STATISTICA, anno LXIII, n. 3, 2003

IL TEST BOOTSTRAP ESTERNO PER LA RICERCA DI RADICI UNITARIE IN PRESENZA DI OUTLIERS

C. Pizzi, I. Procidano, S. Rigatti Luchini

1. INTRODUZIONE

In questi ultimi anni si sta assistendo ad un crescente interesse da parte degli studiosi riguardo al problema della presenza di radici unitarie in una serie storica. Di conseguenza numerosi sono stati i test proposti per la verifica di radici unitarie (Dickey e Fuller, 1979; Fuller, 1996; Procidano e Rigatti Luchini, 1999, 2000, 2001; Phillips e Peron, 1988; Moreno e Romo, 2000).

Tra essi certamente il più utilizzato è il test Dickey-Fuller, DF, o la sua variante Augmented Dickey Fuller, ADF, proposta per tenere conto della violazione, frequente in molte serie storiche economiche, dell'ipotesi di incorrelazione del termine di errore.

L'applicabilità del test DF poggia altresì su di una seconda ipotesi: quella di omoschedasticità del termine di errore. Solo di recente in letteratura si è prestata attenzione alla violazione di questa assunzione valutandone gli effetti e proponendo un test bootstrap, d'ora in avanti indicato come test B, robusto rispetto a tale violazione (Procidano e Rigatti Luchini, 1999, 2000, 2001).

Alla luce delle buone performance ottenute da questo test in specifiche ipotesi di eteroschedasticità, si è pensato di valutarne la robustezza, tramite simulazioni, in alcune situazioni particolari, tipiche delle serie finanziarie o economiche. In primo luogo si è analizzato il caso in cui si abbia una violazione all'ipotesi di omoschedasticità in serie storiche rilevate ad alta frequenza. Successivamente si è studiata la robustezza del test B in serie storiche contaminate dalla presenza di outliers sia di tipo additivo sia di tipo innovativo.

Il lavoro è organizzato nel modo seguente: nel primo paragrafo si richiamerà brevemente il test B e l'algoritmo che è stato implementato per la sua costruzione. Seguiranno due paragrafi in cui verranno presentati i metodi ed i risultati ottenuti nelle situazioni descritte sopra. Infine il lavoro si concluderà con alcune osservazioni e proposte di lavoro.

Lavoro svolto nell'ambito del progetto di ricerca cofinanziato MURST (1999-2001) "Metodi di inferenza statistica per problemi complessi", coordinatore nazionale Prof. Fortunato Pesarin.

2. IL TEST PER RADICI UNITARIE BASATO SUL BOOTSTRAP ESTERNO

Si supponga di possedere T realizzazioni generate dal modello:

$$x_t = \alpha x_{t-1} + \varepsilon_t \qquad t = 1,...,T \tag{1}$$

dove ε_t è un white-noise con varianza pari a σ^2.

L'ipotesi che si vuole sottoporre a verifica è H_0: $\alpha=1$. La statistica utilizzata per il test DF è la seguente:

$$\tau = (\hat{\alpha} - 1)S^{-1}\left(\sum_{t=1}^{T} x_{t-1}^2\right)^{1/2},$$

dove $\hat{\alpha} = \left(\sum_{t=1}^{T} x_t x_{t-1}\right)\left(\sum_{t=1}^{T} x_{t-1}^2\right)^{-1}$, $S^2 = (T-2)^{-1}\sum_{t=1}^{T}(x_t - \hat{\alpha}x_{t-1})^2$.

Poiché la statistica τ, quando l'ipotesi nulla è vera, segue una distribuzione non standard, i percentili della coda di sinistra della distribuzione sono stati tabulati da Fuller (Fuller, 1976) e, successivamente, da Mackinnon (Mackinnon, 1991).

