

Università degli Studi di Padova
Dipartimento di Scienze Statistiche
Corso di Laurea Triennale in
Statistica Economia e Finanza



**RELAZIONE FINALE
UN TEST PER L'AUTOCORRELAZIONE BASATO SULLO
STIMATORE DI CAUCHY**

Relatore: Prof. Luisa Bisaglia
Dipartimento di Scienze Statistiche

Laureanda: Chiara Pastrello
Matricola N. 1033406

Anno Accademico 2013/2014

Sommario

Introduzione	5
Capitolo 1 – Un test per l'autocorrelazione basato sullo stimatore di Cauchy	7
1.1 Introduzione	7
1.2 Il test di Cauchy	9
1.3 Altri due test per la correlazione	14
1.3.1 Il test di Ljung-Box.....	15
1.3.2 Il test di Monti	16
Capitolo 2 – Un test per la bontà di adattamento di modelli ARMA sottostimati	19
2.1 Introduzione	19
2.2 Una variante del test di Cauchy	21
Capitolo 3 – Studio di simulazione	23
3.1 Introduzione	23
3.2 Livello empirico dei test	24
3.2.1 Livello empirico con distribuzioni asintotiche	24
3.2.2 Livello empirico con metodo Monte Carlo	33
3.3 Potenza empirica dei test	37
3.3.1 Potenza empirica per AR(1)	37
3.3.2 Potenza empirica per ARMA sottostimati	44
Conclusioni	51
Appendice	52
A.1 Codice R per livello empirico con distribuzioni asintotiche e dati WN.....	52
A.2 Codice R per livello empirico con algoritmo MC e dati AR(1)	54
A.3 Codice R per potenza empirica per dati AR(2) sottostimati con AR(1).....	58
Bibliografia	63

Introduzione

Questo lavoro trae spunto dall'articolo “*A Cauchy estimator test for autocorrelation*” di Colin M. Gallagher, Thomas J. Fisher & Jie Shen (2014), in cui gli autori propongono un test basato sullo stimatore di Cauchy per verificare l'ipotesi nulla di incorrelazione seriale.

Nella mia tesi viene presentato questo nuovo test per la correlazione seriale in una serie temporale stazionaria osservata.

Quando si ha a che fare con serie di numerosità limitata, ad esempio pochi anni di dati mensili, il test più comunemente implementato di Ljung-Box (LB) per l'autocorrelazione ha poco potere; pertanto può essere utile avere a disposizione dei test più efficaci nel riconoscere la presenza di correlazione significativa con ridotte numerosità campionarie.

Nell'articolo viene sviluppato un test che soddisfa proprio queste richieste; gli autori sostengono infatti che il test basato sullo stimatore di Cauchy sia più potente nel determinare la correlazione in piccoli campioni rispetto ad altri test tradizionali basati sulla funzione di autocorrelazione campionaria.

Lo stimatore di Cauchy era già stato considerato precedentemente in letteratura per la stima di parametri in processi autoregressivi (Gallagher, 2001), per l'identificazione del modello (Gallagher, 2002), e per calcolare intervalli di confidenza per parametri autoregressivi (So e Shin, 1999, Gallagher e Tunno, 2008).

Nell'articolo questo test di Cauchy (abbreviato in seguito con CT, per *Cauchy Test*) per la correlazione seriale viene poi modificato per ricavare una statistica di bontà di adattamento per modelli ARMA sottostimati.

I comportamenti in campioni finiti del test di Cauchy e della sua variante sono analizzati con uno studio di simulazione e vengono confrontati con la statistica di Ljung-Box (LB) e il test di Monti.

Lo scopo di questa tesi è, dunque, quello di replicare quanto eseguito dagli autori nell'articolo sopra citato e confermare i risultati da essi ottenuti, verificando che la nuova statistica proposta sia effettivamente più potente rispetto alle altre.

La tesi è suddivisa in tre capitoli: i primi due costituiscono la parte teorica, mentre l'ultimo riporta i risultati delle simulazioni.

Nel primo capitolo si introduce la statistica basata sullo stimatore di Cauchy e ne viene derivata la distribuzione asintotica sotto l'ipotesi nulla di dati indipendenti e identicamente distribuiti (iid). L'ultima parte del capitolo è dedicata alla descrizione degli altri due test, LB e Monti, che verranno confrontati col CT.

Nel secondo capitolo si accenna al problema della bontà di adattamento ai dati di un modello ARMA, si descrive la variante del CT da applicare come test di correlazione ai residui di ARMA sottostimati, e viene derivata la distribuzione asintotica.

Nel terzo capitolo sono, infine, sintetizzati gli esiti degli esperimenti condotti per confrontare il comportamento delle statistiche proposte; i livelli e le potenze empiriche dei test sono stati ricavati utilizzando sia le rispettive distribuzioni asintotiche, sia un algoritmo di tipo Monte Carlo.

Capitolo 1 – Un test per l'autocorrelazione basato sullo stimatore di Cauchy

1.1 Introduzione

Lo studio della correlazione nell'analisi empirica delle serie temporali è importantissima per poter identificare il modello che meglio si adatta ai dati, per migliorare l'accuratezza delle previsioni e per evitare inferenza errata sui parametri di regressione. Tuttavia, se la correlazione non è presente c'è il rischio di aggiungere inutili complicazioni al modello, nonché errori nella fase di stima. È, quindi, un passo importante verificare prima di tutto l'esistenza di correlazione seriale nel processo di studio.

Diamo allora una breve definizione di *autocorrelazione*.

In un processo stocastico stazionario, $\{X_t\}$ con $t \in T$, la funzione di autocovarianza dipende solo dalla distanza temporale (lag) k tra due variabili X_t e X_{t+k} e non dagli istanti t e $t+k$, e può essere scritta nel seguente modo:

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

Nei processi stocastici stazionari l'autocovarianza assume un ruolo fondamentale, in quanto è un indice delle relazioni lineari esistenti tra coppie di v.c. componenti il processo stocastico prese con uno sfasamento pari a k . Inoltre, l'autocovarianza di un processo stocastico stazionario è una funzione pari di k , cioè $\gamma_k = \gamma_{-k}$.

Dato un processo stocastico stazionario X_t di media μ , varianza σ^2 e funzione di autocovarianza γ_k , la *funzione di autocorrelazione* (o *ACF*, da *AutoCorrelation Function*) è definita come:

$$\rho_{t,t+k} = \rho_k = \frac{\gamma_k}{\gamma_0}, \quad k = 0, \pm 1, \pm 2, \dots$$

L'ACF è una misura della correlazione tra X_t e X_{t+k} e fornisce le stesse informazioni della funzione di autocovarianza, in quanto ne è la normalizzazione, ma ha il vantaggio di essere un numero puro, indipendente dall'unità di misura di X_t .

La funzione di autocorrelazione soddisfa le seguenti proprietà:

$$\rho_0 = 1$$

$$\rho_k = \rho_{-k} \text{ (è una funzione pari)}$$

$$|\rho_k| \leq 1$$

Si può, inoltre, definire la *funzione di autocorrelazione parziale (PACF)* per tenere in considerazione che la correlazione tra due variabili può essere dovuta al fatto che esiste effettivamente un legame lineare diretto tra le variabili o al fatto che queste ultime sono correlate con una terza variabile. Molto spesso nelle serie storiche la correlazione tra X_t e X_{t-k} può essere dovuta alla correlazione che queste variabili hanno con $X_{t-1}, X_{t-2}, \dots, X_{t-k+1}$. La PACF misura l'autocorrelazione tra X_t e X_{t-k} al netto delle variabili intermedie. La funzione di autocorrelazione parziale è dunque pari alla correlazione condizionata

$$P_k = \text{Corr}(X_t, X_{t-k} \mid X_{t-1}, \dots, X_{t-k+1}).$$

1.2 Il test di Cauchy

Sia $\{X_t\}$ una serie temporale debolmente stazionaria osservata al tempo $t = 1, 2, \dots, n$. Dal momento che possiamo sempre sottrarre la media campionaria, senza perdita di generalità assumeremo che $E(X_t) = 0$.

Consideriamo l'equazione

$$X_t = \rho(k)X_{t-k} + Z_t^{(k)}, \quad k = 1, 2, \dots, m, \quad (1)$$

dove $\rho(k)$ è la correlazione al lag k , e $Z_t^{(k)}$ ha media 0 e varianza $\sigma^2 = E(X_t - \rho(k)X_{t-k})^2$.

Per esempio, se $\{X_t\}$ è un processo autoregressivo di ordine 1, l'equazione (1) assumerà la forma:

$$X_t = \phi X_{t-1} + \varepsilon_t,$$

con $\{\varepsilon_t\}$ che consiste di una successione di v.c. indipendenti e identicamente distribuite (iid) secondo una $N(0, \sigma^2)$, e pertanto $Z_t^{(k)}$ si ricava per differenza come $Z_t^{(k)} = X_t - \phi^k X_{t-k}$.

D'altra parte, se $\{X_t\}$ è una successione iid allora per ogni k vale che $Z_t^{(k)} = Z_t = X_t$.

Si noti che per costruzione $Z_t^{(k)}$ e X_{t-k} sono incorrelate e $\rho(k)X_{t-k}$ è il miglior predittore lineare di X_t basato su X_{t-k} .

Se $\{X_t\}$ è un processo gaussiano, allora

$$E(X_t | X_{t-k}) = \rho(k)X_{t-k} \quad (1.1)$$

Dall'equazione (1.1) si nota che, per ogni funzione g misurabile,

$$E(X_t g(X_{t-k})) = \rho(k)E(X_{t-k} g(X_{t-k})).$$

In particolare, sia $g(X_{t-k}) = |X_{t-k}|^r S_{t-k}$, con $r \geq 0$ e $S_t = \text{sign}(X_t)$.

Notiamo che

$$E(X_t | X_{t-k} |^r S_{t-k}) = \rho(k)E |X_{t-k}|^{r+1},$$

o equivalentemente

$$\rho(k) = \frac{E(X_t | X_{t-k} |^r S_{t-k})}{E | X_{t-k} |^{r+1}}. \quad (2)$$

Prendendo $r = 1$, l'equazione (2) corrisponde all'usuale autocorrelazione al ritardo k ; altri valori di r danno la correlazione tra X_t e $|X_{t-k}|^r S_{t-k}$.

Per ogni successione di v.c. indipendenti, $\rho(k) = 0$ per ciascun valore di k diverso da zero.

Il caso speciale di $r = 0$ corrisponde alla correlazione tra $\{X_t\}$ e la serie dei segni $\{S_t\}$; da qui in poi, consideriamo $\rho(k)$ data dall'equazione (2) con $r = 1$.

Per un qualsiasi processo ARMA invertibile, con innovazioni iid, la funzione $\rho(k)$ data dall'equazione (2) soddisfa le stesse proprietà – e quindi assume anche gli stessi valori – della funzione di autocorrelazione.

Per processi ARMA non invertibili, consideriamo la correlazione tra valori presenti del processo e valori passati della serie dei segni.

Per verificare se è presente correlazione significativa in almeno uno dei primi m ritardi, il sistema di ipotesi che andiamo a verificare è:

$$\begin{cases} H_0 : \rho(k) = 0, & k = 1, 2, \dots, m \\ H_1 : \rho^2(k) > 0 & \text{per qualche } 1 \leq k \leq m. \end{cases}$$

In modo analogo si può valutare se la norma euclidea del vettore $\boldsymbol{\rho} = (\rho(1), \dots, \rho(m))^t$ è zero.

Ora deriviamo una nuova statistica test per il sistema d'ipotesi considerato e ne ricaviamo la distribuzione asintotica sotto l'ipotesi nulla di v.c. iid.

Per stimare $\rho(k)$ dai dati, possiamo ricorrere allo stimatore dei momenti

$$\hat{\rho}_r(k) = \frac{\sum_{t=k+1}^n X_t |X_{t-k}|^r S_{t-k}}{\sum_{t=k+1}^n |X_{t-k}|^{r+1}}, \quad (2.1)$$

dove per $r = 1$ la (2.1) corrisponde alla usuale funzione di autocorrelazione campionaria e per $r = 0$ dà l'espressione dello stimatore di Cauchy per l'autocorrelazione al ritardo k . Diversamente dalle solite funzioni di covarianza e correlazione, la versione di Cauchy non è simmetrica nei suoi argomenti.

Per il test di correlazione presentato in questa sezione, vengono usati solo valori positivi di k , mentre per il test di bontà di adattamento del Capitolo 2 useremo valori di k sia positivi che negativi.

Ora consideriamo la seguente espressione:

$$\sum Z_t^{(k)} S_{t-k} = \sum |X_{t-k}| (\hat{\rho}_0(k) - \rho(k)).$$

Per dati iid (ipotesi nulla), gli addendi a sinistra dell'uguaglianza formano una successione dipendente da k tale che

$$(n-k)^{-1/2} \sum Z_t S_{t-k} \Rightarrow N(0, \sigma_k^2),$$

con \Rightarrow che indica convergenza debole e $\sigma_k^2 = \sigma^2$ per $k = 1, 2, \dots, m$.

Per successioni simmetriche e iid, il tasso di convergenza è abbastanza veloce e la distribuzione congiunta di $(Z_{k+1}S_1, \dots, Z_n S_{n-k})$ è uguale a quella di (Z_{k+1}, \dots, Z_n) (si veda il Teorema 3.1 di Gallagher e Tunno, 2008). Inoltre, per dati gaussiani la distribuzione della sommatoria è esattamente normale per ogni n .

È facile vedere che, sotto l'ipotesi nulla di dati iid, per ogni k

$$\sum Z_t S_{t-k} = \hat{\rho}_0(k) \sum |X_{t-k}| = \sum X_t S_{t-k},$$

e dal momento che per ogni t, s ,

$$\text{Cov}(Z_t S_{t-k}, Z_s S_{s-l}) = \begin{cases} \text{Var}(Z_t) & t = s; k = l, \\ 0 & \text{altrimenti,} \end{cases}$$

tramite il teorema di Cramér–Wold si giunge al seguente risultato.

