Advances and Applications in Statistics
Volume 18, Number 1, 2010, Pages 57-72
Published Online: November 23, 2010
This paper is available online at http://pphmj.com/journals/adas.htm

MOMENTS OF AR(2)-MODEL PARAMETER ESTIMATORS

JAN VRBIK

Department of Mathematics
Brock University
St. Catharines, L2S 3A1
Canada

Abstract

We describe a simple technique for computing the first few moments of estimators of the α_1 and α_2 parameters of the AR(2) model. This enables us to find a good approximation to the joint distribution of the two estimators, which represents a major improvement over the Central Limit Theorem, and can be successfully used even when sample size is relatively small. Furthermore, we can also construct a transformation of the two estimators which eliminates their skewness and thus makes the corresponding approximate distribution simpler and more accurate.

1. Introduction

The AR(2) (also known as Yule [1]) model consists of a stationary sequence of autocorrelated random variables, such that

$$X_i - \mu = \alpha_1(X_{i-1} - \mu) + \alpha_2(X_{i-2} - \mu) + \varepsilon_i, \tag{1}$$

where the ε_i's are independent, normally distributed random variables with the mean of zero and the standard deviation equal to σ. This implies that X_i's are normally distributed, with the mean of μ and the variance of

2010 Mathematics Subject Classification: 62E20.

Keywords and phrases: AR(2) model, autocorrelation, parameter estimators, sampling distribution, normal approximation.

Received August 10, 2010

$$V \equiv \frac{(1-\alpha_2)\sigma^2}{(1+\alpha_1-\alpha_2)(1+\alpha_2)(1-\alpha_1-\alpha_2)}. \tag{2}$$

The correlation matrix of n consecutive X_i's is

$$\mathbb{C}_0 = \begin{bmatrix} \rho_0 & \rho_1 & \rho_2 & \cdots & \rho_{n-1} \\ \rho_1 & \rho_0 & \rho_1 & \cdots & \rho_{n-2} \\ \rho_2 & \rho_1 & \rho_0 & \cdots & \rho_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{n-1} & \rho_{n-2} & \rho_{n-3} & \cdots & \rho_0 \end{bmatrix}, \tag{3}$$

where $\rho_0 = 1$, $\rho_1 = \dfrac{\alpha_1}{1-\alpha_2}$, and $\rho_k = \alpha_1\rho_{k-1} + \alpha_2\rho_{k-2}$ for $k > 1$.

$\mathbb{C}_0$ has a relatively simple band-matrix inverse equal to

$$\mathbb{A}_0 = \frac{(1-\alpha_2)}{(1+\alpha_1-\alpha_2)(1+\alpha_2)(1-\alpha_1-\alpha_2)} \times$$

$$\begin{bmatrix} 1 & -\alpha_1 & -\alpha_2 & 0 & \cdots & 0 & 0 & 0 \\ -\alpha_1 & 1+\alpha_1^2 & -\alpha_1+\alpha_1\alpha_2 & -\alpha_2 & \cdots & 0 & 0 & 0 \\ -\alpha_2 & -\alpha_1+\alpha_1\alpha_2 & 1+\alpha_1^2+\alpha_2^2 & -\alpha_1+\alpha_1\alpha_2 & \ddots & 0 & 0 & 0 \\ 0 & -\alpha_2 & -\alpha_1+\alpha_1\alpha_2 & 1+\alpha_1^2+\alpha_2^2 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & -\alpha_1+\alpha_1\alpha_2 & -\alpha_2 & 0 \\ 0 & 0 & 0 & 0 & \ddots & 1+\alpha_1^2+\alpha_2^2 & -\alpha_1+\alpha_1\alpha_2 & -\alpha_2 \\ 0 & 0 & 0 & 0 & \ddots & -\alpha_1+\alpha_1\alpha_2 & 1+\alpha_1^2 & -\alpha_1 \\ 0 & 0 & 0 & 0 & \cdots & -\alpha_2 & -\alpha_1 & 1 \end{bmatrix} \tag{4}$$

whose determinant is

$$D \equiv \frac{(1-\alpha_2)^n}{(1+\alpha_1-\alpha_2)^{n-1}(1+\alpha_2)^{n-2}(1-\alpha_1-\alpha_2)^{n-1}}. \tag{5}$$

The corresponding multivariate probability density function of the standardized random variables $Y_i \equiv \dfrac{X_i - \mu}{\sqrt{V}}$ is thus

$$f(\mathbf{y}) = \frac{\exp\left(-\dfrac{\mathbf{y}^T\mathbb{A}_0\mathbf{y}}{2}\right)}{(2\pi)^{\frac{n}{2}}\sqrt{D}}. \tag{6}$$

