

Un test di omogeneità per modelli alle differenze frazionarie

Angela D'Elia

Dipartimento di Scienze Statistiche, Università di Napoli Federico II
E-mail: angela.delia@unina.it

Summary: The aim of the paper is to develop a homogeneity test about the fractional differencing parameter among two or more blocks (subseries) of data. In fact, it is shown that, especially in finite samples, the distribution of the test statistic, obtained by means of asymptotic results, need a correction since the bias and variance are greater than those expected on the basis of the asymptotic theory. A Monte Carlo study is performed to show the gain obtained in terms of size of the critical region when the proposed correction is used.

Key words: Fractional differencing parameter, Homogeneity test, Whittle likelihood.

1. Introduzione

Negli ultimi decenni si è assistito ad un crescente interesse della letteratura statistica ed econometrica nei confronti dei processi cosiddetti a “lunga memoria” e, in corrispondenza, dei modelli a differenze frazionarie. Tale attenzione è essenzialmente derivata dalla constatazione che numerose serie reali presentano funzioni di autocorrelazione persistenti, cioè che decadono più lentamente di quanto avviene nella classe dei processi ARMA. Il comportamento è stato inizialmente osservato in serie storiche di carattere idrologico (Hurst, 1951; Lawrance e Kottegoda, 1977), ma successivamente è stato individuato ed analizzato anche in settori

di diversa natura, quali economia e finanza (si veda Baillie, 1996, per un'ampia panoramica), reti di traffico (Willinger *et al.*, 1998), ecc.

Nell'ambito della modellistica a differenze frazionarie, una problematica particolarmente interessante è quella relativa alla verifica della presenza costante dell'effetto lunga memoria. Infatti, nell'analisi modellistica di osservazioni relative ad un gran numero di dati (cioè, raccolti su lunghi archi di tempo oppure con frequenza elevata) è importante verificare la costanza delle relazioni trovate, poiché la struttura di dipendenza a lunga memoria può variare nel tempo, sia per mutamenti nel meccanismo fisico che genera i dati, sia perché cambia il modo con cui sono raccolte le osservazioni (come è avvenuto, ad esempio, per il criterio di misura del livello minimo della portata del fiume Nilo accertato a Roda negli anni dal 622 al 1284: Tousson, 1925).

La verifica della stabilità del parametro alle differenze stazionarie ha conseguenze importanti sulla struttura del processo (perché la sua variazione modifica le proprietà probabilistiche), sull'inferenza per i parametri (perché i tassi di convergenza degli stimatori dei parametri si modificano con d) e sull'uso del modello (per esempio, si modificano gli intervalli di previsione). A tal fine, Beran e Terrin (1996) hanno proposto un test per verificare se, al variare degli intervalli temporali considerati, il parametro di lunga memoria risulti costante; la statistica test è derivata da un'estensione funzionale del teorema limite centrale per forme quadratiche e dalla teoria dei ponti browniani, per archi di tempo adiacenti: tuttavia, anche per serie di lunghezza notevole, tale test mostra una sostanziale difformità tra livelli nominali e livelli effettivi. Alcuni recenti sviluppi nell'ottica di discriminare tra cambi di livello e variazioni nel parametro d in un modello ARFIMA, mediante un approccio di tipo Bayesiano, sono invece stati elaborati da Ray e Tsay (2001).

Scopo di questo lavoro è sviluppare un test asintotico (alternativo rispetto alla proposta di Beran e Terrin, 1996) che consenta di verificare se, per uno specificato fenomeno di studio, il valore del parametro alle differenze frazionarie d sia *omogeneo* su diversi archi temporali: tale test viene definito di *omogeneità*, a significare che sottoserie di dati anche non adiacenti (e quindi permutabili) hanno lo stesso comportamento per quanto concerne la dinamica di lunga memoria.

Il lavoro è organizzato come segue. Nel paragrafo 2 sono richiamate la specificazione dei modelli a differenze frazionarie e i principali metodi di stima del parametro d , con particolare attenzione al metodo della massima verosimiglianza e alle sue proprietà. Nel paragrafo 3 si presenta una prima proposta di test di omogeneità con i suoi limiti. Nel paragrafo 4 si conduce uno studio Monte Carlo sulla distorsione e varianza dello stimatore del parametro d , e su tale base, nel paragrafo 5 si perviene ad una correzione della distribuzione della statistica test precedentemente proposta. Il guadagno ottenuto in termini di livello del test è illustrato mediante un ampio esperimento di simulazioni (paragrafo 6). Alcune proposte per ulteriori sviluppi della problematica concludono l'articolo.

