Journal Afrika Statistika

ISSN 0852-0305

Journal Afrika Statistika
Vol. 5, N°5, 2010, page 228–237.

On the stability of the unit root test

Lynda Atil and Hocine Fellag

Department of Mathematics, Mouloud Mammeri University of Tizi Ouzou, Tizi-Ouzou, 15000, Algeria

Received 28 June 2010; Accepted 7 October 2010

Abstract. We propose to study, by two different approaches, Bayesian and classical, the test of the hypothesis of stationary first order autoregressive model against the random walk model. Therefore, we are going to present the classical approach and the Bayesian approach of this test says the unit root test. The principal aim is to improve the unit root test, either with proposing a better statistic, or with proposing an adequate prior in order to make it more powerful.

Résumé. Nous nous proposons d'étudier, par deux types d'approches différentes, Bayésienne et classique, le test de l'hypothèse de stationnarité d'un processus autorégressif d'ordre un contre le modèle de marche aléatoire dit test de la racine unité. Deux approches sont présentes, l'approche classique et l'approche bayésienne. L'objectif essentiel est d'améliorer les performances du test en proposant une statistique adéquate ou en proposant une loi a priori appropriée pouvant le rendre plus puissant.

Key words: Autoregressive model; Dickey-Fuller test; Posterior odds; Power; Unit root
AMS 2000 Mathematics Subject Classification : Primary 62F11; Secondary 62M10.

1. Introduction

Unit root tests is one of the standard tools for time series econometricians. Traditional test procedures are typically based on least-squares estimation which makes them sensitive to outliers and non normalities in finite samples. To solve this drawback, many number of unit root tests have been proposed by various authors as Herce [8], Shin and So [18], which are, in some way, better protected against these deviations from normality. The aim is to classify the effect these outlying observations in the data have on both classical and Bayesian unit root inference and then to find a robust unit root test. In this paper, two different approaches are considered, the classical approach (Section 2) and the Bayesian approach (Section 3). An exhaustive simulation study is presented in the Section 4.

2. Classical approach

Consider a time series $\{x_t\}$ which follows the model

$$(1-\rho B)x_t = \epsilon_t, \qquad t = \ldots, -1, 0, 1, \ldots, n$$

where $\{\epsilon_t\}_{t=1,\ldots,n}$ is a sequence of independent normally distributed random variables with mean zero and variance 1 and B denotes the backshift operator such that $Bx_t = x_{t-1}$. We assume that $x_0 = 0$. Suppose that all what we observe is the segment of observations $x_1, x_2, \ldots, x_n$. To test the non stationarity (unit root test) means testing one of the following tests :

$$\text{bilateral case:} \qquad H_0 : \rho = 1 \qquad \text{vs} \qquad H_1 : \rho \neq 1$$

Lynda Atil: atillynda@mail.ummto.dz
Hocine Fellag: hfellag@mail.ummto.dz

or

$$\text{unilateral case:} \qquad H_0 : \rho = 1 \qquad \text{vs} \qquad H_1 : \rho < 1$$

at the significant level α.

Various authors treated the problem of unit root test. Dickey and Fuller [5] wrote a pioneer paper where they proposed their well known Dickey-Fuller statistic defined as follows :

The general principle of the test of Dickey and Fuller [5] consists in testing the null hypothesis of the presence of a unit root:

$$H_0 : \rho = 1 \qquad \text{vs} \qquad H_1 : |\rho| < 1.$$

The most important aspect of the Dickey & Fuller is the use of the Student statistic associated to the test $\rho = 1$ and the survey of its asymptotic behavior:

$$t_{\hat{\rho}=1} = \frac{(\hat{\rho} - 1)}{\sigma_{\hat{\rho}}} = (\hat{\rho} - 1) \left[\frac{\sum_{t=1}^{T} x_{t-1}^2}{s_T^2} \right]^{\frac{1}{2}}$$

where $\hat{\rho}$ is the well known least squares estimator of ρ and s_T^2 is the estimated variance of the residuals:

$$s_T^2 = \frac{1}{T-1} \sum_{t=1}^{T} \hat{\varepsilon}_t^2 = \frac{1}{T-1} \sum_{t=1}^{T} (x_t - \hat{\rho} x_{t-1})^2.$$

Under the null hypothesis H_0 of non stationarity in the model (x_t), that is $\rho = 1$, the asymptotic distribution of the t-statistic associated to the test of Dickey & Fuller is the following:

$$t_{\hat{\rho}=1} \longrightarrow \frac{1}{2} \frac{\left[W(1)^2 - 1\right]}{\int_0^1 [W(r)]^2 dr}$$

where $W(.)$ is a standard Brownian motion.