To estimate the α_1 and α_2 parameters, we must first compute the (first and second) sample serial correlation coefficients, namely,

$$\hat{\rho}_1 = \frac{\dfrac{\sum_{i=2}^{n} X_i X_{i-1}}{n-1} - \dfrac{\sum_{i=2}^{n} X_i}{n-1} \cdot \dfrac{\sum_{i=2}^{n} X_{i-1}}{n-1}}{\dfrac{\sum_{i=1}^{n} X_i^2}{n} - \left(\dfrac{\sum_{i=1}^{n} X_i}{n}\right)^2}$$

$$\equiv \frac{\dfrac{\sum_{i=2}^{n} Y_i Y_{i-1}}{n-1} - \dfrac{\sum_{i=2}^{n} Y_i}{n-1} \cdot \dfrac{\sum_{i=2}^{n} Y_{i-1}}{n-1}}{\dfrac{\sum_{i=1}^{n} Y_i^2}{n} - \left(\dfrac{\sum_{i=1}^{n} Y_i}{n}\right)^2} \tag{7}$$

and

$$\hat{\rho}_2 = \frac{\dfrac{\sum_{i=3}^{n} X_i X_{i-2}}{n-2} - \dfrac{\sum_{i=3}^{n} X_i}{n-2} \cdot \dfrac{\sum_{i=3}^{n} X_{i-2}}{n-2}}{\dfrac{\sum_{i=1}^{n} X_i^2}{n} - \left(\dfrac{\sum_{i=1}^{n} X_i}{n}\right)^2}$$

$$\equiv \frac{\dfrac{\sum_{i=3}^{n} Y_i Y_{i-2}}{n-2} - \dfrac{\sum_{i=3}^{n} Y_i}{n-2} \cdot \dfrac{\sum_{i=3}^{n} Y_{i-2}}{n-2}}{\dfrac{\sum_{i=1}^{n} Y_i^2}{n} - \left(\dfrac{\sum_{i=1}^{n} Y_i}{n}\right)^2}. \tag{8}$$

Based on these, the usual way of estimating α_1 and α_2 is to take

$$\hat{\alpha}_1 = \hat{\rho}_1 \frac{1 - \hat{\rho}_2}{1 - \hat{\rho}_1^2} \tag{9}$$

and

$$\hat{\alpha}_2 = \frac{\hat{\rho}_2 - \hat{\rho}_1^2}{1 - \hat{\rho}_1^2}. \tag{10}$$

Note that we can design better estimators (ideally, their *range* should be $-1 < \alpha_2 < 1$ and $\alpha_2 - 1 < \alpha_1 < 1 - \alpha_2$), but our choice is adequate for the purpose of this article. The technique described below can be applied with equal ease to more complicated (including maximum-likelihood) estimators, which were investigated, in the context of the AR(1) model, by an earlier study [2].

The distribution of the two estimators is, in the $n \to \infty$ limit, bivariate Normal, whose moments (elements of the variance-covariance matrix in particular) can be easily computed, when expanded in powers of $\frac{1}{n}$. The purpose of this article is to first explain the details of such a computation, and then extend it to higher moments of the sampling distribution, with the ultimate objective of constructing $\frac{1}{\sqrt{n}}$ and $\frac{1}{n}$-accurate corrections to the Normal approximation. Finally, we demonstrate empirically the substantial improvement achieved by these corrections.

2. Moment Generating Function

The $\hat{\alpha}_1$ and $\hat{\alpha}_2$ estimators are functions of eight basic sample statistics, namely

$$U_1 \equiv \sum_{i=1}^{n} Y_i^2,$$

$$U_2 \equiv \sum_{i=2}^{n} Y_i Y_{i-1},$$

$$U_3 \equiv \sum_{i=3}^{n} Y_i Y_{i-2},$$

$$U_4 \equiv \sum_{i=1}^{n} Y_i,$$

$$U_5 = Y_1,$$

$$U_6 = Y_n,$$

$$U_7 = Y_1 + Y_2,$$

$$U_8 = Y_{n-1} + Y_n. \tag{11}$$

It is easy to construct the corresponding joint moment generating function:

$$\begin{aligned} M(\mathbf{t}) &= \int \cdots \int f(\mathbf{y}) \exp\left(\sum_{i=1}^{8} t_i U_i\right) dy_1 \cdots dy_n \\ &= \sqrt{\frac{\det(\mathbb{A}_0)}{\det(\mathbb{A})}} \cdot \exp\left(\frac{\mathbf{t}^T \mathbb{A}^{-1} \mathbf{t}}{2}\right), \end{aligned} \tag{12}$$

where

$$\mathbf{t} \equiv \begin{bmatrix} t_4 + t_5 + t_7 \\ t_4 + t_5 \\ t_4 \\ t_4 \\ \vdots \\ t_4 \\ t_4 + t_6 \\ t_4 + t_6 + t_8 \end{bmatrix} \tag{13}$$

