



UNIVERSITA' DEGLI STUDI DI PADOVA
DIPARTIMENTO DI SCIENZE ECONOMICHE ED AZIENDALI
"M.FANNO"

CORSO DI LAUREA IN ECONOMIA

PROVA FINALE

**"ANALISI DELLA SERIE STORICA DEL PREZZO DEL TITOLO
MEDIOBANCA"**

RELATORE:

CH.MO PROF. NUNZIO CAPPUCCIO

LAUREANDO: EDDIE ROSSI

MATRICOLA N. 1160954

ANNO ACCADEMICO 2019 – 2020

INDICE

INTRODUZIONE	3
CAPITOLO 1	5
1.1 Stazionarietà e trend stocastici	5
1.2 Test di radice unitaria: <i>Dickey-Fuller</i> e <i>Dickey-Fuller Aumentato</i>	6
1.3 Funzione di autocorrelazione (<i>ACF</i>) e di autocorrelazione parziale (<i>PACF</i>)	7
1.4 Test di <i>Box-Ljung</i>	8
1.5 Modello <i>ARCH</i>	9
1.6 Modello <i>GARCH</i>	9
1.7 Previsioni con il modello <i>GARCH (1,1)</i>	10
1.8 Effetto <i>leverage</i> e test di asimmetria	11
1.9 Modello <i>E-GARCH</i>	13
1.10 Modello <i>T-GARCH</i>	14
CAPITOLO 2	15
2.1 Serie storica dei prezzi e dei rendimenti giornalieri	15
2.2 Specificazione modello <i>ARCH</i>	20
2.3 Specificazione modello <i>GARCH</i>	22
2.4 Valutazione della bontà delle stime sulla volatilità delle osservazioni campionarie	24
2.5 Previsioni con <i>GARCH (1,1)</i>	27
2.6 Test di asimmetria: <i>Sign Bias Test</i>	28
2.7 Specificazione modello <i>E-GARCH</i>	29
2.8 Specificazione modello <i>T-GARCH</i>	29
2.9 Serie storica dei rendimenti mensili	31
BIBLIOGRAFIA	35
SITOGRAFIA	36

ABSTRACT

Nel presente elaborato viene analizzata la serie storica dei prezzi e quindi dei rendimenti, prima giornalieri e poi mensili, del titolo *Mediobanca* dal 4/1/2010 al 29/5/2020, per un totale di 2642 osservazioni, allo scopo di studiarne la volatilità.

Mediobanca è una delle principali banche d'investimento italiane. È stata fondata nel 1946 ed è quotata nell'indice *FTSE MIB* della Borsa di Milano, con una capitalizzazione di quasi 6 miliardi di euro.

Da sempre attiva nelle attività di *corporate & investment banking*, si è specializzata nel tempo anche in servizi di credito al consumo, attraverso la controllata *Compass Banca*, e di mercato bancario retail con la controllata *CheBanca!*.

Lo studio è preceduto da un'introduzione che fornisce alcune nozioni di base sul concetto di volatilità.

Nel primo capitolo vengono presentati alcuni concetti chiave, utili a comprendere il successivo procedimento di analisi e vengono descritti i principali modelli e test statistici utilizzati in seguito.

Nel secondo capitolo viene svolta l'analisi vera e propria.

INTRODUZIONE

Studiare una serie storica significa individuare dei pattern, delle ricorrenze nel comportamento delle osservazioni passate e cercare di spiegarne le cause e di prevederne gli effetti attraverso opportuni modelli.

Guardare al passato per prevedere il futuro è il lavoro di molti analisti nel mondo della finanza e non solo. Si possono, per esempio, studiare le serie storiche delle precipitazioni nelle stagioni passate per ottimizzare il raccolto oppure quelle sulla domanda di un particolare bene per calcolarne la quantità da produrre.

Tuttavia, l'idea di riuscire a prevedere l'andamento futuro di un titolo azionario, osservandone le realizzazioni passate, rimane un'illusione.

Secondo la teoria dei mercati efficienti (Malkiel 2003), il prezzo di un titolo oggi è la miglior previsione del suo valore domani e il concetto che sta dietro a questo assunto è molto intuitivo: se tutti possiedono lo stesso insieme di informazioni e un titolo si scopre essere sopravvalutato o sottovalutato, gli investitori venderanno o compreranno fintanto che il prezzo non si stabilirà al valore corretto.

Tuttavia, sebbene i prezzi non presentino pattern prevedibili, si possono osservare dei comportamenti ricorrenti per quanto riguarda la volatilità σ_t , ossia la deviazione standard condizionata dei rendimenti:

$$\mu_t = E(r_t | I_{t-1})$$

$$\sigma_t^2 = \text{Var}(r_t | I_{t-1}) = E[(r_t - \mu_t)^2 | I_{t-1}]$$

$$\sigma_t = \sqrt{\sigma_t^2}$$

dove μ_t è il valore atteso dei rendimenti r_t per il tempo t , condizionato a I_{t-1} , cioè l'insieme delle informazioni disponibili al tempo $t - 1$.

Ci sono periodi in cui i prezzi e quindi i rendimenti sono più "volatili", cioè presentano delle variazioni più marcate in termini assoluti da un giorno all'altro e giorni in cui lo sono meno.

Questo vuol dire che se oggi si è registrata un'importante variazione rispetto al prezzo di ieri è probabile che domani succeda altrettanto, anche se non si può prevedere se in positivo o in negativo.

Il fenomeno secondo cui la volatilità tende a manifestarsi a gruppi (*cluster*) prende il nome di raggruppamenti di volatilità (in inglese *volatility clustering*) o, in termini statistici, di eteroschedasticità condizionata dei rendimenti.

Prevedere la variabilità di una serie può essere interessante per molti motivi. La volatilità di un titolo è un indicatore del rischio di detenere quel titolo, quindi è fondamentale per attività di costruzione e gestione di portafoglio e di risk management. Inoltre il valore di alcuni prodotti finanziari derivati, come ad esempio le opzioni, dipende dalla volatilità attesa del titolo sottostante.

Purtroppo la volatilità non è direttamente osservabile, ma dev'essere stimata a partire dai prezzi dei titoli o dei loro derivati.

La volatilità ottenuta a partire dalla serie storica dei rendimenti del titolo tramite modelli *ARCH* o *GARCH* è detta “volatilità storica” (in inglese *historical volatility*) ed è quella che questo elaborato si propone di studiare.

Quella che si ottiene invece con metodi di *reverse engineering* applicando la formula di *Black-Scholes* (Black e Scholes 1973) al prezzo di contratti derivati, tra cui le opzioni, che hanno come sottostante il relativo titolo azionario, prende il nome di “volatilità implicita” (in inglese *implied volatility*).

Sebbene la volatilità non sia direttamente osservabile, l'evidenza empirica che deriva dall'osservazione delle serie storiche dei rendimenti suggerisce che essa possieda le seguenti proprietà, che la rendono spesso stazionaria: la già accennata presenza di *cluster*, l'evoluzione continua nel tempo (i “salti” improvvisi sono rari) e la variazione all'interno di un range fissato, cioè essa non diverge all'infinito.

Un'altra proprietà è il cosiddetto “effetto *leverage*”, ossia la diversa influenza che hanno su di essa shock positivi e shock negativi, di cui si parlerà meglio nel primo capitolo.

Nel presente elaborato si verificherà innanzitutto che i rendimenti siano linearmente indipendenti, ossia che la serie dei rendimenti sia stazionaria, ma che vi sia dipendenza in valore assoluto o nei quadrati, cioè che la serie presenti eteroschedasticità condizionata di tipo autoregressivo. Se i test danno esito positivo si passerà successivamente alla specificazione di modelli appropriati per lo studio e la previsione della volatilità.

CAPITOLO 1

1.1 Stazionarietà e trend stocastici

La stazionarietà è una condizione fondamentale per poter modellizzare una serie storica. Il concetto di stazionarietà coincide con l'idea che il futuro sia uguale al passato. In termini statistici, una serie storica è stazionaria se la sua distribuzione di probabilità è costante nel tempo.

Tale condizione risulta violata se la serie presenta un trend, ossia una tendenza persistente nel tempo, sia essa di tipo deterministico, quindi una funzione non aleatoria del tempo, o stocastico.

Il modello più semplice di trend stocastico è la cosiddetta “*passeggiata aleatoria*” (in inglese *random walk*), in cui una variabile casuale Y_t viene espressa come somma del suo ritardo primo, Y_{t-1} , e di un'altra variabile casuale, u_t , ad essa incorrelata e con media condizionata nulla, cioè $E(u_t | Y_{t-1}, Y_{t-2}, \dots) = 0$.

$$Y_t = Y_{t-1} + u_t$$

Poiché $E(u_t | Y_{t-1}, Y_{t-2}, \dots) = 0$, la media condizionata di Y_t sulla base dei dati disponibili al tempo $t - 1$ è Y_{t-1} , ossia $E(Y_t | Y_{t-1}, Y_{t-2}, \dots) = Y_{t-1}$.