Following, Nelson and Plosser [12] the most popular classical unit root test has been the Dickey-Fuller test. The Dickey-Fuller statistic is traditionally obtained by estimating an autoregressive (AR) model by ordinary least squares. However, it is argued that the ordinary least square (OLS) estimator is non robust against additive outlier (AO). A test statistic based on the estimator also might therefore also be non robust (for more details, see Fellag [6]). So, the outlier sensitivity of the standard Dickey-Fuller statistic is caused by the non robustness of the OLS estimator.

Franses and Haldrup [7] studied effects of additive outliers on unit root Dickey-Fuller tests. They showed that there is over rejection of the unit root hypothesis when additive outliers occur. Also, Shin et al. [17] investigated effects of outliers on unit root tests in an AR(1) and more. They proved that the limiting distribution of the statistic of Dickey-Fuller is affected by an additive outlier. Also, they proposed a method to detect outliers and to adjust the observations. Maddala and Rao [10] showed that, when n goes to infinity the impacts of finite additive outliers will go to zero. Vogelsang [23] proposed two robust procedures to detect outliers and adjust the observations.

In case of small samples, very little is known to compute the power of the unit root test. Fellag [6] derived a formula for the size and the power of the unit root test when a single (AO) contaminant occurs and when the statistic $n(\hat{\rho_{LS}} - 1)$ is used.

In Atil et al. [2], three statistical tests were compared; two usual and one new. The authors derived formulas for computing the size and the power of the three tests when an innovation outlier (IO) occurs at a specified time, k say. Using this comparative study, they showed that their statistic performs better under contamination. But, the small sample case was considered only. However, there are robust unit root tests which seem to be especially well-suited for fat-tailed error distributions, they are often encountered in financial time series, but also for outliers of different types (Atil [1]). Cartensen [4] tries to do a comparative small sample simulation study that analyzes size and power behavior of different tests in a unified framework in order to advise practitioners. Unfortunately, he noticed that there is no robust test that generally out-performs all its competitors, the values of using robust unit root tests depends heavily of the type of non normality.

3. Bayesian approach

Despite the apparent advantages of the Bayesian approach over the classical approach in unit root testing, only a relatively small number of studies have appeared using the Bayesian approach. The reasons may be that the Bayesian approach requires a likelihood function and the use of prior information. Phillips [13, 14] identified the need for priors as the biggest obstacle to Bayesian analysis and argued for more objective Bayesian analysis in time series.

The modeling objective of the Bayesian approach is not to reject a hypothesis based on a pre-determined level of significance, but to determine how probable a hypothesis is relative to other competing hypotheses. There are several ways of comparing hypotheses using methods. The most common method is to calculate posterior odds ratio for various competing hypotheses based on prior sample information. So, Schotman and Van Dijk [15] propose a posterior odds analysis of the hypothesis of a unit root in real exchange rates, because nominal and real exchange rates almost behave like random walks. This conclusion emerges from much of the recent empirical literature on exchange rate models.

Formal statistical tests do not reject the null hypothesis of a unit root against the alternative of a stationary autoregressive time series model.

Schotman and Van Dijk [15] develop a posterior odds ratio for choosing between a random walk and a stationary AR(1) model. The purpose of their study is to reexamine the random walk results for real exchange rates. Is the random walk still the most favored model if compared directly to a simple plausible alternative?

In order to concentrate on the differences between the classical unit root tests and the Bayesian procedure they start off with the simplest possible model, a first order autoregressive process with mean zero. Suppose that we have a sample of T consecutive observations on a time series y_t generated by:

$$y_t = \rho y_{t-1} + \mu_t \tag{1}$$

with,

i) y_0 is a known constant.
ii) μ_t are identically and independently (i.i.d) normally distributed with mean zero and unknown variance σ^2.
iii) $\rho \in S \cup \{1\}$, $S = \{\rho / -1 < a \leq \rho < 1\}$.

The econometric analysis aims at discriminating between a stationary model (here defined as $a \leq \rho < 1$) and the nonstationary model with $\rho = 1$. The lower bound a in assumption (iii) largely determines the specification of the prior for ρ. We said that the principal Bayesian tool to compare a sharp null hypothesis with a composite alternative hypothesis is the posterior odds ratio which is defined as

$$K_1 = K_0 \frac{\int_0^\infty p(\sigma) L(y \mid \rho = 1, \sigma, y_0) d\sigma}{\int_S \int_0^\infty p(\sigma) p(\rho) L(y \mid \rho, \sigma, y_0) d\sigma d\rho} = \frac{p(\rho = 1 \mid y)}{p(\rho \in S \mid y)}. \tag{2}$$

K_0 and K_1 are the prior odds and the posterior odds in favor of the hypothesis $\rho = 1$, respectively. $p(\rho)$ represents the prior density of $\rho \in S$, $p(\sigma)$ the prior density of σ.

