# Some aspects of stability in time series small sample case

**Hocine Fellag**

Laboratory of Pure and Applied Mathematics, Mouloud Mammeri University of Tizi-Ouzou, Tizi-Ouzou, 15000, Algeria

**Abstract.** In this paper, we consider the problem of stability of the estimation in autoregressive models for the finite sample case. A Monte Carlo comparison of the least square estimator and the Hurwicz estimator is performed in various contaminated models. The paper shows that, the least square estimator is very sensitive to contamination and, despite the median-unbiasedness of the Hurwicz estimator, an optimal and stable estimator in small sample case is to be constructed.

**Résumé.** Dans cet article, nous considérons le problème de la stabilité de l'estimation des paramètres d'un processus autoregressif dans le cas des échantillons finis. Une comparaison par les méthodes de Monte Carlo des estimateurs des moindres carrés et d'Hurwicz est ainsi présentée pour divers types de contamination. Cette étude montre que l'estimateur des moindres carrés est très sensible aux perturbations et que, même si l'estimateur d'Hurwicz est sans biais par rapport  la médiane, un estimateur stable reste encore à construire dans le cas fini.

## 1. Introduction

The concept of robustness has been introduced for the first time by Box and Andersen [6]. In their pioneer paper, they called "robust" a procedure which is insensitive to changes, of a magnitude likely to occur, in practice, in extraneous factors. For Tukey [22], an efficient robust estimate is one whose efficiency is "high" at a variety of strategically chosen distributions. Huber [16] constructed a min-max robust estimate obtained after minimization of the maximum of the asymptotic variance over certain families of distributions. Hampel [15] proposed the notion of qualitative robustness in terms of equicontinuity of some sequences of maps of distributions. Box and Tiao [7] defined two distinct concepts of robustness of statistical techniques in terms of insensitivity to non-normality.

Generally, in statistics, when we are confronted with a statistical problem, one has to adopt a model, establish aims and criteria and calculate optimal solution. However, we must not stop at this point. We build a supermodel, taking account plausible departures from the original model and in the context of the supermodel, we study the performance of our optimal solution, to get the better performance and correct the solution if it is needed. Zieliński [24] introduced a new idea of robustness using a robustness function defined as a real valued function on a parameter space. This function describes the oscillation or the stability of the studied property in the supermodel. In this paper, we consider the stability of the estimation in autoregressive models when the sample size is finite. We show that, despite the popularity of estimators of autoregressive models, when the sample size is small, a new stable estimator of the autoregressive parameter is needed. The paper is outlined as follows.In the second section, we present some references on the estimation of autoregressive models. The third section is devoted to the presentation of some aspects of stability in this kind of models. A comparative study of two interesting estimators is presented in the last section.

Hocine Fellag: hfellag@mail.ummto.dz

## 2. Inference in small sample case

The problem of the estimation the autoregressive model is studied by many authors in the literature. Anderson [3] studied the problem of exact sampling distribution of the least square estimator (LSE) of this autoregressive model. He showed the difficulty to obtain the explicit formula of estimators in small sample case. Reeves [20] proposed a method to obtain values of the distribution of the LSE in the Gaussian case. Ensor and Newton [12] proposed a recursive algorithm for LSE in AR(p) process using Yule Walker estimation. Mikulski and Monsour [19] discussed the optimality of the mean likelihood estimator (MLE) estimator of the autoregressive parameter. Andrews [4] considered the problem of the standard estimators like LSE which are downward biased. He said that this estimator is a misleading indicator of the true value of the parameter. So, he proposed median-unbiased estimators. Suppose that $m(\rho)$ is median function . Andrews proposed

$$\hat{\rho}_{med} = \begin{cases} -1 & \text{if } \hat{\rho}_{LS} \leq -1 \\ m^{-1}(\rho) & \text{if } m(-1) < \hat{\rho}_{LS} \leq m(1) \\ 1 & \text{if } \hat{\rho}_{LS} > m(1) \end{cases}$$

is a median unbiased estimator of $\rho$ with

$$m(1) = \lim_{\rho \to -1} m(\rho)$$

But, there are some drawbacks : this correction method has no explicit formula and does not apply to AR(p) when $p > 1$. The distribution of innovations must be known exactly.

Akkaya and Tiku [1] considered an AR(1) process with iid innovations having specific non-normal, but symmetrical distributions. The small sample estimation case was considered. The authors derived the modified maximum likelihood estimator as solutions of equations after linearizing functions. The problem is that, explicit solutions are not possible. However, they showed that their estimator is more efficient than the LSE estimator.

