Journal of Statistical and Econometric Methods, vol.6, no.1, 2017, 101-108
ISSN: 1792-6602 (print), 1792-6939 (online)
Scienpress Ltd, 2017

Asymptotic relationship between sample mean and sample variance for autoregressive processes of order 1

Athanase Polymenis[1]

Abstract

Autoregressive processes of order 1 (or AR(1) processes) have been extensively used in econometrics and time series literature. Noting that an early important result concerning the sample mean $\overline{U}$ and variance S_U^2 of independent normally distributed random variables U with equal means and variances is that $\overline{U}$ and S_U^2 are independent, the present article investigates whether this result can be extended to AR(1) non-stationary processes as the sample size becomes very large. To this end, a property called "asymptotic stationarity" is used for algebraic calculations. A result for asymptotic independence concerning the sample mean and variance is then adequately derived for these types of processes.

Mathematics Subject Classification: 62E20; 62M10

Keywords: Autoregressive process; Asymptotic stationarity; Asymptotic independence

[1] Department of Economics, University of Patras, University Campus at Rio, 26504 Rio-Patras, Greece. E-mail: athanase@upatras.gr

Article Info: *Received* : November 14, 2016. *Revised* : December 11, 2016.
Published online : January 25, 2017.

1 Introduction

In the present work we consider AR(1) autoregressive random processes of the form

$$X_{n+1} = aX_n + s\varepsilon_{n+1} \quad (1)$$

where ε_i are independently $N(0,1)$ distributed so that $s\varepsilon_i \sim N(0,s^2)$, s is a positive constant, and $|a|<1$.

The above process is usually encountered in econometrics; in this case X_n could represent the error terms of a regression model, also called innovations, which are then correlated with each other, and this creates a serious problem known as autocorrelation problem. On the other hand the AR(1) process is also encountered in time series analyses where X_n denotes an economic variable and index n represents time; the process is then very useful for forecasting future values of X. As noted by Davidson and MacKinnon 1993 (p.329) " this is a most commonly encountered error process in applied work". The aforementioned condition $|a|<1$ is called stationarity condition. We now focus on processes like (1), setting $X_0 = 0$ as initial condition for solving the equation defined by (1). Condition $|a|<1$ implies that the variance

$$Var(X_n) = s^2(1-a^{2n})/(1-a^2)$$

exists as n tends to infinity, so that, for large n, we have

$$Var(X_n) \approx s^2/(1-a^2) = \sigma^2,$$

showing that the variance is constant. Similarly, the covariance $Cov(X_n, X_{n+k}) = a^k s^2(1-a^{2n})/(1-a^2) = a^k\ Var\ (X_n)$ -and $k > 0$ - exists for n large, and so, $Cov(X_n, X_{n+k}) \approx a^k\sigma^2$ (see also [1], pp. 327-328), which does not depend on n , but only on the lag k . Note that a first order autoregressive random process $\{X_n\}$ with these properties is called " asymptotically stationary up to order 2" (see [2], p. 119). As argued in the latter reference, instead of considering this asymptotically stationary process one could equivalently consider the stationary process $X_n = s\sum_{i=0}^{\infty} a^i\, \varepsilon_{n-i}$ (see [2], pp. 121-122), which is a moving average representation, noted as MA(∞), where ∞ stands for infinite order. Thus the property of asymptotic stationarity allows for treating the AR(1) process in equation (1), with initial condition $X_0 = 0$, as a "usual" stationary process by letting n tend to infinity. The reason for considering the model represented by equation (1), with initial condition $X_0 = 0$, is that, as underlined by Spanos (see [3], p. 153), assuming "usual" stationarity is unrealistic, thus implying that for econometric modelling it is more useful to consider index $n = 0, 1, 2, \ldots$ (instead of $n = 0, \pm 1, \pm 2, \ldots$ used for "usual" stationarity). An application example about that is provided by Spanos ([3], pp. 279-281).

Also note that considering the fact that $E(X_n) = 0$ in (1) will not affect covariance properties since if $E(X_n) = \mu \neq 0$ one can write $X_n - \mu = a(X_{n-1} - \mu) + s\varepsilon_n$, with $E(X_n - \mu) = 0$ (see [4], p. 13).