TEOREMA 2.1

Se $\{X_t\}$ è una successione iid, H_0 è vera e

$$\left(\frac{\sum X_t S_{t-1}}{\sqrt{n-1}}, \dots, \frac{\sum X_t S_{t-m}}{\sqrt{n-m}} \right) \Rightarrow N_m(0, D),$$

dove \mathbf{D} è una matrice diagonale con k -esimo elemento $\sigma_k^2 = \sigma^2$.

Notiamo che se i dati sono iid e gaussiani, allora le componenti del vettore hanno ciascuna distribuzione marginale normale e sono tra loro incorrelate.

Tuttavia, se le componenti non sono indipendenti, la normalità della distribuzione congiunta vale solo al limite anche in quel caso.

Per verificare il nostro sistema di ipotesi possiamo usare la norma euclidea del vettore delle correlazioni campionarie di Cauchy standardizzate,

$$\lambda(k) = \frac{\sum X_t S_{t-k}}{\sqrt{(n-k)\hat{\sigma}_k^2}},$$

dove $\hat{\sigma}_k$ è un qualsiasi stimatore consistente di σ_k .

Usando risultati standard¹, si ottiene, quindi, il seguente risultato.

COROLLARIO 2.2

Se $\{X_t\}$ è una successione iid, H_0 è vera e

$$\sum_{k=1}^m (\lambda(k))^2 \Rightarrow \chi_m^2,$$

dove χ_m^2 indica una variabile casuale con distribuzione chi-quadrato con m gradi di libertà.

La varianza σ_k^2 può essere stimata sotto l'ipotesi nulla sfruttando i valori osservati di $\{X_t\}$, ma, per migliorare la performance delle nostre simulazioni, in seguito stimeremo σ_k^2 sotto l'ipotesi alternativa con

$$\hat{\sigma}_k^2 = \frac{\sum_{t=k+1}^n (X_t - \hat{\rho}_1(k)X_{t-k})^2}{(n-k)}. \quad (3)$$

Useremo poi la correlazione campionaria per stimare ρ , dal momento che ha un errore quadratico medio (MSE) minore rispetto allo stimatore di Cauchy dell'autocorrelazione.

La statistica test del Corollario 2.2 è il nostro test di Cauchy¹, o CT, con cui ci proponiamo di verificare un'ipotesi nulla di dati iid contro una generale alternativa di correlazione seriale.

¹Per ulteriori dettagli, dimostrazioni e approfondimenti circa la statistica rimando all'articolo di riferimento (Gallagher, Fisher e Shen, 2014).

1.3 Altri due test per la correlazione

In questa sezione vengono presentati altri due test per verificare la presenza di autocorrelazione nei dati, il test di Ljung-Box e il test di Monti, i quali saranno poi messi a confronto nel Capitolo 3 con il test di Cauchy appena ricavato.

Recentemente, sono state introdotte in letteratura molte nuove statistiche, tra cui test asimmetrici basati sul determinante (Peña e Rodríguez, 2002-2006) e sulla traccia (Fisher e Gallagher, 2012) della matrice di autocorrelazione campionaria, test asimmetrici portmanteau, come una versione asimmetrica del test dello stimatore di Cauchy (simile a quella di Mahdi e McLeod, 2012), e varianti pesate di Ljung-Box e di Monti (Fisher, 2011). Tuttavia, per brevità, nelle prossime analisi saranno tralasciati confronti con questi ultimi, e la nostra attenzione sarà focalizzata solo sui test CT, LB e Monti tradizionali.

I test LB e Monti vengono entrambi considerati statistiche portmanteau.

Un test portmanteau è un tipo di test usato per verificare ipotesi statistiche in cui l'ipotesi nulla è ben specificata, mentre l'ipotesi alternativa è più libera. I test così costruiti possono avere la proprietà di essere almeno moderatamente potenti contro una generale e vasta gamma di possibili scostamenti dall'ipotesi nulla.

Così facendo, nella statistica applicata, un test portmanteau fornisce una via di procedere come un check generale in una situazione dove ci sono molti modi in cui un modello può discostarsi dal processo generatore dei dati. L'uso di questi test evita di dover essere troppo specifici circa il particolare tipo di scostamento che viene riscontrato.

La statistica di Ljung-Box, che è una versione migliorata del test di Box-Pierce, e quella di Monti sono esempi di test portmanteau utilizzati nell'analisi delle serie temporali per verificare la presenza di autocorrelazione.

Procediamo dunque con la loro descrizione.

1.3.1 Il test di Ljung-Box

Il test di Ljung-Box (che prende il nome da Greta M. Ljung e George E. P. Box, 1978) è un tipo di test statistico utilizzato per valutare se in una serie temporale ci sono alcune autocorrelazioni significativamente diverse da zero. Invece di procedere analizzando ogni singolo lag distintamente, consiste in una verifica complessiva di assenza di correlazione "globale" sulla base di un certo numero di ritardi, contro una generica alternativa di correlazione, e per questo è considerato un test portmanteau.

Questo test è spesso conosciuto come Ljung-Box "Q" test, ed è strettamente legato al test di Box-Pierce (da George E. P. Box e David A. Pierce, 1970). In verità, la statistica test Ljung-Box fu descritta in modo esplicito nel documento che portò all'uso del Box-Pierce. Quest'ultimo è una versione semplificata del test di Ljung-Box, da cui è stato sostituito nella maggior parte delle applicazioni moderne, perché i successivi studi di simulazione ne mostrarono il scarso rendimento.

Sempre attraverso studi di simulazione è stato dimostrato che la statistica LB è preferita a BP per avere performance migliori soprattutto in piccoli campioni.

Diamo allora una definizione formale di questa statistica.

Il test di Ljung-Box ha lo scopo di verificare il seguente sistema di ipotesi:

$$\begin{cases} H_0 : \rho_1 = \rho_2 = \dots = \rho_k = \dots = \rho_m = 0 \\ H_1 : \rho_i \neq 0 \text{ per almeno un } i \text{ } 1 \leq i \leq m, m = 1, 2, \dots \end{cases} \Rightarrow \begin{cases} H_0 : \text{dati incorrelati} \\ H_1 : \text{dati correlati} \end{cases}$$

La statistica test è

$$Q = n(n+2) \sum_{k=1}^m \frac{\hat{\rho}_k^2}{n-k}$$

dove n è la dimensione del campione di dati, $\hat{\rho}_k$ è la funzione di autocorrelazione campionaria al ritardo k , e m è il numero di autocorrelazioni prese in esame.

Per dati indipendenti e identicamente distribuiti, la statistica test si distribuisce approssimativamente come un χ_m^2 .

Tuttavia, è opportuno sottolineare che il test di Ljung-Box è comunemente utilizzato lavorando con modelli ARIMA come controllo diagnostico per valutarne la bontà di adattamento. In questo caso, viene applicato per verificare la correlazione dei residui del modello stimato, e non della serie originale, utilizzando come $\hat{\rho}_k$ la funzione di autocorrelazione campionaria dei residui osservati.

Quando applicato ai residui di un modello ARIMA stimato, i gradi di libertà del chi-quadro devono essere aggiustati applicando un correttivo, per tenere in considerazione la stima dei parametri. Dato un ARIMA(p,0,q), sotto l'ipotesi di corretta specificazione, la distribuzione asintotica del test LB è χ_{m-p-q}^2 , con gradi di libertà pari a $m-p-q$, dove p e q sono gli ordini, rispettivamente, della parte autoregressiva e di quella a media mobile nel modello stimato.

Ad esempio, studiando i residui di un processo AR(1), come faremo in seguito, il test avrà distribuzione χ_{m-1}^2 .

1.3.2 Il test di Monti

Una statistica simile al test di Ljung-Box fu introdotta da Monti (1994).

Anche questo test è considerato un test portmanteau, e si pone l'obiettivo di verificare il medesimo sistema di ipotesi; tuttavia, il test di Monti si differenzia dal precedente in quanto al posto dell'autocorrelazione campionaria, $\hat{\rho}_k$, utilizza la funzione di autocorrelazione parziale, \hat{P}_k , per $k = 1, 2, \dots, m$.

La statistica test è

$$M = n(n+2) \sum_{k=1}^m \frac{\hat{P}_k^2}{n-k}$$

dove n è la dimensione del campione, \hat{P}_k è l'autocorrelazione parziale al ritardo k , e m è il numero di ritardi considerati.

Per quanto riguarda la sua distribuzione asintotica, vale quanto detto per il test di Ljung-Box: sotto l'ipotesi nulla di dati incorrelati, la statistica test è approssimativamente distribuita come un χ_m^2 , mentre se viene applicata ai residui di un modello ARIMA stimato e non alla serie originale avrà distribuzione χ_{m-p-q}^2 , posto che il modello ARIMA sia correttamente stimato.

Precedenti studi di simulazione hanno dimostrato che la performance del test di Monti è comparabile a quella di Ljung-Box, ma è più potente per processi a media mobile sottostimati; viceversa, LB ha prestazioni migliori se viene sottostimato l'ordine della parte autoregressiva.

Il test di Monti è stato scelto nell'articolo di riferimento perché, diversamente dal solito, si basa sulla funzione di autocorrelazione parziale, ed inoltre fornisce un altro termine di confronto.

Capitolo 2 – Un test per la bontà di adattamento di modelli ARMA sottostimati

2.1 Introduzione

Nell'articolo di Gallagher, Fisher e Shen a cui facciamo riferimento viene presentata una variante del test di Cauchy da applicare ai residui per valutare la bontà di adattamento, quando si lavora con modelli ARMA sottostimati.

Dopo l'identificazione del modello e la stima dei parametri, il passo finale è quello di valutarne l'adeguatezza mediante opportune analisi e test diagnostici. Tra le assunzioni fondamentali da esaminare ci sono quelle che riguardano i residui.

Sia $\{X_t\}$ un processo autoregressivo a media mobile stazionario ed invertibile, scritto nella forma

$$X_t - \sum_{i=1}^p \phi_i X_{t-i} = \phi_0 + \varepsilon_t - \sum_{j=1}^q \theta_j \varepsilon_{t-j}, \quad (4)$$

con $\{\varepsilon_t\}$ processo *white noise* (cioè v.c. a media nulla, varianza costante e incorrelate). Per ogni modello stimato possiamo calcolare i residui stimati, $\hat{\varepsilon}_t$, che sono stime dei veri ε_t non osservabili.

Il controllo dell'adeguatezza del modello ai dati è effettuato attraverso un'attenta analisi della serie dei residui $\hat{\varepsilon}_t$, alla base di cui c'è la considerazione che se il modello è stato correttamente identificato e stimato, allora sui residui $\hat{\varepsilon}_t$ devono potersi riscontrare le ipotesi fatte a priori sui disturbi ε_t . Tra queste la più importante riguarda l'*incorrelazione seriale*; infatti, se i residui risultano incorrelati, significa che abbiamo modellato adeguatamente la struttura di dipendenza presente nei dati.

La serie dei residui può essere trattata come una serie storica a sé stante per la quale è possibile calcolare le funzioni di autocorrelazione empiriche.

Per verificare l'assenza di correlazione si può procedere in diversi modi, per esempio, esaminando le funzioni di autocorrelazione empiriche (ACF e PACF) dei residui, oppure attraverso dei test.

E' pratica comune valutare la bontà di adattamento di un modello ARMA tramite un test portmanteau di significativa correlazione. A tal proposito si possono applicare ai residui del modello stimato i test portmanteau tipicamente utilizzati nell'analisi delle serie storiche, come il test di Ljung-Box o simili.

Se il processo ARMA viene correttamente identificato, o eventualmente sovrastimato, il valore di ciascuna delle autocorrelazione tra i residui sarà approssimativamente uguale a zero. Se invece il modello ARMA è sottostimato, le autocorrelazioni si discosteranno da zero e assumeranno valori fino a ± 1 .

Un test di corretta specificazione per valutare la bontà di adattamento del modello ai dati risulterà, pertanto, equivalente a un test che verifica l'ipotesi d'incorrelazione dei residui.

2.2 Una variante del test di Cauchy

Il CT descritto nel capitolo precedente può essere modificato per determinare la bontà di adattamento di un processo ARMA così specificato:

$$x_t = \phi_1 x_{t-1} + \dots + \phi_p x_{t-p} + z_t + \theta_1 z_{t-1} + \dots + \theta_q z_{t-q}, \quad (5)$$

dove $\{z_t\}$ è una successione di v.c. iid di media zero e varianza σ^2 con $P(z_t = 0) = 0$.

A tal fine, stimiamo il modello (5) usando il metodo della massima verosimiglianza o una qualsiasi altra procedura di stima con lo stesso comportamento asintotico del metodo dei minimi quadrati.

Siano ora $\{\hat{z}_t\}$ i residui stimati del modello e $\hat{s}_t = \text{sign}(\hat{z}_t)$ i rispettivi segni.

A differenza del CT del Capitolo 1 per cui sono stati usati solo valori positivi di k , distinguiamo tra valori di k positivi e negativi, e modifichiamo il CT derivando una statistica costruita coi residui stimati \hat{z}_t invece che con la serie x_t .

Per $k > 0$, consideriamo le due statistiche

$$\hat{\lambda}(k) = \frac{\sum_{t=k+1}^n \hat{z}_t \hat{s}_{t-k}}{\hat{\sigma} \sqrt{(n-k)}} \quad \text{e} \quad \hat{\lambda}(-k) = \frac{\sum_{t=k+1}^n \hat{s}_t \hat{z}_{t-k}}{\hat{\sigma} \sqrt{(n-k)}},$$

dove $\hat{\sigma}$ è uno stimatore consistente di σ .

