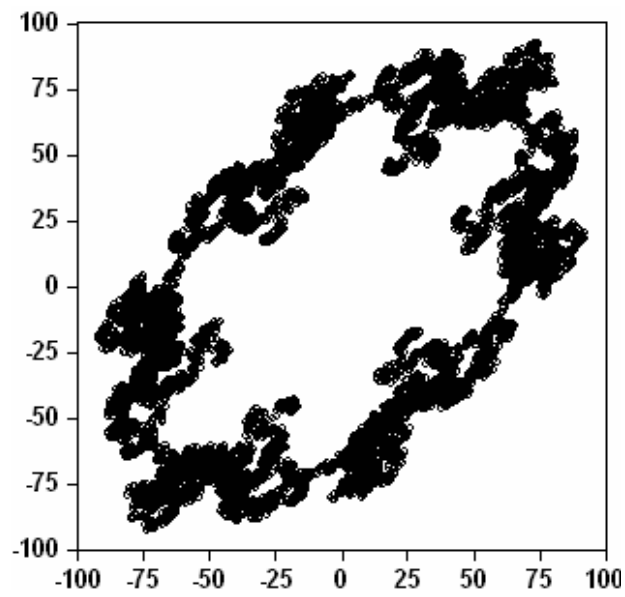


GIUSEPPE STORTI – COSIMO D. VITALE

ANALISI STATISTICA DEI MERCATI MONETARI E
FINANZIARI
ANALISI UNIVARIATA



*Nell'acqua del fiume
si può entrare una sola volta.
Panta rei.
Eraclito*

Indice

CAPITOLO I

Cenni sui mercati monetari e finanziari

1.1 Introduzione	1
1.2 I mercati finanziari e monetari più rilevanti	3
1.3 Alcuni strumenti finanziari	5

CAPITOLO II

Analisi esplorativa descrittiva dei mercati monetari e finanziari

2.1 I prezzi degli attivi	9
2.2 I rendimenti degli attivi	27
2.3 La volatilità degli attivi	53
2.4 Il Valore a Rischio (VaR) degli attivi	60
2.5 Una applicazione a dati reali	68

CAPITOLO III

Variabili casuali e processi stocastici

3.1 Variabili casuali univariate	85
(A) Variabile casuale di Bernoulli	85
(B) Variabile casuale Normale	87
(C) Variabile casuale Chi quadrato	92
(D) Variabile casuale T di Student	94
(E) Variabile casuale F di Fisher	97
(F) Variabile casuale log Normale	99
3.2 Variabili casuali doppie	101
(A) Variabile casuale di Bernoulli doppia	103
(B) Variabile casuale Normale doppia	105
3.3 Alcuni risultati limite	108
3.4 Cenni alle stime di quasi massima verosimiglianza	116
3.5 Alcuni test delle ipotesi basati sulla funzione di verosimiglianza	121
3.6 Serie storiche e processi stocastici	130
3.7 Alcuni particolari processi stocastici	134
(A) Processo indipendente ed identicamente distribuito	135

(B) Processo white noise	135
(C) Medie progressive di processi iid	136
(D) Processo di Wiener o moto Browniano	140
(E) Medie mobili di processi iid	142
(F) Processo random walk	143
CAPITOLO IV	
Analisi dei livelli di attivi monetari e finanziari	
4.1 Introduzione	137
4.2 Analisi dei rendimenti di attivi finanziari	139
4.3 Uso dei modelli ARMA per l'analisi dei livelli dei rendimenti	142
4.4 Il modello AR(1) per i rendimenti	150
4.5 Test per la presenza di radici unitarie	162
4.6 Il modello MA(1) per i rendimenti	167
4.7 Il modello ARMA(1,1) per i rendimenti	174
4.8 Il modello TARMA per i rendimenti	179
4.9 La scelta del modello ottimale	192
4.10 La previsione dei rendimenti	194
(A) Previsioni da un AR(1)	195
(B) Previsione da un MA(1)	200
(C) Previsione da un ARMA(1,1)	202
(D) Previsioni da un TAR(2;1)	205
(E) Previsioni da un TMA(2;1)	208
CAPITOLO V	
Analisi della volatilità di attivi monetari e finanziari	
5.1 Il modello ARCH classico	219
5.2 Il modello ARCH con struttura nei livelli	227
5.3 Identificazione e stima di un ARCH(s)	230
5.4 Il modello GARCH	240
5.5 Varianti del modello GARCH	245
(A) Il modello IGARCH	245
(B) Il modello GARCH-M	247
(C) Il modello CGARCH	249
5.6 La previsione della volatilità	250
5.7 L'effetto leverage nella volatilità	254
(A) Il modello EGARCH	255
(B) Il Modello TGARCH	259
(C) Il modello PGARCH	265
5.8 Stime QMV per modelli ARMA-GARCH	269

CAPITOLO VI	
La valutazione del rischio nei mercati monetari e finanziari	
6.1 Introduzione	275
6.2 Il Valore a Rischio	275
6.3 Il VaR di Riskmetrics	282
6.4 Il VaR per modelli ARMA-GARCH	287
(A) Il modello MA(0)-GARCH(1,1)	288
(B) Il modello MA(1)-GARCH(1,1)	290
(C) Il modello MA(q)-GARCH(1,1)	290
(D) Il modello AR(1)-GARCH(1,1)	293
6.5 Un intervallo di confidenza per il VaR	294
6.6 Stima del percentile $a_{(\alpha)}$ e calcolo del VaR	296
(A) Il VaR con residui normali	296
(B) Il VaR con residui t di Student	297
(C) Il VaR e gli ordinamenti statistici	301
6.7 Il VaR nei modelli TMA-TGARCH	303
6.8 La stima del VaR di medio lungo periodo ($h>1$)	316
6.9 L'Expected Shortfall	324
(A) Misure di rischio coerenti ed Expected Shortfall	324
(B) La stima dell'ES per modelli ARMA-GARCH	327
(C) La stima dell'ES per $h>1$	331
6.10 Valutazione della qualità delle stime VaR: il <i>backtesting</i>	334
(A) Il backtesting	334
(B) Tests di copertura	336
<i>Tavole statistiche</i>	353
<i>Bibliografia</i>	365
<i>Indice analitico</i>	369

CAPITOLO I

Cenni sui mercati monetari e finanziari

1.1 Introduzione

Scopo di questo libro è presentare alcuni semplici strumenti probabilistico – statistici che permettano all'utilizzatore una immediata analisi ed interpretazione dell'evoluzione temporale di *prezzi, rendimenti e volatilità di attivi finanziari e monetari*. In questo volume le analisi riportate sono solo univariate (analisi di singoli attivi), estensioni al multivariato (analisi di più attivi o di interi portafogli) sono rinviate a futuri lavori.

I prezzi degli attivi sono determinati in particolari mercati dettagliatamente regolamentati. In questo volume non tratteremo degli aspetti giuridico – normativi, né di come tecnicamente i prezzi dei diversi attivi si formano su quei mercati, ma solo di particolari analisi empirico – statistiche atte a quantificarne il *rischio di mercato*. Nel volume, utilizzando *serie storiche* osservate, un approccio essenzialmente empirico e l'uso di strumenti propri del Calcolo delle Probabilità e della Statistica, ne analizzeremo l'evoluzione temporale cercando di cogliere le eventuali regolarità empiriche in esse presenti e prevederne gli andamenti futuri. A questo scopo verranno analizzati, stimati e verificati appositi modelli che ci permetteranno, fra l'altro, di costruire *previsioni "ottimali"* di breve o brevissimo periodo. Inoltre verranno presentati strumenti matematico – statistici che permetteranno di misurare il *rischio di mercato*. Gli strumenti che elaboreremo saranno applicati a serie di prezzi di attivi effettivamente osservati sui mercati sia italiani che internazionali.

Più in particolare, nel volume verranno studiate, con dettaglio, le quotazioni di *azioni, obbligazioni, monete* rilevati con *cadenza* giornaliera o superiore (settimanale, mensile). Non verranno trattate serie con cadenze infra-giornaliera, che necessitano di strumenti particolari riconducibili ai processi stocastici a tempo continuo e che esulano dagli scopi di questo lavoro.

Le analisi statistiche verranno effettuate tramite software largamente noti ed utilizzati ed in particolare *Excel*, ormai presente su tutti i computer, ed *EViews 6* (ma possono essere usate anche le versioni precedenti di questo

software). Naturalmente non è per nulla escluso che queste analisi possano essere fatte con altri pacchetti applicativi.

I mercati monetari e finanziari hanno assunto, negli ultimi anni, un peso ed un ruolo sempre più rilevante nelle economie di tutti i paesi ed in particolare in quelli più sviluppati. Sono mercati fortemente interconnessi fra di loro e con l'economia reale. La grave crisi finanziaria del 2007-2010 è nata negli USA, collegata ai mutui del mercato immobiliare, si è diffusa rapidamente agli altri paesi, passando in fine all'economia reale con conseguente crollo della produzione e della domanda di beni e servizi, fallimento di banche e di imprese industriali e commerciali, forte frenata del commercio internazionale ed impennata dei tassi di disoccupazione. Per tamponare la crisi, normalizzare i mercati ed invertire il ciclo negativo, tutte le istituzioni finanziarie centrali ed i governi dei diversi paesi sono dovuti intervenire con normative specifiche in favore delle banche, delle altre strutture finanziarie, delle imprese e dei consumatori immettendo nel sistema flussi monetari molto elevati. Tutto questo è un chiaro indicatore di come i mercati finanziari sino fortemente collegati all'economia reale e la condizionino sistematicamente. Ma questo implica anche che tali mercati necessitano di una nuova e più penetrante normativa atta a contenerne gli eccessi ed evitare che nel futuro si generino crisi di questa profondità e gravità. Le semplici constatazioni sopra riportate dimostrano come la finanza sia interconnessa all'economia reale e sia cruciale per lo sviluppo ed il funzionamento delle moderne economie.

Gli strumenti finanziari hanno il ruolo fondamentale sopra accennato e sono pertanto largamente utilizzati nelle attività economiche, perché permettono di trasferire ricchezza, nel tempo e nello spazio, e frantumare i rischi di impresa distribuendoli su platee le più larghe possibili che possono coinvolgere sia le generazioni presenti che quelle future. Di qui la facilità di diffusione delle crisi finanziarie, che in un mondo fortemente globalizzato difficilmente possono essere ristrette ad un solo paese, e la necessità di regolamentare questi mercati sia a livello locale che internazionale.

Come già accennato i mercati finanziari, per la loro complessità, possono essere studiati ed analizzati da differenti punti di vista: giuridico – normativo, economico – sociale, statistico – matematico, ecc.. In questo volume l'approccio di analisi è esclusivamente di tipo statistico – matematico e lo sforzo è quello di derivare modelli che siano facilmente interpretabili ed analizzabili, ma allo stesso tempo siano in grado di catturare le particolari regolarità empiriche che posseggono singoli attivi, portafogli di attivi o interi mercati. Le conoscenze così ottenute, assieme a quelle di tipo giuridico – normativo, economico – sociale, tecnico – finanziario permetteranno agli utilizzatori (investitori, operatori finanziari, responsabili della politica economica, ecc.) di avere visioni ed informazioni le più articolate possibili che

permettano loro di operare, nei diversi campi di interesse con maggiore consapevolezza ed accuratezza. In altri termini, gli strumenti quantitativi che presenteremo possono far sì che operatori più informati assumano decisioni più rapide e congrue rispetto a chi è privo di dette conoscenze.

Per la loro complessità questi mercati possono essere assimilati a sistemi dinamici che evolvono nel tempo, molto sensibili alle condizioni iniziali del sistema incognite agli operatori. Questo implica, fra l'altro, che sono praticamente non prevedibili se non per particolari aspetti e per orizzonti di breve e brevissimo periodo.

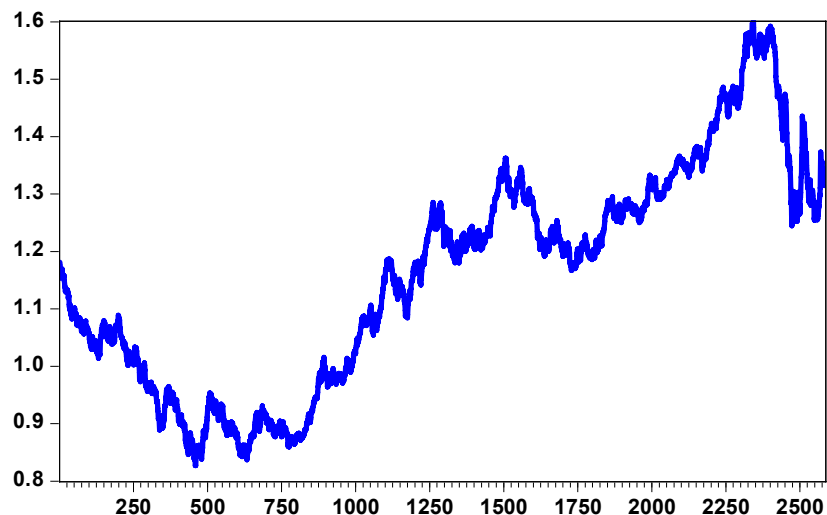
1.2 I mercati finanziari e monetari più rilevanti

Qui di seguito riportiamo alcuni rapidissimi cenni sui mercati finanziari più importanti indicando anche quale è il relativo rischio di mercato. In generale, maggiore è la rischiosità dello strumento finanziario, maggiore potrebbe essere il profitto (o la perdita) per il detentore di quello strumento. In altri termini, chi opera in questi mercati investe capitali presenti o futuri nella speranza che un evento rischioso si verifichi (o non si verifichi) ricavando profitti ed evitando di subire perdite.

(A) IL MERCATO DEI CAMBI: FOREIGN EXCHANGE (FX) MARKET

Su questo mercato vengono trattati i prezzi delle monete dei diversi paesi o gruppi di paesi (come quelli aderenti all'Euro).

Figura 1 – *Cambio Euro/Dollaro USA, 1999 – 2010*



La contrattazione e la formazione del prezzo di una moneta rispetto ad una di riferimento (di solito il dollaro) è continua. Le analisi che qui presenteremo si riferiscono a quelle di chiusura giornaliera. Il rischio di mercato è quello di cambio: compro una moneta ad un dato prezzo e la rivendo o mi impegno di rivenderla successivamente ad un prezzo differente. La differenza di prezzo può essere positiva o negativa: la *speranza* dell'operatore è che sia positiva. Il rischio di cambio può essere rilevato e misurato a partire da una serie temporale osservata su una data moneta. Nella figura 1 è riportato il cambio di chiusura giornaliero Dollaro USA/Euro dal 4 gennaio 1999 all'11 giugno 2010.

(B) IL MERCATO DELLA MONETA: MONEY MARKET (MM)

Su questo mercato operano le banche che prendono in prestito o depositano moneta. Il prezzo del prestito o del deposito è rappresentato dal tasso di interesse annuale quotato per scadenze con diversa frequenza: overnight, una settimana, due settimane, fino ad un massimo di 12 mesi. Il danaro prestato viene restituito (o incassato) alla scadenza del contratto assieme agli interessi maturati. I tassi di riferimento a breve sono quelli *LIBOR* (*London Interbank Offered Rate*) riferiti a prestiti nelle più importanti divise mondiali (Dollaro USA, Euro, Yen Giapponese, Lira Sterlina, ecc.). Il rischio è legato al variare dei tassi di cambio.

(C) INTEREST RATE SWAP (IRS) MARKET

Consiste in un accordo fra due parti che si obbligano ad eseguire una sequenza di pagamenti in date future pre-specificate. Una delle due parti corrisponderà pagamenti fissi stabiliti nel contratto. L'altra eseguirà pagamenti con ammontare incognito al momento della stipula del contratto dato che dipende dai tassi di interesse che si formeranno nel futuro.

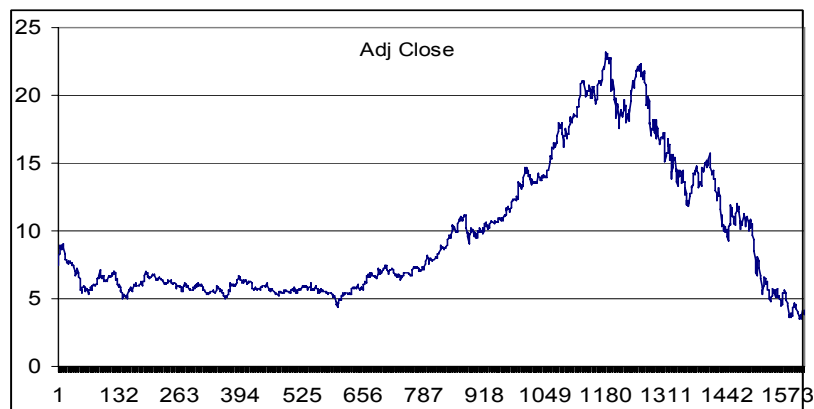
La logica di questi contratti è la stessa di quelle delle assicurazioni contro un evento avverso in cui una parte paga una somma fissa (il premio) e l'altra una variabile legata al verificarsi dell'evento. A differenza delle assicurazioni nei contratti swap le parti si scambiano danaro, una ad interesse fisso e l'altra ad interesse variabile. Si tratta, pertanto, di una assicurazione contro le variazioni dei tassi di interesse ed il rischio di questo mercato è collegato alla variazione dei tassi: se sono superiori o inferiori a quello fisso una parte consegue un vantaggio e l'altra uno svantaggio.

(D) EQUITY MARKET

In questo mercato si formano i prezzi dei titoli azionari. Anche in questo caso il flusso dei prezzi delle azioni è continuo (*tick by tick*) durante i giorni di apertura dei mercati borsistici. A fine giornata viene pubblicato un

prezzo di chiusura. Nel seguito i dati che analizzeremo si riferiranno ai *prezzi di chiusura aggiustati* per tener conto dei pagamenti di eventuali dividendi o di trasformazioni dell'assetto azionario (*splitting*, per esempio). Il rischio di questi mercati è legato al variare della quotazione delle azioni nel tempo. Se posseggo azioni ed i loro prezzi flettono subisco una perdita, se salgono realizzo un profitto. Nella figura 2 sono riportate le quotazioni giornaliere di chiusura aggiustate per il titolo Fiat ordinario osservate dal gennaio 2003 al marzo 2009.

Figura 2 – *Quotazioni giornaliere di chiusura aggiustate Fiat*



1.3 *Alcuni strumenti finanziari*

In questo paragrafo riportiamo alcuni dei più importanti strumenti finanziari. La presentazione sarà sintetica e serve solo per delineare il contesto in cui verranno presentate nel seguito le analisi statistiche – econometriche proposte. Approfondimenti e dettagli su questi strumenti esulano dagli scopi di questo lavoro.

(A) LE AZIONI

Le azioni sono dei documenti che attestano la proprietà di una quota di capitale sociale di società per azioni (S.p.A). I prezzi di questi attivi, se quotati su mercati ufficiali, si formano nelle Borse Valori. Il prezzo di ogni azione viene aggiornato al verificarsi di ogni nuova contrattazione e viene reso disponibile al pubblico in tempo reale. I dati raccolti sotto forma di serie storiche costituiscono delle Banche Dati che possono essere, spesso a pagamento, consultate ed analizzate. Nel seguito analizzeremo dati con cadenza giornaliera che sono, oltre che pubblici, gratuiti.

Vi sono diversi siti web da cui possono essere liberamente scaricati dati finanziari sia nazionali che internazionali. Per citarne alcuni, ricordiamo:

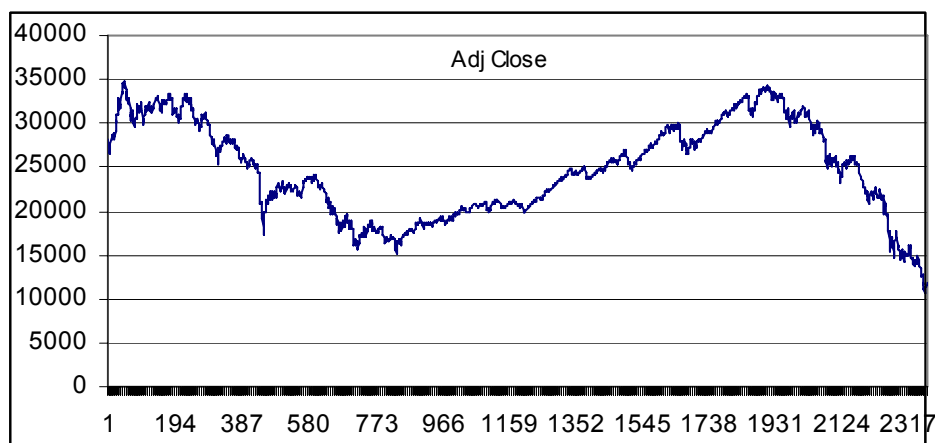
<http://finance.yahoo.com>

<http://it.finance.yahoo.com>

<http://research.stlouisfed.org/fred2>

Oltre alle quotazioni delle singole azioni delle società quotate, vengono costruiti indici sintetici borsistici che permettono di valutare l'andamento complessivo di un dato mercato. Anche questi indici sono disponibili gratuitamente con cadenza almeno giornaliera. Per l'Italia i più noti indici sintetici sono *MIB30*, *MIB*, *MIBTEL*, per il mercato USA sono le diverse varianti degli indici *DOW JONES* (DJ), *NASDAQ*, *Standard & Poors* (S&P), per la Francia il *CAC*, per la Germania il *DAX*. Gli indici sintetici, con l'eccezione del DJ che è una media semplice, sono medie ponderate corrette di un paniere rappresentativo dei prezzi dei titoli trattati in una determinata piazza borsistica. Nella figura sono riportate le quotazioni giornaliere di chiusura aggiustate dell'indice *MIBTEL* per il periodo gennaio 2000, marzo 2009.

Figura 3 – *Indice MIBTEL giornaliero dal gennaio 2000 al marzo 2009*

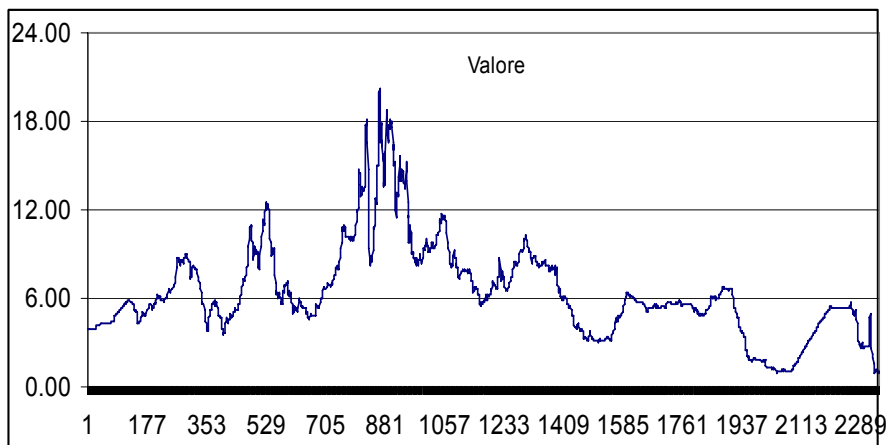


(B) LE OBBLIGAZIONI

Al contrario delle azioni, le obbligazioni rappresentano un titolo di credito nei confronti delle organizzazioni che le hanno emesse. Queste organizzazioni possono essere sia private (Società, Banche, ecc) che

pubbliche (Stati, Regioni, Comuni, ...). L'ente che emette le obbligazioni si impegna a rimborsare il capitale ad una determinata scadenza ed a pagare, eventualmente, cedole periodiche di cui è definito il metodo di calcolo dell'importo al momento dell'emissione. Le cedole possono essere calcolate sia con un tasso fisso che con uno variabile calcolato in base ad un metodo prestabilito. Nella figura 4 è riportato il valore settimanale dei Certificati di Deposito a 3 mesi della Federal Reserve USA.

Figura 4 – *Valore settimanale dei Certificati di Deposito a 3 mesi della Federal Reserve USA*



(C) I PRODOTTI DERIVATI

I prodotti derivati sono strumenti finanziari il cui prezzo è legato a quello di altri strumenti (titoli azionari, indici azionari, valute, tassi di interesse, ecc.) detto *sottostante*. Si tratta di contratti a termine che stabiliscono la compravendita futura di una attività finanziaria a predefinite condizioni. Di questi derivati i contratti più noti ed utilizzati sono i futures e le options.

- *Futures*

Con questi contratti le parti si impegnano a scambiarsi, ad una data prefissata e a un determinato prezzo, un ammontare di attivi finanziari o di un bene reale. Vi è per tanto l'impegno di vendere o acquistare una qualche attività finanziaria o bene reale ad un prezzo prefissato. Il rischio, per esempio per l'acquirente, è che alla scadenza il prezzo pattuito sia superiore a quello corrente e quindi di incorrere in una perdita. Il viceversa si ha per il venditore.

- *Options*

A differenza dei futures, nelle options si acquista o si vende, dietro il pagamento di un prezzo, il *diritto* ad acquistare o vendere attività finanziarie. Questo diritto può essere esercitato oppure può essere lasciato cadere.

Questi ultimi tipi di contratti sono simili a quelli che si stipulano quotidianamente quando si dà un anticipo (*caparra*) per acquistare un bene o un servizio il cui prezzo è pattuito alla stipula del contratto. Alla scadenza del contratto l'acquirente può decidere se perdere l'anticipo o comprare il bene pagando il residuo fra anticipo e prezzo. L'anticipo rappresenta il prezzo dell'opzione e quindi il relativo rischio. Esistono diversi tipi di options, quelle più utilizzate sono quella americana e quella europea.

- *Option americana*

Il compratore può esercitare il diritto ad acquistare in un momento qualsiasi dell'intervallo che va dalla stipula del contratto alla sua scadenza prefissata.

- *Option europea*

Il compratore può esercitare il diritto ad acquistare solo alla scadenza del contratto.

In ogni caso il valore dell'opzione è determinato dal prezzo del *sottostante*, sia esso un bene reale o una attività finanziaria.

I detentori di contratti derivati, così come avviene per qualsiasi bene o attività finanziaria, possono cederli ad altri soggetti prima che il diritto o l'obbligo stabilito venga esercitato.

CAPITOLO II

Analisi esplorativa descrittiva dei mercati monetari e finanziari

2.1 I prezzi degli attivi

Nel seguito indicheremo con X_t la rilevazione di una determinata caratteristica quantitativa di un attivo X al tempo t . La serie analizzata può essere il *prezzo* (P), il *rendimento* (r), il *tasso di interesse* (i), la *variabilità* (v), mentre t rappresenta il tempo in cui la rilevazione è stata effettuata. Nel nostro caso t rappresenta il giorno (la settimana, il mese, ecc.) a cui la rilevazione si riferisce. Inoltre, per semplificare la notazione, una volta indicato l'intervallo di tempo a cui la rilevazione si riferisce e la relativa cadenza, il contatore t assumerà i primi N (tante quante sono le rilevazioni effettuate) valori interi: $t = 1, 2, \dots, N$. Per esempio, se la rilevazione si riferisce ai prezzi giornalieri dell'attivo considerato ed è stata effettuata dal 16/02/09 al 6/03/09, P_1 rappresenterà il prezzo del 16/02/09, P_2 il prezzo del 17/02/09, ..., P_{15} il prezzo del 06/03/09. A titolo di esempio nella Tabella 1 sono riportate le $N = 15$ quotazioni giornaliere di chiusura aggiustate del titolo Campari osservate dal 16/02/09 al 06/03/09 sulla Borsa di Milano.

Tabella 1 – *Quotazioni giornaliere di chiusura aggiustate del titolo Campari*

Data	P_t	t	Data	P_t	t	Data	P_t	t
16/02/09	4.26	1	23/02/09	4.12	6	02/03/09	4.09	11
17/02/09	4.15	2	24/02/09	4.15	7	03/03/09	3.92	12
18/02/09	4.05	3	25/02/09	4.22	8	04/03/09	4.02	13
19/02/09	4.20	4	26/02/09	4.14	9	05/03/09	4.04	14
20/02/09	4.19	5	27/02/09	4.30	10	06/03/09	3.88	15

Come si può notare dalla tabella, le date sono discontinue visto che il sabato e la domenica le Borse sono chiuse e non esiste la relativa quotazione, mentre il contatore t non tiene conto di questa discontinuità. Nelle quotazioni giornaliere le Borse sono chiuse anche nei giorni di festa, diversi dal sabato e la

domenica. Per far sì che la serie osservata abbia sempre 5 osservazioni a settimana, in corrispondenza di questi giorni di festa si ripete l'ultima quotazione disponibile.

I dati giornalieri possono essere aggregati per ottenere quelli settimanali, mensili, ecc. Le quotazioni di attivi sono delle *variabili di stock* e per le aggregazioni settimanali è necessario considerare le quotazioni giornaliere riferite all'ultimo giorno della settimana (il venerdì). Questo vuol dire che la serie settimanale delle quotazioni di chiusura aggiustate è data dalle quotazioni di chiusura aggiustate di tutti i venerdì. In simboli, se indichiamo con $P_t^{(5)}$ il prezzo di chiusura del venerdì e con P_t quello giornaliero, sarà $P_t^{(5)} = P_{5t}$. Nella tabella 2 sono riportate le quotazioni di chiusura aggiustate del titolo Campari riferite allo stesso periodo di osservazione della tabella 1, ma con cadenza settimanale.

Tabella 2 – *Quotazioni settimanali del titolo Campari*

Data	$P_t^{(5)}$	t
20/02/09	4.19	1
27/02/09	4.30	2
06/03/09	3.88	3

Invece che considerare come valore rappresentativo della settimana quello riferito all'ultimo giorno, si potrebbe considerare la media delle quotazioni in ciascuna settimana. In tal modo si otterrebbero le *quotazioni medie* settimanali riportate nella tabella 3

Tabella 3 – *Quotazioni medie settimanali del titolo Campari*

Data	$\bar{P}_t^{(5)}$	t
20/02/09	4.170	1
27/02/09	4.186	2
06/03/09	3.990	3

dove si è posto:

$$\bar{P}_t^{(5)} = \frac{1}{5} \left(P_{5t} + P_{5t-1} + P_{5t-2} + P_{5t-3} + P_{5t-4} \right)$$

Come si può notare, i dati della tabella 2 sono effettivamente osservati ogni venerdì di ciascuna settimana, mentre quelli della tabella 3 sono una

manipolazione (in questo caso una media) dei dati osservati nella tabella 1. Questo è dovuto, come detto, al fatto che le quotazioni sono delle *variabili di stock* e quindi riferite ad un determinato istante temporale. I dati della tabella 2 infatti ci informano di come si è chiusa la quotazione del titolo Campari nelle tre settimane di osservazione. Al contrario i dati della tabella 3 ci informano su come, in media, il titolo Campari si è comportato in ciascuna delle tre settimane prese in considerazione.

Nel seguito, per le quotazioni, le analisi saranno effettuate su dati simili a quelli riportati nella tabella 1, per dati giornalieri, e nella tabella 2 per quelli settimanali.

Per le quotazioni giornaliere dei titoli azionari, ma anche per quelle dei cambi, per ciascun giorno di contrattazione vengono forniti:

- la data della quotazione (*Data*)
- la quotazione di apertura del mercato (*Open*)
- la quotazione massima del giorno (*High*)
- la quotazione minima del giorno (*Low*)
- la quotazione di chiusura del giorno (*Close*)
- il numero di titoli trattati (*Volume*)
- la quotazione di chiusura aggiustata (*Adj Close*).

Nel seguito, se non diversamente indicato, le analisi verranno effettuate sulla serie delle quotazioni *Adj Close* che indicheremo semplicemente con P_t . Ricordiamo che i dati delle tabelle 1 – 3 sono riferiti alle quotazioni *Adj Close*. Come accennato, l'aggiustamento delle quotazioni di chiusura è necessario ogni volta che viene pagato un dividendo, dato che nella valutazione dei prezzi degli attivi dei giorni precedenti è inclusa l'aspettativa sul dividendo. Allo stesso modo, l'aggiustamento si rende necessario quando avvengono operazioni di *splitting* o di *accorpamento*: le azioni in circolazione, a seguito di ristrutturazione del capitale sociale, vengono ritirate e sostituite da nuove azioni con un rapporto di conversione diverso da quello uno ad uno. L'operazione di aggiustamento viene effettuata per avere una serie storica omogenea di modo che sia possibile analizzarne l'evoluzione nel tempo.

Nella tabella 4 sono riportate le quotazioni giornaliere del titolo Campari rilevate dal 15/10/07 al 30/10/07.

Tabella 4 – *Quotazioni di Campari dal 15/10/07 al 30/10/07*

Data	t	Open	High	Low	Close	Volume	Adj Close
15/10/07	1	7.46	7.54	7.38	7.43	772700	7.29
16/10/07	2	7.45	7.45	7.32	7.32	548800	7.19

17/10/07	3	7.39	7.57	7.32	7.49	1132800	7.35
18/10/07	4	7.49	7.55	7.32	7.39	984200	7.25
19/10/07	5	7.34	7.47	7.28	7.34	381600	7.19
22/10/07	6	7.21	7.28	7.12	7.22	622400	7.08
23/10/07	7	7.33	7.36	7.16	7.22	872000	7.08
24/10/07	8	7.25	7.28	7.15	7.18	325500	7.05
25/10/07	9	7.16	7.34	7.16	7.24	666800	7.11
26/10/07	10	7.23	7.29	7.18	7.20	361700	7.06
29/10/07	11	7.17	7.29	7.17	7.26	507000	7.12
30/10/07	12	7.46	7.54	7.38	7.43	772700	7.29

Nella figura 1(a) sono riportati i grafici di Close e Adj Close ed in quella 1(b) le sei serie (open, high, low, close, adj close, volume) per le quotazioni giornaliere del titolo Campari sulla Borsa di Milano osservate dal 06/07/01 al 09/03/09.

Figura1(a) – *Quotazioni giornaliere Close e Adj Close per Campari dal 06/01/01 al 09/03/09*

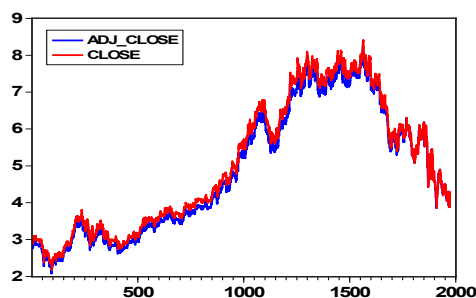
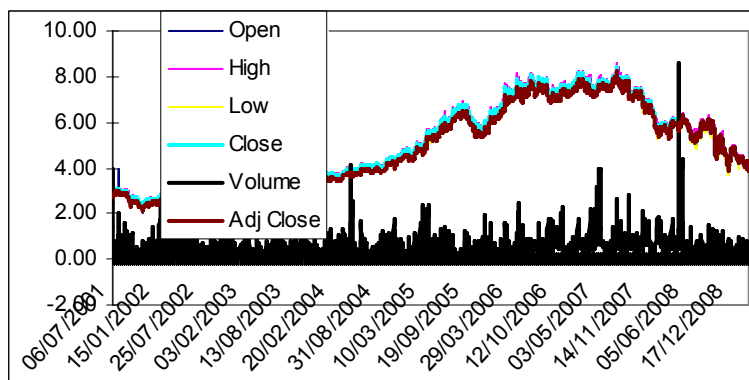


Figura 1(b) – *Quotazioni giornaliere Open, High, Low, Close, Volume in milioni di unità, Adj Close per il titolo Campari dal 06/01/01 al 09/03/09*



Nelle figure 2 sono riportati gli stessi dati con cadenza settimanale.

Figura 2(a) – *Quotazioni settimanali Close e Adj Close per Campari dal 06/07/01 al 09/03/09*

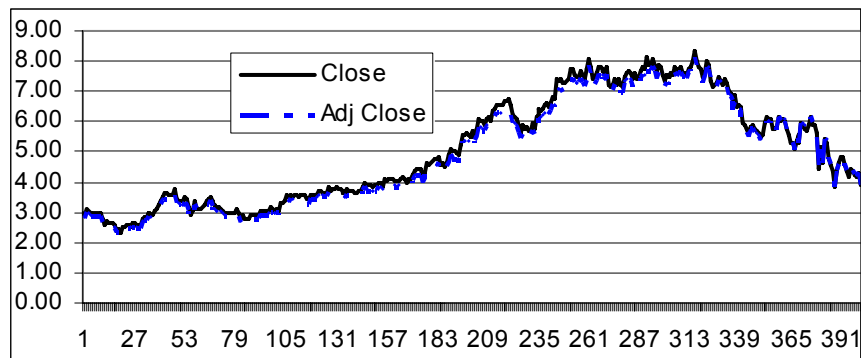
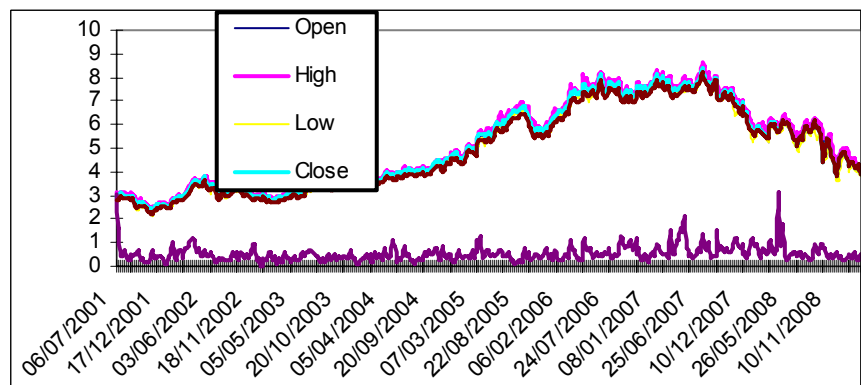


Figura 2(b) – *Quotazioni settimanali Open, High, Low, Close, Volume in milioni di unità, Adj Close per il titolo Campari dal 06/01/01 al 09/03/09*



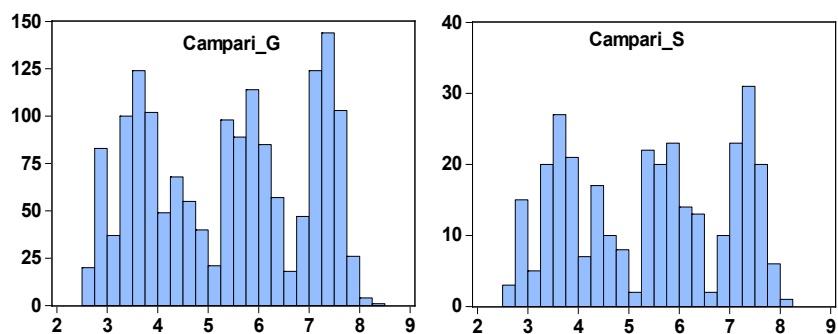
Come possiamo osservare, l'andamento del fenomeno osservato con cadenza giornaliera è praticamente lo stesso di quello osservato con cadenza settimanale. In altri termini il fenomeno è *auto simile* (una parte è simile al tutto). Questo fenomeno è tipico delle serie dei prezzi di attivi finanziari e non solo di quello azionario.

Della serie osservata dei prezzi è possibile costruire l'*istogramma di frequenze* che fornisce indicazioni grafiche sulla forma della distribuzione ed in particolare l'esistenza di massimi, di valori eccezionali, di asimmetria rispetto al baricentro (la media) della rilevazione. Per costruire l'istogramma si ordina la serie osservata in senso non decrescente, partendo dal valore più piccolo che

indichiamo con $P_{(1)}$, fino a quello più grande indicato con $P_{(N)}$, si calcola l'intervallo di variazione dato da $P_{(N)} - P_{(1)}$, si divide questo intervallo in un numero dispari di sotto intervalli (di solito 21 – 25) tutti della stessa ampiezza, si conta il numero delle quotazioni n_i che cadono nell'intervallo $[a_i; a_{i+1})$, si costruiscono tanti rettangoli affiancati fra di loro con base l'ampiezza di ciascun intervallo $a_{i+1} - a_i$ ed altezza pari ad n_i . Il grafico così ottenuto prende il nome di *istogramma di frequenza*.

Nella figura che segue è riportato l'istogramma delle quotazioni Campari, sia giornaliere che settimanali, riferite al periodo dal 06/01/01 - 09/03/09.

Figura 3 – Iistogramma delle quotazioni giornaliere e settimanali del titolo Campari dal 06/01/01 al 09/03/09



La forma dell'istogramma è molto simile per dati giornalieri e settimanali per l'autosimilarità delle quotazioni di borsa.

Sulla serie P_t possono essere calcolati dei semplici indici descrittivi globali che permettono di evidenziare aspetti generali del fenomeno. In particolare:

- la *media*

$$\bar{P} = \frac{1}{N} \sum_{t=1}^N P_t$$

ove N è la numerosità del fenomeno. Questo indice misura la quotazione media dell'attivo nel periodo considerato;

- la *mediana*:

$$Q_2 = M_2 = \left\{ \text{Valore centrale di } P_{(t)} \right\}$$

dove $P_{(t)}$ è la serie ordinata dei prezzi in senso non decrescente;

- il *valore massimo* $P_{(N)}$ e quello *minimo* $P_{(1)}$ di P_t nel periodo analizzato

$$P_{(N)} = \max\{P_t\}; \quad P_{(1)} = \min\{P_t\}$$

che forniscono informazione sui valori estremi delle quotazioni nel periodo considerato;

- la *deviazione standard* (*scarto quadratico medio* o *standard deviation* (*sd*)):

$$S = \left[\frac{1}{N} \sum_{t=1}^N (P_t - \bar{P})^2 \right]^{1/2}$$

che fornisce una misura della variabilità media delle quotazioni nel periodo analizzato;

- l'*asimmetria*:

$$\gamma_1 = \frac{1}{N} \sum_{t=1}^N \left(\frac{P_t - \bar{P}}{S} \right)^3$$

che fornisce indicazioni sulla forma dell'istogramma di P_t intorno al suo baricentro dato dalla media;

- la *curtosi*:

$$\gamma_2 = \frac{1}{N} \sum_{t=1}^N \left(\frac{P_t - \bar{P}}{S} \right)^4 - 3$$

che fornisce indicazioni sulle code dell'istogramma confrontate con quelle della distribuzione normale.

Nelle tabelle seguenti riportiamo questi indici descrittivi riferiti alle quotazioni Campari sopra rappresentate.

Tabella 5 – *Indici descrittivi delle quotazioni Campari giornaliere e settimanali*

	CAM_G	CAM_S
μ = Media (mean)	5.401	5.371
Q_2 = Mediana	5.535	5.530
$P_{(N)}$ = Massimo	8.190	8.250
$P_{(1)}$ = Minimo	2.680	2.620
S = Standard deviation	1.564	1.576
γ_1 = Asimmetria	-0.047	-0.049
γ_2 = Curtosi	-1.331	-1.341
N = Numero osservazioni	1609	320

La serie P_t può essere rappresentata anche con un *box-plot*. Per costruire questo grafico è necessario calcolare:

- la *mediana*: Q_2 ,
- il *primo quartile*, come mediana della prima metà della rilevazione ordinata: Q_1 ,
- il *terzo quartile*, come mediana della seconda metà della rilevazione ordinata: Q_3 ,
- il *cardine inferiore*:

$$h_1 = Q_1 - 1.5(Q_3 - Q_1)$$

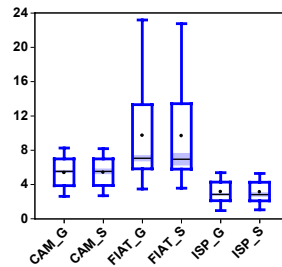
- il *cardine superiore*:

$$h_2 = Q_3 + 1.5(Q_3 - Q_1)$$

Ricordiamo che in $(Q_3 - Q_1)$ cadono, per costruzione, il 50% delle osservazioni.

Nella figura che segue riportiamo i box-plot riferiti alle quotazioni giornaliere e settimanali dei titoli Campari, Fiat, Banca Intesa S. Paolo. Come si può notare dal grafico, le quotazioni Fiat sono quelle che hanno avuto maggiori oscillazioni nel periodo considerato e sono le più asimmetriche.

Figura 4 – Box-plot delle quotazioni giornaliere e settimanali di Campari, Fiat, Banca Intesa S. Paolo

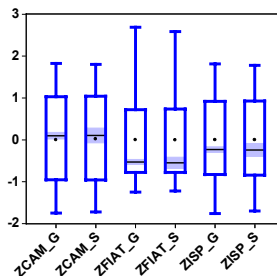


Quando si confrontano k rilevazioni riferite ad attivi diversi è necessario eliminare la diversa unità di misura dei differenti fenomeni. Per far questo dalle serie osservate $P_{i,t}$, $i=1, 2, \dots, k$, si ricavano quelle *standardizzate*

$$ZP_{i,t} = \frac{P_{i,t} - \bar{P}_i}{S_i}, \quad i=1, 2, \dots, k; \quad t=1, 2, \dots, N$$

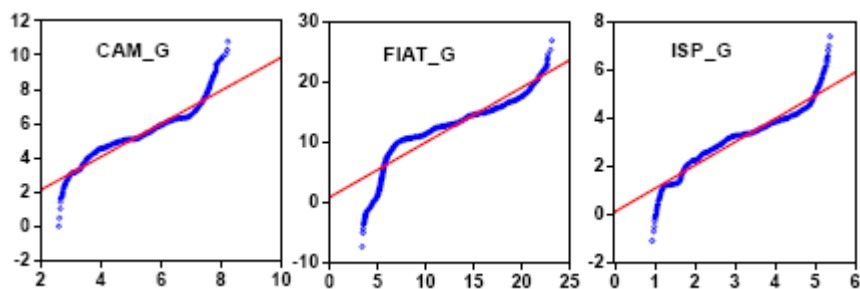
Le standardizzate, per costruzione, hanno media nulla e standard deviation pari ad uno e sono indipendenti dalle unità di misura utilizzate per rilevare le diverse serie (in questo caso il valore nominale delle azioni). Ciò permette di confrontare, a parità di media e variabilità, le diverse serie. Nella figura 5 sono riportati i box-plot per Campari, Fiat, Banca Intesa S. Paolo, con osservazioni sia giornaliere che settimanali, ma standardizzate. Come si può notare, i grafici sono molto diversi dal caso non standardizzato riportato nella figura 4 anche se le oscillazioni di Fiat risultano ancora un po' più marcate di quelle degli altri due titoli

Figura 5 – Box-plot per le serie giornaliere e settimanali standardizzate di Campari, Fiat, Banca Intesa S. Paolo



Un modo alternativo di rappresentare una serie di prezzi è quello di confrontarli con una distribuzione tipo, in questo caso la *Normale standardizzata*. Per far questo si calcolano i percentili empirici della serie standardizzata ZP_t e gli stessi percentili della Normale standardizzata, le coppie di punti così ottenuti si riportano in un diagramma scatter. Nel grafico, inoltre, viene tracciata una retta corrispondente alla diagonale di lato (-4; 4) questo perché nella Normale standardizzata praticamente tutto il fenomeno si esaurisce in quell'intervallo. Il grafico così ottenuto prende il nome di *Quantile-Quantile plot (Q-Qplot)*. Più i percentili osservati si discostano dalla retta, più la distribuzione di ZP_t è lontana da quella della Normale.

Figura 6 – *Q-Qplot delle quotazioni giornaliere di Campari, Fiat, Intesa S. Paolo*



Dal grafico 6 si può osservare come le quotazioni dei tre titoli siano molto lontane dalla distribuzione Normale.

Abbiamo visto che se la serie dei prezzi viene standardizzata si elimina da questa l'unità di misura, ma si perdono le informazioni sulla media e la variabilità del fenomeno. Per evitare questo inconveniente ed ottenere dalla serie osservata una nuova serie che sia indipendente solo dall'unità di misura utilizzata si può costruire la serie dell'*indice a base fissa*:

$${}_1(IP)_t = \frac{P_t}{P_1}, \quad t=1, 2, \dots, N$$

Questa nuova serie preserva l'evoluzione, la diversità fra le medie e la diversità delle variabilità degli attivi analizzati, ma è indipendente dall'unità di misura del fenomeno. A titolo di esempio nella figura 7 sono riportate le serie degli indici a base fissa dei titoli Campari, Fiat, B. Intesa S. Paolo dal 01/01/03 al 10/03/09.

Figura 7 – Serie degli indici giornalieri a base fissa per Campari, Fiat, Intesa S. Paolo

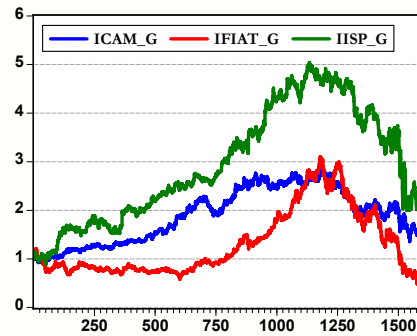
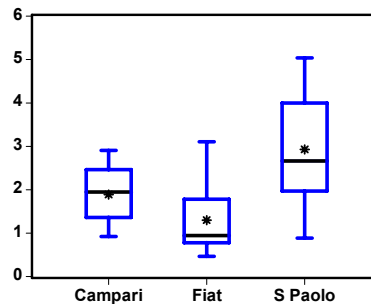


Figura 8 – Box-plot per la serie degli indici di Campari, Fiat, B. Intesa S. Paolo



Come si può osservare da questi grafici, il titolo che è cresciuto ed ha poi subito la più forte contrazione è quello di Banca Intesa S. Paolo. Questo titolo ha raggiunto la quotazione massima il 07/05/07 quando è stato quotato più di 5 volte rispetto al periodo iniziale del 01/01/03, per poi contrarsi repentinamente negli anni 2007 e 2008 e toccare il minimo di 1.4 volte, rispetto al periodo iniziale nel 09/03/09. Queste forti oscillazioni sul titolo di una Banca, in confronto a quanto avvenuto nell'economia reale mostrano come siano state le attività finanziarie ad aver sofferto maggiormente nella crisi iniziata nella tarda primavera del 2007 ed ancora lungi, nel 2009, dal ritenersi esaurita. La crisi finanziaria ha colpito per prima le banche (il massimo di Banca Intesa è stato registrato il 07/05/07) per poi iniziare ad estendersi alle attività reali con alcuni mesi medi di ritardo.

Dalle serie giornaliere possono essere derivate cinque sottoserie settimanali una per ogni giorno di apertura delle borse: da quella di tutti i lunedì fino a quella di tutti i venerdì. Nelle figure seguenti questa operazione è stata fatta per le quotazioni Campari e serve per capire se fra i diversi giorni della settimana vi sono differenze strutturali nelle quotazioni del titolo.

Figura 9 – Box-plot dei 5 giorni per settimana delle quotazioni Campari

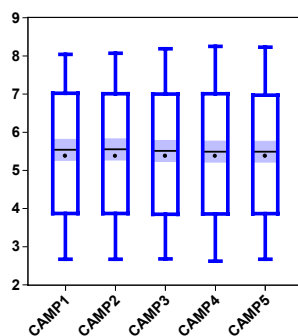


Figura 10 – Istogrammi dei 5 giorni per settimana delle quotazioni Campari

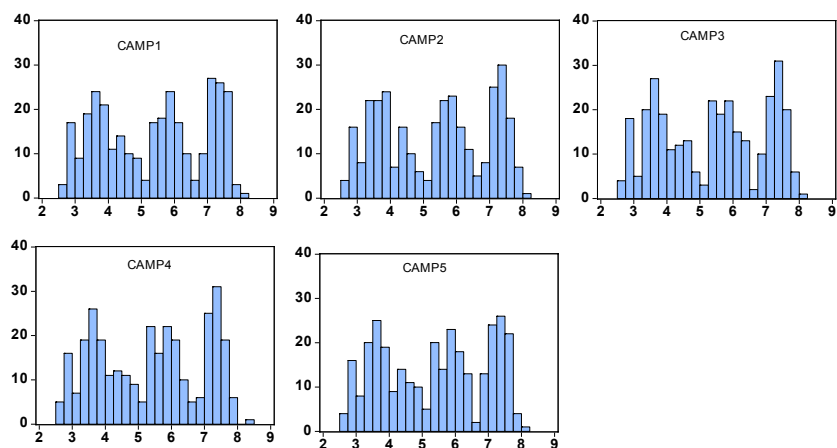


Tabella 6 – Indici descrittivi delle sottoserie di ciascun giorno della settimana per le quotazioni Campari

	CAMP_1	CAMP_2	CAMP_3	CAMP_4	CAMP_5
μ	5.371677	5.371056	5.373416	5.368944	5.371994
M_e	5.540000	5.550000	5.510000	5.490000	5.490000
$r_{(1)}$	8.040000	8.070000	8.190000	8.250000	8.230000
$r_{(N)}$	2.670000	2.670000	2.680000	2.620000	2.670000
S	1.578841	1.582502	1.582691	1.573520	1.572836
γ_1	-0.051089	-0.049721	-0.044720	-0.049555	-0.051563
γ_2	1.649817	1.654419	1.662677	1.666441	1.663763

Come si può notare dalle figure 9 e 10 e dalla tabella 6, e come era atteso, fra i diversi giorni non si rilevano differenze strutturali di nessun tipo nelle quotazioni né dai box-plot, né dagli istogrammi, né dagli indici descrittivi.

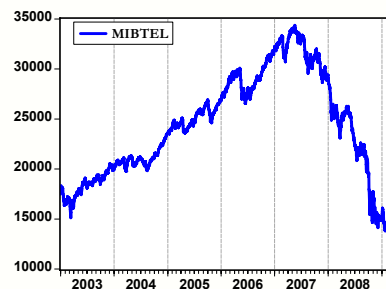
Le elaborazioni e le rappresentazioni grafiche che abbiamo presentato possono essere utilizzate su indici che descrivono l'evoluzione di un *portafoglio* di attività finanziarie o di un intero mercato borsistico.

Supponiamo di avere M attivi ed indichiamo con P_{it} , $i=1, 2, \dots, M$, la quotazione al tempo $t=1, 2, \dots, N$ dell'attivo i -simo e con q_{it} il numero di attivi scambiati tra $t-1$ e t . L'evoluzione temporale congiunta degli M attivi può essere analizzata costruendo il relativo indice sintetico di *Paasche*:

$${}_1(IP)_t = \frac{\sum_{i=1}^M P_{it} q_{it}}{\sum_{i=1}^M P_{i1} q_{i1}} = \frac{\text{Media delle quotazioni al tempo } t}{\text{Media delle quotazioni al tempo 1 scelto come base}}$$

Di solito l'indice viene moltiplicato per un fattore di aggiustamento per tener conto delle ristrutturazioni del portafoglio (uscita di attivi e/o ingresso di nuovi) che avvengono nel tempo. In molti casi, per semplicità di elaborazione, al posto della formula di *Paasche* si usa quella di *Laspeyres* che pondera con le quantità del tempo base e la cui evoluzione è, di solito, abbastanza simile a quella ottenuta con la formula di *Paasche*. Nella figura 11 riportiamo il grafico dei valori dell'indice di chiusura aggiustato MIBTEL (indice telematico della borsa di Milano) osservato giornalmente dal 01/01/03 al 18/03/09 e costruito con una formula simile a quella sopra descritta.

Figura 11 – *Indice MIBTEL delle quotazioni giornaliere aggiustate dal 01/01/03 al 18/03/09*



Da questo grafico emerge chiaramente la grave crisi finanziaria che ha interessato tutte le piazze e mercati con inizio nel maggio 2007. Gli indici sintetici delle quotazioni possono essere analizzati con gli stessi strumenti e

rappresentazioni grafiche che sono stati utilizzati per le quotazioni dei singoli attivi.

Un modo alternativo per estrarre informazioni sull'evoluzione delle quotazioni di un attivo è supporre che queste siano la somma di un *segnale* μ_t ed un *errore* ε_t :

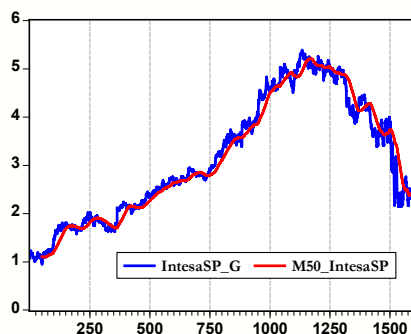
$$P_t = \mu_t + \varepsilon_t, \quad t = 1, 2, \dots, N$$

Le componenti μ_t ed ε_t non sono direttamente osservabili e bisogna derivarle con qualche tecnica. Un modo molto semplice per avere una stima del segnale è calcolare sulla serie delle quotazioni una *media mobile non centrata* di k termini, con k molto più piccolo di N :

$$\mu_{k,t} = \frac{1}{(k-1)+1} \sum_{j=0}^{k-1} P_{t-j} = \frac{1}{k} \sum_{j=0}^{k-1} P_{t-j}, \quad t = k+1, k+2, \dots, N$$

Naturalmente, per $k = N$ si ottiene la media delle quotazioni nel periodo considerato, mentre per $k = 1$ si ha la serie originaria P_t . In pratica, k è molto più piccolo di N e di solito varia fra 20 e 30. La stima del *segnale* $\mu_{k,t}$ ha un andamento più smussato, più liscio di P_t ma è sfasata rispetto a questa nel senso che i suoi picchi e le depressioni sono spostati più a destra rispetto a quelli osservati su P_t . Nella figura 12 è riportato il grafico delle quotazioni e della relativa media mobile con $k=50$ riferito alle quotazioni di Banca Intesa S. Paolo

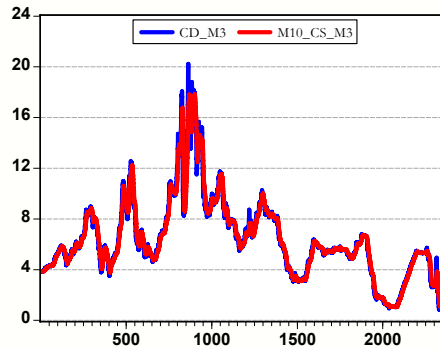
Figura 12 – *Quotazioni giornaliere e media mobile non – centrata (in rosso) con $k=50$ per Banca Intesa S. Paolo nel periodo 01/01/03 – 10/03/09*



Come si può osservare dalla figura 12, la serie $\mu_{k,t}$ è più liscia di P_t ed è sfasata rispetto a questa, inoltre si perdono le prime $(k-1)$ (in questo caso $50-1=49$) osservazioni. Lo sfasamento in $\mu_{k,t}$ è usato dagli analisti tecnici per prevedere, noto P_t , la quotazione P_{t+1} . Se, in fatti, P_t taglia $\mu_{k,t}$ dal basso verso l'alto sono attesi valori crescenti delle quotazioni, mentre se la taglia dall'alto verso il basso sono attesi valori decrescenti delle quotazioni.

Nella figura 13 è riportato il grafico delle quotazioni settimanali e della relativa media mobile non centrata, con $k=10$, dei Certificati di Deposito USA a tre mesi osservati nel periodo 06/12/64 – 10/04/09.

Figura 13 – Certificati di Deposito USA a tre mesi e relativa media mobile (in rosso) non centrata con $k=10$ osservati nel periodo 06/12/64 – 10/04/09

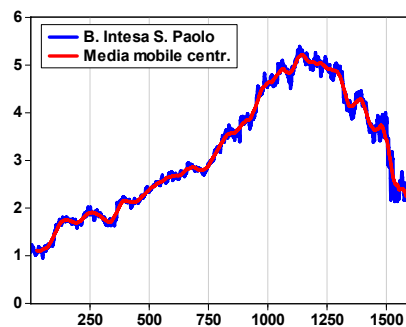


Invece che una *media mobile non centrata* è possibile costruirne una *centrata* che non sfasa ma fa perdere k informazioni sia nel periodo iniziale che in quello finale di osservazione:

$$\mu_{kc,t} = \frac{1}{2(k-1)+1} \sum_{j=-(k-1)}^{k-1} P_{t-j} = \frac{1}{2k-1} \sum_{j=-k+1}^{k-1} P_{t-j}, \quad t = k+1, k+2, \dots, N-k$$

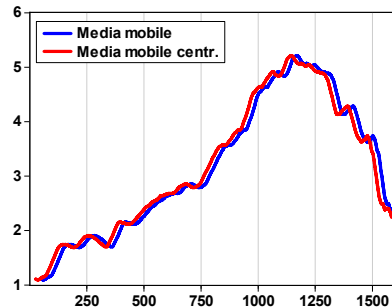
Nella figura 14 è riportata la media mobile centrata per le quotazioni Banca Intesa S. Paolo con $k=25$ per lo stesso periodo di osservazione della figura 12. Come si può notare dalla figura 14, la stima con media mobile centrata non sfasa ed è una stima più precisa del segnale, ma fa perdere $(k-1)$ informazioni sia all'inizio che alla fine della serie e queste ultime sono molto importanti per descrivere possibili scenari futuri.

Figura 14(a) – *Quotazioni giornaliere e media mobile centrata (in rosso) con $k=25$ per Banca Intesa S. Paolo*



Nella figura 14(b) sono riportate le due medie mobili, quella centrata con $k=25$ e quella non centrata con $k=50$. Come si può osservare l'andamento di fondo è simile, ma quella non centrata è sfasata: sale più velocemente e scende più lentamente di quella centrata.

Figura 14(b) – *Media mobile centrata (in rosso) e non centrata per le quotazioni giornaliere di Banca Intesa S. Paolo*



Dalle quotazioni e dalla stima del segnale è possibile ottenere una stima dell'errore non centrato e centrato:

$$\varepsilon_{k,t} = P_t - \mu_{k,t}; \quad \varepsilon_{kc,t} = P_t - \mu_{kc,t}$$

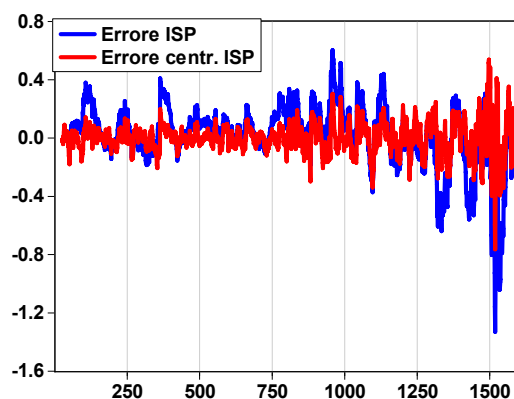
Tabella 7 – *Indici descrittivi per le quotazioni giornaliere di Banca Intesa S. Paolo e sue componenti*

	$\mu_{k,t}$	$\varepsilon_{k,t}$	$\mu_{kc,t}$	$\varepsilon_{kc,t}$	P_t
μ	3.1854	0.0150	3.1863	0.0011	3.1367

M_e	2.8574	0.0438	2.8581	0.0013	2.8500
$x_{(N)}$	5.2126	0.6030	5.2122	0.5385	5.3900
$x_{(1)}$	1.0982	-1.3304	1.0984	-0.7632	0.9500
S	1.2172	0.2259	1.2169	0.1084	1.2427
γ_1	0.1444	-1.4624	0.1440	-0.1520	0.1759
γ_2	-1.2802	4.0435	1.2807	4.1356	-1.2468
N	1560	1560	1559	1559	1609

La stima dell'errore è utile per capire quanto è forte la struttura presente nella serie delle quotazioni e se in queste esistono altre eventuali componenti non catturate dalla media mobile. Nella tabella 7 sono riportati i relativi indici descrittivi.

Figura 15 – Errori non centrati e centrati (in rosso) per le quotazioni di Banca Intesa S. Paolo



Nella figura 16 è riportato l'istogramma dei due errori, quello non centrato (a sinistra) e quello centrato, mentre nella figura 17 è riportato il Q-Qplot delle due serie di errori. Come si può osservare dalla tabella 7 e dai grafici 15 – 18, la serie degli errori non centrati ha una media ed una variabilità più elevata di quella centrata, inoltre possiede una forte asimmetria negativa e questa è una conseguenza del più accentuato sfasamento della serie. Entrambe le serie degli errori hanno una forte curtosi il che le allontana dall'ipotesi di normalità e quindi di pura accidentalità.

Figura 16 – Istogrammi dell'errore non centrato (a sinistra) e di quello centrato per le quotazioni Banca Intesa S. Paolo

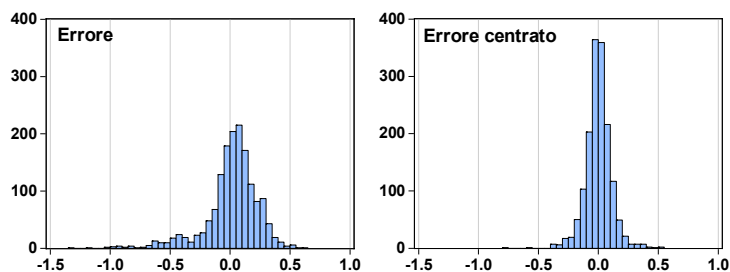
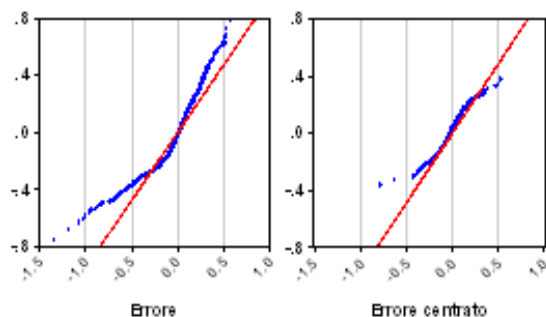
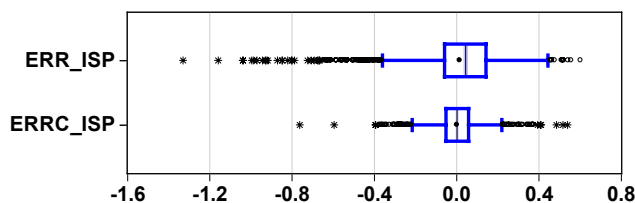


Figura 17 – $Q-Q$ plot delle due serie degli errori delle quotazioni giornaliere di Banca Intesa S. Paolo



Negli errori esistono altre strutture, oltre ai livelli medi, alla variabilità, alla asimmetria, ecc. ed in particolare quella che viene chiamata *volatilità* che analizzeremo diffusamente nel seguito. Osserviamo ancora, dalla figura 15, le forti oscillazioni degli errori rilevabili a partire dalla seconda metà del 2007 in concomitanza con la crisi mondiale dei mercati.

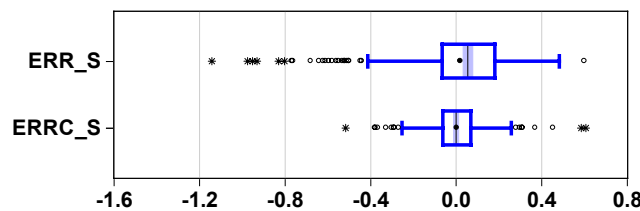
Figura 18 – Box -plot delle due serie degli errori delle quotazioni giornaliere di Banca Intesa S. Paolo



Le medie mobili possono essere invece che semplici, come le due che

abbiamo qui utilizzate, anche ponderate, inoltre possono essere calcolate sullo stesso fenomeno ma rilevato con cadenza diversa (settimanale, mensile, ...). Per l'autosimilarità delle quotazioni degli attivi, le conclusioni a cui si arriva non sono molto diverse da quelle derivabili analizzando dati giornalieri. Nella figura 19, a titolo di esempio, riportiamo il box-plot degli errori, non centrati e centrati, per la serie delle quotazioni settimanali riferite allo stesso periodo di osservazione di Banca Intesa S. Paolo.

Figura 19 – Box-plot dei residui non centrati (a sinistra) e centrati delle rilevazioni settimanali di Banca Intesa S. Paolo



Confrontando la figura 18 con la 19 si deducono informazioni molto simili.

2.2 I rendimenti degli attivi finanziari

Nelle analisi finanziarie e monetarie, oltre ai livelli dei prezzi degli attivi, sono di estremo interesse i relativi *rendimenti*. Se indichiamo con r_t il rendimento di un attivo quotato P_t al tempo $t = 1, 2, \dots, N$, è ragionevole ritenere che sia

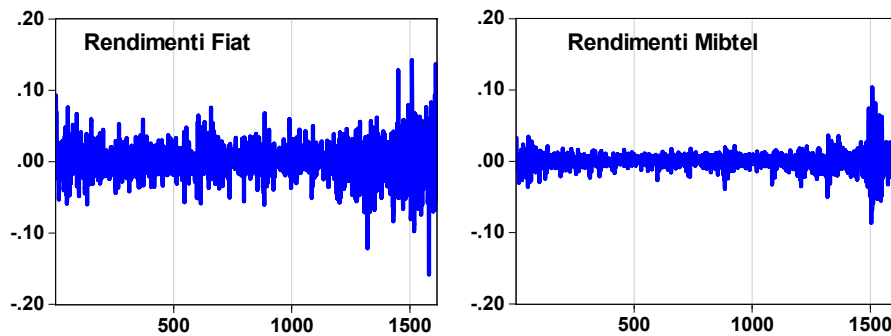
$$P_t = P_{t-1}(1 + r_t^a), \quad t = 2, 3, \dots, N \quad (1)$$

e quindi che il *prezzo* di oggi sia pari a quello di ieri moltiplicato per il fattore $(1 + r_t^a)$. La quantità r_t^a è una percentuale che può essere sia positiva che negativa a seconda che P_t sia cresciuto rispetto al giorno precedente o sia diminuito. Questa quantità è proprio il *rendimento* del giorno considerato. Dalla (1) si ricava

$$r_t^a = \frac{P_t}{P_{t-1}} - 1, \quad t = 2, 3, \dots, N \quad (2)$$

che mostra come r_t^a sia un numero puro indipendente dall'unità di misura utilizzata per rilevare i prezzi. Inoltre, il rapporto P_t / P_{t-1} non è altro che l'*indice a base mobile* della serie dei prezzi dell'attivo, per tanto i rendimenti possono essere interpretati come la serie degli indici a base mobile dei prezzi a meno di uno. Il rendimento, a differenza del prezzi, è un *flusso* e quindi sempre riferito ad un prestabilito intervallo temporale che, in questo caso, coincide con la cadenza (il giorno) con cui i prezzi sono rilevati. Nella figura 20, utilizzando una scala identica, riportiamo i rendimenti giornalieri del titolo Fiat e quelli medi ottenuti dall'indice Mibtel nel periodo 02/01/03 – 10/03/09.

Figura 20 – Rendimenti giornalieri di Fiat e Mibtel dal 02/01/03 al 10/03/09

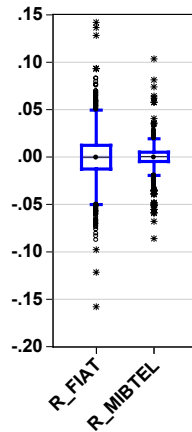


Su questo grafico è possibile fare alcune considerazioni:

- i rendimenti Fiat hanno oscillato, in tutto il periodo analizzato, in modo nettamente più marcato rispetto al rendimento medio del mercato borsistico di Milano rappresentato dall'indice Mibtel;
- le oscillazioni, causa la crisi finanziaria, di entrambe le serie sono più marcate nell'ultimo periodo analizzato e sembrano in via di normalizzazione negli ultimi due mesi considerati;
- le oscillazioni avvengono intorno allo zero e si presentano a grappoli: a periodi di forti turbolenze seguono periodi di più moderata oscillazione;
- le oscillazioni negative sono spesso inferiori di numero a quelle positive, ma più marcate di queste.

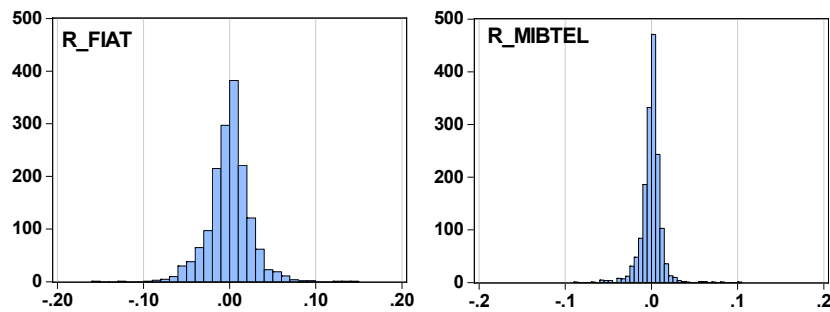
Nella figura 21 riportiamo il Box-plot delle due serie di rendimenti.

Figura 21 – Box-plot dei rendimenti giornalieri di Fiat e Mibtel



che conferma le considerazioni fatte in precedenza. In fine, nella figura 22 riportiamo gli istogrammi delle due serie di rendimenti.

Figura 22 – Istogrammi dei rendimenti giornalieri di Fiat e Mibtel



Nella tabella 8 riportiamo gli indici descrittivi relativi ai rendimenti giornalieri del titolo Fiat e dell'intero mercato borsistico italiano con riferimento al periodo 02/01/03 – 10/03/09. Questi grafici e la tabella confermano le considerazioni che avevamo fatto in generale: le medie sono praticamente nulle, la variabilità dei rendimenti Fiat è circa il doppio di quella Mibtel, presentano entrambi una lieve asimmetria positiva ed una marcata curtosi con quella di Mibtel (γ_{2M}) pari a circa il triplo di quella di Fiat (γ_{2F}):

$$\gamma_{2M} = 15.58 - 3 = 12.578; \quad \gamma_{2F} = 7.16 - 3 = 4.16$$

Tabella 8 – *Indici descrittivi dei rendimenti giornalieri di Fiat e Mibtel*

	FIAT	MIBTEL
μ	-6.51E-05	-0.000207
M_e	0.000000	0.000500
$r_{(N)}$	0.152700	0.109300
$r_{(1)}$	-0.146100	-0.082400
S	0.024480	0.011994
γ_1	0.248157	0.134686
γ_2	4.155962	12.57688
N	1608	1608

Un aspetto negativo dei rendimenti definiti nella (2) è che il suo campo di variazione è $[-1; \infty)$:

$$r_t^a = \frac{P_t}{P_{t-1}} - 1 = \begin{cases} -1 & \text{se } P_t = 0 \\ \infty & \text{se } P_{t-1} \rightarrow 0 \end{cases}$$

e non è compatibile con la distribuzione degli errori accidentali rappresentata dalla Normale che, come noto, è definita nell'intervallo $(-\infty; \infty)$. D'altra parte $(1 + r_t^a)$ assume valori nell'intervallo $(0; \infty)$ ed è compatibile con una log Normale (vedi Cap. 3). Per superare questo inconveniente un modello alternativo alla (1) utilizzato correntemente per descrivere i *prezzi* è

$$P_t = P_{t-1} e^{r_t}, \quad t = 2, 3, \dots, N \quad (3)$$

dalla (3), utilizzando i logaritmi, si ottiene

$$r_t = \log(P_t) - \log(P_{t-1}) \quad (4)$$

e posto $p_t = \log(P_t)$ e $\nabla p_t = p_t - p_{t-1}$ si ricava equivalentemente $r_t = \nabla p_t$. I *rendimenti* così ottenuti vengono anche chiamati *log rendimenti* e sono quelli che utilizzeremo nel seguito. Nella figura 23 riportiamo i rendimenti ottenuti con la (2) e con la (4) riferiti alle quotazioni giornaliere Fiat osservati dal 01/11/06 al 10/03/09.

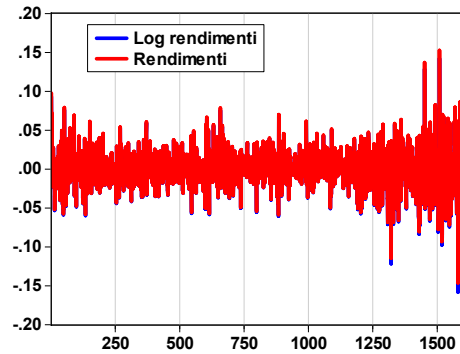
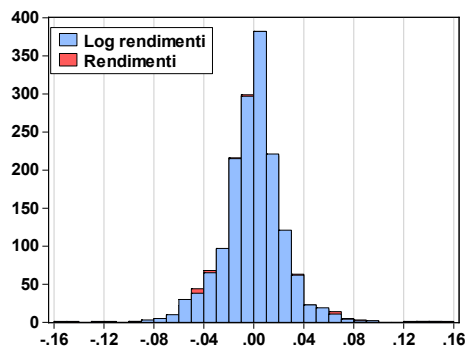
Figura 23 – Rendimenti giornalieri di Fiat calcolati con la (2) (RA_Fiat) e con la (4)

Figura 24 – Istogrammi dei rendimenti giornalieri di Fiat ottenuti con i due metodi di calcolo



Come si può notare dalla figura 23, i rendimenti ottenuti con la (2) sono indistinguibili (almeno graficamente) da quelli generati con la (4). Nella figura 24 sono riportati gli istogrammi dei due rendimenti, riferiti al periodo di osservazione 01/01/03 - 10/03/09, che mostrano più chiaramente che le differenze fra i due metodi di calcolo dei rendimenti sono praticamente trascurabili.

Nella tabella 9 riportiamo gli indici sintetici dei rendimenti del titolo Fiat ottenuti con i due metodi di calcolo e riferiti al periodo di osservazione 01/01/03 - 10/03/09. Come si può notare, le due serie presentano indici pressoché indistinguibili: il calcolo dei rendimenti con i due diversi metodi porta, praticamente, agli stessi risultati dato che, per una proprietà dei logaritmi, se $|x|$ è piccolo, risulta $\log(1+x) \approx x$.

Tabella 9 – *Indici descrittivi dei rendimenti giornalieri Fiat ottenuti con i due metodi di calcolo*

	r_t	r_t^a
μ	-0.000367	-0.000065
M_e	0.000000	0.000000
$r_{(N)}$	0.142100	0.152700
$r_{(1)}$	-0.157900	-0.146100
S	0.024454	0.024480
γ_1	0.028582	0.248157
γ_2	3.963698	4.155962
N	1608	1608

Infatti, applicando i logaritmi ad ambo i membri della (1) ed utilizzando la (4) si ottiene

$$r_t = \log(P_t) - \log(P_{t-1}) = \log(1 + P_t / P_{t-1} - 1) = \log(1 + r_t^a) \approx r_t^a$$

che giustifica quanto visto empiricamente nella tabella 9 e nelle figure 23 e 24. Come già sottolineato, nel seguito utilizzeremo sistematicamente i rendimenti ottenuti con la (4) perché questi sono compatibili con le v.c. *Normale* e *t di Student* (vedi Cap. 3).

Abbiamo già osservato che, mentre i prezzi sono variabili di stock e quindi vengono rilevati in un dato istante (fine giornata, fine settimana, fine mese, ecc.), i rendimenti sono flussi e quindi si riferiscono ad un intervallo di tempo (in un giorno, in una settimana, in un mese, ecc.). Questa considerazione implica che i rendimenti riferiti ad un intervallo di tempo sono uguali alla somma dei rendimenti dei sotto intervalli in cui quello viene diviso. Per esempio, l'intervallo temporale $[t-2, t)$ è equivalente all'intervallo $[t-2, t-1)$ più $[t-1, t)$ e quindi i rendimenti rilevati nell'intervallo $[t-2, t)$ sono dati dai rendimenti osservati in $[t-2, t-1)$ sommati a quelli osservati in $[t-1, t)$. D'altra parte, per la (3) si ha

$$\begin{cases} P_{t-1} = P_{t-2} e^{r_{t-1}} \\ P_t = P_{t-1} e^{r_t} \end{cases}$$

e sostituendo la prima di queste equazioni nella seconda si ricava

$$P_t = \left(P_{t-2} e^{r_{t-1}} \right) e^{r_t} = P_{t-2} e^{r_{t-1} + r_t}$$

considerando i logaritmi di ambo i membri si ottiene infine

$$\log(P_t) - \log(P_{t-2}) = r_{t-1} + r_t$$

Osserviamo che questa proprietà non è verificata se si utilizza la (1) per definire i rendimenti, infatti in questo caso si ottiene

$$P_t = P_{t-2} (1 + r_{t-1}^a)(1 + r_t^a) = P_{t-2} + r_{t-1}^a + r_t^a + r_t^a r_{t-1}^a$$

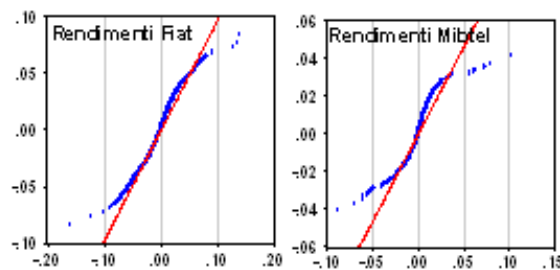
dalla quale risulta

$$\frac{P_t}{P_{t-2}} - 1 = r_{t-1}^a + r_t^a + r_t^a r_{t-1}^a \approx r_{t-1}^a + r_t^a$$

se $|r_{t-1}^a|$ ed $|r_t^a|$ sono “piccoli”, cosa che effettivamente succede per dati giornalieri. In questo caso, effettivamente, $|r_t^a r_{t-1}^a|$ è una quantità molto più piccola sia di $|r_t^a|$ che di $|r_{t-1}^a|$. Questo è un altro motivo che ci ha portato a definire i prezzi con la (3) ed i relativi rendimenti con la (4).

Nella figura seguente riportiamo il Q-Qplot dei rendimenti giornalieri, calcolati con la (4), riferiti alle quotazioni di chiusura aggiustate di Fiat e Mibtel. Da questo grafico notiamo come entrambe le distribuzioni dei rendimenti siano lontane dalla normalità, ma questa diversità è più accentuata per Mibtel, in accordo a quanto già visto nella figura 22 e nella tabella 8.

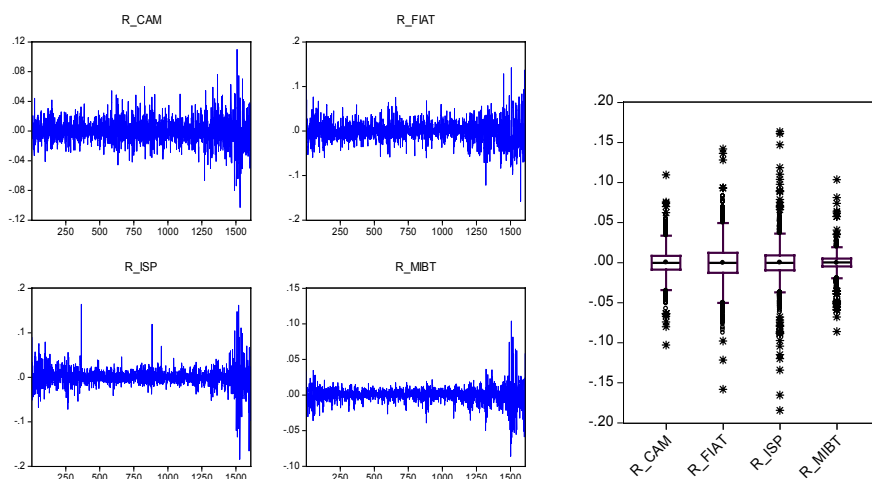
Figura 25 – Q-Qplot per i rendimenti giornalieri aggiustati di Fiat e Mibtel



La figura 25 conferma che l'allontanamento dalla normalità è molto più marcato nei rendimenti Mibtel che in quelli Fiat.

In conclusione, i rendimenti di attivi finanziari sono sistematicamente diversi da quelli ottenibili da una normale. Questo è dovuto, sostanzialmente, alla loro elevata curtosi che è sempre presente in questi fenomeni e li caratterizza. Naturalmente, come fatto per i prezzi, sui rendimenti è possibile calcolare gli indici $(h_1, Q_1, Q_2, Q_3, H_2)$ e derivare il relativo box-plot. D'altra parte i rendimenti sono numeri puri e quindi i confronti fra rendimenti di attivi diversi può essere fatto direttamente senza doverli preventivamente normalizzare. Questo è già stato fatto nella figura 21 e nella tabella 8 in cui sono stati messi a confronto grafici ed indici riferiti ai rendimenti di Fiat e Mibtel. Ovviamente, questi confronti possono essere fatti sui rendimenti di più di due attivi, in teoria, per un numero qualsiasi. In pratica, per non ottenere figure illeggibili, è buona norma non confrontarne simultaneamente più di una decina. Nella figura 26 riportiamo i grafici lineari, i box-plot paralleli, ed i Q-Qplot dei rendimenti giornalieri riferiti alle quotazioni di chiusura di Banca Intesa S. Paolo, Campari, Fiat e Mibtel, mentre nella tabella 10 riportiamo i relativi indici descrittivi. Questi grafici ed indici confermano che i rendimenti posseggono una modesta asimmetria, negativa o positiva, ma una marcata curtosi che li allontana considerevolmente dall'ipotesi di normalità. Da questo punto di vista i rendimenti medi, calcolati sull'indice Mibtel, riferiti all'intera piazza borsistica di Milano, sono quelli che maggiormente si discostano dall'ipotesi di normalità come si può osservare dai relativi Q-Qplot.

Figura 26 – Grafico, box-plot e Q-Qplot per i rendimenti giornalieri Campari, Fiat, Banca Intesa S. Paolo, Mibtel da 01/01/03 a 10/03/09



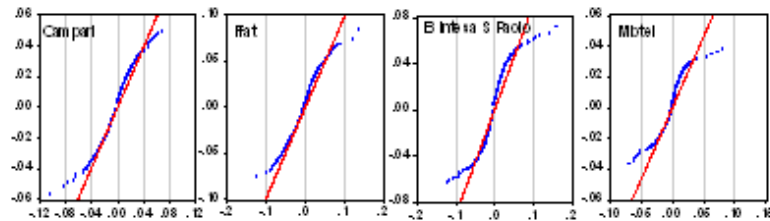


Tabella 10 – Indici descrittivi dei rendimenti giornalieri di Campari, Fiat, Banca Intesa S. Paolo, Mibtel da 01/01/03 a 10/03/09

	R_CAM	R_FIAT	R_ISP	R_MIBT
μ	0.000216	-0.000367	0.000167	-0.000277
M_e	0.000000	0.000000	0.000000	0.000500
$r_{(N)}$	0.109727	0.142100	0.163629	0.103700
$r_{(1)}$	-0.102760	-0.157900	-0.184304	-0.086000
S	0.016872	0.024454	0.022421	0.011998
γ_1	0.074861	0.028582	-0.108898	-0.120741
γ_2	3.885324	3.963698	13.19821	11.80354
N	1607	1608	1607	1608

Dai grafici, dalle tabelle e dalle considerazioni fatte possiamo affermare che i rendimenti definiti secondo la (4) sono compatibili:

- a) con una normale

$$r_t = \mu + \sigma a_t$$

con a_t , $t = 1, 2, \dots$ normali standardizzate indipendenti. Il limite di questo modello è che non è in grado di catturare l'eccesso di curtosi che sistematicamente si osserva nei rendimenti di attivi finanziari;

- b) con una t di Student con numero di gradi di libertà superiore a 4

$$r_t = \mu + \sigma a_t$$

con a_t , $t = 1, 2, \dots$ di Student indipendente con gradi di libertà $g > 4$. Questo modello è in grado di catturare eccesso di curtosi, ma

non eventuale asimmetria;

c) con una *mistura* di variabili indipendenti

$$r_t = p\mu_1 + (1-p)\mu_2 + p\sigma_1 a_{1t} + (1-p)\sigma_2 a_{2t}$$

dove a_{it} è t di Student, con g_i , $i = 1, 2$, gradi di libertà, indipendenti per ogni i e ogni t . Questo modello è in grado di catturare asimmetria ed eccesso di curtosi;

d) con un *modello a soglia* a due regimi

$$r_t = I_{t-1} (\mu_1 + \sigma_1 a_{1t}) + (1 - I_{t-1}) (\mu_2 + \sigma_2 a_{2t})$$

$$I_{t-1} = \begin{cases} 1 & \text{se } r_{t-1} < 0 \\ 0 & \text{se } r_{t-1} \geq 0 \end{cases}$$

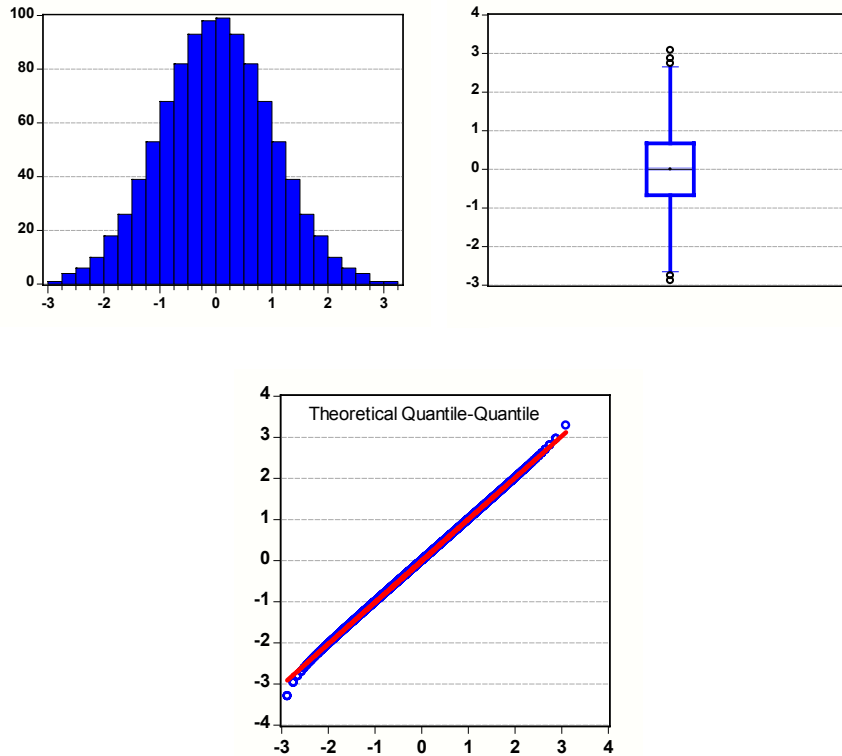
con a_{it} normali standardizzate, oppure t di Student, indipendenti per ogni t . Questo modello è in grado di catturare eccesso di curtosi, eventuale asimmetria presente nei rendimenti e permette di dare una interpretazione più puntuale sulla struttura del mercato distinguendo fra rendimenti *globali* r_t e quelli locali dei singoli regimi: per esempio primo regime recessivo (*mercato orso*) e secondo espansivo (*mercato toro*).

Su alcuni di questi modelli torneremo più in dettaglio nel seguito, per ora mostriamo con alcune simulazioni come si comportano.

MODELLO A)

Nella figura 27 riportiamo l'istogramma di frequenza, il box-plot ed il Q-Qplot della simulazione di $N=10000$ osservazioni, con $\mu = -0.000065$ e $\sigma = 0.0189$, sotto l'ipotesi di normalità. La media e lo scarto quadratico medio sopra indicati sono pari alla media di quelli riportati nella tabella 10.

Figura 27 – Istogramma, box-plot e Q-Qplot della simulazione di $N=10000$ osservazioni normali con $\mu = -0.000065$ e $\sigma = 0.0189$

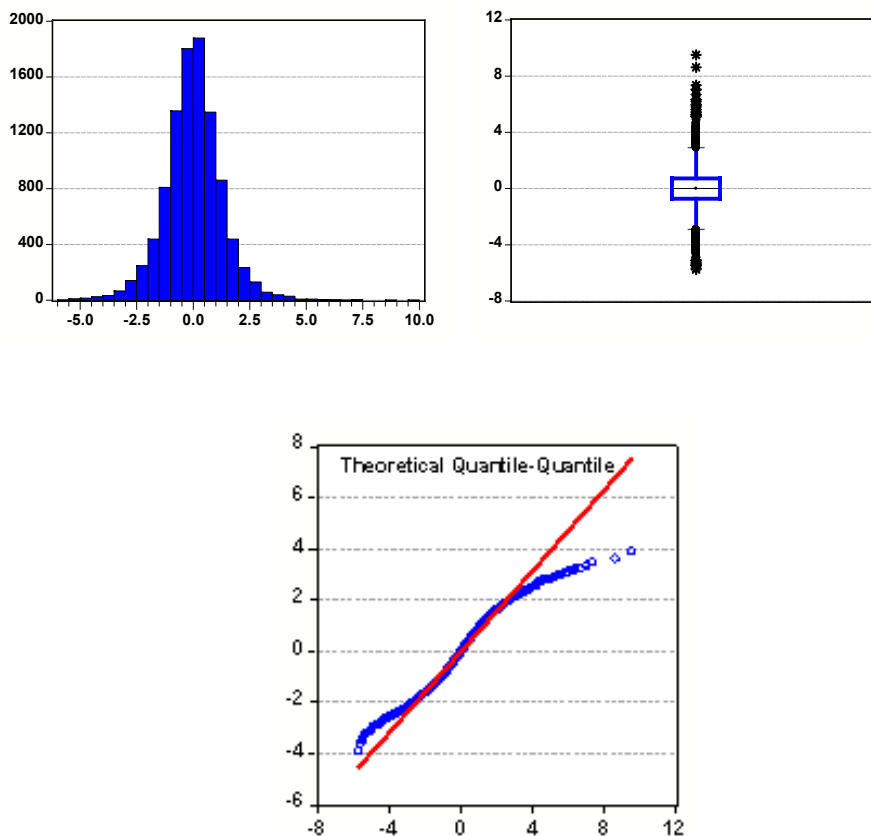


Confrontando queste figure con quelle ottenute su dati reali, si può notare come il modello riproduca le caratteristiche più rilevanti dei rendimenti di attivi in modo grossolano.

MODELLO B)

Nella figura 28 riportiamo l'istogramma di frequenza, il box-plot ed il Q-Qplot della simulazione di $N=10000$ osservazioni con $\mu = -0.000065$, $\sigma = 0.0189$ ed a distribuito come una t di Student con $g = 5$ gradi di libertà. Confrontando queste figure con quelle ottenute su dati reali, si può notare come il modello approssimi i rendimenti di attivi in modo nettamente più soddisfacente rispetto a quello normale. Questo spiega perché è molto utilizzato nelle pratiche applicazioni. Notiamo che il modello t di Student è più flessibile del precedente visto che è funzione di tre parametri (μ, σ, g) mentre l'altro era funzione solo dei primi due.

Figura 28 – Istogramma, box-plot e Q-Qplot della simulazione di $N=10000$ osservazioni t di Student con $g = 5$ gradi di libertà, $\mu = -0.000065$, $\sigma = 0.0189$



MODELLO C)

E' un modello funzione dei sette parametri $(g_1, g_2, p, \mu_1, \mu_2, \sigma_1, \sigma_2)$ e

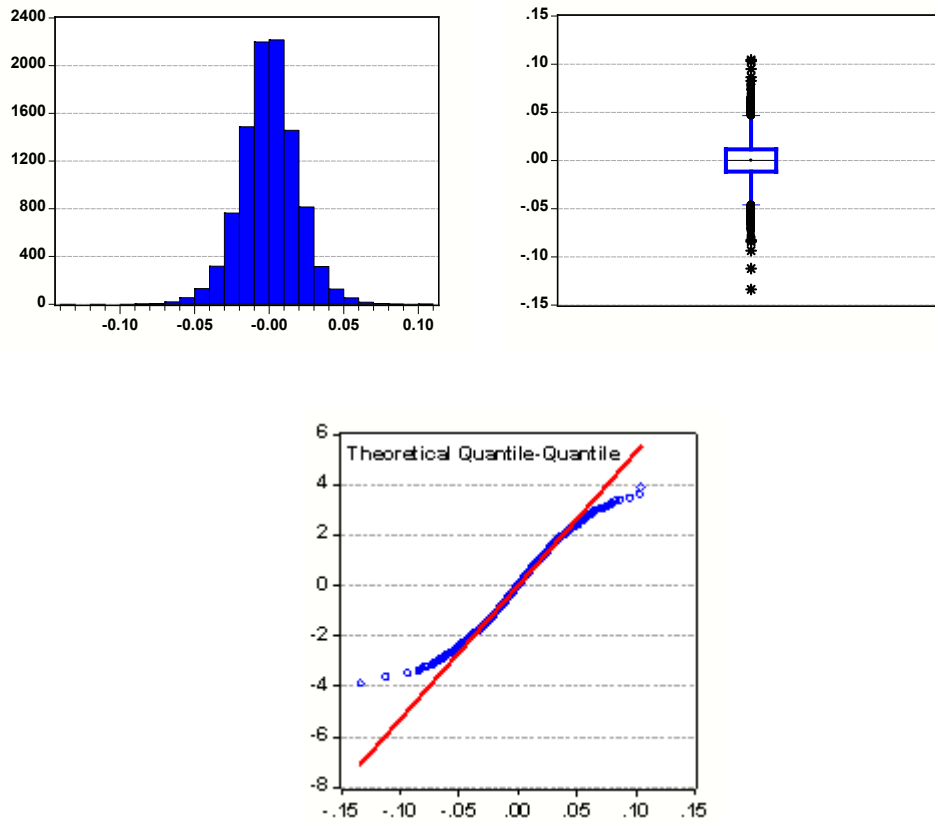
quindi più flessibile dei precedenti. Nella figura 29 riportiamo l'istogramma di frequenza, il box-plot ed il Q-Qplot della simulazione di $N=10000$ osservazioni con

$$g_1 = 10, g_2 = 6, p = 0.6, \mu_1 = 0.003558, \mu_2 = -0.0055, \sigma_1 = 0.015$$

$$\sigma_2 = 0.032657 \text{ di modo che sia } \mu = -0.000065 \text{ e } \sigma = 0.0189$$

Se si confrontano queste figure con quelle relative a dati reali, arriviamo alla conclusione che anche questo modello è capace di cogliere gli aspetti peculiari più rilevanti dei rendimenti di attivi finanziari e monetari.

Figura 29 – Istogramma di frequenza, box-plot e Q-Qplot della simulazione di $N=10000$ osservazioni con $g_1 = 10$, $g_2 = 6$, $p = 0.6$, $\mu_1 = 0.003558$, $\mu_2 = -0.0055$, $\sigma_1 = 0.015$, $\sigma_2 = 0.032657$



MODELLO D)

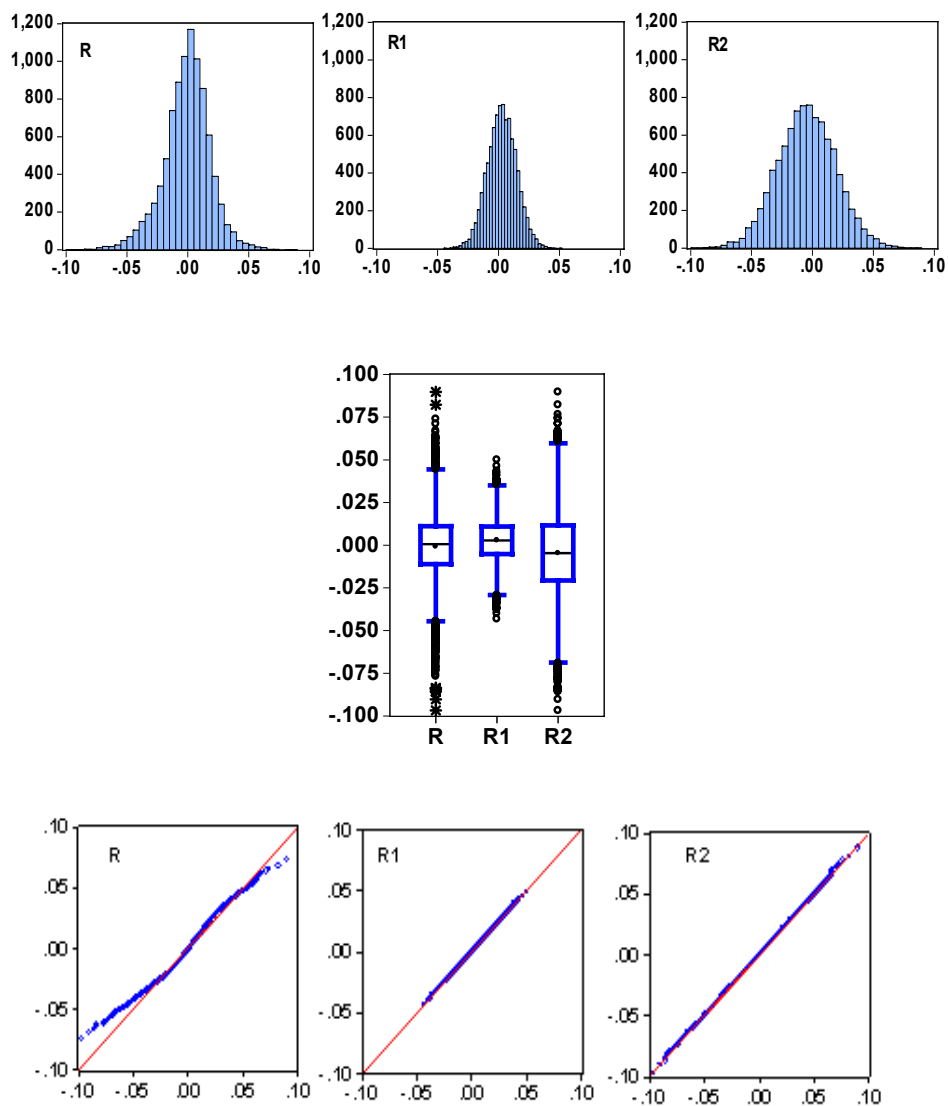
Consideriamo in primo luogo il caso in cui ciascun regime è normale (normalità locale) e risulti

$$p = E(I_{t-1}) = 0.52; \quad \mu_1 = 0.0032; \quad \mu_2 = -0.004$$

$$\sigma_1 = 0.012; \quad \sigma_2 = 0.0024$$

I risultati della simulazione sono riportati nella figura 30.

Figura 30 – Istogramma, box-plot e Q-Qplot di una simulazione di $N=10000$ osservazioni ottenuti da un modello a soglia a due regimi normali



Come si può osservare da questi grafici, anche se i due regimi sono (localmente) normali, i rendimenti mostrano (globalmente) un eccesso di curtosi ed asimmetria negativa come si può anche osservare dalla tabella 11 in cui sono riportati gli indici descrittivi (locali e globali) della serie (globale) dei rendimenti e dei suoi regimi (locale).

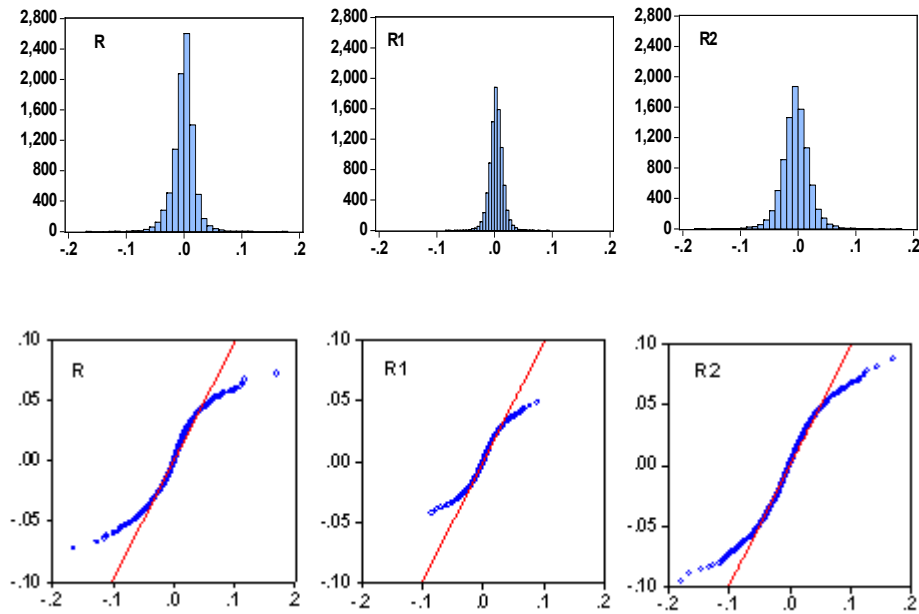
Tabella 11 – *Indici descrittivi della serie e dei suoi due regimi*

	R	R1	R2
μ	-0.000664	0.002901	-0.004597
M_e	0.000680	0.002874	-0.004652
$r_{(N)}$	0.089810	0.050105	0.089810
$r_{(1)}$	-0.096749	-0.043174	-0.096749
S	0.019081	0.011922	0.023845
γ_1	-0.361113	-0.005901	-0.005901
γ_2	1.231331	0.031586	0.031586

Naturalmente, se si suppone che, localmente, i singoli regimi sono compatibili con una t di Student ciascuna con $g > 4$ gradi di libertà l'eccesso di curtosi in r_t sarà ancora più marcato. Con gli stessi parametri del caso precedente, utilizzando una t di Student standardizzata con $g=5$ gradi di libertà si sono ottenuti i risultati della tabella 12 che evidenziano, per l'appunto, come la curtosi sia cresciuta sostanzialmente rispetto al caso normale e sia maggiore di quella dei singoli regimi. Le stesse considerazioni valgono per la figura 31 in cui sono riportati gli istogrammi ed i Q-Qplot.

Tabella 12 – *Indici descrittivi di r_t ed i suoi regimi sotto l'ipotesi $t_{(5)}$*

	R	R1	R2
μ	-0.000130	0.003247	-0.003907
M_e	0.001069	0.003206	-0.003988
$r_{(N)}$	0.236922	0.123661	0.236922
$r_{(1)}$	-0.231753	-0.110677	-0.231753
S	0.019117	0.012036	0.024072
γ_1	-0.361964	0.141609	0.141609
γ_2	11.14597	6.213668	6.213668

Figura 31 – Istogramma e Q-Qplot di r_t ed i suoi regimi sotto l'ipotesi $t_{(5)}$ 

Anche sui rendimenti è possibile calcolare medie mobili di modo che possano essere interpretati come somma di un *segnale* più un *errore*:

$$r_t = \mu_{r,t} + \varepsilon_t, \quad t = 1, 2, \dots, N$$

in particolare, la media mobile non centrata a k termini dei rendimenti, che può essere considerata una stima del segnale, è data da

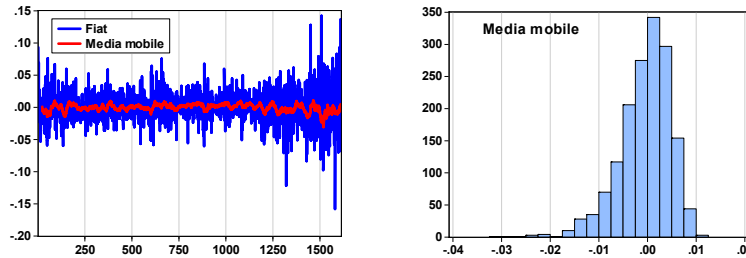
$$\mu_{kr,t} = \frac{1}{k} \sum_{j=0}^{k-1} r_{t-j}, \quad t = k, k+1, \dots, N$$

Questa ci permette di ottenere una stima dell'errore:

$$\varepsilon_{k,t} = r_t - \mu_{kr,t}, \quad t = k, k+1, \dots, N$$

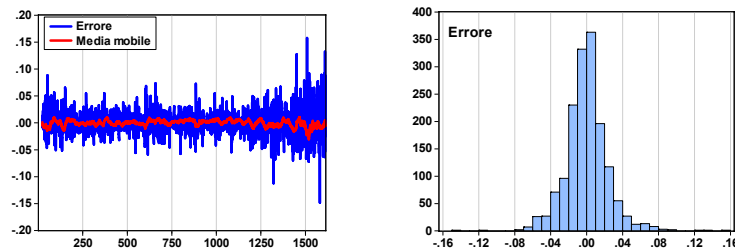
Nella figura 32 è riportato il grafico e l'istogramma di $\mu_{kr,t}$ per i rendimenti giornalieri Fiat con $k=22$ pari a circa un mese borsistico.

Figura 32 – Rendimenti ed istogramma della media mobile non centrata di Fiat



Nella figura 33 è riportato il grafico degli errori $\varepsilon_{k,t}$ ed il relativo istogramma.

Figura 33 – L'errore ed l'istogramma per i rendimenti giornalieri Fiat



Da questi grafici si nota che il segnale presente nei rendimenti di attivi è molto debole anche se, in questo caso, asimmetrico positivo, mentre gli errori sono asimmetrici negativi ed entrambi lontani dal caso normale come si può notare dagli indici riportati nella tabella 13.

Tabella 13 – Indici descrittivi dei rendimenti giornalieri Fiat, e delle stime del segnale e dell' errore

	Errore	Segnale	Rendimenti
μ	0.000171	-0.000364	-0.000364
M_e	-0.000212	9.01E-05	0.000000
$r_{(N)}$	0.148172	0.006516	0.142102
$r_{(1)}$	-0.149970	-0.016172	-0.157903
S	0.023953	0.003837	0.024416
γ_1	0.303817	-1.026562	0.028447
γ_2	4.212097	1.403379	3.984774

Nella figura 34 è riportato il grafico della serie dei rendimenti Fiat e della sua media mobile centrata con il relativo istogramma di frequenze. Come si può notare, questi grafici sono molti simili a quelli della figura 32 in cui era stata utilizzata una media mobile non centrata. Questo è dovuto al fatto che, nei rendimenti, la struttura del segnale è molto debole e pressoché costante. Anche il grafico e l'istogramma dell'errore centrato sono molto simili a quelli non centrati. Nella figura 35 riportiamo le medie mobili non centrate e centrate dei rendimenti Fiat e come si può osservare, il loro andamento è molto simile anche se, come atteso, quello non centrato è sfasato.

Figura 34 – Grafico ed istogramma della media mobile centrata, $k = 25$, dei rendimenti Fiat

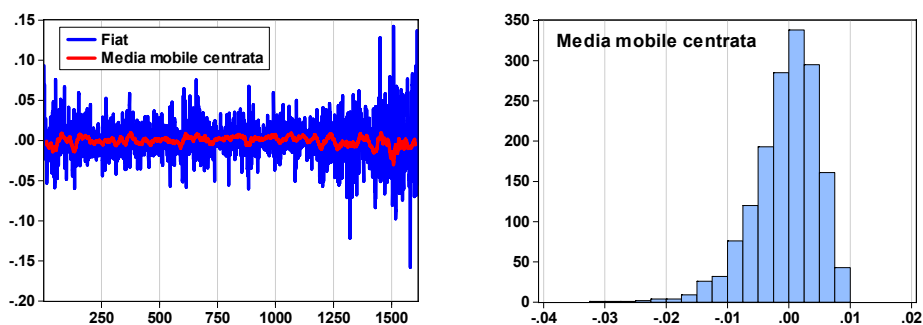
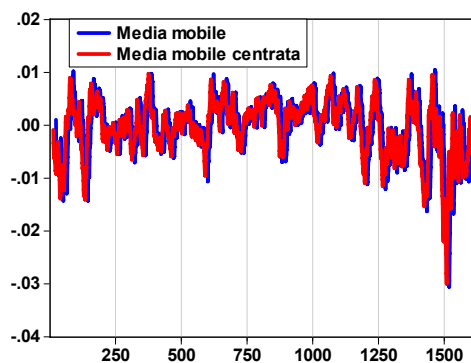


Figura 35 – Medie mobili centrate (in rosso) e non centrate dei rendimenti Fiat



La serie dei rendimenti globali può essere decomposta in due sottoserie (rendimenti locali), ciascuna delle quali corrisponde ad un particolare regime, più precisamente:

$$r_t = I_{t-1}r_t + (1 - I_{t-1})r_t, \quad \text{con } I_{t-1} = \begin{cases} 1 & \text{se } r_{t-1} < 0 \\ 0 & \text{se } r_{t-1} \geq 0 \end{cases} \quad (5)$$

E' possibile quindi analizzare separatamente $r_{1t} = I_{t-1}r_t$ ed $r_{2t} = (1 - I_{t-1})r_t$ in modo da avere informazioni più dettagliate sui rendimenti. Naturalmente, perché queste elaborazioni abbiano senso è necessario che la media di I_{t-1} non sia molto diversa da 0.5 e quindi in ciascun regime vi siano un numero di informazioni sufficienti per poter condurre le elaborazioni. A titolo di esempio, questa tecnica è applicata ai rendimenti giornalieri Fiat osservati dal 02/01/03 al 10/03/09 dove è risultato $p = E(I_{t-1}) = 0.473$.

