Institute for Empirical Research in Economics
University of Zurich

Working Paper Series
ISSN 1424-0459

Working Paper No. 480

Central Limit Theorems When Data Are Dependent: Addressing the Pedagogical Gaps

Timothy Falcon Crack and Olivier Ledoit

February 2010

Central Limit Theorems When Data Are Dependent: Addressing the Pedagogical Gaps

Timothy Falcon Crack[1]
University of Otago

Olivier Ledoit[2]
University of Zurich

Version: August 18, 2009

[1]Corresponding author, Professor of Finance, University of Otago, Department of Finance and Quantitative Analysis, PO Box 56, Dunedin, New Zealand, tcrack@otago.ac.nz
[2]Research Associate, Institute for Empirical Research in Economics, University of Zurich, oledoit@iew.uzh.ch

Central Limit Theorems When Data Are Dependent: Addressing the Pedagogical Gaps

ABSTRACT

Although dependence in financial data is pervasive, standard doctoral-level econometrics texts do not make clear that the common central limit theorems (CLTs) contained therein fail when applied to dependent data. More advanced books that are clear in their CLT assumptions do not contain any worked examples of CLTs that apply to dependent data. We address these pedagogical gaps by discussing dependence in financial data and dependence assumptions in CLTs and by giving a worked example of the application of a CLT for dependent data to the case of the derivation of the asymptotic distribution of the sample variance of a Gaussian AR(1). We also provide code and the results for a Monte-Carlo simulation used to check the results of the derivation.

INTRODUCTION

Financial data exhibit dependence. This dependence invalidates the assumptions of common central limit theorems (CLTs). Although dependence in financial data has been a high-profile research area for over 70 years, standard doctoral-level econometrics texts are not always clear about the dependence assumptions needed for common CLTs. More advanced econometrics books are clear about these assumptions but fail to include worked examples of CLTs that can be applied to dependent data. Our anecdotal observation is that these pedagogical gaps mean that doctoral students in finance and economics choose the wrong CLT when data are dependent.

In what follows, we address these gaps by discussing dependence in financial data and dependence assumptions in CLTs, giving a worked example of the application of a CLT for dependent data to the case of the derivation of the asymptotic distribution of the sample variance of a Gaussian AR(1), and presenting a Monte-Carlo simulation used to check the results of the derivation. Details of the derivations appear in Appendix A, and MATLAB code for the Monte-Carlo simulation appears in Appendix B.

DEPENDENCE IN FINANCIAL DATA

There are at least three well-known explanations for why dependence remains in financial data, even though the profit-seeking motives of thousands of analysts and traders might naively be expected to drive dependence out of the data: microstructure effects, rational price formation that allows for dependence, and behavioral biases. First, microstructure explanations for dependence include robust findings such as thin trading induced index autocorrelation [Fisher, 1966, p. 198; Campbell, Lo, and MacKinlay, 1997, p. 84], spurious cross-autocorrelations [Campbell, Lo, and MacKinlay, 1997, p. 129], genuine cross-autocorrelations [Chordia and Swaminathan, 2000], and bid-ask bounce induced autocorrelation [Roll, 1984; Anderson et al., 2006]. Second, we may deduce from Lucas [1978], LeRoy [1973], and Lo and MacKinlay [1988] that, even if stock market prices satisfy the “efficient markets hypothesis,” rational prices need not follow random walks. For example, some residual predictability will remain in returns if investor risk aversion is high enough that strategies to exploit this predictability are considered by investors to be too risky to undertake. Third, behavioral biases like “exaggeration, oversimplification, or neglect” as identified by Graham and Dodd [1934, p. 585] are robust sources of predictability. Popular examples of these include DeBondt and Thaler [1985, 1987], who attribute medium-term reversal to investor over-reaction to news, and Jegadeesh and Titman [1993], who attribute short-term price momentum to investor under-reaction to news. More recently, Frazzini [2006] documents return predictability driven by the "disposition effect" (i.e., investors holding losing positions, selling winning positions, and therefore under-reacting to news).

Dependence in financial data causes problems for statistical tests. Time series correlation "…is known to pollute financial data…and to alter, often severely, the size and power of testing procedures when neglected" [Scaillet and Topaloglou, 2005, p. 1]. For example, Hong et al. [2007] acknowledge the impact of time series dependence in the form of both volatility clustering and weak autocorrelation for stock portfolio returns. They use a CLT for dependent data from White [1984] to derive a test statistic for asymmetry in the correlation between portfolio and market returns depending upon market direction. Cross-sectional correlation also distorts test statistics and the use of CLTs. For example, Bollerslev et al. [2007] discuss cross-correlation in stock returns as their reason for abandoning CLTs altogether when trying to derive an asymptotic test statistic to detect whether intradaily jumps in an index are caused by co-jumps in individual index constituents. Instead they choose a bootstrapping technique. They argue that the form of the dependence is unlikely to satisfy the conditions of any CLT, even one for dependent data. Other authors assume independence in order to get a CLT they can use. For example, Carrera and Restout [2008, p. 8], who admit their "assumption of independence across individuals is quite strong but essential in order to apply the Lindberg-Levy central limit theorem that permits [us] to derive limiting distributions of tests."

Barbieri et al. [2008] discuss the importance of dependence in financial data. They discuss CLTs and use their discussion to motivate discussion of general test statistics that are robust to dependence and other violations of common CLTs (e.g., infinite variance and non-stationarity). Barbieri et al. [2009] discuss CLTs in finance and deviations from the assumptions of standard CLTs (e.g., time series dependence and time-varying variance). They even go so far as to suggest that inappropriate use of CLTs that are not robust to violations of assumptions may have led to risk-management practices (e.g., use of Value at Risk [VaR]) that failed to account for extreme tail events and indirectly led to the global recession that began in 2007.

Brockett [1983] also discusses misuse of CLTs in risk management. This is, however, an example of the "large deviation" problem (rather than a central limit problem) discussed in Feller [1971, pp. 548–553]. Cummins [1991] provides an excellent explanation of Brockett's work, and Lamm-Tennant et al. [1992] and Powers et al. [1998] both warn the reader about the problem.

Carr and Wu [2003] are unusual in that they deliberately build a model of stock returns that violates the assumptions of a CLT. They do so because they observe patterns in option implied volatility smiles that are inconsistent with the CLT assumptions being satisfied. The assumption they violate is, however, finiteness of second moments rather than independence.

Research interest in dependence in financial data is nothing new. There has been a sustained high level of research into dependence in financial data stretching, for example, from Cowles and Jones [1937] to Fama [1965], to Lo and MacKinlay [1988], to Egan [2008], to Bajgrowicz and Scaillet [2008], to Barbieri et al. [2008, 2009], and beyond.

Given that dependence in financial data is widespread, causes many statistical problems, and is the topic of much research, careful pedagogy in the area of the application of CLTs to dependent data is required.