2. I modelli alle differenze frazionarie

Una classe molto ampia e flessibile di modelli per processi X_t a lunga memoria è costituita dai modelli ARFIMA(p, d, q), introdotti da Granger e Joyeux (1980) e Hosking (1981), e così specificati:

$$\phi(B)(1 - B)^d(X_t - \mu) = \theta(B)a_t,$$

dove μ è il valor medio di X_t , $a_t \sim WN(0, \sigma^2)$, B è l'operatore ritardo tale che $B^k X_t = X_{t-k}$, $k = 0, \pm 1, \dots$, e

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

I polinomi ARMA $\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ e $\theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$ non presentano fattori in comune e ammettono tutte le $(p + q)$ radici al di fuori del cerchio unitario; inoltre, $d \in (-0.5, 0.5)$. Perciò X_t è stazionario ed invertibile (Brockwell e Davis, 1991).

La funzione di densità spettrale di un processo alle differenze frazionarie $X_t \sim \text{ARFIMA}(p, d, q)$ risulta:

$$f(\omega) = |1 - e^{-i\omega}|^{-2d} f_z(\omega), \quad \omega \in [-\pi, \pi],$$

dove $f_z(\omega)$ è la densità spettrale del processo $\nabla^d X_t = Z_t \sim \text{ARMA}(p, q)$.

Il più semplice modello a lunga memoria è il modello ARFIMA $(0,d,0)$ per il quale la componente di lunga memoria dipende dal parametro alle differenze frazionarie d . Se $0 < d < 0.5$, allora X_t è un processo stazionario a lunga memoria con funzione di autocorrelazione che decade a zero con tasso iperbolico:

$$\rho(k) \sim ak^{2d-1}, \quad k \rightarrow \infty,$$

dove a è una costante che non dipende da k ; in tal caso la densità spettrale presenta un picco all'origine. Se $-0.5 < d < 0$ il processo è detto "antipersistente", con densità spettrale tale che $f(0) = 0$ (Beran, 1994).

Nel presente lavoro l'interesse è rivolto a sviluppare un test di omogeneità per il parametro d , esclusivamente per il caso di modelli ARFIMA $(0,d,0)$ con $d > 0$, essendo i fenomeni di antipersistenza di minore interesse e di scarso impatto nelle applicazioni reali. Inoltre, poiché molti risultati asintotici dipendono dalla conoscenza di μ , dato il carattere propositivo di questo lavoro, ci limitiamo a presupporre noto e prefissato $\mu = 0$.

2.1. La stima del parametro d

La stima del parametro alle differenze frazionarie d è stata oggetto di studio da parte di numerosi Autori, mediante approcci e soluzioni abbastanza differenziati. Sono stati infatti proposti stimatori semiparametrici (Geweke e Porter-Hudak, 1983; Robinson, 1995); stimatori basati sul criterio della minima distanza (Tieslau *et al.*, 1996; Corduas, 2000; Galbraith e Zinde-Walsh, 2001); stimatori parametrici nel dominio temporale (Sowell, 1992; Beran, 1995) e nel dominio delle frequenze (Fox e Taqqu, 1986; Cheung e Diebold, 1994); infine, un approccio bayesiano è stato sviluppato da Pay e Ravishanker (1998).

Al fine di sviluppare un test di omogeneità sul parametro alle differenze frazionarie d , nel prosieguo del lavoro si farà riferimento esclusivo allo stimatore di massima verosimiglianza (nella formulazione approssimata di Whittle) e alle sue proprietà asintotiche.

Sia $X_t \sim \text{ARFIMA}(0, d, 0)$ un processo Gaussiano di valor medio 0;

allora, la funzione di log-verosimiglianza risulta (Sowell, 1992):

$$\log L(\mathbf{x}; d) = -\frac{n}{2} \log(2\pi) - \frac{1}{2} \log |\Sigma| - \frac{1}{2} \mathbf{x}' \Sigma^{-1} \mathbf{x},$$

dove $\mathbf{x} = (x_1, x_2, \dots, x_n)'$ è un vettore di osservazioni sul processo, e Σ è la matrice $(n \times n)$ di varianze e covarianze.

