

Università degli Studi di Padova
Dipartimento di Scienze Statistiche
Corso di Laurea Triennale in
Statistica e Gestione delle Imprese



RELAZIONE FINALE
**GLI EFFETTI DEL VALORE INIZIALE SUI TEST DI
RADICE UNITARIA**

Relatore Prof. Luisa Bisaglia
Dipartimento di Scienze Statistiche

Laureando: Serena Pavan
Matricola 1010935

Anno Accademico 2012/2013

Sommario:

| | |
|--|----|
| Introduzione..... | 5 |
| 1. Serie storiche stazionarie e non stazionarie..... | 7 |
| 1.1 Processi stocastici stazionari..... | 7 |
| 1.2 Processi stocastici non stazionari..... | 8 |
| Processi trend-stazionari e differenza-stazionari..... | 8 |
| 1.2.1 Processi stocastici non stazionari più comuni..... | 9 |
| Random walk..... | 9 |
| Processo autoregressivo integrato a media mobile, ARIMA(p, d, q)..... | 11 |
| 2. Test di radice unitaria | 13 |
| 2.1 Test Dickey-Fuller..... | 15 |
| Modello con drift..... | 17 |
| Modello con trend lineare..... | 18 |
| Estensioni del test DF: il test ADF (Augmented DF)..... | 19 |
| 2.2 Test NIV (Nonlinear Instrumental Variable, Yooson Chang, 2002)..... | 22 |
| 3. Gli effetti del valore iniziale sui test di radice unitaria..... | 25 |
| 3.1 Procedure ricorsive di <i>demeaning</i> e <i>detrending</i> | 25 |
| 3.2 Gli effetti del valore iniziale: l'esperimento Monte Carlo..... | 27 |
| Conclusione..... | 47 |
| Appendice A: Codice R utilizzato..... | 49 |
| Bibliografia e sitografia..... | 55 |

Introduzione

Nei test di radice unitaria è frequente l'uso di procedure ricorsive, quali *demeaning* e *detrending*, per ottenere stime più precise dei parametri e una maggiore potenza dei test.

In questo lavoro s'intendono esaminare gli effetti del valore iniziale sui test di radice unitaria quando alla serie originaria sono state applicate procedure ricorsive di *demeaning* o *detrending*.

Mediante esperimenti Monte Carlo verrà esaminato, infatti, come cambi la potenza del test al variare del valore iniziale. Verranno applicati sia dei test di radice unitaria di tipo Dickey-Fuller basati sulla stima del parametro col metodo dei minimi quadrati ordinari, sia dei test NIV (Nonlinear Instrumental Variable) basati sulla stima del parametro tramite il metodo delle variabili strumentali.

Tale analisi muove principalmente dall'articolo di Ming Meng, Hyejin Lee, Myeong Hyeon Cho, Junsoo Lee "Impact of the initial observation on unit root tests using recursive demeaning and detrending procedures" (2013), al quale ci si attiene anche per il valore ottimale di c da usare nel calcolo della variabile strumentale e per i valori critici dei test NIV.

Questo studio è quindi articolato come segue. Nel primo capitolo s'introdurranno i concetti di stazionarietà e non stazionarietà, passando in rassegna i principali processi non stazionari. Sfruttando tali concetti si descriveranno poi nel secondo capitolo i test di radice unitaria, nello specifico i test di tipo Dickey-Fuller e i test di tipo NIV. Utilizzando tali test nel capitolo tre verrà formulato l'esperimento Monte Carlo i cui risultati permetteranno di osservare l'effetto del valore iniziale della serie storica, a cui sono state applicate determinate procedure ricorsive di *demeaning* e *detrending*, su tali test di radice unitaria.

1. Serie storiche stazionarie e non stazionarie

1.1 Processi stocastici stazionari

Un processo stocastico è stazionario se la sua dinamica non presenta variazioni strettamente periodiche e se né la sua media né la sua variabilità presentano cambiamenti di natura sistematica.

Da un punto di vista più formale, un processo stocastico si dice *stazionario in senso stretto* se le distribuzioni congiunte, rispettivamente, di

$$(Y_{t_1}, \dots, Y_{t_n})$$

e di

$$(Y_{t_1+\tau}, \dots, Y_{t_n+\tau})$$

per ogni insieme di istanti t_1, \dots, t_n e per ogni τ , sono uguali.

Per $n=1$ la stazionarietà in senso stretto implica che la distribuzione di Y_t sia la stessa per ogni t e perciò, tenuto conto che i primi due momenti siano finiti, media e varianza sono entrambe costanti e non dipendono dall'istante t :

$$\mu_t = \mu \quad \forall t$$

$$\sigma_t^2 = \sigma^2 \quad \forall t$$

La *funzione di autocovarianza* risulta invece:

$$\gamma_{t_1, t_2} = E\{(Y_{t_1} - \mu)(Y_{t_2} - \mu)\} = \gamma_{|t_2 - t_1|}$$

essa dipende cioè solo dalla distanza temporale tra le due variabili, ma non dagli istanti t_1 e t_2 . Indicando quest'ultimi con $t_1=t$ e $t_2=t+k$ essa è riscrivibile come:

$$\gamma_{t, t+k} = E\{(Y_t - \mu)(Y_{t+k} - \mu)\} = \gamma_k, k = 0, \pm 1, \pm 2, \dots \quad (1)$$

La condizione di stazionarietà in senso stretto, ponendo vincoli sull'intera distribuzione del processo e quindi sui momenti di qualsiasi ordine, è piuttosto forte. Risulta quindi utile, in pratica, definire la stazionarietà in modo meno restrittivo del precedente.

Un processo stocastico Y_t è detto *stazionario in senso debole* (o *stazionario del secondo ordine*) se:

- $E[Y_t] = \mu < \infty \quad \forall t$
- $\text{Var}(Y_t) = \sigma^2 < \infty \quad \forall t$
- $\text{Cov}(Y_t, Y_{t+k}) = E[(Y_t - \mu)(Y_{t+k} - \mu)] = \gamma_k \quad \forall t, \forall k$

Sui momenti di ordine superiore non viene fatta alcuna assunzione.

Le condizioni sopra esposte affermano quindi che un processo stocastico è stazionario se i suoi momenti sono finiti ed indipendenti dal tempo.

La nozione di stazionarietà debole è quella più usata in pratica, in parte per la comodità, e in parte perché per alcuni processi la stazionarietà del secondo ordine è già soddisfacente.

1.2 Processi stocastici non stazionari

La teoria sottostante i modelli fin qui considerati è basata sull'ipotesi di stazionarietà. Tuttavia, in economia, ma non solo, è piuttosto comune la presenza di serie provenienti da un processo generatore dei dati non stazionario.

Limitandosi ai primi due momenti di un processo, la non stazionarietà può dipendere dal fatto che la media e/o la varianza non sono costanti. Nel primo caso si tratta di *non stazionarietà in media*, che tipicamente si verifica in presenza di un trend, nel secondo di *non stazionarietà in varianza*. In questo lavoro si farà riferimento al primo caso di non stazionarietà.

Processi trend-stazionari e differenza-stazionari

Come sopraccitato una delle principali cause di non stazionarietà è la presenza di un trend, ovvero di un movimento persistente di lungo periodo di una variabile nel corso del tempo, la tendenza di fondo del fenomeno considerato.

Nell'ambito economico si incontrano generalmente due tipi di trend:

1. *Trend deterministico*
2. *Trend stocastico*.

Un *trend deterministico* è descritto da una funzione deterministica (non aleatoria) del tempo $f(t)$.

Se la funzione $f(t)$ è lineare il trend sarà del tipo:

$$f(t) = \alpha_0 + \alpha_1 t$$

Un trend di tipo deterministico è quindi completamente "prevedibile" una volta noti i coefficienti che lo specificano.

In generale, i modelli con tale trend sono del tipo

$$Y_t = f(t) + u_t$$

composti cioè da $f(t)$, funzione deterministica, e da una componente di errore u_t assunta solitamente come un processo stocastico stazionario a media nulla. Nella sua forma più comune è $u_t \sim WN(0, \sigma_u^2)$.

Viene infatti considerata la presenza nel processo di una serie di oscillazioni casuali intorno ad un trend (in questo caso ci limitiamo ad un trend lineare), in cui la media del processo è una funzione lineare del tempo:

$$E[Y_t] = \mu_t = \alpha_0 + \alpha_1 t \quad (3)$$

Processi di questo tipo sono detti *trend-stazionari*, perché dopo aver determinato $f(t)$ e averlo sottratto alla serie, ciò che si ottiene è un processo stazionario.

Nella seconda tipologia di trend invece, la componente di fondo varia nel tempo in maniera aleatoria e quindi non completamente prevedibile.

Questa specificazione è detta *processo a radice unitaria*:

$$(1 - B)Y_t = \beta + \psi(B)\varepsilon_t \quad (4)$$

$$\text{con } \psi(B) = \sum_{i=0}^{\infty} \psi_i B^i, \quad \psi_0 = 1$$

poiché la sua equazione caratteristica $1-B = 0$ implica la soluzione unitaria $B=1$. Per rendere stazionario tale processo la procedura corretta non è più quella di rimozione del trend stimato, come nel caso precedente, ma quella di applicazione dell'operatore differenza $\Delta=1-B$. Tali modelli infatti sono chiamati *differenza-stazionari* proprio perché possono essere resi stazionari tramite l'applicazione dell'operatore differenza.

Un esempio tipico di processo di radice unitaria, quindi con trend stocastico è il random walk.

1.2.2 Processi stocastici non stazionari più comuni

Random walk

Sia $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$. Si dice che $\{Y_t\}$ è una *passeggiata casuale (random walk)* se:

$$Y_t = Y_{t-1} + \varepsilon_t \quad Y_0 = \mu.$$

Essendo

$$Y_t = \mu + \sum_{i=1}^t \varepsilon_i,$$

il quale è governato da un trend stocastico prodotto dal cumulo degli errori fino al tempo t con pesi tutti uguali ad uno (processo imprevedibile), si ha:

$$E[Y_t] = \mu,$$

$$Var[Y_t] = t\sigma_\varepsilon^2.$$

Poiché la varianza cambia nel tempo e, in particolare, cresce con esso, il processo è non stazionario. Non stazionarietà che può anche essere desunta dal fatto che il random walk definisce un processo AR(1) con $\phi_1 = 1$ che non rispetta la condizione di stazionarietà $|\phi_1| < 1$.

La funzione di autocorrelazione è pari a

$$\rho_k = \sqrt{\frac{t}{t+k}} \quad \text{con } k = 0, 1, 2, \dots$$

dunque per ogni prefissato valore di k $\rho_k \rightarrow 1$ al crescere di t ; per questo motivo si dice anche che il processo ha memoria infinita.

In modo analogo la funzione di autocorrelazione parziale è:

$$P_k = \begin{cases} 1 & k = 1 \\ 0 & k > 1 \end{cases}$$

Si osserva infine, che pur essendo non stazionario, questo processo è riconducibile tramite differenziazione ad un white noise:

$$\Delta Y_t = Y_t - Y_{t-1} = \varepsilon_t \quad \text{ed è stato così ottenuto un processo stazionario.}$$

Il processo in questione viene anche detto *integrato di ordine uno*, I(1) perché Y_t ha una radice unitaria ovvero la sua equazione caratteristica $1-B=0$ implica la soluzione unitaria $B=1$. Qualora, più un generale, un processo X_t fosse I(d), allora X_t sarebbe un processo a d radici unitarie perché l'equazione $(1-B)^d=0$ ha la soluzione $B=1$ con molteplicità d .

Aggiungendo un drift al processo random walk si ottiene invece:

$$Y_t = \mu + Y_{t-1} + \varepsilon_t.$$

Tale processo inizia solitamente in zero per $t = 0$ così che $Y_1 = \varepsilon_1$ e

$$Y_t = t\mu + \sum_{i=1}^t \varepsilon_i.$$

Essendo

$$E[Y_t] = t\mu,$$

$$Var[Y_t] = t\sigma_\varepsilon^2$$

il processo è non stazionario, sia in media sia in varianza.

Tuttavia anche in questo caso la differenza prima del processo

$$\Delta Y_t = (1-B)Y_t = Y_t - Y_{t-1} = \mu + \varepsilon_t$$

dà luogo ad un processo puramente casuale e quindi è stazionaria.