$L(y \mid .)$ is likelihood function of the observed data $y = (y_1 \dots y_T)'$ and $Y = (y_0, y')'$ is all observed data.
The posterior odds K_1 are equal to the prior odds K_0 times the Bayes factor. The Bayes factor is the ratio of the marginal posterior density of ρ under the null hypothesis $\rho = 1$ to a weighted average of the marginal posterior under the alternative using the prior density of ρ as a weight function. The prior odds express the special weight given to the null hypothesis, the point $\rho = 1$ is given the discrete prior probability $\vartheta = K_0/(1 + K_0)$. From the posterior odds one can compute the posterior probability of the null hypothesis as $K_1/(1 + K_1)$.

For the complete specification of the marginal prior of ρ and σ we assume that:

$$Pr(\rho = 1) = v, \tag{3}$$

$$p(\rho \mid \rho \in S) = \frac{1}{1 - a}, \tag{4}$$

$$p(\sigma) \propto \frac{1}{\sigma}. \tag{5}$$

The prior of ρ is uniform on S but has a discrete probability ϑ that $\rho = 1$. The likelihood function for the vector of T observations y is:

$$L(y \mid \rho, \sigma, y_0) = (2\pi\sigma^2)^{-T/2} \exp\left\{-\frac{1}{2\sigma^2}\mu'\mu\right\} \tag{6}$$

where $\mu = y - y_{-1}\rho$, and $y_{-1} = (y_0, \ldots, y_{T-1})'$.

Having computed the relevant integrals in (2) the posterior odds ratio becomes:

$$K_1 = \frac{C_T^{-1}}{(T-1)^{1/2}} \frac{v}{1-v} \left(\frac{\sigma_0^2}{\hat{\sigma}^2}\right)^{-T/2} \left(\frac{1-a}{s_{\hat{\rho}}}\right) \left[F\left(\frac{1-\hat{\rho}}{s_{\hat{\rho}}}\right) - F\left(\frac{a-\hat{\rho}}{s_{\hat{\rho}}}\right)\right]^{-1} \tag{7}$$

where $\sigma_0^2 = \frac{1}{T-1}(y - y_{-1})'(y - y_{-1})$ $\hat{\sigma}^2 = \frac{1}{T-1}\left(y'y - \frac{(y'_{-1}y)^2}{(y'_{-1}y_{-1})}\right)$

$$s_{\hat{\rho}}^2 = \hat{\sigma}^2(y'_{-1}y_{-1})^{-1} \qquad \hat{\rho} = \frac{y'_{-1}y}{y'_{-1}y_{-1}} \qquad C_T = \frac{\Gamma((T-1)/2)\Gamma(1/2)}{\Gamma(T/2)}.$$

The empirical lower bound a^* is given by:

$$a^* = \hat{\rho} + s_{\hat{\rho}} F^{-1}(\alpha F(-\hat{\tau})) \tag{8}$$

$F(.)$ is the cumulative t-distribution with $(T-1)$ df. $\hat{\tau} = \frac{\hat{\rho}-1}{s_{\hat{\rho}}}$ is the Dickey-Fuller test statistic. The unit root model is preferred if $K_1 > 1$ or $P(\rho = 1|y, y_0) \geq 0.50$, treating thus the null and the alternative in a symmetric way.

After fixing numerical values for ϑ and α the posterior odds are just a function of the data like any other test statistic. Due to a specific way that the lower bound bas been constructed, the posterior odds are directly related to the Dickey-Fuller test. Setting the prior odds equal to one and for large T, Schotman and Van Dijk approximate $F(.)$ to the cumulative normal distribution. The posterior odds become a function of the Dickey-Fuller statistic $\hat{\tau}$

$$\ln K_1 = -\frac{1}{2}\ln(2\pi) - \frac{1}{2}\hat{\tau}^2 + \ln\left(\frac{-\hat{\tau} - F^{-1}(\alpha F(-\hat{\tau}))}{F(-\hat{\tau})}\right). \tag{9}$$

Since the posterior odds are a function of the Dickey-Fuller test statistic its sampling properties correspond exactly to those of the Dickey-Fuller test.

It seems natural to explore the Bayesian approach to the comparison of stationary models with those involving a unit autoregressive root, and there has been interest in this possibility in the econometric literature, dating from Sims [19] and Sims and Uhlig [20]. Marriott and Newbold [11] consider inference from the perspective of an analyst with a single time series, requiring to determine posterior odds for a unit root model compared with a stationary competitor.

However, a critical issue is the specification of a prior for the autoregressive parameter under stationarity. The literature has devoted great attention to the nature of suitable noninformative priors for the autoregressive coefficients. As a few examples consider the work of Sims [19] and Sims and Uhlig [20], who advocate the use of flat priors. Phillips [13] finds that flat priors bias the inference toward stationary models and suggests instead using Jeffrey priors derived from conditional likelihood functions. Also, Uhlig [21] determines the Jeffreys priors for an AR(1) process from the exact likelihoods and justifies the use of flat priors only in some specific cases. Uhlig [22] summarizes most of the Bayesian contribution to the unit root problem and discusses the sensitivity of the tails of the predictive densities on the prior treatment of explosive roots. Schotman and Van Dijk [15] stress the sensitivity of the posterior odds to the size of the stationary region and suggest restricting this size. Berger and Yang [3] consider a reference prior approach for the AR(1) model. It is particularly interesting to notice the paper by Marriott and Newbold [11], who criticize the use of priors such as the uniform or the Jeffreys prior for the autoregressive coefficients in this setting and advocate the use of sharp, informative prior distributions. However, for the simple problem of testing for a unit root in a first order autoregressive process, they find that the prior distribution for the autoregressive coefficient has a substantial impact on the posterior odds, so that, a very sharp beta prior performs extremely well when the generating process is stationary autoregressive, but the uniform prior is preferable when the true model is non stationary. To better clarify this point, consider the two models