For the case where the innovations are distributed according to exponential distribution, Andel [2] obtained an approximation of the moments of the MLE estimator $\hat{\rho}_{MLE} = \min_{2 \leq t \leq n} X_t / X_{t-1}$. Cabrera and Watson [8] proposed in iid case some simulation procedures in order to reduce the mean bias and the median bias in parametric estimation. Luger [17] proposed a median-unbiased procedure when the innovations are distributed according to ARCH model. He could prove that the Hurwicz estimator remains more general assumptions than in the paper of Zieliński [25], without assuming statistical independence. For finite samples, a computationally efficient method was then proposed.

## 3. Some aspects of stability in autoregressive models

To study the problem of robustness in autoregressive models, many computational and approximation methods are given in the literature. Campbell [9] proposed a recursive computation of M-estimators of the parameters of autoregressive processes. Haddad [14] constructed an asymptotic robust estimator of the AR(1) model. Also, some particular aspects are used in Berkoun et al. [5] for the case of the independence test ($H_0 : \rho = 0$ vs $H_1 : \rho \neq 0$). Carstensen [10] investigated the unit root problems in small sample case. He showed that the Dickey-Fuller test based on the least square estimator has lower power. In his paper, he proved that the robustness depends on the type of non-normality at hand explaining then why there is no uniformly most powerful unit root test. To present the problematic of stability in time series, let us present, as an illustration, the following estimators.

### 3.1. The Least Square Estimator

Consider the stationary first-order autoregressive process of the form

$$X_t = \rho X_{t-1} + Y_t, \ t = \ldots, -1, 0, 1, \ldots$$

where $Y_t, t = \ldots, -1, 0, 1, \ldots$, are iid N(0,1).

Suppose that all what we observe is a segment of the process

$$X_1, X_2, \ldots, X_n, \ n \text{ fixed},$$

and take as an estimator of $\rho$ the classical LSE

$$\hat{\rho}_{LS} = \frac{X_1 X_2 + \ldots + X_{n-1} X_n}{X_1^2 + \ldots + X_{n-1}^2}$$

The main reason for our choice of $\hat{\rho}_{LS}$ is that, opposite to the maximum likelihood, mean likelihood, or other estimators, $\hat{\rho}_{LS}$ has that practical advantage that it can be easily calculated, also recursively when it is needed.

Now suppose that due to measurement techniques gross errors may appear so that $Y_1, Y_2, \ldots, Y_n$ are distributed acording to the Tukey Model $T(\epsilon, \sigma)$, i.e.

$$T(\epsilon, \sigma) = (1 - \epsilon)N(0, 1) + \epsilon N(0, \sigma^2)$$

where $\epsilon$ and $\sigma$ are some constants.

Sawa [21] gave the following formula of the bias of the LSE in gaussian case. Unfortunately his formula was too complicated to admit theoretical investigations though numerical computations can be performed without too much difficulty. Fellag and Zieliński [13] proposed to perform the bias of this estimator in contaminated models. Using their formula, the authors said that, the bias under Tukey contamination does not change monotonically either when $\epsilon$ varies in the interval $[0, 1/2]$ or when $\sigma$ is changing in $[1, \infty)$. The lack of monotonicity may be considered as a rather unexpected fact which is worthy of a more detailed study. Hence, one can notice that the variability of the LSE estimator in small samples is very high which makes the estimator rather useless.

### 3.2. The Hurwicz Estimator

In the previous subsection, the well known least square estimator has appeared to be rather unsatisfactory for two reasons :it is biased and it is highly inefficient in small samples. The problem of efficiency for small $n$ is connected with the fact that $\hat{\rho}_{LS}$ has the heavy tail Cauchy distribution if $n = 2$ and the tails get slimmer very slowly when $n$ is growing. E.g. if $\rho = 0$ then $P\{|\hat{\rho}_{LS}| > 1\} = 0.19$ if $n = 3$ and $P\{|\hat{\rho}_{LS}| > 0.5\} = 0.10$ if $n = 10$ (Fellag and Zieliński [13]).

An attractive competitor for $LSE$ is the Hurwicz estimator $\hat{\rho}_{HUR}$ defined as follows: for the same given segment of the process, consider the sequence of ratios

$$X_2/X_1, X_3/X_2, \ldots, X_n/X_{n-1}.$$

To avoid some technicalities we assume that $n$ is even so that the median of the ratios is uniquely determined.