Figura 36 – Rendimenti della Fiat e quelli dei due regimi

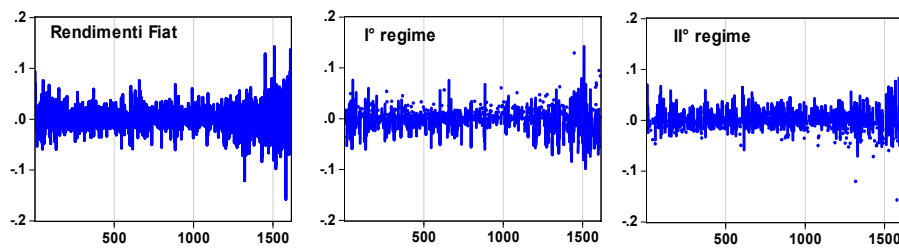


Tabella 14 – Indici sintetici descrittivi dei rendimenti Fiat e dei suoi due regimi

	RF	RF1	RF2
μ	-0.000364	-0.000702	-0.000171
M_e	0.000000	0.000000	0.000000
$r_{(N)}$	0.142102	0.142102	0.136634
$r_{(1)}$	-0.157903	-0.097638	-0.157903
S	0.024416	0.026140	0.022555
γ_1	0.028447	0.159669	-0.210512
γ_2	3.984774	2.762935	5.569118
N	1613	763	849

Nella figura 36 sono riportati i grafici dei rendimenti globali di Fiat e quelli locali dei due regimi e, come si può notare, questi sono sostanzialmente differenti fra di loro e dalla serie originaria. Il primo regime rappresenta

l'andamento del mercato quando il giorno prima si è avuto un risultato negativo ed il secondo regime quando il giorno prima vi è stato un risultato non negativo. Nella tabella 14 sono riportati e confrontati gli indici sintetici descrittivi relativi ai rendimenti giornalieri Fiat ed ai suoi due regimi. Da questa tabella si può osservare come nel periodo esaminato la Fiat ha avuto in media un rendimento negativo e come sia più marcato nel primo regime. Inoltre, i due regimi si differenziano sostanzialmente fra di loro e con la serie dei rendimenti per la diversa asimmetria, negativa nel secondo e positiva nel primo, e la marcata curtosi, molto più elevata nel secondo rispetto al primo.

Nella figura 37 sono riportate le stime delle densità dei rendimenti globali giornalieri di Fiat e dei suoi regimi (locali), nella figura 38 i boxplot e nella figura 39 i Q-Qplot. In particolare da quest'ultimo grafico si può osservare come i rendimenti siano molto più lontani dalla normalità rispetto ad i due regimi.

Figura 37 – Istogrammi della serie dei rendimenti giornalieri Fiat e dei suoi regimi

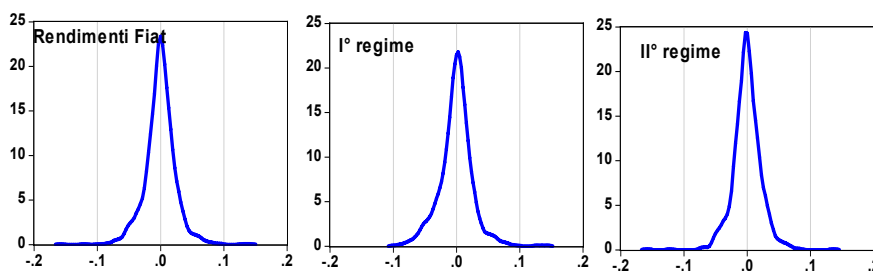


Figura 38 – Box-plot della serie dei rendimenti giornalieri Fiat e dei suoi due regimi

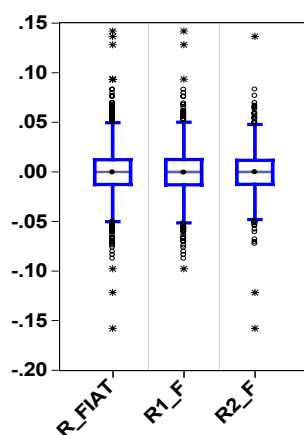
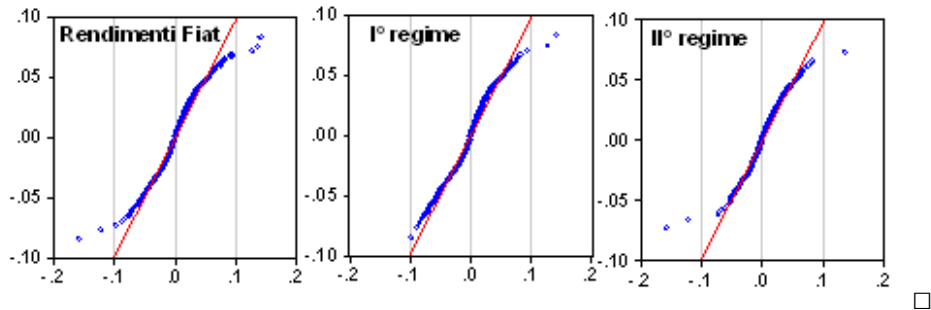


Figura 39 – *Q-Qplot della serie dei rendimenti giornalieri Fiat e dei suoi due regimi*



Della serie dei rendimenti r_t , $t = 1, 2, \dots, N$, di un attivo è possibile derivare la nuova serie $\rho(h)$, $h = 1, 2, \dots, K$, con $K \leq N/4$, che descrive i legami lineari esistenti all'interno dei rendimenti. La serie $\rho(h)$ è detta *funzione di autocorrelazione* dei rendimenti ed è definita da

$$\rho(h) = \frac{\sum_{t=1}^{N-h} (r_t - \bar{r})(r_{t+h} - \bar{r})}{\sum_{t=1}^N (r_t - \bar{r})^2}, \quad h = 1, 2, \dots, K \quad (6)$$

La (6) può essere interpretata anche come

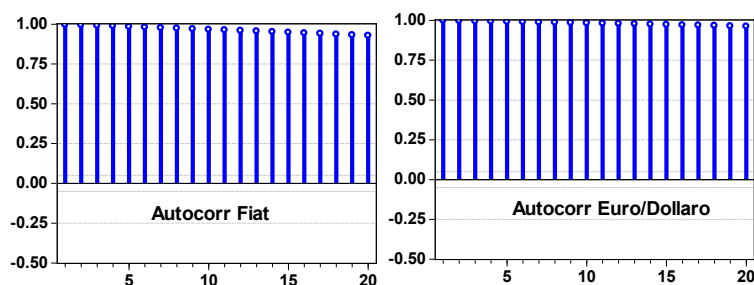
$$\rho(h) = \frac{\text{covarianza}(r_t, r_{t+h})}{[\text{varianza}(r_t) \text{varianza}(r_{t+h})]^{1/2}} = \text{correlazione}(r_t, r_{t+h})$$

e quindi è sempre $-1 \leq \rho(h) \leq 1$. Al crescere di h la funzione $\rho(h)$ decresce a zero e misura l'intensità dei legami lineari esistenti all'interno di rendimenti distanti h tempi. Di solito, nei rendimenti i legami lineari sono praticamente assenti. Naturalmente, questa funzione può essere calcolata per rendimenti rilevati con qualsiasi cadenza: giornaliera, settimanale, mensile, ecc. Un indice simile può essere calcolato anche per le quotazioni di un attivo, ma in tal caso la funzione di autocorrelazione decresce molto lentamente a zero e questo è un sintomo che il fenomeno ha una struttura lineare del tutto erratica (passeggiata aleatoria) tipica delle quotazioni di attivi: $P_t = P_{t-1} + \varepsilon_t$.

Nella figura 40 è riportata la funzione $\rho(h)$, $h=1, 2, \dots, 20$, per le

quotazioni giornaliere del titolo Fiat e per quelle Euro/Dollaro riferite al periodo 01/01/03 – 10/03/09. Come si può osservare per entrambe le serie l'autocorrelazione parte da un valore prossimo ad uno e decresce molto lentamente. Questo andamento è tipico di tutte le quotazioni di attivi finanziari di qualsiasi tipo.

Figura 40 – *Autocorrelazioni delle quotazioni giornaliere di Fiat ed Euro/Dollaro*

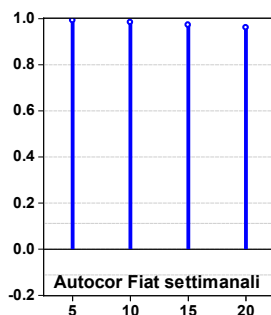


Se la cadenza invece che giornaliera è settimanale, ricordando che in tal caso la relazione fra quotazione giornaliera, P_t , e quotazioni settimanali $P_t^{(5)}$ (cinque giorni a settimana) è $P_t^{(5)} = P_{5t}$, per le rispettive funzioni di autocorrelazioni si ha allo stesso modo

$$\hat{\rho}^{(5)}(h) = \rho(5h), h=1, 2, \dots, K/5$$

In altri termini le autocorrelazioni delle serie settimanali sono date dalla quinta, decima, quindicesima, ecc., autocorrelazione della serie giornaliera. Questo implica che anche le autocorrelazioni delle quotazioni settimanali decadono anche loro a zero molto lentamente.

Figura 41 – *Autocorrelazioni delle quotazioni settimanali di Fiat*



Nella figura 41 sono riportate le prime 4 autocorrelazioni delle quotazioni Fiat settimanali riferite allo stesso periodo di osservazione della figura 40.

Calcoliamo ora le stesse autocorrelazioni per i rendimenti Fiat ottenuti sia con cadenza giornaliera che settimanale.

Figura 42 – *Autocorrelazioni dei rendimenti giornalieri di Fiat*

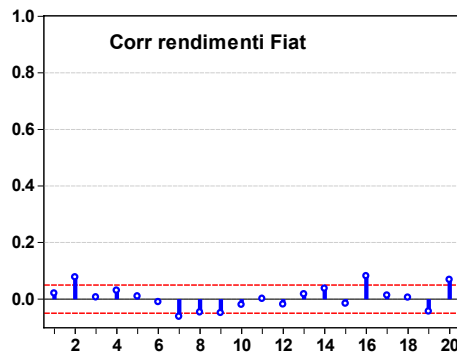
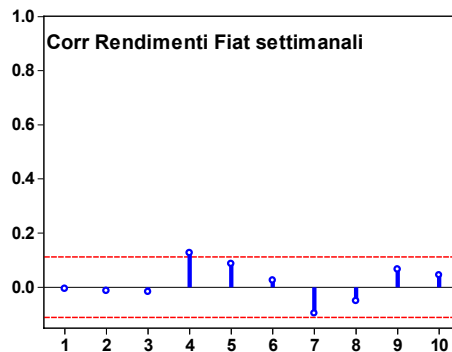


Figura 43 – *Autocorrelazioni settimanali dei rendimenti di Fiat*



e, come si può osservare, le autocorrelazioni, siano le rilevazioni giornaliere o settimanali, sono tutte piccole e vicine allo zero.

□

Analizziamo ora le autocorrelazioni dei rendimenti quando si è utilizzato un modello di struttura a soglia come quello riportato nella (5). In questo caso, se esistono forti legami lineari locali all'interno dei regimi questi saranno molto diversi nei due regimi. In generale, quindi ci si attende che le funzioni di autocorrelazioni locali dei singoli regimi siano molto diverse fra di loro e dai

rendimenti globali dell'attivo considerato. Più precisamente, indichiamo con $\rho_1(h)$ e $\rho_2(h)$ le autocorrelazioni del primo regime $r_{1t} = I_{t-1}r_t$ e del secondo regime $r_{2t} = (1 - I_{t-1})r_t$; posto: $p = \text{media}(I_{t-1}) = \text{media}(I_{t-h-1})$, $p_h = \text{media}(I_{t-1}I_{t-h-1})$ si avrà

$$\rho_1(h) = p_h \rho(h), \quad \rho_2(h) = (1 - 2p + p_h) \rho(h)$$

e risulta

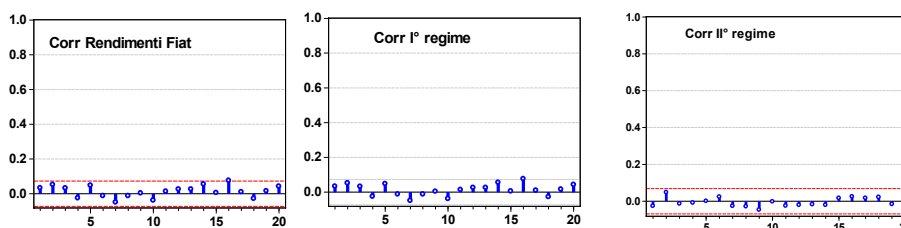
$$\rho(h) = p_h \rho(h) + 2(p - p_h) \rho(h) + (1 - 2p + p_h) \rho(h)$$

da cui si ottiene

$$\rho(h) = \frac{\rho_2(h) - \rho_1(h)}{1 - 2p} \quad (7)$$

Come si può osservare, l'autocorrelazione di r_t globale è una particolare funzione delle autocorrelazioni dei regimi (locali) e della proporzione p . Questo giustifica perché le autocorrelazioni dei singoli regimi possano, in via di principio, essere diverse fra di loro e da quelle dei rendimenti dell'attivo. D'altra parte, abbiamo già sottolineato come nei rendimenti, e quindi anche nei singoli regimi, vi sia pratica assenza di una qualsiasi struttura lineare. Questa considerazione ci porta a ritenere che anche le autocorrelazioni presenti nei regimi siano, escluso qualche caso particolare, tutte praticamente vicine allo zero. Nella figura 44, a titolo di esempio, riportiamo le autocorrelazioni dei rendimenti giornalieri e dei singoli regimi per Fiat.

Figura 44 – Autocorrelazioni dei rendimenti giornalieri di Fiat e quelle dei regimi



Come si può osservare da questi grafici, le autocorrelazioni dei rendimenti di ciascun regime sono effettivamente localmente diverse fra di loro e da quelle di r_t ma tutte praticamente nulle. In particolare, in questo caso, quelle del

secondo presentano un valore leggermente più marcato per $h=1$ mentre le altre sono molto piccole e vicine allo zero come avviene per il secondo regime e per r_t .

Le analisi sui rendimenti che abbiamo fatto per un singolo attivo possono essere estese ad un *portafoglio* o ad un intero mercato. Le elaborazioni riguarderanno il portafoglio o il mercato complessivamente considerato, ma non le relazioni che intercorrono tra i singoli attivi che lo compongono. Queste ultime possono essere effettuate solo in ambito multivariato.

Supponiamo, pertanto, che il nostro portafoglio sia composto da k attivi con quotazioni P_{it} , $i=1, 2, \dots, k$ per ciascuno di essi. Indichiamo con f_i , $i=1, 2, \dots, k$, il peso che l'attivo i -esimo ha sul portafoglio di modo che sia $f_i \geq 0$, $\sum_{i=1}^k f_i = 1$. Il valore complessivo del portafoglio è dato dalla media geometrica ponderata delle singole quotazioni:

$$P_t = \prod_{i=1}^k P_{it}^{f_i}$$

Ricordando la (3) si ha anche

$$P_t = \prod_{i=1}^k P_{it}^{f_i} = \left(\prod_{i=1}^k P_{it-1}^{f_i} \right) e^{r_t}$$

utilizzando i logaritmi di ambo i membri si ricavano i rendimenti

$$r_t = \log \left(\prod_{i=1}^k \frac{P_{it}^{f_i}}{P_{it-1}^{f_i}} \right) = \sum_{i=1}^k \left[\log(P_{it}) - \log(P_{it-1}) \right] f_i = \sum_{i=1}^k r_{it} f_i, \quad t = 1, 2, \dots, N \quad (8)$$

Dalla (8) si ha che il rendimento del portafoglio è una media aritmetica ponderata dei rendimenti dei singoli attivi che lo compongono. Da questa si ha anche che il rendimento medio del portafoglio è dato dalla media dei rendimenti dei singoli attivi:

$$\bar{r} = \sum_{i=1}^k \bar{r}_i f_i$$

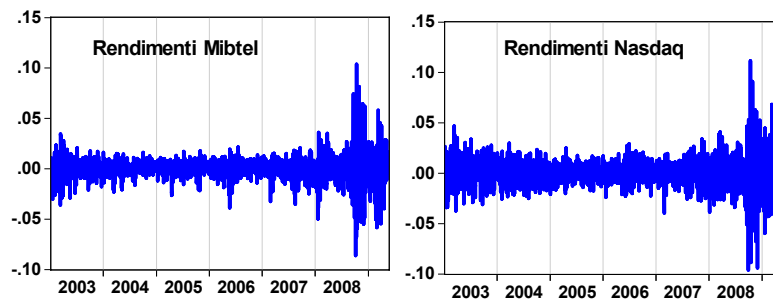
mentre la varianza del rendimento del portafoglio è pari a

$$\begin{aligned} S_r^2 &= \frac{1}{N} \sum_{t=1}^N (r_t - \bar{r})^2 = \frac{1}{N} \sum_{t=1}^N \left[\left(\sum_{i=1}^k r_{it} f_i - \sum_{i=1}^k \bar{r}_i f_i \right)^2 \right] = \\ &= \frac{1}{N} \sum_{t=1}^N \left[\sum_{i=1}^k (r_{it} - \bar{r}_i) f_i \right]^2 = \sum_{i=1}^k \sum_{j=1}^k \left[\frac{1}{N} \sum_{t=1}^N (r_{it} - \bar{r}_i) f_i (r_{jt} - \bar{r}_j) f_j \right] = \\ &= \sum_{i=1}^k \sum_{j=1}^k S_{ij} = \sum_{i=1}^k \sum_{j=1}^k \rho_{ij} S_i S_j \end{aligned} \quad (9)$$

dove con S_{ij} si è indicata la covarianza fra r_{it} ed r_{jt} , e con ρ_{ij} la relativa correlazione, S_i lo scarto quadratico medio di r_{it} . Dalla (9) si ha che la varianza del portafoglio è funzione degli scarti quadratici medi dei singoli attivi che lo compongono e dalle relazioni lineari che esistono fra questi e misurate dalle rispettive correlazioni.

Nella figura 45 riportiamo i rendimenti giornalieri medi dei mercati italiano ed USA misurati rispettivamente dall'indice *Mibtel* e *Nasdaq Composite* riferiti al periodo 01/01/03 – 10/03/09.

Figura 45 – Rendimenti medi di Mibtel e Nasdaq



Dalla figura si nota come il mercato milanese, se si esclude l'ultimo periodo, sia

stato meno variabile di quello USA. Per entrambi è chiara la crisi finanziaria iniziata nel 2007 ed ancora attiva nel 2009 visibile con le più ampie oscillazioni che si registrano in questo ultimo periodo di osservazione.

2.3 La volatilità degli attivi

Nel paragrafo precedente abbiamo mostrato che le caratteristiche più salienti dei rendimenti degli attivi finanziari sono:

- il livello medio vicino allo zero,
- l'asimmetria debole e spesso negativa,
- le relazioni lineari interne generalmente trascurabili,
- una marcata curtosi,
- le oscillazioni che si presentano a grappoli.

Le ultime due caratteristiche implicano che nei rendimenti di attivi devono esistere relazioni non lineari. Tale forma di non linearità è connessa alla variabilità del fenomeno e prende il nome di *volatilità*. La volatilità può essere interpretata come *l'attitudine degli attivi ad avere una variabilità non costante del tempo*. D'altra parte la variabilità è un aspetto non lineare dei fenomeni perché legata ai quadrati dei rendimenti e per tanto non può essere catturata dalle autocorrelazioni che, per costruzione, rilevano solo legami lineari.

Come i rendimenti di un attivo non sono direttamente osservabili, ma sono derivati dalle loro quotazioni, anche la volatilità non è direttamente osservabile, ma va ricavata tramite un qualche semplice modello dai rendimenti. Più precisamente, data la serie storica delle quotazioni di un attivo P_t , $t = 1, 2, \dots, N$, da questa si ricava la serie dei rendimenti r_t , $t = 2, 3, \dots, N$. Se indichiamo con \mathfrak{F}_{t-1} le informazioni che si hanno sul fenomeno al tempo $t-1$, la serie delle volatilità è data da:

$$\sigma_t^2 = \text{media}\left(\varepsilon_t^2 \mid \mathfrak{F}_{t-1}\right) = \text{media}\left((r_t - \bar{r})^2 \mid \mathfrak{F}_{t-1}\right) \quad (10)$$

o, in alternativa, da

$$\sigma_t = \sqrt{\text{media}\left(\varepsilon_t^2 \mid \mathfrak{F}_{t-1}\right)} = \sqrt{\text{media}\left((r_t - \bar{r})^2 \mid \mathfrak{F}_{t-1}\right)}$$

La volatilità può essere misurata, in prima approssimazione, con il quadrato dei rendimenti:

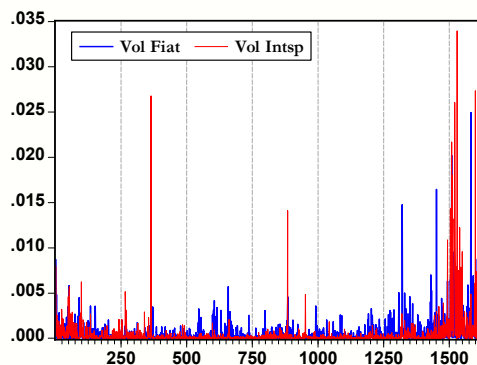
$$v_t = r_t^2, \quad t = 2, 3, \dots, N \quad (11)$$

che, ricordando la (3), permette di costruire una relazione fra prezzi degli attivi, rendimenti e volatilità:

$$v_t = r_t^2 = \left[\log\left(\frac{P_t}{P_{t-1}}\right) \right]^2$$

Nella figura 46 riportiamo le serie della volatilità giornaliera per Fiat e Banca Intesa S. Paolo per il periodo 01/01/03 – 10/03/09 misurata con la (11).

Figura 46 – *Volatilità giornaliera di Fiat e Banca Intesa S. Paolo*



Come si può osservare dalla figura 46, la volatilità di Banca Intesa è più marcata di quella di Fiat e questo in particolare nell'ultimo periodo per effetto della grave crisi finanziaria che, partita dalla finanza, si è estesa a tutto il tessuto produttivo. Dalla figura sopra riportata si può osservare che i valori elevanti della volatilità si presentano a grappoli: a periodi turbolenti si susseguono periodi di relativa calma dei mercati. In altri termini, nella volatilità esistono delle regolarità che possono essere catturate ed eventualmente previste.

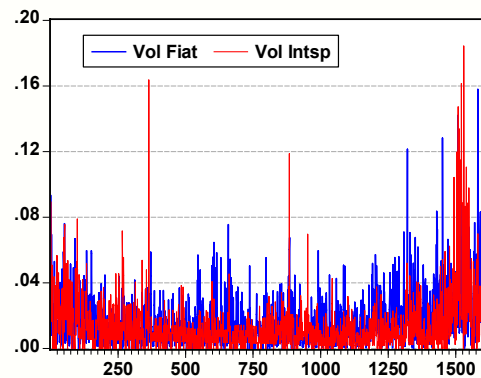
Spesso, la volatilità invece che essere misurata con i quadrati dei rendimenti, lo è con il loro valore assoluto:

$$v_t = |r_t|, \quad t = 2, 3, \dots, N \quad (12)$$

Un vantaggio della (12) è che è meno sensibile, rispetto alla (11), alla presenza nei rendimenti di valori eccezionali.

Nella figura 47 riportiamo la volatilità di Fiat e Banca Intesa S. Paolo misurata con la (12). Come si può notare, l'andamento complessivo della volatilità è lo stesso di quello della figura precedente, ma è più chiaramente leggibile.

Figura 47 – *Volatilità di Fiat e Banca Intesa S. Paolo calcolata con la (11)*

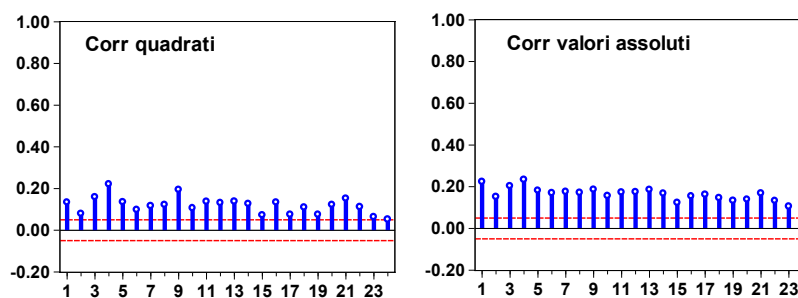


□

Della serie della volatilità, ottenuta con la (11) o la (12), è possibile derivare la serie della autocorrelazioni e, a differenza di quel che avviene per i rendimenti, nella volatilità esistono forti legami lineari, spesso persistenti.

Nella figura 48 sono riportate le prime $k=24$ autocorrelazioni per la volatilità giornaliera di Fiat calcolate con la (11) e con la (12). Come si può notare le autocorrelazioni sono, in ogni caso, significative e persistenti.

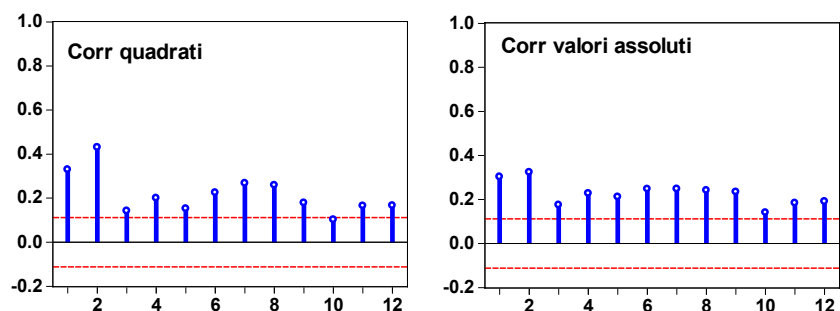
Figura 48 – *Autocorrelazioni della volatilità giornaliera di Fiat calcolata con la (11) (destra) e la (12) (sinistra)*



Naturalmente, le stesse autocorrelazioni per la volatilità possono essere

calcolate su dati con cadenza diversa da quella giornaliera, per esempio settimanale. Anche queste sono ancora significative e persistenti.

Figura 49 – Autocorrelazioni della volatilità settimanale di Fiat



Nella figura 49 sono riportate le prime 12 autocorrelazioni della volatilità settimanale di Fiat calcolata con la (11) e la (12). Come si può osservare, queste autocorrelazioni sono, in questo esempio, ancora più marcate di quelle per i dati giornalieri e non vi è grande differenza se queste sono calcolate con la (11) o la (12).

Sulla serie della volatilità v_t è possibile effettuare tutte le elaborazioni e le rappresentazioni grafiche che sono state fatte per le quotazioni degli attivi e per i rendimenti. In particolare, da $r_t = \bar{r} + \varepsilon_t$ si ricava $\varepsilon_t = r_t - \bar{r}$ e una stima più attendibile di σ_t^2 può essere ottenuta considerando una media mobile della serie $(r_t - \bar{r})^2$. Anche in questo caso le medie mobili possono essere sia centrate che non centrate. In particolare si può utilizzare la varianza mobile non centrata a k termini data da:

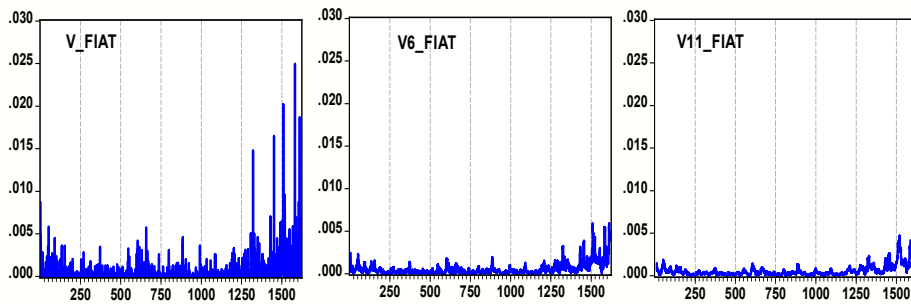
$$\sigma_{k,t}^2 = \frac{1}{k} \sum_{j=0}^{k-1} (r_{t-j} - \bar{r})^2, \quad t = k, k+1, \dots, N \quad (13)$$

La (13) per $k=1$ si riduce a v_t mentre per $k=N$ si ottiene la varianza globale σ^2 dei rendimenti. Il grafico di $\sigma_{k,t}^2$ o della sua radice quadrata $\sigma_{k,t}$ ci fornisce informazioni sull'evoluzione della volatilità dell'attivo considerato. La (13) è una *varianza mobile non centrata* e quindi sfasa. Al suo posto è possibile

calcolare, come è stato fatto per i rendimenti, una varianza mobile centrata che non sfasa, ma fa perdere sia le prime che le ultime $(k-1)$ informazioni. Queste ultime sono utili per effettuare previsioni sul fenomeno. In particolare, utilizzando la (13) si può prevedere σ_{N+1}^2 con $\sigma_{k,N}^2$.

Nella figura 50 è riportata la volatilità giornaliera di Fiat utilizzando varianze mobili non centrate con $k=1, 6, 11$. Naturalmente queste ultime sono più lisce e stabili. Nella figura 51 sono riportati i grafici della deviazione standard mobile giornaliera per Fiat, con $k=1, 6, 11$, che mettono in evidenza come la volatilità ha avuto un andamento prima decrescente per poi crescere marcatamente negli ultimi tempi.

Figura 50 – Volatilità e varianze mobili giornaliere con $k=1, 6, 11$ per Fiat



Come si può osservare dalla figura 51, al crescere di k , nel nostro caso $k=1, 6, 11$, rispettivamente, la serie della volatilità diviene sempre più regolare e, in questo caso, crescente e le autocorrelazioni di conseguenza più elevate (vedi figura 52).

Figura 51 – Deviazione standard e deviazione standard mobili giornaliere con $k=1, 6, 11$ per Fiat

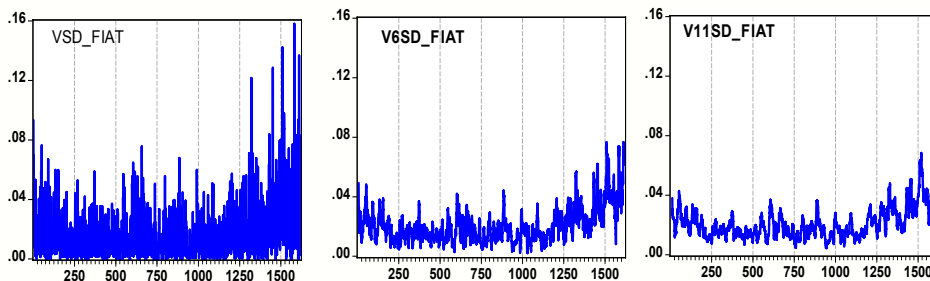
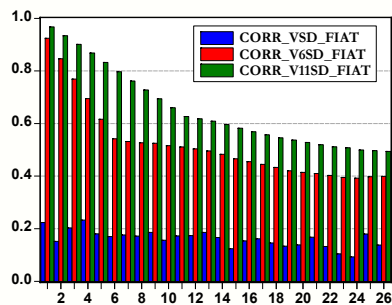


Figura 52 – Autocorrelazioni della deviazione standard e delle deviazioni standard mobili giornaliere con $k=1, 6, 11$ di Fiat



□

Sotto l'ipotesi, abbastanza ragionevole, che sia $r_t = \bar{r} + \sigma_t a_t$ con a_t errore standardizzato (media zero e varianza pari ad uno). E' possibile ottenere una stima di a_t tramite la seguente

$$a_{k,t} = \frac{r_t - \bar{r}}{\sigma_{k,t}} \quad (14)$$

L'analisi di $a_{k,t}$ fornisce ulteriori indicazioni sulla struttura e la stabilità della volatilità, inoltre la (14) verrà utilizzata nel paragrafo seguente per calcolare un indice per la misura della rischiosità di un attivo. Nella determinazione di $\sigma_{k,t}$, e quindi di $a_{k,t}$, è necessario individuare k . Di solito, nelle pratiche applicazioni è $10 \leq k \leq 25$.

Tabella 15 – Indici sintetici per $a_{k,t}$ con $k=1, 6, 11, 16, 21$ per i rendimenti Fiat

	v1	v6	v11	v16	v21
μ	4.63E-17	0.061078	0.049079	0.030080	0.026580
M_e	0.014925	0.021581	0.020611	0.020558	0.020845
$r_{(N)}$	5.835060	19.80473	14.60306	5.661493	5.448826
$r_{(1)}$	-6.452364	-6.272423	-5.126429	-4.295621	-4.780052
S	1.000000	1.399125	1.295267	1.120450	1.102667
γ_1	0.028447	2.236798	1.234825	0.174313	0.065601
γ_2	3.984774	27.28918	11.88711	1.870116	1.722586

Nella tabella 15 riportiamo gli indici descrittivi di $a_{k,t}$ per $k=1, 6, 11, 16, 21$ per i rendimenti giornalieri di Fiat. In questo caso, come si può osservare, i risultati migliori, in termini congiunti di asimmetria, curtosi, varianza e perdita di informazioni, si ottengono per $k=21$. Questa conclusione è confermata dagli istogrammi riportati nella figura 53.

Figura 53 – Istogrammi di $a_{k,t}$ con $k = 1, 6, 11, 16, 21$ riferiti ai rendimenti Fiat

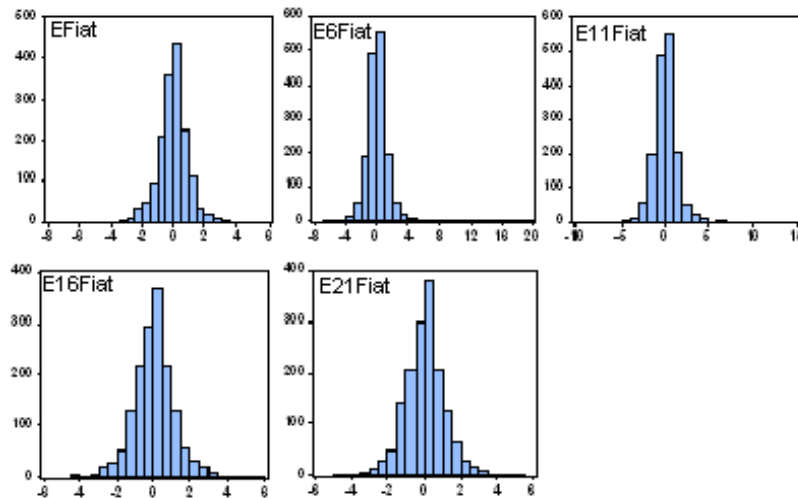


Tabella 16 – Indici descrittivi della stima della volatilità giornaliera di Fiat e dei suoi due regimi con $k=16$

	V16	V16_R1	V16_R2
μ	0.021616	0.022248	0.021049
M_e	0.018741	0.019158	0.018394
$r_{(1)}$	0.063234	0.060797	0.063234
$r_{(N)}$	0.004255	0.005489	0.004255
S	0.010162	0.010405	0.009912
γ_1	1.395536	1.303813	1.485601
γ_2	2.072629	1.547887	2.644572
N	1598	756	842

□

Anche per la volatilità è possibile derivare due *regimi* per analizzarla globalmente o localmente (nei singoli regimi) ed ottenere così informazioni

sulla struttura interna del fenomeno. In particolare, fissato k , da $\sigma_{k,t}$ si può ottenere una stima della volatilità locale (dei singoli regimi) tramite

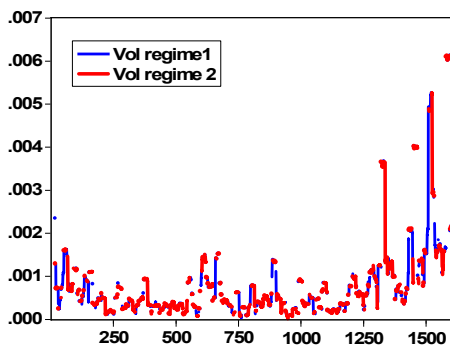
$$\sigma_{k,t} = \sigma_{1k,t} + \sigma_{2k,t}; \quad \sigma_{1k,t} = I_{t-1} \sigma_{k,t}; \quad \sigma_{2k,t} = (1 - I_{t-1}) \sigma_{k,t} \quad (15)$$

dove, come al solito, è

$$I_{t-1} = \begin{cases} 1 & \text{se } r_{t-1} < 0 \\ 0 & \text{se } r_{t-1} \geq 0 \end{cases}$$

A titolo di esempio, nella tabella 16 riportiamo gli indici descrittivi della stima della volatilità globale giornaliera di Fiat e quella locale dei suoi due regimi. Come si può osservare dalla tabella, il secondo regime (quello ottenuto con $r_{t-1} \geq 0$) è leggermente più numeroso ed ha una curtosi più elevata del primo (quello ottenuto con $r_{t-1} < 0$). Nella figura 54 è riportata la stima di $\sigma_{1k,t}$ e di $\sigma_{2k,t}$.

Figura 54 – *Stima della volatilità giornaliera dei due regimi di Fiat*



2.4 Il Valore a Rischio (VaR) degli attivi

Uno strumento molto utilizzato per valutare la rischiosità di un attivo o di un portafoglio di attivi è un indice sintetico che prende il nome di *Valore a Rischio* ed è comunemente indicato con la sigla VaR. Il VaR è un indice che *misura al tempo N la perdita massima, con probabilità α , $0 < \alpha < 1$, in cui si può*

incurrere al tempo $N+h$ nel detenere un attivo o un portafoglio di attivi, date le informazioni che si posseggono fino a quella data. Come si può notare, questo indice dipende:

- (a) dalle informazioni \mathfrak{F}_N che si hanno al tempo N . Di solito queste informazioni sono date dalla serie storica dei prezzi dell'attivo P_t , $t = 1, 2, \dots, N$;
- (b) dal livello di probabilità α scelto. Di solito è $\alpha = 0.05$ oppure $\alpha = 0.01$;
- (c) dall'orizzonte temporale $h = 1, 2, \dots, M$ che si vuole prevedere. Di solito si pone $M=10$.

Nel seguito, per sintetizzare che il Valore a Rischio dipende dalle informazioni (a) – (c) sopra definite, scriveremo $\text{VaR}_N(h, \alpha | \mathfrak{F}_N)$. Detto questo vediamo come operativamente è possibile calcolare il VaR. Nel paragrafo precedente abbiamo detto che ha senso supporre che sia

$$r_t = \bar{r} + \sigma_t a_t, \quad t = 2, 3, \dots, N$$

e che una stima di σ_t può essere ottenuta, fissato k , con $\sigma_{k,t}$ che permette di ricavare l'errore standardizzato

$$a_{k,t} = \frac{r_t - \bar{r}}{\sigma_{k,t}}, \quad t = k+2, k+3, \dots, N$$

e per $t = N+1$ ottenere

$$r_{N+1} = \bar{r} + \sigma_{k,N+1} a_{N+1} \tag{16}$$

Dalla distribuzione di $a_{k,t}$ possiamo derivare il percentile che “stacca” alla sua sinistra l' α per cento dei casi (le perdite) e che, sostituito nella (16), ci permette di ottenere una stima di $\text{VaR}_N(1, \alpha | \mathfrak{F}_N)$. Per calcolare il percentile di $a_{k,t}$ ordiniamo questi errori in senso non decrescente:

$$a_{k,(k+2)} \leq a_{k,(k+3)} \leq \dots \leq a_{k,(N)}$$

da questa serie ordinata, fissato α , per esempio al 5%, deriviamo il relativo percentile di posto $(N - k - 1)\alpha$. Di solito, $(N - k - 1)\alpha$ non è un numero

intero, ma è compreso nell'intervallo $[(N-k-1)\alpha] - [(N-k-1)\alpha] + 1$, dove con $[x]$ abbiamo indicato il massimo intero di x , ed una semplice stima del percentile è data da

$$a_{(\alpha)} = \frac{a_{k[(N-k-1)\alpha]} + a_{k[(N-k-1)\alpha] + 1}}{2}$$

che permette di ottenere una stima del VaR per $h = 1$ nell'ipotesi, ragionevole, che sia $\sigma_{k,N+1}^2 = \sigma_{k,N}^2$:

$$\text{VaR}_N(1, \alpha | \mathfrak{F}_N) = \bar{r} + \sigma_{k,N} a_{(\alpha)} \quad (17)$$

Se si vuole costruire il VaR per $h \geq 1$ tempi in avanti bisogna ricordare che i rendimenti sono una variabile di flusso e quindi quelli riferiti all'intervallo $[N, N+h]$ di lunghezza h sono dati da

$$r_{N+[h]} = r_{N+1} + \dots + r_{N+h}$$

D'altra parte, sappiamo che i rendimenti sono praticamente incorrelati e quindi risulta

$$\text{var}(r_{N+[h]}) = \text{var}(r_{N+1}) + \dots + \text{var}(r_{N+h}) = \sigma_{N+1}^2 + \dots + \sigma_{N+h}^2$$

che ci permette di ottenere una stima del VaR h passi in avanti:

$$\text{VaR}_N(h, \alpha | \mathfrak{F}_N) = h\bar{r} + a_{(\alpha)} \sqrt{\sigma_{k,N+1}^2 + \dots + \sigma_{k,N+h}^2}$$

Se si ipotizza che sia $\sigma_{k,N}^2 = \dots = \sigma_{k,N+j}^2$, $j=1, 2, \dots, h$, ipotesi abbastanza ragionevole, la precedente equazione si semplifica nella seguente

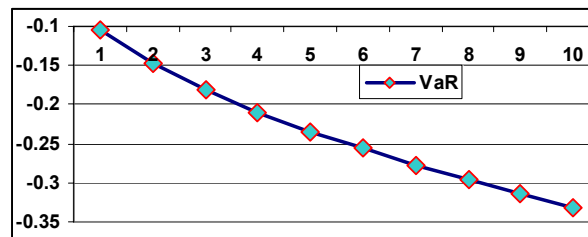
$$\text{VaR}_N(h, \alpha | \mathfrak{F}_N) = h\bar{r} + a_{(\alpha)} \sqrt{h\sigma_{k,N}^2} \quad (18)$$

La (18), che per $h=1$ si riduce alla (17), è operativa dato che tutte le quantità che la compongono sono calcolabili e può essere utilizzata nelle pratiche applicazioni.

Qui di seguito, a scopo esemplificativo, utilizziamo la (18) per calcolare il

valore a rischio, per i dieci giorni successivi a partire dall'ultima informazione del 10/03/09, per le quotazioni Fiat. Nelle analisi si è posto $k=16$ ed i risultati sono riportati nella figura 55.

Figura 55 – *Stima del VaR al 5% di Fiat per un orizzonte temporale massimo di $b=10$ giorni*



Come si può osservare dal grafico, la perdita massima prevista a 10 giorni con una probabilità $\alpha = 0.05$ per le quotazioni giornaliere Fiat va dal 10%, ad un giorno, e si avvicina al 35%, a dieci giorni lavorativi, a partire dal 10/03/09.

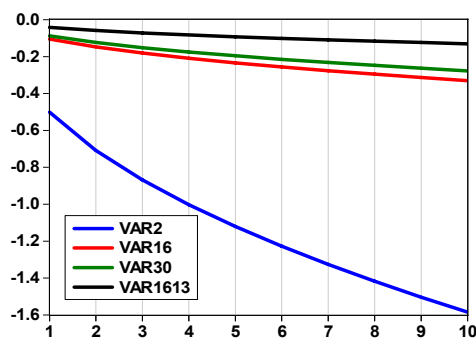
Uno dei punti critici per utilizzare la (18) è individuare k . Mostriamo empiricamente che se la scelta di k è oculata, la sua variazione ha una influenza limitata sul VaR. Naturalmente, al crescere di k la stima della volatilità $\sigma_{k,N}^2$ tende alla varianza dei rendimenti σ^2 ed inevitabilmente anche il VaR assumerà valori differenti al variare di k .

Calcoliamo il VaR per le quotazioni Fiat con diversi valori di k in modo da poter valutare l'influenza che k esercita su di esso. Nella tabella 17 sono riportati i valori del VaR, con $\alpha = 0.05$, rispettivamente per $k = 2, 16, 30, 1613$ e $h = 1, 2, \dots, 10$ giorni a partire dal 10/03/09. Osserviamo che per $k = 1613=N$ la varianza mobile si riduce al solo valore $\sigma_{N,N}^2 = \sigma^2$. Il confronto fra i VaR è riportato nella tabella 17 e nella figura 56. Come atteso, specie in periodi di elevata volatilità, i VaR differiscono fra di loro al variare di k , ma in modo sostanzialmente contenuto se k non è né troppo piccolo e vicino ad uno, né troppo grande e vicino ad N .

Per interpretare il VaR supponiamo che il 10/03/09 sia stato investito in titoli Fiat il capitale $C = 1.000.000$ di euro. Il VaR misura la perdita massima che si può avere, con una probabilità di 0.05, negli $h = 1, 2, \dots, 10$ giorni successivi e tenendo conto delle informazioni che si hanno sul fenomeno. Nei casi da noi illustrati si avrebbero i risultati della tabella 18.

Tabella 17 – *VaR per $\alpha = 0.05$ e $k=2, 16, 30, 1613$ per le quotazioni Fiat*

h	VaR2	VaR16	VaR30	VaR1613
1	-0.502	-0.105	-0.088	-0.042
2	-0.709	-0.148	-0.124	-0.059
3	-0.868	-0.181	-0.152	-0.072
4	-1.003	-0.209	-0.176	-0.083
5	-1.121	-0.234	-0.196	-0.093
6	-1.228	-0.256	-0.215	-0.102
7	-1.326	-0.277	-0.232	-0.110
8	-1.418	-0.296	-0.248	-0.117
9	-1.504	-0.314	-0.263	-0.124
10	-1.585	-0.331	-0.278	-0.131

Figura 56 – *VaR per $\alpha = 0.05$ e $k=2, 16, 30, 1613$ per le quotazioni Fiat*Tabella 18 – *Perdite massime, con probabilità del 5%, che si potrebbero avere, nei dieci giorni successivi al 10/03/09 investendo sul ti Fiat 1.000.000€*

h	VaR16	VaR30
1	-104881.4	-88052.5
2	-148173.9	-124374.3
3	-181393.4	-152244.9
4	-209398.8	-175741.0
5	-234072.0	-196441.4
6	-256378.3	-215156.1
7	-276891.0	-232366.0
8	-295983.8	-248384.5
9	-313916.2	-263429.5
10	-330877.0	-277659.4

□

In generale, la perdita massima con probabilità α che si potrebbe avere al tempo $N+h$ investendo al tempo N un capitale C su un dato attivo finanziario è data da

$$C \times \text{VaR}_N(h, \alpha | \mathfrak{F}_N)$$

Ricapitolando, per calcolare il VaR per un attivo è necessario:

1. rilevare i prezzi P_t , $t = 1, 2, \dots, N$. Per avere stime accettabili il valore di N deve essere sufficientemente elevato. L'evidenza empirica suggerisce $N \geq 600$;
2. calcolare i relativi rendimenti $r_t = \log(P_t / P_{t-1})$, $t = 2, 3, \dots, N$;
3. fissare k , di solito è $k = 22$ pari ai giorni di un mese borsistico, e calcolare media $\bar{r}_{k,t}$ e deviazione standard $\sigma_{k,t}$, $t = k+1, k+2, \dots, N$, mobile dei rendimenti;
4. calcolare gli errori standardizzati $a_{k,t}$ tramite la (17);
5. ordinare $a_{k,t}$ in senso non decrescente ottenendo la serie degli errori ordinati $a_{k,(t)}$, $t = k+1, k+2, \dots, N$;
6. fissato α , di solito al 5%, calcolare i percentili di $a_{k,(t)}$ di posto $[(N - k - 1)\alpha]$ e $[(N - k - 1)\alpha] + 1$ e calcolare $a_{(\alpha)}$ come media di questi;
7. utilizzare la (18) per ottenere una stima del VaR per ciascun orizzonte $h = 1, 2, \dots, 10$.

Per avere maggiori informazioni sulla struttura interna della volatilità di un attivo è possibile calcolare il VaR locale riferito a ciascuno dei due regimi in cui le osservazioni complessive possono essere distinte. Fissato k le deviazioni standard mobili locali riferiti a ciascun regime sono date da

$$\sigma_{1k,t} = I_{t-1} \sigma_{k,t}; \quad \sigma_{2k,t} = (1 - I_{t-1}) \sigma_{k,t}; \quad t = k+1, k+2, \dots, N$$

che ci permettono di calcolare gli errori standardizzati locali riferiti a ciascun regime :

$$a_{1k,t} = I_{t-1} \frac{r_t - \bar{r}}{\sigma_{k,t}} = I_{t-1} a_{k,t}; \quad t = k+1, k+2, \dots, N$$

$$a_{2k,t} = (1 - I_{t-1}) \frac{r_t - \bar{r}}{\sigma_{k,t}} = (1 - I_{t-1}) a_{k,t}; \quad t = k+2, k+3, \dots, N$$

che, fissato α , ci permettono di calcolare i percentili locali di ciascun regime

$$a_{1(\alpha)} = \frac{a_{1[(N_1)\alpha]} + a_{1[(N_1)\alpha]+1}}{2}; \quad a_{2(\alpha)} = \frac{a_{2[(N_2)\alpha]} + a_{2[(N_2)\alpha]+1}}{2}$$

dove è $N_1 + N_2 = N - k - 1$ e N_i è il numero delle quotazioni che appartiene a ciascun regime. Questo ci permette di calcolare il VaR locale per ciascun regime:

$$\text{VaR}_{1N}(h; \alpha, \mathfrak{F}_N) = h\bar{r}_1 + a_{1(\alpha)} \sqrt{h} \sigma_{1k,N}$$

$$\text{VaR}_{2N}(h; \alpha, \mathfrak{F}_N) = h\bar{r}_2 + a_{2(\alpha)} \sqrt{h} \sigma_{2k,N}$$

con $\bar{r}_1 = \text{media}(r_{1t})$ e $\bar{r}_2 = \text{media}(r_{2t})$. Ed, in fine, calcolare quello globale come media ponderata dei VaR di ciascun regime:

$$\text{VaRp}_N(h; \alpha, \mathfrak{F}_N) = p \text{VaR}_{1N}(h; \alpha, \mathfrak{F}_N) + (1 - p) \text{VaR}_{2N}(h; \alpha, \mathfrak{F}_N) \quad (19)$$

La (19) è stata applicata alle quotazioni Fiat per $k = 22$ ed i risultati sono stati confrontati con quelli ottenuti utilizzando la (18). In questo caso si ha:

$$N_1 = 752; \quad N_2 = 840; \quad N - k - 1 = 1592; \quad p = 763/1613 = 0.473; \quad 1 - p = 0.527$$

$$\bar{r}_1 = -0.000702; \quad \bar{r}_2 = -0.000171; \quad \sigma_{2k,N} = 0.0521; \quad \sigma_{1k,N} = 0.0509$$

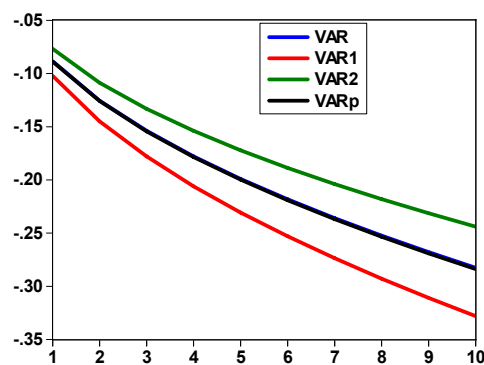
$$a_{(\alpha)} = -1.692; \quad a_{1(\alpha)} = -1.948; \quad a_{2(\alpha)} = -1.469$$

Tabella 19 – VaR con $k=22$ ed $\alpha = 0.05$ per le quotazioni Fiat calcolati con la (18) e la (19) e per ciascuno dei due regimi

h	VaR con (18)	VaR1	VaR2	VaRp con (19)
1	-0.089	-0.102	-0.077	-0.089
2	-0.125	-0.145	-0.109	-0.126
3	-0.154	-0.178	-0.133	-0.154
4	-0.178	-0.206	-0.154	-0.179
5	-0.199	-0.231	-0.172	-0.200
6	-0.218	-0.253	-0.189	-0.219
7	-0.236	-0.274	-0.204	-0.237
8	-0.252	-0.293	-0.218	-0.253
9	-0.268	-0.311	-0.231	-0.269
10	-0.283	-0.328	-0.244	-0.284

Questi dati ci permettono di calcolare il VaR locale riferito a ciascun regime e quello globale. I risultati finali sono riportati nella tabella 19. Come si può osservare il VaR ottenuto con la (19) è praticamente uguale a quello ottenuto con la (18), ma l'analisi per regimi ci fornisce un'ulteriore informazione: di quanto varia la rischiosità di un attivo se il giorno precedente la quotazione ha fatto registrare un segno positivo o negativo. Nel caso delle quotazioni Fiat il primo regime è più rischioso del secondo e la differenza va da un minimo di 2.5 punti per $h=1$, ad un massimo di 8.4 punti per $h=10$. Gli stessi risultati sono riportati, per maggiore completezza illustrativa, anche nella figura 57.

Figura 57 – VaR con $k=22$ ed $\alpha = 0.05$ per le quotazioni Fiat calcolati con la (18) e la (19) e per ciascuno dei due regimi



E' da sottolineare che quanto abbiamo visto, a proposito del

comportamento dei singoli regimi, nel caso delle quotazioni Fiat (un regime è molto più rischioso dell'altro) non è generalizzabile ad altri attivi o portafogli di attivi. Può succedere, anche se più raramente, che fra i due regimi di un dato attivo non siano riscontrabili, in termini di rischiosità, differenze sostanziali ed anche questa è una informazione utile per gli operatori dei mercati di riferimento.

Nelle applicazioni precedenti il VaR è stato calcolato con $\alpha = 0.05$, ma può essere calcolato per valori inferiori ed in particolare per $\alpha = 0.01$. E' da sottolineare che per valori di α inferiori a 0.05 la stima del percentile dei residui standardizzati con il metodo sopra riportato è scadente e la procedura fornisce, di solito, risultati modesti.

2.5 Una applicazione a dati reali

Gli strumenti che sono stati illustrati nelle pagine precedenti verranno applicate a cinque serie di quotazioni sintetiche: Nasdaq financial (Nasdaq_f), Nasdaq Composite (Nasdaq), Mibtel, Dax, Cac. Le prime due serie riguardano il mercato USA e precisamente Nasdaq Composite è un indice sintetico rappresentativo dell'andamento dei prezzi dei titoli quotati nel mercato NASDAQ (*National Association of Securities Dealers Automated Quotation*) prevalentemente riconducibili a società operanti nel settore dell'information technology, mentre l'indice Nasdaq_f è riferito ai soli prezzi delle azioni di società finanziarie. Mibtel è l'indice sintetico telematico della borsa italiana, Dax quello della borsa tedesca e Cac di quella francese. Le rilevazioni riguardano le chiusure aggiustate e sono riferite al periodo che va dal 06/01/03 al 08/05/09. Per la presenza delle festività nazionali, le rilevazioni osservate di ciascuna serie non sono, in generale, cinque a settimana (dal lunedì al venerdì), inoltre queste festività non cadono negli stessi giorni nei vari paesi. Per rendere confrontabili le cinque serie è stato necessario, in primo luogo, far sì che in ogni settimana fossero presenti esattamente cinque osservazioni. Per ottenere questo risultato si è proceduto, per ogni serie, ad individuare i giorni di chiusura infrasettimanali delle borse e sostituire ciascun dato mancante con quello del giorno immediatamente precedente. Inoltre, le cinque serie hanno unità di misura differenti e per eliminare questo ulteriore inconveniente si è proceduto a trasformare le serie originarie nelle serie degli indici a base fissa pari a 100 nel primo giorno di osservazione del 06/01/03. Nel seguito tutte le applicazioni verranno fatte per il periodo di osservazione sopra indicato e sulle serie così trasformate.

Nella figura 58 sono riportate le cinque serie così ottenute dalle quali è evidente l'inversione di tendenza e l'inizio della grande crisi nel maggio – luglio del 2007. L'indice Mibtel ha iniziato la sua discesa qualche mese prima degli altri ed in modo più marcato, con l'eccezione di Nasdaq_f che è cresciuto di meno di Mibtel ed ha avuto un crollo simile a Mibtel.

Figura 58 – Le cinque serie analizzate

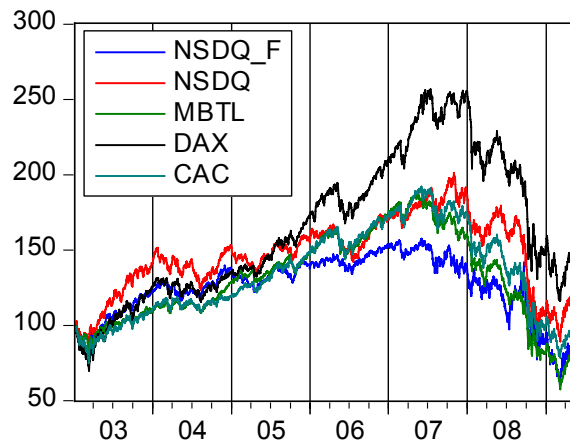


Figura 59 – Box-plot delle cinque serie analizzate

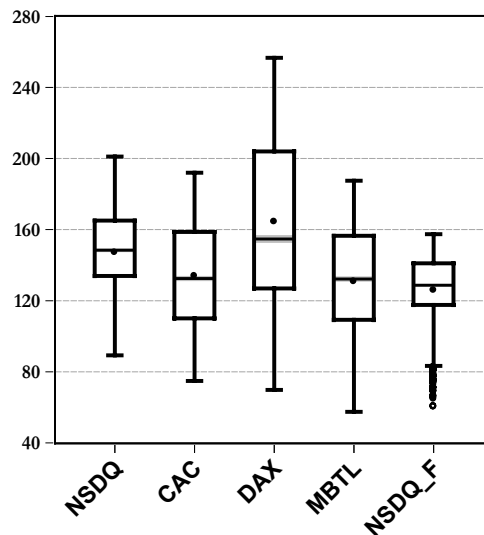


Tabella 20 – *Indici sintetici descrittivi delle cinque serie*

	NSDQ_F	NSDQ	MBTL	DAX	CAC
μ	125.9334	147.1677	130.8771	164.4496	133.8791
M_e	128.7750	148.4150	132.2760	154.7710	132.4730
$r_{(N)}$	157.4680	201.1590	187.5510	256.7330	192.1380
$r_{(1)}$	60.14700	89.25800	57.54500	69.77500	74.85500
S	20.07380	25.01800	30.08397	47.30554	29.77809
γ_1	-0.871132	-0.357391	-0.066138	0.259859	0.117592
γ_2	0.21273	-0.48158	-0.88481	-1.00487	-1.15785
N	1655	1655	1655	1655	1655

L'andamento generale di Mibtel è simile a Nasdaq e Cac, ma differente da Dax e Nasdaq_f come è messo in rilievo anche dalla figura 59 in cui sono riportati i box-plot paralleli delle cinque serie in cui si rileva, fra l'altro, come il Mibtel assume valori minimi e massimi inferiori a quelli degli altri paesi.

A conferma di quanto detto riportiamo nella tabella 20 gli indici descrittivi delle cinque serie.

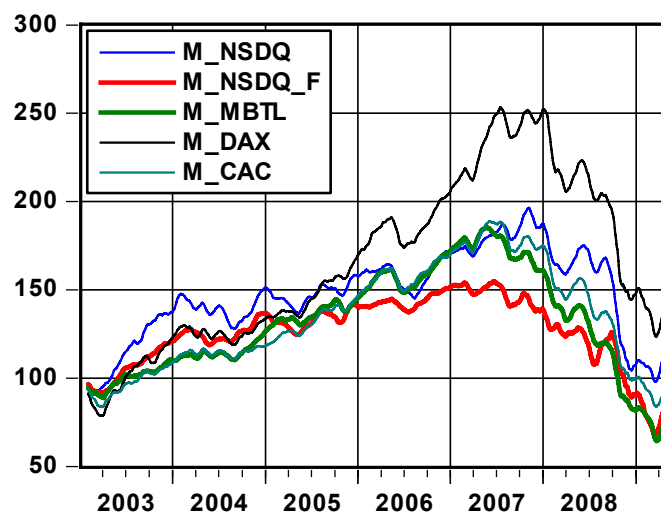
Figura 60 – *Medie mobili non centrate, con $k=22$, delle cinque serie*

Figura 61 – Errori stimati come differenza fra serie osservata e sua media mobile

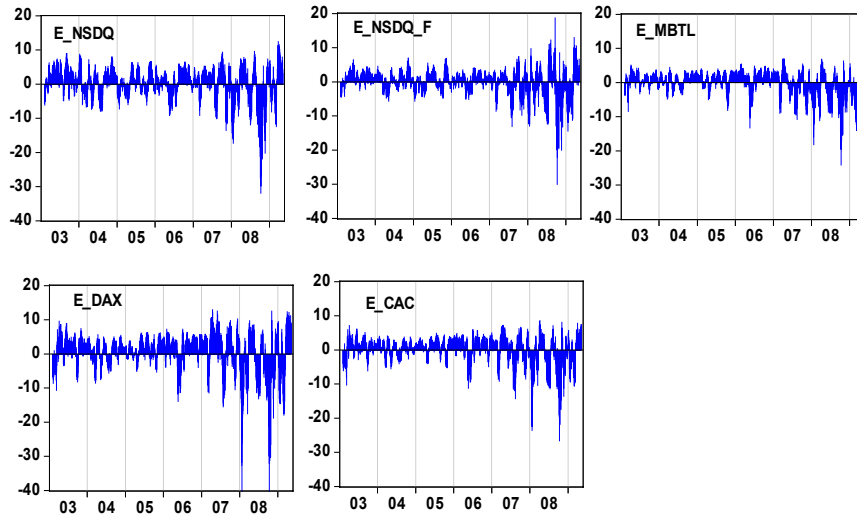
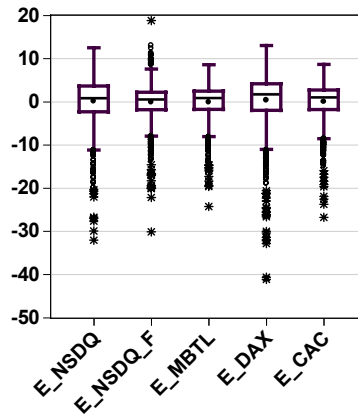


Figura 62 – Box-plot degli errori stimati



Di ciascuna serie calcoliamo la media mobile asimmetrica con $k=22$, approssimativamente pari ai giorni di contrattazione borsistica presenti in un mese. Le serie così ottenute sono riportate nella figura 60 in cui sono più evidenti le considerazioni fatte in precedenza. Nella figura le medie mobili di Mibtel e Nasdaq_f sono state rappresentate con linee più marcate delle altre al fine di metterne in evidenza le similitudini. Dal grafico si intravede anche una inversione di tendenza nel secondo trimestre 2009 che, per la natura sostanzialmente imprevedibile di questi fenomeni, non possiamo identificare

come persistente (fine del ciclo negativo) o transitoria. Dalla figura 61 si nota come gli errori negativi siano molto più marcati di quelli positivi, questo è ben messo in evidenza dai box-plot paralleli riportati nella figura 62 dove i valori eccezionali negativi sono nettamente prevalenti su quelli positivi. In particolare questo fenomeno si osserva in modo più netto dalla metà del 2007 in poi. Nel complesso la distribuzione degli errori presenta una marcata asimmetria negativa. I rendimenti dei cinque mercati analizzati sono riportati nella figura 63 mentre nella tabella 21 sono riportati gli indici descrittivi.

Figura 63 – Rendimenti delle cinque serie analizzate

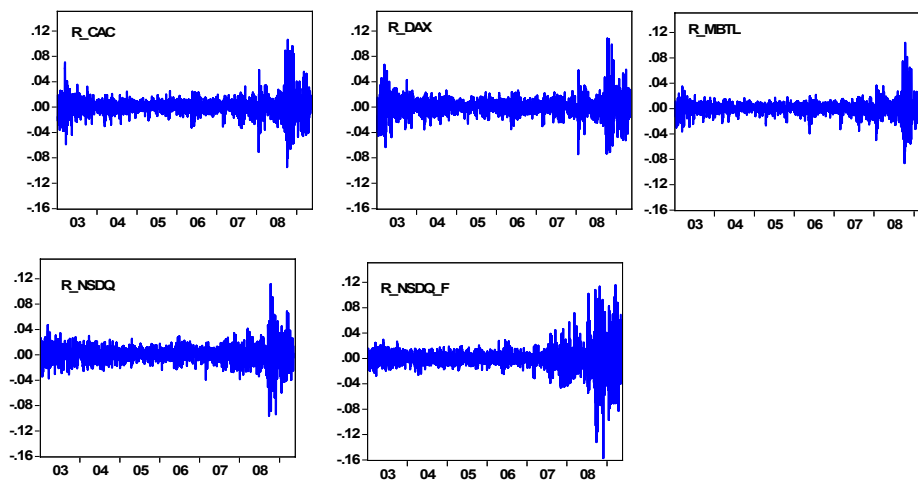
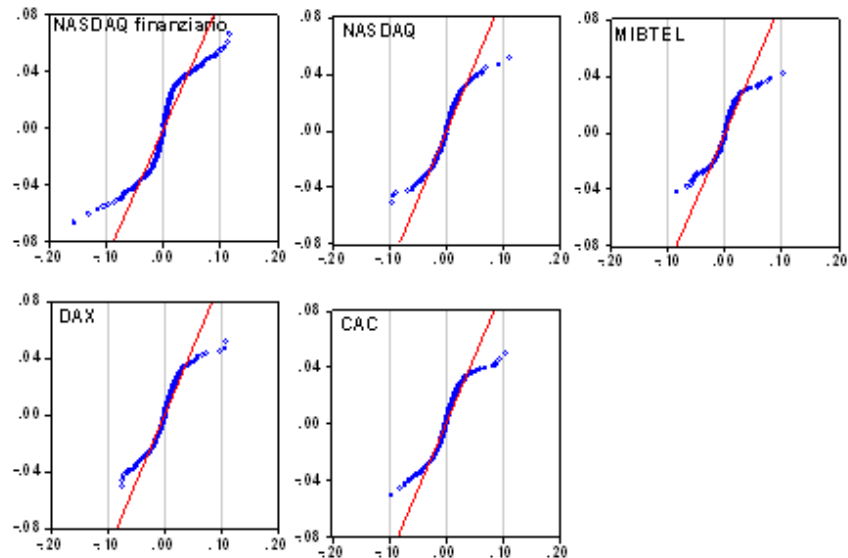


Tabella 21 – Indici descrittivi dei rendimenti delle cinque serie analizzate

	R_NSDQ	R_NSDQ_F	R_MBTL	R_DAX	R_CAC
μ	0.000122	-5.99E-05	-7.89E-05	0.000267	1.90E-05
M_e	0.000353	0.000000	0.000486	0.000687	0.000229
$r_{(N)}$	0.111592	0.115348	0.103707	0.107974	0.105940
$r_{(1)}$	-0.095877	-0.157203	-0.086032	-0.074334	-0.094714
S	0.014921	0.019375	0.012266	0.015005	0.014629
γ_1	-0.118822	-0.153793	-0.109449	0.109910	0.101749
γ_2	7.21697	11.41056	10.81905	7.50996	8.75454
N	1654	1654	1654	1654	1654

Figura 64 – Q-Qplot dei rendimenti delle cinque serie analizzate



Come atteso, i rendimenti osservati per le cinque serie hanno media prossima allo zero, debole asimmetria e marcata curtosi. In ogni caso, sono molto lontani dalla normalità come si può osservare dai Q-Qplot riportati nella figura 64. Le autocorrelazioni sui livelli dei rendimenti, come atteso e come evidente dalla figura 65, sono praticamente inesistenti.

Figura 65 – Autocorrelazioni dei rendimenti delle cinque serie analizzate

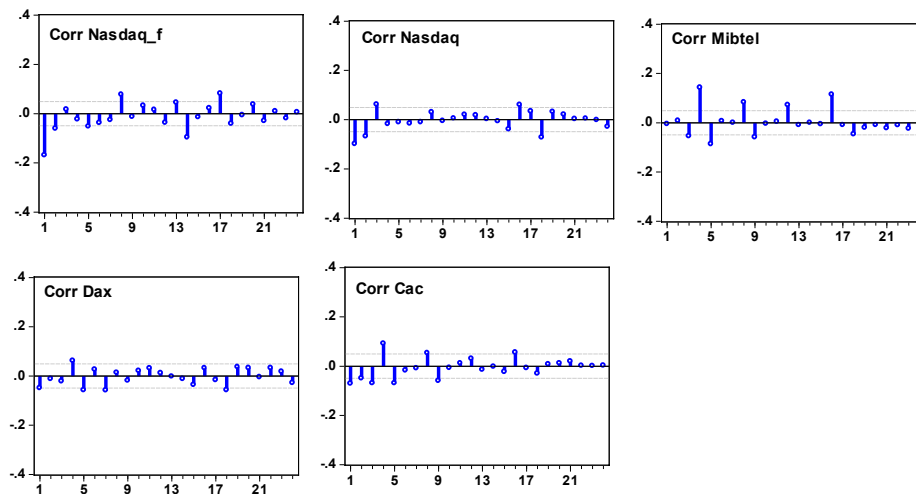


Figura 66 – Rendimenti al quadrato dei cinque attivi analizzati

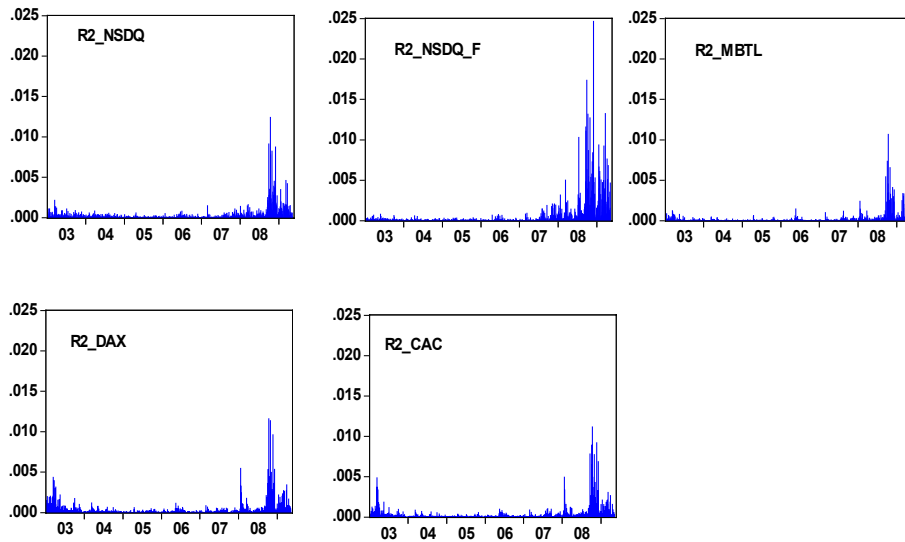
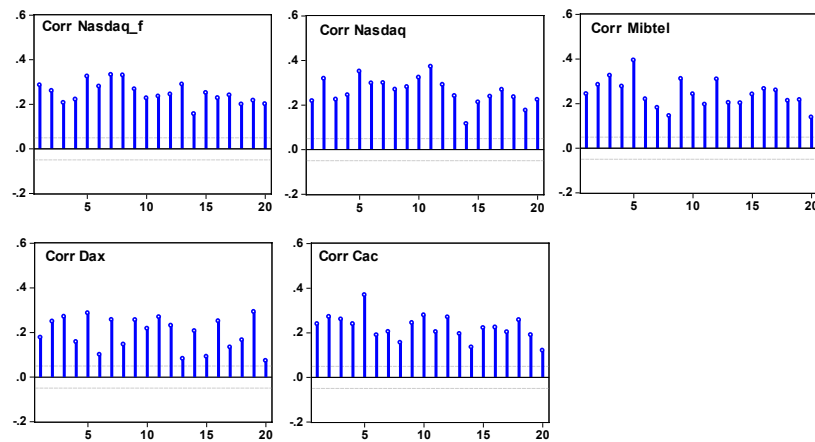


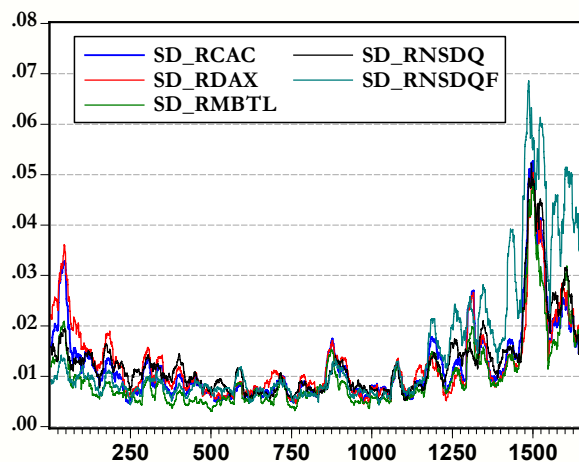
Figura 67 – Prime 50 autocorrelazioni dei rendimenti al quadrato delle cinque serie



Per mettere in evidenza la volatilità degli attivi analizzati, nella figura 66 sono riportati i rendimenti al quadrato delle cinque serie dalle quali si possono osservare le forti oscillazioni dell'ultimo periodo in corrispondenza della crisi e come queste siano nettamente più marcate nell'indice finanziario Nasdaq_f. Osserviamo ancora come i rendimenti al quadrato oscillino a grappoli: a periodi di più forti turbolenze seguono periodi di maggiore calma. Nella figura 67 riportiamo le prime 20 autocorrelazioni dei rendimenti al quadrato delle

cinque serie e, come si può osservare, presentano valori marcati e decadono a zero lentamente. Ciò implica che nei rendimenti esistono forti legami non lineari ed in particolare quadratici. Con $k=22$ si è proceduto a calcolare le deviazioni standard mobili di ciascuna delle cinque serie per utilizzarle come stima delle rispettive volatilità. Nella figura 68 sono riportate le deviazioni standard mobili $\sigma_{22,t}$ delle serie.

Figura 68 – Deviazioni standard mobile con $k = 22$ dei rendimenti delle serie



Come abbiamo già osservato, la volatilità si presenta a grappoli ed ha avuto una forte impennata a partire dalla seconda metà del 2007 per effetto della crisi. Le oscillazioni, in questo ultimo periodo, sono nettamente più marcate per l'indice finanziario (Nasdaq_f). Dalle serie delle deviazioni standard mobili rileviamo il valore di posto $t=N=1654$ e lo riportiamo qui di seguito.

Cac	Dax	Mibtel	Nasdaq	Nasdaq_f
0.0156	0.0174	0.0162	0.0176	0.0425

Per ciascuna serie dei rendimenti calcoliamo quella dei residui standardizzati

$$a_{22,t} = \frac{r_t - \bar{r}}{\sigma_{22,t}}, \quad t = 24, 25, \dots, 1654$$

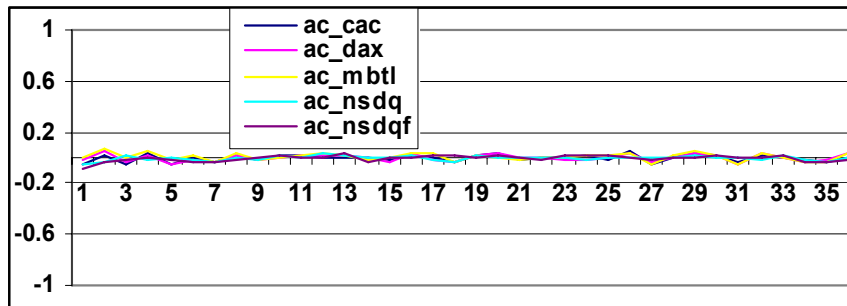
Ordinando questi errori standardizzati in senso non decrescente e tenuto conto che $(1654-24)*0.05=81.5$ si è provveduto a calcolare il relativo percentile

$a_{(0.05)}$ riportato qui di seguito

Cac	Dax	Mibtel	Nasdaq	Nasdaq_f
-1.811	-1.854	-1.834	-1.723	-1.735

Per mostrare che la procedura utilizzata per catturare la volatilità degli attivi analizzati è ragionevolmente efficace, nei prossimi due grafici riportiamo le autocorrelazioni (figura 69) e quelle dei quadrati (figura 70) di $a_{22,t}$ e come si può osservare risultano tutte prossime a zero e del tutto irrilevanti.

Figura 69 – Autocorrelazioni degli errori standardizzati $a_{22,t}$



Le analisi fatte permettono di calcolare il relativo VaR, al livello di rischio del 5%, per un orizzonte temporale di 10 giorni. I risultati del VaR così ottenuti sono riportati nella tabella 22. Da questa tabella rileviamo che i mercati dei maggiori paesi occidentali hanno praticamente la stessa rischio, e questo per la forte integrazione sia economica che normativa che esiste fra di loro, mentre il mercato finanziario americano risulta circa due volte e mezzo più rischioso degli altri mercati qui considerati.

Figura 70 – Autocorrelazioni degli errori standardizzati $a_{22,t}$ al quadrato

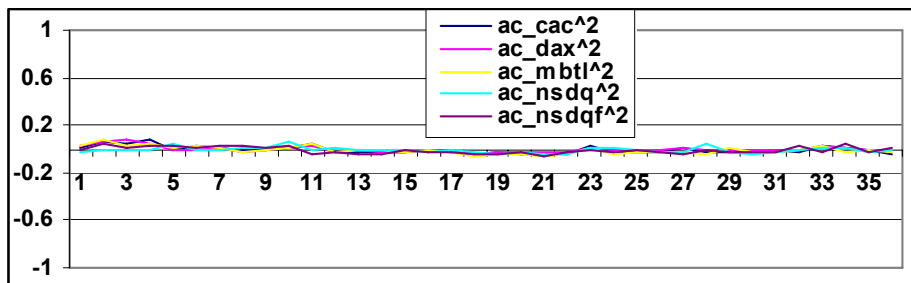
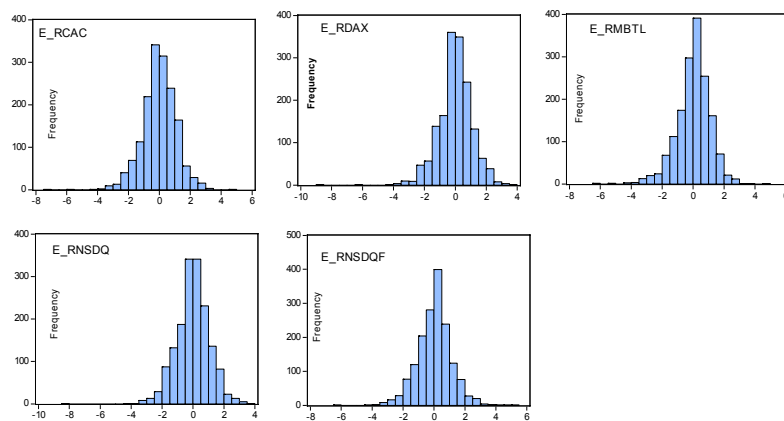


Tabella 22 – VaR con $k=22$ per ciascuno degli attivi analizzati

h	VaR_cac	VaR_dax	VaR_mbtl	VaR_nsdq	VaR_nsdq_f
1	-0.029	-0.032	-0.030	-0.030	-0.074
2	-0.041	-0.045	-0.042	-0.043	-0.104
3	-0.050	-0.055	-0.052	-0.052	-0.128
4	-0.057	-0.063	-0.060	-0.060	-0.148
5	-0.064	-0.071	-0.067	-0.067	-0.165
6	-0.070	-0.077	-0.074	-0.074	-0.181
7	-0.076	-0.084	-0.079	-0.080	-0.196
8	-0.081	-0.089	-0.085	-0.085	-0.209
9	-0.086	-0.094	-0.090	-0.090	-0.222
10	-0.091	-0.099	-0.095	-0.095	-0.234

Figura 71 – Istogrammi di frequenza dei residui standardizzati $a_{22,t}$ 

Nella figura 71 sono riportati gli istogrammi di frequenza degli errori standardizzati $a_{22,t}$ che mostrano una leggera asimmetria ed una moderata curtosi.

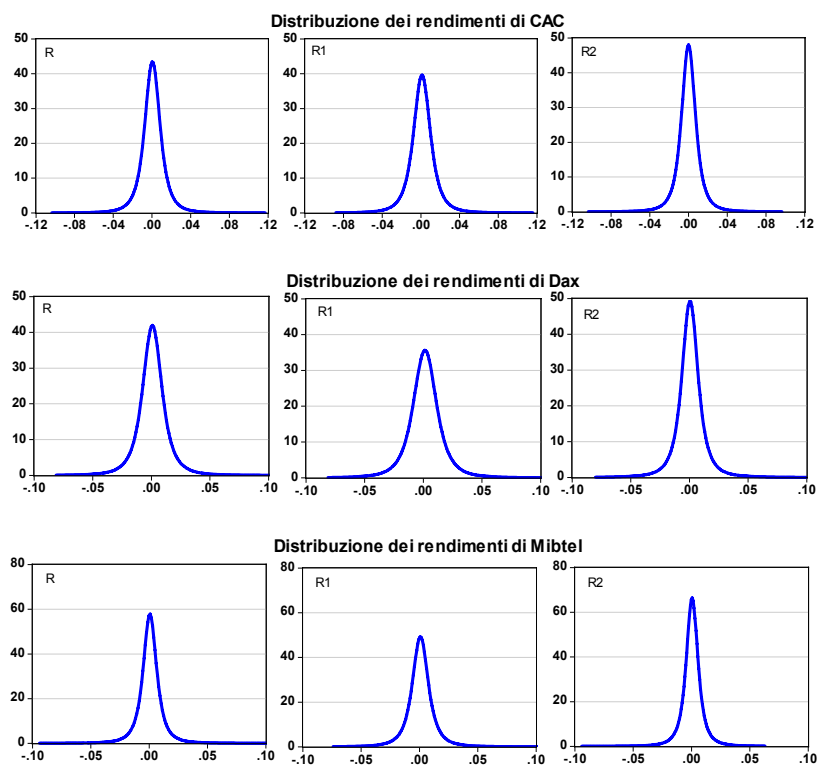
Utilizziamo ora la tecnica dei regimi per ottenere maggiori informazioni sulla struttura interna dei cinque mercati considerati. Dato che l'analisi è condotta su indici sintetici che misurano l'evoluzione media di interi mercati, è da attendersi che le differenze, in termini di VaR e quindi di rischiosità, fra i due regimi siano molto più limitate rispetto alle stesse analisi fatte su singoli attivi (vedi il caso Fiat). Inoltre, non è necessariamente detto che il primo regime sia più rischioso del secondo dato che le attese ribassiste sono, a priori, simili a quelle rialziste. Nella tabella 23 riportiamo, con $k=22$ ed $\alpha = 0.05$, gli

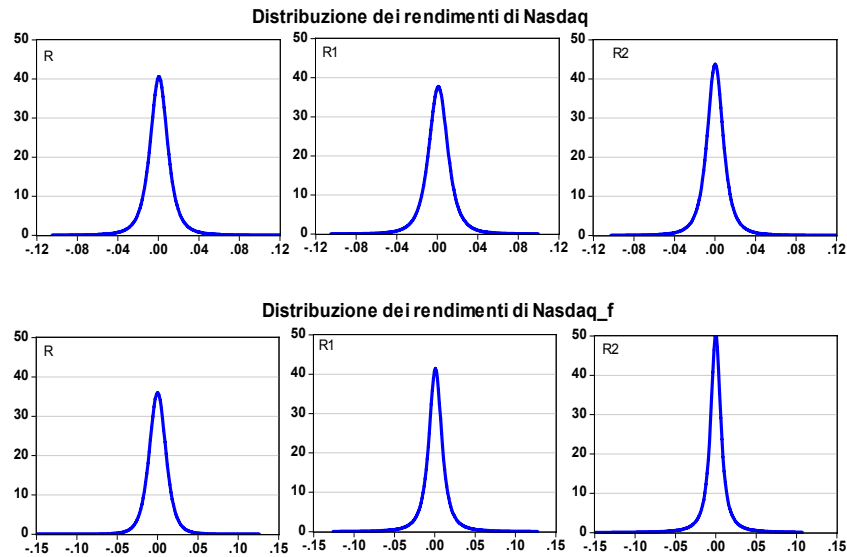
elementi idonei per calcolare la rischiosità locale riferita a ciascun regime per ciascuno dei cinque mercati presi in considerazione.

Tabella 23 – *Medie, standard deviation, numerosità e percentili dei singoli regimi nei cinque mercati analizzati*

	Cac	Dax	Mbtl	Nsdq	Nsdq_f
\bar{r}_1	-0.00057	-0.00011	-0.00009	-0.00044	-0.00200
\bar{r}_2	0.00069	0.00075	-0.00005	0.00080	0.00226
S_1	0.01597	0.01742	0.01620	0.01822	0.04128
S_2	0.01623	0.01782	0.01665	0.01806	0.04353
N_1	782	724	734	734	743
N_2	871	909	919	904	890
$a_{1(\alpha)}$	-1.839	-1.921	-1.933	-1.706	-1.655
$a_{2(\alpha)}$	-1.765	-1.789	-1.746	-1.743	-1.782

Figura 72 – *Distribuzione dei rendimenti delle cinque serie e dei rispettivi regimi*





Come si può osservare dalla figura 72, il primo regime ha, in tutte le serie analizzate, un massimo inferiore a quello del secondo e questo ne giustifica l'esistenza.

Tabella 24 – VaR nei singoli regimi dei cinque mercati analizzati

h	Cac		Dax		Mibtel		Nasdaq		Nasdaq_f	
	VaR ₂	VaR ₁	VaR ₂	VaR ₁	VaR ₂	VaR ₁	VaR ₂	VaR ₁	VaR ₂	VaR ₁
1	-0.029	-0.028	-0.031	-0.033	-0.028	-0.031	-0.031	-0.029	-0.078	-0.068
2	-0.041	-0.040	-0.044	-0.046	-0.040	-0.045	-0.044	-0.041	-0.111	-0.095
3	-0.050	-0.048	-0.054	-0.056	-0.049	-0.055	-0.055	-0.050	-0.137	-0.115
4	-0.058	-0.056	-0.063	-0.064	-0.057	-0.063	-0.063	-0.057	-0.160	-0.132
5	-0.065	-0.062	-0.070	-0.071	-0.064	-0.071	-0.071	-0.063	-0.179	-0.146
6	-0.072	-0.067	-0.077	-0.077	-0.070	-0.077	-0.078	-0.069	-0.198	-0.159
7	-0.078	-0.072	-0.083	-0.083	-0.076	-0.084	-0.084	-0.074	-0.214	-0.170
8	-0.084	-0.077	-0.089	-0.089	-0.081	-0.089	-0.091	-0.079	-0.230	-0.181
9	-0.089	-0.081	-0.094	-0.094	-0.086	-0.095	-0.096	-0.083	-0.245	-0.191
10	-0.094	-0.085	-0.100	-0.098	-0.091	-0.100	-0.102	-0.087	-0.260	-0.200

Nella tabella 24 sono riportati i VaR locali riferita ai due regimi, per dieci giorni a partire dal 09/05/09, di ciascuna delle cinque serie prese in considerazione, con $k=22$ e $\alpha = 0.05$. In fine, nella tabella 25 è riportato il VaR globale come

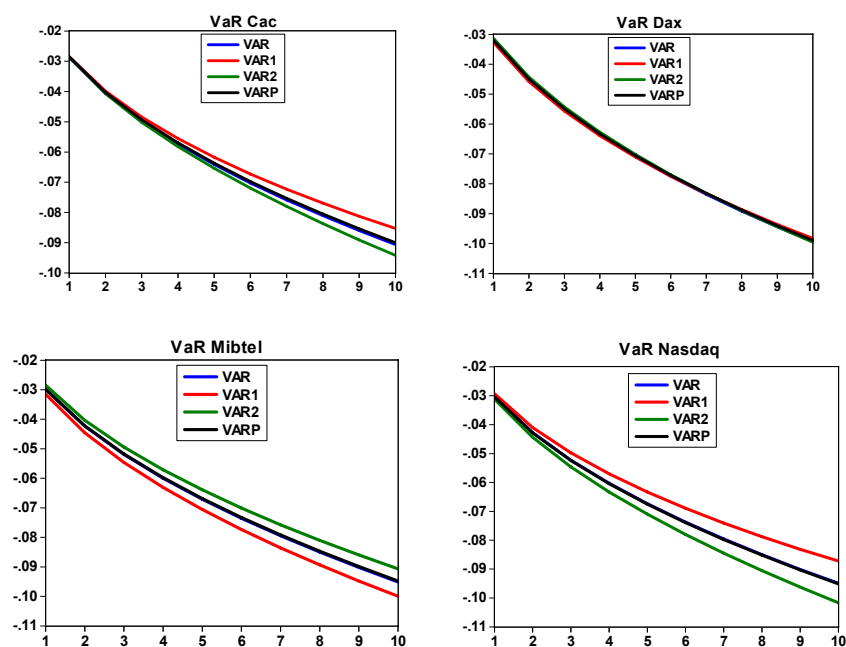
media ponderata di quelli dei singoli regimi e, come si può osservare, sono molto simili a quelli della tabella 22.

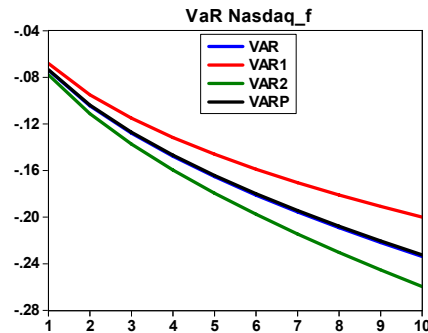
Tabella 25 – *VaR dei cinque mercati come media ponderata dei VaR dei regimi*

h	VaR Cac	VaR Dax	VaR Mibtel	VaR Nsdq	VaR Nsdq_f
1	-0.029	-0.032	-0.030	-0.030	-0.073
2	-0.040	-0.045	-0.042	-0.043	-0.104
3	-0.049	-0.055	-0.052	-0.052	-0.127
4	-0.057	-0.063	-0.060	-0.060	-0.147
5	-0.064	-0.071	-0.067	-0.068	-0.164
6	-0.070	-0.077	-0.073	-0.074	-0.180
7	-0.075	-0.083	-0.079	-0.080	-0.194
8	-0.080	-0.089	-0.085	-0.085	-0.208
9	-0.085	-0.094	-0.090	-0.090	-0.220
10	-0.090	-0.099	-0.095	-0.095	-0.232

Inoltre, come atteso, le differenze fra i singoli regimi sono molto contenute. Nella figura 73 sono riportati i VaR delle cinque serie e quelle dei propri regimi.

Figura 73 – *VaR dei cinque attivi e dei loro regimi*





La differenza fra i VaR locali riferiti ai due regimi può essere interpretata come un indice di affidabilità dell'attivo, del portafoglio o del mercato esaminato, nel senso che maggiore è tale differenza meno affidabile, solido è il relativo attivo, portafoglio o mercato. Da questo punto di vista, dalla tabella 25 e dalla figura 73 osserviamo che il mercato tedesco, descritto dal Dax, è quello più affidabile dato che i VaR dei due regimi sono praticamente coincidenti fra di loro e con i VaR globali calcolati con la (18) e la (19), mentre quello meno affidabile è, come atteso, il mercato finanziario descritto dall'indice Nasdaq_f. Inoltre, il VaR riferito al mercato finanziario americano (Nadasq_f) ha una rischiosità più che doppia rispetto agli altri mercati la cui rischiosità varia da un minimo di circa il 3%, ad un giorno, fino ad un massimo del 10% del capitale investito a dieci giorni. Il VaR calcolato con la (18) è praticamente indistinguibile da quello calcolato con la (19). Osserviamo infine che fra i regimi esiste, con l'esclusione del mercato finanziario descritto dal Nasdaq_f, una limitata differenziazione sintomo di buona affidabilità di questi mercati.

Qui di seguito riportiamo il listato che permette di calcolare il VaR di un attivo finanziario con EViews 6.

'Stima del VaR dell'attivo e quello dei regimi

```

smpl 1 N
genr r=d(log(x))
for !i=3 to N
  smpl !i !i
  if r(!i-1)<0 then genr i1=1
    genr r1=r
    genr it=1
  else genr i2=1
    genr r2=r
    genr it=0

```



```

endif
next
  smpl 1 N
  scalar p=@mean(it)
  genr rm=@mav(r, 22)
  genr vm=@mvar(r, 22)
  smpl 23 N
  genr rm1=rm*i1
  genr rm2=rm*i2
  genr vm1=vm*i1
  genr vm2=vm*i2
  smpl 4 N
  genr z=(r-@mean(r))/sqr(vm)
  genr z1=z*i1
  genr z2=z*i2
  smpl 1 N
  genr h=@obsnum
  scalar zq=@quantile(z, 0.05)
  scalar zq1=@quantile(z1, 0.05)
  scalar zq2=@quantile(z2, 0.05)
  scalar m_r=@mean(r)
  scalar m_r1=@mean(r1)
  scalar m_r2=@mean(r2)
  smpl 1 10
      genr VaR=h*m_r+zq*sqr(@elem(vm, N)*h)
      genr VaR1=h*m_r1+zq1*sqr(@elem(vm, N)*h)
      genr VaR2=h*m_r2+zq2*sqr(@elem(vm, N)*h)
      genr VaRp=p*VaR1+(1-p)*VaR2
  line VaR VaR1 VaR2 VaRp
  smpl 1 N
  group serie r r1 r2 rm rm1 rm2 z z1 z2
  serie.stats(i)
  delete h m_r m_r1 m_r2

```

Per far partire il programma è necessario:

- che sia attivo il workfile con la serie (qui indicata con x nella terza riga del listato) dei prezzi dell'attivo di cui si vuole calcolare il VaR;
- che sia sostituito ad N il numero delle osservazioni di cui la serie x che si vuole analizzare è composta;
- che il renga nel workfile sia almeno pari ad N+10 per permettere di calcolare le 10 previsioni del VaR.

Nel listato il VaR globale è stato calcolato ponendo $\alpha = 0.05$, se si vuole calcolarlo per $\alpha = 0.01$ basta sostituire questo valore nelle istruzioni che

definiscono gli scalari z_q , z_{q1} , z_{q2} corrispondenti, rispettivamente, $a_{(0.05)}$, $a_{1(0.05)}$, $a_{2(0.05)}$.

Output di interesse:

- il grafico contenente: il VaR globale dell'attivo, quello locale dei suoi due regimi VaR1 e VaR2 e quello ottenuto come media ponderata di questi VaRp;
- le serie VaR, VaR1, VaR2, VaRp che contengono i valori numerici dei relativi Valori a Rischio;
- le statistiche descrittive dei rendimenti della serie e di quelle dei regimi: r , $r1$, $r2$; della media mobile dei rendimenti e dei regimi: rm , $rm1$, $rm2$; della standardizzata dei rendimenti e dei regimi: z , $z1$, $z2$, che permettono di avere informazioni su aspetti strutturali dei rendimenti;
- la serie della volatilità mobile globale dell'attivo e quella locale dei suoi regimi: vm , $vm1$, $vm2$, che permette di valutare l'andamento della volatilità nell'intervallo di tempo analizzato;
- lo scalare p che definisce la percentuale di rendimenti che cade nel primo regime.

Le serie sopra discusse possono essere rappresentate graficamente nei modi ritenuti più idonei e possono essere utilizzate per effettuare altre eventuali elaborazioni.

CAPITOLO III

Variabili casuali e processi stocastici

Nel capitolo precedente abbiamo studiato prezzi, rendimenti e volatilità utilizzando esclusivamente strumenti di tipo *descrittivo – esplorativo*. D'altro lato, questi strumenti permettono di ottenere risultati di facile comprensione ed applicazione, ma non sono generalizzabili. Se si vogliono effettuare analisi per ottenere informazioni che siano sia generalizzabili oltre che più dettagliate e precise è indispensabile un approccio di tipo *probabilistico – inferenziale*. In questo capitolo vengono forniti gli strumenti idonei per poter applicare a casi concreti questo secondo approccio. Nel seguito richiameremo brevemente i concetti di

- *variabile casuale* (v.c.)
- metodo di *stima*
- proprietà degli *stimatori*
- *test delle ipotesi*
- *processo stocastico*

A ciascuno di questi argomenti dedicheremo alcuni cenni rinviando per gli approfondimenti ai testi specialistici.

3.1 Variabili casuali univariate

Le variabili casuali che prenderemo in considerazione sono quelle che assumono valori sulla retta reale o nel piano reale. Queste possono essere sia discrete (assumono un numero finito o numerabile di valori), che continue (assumono tutti i valori di un intervallo) con una ben definita probabilità.

(A) VARIABILE CASUALE DI BERNOULLI

La più semplice variabile casuale discreta è quella di Bernoulli che assume i due soli valori 0 ed 1 rispettivamente con probabilità $(1-p)$ e p , con $0 < p < 1$. In sintesi si dirà che X è una v.c. di Bernoulli, e scriveremo $X \sim B(1, p)$, se $P(X=0)=1-p$, mentre $P(X=1)=p$. Un modo alternativo di rappresentare una v.c. di Bernoulli è il seguente

$$X = \begin{cases} 0 & 1-p \\ 1 & p \end{cases} \quad \text{o equivalentemente} \quad \pi(x) = p^x(1-p)^{1-x}, \quad x=0, 1$$

E' immediato verificare che: è sempre $X^k = X$, con k un numero reale non negativo qualsiasi; la variabile casuale $Y=1-X$ è essa stessa di Bernoulli; risulta $X(1-X)=0$. Se indichiamo con $E(X^k)$ il momento k -simo e con $\text{Var}(X)$ la varianza della generica v.c. X di Bernoulli non è difficile verificare che

$$E(X^k) = E(X) = p; \quad \text{Var}(X) = p(1-p)$$

$$\gamma_1 = E \left[\left(\frac{X-p}{\sqrt{p(1-p)}} \right)^3 \right] = \frac{1-2p}{(p(1-p))^{1/2}}$$

$$\gamma_2 = k_2 - 3 = E \left[\left(\frac{X-p}{\sqrt{p(1-p)}} \right)^4 \right] - 3 = \frac{1}{p(1-p)} - 6$$

e la v.c. di Bernoulli è asimmetrica negativa se è $0.5 < p < 1$ ed asimmetrica positiva altrimenti; è platicurtica se risulta $(3 - \sqrt{3})/6 < p < (3 + \sqrt{3})/6$, e leptocurtica altrimenti.

In molte pratiche applicazioni è noto che X è una Bernoulli, ma è incognito il parametro p che la descrive. Di p è possibile costruire una stima se si dispone di un campione casuale di ampiezza N che equivale a supporre che da X siano state fatte N estrazioni casuali con reimmissione (estrazioni bernoulliane) ottenendo (x_1, x_2, \dots, x_n) dove ciascuna x_i , per costruzione, può assumere solo uno dei due valori: 0, oppure 1. Al variare del campione ciascuna x_i descriverà una v.c. X_i , $i=1, 2, \dots, n$, che si distribuirà ancora come una $B(1, p)$. Inoltre, dato che le estrazioni sono state fatte con reimmissione queste v.c. X_i sono fra di loro indipendenti ed una stima di p può essere ottenuta come la media delle osservazioni campionarie x_i :

$$\hat{p} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{n_1}{n} \quad (1)$$

dove deve essere $0 \leq n_1 \leq n$ ed n_1 rappresenta il numero delle x_i che sono pari ad uno. D'altra parte, al variare del campione \hat{p} è una v.c., inoltre per l'indipendenza delle x_i lo *stimatore* \hat{p} possiede tutte le proprietà delle *medie campionarie*. In particolare risulta

$$E(\hat{p}) = E\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} \sum_{i=1}^n E(X_i) = \frac{1}{n} \sum_{i=1}^n p = p$$

e quindi \hat{p} è una *stima non distorta* del parametro vero p . Inoltre, tenendo conto che le estrazioni sono con reimmissione si ha

$$\text{Var}(\hat{p}) = \text{Var}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \text{Var}\left(\sum_{i=1}^n X_i\right) = \frac{1}{n^2} \sum_{i=1}^n p(1-p) = \frac{1}{n} p(1-p)$$

e per $n \rightarrow \infty$ risulta $\text{Var}(\hat{p}) \rightarrow 0$. In conclusione, si può affermare che lo stimatore di p ottenuto con la (1) è, quanto meno, *non distorto e consistente*.

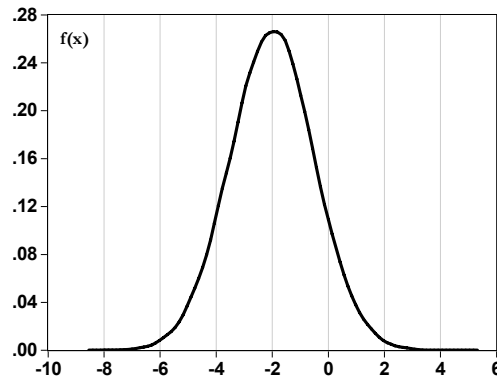
(B) VARIABILE CASUALE NORMALE

La v.c. *normale* ha un ruolo centrale nelle analisi teoriche e nelle applicazioni statistiche, essa viene anche chiamata v.c. di Gauss o di Laplace – Gauss che sono stati i primi autori a proporla ed a studiarne alcune proprietà analitiche. Questa v.c. è continua, assume tutti i valori in $(-\infty, \infty)$ ed è funzione di due parametri: la media μ e la varianza σ^2 . Per dire che la v.c. X è una normale con media μ e varianza σ^2 scriveremo $X \sim N(\mu, \sigma^2)$. Dato che la v.c. normale è continua, è necessariamente descritta da una *funzione di densità* (nel seguito f.d.) che per la normale è data da

$$f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/(2\sigma^2)} \equiv \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\} \quad (2)$$

Nel seguito, quando non vi è ambiguità, scriveremo semplicemente $f(x)$ per indicare una f.d. Nella figura 1 riportiamo la funzione di densità della v.c. normale con $\mu = -2$ e $\sigma^2 = (1.5)^2$

Figura 1 – Funzione di densità della v.c. normale con $\mu = -2$ e $\sigma^2 = (1.5)^2$



La v.c. normale possiede una serie di proprietà ed in particolare:

1. è simmetrica intorno alla propria media: $f(\mu - x) = f(\mu + x)$ per ogni x dell'asse reale;
2. possiede due flessi in corrispondenza di $x = \mu - \sigma$ ed $x = \mu + \sigma$
3. ha un unico massimo in $x = \mu$ ed è pari a $f(\mu) = \frac{1}{\sqrt{2\pi\sigma^2}}$
4. ha per asintoto l'asse delle x e quindi per $\pm x \rightarrow \infty$ risulta $f(x) \rightarrow 0$
5. per ogni v.c. normale risulta $\gamma_1 = \gamma_2 = 0$
6. qualsiasi trasformazione lineare di una v.c. normale è ancora una v.c. normale. In altri termini, se è $X \sim N(\mu, \sigma^2)$ allora la v.c. $Y = a + bX$, con a e b costanti reali qualsiasi, è ancora $N(\mu_y, \sigma_y^2)$ con

$$\mu_y = a + b\mu; \quad \sigma_y^2 = b^2 \sigma^2$$

7. la particolare trasformazione lineare con

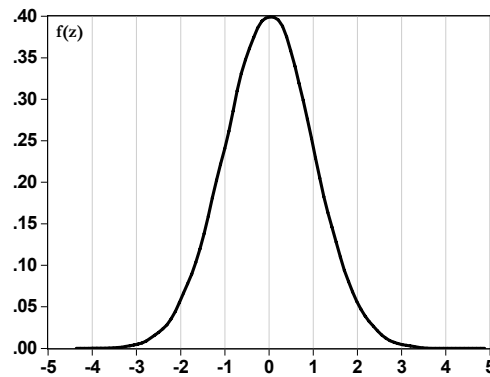
$$a = -\frac{\mu}{\sigma}; \quad b = \frac{1}{\sigma}$$

e quindi

$$Z = -\frac{\mu}{\sigma} + \frac{1}{\sigma}X = \frac{X - \mu}{\sigma}$$

prende il nome di *normale standardizzata* e, per costruzione, è sempre $E(Z)=0$, $\text{Var}(Z)=1$. Nella figura 2 riportiamo la funzione di densità della normale standardizzata

Figura 2 – Funzione di densità della v.c. normale standardizzata



8. date le k v.c. X_1, X_2, \dots, X_k indipendenti fra di loro e ciascuna distribuita come $X_i \sim N(\mu_i, \sigma_i^2)$, è anche

$$Y = a_o + \sum_{i=1}^k a_i X_i \sim N\left(a_o + \sum_{i=1}^k a_i \mu_i, \sum_{i=1}^k a_i^2 \sigma_i^2\right) \quad (3)$$

In altri termini, combinazioni lineari di normali indipendenti sono ancora distribuite come v.c. normali.

Data la v.c. $X \sim N(\mu, \sigma^2)$, noto μ e σ^2 , è possibile calcolare la probabilità che assuma valori in un determinato intervallo (a, b) tramite la seguente:

$$\begin{aligned} P(a < X < b) &= P\left(\frac{a-\mu}{\sigma} < \frac{X-\mu}{\sigma} < \frac{b-\mu}{\sigma}\right) = P\left(\frac{a-\mu}{\sigma} < Z < \frac{b-\mu}{\sigma}\right) = \\ &= \int_{z_1}^{z_2} (2\pi)^{1/2} \exp(-z^2/2) dz \end{aligned}$$

ove si è posto $z_1 = \frac{a-\mu}{\sigma}$ e $z_2 = \frac{b-\mu}{\sigma}$. Dato che il fenomeno è continuo, la probabilità in singoli punti è sempre pari a zero, questo vuol dire che è

$$P(a < X < b) = P(a \leq X < b) = P(a < X \leq b) = P(a \leq X \leq b) \quad (4)$$

E' possibile dimostrare che l'integrale sopra indicato non esiste in forma chiusa, ma può essere sempre calcolato in modo approssimato per via numerica con l'approssimazione desiderata. Per facilitare le applicazioni, sono state costruite delle apposite tavole che permettono di calcolare quelle probabilità. Alcune di queste tavole sono riprodotte in appendice dove vengono riportate probabilità del tipo $P(-\infty < Z < a)$ con $a \geq 0$, da queste si possono ricavare per standardizzazione, tenendo conto della simmetria della Z intorno allo zero, tutte le probabilità del tipo (4). Dalla figura 2 si può notare che, anche se teoricamente Z è definita su tutto l'asse reale, praticamente risulta $P(-4 < Z < 4) \approx 1$.

Spesso è plausibile supporre che un fenomeno X sia distribuito come una v.c. normale, ma di questa v.c. non sono noti i due parametri μ, σ^2 . Questi parametri possono essere stimati a partire da un campione casuale (x_1, x_2, \dots, x_n) osservato su X . In questo caso, dato che il fenomeno è continuo, non esiste più la differenza fra estrazione con reimmissione e senza reimmissione che risultano in ogni caso indipendenti. La stima dei due parametri μ, σ^2 , a partire dal campione osservato, può essere ottenuta con il metodo della *massima verosimiglianza* (MV). Posto $\mathbf{x} = (x_1, x_2, \dots, x_n)$, la *funzione di verosimiglianza* è data da

$$L(\mu, \sigma^2; \mathbf{x}) = \prod_{i=1}^n f(x_i; \mu, \sigma^2) = \left(\frac{1}{2\pi\sigma^2} \right)^{n/2} \exp \left\{ -\frac{1}{2} \sum_{i=1}^n \left(\frac{x_i - \mu}{\sigma} \right)^2 \right\} \quad (5)$$

e bisogna individuare $\hat{\mu}$ e $\hat{\sigma}^2$ di modo che la (5) sia massima. Se consideriamo il logaritmo di ambo i membri della (5), tenendo conto che la funzione logaritmica è monotona non decrescente, questa trasformazione è ancora massimizzata da $\hat{\mu}$ e $\hat{\sigma}^2$. Se al logaritmo della (5) cambiamo di segno, invece del massimo dobbiamo individuare il minimo di tale funzione:

$$-2 \log(L(\mu, \sigma^2; \mathbf{x})) = n \log(2\pi\sigma^2) + \sum_{i=1}^n \left(\frac{x_i - \mu}{\sigma} \right)^2 = \min_{\mu, \sigma^2} \quad (6)$$

Derivando la (6) rispetto a μ e σ^2 ed uguagliando il risultato a zero si ottiene

un sistema di due equazioni nelle due incognite μ, σ^2 :

$$\begin{cases} \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \mu) = 0 \\ -\frac{n}{2} \frac{1}{\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^n (x_i - \mu)^2 = 0 \end{cases}$$

la cui soluzione

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i; \quad \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \hat{\mu})^2 \quad (7)$$

ci fornisce la stima di massima verosimiglianza di media e varianza di un fenomeno distribuito come una normale. Spesso, e verrà fatto anche qui nel seguito, le (7) vengono indicate anche con $\hat{\mu} = \bar{x}$, $\hat{\sigma}^2 = S^2$. Le (7) non sono altro che la media e la varianza calcolate sul campione (x_1, x_2, \dots, x_N) . Dalle (7) si può dimostrare che:

$$E(\bar{x}) = \mu; \quad \text{Var}(\bar{x}) = \frac{\sigma^2}{n}; \quad \gamma_{1\bar{x}} = \gamma_{2\bar{x}} = 0 \quad (8)$$

e quindi \bar{x} è non distorto, consistente, simmetrico e con assenza di curtosi. Ricordando le proprietà della normale si ha anche che

$$\bar{x} \sim N(\mu, \sigma^2/n); \quad Z = \frac{(\bar{x} - \mu)\sqrt{n}}{\sigma} \sim N(0, 1) \quad (9)$$

Inoltre risulta

$$E(S^2) = (n-1) \frac{\sigma^2}{n}; \quad \text{Var}(S^2) = 2(n-1) \left(\frac{\sigma^2}{n} \right)^2$$

$$\gamma_{1S} = \left(\frac{8}{n-1} \right)^{1/2}; \quad \gamma_{2S} = \frac{12}{n-1}$$

e S^2 è asintoticamente non distorto e consistente per σ^2 , ha una distribuzione asimmetrica positiva e leptocurtica. L'asimmetria e la leptocurtosi tendono a

zero al crescere della numerosità campionaria n .

Nel caso normale è possibile dimostrare anche che \bar{x} e S^2 sono anche indipendenti fra di loro e quindi risulta $f(\bar{x}, S^2) = g(\bar{x})h(S^2)$ dove $f(\bar{x}, S^2)$ è la funzione di densità doppia della v.c. (\bar{x}, S^2) mentre $g(\cdot)$ ed $h(\cdot)$ sono le f.d. di ciascuna delle marginali.

Dalla v.c. normale è possibile derivare altre v.c. di utilizzo pratico. Qui di seguito riportiamo le più rilevanti che saranno utilizzate nei prossimi capitoli.

(C) VARIABILE CASUALE CHI QUADRATO

Questa v.c., che nel seguito indicheremo con $\chi_{(k)}^2$, può essere definita come segue. Supponiamo di avere le k v.c. normali standardizzate indipendenti Z_1, Z_2, \dots, Z_k . La nuova v.c.

$$Y = \sum_{i=1}^k Z_i^2 \quad (10)$$

si chiama *v.c. chi quadrato* e si scrive $Y \sim \chi_{(k)}^2$, dove il parametro k individua i *gradi di libertà*. I gradi di libertà si possono anche definire come la differenza fra il numero delle variabili ed il numero dei vincoli. La v.c. chi quadrato è continua, è definita sul semi asse positivo della retta, cioè $[0, \infty)$, ed ammette tutti i momenti. In particolare si può dimostrare che

$$E(\chi_{(k)}^2) = k; \quad \text{Var}(\chi_{(k)}^2) = 2k; \quad \gamma_1 = \left(\frac{8}{k}\right)^{1/2}; \quad \gamma_2 = \frac{12}{k}$$

Pertanto la v.c. chi quadrato è sempre *asimmetrica* positiva e *leptocurtica* ed entrambe, asimmetria e leptocurtosi, tendono a zero al crescere di k . Dalla (10) si ha immediatamente che se è $Y_1 \sim \chi_{(k)}^2$ e $Y_2 \sim \chi_{(h)}^2$ e sono indipendenti risulta

$$Y = Y_1 + Y_2 \sim \chi_{(k+h)}^2$$

Dalla (10) è possibile ricavare la f.d. della v.c. $\chi_{(k)}^2$ ed ottenere:

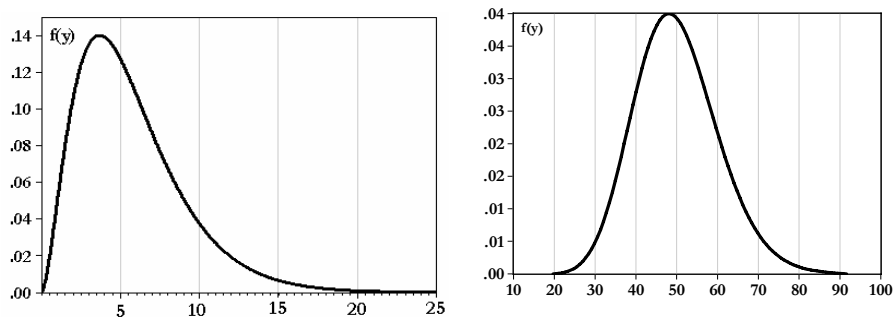
$$f(y) = \frac{e^{-y/2} y^{(k/2)-1}}{2^{k/2} \Gamma(k/2)}, \quad \text{per } y \geq 0 \quad (11)$$

con $\Gamma(k)$, detta *funzione gamma*, data da

$$\Gamma(r) = \int_0^{\infty} x^{r-1} e^{-x} dx \quad \text{con } r > 0 \text{ ed è } \Gamma(r+1) = r \Gamma(r), \quad \Gamma(1/2) = \pi$$

Dove r può anche essere un numero reale positivo qualsiasi. E' immediato verificare che se r è intero risulta $\Gamma(r+1) = r!$. Da queste considerazioni si ha che la v.c. chi quadrato può essere generalizzata al caso in cui k non necessariamente è un intero positivo, ma può essere un numero reale positivo qualsiasi. Nella figura 3 è riportato il grafico delle f.d. della v.c. chi quadrato con $k = 5.7$ e 50 gradi di libertà rispettivamente.