In altre parole, se Y_t segue una passeggiata aleatoria, il valore di Y domani corrisponde al valore di Y oggi più un cambiamento imprevedibile u e proprio per questo il valore di Y oggi è la migliore previsione del valore di Y domani.

Si dimostra che se una serie storica segue una passeggiata aleatoria, allora non si può dire stazionaria.

Infatti, poiché u_t e Y_{t-1} sono per ipotesi incorrelate, si può scrivere

$$\text{var}(Y_t) = \text{var}(Y_{t-1}) + \text{var}(u_t)$$

Si noti che l'uguaglianza $\text{var}(Y_t) = \text{var}(Y_{t-1})$, necessaria affinché la serie sia stazionaria, è verificata solo per $\text{var}(u_t) = 0$, condizione che di fatto annulla il termine di innovazione.

La passeggiata aleatoria è un caso particolare del modello autoregressivo di ordine m , con $m = 1$, $\beta_0 = 0$ e β_1 , coefficiente associato al primo ritardo, fissato uguale a 1.

$$Y_t = \beta_0 + \sum_{i=1}^m \beta_i Y_{t-i} + u_t$$

Si può dimostrare che se $|\beta_i| < 1$ per ogni i e u_t è stazionaria, allora anche Y_t è stazionaria.

Se invece esiste anche solo un coefficiente β pari a 1, allora si dice che la serie Y_t presenta una radice autoregressiva unitaria e che quindi contiene un trend stocastico, ossia non è stazionaria.

Esistono dei test, chiamati di radici unitarie, che ne indagano l'eventuale presenza e si pronunciano per la stazionarietà o meno della serie.

1.2 Test di radice unitaria: *Dickey-Fuller e Dickey-Fuller Aumentato*

Uno dei più famosi test di radice unitaria è il test di *Dickey-Fuller* (Dickey e Fuller 1979).

Il test si basa su un modello autoregressivo di primo ordine e ha come ipotesi nulla la presenza di una radice unitaria, contro l'ipotesi alternativa di stazionarietà.

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + u_t$$

$$H_0: \beta_1 = 1$$

$$H_1: \beta_1 < 1$$

Il modello può essere modificato sottraendo Y_{t-1} ad entrambi i membri, ottenendo

$$\Delta Y_t = \beta_0 + \delta Y_{t-1} + u_t$$

$$H_0: \delta = 0$$

$$H_1: \delta < 0$$

dove $\delta = \beta_1 - 1$.

In questo modo la statistica di *Dickey-Fuller* risulta essere semplicemente la statistica t degli *OLS* per la verifica di $\delta = 0$.

Si noti che il test di *Dickey-Fuller* è unilaterale, poiché l'alternativa rilevante è che Y_t sia stazionaria e quindi $\beta_1 < 1$ o, in alternativa, $\delta < 0$.

Talvolta un modello autoregressivo di primo ordine non è sufficiente a catturare tutta la correlazione seriale in Y_t e risulta necessario "aumentare" il modello alla base del test con dei ritardi di ΔY_t , da qui il nome *Dickey-Fuller Aumentato*.

$$\Delta Y_t = \beta_0 + \delta Y_{t-1} + \gamma_1 \Delta Y_{t-1} + \gamma_2 \Delta Y_{t-2} + \dots + \gamma_p \Delta Y_{t-p} + u_t$$

Le ipotesi nulla e alternativa non cambiano, e la statistica ADF rimane la statistica t degli OLS per la verifica di $\delta = 0$.

Tuttavia la statistica ADF non ha una distribuzione *Normale*, nemmeno per campioni sufficientemente grandi. Si riportano quindi di seguito i valori critici per il test ADF unilaterale.

Livelli di significatività	10%	5%	1%
Valori critici	-2,57	-2,86	-3,43

1.3 Funzione di autocorrelazione (ACF) e di autocorrelazione parziale (PACF)

L'autocovarianza e l'autocorrelazione sono rispettivamente la covarianza e la correlazione di una variabile con i propri valori ritardati. Quindi la j -esima autocovarianza e la j -esima autocorrelazione di una serie temporale Y_t sono rispettivamente la covarianza e la correlazione tra Y_t e Y_{t-j} .

$$\text{Autocovarianza}_j = \text{cov}(Y_t, Y_{t-j})$$

$$\text{Autocorrelazione}_j = \text{corr}(Y_t, Y_{t-j}) = \frac{\text{cov}(Y_t, Y_{t-j})}{\sqrt{\text{var}(Y_t)\text{var}(Y_{t-j})}}$$

Entrambe possono essere stimate attraverso i corrispettivi stimatori campionari $\text{cov}(\widehat{Y}_t, \widehat{Y}_{t-j})$ e $\widehat{\rho}_j$.

$$\text{cov}(\widehat{Y}_t, \widehat{Y}_{t-j}) = \frac{1}{T} \sum_{t=j+1}^T (Y_t - \bar{Y}_{j+1,T})(Y_{t-j} - \bar{Y}_{1,T-j})$$

$$\widehat{\rho}_j = \frac{\text{cov}(\widehat{Y}_t, \widehat{Y}_{t-j})}{\widehat{\text{var}}(Y_t)}$$

dove $\bar{Y}_{j+1,T}$ indica la media campionaria di Y_t , calcolata sulle osservazioni $t = j + 1, \dots, T$ e $\widehat{\text{var}}(Y_t)$ la varianza campionaria di Y .

Chiamiamo funzione di autocorrelazione (in inglese *Autocorrelation Function, ACF*) la funzione che associa ad ogni ritardo j di Y_t , l'autocorrelazione tra Y_t e Y_{t-j} .

La correlazione parziale tra due variabili, invece, è la correlazione che rimane dopo che l'eventuale impatto di tutte le altre variabili casuali è stato eliminato.

Per definire il coefficiente di autocorrelazione parziale consideriamo la seguente regressione

$$Y_t = \varphi_{1k}Y_{t-1} + \varphi_{2k}Y_{t-2} + \dots + \varphi_{jk}Y_{t-j} + \dots + \varphi_{kk}Y_{t-k} + u_t$$

dove φ_{jk} è il coefficiente relativo alla variabile di ritardo j , se il processo ha ordine k .

I coefficienti φ_{jk} sono i coefficienti di autocorrelazione parziale (di ordine k). L'autocorrelazione parziale misura la correlazione tra Y_t e Y_{t-j} che rimane quando l'influenza di $Y_{t-2}, Y_{t-3}, \dots, Y_{t-k}$ è stata eliminata.

Chiamiamo funzione di autocorrelazione parziale (in inglese *Partial Autocorrelation Function*, *PACF*) la funzione che associa ad ogni ritardo j di Y_t , l'autocorrelazione parziale tra Y_t e Y_{t-j} .

1.4 Test di *Box-Ljung*

Il test di *Box-Ljung* (Ljung e Box 1978) viene utilizzato per verificare se l'autocorrelazione di una serie storica con i suoi k ritardi è diversa da zero. Invece di testare ogni ritardo singolarmente, verifica congiuntamente la presenza di autocorrelazione.

Secondo l'ipotesi nulla le osservazioni sono distribuite in modo indipendente, mentre l'ipotesi alternativa stabilisce la presenza di autocorrelazione significativa della serie con almeno uno dei ritardi.

La statistica test può essere scritta come

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

dove k è il numero dei ritardi testati, n è la dimensione del campione e $\hat{\rho}_t^2$ è la k -esima autocorrelazione campionaria.

Sotto l'ipotesi nulla di assenza di autocorrelazione, la statistica $Q(k)$ si distribuisce come una variabile casuale *chi-quadrato* $\chi_{(k)}^2$ con k gradi di libertà.

1.5 Modello ARCH

Il modello *ARCH* (*Autoregressive Conditional Heteroskedasticity*), introdotto da Engel (1982), si propone di studiare la dipendenza non lineare delle osservazioni di una serie storica, esprimendo la varianza dei residui di una regressione principale (*mean equation*) in funzione del quadrato dei residui ritardati.

Nel presente elaborato la *mean equation* consiste in una semplice equazione:

$$r_t = \mu_t + a_t$$

$$\mu_t = E(r_t | I_{t-1})$$

dove μ_t è il valore atteso dei rendimenti r_t per il tempo t , condizionato a I_{t-1} , cioè l'insieme delle informazioni disponibili al tempo $t - 1$ e a_t è l'innovazione al tempo t .