Vale il seguente Teorema:

TEOREMA 3.1

Se $\{X_t\}$ è un processo che si può scrivere nella forma (5), allora per $n \rightarrow \infty$,

$$(\hat{\lambda}(-1), \dots, \hat{\lambda}(-m), \hat{\lambda}(1), \dots, \hat{\lambda}(m))' \Rightarrow N_m(0, W)$$

con

$$W = \begin{pmatrix} I & \delta(I - H) \\ \delta(I - H) & I - \delta H \end{pmatrix},$$

dove \mathbf{I} è la matrice identica $m \times m$, \mathbf{H} una matrice di proiezione di rango $p + q$, e $\delta = (E|z_t|^2) / E(z_t^2)$.

Il corollario che segue può essere utilizzato per sviluppare un test di bontà di adattamento usando $\hat{\lambda}(k)$.

COROLLARIO 3.2

Se $\{X_t\}$ è un processo che si può scrivere nella forma (5), allora per $n \rightarrow \infty$,

$$\sum_{k=-m}^{-1} \hat{\lambda}(k)^2 + \sum_{k=1}^m \hat{\lambda}(k)^2 \Rightarrow Y_1 + (1 + \delta)Y_2 + (1 - \delta)Y_3,$$

dove Y_1 , Y_2 , e Y_3 sono variabili casuali indipendenti con distribuzione chi-quadrato con, rispettivamente, $p + q$, $m - p - q$, e $m - p - q$ gradi di libertà.²

²Per le dimostrazioni del Teorema e Corollario sopra enunciati si veda l'articolo di riferimento (Gallagher, Fisher e Shen, 2014).

Capitolo 3 – Studio di simulazione

3.1 Introduzione

In questo capitolo sono sintetizzati gli esiti delle simulazioni condotte per confrontare il comportamento delle statistiche proposte nei precedenti capitoli.

L'attenzione è rivolta ad indagare i livelli empirici (*size*, in inglese) e le potenze empiriche dei test ad un livello di significatività del 5%.

Vengono presentati i risultati che mostrano le situazioni nelle quali il CT fornisce dei miglioramenti rispetto agli altri metodi.

Il risultato principale è che per piccoli campioni la statistica proposta è in grado di determinare più facilmente l'autocorrelazione rispetto agli altri test basati sull'autocorrelazione campionaria. Inoltre, la statistica di Cauchy funziona relativamente bene al crescere della numerosità campionaria.

Tipicamente, le regole di decisione per un test si basano sui quantili della relativa distribuzione asintotica; in questo capitolo studieremo il comportamento delle nostre statistiche utilizzando sia le loro distribuzioni asintotiche, sia un metodo di tipo Monte Carlo esposto da Lin e McLeod (2006).

I programmi per le simulazioni sono stati scritti con il software statistico R (www.r-project.org).

Alcuni dei comandi di R utilizzati per le analisi statistiche sono riportati in **Appendice**.

In generale, i risultati ottenuti sono in linea con quelli suggeriti dall'articolo di riferimento Gallagher, Fisher e Shen (2014).

3.2 Livello empirico dei test

Un aspetto importante da tenere in considerazione quando si vuole verificare la presenza di correlazione in una serie temporale è la scelta del ritardo m al quale condurre i test.

Sono stati proposti vari suggerimenti in letteratura, ma non è noto un ritardo ottimale.

Negli studi che seguono, useremo i ritardi $m = \sqrt{n}$ e $m = n/3$, basandoci su alcune regole del pollice in letteratura.

Come già discusso riguardo al Corollario 2.2 del primo capitolo, nella nostra statistica CT può essere usato un qualsiasi stimatore consistente per la deviazione standard σ_k .

Per processi *white noise*, l'errore quadratico medio (MSE) calcolato dai dati funziona bene, ma si possono migliorare le prestazioni dei test utilizzando lo stimatore $\hat{\sigma}_k^2$ dato dall'equazione (3).

Iniziamo la nostra analisi studiando il livello empirico delle varie statistiche test in campioni finiti, al variare della numerosità campionaria.

Per “*empirical size*” o livello empirico si intende una stima della probabilità di rifiutare H_0 quando è vera, ovvero della probabilità effettiva di commettere un errore del I tipo.

In questo studio H_0 è l'ipotesi di assenza di correlazione nella serie storica oggetto di analisi; il livello di significatività considerato è pari ad $\alpha=0.05$, dove α è la massima probabilità d'errore del I tipo.

Vogliamo che la *size* stimata non risulti distorta rispetto a quella nominale del 5%; un test che rifiuta l'ipotesi nulla, sbagliando, all'incirca nel 5% dei casi, sarà un test con capacità discriminatorie, tra ipotesi nulla ed alternativa, ottimali.

3.2.1 Livello empirico con distribuzioni asintotiche

Per valutare se la distribuzione asintotica delle statistiche fornisce una buona approssimazione della vera distribuzione dei test sotto l'ipotesi nulla, abbiamo ricavato il livello empirico dei test. A tal fine sono state generate $s=10000$ serie di dati iid distribuiti secondo una $N(0,1)$ al variare della numerosità campionaria $n=(20, 30, 50, 100, 200, 500)$, e sono stati calcolati i quattro test: CT con MSE, CT con deviazione standard stimata con l'equazione (3), LB e Monti.

Successivamente i risultati ottenuti sono stati confrontati con i valori critici ricavati dalle distribuzioni asintotiche dei test, facendo incrementare di un'unità i contatori relativi a ciascun metodo ogni volta che i rispettivi valori critici venivano superati. Con dati WN, tutti i test si approssimano ad un χ_m^2 , quindi, ad un livello di significatività $\alpha=0.05$, si rifiuta l'ipotesi nulla se il valore del test è maggiore del 95-esimo quantile di un chi-quadrato con m gradi di libertà.

Infine, si è ricavata la proporzione di statistiche test che superava il valore critico, calcolando i rapporti tra i valori dei contatori e il numero totale s di serie storiche simulate, ottenendo così una stima della probabilità effettiva (o empirica) dell'errore di I tipo dei test. Per facilitare l'interpretazione, il numero di volte che nelle 10.000 serie simulate è stata rifiutata erroneamente H_0 è stato espresso in percentuale.

I comandi R utilizzati compaiono in **Appendice A.1**, mentre i risultati delle simulazioni con $m = \sqrt{n}$ sono riportati in Tabella 1.1 e quelli con $m = n/3$ in Tabella 1.2.

Test/n	20	30	50	100	200	500
$m = \sqrt{n}$	4	5	7	10	14	22
CT(MSE)	5.41%	5.42%	5.11%	5.49%	5.9%	5.3%
CT(3)	8.43%	7.89%	6.66%	6.54%	6.24%	5.55%
LB	5.42%	5.63%	5.79%	6.18%	6.21%	5.33%
Monti	6.36%	5.55%	5.48%	5.50%	5.72%	5.08%

Tabella 1.1 Livello empirico dei test per dati WN calcolato utilizzando le distribuzioni asintotiche e con $m = \sqrt{n}$.

Test/n	20	30	50	100	200	500
$m = n/3$	7	10	17	33	67	167
CT(MSE)	6.68%	6.96%	6.74%	7.44%	8.56%	8.88%
CT(3)	11.68%	10.4%	9.64%	8.79%	9.27%	9.41%
LB	7.1%	7.58%	8.42%	8.75%	9.85%	9.79%
Monti	5.68%	5.78%	4.95%	4.01%	3.81%	2.67%

Tabella 1.2 Livello empirico dei test per dati WN calcolato utilizzando le distribuzioni asintotiche e con $m = n/3$.

Dall'osservazione dei risultati possiamo notare che quando la varianza è calcolata con l'equazione (3), in entrambe le tabelle l'ampiezza effettiva del test di Cauchy si discosta di molto dal valore teorico del 5%, mentre stimando la varianza con l'MSE, la *size* è più soddisfacente, specialmente per $m = \sqrt{n}$. Tuttavia, il livello empirico del CT con equazione (3) migliora al crescere della dimensione campionaria, abbassandosi in entrambi gli studi e avvicinandosi al 5% nella prima tabella per $n=500$.

I risultati delle simulazioni per il test di Ljung-Box sono abbastanza soddisfacenti nella prima tabella, mentre nella seconda si allontanano dal 5% all'aumentare di m .

La statistica Monti sembra avere complessivamente buone capacità discriminatorie tra H_0 e H_1 , con probabilità di commettere un errore del I tipo che si aggirano attorno al 5%, pur essendo sensibile a numerosità campionarie ridotte ($n=20$, Tab. 1.1).

Vogliamo ora verificare la presenza di correlazione nella serie dei residui nel caso di un modello ARMA stimato. In questo tipo di analisi, ha senso utilizzare solo l'equazione (3) per stimare la varianza, quindi qui procederemo calcolando un unico CT.

Come spiegato in precedenza nella parte teorica della tesi, se un modello del tipo (4) è correttamente specificato, i residui risulteranno indipendenti e identicamente distribuiti.

In questo paragrafo non ci occuperemo di valutare la bontà di adattamento del modello ai dati tramite l'analisi dei residui, ma ci limiteremo a considerare i residui stimati come una serie di dati iid su cui verificare la presenza o meno di correlazione. Rimandiamo i commenti circa l'adeguatezza di modelli stimati al paragrafo 3.3.2, con l'applicazione della variante del CT del Capitolo 2.

Procediamo dunque in modo analogo a prima per ricavare la *size* delle statistiche test usando le distribuzioni asintotiche.

Ho generato 10.000 serie da un modello AR(1) al variare dei valori assunti dal parametro autoregressivo pari a $\phi=(0.3, 0.6, 0.9)$; ho stimato il modello e ho estratto i residui. Ho proceduto calcolando i tre test CT, LB e Monti sui residui. In seguito ho confrontato i risultati ottenuti con i valori critici ricavati dalle distribuzioni asintotiche di ciascun test (facendo attenzione che in questo caso, trattandosi di residui, i gradi di libertà del chi-quadrato devono esser corretti da m a $m-p-q$, dunque $m-1$) e infine ho calcolato la proporzione di valori assunti dai test maggiori dei valori critici. Per facilità di interpretazione ho espresso in percentuale il numero di volte che nelle 10.000 serie simulate H_0 è stata rifiutata in modo errato.

Le probabilità di commettere un errore del I tipo dei test al variare di $n = (20, 30, 50, 100, 200, 500)$ sono riportate in Tabella 2.1 con $m = \sqrt{n}$ e in Tabella 2.2 con $m = n/3$.

	n	20	30	50	100	200	500
	$m = \sqrt{n}$	4	5	7	10	14	22
AR(1) $\phi=0.3$	CT(3)	9%	7.78%	7.63%	6.46%	6.54%	5.69%
	LB	4.63%	4.77%	4.96%	4.85%	5.6%	5.43%
	Monti	6.51%	5.84%	5.71%	4.89%	5.57%	5.04%
AR(1) $\phi=0.6$	CT(3)	8.81%	8.49%	7.14%	6.45%	6.04%	5.98%
	LB	4.76%	5.42%	4.81%	5.3%	5.68%	5.6%
	Monti	5.99%	5.81%	5.24%	5.49%	5.29%	5.41%
AR(1) $\phi=0.9$	CT(3)	8.56%	8.66%	7.39%	7.08%	6.56%	6.28%
	LB	5.5%	5.46%	5.36%	5.94%	5.78%	5.6%
	Monti	6.66%	6.32%	5.39%	5.45%	5.1%	5.33%

Tabella 2.1 Livello empirico dei test per residui estratti da modelli AR(1) stimati correttamente, calcolato utilizzando le distribuzioni asintotiche e con $m = \sqrt{n}$.

	n	20	30	50	100	200	500
	$m = n/3$	7	10	17	33	67	167
AR(1) $\phi=0.3$	CT(3)	10.96%	9.26%	8.91%	8.86%	8.83%	8.9%
	LB	5.35%	5.74%	6.61%	7.26%	8.7%	9.56%
	Monti	5.62%	5.55%	4.94%	4.17%	3.66%	2.78%
AR(1) $\phi=0.6$	CT(3)	11%	9.84%	9.2%	8.89%	8.87%	8.97%
	LB	5.63%	5.91%	7.04%	7.4%	8.74%	9.48%
	Monti	5.48%	5.19%	4.81%	4.09%	3.5%	2.61%
AR(1) $\phi=0.9$	CT(3)	10.91%	9.83%	9.3%	8.92%	8.96%	9.17%
	LB	6.94%	6.77%	7.27%	7.72%	8.66%	9.74%
	Monti	5.58%	5.33%	4.66%	3.84%	3.52%	2.67%

Tabella 2.2 Livello empirico dei test per residui estratti da modelli AR(1) stimati correttamente, calcolato utilizzando le distribuzioni asintotiche e con $m = n/3$.

I livelli empirici ottenuti dalle statistiche LB e Monti si discostano dal valore nominale del 5% per basse numerosità campionarie, ma generalmente migliorano all'aumentare di n .

Il test di Cauchy proposto in alcuni casi sembra avere probabilità di rifiutare erroneamente H_0 molto maggiori del 5%, ma in generale con l'incremento di n si stabilizza e risulta comparabile agli altri due test.

Deduciamo allora che le performance di tutti e tre i test migliorano al crescere di n ; infatti si può notare che nella Tabella 2.1 con $m = \sqrt{n}$ per $n=500$ le *size* di tutti i test sono vicine al 5% per ogni valore di ϕ .

Tuttavia, dalla Tabella 2.2 emerge anche che tutte le statistiche mostrano una certa sensibilità per valori elevati di m .

Fortunatamente, le moderne tecniche computazionali rendono questo problema quasi irrilevante poiché valori critici e p -value possono essere determinati dalla vera distribuzione delle statistiche test, tramite metodi Monte Carlo.

Inoltre, utilizzando la distribuzione Monte Carlo non c'è bisogno di preoccuparsi di decidere se usare l'equazione (3) o il MSE per stimare la varianza nel Cauchy test per processi WN; infatti, per ottenere risultati migliori, si raccomanda di usare l'equazione (3).