$$X_t = X_{t-1} + \varepsilon_t \quad \text{versus} \quad (X_t - \mu) = \phi(X_{t-1} - \mu) + \varepsilon_t$$

where $\{\varepsilon_t\}$ is a Gaussian white noise process, and $|\phi| < 1$.

Much of the published discussion of this problem has concerned the choice of priors for ϕ and μ.The parameter μ represents the mean of the process under stationarity, but is undefined under $\phi = 1$, leading to difficulties noticed by Schotman and Van Dijk [16]. In many practical applications, however, the analyst may be reluctant to specify an informative prior on μ. Indeed, it is difficult to contemplate a situation where the analyst simultaneously feels able to specify a sharp prior for μ while entertaining non zero prior probability for the random walk model where that parameter is undefined. The inclination here is the use of an improper prior for μ. This causes no difficulty in deriving a posterior for ϕ under the stationary autoregressive model. However, when improper priors are used for parameters occurring in one model and not the other, posterior odds ratio are undefined. Marriott and Newbold [11] remove this as a problem by formulating both of the above models in term of the first differences $W_t = X_t - X_{t-1}$, so that the random walk model and first order autoregressive model we consider here are

$$\begin{aligned} M_1 &: W_t = \varepsilon_t \\ M_2 &: W_t - \phi W_{t-1} = \varepsilon_t - \varepsilon_{t-1}. \end{aligned}$$

Given a sample $W = (W_1, \cdots, W_n)$, the Bayesian comparison of the two models proceeds by computing the posterior model probabilities, which are given by Bayes's theorem as

$$P(M_i|W) = \frac{P(M_i)P(W|M_i)}{\sum_{j=1}^{2} P(M_j)P(W|M_j)}. \tag{10}$$

In (10), $P(M_i)$ is the prior probability assigned to model M_i, $P(W|M_i) = \int_{-1}^{1}\int_{0}^{\infty} p(\phi,\sigma|M_i)p(W|\phi,\sigma,M_i)d\sigma d\phi$ is the integrated joint density of (ϕ,σ,W), $p(\phi,\sigma|M_i)$ is the joint prior density for the parameters, and $p(W|\phi,\sigma,M_i)$ is the likelihood. For this approach, Marriott and Newbold [11] take uniform prior for the models so that $P(M_1) = P(M_2) = 0.5$, they consider adopting the simple decision rule that we would accept model M_i if $P(M_i|W) > 0.5$.

The likelihood for M_1 can be written in terms of the differences W as

$$p(W|\sigma, M_1) = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{t=1}^{n} W_t^2\right\}.$$

For model M_2 the likelihood can be shown to be

$$p(W|\phi,\sigma,M_2) = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} A^{\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{t=1}^{n+2} u_t^2\right\}$$

where

$$\begin{aligned} A &= [1 + n(1-\phi)(1+\phi)^{-1}]^{-1}, \\ u_1 &= -AC, \\ u_2 &= \phi(1-\phi^2)^{-\frac{1}{2}}AC, \\ u_3 &= W_1 - (1+\phi)^{-1}AC, \end{aligned}$$

$$u_t = W_{t-2} + (1-\phi)\sum_{j=1}^{t-3} W_j - (1+\phi)^{-1}AC, \quad t = 4, \cdots, n+2$$

$$C = (1-\phi)\sum_{t=1}^{n} W_t + (1-\phi)^2 \sum_{t=1}^{n}(n-t)W_t.$$

Adopting the usual noninformative prior for σ, we write $p(\phi,\sigma) = \sigma^{-1}p(\phi)$. The joint density of (ϕ,σ,W) for M_2 can be written as

$$p(\phi,\sigma,W|M_2) = \frac{1}{\sigma^{n+1}(2\pi)^{\frac{n}{2}}} A^{\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{t=1}^{n+2} u_t^2\right\} p(\phi).$$