Segue la specificazione di un modello *ARCH* vero e proprio di ordine p :

$$a_t = \sigma_t \epsilon_t$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i a_{t-i}^2$$

dove ϵ_t è una sequenza di variabili casuali indipendenti e identicamente distribuite (*iid*) con media 0 e varianza 1, p è il numero di innovazioni ritardate al quadrato a_{t-i}^2 incluse nel modello, $\alpha_0 > 0$ e $\alpha_i \geq 0$ per $i > 0$ e tali per cui la varianza condizionata di a_t sia finita.

Solitamente si assume che ϵ_t segua una distribuzione *Normale* o una *t di Student*.

σ_t rappresenta la volatilità dei rendimenti e il modello è costruito in modo tale per cui tanto più grandi sono, in valore assoluto, le innovazioni passate, tanto più probabile è osservare al tempo presente un elevato valore di σ_t e ciò è in linea con il fenomeno dei raggruppamenti di volatilità (*volatility clustering*) che si osserva nei mercati finanziari.

1.6 Modello GARCH

Il modello *GARCH* (*Generalized Autoregressive Conditional Heteroskedasticity*), introdotto da Bollerslev (1986), è un'estensione del modello *ARCH* in cui la varianza dei residui della regressione principale viene espressa in funzione, oltre che del quadrato dei residui ritardati, anche dei ritardi della varianza stessa dei residui.

Ciò permette di superare un grosso limite dei modelli *ARCH*, che è quello di dover includere molti ritardi per poter ottenere stime appropriate della volatilità.

La specificazione e i vincoli sui parametri sono simili a quelli del modello *ARCH*:

$$r_t = \mu_t + a_t$$

$$\mu_t = E(r_t | I_{t-1})$$

$$a_t = \sigma_t \epsilon_t$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i a_{t-i}^2 + \sum_{i=1}^q \beta_i \sigma_{t-i}^2$$

dove p e q sono rispettivamente il numero di innovazioni ritardate al quadrato a_{t-i}^2 e il numero dei valori ritardati della varianza delle innovazioni σ_{t-i}^2 inclusi nel modello e $\beta_i \geq 0$ e tali per cui la varianza condizionata di a_t sia finita.

1.7 Previsioni con il modello *GARCH* (1,1)

Una volta stimato un modello *GARCH* (1,1) è possibile utilizzarlo per fare delle previsioni sulla volatilità futura.

$$a_t = \sigma_t \epsilon_t$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

Supponiamo di trovarci al tempo h . La volatilità prevista per la successiva seduta di borsa, quella al tempo $h + 1$, sarà

$$\sigma_{h+1}^2 = \alpha_0 + \alpha_1 a_h^2 + \beta_1 \sigma_h^2$$

dove a_h e σ_h^2 sono note al tempo h e α_0 , α_1 e β_1 sono i coefficienti stimati del modello *GARCH* (1,1).

Perciò, in generale, la previsione sul primo periodo successivo si esprime come

$$\sigma_h^2(1) = \alpha_0 + \alpha_1 a_h^2 + \beta_1 \sigma_h^2$$

Per fare previsioni su più periodi futuri successivi al primo, riscriviamo l'equazione ricordando che $a_t^2 = \sigma_t^2 \epsilon_t^2$:

$$\sigma_{t+1}^2 = \alpha_0 + (\alpha_1 + \beta_1) \sigma_t^2 + \alpha_1 \sigma_t^2 (\epsilon_t^2 - 1)$$

Quando $t = h + 1$ l'equazione diventa

$$\sigma_{h+2}^2 = \alpha_0 + (\alpha_1 + \beta_1)\sigma_{h+1}^2 + \alpha_1\sigma_{h+1}^2(\epsilon_{h+1}^2 - 1)$$

Dato che $E(\epsilon_{h+1}^2 - 1|F_h) = 0$, la previsione in h sul secondo periodo successivo diventa

$$\sigma_h^2(2) = \alpha_0 + (\alpha_1 + \beta_1)\sigma_h^2(1)$$

In generale, si può esprimere la previsione in h sulla realizzazione dell' (l) -esimo periodo successivo, $\sigma_h^2(l)$, in funzione dell' $(l - 1)$ -esimo periodo futuro, $\sigma_h^2(l - 1)$:

$$\sigma_h^2(l) = \alpha_0 + (\alpha_1 + \beta_1)\sigma_h^2(l - 1), \quad l > 1$$

Sostituendo ricorsivamente possiamo esprimere la previsione dell' (l) -esimo periodo successivo, $\sigma_h^2(l)$, in funzione della previsione sul primo periodo successivo, $\sigma_h^2(1)$:

$$\sigma_h^2(l) = \frac{\alpha_0[1 - (\alpha_1 + \beta_1)^{l-1}]}{1 - \alpha_1 - \beta_1} + (\alpha_1 + \beta_1)^{l-1}\sigma_h^2(1)$$

Si nota che le previsioni di un *GARCH* $(1,1)$ convergono a una varianza non condizionata di a_t man mano che l'orizzonte di previsione aumenta.

Infatti:

$$\lim_{l \rightarrow \infty} \sigma_h^2(l) = \frac{\alpha_0}{1 - \alpha_1 - \beta_1}$$

dove si suppone che $\alpha_1 + \beta_1 < 1$ e che $var(a_t)$ esista.

1.8 Effetto *leverage* e test di asimmetria

Si osserva spesso nei mercati azionari che i movimenti al ribasso tendono ad essere seguiti da una volatilità maggiore di quella generata da movimenti al rialzo di pari grandezza in termini assoluti.

Ciò significa che gli shock (o innovazioni) hanno un impatto asimmetrico sulla volatilità a seconda del segno con cui si verificano.

Questo fenomeno è conosciuto come effetto *leverage* (Black 1976) e in termini statistici potrebbe essere definito come la presenza di correlazione negativa tra il rendimento al tempo $t - 1$, r_{t-1} , e il quadrato della volatilità al tempo t , σ_t^2 , o, similmente, tra il ritardo primo dell'innovazione, a_{t-1} , e il quadrato dell'innovazione presente, a_t^2 .

Secondo il fenomeno dei *volatility clustering*, elevate innovazioni in valore assoluto tendono ad essere seguite da innovazioni altrettanto elevate in valore assoluto, perciò le innovazioni con segno positivo saranno positivamente correlate con il quadrato delle innovazioni del periodo successivo.

Per le innovazioni con segno negativo, invece, la correlazione è negativa: tanto minore (ossia maggiore in termini di valore assoluto) è l'innovazione oggi, tanto maggiore è il quadrato dell'innovazione domani.

Quello che ci si aspetta in presenza di asimmetria è che il secondo effetto prevalga sul primo è che quindi, complessivamente, la correlazione tra il ritardo primo dell'innovazione, a_{t-1} , e il quadrato dell'innovazione presente, a_t^2 , sia negativa.

Il modello *GARCH* standard non è in grado di cogliere questo effetto, perché tratta in modo simmetrico le innovazioni, che infatti figurano nel modello elevate al quadrato.

Per catturare l'effetto *leverage* sono state introdotte, a partire dal modello *GARCH*, delle modifiche, che hanno portato alla nascita dei cosiddetti modelli *GARCH* asimmetrici.

Uno dei principali test di asimmetria è il *Sign Bias Test*, proposto da Engle e Ng (1993).

Esso consiste nel condurre una semplice regressione *OLS* del quadrato dei residui standardizzati di un modello *GARCH* simmetrico su una variabile dummy che rappresenta il segno del ritardo primo delle innovazioni e può essere scritto come

$$\frac{a_t^2}{\sigma_t^2} = \alpha + \beta D_{t-1} + u_t$$

$$D_{t-1} = \begin{cases} 1, & a_{t-1} < 0 \\ 0, & a_{t-1} \geq 0 \end{cases}$$

dove u_t è una variabile casuale con media zero, $\frac{a_t^2}{\sigma_t^2}$ è la serie dei residui standardizzati al quadrato e D_{t-1} è una variabile dummy che assume il valore 1 se $a_{t-1} < 0$ e 0 altrimenti.

Se $\beta \neq 0$, l'effetto medio su $\frac{a_t^2}{\sigma_t^2}$ è differente a seconda che il rendimento del periodo precedente sia negativo, $\alpha + \beta$, o positivo, α . In particolare l'effetto leverage si manifesta se $\beta > 0$.