Il test B fa riferimento al metodo di ricampionamento proposto da Wu (Wu, 1996). Si tratta di un metodo robusto rispetto all'eteroschedaticità che richiede, per la sua implementazione, l'utilizzo di una distribuzione esterna, la cui scelta costituisce un elemento di soggettività. Nel presente lavoro si sono considerate due diverse distribuzioni esterne. La prima, V, è caratterizzata dalla asimmetria:

$$V = \left(\nu_1 + \frac{Z}{\sqrt{2}}\right)\left(\nu_2 + \frac{Z}{\sqrt{2}}\right) - \nu_1\nu_2$$

dove Z~N(0,1), $\nu_1 = \left(\frac{3}{4} + \frac{\sqrt{17}}{12}\right)^{\frac{1}{2}}$, $\nu_2 = \left(\frac{3}{4} - \frac{\sqrt{17}}{12}\right)^{\frac{1}{2}}$.

Tale distribuzione è stata scelta in ragione dei vantaggi derivanti dall'impiego di una variabile casuale con momento terzo unitario (Liu, 1988, e Mammen, 1993). In questo caso il test bootstrap è indicato B_A.

La seconda è una distribuzione uniforme U(-√3, √3). Si presenta, quindi, simmetrica rispetto allo 0 ed ha il vantaggio di avere supporto compatto. In questo caso il test bootstrap è indicato B_S.

Schematicamente le simulazioni sono state condotte in base al seguente algoritmo:

1) si genera una serie temporale x_t di lunghezza T dal modello (1) sotto l'ipotesi nulla $\alpha=1$;

2) si calcola il valore della statistica $\tau = (\hat{\alpha}-1)S^{-1}\left(\sum_{t=1}^{T} x_{t-1}^2\right)^{1/2}$;

3) si genera una replicazione della serie storica x_t mediante il bootstrap esterno[1] e su questa si calcola il valore della statistica τ che si indicherà con τ^*;

4) si ripetono N volte i passi 1, 2, 3: si ottengono N valori di τ e τ^*;

5) sulla base degli N valori di τ^* si calcola il percentile corrispondente al livello del test $(1-\delta)$: si è ottenuta in questo modo l'estremo della regione di accettazione con il metodo bootstrap;

6) si calcola la proporzione δ_{DF} di volte in cui gli N valori di τ non appartengono alla regione di accettazione per il test DF con valori critici tabulati da Mackinnon (Mackinnon 1991);

7) si calcola la proporzione δ_B di volte in cui gli N valori di τ non appartengono alla regione di accettazione calcolata con il metodo bootstrap esterno.

Per ottenere stime accurate si è scelto un valore per N pari a 10000.

3. ETEROSCHEDASTICITÀ E SERIE AD ALTA FREQUENZA

Si analizza il comportamento del test DF in serie di lunghezza finita generate da un random walk con errori di tipo GARCH(1,1), (Bollerslev, 1986):

$$x_t = x_{t-1} + \varepsilon \quad \mathrm{t} = 1,2,\ldots,\mathrm{T} \qquad x_0 = 0$$

dove ε_t sono variabili casuali Gaussiane con media zero, e varianza

$$\sigma_t{}^2 = \phi_0 + \phi_1 \varepsilon_{t-1}^2 + \phi_2 \sigma_{t-1}^2\ , \quad \sigma_0^2 = 1, \quad \varepsilon_0 = 0\,.$$

Nelle simulazioni, le quantità variabili, ovvero i parametri del modello GARCH(1,1) ed il numero delle osservazioni sono uguali a quelle considerate da Procidano e Rigatti Luchini (1999, 2000). Più precisamente, si sono considerate le numerosità T=25, 50, 100, 200 e per alcune elaborazioni anche la numerosità T=500. Per quanto riguarda i parametri ϕ_0, ϕ_1, ϕ_2, sono state utilizzate le seguenti terne: (0.1, 0.3, 0.7), (0.0009, 0.09, 0.91), (0.0001, 0.03, 0.97).

Le serie ad alta frequenza sono state generate utilizzando un risultato di Nelson (Nelson, 1991), il quale dimostra che, posto $\gamma>0$ l'intervallo con il quale avvengono le rilevazione di una serie storica, un modello GARCH con parametri:

$$\phi_0 = \omega\gamma, \qquad \phi_1 = \psi\gamma^{1/2}, \qquad \phi_2 = 1 - \phi_1$$

con ω e ψ opportunamente fissati, approssima un processo di diffusione a tempo continuo quando γ tende a zero.