and

$$\mathbb{A} \equiv \mathbb{A}_0 - \begin{bmatrix} 2t_1 & t_2 & t_3 & 0 & \cdots & \cdots & 0 & 0 \\ t_2 & 2t_1 & t_2 & t_3 & \ddots & \cdots & 0 & 0 \\ t_3 & t_2 & 2t_1 & t_2 & \ddots & \ddots & 0 & 0 \\ 0 & t_3 & t_2 & 2t_1 & \ddots & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & 2t_1 & t_2 & t_3 \\ 0 & 0 & 0 & 0 & \ddots & t_2 & 2t_1 & t_2 \\ 0 & 0 & 0 & 0 & \cdots & t_3 & t_2 & 2t_1 \end{bmatrix}. \tag{14}$$

The last line of (12) follows from

$$\mathbf{y}^T \mathbb{A} \mathbf{y} - 2\mathbf{y}^T \mathbf{t} = (\mathbf{y} - \mathbf{y}_0)^T \mathbb{A} (\mathbf{y} - \mathbf{y}_0) + \mathbf{y}_0^T \mathbb{A} \mathbf{y}_0, \tag{15}$$

where $\mathbf{y}_0 \equiv \mathbb{A}^{-1} \mathbf{t}$.

For any specific (i.e., numerical) value of n, the moment generating function can be easily (Taylor-) expanded in powers of $t_1, t_2, \ldots, t_8$, provided we make the following simplifications to reduce the complexity of the answer:

- Terms involving large powers of α_1 and α_2 (those which would disappear in the $n \to \infty$ limit) are discarded.
- Since we are interested in $O\left(\frac{1}{n^k}\right)$-accurate results (where k is usually quite small - in our case, it does not exceed 3), we can also discard terms whose combined t_i-power is greater than $2k$.
- The $\mathbb{A}^{-1}$ matrix is efficiently constructed by
$$\mathbb{C}_0 + \mathbb{C}_0(\mathbb{A}_0 - \mathbb{A})\mathbb{C}_0 + \mathbb{C}_0(\mathbb{A}_0 - \mathbb{A})\mathbb{C}_0(\mathbb{A}_0 - \mathbb{A})\mathbb{C}_0 + \cdots + [\mathbb{C}_0(\mathbb{A}_0 - \mathbb{A})]^{2k}\mathbb{C}_0. \tag{16}$$
- Finally, n itself does not have to be very large (all our results were obtained with $n \le 56$).

3. Computing Moments

First we replace, in each of the two expressions (7) and (8), every sample mean $\overline{Q}$ by
$$\varepsilon(\overline{Q} - \mu_Q) + \mu_Q, \tag{17}$$
where μ_Q is the corresponding expected value. The expression for $\hat{\rho}_1$ thus reads
$$\frac{\varepsilon\left(\frac{U_2}{n-1} - \frac{\alpha_1}{1-\alpha_2}\right) + \frac{\alpha_1}{1-\alpha_2} - \varepsilon^2 \frac{U_4 - U_5}{n-1} \cdot \frac{U_4 - U_6}{n-1}}{\varepsilon\left(\frac{U_1}{n} - 1\right) + 1 - \varepsilon^2\left(\frac{U_4}{n}\right)^2} \tag{18}$$
and $\hat{\rho}_2$ becomes
$$\frac{\varepsilon\left(\frac{U_3}{n-2} - \frac{\alpha_1^2}{1-\alpha_2} - \alpha_2\right) + \frac{\alpha_1^2}{1-\alpha_2} + \alpha_2 - \varepsilon^2 \frac{U_4 - U_7}{n-2} \cdot \frac{U_4 - U_8}{n-2}}{\varepsilon\left(\frac{U_1}{n} - 1\right) + 1 - \varepsilon^2\left(\frac{U_4}{n}\right)^2}. \tag{19}$$

We can then expand *any function* of these two expressions (and its first k powers) with respect to ε (up to and including the ε^{2k} term), setting $\varepsilon = 1$ in the end. With the help of the moment generating function of the previous section, we then compute the expected value of each such expansion, using $k+1$ consecutive (numerical) values of n. Note that n, $n-1$ and $n-2$ in the denominators of (18) and (19) remain *unevaluated*. To derive the corresponding general formula (valid for any n), we realize that each such result is a polynomial function of n, whose degree is not bigger than k. Having $k+1$ consecutive values of this polynomial enables us to establish the value of its coefficients.

Once we know the expected value of the first k powers of a random variable, we can easily convert these into the first few *cumulants* of the corresponding distribution.