Then $\hat{\rho}_{HUR}$ is defined as

$$\hat{\rho}_{HUR} = \text{Med} \left\{ \frac{X_2}{X_1}, \frac{X_3}{X_2}, \ldots, \frac{X_n}{X_{n-1}} \right\}$$

where $\text{Med}(\xi_1, \xi_2, \ldots, \xi_N)$ denotes the sample median of the observations $\xi_1, \xi_2, \ldots, \xi_N$. , i.e. if $\xi_{1:N} \leq \xi_{2:N} \leq \ldots \leq \xi_{N:N}$ and $N = 2k - 1$ then $\text{Med}(\xi_1, \xi_2, \ldots, \xi_N) = \xi_{k:N}$.

Zieliński [25] proved that $\hat{\rho}_{HUR}$ is a median-unbiased estimator of $\rho$ whenever the medians of the innovations $Y_1, Y_2, \ldots, Y_n$ are equal to zero. The result is rather general: the innovations do not need to be identically distributed, they may have heavy tails, and the median-unbiasedness of $\hat{\rho}_{HUR}$ remains to hold also under $\epsilon$-contamination of the innovations when the contaminants are distributed symmetrically around zero.

The problem is what is the price for such nice properties of the Hurwicz estimator. In what follows, we present results of an exhaustive simulation study of Mean Absolute Deviation ($MAD$) of $\hat{\rho}_{HUR}$ and compare it with that of the least square estimator $\hat{\rho}_{LS}$. However, it appears in small samples that $MAD$ of $LSE$ is very unstable if the distribution of innovations is heavy tailed. For that reason, we decided to calculate also the Pitman Closeness Criterion ($PCC$) defined as :

$$PCC(\rho) = P_\rho\{|\hat{\rho}_{HUR} - \rho| > |\hat{\rho}_{LS} - \rho|\}$$

which is more stable. We report some results concerning the $PCC$ for both estimators. To a surprise it appeared that despite the median-unbiasedness of $\hat{\rho}_{HUR}$, the estimator $\hat{\rho}_{LS}$ performs better in Gaussian models and in models with innovations whose distributions have not very heavy tails.

### 3.3. Gaussian Model

We define the $MAD$ of estimators and we introduce an appropriate notation as follows

$$MAD_{LS}(\rho) = E_\rho|\hat{\rho}_{LS} - \rho|, \quad MAD_{HUR}(\rho) = E_\rho|\hat{\rho}_{HUR} - \rho|$$

Despite the median-unbiasedness of $\hat{\rho}_{HUR}$ and despite the bias and a large variability of $\hat{\rho}_{LS}$, it appears that $MAD_{LS}(\rho) < MAD_{HUR}(\rho)$. In the investigated cases, we have observed that $MAD$ is a symmetric (around zero) function of $\rho$. Consequently, we confine ourselves to study $MAD's$ for $\rho$ in $[0,1)$ only. The basic tool we have used is Monte Carlo study. The random numbers were generated using the program ULTRA version 1.01 (Marsaglia and Zaman [18]). The Gaussian distributions were generated using the ratio–of–uniform–method (Devroye [11]). Some numerical results of $MAD$ for $n = 4, 20$ are presented in TABLE I. To get an idea of the error of simulation, we present the results from 3 independent series of $10^6$ runs.

| n | $\rho$ | $MAD_{LS}(\rho)$ | | | $MAD_{HUR}(\rho)$ | | |
|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 1 | 2 | 3 |
| | 0 | 0.439 | 0.437 | 0.474 | 0.863 | 0.865 | 0.865 |
| | 0.25 | 0.467 | 0.468 | 0.468 | 0.840 | 0.849 | 0.842 |
| | 0.5 | 0.456 | 0.456 | 0.456 | 0.771 | 0.773 | 0.770 |
| | 0.75 | 0.426 | 0.426 | 0.426 | 0.644 | 0.646 | 0.643 |
| 4 | 0.9 | 0.355 | 0.356 | 0.354 | 0.492 | 0.492 | 0.490 |
| | 0.95 | 0.295 | 0.296 | 0.295 | 0.397 | 0.394 | 0.393 |
| | 0.99 | 0.175 | 0.176 | 0.175 | 0.224 | 0.225 | 0.225 |
| | 0.9999 | 0.029 | 0.029 | 0.029 | 0.036 | 0.036 | 0.036 |
| | 0 | 0.177 | 0.177 | 0.177 | 0.282 | 0.282 | 0.281 |
| | 0.25 | 0.172 | 0.173 | 0.173 | 0.275 | 0.275 | 0.275 |
| | 0.5 | 0.159 | 0.159 | 0.160 | 0.252 | 0.256 | 0.251 |
| | 0.75 | 0.134 | 0.134 | 0.134 | 0.202 | 0.202 | 0.202 |
| 20 | 0.9 | 0.107 | 0.107 | 0.107 | 0.148 | 0.148 | 0.148 |
| | 0.95 | 0.091 | 0.091 | 0.091 | 0.119 | 0.119 | 0.119 |
| | 0.99 | 0.059 | 0.059 | 0.059 | 0.071 | 0.071 | 0.071 |
| | 0.9999 | 0.011 | 0.011 | 0.011 | 0.013 | 0.013 | 0.013 |