The development of the theory provided in the next sections aims at investigating whether a well known and important result from statistical theory concerning independently normally distributed random variables is still asymptotically valid. This result is given (for instance) by the first part of Theorem (1) of [5], p. 119), namely that if U_i are independently normally distributed random variables (μ, σ^2), then their sample mean $\bar{U}$ and variance $S_U^2 = \sum_{i=1}^{n} (U_i - \bar{U})^2 / (n-1)$ are independent. Noting that since the case where U_i are independent can be considered as a particular case of the AR(1) process previously presented, for it simply amounts to taking $a = 0$ in (1), the aim of the present article will be to extend the independence result concerning $\bar{U}$ and S^2 to an asymptotic independence for the process generated by (1), with initial condition $X_0 = 0$, since, to the best of our knowledge, this result does not appear in existing literature.

In section 2, approximate expressions for some useful covariances are calculated when n is large. Section 3 provides approximate results, for n large, concerning independence properties.

2 Asymptotic Theory for Second Order Moments

Lemma 1 *For large n, the covariance $Cov\ (\bar{X}\ ,\ X_i - \bar{X})$ is $O(\frac{1}{n})$, and hence $Cov\left(\bar{X}, \frac{X_i - \bar{X}}{\sqrt{n-1}}\right)$ is $O(\frac{1}{n\sqrt{n}})$.*

Proof. Since according to the AR(1) random process
$X_i = a^{i-1}s\varepsilon_1 + a^{i-2}s\varepsilon_2 + \cdots as\varepsilon_{i-1} + s\varepsilon_i$, we now compute $\bar{X}$ as follows.
$\sum_{i=1}^{n} X_i = (1 + a + \cdots + a^{n-1})s\varepsilon_1 + (1 + a + \cdots + a^{n-2})s\varepsilon_2 + \cdots + (1 + a)s\varepsilon_{n-1} + s\varepsilon_n$
$= \frac{1-a^n}{1-a}s\varepsilon_1 + \frac{1-a^{n-1}}{1-a}s\varepsilon_2 + \cdots + \frac{1-a^2}{1-a}s\varepsilon_{n-1} + \frac{1-a}{1-a}s\varepsilon_n = \frac{s}{1-a}[(1 - a^n)\varepsilon_1 + (1 - a^{n-1})\varepsilon_2 + \cdots + (1 - a^2)\varepsilon_{n-1} + (1 - a)\varepsilon_n]$,
and thus

$$\bar{X} = \frac{s}{(1-a)n}\ [(1 - a^n)\varepsilon_1 + (1 - a^{n-1})\varepsilon_2 + \cdots + (1 - a^2)\varepsilon_{n-1} + (1 - a)\varepsilon_n]. \quad (2)$$

Thus $E(\bar{X}^2) = \frac{s^2}{(1-a)^2 n^2}[(1 - a^n)^2 + \cdots + (1 - a)^2]$, since $E\left(\varepsilon_i\varepsilon_j\right) = 0$ for $i \neq j$ and $E(\varepsilon_i^2) = 1$. A direct development of $(1 - a^n)^2 + \cdots + (1 - a)^2$ leads then to $n - \frac{2a(1-a^n)}{1-a} + \frac{a^2(1-a^{2n})}{1-a^2}$, and so

$$E(\bar{X}^2) = \frac{s^2}{(1-a)^2 n^2}\left(n - \frac{2a(1-a^n)}{1-a} + \frac{a^2(1-a^{2n})}{1-a^2}\right),$$

and thus, for large n, $E(\bar{X}^2) \approx \frac{s^2}{(1-a)^2 n}$.

On the other hand we have

$$Cov(\bar{X},\ X_i - \bar{X}) = E(\bar{X}.\ (X_i - \bar{X})) - E(\ X_i - \bar{X})E(\bar{X}), \text{ with } E(\bar{X}) = 0\ ,$$

since this is a linear combination of $E(\varepsilon_i)$ which all equal 0 (i =1,…, n).

