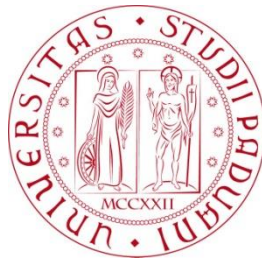


Università degli Studi di Padova
Dipartimento di Scienze Statistiche
Corso di Laurea Triennale in
Statistica e Gestione delle Imprese



RELAZIONE FINALE
**VERIFICA DELL'IPOTESI DI CAMBIAMENTI NEL
PARAMETRO DI INTEGRAZIONE FRAZIONARIA: UNA
VALUTAZIONE MEDIANTE IL METODO MONTE CARLO**

Relatore Prof. Luisa Bisaglia
Dipartimento di Scienze Statistiche

Laureanda: Andrée Giuseppina Tono
Matricola N. 1032710

Anno Accademico 2013/2014

Indice

Introduzione	1
1 I processi a memoria lunga	5
1.1 Origini	5
1.2 Caratteristiche e definizioni	6
1.3 I modelli ARFIMA	8
1.3.1 Dai modelli ARMA ai modelli ARFIMA	8
1.3.2 Proprietà	10
2 Test per identificare cambiamenti di persistenza nei modelli	
<i>ARFIMA</i>	13
2.1 Introduzione	13
2.2 Cambiamenti nel parametro di persistenza	15
2.3 Test per i cambiamenti di persistenza	16
2.3.1 Distribuzioni limite sotto l'ipotesi nulla	17
2.4 Lo stimatore di Whittle	20
2.5 Processo di Wiener	21
3 Esperimento Monte Carlo	23
3.1 Valori critici	24
3.2 <i>Size</i> e potenza dei test	28
3.2.1 <i>Size</i>	29
3.2.2 Potenza	29
3.3 Considerazioni	30

Conclusioni	35
A Codice sorgente R	37
Bibliografia	49
Elenco delle tabelle	51

Introduzione

Le serie storiche di fenomeni economici raramente possono considerarsi generate da processi stocastici stazionari, ovvero processi che non presentano cambiamenti sistematici nella media e/o nella varianza. Di solito esse presentano una tendenza di fondo, per cui non è accettabile quanto meno l'ipotesi di media costante. Inoltre le variabili economiche osservate a cadenza mensile e triennale presentano spesso andamenti caratterizzati da periodicità riconducibile a fattori stagionali (ad esempio nel periodo estivo possono verificarsi dei picchi nei consumi dell'elettricità, legati all'utilizzo per periodi prolungati del condizionatore).

Esistono due condizioni di stazionarietà: quella *in senso stretto* che risulta piuttosto forte in quanto impone dei vincoli sull'intera distribuzione del processo e quindi sui momenti di qualsiasi ordine, e una meno restrittiva cosiddetta *debole* così chiamata perché impone condizioni soltanto sui primi due momenti. La nozione di stazionarietà *debole* è quella più utilizzata, in parte per comodità, e in parte perché la stazionarietà del secondo ordine è già soddisfacente. Una serie si dice stazionaria *in senso debole* se la sua media e la sua varianza sono finite e costanti nel tempo, e se le covarianze tra gli elementi del processo sono finite e dipendono solo dalla distanza k tra gli indici temporali. La stazionarietà in media garantisce che il livello della serie sempre lo stesso, mentre quella in varianza garantisce che le oscillazioni abbiano sempre la stessa struttura di fondo; queste due condizioni rendono una serie più facile da analizzare e garantiscono previsioni più robuste.

Per non perdere quindi i vantaggi assicurati dalla stazionarietà si può rimuovere la non stazionarietà ricorrendo a svariati metodi, uno dei principali è la differenziazione. Una serie che deve essere differenziata d volte

per diventare stazionaria si dice integrata di ordine d , e si indica con $I(d)$. Spesso le procedure utilizzate per identificare il numero di volte in cui una serie deve essere differenziata per raggiungere la stazionarietà sono intuitive ma poco formali, per esempio l'analisi della funzione di autocorrelazione empirica della serie differenziata. Tuttavia, per alcune serie, ricorrere alla differenza prima o seconda sembra eccessivo. Per cercare di ovviare a questo problema [Granger and Joyeux, 1980] e [Hosking, 1981] hanno proposto i modelli ARIMA frazionari, anche detti ARFIMA, per i quali il parametro di differenziazione può essere anche reale e quindi non necessariamente intero. Tali modelli hanno ricevuto una crescente attenzione per la loro flessibilità di applicazione e hanno trovato applicazione in svariati ambiti disciplinari. Vengono utilizzati per descrivere la dinamica di fenomeni che presentano una struttura di dipendenza che persiste nel tempo.

Un argomento di particolare rilievo nell'ambito dei modelli a memoria lunga riguarda la verifica della costanza del parametro di memoria lunga. Infatti, quando si analizzano modelli relativi ad un gran numero di dati, con frequenza elevata o raccolti in un lungo arco temporale, la struttura di dipendenza a memoria lunga può variare nel tempo. Tale comportamento può essere legato sia a mutamenti nel processo generatore dei dati, sia a mutamenti nella procedura di raccolta delle osservazioni. Verificare quindi l'eventuale presenza di cambiamenti del parametro di differenziazione frazionaria, d , risulta molto importante, poiché le variazioni hanno conseguenze rilevanti sulla struttura del processo in quanto potrebbero modificarne le proprietà.

In questo studio viene proposto un nuovo metodo per scoprire cambiamenti di persistenza nei modelli frazionalmente integrati, basato su una stima regressiva e progressiva della regressione. Questa procedura, che segue quanto proposto nell'articolo di Martins and Rodrigues [2014], generalizza gli approcci per i convenzionali processi integrati al contesto dell'integrazione frazionaria. In particolare viene descritta il metodo utilizzato per implementare i test volti a verificare i cambiamenti di persistenza. Vengono derivati risultati asintotici e la performance dei test è valutata tramite un esperimento di Monte Carlo. Tutti i risultati proposti sono stati ottenuti mediante l'utilizzo del software statistico R, scaricabile all'indirizzo www.r-project.org. In appen-

dice A si può trovare il codice sorgente di R utilizzato per l'implementazione dei test.

Questo lavoro è quindi strutturato come segue. Nel primo capitolo sono descritti i processi a memoria lunga ed introdotti i processi ARFIMA(p,d,q). Nel secondo capitolo viene delineata la struttura del test utilizzato per indagare i cambiamenti di persistenza. Nel terzo capitolo dopo aver introdotto l'esperimento di Monte Carlo che si andrà ad eseguire, ne vengono riportati i risultati e le considerazioni che si possono trarre a riguardo. Nell'ultima sezione di questo elaborato vengono proposte le conclusioni del lavoro svolto.

Capitolo 1

I processi a memoria lunga

1.1 Origini

Negli ultimi decenni si è assistito a un crescente interesse della letteratura statistica ed econometrica nei confronti dei processi a memoria lunga e dei modelli a differenze stazionarie. Alla base di tale attenzione vi è la constatazione che diverse serie reali presentano funzioni di autocorrelazione persistenti, ossia che decrescono più lentamente rispetto a quanto avviene per i processi *AutoRegressive Moving Average* (ARMA). Anche se gli econometrici hanno iniziato ad utilizzarli solo a partire dagli anni '80 circa, i processi a memoria lunga hanno avuto con gli statistici un ruolo fondamentale a partire dagli anni '50.

Il primo a rivolgere il proprio interesse verso i processi a memoria lunga fu Hurst, un idrologo inglese che negli anni '50 lavorava ad un progetto di una diga sul Nilo: il suo obiettivo era quello di prevedere l'afflusso idrico e quindi di riuscire a controllare le riserve d'acqua. Hurst voleva verificare se la quantità di pioggia caduta (fattore che solitamente influenza l'innalzamento del livello d'acqua in una diga) seguisse un andamento casuale e quindi modellabile mediante un processo *Random Walk*. Egli realizzò uno strumento statistico noto come *esponente di Hurst* che permetteva di classificare le serie storiche attraverso un'analisi del valore assunto da tale esponente, e non richiedeva alcuna restrizione sulla distribuzione dei dati in esame.

Negli anni '60 il concetto di memoria lunga si è ulteriormente ampliato con Mandelbrot e Wallis. A partire da quel momento, il tema della memoria lunga si è rapidamente sviluppato e, grazie alla diversità delle applicazioni che offre, si è esteso a svariati campi quali l'ingegneria, la fisica, la geologia, la climatologia e le scienze matematiche.

1.2 Caratteristiche e definizioni

Le serie temporali generate da processi a memoria lunga possiedono delle caratteristiche ben precise [Bisaglia, 1998], che possono essere riassunte nei seguenti punti:

- Le osservazioni tendono ad assumere valori al di sopra o al di sotto della media, per periodi di tempo relativamente lunghi: valori elevati tendono ad essere seguiti da valori elevati, mentre valori bassi sono tendenzialmente seguiti da valori bassi.
- Sono presenti cicli non periodici in successione casuale, così se non si hanno abbastanza osservazioni e si osserva la serie solo per un breve periodo di tempo, può sembrare che la serie presenti un ciclo o un *trend* locale. Al contrario, se si guarda all'intera serie, risulta evidente l'assenza di *trend* persistenti o di cicli.
- Le serie appaiono complessivamente stazionarie sia in media che in varianza.
- La varianza della media campionaria converge a zero più lentamente di n^{-1} . Asintoticamente si ha:

$$\text{Var}(\bar{X}_n) \sim C_{var} n^\alpha \quad \text{quando} \quad n \rightarrow \infty \quad (1.1)$$

dove $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$, $\alpha \in (0, 1)$, e C_{var} è una costante positiva.

- La funzione di autocorrelazione (autocovarianza), a differenza per esempio dei processi ARMA, converge molto lentamente a zero, in modo

iperbolico, e quindi la dipendenza tra osservazioni sempre più lontane diminuisce molto lentamente all'aumentare del ritardo k . Asintoticamente:

$$\rho(k) \sim C_\rho k^\alpha \quad \text{quando} \quad k \rightarrow \infty \quad (1.2)$$

dove C_ρ è una costante non nulla e $\alpha \in (0, 1)$.

Spesso è proprio la natura dell'esperimento a generare osservazioni che rimangono correlate per lunghi periodi di tempo. In alcuni casi ci si può aspettare a priori una dipendenza tra le diverse osservazioni ad esempio quando non è possibile pianificare l'esperimento ma bisogna attendere il verificarsi di un qualche fenomeno naturale per poter effettuare l'osservazione, situazione questa che si presenta in svariate discipline quali l'idrologia, la geofisica e la climatologia. Talvolta, contrariamente a quanto appena detto, ci si attende che le osservazioni siano indipendenti: esempio tipico di questa situazione sono gli esperimenti in cui si cerca di eliminare i fattori che potrebbero causare dipendenza tra le osservazioni.

Assumendo che il processo osservato sia stazionario, un processo a memoria lunga può essere definito nei seguenti modi:

Definizione 1.2.1 *Il processo $X = (X_t, t \in \mathbb{Z})$ viene chiamato processo stazionario con memoria lunga o dipendenza di lungo periodo, se è stazionario e se esiste un numero reale $H^1 \in (1/2, 1)$ e una funzione $L_{\rho(k)}$ che varia lentamente all'infinito, tale che:*

$$\rho(k) \sim L_{\rho(k)} k^{2H-2} \quad (1.3)$$

quando k tende a ∞ .

Bisogna sottolineare che le definizioni di processo a memoria lunga sono definizioni asintotiche. Tutto ciò che interessa è il comportamento asintotico della funzione di autocorrelazione al tendere del ritardo all'infinito. Per determinare quindi se il processo generatore della serie storica è a memoria lunga non ci si può soffermare soltanto sullo studio di pochi coefficienti di

¹Tra il parametro H e il parametro d esiste una relazione lineare del tipo $H = d + \frac{1}{2}$

autocorrelazione, ma bisogna tentare di valutare in che modo la funzione di autocorrelazione converge a zero.

Le definizioni precedentemente enunciate possono essere estese anche al caso in cui $H \in (0, 1)$, ma la dipendenza di lungo periodo in senso forte richiede che $H \in (1/2, 1)$.

Esistono diverse classi di modelli che risultano soddisfare le condizioni proposte nelle precedenti definizioni, una di queste è costituita dai modelli ARIMA frazionari, anche detti *AutoRegressive Fractional Integrated Moving Average* (ARFIMA), che verranno presentati nel successivo paragrafo.