PEDAGOGICAL GAPS

We have identified two pedagogical gaps in the area of the application of CLTs to dependent data. First, standard doctoral-level econometrics texts do not always make clear the assumptions required for common CLTs, and they may, by their very nature, fail to contain more advanced CLTs. For example, looking at the Lindberg-Levy and Lindberg-Feller CLTs in Greene [2008], it is not at all clear that they do not apply to dependent data [see Theorems D.18A and D.19A in Greene, 2008, pp. 1054–1055]. Only very careful reading of earlier material in the book, combined with considerable inference, reveals the full assumptions of these theorems. The assumptions for these two theorems are, however, clearly stated in more advanced books [see DasGupta, 2008, p. 63; Davidson, 1997, Theorems 23.3 and 23.6; Feller, 1968, p. 244; Feller, 1971, p. 262; and White, 1984 and 2001, Theorems 5.2 and 5.6]. Second, even where the assumptions for the simple CLTs do appear clearly and where the more advanced CLTs for dependent data are present, we have been unable to find any worked example showing the application of the more advanced CLTs to concrete problems. For example, although Hong et al. [2007] use a CLT for dependent data from White [1984], they gloss over the implementation details because theirs is a research paper, not a pedagogical one.

These pedagogical gaps make the area of the application of advanced CLTs to cases of dependent data poorly accessible to many doctoral students. We believe that the best way to address this problem is by providing a worked example using a CLT for dependent data in a simple case. So, in what follows, we derive the asymptotic distribution of the sample variance of a Gaussian AR(1) process using a CLT from White [1984, 2001]. We also derive the asymptotic distribution of the sample mean for the process. This latter derivation does not need a CLT, but the result is needed for the asymptotic distribution of the sample variance.

WORKED EXAMPLE OF A CLT FOR DEPENDENT DATA

We assume that the random variable X_t follows a Gaussian AR(1) process:

$$X_t = \mu + \rho(X_{t-1} - \mu) + \varepsilon_t, \tag{1}$$

where $\varepsilon_t \sim IID\ N(0, \sigma_\varepsilon^2)$, "IID" means independent and identically distributed, and "$N(a,b)$" denotes a Normal distribution with mean a and variance b. The only other assumption we make in the paper is that $|\rho| < 1$ (so that X_t is stationary).

The functional form of (1) is the simplest example of a non-IID data-generating process. By restricting our attention to an AR(1), we minimize the complexity of the dependence in the data while still being able to demonstrate the use of a CLT for dependent data. Our asymptotic results may be derived without our assumption of Gaussian increments [e.g., using theorems in Fuller, 1996, Section 6.3; or Brockwell and Davis, 1991, Section 6.4]. The Gaussian specification of the problem allows, however, for a cleaner pedagogical illustration using an elegant CLT from White [1984, 2001]. It also allows for a cleaner specification of the Monte-Carlo simulation we perform.

The Gaussian AR(1) process X_t is stationary and ergodic by construction (see the proof of Lemma 4 in Appendix A). Stationarity and ergodicity are strictly weaker than the IID assumption of the classical theorems in probability theory (e.g., the Lindberg-Levy and Lindberg-Feller CLTs). Thus, these theorems do not apply. Stationarity and ergodicity are sufficient, however, for us to derive asymptotic results analogous to those available in the case where X_t is IID.

Let $\hat{\mu}$, and $\hat{\sigma}^2$ denote the usual sample mean and variance of the X_t's,

$$\hat{\mu} \equiv \frac{1}{n}\sum_{t=1}^{n} X_t, \ \ and \ \ \hat{\sigma}^2 \equiv \frac{1}{n-1}\sum_{t=1}^{n}(X_t - \hat{\mu})^2. \tag{2}$$

The following two lemmas and theorem give the asymptotic distribution of the sample mean $\hat{\mu}$ of the Gaussian AR(1) process.

Lemma 1 *We have the following exact distributional result for a Gaussian AR(1):*

$$\sqrt{n}(\hat{\mu} - \mu) + \left[\frac{\rho}{1-\rho} \cdot \frac{X_n - X_0}{\sqrt{n}}\right] \sim N\left(0, \frac{\sigma_\varepsilon^2}{(1-\rho)^2}\right). \tag{3}$$

Proof:

See Appendix A.

Lemma 2 *The following probability limit result holds for the second term on the left-hand side of (3):*

$$plim\left[\frac{\rho}{1-\rho} \cdot \frac{X_n - X_0}{\sqrt{n}}\right] = 0. \tag{4}$$

Proof:

See Appendix A.

Theorem 1 *We have the following asymptotic distributional result for the sample mean of a Gaussian AR(1) process:*[1]

$$\sqrt{n}(\hat{\mu} - \mu) \overset{A}{\sim} N\left(0, \frac{\sigma^2(1+\rho)}{1-\rho}\right), \tag{5}$$

where σ^2 is the variance of X_t.

Proof:

Apply Lemma 2 to (3) in Lemma 1 to deduce the asymptotic Normality of $\sqrt{n}(\hat{\mu}-\mu)$. Then use the stationarity of X_t (recall $|\rho|<1$) to replace σ_ε^2 by $\sigma^2(1-\rho^2)$, thus completing the proof. This proof does not require a CLT, but one is needed in the proof of Lemma 4. See van Belle [2002, p. 8] for a related result and DasGupta [2008, p. 127] for a related exercise.

The following two lemmas and theorem give the asymptotic distribution of the sample variance $\hat{\sigma}^2$ of the Gaussian AR(1) process.

Lemma 3 *We may rewrite the term* $\sqrt{n}(\hat{\sigma}^2-\sigma^2)$ *as follows:*

$$\sqrt{n}(\hat{\sigma}^2-\sigma^2)=\sqrt{n}(s^2-\sigma^2)-\sqrt{n}\left(s^2-\left(\frac{n-1}{n}\right)\hat{\sigma}^2\right)+\frac{\hat{\sigma}^2}{\sqrt{n}}, \tag{6}$$

where $s^2 \equiv \frac{1}{n}\sum_{t=1}^{n}(X_t-\mu)^2$.

Proof:

Direct algebraic manipulation and cancellation of terms.

Lemma 4 *The following asymptotic distributional and probability limit results hold for the three terms on the right-hand side of (6):*

$$\sqrt{n}(s^2-\sigma^2)\overset{A}{\sim}N\left(0,\frac{2\sigma^4(1+\rho^2)}{(1-\rho^2)}\right), \tag{7}$$

$$plim\left[\sqrt{n}\left(s^2-\left(\frac{n-1}{n}\right)\hat{\sigma}^2\right)\right]=0,\ \ and \tag{8}$$

$$plim\left[\frac{\hat{\sigma}^2}{\sqrt{n}}\right]=0. \tag{9}$$

Proof:

This is the most difficult derivation. It requires a CLT for dependent data. See Appendix A.