Non essendo del tutto soddisfatti degli errori del I tipo ottenuti nelle simulazioni precedenti con le distribuzioni asintotiche, sfruttando il metodo Monte Carlo nell'algoritmo proposto da Lin e McLeod (2006), nel prossimo paragrafo cercheremo di ottenere una valida approssimazione della distribuzione campionaria delle statistiche test e performance migliori.

Riportiamo alla fine di questo paragrafo dei grafici di confronto tra distribuzione asintotica (sotto H_0) e distribuzione empirica dei test per dati WN (Figura 1.1, 1.2, 1.3) e per serie di residui di AR(1) (Figura 2.1, 2.2, 2.3), per alcuni valori di m , n e ϕ .

Si può notare dalle figure che al crescere di n le due distribuzioni tendono a coincidere.

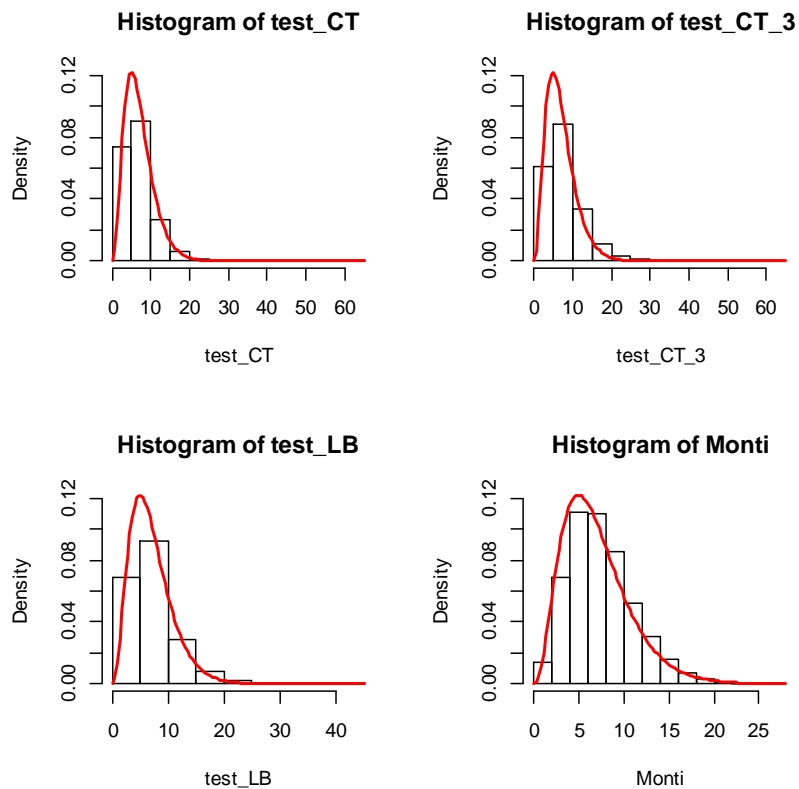


Figure 1.1 *Dati WN, $s=10000, n=20, m=n/3=7$*

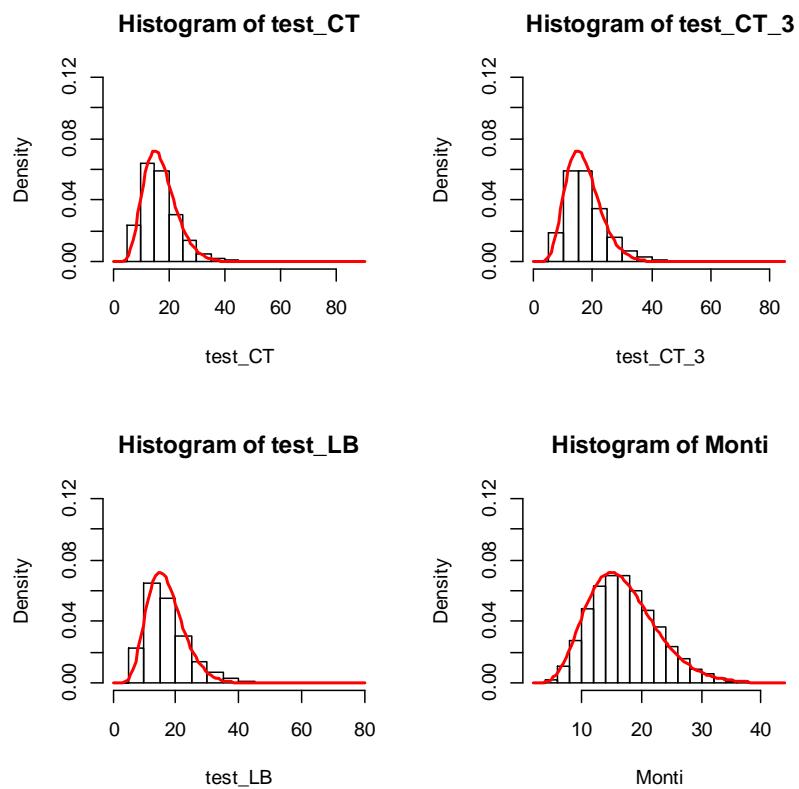


Figure 1.2 *Dati WN, $s=10000, n=50, m=n/3=17$*

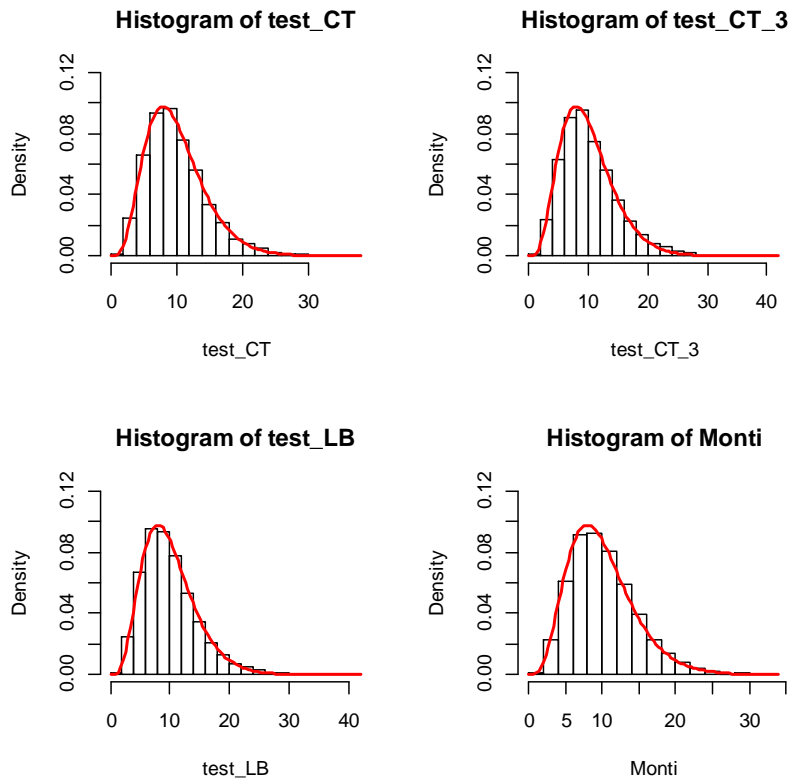


Figura 1.3 Dati WN, $s=10000$, $n=100$, $m=\sqrt{n}=10$

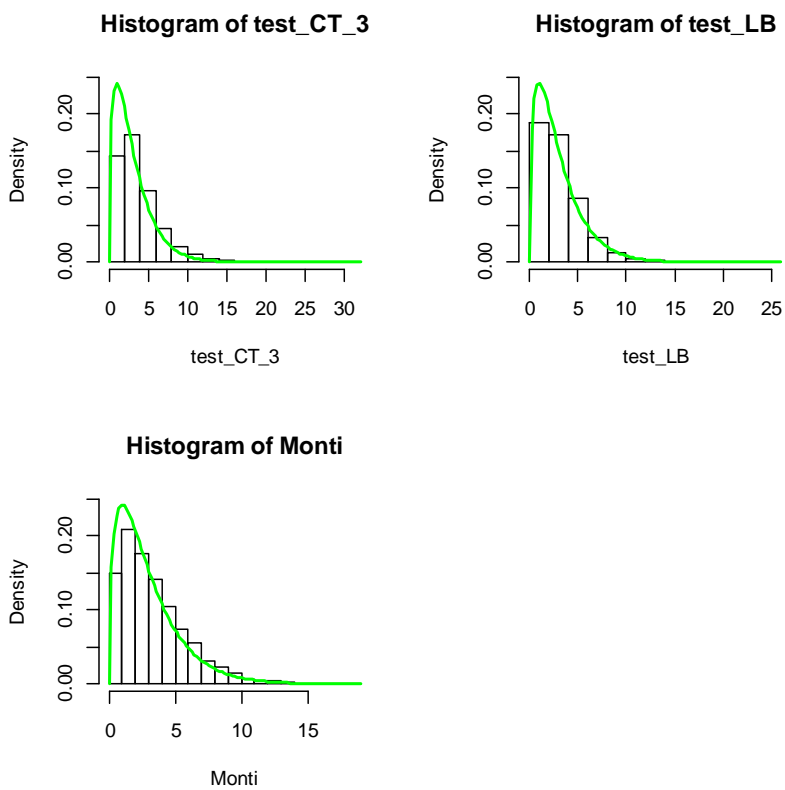


Figura 2.1 Dati da AR(1), $s=10000$, $\phi=0.3$, $n=20$, $m=\sqrt{n}=4$

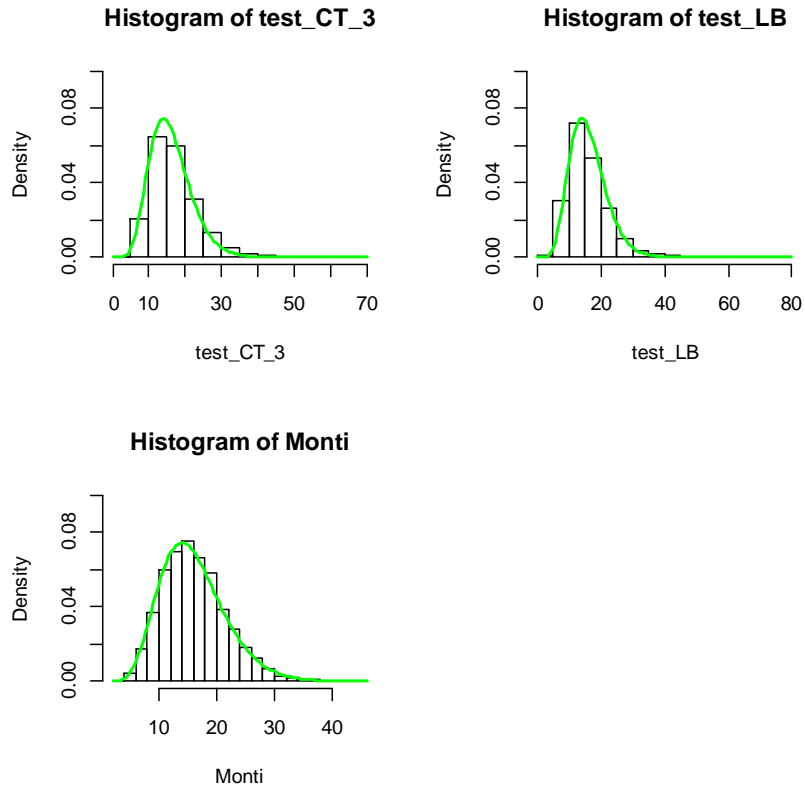


Figura 2.2 Dati da $AR(1)$, $s=10000$, $\phi=0.6$, $n=50$, $m=n/3=17$

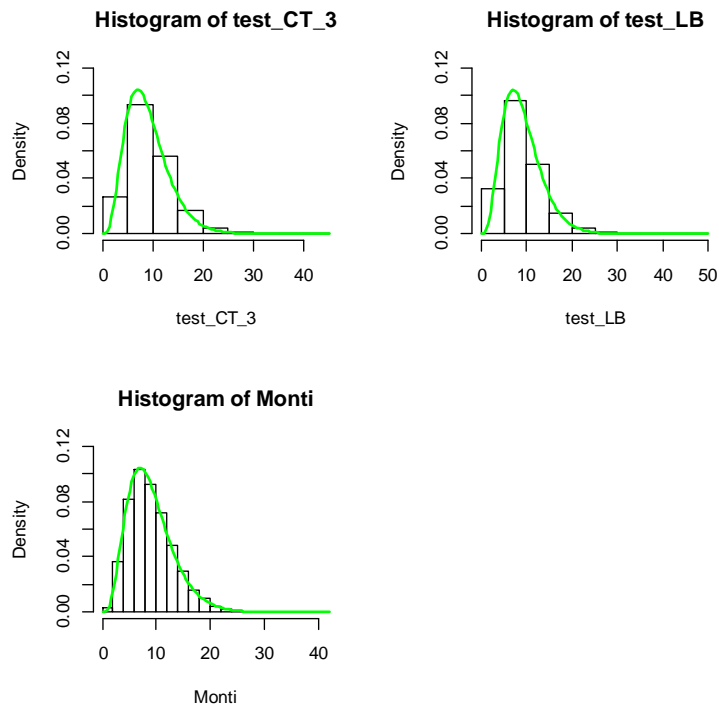


Figura 2.3 Dati da $AR(1)$, $s=10000$, $\phi=0.9$, $n=100$, $m=\sqrt{n}=10$

3.2.2 Livello empirico con metodo Monte Carlo

In questa sezione analizziamo il comportamento delle statistiche quando i p -value sono calcolati sfruttando l'algoritmo descritto da Lin e McLeod in "*Improved Peña-Rodriguez portmanteau test.*" (2006).