**Table I.** Simulated values of $MAD_{LS}(\rho)$ and $MAD_{HUR}(\rho)$.

The table I should be read as follows. In the first of three series of $10^6$ simulations, the observed value of $MAD_{LS}(\rho)$ for $n = 4$ and $\rho = 0$ is equal to 0.439, that in the second series is equal to 0.437, and so on. To compare the estimators, we calculated the ratios defined as

$$RATIO(\rho) = MAD_{HUR}(\rho)/MAD_{LS}(\rho)$$

as well as $PCC(\rho)$. The results for 3 independent series of $10^6$ simulations each are presented in TABLE II.

| n | $\rho$ | $RATIO(\rho)$ | | | $PCC(\rho)$ | | |
|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 1 | 2 | 3 |
| | 0 | 1.821 | 1.827 | 1.824 | 0.674 | 0.674 | 0.675 |
| | 0.25 | 1.792 | 1.812 | 1.796 | 0.673 | 0.673 | 0.673 |
| | 0.5 | 1.692 | 1.695 | 1.690 | 0.665 | 0.665 | 0.665 |
| | 0.75 | 1.510 | 1.515 | 1.510 | 0.638 | 0.638 | 0.639 |
| 4 | 0.9 | 1.382 | 1.383 | 1.382 | 0.614 | 0.614 | 0.614 |
| | 0.95 | 1.342 | 1.331 | 1.332 | 0.603 | 0.603 | 0.603 |
| | 0.99 | 1.275 | 1.274 | 1.280 | 0.592 | 0.591 | 0.592 |
| | 0.9999 | 1.217 | 1.223 | 1.225 | 0.587 | 0.587 | 0.588 |
| | 0 | 1.594 | 1.594 | 1.592 | 0.680 | 0.679 | 0.680 |
| | 0.25 | 1.593 | 1.591 | 1.589 | 0.680 | 0.679 | 0.679 |
| | 0.5 | 1.579 | 1.578 | 1.574 | 0.678 | 0.678 | 0.678 |
| | 0.75 | 1.511 | 1.509 | 1.508 | 0.674 | 0.673 | 0.673 |
| 20 | 0.9 | 1.387 | 1.388 | 1.387 | 0.663 | 0.662 | 0.663 |
| | 0.95 | 1.303 | 1.303 | 1.302 | 0.648 | 0.647 | 0.647 |
| | 0.99 | 1.201 | 1.199 | 1.200 | 0.613 | 0.613 | 0.613 |
| | 0.9999 | 1.189 | 1.187 | 1.191 | 0.609 | 0.608 | 0.608 |

**Table II.** Simulated values of $RATIO(\rho)$ and $PCC(\rho)$

It can be seen that, despite the median-unbiasedness of $\hat{\rho}_{HUR}$, the $LSE$ estimator performs better. One can hope that $LSE$ modified according to the Andrew's [4] suggestion in such a way that it becomes a median-unbiased estimator would perform even better.

*3.4. Symmetric innovations models*

The Hurwicz estimator $\hat{\rho}_{HUR}$ is median-unbiased under all innovations which are distributed symmetrically around zero. It is interesting to see how efficient (in comparison with $LSE$) it is in such situations. To study the problem we have chosen the following distributions:

1. Symmetric stable distributions with the characteristic functions of the form $\exp(-|t|^\gamma)$ for $\gamma \in [1, 2]$. For $\gamma = 2$ we have the normal distribution and for $\gamma = 1$ the heavy tailed Cauchy distribution with probability density function $[\pi(1+x^2)]^{-1}$;
2. Laplace distribution with probability density function (pdf) $(1/2)\exp(-|x|)$;
3. Logistic distribution with cumulative distribution function (cdf) $(1 + e^{-x})^{-1}$;
4. Uniform distribution $U(-1/2, 1/2)$.