Theorem 2 *We have the following asymptotic distributional result for the sample variance of a Gaussian AR(1) process:*

$$\sqrt{n}(\hat{\sigma}^2 - \sigma^2) \overset{A}{\sim} N\left(0, \frac{2\sigma^4(1+\rho^2)}{(1-\rho^2)}\right). \tag{10}$$

Proof:

Apply the three results in Lemma 4 to the three right-hand side terms, respectively, appearing in Lemma 3, and deduce the result directly.

The asymptotic results for $\hat{\mu}$ in (5) of Theorem 1 and for $\hat{\sigma}^2$ in (10) of Theorem 2 have elegant interpretations. The higher is the degree of positive autocorrelation ρ, the larger is the standard error of both $\hat{\mu}$ and $\hat{\sigma}^2$—higher positive ρ means fewer effectively independent observations of X_t. Similarly, the higher is the degree of negative autocorrelation, then the larger is the standard error of $\hat{\sigma}^2$. We leave the reader with a small challenge: Deduce the qualitative explanation for why larger negative autocorrelation *reduces* the standard error of $\hat{\mu}$.

MONTE-CARLO SIMULATION

We have found that a Monte-Carlo simulation of the process and of the asymptotic distributions of the sample estimators aids doctoral student understanding significantly. We therefore present MATLAB code for a Monte-Carlo simulation, and we plot the resulting theoretical and simulated empirical asymptotic distributions.

In the case of the Gaussian AR(1), doctoral students who incorrectly use CLTs for independent data invariably conclude that the variance on the left-hand side of (10) is $2\sigma^4$ rather than $\frac{2\sigma^4(1+\rho^2)}{(1-\rho^2)}$. You may then ask your students to perform a Monte-Carlo simulation of the Gaussian AR(1) process with $\rho \neq 0$, so that they can demonstrate for themselves that they have statistically significantly underestimated the true standard error.

A portion of our MATLAB code for the Monte-Carlo simulation appears in Appendix B. We choose the values $\mu = 0$, $\rho = 0.90$, and $\sigma_\varepsilon = 0.50$. Figures 1 and 2 compare the realized empirical distribution to the theoretical results for both the asymptotic distribution of $\hat{\sigma}^2$ and the actual large sample distribution of $\hat{\sigma}^2$ (they are scaled versions of each other because we use the same random seed). We do not show the analogous results for $\hat{\mu}$.

Two pedagogical purposes are served by the Monte-Carlo simulation. First, our experience is that when a doctoral student simulates the process, repeatedly collects the asymptotic sample statistics, and then forms a distribution, he or she only then attains a clear concrete notion of what an asymptotic distribution actually is. Second, by comparing the realized asymptotic distribution to the derived theoretical one, the students understand the power of a Monte-Carlo in attempting to confirm or deny the consistency of a difficult analytical result—each of Figures 1 and 2 clearly distinguishes between the competing asymptotic distributions.

[Insert Figures 1 and 2 about here]

CONCLUSIONS

In our experience, finance and economics doctoral students have limited exposure to the use of central limit theorems for dependent data. Given that dependence in financial data is widespread, causes many statistical problems, and is the topic of much research, careful pedagogy in the area of the application of CLTs to dependent data is required. We identify, however, two pedagogical gaps in the area. We fill these gaps by discussing dependence in financial data and dependence assumptions for CLTs and by showing how to use a CLT for dependent data to derive the asymptotic distribution of the sample estimator of the variance of a Gaussian AR(1) process. We also present a Monte-Carlo simulation to aid student understanding of asymptotic distributions and to illustrate the use of a Monte-Carlo in attempting to confirm or deny an analytical result.

ENDNOTES

1. If a sequence b_n of random variables converges in distribution to a random variable Z (often written "$b_n \xrightarrow{d} Z$"), then b_n is said to be *asymptotically distributed* as F_Z, where F_Z is the distribution of Z. This is denoted here by "$b_n \overset{A}{\sim} F_Z$" [as in White, 2001, p. 66].

2. Note that White's "stationarity" is *strict stationarity*. That is, $\{Z_t\}_{t=1}^{\infty}$ and $\{Z_{t-k}\}_{t=1}^{\infty}$ have the same joint distribution for every $k > 0$ [see White, 2001, p. 43; and Davidson 1997, p. 193].

REFERENCES

Anderson, R. M., K. S. Eom, S. B. Hahn, and J. H. Park. “Stock Return Autocorrelation Is Not Spurious,” Working Paper, UC Berkeley and Sunchon National University, (May 2008).

Bajgrowicz, P. and O. Scaillet. “Technical Trading Revisited: Persistence Tests, Transaction Costs, and False Discoveries,” *Swiss Finance Institute Research Paper No. 08-05*, (January 1, 2008). Paper available at SSRN: http://ssrn.com/abstract=1095202.

Barbieri, A., V. Dubikovsky, A. Gladkevich, L. R. Goldberg, and M. Y. Hayes. “Evaluating Risk Forecasts with Central Limits,” (July 9, 2008). Available at SSRN: http://ssrn.com/abstract=1114216.

Barbieri, A., V. Dubikovsky, A. Gladkevich, L. R. Goldberg, and M. Y. Hayes. “Central Limits and Financial Risk,” (March 11, 2009). MSCI Barra Research Paper No. 2009-13. Available at SSRN: http://ssrn.com/abstract=1404089.

Bollerslev, T., T.H. Law, and G. Tauchen, “Risk, Jumps, and Diversification,” (August 16, 2007). CREATES Research Paper 2007-19. Available at SSRN: http://ssrn.com/abstract=1150071.

Brockett, P. L. “On the Misuse of the Central Limit Theorem in Some Risk Calculations,” *The Journal of Risk and Insurance* 50(4) (1983), 727–731.

Brockwell, P. J. and R. A. Davis. *Time Series: Theory and Methods* (New York, 1991), 2nd Edition, Springer.

Campbell, J. Y., A. W. Lo, and A. C. MacKinlay. *The Econometrics of Financial Markets*, (Princeton, 1997), Princeton University Press.

Carr, P. and Liuren Wu, "The Finite Moment Log Stable Process and Option Pricing," *Journal of Finance* 58(2) (2003), 753–777.

Carrera, J. E. and R. Restout. "Long Run Determinants of Real Exchange Rates in Latin America" (April 1, 2008). GATE Working Paper No. 08-11. Available at SSRN: http://ssrn.com/abstract=1127121.

Chordia, T. and B. Swaminathan, "Trading Volume and Cross-Autocorrelations in Stock Returns," *Journal of Finance* 55(2) (2000), 913–935.

Cowles, A. and H. Jones. "Some A Posteriori Probabilities in Stock Market Action," *Econometrica* 5 (1937), 280–294.

Cummins, J. D. "Statistical and Financial Models of Insurance Pricing and the Insurance Firm," *The Journal of Risk and Insurance* 58(2) (June, 1991), 261–302.

DasGupta, A. *Asymptotic Theory of Statistics and Probability*, (New York, 2008), Springer.

Davidson, J. *Stochastic Limit Theory* (New York, 1997), Oxford University Press.