Processo autoregressivo integrato a media mobile, ARIMA (p, d, q)

Un'estensione dei modelli ARMA per processi non stazionari, ma omogenei di ordine d e che possono quindi essere resi stazionari tramite opportune trasformazioni, sono i modelli ARIMA(p, d, q).

Sia $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$. Si dice che $\{Y_t\}$ è un processo autoregressivo integrato a media mobile di ordine (p, d, q), e lo si indica con ARIMA(p, d, q), se:

- ✓ $(1 - B)^d Y_t = X_t$ (X_t è la d -esima differenza di Y_t),
- ✓ X_t è modellabile con un ARMA(p, q).

Utilizzando l'operatore differenza ed i polinomi, rispettivamente, autoregressivo ed a media mobile esso può essere riscritto, in termini di Y_t :

$$\phi(B)(1 - B)^d Y_t = \phi_0 + \theta(B)\varepsilon_t$$

dove gli operatori $\phi(B)$ e $\theta(B)$ non hanno zeri in comune. Per le caratteristiche di tale processo ci si riconduce quindi poi ad un ARMA(p, q).

Da quest'analisi emerge quindi quanto sia importante capire se la non stazionarietà dei dati in questione è dovuta alla presenza di un processo generatore trend-stazionario o differenza-stazionario.

Dal punto di vista economico infatti si è interessati, ad esempio, a sapere se le conseguenze di periodi di recessione avranno effetti duraturi sui livelli futuri di alcune variabili chiave del sistema economico (il PIL ad esempio). Se il processo fosse random walk, per ogni k fissato

$$\text{Corr}[Y_t, Y_{t+k}] \rightarrow 1 \text{ al crescere di } t,$$

quindi per tale processo un qualsiasi shock aleatorio avrebbe un effetto permanente. Capire se si tratta di tale processo avrebbe quindi una grossa rilevanza.

Riuscire a distinguere la natura della non stazionarietà permette anche di adottare la trasformazione dei dati ideale che consenta di ricondursi correttamente al caso stazionario. Un altro aspetto con importanti risvolti pratici, che giustificano l'interesse per questa distinzione è quello previsivo.

Per prevedere un processo trend-stazionario del tipo (3) infatti è sufficiente aggiungere la componente deterministica ($\alpha_0 + \alpha_1 t$), assunta

nota, alla previsione della componente stocastica stazionaria, e dunque la previsione k passi in avanti al tempo t sarà:

$$\hat{y}_{t+k} = \alpha_0 + \alpha_1(t+k) + \hat{u}_{t+k|t} \quad (5)$$

con \hat{u}_{t+k} previsione di u_{t+k} al tempo t .

È possibile mostrare (Hamilton, 1994, p.439) che, al crescere dell'orizzonte temporale k , tale previsione converge in media quadratica al trend:

$$\lim_{k \rightarrow \infty} E[\hat{Y}_{t+k} - \alpha_0 - \alpha_1(t+k)]^2 = 0.$$

La previsione di un processo a radice unitaria come

$$Y_t = \beta + Y_{t-1} + \varepsilon_t \quad (6)$$

è data da:

$$\hat{y}_{t+k} = \beta + y_t. \quad (7)$$

Dunque per un random walk con drift (6) la crescita attesa si manifesta ad un tasso costante e pari a β , qualunque sia il punto di partenza y_t .

Con entrambe le specificazioni quindi la previsione \hat{y}_{t+k} in (5) e (7) è data da una funzione lineare dell'orizzonte di previsione k , la differenza sta nell'intercetta. Per un processo trend-stazionario la previsione è una retta la cui intercetta è la stessa indipendentemente dal valore di y_t . Per un processo differenza-stazionario, invece, l'intercetta cambia continuamente ad ogni nuova osservazione y_t .

2. Test di radice unitaria

Come spiegato nel capitolo precedente molte serie storiche economiche non sono stazionarie. È quindi importante riuscire ad individuare la forma più appropriata del trend presente nei dati, sia per l'interpretazione economica sia per la modellazione, in quanto modellare, ad esempio, una serie secondo un modello ARMA(p,q) è possibile solo se la serie è stata prima resa stazionaria con trasformazioni appropriate.

I due metodi visti per rimuovere il trend sono: la rimozione del trend stimato e l'applicazione dell'operatore differenza.

La rimozione del trend è appropriata per serie trend-stazionarie, $I(0)$, mentre l'operatore differenza di ordine d è applicabile ad una serie con trend stocastico e quindi differenza-stazionaria $I(d)$. I test di radice unitaria possono essere usati proprio per determinare se i dati devono essere resi stazionari tramite l'una o l'altra procedura.

- Se la serie è trend-stazionaria, il processo generatore dei dati risulta (limitandoci ad un trend lineare):

$$Y_t = \gamma_0 + \gamma_1 t + e_t$$

dove e_t è un processo stazionario e invertibile a media nulla.

- Se la serie è differenza-stazionaria, il processo generatore dei dati risulta (limitandoci ai processi $I(1)$):

$$Y_t = \alpha_0 + Y_{t-1} + e_t$$

dove, ancora, e_t è un processo stazionario e invertibile a media nulla. Se $e_t \sim WN$, Y_t è un random walk con drift.

- Se combiniamo i due modelli nell'unico modello:

$$Y_t = \gamma_0 + \gamma_1 t + u_t$$

con

$$u_t = \rho u_{t-1} + e_t \quad (u_t \text{ è quindi ipotizzato un AR}(1))$$

possiamo scrivere:

$$\begin{aligned} Y_t &= \gamma_0 + \gamma_1 t + \rho u_{t-1} + e_t \\ &= \gamma_0 + \gamma_1 t + \rho[Y_{t-1} - \gamma_0 - \gamma_1(t-1)] + e_t \\ &= \gamma_0(1-\rho) + \gamma_1\rho + \gamma_1(1-\rho) + \rho Y_{t-1} + e_t \\ &= \beta_0 + \beta_1 t + \rho Y_{t-1} + e_t \end{aligned}$$

dove $\beta_0 = \gamma_0(1-\rho) + \gamma_1\rho$ e $\beta_1 = \gamma_1(1-\rho)$.

In tal caso se:

- ✓ $\rho = 1$ allora $\beta_0 = \gamma_1$ e $\beta_1 = 0$, quindi il modello diventa

$$Y_t = \gamma_1 + Y_{t-1} + e_t$$

che è un processo differenza-stazionario.

- ✓ $\rho < 1$ allora diventa

$$Y_t = \beta_0 + \beta_1 t + \rho Y_{t-1} + e_t$$

che è un processo trend-stazionario.

Discriminare fra i due modelli significa quindi verificare se $\rho = 1$ contro l'alternativa $\rho < 1$.

- Se il modello di partenza è

$$Y_t = \gamma_0 + u_t$$

$$u_t = \rho u_{t-1} + e_t$$

Procedendo come sopra si ottiene:

$$Y_t = \gamma_0(1 - \rho) + \rho Y_{t-1} + e_t$$

$$= \beta_0 + \rho Y_{t-1} + e_t$$

dove $\beta_0 = \gamma_0(1 - \rho)$.

- ✓ Se $\rho = 1$, allora $\beta_0 = 0$ e il modello si riduce a

$$Y_t = Y_{t-1} + e_t, \text{ cioè un modello differenza-stazionario.}$$

- ✓ Se $\rho < 1$, il modello diventa

$$Y_t = \beta_0 + \rho Y_{t-1} + e_t, \text{ cioè } Y_t \text{ è stazionario con media diversa da zero.}$$

Discriminare tra i due modelli significa pertanto verificare l'ipotesi nulla $H_0: \rho = 1$ contro l'alternativa $H_1: \rho < 1$.

- Se il processo generatore dei dati è

$$Y_t = \gamma_0 + \gamma_1 t + \gamma_2 t^2 + u_t$$

$$u_t = \rho u_{t-1} + e_t$$

allora, procedendo come sopra, si ottiene

$$Y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \rho Y_{t-1} + e_t,$$

dove: $\beta_0 = \gamma_0(1 - \rho) + (\gamma_1 - \gamma_2)\rho$,

$$\beta_1 = \gamma_1(1 - \rho) + 2\gamma_2\rho,$$

$$\beta_2 = \gamma_2(1 - \rho).$$

Da notare che:

- ✓ Se $\rho = 1$, allora $\beta_0 = \gamma_1 - \gamma_2 = \alpha_0$, $\beta_1 = 2\gamma_2 = \alpha_1$ e $\beta_2 = 0$ e il modello si riduce a

$$Y_t = \alpha_0 + \alpha_1 t + Y_{t-1} + e_t,$$

che è un modello differenza-stazionario: esso sottende infatti un trend deterministico quadratico e un trend stocastico.

✓ Se $\rho < 1$, il modello è

$$Y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \rho Y_{t-1} + e_t,$$

cioè Y_t è un modello trend-stazionario, costituito da un trend deterministico polinomiale di secondo grado più un errore stazionario.

In conclusione, per tutte le specificazioni sopra considerate, si tratta di verificare l'ipotesi $H_0 : \rho = 1$ contro l'alternativa unidirezionale $\rho < 1$. Il test che ne consegue si chiama *test di radice unitaria*.

2.1 Test Dickey-Fuller

I test di radice unitaria noti come *test di Dickey-Fuller* si basano nel verificare l'ipotesi nulla $H_0 : \rho = 1$ di non stazionarietà sulle seguenti regressioni:

- $Y_t = \rho Y_{t-1} + e_t$
- $Y_t = \beta_0 + \rho Y_{t-1} + e_t$
- $Y_t = \beta_0 + \beta_1 t + \rho Y_{t-1} + e_t$

Le statistiche relative ai test di radice unitaria hanno generalmente distribuzioni asintotiche non normali e non standard poiché Y_t è non stazionario sotto H_0 . Tali distribuzioni sono funzioni di particolari processi stocastici a tempo continuo (moti Browniani standard o processi di Weiner) e non possiedono una conveniente forma chiusa. I loro valori critici sono quindi calcolati con tecniche simulative. Le distribuzioni delle statistiche dipendono dalla specificazione della parte deterministica, perciò c'è necessità di valori critici differenti a seconda della specificazione utilizzata. Dopo queste considerazioni generali si specificherà ora meglio il test Dickey-Fuller.

Si considera il modello AR(1)

$$Y_t = \rho Y_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$$

Le ipotesi di interesse sono

$$H_0 : \rho = 1 \quad (Y_t \sim I(1))$$

$$H_1 : \rho < 1 \quad (Y_t \sim I(0)).$$

Essendo l'ipotesi alternativa unidirezionale, in tutti i test di radice unitaria per sistemi d'ipotesi di questo tipo, l'ipotesi nulla viene rifiutata sempre per valori bassi della statistica test.

Si stima ρ con il metodo dei minimi quadrati ordinari ottenendo:

$$\hat{\rho} = \frac{\sum_{t=1}^n y_t y_{t-1}}{\sum_{t=1}^n y_{t-1}^2}$$

con

$$se(\hat{\rho}) = \sqrt{\frac{\hat{\sigma}^2}{\sum_{t=1}^n y_{t-1}^2}}$$

dove

$$\hat{\sigma}^2 = \frac{\sum_{t=1}^n (y_t - \hat{\rho} y_{t-1})^2}{n - 1}$$

Per verificare le usuali ipotesi su ρ si può utilizzare la statistica τ formalmente uguale all'usuale statistica t

$$\tau = \frac{\hat{\rho} - \rho}{se(\hat{\rho})}$$

Sotto H_1 si ha

$$\sqrt{n}(\hat{\rho} - \rho) \rightarrow N(0, 1 - \rho^2)$$

da cui per n grande,

$$\hat{\rho} \sim N\left(\rho, \frac{1}{n}(1 - \rho^2)\right)$$

quindi $\tau_{H_1} \rightarrow N(0, 1)$.

