UNIT ROOT TEST UNDER CONTAMINATION SMALL SAMPLE CASE

Hocine FELLAG and Lynda ATIL

Deparment of Mathematics
University of Tizi-Ouzou
Algeria

Luzern, November 14-16, 2007

outline

1. INTRODUCTION

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

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

where $\epsilon_1, \epsilon_2, \ldots, \epsilon_n$ *iid* and $N(0, \sigma^2)$
B: backshift operator such that $Bx_t = x_{t-1}$ and $x_0 = 0$
Suppose that all what we observe is the segment of observations

$$x_1, x_2, \ldots, x_n \tag{2}$$

$$H_0 : \rho = 1 \quad \text{vs} \quad H_1 : \rho < 1 \qquad \text{at the significance level } \alpha \tag{3}$$

Two statistics are studied.
First : Dickey-Fuller statistic

$$T_{DF} = n(\hat{\rho}_{LS} - 1) \tag{4}$$

where

$$\hat{\rho}_{LS} = \left[\sum_{t=2}^{n} x_t x_{t-1}\right] \left[\sum_{t=2}^{n} x_{t-1}^2\right]^{-1} \tag{5}$$

$\hat{\rho}_{LS}$:well known least squares estimator of ρ.

Second : Symmetrical Statistic

$$T_{SYM} = -(n-2)^{1/2}(1+\hat{\rho}_S)^{-1/2}(1-\hat{\rho}_S)^{1/2} \tag{6}$$

where $\hat{\rho}_S$ is the simple symmetrical estimator of ρ defined by

$$\hat{\rho}_S = \frac{\sum_{t=2}^{n} x_{t-1}x_t}{\frac{1}{2}(x_1^2 + x_n^2) + \sum_{t=2}^{n-1} x_t^2} \tag{7}$$

T_{SYM} is the corresponding t-statistic given by Fuller (1996).

outline

2. TDF under GAUSSIAN MODEL

Computation of $\alpha = P_{H_0}(n(\hat{\rho}_{LS} - 1) < c)$ for a given value of the constant c

Notations : $X = (x_1, x_2, \ldots, x_n)^T$ - vector of the observations

$E = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)^T$: vector of innovations

A : the nxn-matrix generated by the coefficient ρ defined by

$$a_{ij} = \begin{cases} \rho^{i-j} & \text{if} \quad i \geq j \quad i,j = 1, \ldots, n \\ 0 & \text{elsewhere} \end{cases}$$

Then, we have the relation $X = AE$. $\hat{\rho}_{LS}$ can be written as a ratio of two quadratic forms :$\hat{\rho}_{LS} = \left[X^T R_1 X\right]\left[X^T R_2 X\right]^{-1}$ with

$$R_1 = \begin{pmatrix} 0 & 1/2 & 0 & \dots & 0 \\ 1/2 & 0 & \ddots & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \ddots & 0 & 1/2 \\ 0 & 0 & \dots & 1/2 & 0 \end{pmatrix}$$

and

$$R_2 = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & 0 & 0 \end{pmatrix}$$

Since $X = A.E$, we obtain $\hat{\rho}_{LS} = \left[E^T LE\right]\left[E^T ME\right]^{-1}$ with $L = A^T R_1 A$ and $M = A^T R_2 A$. If we use the notation $c' = 1 + c/n$, we have,

$$P_{H_0}(n(\hat{\rho}_{LS} - 1) < c) = P_{H_0}(E^T.B.E < 0)$$

with $B = A^T(R_1 - c' R_2).A$.

To compute $P_{H_0}(n(\hat{\rho}_{LS} - 1) < c)$, we use the formula given by Imhof (1961) which gives

$$P_{H_0}(n(\hat{\rho}_{LS} - 1) < c) = 0.5 - \frac{1}{\pi} \int_0^\infty \frac{\sin f(u)}{ug(u)} du \qquad (6)$$

where

$$f(u) = \frac{1}{2} \sum_{i=1}^{n} Tan^{-1}(\lambda_i u) \quad \text{and} \quad g(u) = \prod_{1}^{n} \left(1 + \lambda_i^2 u^2\right)^{1/4}$$

$\lambda_1, \lambda_2, \ldots, \lambda_n$ being the eigenvalues of the matrix B. The value of the integral can be obtained using a numerical method.