DeBondt, W. and R. Thaler. "Does the Stock Market Overreact?" *Journal of Finance* 40 (1985), 793–805.

DeBondt, W. and R. Thaler. "Further Evidence on Investor Overreaction and Stock Market Seasonality," *Journal of Finance* 42 (1987), 557–582.

Egan, W. J. "Six Decades of Significant Autocorrelation in the U.S. Stock Market" (January 20, 2008). Available at SSRN: http://ssrn.com/abstract=1088861.

Fama, E. F. "The Behavior of Stock Market Prices," *Journal of Business* 38 (1965), 34–105.

Feller, W. *An Introduction to Probability Theory and Its Applications* (New York, 1968), Volume I, 3rd Edition, John Wiley and Sons.

Feller, W. *An Introduction to Probability Theory and Its Applications* (New York, 1971), Volume II, 2nd Edition, John Wiley and Sons.

Fisher, L., "Some New Stock Market Indexes," *Journal of Business* 39 (1966), 191–225.

Fuller, W. A. *Introduction to Statistical Time Series* (New York, 1996), 2nd Edition, John Wiley and Sons.

Frazzini, A., “The Disposition Effect and Underreaction to News,” *Journal of Finance* 61(4), 2017–2046.

Graham, B. and D. Dodd, *Security Analysis: The Classic 1934 Edition*, (New York, 1934), McGraw-Hill.

Greene, W. H. *Econometric Analysis* (Upper Saddle River, 2008), 6th Edition, Prentice Hall.

Hamilton, J. D. *Time Series Analysis* (Princeton, 1994), Princeton University Press.

Hong, Y., J. Tu, and G. Zhou. “Asymmetries in Stock Returns: Statistical Tests and Economic Evaluation,” *Review of Financial Studies* 20(5) (2007), 1547–1581.

Ibragimov, I. A. and Y. V. Linnik *Independent and Stationary Sequences of Random Variables* (The Netherlands, 1971), ed. by J. F. C. Kingman, Wolters-Noordhoff Publishing Groningen.

Jegadeesh, N. and S. Titman. “Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency,” *Journal of Finance* 48 (1993), 65–91.

Lamm-Tennant, J., L. T. Starks, and L. Stokes. “An Empirical Bayes Approach to Estimating Loss Ratios,” *Journal of Risk and Insurance* 59(3) (1992), 426–442.

LeRoy, S. F. “Risk Aversion and the Martingale Property of Stock Prices,” *International Economic Review* 14(2) (1973), pp. 436–446

Lo, A. W. and A. C. MacKinlay. “Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test,” *Review of Financial Studies* 1(1) (1988), 41–66.

Lucas, R. E., “Asset Prices in an Exchange Economy,” *Econometrica* 46(6) (1978), 1429–1445.

Powers, M.R., M. Shubik, and S.T. Yao, “Insurance Market Games: Scale Effects and Public Policy,” *Journal of Econometrics* 67(2) (1998), 109–134.

Roll, R. “A Simple Implicit Measure of the Effective Bid Ask Spread in an Efficient Market,” *Journal of Finance* 39(4) (1984), 1127–1139.

Rosenblatt, M. “Dependence and Asymptotic Dependence for Random Processes,” *Studies in Probability Theory* (Washington, D.C., 1978), Murray Rosenblatt (ed.), Mathematical Association of America, 24–45.

Scaillet, O. and N. Topaloglou, “Testing for Stochastic Dominance Efficiency” (July 2005). FAME Research Paper No. 154. Available at SSRN: http://ssrn.com/abstract=799788.

van Belle, Gerald, *Statistical Rules of Thumb* (New York, 2002), Wiley Series in Probability and Statistics.

White, H. *Asymptotic Theory for Econometricians* (San Diego, 1984), Academic Press.

White, H. *Asymptotic Theory for Econometricians* (San Diego, 2001), Revised 2nd Edition, Academic Press.

APPENDIX A. DERIVATIONS

Proof of Lemma 1: Rewrite the left-hand side of (3) in terms of the residual ε_t (the exact distribution of which is known).

$$\sqrt{n}(\hat{\mu}-\mu)+\left[\frac{\rho}{1-\rho}\cdot\frac{X_n-X_0}{\sqrt{n}}\right]$$

$$=\frac{\sqrt{n}}{(1-\rho)}\left[(1-\rho)(\hat{\mu}-\mu)+\rho\left(\frac{X_n-X_0}{n}\right)\right]$$

$$=\frac{\sqrt{n}}{(1-\rho)}\left[(\hat{\mu}-\mu)-\rho\left[(\hat{\mu}-\mu)-\left(\frac{X_n-X_0}{n}\right)\right]\right]$$

$$=\frac{\sqrt{n}}{(1-\rho)}\left[\frac{1}{n}\sum_{t=1}^{n}(X_t-\mu)-\rho\left[\frac{1}{n}\sum_{t=1}^{n}(X_t-\mu)-\left(\frac{X_n-X_0}{n}\right)\right]\right]$$

$$=\frac{1}{\sqrt{n}(1-\rho)}\left[\sum_{t=1}^{n}(X_t-\mu)-\rho\left[\sum_{t=1}^{n}(X_t-\mu)-(X_n-X_0)\right]\right]$$

$$=\frac{1}{\sqrt{n}(1-\rho)}\left[\sum_{t=1}^{n}(X_t-\mu)-\rho\sum_{t=1}^{n}(X_{t-1}-\mu)\right]$$

$$=\frac{1}{\sqrt{n}(1-\rho)}\sum_{t=1}^{n}\left[(X_t-\mu)-\rho(X_{t-1}-\mu)\right]$$

$$=\frac{1}{\sqrt{n}(1-\rho)}\sum_{t=1}^{n}\varepsilon_t,$$

where the last line uses the definition of ε_t implicit within (1). We may now use

$\varepsilon_t \sim IID\ N(0,\sigma_\varepsilon^2)$ to deduce

$$\frac{1}{\sqrt{n}(1-\rho)}\sum_{t=1}^{n}\varepsilon_t \sim N\left(0,\frac{\sigma_\varepsilon^2}{(1-\rho)^2}\right),$$

thus proving the lemma.

Proof of Lemma 2:

Let " $var(\cdot)$, " " $cov(\cdot,\cdot)$, " and " $corr(\cdot,\cdot)$, " denote the unconditional variance, covariance, and correlation operators, respectively. Let σ^2 denote $var(X_t)$. The term $\rho(X_n - X_0)/[(1-\rho)\sqrt{n}]$ is shown to have variance of order $O(1/n)$ as follows:

$$var\left[\frac{\rho}{1-\rho}\cdot\frac{(X_n - X_0)}{\sqrt{n}}\right] = \frac{1}{n}\left(\frac{\rho}{1-\rho}\right)^2 var(X_n - X_0)$$

$$= \frac{1}{n}\left(\frac{\rho}{1-\rho}\right)^2 \left[var(X_n) + var(X_0) - 2cov(X_n, X_0)\right]$$

$$= \frac{1}{n}\left(\frac{\rho}{1-\rho}\right)^2 \left[\sigma^2 + \sigma^2 - 2corr(X_n, X_0)\sigma\sigma\right]$$

$$\leq \frac{4\sigma^2}{n}\left(\frac{\rho}{1-\rho}\right)^2. \tag{11}$$

This derivation assumes $|\rho| < 1$ (so that stationarity of X_t gives $var(X_n) = var(X_0) = \sigma^2$). We also use $corr(X_n, X_0) \geq -1$ at the last step.