È raccomandato l'uso di un test Monte Carlo quando si ha a che fare con campioni di dimensione $n < 1000$.

L'algoritmo proposto è essenzialmente equivalente ad un bootstrap parametrico, che si rivela particolarmente utile quando si tratta di approssimare la funzione di distribuzione campionaria di uno stimatore o di una statistica d'interesse.

I passi da compiere rifacendosi a tale procedura sono i seguenti:

1. Data una serie storica, identificare il modello ARMA più appropriato ai dati, stimarlo e calcolare le statistiche test, T , sui residui, che si suppone siano iid.
2. Scegliere il numero di replicazioni Monte Carlo, B . Tipicamente $100 \leq B \leq 1000$.
3. Iniziare il bootstrap simulando un'altra serie da un modello ARMA con le stesse caratteristiche della serie originale usando come parametri i coefficienti stimati ottenuti al Passo 1 e, come innovazioni, i residui stimati sempre al Passo 1 ricampionati con reinserimento; dopo aver stimato anche questo modello e averne estratto i residui, calcolare le statistiche test, T^* , su di essi.
4. Ripetere il Passo 3 per B volte, contando il numero di volte k in cui il valore delle statistiche test T^* è maggiore o uguale a quello delle T ottenute al Passo 1.
5. Il p -value per i test è dato da $(k + 1)/(B + 1)$.
6. Rifiutare l'ipotesi nulla se il p -value è minore di un prestabilito livello di significatività (es. $\alpha=0.05$).

Il tutto deve essere ripetuto per ciascuna delle s serie considerate; infine, se si desidera ricavare la *size* empirica dei test, calcolare la proporzione di volte che nelle s simulazioni viene rifiutata l'ipotesi nulla.

Nelle nostre analisi simuliamo $s=1000$ serie indipendenti e utilizziamo $B=1000$ iterazioni bootstrap.

Poiché per numerosità campionarie elevate ($n=200$ e $n=500$) avevamo ottenuto risultati soddisfacenti, prossimi al 5%, applichiamo l'algoritmo solo a numerosità più basse.

Calcoliamo dunque la size empirica applicando il metodo appena esposto a 1000 serie di dati WN e processi AR(1) correttamente specificati, con parametro autoregressivo $\phi=(0.3, 0.6, 0.9)$.

I comandi utilizzati per simulare, ad esempio, 1000 serie storiche indipendenti da un AR(1) con parametro autoregressivo pari a 0.6 e composte da $n=48$ osservazioni sono riportati in **Appendice A.2**.

In Tabella 3.1 sono mostrati i risultati delle simulazioni con $m = \sqrt{n}$ e in Tabella 3.2 quelli con $m = n/3$, al variare di $n = (12, 24, 36, 48, 96)$.

	n	12	24	36	48	96
	$m = \sqrt{n}$	3	5	6	7	10
WN	CT(3)	5.3%	5%	4.2%	5.2%	4%
	LB	4.4%	6.3%	4.5%	5.7%	4.7%
	Monti	4.3%	5.8%	5.3%	5.1%	4.6%
AR(1) $\phi=0.3$	CT(3)	4.5%	5%	5.9%	5.5%	5.3%
	LB	4.6%	4.6%	5.4%	4.8%	4%
	Monti	3.9%	5.1%	5.8%	6.1%	5%
AR(1) $\phi=0.6$	CT(3)	3.2%	4.5%	5.6%	4.2%	4.5%
	LB	4.7%	6.2%	4.4%	4.5%	5.3%
	Monti	4.6%	5.5%	4.6%	4.6%	3.8%
AR(1) $\phi=0.9$	CT(3)	4.5%	5.1%	5.4%	4.1%	4.6%
	LB	5.6%	5.4%	5.2%	5.7%	5.3%
	Monti	5.8%	5.8%	5.7%	5.1%	4.8%

Tabella 3.1 *Livello empirico dei test per dati WN e modelli AR(1) stimati correttamente, calcolato utilizzando il metodo Monte Carlo e con $m = \sqrt{n}$.*

	n	12	24	36	48	96
	$m = n/3$	4	8	12	16	32
WN	CT(3)	4.7%	5.3%	5.1%	4.7%	4.5%
	LB	5.1%	5.2%	4.7%	4.8%	5.2%
	Monti	4.4%	4.6%	4.8%	4.6%	4.9%
AR(1) $\phi=0.3$	CT(3)	4.4%	3.6%	4.2%	5.5%	3.9%
	LB	5.3%	5%	4.4%	5.3%	5.1%
	Monti	4.2%	5.1%	4.2%	5.2%	4.5%
AR(1) $\phi=0.6$	CT(3)	3.9%	5.4%	5.8%	3.9%	4.5%
	LB	4.5%	5.7%	4.9%	5%	4%
	Monti	3.9%	5.7%	5%	5.2%	5.5%
AR(1) $\phi=0.9$	CT(3)	3.9%	4.7%	5.2%	6.2%	4.9%
	LB	6.8%	4.8%	3.8%	4.6%	4.6%
	Monti	5.4%	5.4%	4.3%	4.4%	5.2%

Tabella 3.2 Livello empirico dei test per dati WN e modelli AR(1) stimati correttamente, calcolato utilizzando il metodo Monte Carlo e con $m = n/3$.

Come ci aspettavamo e possiamo notare dalle Tabelle 3.1 e 3.2, tutte le statistiche test hanno un errore del I tipo soddisfacente quando i p -value sono calcolati attraverso il metodo Monte Carlo; al variare di ogni valore di ϕ e n , il livello empirico è, infatti, quasi sempre vicino al 5%, dando evidenza di un notevole miglioramento nella stima della *size* effettiva dei test applicando il bootstrap rispetto alle tabelle precedenti usando le distribuzioni asintotiche.

3.3 Potenza empirica dei test

Consideriamo ora la potenza empirica (*empirical power*, in inglese) dei test presi in esame.

La potenza empirica di un test statistico è la probabilità di rifiutare H_0 quando è falsa, ovvero è la probabilità di non commettere un errore del II tipo (accettare H_0 quando è falsa); la potenza del test sarà tanto maggiore quanto più piccolo è l'errore del II tipo, e viceversa.

La potenza empirica è dunque la capacità di un test statistico di riconoscere la falsità di H_0 quando essa è effettivamente falsa.

In questa sezione sottoponiamo a verifica ancora una volta lo stesso sistema d'ipotesi. Per ricavare le potenze è necessario mettersi sotto l'ipotesi alternativa H_1 ; ci aspettiamo di rifiutare H_0 molte volte e di ottenere valori il più alti possibile, prossimi al 100%, ossia che i test abbiano buone capacità discriminatorie di riconoscere la falsità dell'ipotesi nulla.

Nelle prossime pagine calcoleremo la potenza empirica dei test per modelli AR(1) utilizzando di nuovo sia distribuzioni asintotiche sia l'algoritmo MC, per poi passare alla potenza dei test di bontà di adattamento per ARMA sottostimati.

L'obiettivo è confermare la tesi degli autori dell'articolo che il test dello stimatore di Cauchy (e la sua variante) è più potente nel determinare la correlazione in piccoli campioni rispetto ai test LB e Monti.

3.3.1 Potenza empirica per AR(1)

Nel nostro primo studio analizziamo un modello AR(1) con $\phi=(0.3, 0.6, 0.9)$.

Per far ciò incrementiamo la numerosità campionaria n da 10 a 100 con passo pari a 10 e calcoliamo le statistiche test al ritardo $m = n/3$.

Iniziamo simulando 1000 serie e utilizzando le distribuzioni asintotiche per il calcolo delle potenze empiriche.

Abbiamo generato le serie con le caratteristiche desiderate e calcolato le statistiche test direttamente su di esse, invece che sui residui estratti dal modello stimato come fatto in precedenza. In questo modo ci mettiamo sotto l'ipotesi alternativa di dati non più iid, ma correlati e ci aspettiamo pertanto che i test rifiutino H_0 con probabilità elevate.

Si è proceduto confrontando gli esiti dei test con i valori critici delle rispettive distribuzioni asintotiche.

Abbiamo infine espresso in percentuale la proporzione con cui i test hanno rifiutato H_0 quando era falsa.

I risultati ottenuti sono riportati nella Tabella 4.1 e nei grafici in Figura 3.1, 3.2 e 3.3.

La potenza empirica assume valori compresi tra 0 e 1, ma per comodità qui sarà indicata in termini percentuali.

	n	10	20	30	40	50	60	70	80	90	100
	m=n/3	3	7	10	13	17	20	23	27	30	33
AR(1) $\phi=0.3$	CT(3)	17.8%	21.6%	22.6%	24.4%	26.2%	30%	31.4%	32.9%	35.7%	35.7%
	LB	4.7%	13.9%	17.7%	21.7%	25.4%	29.7%	31.4%	35.2%	39.2%	40.2%
	Monti	5.7%	9.2%	11.4%	14%	15.3%	17.3%	17.1%	19.9%	21.7%	22.8%
AR(1) $\phi=0.6$	CT(3)	41.8%	57.5%	71%	75.8%	82.7%	87.2%	88%	92.2%	93.3%	95.8%
	LB	8%	39.2%	57.3%	72%	82.7%	86.4%	91.2%	94.2%	96.7%	97.4%
	Monti	12.7%	28.8%	44.6%	57.4%	67.8%	76.8%	80.3%	84.8%	90.6%	92.9%
AR(1) $\phi=0.9$	CT(3)	80.7%	93.7%	98.1%	99.5%	99.9%	99.9%	100%	100%	100%	100%
	LB	15.6%	73.8%	93%	98.1%	99.6%	99.8%	100%	100%	100%	100%
	Monti	19.8%	61.7%	87.2%	96.3%	99.1%	99.7%	99.9%	100%	100%	100%

Tabella 4.1 Potenza empirica dei test per modelli AR(1), calcolata utilizzando le distribuzioni asintotiche.

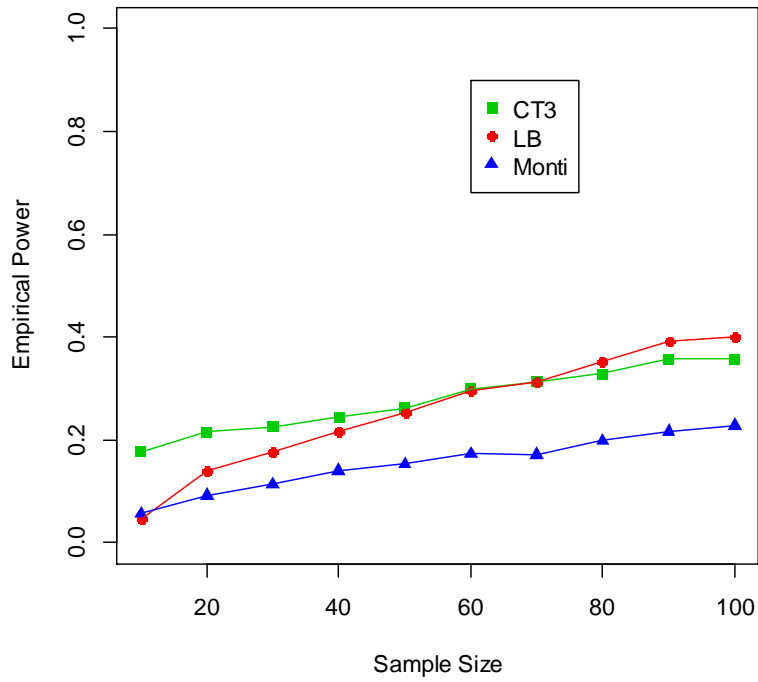


Figura 3.1 Potenza empirica dei test per AR(1) con $\phi=0.3$, utilizzando le rispettive distribuzioni asintotiche.

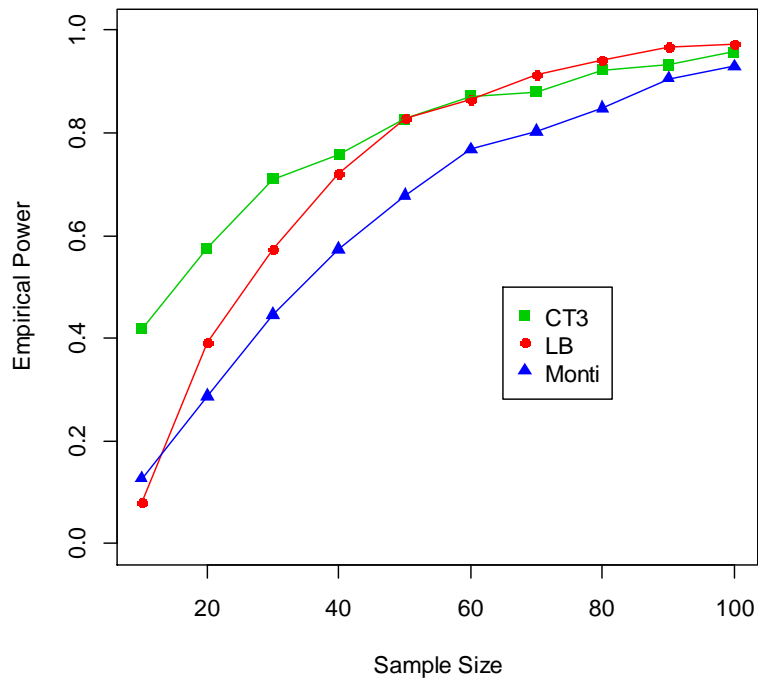


Figura 3.2 Potenza empirica dei test per AR(1) con $\phi=0.6$, utilizzando le rispettive distribuzioni asintotiche.

