Rep. Univ. Electro-Comm. 29-2, (Sci. & Tech. Sect.), pp. 213-218 February, 1979

On the Second Order Asymptotic Optimality of Estimators in an Autoregressive Process*

Masafumi AKAHIRA**

Abstract

Let $\{X_t\}$ be defined by $X_t=\theta X_{t-1}+U_t$ $(t=1, 2, \ldots)$, where $\{U_t\}$ is a sequence of independent identically distributed random variables with mean 0 and variance 1 and X_0 is a random variable with mean 0 and variance σ^2 and for each t X_0 is independent of U_t. We assume that $|\theta|<1$. It is shown that the stochastic expansion of the maximum likelihood estimator (MLE) of θ is obtained and the MLE is second order asymptotically efficient and that the least squares estimator is second order asymptotically efficient when U_t $(t=1, 2, \ldots)$ and X_0 are normally distributed and $\sigma^2=1/(1-\theta^2)$ in the above process. It is noted that the initial condition of X_0 does not affect the second order asymptotic efficiency of the MLE.

1. Introduction

It is shown by Akahira ([1]) that the bound of the second order asymptotic distributions of second order asymptotically median unbiased (AMU) estimators of θ and the estimator of θ attaining it, that is, the second order asymptotically efficient estimator of θ are obtained in an autoregressive process $\{X_t\}$ which is defined recursively by $X_t=\theta X_{t-1}+U_t$ $(t=1, 2, \ldots)$, where $X_0=0$, $|\theta|<1$ and $\{U_t\}$ is a sequence of independent identically distributed (i. i. d) random variables with mean 0 and variance σ^2.

In this paper we consider an autoregressive process $\{X_t\}$ which is defined by $X_t=\theta X_{t-1}+U_t$ $(t=1, 2, \ldots)$, where $\{U_t\}$ is a sequence of i. i. d. random variables with mean 0 and variance 1 and X_0 is a random variable with mean 0 and variance σ^2 and for each t X_0 is independent of U_t.

We assume that $|\theta|<1$. It is shown that the stochastic expansion of the maximum likelihood estimator (MLE) of θ is obtained and the MLE is second order asymptotically efficient and that the least squares estimator is second order asymptotically efficient when U_t $(t=1, 2, \ldots)$ and X_0 are normally distributed and $\sigma^2=1/(1-\theta^2)$ in the above process. It is noted that the initial condition of X_0 does not affect the second order asymptotic efficiency of the MLE.

2. Results

Let $(\mathcal{X}, \mathcal{B})$ be a sample space and Θ a parameter space, which is assumed to be an open set in a Euclidean 1-space R^1. We shall denote by $(\mathcal{X}^{(T)}, \mathcal{B}^{(T)})$ the T-fold direct products of

* Received on October 24, 1978
This research was supported by Japan Ministry of Education.
** Statistical Laboratory, University of Electro-Communications

$(\mathcal{X}, \mathcal{B})$. For each $T=1, 2, \ldots$, the points of $\mathcal{X}^{(T)}$ will be denoted by $\tilde{x}_T=(x_1, \ldots, x_T)$. We consider a sequence of classes of probability measures $\{P_{T,\theta} : \theta \in \Theta\}$ $(T=1, 2, \ldots)$ each defined on $(\mathcal{X}^{(T)}, \mathcal{B}^{(T)})$ such that for each $T=1, 2, \ldots$ and each $\theta \in \Theta$ the following holds:

$$P_{T,\theta}(B^{(T)})=P_{T+1,\theta}(B^{(T)} \times \mathcal{X})$$

for all $B^{(T)} \in \mathcal{B}^{(T)}$.

An estimator of θ is defined to be a sequence $\{\hat{\theta}_T\}$ of $\mathcal{B}^{(T)}$-measurable functions $\hat{\theta}_T$ on $\mathcal{X}^{(T)}$ to Θ $(T=1, 2, \ldots)$. For simplicity we may denote an estimator $\hat{\theta}$ instead of $\{\hat{\theta}_T\}$. For increasing sequence of positive numbers $\{c_T\}$ (c_T tending to infinity) an estimator $\hat{\theta}$ is called consistent with order $\{c_T\}$ (or $\{c_T\}$-consistent for short) if for every $\varepsilon>0$ and every $\vartheta \in \Theta$ there exist a sufficiently small positive number δ and a sufficiently large positive number L satisfying the following:

$$\lim_{T\to\infty} \sup_{\theta : |\theta-\vartheta|<\delta} P_{T,\theta}\{c_T|\hat{\theta}_T-\theta| \geq L\} < \varepsilon \quad ([2]).$$