Tchebychev's Inequality [Greene 2008, p. 1040] says that for random variable V and small $\delta > 0$,

$$P(|V - E(V)| > \delta) \leq \frac{var(V)}{\delta^2}.$$

We may apply Tchebychev's Inequality to $V_n \equiv \rho(X_n - X_0)/[(1-\rho)\sqrt{n}]$, and use (11) to find

$$P\left(\left|\frac{\rho}{1-\rho}\cdot\frac{(X_n - X_0)}{\sqrt{n}}\right| > \delta\right) \leq \frac{var(V_n)}{\delta^2}$$

$$\leq \frac{4\sigma^2}{n\delta^2}\left(\frac{\rho}{1-\rho}\right)^2.$$

Thus, for any $\delta > 0$, we have $\lim_{n\to\infty} P(|V_n| > \delta) = 0$. That is, $plim V_n = 0$, thus proving the lemma.

Proof of Lemma 4: We demonstrate each of Equations (7), (8), and (9) in turn. We begin with the proof of the asymptotic result in (7):

$$\sqrt{n}(s^2 - \sigma^2) \overset{A}{\sim} N\left(0, \frac{2\sigma^4(1+\rho^2)}{(1-\rho^2)}\right),$$

where $s^2 \equiv \frac{1}{n}\sum_{t=1}^{n}(X_t - \mu)^2$, and $\sigma^2 = var(X_t)$. To derive this result, we apply the following CLT for non-IID data adapted directly from White [1984].

Theorem [from White 1984, Theorem 5.15, p. 118]

Let $\Im_t$ be the sigma-algebra generated by the entire current and past history of a stochastic variable Z_t; let $\Re_{t,j}$ be the revision made in forecasting Z_t when information becomes available at time $t-j$, that is, $\Re_{t,j} \equiv E(Z_t \mid \Im_{t-j}) - E(Z_t \mid \Im_{t-j-1})$; let $\bar{Z}_n$ denote the sample mean of $Z_1, \ldots, Z_n$; and let $\bar{\sigma}_n^2 \equiv var(\sqrt{n}\bar{Z}_n)$. Then, if the sequence $\{Z_t\}$ satisfies the following conditions: 1. $\{Z_t\}$ is stationary;[2] 2. $\{Z_t\}$ is ergodic; 3. $E(Z_t^2) < \infty$; 4. $E(Z_0 \mid \Im_{-m}) \overset{q.m.}{\to} 0$ as $m \to \infty$; and 5. $\sum_{j=0}^{\infty}\left[var\left(\Re_{0,j}\right)\right]^{1/2} < \infty$, we obtain the results $\bar{\sigma}_n^2 \to \bar{\sigma}^2$, as $n \to \infty$, and if $\bar{\sigma}^2 > 0$, then $\frac{\sqrt{n}\bar{Z}_n}{\bar{\sigma}} \overset{A}{\sim} N(0,1)$.

We apply the theorem to $Z_t \equiv (X_t - \mu)^2 - \sigma^2$. With this definition of Z_t, we obtain $\bar{Z}_n = (1/n)\sum_{t=1}^{n} Z_t = s^2 - \sigma^2$, and, thus, $\sqrt{n}\bar{Z}_n = \sqrt{n}(s^2 - \sigma^2)$. However, before we can apply the theorem, we must check that its five conditions are satisfied, and we must calculate $\lim_{n\to\infty} \bar{\sigma}_n^2 = \lim_{n\to\infty} var(\sqrt{n}\bar{Z}_n)$. We begin by checking the five conditions.

Condition 1: We have assumed $|\rho|<1$. Thus, our Gaussian AR(1) process X_t is stationary. Stationarity of X_t yields stationarity of Z_t immediately (by definition of Z_t).

Condition 2: White [2001, p. 48] uses Ibragimov and Linnik [1971, pp. 312–313] to deduce that a Gaussian AR(1) with $|\rho|<1$ is strong mixing. White [2001, p. 48] then uses Rosenblatt [1978] to state that strong mixing plus stationarity (recall $|\rho|<1$) implies ergodicity. It follows that X_t is ergodic. This yields ergodicity of Z_t immediately (by definition of Z_t).

Condition 3: We note first that since ε_t is Gaussian, then so too is X_t [Hamilton 1994, p. 118]. It is well known that if $X_t \sim N(\mu,\sigma^2)$, then $E[(X_t-\mu)^4] = 3\sigma^4$. It follows that

$$E(Z_t^2) = E[((X_t-\mu)^2-\sigma^2)^2]$$

$$= E[(X_t-\mu)^4 - 2\sigma^2(X_t-\mu)^2+\sigma^4]$$

$$= 3\sigma^4 - 2\sigma^4 + \sigma^4 = 2\sigma^4 < \infty. \qquad (12)$$

Condition 4: To show that $E(Z_0 \mid \Im_{-m}) \xrightarrow{q.m.} 0$ as $m\to\infty$, we must show that $E(Z_t \mid \Im_{t-m}) \xrightarrow{q.m.} 0$ as $m\to\infty$ in the special case $t=0$. In fact, we can prove convergence in quadratic mean for *any* t if we can show $E([E(Z_t \mid \Im_{t-m})]^2) \to 0$ as $m\to\infty$ [see White, 1984, p. 117]. To derive $E(Z_t \mid \Im_{t-m})$, we first consider the term $Z_t + \sigma^2 = (X_t-\mu)^2$ as follows:

$$X_t - \mu = \rho(X_{t-1} - \mu) + \varepsilon_t$$

$$\vdots$$

$$= \rho^m (X_{t-m} - \mu) + \sum_{k=0}^{m-1} \rho^k \varepsilon_{t-k}. \tag{13}$$

With $Z_t + \sigma^2 = (X_t - \mu)^2$, it follows from (13) that

$$E(Z_t + \sigma^2 \mid \Im_{t-m}) = E\left[\left(\rho^m (X_{t-m} - \mu) + \sum_{k=0}^{m-1} \rho^k \varepsilon_{t-k}\right)^2 \middle| \Im_{t-m}\right]$$

$$= \rho^{2m} (X_{t-m} - \mu)^2 + 0 + \sum_{k=0}^{m-1} \rho^{2k} \sigma_\varepsilon^2$$

$$= \rho^{2m} (X_{t-m} - \mu)^2 + \left(\frac{1-\rho^{2m}}{1-\rho^2}\right) \sigma_\varepsilon^2$$

$$= \rho^{2m} (X_{t-m} - \mu)^2 + \left(\frac{1-\rho^{2m}}{1-\rho^2}\right) [\sigma^2 (1-\rho^2)]$$

$$= \rho^{2m} (X_{t-m} - \mu)^2 + \sigma^2 (1-\rho^{2m}). \tag{14}$$

If we now cancel σ^2 from both sides of (14), we find

$$E(Z_t \mid \Im_{t-m}) = \rho^{2m} [(X_{t-m} - \mu)^2 - \sigma^2] = \rho^{2m} Z_{t-m}. \tag{15}$$