Nonostante alcuni recenti sviluppi (Pai e Ravishanker, 1995; Bertelli e Caporin, 2002; Doornik e Ooms, 2002), è noto che la ricerca del massimo di tale funzione di log-verosimiglianza risulta molto pesante dal punto di vista computazionale, specialmente perché l'inversione della matrice Σ di dimensioni $(n \times n)$ richiede il calcolo di funzioni ipergeometriche ad ogni iterazione. Perciò si preferisce massimizzare, invece, la funzione di log-verosimiglianza approssimata proposta da Whittle (Beran, 1994):

$$\log L(\mathbf{x}; d) = -\frac{n}{2} \log(2\pi) - \frac{n}{2} \frac{1}{2\pi} \int_{-\pi}^{\pi} \log f(\omega; d) d\omega - \frac{1}{2} \mathbf{x}' A(d) \mathbf{x} \quad (1)$$

dove la matrice $A(d)$ ha come elemento generico

$$\{a(j-l)\}_{j,l=1,\dots,n} = (2\pi)^{-2} \int_{-\pi}^{\pi} \frac{1}{f(\omega; d)} e^{i(j-l)\omega} d\omega.$$

Nel dominio delle frequenze la log-verosimiglianza profilo della (1) risulta (Boes *et al.*, 1989):

$$\log L(d, \hat{\sigma}^2) \propto -\frac{n}{2} \log \left(\sum_{j=1}^{[(n-1)/2]} \frac{I(\omega_j)}{f(\omega_j)} \right) - \frac{1}{2} \sum_{j=1}^{[(n-1)/2]} \log f(\omega_j), \quad (2)$$

dove $I(\omega_j) = (2\pi n)^{-1} \left| \sum_t x_t^* \exp(-i\omega_j t) \right|^2$ è il periodogramma osservato alle frequenze $\omega_j = 2\pi j/n$, $j = 1, 2, \dots, [(n-1)/2]$, x_t^* è la serie centrata rispetto alla sua media campionaria e $\hat{\sigma}^2 = n^{-1} \sum_j I(\omega_j)/f(\omega_j)$.

Sia \hat{d}_{ML} lo stimatore di massima verosimiglianza ottenuto dalla (1) per il parametro d di un modello ARFIMA(0, d ,0) con $d > 0$. Allora, risulta (Yajima, 1985): $\sqrt{n}(\hat{d}_{ML} - d) \overset{\mathcal{L}}{\sim} N(0, 6/\pi^2)$.

Fox e Taqqu (1986) hanno dimostrato che per lo stimatore ottenuto dalla massimizzazione della log-verosimiglianza profilo (2) - che indicheremo semplicemente con \hat{d} - vale la stessa distribuzione asintotica ottenuta da Yajima (1985), quando $d > 0$. Quindi si ha:

$$\sqrt{n}(\hat{d} - d) \overset{\mathcal{L}}{\sim} N(0, 6/\pi^2).$$

Come è stato spesso sottolineato nella letteratura (Li e McLeod, 1986; Beran, 1995; Baillie 1996), per tale processo la varianza dello stimatore *non* dipende dal parametro d .

3. Un test di omogeneità per il parametro d

La distribuzione asintotica dello stimatore di massima verosimiglianza del parametro d , sopra presentata, diviene il principale punto di riferimento per lo sviluppo di un test di omogeneità su d .

Si consideri una serie \mathbf{x} generata da un processo $X_t \sim \text{ARFIMA}(0, d, 0)$ tale che

$$\nabla^d X_t = a_t,$$

dove $a_t \sim WN(0, \sigma^2)$, e per semplicità di sviluppi si ponga $\sigma^2 = 1$. Siano inoltre $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_s$, s sottoserie di \mathbf{x} , ciascuna di n osservazioni. Al fine di verificare l'esistenza di un comportamento di lunga memoria omogeneo su archi temporali differenti, sottoponiamo a test l'ipotesi:

$$H_0 : d_1 = d_2 = \dots = d_s = d$$

dove d_1, d_2, \dots, d_s sono i parametri alle differenze frazionarie dei processi $\text{ARFIMA}(0, d_i, 0)$ $i = 1, 2, \dots, s$, dai quali si assume siano generate le s sottoserie, e d è noto e prefissato.