- Sotto H_0 , Y_t è non stazionario, e gli usuali momenti campionari non convergono a costanti fisse, ma convergono a funzioni di moti Browniani, per cui la statistica τ sotto H_0 converge a una distribuzione detta distribuzione di Dickey-Fuller (DF). I quantili di interesse della distribuzione DF sono calcolati generalmente tramite simulazione o approssimazione numerica e sono:

| ρ | $q(N(0, 1))$ | $q(DF)$ |
|--------|--------------|---------|
| 0.01 | -2.326 | -2.565 |
| 0.05 | -1.645 | -1.941 |
| 0.10 | -1.282 | -1.617 |

Nel verificare la presenza di radici unitarie è importante specificare correttamente le ipotesi H_0 e H_1 , che devono caratterizzare in maniera appropriata la tendenza (trend) dei dati. Se i dati osservati infatti non esibiscono particolari trend crescenti o decrescenti, allora le ipotesi da verificare dovrebbero riflettere questo andamento. Le caratteristiche di trend dei dati sotto l'ipotesi alternativa dovrebbero aiutare a determinare la specificazione della regressione da usare. È infatti importante specificare correttamente il tipo di termine deterministico impiegato nella regressione, poiché questo influenza le distribuzioni asintotiche delle statistiche usate per la verifica della presenza di radici unitarie.

Modello con drift

Viene considerata la regressione

$$Y_t = \alpha + \rho Y_{t-1} + \varepsilon_t$$

ed il sistema di ipotesi da verificare è sempre:

$$H_0 : \rho = 1$$

contro l'alternativa

$$H_1 : \rho < 1.$$

Essendo $\alpha = \mu(1 - \rho)$ con $\mu = E[Y_t]$, una corretta specificazione delle regressioni sotto entrambe le ipotesi dovrebbe portare a

$$H_0 : \rho = 1 \text{ e } \alpha = 0 \text{ (altrimenti ci sarebbe un trend crescente anche sotto } H_0)$$

$$H_1 : \rho < 1 \text{ e/o } \alpha \neq 0$$

ovvero, sotto H_0 Y_t è un random walk senza drift mentre sotto H_1 è un processo stazionario con eventuale media diversa da zero.

La statistiche test calcolata con il metodo dei minimi quadrati è ancora

$$\tau_c = \frac{\hat{\rho} - 1}{se(\hat{\rho})}$$

Sotto H_0 , la distribuzione asintotica è influenzata dalla presenza, ma non dal valore, della costante nella regressione. Anche in questo caso i valori critici della distribuzione di τ_c sono ottenuti per approssimazione numerica o simulazione e sono:

| p | $q(N(0, 1))$ | $q(DF)$ | $q(DF_c)$ |
|------|--------------|---------|-----------|
| 0.01 | -2.326 | -2.565 | -3.430 |
| 0.05 | -1.645 | -1.941 | -2.861 |
| 0.10 | -1.282 | -1.617 | -2.567 |

Si può inoltre verificare l'ipotesi nulla congiunta

$$H_0 : (\alpha, \rho) = (0, 1)$$

Con un test F del tipo:

$$F = \frac{(SSR_1 - SSR_0)/k}{SSR_0/(n - k)}$$

dove:

- SSR_1 indica la somma dei quadrati dei residui del modello vincolato,
- SSR_0 indica la somma dei quadrati dei residui ottenuti dal modello non vincolato,
- K indica in numero di vincoli.

La distribuzione asintotica di F è non standard, ma Dickey e Fuller ne hanno tabulato tramite simulazioni valori critici.

Modello con trend lineare

Si consideri la regressione

$$Y_t = \alpha + \delta t + \rho Y_{t-1} + \varepsilon_t.$$

In questo caso il sistema di ipotesi è

$$H_0 : (\delta, \rho) = (0, 1)$$

$H_1 : Y_t$ è trend-stazionario.

Sotto H_0 , la distribuzione asintotica della statistica di interesse, τ_t (strutturata come consueto), è influenzata dalla presenza ma non dal valore del termine di trend nella regressione.

Anche in questo caso i valori critici della distribuzione di τ_t sono ottenuti per approssimazione numerica o simulazione. Essi sono:

| p | q(N(0, 1)) | q(DF) | q(DF _c) | q(DF _t) |
|------|------------|--------|---------------------|---------------------|
| 0.01 | -2.326 | -2.565 | -3.430 | -3.958 |
| 0.05 | -1.645 | -1.941 | -2.861 | -3.410 |
| 0.10 | -1.282 | -1.617 | -2.567 | -3.127 |

È possibile considerare anche in questo caso un test di tipo F per la verifica di ipotesi congiunte sui parametri:

$$H_0 : (\delta, \rho) = (0, 1) \text{ } \alpha \text{ libero}$$

Oppure

$$H_0 : (\alpha, \delta, \rho) = (0, 0, 1)$$

Le statistiche a test corrispondenti hanno formulazione consueta, ma distribuzioni asintotiche e valori critici ottenuti da Dickey-Fuller con simulazioni.

Estensioni del test DF: il test ADF (Augmented DF)

Serie storiche economiche, ma non solo, hanno spesso struttura più complessa di quella dell'AR(1) su cui si basano i test DF appena visti. Esistono pertanto test di tipo DF basati su modelli AR di ordine superiore al primo, o su più generali modelli ARMA, chiamati test ADF.

Si consideri quindi il modello AR(p)

$$Y_t = D_t + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \varepsilon_t, \quad \varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$$

dove D_t è la parte deterministica ($D_t = 0, D_t = \alpha, D_t = \alpha + \delta t, \dots$).

Una formulazione equivalente è:

$$Y_t = D_t + \rho Y_{t-1} + \sum_{j=1}^{p-1} \beta_j \Delta Y_{t-j} + \varepsilon_t$$

dove

$$\rho = \phi_1 + \phi_2 + \dots + \phi_p$$

e

$$\beta_j = - \sum_{i=j+1}^p \phi_i.$$

In generale, per il modello AR(p), l'ipotesi nulla di radice unitaria è:

$$H_0 : \phi_1 + \phi_2 + \dots + \phi_p = \rho = 1$$

mentre l'ipotesi alternativa è $H_1 : \rho < 1$ ossia $\phi_1 + \phi_2 + \dots + \phi_p < 1$ che è una condizione necessaria affinché le soluzioni di $\phi(B) = 0$ siano in modulo maggiori di uno.

Sotto H_0 la statistica

$$\tau_{ADF} = \frac{\hat{\rho} - 1}{se(\hat{\rho})}$$

ha la stessa distribuzione asintotica di quella ottenuta per il caso precedente AR(1), a parità di specificazione di D_t . Lo stesso vale per i test di tipo F.

Se il modello di partenza possiede una componente MA invertibile, cioè

$$\phi(B)Y_t = \theta(B)\varepsilon_t$$

con le radici di $\theta(B) = 0$ tutte in modulo maggiori di uno allora Y_t ammette la rappresentazione AR(∞):

$$\pi(B)Y_t = \varepsilon_t, \pi(B) = \frac{\phi(B)}{\theta(B)}.$$

Data la condizione di invertibilità, il modello AR(∞) può essere approssimato da un modello AR(p), con p finito e opportunamente scelto.

Se p è troppo piccolo, la correlazione seriale residua negli errori può provocare distorsione nel test. Se invece è troppo grande la potenza del test può diminuire.

Pertanto, una procedura per determinare il valore di p più opportuno, è quella suggerita da Ng e Perron (1995):

- Si sceglie un limite superiore per p, ovvero p_{max} , e si stima la regressione con $p = p_{max}$.
- Si verifica la significatività del coefficiente relativo al ritardo $p = p_{max}$. Se il valore della statistica $|t| > 1.6$ allora si sceglie $p = p_{max}$. In caso contrario si riduce la lunghezza del ritardo massimo di 1 e si ripete la procedura.

Una regola pratica utile per determinare p_{max} è

$$p_{max} = [12 * (\frac{n}{100})^{1/4}]$$

dove $[x]$ intende la parte intera di x.

In alternativa a tale metodo è possibile scegliere p tramite l'uso di qualche criterio di identificazione automatica dell'ordine del modello AR (AIC,BIC...).

Una scelta molto rilevante è quella del modello per verificare l'ipotesi di radice unitaria, o meglio della parte deterministica del modello, perché in base a questa varia il valore della statistica test e i quantili di riferimento.

È utile a tale scopo esaminare il grafico della serie, per notare la presenza di un eventuale trend nei dati e provvedere poi ad un'appropriata specificazione di entrambe le ipotesi, in particolare stimare il modello sotto l'alternativa ed imporre eventualmente le restrizioni implicate dalla nulla.

Si comincia dal modello meno restrittivo fra i modelli plausibili (generalmente quello che include un trend e un drift)

$$\Delta Y_t = \alpha + \delta t + \gamma Y_{t-1} + \sum_{i=1}^{p-1} \beta_i \Delta Y_{t-i} + \varepsilon_t \quad (8)$$

e si usa la statistica τ per verificare l'ipotesi nulla $\gamma = 0$. I test di radice unitaria hanno una potenza (probabilità di rifiutare H_0 quando questa è falsa) bassa, ossia portano ad indicare troppo spesso che una serie abbia radice unitaria. Se l'ipotesi nulla di radice unitaria viene rigettata non c'è quindi bisogno di procedere ulteriormente e si può concludere che la serie Y_t non contiene radici unitarie.

Se H_0 non viene rigettata, è necessario determinare se i regressori deterministici presenti nel modello (8) sono tutti significativi o meno. Si tratta quindi di verificare $H_0 : \gamma = \delta = 0$. Se il trend risulta significativo bisogna verificare nuovamente la presenza di radice unitaria, $H_0 : \gamma = 0$ usando la distribuzione normale standard. Quindi se H_0 viene rigettata ci si ferma concludendo che la serie storica non contiene una radice unitaria. Altrimenti si conclude che la serie possiede una radice unitaria. Se invece il trend non risulta significativo si procede con la stima del modello

$$\Delta Y_t = \alpha + \gamma Y_{t-1} + \sum_{i=1}^{p-1} \beta_i \Delta Y_{t-i} + \varepsilon_t$$

e si verifica l'ipotesi nulla $H_0 : \gamma = 0$. Se H_0 è rifiutata si conclude che il modello non contiene una radice unitaria. Se, invece, non si può rigettare H_0 si verifica $H_0 : \alpha = 0$ condizionatamente a $\gamma = 0$ utilizzando la normale standard e se si accetta H_0 si conclude che la serie ha una radice unitaria.

Se, invece, il drift non è significativo si procede con la stima del modello senza trend e senza drift:

$$\Delta Y_t = \gamma Y_{t-1} + \sum_{i=1}^{p-1} \beta_i \Delta Y_{t-i} + \varepsilon_t$$

e si verifica l'ipotesi nulla $H_0 : \gamma = 0$ utilizzando la statistica τ . Se l'ipotesi nulla di radice unitaria è rigettata si conclude che la serie non contiene una radice unitaria. Altrimenti si conclude che possiede una radice unitaria.

2.2 Test NIV (Nonlinear Instrumental Variable) (Yooson Chang, 2002)

I test di radice unitaria di questo tipo prendono in considerazione la possibilità di correlazione tra le variabili esplicative del modello e la componente non spiegata, o disturbo. In base a tale ipotesi l'usale metodo di stima dei minimi quadrati ordinari (OLS), usato anche per i test DF, non consente di ottenere stime consistenti dei parametri (cioè asintoticamente corrette e con varianza asintoticamente nulla). Con l'utilizzo di una variabile strumentale è tuttavia ancora possibile ottenere stime consistenti con la formulazione che vedremo in (10). Questi test presentano quindi la stessa struttura dei test DF precedentemente visti, ma sono basati sulla stima non lineare del coefficiente autoregressivo tramite l'uso di una variabile strumentale. Le procedure di questo test sono estendibili in modo del tutto analogo anche a processi più complessi dell'AR(1) che qui ci si limita ad analizzare.

Si considera un processo autoregressivo a radice unitaria

$$Y_t = \alpha Y_{t-1} + u_t \quad \text{con } \alpha = 1 \quad (9)$$

con valore iniziale arbitrario $y_0 = O_p(1)$. Il termine di errore u_t è assunto avere media zero e soddisfare una delle seguenti condizioni.

Condizione 1 : $\{u_t\}$ è un processo generato dalla differenza prima di una martingala con $\sigma^2 = E(u_t^2|F_t)$ e $\sup_{t>0} E(|u_t|^p|F_t) < \infty$ per $p > 2$, con $F_t = \sigma(\{u_s\}_{s<t})$. Dove un martingala è un processo stocastico X_t indicizzato da un parametro crescente t (spesso il tempo) con la seguente proprietà: per ogni $s \leq t$ il valore atteso di X_t condizionato ai valori X_r , $r \leq s$, è uguale ad X_s (esempio il moto browniano).