For the $\hat{\alpha}_1$ estimator, this procedure yields the following mean:

$$m_1 \equiv \alpha_1 - \frac{1+\alpha_1+\alpha_2}{n} + \cdots \tag{20}$$

variance

$$\begin{aligned} v_1 = {} & \frac{1-\alpha_2^2}{n} + \frac{2}{n^2[(1-\alpha_2)^2-\alpha_1^2]^2}[(1-\alpha_2)^4(2+4\alpha_2+5\alpha_2^2) \\ & + \alpha_1(1-\alpha_2)^4(1+\alpha_2) - 2\alpha_1^2(1-\alpha_2)^2(2+14\alpha_2+7\alpha_2^2) \\ & - 2\alpha_1^3(1-\alpha_2)^2(1+\alpha_2) - \alpha_1^4(1-12\alpha_2-3\alpha_2^2) + \alpha_1^5(1+\alpha_2) + \alpha_1^6] + \cdots \end{aligned} \tag{21}$$

skewness

$$\kappa_{30} = \frac{6\alpha_1\alpha_2}{(1-\alpha_2)\sqrt{n(1-\alpha_2^2)}} + \cdots \tag{22}$$

and kurtosis

$$3 + 6\frac{1+3\alpha_2+\alpha_2^2-\alpha_1^2(2+6\alpha_2)-5\alpha_1^3}{n(1-\alpha_2)^2(1+\alpha_2)}\cdots \tag{23}$$

(the second term of this expression is the so called fourth normalized cumulant κ_{40}).

Similarly, $\hat{\alpha}_2$ has the following mean:

$$m_2 = \alpha_2 - \frac{2(1+2\alpha_2)}{n} + \cdots \tag{24}$$

variance

$$v_2 = \frac{1-\alpha_2^2}{n} +$$

$$2\frac{2\alpha_2(1-\alpha_2)^4(2+5\alpha_2) + \alpha_1^2(1-\alpha_2)^2(1-8\alpha_2-16\alpha_2^2) - \alpha_1^4(1-4\alpha_2-8\alpha_2^2)}{n^2[(1-\alpha_2)^2-\alpha_1^2]^2} + \cdots \tag{25}$$

skewness

$$\kappa_{03} = \frac{-6\alpha_2}{\sqrt{n(1-\alpha_2^2)}} + \cdots \tag{26}$$

and fourth normalized cumulant

$$\kappa_{04} = -6\frac{1-11\alpha_2^2}{n(1-\alpha_2^2)} + \cdots. \tag{27}$$

The covariance between $\hat{\alpha}_1$ and $\hat{\alpha}_2$ equals

$$\begin{aligned} m_{11} = {} & \frac{-\alpha_1(1+\alpha_2)}{n} + \frac{1}{n^2[(1-\alpha_2)^2-\alpha_1^2]^2}[(1-\alpha_2)^4(1+4\alpha_2+3\alpha_2^2) \\ & + \alpha_1(1-\alpha_2)^3(3+8\alpha_2-13\alpha_2^2) - 2\alpha_1^2(1-\alpha_2)^2(1+4\alpha_2+3\alpha_2^2) \\ & - 2\alpha_1^3(4+\alpha_2-15\alpha_2^2+9\alpha_2^3) + \alpha_1^4(1+4\alpha_2+3\alpha_2^2) + \alpha_1^5(5+9\alpha_2)] + \cdots \end{aligned} \tag{28}$$

and the remaining (*joint*) normalized cumulants are

$$\begin{aligned} \kappa_{21} &\equiv \frac{\mathbb{E}[(\hat{\alpha}_1-\mu_1)^2(\hat{\alpha}_2-\mu_2)]}{v_1\sqrt{v_2}} = \frac{-2(1+\alpha_2-2\alpha_2^2-\alpha_1^2)}{(1-\alpha_2)\sqrt{n(1-\alpha_2^2)}} + \cdots, \\ \kappa_{12} &\equiv \frac{\mathbb{E}[(\hat{\alpha}_1-\mu_1)(\hat{\alpha}_2-\mu_2)^2]}{v_2\sqrt{v_1}} = \frac{6\alpha_1\alpha_2}{(1-\alpha_2)\sqrt{n(1-\alpha_2^2)}} + \cdots \end{aligned} \tag{29}$$

and

$$\kappa_{31} \equiv \frac{\mathbb{E}[(\hat{\alpha}_1 - \mu_1)^3(\hat{\alpha}_2 - \mu_2)]}{v_1\sqrt{v_1 v_2}} = \frac{6\alpha_1(2 - 2\alpha_2 - 10\alpha_2^2 - \alpha_1^2)}{n(1-\alpha_2)(1-\alpha_2^2)} + \cdots,$$

$$\kappa_{22} \equiv \frac{\mathbb{E}[(\hat{\alpha}_1 - \mu_1)^2(\hat{\alpha}_2 - \mu_2)^2]}{v_1 v_2}$$