1.3 I modelli ARFIMA

1.3.1 Dai modelli ARMA ai modelli ARFIMA

[Box and Jenkins, 1976] hanno proposto un'estensione dei processi ARMA che riguarda certi tipi di processi non stazionari. Fino a quel momento lo strumento d'elezione per lo studio delle serie storiche economiche era la modellistica ARMA. Nel loro studio Box e Jenkins introducono processi noti come *AutoRegressive Integrated Moving Average* (ARIMA), ossia processi non stazionari omogenei di grado d , che possono essere resi stazionari tramite opportune trasformazioni. È utile quindi richiamare la nozione di stazionarietà: un processo stocastico è stazionario se la sua media e la sua variabilità non presentano cambiamenti di natura sistematica e se la sua dinamica non presenta variazioni strettamente periodiche. In molti casi è tuttavia possibile, mediante opportune trasformazioni dei dati, ricondurre alla stazionarietà anche serie originariamente non stazionarie. Una strada che si può percorrere per rendere stazionaria una serie è quella della differenziazione [Di Fonzo and Lisi, 2005]. Se X_t è la d -esima differenza di Y_t , $X_t = (1 - B)^d Y_t$, con $d \in \mathbb{N}$ si dice che Y_t è un processo integrato di ordine d e lo si indica con ARIMA(p,d,q), dove d è il parametro di integrazione o differenziazione, se l'equazione corrispondente è:

$$\Phi(B)(1 - B)^d Y_t = \Theta(B) \varepsilon_t \tag{1.4}$$

dove ε_t è un *white noise* di media 0 e varianza σ_ε^2 e $\Phi(\cdot)$ e $\Theta(\cdot)$ sono due polinomi di grado rispettivamente p e q .² Tali modelli sono in grado di cogliere soltanto la dipendenza di breve periodo, pertanto, poichè molto frequentemente in ambito economico e finanziario è presente una forte dipendenza tra le osservazioni che persiste nel tempo, [Granger and Joyeux, 1980] e [Hosking, 1981] hanno introdotto i modelli ARFIMA(p,d,q). Sin dalla loro introduzione sono stati oggetto di innumerevoli studi a causa del gran numero di serie economiche che presentano memoria lunga e che quindi necessitano di essere modellate mediante tali modelli. Le serie in cui si verificano queste situazioni includono i tassi di scambio, i tassi di interesse, le serie dei consumi, della disoccupazione e molte altre. I modelli ARFIMA(p,d,q) si propongono come generalizzazione dei modelli ARIMA(p,d,q), e risultano utili per modellare processi le cui realizzazioni evidenziano correlazioni seriali significative anche se misurate in tempi molto distanti, al contrario di quanto avviene nella modellistica più tradizionale in cui si presuppone indipendenza tra osservazioni molto distanti nel tempo.

Volendo evidenziare un collegamento tra i processi ARMA ARIMA e ARFIMA si può definire un modello ARMA(p,q), ossia un processo che soddisfa l'equazione alle differenze

$$X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} \quad (1.5)$$

dove $\varepsilon_t \sim w.n(0, \sigma^2)$, come un particolare ARIMA in cui d è uguale a 0, ovvero un ARIMA(p,0,q). La famiglia di modelli ARIMA prevede l'utilizzo di un valore intero per descrivere l'ordine di integrazione di un processo. Gli ARFIMA costituiscono quindi una generalizzazione dei modelli ARIMA in quanto il parametro d è autorizzato ad assumere un valore reale, non intero.

Se d è intero, allora:

$$(1 - B)^d = \sum_{k=0}^d \binom{d}{k} (-1)^k B^k \quad (1.6)$$

²I due polinomi $\Phi(B)$ e $\Theta(B)$ sono tali per cui $\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$, $\Theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q$, mentre l'operatore ritardo B è tale che $B^j X_t = X_{t-j}$ per $j = \dots, -1, 0, 1, \dots$

dove il coefficiente binomiale è pari a:

$$\binom{d}{k} = \frac{d!}{k!(d-k)!} = \frac{\Gamma(d+1)}{\Gamma(k+1)\Gamma(d-k+1)} \quad (1.7)$$

e $\Gamma(\cdot)$ è la funzione Gamma. Dal momento che la funzione Gamma è definita per tutti i valori reali, è possibile estendere la definizione del coefficiente binomiale al caso $d \in \mathbb{R}$, come:

$$(1-B)^d = \sum_{k=0}^{\infty} \binom{d}{k} (-1)^k B^k = \sum_{k=0}^{\infty} \pi_k B^k \quad (1.8)$$

dove

$$\pi_k = \frac{\Gamma(k-d)}{\Gamma(k+1)\Gamma(-d)} = \prod_{0 < j \leq k} \frac{j-1-d}{j}, j = 0, 1, 2, \dots \quad (1.9)$$

1.3.2 Proprietà

Definizione 1.3.1 $X = (X_t, t \in \mathbb{Z})$ è un processo ARFIMA(p, d, q) con $d \in (-1/2, 1/2)$ se è stazionario³ e se soddisfa l'equazione alle differenze:

$$\Phi(B) \Delta^d X_t = \Theta(B) \varepsilon_t \quad (1.10)$$

dove $\varepsilon_t \sim w.n(0, \sigma_\varepsilon^2)$ e $\Phi(\cdot)$ e $\Theta(\cdot)$ sono polinomi di grado p e q rispettivamente.

Quando $d \in (0, 1/2)$ il processo ARFIMA(p, d, q) è un processo stazionario con memoria lunga. In questo caso il processo $FI(d)$ ha una funzione di autocorrelazione che tende a zero molto lentamente. Quando $d = 0$ si ritorna all'usuale modello ARMA(p, q). Quando $d \in (-1/2, 0)$ il processo ARFIMA(p, d, q) è un processo stazionario con memoria intermedia o dipendenza negativa o antipersistente. Il parametro d misura l'intensità della dipendenza. Se $d \in (0, 1/2)$ il processo è stazionario e invertibile.

³Il processo è stazionario se le radici di $\Phi(B) = 0$ cadono fuori del raggio di cerchio unitario

Testare l'ipotesi nulla di dipendenza debole contro l'alternativa di memoria lunga significa verificare l'ipotesi che d sia uguale a zero contro l'alternativa che d sia diverso da zero.

Capitolo 2

Test per identificare cambiamenti di persistenza nei modelli *ARFIMA*

2.1 Introduzione

Eseguire dei test per verificare la presenza di radici unitarie è una procedura di routine nelle ricerche empiriche, a causa delle diverse implicazioni economiche e statistiche nel classificare una serie come stazionaria o non stazionaria. Quando si verificano degli *shocks* esterni sulle variabili economiche e finanziarie, a seconda che la natura delle serie in esame sia stazionaria o meno, si avranno risposte diverse. L'impatto degli *shocks* è transitorio per le serie stazionarie, pertanto il loro comportamento nel lungo periodo tenderà ad avvicinarsi ad un livello medio e lo shock tenderà ad essere riassorbito. Per quanto riguarda quelle non stazionarie gli *shocks* casuali hanno effetti persistenti che impediscono alla serie di ritornare a un qualsiasi livello definito.

In anni recenti è stata redatta una vasta letteratura che documenta i cambiamenti delle serie economiche temporali. È stato osservato che le variabili macroeconomiche possono manifestare cambiamenti nella persistenza all'interno di un determinato periodo. Infatti molte serie sembrano cambiare da

un comportamento $I(0)$ a un comportamento $I(1)$, o viceversa. Esistono vari test che si propongono di verificare quanto appena detto, ad esempio quelli realizzati da [Kim and Amador, 2002] che studiano come identificare cambiamenti strutturali dovuti a variazioni nella persistenza di una serie temporale e quelli di [Busetto and Taylor, 2004] che cercano di valutare i cambiamenti, che accadono in un momento ignoto, sia da $I(0)$ a $I(1)$ che viceversa.

In un lavoro recente, [Hassler and Scheithauer, 2011] hanno considerato i test proposti da [Kim and Amador, 2002], [Busetto and Taylor, 2004] che verificavano l'ipotesi nulla di memoria corta contro l'alternativa di un cambio verso una situazione di non stazionarietà $I(1)$ e hanno dimostrato che questi test sono adeguati anche per identificare cambiamenti da $I(d_0)$, dove $d_0 = 0$, a $I(d)$, dove $d > 0$ (memoria lunga). Tuttavia, considerare $I(0)$ come ipotesi nulla potrebbe essere restrittivo perché qualcuno potrebbe voler considerare come ipotesi nulla un ordine di integrazione diverso da zero, e non necessariamente un numero intero. Un ulteriore sviluppo è stato proposto da [Sibbertsen and Kruse, 2009] che seguono il lavoro di [Leybourne et al., 2007]. Loro adattano le statistiche del loro test Cusum basato sui quadrati, calcolato attraverso una valutazione delle serie storiche progressiva e regressiva, al contesto della memoria lunga, e dimostrano che lo stimatore del punto di rottura proposto da [Leybourne et al., 2007] è consistente sotto l'ipotesi di memoria lunga anche se presenta un tasso di convergenza più lento (che dipende da d). [Sibbertsen and Kruse, 2009] hanno osservato che la procedura di [Leybourne et al., 2007] soffre di gravi distorsioni del livello di significatività del test se il processo generatore dei dati ha memoria lunga, pertanto forniscono nuovi valori critici appropriati per il contesto $I(d)$ che dipendono dal parametro di memoria d .

Questo studio illustra procedure in grado di individuare cambiamenti nella persistenza in un processo frazionalmente integrato, argomento già trattato da [Beran and Terrin, 1996]. Viene proposto un nuovo metodo basato sulla stima ricorsiva, progressiva e regressiva, del test di [Breitung and Hassler, 2002] seguendo l'approccio di [Leybourne et al., 2003]. Vengono derivati dei risultati asintotici e la performance delle nuove procedure è valutata tramite un esperimento di Monte Carlo. In particolare viene prestata un'attenzione

speciale al caso in cui d è sconosciuto, per il quale esistono pochi risultati in letteratura. La performance del test, unita alla sua semplicità di applicazione, lo rende un approccio interessante per le analisi empiriche.

2.2 Cambiamenti nel parametro di persistenza

Si considerino i dati generati da un processo frazionalmente integrato di ordine d_t , tale che:

$$(1 - L)^{d_t} y_t = \varepsilon_t \quad (2.1)$$

dove $y_t = 0$ per $t \leq 0$, e ε_t soddisfa le assunzioni che verranno riportate sotto. Sotto l'ipotesi nulla si assume che il parametro di integrazione frazionaria d_t sia costante nel campione, $d_t = d_0$. Sotto l'ipotesi alternativa, considerando due parametri di integrazione frazionaria, tali che d_0 corrisponde al primo sottocampione e d_1 al secondo, vengono considerate due situazioni:

- una diminuzione nella persistenza con $d_0 > d_1$, che si denota con H_{01} ,
- un aumento nella persistenza con $d_0 < d_1$, che si denota con H_{10} .

Sotto entrambe le alternative il cambiamento nella persistenza avviene all'istante $[\tau^*N]$, dove τ^* assume valori nell'intervallo $[\Lambda_l, \Lambda_u] \subset (0, 1)$ e $\Lambda_u = 1 - \Lambda_l$. Per quanto riguarda le assunzioni relative a ε_t si considerino le seguenti:

($\mathcal{A}.1$) Il processo delle innovazioni ε_t è indipendente e identicamente distribuito con $E(\varepsilon_t^2) = \sigma^2 < \infty$ e $E(|\varepsilon_t^4|^{1+r})$ è limitato per alcuni $r > 0$.

($\mathcal{A}.2$) Il processo ε_t soddisfa $a(L)\varepsilon_t = v_t$, dove $a(L) = 1 - \sum_j^p a_j L^j$, $p \geq 0$, tale che $a(z)$ ha tutte le sue radici fuori del cerchio di raggio unitario e v_t soddisfa l'assunzione ($\mathcal{A}.1$).

2.3 Test per i cambiamenti di persistenza

Per introdurre i test per verificare cambiamenti nella persistenza si considerino i dati generati da (2.1) con $d_t = d_0$ e con ε_t che soddisfa l'assunzione (A.1). In questo caso, per ogni fissato $\tau \in [\Lambda_l, \Lambda_u]$ la regressione ausiliaria è semplicemente:

$$x_t = \phi(\tau) x_{t-1}^* + e_t, \quad t = 2, \dots, [\tau N] \quad (2.2)$$

dove $x_t = (1 - L)^{d_0} y_t$ e $x_{t-1}^* = \sum_{j=1}^{t-1} \frac{x_{t-j}}{j}$, per i dettagli sull'impostazione della regressione per i test di integrazione frazionaria si veda [Breitung and Hassler, 2002]. Noi siamo interessati ad individuare cambiamenti nel parametro di memoria stimando ricorsivamente la (2.2) sul campione completo. In particolare, considerando la regressione ausiliaria (2.2), la statistica t ai minimi quadrati per $\hat{\phi}(\tau) = 0$, che denotiamo con $t_{\phi f}(\tau)$, è calcolata per ciascun $\tau \in [\Lambda_l, \Lambda_u]$ ed è tale che:

$$t_{\phi f}(\tau) = \frac{\sum_{t=2}^{\tau N} x_t x_{t-1}^*}{\hat{\sigma}_e(\tau) \sqrt{\sum_{t=2}^{\tau N} x_{t-1}^{*2}}} \quad (2.3)$$

dove $\hat{\sigma}_e(\tau) = \sqrt{\frac{1}{\tau N - 2} \sum_{t=2}^{\tau N} \hat{e}_t^2}$ e \hat{e}_t è il residuo ai minimi quadrati di (2.2). Il parametro Λ_l che definisce il limite inferiore dell'insieme di valori per τ è un valore arbitrario. Generalmente in letteratura $\Lambda_l = 0.15$ oppure $\Lambda_l = 0.20$.

