

Identificazione e verifica per modelli di serie storiche: una prima rassegna critica¹

Domenico Piccolo

Dipartimento di Scienze Statistiche, Università di Napoli Federico II
Centro di Specializzazione e Ricerche, Portici
e-mail: dopiccol@unina.it

Summary: In this paper we propose a critical review on testing statistical hypotheses in time series analysis for the identification and the model checking steps. After a preliminary discussion about the stochastic foundations of linear modelling, the related issues of Normality and stationarity are raised in frequential and time domains. Moreover, the problems of outliers and influential data are considered in a modelling framework. Some general comments and final remarks conclude the paper.

Key words: Time Series Models, Identification Tests, Model Checking.

¹Questo lavoro costituisce un parziale adattamento della relazione invitata presentata al Convegno della Commissione Scientifica della SIS su “Metodi Statistici ed Econometria” il 28 maggio 1993, presso la Facoltà di Scienze Statistiche dell’Università di Bologna. Tale contributo non è mai stato pubblicato per cui, nonostante abbisogni di un un aggiornamento bibliografico, ci sembra ancora valido sia per la prospettiva metodologica che per le considerazioni critiche che contiene. Soprattutto, ci dispiaceva disperdere un patrimonio bibliografico, notevole ed importante per ogni eventuale sviluppo successivo.

1. Introduzione

La identificazione e la verifica di modelli per serie storiche sono temi riguardanti il nucleo centrale della ricerca statistica ed econometrica su dati temporali. In tale ambito, questo lavoro intende essere una rassegna critica, circoscritta nelle tematiche affrontate e nel periodo temporale esaminato: per sua natura, quindi, esso si configura come una libera scelta dell'Autore che potrà solo sfiorare tematiche ed aspetti pure importanti, per i cui approfondimenti il ricorso alla bibliografia è necessario. Se ciò aiuta a circoscrivere i danni, resta ancora vero che gli argomenti esclusi sono certo più numerosi di quelli trattati². In effetti, l'articolo intende esaminare, in un'ottica unitaria, le questioni derivanti dalla verifica di ipotesi statistiche sia nella fase di specificazione di un modello che in quella di conferma a posteriori del medesimo, dopo essere stato stimato dalle osservazioni. La procedura si caratterizza in modo dialettico perchè la costruzione di un modello statistico avviene, progressivamente, confrontando le informazioni presenti nei dati con l'aderenza di questi a schemi precostituiti. Poiché è convinzione unanime che "tutti i modelli sono intrinsecamente sbagliati", il test si configura come una verifica di ipotesi individuate all'interno di un insieme arbi-

²Tranne brevi cenni, sono esclusi dalla nostra trattazione i modelli non lineari e quelli caotici (per questi ultimi, si vedano Vitale, 1993 ed il numero monografico del JRSS, B, 1992, vol.54, n.2, che costituisce una aggiornata introduzione statistica al tema), i modelli spazio-temporali (per i quali si possono consultare i lavori contenuti in Zani, 1993), i modelli multivariati di serie storiche (per i quali un testo pregevole è quello di Lütkepohl, 1991). Inoltre, in questo lavoro non saranno esaminati test e procedure di tipo Bayesiano, ne' le altre problematiche statistiche nate con interessi prevalentemente econometrici, come i test di causalità, l'analisi di cointegrazione e sulle radici unitarie, i modelli a ritardi distribuiti, quelli ad equazioni simultanee e quelli con cambiamenti di regime (Un ottimo testo per lo studio econometrico delle serie storiche è quello di Hamilton, 1995). Infine, la problematica della destagionalizzazione -che per sua natura include un numero notevole di test statistici su serie temporali- può essere ragionevolmente riassunta mediante i lavori di Findley et al. (1998), Gomez e Maravall (1997), Piccolo (1999) e mediante l'esame della cospicua bibliografia richiamata in questi lavori.

trario di possibili specificazioni: queste ultime sono delimitate dalla convenienza operativa, dalla tipologia delle applicazioni, dal software disponibile, dal livello di sofisticazione degli utenti, dalla frequenza delle modifiche da effettuare, e così' via³.

L'articolo si snoda secondo tre livelli interconnessi: anzitutto, si introduce la tipologia dei modelli e dei test per serie storiche; quindi, si discute della specificazione statistica del modello e della verifica delle condizioni preliminari perché essa risulti corretta; infine, si esaminano i test sulla bontà di adattamento (goodness-of-fit) del modello stimato. Le considerazioni finali si soffermano brevemente sulla validità di una inferenza condotta su fenomeni dinamici e sulla necessità che essa si svolga utilizzando sia la teoria specifica dello scienziato che conosce ed interpreta i fenomeni reali, che lo strumentario statistico più efficiente per i dati disponibili.

Infine, anche se obiettivo prioritario di questo contributo è il tentativo di giungere ad una sistematizzazione unitaria di numerose proposte presenti nella letteratura, secondo una chiave di lettura personale, esso contiene anche qualche suggerimento innovativo sul quale varrebbe la pena di indagare successivamente.

2. Modelli e test per serie storiche

In questo paragrafo, richiamiamo brevemente simbologia e notazione della modellistica che utilizzeremo, accennando poi alla tipologia dei test cui faremo riferimento.

Ogni ricerca empirica su dati reali postula un meccanismo generatore (che è e resta incognito per il ricercatore) ma che conviene circoscrivere per finalità operative o di ricerca ulteriore. In quest'ottica, il modello è un fatto strumentale, progressivamente perfezionabile ma sempre imperfetto, utile e necessario per "mimare" il meccanismo generatore delle osservazioni. In quanto tale, va controllato e sostituito.

³Per una discussione generale sui test, con particolare riferimento alle tematiche affrontate in questa rassegna, si vedano i lavori di Monari (1993), Monfort (1993), Orsi (1993).

tuito non appena la conoscenza scientifica migliora, la metodologia statistica si raffina, i mezzi di calcolo divengono più potenti, le informazioni sperimentali si arricchiscono. Per questo, da circa trent'anni, abbiamo diffuso nel nostro Paese, sia nella didattica che nella ricerca, un paradigma di analisi della modellistica per dati dinamici che si snoda sul percorso logico:

$$\boxed{\text{Serie}} \rightarrow \boxed{\text{Modello}} \rightarrow \boxed{\text{Processo}}$$

Questa impostazione, che deriva da procedure ben consolidate in ogni settore della conoscenza sperimentale, ha successivamente ricevuto cittadinanza e formalizzazione in Econometria (Spanos, 1986 ed i riferimenti citati) mediante lo schema corrispondente :

$$\boxed{\text{Data Generating Process}} \Leftrightarrow \boxed{\text{Model}} \Leftrightarrow \boxed{\text{Data}}$$

In effetti, noi consideriamo il processo stocastico X_t come la struttura probabilistica che genera la serie storica $\{x_t, t=1, 2, \dots, n\}$; ebbene, x_t è solo una parte finita delle realizzazioni $\{x_t, t=\dots, -1, 0, +1, \dots\}$ del processo X_t .

La famiglia delle ripartizioni finite di X_t costituisce, sotto ipotesi convenienti, la conoscenza necessaria e sufficiente per la sua identificazione; ma la sua complessità analitica è in generale così elevata che solo restringendo la classe di processi "utili" sarà possibile pervenire dalla specifica serie osservata a quel particolare processo che, con elevata probabilità, potrebbe averla generata.

Quindi, per inferire su X_t occorre una ipotesi distributiva sulla famiglia di variabili casuali (v.c.) X_t . Tale conoscenza non è senza conseguenze pratiche; infatti, solo in tal modo è corretto rispondere a domande riguardanti X_{t+k} (se $k > 0$, trattasi di previsione; se $k < 0$, trattasi di dati mancanti; se $k \neq 0$, trattasi di interpolazione), alle sue eventuali componenti (estrazione del segnale, destagionalizzazione), alla misura del legame con altri fenomeni (modelli econometrici, analisi multivariata, rappresentazioni multidimensionali), e così via.

D'altra parte, è possibile risalire dalla serie al processo solo se si ammette che nei dati vi siano sufficienti informazioni per costruire

stimatori almeno consistenti dei parametri di interesse⁴: ciò richiede la stazionarietà, la ergodicità e la Gaussianità di X_t . Rinviano all'abbondante bibliografia sul tema per un dettagliato esame di tali proprietà (tutt'altro che coincidenti), va evidenziato come il passaggio dal processo ad una classe di modelli che ne costituisca una parametrizzazione (efficace sul piano operativo ed efficiente sul piano statistico) avviene grazie al teorema di Wold⁵. Nelle sue linee essenziali, questo fondamentale risultato afferma che la stazionarietà del processo consente una rappresentazione lineare tramite un processo White Noise (una successione incorrelata di v.c. di media zero e varianza costante) ed un insieme di parametri che, se il processo è Gaussiano, ne determinano sia la variabilità che la correlazione seriale.

Nel seguito, per ragioni di natura statistica, delimitiamo la discussione ai processi lineari Gaussiani (ed a quelli ad essi riconducibili); infatti, tale struttura parametrica discende dal teorema di Wold tramite la stazionarietà. Inoltre, la Gaussianità consente la sua identificabilità tramite opportune funzioni-momento, connesse in modo semplice ai parametri del modello e, a loro volta, stimabili in modo efficiente dalle osservazioni.

Tale è l'ambito nel quale avviene nel 1970 una sorta di "rivoluzione culturale" grazie al lavoro di Box e Jenkins, il cui merito incontestabile è stato quello di proporre una procedura operativa mediante la quale, data una qualunque successione temporale, si costruisce per essa in modo efficiente un modello statistico lineare finalizzato alla

⁴Una situazione completamente differente è quella della verifica di ipotesi statistiche su numerose serie brevi, come trattato da Cox & Solomon (1988), ovvero su una collezione di processi, come trattato da Azzalini (1984).

⁵Il teorema di Wold rappresenta uno dei momenti centrali della analisi statistica delle serie storiche perché consente il passaggio da un'analisi ad hoc sui fenomeni dinamici alla possibilità di costruire modelli a partire dalle osservazioni in modo completamente generale. In fondo, esso rientra nella filosofia tipica della scienza statistica, perché consente la migliore riduzione delle informazioni presenti nei dati ed indica, parimenti, lo strumento efficace per operare quella riduzione. Ciò resta valido anche per le successive numerose generalizzazioni e varianti del teorema di Wold.

previsione e al controllo. Gli anni successivi hanno mostrato che tale modellistica può avere innumerevoli altre finalizzazioni essendo, proprio grazie al teorema di Wold, la struttura fondamentale e primaria per qualsiasi sviluppo sull'analisi delle serie storiche.

Nella sua essenza, come è già evidente sin dai lavori di Wold (1954), Grenander e Rosenblatt (1957), Whittle (1963), la modellistica ARMA nasce come approssimazione di tipo razionale⁶ dell'operatore lineare (e quindi dello spettro) di un processo lineare Gaussiano. Su tale risultato, secondo i canoni propri dell'analisi numerica delle funzioni limitate, si basa la possibilità di pervenire ad una formalizzazione parsimoniosa e malleabile.

Nel seguito, sia $Z_t = X_t - \mu_t$ un processo Gaussiano stazionario di media zero e dal quale è stata eliminata qualsiasi componente predeterminata. Diremo che $Z_t \sim \text{ARMA}(p, q)$ se esso è generato dalle relazioni ricorsive: $\phi(B)Z_t = \theta(B)a_t$, ove a_t è una successione incorrelata di v.c. $\mathcal{N}(0, \sigma^2)$, siglata con WN. Gli operatori in B (definiti da: $B^k Z_t = Z_{t-k}$, $\forall k=0, \pm 1, \pm 2, \dots$) sono tali che $\phi(B)\theta(B)=0$ possiede tutte le $(p+q)$ radici esterne al cerchio unitario, il che garantisce stazionarietà alla componente AR ed invertibilità alla componente MA; inoltre, la rappresentazione è minimale nel senso che $\phi(B)$ e $\theta(B)$ non hanno radici comuni. Il processo ARMA possiede funzione di autocovarianza $\gamma_k = E(Z_t Z_{t-k})$ e funzione di autocorrelazione globale $\rho_k = \gamma_k / \gamma_0$ assolutamente sommabili e son ben definite le loro trasformate di Fourier più utili: lo spettro $f(\omega)$ e lo spettro normalizzato $f(\omega)/\gamma_0$, rispettivamente.

Data una realizzazione finita $\{z_t, t=1, 2, \dots, n\}$, le problematiche che concernono il test delle ipotesi statistiche intervengono in almeno due momenti essenziali: quello iniziale, che riguarda la specificazione del modello, e quello finale, che riguarda la verifica successiva quando tale modello è stato adattato alle osservazioni. Poiché questi aspetti saranno discussi con ampiezza nei paragrafi successivi, esaminiamo ora la tipologia dei test in rapporto al loro utilizzo nell'ambito della modellistica per serie storiche⁷.

⁶Tale approccio è stato discusso con ampiezza da Piccolo (1990 a), 123-146.

⁷Per una buona introduzione alla problematica dei test segnaliamo le trat-

In primo luogo, è abbastanza raro trovare test esatti nello studio delle serie storiche, e ciò avviene per almeno due motivi:

i) non sono rispettate le condizioni usuali dell'inferenza, cioè quella di distribuzioni identicamente e indipendentemente distribuite per Z_t ;

ii) la numerosità dei dati necessari per la costruzione dei modelli su serie storiche è tale da legittimare, nella gran parte dei casi, test di tipo asintotico⁸.

Esaminando le varie tipologie di test, osserviamo che la verifica dei parametri stimati si conduce spesso tramite l'approccio detto di pura significatività (Cox e Hinkley, 1974) rapportando il parametro stimato al doppio del suo errore standard. Ciò andrebbe esaminato con cura nella modellistica di numerose serie storiche perché, se utilizzato come criterio decisionale automatico, esso prevede comunque un errato rifiuto di H_0 al 5%, cioè una volta su 20 test indipendenti. La diffusione massiccia di test (peraltro, non-indipendenti fra loro), offerta in gran numero dai moderni software econometrici, suggerisce qualche rilievo critico sulla pratica di esaminare "tutti" i test per giudicare di un modello: anche se i dati sono statisticamente coerenti con il modello, in tali casi, almeno qualche test produrrà il rifiuto dell'ipotesi nulla, senza che ciò implichi alcuna valutazione critica sul modello.

Una distinzione importante avviene tra test "nested" e "non-nested", dove i primi includono ipotesi che possono essere considerate casi particolari dell'alternativa, mentre nei secondi si tratta di confronto tra modelli sostanzialmente differenti.

tazioni di Harvey (1981), 144-188, e di Cappuccio e Orsi (1991), le ampie rassegne di Engle (1984) e Rothenberg (1984) e gli articoli di Breusch e Pagan (1980), di Hosking (1980) e di Orsi (1990).

⁸In tali situazioni, si deriva un test di tipo asintotico e, successivamente, se ne controlla la validità per n finito (in termini di livello nominale e di potenza) mediante studi di simulazione. Per una introduzione ai test asintotici si vedano: Buse (1982) e Piccolo (2000), 639-653. Più ampi riferimenti (con una notevole enfasi per gli interessi econometrici) si ritrovano nei volumi di Gourieroux e Monfort (1995), che affrontano le problematiche inferenziali dei test per modelli dinamici da molteplici punti di vista e con notevole rigore formale.

Per quanto riguarda i test “nested”, la letteratura annovera alcuni test asintotici tra i quali i più diffusi sono: LR (Likelihood Ratio test), LM (Lagrange Multiplier oppure Scoring test), W (Wald test). Nonostante alcuni risultati parziali, e alcune graduazioni sotto certe ipotesi, è complesso stabilire in generale criteri di preferenza tra essi, essendo tutti distribuiti asintoticamente come una v.c. χ^2 con pari gradi di libertà.

In effetti, la scelta a favore di uno di essi dipende dal problema e dalla convenienza computazionale. Per questo, il LM test si ritrova più spesso degli altri in quanto, per la sua determinazione, occorre la sola stima parametrica sotto H_0 (il che è generalmente più semplice); studi empirici mostrano anche soddisfacenti livelli di potenza rispetto ad alternative di tipo “locale”. Inoltre, il LM test presenta l’indubbio vantaggio di essere spesso riconducibile al valore $n \cdot R^2$ (cioè in funzione del tradizionale indice di determinazione per alcuni modelli di regressione) per cui, anche con i software più comuni, esso può essere agevolmente dedotto tramite una opportuna ri-stima di regressioni “ampliate” sui residui (Breusch e Pagan, 1980).

Meno applicati ci risultano sia il test di Wald, che richiede la stima sotto l’ipotesi alternativa (spesso più complessa e, talora, nemmeno specificata in dettaglio), che il LR test, che richiede la stima sotto le due ipotesi, ristretta e non-ristretta, assommando quindi le difficoltà dei due test precedenti senza produrre, generalmente, vantaggi in termini di potenza.

Per quanto riguarda i test “non-nested”, non ci sembrano diffuse -nel caso delle serie storiche- le proposte di Cox (1962) e di Atkinson (1970). Invece, come sarà discusso in seguito, a seguito del lavoro di Akaike (1974), hanno riscosso maggiore adesione gli approcci derivanti dal confronto tra modello dei dati e modello ipotizzato mediante misure di informazione, come la divergenza di Kullback-Leibler.

3. Identificazione e verifica nei modelli per serie storiche

La costruzione di un modello statistico avviene mediante la interazione progressiva tra distribuzione a priori, verosimiglianza, approssimazioni numeriche e nuove informazioni. Peraltro, il progresso metodologico (in Statistica, Econometria, Analisi numerica) costringe a rivedere alcune priorità sia nella specificazione che nella verifica dei modelli.

In sintesi, la identificazione⁹ di un modello per serie storiche consiste nella determinazione delle sue caratteristiche di rango nell'ambito di una certa classe; in altri termini, occorre determinare le condizioni essenziali per definire la funzione di verosimiglianza che configura la tappa successiva come stima efficiente, consistente, asintoticamente Normale, etc. Per questo, i test di identificazione costituiscono una successione piuttosto numerosa di verifiche di sotto-ipotesi particolari.

Per contro, la fase di verifica di un modello è successiva alla sua stima e costituisce, quindi, una conferma derivata dai dati della parametrizzazione ricercata (per rango e per valore dei parametri), sia di per se' che rispetto a formulazioni alternative.

Sul piano formale, nella fase di identificazione ci poniamo problemi che riguardano Z_t , $\psi(B)$ e a_t nel processo lineare: $Z_t = \psi(B)a_t$. Invece, nella fase di verifica ci poniamo problemi che riguardano $\hat{\psi}(B)$ e $\hat{a}_t$ nel modello stimato: $z_t = \hat{\psi}(B) \hat{a}_t$. Ora, se $\hat{a}_t$ è stima consistente di a_t , ogni affermazione su $\hat{\psi}$ è probabilisticamente "vicina" alla corrispondente affermazione su ψ e, quindi, sul processo generatore dei dati: il che può essere utilizzato nei test di verifica sul modello stimato.

Anche se le due fasi sono logicamente differenti, sul piano operativo sono regolarmente connesse. Per esempio, la individuazione

⁹Continueremo a parlare di "identificazione" anche se sarebbe più corretto definire tale fase "specificazione". La terminologia risale al lavoro di Box e Jenkins (1970) che la circoscrivono alla sola "identificazione" degli ordini (p,d,q) di un modello ARIMA; tuttavia, essa si è diffusa in altri contesti e viene spesso utilizzata anche in frasi come "Model Identification" di contenuto più generale.

di una componente MA(1) è un test di specificazione se applicata alla serie originaria ma costituisce il rifiuto dell'ipotesi H_0 nel test di adattamento se applicata alla serie dei residui da un modello già stimato; eppure, avviene con i medesimi strumenti e con lo stesso software, anche se con differenti gradi di libertà.