Integrating this with respect to σ gives

$$p(\phi,W|M_2) = (2\pi)^{-\frac{n}{2}}\Gamma\left(\frac{n}{2}\right) 2^{(n-2)/2} A^{\frac{1}{2}} \left[\sum_{t=1}^{n+2} u_t^2\right]^{-\frac{n}{2}} p(\phi)$$

and then

$$P(W|M_2) = \int_{-1}^{1} p(\phi,W|M_2)d\phi.$$

It is straightforward to show that the integrated joint density in the case of M_1 is given by

$$P(W|M_1) = (2\pi)^{-\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)2^{(n-2)/2}\left[\sum_{t=1}^{n+2}W_t^2\right]^{-\frac{n}{2}}.$$

Equation (10) can now be used to obtain the posterior model probabilities as

$$P(M_1|W) = \frac{1}{1+K} \qquad \text{and} \qquad P(M_2|W) = \frac{K}{1+K}$$

where

$$K = \frac{\int_{-1}^{1} A^{\frac{1}{2}}\left[\sum_{t=1}^{n+2}u_t^2\right]^{-\frac{n}{2}}p(\phi)d\phi}{[\sum_{t=1}^{n}W_t^2]^{-\frac{n}{2}}}.$$

It should be noticed that the analysis of first differences involves no information loss about the parameter of the autoregressive model compared with an analysis of levels with an improper prior on μ.
An important consideration is how much prior information is actually available for ϕ. Can a uniform really be considered a sensible choice if the investigator seriously believes that a random walk model could provide a reasonable explanation of the behavior of a time series? Marriott and Newbold use for purposes of comparison the uniform prior for ϕ, together with two sharper, beta priors with densities,

$$p(\phi) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)2^{\alpha+\beta-1}}(1+\phi)^{\alpha-1}(1-\phi)^{\beta-1}, \quad |\phi| < 1.$$

Marriott and Newbold [11] generated 1000 samples of 100 observations from first order autoregressive with $\phi = 1, 0.95, 0.90, 0.85$ and 0.80, they computed the proportion of the samples for which $P(M_1|W) > 0.5$, that is, the proportion of the time that the random walk model would have been chosen.

	prior for ϕ		
ϕ	uniform	beta1	beta2
1	0.968	0.875	0.677
0.95	0.827	0.554	0.244
0.90	0.554	0.180	0.017
0.85	0.212	0.018	0.000
0.80	0.041	0.000	0.000

It appeared that,when the true process is stationary autoregressive ($\phi \neq 1$), the prior beta2 performs well and would be the best of the three in such situation. However, the opposite is the case for a random walk process ($\phi = 1$). The uniform prior seems the best.

4. A Monte Carlo study

An empirical comparison is made in the setting of the beta generalised prior of Libby and Novick [9] defined by the following density

$$f(\phi, a, b, \lambda) = \frac{\lambda^a\phi^{a-1}(1-\phi)^{b-1}}{B(a,b)\{1-(1-\lambda)\phi\}^{a+b}}$$

for $0 \leq \phi \leq 1$, where $a > 0$, $b > 0$ and $\lambda > 0$. When $\lambda = 1$, this reduces to the standard beta.
The following density represents the beta generalised distribution for $-1 \leq \phi \leq 1$, where $a > 0$, $b > 0$ and $\lambda > 0$:

$$p(\phi) = \frac{2\lambda^a(\phi+1)^{a-1}(1-\phi)^{b-1}}{B(a,b)\{(1-\phi)+\lambda(\phi+1)\}^{a+b}}.$$

We replace the prior of ϕ in the formula of the posterior odds ratio

$$K = \frac{\int_{-1}^{1} A^{\frac{1}{2}}\left[\sum_{t=1}^{n+2}u_t^2\right]^{-\frac{n}{2}}p(\phi)d\phi}{[\sum_{t=1}^{n}W_t^2]^{-\frac{n}{2}}}.$$

and we effected a comparison for different values of a, b and λ.

This simulation study is established in order to be able to improve the results obtained by Marriott and Newbold [11](1998) who use a beta prior that they judge more adequate than the uniform prior.

Marriott and Newbold [11] use three priors, the uniform law $U[0,1]$, the law beta1=Beta(5, 0.5) and the law beta2=Beta(50, 0.5). They notice that when the data are generated by a stationary autoregressive process of order one, the priori law which performs the best is the beta2, however this state changes under the hypothesis of the unit root and the uniform distribution is the more adequate one in this case. In the following, we present results obtained in the setting of our empiric exhaustive survey. All the tables and figures are given in the annex. In Table 1, we compute the variation, for different values of λ, of the proportion of the samples for which $P(M_1|W) > 0.5$ under the hypothesis $\phi = 0.95$, that is in the case of a stationary autoregressive model of order one with $a = 50$, $b = 0.5$, the values of λ vary from 1 to 5. We remark that this proportion decreases as λ increases until a certain value "that is $\lambda = 1.3$" for taking again its ascension.
When $\lambda = 1$, we recover approximatively the results of Marriott and Newbold [11], one notices that one can find for a given λ, a Beta generalized distribution which performs better than the Beta standard distribution. These simulations are made in the setting of 1000 samples of 100 observations, that is, under the same conditions of those of Marriott and Newbold [11] in order to be able to establish an empirical comparison. The figure 1 illustrates the results of the Table 1.