A $\{c_T\}$-consistent estimator $\hat{\theta}$ is second order asymptotically median unbiased (or second order AMU) if for any $\vartheta \in \Theta$, there exists a positive number δ such that

$$\lim_{T\to\infty} \sup_{\theta : |\theta-\vartheta|<\delta} c_T \left| P_{T,\theta}\{\hat{\theta}_T \leq \theta\} - \frac{1}{2} \right| = 0;$$

$$\lim_{T\to\infty} \sup_{\theta : |\theta-\vartheta|<\delta} c_T \left| P_{T,\theta}\{\hat{\theta}_T \geq \theta\} - \frac{1}{2} \right| = 0.$$

We have defined a second order AMU estimator to be second order asymptotically efficient if the second order asymptotic distribution of it attains uniformly the bound of the second order asymptotic distributions of the second order AMU estimators ([3]). In the subsequent discussions we shall deal only with the case when $c_T=\sqrt{T}$.

Let X_t $(t=1, 2, \ldots)$ be defined recursively by

(1) $\quad X_t=\theta X_{t-1}+U_t \quad (t=1, 2, \ldots),$

where $\{U_t\}$ is a sequence of i.i.d. random variables with mean 0 and variance 1 and X_0 is a random variable with mean 0 and variance σ^2 and for each t X_0 is independent of U_t. We assume that $\Theta=(-1, 1)$, i.e., $|\theta|<1$. It is seen that the MLE of θ is $\{\sqrt{T}\}$-consistent.

We denote a joint density of $(X_0, X_1, \ldots, X_T)$ by $L(\theta)=L(\theta : x_0, x_1, \ldots, x_T)$. Suppose that the joint density satisfies the following:

(i) $\{(x_0, \tilde{x}_T) : L(\theta ; x_0, \tilde{x}_T)>0\}$ does not depend on θ;

(ii) For almost all $(x_0, \tilde{x}_T)$, $L(\theta ; x_0, \tilde{x}_T)$ is three times differentiable in θ;

(iii) For each $\theta \in \Theta$

$$0<I(\theta)=\frac{1}{T}E_\theta\left[\left\{\frac{\partial}{\partial\theta}\log L(\theta)\right\}^2\right]=-\frac{1}{T}E_\theta\left[\frac{\partial^2}{\partial\theta^2}\log L(\theta)\right]<\infty;$$

(iv) There exist

$$J(\theta)=\frac{1}{T}E_\theta\left[\left\{\frac{\partial^2}{\partial\theta^2}\log L(\theta)\right\}\left\{\frac{\partial}{\partial\theta}\log L(\theta)\right\}\right]$$

and

$$K(\theta)=\frac{1}{T}E\left[\left\{\frac{\partial}{\partial\theta}\log L(\theta)\right\}^3\right]$$

and the following holds:

$$\frac{1}{T}E_\theta\left[\frac{\partial^3}{\partial\theta^3}\log L(\theta)\right]=-3J(\theta)-K(\theta).$$

By the same way as in [3] it is shown that the MLE is second order asymptotically efficient. Let $\hat{\theta}_{ML}$ be the MLE of θ. By Taylor expansion we have

$$0=\frac{\partial}{\partial\theta}\log L(\hat{\theta}_{ML})$$

$$=\frac{\partial}{\partial\theta}\log L(\theta)+\left\{\frac{\partial^2}{\partial\theta^2}\log L(\theta)\right\}(\hat{\theta}_{ML}-\theta)+\frac{1}{2}\left\{\frac{\partial^3}{\partial\theta^3}\log L(\theta^*)\right\}(\hat{\theta}_{ML}-\theta)^2,$$

where $|\theta^*-\theta|\leq|\hat{\theta}_{ML}-\theta|$.