Figura 3 – Funzioni di densità della chi quadrato con $k = 5.7$ e $k = 50$



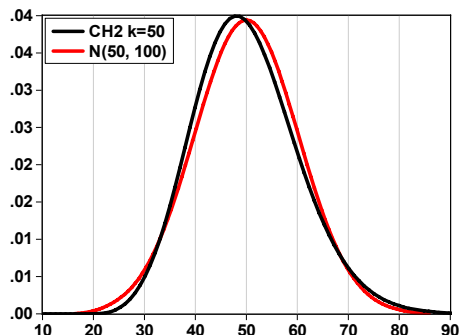
Una proprietà della v.c. chi-quadrato è che al crescere di k la sua asimmetria ed eccesso di curtosi (*leptocurtosi*) si riducono progressivamente e per $k \geq 50$ risulta ben approssimata da una v.c. normale $N(k, 2k)$. Questa proprietà è messa in evidenza dalla figura 3 che mostra come con $k=50$ la f.d. della v.c. $\chi_{(k)}^2$ diventa molto simile quella di una $N(50; 100)$ ed è confermata dalla figura

4 in cui sono confrontate le densità delle v.c. $\chi_{(50)}^2$ e $N(50, 100)$.

Tenendo conto della (7) e della (10), per l'indipendenza fra \bar{x} ed S^2 si può dimostrare che

$$\frac{n}{\sigma^2} S^2 \sim \chi_{(n-1)}^2; \quad S^2 \sim \frac{\sigma^2}{n} \chi_{(n-1)}^2 \quad (12)$$

Questo risultato è utile per costruire test delle ipotesi su σ^2 .

Figura 4 – Funzione di densità del chi quadrato con $k=50$ e della $N(50, 100)$ (D) VARIABILE CASUALE t DI STUDENT

La v.c. t di Student, che nel seguito indicheremo con $t_{(k)}$, si può ottenere come una trasformata delle v.c. normale e chi quadrato. Più precisamente, se risulta $Z \sim N(0, 1)$ ed $Y \sim \chi_{(k)}^2$ e sono indipendenti, la nuova v.c.

$$T = \frac{Z}{(Y/k)^{1/2}} = \frac{N(0, 1)}{(\chi_{(k)}^2/k)^{1/2}} \quad (13)$$

viene chiamata *v.c. t di Student* con k gradi di libertà. Per costruzione, il denominatore della (13) è sempre non negativo e pertanto la v.c. t assume tutti i valori assunti dalla normale. Ricordando le proprietà della v.c. chi quadrato è immediato dimostrare che

$$E\left(\frac{\chi_{(k)}^2}{k}\right) = \frac{1}{k} E(\chi_{(k)}^2) = 1; \quad \text{Var}\left(\frac{\chi_{(k)}^2}{k}\right) = \frac{1}{k^2} \text{Var}(\chi_{(k)}^2) = \frac{2}{k}$$

Questo risultato implica che la v.c. che è al denominatore della (13) in media vale 1 indipendentemente da k , mentre la sua varianza tende a zero al crescere di k . In altri termini la v.c. che è al denominatore della (13) al crescere di k tende ad assumere il valore 1 con certezza. Una conseguenza di questo risultato è che al crescere di k la v.c. $t_{(k)}$ si avvicina sempre di più alla v.c. che è al suo numeratore e cioè alla normale standardizzata. In pratica basta che sia $k \geq 30$

perchè l'approssimazione alla normale standardizzata sia molto buona. Si può dimostrare che per una v.c. $t_{(k)}$ risulta

$$E(t_{(k)}) = 0 \text{ se } k > 1; \quad \text{Var}(t_{(k)}) = \frac{k}{k-2} \text{ se } k > 2$$

$$\gamma_1 = 0 \text{ se } k > 3; \quad \gamma_2 = \frac{6}{k-4} \text{ se } k > 4. \quad (14)$$

Più in generale la v.c. $t_{(k)}$ ammette momento di ordine r se è $k > r$. Per $k=1$ la t di Student non ammette momenti di nessun tipo e viene chiamata v.c. di *Cauchy*. Da quanto abbiamo osservato si ha che se si vuole stimare il vettore $(\mu, \sigma^2, \gamma_1, \gamma_2)$ deve necessariamente essere $k > 4$. Nel seguito supporremo che questa condizione sia sempre verificata. Dalla (14) si ha che la v.c. t di Student è sempre simmetrica intorno allo zero, è leptocurtica e la sua curtosi tende a zero al crescere di k . Questa considerazione implica che i *valori estremi* (valori troppo piccoli e lontani da zero in negativo, o valori troppo grandi e lontani da zero in positivo) sono più frequenti nella v.c. t di Student rispetto a quanto avviene nella v.c. Normale standardizzata. In altri termini la v.c. $t_{(k)}$ ha code più "pesanti" di quelle della normale standardizzata.

Utilizzando la (13) si può dimostrare che la f.d. della t di Student con k gradi di libertà è pari a

$$f(t) = \frac{1}{(k\pi)^{1/2}} \frac{\Gamma((k+1)/2)}{\Gamma(k/2)} \left(1 + t^2/k\right)^{-(k+1)/2} \quad (15)$$

Anche nella v.c. $t_{(k)}$ la costante k non necessariamente deve essere un intero, ma può essere un numero reale positivo qualsiasi. Nella figura 5, a titolo di esempio riportiamo la funzione di densità (15) quando è $k = 5.7$ e 30 , rispettivamente.

Come si può osservare dalla figura 5 la $t_{(k)}$ con $k = 5.7$ gradi di libertà ha code più marcate rispetto a quella con $k = 30$ gradi di libertà. Nella figura 6 confrontiamo la f.d. della $t_{(k)}$ con 5.7 gradi di libertà con quella della normale standardizzata.

Figura 5 – Funzione di densità della v.c. $t_{(k)}$ di Student con $k = 5.7$ e $k = 30$

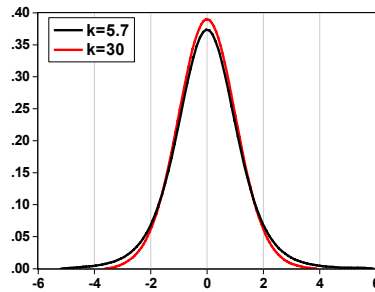
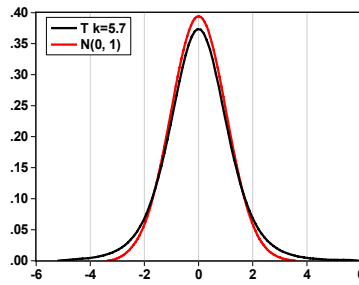
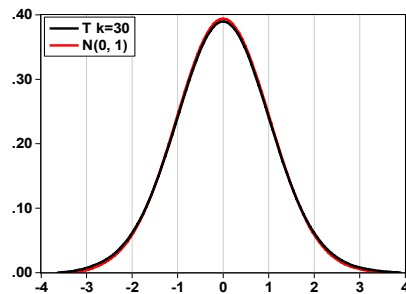


Figura 6 – Funzione di densità delle v.c. $t_{(k)}$, con $k = 5.7$, e $N(0, 1)$



Come si può osservare confrontando le figure 5 e 6, queste sono praticamente equivalenti il che giustifica che, per $k \geq 30$, la v.c. $t_{(k)}$ può essere ben approssimata dalla normale standardizzata. Quanto abbiamo appena affermato è messo ancora meglio in evidenza dalla figura 7 dove sono riportate le f.d. della $t_{(30)}$ e quella della $N(0, 1)$.

Figura 7 – Funzioni di densità della v.c. $t_{(30)}$ e della $N(0, 1)$



La v.c. $t_{(k)}$ ha un ruolo molto rilevante nelle pratiche applicazioni ed in particolare nel test delle ipotesi sulla media di una data popolazione. Infatti, se \bar{x} ed S^2 sono, rispettivamente, la stima di *massima verosimiglianza* (MV) della media e della varianza di un fenomeno che si distribuisce come una v. c. normale, ricordando la (9) e la (12) e che \bar{x} ed S^2 sono indipendenti risulta

$$\begin{aligned} \frac{\sqrt{N}(\bar{x} - \mu)}{\sigma} \cdot \left(\frac{NS^2}{\sigma^2} \frac{1}{N-1} \right)^{1/2} &= \frac{N(0,1)}{\left(\chi_{(N-1)}^2 / (N-1) \right)^{1/2}} = \\ &= \frac{\sqrt{N-1}(\bar{x} - \mu)}{S} \sim t_{(N-1)} \end{aligned} \quad (16)$$

La (15) verrà utilizzata sistematicamente per sottoporre a test i parametri di modelli e verificarne la loro validità.

(E) VARIABILE CASUALE F DI FISHER

La v.c. *F di Fisher* è anche essa derivata come una trasformazione di v.c. riconducibili alla normale. Più in particolare, supponiamo di avere le due v.c. chi quadrato $X \sim \chi_{(h)}^2$ ed $Y \sim \chi_{(k)}^2$ fra di loro indipendenti, la nuova v.c.

$$F = \frac{\chi_{(h)}^2 / h}{\chi_{(k)}^2 / k} \quad (17)$$

si chiama v.c. *F di Fisher* con (h, k) gradi di libertà e si scrive $F \sim F_{(h, k)}$. Dalla (17) si ha immediatamente che l'inverso di una *F* è ancora una v.c. *F di Fisher* dato che

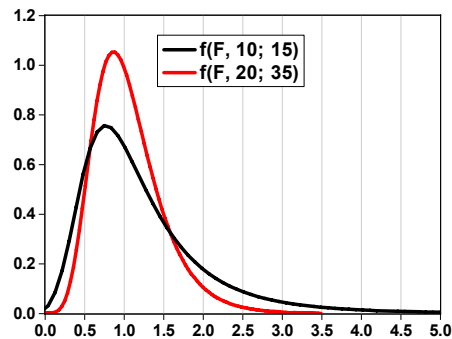
$$\frac{1}{F} = \frac{\chi_{(k)}^2 / k}{\chi_{(h)}^2 / h} \sim F_{(k, h)}$$

Inoltre, dalla (13) risulta

$$t_{(k)}^2 = \frac{(N(0,1))^2 / 1}{\chi_{(k)}^2 / k} \sim F_{(1, k)}$$

In questo senso la v.c. F può essere considerata una generalizzazione della v.c. $t_{(k)}$. Ricordando che per k che tende all'infinito la v.c. $\chi_{(k)}^2/k$ tende ad assumere il valore 1 con certezza, segue che al crescere di k la v.c. $h F_{(h,k)}$ tende alla v.c. $\chi_{(h)}^2$. Nella figura 8 riportiamo la f.d. della F per le due coppie di gradi di libertà (10; 15) e (20; 35).

Figura 8 – Funzioni di densità delle v.c. $F_{(10, 15)}$ e $F_{(20, 35)}$



E' possibile dimostrare che è

$$E(F_{(h,k)}) = \frac{k}{k-2} \quad \text{se } k > 2; \quad \text{Var}(F_{(h,k)}) = \frac{2k^2(h+k-2)}{h(k-2)^2(k-4)} \quad \text{se } k > 4$$

$$\gamma_1 = \left(\frac{8(k-4)}{h(h+k-2)} \right)^{1/2} \frac{2h+k-2}{k-6} \quad \text{se } k > 6$$

$$\gamma_2 = \frac{12 \left[(k-2)^2(k-4) + h(h+k-2)(5k-22) \right]}{h(k-6)(k-8)(h+k-2)} \quad \text{se } k > 8$$

Dalle formule sopra riportate si ha che l'esistenza di questi indici dipende esclusivamente dai gradi di libertà del denominatore. Inoltre, si verifica immediatamente che è sempre $\gamma_1 > 0$ e $\gamma_2 > 0$. Pertanto la F di Fisher è sempre asimmetrica positiva e leptocurtica.

(F) VARIABILE CASUALE LOG-NORMALE

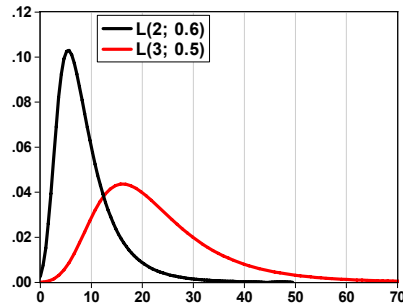
Supponiamo di avere la v.c. $X \sim N(\lambda, \delta)$, la nuova v.c. Y

$$Y = e^X \quad (18)$$

si chiama v.c. *log-normale* con parametri (λ, δ) e si scrive $Y \sim L(\lambda, \delta)$. Dalla (18) si ha che la v.c. log-normale è definita sul semiasse positivo della retta reale; inoltre, se si considera il logaritmo di ambo i membri della (18) risulta

$$\log(Y) = X \sim N(\lambda, \delta) \quad (19)$$

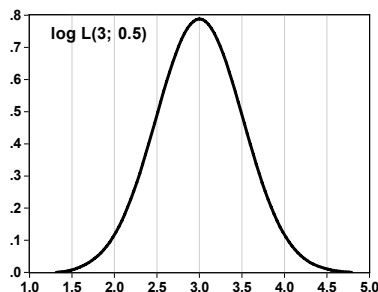
Figura 9 – Funzioni di densità delle log-normali con $(2; 0.6)$ e $(3; 0.5)$



La (19) implica che se un fenomeno si distribuisce come una log-normale, per analizzarlo basta utilizzare la normale calcolata sul suo logaritmo. Nella figura 9 sono riportate le funzioni di densità delle log-normali con parametri $(2; 0.6)$ e $(3; 0.5)$.

Nella figura 10 è riportata la f.d. del logaritmo della log normale con parametri $(3; 0.5)$ e, come si può osservare, si distribuisce come una $N(3; 0.5)$.

Figura 10 – Funzione di densità del logaritmo della log-normale con $(3; 0.5)$



Della v.c. log normale è possibile calcolare tutti i momenti, in particolare, si può dimostrare che è:

$$E(Y) = e^{\lambda + \delta^2/2}; \quad \text{Var}(Y) = e^{2\lambda + \delta^2} \left(e^{\delta^2} - 1 \right) \quad (20)$$

$$\gamma_1 = \left(e^{\delta^2} + 2 \right) \left(e^{\delta^2} - 1 \right)^{1/2}; \quad \gamma_2 = e^{4\delta^2} + 2e^{3\delta^2} + 3e^{2\delta^2} - 6$$

Da queste espressioni è immediato verificare che è sempre $E(Y) > 0$, $\gamma_1 > 0$ e $\gamma_2 > 0$, pertanto la v.c. log-normale è sempre asimmetrica positiva e leptocurtica.

Dalle (20), ricordando che $\text{Var}(Y) = E(Y^2) - [E(Y)]^2$, si ha

$$E(Y) = e^{\lambda + \delta^2/2}; \quad E(Y^2) = e^{2\lambda + 2\delta^2}$$

e, considerando i logaritmi di ambo i membri e risolvendo il relativo sistema, si ottiene

$$\begin{cases} \lambda = 2 \log[E(Y)] - \frac{1}{2} \log[E(Y^2)] \\ \delta^2 = \log[E(Y^2)] - \log[E(Y)] \end{cases}$$

che mette in relazione la media e la varianza della v.c. normale $X \sim N(\lambda, \delta)$ con i primi due momenti della log normale $L(\lambda, \delta)$. Questa relazione permette di ottenere, nota una stima dei primi due momenti di Y , una stima di λ e δ .

3.2 Variabili casuali doppie

Le v.c. semplici possono essere generalizzate al caso multiplo quando si vogliono analizzare simultaneamente due o più fenomeni. Qui di seguito trattiamo brevemente delle v.c. doppie (X, Y) . In questo caso, nota la legge di distribuzione della doppia (X, Y) è sempre possibile derivare le leggi di distribuzione delle v.c. semplici X ed Y che vengono chiamate *marginali*. In genere non è vero il viceversa e cioè che, note le marginali X ed Y , sia possibile derivare la doppia (X, Y) a meno che X ed Y non siano *indipendenti*. Inoltre, da una v.c. doppia, oltre alle marginali, è possibile derivare le v.c. *condizionate*. Di

queste ne esistono due classi: le X dato che è $Y=y$, indicate con $(X | Y = y)$, e le Y dato che è $X=x$, indicate con $(Y | X = x)$. Nota una delle due classi di v.c. *condizionate* è sempre possibile derivare la v.c. doppia.

Della v.c. doppia (X, Y) è possibile derivare, quando esistono, alcuni indici sintetici ed in particolare:

- le *medie* delle due marginali: $E(X) = \mu_x$; $E(Y) = \mu_y$
- le *varianze* delle due marginali: $\text{Var}(X) = \sigma_x^2$; $\text{Var}(Y) = \sigma_y^2$
- la *covarianza* fra X ed Y : $\text{cov}(X, Y) = E[(X - \mu_x)(Y - \mu_y)] = \sigma_{xy}$

L'indice σ_{xy} misura l'esistenza di legami lineari fra le marginali della doppia.

Queste espressioni possono essere scritte anche con vettori e matrici ed ottenere il *vettore delle medie*

$$\boldsymbol{\mu}' = (\mu_x, \mu_y)$$

e la *matrice di varianze-covarianze*

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{bmatrix}$$

La matrice di varianze-covarianze possiede una serie di proprietà ed in particolare:

- è simmetrica e quindi $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}'$
- è semi definita positiva e quindi il suo *determinante* è tale che

$$|\boldsymbol{\Sigma}| = \sigma_x^2 \sigma_y^2 - \sigma_{xy}^2 \geq 0 \quad (21)$$

- tenendo conto della (21), si può costruire il nuovo indice

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \quad (22)$$

che prende il nome di *correlazione* e misura l'*intensità* dei legami lineari esistenti fra X ed Y e risulta $-1 \leq \rho_{xy} \leq 1$

- se è $|\boldsymbol{\Sigma}| > 0$, della matrice $\boldsymbol{\Sigma}$ è possibile costruire la sua *matrice inversa*

Σ^{-1} per cui $\Sigma^{-1} \Sigma = \Sigma \Sigma^{-1} = \mathbf{I}$, dove \mathbf{I} è una matrice che assume valori uno sulla diagonale principale e zero fuori la diagonale principale:

$$\mathbf{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

la matrice \mathbf{I} viene chiamata *matrice identica*;

- dalla matrice

$$\Sigma = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{bmatrix} = \begin{bmatrix} \sigma_x & 0 \\ 0 & \sigma_y \end{bmatrix} \begin{bmatrix} 1 & \sigma_{xy}/\sigma_x\sigma_y \\ \sigma_{xy}/\sigma_x\sigma_y & 1 \end{bmatrix} \begin{bmatrix} \sigma_x & 0 \\ 0 & \sigma_y \end{bmatrix}$$

è possibile derivare la nuova matrice \mathbf{R} , detta matrice di *correlazione*, data da

$$\mathbf{R} = \begin{bmatrix} 1 & \sigma_{xy}/\sigma_x\sigma_y \\ \sigma_{xy}/\sigma_x\sigma_y & 1 \end{bmatrix} = \begin{bmatrix} 1 & \rho_{xy} \\ \rho_{xy} & 1 \end{bmatrix}$$

anche questa matrice è simmetrica e semi definita positiva e risulta $|\mathbf{R}| = 1 - \rho_{xy}^2 \geq 0$.

Della v.c. condizionata ($X | Y = y$) si possono calcolare, se esistono, media e varianza che saranno, in generale, una funzione di y :

$$E(X|Y = y) = \mu_{x/y}; \quad \text{Var}(X|Y = y) = \sigma_{x/y}^2$$

lo stesso si ha per la media e la varianza di ($Y | X = x$) che, in generale, saranno una funzione di x :

$$E(Y|X = x) = \mu_{y/x}; \quad \text{Var}(Y|X = x) = \sigma_{y/x}^2$$

Queste funzioni permettono di evidenziare i legami, in media, che esistono fra X ed Y e fra Y ed X .

Presentiamo due v.c. doppie che verranno utilizzate nel seguito.

(A) VARIABILE CASUALE DI BERNOULLI DOPPIA

La Bernoulli doppia (B_1, B_2) è una v.c. discreta che assume quattro coppie di valori con le rispettive probabilità come riportato nella tabella 1:

Tabella 1 – Distribuzione di probabilità della Bernoulli doppia

$B_2 \backslash B_1$	0	1	Totale
0	$1-p-q+\tau$	$p-\tau$	$1-q$
1	$q-\tau$	τ	q
Totale	$1-p$	p	1

dove è $p \geq \tau$; $q \geq \tau$; $\tau = P(B_1 = 1, B_2 = 1) \geq p + q - 1 \geq 0$. Dalla tabella 1 si ha che noti τ e le marginali B_1 e B_2 è possibile derivare la doppia e viceversa. In molti casi è $p = q$, tale distribuzione di (B_1, B_2) è riportata nella tabella 2.

Tabella 2 – Distribuzione di probabilità della Bernoulli doppia con $p = q$

$B_2 \backslash B_1$	0	1	Totale
0	$1-2p+\tau$	$p-\tau$	$1-p$
1	$p-\tau$	τ	p
Totale	$1-p$	p	1

Dalla tabella 1 si ha che B_1 e B_2 sono indipendenti se è $\tau = pq$, inoltre risulta:

$$E(B_1) = p; \quad \text{Var}(B_1) = p(1-p); \quad E(B_2) = q; \quad \text{Var}(B_2) = q(1-q)$$

$$E(B_1 B_2) = \tau; \quad \text{Cov}(B_1, B_2) = \tau - pq$$

ed infine

$$\Sigma = \begin{bmatrix} p(1-p) & \tau - pq \\ \tau - qp & q(1-q) \end{bmatrix}; \quad |\Sigma| = p(1-p)q(1-q) - (\tau - pq)^2 \geq 0$$

e quindi

$$\rho = \frac{\tau - pq}{[p(1-p)q(1-q)]^{1/2}}, \quad \text{per } 0 < p < 1, \quad 0 < q < 1$$

che quando è $p = q$ diviene

$$\rho = \frac{\tau - p^2}{p(1-p)}, \quad \text{per } 0 < p < 1$$

Per la v.c. condizionata $(B_1 | B_2 = y)$ in questo caso è $y = 0, 1$ e si ha rispettivamente

$$E(B_1 | B_2 = 0) = \frac{p - \tau}{1 - q}; \quad E(B_1 | B_2 = 1) = \frac{\tau}{q}$$

che implicano

$$E(B_1) = (1 - q)E(B_1 | B_2 = 0) + qE(B_1 | B_2 = 1) = p$$

cioè $E(B_1)$ è la media ponderata di $E(B_1 | B_2 = 0)$ ed di $E(B_1 | B_2 = 1)$. Un risultato analogo si ottiene per $E(B_2)$, mentre si ha

$$\text{Var}(B_1 | B_2 = 0) = \frac{p - \tau}{1 - q} \left(1 - \frac{p - \tau}{1 - q} \right); \quad \text{Var}(B_1 | B_2 = 1) = \frac{\tau}{q} \left(1 - \frac{\tau}{q} \right)$$

Mentre $\text{Var}(B_1)$ non è data dalla media ponderata fra $\text{Var}(B_1 | B_2 = 0)$ e $\text{Var}(B_1 | B_2 = 1)$. Risultati analoghi si hanno per la condizionata $(B_2 | B_1 = x)$ e quando è $p = q$.

(B) VARIABILE CASUALE NORMALE DOPPIA

La v.c. normale può essere generalizzata al caso multiplo ed in particolare a quello doppio. Più precisamente, la v.c. doppia $\mathbf{X}' = (X, Y)$ è normale, e si

indica con $\mathbf{X}' = (X, Y) \sim N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, se la sua funzione di densità è data da

$$f(x, y) = (2\pi)^{-2/2} |\boldsymbol{\Sigma}|^{-1/2} \exp\left\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})\right\}$$

ove si è posto

$$\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}; \quad \boldsymbol{\mu} = \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}; \quad \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{bmatrix}; \quad \mathbf{R} = \begin{bmatrix} 1 & \rho_{xy} \\ \rho_{xy} & 1 \end{bmatrix}$$

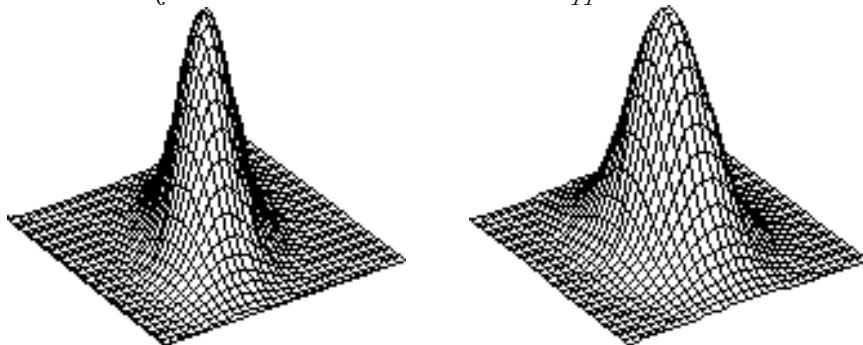
$$\boldsymbol{\Sigma}^{-1} = \frac{1}{\sigma_x^2 \sigma_y^2 - \sigma_{xy}^2} \begin{bmatrix} \sigma_y^2 & -\sigma_{xy} \\ -\sigma_{xy} & \sigma_x^2 \end{bmatrix}$$

Come si può osservare, la normale doppia è funzione dei cinque parametri $(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho_{xy})$ con $\rho_{xy} = \sigma_{xy} / \sigma_x \sigma_y$. Nella figura 11 sono riportate le f.d di due normali doppie con parametri, rispettivamente, pari a

$$(\mu_x = 0, \mu_y = 0, \sigma_x^2 = 1, \sigma_y^2 = 2, \rho_{xy} = 0.5)$$

$$(\mu_x = 0, \mu_y = 0, \sigma_x^2 = 1, \sigma_y^2 = 2, \rho_{xy} = 0)$$

Figura 11 – Funzione di densità delle due v.c. normali doppie



La v.c. normale doppia possiede, fra l'altro, le seguenti proprietà:

1. se (X, Y) è una normale doppia allora X ed Y sono indipendenti se e solo se risulta $\rho_{xy} = 0$. In altri termini, nel mondo delle normali esistono solo legami lineari, pertanto X ed Y sono indipendenti *se e solo se sono incorrelate*;
2. le marginali di una normale doppia $(X, Y) \sim N(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho_{xy})$ sono ancora normali univariate: $X \sim N(\mu_x, \sigma_x^2)$ ed $Y \sim N(\mu_y, \sigma_y^2)$
3. le condizionate di una normale doppia sono normali univariate:

$$(X|Y = y) \sim N\left(\mu_x + \frac{\sigma_{xy}}{\sigma_y^2}(y - \mu_y); \sigma_x^2(1 - \rho_{xy}^2)\right)$$

$$(Y|X = x) \sim N\left(\mu_y + \frac{\sigma_{xy}}{\sigma_x^2}(x - \mu_x); \sigma_y^2(1 - \rho_{xy}^2)\right)$$

è di interesse osservare che le medie condizionate della normale doppia sono delle particolari rette;

4. la funzione di densità di una normale doppia possiede un unico massimo in $(x = \mu_x, y = \mu_y)$;
5. se (X, Y) è una v.c. Normale doppia con $\mu_x = \mu_y = 0$ si dimostra che

$$\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X} = \frac{1}{\sigma_x^2\sigma_y^2 - \sigma_{xy}^2} \left[X^2\sigma_y^2 - 2XY\sigma_{xy} + Y^2\sigma_x^2 \right] \sim \chi_{(2)}^2$$

Più in generale, data la matrice \mathbf{A} tale che $\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}' = \mathbf{A}$, è sempre $\mathbf{X}'\mathbf{A}\mathbf{X} \sim \chi_{(2)}^2$;

6. se \mathbf{X} si distribuisce come una normale doppia allora una sua qualsiasi trasformazione lineare è ancora normale:

$$Y = \mathbf{A}\mathbf{X} + \mathbf{a} \sim N(\mathbf{A}\boldsymbol{\mu} + \mathbf{a}; \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}')$$

dove \mathbf{A} è una matrici di costanti ed \mathbf{a} un vettore di costanti qualsiasi. In particolare se \mathbf{A} è il vettore riga $\mathbf{A} = (b, c)$ ed \mathbf{a} uno scalare risulta

$$\begin{aligned}
 a + (b, c) \begin{pmatrix} X \\ Y \end{pmatrix} &= a + bX + cY \sim N \left(a + (b, c) \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}; (b, c) \begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{bmatrix} \begin{pmatrix} b \\ c \end{pmatrix} \right) \\
 &= N \left(a + b\mu_x + c\mu_y; b^2\sigma_x^2 + 2bc\sigma_{xy} + c^2\sigma_y^2 \right).
 \end{aligned}$$

In altri termini, combinazioni lineari di normali doppie, siano esse dipendenti o indipendenti, si distribuiscono ancora come normali;

7. dato il campione casuale (x_i, y_i) , $i = 1, 2, \dots, N$, estratto da una normale doppia questo può essere utilizzato per stimare media e varianza delle marginali X ed Y con gli strumenti che abbiamo visto nel paragrafo precedente e tali proprietà possono essere utilizzate per avere una stima della correlazione ρ_{xy} tramite la seguente

$$r = \frac{S_{xy}}{S_x S_y} = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\left[\sum_{i=1}^N (x_i - \bar{x})^2 \sum_{i=1}^N (y_i - \bar{y})^2 \right]^{1/2}}$$

Si può dimostrare che, sotto l'ipotesi che sia $\rho_{xy} = 0$ e quindi che X ed Y siano indipendenti, risulta

$$r \left(\frac{N-2}{1-r^2} \right)^{1/2} \sim t_{(N-2)}$$

tale risultato viene frequentemente utilizzato in molte applicazioni per sottoporre a test l'indipendenza di caratteri ben approssimati da normali.

I risultati sopra illustrati per il caso di normali bivariate possono essere estesi, abbastanza agevolmente, alle normali multivariate.

3.3 Alcuni risultati limite

In questo paragrafo presentiamo alcuni risultati rilevanti riguardanti il comportamento limite di successioni di v.c. Più precisamente data la successione di v.c. $X_1, X_2, \dots, X_n, \dots = \{X_n\}$ si vuole individuare, sotto

specificate condizioni, il suo comportamento o quello di una sua opportuna trasformazione quando n tende ad infinito. A questo proposito diamo alcune definizioni.

Definizione 1

Si dice che la successione di v.c. $\{X_n\}$ converge in probabilità alla v.c. X se per ogni $\varepsilon > 0$ risulta

$$\lim_{n \rightarrow \infty} P(|X_n - X| < \varepsilon) = 1 \quad (23)$$

Spesso per indicare la convergenza in probabilità si usano anche le seguenti notazioni

$$P \lim_{n \rightarrow \infty} X_n = X; \quad X_n \xrightarrow{P} X$$

Definizione 2

La successione $\{X_n\}$ converge quasi certamente (a.s.) ad X se

$$\lim_{N \rightarrow \infty} P\left\{\sup_{n \geq N} |X_n - X| > \varepsilon\right\} = 0 \text{ per ogni } \varepsilon > 0$$

e scriveremo $(X_n - X) \xrightarrow{\text{a.s.}} 0$.

Definizione 3

La successione $\{X_n\}$ converge in media quadratica (q. m.) ad X se

$$\lim_{n \rightarrow \infty} E[(X_n - X)^2] = 0$$

e scriveremo $(X_n - X) \xrightarrow{\text{q.m.}} 0$

Si può dimostrare che la convergenza a.s. e quella q.m. implicano quella in probabilità che, pertanto, viene anche chiamata *convergenza debole*.

Data la successione di v.c. $\{X_n\}$, supponiamo che i momenti $\mu_i = E(X_i)$ e $\sigma_i^2 = \text{Var}(X_i)$ esistono finiti per ogni i e calcoliamo la media campionaria

$$\bar{X}_n = \frac{1}{n} \sum_{j=1}^n X_j$$

E' immediato verificare che è sempre

$$E(\bar{X}_n) = \frac{1}{n} \sum_{i=1}^n \mu_i; \quad \text{Var}(\bar{X}_n) = \frac{1}{n^2} \sum_{i=1}^n \sigma_i^2$$

Questo ci permette di dare la seguente

Definizione 4

Si dice che la successione $\{\bar{X}_n\}$ soddisfa la *legge dei grandi numeri* se è

$$\bar{X}_n - E(\bar{X}_n) \xrightarrow{P} 0$$

Se invece è

$$\bar{X}_n - E(\bar{X}_n) \xrightarrow{\text{a.s.}} 0$$

allora $\{\bar{X}_n\}$ soddisfa la *legge forte dei grandi numeri*.

Alcuni risultati importanti sono i seguenti:

1. se è $\lim_{n \rightarrow \infty} \frac{1}{n^2} \sum_{i=1}^n \sigma_i^2 = 0$

allora è anche $\bar{X}_n - E(\bar{X}_n) \xrightarrow{P} 0$ e quindi $\{X_n\}$ soddisfa la legge dei grandi numeri.

Esempio

Supponiamo che sia

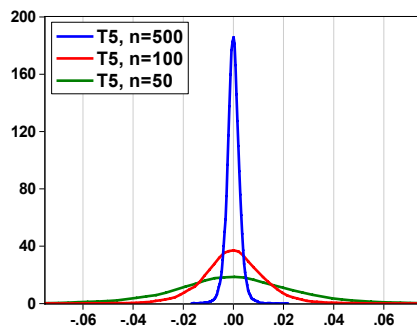
$$X_n = \frac{1}{n} t_{(k)}$$

dove con $t_{(k)}$ si è indicata la v.c. t di Student con $k > 2$ gradi di libertà. In tal caso è facile verificare che

$$\bar{X}_n = \frac{1}{n} \sum_1^n \left(\frac{1}{n} t_{(k)} \right) = \frac{1}{n} t_{(k)}; \quad E(\bar{X}_n) = 0; \quad \text{Var}(\bar{X}_n) = \frac{k}{n^2(k-2)}$$

e quindi $\{\bar{X}_n\}$ soddisfa la *legge dei grandi numeri*.

Figura 12 – Distribuzione di \bar{X}_n per $n=50, 100, 500$



Nella figura 12 è riportata la distribuzione di $\left\{\frac{1}{n}t_{(5)}\right\}$ con $n = 50, 100, 500$, rispettivamente, ed è evidente come la successione tende, al crescere di n , ad assumere il valore zero con certezza.

□

2. Se è $\lim_{n \rightarrow \infty} \sum_{i=1}^n \frac{\sigma_i^2}{i^2} = 0$

allora è anche $(\bar{X}_n - E(\bar{X}_n)) \xrightarrow{\text{a.s.}} 0$

ed $\{\bar{X}_n\}$ soddisfa la *legge forte dei grandi numeri*.

3. se $g(\cdot)$ è una funzione continua e se $X_n \xrightarrow{P} X$ allora è anche $g(X_n) \xrightarrow{P} g(X)$

Definizione 5

Data la successione $\{X_n\}$, sia $F_n(x)$ la funzione di ripartizione di X_n ed $F(x)$ quella della v.c. X , si dice che $\{X_n\}$ converge in legge o in distribuzione ad X se

$$\lim_{n \rightarrow \infty} F_n(x) = F(x)$$

in tutti i punti in cui $F(x)$ è continua.

Per dire che $\{X_n\}$ converge in legge ad X spesso si usa la notazione

$$X_n \xrightarrow{L} X$$

Nella convergenza in distribuzione la v.c. limite X viene anche chiamata v.c. *asintotica* e la sua funzione di ripartizione, $F(x)$, distribuzione *asintotica*. Se la funzione di ripartizione $F(x)$ di X è continua per cui ammette la f.d. $f(x)$ e se $X_n \xrightarrow{L} X$ allora si dice che X_n converge ad X *uniformemente in legge*.

E' possibile dimostrare i seguenti risultati notevoli:

1. la convergenza in probabilità implica quella in distribuzione, ma non vale il viceversa a meno che X non sia una costante nel qual caso le due convergenze sono equivalenti;
2. se $X_n \xrightarrow{L} X$, se $g(\cdot)$ è una funzione continua e se X_n ed X ammettono i momenti di ordine k allora si ha che

$$E[g(X_n)^k] \longrightarrow E[g(X)^k]$$

ed in particolare i momenti di ordine k di X_n convergono ai momenti di ordine k di X ;

3. date le successioni $\{X_n\}$ e $\{Y_n\}$

- se $|X_n - Y_n| \xrightarrow{P} 0$ ed $Y_n \xrightarrow{L} Y$, allora $X_n \xrightarrow{L} Y$
- se $X_n \xrightarrow{L} X$ ed $Y_n \xrightarrow{P} 0$, allora $X_n Y_n \xrightarrow{P} 0$
- se $X_n \xrightarrow{L} X$ ed $Y_n \xrightarrow{P} c$, allora

$$X_n + Y_n \xrightarrow{L} X + c, \quad X_n Y_n \xrightarrow{L} Xc$$

- se $X_n \xrightarrow{L} X$ e $Y_n \xrightarrow{P} 1$, allora $X_n \frac{1}{Y_n} \xrightarrow{P} X$

Esempio

Sia

$$X_n = Z + \frac{1}{n} t_{(k)}$$

con $Z \sim N(0, 1)$ e $t_{(k)}$ la v.c. t di Student con $k > 2$ gradi di libertà indipendenti fra di loro, in questo caso si ha

$$\bar{X}_n = \frac{1}{n} \sum_{j=1}^n X_j = Z + \frac{1}{n} t_{(k)}$$

che implica

$$E(\bar{X}_n) = 0; \quad \text{Var}(\bar{X}_n) = 1 + \frac{k}{n^2(k-2)}$$

e quindi

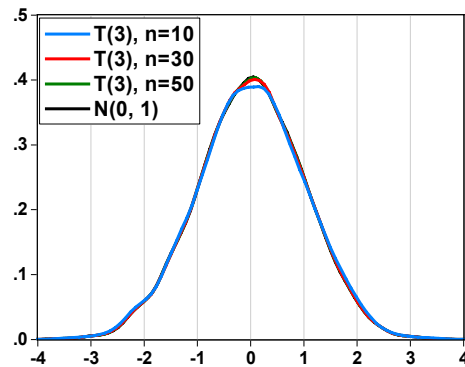
$$\frac{1}{n} t_{(k)} \xrightarrow{P} 0$$

ma allora

$$X_n = \bar{X}_n \xrightarrow{P} Z$$

Nella figura 13 illustriamo questa convergenza per $k=3$.

Figura 13 – Distribuzione di X_n con $k=3$ ed $n=10, 30, 50$



Come si può osservare dalla figura la convergenza è molto rapida. □

Possiamo ora dare l'enunciato di un teorema molto rilevante nell'inferenza statistica. Questo teorema va sotto il nome di *Teorema del Limite Centrale* ed è spesso indicato con la sigla CLT (*Central Limit Theorem*).

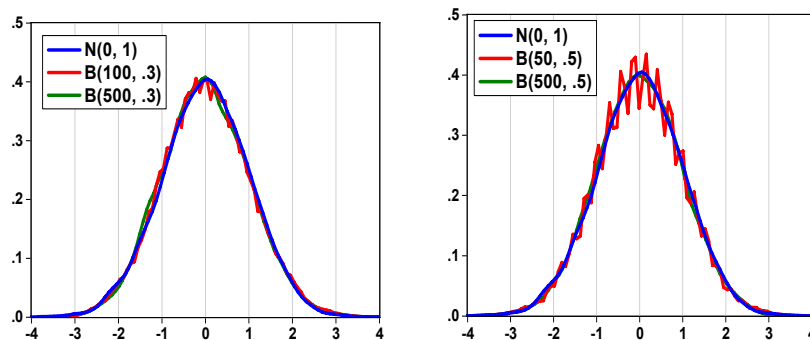
Teorema

Data la successione $\{X_n\}$ di v.c. a componenti indipendenti ed identicamente distribuite, per cui risulta $F_n(x) = F(x)$ per ogni n , se $E(X_n) = \mu$ e $\text{Var}(X_n) = \sigma^2$ sono finiti, allora si dimostra che

$$Z_n = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{L} N(0, 1) \quad (24)$$

Da questo teorema si ha che la convergenza è verificata indipendentemente dal fatto che le X_n siano v.c. discrete o continue; inoltre, se n è sufficientemente grande si ha $\bar{X}_n \approx N(\mu, \sigma^2/n)$. Un caso di questo tipo lo abbiamo già visto a proposito della v.c. chi quadrato con n gradi di libertà che può essere approssimata da una normale $N(n, 2n)$ se n è sufficientemente grande (in questo caso deve essere $n > 50$). Nella figura 14 mostriamo come opera la (24) se $X_n = B(1, p)$ con $p=0.3$ e $p=0.5$. Vengono considerati diversi valori di n e le distribuzioni ottenute vengono confrontate con una $N(0,1)$. Come si può osservare da questi grafici, l'approssimazione è accettabile quando è $n > 100$.

Figura 14 – Distribuzione di Z_n , con $X_n = B(1, p)$, $p=0.3$ e 0.5 , confrontata con la $N(0, 1)$ e diversi valori di n



Il CLT è stato generalizzato in diversi modi imponendo condizioni diverse sui momenti delle X_i , qui di seguito riportiamo una versione che presuppone l'esistenza dei primi tre momenti.

Teorema

Data la successione $\{X_n\}$ di v.c. a componenti indipendenti tali che

$$E(X_i) = \mu_i < \infty; \quad \text{Var}(X_i) = \sigma_i^2 < \infty; \quad E\left[|X_i - \mu_i|^3\right] = \beta_i < \infty,$$

posto

$$B_n = \left(\sum_{i=1}^n \beta_i\right)^{1/3}; \quad C_n = \left(\sum_{i=1}^n \sigma_i^2\right)^{1/2}$$

se risulta

$$\lim_{n \rightarrow \infty} \frac{B_n}{C_n} = 0$$

allora è

$$Z_n = \frac{\sum_{i=1}^n (X_i - \mu_i)}{C_n} \xrightarrow{L} N(0, 1)$$

Come si può notare il CLT è stato generalizzato al caso di successioni di v.c. indipendenti, ma con diversa distribuzione, imponendo però alcune condizioni sull'esistenza ed il comportamento dei primi tre momenti di ciascuna X_i .

Un caso di applicazione del CLT riguarda il comportamento asintotico delle stime di massima verosimiglianza. Si può infatti dimostrare che, se X è una v.c. con f.d. $f(x; \theta)$, con θ incognito e da stimare e se $\hat{\theta}_n$ è lo stimatore di massima verosimiglianza ottenuto sulla base del campione casuale $\mathbf{x} = (x_1, x_2, \dots, x_n)$, allora è anche

$$(\hat{\theta}_n - \theta)(nI(\theta))^{1/2} \xrightarrow{L} N(0, 1) \quad (25)$$

dove $I(\theta)$ è detta *informazione di Fisher* e risulta

$$nI(\theta) = -E\left[\frac{\partial^2 \log L(\theta; \mathbf{x})}{\partial \theta^2}\right] = E\left[\left(\frac{\partial \log L(\theta; \mathbf{x})}{\partial \theta}\right)^2\right]$$

La (25) è molto utilizzata nelle pratiche applicazioni per fare inferenza.

SVILUPPO IN SERIE DI EDGEWORTH

Questo sviluppo permette di approssimare, con un ordine prefissato, la funzione di ripartizione di una qualsiasi v.c. Sia $G(x)$ la funzione di ripartizione della v.c. X e $\Phi(x)$ quella della v.c. normale standardizzata con f.d. $\phi(x)$. $G(x)$ può essere sviluppata in *serie di Edgeworth* intorno ad $\Phi(x)$ ottenendo

$$G(x) = \Phi(x) + c_1 \Phi^{(1)}(x) + c_2 \Phi^{(2)}(x) + c_3 \Phi^{(3)}(x) + \dots$$

dove $\Phi^{(i)}(x)$ è la derivata i -sima di $\Phi(x)$ rispetto ad x . In particolare, per la v.c.

$$T_n = \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma}$$

se indichiamo con $F_n(t)$ la funzione di ripartizione di T_n , si può dimostrare che

$$F_n(t) = \Phi(t) - \phi(t) \frac{\gamma_{1n}}{6} H_2(t) + \phi(t) \frac{\gamma_{2n}}{24} H_3(t) + \phi(t) \frac{\gamma_{1n}^2}{72} H_5(t) + R_n$$

dove γ_{1n} e γ_{2n} sono gli indici di asimmetria e di curtosi di T_n , R_n è il resto ed è dell'ordine $O(n^{-3/2})$, cioè converge a zero con la stessa velocità di $1/n^{3/2}$, $H_j(t)$ sono i polinomi di *Hermite* ed è

$$H_2 = t^2 - 1; \quad H_3 = t^3 - 3t; \quad H_5 = t^5 - 10t^3 + 15t$$

3.4 Cenni alle stime di quasi massima verosimiglianza

In talune applicazioni, il ricorso al metodo della massima verosimiglianza è spesso inibito da due ordini di problemi imputabili a:

- i) difficoltà nell'identificare la famiglia parametrica di distribuzioni alla quale è riconducibile la funzione di verosimiglianza.
- ii) scarsa trattabilità, sul piano computazionale, della funzione di verosimiglianza rilevante per il problema di stima in esame.

In tali situazioni, un'alternativa praticabile alla stima di massima verosimiglianza è data dal metodo della *Quasi Massima Verosimiglianza* (QMV), originariamente proposto da Weddeburn (1974) nell'ambito della stima dei modelli lineari generalizzati (vedi McCullagh e Nelder (1989)) e successivamente applicato in maniera diffusa in diversi campi della statistica. I

principali vantaggi della stima di QMV, rispetto al classico metodo della MV, risiedono senz'altro nella possibilità di ottenere delle stime dei parametri di interesse senza la necessità di dover definire un modello stocastico completo per i dati osservati. Per poter applicare il metodo della QMV è infatti sufficiente procedere solo ad una specificazione *parziale* del modello. Il prezzo da pagare per tale flessibilità è una perdita in termini di efficienza. E' infatti possibile dimostrare, sotto condizioni abbastanza generali, che le stime di QMV sono consistenti ed asintoticamente normali, ma asintoticamente meno efficienti di quelle di MV.

Di seguito, al fine di presentare in maniera didatticamente efficace la struttura della procedura di stima, ne discutiamo l'applicazione nel campo della stima di un modello di regressione. Nello specifico, sia (Y_1, \dots, Y_n) un campione di variabili casuali indipendenti con

$$Y_i = \mu + \varepsilon_i \quad (26)$$

dove $\varepsilon_i \stackrel{\text{iid}}{\sim} (0, \sigma^2)$, per $i = 1, \dots, n$. Si assuma poi che la varianza $\text{var}(Y_i)$ sia legata al valore atteso μ da una funzione del tipo

$$\sigma^2 = \text{var}(Y_i) = \psi V(\mu)$$

dove la forma funzionale di $V(\cdot)$ è nota. Lo scopo di questa ipotesi, come si avrà modo di riscontrare analiticamente nel seguito, è quello di rendere il problema della stima della media indipendente dal problema della stima della varianza. Si consideri quindi la variabile casuale

$$u_i = u(\mu, Y_i) = \frac{Y_i - \mu}{\psi V(\mu)}$$

E' facile dimostrare che le u_i soddisfano le seguenti condizioni

$$\begin{aligned} E(u_i) &= 0 \\ \text{var}(u_i) &= \frac{\psi V(\mu)}{\psi^2 V(\mu)^2} = \frac{1}{\psi V(\mu)} = -E\left(\frac{\partial u_i}{\partial \mu}\right) \end{aligned}$$

Analogamente, sia

$$s_i = s(Y_i; \mu) = \frac{\partial \log(L(\mu; Y_i))}{\partial \mu}$$

dove $L(\mu; Y_i)$ è la funzione di verosimiglianza, incognita, delle Y_i . La funzione $s(\cdot)$ è anche nota, nella letteratura anglosassone, con il nome di *score function*. La teoria dell'inferenza (si veda, fra gli altri, Azzalini 2001) ci suggerisce che

$$\begin{aligned} E(s_i) &= 0 \\ \text{var}(s_i) &= -E\left(\frac{\partial s_i}{\partial \mu}\right) \end{aligned}$$

in maniera analoga a quanto riscontrato per le u_i . In virtù di tali analogie la funzione $u(\cdot)$ viene definita *quasi-score*.

La logica alla base della stima di QMV risiede nell'intuizione che, se la funzione quasi-score $u(\cdot)$ si comporta in maniera simile alla derivata della funzione di log-verosimiglianza, ci si attende che l'integrale di $u(\cdot)$

$$Q(\mu, Y_i) = \int_{-\infty}^{\mu} u(z, Y_i) dz = \int_{-\infty}^{\mu} \frac{Y_i - z}{\psi V(z)} dz$$

si comporti in maniera simile ad una funzione di log-verosimiglianza. Si definisce quindi la funzione di *Quasi (log) Verosimiglianza* aggregata come

$$\bar{Q}(\mu) = \sum_{i=1}^n Q(\mu, Y_i) \quad (27)$$

Una stima di QMV di μ può essere quindi ottenuta massimizzando la (27) rispetto alla media μ . A tal fine, è sufficiente verificare che siano soddisfatte le condizioni del primo ordine ottenute uguagliando a zero la derivata della funzione di quasi verosimiglianza rispetto a μ :

$$\frac{\partial \bar{Q}(\mu)}{\partial \mu} = \sum_{i=1}^n u_i = \frac{1}{\psi} \sum_{i=1}^n \frac{Y_i - \mu}{V(\mu)} = 0 \quad (28)$$

Si noti che, nell'equazione (28), il valore della costante ψ non influisce in alcun modo sulla soluzione del sistema ovvero sulla stima dei coefficienti di

regressione. L'equazione (28) può quindi essere ulteriormente semplificata omettendo la costante $1/\psi$

$$\sum_{i=1}^n \frac{Y_i - \mu}{V(\mu)} = 0 \quad (29)$$

Si noti che, se si pone $V(\mu)=1$, è facile dimostrare che la funzione di quasi verosimiglianza $Q(\mu; Y_i)$ è data da

$$Q(\mu; Y_i) = -\frac{(Y_i - \mu)^2}{2\psi} = -\frac{(Y_i - \mu)^2}{2\sigma^2}$$

In forma aggregata, diviene

$$\bar{Q}(\mu) = -\frac{1}{2\sigma^2} \sum_{i=1}^n (Y_i - \mu)^2 \quad (30)$$

che, a meno di una costante, è uguale alla funzione di log-verosimiglianza indotta dalle ipotesi di normalità delle Y_i . In tal caso, quindi, la stima di QMV della media μ viene ottenuta massimizzando una funzione di log-verosimiglianza di tipo normale pur sempre prescindendo dall'ipotesi che la distribuzione, ricordiamo incognita, delle Y_i sia effettivamente normale. Da un diverso punto di vista, massimizzare la (30) equivale a minimizzare la somma dei quadrati degli errori il che implica che la stime di QMV basate sulla (30) si riconducono alle stime dei Minimi Quadrati Ordinari (OLS).

Le stesse argomentazioni sviluppate con riferimento al semplice caso della stima della media possono essere facilmente estese al caso di un modello di regressione non-lineare. A tal fine, nella (26) si ponga:

$$\mu = \mu_i = E(Y_i) = \mu(\mathbf{X}_i; \boldsymbol{\beta})$$

dove $\mathbf{X}_i = (X_{i1}, \dots, X_{ip})$ è un vettore di variabili esplicative opportunamente scelte e $\boldsymbol{\beta}$ un vettore di parametri incogniti da stimare. Nel caso del modello di regressione lineare la funzione $\mu(\cdot)$ viene ad essere data da

$$\mu(\mathbf{X}_i; \boldsymbol{\beta}) = \beta_0 + \sum_{j=1}^{p-1} \beta_j X_{ij}$$

con $\beta = (\beta_0, \beta_1, \dots, \beta_{p-1})$. La nostra trattazione non è tuttavia in alcun modo vincolata all'ipotesi di linearità del modello e la procedura che viene di seguito discussa può essere efficacemente applicata anche a modelli di natura non lineare. In generale, le stime di QMV dei parametri β_j possono essere ottenute risolvendo il sistema di p equazioni in p incognite che si ottiene uguagliando a zero le derivate della funzione di quasi-verosimiglianza aggregata, cioè le u_i , rispetto ai singoli parametri β_j

$$\frac{1}{\Psi} \sum_{i=1}^n \frac{(X_i - \mu_i)}{V(\mu_i)} \frac{\partial \mu_i}{\partial \beta_j} = 0 \quad (31)$$

per $j = 0, \dots, p-1$. Come già riscontrato nel caso della (28), il valore della costante $1/\Psi$ non influisce sulle stime dei coefficienti di regressione β_j .

Ancora una volta, inoltre, se si pone $V(\mu_i) = 1$, è immediato notare come le (31) divengano le equazioni normali dei minimi quadrati e le stime di QMV coincidano con quelle OLS. L'approccio presentato non permette di ottenere una stima della varianza σ^2 che può tuttavia essere derivata ricorrendo al metodo dei momenti. E' però possibile generalizzare il metodo della QMV al caso in cui l'obiettivo sia non solo quello di stimare i parametri da cui dipende il livello medio μ della variabile di risposta Y ma anche la sua varianza σ^2 ed eventuali parametri incogniti che ne caratterizzano la specificazione, nel caso in cui il valore di σ^2 venga a dipendere da un'insieme di regressori esogeni o endogeni. Come vedremo, quest'ultima situazione ricorre frequentemente nell'analisi di dati finanziari. In questo campo, infatti, la varianza dei rendimenti su un dato attivo viene utilizzata come misura del rischio di mercato collegato alla detenzione di quel particolare attivo.

In conclusione è importante notare come, in base a quanto emerso dalla discussione finora presentata, non esista un modo univoco di costruire una funzione di quasi verosimiglianza. Diverse specificazioni della *funzione di varianza* $V(\mu)$ possono infatti condurre a diverse funzioni di quasi verosimiglianza.

Sul piano teorico, Heyde (1997) dimostra che è possibile identificare una funzione di quasi verosimiglianza ottimale massimizzando la correlazione fra le sue derivate, cioè le u_i , e la funzione score, cioè le s_i . Sul piano pratico, l'implementazione di questo criterio è però limitata dalla considerazione che, nella maggior parte dei casi in cui si è costretti a ricorrere alla stima di QMV, la

forma della funzione di verosimiglianza e, quindi, della funzione score associata sono incognite. Per ulteriori dettagli ed approfondimenti sulle stime di QMV e sulle loro proprietà si rimanda il lettore interessato al volume di Heyde (1997) .

3.5 Alcuni test delle ipotesi basati sulla funzione di verosimiglianza

In questo paragrafo verranno presentati tre differenti approcci alla costruzione di tests basati sulla funzione di verosimiglianza. In particolare, la nostra attenzione si soffermerà sulle seguenti tipologie di tests:

- A) tests del rapporto di verosimiglianza
- B) tests dei moltiplicatori di Lagrange; tests di Wald.

E' importante sottolineare che questi tests sono di carattere generale e, in linea di principio, applicabili a qualsiasi modello per cui sia possibile definire e massimizzare, analiticamente o avvalendosi di metodi numerici, la funzione di verosimiglianza. Nei capitoli che seguono, le tecniche che vengono qui discusse verranno diffusamente applicate per la costruzione di tests di identificazione e diagnostica per modelli ad eteroschedasticità condizionata e nel campo del risk management.

A) TESTS DEL RAPPORTO DI VEROSIMIGLIANZA

Il principale fondamento teorico su cui si fonda la diffusa applicazione del test del rapporto di verosimiglianza è dato dal *Lemma di Neyman-Pearson*.

Sia X una variabile casuale continua con funzione di densità data da $f(x; \boldsymbol{\theta})$, dove con $\boldsymbol{\theta} \in \Theta$ si è indicato il vettore di parametri da cui dipende la $f(.,.)$. Si indichi poi con $\mathbf{x} = (x_1, x_2, \dots, x_n)$ un campione casuale osservato su X e si assuma di voler verificare il seguente sistema di *ipotesi semplici*:

$$\begin{cases} H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0 \\ H_1 : \boldsymbol{\theta} = \boldsymbol{\theta}_1 \end{cases} \quad (32)$$

La funzione di verosimiglianza è data da

$$L(\boldsymbol{\theta}; \mathbf{x}) = \prod_{i=1}^n f(x_i; \boldsymbol{\theta})$$

con $\mathbf{x} = (x_1, x_2, \dots, x_n)$. Un *test del Rapporto di Verosimiglianza Semplice* (RVS) per il sistema di ipotesi in (32) è basato sulla seguente regola di decisione:

$$\text{si rifiuta } H_0 \text{ se } \lambda_n < \lambda_\alpha \text{ e si accetta } H_0 \text{ se } \lambda_n \geq \lambda_\alpha \quad (33)$$

dove λ_n è il *rapporto di verosimiglianza semplice* dato da

$$\lambda_n = \lambda(\mathbf{x}) = \frac{\prod_{i=1}^n f(x_i; \theta_0)}{\prod_{i=1}^n f(x_i; \theta_1)} = \frac{L_0}{L_1}$$

e λ_α è una costante non negativa tale che $P = \{\lambda_n < \lambda_\alpha \mid H_0\} = 1 - \alpha$. Il Lemma di Neyman-Pearson (per una dimostrazione vedi, fra gli altri, Piccolo (2001)) ci suggerisce che il test del RVS è un test ottimale per il sistema di ipotesi in (32). La costante α rappresenta la probabilità di commettere l'errore rifiuto H_0 mentre H_0 è vera (errore del primo tipo).

Lemma (di Neyman e Pearson)

Per il sistema di ipotesi (32) sul vettore di parametri θ della funzione di densità $f(\mathbf{x}; \theta)$, il test RVS definito dalla (33) ha la *massima potenza* fra tutti i test con livello di significatività α . Dove per potenza si intende la probabilità di rifiutare H_0 ed H_0 è falsa.

Sebbene questo risultato assuma una indubbia rilevanza sul piano teorico, nelle pratiche applicazioni raramente ci si trova a dover lavorare con sistemi di ipotesi semplici del tipo in (32). Molto più frequentemente ci si trova a dover verificare dei sistemi di ipotesi *composte* del tipo

$$\begin{cases} H_0 : \theta \in \Theta_0 \\ H_1 : \theta \notin \Theta_0 \end{cases} \quad (34)$$

dove con $\Theta_0 \subseteq \Theta$ si è indicato quel sottoinsieme dello spazio parametrico tale che sia verificata l'ipotesi H_0 .

Una importante famiglia di test per la verifica di ipotesi composte del tipo (34) è basata sulla statistica del *Rapporto di Verosimiglianza Generalizzato* o anche, semplicemente, detta del *Rapporto di Verosimiglianza* (RV)

$$\lambda_n^* = \lambda^*(\mathbf{x}) = -2 \ln \left(\frac{\sup_{\theta \in \Theta_0} \prod_{i=1}^n f(x_i; \theta)}{\sup_{\theta \in \Theta} \prod_{i=1}^n f(x_i; \theta)} \right) = -2 \ln \left(\frac{ML_0}{ML_1} \right) = -2 \ln RV \quad (35)$$

Si noti che sarà sempre verificata la condizione $\lambda_n^* \geq 0$ dal momento che $0 \leq RV \leq 1$, essendo RV dato dal rapporto di due quantità non negative, ML_0 e ML_1 , con $ML_0 \leq ML_1$ per costruzione.

Un test del RV sarà quindi basato sulla seguente regola di decisione

$$\text{si rifiuti } H_0 \text{ se } \lambda_n^* > \lambda_\alpha^* \text{ e si accetti } H_0 \text{ se } \lambda_n^* \leq \lambda_\alpha^* \quad (36)$$

dove $P\{\lambda_n^* \leq \lambda_\alpha^*\} = 1 - \alpha$. Si può dimostrare sotto condizioni generali di regolarità e per campioni di grande dimensione che la distribuzione di λ_n^* è ben approssimata da una v.c. $\chi_{(g)}^2$, ovvero

$$\lambda_n^* \approx \chi_{(g)}^2$$

per $\chi_{(g)}^2 \rightarrow \infty$, dove g corrisponde al numero di parametri soggetti a vincolo sotto H_0 . Un caso particolare della (34), che ricorre molto frequentemente nelle applicazioni pratiche, è quello in cui si vuole sottoporre a verifica il sistema di ipotesi bidirezionale

$$\begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta \neq \theta_0 \end{cases} \quad (37)$$

In questo caso, la statistica del test del RV è data da

$$\lambda_n^* = \lambda^*(\mathbf{x}; \theta_0) = -2(\ell(\theta_0; \mathbf{x}) - \ell(\hat{\theta}; \mathbf{x}))$$

che, per n sufficientemente grande, sotto H_0 , si distribuisce, in maniera approssimata, come una v.c. $\chi_{(g)}^2$ dove g corrisponde alla dimensione del vettore $\boldsymbol{\theta}$. Nelle applicazioni molto spesso si ha necessità di verificare delle ipotesi formulate su un sottoinsieme dei parametri del modello, $\boldsymbol{\psi} \subseteq \boldsymbol{\theta}$, piuttosto che sull'intero vettore dei parametri. In questa situazione il vettore di parametri $\boldsymbol{\theta}$ può essere partizionato in due componenti: il vettore dei parametri *di interesse*, $\boldsymbol{\psi}$, ed il vettore dei parametri cosiddetti *di disturbo*, $\boldsymbol{\omega}$. In questo caso, la statistica del test del RV per il sistema di ipotesi

$$\begin{cases} H_0 : \boldsymbol{\psi} = \boldsymbol{\psi}_0 \\ H_1 : \boldsymbol{\psi} \neq \boldsymbol{\psi}_0 \end{cases} \quad (38)$$

viene è data da

$$\lambda_n^* = \lambda^*(\mathbf{x}; \boldsymbol{\psi}_0 | \boldsymbol{\omega}) = -2(\ell(\boldsymbol{\psi}_0, \hat{\boldsymbol{\omega}}_{\boldsymbol{\psi}_0}; \mathbf{x}) - \ell(\hat{\boldsymbol{\psi}}, \hat{\boldsymbol{\omega}}; \mathbf{x}))$$

dove $\hat{\boldsymbol{\omega}}_{\boldsymbol{\psi}_0}$ rappresenta la stima di massima verosimiglianza di $\boldsymbol{\omega}$ ottenuta fissando $\boldsymbol{\psi} = \boldsymbol{\psi}_0$ mentre $\hat{\boldsymbol{\psi}}$ e $\hat{\boldsymbol{\omega}}$ sono le stime di massima verosimiglianza, non condizionate, di $\boldsymbol{\psi}$ e $\boldsymbol{\omega}$, rispettivamente. Come nel caso precedente, per N sufficientemente grande e sotto H_0 , $\lambda_n^* \approx \chi_{(g)}^2$ dove g corrisponde alla dimensione del vettore $\boldsymbol{\psi}$. La funzione $\ell(\boldsymbol{\psi}, \hat{\boldsymbol{\omega}}_{\boldsymbol{\psi}}; \mathbf{x})$ viene anche detta funzione di log-verosimiglianza profilo per il vettore di parametri $\boldsymbol{\psi}$.

Esempio (v.c. di Bernoulli)

Sia $\mathbf{x} = (x_1, x_2, \dots, x_n)$ un campione casuale osservato su una v.c. X di Bernoulli con funzione di distribuzione di probabilità data da

$$\pi(x) = p^x (1-p)^{1-x}$$

La funzione di verosimiglianza è quindi data

$$L(p; \mathbf{x}) = \prod_{i=1}^n \pi(x_i) = \prod_{i=1}^n p^{x_i} (1-p)^{1-x_i} = p^{\sum_{i=1}^n x_i} (1-p)^{n - \sum_{i=1}^n x_i}$$

dalla quale, passando ai logaritmi, si ottiene la funzione di log-verosimiglianza

$$\ell(\mathbf{p}; \mathbf{x}) = \ln L(\mathbf{p}; \mathbf{x}) = \sum_{i=1}^n x_i \ln(p) + \left(n - \sum_{i=1}^n x_i\right) \ln(1-p) \quad (39)$$

Differenziando la (39) rispetto a p ed uguagliando a zero la derivata così ottenuta si ha

$$\frac{\partial \ell(\mathbf{p}; \mathbf{x})}{\partial p} = \frac{\sum_{i=1}^n x_i}{p} - \frac{n - \sum_{i=1}^n x_i}{1-p} = 0$$

Risolvendo per p , si ottiene lo stimatore di massima verosimiglianza

$$\hat{p} = \frac{\sum_{i=1}^n x_i}{n} = \frac{n_1}{n} \quad (40)$$

Che, in questo caso, coincide con lo stimatore basato sul metodo dei momenti le cui proprietà sono state già analizzate nella sezione 3.1.

Si assuma ora di voler verificare il sistema di ipotesi $H_0: p = p_0$ in alternativa a $H_1: p \neq p_0$, utilizzando il test del rapporto di verosimiglianza. Il valore del massimo della funzione di log-verosimiglianza sotto l'ipotesi alternativa H_1 , ovvero in assenza di vincoli su p , può essere calcolato sostituendo la (40) nella (39)

$$\begin{aligned} \ln(\text{ML}_1) &= n_1 \ln\left(\frac{n_1}{n}\right) + (n_1 - n) \ln\left(1 - \frac{n_1}{n}\right) = \\ &= n_1 \ln(n_1) + (n - n_1) \ln(n - n_1) - n \ln(n). \end{aligned}$$

Il massimo in corrispondenza dell'ipotesi nulla H_0 viene invece ottenuto sostituendo il valore p_0 nella (39):

$$\ln(\text{ML}_0) = n_1 \ln(p_0) + (n - N_1) \ln(1 - p_0)$$

La statistica del RV è data da

$$\lambda_n^*(p) = -2(\ln(\text{ML}_0) - \ln(\text{ML}_1))$$

che, sotto H_0 e per n grande ($n \rightarrow \infty$), risulterà approssimativamente distribuita come una v.c. $\chi_{(1)}^2$. Si rifiuterà H_0 , al livello di significatività α , per $\lambda_n^*(p) > \chi_{(1;1-\alpha)}^2$.

□

Esempio (v.c. Normale)

Sia $\mathbf{x} = (x_1, x_2, \dots, x_n)$ un campione casuale osservato su una v.c. Normale $X \sim N(\mu, \sigma^2)$. Si assuma di voler verificare il sistema di ipotesi: $H_0: \mu = \mu_0$ in alternativa a $H_1: \mu \neq \mu_0$. Applicando la (6) si ottiene

$$-2 \ln(\text{ML}_1) = n \log(2\pi \hat{\sigma}^2) + \sum_{i=1}^n \left(\frac{x_i - \hat{\mu}}{\hat{\sigma}} \right)^2 \quad (41)$$

dove i valori di $\hat{\mu}$ e $\hat{\sigma}$ possono essere ottenuti applicando le formule nell'equazione (7). Similmente

$$-2 \ln(\text{ML}_0) = n \log(2\pi \hat{\sigma}_0^2) + \sum_{i=1}^n \left(\frac{x_i - \mu_0}{\hat{\sigma}_0} \right)^2 \quad (42)$$

dove

$$\hat{\sigma}_0^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu_0)^2.$$

La statistica del RV viene quindi ottenuta sottraendo la (41) dalla (42)

$$\lambda_n^*(\mu | \sigma^2) = -2 \ln(\text{ML}_0) + 2 \ln(\text{ML}_1).$$

Sotto H_0 e per n grande ($n \rightarrow \infty$), $\lambda_n^*(\mu | \sigma^2) \approx \chi_{(1)}^2$. Si rifiuterà H_0 , al livello di significatività α , per $\lambda_n^*(\mu | \sigma^2) > \chi_{(1;1-\alpha)}^2$.

□

Esempio (regressione lineare semplice con errori normali)

Sia $\mathbf{Y}' = (y_1, y_2, \dots, y_n)$ un campione casuale osservato su un dato fenomeno con $y_i \sim N(\mu_i, \sigma^2)$ e $\mu_i = b_0 + b_1 x_i$ dove con x_i si è indicato il

valore osservato di una data variabile esplicativa X in corrispondenza dell'unità i -esima del campione. Ai fini dell'analisi si assume che le x_i siano predeterminate ovvero che non abbiano natura stocastica. In pratica, quanto affermato equivale ad assumere che le y_i siano spiegate da un modello di regressione lineare semplice del tipo

$$y_i = b_0 + b_1 x_i + e_i$$

con $e_i \stackrel{\text{iid}}{\sim} N(0, \sigma^2)$, per $i=1, \dots, n$. La funzione di densità delle y_i è data da

$$f(y_i; b_0, b_1, \sigma^2 | x_i) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2}\left(\frac{y_i - b_0 - b_1 x_i}{\sigma}\right)^2\right\}$$

per $i=1, \dots, n$. La funzione di verosimiglianza che ne deriva è

$$\begin{aligned} L(b_0, b_1, \sigma^2; \mathbf{y} | \mathbf{x}) &= \prod_{i=1}^n f(y_i; b_0, b_1, \sigma^2 | x_i) = \\ &= \prod_{i=1}^n \left(\frac{1}{2\pi\sigma^2}\right)^{N/2} \exp\left\{-\frac{1}{2}\sum_{i=1}^n \left(\frac{y_i - b_0 - b_1 x_i}{\sigma}\right)^2\right\} \end{aligned}$$

Passando ai logaritmi si ottiene

$$-2\ell(b_0, b_1, \sigma^2; \mathbf{y} | \mathbf{x}) = n \log(2\pi\sigma) + \sum_{i=1}^n \left(\frac{y_i - b_0 - b_1 x_i}{\sigma}\right)^2$$

Nelle applicazioni pratiche, risulta spesso di interesse sottoporre a verifica il sistema di ipotesi: $H_0: b = 0$ in alternativa a $H_1: b \neq 0$ il che equivale a verificare se il livello medio di un dato fenomeno, il livello di output di un processo produttivo o il rendimento atteso di un titolo, è significativamente correlato con il livello di una data variabile esplicativa X . E' opportuno notare che la specificazione di X non è soggetta a vincolo alcuno: essa potrebbe essere una variabile continua come anche una variabile dicotomica che esprime l'appartenenza o meno ad un dato gruppo di osservazioni assumendo valore 1

se l' i -esima osservazione presenta una determinata caratteristica e 0 altrimenti (ad esempio, processi produttivi che utilizzano una data tecnologia o titoli azionari di un determinato settore). Sotto H_0 , il massimo della funzione di log-verosimiglianza è proporzionale a:

$$-2 \ln(\text{ML}_0) = n \log(2\pi\hat{\sigma}_0^2) + \sum_{i=1}^n \left(\frac{y_i - \hat{\mu}_y}{\hat{\sigma}_0} \right)^2 \quad (43)$$

dove

$$\hat{\sigma}_0^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{\mu}_y)^2; \quad \hat{\mu}_y = \hat{b}_0 = \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$$

Sotto H_1 , invece, si ottiene:

$$-2 \ln(\text{ML}_1) = n \log(2\pi\hat{\sigma}^2) + \sum_{i=1}^n \left(\frac{y_i - \hat{b}_0 - \hat{b}_1 x_i}{\hat{\sigma}} \right)^2 \quad (44)$$

dove

$$\hat{b}_1 = \frac{\sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2}; \quad \hat{b}_0 = \bar{y} - \hat{b}_1 \bar{x}; \quad \hat{\sigma}^2 = \frac{\sum_{i=1}^n (y_i - \hat{b}_0 - \hat{b}_1 x_i)^2}{n}$$

sono gli stimatori di massima verosimiglianza di (b_0, b_1, σ^2) ottenuti risolvendo il seguente sistema di equazioni rispetto a (b_0, b_1, σ^2) :

$$\begin{aligned} \frac{1}{\sigma} \sum_{i=1}^n (y_i - b_0 - b_1 x_i) &= 0 \\ \frac{1}{\sigma} \sum_{i=1}^n (y_i - b_0 - b_1 x_i) x_i &= 0 \\ -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^n (y_i - b_0 - b_1 x_i)^2 &= 0 \end{aligned}$$

□

Analogamente a quanto riscontrato nell'esempio precedente, la statistica del RV è data dalla differenza fra le quantità calcolate nelle equazioni (42) e (41)

$$\lambda_n^*(b_1 | b_0, \sigma^2) = -2 \ln(\text{ML}_0) + 2 \ln(\text{ML}_1)$$

Se H_0 è verificata e per n grande ($n \rightarrow \infty$), $\lambda_n^*(b_1 | b_0, \sigma^2) \approx \chi_{(1)}^2$. Si rifiuterà quindi H_0 , al livello di significatività α , per $\lambda_n^*(b_1 | b_0, \sigma^2) > \chi_{(1,1-\alpha)}^2$

□

Esempio (regressione lineare multipla con errori normali)

Il test illustrato nell'esempio precedente può essere facilmente generalizzato al caso in cui si assume che il livello medio delle y_i possa potenzialmente dipendere da una pluralità di variabili esplicative predeterminate ($x_{1i}, x_{2i}, \dots, x_{pi}$), con $i=1, \dots, n$. Generalizzando quanto precedentemente osservato per il caso di una singola variabile esplicativa, è semplice dimostrare che il modello probabilistico descritto equivale ad un modello di regressione lineare multipla del tipo

$$y_i = b_0 + b_1 x_{1i} + \dots + b_p x_{pi} + e_i$$

con $e_i \stackrel{\text{iid}}{\sim} N(0, \sigma^2)$, per $i=1, \dots, n$. La funzione di log-verosimiglianza è data da

$$-2\ell(\mathbf{b}, \sigma^2; \mathbf{y} | \mathbf{x}) = n \log(2\pi\sigma) + \sum_{i=1}^n \left(\frac{y_i - \mathbf{x}_i' \mathbf{b}}{\sigma} \right)^2$$

dove $\mathbf{x}_i' = (1, x_{1i}, x_{2i}, \dots, x_{pi})$ e $\mathbf{b}' = (b_0, b_1, \dots, b_p)$. Le stime di massima verosimiglianza dei coefficienti \mathbf{b} e σ^2 possono essere ottenute risolvendo per \mathbf{b} e σ^2 il seguente sistema di $(p+1)$ equazioni lineari in $(p+1)$ incognite:

$$\frac{1}{\sigma} \sum_{i=1}^n (y_i - \mathbf{x}_i' \mathbf{b}) \mathbf{x}_j = 0, \quad j=1, 2, \dots, p$$

$$-\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^n (y_i - \mathbf{x}_i' \mathbf{b})^2 = 0$$

le cui soluzioni sono date da

$$\hat{\mathbf{b}} = \left(\sum_{i=1}^n \mathbf{x}_i \mathbf{x}_i' \right)^{-1} \sum_{i=1}^n \mathbf{x}_i y_i ; \quad \hat{\sigma}^2 = \frac{\sum_{i=1}^n (y_i - \mathbf{x}_i' \hat{\mathbf{b}})^2}{n}$$

Il massimo della log-verosimiglianza sotto H_0 si calcola in maniera analoga a quanto illustrato nella (41) mentre il massimo della log-verosimiglianza sotto H_1 è proporzionale a

$$-2 \ln(\text{ML}_1) = n \log(2\pi \hat{\sigma}^2) + \sum_{i=1}^n \left(\frac{y_i - \mathbf{x}_i' \hat{\mathbf{b}}}{\hat{\sigma}} \right)^2$$

La statistica del RV viene calcolata nel modo consueto

$$\lambda_n^*(b_1, \dots, b_p \mid b_0, \sigma^2) = -2 \ln(\text{ML}_0) + 2 \ln(\text{ML}_1)$$

Per campioni di dimensione elevata, se l'ipotesi nulla è verificata, la sua distribuzione sarà ben approssimata da una v.c. $\chi_{(p)}^2$ ovvero

$$\lambda_N^*(b_1, \dots, b_p \mid b_0, \sigma^2) \approx \chi_{(p)}^2$$

□

La procedura proposta può essere facilmente generalizzata al caso in cui si voglia sottoporre a verifica solo un sottoinsieme di coefficienti di regressione. In questo particolare caso, il ricercatore vuole sottoporre a verifica l'ipotesi che un dato gruppo di variabili di regressione sia congiuntamente significativo nello spiegare le variazioni della y . In pratica si viene ad effettuare un test per l'omissione di variabili rilevanti dal modello o, da un diverso punto di vista, per l'inclusione di variabili ridondanti. La distribuzione della statistica test del RV, in quest'ultimo caso, sarà ben approssimata da una v.c. $\chi_{(p_0)}^2$ con $p_0 \leq p$ pari al numero di coefficienti da sottoporre a test.

B) TEST DI WALD E TEST DEI MOLTIPLICATORI DI LAGRANGE (LM)

Presentiamo ora due statistiche test basate sulla funzione di verosimiglianza che possono essere utilizzate in alternativa al test del rapporto di verosimiglianza: il test di Wald ed il test dei Moltiplicatori di Lagrange (LM). Nello specifico, si assuma di voler sottoporre a verifica il sistema di ipotesi ad alternativa bidirezionale definite nelle (37). Il test del RV può essere interpretato come la distanza fra il valore ipotizzato per il vettore dei parametri (θ_0) ed il valore di θ maggiormente compatibile con l'evidenza empirica

contenuta nei dati, ovvero $\hat{\theta}$. La metrica scelta per misurare tale distanza si basa sul valore della funzione di verosimiglianza in corrispondenza dei due vettori. Al pari del test del RV, la statistica del test di Wald può anche essere interpretata come una misura di distanza fra $\hat{\theta}$ e θ_0 , seppure basata su una metrica differente. La principale differenza rispetto al test del rapporto di verosimiglianza può essere individuata nel fatto che la metrica utilizzata per misurare la distanza fra $\hat{\theta}$ e θ_0 viene ora definita direttamente sullo spazio parametrico piuttosto che sulla funzione di verosimiglianza. In particolare, la statistica test di Wald è data dalla seguente forma quadratica

$$W_N(\theta_0) = (\hat{\theta} - \theta_0)' \mathbf{I}(\hat{\theta})(\hat{\theta} - \theta_0)$$

dove

$$\mathbf{I}(\hat{\theta}) = -E \left[\frac{\partial^2 \ell(\theta; \mathbf{x})}{\partial \theta \partial \theta'} \right]_{\theta = \hat{\theta}}$$

rappresenta la matrice (p, p) di informazione attesa calcolata in corrispondenza dello stimatore di massima verosimiglianza. Al riguardo è interessante notare come il test di Wald possa essere interpretato come un test sulla media. Infatti si può dimostrare, sulla base della teoria classica della verosimiglianza, per valori di N sufficientemente elevati, che la distribuzione campionaria di $\hat{\theta}$ tende ad una normale multivariata con media θ^* e matrice di varianze e covarianza data da $\mathbf{I}(\theta^*)^{-1}$. In pratica, si ha

$$\hat{\theta} \approx N(\theta^*, \mathbf{I}(\theta^*)^{-1})$$

dove, con θ^* si è indicato il valore “vero” del vettore di parametri θ . Non conoscendo il valore di θ^* , la matrice di informazione attesa $\mathbf{I}(\theta^*)^{-1}$ può essere consistentemente stimata sostituendo $\hat{\theta}$ a θ^* . Il test di Wald può quindi essere inteso come un test sulla media di $\hat{\theta}$ basato sulla distribuzione asintotica dello stimatore. Da quanto detto segue che, nel semplice caso in cui il vettore dei parametri sia dato da uno scalare ($p=1$) e sotto alcune condizioni generali di regolarità, per n sufficientemente grande si avrà

$$(\hat{\theta} - \theta_0) \sqrt{\mathbf{I}(\hat{\theta})} \approx N(0, 1)$$

per cui

$$W_n(\theta_0) = (\hat{\theta} - \theta_0)^2 I(\hat{\theta}) \approx \chi_{(1)}^2$$

Generalizzando quanto detto al caso di un vettore di parametri p -dimensionale, si può dimostrare che sotto H_0 il test del rapporto di verosimiglianza ed il test di Wald sono equivalenti ovvero tendono ad avere la medesima distribuzione asintotica. Quindi, per valori di N sufficientemente elevati, si avrà che

$$W_n(\theta_0) \approx \chi_{(p)}^2$$

dove p coincide con la dimensione di θ . Si rifiuterà H_0 al livello α se risulta $W_n > \chi_{(p; 1-\alpha)}^2$.

Nel caso in cui siano presenti dei parametri di disturbo, il sistema di ipotesi di riferimento è dato dalle (16). In tal caso la statistica test di Wald può essere ancora costruita nel modo appena descritto facendo però riferimento alla funzione di log-verosimiglianza profilo rispetto al vettore dei parametri di interesse ψ piuttosto che alla funzione di log-verosimiglianza complessiva. La distribuzione asintotica sarà ancora data da una v.c. χ^2 con numero di gradi di libertà pari alla dimensione di ψ .

Una seconda alternativa al test del RV è rappresentata dal test LM o score test. La statistica del test LM è data da

$$LM_n(\theta_0) = \mathbf{s}(\theta_0)' \mathbf{I}(\theta_0)^{-1} \mathbf{s}(\theta_0)$$

dove con

$$\mathbf{s}(\theta_0) = \sum_{i=1}^n s_i(\theta_0) = \sum_{i=1}^n \left. \frac{\partial}{\partial \theta} \log(f(x_i; \theta)) \right|_{\theta=\theta_0}$$

si è indicato il vettore score valutato assumendo verificati i vincoli imposti dall'ipotesi nulla ovvero ponendo $\theta = \theta_0$. Come le due precedenti statistiche test (RV e Wald) anche il test LM può essere interpretato come una misura di distanza fra $\hat{\theta}$ e θ_0 . La metrica utilizzata viene però ora definita in termini della derivata della funzione di verosimiglianza: il vettore score. Si noti poi che, se θ_0 costituisce il “vero” valore del parametro, si può dimostrare che $E(\mathbf{s}(\theta_0)) = \mathbf{0}$.

Inoltre, per il teorema del limite centrale, per campioni di numerosità sufficientemente elevata, si può dimostrare che il vettore score ha la seguente distribuzione approssimata

$$\mathbf{s}(\boldsymbol{\theta}_0) \approx N(\mathbf{0}, \mathbf{I}(\boldsymbol{\theta}_0))$$

ne segue che, al pari del test di Wald, anche il test LM, può essere interpretato come un test sulla media di una variabile casuale normale. A differenza del test di Wald, il valore di $\mathbf{I}(\boldsymbol{\theta}^*)$ viene però approssimato in termini di $\boldsymbol{\theta}_0$ piuttosto che di $\hat{\boldsymbol{\theta}}$. Si può dimostrare infine che, se i vincoli imposti sotto l'ipotesi nulla H_0 sono verificati, il test LM è asintoticamente equivalente ai precedenti e, quindi, per un vettore di parametri p -dimensionale e per valori di n sufficientemente elevati, si ha che $LM_n(\boldsymbol{\theta}_0) \approx \chi_{(p)}^2$ e pertanto si rifiuterà H_0 al livello α per valori della statistica $LM > \chi_{(p; 1-\alpha)}^2$. Anche in questo caso, analogamente a quanto discusso per il test di Wald ed il test del RV, in presenza di parametri di disturbo, la statistica del test LM può essere costruita facendo riferimento alla funzione di log-verosimiglianza profilo rispetto a ψ .

La disponibilità di tre statistiche test caratterizzate dalla medesima distribuzione asintotica può sicuramente contribuire a creare delle incertezze nei practitioners che hanno necessità di utilizzare questi strumenti in diversi contesti applicativi. In generale, va senz'altro sottolineato che non è possibile stabilire un ordinamento univoco delle tre statistiche presentate (RV, Wald, LM) sulla base della potenza del test e della velocità di convergenza alla distribuzione nulla. E' però possibile fare delle considerazioni di carattere pratico che possono fungere da guida nella scelta della statistica test da applicare in uno specifico problema.

Il principale vantaggio della statistica di Wald va senz'altro individuato nella sua facile interpretabilità: una metrica basata sullo spazio parametrico risulta infatti maggiormente intuitiva di una basata sulla funzione di verosimiglianza o sulle sue derivate. D'altra parte la statistica di Wald presenta il limite di non essere invariante rispetto a riparametrizzazioni del modello. Quindi, diverse rappresentazioni parametriche di uno stesso modello statistico possono potenzialmente condurre a risultati del test differenti. Il test LM presenta, invece, il vantaggio di non richiedere la stima del modello non vincolato, ovvero quello dedotto sotto H_1 , che, in alcuni casi, può risultare particolarmente onerosa sul piano computazionale o, in casi estremi, non praticabile.

Di seguito riportiamo due esempi di applicazione dei tests Wald ed LM. Il

primo si riferisce alla stima della probabilità di successo in una variabile casuale di Bernoulli mentre il secondo si riferisce alla stima della media di una variabile casuale normale nell'ipotesi che la varianza sia nota. In questo secondo caso, vedremo come le tre statistiche test (RV, Wald ed LM) siano identiche a conferma delle forti connessioni e similarità esistenti fra questi tre approcci alternativi alla verifica di ipotesi.

Esempio (variabile casuale di Bernoulli).

Si consideri un campione casuale $\mathbf{x} = (x_1, x_2, \dots, x_n)$ osservato su una v.c. X di Bernoulli con funzione di distribuzione di probabilità data da

$$\pi(x) = p^x (1-p)^{1-x}$$

e sottoponiamo a verifica $H_0: p = p_0$ verso $H_1: p \neq p_0$ utilizzando il test di Wald ed il test LM. A tal fine è necessario calcolare le derivate prima e seconda della funzione di log-verosimiglianza rispetto al parametro incognito p . Dall'esempio 1 ricordiamo che la derivata della funzione di log-verosimiglianza, funzione *score*, è data da

$$\ell'(\mathbf{x}; p) = s(p) = \frac{\sum_{i=1}^n x_i}{p} - \frac{n - \sum_{i=1}^n x_i}{1-p} = \frac{n\hat{p}}{p} - \frac{n - n\hat{p}}{1-p} = n \frac{\hat{p} - p}{p(1-p)}$$

mentre la derivata seconda è data da

$$\ell''(\mathbf{x}; p) = -\frac{\sum_{i=1}^n x_i}{p^2} - \frac{n - \sum_{i=1}^n x_i}{(1-p)^2}$$

Possiamo quindi calcolare l'informazione attesa di Fisher come

$$\begin{aligned} I(p) &= E(-\ell''(\mathbf{x}; p)) = \frac{\sum_{i=1}^n E(x_i)}{p^2} + \frac{n - \sum_{i=1}^n E(x_i)}{(1-p)^2} = \frac{np}{p^2} + \frac{n(1-p)}{(1-p)^2} \\ &= \frac{n}{p} + \frac{n}{1-p} = \frac{n}{p(1-p)} \end{aligned}$$

La statistica di Wald è quindi data da

$$W_n(p_0) = (\hat{p} - p_0)^2 I(\hat{p}) = \frac{n(\hat{p} - p_0)^2}{\hat{p}(1 - \hat{p})}$$

che, al crescere di n , tende in distribuzione ad una v.c. $\chi_{(1)}^2$, mentre la statistica del test LM è data da

$$LM_n(p_0) = \frac{s(p_0)^2}{I(p_0)} = \left(\frac{n(\hat{p} - p_0)}{p_0(1 - p_0)} \right)^2 \frac{p_0(1 - p_0)}{N} = \frac{n(\hat{p} - p_0)^2}{p_0(1 - p_0)}$$

che, se H_0 è verificata, come la precedente, al crescere di n , tende in distribuzione ad una v.c. $\chi_{(1)}^2$. Si noti che l'unica differenza fra le due statistiche sta nel denominatore dove, nel test di Wald, figura lo stimatore di MV, \hat{p} , mentre nel test LM ritroviamo p_0 . Dal momento che, al crescere di n , sotto H_0 , \hat{p} tende al valore "vero" del parametro, ovvero p_0 , risulta pressoché immediato dimostrare l'equivalenza asintotica delle due statistiche test con riferimento alla loro distribuzione nulla.

□

Esempio (variabile casuale Normale con σ^2 nota)

Si consideri un campione casuale $\mathbf{x} = (x_1, x_2, \dots, x_n)$ osservato su una v.c. Normale $X \sim N(\mu, \sigma^2)$ dove si assume sia $\sigma^2 = \theta$ dove θ è un valore noto. Sottoponiamo a verifica il sistema di ipotesi $H_0: \mu = \mu_0$ in alternativa a $H_1: \mu \neq \mu_0$. Si noti che

$$ML_0 = \ell(\mu_0) = -\frac{\sum_{i=1}^n (x_i - \mu_0)^2}{2\sigma^2}$$

$$ML_1 = \ell(\hat{\mu}) = -\frac{\sum_{i=1}^n (x_i - \hat{\mu})^2}{2\sigma^2}$$

con $\hat{\mu} = \bar{X}$, dalle quali segue che la statistica del test del RV può essere calcolata come

$$\begin{aligned}
\lambda_n^*(\mu_0) &= 2(\text{ML}_1 - \text{ML}_0) = \frac{1}{\sigma^2} \left(-\sum_{i=1}^n (x_i - \hat{\mu})^2 + \sum_{i=1}^n (x_i - \mu_0)^2 \right) = \\
&= \frac{1}{\sigma^2} \sum_{i=1}^n (-x_i^2 - \hat{\mu}^2 + 2x_i\hat{\mu} + x_i^2 + \mu_0^2 - 2x_i\mu_0) = \\
&= \frac{n}{\sigma^2} (-\hat{\mu}^2 + 2\hat{\mu}^2 - 2\mu_0\hat{\mu} + \mu_0^2) = \frac{n}{\sigma^2} (\hat{\mu} - \mu_0)^2
\end{aligned}$$

La statistica del test di Wald è invece data da

$$W_n(\mu_0) = (\hat{\mu} - \mu_0)^2 I(\hat{\mu}) \quad (23)$$

È semplice dimostrare che

$$\ell'(\mathbf{x}; \mu) = s(\mathbf{x}; \mu) = \frac{\sum_{i=1}^n (x_i - \mu)^2}{\sigma^2} = \frac{n(\hat{\mu} - \mu)}{\sigma^2}$$

da cui segue che

$$\ell''(\mathbf{x}; \mu) = -\frac{n}{\sigma^2}; \quad I(\mu) = \frac{n}{\sigma^2}, \quad \forall \mu$$

Sostituendo nella (23) si ottiene

$$W_n(\mu_0) = \frac{N(\hat{\mu} - \mu_0)^2}{\sigma^2}$$

La statistica del test LM può essere quindi calcolata come

$$\text{LM}_n(\mu_0) = \frac{s(\mu_0)^2}{I(\mu_0)} = \left(\frac{n(\hat{\mu} - \mu_0)}{\sigma^2} \right)^2 \frac{\sigma^2}{n} = \frac{n(\hat{\mu} - \mu_0)^2}{\sigma^2}$$

ovvero le tre statistiche test coincidono

$$\lambda_n^*(\mu_0) = W_n(\mu_0) = \text{LM}_n(\mu_0)$$

Un'ulteriore considerazione da farsi è che, in questo caso particolare, le tre

statistiche test risultano tutte distribuite come una v.c. $\chi_{(1)}^2$ anche per campioni di numerosità finita. In altre parole, il risultato relativo alla distribuzione nulla dei test è di tipo esatto e non asintotico.

□

3.6 Serie storiche e processi stocastici

Le serie storiche possono essere studiate, oltre che con strumenti descrittivi – esplorativi, come visto nel capitolo II, anche utilizzando quelli offerti dal Calcolo delle Probabilità e dall'Inferenza Statistica. Questo secondo approccio permette di ottenere risultati più facilmente generalizzabili, più precisi e meglio interpretabili quando applicati a fenomeni reali. A questo scopo, supponiamo che la serie storica (x_1, x_2, \dots, x_N) osservata sul fenomeno casuale X sia una realizzazione finita della variabile casuale N -pla (X_1, X_2, \dots, X_N) che a sua volta è una porzione finita di una successione di variabili casuali $\{X_t\}$ indicizzate con il parametro, $t = 1, 2, \dots$, rappresentante il tempo. A differenza delle generiche successioni di v.c. le cui componenti X_i possono essere riferite a fenomeni differenti, nel caso dei *processi stocastici temporali* le v.c. componenti X_t sono riferite allo stesso fenomeno in tempi diversi: per esempio il prezzo di un dato attivo finanziario, il suo rendimento, il volume delle transazioni, ecc. osservati in giorni diversi, settimane diverse, mesi diversi, ecc.

Più precisamente vale la seguente

Definizione 6

Una successione di v.c. $\{X_t\}$, $t = 1, 2, \dots$, si chiama *processo stocastico* se le v.c. X_t sono riferite allo stesso fenomeno X . Se t è il tempo si parla di *processo stocastico temporale*.

Nel seguito, il processo stocastico $\{X_t\}$ lo indicheremo più semplicemente con X_t . Possiamo ora fare alcune importanti osservazioni:

1. nella definizione si è supposto che il tempo fosse discreto e con cadenza costante (oraria, giornaliera, settimanale, ecc.) Ma i processi stocastici temporali possono essere estesi anche al caso di tempo continuo in cui t varia con continuità dell'intervallo $[0, \infty)$. Per indicare questi processi di solito si usa la notazione $X(t)$. Nel seguito analizzeremo ed applicheremo a

- fenomeni finanziari solo processi temporali a parametro discreto;
2. a seconda che il fenomeno X sia discreto o continuo, i processi stocastici temporali possono essere anche distinti in processi discreti e processi continui (il fenomeno X preso in considerazione è, almeno teoricamente, continuo);
 3. i prezzi, i rendimenti, i tassi di cambio, la volatilità sono fenomeni continui rilevati, molto spesso, a tempi costanti (ogni ora, ogni giorno, ogni settimana, ogni mese, ecc.) e quindi per il loro studio ha senso utilizzare i processi *stocastici temporali continui a tempo discreto*;
 4. dato che il tempo è irreversibile (il presente è funzione del passato, ma non del futuro) ha senso supporre che sia

$$X_t = g(X_{t-1}, X_{t-2}, \dots, a_t) \quad (45)$$

con a_t un processo stocastico temporale puramente casuale e non direttamente osservabile e $g(\cdot)$ una qualche funzione nota;

5. in pratica è sensato supporre che la dipendenza di X_t dal suo passato arrivi fino ad una qualche data finita $t-p$, con $0 < p < \infty$ e quindi è possibile sostituire la (45) con la seguente

$$X_t = g(X_{t-1}, X_{t-2}, \dots, X_{t-p}, a_t)$$

Visto che abbiamo supposto che il fenomeno in analisi è continuo, vuol dire che la v.c. k -pla $(X_{t-1}, \dots, X_{t-h})$ ammette, per ogni t e h , la funzione di densità

$$f_{t-1, t-2, \dots, t-h}(x_1, x_2, \dots, x_h) \quad (46)$$

che, in generale, sarà funzione di (t, h) e quindi la f.d. di X_t sarà una funzione di $g(\cdot)$, $f(\cdot)$, p , t , h e se non si impone una qualche condizione di “regolarità” diviene impossibile procedere ad una qualsiasi analisi del fenomeno. Nel seguito l’analisi sarà riferita ai processi stocastici, o ad una loro semplice trasformata $g(\cdot)$, che siano *stazionari*.

Definizione 7

Il processo stocastico X_t è *stazionario in senso forte* (detto anche in *senso stretto*) se la funzione di densità della v.c. $(X_t, X_{t-1}, \dots, X_{t-h})$ è la stessa di quella della v.c. $(X_s, X_{s-1}, \dots, X_{s-h})$, per ogni t ed s e qualsiasi h .

Questa definizione di stazionarietà implica che le funzioni di densità di processi stazionari in senso stretto non dipendono dal tempo (s oppure t), ma solo da h e pertanto il processo deve essere, in questo senso, “*stabile*” nel tempo. La stazionarietà forte possiede una serie di proprietà probabilistiche che la rendono teoricamente attraente, ma da un punto di vista pratico è molto restrittiva e di difficile verifica. Una forma di stazionarietà che può essere facilmente verificata, partendo da una serie storica osservata, è quella riportata nella definizione seguente.

Definizione 8

Il processo stocastico X_t è *stazionario in senso debole* o del *secondo ordine* (nel seguito diremo più semplicemente *stazionario*) se ammette i primi due momenti ed è:

$$\bullet \quad -\infty < E(X_t) = \mu < \infty, \quad t = 1, 2, \dots \quad (47)$$

$$\bullet \quad 0 < \text{Var}(X_t) = \sigma^2 < \infty, \quad t = 1, 2, \dots \quad (48)$$

$$\bullet \quad -\infty < \text{Cov}(X_t, X_{t-h}) = \gamma(h) < \infty, \quad t = 1, 2, \dots; h = 0, \pm 1, \pm 2, \dots \quad (49)$$

Notiamo che:

- a) un processo stazionario in senso stretto che ammette i primi due momenti è anche stazionario in senso debole;
- b) se il processo è stazionario in senso debole ed è normale lo è anche in senso forte. Questo perché nel mondo normale gli unici legami esistenti sono quelli lineari misurati dalla (49);
- c) nel seguito per un processo stazionario supporremo che, oltre alle (47)-(49), sia anche

$$\sum_{h=0}^{\infty} \gamma^2(h) < \infty \quad (50)$$

e quindi che i legami lineari esistenti nel processo decrescono a zero abbastanza rapidamente. Se la (50) è verificata si dice che il processo, oltre a stazionario, è anche *ergodico*;

- d) un modo semplice per definire l'ergodicità di un processo stocastico X_t è il seguente: X_t è *ergodico* se, per ogni funzione continua $g(\cdot)$ risulta

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N g(x_t) = E[g(X_t)]$$

dove (x_1, x_2, \dots, x_N) è una qualsiasi serie storica osservata su X_t . Da questa definizione segue che, se il processo è *ergodico*, tutti i suoi momenti possono essere, in pratica, ottenuti da una sua serie storica con numerosità “sufficientemente” elevata;

e) la funzione $\gamma(h)$ è detta *funzione di autocovarianza* o più semplicemente *autocovarianza* ed è tale che $\gamma(0) = \text{Var}(X_t) = \sigma^2$, $\gamma(h) = \gamma(-h)$ per ogni h .

Dalla (47) è possibile derivare la funzione

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)}, \quad h = 0, \pm 1, \pm 2, \dots \quad (51)$$

che è chiamata funzione di *autocorrelazione* del processo e per costruzione risulta $-1 \leq \rho(h) = \rho(-h) \leq 1$ per ogni h .

In pratica la (51) non è nota ed è necessario stimarla a partire da una serie storica (x_1, x_2, \dots, x_N) osservata sul processo. Se il processo è *stazionario ed ergodico* si può dimostrare che:

$$\hat{\rho}(h) = \frac{\sum_{t=1}^{N-h} (x_t - \bar{x})(x_{t+h} - \bar{x})}{\sum_{t=1}^N (x_t - \bar{x})^2}, \quad h = 1, 2, \dots, M \quad (52)$$

con M non superiore ad $N/4$, è una stima di $\rho(h)$ asintoticamente non distorta e consistente, così come lo sono le stime

$$\bar{x} = \frac{1}{N} \sum_{t=1}^N x_t; \quad \hat{\sigma}^2 = S^2 = \frac{1}{N} \sum_{t=1}^N (x_t - \bar{x})^2$$

rispettivamente per μ e σ^2 . Inoltre, si può dimostrare che, sotto l'ipotesi che sia $\rho(h) = 0$, un *intervallo di confidenza* approssimato al 95% per $\rho(h)$ è dato da

$$-2/\sqrt{N} \leq \rho(h) \leq 2/\sqrt{N} \quad (53)$$

che permette di capire se una determinata autocorrelazione $\rho(h)$ sia da considerare pari a zero o meno a seconda che la corrispondente stima $\hat{\rho}(h)$ sia contenuta o meno nell'intervallo $[-2/\sqrt{N}; 2/\sqrt{N}]$.

Le stime $\hat{\rho}(h)$ vengono riportate in un grafico in funzione di h che prende il nome di *correlogramma* che è utile per illustrare i legami lineari presenti in una

dato processo stocastico e la loro intensità. Sulla funzione di autocorrelazione $\rho(h)$ sono stati costruiti dei test globali che permettono di verificare l'ipotesi $H_0 : \rho(1) = \rho(2) = \rho(3) = \dots = 0$ verso $H_1 : \rho(h) \neq 0$, per almeno un $h > 0$. Il più utilizzato di questi test è quello di *Ljung e Box* i quali dimostrano che se, H_0 è vera, per N sufficientemente grande, risulta

$$Q_c = N(N+2) \sum_{h=1}^M \frac{\hat{\rho}^2(h)}{N-h} \sim \chi_{(M-k)}^2$$

dove k rappresenta il numero dei parametri che sono stati stimati sul processo prima di calcolare la (44). Si rifiuta H_0 , al livello α , se è

$$Q_c > \chi_{(M-k; \alpha)}^2$$

dove $\chi_{(M-k; \alpha)}^2$ è il percentile calcolato sulla v.c. chi quadrato con $(M-k)$ gradi di libertà che stacca alla sua destra la probabilità α .

3.7 Alcuni particolari processi stocastici

In questo paragrafo presentiamo alcuni particolari processi stocastici che incontreremo nei capitoli successivi e che hanno rilevanza nelle pratiche applicazioni.

(A) PROCESSO INDIPENDENTE ED IDENTICAMENTE DISTRIBUITO

Indichiamo con a_t un processo indipendente ed identicamente distribuito (*processo iid*). In questo caso la sua funzione di densità, che indichiamo con $f(x)$, è la stessa per ogni t (identicamente distribuito); inoltre per l'indipendenza delle componenti la funzione di densità di ogni v.c. $(k+1)$ -pla $(a_t, a_{t-1}, \dots, a_{t-k})$ è data da

$$f_{t, t-1, t-2, \dots, t-k}(x_1, x_2, \dots, x_k) = (f(x))^k, \text{ per ogni } t \text{ ed ogni } k$$

ed il processo risulta stazionario in senso forte. Se il processo ammette i primi due momenti deve essere

$$E(a_t) = \mu; \quad \text{Var}(a_t) = \sigma^2; \quad \gamma(h) = \text{cov}(a_t, a_{t-h}) = 0$$

ed il processo è anche *stazionario in senso debole*. Per indicare che a_t è *indipendente ed identicamente distribuito* ed ammette i primi due momenti scriveremo $a_t \stackrel{\text{iid}}{\sim} (\mu, \sigma^2)$. In molte applicazioni è giustificato supporre che sia $\mu = 0$ ed in tal caso scriveremo $a_t \stackrel{\text{iid}}{\sim} (0, \sigma^2)$.

(B) PROCESSO WHITE NOISE

Il processo *white noise* a_t è un particolare processo stazionario in senso debole con

$$E(a_t) = 0; \quad \text{Var}(a_t) = \sigma^2; \quad \gamma(h) = \text{cov}(a_t, a_{t-h}) = 0 \quad (54)$$

e si scrive $a_t \sim \text{WN}(0, \sigma^2)$. In generale, se è $a_t \sim \text{WN}(0, \sigma^2)$ questo non implica che è anche $a_t \stackrel{\text{iid}}{\sim} (0, \sigma^2)$, mentre abbiamo visto che è vero il viceversa. Nel caso particolare in cui a_t è normale ed è white noise, allora è anche iid.

Per indicare che a_t è un white noise normale scriveremo $a_t \sim \text{NWN}(0, \sigma^2)$.

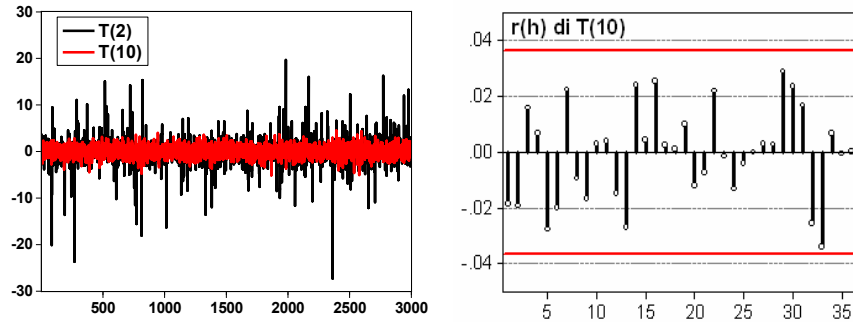
Il processo white noise è molto utilizzato per descrivere i residui di processi stocastici rappresentati con particolari modelli statistici. Nella figura 15 riportiamo la serie storica di $N=3000$ osservazioni generata dal processo iid con a_t una v.c. t di Student con $k = 2$ e 10 gradi di libertà, rispettivamente, e le prime 36 autocorrelazioni stimate per $t_{(10)}$. Osserviamo che quando è $k=2$ il processo è iid ma non è un white noise dato che in tal caso la varianza non esiste e perde di significato stimarla, come perde di significato stimarne le autocorrelazioni.

Come si può osservare dalla figura 15 tutte le autocorrelazioni stimate sono nella banda di confidenza

$$\left[-2/\sqrt{3000}; \quad 2/\sqrt{3000} \right]$$

(riportata in rosso) e quindi non sono, con una probabilità del 95%, significativamente diverse dal valore atteso che è zero.

Figura 15 – Serie storiche generate da un processo iid t di Student con $k=2$ e 10 (in rosso) gradi di libertà e prime 36 autocorrelazioni per $t_{(10)}$



(C) MEDIE PROGRESSIVE DI PROCESSI IID

Sia $a_t \stackrel{iid}{\sim} (0, \sigma^2)$ e consideriamo il processo stocastico:

$$X_t = \frac{1}{t} \sum_{j=1}^t a_j \quad (55)$$

È agevole verificare che

- $E(X_t) = \frac{1}{t} \sum_{j=1}^t E(a_j) = 0, t=1, 2, \dots$
- $\text{Var}(X_t) = \text{Var}\left(\frac{1}{t} \sum_{j=1}^t a_j\right) = \frac{1}{t^2} \sum_{j=1}^t \text{Var}(a_j) = \frac{\sigma^2}{t}, t=1, 2, \dots$
- $\text{cov}(X_t, X_{t+h}) = E(X_t X_{t+h}) = \frac{1}{t(t+h)} E\left[\sum_{j=1}^t a_j \left(\sum_{j=1}^t a_j + \sum_{j=t+1}^{t+h} a_j\right)\right] =$

$$= \frac{1}{t(t+h)} t\sigma^2 = \frac{\sigma^2}{t+h}, t=1, 2, \dots; h=0, 1, 2, \dots$$
- $\text{cov}(X_t, X_{t-h}) = E(X_t X_{t-h}) = \frac{\sigma^2}{t}, t=1, 2, \dots; h=0, 1, 2, \dots$

ed il processo non è stazionario né in varianza, né in covarianza, inoltre risulta

$\text{cov}(X_t, X_{t+h}) \neq \text{cov}(X_t, X_{t-h})$. Se al posto del processo (55) consideriamo il seguente

$$W_t = \frac{1}{\sqrt{N}} \sum_{j=1}^t a_j \quad (56)$$

si verifica immediatamente che è

$$E(W_t) = 0; \quad \text{Var}(W_t) = \frac{t}{N} \sigma^2; \quad \text{cov}(W_t, W_{t-h}) = \frac{t-h}{N} \sigma^2$$

$$\text{cov}(W_t, W_{t+h}) = \frac{t}{N} \sigma^2$$

ed anche W_t non è stazionario, ma a differenza del (55) degenera per $N \rightarrow \infty$. Il processo W_t ha una serie di proprietà che lo rendono utile per rappresentare i prezzi di attivi finanziari, in particolare è agevole verificare che:

- se dal (56) consideriamo il nuovo processo $\nabla W_t = W_t - W_{t-1}$ risulta

$$\nabla W_t = \frac{1}{\sqrt{N}} a_t \stackrel{\text{iid}}{\sim} \left(0, \frac{\sigma^2}{N} \right)$$

e quindi le variazioni ∇W_t sono iid;

- se è $a_t \stackrel{\text{iid}}{\sim} N(0, \sigma^2)$, $t = 1, 2, \dots$, allora è anche

$$W_t \sim N\left(0, \frac{t}{N} \sigma^2\right); \quad (W_t, W_{t-h}) \sim N_2\left(0, 0; \frac{t}{N} \sigma^2; \frac{t-h}{N} \sigma^2; \frac{t-h}{N} \sigma^2\right)$$

- se è semplicemente $a_t \sim WN(0, \sigma^2)$ allora W_t è ad *incrementi incorrelati* (detti anche *ortogonali*);
- posto $v < s < t$ i processi

$$(W_t - W_s) = \frac{1}{\sqrt{N}} \sum_{j=s+1}^t a_j; \quad W_v = \frac{1}{\sqrt{N}} \sum_{j=1}^v a_j$$

sono indipendenti se è $a_t \stackrel{\text{iid}}{\sim} (0, \sigma^2)$, mentre sono solo incorrelati se è $a_t \sim \text{WN}(0, \sigma^2)$;

- se a_t , oltre ad essere iid, è anche normale si avrà

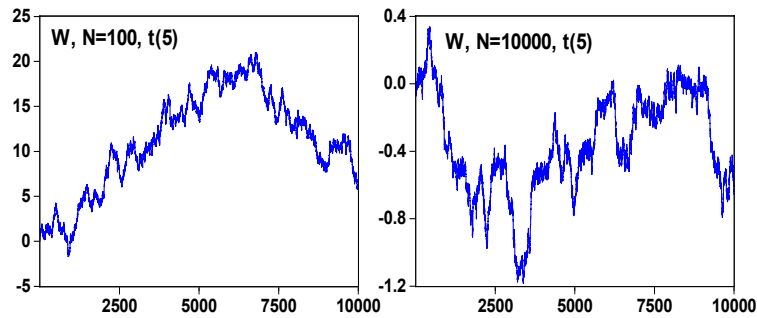
$$(W_t - W_s) \sim N\left(0, \frac{t-s}{N} \sigma^2\right)$$

- se è $a_t \sim \text{WN}(0, \sigma^2)$ e consideriamo le variazioni $(W_{t+1} - W_{s+1})$ e $(W_{t-s+1} - W_1)$ queste hanno la stessa media, la stessa varianza e la stessa autocovarianza, mentre se è $a_t \stackrel{\text{iid}}{\sim} N(0, \sigma^2)$ risulta

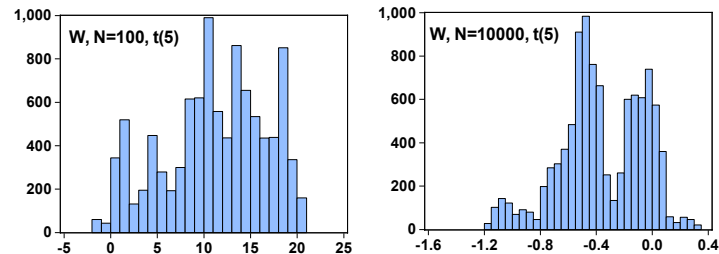
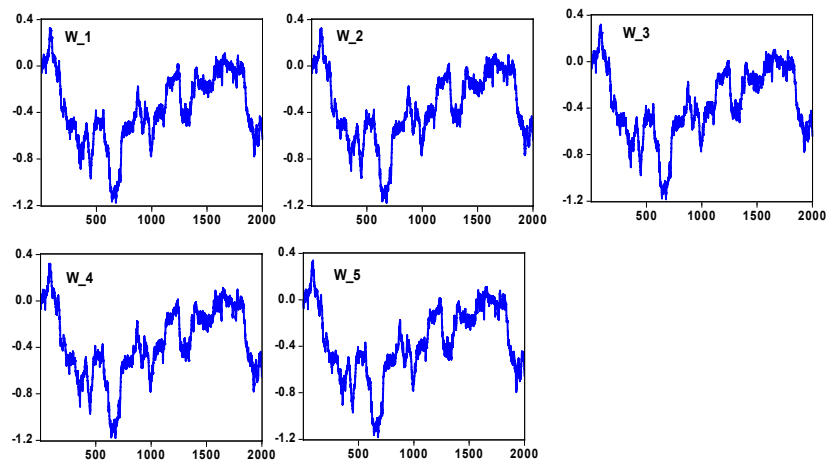
$$(W_{t+1} - W_{s+1}) = \frac{1}{\sqrt{N}} \sum_{j=s+2}^{t+1} a_j \sim N\left(0, \frac{t-s}{N} \sigma^2\right)$$

$$(W_{t-s+1} - W_1) = \frac{1}{\sqrt{N}} \sum_{j=2}^{t-s+1} a_j \sim N\left(0, \frac{t-s}{N} \sigma^2\right)$$

Figura 16 – W_t con $a_t = t_{(5)}$ ed $N=100$ e 10000



Nella figura 16 è riportata una serie storica di 10000 osservazioni generate dal processo W_t quando è $a_t = t_{(5)}$ ed $N=100$ ed $N=10000$ rispettivamente.