A partire da questo risultato, si valuta se e quali distorsioni si hanno sui test DF e B al tendere di γ a zero. I parametri del modello GARCH sono stati scelti in modo tale da considerare processi integrati, mentre a ϕ_0 sono stati attribuiti tre differenti valori che individuano processi via via più simili ad un processo GARCH degenere. Le simulazione (Tab. 1) sono state ottenute fissando per ω=0.01, ψ=0.3, γ=1, 0.09, 0.01.

---

[1] La procedura seguita è descritta in Procidano e Rigatti Luchini (Procidano e Rigatti Luchini, 2001).

Il comportamento del test DF risulta sostanzialmente differente da quanto si era osservato nelle simulazioni condotte in ipotesi di eteroschedasticità con processi GARCH integrati o quasi integrati. In quel caso, infatti, si era registrato un aumento della distorsione al crescere della numerosità campionaria T (Procidano e Rigatti Luchini, 1999, 2000, 2001). Ora, invece, i risultati ottenuti indicano che, con processi GARCH integrati, al decrescere di γ il test DF non è significativamente distorto. Ciò è probabilmente da imputare ai piccoli valori attribuiti a ϕ_0.

TABELLA 1

Proporzione di rigetti nell'ipotesi nulla α=1 su N=10.000 serie storiche;
ω=0.01, ψ=0.3, livello nominale 1-δ=0.95

Test	T=25	T=50	T=100	T=200	T=500
		$\gamma = 1$	$\phi_0 = 0.1$ $\phi_1 = 0.3$ $\phi_2 = 0.7$		
DF	**0.0632**	**0.0698**	**0.0786**	**0.0761**	**0.0861**
B_A	*0,0559*	*0.0535*	*0.0553*	*0.0512*	*0.0542*
B_S	*0.0588*	*0.0621*	*0.0593*	*0.0547*	*0.0543*
		$\gamma = 0.09$	$\phi_0 = 0.0009$ $\phi_1 = 0.09$ $\phi_2 = 0.91$		
DF	**0.0471**	**0.0576**	**0.0554**	**0.0648**	**0.0681**
B_A	*0.0555*	*0.0496*	*0.0521*	*0.0476*	*0.0476*
B_S	*0.0536*	*0.0605*	*0.0524*	*0.0519*	*0.0493*
		$\gamma = 0.01$	$\phi_0 = 0.0001$ $\phi_1 = 0.03$ $\phi_2 = 0.97$		
DF	**0.0464**	**0.0539**	**0.0522**	**0.0569**	**0.0492**
B_A	*0.0470*	*0.0548*	*0.0522*	*0.0556*	*0.0455*
B_S	*0.0535*	*0.0592*	*0.0556*	*0.0541*	*0.0476*

4. PRESENZA DI OUTLIERS

Si analizzano ora le prestazione dei test DF e B nel caso in cui siano presenti nella serie storica degli outliers di tipo additivo o innovativo. Le performance del test B vengono confrontate con quelle del test DF tramite simulazione di serie storiche di numerosità T=25, 50, 100, 200.

Un outlier viene definito come additivo quando si presenta come uno shock che colpisce le realizzazioni di un processo stocastico solo in un determinato istante. Formalmente, se indichiamo con y_t un processo stocastico di tipo random-walk, ossia:

$$y_t = y_{t-1} + \varepsilon_t$$

dove ε_t è un processo white noise, allora un processo x_t affetto da un outlier di tipo additivo può essere espresso da:

$$x_t = y_t + \delta_t \beta_A$$

dove β_A rappresenta l'ampiezza dello shock e

$$\delta_t = \begin{cases} 0 & \text{se all'istante t non è presente un outlier} \\ 1 & \text{altrimenti} \end{cases}$$

Un outlier si definisce innovativo quando, anziché colpire le realizzazioni di un processo stocastico, si manifesta come uno shock che colpisce il termine di errore ε_t, formalmente:

$$x_t = x_{t-1} + a_t \tag{2}$$

$$a_t = \varepsilon_t + \delta_t \beta_I \qquad \delta_t = \begin{cases} 0 & \text{per } t < T_1 \\ 1 & \text{per } t \geq T_1 \end{cases}$$

Un outlier di tipo innovativo produce, quindi, un salto permanente di ampiezza β_I nelle realizzazioni del processo stocastico non contaminato.

