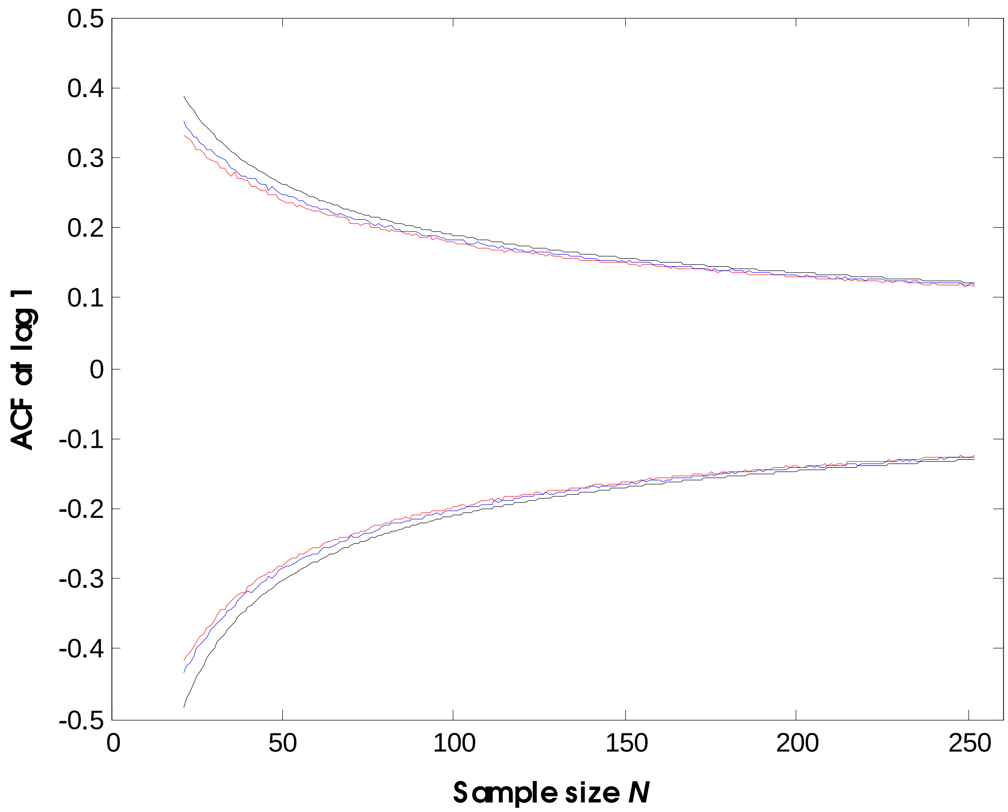

Figure 3. Same as Figure 2 but with the theoretical 95% confidence limits (black) computed using the "full" approximation ($-1/N \pm 2/N^{1/2}$). The extra $-1/N$ term shifts the theoretical limits.