La differenza, sul piano logico e computazionale, risiede in questo: un test di specificazione è sempre una verifica della dinamica di Z_t a partire da z_t nel quale a_t , che determina Z_t tramite $\psi(B)$, non è osservabile; invece, un test di adattamento è un raffronto tra z_t osservata e $\hat{a}_t$ dedotta condizionatamente ad un modello (specificato prima e stimato poi, a partire da z_t). Come osservato da Whittle (1952), nel primo caso la specificazione avviene discriminando tra tante ipotesi, nel secondo il “fitting” è valutato prioritariamente verso una ipotesi di WN dei residui stimati.

Per accentuare questa analogia, che giustifica anche la sequenzialità dei successivi paragrafi, supponiamo di possedere una procedura $\tau(\cdot)$ per verificare l'ipotesi che una serie sia generata da un processo WN contro l'alternativa che sia generata da un modello specificato da ψ . Ebbene, nella fase di identificazione il test opera mediante $\tau(z_t)$; invece, nella fase di diagnostica il test opera mediante $\tau(\hat{a}_t) = \tau(\hat{\psi}^{-1}(B)z_t)$ valutando la capacità di $\hat{\psi}^{-1}$ di trasformare z_t in una serie $\hat{a}_t$, il più possibile simile alle realizzazioni di un processo WN. Pertanto, alcuni test sulla serie originaria si traducono in modo quasi identico sulla serie dei residui stimati da un modello sotto H_0 , anche se si modifica lo spazio in cui essi si muovono (e quindi i gradi di libertà della distribuzione di riferimento).

Nel seguito, e in modo trasversale rispetto alle due fasi elencate, la nostra analisi si snoderà nel modo seguente:

i) occorre verificare, anzitutto, se il processo è Gaussiano e stazionario, ovvero quali varianti sostanziali a tale ipotesi devono essere prese in considerazione;

ii) quindi, il teorema di Wold assicura la rappresentazione parametrica di tipo lineare sia per una eventuale componente deterministica (da individuare ed eliminare con metodi regressivi) che per una componente essenzialmente stocastica (per la quale una approssimazione

parametrica di tipo razionale è efficiente sia sul piano statistico che su quello computazionale);

iii) infine, occorre identificare gli ordini del modello e, successivamente, verificarne la validità statistica complessiva.

Le proprietà di stazionarietà, linearità, memoria finita e reversibilità temporale sono concetti strettamente connessi alla Gaussianità del processo che appare, quindi, come una delle ipotesi principali¹⁰ da verificare, o a cui ricondursi, anche per le conseguenze derivanti dalla stima. In tale ambito, si giustifica lo spazio che daremo alla verifica e alla ricerca della stazionarietà, una tematica la cui rilevanza si è recentemente accresciuta anche in ambito econometrico, con la nota problematica delle “radici unitarie”.

A tale schema premetteremo una discussione circa la possibilità di estrarre informazioni da una serie storica in termini di capacità previsiva poiché, infatti, il rifiuto di tale ipotesi è preliminare rispetto a tutte le analisi successive.

4. Il contenuto previsivo intrinseco di una serie

La disponibilità di software potente e poco costoso ha alterato la prospettiva tradizionale (Kendall, 1973) secondo la quale, prima di ogni analisi, di tipo parametrico o non-parametrico, occorre controllare se la serie fosse assolutamente casuale; in effetti, tale problematica viene oggi relegata alla verifica di alcuni test sui residui ottenuti da un modello stimato. Poiché non si ha casualità in assoluto, ma solo in rapporto ad alternative prefissate, il rifiuto dell'ipotesi nulla di casualità nelle osservazioni indica di fatto, nel contesto delle serie storiche, una preferenza verso un modello con contenuto previsivo specificato.

¹⁰La Gaussianità costituisce un punto di riferimento abbastanza usuale, anche quando non è strettamente imposta al processo. Per esempio, nella terminologia di Whittle (1953), si parla di stimatori Gaussiani (che sono approssimazioni spettrali) quando si derivano stimatori per i parametri mediante la massimizzazione della verosimiglianza determinata sotto l'ipotesi di Normalità, anche se essa non è presupposta per il processo.

D'altra parte, va distinta la possibilità di migliorare le previsioni, utilizzando modelli "superiori" e metodologie più raffinate, dalla possibilità di individuare -in rapporto ad una certa classe- la capacità previsiva "massima" compatibile con le osservazioni¹¹. Parlando di contenuto previsivo intrinseco, intendiamo sottolineare come ogni successione temporale possieda una "sua" capacità previsiva, che peraltro può essere stimata dai dati.

E' noto che l'indice di capacità previsiva¹² per un solo orizzonte temporale, condizionato alla storia passata, è espresso da:
 $R^2 = 1 - \text{Var}(a_t) / \text{Var}(Z_t)$. Se il processo è lineare, esso è funzione esclusiva dei parametri del processo essendo $R^2 = (1 + (\Sigma \psi_j^2)^{-1})^{-1} = (1 + \mathcal{A}^{-1})^{-1}$, avendo posto: $\mathcal{A} = \Sigma \psi_j^2$.

Ebbene, grazie alla formula di Kolmogorov, che esprime la varianza minimale delle innovazioni σ^2 tramite lo spettro $f(\omega)$ di Z_t , si può scrivere:

$$R^2 = 1 - (2\pi)^{-1} \int_{-\pi}^{\pi} \log f(\omega) d\omega / \text{Var}(Z_t)$$

Poiché per ogni successione stazionaria sia $\text{Var}(Z_t)$ che $f(\omega)$ sono esprimibili tramite stimatori non-parametrici, è possibile pervenire ad una stima $\hat{R}^2$ da utilizzare come riferimento in rapporto alla analoga quantità dedotta dal modello. Tale constatazione da un lato consente di verificare l'effettiva bontà del modello, dall'altro informa circa le aspettative della serie in rapporto alle previsioni, costituendo un test preliminare non vincolato ad un solo modello¹³.

¹¹Questo aspetto è discusso con ampiezza in Piccolo (1984 a).

¹²Sul piano formale, un approccio più generale trovasi in Jewell e Bloomfield (1983) e Jewell et al. (1983) mentre le tematiche dell'indice R^2 per processi lineari sono state affrontate da Nelson (1976) e Piccolo (1982 b; 1984 a) e sono state riprese, di recente, anche da Bhansali (1993). Qui stiamo sintetizzando argomenti sviluppati da Corduas (1990) e Piccolo (1990 a), 146-151.

¹³Se invece il modello è specificato, allora, utilizzando il lavoro di Nagelkerke (1991) che si basa su precedenti proposte di Maddala, Cox e Snell e Magee, si può pervenire ad una definizione completamente generale di un coefficiente di determinazione, anche per modelli di serie storiche, mediante l'espressione: $R^2 = 1 - [L(0)/L(\hat{\beta})]^{2/n}$, dove $L(0)$ e $L(\hat{\beta})$ sono rispettivamente le verosimiglianze

La problematica della stima spettrale non-parametrica di σ^2 , cui abbiamo fatto riferimento, è affrontata da Davis & Jones (1968), Bhansali (1974), Janacek (1975), Hannan e Nicholls (1977) e generalizzata al caso multivariato da Jones (1976). Inoltre, Pukkila e Nyquist (1985) hanno suggerito di utilizzare come stimatore spettrale della varianza delle innovazioni la formula di Kolmogorov, ma modificando il periodogramma usuale con una procedura di “tapering” (a causa dell’effetto di distorsione della sua stima¹⁴): questa operazione risulta tanto più efficace per quanto più si è vicini a situazioni non-stazionarie. Nel loro lavoro, pertanto, viene proposto di stimare σ^2 tramite lo stimatore non-parametrico:

$$\hat{\sigma}^2 = 2\pi * \exp\left(\frac{1}{n} \sum_{j=1}^n \log I_{n,T}(\omega_j) + \gamma - \frac{\pi^2}{12n}\right),$$

dove $\gamma=0.57721$ è la costante di Eulero, mentre $I_{n,T}(\cdot)$ è il periodogramma “tapered”.

Pertanto, la nostra proposta è così schematizzabile: prima di qualsiasi modellistica, si può stimare l’indice R^2 (utilizzando la stima spettrale $\hat{\sigma}^2$ per la varianza del processo WN e l’usuale varianza campionaria per la stima per processo Z_t), e poi confrontarlo con quello deducibile dal modello (in modo parametrico o non-parametrico). In altra ottica, si può utilizzare la stessa procedura sui residui per verificare l’adattamento di un modello stimato.

Invece, la stima della varianza delle innovazioni in un contesto parametrico (generalmente AR) e le sue connessioni con la stima dell’ordine AR sono affrontate da Paulsen e Tjøstheim (1985), che insistono sui problemi posti dalla distorsione degli stimatori AR ottenuti dalle equazioni di Yule-Walker.

A tale riguardo, poiché il calcolo di $\hat{R}^2 = 1 - \Sigma(\hat{a}_t)^2 / \Sigma(z_t - \bar{z})^2$ è una misura onnipresente negli output di pacchetti statistici ed econo-

nel modello adattato alle osservazioni e quello “nullo” sotto H_0 . Tale misura possiede tutte le proprietà del classico indice R^2 nella regressione, e può ancora essere normalizzato.

¹⁴Per una discussione sulla riduzione della distorsione sia nei parametri che nella varianza del WN, apportata dal “tapering”, si può consultare Zhang (1992).

metrici, va detto che tale stimatore ammette una distribuzione asintoticamente Normale, derivata -nel caso dei modelli ARMA- da Hosking (1979). In tal modo, si può predisporre un test di significatività utilizzando tale indicatore poich :

$$\frac{\sqrt{n}(\hat{R}^2 - R^2)}{2\sqrt{\nu(1 - R^2)}} \rightarrow \mathcal{N}(0, 1),$$

dove si   posto: $\nu = \sum_{k=0}^{\infty} \rho_k^2$. Quest'ultima quantit  pu  essere valutata sulla base di un modello oppure stimata dai dati.

5. Test di Normalit  su serie storiche

L'affermazione secondo cui z_t   generata da un processo Z_t Gaussiano  , assieme alla stazionariet , una delle ipotesi pi  importanti¹⁵ delle serie storiche perch  agevola la costruzione di un modello, nelle sue varie fasi, rendendo anche ottimali le previsioni lineari basate sulla storia passata di Z_t . Il problema   complicato dal fatto che dovendo verificare la Gaussianit  di a_t (che non   osservabile) occorre esaminare z_t . Anche per questo   necessario eliminare dalla serie, prioritariamente, ogni componente deterministica che impedisca alle realizzazioni di essere generate da un processo Gaussiano: trend, valori eccezionali, cambi di livello, non-linearit , etc.

D'altra parte, va ricordato che ogni combinazione lineare di v.c. indipendenti e identicamente distribuite tende velocemente a produrre v.c. Normali, grazie ad una generalizzazione del teorema limite centrale (Mallows, 1967). Pertanto, l'assenza di Normalit  deve

¹⁵Vi sono altri legami tra identificabilit  e Gaussianit  di un processo stocastico, tra cui la reversibilit  e la direzionalit , rispetto a t , delle serie storiche. Infatti, Weiss (1975) ha mostrato che processi ARMA(p, q) con $p > 0$ sono reversibili se e solo se sono Gaussiani ed il risultato giustifica molte delle varianti non-lineari, ampiamente discusse, fra gli altri, da Tong (1990). A tale riguardo segnaliamo l'aggiornata rassegna di Lawrance (1991) e il lavoro di Breidt e Davis (1990), mentre Lawrance e Lewis (1992) sfruttano la irreversibilit  temporale per la identificazione di strutture non-lineari e non-Gaussiane.

indurre dubbi sulla linearità del processo oppure suggerire una correlazione seriale di lag molto basso. Per esempio, è pressoché impossibile, partendo da una qualunque distribuzione di a_t con varianza finita, generare processi AR che non siano Normali per cui la non-Gaussianità può essere indice di componenti MA di basso ordine (Lusk e Wright, 1982) generate da processi WN non-Gaussiani ovvero di dinamiche non-reversibili (Weiss, 1975) e quindi non-lineari nel processo.

Circa il ruolo della Normalità nelle serie storiche, si possono individuare tre momenti:

- i) si verifica se la serie è Gaussiana;
- ii) se non è Gaussiana, si può individuare quale trasformazione avvicina le osservazioni alla Normalità in rapporto alla simmetria, unimodalità, curtosi, verificando poi sulla serie trasformata se tale Normalità è stata adeguatamente raggiunta;
- iii) si utilizza la verosimiglianza Gaussiana anche se la serie non è strettamente Gaussiana, grazie ai risultati di Whittle (1953) che mostrano che gli stimatori così derivati sono comunque asintoticamente efficienti, Normali e consistenti (Dzhaparidze, 1986).

La teoria classica utilizzerebbe a tale riguardo gli usuali test di adattamento confrontando frequenze osservate e frequenze teoriche sotto l'ipotesi di Normalità mediante il test χ^2 . Mentre è sicuramente utile confrontare istogrammi¹⁶ dedotti dalla serie con una distribuzione Normale, occorre molta cautela nell'applicazione dei test χ^2 di adattamento su dati correlati. A tale riguardo, sotto ipotesi abbastanza generali sui processi lineari, Gasser (1975) ritrova le distribuzioni asintotiche per gli indici di asimmetria e curtosi, a meno di $\mathcal{O}(n^{-2})$, già precedentemente ottenute da Lomnicki (1961).

Una rassegna molto ampia sui test di Normalità per dati indipendenti trovasi in Mardia (1980) e D'Agostino (1982): tale disamina

¹⁶A tale riguardo, ci sembra preferibile l'uso di istogrammi perequati, meno vincolati di quelli tradizionali all'ampiezza delle classi e a dati estremi: i software oggi disponibili consentono una buona qualità grafica ed una malleabilità di esecuzione che si rivelano importanti per individuare sottoserie, cambiamenti di regimi, valori anomali, e così via.

mostra che studi teorici ed esperienze di simulazioni, anche vaste (Shapiro et al., 1968), suggeriscono una certa preferenza per i test basati su osservazioni ordinate, soprattutto quando si tratta di campioni non eccessivamente numerosi. In tale circostanza, si preferisce spesso il test W_n di Shapiro e Wilk (1965) definito come il rapporto tra due stimatori della varianza della popolazione: il primo, ottenuto mediante combinazione lineare dei dati ordinati; il secondo, nel modo usuale. La significatività del test abbisogna di tavole, predisposte dagli stessi Autori. Successivamente, Royston (1982) ha ampliato quelle tavole per $3 \leq n \leq 2000$ pervenendo, quindi, ad una trasformazione di W_n che approssima molto fedelmente una v.c. Normale standardizzata.

Poiché la difficoltà del test deriva dalla necessità di disporre di tavole per i coefficienti della combinazione lineare ovvero per i livelli critici usuali, nel caso delle numerosità coerenti nell'analisi di serie storiche, si può sostituire al test W_n uno dei tanti proposti con la stessa logica da D'Agostino (1971), Shapiro e Francia (1972), Filliben (1975): la loro equivalenza asintotica con quello di Shapiro e Wilk è stata dimostrata da Verril e Johnson (1987), anche se resta difficile esplicitarne la distribuzione esatta (Leslie et al., 1986).

Fra questi, il test di D'Agostino (1971) può essere calcolato tramite la statistica:

$$D_n = 33.349 * n^{-1/2} \left(\sum_{j=1}^n \{j - 0.5(n+1)\} z_{(j)} - 0.282 \right).$$

Poiché, sotto l'ipotesi nulla, $D_n \rightarrow \mathcal{N}(0,1)$, la sua utilizzazione è immediata e non richiede ulteriori elaborazioni o tavole statistiche.

Tali risultati (derivati nel caso di campioni casuali di v.c. indipendenti e somiglianti) possono essere utilizzati anche nel caso di serie storiche grazie ad un risultato fondamentale di Pierce (1985), il quale dimostra che i test di Normalità basati sui residui da un modello AR possiedono la stessa distribuzione asintotica dei test basati su osservazioni indipendenti ed identicamente distribuite. Ciò legittima, sotto ampie condizioni, l'applicabilità dei test di Normalità ai pro-

cessi lineari, essendo questi approssimabili con modelli AR di ordine elevato.

A tale riguardo, soprattutto nella letteratura econometrica, si è diffuso un test che si attribuisce a Jarque e Bera (1987). Assumendo che osservazioni e residui da un modello di regressione derivino da v.c. appartenenti alla famiglia di Pearson, Jarque e Bera sottopongono a verifica la Normalità delle osservazioni (e poi dei residui) tramite il LM test, pervenendo alla statistica:

$$JB_N = n[(\sqrt{\hat{\beta}_1})^2/6 + (\hat{\beta}_2 - 3)^2/24],$$

essendo $\hat{\beta}_1$ e $\hat{\beta}_2$, rispettivamente, le stime dei coefficienti di asimmetria e curtosi¹⁷. La v.c. JB_N è distribuita asintoticamente come una v.c. χ^2 con $g=2$ gradi di libertà.

Bowman e Shenton (1975) proposero lo stesso test “omnibus” senza ricavarlo dalla teoria dei LM test mentre, per processi lineari, Lomnicki (1961) aveva già analizzato, separatamente, gli indici di asimmetria e curtosi. Un risultato esplicito per processi AR(p) è proposto da Lütkepohl e Schneider (1989) che applicano con soddisfazione risultati il test JB_N alle osservazioni z_t e ai residui $\hat{a}_t$ (da modelli semplici per serie storiche): i risultati sono migliori sui residui. Data la semplicità di calcolo, crediamo che un test del genere potrebbe costituire la regola in tutti i programmi di identificazione e stima¹⁸.

¹⁷Si osservi che, per un processo lineare di cui è nota la funzione di autocorrelazione ρ_k , è possibile migliorare il test JB_N utilizzando le varianze asintotiche degli stimatori di asimmetria e curtosi che Lomnicki (1961) ha mostrato essere, rispettivamente, pari a: $n * Var(\hat{\beta}_1) = 6 \sum_{k=-\infty}^{\infty} \rho_k^3$; $n * Var(\hat{\beta}_2) = 24 \sum_{k=-\infty}^{\infty} \rho_k^4$; $n * Cov(\hat{\beta}_1, \hat{\beta}_2) = 0$. Evidentemente, JB_N è derivato sotto l'ipotesi $H_0: \rho_k \equiv 0, \forall k \neq 0$.

¹⁸Successivamente, Jarque e Bera (1990) hanno generalizzato JB_N mediante un test congiunto L_N per la verifica di Normalità, indipendenza e omoschedasticità dei residui dalla regressione, agevolmente adattabile ai residui da un modello per serie storiche. Il test è ancora derivato nell'ambito dei LM test e deve la sua utilità alla scissione in tre componenti indipendenti, di cui la prima $L_{1,N} \equiv JB_N$ è un test di Normalità, la seconda $L_{2,N}$ è il test di Breusch e Pagan (1979) per la verifica di

La letteratura annovera altri test di Normalità, come quelli basati sulle analisi bi-spettrali (Hinich, 1982), oppure su aspetti specifici della Normalità multivariata. Fra questi ultimi, ricordiamo il contributo di Cox e Small (1978), di cui si usa talvolta un test di Normalità bivariata, per controllare la Normalità delle coppie di realizzazioni; quello di Lin e Mudholkar (1980) per alternative asimmetriche, e di bimodalità proposto da Oja (1983) in alternativa alla unimodalità Normale; di Srivastava e Hui (1987) per studiare la Normalità multivariata tramite il test di Shapiro-Wilk.