Generalizzando i risultati di [Leybourne et al., 2007] segue che, a causa dei problemi di consistenza della statistica $t_{\phi f}(\tau)$, viene calcolata anche la statistica invertita $t_{\phi r}(\tau)$, dove x_t è rimpiazzata dalla serie invertita nel tempo $w_t = x_{N-t+1}$. Quindi, per le restanti $(1 - \tau)N$ osservazioni, segue che:

$$w_t = \phi(\tau) w_{t-1}^* + u_t, \quad t = 2, \dots, [(1 - \tau)N] \quad (2.4)$$

dove $w_t = x_{N-t+1}$, $w_{t-1}^* = \sum_{j=1}^{t-1} \frac{w_{t-j}}{j} = \sum_{j=1}^{t-1} \frac{x_{N-t+j+1}}{j} = x_{N-t+2}^*$ e la risultan-

te statistica t necessaria a verificare $\hat{\phi}(\tau) = 0$ è:

$$t_{\phi r}(\tau) = \frac{\sum_{t=2}^{(1-\tau)N} w_t w_{t-1}^*}{\widehat{\sigma}_u(\tau) \sqrt{\sum_{t=2}^{(1-\tau)N} w_{t-1}^{*2}}} = \frac{\sum_{t=2}^{(1-\tau)N} x_{N-t+1} x_{N-t+2}^*}{\widehat{\sigma}_u(\tau) \sqrt{\sum_{t=2}^{(1-\tau)N} x_{N-t+2}^{*2}}} \quad (2.5)$$

dove $\widehat{\sigma}_u(\tau) = \sqrt{\frac{1}{(1-\tau)N} \sum_{t=2}^{(1-\tau)N} \widehat{u}_t^2}$ e \widehat{u}_t è il residuo ai minimi quadrati della regressione (2.4).

Dato che generalmente quando si realizza il test per i cambiamenti nella persistenza il momento in cui avviene il cambiamento è ignoto, è utile fornire anche il valore inferiore della statistica t calcolata dalla (2.2) e dalla (2.4) per ogni $\tau \in [\Lambda_l, \Lambda_u]$, tale che:

$$\mathcal{T}_k \equiv \inf_{\tau \in [\Lambda_l, \Lambda_u]} t_{\phi k}(\tau) \quad \text{con} \quad k = f, r \quad (2.6)$$

Se sotto l'alternativa la direzione del cambiamento è nota possiamo calcolare sia \mathcal{T}_f sia \mathcal{T}_r . Tuttavia se la direzione del cambiamento è sconosciuta, come di solito accade, si raccomanda l'utilizzo di $\mathcal{T}_{min} = \min(\mathcal{T}_f, \mathcal{T}_r)$.

Osservazione 1. Per investigare possibili cambiamenti nel parametro di memoria è possibile considerare anche il quadrato della statistica t calcolata in (2.3) e in (2.5) e individuarne il massimo, tale che:

$$\mathcal{T}_k^2 \equiv \sup_{\tau \in [\Lambda_l, \Lambda_u]} t_{\phi k}^2(\tau) \quad \text{con} \quad k = f, r \quad (2.7)$$

Inoltre, come suggerito per le statistiche \mathcal{T}_f e \mathcal{T}_r , se la direzione del cambiamento è sconosciuta, si può utilizzare $\mathcal{T}_{max}^2 = \max(\mathcal{T}_f^2, \mathcal{T}_r^2)$. Una proprietà interessante delle statistiche introdotte in questa osservazione è che sotto l'ipotesi alternativa il segno della differenza tra d_0 e d_1 non è rilevante, contrariamente a quanto avviene per le statistiche t (2.3) e (2.5).

2.3.1 Distribuzioni limite sotto l'ipotesi nulla

Ai fini di analizzare le distribuzioni limite del test statistico, il processo generatore dei dati in (2.1) è considerato sotto l'ipotesi nulla che $d_t = d_0$.

Così si può enunciare il seguente teorema per $x_t = (1-L)^{d_0} y_t$ che caratterizza le distribuzioni limite di $t_{\phi_f}(\tau)$ e $t_{\phi_r}(\tau)$.

Teorema 2.1 *Considerando i dati generati da (2.1) con $d_t = d_0$, $x_t = (1-L)^{d_0} y_t$ e con ε_t che soddisfa l'assunzione (A.1) durante il periodo di campionamento, ne consegue che per $N \rightarrow \infty$ per ogni fissato, noto $\tau \in [\Lambda_l, \Lambda_u]$ le statistiche proposte in (2.3) e (2.5) sono distribuite come normali standard, tali che:*

$$t_{\phi_k}(\tau) \xrightarrow{d} N(0, 1), \quad \text{con} \quad k = f, r \quad (2.8)$$

Inoltre, per quanto riguarda le statistiche test in (2.6) e riferendosi alla convergenza uniforme piuttosto che alla convergenza puntuale come nel teorema 2.1, le corrispondenti distribuzioni asintotiche sono proposte nel seguente teorema.

Teorema 2.2 *Sotto le assunzioni del teorema 2.1, segue che per $N \rightarrow \infty$:*

$$\mathcal{T}_f \Rightarrow \inf_{\tau \in [\Lambda_l, \Lambda_u]} \Phi_{f\tau} \quad (2.9a)$$

$$\mathcal{T}_r \Rightarrow \inf_{\tau \in [\Lambda_l, \Lambda_u]} \Phi_{r\tau} \quad (2.9b)$$

dove $\Phi_{f\tau} = W(\tau) / \sqrt{\tau}$, $\Phi_{r\tau} = [W(1) - W(\tau)] / \sqrt{1-\tau}$ e $W(r)$ rappresenta un processo di Wiener. Inoltre,

$$\mathcal{T}_{min} \Rightarrow \min \left\{ \inf_{\tau \in [\Lambda_l, \Lambda_u]} \Phi_{f\tau}, \inf_{\tau \in [\Lambda_l, \Lambda_u]} \Phi_{r\tau} \right\} \quad (2.10)$$

Tuttavia, poiché l'ordine di integrazione d_0 è generalmente sconosciuto, per implementare il test bisogna considerare una sua stima consistente. Quindi si considera uno stimatore \widehat{d}_N consistente con tasso di convergenza pari a \sqrt{N} (si veda [Fox and Taqqu, 1986]) del parametro d_0 di un modello ad integrazione frazionaria e si definisce $\widehat{x}_t = (1-L)^{\widehat{d}_N} y_t$ da utilizzare nell'impostazione delle regressioni ausiliarie (2.2) e (2.4), dove \widehat{d}_N è ottenuto dal campione completo, $t = 1, \dots, N$. L'uso di \widehat{d}_N piuttosto che del vero d_0 per calcolare \widehat{x}_t ha implicazioni sulle distribuzioni limite; si veda [Hassler and

Breitung, 2006]. Il seguente teorema stabilisce quindi le nuove distribuzioni limite delle statistiche test in questo contesto.

Teorema 2.3 *Assumendo che i dati siano generati da (2.1) sotto l'assunzione (A.1) e che venga utilizzato uno stimatore d_0 \sqrt{N} -consistente, detto \widehat{d}_N , per ottenere la variabile filtrata $\widehat{x}_t = (1 - L)^{d_0} y_t$ usata per impostare la regressione (2.2) e (2.4), segue che per $N \rightarrow \infty$ le statistiche test in (2.3), (2.5), (2.6) e la statistica \mathcal{T}_{min} convergono rispettivamente a:*

$$t_{\phi_f}^{\widehat{d}_N} \Rightarrow \left(\frac{1}{\sqrt{\pi}} - \frac{\pi}{\sqrt{6}} \sqrt{\vartheta(d_0)} \right) W(\tau) \equiv \Phi_{f\tau}(d_0) \quad (2.11a)$$

$$t_{\phi_r}^{\widehat{d}_N} \Rightarrow \left(\frac{1}{\sqrt{1-\pi}} - \frac{\pi}{\sqrt{6}} \sqrt{\vartheta(d_0)} \right) [W(1) - W(\tau)] \equiv \Phi_{r\tau}(d_0) \quad (2.11b)$$

$$\mathcal{T}_f \Rightarrow \inf_{\tau \in [\Lambda_l, \Lambda_u]} \Phi_{f\tau}(d_0) \quad (2.11c)$$

$$\mathcal{T}_r \Rightarrow \inf_{\tau \in [\Lambda_l, \Lambda_u]} \Phi_{r\tau}(d_0) \quad (2.11d)$$

$$\mathcal{T}_{min} \Rightarrow \min \left\{ \inf_{\tau \in [\Lambda_l, \Lambda_u]} \Phi_{f\tau}(d_0), \inf_{\tau \in [\Lambda_l, \Lambda_u]} \Phi_{r\tau}(d_0) \right\} \quad (2.11e)$$

dove $\sqrt{T}(\widehat{d}_t - d_0) \xrightarrow{d} N(0, \vartheta(d_0))$ e $\vartheta(d_0)$ corrisponde alla varianza dello stimatore.

Corollario 2.1 *Sotto le assunzioni del teorema (2.3), se per stimare d_0 è usato uno stimatore di Whittle o uno stimatore di Massima Verosimiglianza, allora le distribuzioni limite delle statistiche test proposte in (2.3), (2.5), (2.6) e la statistica \mathcal{T}_{min} convergono rispettivamente a:*

$$t_{\phi_f}^{\widehat{d}_N}(\tau) \Rightarrow \left(\frac{1}{\sqrt{\tau}} - 1 \right) W(\tau) \equiv \Phi_{f\tau}^* \quad (2.12a)$$

$$t_{\phi_r}^{\widehat{d}_N} \Rightarrow \left(\frac{1}{\sqrt{1-\pi}} - 1 \right) [W(1) - W(\tau)] \equiv \Phi_{r\tau}^* \quad (2.12b)$$

$$\mathcal{T}_f \Rightarrow \inf_{\tau \in [\Lambda_l, \Lambda_u]} \Phi_{f\tau}^* \quad (2.12c)$$

$$\mathcal{T}_r \Rightarrow \inf_{\tau \in [\Lambda_l, \Lambda_u]} \Phi_{r\tau}^* \quad (2.12d)$$

$$\mathcal{T}_{min} \Rightarrow \min \left\{ \inf_{\tau \in [\Lambda_l, \Lambda_u]} \Phi_{f\tau}^*, \inf_{\tau \in [\Lambda_l, \Lambda_u]} \Phi_{r\tau}^* \right\} \quad (2.12e)$$

2.4 Lo stimatore di Whittle

Lo stimatore di Whittle si basa su un'approssimazione della funzione di verosimiglianza. Esistono in letteratura diversi approcci per determinare una buona approssimazione della funzione di verosimiglianza, basati essenzialmente sul lavoro di [Whittle, 1951], il quale ha introdotto tale approssimazione nel caso di variabili casuali debolmente indipendenti.

Si consideri una realizzazione, $X = (X_t, t = 1, \dots, N)$, di un processo gaussiano lineare, invertibile e casuale la cui funzione di verosimiglianza si presenta nella seguente forma:

$$\mathcal{L}_N(X, \theta) = -\frac{N}{2} \log 2\pi - \frac{1}{2} \log |\Sigma_N(\theta)| - \frac{1}{2} X' \Sigma_N^{-1}(\theta) X \quad (2.13)$$

dove $\theta = (\sigma^2, H, \theta_3, \dots, \theta_M)$ è il vettore di parametri sconosciuti e $\Sigma_N(\theta)$ è la matrice di varianze e covarianze.

I due termini $\log |\Sigma_N(\theta)|$ e $\Sigma_N^{-1}(\theta)$, al fine di semplificare i calcoli ed evitare di incorrere in una serie di problemi, possono essere approssimati mediante le seguenti espressioni:

$$\log |\Sigma_N(\theta)| \sim \frac{N}{2\pi} \int_{-\pi}^{\pi} \log f(\lambda, \theta) d\lambda, \quad (2.14)$$

e

$$\Sigma_N^{-1}(\theta) \sim A(\theta) = [\alpha_{j-k}(\theta)]_{j,k=1,2,\dots,N}, \quad (2.15)$$

dove

$$\alpha_{j-k}(\theta) = \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} \frac{e^{i(j-k)\lambda}}{f(\lambda, \theta)} d\lambda \quad (2.16)$$

In tal modo è possibile ottenere un'approssimazione dello stimatore di massima verosimiglianza, del vettore di parametri θ , minimizzando, rispetto a θ , la funzione di log-verosimiglianza approssimata:

$$\mathcal{L}_W(X, \theta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log f(\lambda, \theta) d\lambda + \frac{X' A(\theta) X}{N}. \quad (2.17)$$

Si osservi che:

$$\frac{X' A(\theta) X}{N} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{I_N(\lambda)}{f(\lambda, \theta)} d\lambda \quad (2.18)$$

dove $I_N(\lambda)$ è il periodogramma, cioè:

$$I_N(\lambda) = \frac{1}{2\pi N} \left| \sum_{j=1}^N e^{ij\lambda} (X_j - \bar{X}_N)^2 \right| \quad (2.19)$$

La funzione di log-verosimiglianza diventa così:

$$\mathcal{L}_W(X, \theta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\log f(\lambda, \theta) + \frac{I_N(\lambda)}{f(\lambda, \theta)} \right) d\lambda. \quad (2.20)$$

L'indice W denota lo stimatore di Whittle, mentre $\widehat{\theta}_W$ il valore di θ che minimizza la funzione di cui sopra.

Lo stimatore di Whittle è computazionalmente semplice, tuttavia questo stimatore ha lo svantaggio di richiedere una semplificazione a priori della forma parametrica di $f(\lambda, \theta)$.