$$= \frac{2(1 + 11\alpha_2 + 5\alpha_2^2 - 17\alpha_2^3 - 4\alpha_1^2(1 + 4\alpha_2))}{n(1-\alpha_2)(1-\alpha_2^2)} + \cdots,$$

$$\kappa_{13} \equiv \frac{\mathbb{E}[(\hat{\alpha}_1 - \mu_1)(\hat{\alpha}_2 - \mu_2)^3]}{v_2\sqrt{v_1 v_2}} = \frac{6\alpha_1(1 - 11\alpha_2^2)}{n(1-\alpha_2^2)(1-\alpha_2)} + \cdots. \tag{30}$$

Using the same approach, we can find the first few moments of $F(\hat{\alpha}_2)$, where F is an arbitrary function. This shows that the leading $\left(\text{i.e., } \propto \frac{1}{\sqrt{n}}\right)$ term of the corresponding skewness is proportional to $2\alpha_2 F'(\alpha_2) - (1-\alpha_2^2)F''(\alpha_2)$. To eliminate it, we take the simplest solution of the corresponding differential equation, namely $F \equiv \text{ArcTanh}$. The resulting sample statistic $\hat{\Theta}_2 \equiv \text{ArcTanh}(\hat{\alpha}_2)$ has the following mean:

$$M_2 = \text{ArcTanh}(\alpha_2) - \frac{2 + 3\alpha_2}{n(1-\alpha_2^2)} + \cdots \tag{31}$$

variance

$$V_2 = \frac{1}{n(1-\alpha_2^2)} + \frac{2[(1-\alpha_2)^4 - \alpha_1^2(1-\alpha_2)^2(1-4\alpha_2^2) - 2\alpha_1^4\alpha_2^2]}{n^2[(1-\alpha_2)^2 - \alpha_1^2]^2(1-\alpha_2^2)^2} + \cdots \tag{32}$$

and the fourth normalized cumulant

$$K_{04} \equiv \frac{\mathbb{E}[(\hat{\Theta}_2 - M_2)^4]}{V_2^4} - 3 = \frac{2 - 6\alpha_2^2}{n(1-\alpha_2^2)} + \cdots. \tag{33}$$

Similarly, we can find the first few moments of $G(\hat{\alpha}_1, \hat{\alpha}_2)$ and eliminate the leading term of its skewness. This leads to a more complicated partial differential equation for G, whose simplest bivariate solution is $\hat{\Theta}_1 \equiv \text{ArcTanh}\left(\frac{\hat{\alpha}_1}{1-\hat{\alpha}_2}\right)$.

The mean, variance and 4th normalized cumulant of $\hat{\Theta}_1$ are

$$M_1 \equiv \operatorname{ArcTanh}\left(\frac{\alpha_1}{1-\alpha_2}\right) - \frac{(1+\alpha_2)(1-\alpha_2+2\alpha_1)}{n[(1-\alpha_2)^2-\alpha_1^2]} + \cdots, \tag{34}$$

$$V_1 \equiv \frac{1-\alpha_2^2}{n[(1-\alpha_2)^2-\alpha_1^2]} + \frac{(1-\alpha_2)^2(1-6\alpha_2-\alpha_2^2)-\alpha_1^2(1+2\alpha_2+3\alpha_2^2)}{n^2[(1-\alpha_2)^2-\alpha_1^2]^2} + \cdots \tag{35}$$

and

$$K_{40} = 2\frac{(1-\alpha_2)^2(1-4\alpha_2+\alpha_2^2)-3\alpha_1^2(1+\alpha_2^2)}{n[(1-\alpha_2)^2-\alpha_1^2](1-\alpha_2^2)} + \cdots, \tag{36}$$

respectively.

To deal with the bivariate case, we first need the covariance between $\hat{\Theta}_1$ and $\hat{\Theta}_2$, namely

$$M_{11} \equiv \frac{2\alpha_2}{n^2[(1-\alpha_2)^2-\alpha_1^2]^2(1-\alpha_2^2)}$$

$$\times[(1-\alpha_2)^3(1+\alpha_2)+\alpha_1(1-\alpha_2)^2(4+\alpha_2)-\alpha_1^2(1-\alpha_2^2)-\alpha_1^3(2+\alpha_2)]+\cdots \tag{37}$$

(note that there is no $\frac{1}{n}$ term, which implies that the two estimators are, in the $n \to \infty$ limit, independent of each other).