Sul fronte delle trasformazioni Normalizzanti¹⁹, va citata anzitutto quella di Box e Cox (1964), che spesso si risolve in quella logaritmica (e di cui ben raramente si utilizza una verifica, anche grossolana, della sua significatività). Tale problematica è molto ampia e dibattuta, per cui ci limitiamo a segnalare il contributo di Linnet (1988), che verifica l'effetto della trasformazione su alcuni test di Normalità; i problemi posti dagli outliers nella scelta della trasformazione (Carroll, 1982) e l'uso di metodi resistenti (D'Esposito, 1984) che hanno ricevuto recenti conferme da estesi esperimenti di simulazione (Atanasio e Wong, 1995). In un contesto di modello di regressione, Atkinson e Lawrance (1989) mostrano che un LM test proposto da Lawrance (1987) è asintoticamente preferibile rispetto a numerosi altri test alternativi, ma non ci risultano né test né confronti analoghi per modelli di serie storiche.

Nella sua formulazione standard, la trasformazione di Box e Cox cerca di pervenire ad una serie $z_t(\lambda)$ che sia più Normale, più omoschedastica e con effetti additivi sulla serie z_t (che supponiamo pos-

omoschedasticità su residui Normali e indipendenti (con r vincoli di specificazione della eteroschedasticità) ed infine $L_{3,n}$ è il test di Box e Pierce (1970) sulla somma dei quadrati delle prime m autocorrelazioni stimate sui residui. Se è vera H_0 , $L_n = L_{1,n} + L_{2,n} + L_{3,n} \rightarrow \chi^2$ con gradi di libertà $g = 2 + r + m$, con addendi perfettamente corrispondenti.

¹⁹Va segnalata, per osservazioni $z_t \in [0,1]$, la proposta di Wallis (1987) di considerare la trasformazione $\log(z_t/(1-z_t))$ che fa coincidere modelli lineari, funzioni di autocorrelazioni e trasformate sia per z_t che per il suo complemento $(1-z_t)$. Ciò appare ragionevole nel trattamento di rapporti di composizione: maschi e femmine, occupati e disoccupati, favorevoli e contrari, etc.

itiva o traslata in modo che lo diventi). Ciò avviene, stimando con il metodo della massima verosimiglianza, entro un certo modello, un parametro λ tale che:

$$z_t(\lambda) = \begin{cases} x_t^\lambda - 1)/\lambda, & \text{se } \lambda \neq 0; \\ \log(z_t), & \text{se } \lambda = 0. \end{cases}$$

Poiché il logaritmo della verosimiglianza è, a meno di costanti non essenziali:

$$\log[L(\lambda)] \propto (-n/2) * \hat{\sigma}^2(\lambda) + (\lambda - 1) * \sum_{t=1}^n \log(z_t),$$

ove $\hat{\sigma}(\lambda) = \sum_{t=1}^n [z_t(\lambda) - \bar{z}_t(\lambda)]^2/n$, $\bar{z}_t(\lambda) = \sum_{t=1}^n z_t(\lambda)/n$, spesso si suggerisce di trasformare z_t dividendo tutte le osservazioni per la loro media geometrica elevata a $(\lambda - 1)$. Per massimizzare la verosimiglianza delle osservazioni così trasformate (che, ora, sono comparabili tra loro), basta cercare il valore di λ che minimizza la loro varianza.

Draper & Cox (1969) approssimano la varianza dello stimatore $\hat{\lambda}$ mediante:

$$nVar(\hat{\lambda}) = (2/(3\Delta^2)) * (1 - \gamma_1^2/3 + 7\gamma_2/18)^{-1},$$

ove γ_1 è il momento terzo standardizzato e γ_2 è il momento quarto standardizzato meno 3, mentre $\Delta = \lambda\sigma/(1 + \lambda\mu)$.

In pratica, oltre che disegnare la verosimiglianza attorno a $\hat{\lambda}$, si può (Draper e Smith, 227-228) tracciare una linea parallela all'ascissa ad altezza $L(\hat{\lambda}) - 1.96$ (cioè, circa $L(\hat{\lambda}) - 2$) individuando i due valori corrispondenti sulle ascisse, λ_1 e λ_2 : ebbene, $[\lambda_1, \lambda_2]$ costituisce, con buona approssimazione, un intervallo di confidenza al 95% per λ .

6. La verifica e la ricerca della stazionarietà in media

La stazionarietà viene intesa, nei processi Gaussiani, come omogeneità rispetto al tempo del valor medio, della varianza e della fun-

zione di autocovarianza. In questo paragrafo, esamineremo esclusivamente la verifica e la ricerca della stazionarietà in media²⁰.

La stazionarietà viene regolarmente violata in quasi tutte le serie economiche, finanziarie, monetarie e aziendali, per cui assume rilievo la sua verifica e la eventuale successiva trasformazione per spiegare un trend in livello. Peraltro, in questa ricerca, si assiste ad un significativo passaggio dall'uso di trend deterministici a trend di tipo stocastico, questi ultimi sempre più raffinati e con numerose valenze interpretative (Harvey, 1989).

La letteratura econometrica, di fronte all'elevata capacità previsiva della modellistica ARIMA, ha acquisito la consapevolezza che trend lineare secondo una retta e trend stocastico mediante operatori differenza implicano comportamenti previsivi, test di specificazione e procedure di stima profondamente differenti. Nella letteratura statistica, questa consapevolezza deriva, sin dai lavori di Wold, Yaglom e Whittle, dalla necessità e dalla opportunità di utilizzare operatori stocastici per il trend esattamente come per le altre componenti di una serie, per spiegare componenti spettrali rilevanti nelle serie reali: il che ha condotto con naturalezza alla introduzione della modellistica ARIMA. Analogo percorso è stato seguito da studiosi di dati economici che hanno constatato come, a fini previsivi, interpretativi e di controllo, l'applicazione di un polinomio in t creava risultati e conseguenze ben differenti dall'uso di differenze successive²¹.

²⁰Per la verifica della omoschedasticità e la modellistica più recente che cerca di spiegarne le deviazioni (tramite modelli ARCH e varianti), rinviamo a Judge et al. (1985), e alla letteratura ivi citata. Similmente, non possiamo entrare nei dettagli della stabilità della struttura di correlazione seriale essendo il problema connesso alla costanza dei parametri e del modello rispetto al tempo. In tale ambito, ci limitiamo a segnalare l'approccio dei modelli AR con coefficienti stocastici, studiati da Nicholls e Quinn (1982). Una classe molto più generale per lo studio di tali dinamiche è quella dei modelli "State Dependent", introdotti da Priestley (1980). Un problema specifico consiste nel costruire test per la rilevazione di "rottture" (cioè modifiche della dinamica del sistema) come affrontato da Fasso' (1989, 1990), in ambito univariato e multivariato, mediante test analoghi a quello di Box e Pierce, e derivato nell'ambito della teoria dei LM test. Un test aggiornato e completo è quello di Hamilton (1995).

²¹Fra le conseguenze di una errata specificazione è ben noto che una periodicità

Una considerazione più generale, esplicitamente segnalata da Hilmer e Tiao (1979) e formalmente evidenziata da Lütkepohl (1982) anche mediante evidenze empiriche, riguarda la prassi non corretta di differenziare simultaneamente serie non stazionarie prima di sottoporle a test o modellistica multivariata. Ciò crea non-invertibilità nelle componenti MA, distrugge alcuni legami di tipo AR e non aiuta ad evidenziare la presenza di “cointegrazione” tra le serie.

D'altra parte, poiché un trend deterministico, se esiste, è implicito nell'uso di un operatore alle differenze, tramite una costante significativamente diversa da zero (Dickey et al. 1986), bisognerebbe in linea di principio optare per le differenze successive (stagionali e non) sia in funzione della varianza da esse spiegata che della evidenza grafica di rendere le serie stazionarie nel livello. Anche in tale ambito è sconsigliabile affidarsi ad una procedura automatica poiché essa, per n finito, condurrebbe a differenziare le serie anche quando sono necessarie componenti AR con una componente inerziale di un certo peso. Sul piano frequenziale, ciò equivale a confondere uno spettro con picco infinito nell'origine con uno spettro avente concentrazione di varianza attorno all'origine.

In ambito non-parametrico, Brillinger (1989) propone un test di verifica per un trend monotono sovrapposto ad una serie stazionaria fondato sul rapporto tra combinazioni lineari delle osservazioni rispetto al suo errore standard e ne mostra la Normalità asintotica e l'elevata potenza per un'ampia classe di alternative. In effetti, nel modello $z_t = S_t + E_t$ (ove S_t è il trend da stimare mentre E_t è un processo stazionario di disturbo) si cerca di verificare: $H_0: S_t$ è costante, contro l'alternativa $H_1: S_t \leq S_{t+1}$, con almeno un t tale che $S_t < S_{t+1}$. Il test è fondato sulla statistica:

$$B_t = \frac{\sum c_t z_t}{\sqrt{2\pi \hat{f}_E(0) \sum (c_t)^2}}$$

spuria di tipo stocastico è creata da una regressione su t per serie generate da processi ARIMA(0,1,0), cioè “random walk”, come mostrato nel lavoro di Nelson e Kang (1981) che evidenziano il rischio di ritrovare nei dati un periodo “indotto” pari a circa $0.83 \cdot n$.

nella quale i coefficienti c_t , per una serie di lunghezza n , sono stati determinati secondo criteri di ottimo da Abelson & Tukey (1963) e sono espressi da:

$$c_t = \sqrt{t(1 - t/n)} - \sqrt{(t+1)(1 - (t+1)/n)},$$

mentre $\hat{f}_E(0)$ esprime la stima dello spettro del processo Noise nell'origine; essa si può ottenere in molti modi, anche robusti²².

Nelle serie storiche reali è praticamente impossibile trovare operatori ∇^d con ordini diversi da $d=0,1,2$; ma è recente l'interesse per processi generati da differenze frazionarie, cioè ove d può non essere intero. La presenza di basse ma persistenti correlazioni seriali anche per lag assai distanti sono stati notati in molti fenomeni di origine fisica (metereologici, idrologici, etc.) e, poi, anche in quelli economici e finanziari non appena si possiedono un numero elevato di osservazioni (per esempio, settimanali o giornalieri su serie finanziarie). Occorrevano, quindi, modelli nei quali fosse contemporaneamente presente una autocorrelazione che tendeva velocemente a zero (e qui la classe dei processi ARMA con le opportune varianti è più che adeguata) ed una autocorrelazione che si annullava con un tasso di decadimento molto più basso di quello esponenziale (e qui la classe dei modelli ARIMA tradizionali non era sufficiente). Queste esigenze, anche in connessione alla teoria dell'informazione, sono illustrate con grande efficacia da Parzen (1992).

Tale aspetto può farsi risalire alla presenza di una componente ω^{-d} nello spettro, quando $\omega \rightarrow 0$ (Cox, 1977), il che ne spiega la difficoltà di individuazione nelle serie economiche, per le quali una numerosità

²²Il suggerimento è di Brillinger (1989), 25-26. Il test ha rilevanza per serie molto lunghe ed è stato applicato in ecologia, astronomia, demografia, etc. (Brillinger si occupa della deforestazione del bacino Amazzonico partendo da oltre 30mila osservazioni giornaliere), ma affinché diventi realisticamente applicabile ci sembra opportuno adattarlo con stimatori della varianza del processo WN meno esigenti, fondati su ipotesi parametriche semplici. In tal caso, potrebbe essere utilizzato sistematicamente sulle serie differenziate nelle procedure usuali verificandone significatività e potenza, anche per campioni di numerosità più moderata.

modesta non consente un buon grado di discriminazione spettrale. Mentre per un processo stazionario ARMA la funzione di autocorrelazione ρ_k tende a zero come θ^k , $|\theta| < 1$, un processo frazionario ρ_k tende a zero come k^{2d-1} , $d < 0.5$. Granger e Joyeux (1980) introducono dal punto di vista statistico e nell'ottica dei modelli lineari tale problematica, mentre una sintesi efficace è contenuta in Brockwell e Davis (1987), 464-478. Inoltre, Gray et al. (1989) propongono una generalizzazione dei processi ARMA frazionari sostituendo all'operatore $(1-B)^d$ l'operatore di Gegenbauer $(1-2\delta B+B^2)^\lambda$ che estende il comportamento spettrale dei processi frazionari a qualsiasi frequenza angolare $0 \leq \omega \leq \pi$.

Un processo con differenza frazionaria è formalizzato da Hosking (1981) che, anche sulla base di precedenti lavori: Cox (1977), Granger (1978, 1980), Lawrance e Kottegoda (1977), Granger e Joyeux (1980), introduce un modello ARIMA(p,d,q) per il quale $-0.5 < d < +0.5$ implica sia stazionarietà che invertibilità per Z_t ; in particolare, mostra come $d < +0.5$ implica la stazionarietà di Z_t , mentre $d > -0.5$ implica la invertibilità di Z_t .

Pertanto, il processo $\nabla^d Z_t = a_t$ copre un ampio spettro di possibilità:

- i) per $d \in (0, .5)$ è stazionario con effetti persistenti (cioè “long memory”: funzioni di autocorrelazione globale e parziale tutte positive e monotonicamente decrescenti a zero in modo iperbolico con spettro concentrato alle bassissime frequenze);
- ii) quando $d=0$ è un processo WN;
- iii) quando $d \in (-.5, 0)$ il processo ha funzioni di autocorrelazione globale e parziale tutte negative per lags $k \neq 0$ con spettro concentrato alle altissime frequenze.

Allora, un modello ARIMA(p,d,q) con $d \in (-.5, +.5)$ può essere scelto in modo da utilizzare i parametri ARMA per spiegare la struttura di correlazione ai lag più bassi e la differenza d per spiegare la struttura di correlazione persistente ai lag più elevati. A partire da Janacek (1982), che propone di stimare d mediante una regressione sul logaritmo del periodogramma, Geweke e Porter-Hudak (1983)

confrontano vari stimatori alternativi dello stesso tipo e Kashyap e Eom (1988) ne dimostrano non-distorsione e consistenza. Tenuto conto che le funzioni di autocorrelazioni parziali di un processo frazionario sono pari a: $\pi_k = d/(k-d)$, $k > 1$, abbiamo proposto stimatori di d anche sul piano temporale, per i quali alcuni semplici esperimenti preliminari ci hanno confermato l'utilità.

In tale ambito, un test è derivato da Li e McLeod (1986) che, nello specifico, confermano che gli sviluppi asintotici sono validi anche per il modello ARIMA con differenza frazionaria aggiungendo che, per media nota (o supposta nota), vale ancora la Normalità asintotica degli stimatori ML dei parametri ARMA e dell'ordine d della differenza frazionaria. In particolare, si ottiene: $\sqrt{n}(\hat{d} - d) \rightarrow \mathcal{N}(0, 6/\pi^2)$ evidenziando una distribuzione asintotica svincolata da quella dei parametri ARMA.

Similmente, viene confermato il risultato asintotico per il test di Ljung e Box (oltre che quello di Box e Pierce, ovviamente) che, se calcolato su m lags, va confrontato con un livello critico per la v.c. χ^2 con $g=m-p-q-1$ gradi di libertà. Inoltre, si dimostra che la varianza dell'autocorrelazione dei residui stimati dal modello frazionario $\nabla^d Z_t = a_t$ è approssimata²³ da $Var(\hat{r}_k) = n^{-1}(1 - 6/(\pi k)^2)$.

Purtroppo, le conclusioni di Davies e Harte (1984) mostrano che, se le osservazioni non superano il migliaio, anche un test con proprietà ottimali, difficilmente discrimina tra una componente AR(1) con basso valore del parametro (o ARMA(1,1) con basso valore di $|\phi - \theta|$) ed un processo frazionario.

Beran (1992) ha proposto un test di "long-range dependence" su processi Gaussiani di tipo generale, mediante il rapporto di funzioni del periodogramma, precedentemente introdotto da Milhoj (1981)

²³Si noti che, in pratica, ciò crea un intervallo di confidenza attorno alle autocorrelazioni stimate più ridotto solo per il primo lag. Difatti, al 95%, invece di $\pm 2 \cdot n^{-0.5}$, per il lag 1, si ha: $\pm 2 \cdot 0.626 \cdot n^{-0.5}$; per il lag 2 si ha: $\pm 2 \cdot 0.921 \cdot n^{-0.5}$, e così' via.

come test di adattamento. Esso può essere espresso tramite:

$$T_n = \frac{4 \pi n^{-1} \sum_{j=1}^{n_0} (\nu_j)^2}{\left(4 \pi n^{-1} \sum_{j=1}^{n_0} \nu_j \right)^2} = (2\pi)^{-1} \left[1 + \sum_{k=1}^{n-1} \hat{r}_k^2 \right],$$

dove $\hat{r}_k$ rappresentano le autocorrelazioni stimate dai residui del modello adattato a z_t e dove, per semplicità notazionale, abbiamo posto: $\nu_j = I(\omega_j)/f(\omega_j)$, $\omega_j = 2\pi j/n$, $j=1,2,\dots,n_0$, mentre: $n_0 = (n-1)/2$ per n dispari, $n_0 = n/2 - 1$ per n pari. Il periodogramma è, al solito, definito tramite: $I_n(\omega_j) = (2\pi n)^{-1} \left| \sum_{t=1}^n z_t \exp(-i\omega_j t) \right|^2$.

La ragionevolezza del test, come sostiene Beran (1992), 752, deriva dal fatto che T_n tiene conto di tutte le autocorrelazioni dei residui stimabili (e non delle prime m , come nei test di adattamento più comuni); infatti, nei processi con differenze frazionarie non tendono esponenzialmente a zero le autocorrelazioni teoriche (come invece avviene per i processi ARMA), per cui tutte le stime apportano utili elementi di informazione circa il rifiuto dell'ipotesi nulla.

Il test è semplice da implementare perché, sotto H_0 , è distribuito asintoticamente come una v.c. $\mathcal{N}(\pi^{-1}, 2\pi^{-2}n^{-1})$. Poiché anche per valori di n di poco superiore al centinaio è confermata l'approssimazione: $P(T_n \leq t_0) = \Phi(\sqrt{n}(t_0\pi - 1)/\sqrt{2})$, se z_α è il percentile $100(1-\alpha)\%$ della v.c. $\mathcal{N}(0,1)$, allora si rifiuta H_0 quando $T_n > \pi^{-1}(1 + z_\alpha\sqrt{2/n})$. Le esperienze riportate confermano l'elevata potenza di questo test contro alternative di autocorrelazioni decadenti a zero con estrema lentezza e la opportunità (per la sua semplicità computazionale) di includerlo nella prassi delle procedure statistiche standard.

7. Test sulle componenti spettrali

L'analisi spettrale si è sviluppata grazie all'apporto che alcuni grandi probabilisti hanno avuto sullo studio di fenomeni stocastici. Essa deve il suo "successo" metodologico alla possibilità di scomporre un processo stazionario in componenti indipendenti, ciascuna delle

quali possiede un contenuto proprio di varianza spiegata per cui è agevole costruire oggettivamente una graduatoria di importanza e significatività statistica. Per esempio, un test di incorrelazione della serie -per tutti i lag temporali- si traduce, sul piano spettrale, in un test $H_0: \beta_1=0$ nel modello di regressione: $I_n(\omega_j)=\beta_0 + \beta_1*\omega_j + \epsilon_j$, essendo $I_n(\omega_j)$ l'ordinata del periodogramma alla frequenza angolare $\omega_j=\pi j/m$, $j=0,1,\dots,m$.