Condizione 2 : $\{u_t\}$ è una successione di variabili casuali i. i. d $(0, \sigma_\varepsilon^2)$ con $E|u_t|^p < \infty$ per un certo $p > 4$, e la sua distribuzione è assolutamente continua secondo la misura di Lebesgue e ha funzione caratteristica ϕ tale che $\lim_{\lambda \rightarrow \infty} \lambda^r \phi \hat{\lambda} = 0$ per un certo $r > 0$.

Entrambe queste condizioni implicano che Y_{t-1} sia incorrelato con u_t e che la regressione (9) sia quindi correttamente specificata.

Tenuto conto di ciò si considera lo stimatore IV (Instrumental Variable) dell'autoregressione (9) tramite l'uso di una variabile strumentale data da $z_t = F(y_{t-1})$ con F funzione non lineare. Questa trasformazione F verrà detta *funzione generatrice degli strumenti* (IGF: Instrument Generating Function). Lo stimatore in questione è quindi :

$$\hat{\alpha} = \frac{\sum_{t=1}^n F(y_{t-1})y_t}{\sum_{t=1}^n F(y_{t-1})y_{t-1}} \quad (10)$$

In questa forma generale, la classe degli stimatori IV, che può essere rappresentata dalla (10), include anche gli stimatori OLS come caso particolare, quando cioè la IGF è lineare ed è la funzione identità ($F(x) = x$).

Per testare la presenza di radice unitaria nell'autoregressione (9) verificando il sistema d'ipotesi

$$H_0 : \alpha = 1$$

$$H_1 : \alpha < 1$$

è sufficiente basarsi sulla stima IV nella (10) e costruire la statistica τ , strutturata come consueto:

$$\tau_{IV} = \frac{\hat{\alpha} - 1}{s(\hat{\alpha})} \quad (11)$$

dove $s(\hat{\alpha})$ è l'errore standard dello stimatore IV $\hat{\alpha}$:

$$s(\hat{\alpha})^2 = \hat{\sigma}^2 [(\sum_{t=1}^n F(y_{t-1})y_{t-1})(\sum_{t=1}^n F(y_{t-1})^2)^{-1}(\sum_{t=1}^n y_{t-1}F(y_{t-1}))]^{-1}$$

$$s(\hat{\alpha})^2 = \hat{\sigma}^2 \frac{\sum_{t=1}^n F(y_{t-1})^2}{(\sum_{t=1}^n F(y_{t-1})y_{t-1})^2}$$

con il comune stimatore della varianza dell'errore

$$\hat{\sigma}^2 = n^{-1} \sum_{t=1}^n \hat{u}_t^2$$

dove $\hat{u}_t = y_t - \hat{\alpha}y_{t-1}$.

La distribuzione della statistica τ_{IV} è normale standard, se la funzione generatrice degli strumenti IGF è regolarmente integrabile. Non è quindi necessario calcolare dei quantili ad hoc tramite simulazioni, ma è sufficiente confrontare il valore della statistica con i quantili tabulati della normale standard.

3. Gli effetti del valore iniziale sui test di radice unitaria

L'applicazione di procedure ricorsive di *demeaning*, scarto dalla media, e *detrending*, scarto dal trend stimato, ad una serie storica è frequente.

Queste procedure possono infatti portare vantaggi utili alle analisi che si intendono effettuare sulle serie in questione. So e Shin (1999 e 2001) dimostrano infatti come l'usuale pratica di *demeaning* e *detrending* dei dati possa portare alla distorsione nelle stime dei coefficienti AR ed avere anche impatto sulla potenza dei test di radice unitaria. Essi hanno notato come la distorsione provenisse da una possibile correlazione tra y_{t+k} e e_t , per $k > 0$. Perciò hanno proposto una procedura ricorsiva, usando i dati fino al tempo $t-1$ invece che fino al tempo t , per eliminare i termini che rappresentano tale correlazione. Questo trattamento ha portato a risultati soddisfacenti permettendo infatti stime più precise dei parametri e un aumento della potenza nei test di radice unitaria, per questo tali metodi sono stati ampiamente usati.

Tuttavia, malgrado il successo che si ottiene adottando le procedure ricorsive sopraindicate, rimane l'importante questione di valutare l'impatto del valore iniziale sui test di radice unitaria quando sono state impiegate tali procedure e quando no.

Nella parte seguente di questo lavoro verrà analizzato proprio tale problema, tramite l'uso di simulazioni Monte Carlo e facendo riferimento ai test introdotti nel capitolo precedente. Si procede ora con l'introduzione delle procedure ricorsive che verranno utilizzate nell'esperimento e successivamente con l'elenco delle fasi e l'analisi dei risultati di quest'ultimo.

3.1 Procedure ricorsive di *demeaning* e *detrending*

Si consideri un generico modello AR(1) con media μ diversa da zero (la discussione è del tutto analoga anche per un modello ARMA(p , q)):

$$Y_t - \mu = \beta(Y_{t-1} - \mu) + e_t$$

con $e_t \sim WN(0, \sigma_\varepsilon^2)$ e β coefficiente AR.

Volendo ridurre l'effetto dell'ignota media μ , attenendosi come sopraccitato a quanto dimostrato da So e Shin (1999) per quanto concerne le procedure di *demeaning* più opportune, si applica :

$$Y_t^* = \beta Y_{t-1}^* + e_t \quad (12)$$

dove:

(Metodo 1)

$$Y_t^* = Y_t - \bar{y}_{t-1}$$

$$Y_{t-1}^* = Y_{t-1} - \bar{y}_{t-1},$$

dove $\bar{y}_{t-1} = \frac{1}{t-1} \sum_{k=1}^{t-1} y_k$.

I test di radice unitaria basati sulle procedure di *demeaning* ricorsivo sono invarianti rispetto a qualsiasi parametro di disturbo sotto l'ipotesi nulla.

Per quanto riguarda le procedure di *detrending*, da applicare a una serie con trend lineare, si fa riferimento a quelle che consentono di evitare che i test di radice unitaria dipendano dal parametro di disturbo (che magari riflette la grandezza del coefficiente del trend).

In tal modo, infatti, i test risultanti non sarebbero delle quantità pivotali sotto l'ipotesi nulla, pertanto ci si focalizzerà su questi metodi di *detrending* ricorsivo che permettono di evitare tale problema.

Per prima, verrà esaminata la procedura di *detrending* ricorsivo proposta da Chang e Park (2004).

(Metodo 2)

$$\tilde{y}_t = y_t - y_0 - \frac{1}{T} \sum_{t=1}^T \Delta y_t - \sum_{k=1}^{t-1} \frac{1}{k} (y_k - y_0)$$

$$\tilde{y}_{t-1} = y_{t-1} - y_0 - \sum_{k=1}^{t-1} \frac{1}{k} (y_k - y_0)$$

Per seconda, verrà presa in considerazione una procedura proposta da Chang (2002).

(Metodo 3)

$$\tilde{y}_t = y_t + \frac{2}{t-1} \sum_{k=1}^{t-1} y_k - \frac{1}{T} \sum_{t=1}^T \Delta y_t - \frac{6}{t(t-1)} \sum_{k=1}^{t-1} k y_k$$

$$\tilde{y}_{t-1} = y_{t-1} + \frac{2}{t-1} \sum_{k=1}^{t-1} y_k - \frac{6}{t(t-1)} \sum_{k=1}^{t-1} k y_k$$

Infine, verrà considerata la procedura di Taylor (2002), che può funzionare meglio nei test basati sulla stima con gli OLS, in accordo con quanto dimostrato da Rodrigues (2006).

(Metodo 4)

$$\tilde{y}_t = y_t + \frac{2}{t} \sum_{k=1}^t y_k - \frac{6}{t(t+1)} \sum_{k=1}^t k y_k$$

$$\tilde{y}_{t-1} = y_{t-1} + \frac{2}{t-1} \sum_{k=1}^{t-1} y_k - \frac{6}{t(t-1)} \sum_{k=1}^{t-1} k y_k$$

Lo scopo di queste alternative è quindi quello di liberarsi della dipendenza dal parametro di disturbo. Il Metodo 4 è ottenuto da un *detrending* ricorsivo di y_t con i dati finiti al tempo t , invece che al tempo $t-1$, senza sottrarre la media di Δy_t (differenza di y_t) come viene fatto invece nei metodi 2 e 3. Usando una di queste procedure ricorsive, si può verificare l'ipotesi nulla di radice unitaria, $H_0 : \beta = 1$, usando il modello (12).

3.2 Effetti del valore iniziale: l'esperimento Monte Carlo

Si procede ora, tramite esperimento Monte Carlo con l'analisi degli effetti del valore iniziale sui test di radice unitaria, utilizzando le procedure di *demeaning* e *detrending* descritte sopra. I comandi del software statistico R utilizzati per effettuare le simulazioni ed ottenere i risultati che seguiranno (riportati nelle varie tabelle), sono riportati nell'appendice.

Si consideri ora una serie storica

$$Y_t = z_t' \delta + v_t \quad t = 1, \dots, T, \tag{13}$$

$$v_t = \beta v_{t-1} + e_t, \tag{14}$$

dove $z_t = 1$ per un modello con una costante, oppure $z_t = (1, t)'$ per un modello con un trend con $\delta = (\delta_0, \delta_1)'$, $v_0 = \xi$, e e_t segue una distribuzione normale standard, $e_t \sim N(0, 1)$ (usando una distribuzione diversa dalla normale si ottengono risultati simili).

Per cominciare, facciamo assumere al valore iniziale ξ alcuni valori fissati : $\xi = 0, 5, 10, 20, 100$. Pertanto la prima osservazione di Y_t sarà data da

¹ $\delta = (\delta_0, \delta_1)$ verrà assunto arbitrariamente pari a $(1, 0.5)$ essendo la distribuzione asintotica della statistica test sotto H_0 influenzata solo dalla presenza ma non dal valore della costante e del trend nella regressione.

$y_1 = z_1' \delta + v_1$, con $v_1 = \beta v_0 + e_1 (= \beta \xi + e_1)$, e continuando dalla (13) e (14) si ottiene y_t , $t = 2, \dots, T$. Poiché è insolito osservare valori estremi nelle applicazioni empiriche, quali $\xi = 100$, questi risultati possono fornire utili informazioni sul comportamento asintotico.

Verranno qui considerati i risultati delle simulazioni dove sono stati usati campioni di dimensioni limitate, $T = 100$ e $T = 300$, ma i risultati utilizzando campioni più grandi sono simili.

Come già precedentemente detto verranno esaminati i test di radice unitaria basati sia sulla stima del parametro autoregressivo con gli OLS che col NIV .