outline

3.TDF UNDER ADDITIVE OUTLIER MODEL

Additive outlier of magnitude Δ occurs at a time k $(1 < k < n)$. At a position $k \in]1, n[$, a single additive outlier of magnitude Δ occurs. Hence, instead of the segment $x_1, x_2, \ldots, x_n$, we observe $z_1, z_2, \ldots, z_n$ where

$$z_t = x_t \quad \forall t \neq k \quad \text{and} \quad z_k = x_k + \Delta$$

Problem :Computation of $P_{H_0}(n(\hat{\rho}^*_{LS} - 1) < c)$ in contaminated model where

$$\begin{aligned}
\hat{\rho}^*_{LS} &= \left[\sum_{t=2}^{n} z_t z_{t-1}\right]\left[\sum_{t=2}^{n} z_{t-1}^2\right]^{-1} \\
&= \left[\sum_{t=2}^{n} x_t x_{t-1} + \Delta(x_{k-1} + x_{k+1})\right]\left[\sum_{t=2}^{n} x_{t-1}^2 + 2\Delta x_k + \Delta^2\right]^{-1}.
\end{aligned}$$

Notations : $Z = (z_1, z_2, \ldots, z_n)^T$ and $Y = (y_1, y_2, \ldots, y_n)^T$
where

$$y_t = \varepsilon_t \quad \forall t \neq k, k+1 \quad \text{and} \quad y_k = \varepsilon_k + \Delta \quad , \quad y_{k+1} = \varepsilon_{k+1} - \Delta$$

Then, we have $Z = AY$ and

$$P_{H_0}(n(\hat{\rho}^*_{LS} - 1) < c) = P_{H_0}(\hat{\rho}^*_{LS} < c') = P_{H_0}(Y^T A^T (R_1 - c' R_2) AY < 0)$$

where c' is defined above.
Using Imhof's formula, we obtain the value of $P_{H_0}(n(\hat{\rho}^*_{LS} - 1) < c)$ in the presence of an additive outlier contamination

$$P_{H_0}(n(\hat{\rho}^*_{LS} - 1) < c) = 0.5 - \frac{1}{\pi}\int_0^{\infty} \frac{\sin f^*(u,\Delta)}{ug^*(u,\Delta)}du \qquad (8)$$

where

$$f^*(u,\Delta) = f(u) + \frac{\Delta^2 u}{2}\sum_{i=1}^{n} \frac{\lambda_i(Q_{k,i} - Q_{k+1,i})^2}{1+\lambda_i^2 u^2}$$

and

$$g^*(u,\Delta) = g(u)exp\left\{\frac{\Delta^2 u^2}{2}\sum_{i=1}^{n} \frac{\lambda_i^2(Q_{k,i} - Q_{k+1,i})^2}{1+\lambda_i^2 u^2}\right\}$$

$\lambda_1, \lambda_2, \ldots, \lambda_n$: the same eigenvalues of the matrix B .
$Q_{k,i}$ and $Q_{k+1,i}$: the (k,i) and $(k+1,i)$ element of the orthogonal matrix Q containing the normalized eigenvectors of B respectively.

Using numerical integration $\longrightarrow$ numerical approximations of the size in the contaminated model.

REMARKS.

1) The power of the test for a given $\rho = \rho_0$ is easily obtained in the same way with assuming that $y_{k+1} = \varepsilon_{k+1} - \rho_0 \Delta$. Then, we derive the formula given above with replacing $(Q_{k,i} - Q_{k+1,i})^2$ by $(Q_{k,i} - \rho_0 Q_{k+1,i})^2$ in the expression of $f^*(u,\Delta)$ and $g^*(u,\Delta)$.

2) Since the formula (8) depends on Δ^2 only, we can say that, positive and negative values of Δ have the same effect on the size and the power.

APPLICATION FOR AO MODEL. The following Table presents some values of the power of the 5 % level test when
$\rho = 0.5, 0.7, 0.8, 0.9, 0.95, 1.0$
and $\Delta = 0, 1, 2, 3, 4, 5$.

Critical values (5 % level) for $n = 5, 10, 25$, equal to -5.612, -6.5575 and -7.3800 respectively. ($\rho = 1.0$, size of unit root test (Fuller,1996,p.641)).

Variation of the power of the test with Δ

n	ρ	Δ					
		0	1	2	3	4	5
10	0.5	0.3606	0.4093	0.5235	0.6455	0.7447	0.8187
	0.7	0.1799	0.2239	0.3352	0.4709	0.5946	0.6932
	0.8	0.1202	0.1554	0.2510	0.3747	0.4967	0.6011
	0.9	0.0781	0.1051	0.1803	0.2824	0.3894	0.4886
	0.95	0.0625	0.0857	0.1507	0.2403	0.3362	0.4279
	1.0	0.0500	0.0696	0.1250	0.2020	0.2856	0.3674
25	0.5	0.9339	0.9441	0.9659	0.9846	0.9948	0.9986
	0.7	0.5917	0.6299	0.7245	0.8314	0.9141	0.9628
	0.8	0.3392	0.3761	0.4769	0.6126	0.7460	0.8509
	0.9	0.1490	0.1717	0.2387	0.3425	0.4665	0.5908
	0.95	0.0848	0.1002	0.1468	0.2223	0.3183	0.4227
	1.0	0.0500	0.0582	0.0883	0.1371	0.2014	0.2744

outline

3. MONTE CARLO STUDY UNDER SYMMETRICAL CONTAMINATION

AIM

What is the behavior of the test when the innovations are distributed according to $(1-\epsilon).N(0,1)+\epsilon.F$ where F some given distribution.