Pertanto, una volta stimati i parametri della regressione, si calcola la statistica t per il coefficiente β relativo alla variabile dummy, D_{t-1} , verificando il seguente sistema d'ipotesi:

$$\begin{cases} H_0: \beta = 0 \\ H_1: \beta > 0 \end{cases}$$

1.9 Modello *E-GARCH*

Uno dei più famosi modelli asimmetrici è il cosiddetto modello *E-GARCH* (*Exponential GARCH*) introdotto da Nelson (1991).

Un modello *E-GARCH* (m, s) può essere specificato come segue:

$$a_t = \sigma_t \epsilon_t$$

$$\ln(\sigma_t^2) = \alpha_0 + \frac{1 + \beta_1 L + \dots + \beta_{s-1} L^{s-1}}{1 - \alpha_1 L - \dots - \alpha_m L^m} g(\epsilon_{t-1})$$

$$g(\epsilon_t) = \begin{cases} (\theta - \gamma)\epsilon_t, & \epsilon_t < 0 \\ (\theta + \gamma)\epsilon_t, & \epsilon_t \geq 0 \end{cases}$$

dove L è l'operatore "ritardo" tale che $Lg(\epsilon_t) = g(\epsilon_{t-1})$ e α_i, β_j sono parametri non negativi che soddisfano condizioni simili a quelle dei modelli *GARCH*.

Per semplicità nel processo di stima alcuni software statistici usano la seguente versione leggermente modificata del modello:

$$a_t = \sigma_t \epsilon_t$$

$$\ln(\sigma_t^2) = \alpha_0 + \sum_{i=1}^m \alpha_i \frac{|a_{t-i}| + \gamma_i a_{t-i}}{\sigma_{t-i}} + \sum_{j=1}^s \beta_j \ln(\sigma_{t-j}^2)$$

Un'innovazione a_{t-i} positiva contribuisce $\alpha_i(1 + \gamma_i)|\epsilon_{t-i}|$ al logaritmo della varianza condizionata $\ln(\sigma_t^2)$, mentre un'innovazione a_{t-i} negativa vi contribuisce $\alpha_i(1 - \gamma_i)|\epsilon_{t-i}|$, dove $\epsilon_{t-i} = a_{t-i}/\sigma_{t-j}$. In questo modo, il parametro γ_i è una stima dell'effetto *leverage* e ci si aspetta che assuma valori negativi.

1.10 Modello *T-GARCH*

Un altro tra i più importanti modelli asimmetrici è il cosiddetto modello *T-GARCH* (*Threshold GARCH*) (Glosten, Jagannathan, e Runkle 1993; Zakoian 1994).

Un modello *T-GARCH* (m, s) può essere specificato come segue:

$$a_t = \sigma_t \epsilon_t$$
$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^m (\alpha_i + \gamma_i D_{t-i}) a_{t-i}^2 + \sum_{j=1}^s \beta_j \sigma_{t-j}^2$$
$$D_{t-i} = \begin{cases} 1, & a_{t-1} < 0 \\ 0, & a_{t-1} \geq 0 \end{cases}$$

dove α_i , γ_i , β_j sono parametri non negativi che soddisfano condizioni simili a quelle dei modelli *GARCH* e D_{t-i} è una variabile dummy che assume il valore 1 se $a_{t-1} < 0$ e 0 altrimenti.

Il modello è costruito in modo da pesare diversamente le innovazioni positive, α_i , da quelle negative, $\alpha_i + \gamma_i$, così che il parametro γ_i è una stima dell'effetto *leverage* e ci si aspetta che assuma valori positivi.

CAPITOLO 2

2.1 Serie storica dei prezzi e dei rendimenti giornalieri

Punto di partenza della presente analisi è la serie storica dei prezzi giornalieri del titolo *Mediobanca* dal 4/1/2010 al 29/5/2020, per un totale di 2642 osservazioni. Per ogni giornata viene considerato il prezzo di chiusura aggiustato per dividendi e frazionamenti. I dati sono stati raccolti dalla piattaforma *Yahoo finance*.

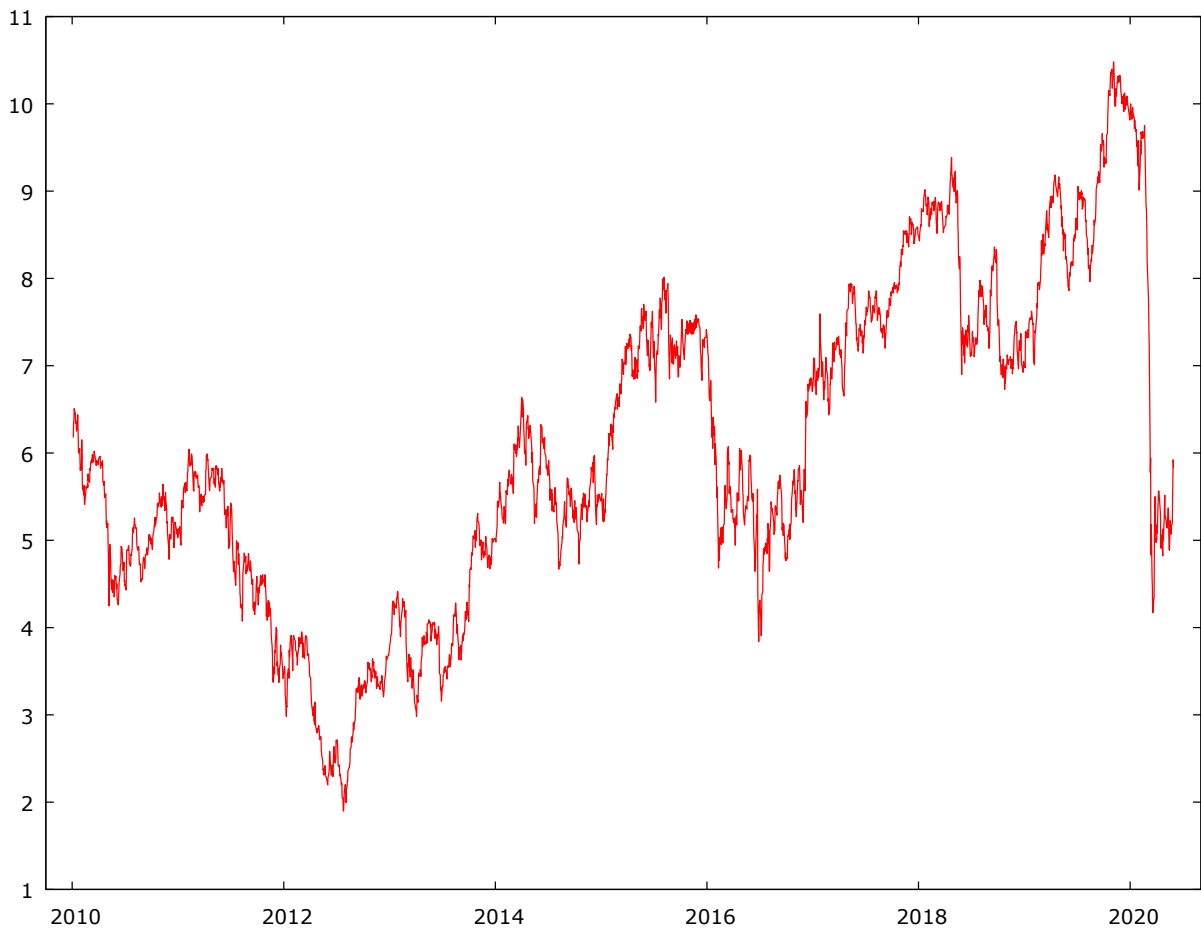


Figura 1: Grafico della serie storica dei prezzi di chiusura giornalieri aggiustati del titolo Mediobanca. Osservazioni dal 4/1/2010 al 29/5/2020.

Abbiamo definito la volatilità come la deviazione standard condizionata dei rendimenti giornalieri.