L'esperimento di simulazione è stato impostato nel modo seguente: in primo luogo si sono generate serie storiche contaminate da un solo outlier e successivamente da un numero *n* pari al 4% della numerosità. Le ampiezze degli outliers considerati sono state fatte variare da $\pm 3\sigma$ a $\pm 5\sigma$ dove σ^2 è la varianza del termine di errore ε_t. Le tabelle 2 e 3 riassumono i risultati delle simulazioni nelle due differenti ipotesi.

TABELLA 2

Proporzione di rigetti dell'ipotesi nulla α=1 su N=10.000 serie storiche contaminate con 1 outlier additivo; livello nominale 1-δ=0.95

Test	T=25	T=50	T=100	T=200
		β_A=3		
DF	**0.1350**	**0.0947**	**0.0781**	**0.0612**
B_A	*0.1332*	*0.0948*	*0.0842*	*0.0640*
B_S	*0.1080*	*0.0848*	*0.0753*	*0.0612*
		β_A=4		
DF	**0.1945**	**0.1337**	**0.0979**	**0.0709**
B_A	*0.1858*	*0.1283*	*0.0970*	*0.0737*
B_S	*0.1361*	*0.1051*	*0.0895*	*0.0734*
		β_A=5		
DF	**0.2662**	**0.1778**	**0.1227**	**0.0833**
B_A	*0.2384*	*0.1625*	*0.1194*	*0.0841*
B_S	*0.1556*	*0.1310*	*0.1105*	*0.0828*

TABELLA 3

Proporzione di rigetti dell'ipotesi nulla α=1 su N=10.000 serie storiche contaminate con N=0.04 T outliers additivi; livello nominale 1-δ=0.95

Test	T=25	T=50	T=100	T=200
		β_A=3		
DF	**0.1350**	**0.1453**	**0.1421**	**0.1449**
B_A	*0.1332*	*0.1475*	*0.1539*	*0.1471*
B_S	*0.1080*	*0.1373*	*0.1514*	*0.1440*
		β_A=4		
DF	**0.1945**	**0.2109**	**0.2209**	**0.2168**
B_A	*0.1858*	*0.2059*	*0.2262*	*0.2292*
B_S	*0.1361*	*0.1777*	*0.2182*	*0.2209*
		β_A=5		
DF	**0.2662**	**0.2887**	**0.2986**	**0.2952**
B_A	*0.2384*	*0.2772*	*0.3010*	*0.3046*
B_S	*0.1556*	*0.2341*	*0.2863*	*0.3016*

Nel caso di una contaminazione con un solo outlier additivo (Tab. 2), i test manifestano una chiara tendenza a sovrarigettare l'ipotesi nulla quando è vera. La distorsione tende a ridursi al crescere della numerosità campionaria *T*, confermando l'intuizione che la presenza di un solo outlier ha effetti sempre minori quanto maggiore è la lunghezza di una serie storica. Tra le due distribuzioni bootstrap, le performance migliori si hanno con quella simmetrica.

Mentre la posizione degli outliers additivo non influenza i risultati delle simulazioni, la posizione è rilevante per gli outliers innovativi, poiché il loro effetto non rimane isolato ma permane nel tempo. I primi risultati ottenuti si riferiscono al caso in cui gli outliers innovativi si trovino equidistanziati gli uni dagli altri.

Le tabelle 4 e 5 riportano i risultati delle simulazioni nel caso di outliers innovativi pari a $n= 0.04\ T$, di segno concorde e successivamente di segno discorde.

Nella situazione di contaminazione con outliers innovativi di segno concorde, sia il test DF che il test B denotano una tendenza a sottorigettare l'ipotesi nulla, che si accentua al crescere della numerosità campionaria. Quando si considerano outlilers innovativi di segno alterno si assiste viceversa ad una tendenza per entrambi i test a sovrarigettare l'ipotesi nulla.