Figura 17 – Istogramma di W_t con $a_t = t_{(5)}$ ed $N=100$ e 10000 Figura 18 – Sottoserie W_{5t} di W_t con $N=10000$ 

Come si può osservare dalla figura 16, l'andamento di W_t è molto simile a quello delle quotazioni di attivi finanziari. Le stesse considerazioni valgono per la figura 17 in cui sono riportati gli istogrammi di frequenza delle serie. La figura 18 riporta le 5 sottoserie di W_t estratte con cadenza $5t$ ed $N=10000$ e mostra, confrontata con la parte destra della figura 16, la sua autosimilarità (una parte è simile al tutto) proprietà tipica dei prezzi degli attivi finanziari.

(D) PROCESSO DI WIENER O MOTO BROWNIANO

Questo processo è una generalizzazione, al caso di parametro continuo, del processo riportato nella (56) sotto l'ipotesi che le a_t siano iid e normali. Formalmente il processo viene indicato con $W(t)$, per $0 \leq t \leq 1$, ed è un processo stocastico di Wiener se è ad incrementi normali indipendenti e stazionari e sono, per tanto, soddisfatte le condizioni seguenti:

- $W(t)$ varia con continuità con t in $[0; 1]$;
- per $v \leq s \leq t$ il processo $W(t)-W(s)$ è indipendente da $W(v)$;
- per $s \leq t$ gli incrementi $W(t)-W(s)$ e $W(t-s)-W(0)$ hanno la stessa distribuzione di probabilità;
- $W(t)-W(s) \sim N(0; \sigma^2(t-s))$

Come si può notare, fra il processo di Wiener e quello definito nella (54) vi sono molte proprietà in comune. Questa somiglianza è giustificata dal teorema di Donsker (1951):

Teorema

Se le v.c. a_j sono $N(0, \sigma^2)$, per $t \in [0; 1]$, indicando con $[Nt]$ la parte intera di Nt , allora per $N \rightarrow \infty$ si ha

$$W_{N,t} = \frac{1}{\sqrt{N}} \sum_{j=1}^{[Nt]} a_j \xrightarrow{L} W(t); \quad t \in [0; 1]$$

Il processo di Wiener è facilmente estensibile a caso di $a_t \stackrel{\text{iid}}{\sim} N(\mu, \sigma^2)$ ottenendo $(W(t)-W(s)) \sim N(\mu(t-s), \sigma^2(t-s))$. Se è $W(0) = 0$ con probabilità 1, $\mu = 0$ e $\sigma^2 = 1$ si dice che il processo di Wiener è *standard*. Inoltre, si dimostra che $W(t)$ non è differenziabile quasi certamente nel senso che $dW(t)/dt$ esiste solo in un sottoinsieme del suo campo di esistenza e questo sotto insieme ha misura nulla. Tale risultato implica che il processo di Wiener non può essere analizzato con l'usuale calcolo integrale, ma è necessario usare quella che viene chiamata *integrazione di Ito*. Utilizzando tale integrazione si dimostra che se il processo dei prezzi di un attivo soddisfa l'equazione:

$$dP_t = \mu P_t dt + \sigma P_t dW(t)$$

allora risulta

$$d \log(P_t) = \left(\mu - \frac{\sigma^2}{2} \right) dt + \sigma dW(t)$$

ed i rendimenti in $[Nt]$ sono dati da

$$r_{[t,N]} = \log(P_N) - \log(P_t) \sim N \left[\left(\mu - \frac{\sigma^2}{2} \right) (N-t); \sigma^2 (N-t) \right]$$

e posto $N-t = \Delta$ si ottiene

$$r_{[t, t+\Delta]} = \log(P_{t+\Delta}) - \log(P_t) \sim N \left[\left(\mu - \frac{\sigma^2}{2} \right) \Delta; \sigma^2 \Delta \right].$$

Data la serie storica (r_1, r_2, \dots, r_N) è possibile ottenere le stime della media e della varianza tramite

$$\bar{r} = \frac{1}{N} \sum_{t=1}^N r_t; \quad S^2 = \frac{1}{N} \sum_{t=1}^N (r_t - \bar{r})^2$$

e dalle uguaglianze

$$\bar{r} = \left(\mu - \frac{\sigma^2}{2} \right) \Delta; \quad S^2 = \sigma^2 \Delta$$

ottenere le stime

$$\hat{\mu} = \frac{\bar{r}}{\Delta} + \frac{S^2}{2\Delta}; \quad \hat{\sigma}^2 = \frac{S^2}{\Delta}$$

che, nel caso particolare di $\Delta = 1$, si riducono a

$$\hat{\mu} = \bar{r} + \frac{S^2}{2}; \quad \hat{\sigma}^2 = S^2$$

Questi risultati sono ottenuti sotto l'ipotesi che le a_t siano normali e indipendenti, cosa tutta da verificare nelle pratiche applicazioni.

(E) MEDIE MOBILI DI PROCESSI IID

Consideriamo il processo stocastico

$$M_t^{(k)} = c + \frac{1}{k} \sum_{j=0}^{k-1} a_{t-j} \quad (57)$$

con c costante, k intero prefissato ed $a_t \stackrel{\text{iid}}{\sim} (0, \sigma^2)$, risulta

$$E(M_i^{(k)}) = c; \quad \text{Var}(M_t^{(k)}) = \frac{\sigma^2}{k}$$

$$\text{Cov}(M_t^{(k)}, M_{t+h}^{(k)}) = \frac{k-h}{k^2} \sigma^2, \quad \text{per } h < k$$

per $h \geq k$ si ha $\text{cov}(M_t^{(k)}, M_{t+h}^{(k)}) = 0$ e quindi

$$\rho(h) = \begin{cases} 1 - h/k & \text{per } h < k \\ 0 & \text{altrove} \end{cases}$$

In sintesi il processo (57) è sempre stazionario ed ha una memoria limitata a $h < k$. Nella figura 19 sono riportate le serie storiche di $N=10000$ osservazioni del processo (57) per $c=1$, $k=10$, $a_t \sim t_{(5)}$ e le stime delle loro funzioni di densità.

Figura 19 – Serie storica di $M_t^{(10)}$ e di $a_t \sim t_{(5)}$ e loro f.d

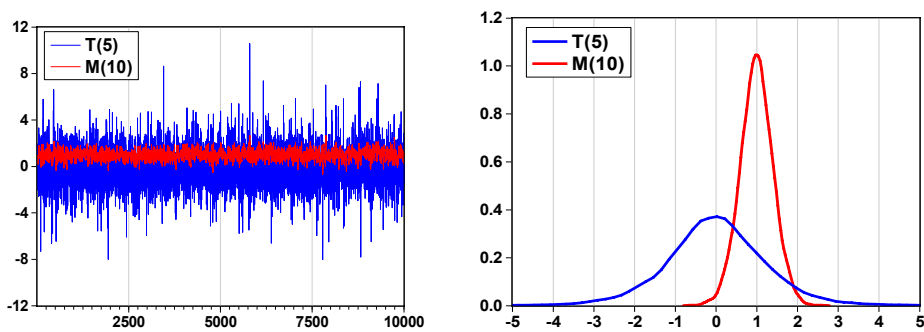
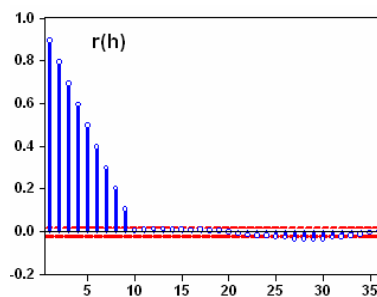


Figura 20 – Autocorrelazioni di $M_t^{(10)}$



Nella figura 20 sono riportate le prime 36 autocorrelazioni stimate di $M_t^{(10)}$ con la relativa banda di confidenza (in rosso). Come si può osservare, solo le prime $k-1=9$ autocorrelazioni sono significativamente diverse da zero e decrescono linearmente al crescere di h .

(F) PROCESSO RANDOM WALK

Un processo Random Walk X_t , che di solito si indica con $X_t \sim RW$, è definito come

$$X_t = c + X_{t-1} + a_t; \quad a_t \sim \text{WN}(0, \sigma^2) \quad (58)$$

dove la costante c viene chiamata *drift*. Se x_0 è un valore iniziale di X_t , per sostituzioni successive, alla t -sima iterazione si ottiene

$$X_t = (x_0 + ct) + \sum_{j=0}^{t-1} a_{t-j} \quad (59)$$

ed il fenomeno evolve lungo una retta con intercetta data dal valore iniziale x_0 e coefficiente angolare dato dal *drift*. E' immediato verificare che è

$$E(X_t) = X_0 + ct; \quad \text{Var}(X_t) = t\sigma^2$$

$$\text{cov}(X_t, X_{t+h}) = \min(t, t+h)\sigma^2; \quad \rho(h) = 1$$

ed il processo non è stazionario. Nella figura 21 riportiamo una traccia del processo (58) con $c = 0.1$, $x_0 = 5$; $a_t \stackrel{\text{iid}}{\sim} t_{(5)}$ e la stima delle prime 20 autocorrelazioni.

Ne caso in cui nella (58) è $c=0$, nella (59) scompare il trend lineare ed il processo è fortemente influenzato dal valore iniziale, che di solito, è incognito. In questo caso le tracce del processo sono molto diverse da quella rappresentata nella figura 21. Nella figura 22, a titolo di esempio, riportiamo la serie storica della (58), con $c=0$ ed $a_t \stackrel{\text{iid}}{\sim} t_{(5)}$, e le prime 20 autocorrelazioni stimate.

Figura 21 – Traccia del Random Walk con $c = 0.1$, $x_0 = 5$; $a_t \stackrel{iid}{\sim} t_{(5)}$ e prime 20 autocorrelazioni stimate con banda di confidenza (in rosso)

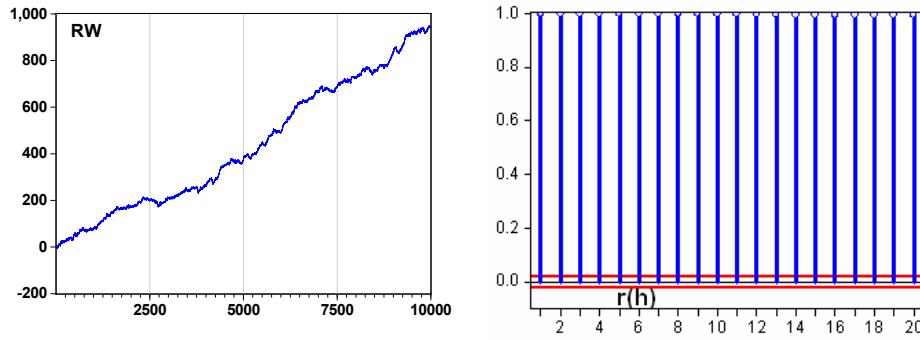
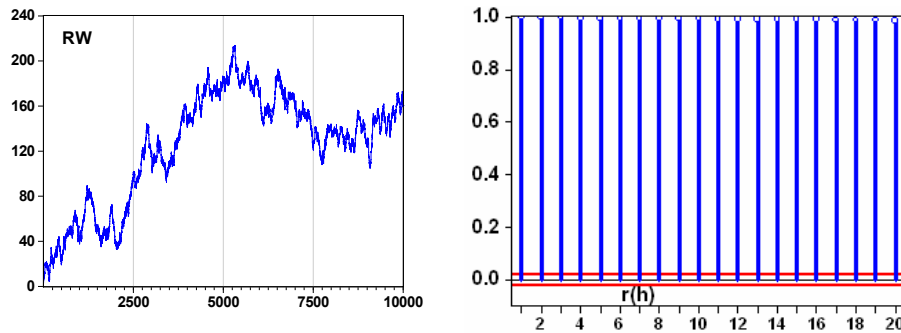


Figura 22 – Traccia ed autocorrelazioni del RW senza drift



Dalla (56), se calcoliamo le variazioni, chiamate anche *differenze prime*, si ottiene:

$$X_t - X_{t-1} = c + a_t \sim \text{WN}(c, \sigma^2)$$

Il processo stocastico che ne risulta $\nabla X_t = X_t - X_{t-1}$ è stazionario. L'esperienza empirica ci mostra che le serie storiche dei prezzi degli attivi finanziari sono ben rappresentate da processi stocastici *random walk* con $c=0$.

CAPITOLO IV

Analisi dei livelli di attivi finanziari e monetari

4.1 Introduzione

Nel capitolo precedente abbiamo sottolineato come un possibile modello per l'analisi dei prezzi di attivi finanziari sia la soluzione della seguente equazione differenziale stocastica:

$$dP_t = \mu P_t dt + \sigma P_t dW(t) \quad (1)$$

che è un caso particolare di quelli che vengono chiamati *processi stocastici di diffusione* (vedi, fra gli altri, Iacus, 2008). Dalla (1), utilizzando i logaritmi di ambo i membri e l'integrazione di Ito, si può dimostrare che è

$$d \log(P_t) = \left(\mu - \frac{\sigma^2}{2} \right) dt + \sigma dW(t)$$

che permette di derivare i rendimenti riferiti all'intervallo temporale Δ

$$r_{t-1}[\Delta] = \log(P_{t-1+\Delta}) - \log(P_{t-1}) \sim N \left(\left(\mu - \frac{\sigma^2}{2} \right) \Delta; \sigma^2 \Delta \right),$$

che per $\Delta = 1$ diventa

$$r_t = \log(P_t) - \log(P_{t-1}) \sim N \left(\left(\mu - \frac{\sigma^2}{2} \right); \sigma^2 \right)$$

Questo implica che il modello (1) non è in grado di catturare l'eccesso di curtosi che, sistematicamente, si osserva nei rendimenti di attivi. Tale problema può essere risolto in diversi modi, gli approcci più utilizzati nelle pratiche

applicazioni sono due:

1. si generalizza la (1) di modo che il relativo modello sia in grado di catturare l'eccesso di curtosi;
2. si utilizzano appositi modelli statistico – econometrici, costruiti con un approccio essenzialmente empirico, che siano in grado di spiegare l'eccesso di curtosi.

Nel seguito la problematica della modellizzazione degli attivi verrà affrontata con questo secondo approccio e verrà utilizzata per analizzare i prezzi di attivi ed i connessi rendimenti osservati con cadenza giornaliera, settimanale o mensile.

Da quanto detto nel capitolo II, il modello empirico più plausibile per i *prezzi* degli attivi è

$$P_t = P_{t-1} e^{r_t}, \quad t = 1, 2, \dots \quad (2)$$

che implica

$$r_t = \log(P_t) - \log(P_{t-1})$$

Inoltre, se si impone che per i *rendimenti* sia

$$\begin{cases} r_t = \mu_t + \varepsilon_t \\ \varepsilon_t = \sigma_t a_t \end{cases} \quad (3)$$

con

$$\mu_t = f(r_{t-1}, r_{t-2}, \dots; \varepsilon_t, \varepsilon_{t-2}, \dots); \quad \sigma_t = g(\sigma_{t-1}, \sigma_{t-2}, \dots; \varepsilon_{t-1}^2, \varepsilon_{t-2}^2, \dots) > 0$$

$$(a_t | \mathfrak{F}_{t-1})^{\text{iid}} \sim (0, 1); \quad \varepsilon_t \sim \text{WN}(0, \sigma_\varepsilon^2); \quad \mathfrak{F}_{t-1} = (r_{t-1}, r_{t-2}, \dots)$$

si ottiene

$$P_t = P_{t-1} e^{\mu_t + \sigma_t a_t}$$

Dalla (2), posto $p_t = \log(P_t)$ si ricava

$$r_t = p_t - p_{t-1} = \mu_t + \sigma_t a_t \quad (4)$$

Ovvero i rendimenti sono dati dalla somma fra un livello medio μ_t , che rappresenta il *segnale*, ed un *errore* che possiede una varianza σ_t^2 funzione del tempo. La (4) implica che, se è $\mu_t = 0$, come succede in molte pratiche applicazioni riguardanti gli attivi finanziari, i logaritmi dei prezzi si comportano come un processo *random walk*. In alcuni casi è $\mu_t \approx 0$ e p_t diviene un *quasi random walk*, cioè una random walk a meno del segnale.

Osservando la (4) si può concludere che nel seguito è necessario affrontare e risolvere i seguenti problemi:

1. identificare un modello per μ_t che rappresenti il livello medio dei rendimenti, stimarne i relativi parametri e costruirne le previsioni;
2. derivare la distribuzione probabilistica di ε_t e verificare se questa è riconducibile alla normale o a qualche altra distribuzione nota;
3. identificare un modello per σ_t^2 , stimarne i relativi parametri e costruirne le previsioni;
4. derivare la distribuzione di a_t e verificare se questa è compatibile con quella della normale standardizzata o con qualche altra distribuzione nota;
5. stimare e prevedere il rischio di mercato degli attivi o dei portafogli di attivi che sono in analisi.

Queste problematiche verranno affrontate in questo e nei prossimi capitoli.

4.2 *Analisi dei rendimenti di attivi finanziari*

Spesso, per poter fare inferenza, si ipotizza che il fenomeno allo studio, o una sua qualche semplice trasformata, si distribuisce come una v.c. normale. Naturalmente, questa ipotesi deve essere verificata a partire dalle informazioni campionarie che si hanno sul fenomeno allo studio. Un modo semplice per

ottenere questo risultato è stimare gli indici di asimmetria e curtosi, γ_1 e γ_2 , e verificare, sulla base delle stime ottenute, che sia $\gamma_1 = \gamma_2 = 0$, che è quello che si avrebbe se effettivamente il fenomeno fosse distribuito come una normale. In altri termini, data una serie storica osservata sul fenomeno, è necessario sottoporre a test

$$\begin{cases} H_0 : \gamma_1 = \gamma_2 = 0 \\ H_1 : \gamma_1 \text{ e/o } \gamma_2 \neq 0 \end{cases}$$

Sia (r_1, r_2, \dots, r_N) la serie storica dei rendimenti osservata ai tempi $t = 1, 2, \dots, N$, utilizziamo queste informazioni per ottenere, con le formule viste nel capitolo II, la stima $(\hat{\gamma}_1, \hat{\gamma}_2)$ di (γ_1, γ_2) e costruiamo il test tenendo conto del seguente teorema (Snedecor e Cochran (1980)).

Teorema

Se è $r_t \sim N(\mu; \sigma^2)$ allora, per $N \rightarrow \infty$, si ha

$$\hat{\gamma}_1 \xrightarrow{L} N\left(0; \frac{6}{N}\right); \quad \hat{\gamma}_2 \xrightarrow{L} N\left(0; \frac{24}{N}\right)$$

e sono indipendenti.

Da questo teorema, standardizzando $\hat{\gamma}_1$ e $\hat{\gamma}_2$ e ricordando la definizione della v.c. chi quadrato risulta

$$JB_c = \frac{\hat{\gamma}_1^2}{6/N} + \frac{\hat{\gamma}_2^2}{24/N} = \frac{N}{6} \left[\hat{\gamma}_1^2 + \frac{\hat{\gamma}_2^2}{4} \right] \xrightarrow{H_0} \chi_{(2)}^2 \quad (5)$$

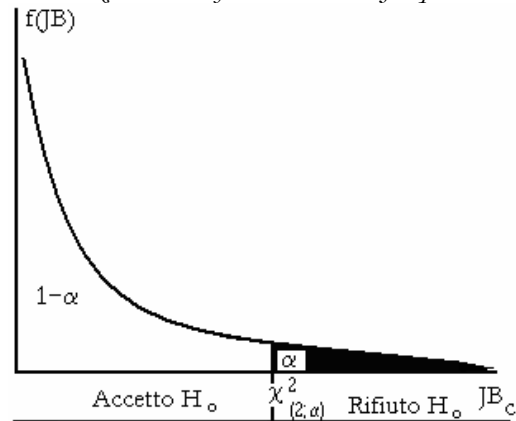
La (5) può essere utilizzata per costruire il test di normalità sul fenomeno in studio. Questo test prende il nome di *test di Jarque – Bera* e si rifiuta $H_0 : \gamma_1 = \gamma_2 = 0$, al livello α se è

$$JB_c > \chi_{(2; \alpha)}^2$$

dove $\chi_{(2; \alpha)}^2$ è il percentile della v.c. chi quadrato che stacca alla sua destra l' α

per cento dei casi. Graficamente la ragione di accettazione e di rifiuto di H_0 e della relativa normalità è schematizzata nella figura 1.

Figura 1 – Regione di accettazione e di rifiuto del test di Jarque – Bera



Nella tabella 1 riportiamo gli indici descrittivi stimati, il valore JB_c calcolato ed il valore soglia $\chi_{(2; \alpha)}^2$ per le serie delle quotazioni giornaliere dei prezzi di Campari, Fiat, Intesa S. Paolo, Mibtel, Nasdaq, Standard & Poor 500 osservati dal 2003 al 2009.

Tabella 1 – Indici descrittivi e test di Jarque – Bera per i prezzi degli attivi di Campari, Fiat, Intesa S. Paolo, Mibtel, Nasdaq, Standard & Poor500

	CAMP	FIAT	INTSP	MBTL	NSDQ	S&P
\bar{P}	5.3704	9.7020	3.1371	24202	1576.4	1204.7
M_c	5.5200	7.0500	2.8500	24385	1565.0	1214.1
S	1.5739	5.0134	1.2424	5359.1	275.11	192.94
\hat{Y}_1	-0.0475	1.0190	0.1751	-0.0459	-0.0638	-0.3178
\hat{Y}_2	-1.37	-0.19	-1.25	-2.88	-0.48	-0.53
JB_c	120.76	281.78	112.45	52.459	16.732	46.189
$\chi_{(2; 0.05)}^2$	5.99	5.99	5.99	5.99	5.99	5.99
N	1614	1614	1610	1620	1620	1620

Come si può osservare dalla tabella, e come era largamente atteso, per tutte le sei serie analizzate si rifiuta l'ipotesi che queste si distribuiscano come una normale. Nella tabella 2 riportiamo le stesse elaborazioni riferite ai rendimenti di quelle serie.

Tabella 2 – *Indici descrittivi e test di Jarque – Bera per i rendimenti delle sei serie considerate*

	CAMP	FIAT	INTSP	MIBTEL	NSDQ	S&P
\bar{r}	0.0002	-0.0003	0.0002	-0.0002	0.0000	0.0000
M_e	0.0000	0.0000	0.0000	0.0004	0.0004	0.0004
S	0.0169	0.0244	0.0226	0.0119	0.0155	0.0133
$\hat{\gamma}_1$	0.1018	0.0284	0.0480	-0.1179	-0.0442	-0.3329
$\hat{\gamma}_2$	3.9184	3.9847	13.455	11.760	7.3630	13.232
JB_c	1034.7	1067.3	12138	9333.0	3657.7	11842
$\chi^2_{(2;0.05)}$	5.99	5.99	5.99	5.99	5.99	5.99
N	1613	1613	1609	1619	1619	1619

Ed anche in questo caso viene rifiutata l'ipotesi di normalità. Osserviamo che, nel caso dei prezzi, dato che non sono in generale stazionari, le stime degli indici descrittivi che si ottengono sono distorte e non consistenti. Pertanto perde di significato costruire su di essi i classici test delle ipotesi. In particolare, il test di Jarque – Bera ha senso costruirlo se la numerosità è elevata, come nei casi qui analizzati, ed il processo stocastico generatore dei dati è stazionario come avviene per i rendimenti.

4.3 *Uso dei modelli ARMA per l'analisi dei livelli dei rendimenti*

Nel paragrafo precedente abbiamo sottolineato che ha senso supporre che il livello medio dei rendimenti sia una qualche funzione dei rendimenti passati e degli errori presenti e passati. Per rendere operativo questo concetto è necessario definire la funzione che lega i rendimenti al suo passato e stimarne, eventualmente, i relativi parametri. Esistono diversi modi per ottenere una stima del livello medio dei rendimenti e nel capitolo II abbiamo illustrato ed utilizzato quello basato sulle medie mobili. In questo paragrafo supponiamo, invece, che sia

$$r_t = \phi_0 + \phi_1 r_{t-1} + \phi_2 r_{t-2} + \dots + \phi_p r_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \quad (6)$$

dove

$$\varepsilon_t \sim \text{WN}(0, \sigma^2), \text{ mentre } (\phi_0, \phi_1, \phi_2, \dots, \phi_p, \theta_1, \dots, \theta_q, \sigma^2)$$

sono $(p+q+2)$ parametri incogniti da stimare. Il modello definito nella (6) viene chiamato ARMA(p, q) (*Autoregressivo* di ordine p e *Media mobile* di ordine q) e siglato con $r_t \sim \text{ARMA}(p, q)$.

E' importante osservare che, a differenza di altri contesti, nelle analisi dei rendimenti di attivi finanziari p e q sono, di solito, molto piccoli. Infatti per le analisi di attivi finanziari i casi più rilevanti sono quelli con $p \leq 1$ e $q \leq 5$. Questo perché, nell'ipotesi di efficienza e trasparenza dei mercati, la serie dei livelli medi μ_t dei rendimenti deve essere vicina allo zero se non proprio nulla. Osserviamo che per $p = q = 0$ si ha $r_t = \phi_0 + \varepsilon_t$ e quindi i rendimenti, a meno del *drift* ϕ_0 , sono un *white noise*, mentre per il logaritmo dei prezzi risulta $p_t = \phi_0 + p_{t-1} + \varepsilon_t$ che è un *random walk* con *drift*. Osserviamo ancora che se fosse $p = q = 0$ e risultasse $\phi_0 \neq 0$, si avrebbe $E(r_t) = E(\phi_0 + \varepsilon_t) = \phi_0$ che implicherebbe l'esistenza di *arbitraggio* in quel mercato. D'altro lato, nei mercati degli attivi finanziari l'arbitraggio può essere giustificato solo per brevissimi periodi e non in modo sistematico. In caso contrario verrebbe a cadere l'ipotesi di efficienza e trasparenza di tali mercati. Dalla prima riga della tabella 2 si può notare come la stima del livello medio dei sei attivi analizzati sia praticamente nulla a conferma del fatto che nel modello per i rendimenti il *drift* deve essere sostanzialmente pari a zero.

Se è $p = q = 0$ gli unici parametri da stimare sono ϕ_0 e σ^2 e dato che è

$$E(r_t) = \phi_0; \quad \text{Var}(r_t) = \text{Var}(\phi_0 + \varepsilon_{3t}) = \sigma^2$$

queste stime possono essere ottenute, a partire dalla serie storica (r_1, r_2, \dots, r_N) osservata sui rendimenti, con

$$\hat{\phi}_0 = \bar{r} = \frac{1}{N} \sum_{t=1}^N r_t; \quad S^2 = \frac{1}{N} \sum_{t=1}^N (r_t - \bar{r})^2$$

Dato che r_t è stazionario, queste stime sono asintoticamente non distorte e

consistenti. Inoltre se è $\varepsilon_t \stackrel{\text{iid}}{\sim} N(0, \sigma^2)$, è anche

$$\hat{\phi}_0 \sim N\left(\phi_0; \frac{\sigma^2}{N}\right); \quad \frac{NS^2}{\sigma^2} \sim \chi_{(N-1)}^2 \quad (7)$$

e sono fra di loro indipendenti. Questo risultato ci permette di derivare la statistica test sul *drift* ϕ_0

$$t_c = \frac{\sqrt{N-1}(\hat{\phi}_0 - \phi_0)}{S} \sim t_{(N-1)} \quad (8)$$

che ci permette di sottoporre a test $H_0 : \phi_0 = 0$ verso $H_1 : \phi_0 \neq 0$ e rifiutare H_0 se risulta

$$|t_c| > t_{(N-1; \alpha/2)} \quad (9)$$

dove $t_{(N-1; \alpha/2)}$ è il percentile della v.c. t di Student con $(N-1)$ gradi di libertà che stacca, rispettivamente, nella coda destra ed in quella sinistra una probabilità di $\alpha/2$. D'altro lato sappiamo che la v.c. t di Student, con gradi di libertà superiore a 30, è ben approssimata dalla normale standardizzata. Nel caso di attivi finanziari N è molto grande, come si può osservare anche dalle tabelle 1 e 2, e quindi la (9) può essere sostituita dalla seguente

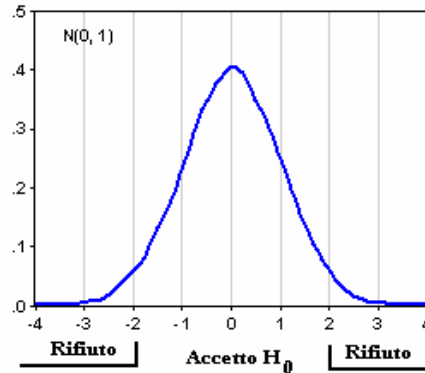
$$|t_c| > z_{(\alpha/2)}$$

dove $z_{(\alpha/2)}$ è il percentile calcolato sulla normale standardizzata. Se si tiene conto che di solito si pone $\alpha=0.05$ e che in questo caso è $z_{(0.025)} \approx 2$, si rifiuta H_0 al 5% più semplicemente se

$$|t_c| > 2 \quad \text{o equivalentemente} \quad |\hat{\phi}_0 - \phi_0| > 2 \frac{S}{\sqrt{N-1}}$$

La regione critica che ne deriva viene illustrata graficamente in figura 2.

Figura 2 – Regione di significatività per ϕ_0 con $\alpha = 0.05$



In generale è possibile costruire il seguente *intervallo di confidenza* per ϕ_0 , al livello $1-\alpha$

$$\hat{\phi}_0 - \frac{S}{\sqrt{N-1}} t_{((N-1); \alpha/2)} \leq \phi_0 \leq \hat{\phi}_0 + \frac{S}{\sqrt{N-1}} t_{((N-1); \alpha/2)}$$

Nel caso delle serie analizzate nella tabella 2, si ottengono i risultati della tabella 3. Da questa tabella si osserva che nessuno dei livelli medi dei rendimenti delle sei serie è, come atteso, significativamente diverso da zero.

Tabella 3 – *Sima dei livelli medi, dell'errore standard e della t_c per le sei serie*

	CAMP	FIAT	INTSP	MIBTEL	NSDQ	S&P
$\hat{\phi}_0$	0.0002	-0.0003	0.0002	-0.0002	0.0000	0.0000
$S/\sqrt{N-1}$	0.00042	0.00061	0.00057	0.00029	0.00038	0.00033
t_c	0.51541	-0.60282	0.45552	-0.83103	0.12507	-0.45878

I risultati (7) – (9) sono stati ottenuti sotto l'ipotesi che i rendimenti si distribuiscano come normali indipendenti. D'altra parte, come più volte sottolineato, una delle caratteristiche dei rendimenti di attivi finanziari è che risulta sistematicamente $\gamma_2 > 0$ e quindi c'è un eccesso di curtosi che non può, ovviamente, essere catturato dal modello normale. Di solito, una ipotesi più realistica è che sia $r_t \stackrel{iid}{\sim} t_{(k)}$ con $k > 4$. Questo implica che le (7)-(9) non sono

verificate. D'altra parte, se la numerosità N della serie è sufficientemente elevata, come generalmente succede in questi casi, per il *Teorema del Limite Centrale* le (7) - (9) continuano a valere in modo approssimato con un errore di approssimazione che di solito è dell'ordine $O(N^{-1/2})$. In definitiva, nelle analisi dei prezzi di attivi finanziari con cadenze giornaliere, settimanali o mensili i risultati sopra presentati, anche se non esatti, possono essere utilizzati nelle pratiche applicazioni senza eccessive preoccupazioni.

La (7) può essere utilizzata anche per sottoporre a test la *varianza* dei rendimenti ed in particolare è possibile testare $H_0 : \sigma^2 = \sigma_0^2$ verso $H_1 : \sigma^2 > \sigma_0^2$, dove σ_0^2 è un valore prefissato. In tal caso risulta

$$\chi_c^2 = \frac{NS^2}{\sigma_0^2} \sim \chi_{(N-1)}^2 \quad (10)$$

e si rifiuta H_0 al livello α se

$$\chi_c^2 > \chi_{(N-1; \alpha)}^2 \quad (11)$$

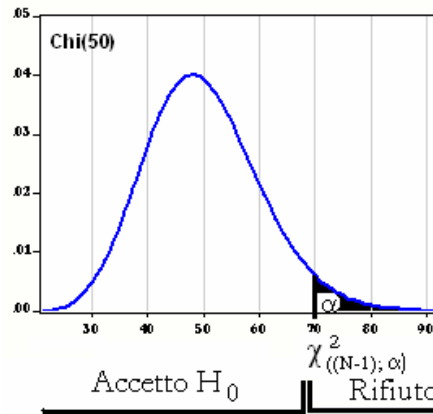
dove $\chi_{(N-1; \alpha)}^2$ è il percentile della v.c. chi quadrato che stacca alla sua destra una probabilità pari ad α . D'altra parte, abbiamo già sottolineato come la v.c. $\chi_{(N-1)}^2$ sia ben approssimata da una normale $N[(N-1); 2(N-1)]$ e quindi al posto della (11) si può considerare la regione di rifiuto approssimata

$$\chi_c^2 > (N-1) + z_{(\alpha)} \sqrt{2(N-1)}$$

Dove $z_{(\alpha)}$ è il percentile della normale standardizzata che stacca alla sua destra una probabilità pari ad α . Per $\alpha = 0.05$ si ha $z_{(0.05)} = 1.645$ ed in questo caso la regione di rifiuto approssimata diventa più semplicemente

$$\chi_c^2 > (N-1) + 1.645 \sqrt{2(N-1)}$$

Figura 3 – Regione di significatività per il test sulla varianza dei rendimenti



Nella tabella 4 riportiamo il χ_c^2 per le sei serie analizzate sotto l'ipotesi che, per ciascuna di essa, sia $H_0 : \sigma_0^2 = 0.004$.

Tabella 4 – Test sulla varianza dei rendimenti delle sei serie per $\sigma_0^2 = 0.004$

	CAMP	FIAT	INTSP	MIBTEL	NSDQ	S&P
S	0.0169	0.0244	0.0226	0.0119	0.0155	0.0133
S^2	0.0002856	0.000595	0.000511	0.000142	0.00024	0.000177
N	1613	1613	1609	1619	1619	1619
χ_c^2	1151.72	2400.79	2054.53	573.17	972.41	715.96
$\chi_{(N-1; 0.05)}^2$	1705.40	1705.40	1701.29	1711.58	1711.58	1711.58

Come si può osservare, l'ipotesi $H_0 : \sigma_0^2 = 0.004$ viene accettata per le serie Campari, Mibtel, Nasdaq, S&P, mentre viene rifiutata per Fiat e Banca Intesa S. Paolo.

Ritornando al modello ARMA(p, q) definito nella (6), se è $\phi_0 = 0$, come succede in pratica per i rendimenti degli attivi finanziari, questa equazione si può scrivere

$$r_t - \phi_1 r_{t-1} - \phi_2 r_{t-2} - \dots - \phi_p r_{t-p} = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}$$

Se introduciamo l'operatore B tale che $BX_t = X_{t-1}$, si avrà $B^k X_t = X_{t-k}$ e con queste notazioni la (6) diviene

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) r_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) \varepsilon_t$$

Posto

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) = \phi_p(B); \quad (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) = \theta_q(B)$$

il modello ARMA(p, q) può essere scritto in modo compatto come

$$\phi_p(B) r_t = \theta_q(B) \varepsilon_t \quad (12)$$

e si scrive $r_t \sim \text{ARMA}(p, q)$, dove $\phi_p(B) = 0$ è il polinomio *autoregressivo* in B di grado p e $\theta_q(B) = 0$ è quello *media mobile* di grado q . Si può dimostrare (vedi Box e Jenkins (1976)) che il processo r_t rappresentabile con i modelli (6) e (12) è *stazionario ed ergodico* se tutte le soluzioni radici di $\phi_p(B) = 0$ (radici del polinomio $\phi_p(B)$ che nel seguito indicheremo con λ_i , $i=1, 2, \dots, p$) sono in modulo maggiori di uno (fuori dal cerchio unitario) e quindi le loro inverse λ_i^{-1} minori di uno (dentro al cerchio unitario). Questo risultato è molto importante perché permette di stabilire abbastanza agevolmente la stazionarietà o meno di un processo stocastico rappresentabile con un ARMA. Se il processo è stazionario ed ergodico è possibile, utilizzando una serie storica di lunghezza N osservata sul fenomeno, stimarne i parametri e fare inferenza su questi. In particolare è possibile stabilire, con una probabilità di errore α , quali parametri possono considerarsi significativamente diversi da zero e quindi mantenerli nella (6) o nella (12), oppure non significativamente diversi da zero ed eliminarli in modo da utilizzare un modello che abbia il più piccolo numero di parametri possibile (*principio di parsimonia* del modello).

Se tutte le q soluzioni di $\theta_q(B) = 0$ (le *radici* di $\theta_q(B)$) sono in modulo maggiori di uno il processo r_t rappresentabile con la (6) o la (12) è *invertibile*. L'invertibilità è una condizione molto importante perché permette di costruire *previsioni ottimali* sul fenomeno oggetto di studio. Nel seguito si verificherà sempre che il processo stocastico oggetto di studio sia stazionario ed invertibile

e se questo non dovesse accadere si utilizzeranno particolari trasformazioni che lo rendono tale.

Osserviamo che:

- se è $q = 0$ la (12) diviene $r_t \sim \text{ARMA}(p, 0) = \text{AR}(p)$ ed r_t è sempre *invertibile*;
- se è $p = 0$ la (12) diviene $r_t \sim \text{ARMA}(0, q) = \text{MA}(q)$ ed r_t è sempre *stazionario*.

Nel seguito si imporrà anche che tutte le radici di $\phi_p(B)$ siano diverse da quelle di $\theta_q(B)$. Questo vincolo è necessario per evitare che nella (12) le radici comuni si cancellino ed il modello risulti mal specificato e le sue stime non attendibili. Infatti se $\phi_p(B)$ e $\theta_q(B)$ avessero, per esempio, una radice uguale che indichiamo con λ , si avrebbe

$$(1 - \lambda^{-1}B)\phi_{p-1}^*(B) = \phi_p(B); \quad (1 - \lambda^{-1}B)\theta_{q-1}^*(B) = \theta_q(B)$$

che sostituite nella (12) porterebbero alla cancellazione di $(1 - \lambda^{-1}B)$ e darebbero $\phi_{p-1}^*(B)r_t = \theta_{q-1}^*(B)\varepsilon_t$ che implica un modello ARMA(p-1, q-1).

E' noto che dati due polinomi di grado finito, in questo caso $\phi_p(B)$ e $\theta_q(B)$, il loro rapporto è ancora un polinomio MA di grado infinito:

$$\frac{\theta_q(B)}{\phi_p(B)} = 1 + \psi_1 B + \psi_2 B^2 + \dots + \psi_k B^k + \dots = \psi(B)$$

$$\frac{\phi_p(B)}{\theta_q(B)} = 1 + \pi_1 B + \pi_2 B^2 + \dots + \pi_k B^k + \dots = \pi(B)$$

Ne segue che la (12), se è stazionaria ed invertibile, ammette le seguenti due rappresentazioni alternative:

1. quella *media mobile* $r_t \sim \text{MA}(\infty)$:

$$r_t = \frac{\phi_0}{\phi(B)} + \frac{\theta_q(B)}{\phi_p(B)} \varepsilon_t = \frac{\phi_0}{\phi(1)} + \sum_{j=1}^{\infty} \psi_j \varepsilon_{t-j} + \varepsilon_t \quad (13)$$

con

$$\sum_{j=1}^{\infty} \psi_j^2 < \infty$$

2. quella *autoregressiva* $r_t \sim \text{AR}(\infty)$:

$$r_t = \frac{\phi_0}{\theta(1)} + \varepsilon_t - \sum_{j=1}^{\infty} \pi_j r_{t-j} \quad (14)$$

con

$$\sum_{j=1}^{\infty} \pi_j^2 < \infty$$

In definitiva, se il processo è stazionario ed invertibile questo può essere rappresentato alternativamente con le (12) - (14) a seconda delle convenienze contingenti per cui si opera.

4.4 Il modello $AR(1)$ per i rendimenti

Nel paragrafo precedente abbiamo studiato, fra l'altro, il caso particolare in cui è $p = q = 0$ e quindi i rendimenti si comportano come un white noise. In questo supponiamo che sia $p = 1$ e $q = 0$ e quindi i rendimenti si comportano come un $AR(1)$:

$$(1 - \phi B)r_t = \phi_0 + \varepsilon_t \Leftrightarrow r_t = \phi_0 + \phi r_{t-1} + \varepsilon_t \quad (15)$$

di modo che per i logaritmi dei prezzi $p_t = \log(P_t)$ si abbia

$$p_t = \phi_0 + p_{t-1} + \phi r_{t-1} + \varepsilon_t$$

o equivalentemente

$$\nabla p_t = \phi_0 + \phi(\nabla p_t) + \varepsilon_t$$

dove si è posto $\nabla p_t = p_t - p_{t-1}$ ovvero la differenza prima dei logaritmi dei prezzi dell'attivo. Se i rendimenti si comportano come la (15) è necessario stimare i tre parametri $(\phi_0; \phi; \sigma^2)$. Dalla (15), considerando il valore medio di ambo i membri e ricordando che r_t è stazionario mentre $\varepsilon_t \sim \text{WN}(0; \sigma^2)$ si ha

$$\phi_0 = (1 - \phi)E(r_t) \quad (16)$$

Per stimare il *drift* ϕ_0 è necessario stimare la media dei rendimenti ed il parametro autoregressivo ϕ . Inoltre, se i rendimenti sono in media nulli anche il *drift* è pari a zero. Ancora, dalla (15) si ha che i rendimenti sono *stazionari* se la radice del polinomio $(1 - \phi B) = 0$ è, in valore assoluto, maggiore di 1. Ma l'unica radice di $(1 - \phi B)$ è data da $\lambda = 1/\phi$ ed è $|\lambda| > 1$ se e solo se risulta $|\phi| < 1$. In conclusione il processo dei rendimenti rappresentato dal modello AR(1) riportato nella (15) è stazionario se e solo se è $|\phi| < 1$ che è equivalente a $-1 < \phi < 1$. La rappresentazione MA(∞) di un AR(1) stazionario si ottiene utilizzando la (13) e quindi risolvendo rispetto ai ψ_j la seguente uguaglianza fra polinomi:

$$1 + \psi_1 B + \psi_2 B^2 + \psi_3 B^3 + \dots = \frac{1}{1 - \phi B}$$

In questo caso è agevole dimostrare che è $\psi_j = \phi^j, j=1, 2, 3, \dots$, ed ottenere

$$r_t = \frac{\phi_0}{1 - \phi} + \frac{1}{1 - \phi B} \varepsilon_t = \frac{\phi_0}{1 - \phi} + \varepsilon_t + \sum_{j=1}^{\infty} \phi^j \varepsilon_{t-j} \quad (17)$$

Da questa ultima uguaglianza, ricordando che $\varepsilon_t \sim \text{WN}(0; \sigma^2)$, si ricava

$$E(r_t) = \frac{\phi_0}{1 - \phi}; \quad \text{Var}(r_t) = \sigma_r^2 = \sigma^2 + \sum_{j=0}^{\infty} \phi^{2j} = \frac{\sigma^2}{1 - \phi^2} \quad (18)$$

ed entrambe sono definite se il processo è stazionario cioè se risulta $-1 < \phi < 1$. Per quanto riguarda le *autocovarianze*, dalla (17), tenendo conto che per costruzione è $\varepsilon_t \sim \text{WN}(0; \sigma^2)$, si ha $\text{cov}(r_{t-s}, \varepsilon_t) = 0$, e quindi

$$\gamma(h) = \text{Cov}(r_t, r_{t-h}) = \phi \gamma(h-1)$$

che per sostituzioni successive diviene

$$\gamma(h) = \begin{cases} \sigma^2 & \text{per } h = 0 \\ 1 - \phi^2 & \\ \phi^h \gamma(0) & \text{per } h > 0 \end{cases} \quad (19)$$

La (19) conferma che ha senso derivare $\gamma(h)$ solo se il processo è stazionario. Dalla (19) è possibile ottenere la funzione di *autocorrelazione* di un AR(1)

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \begin{cases} 1 & \text{per } h = 0 \\ \phi^h & \text{per } h > 0 \end{cases} \quad (20)$$

Dalla (20) si deriva che se il processo AR(1) è stazionario allora è anche

$$\sum_{h=0}^{\infty} \rho^2(h) = \sum_{h=0}^{\infty} \phi^{2h} = \frac{1}{1 - \phi^2}$$

ed il processo è *ergodico*. Inoltre più ϕ si avvicina ad 1 oppure a -1, più è lenta la tendenza di $\rho(h)$ a zero al crescere di h . In particolare come caso limite, almeno formalmente, risulta

$$\begin{cases} \rho(h) = 1 & \text{per } \phi = 1 \\ |\rho(h)| = 1 & \text{per } \phi = -1. \end{cases}$$

In conclusione, se è $r_t \sim \text{AR}(1)$, l'evoluzione dinamica nel tempo di r_t è determinata dai tre parametri (ϕ_0, ϕ, σ^2) che, se il processo è stazionario, possono essere stimati a partire da una serie storica (r_1, r_2, \dots, r_N) osservata su

r_t . In particolare questa serie storica può essere utilizzata per stimare le prime M autocorrelazioni, in pratica è $M < N/4$, tramite la seguente

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)} = \frac{\sum_{t=1}^{N-h} (r_t - \bar{r})(r_{t+h} - \bar{r})}{\sum_{t=1}^N (r_t - \bar{r})^2}, \quad h=1, 2, \dots, M \quad (21)$$

Si può dimostrare che se il processo AR(1) è stazionario, la (21) fornisce una stima asintoticamente non distorta e consistente della funzione di autocorrelazione, mentre se il processo non è stazionario la (21) sottostima sistematicamente $\rho(h)$.

Dalla (21) per $h=1$ si ha $\rho(1) = \phi$ e quindi, ricordando le (18), una stima consistente per i tre parametri (ϕ_0, ϕ, σ^2) è data, rispettivamente, da

$$\hat{\phi} = \hat{\rho}(1); \quad \hat{\phi}_0 = \bar{r}(1 - \hat{\phi}); \quad S^2 = \hat{\gamma}(0)(1 - \hat{\phi}^2) \quad (22)$$

dove $\hat{\gamma}(0)$ è la stima campionaria della varianza di r_t . Le (20) permettono anche di stabilire se effettivamente il processo è rappresentabile con un AR(1). Infatti per $0 < \phi < 1$ la funzione $\rho(h)$ decresce esponenzialmente a zero al crescere di h , mentre per $-1 < \phi < 0$ decresce esponenzialmente a zero assumendo alternativamente valori negativi o positivi a seconda che h sia dispari o pari. Naturalmente un comportamento simile è atteso per la sua stima $\hat{\rho}(h)$. Quanto detto è sostanzialmente confermato dalle simulazioni riportate nella figura 6, in particolare quando ϕ non è troppo vicino allo zero.

Esempio 1

Supponiamo che r_t sia rappresentabile con il modello AR(1):

$$r_t = 0.2 + \phi r_{t-1} + e_t; \quad e_t \sim 0.3t_{(5)}; \quad \text{Var}(e_t) = \sigma^2 = 0.15$$

ed assuma i seguenti diversi valori del parametro ϕ :

$$\phi = (-1.0; -0.7; -0.3; 0.3; 0.7; 1.0)$$

Da queste posizioni, ricordando che $\text{Var}(t_{(k)}) = k/(k-2)$, si ha anche

Nella tabella 5 sono riportati gli indici descrittivi che mettono in evidenza le differenze strutturali, in media, variabilità, asimmetria e curtosi, esistenti fra le sei serie ottenute per i sei valori di ϕ . Nella figura 6 riportiamo le stime delle prime 20 autocorrelazioni ottenute con la (21) al variare di ϕ . Ricordiamo che in questa simulazione la banda di confidenza per $\rho(h) = 0$ è

$$\left(-2/\sqrt{1500} = -0.052; 2/\sqrt{1500} = 0.052\right)$$

qualsiasi sia ϕ . Come si può notare da questa figura, le autocorrelazioni per $\phi = -1; -0.7; 0.7; 1$ sono tutte fuori di tale banda mentre per $\phi = -0.3$ e 0.3 lo sono solo le prime e questo a conferma del fatto che al tendere di ϕ a zero il processo è sempre meno distinguibile da un WN.

Figura 5 – Istogramma di frequenze delle sei serie $AR(1)$ al variare di ϕ

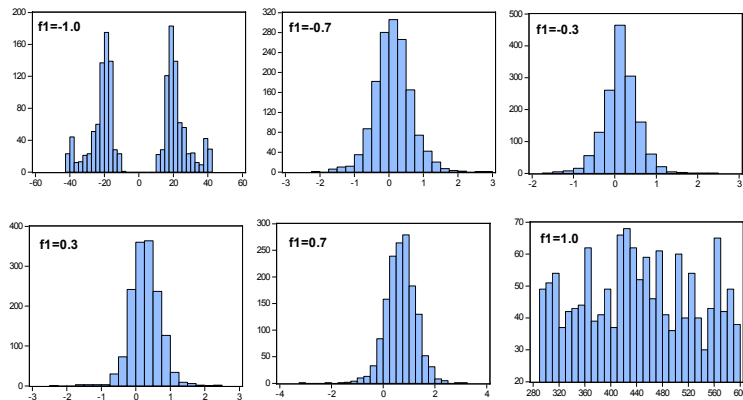
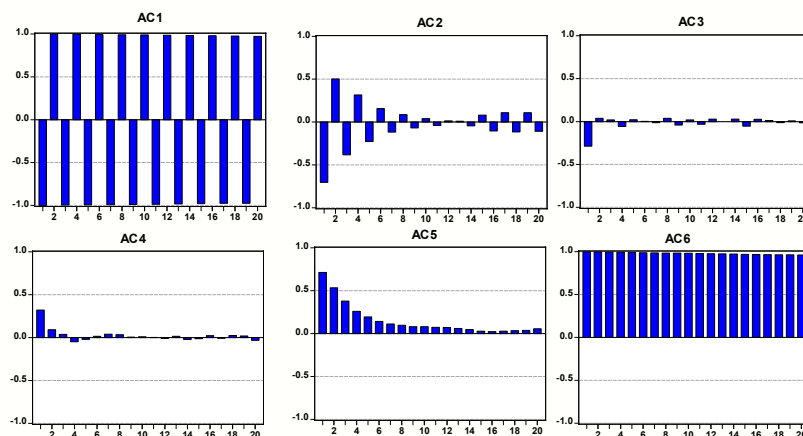


Figura 6 – Autocorrelazioni delle sei serie $AR(1)$ al variare di ϕ



Quanto detto per la tabella 5 sulla diversità delle serie storiche ed in particolare delle due non stazionarie è confermato dalla struttura degli istogrammi di frequenza riportati nella figura 5. Ricordiamo che le serie indicate con r1 ed r6 sono non stazionarie e le relative stime dei parametri, degli istogrammi e delle autocorrelazioni sono distorte ed inconsistenti, mentre le altre sono asintoticamente non distorte e consistenti. Utilizzando i dati della tabella 5 e le (22) sono state elaborate le stime preliminari e confrontate con i valori veri nella tabella 6. Queste stime confermano empiricamente come effettivamente esse siano accettabili solo quando il processo è stazionario.

Tabella 6 – *Stime dei momenti e valori veri del modello AR(1) per le sei serie*

	R1		r2		r3		r4		r5		r6	
	veri	stim	veri	stim	veri	stim	veri	stim	veri	stim	veri	stim
ϕ_0	0.2	0.220	0.2	0.203	0.2	0.189	0.2	0.190	0.2	0.205	0.2	0.89
ϕ	-1.0	-0.99	-0.7	-0.7	-0.3	-0.29	0.3	0.32	0.7	0.71	1.0	0.99
σ^2	0.15	1.11	0.15	0.144	0.15	0.165	0.15	0.159	0.15	0.168	0.15	30.6
μ		0.109	0.118	0.119	0.154	0.147	0.286	0.279	0.667	0.684		443
σ_x^2		556	0.294	0.285	0.165	0.170	0.165	0.177	0.294	0.329		7654

□

Le stime che abbiamo riportato nell'esempio precedente sono quelle dei momenti che, si dimostra, non sono le ottimali. Se le stime vengono derivate con il *metodo dei minimi quadrati*, esse risultano più efficienti. Le stime dei minimi quadrati si ottengono calcolando dalla (15) la somma dei quadrati dei residui

$$\sum_{t=2}^N \varepsilon_t^2 = \sum_{t=2}^N (r_t - \phi_0 - \phi r_{t-1})^2$$

minimizzandola rispetto ai due parametri incogniti (ϕ_0, ϕ) ed ottenendo $(\hat{\phi}_0, \hat{\phi})$. Si derivano quindi i residui stimati $\hat{\varepsilon}_t = r_t - \hat{\phi}_0 - \hat{\phi} r_{t-1}$ che si utilizzano per ottenere una stima di $\sigma^2 = \text{Var}(\varepsilon_t)$ tramite la seguente

$$S^2 = \frac{1}{N-1} \sum_{t=2}^N \hat{\varepsilon}_t^2$$

In altri termini si procede come per la stima dei parametri del modello di regressione lineare, l'unica differenza è che in questo contesto l'unica variabile esplicativa è rappresentata dalla ritardata r_{t-1} . Inoltre, così come avviene nel modello di regressione, è possibile derivare l'errore standard della stima di ciascun parametro ed utilizzarlo per costruire il relativo test. Anche nell'ambito delle serie storiche si costruisce l'indice di accostamento R^2 e si procede ad una analisi dei residui per stabilire se effettivamente il modello selezionato è quello adeguato.

Esempio 1 (continua)

Riprendiamo l'esempio precedente. Le stime dei parametri con il metodo dei minimi quadrati sono riportate nella tabella 7 per ciascuna delle sei serie AR(1) ottenute al variare di ϕ . Nell'ultima riga è riportato l'inverso della radice di $(1-\phi B)$ che abbiamo indicato con λ^{-1} . Come si può notare da questa tabella, le stime dei minimi quadrati dei tre parametri (ϕ, ϕ_0, σ^2) sono sempre molto buone e molto vicine al valore vero anche quando il processo non è stazionario. In ogni caso i valori veri di ϕ e ϕ_0 sono sempre all'interno del relativo intervallo di confidenza che non contiene mai lo zero. Tutto questo è anche dovuto al fatto che il numero delle osservazioni N risulta, in questo esempio, piuttosto elevato come d'altronde succede nelle applicazioni ad attivi finanziari.

Tabella 7 – *Stime LS dei parametri dei modelli AR(1). In parentesi gli errori standard delle stime*

	r1 $\phi = -1$	r2 $\phi = -0.7$	r3 $\phi = -0.3$	r4 $\phi = 0.3$	r5 $\phi = 0.7$	r6 $\phi = 1$
$\hat{\phi}_0$	0.203 (.0105)	0.203 (.0100)	0.189 (.0108)	0.190 (.0124)	0.197 (.0162)	0.165 (.0521)
$\hat{\phi}$	-1.00 (.000)	-0.703 (.018)	-0.287 (.025)	0.320 (.024)	0.710 (.018)	1.00 (.000)
$\hat{\sigma}^2$	0.1647	0.1440	0.1559	0.1589	0.1620	0.1528
R^2	0.999	0.494	0.082	0.102	0.508	0.999
$\hat{\lambda}^{-1}$	-1.00	-0.70	-0.29	0.32	0.71	1.00

□

Spesso nelle pratiche applicazioni invece che adattare i modelli ARMA

Figura 7 – Funzioni di densità stimate delle serie dei residui

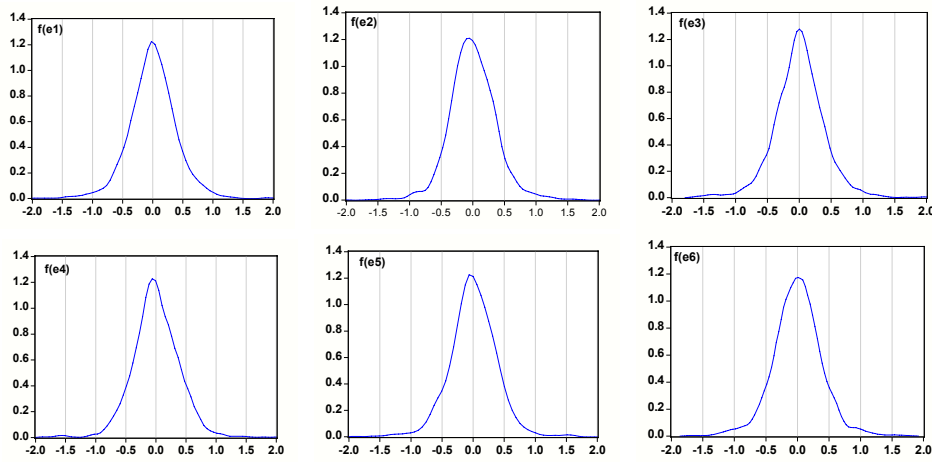
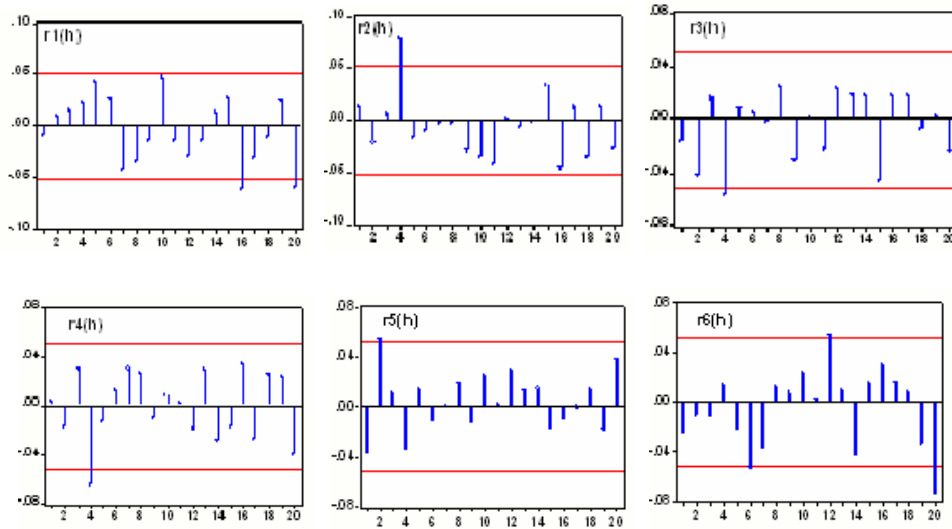


Figura 8 – Prime 20 autocorrelazioni dei residui e relativa banda di confidenza (in rosso) al 5%



Come si può osservare dalla tabella 8 i residui hanno tutti media nulla ed un marcato eccesso di curtosi come era da attendersi dato che sono stati generati da una *t* di Student con 5 gradi di libertà. La non normalità dei residui è anche confermata dal test di *Jarque – Bera* che rifiuta in tutti e sei i casi l'ipotesi di normalità. Dalla figura 8 emerge che, sostanzialmente, tutti i legami lineari sono stati catturati dal modello dato che quasi tutte le 20 autocorrelazioni stimate dei residui ricadono nella banda di confidenza con un errore del 5% (una ogni 20).

□

Vediamo cosa succede se il modello è mal specificato ed in particolare si aggiunge ad esso una variabile esplicativa che non dovrebbe essere presente. Supponiamo che il modello vero per r_t sia un AR(1), ma erroneamente si usa un AR(2). Questo vuol dire che si stima il modello

$$r_t = \phi_0 + \phi_1 r_{t-1} + \phi_2 r_{t-2} + \varepsilon_t \quad (23)$$

Osserviamo che se nella (23) è $\phi_2 = 0$ si ritorna al caso dell'AR(1). Dalla teoria sappiamo che le stime dei parametri $(\phi_0, \phi_1, \phi_2, \sigma^2)$ con le osservazioni generate dal modello (15) sono ancora consistenti, ma non le più efficienti. Questo vuol dire che le stime di $(\phi_0, \phi_1, \sigma^2)$ nel modello (23) sono simili a quelle ottenute con il (15), mentre quella di ϕ_2 dovrebbe risultare, sottoponendola a test, non significativamente diversa da zero di modo che può essere eliminata per riottenere la (15).

Esempio 1 (continua)

Nella tabella 9 sono riportate le stime dei parametri del modello (23) e confrontate con i parametri veri da cui il modello (15) è stato generato. Come si può osservare, la stima di ϕ_2 è per tutte e sei le serie non significativamente diversa da zero e la relativa variabile deve essere eliminata dal modello. Questo ci permette di identificare in modo corretto il modello "vero" (15). Le stime degli altri parametri sono molto simili a quelle della tabella 7.

Vediamo ora cosa succede se il modello stimato è mal specificato nel senso di eliminare una ritardata che invece dovrebbe esserci. Ritornando all'esempio precedente adattiamo ai dati generati dal modello AR(1) un white noise. In altri termini alle osservazioni generate da

$$r_t = 0.2 + \phi r_{t-1} + e_t; \quad e_t \sim 0.3 t_{(5)}; \quad \text{Var}(e_t) = \sigma^2 = 0.15$$

adattiamo il modello mal specificato

$$r_t = \phi_0 + e_t \quad (24)$$

Il modello è mal specificato per aver omesso la variabile r_{t-1} . Stimiamo i due

parametri (ϕ_0, σ^2) di questo modello con il metodo dei minimi quadrati. I risultati della stima sono presentati nella tabella 10.

Tabella 9 – *Stime dei parametri del modello (23) quando quello vero è il (15)*

$\phi_0; \phi_1$	r1 0.2; -1.0	r2 0.2; -0.7	r3 0.2; -0.3	r4 0.2; 0.3	r5 0.2; 0.7	r6 0.2; 1.0
$\hat{\phi}_0$	0.205 (.0117)	0.20 (.0113)	0.199 (.0118)	0.192 (.0133)	0.187 (.0169)	0.169 (.0523)
$\hat{\phi}_1$	-1.01 (.000)	-0.69 (.026)	-0.302 (.026)	0.325 (.026)	0.676 (.026)	0.977 (.026)
$\hat{\phi}_2$	-0.010 (.026)	0.018 (.026)	-0.052 (.026)	-0.013 (.026)	0.051 (.026)	0.023 (.026)
S^2	0.1647	0.1441	0.1555	0.1590	0.1617	0.1528
R^2	0.999	0.494	0.085	0.103	0.509	0.999
$ \hat{\lambda}_1 $	1.00	0.72	0.23	0.28	0.74	1.00
$ \hat{\lambda}_2 $	0.01	0.02	0.23	0.05	0.07	0.02

In questo caso le stime ottenute sono molto lontane dai parametri veri indipendentemente dal fatto che il processo sia stazionario o meno. Più precisamente le stime ottenute sono distorte ed inconsistenti. Questo vuol dire che nella formulazione di un modello, quando non si hanno informazioni teoriche su come deve essere specificato, è meglio utilizzare modelli con un numero maggiore di parametri per poi eliminarli, se sono non significativamente diversi da zero, che utilizzare modelli mal specificati per aver eliminato qualche parametro che invece doveva essere preso in considerazione.

Tabella 10 – *Stime dei parametri del modello (24) quando il vero è il (15)*

$\phi_0; \phi_1$	r1 0.2; -1.0	r2 0.2; -0.7	r3 0.2; -0.3	r4 0.2; 0.3	r5 0.2; 0.7	r6 0.2; 1.0
$\hat{\phi}_0$	0.109 (.609)	0.119 (.0134)	0.147 (.0106)	0.279 (.0109)	0.684 (.0148)	443.49 (.2259)
S^2	555.9	0.2845	0.1697	0.1769	0.3106	765.4
R^2	0.000	0.000	0.000	0.000	0.000	0.000

□

Le analisi ed i risultati statistici del modello AR(1) possono essere estesi al più generale AR(p), con $p > 0$. In questo caso è necessario stimare $p+2$ parametri e verificare che il processo sia stazionario calcolando le p radici del polinomio $1 - \phi_1 B - \dots - \phi_p B^p = 0$. La stima dei parametri $(\phi_0, \phi_1, \dots, \phi_p)$ si ottiene minimizzando la somma dei quadrati dei residui:

$$\sum_{t=p+1}^N \varepsilon_t^2 = \sum_{t=p+1}^N (r_t - \phi_0 - \phi_1 r_{t-1} - \dots - \phi_p r_{t-p})^2 = \min_{\phi_0, \phi_1, \dots, \phi_p}$$

ottenendo $(\hat{\phi}_0, \hat{\phi}_1, \dots, \hat{\phi}_p)$; ciò fatto si calcolano i residui stimati

$$\hat{\varepsilon}_t = r_t - \hat{r}_t = r_t - \hat{\phi}_0 - \hat{\phi}_1 r_{t-1} - \dots - \hat{\phi}_p r_{t-p}, \quad t=p+1, p+2, \dots, N$$

e si stima σ^2 con la seguente:

$$S^2 = \frac{1}{N-p} \sum_{t=p+1}^N \hat{\varepsilon}_t^2$$

Ricordiamo che in questo caso i gradi di libertà sono $g = N - p - 1$. Si procede, quindi all'analisi dei residui stimati.

4.5 Test sulla presenza di radici unitarie

I risultati presentati nelle pagine precedenti possono essere utilizzati per costruire un test che permetta, formalmente, di verificare la presenza di *radici unitarie* ($\lambda = \lambda^{-1} = 1$) nel processo generatore dei dati osservati. Questa problematica è stata affrontata in Dickey e Fuller (1979) ed approfondita, fra gli altri, in Phillips (1987), Chan e Wei (1988), Phillips e Perron (1988). Supponiamo che $p_t = \log(P_t)$ sia rappresentabile dal modello AR(1)

$$p_t = \phi p_{t-1} + \varepsilon_t; \quad \varepsilon_t \stackrel{\text{iid}}{\sim} (0; \sigma^2) \quad (25)$$

e si voglia sottoporre a test l'ipotesi di radice unitaria e quindi che sia $H_0 : \phi = 1$ verso $H_1 : \phi < 1$. Sotto H_0 il parametro ϕ scompare dalla (25) che diventa un *random walk*. In altri termini, sotto H_0 la (25) diviene $p_t = p_{t-1} + \varepsilon_t$ che, come sappiamo non è stazionario. In queste condizioni non esiste una distribuzione standard per p_t e risulta difficile fare inferenza sulla sua dinamica. La (25), sottraendo ad ambo i membri p_{t-1} , si può anche scrivere

$$\nabla p_t = (\phi - 1)p_{t-1} + \varepsilon_t \tag{26}$$

e le ipotesi di radici unitarie diventano $H_0 : \alpha = (\phi - 1) = 0$ verso $H_1 : \alpha < 0$. Dato che nella (25) si è supposto $E(p_t) = \mu = 0$, la stima dei minimi quadrati di ϕ e del suo errore standard sono dati rispettivamente da

$$\hat{\phi} = \hat{\rho}(1) = \frac{\hat{\gamma}(1)}{\hat{\gamma}(0)} = \frac{\sum_{t=2}^N p_t p_{t-1}}{\sum_{t=1}^N p_t^2}; \quad \sigma_{\hat{\phi}}^2 = \frac{\sigma^2}{\sum_{t=1}^N p_t^2}$$

che permettono, sotto H_0 , di derivare la statistica test

$$\begin{aligned} DF_c &= \frac{\hat{\phi} - 1}{\hat{\sigma}_{\hat{\phi}}} = \frac{1}{\hat{\sigma}_{\hat{\phi}}} \left[\frac{\sum_{t=2}^N p_t p_{t-1}}{\sum_{t=1}^N p_t^2} - 1 \right] = \frac{1}{\hat{\sigma}_{\hat{\phi}}} \left[\frac{\sum_{t=2}^N (p_t - p_{t-1}) p_{t-1}}{\sum_{t=1}^N p_t^2} - \frac{p_N^2}{\sum_{t=1}^N p_t^2} \right] \approx \\ &\approx \frac{\sum_{t=2}^N \varepsilon_t p_{t-1}}{\hat{\sigma} \left(\sum_{t=1}^N p_t^2 \right)^{1/2}} \end{aligned} \tag{27}$$

Per poter individuare la regione critica del test è necessario derivare, almeno in via asintotica, la distribuzione di DF_c . Gli Autori citati dimostrano che la (27) non ha una distribuzione standard (come la t di Student), ma è funzione del processo stocastico di Wiener. Più in particolare Dickey e Fuller dimostrano quanto riportato nella tabella 11.

Tabella 11 – Distribuzione asintotica di DF_c sotto l'ipotesi $\phi - 1 = 0$

Modello vero	Modello stimato	T. convergenza	Distribuzione asintotica
$p_t = \phi p_{t-1} + \varepsilon_t$	$p_t = \phi p_{t-1} + \varepsilon_t$	$1/N$	$\frac{W^2(1) - \sigma_r^2 / \sigma^2}{2 \int_0^1 W^2(s) ds}$
$p_t = \phi p_{t-1} + \varepsilon_t$	$p_t = \phi_0 + \phi p_{t-1} + \varepsilon_t$	$1/N$	$\frac{W^2(1) - \sigma_r^2 / \sigma^2 - 2W(1) \int_0^1 W(s) ds}{2 \int_0^1 W^2(s) ds - 2 \left[\int_0^1 W(s) ds \right]^2}$

Questi risultati sono stati estesi al caso più generale in cui p_t ha una radice unitaria, ma non è un random walk ed in particolare quando è

$$p_t = \phi p_{t-1} + \alpha_1 (\nabla p_{t-1}) + \alpha_2 (\nabla p_{t-2}) + \dots + \alpha_k (\nabla p_{t-k}) + e_t$$

dove è $e_t \stackrel{iid}{\sim} (0; \sigma^2)$. Sottraendo ad ambo i membri di questa equazione p_{t-1} si ottiene la generalizzazione della (26):

$$\nabla p_t = (\phi - 1) p_{t-1} + \alpha_1 (\nabla p_{t-1}) + \alpha_2 (\nabla p_{t-2}) + \dots + \alpha_k (\nabla p_{t-k}) + e_t$$

che permette di costruire la nuova statistica test ADF_c (*Augmented Dickey Fuller*) simile alla (27). I due Autori dimostrano che, sotto H_0 , ADF_c si distribuisce asintoticamente con la stessa distribuzione riportata nella tabella 11. Fissato α , i valori critici per il test $ADF_{(\alpha)}$ sono stati derivati per simulazione e risulta:

(A) se è $\phi_0 = 0$

$$ADF_{(0.01)} = -2.57; \quad ADF_{(0.05)} = -1.94$$

(B) se è $\phi_0 \neq 0$

$$ADF_{(0.01)} = -3.43; \quad ADF_{(0.05)} = -2.86$$

Al livello α

- si rifiuta H_0 se

$$ADF_c < ADF_{(\alpha)}$$

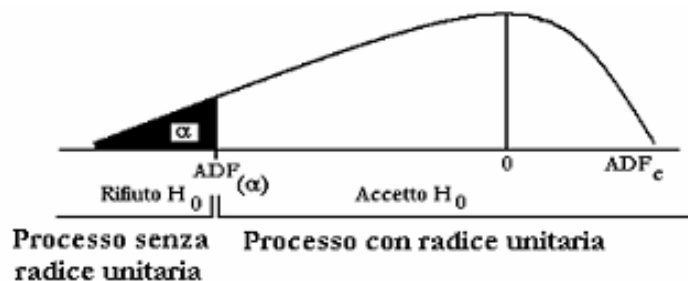
ed in tal caso il processo non ha una radice pari ad 1, ma comunque può essere non stazionario,

- si accetta H_0 se

$$ADF_c \geq ADF_{(\alpha)}$$

ed il processo ha una radice pari ad 1 e quindi è anche non stazionario.

Figura 9 – Regione di accettazione e di rifiuto del test ADF



La figura 9 schematizza le regioni di accettazione e di rifiuto del test ADF. Osserviamo che dire che il processo non ha una radice unitaria non vuol dire necessariamente che sia stazionario, infatti se è $\lambda = -1$, il processo non è stazionario, ma non ha una radice unitaria e quindi la statistica test ADF_c cadrà nella regione di rifiuto di H_0 .

Esempio 1 (continua)

Qui di seguito applichiamo il test ADF alle sei serie dell'esempio 1 ed alle loro

differenze: $r_{it} - r_{it-1}$, $i=1, 2, \dots, 6$. In questo esempio la $r1$ è non stazionaria, ma non ha radice pari ad 1, le serie da $r2$ a $r5$ sono stazionarie, la serie $r6$ ha una radice unitaria e quindi è anche non stazionaria. Da quanto detto in precedenza, ci attendiamo che il test ci faccia rifiutare H_0 per le serie da $r1$ ad $r5$ e ce la faccia accettare solo per $r6$. I risultati del test sono riportati nella tabella 12.

Tabella 12 – Test ADF per le sei serie dell'esempio 1 ove è $\phi_0 \neq 0$

	r1	r2	r3	r4	r5	r6
ADF_c	-4496.0	-92.67	-52.01	-27.74	-15.86	0.7607
$ADF_{(0.01)}$	-3.43	-3.43	-3.43	-3.43	-3.43	-3.43
$ADF_{(0.05)}$	-2.86	-2.86	-2.86	-2.86	-2.86	-2.86
Decisione	Rifiuto	Rifiuto	Rifiuto	Rifiuto	Rifiuto	Accetto

Quanto riportato nella tabella 12 è del tutto conforme alle attese. Lo stesso test è stato utilizzato per le differenze prime di ciascuna delle sei serie. In questo caso nessuna delle sei serie ha una radice pari ad uno e ci attendiamo che il test ci porti sempre al rifiuto dell'ipotesi nulla di radice unitaria. I risultati, riportati nella tabella 13, sono perfettamente conformi alle attese.

Tabella 13 – Test ADF per le differenze delle sei serie AR(1)

	$\nabla r1$	$\nabla r2$	$\nabla r3$	$\nabla r4$	$\nabla r5$	$\nabla r6$
ADF_c	-18.06	-18.87	-19.95	-18.73	-19.04	-39.64
$ADF_{(0.01)}$	-2.57	-2.57	-2.57	-2.57	-2.57	-3.42
$ADF_{(0.05)}$	-1.94	-1.94	-1.94	-1.94	-1.94	-2.86
Decisione	Rifiuto	Rifiuto	Rifiuto	Rifiuto	Rifiuto	Rifiuto

Osserviamo che nella tabella 13 si è tenuto conto che in $\nabla r6$ è conservata l'intercetta $\phi_0 = 0.2$, mentre nelle altre cinque serie non lo è.

Lo stesso test è stato applicato alle serie degli indici dei prezzi degli attivi: Cac, Dax, Mibtel, Nasdaq, Nasdaq_f ed ai relativi rendimenti. I risultati sono sintetizzati, rispettivamente, nelle tabelle 14 e 15.

Tabella 14 – Test ADF per le serie Cac, Dax, Mibtel, Nasdaq, Nasdaq_f

	Cac	Dax	Mibtel	Nasdaq	Nasdaq_f
ADF_c	-1.181	-1.451	-0.650	-1.883	-1.404
$ADF_{(0.01)}$	-2.57	-2.57	-2.57	-2.57	-2.57
$ADF_{(0.05)}$	-1.94	-1.94	-1.94	-1.94	-1.94
Decisione	Accetto	Accetto	Accetto	Accetto	Accetto

Tabella 15 – Test ADF per le serie dei rendimenti di Cac, Dax, Mibtel, Nasdaq, Nasdaq_f

	rCac	rDax	rMibtel	rNasdaq	rNasdaq_f
ADF_c	-20.69	-42.66	-18.14	-32.58	-34.04
$ADF_{(0.01)}$	-2.57	-2.57	-2.57	-2.57	-2.57
$ADF_{(0.05)}$	-1.94	-1.94	-1.94	-1.94	-1.94
Decisione	Rifiuto	Rifiuto	Rifiuto	Rifiuto	Rifiuto

Come atteso, mentre le serie degli indici dei prezzi hanno tutte una radice unitaria, quelle dei rendimenti ne sono sempre prive.

□

4.6 Il modello MA(1) per i rendimenti

La presenza di una struttura autoregressiva nei rendimenti è molto rara visto che implica una “memoria” infinita del fenomeno, mentre è più frequente quella media mobile con, di solito, $q \leq 5$. Le strutture media mobile, a differenza di quelle autoregressive, hanno una “memoria” del passato che non supera q . In questo senso un modello spesso utilizzato per rappresentare i livelli dei rendimenti è l'MA(q) con $q \leq 5$ (memoria massima della settimana borsistica) ed in particolare l'MA(1) (memoria del giorno precedente):

$$r_t = \phi_0 + \varepsilon_t - \theta\varepsilon_{t-1}; \quad \varepsilon_t \sim \text{WN}(0; \sigma^2) \quad (28)$$

Se i rendimenti si comportano come la (28) si ha

$$E(r_t) = \phi_0; \quad \gamma(0) = \text{Var}(r_t) = \sigma^2(1 + \theta^2) \quad (29)$$

mentre è

$$\gamma(h) = \begin{cases} -\theta\sigma^2 & \text{per } h = 1 \\ 0 & \text{per } h > 1 \end{cases}$$

e per l'autocorrelazione risulta

$$\rho(h) = \begin{cases} -\frac{\theta}{1+\theta^2} & \text{per } h = 1 \\ 0 & \text{per } h > 1. \end{cases} \quad (30)$$

Da queste relazioni si ha che il processo MA(1) è sempre stazionario. Dalla (30) risulta, come già sottolineato, che l'autocorrelazione dell'MA(1) ha memoria solo fino a $q=1$. Per i rendimenti, nella plausibile ipotesi di assenza di arbitraggio, di solito è $\phi_0 = 0$. Inoltre il polinomio $(1-\theta B)$ possiede l'unica radice $\lambda = 1/\theta$ ed il processo è *invertibile* solo se è $-1 < \theta < 1$. Se il processo MA(1) è invertibile può essere rappresentato dall'AR(∞):

$$\varepsilon_t = -\frac{\phi_0}{1-\theta} + \sum_{j=1}^{\infty} \theta^j r_{t-j}$$

Dalla (30) per $h=1$ si ottiene $\rho(1)\theta^2 + \theta + \rho(1) = 0$ e risolvendo rispetto a θ si ottiene

$$\theta = \frac{-1 \pm \sqrt{1-4\rho^2(1)}}{2\rho(1)} \quad (31)$$

Delle due soluzioni della (31) si sceglie quella, se esiste, con $-1 < \theta < 1$. Osserviamo che la (31) ha senso solo se è $1-4\rho^2(1) > 0$ che implica $-0.5 < \rho(1) < 0.5$ e questa è una condizione necessaria perché un processo sia rappresentabile con un MA(1) invertibile. La (29) e la (31) possono essere utilizzate per derivare stime dei momenti dei tre parametri $(\phi_0; \theta; \sigma^2)$ tramite le seguenti

$$\hat{\phi}_0 = \bar{r}; \quad \hat{\theta} = \frac{1 \pm \sqrt{1-4\hat{\rho}^2(1)}}{2\hat{\rho}(1)}; \quad S^2 = \frac{\hat{\gamma}(0)}{1+\hat{\theta}^2} \quad (32)$$

Di solito queste stime non sono molto efficienti. Stime alternative alle (32) si ottengono utilizzando iterativamente il metodo dei minimi quadrati a partire dalla (28):

1. nella (28) si pone $\hat{\phi}_0^{(0)} = \bar{r}$, $\varepsilon_0 = 0$ e si ottiene $\hat{\varepsilon}_t^{(1)} = r_t - \bar{r}$;
2. si considera $\varepsilon_t = r_t - \phi_0 + \theta \hat{\varepsilon}_{t-1}^{(1)}$, si stimano $(\phi_0; \theta)$ minimizzando

$$\sum_{t=2}^N (r_t - \phi_0 + \theta \hat{\varepsilon}_{t-1}^{(1)})^2$$

si ottengono le nuove stime $(\hat{\phi}_0^{(1)}; \hat{\theta}^{(1)})$ che permettono di derivare i nuovi residui

$$\hat{\varepsilon}_t^{(2)} = r_t - \hat{\phi}_0^{(1)} + \hat{\theta}^{(1)} \hat{\varepsilon}_{t-1}^{(1)}, \quad t=2, 3, \dots, N$$

3. si considera $\varepsilon_t = r_t - \phi_0 + \theta \hat{\varepsilon}_{t-1}^{(2)}$, si stimano $(\phi_0; \theta)$ minimizzando

$$\sum_{t=2}^N (r_t - \phi_0 + \theta \hat{\varepsilon}_{t-1}^{(2)})^2$$

si ottengono le nuove stime $(\hat{\phi}_0^{(2)}; \hat{\theta}^{(2)})$ che permettono di generare i nuovi residui

$$\hat{\varepsilon}_t^{(3)} = r_t - \hat{\phi}_0^{(2)} + \hat{\theta}^{(2)} \hat{\varepsilon}_{t-1}^{(2)}, \quad t=2, 3, \dots, N$$

4. si ripete la procedura (2)-(3) m-volte fino a che le stime dei parametri convergono. Di solito bastano poche iterazioni, di solito non più di 15, per ottenere le stime finali;
5. si derivano i residui finali

$$\hat{\varepsilon}_t^{(m)} = r_t - \hat{\phi}_0^{(m-1)} + \hat{\theta}^{(m-1)} \hat{\varepsilon}_{t-1}^{(m-1)}, \quad t=2, 3, \dots, N$$

e si utilizzano per ottenere una stima della varianza σ^2 di ε_t :

$$S^2 = \frac{1}{N-1} \sum_{t=2}^N (\hat{\varepsilon}_t^{(m)})^2$$

Anche in questo caso, per le stime finali dei parametri, è possibile calcolare la stima dello standard error ed utilizzarla per sottoporre a test le ipotesi $H_0 : \theta = 0$ verso $H_1 : \theta \neq 0$.

Esempio 2

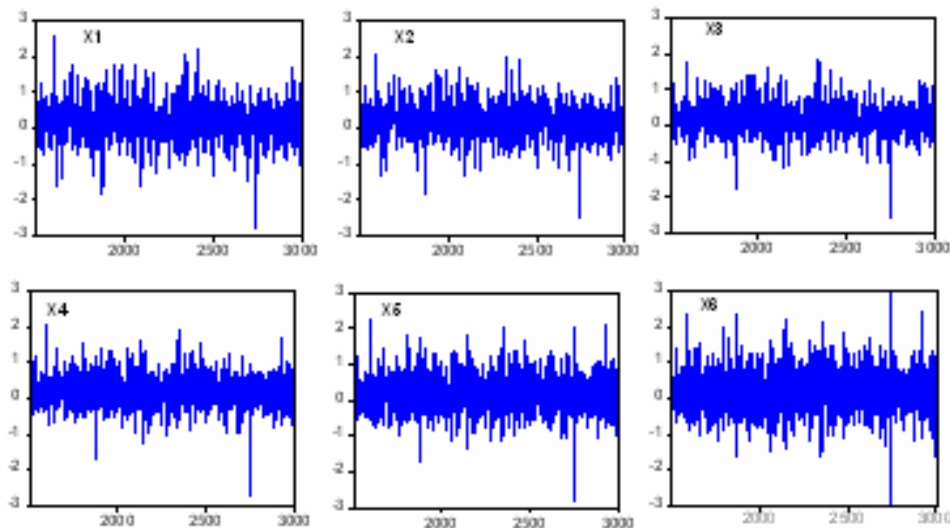
Supponiamo che r_t sia rappresentabile con il modello MA(1):

$$r_t = 0.2 + \varepsilon_t - \theta \varepsilon_{t-1}; \quad \varepsilon_t \sim 0.3t_{(5)}; \quad \text{Var}(\varepsilon_t) = \sigma^2 = 0.15$$

$$\text{Var}(r_t) = 0.15(1 + \theta^2); \quad \theta = -1.0; -0.7; -0.3; 0.3; 0.7; 1.0$$

Dal modello, per i diversi valori di θ sopra riportati, generiamo sei serie storiche di 3000 osservazioni ciascuna, di queste eliminiamo le prime 1500 per ridurre l'effetto del valore iniziale.

Figura 10 – Serie storiche dei sei MA(1) per i diversi valori di θ

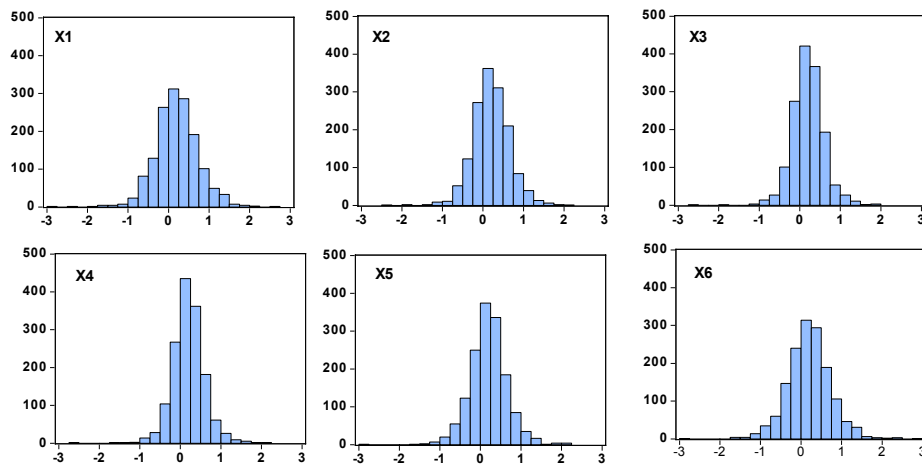


Le susseguenti analisi sono effettuate con le rimanenti $N=1500$ osservazioni. Nella figura 10 riportiamo i grafici delle sei serie storiche MA(1) sopra descritte.

Dalla tabella 16 osserviamo che tutte e sei le serie, come atteso, sono lontane dal caso di normalità dato che ad una trascurabile asimmetria corrisponde una accentuata curtosi, inoltre la r_1 e la r_2 non sono invertibili.

Tabella 16 – Indici sintetici descrittivi delle sei serie $MA(1)$

	r1	r2	r3	r4	r5	r6
\bar{r}	0.201657	0.201363	0.200970	0.200381	0.199988	0.199693
M_e	0.182621	0.185218	0.186592	0.190655	0.196895	0.192512
$\hat{\gamma}(0)$	0.524689	0.453103	0.388893	0.392780	0.460861	0.534255
$\hat{\rho}(1)$	0.50667	0.47457	0.26875	-0.29857	-0.49382	-0.52390
$\hat{\gamma}_1$	-0.053822	-0.073333	-0.118447	-0.081051	0.018203	0.075106
$\hat{\gamma}_2$	1.768941	1.870312	2.999295	3.462424	2.345165	2.000826
JB_c	196.2962	219.9736	565.7431	750.9162	343.8203	251.6168
p	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
N	1500	1500	1500	1500	1500	1500

Figura 11 – Istogrammi delle sei serie $MA(1)$ 

Queste considerazioni sono anche confermate dagli istogrammi di frequenza riportati nella figura 11. Nella figura 12 sono riportate le prime 20 autocorrelazioni stimate per le sei serie $MA(1)$ in cui sono indicate, in rosso, le relative bande di confidenza al 95%. Le sei serie sono compatibili con il modello $MA(1)$ dato che solo $\hat{\rho}(1)$ è fuori della banda di confidenza. Nella tabella 17 riportiamo le stime preliminari derivate dalla tabella 16 utilizzando le (32). Dalla tabella 17 si può osservare che, se è $|\hat{\rho}(1)| > 0.5$, il processo non è invertibile, e non esistono le stime preliminari di θ e σ^2 . Notiamo ancora che

la stima di ϕ_0 è sempre accettabile mentre quella di $\sigma^2 = \gamma(0)$ è abbastanza imprecisa. Nella tabella 18 riportiamo le stime degli stessi parametri ottenute con il metodo LS iterativo utilizzando la procedura sopra descritta. Confrontando le stime preliminari riportate nella tabella 17 con quelle dei minimi quadrati iterativi riportate nella tabella 18 si può osservare che, quando quelle preliminari esistono, sono fra di loro equivalenti. Le stime dei parametri ottenute con il metodo iterativo esistono sempre anche quando il modello è non invertibile, inoltre consentono di derivare l'errore standard delle stime (in parentesi nella tabella 18) che permette di costruire test e/o intervalli di confidenza. Nella figura 13 sono riportate le prime 20 autocorrelazioni stimate dei residui delle sei serie MA(1) con le relative (in rosso) bande di confidenza al 95%.

Figura 12 – Autocorrelazioni delle sei serie MA(1)

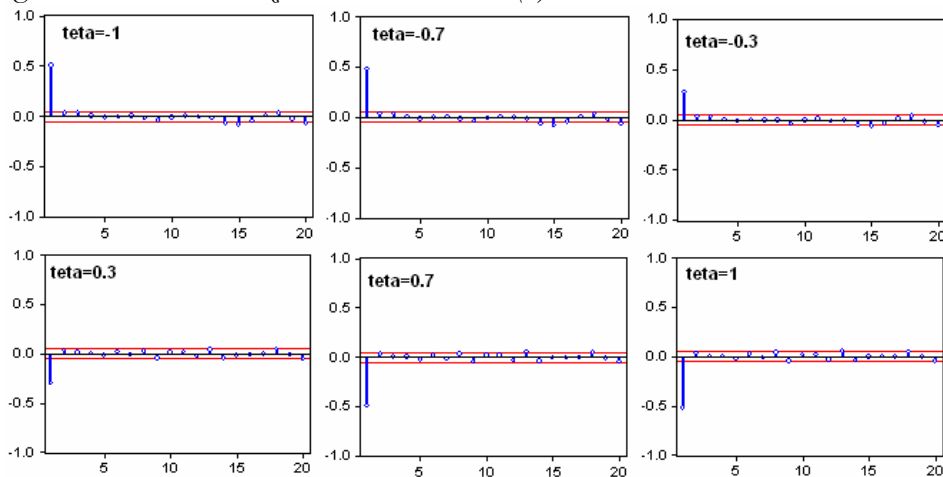
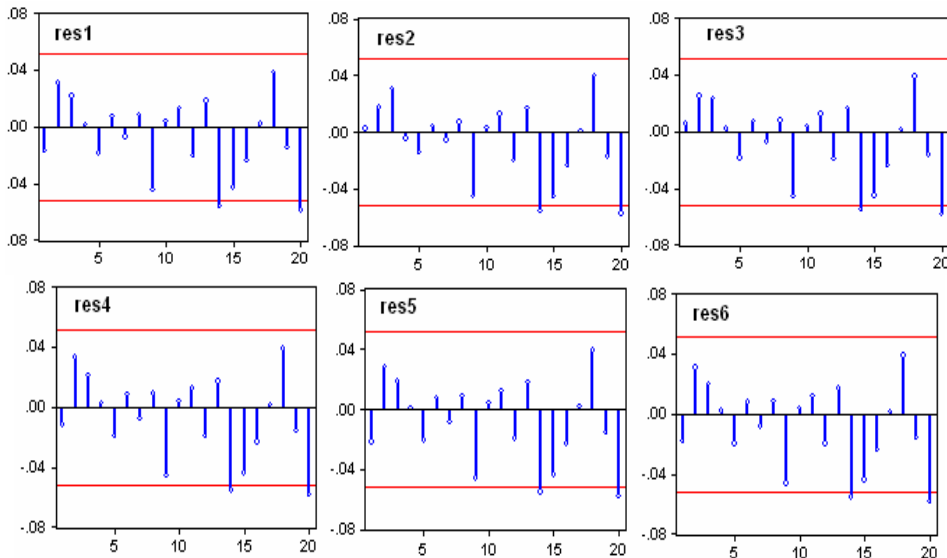


Tabella 17 – Stime preliminari del modello MA(1) con le (32)

	r1		r2		r3		r4		r5		r6	
	veri	stim	veri	stim	veri	stim	veri	stim	veri	stim	veri	stim
ϕ_0	0.2	.202	0.2	.201	0.2	.201	0.2	.200	0.2	.200	0.2	.200
θ	-1.0		-0.7	-.72	-0.3	-.29	0.3	.330	0.7	.850	1.0	
σ^2	0.15		0.15	.135	0.15	.139	0.15	.139	0.15	.123	0.15	
$\gamma(0)$	0.30	.275	0.22	.205	0.16	.151	0.16	.154	0.22	.212	0.30	.285

Come si può osservare, praticamente, tutte le autocorrelazioni dei residui sono all'interno della banda di confidenza e quindi non distinguibili da zero. Questo

Figura 13 – Autocorrelazioni dei residui dei sei MA(1)



□

Come già osservato, in un numero significativo di serie di rendimenti queste sono ben adattate da un MA(q) con q di solito non superiore a 5. In questo contesto più generale, le analisi statistiche possono essere fatte solo per via numerica utilizzando software specialistici.

4.7 Il modello ARMA(1, 1) per i rendimenti

In qualche raro caso i rendimenti di attivi possono essere rappresentati con il modello ARMA(1, 1):

$$r_t = \phi r_{t-1} + \varepsilon_t - \theta \varepsilon_{t-1}; \quad \varepsilon_t \sim \text{WN}(0; \sigma^2) \quad (33)$$

in cui è presente sia la componente AR, descritta dal parametro ϕ , che quella MA riferita al parametro θ . In questo caso è facile verificare che il processo è stazionario se risulta $-1 < \phi < 1$, mentre è invertibile se è $-1 < \theta < 1$. Inoltre si ha:

1. $E(r_t) = 0$

$$2. \quad \gamma(0) = \text{Var}(r_t) = \sigma^2 \frac{1 - 2\phi\theta + \theta^2}{1 - \phi^2}$$

che esiste se il processo è stazionario;

$$3. \quad \gamma(1) = \phi\gamma(0) - \theta\sigma^2$$

che esiste se il processo è stazionario;

$$4. \quad \gamma(h) = \phi\gamma(h-1)$$

e quindi a partire dal ritardo $h=2$ si comporta come un AR(1).

Questi risultati permettono di derivare la funzione di autocorrelazione del modello ARMA(1, 1) stazionario ed invertibile:

$$\rho(h) = \begin{cases} \frac{(1-\phi\theta)(\phi-\theta)}{1-2\phi\theta+\theta^2} & \text{per } h = 1 \\ \phi^{h-1} \rho(1) & \text{per } h = 2, 3, \dots \end{cases} \quad (34)$$

Dalla (34) si ha immediatamente che se $\phi = \theta$ allora è $\rho(h) = 0$, per $h > 0$, ed i rendimenti si comportano come un white noise. Possiamo così concludere che una condizione necessaria perché il processo si comporti come un ARMA(1, 1) è che sia $\phi \neq \theta$. La (33) si può scrivere

$$(1 - \phi B)r_t = (1 - \theta B)\varepsilon_t$$

che è stazionario ed invertibile solo se risulta contemporaneamente $-1 < \phi < 1$, $-1 < \theta < 1$, in tal caso la (33) ammette la rappresentazione MA(∞)

$$r_t = \frac{(1 - \theta B)}{(1 - \phi B)} \varepsilon_t = \varepsilon_t + \sum_{j=1}^{\infty} (\phi - \theta)\phi^{j-1} \varepsilon_{t-j}$$

e la rappresentazione AR(∞)

$$\varepsilon_t = \frac{(1 - \phi B)}{(1 - \theta B)} r_t = r_t + \sum_{j=1}^{\infty} (\theta - \phi)\theta^{j-1} r_{t-j}$$

ed anche da queste equazioni si vede immediatamente che se è $\phi = \theta$ il

modello ARMA(1, 1) si riduce al white noise ε_t . Queste considerazioni implicano che quando ad una serie storica osservata si adatta un modello ARMA(1, 1) bisogna verificare che le stime dei due parametri ϕ e θ siano significativamente diverse fra di loro.

Anche in questo caso è possibile derivare delle stime iniziali dei momenti dei tre parametri (ϕ, θ, σ^2) tenendo conto che è

$$\sigma^2 = \frac{\gamma(0)(1-\phi^2)}{1-2\phi\theta+\theta^2}; \quad \rho(1) = \frac{(1-\phi\theta)(\phi-\theta)}{1-2\phi\theta+\theta^2}; \quad \rho(2) = \phi\rho(1)$$

che formano un sistema, non lineare, di tre equazioni nelle tre incognite (ϕ, θ, σ^2) . Risolvendo il sistema e sostituendo ai valori teorici le stime $(\hat{\gamma}(0), \hat{\rho}(1), \hat{\rho}(2))$ si ottengono quelle iniziali dei momenti:

$$\hat{\phi} = \frac{\hat{\rho}(2)}{\hat{\rho}(1)}; \quad \hat{\theta} = \frac{-(1-2\hat{\phi}\hat{\rho}(1)+\hat{\phi}^2)}{2(\hat{\rho}(1)-\hat{\phi})} \pm \left[\left(\frac{1-2\hat{\phi}\hat{\rho}(1)+\hat{\phi}^2}{2(\hat{\rho}(1)-\hat{\phi})} \right)^2 - 1 \right]^{1/2}$$

$$\hat{\sigma}^2 = \frac{\hat{\gamma}(0)(1-\hat{\phi}^2)}{1-2\hat{\phi}\hat{\theta}+\hat{\theta}^2} \quad (35)$$

e delle due soluzioni che si derivano per $\hat{\theta}$ si sceglie quella con $-1 < \hat{\theta} < 1$ ed ha senso se

$$(1-2\hat{\phi}\hat{\rho}(1)+\hat{\phi}^2)^2 \geq (2(\hat{\rho}(1)-\hat{\phi}))^2$$

Esempio 3

Supponiamo che il processo r_t sia generato dal modello ARMA(1, 1)

$$r_t = \phi r_{t-1} + \varepsilon_t - \theta \varepsilon_{t-1}; \quad \varepsilon_t \stackrel{\text{iid}}{\sim} t_{(5)}$$

Partendo dal valore iniziale $r_0 = 0.2$ generiamo sei serie storiche ARMA(1,1) di 3000 osservazioni, una per ciascuna delle seguenti coppie di parametri (θ, ϕ) :

$$(-1; 0.8); (-.8; 0.5); (0.5; -0.5); (-0.5; -0.8); (0.5; 0.8); (1; -0.5)$$

di queste 3000 osservazioni utilizziamo nel seguito le ultime $N=1500$. In questo caso deve essere $\text{Var}(\varepsilon_t) = \text{Var}(t_{(5)}) = 5/3$. Nella tabella 20 riportiamo le stime delle prime due autocorrelazioni di ciascuno dei sei modelli ARMA(1,1), la stima della varianza della serie e quindi, usando la prima delle (35), quella del parametro autoregressivo ϕ .

Se si confronta la stima preliminare riportata nella penultima riga della tabella 20 con i valori veri dell'ultima riga si osserva che i risultati ottenuti sono sostanzialmente dignitosi, ma certamente migliorabili. Partendo dalla tabella 20, utilizzando le seconde due della (35), nella tabella 21 riportiamo le stime preliminari di $(\theta; \sigma^2)$ e le confrontiamo con il valore vero dei due parametri.

Tabella 20 – Stime di $\rho(1)$, $\rho(2)$, $\gamma(0)$ e di ϕ per i sei ARMA(1, 1)

	r1	r2	r3	r4	r5	r6
$\hat{\rho}(1)$	0.907	0.755	-0.724	-0.414	0.414	-0.760
$\hat{\rho}(2)$	0.741	0.399	0.380	0.344	0.351	0.397
$\hat{\gamma}(0)$	16.923	5.25	3.814	2.038	2.050	6.546
$\hat{\phi}$	0.817	0.528	-0.525	-0.831	0.848	-0.487
ϕ	0.80	0.50	-0.50	-0.80	0.80	-0.50

Tabella 21 – Stime preliminari di θ e σ^2 negli ARMA(1, 1)

	r1	r2	r3	r4	r5	r6
θ	-1.00	-0.80	0.50	-0.50	0.50	1.00
$\hat{\theta}$	-0.78	-0.71	0.47	-0.54	0.56	
σ^2	1.667	1.667	1.667	1.667	1.667	1.667
S^2	1.947	1.685	1.607	1.588	1.586	

Dalla tabella 21 si osserva che le stime preliminari dei due parametri (θ, σ^2) perdono di significato quando il modello non è invertibile come è per le due serie r1 ed r6. Nei casi in cui il modello è stazionario ed invertibile i risultati sono modesti anche con una numerosità abbastanza elevata di $N=1500$ come è in questo caso. Nella figura 14 riportiamo le autocorrelazioni stimate per le sei serie.

La figura 14 conferma quanto visto teoricamente: le autocorrelazioni di un ARMA(1,1) si comportano a partire dal ritardo $h=2$ come quelle di un

AR(1). Le stime dei tre parametri delle sei serie ARMA(1, 1) ottenute con il metodo dei minimi quadrati iterativo sono riportate nella tabella 22.

Figura 14 – Autocorrelazioni delle sei serie ARMA(1, 1)

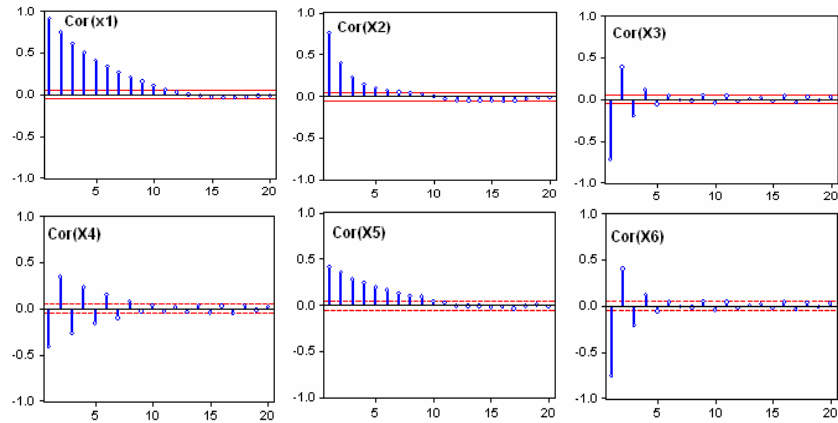


Tabella 22 – Stime LS dei parametri dei sei modelli ARMA(1, 1)

$\phi; \theta$	r1	r2	r3	r4	r5	r6
	0.8; -1.0	0.5; -0.8	-0.5; 0.5	-0.8; -0.5	0.8; 0.5	-0.5; 1.0
$\hat{\phi}$	0.812 (.0151)	0.515 (.0293)	-0.525 (.0274)	-0.807 (.0293)	0.825 (.0274)	-0.510 (.0222)
$\hat{\theta}$	-0.997 (.0000)	-0.788 (.0171)	0.477 (.0283)	-0.498 (.0431)	-0.525 (.0414)	0.997 (.0011)
S^2	1.598	1.596	1.597	1.597	1.597	1.601
R^2	.905	0.697	0.583	0.217	0.221	0.756
$\hat{\lambda}_{\phi}^{-1}$	0.81	0.51	-0.53	-0.81	0.83	-0.51
$\hat{\lambda}_{\theta}^{-1}$	-1.00	-0.79	0.48	-0.50	0.52	1.00

Come si può osservare queste stime sono nettamente migliori di quelle preliminari riportate nelle tabelle 20 e 21. Nella tabella 23 riportiamo i valori degli indici descrittivi di $\varepsilon_t \stackrel{iid}{\sim} t_{(5)}$ simulato e dei residui $\hat{\varepsilon}_t$ stimati nei sei modelli ARMA(1,1). Dai risultati riportati nella tabella 23 deduciamo che i residui stimati per le sei serie ARMA(1, 1) hanno una struttura molto simile a quelli simulati utilizzati per derivarle. Questo implica che le stime LS iterative dei parametri del modello sono molto buone.

Tabella 23 – Indici descrittivi dei residui simulati e di quelli stimati

	T5	RES1	RES2	RES3	RES4	RES5	RES6
$\bar{\epsilon}$	-0.00067	-0.00049	-0.00086	-5.33E-05	-0.00089	0.00031	-0.00796
M_c	-0.00361	-0.00321	-0.01260	-0.00859	-0.00526	-0.00765	-0.02209
Max	5.65234	5.75155	5.67036	5.67259	5.61600	5.67140	5.62823
Min	-6.03680	-5.92163	-6.03237	-6.03736	-6.03924	-5.98595	-6.02493
S	1.26364	1.26379	1.26306	1.26312	1.26338	1.26313	1.26485
$\hat{\gamma}_1$	0.05425	0.05242	0.05912	0.05798	0.05554	0.05502	0.05361
$\hat{\gamma}_2$	1.27628	1.27025	1.27624	1.27011	1.27521	1.25881	1.26532
JB_c	102.542	101.533	102.673	101.665	102.407	99.7958	100.783
p	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
N	1500	1500	1500	1500	1500	1500	1500

□

Le proprietà dei modelli e la procedura di stima dei parametri usata per l'ARMA(1, 1) possono essere estese, con le relative complicazioni formali, al caso generale degli ARMA(p, q). In questo caso le stime preliminari possono essere ottenute solo ricorrendo a software dedicati. E' comunque da ricordare che nelle applicazioni ai rendimenti di attivi ben difficilmente si utilizzano modelli di questa complessità.

4.8 Il modello TARMA per i rendimenti

Nei paragrafi precedenti abbiamo supposto che i livelli medi dei rendimenti fossero rappresentabili con modelli lineari ed in particolare con modelli della classe ARMA. D'altra parte gli attivi finanziari sono fenomeni molto complessi e le approssimazioni lineari, spesso, risultano grossolane. Per avere analisi più precise è necessario ricorrere ad una modellistica non lineare. Ma di modelli non lineari ne esistono svariate tipologie ciascuna delle quali possiede specifiche proprietà che ne rendono l'uso interessante. Fra le molteplici classi di modelli non lineari ne illustreremo, di seguito, una particolare che è una diretta generalizzazione degli ARMA. I modelli a cui ci riferiamo sono i modelli *Threshold* ARMA, in sigla TARMA o ARMA a soglia. Questi modelli sono molto flessibili, ne sono note le proprietà probabilistiche e, in particolare, le condizioni di stazionarietà ed invertibilità, le stime LS o quelle di massima verosimiglianza si ottengono abbastanza facilmente con le

ben note proprietà asintotiche.

I TARMA sono una generalizzazione dei TAR (Threshold AutoRegressive) introdotti da Tong (1978) ed approfonditi dallo stesso Tong (1980). Le caratteristiche probabilistiche dei TARMA sono state analizzate da numerosi autori, fra gli altri ricordiamo i lavori di Brockwell et al. (1992), Liu e Susko (1992), Ling (1999), Kapetanios e Shin (2006), Amendola et al. (2006) e (2009) che ne derivano momenti e condizioni sufficienti di stazionarietà. Questi modelli sono stati applicati sia a fenomeni finanziari che non. Fra gli altri ricordiamo i lavori di Li e Lam (1995) e Ling (1999) che li usano per lo studio delle asimmetrie e della volatilità nei mercati finanziari; Kräger e Kugler (1993), Peel e Speight (1994) e Chappel et al. (1996) che li applicano allo studio dei tassi nominali di cambio; Obstfeld e Taylor (1997) e O'Connell (1998) che li utilizzano per l'analisi dei tassi di cambio reali.

Il processo dei rendimenti r_t globali si distribuisce come un TARMA(2; r, q) se

$$r_t = \begin{cases} c_1 + \phi_{11}r_{t-1} + \dots + \phi_{1r}r_{t-r} + \varepsilon_t - \theta_{11}\varepsilon_{t-1} - \dots - \theta_{1q}\varepsilon_{t-q} & \text{se } X_{t-d} < v \\ c_2 + \phi_{21}r_{t-1} + \dots + \phi_{2r}r_{t-r} + \varepsilon_t - \theta_{21}\varepsilon_{t-1} - \dots - \theta_{2q}\varepsilon_{t-q} & \text{se } X_{t-d} \geq v \end{cases}$$

con X_t il processo soglia e $\varepsilon_t \stackrel{\text{iid}}{\sim} (0; \sigma^2)$. Il modello sopra riportato è il TARMA con due regimi, ma i risultati che presenteremo possono essere facilmente estesi al caso di $k > 2$ regimi. C'è da osservare che nelle pratiche applicazioni i regimi di solito sono al massimo tre. La presenza della variabile soglia che regola il passaggio da un regime all'altro fa sì che il fenomeno è il risultato di un *rimescolamento* continuo di due modelli lineari e proprio da questo rimescolamento scaturisce la non linearità.

Se si introduce il processo stocastico bernoulliano

$$I_{t-d} = \begin{cases} 1 & \text{se } X_{t-d} < v \\ 0 & \text{se } X_{t-d} \geq v \end{cases} \quad (36)$$

e si pone

$$r_t^{(i)} = c_i + \phi_{i1}r_{t-1} + \dots + \phi_{ir}r_{t-r} + \varepsilon_{it} - \theta_{i1}\varepsilon_{t-1} - \dots - \theta_{iq}\varepsilon_{t-q}$$

il TARMA può anche essere scritto in modo più compatto

$$\mathbf{r}_t = I_{t-d} \mathbf{r}_t^{(1)} + (1 - I_{t-d}) \mathbf{r}_t^{(2)} \quad (37)$$

o anche

$$\begin{aligned} I_{t-d}(c_1 + \phi_1(B)\mathbf{r}_t) + (1 - I_{t-d})(c_2 + \phi_2(B)\mathbf{r}_t) = \\ = I_{t-d}\theta_1(B)\boldsymbol{\varepsilon}_t + (1 - I_{t-d})\theta_2(B)\boldsymbol{\varepsilon}_t \end{aligned} \quad (38)$$

dove $\phi_i(B)$ è il polinomio AR di grado r e $\theta_i(B)$ quello MA di grado q del regime $i=1, 2$. Dalla (37) per $I_{t-d} = 1$, oppure per $I_{t-d} = 0$, $t = 1, 2, \dots$, il modello TARMA(2; r, q) si riduce al classico ARMA(r, q). Questo mostra come effettivamente i TARMA sono una generalizzazione degli ARMA al non lineare. Una ulteriore utile generalizzazione dei TARMA si ottiene supponendo che ciascun regime abbia un diverso processo di errore ed in particolare sia

$$\boldsymbol{\varepsilon}_t = I_{t-d}\boldsymbol{\varepsilon}_{1t} + (1 - I_{t-d})\boldsymbol{\varepsilon}_{2t}$$

con $\boldsymbol{\varepsilon}_{1t}$ ed $\boldsymbol{\varepsilon}_{2t}$ processi stocastici iid ed indipendenti fra di loro. Questo fa sì che ogni regime può avere un errore con distribuzione di probabilità differente rendendo ancora più flessibile il modello. Nel seguito si supporrà che le innovazioni $\boldsymbol{\varepsilon}_{1t}$ ed $\boldsymbol{\varepsilon}_{2t}$ siano assolutamente continue ed ammettano almeno

momenti di ordine quattro. Se si tiene conto che $I_{t-d}^2 = I_{t-d}$ e quindi che $I_{t-d}(1 - I_{t-d}) = 0$, si ha anche

$$\begin{aligned} I_{t-d}(c_1 + \phi_1(B)\mathbf{r}_t) + (1 - I_{t-d})(c_2 + \phi_2(B)\mathbf{r}_t) = \\ = I_{t-d}\theta_1(B)\boldsymbol{\varepsilon}_{1t} + (1 - I_{t-d})\theta_2(B)\boldsymbol{\varepsilon}_{2t} \end{aligned} \quad (38b)$$

che mette in evidenza come ciascun regime può avere una innovazione con distribuzione differente.

Si può dimostrare (vedi Niglio e Vitale (2009)) che, se almeno uno dei regimi è stazionario (stazionarietà locale), il TARMA è globalmente stazionario. Possono tuttavia esistere casi in cui il TARMA è globalmente stazionario, ma è localmente non stazionario in tutti i suoi regimi. In altri termini, la regione di stazionarietà di questi modelli è molto più ampia di quella degli ARMA. Questo contribuisce a rendere i TARMA molto flessibili e capaci di rappresentare

fenomeni complessi come lo sono quelli finanziari. D'altra parte, per poter applicare i TARMA è necessario, non solo individuare gli ordini (r, q) di ciascun regime, ma, nota la variabile soglia X_t , derivare (d, v) e quindi stimare i parametri AR ed MA di ciascun regime. In molte applicazioni la variabile soglia è una semplice funzione di r_{t-d} per esempio

$$I_{t-d} = \begin{cases} 1 & \text{se } r_{t-d} - r_{t-d-1} < v \\ 0 & \text{se } r_{t-d} - r_{t-d-1} \geq v \end{cases}$$

In questi casi si dice che la (38) è un SETARMA (Self Exciting TARMA). Nel seguito i modelli che utilizzeremo saranno, per l'appunto, dei SETARMA.