Ricaviamo dunque la serie storica dei rendimenti giornalieri come differenza prima del logaritmo naturale dei prezzi (che corrisponde circa alla variazione percentuale giornaliera dei prezzi, quindi al rendimento percentuale giornaliero).

$$r_t = \ln(p_t) - \ln(p_{t-1}) \approx \frac{p_t - p_{t-1}}{p_t}$$

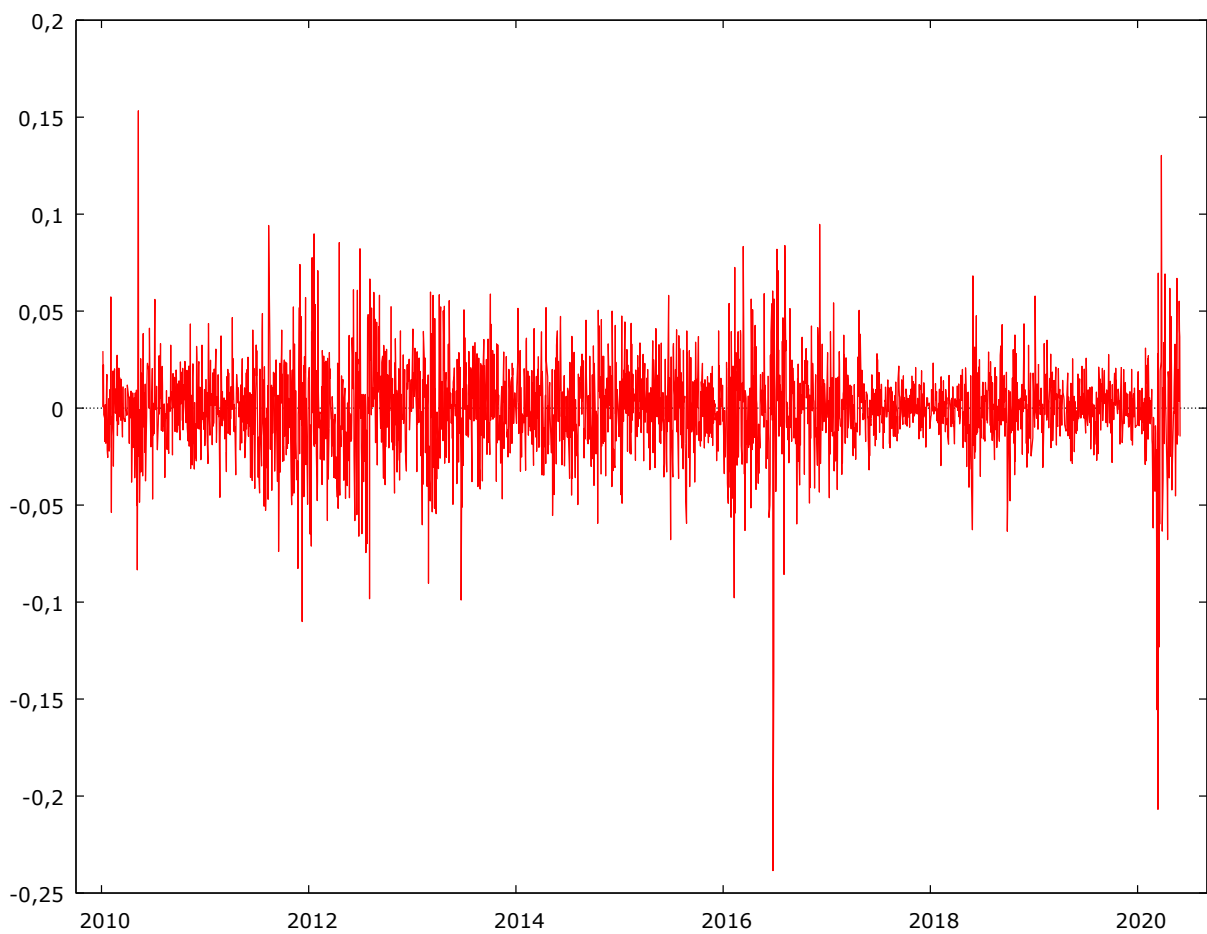


Figura 2: Grafico della serie storica dei rendimenti percentuali giornalieri del titolo Mediobanca.

Il motivo per cui utilizziamo la serie storica dei rendimenti e non quella dei prezzi è che quest'ultima non è stazionaria.

Osservando infatti il correlogramma relativo al logaritmo dei prezzi e si nota un'autocorrelazione significativa all'1% anche oltre il ventesimo ritardo.

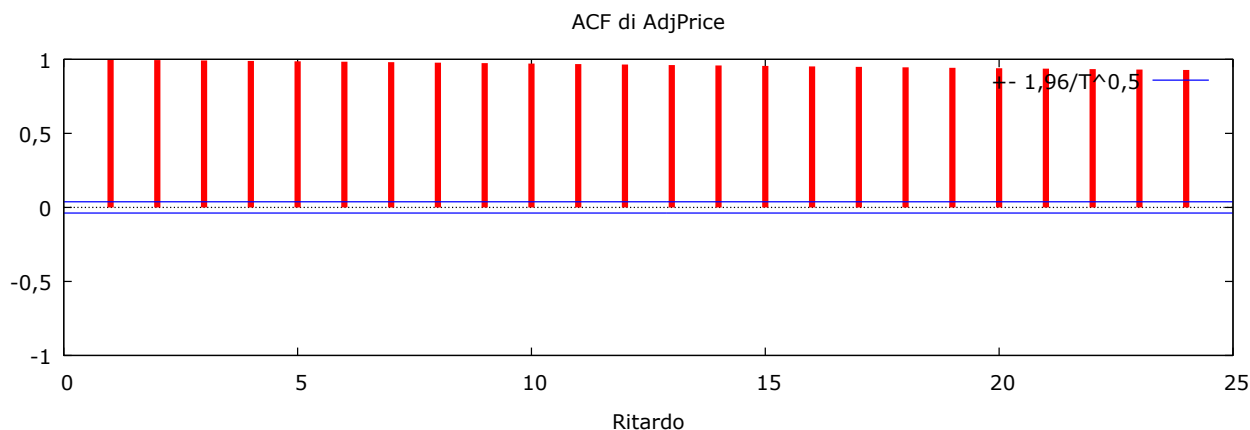


Figura 3: Correlogramma con funzione di autocorrelazione (ACF) della serie storica dei prezzi di chiusura giornalieri aggiustati. Vengono considerati i primi 24 ritardi.

Eseguiamo inoltre un test di radice unitaria sulla serie dei logaritmi dei prezzi, in particolare il *test ADF (Augmented Dickey-Fuller)*.

Includiamo i primi 24 ritardi di ΔY_t .

Il test restituisce un *p-value* di 0,3614: accettiamo dunque l'ipotesi nulla di presenza di radici unitarie.

La stazionarietà è un requisito fondamentale per la costruzione di un modello sulla volatilità.

Prima di procedere assicuriamoci che la serie dei rendimenti rispetti questo criterio, verificando che le osservazioni non presentino autocorrelazioni significative.

Osservando il correlogramma notiamo una lieve correlazione con i primi due ritardi (nessuna della quali tuttavia significativa all'1%), mentre per i ritardi successivi l'ACF va a 0.

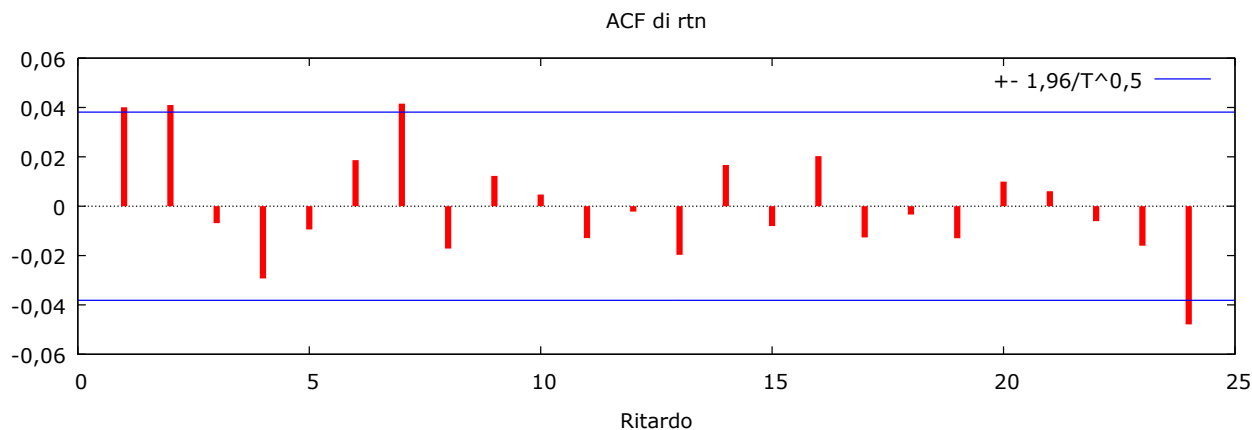


Figura 4: Correlogramma con funzione di autocorrelazione (ACF) della serie storica dei rendimenti percentuali giornalieri. Vengono considerati i primi 24 ritardi.

L'ipotesi che sta alla base della costruzione di un modello sulla volatilità è che i rendimenti non siano linearmente autocorrelati, ma che vi sia una qualche dipendenza tra di loro. Se infatti osserviamo il correlogramma relativo al valore assoluto e al quadrato dei rendimenti notiamo una significativa autocorrelazione rispettivamente oltre il ventesimo e fino al dodicesimo ritardo.