It results that $Cov(\bar{X},\ X_i - \bar{X}) = E(\bar{X}.\ (X_i - \bar{X}))= E(X_i\bar{X} - \bar{X}^2)$. We now compute $E(X_i\bar{X})$, using (2).

$$E(X_i\bar{X}) = \frac{s^2}{(1-a)n}E\{\left[a^{i-1}\varepsilon_1 + \cdots + a\varepsilon_{i-1} + \varepsilon_i\right]\left[\left(1 - a^i\right)\varepsilon_1 + \cdots + (1 - a^2)\varepsilon_{i-1} + (1 - a)\varepsilon_i\right]\}$$

$$= \frac{s^2}{(1-a)n}\left[a^{i-1}(1-a^i) + a^{i-2}(1-a^{i-1}) \ldots + a(1-a^2) + (1-a)\right]$$

since $E(\varepsilon_\iota \varepsilon_j) = 0$ for $i \neq j$ and $E(\varepsilon_i^2) = 1$. It results that

$$E(X_i\bar{X}) = \frac{s^2}{(1-a)n}\left[\frac{1-a^i}{1-a} - a\frac{1-a^{2i}}{1-a}\right] = \frac{s^2}{(1-a)^2 n}\left[(1-a^i) - a(1-a^{2i})\right].$$

We finally obtain $Cov(\bar{X}, X_i - \bar{X})$, for large n,

$Cov(\bar{X}, X_i - \bar{X}) = E(X_i\bar{X}) - E(\bar{X}^2) \approx \frac{s^2}{(1-a)^2 n}\left[(1-a^i) - a(1-a^{2i})\right] - \frac{s^2}{(1-a)^2 n} = \frac{s^2}{(1-a)^2 n}(a^{2i+1} - a^i - a).$

Since |a|<1, It results that $Cov(\bar{X}, X_i - \bar{X})$ is $O(\frac{1}{n})$, and thus $Cov(\bar{X}, \frac{X_i-\bar{X}}{\sqrt{n-1}})$ is $O(\frac{1}{n\sqrt{n}})$. This completes the proof of Lemma 1.

Lemma 2 *For large n, the covariance*

$$Cov\left(\frac{X_i-\bar{X}}{\sqrt{n-1}}, \frac{X_j-\bar{X}}{\sqrt{n-1}}\right) = \frac{1}{n-1}\, Cov(X_i, X_j) + O(\frac{1}{n^2}),\ for\ i \neq j,$$

and thus $\left(\frac{X_i-\bar{X}}{\sqrt{n-1}}, \frac{X_j-\bar{X}}{\sqrt{n-1}}\right) \approx \frac{1}{n-1}\, Cov(X_i, X_j)$.

Proof. $Cov(X_i - \bar{X}, X_j - \bar{X}) = E(X_i - \bar{X}, X_j - \bar{X}) - E(X_i - \bar{X})E(X_j - \bar{X}) = E(X_i - \bar{X}, X_j - \bar{X})$

since $E(X_i) = E(X_j) = E(\bar{X}) = 0$. Hence

$$Cov(X_i - \bar{X}, X_j - \bar{X}) = E(X_iX_j) - E(X_i\bar{X}) - E(X_j\bar{X}) + E(\bar{X}^2).$$

From the results obtained in the previous lemma, $E(X_i\bar{X})$, $E(X_j\bar{X})$, and $E(\bar{X}^2)$ are all $O(\frac{1}{n})$ for large n, whereas $E(X_i X_j) = a^k(1-a^{2i})s^2/(1-a^2)$, for $i < j$, which does not depend on n, and the lemma is proved.

Lemma 3 *For large n, the variance* $Var(\frac{X_i-\bar{X}}{\sqrt{n-1}}) = \frac{1}{n-1} Var(X_i) + O(\frac{1}{n^2})$, *and thus*

$$Var(\frac{X_i-\bar{X}}{\sqrt{n-1}}) \approx \frac{1}{n-1} Var(X_i).$$

Proof.

$$Var\left(\frac{X_i - \bar{X}}{\sqrt{n-1}}\right) = E\left[\frac{(X_i - \bar{X})^2}{n-1}\right]$$

since $E(X_i - \bar{X}) = 0$. Thus

$\frac{1}{n-1}Var(X_i - \bar{X}) = \frac{1}{n-1}E(X_i^2) - \frac{2}{n-1} E(X_i\bar{X})+\frac{1}{n-1}E(\bar{X}^2)$. The lemma follows since, as obtained for the previous lemmas, $E(X_i\bar{X})$ and $E(\bar{X}^2)$ are $\left(\frac{1}{n}\right)$, for large n , and $E(X_i^2) = Var(X_i) = (1-a^{2i})s^2/(1-a^2)$, which does not depend on n.