some simulation results

1. $F \sim N(0, \sigma^2)$ $(TukeyModel)$ $n = 25$ (TDF in brackets)

epsilon	sigma					
	0	1	2	3	5	10
0.00	0.049	0.049	0.049	0.049	0.049	0.049
0.10	0.050 (0.051)	0.049 (0.049)	0.050 (0.051)	0.049 (0.051)	0.049 (0.053)	0.045 (0.053)
0.20	0.051 (0.049)	0.049 (0.049)	0.048 (0.049)	0.054 (0.050)	0.049 (0.050)	0.049 (0.056)
0.30	0.051 (0.052)	0.049 (0.049)	0.050 (0.050)	0.051 (0.055)	0.048 (0.054)	0.047 (0.056)
0.40	0.051 (0.052)	0.049 (0.049)	0.050 (0.051)	0.050 (0.054)	0.051 (0.050)	0.049 (0.049)
0.50	0.045 (0.049)	0.049 (0.049)	0.050 (0.053)	0.048 (0.050)	0.049 (0.051)	0.048 (0.053)

2.F symmetric α-stable for n=25 (TDF in brackets)

ϵ	α				
	1.00	1.25	1.50	1.75	2.00
0.00	0.049 (0.051)	0.049 (0.051)	0.049 (0.051)	0.049 (0.051)	0.049 (0.051)
0.10	0.049 (0.052)	0.049 (0.050)	0.048 (0.053)	0.046 (0.054)	0.050 (0.051)
0.20	0.048 (0.053)	0.049 (0.051)	0.053 (0.053)	0.049 (0.049)	0.051 (0.051)
0.30	0.051 (0.048)	0.051 (0.050)	0.051 (0.050)	0.051 (0.047)	0.051 (0.051)
0.40	(0.051 (0.048)	0.056 (0.052)	0.046 (0.050)	0.051 (0.051)	0.049 (0.056)
0.50	0.052 (0.051)	0.049 (0.052)	0.049 (0.054)	0.049 (0.049)	0.053 (0.052)

3. F $\sim$ U[-a,a],a=π, 1.0,0.5 : n=25 (TDF in brackets)

ϵ	a		
	3.14	1.00	0.50
0.00	0.049 (0.048)	0.049 (0.048)	0.049 (0.048)
0.10	0.049 (0.051)	0.052 (0.047)	0.049 (0.050)
0.20	0.048 (0.050)	0.048 (0.050)	0.052 (0.052)
0.30	0.049 (0.045)	0.050 (0.048)	0.050 (0.054)
0.40	0.048 (0.053)	0.053 (0.049)	0.053 (0.054)
0.50	0.048 (0.051)	0.050 (0.050)	0.047) (0.049)

4. F $\sim$ EXP(1) for n=25,50,100 (TDF in brackets)

ϵ	n=25	n=50	n=100
0.00	0.051 (0.050)	0.050 (0.050)	0.051 (0.051)
0.10	0.071) (0.066)	0.090 (0.082)	0.128 (0.116)
0.20	0.121 (0.099)	0.206 (0.161)	0.373 (0.264)
0.30	0.210 (0.163)	0.401 (0.264)	0.680 (0.423)
0.40	0.337 (0.227)	0.621 (0.378)	0.899 (0.549)
0.50	0.483 (0.299)	0.814 (0.483)	0.984 (0.640)

5. F $\sim$ U[0,1] for n=25,50,100 (TDF in brackets)

epsilon	n=25	n=50	n=100
0.00	0.050 (0.053)	0.050 (0.048)	0.050 (0.051)
0.10	0.058 (0.056)	0.064 (0.059)	0.079 (0.067)
0.20	0.081 (0.073)	0.111 (0.092)	0.167 (0.129)
0.30	0.129 (0.106)	0.191 (0.153)	0.334 (0.239)
0.40	0.202 (0.144)	0.333 (0.245)	0.572 (0.368)
0.50	0.319 (0.214)	0.525 (0.342)	0.800 (0.499)

outline

4. CONCLUSION AND OUTLOOKS

- The unit root test is very sensitive to Additive Outlier contamination
- The size of the unit root seems to be stable under symmetrical contamination around zero
- Theoretical study of the power test is needed