2.5 Processo di Wiener

Un processo di Wiener è un processo stocastico Gaussiano in tempo continuo. Un processo stocastico $(W_t)_{t \geq 0}$ definito su uno spazio di probabilità (Ω, \mathcal{F}, P) , si dice processo di Wiener, o moto Browniano, se verifica le seguenti proprietà:

- il processo parte da zero, ovvero $W_0=0$ quasi certamente,
- ha *incrementi indipendenti*, cioè per ogni scelta di $0 \leq t_1 < t_2 < \dots < t_n$, con $n \in \mathbb{N}$, le variabili aleatorie $(W_{t_1}, W_{t_2} - W_{t_1}, \dots, W_{t_n} - W_{t_{n-1}})$ sono indipendenti,
- ha *incrementi gaussiani*, cioè presi $0 \leq s < t$, $(W_t - W_s)$ ha legge gaussiana $N(0, t - s)$,
- le traiettorie del processo (W_t) , $t \rightarrow W_t$, sono quasi certamente continue.

Capitolo 3

Esperimento Monte Carlo

Il metodo Monte Carlo appartiene alla famiglia dei test statistici non parametrici e risulta particolarmente utile per affrontare i problemi computazionali legati ai test. Lo scopo del nostro esperimento è quello di verificare se i test, descritti nel capitolo precedente, si comportano bene in termini di livello e potenza empirici, in campioni di numerosità finita. Per fare ciò verranno seguite le procedure proposte in modo dettagliato nel capitolo 2.

In questo capitolo vengono proposti i valori critici per:

$$\mathcal{I}_f, \mathcal{I}_r, \mathcal{I}_{r,Rov}, \min(\mathcal{I}_f, \mathcal{I}_r), \min(\mathcal{I}_f, \mathcal{I}_{r,Rov}), \mathcal{I}_f^2, \mathcal{I}_r^2, \mathcal{I}_{r,Rov}^2, \\ \max(\mathcal{I}_f^2, \mathcal{I}_r^2), \max(\mathcal{I}_f^2, \mathcal{I}_{r,Rov}^2)$$

per il contesto in cui d_0 è stimato attraverso il metodo di Fox e Taqu(1986). Vengono inoltre fornite il livello e la potenza empirici dei test.

Per realizzare l'esperimento di Monte Carlo, in cui viene considerato come DGP un processo FI(d), abbiamo generato i dati a partire da un modello ARFIMA(0,d,0), tale che: $(1 - L)^{d_0}y_t = \varepsilon_t$, dove $y_t = 0$ per $t \leq 0$ ed $\varepsilon_t \sim n.i.d(0, 1)$. Al fine di determinare i valori critici l'esperimento è stato ripetuto per cinque diverse numerosità campionarie delle serie simulate, pari rispettivamente a $N = 100, 250, 500, 750, 1000$, e il numero delle replicazioni realizzate, per DGP, è stato pari a 10000; per verificare la performance dei test in termini di *size* e potenza dei test effettuati abbiamo considerato soltanto due diverse numerosità delle serie, pari rispettivamente a $N = 250$ e

$N = 500$. Il software utilizzato per la realizzazione delle simulazioni è stato **R**, i cui comandi sono riportati nell'appendice.

Tutti i dati presenti nelle tabelle che verranno inserite in seguito, relative ai valori critici, alla *size* e alla potenza dei test, sono stati arrotondati alla quarta cifra decimale.

3.1 Valori critici

Poiché siamo interessati soltanto ai processi che presentano memoria lunga abbiamo considerato soltanto valori di d_0 appartenenti all'intervallo $(0, 1/2)$, $d_0 = 0.0, 0.1, 0.2, 0.3, 0.4, 0.45, 0.49$.

La determinazione dei valori critici serve per distinguere la regione di rifiuto dalla regione di accettazione, e ci permette di determinare la *size* la potenza del test.

Per il calcolo dei valori critici abbiamo prima simulato, mediante il software **R**, una serie $FI(d_0)$ attraverso la funzione `armaSim` disponibile nel pacchetto `fArma`; in seguito abbiamo stimato d a partire dalla serie precedentemente simulata e quindi abbiamo calcolato i residui della serie simulata con il parametro d stimato. Per stimare d abbiamo utilizzato il metodo descritto in dettaglio nel capitolo 2.4 relativo allo stimatore di Whittle. Abbiamo poi impostato le tre diverse regressioni ausiliarie necessarie a ricavare i diversi valori di \mathcal{T} . Rispetto a quanto fatto nell'articolo [Martins and Rodrigues, 2014] abbiamo ritenuto opportuno includere anche i valori di \mathcal{T} che sono stati calcolati a partire dalla serie invertita nel tempo, procedimento che segue quanto proposto nella parte teorica; è per questo motivo che ci ritroviamo ad avere tre diverse regressioni anziché due (la terza è quella utilizzata per calcolare i valori sulla serie invertita nel tempo). Pertanto, oltre ai valori che vengono calcolati nell'articolo, in questo lavoro sono stati aggiunti $\mathcal{T}_{r,Rov}$, $\min(\mathcal{T}_f, \mathcal{T}_{r,Rov})$, $\mathcal{T}_{r,Rov}^2$ e $\max(\mathcal{T}_f^2, \mathcal{T}_{r,Rov}^2)$. Poiché nelle diverse regressioni ausiliarie è necessario definire anche dei valori per τ , abbiamo ritenuto opportuno utilizzare un range di valori compreso tra 0.2 e 0.8 con un passo di incremento dello 0.05. Quindi, dopo aver impostato le tre diverse regressioni

necessarie a calcolare:

$$\mathcal{I}_f, \mathcal{I}_r, \mathcal{I}_{r,Rov}, \min(\mathcal{I}_f, \mathcal{I}_r), \min(\mathcal{I}_f, \mathcal{I}_{r,Rov}), \mathcal{I}_f^2, \mathcal{I}_r^2, \mathcal{I}_{r,Rov}^2, \\ \max(\mathcal{I}_f^2, \mathcal{I}_r^2) \text{ e } \max(\mathcal{I}_f^2, \mathcal{I}_{r,Rov}^2)$$

abbiamo individuato i valori critici corrispondenti ai diversi livelli di significatività. Nello specifico: per $\mathcal{I}_f, \mathcal{I}_r, \mathcal{I}_{r,Rov}, \min(\mathcal{I}_f, \mathcal{I}_r)$ e $\min(\mathcal{I}_f, \mathcal{I}_{r,Rov})$, poiché si trovano nella coda sinistra della distribuzione, abbiamo calcolato i valori critici ad un livello di significatività del 1%, 5% e 10%, mentre per $\mathcal{I}_f^2, \mathcal{I}_r^2, \mathcal{I}_{r,Rov}^2, \max(\mathcal{I}_f^2, \mathcal{I}_r^2)$ e $\max(\mathcal{I}_f^2, \mathcal{I}_{r,Rov}^2)$, poiché si trovano nella coda destra della distribuzione, li abbiamo calcolati ad un livello di significatività del 90%, 95% e 99%.

I valori critici proposti nelle tabelle 3.1 e 3.2 sono validi per i test di regressione con e senza termine deterministico. Tuttavia, tutti i risultati proposti in questa sezione sono stati calcolati senza includere il termine deterministico.

Tabella 3.1: Valori critici per $\mathcal{F}_f, \mathcal{F}_r, \mathcal{F}_{r,Rov}, \min(\mathcal{F}_f, \mathcal{F}_r), \min(\mathcal{F}_f, \mathcal{F}_{r,Rov})$

d	$N = 100$				$N = 250$				$N = 500$				$N = 750$				$N = 1000$			
	0.01	0.05	0.1	0.1	0.01	0.05	0.1	0.1	0.01	0.05	0.1	0.1	0.01	0.05	0.1	0.01	0.05	0.1		
0.0	\mathcal{F}_f	-3.0579	-2.3618	-2.0242	-3.0824	-2.3903	-2.0114	-2.0114	-2.9172	-2.3204	-1.9858	-1.9858	-2.9259	-2.2778	-1.9727	-2.9605	-2.3166	-1.9723		
	\mathcal{F}_r	-3.1536	-2.4214	-2.0684	-3.0077	-2.3693	-2.0142	-2.0142	-3.0196	-2.3161	-1.9888	-1.9888	-2.9392	-2.3307	-2.0049	-2.9761	-2.3138	-2.0085		
	$\mathcal{F}_{r,Rov}$	-0.7588	-0.0752	0.2691	0.4367	1.0978	1.4541	1.6705	1.6705	2.3871	2.7643	2.7333	3.3692	3.3692	3.7426	3.5032	4.1960	4.5284		
	min	-3.3351	-2.7163	-2.3671	-3.3088	-2.6646	-2.3445	-2.3445	-3.1613	-2.5849	-2.2971	-2.2971	-3.1667	-2.5674	-2.2789	-3.1652	-2.5817	-2.2827		
0.1	minRov	-3.0579	-2.3627	-2.0250	-3.0824	-2.3903	-2.0114	-2.0114	-2.9172	-2.3204	-1.9858	-1.9858	-2.9259	-2.2778	-1.9727	-2.9605	-2.3166	-1.9723		
	\mathcal{F}_f	-2.7083	-2.0163	-1.6681	-2.5076	-1.8405	-1.5145	-1.5145	-2.3877	-1.7666	-1.4802	-1.4802	-2.3014	-1.7555	-1.4694	-2.3854	-1.8253	-1.5133		
	\mathcal{F}_r	-2.7434	-2.0060	-1.6908	-2.4295	-1.8650	-1.5609	-1.5609	-2.4088	-1.8147	-1.5076	-1.5076	-2.4353	-1.7978	-1.4917	-2.4149	-1.8064	-1.4959		
	$\mathcal{F}_{r,Rov}$	-0.5824	0.1104	0.4632	0.6588	1.3772	1.7305	1.7305	1.9712	2.6507	2.9991	2.9991	2.9706	3.6354	3.9732	3.7789	4.4417	4.7808		
0.2	min	-3.0063	-2.3339	-2.0058	-2.7061	-2.1523	-1.8514	-1.8514	-2.6233	-2.0753	-1.7844	-1.7844	-2.6126	-2.0399	-1.7783	-2.6044	-2.0902	-1.8151		
	minRov	-2.7083	-2.0173	-1.6704	-2.5076	-1.8405	-1.5145	-1.5145	-2.3877	-1.7666	-1.4802	-1.4802	-2.3014	-1.7555	-1.4694	-2.3854	-1.8253	-1.5133		
	\mathcal{F}_f	-2.5436	-1.8717	-1.5207	-2.3959	-1.7715	-1.4756	-1.4756	-2.3704	-1.7723	-1.4755	-1.4755	-2.4280	-1.7749	-1.4975	-2.3603	-1.7749	-1.4765		
	\mathcal{F}_r	-2.6326	-1.9397	-1.5782	-2.4445	-1.8607	-1.5453	-1.5453	-2.4208	-1.8529	-1.5340	-1.5340	-2.3588	-1.8429	-1.5409	-2.3399	-1.7982	-1.5001		
0.3	$\mathcal{F}_{r,Rov}$	-0.5015	0.1738	0.5305	0.7272	1.3653	1.7216	1.7216	1.9901	2.6353	2.9983	2.9983	2.9941	3.6154	3.9921	3.7478	4.4289	4.7660		
	min	-2.8749	-2.2041	-1.9033	-2.6626	-2.0954	-1.8145	-1.8145	-2.6425	-2.0634	-1.8103	-1.8103	-2.5861	-2.0744	-1.8080	-2.5895	-2.0448	-1.7882		
	minRov	-2.5446	-1.8721	-1.5225	-2.3959	-1.7715	-1.4756	-1.4756	-2.3704	-1.7723	-1.4755	-1.4755	-2.4280	-1.7749	-1.4975	-2.3603	-1.7749	-1.4765		
	\mathcal{F}_f	-2.5139	-1.8026	-1.4628	-2.3990	-1.7765	-1.4407	-1.4407	-2.3721	-1.7656	-1.4323	-1.4323	-2.3148	-1.7772	-1.4585	-2.3231	-1.7251	-1.4264		
0.4	\mathcal{F}_r	-2.5521	-1.9313	-1.5832	-2.4286	-1.8341	-1.5342	-1.5342	-2.4676	-1.8148	-1.5332	-1.5332	-2.4404	-1.8362	-1.5281	-2.4584	-1.8330	-1.5152		
	$\mathcal{F}_{r,Rov}$	-0.4719	0.2025	0.5357	0.7105	1.3655	1.7205	1.7205	1.9192	2.5891	2.9772	2.9772	2.9045	3.5919	3.9413	3.7190	4.4332	4.7691		
	min	-2.8106	-2.1617	-1.8595	-2.6801	-2.0956	-1.8078	-1.8078	-2.6239	-2.0731	-1.7875	-1.7875	-2.6390	-2.0656	-1.8056	-2.6067	-2.0578	-1.7803		
	minRov	-2.5139	-1.8056	-1.4664	-2.3990	-1.7765	-1.4407	-1.4407	-2.3721	-1.7656	-1.4323	-1.4323	-2.3148	-1.7772	-1.4585	-2.3231	-1.7251	-1.4264		
0.45	\mathcal{F}_f	-2.4749	-1.7312	-1.3695	-2.3043	-1.6626	-1.3475	-1.3475	-2.2767	-1.6941	-1.3593	-1.3593	-2.3138	-1.6955	-1.3566	-2.2733	-1.6636	-1.3552		
	\mathcal{F}_r	-2.6389	-1.9476	-1.5873	-2.5339	-1.8931	-1.5707	-1.5707	-2.4453	-1.8345	-1.5363	-1.5363	-2.4466	-1.8667	-1.5748	-2.5051	-1.8528	-1.5568		
	$\mathcal{F}_{r,Rov}$	-0.5540	0.1588	0.5433	0.6640	1.3606	1.6845	1.6845	2.0201	2.6497	2.9927	2.9927	2.9097	3.5669	3.9182	3.7303	4.3876	4.7493		
	min	-2.8778	-2.1527	-1.8383	-2.7220	-2.0730	-1.7920	-1.7920	-2.6115	-2.0522	-1.7658	-1.7658	-2.5830	-2.0719	-1.7911	-2.6527	-2.0556	-1.7615		
0.49	minRov	-2.4749	-1.7368	-1.3766	-2.3043	-1.6626	-1.3475	-1.3475	-2.2767	-1.6941	-1.3593	-1.3593	-2.3138	-1.6955	-1.3566	-2.2733	-1.6636	-1.3552		
	\mathcal{F}_f	-2.2902	-1.6018	-1.2026	-2.2980	-1.6166	-1.2460	-1.2460	-2.2501	-1.5747	-1.2342	-1.2342	-2.2746	-1.5725	-1.2459	-2.2124	-1.5820	-1.2728		
	\mathcal{F}_r	-2.5913	-1.9039	-1.5420	-2.5365	-1.9023	-1.5464	-1.5464	-2.4810	-1.8502	-1.5345	-1.5345	-2.4167	-1.8142	-1.5284	-2.4750	-1.8679	-1.5563		
	$\mathcal{F}_{r,Rov}$	-0.4937	0.1874	0.5513	0.6258	1.3245	1.6856	1.6856	1.9688	2.5980	2.9480	2.9480	3.0371	3.6049	3.9211	3.6809	4.3067	4.6730		
0.49	min	-2.7638	-2.0707	-1.7671	-2.6778	-2.0797	-1.7565	-1.7565	-2.5865	-2.0335	-1.7327	-1.7327	-2.5574	-2.0040	-1.7178	-2.5960	-2.0384	-1.7480		
	minRov	-2.2920	-1.6086	-1.2157	-2.2980	-1.6166	-1.2460	-1.2460	-2.2501	-1.5747	-1.2342	-1.2342	-2.2746	-1.5725	-1.2459	-2.2124	-1.5820	-1.2728		
	\mathcal{F}_f	-2.0334	-1.2034	-0.7613	-2.0399	-1.2288	-0.8034	-0.8034	-1.9259	-1.2323	-0.8418	-0.8418	-1.9806	-1.2716	-0.8746	-1.9475	-1.2462	-0.8800		
	\mathcal{F}_r	-2.5991	-1.8006	-1.3650	-2.3842	-1.7216	-1.3670	-1.3670	-2.3896	-1.7114	-1.3761	-1.3761	-2.3669	-1.7751	-1.4388	-2.3615	-1.7835	-1.4522		
0.49	$\mathcal{F}_{r,Rov}$	-0.4522	0.2631	0.6202	0.7266	1.3971	1.7404	1.7404	1.9661	2.6105	2.9514	2.9514	2.8721	3.5358	3.8843	3.6307	4.2949	4.6582		
	min	-2.7238	-1.9578	-1.5483	-2.5215	-1.8677	-1.5261	-1.5261	-2.4444	-1.8625	-1.5263	-1.5263	-2.4478	-1.8933	-1.5910	-2.4724	-1.8936	-1.5769		
	minRov	-2.0334	-1.2123	-0.7822	-2.0399	-1.2304	-0.8036	-0.8036	-1.9259	-1.2323	-0.8418	-0.8418	-1.9806	-1.2716	-0.8746	-1.9475	-1.2462	-0.8800		