Il test NIV che verrà qui usato è in particolare quello suggerito da Phillips (et al. 2004) basato sulla statistica t_{IV} (11) per β ottenuto tramite stima IV dall' equazione (12), che risulta:

$$\hat{\beta}_{IV} = \left(\sum_{t=1}^T \tilde{y}_{t-1} F(\tilde{y}_{t-1})^{-1} \right) \sum_{t=1}^T F(\tilde{y}_{t-1}) \tilde{y}_t$$

dove F è la funzione generatrice degli strumenti (IGF), grafico in Figura 1, data da:

$$F(\tilde{y}_{t-1}) = \tilde{y}_{t-1} e^{-c|\tilde{y}_{t-1}|}$$

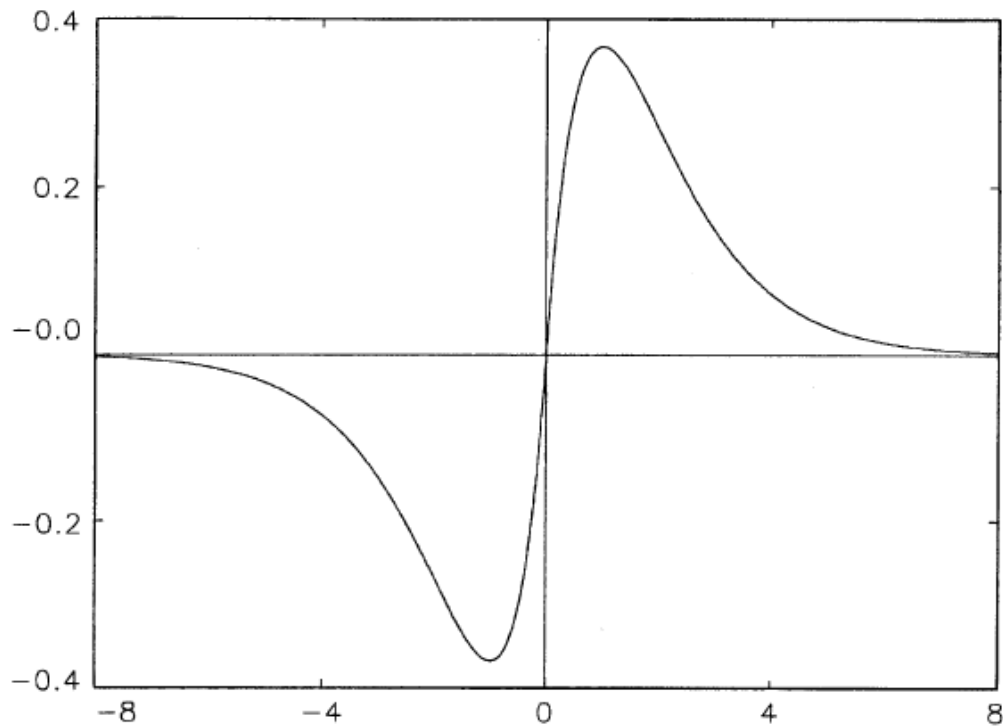


Figura 1: $F(x) = x e^{-|x|}$

Per il valore ottimale di c utilizzato ai fini calcolo della IGF dei vari test, ci si attiene a Meng, Lee, Cho (2013) che basandosi su Lee (e al. 2002) hanno riportato una griglia per il valore ottimale di c che permette di ottenere il 5% delle statistiche t -ratio vicino alla distribuzione normale standard. Tali valori sono riportati in (Tabella A.1) assieme ai valori critici.

(Tabella A.1)

| | T =100 | | | | T = 300 | | | |
|--------------|--------|--------|--------|--------|---------|--------|--------|--------|
| | c | 1% | 5% | 10% | c | 1% | 5% | 10% |
| NR(costante) | 1.66 | -2.271 | -1.639 | -1.309 | 1.03 | -2.290 | -1.649 | -1.312 |
| Metodo 1 | 0.55 | -2.360 | -1.644 | -1.266 | 0.36 | -2.326 | -1.622 | -1.246 |
| NR(trend) | 2.81 | -2.192 | -1.649 | -1.366 | 1.92 | -2.175 | -1.627 | -1.334 |
| Metodo 2 | -0.16 | -2.234 | -1.643 | -1.357 | -1.13 | -2.256 | -1.669 | -1.404 |
| Metodo 3 | 0.25 | -2.315 | -1.641 | -1.284 | 0.23 | -2.284 | -1.607 | -1.250 |
| Metodo 4 | 2.88 | -2.145 | -1.652 | -1.397 | 2.16 | -2.122 | -1.654 | -1.412 |

Si procede ora con la computazione di 10.000 simulazioni con il software statistico R, raccogliendo *size* e potenza empiriche dei due tipi di test, nel caso siano applicati alla serie originaria e nel caso vengano applicati alla stessa serie sulla quale è stata usata la procedura di *demeaning* ricorsivo (Metodo 1) .

I risultati sono riportati in (Tabella 1.1) e (Tabella 1.2).

La (Tabella 1.1) fornisce *size* e potenza empiriche dei test basati sulla stima del coefficiente autoregressivo col metodo OLS, mentre la (Tabella 1.2) quelli riguardanti i test basati sul metodo NIV.

Le *size* empiriche (con $\beta = 1$) di tutti i test non risultano particolarmente influenzate dai differenti valori iniziali per entrambi i test DF ed anche per i test dove sono state applicate le procedure del (Metodo 1). La questione rilevante è l'effetto sulla potenza del test. Quando il valore iniziale è zero, si nota infatti come i test dov'è stato usato il (Metodo 1) abbiano potenza maggiore rispetto al test DF usuale. Questi sono i risultati riportati in letteratura a sostegno dell'uso di procedure ricorsive. Emerge tuttavia chiaramente come questo test tenda invece a perdere potenza in modo significativo all'aumentare del valore iniziale.

Questo risultato è una novità per la letteratura.

Nei casi estremi, quando cioè il valore iniziale diventa molto grande $\xi \rightarrow \infty$, la potenza dei test tende a zero anche per campioni grandi come $T = 300$. Al contrario, la potenza degli usuali test DF tende ad uno sotto le stesse condizioni. All'aumentare del valore iniziale i test ai quali è stata applicata la procedura di *demeaning* diventano molto meno potenti rispetto ai convenzionali test DF.

Risultati simili vengono riscontrati anche per i test di tipo DF basati sul metodo NIV. I test NIV usando il (Metodo 1) tendono infatti anch'essi a perdere potenza all'aumentare del valore iniziale. Contrariamente ai test di tipo DF basati sul metodo OLS, i test NIV senza fare uso della procedura ricorsiva (indicati in (Tabella 1.2) con NR ovvero *not recursive*) tendono anch'essi a perdere potenza a mano a mano che il valore iniziale diventa più grande. Questo è un risultato inaspettato. Sembra che la normalità asintotica dei test NIV non richieda necessariamente l'uso di metodi ricorsivi.

Si può inoltre notare come la *size* dei test NIV per i quali è stato usato il (Metodo 1) risulti in modo anomalo essere circa zero per qualsiasi valore iniziale. Analizzando la distribuzione delle statistiche a test coinvolte, istogramma vedi Figura 2, si nota che essa noi coincide con quella della normale standard. Tale andamento fa supporre che la normalità asintotica di tale statistica non sia in questo caso verificata.

Histogram of test

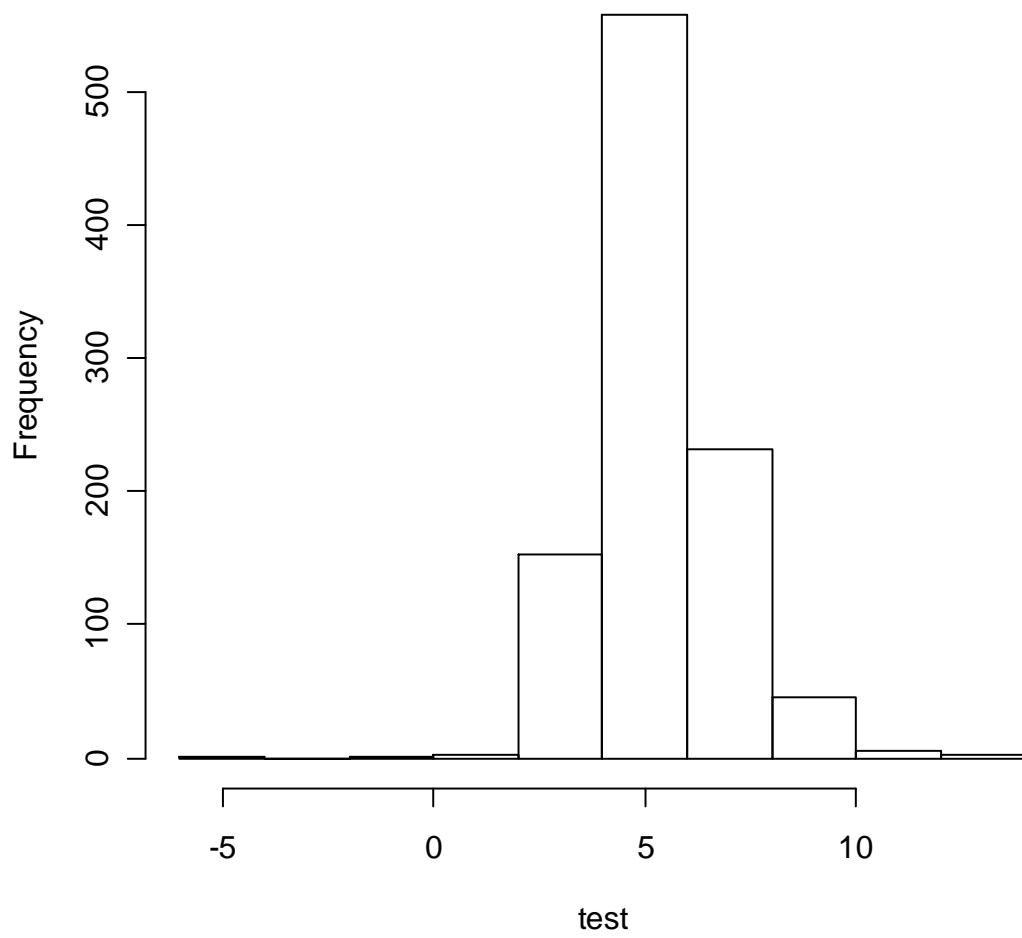


Figura 2: Istogramma delle statistiche a test nel caso di test NIV quando è stato applicato il (Metodo 1).

(Tabella 1.1): Size e potenza empiriche dei test DF al variare del valore iniziale ξ , quando alla serie è stato applicato il (Metodo 1) o meno.

| T | β | Metodo | OLS | | | | | |
|-----|---------|---------|-----------|-----------|------------|------------|-------------|-------|
| | | | $\xi = 0$ | $\xi = 5$ | $\xi = 10$ | $\xi = 20$ | $\xi = 100$ | |
| 100 | 1.0 | DF | 0.054 | 0.054 | 0.045 | 0.047 | 0.056 | |
| | | Metodo1 | 0.054 | 0.048 | 0.056 | 0.051 | 0.056 | |
| | 0.9 | DF | 0.279 | 0.189 | 0.196 | 0.263 | 1.000 | |
| | | Metodo1 | 0.820 | 0.696 | 0.313 | 0.001 | 0.000 | |
| | 0.8 | DF | 0.636 | 0.631 | 0.613 | 0.962 | 1.000 | |
| | | Metodo1 | 0.992 | 0.986 | 0.911 | 0.357 | 0.135 | |
| | 300 | 1.0 | DF | 0.052 | 0.049 | 0.049 | 0.052 | 0.049 |
| | | | Metodo1 | 0.055 | 0.054 | 0.054 | 0.051 | 0.051 |
| | | 0.9 | DF | 0.977 | 0.934 | 0.915 | 0.982 | 1.000 |
| | | | Metodo1 | 1.000 | 1.000 | 1.000 | 0.997 | 0.000 |
| 0.8 | | DF | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | |
| | | Metodo1 | 1.000 | 1.000 | 1.000 | 0.965 | 0.000 | |

(Tabella 1.2) : Size e potenza empiriche dei test NIV al variare del valore iniziale ξ , quando alla serie è stato applicato il (Metodo 1) o meno.

| T | β | Metodo | NIV | | | | |
|-----|---------|---------|-----------|-----------|------------|------------|-------------|
| | | | $\xi = 0$ | $\xi = 5$ | $\xi = 10$ | $\xi = 20$ | $\xi = 100$ |
| 100 | 1.0 | NR | 0.050 | 0.051 | 0.051 | 0.050 | 0.050 |
| | | Metodo1 | 0.005 | 0.002 | 0.005 | 0.005 | 0.002 |
| | 0.9 | NR | 0.256 | 0.271 | 0.283 | 0.259 | 0.000 |
| | | Metodo1 | 0.270 | 0.717 | 0.820 | 0.286 | 0.004 |
| | 0.8 | NR | 0.582 | 0.604 | 0.573 | 0.568 | 0.332 |
| | | Metodo1 | 0.980 | 0.984 | 0.977 | 0.568 | 0.000 |
| 300 | 1.0 | NR | 0.051 | 0.049 | 0.049 | 0.051 | 0.049 |
| | | Metodo1 | 0.003 | 0.002 | 0.003 | 0.002 | 0.000 |
| | 0.9 | NR | 0.513 | 0.490 | 0.515 | 0.489 | 0.505 |
| | | Metodo1 | 0.932 | 0.989 | 0.997 | 0.949 | 0.024 |
| | 0.8 | NR | 0.933 | 0.915 | 0.922 | 0.936 | 0.876 |
| | | Metodo1 | 1.000 | 1.000 | 1.000 | 0.999 | 0.000 |

Successivamente, si procede ad esaminare gli effetti sulla potenza dei test quando vengono usati differenti valori della varianza di ξ , ovvero con $\sigma_{\xi}^2 = 1, 5, 10, 20, 100$, e quando la media di ξ ($E(\xi) = E_{\xi}$) assume due differenti valori: $E_{\xi} = 0$ o $E_{\xi} = 10$. Si assume cioè che $v_0 = \xi \sim N(E_{\xi}, \sigma_{\xi}^2)$. Usando altre distribuzioni di ξ non si ottengono risultati molto differenti. Ne segue quindi che la prima osservazione di Y_t è dato da $y_1 = z_1\delta + v_1$ con $v_1 = \beta v_0 + e_1$.