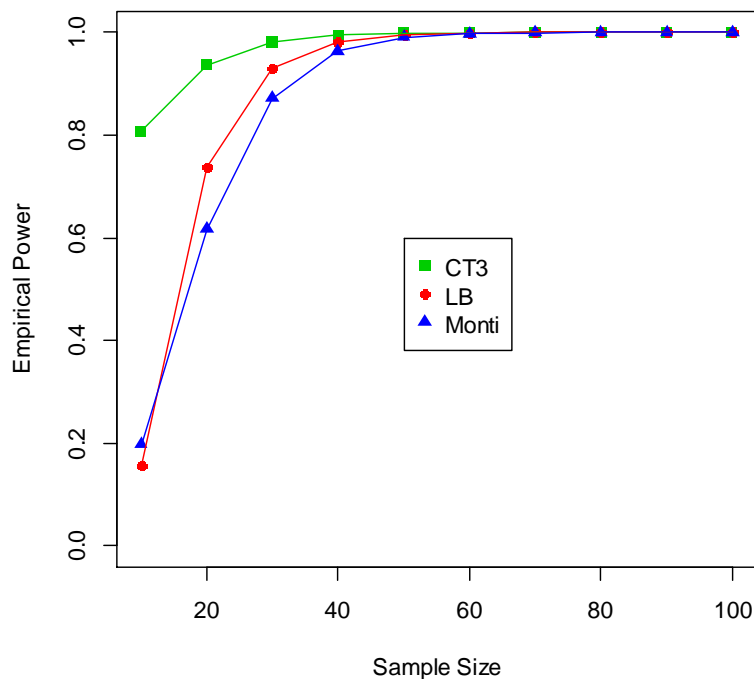


Figura 3.3 Potenza empirica dei test per AR(1) con $\phi=0.9$, utilizzando le rispettive distribuzioni asintotiche.

Come si nota dalla Tabella 4.1 e dalle Figure 3.1, 3.2 e 3.3, dal comportamento asintotico delle statistiche test sotto l'ipotesi alternativa si ricava che per $\phi=0.3$ e n da 10 a 40 il CT registra potenze maggiori rispetto agli altri due; da $n=50$ in poi il suo comportamento è simile al LB, fino ad essere oltrepassato da questo a $n=80$; entrambi i test hanno potenze superiori a Monti, arrivando per $n=100$ a riconoscere la falsità di H_0 con probabilità attorno al 40%, contro un 23%.

Ricordando che l'autocorrelazione di un processo AR(1) è pari a $\rho_k = \phi^k$, aumentando il valore di ϕ da 0.3 a 0.6 e a 0.9, aumenterà anche la correlazione nella serie.

Con parametro autoregressivo pari a 0.6, infatti, tutte le statistiche rifiutano molto più spesso H_0 ; il CT ha ancora potenza maggiore rispetto agli altri fino a $n=40$, per $n=50$ e 60 i test CT e LB sono simili e maggiori di Monti, da 70 in poi LB supera CT fino a che per $n=100$ le tre statistiche riportano circa la stessa potenza attorno al 95%.

Infine, per $\phi=0.9$ e numerosità campionaria bassa ($n=10$) il CT risulta di molto più potente rispetto agli altri due test, arrivando a rifiutare H_0 falsa per l'80% delle volte, contro il 20% degli altri.

Da qui è ben evidente la maggiore potenza del test proposto in piccoli campioni, caratteristica che vogliamo sottolineare in questo lavoro.

A partire dalla numerosità $n=20$ anche le statistiche LB e Monti hanno potenze elevate e sempre più simili al CT, tanto che già da $n=50$ in poi tutti e tre i test riconoscono la falsità di H_0 con probabilità del 100%.

Facciamo ora la stessa indagine, ma calcolando i p -value delle statistiche utilizzando l'algoritmo di Lin e McLeod.

Replichiamo il procedimento del metodo descritto al paragrafo 3.2.2, con l'unica differenza che al Passo 1 i test T si calcolano direttamente sulle serie generate e non sui residui, mettendoci sotto l'ipotesi alternativa H_1 di dati correlati; il bootstrap del Passo 3 rimane invariato (sotto H_0): si calcolano i test T^* sempre sui residui stimati, che si suppone siano incorrelati.

Gli esiti degli esperimenti sono i seguenti:

	n	10	20	30	40	50	60	70	80	90	100
	m=n/3	3	7	10	13	17	20	23	27	30	33
AR(1) $\phi=0.3$	CT(3)	14.2%	17%	19.8%	19.8%	22.9%	25.1%	26.4%	26%	29.2%	28%
	LB	10.5%	15.3%	20.7%	21.4%	27.8%	27.8%	30.1%	31.3%	34.5%	35.3%
	Monti	11.2%	15.8%	14.9%	16.4%	21.5%	22.1%	22.9%	25.5%	24.8%	28%
AR(1) $\phi=0.6$	CT(3)	34.2%	52%	64%	71.3%	78.6%	85.3%	85.4%	91.2%	91.6%	94%
	LB	16.4%	44.7%	62.9%	71.6%	80.6%	87.1%	90.4%	94.1%	95.2%	97%
	Monti	19.6%	36.6%	53%	63.8%	75.3%	82.7%	84.9%	92.3%	92.9%	94.8%
AR(1) $\phi=0.9$	CT(3)	76.7%	93.5%	97.4%	99.6%	99.8%	99.8%	100%	100%	100%	100%
	LB	27.6%	77.6%	93%	98.2%	99.6%	99.8%	100%	100%	100%	100%
	Monti	29.4%	70.9%	88.7%	96.6%	99.2%	99.8%	100%	100%	100%	100%

Tabella 4.2 Potenza empirica dei test per modelli AR(1), calcolata utilizzando l'algoritmo MC.

Le Figure 4.1, 4.2 e 4.3 rappresentano graficamente le potenze empiriche della Tabella 4.2 in funzione di n .

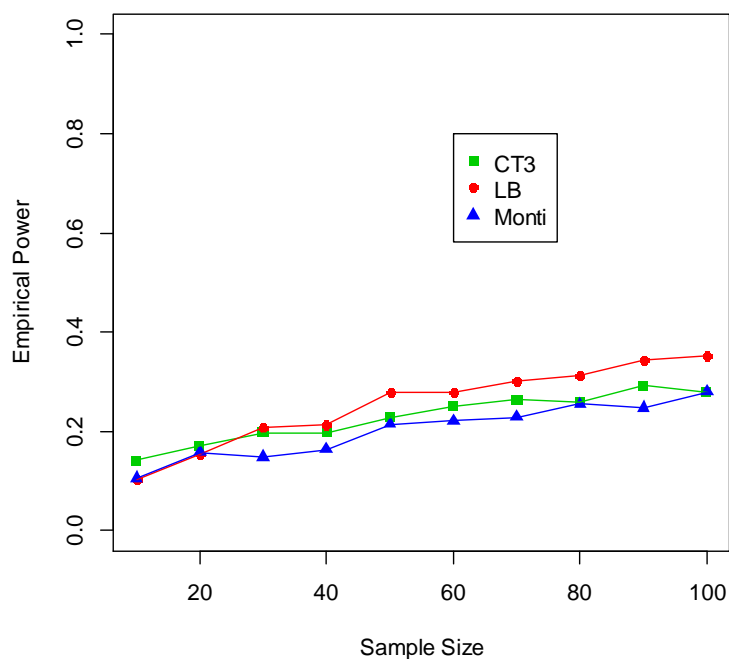


Figura 4.1 Potenza empirica dei test per AR(1) con $\phi=0.3$, utilizzando l'algoritmo Monte Carlo.

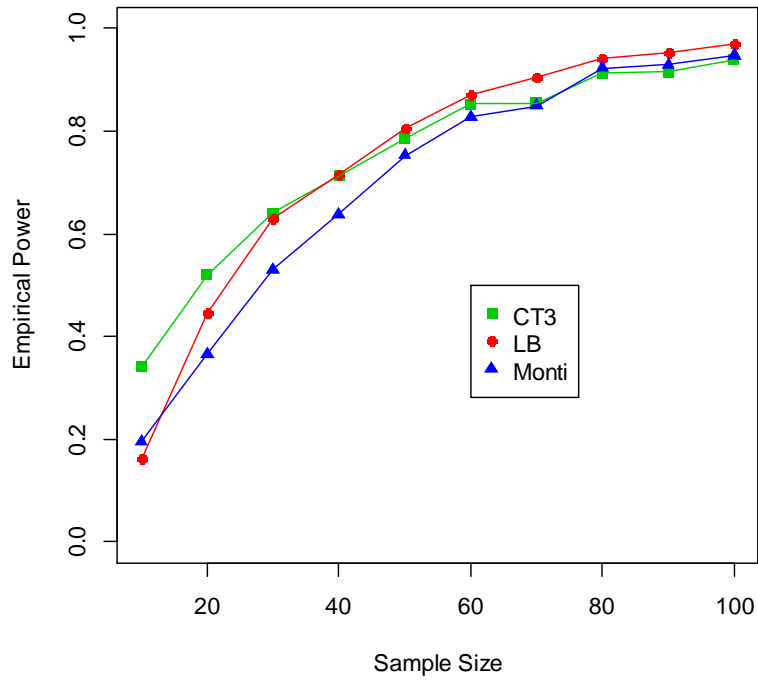


Figura 4.2 Potenza empirica dei test per $AR(1)$ con $\phi=0.6$, utilizzando l'algoritmo Monte Carlo.

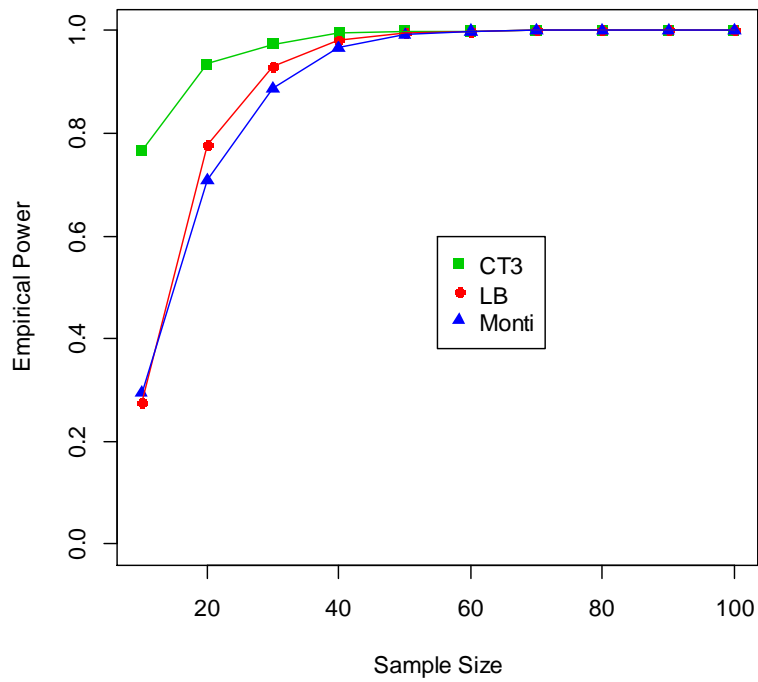


Figura 4.3 Potenza empirica dei test per $AR(1)$ con $\phi=0.9$, utilizzando l'algoritmo Monte Carlo.

Dalla Tabella 4.2 e dalle Figure 4.1, 4.2 e 4.3 possiamo trarre informazioni analoghe a quelle della Tabella 4.1; il CT ha una maggiore capacità di riconoscere la falsità di H_0 per piccoli campioni rispetto alle altre due procedure portmanteau, mentre tutte e tre le statistiche sono comparabili per dimensioni campionarie sempre più grandi.

È da ribadire qui che sono stati scelti proprio questi esempi per mostrare le situazioni in cui il CT proposto ha una buona performance.

Ad esempio, per processi MA sottostimati, test basati sulle autocorrelazioni parziali (come il test di Monti) generalmente hanno prestazioni migliori del CT; tuttavia, una sua versione costruita con le correlazioni parziali potrebbe competere in questo caso.

3.3.2 Potenza empirica per ARMA sottostimati

Vogliamo ora ricavare la potenza empirica dei test usati per valutare per la bontà di adattamento di un modello ARMA.

Come descritto nell'Introduzione del Capitolo 2, una parte fondamentale del controllo dell'adeguatezza del modello ai dati consiste in un'attenta analisi della serie dei residui. Ricordiamo, infatti, che se il modello è stato correttamente identificato e stimato, i residui dovranno risultare incorrelati; se invece il modello ARMA è sottostimato, le autocorrelazioni si discosteranno da zero e assumeranno valori prossimi a ± 1 .

Verificando la presenza di autocorrelazione nei residui, si può quindi dedurre se il modello stimato rappresenta adeguatamente il processo generatore dei dati oppure no.

In questa sezione, come fatto nella precedente, vogliamo esaminare la potenza dei vari test, confrontando la loro capacità di riconoscere la falsità dell'ipotesi di incorrelazione, dal momento che stiamo lavorando con dati che sappiamo a priori essere dipendenti proprio per nostra costruzione.

Per fare ciò, simuleremo due processi e vi adatteremo un modello sbagliato, sottostimandoli; i test di correlazione che compareremo sono la variante del CT per modelli ARMA sottostimati descritta al Capitolo 2, il test LB e quello di Monti.

Ancora una volta, miriamo a confermare la tesi dell'articolo che il CT è più potente degli altri per piccoli campioni.

Nel primo esperimento, la potenza delle tre statistiche è studiata in funzione di n che varia da 10 a 100 con passo pari a 10 e ritardo $m = n/3$; i p -value sono ottenuti con l'algoritmo di Monte Carlo.

Generiamo $s=1000$ serie indipendenti da un processo AR(2) con parametri autoregressivi $\phi_1 = 0.1$ e $\phi_2 = 0.8$, e lo stimiamo non correttamente adattandovi un AR(1). Estraiamo poi i residui che, provenendo da un modello sottostimato, saranno correlati; ci siamo dunque messi sotto l'ipotesi alternativa e calcoliamo su di essi i test d'interesse.