Tabella 3.2: Valori critici per $\mathcal{I}_f^2, \mathcal{I}_r^2, \mathcal{I}_{r,Rov}^2, \max(\mathcal{I}_f^2, \mathcal{I}_r^2), \max(\mathcal{I}_f^2, \mathcal{I}_{r,Rov}^2)$

d		$N = 100$			$N = 250$			$N = 50$			$N = 750$			$N = 1000$		
		0.9	0.95	0.99	0.9	0.95	0.99	0.9	0.95	0.99	0.9	0.95	0.99	0.9	0.95	0.99
0.0	\mathcal{I}_f^2	4.4326	5.8768	9.5226	4.3167	5.9126	9.6186	4.2361	5.5482	8.5174	4.1397	5.3953	8.6289	4.2425	5.5317	9.0597
	\mathcal{I}_r^2	4.5908	6.0929	9.9977	4.3183	5.7475	9.0800	4.2172	5.5199	9.2349	4.2688	5.5917	8.7389	4.2337	5.5230	8.9345
	$\mathcal{I}_{r,Rov}^2$	24.3542	27.7477	34.6943	48.3434	52.4562	60.4492	87.8754	93.1697	103.8807	126.2837	132.9353	146.3795	164.0743	171.0574	185.7793
	\max	5.8810	7.5342	11.2843	5.6933	7.1972	11.0113	5.4155	6.8338	10.0282	5.3499	6.6795	10.0952	5.3699	6.7751	10.0986
	\max_{Rov}	24.3542	27.7477	34.6943	48.3434	52.4562	60.4492	87.8754	93.1697	103.8807	126.2837	132.9353	146.3795	164.0743	171.0574	185.7793
0.1	\mathcal{I}_f^2	4.2077	5.6402	9.9491	3.6503	4.8720	8.1815	3.3160	4.3361	6.9995	3.2891	4.2135	6.5185	3.3233	4.2818	6.3382
	\mathcal{I}_r^2	4.2077	5.6402	9.9491	3.5482	4.6925	7.6544	3.3729	4.4513	6.9898	3.2696	4.2676	6.9212	3.3267	4.2646	6.7387
	$\mathcal{I}_{r,Rov}^2$	27.0074	30.4440	36.4273	51.8978	55.9695	64.2129	92.0533	97.3819	107.5773	131.2600	137.4613	150.5303	170.7167	177.1339	191.6006
	\max	5.4470	7.1753	12.2594	4.5479	5.8336	9.3006	4.2193	5.3597	8.2504	4.0719	5.1080	7.6557	4.0924	5.1475	7.7161
	\max_{Rov}	27.0074	30.4440	36.4273	51.8978	55.9695	64.2129	92.0533	97.3819	107.5773	131.2600	137.4613	150.5303	170.7167	177.1339	191.6006
0.2	\mathcal{I}_f^2	4.9917	7.1418	13.0636	3.8434	5.2261	9.1181	3.5098	4.5761	7.4862	3.4536	4.4763	7.1790	3.3366	4.3465	6.8697
	\mathcal{I}_r^2	4.7615	6.7982	13.3784	3.7577	4.9386	8.4260	3.4551	4.4656	7.3061	3.3714	4.3555	6.4775	3.2337	4.2030	6.4392
	$\mathcal{I}_{r,Rov}^2$	28.0398	31.3968	39.1059	51.8615	56.1096	64.0216	91.7910	96.6377	106.7405	130.9571	136.8497	149.0311	169.7471	176.7317	189.9574
	\max	6.3156	8.5325	15.1260	4.7978	6.1468	10.0409	4.3415	5.5138	8.5773	4.2495	5.2565	8.0978	4.0788	5.1066	7.6504
	\max_{Rov}	28.0398	31.3968	39.1059	51.8615	56.1096	64.0216	91.7910	96.6377	106.7405	130.9571	136.8497	149.0311	169.7471	176.7317	189.9574
0.3	\mathcal{I}_f^2	5.9254	9.1432	20.5153	4.5002	6.1094	12.1199	3.9475	5.2469	8.8495	3.8388	4.9152	8.0666	3.5867	4.7699	7.5952
	\mathcal{I}_r^2	5.0891	7.6901	18.6598	3.8142	5.1214	9.3764	3.4875	4.5081	7.4094	3.3661	4.3707	6.8974	3.2648	4.2446	6.7544
	$\mathcal{I}_{r,Rov}^2$	28.8410	32.6879	45.1575	51.6131	55.6875	63.9078	91.5705	96.5421	106.9015	130.9944	136.9727	147.6272	169.2115	176.0893	189.6368
	\max	7.1295	10.4465	22.2404	5.2032	7.0183	12.3779	4.6913	6.0781	9.5347	4.4958	5.6173	8.8910	4.3468	5.5664	8.4205
	\max_{Rov}	28.8456	32.7246	45.1594	51.6131	55.6875	63.9078	91.5705	96.5421	106.9015	130.9944	136.9727	147.6272	169.2115	176.0893	189.6368
0.4	\mathcal{I}_f^2	9.3660	14.7792	37.4898	6.9386	10.3876	21.6446	5.8145	8.4229	15.9782	5.4762	7.8418	15.1374	5.1136	7.2900	14.1477
	\mathcal{I}_r^2	6.2298	10.2659	30.1237	4.4786	6.3220	12.4240	3.6875	5.0095	8.5910	3.7426	4.8359	7.8464	3.5391	4.5784	7.3455
	$\mathcal{I}_{r,Rov}^2$	29.8567	34.5911	57.2075	51.4612	55.6758	65.1931	90.5328	95.5496	105.4845	129.4079	135.9307	147.7036	167.5357	174.8160	188.8530
	\max	10.0885	15.5162	38.6559	7.4819	10.6893	22.1677	6.2991	8.6391	16.1166	5.9676	8.1333	15.1924	5.6921	7.6422	14.2758
	\max_{Rov}	29.9126	34.6369	57.6775	51.4612	55.6758	65.1931	90.5328	95.5496	105.4845	129.4079	135.9307	147.7036	167.5357	174.8160	188.8530
0.45	\mathcal{I}_f^2	18.1006	28.9274	68.2023	12.9367	20.5551	44.3951	11.6559	17.4198	35.5686	10.7268	16.6459	32.7118	9.7072	14.8553	29.7083
	\mathcal{I}_r^2	10.0791	17.6208	51.8608	5.9384	8.9878	21.9174	4.7358	6.6593	14.2653	4.4294	5.9682	11.5241	4.1618	5.6104	9.8223
	$\mathcal{I}_{r,Rov}^2$	33.2386	41.5676	86.1946	52.1573	56.6238	67.6016	90.1455	95.3582	106.1175	128.4825	134.1389	146.3741	166.4256	173.1637	186.7888
	\max	18.4792	29.2106	70.3861	13.0874	20.6285	44.3951	11.7449	17.4323	35.5686	10.8017	16.6459	32.7118	9.8084	14.9258	29.7083
	\max_{Rov}	33.6284	42.5402	86.8162	52.2006	56.7065	68.4957	90.1455	95.3582	106.1175	128.4825	134.1389	146.3741	166.4256	173.1637	186.7888
0.49	\mathcal{I}_f^2	80.6167	125.4279	241.3720	68.3766	107.4526	211.8803	67.6545	105.3709	189.1685	65.4880	97.3614	198.1288	63.2042	96.3603	184.9993
	\mathcal{I}_r^2	45.6690	84.7404	234.6658	27.0425	45.6257	112.8362	19.7859	33.7583	75.9676	16.0963	27.7489	66.7492	14.1441	24.1794	56.3271
	$\mathcal{I}_{r,Rov}^2$	73.9508	122.7369	303.5430	68.6736	90.2783	170.5367	96.5149	107.5351	155.4517	129.4071	137.4622	171.8239	164.3723	171.5966	193.9100
	\max	80.7141	125.7580	253.2365	68.3766	107.4526	211.8803	67.6545	105.3709	189.1685	65.4880	97.3614	198.1288	63.2042	96.3603	184.9993
	\max_{Rov}	86.5939	135.3438	305.2113	75.8774	110.7574	212.3735	98.2158	114.6577	189.1686	129.8624	139.0959	198.9826	164.7096	172.3484	199.7330

3.2 Size e potenza dei test

Con il proposito di valutare *size* e potenza delle statistiche proposte in questo lavoro in un contesto in cui l'ordine di integrazione d è sconosciuto si considerino i dati considerati dai seguenti DGP:

$$\begin{aligned} (1-L)^{d_0}y_t &= \varepsilon_t & , & & t = 1, \dots, [\tau^*N] \\ (1-L)^{d_1}y_t &= \varepsilon_t & , & & t = [\tau^*N] + 1, \dots, N \end{aligned} \quad (3.1)$$

dove $\tau^* = 0.5$, $d_0 = \{0, 0.3, 0.49\}$, $d_1 = \{0, 0.1, 0.2, 0.3, 0.4, 0.45, 0.49\}$, $y_t = 0$ per $t \leq 0$ e $\varepsilon_t \sim n.i.d(0, 1)$. La lunghezza delle serie storiche considerate nelle simulazioni è $N = 250$ e $N = 500$. L'ordine di integrazione è stato stimato attraverso il metodo di [Fox and Taqqu, 1986].