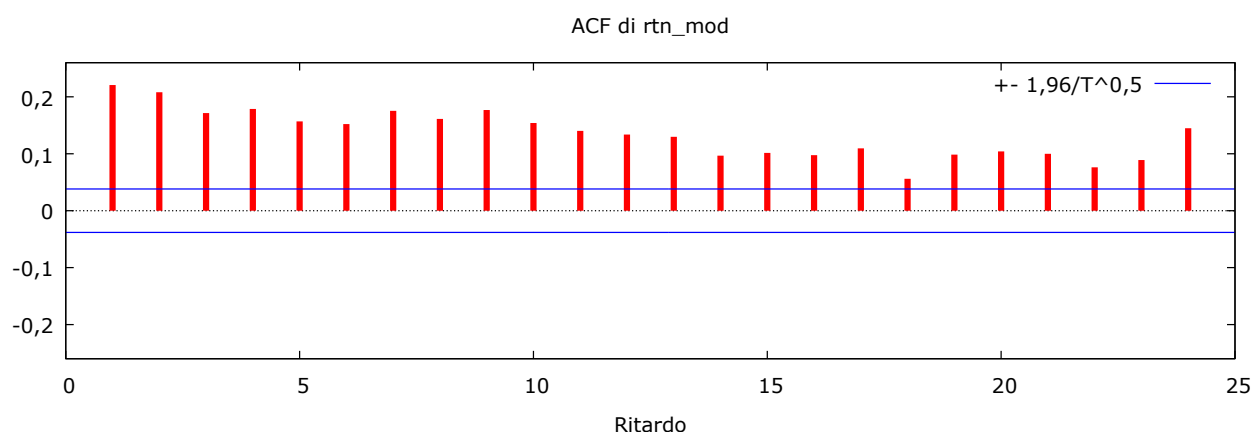


Figura 5: Correlogramma con funzione di autocorrelazione (ACF) della serie storica dei rendimenti percentuali giornalieri in valore assoluto. Vengono considerati i primi 24 ritardi.

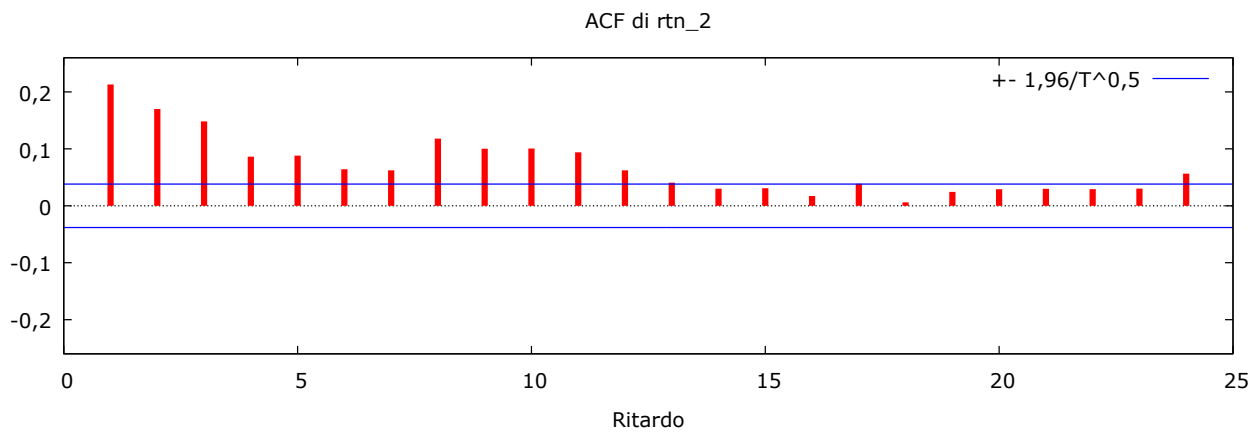


Figura 6: Correlogramma con funzione di autocorrelazione (ACF) della serie storica dei quadrati dei rendimenti percentuali giornalieri. Vengono considerati i primi 24 ritardi.

La presenza di eteroschedasticità condizionata autoregressiva (ARCH) è verificata anche attraverso il test di *Box-Ljung*, che mostra forti effetti ARCH con una statistica test $Q(12) = 439,96$ e *p-value* dell'ordine di 10^{-16} .

2.2 Specificazione modello ARCH

Sviluppiamo un modello ARCH a partire dalla serie r_t dei rendimenti giornalieri.

La *mean equation* stimata è la seguente:

$$r_t = 2,14625 \cdot 10^{-4} + a_t$$

da cui ricaviamo la serie delle innovazioni a_t .

Applichiamo la funzione PACF (*Partial Autocorrelation Function*) alla serie a_t^2 delle innovazioni al quadrato per determinare l'ordine p del modello.

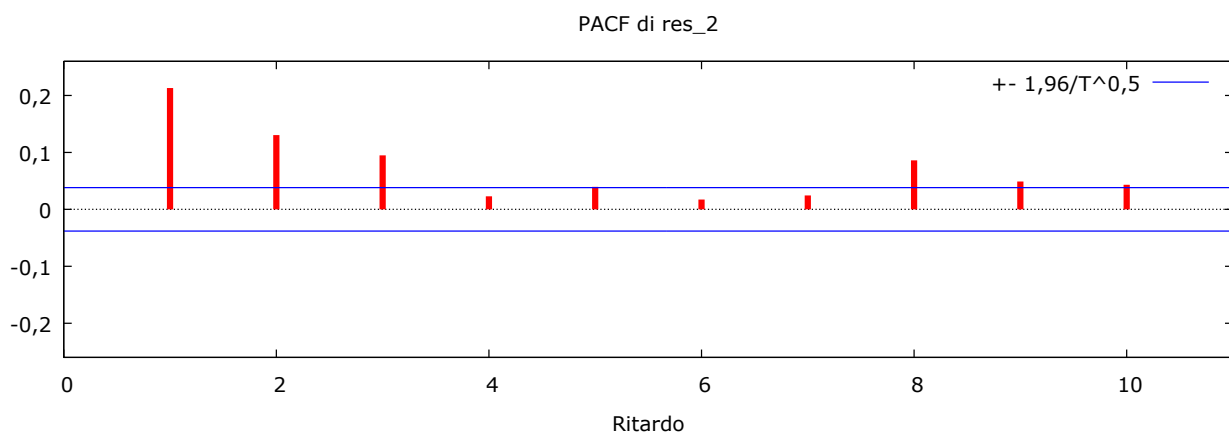


Figura 7: Correlogramma con funzione di autocorrelazione parziale (PACF) della serie storica dei quadrati delle innovazioni. Vengono considerati i primi 10 ritardi.

La correlazione parziale risulta essere significativa con i primi tre ritardi $a_{t-1}^2, a_{t-2}^2, a_{t-3}^2$.

Specifichiamo quindi il modello ARCH (3)

$$a_t = \sigma_t \epsilon_t$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \alpha_2 a_{t-2}^2 + \alpha_3 a_{t-3}^2$$

	Coefficiente	Errore std.	Statistica t	p-value
α_0	0,000255453	1,47601e-05	17,308	< 2e-16 ***
α_1	0,263165	0,0331981	7,929	2,24e-015 ***
α_2	0,173160	0,0278297	6,221	4,95e-010 ***
α_3	0,162827	0,0261660	6,223	4,88e-010 ***

I coefficienti stimati sono tutti positivi e significativi all'1%.

Perciò il modello può essere scritto come

$$\sigma_t^2 = 0,000218679 + 0,263165 a_{t-1}^2 + 0,173160 a_{t-2}^2 + 0,162827 a_{t-3}^2$$

Indaghiamo un'eventuale autocorrelazione residua nella serie

$$\tilde{a}_t^2 = \left(\frac{a_t}{\sigma_t}\right)^2$$

dei residui standardizzati al quadrato, che indicherebbe l'incapacità del modello di spiegare completamente gli effetti *ARCH*.

Lo facciamo innanzitutto osservandone il correlogramma, che, a eccezione dei primi tre ritardi inclusi nel modello, mostra un'autocorrelazione significativa con gran parte dei rimanenti.