In addition to zero skewness (i.e., $K_{30} = 0$ and $K_{03} = 0$), we also get $K_{21} = 0$, whereas

$$K_{21} = \frac{-2\alpha_2}{\sqrt{n(1-\alpha_2^2)}} + \cdots. \tag{38}$$

Finally, two of the remaining fourth-order normalized cumulants, namely, K_{31} and K_{13} are also equal to zero, and

$$K_{22} \equiv \frac{\mathbb{E}\{(\hat{\Theta}_1-M_1)^2(\hat{\Theta}_2-M_2)^2\} - 2\mathbb{E}\{(\hat{\Theta}_1-M_1)(\hat{\Theta}_2-M_2)\}^2}{V_1V_2} - 1$$

$$= \frac{4\alpha_2^2}{n(1-\alpha_2^2)} + \cdots. \tag{39}$$

4. Approximate Sampling Distributions

Based on these results, we can now improve the basic Normal approximation, by adding the corresponding corrections.

The first thing we need to do is to *standardize* both $\hat{\alpha}_1$ and $\hat{\alpha}_2$ by introducing

$$Z_1 = \frac{\hat{\alpha}_1 - m_1}{\sqrt{v_1}} \tag{40}$$

and

$$Z_2 = \frac{\hat{\alpha}_2 - m_2}{\sqrt{v_2}}. \tag{41}$$

The basic Normal approximation takes $m_1 = \alpha_1$, $m_2 = \alpha_2$ and

$$v_1 = v_2 = \frac{1 - \alpha_2^2}{n} \tag{42}$$

and uses

$$f_0(z) \equiv \exp\left(-\frac{z^2}{2}\right) \Big/ \sqrt{2\pi} \tag{43}$$

as the (approximate) probability density function of each Z_1 and Z_2.

In our approximation, we must use the complete, two-term expression (20) for the mean and (21) for the variance, and we must also modify (43) to

$$f_0(z) \cdot \left[1 + \frac{\kappa_3}{3!} H_3(z) + \frac{1}{2}\left(\frac{\kappa_3}{3!}\right)^2 H_6(z) + \frac{\kappa_4}{4!} H_4(z) + \cdots\right], \tag{44}$$

where κ_3 and κ_4 are the corresponding third and fourth cumulants, and

$$H_i(z) = \frac{(-1)^i}{f_0(z)} \cdot \frac{d^i f_0(z)}{dz^i} \tag{45}$$

(each is a simple, *i*th degree polynomial).

We can show that the error of this approximation is of the $O(n^{-3/2})$ type [3].

To demonstrate the improvement achieved over the Normal approximation, we have generated one hundred thousand random sequences of 60 observations from the AR(2) model with $\alpha_1 = 1.3$, $\alpha_2 = -0.8$ ($\mu = 0$ and $\sigma = 1$; this choice is inconsequential), computed the corresponding values of $\hat{\alpha}_1$, and displayed these in a histogram (thus getting a fairly accurate idea about the *exact* distribution). This is plotted together with the Normal and $\frac{1}{n}$-accurate approximations of (43) and (44), respectively, both transformed back into the $\hat{\alpha}_1$ scale (see Figure 1).

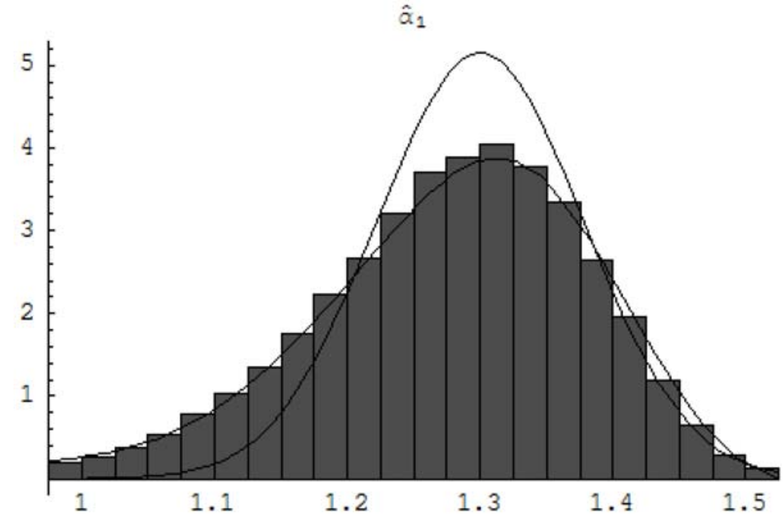

Figure 1

We can see that the Normal approximation (the symmetric curve) is still extremely inaccurate, whereas the new approximation, given by

$$f(\hat{\alpha}_1) = \frac{\exp\left(-\frac{z^2}{2}\right)}{0.27867}$$

$$\times (1.0101 + 0.3730z + 0.0957z^2 - 0.1243z^3 - 0.0739z^4 + 0.0077z^6), \quad (46)$$

where

$$z = \frac{\hat{\alpha}_1 - 1.275}{0.1112} \quad (47)$$

appears more than adequate. Note that κ_3 and κ_4 of (44) are to be identified with κ_{30} and κ_{40} of the previous section.