To simulate the sample (2) we have to start with simulating $X_0$. The basic formula for $X_0$ is

$$X_0 = \sum_{j=0}^{\infty} \rho^j Y_{-j}$$

If the innovations $Y_j$ are normal $N(0,1)$ then $X_0$ is normal $N(0,\sigma^2)$ with $\sigma^2 = (1 - \rho^2)^{-1}$. If $Y_j$ have a stable distribution with the characteristic function $\exp(-|t|^\gamma)$ then $X_0$ has the stable distribution with characteristic function $\exp(-|t|^\gamma/(1 - \rho^\gamma))$. Hence to get $X_0$ one has to simulate a random variable according to the characteristic function $\exp(-|t|^\gamma)$ and divide the result by the scale parameter equal to $(1 - \rho^\gamma)^{1/\gamma}$. The symmetric stable distributions are simulated using the method given by Weron [23]. To obtain $X_0$ in the Laplace, Logistic, and Uniform cases we used the Central Limit Theorem: first we simulated a normal $N(0,1)$ variate and then divide the result by the appropriate scale parameter: $\sqrt{(1-\rho^2)/2}$ for Laplace, $\sqrt{3(1-\rho^2)}/\pi$ for Logistic, and $\sqrt{12(1-\rho^2)}$ for uniform distribution. In TABLE III and TABLE IV, mean values of 3 independent series of $10^6$ runs of respectively $RATIO(\rho)$ and $PCC(\rho)$ are presented.

| n | $\rho$ | Stable Distributions | | | | | Laplace | Logistic | Uniform |
|---|---|---|---|---|---|---|---|---|---|
| | | 2 normal | 1.75 | 1.5 | 1.25 | 1 Cauchy | | | |
| | 0 | 1.824 | 1.769 | 1.600 | 1.292 | 0.447 | 1.888 | 1.832 | 1.781 |
| | 0.25 | 1.800 | 1.711 | 1.530 | 1.146 | 0.360 | 1.713 | 1.777 | 1.785 |
| | 0.5 | 1.693 | 1.585 | 1.389 | 1.001 | 0.329 | 1.524 | 1.635 | 1.758 |
| | 0.75 | 1.512 | 1.402 | 1.219 | 0.862 | 0.294 | 1.323 | 1.428 | 1.662 |
| 4 | 0.9 | 1.383 | 1.279 | 1.104 | 0.801 | 0.272 | 1.209 | 1.298 | 1.583 |
| | 0.95 | 1.335 | 1.224 | 1.041 | 0.747 | 0.258 | 1.163 | 1.256 | 1.559 |
| | 0.99 | 1.277 | 1.152 | 0.955 | 0.682 | 0.229 | 1.106 | 1.203 | 1.474 |
| | 0.9999 | 1.222 | 1.061 | 0.849 | 0.537 | 0.212 | 1.033 | 1.151 | 1.417 |
| | 0 | 1.593 | 1.576 | 1.536 | 1.448 | 1.089 | 1.349 | 1.536 | 1.852 |
| | 0.25 | 1.591 | 1.547 | 1.463 | 1.302 | 0.879 | 1.307 | 1.522 | 1.865 |
| | 0.5 | 1.577 | 1.491 | 1.353 | 1.132 | 0.681 | 1.233 | 1.487 | 1.894 |
| | 0.75 | 1.509 | 1.381 | 1.196 | 0.936 | 0.500 | 1.131 | 1.407 | 1.863 |
| 20 | 0.9 | 1.388 | 1.236 | 1.030 | 0.762 | 0.378 | 1.011 | 1.278 | 1.733 |
| | 0.95 | 1.303 | 1.141 | 0.928 | 0.660 | 0.326 | 0.937 | 1.187 | 1.641 |
| | 0.99 | 1.200 | 1.009 | 0.773 | 0.500 | 0.228 | 0.851 | 1.084 | 1.530 |
| | 0.9999 | 1.189 | 0.919 | 0.622 | 0.337 | 0.115 | 0.813 | 1.070 | 1.526 |