Risulta opportuno richiamare la nozione di funzione di potenza. La funzione di potenza di un test con regione di rifiuto R è

$$\pi(\theta) = Pr_{\theta}(Y \in R), \quad \theta \in \Theta \quad (3.2)$$

e permette di ottenere la probabilità con cui si può incorrere nei due tipi di errore: l'errore del primo tipo che si definisce come la probabilità di rifiutare l'ipotesi nulla quando è vera e quello del secondo tipo che si definisce come la probabilità di accettare l'ipotesi nulla quando è falsa. Indicando con $\alpha(\cdot)$ la probabilità d'errore del primo tipo e con $\beta(\cdot)$ la probabilità d'errore del secondo tipo, si ha infatti $\pi(\theta) = \alpha(\theta)$ se $\theta \in \Theta_0$ e $1 - \pi(\theta) = \beta(\theta)$ se $\theta \in \Theta \setminus \Theta_0$. Il test ideale non è costruibile, salvo che in casi banali.

Le tabelle 3.3 e 3.4 presentano la *size* e la potenza empiriche dei test per le due diverse numerosità campionarie, $N = 250$ e $N = 500$. Data la natura dell'ipotesi nulla ($d_0 = d_1$) e le alternative considerate ($d_0 \neq d_1$), ci aspettiamo che le statistiche dipendano dalle due alternative: $d_0 < d_1$ e da $d_0 > d_1$.

3.2.1 *Size*

Nello specifico, per il calcolo della *size* del test abbiamo creato, per ciascun valore di d_1 , il corrispondente vettore contenente i valori critici dei vari \mathcal{T} ad un livello di significatività del 5% e abbiamo inizializzato, ponendoli pari a zero, dieci contatori relativi ai vari \mathcal{T} . In seguito, dopo aver simulato la serie e stimato il valore del parametro d , abbiamo selezionato gli appropriati valori critici da utilizzare in ciascuna iterazione tenendo in considerazione il valore assunto dalla stima del parametro d . Ovvero, qualora $d \in [0, 0.05]$ abbiamo selezionato i valori critici corrispondenti a $d_0 = 0$, qualora $d \in (0.05, 0.15]$ abbiamo selezionato i valori critici corrispondenti a $d_0 = 0.1$, e così via fino al caso in cui $d > 0.45$ per il quale abbiamo selezionato i valori critici corrispondenti a $d_0 = 0.49$. Successivamente, dopo aver calcolato i valori della statistica t sulle tre diverse regressioni impostate, abbiamo individuato i minimi e i massimi necessari, seguendo quanto proposto nella parte teorica. Per concludere, per valutare quanti \mathcal{T} cadessero nella zona di rifiuto, abbiamo confrontato il valore di ogni \mathcal{T} con il corrispettivo valore critico: per valori piccoli di \mathcal{T} la zona di rifiuto è costituita da tutti i \mathcal{T} minori dei rispettivi valori critici, mentre per valori grandi di \mathcal{T} la regione di rifiuto è costituita da tutti i \mathcal{T} maggiori dei rispettivi valori critici. A ciascuna iterazione i contatori inseriti precedentemente, qualora il valore dei diversi \mathcal{T} cadesse nella zona di rifiuto, hanno subito un incremento unitario.

3.2.2 *Potenza*

Per il calcolo della potenza del test, invece, abbiamo simulato due diverse serie entrambe di lunghezza pari a $N/2$ ($\tau^* = 0.5$): in una delle due abbiamo impostato come valore per il parametro d il valore assunto da d_0 , mentre nell'altra il valore assunto da d_1 . In seguito, dopo aver creato con la funzione `ts` un oggetto di formato *time series*, composto da entrambe le serie che abbiamo simulato abbiamo seguito lo stesso procedimento utilizzato per calcolare la *size* del test. Quindi, anche in questo caso, per la selezione dei valori critici da utilizzare in ciascuna iterazione abbiamo tenuto in considerazione il valore assunto dalla stima del parametro d e per la valutazione del

numero di \mathcal{T} appartenenti alla zona di rifiuto abbiamo confrontato il valore di questi ultimi con i rispettivi valori critici.

3.3 Considerazioni

Dai risultati riportati nelle tabelle 3.3 e 3.4, si può constatare che la *size* effettiva stimata (ovvero il livello di significatività effettivo stimato) risulta essere in linea rispetto al valore nominale, o teorico, del 5% per entrambe le numerosità campionarie. Emerge però, per \mathcal{T}_f^2 , \mathcal{T}_r^2 , $\mathcal{T}_{r,Rov}^2$, $\max(\mathcal{T}_f^2, \mathcal{T}_r^2)$ e $\max(\mathcal{T}_f^2, \mathcal{T}_{r,Rov}^2)$, una distorsione quando d_0 e d_1 sono pari a 0.49 (che potrebbe essere dovuto al fatto che il valore 0.49 è molto vicino al limite di non stazionarietà) leggermente più accentuata per serie storiche di numerosità campionaria più limitata ($N = 250$). Entrambe le tabelle ci mostrano la frazione di valori che rifiutiamo sotto l'ipotesi nulla a un livello nominale del 5%, la quale approssima in modo consistente la probabilità di rifiutare H_0 , condizionatamente al fatto che H_0 sia vera, dove H_0 indica la costanza di d nel campione considerato.

Per quanto riguarda la potenza dei test effettuati ci si aspetta che cresca man mano che ci si allontana dall'ipotesi nulla, ovvero all'aumentare della diversità del valore di d sotto l'ipotesi nulla e sotto l'ipotesi alternativa.

Quando l'ipotesi nulla è $d_0 = 0$, sia per $N = 250$ che per $N = 500$, si nota che effettivamente la potenza aumenta all'aumentare di d_1 , anche se \mathcal{T}_r ha un andamento leggermente anomalo in quanto presenta sempre valori molto bassi, e l'andamento crescente non è ben evidente. In questo caso sembrerebbe che il comportamento di $\mathcal{T}_{r,Rov}$ sia migliore di \mathcal{T}_r , nonostante anche la potenza di $\mathcal{T}_{r,Rov}$ aumenti molto lentamente. Tra tutti i diversi \mathcal{T} , quando $d_0 = 0$, quello migliore risulta essere \mathcal{T}_f , che converge a 1 più velocemente rispetto agli altri, e ciò risulta ancor più evidente per $N = 500$. Ciò potrebbe significare che quando l'alternativa considerata è $d_0 < d_1$ il test più opportuno da utilizzare potrebbe essere \mathcal{T}_f .

Quando l'ipotesi nulla è $d_0 = 0.3$, per entrambe le numerosità, ci si aspetta che la potenza segua un andamento crescente all'aumentare e al diminuire di d_1 . Stando ai risultati che abbiamo ottenuto, tranne per qualche valore,

Tabella 3.3: Size empirica e potenza dei test per i cambiamenti di persistenza ($N = 250$)

d_0	d_1	\mathcal{I}_f	\mathcal{I}_r	$\mathcal{I}_{r,Rov}$	min	min _{Rov}	\mathcal{I}_f^2	\mathcal{I}_r^2	$\mathcal{I}_{r,Rov}^2$	max	max _{Rov}
0	0	0.0477	0.0497	0.0492	0.0481	0.0477	0.0446	0.0495	0.0454	0.0468	0.0454
	0.1	0.1209	0.0106	0.0209	0.0677	0.1209	0.0756	0.055	0.1016	0.0607	0.1016
	0.2	0.4127	0.0046	0.0116	0.2828	0.4127	0.2592	0.2018	0.1423	0.2506	0.1423
	0.3	0.7673	0.0044	0.0085	0.6452	0.7673	0.5738	0.4239	0.1877	0.5456	0.1877
	0.4	0.9573	0.0035	0.0093	0.912	0.9573	0.8397	0.6402	0.2281	0.8147	0.2281
	0.45	0.9877	0.0048	0.0144	0.9713	0.9877	0.813	0.6273	0.2195	0.7979	0.2196
	0.49	0.9982	0.055	0.0966	0.9948	0.9982	0.4104	0.3216	0.1202	0.4052	0.1207
0.3	0	0.0021	0.7626	0.3976	0.6638	0.0024	0.4297	0.63	0.0053	0.5826	0.0053
	0.1	0.0064	0.4616	0.2273	0.3459	0.0064	0.1993	0.2904	0.0114	0.2536	0.0114
	0.2	0.0178	0.189	0.1116	0.1262	0.0178	0.0787	0.088	0.0218	0.0865	0.0218
	0.3	0.0548	0.0545	0.051	0.0545	0.0548	0.0584	0.0511	0.0494	0.0571	0.0494
	0.4	0.2964	0.0255	0.0392	0.1853	0.2964	0.302	0.0473	0.0573	0.0317	0.0573
	0.45	0.5665	0.0274	0.0439	0.4177	0.5664	0.0324	0.0444	0.0473	0.0358	0.0473
	0.49	0.8478	0.1253	0.1583	0.7547	0.8477	0.0177	0.0255	0.0239	0.0187	0.0239
0.49	0	0.0012	0.9972	0.889	0.9929	0.0922	0.2297	0.4275	0	0.3857	0
	0.1	0.0016	0.9799	0.8035	0.9645	0.0499	0.118	0.3339	0.0003	0.2477	0.0003
	0.2	0.0021	0.9127	0.675	0.8754	0.0261	0.0435	0.176	0.0012	0.0922	0.0012
	0.3	0.0063	0.7733	0.5205	0.7236	0.0176	0.0189	0.0516	0.0029	0.0261	0.0029
	0.4	0.0334	0.5785	0.3709	0.5275	0.036	0.0183	0.0108	0.0043	0.0181	0.0043
	0.45	0.1006	0.4769	0.3104	0.4524	0.102	0.0227	0.0115	0.0061	0.0218	0.0065
	0.49	0.0429	0.0479	0.0513	0.0475	0.0427	0.1846	0.1581	0.1152	0.1838	0.1205

Tabella 3.4: Size empirica e potenza dei test per i cambiamenti di persistenza ($N = 500$)

d_0	d_1	\mathcal{F}_f	\mathcal{F}_r	$\mathcal{F}_{r,Rov}$	min	min $_{Rov}$	\mathcal{F}_f^2	\mathcal{F}_r^2	$\mathcal{F}_{r,Rov}^2$	max	max $_{Rov}$
0	0	0.0514	0.0508	0.0466	0.0498	0.0514	0.0504	0.0508	0.0484	0.0495	0.0484
	0.1	0.2147	0.0041	0.0095	0.1297	0.2147	0.1485	0.1038	0.1148	0.1239	0.1148
	0.2	0.7255	0.0030	0.0076	0.5944	0.7255	0.5909	0.4706	0.1524	0.5514	0.1524
	0.3	0.9701	0.0008	0.0029	0.9429	0.9701	0.9354	0.8596	0.2206	0.9175	0.2206
	0.4	0.9995	0.0013	0.0042	0.9988	0.9995	0.9975	0.9876	0.2824	0.9960	0.2824
0.3	0.45	0.9999	0.0017	0.0065	0.9996	0.9999	0.9538	0.9516	0.2700	0.9537	0.2700
	0.49	0.9999	0.0332	0.0753	0.9999	0.9999	0.5326	0.5167	0.1596	0.5326	0.1657
	0	0.0014	0.9703	0.5986	0.9435	0.0014	0.8629	0.9384	0.0024	0.9181	0.0024
	0.1	0.0020	0.7458	0.3401	0.6520	0.0020	0.4698	0.6297	0.0055	0.5734	0.0055
	0.2	0.0084	0.3087	0.1460	0.2110	0.0084	0.1316	0.1924	0.0206	0.1620	0.0206
0.49	0.3	0.0486	0.0518	0.0455	0.0503	0.0486	0.0570	0.0509	0.0500	0.0521	0.0500
	0.4	0.3907	0.0157	0.0282	0.2606	0.3907	0.0820	0.0994	0.0717	0.0815	0.0716
	0.45	0.7113	0.0156	0.0331	0.5770	0.7113	0.1358	0.1609	0.0681	0.1377	0.0678
	0.49	0.9327	0.1100	0.1402	0.8813	0.9327	0.0795	0.0933	0.0352	0.0808	0.0344
	0	0.0001	1.0000	0.9802	1.0000	0.0635	0.5111	0.8029	0.0000	0.5216	0.0000
0.49	0.1	0.0002	0.9998	0.9241	0.9987	0.0237	0.3761	0.6087	0.0000	0.4515	0.0000
	0.2	0.0008	0.9884	0.8107	0.9781	0.0072	0.1748	0.4184	0.0005	0.3032	0.0005
	0.3	0.0027	0.9018	0.6277	0.8650	0.0042	0.0488	0.1807	0.0026	0.0912	0.0026
	0.4	0.0231	0.6841	0.4196	0.6364	0.0235	0.0188	0.0222	0.0034	0.0214	0.0030
	0.45	0.1044	0.5421	0.3334	0.5191	0.1045	0.0183	0.0105	0.0044	0.0185	0.0041
0.49	0.0442	0.0530	0.0516	0.0505	0.0442	0.1310	0.1108	0.0632	0.1312	0.0636	

si verifica quanto appena enunciato. Tuttavia, nella maggior parte dei casi l'aumento risulta più evidente in una sola direzione, ossia in alcuni casi l'aumento della potenza è più accentuato all'aumentare di d_1 piuttosto che al diminuire di quest'ultimo e in altri casi invece avviene il contrario. Quanto appena detto vale per entrambe le numerosità, $N = 250$ e $N = 500$. L'unico che presenta un sostanziale aumento in entrambe le direzioni è $\min(\mathcal{I}_f, \mathcal{I}_r)$, aumento che è più accentuato per serie di numerosità pari a 500. Da questa considerazione segue che, quando non si conosce l'ipotesi alternativa da considerare, ovvero non è noto se l'ipotesi alternativa sia $d_0 < d_1$ o $d_0 > d_1$, il test da preferire sembrerebbe essere proprio $\min(\mathcal{I}_f, \mathcal{I}_r)$. Quanto appena detto conferma ciò che già era stato proposto nel capitolo 2.3, nel quale si raccomandava appunto l'utilizzo di $\min(\mathcal{I}_f, \mathcal{I}_r)$ quando la direzione del cambiamento era sconosciuta. Nel caso in cui la direzione del cambiamento sia nota, sembrerebbe raccomandabile l'utilizzo di \mathcal{I}_r qualora l'ipotesi alternativa sia $d_0 > d_1$ e l'utilizzo di \mathcal{I}_f qualora l'ipotesi alternativa sia $d_0 < d_1$.