Sul piano analitico, inoltre, l'analisi spettrale consente di distinguere se le componenti che apportano spiegazione alla dinamica del processo sono in numero discreto (e, allora, una o più componenti deterministiche sono necessarie e sufficienti per spiegare una dinamica fondamentalmente regressiva, in senso classico) oppure corrispondono all'intervallo delle frequenze angolari comprese tra $-\pi$ e $+\pi$ (e, allora, solo una struttura stocastica -come per esempio quella dei processi lineari- può darne ragione). Tuttavia, si identifica a volte l'onda periodica presente nella struttura di correlazione (che è quindi di tipo deterministico perché deriva da una equazione alle differenze finite su ρ_k) con la moda della densità spettrale (che è di tipo stocastico perché deriva dall'intervallo infinitesimo sulle frequenze angolari che possiede massima percentuale di varianza spiegata). Come rilevato da Piccolo (1978, 1979), tali concetti sono coincidenti solo ai limiti della regione ammissibile mentre, nella maggior parte dei casi, ciò induce una distorsione nella valutazione dell'onda predominante.

E' importante, quindi, discriminare tra componente deterministica e stocastica perché, per esempio, nella fase di previsione le differenze tra le due scelte sono rilevanti. Infatti, una componente deterministica riproduce rigorosamente un comportamento periodico mentre una componente stocastica, pur se maggiormente sensibile agli ultimi dati, tende comunque verso il livello medio del fenomeno. Nel seguito, quindi, distingueremo fra test spettrali su componenti deterministiche (armoniche) e test che utilizzano lo spettro di processi lineari.

Una prima proposta di verifica delle componenti deterministiche risale a Fisher (1929) che propose un test sulla presenza di una singola componente armonica tramite il confronto tra l'ammontare

della varianza spiegata dall'ordinata massima del periodogramma e l'ammontare complessivo, cioè la varianza della serie. In questa linea, Siegel (1980) esamina tutte le ordinate del periodogramma che superano un certo valore-soglia e sfrutta la dualità di tale problema con quello di coprire un cerchio tramite archi casuali di lunghezza g prefissata, mostrando poi una potenza più elevata del suo test. Reschenhofer (1989), invece, suggerisce una procedura "adattiva" (così definita perché determina le frequenze da sottoporre a test sulla parte reale della trasformata di Fourier delle osservazioni e ne verifica poi la significatività sulla parte immaginaria) e dimostra anche che il test di Siegel può essere utilizzato per alternative di tipo continuo, pur essendo stato introdotto per componenti deterministiche. Successivamente, Reschenhofer e Bomze (1992) hanno combinato il precedente risultato di Reschenhofer utilizzando, in un'unica procedura, i test sulle parti reale e immaginaria. Sulla stessa scia, Damsleth e Spiotvoll (1982) propongono una procedura stepwise per la determinazione di componenti multiple.

Chiu (1989) rileva i difetti del test di Fisher perché media il periodogramma in presenza di ordinate notevolmente più elevate, le quali possono essere considerate quando non è vera H_0 , come degli outliers in "dati" generati da una distribuzione esponenziale. Per tale motivo, sarebbe preferibile utilizzare procedure più resistenti della media aritmetica: Chiu propone una media "trimmed" confrontando le ordinate del periodogramma con la somma delle prime $100 \cdot \beta\%$ di esse (ordinate in senso non-decrescente). Le scelte $\beta=0.9$ oppure $\beta=0.95$ sembrano fornire buoni risultati contro l'alternativa di componenti periodiche multiple.

Sullo stesso problema, Quinn e Thomson (1991) derivano un LR test per sottoporre a verifica p armoniche, a loro volta stimate dalle osservazioni, tramite un test asintoticamente χ^2 , mentre una procedura consistente per individuare armoniche è quella di Zhao-Guo (1988). Similmente, Lii e Tsou (1992) generalizzano alcuni test classici (Bartlett, Whittle, Priestley) sulla base di cumulanti quarti e funzioni bi-spettrali. Va anche citato il test $P(\lambda)$ di Priestley (1981), 626-642, proposto sin dal 1962, e valido per processi con spettri misti,

ma la cui complessità computazionale ne ha scoraggiato l'uso.

Nello stesso ambito, Hannan e Quinn (1989) affrontano il problema della separabilità tra componenti "vicine" che diventa importante negli studi econometrici per distinguere tra periodi stocastici molto lunghi e trend evolutivo. Circa il numero di componenti armoniche da esaminare, Quinn (1989) propone un criterio tipo AIC nel quale le successive "varianze residue" sono determinate sottraendo alla varianza della serie le varianze ordinatamente spiegate dalle componenti deterministiche.

Per quanto riguarda la verifica su componenti stocastiche, Bartlett (1955, 1966) propose di applicare il test di Kolmogorov-Smirnov al periodogramma cumulato che, sotto $H_0: Z_t \sim WN$ e per $0 \leq \omega \leq \pi$, è la realizzazione di una funzione di ripartizione definita come la retta passante tra l'origine $(0,0)$ e il punto $(\pi,1)$. Tale test -spesso richiamato in molti manuali ed utilizzato in numerose applicazioni- ha scarsa potenza contro alternative contenenti spettri plurimodali. In più, nella forma usualmente applicata, Dahlhaus (1985) ha mostrato che esso non tende nemmeno alla distribuzione di Kolmogorov-Smirnov e, perché ciò avvenga, ha suggerito di modificarlo tramite la stima delle innovazioni da un processo ARMA. In alternativa, Reschenhofer e Bomze (1991) hanno proposto la lunghezza del periodogramma cumulato, interpolato linearmente, come misura di deviazione rispetto all'ipotesi di WN.

Una impostazione più globale è presente in Newton e Pagano (1984) -e nei lavori collegati- i quali derivano intervalli simultanei di confidenza per l'intero spettro stimato in modo parametrico tramite processi AR; tali risultati sono stati, poi, migliorati da Tomasek (1987).

Un filone importante connesso all'analisi spettrale è quello della similarità e discriminazione tra spettri di serie storiche differenti, per il quale ci limitiamo, qui, a ricordare i test che si riconducono, rispettivamente, a confronto tra gli spettri, le funzioni di autocorrelazione ed i modelli per serie storiche.

Per la discriminazione tra spettri, segnaliamo i contributi di Dargahi-Noubary e Laycock (1981) che utilizzano un test di confronto

tra le stime spettrali per distinguere tra eventi sismici ed esplosioni nucleari nel sottosuolo. Inoltre, Coates e Diggle (1986) confrontano test parametrici e non-parametrici, anche in alternativa a processi AR, proponendo un test basato sul campo di variazione del logaritmo dei rapporti tra i rispettivi periodogrammi cumulati delle due serie; la loro proposta è stata generalizzata da Pötscher e Reschenhofer (1988) nel caso di osservazioni ripetute.

Tale famiglia di test è generata dai contributi di Shumway (1982) e Shumway e Unger (1974) che introducono come miglior criterio discriminante il rapporto tra i periodogrammi (meglio se perequati con finestre spettrali) i quali, sotto H_0 , si distribuiscono come una v.c. di Fisher. In tale contesto, viene anche mostrato come una funzione discriminante quadratica tra le differenze delle inverse delle matrici di covarianza equivale ad una forma quadratica tra gli spettri inversi, come spesso è stato proposto ed utilizzato nella letteratura sia a fini descrittivi che di classificazione.

Un approccio differente è quello di Mélard e Roy (1984, 1987) che confrontano serie indipendenti mediante test di uguaglianza delle funzioni di autocorrelazioni: la statistica da loro usata è una forma quadratica nella quale l'usuale stimatore delle funzioni di autocorrelazione è ponderato per assicurarne la consistenza.

Infine, l'intera problematica si riconnette a quella affrontata, con maggiore enfasi sul piano temporale, nella monografia di Corduas (1992 a). A tale riguardo, Piccolo (1983, 1990 b) ha proposto una metrica per i modelli ARIMA, per la quale -grazie a numerose verifiche empiriche ed alcuni risultati inferenziali (Corduas, 1992 b)- è possibile una utilizzazione operativa sia nella fase di stima dei modelli (Corduas, 1997) che nelle indagini congiunturali (Piccolo, 1999).

8. Identificazione degli ordini di un modello lineare

La specificazione di un modello lineare avviene mediante l'approssimazione della forma razionale $ARMA(p,q)$, per cui è necessario sot-

toporre a test ipotesi convenienti sui parametri connessi a p e q . A tale riguardo, due sono gli approcci generalmente usati: quello dei criteri di selezione, basati su quantità da minimizzare in funzione diretta della varianza residua e in funzione inversa del numero dei parametri coinvolti; e criteri direttamente orientati ad una verifica statistica delle ipotesi sui modelli.

Punto di partenza della discussione sono i risultati sulla distribuzione degli stimatori della funzione di autocorrelazione globale²⁴, la cui varianza fu derivata da Bartlett (1946), e quelli sulla autocorrelazione parziale, la cui varianza fu derivata da Quenouille (1949), mentre analoghi approcci operativi sono quelli sulle funzioni di autocorrelazione inverse e inverse parziali (discusse ed applicate da Hipel et al., 1977; McLeod et al., 1977). Per contro, Bhansali (1983) cerca di valutare l'ordine della componente MA mediante la stima della funzione di autocorrelazione inversa, come già in McClave (1978). Ancora, Godolphin (1980) propone test formali per l'ordine di un processo ARMA, basati su un gruppo di autocorrelazioni stimate sui residui, inferiore a quello usuale $m = \mathcal{O}(n^{1/2})$, che si dimostra più sensibile per discriminare tra modelli alternativi.

Al riguardo, un lavoro di Abraham e Ledolter (1984), le cui risultanze andrebbero più ampiamente sperimentate e discusse, mostra che -in termini di potenza- la funzione di autocorrelazione parziale è preferibile rispetto a quella inversa per processi AR, ma che la funzione di autocorrelazione inversa parziale è preferibile rispetto a quella globale per processi MA. Tali fatti suggerirebbero strategie composite differenti da quelle usualmente presenti nei pacchetti statistici, anche se tutte le funzioni di cui si discute si determinano sempre dalla stima non-parametrica della funzione di autocovarianza o della sua trasformata di Fourier (spettro).

Siamo convinti che le funzioni di autocorrelazione vadano sottoposte ad analisi "grafica" perché ciò fornisce maggiori informazioni di un test di significatività in quanto, per ciascun lag, cambia la distribuzione a priori dello statistico. Alla fine, si afferma spesso

²⁴Riferimenti obbligati sono anche gli sviluppi successivi di McLeod (1978) e Hosking (1978).

che: “la identificazione di un modello ARMA è un’arte più che un metodo”. Ma ogni arte si esercita mediante strumenti artigianali e le risultanze di alcuni decenni di esperimenti su serie reali e su dati simulati consentono ora di fornire crediti differenti alle varie proposte.

Il recente esperimento di Koreisha e Yoshimoto (1991) indica chiaramente una preferenza per criteri di tipo automatico generati da stime lineari in due tempi: la prima mediante un modello long-AR, e la seconda mediante un criterio di minimi quadrati generalizzati sulla formulazione ARMA, in cui però le componenti non osservabili a_t sono sostituite mediante quelle stimate $\hat{a}_t$ tramite il modello long-AR. Tale approccio si ritrova in numerosi lavori di Pukkila, ed è ampiamente discusso e sperimentato in Pukkila et al. (1990). Si noti comunque che, a causa dell’elevata distorsione degli stimatori lineari derivati dalle equazioni di Yule-Walker, viene sempre preferito l’algoritmo di Burg (1978) che supera tali inconvenienti. In particolare, la scelta degli ordini AR avviene tramite il criterio di Schwarz (1978), e non quello di Akaike (1974) che sovrapparametrizza o di Hannan e Quinn (1979) che tende a sottovalutare il numero di parametri.

Rinviando alla elencazione delle più diffuse proposte in de Gooijer et al. (1985), ci limitiamo ora ad esaminare alcuni risultati generali che, seguendo i criteri del lavoro di Koreisha e Yoshimoto (1991), suddividono i metodi di identificazione in tre gruppi:

i) quelli connessi alla singolarità di matrici deducibili dalle equazioni ricorrenti che definiscono il processo ARMA e la sua struttura di correlazione sino al secondo ordine;

ii) quelli derivanti da approcci regressivi di tipo ricorsivo;

iii) quelli inerenti la formulazione di un criterio statistico di tipo automatico fondato sulle stime di una ampia classe di modelli e di un indice la cui minimizzazione è criterio di giudizio per la scelta.

Fra i criteri basati sulla singolarità di determinanti di vario tipo, va segnalato il cosiddetto “corner method” proposto da Béguin et al. (1980) e successivamente ampliato e generalizzato in molte formulazioni. Ebbene, come Piccolo e Tunnicliffe Wilson (1984) hanno dimostrato, essi derivano tutti da alcune equazioni alle differenze, presenti nel test di Box e Jenkins (1970): le equazioni di Yule-Walker,

quelle della funzione di previsione, le relazioni tra parametri $AR(\infty)$ e $MA(\infty)$ nel processo ARMA. Alcune relazioni tra tali espressioni, soprattutto per il loro aspetto ricorsivo e le implicazioni computazionali, sono riportate in Tuan (1984). Tale singolarità è unitariamente accertabile mediante quella della matrice delle varianze e covarianze delle previsioni condizionate al passato le quali si collocano in uno spazio di dimensione $\max(p,q)$, come dimostrato da Akaike (1976) e come utilizzato con particolare efficacia²⁵ da Mazzali (1982). Contemporaneamente, Jewell e Bloomfield (1983) e Jewell et al. (1983) mostrano che tale derivazione può collegarsi alla correlazione canonica tra passato e presente, mentre Battaglia (1983) e Corduas (1990) evidenziano in tali legami degli aspetti strutturali di fondamentale importanza per lo studio e l'inferenza sui processi lineari.

Un approccio di tipo regressivo è stato proposto da Tsay e Tiao (1984) che introducono la stima della funzione di autocorrelazione estesa calcolata sui residui da modelli AR di ordini crescenti; essa, tuttavia, non ci è parsa molto selettiva nella prassi e nelle simulazioni riportate.

Con riguardo, poi, ai criteri cosiddetti automatici, l'esperienza ha indicato una netta preferenza per il criterio $SIC = n * \log(\hat{\sigma}^2) + 2\log(p)$ di Schwarz (1978), anche per processi multivariati. In effetti, le numerose proposte esistenti sono tutte connesse fra loro (Piccolo, 1980) in quanto risultano funzione della varianza dei residui e del numero dei parametri, cui applicano una penalizzazione che costituisce la vera discriminante tra di loro; anche il criterio CAT di Parzen (1974), concettualmente differente, si è dimostrato, poi, asintoticamente equivalente ad AIC.

Hurvich e Tsai (1989, 1991) hanno migliorata la bassa performance di AIC, che è distorto per le stime su modelli $AR(p)$. Anziché considerare:

$$AIC = n * \log(\hat{\sigma}^2) + 2(p + 1),$$

²⁵Sulle relazioni tra passato, presente e futuro, nonché su quelle di causalità, sono anche importanti i contributi di Faliva (1991) e di Faliva e Zoia (1993).

come usuale, essi invece propongono:

$$AIC_U = n * \log(\hat{\sigma}^2) + n(1 + p/n)[1 - (p + 2)/n]^{-1}.$$

Recentemente, Hurvich et al. (1990) e Hurvich (1992) hanno ancora modificato tale espressione ottenendo ulteriori miglioramenti per la scelta di modelli in piccoli campioni.

Kavalieris (1989) ha mostrato che, nel caso dei processi AR, i criteri che minimizzano la somma dei quadrati dei residui ricorsivi, la lunghezza della descrizione predittiva e il criterio di Schwarz coincidono. Inoltre, Lii (1985) collega la determinazione dell'ordine del modello all'approssimazione di Padé, la cui stima iniziale dei parametri è consistente anche se non pienamente efficiente. Infine, sull'ampio fronte della scelta di sottoinsiemi di tipo AR, ci limitiamo a segnalare i lavori di Haggan e Oyetunji (1984) e di Yu e Lin (1991).

Nell'ambito della identificazione, intendiamo anche discutere un approccio raramente utilizzato su dati economici e che, a partire dal lavoro di Kedem (1981), presenta qualche elemento di interesse. L'Autore generando da z_t delle serie storiche "binarie" (cioè serie ridefinite come una successione di 0 e 1 a seconda che z_t sia inferiore o non inferiore ad una certa soglia δ), cerca di inferire sul meccanismo generatore delle osservazioni a partire dagli "attraversamenti dell'ascissa delle differenze di ordine successivo di una serie" detti HOC (=Higher-Order Crossings). La razionalità dell'analisi si basa su alcuni teoremi, tra cui quello più rilevante (Kedem e Slud, 1981) afferma che, indicato con $D_{n,k}$ il numero di attraversamenti dell'asse t ("axis crossing") della $(k-1)$ differenza $\nabla^{k-1} Z_t$, $k=1,2,3,\dots$ allora, sotto l'ipotesi Gaussiana per Z_t , si derivano relazioni notevoli tra il valor medio di tali HOC e le funzioni di autocorrelazione. In generale:

$$\cos\left(\frac{\pi E(D_{n,k+1})}{n-1}\right) = \frac{\nabla^{2k} \rho_{k-1}}{\nabla^{2k} \rho_k}, \quad k = 1, 2, \dots$$

nella quale l'operatore ∇ opera qui su k . In particolare, si ha:

$$\cos(\pi E(D_{n,1})/(n-1)) = \rho_1.$$

Ciò mostra, per processi Gaussiani, una corrispondenza tra funzioni di autocorrelazione (e tutte quelle derivabili da ρ_k) e valor medio di HOC, e di conseguenza suggerisce un utilizzo inferenziale di tali quantità (empiricamente determinabili dalle osservazioni) per lo studio del modello generatore di Z_t . Infatti, grazie alla conoscenza della varianza asintotica delle stime di $D_{n,k}$, Kedem (1987 b) punta alla identificazione dei modelli; Kedem e Slud (1981) propongono un test sulla bontà di adattamento tra frequenze attese e teoriche; Kedem e Slud (1982) utilizzano HOC quale criterio di discriminazione tra serie.

Tra questi risultati, segnaliamo quello proposto da Kedem e Reed (1986) che pervengono ad un “fast WN test” basato sugli HOC. Infatti, è possibile derivare un intervallo di confidenza al 95% per l’ipotesi che Z_t sia un processo WN poiché, asintoticamente,

$$\frac{D_{n,k} - (n-1)(0.5 + \nu_k^2)}{\sqrt{(n-1)(0.25 - \nu_k^2)}} \rightarrow \mathcal{N}(0, 1),$$

ove: $\nu_k = 0.5 - \pi^{-1} \arcsen(1 - k^{-1})$, $k=1, 2, \dots$. Gli Autori suggeriscono di limitarsi ai valori di $k=1, 2, \dots, 8$ e ne confermano l’utilità sia su serie reali che mediante esperimenti di simulazione.