It follows that $E([E(Z_t \mid \Im_{t-m})]^2) = E([\rho^{2m} Z_{t-m}]^2) = \rho^{4m} E(Z_{t-m}^2) = \rho^{4m} 2\sigma^4$ (using (12) and stationarity of Z_t). With $|\rho| < 1$, we deduce that $E([E(Z_t \mid \Im_{t-m})]^2) \to 0$ as $m \to \infty$, and, thus, that $E(Z_t \mid \Im_{t-m}) \xrightarrow{q.m.} 0$ as $m \to \infty$ [using White, 1984, p. 117], as required.

Condition 5: Applying (15) to the definition of $\Re_{t,j}$ yields

$$\Re_{t,j} \equiv E(Z_t \mid \Im_{t-j}) - E(Z_t \mid \Im_{t-j-1})$$

$$= \rho^{2j} Z_{t-j} - \rho^{2(j+1)} Z_{t-(j+1)}. \tag{16}$$

By definition, $E(Z_t) = 0$, so $E(\Re_{t,j}) = 0$, and, thus, $var(\Re_{t,j}) = E(\Re_{t,j}^2)$. Manipulating (16), we get

$$var(\Re_{t,j}) = E(\Re_{t,j}^2)$$

$$= E([\rho^{2j} Z_{t-j} - \rho^{2(j+1)} Z_{t-(j+1)}]^2)$$

$$= \left(\rho^{4j} + \rho^{4(j+1)}\right) 2\sigma^4 - 2\rho^{4j+2} E(Z_{t-j} Z_{t-(j+1)}),$$

$$= \left(\rho^{4j} + \rho^{4(j+1)}\right) 2\sigma^4 - 2\rho^{4j+2} E(Z_t Z_{t-1}), \tag{17}$$

where we used (12) and the fact that $E(Z_{t-j}) = E(Z_{t-(j+1)}) = 0$. We also used stationarity of Z_t to rewrite $E(Z_{t-j} Z_{t-(j+1)})$ as $E(Z_t Z_{t-1})$.

The term $E(Z_t Z_{t-1})$ in (17) may be expanded as follows:

$$E(Z_t Z_{t-1}) = E[((X_t - \mu)^2 - \sigma^2)((X_{t-1} - \mu)^2 - \sigma^2)]$$

$$= E[(X_t - \mu)^2 (X_{t-1} - \mu)^2] - \sigma^4.$$

Plugging this expression for $E(Z_t Z_{t-1})$ into (17) gives

$$var(\Re_{t,j}) = \left(\rho^{4j} + \rho^{4(j+1)}\right) 2\sigma^4$$

$$-2\rho^{4j+2}\left(E[(X_t - \mu)^2 (X_{t-1} - \mu)^2] - \sigma^4\right)$$

$$= \left(\rho^{4j} + \rho^{4(j+1)}\right) 2\sigma^4 - 2\rho^{4j+2}\left(E(Y_t^2 Y_{t-1}^2) - \sigma^4\right), \tag{18}$$

where $Y_t \equiv (X_t - \mu)$. The term $E(Y_t^2 Y_{t-1}^2)$ is a special case of a more general term $E(Y_t^2 Y_{t-d}^2)$, which we now evaluate (we need the general term later in the proof). From the definition of the Gaussian AR(1) (1) and from (13), we deduce that $Y_t = \rho^d Y_{t-d} + \sum_{k=0}^{d-1} \rho^k \varepsilon_{t-k}$ and that Y_t is Gaussian with zero-mean. It follows that

$$E(Y_t^2 Y_{t-d}^2) = E\left[\left(\rho^d Y_{t-d} + \sum_{k=0}^{d-1}\rho^k \varepsilon_{t-k}\right)^2 Y_{t-d}^2\right]$$

$$= \rho^{2d} E(Y_{t-d}^4) + 2\rho^d E\left(\sum_{k=0}^{d-1}\rho^k \varepsilon_{t-k}\right)^2 E(Y_{t-d}^3)$$

$$+ E\left[\left(\sum_{k=0}^{d-1}\rho^k \varepsilon_{t-k}\right)^2\right] E(Y_{t-d}^2)$$

$$= 3\rho^{2d}\sigma^4 + 0 + \sigma^2 \sum_{k=0}^{d-1}\rho^{2k}\sigma_\varepsilon^2$$

$$= 3\rho^{2d}\sigma^4 + \sigma^2\left(\frac{1-\rho^{2d}}{1-\rho^2}\right)[\sigma^2(1-\rho^2)]$$

$$= \sigma^4(1+2\rho^{2d}), \tag{19}$$

where we used independence of Y_{t-d} and ε_{t-k} for $k < d$ to separate expectations in the cross-product term. We also used the mean-zero Normality of Y_{t-d} to write $E(Y_{t-d}^3) = 0$, and $E(Y_{t-d}^4) = 3\sigma^4$. If we now set $d = 1$ in (19) and plug this into (18), we obtain

$$var(\Re_{t,j}) = \left(\rho^{4j} + \rho^{4(j+1)}\right)2\sigma^4 - 2\rho^{4j+2}\left(\sigma^4(1+2\rho^2) - \sigma^4\right)$$

$$= 2\sigma^4(1-\rho^4)\rho^{4j}. \tag{20}$$

Thus, $[var(\Re_{t,j})]^{1/2} = \sqrt{E(\Re_{t,j}^2)} = \sqrt{2\sigma^4(1-\rho^4)}\,\rho^{2j}.$ It follows that

$$\sum_{j=0}^{\infty}\left(var\left(\Re_{t,j}\right)\right)^{1/2} = \sqrt{2\sigma^4(1-\rho^4)}\sum_{j=0}^{\infty}\rho^{2j}$$

$$= \frac{\sqrt{2\sigma^4(1-\rho^4)}}{1-\rho^2}$$

$$= \frac{\sqrt{2\sigma^4(1+\rho^2)(1-\rho^2)}}{1-\rho^2}$$

$$= \sqrt{\frac{2\sigma^4(1+\rho^2)}{1-\rho^2}} < \infty.$$

This latter result holds in the special case $t = 0$, so the fifth and final prerequisite for applying White's Theorem to Z_t is satisfied.