Poiché $\sqrt{n}(\hat{d}_i - d) \stackrel{a}{\sim} N(0, 6/\pi^2)$, $\forall i = 1, 2, \dots, s$, allora, sotto H_0 , risulta¹:

$$T_n = \frac{n\pi^2}{6} \sum_{i=1}^s (\hat{d}_i - d)^2 \stackrel{a}{\sim} \chi_s^2,$$

dove T_n è la statistica test per l'ipotesi H_0 di omogeneità. Ne consegue che, posto $\alpha = Pr(T_n \geq \chi_{\alpha, s}^2 \mid H_0)$, $\chi_{\alpha, s}^2$ risulta il valore critico in base al quale rifiutare o meno l'ipotesi di omogeneità.

Come è agevole constatare, la statistica test ora ottenuta è caratterizzata da due evidenti limiti, dovuti alla sua natura asintotica.

¹Si noti che trattandosi di un test di omogeneità, e non di stabilità dinamica, è importante assumere che gli stimatori dei parametri alle differenze frazionarie delle s sottoserie siano tra loro indipendenti.

Infatti, in primo luogo essa si basa sull'assunzione che lo stimatore di massima verosimiglianza \hat{d}_i sia asintoticamente non distorto: in realtà, in campioni di numerosità finita lo stimatore \hat{d}_i può presentare una notevole distorsione, e rendere quindi necessaria una correzione della distribuzione (come proposto fra gli altri da Lieberman, 2001; Lieberman e Phillips, 2001). Analogamente, il risultato per cui $\lim_{n \rightarrow \infty} \{n \text{Var}(\hat{d}_i)\} = 6/\pi^2 \simeq 0.6079$ induce un ulteriore elemento di approssimazione per campioni finiti.

Quando tali fonti di approssimazione non sono tenute in debito conto nella distribuzione della statistica test T_n , ne consegue che il test sopra presentato conduca ad un livello effettivo di gran lunga maggiore di quello nominale (come è ampiamente confermato dallo studio di simulazione illustrato nel paragrafo 6).

4. Distorsione e varianza di \hat{d} in campioni di numerosità finita

Al fine di correggere la distribuzione della statistica test sopra illustrata, è stato condotto un esperimento Monte Carlo, per vari valori di d e n , per studiare il comportamento della distorsione e della varianza di \hat{d} in campioni di numerosità finita.

In particolare, sono state generate 1000 repliche Monte Carlo di serie a differenza frazionaria, per ciascuno dei seguenti valori di $n = 50, 100, 150, 200, 250, 300, 400, 500, 750, 1000, 1500$, e di $d = 0.1, 0.2, 0.3, 0.4$. Per ognuna delle 44 (11×4) coppie di valori (n, d) è stata ottenuta la distribuzione Monte Carlo dello stimatore di massima verosimiglianza, e ne sono state calcolate distorsione e varianza. È stato inoltre calcolato il *divario* tra la varianza Monte Carlo e la varianza asintotica ($0.6079/n$).

La Figura 1 mostra l'andamento della distorsione e del divario al variare di n , per diversi valori di d .

Risulta, quindi, evidente che nella distribuzione Monte Carlo dello stimatore \hat{d} la distorsione e il divario rispetto alla varianza asintotica hanno un comportamento sostanzialmente costante al variare di d , sia in termini di andamento che di valori numerici; come è atteso, tendono entrambi a 0 al crescere della lunghezza n della serie, mentre non sono affatto trascu-

rabili per campioni di numerosità più limitata.

Tale comportamento ci ha indotto a valutare una relazione asintotica tra distorsione e divario da un lato e lunghezza della serie dall'altro. In particolare, utilizzando le 44 realizzazioni della distorsione e del divario dello stimatore, abbiamo ottenuto le seguenti relazioni sulla base di un criterio di adattamento dei Minimi Quadrati².