I risultati considerando l'uso dei differenti valori della varianza di ξ sono mostrati in (Tabella 2.1) e in (Tabella 2.2). Sarebbe insolito osservare valori estremi quali $\sigma_{\xi}^2 > 5$, ma questi risultati sono utili poiché possono mostrare il comportamento asintotico. Dai dati ottenuti si nota come tutti i test effettuati con l'uso di procedure di *demeaning* ricorsivo tendano a perdere potenza quando la varianza del valore iniziale aumenta. Quindi la

perdita di potenza aumenta all'aumentare della varianza del valore iniziale.

(Tabella 2.1): Size e potenza empiriche dei test DF al variare della varianza del valore iniziale $\xi \sim N(E_\xi, \sigma_\xi^2)$, quando alla serie è stato applicato il (Metodo 1) o meno.

| T | β | Metodo | OLS | | | | | |
|---|--|---------|--------------------|--------------------|---------------------|---------------------|----------------------|-------|
| | | | $\sigma_\xi^2 = 1$ | $\sigma_\xi^2 = 5$ | $\sigma_\xi^2 = 10$ | $\sigma_\xi^2 = 20$ | $\sigma_\xi^2 = 100$ | |
| 100 ($E_\xi=0$) | 1.0 | DF | 0.050 | 0.050 | 0.050 | 0.053 | 0.050 | |
| | | Metodo1 | 0.055 | 0.055 | 0.056 | 0.050 | 0.055 | |
| | 0.9 | DF | 0.274 | 0.298 | 0.300 | 0.377 | 0.567 | |
| | | Metodo1 | 0.558 | 0.516 | 0.470 | 0.403 | 0.244 | |
| | 100 ($E_\xi=10$) | 1.0 | DF | 0.054 | 0.051 | 0.050 | 0.051 | 0.049 |
| | | | Metodo1 | 0.050 | 0.045 | 0.051 | 0.052 | 0.052 |
| 0.9 | | DF | 0.194 | 0.177 | 0.175 | 0.173 | 0.247 | |
| | | Metodo1 | 0.277 | 0.364 | 0.323 | 0.480 | 0.135 | |

(Tabella 2.2): Size e potenza empiriche dei test NIV al variare della varianza del valore iniziale $\xi \sim N(E_\xi, \sigma_\xi^2)$, quando alla serie è stato applicato il (Metodo 1) o meno.

| T | β | Metodo | NIV | | | | |
|---|--|---------|--------------------|--------------------|---------------------|---------------------|----------------------|
| | | | $\sigma_\xi^2 = 1$ | $\sigma_\xi^2 = 5$ | $\sigma_\xi^2 = 10$ | $\sigma_\xi^2 = 20$ | $\sigma_\xi^2 = 100$ |
| 100 ($E_\xi=0$) | 1.0 | NR | 0.039 | 0.025 | 0.046 | 0.043 | 0.031 |
| | | Metodo1 | 0.005 | 0.005 | 0.003 | 0.002 | 0.002 |
| | 0.9 | NR | 0.261 | 0.268 | 0.260 | 0.266 | 0.270 |
| | | Metodo1 | 0.326 | 0.271 | 0.551 | 0.416 | 0.231 |
| | | | | | | | |
| | 100 ($E_\xi=10$) | 1.0 | NR | 0.009 | 0.023 | 0.019 | 0.013 |
| Metodo1 | | | 0.001 | 0.002 | 0.001 | 0.000 | 0.000 |
| 0.9 | | NR | 0.249 | 0.262 | 0.279 | 0.309 | 0.251 |
| | | Metodo1 | 0.828 | 0.731 | 0.834 | 0.824 | 0.105 |

In (Tabella 3.1) e (Tabella 3.2) viene invece fornita la media delle stime $\hat{\beta}$ del parametro β sotto l'ipotesi nulla ($\beta = 1$) e sotto l'alternativa ($\beta = 0.9$ e $\beta = 0.2$), rispettivamente. Sotto l'ipotesi nulla, i valori stimati dei coefficienti di tutti i test non cambiano significativamente al variare del valore iniziale. Tuttavia, sotto l'alternativa, si osserva che il parametro stimato diverge dal valore del vero parametro quando viene usato il metodo ricorsivo. Dall'altro lato, la stima del parametro β si avvicina al vero valore del parametro per gli usuali test senza l'uso della procedura ricorsiva all'aumentare del valore iniziale.

(Tabella 3.1): Media delle stime del parametro autoregressivo β per i test DF, al variare del valore iniziale ξ , quando alla serie è stato applicato il (Metodo 1) o meno.

| T | B | Metodo | OLS | | | | |
|------------|-----|---------|-----------|-----------|------------|------------|-------------|
| | | | $\xi = 0$ | $\xi = 5$ | $\xi = 10$ | $\xi = 20$ | $\xi = 100$ |
| 100 | 1.0 | DF | 0.999 | 0.999 | 0.999 | 0.999 | 0.999 |
| | | Metodo1 | 0.997 | 0.997 | 0.997 | 0.997 | 0.997 |
| | 0.9 | DF | 0.879 | 0.862 | 0.855 | 0.875 | 0.900 |
| | | Metodo1 | 0.862 | 0.855 | 0.847 | 0.857 | 0.960 |
| | 0.2 | DF | 0.179 | 0.177 | 0.193 | 0.212 | 0.200 |
| | | Metodo1 | 0.182 | 0.188 | 0.193 | 0.208 | 0.754 |
| 300 | 1.0 | DF | 0.999 | 0.999 | 0.999 | 0.999 | 0.999 |
| | | Metodo1 | 0.999 | 0.999 | 0.999 | 0.999 | 0.999 |
| | 0.9 | DF | 0.889 | 0.888 | 0.888 | 0.889 | 0.900 |
| | | Metodo1 | 0.893 | 0.889 | 0.883 | 0.895 | 0.985 |
| | 0.2 | DF | 0.194 | 0.196 | 0.193 | 0.204 | 0.200 |
| | | Metodo1 | 0.196 | 0.196 | 0.196 | 0.205 | 0.432 |

(Tabella 3.2): Media delle stime del parametro autoregressivo β per i test NIV, al variare del valore iniziale ξ , quando alla serie è stato applicato il (Metodo 1) o meno.

| T | B | Metodo | NIV | | | | |
|-----|-----|---------|-----------|-----------|------------|------------|-------------|
| | | | $\xi = 0$ | $\xi = 5$ | $\xi = 10$ | $\xi = 20$ | $\xi = 100$ |
| 100 | 1.0 | NR | 1.045 | 1.066 | 1.065 | 1.047 | 1.000 |
| | | Metodo1 | 1.139 | 1.135 | 1.135 | 1.138 | 1.375 |
| | 0.9 | NR | 0.903 | 0.900 | 0.905 | 0.905 | 0.899 |
| | | Metodo1 | 0.946 | 0.896 | 0.862 | 0.947 | 1.273 |
| | 0.2 | NR | 0.202 | 0.197 | 0.200 | 0.195 | 0.204 |
| | | Metodo1 | 0.195 | 0.196 | 0.216 | 0.260 | 0.537 |
| 300 | 1.0 | NR | 1.028 | 1.048 | 1.060 | 1.052 | 1.007 |
| | | Metodo1 | 1.136 | 1.138 | 1.134 | 1.137 | 1.136 |
| | 0.9 | NR | 0.903 | 0.904 | 0.898 | 0.897 | 0.903 |
| | | Metodo1 | 0.918 | 0.897 | 0.889 | 0.918 | 1.005 |
| | 0.2 | NR | 0.200 | 0.198 | 0.199 | 0.199 | 0.198 |
| | | Metodo1 | 0.198 | 0.200 | 0.207 | 0.226 | 0.344 |

Si procede ora prendendo in esame gli effetti del valore iniziale sulla potenza test di radice unitaria quando vengono usate le procedure di *detrending* ricorsivo (Metodo 2), (Metodo 3) e (Metodo 4). I modelli con trend lineare sono utilizzati spesso negli studi empirici, per questo è molto utile analizzare gli effetti del valore iniziale anche in tali situazioni. I risultati della simulazione con R sono riportati in (Tabella 4.1) e (Tabella 4.2).

Nel complesso, questi risultati mostrano che la *size* e la potenza dei test di radice unitaria utilizzando qualsiasi procedura di *detrending*, tra le proposte, mostrano un andamento simile a quelli ottenuti utilizzando la procedura di *demeaning* ricorsivo. Infatti anch'essi nella maggior parte dei casi perdono potenza all'aumentare del valore iniziale.

(Tabella 4.1): Size e potenza empiriche dei test DF al variare del valore iniziale ξ , quando alla serie sono stati applicati (Metodo 2), (Metodo 3), (Metodo 4) o meno.

| T | B | Metodo | OLS | | | | |
|-----|-----|---------|-----------|-----------|------------|------------|-------------|
| | | | $\xi = 0$ | $\xi = 5$ | $\xi = 10$ | $\xi = 20$ | $\xi = 100$ |
| 100 | 1.0 | DF | 0.054 | 0.054 | 0.054 | 0.054 | 0.054 |
| | | Metodo2 | 0.075 | 0.079 | 0.077 | 0.077 | 0.067 |
| | | Metodo3 | 0.078 | 0.073 | 0.070 | 0.072 | 0.075 |
| | | Metodo4 | 0.057 | 0.057 | 0.057 | 0.057 | 0.057 |
| | 0.9 | DF | 0.183 | 0.175 | 0.256 | 0.627 | 1.000 |
| | | Metodo2 | 0.309 | 0.197 | 0.044 | 0.000 | 0.000 |
| | | Metodo3 | 0.295 | 0.209 | 0.043 | 0.000 | 0.000 |
| | | Metodo4 | 0.273 | 0.176 | 0.049 | 0.000 | 0.000 |
| | 0.8 | DF | 0.560 | 0.642 | 0.808 | 0.997 | 1.000 |
| | | Metodo2 | 0.767 | 0.422 | 0.040 | 0.000 | 0.000 |
| | | Metodo3 | 0.757 | 0.420 | 0.042 | 0.000 | 0.000 |
| | | Metodo4 | 0.791 | 0.663 | 0.276 | 0.000 | 0.000 |
| 300 | 1.0 | DF | 0.044 | 0.054 | 0.050 | 0.065 | 0.056 |
| | | Metodo2 | 0.081 | 0.073 | 0.080 | 0.081 | 0.069 |
| | | Metodo3 | 0.086 | 0.080 | 0.076 | 0.056 | 0.056 |
| | | Metodo4 | 0.056 | 0.057 | 0.056 | 0.050 | 0.049 |
| | 0.9 | DF | 0.912 | 0.930 | 0.968 | 1.000 | 1.000 |
| | | Metodo2 | 0.732 | 0.073 | 0.005 | 0.000 | 0.000 |
| | | Metodo3 | 0.947 | 0.945 | 0.902 | 0.401 | 0.000 |
| | | Metodo4 | 0.974 | 0.965 | 0.900 | 0.357 | 0.000 |
| | 0.8 | DF | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | Metodo2 | 0.896 | 0.352 | 0.002 | 0.000 | 0.000 |
| | | Metodo3 | 1.000 | 1.000 | 1.000 | 1.000 | 0.000 |
| | | Metodo4 | 1.000 | 1.000 | 1.000 | 1.000 | 0.000 |

(Tabella 4.2): Size e potenza empiriche dei test NIV al variare del valore iniziale ξ , quando alla serie sono stati applicati (Metodo 2), (Metodo 3), (Metodo 4) o meno.