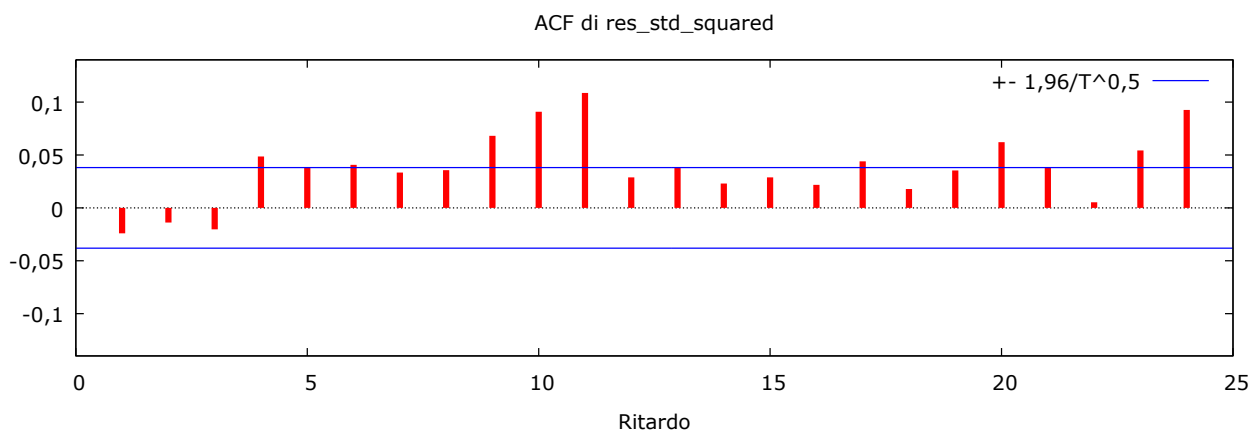


Figura 8: Correlogramma con funzione di autocorrelazione (ACF) della serie storica dei quadrati dei residui standardizzati del modello ARCH (3). Vengono considerati i primi 24 ritardi.

Eseguiamo inoltre il test di *Box-Ljung* considerando i primi 10 ritardi di \tilde{a}_t^2 . Otteniamo una statistica $Q(10) = 58,26125$ a cui corrisponde un *p-value* di $7,71869 \cdot 10^{-9}$, che ci porta a rifiutare l'ipotesi nulla di assenza di eteroschedasticità condizionata autoregressiva residua, confermando quanto si era ipotizzato a partire dal correlogramma.

Il modello non è quindi in grado di spiegare completamente l'eteroschedasticità condizionata dei residui.

2.3 Specificazione modello *GARCH*

Nella pratica si osserva che un modello *GARCH* (1,1) è spesso sufficiente a spiegare con buona precisione la volatilità.

La *mean equation* è la stessa specificata in precedenza:

$$r_t = 2,14625 \cdot 10^{-4} + a_t$$

Specifichiamo quindi il modello *GARCH* (1,1):

$$a_t = \sigma_t \epsilon_t$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

	<i>Coefficiente</i>	<i>Errore std.</i>	<i>Statistica t</i>	<i>p-value</i>
α_0	4,56068e-06	1,51624e-06	3,046	0,00232 ***
α_1	0,104188	0,0108085	9,660	<2e-16 ***
β_1	0,895332	0,0103669	86,928	<2e-16 ***

I coefficienti stimati sono tutti positivi e significativi all'1%.

Perciò il modello può essere scritto come

$$\sigma_t^2 = 4,56068 \cdot 10^{-6} + 0,104188 a_{t-1}^2 + 0,895332 \sigma_{t-1}^2$$

Indaghiamo infine un'eventuale autocorrelazione residua nella serie

$$\tilde{a}_t^2 = \left(\frac{a_t}{\sigma_t} \right)^2$$

dei residui standardizzati al quadrato.

Il correlogramma suggerisce un'autocorrelazione praticamente assente per la quasi totalità dei ritardi.

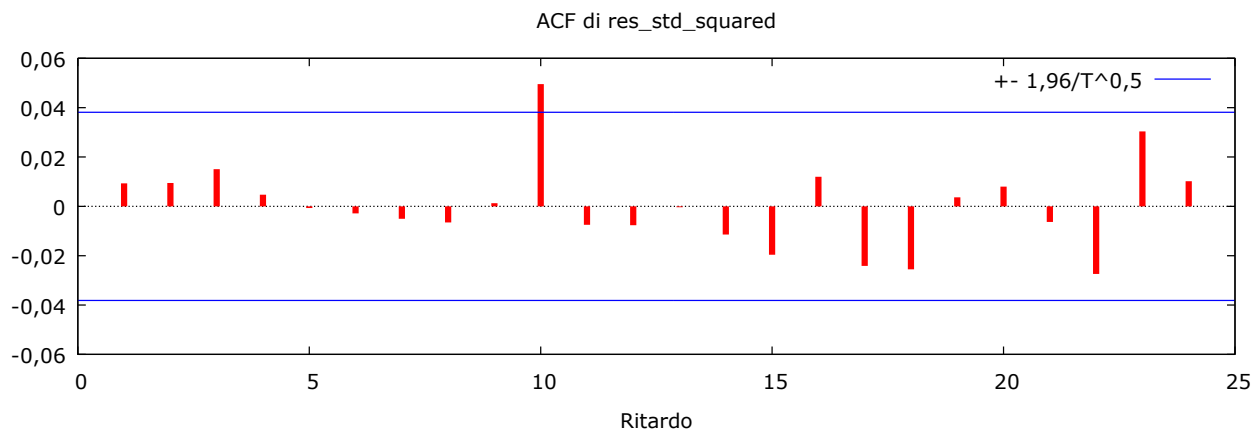


Figura 9: Correlogramma con funzione di autocorrelazione (ACF) della serie storica dei quadrati dei residui standardizzati del modello GARCH (1,1). Vengono considerati i primi 24 ritardi.

Per confermare questa ipotesi applichiamo il test di *Box-Ljung*, considerando i primi 10 ritardi di $\tilde{\alpha}_t^2$. Otteniamo una statistica $Q(10) = 7,842908$ a cui corrisponde un *p-value* di 0,6442, che ci porta ad accettare l'ipotesi nulla di assenza di eteroschedasticità condizionata autoregressiva residua.

Il modello cattura quindi completamente l'effetto *ARCH* dei residui.

2.4 Valutazione della bontà delle stime sulla volatilità delle osservazioni campionarie

A partire dal modello $GARCH(1,1)$, ricaviamo la serie storica delle stime sulla volatilità σ_t .

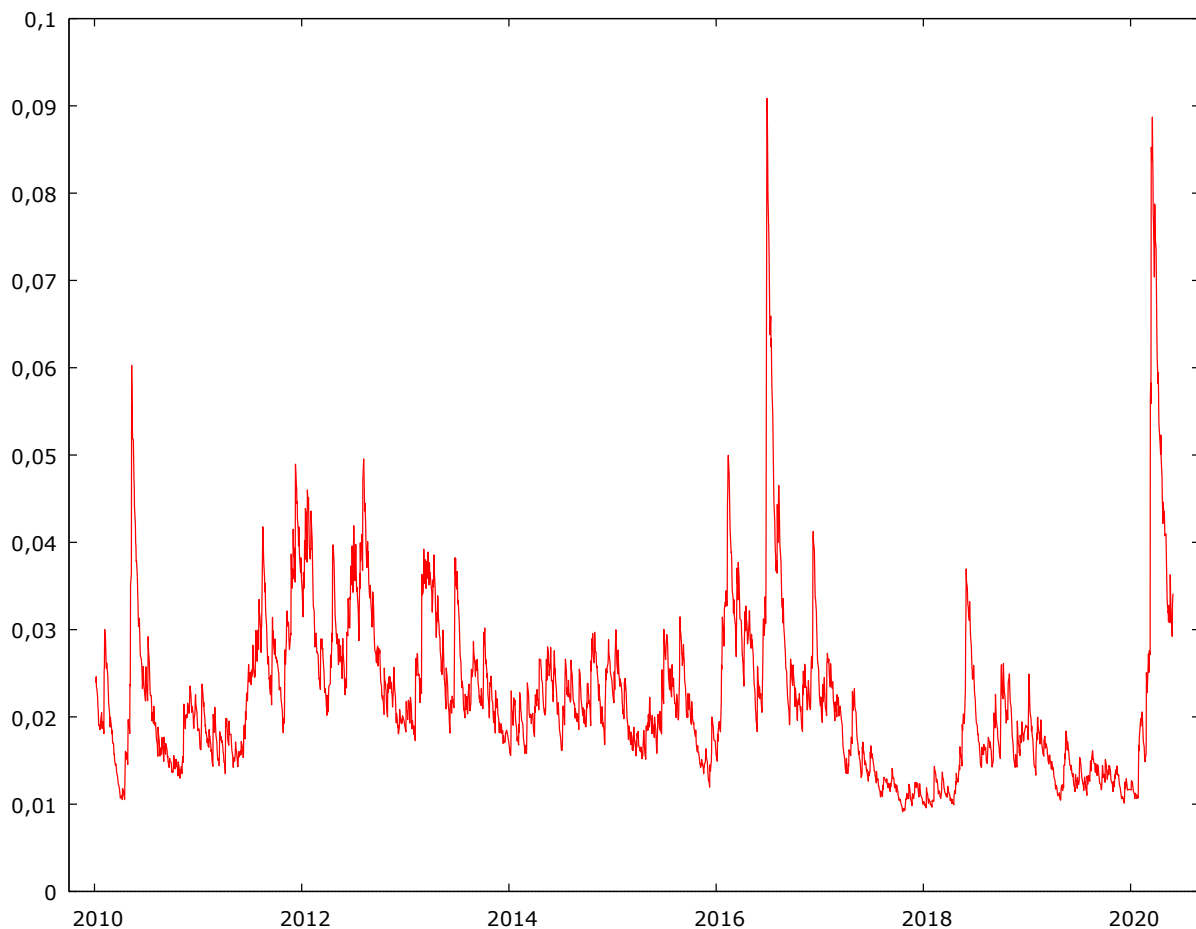


Figura 10: Serie storica della volatilità stimata con il modello $GARCH(1,1)$.