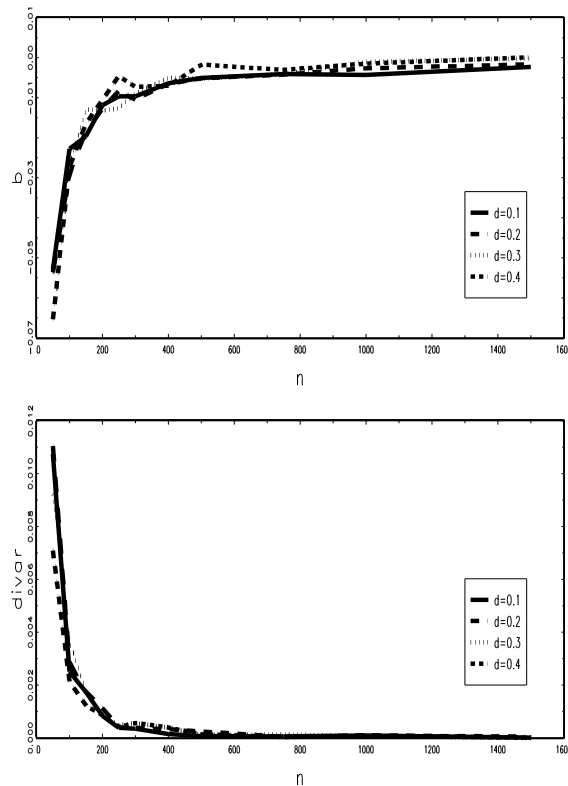


Figura 1. Distorsione e divario dello stimatore \hat{d} .

La distorsione nella stima di massima verosimiglianza è risultata ben approssimata dalla relazione (Figura 2):

$$b_n \simeq \beta_0 + \beta_1 \frac{1}{n}$$

²I risultati qui presentati sottintendono che essi sono stati i migliori e i più coerenti in termini di adattamento tra quelli ottenuti mediante il ricorso a diverse funzioni elementari.

dove $\beta_0 = 0.001137$ ($p\text{-value}=0.03$), $\beta_1 = -2.8290861$ ($p\text{-value} < 0.001$), $R^2 = 0.976$. Essa conferma che la distorsione è dell'ordine $O(n^{-1})$.

Il divario tra la varianza Monte Carlo e quella asintotica è risultata ben approssimata dalla relazione (Figura 2):

$$divar_n \simeq \gamma_1 \frac{1}{n^2}$$

con $\gamma_1 = 24.2386438$ ($p\text{-value} < 0.001$); $\rho^2(divar_n, \widehat{divar}_n) = 0.961$. Essa conferma che la varianza asintotica, al di là di una costante, è dell'ordine $O(n^{-2})$, come ampiamente discusso e dimostrato da Lieberman (2001).

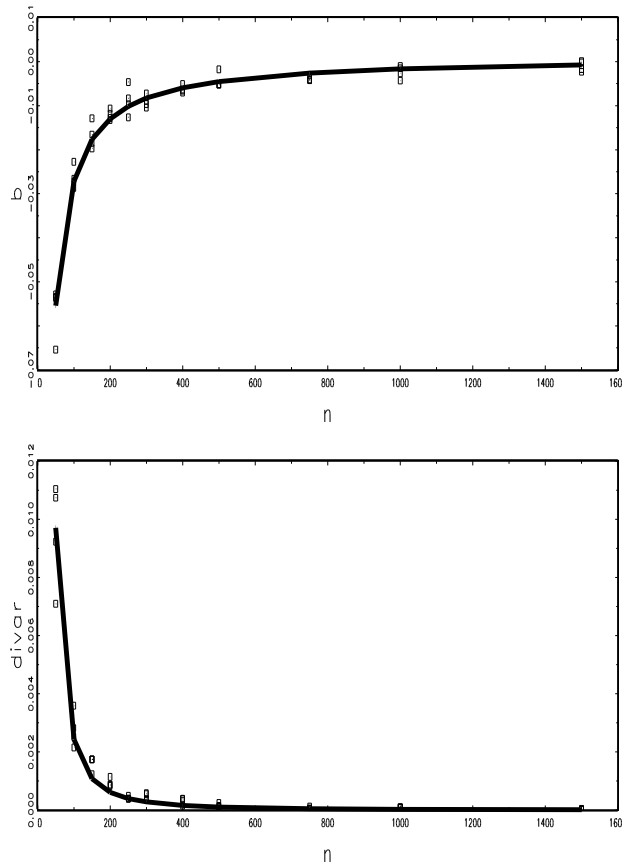


Figura 2. Approssimazioni della distorsione e del divario.

5. Un test di omogeneità per d con distribuzione corretta

Illustriamo, ora, come sia possibile procedere ad una modifica della distribuzione della statistica test T_n , presentata nel paragrafo 3, sulla base delle risultanze del precedente studio Monte Carlo.