| T | β | Metodo | NIV | | | | |
|-----|---------|---------|-----------|-----------|------------|------------|-------------|
| | | | $\xi = 0$ | $\xi = 5$ | $\xi = 10$ | $\xi = 20$ | $\xi = 100$ |
| 100 | 1.0 | NR | 0.055 | 0.055 | 0.055 | 0.055 | 0.055 |
| | | Metodo2 | 0.925 | 0.931 | 0.927 | 0.925 | 0.931 |
| | | Metodo3 | 0.525 | 0.628 | 0.627 | 0.160 | 0.000 |
| | | Metodo4 | 0.155 | 0.160 | 0.154 | 0.165 | 0.159 |
| | 0.9 | NR | 0.150 | 0.132 | 0.079 | 0.010 | 0.000 |
| | | Metodo2 | 0.878 | 0.994 | 1.000 | 0.877 | 0.000 |
| | | Metodo3 | 0.171 | 0.755 | 0.915 | 0.015 | 0.000 |
| | | Metodo4 | 0.159 | 0.244 | 0.272 | 0.159 | 0.000 |
| | 0.8 | NR | 0.315 | 0.300 | 0.242 | 0.105 | 0.000 |
| | | Metodo2 | 1.000 | 1.000 | 1.000 | 0.348 | 0.000 |
| | | Metodo3 | 0.698 | 0.990 | 0.779 | 0.001 | 0.000 |
| | | Metodo4 | 0.367 | 0.428 | 0.366 | 0.087 | 0.000 |
| 300 | 1.0 | NR | 0.046 | 0.046 | 0.046 | 0.046 | 0.046 |
| | | Metodo2 | 0.998 | 0.998 | 0.997 | 0.996 | 1.000 |
| | | Metodo3 | 0.376 | 0.412 | 0.447 | 0.189 | 0.000 |
| | | Metodo4 | 0.173 | 0.157 | 0.155 | 0.157 | 0.152 |
| | 0.9 | NR | 0.400 | 0.398 | 0.358 | 0.300 | 0.000 |
| | | Metodo2 | 1.000 | 1.000 | 1.000 | 1.000 | 0.047 |
| | | Metodo3 | 0.288 | 0.847 | 0.988 | 0.084 | 0.000 |
| | | Metodo4 | 0.304 | 0.355 | 0.396 | 0.295 | 0.000 |
| | 0.8 | NR | 0.854 | 0.853 | 0.816 | 0.785 | 0.087 |
| | | Metodo2 | 1.000 | 1.000 | 1.000 | 1.000 | 0.020 |
| | | Metodo3 | 0.882 | 0.999 | 0.890 | 0.009 | 0.000 |
| | | Metodo4 | 0.634 | 0.684 | 0.626 | 0.393 | 0.001 |

La questione è del tutto analoga anche nel caso in cui vengano analizzate size e potenza empiriche al variare della varianza del valore iniziale $\xi \sim N(E_\xi, \sigma_\xi^2)$. I risultati sono riportati in (Tabella 5.1) e (Tabella 5.2).

(Tabella 5.1): Size e potenza empiriche dei test DF al variare della varianza del valore iniziale $\xi \sim N(E_\xi, \sigma_\xi^2)$, quando alla serie sono stati applicati (Metodo 2), (Metodo 3), (Metodo 4) o meno.

| T | β | Metodo | OLS | | | | |
|--|---------|---------|--------------------|--------------------|---------------------|---------------------|----------------------|
| | | | $\sigma_\xi^2 = 0$ | $\sigma_\xi^2 = 5$ | $\sigma_\xi^2 = 10$ | $\sigma_\xi^2 = 20$ | $\sigma_\xi^2 = 100$ |
| 100 ($E_\xi=0$) | 1.0 | DF | 0.050 | 0.053 | 0.044 | 0.058 | 0.062 |
| | | Metodo2 | 0.074 | 0.086 | 0.066 | 0.082 | 0.077 |
| | | Metodo3 | 0.053 | 0.048 | 0.043 | 0.053 | 0.063 |
| | | Metodo4 | 0.941 | 0.948 | 0.929 | 0.947 | 0.939 |
| | 0.9 | DF | 0.152 | 0.189 | 0.167 | 0.179 | 0.201 |
| | | Metodo2 | 0.171 | 0.204 | 0.166 | 0.161 | 0.212 |
| | | Metodo3 | 0.243 | 0.305 | 0.314 | 0.246 | 0.312 |
| | | Metodo4 | 0.035 | 0.042 | 0.030 | 0.037 | 0.037 |
| 300 ($E_\xi=10$) | 1.0 | DF | 0.045 | 0.056 | 0.049 | 0.046 | 0.048 |
| | | Metodo2 | 0.078 | 0.077 | 0.077 | 0.073 | 0.073 |
| | | Metodo3 | 0.055 | 0.049 | 0.043 | 0.045 | 0.061 |
| | | Metodo4 | 0.907 | 0.901 | 0.928 | 0.914 | 0.898 |
| | 0.9 | DF | 0.274 | 0.228 | 0.266 | 0.207 | 0.171 |
| | | Metodo2 | 0.010 | 0.013 | 0.011 | 0.007 | 0.210 |
| | | Metodo3 | 0.052 | 0.112 | 0.069 | 0.219 | 0.015 |
| | | Metodo4 | 0.025 | 0.033 | 0.031 | 0.033 | 0.045 |

(Tabella 5.2): Size e potenza empiriche dei test DF al variare della varianza del valore iniziale $\xi \sim N(E_\xi, \sigma_\xi^2)$, quando alla serie sono stati applicati (Metodo 2), (Metodo 3), (Metodo 4) o meno.

| T | β | Metodo | NIV | | | | |
|--|---------|---------|--------------------|--------------------|---------------------|---------------------|----------------------|
| | | | $\sigma_\xi^2 = 0$ | $\sigma_\xi^2 = 5$ | $\sigma_\xi^2 = 10$ | $\sigma_\xi^2 = 20$ | $\sigma_\xi^2 = 100$ |
| 100 ($E_\xi=0$) | 1.0 | NR | 0.023 | 0.013 | 0.018 | 0.020 | 0.015 |
| | | Metodo2 | 0.942 | 0.937 | 0.943 | 0.938 | 0.938 |
| | | Metodo3 | 0.450 | 0.156 | 0.407 | 0.461 | 0.000 |
| | | Metodo4 | 0.015 | 0.001 | 0.002 | 0.000 | 0.000 |
| | 0.9 | NR | 0.143 | 0.140 | 0.137 | 0.128 | 0.100 |
| | | Metodo2 | 0.008 | 0.002 | 0.023 | 0.011 | 0.040 |
| | | Metodo3 | 0.000 | 0.000 | 0.003 | 0.001 | 0.000 |
| | | Metodo4 | 0.002 | 0.001 | 0.001 | 0.000 | 0.000 |
| 300 ($E_\xi=10$) | 1.0 | NR | 0.004 | 0.009 | 0.008 | 0.003 | 0.008 |
| | | Metodo2 | 0.999 | 0.999 | 0.998 | 0.999 | 0.999 |
| | | Metodo3 | 0.311 | 0.315 | 0.168 | 0.193 | 0.032 |
| | | Metodo4 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| | 0.9 | NR | 0.070 | 0.073 | 0.072 | 0.072 | 0.069 |
| | | Metodo2 | 0.006 | 0.002 | 0.000 | 0.000 | 0.000 |
| | | Metodo3 | 0.005 | 0.000 | 0.002 | 0.000 | 0.000 |
| | | Metodo4 | 0.001 | 0.001 | 0.001 | 0.000 | 0.000 |

Sono riscontrabili anche in questo caso nei test di tipo NIV valori anomali della size, quando viene fatto uso di procedure ricorsive. A tal proposito viene riportato, in Figura 3, l'istogramma rappresentante la distribuzione delle statistiche test quando viene applicato il (Metodo 2) che mostra un andamento che si scosta da quello della normale standard con la quale invece dovrebbe essere confrontabile.

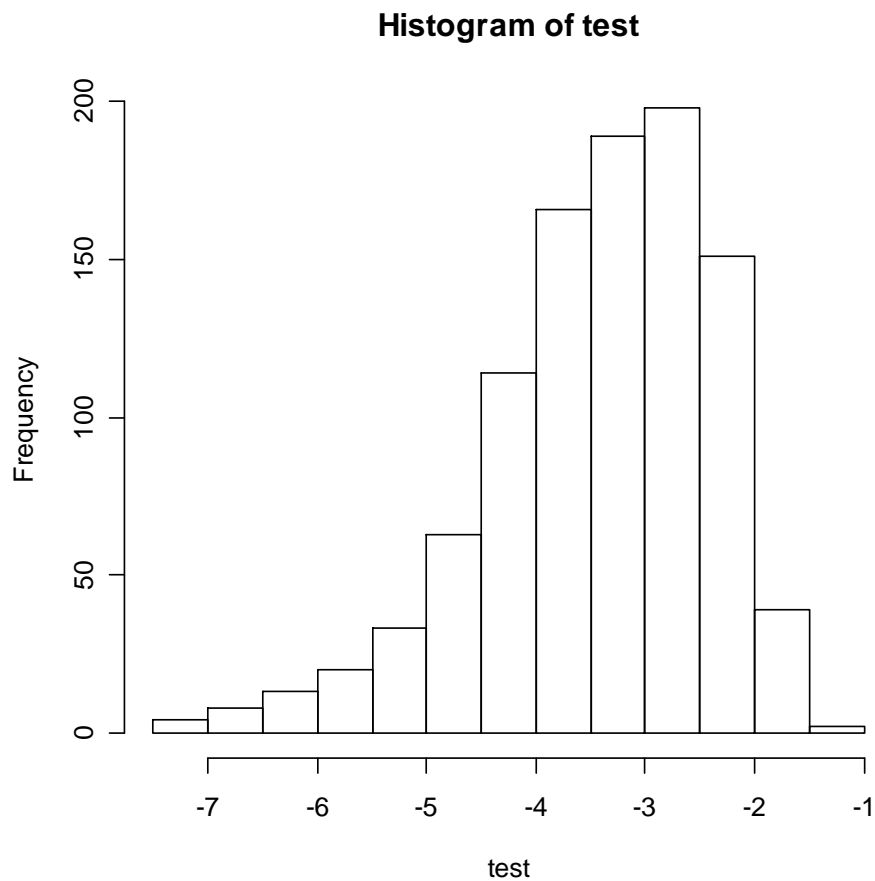


Figura 3: Istogramma delle statistiche a test nel caso di test NIV quando è stato applicato il (Metodo 2).

I risultati delle simulazioni effettuate per ottenere le stime del parametro β , riportate in (Tabella 6.1) e (Tabella 6.2) sono simili a quelle ottenute per i casi in cui si è fatto uso del (Metodo 1).