Per quanto riguarda il bootstrap, consideriamo un numero di replicazioni $B=500$ e non più 1000 per ridurre difficoltà legate al tempo e al carico computazionale richiesto da tale procedura. In ciascun'iterazione abbiamo simulato una nuova serie da un modello AR(1) con parametro ϕ pari a quello stimato prima e con innovazioni i residui stimati sempre nel modello precedente ma ricampionati, in modo che perdessero la loro correlazione; l'abbiamo stimata correttamente, estratto i residui e calcolato i test su questi. Nel bootstrap siamo infatti ancora sotto l'ipotesi nulla di dati incorrelati.

Per i comandi del programma R usato qui rimandiamo ancora una volta all'**Appendice A.3**, mentre i risultati degli esperimenti si trovano in Tabella 5.1, con relativo grafico in Figura 5.1.

	n	10	20	30	40	50	60	70	80	90	100
	m=n/3	3	7	10	13	17	20	23	27	30	33
AR(2) stimati come AR(1) $\phi_1=0.1$ $\phi_2=0.8$	CT(3) variant	4.7%	32.6%	64.1%	80.3%	91.1%	95.5%	97.7%	98.9%	99.3%	99.6%
	LB	3%	33.3%	68.3%	84%	93.8%	97%	98.3%	99.4%	99.8%	99.9%
	Monti	4.8%	28.7%	61.7%	80.1%	92.2%	95.6%	97.7%	99.1%	99.6%	99.9%

Tabella 5.1 Potenza empirica dei test per modelli AR(2) sottostimati, calcolata utilizzando l'algoritmo MC.

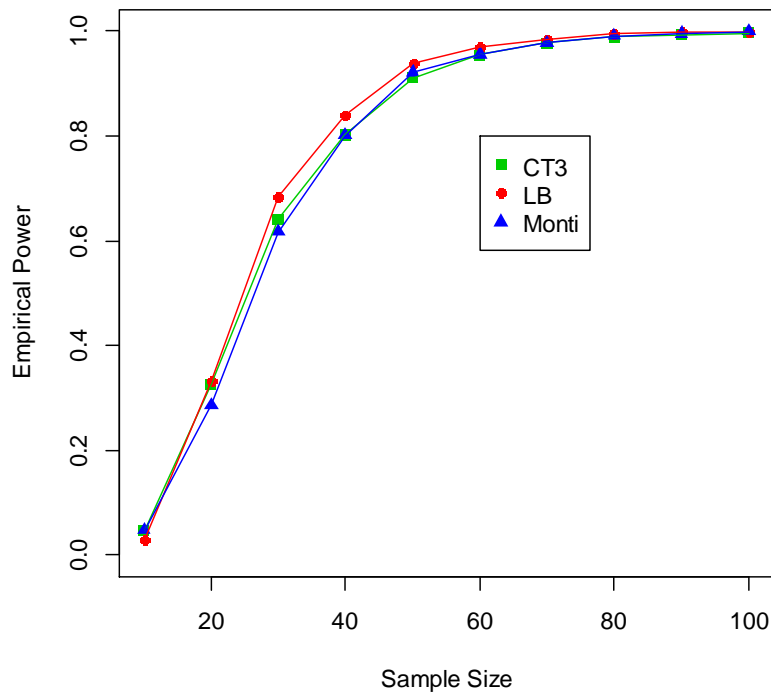


Figura 5.1 Potenza empirica per AR(2) sottostimati come AR(1), $\phi_1 = 0.1$ e $\phi_2 = 0.8$.

Da questi risultati possiamo notare che per un numero ridotto di osservazioni ($n=10$) la potenza del CT è maggiore rispetto a LB e prossima a Monti; via via che la numerosità campionaria, n , aumenta il CT ha un andamento simile a LB, venendo anche superato da quest'ultimo, ed entrambi i test dominano Monti, che è solitamente migliore nel determinare componenti sottostimate di processi a media mobile. Da $n=50$ in poi, le tre statistiche hanno prestazioni pressoché uguali e rifiutano H_0 da oltre il 90% a circa il 100% delle volte, riconoscendo nella quasi totalità dei casi che i residui sono correlati e dunque che il modello da noi stimato non spiega adeguatamente la reale struttura di dipendenza dei dati.

Ripetiamo ora lo stesso procedimento per esplorare la performance della potenza dei tre test in un modello ARMA con ordine maggiore rispetto al caso appena affrontato.

I dati sono generati da un processo ARMA(2,1) con parametri $\phi_1 = 0.2$, $\phi_2 = 0.7$ e $\theta = -0.5$, che viene sottostimato sempre con un AR(1).

In questa seconda analisi, le statistiche test sono calcolate di nuovo usando l'algoritmo di Monte Carlo con 500 iterazioni bootstrap e numerosità campionaria che varia da $n=10$ a $n=100$ con passo 10, ma questa volta al ritardo $m = \sqrt{n}$.

Ecco le potenze ottenute:

	n	10	20	30	40	50	60	70	80	90	100
	$m=\sqrt{n}$	3	4	5	6	7	8	8	9	9	10
ARMA (2,1) stimati come AR(1) $\phi_1=0.2$ $\phi_2=0.7$ $\theta=-0.5$	CT(3) variant	3%	13.3%	39%	58%	73.2%	85.9%	88.2%	94.3%	96.6%	97.7%
	LB	2.7%	12.4%	43.8%	63.6%	77.7%	89.5%	92.1%	95.8%	97.8%	98%
	Monti	3.5%	8.4%	34.2%	54.3%	71.2%	84.8%	87.8%	94.1%	96.3%	97.7%

Tabella 5.2 Potenza empirica dei test per modelli ARMA(2,1) sottostimati, calcolata utilizzando l'algoritmo MC.

La Figura 5.2 mostra graficamente i risultati raggiunti.

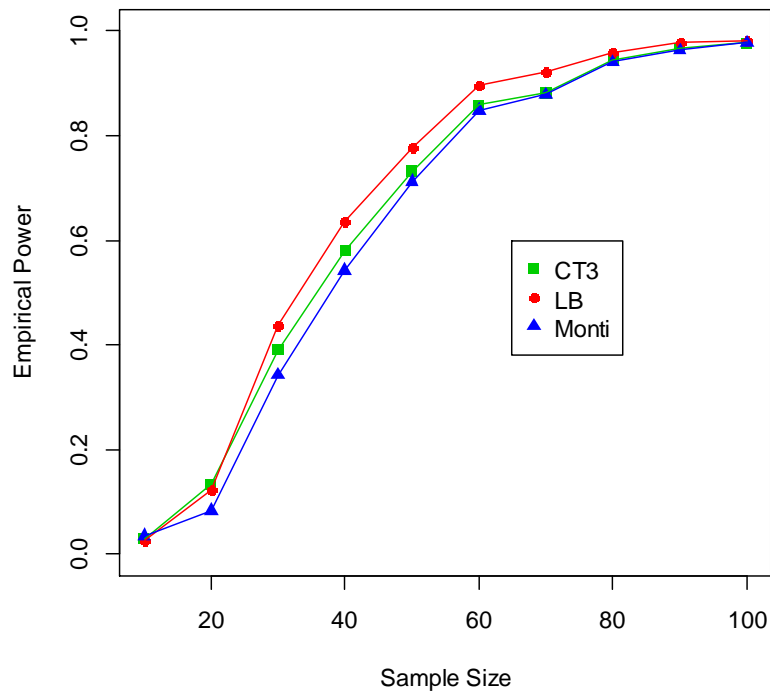


Figura 5.2 Potenza empirica per $ARMA(2,1)$ sottostimati come $AR(1)$, $\phi_1 = 0.2$ e $\phi_2 = 0.7$, $\theta = -0.5$.

Anche in questo studio la variante del test basato sullo stimatore di Cauchy risulta il metodo più potente per alcune piccole numerosità campionarie ($n=20$), è paragonabile a LB per valori intermedi di n ed eventualmente superato da questo al crescere di n .

Da $n=50$ in poi le tre statistiche assumono circa gli stessi valori, arrivando a riconoscere quasi al 100% che l'ipotesi di incorrelazione dei residui non viene rispettata e quindi che il modello da noi sottostimato non si adatta bene ai dati.

Infine, per sottolineare ulteriormente le prestazioni del test di Cauchy, è stato condotto un ultimo esperimento riassuntivo e comprensivo di vari metodi adottati nelle simulazioni precedenti.

Il comportamento della statistica in piccoli campioni è sintetizzato in Tabella 6.

Abbiamo generato 1000 serie di lunghezza $n=18$ da un processo AR(2) con parametri $\phi_1 = 0.3$ e $\phi_2 = 0.6$; per ciascuna serie simulata sono stati calcolati i tre test CT, LB e Monti.

Per prima cosa abbiamo verificato la presenza di correlazione nei processi, ripetendo il procedimento descritto al paragrafo 3.3.1 per ricavare la potenza empirica con algoritmo Monte Carlo, applicandolo in questo caso ad un AR(2). I risultati sono riportati nella prima colonna della Tabella 6.

Possiamo vedere chiaramente che il Cauchy test ha la più elevata capacità di rilevare la correlazione in un processo simulato autoregressivo e per n ridotto. Rifiuta H_0 quando è falsa e riconosce la presenza di correlazione nei dati per oltre l'80% delle volte, contro il 54% di LB e 39% di Monti.

In seconda analisi, abbiamo adattato ai nostri dati un modello sbagliato AR(1), come effettuato nel paragrafo 3.3.2. Le potenze empiriche ottenute compaiono nella seconda colonna della Tabella 6. Notiamo che il test proposto è comparabile favorevolmente agli altri metodi.

Infine, abbiamo stimato il corretto modello AR(2) e abbiamo ricavato il livello empirico dei test, sfruttando quanto esposto al paragrafo 3.2.2. Gli esiti relativi a questi errori del I tipo appaiono nell'ultima colonna della tabella; anche in questo caso, tutti e tre i test sono soddisfacenti: sbagliano rifiutando H_0 vera attorno al valore nominale del 5%.

Metodo	Potenza empirica	Potenza stimando un AR(1)	Livello empirico
CT	82%	11.5%	5%
LB	54%	13%	6%
Monti	39%	10.3%	4%

Tabella 6 Potenze empiriche e size per un AR(2), $n=18$, $m=6$.

Conclusioni

In questa tesi sono stati introdotti due nuovi test: uno di tipo portmanteau per la correlazione e l'altro per valutare la bontà di adattamento di modelli ARMA sottostimati. Entrambe le statistiche si basano sullo stimatore di Cauchy per l'autocorrelazione.

La distribuzione asintotica delle statistiche è stata quantificata e poi studiata via simulazione.

Gli esperimenti condotti hanno dimostrato che l'uso delle distribuzioni asintotiche può portare a commettere errori del I tipo per piccole dimensioni campionarie. Per questo si raccomanda l'utilizzo di metodi Monte Carlo come quello descritto da Lin e McLeod (2006) per determinare valori critici e p -value, la cui applicazione nelle simulazioni ha riportato un notevole miglioramento nella stima della *size* effettiva dei test.

Per aumentare ulteriormente la performance delle potenze, si suggerisce inoltre di stimare il termine di varianza del CT con l'equazione (3).

Negli studi effettuati il CT e la sua variante sono stati confrontati con il test di Ljung-Box e con quello di Monti (1994); in generale, le nuove statistiche proposte si sono dimostrate le più efficaci nel riconoscere la presenza di correlazione significativa in piccoli campioni.

Complessivamente, i risultati ottenuti in questo lavoro sono in linea con quelli suggeriti dall'articolo di riferimento Gallagher, Fisher e Shen (2014), da cui muove principalmente l'intera analisi.

Come i test basati sulla funzione di autocorrelazione campionaria dei residui, il CT proposto è ritenuto dagli autori migliore nel riconoscere processi autoregressivi sottostimati rispetto a componenti a media mobile mancanti; una sua versione realizzata con le correlazioni parziali di Cauchy, simile al test di Monti, sarebbe forse più potente anche per modelli MA.

Inoltre si potrebbe costruire un test asimmetrico basato sulla matrice delle correlazioni di Cauchy simile a quelli proposti da Peña e Rodríguez (2002 e 2006), Mahdi e McLeod (2012) o Fisher e Gallagher (2012). Un test di questo tipo potrebbe essere anche migliore di quelli introdotti in questo lavoro.