E' emerso, infatti, che per numerosità finite: $E(\hat{d}_i) = d_i + b_n$, $b_n \neq 0$; $Var(\hat{d}_i) = (6/(n\pi^2)) + divar_n = v_n$. Allora, sotto l'ipotesi $H_0 : d_i = d$, $\forall i = 1, 2, \dots, s$, si ha:

$$\frac{\sum_{i=1}^s (\hat{d}_i - d)^2}{v_n} - \frac{sb_n^2}{v_n} \stackrel{a}{\sim} \chi_s^2,$$

da cui, mediante agevoli passaggi, si ottengono la statistica test

$$\delta_n^2 = \sum_{i=1}^s (\hat{d}_i - d)^2,$$

e il corrispondente percentile $100(1 - \alpha)\%$:

$$\delta(\alpha, s, n) = \chi_{\alpha, s}^2 v_n + s(b_n)^2.$$

Quindi, sostituendo le approssimazioni ottenute nel paragrafo precedente per la distorsione e il divario, si perviene al seguente percentile modificato:

$$\begin{aligned} \delta(\alpha, s, n) &= \chi_{\alpha, s}^2 \left(\frac{6}{n\pi^2} + \gamma_1 \frac{1}{n^2} \right) + s \left(\beta_0 + \beta_1 \frac{1}{n} \right)^2 = \\ &= s\beta_0^2 + \frac{1}{n} \left(\chi_{\alpha, s}^2 \frac{6}{\pi^2} + 2s\beta_0\beta_1 \right) + \frac{1}{n^2} \left(\chi_{\alpha, s}^2 \gamma_1 + s\beta_1^2 \right). \end{aligned}$$

Dunque, si rifiuta l'ipotesi di omogeneità H_0 tutte le volte che $\delta_n^2 > \delta(\alpha, s, n)$.

Ad esempio, se $n = 100$ e $s = 2$, si rifiuta l'ipotesi $H_0 : d_1 = d_2 = d$ quando

$$(\hat{d}_1 - d)^2 + (\hat{d}_2 - d)^2 \geq (0.2289)^2.$$

6. Uno studio Monte Carlo sul livello del test

Al fine di valutare il livello effettivo del test proposto nel precedente paragrafo, e di quantificare il guadagno ottenuto rispetto alla soluzione illustrata nel paragrafo 3, abbiamo sviluppato uno studio di simulazione.

In particolare, abbiamo simulato, mediante 1000 repliche, la distribuzione Monte Carlo delle statistiche test T_n e δ_n^2 , per diversi valori di n (50, 100, 150, 200, 250, 300, 400, 500, 750, 1000, 1500), di d (0.1, 0.2, 0.3, 0.4), e per un numero di sottoserie s pari a 2.

Posto il livello nominale $\alpha = 0.05$, per ogni coppia di valori (n, d) è stato calcolato il livello effettivo corrispondente alle due statistiche test. Nel prosieguo indicheremo con α_{as} il livello del test basato su T_n (cioè basato sulla distribuzione asintotica dello stimatore \hat{d}), e con α^* il livello effettivo del test basato sulla statistica δ_n^2 (che prevede l'uso del percentile modificato).

Nella Tabella 1 sono illustrati i valori di α_{as} e α^* ottenuti in corrispondenza dei diversi n e d . Appare così evidente che il livello α^* associato alla statistica test δ_n^2 è nettamente inferiore a quello (α_{as}) ottenuto con la statistica test T_n , e molto più vicino al livello nominale $\alpha = 0.05$, soprattutto per serie di lunghezza $n < 1000$. In particolare, quando $n \leq 100$, α^* risulta essere all'incirca pari a $\alpha_{as}/3$.

Si noti, inoltre, che in uno studio Monte Carlo preliminare, qui non riportato, è emerso che la correzione dello stimatore \hat{d} in termini di distorsione, proposta da Lieberman e Phillips (2001), conduce a livelli effettivi di α intermedi rispetto ad α_{as} e α^* .

Come risulta evidente dall'espressione ottenuta per il percentile modificato $\delta(\alpha, s, n)$, il miglioramento nel livello del test dovrebbe essere sensibile anche alla numerosità s delle sottoserie considerate.