Putting

$$T_n=\sqrt{T}(\hat{\theta}_{ML}-\theta)\,;$$

$$Z_1(\theta)=\frac{1}{\sqrt{T}}\frac{\partial}{\partial\theta}\log L(\theta)\,;$$

$$Z_2(\theta)=\frac{1}{\sqrt{T}}\left\{\frac{\partial^2}{\partial\theta^2}\log L(\theta)-E_\theta\left[\frac{\partial^2}{\partial\theta^2}\log L(\theta)\right]\right\},$$

we obtain

$$(2)\qquad 0=\frac{1}{\sqrt{T}}\frac{\partial}{\partial\theta}\log L(\theta)+\frac{1}{T}\left\{\frac{\partial^2}{\partial\theta^2}\log L(\theta)\right\}T_n+\frac{1}{2T\sqrt{T}}\left\{\frac{\partial^3}{\partial\theta^3}\log L(\theta^*)\right\}T_n{}^2.$$

$$=Z_1(\theta)+\frac{1}{\sqrt{T}}Z_2(\theta)T_n-I(\theta)T_n+\frac{1}{2T\sqrt{T}}\left\{\frac{\partial^3}{\partial\theta^3}\log L(\theta^*)\right\}T_n{}^2.$$

Since $(1/T)(\partial^3/\partial\theta^3)\log L(\theta)$ converges in probability to $-3J(\theta)-K(\theta)$ it follows from (2) that

$$0=Z_1(\theta)-\left\{I(\theta)-\frac{1}{\sqrt{T}}Z_2(\theta)\right\}T_n-\frac{1}{2\sqrt{T}}\{3J(\theta)+K(\theta)\}T_n{}^2.$$

Hence we have

$$T_n=\frac{Z_1(\theta)}{I(\theta)}+\frac{Z_1(\theta)Z_2(\theta)}{I(\theta)^2\sqrt{T}}-\frac{3J(\theta)+K(\theta)}{2I(\theta)^3\sqrt{T}}Z_1(\theta)^2+o_p\left(\frac{1}{\sqrt{T}}\right).$$

Therefore we have established:

Theorem 1. Under the conditions (i)～(iv)

$$\sqrt{T}(\hat{\theta}_{ML}-\theta)=\frac{Z_1(\theta)}{I(\theta)}+\frac{Z_1(\theta)Z_2(\theta)}{I(\theta)^2\sqrt{T}}-\frac{3J(\theta)+K(\theta)}{2I(\theta)^3\sqrt{T}}Z_1(\theta)^2+o_p\left(\frac{1}{\sqrt{T}}\right).$$

We modify $\hat{\theta}_{ML}$ to be second order AMU and denote it by $\hat{\theta}_{ML}{}^*$. By a similar way as the i. i. d. case ([3]) we have the following:

Theorem 2. Under the conditions (i)～(iv) $\hat{\theta}_{ML}{}^*$ is second order asymptotically efficient.

We further assume that U_t $(t=1, 2, \ldots)$ and X_0 are normally distributed and $\sigma^2=1/(1-\theta^2)$ in the process (1). Then it is easily seen that the process (1) is stationary. Then the joint density of $(X_0, X_1, \ldots, X_T)$ is given by

$$L(\theta)=L(\theta : x_0, x_1, \ldots, x_T)$$

$$=\frac{1}{(2\pi)^{(T+1)/2}}\sqrt{1-\theta^2}\exp\left\{-\frac{1}{2}\sum_{t=1}^{T}(x_t-\theta x_{t-1})^2-\frac{1}{2}(1-\theta^2)x_0{}^2\right\}.$$

Since

$$\frac{\partial\log L}{\partial\theta}=\sum_{t=1}^{T}(X_t-\theta X_{t-1})X_{t-1}+\theta X_0{}^2-\frac{\theta}{1-\theta^2}\,;$$

$$\frac{\partial^2\log L}{\partial\theta^2}=-\sum_{t=1}^{T}X_{t-1}{}^2+X_0{}^2-\frac{1-\theta^2}{(1-\theta^2)^2},$$

it follows by the stationarity of $\{X_t\}$ that

$$Z_1(\theta)=\frac{1}{\sqrt{T}}\sum_{t=1}^{T}U_tX_{t-1}+\frac{1}{\sqrt{T}}\theta X_0{}^2-\frac{\theta}{1-\theta^2}\frac{1}{\sqrt{T}}\,;$$

$$Z_2(\theta)=\frac{1}{\sqrt{T}}\left(-\sum_{t=2}^{T}X_{t-1}^2+\frac{T}{1-\theta^2}\right);$$