TABELLA 4

Proporzione di rigetti dell'ipotesi nulla α=1 su N=10.000 serie storiche contaminate con N=0.04 T outliers innovativi; livello nominale 1-δ=0.95

Test	T=25	T=50	T=100	T=200
		β_I=3		
DF	**0.0498**	**0.0448**	**0.0371**	**0.0221**
B_A	*0,0627*	*0.0509*	*0.0397*	*0.0252*
B_S	*0.0541*	*0.0477*	*0.0393*	*0.0226*
		β_I=4		
DF	**0.0440**	**0.0428**	**0.0300**	**0.0106**
B_A	*0.0557*	*0.0491*	*0.0341*	*0.0121*
B_S	*0.0511*	*0.0491*	*0.0325*	*0.0109*
		β_I=5		
DF	**0.0398**	**0.0351**	**0.0193**	**0.0025**
B_A	*0.0528*	*0.0417*	*0.0215*	*0.0027*
B_S	*0.0523*	*0.0416*	*0.0208*	*0.0026*

TABELLA 5

Proporzione di rigetti dell'ipotesi nulla α=1 su N=10.000 serie storiche contaminate con N=0.04 T outliers innovativi di segno alterno; livello nominale 1-δ=0.95

Test	T=25	T=50	T=100	T=200
		β_I=3		
DF	**0.0498**	**0.0524**	**0.0618**	**0.0746**
B_A	*0.0627*	*0.0641*	*0.0680*	*0.0810*
B_S	*0.0541*	*0.0576*	*0.0634*	*0.0711*
		β_I=4		
DF	**0.0440**	**0.0513**	**0.0679**	**0.0951**
B_A	*0.0557*	*0.0636*	*0.0795*	*0.1075*
B_S	*0.0541*	*0.0617*	*0.0783*	*0.0939*
		β_I=5		
DF	**0.0398**	**0.0494**	**0.0734**	**0.1189**
B_A	*0.0528*	*0.0638*	*0.0926*	*0.1357*
B_S	*0.0523*	*0.0631*	*0.0900*	*0.1216*

5. CONCLUSIONI

I risultati degli esperimenti di simulazione hanno evidenziato che il test DF sembra essere robusto quando viene utilizzato su serie campionate ad alta frequenza. Il livello di distorsione seppur relativamente maggiore a quello ottenuto con il test B risulta contenuto. In particolare tende a crescere all'aumentare della numerosità della serie storica ad eccezione del caso con disturbo GARCH di parametri (ϕ_0=0.1, ϕ_1=0.3, ϕ_2=0.7). Il comportamento del test B, in entrambe le distribuzioni esterne, conferma i risultati ottenuti in altri esperimenti di simulazione (Procidano e Rigatti Luchini, 1999, 2000) confermandone la robustezza rispetto alla frequenza di rilevazione e alla caratteristica di processo GARCH degenere.

La presenza di uno o più outlier additivi produce distorsione nel test DF e, in misura minore, nel test B. La distorsione aumenta all'aumentare del numero di outliers che intervengono nella serie storica: si ottengono risultati migliori nel caso di numerosità contenute e quando si utilizza la distribuzione bootstrap simmetrica.

Per quanto riguarda gli outliers innovativi, non si sono riscontrate evidenti differenze tra test DF e B. A priori, e limitatamente alle simulazioni effettuate, non pare possibile stabilire il segno della distorsione. Tuttavia maggiormente dannosi si sono rivelati gli outlier di tipo additivo, come era presumibile aspettarsi, visto che la presenza di un outlier di tipo innovativo in un random walk non è distinguibile da un cambio di livello stocastico in una serie storica. Outlier additivi di modesta ampiezza, quindi non facilmente identificabili in una serie storica anche di elevata numerosità campionaria, hanno prodotto distorsioni significative.

Complessivamente, il test DF appare meno adeguato del test B per la verifica della presenza di radici unitarie in caso di contaminazioni con outlier anche se, a volte, ha mostrato una buona performance.