Quando invece l'ipotesi nulla è $d_0 = 0.49$ il comportamento rispecchia quanto dovrebbe accadere, ossia un aumento della potenza man mano che d_1 diminuisce, tranne che per $\mathcal{I}_{r,Rov}^2$ e $\max(\mathcal{I}_f^2, \mathcal{I}_{r,Rov}^2)$ i quali dimostrano un comportamento inverso, infatti la potenza, oltre ad essere molto bassa, diminuisce anziché aumentare. Anche \mathcal{I}_f , che sotto l'ipotesi nulla $d_0 = 0$ sembrava essere il migliore, è caratterizzato da valori molto bassi che decrescono man mano che ci si allontana dall'ipotesi nulla. In questo caso il test che presenta il comportamento migliore è \mathcal{I}_r , il quale converge a 1 più velocemente rispetto a tutti gli altri. Ancora una volta la convergenza è più rapida per serie di numerosità pari a 500. Ciò conferma quanto si è appena visto, ossia che quando l'alternativa da considerare è $d_0 > d_1$, e quindi la direzione del cambiamento è nota, il test da preferire sembrerebbe appunto essere \mathcal{I}_r .

Volendo trarre delle considerazioni a livello generale, la cosa più evidente è che, come ci si aspettava, i test sembrano comportarsi meglio considerando serie di numerosità più elevata. Inoltre, per entrambe le numerosità considerate, sotto le tre diverse ipotesi nulle, i test che manifestano il comportamento

peggiore sono $\mathcal{I}_{r,Rov}^2$ e $\max(\mathcal{I}_f^2, \mathcal{I}_{r,Rov}^2)$ per i quali la potenza mantiene sempre valori molto bassi. Emerge anche che, a seconda che si conosca o meno la direzione del cambiamento i test preferibili cambiano: se si conosce l'alternativa da considerare, quando l'ipotesi alternativa è $d_0 > d_1$ il test che risulta essere migliore è \mathcal{I}_r , al contrario quando l'ipotesi alternativa è $d_0 < d_1$ il test che risulta essere migliore è \mathcal{I}_f ; mentre, se non si conosce la direzione del cambiamento il test da prediligere risulta essere $\min(\mathcal{I}_f, \mathcal{I}_r)$.

I risultati derivanti dalle simulazioni realizzate sono in linea con quelli ottenuti nell'articolo di riferimento [Martins and Rodrigues, 2014]. È chiaro che per i test $\mathcal{I}_{r,Rov}$, $\min(\mathcal{I}_f, \mathcal{I}_{r,Rov})$, $\mathcal{I}_{r,Rov}^2$ e $\max(\mathcal{I}_f^2, \mathcal{I}_{r,Rov}^2)$ non possiamo avere un riscontro nell'articolo, poiché si tratta di test aggiuntivi non presenti nel lavoro di [Martins and Rodrigues, 2014] per le ragioni già esposte nella sezione 3.1.

Conclusioni

L'obiettivo di questo lavoro è quello di identificare dei test in grado di individuare cambiamenti di persistenza nei modelli a differenze frazionarie. Per introdurre l'argomento di riferimento, è stato presentato il concetto di memoria lunga e sono state definite le caratteristiche dei modelli ARFIMA(p,d,q). In seguito abbiamo delineato gli opportuni test per scoprire eventuali cambiamenti nel parametro di integrazione frazionaria e, con l'ausilio del software statistico R, abbiamo realizzato le necessarie simulazioni volte ad indagare la *size* e la potenza dei test introdotti.

Nella sezione 2.3 abbiamo visto come vi sia una sostanziale modifica nella selezione dei test da utilizzare qualora si disponga o meno di indicazioni sulla direzione del cambiamento. Dopo un'analisi dei risultati ottenuti è emerso, a conferma di quanto appena richiamato, che i test da preferire cambiano in relazione alle informazioni che abbiamo a disposizione. In particolare, se non si conosce la direzione del cambiamento, ovvero non è noto se l'alternativa da considerare sia $d_0 < d_1$ o $d_0 > d_1$, è raccomandabile l'utilizzo del test $\min(\mathcal{T}_f, \mathcal{T}_r)$. Se invece si è a conoscenza dell'alternativa da considerare i test che risultano preferibili sono: \mathcal{T}_r se l'alternativa sotto esame è $d_0 > d_1$ e \mathcal{T}_f se l'alternativa sotto esame è $d_0 < d_1$. Sempre nella sezione 2.3 si erano introdotti anche i quadrati delle statistiche t che presentavano il vantaggio di non dover considerare il segno della differenza tra d_0 e d_1 . Tuttavia tali test sembrano essere meno performanti soprattutto per la numerosità $N = 250$. Ci si aspetta infatti che la potenza converga a 1 ma per la numerosità appena menzionata la potenza mantiene valori bassi, salvo che per il caso in cui l'ipotesi nulla è $d_0 = 0$. Situazione che però migliora quando si considerano serie di lunghezza maggiore.

Dando uno sguardo generale a tutti i test realizzati, si evince che la potenza aumenta all'aumentare della numerosità delle serie considerate, ossia aumenta all'aumentare di N , e ciò sta ad indicare la consistenza dei test.

Volendo fare un confronto tra questo lavoro e l'articolo di riferimento [Martins and Rodrigues, 2014] emerge che i risultati relativi alla *size* e alla potenza dei test ottenuti attraverso le simulazioni realizzate sono in linea con quelli proposti nell'articolo.

Appendice A

Codice sorgente R

```
## 1) CALCOLO DEI VALORI CRITICI
```

```
# Calcolo i residui di un modello FI(d)
arfima.res<-function(y, d)
{
  t=length(y)
  ar.coef<-array(1,(t-1))
  residui<-array(1,t)
  residui[1]<-y[1]
  coef<-1;
  for ( k in 1:(t-1))
    {
      coef<-coef*(k-1-d)/k
      ar.coef[k]<-coef
      residui[k+1]<-y[k+1]+ar.coef[1:k]%*%y[k:1]
    }
  return(residui)
}
```

```
# Stimo d mediante il metodo di Fox e Taqqu(1986). Si veda il capitolo
2.4.
#funzione per stimare d
whittle<-function(x)
{
  result <- list()
```

```

t <- length(x)
FFT<-Mod(fft(x))^2/(2*pi*t)
period <- FFT[2:(t%/2)]
spec <- spettro(t)
q<-function(d) t(period)%*%spec^(2*d)
min<-optimize(q,c(0,1))
dhat<-min$minimum
sigma2 <- q(dhat)/(2*(2*pi)^2)
result$d <- dhat
result$sigma <- sigma2
return(result)
}

# Calcolo lo spettro
spettro<-function(t)
{
tt<-trunc((t-1)/2)
y<-array (1,tt)
i<-c(1:t)
for (j in 1:tt)
{
y [j]<- abs(2*sin(pi*j/t))
}
return(y)
}

# Simulazione serie FI(d)
library(fArma)      #richiamo la libreria
S=10000             #numero di simulazioni
N=...               #lunghezza della serie storica.
    N=100,250,500,750,1000
d0=...
#valore del parametro d0. d0= 0,0.1,0.2,0.3,0.4,0.45,0.49
Tau_f=NULL
Tau_r=NULL
Tau_rRov=NULL      #Tau_r calcolato sulla serie invertita nel tempo
Tau_min=NULL       #valore minimo tra Tau_f e Tau_r

```

Capitolo A. Codice sorgente R

```
Tau_minrov=NULL      #valore minimo tra Tau_f e Tau_rRov

#I seguenti Tau sono quelli elevati al quadrato
Tau_f2=NULL
Tau_r2=NULL
Tau_rRov2=NULL       #(Tau_r)^2 calcolato sulla serie invertita nel tempo
Tau_max=NULL        #valore massimo tra Tau_f2 e Tau_r2
Tau_maxrov=NULL     #valore massimo tra Tau_f2 e Tau_rRov2
tau = seq(0.2, 0.8, by=0.05)
l_tau=length(tau)

for(i in 1:S){
y = armaSim(model = list(ar = 0, d = d0, ma = 0), n = N)
d_stima=whittle(y)$d
x = arfima.res(y, d_stima)
test1=NULL
test2=NULL
test3=NULL
test4=NULL
test5=NULL
test6=NULL

for(j in 1:l_tau){      #ciclo 2
x_star=NULL
N_tau=trunc(tau[j]*N)
indice1= c(1:(N-1))
for (l in 1:(N-1)){    #ciclo3
x_star[l]=sum(x[l:1]/(indice1[1:l])) } #chiude il ciclo 3

mod1=lm(x[2:N_tau] ~ -1 + x_star[1:(N_tau-1)])
#modello utilizzato per ricavare Tau_f

mod2=lm(x[(N-1):(N_tau+1)] ~ -1 + x_star[N:(N_tau+2)])
#modello utilizzato per ricavare Tau_rRov

mod3=lm(x[(N_tau+2):N] ~ -1 + x_star[(N_tau+1):(N-1)])
#modello utilizzato per ricavare Tau_r

test1[j]=summary(mod1)$coef[,3]
```

```

test2[j]=summary(mod2)$coef[,3]
test3[j]=summary(mod3)$coef[,3]
test4[j]=(summary(mod1)$coef[,3])^2
test5[j]=(summary(mod2)$coef[,3])^2  #test per i quadrati
test6[j]=(summary(mod3)$coef[,3])^2
} #chiude il ciclo 2

#calcolo dei test
Tau_f[i] = min(test1)
Tau_rRov[i] = min(test2)
Tau_r[i] = min(test3)
Tau_f2[i] = max(test4)
Tau_rRov2[i] = max(test5)
Tau_r2[i] = max(test6)
Tau_min[i]=min(Tau_f[i],Tau_r[i])
Tau_minrov[i]= min(Tau_f[i],Tau_rRov[i])
Tau_max[i]=max(Tau_f2[i],Tau_r2[i])
Tau_maxrov[i]= max(Tau_f2[i],Tau_rRov2[i])
} #chiude il ciclo grande

#calcolo i valori critici
vc_test1=quantile(Tau_f, probs=c(0.01, 0.05, 0.1))
vc_test2=quantile(Tau_rRov, probs=c(0.01, 0.05, 0.1))
vc_test3=quantile(Tau_r, probs=c(0.01, 0.05, 0.1))
vc_test4=quantile(Tau_f2, probs=c(0.9, 0.95, 0.99))
vc_test5=quantile(Tau_rRov2, probs=c(0.9, 0.95, 0.99))
vc_test6=quantile(Tau_r2, probs=c(0.9, 0.95, 0.99))
vc_test7=quantile(Tau_min, probs=c(0.01, 0.05, 0.1))
vc_test8=quantile(Tau_minrov, probs=c(0.01, 0.05, 0.1))
vc_test9=quantile(Tau_max, probs=c(0.9, 0.95, 0.99))
vc_test10=quantile(Tau_maxrov, probs=c(0.9, 0.95, 0.99))

vc_test1
vc_test2
vc_test3
vc_test4
vc_test5
vc_test6
vc_test7