In the Table 2, one uses the value of λ selected from the table 1 either $\lambda = 1.3$. One calculate then with varying the value of ϕ, the proportion of times where $P(M_1|W) > 0.5$. One notices that for this distribution of Beta generalized, the results prove to be satisfactory under the H_0 hypothesis and the H_1 hypothesis.
For example, when we generate an AR(1) process under the stationary condition $\phi = 0.99$, the proportion of times where $P(M_1|W) > 0.5$ is 0.4670, for $\phi = 0.95$ this proportion is equal to 0.1450, which means that we reject the unit root hypothesis H_0. But, when $\phi = 1$ the proportion is equal to 0.6040, which is not a bad result. This is illustrated in the figure 2.

The results obtained in the Table 3 are analogues to those obtained in the Table 1, that is, they go in the same sense but for different values of a and b, $a = 5$, $b = 0.5$. The values of λ vary from 0.07 to 5.00, and we select the one which gives the smallest proportion. The figure 3 illustrates these results. The value of λ which is selected is $\lambda = 0.1$ from the Table 3. Then we proceed from analogous manner to the Table 2, when we generate an AR(1) process under the stationary condition $\phi = 0.99$, the proportion of times where $P(M_1|W) > 0.5$ is of 0.4760, for $\phi = 0.95$ this proportion is equal to 0.1650, which means that we reject the unit root hypothesis H_0. The figure 4 illustrates the results of the Table 4.

The results obtained in the table 5 are analogues to those obtained in the table 1 and table 3, that is, they go in the same sense but for different values of a and b, $a = 1$, $b = 1$. The values of λ vary from 0.004 to 5.00, and we select the one which gives the smallest proportion. The figure 5 illustrates the results of the Table 5. The value of λ which is selected is $\lambda = 0.008$ from the table 5. Then we proceed from analogous manner to the Tables 2 and 4, when we generate an AR(1) process under the stationary condition $\phi = 0.99$, the proportion of times where $P(M_1|W) > 0.5$ is of 0.5220, for $\phi = 0.95$ this proportion is equal to 0.1570. The figure 6 illustrates the results of Table 6. These results are obtained in the setting of a simulation of 100 observations for 1000 repetitions.

5. Conclusion

Marriott and Newbold [11] explored the use of the distribution beta as the specification of an prior. They demonstrated how the Bayesian calculations could be shaped, noting to the passage the importance for the analyst to pay a particular attention on the question of the choice of an adequate distribution a priori.

However, we noticed that the results of Marriott and Newbold [11] can be improved, for this, we proposed the law beta generalized of Libby and Novick, that are provided with three parameters λ, a and b. We notice that for a certain value of λ, we can find an a and b, that allow us to take the good decision.

Therefore, the construction of a stable test of the unit root constitutes an important difficulty in the small samples case. Indeed, when we don't benefit from the nice theorems of convergence who facilitate the mathematical and statistical treatments, we can have recourse to numeric techniques or to monte carlo methods. Generally, the authors agree with the idea that the Bayesian approach offers a more satisfactory alternative than the classic approach in the empiric modelisation. In the unit root tests Sims [19], Sims and Uhlig [20], and other authors recommended the Bayesian approach.

This simulation study is limited to the simplest possible unit root model since it's aim is the behavior of different unit root test under departures from normality. However, there is evidence, from previous works, that more complex models can lead to different conclusions.

Then, can we say what is the best approach in the unit root tests, the Bayesian approach or the classic approach? Intuitively, it would be hazardous to discriminate between two big theories on the basis of a univariate model, the problem remains therefore unsolved.

Acknowledgment
Our sincere thanks to the Editors and the anonymous reviewers for their comments and suggestions.

Appendix A: Tables

Proportion of times that $P(M_1|W) > 0.5$.

λ	proportion	λ	proportion
5.0	0.2670	1.7	0.1460
4.0	0.2270	1.6	0.1460
3.5	0.2120	1.5	0.1450
3.0	0.1980	1.4	0.1460
2.5	0.1770	1.3	0.1450
2.0	0.1580	1.2	0.1480
1.9	0.1540	1.1	0.1530
1.8	0.1500	1.0	0.1620

Table 1. $\phi = 0.95$, $a = 50$, $b = 0.5$.

ϕ	proportion
1.00	0.6040
0.99	0.4670
0.98	0.3530
0.97	0.2710
0.96	0.2030
0.95	0.1450
0.90	0.0200
0.85	0.0030
0.80	0

Table 2. $\lambda = 1.3$, $a = 50$, $b = 0.5$.

Proportion of times that $P(M_1|W) > 0.5$.