Infine, Kedem (1987 a) utilizza lo stesso approccio per la individuazione di periodicità e mostra che la successione $\pi D_{n,k}/(n-1)$ può essere utilizzata per la determinazione delle frequenze dominanti dello spettro. Infine, He e Kedem (1990) ne studiano l’uso per un particolare processo AR(2) anche in relazione ad eventuali radici unitarie. In quest’ultimo lavoro, non si considera più la quantità HOC ma ZCR (zero-crossing rate) che è invariante con n . Per esempio, per un processo AR(2) del tipo: $Z_t = 2\cos(\theta)Z_{t-1} - Z_{t-2} + a_t$, si ottiene:

$$\lim_{n \rightarrow \infty} D_{n,k}/n = \theta/\pi.$$

La nostra opinione è che tale approccio, essendo di immediata applicazione, può essere inserito come routine di calcolo nei pacchetti statistici esistenti con finalità descrittive ma, sul piano inferenziale, ci sembra permangono tuttora due debolezze:

i) tranne eccezioni, gli esperimenti riportati riguardano serie con un migliaio di osservazioni;

ii) non ci sembra immediato individuare comportamenti univocamente ben definiti a partire dal grafico degli HOC.

Concludiamo questo paragrafo dedicato alla identificazione con due considerazioni sintetiche: anzitutto, un modello ARMA ha senso se è parsimonioso, altrimenti la collocazione dei parametri diventa dubbia e assai instabile nello spazio ammissibile (Piccolo, 1982 a). In secondo luogo, la scelta di un modello AR dipende dalla funzione di autocorrelazione parziale e, dualmente, la scelta di un modello MA dalla funzione di autocorrelazione globale; pertanto, in fin dei conti, tutti i criteri dipendono solo dai valori di tali funzioni in alcuni lags più che dalla intera struttura di correlazione seriale (Piccolo, 1980). Per questo, è necessario studiare la sensibilità di tali funzioni rispetto a valori anomali (Franses, 1991) considerando anche versioni robuste, come in Masarotto (1987, 1988).

9. Test diagnostici sul modello stimato

La fase di validazione statistica del modello consta nell'accertamento della rispondenza tra aspettative e osservazioni dedotte sotto il modello, per cui il rifiuto di H_0 è contemporaneamente il rifiuto del modello identificato in precedenza, almeno in sotto-ipotesi specifiche. L'aspetto interattivo della procedura, tuttavia, consente anche di orientarsi verso alternative ben definite²⁶ anche se, formalmente, il

²⁶Esistono interazioni fortissime tra acquisizioni metodologiche e risultanze operative che non sono codificabili in test di tipo classico ma che fanno parte dello strumentario di chi analizza dati temporali. Esemplifichiamo solo un caso standard: se l'analisi dei residui evidenzia una forte autocorrelazione negativa al lag 1, attorno o superiore a -0.5 , è praticamente certo che si è operata una sovra-differenziazione; ma se essa è accompagnata da una autocorrelazione al lag 2 di segno positivo (e notevolmente inferiore in valore assoluto), è certo più probabile orientarsi verso la identificazione di una componente AR(1) con segno negativo. Ebbene, la osservabilità (Piccolo, 1981 a) di tale situazione richiede tra 100 e 400 osservazioni, a meno che non si disponga di una distribuzione a priori molto informativa sulla funzione di autocorrelazione, ovvero equivalentemente non si

test è costruito attorno al rifiuto di H_0 .

Nella fase di diagnostica del modello stimato, molte scelte operative sono riconducibili alla teoria delle decisioni statistiche: significatività dei parametri, test sui residui, ipotesi distributiva per gli errori, presenza di valori aberranti, conformità del modello a classi predefinite, etc. Nel seguito, tuttavia, insisteremo sui test basati sulla correlazione seriale dei residui che sottopongono a verifica un adattamento complessivo tramite funzioni di autocorrelazione di $\hat{a}_t$.

Poiché lo studio della distribuzione della funzione di autocorrelazione appare assai complesso²⁷, come ampiamente discusso in Anderson (1971), ci si orienta verso misure sintetiche di tutte o di alcune autocorrelazioni stimate tramite un test "portmanteau" che le sintetizza.

Fondamentale per tali lavori è stato il contributo di Box e Pierce (1970) che esaminano la distribuzione dell'autocorrelazione stimata sui residui da un modello ARMA. Se indichiamo, rispettivamente, con a_t e $\hat{a}_t$ le innovazioni vere e quelle stimate da un modello ARMA, le corrispondenti funzioni di autocorrelazione saranno:

$$r_k = \frac{\sum_{t=k+1}^n a_t a_{t-k}}{\sum_{t=1}^n a_t^2}; \quad \hat{r}_k = \frac{\sum_{t=k+1}^n \hat{a}_t \hat{a}_{t-k}}{\sum_{t=1}^n \hat{a}_t^2}.$$

Ora, sotto H_0 , il vettore delle v.c. r_k è distribuito come una v.c. Normale multivariata con vettore medio nullo e matrice di varianze e covarianze diagonale tale che: $\text{var}(r_k) = n^{-1} * (n-k) / (n+2) \simeq n^{-1}$. Allora, per un prefissato m , Box e Pierce propongono il test: $Q'_m = n \sum_{k=1}^m r_k^2$ che si distribuisce asintoticamente come una v.c. χ^2 con m gradi di libertà. Quando i parametri sono stimati tramite un modello ARMA(p,q) il test equivalente calcolato tramite i residui $\hat{a}_t$,

lavori in un contesto di serie simili ove tale comportamento è stato spesso osservato ed è ragionevole sul piano della dinamica dei fenomeni osservati.

²⁷La complessità è ben evidente nei lavori citati in Anderson (1971) e non ci pare realistico riferirsi all'ipotesi di circolarità con serie reali, in quanto tale assunto formale non trova altra giustificazione che la semplificazione analitica delle distribuzioni e degli autovalori coinvolti. Una idea delle difficoltà computazionali ed algebriche di risultati esatti, poi spesso "corretti", si può avere consultando, fra gli altri, i lavori di De Gooijer (1980) e di Dufour e Roy (1984).

cioè: $Q'_m = n \sum_{k=1}^m \hat{r}_k^2$ si distribuisce asintoticamente come una v.c. χ^2 con $m - (p+q)$ gradi di libertà.

Grazie ai lavori di Prothero & Wallis (1976) e Davies et al. (1977), e ad una consistente evidenza empirica, si constatò che i livelli effettivi di significatività del test Q'_m erano regolarmente più bassi di quelli asintotici. Per questo, Ljung e Box (1978) hanno proposto di inserire il fattore $(n+2)/(n-k)$ nel test Q'_m pervenendo, quando si tratta di autocorrelazioni stimate dai residui di un modello ARMA(p,q) al test:

$$Q_m = n(n+2) \sum_{k=1}^m \hat{r}_k^2 / (n-k)$$

il quale si distribuisce asintoticamente ancora come una v.c. χ^2 con $m - (p+q)$ gradi di libertà.

Il test di Ljung e Box presenta livelli di significatività nominali più vicini a quelli teorici ed è decisamente più potente del test Q'_m , realizzando anche una maggiore fedeltà complessiva tra distribuzione empirica e densità della corrispondente v.c. χ^2 : esso, pertanto, va preferito quale test sui residui da un modello per serie storiche²⁸. Successive modifiche di Ljung (1986) includono l'approssimazione della distribuzione di Q_m con quella di $a\chi_b^2$ ove $a = \Sigma \lambda_i^2 / \Sigma \lambda_i$, $b = (\Sigma \lambda_i)^2 / \Sigma \lambda_i^2$, essendo λ_i gli autovalori della matrice di varianze e covarianze degli r_k , riportata in McLeod (1986) e Ljung (1986). Anche se le elaborazioni occorrenti per implementare tali risultati non sono complesse, queste non ci sembrano giustificate nelle serie con sufficiente numerosità.

Il test Q_m è stato oggetto di varie proposte di modifica, anzitutto, migliorando la distribuzione di $\hat{r}_k$, come in Ali (1984) che adatta densità della famiglia di Pearson (preferibili anche rispetto allo sviluppo in serie di Edgeworth). In alternativa, Newbold (1980)

²⁸La raccomandazione non è proprio banale visto che sia software che articoli, in ambito statistico ed econometrico, senza alcuna giustificazione, continuano a presentare e pubblicare i risultati del test di Box e Pierce, anche dopo più di venti anni dalla proposta del test di Ljung e Box.

ha proposto un test per bassi valori di m fondato su $S_m = n \hat{r}' \hat{C}^{-1} \hat{r}$, (essendo C la matrice calcolata da McLeod, 1978 nel caso di modelli ARMA), ma esso presenta l'inconveniente di diventare presto quasi singolare. Mélard e Roy (1987) considerano la distribuzione asintotica di $\hat{r}' \Sigma^{-1} \hat{r}$ (che è di tipo χ^2) ove gli elementi di Σ sono deducibili dalla ben nota formula di Bartlett (1946) ma sono stimati mediante covarianze perequate con la finestra di Bartlett modificata o di Parzen.

Godfrey (1979) rileva i difetti del test di Box e Pierce e si pone il problema della adeguatezza di un modello ARMA contro le due alternative²⁹: $H_A: \text{ARMA}(p+r, q)$ e $H_B: \text{ARMA}(p, q+r)$ e, grazie al LM test, ottiene una procedura piuttosto semplice per l'analisi.

Se $\hat{a}_t$ sono i residui stimati sotto $H_0: \text{ARMA}(p, q)$, allora il test LM contro $H_A: \text{ARMA}(p+r, q)$ equivale al calcolo di $n^* R^2$ nella regressione stimata con il metodo dei minimi quadrati per il modello seguente:

$$\hat{a}_t = \alpha_1 x_{t-1} + \dots + \alpha_{p+r} x_{t-p-r} + \beta_1 y_{t-1} + \dots + \beta_q y_{t-q} + \epsilon_t, t = 1, 2, \dots, n$$

mentre il test LM contro $H_B: \text{ARMA}(p, q+r)$ equivale al calcolo di $n^* R^2$ nella regressione stimata con il metodo dei minimi quadrati per il modello seguente:

$$\hat{a}_t = \alpha_1 x_{t-1} + \dots + \alpha_p x_{t-p} + \beta_1 y_{t-1} + \dots + \beta_{q+r} y_{t-q-r} + \eta_t, t = 1, 2, \dots, n.$$

Nei due modelli precedenti, abbiamo definito: $x_t = y_t = 0$, per $t < 0$ ed inoltre:

$$x_t = -z_t + \hat{\theta}_1 x_{t-1} + \dots + \hat{\theta}_q x_{t-q}, t = 1, 2, \dots, n$$

²⁹Si osservi che Hannan (1970), 388-389; 409-414, ha dimostrato che non si può verificare l'ipotesi di un modello ARMA di ordini prefissati contro l'alternativa di ordini ARMA entrambi differenti perché è singolare, in tal caso, la matrice di informazione. Tuttavia, Poskitt & Tremayne (1980) hanno imposto talune restrizioni su un test LM che consentono anche in tal caso un suo utilizzo, purché lo si interpreti come un test di pura significatività, essendo valido per una classe molto ampia di alternative.

$$y_t = \hat{a}_t + \hat{\theta}_1 y_{t-1} + \dots + \hat{\theta}_q y_{t-q}, t = 1, 2, \dots, n.$$

Godfrey osserva che se il processo Z_t è AR(p) allora: $y_t = -x_t + \hat{\phi}_1 x_{t-1} + \dots + \hat{\phi}_p x_{t-p}$, per cui il coefficiente di determinazione R^2 per i due precedenti modelli di regressione coincide, e il test LM di H_0 : AR(p) contro H_A : AR(p+r) equivale³⁰ a quello contro l'alternativa H_B : ARMA(p,r).

Tuttavia, Newbold (1980) ha dimostrato che il test di Godfrey e il test di Box e Pierce (e, quindi, quello di Ljung e Box) sono perfettamente equivalenti³¹: cioè esaminare le prime r autocorrelazioni dei residui da un modello ARMA(p,q) mediante un test portmanteau ovvero adattarvi r parametri aggiuntivi mediante un test LM equivale asintoticamente alla stessa regione critica: Ljung (1988). Una derivazione unitaria del test di Box e Pierce e del test di Quenouille (1947) per processi AR(p) era già presente in Hosking (1978) quale test LR contro particolari ipotesi alternative.

Un problema rilevante è la determinazione di m, il lag massimo cui applicare il test di Ljung e Box e su esso agiscono considerazioni contrapposte. Da un lato, si cerca di includere un certo numero di $\hat{r}_k$ per tener conto di specificazioni alternative rilevanti ai fini della costruzione di un modello (per esempio, su dati mensili, si cerca di giungere sino a m=24 per tener conto di un effetto stagionale persistente): peraltro, sono comuni i pacchetti statistici che indicano in m=20 una opzione "opportuna". D'altra parte, se m=1 il test

³⁰Il lettore si renderà conto della contemporanea correttezza e debolezza di questa affermazione in un contesto di analisi delle serie storiche. Difatti, mentre è ovvio che ciò avviene perché il test LM tiene conto nella sua derivazione solo dell'ipotesi H_0 , è limitante -solo per fare un esempio- affermare che un test di H_0 :AR(1) contro H_A :AR(2) equivale ad un test di H_0 :AR(1) contro H_B :ARMA(1,1) perché nessuno statistico -con un minimo di esperienza di funzioni di autocorrelazione (globale e parziale)- potrebbe ritenere equivalenti le due situazioni.

³¹L'equivalenza anche per un test spettrale del tipo LM trovasi in Tuan (1986) mentre il lavoro di Saikkonen (1983) confronta alcuni test di adattamento basati sulle frequenze i quali, tuttavia, presentano efficienza relativa asintotica nulla rispetto al corrispondente test sulle autocorrelazioni (Box e Pierce), come alcune osservazioni di Milhoj (1981) sembrano confermare.

è pressoché perfetto essendo equivalente alla significatività di un parametro aggiuntivo di una regressione corrispondente ad un modello alternativo; inoltre, tanto più è vera H_0 tanto più è dannosa la considerazione di molti $(\hat{r}_k)^2$ nella sommatoria che definisce Q_m , aggiungendo essi una inutile variabilità campionaria.

In effetti, i lavori citati e gli esperimenti di simulazione riportati sono concordi nel ritenere che m deve essere sufficientemente piccolo per le analisi. Quando si vogliono verificare ipotesi precise (connesse alla stagionalità per esempio) allora è preferibile indagare sulle singole autocorrelazioni, come in McLeod (1978, 301-302), o su gruppi di autocorrelazioni, come in Godolphin (1978). Conclusioni analoghe sono state ottenute da Battaglia (1990) che, determinando m in funzione della potenza approssimata dei test portmanteau, valuta in $m=5$ una sorta di valore-soglia al di sopra del quale la potenza sembra diminuire per varie ipotesi alternative³². Va, infine, citata la regola $m=\mathcal{O}(\sqrt{n})$, spesso usata negli sviluppi formali, ma ampiamente disattesa nell'analisi di serie reali.

Gli approcci precedenti fondano i test sulle funzioni di autocorrelazione stimate sui residui dal modello non solo perché ciò deriva dalla teoria asintotica dei test ma anche perché nella procedura di Box e Jenkins la quantità $\hat{r}_k$ è comunque calcolata per orientare verso una specificazione alternativa. Tuttavia, alcune semplici varianti possono aiutare nella fase decisionale.

In primo luogo, si potrebbe utilizzare la trasformazione $\mathcal{Z}$ di Fisher per le autocorrelazioni. Poiché $\hat{\mathcal{Z}}_k = 0.5 * \log[(1 + \hat{r}_k)/(1 - \hat{r}_k)]$ possiede distribuzione più simmetrica e vicina alla Normalità degli originari $\hat{r}_k$ si può ipotizzare che quantità del tipo $n * \sum_0^\infty \hat{\mathcal{Z}}_i^2$ siano più vicine, per campioni finiti, alla distribuzione asintotica χ^2 . Ciò è stato recentemente applicato da Chen & Mudholkar (1990) per un

³²poiché molti risultati sembrano dipendere strettamente dalla scelta di m , ci sorprende -come rilevato da Battaglia (1990), 337-338- che all'argomento sia stato dedicato così poco spazio nella letteratura. Inoltre, manca una analisi sistematica delle connessioni tra m e la collocazione dei parametri sulla regione ammissibile mentre ciò sembra essere un dato rilevante sia nelle analisi teoriche che nei risultati sperimentali. Una soluzione potrebbe essere quella di individuare proposte "adattive" per la determinazione di un m ottimale.

test di completa indipendenza in dati multinomiali. Tuttavia, le risultanze empiriche mostrano un quasi perfetto parallelismo tra questo approccio e quello dei test di Box e Pierce e di Ljung e Box, per cui sia in termini di significatività che di potenza, tale nuova misura non sembra giustificata.

Un test analogo a quello di Ljung e Box, ma basato sulla stima della funzione di autocorrelazione parziale, è stato introdotto da Monti (1994) e si rivela preferibile soprattutto in termini di potenza per le componenti MA. In effetti, la proposta trova giustificazione nel fatto che le stime della funzione di autocorrelazione parziale (derivabili dall'algoritmo di Durbin-Levinson) costituiscono un insieme di stimatori asintoticamente indipendenti e consentono una parametrizzazione efficiente del processo: Hannan (1960), 93-97; Barndorff-Nielsen & Schou (1973); Box, et al. (1972). Con finalità analoghe, Battaglia e Baragona (1993) utilizzano l'indice di determinismo lineare sfruttando, quindi, sia le stime delle funzioni di autocorrelazione globale che quelle inverse dei residui (per la distribuzione di quest'ultima, che replica quella di r_k , si veda Hosking (1980) nonché la monografia di Battaglia, 1984).

La nostra convinzione di fronte a tali molteplici proposte, tutte basate sulla correlazione seriale dei residui stimati da un modello, è che il problema di fondo sia quello di pervenire alla migliore sintesi (in termini di efficienza, potenza, parsimonia, etc.) delle informazioni. Queste sono tutte racchiuse nella funzione di autocorrelazione globale che, va sottolineato, consente una stima non-parametrica consistente e notevolmente efficiente a partire dai dati originari. Tali proposte, in vario modo, modificano la sintesi in funzione dell'ipotesi alternativa (componenti AR, MA, miste, etc.) ma si potrebbe pervenire ad una soluzione "adattiva" che le includa tutte e lasci alle osservazioni la specificazione del test più idoneo.

D'altra parte, la stessa funzione di autocorrelazione dei residui contiene ulteriori informazioni, come è dimostrato da altre proposte. Per la sua semplicità, ricordiamo quello di Hokstad (1983), che esamina la funzione di cross-correlazione tra residui stimati e serie ori-

ginaria mediante: $s_k = \text{Corr}(a_t, z_{t-k})$, la quale viene stimata da:

$$\hat{s}_k = \frac{\sum_{t=k+1}^n \hat{a}_t (z_{t-k} - \bar{z})}{\sum_{t=1}^n \hat{a}_t^2 \sum_{t=1}^n (z_t - \bar{z})^2}.$$

Essa va utilizzata con cautela essendo l'intervallo $\pm 2/\sqrt{n}$ solo un limite superiore della varianza asintotica.

Fra i numerosi altri lavori sull'argomento, ricordiamo quelli sul confronto di potenza di alcuni test di bontà di adattamento condotto da Clarke & Godolphin (1982), e nel caso multivariato da Poskitt & Tremayne (1986), su test diagnostici di tipo bi-spettrale per una dipendenza non-lineare da Ashley et al. (1986), sul calcolo dei LM test con procedure di stima efficienti da Pham (1987).