$$I(\theta)=\frac{1}{1-\theta^2}+o(1).$$

By Theorem 1 we have

$$\begin{aligned}
&\sqrt{T}(\hat{\theta}_{ML}-\theta)\\
&=\frac{1}{I(\theta)\sqrt{T}}\sum_{t=1}^{T}U_tX_{t-1}+\frac{\theta}{I(\theta)\sqrt{T}}X_0^2-\frac{\theta}{(1-\theta^2)I(\theta)\sqrt{T}}+\frac{1}{I(\theta)^2\sqrt{T}}\\
&\quad\cdot\left[\left\{\frac{1}{\sqrt{T}}\sum_{t=1}^{T}U_tX_{t-1}+\frac{1}{\sqrt{T}}\theta X_0^2-\frac{\theta}{(1-\theta^2)\sqrt{T}}\right\}\right.\\
&\quad\left.\cdot\left\{\frac{1}{\sqrt{T}}\left(-\sum_{t=2}^{T}X_{t-1}^2+\frac{T-1}{1-\theta^2}\right)\right\}\right]\\
&\quad+\frac{3J(\theta)+K(\theta)}{2I(\theta)^3\sqrt{T}}\left\{\frac{1}{\sqrt{T}}\sum_{t=1}^{T}U_tX_{t-1}+\frac{1}{\sqrt{T}}\theta X_0^2-\frac{\theta}{(1-\theta^2)\sqrt{T}}\right\}^2+o_p\left(\frac{1}{\sqrt{T}}\right)\\
&=\frac{1-\theta^2}{\sqrt{T}}\sum_{t=1}^{T}U_tX_{t-1}+\frac{(1-\theta^2)\theta}{\sqrt{T}}X_0^2-\frac{\theta}{\sqrt{T}}\\
&\quad+\frac{(1-\theta^2)^2}{T\sqrt{T}}\left\{-\left(\sum_{t=1}^{T}X_{t-1}^2\right)\sum_{t=1}^{T}U_tX_{t-1}-\theta X_0^2\sum_{t=1}^{T}X_{t-1}^2+\frac{\theta}{1-\theta^2}\sum_{t=1}^{T}X_{t-1}^2\right.\\
&\quad\left.+\frac{T}{1-\theta^2}\sum_{t=1}^{T}U_tX_{t-1}+\frac{\theta(T-1)}{1-\theta^2}X_0^2-\frac{\theta(T-1)}{(1-\theta^2)^2}\right\}\\
&\quad+\frac{(1-\theta^2)^3\{3J(\theta)+K(\theta)\}}{2T\sqrt{T}}\left\{\left(\sum_{t=1}^{T}U_tX_{t-1}\right)^2+\theta^2X_0^2+\frac{\theta^2}{(1-\theta^2)^2}\right.\\
&\quad\left.+2\theta X_0^2\sum_{t=1}^{T}U_tX_{t-1}-\frac{2\theta^2X_0^2}{1-\theta^2}-\frac{2\theta}{1-\theta^2}\sum_{t=1}^{T}U_tX_{t-1}\right\}+o_p\left(\frac{1}{\sqrt{T}}\right).
\end{aligned}$$

Hence we obtain

$$\begin{aligned}
(2)\qquad \sqrt{T}(\hat{\theta}_{ML}-\theta)=&-\frac{2\theta}{\sqrt{T}}+\frac{2\theta(1-\theta^2)}{\sqrt{T}}X_0^2+\frac{2(1-\theta^2)}{\sqrt{T}}\sum_{t=1}^{T}U_tX_{t-1}\\
&-\frac{(1-\theta^2)^2}{T\sqrt{T}}\left(\sum_{t=1}^{T}X^2_{t-1}\right)\sum_{t=1}^{T}U_tX_{t-1}\\
&+\frac{(1-\theta^2)^3\{3J(\theta)+K(\theta)\}}{2T\sqrt{T}}\left(\sum_{t=1}^{T}U_tX_{t-1}\right)^2+o_p\left(\frac{1}{\sqrt{T}}\right).
\end{aligned}$$

Since $E(U_t^3)=0$, it is seen that the fifth term of the right-hand side of (2) does not affect the asymptotic distribution of $\hat{\theta}_{ML}$ up to order $T^{-1/2}$ ([1]).