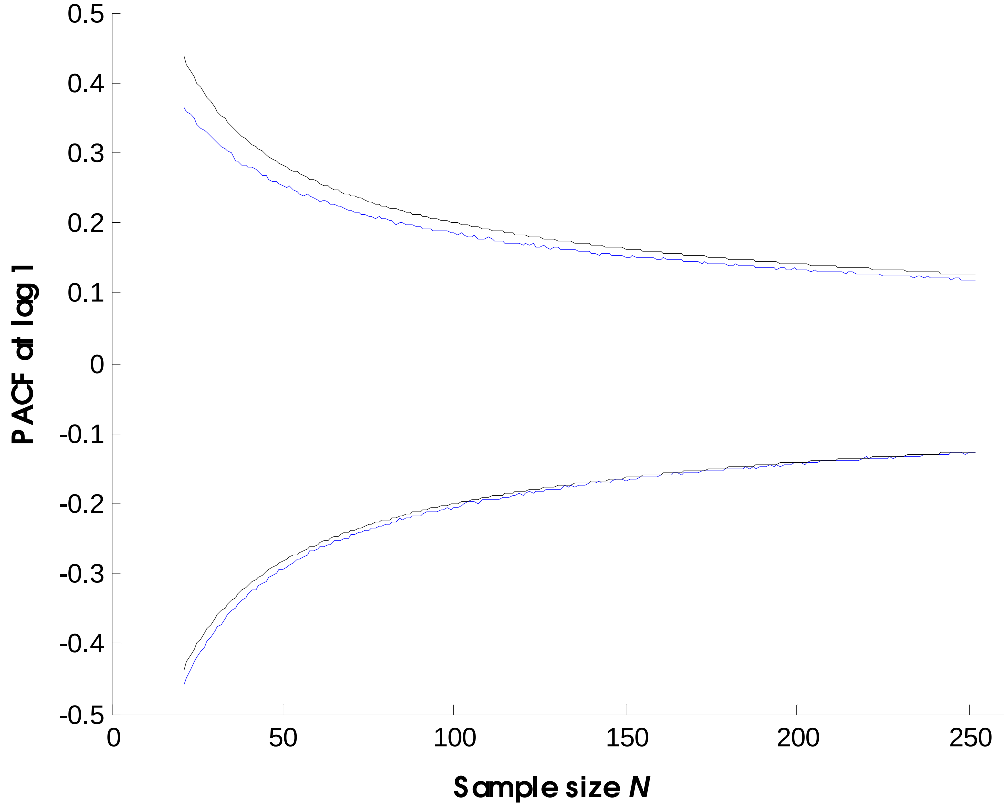

Figure 4. Plot of partial autocorrelation function vs. sample size for lag = 1. Black lines are the theoretical 95% confidence limits ($\pm 2/N^{½}$) and the blue lines are the Monte Carlo 95% confidence limits from the simulation using normally distributed data.

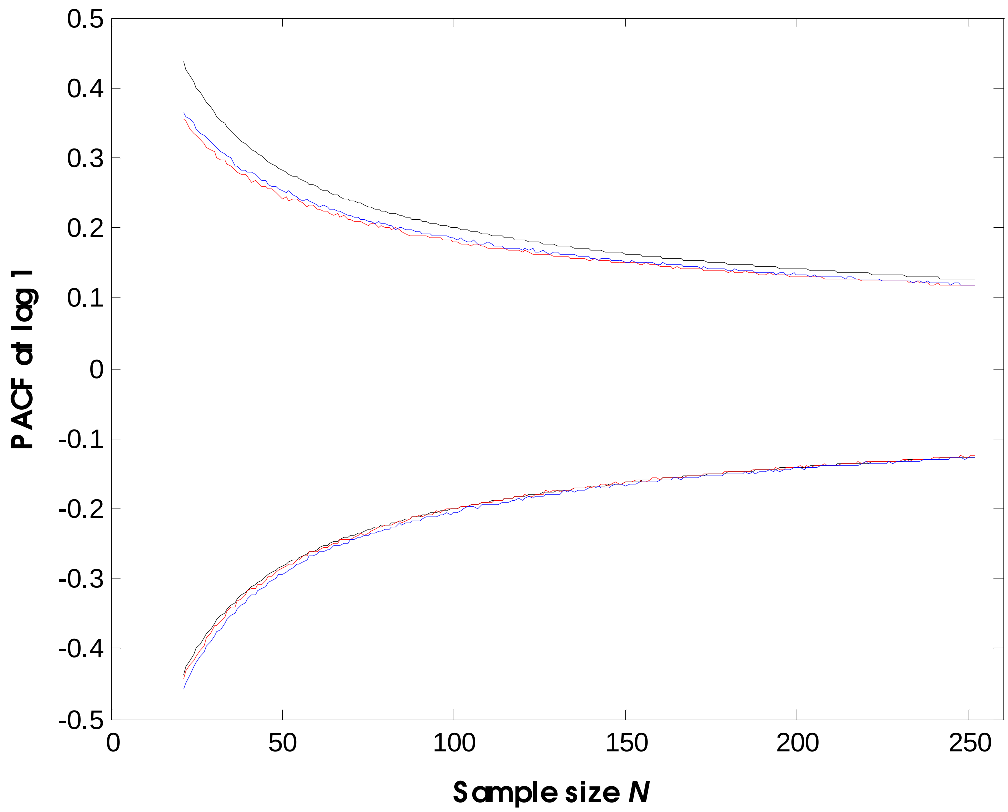

Figure 5. Same as Figure 4 with added red lines representing the Monte Carlo confidence limits generated using data from the t-distribution