Non ci occuperemo dei test di correlazione seriale basati sui ranghi (ottenuti sostituendo i ranghi ai valori di z_t nelle usuali formule di autocorrelazione) ma ci limitiamo a citare Bartels (1982), che esamina una versione sui ranghi del rapporto di von Neumann, e Chan e Tran (1992), nel quale l'ipotesi di dipendenza seriale è sottoposta a test in modo non parametrico mediante il confronto tra la distribuzione empirica doppia e quelle marginali (il cui prodotto, sotto H_0 , non dovrebbe differire significativamente dalla distribuzione doppia). Sul piano analitico, la ottimalità di tale classe di test (sia in generale che per verificare ipotesi su modelli ARMA) è stata esaminata rigorosamente in Hallin et al. (1985) e in Hallin e Puri (1988) anche mediante il confronto con i test usuali basati sulla correlazione seriale dei valori osservati. Una rassegna sul tema, piuttosto ampia, è quella di Hallin e Puri (1992).

Infine, anche se non affronteremo le problematiche dei test per modelli non-lineari³³, non ci possiamo esimere dal ricordare la sem-

³³Una rassegna critica ed unitaria dei test più diffusi in ambito non-lineare è contenuta in Corduas (1994) mentre i riferimenti essenziali per lo studio dei processi non-lineari includono i contributi di Granger e Andersen (1978), Subba Rao (1981), Priestley (1980), Tong (1990), Tsay (1986, 1989) e le recenti proposte di Chen e Tsay (1993) nonché i numerosi test e confronti riportati in tali lavori. Va tuttavia rilevato che la letteratura sui modelli non-lineari è immensa e diffusa nei più diversi settori di applicazione, generando spesso una modellistica ad hoc, difficile da generalizzare.

plicità di implementazione di test sulle autocorrelazioni dei residui al quadrato, come in Granger e Andersen (1978) e Maravall (1983). Inoltre, McLeod e Li (1983) mostrano che tale test è asintoticamente distribuito come una v.c. χ_m^2 perché le stime delle funzioni di autocorrelazione globale dei residui al quadrato (a differenze delle omologhe quantità senza i quadrati) sono asintoticamente multinormali e indipendenti con varianza n^{-1} . Pertanto, un test del tipo:

$$Q_{aa} = n(n+2) \sum_{k=1}^m (\hat{r}_{aa,k})^2 / (n-k)$$

ove: $\hat{r}_{aa} = \sum_{t=k+1}^n (\hat{a}_t^2 - \hat{\sigma}^2)(\hat{a}_{t-k}^2 - \hat{\sigma}^2) / \sum_{t=1}^n (\hat{a}_t^2 - \hat{\sigma}^2)^2$, con $\hat{\sigma}^2 = \sum_{t=1}^n \hat{a}_t^2 / n$, si può utilizzare come uno dei più semplici test di non-linearità dei residui.

10. Test su dati anomali e influenti in serie storiche

Avendo escluso dalla nostra trattazione i test su modelli non-lineari, non possiamo approfondire tale questione se non per evidenziare il rischio di confondere un gruppo di dati anomali con un meccanismo generatore di tipo non-lineare: la differenza interpretativa e previsiva delle due situazioni è, difatti, abbastanza netta.

In presenza di outliers, gli approcci da seguire sono almeno due: sottoporre a test i dati per individuarli, modificarli e/o spiegarli oppure modificare le procedure di stime, rivolgendosi a tecniche robuste. Nello specifico delle serie temporali, ciò ha conseguenze sia sui test di identificazione che su quelli di verifica, sempre nella direzione di una inflazione di varianza e in una decrescita di correlazione seriale. Aderendo ad una impostazione operativa, ci limitiamo a distinguere, nel seguito, solo alcuni test più semplici distinguendo, per comodità espositiva, tra test su dati anomali (outliers) e ricerca di dati influenti (influential data).

Per il primo tipo di test, ricordiamo i lavori pioneristici di Fox (1972), Box e Tiao (1975) e gli importanti contributi della scuola di Tiao, fra cui il lavoro di Tsay (1988), e inoltre quelli di Baragona

(1990) e Barbieri (1998); sul versante multivariato, ci limitiamo a segnalare l'articolo di Abraham (1980) e le proposte di Barbieri (1991). Per la sua immediatezza, presentiamo un test di Abraham e Yatawara (1988) che può essere agevolmente implementato.

Dato un processo ARMA, la presenza di un outlier al tempo $t=k$ può essere accertata ottenendo anzitutto la sua formulazione $AR(\infty)$: $\pi(B)Z_t = a_t$ e quindi calcolando la statistica

$$T_{n,k} = a_k^2/\sigma^2 + \left(\sum_{j=1}^{\infty} \pi_j a_{k+j} \right)^2 / \left(\sigma^2 \sum_{j=1}^{\infty} \pi_j^2 \right), k = 1, 2, \dots, n$$

che, sotto H_0 , è distribuita asintoticamente come una v.c. χ^2 con 2 gradi di libertà. Se invece k è incognito, gli Autori suggeriscono di considerare la statistica: $T_n = \max(T_{n,k})$ che, per un processo $AR(p)$ e un livello di significatività α , va confrontata con il livello critico $t_{\alpha,p}$ ben approssimato da: $\hat{t}_{\alpha,p} = 2[\log(n-2p) - \log(-\log(1-\alpha)) + \log(0.8)]$. Il test può essere utilizzato anche per discriminare fra un outlier additivo e innovativo.

Infine, Ledolter (1990) ha proposto un test su outliers molto semplice ed efficace suggerendo di esaminare, rispetto a $t=1,2,\dots,n$, il grafico di $W_{(t)}$ definito da

$$W_{(t)} = n \left(\frac{\hat{\sigma}^2}{\hat{\sigma}_{(t)}^2} - 1 \right)$$

e basato sul confronto tra due varianze campionarie: quella stimata da tutte le osservazioni e quella stimata considerando z_t come "dato mancante". Ledolter mostra che, asintoticamente, un test sul singolo outlier può essere ottenuto considerando $W_{(t)}$ come un v.c. χ^2 con 1 grado di libertà, per cui il grafico di $W_{(t)}$ con una linea orizzontale posta a $\chi_{\alpha}^2(1)$ può essere una sorta di segnale per z_t . Invece, per un test complessivo su un centinaio di osservazioni (sulla base di una estesa simulazione), viene suggerito di porre tale linea a circa 13.

I lavori che misurano l'effetto dei singoli dati sulla stima dei parametri, si collocano nell'ambito delle metriche tra processi lineari

come in Peña (1984, 1987, 1989) e Corduas (1989) che, entrambi, propongono di stimare l'effetto di un singolo dato trattandolo come "dato mancante": Peña utilizza una metrica di Mahalanobis e Corduas una misura di divergenza tra processi lineari. I due approcci sembrano convergere nei risultati ed entrambi, asintoticamente, possono far riferimento a distribuzioni del tipo χ^2 , anche se i due Autori affrontano il problema, come è tipo dell'analisi di influenza, in un'ottica più descrittiva che inferenziale.

Un approccio differente si fonda sulla introduzione dei criteri di robustezza nell'analisi delle serie storiche, a partire dai lavori di Martin (1981). Va aggiunto che, nonostante un certo numero di contributi, la prassi non è ancora orientata verso l'utilizzazione di misure resistenti agli outliers, forse anche per la difficoltà computazionale di alcune procedure, generalmente di tipo iterativo, nelle quali persiste una componente di soggettività nella scelta delle funzioni-peso. In tale ambito, Masarotto (1987, 1988) ha proposto stimatori robusti delle funzioni di autocorrelazione per una identificazione meno determinata da valori aberranti, mentre Li (1988) ripercorre tutte le tappe dei portmanteau test e dei LM test per dimostrare che un test robusto (equivalente di quello di Ljung-Box) è utilizzabile sostituendo alla tradizionale stima della funzione di autocorrelazione dei residui la quantità, proposta da Bustos & Yohai (1986):

$$\hat{r}_k = \Sigma_t \eta(\hat{a}_t/\hat{\sigma}, \hat{a}_{t-k}/\hat{\sigma})$$

dove $\hat{\sigma}$ è uno stimatore robusto di scala mentre $\eta(.,.)$ è una funzione dispari dei suoi argomenti. In tal modo, la versione robusta del test di Ljung-Box diventa:

$$Q_m^* = c^{-1} n^2 \sum_{j=1}^m (\hat{r}_j)^2 / (n-j),$$

ove $c = E(\eta^2(a_t/\sigma, a_{t-1}/\sigma))$. Li dimostra che Q_m^* si distribuisce asintoticamente come una v.c. χ^2 con gradi di libertà pari a $m-p-q$. Sulla stessa linea, ma in completa generalità, si collocano le proposte

di Basawa et al. (1985), che presentano versioni robuste dei LM e Wald test in modelli di serie storiche rendendo robuste le corrispondenti equazioni di stima. I risultati riportati per il caso di processi AR e le simulazioni di potenza suggeriscono una certa preferenza per la versione robusta del test di Wald. Infine, Kassam (1982) affronta il problema del test robusto nel contesto della regressione ed interpolazione di serie storiche.

11. La verifica del modello come test globale

Quando si sottopone a verifica un modello econometrico, semplice o complesso, alla significatività dei coefficienti e al controllo del rispetto delle ipotesi su cui è stato stimato, si premette una accurata disamina del significato delle variabili coinvolte, del loro peso reciproco, della giustezza del legame osservato con quello previsto dalla teoria e della sensibilità dei reciproci effetti in scenari previsivi o simulativi. Nei modelli per serie storiche tale controllo sembra assente poiché, si afferma, non sussiste una teoria esplicativa del ruolo delle variabili coinvolte. Per questo, in assenza di informazioni a priori, la fedeltà dei dati al modello viene rigorosamente esaminata mediante una serie articolata e lunga di controlli statistici sulla Normalità, sulla omoschedasticità, sulla incorrelazione dei residui, sulla assenza di valori anomali, sulla stabilità del meccanismo generatore, sulla coerenza delle ipotesi di linearità, e così' via.

Questo atteggiamento ha i suoi rischi perché in una lunga successione di test diventa molto probabile rifiutare incorrettamente H_0 in uno o più casi. Pertanto, esattamente come in Econometria, la decisione circa un modello non dovrebbe essere affidata esclusivamente ad una successione di regioni critiche determinate asintoticamente. In effetti, ogni modello per serie storiche mira a riprodurre la dinamica del fenomeno osservato cogliendone fatti salienti ed esplicitando in una forma parametrica efficienti comportamenti tendenziali, ciclici, stagionali, alternanze, flessi, cambiamenti di livelli. Quindi, ogni modello deve contenere esplicitazioni su tali dinamiche e ciò non può

derivare dal confronto tra statistiche campionarie e regioni critiche ma consiste in una valutazione complessiva, quindi anche qualitativa, nella quale i test presentati intervengono in modo decisivo ma non assoluto³⁴.

In tale ottica si pone anche il recente contributo di Harvey (1992) che, assieme ai residui ordinari da un modello esamina dei residui “ausiliari”, ottenuti dalla scissione di un fenomeno nelle sue componenti strutturali. L’applicazione di test tradizionali a tale molteplicità di residui assume poco senso se interpretati in un’ottica classica (essi sono, infatti, strettamente dipendenti e deducibili dal modello e dai residui tradizionali) ma veicolano, invece, informazioni importanti se utilizzati per giudicare della adeguatezza della decomposizione in componenti strutturali predefinite.

12. Considerazioni finali

E’ noto che la teoria classica dell’inferenza statistica non può essere applicata all’analisi delle serie storiche perché una successione di dati temporali non può essere concepita come realizzazione di campioni casuali con componenti indipendenti e identicamente distribuiti: ciò costituisce l’atto di nascita della “Analisi delle serie storiche” come settore ampio e tipico della disciplina statistica.

Tale constatazione può apparire riduttiva se il vincolo dinamico viene interpretato solo come una difficoltà della ricerca metodologica mentre invece esso costituisce uno degli aspetti più costruttivi e rilevanti sia della vita quotidiana che della ricerca scientifica. Difatti, è solo perché i dati sono correlati che possiamo parlare di controllo, previsione e simulazione con intervalli di confidenza più ristretti di quelli deducibili dalla media di osservazioni bernoulliane; è solo perché i dati non sono affatto indipendenti che riusciamo a sostituire dati

³⁴Dovrebbe essere evidente che noi interpretiamo tale momento di verifica del modello come l’equivalente dell’esame critico di una relazione causale di un modello econometrico e della sua coerenza esplicativa in rapporto alla teoria economica.

mancauti con valori molto più pregnanti e interpretabili della media delle osservazioni; è perché i dati costituiscono parte finita di un processo stocastico (che auspichiamo sia differente da un processo WN) che noi continuiamo a credere che destagionalizzazione ed analisi congiunturali (operazioni intrinsecamente arbitrarie) abbiano una valenza operativa ed una utilità oggettiva per decisioni di politica economica e finanziaria; e gli esempi potrebbero continuare.

Quindi, i test statistici analizzati (e i tanti appena sfiorati) sono solo una modalità di sintesi decisionale su fenomeni strettamente connessi in rapporto ad una struttura semplificatrice, quale è il modello statistico. La numerosità e diversità dei test presentati è giustificata dalla gamma dei modelli che si utilizzano, dal basso costo computazionale che oggi si sostiene anche per elaborazioni complesse, dalla necessità di offrire molte informazioni ad utenti inesperti e troppo spesso “neutrali”.

Tuttavia, l'intera gamma dei test che si applicano dipende dagli stessi dati, cioè dalle stesse informazioni e dal confronto tra informazioni possedute e riferimento modellistico entro cui racchiudere tali informazioni. Ciò che realmente cambia da un test all'altro non è l'insieme delle osservazioni ma il contesto entro il quale gli stessi dati si ritengono spiegati. Nell'attuale momento dell'indagine statistica, che si ricollega non casualmente al crollo di teorie e modelli onnicomprensivi, non è negativo che ciascuno elabori e rielabori le proprie informazioni costruendo nuovi e più raffinati schemi di riferimento con i quali confrontare le proprie osservazioni. Andrebbe tuttavia sottolineato che il nucleo di informazioni posseduto è limitato e che la sua validità informativa non cresce con la ricchezza delle ipotesi ma viene solo non-rifiutata rispetto a certe alternative. Stiamo dicendo, per banalizzare, che per 100 osservazioni temporali esiste almeno una funzione polinomiale che le spiega tutte ma nessuno è interessato ad utilizzarla se non per esercitarsi sui polinomi di Lagrange. Ci interessa, invece, descrivere la realtà per prevederla, collegarla ad altri contesti, sezionarla in componenti non osservabili, in breve: racchiuderla in schemi semplici ed utilizzabili.

Alla fine, la grande pletora di test statistici è sempre funzione

dei primissimi momenti semplici e misti della serie osservata $\{z_t, t=1,2,\dots,n\}$; se Z_t è Gaussiano, questo implica che ogni analisi sulla funzione di autocorrelazione ρ_k (comunque trasformata: parziale, inversa, spettro, etc.) è idonea in linea di principio a sintetizzare Z_t . Se esiste una moltitudine di test è solo perché ci interessano alternative specifiche, oppure perché il rifiuto di una ipotesi ci sembra più convincente quando deriva da angolature differenti. La teoria asintotica non fa che confermare questo principio elementare individuando medesime distribuzioni per test che nascono e si calcolano in modo differente. Nella sua banalità ciò deriva dal fatto che i dati non generano nuove informazioni.

E' tempo, allora, di ricercare procedure e test derivati dalle serie e confrontati direttamente con esse, come nel contributo di Tsay (1992). Un modello non è un fatto oggettivo ma solo uno strumento predisposto per uno specifico obiettivo da perseguire utilizzando i dati temporali. Per questo la definizione degli obiettivi, che sono spesso sintetizzabili tramite funzionali statistici, precede e specifica il modello statistico per i dati, ne determina le caratteristiche, individua i test più convenienti per la identificazione e consegue il suo controllo mediante una procedura di bootstrap parametrico.

Crediamo che il futuro della ricerca sia nella predisposizione di strumenti di analisi, e quindi di test, direttamente connessi alle ripartizioni empiriche delle osservazioni e a tecniche di ricampionamento parametrico che confrontino, senza sovrastrutture, i modelli con le osservazioni. Alla base vi è la consapevolezza che se il modello è adeguato per generare le osservazioni che possediamo esso, replicato, deve riprodurre le caratteristiche temporali, frequenziali, strutturali. In tal modo, i test delle ipotesi statistiche sui modelli per serie storiche diverranno strumenti sempre più utili nella fase operativa della ricerca.

Ringraziamenti: La ricerca è stata realizzata grazie a parziali finanziamenti del CNR e del MURST nonché ai progetti di ricerca dell'Ateneo Federico II nei quali l'Autore è coinvolto. La presente stesura ha usufruito

di numerosi commenti costruttivi da parte di numerosi statistici ed econometrici italiani fra i quali ci piace segnalare l'impegno del prof. Mario Faliva che, per un intero triennio, ha stimolato molti di noi ad occuparci del fecondo ed incessante legame esistente tra Statistica ed Econometria.