```

Capitolo A. Codice sorgente R

```
vc_test8
vc_test9
vc_test10
```

```
## 2) CALCOLO SIZE DEL TEST PER N=250
```

```
library(fArma)
#numero di simulazioni
S=10000
#lunghezza della serie storica
N=250
#valore del parametro d0. Per il calcolo della size d0= 0.0, 0.3, 0.49.
d0=...

Tau_f=NULL
Tau_r=NULL
Tau_rRov=NULL
Tau_min=NULL
Tau_minrov=NULL
Tau_f2=NULL
Tau_r2=NULL
Tau_rRov2=NULL
Tau_max=NULL
Tau_maxrov=NULL

#sostituisco i valori critici trovati precedentemente al livello del 5%,
  N=250
vc_0=c(-2.390345, -2.369255, 1.0978154, -2.664596, -2.390345, 5.912636,
        5.747492, 52.45617, 7.197152, 52.45617)
vc_01=c(-1.840509, -1.865043, 1.3772163, -2.152303, -1.840509, 4.872031,
        4.692537, 55.96952, 5.833572, 55.96952)
vc_02=c(-1.771504, -1.860679, 1.3653231, -2.095371, -1.771504, 5.226095,
        4.938591, 56.10956, 6.146754, 56.10956)
vc_03=c(-1.776532, -1.834094, 1.3655048, -2.095557, -1.776532, 6.109423,
        5.12141, 55.68748, 7.018259, 55.68748)
vc_04=c(-1.662569, -1.89305, 1.3605895, -2.072972, -1.662569, 10.387606,
        6.32198, 55.67582, 10.68928, 55.67582)
```

```
vc_045=c(-1.616618, -1.902328, 1.3244899, -2.079666, -1.616618, 20.55505,  
8.987765, 56.62381, 20.62849, 56.70649)  
vc_049=c(-1.2287803, -1.721612, 1.3970773, -1.8677, -1.2304207,  
107.45257, 45.62574, 90.27827, 107.45257, 110.75742)
```

```
#inserisco un contatore
```

```
cont1=0  
cont2=0  
cont3=0  
cont4=0  
cont5=0  
cont6=0  
cont7=0  
cont8=0  
cont9=0  
cont10=0
```

```
#simulazione serie FI(d)
```

```
for(i in 1:S){  
y = armaSim(model = list(ar = 0, d = d0, ma = 0), n = N)  
d_stima=whittle(y)$d  
x = arfima.res(y, d_stima)
```

```
if(d_stima>=0 && d_stima<= 0.05) {vc=vc_0}  
if(d_stima>0.05 && d_stima<=0.15) {vc=vc_01}  
if(d_stima>0.15 && d_stima<=0.25) {vc=vc_02}  
if(d_stima>0.25 && d_stima<=0.35) {vc=vc_03}  
if(d_stima>0.35 && d_stima<0.45) {vc=vc_04}  
if(d_stima==0.45) {vc=vc_045}  
if(d_stima>0.45) {vc=vc_049}
```

```
#I valori critici che devono essere utilizzati ad ogni iterazione  
dipendono dal valore di d stimato
```

```
for(j in 1:l_tau){ #ciclo 2  
x_star=NULL  
N_tau=trunc(tau[j]*N)  
indice1= c(1:(N-1))
```

```
for (l in 1:(N-1)){
x_star[l]=sum(x[1:l]/(indice1[1:l])) }

mod1=lm(x[2:N_tau] ~ -1 + x_star[1:(N_tau-1)])
mod2=lm(x[(N-1):(N_tau+1)] ~ -1 + x_star[N:(N_tau+2)])
mod3=lm(x[(N_tau+2):N] ~ -1 + x_star[(N_tau+1):(N-1)])

test1[j]=summary(mod1)$coef[,3]
test2[j]=summary(mod2)$coef[,3]
test3[j]=summary(mod3)$coef[,3]
test4[j]=(summary(mod1)$coef[,3])^2
test5[j]=(summary(mod2)$coef[,3])^2
test6[j]=(summary(mod3)$coef[,3])^2

} #chiude il ciclo 2

#calcolo dei test
Tau_f = min(test1)
Tau_rRov = min(test2)
Tau_r = min(test3)
Tau_f2=max(test4)
Tau_rRov2=max(test5)
Tau_r2=max(test6)
Tau_min=min(Tau_f,Tau_r)
Tau_minrov=min(Tau_f, Tau_rRov)
Tau_max=max(Tau_f2,Tau_r2)
Tau_maxrov=max(Tau_f2,Tau_rRov2)

if (Tau_f < vc[1]) {cont1=cont1+1}
if (Tau_r < vc[2]) {cont2=cont2+1}
if (Tau_rRov < vc[3]) {cont3=cont3+1}
if(Tau_min<vc[4]) {cont4=cont4+1}
if(Tau_minrov<vc[5]) {cont5=cont5+1}
if (Tau_f2 > vc[6]) {cont6=cont6+1}
if (Tau_r2 > vc[7]) {cont7=cont7+1}
if (Tau_rRov2 > vc[8]) {cont8=cont8+1}
if(Tau_max>vc[9]) {cont9=cont9+1}
```

```
if(Tau_maxrov>vc[10]) {cont10=cont10+1}  
} #chiude il ciclo grande
```

```
print(cont1/S)  
print(cont2/S)  
print(cont3/S)  
print(cont4/S)  
print(cont5/S)  
print(cont6/S)  
print(cont7/S)  
print(cont8/S)  
print(cont9/S)  
print(cont10/S)
```

```
## 3) CALCOLO LA POTENZA DEL TEST PER N=250
```

```
library(fArma)  
#numero di simulazioni  
S=10000  
#lunghezza della serie storica  
N=250  
  
#valore del parametro d  
d0=... #d0 assume i valori: 0.0, 0.3, 0.49  
d1=... #d1 assume i valori: 0.0, 0.1, 0.2, 0.3, 0.4, 0.45, 0.49  
  
Tau_rRov=NULL  
Tau_f=NULL  
Tau_r=NULL  
Tau_rRov2=NULL  
Tau_f2=NULL  
Tau_r2=NULL  
Tau_min=NULL  
Tau_minrov=NULL  
Tau_max=NULL  
Tau_maxrov=NULL
```



```
cont1=0
cont2=0
cont3=0
cont4=0
cont5=0
cont6=0
cont7=0
cont8=0
cont9=0
cont10=0

#simulazione serie FI(d)
for(i in 1:S){
y1 = armaSim(model = list(ar = 0, d = d0, ma = 0), n = N/2)
y2 = armaSim(model = list(ar = 0, d = d1, ma = 0), n = N/2)
y=ts(c(y1,y2))

d_stima=whittle(y)$d
x = arfima.res(y, d_stima)

if(d_stima>=0 && d_stima<= 0.05) {vc=vc_0}
if(d_stima>0.05 && d_stima<=0.15) {vc=vc_01}
if(d_stima>0.15 && d_stima<=0.25) {vc=vc_02}
if(d_stima>0.25 && d_stima<=0.35) {vc=vc_03}
if(d_stima>0.35 && d_stima<0.45) {vc=vc_04}
if(d_stima==0.45) {vc=vc_045}
if(d_stima>0.45) {vc=vc_049}

for(j in 1:l_tau){
x_star=NULL
N_tau=trunc(tau[j]*N)
indice1= c(1:(N-1))
for (l in 1:(N-1)){
x_star[l]=sum(x[1:l]/(indice1[1:l])) } #chiude ciclo 3

mod1=lm(x[2:N_tau] ~ -1 + x_star[1:(N_tau-1)])
mod2=lm(x[(N-1):(N_tau+1)] ~ -1 + x_star[N:(N_tau+2)])
```

```
mod3=lm(x[(N_tau+2):N] ~ -1 + x_star[(N_tau+1):(N-1)])
```

```
test1[j]=summary(mod1)$coef[,3]
test2[j]=summary(mod2)$coef[,3]
test3[j]=summary(mod3)$coef[,3]
test4[j]=(summary(mod1)$coef[,3])^2
test5[j]=(summary(mod2)$coef[,3])^2
test6[j]=(summary(mod3)$coef[,3])^2
} #chiude il ciclo 2
```

```
#calcolo dei test
```

```
Tau_f = min(test1)
Tau_rRov = min(test2)
Tau_r = min(test3)
Tau_f2=max(test4)
Tau_rRov2=max(test5)
Tau_r2=max(test6)
Tau_min=min(Tau_f,Tau_r)
Tau_minrov=min(Tau_f, Tau_rRov)
Tau_max=max(Tau_f2,Tau_r2)
Tau_maxrov=max(Tau_f2,Tau_rRov2)
```

```
if (Tau_f < vc[1]) {cont1=cont1+1}
if (Tau_r < vc[2]) {cont2=cont2+1}
if (Tau_rRov < vc[3]) {cont3=cont3+1}
if (Tau_min<vc[4]) {cont4=cont4+1}
if (Tau_minrov<vc[5]) {cont5=cont5+1}
if (Tau_f2 > vc[6]) {cont6=cont6+1}
if (Tau_r2 > vc[7]) {cont7=cont7+1}
if (Tau_rRov2 > vc[8]) {cont8=cont8+1}
if (Tau_max>vc[9]) {cont9=cont9+1}
if (Tau_maxrov>vc[10]) {cont10=cont10+1}
} #chiude il ciclo grande
```

```
print(cont1/S)
print(cont2/S)
print(cont3/S)
print(cont4/S)
```

```
print(cont5/S)
print(cont6/S)
print(cont7/S)
print(cont8/S)
print(cont9/S)
print(cont10/S)
```

```
## 4) CALCOLO SIZE E POTENZA DEL TEST PER N=500
```

```
#Per il calcolo della size e della potenza del test per N=500 si utilizza
  lo stesso procedimento utilizzato per N=250 (si vedano i punti 2 e 3
  di tale appendice). In questo caso bisogna sostituire, all'interno
  dei vettori di valori critici, i valori critici corrispondenti alla
  serie di lunghezza 500 al livello del 5%. Inoltre al posto di N=250
  deve essere posto N=500.
```

```
vc_0=c(-2.320429, -2.316113, 2.387138, -2.584929, -2.320429, 5.548231,
        5.519902, 93.16966, 6.833806, 93.16966)
vc_01=c(-1.76656, -1.814705, 2.650661, -2.075304, -1.76656, 4.336096,
        4.451295, 97.38186, 5.359705, 97.38186)
vc_02=c(-1.772283, -1.852942, 2.635335, -2.063445, -1.772283, 4.576082,
        4.465563, 96.63768, 5.513834, 96.63768)
vc_03=c(-1.765634, -1.814815, 2.589113, -2.073084, -1.765634, 5.246854,
        4.508095, 96.54205, 6.078084, 96.54205)
vc_04=c(-1.694073, -1.83452, 2.649674, -2.05221, -1.694073, 8.422928,
        5.009535, 95.54956, 8.639106, 95.54956)
vc_045=c(-1.574659, -1.850202, 2.597962, -2.033461, -1.574659, 17.41983,
        6.659259, 95.35821, 17.43231, 95.35821)
vc_049=c(-1.2322738, -1.711381, 2.610477, -1.862487, -1.2322738,
        105.3709, 33.7583, 107.53512, 105.3709, 114.65772)
```


Bibliografia

- Beran, J. and Terrin, N. (1996). Testing for a change of the long-memory parameter. *Biometrika*, 83(3):627–638.
- Bisaglia, L. (1998). *Processi a memoria lunga: problemi di stima, identificazione e previsione*. Tesi di dottorato, Università degli Studi di Padova.
- Box, G. E. P. and Jenkins, G. M. (1976). *Time series analysis forecasting and control*. San Francisco: Holden-Day.
- Breitung, J. and Hassler, U. (2002). Inference on the cointegration rank in fractionally integrated processes. *Journal of Econometrics*, 110(2):167 – 185. Long memory and nonlinear time series.
- Busetti, F. and Taylor, A. (2004). Tests of stationarity against a change in persistence. *Journal of Econometrics*, 123(1):33 – 66.
- Di Fonzo, T. and Lisi, F. (2005). *Serie storiche economiche - Analisi statistiche e applicazioni*. Carrocci editore.
- Fox, R. and Taqqu, M. S. (1986). Large-sample properties of parameter estimates for strongly dependent stationary gaussian time series. *The Annals of Statistics*, 14(2):517–532.
- Granger, C. W. J. and Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis*, 1(1):15–29.

- Hassler, U. and Breitung, J. (2006). A residual-based lm-type test against fractional cointegration. *Econometric Theory*, pages 1091–1111.
- Hassler, U. and Scheithauer, J. (2011). Detecting changes from short to long memory. *Statistical Papers*, 52(4):847–870.
- Hosking, J. R. M. (1981). Fractional differencing. *Biometrika*, 68(1):165–176.
- Kim, J.-Y. and Amador, R. B. (2002). Corrigendum to “detection of change in persistence of a linear time series” [j. econom. 95 (2000) 97–116]. *Journal of Econometrics*, 109(2):389 – 392.
- Leybourne, S., Kim, T.-H., Smith, V., and Newbold, P. (2003). Tests for a change in persistence against the null of difference-stationarity. *Econometrics Journal*, 6(2):291–311.
- Leybourne, S., Taylor, R., and Kim, T.-H. (2007). Cusum of squares-based tests for a change in persistence. *Journal of Time Series Analysis*, 28(3):408–433.
- Martins, L. F. and Rodrigues, P. M. (2014). Testing for persistence change in fractionally integrated models: An application to world inflation rates. *Computational Statistics and Data Analysis*, 76:502 – 522. CFEnetwork: The Annals of Computational and Financial Econometrics 2nd Issue.
- Sibbertsen, P. and Kruse, R. (2009). Testing for a break in persistence under long-range dependencies. *Journal of Time Series Analysis*, 30(3):263–285.
- Whittle, P. (1951). *Hypothesis Testing in Time Series Analysis*. Hafner.

Elenco delle tabelle

3.1	Valori critici per $\mathcal{I}_f, \mathcal{I}_r, \mathcal{I}_{r,Rov}, \min(\mathcal{I}_f, \mathcal{I}_r), \min(\mathcal{I}_f, \mathcal{I}_{r,Rov})$	26
3.2	Valori critici per $\mathcal{I}_f^2, \mathcal{I}_r^2, \mathcal{I}_{r,Rov}^2, \max(\mathcal{I}_f^2, \mathcal{I}_r^2), \max(\mathcal{I}_f^2, \mathcal{I}_{r,Rov}^2)$	27
3.3	<i>Size</i> empirica e potenza dei test per i cambiamenti di persistenza ($N = 250$)	31
3.4	<i>Size</i> empirica e potenza dei test per i cambiamenti di persistenza ($N = 500$)	32