Lemma 4 *For large , the variance* $Var(\bar{X}) = \frac{s^2}{(1-a)^2 n}+ O(\frac{1}{n^2})$ *and thus*

$$Var(\bar{X}) \approx \frac{s^2}{(1-a)^2 n}.$$

Proof.

The proof has already been obtained in Lemma 1. Indeed we found that

$$E(\bar{X}^2)=\frac{s^2}{(1-a)^2 n^2}\left(n - \frac{2a(1-a^n)}{1-a} + \frac{a^2(1-a^{2n})}{1-a^2}\right) = \frac{s^2}{(1-a)^2 n} + O(\frac{1}{n^2}).$$

3 Asymptotic Independence Results

Theorem 1 *For large* n*, the sample mean* $\bar{X}$ *is independent of the vector random variable* $(\frac{X_1-\bar{X}}{\sqrt{n}}, \dots, \frac{X_n-\bar{X}}{\sqrt{n}})$.

Proof.

Lemmas 3 and 2 have respectively shown that, for large n,

$$Var(\frac{X_i-\bar{X}}{\sqrt{n-1}}) \approx \frac{1}{n-1} Var(X_i) \quad \text{and} \quad Cov\left(\frac{X_i-\bar{X}}{\sqrt{n-1}}, \frac{X_j-\bar{X}}{\sqrt{n-1}}\right) \approx \frac{1}{n-1} Cov(X_i, X_j)$$

since the differences

$$Var\left(\frac{X_i-\bar{X}}{\sqrt{n-1}}\right) - \frac{1}{n-1} Var(X_i) \quad \text{and} \quad Cov\left(\frac{X_i-\bar{X}}{\sqrt{n-1}}, \frac{X_j-\bar{X}}{\sqrt{n-1}}\right) - \frac{1}{n-1} Cov(X_i, X_j)$$

are both $O(\frac{1}{n^2})$ or, equivalently, $o(\frac{1}{n})$. In the same spirit $Cov\ (\bar{X}, \frac{X_i - \bar{X}}{\sqrt{n-1}})$ can be considered as negligible since it has been found to be $O(\frac{1}{n\sqrt{n}})$ by the result of Lemma 1, or equivalently, $o(\frac{1}{n})$. Furthermore $Var(\bar{X})$ is approximately equal to $\frac{s^2}{(1-a)^2 n}$ by Lemma 4, since the difference $Var(\bar{X}) - \frac{s^2}{(1-a)^2 n}$ is $O(\frac{1}{n^2})$, or equivalently, $o(\frac{1}{n})$. Theorem 1 then follows using theory of independent variables. First note that the vector $\boldsymbol{V} = (\bar{X}, \frac{X_1 - \bar{X}}{\sqrt{n-1}}, \dots, \frac{X_n - \bar{X}}{\sqrt{n-1}})$ is normally distributed since any linear combination of its components is a linear combination of the X_i (see Definition 4.9(7) of [5], p. 118). Letting Γ be the covariance matrix of $\boldsymbol{V}$, the characteristic function $\Phi_{\boldsymbol{V}}(t_1, t_{2,\dots,}^1 t_2^n)$ of $\boldsymbol{V}$ is thus $exp(-\frac{1}{2}t\Gamma\ t')$, where $t = (t_1, t_{2,\dots,}^1 t_2^n)$, and prime denotes transpose (see [5], p. 187). Since by Lemmas 1 and 4 all entries of row 1 and column 1, except for the first one which refers to $Var(\bar{X})$, are approximately equal to 0, for large n , the expression $exp(-\frac{1}{2}t\Gamma\ t')$ can be written, for large n, as the approximate product $exp\left(-\frac{1}{2}t_1{}^2 Var(\bar{X})\right).exp\left(-\frac{1}{2}t_2\Sigma\ t_2{}'\right)$, where Σ is the covariance matrix of the vector $\boldsymbol{V}$, and hence Theorem 1 is proved.