λ	proportion	λ	proportion
5.0	0.6920	0.6	0.3310
4.0	0.6660	0.5	0.2920
3.5	0.6510	0.4	0.2490
3.0	0.6240	0.3	0.2140
2.0	0.5600	0.2	0.1740
1.0	0.4430	0.1	0.1650
0.9	0.4220	0.09	0.1760
0.8	0.4050	0.08	0.1830
0.7	0.3720	0.07	0.2000

Table 3. $\phi = 0.95$, $a = 5$, $b = 0.5$.

ϕ	proportion
1.00	0.6020
0.99	0.4760
0.98	0.3610
0.97	0.2820
0.96	0.2220
0.95	0.1650
0.90	0.0220
0.85	0.0060
0.80	0

Table 4. $\lambda = 0.1$, $a = 5$, $b = 0.5$.

Proportion of times that $P(M_1|W) > 0.5$.

λ	proportion	λ	proportion
5.0	0.9670	0.06	0.4320
4.0	0.9630	0.05	0.3920
3.0	0.9570	0.04	0.3460
2.0	0.9350	0.03	0.2960
1.0	0.8810	0.02	0.2430
0.50	0.8010	0.01	0.1640
0.40	0.7700	0.009	0.1600
0.30	0.7140	0.008	0.1570
0.20	0.6500	0.007	0.1590
0.10	0.5260	0.006	0.1670
0.09	0.5000	0.005	0.1760
0.08	0.4780	0.004	0.1930
0.07	0.4600		

Table 5. $\phi = 0.95$, $a = 1$, $b = 1$.

ϕ	proportion
1.00	0.6730
0.99	0.5220
0.98	0.4000
0.97	0.3110
0.96	0.2350
0.95	0.1570
0.90	0.0200
0.85	0.0020
0.80	0

Table 6. $\lambda = 0.008$, $a = 1$, $b = 1$.