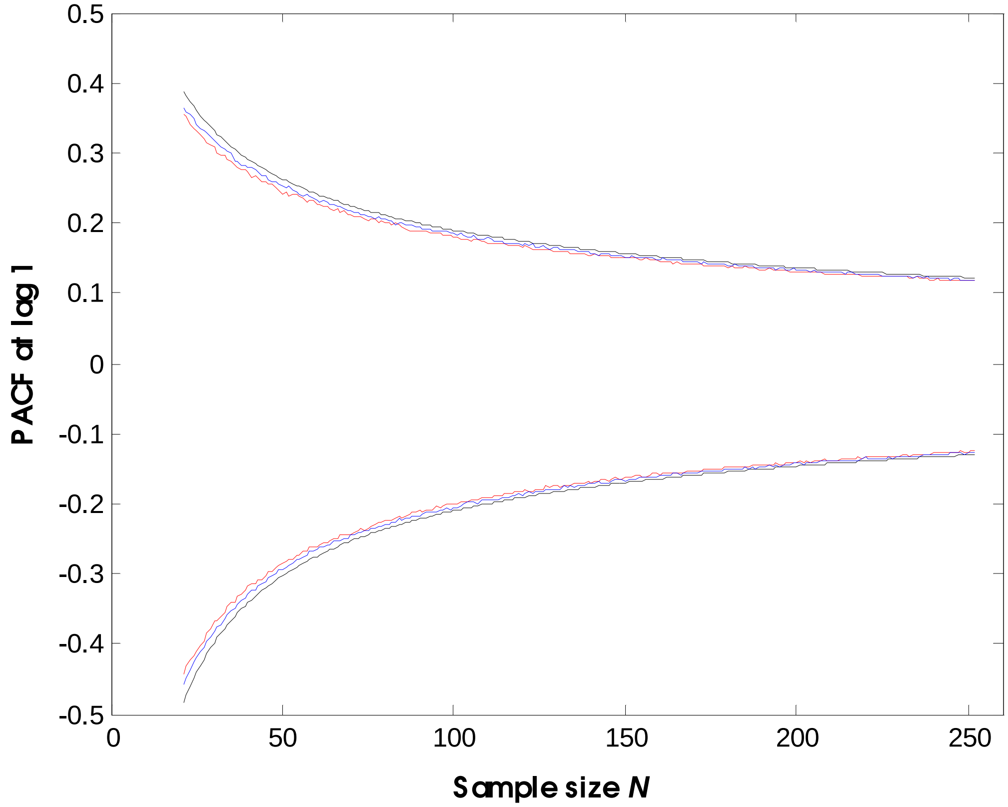

Figure 6. Same as Figure 5 but with the theoretical 95% confidence limits (black) computed using the "full" approximation ($-1/N \pm 2/N^{½}$).

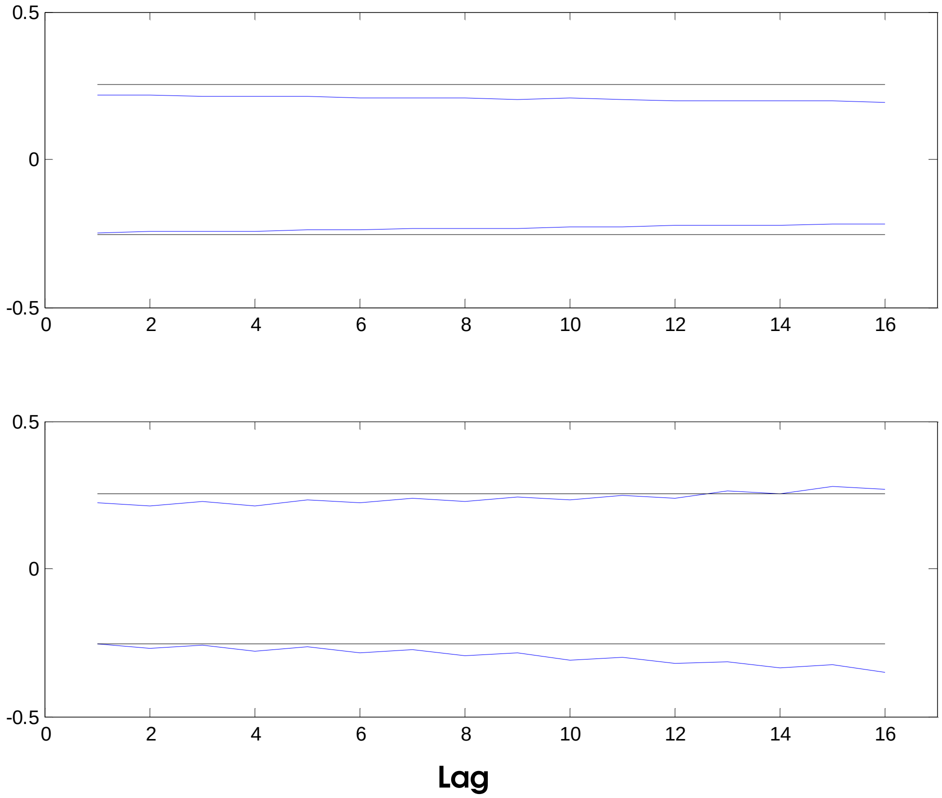

Figure 7. Plot of ACF (top) and PACF(bottom) with $\pm\ 2/N^{½}$ confidence limits in black and Monte Carlo confidence limits in blue for N = 63.