The corresponding results for $\hat{\alpha}_2$ are displayed in Figure 2 (now, of course $\kappa_3 \equiv \kappa_{03}$ and $\kappa_4 \equiv \kappa_{04}$).

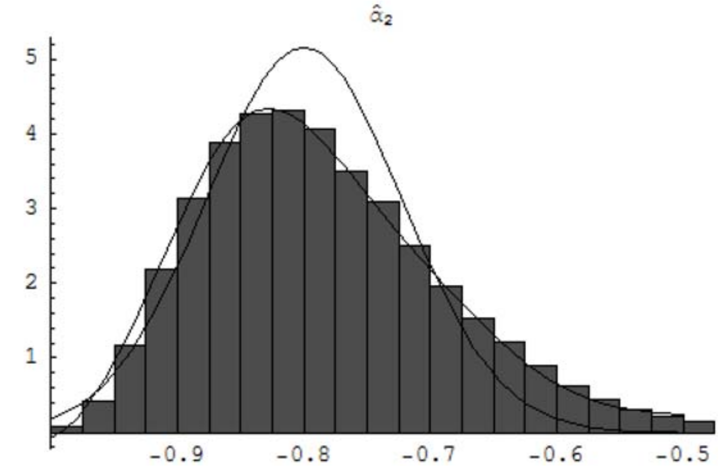

Figure 2

When we replace $\hat{\alpha}_1$ and $\hat{\alpha}_2$ by $\text{ArcTanh}\left(\dfrac{\hat{\alpha}_1}{1-\hat{\alpha}_2}\right)$ and $\text{ArcTanh}\,(\hat{\alpha}_2)$, the two distributions become nearly symmetrical, yet the improvement achieved by the new approximation over the Normal distribution is still quite pronounced (Figures 3 and 4, respectively).

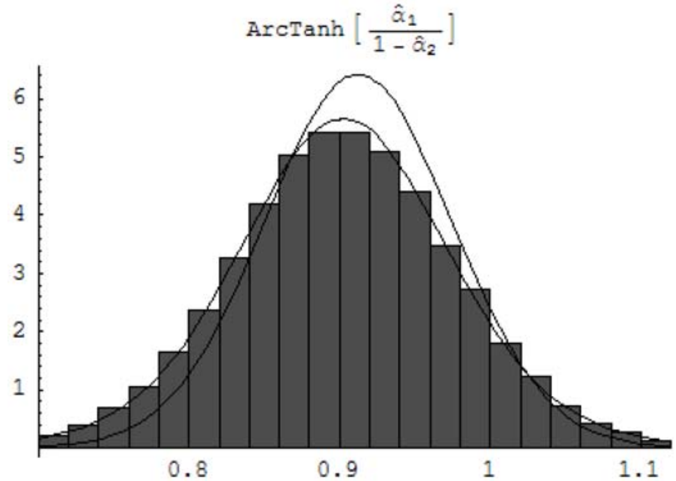

Figure 3

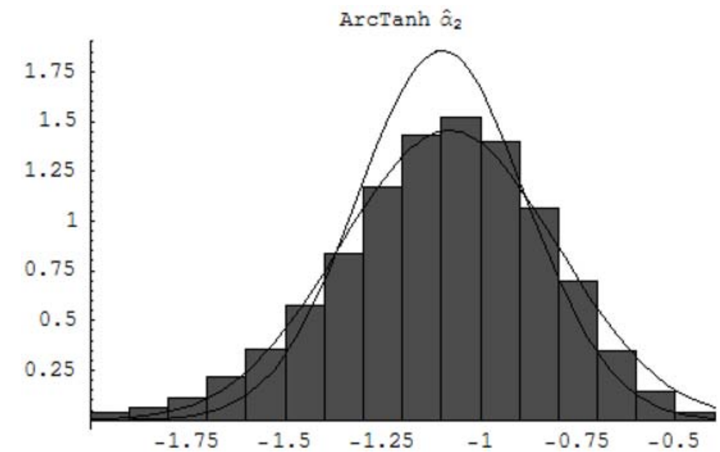

Figure 4

To find an improved approximation to the *bivariate* distribution of Z_1 and Z_2, it is most convenient to start with the corresponding *joint* moment generating function, which is, to the same degree of approximation, given by

$$\exp\left(\frac{t_1^2 + t_2^2 + 2rt_1t_2}{2}\right) \cdot \left[1 + \frac{\kappa_{30}t_1^3 + 3\kappa_{21}t_1^2t_2 + 3\kappa_{12}t_1^2t_2 + \kappa_{03}t_2^3}{3!}\right.$$
$$+ \frac{(\kappa_{30}t_1^3 + 3\kappa_{21}t_1^2t_2 + 3\kappa_{12}t_1^2t_2 + \kappa_{03}t_2^3)^2}{2(3!)^2}$$
$$\left. + \frac{\kappa_{40}t_1^4 + 4\kappa_{31}t_1^3t_2 + 6\kappa_{22}t_1^2t_2^2 + 4\kappa_{13}t_1^3t_2 + \kappa_{04}t_2^4}{4!} + \cdots\right], \tag{48}$$

where r is the correlation coefficient between Z_1 and Z_2.