Infatti, un andamento analogo (Tabella 2) è stato evidenziato mediante un ulteriore studio di simulazione, condotto nel caso di un numero di sottoserie $s=5$, per $n=100, 200, 500$; $d=0.1, 0.2, 0.3, 0.4$. Ci sembra, quindi, che i risultati precedenti possano essere generalizzati $\forall s \geq 2$.

Ne consegue che, per verificare l'ipotesi di omogeneità del comportamento di lunga memoria su s archi temporali, sia preferibile ricorrere alla statistica test δ_n^2 , e quindi al percentile modificato $\delta(\alpha, s, n)$, ottenuto nel

precedente paragrafo.

Tabella 1. Livelli dei test basati su T_n e su δ_n^2 (per $s=2$)

n	d	α_{as}	α^*	n	d	α_{as}	α^*
50	0.1	0.236	0.082	400	0.1	0.089	0.069
	0.2	0.239	0.078		0.2	0.071	0.057
	0.3	0.199	0.066		0.3	0.085	0.063
	0.4	0.205	0.074		0.4	0.105	0.074
100	0.1	0.170	0.076	500	0.1	0.091	0.074
	0.2	0.140	0.064		0.2	0.075	0.059
	0.3	0.158	0.065		0.3	0.081	0.069
	0.4	0.120	0.053		0.4	0.084	0.076
150	0.1	0.114	0.058	750	0.1	0.072	0.054
	0.2	0.113	0.063		0.2	0.075	0.067
	0.3	0.130	0.081		0.3	0.065	0.054
	0.4	0.096	0.062		0.4	0.074	0.063
200	0.1	0.122	0.080	1000	0.1	0.062	0.054
	0.2	0.106	0.064		0.2	0.072	0.068
	0.3	0.109	0.065		0.3	0.071	0.065
	0.4	0.077	0.045		0.4	0.077	0.069
250	0.1	0.102	0.063	1500	0.1	0.059	0.053
	0.2	0.100	0.065		0.2	0.067	0.064
	0.3	0.101	0.072		0.3	0.060	0.054
	0.4	0.088	0.057		0.4	0.056	0.054
300	0.1	0.087	0.070				
	0.2	0.103	0.070				
	0.3	0.099	0.068				
	0.4	0.088	0.064				

Tabella 2. Livelli dei test basati su T_n e su δ_n^2 (per $s=5$)

n	d	α_{as}	α^*
100	0.1	0.242	0.080
	0.2	0.245	0.092
	0.3	0.239	0.090
	0.4	0.164	0.059
200	0.1	0.143	0.073
	0.2	0.145	0.073
	0.3	0.145	0.079
	0.4	0.112	0.060
500	0.1	0.096	0.066
	0.2	0.076	0.049
	0.3	0.087	0.064
	0.4	0.091	0.062

7. Ulteriori sviluppi

Scopo dell'articolo è stato quello di sviluppare un test di omogeneità per il parametro alle differenze frazionarie di modelli ARFIMA(0, d ,0) con $d > 0$. In particolare, in questo lavoro è stato proposto di ricorrere ad una statistica test estremamente agevole dal punto di vista computazionale, e di modificare il percentile teorico del χ^2 , tenendo in debito conto distorsione e varianza dello stimatore di massima verosimiglianza di d , il cui effetto (rispetto ai risultati asintotici) è piuttosto notevole per numerosità campionarie finite, anche elevate.

Ulteriori sviluppi di tale ricerca saranno rivolti in primo luogo ad uno studio Monte Carlo della potenza del test proposto, specificando con cura l'ipotesi alternativa di interesse.

Inoltre, occorre estendere il test di omogeneità qui proposto al caso più generale di modelli ARFIMA(p, d, q), quando $p > 0, q > 0$.

Ringraziamenti: Il lavoro ha beneficiato di fondi PRIN e CNR (responsabile M. Corduas) afferenti al Dipartimento di Scienze Statistiche, Università degli Studi di Napoli Federico II. L'Autore si è avvalso inoltre delle strutture scientifiche del CFEPSR, Portici. Si ringraziano i referees per gli utili commenti al lavoro.

Riferimenti bibliografici

Baillie R. T. (1996) Long memory processes and fractional integration in econometrics, *Journal of Econometrics*, 73, 5-59.

Beran J. (1994) *Statistics for Long-Memory Processes*, Chapman & Hall, New York.

Beran J. (1995) Maximum likelihood estimation of the differencing parameter for invertible short and long memory autoregressive integrated moving average models, *Journal of the Royal Statistical Society, B*, 57, 659-672.