**TABLE III.** Simulated values of $RATIO(\rho)$ for symmetric innovations.

| n | $\rho$ | Stable Distributions | | | | | Laplace | Logistic | Uniform |
|---|---|---|---|---|---|---|---|---|---|
| | | 2 normal | 1.75 | 1.5 | 1.25 | 1 Cauchy | | | |
| | 0 | 0.675 | 0.663 | 0.648 | 0.628 | 0.604 | 0.635 | 0.664 | 0.714 |
| | 0.25 | 0.673 | 0.660 | 0.642 | 0.618 | 0.582 | 0.627 | 0.661 | 0.712 |
| | 0.5 | 0.665 | 0.648 | 0.623 | 0.588 | 0.535 | 0.603 | 0.646 | 0.713 |
| | 0.75 | 0.639 | 0.616 | 0.585 | 0.537 | 0.467 | 0.561 | 0.609 | 0.697 |
| 4 | 0.9 | 0.614 | 0.587 | 0.548 | 0.491 | 0.412 | 0.529 | 0.582 | 0.681 |
| | 0.95 | 0.603 | 0.573 | 0.530 | 0.469 | 0.388 | 0.514 | 0.571 | 0.674 |
| | 0.99 | 0.592 | 0.556 | 0.507 | 0.443 | 0.365 | 0.495 | 0.560 | 0.666 |
| | 0.9999 | 0.588 | 0.544 | 0.493 | 0.431 | 0.358 | 0.478 | 0.553 | 0.693 |
| | 0 | 0.680 | 0.662 | 0.643 | 0.623 | 0.602 | 0.584 | 0.657 | 0.762 |
| | 0.25 | 0.680 | 0.658 | 0.633 | 0.601 | 0.556 | 0.575 | 0.654 | 0.762 |
| | 0.5 | 0.678 | 0.649 | 0.611 | 0.559 | 0.483 | 0.560 | 0.649 | 0.763 |
| | 0.75 | 0.674 | 0.632 | 0.575 | 0.493 | 0.381 | 0.537 | 0.640 | 0.765 |
| 20 | 0.9 | 0.663 | 0.608 | 0.529 | 0.421 | 0.284 | 0.510 | 0.622 | 0.760 |
| | 0.95 | 0.647 | 0.583 | 0.490 | 0.369 | 0.226 | 0.483 | 0.729 | 0.749 |
| | 0.99 | 0.613 | 0.531 | 0.421 | 0.293 | 0.166 | 0.432 | 0.555 | 0.730 |
| | 0.9999 | 0.608 | 0.504 | 0.387 | 0.266 | 0.152 | 0.404 | 0.548 | 0.750 |

**TABLE IV.** Simulated values of $PCC(\rho)$ for symmetric innovations.

To assess the error of calculations, in TABLE V we present the results from 3 independent series of $10^6$ runs each when the innovations are distributed according to the Cauchy distribution.

| | | $MAD_{LS}(\rho)$ | | | $MAD_{HUR}(\rho)$ | | | $RATIO(\rho)$ | | |
|---|---|---|---|---|---|---|---|---|---|---|
| n | $\rho$ | 1 | 2 | 3 | 1 | 2 | 3 | 1 | 2 | 3 |
| | 0 | 2.074 | 1.865 | 9.484 | 1.197 | 1.187 | 1.203 | 0.577 | 0.636 | 0.127 |
| | 0.25 | 1.797 | 1.661 | 9.627 | 0.860 | 0.851 | 0.851 | 0.478 | 0.512 | 0.088 |
| | 0.5 | 1.421 | 1.337 | 8.668 | 0.632 | 0.629 | 0.630 | 0.445 | 0.470 | 0.072 |
| | 0.75 | 0.941 | 0.982 | 6.449 | 0.393 | 0.396 | 0.395 | 0.418 | 0.404 | 0.061 |
| 4 | 0.9 | 0.560 | 0.570 | 3.465 | 0.212 | 0.214 | 0.213 | 0.378 | 0.375 | 0.061 |
| | 0.95 | 0.363 | 0.372 | 1.918 | 0.123 | 0.129 | 0.130 | 0.357 | 0.346 | 0.068 |
| | 0.99 | 0.115 | 0.130 | 0.428 | 0.036 | 0.037 | 0.036 | 0.317 | 0.285 | 0.085 |
| | 0.9999 | 0.002 | 0.002 | 0.005 | 0.0006 | 0.0007 | 0.0007 | 0.272 | 0.240 | 0.122 |
| | 0 | 0.157 | 0.176 | 0.159 | 0.178 | 0.178 | 0.178 | 1.135 | 1.011 | 1.121 |
| | 0.25 | 0.157 | 0.184 | 0.148 | 0.142 | 0.142 | 0.142 | 0.905 | 0.772 | 0.960 |
| | 0.5 | 0.154 | 0.181 | 0.139 | 0.106 | 0.106 | 0.106 | 0.689 | 0.587 | 0.766 |
| | 0.75 | 0.136 | 0.154 | 0.117 | 0.067 | 0.067 | 0.067 | 0.492 | 0.436 | 0.571 |
| 20 | 0.9 | 0.102 | 0.109 | 0.090 | 0.037 | 0.037 | 0.037 | 0.369 | 0.346 | 0.417 |
| | 0.95 | 0.073 | 0.077 | 0.070 | 0.024 | 0.024 | 0.024 | 0.328 | 0.309 | 0.341 |
| | 0.99 | 0.033 | 0.034 | 0.032 | 0.007 | 0.007 | 0.007 | 0.230 | 0.223 | 0.230 |
| | 0.9999 | 0.001 | 0.001 | 0.001 | 0.0002 | 0.0002 | 0.0002 | 0.116 | 0.114 | 0.115 |