Dalla (37), sotto l'ipotesi di stazionarietà del processo, posto $p = E(I_{t-d})$ ed indicando il momento h -simo del regime i -simo con

$$\mu_h^{(i)} = E\left((r_t^{(i)})^h \mid I_{t-d} = j\right)$$

per $(i=1, j=1)$ nel primo regime e $(i=2, j=0)$ nel secondo regime, si ha

$$\mu_h = E(r_t^h) = p\mu_h^{(1)} + (1-p)\mu_h^{(2)}$$

In altri termini, il momento h -simo globale del processo è la media ponderata dei momenti h -simi locali dei singoli regimi. In particolare, per la media risulta

$$\mu = E(r_t) = p\mu_1^{(1)} + (1-p)\mu_1^{(2)} \quad (39)$$

che può essere nulla anche se non lo è quella di entrambi i regimi. Viceversa, se la media di entrambi i regimi è nulla, lo è anche quella di r_t .

La varianza di globale r_t è data da

$$\begin{aligned} \sigma_x^2 &= \text{Var}(r_t) = \\ &= p\text{Var}(r_t \mid I_{t-d} = 1) + (1-p)\text{Var}(r_t \mid I_{t-d} = 0) + p(1-p)\left(\mu_1^{(1)} - \mu_1^{(2)}\right)^2 \end{aligned} \quad (40)$$

che è una funzione delle varianze e delle medie locali dei due regimi. Se le medie condizionate dei due regimi sono uguali, la varianza di r_t è una media

ponderata della varianza condizionata dei singoli regimi.

Allo stesso modo, per le autocovarianze di r_t (Amendola, Niglio, Vitale, 2006), posto

$$\tau_h = E(I_{t-d} I_{t-d-h})$$

$$\gamma_{ij}(h) = \text{cov}(r_t, r_{t-h} | I_{t-d} = 2-i, I_{t-d-h} = 2-j), \text{ per } i, j = 1, 2$$

risulta

$$\begin{aligned} \gamma(h) = \text{cov}(r_t, r_{t-h}) &= \tau_h \gamma_{11}(h) + (p - \tau_h)(\gamma_{12}(h) + \gamma_{21}(h)) + \\ & (1 - 2p + \tau_h) \gamma_{22}(h) + (\tau_h - p^2) (\mu^{(1)} - \mu^{(2)})^2 \end{aligned} \quad (41)$$

e la funzione di autocovarianza dei rendimenti globali con una struttura TARMA è funzione delle auto e cross covarianze locali dei due regimi oltre che delle medie di questi. Se le medie dei due regimi sono uguali fra di loro, l'autocovarianza dei rendimenti con struttura TARMA è una media ponderata fra le auto e le cross covarianze dei regimi. Per $h=0$ si ha $\tau_0 = p$ e la (41) si riduce alla (40), inoltre, dividendo la (41) per la (40) si ottiene la funzione di autocorrelazione del modello a soglia.

Per l'indice di *asimmetria*, indicando con $\gamma_1^{(1)}$ quella locale riferita al primo regime condizionata a $I_{t-d} = 1$ e con $\gamma_1^{(2)}$ quella riferita al secondo regime condizionata a $I_{t-d} = 0$, si ottiene:

$$\begin{aligned} \gamma_1 &= E \left(\left(\frac{r_t - \mu}{\sigma_x} \right)^3 \right) = p \gamma_1^{(1)} + (1-p) \gamma_1^{(2)} + \\ & + 3 \frac{p \mu_2^{(1)} (\mu_1^{(1)} - \mu) + (1-p) \mu_2^{(2)} (\mu_1^{(2)} - \mu)}{\sigma_x^3} - 2 \frac{p (\mu_1^{(1)})^3 + (1-p) (\mu_1^{(2)})^3 - \mu^3}{\sigma_x^3} \end{aligned}$$

dove può essere $\gamma_1 \neq 0$ anche se gli indici di asimmetria dei singoli regimi sono nulli e viceversa. Nel caso in cui le medie condizionate dei regimi sono le stesse l'asimmetria della serie è una media ponderata degli indici di asimmetria dei singoli regimi. In conclusione, l'asimmetria del processo è fortemente legata alle medie, e quindi ai *drift*, dei singoli regimi.

Allo stesso modo per l'indice di *curtosi globale*, indicando con $\gamma_2^{(i)}$ quella locale riferita al regime i , si ottiene

$$\begin{aligned} \gamma_2 = E \left(\left(\frac{r_t - \mu}{\sigma_x} \right)^4 \right) - 3 &= p\gamma_2^{(1)} + (1-p)\gamma_2^{(2)} + 3 \frac{p(\mu_1^{(1)})^4 + (1-p)(\mu_1^{(2)})^4 - \mu^4}{\sigma_x^4} + \\ &+ 4 \frac{p\mu_3^{(1)}(\mu_1^{(1)} - \mu) + (1-p)\mu_3^{(2)}(\mu_1^{(2)} - \mu)}{\sigma_x^4} - \\ &+ 6 \frac{p\mu_2^{(1)}((\mu_1^{(1)})^2 - \mu^2) + (1-p)\mu_2^{(2)}((\mu_1^{(2)})^2 - \mu^2)}{\sigma_x^4} \end{aligned}$$

ed anche in questo caso la *curtosi globale* può essere nulla senza che lo sia quella locale dei regimi e viceversa ed, inoltre, dipende in modo determinante dalle medie, e quindi dai *drift*, dei singoli regimi. Se le medie condizionate dei singoli regimi sono uguali l'indice di *curtosi* è una media ponderata di quelli dei singoli regimi.

Il modello (38) può essere facilmente generalizzato al caso di $k \geq 2$ regimi anche se, come già detto, nelle applicazioni a fenomeni reali non si va oltre a $k = 3$. Come già accennato, per applicare concretamente il modello TARMA è necessario identificare (d, v, p, q) e quindi stimarne i restanti parametri AR ed MA con il metodo dei minimi quadrati o quello della massima verosimiglianza. Se il modello è ben specificato le stime che si ottengono sono asintoticamente non distorte e consistenti. Di solito l'identificazione del modello avviene utilizzando conoscenze a priori sul fenomeno in studio ed effettuando tentativi di stima in una griglia prefissata di valori. Fra i diversi modelli stimati si sceglie quello che adatta meglio i dati con il minor numero di parametri possibile.

Nel caso dei rendimenti è plausibile porre $v = 0$ il che permette di discriminare fra mercato "*orso*" e mercato "*toro*". Inoltre, per queste serie è

sensato supporre $d = 1$ nell'ipotesi che i rendimenti abbiano memoria breve ed al tempo t siano influenzati, al più, dai rendimenti del tempo precedente $t-1$. Inoltre, come per la modellistica ARMA, anche per i TARMA è giustificato supporre che i parametri p e q siano di basso ordine e di solito risulti $p \leq 1$, $q \leq 5$. Tenendo conto di queste considerazioni, in pratica l'applicazione dei modelli TARMA ai rendimenti di attivi si restringe ad un numero piuttosto limitato di alternative.

Esempio 4

Supponiamo che il processo r_t sia generato dal modello TARMA(2;1;0) = TAR(2; 1):

$$r_t = I_{t-1}(c_1 + \phi_1 r_{t-1} + \varepsilon_{1t}) + (1 - I_{t-1})(c_2 + \phi_2 r_{t-1} + \varepsilon_{2t}) \quad (42)$$

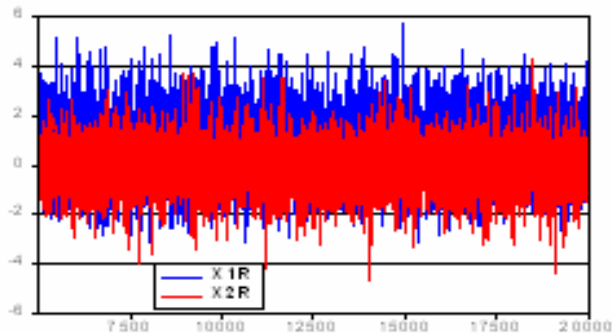
$$I_{t-1} = \begin{cases} 1 & \text{se } r_{t-1} - r_{t-2} \geq 0 \\ 0 & \text{se } r_{t-1} - r_{t-2} < 0 \end{cases}; \quad \varepsilon_{1t} \stackrel{\text{iid}}{\sim} N(0,1); \quad \varepsilon_{2t} \stackrel{\text{iid}}{\sim} 0.5t_{(5)}$$

con

$$c_1 = 0.1; c_2 = -0.1; \phi_1 = 0.7; \phi_2 = -0.5; \sigma_1 = 1; \sigma_2 = .5\sqrt{5/3} = 0.6455$$

Partendo dal valore iniziale $r_0 = 0.1$ è stata simulata una serie di 20000 osservazioni, di queste sono state scartate le prime 5000 per eliminare l'influenza del valore iniziale e le analisi successive sono state effettuate sulle rimanenti $N=15000$. Nella figura 15 è riportato il grafico di r_t e dei suoi due regimi r_{1t} (blue) ed r_{2t} (rosso) e nella 16 la stima delle densità di probabilità.

Figura 15 – Grafico della serie simulata e dei suoi due regimi



Dalle figure 15 e 16 e dalla tabella 24 si può notare come la serie osservata sia molto diversa localmente nei due regimi e questi fra di fra loro. In particolare, mentre i regimi, per il condizionamento rispettivamente a $I_{t-1} = 1$ e $I_{t-1} = 0$, sono lontani dalle due variabili casuali che li hanno generati, r_t ha una asimmetria più elevata di quella di ciascuno dei due regimi e una curtosi intermedia. Nella tabella 25 sono riportate le stime dei parametri del modello che risultano molto vicine ai valori veri e sono altamente significative. La presenza di non linearità in r_t è messa in evidenza dalle figure 18 e 19.

Figura 16 – Densità di probabilità stimata della serie e dei suoi regimi

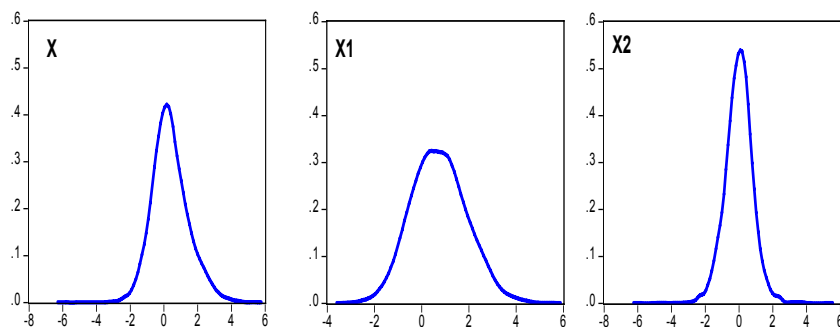


Tabella 24 – Indici descrittivi della serie e dei due regimi

	r_t	r_{1t}	r_{2t}
Media	0.374548	0.736668	-0.007415
Mediana	0.279950	0.688196	0.016755
Max	5.485860	5.485860	5.304914
Min	-6.008774	-3.207231	-6.008774
S_r	1.095812	1.191123	0.828737
$\hat{\gamma}_1$	0.416590	0.197228	-0.102342
$\hat{\gamma}_2$	0.803208	0.041503	2.078515
JB	837.0820	50.47288	1326.811
Prob.	0.000000	0.000000	0.000000
N	15000	7700	7300

Figura 17 – Autocorrelazioni di r_t e dei due regimi

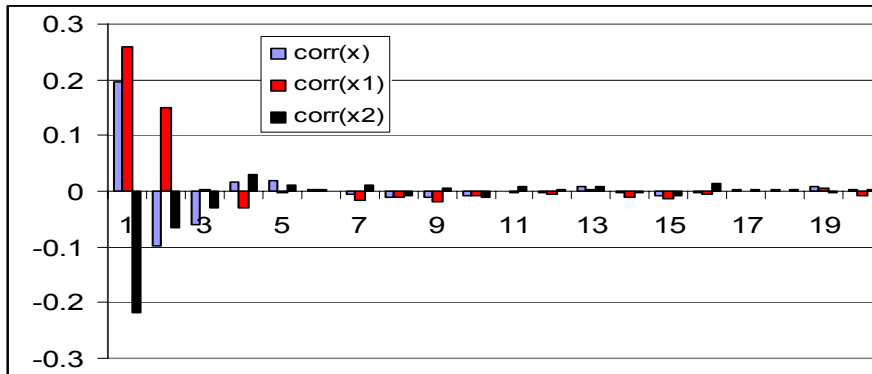


Tabella 25 – Stime LS dei parametri del TARMA

Parametri	Stime	Std. Error	t-Statistic	Prob.
$c_1=0.1$	0.106	0.0137	6.81	0.0000
$\phi_1=0.7$	0.690	0.0108	63.48	0.0000
$c_2=-0.1$	-0.107	0.0102	-10.73	0.0000
$\phi_2=-0.5$	-0.501	0.0102	-50.40	0.0000
$\hat{p} = N_1 / N$	0.513			
$\sigma_1 = 1.0000$	1.009			
$\sigma_2 = 0.6455$	0.651			

Figura 18 – Autocorrelazioni dei quadrati di r_t e dei residui del TAR

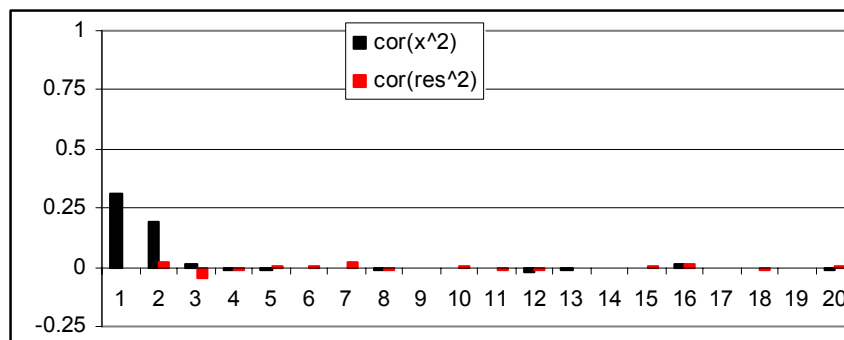
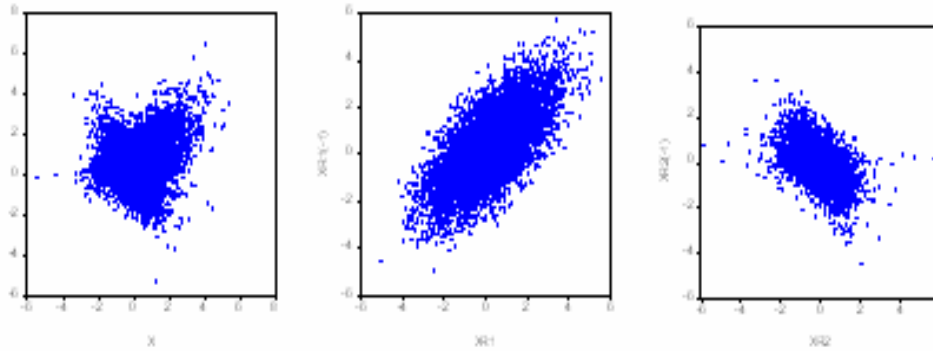


Figura 19 – Scatter di r_t e dei due regimi

In particolare nella figura 18 sono riportate le autocorrelazioni di r_t^2 e quelle dei residui al quadrato del modello. Questa figura mostra la presenza di significativi legami in r_t^2 assenti nei quadrati dei residui e quindi completamente catturati dal modello TAR(2; 1). Nella figura 19 è riportato lo scatter di (r_t, r_{t-1}) , che ha una forma molto lontana da quella ellittica, e quello dei regimi (r_{1t}, r_{1t-1}) e (r_{2t}, r_{2t-1}) che hanno, viceversa, una forma ellittica tipica dei fenomeni lineari o vicini al lineare.

□

Come per la modellistica ARMA un ruolo importante nell'analisi lineare dei rendimenti è giocato dai modelli MA, anche nei TARMA i modelli TMA hanno una certa rilevanza dato che, come vedremo fra breve, posseggono una memoria limitata nel tempo tipica dei rendimenti di attivi finanziari. Qui di seguito presentiamo con un certo dettaglio il caso del TARMA(2; 0, 1) = TMA(2; 1) e quindi con una la componente MA in ciascun regime:

$$r_t = I_{t-1}(c_1 + a_{1t} - \theta_1 a_{1t-1}) + (1 - I_{t-1})(c_2 + a_{2t} - \theta_2 a_{2t-1}) \quad (43)$$

Dalla (43) si possono ricavare gli indici sintetici descrittivi che lo caratterizzano, in particolare

$$\mu = E(r_t) = pc_1 + (1-p)c_2$$

$$\gamma(0) = \text{Var}(r_t) = p\sigma_1^2(1 + \theta_1^2) + (1-p)\sigma_2^2(1 + \theta_2^2) + p(1-p)(c_1 - c_2)^2$$

$$\gamma(1) = (\tau_1 - p^2)(c_1 - c_2)^2 - \left[\tau_1 \theta_1 \sigma_1^2 + (p - \tau_1) \sigma_1 \sigma_2 (\theta_1 + \theta_2) + (1 - 2p + \tau_1) \theta_2 \sigma_2^2 \right]$$

$$\gamma(h) = 0 \text{ per } h = 2, 3, \dots$$

con $\tau_1 = E(I_t I_{t-1})$. Dall'espressione della funzione di *autocovarianza* rileviamo che, come avviene per gli MA(1), anche i TMA(2; 1) hanno un punto di troncamento in $h=1$. La struttura della varianza e dell'autocovarianza, e quindi dell'autocorrelazione, dei TMA è molto più complessa di quella degli MA dato che dipende anche dai livelli medi dei singoli regimi. Per quel che riguarda la stima dei parametri, così come avviene per gli MA, viene ottenuta con una procedura iterativa.

Esempio 5

Supponiamo che r_t sia generato dal TARMA(2; 0, 1) = TMA(2; 1) riportato nella (43) con

$$c_1 = 0.1; c_2 = -0.15; \theta_1 = -0.8; \theta_2 = 0.7; a_{1t} \stackrel{iid}{\sim} N(0,1); a_{2t} \stackrel{iid}{\sim} 0.5t_{(5)}$$

Partendo dal valore iniziale $r_0 = 0.1$ sono state generate 20000 osservazioni e nelle successive analisi ne sono state utilizzate le ultime $N=15000$ per eliminare l'influenza del valore iniziale.

Figura 20 – Serie r_t e suoi regimi

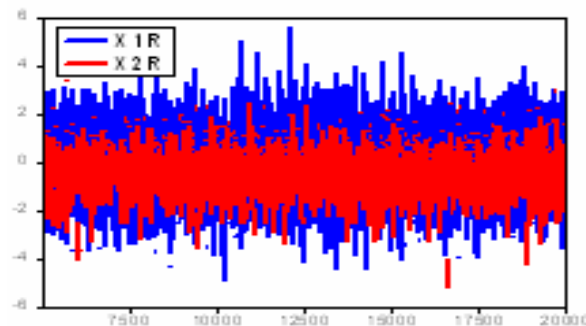
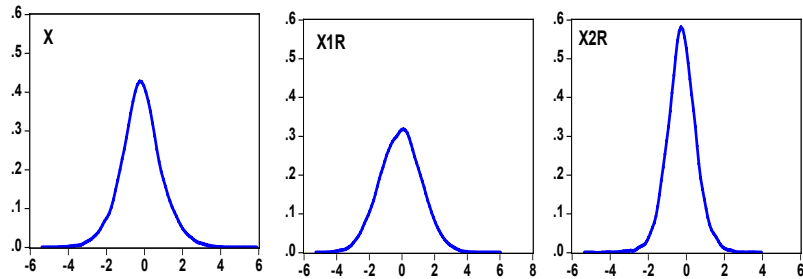


Figura 21 – *Stima delle densità di r_t e dei suoi regimi*

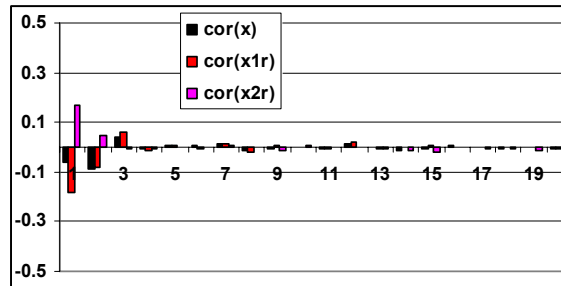
Ricordiamo che, come è per gli MA, anche i TMA sono sempre stazionari. Nella figura 20 riportiamo i grafici dei regimi, mentre nella 21 è riportata la stima della funzione di densità di r_t e dei suoi regimi.

Tabella 26 – *Stime degli indici descrittivi di r_t e dei suoi regimi*

	r_t	r_{1t}	r_{2t}
Media	-0.161663	-0.088670	-0.249283
Mediana	-0.181475	-0.073809	-0.250100
Max	5.620516	5.620516	3.708140
Min	-5.111178	-4.862779	-5.111178
S_r	1.080218	1.272104	0.781747
$\hat{\gamma}_1$	0.085533	0.007784	-0.097405
$\hat{\gamma}_2$	0.775595	0.046986	1.344747
JB	394.2567	0.835350	524.4246
Prob.	0.000000	0.658576	0.000000
N	15000	8183	6817

Nella tabella 26 riportiamo gli indici sintetici descrittivi delle tre serie e, come si può osservare, il primo è compatibile con la normale, mentre non lo sono il secondo e la serie risultante r_t . Nella figura 21 riportiamo le prime 20 autocorrelazioni di r_t e dei suoi regimi e notiamo che per tutte e tre le serie solo le prime due autocorrelazioni sono sostanzialmente diverse da zero. Nella figura 22 riportiamo i diagramma a scatter delle tre serie e solo quello per r_t presenta una evidente non linearità.

Figura 21 – Stime delle autocorrelazioni di r_t e dei suoi regimi



Nella tabella 27 sono riportate le stime LS iterative dei parametri del modello TMA(2;1) e queste sono molto vicine ai valori veri ed altamente significative. Nella figura 23 sono riportate le autocorrelazioni del quadrato di r_t che mostrano un valore significativamente diverso da zero al ritardo 1. Nello stesso grafico sono riportate le autocorrelazioni dei residui del modello stimato e quelle dei loro quadrati e nessuna di queste autocorrelazioni è significativamente diversa da zero. Questo vuol dire che il modello TMA(2;1) ha effettivamente catturato la non linearità presente in r_t .

Figura 22 – Scatter di r_t e dei regimi

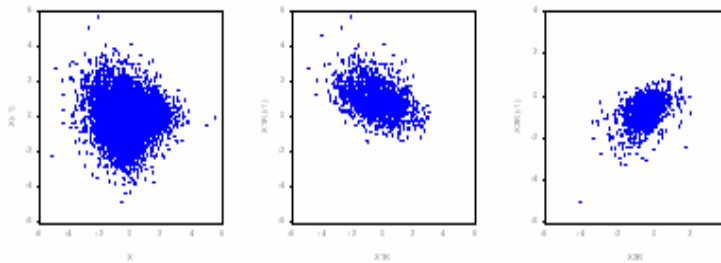


Figura 23 – Autocorrelazioni di: r_t al quadrato, del residuo stimato, del residuo stimato al quadrato

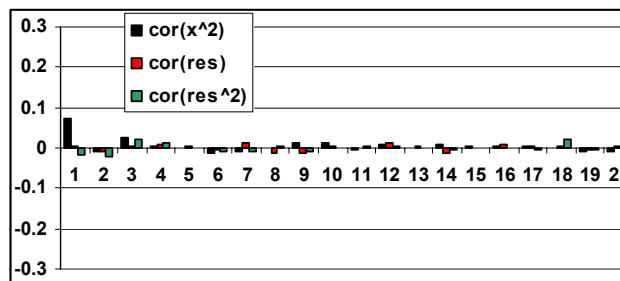


Tabella 27 – Stime dei parametri del TMA(2; 1)

Parametri	Stime	Std. Error	t-Statistic	Prob.
$c_1=0.1$	0.111	0.0097	11.44	0.0000
$\theta_1=-0.8$	-0.804	0.0097	-82.82	0.0000
$c_2=-0.15$	-0.148	0.0106	-14.02	0.0000
$\theta_2=0.7$	0.713	0.0159	44.92	0.0000
$\hat{p} = N_1 / N$	0.546			
$\sigma_1 = 1.0000$	1.003			
$\sigma_2 = 0.6455$	0.6287			

□

4.9 La scelta del modello ottimale

Nelle pratiche applicazioni, alla stessa serie osservata possono essere adattati differenti modelli sia della classe ARMA che di quella TARMA. Sorge quindi il problema di scegliere quello ottimale utilizzando un qualche criterio informativo. Di questi indici ne sono stati proposti in letteratura più di uno, qui di seguito presentiamo quelli più noti ed usati: R^2 , AIC (*Akaike Information Criterion*) e SIC (*Schwartz Information Criterion*). Scelto un criterio informativo, si procede al suo calcolo per ogni modello stimato. Si sceglie il modello con R^2 massimo ed AIC e SIC minimi.

L'indice R^2 è dato da

$$R^2 = \frac{\text{Var}(\hat{X}_t)}{\text{Var}(X_t)} = 1 - \frac{\hat{\sigma}^2}{\text{Var}(X_t)}$$

dove \hat{X}_t rappresenta i valori di X_t quando ai suoi parametri incogniti si sostituiscono le rispettive stime. Gli indici AIC e SIC sono costruiti come somma di due espressioni: la prima è una funzione monotona della varianza stimata del residuo del modello e l'altra una funzione del numero dei parametri stimati. Fissato N , la prima decresce con la varianza stimata S_a^2 del residuo a_t mentre la seconda tende a crescere con il numero M dei parametri stimati. Come intuibile, più la varianza del residuo è piccola più il modello adatta bene i

dati, d'altro lato al crescere del numero dei parametri questa varianza tende comunque a diminuire anche quando quei parametri sono non significativi. Di qui l'esigenza di costruire indici globali che tengano conto delle due opposte esigenze:

ottenere l'adattamento il più elevato possibile con il minor numero di parametri.

Più precisamente AIC è dato da

$$AIC = \log(S_a^2) + \frac{2M}{N} \quad (44)$$

mentre SIC è pari a

$$SIC = \log(S_a^2) + \frac{M \log(N)}{N} \quad (45)$$

dove S_a^2 è la varianza stimata del residuo del modello, M il numero complessivo di parametri stimati ed N il numero di osservazioni utilizzate nella stima. C'è da osservare che non sempre il modello ottimale individuato con AIC è lo stesso di quello di quello individuato da SIC. In questi casi si utilizza quel modello che è più semplice interpretare ed utilizzare.

Esempio 6

Alla serie storica generata dal TAR(2; 1) definita nella (42) dell'esempio 4 adattiamo i seguenti modelli stimandone i parametri con il metodo dei minimi quadrati iterativo: AR(1); AR(2); MA(1); MA(2); ARMA(1;1); TAR(2; 2); TARMA(2; 1, 1); TAR(2; 1). I risultati della stima sono riportati nella tabella 28 assieme agli indici AIC, SIC ed R^2 .

Dalla tabella 28 si osserva che il minimo per gli indici AIC e SIC ed il massimo di accostamento R^2 si ottengono con il modello TARMA(2; 1, 1). Ma in questo modello le stime di $\theta_1^{(1)}$ e $\theta_1^{(2)}$ risultano non significativamente diverse da zero essendo entrambe inferiori a ± 2 volte gli standard error riportati al di sotto di ogni stima. Lo stesso avviene per il TAR(2; 2) in cui i parametri $\phi_2^{(1)}$ e $\phi_2^{(2)}$ non sono significativamente diversi da zero. Questo vuol dire che quei parametri devono essere eliminati dal modello TARMA(2; 1, 1) ed ottenere correttamente il TMA(2; 1) che ha generato i dati.

Tabella 28 – Stime dei parametri dei modelli indicati (in grassetto quelli veri)

Parametri	c_1	c_2	$\phi_1^{(1)}$	$\phi_1^{(2)}$	$\phi_2^{(1)}$	$\phi_2^{(2)}$	$\theta_1^{(1)}$	$\theta_1^{(2)}$	AIC	SIC	R^2
Valori veri	0.1	-0.1	0.7	0.5	0	0	0	0			
AR(1)	.284 (.009)		.197 (.008)						2.93	2.93	0.039
AR(2)	.324 (.009)		.225 (.008)	-.143 (.008)					2.91	2.91	0.058
MA(1)	.354 (.010)						.246 (.008)		2.92	2.92	0.049
MA(2)	.354 (.010)						.218 (.008)	0.073 (.008)	2.91	2.91	0.054
ARMA(1; 1)	.354 (.010)		-.182 (.032)				.414 (.010)		2.91	2.91	0.052
TAR(2; 2)	.113 (.016)	-.108 (.016)	.663 (.014)	-.506 (.010)	.010 (.013)	.004 (.013)			2.51	2.51	0.366
TARMA(2;1,1)	.108 (.014)	-.109 (.010)	.708 (.012)	-.506 (.010)			.006 (.012)	.000 (.008)	2.51	2.50	0.403
TAR(2; 1)	.106 (.014)	-.104 (.010)	.670 (.011)	-.507 (.010)					2.51	2.51	0.366

□

4.10 La previsione dei rendimenti

La scelta e stima di un modello ottimale per i rendimenti è utile per analizzare ed interpretare la dinamica del fenomeno, ma è indispensabile per derivarne anche le *previsioni ottimali*. Supponiamo di disporre della serie dei rendimenti (r_1, r_2, \dots, r_N) a cui sia stato adattato il modello ottimale della classe TARMA. Questo modello può essere utilizzato per derivare le k previsioni $(r_{N+1}, r_{N+2}, \dots, r_{N+k})$. Nel seguito per indicare la previsione di r_{N+j} , note le informazioni (r_1, r_2, \dots, r_N) , si utilizzerà la notazione $r_N(j)$, $j = 1, 2, \dots, k$. Queste previsioni vengono anche dette *ex-ante (forecasted)* perché si prevede r_{N+j} date le informazioni disponibili fino al tempo $t=N$ *prima* che r_{N+j} si sia verificato, altrimenti si tratta di valori interpolati o *ex-post (fitted)*. In generale, si può dimostrare che la previsione lineare ottimale è data dalla media condizionata:

$$r_N(j) = E\left(r_{N+j} \mid r_N, r_{N-1}, \dots, r_1\right)$$

Nel seguito analizzeremo le previsioni da modelli TARMA che hanno una qualche possibilità pratica di presentarsi. Naturalmente, se è $r_t \stackrel{\text{iid}}{\sim} \text{WN}(0, \sigma^2)$ allora è sempre $r_N(j) = 0$ per $j = 1, 2, \dots$. In altri termini, se i rendimenti si comportano come un white noise il previsore ottimale è pari a zero, qualsiasi sia l'orizzonte previsivo j .

(A) PREVISIONI DA UN AR(1)

Supponiamo che i rendimenti abbiano una struttura AR(1) e quindi sia

$$r_t = \phi_0 + \phi r_{t-1} + \varepsilon_t; \quad \varepsilon_t \sim \text{WN}(0; \sigma^2)$$

Siano $\mathfrak{I}_N = (r_1, r_2, \dots, r_N)$ le informazioni disponibili al tempo N e quindi r_N è noto. Si ottiene

$$\begin{aligned} r_N(1) &= E\left(r_{N+1} \mid \mathfrak{I}_N\right) = E\left(\phi_0 + \phi r_N + \varepsilon_{N+1} \mid \mathfrak{I}_N\right) = \\ &= \phi_0 + \phi r_N + E\left(\varepsilon_{N+1} \mid \mathfrak{I}_N\right) = \phi_0 + \phi r_N \end{aligned}$$

Allo stesso modo risulta

$$r_N(2) = E\left(r_{N+2} \mid \mathfrak{I}_N\right) = E\left(\phi_0 + \phi r_{N+1} + \varepsilon_{N+2} \mid \mathfrak{I}_N\right) = \phi_0 + \phi r_N(1).$$

Con le stesse argomentazioni in generale si ha

$$r_N(k) = \phi_0 + \phi r_N(k-1), \quad k = 1, 2, \dots, \quad \text{con } r_N(0) = r_N \quad (46)$$

La (46) con sostituzioni successive si può anche scrivere

$$r_N(k) = \phi_0 \frac{1 - \phi^{k+1}}{1 - \phi} + \phi^k r_N \quad (47)$$

ed al crescere di k , dato che il processo dei rendimenti degli attivi è sempre

stazionario e quindi risulta $|\phi| < 1$, si ha

$$r_N(k) \xrightarrow{k} \frac{\phi_0}{1-\phi} = E(r_t)$$

In altri termini, al crescere dell'orizzonte temporale k , la previsione tende alla media dei rendimenti. In pratica i due parametri (ϕ_0, ϕ) sono incogniti e per ottenere una previsione stimata si sostituiscono nella (47) ai valori veri le rispettive stime (principio del plug-in):

$$\hat{r}_N(k) = \hat{\phi}_0 \frac{1-\hat{\phi}^{k+1}}{1-\hat{\phi}} + \hat{\phi}^k r_N$$

Del previsore $r_N(k)$ è possibile derivare l'*errore di previsione*

$$e_N(k) = r_{N+k} - r_N(k)$$

Ricordando la (17) avremo

$$r_{N+k} = \frac{\phi_0}{1-\phi} + \sum_{j=0}^{\infty} \phi^j \varepsilon_{N+k-j}$$

mentre

$$r_N(k) = E(r_{N+k} | \mathfrak{F}_N) = E\left(\frac{\phi_0}{1-\phi} + \sum_{j=0}^{\infty} \phi^j \varepsilon_{N+k-j} \mid \mathfrak{F}_N\right) = \frac{\phi_0}{1-\phi} + \sum_{j=k}^{\infty} \phi^j \varepsilon_{N+k-j}$$

dove si è tenuto conto del fatto che, per $j < k$, risulta

$$E(\varepsilon_{N+k-j} | \mathfrak{F}_N) = 0$$

Avremo così

$$e_N(k) = \sum_{j=0}^{k-1} \phi^j \varepsilon_{N+k-j}$$

da cui si ottiene

$$E(e_N(k)) = E\left(\sum_{j=0}^{k-1} \phi^j \varepsilon_{N+k-j}\right) = 0$$

$$\text{Var}(e_N(k)) = \sigma_p^2(k) = \text{Var}\left(\sum_{j=0}^{k-1} \phi^j \varepsilon_{N+k-j}\right) = \sigma^2 \sum_{j=0}^{k-1} \phi^{2j} = \sigma^2 \frac{1-\phi^{2k}}{1-\phi^2}$$

Questo vuol dire che in media l'errore di previsione è nullo e quindi il previsore, calcolato con la (47), è non distorto. D'altra parte la varianza dell'errore di previsione è sempre minore della varianza di r_t e tende a questa al crescere di k . Inoltre si può dimostrare che, fra tutti i previsori non distorti, quello basato sulla media condizionata è caratterizzato dalla minima varianza dell'errore di previsione.

Sotto l'ipotesi di normalità dei residui, l'errore di previsione è una combinazione lineare di normali indipendenti e quindi è esso stesso normale con media nulla e varianza $\sigma_p^2(k)$. Questo risultato permette di calcolare l'intervallo di previsione approssimato al livello $1-\alpha$:

$$\hat{r}_N(k) - z_{(\alpha/2)} \hat{\sigma}_p(k) \leq r_{N+k} \leq \hat{r}_N(k) + z_{(\alpha/2)} \hat{\sigma}_p(k)$$

dove $z_{(\alpha/2)}$ è, al solito, il percentile della normale standardizzata che stacca alla sua destra $\alpha/2$ di probabilità mentre è

$$\hat{\sigma}_p^2(k) = \hat{\sigma}^2 \frac{1-\hat{\phi}^{2k}}{1-\hat{\phi}^2}$$

Per il teorema del limite centrale questi risultati continuano a valere in modo approssimato anche quando non è vera l'ipotesi di normalità dei residui.

Esempio 7

Alle serie dei rendimenti delle quotazioni Campari, Fiat, Mibtel, Nasdaq, S&P 500 adattiamo un AR(1), ne analizziamo i residui stimati ed effettuiamo $k = 10$ previsioni ex-ante. I risultati della stima, riferiti al periodo 02/01/03 – 17/02/09 sono riportati nella tabella 29. Come si può notare la stima della costante ϕ_0 è sempre non significativamente diversa da zero e quindi deve essere eliminata dal modello, mentre quella del parametro autoregressivo ϕ è sempre molto piccola ed è non significativamente diversa da zero per i rendimenti Mibtel.

Tabella 29 – *Stime dei parametri (ϕ_0, ϕ) ed indici di accostamento dell'AR(1) per i rendimenti giornalieri dei cinque attivi. In parentesi gli s.e. delle stime*

Serie	ϕ_0	ϕ	R^2	AIC	SIC
Campari	0.0002 (.0004)	-0.0910 (.0249)	0.008	-5.351	-5.345
Fiat	-0.0005 (.0006)	0.0426 (.0205)	0.002	-4.636	-4.630
Mibtel	-0.0002 (.0003)	-0.0132 (.0249)	0.002	-6.069	-6.063
Nasdaq	0.0000 (.0003)	-0.1168 (.0250)	0.014	-5.527	-5.521
S&P	0.0001 (.0003)	-0.1406 (.0247)	0.020	-5.843	-5.841

Tabella 30 – *Stime finali. In parentesi gli standard errors delle stime*

Serie	ϕ	R^2	AIC	SIC
Campari	-0.090 (.025)	0.008	-5.353	-5.349
Fiat	0.043 (.025)	0.001	-4.637	-4.634
Mibtel	0.001 (.025)	0.000	-6.070	-6.067
Nasdaq	-0.117 (.025)	0.014	-5.529	-5.526
S&P	-0.141 (.025)	0.020	-5.844	-5.841

Riadattiamo il modello AR(1) alle 5 serie sopra elencate dopo aver eliminato i parametri non significativamente diversi da zero e utilizziamo i modelli finali stimati per effettuare le previsioni fino a $k = 10$ giorni in avanti. I risultati della stima sono riportati nella tabella 30.

Dalla tabella 30 si nota che il parametro autoregressivo ϕ non è significativamente diverso da zero per i rendimenti Fiat e Mibtel che quindi si comportano, sostanzialmente, come dei white noise. Nella figura 24 sono riportate le prime 12 autocorrelazioni dei residui del modello, per Campari, Nasdaq e S&P 500, e le autocorrelazioni di Fiat e Mibtel e, come si può osservare, alcune autocorrelazioni fuoriescono dalla banda di confidenza e quindi il modello deve essere affinato.

Figura 24 – Autocorrelazione dei residui dell'AR(1) adattato, nell'ordine, ai rendimenti di Campari, Fiat, Nasdaq, Mibtel e S&P 500

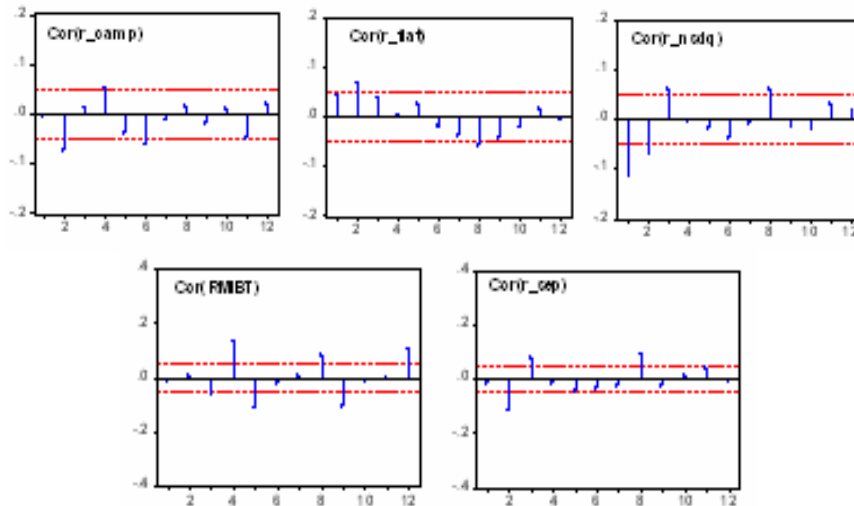
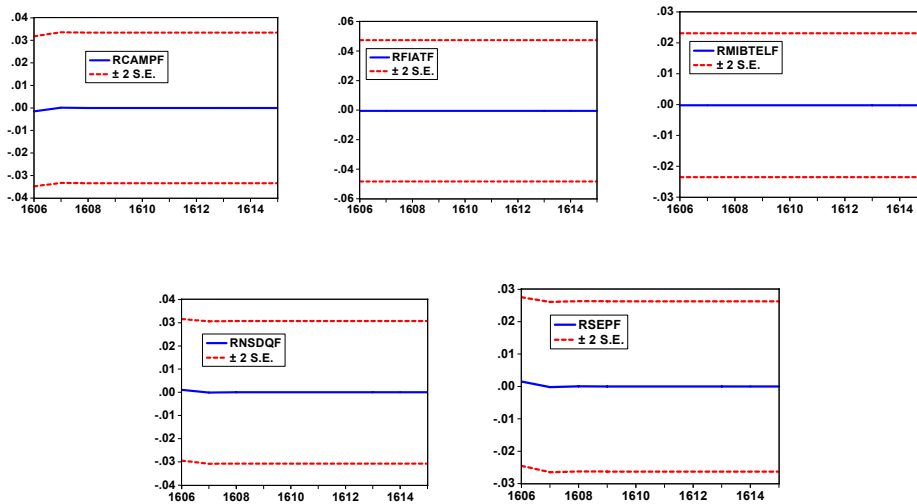


Figura 25 – Previsioni ed intervalli di confidenza delle previsioni dei rendimenti delle cinque serie considerate



Come atteso la previsione dei rendimenti per $k = 1, 2, \dots, 10$ giorni delle cinque serie analizzate è sostanzialmente uguale alla propria media che, essendo $\phi_0 = 0$ in ciascuna di loro, è nulla.

□

(B) PREVISIONI DA UN MA(1)

Supponiamo che il processo generatore dei rendimenti sia un MA(1):

$$r_t = \phi_0 + \varepsilon_t - \theta\varepsilon_{t-1}, \quad \varepsilon_t \sim \text{WN}(0; \sigma^2)$$

risulta

$$r_N(1) = E(r_{N+1} | \mathfrak{F}_N) = \phi_0 + E(\varepsilon_{N+1} | \mathfrak{F}_N) - \theta E(\varepsilon_N | \mathfrak{F}_N) = \phi_0 - \theta\varepsilon_N$$

$$r_N(k) = E(r_{N+k} | \mathfrak{F}_N) = \phi_0 + E(\varepsilon_{N+k} | \mathfrak{F}_N) - \theta E(\varepsilon_{N+k-1} | \mathfrak{F}_N) = \phi_0 \quad \text{per } k > 1.$$

Come si può notare, le previsioni sono costanti e pari alla media ϕ_0 per $k > 1$. L'errore di previsione, in questo caso, è

$$e_N(k) = \begin{cases} \varepsilon_{N+1} & k = 1 \\ \varepsilon_{N+k} - \theta\varepsilon_{N+k-1} & k > 1 \end{cases}$$

e la varianza dell'errore di previsione diviene

$$\sigma_p^2(k) = \text{Var}(e_N(k)) = \begin{cases} \sigma^2 & k = 1 \\ \sigma^2(1 + \theta^2) & k > 1 \end{cases}$$

Le previsioni stimate sono date da

$$\hat{r}_N(k) = \begin{cases} \hat{\phi}_0 - \hat{\theta}\hat{\varepsilon}_N & k = 1 \\ \hat{\phi}_0 & k > 1 \end{cases}$$

con l'intervallo di previsione, al livello $1 - \alpha$, pari a

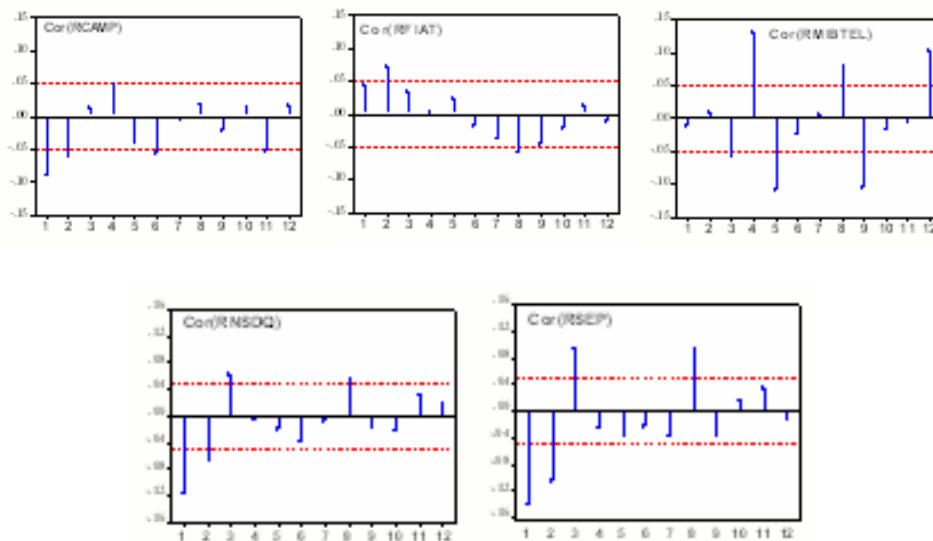
$$\hat{r}_N(k) - z_{\alpha/2} \hat{\sigma}_p(k) \leq r_{N+k} \leq \hat{r}_N(k) + z_{\alpha/2} \hat{\sigma}_p(k)$$

Questi risultati riferiti al modello MA(1) possono essere agevolmente estesi al generico MA(q) ricordando che per i rendimenti di attivi finanziari, di solito è $q \leq 5$.

Esempio 8

Alle serie dei rendimenti dell'esempio 7 applichiamo un modello della classe MA(q) tenendo conto che, se un fenomeno è un MA(q), la sua autocorrelazione ha un punto di troncamento proprio in q. Utilizziamo le stime ottenute per prevedere i rendimenti $k = 10$ giorni in avanti. A ciascuna delle cinque serie è stato adattato un MA(5). Per ciascuno dei modelli stimati sono stati individuati i parametri significativamente diversi da zero ed è stata effettuata una nuova stima con i soli parametri MA significativi. I risultati della stima sono riportati nella tabella 31 ove, al di sotto di ciascuna stima, è indicato il corrispondente errore standard. Come si può osservare dalla tabella, la stima del *drift* è risultata sempre non significativamente diversa da zero. Questo vuol dire che in media i rendimenti di quelle serie, nel periodo analizzato, non sono significativamente diversi da zero. Inoltre, la memoria dei rendimenti giornalieri non va oltre una settimana borsistica (5 giorni). In ogni caso i valori stimati sono tutti molto piccoli come pure la percentuale di variabilità spiegata dal modello e misurata dall'indice R^2 .

Figura 26 – Autocorrelazioni dei rendimenti di Campari, Fiat, Mibtel, Nasdaq, S&P 500 con rispettiva banda di confidenza al 95%



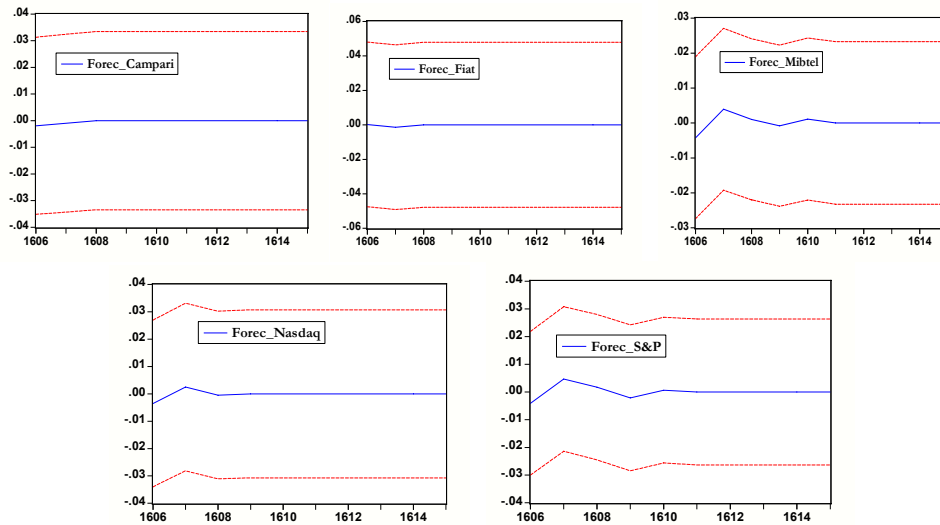
Se confrontiamo gli indici AIC e SIC della tabella 31 con quelli della tabella 30 in cui ai dati era stato adattato un AR(1), notiamo che quelli della tabella 31 sono marginalmente migliori. Questo conferma come effettivamente, se nei rendimenti vi è una qualche struttura, questa è, di solito, di tipo MA(q). Nella

figura 27 sono riportate $k=10$ previsioni stimate che non risultano molto diverse da quelle della figura 25 ottenute utilizzando un AR(1).

Tabella 31 – Stime dei parametri dei modelli $MA(q)$ per ciascuna delle cinque serie

	θ_1	θ_2	θ_3	θ_4	θ_5	R^2	AIC	SIC
Campari	0.096 (.025)	0.058 (.025)				0.013	-5.36	-5.35
Fiat		-0.074 (.025)				0.005	-4.63	-4.63
Mibtel			0.054 (.025)	-0.105 (.025)	0.092 (.025)	0.026	-6.09	-6.08
Nasdaq	0.121 (.025)	0.065 (.025)	-0.059 (.025)			0.022	-5.54	-5.53
S&P	0.149 (.025)	0.103 (.025)	-0.096 (.025)		0.069 (.025)	0.042	-5.87	-5.85

Figura 27 – Previsioni $k = 10$ tempi in avanti delle cinque serie analizzate con il modello $MA(q)$ della tabella 31



□

(C) PREVISIONI DA UN ARMA(1, 1)

Supponiamo che i rendimenti di un dato attivo siano generati da un ARMA(1,1):

$$r_t = \phi_0 + \phi r_{t-1} + \varepsilon_t - \theta \varepsilon_{t-1}$$

con $\phi \neq \theta$. E' abbastanza agevole verificare che

$$r_N(1) = E(r_{N+1} | \mathfrak{F}_N) = E(\phi_0 + \phi r_N + \varepsilon_{N+1} - \theta \varepsilon_N | \mathfrak{F}_N) = \phi_0 + \phi r_N - \theta \varepsilon_N$$

$$r_N(k) = \phi_0 + \phi r_N(k-1), \quad \text{per } k > 1$$

Come si può notare, le previsioni di un ARMA(1,1), dopo il primo tempo $k=1$, si comportano come quelle di un AR(1). Per sostituzioni successive si ottiene

$$r_{N+k} = \frac{\phi_0}{1-\phi} + \varepsilon_{N+k} + \sum_{j=1}^{\infty} (\phi - \theta) \phi^{j-1} \varepsilon_{N+k-j}$$

$$r_N(k) = E \left(\frac{\phi_0}{1-\phi} + \varepsilon_{N+k} + \sum_{j=1}^{\infty} (\phi - \theta) \phi^{j-1} \varepsilon_{N+k-j} \right) =$$

$$= \frac{\phi_0}{1-\phi} + \sum_{j=k}^{\infty} (\phi - \theta) \phi^{j-1} \varepsilon_{N+k-j}$$

che permettono di derivare l'errore di previsione e la sua varianza:

$$e_N(k) = r_{N+k} - r_N(k) = \varepsilon_{N+k} + \sum_{j=1}^{k-1} (\phi - \theta) \phi^{j-1} \varepsilon_{N+k-j}$$

$$\sigma_p^2(k) = \text{Var}(e_N(k)) = \sigma^2 + \sigma^2 (\phi - \theta)^2 \frac{1 - \phi^{2k}}{1 - \phi^2}$$

ed il relativo *intervallo di previsione* al livello $1 - \alpha$

$$\hat{r}_N(k) - z_{\alpha/2} \hat{\sigma}_p(k) \leq r_{N+k} \leq \hat{r}_N(k) + z_{\alpha/2} \hat{\sigma}_p(k)$$

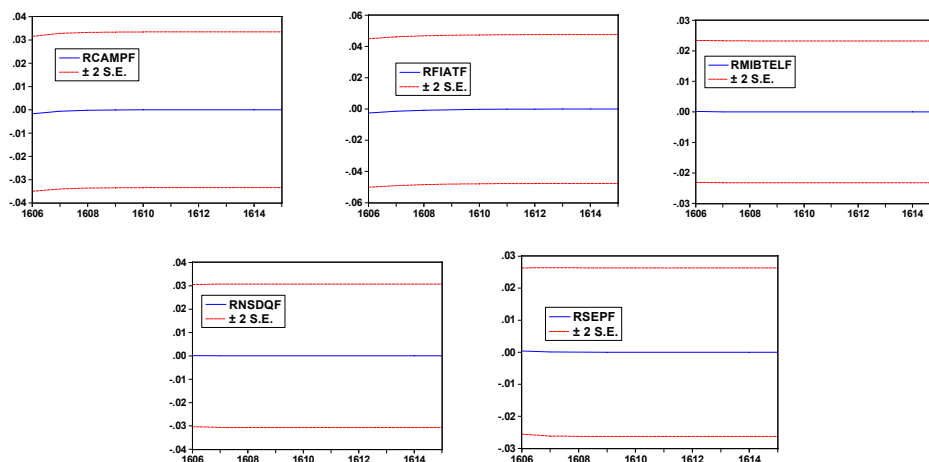
Esempio 9

Alle cinque serie analizzate nell'esempio 8 adattiamo un modello ARMA(1, 1) e deriviamo le relative previsioni per i $k = 10$ tempi successivi. Le stime dei parametri del modello ARMA(1, 1), con $\phi_0 = 0$, sono riportate nella tabella 32.

Tabella 32 – Stime dei parametri del modello $ARMA(1, 1)$ per le cinque serie

	ϕ	θ	R^2	AIC	SIC
Campari	0.354 (.192)	0.456 (.182)	0.012	-5.355	-5.348
Fiat	0.568 (.159)	0.513 (.166)	0.006	-4.640	-4.633
Mibtel	0.189 (.336)	0.202 (.336)	0.000	-6.069	-6.063
Nasdaq	0.270 (.463)	0.399 (.155)	0.018	-5.532	-5.526
S&P	0.293 (.126)	0.458 (.117)	0.029	-5.852	-5.846

Analizzando i risultati della tabella 32 emerge che il modello $ARMA(1, 1)$ non è adeguato per nessuna delle cinque serie. Infatti, per la serie dei rendimenti Campari la stima del parametro ϕ non è significativamente diversa da zero, per i rendimenti Mibtel entrambi i parametri (ϕ, θ) non sono significativamente diversi da zero, per la serie dei rendimenti Nasdaq il parametro ϕ non è significativamente diverso da zero, per la serie dei rendimenti S&P 500 l'indice R^2 dell' $ARMA(1, 1)$ è nettamente più piccolo di quello ottenuto con il modello $MA(5)$ riportato nella tabella 31. Infine, per i rendimenti Fiat le stime dei due parametri ϕ e θ sono entrambe significativamente diverse da zero, ma praticamente uguali fra di loro e quindi cancellabili.

Figura 28 – Previsioni $k = 10$ tempi in avanti delle cinque serie analizzate con il modello $ARMA(1, 1)$ 

Confrontando i risultati delle previsioni ottenute con il modello ARMA(1,1) con quelli dell'MA(q) riportate nella figura 27 si vede come fra di loro vi siano delle contenute differenze ed in particolare per i rendimenti di Mibtel, Nasdaq e S&P 500.

□

(D) PREVISIONI DA UN TAR(2; 1)

Supponiamo che i rendimenti di un attivo siano generati dal modello TAR(2;1):

$$r_t = I_{t-1}(c_1 + \phi_1 r_{t-1} + \varepsilon_{1t}) + (1 - I_{t-1})(c_2 + \phi_2 r_{t-1} + \varepsilon_{2t})$$

Date le informazioni $\mathfrak{F}_N = (r_1, r_2, \dots, r_N)$, è immediato derivare la previsione di r_{N+1} :

$$r_N(1) = E(r_{N+1} | \mathfrak{F}_N) = I_N(c_1 + \phi_1 r_N) + (1 - I_N)(c_2 + \phi_2 r_N) =$$

$$= [I_N c_1 + (1 - I_N) c_2] + [I_N \phi_1 + (1 - I_N) \phi_2] r_N$$

$$= \begin{cases} I_N(c_1 + \phi_1 r_N) & \text{se } r_N \geq 0 \\ (1 - I_N)(c_2 + \phi_2 r_N) & \text{se } r_N < 0 \end{cases}$$

Per le previsioni successive è necessario effettuare alcune assunzioni per poterle praticamente derivare, in particolare assumiamo che per $k > 1$ sia:

- $E(I_{N+k-1} | \mathfrak{F}_N) = p$
- $E(I_{N+k-1} \varepsilon_{jN+k} | \mathfrak{F}_N) = 0, j=1, 2;$

in tal modo risulta

$$r_N(k) = E(r_{N+k} | \mathfrak{F}_N) =$$

$$= c_1 E(I_{N+k-1} | \mathfrak{F}_N) + \phi_1 E(I_{N+k-1} r_{N+k-1} | \mathfrak{F}_N) + E(I_{N+k-1} \varepsilon_{1N+k} | \mathfrak{F}_N) +$$

$$\begin{aligned}
& + c_2 E\left((1 - I_{N+k-1}) \mid \mathfrak{F}_N\right) + \phi_2 E\left((1 - I_{N+k-1}) r_{N+k-1} \mid \mathfrak{F}_N\right) + \\
& + E\left((1 - I_{N+k-1}) \varepsilon_{2N+k} \mid \mathfrak{F}_N\right) = (p c_1 + (1-p)c_2) + (p\phi_1 + (1-p)\phi_2) r_N (k-1), \\
& \qquad \qquad \qquad k=2, 3, \dots
\end{aligned}$$

e le previsioni del TAR(2;1) si comportano, per $k > 1$, come quelle di un AR(1) con

$$c = \phi_0 = (p c_1 + (1-p)c_2); \quad \phi = (p\phi_1 + (1-p)\phi_2); \quad \varepsilon_t = (p\varepsilon_{1t} + (1-p)\varepsilon_{2t})$$

Per quanto riguarda l'errore di previsione si ha

$$e_N(1) = (I_N \varepsilon_{1t} + (1 - I_N) \varepsilon_{2t})$$

$$e_N(k) = \sum_{j=0}^{k-1} \phi^j \varepsilon_{N+k-j} =$$

$$= \sum_{j=0}^{k-1} (p\phi_1 + (1-p)) \left(I_{N+k-1-j} \varepsilon_{1N+k-j} (1 - I_{N+k-1-j}) \varepsilon_{2N+k-j} \right), \quad k=2, 3, \dots$$

che permette di calcolare la varianza dell'errore di previsione

$$\sigma_p^2(k) = \begin{cases} I_N \sigma_1^2 + (1 - I_N) \sigma_2^2 & \text{se } k = 1 \\ \left(p\sigma_1^2 + (1-p)\sigma_2^2 \right) \frac{1 - (p\phi_1 - (1-p)\phi_2)^{2k}}{1 - (p\phi_1 - (1-p)\phi_2)^2} & \text{se } k = 2, 3, \dots \end{cases}$$

Questa si riduce a quella dell'AR(1) se è $p=0$ oppure $p=1$. Le previsioni, gli errori e le relative varianze di previsione sono state derivate supponendo di conoscere i parametri incogniti. Nelle pratiche applicazioni questi parametri vengono sostituiti con loro stime ottenendo le previsioni stimate e le relative varianze di previsione stimate. Una volta calcolata la varianza di previsione è possibile derivare il relativo *intervallo di previsione* al livello $1 - \alpha$:

$$\hat{r}_N(k) - z_{(\alpha/2)} \hat{\sigma}_p(k) \leq r_{N+k} \leq \hat{r}_N(k) + z_{(\alpha/2)} \hat{\sigma}_p(k)$$

dove, ricordiamo, di solito è $\alpha = 5\%$ e risulta $z_{(0.05/2)} \approx 2$.

Esempio 10

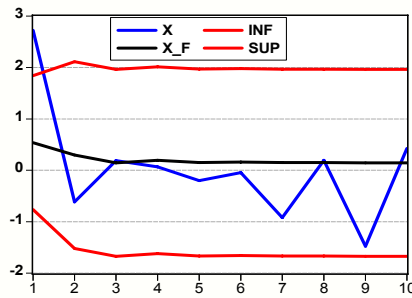
Utilizziamo questi risultati per prevedere le ultime $k=10$ osservazioni del TAR(2;1) riportato nell'esempio 4. In questo caso si ha $N=14990$

$$I_N = 0; \quad r_N = 0.62743; \quad \hat{p} = 0.513; \quad \hat{c}_1 = 0.106; \quad \hat{c}_2 = -0.107$$

$$\hat{\phi}_1 = 0.69; \quad \hat{\phi}_2 = -0.501; \quad \hat{\sigma}_1^2 = 1.1881; \quad \hat{\sigma}_2^2 = 0.4238$$

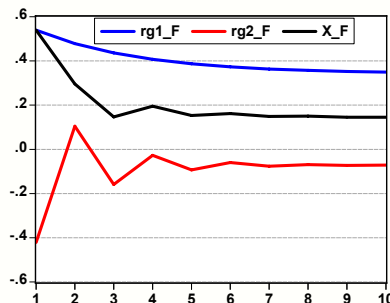
Utilizzando queste stime e le formule sopra riportate abbiamo calcolato le previsioni e le relative bande di previsione e le abbiamo riportate nella figura 29.

Figura 29 – Serie osservata X , prevista X_F e relative bande di confidenza al 95%



Nella figura 30 riportiamo le previsioni globali della serie e quelle locali riferite ai due regimi.

Figura 30 – Previsioni di X e quelle dei suoi due regimi



□

(E) PREVISIONI DA UN TMA(2; 1)

Supponiamo che i rendimenti di un dato attivo siano rappresentabili con il TMA(2;1):

$$r_t = I_{t-1}(c_1 + \varepsilon_t - \theta_1 \varepsilon_{t-1}) + (1 - I_{t-1})(c_2 + \varepsilon_t - \theta_2 \varepsilon_{t-1})$$

$$\varepsilon_t = I_{t-1} \varepsilon_{1t} + (1 - I_{t-1}) \varepsilon_{2t}$$

Date le informazioni $\mathfrak{F}_N = (r_1, r_2, \dots, r_N)$, ricordando quanto detto nel paragrafo 4.8, le previsioni del TMA(2;1) sotto le assunzioni

$$E(\varepsilon_{jN+k} | \mathfrak{F}_N) = 0, \quad j=1, 2; \quad k=1, 2, \dots$$

$$E(I_{N+k-1} | \mathfrak{F}_N) = p, \quad k=2, 3, \dots$$

$$E(I_{N+k-1} \varepsilon_{jN+k-1} | \mathfrak{F}_N) = E(I_{N+k-1} \varepsilon_{jN+k} | \mathfrak{F}_N) = 0, \quad j=1, 2; \quad k=2, 3, \dots$$

sono date da

$$r_N(1) = E(r_{N+1} | \mathfrak{F}_N) = (I_N c_1 + (1 - I_N) c_2) - (I_N \theta_1 \varepsilon_N + (1 - I_N) \theta_2 \varepsilon_N)$$

$$r_N(k) = (p c_1 + (1 - p) c_2), \quad k = 2, 3, \dots$$

Questi risultati permettono di derivare l'errore di previsione

$$e_N(1) = r_{N+1} - r_N(1) = I_N \varepsilon_{N+1} + (1 - I_N) \varepsilon_{N+1}$$

$$e_N(k) = r_{N+k} - r_N(k) =$$

$$= (I_{N+k-1} - p)(c_1 - c_2) + I_{N+k-1}(\varepsilon_{N+k} - \theta_1 \varepsilon_{N+k-1}) +$$

$$+ (1 - I_{N+k-1})(\varepsilon_{N+k} - \theta_1 \varepsilon_{N+k-1}), \quad k = 2, 3, \dots$$

e la varianza dell'errore di previsione:

$$\sigma_p^2(1) = I_N \sigma_1^2 + (1 - I_N) \sigma_2^2$$

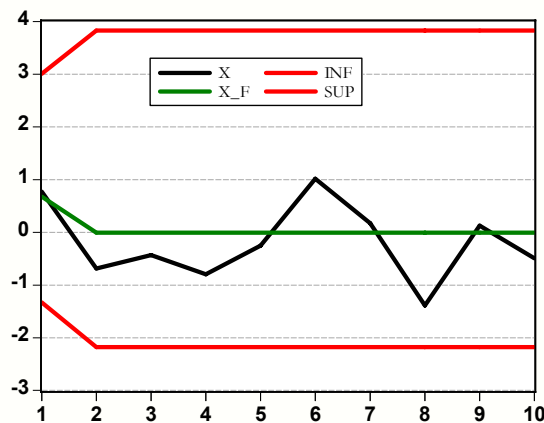
$$\sigma_p^2(k) = \sigma^2 = p(1 + \theta_1^2) \sigma_1^2 + (1 - p)(1 + \theta_2^2) \sigma_2^2, \quad k = 2, 3, \dots$$

Le previsioni e le varianze di previsione possono essere stimate sostituendo ai parametri incogniti le relative stime e quindi utilizzate nelle pratiche applicazioni.

Esempio 11

Utilizziamo questi risultati per prevedere le ultime $k=10$ osservazioni del TMA(2; 1) riportato nell'esempio 4.

Figura 31 – Serie osservata X , prevista X_F e relativa banda di previsione al 95%

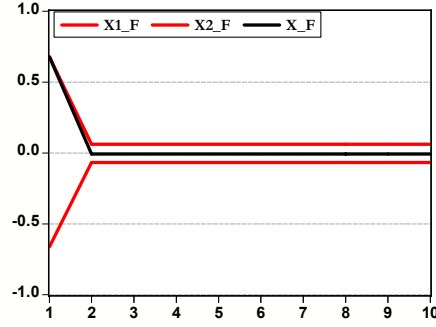


In questo caso si ha

$$N=14990; I_N = 1; \varepsilon_N = 0.704329; \hat{p} = 0.5455; \hat{c}_1 = 0.111; \hat{c}_2 = -0.148$$

$$\hat{\theta}_1 = -0.804; \hat{\theta}_2 = 0.731; \hat{\sigma}_1^2 = 1.006; \hat{\sigma}_2^2 = 0.3952$$

Utilizzando i risultati sopra presentati, le previsioni, i dati osservati e le relative bande di confidenza sono riportati nella figura 31. Nella figura 32 sono riportate le previsioni di X e quelle dei suoi due regimi.

Figura 32 – Previsione di X e dei suoi due regimi

□

Le analisi ed i risultati visti per un TMA(2; 1) possono essere estesi, con qualche complicazione formale al caso generale dei TMA(k;q) che hanno k regimi e q parametri MA in ciascun regime. In particolare, per k=2 regimi, con qualche complicazione formale di tipo puramente algebrico, tenendo conto che

$I_{t-1}^2 = I_{t-1}$ e che $I_{t-1}(1 - I_{t-1}) = 0$ per ogni t, risulta:

$$r_N(1) = -\sum_{j=1}^q (I_N \theta_{1j} + (1 - I_N) \theta_{2j}) \varepsilon_{N-j+1}; \quad e_N(1) = \varepsilon_{N+1}$$

$$E(e_N(1)) = 0; \quad \sigma_p^2(1) = \sigma^2$$

mentre per $k = 2, 3, \dots, q$ si ha

$$r_N(k) = -\sum_{j=k}^q (p \theta_{1j} + (1-p) \theta_{2j}) \varepsilon_{N+k-j}$$

$$e_N(k) = \varepsilon_{N+k} + \sum_{j=1}^{k-1} ((I_{N+k-1} - 1) \theta_{2j} - I_{N+k-1} \theta_{1j}) \varepsilon_{N+k-j} +$$

$$+ (p - I_{N+k-1}) \sum_{j=k}^q (\theta_{1j} - \theta_{2j}) \varepsilon_{N+k-j}$$

$$E(e_N(k)) = 0$$

$$\sigma_p^2(k) = \sigma^2 \left[1 + \sum_{j=1}^{k-1} (p\theta_{1j}^2 + (1-p)\theta_{2j}^2) + p(1-p) \sum_{j=k}^q (\theta_{1j} - \theta_{2j})^2 \right]$$

ed infine, per $k=q+1, q+2, \dots$, risulta

$$r_N(k) = 0; \quad e_N(k) = r_{N+k}$$

$$E(e_N(k)) = 0; \quad \sigma_p^2(k) = \sigma^2 \left[1 + \sum_{j=1}^q (p\theta_{1j}^2 + (1-p)\theta_{2j}^2) \right]$$

Come si può osservare, le previsioni sono sempre non distorte; inoltre, quando è $p=1$ oppure $p=0$ e quindi $\theta_{1j} = \theta_{2j}$, $j=1, 2, \dots, q$, il TMA(2; q) si riduce al classico MA(q) e le previsioni, gli errori di previsione e la varianza di questi si riducono a quella degli MA. In pratica i parametri θ_{1j}, θ_{2j} , $j=1, 2, \dots, q$, sono incogniti e per poter effettivamente calcolare le previsioni e la loro varianza si sostituiscono con le relative stime.

Nell'applicazione seguente mostriamo come è possibile stimare e prevedere un TARMA(2; 1, 5) applicato alle cinque serie analizzate nelle tabelle 31 e 32.

Esempio 12

Consideriamo le serie storiche dei rendimenti giornalieri dei cinque attivi Campari, Fiat, Mibtel, Nasdaq, S&P analizzate nelle tabelle 31 e 32, adattiamo a ciascuna di queste un modello TMA(2; q) con $q \leq 5$, confrontiamo i risultati con quelli della tabella 31 e costruiamo le relative previsioni per i successivi $k=10$ giorni.

Nella tabella 33 sono riportati gli indici descrittivi più rilevanti di ciascuno dei due regimi per le cinque serie. Da questi indici emerge che effettivamente le serie hanno regimi, in particolare per quanto riguarda asimmetria e curtosi, diversi fra di loro e questo potrebbe giustificare l'uso del modello a soglia. Con l'esclusione di Fiat, nelle altre quattro serie la varianza e la curtosi del secondo regime è sempre più elevata di quella del primo. Ai rendimenti giornalieri di ciascuna serie è stato adattato un TMA(2; 5).

La stima, con il metodo dei minimi quadrati, è stata effettuata più volte per eliminare dal modello i parametri che non erano significativamente diversi da zero. Inoltre, come per tutti i modelli con componente MA, le stime sono

state ottenute iterativamente utilizzando una procedura simile a quella degli ARMA.

Tabella 33 – *Indici descrittivi dei regimi delle cinque serie analizzate*

	R_CAMP		R_FIAT		R_MIBTEL		R_NSDQ		R_S&P	
Reg.	1	2	1	2	1	2	1	2	1	2
\bar{x}	-0.0004	0.0010	-0.0002	-0.0009	0.0000	-0.0001	-0.0009	0.0012	-0.0013	0.0013
Me	0.0000	0.0000	0.0000	0.0000	0.0001	0.0007	0.0000	0.0012	0.0000	0.0016
Max	0.1097	0.0743	0.0830	0.1421	0.0814	0.1037	0.0614	0.1184	0.0627	0.1095
Min	-0.0666	-0.1027	-0.1579	-0.0976	-0.0507	-0.0676	-0.0607	-0.1111	-0.0935	-0.0946
S	0.0160	0.0175	0.0217	0.0258	0.0106	0.0135	0.0140	0.0166	0.0118	0.0145
$\hat{\gamma}_1$	0.4463	-0.2511	-0.5069	0.0855	0.1622	0.4482	-0.2784	-0.0108	-1.5553	0.3083
$\hat{\gamma}_2$	4.0172	4.0774	4.8072	2.7171	8.829	11.311	5.5206	13.540	14.559	17.376
JB	618.12	507.04	844.80	233.79	2849.6	3824.9	244.07	3324.0	5344.0	6056.9
Prob	0.0000	0.0000	0.0000	0.00	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
N	876	721	840	757	876	713	879	718	895	702

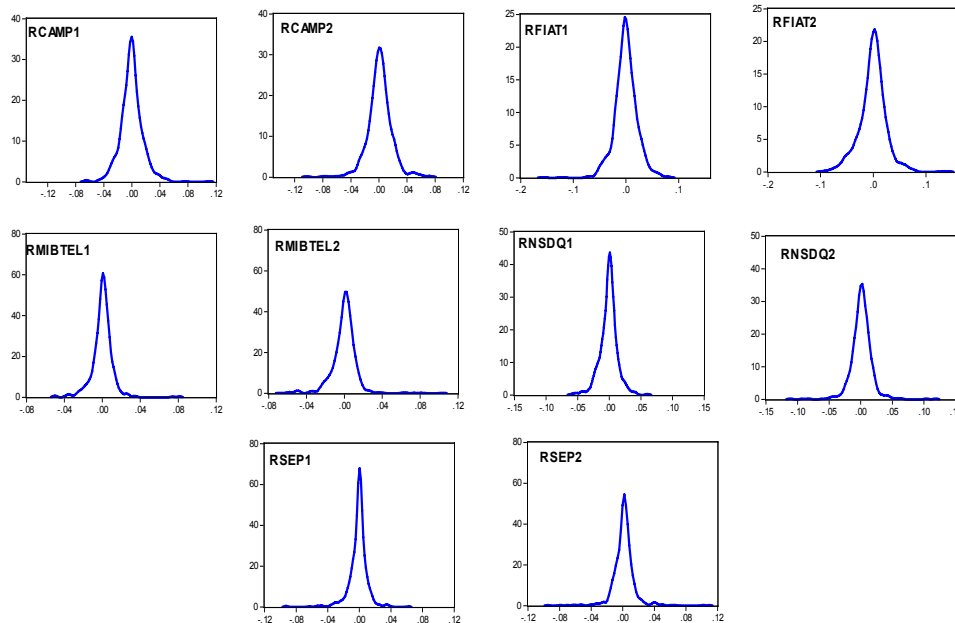
Tabella 34 – *Stime finali dei modelli TARMA(2;1, 5) per le cinque serie*

	$\hat{\theta}_{1i}$	$\hat{\theta}_{2i}$	$\hat{\theta}_{3i}$	$\hat{\theta}_{4i}$	$\hat{\theta}_{5i}$	$\hat{\sigma}_i$	$\hat{\gamma}_{1i}$	$\hat{\gamma}_{2i}$	$\hat{\tau}$	R^2	AIC
Campari	.096	.058				.0166	.012	4.31		.013	
1° reg.	.102	.082				.0159	.415	3.82	.549	.013	-5.35
2° reg.	.090					.0175	-3.47	4.30	.451		
Fiat		-.074				.0237	-.127	3.62		.005	
1° reg.						.0236	-7.64	4.91	.526	.004	-4.64
2° reg.		-.090				.0284	.137	2.69	.474		
Mibtel			.054	-.105	.092	.0114	-.330	9.18		.026	
1° reg.						.0098	-1.18	10.57	.553	.045	-6.11
2° reg.			.121	-1.88	.206	.0130	.139	7.37	.447		
Nasdaq	.121	.065	-.059			.0152	-.203	7.34		.022	
1° reg.	.122					.0164	-2.00	8.80	.550	.026	-5.54
2° reg.	.118	.159				.0140	-2.33	2.48	.450		
S&P	.149	.103	-.096		.069	.0129	-.722	12.49		.042	
1° reg.	.151			.075	.081	.0117	-1.51	11.69	.560	.049	-5.87
2° reg.	.157	.187	-1.05			.0141	-1.20	12.06	.440		

Come atteso, è risultato $c_1 = c_2 = 0 = \bar{r} = 0$. Le stime finali sono riportate nella tabella 34. I parametri qui riportati sono tutti significativamente diversi da

zero, almeno al 5%, e per semplicità espositiva sono stati omessi i relativi standard error. In tale tabella l'indice $i = 1, 2$ indica il corrispondente regime. Per facilitare il confronto con la tabella 31, nella 34 sono riportate le stime (in grassetto) del modello MA(q), $q \leq 5$, adattato alle cinque serie. Nella figura 33 sono riportate le stime delle funzioni di densità locali riferite ai rendimenti dei regimi nelle cinque serie. Dalle tabelle 33 e 34 e dalla figura 33 si può osservare che, con l'esclusione di Fiat in cui la presenza di regimi e quindi di non linearità nei livelli è labile, nelle altre quattro è abbastanza significativa tanto che la struttura MA dei due regimi è molto diversa fra di loro ed il secondo risulta sistematicamente più complesso del primo.

Figura 33 – Stime delle funzioni di densità dei due regimi delle cinque serie analizzate



□

Per effettuare le stime è stato costruito un programma in EViews ed il relativo listato è riportato qui di seguito.

'Stima TARMA(2; 1, 5)

smpl 2 **N**

genr **r** =d(log(**x**))

genr dt=1

genr res=0

genr r1=1

```

genr r2=1
delete dt r1 r2 res
equation eq0.ls r=c(1)+c(2)*r(-1)
genr res=resid
for li=2 to N
    simpl li li
    if r(li-1)>=0 then genr dt=1
    genr r1=r
    else genr dt=0
    genr r2=r
endif
next
line r1 r2
simpl 3 N
for li=1 to M
    simpl 8 N
    equation eq1.ls r=dt*(c(1)+c(2)*r(-1)-c(3)*res(-1)-c(4)*res(-2)-
c(5)*res(-3)-c(6)*res(-4)-c(7)*res(-5))+(1-dt)*(c(8)+c(9)*r(-1)-c(10)*res(-1)-
c(11)*res(-2)-c(12)*res(-3)-c(13)*res(-4)-c(15)*res(-5))
    genr res=resid
next

```

Nota

Nel programma è necessario sostituire a:

1. **x**, il nome della serie dei prezzi dell'attivo di cui si vuole calcolare i rendimenti, qui chiamati **r**, ed adattare il modello TARMA(2; 1, 5). La serie **x** deve essere contenuta nel workfile e questo deve essere attivo;
2. **N**, la numerosità della serie. Deve essere non inferiore a 500. Nell'applicazione è $N = 1604$;
3. **M**, il numero delle iterazioni per la stima della parte MA. Deve essere non inferiore a 100. Nell'applicazione è $M=200$.

Il valore 8 nella istruzione “simpl 8 **N**” è dovuto al fatto che, in questo caso, per stimare eq1 si perdono le prime 7 informazioni: una per calcolare i rendimenti, una per stimare l'equazione iniziale eq0 che contiene un ritardo, cinque per stimare l'equazione finale eq1 che contiene cinque ritardi.

L'output di interesse:

1. **dt**, la dicotomica necessaria per derivare una stima di $\tau = \Pr(dt = 1)$ calcolandone la media. In generale questa media deve essere compresa in

- 0.40 e 0.60;
2. r_1 , la serie del primo regime;
 3. r_2 , la serie del secondo regime che, se strutturalmente diversa da r_1 , giustifica l'adozione del modello TARMA;
 4. res , i residui stimati del TARMA utili per ottenere stime delle varianze dei singoli regimi;
 5. eq_1 l'equazione che contiene la stima finale dei parametri, la loro significatività, gli indici di accostamento, l'analisi dei residui, ecc.

Se il TARMA da stimare ha (p, q) diverso da $(1, 5)$ qui usato, il listato può essere facilmente modificato di conseguenza. Per i rendimenti di attivi, i parametri (p, q) qui utilizzati sono, di solito, sufficienti. Osserviamo che nella stima non si sono utilizzate le ultime 10 osservazioni proprio per poterle confrontare con le previsioni ottenute dal modello e quindi per verificare fino a che punto il modello TARMA fornisce informazioni utili ai fini della previsione ed in particolare se tali previsioni sono più o meno informative di quelle ottenute nel caso lineare.

□

Esempio 12 (continua)

Utilizziamo i risultati e le stime sopra riportate per il TMA(2; q), costruiamo le previsioni per $k=10$ giorni e confrontiamole con gli ultimi 10 valori effettivamente osservati, ma non utilizzati nella stima. Nella tabella 35 sono riportati i valori osservati dei rendimenti delle cinque serie analizzate e quelli previsti (in grassetto).

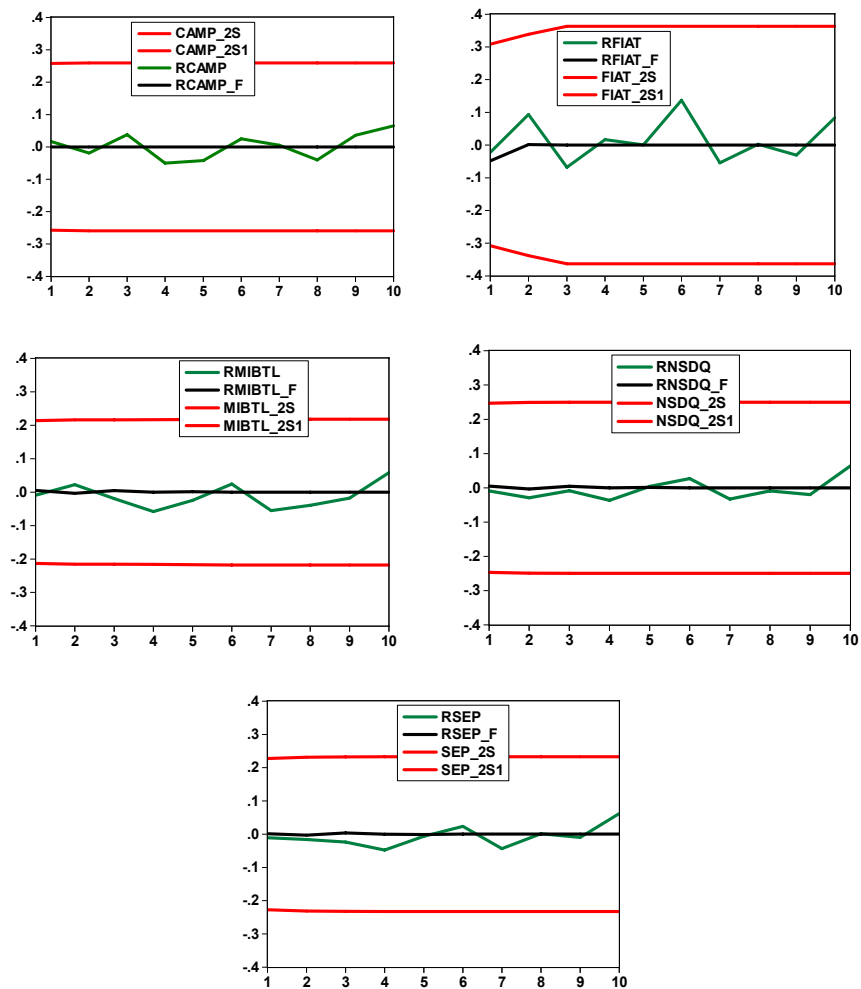
Tabella 35 – Valori osservati e previsti (in grassetto) dei rendimenti con il TMA(2; q)

k	camp	camp_f	fiat	fiat_f	mibt1	mibt1_f	nsdq	nsdq_f	s&p	s&p_f
1	0.017	0.000	-0.023	-0.049	-0.010	0.005	-0.009	0.005	-0.011	0.001
2	-0.019	0.000	0.093	0.002	0.022	-0.004	-0.029	-0.004	-0.016	-0.004
3	0.038	0.000	-0.068	0.000	-0.020	0.005	-0.009	0.005	-0.024	0.004
4	-0.050	0.000	0.017	0.000	-0.058	0.000	-0.037	0.000	-0.048	0.000
5	-0.042	0.000	0.000	0.000	-0.025	0.001	0.004	0.001	-0.006	-0.001
6	0.025	0.000	0.137	0.000	0.024	0.000	0.027	0.000	0.023	0.000
7	0.005	0.000	-0.054	0.000	-0.055	0.000	-0.033	0.000	-0.043	0.000
8	-0.040	0.000	0.003	0.000	-0.039	0.000	-0.009	0.000	0.001	0.000
9	0.035	0.000	-0.031	0.000	-0.018	0.000	-0.020	0.000	-0.010	0.000
10	0.065	0.000	0.083	0.000	0.058	0.000	0.064	0.000	0.062	0.000

Nelle figure seguenti riportiamo i valori osservati e previsti per le cinque serie e le relative bande di confidenza al livello $1 - \alpha = 0.95$. Per rendere possibile il confronto è stata utilizzata una identica scala per i cinque grafici. Come si può

osservare, le bande di confidenza sono molto larghe indice di una forte volatilità del fenomeno. Questa volatilità è più marcata per i rendimenti di Fiat e Campari, ed in particolare per Fiat, e meno come atteso, per le altre tre serie che sono riferite ai rendimenti medi di interi mercati. Si osserva anche che nel confronto dei tre mercati Mibtel, Nasdaq e S&P500, quello Mibtel riferito al mercato borsistico italiano risulta marginalmente meno volatile.

Figura 34 – Valori osservati, previsti e bande di confidenza (in rosso) per i rendimenti delle cinque serie analizzate



□

Dall'esempio sopra riportato si ha conferma come nei livelli dei rendimenti

vi siano strutture di dipendenza, siano esse lineari che non, molto labili e praticamente trascurabili. Viceversa i rendimenti globali presentano una struttura non lineare nella volatilità come si può desumere dalla sostanziale diversità della variabilità locale e della curtosi locale riferita ai due regimi. Questi aspetti verranno analizzati e messi in evidenza nei prossimi due capitoli.

CAPITOLO V

Analisi della volatilità di attivi finanziari e monetari

5.1 *Il modello ARCH classico*

Nel capitolo precedente abbiamo analizzato e previsto, utilizzando la modellistica lineare ARMA e la sua generalizzazione TARMA, i rendimenti di attivi monetari e finanziari e, come atteso, abbiamo evidenziato che l'incidenza dei livelli medi sulla struttura evolutiva di questi fenomeni è assente o molto modesta. In quelle analisi abbiamo anche sottolineato che, quando ha senso supporre in essi dei regimi, questi si differenziano, più che per i diversi livelli medi, per la loro diversa variabilità e curtosi.

In questo capitolo, partendo dalla constatazione empirica che i rendimenti di attivi mostrano sistematicamente un eccesso di curtosi, verrà analizzato un loro aspetto caratterizzante che va sotto il nome di *volatilità*. La volatilità nasce dal fatto che la variabilità dei rendimenti non è costante nel tempo, ma evolve con questo ed a periodi di bassa variabilità ne seguono, di solito, altri con più elevata variabilità. Da questo punto di vista la volatilità non è altro che l'attitudine dei rendimenti a possedere una variabilità evolutiva (funzione del tempo) che si presenta in forme clusterizzate, cioè a grappoli. D'altra parte la variabilità di questi fenomeni può essere anche interpretata come una misura della *rischiosità* del relativo attivo: più la variabilità, la volatilità, è elevata, minore è la sua affidabilità e quindi maggiore il *rischio di mercato* di quell'attivo.

L'analisi della volatilità degli attivi finanziari può essere condotta ricorrendo a strumenti diversi. Uno di questi, di tipo non parametrico e basato sulla varianza mobile, è stato illustrato nel capitolo II. Qui di seguito utilizziamo uno strumento alternativo di tipo parametrico basato sull'uso di semplici modelli statistici della classe ARCH e sue generalizzazioni. Per analizzare e prevedere la volatilità di attivi Engle (1982) propose questa classe di modelli che si è dimostrata molto flessibile per la sua adattabilità ai fenomeni reali ed in possesso delle peculiari proprietà statistiche e probabilistiche proprie dei rendimenti di attivi finanziari. Consideriamo, per il momento, il caso più

semplice dell'ARCH(1). Più precisamente supponiamo che:

1. i rendimenti r_t sono stazionari, non hanno alcuna struttura nei livelli e si comportano come un white noise: $r_t \sim \text{WN}(0, \sigma^2)$
2. posto $\sigma_t > 0$, sia:

$$r_t = \sigma_t a_t \quad (1)$$

3. $a_t \stackrel{\text{iid}}{\sim} (0, 1)$ è continua con funzione di densità $f(a) \geq 0$ definita su tutta la retta reale;
4. si abbia:
- 5.

$$\sigma_t^2 = w + \alpha r_{t-1}^2, \quad \text{con } w > 0; 0 \leq \alpha < 1 \quad (2)$$

Studiamo le proprietà statistiche di questo modello e verifichiamo se è idoneo a descrivere i rendimenti e quindi se possiede le proprietà tipiche di questi: varianza non costante nel tempo ed eccesso di curtosi. Per far questo calcoliamo, sotto le condizioni sopra riportate, il vettore dei parametri $(\mu, \sigma^2, \gamma_1, \gamma_2)$. Indichiamo con \mathfrak{F}_{t-1} le informazioni che si hanno su r_t al tempo t , di solito è $\mathfrak{F}_{t-1} = (r_1, r_2, \dots, r_{t-1})$, avremo:

(A) **media** di r_t :

$$\begin{aligned} E(r_t) &= E[E(r_t | \mathfrak{F}_{t-1})] = E[E(\sigma_t a_t | \mathfrak{F}_{t-1})] = E[E((w + \alpha r_{t-1}^2)^{1/2}) a_t | \mathfrak{F}_{t-1}] = \\ &= E\left[(w + \alpha r_{t-1}^2)^{1/2} E(a_t | \mathfrak{F}_{t-1}) \right] = E\left[(w + \alpha r_{t-1}^2)^{1/2} E(a_t) \right] = 0 \end{aligned} \quad (3)$$

e come atteso, i rendimenti descritti da quel modello sono in media nulli;

(B) **media condizionata** di r_t alle informazioni disponibili \mathfrak{F}_{t-1} :

$$E(r_t | \mathfrak{F}_{t-1}) = E\left[(w + \alpha r_{t-1}^2)^{1/2} a_t | \mathfrak{F}_{t-1} \right] = (w + \alpha r_{t-1}^2)^{1/2} E(a_t | \mathfrak{F}_{t-1}) =$$

$$= \left(w + \alpha r_{t-1}^2 \right)^{1/2} E(a_t) = 0$$

ed anche questi, come atteso, sono nulli;

(C) **varianza** di r_t :

$$\begin{aligned} \sigma^2 &= \text{Var}(r_t) = E(r_t^2) = E\left[E(r_t^2 | \mathfrak{F}_{t-1})\right] = E\left[E(\sigma_t^2 a_t^2 | \mathfrak{F}_{t-1})\right] = \\ &= E\left[E\left((w + \alpha r_{t-1}^2) a_t^2 | \mathfrak{F}_{t-1}\right)\right] = E\left[(w + \alpha r_{t-1}^2) E(a_t^2 | \mathfrak{F}_{t-1})\right] = \\ &= E\left[(w + \alpha r_{t-1}^2) \text{Var}(a_t^2)\right] = E(w + \alpha r_{t-1}^2) = w + \alpha E(r_{t-1}^2) = w + \alpha \sigma^2 \end{aligned}$$

che implica

$$\sigma^2 = \frac{w}{1 - \alpha} \quad (4)$$

e la (4) giustifica le condizioni sui parametri: $w > 0$, $0 \leq \alpha < 1$;

(D) **varianza condizionata** di r_t alle informazioni disponibili \mathfrak{F}_{t-1} :

$$\begin{aligned} \sigma_t^2 &= \text{Var}(r_t | \mathfrak{F}_{t-1}) = \text{Var}(\sigma_t a_t | \mathfrak{F}_{t-1}) = \text{Var}\left((w + \alpha r_{t-1}^2)^{1/2} a_t | \mathfrak{F}_{t-1}\right) = \\ &= (w + \alpha r_{t-1}^2) \text{Var}(a_t | \mathfrak{F}_{t-1}) = (w + \alpha r_{t-1}^2) \text{Var}(a_t) = w + \alpha r_{t-1}^2 \end{aligned}$$

che, mostra come la (2), non è altro che la varianza dei rendimenti condizionata alle informazioni \mathfrak{F}_{t-1} che, per le condizioni poste sulla (4), è sempre positiva, mentre la (4) è la varianza non condizionata dei rendimenti che, per le condizioni poste sui due parametri (w , α), è anche essa sempre positiva;

(E) **asimmetria** di r_t :

tenendo conto della (3), l'indice di asimmetria dei rendimenti è dato da

$$\gamma_1 = \frac{E(r_t^3)}{\sigma^3}$$

Indichiamo con γ_{1a} l'indice di asimmetria di a_t , ricordando che si è posto $\text{Var}(a_t) = 1$, utilizzando una dimostrazione simile a quelle che ha permesso di derivare la (3) e la (4), si ha

$$\begin{aligned} E(r_t^3) &= E[E(r_t^3 | \mathfrak{F}_{t-1})] = E\left[\left(w + \alpha r_{t-1}^2\right)^{3/2} E(a_t^3 | \mathfrak{F}_{t-1})\right] = \\ &= E\left[\left(w + \alpha r_{t-1}^2\right)^{3/2} E(a_t^3)\right] = \gamma_{1a} E\left[\left(w + \alpha r_{t-1}^2\right)^{3/2}\right] \end{aligned}$$

che ci permette di ottenere

$$\gamma_1 = \gamma_{1a} E\left[\left(\frac{w + \alpha r_{t-1}^2}{w + \alpha \sigma^2}\right)^{3/2}\right] \quad (5)$$

Ma, per costruzione, è $\frac{w + \alpha r_{t-1}^2}{w + \alpha \sigma^2} > 0$ e quindi il segno dell'asimmetria dei

rendimenti dipende da quella del processo $a_t \stackrel{\text{iid}}{\sim} (0, 1)$. In particolare, se l'asimmetria di a_t è nulla lo è anche quella dei rendimenti e viceversa, questo è

sicuramente vero se è $a_t \sim N(0, 1)$, oppure $a_t \sim \left(\frac{k-2}{k}\right)^{1/2} t_{(k)}$ con $k > 3$.

(F) **curtosi** di r_t :

tenendo conto della (3), l'indice di curtosi dei rendimenti è dato da

$$\gamma_2 = \frac{E(r_t^4)}{\sigma^4} - 3 = \frac{\mu_4}{\sigma^4} - 3$$

Se indichiamo con γ_{2a} l'indice di curtosi di a_t , con un ragionamento simile a quello utilizzato per derivare la (5) si ottiene:

$$\begin{aligned}\mu_4 &= E(r_t^4) = E[E(r_t^4 | \mathfrak{F}_{t-1})] = E[E(\sigma_t^4 a_t^4 | \mathfrak{F}_{t-1})] = E[(w + \alpha r_{t-1}^2)^2 E(a_t^4)] = \\ &= (\gamma_{2a} - 3) E[(w + \alpha r_{t-1}^2)^2] = (\gamma_{2a} - 3) E[w^2 + \alpha^2 r_{t-1}^4 + 2w\alpha r_{t-1}^2] = \\ &= (\gamma_{2a} - 3) [w^2 + \alpha^2 \mu_4 + 2w\alpha \text{Var}(r_{t-1})]\end{aligned}$$

sostituendo in $\text{Var}(r_{t-1})$ la (4) si ha

$$\mu_4 = \frac{\gamma_{2a} + 3}{1 - (\gamma_{2a} + 3)\alpha^2} \frac{w^2(1 + \alpha)}{1 - \alpha}$$

che permette di calcolare l'indice di curtosi dei rendimenti con struttura ARCH(1):

$$\gamma_2 = \frac{\mu_4}{\sigma_r^4} - 3 = \frac{(\gamma_{2a} + 3)(1 + 2\alpha^2) - 3}{1 - (\gamma_{2a} + 3)\alpha^2} \quad (6)$$

Se è $a_t \stackrel{\text{iid}}{\sim} N(0, 1)$ e quindi $\gamma_{2a} = 0$, la (6) diventa

$$\gamma_2 = \frac{6\alpha^2}{1 - 3\alpha^2}$$

e risulta $\gamma_2 > 0$ per $0 \leq \alpha < 1/\sqrt{3} \approx 0.57735$. In altri termini, i rendimenti possono avere un eccesso di curtosi anche se il residuo standardizzato a_t è distribuito normalmente.

In conclusione, il modello descritto nelle (1) - (2) è effettivamente in grado di catturare gli aspetti più rilevanti dei rendimenti di attivi finanziari: variabilità non costante nel tempo ed eccesso di curtosi.

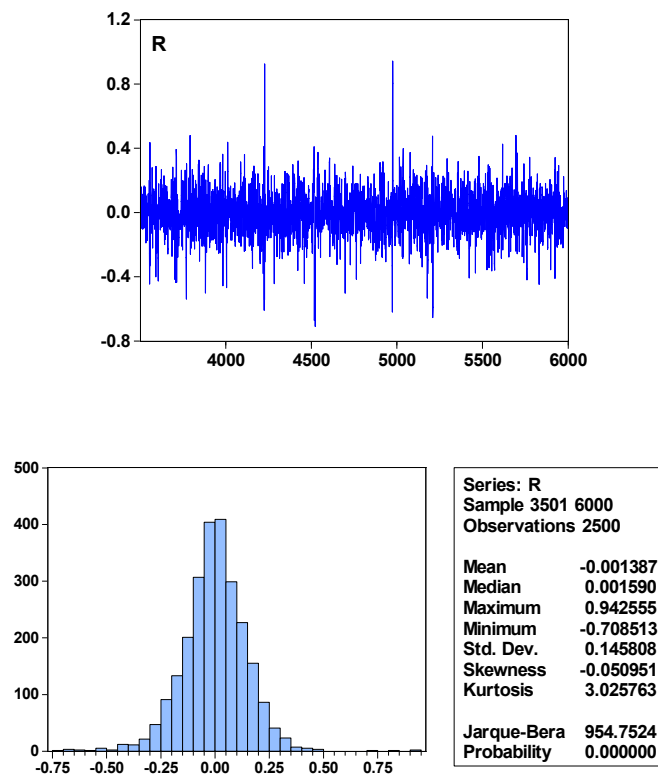
Esempio 1

Simuliamo una serie dei rendimenti con struttura white noise per i livelli e ARCH(1) per la volatilità:

$$\begin{cases} r_t = \sigma_t a_t \\ \sigma_t^2 = 0.01 + 0.5r_{t-1}^2 \end{cases}$$

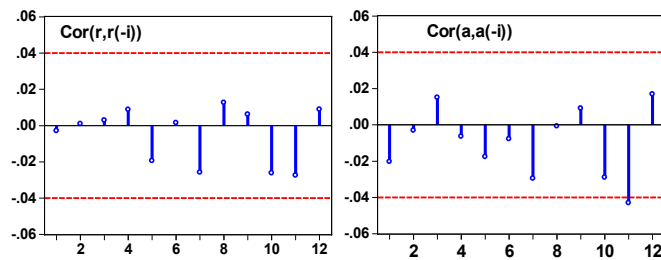
con $a_t \stackrel{iid}{\sim} N(0, 1)$. I valori iniziali utilizzati per la simulazione sono $r_0 = 0.03$ e $\sigma_0^2 = 0.0002$. Il numero delle osservazioni simulate è stato pari a 6000, di queste sono state utilizzate le ultime 2500 per eliminare l'effetto dei valori iniziali. Nella figura 1 è riportato il grafico di r_t ed il suo istogramma di frequenze con le relative statistiche descrittive.

Figura 1 – Serie r_t e suo istogramma di frequenze con relativi indici descrittivi



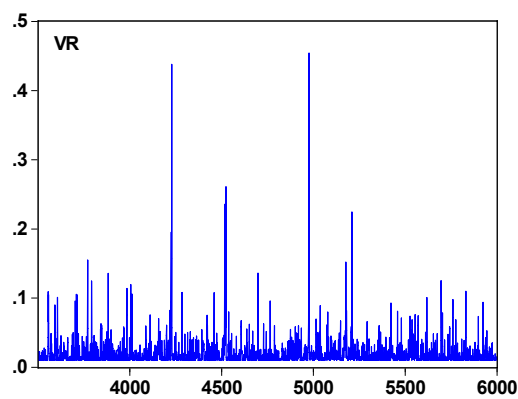
Da questo grafico emerge chiaro l'eccesso di curtosi in r_t anche se si è posto $a_t \stackrel{iid}{\sim} N(0, 1)$ e quindi a_t è distribuita come una normale standardizzata per ogni t . Osserviamo inoltre che dalla (4) deve essere $\sigma = 0.14142$, mentre dall'istogramma risulta $\hat{\sigma} = 0.1458$ che è molto vicino a quello vero.

Figura 2 – Prime 12 autocorrelazioni di r_t ed a_t



Nella figura 2 sono riportate le prime 12 autocorrelazioni di r_t ed a_t che sono interne alla relativa banda di confidenza (in rosso) al 95% e quindi sono entrambe compatibili, come atteso, con un white noise. Nella figura 3 è riportato il grafico di σ_t^2 in cui è evidente che la varianza non è costante nel tempo e si presenta a *cluster*: periodi di alta volatilità si susseguono a periodi di bassa volatilità.

Figura 3 – Grafico di σ_t^2 nella serie simulata



□

Quanto visto per un ARCH(1) può essere generalizzato, con qualche complicazione formale, al modello ARCH(s), con $s \geq 1$, in cui si suppone che la varianza condizionata sia funzione di s ritardi al quadrato dei rendimenti:

$$\sigma_t^2 = w + \alpha_1 r_{t-1}^2 + \alpha_2 r_{t-2}^2 + \dots + \alpha_s r_{t-s}^2 \quad (7)$$

In questo caso, note le informazioni $\mathfrak{I}_{t-s} = (r_1, r_2, \dots, r_{t-s})$, con una procedura simile a quella dell'ARCH(1) si dimostra che è

$$E(r_t) = 0$$

$$\sigma^2 = \text{Var}(r_t) = \frac{w}{1 - (\alpha_1 + \alpha_2 + \dots + \alpha_s)} \quad (8)$$

che implica che le condizioni sufficienti di esistenza di un ARCH(s) sono

$$w > 0; \quad 0 \leq \alpha_i < 1, i=1, 2, \dots, s; \quad 0 \leq \alpha_1 + \alpha_2 + \dots + \alpha_s < 1$$

Se poniamo

$$\mathbf{r}'_{t-1} = (r_{t-1}, r_{t-2}, \dots, r_{t-s}); \quad \mathbf{D}_\alpha = \begin{bmatrix} \alpha_1 & 0 & \dots & 0 \\ 0 & \alpha_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \alpha_s \end{bmatrix}$$

la (7) si può scrivere

$$\sigma_t^2 = w + \mathbf{r}'_{t-1} \mathbf{D}_\alpha \mathbf{r}_{t-1} \quad (9)$$

La (9) mostra che la varianza condizionata σ_t^2 è positiva se è $w > 0$ e \mathbf{D}_α una matrice semi definita positiva, ma dato che \mathbf{D}_α è diagonale, questo è possibile solo se risulta $\alpha_i \geq 0$, $i = 1, 2, \dots, s$, d'altra parte, per la (8) deve essere $0 \leq \alpha_1 + \alpha_2 + \dots + \alpha_s < 1$. L'idea di scrivere un ARCH(s) come la (9) permette

di definire un $ARCH(s)$ generalizzato:

$$\sigma_t^2 = w + \mathbf{r}'_{t-1} \Sigma \mathbf{r}_{t-1}$$

con Σ matrice simmetria e semi definita positiva. L'inconveniente di questa generalizzazione è il numero di parametri che bisogna stimare che cresce rapidamente con s . Infatti i parametri di questo modello sono pari al numero degli elementi diversi di Σ più w e quindi in totale $m=1+s(s+1)/2$. Per esempio è $s=6$ i parametri da stimare sono $m=22$ che per un modello univariato sono effettivamente un numero eccessivo.

5.2 Il modello ARCH con struttura nei livelli

Nel paragrafo precedente abbiamo supposto che i rendimenti non possedessero una struttura nei livelli medi e quindi potessero essere rappresentati con un processo white noise. D'altro lato, nel capitolo precedente abbiamo visto che nei livelli è qualche volta presente una struttura del tipo ARMA anche se molto debole e semplice.

Supponiamo che r_t sia rappresentabile con un ARMA(p,q):

$$r_t = \phi_0 + \phi_1 r_{t-1} + \phi_2 r_{t-2} + \dots + \phi_p r_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$$

con $\varepsilon_t \sim WN(0; \sigma^2)$. In questo caso la (1) e la (7) diventano

$$\varepsilon_t = \sigma_t a_t; \quad \sigma_t^2 = w + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \dots + \alpha_s \varepsilon_{t-s}^2$$

e la volatilità va cercata sui residui del modello ARMA per r_t . In tal caso ai rendimenti è necessario adattare simultaneamente il modello ARMA(p, q)-ARCH(s) dato dalle seguenti tre equazioni:

$$r_t = \phi_0 + \phi_1 r_{t-1} + \phi_2 r_{t-2} + \dots + \phi_p r_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$$

$$\varepsilon_t = \sigma_t a_t$$

$$\sigma_t^2 = w + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \dots + \alpha_s \varepsilon_{t-s}^2 \tag{10}$$

E' interessante osservare che se nella (10) si pone $\sigma_t^2 = \varepsilon_t^2 - \eta_t$ questa si può scrivere:

$$\varepsilon_t^2 = w + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \dots + \alpha_s \varepsilon_{t-s}^2 + \eta_t$$

ed il modello ARCH(s) si trasforma in un particolare AR(s) sui quadrati dei residui ε_t^2 dei rendimenti. Osserviamo che il processo η_t non è iid ma è solo un white noise. Per verificarlo consideriamo il modello ARCH(1), in questo caso risulta

$$\varepsilon_t^2 = w + \alpha \varepsilon_{t-1}^2 + \eta_t$$

da cui

$$\eta_t = \varepsilon_t^2 - w - \alpha \varepsilon_{t-1}^2 = \varepsilon_t^2 - \sigma_t^2$$

Per la **media** di η_t , per quanto visto nei punti (C) - (D), si ha

$$E(\eta_t) = E(\varepsilon_t^2 - \sigma_t^2) = E(\varepsilon_t^2) - E(\sigma_t^2) = \sigma_\varepsilon^2 - \sigma_\varepsilon^2 = 0$$

Per la **varianza** di η_t , ricordando la seconda delle (10), risulta

$$\begin{aligned} \text{Var}(\eta_t) &= \text{Var}(\varepsilon_t^2 - \sigma_t^2) = E(\varepsilon_t^4 + \sigma_t^4 - 2\varepsilon_t^2 \sigma_t^2) = \\ &= E(\varepsilon_t^4 + \sigma_t^4 - 2a_t^2 \sigma_t^4) = E(\varepsilon_t^4) + E(\sigma_t^4) - 2E(a_t^2 \sigma_t^4) = E(\varepsilon_t^4) - E(\sigma_t^4) \end{aligned}$$

D'altra parte, posto $\gamma_{2a} = E(a_t^4) - 3$, si ottiene

$$E(\varepsilon_t^4) = E(\sigma_t^4 a_t^4) = E(\sigma_t^4) E(a_t^4) = E(\sigma_t^4) (\gamma_{2a} + 3)$$

che implica

$$\begin{aligned} E(\sigma_t^4) &= E\left((w + \alpha \varepsilon_{t-1}^2)^2\right) = w^2 + \alpha^2 E(\varepsilon_{t-1}^4) + 2w\alpha E(\varepsilon_{t-1}^2) = \\ &= w^2 + \alpha^2 (\gamma_{2a} + 3) E(\sigma_t^4) + 2w\alpha \frac{w}{1-\alpha} \end{aligned}$$

da questa si ricava

$$E(\sigma_t^4) = \frac{w^2(1+\alpha)}{(1-\alpha)(1-\alpha^2(\gamma_{2a}+3))};$$

Sostituendo $E(\varepsilon_t^4)$ e $E(\sigma_t^4)$ in $\text{Var}(\eta_t)$, si ottiene

$$\begin{aligned} \text{Var}(\eta_t) &= (\gamma_{2a} + 3)E(\sigma_t^4) - E(\sigma_t^4) = E(\sigma_t^4)(\gamma_{2a} + 2) = \\ &= \frac{w^2(1+\alpha)(\gamma_{2a} + 2)}{(1-\alpha)(1-\alpha^2(\gamma_{2a} + 3))} \end{aligned}$$

e dato che deve essere $0 \leq \alpha < 1$, questa ha senso se è $0 \leq \alpha^2(\gamma_{2a} + 3) < 1$ ed è costante rispetto al tempo e quindi η_t è stazionario in varianza.

Per l'*autocovarianza* di η_t risulta

$$\begin{aligned} \text{Cov}(\eta_t, \eta_{t-1}) &= E(\eta_t \eta_{t-1}) = E\left((\varepsilon_t^2 - \sigma_t^2)(\varepsilon_{t-1}^2 - \sigma_{t-1}^2)\right) = \\ &= E\left((\varepsilon_t^2 - w - \alpha \varepsilon_{t-1}^2)(\varepsilon_{t-1}^2 - w - \alpha \varepsilon_{t-2}^2)\right) = \\ &= E\left(\varepsilon_t^2 \varepsilon_{t-1}^2 - w \varepsilon_t^2 - \alpha \varepsilon_t^2 \varepsilon_{t-2}^2 - w \varepsilon_{t-1}^2 + w^2 + w \alpha \varepsilon_{t-2}^2\right) + \\ &\quad + E\left(-\alpha \varepsilon_{t-1}^4 + w \alpha \varepsilon_{t-1}^2 + \alpha^2 \varepsilon_{t-1}^2 \varepsilon_{t-2}^2\right) \end{aligned}$$

per la stazionarietà di ε_t si ha

- $E(\varepsilon_t^2) = E(\varepsilon_{t-1}^2) = E(\varepsilon_{t-2}^2) = \sigma_\varepsilon^2$
- $E(\varepsilon_t^2 \varepsilon_{t-1}^2) = E(\varepsilon_{t-1}^2 \varepsilon_{t-2}^2) = E(\sigma_t^2 a_t^2 \varepsilon_{t-1}^2) =$

$$= E\left((w + \alpha \varepsilon_{t-1}^2) a_t^2 \varepsilon_{t-1}^2\right) = w\sigma_\varepsilon^2 + \alpha E(\varepsilon_t^4)$$
- $E(\varepsilon_t^2 \varepsilon_{t-2}^2) = E\left((w + \alpha \varepsilon_{t-1}^2) a_t^2 \varepsilon_{t-2}^2\right) = w(1 + \alpha)\sigma_\varepsilon^2 + \alpha^2 E(\varepsilon_t^4)$

Sostituendo queste espressioni in $\text{Cov}(\eta_t, \eta_{t-1})$ e tenendo conto della (4) si ricava il risultato cercato:

$$\gamma_\eta(1) = \text{Cov}(\eta_t, \eta_{t-1}) = 0$$

Con una dimostrazione simile è possibile provare che è $\gamma_\eta(h) = 0$ per qualsiasi $h > 0$. Questo dimostra che il processo stocastico η_t , sotto le condizioni sopra ricordate, è un white noise.

5.3 Identificazione e stima di un ARCH(s)

Il modello ARMA-ARCH definito nella (10) deve essere identificato e stimato utilizzando la serie storica (r_1, r_2, \dots, r_N) dei rendimenti osservata per un dato attivo finanziario. Questo ci permette di costruire previsioni ottimali sia per i livelli del fenomeno che per la relativa volatilità. Per l'identificazione della presenza di componenti ARMA si stima il *correlogramma* dei rendimenti con la relativa banda di confidenza. Esiste una struttura ARMA nei rendimenti se vi sono autocorrelazioni, in particolare qualcuna delle prime, che cadono al di fuori di tale banda. Per l'identificare la presenza di strutture ARCH si stima il *correlogramma* dei quadrati dei rendimenti r_t^2 con la relativa banda di confidenza. Esiste volatilità che può essere catturata con un ARCH se qualcuna di queste autocorrelazioni, ed in particolare le prime, cade al di fuori della banda di confidenza.

Esempio 2

Consideriamo la serie dei rendimenti dell'esempio simulato 1 e stimiamo le prime 12 autocorrelazioni di r_t e quella dei suoi quadrati. Il risultato è riportato nella figura 4. Come si può osservare dalla figura 4, e come atteso, dalle autocorrelazioni di r_t si desume che questa serie ha il comportamento di un white noise in cui non esiste alcuna struttura ARMA significativa nei livelli, mentre dal *correllogramma* di r_t^2 emerge la presenza di un chiaro effetto ARCH(1) che è necessario stimare.