Appendice

A.1 Codice R per livello empirico con distribuzioni asintotiche e dati WN

```
library("portes")

s=10000    #n. di simulazioni

n=50      #lunghezza della serie   t=1,2,...,n

#scelgo il lag

m=round(n/3)
#oppure
m=round(sqrt(n)) #lag temporale k=1,2,...,m

test_CT=NULL
test_CT_3=NULL
test_LB=NULL
Monti=NULL

for(i in 1:s){
    #serie simulate
    x=rnorm(n)

    lambda=NULL
    lambda_3=NULL

    S=sign(x)

    sigma=var(x)

    sigma_3=NULL
    p1=NULL

    for(k in 1:m){
p1[k]=(sum(x[(k+1):n]*abs(x[1:(n-k)])*S[1:(n-k)]))/(sum(x[1:(n-k)]^2))
sigma_3[k]=sum((x[(k+1):n]- p1[k] * x[1:(n-k)])^2)/(n-k)
    }
}
```

```

#calcolo statistiche test:

#CT_MSE
for(k in 1:m){
lambda[k]=(sum(x[(k+1): n]* S[1:(n-k)]))/(sqrt((n-k)*sigma))
}

test_CT[i]=sum((lambda)^2)

#CT_3
for(k in 1:m){
lambda_3[k]=(sum(x[(k+1): n]* S[1:(n-k)]))/(sqrt((n-k)*sigma_3[k]))
}

test_CT_3[i]=sum((lambda_3)^2)

#LB
test_LB[i]= LjungBox(x, lags=m)[2]

#Monti
pacf=pacf(x, lag.max=m, plot=F)$acf #estraggo

somma=NULL

for(k in 1:m){
somma[k]=(pacf[k]^2/(n-k))
}

Monti[i]=n*(n+2)*sum(somma)
}

#valore critico ricavato dalla distribuzione asintotica del test
val_test=qchisq(0.95,m)

count_CT=0 # % di Rifiuto
count_CT_3=0
count_LB=0
count_M=0

for(i in 1:s){
if(test_CT[i] > val_test)
{count_CT=count_CT+1}

```

```

if(test_CT_3[i] > val_test)
{count_CT_3=count_CT_3+1}
if(test_LB[i] > val_test)
{count_LB=count_LB+1}
if(Monti[i] > val_test)
{count_M=count_M+1}
      }

#percentuale delle volte che rifiuto erroneamente H0
perc_CT=count_CT/s*100
perc_CT_3=count_CT_3/s*100
perc_LB=count_LB/s*100
perc_M=count_M/s*100

print(perc_CT)
print(perc_CT_3)
print(perc_LB)
print(perc_M)

```

A.2 Codice R per livello empirico con algoritmo MC e dati AR(1)

```

library("portes")
library("parallel")

s=1000 #n. di simulazioni

n=48 #lunghezza della serie t=1,2,...,n

#scelgo il lag
m=round(n/3)
#oppure
m=round(sqrt(n)) #lag temporale k=1,2,...,m

p_value_CT_3=NULL
p_value_LB=NULL
p_value_M=NULL
test_CT_3=NULL
test_CT_3_star=NULL
test_LB=NULL
test_LB_star=NULL

```

```

Monti=NULL
Monti_star=NULL

for(i in 1:s){

    #serie simulate

    phi <- 0.6

    #genero i dati
    y <- arima.sim(list(order=c(1,0,0), ar=phi),n=n)

    #stimo il modello
    mod = arima(y, order = c(1, 0, 0), method="ML")

    #estraggo i residui
    x=residuals(mod)

    #calcolo i test sui residui
    lambda_3=NULL

    S=sign(x)

    sigma_3=NULL
    p1=NULL

    for(k in 1:m){
p1[k]=(sum(x[(k+1):n]*abs(x[1:(n-k)])*S[1:(n-k)]))/(sum(x[1:(n-k)]^2))
sigma_3[k]=sum((x[(k+1):n]- p1[k] * x[1:(n-k)])^2)/(n-k)

lambda_3[k]=(sum(x[(k+1): n] * S[1:(n-k)]))/(sqrt((n-k)*sigma_3[k]))
    }

    #calcolo statistica test:

    #CT_3
    test_CT_3[i]=sum((lambda_3)^2)

    #LB
    test_LB[i]= LjungBox(x, lags=m)[2]    #statistica test

```

```

#Monti
pacf=pacf(x, lag.max=m, plot=F)$acf #estraggo

somma=NULL

for(k in 1:m){
  somma[k]=(pacf[k]^2/(n-k))
}

Monti[i]=n*(n+2)*sum(somma)

#inizio ciclo MC (bootstrap)
cont_star_CT_3=0
cont_star_LB=0
cont_star_M=0

#estraggo il parametro stimato in mod
phi_star=mod$coef[1]

B=1000
for(j in 1:B){

#simulo lo stesso modello con parametro stimato prima e ricampionando
i residui
y_star=arima.sim(list(order=c(1,0,0),ar=phi_star),innov=sample(x),n=n)

#stimo il modello
mod_star = arima(y_star, order = c(1, 0, 0), method="ML")

#estraggo i residui e calcolo i test su di essi
x_star =residuals(mod_star)

lambda_3_star = NULL

S_star=sign(x_star)

sigma_3_star=NULL
p1_star=NULL

for(k in 1:m){

```



```

p1_star[k]=(sum(x_star[(k+1):n]*abs(x_star[1:(n-k)])*S_star[1:(n-
k)]))/(sum(x_star[1:(n-k)]^2))

sigma_3_star[k]=sum((x_star[(k+1):n]-p1_star[k]*x_star[1:(n-
k)]^2)/(n-k)

lambda_3_star[k]=(sum(x_star[(k+1):n]*S_star[1:(n-k)]))/(sqrt((n-
k)*sigma_3_star[k]))
    }

#CT_3
test_CT_3_star[j] = sum((lambda_3_star)^2)

if (test_CT_3_star[j] >= test_CT_3[i]){
cont_star_CT_3=cont_star_CT_3 + 1}

#LB
test_LB_star[j]= LjungBox(x_star, lags=m)[2]

if (test_LB_star[j] >= test_LB[i]){
cont_star_LB=cont_star_LB + 1}

#Monti
pacf=pacf(x_star, lag.max=m, plot=F)$acf #estraggo

somma_star=NULL

for(k in 1:m){
    somma_star[k]=(pacf[k]^2/(n-k))
    }
Monti_star[j]=n*(n+2)*sum(somma_star)

if (Monti_star[j] >= Monti[i]){
    cont_star_M=cont_star_M + 1}

} #fine ciclo MC

#calcolo i p_value:
p_value_CT_3[i] = (cont_star_CT_3 + 1)/(B+1)
p_value_LB[i]= (cont_star_LB + 1)/ (B+1)
p_value_M[i]= (cont_star_M + 1)/(B+1)
}

```

```

count_pv_CT_3=0 # % di Rifiuto
count_pv_LB=0
count_pv_M=0

for(i in 1:s){
if(p_value_CT_3[i] < 0.05) #rifiuto
{count_pv_CT_3=count_pv_CT_3 +1}
if(p_value_LB[i] < 0.05) #rifiuto
{count_pv_LB=count_pv_LB +1}
if(p_value_M[i] < 0.05) #rifiuto
{count_pv_M=count_pv_M +1}
}

#percentuale delle volte che rifiuto H0 quando è vera
perc_pv_CT_3=count_pv_CT_3/s*100
print(perc_pv_CT_3)

perc_pv_LB=count_pv_LB/s*100
print(perc_pv_LB)

perc_pv_M=count_pv_M/s*100
print(perc_pv_M)

```

A.3 Codice R per potenza empirica per dati AR(2) sottostimati con AR(1)

```

library("portes")
library("parallel")

s=1000 #n. di simulazioni

n=60 #lunghezza della serie t=1,2,...,n

#scelgo il lag
m=round(n/3)
#oppure
m=round(sqrt(n)) #lag temporale k=1,2,...,m

p_value_CT_3=NULL
p_value_LB=NULL
p_value_M=NULL

```

```

test_CT_3=NULL
test_CT_3_star=NULL
test_LB=NULL
test_LB_star=NULL
Monti=NULL
Monti_star=NULL

for(i in 1:s){

    #serie simulate sotto H1
    phi1 <- 0.1
    phi2 <- 0.8

    y <- arima.sim(list(order=c(2,0,0), ar = c(phi1, phi2)),n=n)

    #stimo un modello sbagliato
    mod = arima(y, order = c(1, 0, 0), method="ML")

    #i residui del modello sbagliato saranno correlati
    x=residuals(mod)

    #calcolo i test su x, residui sotto H1

#utilizzo la variante del CT per modelli ARMA sottostimati
#mostro come si costruisce il test bontà di adattamento

    lambda_3_pos= NULL
    lambda_3_neg= NULL

    S=sign(x)

    sigma_3=NULL
    p1=NULL

    for(k in 1:m){

p1[k]=(sum(x[(k+1):n]*abs(x[1:(n-k)])*S[1:(n-k)]))/(sum(x[1:(n-k)]^2))
        sigma_3[k]=sum((x[(k+1):n]- p1[k] * x[1:(n-k)])^2)/(n-k)

```

```

lambda_3_pos[k]=(sum(x[(k+1):n]*S[1:(n-k)]))/(sqrt((n-k)*sigma_3[k]))

lambda_3_neg[k]=(sum(S[(k+1):n]* x[1:(n-k)]))/(sqrt((n-k)*sigma_3[k]))
    }

    #calcolo statistica test:

    #CT_3
    #BONTA DI ADATTAMENTO

    test_CT_3[i]=sum((lambda_3_pos)^2)+sum((lambda_3_neg)^2)

    #LB
    test_LB[i]= LjungBox(x, lags=m)[2]    #statistica test

    #Monti
    pacf=pacf(x, lag.max=m, plot=F)$acf    #estraggo

    somma=NULL

    for(k in 1:m){
        somma[k]=(pacf[k]^2/(n-k))
    }

    Monti[i]=n*(n+2)*sum(somma)

#inizio ciclo MC (bootstrap)
    cont_star_CT_3=0
    cont_star_LB=0
    cont_star_M=0

    phi_star=mod$coef[1]

    B=500
    for(j in 1:B){

y_star=arima.sim(list(order=c(1,0,0),ar=phi_star),innov=sample(x),n=n)
        #ricampionando i residui, perdono la correlazione

    mod_star = arima(y_star, order = c(1, 0, 0), method="ML")

```

```

x_star=residuals(mod_star)

#calcolo i test sui residu, che sotto H0 saranno incorrelati

lambda_3_star_pos = NULL
lambda_3_star_neg = NULL

S_star=sign(x_star)

sigma_3_star=NULL
p1_star=NULL

for(k in 1:m){
p1_star[k]=(sum(x_star[(k+1):n]*abs(x_star[1:(n-k)])*S_star[1:(n-
k)]))/(sum(x_star[1:(n-k)]^2))

sigma_3_star[k]=sum((x_star[(k+1):n]-p1_star[k]*x_star[1:(n-
k)])^2)/(n-k)

lambda_3_star_pos[k]=(sum(x_star[(k+1):n]*S_star[1:(n-k)]))/(sqrt((n-
k)*sigma_3_star[k]))

lambda_3_star_neg[k]=(sum(S_star[(k+1):n]*x_star[1:(n-k)]))/(sqrt((n-
k)*sigma_3_star[k]))
}

#CT_3
test_CT_3_star[j]=sum((lambda_3_star_pos)^2)+sum((lambda_3_star_neg)^2
)

if (test_CT_3_star[j] >= test_CT_3[i]){
cont_star_CT_3=cont_star_CT_3 + 1}

#LB
test_LB_star[j]= LjungBox(x_star, lags=m)[2]

if (test_LB_star[j] >= test_LB[i]){
cont_star_LB=cont_star_LB + 1}

#Monti
pacf=pacf(x_star, lag.max=m, plot=F)$acf #estraggo

somma_star=NULL

```

```

        for(k in 1:m){
            somma_star[k]=(pacf[k]^2/(n-k))
        }

    Monti_star[j]=n*(n+2)*sum(somma_star)

    if (Monti_star[j] >= Monti[i]){
        cont_star_M=cont_star_M + 1}

} #fine ciclo MC

p_value_CT_3[i] = (cont_star_CT_3 +1)/(B+1)
p_value_LB[i]= (cont_star_LB +1)/ (B+1)
p_value_M[i]= (cont_star_M +1)/(B+1)

}

count_pv_CT_3=0 # % di Rifiuto
count_pv_LB=0
count_pv_M=0

for(i in 1:s){
if(p_value_CT_3[i] < 0.05) #rifiuto
{count_pv_CT_3=count_pv_CT_3 +1}
if(p_value_LB[i] < 0.05) #rifiuto
{count_pv_LB=count_pv_LB +1}
if(p_value_M[i] < 0.05) #rifiuto
{count_pv_M=count_pv_M +1}
}

#percentuale delle volte che rifiuto H0 quando è falsa
perc_pv_CT_3=count_pv_CT_3/s*100
print(perc_pv_CT_3)

perc_pv_LB=count_pv_LB/s*100
print(perc_pv_LB)

perc_pv_M=count_pv_M/s*100
print(perc_pv_M)

```

Bibliografia

Box G.E.P., Pierce D.A. (1970). *Distribution of residual autocorrelations in autoregressive-integrated moving average time series models*. Journal of the American Statistical Association; 65(332):1509–1526.

Fisher T.J. (2011). *Testing Adequacy of ARMA Models using a Weighted Portmanteau Test on the Residual Autocorrelations*. Contributed Paper 327-2011 to 2011 SAS Global Forum.

Fisher T.J., Gallagher C.M. (2012). *New weighted portmanteau statistics for time series goodness of fit testing*. Journal of the American Statistical Association; 107(498):777–787.

Gallagher C.M. (2001). *A method for fitting stable autoregressive models using the autocovariation function*. Statistics & Probability Letters; 53(4):381–390.

Gallagher C.M. (2002). *Order identification for Gaussian moving averages using the covariation*. Journal of Statistical Computation and Simulation; 72(4):279–284.

Gallagher C.M., Fisher T.J., Shen J. (2014). *A Cauchy estimator test for autocorrelation*. Journal of Statistical Computation and Simulation.

Gallagher C.M., Tunno F. (2008). *A small sample confidence interval for autoregressive parameters*. Journal of Statistical Planning and Inference; 138(12):3858–3868.

Lin J., McLeod A.I. (2006). *Improved Peña-Rodriguez portmanteau test*. Computational Statistics and Data Analysis; 51(3):1731–1738.

Ljung G.M., Box G.E.P. (1978). *On a measure of lack of fit in time series models*. Biometrika; 65(2):297–303.

Mahdi E., McLeod I.A. (2012). *Improved multivariate portmanteau test*. Journal of Time Series Analysis; 33(2):211–222.

Monti A.C. (1994). *A proposal for a residual autocorrelation test in linear models*. Biometrika; 81(4):776–780.

Peña D., Rodríguez J. (2002). *A powerful portmanteau test of lack of fit for time series*. Journal of the American Statistical Association; 97(458): 601–610.

Peña D., Rodríguez J. (2006). *The log of the determinant of the autocorrelation matrix for testing goodness of fit in time series*. Journal of Statistical Planning and Inference; 136(8):2706–2718.

So B.S., Shin D.W. (1999). *Cauchy estimators for autoregressive processes with applications to unit root tests and confidence intervals*. Econometric Theory; 15:165–176.