Table 1. Comparison of results from algorithms used in this paper to compute ACF and PACF vs. published results for the batch dataset in BJR. Results are in good agreement.

Lag	published BJR ACF	ACF this paper	published BJR PACF	PACF this paper
1	-0.39	-0.39	-0.40	-0.42
2	0.30	0.30	0.19	0.19
3	-0.17	-0.17	0.01	0.01
4	0.07	0.07	-0.07	-0.06
5	-0.10	-0.10	-0.07	-0.07
6	-0.05	-0.05	-0.15	-0.14
7	0.04	0.04	0.05	0.05
8	-0.04	-0.04	0.00	0.00
9	-0.01	0.00	-0.10	-0.10
10	0.01	0.01	0.05	0.05
11	0.11	0.11	0.18	0.18
12	-0.07	-0.07	-0.05	-0.05
13	0.15	0.15	0.09	0.09
14	0.04	0.04	0.18	0.18
15	-0.01	-0.01	0.01	0.01

Table 2. Differences between 95% Monte Carlo confidence limits and theoretical constant cutoffs for N = 63.

Lag	Lower 95% simulation	Upper 95% simulation	Lower -1/N-2/sqrt(N)	Upper -1/N+2/sqrt(N)	Lower -2/sqrt(N)	Upper 2/sqrt(N)
1	-0.248	0.220	-0.268	0.236	-0.252	0.252
2	-0.248	0.219	-0.268	0.236	-0.252	0.252
3	-0.244	0.215	-0.268	0.236	-0.252	0.252
4	-0.244	0.216	-0.268	0.236	-0.252	0.252
5	-0.239	0.213	-0.268	0.236	-0.252	0.252
6	-0.241	0.211	-0.268	0.236	-0.252	0.252
7	-0.237	0.207	-0.268	0.236	-0.252	0.252
8	-0.235	0.206	-0.268	0.236	-0.252	0.252
9	-0.233	0.206	-0.268	0.236	-0.252	0.252
10	-0.230	0.204	-0.268	0.236	-0.252	0.252
11	-0.228	0.201	-0.268	0.236	-0.252	0.252
12	-0.226	0.201	-0.268	0.236	-0.252	0.252
13	-0.222	0.198	-0.268	0.236	-0.252	0.252
14	-0.223	0.196	-0.268	0.236	-0.252	0.252
15	-0.220	0.193	-0.268	0.236	-0.252	0.252

Table 3. Differences between 95% Monte Carlo confidence limits and theoretical constant cutoffs for N = 252.

Lag	Lower 95% simulation	Upper 95% simulation	Lower -1/N-2/sqrt(N)	Upper -1/N+2/sqrt(N)	Lower -2/sqrt(N)	Upper 2/sqrt(N)
1	-0.124	0.118	-0.130	0.122	-0.126	0.126
2	-0.124	0.117	-0.130	0.122	-0.126	0.126
3	-0.124	0.119	-0.130	0.122	-0.126	0.126
4	-0.124	0.117	-0.130	0.122	-0.126	0.126
5	-0.124	0.117	-0.130	0.122	-0.126	0.126
6	-0.125	0.118	-0.130	0.122	-0.126	0.126
7	-0.123	0.116	-0.130	0.122	-0.126	0.126
8	-0.124	0.116	-0.130	0.122	-0.126	0.126
9	-0.123	0.114	-0.130	0.122	-0.126	0.126
10	-0.122	0.115	-0.130	0.122	-0.126	0.126
11	-0.123	0.116	-0.130	0.122	-0.126	0.126
12	-0.122	0.116	-0.130	0.122	-0.126	0.126
13	-0.123	0.115	-0.130	0.122	-0.126	0.126
14	-0.124	0.114	-0.130	0.122	-0.126	0.126
15	-0.122	0.113	-0.130	0.122	-0.126	0.126

References

1. G. E. P. Box, G. M. Jenkins, G. C. Reinsel *Time Series Analysis: Forecasting & Control, 3rd ed.*, Prentice Hall, 1994.
2. C. Chatfield, *The Analysis of Time Series, An Introduction, 4th ed.*, Chapman and Hall, 1989.
3. A. Pankratz, *Forecasting with Univariate Box-Jenkins Models: Concepts and Cases*, Wiley, 1983.
4. W. Enders, *Applied Econometric Time Series, 2^{nd} ed.*, Wiley, 2004.
5. Professor Robert Nau's Decision 411 Forecasting course, Duke University, Fuqua School of Business, http://www.duke.edu/~rnau/411home.htm
6. W. J. Egan, "The Distribution of S&P 500 Index Returns," posted online Jan 10, 2007 at http://ssrn.com/abstract=955639