Dipartimento di Statistica CLAUDIO PIZZI
Università Ca' Foscari, Venezia ISABELLA PROCIDANO

Dipartimento di Scienze Statistiche SILIO RIGATTI LUCHINI
Università di Padova

BIBLIOGRAFIA

T. BOLLERSLEV (1987), *A conditionally heteroschedasticity time series model for speculative prices an rates of return*. "The Review of Economics and Statistics", 69, pp. 542-547.

D.A. DICKEY, W.A. FULLER (1979), *Distribution of the estimators for autoregressive time series with unit root*, "Journal of American Statistical Association", 74, pp. 427-431.

W.A. FULLER (1996), *Introduction to statistical time series,* Willey, New York.

R.Y. LIU (1988) *Bootstrap procedures under some non-i.i.d. models.* "Annals of Statistics", 16, pp. 1697-1708.

J.C. MACKINNON (1991) *Critical values for cointegration tests*, in R.F. Engle and C.W.J. Granger (eds.) "Long Run Economic Relationships", Oxford University Press, pp. 267-276.

E. MAMMEN (1993) *Bootstrap and wild bootstrap for high dimensional linear models*, "Annals of Statistics", 21, pp. 255-285.

M. MORENO, J. ROMO (2000), *Bootstrap tests for unit roots based on LAD estimation*, "Journal of statistical Planning and Inference", 83, pp. 347-367.

D.B. NELSON (1991), *ARCH model as diffusion approximations*, "Journal of Econometrics", 45, pp. 7-38.

P.C.B. PHILLIPS, P. PERRON (1988), *Testing for a unit root in time series regression* , "Biometrika", 75, pp. 335-346.

I. PROCIDANO, S. RIGATTI LUCHINI (1999), *Verifica di radici unitarie mediante bootstrap,* Convegno S.CO. 99, Venezia 27-29 settembre 1999, pp. 242-247.

I. PROCIDANO, S. RIGATTI LUCHINI (2000), *Il bootstrap esterno nella verifica di radici unitarie in condizioni di eteroschedasticità*, XL Riunione scientifica della Società italiana di Statistica, Firenze, 26-29 aprile 2000, pp. 303-306.

I. PROCIDANO, S. RIGATTI LUCHINI (2002), *Testing unit roots by bootstrap*, "Metron", 60, 1-2, pp. 175-189.

C.F.J. WU (1986), *Jackknife, bootstrap and other resampling methods in regression analysis*, *(with discussion)*, "Annals of Statistics", 14, pp. 1261-1295.

RIASSUNTO

Il test bootstrap esterno per la ricerca di radici unitarie in presenza di outliers

Il test Dickey-Fuller è uno strumento statistico utile ad individuare la presenza di radici unitarie nelle serie storiche. Il presente lavoro confronta la robustezza del test Dickey-Fuller con quella della sua estensione bootstrap da due diversi punti di vista. Da un lato si studiano le conseguenze della violazione delle ipotesi di omoschedasticità e di incorrelazione dei residui, dall'altro si analizzano gli effetti di una contaminazione di outliers additivi e innovativi della serie storica. Il comportamento sotto queste due particolari situazioni del test Dickey-Fuller è confrontato con i risultati ottenuti dalla versione bootstrap dello stesso test.

I risultati ottenuti evidenziano la necessità di utilizzare la versione bootstrap nel caso di eteroschedasticità dei residui e nel caso in cui si possa sospettare la presenza di outliers nella serie storica.

SUMMARY

The wild bootstrap for unit root test in presence of outliers

The Dickey-Fuller test is a usefull statistical tool to detect unit roots in time series. This paper compare the robustness of the Dickey-Fuller test and its bootstrap extention from two different point of view. On one hand we study the conseguences of the violations of the residuals omoschedasticity and incorrelation hypotheses, on the other hand we analize the effect of the additive and innovative outliers contamination of time series. The results point out that the bootstrap test give us better performace than the Dickey-Fuller one when the time series is contaminated by outliers and in some cases also when the residuals are correlated or heteroschedastic.