We must now find $\bar{\sigma}^2 \equiv \lim_{n\to\infty} \bar{\sigma}_n^2 = \lim_{n\to\infty} var(\sqrt{n}\bar{Z}_n)$. Recall that we have $Z_t = (X_t - \mu)^2 - \sigma^2$, so that $\sqrt{n}\bar{Z}_n = \sqrt{n}(s^2 - \sigma^2)$, where $s^2 \equiv (1/n)\sum_{t=1}^n (X_t - \mu)^2 = (1/n)\sum_{t=1}^n Y_t^2$. With σ^2 a constant, we know that $var(\sqrt{n}\bar{Z}_n) = var(\sqrt{n}s^2)$. It is easier to work with $var(ns^2)$, so we do that and then adjust the result.

$$var(ns^2) = var\left(\sum_{t=1}^n Y_t^2\right)$$

$$= E\left[\left(\sum_{t=1}^n Y_t^2\right)^2\right] - \left[E\left(\sum_{t=1}^n Y_t^2\right)\right]^2$$

$$= E\left[\sum_{t=1}^n \sum_{s=1}^n Y_t^2 Y_s^2\right] - (n\sigma^2)^2$$

$$= \left[2\sum_{t=2}^n \sum_{d=1}^{t-1} E\left(Y_t^2 Y_{t-d}^2\right) + nE(Y_t^4)\right] - (n\sigma^2)^2$$

$$= \left[2\sigma^4 \sum_{t=2}^n \sum_{d=1}^{t-1} \left(1 + 2\rho^{2d}\right) + 3n\sigma^4\right] - (n\sigma^2)^2, \qquad (21)$$

where we used (19) to replace $E\left(Y_t^2 Y_{t-d}^2\right)$. If we divide (21) by σ^4 and combine the final two terms, we get

$$\frac{var(ns^2)}{\sigma^4} = 2\sum_{t=2}^n \sum_{d=1}^{t-1} \left(1 + 2\rho^{2d}\right) + n(3-n)$$

$$= 2\sum_{t=2}^n \left[(t-1) + 2\rho^2 \left(\frac{1-\rho^{2(t-1)}}{1-\rho^2}\right)\right] + n(3-n)$$

It is easily shown that $2\sum_{t=2}^{n}(t-1) = -n(3-n)+2n$, so we get some cancellation as follows:

$$\frac{var(ns^2)}{\sigma^4} = \frac{4\rho^2}{1-\rho^2}\left[(n-1)-\sum_{t=2}^{n}\rho^{2(t-1)}\right]+2n$$

$$= \frac{4\rho^2(n-1)+2n(1-\rho^2)}{1-\rho^2} - \frac{4\rho^2}{1-\rho^2}\cdot\rho^2\left(\frac{1-\rho^{2(n-1)}}{1-\rho^2}\right)$$

$$= \frac{4\rho^2 n+2n-2n\rho^2}{1-\rho^2} - \frac{4\rho^2}{1-\rho^2}\left[1+\rho^2\left(\frac{1-\rho^{2(n-1)}}{1-\rho^2}\right)\right]$$

$$= \frac{2n(1+\rho^2)}{(1-\rho^2)} - \frac{4\rho^2(1-\rho^{2n})}{(1-\rho^2)^2}. \qquad (22)$$

It follows immediately that $var(\sqrt{n}s^2) \to \frac{2\sigma^4(1+\rho^2)}{(1-\rho^2)}$ as $n \to \infty$. Using this result in the last part of White's theorem yields

$$\sqrt{n}(s^2-\sigma^2) \overset{A}{\sim} N\left(0, \frac{2\sigma^4(1+\rho^2)}{(1-\rho^2)}\right),$$

thus proving (7)—the first of the three parts of Lemma 4.

To demonstrate (8)—the second of the three parts of Lemma 4—we need the probability limit of $\sqrt{n}\left(s^2-((n-1)/n)\hat{\sigma}^2\right)$. Direct algebraic manipulation yields

$$\left[\sqrt{n}\left(s^2-\left(\frac{n-1}{n}\right)\hat{\sigma}^2\right)\right] = \sqrt{n}(\hat{\mu}-\mu)^2$$

$$= \frac{1}{\sqrt{n}}\left[\frac{\sigma^2(1+\rho)}{1-\rho}\right]\cdot\left[\frac{\sqrt{n}(\hat{\mu}-\mu)}{\sqrt{\frac{\sigma^2(1+\rho)}{1-\rho}}}\right]^2$$

$$= \frac{1}{\sqrt{n}}\left[\frac{\sigma^2(1+\rho)}{1-\rho}\right]Q_n^2, \qquad (23)$$

where $Q_n \equiv \left[\frac{\sqrt{n}(\hat{\mu}-\mu)}{\sqrt{\frac{\sigma^2(1+\rho)}{1-\rho}}} \right]$ is asymptotically standard Normal (a consequence of Theorem 1). We may now apply a result analogous to Slutsky's Theorem for probability limits [see Greene, 2008, p. 1045] to deduce that $Q_n^2 \overset{A}{\sim} \chi_1^2$ (that is, Q_n^2 is asymptotically chi-square with one degree of freedom). Thus, Q_n^2 is of bounded variance. It follows that one application of Tchebychev's Inequality to (23) produces the result:

$$plim\left[\sqrt{n}\left(s^2 - \left(\frac{n-1}{n}\right)\hat{\sigma}^2\right)\right] = 0,$$

thus proving (8)—the second of the three parts of Lemma 4.

To demonstrate (9)—the third and final part of Lemma 4—we need the probability limit of $(\hat{\sigma}^2/\sqrt{n})$. Algebraic manipulation gives

$$\hat{\sigma}^2 = \left(\frac{n}{n-1}\right)s^2 - \left(\frac{n}{n-1}\right)(\hat{\mu}-\mu)^2. \quad (24)$$

The variance of s^2 goes to zero as $n \to \infty$ (a consequence of (22)). The variance of $(\hat{\mu}-\mu)^2$ goes to zero as $n \to \infty$ (a consequence of $Q_n^2 \overset{A}{\sim} \chi_1^2$, from above). In (24), the coefficients $n/(n-1) \to 1$ as $n \to \infty$. It follows that $var(\hat{\sigma}^2) \to 0$ with n. An application of Tchebychev's Inequality yields immediately