Riferimenti Bibliografici

Abelson R.P., Tukey J.W. (1963) Efficient utilization of non-numerical information in quantitative analysis: general theory and the case of simple order, *Annals of Mathematical Statistics*, 34, 1347-1369

Abraham B. (1980) Intervention Analysis and Multiple Time Series, *Biometrika*, 67, 1, 73-78

Abraham B., Yatawara N. (1988) A Score Test for Detection of Time Series Outliers, *Journal of Time Series Analysis*, 9, 2, 109-119

Akaike H. (1974) A New Look at the Statistical Model Identification, *IEEE Transaction on Automatic Control*, 19, 716-723

Akaike H. (1976) Canonical Correlation Analysis of Time Series and the Use of an Information Criterion, in: R.K. Mehra & D.G. Lianiotis (editors), *System Identification: Advances and Case Studies*, Academic Press, London & New York, 27-96

Alagon J. (1989) Spectral Discrimination for two Groups of Time Series, *Journal of Time Series Analysis*, 10, 3, 203-214

Ali M.M. (1984) Distribution of the Sample Autocorrelations when the Observations are from a Stationary Autoregressive-Moving-Average Process, *Journal of Business & Economic Statistics*, 2, 3, 271-278

Anderson T.W. (1971) *The Statistical Analysis of Time Series*, J. Wiley & Sons, New York

Ansley C.F. e Newbold P. (1979); On the finite sample distribution of residual autocorrelations in autoregressive-moving average models, *Biometrika*, 66, 547-53

Ashley R. A., Patterson D.M. e Hinich M.J. (1986) A Diagnostic Test for Nonlinear Serial Dependence in Time Series Fitting Errors,

Journal of Time Series Analysis, 7, 3, 165-178

Atkinson A.C. (1970) A Method for Discriminating between Models, *Journal of the Royal Statistical Society, B*, 32, 323-344

Atkinson A.C., Lawrence A.J. (1989) A Comparison of Asymptotically Equivalent Test Statistics for Regression Transformation, *Biometrika*, 76, 2, 223-229

Attanasio M. e Wong W.K. (1995) Transformations for Stabilizing Spread in a Time Series Model: A Comparative Study, *Journal of Statistical Computations and Simulations*, 22, 55-70

Azzalini A. (1984) Estimation and Hypothesis Testing for Collections of Autoregressive ime Series, *Biometrika*, 71, 1, 85-90

Baragona R. (1990) Diverse impostazioni inferenziali per lo studio di dati anomali nelle serie temporali, *Pubblicazioni del Dipartimento di Statistica, Probabilità e Stat. Applic., Serie A*, n.31, Roma

Barbieri M.M. (1991) Outliers in serie temporali multivariate, *Quaderni di Statistica e Econometria*, XIII, 55-67

Barbieri M.M. (1998) Additive and Innovational Outliers in Autoregressive Time Series: A Unified Bayesian Approach, *STATISTICA*, LVIII, 3, 394-409

Bardoff-Nielsen O., Schou G. (1973) On the Parametrization of Autoregressive Models by Partial Autocorrelations, *Journal of Multivariate Analysis*, 3, 408-419

Bartels R. (1982) The Rank Version of von Neumann's Ratio Test for Randomness, *Journal of the American Statistical Association*, 77, 40-46

Bartlett M.S. (1946) On the Theoretical Specification and Sampling Properties of Autocorrelated Time Series, *Journal of the Royal Statistical Society, B*, 8, 27-41;85-95 (corrigenda:1948, 10, 100)

Basawa I.V., Higgins R.M. e Staudte R.G. (1985) Robust Tests for Time Series with an Application to First-order Autoregressive Processes, *Biometrika*, 72, 3, 559-571

Battaglia F. (1983) Inverse Autocovariances and a Measure of Linear Determinism for a Stationary Process, *Journal of Time Series Analysis*, 4, 2, 79-87

Battaglia F., Baragona R. (1992) A White Noise Test for Time

Series based on the Index of Linear Determinism, Progetto MURST 40di Roma "La Sapienza", *manoscritto*

Béguin J.M., Gouriéroux C. e Monfort A. (1980) Identification of Mixed Autoregressive-Moving Average Process: the Corner Method, in: O.D.Anderson (editor), *Time Series*, North-Holland, Amsterdam, 423-426

Beran J. (1992) A Goodness-of-fit Test for Time Series with Long Range Dependence, *Journal of the Royal Statistical Society, B*, 54, 3, 749-760

Bhansali R.J. (1974) Asymptotic Properties of the Wiener-Kolmogorov Predictor, I., *Journal of the Royal Statistical Society, B*, 36, 61-73

Bhansali R.J. (1983) Estimation of the Order of a Moving Average Model from Autoregressive and Window Estimates of the Inverse Correlation Function, *Journal of Time Series Analysis*, 4, 3, 137-162

Bhansali R.J. (1993) Estimation of the Prediction Error Variance and an R^2 Measure by Autoregressive Model Fitting, *Journal of Time Series Analysis*, 14, 2, 125-146

Bowman K.O., Shenton L.R. (1975) Omnibus Contours for Departures from Normality Based on b_1 and b_2 , *Biometrika*, 62, 243-250

Box G.E.P., Cox D.R. (1964) An Analysis of Transformations (with discussion), *Journal of the Royal Statistical Society, B*, 26, 211-243

Box G.E.P., Jenkins G.M. (1970) *Time Series Analysis: Forecasting and Control*, Holden-Day, S.Francisco (2nd revised edition, 1976)

Box G.E.P., Jenkins G.M. e Guttman I. (1972) Partial Autocorrelations from a Bayesian Viewpoint and Orthogonal Parametrization, *Metron*, XXX, 1-4, 87-112

Box G.E.P., Pierce D.A. (1970) Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Models, *Journal of the American Statistical Association*, 65, 232, 1509-1529

Box G.E.P., Tiao G.C. (1975) Intervention Analysis with Applications to Economic and Environmental Problems, *Journal of the American Statistical Association*, 70, 70-79

Breidt F.J., Davis R.A. (1992) Time-Reversibility and Indepen-

dence of Innovations for Stationary Time Series, *Journal of Time Series Analysis*, 13, 5, 377-390

Breusch T.S., Pagan A.R. (1980) The Lagrange Multiplier Test and its Applications to Model Specification in Econometrics, *Review of Economic Studies*, 47, 239-257

Brillinger D.R. (1989) Consistent Detection of a Monotonic Trend Superposed on a Stationary Time Series, *Biometrika*, 76, 1, 23-30

Brockwell P.J., Davis R.A. (1987) *Time Series: Theory and Methods*, Springer-Verlag, New York

Burg J.P. (1978) A New Analysis Technique for Time Series Data, in: *Modern Spectrum Analysis*, D.G.Childers (editor), IEEE Press, New York, 42-48

Buse A. (1982) The Likelihood Ratio, Wald, and Lagrange Multiplier Tests: an Expository Note, *The American Statistician*, 36, 3, 153-157

Bustos O.H., Yohai V.J. (1986) Robust Estimates for ARMA Models, *Journal of the American Statistical Association*, 81, 155-168

Cappuccio N., Orsi R. (1991) *Econometria*, Il Mulino, Bologna

Carroll R.J. (1982) Two Examples of Transformations when there are Possible Outliers, *Applied Statistics*, 31, 2, 149-152

Chan N.H., Tran L.T. (1992) Nonparametric Tests for Serial Dependence, *Journal of Time Series Analysis*, 13, 1, 19-28

Chen R., Tsay R.S. (1993) Functional-Coefficient Autoregressive Models, *Journal of the American Statistical Association*, 88, 241, 298-308

Chen S., Mudholkar G. S. (1990) Null Distribution of the Sum of Squared z-Transforms in Testing Complete Independence, *Annals of the Institute of Statistical Mathematics*, 42, 1, 149-155

Chiu S.T. (1989) Detecting Periodic Components in a White Gaussian Time Series, *Journal of the Royal Statistical Society, B*, 51, 2, 249-259

Clarke B.R., Godolphin E.J. (1982) Comparative Power Studies for Goodness of Fit Tests of Time Series Models, *Journal of Time Series Analysis*, 3, 3, 141-151

Coates D.S., Diggle P.J. (1986) Tests for Comparing two Esti-

mated Spectral Densities, *Journal of Time Series Analysis*, 7, 1, 7-20

Corduas M. (1989) Detecting Influential Observations in ARIMA Models, *Technical Report*, Centro di Specializzazione e Ricerche, Portici

Corduas M. (1990) Indicatori Strutturali per Processi Stocastici Lineari, *STATISTICA*, L, 2, 199-214

Corduas M. (1992 a) Misure di Distanza tra Serie Storiche e Modelli Parametrici, *Quaderni dell'Istituto Economico-Finanziario*, n.3, Università di Napoli Federico II

Corduas M. (1992 b) On the Distributional Properties of a Distance Criterion for Time Series Models, *Atti del Convegno DISTANCIA '92*, Rennes, 325-328

Corduas M. (1994) Nonlinearity Tests in Time Series Analysis, *Journal of the Italian Statistical Society*, 3, 291-313

Corduas M. (1997) Indirect Inference for Fractional Time Series Models, *Journal of Statistical Computation and Simulation*, 59, 221-232

Cox D. R. (1972) Some further results on separate families of hypotheses, *Journal of the Royal Statistical Society, B*, 38, 45-53

Cox D.R. (1977) Contribution to discussion of paper by A.J. Lawrance and N.T.Kottegoda (1977), *Journal of the Royal Statistical Society A*, 140, 34

Cox D.R., Hinkley D.V. (1974) *Theoretical Statistics*, Chapman & Hall, London

Cox D.R., Solomon P.J. (1988) On Testing for Serial Correlation in Large Numbers of Small Samples, *Biometrika*, 75, 1, 145-148

Cox D.R., Small N.J.H. (1978) Testing Multivariate Normality, 65, 2, 263-272

D'Agostino R.B. (1982) Test for Departures from Normality, in: S.Kotz, N.L.Johnson & Read, C.B. (editors), J.Wiley & Sons, 315-324

D'Agostino R.B. (1971) An Omnibus Test of Normality for Moderate and Large Sample Sizes, *Biometrika*, 58, 341-348

Dahlhaus R. (1985) On the Asymptotic Distribution of Bartlett's Up-Statistic, *Journal of Time Series Analysis*, 6, 4, 213-227

Damsleth E., Spjotvoll E. (1982) Estimation of Trigonometric Components in Time Series, *Journal of the American Statistical Association*, 77, 378, 381-387

Dargahi-Noubary G.R., Laycock P.J. (1981) Spectral Ratio Discriminants and Information Theory, *Journal of Time Series Analysis*, 2, 2, 71-86

Davies R.B., Hart D.S. (1987) Tests for Hurst Effect, *Biometrika*, 74, 1, 95-101

Davies N., Newbold P. (1979) Some Power Studies of a Portmanteau Test of Time Series Model Specification, *Biometrika*, 66, 153-155

Davies N., Triggs C.M. e Newbold P. (1977) Significance Levels of the Box-Pierce Portmanteau Statistic in Finite Sample, *Biometrika*, 64, 517-522

Davis H.T., Jones R.H. (1968) Estimation of the Innovation Variance of a Stationary Time Series, *Journal of the American Statistical Society*, 63, 141-149

de Gooijer J.G. (1980) Exact Moments of the Sample Autocorrelations from Series generated by general ARIMA processes of order (p,d,q) , $d=0$ or 1 , *Journal of Econometrics*, 14, 365-379

de Gooijer J.G., Abraham B. , Gould A. e Robinson L. (1985) Methods for Determining the Order of an Autoregressive-Moving Average Process: a Survey, *International Statistical Review*, 53, 301-329

D'Esposito M.R. (1984) Metodi esplorativi e scelta di una trasformazione per dati stagionali, *Atti XXXII Riunione Scientifica SIS*, Sorrento, I, 359-369

Dickey D.A., Bell W.R. e Miller R.B. (1986) Unit Roots in Time Series Models: Tests and Implications, *The American Statistician*, 40, 1, 12-26

Draper N.R., Cox D.R. (1969) On Distributions and their Transformations to Normality, *Journal of the Royal Statistical Society, B*, 31, 472-476

Draper N.R., Smith H. (1981) *Applied Regression Analysis*, 2nd edition, J.Wiley & Sons, New York

Dufour J.M., Roy R. (1984) Some Robust Exact Results on Sam-

ple Autocorrelations and Tests of Randomness, *Proceedings of the ASA Meeting*, Philadelphia, Pennsylvania

Dzhaparidze K. (1986) *Parameter Estimation and Hypothesis Testing in Spectral Analysis of Stationary Time Series*, Springer-Verlag, New York

Engle R. (1984) Wald, Likelihood Ratio, and Lagrange Multiplier Tests in Econometrics, in: Z.Griliches and M.D.Intrilligator (editors), *Handbook of Econometrics*, II, Elsevier Science Publ. BV, 775-826

Faliva M. (1991) Sur les relations entre passé-present et future des suites économiques, *Contributi alla Ricerca in Econometria*, n.3, Università Cattolica, Milano

Faliva M., Zoia G. (1993) La verifica della struttura causale dei modelli lineari, Commissione SIS su Metodi Statistici ed Econometria, Convegno di Bologna, 27-28 maggio, *manoscritto*

Fassò A. (1989) Rilevazione di rotture in sistemi fisici a componenti aleatorie, *STATISTICA*, XLIX, 4, 569-583

Fassò A. (1990) Su un test portmanteau per la rilevazione on-line di rotture, *Atti XXXV Riunione Scientifica SIS*, Padova, 2, 317-324

Filliben J.J. (1975) The Probability Plot Correlation Coefficient Test for Normality, *Technometrics*, 17, 111-117

Findley D.F., Monsell B.C., Bell W.R. e Otto M.C. (1998) New Capabilities and Methods of the X-12-ARIMA Seasonal Adjustment Program, *Journal of Business and Economic Statistics*, 16, 127-152

Fisher R.A. (1929) Test of Significance in Periodogram Analysis, *Proceeding of the Royal Society of Edinburgh, A*, 125, 54-59

Fox A.J. (1972) Outliers in Time Series, *Journal of the Royal Statistical Society, B*, 34, 350-363

Franses P.H. (1991) The Detection of Observations possibly Influential for Model Selection, *Journal of Time Series Analysis*, 11, 321-325

Gasser T. (1975) Goodness-of-fit tests for Correlated Data, *Biometrika*, 62, 3, 563-570

Geweke J., Porter-Hudak S. (1983) The Estimation and Application of Long Memory Time Series Models, *Journal of Time Series Analysis*, 4, 4, 221-238

Godfrey L.G. (1979) Testing the Adequacy of a Time Series Model, *Biometrika*, 66, 1, 67-72

Godolphin E.J. (1978) A Large-sample Test for Detecting Gaps in Moving Average Models, *Journal of the Royal Statistical Society*, 40, 3, 290-295

Godolphin E.J. (1980) A Method for Testing the Order of an Autoregressive-Moving Average Process, *Biometrika*, 67, 3, 699-703

Gourieroux, C., Monfort A. (1995) *Statistics and Econometric Models*, voll.I e II, Cambridge University Press, Cambridge

Gomez V., Maravall A. (1997) *Programs TRAMO and SEATS*, Instructions for the Users, Updated Version, Banco de Espana, Servicio de Estudios, Madrid

Granger C.W.J. (1978) New Classes of Time Series Models, *The Statistician*, 27, 237-253

Granger C.W.J., Andersen A.P. (1978) *Introduction to Bilinear Time Series Models*, Vandenhoeck and Ruprecht, Gottingen

Granger C.W.J., Joyeux R. (1980) An Introduction to Long-Range Time Series Models and Fractional Differencing, *Journal of Time Series Analysis*, 1, 1, 15-30

Gray H.L., Zhang N.F e Woodward W.A. (1989) On Generalized Fractional Processes, *Journal of Time Series Analysis*, 10, 3, 233-257

Grenander U., Rosenblatt M. (1957) *Statistical Analysis of Stationary Time Series*, J.Wiley & Sons, New York

Haggan, V., Oyetunji O.B. (1984) On the Selection of Subset Autoregressive Time Series Models, *Journal of Time Series Analysis*, 5, 2, 103-113

Hallin M., Ingebleek J.F. e Puri M.L. (1985) Linear Serial Rank Tests for Randomness against ARMA Alternatives, *The Annals of Statistics*, 13, 3, 1156-1181

Hallin M., Puri M.L. (1988) Optimal Rank-Based Procedures for Time Series Analysis Testing an ARMA Model against other ARMA Models, *The Annals of Statistics*, 16, 1, 402-432

Hallin M., Puri M.L. (1992) Rank Tests for Time Series Analysis: a Survey, in D. Brillinger *et al.* (eds.), *New Directions in Time Series Analysis*, part I, 111-153.

- Hamilton J.D. (1995) *Econometria delle serie storiche*, edizione italiana a cura di B. Sitzia, Monduzzi editore, Bologna
- Hannan E.J. (1960) *Time Series Analysis*, Methuen & Co. Ltd, London
- Hannan E.J. (1970) *Multiple Time Series*, J.Wiley & Sons, New York
- Hannan E.J., Nicholls D.F. (1977) The Estimation of the Prediction Error Variance, *Journal of the American Statistical Association*, 72, 834-840
- Hannan E.J., Quinn B.G. (1979) The Determination of the Order of an Autoregression, *Journal of the Royal Statistical Society, B*, 41, 190-195
- Hannan E.J., Quinn B.G. (1989) The Resolution of Closely Adjacent Spectral Lines, *Journal of Time Series Analysis*, 10, 1, 13-31
- Harvey A.C. (1981) *The Econometric Analysis of Time Series*, P.Allan Publisher, Oxford
- Harvey A.C. (1989) *Forecasting, Structural Time Series Analysis Models and the Kalman Filter*, Cambridge University Press, Cambridge
- Harvey A.C., Koopman S.J. (1992) Diagnostic Checking of Unobserved-Components Time Series Models, *Journal of Business and Economic Statistics*, 10, 4, 377-389
- He S., Kadem B. (1990) The Zero-crossing Rate of Autoregressive Processes and its Link to Unit Roots, *Journal of Time Series Analysis*, 11, 3, 201-213
- Hillmer S.C., Tiao G.C. (1979) Likelihood Function of Stationary Multiple Autoregressive Moving Average Models, *Journal of the American Statistical Association*, 74, 652-660
- Hinich M.J. (1982) Testing for Gaussianity and Linearity of a Stationarity Time Series, *Journal of Time Series Analysis*, 3, 3, 169-176
- Hinich M.J., Patterson D.M. (1992) A New Diagnostic Test of Model Inadequacy which uses the Martingale Difference Criterion, *Journal of Time Series Analysis*, 13, 3, 233-252
- Hipel K.W., McLeod A.I. e Lennox W.C. (1977) Advances in Box-

- Jenkins Modeling. 1. Model Construction, *Water Resources Research*, 13, 3, 567-575
- Hokstad P. (1983) A Method for Diagnostic Checking of Time Series Models, *Journal of Time Series Analysis*, 4, 3, 177-183
- Hosking J.R.M. (1978) A Unified Derivation of the Asymptotic Distributions of Goodness-of-fit Statistics for Autoregressive Time-series Models, *Journal of the Royal Statistical Society, B*, 40, 3, 341-349
- Hosking J.R.M. (1980 a) Lagrange Multiplier Tests of Time Series Models, *Journal of the Royal Statistical Society, B*, 42, 170-181
- Hosking J.R.M. (1980 b) The Asymptotic Distribution of the Sample Inverse Autocorrelations of an Autoregressive-Moving Average Process, *Biometrika*, 67, 223-226
- Hosking J.R.M. (1981) Fractional Differencing, *Biometrika*, 68, 1, 165-176
- Hurvich C.M. (1992) Selection of Time Series Models and Spectrum Estimates using a Bias-corrected Generalization of AIC, in: D.Brillinger *et al.* (editors), *New Directions in Time Series Analysis*, part I, Springer-Verlag, New York
- Hurvich C.M., Shumway R. e Tsai C.L. (1990) Improved Estimators of Kullback-Leibler Information for Autoregressive Model Selection in Small Samples, *Biometrika*, 77, 4, 709-719
- Hurvich C.M., Tsai C.L. (1989) Regression and Time Series Model Selection in Small Samples, *Biometrika*, 76, 2, 297-307
- Hurvich C.M., Tsai C.L. (1991) Bias of the Corrected AIC Criterion for Underfitted Regression and Time Series Models, *Biometrika*, 78, 3, 499-509
- Janacek G.J. (1975) Estimation of the Mean Square Error of prediction, *Biometrika*, 62, 175-180
- Janacek G.J. (1982) Determining the Degree of Differencing for Time Series via the Log Spectrum, *Journal of Time Series Analysis*, 3, 3, 177-183
- Jarque C.M., Bera A.K. (1987) A Test for Normality of Observations and Regression Residuals, *International Statistical Review*, 55, 2, 163-172

Jarque C.M., Bera A.K. (1990) Efficient Tests for Normality, Heteroschedasticity and Serial Independence of Regression Residuals, *Economic Letters*, 6, 255-259

Jewell N.P., Bloomfield P. (1983) Canonical Correlations of Past and Future for Time Series: Definitions and Theory, *The Annals of Statistics*, 11, 3, 837-847

Jewell N.P., Bloomfield, P. e Bartmann F.C. (1983) Canonical Correlations of Past and Future for Time Series: Bounds and Computations, *The Annals of Statistics*, 11, 3, 848-855

Jones R.H. (1976) Estimation of the Innovation Generating Variance of a Multivariate Stationary Time Series, *Journal of the American Statistical Association*, 71, 386-388