(Tabella 6.1): Media delle stime del parametro autoregressivo β per i test DF, al variare del valore iniziale ξ , quando alla serie sono stati applicati (Metodo 2), (Metodo3), (Metodo 4) o meno.

| T | β | Metodo | OLS | | | | |
|-----|---------|---------|-----------|-----------|------------|------------|-------------|
| | | | $\xi = 0$ | $\xi = 5$ | $\xi = 10$ | $\xi = 20$ | $\xi = 100$ |
| 100 | 1.0 | DF | 0.999 | 0.999 | 0.999 | 0.999 | 0.999 |
| | | Metodo2 | 0.920 | 0.914 | 0.917 | 0.921 | 0.919 |
| | | Metodo3 | 0.919 | 0.918 | 0.917 | 0.917 | 0.915 |
| | | Metodo4 | 1.031 | 1.030 | 1.029 | 1.028 | 1.024 |
| | 0.9 | DF | 0.997 | 1.001 | 1.005 | 1.011 | 0.911 |
| | | Metodo2 | 0.853 | 0.877 | 0.917 | 0.964 | 0.993 |
| | | Metodo3 | 0.856 | 0.877 | 0.918 | 0.964 | 0.994 |
| | | Metodo4 | 1.031 | 1.031 | 1.030 | 1.029 | 1.026 |
| | 0.2 | DF | 0.194 | 0.194 | 0.195 | 0.198 | 0.199 |
| | | Metodo2 | 0.275 | 0.298 | 0.383 | 0.602 | 0.949 |
| | | Metodo3 | 0.268 | 0.309 | 0.384 | 0.599 | 0.949 |
| | | Metodo4 | 1.030 | 1.031 | 1.031 | 1.031 | 1.031 |
| 300 | 1.0 | DF | 0.999 | 0.999 | 0.999 | 0.999 | 0.999 |
| | | Metodo2 | 0.973 | 0.972 | 0.971 | 0.972 | 0.972 |
| | | Metodo3 | 0.972 | 0.972 | 0.973 | 0.972 | 0.972 |
| | | Metodo4 | 1.010 | 1.010 | 1.010 | 1.010 | 1.009 |
| | 0.9 | DF | 0.999 | 1.000 | 1.000 | 1.001 | 1.004 |
| | | Metodo2 | 0.891 | 0.909 | 0.943 | 0.978 | 0.997 |
| | | Metodo3 | 0.892 | 0.911 | 0.944 | 0.978 | 0.997 |
| | | Metodo4 | 1.010 | 1.010 | 1.010 | 1.010 | 1.010 |
| | 0.2 | DF | 0.194 | 0.194 | 0.195 | 0.198 | 0.199 |
| | | Metodo2 | 0.289 | 0.316 | 0.401 | 0.616 | 0.967 |
| | | Metodo3 | 0.287 | 0.287 | 0.403 | 0.613 | 0.968 |
| | | Metodo4 | 1.010 | 1.030 | 1.030 | 1.030 | 1.030 |

(Tabella 6.2): Media delle stime del parametro autoregressivo β per i test DF, al variare del valore iniziale ξ , quando alla serie sono stati applicati (Metodo 2), (Metodo3), (Metodo 4) o meno.

| T | β | Metodo | NIV | | | | | |
|-----|---------|---------|-----------|-----------|------------|------------|-------------|-------|
| | | | $\xi = 0$ | $\xi = 5$ | $\xi = 10$ | $\xi = 20$ | $\xi = 100$ | |
| 100 | 1.0 | NR | 2.175 | 1.329 | 1.170 | 1.085 | 1.016 | |
| | | Metodo2 | 0.946 | 0.944 | 0.946 | 0.944 | 0.943 | |
| | | Metodo3 | 0.949 | 0.937 | 0.929 | 0.941 | 0.951 | |
| | | Metodo4 | 0.888 | 0.881 | 0.883 | 0.889 | 0.883 | |
| | 0.9 | NR | 2.199 | 1.268 | 1.122 | 1.053 | 1.016 | |
| | | Metodo2 | 0.994 | 0.995 | 0.996 | 0.996 | 0.998 | |
| | | Metodo3 | 0.995 | 0.986 | 0.961 | 0.969 | 0.935 | |
| | | Metodo4 | 1.177 | 1.203 | 1.232 | 1.260 | 1.851 | |
| | 0.2 | NR | 1.976 | 1.958 | 1.887 | 1.787 | 1.647 | |
| | | Metodo2 | 0.579 | 0.907 | 0.971 | 0.989 | 0.997 | |
| | | Metodo3 | 0.449 | 0.472 | 0.591 | 0.814 | 0.967 | |
| | | Metodo4 | 0.203 | 0.195 | 0.232 | 0.293 | 0.739 | |
| | 300 | 1.0 | NR | 2.615 | 1.339 | 1.167 | 1.082 | 1.016 |
| | | | Metodo2 | 0.979 | 0.980 | 0.979 | 0.979 | 0.979 |
| | | | Metodo3 | 0.983 | 0.979 | 0.976 | 0.981 | 0.984 |
| | | | Metodo4 | 0.937 | 0.939 | 0.937 | 0.934 | 0.929 |
| 0.9 | | NR | 2.282 | 1.259 | 1.121 | 1.053 | 1.015 | |
| | | Metodo2 | 0.996 | 0.997 | 0.997 | 0.997 | 0.999 | |
| | | Metodo3 | 0.998 | 0.994 | 0.984 | 0.987 | 0.972 | |
| | | Metodo4 | 1.077 | 1.096 | 1.121 | 1.191 | 1.825 | |
| 0.2 | | NR | 1.714 | 2.039 | 1.811 | 1.859 | 1.684 | |
| | | Metodo2 | 0.589 | 0.907 | 0.972 | 0.989 | 0.998 | |
| | | Metodo3 | 0.452 | 0.489 | 0.623 | 0.853 | 0.988 | |
| | | Metodo4 | 0.202 | 0.195 | 0.208 | 0.236 | 0.389 | |

Riassumendo, si è scoperto come tutti i test, sia i test basati sugli OLS che quelli basati sul NIV, se sono state usate procedure ricorsive, tendono a perdere potenza all'aumentare del valore iniziale.

Questi risultati non devono tuttavia necessariamente implicare un divieto dell'uso dei test effettuati con questi metodi ricorsivi. Tali test infatti hanno potenza maggiore rispetto ai test usuali sotto certe condizioni, cioè quando il valore iniziale è piccolo. Pertanto si può osservare una situazione di *trade-off* dove il miglioramento in termini di potenza del test usando metodi ricorsivi si può ottenere a costo di introdurre altre questioni.

Conclusione

In questo lavoro si sono esaminati, dopo un'introduzione sulle serie storiche stazionarie e non e sui test di radice unitaria di interesse, i test di radice unitaria quando sono state applicate procedure di *demeaning* e *detrending* ricorsivo, cosa molto popolare in letteratura.

Nonostante si riscontri che essi conducono a stime più precise del coefficiente autoregressivo e ad un miglioramento della potenza del corrispondente test di radice unitaria quando il valore iniziale della serie storica è piccolo, queste caratteristiche desiderabili scompaiono all'aumentare del valore iniziale. Emerge così, come le procedure ricorsive siano utili e chiaramente migliori in alcune situazioni, mentre non siano preferibili in altre.

Sarebbe quindi opportuno considerare una strategia che sfrutti le proprietà desiderabili di entrambi i test, uno con l'uso di procedure ricorsive e uno che non utilizza tali procedure.

Appendice: Codice del software R utilizzato

Si riportano di seguito i principali comandi R utilizzati per le simulazioni dell'esperimento Monte Carlo:

```
T=100, T=300
#valore costante (drift)
const=1
#valore coefficiente del trend:
coeftrend= 0.5
#trend:
trend=c(1:T)
#valore iniziale
vi = 0, 5, 10, 20, 100
oppure
rnorm(1, mean= $\mu$ , sqrt( $\sigma$ ))  $\mu = 0, 10$   $\sigma = 1, 5, 10, 20, 100$ .
y_0 = vi
phi=0.9, 0.8, 0.2
cont=0
s=1000
pv=0.05
for (i in 1:s){

#modello con costante sotto l'ipotesi nulla:
x=rnorm(T)
v_t = diffinv(x, xi=y_0)
y_t=ts(v_t + trend*const)[1:T]

#modello con costante e trend sotto l'ipotesi nulla:
x=rnorm(T)
v_t = diffinv(x, xi=y_0)
y_t=ts(v_t + trend*const + coeftrend*trend)[1:T]

#modello con costante sotto l'alternativa:
x=rnorm(T)
```

```

v_t = NULL
v_t[1] = phi*vi + x[1]
for (j in 2:T){
v_t[j] = phi*v_t[j-1] + x[j]}
y_t = ts(const+ v_t)

# modello con costante e trend sotto l'alternativa :
x=rnorm(T)
v_t = NULL
v_t[1] = phi*vi + x[1]
for (j in 2:T){
v_t[j] = phi*v_t[j-1] + x[j]}
y_t=ts(const + coeftrend*trend + v_t)

#procedura di demeaning:
y = embed(y_t,2)
m.y_t=NULL
for (j in 1:T-1){
m.y_t[j] = mean(y[,2][1:j])
}
y_tstar=(y[,1]-m.y_t)

#procedura di detrending metodo 2:
y_diff_media = mean(diff(y_t))
yd[1]=y_t[1]-y_0 - y_diff_media
for (t in 2:T){
yd[t] = yd[t-1] + (y_t[t] - y_t[t-1]) - y_diff_media
yd[t] = yd[t-1] +y_t[t] - y_t[t-1] - y_diff_media - (y_t[t-1] - y_0)/(t-1)
}
y_tstar=ts(yd)

#procedura di detrending metodo 3:
y_diff_media = mean(diff(y_t))
yd[1]=y_t[1] - y_diff_media

```

```

yd[2]= y_t[2] - y_t[1] - y_diff_media
for (t in 3:T){
yd[t] = yd[t-1] + (y_t[t] - y_t[t-1]) - y_diff_media
}
y_tstar=ts(yd)

```

#procedura di *detrending* metodo 4:

```

yd[1]= 0 #y1d = y1 + 2y1 -3y1=0
for(k in 1:T){
b[k]=(sum((1:k)*y_t[1:k]))
}
for(j in 1:T){
m[j]= mean(y_t[1:j])
}
for(t in 1:T){
yd[t]= y_t[t] +2*m[t] - (6/(t*(t+1)))*b[t]
}
y_tstar=ts(yd)

```

#test Dickey-Fuller per la serie originaria con trend:

```

test_df1=adfTest(y_tstar, lags=1,type="ct")
p_value<-test_df1@test$p.value
if (p_value < pv){cont=cont+1}
}
power<-cont/s #size se sotto l'ipotesi nulla
power

```

#test Dickey-Fuller per la serie a cui sono state applicate procedure ricorsive:

```

test_df1=adfTest(y_tstar, lags=1,type="nc")
p_value<-test_df1@test$p.value
if (p_value < pv){cont=cont+1}
}
power<-cont/s #size se sotto l'ipotesi nulla

```

power

#Test NIV per la serie alla quale sono state applicate le procedure
ricorsive:

b = embed(y_tstar, 2)

Fy_t_1star= b[,2]*exp(-c*abs(b[,2]))

#stima di beta:

beta= (sum(Fy_t_1star*b[,1])/sum(b[,2]*Fy_t_1star))

#calcolo dello se:

u_t= b[,1] -(beta*b[,2])

sigma2= (sum(u_t^2))/T

se=sqrt(sigma2*((sum(Fy_t_1star^2))/(sum((Fy_t_1star*b[,2])^2)))/T)

#calcolo della t_ratio:

t_ratio=(beta-1)/se

#controllo accetto-rifiuto

if (t_ratio < cv){

cont=cont+1

}

size<-cont/s #power nel caso dell'ipotesi alternativa

size

#Test NIV per la serie originaria:

#stima di beta:

y=embed(y_t,2)

y_t_1=y[,2]

Fy_t_1= y_t_1*exp(-c*abs(y_t_1))

beta<-(sum(Fy_t_1*y_t[1:T+1]))/(sum(y_t_1*Fy_t_1))

#calcolo dello se:

u_t= y_t[1:T+1] -(beta*y_t_1)

sigma2= (sum(u_t^2))/T

se=sqrt(sigma2*((sum(Fy_t_1^2))/(sum((Fy_t_1*y_t_1)^2)))/T)

#calcolo della t_ratio:

t_ratio=(beta-1)/se

#controllo accetto-rifiuto

```
if (t_ratio < cv){  
cont=cont+1  
}  
size<-cont/s #power nel caso dell'ipotesi alternativa  
size
```


Bibliografia e sitografia

Bibliografia:

- BERNHARD PFAFF, (2008), *Analysis of Integrated and Cointegrated Time Series with R*, Springer, New York.
- CHANG, Y., (2002), *Nonlinear IV unit root tests in panels with cross-sectional dependency*, Journal of Econometrics 110, 261-292.
- DI FONZO T., LISI F., (2005), *Serie Storiche Economiche*, Carrocci, Roma.
- MING MENG, HYEJIN LEE, MYEONG HYEON CHO, JUNSOOLEE, (2013), *Impacts of the initial observation on unit root test using recursive demeaning and detrending procedures*, Economics Letters 120, 195-199.
- PAULO M. M. RODRIGUES (2004), *Properties of Recursive Trend-Adjusted Unit Root Tests*, European University Institute Working Paper.
- PHILLIPS, P.C.B., PARK, J.Y., CHANG, Y., (2004), *Nonlinear instrumental variable estimation of an autoregression*, Journal of Econometrics 118, 219-246.

Sitografia:

- ERIC ZIVOT, *Unit root tests*,
<http://faculty.washington.edu/ezivot/econ584/notes/unitroot.pdf>.