**TABLE V.** Simulated values of $MAD_{LS}(\rho)$, $MAD_{HUR}(\rho)$ and $RATIO(\rho)$ for Cauchy distribution.

From TABLE V, one can notice that, when $\gamma = 1$, the error of simulation of $RATIO(\rho)$ is high. This is due to high instability of $\hat{\rho}_{LS}$ in small samples when the innovations are distributed according to Cauchy distribution. However, if $\gamma > 1$, the variability of $MAD_{LS}(\rho)$ becomes smaller. For example, if $\gamma = 1.25$, n=4 and $\rho = 0$, the $MAD$ of $\hat{\rho}_{LS}$ varies between 0.7595 and 0.8078. Finally, one can say that, the closer $\gamma$ to 1, the better is $\hat{\rho}_{HUR}$ and the closer $\gamma$ to 2, the better is $\hat{\rho}_{LS}$.

### 3.5. Gaussian Model with Tukey Contamination

The behavior of $RATIO(\rho)$ under $\epsilon$-contamination of the Tukey type for the distribution of innovations needs a comment. Some numerical results of simulations are exhibited in TABLE VI.

H. Fellag, Journal Afrika Statistika, Vol. 5, N°8, 2010, page 252–259.

Some aspects of stability in time series small sample case
258

| $\epsilon$ | $\rho = 0.0$ | | | | | | $\rho = 0.95$ | | | | | |
| | $\sigma$ | | | | | | $\sigma$ | | | | | |
| | 1 | 2 | 4 | 8 | 16 | 32 | 1 | 2 | 4 | 8 | 16 | 32 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.001 | 1.82 | 1.82 | 1.82 | 1.82 | 1.81 | 1.80 | 1.32 | 1.32 | 1.32 | 1.32 | 1.31 | 1.30 |
| 0.005 | 1.82 | 1.82 | 1.82 | 1.80 | 1.76 | 1.69 | 1.32 | 1.32 | 1.31 | 1.30 | 1.27 | 1.23 |
| 0.01 | 1.82 | 1.82 | 1.81 | 1.79 | 1.71 | 1.65 | 1.32 | 1.32 | 1.31 | 1.29 | 1.24 | 1.15 |
| 0.05 | 1.82 | 1.82 | 1.80 | 1.71 | 1.47 | 1.14 | 1.32 | 1.31 | 1.27 | 1.17 | 1.02 | 0.81 |
| 0.1 | 1.82 | 1.82 | 1.78 | 1.64 | 1.34 | 1.05 | 1.32 | 1.30 | 1.22 | 1.09 | 0.87 | 0.64 |
| 0.2 | 1.82 | 1.83 | 1.78 | 1.62 | 1.34 | 1.06 | 1.32 | 1.28 | 1.17 | 0.97 | 0.71 | 0.51 |
| 0.3 | 1.82 | 1.83 | 1.82 | 1.70 | 1.49 | 1.27 | 1.32 | 1.29 | 1.15 | 0.95 | 0.71 | 0.51 |
| 0.4 | 1.82 | 1.83 | 1.83 | 1.81 | 1.67 | 1.41 | 1.32 | 1.28 | 1.14 | 0.96 | 0.74 | 0.56 |
| 0.5 | 1.82 | 1.84 | 1.87 | 1.92 | 1.87 | 1.74 | 1.32 | 1.29 | 1.17 | 0.99 | 0.81 | 0.64 |

**TABLE VI.** Simulated values of $\rho$ for n=4 and $\rho = 0, 0.95$.

When $\epsilon$ is fixed, greater $\sigma$ amounts to heavier tails of the mixture and heavy tails do influence the ratio $RATIO(\rho)$ as it has been shown in 2.3. Hence, the ratio is decreasing monotonically in each row of TABLE VI. However, the ratio does not change monotonically with $\epsilon \in (0, 0.5)$ for any fixed $\sigma$. One can suspect that the main reason for the nonmonotonical behavior of the ratio is the kurtosis of the mixture but that explains the phenomenon only partially.