Appendix B: Figures

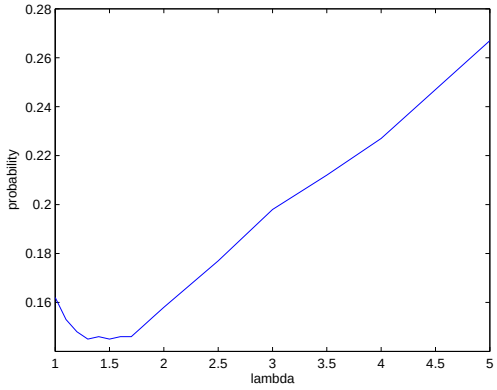

Fig. B1. $\phi = 0.95$, $a = 50$, $b = 0.5$

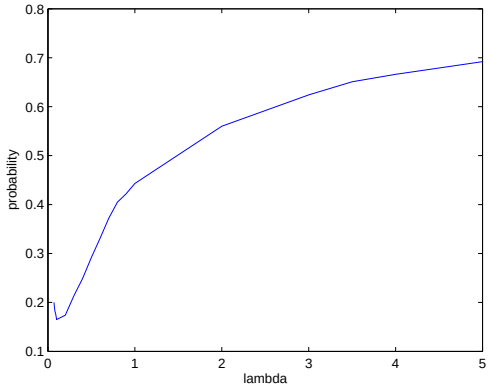

Fig. B2. $\lambda = 0.008$, $a = 1$

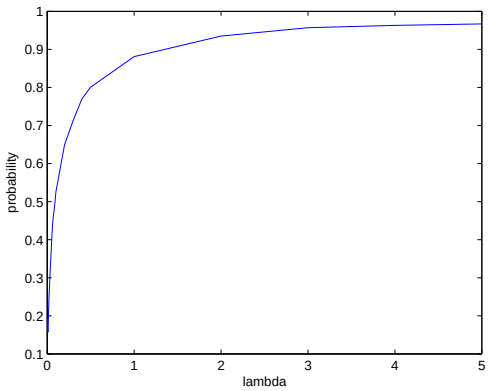

Fig. B3. $\phi = 0.95$, $a = 5$, $b = 0.5$

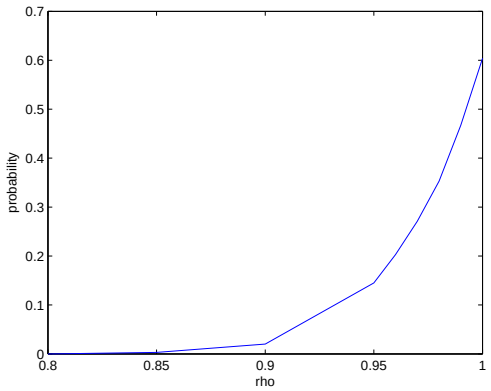

Fig. B4. $\lambda = 0.1$, $a = 5$

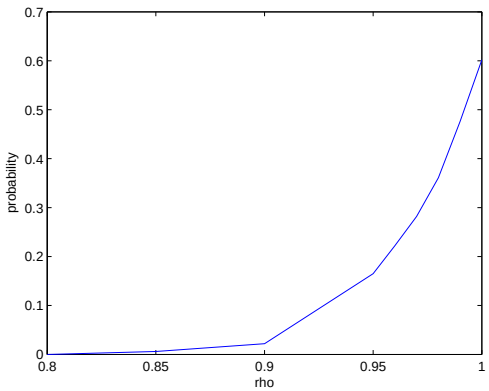

Fig. B5. $\phi = 0.95$, $a = 1$, $b = 1$

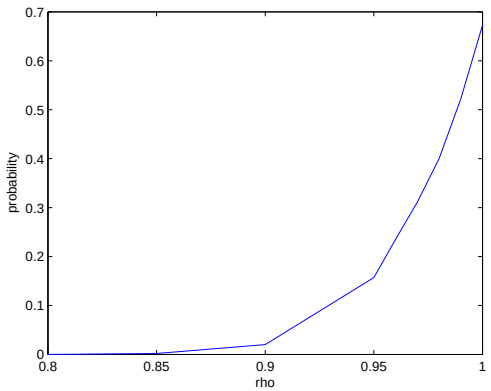

Fig. B6. $\lambda = 0.008$, $a = 1$

References

[1] Atil, L., 2010. Stabilité des modèles ARCH et du test de la racine unité.Thèse de Doctorat. Université de Tizi-Ouzou.

[2] Atil, L., Fellag, H. and Nouali, K., 2006. Unit root test under innovation outlier contamination. Small sample case. *Discuss. Math. Probab. Stat.* , **26**, 5-17.

[3] Berger, J.O. and Yang, R.Y., 1994. Noninformative priors and Bayesian testing for the AR(1) model. *Econometric Theory* **10**, 461-482.

[4] Cartensen, K., 2003. The finite-sample performance of robust unit root tests. *Statistical Papers* **44**, 469-482.

[5] Dickey, D.A. and Fuller, W.A., 1979. Distribution of estimators for autoregressive time series with unit root. *J. Amer. Statist. Assoc.*, **74**, 427-431.

[6] Fellag, H., 2001. Unit root test in the presence of a single additive outlier. Small sample case. *Discuss. Math. Probab. Stat.* **21**, 89-97.
[7] Franses, P.H. and Haldrup, N., 1994. The effects of additive outliers on tests for unit root and cointegration. *J. Bus. Econom. Statist.***12**, 471-478.
[8] Herce, M.A., 1996. Asymptotic theory of LAD estimation in a unit root with finite variance errors. *Econometric Theory.* **12**, 129-153.
[9] Libby, D.L., and Novick, M.R., 1982. Multivariate generalized beta-distributions with applications to utility assessment. *Journal of Educational Statistics.* **7**, 271-294.
[10] Maddala, G.S. and Rao, C.R., 1997. Handbook in statistics, *Elsevier Sciences.* New York. Volume **15**, 237-266, .
[11] Marriott, J. and Newbold, P., 1998. Bayesian comparison of ARIMA and stationary ARMA models. *International Statistical Review.* **66**(3), 323-336.
[12] Nelson, C.R. and Plosser, C.I., 1982. Trends and random walk in macroeconomic time series: Some evidence and implication. *Journal of Monetary Economics.* **10**, 139-162.
[13] Phillips, P.C.B., 1991a. To criticize the critics: An objective Bayesian analysis of stochastics trends. *J. Appl. Econometrics* **6**, 333-364.
[14] Phillips, P.C.B., 1991b. Bayesian routes and unit roots:De rebus priorbus semper est disputandum. *J. Appl. Econometrics* **6**, 435-474.
[15] Schotman, P. and Van Dijk, H.K., 1991a. A Bayesian analysis of the unit root in real exchange rates. *J. Appl. Econometrics* **49**, 195-238.
[16] Schotman, P. and Van Dijk, H.K., 1991b. On Bayesian routes to unit roots. *J. Appl. Econometrics.* **6**, 461-464.
[17] Shin, D.W., Sarkar, S. and Lee, J.H., 1996. Unit root tests for time series with outliers. *Statist. Proba. Letters.* **30**, 189-197.
[18] Shin, D.W., and So, B.S., 1999. Unit root tests based on adaptive maximum likelihood estimation. *Econometric Theory.* **15**, 1-23.
[19] Sims, C.A., 1988. Bayesian skepticism on unit root econometrics. *J. Econom. Dynam. Control* **12**, 463-474.
[20] Sims, C.A. and Uhlig, H., 1991. Understanding unit rooters: A helicopter tour. *Econometrica.* **59**, 1591-1599.
[21] Uhlig, H, 1994a. On Jeffreys prior when using the exact likelihood function. *Econometric Theory.* **10**, 633-644.
[22] Uhlig, H, 1994b. What macroeconomists should know about unit roots: A Bayesian perspective. *Econometric Theory.* **10**, 645-671.
[23] Vogelsang, T.J., 1999. Two simple procedures for testing for a unit root when there are additive outliers. *J. Time Ser. Anal.* **20**(2), 237-252.