Judge G.G., Griffiths W.E., Hill R.C., Lütkepohl H. e Lee T.-C. (1985) *The Theory and Practice of Econometrics* (2nd edition), J.Wiley & Sons, New York

Kashyap R.L., Eom, K.B. (1988) Estimation in Long-Memory Time Series Model, *Journal of Time Series Analysis*, 9, 1, 35-41

Kassam S.A. (1982) Robust Hypothesis Testing and Robust Time Series Interpolation and Regression, *Journal of Time Series Analysis*, 3, 3, 185-194

Kavalieris L. (1989) The Estimation of the Order of an Autoregression using Recursive Residuals and Cross-Validation, *Journal of Time Series Analysis*, 10, 3, 271-281

Kedem B. (1980) *Binary Time Series*, M.Dekker, New York

Kedem B. (1987 a) Detection of Periodicities by Higher-Order Crossings, *Journal of Time Series Analysis*, 8, 1, 39-50

Kedem B. (1987 b) Higher-Order Crossings in Time Series Model Identification, *Technometrics*, 29, 2, 193-204

Kedem B., Reed G. (1986) On the Asymptotic Variance of Higher Order Crossings with Special Reference to a Fast White Noise Test, *Biometrika*, 73, 1, 143-149

Kedem B., Slud E. (1981) On Goodness of Fit of Time Series Models: an Application of Higher Order Crossings, *Biometrika*, 68, 2, 551-556

Kedem B., Slud E. (1982) Time Series Discrimination by Higher

Order Crossings, *The Annals of Statistics*, 10, 3, 786-794

Koreisha S., Yoshimoto G. (1991) A Comparison among Identification Procedures for Autoregressive Moving Average Models, *International Statistical Review*, 59, 1, 37-57

Kendall M.G. (1973) *Time Series*, C.Griffin and Co., London

Lawrance A.J. (1987) The Score Statistic for Regression Transformation, *Biometrika*, 74, 275-279

Lawrance A.J. (1991) Directionality and Reversibility in Time Series, *International Statistical Review*, 59, 1, 67-79

Lawrance A.J., Kottegoda N.T. (1977) Stochastic Modelling of Riverflow Time Series (with discussion), *Journal of the Royal Statistical Society, A*, 140, 1-47

Lawrance A.J., Lewis P.A.W. (1992) Reverse Residuals in Autoregressive Time Series Analysis, *Journal of Time Series Analysis*, 13, 3, 253-266

Ledolter J. (1990) Outliers Diagnostics in Time Series Analysis, *Journal of Time Series Analysis* 11, 4, 317-324

Leslie J.R., Stephens M.A. e Fotopoulos S. (1986) Asymptotic Distribution of the Shapiro-Wilk W for Testing for Normality, *The Annals of Statistics*, 14, 4, 1497-1506

Li W.K. (1988) A Goodness-of-fit test in Robust Time Series Modelling, *Biometrika*, 75, 2, 355-361

Li W.K., McLeod A.I. (1986) Fractional Time Series Modelling, *Biometrika*, 73, 1, 217-221

Lii K.S. (1985) Transfer Function Model Order and Parameter Estimation, *Journal of Time Series Analysis*, 6, 3, 153-169

Lii K.S., Tsou, T.-H. (1992) Detecting Sinusoids in Non-Gaussian Noise, *Journal of Time Series Analysis*, 13, 5, 391-409

Lin C.C., Mudholkar G.S. (1980) A Simple Test for Normality against Asymmetric Alternatives, *Biometrika*, 67, 2, 455-461

Linnet K. (1988) Testing Normality of Transformed Data, *Applied Statistics*, 37, 2, 180-186

Ljung G.M. (1986) Diagnostic Testing of Univariate Time Series Models, *Biometrika*, 73, 3, 725-730

Ljung G.M. (1988) On the Lagrange Multiplier Test for Autore-

gressive Moving-Average Models, *Journal of Time Series Analysis*, 9, 4, 355-359

Ljung G.M., Box G.E.P. (1978) On a Measure of Lack of Fit in Time Series Models, *Biometrika*, 65, 2, 297-303

Lusk E.J., Wright H. (1982) Non-Gaussian Series and Series with Non-zero Means Practical Implications for Time Series Analysis, *Statistics & Probability Letters*, 1, 2-6

Lütkepohl H. (1982) Differencing Multiple Time Series: Another Look at Canadian Money and Income Data, *Journal of Time Series Analysis*, 3, 4, 235-243

Lütkepohl H. (1991) *Introduction to Multiple Time Series Analysis*, Springer-Verlag, Berlin

Lütkepohl H., Schneider W. (1989) Testing for Nonnormality of Autoregressive Time Series, *Computational Statistics Quarterly*, 2, 151-168

McClave J.T. (1978) Estimating the Order of Moving Average Models: the Max 2 Method, *Communications in Statistics, Theory and Methods*, A7(3), 259-276

McLeod A.I. (1978) On the Distribution of Residual Autocorrelations in Box-Jenkins Models (1978), *Journal of the Royal Statistical Society, B*, 40, 3, 296-302

McLeod A.I., Hipel K.W. e Lennox W.C. (1977) Advances in Box-Jenkins Modeling. 2.Applications, *Water Resources Research*, 13, 3, 577-586

McLeod A.I., Li W.K. (1983) Diagnostic Checking ARMA Time Series Models Using Squared-Residual Autocorrelations, *Journal of Time Series Analysis*, 4, 4, 269-273

Mallows C.L. (1967) Linear Processes are Nearly Gaussian, *Journal of Applied Probability*, 4, 313-329

Maravall A. (1983) An Application of Nonlinear Time Series Forecasting, *Journal of Business & Economic Statistics*, 1, 1, 66-74

Mardia K.V. (1980) Tests of Univariate and Multivariate Normality, in: Krishnaiah P.R. (editor), *Handbook of Statistics*, North-Holland P.Co., 1, 279-320

Masarotto G. (1987) Robust Identification Of Autoregressive Mov-

ing Average Models, *Applied Statistics*, 36, 2, 214-220

Masarotto G. (1988) Stimatori robusti ad un passo della funzione di autocorrelazione, *Atti XXXIV Riunione Scientifica SIS*, Siena, 2/1, 67-74

Mélard G., Roy R. (1984) Sur un Test d'égalité des autocovariances de deux séries chronologiques, *Canadian Journal of Statistics*, 12, 333-342

Mélard G., Roy R. (1987) On Confidence Intervals and Tests for Autocorrelations, *Computational Statistics and Data Analysis*, 5, 31-44

Milhoj A. (1981) A Test of Fit in Time Series Models, *Biometrika*, 68, 177-188

Monari P. (1993) Falsificazione e scelta tra alternative: il dilemma non risolto dei test di ipotesi, Commissione Scientifica della SIS, Convegno di Bologna, 27-28 maggio, *manoscritto*

Monfort A. (1993) Simulation Methods for Encompassing Tests and Indirect Inference, Commissione Scientifica della SIS, Convegno di Bologna, 27-28 maggio, *manoscritto*

Monti A.C. (1994) A Proposal for a Residual Autocorrelations Test in Linear Models, *Biometrika*, 81, 4, 776-780

Mazzali A. (1982), Gli spazi di previsione dei modelli ARMA, *Rivista di Statistica Applicata*, 15, 4, 295-308

Nagelkerke N.J.D. (1991) A Note on the General Definition of the Coefficient of Determination, *Biometrika*, 78, 3, 691-692

Nelson C.R. (1976) The Interpretation of R^2 in Autoregressive-Moving Average Time Series Models, *The American Statistician*, 30, 4, 175-180

Nelson C.R., Kang H. (1981) Spurious Periodicity in Inappropriately Detrended Time Series, *Econometrica*, 49, 3, 741-751

Newbold P. (1980) The equivalence of two tests of time series model adequacy, *Biometrika*, 67, 2, 463-465

Newton H.J., Pagano, M. (1984) Simultaneous Confidence Bands for Autoregressive Spectra, *Biometrika*, 71, 1, 197-202

Nicholls D.F., Quinn B.G. (1982) *Random Coefficient Autoregressive Models: an Introduction*, Springer-Verlag, New York

- Oja H. (1983) New test for Normality, *Biometrika*, 70, 1, 297-299
- Orsi R. (1990) Test per la specificazione di modelli econometrici, *Atti XXXV Riunione Scientifica SIS*, Padova, I, 137-148
- Orsi R. (1993) Sull'uso dei test in econometria, Commissione Scientifica della SIS, Convegno di Bologna, 27-28 maggio, *manoscritto*
- Parzen E. (1974) Some Recent Advances in Time Series Modeling, *IEEE Transactions on Automatic Control*, AC-19, 6, 723-730
- Parzen E. (1992) Time Series, Statistics, and Information, in: D.Brillinger et al. (editors), *New Directions in Time Series Analysis*, part I, Springer-Verlag, New York, 265-286
- Paulsen J., Tjøstheim D. (1985) On the Estimation of residual Variance and Order Determination in Autoregressive Time Series, *Journal of the Royal Statistical Society, B*, 47, 2, 216-228
- Peña D. (1984) Influential Observations in Time Series, Mathematics research Center, *Technical Report* 2718, University of Wisconsin-Madison
- Peña D. (1984) Measuring Importance of Outliers in ARIMA Models, in: M.L.Puri et al. (editors) *New Perspectives in Theoretical and Applied Statistics*, J.Wiley & Sons, New York
- Peña D. (1989) Influential Observations in Time Series, *manuscript*
- Pham D. T. (1987) Exact Maximum Likelihood Estimate and Lagrange Multiplier Test Statistic for ARMA Models, *Journal of Time Series Analysis*, 8, 1, 61-78
- Piccolo D. (1978 a) Efficienza Asintotica di Stimatori Alternativi del Periodo, *Metron*, XXXVI, 1/2, 155-171
- Piccolo D. (1978 b) Problemi Statistici nella Stima del Periodo, *Quaderni di Statistica e Econometria*, I, 92-122
- Piccolo D. (1978 c) La metodologia Box-Jenkins per l'analisi di fenomeni monetari e creditizi, *Rivista di Statistica Applicata*, 11, 3, 161-179
- Piccolo D. (1980) Un confronto fra alcuni criteri statistici di identificazione automatica, *Quaderni di Statistica e Econometria*, III, 117-167
- Piccolo D. (1981) Regioni osservabili per modelli AR e MA, *STA-*

STATISTICA, XLI, 1, 11-22

Piccolo D. (1982 a) The Size of the Stationarity and Invertibility Region of an Autoregressive-Moving Average Process, *Journal of Time Series Analysis*, 3, 4, 245-247

Piccolo D. (1982 b) Capacità previsiva dei modelli stocastici lineari, *Atti XXXI Riunione Scientifica SIS*, Torino, II, 480-503

Piccolo D. (1984 a) Analisi canonica e capacità previsiva per processi multivariati, *Rivista di Statistica Applicata*, 17, 1, 3-12

Piccolo D. (1984 b) Una topologia per la classe dei modelli ARIMA, *STATISTICA*, XLIV, 47-59

Piccolo D. (1990 a) *Introduzione all'Analisi delle Serie Storiche*, Nuova Italia Scientifica, Roma

Piccolo D. (1990 b) A Distance Measure for Classifying ARIMA Models, *Journal of Time Series Analysis*, 11, 153-164

Piccolo D. (1999) Seasonal Adjustment Research Appraisal (SARA): Final Statistical Recommendations of the Scientific Committee, *Atti dei Contributi scientifici Progetto SARA*, Annali di Statistica, ISTAT, Roma

Piccolo D. (2000) *Statistica*, Seconda edizione, Il Mulino, Bologna

Piccolo D., Tunncliffe Wilson G. (1984) A Unified Approach to ARMA Model Identification and Preliminary Estimation, *Journal of Time Series Analysis*, 5, 3, 183-204

Pierce D. A. (1985) Testing Normality in Autoregressive Models, *Biometrika*, 72, 293-297

Poskitt D.S., Tremayne A.R. (1980) Testing the Specification of a Fitted Autoregressive-Moving Average Model, *Biometrika*, 67, 2, 359-363

Poskitt D.S., Tremayne A.R. (1986) Some Aspects of the Performance of Diagnostic Checks in Bivariate Time Series Models, *Journal of Time Series Models*, 7, 3, 217-233

Pötscher B.M., Reschenhofer E. (1988) Discriminating between two Spectral Densities in case of Replicated Observations, *Journal of Time Series Analysis*, 9, 3, 221-224

Priestley M.B. (1980) State-Dependent Models: A General Approach to Nonlinear Time Series Analysis, *Journal of Time Series*

Analysis, 1, 47-71

Priestley M.B. (1981) *Spectral Analysis and Time Series*, voll. 1 e 2, Academic Press, London

Prothero D.L., Wallis K.F. (1976) Modelling Macroeconomic Time Series (with discussion), *Journal of the Royal Statistical Society, A*, 139, 468-500

Pukkila T., Koreisha S. e Kallinen A. (1990) The Identification of ARMA Models, *Biometrika*, 77, 3, 537-548

Pukkila T., Nyquist T. (1985) On the Frequency Domain Estimation of the Innovation Variance of a Stationary Univariate Time Series, *Biometrika*, 75, 2, 317-323

Quenouille M.H. (1947) A Large-sample Test for the Goodness of Fit of Autoregressive Schemes, *Journal of the Royal Statistical Society, A*, 110, 123-129

Quenouille M.H. (1949) Approximate Tests of Correlation in Time Series, *Journal of the Royal Statistical Society, B*, 11, 68-84

Quinn B.G. (1989) Estimating the Number of Terms in a Sinusoidal Regression, *Journal of Time Series Analysis*, 10, 1, 71-75

Quinn B.G., Thomson P.J. (1991) Estimating the Frequency of a Periodic Function, *Biometrika*, 78, 1, 65-74

Reschenhofer E. (1989) Adaptive Test for White Noise, *Biometrika*, 76, 3, 629-632

Reschenhofer E., Bomze I. M. (1991) Length Tests for Goodness-of-fit, *Biometrika*, 78, 629-632

Reschenhofer E., Bomze I.M. (1992) Testing for White Noise against Multimodal Spectral Alternatives, *Journal of Time Series Analysis*, 13, 5, 435-439

Royston J.P. (1982) An Extension of Shapiro and Wilk's W test for Normality to Large Samples, *Applied Statistics*, 31, 2, 115-124

Royston J.P. (1982) The W Test for Normality, Algorithm AS 181, *Applied Statistics*, 31, 3, 176-180

Rothenberg T.J. (1984) Approximating the Distributions of Econometric Estimators and Test Statistics in: Z.Griliches and M.D.Intriligator (editors), *Handbook of Econometrics*, II, Elsevier Science Publ. BV, 775-826

Saikkonen P. (1983) Asymptotic Relative Efficiency of Some Tests of Fit in Time Series Models, *Journal of Time Series Analysis*, 4, 1, 69-78

Schwarz G. (1978) Estimating the Dimension of a Model, *The Annals of Statistics*, 6, 461-464

Shapiro S.S., Francia R.S. (1972) An Approximate Analysis of Variance Test for Normality, *Journal of the American Statistical Association*, 67, 215-216

Shapiro S.S., Wilk M.B. (1965) An Analysis of Variance Test for Normality (complete samples), *Biometrika*, 52, 3 and 4, 591-611

Shapiro S.S., Wilk M.B. e Chen H.J. (1968) A Comparative Study of Various Tests for Normality, *Journal of the American Statistical Association*, 63, 1343-1372

Shumway R.H. (1982) Discriminant Analysis for Time Series, in: *Handbook of Statistics*, 2, (P.R.Krishnaiah and L.N.Kanal, editors), North-Holland, Amsterdam, 1-46

Shumway R.H., Unger A.N. (1974) Linear Discriminant Functions for Stationary Time Series, *Journal of the American Statistical Association*, 65, 1527-1546

Siegel A.F. (1980) Testing for Periodicity in a Time Series, *Journal of the American Statistical Association*, 75, 370, 345-348

Spanos A. (1986) *Statistical Foundations of Econometric Modelling*, Cambridge University Press, Cambridge

Srivastava M.S., Hui T.K. (1987) On Assessing Multivariate Normality Based on Shapiro-Wilk W Statistics, *Statistics and Probability Letters*, 5, 15-18

Subba Rao T. (1981) On the Theory of Bilinear Time Series Models, *Journal of the Royal Statistical Society, B*, 43, 244-255

Tomasek L. (1987) Asymptotic Simultaneous Confidence Bands for Autoregressive Spectral Density, *Journal of Time Series Analysis*, 8, 4, 469-477

Tong H. (1990) *Nonlinear Time Series Analysis: a Dynamical System Approach*, Clarendon Press, Oxford

Tsay R.S. (1986) Nonlinearity Tests for Time Series, *Biometrika*, 73, 461-466

- Tsay R.S. (1988) Outliers, Level Shifts and Variance Changes in Time Series, *Journal of Forecasting*, 7, 1-20
- Tsay R.S. (1989) Testing and Modeling Nonlinearity in Univariate Time Series Analysis, *Journal of the American Statistical Association*, 84, 231-240
- Tsay R.S. (1992) Model Checking via Parametric Bootstraps in Time Series Analysis, *Applied Statistics*, 41, 1, 1-15
- Tsay R.S., Tiao G.C. (1984) Consistent Estimates of Autoregressive Parameters and Extended Sample Autocorrelation Function for Stationary and Nonstationary ARMA Models, *Journal of the American Statistical Association*, 79, 84-96
- Tuan P.D. (1984) A Note on Some Statistics Useful in Identifying the Order of Autoregressive Moving Average Model, *Journal of Time Series Analysis*, 5, 4, 273-279
- Tuan P.D. (1986) A Frequency Domain Approach to Lagrange Multiplier Test for Autoregressive Moving Average Models, *Journal of Time Series Analysis*, 7, 1, 73-78
- Verrill S., Johnson R.A. (1987) The Asymptotic Equivalence of Some Modified Shapiro-Wilk Statistics - complete and censored cases, *The Annals of Statistics*, 15, 1, 413-419
- Vitale C. (1993) *Non linearità, caos e caso*, Centro di Specializzazione e Ricerche, Portici
- Yu G.H., Lin Y.C. (1991) Methodology for Selecting Subset Autoregressive Time Series Models, *Journal of Time Series Analysis*, 12, 4, 363-373
- Wallis K.F. (1987) Time Series Analysis of Bounded Economic Variables, *Journal of Time Series Analysis*, 8, 1, 115-123
- Weiss G. (1975) Time Reversibility of Linear Stochastic Processes, *Journal of Applied Probability*, 12, 831-836
- Whittle P. (1952) Test of Fit in Time Series, *Biometrika*, 39, 309-318
- Whittle P. (1953) Estimation and Information in Stationary Time Series, *Arkiv fur Matematik*, 2, 423-434
- Whittle P. (1963) *Prediction and Regulation by Linear Least-Squares Methods* (2nd revised edition, 1984), Basil Blackwell, Oxford

Wold H. (1954) *A Study in the Analysis of Stationary Time Series* (1st edition, 1938), Almqvist & Wicksell, Uppsala

Zani S., editor (1993), *Metodi Statistici per le Analisi Territoriali*, F. Angeli, Milano

Zhang H.C. (1992) Reduction of the Asymptotic Bias of Autoregressive and Spectral Estimators by Tapering, *Journal of Time Series Analysis*, 13, 5, 451-469

Zhao-Guo C. (1988) An Alternative Consistent Procedure for Detecting Hidden Frequencies, *Journal of Time Series Analysis*, 9, 3, 301-317