Theorem 2 *For large n, $\bar{X}^2$ is independent of $S_X^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2$.*

Proof. The proof is based on well known general results from theory of independent random variables and on Theorem 1. We use functions g and h, $g: \mathbb{R} \to \mathbb{R}$ and $h: \mathbb{R}^n \to \mathbb{R}$, with $g(\bar{X}) = \bar{X}^2$, and for $\boldsymbol{Y} = \left(\frac{X_1 - \bar{X}}{\sqrt{n-1}}, \dots, \frac{X_n - \bar{X}}{\sqrt{n-1}}\right)$, $h(\boldsymbol{Y}) = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2$, where $g(\bar{X})$ maps the sample space Ω into $\mathbb{R}$ by $g(\bar{X})(\omega) = g(\bar{X}(\omega)) = \bar{X}^2(\omega)$, for $\omega \in \Omega$, and $h(\boldsymbol{Y})$ maps $\boldsymbol{\Omega} = \Omega_1 \mathrm{x} \Omega_2 \mathrm{x} \dots \mathrm{x} \Omega_n$ into $\mathbb{R}$ by $h(\boldsymbol{Y})(\boldsymbol{\omega}) = h(\boldsymbol{Y}(\boldsymbol{\omega})) = \frac{1}{n-1}\sum_{i=1}^n [X_i(\omega_i) - \frac{1}{n}\sum_{i=1}^n X_i(\omega_i)]^2$, for $\boldsymbol{\omega} = (\omega_1, \omega_2, \dots, \omega_n) \in \boldsymbol{\Omega}$. Let $g^{-1}((-\infty, x])$ and $h^{-1}((-\infty, y])$ denote the preimages of $(-\infty, x]$ and $(-\infty, y]$ under g and h respectively, and consider the joint distribution function of the random vector $(g(\bar{X}), h(\boldsymbol{Y}))$, that is, $P(g(\bar{X}) \leq x, h(\boldsymbol{Y}) \leq y)$, with $x \in \mathbb{R}$ and $y \in \mathbb{R}$. Then, $P(g(\bar{X}) \leq x, h(\boldsymbol{Y}) \leq y) = P\{g(\bar{X}) \in$

$(-\infty, x]), h(\boldsymbol{Y}) \in (-\infty, y])\} = P\{\bar{X} \in g^{-1}(-\infty, x]), \boldsymbol{Y} \in h^{-1}(-\infty, y]\}$ which, for large $n,$ is approximately equal to $P\{\bar{X} \in g^{-1}(-\infty, x])\}. P\{\boldsymbol{Y} \in h^{-1}(-\infty, y]\},$ by the result of Theorem 1.

Finally $P\{\bar{X} \in g^{-1}(-\infty, x])\}. P\{\boldsymbol{Y} \in h^{-1}(-\infty, y]\} = P\{g(\bar{X}) \in (-\infty, x]\}.$
$P\{h(\boldsymbol{Y}) \in (-\infty, y]\} = P(g(\bar{X}) \leq x). P(h(\boldsymbol{Y}) \leq y).$ Thus, for large n, the joint distribution of $\left(g(\bar{X}), h(\boldsymbol{Y})\right)$ is approximately equal to the product of the marginal distributions of $g(\bar{X})$ and $h(\boldsymbol{Y})$, thus proving Theorem 2.

References

[1] R. Davidson and J. G. MacKinnon, *Estimation and inference in econometrics*, Oxford University Press, 1993.

[2] M. B. Priestley, *Spectral analysis and time series*, *Volume 1- Univariate time series*, Academic Press, 1981.

[3] A. Spanos, *Statistical foundations of econometric modelling*, Cambridge University Press, 1986.

[4] C.W. J. Granger and P. Newbold, *Forecasting economic time series*, Academic Press, New York, 1977.

[5] G. Grimmett and D. Stirzaker, *Probability and random processes*, Oxford University Press, 2001.