Figura 4 – Prime 12 autocorrelazioni stimate di r_t e di r_t^2

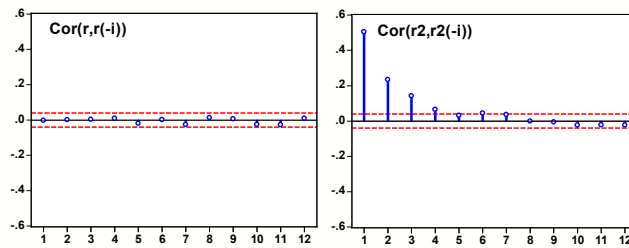
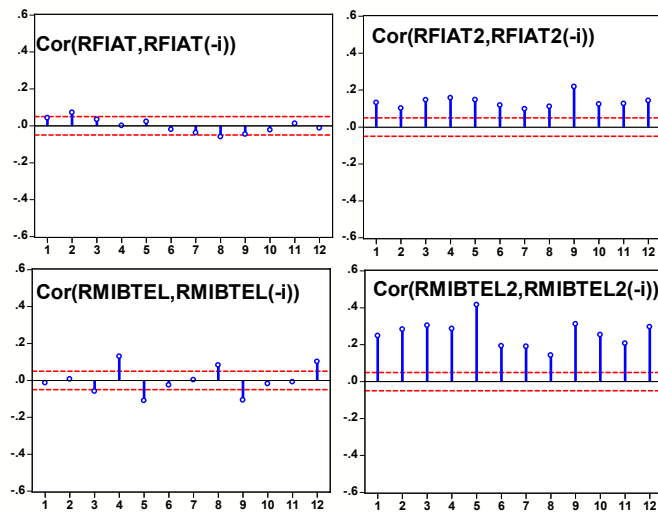


Figura 5 – Prime 12 autocorrelazioni stimate dei rendimenti di Fiat e Mibtel e dei loro quadrati



Consideriamo le serie dei rendimenti di Fiat e Mibtel e stimiamo le prime 12 autocorrelazioni e quelle dei loro quadrati. I risultati sono riportati nella

figura 5. Come si può osservare dalla figura 5 i rendimenti di Fiat sono praticamente quelli di un white noise, mentre in quelli di Mibtel potrebbe esistere una debole struttura di tipo MA. Viceversa, sia nei rendimenti Fiat che Mibtel è presente una marcata *volatilità* di tipo ARCH(s), questa è più accentuata in Mibtel che in Fiat.

□

Una volta stabilito che nel fenomeno esistono componenti ARMA e/o ARCH il modello va identificato e stimato simultaneamente con il metodo della massima verosimiglianza. Il modello (10) così stimato deve essere sottoposto a verifica:

- tutti i parametri, sia della parte ARMA che di quella ARCH, devono essere significativamente diversi da zero. Se qualcuno di questo non lo è bisogna eliminarlo dal modello ed effettuare una nuova stima;
- le autocorrelazioni dei residui ε_t devono cadere all'interno della banda di confidenza al 95% data da $(-2/\sqrt{N}; 2/\sqrt{N})$;
- le autocorrelazioni di a_t e quelle di a_t^2 devono cadere all'interno della banda di confidenza al 95% data da $(-2/\sqrt{N}; 2/\sqrt{N})$;
- deve risultare $w > 0$, $0 \leq \alpha_i < 1$, $\alpha_1 + \alpha_2 + \dots + \alpha_s < 1$;
- l'indice di curtosi di a_t deve essere non superiore a quello di ε_t ;
- è necessario costruire un test per verificare che nei residui a_t non sia rimasta ulteriore volatilità.

Se qualcuno di questi requisiti non è soddisfatto il modello ARMA-ARCH deve essere riformulato e ristimato. La procedura va ripetuta fino a che tutti questi requisiti sono sostanzialmente verificati.

La stima dei parametri del modello ARMA(p,q)-ARCH(s) riportato nella (10), sotto l'ipotesi che sia $a_t \stackrel{iid}{\sim} N(0, 1)$ può essere ottenuta simultaneamente con il metodo della massima verosimiglianza. Indicando con δ il vettore di tutti gli $m=p+q+s+1$ parametri la *funzione di verosimiglianza* diviene

$$L(\theta) = f(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_N | \delta) =$$

$$= f(\varepsilon_N | \mathfrak{F}_{N-1}) f(\varepsilon_{N-1} | \mathfrak{F}_{N-2}) \dots f(\varepsilon_{N-m} | \mathfrak{F}_{N-m-1}) f(a_1, a_2, \dots, a_m | \delta) =$$

$$\begin{aligned}
 &= \prod_{t=m+1}^N \left[\left(1/2\pi\sigma_t^2 \right)^{1/2} \exp \left\{ -\varepsilon_t^2 / 2\sigma_t^2 \right\} \right] f(a_1, a_2, \dots, a_m | \delta) \approx \\
 &\approx \prod_{t=m+1}^N \left[\left(1/2\pi\sigma_t^2 \right)^{1/2} \exp \left\{ -\varepsilon_t^2 / 2\sigma_t^2 \right\} \right]
 \end{aligned}$$

e la relativa log-verosimiglianza approssimata è

$$-2 \log L(\delta) = \sum_{t=m+1}^N \left[\log(2\pi) + \log(\sigma_t^2) + \varepsilon_t^2 / \sigma_t^2 \right] \tag{11}$$

dove

$$\varepsilon_t = r_t - (\phi_1 r_{t-1} + \phi_2 r_{t-2} + \dots + \phi_p r_{t-p}) + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}$$

$$\sigma_t^2 = w + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \dots + \alpha_s \alpha_{t-s}^2$$

La minimizzazione, con metodi iterativi, della (11) rispetto ai parametri incogniti permette di ottenere le stime di massima verosimiglianza. Se a_t invece che come una normale si distribuisce come una $t_{(k)}$ standardizzata, e precisamente $a_t \stackrel{iid}{\sim} ((k-2)/k)^{1/2} t_{(k)}$, può essere derivata una espressione, qui non riportata, analoga alla (11) e le relative stime ML possono essere ottenute in maniera analoga. Per far sì che in questo caso a_t possenga i primi quattro momenti si impone che sia $k > 4$. Ottenute le stime dei parametri, sia della struttura ARMA che ARCH, si sostituiscono nelle (10) e si derivano:

- i residui stimati $\hat{\varepsilon}_t = r_t - \hat{r}_t$ del modello ARMA che permettono di verificare la loro compatibilità con un white noise, ottenere una stima non condizionata della varianza di ε_t e derivare le previsioni dei livelli h tempi in avanti;
- una stima della volatilità $\hat{\sigma}_t^2$ che permette di derivarne le previsioni h tempi in avanti e calcolare indici sintetici che misurino la rischiosità dell'attivo finanziario;

- i residui standardizzati $\hat{a}_t = \hat{\varepsilon}_t / \hat{\sigma}_t$ che permettono di verificare, tramite la stima delle autocorrelazioni di \hat{a}_t e \hat{a}_t^2 ed il calcolo di indici sintetici e grafici, che il modello ha catturato tutta la *volatilità* dell'attivo. Quest'ultimo aspetto può essere confermato dall'uso di un ARCH test di cui trattiamo brevemente qui di seguito.

In alternativa si può utilizzare una stima di QMV, come mostreremo nel paragrafo 5.8

ARCH test

Consideriamo il modello di regressione

$$\hat{a}_t^2 = \alpha_0 + \alpha_1 \hat{a}_{t-1}^2 + \alpha_2 \hat{a}_{t-2}^2 + \dots + \alpha_m \hat{a}_{t-m}^2 + e_t, \quad m \geq 5, t=1, 2, \dots$$

stimiamo i suoi parametri con il metodo dei minimi quadrati ottenendo $(\hat{\alpha}_0, \dots, \hat{\alpha}_m)$ che permettono di derivare i residui stimati \hat{e}_t , calcoliamo la varianza campionaria di \hat{a}_t e quella di \hat{e}_t :

$$S_a^2 = \frac{1}{N} \sum_{t=2m+1}^N (\hat{a}_t - \bar{a})^2; \quad S_e^2 = \frac{1}{N} \sum_{t=2m+1}^N (\hat{e}_t - \bar{e})^2$$

e costruiamo la statistica test

$$F_c = \frac{(S_a^2 - S_e^2) / m}{S_e^2 / (N - 2m - 1)}$$

Sotto l'ipotesi $H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_m = 0$, F_c si distribuisce come una F di Fisher con $(m, N-m-1)$ gradi di libertà. Si rifiuta H_0 al livello α se risulta:

$$F_c > F_{(m; N-2m-1; \alpha)}$$

dove, al solito $F_{(m; N-2m-1; \alpha)}$ è il percentile della v.c. F di Fisher con $(m, N-2m-1)$ gradi di libertà che stacca alla sua destra una probabilità pari ad α . Nelle applicazioni ad attivi finanziari di solito N è molto grande in assoluto e molto

più grande di m e vale l'approssimazione

$$F_{(m; N-2m-1)} \approx \frac{1}{m} \chi_m^2$$

che viene utilizzata per ottenere la regione di rifiuto (regione critica) al livello α :

$$mF_c > \chi_{(m; \alpha)}^2$$

dove $\chi_{(m; \alpha)}^2$ è il percentile della v.c. chi quadrato che stacca alla sua destra una probabilità pari ad α .

Esempio 3

Riprendiamo l'esempio 1 in cui si è supposto r_t white noise, con volatilità data da un ARCH(1) ed $a_t \stackrel{iid}{\sim} N(0, 1)$. Simuliamo il modello per $B=500$ volte e stimiamo i due parametri (w, α) con il metodo della massima verosimiglianza sopra illustrato. I risultati della stima sono riportati nella tabella 1.

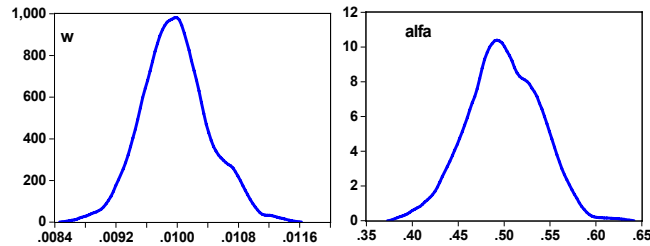
Tabella 1 – Indici descrittivi delle 500 stime di (w, α) con a_t normale

	$w=0.01$	$\alpha=0.5$
μ	0.009982	0.499386
Me	0.009963	0.497728
Max	0.011392	0.619560
Min	0.008697	0.395161
σ	0.000435	0.038308
γ_1	0.236255	-0.043695
γ_2	0.210908	-0.138998
JB	5.578065	0.561613
Prob	0.061481	0.755174
B	500	500

Nella figura 6 sono riportate le stime delle funzioni di densità delle due stime. Come si può osservare da tabella e figura le stime ottenute per i due parametri sono, sostanzialmente, non distorte, consistenti e distribuite normalmente. Le elaborazioni sono state rifatte utilizzando gli stessi parametri, identici valori iniziali, ma innovazioni distribuite come t di Student standardizzata con $k=5$

gradi di libertà e precisamente $a_t \stackrel{\text{iid}}{\sim} (3/5)^{1/2} t_{(5)}$.

Figura 6 – Funzione di densità delle stime dei due parametri con a_t normale

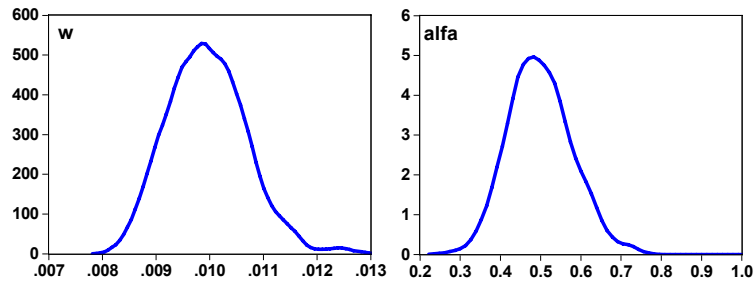


I risultati della simulazione sono riportati nella tabella 2 e nella figura 7. Possiamo notare che le stime sono ancora non distorte e consistenti ed asintoticamente normali, ma rispetto al caso precedente presentano una più elevata variabilità e l'esistenza di valori eccezionali nella coda destra come è anche testimoniato dall'alto valore dell'indice di asimmetria.

Tabella 2 – Indici descrittivi delle stime 500 di (w, α) con $a_t \sim t_{(5)}$ di Student

	$w=0.01$	$\alpha=0.5$
μ	0.009970	0.500894
Me	0.009948	0.494228
Max	0.014579	1.467968
Min	0.008238	0.266788
σ	0.000781	0.089285
γ_1	1.050956	2.676166
γ_2	6.942616	29.40810
JB	415.8803	15125.73
Prob	0.000000	0.000000
B	500	500

Qui di seguito riportiamo il listato della simulazione e stima dell'ARCH(1) con residui dati da una $t_{(5)}$ standardizzata. Naturalmente, con gli aggiustamenti del caso, è possibile utilizzare il programma per simulare e/o stimare un ARCH(s) con $s \geq 1$.

Figura 7 – Funzione di densità delle stime dei due parametri con $a_t \sim t_{(5)}$ di Student

Listato per la simulazione, con $B=500$, e stima di un ARCH(1) con innovazioni $t_{(5)}$.

'Modello ARCH1

```

for lj=1 to 500
    smpl 1 6000
    genr r=0.03
    genr vr=0.0002
    genr res=@rtdist(5)
    for li=2 to 6000
        smpl li li
        genr vr=.01+.5*(r(-1))^2
        genr r=(vr^.5)*((3/5)^.5)*res
    next
    smpl 3501 6000
    equation eq_vrt.arch(1,0) r
    smpl lj lj
    genr c1t=c(1)
    genr c2t=c(2)
next
smpl 1 500

```

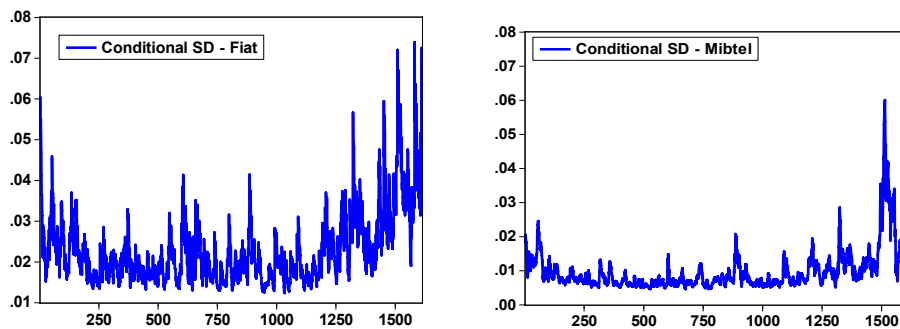
Consideriamo la serie dei rendimenti degli attivi Fiat e Mibtel, più volte presentata, ed adattiamole un modello della classe ARMA-ARCH. Abbiamo già visto, nell'esempio 8 del capitolo IV, che per i livelli il modello ARMA(p,q) che meglio adatta i rendimenti di questi due attivi si ha con $p=0$ e $q=2$ per Fiat e con $p=0$ e $q=5$ per Mibtel. Per la *volatilità*, tenendo conto delle autocorrelazioni dei quadrati dei rendimenti riportati nella figura 5 che ne evidenziano un comportamento molto complesso, consideriamo un ARCH(s) con $s \leq 9$. Dove $s = 9$ è il valore massimo che EVIEWS 6 riesce a stimare per un ARCH. I

risultati della stima sono riportati nella tabella 3. Da questa tabella si può osservare che il parametro θ_1 per Fiat non è significativamente diverso da zero e deve essere eliminato, mentre per Mibtel nessuno dei cinque parametri θ risulta significativamente diverso da zero al 5%. Sia per Fiat che per Mibtel non è significativo il parametro α_7 nell'equazione riferita alla volatilità e dovrebbe essere eliminato, ma il pacchetto EViews 6 non permette di effettuare queste operazioni. Infine osserviamo che per tutte e due gli attivi è

$$w > 0; \quad 0 \leq \alpha_i < 1, i=1, 2, \dots, 9; \quad \sum_{i=1}^9 \alpha_i < 1$$

e quindi sono soddisfatte le condizioni sui parametri del modello ARCH. Ripuliamo il modello eliminando dai livelli i parametri MA che sono risultati non significativamente diversi da zero. Le stime finali sono riportati nella tabella 4. Questa volta tutti i parametri sono significativamente diversi da zero al 5% con l'eccezione di α_7 per entrambi gli attivi. Il valore dei parametri per la volatilità è, praticamente, uguale a quello della tabella 3 ed anche in questo caso sono soddisfatte le condizioni sui parametri dell'ARCH. Il modello così ottenuto può essere considerato "ottimale" ed essere utilizzato per prevedere il rischio futuro.

Figura 8 – *Stima di σ_t per le serie degli attivi Fiat e Mibtel*



Nella figura 8 sono riportati i grafici delle stime della volatilità σ_t per Fiat e Mibtel ottenuti dal modello MA-ARCH adattato alle due serie. Da questo grafico si evince che:

Tabella 3 – Stima dei parametri del modello ARMA-ARCH per Fiat e Mibtel

	Fiat			Mibtel		
	Coeff.	SE	Prob.	Coeff.	SE	Prob.
ϕ_0	0.0011	0.0005	.038	0.0006	0.0002	.002
θ_1	-0.0056	0.0278	.842	0.0305	0.0268	.255
θ_2	-0.0601	0.0267	.023	-0.0325	0.0284	.253
θ_3				0.0190	0.0262	.468
θ_4				-0.0291	0.0259	.260
θ_5				0.3300	0.0249	.186
w	0.00015	0.0000	.000	0.00002	0.0000	.000
α_1	0.1796	0.0348	.000	0.0504	0.0159	.002
α_2	0.0808	0.0254	.002	0.1562	0.0317	.000
α_3	0.0860	0.0297	.004	0.0728	0.0229	.002
α_4	0.1236	0.0282	.000	0.1339	0.0205	.000
α_5	0.1111	0.0238	.000	0.1678	0.0237	.000
α_6	0.0560	0.0283	.047	0.1256	0.0273	.000
α_7	0.0222	0.0187	.234	0.0409	0.0210	.052
α_8	0.0570	0.0249	.020	0.0479	0.0248	.019
α_9	0.0693	0.0217	.001	0.0549	0.0228	.016
$\sum \alpha_i$	0.7856			0.8604		

1. le quotazioni Fiat sono molto più volatili di quelle Mibtel. Questo è un risultato atteso dato che Mibtel è un indice medio ponderato che misura l'evoluzione media del mercato borsistico italiano;
2. entrambe le serie mostrano una volatilità non costante nel tempo ed in particolare molto accentuata nel periodo finale che, ricordiamo, è riferito all'inverno – primavera 2009 e quindi in piena crisi finanziaria.

Tabella 4 – *Stima finale dei parametri del modello ARMA-ARCH per Fiat e Mibtel*

	Fiat			Mibtel		
	Coeff	SE	Prob.	Coeff	SE	Prob.
ϕ_0	0.0011	0.0005	.033	0.0006	0.0002	.002
θ_2	-0.0623	0.0267	.019			
w	0.00015	0.00001	.000	0.00002	0.0000	.000
α_1	0.1804	0.0344	.000	0.0506	0.0153	.001
α_2	0.0799	0.0252	.002	0.1564	0.0310	.000
α_3	0.0857	0.0295	.004	0.0726	0.0219	.001
α_4	0.1244	0.0283	.000	0.1346	0.0203	.000
α_5	0.1106	0.0233	.000	0.1681	0.0228	.000
α_6	0.0564	0.0283	.046	0.1295	0.0271	.000
α_7	0.0222	0.0187	.234	0.0406	0.0209	.053
α_8	0.0567	0.0248	.022	0.0577	0.0242	.017
α_9	0.0688	0.0216	.001	0.0546	0.0223	.014
$\sum \alpha_i$	0.7851			0.8647		

□

5.4 Il modello GARCH

I modelli ARCH richiedono, come constatato nelle tabelle 3 e 4 dell'esempio 3, un numero di parametri piuttosto elevato per poter rappresentare in modo adeguato la volatilità di un dato attivo finanziario. Nell'ultimo esempio del paragrafo precedente abbiamo visto che i parametri necessari devono essere almeno nove. Per superare questo inconveniente Bollerslev (1986), utilizzando una logica simile a quella che ha permesso di passare dai modelli AR a quelli ARMA, ha proposto di generalizzare i modelli ARCH nei GARCH (*Generalized ARCH*). Più in dettaglio, i rendimenti r_t di un dato attivo hanno una struttura ARMA(p, q) nei livelli e GARCH(s, v) nella volatilità se sono soddisfatte le seguenti tre equazioni:

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) r_t = \phi_0 + (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) \varepsilon_t$$

$$\varepsilon_t = \sigma_t a_t$$

$$\sigma_t^2 = w + \alpha_1 \varepsilon_{t-1}^2 + \dots + \alpha_s \varepsilon_{t-s}^2 + \beta_1 \sigma_{t-1}^2 + \dots + \beta_v \sigma_{t-v}^2 \quad (12)$$

Della struttura dei livelli dei rendimenti, che si solito sono ben rappresentati da qualche semplice modello MA, ne abbiamo già parlato nelle pagine precedenti. Qui segnaliamo che, nelle pratiche applicazioni, anche la struttura del GARCH per la volatilità è molto semplice. Infatti, di solito, basta il GARCH(1, 1)

$$\sigma_t^2 = w + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \quad (13)$$

per ben rappresentare questo fenomeno. Con questa constatazione pratica, analizziamo qui di seguito con un certo dettaglio il modello (13). Se in questo si pone $\sigma_t^2 = \varepsilon_t^2 - \eta_t$ si ottiene

$$\varepsilon_t^2 = w + (\alpha + \beta) \varepsilon_{t-1}^2 + \eta_t - \beta \eta_{t-1} \quad (14)$$

che è un ARMA(1,1) sugli ε_t^2 . Come per i modelli ARCH anche in questo caso il processo η_t non è iid ma più semplicemente un white noise.

Se si tiene conto che è $\varepsilon_t = \sigma_t a_t$, la (13) si può scrivere

$$\sigma_t^2 = w + \alpha \sigma_{t-1}^2 a_{t-1}^2 + \beta \sigma_{t-1}^2$$

ed utilizzando le stesse argomentazioni usate per derivare i momenti del modello ARCH, si ha

$$\sigma_\varepsilon^2 = E(\sigma_t^2) = \frac{w}{1 - (\alpha + \beta)} \quad (15)$$

e, condizioni sufficienti, perché il GARCH(1, 1) abbia senso deve essere

$$w > 0; \alpha \geq 0; \beta \geq 0; 0 \leq \alpha + \beta < 1$$

Supponiamo che a_t ammetta i primi quattro momenti e poniamo

$$k_{2a} = E(a_t^4) = \gamma_{2a} + 3, \text{ si ottiene}$$

$$E(\varepsilon_t^4)E(\sigma_t^4 a_t^4) = E(a_t^4)E(\sigma_t^4) = (\gamma_{2a} + 3)E(\sigma_t^4)$$

per la (13) e la stazionarietà dei processi coinvolti risulta

$$\begin{aligned} E(\sigma_t^4) &= E\left((w + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2)^2\right) = \\ &= w^2 + \alpha^2(\gamma_{2a} + 3)E(\sigma_{t-1}^4) + \beta^2E(\sigma_{t-1}^4) + 2w\alpha\sigma_\varepsilon^2 + 2w\beta\sigma_\varepsilon^2 + 2\alpha\beta E(\sigma_{t-1}^4) = \\ &= w^2 + \alpha^2(\gamma_{2a} + 3)E(\sigma_t^4) + \beta^2E(\sigma_t^4) + 2w\alpha\sigma_\varepsilon^2 + 2w\beta\sigma_\varepsilon^2 + 2\alpha\beta E(\sigma_t^4) \end{aligned}$$

da cui

$$\begin{aligned} E(\sigma_t^4) &= \frac{w^2 + 2w\sigma_\varepsilon^2(\alpha + \beta)}{1 - \alpha^2(\gamma_{2a} + 2) - (\alpha + \beta)^2} = \\ &= \frac{w^2(1 + \alpha + \beta)}{(1 - \alpha - \beta)\left(1 - \alpha^2(\gamma_{2a} + 2) - (\alpha + \beta)^2\right)} \end{aligned}$$

ed infine

$$\begin{aligned} \gamma_{2\varepsilon} &= \frac{E(\varepsilon_t^4)}{\left(E(\varepsilon_t^2)\right)^2} - 3 = \frac{(\gamma_{2a} + 3)w^2(1 + \alpha + \beta)}{(1 - \alpha - \beta)\left(1 - \alpha^2(\gamma_{2a} + 2) - (\alpha + \beta)^2\right)} \frac{(1 - \alpha - \beta)}{w^2} - 3 = \\ &= \frac{(\gamma_{2a} + 3)\left(1 - (\alpha + \beta)^2\right)}{1 - \alpha^2(\gamma_{2a} + 2) - (\alpha + \beta)^2} - 3 \end{aligned} \quad (16)$$

che nel caso di normalità $a_t \stackrel{\text{iid}}{\sim} N(0, 1)$ si riduce a

$$\gamma_{2\varepsilon}^{(g)} = \frac{6\alpha}{1 - 2\alpha^2(\alpha + \beta)^2} \quad (17)$$

In definitiva, nel caso di normalità, possiamo osservare che:

- se è $1 - 2\alpha^2(\alpha + \beta)^2 > 0$ il processo ε_t possiede comunque eccesso di *curtosi*;
- se è $\alpha = 0$ si ha anche $\gamma_{2\varepsilon}^{(g)} = 0$ ed in tal caso ε_t non ha eccesso di *curtosi*;
- se è $\beta = 0$ le (16) e (17) si riducono, come atteso, a quelle dell' ARCH(1).

In generale Bai et al. (2001) dimostrano che la (16), tenendo conto della (17), si può scrivere

$$\gamma_{2\varepsilon} = \frac{\gamma_{2a} + \gamma_{2\varepsilon}^{(g)} + \frac{5}{6}\gamma_{2a}\gamma_{2\varepsilon}^{(g)}}{1 - \frac{1}{6}\gamma_{2a}\gamma_{2\varepsilon}^{(g)}} \quad (18)$$

questa ci permette, conoscendo la distribuzione di a_t , di derivare rapidamente l'indice di *curtosi* per ε_t . In particolare se si suppone che a_t sia una $t_{(k)}$, con $k > 4$, in tal caso sappiamo che è $\gamma_{2a} = 6/(k-4)$ e la (18) diventa

$$\gamma_{2\varepsilon}^{(t)} = \frac{6 + (k+1)\gamma_{2\varepsilon}^{(g)}}{(k-4) - \gamma_{2\varepsilon}^{(g)}} = \frac{6(1 - 2\alpha^2 - (\alpha + \beta)^2 + \alpha^2(k+1))}{(k-4)(1 - 2\alpha^2 - (\alpha + \beta)^2) - 6\alpha^2}$$

Esempio 4

Applichiamo un modello della classe MA-GARCH alle quotazioni giornaliere e settimanali degli attivi Campari, Fiat e Banca Intesa S. Paolo osservate dalla prima settimana di gennaio 2003, all'ultima settimana di febbraio 2009. Le stime saranno effettuate sia sotto l'ipotesi che le innovazioni a_t siano distribuite normalmente, sia che siano $t_{(k)}$ standardizzate. Nella tabella 5 riportiamo le stime del modello MA-GARCH sotto l'ipotesi di innovazioni normali e per dati sia giornalieri che settimanali. Come si può osservare dalla tabella 5, per le rilevazioni giornaliere (in grassetto) il modello può essere accettato per gli attivi Campari e Fiat, mentre va rifiutato per Banca Intesa S. Paolo dato che non è rispettata la condizione $\alpha + \beta < 1$. Per le rilevazioni settimanali il modello stimato va rifiutato per tutti e tre gli attivi considerati dato che una o più delle condizioni che i parametri GARCH devono possedere non sono rispettate: w è negativo o non significativamente diverso da zero, risulta $\alpha + \beta \geq 1$.

Tabella 5 – *Stime dei modelli ARMA-GARCH per gli attivi di Campari, Fiat e B. Intesa S. Paolo per dati giornalieri e settimanali. Ipotesi di normalità*

		ϕ_0	θ_1	θ_2	w	α	β	$\alpha + \beta$
Campari giornal. Settim.	Stima Prob.		0.113 (.000)	0.052 (.046)	0.000002 (.016)	0.039 (.000)	0.956 (.000)	0.995
	Stima Prob.	0.005 (.000)	0.106 (.036)		-0.00002 (.000)	0.021 (.004)	1.007 (.000)	1.028
Fiat giornal. Settim.	Stima Prob.				0.00009 (.000)	0.086 (.000)	0.899 (.000)	0.985
	Stima Prob.				0.00002 (.588)	0.085 (.003)	0.916 (.000)	1.001
Int. San P giornal. settim.	Stima Prob.				0.000004 (.000)	0.119 (.000)	0.881 (.000)	1.000
	Stima Prob.	0.005 (.022)			0.000006 (.536)	0.103 (.000)	0.919 (.000)	1.022

Tabella 6 – *Stime dei modelli ARMA-GARCH per gli attivi di Campari, Fiat e B. Intesa S. Paolo per dati giornalieri e settimanali con innovazioni $t_{(k)}$*

		ϕ_0	θ_1	θ_2	k	w	α	β	$\alpha + \beta$
Campari giornal. settim.	Stima Prob.		0.134 (.000)	0.053 (.031)	5.0 (.000)	0.000003 (.000)	0.047 (.000)	0.943 (.000)	0.990
	Stima Prob.	0.003 (.022)	0.150 (.002)		5.6 (.013)	-0.00001 (.008)	.010 (.482)	1.012 (.000)	1.022
Fiat giornal. settim.	Stima Prob.				6.3 (.000)	0.00001 (.008)	0.098 (.000)	0.890 (.000)	0.988
	Stima Prob.				24.6 (.519)	0.00002 (.611)	0.082 (.008)	0.918 (.000)	1.000
Int. S. Paolo giornal. settim.	Stima Prob.				5 (.000)	0.000008 (.001)	0.117 (.000)	0.868 (.000)	0.985
	Stima Prob.				3.7 (.000)	0.000027 (.338)	0.084 (.036)	0.915 (.000)	0.999

Le stesse stime sono state fatte sotto l'ipotesi che le innovazioni siano $t_{(k)}$ standardizzate, in questo caso è necessario stimare anche i gradi di libertà rappresentati dal parametro k. I risultati della stima sono riportati nella tabella 6. Anche in questo caso i modelli adattati ai dati giornalieri, in grassetto,

possono essere accettati per le tre serie considerate, mentre questo non è possibile per i dati settimanali. Osserviamo infine che i risultati ottenuti sono molto simili a quelli del caso normale.

□

5.5 Varianti del modello GARCH

Nel paragrafo precedente abbiamo illustrato le caratteristiche più rilevanti dei modelli ARMA-GARCH per l'analisi dei livelli e della *volatilità* dei rendimenti di attivi. Questa classe di modelli si è mostrata abbastanza flessibile in molte applicazioni pratiche, ma non nella loro totalità. Si è pertanto reso necessario elaborarne varianti che rendessero il modello più flessibile nelle pratiche applicazioni. Qui di seguito presentiamo alcune di queste.

(A) MODELLO IGARCH

Nelle tabelle 5 e 6 abbiamo visto che in molte applicazioni, ed in particolare per i dati con cadenza settimanale, la condizione $\alpha + \beta < 1$ non è rispettata, ma è $\alpha + \beta \geq 1$ ed il classico modello GARCH perde di significato. Per poter superare questo inconveniente è necessario imporre nella stima che sia contemporaneamente $w = 0$ ed $\alpha + \beta = 1$. Il relativo modello, a somiglianza degli ARIMA quando hanno una radice AR pari ad uno, prende il nome di IGARCH (*Integrated GARCH*). La quantità $\alpha + \beta$ viene anche chiamata *persistenza* della volatilità, la volatilità passata dell'attivo ha un impatto persistente nel presente:

$$r_t = \mu_t + \varepsilon_t; \quad \varepsilon_t = \sigma_t a_t; \quad a_t \stackrel{\text{iid}}{\sim} (0, 1); \quad \sigma_t^2 = w + (1 - \beta)\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2 \quad (19)$$

In questo caso la varianza di ε_t non è definita a meno che non sia $w=0$ ed è quello che viene fatto nelle pratiche applicazioni. Nel lungo periodo la volatilità di un attivo ben difficilmente si comporterà come l'IGARCH, ma può essere osservato per brevi periodi per slittamenti occasionali della volatilità. Per esempio questo comportamento con $w=0$ si può avere quando si analizzano dati relativi a periodi di normale evoluzione delle borse seguiti o preceduti da periodi di gravi crisi finanziarie e quindi di particolare accentuata *volatilità*.

La (19) si può anche scrivere

$$\varepsilon_t^2 = w + \varepsilon_{t-1}^2 + \eta_t - \beta\eta_{t-1}$$

che, per quanto visto nel Capitolo III, evolve lungo una retta a cui si

sovrappone un errore legato al processo di *Wiener*, più precisamente risulta:

$$\varepsilon_t^2 = \varepsilon_0^2 + wt + \sum_{j=1}^t (\eta_j - \beta \eta_{j-1})$$

che implica

$$\text{Var}(\varepsilon_t) = E(\varepsilon_t^2) = \varepsilon_0^2 + wt$$

e mostra come la *volatilità* dell'attivo sia non stazionaria ed evolva come una retta crescente. Inoltre, la volatilità risulta fortemente sensibile alla condizione iniziale ε_0^2 . Se, come avviene in pratica, si pone $w = 0$, la volatilità coincide con il valore iniziale. Le considerazioni sopra esposte spiegano perché i modelli IGARCH possono avere senso solo nel breve periodo e con $w=0$.

Esempio 5

Agli attivi Campari, Fiat, Banca Intesa S. Paolo, osservati con cadenza sia giornaliera che settimanale adattiamo un modello MA-IGARCH.

Tabella 7 – *Stime dell'MA-IGARCH per gli attivi di Campari, Fiat e Banca Intesa S. Paolo per dati giornalieri e settimanali con residui normali*

		ϕ_0	θ_1	θ_2	θ_5	β	$1-\beta$	AIC
Campari giornal. settim.	Stima Prob.	0.0006 (.025)	0.106 (.000)	0.058 (.012)		0.029 (.000)	0.971 (.000)	-5.499
	Stima Prob.	0.0031 (.006)	0.113 (.023)			0.052 (.004)	0.948 (.000)	-3.992
Fiat giornal. settim.	Stima Prob.	0.0009 (.046)		-0.047 (.043)		0.053 (.000)	0.947 (.000)	-4.818
	Stima Prob.				-0.108 (.040)	0.069 (.003)	0.931 (.000)	-3.201
Int. S. Paolo giornal. settim.	Stima Prob.	.0008 (.008)				0.076 (.000)	0.924 (.000)	-5.280
	Stima Prob.	.0050 (.000)				0.070 (.000)	0.930 (.000)	-3.626

Le stime ottenute sono riportate nella tabella 7. Nella figura 9 sono riportate le prime 12 autocorrelazioni dei residui standardizzati e dei residui standardizzati al quadrato dei tre attivi con le relative bande di confidenza al 95%. Dalla tabella 7 e dalla figura 9 si può osservare come il modello MA-IGARCH è adeguato per descrivere le serie dei dati giornalieri di Campari e Banca Intesa S.

Paolo, ma non lo è per Fiat a cui si adatta meglio il modello MA-GARCH. Dalla tabella 7 e dalla figura 10 si può concludere che il modello MA-IGARCH è soddisfacentemente adeguato per tutte e tre le serie settimanali degli attivi considerati.

Figura 9 – Autocorrelazioni dei residui e dei residui al quadrato per Campari, Fiat e Banca Intesa S. Paolo. Dati giornalieri e residui normali

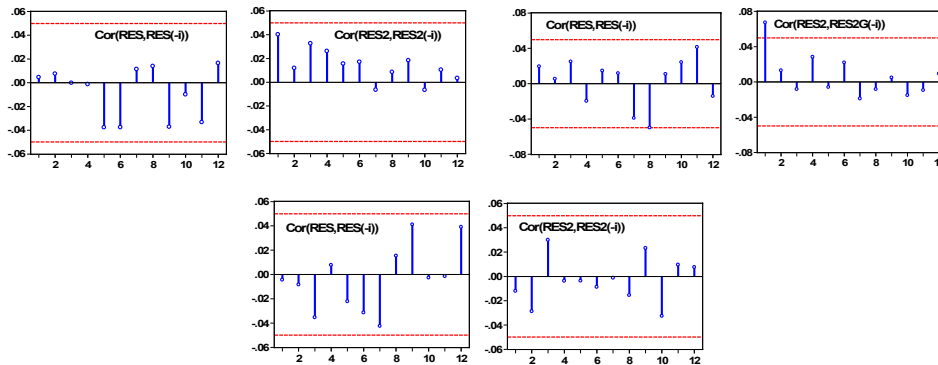
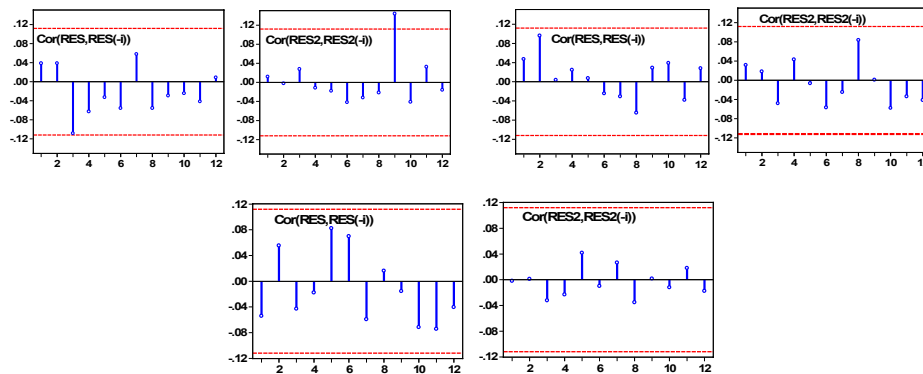


Figura 10 – Autocorrelazioni dei residui e dei residui al quadrato per Campari, Fiat ed Intesa S. Paolo. Dati settimanali e residui normali



□

(B) MODELLO GARCH-M

In alcuni casi il livello dei rendimenti è una qualche funzione, di solito monotona, della sua volatilità. Se si verifica una situazione di questo tipo vuol dire che i livelli dei rendimenti risultano prevedibili ed esiste una possibilità di *premio al rischio*. Questa considerazione e l'evidenza empirica ci portano ad escludere che, in generale, i rendimenti di attivi finanziari in un mercato abbastanza efficiente e trasparente possano essere ben spiegati da un modello GARCH-M se non in situazioni molto particolari e di breve periodo.

Il più semplice modello utilizzato è il GARCH(1, 1)-M dato da:

$$r_t = \mu_t + c g(\sigma_t^2) + \varepsilon_t; \quad \varepsilon_t = \sigma_t a_t; \quad a_t \stackrel{iid}{\sim} (0, 1); \quad \sigma_t^2 = w + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2$$

Di solito la funzione $g(\cdot)$ è monotona ed i casi più utilizzati sono: $g(\sigma_t^2) = \sigma_t^2$; $g(\sigma_t^2) = \sigma_t$; $g(\sigma_t^2) = \log(\sigma_t^2)$. Naturalmente, se nella stima del modello la costante c non è significativamente diversa da zero vuol dire che non vi è premio al rischio e bisogna riformulare il modello e ristimarlo.

Esempio 6

Il modello MA-IGARCH-M è stato adattato ai rendimenti delle quotazioni sia giornaliere che settimanali del titolo Campari già incontrato in precedenti esempi. I risultati della stima sono riportati nella tabella 8.

Tabella 8 – *Modello MA-IGARCH-M per i rendimenti giornalieri e settimanali di Campari*

Gionr.	ϕ_0	θ_1	θ_2	σ_t	β	$1-\beta$	AIC
		0.106 (.000)	0.058 (.013)	0.034 (.077)	0.029 (.000)	0.971 (.000)	-5.499
	ϕ_0	θ_1	θ_2	σ_t^2	β	$1-\beta$	AIC
	.0013 (.017)	0.107 (.000)	0.059 (.011)	-3.395 (.158)	0.029 (.000)	0.971 (.000)	-5.499
	ϕ_0	θ_1	θ_2	$\log(\sigma_t^2)$	β	$1-\beta$	AIC
		0.106 (.000)	0.058 (.012)	-0.00007 (.020)	0.029 (.000)	0.971 (.000)	-5.500
Settim.	ϕ_0	θ_1	θ_2	σ_t	β	$1-\beta$	AIC
		0.109 (.029)		0.099 (.021)	0.052 (.000)	0.948 (.000)	-3.988
	ϕ_0	θ_1	θ_2	σ_t^2	β	$1-\beta$	AIC
	.0047 (.025)	0.116 (.022)		-2.056 (.384)	0.051 (.000)	0.949 (.000)	-3.987
	ϕ_0	θ_1	θ_2	$\log(\sigma_t^2)$	β	$1-\beta$	AIC
		0.114 (.020)		-0.0004 (.005)	0.052 (.000)	0.948 (.000)	-3.992

Come si può notare dalla tabella 8 nei rendimenti di Campari esiste premio al rischio nel periodo analizzato (gennaio 2003, febbraio 2009) se si utilizza in entrambi i casi la funzione $g(\sigma_t^2) = \log(\sigma_t^2)$. In entrambi i casi, come atteso, fra volatilità e rendimenti esiste un legame inverso: l'aumento della volatilità implica una riduzione dei rendimenti anche se l'impatto, pur risultando significativamente diverso da zero, è molto piccolo: 7 su 100.000 per i dati giornalieri e 4 su 10.000 per quelli settimanali.

□

(C) MODELLO CGARCH

Il modello CGARCH (Component GARCH), raramente utilizzato nelle pratiche applicazioni, descrive la volatilità di un attivo tramite due equazioni:

$$\begin{cases} \sigma_t^2 = m_t + \alpha(\varepsilon_{t-1}^2 - m_{t-1}) + \beta(\sigma_{t-1}^2 - m_{t-1}) \\ m_t = w + \gamma(m_{t-1} - w) + \delta(\varepsilon_{t-1}^2 - \sigma_{t-1}^2) \end{cases}$$

La prima è detta *componente transitoria* della volatilità, descrive l'evoluzione di breve periodo di questa e converge al livello di lungo periodo della volatilità, m_t , con la potenza della persistenza $(\alpha + \beta)$. La seconda equazione descrive il comportamento di lungo periodo della volatilità, è detta *componente permanente* e converge a w con la potenza di γ .

Di solito è $0.99 \leq \gamma < 1$ e la componente m_t converge alla costante w molto lentamente. Sostituendo la seconda equazione nella prima due volte si ottiene:

$$\begin{aligned} \sigma_t^2 = & \left(1 - \gamma^2 - (\alpha + \beta)(1 - \gamma)\right)w + (\alpha + \delta)\varepsilon_{t-1}^2 - \delta(\gamma - (\alpha + \beta))\varepsilon_{t-2}^2 + \\ & + (\beta - \delta)\sigma_{t-1}^2 - \delta(\gamma - (\alpha + \beta))\sigma_{t-2}^2 + \gamma(\gamma - (\alpha + \beta))m_{t-2} \end{aligned}$$

e dato che $\gamma(\gamma - (\alpha + \beta))$ è, di solito, molto piccolo il CGARCH si riduce, sostanzialmente, ad un caso particolare di GARCH(2,2). Questo risultato giustifica perché questo modello ha scarso utilizzo nella pratica. Visto che il CGARCH è, praticamente, un particolare GARCH(2,2) le sue proprietà statistiche, la stima e le previsioni seguono le regole generali dei GARCH.

Esempio 7

Ai rendimenti di Campari, Fiat e Banca Intesa S. Paolo, sia giornalieri che

settimanali, adattiamo un modello della classe CGARCH. Nella tabella 9 riportiamo i risultati della stima e per nessuna delle sei serie il modello MA-CGARCH-M risulta adeguato.

Tabella 9 – *Stima del modello MA-CGARCH-M per i rendimenti Campari, Fiat e Banca Intesa S. Paolo. Dati giornalieri e settimanali*

Campari	θ_1	$\log(\sigma_t^2)$	w	γ	δ	α	β	AIC
Giorn.	0.119 (.000)	-0.00008 (.021)	0.0015 (.872)	1.000 (.000)	0.024 (.000)	0.057 (.000)	0.074 (.000)	-5.507
Settim.			0.002 (.379)	0.988 (.000)	0.090	0.103 (.000)	-0.948 (.000)	-4.048
Fiat	θ_1	$\log(\sigma_t^2)$	w	γ	δ	α	β	AIC
Giorn.	0.993 (.000)		0.0008 (.027)	0.993 (.000)	0.071 (.000)	0.075 (.014)	-0.089 (.754)	-4.839
Settim.	0.973 (.000)		0.0058 (.691)	0.994 (.000)	0.090 (.019)	0.043 (.535)	0.045 (.977)	-3.184
Int. S. P.	θ_1	$\log(\sigma_t^2)$	w	γ	δ	α	β	AIC
Giorn.	0.986 (.000)	0.000002 (.031)	0.00093 (.069)	0.993 (.000)	0.116 (.000)	-0.134 (.000)	0.073 (.000)	-5.278
Settim.	0.990 (.000)	0.000013 (.005)	0.037 (.865)	0.999 (.000)	0.091 (.000)	-0.054 (.000)	-0.489 (.000)	-3.663

□

5.6 La previsione della volatilità

Il modello ARMA-GARCH adattato ai rendimenti di un attivo può essere utilizzato per ottenere simultaneamente *previsioni* ottimali dei livelli e della volatilità. Della previsione dei livelli ne abbiamo trattato nel capitolo IV, qui di seguito, utilizzando una procedura simile, discutiamo di quella della volatilità. Per non rendere l'esposizione eccessivamente macchinosa analizziamo e costruiamo le previsioni del solo modello GARCH(1,1) che è quello generalmente utilizzato nelle pratiche applicazioni. Supponiamo per il momento di conoscere i tre parametri (w, α, β) . Date le informazioni disponibili (r_1, r_2, \dots, r_N) e quindi

$$\mathfrak{F}_N = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_N; \sigma_1^2, \sigma_2^2, \dots, \sigma_N^2)$$

dalla (13) si ha immediatamente che la previsione di σ_{N+1}^2 è data da

$$\sigma_N^2(1) = w + \alpha \varepsilon_N^2 + \beta \sigma_N^2 \tag{20}$$

Per la previsione di σ_{N+2}^2 , ricordando che è $\varepsilon_t = \sigma_t a_t$, si ottiene

$$\begin{aligned} \sigma_{N+2}^2 &= w + \alpha \varepsilon_{n+1}^2 + \beta \sigma_{N+1}^2 = w + \alpha \sigma_{N+1}^2 a_{N+1}^2 + \beta \sigma_{N+1}^2 = \\ &= w + (\alpha a_{N+1}^2 + \beta) \sigma_{N+1}^2 \end{aligned}$$

e la relativa previsione è data da

$$\begin{aligned} \sigma_N^2(2) &= E(\sigma_{N+2}^2 | \mathfrak{F}_N) = w + E((\alpha a_{N+1}^2 + \beta) \sigma_{N+1}^2 | \mathfrak{F}_N) = \\ &= w + (\alpha + \beta) E(\sigma_{N+1}^2 | \mathfrak{F}_N) = w + (\alpha + \beta) \sigma_N^2(1) \end{aligned}$$

In definitiva, iterando la procedura h volte si ottiene che la previsione di σ_{N+h}^2 è

$$\sigma_N^2(h) = \begin{cases} w + \alpha \varepsilon_N^2 + \beta \sigma_N^2 & \text{per } h = 1 \\ w + (\alpha + \beta) \sigma_N^2(h-1) & \text{per } h > 1 \end{cases} \tag{21}$$

La (21), sostituendo iterativamente $\sigma_N^2(h-1)$, può essere scritta in funzione della (20):

$$\sigma_N^2(h) = w + (\alpha + \beta)w + (\alpha + \beta)^2 w + \dots + (\alpha + \beta)^{h-2} w + (\alpha + \beta)^{h-1} \sigma_N^2(1) =$$

$$\begin{aligned}
&= w \frac{1 - (\alpha + \beta)^{h-1}}{1 - (\alpha + \beta)} + (\alpha + \beta)^{h-1} \sigma_N^2(1) = \sigma_\varepsilon^2 \left(1 - (\alpha + \beta)^{h-1} \right) + (\alpha + \beta)^{h-1} \sigma_N^2(1) = \\
&= \sigma_\varepsilon^2 + (\alpha + \beta)^{h-1} \left(\sigma_N^2(1) - \sigma_\varepsilon^2 \right) \quad (22)
\end{aligned}$$

e se è $(\alpha + \beta) < 1$, per $h \rightarrow \infty$, si ottiene

$$\sigma_N^2(h) \rightarrow \sigma_\varepsilon^2$$

e la previsione della volatilità converge alla varianza non condizionata σ_ε^2 se questa esiste.

Nel caso del modello IGARCH(1,1) abbiamo visto che deve essere $w = 0$ ed $(\alpha + \beta) = 1$ e per ogni h otteniamo

$$\sigma_N^2(h) = \sigma_N^2(1)$$

La previsione dei modelli GARCH-M si ottiene allo stesso modo, ma in questo caso la previsione della volatilità ha un impatto diretto anche su quella dei livelli:

$$r_N(h) = \mu_N(h) + c g(\sigma_N^2(h))$$

In pratica i parametri ARMA e GARCH del modello sono incogniti ed è necessario stimarli con le informazioni disponibili. La stima viene fatta, di solito, con il metodo della massima verosimiglianza sotto l'ipotesi che siano note le distribuzioni delle innovazioni ε_t ed a_t . Come abbiamo visto per gli ARCH la stima è ottenuta in modo iterativo ed utilizzando appropriati pacchetti statistici. Ottenute le stime dei parametri incogniti la previsione stimata della volatilità è ottenuta sostituendo queste stime nella (21) ed ottenendo

$$\hat{\sigma}_N^2(h) = \begin{cases} \hat{w} + \hat{\alpha} \hat{\varepsilon}_N^2 + \hat{\beta} \hat{\sigma}_N^2 & \text{per } h = 1 \\ \hat{w} + (\hat{\alpha} + \hat{\beta}) \hat{\sigma}_N^2(h-1) & \text{per } h > 1 \end{cases}$$

Come già detto, il modello stimato, prima di effettuare le previsioni, va

verificato ed in particolare va accertato che le autocorrelazioni di

$$\hat{a}_t = \hat{\varepsilon}_t / \hat{\sigma}_t, \quad \hat{a}_t^2 = (\hat{\varepsilon}_t / \hat{\sigma}_t)^2$$

siano tutte, o almeno le prime 10, all'interno della banda di confidenza $(-2/\sqrt{N}, 2/\sqrt{N})$, che la distribuzione di \hat{a}_t sia compatibile con quella di una normale o di una $t_{(k)}$ e che la curtosi di \hat{a}_t sia non superiore a quella di $\hat{\varepsilon}_t$.

Esempio 7

Consideriamo gli attivi Campari e Fiat osservati sia giornalmente che settimanalmente, nel periodo sopra specificato, adattiamo a questi attivi modelli della classe MA-GARCH-M e costruiamo le previsioni per $h=10$ tempi in avanti.

Dalla tabella 9 osserviamo che un modello che può rappresentare i rendimenti delle quotazioni giornaliere di Campari è un MA(2)-GARCH(1,1)-M mentre per quelli settimanali un MA(1)-GARCH(1,1)-M con, in entrambi i casi, $g(\sigma_t^2) = \log(\sigma_t^2)$. I risultati della previsione per i rendimenti giornalieri e settimanali, sia dei livelli che della volatilità, sono riportati nella figura 11.

Figura 11 – *Previsioni dei rendimenti dei livelli e della volatilità di Campari. Osservazioni giornaliere e settimanali*

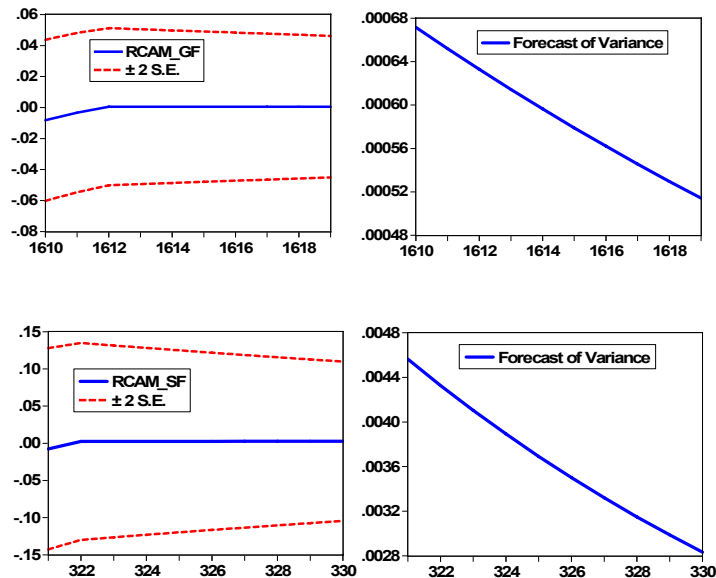
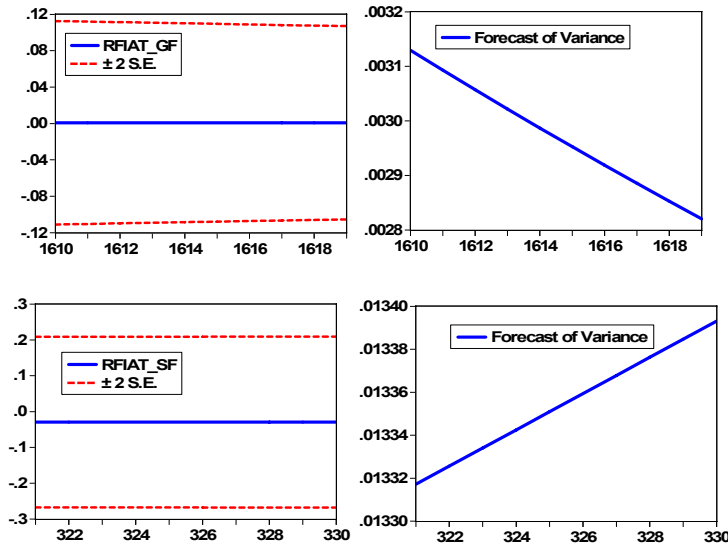


Figura 12 – Previsioni dei rendimenti dei livelli e della volatilità di Fiat. Osservazioni giornaliere e settimanali



Nella figura 12 sono riportate le previsioni dei livelli e della volatilità dei rendimenti giornalieri e settimanali Fiat. Ai dati giornalieri è stato adattato un modello GARCH(1,1)-M, mentre a quelli settimanali un IGARCH(1,1)-M. In entrambi i casi si è posto $g(\sigma_t^2) = \log(\sigma_t^2)$. Da questi grafici si può osservare come la volatilità di Fiat, e quindi la rischiosità di questo attivo, sia nettamente più elevata di quella di Campari.

□

5.7 L'effetto leverage nella volatilità

I modelli ARCH e GARCH sono funzioni dei rendimenti al quadrato o dei residui al quadrato del modello ARMA adattato ai rendimenti. Questo vuol dire che tali modelli trattano allo stesso modo sia rendimenti positivi che negativi. In pratica, nella maggior parte dei casi, si osserva che i rendimenti negativi sono più accentuati, anche se in numero inferiore, di quelli positivi. D'altro lato, come si può ben comprendere, sono le flessioni delle quotazioni (mercato orso), con conseguenti rendimenti negativi, che generalmente suscitano negli operatori più apprensione rispetto ai rialzi (mercato toro).

Per catturare questo effetto asimmetrico, che viene chiamato *leverage*, sono state proposte una serie di varianti dei GARCH. Qui di seguito presentiamo quelli più noti ed utilizzati nelle pratiche applicazioni.

(A) IL MODELLO EGARCH

E' stato proposto in Nelson (1991), è uno dei più utilizzati per catturare il *leverage* degli attivi e si basa su una funzione esponenziale da cui prende il nome. Qui di seguito ne presentiamo una variante che generalizza quella originaria conservandone le relative proprietà statistiche ed interpretative. I rendimenti di un dato attivo hanno una struttura EGARCH(p,q,v) se è:

$$r_t = \mu_t + \varepsilon_t; \quad \varepsilon_t = \sigma_t a_t; \quad \text{con } a_t \stackrel{\text{iid}}{\sim} (0, 1)$$

$$\log(\sigma_t^2) = w + \sum_{j=1}^q \beta_j \log(\sigma_{t-j}^2) + \sum_{j=1}^p \alpha_j \frac{|\varepsilon_{t-j}|}{\sigma_{t-j}} + \sum_{j=1}^v \gamma_j \frac{\varepsilon_{t-j}}{\sigma_{t-j}} \quad (23)$$

con μ_t espressa, come al solito, con qualche semplice modello della classe ARMA. In pratica, quasi sempre viene utilizzato un modello con $p=q=v=1$ di modo che la (23) diviene un EGARCH(1,1,1) e risulta:

$$\log(\sigma_t^2) = w + \beta \log(\sigma_{t-1}^2) + \alpha \frac{|\varepsilon_{t-1}|}{\sigma_{t-1}} + \gamma \frac{\varepsilon_{t-1}}{\sigma_{t-1}}. \quad (24)$$

L'effetto asimmetrico è dovuto al fatto che $\log(\sigma_t^2)$ risponde in modo differente a seconda che ε_t sia positivo o negativo, infatti:

$$\log(\sigma_t^2) = \begin{cases} w + \beta \log(\sigma_{t-1}^2) + (\alpha + \gamma) \frac{|\varepsilon_{t-1}|}{\sigma_{t-1}} & \text{se } \varepsilon_{t-1} \geq 0 \\ w + \beta \log(\sigma_{t-1}^2) + (\alpha - \gamma) \frac{|\varepsilon_{t-1}|}{\sigma_{t-1}} & \text{se } \varepsilon_{t-1} < 0 \end{cases}$$

ed è il parametro γ che permette di discriminare fra la volatilità indotta da residui non negativi $\varepsilon_{t-1} \geq 0$ da quella di residui negativi. In genere ci si aspetta che deviazioni ε_{t-1} negative facciano crescere la volatilità e che quindi debba essere $\gamma < 0$. Se si considera l'esponenziale di ambo i membri della (24) si ottiene

$$\sigma_t^2 = e^w \left(\sigma_{t-1}^2 \right)^\beta e^{(\alpha |\varepsilon_{t-1}| + \gamma \varepsilon_{t-1}) / \sigma_{t-1}} \quad (25)$$

che ne giustifica il nome di EGARCH. Se si utilizza l'operatore B, la (24) diviene

$$(1 - \beta B) \log(\sigma_t^2) = \frac{\alpha |\varepsilon_{t-1}| + \gamma \varepsilon_{t-1}}{\sigma_{t-1}}$$

ed il modello ha senso se è $0 \leq \beta < 1$, indipendentemente dal valore e dal segno degli altri parametri. In conclusione l'EGARCH modella la volatilità in modo non lineare e questo permette di catturare l'esistenza del *leverage*.

Come visto per i GARCH, anche per quelli EGARCH esiste una versione EGARCH-M per l'eventuale *premio al rischio* presente nei dati e, quello che è più rilevante, la stima del modello permette di derivare previsioni h tempi in avanti. In particolare, le previsioni della volatilità da un EGARCH(1,1,1) si ottengono utilizzando la (25). Per $h=1$ la previsione si ottiene direttamente:

$$\sigma_N^2(1) = \begin{cases} e^w \left(\sigma_N^2 \right)^\beta e^{(\alpha + \gamma) \varepsilon_N / \sigma_N} & \text{se } \varepsilon_N \geq 0 \\ e^w \left(\sigma_N^2 \right)^\beta e^{(\alpha - \gamma) \varepsilon_N / \sigma_N} & \text{se } \varepsilon_N < 0 \end{cases}$$

Per $h > 1$ passi in avanti, tenendo conto che è

$$\begin{aligned} \sigma_{N+h}^2 &= e^w \left(\sigma_{N+h-1}^2 \right)^\beta \exp \left\{ \left(\alpha |\varepsilon_{N+h-1}| + \gamma \varepsilon_{N+h-1} \right) / \sigma_{N+h-1} \right\} = \\ &= e^w \left(\sigma_{N+h-1}^2 \right)^\beta \exp \left\{ \left(\alpha |a_{N+h-1}| + \gamma a_{N+h-1} \right) \right\} \end{aligned}$$

risulta

$$\sigma_N^2(h) = E \left(\sigma_{N+h}^2 \mid \mathfrak{F}_N \right) = e^w \left(\sigma_N^2(h-1) \right)^\beta E \left(\exp \left(\alpha |a_{N+h-1}| + \gamma a_{N+h-1} \right) \right)$$

e la media $E \left(\exp \left(\alpha |a_{N+h-1}| + \gamma a_{N+h-1} \right) \right)$, di solito, può essere calcolata solo per

via numerica se è nota la distribuzione di a_t . Infatti, se indichiamo con $f(z)$ la funzione di densità di a_t si ottiene:

$$\begin{aligned} E\left(\exp(\alpha |a_{N+h-1}| + \gamma a_{N+h-1})\right) &= \int_{-\infty}^{\infty} \exp(\alpha |z| + \gamma z) f(z) dz = \\ &= \int_{-\infty}^0 \exp(z(\alpha - \gamma)) f(z) dz + \int_0^{\infty} \exp(z(\alpha + \gamma)) f(z) dz \end{aligned}$$

che nel caso di $a_t \stackrel{\text{iid}}{\sim} N(0, 1)$ diviene

$$\begin{aligned} E\left(\exp(\alpha |a_{N+h-1}| + \gamma a_{N+h-1})\right) &= \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 \exp\left((\alpha - \gamma)z - z^2/2\right) dz + \frac{1}{\sqrt{2\pi}} \int_0^{\infty} \exp\left((\alpha + \gamma)z - z^2/2\right) dz = \\ &= \exp\left((\alpha - \gamma)^2/2\right) \Phi(\gamma - \alpha) + \exp\left((\alpha + \gamma)^2/2\right) \Phi(\gamma + \alpha) \end{aligned}$$

dove $\Phi(z)$ è la funzione di ripartizione della normale standardizzata calcolata nel punto z . In definitiva, sotto l'ipotesi di normalità, la previsione della volatilità $h > 1$ tempi in avanti di un EGARCH(1,1,1) è data da:

$$\sigma_N^2(h) = e^w \left(\sigma_N^2(h-1)\right)^\beta \left[\exp\left((\alpha - \gamma)^2/2\right) \Phi(\gamma - \alpha) + \exp\left((\alpha + \gamma)^2/2\right) \Phi(\gamma + \alpha) \right]$$

Se è

$$a_t \stackrel{\text{iid}}{\sim} ((k-2)/k)^{1/2} t_{(k)}$$

i risultati sono formalmente più complessi anche se si derivano con una procedura simile.

Esempio 8

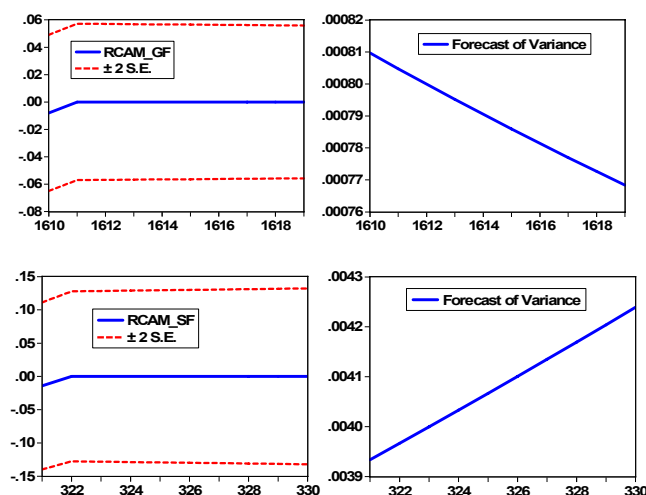
Ai rendimenti di Campari e Fiat, rilevati sia giornalmente che

settimanalmente, adattiamo modelli della classe MA-EGARCH-M e prevediamo $h=10$ valori sia nei livelli che nella volatilità. Nella tabella 10 è riportata la stima del miglior modello della classe.

Tabella 10 – *Stime del modello MA-EGARCH-M per i rendimenti di Campari e Fiat. Dati giornalieri e settimanali*

Campari	θ_1		w	α	γ	β	AIC
Giorn.	0.117 (.000)		-0.168 (.000)	0.101 (.000)	-0.040 (.000)	0.989 (.000)	-5.507
Settim.	0.171 (.000)		0.025 (.000)	-0.021 (.000)	-0.046 (.000)	.999 (.000)	-4.055
Fiat	ϕ_0	$\log(\sigma_t^2)$	w	α	γ	β	AIC
Giorn.	-0.023 (.001)	-0.003 (.000)	-0.271 (.000)	0.159 (.000)	-0.070 (.000)	0.981 (.000)	-4.844
Settim.	-0.111 (.003)	-0.018 (.002)	-0.071 (.350)	0.146 (.001)	-0.001 (.970)	1.006 (.000)	-3.219

Figura 13 – *Previsione dei livelli e della volatilità per $h=10$ dei rendimenti Campari. Dati giornalieri (primi due grafici) e settimanali*

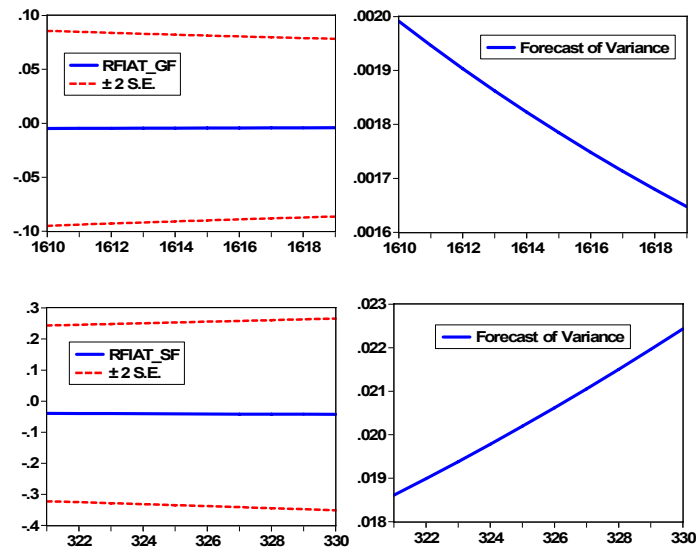


Come si può osservare dalla tabella 10 ai rendimenti di Campari è stato adattato un modello MA(1)-EGARCH(1,1,1), mentre a quelli di Fiat un EGARCH(1,1,1)-M; per tutti i modelli, come atteso, la stima di γ è negativa. Il modello scelto si adatta bene ai dati giornalieri dei rendimenti sia di Campari che di Fiat, mentre per quelli settimanali è sicuramente non idoneo per quelli

Fiat ed ai limiti per quelli Campari.

Nella figura 13 sono riportate le previsioni $h=10$ tempi in avanti per i rendimenti di Campari sia giornalieri che settimanali utilizzando il modello stimato della tabella 10. Queste previsioni sono molto simili a quelle ottenute con il modello MA-GARCH. Nella figura 14 sono riportate le stesse previsioni riferite ai rendimenti di Fiat sia giornalieri che settimanali. Anche queste previsioni sono molto simili a quelle ottenute con il modello GARCH.

Figura 14 – Previsione dei livelli e della volatilità per $h = 10$ dei rendimenti Fiat. Dati giornalieri (primi due grafici) e settimanali



□

(B) IL MODELLO TGARCH

Questo modello è stato proposto da Glosten et al. (1993) e da Zakoian (1994) ed utilizza, come abbiamo già visto per i TARMA, una variabile dicotomica funzione di r_{t-1} per discriminare fra rendimenti non negativi e negativi e catturare così l'eventuale effetto asimmetrico presente nella volatilità dell'attivo. Se indichiamo con I_{t-j} la variabile dicotomica:

$$I_{t-j} = \begin{cases} 0 & \text{se } r_{t-j} \geq 0 \\ 1 & \text{se } r_{t-j} < 0 \end{cases}$$

il modello TGARCH(p, q, v) è dato da

$$\sigma_t^2 = w + \sum_{j=1}^p \alpha_j \varepsilon_{t-j}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2 + \sum_{j=1}^v \gamma_j (\varepsilon_{t-j}^2 I_{t-j})$$

In pratica, il modello più largamente utilizzato è il TGARCH(1,1,1):

$$\sigma_t^2 = w + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 + \gamma (\varepsilon_{t-1}^2 I_{t-1}) \quad (26)$$

che può essere scritto:

$$\sigma_t^2 = \begin{cases} w + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 & \text{se } r_{t-1} \geq 0 \\ w + (\alpha + \gamma) \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 & \text{se } r_{t-1} < 0 \end{cases} \quad (27)$$

e mette in evidenza l'impatto asimmetrico che il segno dei residui ha sulla volatilità dell'attivo nei due regimi. Dalla (27) si ha che il segno di γ può essere sia positivo che negativo, anche se è ragionevole attendersi che questo sia negativo dato che per innovazioni ε_{t-1} negative è plausibile aspettarsi un aumento della volatilità. Da $\varepsilon_t \sim \text{WN}(0, \sigma_\varepsilon^2)$ e dalla stazionarietà e natura binaria di I_{t-1} si ha

$$\begin{aligned} \sigma_\varepsilon^2 &= E(\varepsilon_t^2) = E(\varepsilon_{t-1}^2) = pE(\varepsilon_{t-1}^2 | I_{t-1} = 1) + (1-p)E(\varepsilon_{t-1}^2 | I_{t-1} = 0) = \\ &= p\mu_{2\varepsilon|1} + (1-p)\mu_{2\varepsilon|0} \end{aligned} \quad (28)$$

D'altra parte, tenendo conto della (28), risulta

$$E(\varepsilon_{t-1}^2 I_{t-1}) = pE(\varepsilon_{t-1}^2 | I_{t-1} = 1) = p\mu_{2\varepsilon|1} = \sigma_\varepsilon^2 - (1-p)\mu_{2\varepsilon|0}. \quad (29)$$

Considerando la media di ambo i membri della (26) e ricordando la (29) otteniamo

$$\sigma_\varepsilon^2 = w + \alpha \sigma_\varepsilon^2 + \beta \sigma_\varepsilon^2 + \gamma E(\varepsilon_{t-1}^2 I_{t-1}) =$$

$$= w + \alpha\sigma_{\varepsilon}^2 + \beta\sigma_{\varepsilon}^2 + \gamma p\mu_{2\varepsilon|1} = w + \alpha\sigma_{\varepsilon}^2 + \beta\sigma_{\varepsilon}^2 + \gamma\sigma_{\varepsilon}^2 - \gamma(1-p)\mu_{2\varepsilon|0} \quad (30)$$

In fine, dalla (30) si derivano le due espressioni alternative per σ_{ε}^2 :

$$\sigma_{\varepsilon}^2 = \frac{w + \gamma p\mu_{2\varepsilon|1}}{1 - (\alpha + \beta)} = \frac{w - \gamma(1-p)\mu_{2\varepsilon|0}}{1 - (\alpha + \beta + \gamma)} \quad (31)$$

Dalla (31) si ottengono le condizioni perché un TGARCH(1,1,1) sia considerato accettabile:

$$w + \gamma p\mu_{2\varepsilon|1} > 0; \quad \alpha \geq 0; \quad \beta \geq 0; \quad 1 - (\alpha + \beta) > 0;$$

oppure:

$$w + \gamma p\mu_{2\varepsilon|1} < 0; \quad \alpha \geq 0; \quad \beta \geq 0; \quad 1 - (\alpha + \beta) < 0;$$

oppure:

$$w - \gamma(1-p)\mu_{2\varepsilon|0} > 0; \quad \alpha \geq 0; \quad \beta \geq 0; \quad 1 - (\alpha + \beta + \gamma) > 0$$

oppure:

$$w - \gamma(1-p)\mu_{2\varepsilon|0} < 0; \quad \alpha \geq 0; \quad \beta \geq 0; \quad 1 - (\alpha + \beta + \gamma) < 0.$$

Come si può osservare le condizioni perché un TGARCH sia accettabile sono molto più ampie di quelle di un semplice GARCH. In particolare, il processo è accettabile se i parametri $(w, \alpha, \beta, \gamma)$ sono tutti positivi ed è $(\alpha + \beta) < 1$. Questo implica che un TGARCH può essere localmente non accettabile (in uno dei due regimi), ma può esserlo comunque globalmente. Se è $\gamma = 0$ e $v = 1$, oppure $p=0$, il TGARCH(1, 1, 1) e le condizioni per la sua esistenza si riducono a quelle del classico GARCH(1, 1).

La stima dei parametri di un ARMA-TGARCH si ottiene, con qualche complicazione formale in più, con una metodologia simile a quella utilizzata per gli ARMA-GARCH tenendo conto che $p = E(I_{t-1})$ può essere stimato a priori e facilmente con

$$\hat{p} = \sum_{t=2}^N I_{t-1} / N$$

naturalmente, anche per il TGARCH esiste la variante TGARCH-M.

Calcoliamo ora la previsione della volatilità h passi in avanti di un TGARCH:

- per $h=1$ si ha

$$\begin{aligned}\sigma_N^2(1) &= E(\sigma_{N+1}^2 | \mathfrak{F}_N) = E(w + \alpha \varepsilon_N^2 + \beta \sigma_N^2 + \gamma (\varepsilon_N^2 I_N) | \mathfrak{F}_N) = \\ &= w + \alpha \varepsilon_N^2 + \beta \sigma_N^2 + \gamma \sigma_N^2 I_N\end{aligned}$$

dato che noto \mathfrak{F}_N lo sono anche ε_N^2 e σ_N^2 ;

- per $h>1$ risulta

$$\begin{aligned}\sigma_N^2(h) &= E(\sigma_{N+h}^2 | \mathfrak{F}_N) = E(w + \alpha \varepsilon_{N+h-1}^2 + \beta \sigma_{N+h-1}^2 + \gamma (\varepsilon_{N+h-1}^2 I_{N+h-1}) | \mathfrak{F}_N) = \\ &= w + (\alpha + \beta + \gamma p) \sigma_N^2 (h-1)\end{aligned}$$

Osserviamo che se è $p=0$ queste formule si riducono a quelle già viste per il modello GARCH(1,1).

Il modello (26) può essere ulteriormente generalizzato di modo che l'effetto soglia sia presente non solo su α , ma anche su (w, β) utilizzando un TARGARCH(2;1,1):

$$\sigma_t^2 = (w_1 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2)(1 - I_{t-1}) + (w_2 + \alpha_2 \varepsilon_{t-1}^2 + \beta_2 \sigma_{t-1}^2) I_{t-1} \quad (32)$$

e l'effetto soglia è evidenziato da:

$$\sigma_t^2 = \begin{cases} w_1 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 & \text{se } I_{t-1} = 0 \\ w_2 + \alpha_2 \varepsilon_{t-1}^2 + \beta_2 \sigma_{t-1}^2 & \text{se } I_{t-1} = 1 \end{cases}$$

La (32) si può anche scrivere:

$$\sigma_t^2 = w_1 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 + \left((w_2 - w_1) + (\alpha_2 - \alpha_1) \varepsilon_{t-1}^2 + (\beta_2 - \beta_1) \sigma_{t-1}^2 \right) I_{t-1}$$

che permette di ottenere le previsioni della varianza condizionata:

$$\sigma_N^2(1) = (w_1 + \alpha_1 \varepsilon_N^2 + \beta_1 \sigma_N^2) I_N + (w_2 + \alpha_2 \varepsilon_N^2 + \beta_2 \sigma_N^2) (1 - I_N)$$

$$\sigma_N^2(j) = (w_1 + (\alpha_1 + \beta_1) \sigma_N^2(j-1)) p + (w_2 + (\alpha_2 + \beta_2) \sigma_N^2(j-1)) (1-p)$$

Questi risultati generalizzano quelli visti per i TGARCH e mettono anche in evidenza che il modello ha senso se almeno uno dei parametri del primo regime è sostanzialmente diverso dal corrispondente del secondo regime. Se i tre parametri del primo regime coincidono con quelli del secondo il modello si riduce al classico GARCH(1,1). D'altra parte, tenuto conto che è

$$I_{t-1} + (1 - I_{t-1}) = 1 \text{ e posto } \sigma_t^2 = \varepsilon_t^2 - \eta_t$$

il TARGARCH(2;1,1) si può scrivere

$$\begin{aligned} \varepsilon_t^2 = & (w_1 + (\alpha_1 + \beta_1) \varepsilon_{t-1}^2 + \eta_t - \beta_1 \eta_{t-1}) I_{t-1} + \\ & + (w_1 + (\alpha_1 + \beta_1) \varepsilon_{t-1}^2 + \eta_t - \beta_1 \eta_{t-1}) (1 - I_{t-1}) \end{aligned}$$

che è una TARMA(2;1,1) in ε_t^2 e le proprietà dei TARMA vengono estese immediatamente ai TARGARCH. Anche le previsioni di questi modelli possono essere ottenute con la stessa tecnica vista per i TARMA.

Esempio 9

Ai rendimenti sia giornalieri che settimanali degli attivi Campari, Fiat e Banca Intesa S. Paolo adattiamo un modello della classe MA-TGARCH-M. I risultati della stima ottimale sono riportati nella tabella 11.

Come si può osservare dalla tabella 11 il modello della classe MA-TGARCH-M adatta bene tutte le serie con l'esclusione dei rendimenti settimanali di Campari.

Tabella 11 – *Stima di un modello MA-TGARCH-M per i rendimenti degli attivi Campari, Fiat, B. Intesa S. Paolo. Dati giornalieri e settimanali*

Campari	θ_1			w	α	γ	β	$\alpha + \beta$
Giorn.	0.111 (.000)			.000003 (.001)	0.017 (.013)	0.047 (.000)	0.952 (.000)	0.969

Settim.	0.073 (.000)			0.00023 (.000)	-0.097 (.000)	.542 (.000)	.677 (.000)	
Fiat	ϕ_0	$\log(\sigma_t^2)$		w	α	γ	β	$\alpha + \beta$
Giorn.	-0.016 (.033)	-0.002 (.027)		0.00001 (.000)	0.039 (.003)	0.093 (.000)	0.895 (.000)	0.934
Settim.	-0.088 (.027)	-0.014 (.023)		0.000065 (.046)		0.127 (.000)	0.940 (.000)	0.940
Int. S. P.	θ_1	θ_4	σ_t^2	w	α	γ	β	$\alpha + \beta$
Giorn.	0.989 (.000)	0.008 (.000)	-0.004	0.000002		0.127 (.000)	0.940 (.000)	0.940
Settim.	0.988 (.000)			-0.000000 (.942)		0.148 (.000)	0.942 (.000)	0.942

□

Con le stesse argomentazioni utilizzate per il TGARCH, in questo caso si ottiene:

$$\begin{aligned} \sigma_{\varepsilon}^2 &= \frac{w_1 + p(w_2 - w_1) + ((\alpha_2 - \alpha_1) + (\beta_2 - \beta_1))p\mu_{2\varepsilon|1}}{1 - (\alpha_1 + \beta_1)} = \\ &= \frac{w_1 + p(w_2 - w_1) + ((\alpha_1 - \alpha_2) + (\beta_1 - \beta_2))(1-p)\mu_{2\varepsilon|0}}{1 - (\alpha_2 + \beta_2)} \end{aligned}$$

che definiscono le condizioni per l'esistenza del TARGARCH(2;1,1). Queste condizioni sono molto più generali di quelle di un semplice GARCH ed è sufficiente che uno dei due regimi sia accettabile (accettabilità locale) perché il modello lo sia (accettabilità globale).

La stima dei parametri con il metodo della massima verosimiglianza si ottiene con una procedura iterativa simile a quella vista per gli ARCH ed i GARCH, mentre le previsioni presentano gli stessi problemi per la previsione dei TARMA.

Da un punto di vista pratico, il modello (32) non può essere stimato direttamente con il pacchetto EViews attualmente in uso, ma è necessario elaborare un idoneo programma di calcolo.

(C) IL MODELLO PGARCH

Questo modello è stato introdotto da Ding et al. (1993) e nella sua formulazione generale il PGARCH($p, q; \delta$) (*Power GARCH*) è dato da:

$$\sigma_t^\delta = w + \sum_{j=1}^p \alpha_j \left(\left| \varepsilon_{t-j} \right| - \gamma_j \varepsilon_{t-j} \right)^\delta + \sum_{j=1}^q \beta_j \sigma_{t-j}^\delta \quad (33)$$

con

$$\delta > 0; \quad \left| \gamma_j \right| = \begin{cases} \leq 1 & \text{per } j = 1, 2, \dots, r \leq p \\ 0 & \text{per } j > r \end{cases}$$

Anche questo modello è utilizzato, qualche volta, nel solo caso particolare di $p=q=1$:

$$\sigma_t^\delta = w + \alpha \left(\left| \varepsilon_{t-1} \right| - \gamma \varepsilon_{t-1} \right)^\delta + \beta \sigma_{t-1}^\delta$$

L'effetto asimmetrico si ottiene per $\gamma \neq 0$, mentre per $\delta = 2$ e $\gamma = 0$ ci riconduce al classico modello GARCH(1,1). La stima dei parametri viene ottenuta con il metodo della massima verosimiglianza utilizzando una procedura iterativa e le stime ottenute vengono usate per costruire le previsioni h tempi in avanti. Naturalmente, di questo modello esiste anche la versione PGARCH-M per il quale valgono le considerazioni generali fatte nel paragrafo precedente.

Esempio 10

Ai rendimenti di Campari, Fiat e Banca Intesa S. Paolo adattiamo il miglior modello della classe MA-PGARCH-M. Nella tabella 12 sono riportati i risultati della stima sia per dati giornalieri che settimanali.

Come si può osservare dalla tabella 12, il modello PGARCH-M non è idoneo per nessuna delle sei serie analizzate.

Tabella 12 – *Stima dell' MA-PGARCH-M per i rendimenti di Campari, Fiat e Banca Intesa S. Paolo. Dati giornalieri e settimanali*

Campari	θ_1		w	α	γ	δ	β	AIC
Giorn.	0.113 (.000)		0.000016 (.000)	0.045 (.000)	0.360 (.002)	1.608 (.000)	0.950 (.000)	-5.507

Settim.	0.129	-0.00035 (.005)	-0.0003 (.305)	0.020	0.977 (.000)	0.317 (.000)	0.990 (.000)	-4.045
Fiat	ϕ_0	$\log(\sigma_t^2)$	w	α	γ	δ	β	AIC
Giorn.	-0.020 (.007)	-0.0026 (.005)	0.00025 (.305)	0.086 (.000)	0.425 (.000)	1.179 (.000)	0.988 (.000)	-4.845
Settim.	-0.092 (.000)	-0.0149 (.000)	-0.00058 (.852)	0.082 (.003)	0.045 (.848)	.432 (.390)	0.936 (.000)	-3.221
Int S. P.	θ_1		w	α	γ	δ	β	AIC
Giorn.	0.997 (.000)		0.000002 (.825)	0.120 (.000)	0.142 (.000)	1.403 (.000)	0.908 (.000)	-5.305
Settim.	0.987 (.000)		0.000006 (.836)	0.022 (.787)	0.726 (.723)	3.068 (.051)	0.930 (.000)	-3.726

□

Esempio 11

Alle tre serie viste nell'esempio precedente è stato adattato un TMA(2;2)-TIGARCH(1,1,1) in modo da evidenziare, se esistono, differenze significative sia nei livelli che nella volatilità nei due regimi. Qui di seguito è riportato il listato riguardante i prezzi P di un generico attivo di N osservazioni. Il modello è stato iterato K=50 volte. Modelli con struttura diversa, nei livelli e nella volatilità, possono essere ottenuti modificando opportunamente il listato. Per far girare il programma è necessario che sia attivo un Workfile con dentro la serie storica dei prezzi dell'attivo P e sostituire ad N il numero delle osservazioni.

'Stima TMA_TIGARCH

```

smpl 1 N
genr r=d(log(P))
smpl 1 N
for li=3 to N
    smpl li li
    if r(li-1)>=0 then genr it=1
    else genr it=0
endif
next
smpl 1 N
scalar p=@mean(it)

```

```

equation eqv.arch(1,1,integrated) r c ma(1) ma(2)
fit rf vse v
gener re=resid
equation eq1.ls r=(c(1)+c(2)*re(-1)+c(3)*re(-2))*it+(c(4)+c(5)*re(-1)+c(6)*re(-2))*(1-it)
gener re=resid
for li=1 to 50
    equation eq1.ls r=(c(1)+c(2)*(c(4)+c(5)*re(-1)+c(6)*re(-2))*(1-it)
    gener re=resid
    equation eqv1.ls v=((1-c(3))*re(-1)^2+c(3)*v(-1))*it+((1-c(5))*re(-1)^2+c(5)*v(-1))*(1-it)
    fit rv
    gener rv=v
    simpl 3 N
next
simpl 1 N
gener at=re/v^.5
for li=3 to N
    simpl li li
    if r(li-1)>=0 then gener v1=v
    gener r1=r
    gener a1=at
    else gener v2=v
    gener r2=r
    gener a2=at
endif
next
simpl 1 N

```

Output di interesse:

- eq1 ed eqv1 in cui sono riportate le stime dei parametri dei livelli e della volatilità dei due regimi;
- p in cui è riportata la stima di $E(I_{t-1})$;
- a1 ed a2 in cui sono riportati i residui standardizzati dei due regimi;
- r1 ed r2 in cui sono riportati i rendimenti dei due regimi;
- v1 e v2 che riportano le serie delle volatilità dei due regimi.

Tabella 13 – *Stime dei livelli e della volatilità dei due regimi per ogni le quotazioni giornaliera di Campari, Fiat e Banca Intesa S. Paolo*

	1° regime		2° regime					r_1	r_2	v_1	v_2	
Camp.												
Livello	θ_1	θ_2	ϕ_0	θ_1	θ_2	p	AIC	\bar{r}	.000	.000	.0003	.0003
	.109 (.003)	.097 (.007)		.098 (.008)		.548 (.000)	-5.304	$\hat{\sigma}$.019	.020	.0005	.0004
									$\hat{\gamma}_1$	-.20	-.10	6.8
Volat.	α	$1 - \alpha$		α	$1 - \alpha$			$\hat{\gamma}_2$	15.9	12.1	59.6	41.4
	.031 (.000)	.969 (.000)		.029 (.000)	.971 (.000)							
Fiat												
Livello	θ_1	θ_2	ϕ_0	θ_1	θ_2	p	AIC	\bar{r}	.000	.000	.0005	.0005
					-.08 (.030)	.525 (.000)	-4.607	$\hat{\sigma}$.022	.023	.0005	.0005
									$\hat{\gamma}_1$	-.05	.12	3.38
Volat.	α	$1 - \alpha$		α	$1 - \alpha$			$\hat{\gamma}_2$	7.3	3.5	16.9	11.1
	.052 (.000)	.948 (.000)		.054 (.000)	.946 (.000)							
In. S. P.												
Livello	θ_1	θ_2	ϕ_0	θ_1	θ_2	p	AIC	\bar{r}	.000	.001	.0005	.0005
			.003 (.021)	-.149 (.002)	.194 (.000)	.563 (.000)	-4.812	$\hat{\sigma}$.021	.025	.0009	.0009
									$\hat{\gamma}_1$	-.10	.08	4.2
Volat.	α	$1 - \alpha$		α	$1 - \alpha$			$\hat{\gamma}_2$	15.5	11.6	20.1	16.5
	.076 (.000)	.924 (.000)		.079 (.000)	.921 (.000)							

Nella tabella 13 sono riportati i risultati della stima locale riferita ai due regimi con gli indici descrittivi più rilevanti che caratterizzano sia i livelli locali che la volatilità locale. Come si può osservare da questa tabella, per tutte e tre le serie analizzate i parametri dei livelli nei due regimi sono sostanzialmente differenti fra di loro, mostrano una marcata differenziazione nella loro curtosi con una sistematica leptocurtosi del primo regime rispetto al secondo. I parametri della volatilità nei due regimi sono sostanzialmente identici, per ciascuna delle tre serie considerate. Una marcata diversità si ha solo per una maggiore curtosi della volatilità del primo regime rispetto al secondo.

□

La stima di modelli ARMA-GARCH e sue varianti può essere ottenuta interattivamente da Eviews. Qui di seguito sono riportati alcuni casi notevoli:

A) STIMA DI UN MA(2)-GARCH(1,1)

1. “Cliccare” su **Quick** e quindi su **Estimation Equation...**
2. “Cliccare” su **Methods** e quindi su **ARCH-Autoregressive Conditional Heteroskedasticity**
3. Nel riquadro **Mean equation** inserire **$d(\log(P)) c ma(1) ma(2)$** , dove P è il nome dato alla serie dei prezzi degli attivi. Questo comando permette di calcolare i rendimenti e stimare un MA(2). Premere il pulsante **ok**

B) STIMA DI UN MA(2)-IGARCH(1,1)

Ripetere i passi 1.-3. sopra riportati e “cliccare” su **Restrictions**, scegliere IGARCH e quindi premere **ok**

C) STIMA DI UN MA(2)-TGARCH(1,1,1)

Ripetere i passi 1.-3. e in **Threshold** inserire il numero 1, quindi premere **ok**

D) STIMA DI UN MA(2)-EGARCH(1,1,1)

Ripetere i passi 1.-3. e in **Model** scegliere **EGARCH** e quindi premere **ok**

Le procedure sopra elencate forniscono le stime dei parametri del modello sotto l'ipotesi che i residui siano normali standardizzate, se si vogliono sotto l'ipotesi che questi siano t di Student è necessario, prima di premere **ok**, “cliccare” nella casella **Error distribution** e scegliere **Student's t**. Con procedure simili possono essere stimati modelli più complessi.

5.8 Stime QMV per modelli ARMA-GARCH

Bollerslev e Wooldridge (1992), d'ora in poi BW, propongono un approccio alla stima dei parametri di modelli ad eteroschedasticità condizionata basato sul metodo della QMV. Si assuma che il modello di riferimento per i rendimenti sia del tipo

$$r_t = \mu_t + \varepsilon_t \quad (34)$$

con

$$\varepsilon_t = \sigma_t a_t; \quad a_t \stackrel{\text{iid}}{\sim} (0, 1); \quad \mu_t = E(r_t | \mathfrak{F}_{t-1}) = m(\mathfrak{F}_{t-1}; \theta)$$

$$\sigma_t^2 = \text{var}(r_t | \mathfrak{F}_{t-1}) = h^2(\mathfrak{F}_{t-1}; \boldsymbol{\gamma})$$

dove $\boldsymbol{\theta}$ e $\boldsymbol{\gamma}$ sono due vettori di parametri incogniti da stimare di dimensioni $(m,1)$ e $(v,1)$, rispettivamente. Si noti che la famiglia parametrica a cui appartiene la distribuzione degli errori standardizzati a_t è stata volutamente lasciata non specificata. BW propongono di stimare i parametri del modello (34) massimizzando rispetto al vettore di parametri $\boldsymbol{\xi} = (\boldsymbol{\theta}', \boldsymbol{\gamma}')$ una funzione di quasi-verosimiglianza di tipo Gaussiano

$$\tilde{\boldsymbol{\xi}} = \arg \min_{\boldsymbol{\xi}} Q(\boldsymbol{\xi}; \mathbf{r})$$

dove

$$Q(\boldsymbol{\theta}, \boldsymbol{\gamma}; \mathbf{r}) = \sum_{t=1}^N Q_t(\boldsymbol{\theta}, \boldsymbol{\gamma}; r_t) = - \sum_{t=1}^N \left[\log(\sigma_t^2) + \frac{(r_t - \mu_t)^2}{\sigma_t^2} \right]$$

Se si introduce l'ulteriore ipotesi che la distribuzione degli errori standardizzati sia di tipo normale $a_t \stackrel{\text{iid}}{\sim} N(0,1)$ le stime ottenute sono stime di MV e quindi consistenti ed asintoticamente efficienti. Nel caso questa ipotesi non sia verificata, le stime $\tilde{\boldsymbol{\xi}}$ sono di QMV, ovvero ottenute massimizzando una funzione obiettivo che rappresenta solo un'approssimazione della funzione di verosimiglianza. BW dimostrano che le stime così ottenute sono consistenti anche se asintoticamente meno efficienti di quelle di MV. Per valori di N molto elevati, la distribuzione campionaria dello stimatore è ben approssimata da una variabile casuale normale multivariata

$$\tilde{\boldsymbol{\xi}} \approx N(\tilde{\boldsymbol{\xi}}, \boldsymbol{\Sigma}) \quad (35)$$

con $\boldsymbol{\Sigma} = N^{-1}(\mathbf{A}^{-1} \mathbf{B} \mathbf{A}^{-1})$ e

$$\mathbf{A} = -E(\boldsymbol{\Delta}_2(r_t; \boldsymbol{\xi})) \quad (36)$$

$$\mathbf{B} = E(\boldsymbol{\Delta}_1(r_t; \boldsymbol{\xi}) \boldsymbol{\Delta}_1(r_t; \boldsymbol{\xi})') \quad (37)$$

dove:

i) $\Delta_2(r_t; \xi)$ è la matrice delle derivate seconde della funzione di quasi-verosimiglianza aggregata rispetto agli elementi del vettore $\tilde{\xi}$ il cui generico elemento di posto (i, j) è dato da

$$\Delta_2(i, j) = \frac{\partial^2 \bar{Q}(\xi; r)}{\partial \xi_i \partial \xi_j} \quad (38)$$

dove, con ξ_i si è indicato l'i-esimo elemento del vettore ξ , per $i, j=1, \dots, (m+v)$.

ii) $\Delta_1(r_t; \xi)$ è il vettore, di dimensione $((m+v), 1)$, delle derivate delle funzione di quasi-verosimiglianza aggregata rispetto agli elementi del vettore ξ il cui i-esimo elemento è dato da

$$\Delta_1(i) = \frac{\partial \bar{Q}(\xi; r)}{\partial \xi_i} \quad (39)$$

per $i=1, \dots, (m+v)$.

Una valutazione approssimata degli standard errors degli stimatori di QMV può essere ottenuta calcolando la radice quadrata degli elementi posti sulla diagonale principale di Σ , la matrice di varianze e covarianze (asintotica) di $\tilde{\xi}$, la quale, a sua volta dipende dalle matrici \mathbf{A} e \mathbf{B} . Stime di \mathbf{A} e \mathbf{B} possono essere ottenute

i) sostituendo l'operatore media campionaria all'operatore valore atteso nelle equazioni (36) e (37)

ii) sostituendo $\tilde{\xi}$ a ξ nelle equazioni (38) e (39).

Attraverso queste operazioni si ottengono gli stimatori

$$\tilde{\mathbf{A}} = -\frac{1}{N} \sum_{t=1}^N \Delta_2(r_t; \tilde{\xi}) \quad (40)$$

$$\tilde{\mathbf{B}} = -\frac{1}{N} \sum_{t=1}^N \Delta_1(r_t; \tilde{\xi}) \Delta_1(r_t; \tilde{\xi})' \quad (41)$$

Delle stime degli standard errors degli elementi di $\tilde{\xi}$ possono quindi essere ottenute calcolando la radice quadrata degli elementi situati sulla diagonale principale della matrice

$$\tilde{\Sigma} = \frac{1}{N}(\tilde{\mathbf{A}}^{-1}\tilde{\mathbf{B}}\tilde{\mathbf{A}}^{-1})$$

Gli stimatori degli standard errors così ottenuti sono noti come *standard errors (s.e.) di Bollerslev-Wooldridge*. E' possibile dimostrare che gli s.e. di BW, oltre a fornire una stima consistente degli s.e. dello stimatore di QMV, mantengono la loro validità quando la distribuzione dell'errore standardizzato è di tipo normale e, quindi, lo stimatore di QMV ($\tilde{\xi}$) viene di fatto a coincidere con quello di MV ($\hat{\xi}$). E' infatti possibile dimostrare che, se l'ipotesi di normalità degli errori è verificata, allora $\mathbf{A}=\mathbf{B}$. Sostituendo nella (34), si ottiene

$$\tilde{\xi} \approx N(\xi, N^{-1}(\mathbf{A}^{-1})) \quad (42)$$

o equivalentemente

$$\tilde{\xi} \approx N(\xi, N^{-1}(\mathbf{B}^{-1})) \quad (43)$$

dove $\tilde{\xi} \equiv \hat{\xi}$, ovvero gli stimatori di QMV e MV coincidono. La distribuzione dello stimatore di MV $\hat{\xi}$ può quindi essere ottenuta come caso particolare della (34) sotto l'ipotesi di normalità degli errori. Tecnicamente si dice che gli stimatori degli s.e. di BW sono *robusti* rispetto a violazioni dell'ipotesi di normalità degli errori z_t .

Esempio 12.

Nell'esempio 4 sono stati stimati dei modelli ARMA-GARCH per i rendimenti giornalieri sui prezzi degli attivi Campari, Fiat e Intesa San Paolo attraverso la massimizzazione di una funzione di verosimiglianza di tipo Gaussiano. Le diagnostiche sui residui di tali modelli e, in particolare il test di Jarque-Bera (Tab 14), indicano che non vi è nei dati evidenza empirica sufficiente a confortare l'ipotesi di normalità degli errori.

Ne segue che le stime ottenute non possono essere interpretate come delle stime di MV ma piuttosto come delle stime di QMV. Al fine di ottenere delle stime consistenti degli standard errors dei parametri è quindi necessario utilizzare gli stimatori robusti degli s.e. proposti da BW.

Tabella 14 – Test di Jarque-Bera sui residui del modello ARMA-GARCH per i rendimenti giornalieri di Campari, Fiat e Intesa San Paolo: statistiche test e *p-values*

	Campari	Fiat	Int. San P.
Jarque-Bera	148.3628	136.4112	1971.7510
Prob.	0.0000	0.0000	0.0000

In Eviews questo obiettivo può essere raggiunto specificando un parametro opzionale nel comando utilizzato per la stima di modelli di tipo GARCH. Nel caso dei rendimenti giornalieri dell'esempio 5, i comandi di stima da utilizzare sono

```
equation modcamp_g.arch(1,1,b) dlcamp_g c ma(1) ma(2)
equation modfiat_g.arch(1,1,b) dlfiat_g
equation modintsp_g.arch(1,1,b) dlintsp_g
```

dove il parametro 'b' seleziona l'opzione di stima degli standard errors utilizzando le formule proposte da BW.

Tabella 15 – Stime di QMV dei modelli ARMA-GARCH per Campari, Fiat e B. Intesa S. Paolo per dati giornalieri. s.e. stimati con il metodo di BW

		θ_1	w	α	β	AIC	$\alpha + \beta$
Campari	Stima	0.113	0.000002	0.039	0.952	-5.509	0.993
	Prob.	(.000)	(.056)	(.000)	(.000)		
Fiat	Stima		0.00001	0.089	0.896	-4.832	0.985
	Prob.		(.020)	(.000)	(.000)		
Int. S. Paolo	Stima		0.000004	0.119	0.881	-5.296	1.002
	Prob.		(.031)	(.000)	(.000)		

Alcune osservazioni:

- 1) per quanto riguarda il titolo Campari, il coefficiente θ_2 non è risultato significativamente diverso da zero sulla base degli standard errors di BW. Si è quindi proceduto a ristimare il modello con l'ordine della componente MA ridotto da 2 ad 1;
- 2) si noti che selezionando l'opzione 'b \approx s.e. stimati con il metodo di BW' le stime degli s.e. ottenute saranno consistenti indipendentemente dal fatto che l'ipotesi di normalità degli errori sia verificata o meno. Diversamente, se l'opzione 'h' non viene selezionata, le stime degli s.e. calcolate da Eviews saranno basate sulla (43) e potrebbero risultare inconsistenti nel caso di

disnormalità degli errori standardizzati a_t ;

- 3) come è possibile verificare confrontando i risultati ottenuti in tabella 15 con quelli precedentemente riportati in tabella 5, l'opzione ' b ' non modifica le stime puntuali dei coefficienti ma solo le stime dei relativi standard errors e, quindi, i valori dei p-values.

CAPITOLO VI

La valutazione del rischio nei mercati monetari e finanziari

6.1 Introduzione

La stima e la previsione dei livelli e della volatilità di un attivo, un portafoglio di attivi, o quella di un intero mercato finanziario o monetario, forniscono informazioni sulla struttura del loro rischio di mercato. Ma gli strumenti analitico – statistici che abbiamo presentato nei capitoli precedenti possono essere utilizzati, così come fatto nel capitolo II, anche per definire indici statistici sintetici che permettono di fornire una valutazione futura di tale rischio in modo da poterlo, fra l'altro, confrontare nel tempo e nello spazio.

Nell'evoluzione dei prezzi di attivi una attenzione particolare è rivolta agli *eventi rari* ed alla stima della probabilità del loro verificarsi. L'attenzione per la costruzione di indici sintetici che siano in grado di misurare la rischiosità di un attivo o di un portafoglio di attivi è fortemente correlata al possibile presentarsi di eventi rari che, per definizione, cadono nelle code delle distribuzioni di probabilità che le caratterizzano. In tale ambito l'attenzione degli analisti e degli operatori è particolarmente interessata alla coda sinistra dato che, per un detentore di attivi, in questa si situano gli eventi avversi, quelli che sono connessi con le possibili perdite di mercato.

In generale, gli indici che misurano la rischiosità di attivi sono anche utilizzati per classificare, da parte di Agenzie specializzate possibilmente indipendenti, gli attivi stessi ed attribuire a questi un *rating* di affidabilità (basso, medio, alto, ecc.) che varia nel tempo. Il *rating* fornisce agli investitori, ed agli operatori di mercato in generale, una informazione sulla probabilità che quell'attivo o mercato ha di incorrere in eventi rari avversi. Questi *rating* vengono anche utilizzati dalle Istituzioni Finanziarie per stabilire se e a quale tasso di interesse fornire capitali alle società ed alle istituzioni emittenti attivi finanziari o monetari.

Uno degli indici più utilizzati per misurare il rischio di mercato di attivi e seguirlo nella sua evoluzione temporale è il VaR (*Value at Risk*) che può essere

definito come:

la perdita massima in cui un investitore potrà incorrere, con un prefissato livello di probabilità α , ad un orizzonte temporale futuro $N+h$

Se $\mathfrak{I}_N = (r_1, r_2, \dots, r_N)$ sono le informazioni disponibili al tempo N , il VaR sarà una funzione di \mathfrak{I}_N , α , h e scriveremo $\text{VaR}_N(h; \alpha)$, con $h=1, 2, \dots, M$ ed $0 < \alpha < 1$. Nelle pratiche applicazioni per dati giornalieri è $M=10$ (due settimane di borsa) ed $\alpha = 0.05$ oppure $\alpha = 0.01$.

Nei mercati monetari e finanziari esistono diversi tipi di rischio: *rischio di credito*, di *liquidità*, di *mercato*, ecc. Il VaR è essenzialmente un indice sintetico che misura il *rischio di mercato* dell'attivo o del portafoglio in analisi. Come già accennato, questa misura del rischio può essere utilizzata da istituzioni finanziarie e monetarie, oppure da singoli operatori ed investitori per valutare il rischio di mercato di attivi, ma viene sistematicamente usato anche da Enti Regolatori del Mercato per fissare regole ben definite sul tipo, la qualità e la quantità delle operazioni possibili. L'uso corretto delle informazioni che fornisce il VaR permette anche alle istituzioni finanziarie di poter continuare ad operare pure dopo che un evento raro avverso si è verificato in quel mercato. L'allentamento dei vincoli di mercato e la non curanza degli allarmi sul rischio connesso, segnalati dal crescere abnorme del VaR degli attivi e dei mercati finanziari di tutto il mondo, hanno fatto sì che, della crisi finanziaria iniziata nella tarda primavera del 2007, si è incominciato a prenderne coscienza solo un anno dopo quando gli effetti si erano già riversati sull'economia reale devastandola.

6.2 Il Valore a Rischio

Come accennato nel paragrafo precedente, il VaR è definito in un contesto probabilistico ed è connesso al verificarsi di eventi rari avversi che si situano nella coda sinistra della funzione di densità dei rendimenti. In termini statistici il VaR non è altro che un *percentile* della distribuzione di probabilità dei rendimenti attesi h tempi in avanti. Più precisamente, dato il processo stocastico dei rendimenti r_t , che sappiamo essere stazionario, riferito ai prezzi P_t di un attivo o di un portafoglio di attivi, siano \mathfrak{I}_{t-1} le informazioni

disponibili su r_t al tempo $t-1$, indichiamo con $F(r|\mathfrak{I}_{t-1})$ la distribuzione di ripartizione di r_t condizionata alle informazioni \mathfrak{I}_{t-1} . Dato che i rendimenti sono un processo continuo stazionario, questo ammetterà la funzione di densità condizionata $f(r|\mathfrak{I}_{t-1})$ che supporremo, senza perdere di generalità, essa stessa continua. Fissato il livello di probabilità α esiste un unico valore, diciamo $r_{(\alpha)}$ per cui vale la seguente uguaglianza:

$$F^{-1}(r_{(\alpha)}|\mathfrak{I}_{t-1}) = \alpha \quad (1)$$

La (1) permette di definire analiticamente il VaR calcolato al tempo $t-1$ con un orizzonte temporale $h=1$ ed al livello di probabilità α come:

$$\text{VaR}_{t-1}(1; \alpha) = r_{(\alpha)} \quad (2)$$

La grande diffusione che ha avuto questo indice sintetico fra gli operatori ed utilizzatori è dovuta a due fattori:

- la sua semplice ed immediata interpretazione: misura la perdita massima in cui si può incorrere con probabilità non superiore ad α prevista al tempo t con le informazioni disponibili al tempo precedente $t-1$;
- anche se questa misura del rischio può risultare non coerente con un insieme di postulati (vedi Artzner et al, 1999), deve la sua fortuna agli Accordi di Basilea sulla vigilanza bancaria (*Basle Committee in Banking Supervision*, 1995, 1996, 2003) che hanno imposto agli operatori finanziari e monetari la valutazione del rischio di mercato calcolando, fra l'altro, proprio il VaR fino ad un massimo di $M=10$ giorni.

In ogni caso il VaR è un indice di valutazione e previsione del rischio di mercato degli attivi di brevissimo periodo (massimo 10 giorni) e pertanto non è in grado di prevedere crisi generalizzate e di medio – lungo periodo. D'altro lato, misure alternative al VaR (vedi Dowd e Blake (2006)), che sono coerenti con i postulati di Artzner et al. (1999), sono comunque funzioni di percentili. Inoltre, sotto l'ipotesi che i rendimenti si distribuiscono come una Normale o una t di Student, il VaR soddisfa quei postulati e pertanto risulta coerente.

I percentili, come gli ordinamenti statistici, sono statistiche non regolari (non sono *smooth*), e la loro stima implicherebbe rilevanti problemi di approssi-

mazione anche con dati iid e questo non è il caso dei rendimenti che, abbiamo già visto, non sono mai iid. In generale, è noto (vedi Berkowitz, (2001)), che se X ha funzione di ripartizione $F(x)$ con funzione di densità $f(x)$ continua, dato il campione iid (x_1, \dots, x_N) estratto da X , se indichiamo con $F_N(x)$ la funzione di ripartizione empirica ottenuta sulla base del campione:

$$F_N(x) = \begin{cases} 0 & \text{per } x < x_{(1)} \\ 1/N & \text{per } x_{(1)} \leq x < x_{(2)} \\ 2/N & \text{per } x_{(2)} \leq x < x_{(3)} \\ \dots & \dots\dots\dots \\ 1 & \text{per } x \geq x_{(N)} \end{cases}$$

dove $(x_{(1)}, x_{(2)}, \dots, x_{(N)})$ è il campione osservato ordinato in senso non decrescente (*ordinamento statistico*), risulta

$$F_N(x) \xrightarrow{P} F(x)$$

Inoltre, se $x_{(\alpha)}$ è il *percentile* di ordine α di X calcolato su $F(x)$, mentre $\hat{x}_{(\alpha)}$ è quello calcolato su $F_N(x)$, se $\hat{x}_{(\alpha)}$ è unico risulta:

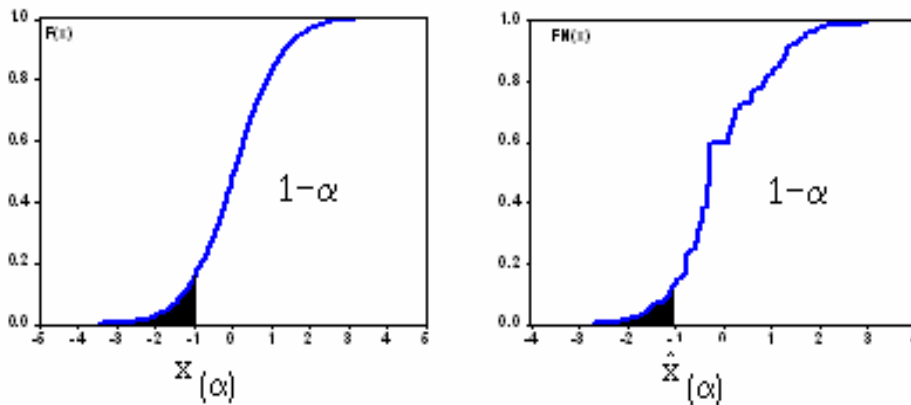
$$\sqrt{N}(\hat{x}_{(\alpha)} - x_{(\alpha)}) \xrightarrow{L} N\left(0; \frac{\alpha(1-\alpha)}{f^2(x_{(\alpha)})}\right), \quad 0 < \alpha < 1 \quad (3)$$

e pertanto l'errore di stima di $x_{(\alpha)}$ con $\hat{x}_{(\alpha)}$ è asintoticamente normale e dell'ordine $O(N^{-1/2})$ cioè tende a zero con \sqrt{N} .

Nella figura 1 è riportato schematicamente il percentile teorico $x_{(\alpha)}$ calcolato su $F(x)$ che stacca alla sua sinistra l' α per cento di probabilità e il

corrispondente empirico campionario $\hat{x}_{(\alpha)}$ calcolato su $F_N(x)$ con $N=200$.

Figura 1 – Percentile teorico di $F(x)$ e quello empirico di $F_N(x)$



Inoltre, date le osservazioni (x_1, x_2, \dots, x_t) supposte iid, per la relativa distribuzione empirica si ha $F_t(x) \stackrel{iid}{\sim} U(0, 1)$ ed equivalentemente

$$Z_t = \Phi^{-1}(F_t(x)) \stackrel{iid}{\sim} N(0, 1)$$

dove $U(0,1)$ è la v.c. uniforme distribuita nell'intervallo $(0,1)$ e $\Phi(\cdot)$ la funzione di ripartizione della Normale standardizzata.

Per N finito e prefissato l'ordine di approssimazione è tanto peggiore quanto più la probabilità α è vicina ai suoi estremi: zero ed uno. In generale, la distorsione della stima, a parità di altre condizioni, è positiva e cresce al decrescere di α ed al crescere della “pesantezza” delle code della funzione di densità $f(x)$ (vedi Inui et al. (2005)). Gli aspetti negativi della stima $\hat{x}_{(\alpha)}$ di $x_{(\alpha)}$ sono propri del VaR dato che è calcolato per valori piccoli di α : di solito per $\alpha = 0.05$ oppure per $\alpha = 0.01$ e per distribuzioni con code pesanti come, ad esempio, la v.c. $t_{(5)}$ utilizzata nei precedenti esempi.

In pratica i rendimenti di un attivo o di un portafoglio sono lontani dal caso iid, ma se si conosce la media condizionata prevista $E(r_{N+1} | \mathfrak{F}_N)$ e la varianza condizionata prevista $Var(r_{N+1} | \mathfrak{F}_N)$ è plausibile supporre che il

processo condizionato

$$\left(a_{N+1}|\mathfrak{F}_N\right) = \frac{\left(r_{N+1}|\mathfrak{F}_N\right) - E\left(r_{N+1}|\mathfrak{F}_N\right)}{\left[\text{Var}\left(r_{N+1}|\mathfrak{F}_N\right)\right]^{1/2}} \quad (4)$$

sia, praticamente, iid e pertanto è possibile ottenere il VaR al tempo N, con livello di probabilità α e ad un orizzonte temporale N+1 come:

$$\begin{aligned} \text{VaR}_N(1; \alpha) &= E\left(r_{N+1}|\mathfrak{F}_N\right) + a_{(\alpha)} \left[\text{Var}\left(r_{N+1}|\mathfrak{F}_N\right)\right]^{1/2} = \\ &= r_N(1) + a_{(\alpha)} \sigma_N(1) \end{aligned}$$

ove $a_{(\alpha)} = F_a^{-1}(\alpha)$ è il percentile di ordine α del processo stocastico condizionato e standardizzato a_t con distribuzione di probabilità $F_a(x)$.