## 4. Conclusion

The LSE estimator is very popular and widely used in applications (regression, prediction, unit roots). It is very sensitive to contamination. It can be seen that, despite the median-unbiasedness of $\hat{\rho}_{HUR}$, the $LSE$ estimator performs better One can hope that $LSE$ modified according to the Andrew's suggestion in such a way that it becomes a median-unbiased estimator would perform better. One can say that an optimal and stable estimator in small sample case is to be constructed.

## References

[1] Akkaya, A.D., Tiku, M.L., 2008. Autoregressive models with short-tailed symmetric distributions. *Statistics* **42**(3), 207-221.

[2] Anděl, J., 1988. On AR(1) process with exponential white noise. *Comm. in Statist., Theory and methods.* **5**(17),1481-1495.

[3] Anderson, R.L., 1942. Distribution of the serial correlation coefficient. *Ann. of Math. Stat.*, **13**, 1-13.

[4] Andrews, D.W.K., 1993. Exactly median unbiased estimation of first order autoregressive / unit root models. *Econometrica.* **1**(61), 139-165.

[5] Berkoun, Y, Fellag, H. and Zielinski, R., 2003. Robust testing serial correlation in AR(1) processes in the presence of a single additive outlier. *Comm. In Statist., Theory and Methods*, 2003, Vol. **32**, issue 8, 1527-1540.

[6] Box,G.E.P. and Andersen, S.L., 1955. Permutation theory in the derivation of robust criteria and the study of departures from asymptions. *J. Royal Stat. Soc.* Ser. B, **17**, 1-34.

[7] Box, G.E.P. and Tiao, G.C., 1964. A note on criterion robustness and inference robustness. *Biometrika* **51**, 169-173.

[8] Cabrera, J. and Watson, G.S., 1997. Simulation methods for mean and median bias reduction in parametric estimation. *J. Statist. Plann. Inference* **57**, 143-152.

[9] Campbell, K., 1982. Recursive computation of m-estimates for the parameters of a finite autoregressive process. *Ann. Stat.* Vol. **10**, 2, 442-453.

[10] Cartensen, K., 2003. The finite sample performance of robust unit root test. *Statistical papers* **44**, 469-482.

[11] Devroye, L., 1986. **Non uniform random variate generation.** Springer Verlag.

[12] Ensor, K.B. and Newton, H.J., 1990. A recursive in order algorithm for least square estimates of a first autoregressive process.*J. Statist. Comput. Simulation.* **37**, 115-126.

[13] Fellag, H. and Zielinski, R., 1996. Bias of the LSE Estimator of the First-Order Autoregressive Model under Tukey Contamination. *Comm. In Statist., Theory and Methods.* **7**(25), 1537-1551.

[14] Haddad, J.N., 2000. On robust estimation in the first-order autoregressive process. *Comm. in statist., Theory and methods.* **1**(29), 45-54.

[15] Hampel, 1971. A general a general qualitative definition of robustness. *Ann. Math. Stat.* **19**, 225-239.

[16] Huber, P.J., 1964. Robust estimation of location parameter. *Ann. Math. Stat.* **35**, 73-101.

[17] Luger, R., 2006. Median unbiased estimation and exact inference methods for first-order autoregressive models with conditional heteroscedasticity of unknown form. *J. Time Ser. Anal.* **1**(27), 119-128.

[18] Marsaglia, G. and Zaman, A., 1991. Program ULTRA, version 1.01.

[19] Mikulsky, P.W. and Monsour, M.J., 1991. Optimality of the maximum likelihood estimator in first-order autoregressive process. *J. Time Ser. Anal.* **3**(12), 237-253.

[20] Reeves, J.E., 1972. The distribution of the maximum likelihood estimator of the parameter in the first-order autoregressive series. *Biometrica* **59**, 387-394.

[21] Sawa, T., 1978. The exact moments of the least square estimator for the autoregressive model. *J. Econometrics*. **8**, 159-172.

[22] Tukey, J.W., 1960. A survey of sampling from contaminated distributions. Contribution to Probability and Statistics. Olkin, Ed., Stanford University Press, Stanford, Calif.

[23] Weron, R., 1996. On the Chambers-Mallows-Stuck Method for simulating skewed stable random variables. *Statist. and Pro. letters* **28**, 165-171.

[24] Zielinski, R., 1977. Robustness : A quantitative approach. Bulletin de l'académie polonaise des sciences. Series Sciences. *Math, Astr.* Vol **XXV**(12), 1281-1286.

[25] Zielinski, R., 1999. A median-unbiased estimator of the AR(1) coefficient. *J. Time Ser. Anal.*, **20**, 477-481.