The previous expression can be easily converted to the corresponding joint probability density function by

$$\exp\left(\frac{t_1^2 + t_2^2 + 2rt_1t_2}{2}\right) \to \frac{1}{2\pi\sqrt{1-r^2}} \exp\left(\frac{z_1^2 + z_2^2 - 2rz_1z_2}{2(1-r^2)}\right) \equiv g_0(z_1, z_2) \tag{49}$$

and

$$t_1^i t_2^j \exp\left(\frac{t_1^2 + t_2^2 + 2rt_1t_2}{2}\right) \to (-1)^{i+j} \cdot \frac{d^{i+j} g_0(z_1, z_2)}{dz_1^i dz_2^j}. \tag{50}$$

For the $\hat{\alpha}_1, \hat{\alpha}_2$ pair, this results in an expression too long to be quoted here, but for $\hat{\Theta}_1 = \mathrm{ArcTanh}\left(\frac{\hat{\alpha}_1}{1-\hat{\alpha}_2}\right)$ and $\hat{\Theta}_2 = \mathrm{ArcTanh}(\hat{\alpha}_2)$, using the same $n = 60$ and $\alpha_1 = 1.3$, $\alpha_2 = -0.8$ as before, we get the following result:

$$f(\hat{\theta}_1, \hat{\theta}_2) = \frac{g_0(z_1, z_2)}{0.020224} \cdot [1.0169 - 0.1782z_1 - 0.2395z_2 - 0.0004z_1^2$$

$$+ 0.1205z_1z_2 + 0.0788z_2^2 + 0.0649z_1^3 + 0.2615z_1^2z_2 + 0.1352z_1z_2^2$$

$$+ 0.0188z_2^3 - 0.0204z_1^4 - 0.1236z_1^3z_2 - 0.1796z_1^2z_2^2 - 0.0794z_1z_2^3$$

$$- 0.0143z_2^4 + 0.0021z_1^6 + 0.0170z_1^5z_2 + 0.0429z_1^4z_2^2 + 0.0366z_1^3z_2^3$$

$$+ 0.0140z_1^2z_2^4 + 0.0025z_1z_2^5 + 0.0002z_2^6], \tag{51}$$

where $r = M_{11}/\sqrt{V_1V_2} = -0.2897$ and

$$z_1 = \frac{\hat{\theta}_1 - 0.90281}{0.074472}, \tag{52}$$

$$z_2 = \frac{\hat{\theta}_2 - 1.08009}{0.27156}. \tag{53}$$

To compare it with the corresponding empirical distribution, we display a set of four contours (boundaries of an *acceptance region* to test H_0: $\alpha_1 = 1.3$, $\alpha_2 = -0.8$ against all alternatives, for four different values of significance level) in Figure 5. Note that the empirical results (dashed lines) are not very accurate.

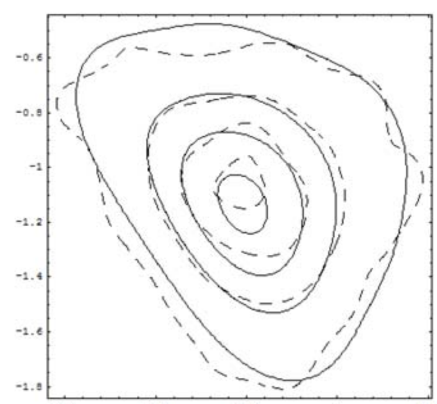

Figure 5

This again represents a marked improvement over the Normal approximation (4 concentric near-circles).

5. Conclusion

We have shown that it is relatively simple to find a good approximation to sampling distributions of estimators related to the AR(2) model. The technique described in this article is fully general, and can be applied to any other model, as long as the sample statistics under consideration meet the assumptions of the Central Limit Theorem.

References

[1] M. G. Kendall and A. Stuart, The Advanced Theory of Statistics, Vol. 3, Hafner Publishing Company, New York, 1968.

[2] Pham Dinh Tuan, Approximate distribution of parameter estimators for first-order autoregressive models, J. Time Series Anal. 13 (1992), 147-170.

[3] J. Vrbik, Finite-*n* corrections to normal approximation, Comm. Statist. Simulation Comput. 34 (2005), 827-837.