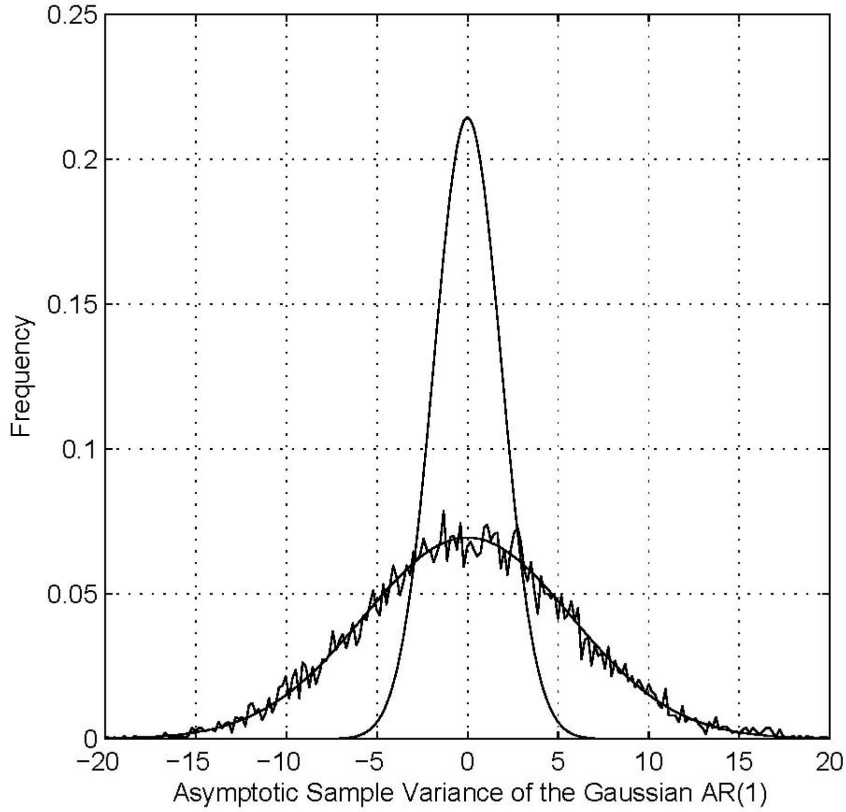
$$plim\left[\frac{\hat{\sigma}^2}{\sqrt{n}}\right] = 0,$$

thus proving the third and final part of Lemma 4.

APPENDIX B. MATLAB MONTE-CARLO CODE

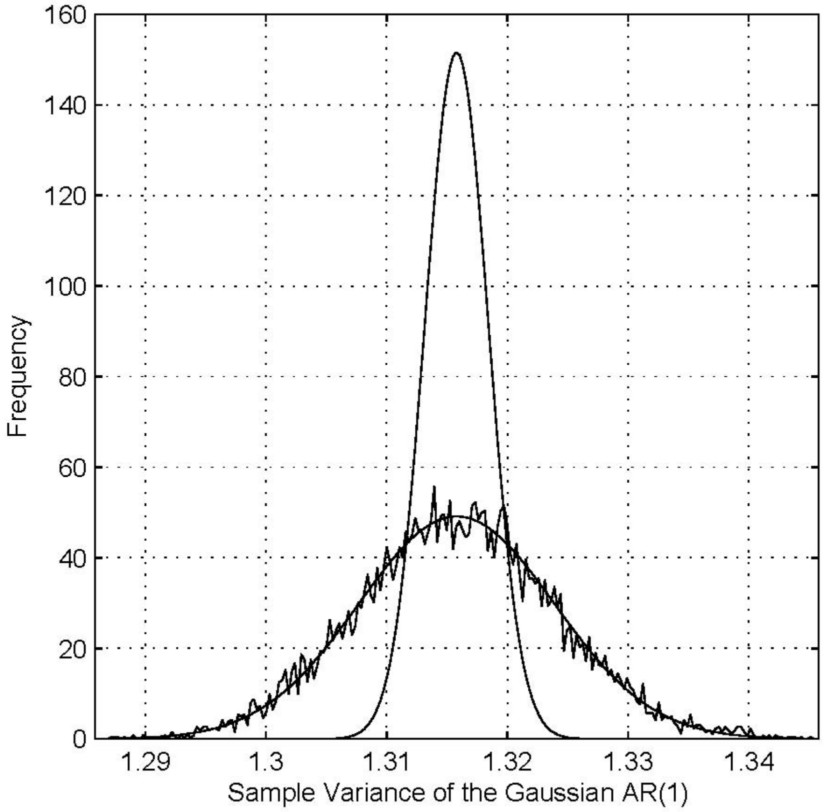
```
clear;
rho=0.90;sigmae=0.50;mu=0;sigma=sigmae/sqrt(1-rho^2);
N=500000;NUMBREPS=10000; rseed=20081103; randn('seed',rseed);
collect=[ ];
for J=1:NUMBREPS
Y=[]; epsilon=randn(N,1); xpf=epsilon*sigmae;
bpf=1; apf=[1 -rho]; Y=filter(bpf,apf,xpf);
collect=[collect' [mean(Y) var(Y)]']';
end
asymeanv=0; asyvarv=2*(sigma^4)*(1+rho^2)/(1-rho^2);
asymeanv1=0; asyvarv1=2*(sigma^4); v=sqrt(N)*(collect(:,2)-sigma^2);
hpdf=[];mynormpdf=[];[M,X]=hist(v,250);M=M';X=X';dx=min(diff(X));
hpdf=M/(sum(M)*dx);
mynormpdf=(1/(sqrt(2*pi)*sqrt(asyvarv))).*exp(
-0.5*((X-asymeanv)/sqrt(asyvarv)).^ 2);
mynormpdf1=(1/(sqrt(2*pi)*sqrt(asyvarv1))).*exp(
-0.5*((X-asymeanv1)/sqrt(asyvarv1)).^2);
plot(X,[hpdf mynormpdf mynormpdf1],'k')
xlabel('Asymptotic Sample Variance of the Gaussian AR(1)');
ylabel('Frequency');
```

Figure 1. Histogram of Simulated Empirical PDF of $\sqrt{n}(\hat{\sigma}^2 - \sigma^2)$



We use MATLAB to simulate a time series of 500,000 observations of the Gaussian AR(1) using $\rho = 0.90$, $\sigma_\varepsilon = 0.50$, and $\mu = 0$. We then record the sample variance $\hat{\sigma}^2$ of the process. We repeat this 10,000 times and plot (the uneven line) the realized density of $\sqrt{n}(\hat{\sigma}^2 - \sigma^2)$. We overlay on the plot the correct theoretical density $N\left(0, \frac{2\sigma^4(1+\rho^2)}{(1-\rho^2)}\right)$ and the most common incorrect student-derived theoretical density $N(0, 2\sigma^4)$. The correct density is the one close to the empirical density; the incorrect density is more peaked.

Figure 2. Histogram of Simulated Empirical PDF of $\hat{\sigma}^2$



We use MATLAB to simulate a time series of 500,000 observations of the Gaussian AR(1) using $\rho = 0.90$, $\sigma_\varepsilon = 0.50$, and $\mu = 0$. We then record the sample variance $\hat{\sigma}^2$ of the process. We repeat this 10,000 times and plot (the uneven line) the realized density of $\hat{\sigma}^2$. We overlay on the plot the correct theoretical density $N\left(\sigma^2, \frac{2\sigma^4(1+\rho^2)}{n(1-\rho^2)}\right)$ and the most common incorrect student-derived theoretical density $N\left(\sigma^2, \frac{2\sigma^4}{n}\right)$. The correct density is the one close to the empirical density; the incorrect density is more peaked.