Next it is shown that the least squares estimator

$$\hat{\theta}_{LS}=\frac{\sum_{t=1}^{T}X_tX_{t-1}}{\sum_{t=1}^{T}X_{t-1}^2}$$

is also second order asymptotically efficient. It is seen that $\hat{\theta}_{LS}$ is $\{\sqrt{T}\}$-consistent.

$$E_\theta\left(\sum_{t=1}^{T}X_{t-1}^2\right)=\frac{T}{1-\theta^2}=\mu \text{ (say)};$$

$$\sum_{t=1}^{T}X_tX_{t-1}-\theta\sum_{t=1}^{T}X_{t-1}^2=\sum_{t=1}^{T}(X_t-\theta X_{t-1})X_{t-1}=\sum_{t=1}^{T}U_tX_{t-1},$$

it follows that

$$\hat{\theta}_{LS}-\theta=\frac{\sum_{t=1}^{T} U_t X_{t-1}}{\sum_{t=1}^{T} X_{t-1}^2}$$

$$=\frac{\sum_{t=1}^{T} U_t X_{t-1}}{\left(1+\frac{\sum_{t=1}^{T} X_{t-1}^2-\mu}{\mu}\right)}$$

$$=\frac{1}{\mu}\left(\sum_{t=1}^{T} U_t X_{t-1}\right)\left\{1-\frac{\sum_{t=1}^{T} X_{t-1}^2-\mu}{\mu}+\frac{\left(\sum_{t=1}^{T} X_{t-1}^2-\mu\right)^2}{\mu^2}\right\}+o_p\left(\frac{1}{\sqrt{T}}\right)$$

$$=\frac{1}{\mu}\sum_{t=1}^{T} U_t X_{t-1}+\frac{1}{\mu^2}\left(\sum_{t=1}^{T} U_t X_{t-1}\right)\left(\sum_{t=1}^{T} X_{t-1}^2-\mu\right)$$

$$+\frac{1}{\mu^3}\left(\sum_{t=1}^{T} U_t X_{t-1}\right)\left(\sum_{t=1}^{T} X_{t-1}^2-\mu\right)^2+o_p\left(\frac{1}{\sqrt{T}}\right)$$

$$=\frac{1-\theta^2}{T}\sum_{t=1}^{T} U_t X_{t-1}-\frac{(1-\theta^2)^2}{T^2}\left(\sum_{t=1}^{T} X_{t-1}U_t\right)\left(\sum_{t=1}^{T} X_{t-1}^2-\frac{T}{1-\theta^2}\right)$$

$$+o_p\left(\frac{1}{\sqrt{T}}\right)$$

Hence we obtain

$$\sqrt{T}(\hat{\theta}_{LS}-\theta)=\frac{2(1-\theta^2)}{\sqrt{T}}\sum_{t=1}^{T} U_t X_{t-1}-\frac{(1-\theta^2)^2}{T\sqrt{T}}\left(\sum_{t=1}^{T} X_{t-1}^2\right)\left(\sum_{t=1}^{T} U_t X_{t-1}\right)+o_p\left(\frac{1}{\sqrt{T}}\right). \tag{3}$$

We modify $\hat{\theta}_{LS}$ to be second order AMU and denote it by $\hat{\theta}_{LS}^*$. Since $E(X_0^2 U_t X_{t-1})=0$ $(t=1, 2, \ldots, T)$, it follows that the second term with X_0 of the right-hand side of (2) does not affect the asymptotic distribution of the MLE up to the order $T^{-1/2}$, that is, the second order asymptotic efficiency of the MLE.

By Theorems 1, 2 and (2) and (3) we have established:

Theorem 3. $\hat{\theta}_{LS}^*$ is second order asymptotically efficient.

Remark: Although it is seen from the above that the second order asymptotic efficiency of the MLE does not depend on the initial condition of X_0, it is noted that the second term with X_0 of the right-hand side of (2) may affect the m-th $(m\geqq 3)$ order asymptotic efficiency of the MLE.

Acknowledgment

The author wishes to thank Professor K. Takeuchi of Tokyo University for suggestive discussions.

References

[1] Akahira, M., "A note on the second order asymptotic efficiency of estimators in an autoregressive process," Rep. Univ. Electro-Comm. **26**, 143–149 (1975).

[2] Akahira, M., "Asymptotic theory for estimation of location in non-regular cases, I: Order of

convergence of consistent estimators," Rep. Stat. Appl. Res., JUSE, **22**, 8-26 (1975).

[3] Takeuchi, K. and Akahira, M., "On the second order asymptotic efficiencies of estimators," Proc. of the Third Japan-USSR Symp. on Prob. Theory, Lecture Notes in Math. **550**, Springer Verlag, 604-638 (1976).