Per calcolare il VaR è necessario stimare, a partire dal campione osservato sui rendimenti dell'attivo (r_1, r_2, \dots, r_N) , le tre quantità

$$r_N(1); \quad a_{(\alpha)}; \quad \sigma_N(1)$$

inoltre, queste devono essere generalizzate per calcolare il $\text{VaR}_N(h; \alpha)$, cioè il *valore a rischio* al tempo N+h, con $h \geq 1$, utilizzando le informazioni \mathfrak{F}_N disponibili al tempo N. Per calcolare $\text{VaR}_N(h; \alpha)$ osserviamo che:

- il rendimento r_t è riferito all'intervallo temporale $[t-1, t)$ e precisamente $r_t = p_t - p_{t-1}$, con $p_t = \log(P_t)$, e quindi la previsione ottimale del rendimento r_{N+j} riferito all'intervallo temporale unitario $[N+j-1, N+j)$ è data da

$$r_N(j) = E\left(r_{N+j}|\mathfrak{F}_N\right)$$

- il rendimento riferito nell'intervallo temporale $[N, N+h)$, e quindi

composto da $h \geq 1$ tempi di osservazione, lo indicheremo con $r_{N+[h]}$ e di conseguenza è dato da

$$r_{N+[h]} = \sum_{j=1}^h r_{N+j}$$

- se indichiamo con $r_N[h]$ la previsione ottimale di $r_{N+[h]}$, risulterà

$$r_N[h] = E(r_{N+[h]} | \mathfrak{F}_N) = \sum_{j=1}^h r_N(j)$$

- di conseguenza poniamo anche

$$\sigma_N^2[h] = \text{Var}(r_{N+[h]} | \mathfrak{F}_N)$$

In definitiva, utilizzando queste notazioni e considerazioni il *valore a rischio*, riferito al tempo $N+h$ (cioè h tempi in avanti), calcolato con le informazioni fino al tempo N , è dato da

$$\begin{aligned} \text{VaR}_N(h; \alpha) &= E(r_{N+[h]} | \mathfrak{F}_N) + a_{(\alpha)} \left[\text{Var}(r_{N+[h]} | \mathfrak{F}_N) \right]^{1/2} = \\ &= r_N[h] + a_{(\alpha)} \sigma_N[h] \end{aligned} \quad (5)$$

Per calcolare concretamente la (5) è necessario derivare analiticamente e quindi stimare dai dati le tre quantità $r_N[h]$, $a_{(\alpha)}$, $\sigma_N[h]$. Di questa problematica ci occuperemo nei paragrafi seguenti.

La (5) è stata derivata per valutare il rischio di un singolo attivo. Con un ragionamento simile può essere utilizzata per valutare il rischio di un portafoglio di attivi o di un intero mercato. Naturalmente questo è possibile se non si è interessati alle interrelazioni che intercorrono fra gli attivi che compongono il portafoglio e la composizione di questo non muta nell'arco di tempo analizzato. In caso contrario è necessario ricorrere ad una analisi multivariata che prenda esplicitamente in considerazione i legami esistenti fra i diversi attivi.

I problemi statistici che implica la derivazione del VaR con la (5) sono

numerosi e possono essere affrontati sia in un ambito non parametrico (come visto nel capitolo II) che parametrico. In particolare è necessario:

- (a) prevedere in modo efficiente media e varianza condizionate dei rendimenti dell'attivo o del portafoglio in analisi riferiti all'intervallo temporale $[N, N + h]$;
- (b) costruire una stima efficiente della standardizzata definita nella (4), verificarne la sua compatibilità con l'ipotesi di dati iid e stimarne la relativa distribuzione di probabilità;
- (c) calcolare il VaR per $h \geq 1$ passi in avanti e derivarne, eventualmente, un intervallo di confidenza.

Nei prossimi paragrafi affronteremo le problematiche sopra enumerate.

6.3 Il VaR di Riskmetrics

E' il modo più semplice per derivare il VaR dato che il modello su cui è calcolato è il seguente:

$$\begin{cases} r_t \stackrel{\text{iid}}{\sim} N(0; \sigma_r^2) \\ r_t = \sigma_t a_t; \quad a_t \stackrel{\text{iid}}{\sim} N(0;1) \\ \sigma_t^2 = \beta \sigma_{t-1}^2 + (1-\beta) r_{t-1}^2; \quad 0 < \beta < 1. \end{cases} \quad (6)$$

In altri termini si suppone che i rendimenti siano iid e distribuiti come una Normale, mentre la volatilità è descritta da un IGARCH(1,1) con $w = 0$. In questo caso risulta

$$r_{N+[h]} = p_{N+h} - p_N = r_{N+h} + r_{N+h-1} + \dots + r_{N+1}$$

e per la prima delle (6) si ottiene

$$\begin{aligned} r_N[h] &= E(p_{N+h} - p_N | \mathfrak{F}_N) = E(r_{N+h} + r_{N+h-1} + \dots + r_{N+1} | \mathfrak{F}_N) = \\ &= E(r_{N+h} | \mathfrak{F}_N) + E(r_{N+h-1} | \mathfrak{F}_N) + \dots + E(r_{N+1} | \mathfrak{F}_N) = 0 \end{aligned}$$

Inoltre, per quanto detto a proposito dei modelli IGARCH(1,1) con $w = 0$, si

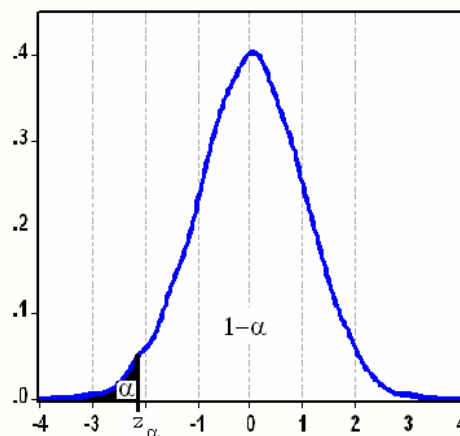
ha

$$\begin{aligned} \sigma_N^2[h] &= \text{Var}(r_{N+h} + r_{N+h-1} + \dots + r_{N+1} | \mathfrak{F}_N) = \\ &= \sigma_N^2(h) + \sigma_N^2(h-1) + \dots + \sigma_N^2(1) = h\sigma_N^2(1) \end{aligned}$$

ed il VaR calcolato h tempi in avanti, a partire da N, diventa

$$\text{VaR}_N(h; \alpha) = z_{(\alpha)} \sqrt{h(\beta\sigma_N^2 + (1-\beta)r_N^2)} \quad (7)$$

Figura 2 – Percentile di ordine α della Normale standardizzata



dove $z_{(\alpha)}$ è il percentile della Normale standardizzata che stacca alla sua sinistra la probabilità α e si è tenuto conto che, come abbiamo visto nelle previsioni dei modelli IGARCH(1,1), si ha

$$\sigma_N^2(1) = \beta\sigma_N^2 + (1-\beta)r_N^2$$

Come si può osservare dalla figura 2, il percentile $z_{(\alpha)}$ è sempre negativo visto che è calcolato sulla coda sinistra (le perdite) della Normale standardizzata.

In pratica le quantità (β, σ_N^2) sono incognite ed a queste vanno sostituite quelle ottenute dalle stime del modello IGARCH(1,1). Inoltre per $\alpha = 0.05$ si

ha $z_{(0.05)} = -1.645$, mentre per $\alpha = 0.01$ è $z_{(0.01)} = -2.330$ e le formule finali che vengono utilizzate per il calcolo del VaR nell'ottica di Riskmetrics sono:

$$\begin{cases} \text{VaR}_N(h;0.05) = -1.645\sqrt{h} \left(\hat{\beta}\hat{\sigma}_N^2 + (1-\hat{\beta})r_N^2 \right)^{1/2} \\ \text{VaR}_N(h;0.01) = -2.33\sqrt{h} \left(\hat{\beta}\hat{\sigma}_N^2 + (1-\hat{\beta})r_N^2 \right)^{1/2} \end{cases} \quad (8)$$

Nel caso in cui la prima delle (6) sia ancora iid ma con $r_t = \mu + \varepsilon_t$, si ha

$$\hat{\mu} = \bar{r} = \frac{1}{N} \sum_{t=1}^N r_t; \quad \hat{\varepsilon}_t = r_t - \bar{r}$$

e le (8) diventano

$$\begin{cases} \text{VaR}_N(h;0.05) = h\bar{r} - 1.645\sqrt{h} \left(\hat{\beta}\hat{\sigma}_N^2 + (1-\hat{\beta})\hat{\varepsilon}_N^2 \right)^{1/2} \\ \text{VaR}_N(h;0.01) = h\bar{r} - 2.33\sqrt{h} \left(\hat{\beta}\hat{\sigma}_N^2 + (1-\hat{\beta})\hat{\varepsilon}_N^2 \right)^{1/2} \end{cases} \quad (9)$$

ed anche queste si calcolano abbastanza facilmente dai dati osservati. Notiamo che la caratteristica principale del VaR secondo Riskmetrics è la sua estrema semplicità di calcolo. Naturalmente se i dati osservati si comportano in modo molto diverso da quanto ipotizzato nelle (6) la stima del VaR secondo Riskmetrics può essere molto diversa da quella vera, in particolare per orizzonti temporali $h > 1$.

Esempio 1

Calcoliamo il VaR, con $h=1, 2, \dots, 10$, per gli attivi i Campari, Fiat, Mibtel e Banca Intesa S. Paolo osservati giornalmente nel periodo 01/01/03 – 04/03/09 utilizzando la procedura Riskmetrics sopra descritta. Nella tabella 1 sono riportate le stime dei parametri $(\mu, \beta, \sigma_N^2, \varepsilon_N)$ necessarie per poter derivare tramite la (9) il relativo Valore a Rischio. Nella tabella 2 è riportato il VaR per $N=1611$ ed $h=1, 2, \dots, 10$ sia con $\alpha = 0.05$ che con $\alpha = 0.01$. I VaR dei quattro attivi così calcolati sono riportati nella figura 3 in modo da cogliere visivamente le differenti rischiosità esistenti fra i quattro attivi considerati.

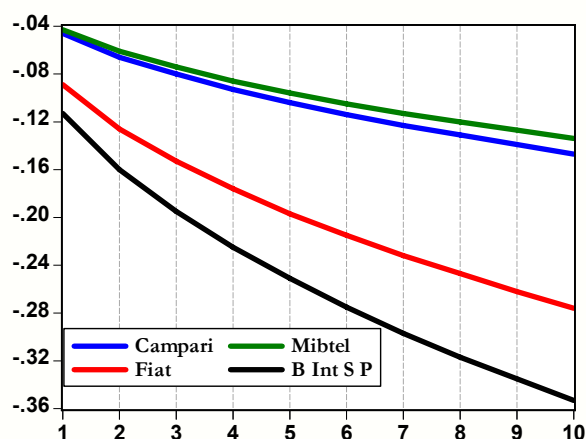
Tabella 1 – *Stime dei parametri del modello IGARCH per gli attivi giornalieri Campari, Fiat, Mibtel, Banca Intesa S. Paolo*

Parametri	Campari	Fiat	Mibtel	B. Int S. P.
\bar{r}	0.000613	0.000880	0.000405	0.00076
$\hat{\beta}$	0.9693	0.9464	0.9303	0.9235
$\hat{\sigma}_N^2$	0.000686	0.0027815	0.0007374	0.003462
$\hat{\epsilon}_N$	0.064392	0.0819808	0.016413	0.145194

Tabella 2 – $Var_N(h;0.05)$ per gli attivi Campari, Fiat, Mibtel, Banca Intesa S. Paolo

h	Campari	Fiat	Mibtel	B. Int S. P.
1	-0.046	-0.089	-0.043	-0.113
2	-0.066	-0.126	-0.061	-0.160
3	-0.080	-0.153	-0.074	-0.195
4	-0.093	-0.176	-0.086	-0.225
5	-0.104	-0.197	-0.096	-0.251
6	-0.114	-0.215	-0.105	-0.275
7	-0.123	-0.232	-0.113	-0.297
8	-0.131	-0.247	-0.120	-0.317
9	-0.139	-0.262	-0.127	-0.335
10	-0.147	-0.276	-0.134	-0.353

Figura 3 – Var per $\alpha = 0.05$ di Campari, Fiat, Mibtel, B. Intesa S. Paolo



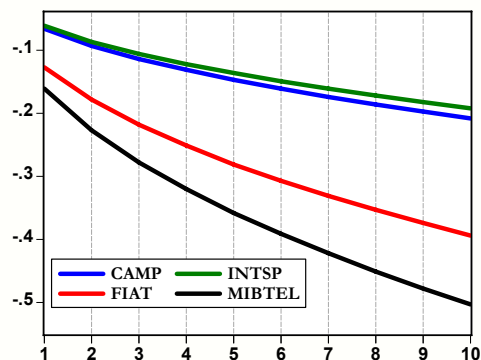
Come si può osservare dalla tabella 2 e dalla figura 3 e come atteso, dei quattro attivi analizzati quello meno rischioso è il Mibtel che descrive l'andamento medio della Borsa di Milano e quindi del mercato mobiliare italiano. La rischiosità di Mibtel passa dal 4.3% ad un giorno, fino al 13.4% a dieci giorni. Delle altre tre società considerate quella più rischiosa, come plausibile visto che le previsioni sono relative al marzo 2009 e quindi in piena crisi finanziaria, è Banca Intesa San Paolo il cui VaR va dall'11.3% ad un giorno, al 35.3% a dieci giorni e quindi circa tre volte di quello Mibtel. In altri termini, per un investitore in azioni di Banca Intesa S. Paolo era prevista il 4 marzo 2009, a dieci giorni, una perdita massima di più di un terzo del capitale investito e questo con una probabilità del 5%.

Nella tabella 3, utilizzando la stessa metodologia, ma con $\alpha = 0.01$, è riportato il VaR dei quattro attivi su menzionati, e quindi ripreso nella figura 4.

Tabella 3 – $\text{VaR}_N(h;0.01)$ per gli attivi Campari, Fiat, Mibtel, Banca Intesa S. Paolo

h	Campari	Fiat	Mibtel	B. Int S. P.
1	-0.066	-0.127	-0.061	-0.161
2	-0.093	-0.178	-0.087	-0.227
3	-0.114	-0.218	-0.106	-0.278
4	-0.131	-0.251	-0.122	-0.320
5	-0.147	-0.281	-0.136	-0.358
6	-0.161	-0.307	-0.149	-0.391
7	-0.174	-0.331	-0.161	-0.422
8	-0.186	-0.353	-0.172	-0.451
9	-0.197	-0.374	-0.182	-0.478
10	-0.208	-0.394	-0.192	-0.503

Figura 4 – VaR per $\alpha = 0.01$ di Campari, Fiat, Mibtel, B. Intesa S. Paolo



Confrontando la figura 3 con la 4 si può osservare che l'andamento del VaR è rimasto lo stesso, ma è quasi raddoppiato il rischio di perdita.

□

6.4 Il VaR per modelli ARMA-GARCH

La derivazione del VaR con la procedura Riskmetrics è molto semplice anche se approssimata. Se si vuole un valore del VaR più preciso è necessario un modello più flessibile, ma anche più complesso. Qui di seguito accenniamo al calcolo del VaR quando ai rendimenti è adattato un modello ARMA(p,q) – GARCH(s,v):

$$\begin{cases} \phi(B)r_t = \phi_0 + \theta(B)\varepsilon_t \\ \varepsilon_t = \sigma_t a_t ; \quad a_t \stackrel{iid}{\sim} (0,1) \\ \alpha(B)\sigma_t^2 = w + \beta(B)\varepsilon_t^2 \end{cases} \quad (10)$$

dove si è posto

$$\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p ; \quad \theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$$

$$\alpha(B) = 1 - \alpha_1 B - \dots - \alpha_s B^s ; \quad \beta(B) = \beta_1 B + \dots + \beta_v B^v$$

Sotto le usuali ipotesi di stazionarietà, invertibilità ed esistenza dei processi stocastici coinvolti, si ha ancora

$$r_N[h] = \sum_{j=1}^h r_N(j)$$

con $r_N(j)$ il previsore di r_{N+j} . Per il previsore di σ_{N+h}^2 risulta

$$\begin{aligned} \sigma_N^2[h] &= \text{Var}(r_{N+h} | \mathfrak{F}_N) = \text{Var}\left(\sum_{j=1}^h e_N(j) | \mathfrak{F}_N\right) = \\ &= \sigma_N^2(h) + (1 + \psi_1)^2 \sigma_N^2(h-1) + (1 + \psi_1 + \psi_2)^2 \sigma_N^2(h-2) + \end{aligned}$$

$$+ \dots + \left(\sum_{j=0}^{h-1} \psi_j \right)^2 \sigma_N^2(1) = \sum_{j=0}^{h-1} \sigma_N^2(h-j) \left(\sum_{i=0}^j \psi_i \right)^2$$

ove si è posto

$$e_N(j) = r_{N+j} - r_N(j) = \sum_{i=0}^{j-1} \psi_i \varepsilon_{N+j-i}$$

$$\psi_0 = 1; \quad 1 + \psi_1 B + \psi_2 B^2 + \psi_3 B^3 + \dots = \frac{\theta(B)}{\phi(B)}$$

mentre $\sigma_N^2(j)$ è definita nella (21) del capitolo V e ripresa nella (12) seguente. Sostituendo queste espressioni nella (5) si ottiene il VaR

$$\text{VaR}_N(h; \alpha) = r_N[h] + a_{(\alpha)} \sigma_N[h] \quad (11)$$

con $a_{(\alpha)}$ l' α -simo percentile della distribuzione di a_t . In generale $\sigma_N[h]$ figurante nella (11) non ammette una espressione esplicita, se si escludono casi particolari in cui si riesce facilmente a derivare una espressione dei pesi ψ_j , ed il suo calcolo viene fatto con metodi numerici o con approssimazioni. Qui di seguito illustriamo, per l'appunto, alcuni casi particolari di rilevanza pratica che ammettono una formula esplicita per i ψ_j e quindi per $\sigma_N[h]$.

(A) IL MODELLO MA(0)-GARCH(1, 1)

In questo caso si suppone che i rendimenti siano ben rappresentati dal modello

$$r_t = \mu + \varepsilon_t; \quad \sigma_t^2 = w + \alpha \varepsilon_t^2 + \beta \sigma_{t-1}^2$$

che implica

$$r_N(j) = \mu; \quad r_N[h] = h\mu; \quad \psi_j = 0, j=1, 2, \dots; \quad \sigma_N^2[h] = \sum_{j=1}^h \sigma_N^2(j), h=1, 2, \dots$$

con

$$\sigma_N^2(j) = \begin{cases} w + \alpha \varepsilon_N^2 + \beta \sigma_N^2 & \text{per } j = 1 \\ w + (\alpha + \beta) \sigma_N^2(j-1) & \text{per } j > 1 \end{cases} \quad (12)$$

La (12), stimati i parametri (μ, α, β) , permette di ottenere una stima della (11):

$$\begin{aligned} \text{VaR}_N(h, \alpha) &= \\ &= \begin{cases} \bar{r} + a_{(\alpha)} \left(\hat{w} + \hat{\alpha} \varepsilon_N^2 + \hat{\beta} \hat{\sigma}_N^2 \right)^{1/2} & \text{per } h = 1 \\ h\bar{r} + a_{(\alpha)} \left(h\hat{w} + \hat{\alpha} \varepsilon_N^2 + \hat{\beta} \hat{\sigma}_N^2 + (\hat{\alpha} + \hat{\beta}) \sum_{j=1}^{h-1} \hat{\sigma}_N^2(j) \right)^{1/2} & \text{per } h > 1 \end{cases} \end{aligned} \quad (13)$$

Nella (13) rimane comunque da stimare $a_{(\alpha)}$. Di questa problematica tratteremo nel prossimo paragrafo.

Esempio 2

Consideriamo i rendimenti giornalieri Fiat trattati nell'esempio precedente, e calcoliamo il VaR adattandovi il modello MA(0)-GARCH(1,1). Le stime dei parametri di interesse di tale modello sono:

$$\begin{aligned} \bar{r} = \hat{\mu} &= 0.000981; \quad \hat{w} = 0.00000999; \quad \hat{\alpha} = 0.0931; \quad \hat{\beta} = 0.892 \\ \hat{\varepsilon}_N &= 0.13565; \quad \hat{\sigma}_N^2 = 0.0019624 \end{aligned}$$

Tabella 4 – VaR per Fiat con $\alpha = 0.05$

h	\bar{r}	$\hat{\sigma}_N^2(j)$	$\hat{\sigma}_N^2[h]$	VaR(h)
1	0.000981	0.003474	0.003474	-0.096
2	0.000981	0.003432	0.006906	-0.136
3	0.000981	0.003390	0.010296	-0.166
4	0.000981	0.003349	0.013644	-0.191
5	0.000981	0.003308	0.016952	-0.213
6	0.000981	0.003268	0.020220	-0.233
7	0.000981	0.003229	0.023449	-0.251
8	0.000981	0.003190	0.026639	-0.268
9	0.000981	0.003152	0.029791	-0.283
10	0.000981	0.003114	0.032905	-0.297

Queste stime, sotto l'ipotesi che a_t sia un processo iid Normale e quindi risulti $a_{(0.05)} = -1.645$, sostituite nella (13) ci hanno permesso di costruire nell'ultima colonna della tabella 4 il relativo VaR. Dato che il modello MA(0)-GARCH(1,1) utilizzato per Fiat ha parametri molto simili a quello dell'IGARCH(1,1) utilizzato da Riskmetrics, i risultati della tabella 5 sono molto simili a quelli riportati nella tabella 2. □

(B) IL MODELLO MA(1)-GARCH(1,1)

In questo caso si suppone che il modello che adatta i rendimenti sia dato da

$$r_t = \mu + \varepsilon_t - \theta\varepsilon_{t-1}; \quad \sigma_t^2 = w + \alpha\varepsilon_t^2 + \beta\sigma_{t-1}^2$$

da cui si deriva

$$r_{N+[h]} = h\mu - \theta\varepsilon_N + (1-\theta)\sum_{j=1}^{h-1}\varepsilon_{N+j} + \varepsilon_{N+h}$$

e quindi

$$r_N[h] = E(r_{N+[h]}|\mathfrak{F}_N) = h\mu - \theta\varepsilon_N, \quad h=1, 2, \dots$$

che implica

$$\sum_{j=1}^h e_N(j) = r_{N+[h]} - r_N[h] = \varepsilon_{N+h} + (1-\theta)\sum_{j=1}^{h-1}\varepsilon_{N+j}$$

da cui si ottiene

$$\sigma_N^2[h] = \text{Var}(r_{N+[h]}|\mathfrak{F}_N) = \sigma_N^2(h) + (1-\theta)^2 \sum_{j=1}^{h-1} \sigma_N^2(j), \quad h=1, 2, \dots$$

con $\sigma_N^2(j)$ definito nella (12). Questi risultati sostituiti nella (11) permettono di calcolare il VaR dopo aver definito una stima del percentile $a_{(\alpha)}$.

(C) IL MODELLO MA(Q)-GARCH(1,1)

Generalizziamo il caso precedente supponendo che sia

$$r_t = \mu + \varepsilon_t - \theta_1\varepsilon_{t-1} - \dots - \theta_q\varepsilon_{t-q}; \quad \sigma_t^2 = w + \alpha\varepsilon_t^2 + \beta\sigma_{t-1}^2$$

da cui si deriva

$$r_N(j) = \mu - \sum_{i=0}^{q-j} \theta_{i+j} \varepsilon_{N-i}, \quad \text{per } j=1, 2, \dots, q$$

$$r_N(q+j) = \mu, \quad \text{per } j>0$$

$$\psi_j = \begin{cases} -\theta_j & \text{per } j = 1, 2, \dots, q \\ 0 & \text{per } j = q+1, q+2, \dots \end{cases}$$

che permettono di ottenere una stima del VaR particolarizzando la (11) ed ottenendo:

$$r_N[h] = h\mu - \sum_{j=1}^h \sum_{i=0}^{q-j} \theta_{i+j} \varepsilon_{N-i}$$

$$\sigma_N^2[h] = \sigma_N^2(h) + \sum_{j=1}^{h-1} \sigma_N^2(h-j) \left(1 - \sum_{i=1}^j \theta_i \right)^2$$

con $\theta_{i+q} = 0$, per $i=1, 2, \dots$, e $\sigma_N^2(j)$ definita nella (12). Nelle pratiche applicazioni è spesso $\psi_j \approx 0$ e le (13) forniscono risultati approssimati apprezzabili in ogni caso.

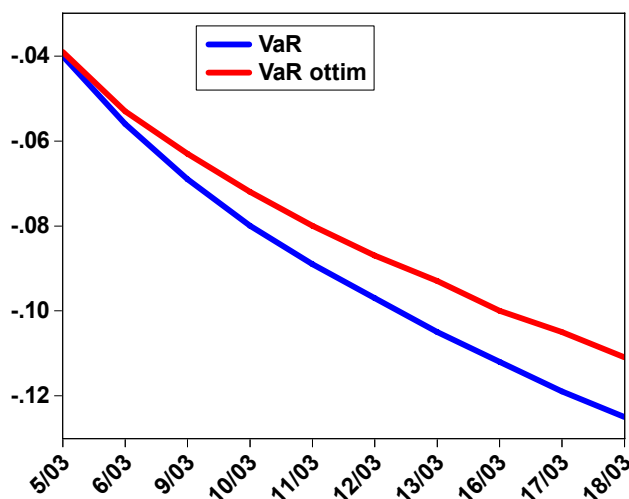
Esempio 3

Calcoliamo il VaR per Fiat e Campari utilizzando sia un MA(0)-GARCH(1,1) che il più generale MA(q)-GARCH(1, 1) ottimale, sotto l'ipotesi di normalità di a_t , in modo da verificare il grado di approssimazione della (13). Per Fiat il modello ottimale, nell'ambito della classe degli MA(q)-GARCH(1,1), è MA(0)-GARCH(1,1) e la stima del VaR ottenuta con la (13) è quella ottimale. Per Campari la stima ottimale, nell'ambito della classe MA(q)-GARCH(1, 1), si ottiene per $q=1$, $\phi_0 = \mu = 0$, $\hat{\theta} = 0.1137$. In questo caso il VaR ottenuto con il modello approssimato MA(0)-GARCH(1, 1) e quello ottimale MA(1)-GARCH(1, 1) portano a risultati leggermente diversi. La differenza massima è dell'1.4% per $h=10$, come evidenziato nelle ultime due colonne tabella 5 ottenuta con $\alpha = 0.05$ e sotto l'ipotesi di normalità di a_t .

Tabella 5 – *VaR per Campari con il modello MA(0)-GARCH(1,1) e quello ottimale MA(1)-GARCH(1,1) con $\alpha = 0.05$*

h	$r_N(j)$	$r_N(j)$ Ottimo	$\sigma_N^2(j)$	$\sigma_N^2(j)$ Ottimo	$r_N[j]$	$r_N[j]$ Ottimo	$\sigma_N^2[j]$	$\sigma_N^2[j]$ Ottimo	VaR _N	VaR _N Ottimo
0	0.0252	0.0198	0.000586	0.000578						
1	0.0000	-0.0022	0.000587	0.000568	0.0000	-0.0022	0.0006	0.00058	-0.040	-0.039
2	0.0000	0.0000	0.000585	0.000566	0.0000	-0.0022	0.0012	0.00102	-0.056	-0.053
3	0.0000	0.0000	0.000583	0.000565	0.0000	-0.0022	0.0018	0.00146	-0.069	-0.063
4	0.0000	0.0000	0.000582	0.000563	0.0000	-0.0022	0.0023	0.00191	-0.080	-0.072
5	0.0000	0.0000	0.000580	0.000561	0.0000	-0.0022	0.0029	0.00235	-0.089	-0.080
6	0.0000	0.0000	0.000579	0.000560	0.0000	-0.0022	0.0035	0.00279	-0.097	-0.087
7	0.0000	0.0000	0.000577	0.000558	0.0000	-0.0022	0.0041	0.00322	-0.105	-0.093
8	0.0000	0.0000	0.000575	0.000556	0.0000	-0.0022	0.0046	0.00366	-0.112	-0.100
9	0.0000	0.0000	0.000574	0.000555	0.0000	-0.0022	0.0052	0.00410	-0.119	-0.105
10	0.0000	0.0000	0.000572	0.000553	0.0000	-0.0022	0.0058	0.00453	-0.125	-0.111

Figura 5 – *VaR per Campari con il modello MA(0)-GARCH(1,1) e quello ottimale*



Come si può osservare dalle ultime due colonne della tabella 6 e dalla figura 5 il VaR ottenuto con il modello approssimato MA(0)-GARCH(1, 1) è leggermente più elevato di quello ottenibile con il modello ottimale per valori grandi di h.

Tabella 6 – VaR di Campari calcolato con il modello MA(0)-GARCH(1,1) e quello di MA(1)-GARCH(1,1)-M con $\alpha = 0.05$

h	$r_N(h)$	$\sigma_N^2(h)$	$r_N[h]$	$\sigma_N^2[h]$	$VaR_N - M(h)$	$VaR_N(h)$
0	0.02519	0.000582				
1	-0.00166	0.000571	-0.00166	0.00064	-0.043	-0.039
2	0.00055	0.000569	-0.00111	0.00114	-0.057	-0.053
3	0.00055	0.000567	-0.00056	0.00164	-0.067	-0.063
4	0.00055	0.000565	-0.00002	0.00214	-0.076	-0.072
5	0.00055	0.000563	0.00053	0.00264	-0.084	-0.080
6	0.00055	0.000562	0.00108	0.00313	-0.091	-0.087
7	0.00055	0.000560	0.00163	0.00363	-0.097	-0.093
8	0.00055	0.000558	0.00217	0.00412	-0.103	-0.100
9	0.00055	0.000556	0.00272	0.00461	-0.109	-0.105
10	0.00055	0.000554	0.00327	0.00510	-0.114	-0.111

Il VaR può essere anche calcolato nell'ipotesi che nei dati vi sia *premio al rischio* ed il modello generatore dei dati sia pertanto un MA(q)-GARCH(1,1)-M. Anche in questo caso i risultati che si ottengono non sono molto diversi da quelli sopra presentati. Nelle ultime due colonne della tabella 6 sono confrontati il VaR-M per Campari quando si usa il modello MA(1)-GARCH(1,1)-M ed il suo corrispondente, VaR ottenuto con MA(1)-GARCH(1,1) e riportato nella tabella 5. In questo caso le differenze sono praticamente trascurabili.

□

(D) IL MODELLO AR(1)-GARCH(1,1)

In questo caso si suppone che il modello per i rendimenti sia

$$r_t = \phi_o + \phi r_{t-1} + \varepsilon_t; \quad \sigma_t^2 = w + \alpha \varepsilon_t^2 + \beta \sigma_{t-1}^2$$

da cui si ottiene

$$r_N(j) = \phi_o + \phi r_N(j-1); \quad r_N[h] = h\phi_o + \phi \sum_{j=1}^h r_N(j-1)$$

con $r_N(0) = r_N$; $\phi_o = \mu(1 - \phi)$; $\psi_j = \phi^j$ che implica

$$\sigma_N^2[h] = \frac{1}{1 - \phi^2} \sum_{j=1}^h (1 - \phi^{h-j+1}) \sigma_N^2(j)$$

dove $\sigma_N^2(j)$ è definito nella (12). Sostituendo ai parametri incogniti le stime ed utilizzando la (11) si ottiene il VaR. Anche in questo caso è necessario definire una stima di $a_{(\alpha)}$. I risultati qui presentati possono essere estesi al modello generale ARMA-GARCH-M, ma le espressioni formali diventano molto complesse e si preferisce, in questi rari casi, calcolare la stima del VaR per via numerica tramite software specialistici.

6.5 Un intervallo di confidenza per il VaR

E' possibile costruire un intervallo di confidenza asintotico per il VaR utilizzando i risultati sui percentili presentati nel paragrafo 6.2 ed in particolare la (3). Se il modello MA(q)-GARCH(1,1) utilizzato per rappresentare i rendimenti è adeguato, un intervallo di confidenza approssimato al livello del 95% per il VaR è dato da

$$\begin{aligned} \hat{\text{VaR}}_N(h; \alpha) - 2 \left[\frac{\alpha(1-\alpha)\hat{\sigma}_N^2[h]}{N f^2(a_{(\alpha)})} \right]^{1/2} &\leq \text{VaR}_N(h; \alpha) \leq \\ &\leq \hat{\text{VaR}}_N(h; \alpha) + 2 \left[\frac{\alpha(1-\alpha)\hat{\sigma}_N^2[h]}{N f^2(a_{(\alpha)})} \right]^{1/2} \end{aligned} \quad (14)$$

Ove si è tenuto conto che nella normale standardizzata il percentile corrispondente a $(0.05/2)$ è $z_{(0.05/2)} = -1.96 \approx -2$ e quello per $(1-0.05/2)$ è $z_{(1-0.05/2)} = 1.96 \approx 2$. Per poter applicare la (14) è necessario, fissato α , conoscere la funzione di densità $f(a)$ di a_t e quindi derivare $a_{(\alpha)}$. Sotto l'ipotesi che il processo dei residui standardizzati a_t sia, oltre che iid, Normale e per $\alpha = 0.05$ si ha $a_{(0.05)} = 1.645$ mentre risulta

$$f^2(1.645) = \frac{1}{2\pi} \left(\exp \left\{ -\frac{(1.645)^2}{2} \right\} \right)^2 = 0.01063696$$

e la (14) diventa

$$\begin{aligned} \hat{\text{VaR}}_N(h;0.05) - 4.2264 \frac{\hat{\sigma}_N[h]}{\sqrt{N}} &\leq \text{VaR}_N(h;0.05) \leq \\ &\leq \hat{\text{VaR}}_N(h;0.05) + 4.2264 \frac{\hat{\sigma}_N[h]}{\sqrt{N}} \end{aligned} \quad (15)$$

Un risultato analogo si ottiene per $\alpha = 0.01$ con la costante moltiplicativa che in questo caso è pari a 7.4647. Naturalmente un risultato simile può essere calcolato sotto l'ipotesi che il processo dei residui standardizzati sia distribuito proporzionalmente alla v.c. t di Student standardizzata $((g - 2) / g)^{1/2} t_{(g)}$ con g stimato dai dati.

Esempio 4

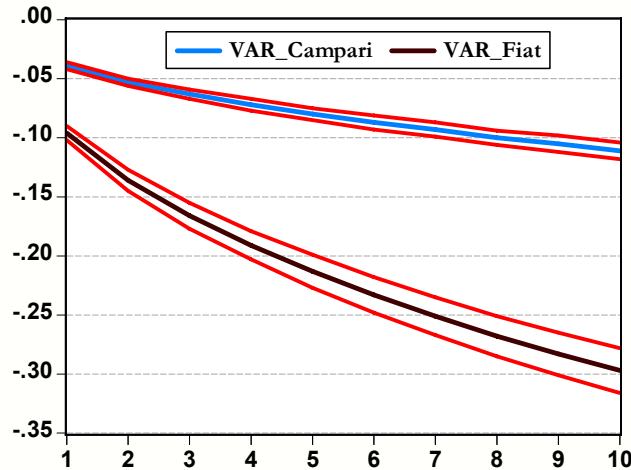
Calcoliamo un intervallo di confidenza al 95% per il VaR riferito ai rendimenti di Campari e Fiat utilizzando il modello ottimale della classe MA(q)-GARCH(1,1). Dai dati delle tabelle 5 e 6 e la (15) si ottengono i risultati della tabella 7.

Tabella 7 – Intervalli di confidenza al 95% del $\text{VaR}_N(h;0.05)$ per Campari e Fiat

Fiat			Campari		
VaR	VaR _{max}	VaR _{min}	VaR	VaR _{max}	VaR _{min}
-0.096	-0.102	-0.090	-0.039	-0.042	-0.036
-0.136	-0.145	-0.127	-0.053	-0.056	-0.050
-0.166	-0.177	-0.155	-0.063	-0.067	-0.059
-0.191	-0.203	-0.179	-0.072	-0.077	-0.067
-0.213	-0.227	-0.199	-0.080	-0.085	-0.075
-0.233	-0.248	-0.218	-0.087	-0.093	-0.081
-0.251	-0.267	-0.235	-0.093	-0.099	-0.087
-0.268	-0.285	-0.251	-0.100	-0.106	-0.094
-0.283	-0.301	-0.265	-0.105	-0.112	-0.098
-0.297	-0.316	-0.278	-0.111	-0.118	-0.104

Notiamo che il titolo Fiat, non solo è molto più rischioso di Campari, ma ha anche un intervallo di confidenza nettamente più ampio di quello di Campari, come è illustrato dalla figura 6.

Figura 6 – VaR per Fiat e Campari e relativo intervallo di confidenza al 95%



□

6.6 Stima del percentile $a_{(\alpha)}$ e calcolo del VaR

Nelle pagine precedenti abbiamo più volte accennato al fatto che per calcolare il VaR è necessario derivare una stima del percentile $a_{(\alpha)}$. Qui di seguito riportiamo i metodi più utilizzati nelle pratiche applicazioni per derivarlo. Nel seguito tratteremo esplicitamente degli MA(q)-GARCH(1, 1), ma quanto detto può essere facilmente esteso alle varianti della classe GARCH che permettono di catturare sia il *leverage*, che l'eventuale *premio al rischio*.

(A) IL VAR CON RESIDUI NORMALI

Come già accennato, uno dei modi per calcolare il VaR è supporre che il processo standardizzato $a_t = \varepsilon_t / \sigma_t$ sia iid e Normale, in queste condizioni risulta:

$$\begin{cases} \text{VaR}_N(h; 0.05) = \hat{r}[h] - 1.645 \hat{\sigma}_N[h] \\ \text{VaR}_N(h; 0.01) = \hat{r}[h] - 2.330 \hat{\sigma}_N[h] \end{cases} \quad (16)$$

D'altra parte, per poter correttamente applicare questa procedura è necessario verificare statisticamente che l'ipotesi di normalità sia accettabile. Infatti, molto

spesso, pur essendo accettabile l'ipotesi che il processo a_t sia simmetrico, standardizzato ed iid, non lo è quella di normalità visto che a_t possiede code molto più "pesanti" di quelle di una normale standardizzata e molto più simili a quelle di una t di Student. Per verificare che il processo dei residui a_t sia normale si utilizzano, dati i residui stimati $(\hat{a}_1, \hat{a}_2, \dots, \hat{a}_N)$ dal modello ARMA-GARCH o sue varianti, due strumenti già trattati nelle pagine precedenti:

- 1) il test di Jarque-Bera per a_t
- 2) il grafico Q-Qplot per \hat{a}_t

Per quel che riguarda l'ipotesi di indipendenza di a_t questa viene verificata costruendo i correlogrammi di \hat{a}_t e di \hat{a}_t^2 e verificando che le relative autocorrelazioni cadono tutte nell'intervallo di confidenza asintotico al 95% dato da $[-2/\sqrt{N}; 2/\sqrt{N}]$.

(B) IL VAR CON RESIDUI T DI STUDENT

Una alternativa più plausibile all'ipotesi di normalità per a_t è quella di supporre che sia proporzionale ad una v.c. t di Student con $g > 4$ gradi di libertà. Ricordando che per una v. c. $t_{(g)}$ è $E(t_{(g)}) = 0$ e $\text{Var}(t_{(g)}) = g/(g-2)$, deve essere $a_t = ((g-2)/g)^{1/2} t_{(g)}$ e la (11) diventa

$$\text{VaR}_N(h; \alpha) = r_N[h] + ((g-2)/g)^{1/2} t_{(g; \alpha)} \sigma_N[h] \quad (17)$$

dove $t_{(g; \alpha)}$ è l' α -simo percentile calcolato sulla v.c. t di Student con g gradi di libertà. Il parametro g deve essere esso stesso stimato dai dati. La stima di g può essere fatta con il metodo della massima verosimiglianza congiuntamente alla stima dei parametri del modello ARMA-GARCH o sue varianti, oppure successivamente utilizzando l'indice di curtosi stimato da \hat{a}_t tenendo conto che per una v.c. $t_{(g)}$ risulta $\gamma_2 = 6/(g-4)$ da cui si ottiene la stima, con il metodo dei momenti $\hat{g} = 4 + 6/\hat{\gamma}_2$. In particolare, se è $g=5$ sostituendo nella (17) si ha

- $(3/5)^{1/2} t_{(5;0.05)} = -1.561$, se si è posto $\alpha = 0.05$;
- $(3/5)^{1/2} t_{(5;0.01)} = -2.606$, se si è posto $\alpha = 0.01$;

che, sostituiti nella (16), permettono di ottenere la stima del VaR sotto l'ipotesi che i residui standardizzati siano t di Student con $g=5$ gradi di libertà:

$$\begin{cases} \hat{\text{VaR}}_N(h;0.05) = \hat{r}[h] - 1.561 \hat{\sigma}_N[h] \\ \hat{\text{VaR}}_N(h;0.01) = \hat{r}[h] - 2.606 \hat{\sigma}_N[h] \end{cases} \quad (18)$$

Nelle pratiche applicazioni le stime dei parametri del modello ARMA-GARCH, o di sue varianti, ottenute con il metodo della massima verosimiglianza sotto l'ipotesi di normalità sono molto simili a quelle ottenute sotto l'ipotesi di t di Student, mentre cambia in modo sostanziale il percentile $a_{(\alpha)}$ come si può osservare confrontando la (18) con la (16). Nell'ipotesi di normalità, a parità di altre condizioni, il VaR con $\alpha = 0.05$ è più elevato rispetto a quello t di Student, mentre avviene il viceversa se è $\alpha = 0.01$. Naturalmente, man mano che g cresce, di solito basta che sia $g > 10$, le stime dei parametri del modello ed il relativo VaR sono sempre più indistinguibili nei due casi sopra esaminati.

Sotto l'ipotesi che i residui standardizzati siano proporzionali ad una $t_{(5)}$ è possibile derivare l'intervallo di confidenza. In particolare, in questo caso si può verificare che è

$$f^2(1.561) = 0.0113056; \quad f^2(2.606) = 0.00033069$$

e quindi

$$\text{VaR}_N(h;0.05) - 4.0995 \frac{\hat{\sigma}_N[h]}{\sqrt{N}} \leq \text{VaR} \leq \text{VaR}_N(h;0.05) + 4.0995 \frac{\hat{\sigma}_N[h]}{\sqrt{N}}$$

$$\text{VaR}_N(h;0.01) - 10.943 \frac{\hat{\sigma}_N[h]}{\sqrt{N}} \leq \text{VaR} \leq \text{VaR}_N(h;0.05) + 10.943 \frac{\hat{\sigma}_N[h]}{\sqrt{N}}$$

che confermano quanto detto a proposito della (16) e della (18).

Esempio 5

Stimiamo il VaR per i rendimenti di Fiat con i dati giornalieri già analizzati

in esempi precedenti.

Tabella 8 – Modelli ottimali per i rendimenti giornalieri di Fiat

	Modello	Condizioni	g	AIC	SIC
Normale	MA(0)- GARCH(1,1)	$\phi_0 \neq 0$		-4.833284	-4.819908
	MA(0)- TGARCH(1,1,1)	$\phi_0 = 0$		-4.845152	-4.831775
	MA(2)- EARCH(1,1,1)	$\phi_0 = 0$ $\theta_1 = 0$		-4.841631	-4.824910
T di Student	MA(0)- GARCH(1,1)	$\phi_0 \neq 0$	6.16	-4.871129	-4.577520
	MA(0)- TGARCH(1,1,1)	$\phi_0 = 0$	6.67	-4.877892	-4.861172
	MA(2)- EARCH(1,1,1)	$\phi_0 = 0$ $\theta_1 = 0$	6.64	-4.874845	-4.854780

Alla serie di $N=1611$ osservazioni sono stati adattati i modelli delle classi MA(q)-GARCH, MA(q)-TGARCH, MA(q)-EGARCH sia sotto l'ipotesi che i residui standardizzati a_t siano normali, sia che siano $t_{(g)}$ con g ulteriore parametro da stimare. I modelli qui riportati sono quelli ottimali nel senso che: a) tutti i loro parametri sono significativamente diversi da zero; b) quelli riguardanti la volatilità rispettano le condizioni di accettabilità; c) le autocorrelazioni dei residui standardizzati e quelle dei loro quadrati ricadono tutte nel relativo intervallo di confidenza al 95%. Questi modelli, sia per il caso normale che quello t di Student, sono stati riportati nella tabella 9 assieme, nelle ultime due colonne, agli indici globali AIC e SIC. Come si può osservare dalla tabella 8 il modello ottimale, sia sotto l'ipotesi di normalità che in quello t di Student, è MA(0)-TGARCH(1,1,1) con $\phi_0 = 0$ le cui stime sono riportate nella tabella 9

Tabella 9 – Stime dei parametri del modello MA(0)-TGARCH(1,1,1) per Fiat

	Normale	p-value	T di Student	p-value
g			6.67	0.000
w	0.0000126	0.000	0.0000135	0.001
α	0.0368	0.002	0.0534	0.007
γ	0.0868	0.000	0.0872	0.002
β	0.8974	0.000	0.8825	0.000

Queste stime sono state utilizzate per derivare le previsioni dei livelli e della volatilità e quindi per costruire il relativo VaR. Nella tabella 10 è riportato il VaR con $\alpha = 0.05$ ed $\alpha = 0.01$ sia sotto l'ipotesi di normalità che di t di Student.

Osserviamo che nel caso qui analizzato è $6 < \hat{g} = 6.67 < 7$ ed il percentile $a_{(\alpha)}$ è stato derivato come media fra il caso di $g=6$ e quello di $g=7$:

$$\hat{a}_{(\alpha)} = \frac{t_{(6;\alpha)}(4/6)^{1/2} + t_{(7;\alpha)}(5/7)^{1/2}}{2}$$

ottenendo

$$\hat{a}_{(0.05)} = -\frac{1.586 + 1.602}{2} = -1.594; \quad \hat{a}_{(0.01)} = -\frac{2.566 + 2.534}{2} = -2.550.$$

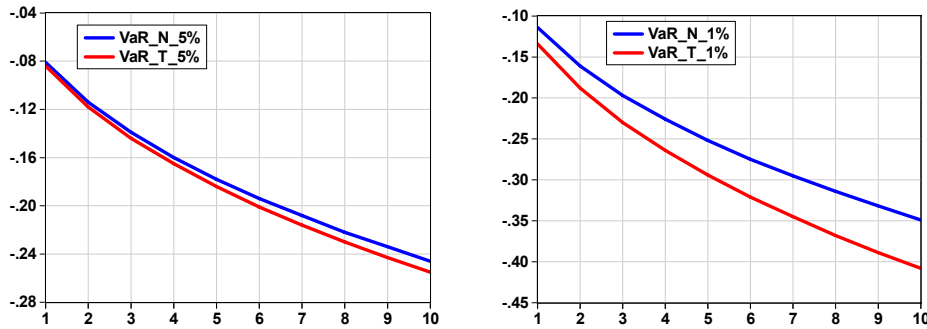
In realtà, con Eviews, può essere calcolato un percentile più preciso utilizzando il comando $@qtdist(df, \alpha)$, dove $df=6.67$ sono i gradi di libertà stimati ed α il percentile prefissato.

Tabella 10 – VaR per i rendimenti Fiat con residui normali e t di Student

h	Normale				T di Student			
	$\sigma_N^2(h)$	$\sigma_N^2[h]$	VaR _N $\alpha = .05$	VaR _N $\alpha = .01$	$\sigma_N^2(h)$	$\sigma_N^2[h]$	VaR _N $\alpha = .05$	VaR _N $\alpha = .001$
1	0.00241	0.00241	-0.081	-0.114	0.00275	0.00275	-0.084	-0.134
2	0.00237	0.00479	-0.114	-0.161	0.00270	0.00545	-0.118	-0.188
3	0.00233	0.00712	-0.139	-0.197	0.00266	0.00811	-0.144	-0.230
4	0.00229	0.00941	-0.160	-0.226	0.00262	0.01073	-0.165	-0.264
5	0.00225	0.01166	-0.178	-0.252	0.00258	0.01331	-0.184	-0.294
6	0.00222	0.01388	-0.194	-0.275	0.00254	0.01585	-0.201	-0.321
7	0.00218	0.01606	-0.208	-0.295	0.00250	0.01835	-0.216	-0.345
8	0.00214	0.01820	-0.222	-0.314	0.00246	0.02082	-0.230	-0.368
9	0.00211	0.02031	-0.234	-0.332	0.00243	0.02324	-0.243	-0.389
10	0.00207	0.02238	-0.246	-0.349	0.00239	0.02564	-0.255	-0.408

Come si può osservare dalla figura 7, in cui sono riportate le stime del VaR con residui Normali e t di Student e per $\alpha=0.05$ ed $\alpha=0.01$, il valore del VaR è sostanzialmente lo stesso se si considera $\alpha=0.05$, mentre è significativamente più elevato nel caso di residui t di Student se è $\alpha=0.01$.

Figura 7 – VaR di Fiat sotto l'ipotesi di residui normali e t di Student (in rosso) con $\alpha = 0.05$ e $\alpha = 0.01$ (grafico a destra)



□

(C) IL VAR E GLI ORDINAMENTI STATISTICI

Il percentile $a_{(\alpha)}$ può essere stimato in modo non parametrico utilizzando l'ordinamento statistico dei residui standardizzati \hat{a}_t , $t=1, 2, \dots, N$. Supponiamo che sui rendimenti (r_1, r_2, \dots, r_N) osservati su un attivo finanziario o monetario sia stato stimato un modello della classe ARMA-GARCH o sue varianti e di aver ottenute le stime dei residui standardizzati $(\hat{a}_1, \hat{a}_2, \dots, \hat{a}_N)$. Questi residui vengono ordinati in senso non decrescente ottenendo l'ordinamento statistico

$$\hat{a}_{(1)} \leq \hat{a}_{(2)} \leq \dots \leq \hat{a}_{(N)}$$

Fissato α , si scelgono $\hat{a}_{([N\alpha])}$ e $\hat{a}_{([N\alpha]+1)}$, cioè $\hat{a}_{(t)}$ di posto $[N\alpha]$ e $[N\alpha]+1$, dove con $[N\alpha]$ si è indicato il massimo intero di $N\alpha$. La stima del percentile $a_{(\alpha)}$ viene ottenuta come media aritmetica di questi due valori selezionati:

$$\hat{a}_{(\alpha)} = \frac{\hat{a}_{([N\alpha])} + \hat{a}_{([N\alpha]+1)}}{2} \tag{19}$$

Di solito la stima (19) di $a_{(\alpha)}$, o sue varianti più sofisticate, è poco efficiente per α piccolo, così come avviene nel calcolo del VaR.

Esempio 6

Ai rendimenti giornalieri di Fiat riferiti al periodo 02/01/03 – 04/03/09

adattiamo il modello ottimale con residui t di Student, costruiamo il VaR utilizzando gli ordinamenti statistici e lo confrontiamo con quello ottenuto usando i residui standardizzati proporzionali ad una v.c. t riportato nelle ultime due colonne della tabella 11. In questo caso, la numerosità di rendimenti è $N=1610$:

- per $\alpha = 0.05$ si ha

$$[N\alpha] = [1610 * 0.05] = 80; \quad [N\alpha] + 1 = [1610 * 0.05] + 1 = 80 + 1 = 81$$

ed ordinando i residui standardizzati del modello MA(0)-TGARCH(1,1,1) stimato nell'ipotesi t di Student si ha

$$\hat{a}_{(80)} = -1.6035; \quad \hat{a}_{(81)} = -1.5987$$

che implicano

$$\hat{a}_{(0.05)} = \frac{-1.6035 - 1.5987}{2} = -1.601$$

che è molto vicino a quello ottenuto nell'ipotesi di residui t di Student con $6 \leq g \leq 7$ utilizzato nella tabella 11 che, ricordiamo, era $\hat{a}_{(0.05)} = -1.594$. In tal caso il VaR è molto simile a quello ottenuto nella penultima colonna della tabella 11;

- per $\alpha = 0.01$ si ha

$$[N\alpha] = [1610 * 0.01] = 16; \quad [N\alpha] + 1 = [1610 * 0.01] + 1 = 17$$

e quindi

$$\hat{a}_{(16)} = -2.3561; \quad \hat{a}_{(17)} = -2.3292$$

che implicano

$$\hat{a}_{(0.01)} = \frac{-2.3561 - 2.3293}{2} = -2.343$$

In questo caso il VaR è sostanzialmente differente da quello ottenuto nell'ultima colonna della tabella 11. Questo è dovuto al fatto che α è molto piccolo e la stima ottenuta con gli ordinamenti statistici abbastanza imprecisa.

□

6.7 Il VaR nei modelli TMA-TGARCH

In alcuni degli esempi visti nel capitolo precedente abbiamo constatato empiricamente come gli attivi finanziari posseggono componenti a soglia nei livelli, ma queste sono molto più deboli nella volatilità. Quando i regimi esistono, questi hanno un impatto, più che sui parametri, sull'asimmetria e la curtosi. D'altro lato, l'esistenza di regimi ha a sua volta conseguenze immediate sul calcolo del VaR. Qui di seguito, per evitare eccessive complicazioni formali, analizziamo il VaR globale e quello locale, riferito ai singoli regimi, in un caso molto semplice con regimi sia nei livelli che nella volatilità. Infatti, il modello che analizzeremo è il TMA(2;1)-TGARCH(1,1,1) che prevede, nei livelli, due regimi che si comportano entrambi come un MA(1) e, nella volatilità, il semplice modello a soglia TGARCH. Più specificatamente si suppone che sia:

$$r_t = (\varepsilon_t + c_1 - \theta\varepsilon_{t-1})I_{t-1} + (\varepsilon_t + c_2 - \delta\varepsilon_{t-1})(1 - I_{t-1}); \text{ con } \varepsilon_t \sim \text{WN}(0; \sigma_\varepsilon^2)$$

$$\varepsilon_t = \sigma_t a_t; \quad \text{con } a_t \stackrel{\text{iid}}{\sim} (0;1)$$

$$\sigma_t^2 = \left(w + (\alpha + \gamma)\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2 \right) I_{t-1} + \left(w + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2 \right) (1 - I_{t-1})$$

dove si è posto

$$I_{t-1} = \begin{cases} 1 & \text{se } r_{t-1} < 0 \\ 0 & \text{se } r_{t-1} \geq 0 \end{cases}$$

Costruendo un apposito programma, per esempio in EViews, è possibile stimare i parametri del modello, costruire le previsioni dei livelli e della volatilità h passi in avanti e derivare il relativo VaR. La stima dei parametri avviene, in modo iterativo, con la stessa tecnica che abbiamo illustrato nel capitolo precedente. Qui di seguito deriviamo le previsioni dei livelli $r_N(j)$ e della volatilità $\sigma_N^2(j)$ e, fissato il livello di probabilità α ed il percentile $a_{(\alpha)}$, come illustrato nelle pagine precedenti, deriviamo il relativo $\text{VaR}_N(h; \alpha)$

$$\text{VaR}_N(h; \alpha) = r_N[h] + a_{(\alpha)}\sigma_N[h]$$

Ricordiamo che

$$r_{N+h} = r_{N+1} + r_{N+2} + \dots + r_{N+h}$$

$$r_N[h] = E(r_{N+h} | \mathfrak{F}_N) = r_N(1) + \dots + r_N(h); \quad \sigma_N^2[h] = \text{Var}(r_{N+h} | \mathfrak{F}_N)$$

Tenendo conto di queste identità e che è $r_t \sim \text{TMA}(2;1)$, per le previsioni dei livelli, ricordando che in questo caso I_N ed ε_N sono noti e che è $p = E(I_{t-1})$, si ottiene

$$r_N(j) = \begin{cases} I_N(c_1 - \theta\varepsilon_N) + (1 - I_N)(c_2 - \delta\varepsilon_N) & \text{per } j = 1 \\ pc_1 + (1-p)c_2 & \text{per } j > 1 \end{cases}$$

e quindi

$$r_N[h] = \begin{cases} I_N(c_1 - \theta\varepsilon_N) + (1 - I_N)(c_2 - \delta\varepsilon_N) & \text{per } h = 1 \\ (h-1)(pc_1 + (1-p)c_2) + r_N(1) & \text{per } h > 1 \end{cases}$$

Il calcolo di $\sigma_N^2[h]$ è più complesso. Se poniamo $e_N(j) = r_{N+j} - r_N(j)$, si ha

$$e_N(1) = \varepsilon_{N+1}$$

$$e_N(j) = \varepsilon_{N+j} + (I_{N+j-1} - p)(c_1 - c_2) - (I_{N+j-1}\theta + (1 - I_{N+j-1})\delta)\varepsilon_{N+j-1}, \quad j > 1$$

che posto $e_N[h] = e_N(1) + e_N(2) + \dots + e_N(h)$ implica

$$e_N[1] = e_N(1) = \varepsilon_{N+1}$$

$$e_N[h] = (h-1)(\bar{p}_N(h) - p)(c_1 - c_2) + \varepsilon_{N+h} + \sum_{j=1}^{h-1} ((1-\delta) + I_{N+j}(\delta - \theta))\varepsilon_{N+j}$$

per $h > 1$, dove è

$$\bar{p}_N(h) = \frac{1}{h-1} \sum_{j=1}^{h-1} I_{N+j} \xrightarrow{P} p$$

Utilizzando queste espressioni si ha

$$\begin{aligned} \sigma_N^2[h] &= \text{Var}(r_{N+[h]} | \mathfrak{F}_N) = E \left[\left(r_{N+[h]} - r_N[h] \right)^2 | \mathfrak{F}_N \right] = \\ &= \text{Var}(e_N(1) + e_N(2) + \dots + e_N(h)) = \text{Var}(e_N[h]) \end{aligned}$$

Se si tiene conto che per ipotesi ε_t è un white noise e quindi con media zero, varianza costante ed a componenti incorrelate e che

$$I_t^2 = I_t; (1 - I_t)^2 = (1 - I_t); I_t(1 - I_t) = 0, \quad \text{per ogni } t$$

dopo alcuni passaggi algebrici si ha

$$\sigma_N^2[1] = \sigma_N^2(1)$$

mentre per $h > 1$ risulta

$$\begin{aligned} \sigma_N^2[h] &= \sigma_N^2(h) + \left(p(1-\theta)^2 + (1-p)(1-\delta)^2 \right) \sum_{j=1}^{h-1} \sigma_N^2(j) + \\ &+ (c_1 - c_2)^2 \left((h-1)p[1 - (h-1)p] + \sum_{i \neq j}^{h-1} \tau_{|i-j|} \right) \end{aligned} \quad (20)$$

dove $\sigma_N^2(j)$ è definito nel capitolo precedente e si è posto

$$\tau_{|i-j|} = E(I_{N+i} I_{N+j}), \quad i \neq j = 1, 2, \dots, h-1$$

Nelle pratiche applicazioni è di solito $c_1 - c_2 \approx 0$ mentre $\bar{p}_N(h-1) \xrightarrow{P} p$ e l'ultimo addendo della (20) può essere trascurato. La sostituzione dei parametri

incogniti con le loro stime permette di ottenere una stima del VaR. Se il modello per i livelli è un TMA(2; q) la procedura di calcolo di $r_N[h]$ e di $\sigma_N^2[h]$ è la stessa anche se diviene, formalmente, molto più complessa man mano che q aumenta.

Esempio 7

Ai rendimenti giornalieri degli attivi Campari, Fiat, Mibtel, Banca Intesa S. Paolo, già analizzati nelle pagine precedenti, adattiamo un modello della classe TMA(2;1)-TGARCH(1,1,1), nell'ipotesi che i residui standardizzati si distribuiscono come una t di Student con g gradi di libertà, e deriviamo il relativo VaR.

Tabella 11 – *Stime dei parametri del modello TMA(2;1)-TGARCH(1,1,1) per i rendimenti giornalieri di Campari, Fiat, Mibtel, Banca Intesa S. Paolo*

	Campari	Fiat	B. S. Paolo	Mibtel
c_1	-0.00033	0.000634	0.002556	-0.000495
c_2	0.00119	0.000082	-0.001386	-0.000527
θ	0.108	-0.080	-0.105	0.033
δ	0.156	0.016	-0.082	-0.031
w	.00000388	.0000121	.00000376	0.0000026
α	0.018	0.046	0.004	0.000
β	0.946	0.896	0.932	0.907
γ	0.049	0.083	0.123	0.186
ρ	0.451	0.4729	0.4363	.4475
ε_N	0.020420	0.136556	0.146571	0.023459
σ_N^2	0.000659	0.001990	0.005030	0.001039
g	5.37	6.92	5.34	9.13
I_N	1	0	1	1

Le stime dei parametri sono riportate nella tabella 11. Come si può osservare per tutte le serie è praticamente $c_1 \approx c_2 \approx 0$ e per calcolare il VaR si può usare la (20) senza il suo secondo addendo. Questi risultati sono stati utilizzati per ottenere le prime h=10 previsioni, riportate nella tabella 12, sia dei livelli che della volatilità dei rendimenti dei 4 attivi analizzati.

Tabella 12 – Previsione dei livelli e della volatilità

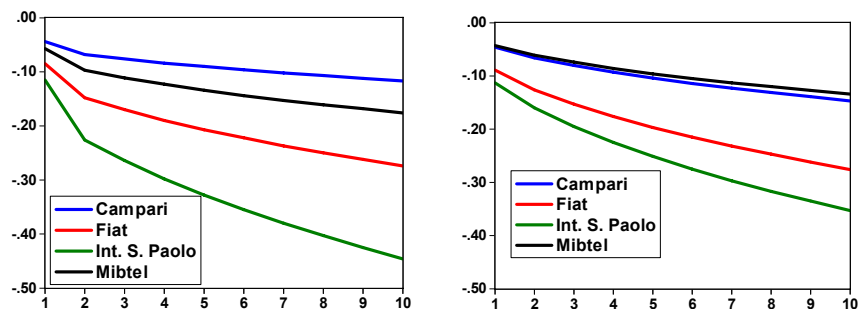
h	Campari		Fiat		Banca Intesa S. Paolo		Mibtel	
	$r_N[j]$	$\sigma_N^2[j]$	$r_N[j]$	$\sigma_N^2[j]$	$r_N[j]$	$\sigma_N^2[j]$	$r_N[j]$	$\sigma_N^2[j]$
1	-0.0025	0.00066	-0.00207	0.00265	0.01802	0.00742	-0.0013	0.00103
2	-0.0020	0.00162	-0.00172	0.00822	0.01835	0.02495	-0.0018	0.00304
3	-0.0015	0.00210	-0.00138	0.01093	0.01868	0.03355	-0.0023	0.00398
4	-0.0010	0.00257	-0.00104	0.01360	0.01902	0.04206	-0.0028	0.00488
5	-0.0005	0.00304	-0.00069	0.01623	0.01935	0.05048	-0.0033	0.00575
6	0.0000	0.00350	-0.00035	0.01883	0.01968	0.05881	-0.0038	0.00660
7	0.0005	0.00396	-0.00001	0.02139	0.02002	0.06706	-0.0043	0.00741
8	0.0010	0.00442	0.00034	0.02392	0.02035	0.07522	-0.0049	0.00820
9	0.0015	0.00487	0.00068	0.02641	0.02068	0.08331	-0.0054	0.00896
10	0.0020	0.00532	0.00102	0.02887	0.02102	0.09131	-0.0059	0.00970

Tabella 13 – VaR con $\alpha = 0.05$ e residui t di Student per i quattro attivi

h	Campari	Fiat	B. Intesa S. Paolo	Mibtel
1	-0.044	-0.085	-0.115	-0.057
2	-0.068	-0.148	-0.226	-0.097
3	-0.076	-0.170	-0.264	-0.111
4	-0.084	-0.190	-0.298	-0.123
5	-0.090	-0.207	-0.328	-0.134
6	-0.096	-0.222	-0.355	-0.144
7	-0.102	-0.237	-0.380	-0.153
8	-0.107	-0.250	-0.403	-0.161
9	-0.112	-0.262	-0.425	-0.168
10	-0.117	-0.274	-0.446	-0.176

Le previsioni hanno permesso di calcolare il relativo VaR per i quattro attivi che è stato riportato nella tabella 13. I risultati ottenuti utilizzando il modello TMA(2;1)-TGARCH(1,1,1) confermano come il titolo nettamente più rischioso dei quattro sia quello di Banca Intesa S. Paolo. Come già evidenziato, questo risultato era atteso visto che le previsioni sono state fatte in piena crisi finanziaria che ha coinvolto pesantemente le banche. Per avere un confronto visivo dell'andamento dei VaR delle quattro serie analizzate questi sono riportati nella figura 7 (parte sinistra).

Figura 8 – VaR per Campari, Fiat, Intesa S. Paolo e Mibtel con TMA-TGARCH (figura sinistra) e con Riskmetrics (figura destra)



E' interessante confrontare i risultati ottenuti nella tabella 13 con gli omologhi della tabella 2 derivata utilizzando Riskmetrics e riportati nella figura 8 (parte destra). Globalmente i risultati della stima del VaR con i due metodi sono molto simili anche se il VaR di Riskmetrics risulta, generalmente, meno elevato, per le quattro serie analizzate ed in particolare per Banca Intesa S. Paolo, rispetto a quello più preciso ottenuto con il TMA-TGARCH.

Per ciascuno dei quattro attivi sopra considerati possiamo calcolare il VaR locale riferito a ciascun regime: $VaR_{N_1}(h; \alpha)$ e $VaR_{N_2}(h; \alpha)$. In questo caso, ricordando che il modello utilizzato è un TMA(2;1)-TGARCH(1,1,1), per i regimi dei livelli si ha:

$$r_{N_1}[h] = hc_1 - \theta \varepsilon_N; \quad r_{N_2}[h] = hc_2 - \delta \varepsilon_N, \quad h=1, 2, \dots$$

mentre per la volatilità risulta:

$$\begin{cases} \sigma_{N_1}^2(1) = w + (\alpha + \gamma)\varepsilon_N^2 + \beta\sigma_N^2 \\ \sigma_{N_1}^2(j) = w + (\alpha + \beta + \gamma)\sigma_{N_1}^2(j-1) \end{cases}; \quad \begin{cases} \sigma_{N_2}^2(1) = w + \alpha\varepsilon_N^2 + \beta\sigma_N^2 \\ \sigma_{N_2}^2(j) = w + (\alpha + \beta)\sigma_{N_2}^2(j-1) \end{cases}$$

che permettono di calcolare

$$\begin{cases} \sigma_{N_1}^2[1] = \sigma_{N_1}^2(1) \\ \sigma_{N_1}^2[h] = (1 - \theta)^2 \sum_{j=1}^{h-1} \sigma_{N_1}^2(j) + \sigma_{N_1}^2(h), \quad h > 1 \end{cases}$$

$$\begin{cases} \sigma_{N2}^2[1] = \sigma_{N2}^2(1) \\ \sigma_{N2}^2[h] = (1 - \delta)^2 \sum_{j=1}^{h-1} \sigma_{N2}^2(j) + \sigma_{N2}^2(h), \quad h > 1 \end{cases}$$

In questo modo il VaR per ciascun regime è dato da

$$\begin{cases} \text{VaR}_{N1}(h; \alpha) = r_{N1}[h] + a_{1(\alpha)} \sigma_{N1}[h] \\ \text{VaR}_{N2}(h; \alpha) = r_{N2}[h] + a_{2(\alpha)} \sigma_{N2}[h] \end{cases}$$

Applichiamo le formule sopra riportate per calcolare il VaR locale riferito a ciascuno dei due regimi degli attivi Campari, Fiat, B. Intesa S. Paolo, Mibtel. Nella tabella 14 riportiamo le previsioni dei livelli e della volatilità in ciascun regime di Campari e Fiat, mentre nella tabella 15 riportiamo quelle di Banca Intesa S. Paolo e Mibtel. Nella tabella 16 riportiamo i VaR dei due regimi con $\alpha = 0.05$ per i quattro attivi.

Tabella 14 – Previsioni dei livelli e della volatilità nei due regimi dei rendimenti di Campari e Fiat

h	Campari				Fiat			
	$r_{N1}[h]$	$\sigma_{N1}^2[h]$	$r_{N2}[h]$	$\sigma_{N2}^2[h]$	$r_{N1}[h]$	$\sigma_{N1}^2[h]$	$r_{N2}[h]$	$\sigma_{N2}^2[h]$
1	-0.0025	0.0010	-0.0021	0.0006	0.0116	0.0194	-0.0021	0.0027
2	-0.0029	0.0019	-0.0009	0.0011	0.0122	0.0425	-0.0020	0.0051
3	-0.0032	0.0027	0.0003	0.0015	0.0129	0.0662	-0.0019	0.0074
4	-0.0035	0.0036	0.0015	0.0019	0.0135	0.0906	-0.0018	0.0096
5	-0.0039	0.0044	0.0027	0.0023	0.0141	0.1156	-0.0017	0.0116
6	-0.0042	0.0053	0.0039	0.0027	0.0148	0.1412	-0.0017	0.0136
7	-0.0045	0.0062	0.0051	0.0030	0.0154	0.1676	-0.0016	0.0154
8	-0.0049	0.0072	0.0063	0.0034	0.0160	0.1946	-0.0015	0.0172
9	-0.0052	0.0081	0.0075	0.0037	0.0167	0.2224	-0.0014	0.0189
10	-0.0055	0.0090	0.0087	0.0041	0.0173	0.2508	-0.0013	0.0205

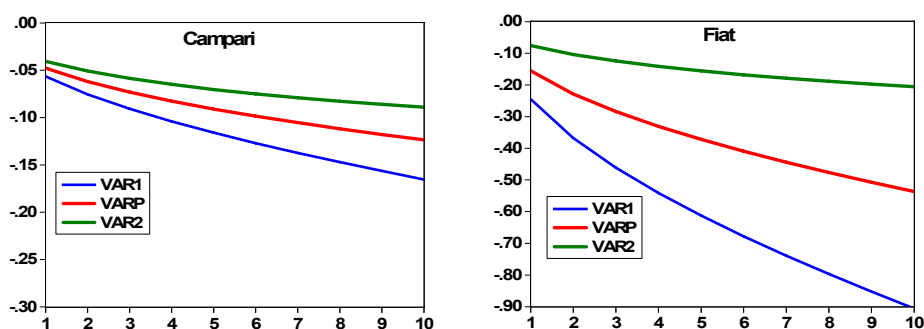
Tabella 15 – Previsioni dei livelli e della volatilità nei due regimi dei rendimenti di Banca Intesa S. Paolo e Mibtel

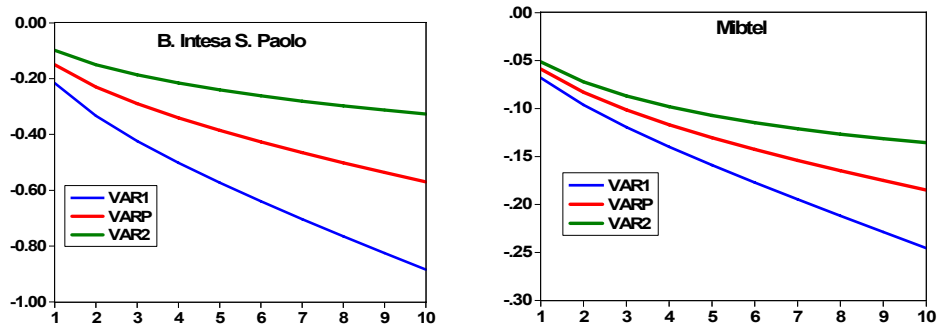
h	Banca Intesa S. Paolo				Mibtel			
	$r_{N1}[h]$	$\sigma_{N1}^2[h]$	$r_{N2}[h]$	$\sigma_{N2}^2[h]$	$r_{N1}[h]$	$\sigma_{N1}^2[h]$	$r_{N2}[h]$	$\sigma_{N2}^2[h]$
1	0.0180	0.0248	0.0107	0.0048	-0.00127	0.0014	0.00012	0.0009
2	0.0206	0.0565	0.0093	0.0101	-0.00177	0.0029	-0.00033	0.0018
3	0.0231	0.0902	0.0079	0.0150	-0.00226	0.0044	-0.00086	0.0026
4	0.0257	0.1257	0.0066	0.0196	-0.00276	0.0060	-0.00138	0.0033
5	0.0282	0.1634	0.0052	0.0240	-0.00325	0.0078	-0.00191	0.0039
6	0.0308	0.2033	0.0038	0.0281	-0.00375	0.0096	-0.00244	0.0044
7	0.0334	0.2456	0.0024	0.0319	-0.00424	0.0116	-0.00296	0.0049
8	0.0359	0.2903	0.0010	0.0354	-0.00474	0.0138	-0.00349	0.0053
9	0.0385	0.3376	-0.0004	0.0388	-0.00523	0.0160	-0.00402	0.0057
10	0.0410	0.3877	-0.0018	0.0419	-0.00573	0.0185	-0.00454	0.0060

Tabella 16 – VaR nei due regimi per gli attivi Campari, Fiat, B. Intesa S. Paolo, Mibtel con $\alpha = 0.05$

h	Campari		Fiat		B. INT. SP		Mibtel	
	VaR ₁	VaR ₂	VaR ₁	VaR ₂	VaR ₁	VaR ₂	VaR ₁	VaR ₂
1	-0.056	-0.041	-0.245	-0.076	-0.216	-0.099	-0.068	-0.051
2	-0.075	-0.051	-0.368	-0.104	-0.333	-0.150	-0.096	-0.072
3	-0.091	-0.059	-0.461	-0.124	-0.423	-0.186	-0.119	-0.087
4	-0.104	-0.065	-0.541	-0.141	-0.502	-0.216	-0.140	-0.098
5	-0.116	-0.070	-0.612	-0.156	-0.573	-0.240	-0.159	-0.107
6	-0.127	-0.075	-0.678	-0.168	-0.640	-0.262	-0.177	-0.115
7	-0.137	-0.079	-0.739	-0.179	-0.703	-0.281	-0.195	-0.121
8	-0.147	-0.083	-0.797	-0.189	-0.765	-0.297	-0.212	-0.127
9	-0.156	-0.086	-0.852	-0.197	-0.825	-0.313	-0.229	-0.131
10	-0.165	-0.089	-0.906	-0.205	-0.885	-0.326	-0.246	-0.136

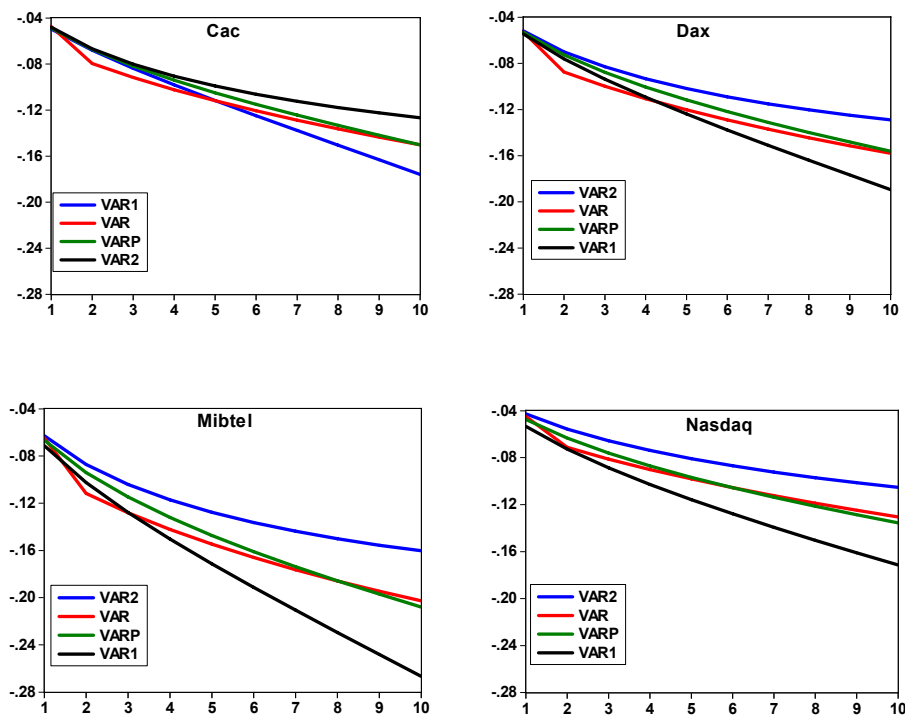
Figura 9 – VaR di Campari, Fiat, B. Int. S. Paolo, Mibtel e dei loro regimi

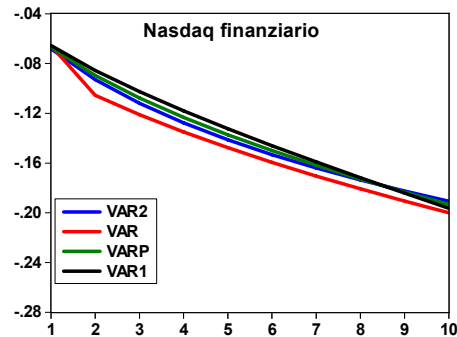




Nella figura 9, in ciascun grafico con $\alpha = 0.05$, sono confrontati, per ciascuno dei quattro attivi, i VaR locali con quello globale (VaRp) ottenuto come media ponderata di quelli dei singoli regimi. Questi grafici forniscono ulteriori informazioni sulla rischiosità dell'attivo in analisi ed in particolare sulla struttura interna del rischio distinguendo fra mercato *orso* ($r_{t-1} < 0$), che in questo caso è sistematicamente quello più rischioso, e mercato *toro* ($r_{t-1} \geq 0$).

Figura 10 – VaR e suoi regimi per Cac, Dax, Mibtel, Nasdaq, Nasdaq Finanziario





Nella figura 10 riportiamo i VaR locali riferiti ai regimi (VaR1 e VaR2), il VaR globale ottenuto come media ponderata di quello dei regimi (VaRp) e quello approssimato (VaR), calcolato come nella tabella 14 quando nella (20) si trascura il secondo addendo, per le serie finanziarie Cac, Dax, Mibtel, Nasdaq e Nasdaq-finanziario già analizzate nel capitolo II.

Qui di seguito riportiamo il listato per la stima con EViews 6 del TMA(2;1)-TIGARCH(1,1,1) ed i relativi VaR.

Stima TMA_TARCH e VaR

```

smpl 1 N
genr r=0
genr rf=0
genr r1=0
genr r2=0
genr re=0
genr re1=0
genr re2=0
genr at=0
genr a1=0
genr a2=0
genr v=0
genr v1=0
genr v2=0
genr ve=0
genr it=0
genr i1=0
genr i2=0
delete r r1 r2 rf re re1 re2 at a1 a2 v v1 v2 ve it i1 i2
genr r=d(log(x))

```

```

    smpl 2 N
for li=3 to N
    smpl li li
    if r(li-1)<0 then genr it=1
        genr r1=r
        else genr it=0
        genr r2=r
    endif
next
    smpl 3 N
    scalar p=@mean(it)
    equation eq0.ls r=(c(1)+c(2)*r(-1)+c(3)*r(-2))*it+(c(4)+c(5)*r(-1)+c(6)*r(-
2))*(1-it)
    genr re=resid
for li=1 to 20
    equation eqm.ls r=(c(1)-c(2)*re(-1))*it+(c(3)-c(4)*re(-1))*(1-it)
    smpl 1 20-li+1
    smpl 5 N
    genr re=resid
next
    scalar c1=c(1)
    scalar teta=c(2)
    scalar c2=c(3)
    scalar delt=c(4)
    smpl 4 N
    equation eqv.arch(1,1, thrsh=1, tdist) re
    smpl N+1 N+M
    forecast fr ve v
    smpl 1 N
    fit rf vef vf
    genr at=re/v^.5
    genr re=resid
    smpl 2 N
    scalar w=c(1)
    scalar a=c(2)
    scalar g=c(3)
    scalar b=c(4)
for li=3 to N
    smpl li li
    if r(li-1)<0 then genr v1=v
        genr a1=at

```

```

        genr i1=1
        genr re1=re
        else genr v2=v
        genr a2=at
        genr i2=1
        genr re2=re
    endif
next
    smpl 1 N
    scalar zq=@quantile(at,.05)
    scalar zq1=@quantile(a1,.05)
    scalar zq2=@quantile(a2,.05)
    scalar vn=@elem(v,"N")
    scalar en=@elem(re,"N")
    scalar in=@elem(it,"N")
    smpl 1 1
    genr h=1
    genr vnt=w+(a+g*in)*en^2+b*vn
    genr vn1=w+(a+b)*en^2+b*vn
    genr vn2=w+a*en^2+b*vn
    genr svn1=vn1
    genr vs1=svn1
    genr svn2=vn2
    genr vs2=svn2
    genr vs=vnt
    genr svn=vnt
    genr rs=in*(c1-teta*en)+(1-in)*(c2-delt*en)
    smpl 2 M
    genr vnt=w+(a+g*p+b)*vnt(-1)
    genr vn1=w+(a+b+g)*vn1(-1)
    genr vn2=w+(a+b)*vn2(-1)
    genr svn1=vn1+svn1(-1)
    genr svn2=vn2+svn2(-1)
    genr svn=vnt+svn(-1)
    genr h=1+h(-1)
    genr vs=vnt+(p*(1-teta)^2+(1-p)*(1-delt)^2)*svn
    genr vs1=vn1+(1-teta)^2*svn1(-1)
    genr vs2=vn2+(1-delt)^2*svn2(-1)
    genr rs=(h-1)*(p*c1+(1-p)*c2)+@elem(rs,"1")
    smpl 1 M
    genr rs1=h*c1-teta*en

```

```

genr rs2=h*c2-delt*en
genr var1=h*c1-teta*en+zq1*vs1^.5
genr var2=h*c2-delt*en+zq2*vs2^.5
genr varp=p*var1+(1-p)*var2
genr var=rs+zq*vs^.5
line(ab=boxplot) var1 varp var var2
delete svn svn1 svn2 fr vnt re re1 re2 ve vef rf vf eq0 h
smpl 1 N+M

```

Per poter eseguire il programma ed ottenere le stime è necessario che sia attivo il workfile in cui è riportata la serie dei prezzi, dell'attivo o del portafoglio, qui indicata con x , in analisi. Nel listato bisogna sostituire ad N il numero delle osservazioni di x ed a M il numero delle previsioni che si vogliono effettuare. Se si è interessati alla stima dei livelli con il più generale TMA(2; q) il listato va modificato di conseguenza.

L'output di interesse del programma è:

- eqm, che riporta le stime dei minimi quadrati dei parametri del modello TMA(2;1) nei due regimi;
- eqv, che riporta le stime di massima verosimiglianza del modello TGARCH(1,1,1) con residui t di Student. Il listato è predisposto per la stima del TGARCH(1,1,1), se si è interessati al TIGARCH(1,1,1) basta **sostituire la riga**
equation eqv.arch(1,1, thrsh=1, tdist) re
con la riga
equation eqv.arch(1,1, integrated, thrsh=1, tdist) re
- p, che permette di ottenere una stima di $p=E(I_{t-1})$;
- le serie $r, r1, r2, re, re1, re2, v, v1, v2, at, a1, a2, it, i1, i2, var1, var2, varp, var$ che permettono di ottenere rispettivamente: i rendimenti complessivi (r) e quelli del primo ($r1$) e del secondo ($r2$) regime; i residui dei livelli (re) e quelli del primo ($re1$) e del secondo ($re2$) regime; la volatilità (v) e quella del primo ($v1$) e del secondo ($v2$) regime; i residui standardizzati (at) e quelli del primo ($a1$) e del secondo ($a2$) regime; la serie dicotomica (it) e quella del primo ($i1$) e del secondo regime ($i2$); il VaR del primo regime ($var1$), del secondo regime ($var2$), quello globale ottenuto come media ponderata dei VaR dei regimi ($varp$) ed il VaR globale approssimato con la (20) senza il secondo addendo;
- il grafico dei quattro VaR sopra menzionati.

□

6.8 La stima del VaR di medio-lungo periodo ($h > 1$)

Nel paragrafo 6.3 si è mostrato come, nel caso del modello Riskmetrics, sia possibile definire una approssimazione analitica per il calcolo del valore a rischio per periodi di detenzione di lunghezza $h > 1$. Nel paragrafo 6.4 si è inoltre illustrato come sia possibile raggiungere un analogo risultato per i modelli ARMA-GARCH e TMA-GARCH. Nel caso di modelli più complessi, tuttavia, stime di medio-lungo periodo del VaR non sono generalmente ottenibili attraverso strumenti analitici ma si rende invece necessario ricorrere a procedure numeriche basate su simulazioni Monte Carlo. Sebbene non esistano pacchetti software commerciali che includano delle funzioni predefinite per l'esecuzione di tali procedure, esse possono essere facilmente implementate in qualsiasi linguaggio statistico (Eviews, Matlab, R, S-plus, Gauss, Stata, etc...) richiedendo come pre-requisito semplicemente la disponibilità di un generatore di numeri casuali. Nel seguito verrà descritto un algoritmo numerico per la stima del VaR per orizzonti temporali di lunghezza $h > 1$. Sebbene, nel corso della trattazione, si farà riferimento a modelli della classe ARMA-GARCH con errori distribuiti come una t di Student, ci preme precisare come tale scelta sia motivata esclusivamente da considerazioni di natura didattica. La procedura di seguito illustrata può infatti essere generalizzata in maniera pressoché immediata a modelli più complessi e generali di quelli che vengono di seguito considerati. E' infatti proprio in questi casi che, vista la difficoltà nel derivare formule analitiche, si avverte in maniera più pressante la necessità di una procedura di stima di tipo numerico.

Sia (r_1, r_2, \dots, r_N) una serie storica di rendimenti relativi ad un dato attivo finanziario. Si assuma poi che i rendimenti r_t , $t=1, \dots, N+h$, siano stati generati da un processo ARMA(p, q)-GARCH(s, v)

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_r B^r) r_t = \phi_0 + (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) \varepsilon_t$$

$$\varepsilon_t = \sigma_t a_t$$

$$\sigma_t^2 = w + \alpha_1 \varepsilon_{t-1}^2 + \dots + \alpha_s \varepsilon_{t-s}^2 + \beta_1 \sigma_{t-1}^2 + \dots + \beta_v \sigma_{t-v}^2 \quad (21)$$

dove $a_t \sim ((g-2)/g)^{1/2} t_{(g)}$

Ricordiamo che il VaR h passi in avanti ($h > 0$) al livello α , $\text{VaR}_N(h; \alpha)$, può

essere definito come il quantile di ordine α della distribuzione dei rendimenti sul periodo che va da $N+1$ ad $N+h$:

$$r_{N+[h]} = \sum_{i=1}^h r_{N+i}$$

condizionata all'informazione disponibile al tempo N , \mathfrak{F}_N .

In termini economici, ci si riferisce ad una situazione in cui un dato investitore, che al tempo N ha acquistato o ha in programma di acquistare un titolo, vuole valutare il rischio a cui sarà esposto detenendo il titolo per il periodo $[N, N+h]$. Sulla base di tale definizione, per il calcolo del $\text{VaR}_N(h; \alpha)$ è possibile utilizzare il seguente algoritmo:

1) utilizzando un generatore di numeri casuali, si generano indipendentemente m vettori h -dimensionali di realizzazioni della variabile casuale errore a_t , $t=N+1, \dots, N+h$. Il risultato di questa operazione sarà una matrice $A_{N,h}$, di dimensioni (h, m) di numeri pseudo-casuali generati da una v.c. t di Student standardizzata con g gradi di libertà:

$$A_{N,h} = \begin{bmatrix} a_{N+1}^{(1)} & \cdots & a_{N+1}^{(i)} & \cdots & a_{N+1}^{(m)} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{N+j}^{(1)} & \cdots & a_{N+j}^{(i)} & \cdots & a_{N+j}^{(m)} \\ \vdots & & \vdots & \ddots & \vdots \\ a_{N+h}^{(1)} & \cdots & a_{N+h}^{(i)} & \cdots & a_{N+h}^{(m)} \end{bmatrix}$$

L' i -esima colonna di questa matrice,

$$\mathbf{a}_{N,h}^{(i)} = (a_{N+1}^{(i)}, \dots, a_{N+j}^{(i)}, \dots, a_{N+h}^{(i)}), \quad i=1, \dots, m$$

può essere interpretata come una delle possibili realizzazioni a cui il processo degli errori a_t può dar luogo negli h giorni successivi alla fine della serie osservata ovvero nell'intervallo $[N, N+h]$, dove l'elemento di posto $\{i, j\}$ della

matrice $A_{N,h}$, $a_{N+j}^{(i)}$, $j=1, \dots, h$, $i=1, \dots, m$, è dato dall' i -esimo valore simulato per il periodo $N+j$;

2) ciascuna delle colonne della matrice $A_{N,h}$ corrisponderà ad una differente realizzazione futura del processo dei rendimenti r_t sull'intervallo $[N, N+h)$. Conseguentemente, per il tempo $N+1$, sarà possibile generare m rendimenti e volatilità simulate attraverso le equazioni

$$r_{N+1}^{(i)} = \phi_0 + \phi_1 r_N + \phi_2 r_{N-1} + \dots + \phi_p r_{N-p+1} + \\ - \theta_1 \varepsilon_N - \theta_2 \varepsilon_{N-1} \dots - \theta_q \varepsilon_{N-q+1} + \varepsilon_{N+1}^{(i)}$$

$$\varepsilon_{N+1}^{(i)} = \sigma_{N+1}^{(i)} a_{N+1}^{(i)}$$

$$\sigma_{N+1}^{(i)2} = w + \alpha_1 \varepsilon_N^2 + \dots + \alpha_s \varepsilon_{N-s+1}^2 + \beta_1 \sigma_N^2 + \dots + \beta_v \sigma_{N-v+1}^2$$

3) il passo 2 viene ripetuto per $t=N+2, \dots, N+h$. Per il tempo $N+2$, le equazioni utilizzate per generare i valori simulati saranno date da

$$r_{N+2}^{(i)} = \phi_0 + \phi_1 r_{N+1}^{(i)} + \phi_2 r_N + \dots + \phi_p r_{N-p+2} + \\ - \theta_1 \varepsilon_{N+1}^{(i)} - \theta_2 \varepsilon_N \dots - \theta_q \varepsilon_{N-q+2} + \varepsilon_{N+2}^{(i)}$$

$$\varepsilon_{N+2}^{(i)} = \sigma_{N+2}^{(i)} a_{N+2}^{(i)}$$

$$\sigma_{N+2}^{(i)2} = w + \alpha_1 \varepsilon_{N+1}^{(i)2} + \alpha_2 \varepsilon_N^2 \dots + \alpha_s \varepsilon_{N-s+2}^2 + \beta_1 \sigma_{N+1}^{(i)2} + \beta_2 \sigma_N^2 \dots + \beta_v \sigma_{N-v+2}^2$$

per $i=1, \dots, m$. Le equazioni per gli istanti successivi si derivano in maniera analoga sostituendo i valori osservati con quelli simulati al crescere di t . Alla fine della procedura, si saranno generate m ipotetiche traiettorie future (scenari), di lunghezza h , del processo dei rendimenti. Tali traiettorie possono essere raggruppate in una matrice di dimensioni (h, m) che sarà data da

$$B_{N,h} = \begin{bmatrix} r_{N+1}^{(1)} & \cdots & r_{N+1}^{(i)} & \cdots & r_{N+1}^{(m)} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ r_{N+j}^{(1)} & \cdots & r_{N+j}^{(i)} & \cdots & r_{N+j}^{(m)} \\ \vdots & & \vdots & \ddots & \vdots \\ r_{N+h}^{(1)} & \cdots & r_{N+h}^{(i)} & \cdots & r_{N+h}^{(m)} \end{bmatrix}$$

4) sommando gli elementi in ciascuna colonna di $B_{N,h}$ si ottiene un vettore (1, m) di ipotetici rendimenti cumulati sul periodo $[N, N+h)$

$$c_{N,h} = \begin{bmatrix} r_{N+[h]}^{(1)} & \cdots & r_{N+[h]}^{(i)} & \cdots & r_{N+[h]}^{(m)} \end{bmatrix}$$

dove

$$r_{N+[h]}^{(i)} = \sum_{j=1}^h r_{N+j}^{(i)}; \quad i=1, \dots, m$$

5) il VaR h passi in avanti al livello α , $VaR_N(h; \alpha)$, può infine essere approssimato calcolando il quantile empirico del vettore dei rendimenti cumulati $c_{N,h}$.

Nelle applicazioni pratiche della procedura illustrata i valori dei parametri del modello saranno generalmente incogniti ed andranno quindi sostituiti con le rispettive stime di massima verosimiglianza. Va infine osservato come la determinazione del numero di simulazioni m è di importanza cruciale e può influenzare in maniera significativa i risultati ottenuti. Ovviamente, l'accuratezza dell'approssimazione tende a migliorare al crescere di m . Asintoticamente, al limite per $m \rightarrow \infty$, il valore approssimato tende a coincidere con quello esatto. D'altra parte, bisogna tener conto del fatto che sussistono vincoli di natura computazionale che limitano il numero di simulazioni che è possibile effettuare. Tali vincoli divengono particolarmente pressanti quando si è costretti ad analizzare un numero molto elevato di serie storiche in tempi relativamente ristretti, come di solito avviene nelle società finanziarie. I tempi di elaborazione sono poi fortemente condizionati dalla complessità del modello con cui ci si trova ad operare.

In sintesi, la scelta del valore di m da utilizzarsi verrà quindi

inevitabilmente a dipendere dai seguenti fattori:

- risorse hardware disponibili
- numero di serie da analizzare
- complessità del modello
- tempi assegnati per l'elaborazione

E' importante tenere conto, infine, del fatto che, a parità di numero di simulazioni effettuate, l'accuratezza dell'approssimazione tenderà a peggiorare al decrescere di α .

Esempio 8

Calcoliamo ora il VaR per i titoli Campari e Fiat utilizzando i modelli MA(q)-GARCH(s, v) identificati e stimati nel paragrafo 6.4 (esempio 3) e considerando diversi valori di h, in particolare $1 \leq h \leq 22$. Rispetto ai modelli utilizzati nel paragrafo 6.4, inoltre, introduciamo l'ipotesi che gli errori a_t siano distribuiti come una t di Student, con g gradi di libertà, standardizzata:

$$a_t \stackrel{\text{iid}}{\sim} \sqrt{(g-2)/g} t_{(g)}$$

I parametri dei modelli utilizzati per la stima del VaR sono stati stimati con il metodo della massima verosimiglianza. Le stime ed i relativi p-values sono riportati in tabella 17.

Tabella 17 – *Stime dei modelli MA-GARCH per i rendimenti giornalieri sugli attivi Campari e Fiat con residui t di Student*

Titolo		θ_1	w	α	β	g	AIC
Fiat	Stima Prob.		1.01* (.006)	0.100 (.000)	0.887 (.000)	6.224 (.000)	-4.866
Campari	Stima Prob.	-0.144 (.000)	0.352* (.034)	0.046 (.000)	0.944 (.000)	5.042 (.006)	-5.552

(*: $\times 10^{-5}$)

Si procede quindi alla stima del VaR per valori crescenti di h ($1 \leq h \leq 22$) utilizzando un numero di simulazioni Monte Carlo $m = 100.000$. Di seguito, a titolo esemplificativo, viene riportato il programma, scritto in linguaggio Eviews, utilizzato per la stima del VaR nel caso di un modello MA(1)-GARCH(1, 1).

'Espando il range della workfile aggiungendo 22 osservazioni (se richiesto)

```
scalar maxh=22
pagestruct(end=@last+maxh) *
```

'Seleziono un sample corrispondente al periodo per cui sono disponibili dati osservati.

```
smp1 @first @last-maxh
```

'Stimo un modello MA(1)-GARCH(1,1) con errori t. "serie" indica la serie dei rendimenti logaritmici sull'attivo di interesse

```
equation modserie.arch(1,1,tdist) serie ma(1) c
```

'Estraggo le volatilità, i residui e i coefficienti del modello stimato

```
modserie.makegarch(v) volserie
modserie.makeresids(o) res
vector coeff=modserie.@coefs
scalar theta1=coeff(1)
scalar omega=coeff(2)
scalar a1=coeff(3)
scalar b1=coeff(4)
scalar df=coeff(5)
```

'Estendo il sample per la previsione

```
smp1 @all
'Inizializzo le serie per la previsione
gener vol_f=0
gener r=serie
gener obs=@trend+1
scalar N=@max(obs)-maxh
scalar q=1
scalar s=1
scalar v=1
```

'Seleziono l'orizzonte da prevedere (maxh), il numero di simulazioni (nsim)

```
scalar nsim=10^5
```

'Preordino le matrici destinate a contenere i rendimenti (B_t), i residui e le volatilità future simulate

```
matrix(maxh+q,nsim) vret
matrix(maxh+q,nsim) vres
matrix(maxh+s,nsim) vvol
```

'Inizializzo l'algoritmo di simulazione

```
rowvector(nsim) r1=r(N)
rowvector(nsim) res1=res(N)
rowvector(nsim) v1=volserie(N)
```

```
rowplace(vret, r1, 1)
rowplace(vres, res1, 1)
rowplace(vvol, v1, 1)
```

'Ciclo 'for' per la generazione delle traiettorie future dei rendimenti

```
for li=N+1 to N+maxh
for lj=1 to nsim
vvol(li+s-N,lj)=omega+a1*vres(li+q-N-1,lj)^2+b1*vvol(li+s-N-1,lj)
scalar u= @sqr(vvol(li+s-N,lj))*@sqr((df-2)/df)*@rtdist(df)
vres(li+q-N,lj)=u
vret(li+q-N,lj)=theta1*vres(li+q-N-1,lj)+u
next
next
```

'Calcolo del VaR a partire dalla matrice dei rendimenti simulati

```
scalar alpha=0.05
matrix(maxh,nsim) cvret
matrix vret=@subextract(vret, q+1,1,maxh+q,nsim)
vector(maxh,1) varmat
```

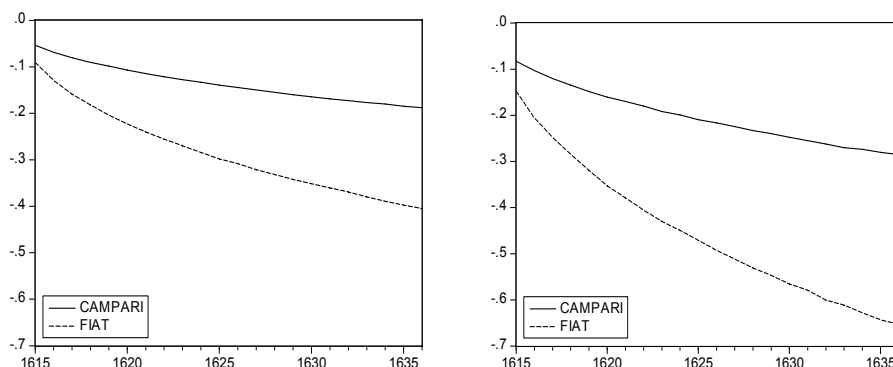
```
for li = 1 to maxh
matrix temp1=@subextract(vret,1,1, li,nsim)
vector temp2=@csum(temp1)
rowvector temp3=@transpose(temp2)
rowplace(cvret,temp3,li)
varmat(li)=@quantile(temp3,alpha)
next
```

'Grafico lineare del VaR stimato rispetto al valore di b=1,2,...,22.

```
smpl @last-maxh+1 @last
mtos(varmat,varmats)
line varmats
```

Le stime del VaR ottenute per i titoli Fiat e Campari, per i livelli di confidenza $\alpha=0.05$ e $\alpha=0.01$, sono state rappresentate graficamente in figura 11.

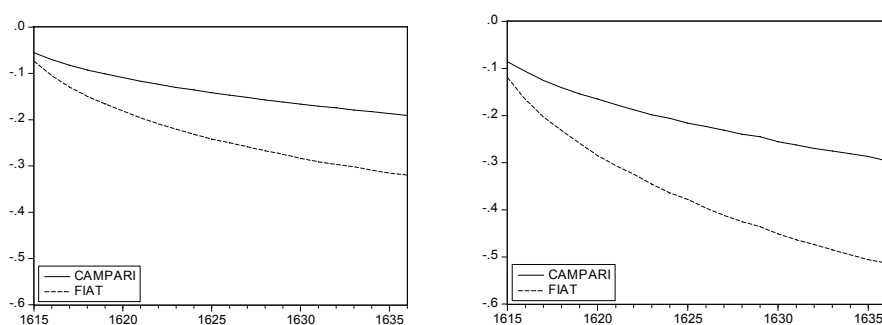
Figura 11 – *VaR stimato per Campari e Fiat con $1 \leq b \leq 22$ per $\alpha=0.05$ (sinistra) e $\alpha=0.01$ (destra) utilizzando modelli MA-GARCH con errori t*



L'andamento dei valori a rischio riportati in figura 11 conferma quanto già riscontrato nei paragrafi 6.3.e 6.4 nelle figure 3 e 4. La detenzione del titolo Fiat comporta un rischio molto più elevato di quello associato al titolo Campari. Per Fiat, inoltre, l'esposizione al rischio tende a crescere molto velocemente al crescere dell'orizzonte temporale di riferimento indicando una notevole incertezza riguardo all'andamento delle quotazioni di medio-lungo periodo del titolo. Al livello $\alpha=0.01$, il rischio associato alla detenzione del titolo Fiat per un mese borsistico è approssimativamente pari a 0.6 ovvero al 60% del capitale investito.

□

Figura 12 – *VaR stimato per Campari e Fiat con $1 \leq b \leq 22$ per $\alpha=0.05$ (sinistra) e $\alpha=0.01$ (destra) utilizzando modelli MA-EGARCH con errori t*



L'approccio, ed il programma, presentati possono essere facilmente generalizzati al caso di modelli, per i livelli e la volatilità, diversi da quelli qui considerati. A tal fine è sufficiente sostituire le equazioni utilizzate per la

simulazione dei livelli e delle volatilità future con quelle relative allo specifico modello di interesse. Se, ad esempio, si volesse assumere, per la volatilità, un modello EGARCH(1,1,1), il ciclo 'for' per la simulazione delle traiettorie future sarebbe sostituito dalle seguenti istruzioni

'Ridefinizione di vvol come matrice dei logaritmi delle volatilità

```
rowvector(nsim) v1=log(volserie(N))
```

```
rowplace(vvol, v1, 1)
```

'Ciclo 'for' per la generazione delle traiettorie future dei rendimenti

```
for li=N+1 to N+maxh
```

```
for lj=1 to nsim
```

```
scalar rtm1=vres(li+q-N-1,lj)
```

```
scalar voltm1=vvol(li+1-N-1,lj)
```

```
vvol(li+1-N,lj)=omega+a1*@abs(rtm1)/@sqr(exp(voltm1))+(b1*voltm1)+c1*
```

```
rtm1/(@sqr(exp(voltm1)))
```

```
scalar u= @sqr(exp(vvol(li+1-N,lj)))*@sqr((df-2)/df)*@rtdist(df)
```

```
vres(li+q-N,lj)=u
```

```
vret(li+q-N,lj)=theta1*vres(li+q-N-1,lj)+u
```

```
next
```

```
next
```

Dopo aver adattato dei modelli MA(1)-EGARCH(1,1,1) e MA(0)-EGARCH(1,1,1) ai rendimenti sui titoli Campari e Fiat, rispettivamente, abbiamo quindi stimato il relativo VaR applicando la procedura di simulazione sopra illustrata. I risultati sono stati riportati in figura 12.

6.9 L'Expected Shortfall

(A) MISURE DI RISCHIO COERENTI ED EXPECTED SHORTFALL

Si assuma di trovarsi in una situazione in cui un generico investitore ha la possibilità di ripartire le sue scelte di investimento all'interno di un insieme di k attivi finanziati A_i , $i=1, \dots, k$. Si indichi inoltre con r_i il valore futuro netto, ad un dato orizzonte, dell'investimento effettuato nell'attivo A_i . È immediato notare che r_i è una variabile aleatoria dal momento che il suo valore è determinato dall'andamento futuro del mercato, naturalmente soggetto ad incertezza. Si consideri ora una funzione $\rho(\cdot)$ avente come dominio l'insieme dei possibili valori assunti dalle r_i e come codominio la retta reale \mathcal{R} . La

funzione $\rho(\cdot)$ viene definita una *misura di rischio monetario* se, per ogni $1 \leq i, j \leq m$, soddisfa le seguenti condizioni:

1. *monotonicità*: se $|r_i| \leq |r_j|$, allora $\rho(r_i) \leq \rho(r_j)$

ovvero, maggiore è il profitto, maggiore è il rischio;

2. *invarianza rispetto a traslazione*: se $a \in \mathfrak{R}$, allora $\rho(a + r_i) = \rho(r_i) - a \Rightarrow \rho(r_i) = \rho(a + r_i) + a$

ovvero, includendo un attivo a reddito fisso ($a > 0$) nel proprio portafoglio, si hanno degli effetti benefici sul rischio. Il rischio associato all'investimento di partenza $\rho(r_i)$, infatti, è maggiore del rischio associato al portafogli integrato con un investimento a reddito fisso, $\rho(r_i + a)$, e la differenza fra le due misure di rischio è la costante a . L'esposizione di questa proprietà è facilmente generalizzabile al caso in cui l'investimento di base non è composto da un solo attivo ma da un portafoglio di attivi.

Una misura di rischio monetario viene invece detta *coerente* (Artzner et al., 1999) se, in aggiunta alle precedenti, soddisfa le seguenti proprietà:

3. *sub-additività*: $\rho(r_i + r_j) \leq \rho(r_i) + \rho(r_j)$;

in altri termini la *diversificazione* degli investimenti ha effetti benefici sull'esposizione al rischio;

4. *omogeneità*: se $a > 0$, allora $\rho(ar_i) = a \rho(r_i)$;

ovvero, ad esempio, raddoppiando il profitto raddoppierà anche il rischio.

Si noti che negli assiomi 1-2 il valore della misura di rischio $\rho(\cdot)$ cresce al crescere dell'effettiva esposizione al rischio dell'investimento in esame. Finora, in questo testo si è adottata come misura di rischio il "rendimento minimo probabile", in altre parole il VaR, il cui valore è inversamente legato all'effettiva esposizione al rischio. Questa apparente incoerenza può tuttavia essere risolta in maniera molto semplice adottando come misura di rischio il VaR moltiplicato per -1 piuttosto che preso con il suo segno. In tal modo si ottiene una misura di rischio coerente con gli assiomi 1-2 e, sul piano informativo, equivalente al VaR.

E' possibile dimostrare che il VaR non è, in generale, una misura di rischio *coerente* in quanto non rispetta il requisito (3) della sub-additività. Una conseguenza peggiore di questo fatto è che, in una società finanziaria, l'utilizzo del VaR quale misura di rischio potrebbe scoraggiare il ricorso alla diversificazione. Si può tuttavia dimostrare che il VaR costituisce una misura di rischio coerente se si formula l'ipotesi aggiuntiva che la distribuzione dei rendimenti sia di tipo *ellittico*. Un importante esempio di variabile casuale ellittica è rappresentato dalla variabile casuale normale. Inoltre, va osservato che il valore del VaR dipende soltanto dal valore di un particolare quantile della distribuzione dei rendimenti e non viene minimamente influenzato da quello che avviene nella coda sinistra della distribuzione sull'intervallo $(-\infty, \text{VaR}]$. In teoria quindi potremmo avere due investimenti alternativi, descritti dalle distribuzioni dei rendimenti associate, caratterizzati dallo stesso VaR ma con comportamenti di coda e quindi rischiosità del tutto differenti. In termini economici, questo viene a significare che potrebbero esistere situazioni caratterizzate da rischiosità marcatamente differenti, ma non discriminabili sulla base del semplice VaR.

Considerazioni di questo tipo rappresentano il principale motore che ha dato impulso alla ricerca finalizzata all'elaborazione di misure di rischio alternative al VaR che siano *coerenti* ed in grado di discriminare situazioni in cui l'informazione desunta dal VaR può essere non esaustiva ai fini della valutazione di scenari di investimento alternativi. Fra le diverse proposte emerse in letteratura, quella che senz'altro assume maggiore rilevanza, sia sul piano pratico che quello teorico, è l'*Expected Shortfall* o, traducendo letteralmente dall'inglese, *Caduta Attesa* (Artzner et al., 1999).

Sia r_t , $t=1, \dots, N$, una serie storica di rendimenti relativi ad una dato attivo finanziario. L'*Expected Shortfall* (ES) al livello α relativo al periodo $[N, N+1)$ è dato dal valore atteso dei rendimenti condizionatamente all'informazione passata ed al fatto che si sia verificata una violazione del VaR, ovvero

$$ES_N(1; \alpha) = E(r_{N+1} \mid r_{N+1} < \text{VaR}_N(1; \alpha), \mathfrak{F}_N) \quad (22)$$

Dalla definizione fornita è facile dedurre che

$$ES_N(1; \alpha) < \text{VaR}_N(1; \alpha)$$

Sul piano economico, l'ES misura la perdita a cui mediamente si andrà incontro nei casi in cui i rendimenti osservati risultino inferiori al VaR.

Analogamente a quanto già visto per il VaR, la definizione (22) è facilmente generalizzabile al caso in cui il periodo di detenzione programmato

sia di lunghezza $h > 1$. La definizione (22) viene ad essere riformulata in termini della distribuzione dei rendimenti cumulati sul periodo $[N, N+k)$ ovvero

$$ES_N(h; \alpha) = E(r_{N+[h]} | r_{N+[h]} < VaR_N(h; \alpha), \mathfrak{F}_N) \quad (23)$$

In aggiunta alla discussione di carattere teorico sulle proprietà dell'*Expected Shortfall* introdotta all'inizio del paragrafo, sul piano più strettamente statistico c'è da osservare che il calcolo dell'ES, rispetto al VaR, presenta il vantaggio di utilizzare anche l'informazione contenuta nella coda sinistra della distribuzione, cosa che invece non avviene con il VaR. Per assurdo, infatti, si potrebbe contaminare l'intera coda sinistra della distribuzione, sostituendo le osservazioni al di sotto del VaR con dei valori eccezionalmente piccoli, a patto che questi siano inferiori al VaR, senza che il valore del VaR stesso ne risulti influenzato. Al tempo stesso, però, ne risulterebbe fortemente influenzato il valore dell'ES. In conclusione, l'ES non va quindi necessariamente visto come un'alternativa al VaR ma come un'utile informazione destinata a completare ed integrare la valutazione del rischio espressa dal VaR.

(B) LA STIMA DELL'ES PER MODELLI ARMA-GARCH

Illustriamo ora delle tecniche per la stima dell'ES quando il modello di riferimento rientra nella classe ARMA-GARCH. Tratteremo dapprima il caso in cui il periodo di detenzione è $h=1$ per poi generalizzare l'esposizione al caso $h>1$.

Si assuma, in generale, che sia $r_t \sim \text{ARMA}(p, q)\text{-GARCH}(s, v)$. Si noti che, per le equazioni (21), si ha $r_{N+[1]} = r_N(1) + \varepsilon_{N+1}$ e $\varepsilon_{N+1} = a_{N+1} \sigma_{N+1}$, ed è possibile scrivere

$$\begin{aligned} ES_N(1; \alpha) &= E(r_{N+1} | \mathfrak{F}_N, r_{N+1} < VaR_N(1; \alpha)) = \\ &= E(r_N(1) + a_{N+1} \sigma_{N+1} | \mathfrak{F}_N, r_{N+1} < VaR_N(1; \alpha)) \end{aligned}$$

Si osservi inoltre che, noti i parametri del modello ARMA, il valore di $r_N(1) = E(r_{N+1} | \mathfrak{F}_N)$ e σ_{N+1} sono determinati solo da \mathfrak{F}_N e quindi noti. Tenendo anche conto dell'indipendenza di a_{N+1} e \mathfrak{F}_N , ne segue che

$$ES_N(1; \alpha) = r_N(1) + E(a_{N+1} \sigma_{N+1} | \mathfrak{F}_N, r_{N+1} < VaR_N(1; \alpha)) =$$

$$= r_N(1) + \sigma_{N+1} E(a_{N+1} | r_{N+1} < \text{VaR}_N(1; \alpha)) \quad (24)$$

La precedente relazione può essere ulteriormente semplificata se si tiene conto del fatto che la disequazione $r_{N+1} < \text{VaR}_N(1; \alpha)$ può essere riformulata nel modo che segue

$$r_{N+1} < \text{VaR}_N(1; \alpha) \approx r_N(1) + a_{N+1} \sigma_{N+1} < r_N(1) + a_\alpha \sigma_{N+1} \approx a_{N+1} < a_\alpha$$

Sostituendo nella (24) si ottiene:

$$ES_N(1; \alpha) = r_N(1) + \sigma_{N+1} c_\alpha \quad (25)$$

dove $c_\alpha = E(a_{N+1} | a_{N+1} < a_\alpha)$. Una stima di $ES_N(1; \alpha)$ può essere ottenuta sostituendo ai parametri incogniti nella (25), ovvero $(\phi_i, \theta_j, \omega, \alpha_h, \beta_k, c_\alpha)$, le corrispondenti stime, per $i = 0, \dots, p$; $j = 1, \dots, q$; $h = 1, \dots, s$; $k = 1, \dots, v$.

Sul piano pratico, la (25) implica che, se si è interessati a stimare $ES_N(1; \alpha)$ a partire da un dato modello ARMA-GARCH precedentemente identificato e stimato, l'unico parametro aggiuntivo da stimare è c_α che viene così a giocare un ruolo analogo a quello del percentile a_α nella stima del VaR. Il problema della stima di c_α viene affrontato diversamente a seconda che la distribuzione dell'errore a_t sia nota oppure no. Nel primo caso inoltre, si presentano difficoltà diverse a seconda che la distribuzione dell'errore sia di tipo gaussiano o t di Student.