Beran J., Terrin N. (1996) Testing for a change of the long-memory parameter, *Biometrika*, 83, 627-638; correction in *Biometrika*, 86, 233.

Bertelli S., Caporin M. (2002) A note on calculating autocovariances of long-memory processes, *Journal of Time Series Analysis*, 23, 503-508.

Boes D. C., Davis R. A., Gupta S. N. (1989) Parameter estimation in low order fractionally differenced ARMA models, *Stochastic Hydrology and Hydraulics*, 3, 97-110.

Brockwell P. J., Davis R. A. (1991) *Time series: Theory and methods* (2nd edition), Springer-Verlag, New York.

Cheung Y. W., Diebold F. X. (1994) On maximum likelihood estimation of the differencing parameter of fractionally integrated white noise with unknown mean, *Journal of Econometrics*, 62, 301-316.

Corduas M. (2000) Preliminary estimation of ARFIMA models, in *COMPSTAT 2000*, Physica-Verlag, Heidelberg, 247-252.

Doornik J. A., Ooms M. (2002) Computational aspects of maximum likelihood estimation of autoregressive fractionally integrated moving average models, *Computational Statistics and Data Analysis*, forthcoming.

Fox R., Taqqu M.S. (1986) Large sample properties of parameter es-

timates for strongly dependent stationary gaussian time series, *Annals of Statistics*, 14, 517-532.

Galbraith J.W., Zinde-Walsh V. (2001) Autoregression-based estimators for ARFIMA models, *Technical Reports*, 11, CIRANO papers, Montreal.

Geweke J., Porter-Hudak S. (1983) The estimation and application of long memory time series models, *Journal of Time Series Analysis*, 4, 221-238.

Granger C. W. J., Joyeux R. (1980) An introduction to long memory time series models and fractional differencing, *Journal of Time Series Analysis*, 1, 15-39.

Hosking J. R. M. (1981) Fractional differencing, *Biometrika*, 68, 165-176.

Hurst H. E. (1951) Long-term storage capacity of reservoirs, *Transactions of the American Society of Civil Engineers*, 116, 770-799.

Lawrance A. J., Kottegoda N. T. (1977) Stochastic modelling of river-flow time series, *Journal of the Royal Statistical Society, A*, 140, 1-47.

Li W. K., McLeod A. I. (1986) Fractional time series modelling, *Biometrika*, 73, 217-221.

Lieberman O. (2001) Penalised maximum likelihood estimation for fractional Gaussian processes, *Biometrika*, 88, 888-894.

Lieberman O., Phillips P. C. B. (2001) Second order expansions for the distribution of the maximum likelihood estimator of the fractional difference parameter, *Cowles Foundation Discussion Paper*, No. 1308, Yale University, Connecticut.

Pai J. S., Ravishanker, N. (1995) Exact likelihood function forms for an ARFIMA process, *Technical Report*, No. 9509, Department of Statistics, University of Connecticut, Storrs.

Pai J. S., Ravishanker, N. (1998) Bayesian analysis of autoregressive fractionally integrated moving-average processes, *Journal of Time Series Analysis*, 19, 99-112.

Ray B. K., Tsay R. S. (2002) Bayesian methods for change-point detection in long-range dependent processes, *Journal of Time Series Analysis*, forthcoming.

Robinson P. M. (1995) Log-periodogram regression of time series

with long range dependence, *Annals of Statistics*, 23, 1048-1072.

Sowell F. B. (1992) Maximum likelihood estimation of stationary univariate fractionally integrated time series models, *Journal of Econometrics*, 53, 165-188.

Tieslau M.A., Schmidt P., Baillie R.T. (1996) A minimum distance estimator for long-memory errors, *Journal of Econometrics*, 71, 249-264.

Tousson O. (1925) Mémoire sur l'histoire du Nil, *Mémoires de l'Institut d'Egypte*, 9, 366-385.

Yajima Y. (1985) On estimation of long memory time series models, *Australian Journal of Statistics*, 27, 303-320.

Willinger W., Paxson V., Taqqu M. S. (1998) Self similarity and heavy tails: structural modeling of network traffic, in *A Practical Guide to Heavy Tails*, (Adler R. J., Feldman R. E. Taqqu M. S. eds.), Birkhauser, Boston, 27-53.