(a) *Stima non-parametrica di c_α nel caso in cui la distribuzione di a_t non sia nota*

La necessità di ricorrere ad uno stimatore non parametrico di c_α si fa sentire in particolar modo nel caso in cui le stime dei parametri del modello ARMA-GARCH siano state ottenute con il metodo della Quasi Massima Verosimiglianza. Tale approccio, pur permettendo di stimare in maniera consistente i parametri dei modelli ARMA e GARCH per le componenti di media e varianza condizionata, non consente di fare inferenza sulla forma della distribuzione di a_t e dei suoi parametri. Una stima di c_α può tuttavia essere ancora ottenuta utilizzando la semplice procedura che viene di seguito descritta:

- si salvano i residui standardizzati del modello ARMA-GARCH stimato per la serie storica dei rendimenti osservati dal tempo 1 al tempo N:

$(\hat{a}_1, \dots, \hat{a}_N)$;

- si calcola il quantile empirico di ordine α della serie storica dei residui standardizzati: \hat{a}_α ;
- si stima c_α come media dei residui \hat{a}_t tali che sia $\hat{a}_t < \hat{a}_\alpha$:

$$\hat{c}_\alpha = \frac{\sum_{t=1}^N \hat{a}_t I(\hat{a}_t < \hat{a}_\alpha)}{\sum_{t=1}^N I(\hat{a}_t < \hat{a}_\alpha)}$$

ove, come di consueto, $I(\cdot)$ rappresenta la funzione indicatrice.

(b) *Stima di c_α nel caso in cui $a_t \stackrel{iid}{\sim} N(0, 1)$*

Nel caso in cui $a_t \sim N(0, 1)$, una stima di c_α può essere ottenuta in maniera analitica. Si può infatti dimostrare che

$$c_\alpha = -\frac{\phi(\Phi^{-1}(\alpha))}{\alpha} \quad (26)$$

dove con $\phi(\cdot)$ e $\Phi(\cdot)$ si sono indicate, rispettivamente, le funzioni di densità e di ripartizione della v.c. normale standardizzata, con $\Phi^{-1}(\alpha)$ si è invece indicato il quantile teorico di ordine α della medesima variabile, pari a -1.645, per $\alpha=0.05$, ed a -2.33, per $\alpha=0.01$. I corrispondenti valori di c_α saranno invece $c_{0.05} = -2.063$ e $c_{0.01} = -2.665$.

(c) *Stima di c_α nel caso in cui $a_t \stackrel{iid}{\sim} ((g-2)/g)^{1/2} t_{(g)}$*

Nel caso in cui la distribuzione dell'errore a_t è proporzionale ad una v.c. t di Student standardizzata, non esiste, in generale, una formula analitica per il calcolo di c_α . Anche se, in teoria, è pur sempre possibile utilizzare la procedura non-parametrica precedentemente illustrata al punto (a). Una stima più efficiente di c_α può essere ottenuta attraverso una semplice procedura basata sulla simulazione Monte Carlo. Tale procedura può essere rappresentata in maniera schematica attraverso i seguenti passi:

- si procede alla stima dei parametri del modello ARMA-GARCH di riferimento utilizzando il metodo della massima verosimiglianza;
- si generano indipendentemente m realizzazioni pseudo-casuali di una variabile $t_{(g)}$. Si procede quindi a standardizzare i valori ottenuti moltiplicandoli per $\sqrt{(g-2)/g}$, ovvero il reciproco della deviazione standard di una v.c. $t_{(g)}$. Il risultato sarà un campione di m osservazioni da una v.c. t di Student standardizzata: $(\tilde{a}_1, \tilde{a}_2, \dots, \tilde{a}_m)$ con g gradi di libertà. Compatibilmente con la capacità di calcolo delle risorse hardware e software disponibili, al fine di garantire una maggiore efficienza delle stime, si consiglia di adottare un valore di m il più possibile elevato ($g > 10^5$);
- si calcola il quantile empirico di ordine α del campione estratto, che indicheremo con \tilde{a}_α ;
- una stima di c_α viene quindi ottenuta calcolando la media aritmetica delle \tilde{a}_i tali che $\tilde{a}_i < \tilde{a}_\alpha$. In simboli

$$\hat{c}_\alpha = \frac{\sum_{t=1}^m \tilde{a}_t I(\tilde{a}_t < \tilde{a}_\alpha)}{\sum_{t=1}^m I(\tilde{a}_t < \tilde{a}_\alpha)}$$

Tabella 18 – Valori simulati di c_α per la v.c. t di Student standardizzata per diversi valori del parametro g ($m=5 \times 10^5$)

g	$c_{0.05}$	$c_{0.01}$	g	$c_{0.05}$	$c_{0.01}$
3	-2.242	-4.022	10	-2.149	-3.007
4	-2.272	-3.713	11	-2.147	-2.966
5	-2.240	-3.460	12	-2.139	-2.942
6	-2.218	-3.285	13	-2.128	-2.927
7	-2.198	-3.166	14	-2.122	-2.895
8	-2.168	-3.097	15	-2.129	-2.890
9	-2.172	-3.040	1000	-2.063	-2.679

Nella tabella 18 abbiamo riportato i valori simulati del parametro c_α

nell'ipotesi che la distribuzione dell'errore a_t sia di tipo t di Student standardizzata per diversi valori dei gradi di libertà. Come atteso, al crescere di g , i valori di c_α tendono ai corrispondenti valori teorici precedentemente calcolati per la v.c. normale standardizzata.

(C) LA STIMA DELL'*ES* PER $h > 1$

La stima dell'*ES* per orizzonti temporali di durata $h > 1$ comporta delle difficoltà aggiuntive rispetto al caso $h=1$. Similmente a quanto visto per il VaR, tali difficoltà scaturiscono essenzialmente dall'incertezza associata alle realizzazioni future del processo dei rendimenti dalle quali viene a dipendere il valore di $ES_N(h; \alpha)$. In generale, non è possibile derivare una formula analitica per il calcolo di $ES_N(h; \alpha)$ e, come per il VaR, per la sua stima si rende necessario il ricorso ad una procedura numerica basata su tecniche di simulazione.

Sul piano operativo, si procede in due passi. Prima si stima $VaR_N(h; \alpha)$ utilizzando l'algoritmo illustrato nel paragrafo 6.8. Quindi, si aggiunge alla procedura l'ulteriore passo seguente:

- sia

$$\mathbf{c}_{N,h} = \left[r_{N+[h]}^{(1)} \quad \cdots \quad r_{N+[h]}^{(i)} \quad \cdots \quad r_{N+[h]}^{(m)} \right]$$

il vettore simulato dei rendimenti cumulati $\mathbf{c}_{N,h}$. Si selezionano, all'interno di $\mathbf{c}_{N,h}$, tutte le osservazioni $r_{N+[h]}^{(i)}$ il cui valore risulti inferiore al VaR stimato ovvero tali che siano

$$r_{N+[h]}^{(i)} < \hat{VaR}_N(h; \alpha)$$

Il valore di $ES_N(h; \alpha)$ viene quindi stimato calcolando la media aritmetica di tali osservazioni

$$\hat{\text{VaR}}_N(h; \alpha) = \frac{\sum_{i=1}^m r_{N+[h]}^{(i)} I\left(r_{N+[h]}^{(i)} < \hat{\text{VaR}}_N(h, \alpha)\right)}{\sum_{i=1}^m I\left(r_{N+[h]}^{(i)} < \hat{\text{VaR}}_N(h, \alpha)\right)}$$

Esempio 9

In questo esempio mostreremo come, partendo dalle stime del VaR ottenute per gli attivi Fiat e Campari nell'esempio 8, sia possibile ottenere con uno sforzo aggiuntivo estremamente contenuto, sia in termini di tempi di programmazione che di elaborazione, anche delle stime dell'*Expected Shortfall*. Per quanto riguarda l'implementazione software della procedura, ai fini del calcolo dell'ES, come viene di seguito illustrato, è sufficiente aggiungere alcune linee di codice, evidenziate in grassetto, al programma in linguaggio Eviews utilizzato per lo svolgimento dei calcoli dell'esempio 8.

Inizializzo il vettore destinato a contenere i valori di ES per $b=1,2,\dots,22$.

vector(maxh,1) esmat

'Calcolo di VaR ed ES per $b=1,2,\dots,22$.

for li =1 to maxh

matrix temp1=@subextract(vret,1,1, li,nsim)

vector temp2=@csum(temp1)

rowvector temp3=@transpose(temp2)

rowplace(cvret,temp3,li)

varmat(li)=@quantile(temp3,alpha)

rowvector temp4=@sort(temp3)

rowvector temp5=@subextract(temp4,1,1,1,nsim*alpha)

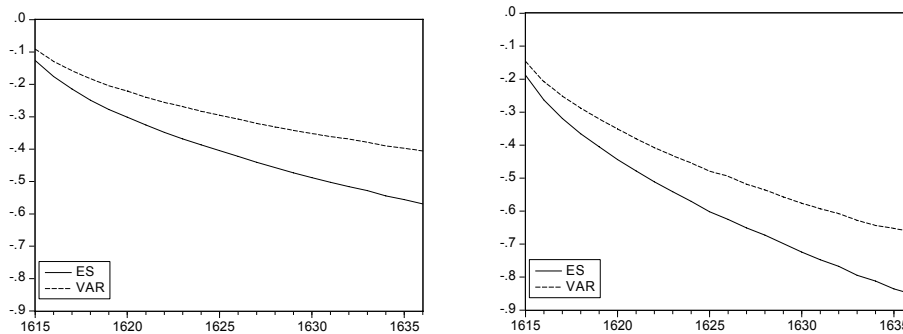
esmat(li)=@mean(temp5)

next

Nelle figure 13 e 14 gli ES stimati per i rendimenti giornalieri sui titoli Campari e Fiat vengono confrontati graficamente con i rispettivi VaR considerando due differenti livelli di copertura, $\alpha=0.05$ e $\alpha=0.01$, ed un orizzonte massimo di previsione pari a 22 giorni approssimativamente equivalenti ad 1 mese di contrattazioni. Come atteso, le stime dell'ES sono sempre inferiori a quelle del VaR e, similmente a quanto visto per il VaR, il valore dell'ES tende a diminuire

al diminuire di α .

Figura 13 – *VaR ed ES stimati per il titolo Fiat, con $1 \leq b \leq 22$, $\alpha=0.05$ (sinistra) e $\alpha=0.01$ (destra), utilizzando un modello MA(0)-GARCH(1,1) con errori t*



I valori stimati dell'ES, riportati in tabella 19 per $h = 1, \dots, 10$, confermano come, anche sulla base dell'ES, il titolo FIAT risulti molto più rischioso del titolo Campari. In situazioni di mercato turbolente in cui il VaR al livello $\alpha=0.05$ viene violato, gli investitori Fiat perderanno in media il 12.57% del capitale investito, in un giorno, ed il 39.16% dopo 10 giorni. Questi valori scendono a 7.5% e 17.58% per il titolo Campari. Analoghe considerazioni valgono per il caso $\alpha=0.01$. Le perdite attese saranno, a 1 e 10 giorni, 18.93% e 56.82% per Fiat, e 10.61% e 24.28% per Campari.

Figura 14 – *VaR ed ES stimati per il titolo Campari, con $1 \leq b \leq 22$, $\alpha=0.05$ (sinistra) e $\alpha=0.01$ (destra), utilizzando un modello MA(1)-GARCH(1,1) con errori di Student*

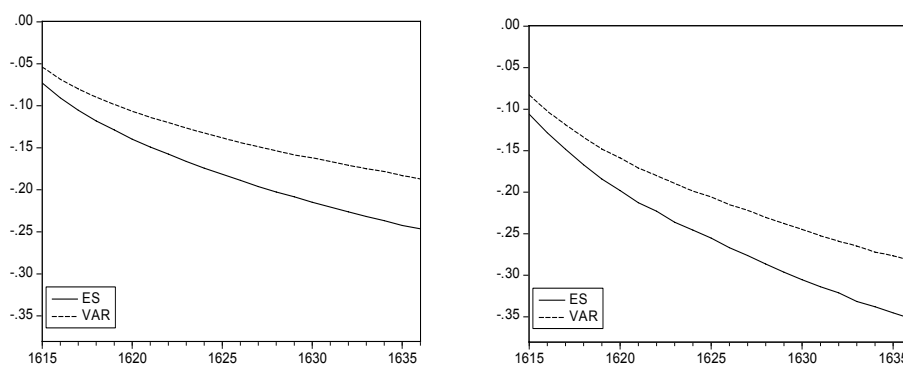


Tabella 19 – *VaR ed ES stimati per il titolo Campari, con $1 \leq b \leq 22$, $\alpha=0.05$ (sinistra) e $\alpha=0.01$ (destra), utilizzando un modello MA(1)-GARCH(1,1) con errori t di Student*

H	ES _{FIAT,0.05}	ES _{CAMP,0.05}	ES _{FIAT,0.01}	ES _{CAMP,0.01}
1	-0.12573	-0.07335	-0.18934	-0.10607
2	-0.17723	-0.09186	-0.26067	-0.12974
3	-0.21663	-0.10693	-0.31337	-0.14914
4	-0.24926	-0.11948	-0.36019	-0.16676
5	-0.27952	-0.1304	-0.39934	-0.18365
6	-0.30715	-0.14149	-0.43889	-0.19657
7	-0.32948	-0.1507	-0.47686	-0.20883
8	-0.35183	-0.15976	-0.50747	-0.2209
9	-0.37158	-0.16788	-0.53988	-0.23299
10	-0.39157	-0.17579	-0.56822	-0.24282

□

Ci sembra opportuno ricordare come il valore del VaR stimato, invece, semplicemente informa l'investitore che, con probabilità pari ad $1-\alpha$, i rendimenti non saranno inferiori ad una determinata soglia: il VaR. Nessuna informazione viene però fornita in merito al probabile andamento delle perdite future nel caso in cui si scenda al di sotto di tale soglia.

6.10 Valutazione della qualità delle stime VaR: il backtesting

(A) IL BACKTESTING

Gli operatori finanziari, per motivi di natura sia normativa che strategica, hanno necessità di avere a disposizione stime del rischio di mercato realistiche e costantemente aggiornate. Il numero di modelli e procedure statistiche potenzialmente utilizzabili per il perseguimento di questo fine è elevatissimo ed in continua crescita. Si rende quindi necessario predisporre degli strumenti statistici che permettano di: i) misurare la qualità delle stime del VaR ottenute attraverso un determinato approccio; ii) confrontare le stime ottenute da diversi approcci alternativi.

Tali valutazioni vanno ovviamente effettuate prima che un dato modello venga utilizzato sul campo. Al fine di rendere possibile la valutazione della qualità delle stime, inoltre, è necessario disporre dei valori dei rendimenti osservati per i periodi relativamente ai quali si vuole misurare il rischio. Ciò è ovviamente possibile solo se la valutazione viene fatta *a posteriori* basandosi su dati storici.

Nello specifico, si considera una serie storica, di lunghezza il più possibile elevata, di rendimenti passati relativa all'attivo o al portafoglio oggetto di interesse. Tale serie viene poi divisa in due parti delle quali la prima viene utilizzata per l'identificazione e la stima dei modelli mentre la seconda viene utilizzata per la stima del VaR e la valutazione della qualità delle stime ottenute. In pratica, i passi della procedura possono essere schematizzati nel modo seguente:

- data una serie storica di rendimenti di lunghezza N , se ne selezionano le prime M osservazioni ($M < N$): (r_1, r_2, \dots, r_M) ;
- a partire dalla serie (r_1, r_2, \dots, r_M) , si identifica un opportuno modello statistico (ad es. ARMA-GARCH) per i rendimenti e si utilizza tale modello per ottenere delle stime per il VaR relativo agli istanti da $(M+1)$ a N :

$$(\text{VaR}_M(1; \alpha), \text{VaR}_{M+1}(1; \alpha), \dots, \text{VaR}_{N-1}(1; \alpha))$$

- le stime del VaR relative al periodo di validazione, da $M+1$ a N , vengono confrontate con i rendimenti relativi allo stesso periodo al fine di valutarne le proprietà statistiche ed economiche.

La procedura illustrata è nota, nella terminologia degli analisti finanziari, con il nome di *backtesting* e può essere generalizzata in maniera pressoché immediata al caso di stime VaR con $h > 1$. Inoltre, al fine di tenere conto della presenza di cambiamenti strutturali nel processo che genera i rendimenti e le loro volatilità, è sufficiente introdurre un semplice correttivo che consiste nel prevedere la possibilità di ripetere le fasi di identificazione e stima del modello ad intervalli costanti, sulla base di una finestra mobile di osservazioni. Ad esempio, il modello viene inizialmente stimato ed identificato sulla base delle prime M osservazioni in $[0, M)$ cioè (r_1, r_2, \dots, r_M) . Utilizzando tale modello si procede a stimare il VaR per il periodo $[M, M+k)$, con $k > 0$. Quindi, si ripete la procedura di identificazione e stima del modello utilizzando le osservazioni relative al periodo $[k, k+M)$: $(r_{k+1}, r_{k+2}, \dots, r_{M+k})$. Il nuovo modello così ottenuto viene utilizzato per stimare il VaR per il periodo $[M+k, M+2k)$ fino a che le fasi di identificazione e stima vengono nuovamente ripetute sulla base della serie $(r_{M+k+1}, r_{M+k+2}, \dots, r_{M+2k})$. I passi descritti vengono ripetuti fino ad esaurimento delle osservazioni disponibili. In tal modo, si prevede implicitamente che il modello ottimale per i dati osservati possa cambiare nel tempo. Di fatto, tale strategia equivale a sostituire un modello fisso nel tempo con un modello flessibile la cui forma funzionale e parametri possono variare

tenendo conto di effetti di natura congiunturale e/o breaks strutturali permanenti.

Nel paragrafo che segue verranno discusse alcune procedure di test che vengono comunemente utilizzate per la valutazione delle proprietà statistiche delle stime VaR. In particolare, saranno oggetto di trattazione alcuni test statistici che vengono spesso sinteticamente raggruppati sotto la denominazione di *tests di copertura*.

(B) TESTS DI COPERTURA

Sul piano statistico, il problema della stima del VaR coincide con la stima di un quantile estremo, di ordine $\alpha \ll 1$, della distribuzione condizionata dei rendimenti. E' quindi perfettamente ammissibile che occasionalmente il valore osservato del rendimento risulti inferiore al VaR senza che questo implichi necessariamente che il VaR non sia stato stimato in maniera accurata. Nello specifico, in corrispondenza di un modello correttamente specificato, ci si attende che, su un periodo sufficientemente lungo, i rendimenti osservati risultino inferiori al VaR all'incirca nell' $\alpha\%$ dei casi.

Formalmente, è utile definire la serie delle variabili dicotomiche

$$L_{M+j} = I(r_{M+j} < \text{VaR}_{M+j-1}(1, \alpha)) = \begin{cases} 1 & \text{se } r_{M+j} < \text{VaR}_{M+j-1}(1, \alpha) \\ 0 & \text{se } r_{M+j} \geq \text{VaR}_{M+j-1}(1, \alpha) \end{cases}$$

che vengono anche dette *eccezioni* o *violazioni*, con $j = 1, \dots, N-M$. E' immediato constatare che ogni L_{M+j} si distribuisce come una variabile casuale di

Bernoulli con parametro (probabilità di successo) pari ad α :

$$L_{M+j} \sim B(1; \alpha), \quad j = 1, \dots, N-M.$$

Si noti poi che, assumendo che il modello di riferimento sia della classe (T)ARMA-GARCH, è possibile scrivere

$$\begin{aligned} L_{M+j} &= I(r_{M+j} < \text{VaR}_{M+j-1}(1; \alpha)) = \\ &= I(r_{M+j-1}(1) + a_{M+j} \sigma_{M+j} < r_{M+j-1}(1) + a_{\alpha} \sigma_{M+j}) = I(a_{M+j} < a_{\alpha}) \end{aligned} \quad (27)$$

Ne segue che le eccezioni L_{M+j} , in quanto funzioni di variabili casuali

indipendenti ed identicamente distribuite, saranno anche esse iid

$$L_{M+j} \stackrel{\text{iid}}{\sim} B(1; \alpha), \quad j = 1, \dots, N-M \quad (28)$$

Nelle pratiche applicazioni, non è possibile osservare le L_{M+j} ma solo delle loro stime, che indicheremo con \hat{L}_{M+j} , ottenute sostituendo a $r_{M+j-1}(1)$, σ_{M+j-1}^2 e a a_α le loro stime $\hat{r}_{M+j-1}(1)$, $\hat{\sigma}_{M+j-1}^2$ ed \hat{a}_α . Se il modello per i rendimenti è correttamente specificato, ci attenderemo che la distribuzione delle \hat{L}_{M+j} non si discosti significativamente da quella delle L_{M+j} . Qualora, invece, il modello per i rendimenti non descriva adeguatamente il processo generatore sottostante, ci attenderemo che anche la distribuzione \hat{L}_{M+j} risulti marcatamente differente da quella teorica attesa per le L_{M+j} sulla base della (28) e, in particolare, potrebbero verificarsi le seguenti ipotesi:

- (h1): $\hat{L}_{M+j} \sim B(1; \alpha')$ con $\alpha' \neq \alpha$
- (h2): le \hat{L}_{M+j} non sono indipendenti.
- (h3): sia la h1 che la h2 risultano verificate.

In letteratura sono state proposte delle statistiche test per sottoporre a verifica, rispettivamente, ciascuna delle precedenti ipotesi:

- test di corretta copertura marginale (h1)
- test di indipendenza (h2)
- test di corretta copertura condizionata (h3)

(h1) Test di corretta copertura marginale

Kupiec (1995) ha proposto una statistica test per verificare l'ipotesi di corretta specificazione del modello contro un'ipotesi alternativa del tipo (h1). Questo test è noto anche come *test di corretta copertura marginale*. Nello specifico il sistema di ipotesi da sottoporre a verifica è:

$$\begin{cases} H_0 : \hat{L}_{M+j} \sim B(1; \alpha_0) \\ H_1 : \hat{L}_{M+j} \sim B(1; \alpha), \text{ con } \alpha \neq \alpha_0 \end{cases}$$

dove con α_0 si è indicato il livello di copertura nominale desiderato. Una statistica test per verificare tale sistema di ipotesi può essere ottenuta attraverso il metodo del rapporto di verosimiglianza.

Assumendo che le \hat{L}_{M+j} siano indipendenti, la funzione di verosimiglianza è data da

$$L(\alpha; \hat{L}_{M+j}) = \alpha^{n_e} (1-\alpha)^{n_e} \quad \text{per } j=1, \dots, N-M \quad (29)$$

dove

$$n_e = \sum_{j=1}^{N-M} I(r_{M+j} < \text{VaR}_{M+j-1}(1; \alpha))$$

rappresenta il numero delle eccezioni del VaR avvenute nel periodo di riferimento. Le statistiche test ottenute attraverso il metodo del rapporto di verosimiglianza possono essere interpretate come delle misure di distanza fra il massimo globale della funzione di verosimiglianza, $L(\hat{\alpha}; \hat{L}_{M+j})$, raggiunto in corrispondenza dello stimatore di massima verosimiglianza $\hat{\alpha}$, ed il massimo vincolato ottenuto sotto l'ipotesi che i vincoli espressi dall'ipotesi nulla siano verificati. Nel nostro caso, quest'ultimo valore viene ottenuto calcolando semplicemente il valore della funzione di verosimiglianza in corrispondenza del punto α_0 :

$$L(\alpha_0; \hat{L}_{M+j}) = \alpha_0^{n_e} (1-\alpha_0)^{n_e} \quad (30)$$

e la statistica test è data da

$$\text{CM}_M = -2 \log \left(\frac{L(\alpha_0; \hat{L}_{M+j})}{L(\hat{\alpha}; \hat{L}_{M+j})} \right) = -2 \left[\log(L(\alpha_0; \hat{L}_{M+j})) - \log(L(\hat{\alpha}; \hat{L}_{M+j})) \right] =$$

$$\begin{aligned}
&= -2[\ell(\alpha_0; \hat{L}_{M+j}) - \ell(\hat{\alpha}; \hat{L}_{M+j})] = \\
&= -2[n_e \log(\alpha_0) + (n - n_e) \log(1 - \alpha_0) - n_e \log(\hat{\alpha}) - (n - n_e) \log(1 - \hat{\alpha})]
\end{aligned}$$

per $j = 1, \dots, n$, con $n=N-M$. Massimizzando la (29) rispetto ad α , inoltre, è semplice verificare che lo stimatore di massima verosimiglianza di α è dato da:

$$\hat{\alpha} = \frac{n_e}{n}$$

Si può dimostrare che, nel caso H_0 risulti verificata, se n è sufficientemente grande e se le \hat{L}_{M+j} sono indipendenti, allora $CM_M \approx \chi_{(1)}^2$. La regione critica del test si colloca nella coda destra della distribuzione ovvero rifiuteremo H_0 al livello β se $CM_M > \chi_{(1;1-\beta)}^2$, dove con $\chi_{(1;1-\beta)}^2$ si è indicato il quantile di ordine $(1-\beta)$ di una variabile casuale chi-quadrato con 1 grado di libertà.

Se viene meno l'ipotesi che le \hat{L}_{M+j} sono indipendenti, la funzione di verosimiglianza non è più data dalla (29) e, conseguentemente, anche i risultati sulla distribuzione asintotica di CM_M non sono più applicabili.

(b2) Test di indipendenza

Al fine di valutare l'attendibilità dei risultati del test di corretta copertura marginale, si rende quindi necessario verificare l'ipotesi di indipendenza delle \hat{L}_{M+j} attraverso un *test di indipendenza* delle \hat{L}_{M+j} (Christoffersen, 1998).

Tale test verifica l'ipotesi nulla di indipendenza delle \hat{L}_{M+j} contro l'ipotesi alternativa che le \hat{L}_{M+j} siano dipendenti e, in particolare, che siano generate da una *Catena di Markov* del primo ordine.

Catena di Markov: Un processo stocastico a valori discreti X_t è una Catena di Markov del primo ordine se soddisfa la seguente condizione nota come *Condizione di Markov*

$$P(X_t = x_t | X_0 = x_0, X_1 = x_1, \dots, X_{t-1} = x_{t-1}) = P(X_t = x_t | X_{t-1} = x_{t-1}) \quad (31)$$

$\forall t \geq 1$ e $\forall \{x_0, x_1, \dots, x_{t-1}\} \in S$, dove S è l'insieme discreto in cui il processo X_t assume valori.

Nel caso di indipendenza delle \hat{L}_{M+j} , si avrà che

$$P(\hat{L}_{M+j} = 1) = P(\hat{L}_{M+j} = 1 | \hat{L}_{M+j-1} = 0) = P(\hat{L}_{M+j} = 1 | \hat{L}_{M+j-1} = 1)$$

ovvero la distribuzione di probabilità delle eccezioni rimane invariata indipendentemente dal fatto che nell'istante precedente si sia avuta una violazione del VaR o meno.

Nel caso in cui le \hat{L}_{M+j} siano state generate da una Catena di Markov, invece, tale uguaglianza non sarà più verificata ma, in generale, si avrà

$$P(\hat{L}_{M+j} = 1) \neq P(\hat{L}_{M+j} = 1 | \hat{L}_{M+j-1} = 0) \neq P(\hat{L}_{M+j} = 1 | \hat{L}_{M+j-1} = 1) \quad (32)$$

In quest'ultimo caso la distribuzione condizionata delle \hat{L}_{M+j} sarà ancora di tipo Bernoulli con parametro che varia a seconda del valore delle \hat{L}_{M+j} osservato al tempo precedente

$$P(\hat{L}_{M+j} = 1 | \hat{L}_{M+j-1} = s) = \alpha_{1s}$$

con $s \in \{0, 1\}$. Il sistema di ipotesi da verificare diviene quindi

$$\begin{cases} H_0 : \alpha = \alpha_{11} = \alpha_{10} \\ H_1 : \alpha_{11} \neq \alpha_{10} \end{cases}$$

Come il precedente, anche questo è un test del rapporto di verosimiglianza. Venendo però meno l'ipotesi di indipendenza delle eccezioni, la funzione di verosimiglianza va costruita in maniera differente da quanto illustrato per il test di corretta copertura marginale. Nel caso di dati dipendenti, analogamente a quanto dimostrato nel capitolo V con riferimento alla stima dei modelli ARCH,

la funzione di verosimiglianza può essere infatti espressa come prodotto delle distribuzioni di probabilità delle singole osservazioni condizionate all'informazione disponibile al tempo t-1

$$\begin{aligned} L(\boldsymbol{\alpha}; \hat{L}_{M+j}) &= p(\hat{L}_{M+1}, \hat{L}_{M+2}, \dots, \hat{L}_{M+n}) \\ &= p(\hat{L}_{M+n} | \hat{L}_{M+1}, \dots, \hat{L}_{M+n-1}) p(\hat{L}_{M+1}, \dots, \hat{L}_{M+n-1}) = \\ &= p(\hat{L}_{M+n} | \mathfrak{S}_{M+n-1}) p(\mathfrak{S}_{M+n-1}) \end{aligned}$$

che per la condizione di Markov può essere riscritta come

$$L(\boldsymbol{\alpha}; \hat{L}_{M+j}) = p(\hat{L}_{M+n} | \hat{L}_{M+n-1}) p(\mathfrak{S}_{M+n-1})$$

con $\boldsymbol{\alpha} = (\alpha_{10}, \alpha_{11})$. Per sostituzioni successive si ottiene

$$L(\boldsymbol{\alpha}; \hat{L}_{M+j}) = p(\hat{L}_M) \prod_{j=1}^n p(\hat{L}_{M+j} | \hat{L}_{M+j-1}) \approx \prod_{j=1}^n p(\hat{L}_{M+j} | \hat{L}_{M+j-1})$$

Dato che

$$p(\hat{L}_{M+j} | \hat{L}_{M+j-1} = 1) = \alpha_{11}^{\hat{L}_{M+j}} (1 - \alpha_{11})^{1 - \hat{L}_{M+j}}$$

ne segue che

$$\begin{aligned} L(\boldsymbol{\alpha}; \hat{L}_{M+j}) &= \prod_{j=1}^n p(\hat{L}_{M+j} | \hat{L}_{M+j-1}) = \prod_{s=0}^1 \alpha_{1s}^{n_{1s}} (1 - \alpha_{1s})^{n_{0s}} = \\ &= \alpha_{10}^{n_{10}} (1 - \alpha_{10})^{n_{00}} \alpha_{11}^{n_{11}} (1 - \alpha_{11})^{n_{01}} \end{aligned} \quad (33)$$

dove con n_{ij} si è indicato il numero di volte che, in un dato istante temporale

t, si è osservato $\hat{L}_t = i$ dopo aver osservato $\hat{L}_{t-1} = j$ nell'istante precedente.

Il valore della (33) in corrispondenza dello stimatore di massima verosimiglianza sarà quindi dato da

$$L(\hat{\boldsymbol{\alpha}}; \hat{L}_{M+j}) = \hat{\alpha}_{10}^{n_{10}} (1 - \hat{\alpha}_{10})^{n_{00}} \hat{\alpha}_{11}^{n_{11}} (1 - \hat{\alpha}_{11})^{n_{01}} \quad (34)$$

con $\hat{\alpha} = (\hat{\alpha}_{10}, \hat{\alpha}_{11})$ e

$$\hat{\alpha}_{10} = \frac{n_{10}}{n_{10} + n_{00}}; \quad \hat{\alpha}_{11} = \frac{n_{11}}{n_{11} + n_{01}}$$

Nel caso di indipendenza, ovvero sotto il vincolo implicato dall'ipotesi nulla H_0 , il massimo vincolato della funzione di verosimiglianza sarà dato da

$$L(\hat{\alpha}_0; \hat{L}_{M+j}) = \hat{\alpha}^{n_{1\cdot}} (1 - \hat{\alpha})^{n_{0\cdot}}$$

dove $\hat{\alpha}_0 = \hat{\alpha}$ e

$$\hat{\alpha} = \frac{n_{1\cdot}}{n_{0\cdot} + n_{1\cdot}}$$

è lo stimatore di massima verosimiglianza di α sotto H_0 , mentre

$$n_{1\cdot} = n_{10} + n_{11} = n_e; \quad n_{0\cdot} = n_{00} + n_{01} = n - n_e$$

rappresentano il numero di volte in cui si sono osservati i valori 1 e 0, rispettivamente, per le \hat{L}_{M+j} , $j = 1, \dots, n$. Il rapporto di verosimiglianza è quindi dato da

$$\begin{aligned} ID_n &= -2 \log \left(\frac{L(\hat{\alpha}_0; \hat{L}_{M+j})}{L(\hat{\alpha}; \hat{L}_{M+j})} \right) = -2 \left[\log \left(L(\hat{\alpha}_0; \hat{L}_{M+j}) \right) - \log \left(L(\hat{\alpha}; \hat{L}_{M+j}) \right) \right] = \\ &= -2 [\ell(\hat{\alpha}_0; \hat{L}_{M+j}) - \ell(\hat{\alpha}; \hat{L}_{M+j})] = \\ &= -2 [n_{1\cdot} \log(\hat{\alpha}) + n_{0\cdot} \log(1 - \hat{\alpha}) - n_{10} \log(\hat{\alpha}_{10}) - n_{00} \log(1 - \hat{\alpha}_{10}) - n_{11} \log(\hat{\alpha}_{11}) + \\ &\quad - n_{01} \log(1 - \hat{\alpha}_{11})] \end{aligned}$$

Anche in questo caso si può dimostrare che, per n sufficientemente grande, sotto H_0 , $ID_n \approx \chi_{(1)}^2$. La regione critica del test si colloca nella coda destra

della distribuzione ovvero rifiuteremo H_0 al livello β se $ID_n > \chi_{(1;1-\beta)}^2$.

(b3) *Test di corretta copertura condizionata*

Il test di corretta copertura condizionata (Christoffersen, 1998) permette di verificare simultaneamente le due ipotesi precedenti. Il sistema delle ipotesi da sottoporre a verifica è

$$\begin{cases} H_0 : \alpha_{11} = \alpha_{10} = \alpha_0 \\ H_1 : \alpha_{11} \neq \alpha_{10} \end{cases}$$

che, come si può facilmente osservare, può essere ottenuto combinando l'ipotesi nulla del test di corretta copertura marginale con l'ipotesi alternativa del test di indipendenza. La statistica test è data dal rapporto di verosimiglianza

$$CC_n = -2 \left[\log(L(\alpha_0; \hat{L}_{M+j})) - \log(L(\hat{\alpha}; \hat{L}_{M+j})) \right]$$

dove $L(\alpha_0; \hat{L}_{M+j})$ ed $L(\hat{\alpha}; \hat{L}_{M+j})$ sono definite come nelle equazioni (30) e (34), rispettivamente. Ne segue che la statistica test è data da

$$CC_M = -2 [n_{1.} \log(\alpha_0) + n_{0.} \log(1 - \alpha_0) - n_{10} \log(\hat{\alpha}_{10}) - n_{00} \log(1 - \hat{\alpha}_{10}) + \\ - n_{11} \log(\hat{\alpha}_{11}) - n_{01} \log(1 - \hat{\alpha}_{11})]$$

E' possibile dimostrare (Christoffersen, 1998) che, per n sufficientemente grande, sotto H_0 , $ID_n \approx \chi_{(2)}^2$. Si rifiuterà H_0 al livello di significatività β per valori di $CC_n > \chi_{(2;1-\beta)}^2$. E' inoltre molto semplice dimostrare che vale la seguente relazione

$$CM_n + ID_n = CC_n \tag{35}$$

con la quale si viene a stabilire un preciso legame teorico fra i tre tests presentati.

Esempio 10

In questo esempio vengono presentati i risultati di un esercizio di

backtesting sui rendimenti giornalieri dei titoli Fiat e Campari. In entrambi i casi, le osservazioni disponibili (1613) vengono preventivamente divise in due sottoserie. La prima sottoserie, che ricomprende le osservazioni dal giorno 1 al giorno 1363, viene utilizzata per l'identificazione e la stima di un opportuno modello ARMA-GARCH per i rendimenti.

Tabella 20 – *Stime dei parametri dei modelli GARCH(1,1) per i rendimenti giornalieri su Fiat e Campari stimati con il metodo della QMV e della MV (osservazioni dal giorno 1 al giorno 1363)*

Titolo		w	α	β	g	AIC
Fiat (QMV)	Stima Prob.	2.47* (.0015)	0.1070 (.0000)	0.8329 (.0000)		-5.053
Fiat (MV)	Stima Prob.	.2.63* (.0012)	0.1285 (.0000)	0.8139 (.0000)	6.1376 (.0000)	-5.088
Campari (QMV)	Stima Prob.	3.82* (.1432)	0.3341 (.0043)	0.9488 (.0000)		-5.691
Campari (MV)	Stima Prob.	3.14* (.0240)	0.0963 (.0030)	0.7589 (.0000)	5.3946 (.0000)	-5.721

(*: $\times 10^{-5}$)

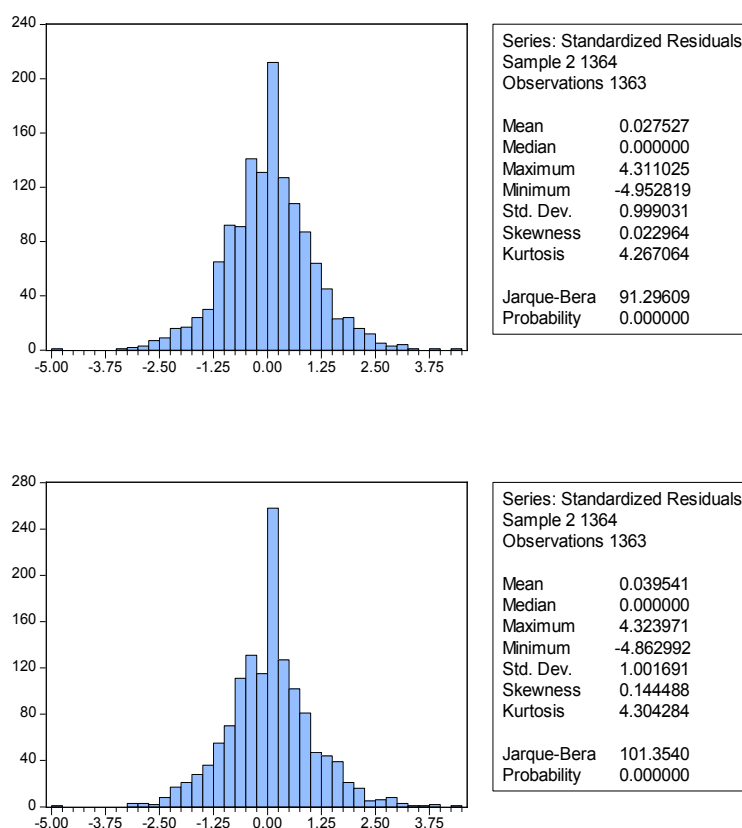
Le rimanenti 250 osservazioni vengono invece utilizzate per la previsione del VaR e la valutazione della qualità delle previsioni ottenute. In pratica, sulla base dei modelli stimati a partire dalla prima sottoserie, si procede a stimare $\text{VaR}_M(1;\alpha)$ per $M=1364, \dots, 1613$. Ad ogni istante viene aggiornato il set informativo condizionante usato per la stima del VaR mentre i parametri del modello di riferimento vengono mantenuti costanti.

Per entrambe le serie viene identificato e stimato un modello ARMA(0, 0) - GARCH(1, 1). La stima dei modelli è stata effettuata attraverso due differenti approcci: i) il metodo della Quasi Massima Verosimiglianza ii) il metodo della Massima Verosimiglianza basato sull'ipotesi che gli errori si distribuiscano come una t di Student standardizzata. Le stime ed i relativi p-value sono stati riportati in tabella 20.

I residui standardizzati dei modelli stimati sono stati riportati nella figura 15, per il metodo della QMV, e figura 16 per, il metodo della MV. In entrambi i casi, i valori della statistica test di Jarque-Bera evidenziano delle significative deviazioni dalla normalità. Il quantile dei residui standardizzati a α è stato

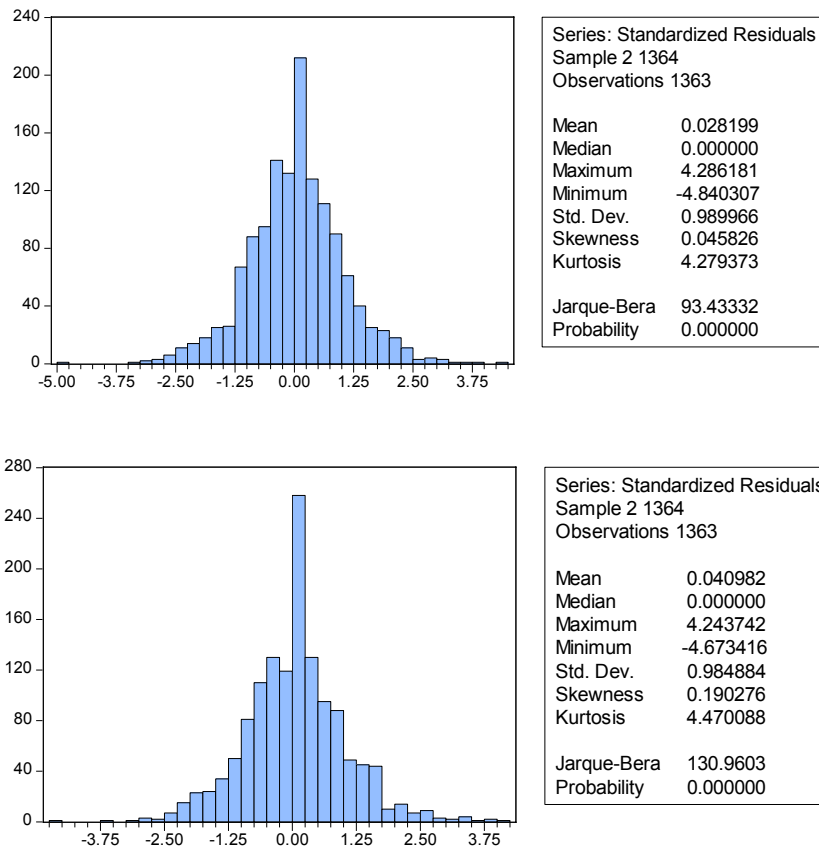
stimato seguendo due differenti approcci. Nel caso in cui le stime dei parametri del modello sono state ottenute con il metodo della QMV, si è ottenuta una stima di α calcolando il quantile empirico dei residui standardizzati.

Figura 15 – Residui standardizzati ottenuti dal modello GARCH stimato con il metodo della QMV per gli attivi FLAT (alto) e CAMPARI (basso)



Diversamente, nel caso delle stime di MV, si è calcolato il quantile di ordine α di una v.c. t di Student standardizzata con numero di gradi di libertà pari al valore stimato. Si è quindi proceduto a stimare il VaR utilizzando le formule derivate nel paragrafo 6.4. Dato il numero relativamente ridotto di osservazioni incluso nel periodo di validazione, 250, si è stimato il VaR solo per $\alpha=0.05$. Al fine di ottenere delle valutazioni attendibili della qualità delle stime del VaR per $\alpha=0.01$, infatti, sarebbe stato necessario prendere in considerazione un periodo di durata molto superiore.

Figura 16 – Residui standardizzati ottenuti dal modello GARCH stimato con il metodo della MV per gli attivi FLAT (alto) e CAMPARI (basso)



La qualità delle stime del VaR ottenute è stata valutata calcolando le statistiche test per verificare le ipotesi di corretta copertura marginale, indipendenza e corretta copertura condizionata. I relativi p-values (tabella 21) mostrano come nessuno dei modelli stimati sia in grado di prevedere adeguatamente l'effettiva esposizione al rischio nel periodo di previsione. Di seguito riportiamo il listato in linguaggio Eviews utilizzato per le elaborazioni

Inizializzazione

scalar maxh=250

smpl @last -maxh+1 @last

scalar alpha0=0.05

genr serie=dlfiat_g

'Genero la serie delle eccezioni e calcolo $\hat{\alpha}$

genr Lt=serie<var05

```
scalar ne=@sum(Lt)
scalar alphahat=ne/maxh
```

'Calcolo il rapporto di verosimiglianza per il test di corretta copertura marginale

```
scalar lnum=ne*log(alpha0)+log(1-alpha0)*(maxh-ne)
scalar lden=ne*log(alphahat)+log(1-alphahat)*(maxh-ne)
scalar lrt_cm=-2*(lnum-lden)
scalar pval_cm=1-@cchisq(lrt_cm,1)
```

'Calcolo del rapporto di verosimiglianza per il test di indipendenza (ID): calcoli preliminari

```
genr count00=0
genr count01=0
genr count10=0
genr count11=0
genr time=@trend+1
scalar end=@max(time)-1
for li=end-(maxh+2) to end
if (Lt(li-1)=0 and Lt(li)=1) then count01(li) =1
endif
if (Lt(li-1)=1 and Lt(li)=0) then count10(li) =1
endif
if (Lt(li-1)=0 and Lt(li)=0) then count00(li)=1
endif
if (Lt(li-1)=1 and Lt(li)=1) then count11(li)=1
endif
next
scalar n01=@sum(count01)
scalar n10=@sum(count10)
scalar n11=@sum(count11)
scalar n00=@sum(count00)
```

```
scalar alpha01=n01/(n00+n01)
scalar alpha11=n11/(n10+n11)
```

'Statistica test ID

```
scalar lnum=(log(1-alphahat)*(n00+n10))+log(alphahat)*(n01+n11)
scalar lden=(log(1-alpha01)*(n00))+log(alpha01)*(n01)+log(1-
alpha11)*(n10)+log(alpha11)*n11
scalar lrt_ind=-2*(lnum-lden)
```

'Pvalue test ID


```
scalar pval_ind=1-@cchisq(lrt_ind,1)
```

```
'Calcolo del rapporto di verosimiglianza per il test di copertura condizionata (CC)
scalar lnum=(ne*log(alpha0))+(log(1-alpha0)*(maxh-ne))
scalar lden=(log(1-alpha01)*n00)+(log(alpha01)*(n01))+(log(1-
alpha11)*n10)+log(alpha11)*n11
```

```
'Statistica test CC
scalar lrt_cc=-2*(lnum-lden)
```

```
'Pvalue test CC
scalar pval_cc=1-@cchisq(lrt_cc,2)
```

Tabella 21 – Test di corretta copertura marginale (CM), indipendenza (ID) e corretta copertura condizionata (CC) per gli attivi Fiat e Campari. Modelli GARCH(1,1) stimati sulle osservazioni da 1 a 1363

FIAT				
Metodo	$\hat{\alpha}$	CM	ID	CC
Quant. emp.	0.108	0.0002	0.5237	0.0009
t di Student	0.104	0.0006	0.6144	0.0021
CAMPARI				
	$\hat{\alpha}$	CM	ID	CC
Quant. emp.	0.084	0.0240	0.3503	0.0463
t di Student	0.128	0.0000	0.5134	0.0000

I valori di $\hat{\alpha}$ mostrano che, sia per Fiat che per Campari, indipendentemente dal metodo di stima utilizzato, il livello di copertura empirico del VaR, $\hat{\alpha}$, è molto più elevato del livello nominale atteso, $\alpha_0=0.05$. L'ipotesi di indipendenza delle "eccezioni" risulta invece sempre verificata indicando che le violazioni del VaR non tendono a "clusterizzarsi" nel tempo. L'assenza di dipendenza nella serie delle eccezioni, inoltre, ha una seconda importante implicazione: la significatività delle differenze fra livello di copertura nominale ed empirico può essere valutata utilizzando sia il test di corretta copertura marginale (CM) che di corretta copertura condizionata (CC). In assenza di dipendenza, per campioni di dimensione sufficientemente elevata, infatti, ci si attende che questi tests diano risultati concordanti. Nel nostro caso, per entrambi gli attivi ed i metodi di stima considerati, entrambi i tests CM e CC ci conducono a rifiutare l'ipotesi che il livello di copertura delle stime VaR sia

effettivamente pari a 0.05.

I risultati ottenuti non debbono sorprenderci dal momento che il periodo utilizzato per la valutazione delle previsioni comprende, approssimativamente, un anno di contrattazioni borsistiche con termine in data 6/3/2009. Va infatti notato come in tale periodo venga a ricadere la fase più acuta della recente crisi finanziaria. I repentini e continui mutamenti che hanno interessato i mercati mondiali nel corso di questo periodo ci inducono a ritenere che sia semplicistico cercare di spiegare l'evoluzione del rischio su questo arco di tempo attraverso un unico modello a parametri fissi. Come già accennato, una soluzione potrebbe essere quella di ristimare il modello ad intervalli di tempo costanti utilizzando una finestra mobile di osservazioni di ampiezza costante. Nel nostro caso, questa strategia è stata implementata ristimando il modello con cadenza giornaliera. In particolare, sono stati utilizzati i primi 1363 rendimenti giornalieri per stimare il VaR alla fine del giorno 1364. Quindi, all'istante successivo, il modello è stato ristimato utilizzando le osservazioni da 2 a 1364. I parametri stimati sono stati quindi utilizzati per stimare il valore del VaR alla fine del giorno 1365, e così via fino all'esaurimento delle osservazioni disponibili. In questo modo le stime dei parametri utilizzati per la stima del VaR vengono aggiornate ogni volta che una nuova osservazione si rende disponibile.

In totale si è proceduto a stimare 250 modelli GARCH(1, 1) per ciascun attivo. A titolo esemplificativo le stime dei parametri ottenute per il titolo Fiat sono state rappresentate graficamente in figura 17. L'analisi dei grafici rivela come tutti i parametri del modello GARCH(1, 1) tendano a variare nel corso del periodo di previsione. In particolare, la persistenza ($\alpha + \beta$) ed il livello di lungo periodo della volatilità $\omega/(1 - \alpha - \beta)$ sono caratterizzati da un marcato trend crescente che, nel caso di $\omega/(1 - \alpha - \beta)$, assume un andamento di tipo esponenziale. Questo dato è particolarmente interessante in quanto implica che l'incertezza di medio lungo-periodo tende a crescere in maniera esponenziale, e quindi potenzialmente esplosiva, nel periodo iniziale dell'anno 2009. Al fine di consentire la replicabilità dell'esercizio, riportiamo schematicamente il listato Eviews utilizzato per la stima dei modelli GARCH(1, 1) su una finestra mobile di osservazioni.

I risultati dei tests di CM, ID e CC effettuati sulle stime VaR basate sulla finestra mobile (tabella 22) mostrano un evidente miglioramento rispetto al caso precedente in cui l'intera serie dei VaR stimati veniva ottenuta facendo riferimento ad un singolo modello stimato su una finestra fissa composta dalle osservazioni iniziali della serie. L'ipotesi nulla di indipendenza delle \hat{L}_t viene accettata per entrambi gli attivi mentre i tests di corretta copertura marginale e condizionata ci conducono ad accettare l'ipotesi nulla che il livello di copertura

empirico non sia significativamente diverso dal suo valore nominale.

Figura 17 – Modello GARCH(1,1) per Fiat: stima su finestra mobile. Da sinistra verso destra e dall'alto verso il basso: valori stimati per $(\alpha + \beta)$, $\omega / (1 - \alpha - \beta)$, ω , α , β , v con il metodo della QMV (linea continua) e della MV con errori t di Student (linea tratteggiata)

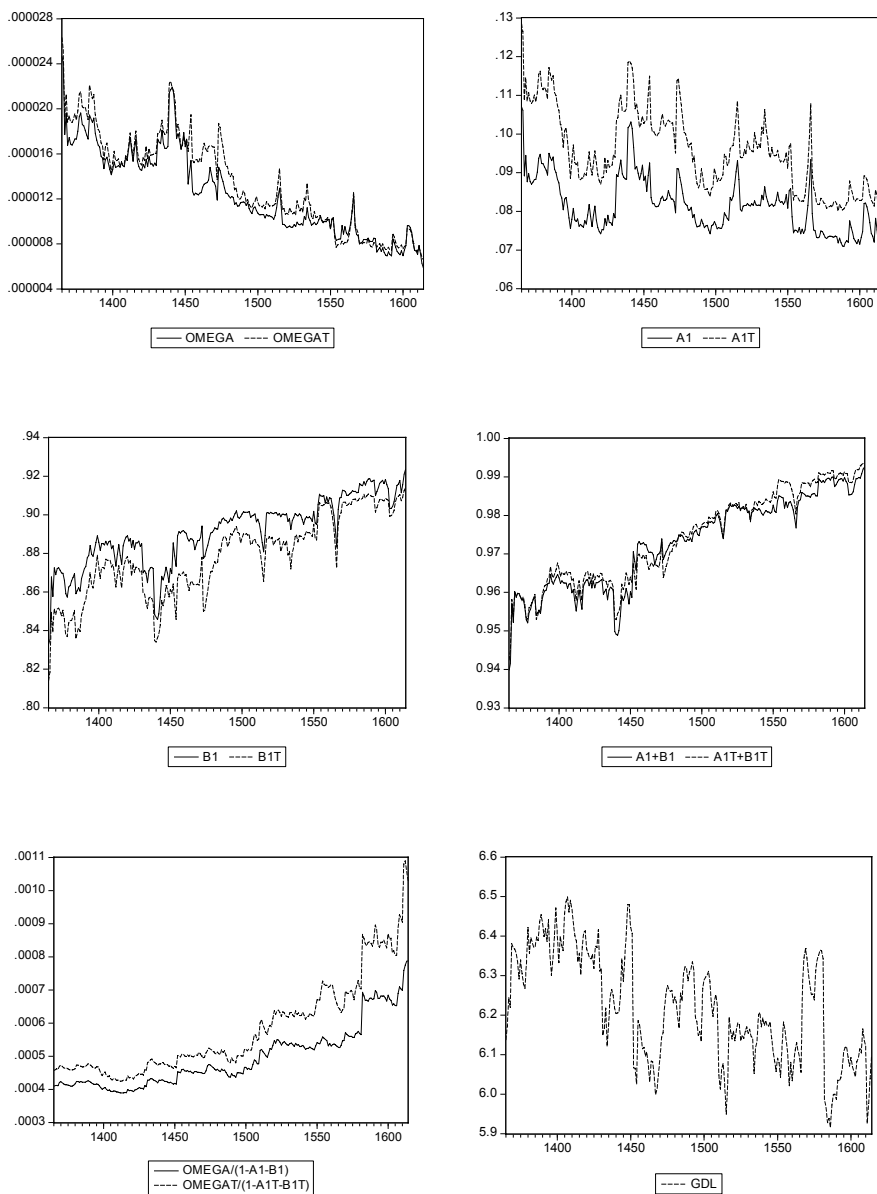


Tabella 22 – Test di corretta copertura marginale (CM), indipendenza (ID) e corretta copertura condizionata (CC) per gli attivi Fiat e Campari: *p-values*. Modelli GARCH(1, 1) stimati su finestra mobile

FIAT				
Metodo	$\hat{\alpha}$	CM	ID	CC
Quant. emp.	0.068	0.2146	0.4416	0.3209
t di Student	0.072	0.1331	0.5356	0.2480
CAMPARI				
	$\hat{\alpha}$	CM	ID	CC
Quant. emp.	0.064	0.3294	0.0811	0.1270
t di Student	0.072	0.1331	0.1588	0.1113

Inizializzazione

scalar maxh=250

smpl @all

genr serie=dlfiat_g

genr z05_e=-1.64

genr z05_t=-1.64

genr var05_e=0

genr var05_t=0

genr omega=0

genr omegat=0

genr a1=0

genr a1t=0

genr b1=0

genr b1t=0

genr gdl=0

genr volserie = @var(serie)

genr serie_fs=0

genr volseriet = @var(serie)

genr serie_fst=0

Stima mobile (rolling) dei modelli GARCH(1,1)

smpl 1 @last -maxh

```

for li=1 to maxh
equation modback.arch(1,1,h) serie
equation modbackt.arch(1,1,tdist) serie
scalar gdl=modbackt.@coefs(4)
modback.makeresids(s) zback
scalar ze=@quantile(zback,0.05)
scalar zt=@qtdist(0.05,gdl)*@sqr((gdl-2)/gdl)
smpl @last-maxh+li @last -maxh+li

```

'Salvo i parametri e le variabili che verranno successivamente utilizzate per la stima del VaR

```

modback.fit serie_fs serie_ses volserie
modbackt.fit serie_fst serie_sest volseriet

```

```

z05_e=ze
z05_t=zt

```

```

omega=modback.@coefs(1)
a1=modback.@coefs(2)
b1=modback.@coefs(3)

```

```

omegat=modbackt.@coefs(1)
a1t=modbackt.@coefs(2)
b1t=modbackt.@coefs(3)
gdlt=modbackt.@coefs(4)

```

'Adegua il sample per la stima successiva

```

smpl li+1 1614-maxh+li
next

```

'Calcolo il VaR per il periodo di previsione

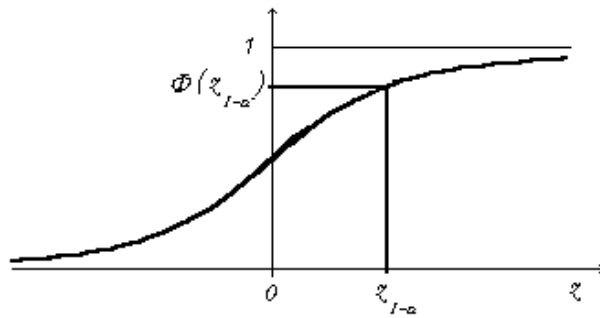
```

smpl 1614-maxh+1 1614
var05_e=z05_e*@sqr(volserie)
var05_t=z05_t*@sqr(volseriet)

```

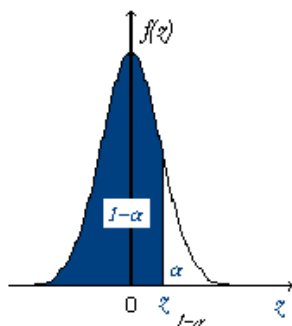
Tavole statistiche

Tavola 1 – Valori di $z_{(1-\alpha)}$ con $1-\alpha = \Phi(z_{(1-\alpha)}) = P\{Z \leq z_{(1-\alpha)}\}$



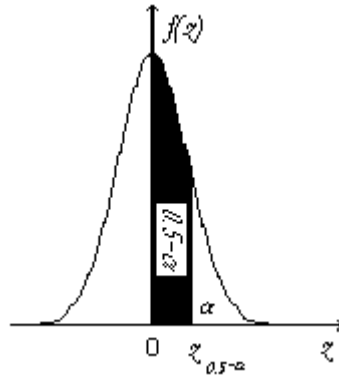
$\Phi(z_{(1-\alpha)})$	$z_{(1-\alpha)}$	$\Phi(z_{(1-\alpha)})$	$z_{(1-\alpha)}$	$\Phi(z_{(1-\alpha)})$	$z_{(1-\alpha)}$
.50	0.000	.91	1.341	.99500	2.576
.55	0.126	.92	1.405	.99900	3.090
.60	0.253	.93	1.476	.99950	3.291
.65	0.385	.94	1.555	.99990	3.719
.70	0.524	.95	1.645	.99995	3.891
.75	0.674	.96	1.751	.99999	4.265
.80	0.842	.97	1.881	.99999	4.417
.85	1.036	.98	2.054	.99999	4.753
.90	1.282	.99	2.326	.99999	5.199

Tavola 2 – Probabilità $P\{Z \leq z_{(1-\alpha)}\} = \int_{-\infty}^{z_{(1-\alpha)}} f(z) dz = 1 - \alpha$ della v.c. normale standardizzata Z per diversi valori di $z_{(1-\alpha)}$ (nella prima colonna è riportato l'intero e la prima cifra decimale, nella prima riga la seconda cifra decimale, all'interno della tabella i valori di $1-\alpha$)



$z_{(1-\alpha)}$	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767

Tavola 3 – Probabilità $P\{0 \leq Z \leq z_{(1-\alpha)}\} = \int_0^{z_{(1-\alpha)}} f(z) dz = 0.5 - \alpha$ nella v.c. normale standardizzata Z per differenti valori di $z_{(1-\alpha)}$



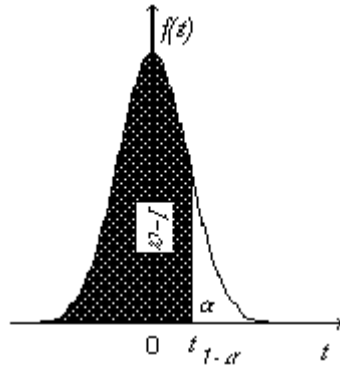
$z_{(0.5-\alpha)}$	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2703	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767

Continua Tavola 3

$z_{(0.5-\alpha)}$	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990

Tavola 4 – Valori $t_{(1-\alpha)}$ della v.c. T di Student per differenti livelli di probabilità e diversi gradi di

$$\text{libertà (prima colonna)} \quad P\left\{T_{(g)} \leq t_{(1-\alpha)}\right\} = \int_{-\infty}^{t_{(1-\alpha)}} f(t) dt = 1 - \alpha$$



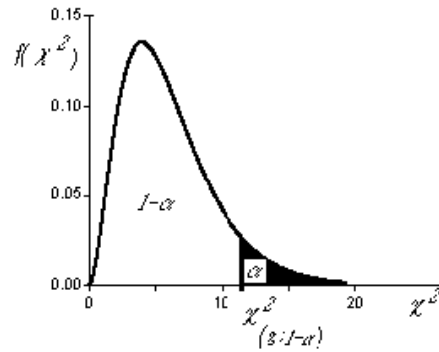
α	0.40	0.25	0.10	0.05	0.025	0.01	0.005	0.0005
$g \setminus 1-\alpha$	0.60	0.75	0.90	0.95	0.975	0.990	0.995	0.9995
1	0.325	1.000	3.087	6.314	12.706	31.821	63.657	636.619
2	0.289	0.816	1.886	2.920	4.303	6.965	9.925	31.598
3	0.277	0.765	1.638	2.353	3.182	4.541	5.841	12.924
4	0.271	0.741	1.533	2.132	2.776	3.747	4.604	8.610
5	0.267	0.727	1.476	2.015	2.571	3.365	4.032	6.869
6	0.265	0.718	1.440	1.943	2.447	3.143	3.707	5.959
7	0.263	0.711	1.415	1.895	2.365	2.998	3.499	5.408
8	0.262	0.706	1.397	1.860	2.306	2.896	3.355	5.041
9	0.261	0.703	1.383	1.833	2.262	2.821	3.250	4.781
10	0.260	0.700	1.372	1.812	2.228	2.764	3.169	4.587
11	0.260	0.697	1.363	1.796	2.201	2.718	3.106	4.437
12	0.259	0.695	1.356	1.782	2.179	2.681	3.055	4.318
13	0.259	0.694	1.350	1.771	2.160	2.650	3.012	4.221
14	0.258	0.692	1.345	1.761	2.145	2.624	2.977	4.140
15	0.258	0.691	1.341	1.753	2.131	2.602	2.947	4.073
16	0.258	0.690	1.337	1.746	2.120	2.583	2.921	4.015
17	0.257	0.689	1.333	1.740	2.110	2.567	2.898	3.965
18	0.257	0.688	1.330	1.734	2.101	2.552	2.878	3.922
19	0.247	0.688	1.328	1.729	2.093	2.539	2.861	3.883
20	0.257	0.687	1.325	1.725	2.086	2.528	2.845	3.850

Continua Tavola 4

α	0.40	0.25	0.10	0.05	0.025	0.010	0.005	0.0005
$g \setminus 1-\alpha$	0.60	0.75	0.90	0.95	0.975	0.905	0.995	0.9995
21	0.257	0.686	1.323	1.721	2.080	2.518	2.831	3.819
22	0.256	0.686	1.321	1.717	2.074	2.508	2.819	3.792
23	0.256	0.685	1.319	1.714	2.069	2.500	2.807	3.767
24	0.256	0.685	1.318	1.711	2.064	2.492	2.797	3.745
25	0.256	0.684	1.316	1.708	2.060	2.485	2.787	3.725
26	0.256	0.684	1.315	1.706	2.056	2.479	2.779	3.707
27	0.256	0.684	1.314	1.703	2.052	2.473	2.771	3.690
28	0.256	0.683	1.313	1.701	2.048	2.467	2.763	3.674
29	0.256	0.683	1.311	1.699	2.045	2.462	2.756	3.659
30	0.256	0.683	1.310	1.697	2.042	2.457	2.750	3.646
40	0.255	0.681	1.303	1.684	2.021	2.423	2.704	3.551
60	0.254	0.679	1.296	1.671	2.000	2.390	2.660	3.460
120	0.254	0.677	1.289	1.658	1.980	2.358	2.617	3.373
∞	0.253	0.674	1.282	1.645	1.960	2.326	2.576	3.291

Tavola 5 – Valori $\chi^2_{(g, \alpha)}$ della v.c. $\chi^2_{(g)}$ per differenti livelli di probabilità e gradi di libertà:

$$P\left\{\chi^2_{(g, \alpha)} \leq \chi^2 < \infty\right\} = \int_{\chi^2_{(g, \alpha)}}^{\infty} f(\chi^2) d\chi^2 = \alpha. \text{ Coda destra}$$



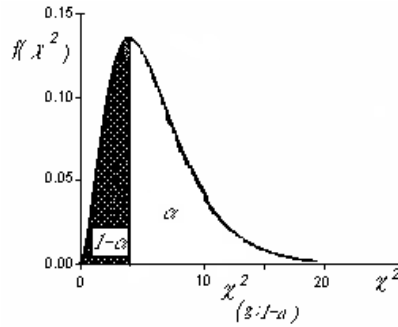
$1-\alpha$	0.750	0.900	0.950	0.975	0.990	0.995	0.999
$g \backslash \alpha$	0.250	0.100	0.050	0.025	0.010	0.005	0.001
1	1.32330	2.70554	3.84146	5.02389	6.63490	7.87944	10.828
2	2.77259	4.60517	5.99147	7.37776	9.21034	10.5966	13.816
3	4.10835	6.25139	7.81473	9.34840	11.3449	12.8381	16.266
4	5.38527	7.77944	9.48773	11.1433	13.2767	14.8602	18.467
5	6.62568	9.23635	11.0705	12.8325	15.0863	16.7496	20.515
6	7.84080	10.6446	12.5916	14.4494	16.8119	18.5476	22.458
7	9.03715	12.0170	14.0671	16.0128	18.4753	20.2777	24.322
8	10.2188	13.3616	15.5073	17.5346	20.0902	21.9550	26.125
9	11.3887	14.6837	16.9190	19.0228	21.6660	23.5893	27.877
10	12.5489	15.9871	18.3070	20.4831	23.2093	25.1882	29.588
11	13.7007	17.2750	19.6751	21.9200	24.7250	26.7569	31.264
12	14.8454	18.5494	21.0261	23.3367	26.2170	28.2995	32.909
13	15.9839	19.8119	22.3621	24.7356	27.6883	29.8194	34.528
14	17.1170	21.0642	23.6848	26.1190	29.1413	31.3193	36.123
15	18.2451	22.3072	24.9958	27.4884	30.5779	32.8013	37.697
16	19.3688	23.5418	26.2962	28.8454	31.9999	34.2672	39.252
17	20.4887	24.7690	27.5871	30.1910	33.4087	35.7185	40.790
18	21.6049	25.9894	28.8693	31.5264	34.8053	37.1564	42.312
19	22.7178	27.2036	30.1435	32.8523	36.1908	38.5822	43.820

Continua Tavola 5

$1-\alpha$	0.750	0.900	0.950	0.975	0.990	0.995	0.999
$g \setminus \alpha$	0.250	0.100	0.050	0.025	0.010	0.005	0.001
20	23.8277	28.4120	31.4104	34.1696	37.5662	39.9908	45.315
21	24.9348	29.6151	32.6705	35.4789	38.9321	41.4010	46.797
22	26.0393	30.8133	33.9244	36.7807	40.2894	42.7956	48.268
23	27.1413	32.0069	35.1725	38.0757	41.6384	44.1813	49.728
24	28.2412	33.1963	36.4151	39.3641	42.9798	45.5585	51.179
25	29.3389	34.3816	37.6525	40.6465	44.3141	46.9278	52.600
26	30.4345	35.5631	38.8852	41.9232	45.6417	48.2899	54.052
27	31.5284	36.7412	40.1133	43.1944	46.9630	49.6449	55.476
28	32.6205	37.9159	41.3372	44.4607	48.2782	50.9933	56.892
29	33.7109	39.0875	42.5569	45.7222	49.5879	52.3356	58.302
30	34.7998	40.2560	43.7729	46.9792	50.8922	53.6720	59.703
40	45.6160	51.8050	55.7585	59.3417	63.6907	66.7659	73.402
50	56.3336	63.1671	67.5048	71.4202	76.1539	79.4900	86.661
60	66.9814	74.3970	79.0819	83.2976	88.3794	91.9517	99.607
70	77.5766	85.5271	90.5312	95.0231	100.425	104.215	112.317
80	88.1303	96.5782	101.879	106.629	112.329	116.321	124.839
90	98.6499	107.565	113.145	118.136	124.116	128.299	137.208
100	109.141	118.498	124.342	129.561	135.807	140.169	149.449

Tavola 6 – Valori $\chi^2_{(g;1-\alpha)}$ della v.c. $\chi^2_{(g)}$ per differenti livelli di probabilità e gradi di libertà:

$$P\left\{0 \leq \chi^2_{(g)} \leq \chi^2_{(g;1-\alpha)}\right\} = 1 - \alpha. \text{ Coda sinistra}$$



α	0.995	0.990	0.980	0.950	0.900	0.750	0.500
$g \setminus 1-\alpha$	0.005	0.010	0.020	0.050	0.100	0.250	0.500
1	3927·10 ⁻⁸	15709·10 ⁻⁸	98207·10 ⁻⁸	39321·10 ⁻⁷	0.01579	0.10153	0.45493
2	0.01002	0.02010	0.05063	0.10259	0.21072	0.57536	1.38629
3	0.07172	0.11483	0.21579	0.35185	0.58437	1.21253	2.36597
4	0.20699	0.29711	0.48442	0.71072	1.06362	1.92255	3.35670
5	0.41174	0.55430	0.83121	1.14547	1.61031	2.67460	4.35146
6	0.67572	0.87208	1.23734	1.63539	2.20413	3.45460	5.34812
7	0.98926	1.23904	1.68987	2.16735	2.83311	4.25485	6.34581
8	1.34441	1.64648	2.17973	2.73264	3.48954	5.07064	7.34412
9	1.73492	2.08791	2.70039	3.32511	4.16816	5.89883	8.34283
10	2.15585	2.55821	3.24697	3.94030	4.86518	6.73720	9.34182
11	2.60321	3.05347	3.81575	4.57481	5.57779	7.58412	10.3410
12	3.07382	3.57056	4.40379	5.22603	6.30380	8.43842	11.3403
13	3.56503	4.10691	5.00874	5.89186	7.04150	9.29906	12.3398
14	4.07468	4.66043	5.62872	6.57063	7.78953	10.1653	13.3393
15	4.60094	5.22935	6.26214	7.26094	8.54675	11.0365	14.3389
16	5.14224	5.81221	6.90766	7.96164	9.31223	11.9122	15.3385
17	5.69724	6.40776	7.56418	8.67176	10.0852	12.2919	16.3381
18	6.26481	7.01491	8.23075	9.39046	10.8649	13.6753	17.3379
19	6.84398	7.63273	8.90655	10.1170	11.6509	14.5620	18.3376
20	7.43386	8.26040	9.59083	10.8508	12.4426	15.4518	19.3374
21	8.03366	8.89720	10.2829	11.5913	13.2396	16.3444	20.3372

Continua Tavola 6

α	0.995	0.990	0.980	0.950	0.900	0.750	0.500
$g \setminus 1-\alpha$	0.005	0.010	0.025	0.050	0.100	0.250	0.500
22	8.64272	9.54249	10.9823	12.3380	14.0415	17.2396	21.3370
23	9.26042	10.1957	11.6885	13.0905	14.8479	18.1373	22.3369
24	9.88623	10.8564	12.4011	13.8484	15.6587	19.0372	23.3367
25	10.5197	11.5240	13.1197	14.6114	16.4734	19.9393	24.3366
26	11.1603	12.1981	13.8439	15.3791	17.2919	20.8434	25.3364
27	11.8076	12.8786	14.5733	16.1513	18.1138	21.7494	26.3363
28	12.4613	13.5648	15.3079	16.9279	18.9392	22.6572	27.3363
29	13.1211	14.2565	16.0471	17.7083	19.7677	23.5666	28.3362
30	13.7867	14.9535	16.7908	18.4926	20.5992	24.4776	29.3360
40	20.7065	22.1643	24.4331	26.5093	29.0505	33.6603	39.3354
50	27.9907	29.7067	32.3574	34.7642	37.6886	42.9421	49.3349
60	35.5346	37.4848	40.4817	43.1879	46.4589	52.2938	59.3347
70	43.2752	45.4418	48.7576	51.7393	55.3290	61.6983	69.3344
80	51.1720	53.5400	57.1532	60.3915	64.2778	71.1445	79.3343
90	59.1963	61.7541	65.6466	69.1260	73.2912	80.6247	89.3342
100	67.3276	70.0648	74.2219	77.9295	82.3581	90.1332	99.3341

Bibliografia

- AKAIKE H, 1973, Information theory and an extension of the maximum likelihood principle, in Petrov, Csaki, *International Symposium on Information Theory*, Akademia Kiado, Budapest
- AKAIKE H, 1974, A new look at the statistical model identification, *IEEE Transactions on Automatic Control*, AC-19, 716-723
- AMENDOLA A, NIGLIO M, VITALE CD, 2009, Statistical properties of SETARMA models, *Communication in Statistics: theory and Methods*, 38, 2479-2497
- AMENDOLA A, NIGLIO M, VITALE CD, 2009, Local unity roots and global stationarity of TARMA models, *Dipartimento di Scienze Economiche e Statistiche, Salerno*
- ARTZNER P, DELBAEN F, EBER J-M, HEALT D, 1999, Coherent measures of risk, *Mathematical Finances*, 9, 203-228
- AZZALINI A, 1996, *Statistical Inference – Based on the Likelihood*, Chapman & Hall.
- BAI X, RUSSEL JR, TIAO GC, 2001, Kurtosis of GARCH and stochastic volatility models, *Working paper, Graduate School of Business*, University of Chicago
- BASLE COMMITTEE IN BANKING SUPERVISION, 1995, An internal model-based approach to market risk capital requirements, <http://www.bis.org>.
- BASLE COMMITTEE IN BANKING SUPERVISION, 1996, Overview of the amendment to the capital accord to incorporate market risk, <http://www.bis.org>
- BASLE COMMITTEE IN BANKING SUPERVISION, 2003, Overview of the new Basel capital accord. Consultative document, Basel: *Bank for International Settlements*
- BERKOWITZ J, 2001, Testing density forecasts with applications to risk management, *Journal of Business & Economic Statistics*, 19, 465-474
- BOLLERSLEV T, 1986, Generalized autoregressive conditional heteroskedasticity, *Jornal of Econometrics*, 31, 307-327
- BOLLERSLEV T, WOOLDRIDGE J, (1992) Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time-varying Covariances, *Econometric Reviews*, 11, 143–172

- BOX JEP, JENKINS GM, 1976, *Time Series Analysis: Forecasting and Control*, S. Francisco, Holden Day
- BROCKWELL PJ, DAVIS RA, 1991, *Time Series: Theory and Methods*, 2° ediz., Springer-Verlag, New York
- BROCKWELL PJ, DAVIS RA, 1996, *Introduction to Time Series and Forecasting*, Springer-Verlag, New York
- CAMPBELL JY, LO AW, MACKINLAY AC, 1997, *The Econometrics of Financial Markets*, Princeton University Press: New Jersey
- CHAN NH, WEI CZ, 1988, Limiting distribution of least square estimates of unstable autoregressive processes, *Annals of Statistics*, 16, 347-401
- CHAPPELL D, PADMORE J, MISTRY P, ELLIS C, 1996, A threshold model for the French/ Deutschmark exchange rate, *Journal of Forecasting*, 15, 155-164
- CHRISTOFFERSEN PF, (1998) Evaluating Interval Forecasts, *International Economic Review*, Vol. 39, No. 4, Symposium on Forecasting and Empirical Methods in Macroeconomics and Finance. (Nov., 1998), pp. 841-862
- COX JC, RUBISTEIN M, 1985, *Options Markets*, Prentice-Hall: Englewood Cliffs, New Jersey
- DICKEY AD, FULLER WA, 1979, Distribution of the estimates for autoregressive time series with a unit root, *JASA*, 74, 427-431
- DING ZC, GRANGER CWJ, ENGLE RF, 1993, A long memory property of stock market return and a new model, *Journal of Empirical Finance*, 1, 83-106
- DONSKER M, 1951, An invariance principle for certain probability limit theorems, *Mem. American Mathematical Society*, N, 6
- DOWD K, BLAKE D, 2006, After VaR: the theory, estimation, and insurance applications of quantile – based risk measures, *The Journal of Risk Insurance*, 73, 193-229
- ENGLE RF, 1982, Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflations, *Econometrica*, 50, 987-1007
- GALLO GM, PACINI B, 2002, *Metodi Quantitativi per i Mercati Finanziari*, Carocci, Roma
- GLOSTEN LR, JAGANNATHAN R, RUNKLE D, 1993, On the relation between the expected value and the volatility of the nominal excess return on stocks, *Journal of Finance*, 48, 1779-1801
- HAMILTON JD, 1994, *Time Series Analysis*, Princeton University Press, New York
- HEYDE CC, (1997) *Quasi-Likelihood and Its Applications*, Springer Series in Statistics, Springer
- HULL JC, 1997, *Options, Futures, and other Derivatives*, 3° ediz. Prentice-Hall, New Jersey
- IACUS SM, 2008, *Simulation and Inference for Stochastic Differential Equation: with R examples*, Springer Series in Statistics, Springer

- INUI K, KIJIMA M, KITANO A, 2005, VaR is subject to a significant positive bias, *Statistics & Probability Letters*, vol. 72, pp. 299-311
- KAPETANIOS G, SHIN Y, 2006, Unit root test in three-regimes SETAR models, *Econometrics Journal*, 9, 252-278
- KRÄGER H, KUGLER P, 1993, Non-linearities in foreign exchange markets: a different perspective, *Journal of International Money and Finance*, 93, 478-493
- KUPIEC P, 1995. Techniques for verifying the accuracy of risk measurement models. *Journal of Derivatives* 3, 73-84.
- LAFRATTA G, 2004, *Metodi Statistici per l'Analisi dei Mercati Finanziari*, Franco Angeli, Milano
- LI WK, LAM K, 1995, Modelling asymmetry in stock returns by a threshold autoregressive conditional heteroscedastic model, *The Statistician*, 44, 333-341
- LING S, 1999, On the probabilistic properties of a double threshold ARMA conditional heteroskedasticity model, *Journal of Applied Probability*, 36, 688-705
- LIU J, SUSKOU E, 1992, On strict stationarity and ergodicity of a non-linear ARMA model, *Journal of Applied Probability*, 29, 363-373
- MCCULLAGH P, NELDER JA, (1989) *Generalized Linear Models* (2nd Edition), Chapman & Hall
- NELSON DB, 1991, Conditional heteroskedasticity in asset returns: a new approach, *Econometrica*, 59, 347-370
- NIGLIO M, VITALE CD, 2009, Invertibility of threshold moving average models, *Dipartimento di Scienze Economiche e Statistiche, Salerno*
- OBSTFELD M, TAYLOR AM, 1997, Nonlinear aspects of good-market arbitrage and adjustment: Heckscher's commodity points revisited, *Journal of the Japanese and International Economie*, 11, 441-479
- O'CONNEL PGJ, 1998, Market frictions and real exchange rates, *Journal of International Money and Finance*, 17, 71-95
- PEEL DA, SPEIGHT AEH, 1994, Testing for non-linear dependence in inter-war exchange rates, *Weltwirtschaftliches Archiv*, 130, 391-417
- PHILLIPS PCB, 1987, Time series regression with a unit root, *Econometrica*, 55, 277-301
- PHILLIPS PCB, PERRON P, 1988, Testing for a unit root in time series regression, *Biometrika*, 75, 335-346
- PICCOLO D, 2001, *Statistica*, Ed. Il Mulino
- PICCOLO D, VITALE CD, 1984, *Metodi Statistici per l'Analisi Economica*, Il Mulino, Bologna
- RISKMETRICS, 1996, Technical Document, quarta ediz., RiskMetrics Group, Inc., New York

- SNEDECOR GW, COCHRAN WG, 1980, *Statistical Methods*, 7th edition, Iowa State University Press: Ames, Iowa
- TONG H, 1978, On a threshold model, in *Pattern Recognition and Signal Processing*, ed. Chen, Sijboff & Noordhoff, Amsterdam
- TONG H, LIM KS, 1980, Threshold autoregression, limit cycle and cyclical data, *Journal of the Royal Statistical Society*, B 42, 245-292
- TONG H, 1990, *Non-Linear Time Series: a Dynamical System Approach*, Oxford University Press, Oxford
- TSAY RS, 2003, *Analysis of Financial Time Series*, John Wiley,
- WEDDERBURN RWM, (1974) Quasi-likelihood Functions, Generalized Linear Models, and the Gauss - Newton Method, *Biometrika*, **61** (3), 439-447
- ZAKOÏAN JM, 1994, Threshold heteroscedastic models, *Journal of Economic Dynamics and Control*, 18, 931-955

Indice analitico

- ADF test 178, 179, 180, 181
- Adj Close 5, 6, 11, 12, 13
- AIC 206, 207, 208, 212, 215, 216, 218, 226, 260, 262, 264, 272, 279, 280, 282, 287, 313, 334, 358
- Akaike 206, 379
- Amendola 194, 197
- AR 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 174, 175, 176, 180, 188, 189, 191, 195, 196, 198, 207, 208, 209, 211, 212, 215, 216, 217, 220, 242, 254, 259, 307
 - GARCH 307
- AR(∞) 189
- Arbitraggio 157, 182
- Autoregressive 157, 162, 165, 181, 191, 194, 211, 212, 283
- AR-GARCH 307
- ARCH 233, 237, 238, 240, 241, 242, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 257, 266, 268, 278, 283, 354, test 248
- ARMA
 - 156, 157, 161, 162, 163, 171, 188, 189, 190, 191, 192, 193, 195, 199, 202, 206, 207, 208, 216, 217, 218, 219, 226, 233, 241, 244, 245, 246, 247, 251, 253, 254, 255, 258, 259, 264, 266, 268, 269, 275, 283, 286, 287, 301, 308, 311, 315, 330, 341, 342, 343, 344, 349, 358
 - ARCH, 241, 244, 246, 251, 253, 254
- GARCH, 258, 259, 264, 275, 283, 286, 287, 301, 311, 312, 315, 330, 341, 342, 343, 344, 349, 358
- GARCH-M 272, 308
- TGARCH 275
- Artzner 291, 339, 340
- Asimmetria 13, 15, 16, 25, 26, 29, 34, 36, 40, 46, 53, 59, 72, 73, 77, 91, 92, 93, 115, 154, 184, 197, 198, 200, 225, 235, 236, 250, 317
 - locale 197
- Asintoticamente
 - efficienti 284
 - equivalente 132
 - meno efficiente 116
 - non distorto 91, 139, 157, 167, 170, 198
 - normale 116, 250, 292
- Attivi finanziari 1, 7, 13, 27, 34, 35, 38, 48, 53, 65, 81, 136, 143, 145, 150, 151, 153, 157, 158, 159, 160, 161, 171, 193, 202, 214, 233, 237, 244, 248, 247, 254, 261, 289, 315, 317, 330, 338, 340
- Autocorrelazione 47, 48, 49, 50, 53, 55, 56, 57, 58, 73, 74, 76, 139, 140, 141, 142, 148, 149, 150, 166, 167, 169, 170, 172, 173, 182, 185, 186, 188, 189, 191, 192, 197, 201, 202, 203, 204, 205, 212, 213, 215, 239, 244, 245, 246, 248, 251, 260, 261, 267, 311, 313
 - globale 50
 - locale 49, 50
 - quadrati 74, 201, 205, 245, 251, 261
 - regimi 49, 50, 201, 204, 205
 - residui 76, 168, 173, 186, 188, 194, 202, 205, 212, 213, 226, 246, 260, 261, 313
 - volatilità 55, 56
- Autocovarianza 139, 144, 166, 197, 203, 243
 - locale 197

- Auto simile 13
- Azioni 1, 4, 5, 6, 11, 17, 68, 300
- Azzalini 117
- Bai 257
- Banca Intesa S. Paolo 16, 17, 18, 19, 22, 23, 24, 25, 26, 27, 34, 35, 54, 55, 155, 161, 257, 258, 260, 261, 263, 264, 277, 279, 282, 286, 287, 298, 299, 300, 320, 321, 322, 323, 324, 325
- Banda
 confidenza, 141, 149, 150, 169, 173, 185, 186, 212, 215, 239, 244, 246, 267
 previsione 223
- Basle Committee 291
- Bera 154, 155, 156, 173, 238, 286, 287, 311, 358, 359, 360
- Backtesting 348, 349, 358
- Berkowitz 292
- Bernoulli 85, 86, 103, 123, 133, 350, 354
- Blake 291
- Bollerslev 254, 283, 286
- Box 140, 162
- Box-plot 17, 19, 20, 26, 27, 29, 46, 69, 71
- Brockwell 194
- BW 283, 284, 286, 287
- CAC 6, 69, 70, 71, 72, 73, 74, 78
- Cadenza 1, 5, 6, 9, 10, 13, 27, 28, 47, 48, 49, 56, 136, 145, 152, 259, 260, 363
- Caduta attesa 340
- Campani 9, 10, 11, 12, 13, 14, 16, 17, 18, 19, 20, 34, 35, 155, 161, 211, 212, 213, 215, 216, 218, 225, 226, 230, 257, 258, 260, 261, 262, 263, 264, 267, 268, 271, 272, 273, 277, 279, 282, 286, 287, 298, 299, 300, 305, 306, 307, 309, 310, 320, 321, 322, 323, 324, 334, 337, 338, 346, 347, 348, 358, 362, 365
- Cardine
 inferiore 16
 superiore 16
- Cauchy 95
- Central limit theorem, 112
- CGARCH, 263, 264
- Chan 176
- Chappel 194
- Christoffersen 353, 357
- Close 11, 12, 13
- CLT 112, 113, 114
- Cluster, 239
- Cochran 154
- Componente
 AR 188
 MA 202, 225
 permanente 263
 transitoria 263
- Convergenza
 a.s. 108
 debole 108
 distribuzione 111, 132
 legge 111
 media quadratica 108
 probabilità 108
- Copertura marginale 351, 353, 354, 357, 360, 361, 362, 363, 365
- Correlazione 47, 52, 101, 102, 107, 119
- Correlogramma 139, 244
- Covarianza 52, 101, 130, 142, 197, 285
- Cross covarianza 197
 locale 197
- Curtosi 15, 16, 25, 29, 34, 35, 36, 40, 41, 46, 53, 59, 60, 73, 77, 91, 93, 95, 115, 151, 152, 154, 159, 169, 173, 184, 198, 200, 225, 231, 233, 234, 236, 237, 239, 246, 257, 267, 282, 311, 317
 globale 198
 locale 198, 231
- Data 11
- DAX 6, 69, 70, 71, 72, 73, 74
- Densità 46, 87, 88, 89, 92, 93, 94, 95, 96, 98, 99, 105, 106, 120, 121, 126,

- 137, 138, 140, 148, 172, 173, 199,
200, 204, 227, 234, 249, 250, 251,
271, 290, 291, 292, 293, 308, 343
- Derivati 7, 8
- Determinante 101
- Deviazione standard 15, 57, 58, 65, 344
- mobile 57, 58
- DF test 177, 178
- Dickey 176, 178
- Differenze prime 150, 165, 180
- Ding 279
- Distribuzione 13, 18, 61, 78, 79, 91,
100, 103, 110, 112, 113, 114, 118,
122, 129, 132, 134, 153, 177, 178,
257, 267, 271, 286, 302, 331, 340,
341, 342, 351, 353, 357
- asintotica 111, 130, 131, 132, 178,
353
 - campionaria 130, 284
 - condizionata 350, 354
 - empirica 293
 - errori accidentali 30, 72, 284, 286,
342, 343, 345
 - normale, 15, 18
 - nulla 134, 136
 - probabilità 103, 123, 133, 146, 153,
195, 290, 294, 296, 354
 - ripartizione 291
- Donsker 146
- Dow Jones 6
- Dowd 291
- Drift 149, 150, 157, 158, 165, 172, 198,
215
- Edgeworth 115
- EGARCH 269, 270, 271, 283, 338
- EGARCH-M 270, 272
- Engle 233
- Equity market 4
- Ergodico 138, 139, 166
- Errore 22, 25, 26, 30, 42, 43, 58, 61, 65,
72, 118, 153, 156, 162, 173, 177, 195,
215, 220, 260, 286, 330, 331, 334,
337, 347, 348, 358, 364
- centrato 25, 26
 - approssimazione 160
 - non centrato 25, 27
 - normale 125, 128, 286, 287
 - previsione 225
 - primo tipo 121
 - standardizzato 65, 75, 76, 77, 284,
287
 - stimato 71
- Evento 3, 4
- raro 289, 290
- Errore standard 177, 184, 207, 215, 227
- Errore di previsione 225
- e.s. 340, 341, 342, 345, 346, 347, 348
- EViews 1, 81, 227, 251, 252, 278, 283,
287, 317, 326, 330, 334, 346, 360,
363
- Ex ante 208, 211
- post 208
- Excell 1
- Expected shortfall 338, 340, 341, 346
- Fiat 5, 16, 17, 18, 19, 28, 29, 30, 31, 32,
33, 34, 35, 42, 43, 44, 45, 46, 47, 48,
49, 50, 54, 55, 56, 57, 58, 59, 60, 63,
64, 66, 67, 68, 77, 155, 156, 159,
161, 211, 212, 213, 215, 216, 218,
225, 226, 227, 229, 230, 245, 246,
251, 252, 253, 254, 257, 258, 260,
261, 263, 264, 267, 268, 271, 272,
273, 277, 278, 279, 280, 282, 286,
287, 298, 299, 300, 303, 304, 305,
309, 310, 312, 313, 314, 315, 320,
321, 322, 323, 324, 334, 337, 338,
346, 347, 358, 359, 362, 363, 364,
365
- Fisher 97, 98, 114, 133, 248
- Fitted 208
- Forecasted 208
- Foreign Exchange Market 3
- Forma quadratica 130
- Fuller, 176, 178
- Funzione di

- autocorrelazione 47, 139, 140, 166, 167, 189, 197
 autocovarianza 139, 197, 203
 densità, 87, 88, 89, 92, 94, 95, 96, 99, 105, 106, 120, 121, 126, 137, 140, 204, 234, 250, 251, 271, 290, 292, 293, 308
 densità locale 227
 densità condizionata 291
 distribuzione 123, 133
 log-verosimiglianza 117, 118, 123, 124, 127, 128, 131, 132, 133
 quasi verosimiglianza 117, 118, 119, 284, 285
 ripartizione 110, 111, 115, 271, 292, 293
 verosimiglianza 90, 115, 117, 120, 123, 126, 129, 130, 131, 132, 246, 284, 286, 352, 353, 354, 356
 Futures 7, 8
 FX Market 3
 GARCH 254, 255, 257, 259, 263, 264, 266, 268, 270, 273, 275, 276, 277, 278, 279, 287, 310, 342, 358, 359, 360, 362, 363, 364, 365
 GARCH-M 261, 262, 266, 268
 Gauss 87, 330
 Glosten 273
 Gradi di libertà 35, 36, 37, 38, 41, 92, 93, 94, 95, 97, 98, 109, 112, 113, 131, 141, 142, 158, 173, 176, 248, 249, 258, 311, 312, 320, 331, 344, 345, 359, 373, 375, 377
 Heyde 119, 120
 Hermite 115
 High 11, 12, 13
 Iacus 151
 IGARCH 259, 260, 266, 283, 296, 297, 299, 304
 IGARCH-M 268
 iid 116, 140, 141, 142, 143, 144, 145, 147, 149, 150, 152, 159, 190, 192, 195, 199, 203, 234, 236, 237, 237, 239, 242, 246, 247, 249, 250, 255, 256, 259, 262, 269, 271, 283, 284, 292, 293, 294, 296, 298, 301, 304, 308, 310, 311, 317, 334, 351
 Incrementi 146
 incorrelati, 143
 normali indipendenti 145
 Indice 14, 21, 47, 58, 60, 61, 68, 81, 101, 187, 253, 290, 291
 base fissa 18, 19, 68
 base mobile 28
 Informazione di Fisher 114
 Inui 293
 Integrazione di Ito 146, 151
 Interest rate swap 4
 Intervallo 14, 28, 32, 85, 89
 confidenza 139, 159, 171, 296, 308, 309, 310, 313
 previsione 217, 220
 Intesa S. Paolo 16, 17, 18, 19, 22, 23, 24, 25, 26, 27, 34, 35, 54, 55, 155, 161, 257, 258, 260, 261, 263, 264, 277, 279, 282, 286, 287, 298, 299, 300, 320, 321, 322, 323, 324, 325
 Invertibile 162, 163, 164, 182, 188, 189, 191
 Invertibilità 193, 301
 Ipotesi 85, 93, 97, 116, 118, 123, 133, 153, 156, 177, 182, 301, 351, 360
 bidirezionale 122, 129
 composte 121
 semplice 120, 121
 IRS market 4
 Istogramma 14, 37, 38, 39, 40, 42, 43, 44, 145, 169, 238
 Jarque 154, 155, 156, 173, 238, 286, 287, 311, 358, 359, 360

- Kapetanios 194
 Krager 194
 Kugler 194
 Kupiec 351
- Lagrange 120, 129
 Lam 194
 Laplace 87
 Laspeyres 21
 Legami lineari locali 49
 Legge
 - distribuzione 100, 110, 111
 - grandi numeri 109, 110
 - forte dei grandi numeri 109, 110
- Leptocurtosi 91, 92, 93, 282
 Leverage 268, 269, 310
 Li 194
 LIBOR 4
 Ling 194
 Liu 194
 Livelli locali 282
 Ljung 140
 LM 129, 131, 132, 133, 134, 135
 Log normale 30, 99, 100,
 - rendimenti 30, 31
 - verosimiglianza 117, 118, 123, 124, 128, 129, 131, 132, 133, 247
- Low 11, 12, 13
 LS 171, 186, 192, 193, 201, 205
- MA 163, 181, 182, 184, 185, 186, 187, 188, 195, 1196, 198, 202, 203, 204, 207, 208, 214, 215, 216, 218, 224, 225, 227, 228, 246, 252, 255, 283, 317
- ARCH 252
 - CGARCH-M 264
 - EGARCH 272, 283, 313, 337, 338
 - EGARCH-M 272
 - GARCH 257, 261, 273, 283, 302, 303, 304, 305, 306, 307, 308, 309, 310, 313, 334, 335, 337, 347, 348
 - GARCH-M 261, 267, 307
 - IGARCH 260, 261, 283
 - IGARCH-M 262
 - PGARCH-M 279
 - TGARCH 283, 313, 316
 - TGARCH-M 277
- MA(∞) 164, 165, 189
 Marginali 92, 100, 101, 103, 106, 107
 Massima potenza 121
 Massima verosimiglianza 90, 91, 97, 114, 115, 123, 124, 127, 128, 130, 194, 198, 246, 247, 249, 266, 278, 279, 311, 312, 329, 333, 334, 344, 352, 353, 355, 356, 358
- Massimo 15, 16
 - intero 62, 315
- Matlab 330
- Matrice 101, 102, 106, 130, 240, 241, 285, 286, 331, 332, 35, 336, 338
 - correlazione 102
 - identica 102
 - inversa 101
 - informazione 130
 - varianze covarianze 101, 130, 285
- McCullagh 115
- Media 6, 10, 11, 13, 14, 15, 16, 17, 18, 21, 22, 36, 45, 46, 50, 51, 53, 58, 65, 66, 73, 79, 80, 83, 86, 87, 88, 91, 94, 97, 100, 130, 132,
 - 144, 169, 173, 196, 200, 210, 215, 234, 242, 325, 326, 329, 344, 346
 - campionaria 91, 108, 285
 - condizionata 102, 104, 208, 234, 293, 296, 342
 - errore di previsione 211
 - mobile centrata 22, 23, 24, 44
 - mobile non centrata 22, 23, 24, 25, 42, 43, 44, 56, 71, 83
 - progressiva di processi iid 142
 - regimi 196, 197, 198
 - stimata 86, 91, 97, 107, 116, 118, 133, 147, 165
- Mediana 15, 16, 200, 204
- Mercato 1, 2, 3, 4, 6, 11, 36, 45, 51, 52, 68, 81, 119, 153, 157, 233, 261, 289, 290, 291, 295, 300, 338, 347, 348

- borsistico 21, 28, 29, 230, 253, 289
 dei cambi 3
 della moneta 4
 finanziario 76, 81
 orso 36, 268, 325
 toro 36, 268, 325
- Metodo**
 stima 85, 186
 BW 287
 massima verosimiglianza 90, 115, 116, 119, 246, 249, 266, 278, 279, 311, 312, 334, 344, 358, 360, 364
 minimi quadrati 170, 171, 175, 183, 186, 191, 198, 207, 225, 246, 248, momenti 119, 124, 311
 QMV 116, 119, 283, 342, 358, 359, 364
- MIB** 6
MIB-30 6
MIBTEL 6, 21, 28, 29, 30, 33, 34, 35, 52, 68, 69, 70, 71, 73, 74, 75, 76, 78, 79, 80, 155, 156, 159, 161, 180, 181, 211, 212, 213, 215, 216, 218, 219, 225, 226, 230, 245, 246, 251, 252, 253, 254, 298, 299, 300, 320, 321, 322, 323, 324, 325, 326
- Minimi quadrati** 118, 119, 170, 171, 175, 177, 183, 186, 192, 198, 206, 207, 225, 246, 248, 329
- Minimo** 15, 16, 19, 67, 81, 90, 187, 207
- Mistura** 36
- ML** 90, 122, 124, 125, 127, 128, 129, 134, 135, 247
- Modello**
 ottimale 206, 305, 306, 309, 313, 316, 349
 a soglia 36, 39, 197, 225, 317
- Momento** 92, 95, 100, 108, 111, 113, 114, 138, 139, 140, 141, 194, 195, 196, 247, 255
 globale 196
 locale 196
- Money market** 4
- Moto Browniano** 145
- MV** 97, 115, 116, 134, 284, 286, 358, 359, 360, 364
- NASDAQ** 6, 52, 68, 73, 74, 75, 76, 79, 80, 155, 161, 180, 181, 211, 212, 213, 215, 216, 218, 219, 225, 226, 230, 325, 326
 finanziario 68, 69, 70, 71, 73, 74, 75, 76, 79, 80, 81, 180, 181, 325, 326
- Nelder** 115
- Nelson** 269
- Neyman** 120, 121
- Niglio** 195
- Non distorto** 87, 91, 211
- Non invertibile** 185, 186, 191
- Non stazionario** 142, 143, 149, 171, 177, 179, 195
- Normale** 15, 18, 30, 32, 34, 35, 37, 41, 43, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 99, 100, 106, 115, 118, 125, 132, 133, 134, 138, 141, 144, 153, 154, 156, 158, 159, 160, 204, 211, 239, 247, 249, 250, 259, 267, 271, 284, 291, 292, 293, 296, 297, 304, 308, 310, 311, 313, 314, 340, 343, 345, 368, 371
 doppia 104, 105, 106, 107
 multivariata 130, 284, 286
 standardizzata 18, 89, 94, 95, 96, 115, 153, 158, 160, 211, 239, 271, 284, 293, 297, 308, 311, 343, 345, 368, 371
- Normalità** 25, 33, 34, 36, 46, 73, 118, 155, 156, 173, 184, 211, 256, 258, 271, 286, 287, 305, 310, 311, 312, 313, 314, 358
- Obbligazioni** 1, 6, 7
- Obstfeld** 194
- O'Connell** 194
- OLS** 118, 119
- Open** 11, 12, 13
- Option Americana** 8
- Option Europea** 8
- Options** 7, 8
- Paasche** 21

- Parametri 37, 38, 41, 87, 90, 97, 99,
 105, 116, 118, 119, 120, 121, 122,
 123, 129, 130, 131, 132, 140, 153,
 156, 157, 162, 165, 166, 167, 170,
 171, 172, 174, 175, 176, 182, 183,
 184, 186, 187, 190, 191, 192, 193,
 196, 198, 199, 200, 201, 203, 205,
 206, 207, 208, 210, 212, 215, 216,
 217, 218, 220, 223, 224, 225, 226,
 229, 234, 235, 241, 246, 247, 248,
 249, 250, 251, 252, 253, 254, 257,
 264, 266, 270, 275, 277, 278, 279,
 281, 282, 283, 284, 286, 298, 299,
 303, 304, 308, 311, 312, 313, 317,
 319, 320, 329, 333, 334, 341, 342,
 344, 349, 358, 359, 363, 366
 Pearson 120, 121
 Peel 194
 Percentile 18, 61, 62, 65, 66, 68, 75, 78,
 140, 154, 158, 160, 211, 248, 249,
 290, 291, 292, 293, 294, 297, 302,
 304, 308, 310, 311, 312, 314, 315,
 317, 342
 locale 66
 Perdita 3, 5, 7, 116, 301, 340
 informazioni 59
 massima 60, 63, 65, 290, 291, 300,
 339, 340
 Perron 176
 Persistenza 259, 263, 363
 PGARCH 279
 PGARCH-M 279
 Phillips 176
 Piccolo 121
 Portafoglio 21, 51, 52, 60, 61, 81, 289,
 290, 293, 295, 296, 329, 349
 Potenza 121, 132, 263
 Premio al rischio 261, 262, 263, 270,
 307, 310
 Prevedere 23, 57, 61, 153, 215, 221, 223,
 225, 233, 252, 291, 296, 335, 349,
 360
 Previsione 57, 153, 208, 210, 212, 213,
 214, 216, 219, 220, 221, 223, 224,
 229, 247, 264, 265, 266, 267, 268,
 270, 272, 273, 289, 321, 323, 346,
 360, 363
 AR 209
 ARMA 216, 217, 218, 219
 ex ante 208, 211
 ex post 208
 GARCH 264, 273
 globale 221
 IGARCH 297
 locale 221
 MA 214
 ottimale 1, 162, 208, 244, 294, 295
 PGARCH 279
 rendimenti 208, 213, 267
 rischio 291
 TARMA 209, 219, 220, 222, 225,
 229, 277, 278, 318, 321, 329
 VaR 82, 358
 volatilità 264, 266, 270, 271, 276,
 314, 317, 321
 Previsore 209, 210, 211, 301
 Prezzo 1, 3, 4, 5, 6, 7, 8, 9, 11, 13, 15,
 18, 27, 28, 30, 32, 33, 34, 54, 61, 65,
 68, 82, 85, 116, 136, 137, 143, 145,
 150, 151, 152, 153, 155, 156, 157,
 160, 164, 165, 180, 181, 228, 280,
 283, 286, 289, 290, 329
 chiusura 5, 10
 chiusura aggiustato 5
 opzione 8
 Primo quartile 16
 Principio di parsimonia 162
 Processo
 AR(1) 166, 167
 ARMA(1, 1) 190, 330
 Bernoulli 194
 condizionato 294
 diffusione 151
 ergodico 138, 139, 166
 errore 195, 331
 iid 140, 141, 142, 236, 304, 310, 311
 invertibile 164, 182
 MA(1) 182
 non stazionario 142, 149, 167, 171,
 179

- random walk 149, 153
 residui 308, 309, 311
 soglia 194
 stazionario 138, 139, 140, 148, 162, 164, 166, 170, 175, 176, 188, 189, 291
 stazionario debole 138, 141
 stocastico 85, 136, 137, 138, 140, 142, 147, 150, 156, 162, 244, 290, 353
 TARMA 194
 white noise 141, 169, 241
 Wiener 145, 146, 178, 260
 Wiener standard
- Quasi score 117
 QMV 115, 116, 117, 118, 119, 120, 248, 283, 284, 285, 286, 287, 358, 359, 364
 Q-Qplot 18, 25, 26, 33, 34, 36, 38, 39, 41, 42, 46, 47, 73, 311
 Quotazione 5, 9, 10, 11, 14, 19, 21, 23, 48, 67
 R 330
 Radice 162, 163, 165, 171, 176, 177, 182, 259
 unitaria 176, 178, 179, 180, 181
 Random walk 149, 150, 153, 157, 177, 178
 Rapporto di verosimiglianza 120, 121, 122, 124, 129, 130, 131, 352, 354, 356, 357, 361, 362
 Rating 289
 Regimi 36, 39, 40, 41, 42, 44, 45, 46, 47, 49, 50, 51, 59, 60, 65, 66, 67, 68, 77, 78, 79, 80, 81, 83, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 221, 223, 224, 225, 226, 227, 229, 231, 233, 274, 275, 277, 278, 280, 281, 282, 317, 322, 323, 324, 325, 326, 329
 Regressione 116, 118, 119, 125, 126, 128, 129, 171, 248
 Rendimenti 1, 9, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 40, 41, 42, 43, 44, 45, 46, 47, 49, 50, 51, 52, 54, 55, 56, 57, 58, 59, 62, 63, 65, 72, 73, 75, 78, 79, 83, 85, 119, 126, 136, 137, 146, 151, 152, 153, 154, 156, 157, 159, 160, 161, 164, 165, 180, 181, 182, 188, 189, 193, 194, 197, 198, 199, 202, 211, 212, 214, 215, 216, 218, 219, 225, 229, 230, 234, 235, 236, 237, 238, 241, 244, 245, 246, 251, 254, 255, 261, 262, 263, 264, 267, 268, 269, 271, 272, 273, 277, 279, 281, 283, 286, 287, 290, 291, 294, 296, 301, 302, 304, 307, 308, 309, 312, 313, 314, 315, 320, 323, 324, 329, 331, 332, 333, 334, 335, 338, 340, 341, 343, 345, 346, 348, 349, 350, 351, 358, 363
 attivi 27, 34, 37, 38, 50, 51, 53, 151, 153, 159, 161, 188, 193, 202, 209, 210, 214, 219, 222, 229, 233, 237, 259, 261, 264, 277, 293, 294, 296, 320, 330
 al quadrato 74, 240, 242, 244, 268
 globali 36, 44, 45, 46, 194, 197, 231
 locali 36, 44, 45, 46
 minimo probabile 339
 previsione 208, 210, 213
 Residui 141, 170, 171, 172, 173, 176, 183, 186, 187, 188, 201, 205, 206, 211, 246
 al quadrato 202, 242, 261, 268, 269, 274, 315, 329
 autocorrelazione 213, 261
 modello 206, 207, 212, 241, 286, 287
 non centrati 27
 normali 260, 261, 283, 310, 311, 313, 314, 315
 simulati 187, 192, 193
 standardizzati 68, 75, 77, 237, 248, 260, 281, 283, 308, 309, 312, 313, 315, 316, 320, 329, 343, 358, 359, 360
 stimati 170, 172, 176, 187, 192, 193, 205, 211, 229, 247, 248, 311, 343
 t di Student 250, 311, 312, 314, 315,

- 316, 320, 321, 329, 334
Rimescolamento 194
Rischio 7, 8, 291, 295, 301, 325, 331, 337, 338, 339, 340, 360, 363
 cambio 4, 290
 coerente 340
 credito 290
 futuro 252, 289
 liquidità 290
 locale 78
 mercato 1, 3, 4, 5, 119, 153, 233, 289, 290, 291, 348
Rischiosità 3, 58, 60, 67, 68, 76, 77, 78, 81, 233, 247, 268, 289, 298, 300, 325, 340
Risk free 339
Riskmetrics 296, 298, 301, 304, 322, 330

Scarto 172
 quadratico medio 15, 36
Schwartz 206
Score 117, 131, 132
 function 117, 119, 120, 133
 test 131
Sd 15,
S.e. 108, 212, 213, 218, 267, 268, 272, 273, 286, 287
Segnale 22, 23, 42, 43, 44, 153
S&P 6, 155, 156, 159, 161, 211, 212, 213, 215, 216, 218, 219, 225, 226, 229
Serie 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 22, 23, 25, 27, 28, 29, 31, 41, 44, 45, 46, 47, 48, 53, 54, 55, 56, 57, 68, 69, 70, 71, 72, 73, 74, 75, 78, 79, 80, 82, 83, 88, 101, 138, 143, 145, 155, 156, 157, 159, 160, 161, 169, 179, 180, 181, 184, 185, 186, 187, 188, 191, 192, 198, 199, 200, 203, 204, 206, 208, 211, 212, 213, 215, 216, 217, 218, 221, 223, 225, 226, 227, 228, 229, 230, 238, 239, 245, 251, 253, 259, 260, 261, 264, 268, 277, 279, 280, 281, 282, 283, 313, 320, 321, 322, 326, 331, 334, 335, 349, 350, 358, 362, 363
 attivi 1, 252
Edgeworth 115,
 errori 25, 26, 65,
 standardizzata 9, 18
 standardizzata locale 65
 storica 1, 4, 5, 11, 61, 136, 138, 139, 141, 142, 144, 147, 148, 149, 150, 154, 157, 162, 166, 167, 168, 170, 171, 172, 173, 184, 190, 207, 225, 244, 280, 330, 333, 340, 343, 348
SETARMA 196
Shin 194
SIC 206, 207, 208, 212, 215, 216, 218, 313
Sistema
 equazioni 91, 100, 117, 119, 127, 128, 190
 ipotesi 120, 121, 122, 123, 124, 125, 126, 129, 131, 134, 351, 352, 354, 357
Smooth 291
Snedecor 154
Soglia 49, 155, 276, 317, 348
 modello 36, 40, 193, 197, 225, 317
 processo 194
Sottostante 7, 8
Speight 194
Splitting 5, 11
S-plus 330
Standard error 184, 201, 206, 207, 227
Standard & Poors 6, 155, 156, 159, 161, 211, 212, 213, 215, 216, 218, 219, 225, 226, 229
Standard deviation 15, 16, 17, 57, 58, 65, 75, 78, 344
 mobile 65
Standard error 159, 171, 177, 184, 186, 207, 212, 227, 285, 286, 287, 288, 358
Standardizzata 17, 18, 35, 36, 41, 58, 61, 83, 89, 92, 94, 95, 96, 115, 153, 158, 160, 211, 237, 239, 247, 249, 250, 257, 258, 271, 283, 286, 293, 294, 296, 297, 308, 309, 310, 311,

Negli ultimi due decenni si è consolidato in campo finanziario l'impiego di tecniche statistiche come strumenti operativi utili in strategie di trading, asset allocation e risk management..

Questo volume raccoglie l'esperienza didattica che, negli ultimi dieci anni, gli autori hanno maturato, nelle analisi finanziarie, in corsi di laurea triennale, magistrale, master e dottorato. Nel testo l'esposizione degli aspetti teorici viene corredata con l'illustrazione di casi studio basati su dati reali. Al fine di garantire la piena ed effettiva replicabilità delle analisi proposte, vengono forniti al lettore i relativi programmi software, in linguaggio Eviews. Ne risulta un percorso di apprendimento guidato praticabile, con diversi livelli di approfondimento: studenti universitari, dottorandi e practitioners.

Più in particolare, il volume si compone di quattro parti: a) nei capitoli I e II vengono illustrati alcuni strumenti utili per l'analisi statistica esplorativo-descrittiva dei mercati finanziari e monetari; b) il capitolo III presenta dei richiami di calcolo delle probabilità ed inferenza statistica; c) nei capitoli IV e V vengono approfonditi i principali modelli stocastici dinamici per l'analisi dei rendimenti finanziari; d) il capitolo VI, infine, presenta applicazioni di tecniche di analisi riferite al risk management.

Giuseppe Storti è professore associato di Statistica Economica presso la Facoltà di Economia dell'Università degli Studi di Salerno. In passato ha insegnato Statistica Economica, Analisi delle Serie Storiche Finanziarie ed Analisi Statistica dei Mercati Monetari e Finanziari in diversi corsi di laurea triennale e magistrale. Ha pubblicato diversi articoli su queste tematiche in riviste scientifiche di livello internazionale.

Cosimo D. Vitale è, dal 1990, professore ordinario di Statistica presso la Facoltà di Economia dell'Università di Salerno ove insegna Statistica, Processi Stocastici ed Inferenza Statistica, Analisi Statistica dei Mercati Monetari e Finanziari in corsi di primo e secondo livello e Dottorato di Ricerca. Nel passato ha insegnato Analisi dei Dati presso la Scuola di Ricerca Operativa della Facoltà di Statistica di Roma "La Sapienza", Statistica Metodologica presso la Facoltà di Scienze Economiche e Sociali dell'Università della Calabria, Statistica e Calcolo delle Probabilità presso la Facoltà di Ingegneria di Napoli "Federico II". E' autore di libri ed articoli apparsi anche in riviste di livello internazionale.

In copertina:

Attrattore di un processo stocastico limitato, non lineare e non stazionario sia localmente che globalmente.