Per poter utilizzare il modello per fare previsioni, siamo interessati a sapere se le stime siano attendibili.

Un approccio comune per valutare la rilevanza pratica di qualsiasi modello è il confronto delle previsioni implicite con le successive realizzazioni. Sfortunatamente, la volatilità non è osservata direttamente, quindi questo approccio non è immediatamente applicabile per la valutazione delle previsioni della volatilità.

Una possibile idea potrebbe essere quella di confrontare la volatilità cosiddetta storica, quella che abbiamo ricavato attraverso l'analisi della serie temporale dei rendimenti, con la volatilità implicita.

Purtroppo questo confronto non è possibile per questo elaborato perché non abbiamo a disposizione la serie storica dei prezzi delle opzioni tra il 2010 e il 2020.

Un'altra possibilità è quella di utilizzare le innovazioni dei rendimenti (che corrispondono ai residui della *mean equation*) al quadrato, a_t^2 , come approssimazione della volatilità realizzata *ex post* e confrontarle, attraverso una semplice regressione OLS, con il quadrato della volatilità stimata, σ_t^2 .

Ricordiamo che

$$a_t = \sigma_t \cdot \epsilon_t$$

dove ϵ_t indica un processo stocastico indipendente di varianza unitaria e media zero, mentre la volatilità latente, σ_t , evolve secondo il modello *GARCH (1,1)* stimato.

Questa scelta è giustificata dal fatto che, se il modello è specificato correttamente, allora $E_{t-1}(a_t^2) = E_{t-1}(\sigma_t^2 \cdot \epsilon_t^2) = \sigma_t^2$

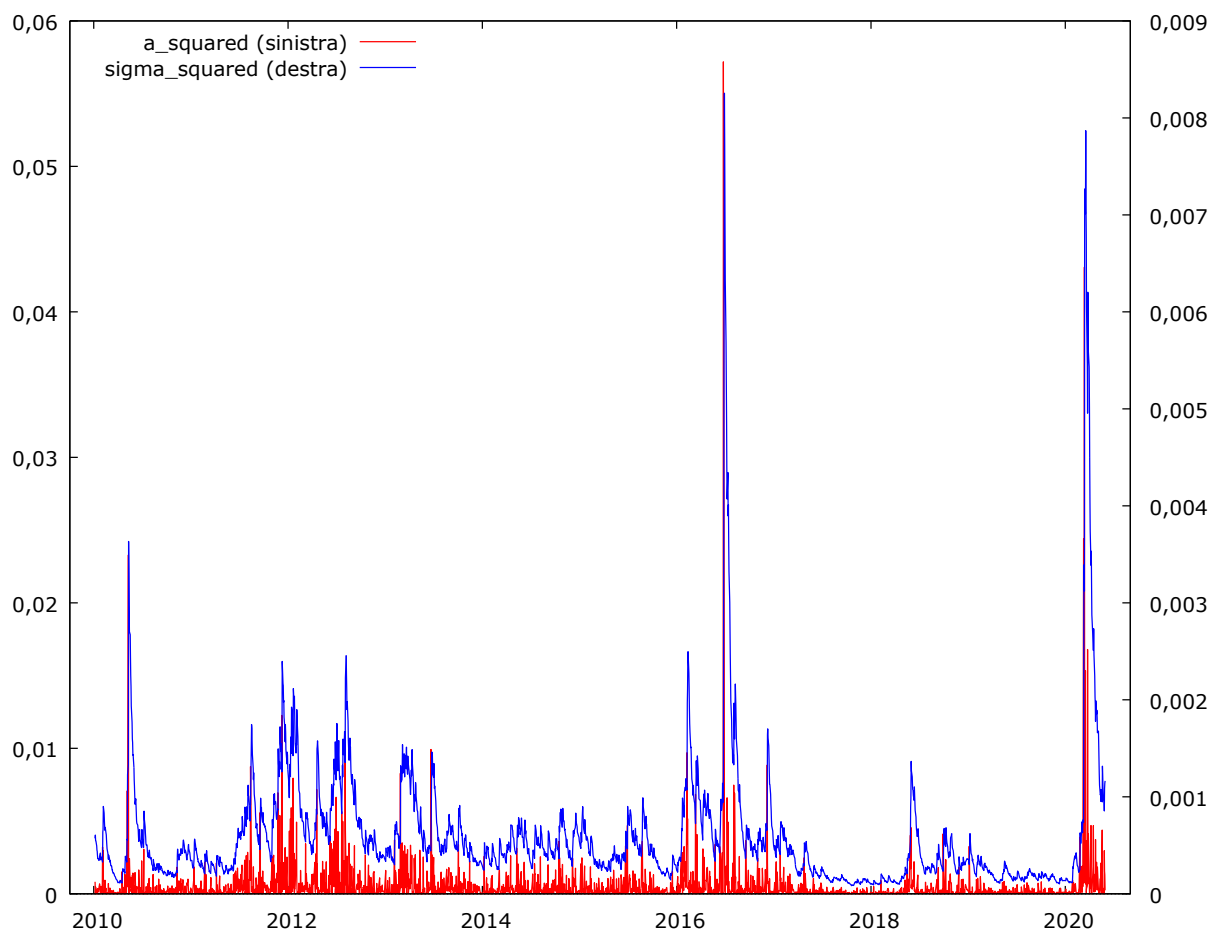


Figura 11: Confronto tra la serie storica della volatilità stimata al quadrato (in rosso) e delle innovazioni al quadrato (in blu).

Conduciamo una semplice regressione OLS delle innovazioni al quadrato, a_t^2 , sul quadrato della volatilità, σ_t^2 :

$$a_t^2 = \alpha + \beta \cdot \sigma_t^2 + u_t$$

Se il modello per la varianza condizionata è correttamente specificato e $E(a_t^2) = \sigma_t^2$, segue che, nella popolazione, $\alpha = 0$ e $\beta = 1$.

L' R^2 della regressione verrà interpretato come un semplice indicatore del grado di prevedibilità nel processo di volatilità, e quindi del potenziale significato economico delle previsioni di volatilità.

I coefficienti stimati sono i seguenti:

	<i>Coefficiente</i>	<i>Errore std.</i>	<i>Statistica t</i>	<i>p-value</i>
α	0,000165079	4,66269e-05	3,540	0,0004 ***
β	0,663969	0,0495402	13,40	<2e-016 ***

Perciò la regressione può essere scritta come

$$a_t^2 = 0,000165079 + 0,663969 \cdot \sigma_t^2 + u_t$$

il cui $R^2 = 0,063730$.

L' R^2 è piuttosto basso, al punto da suggerire che la volatilità non sia stata stimata correttamente.

Un R^2 basso, tuttavia, è linea con l'evidenza nella letteratura per altri titoli e campioni.

Sebbene ciò possa far pensare che i modelli di volatilità standard abbiano un valore pratico limitato, dato che, nonostante le stime campionarie significative dei parametri, spiegano poca della variabilità nei rendimenti quadrati *ex post*, è importante riconoscere che non vi è alcuna contraddizione tra buone previsioni di volatilità e scarso potere predittivo per le innovazioni e quindi per i rendimenti giornalieri al quadrato.

Si può dimostrare (Andersen e Bollerslev 1998) che, date le ipotesi del modello, l' R^2 nella popolazione è limitato superiormente da valori relativamente molto bassi e pertanto lo sarà tanto di più nel campione. Infatti, sebbene l'innovazione al quadrato, a_t^2 , fornisca una stima non distorta per la volatilità, si osservano valori molto dispersi a causa del termine di errore idiosincratice, ϵ_t^2 .

Di conseguenza, lo scarso potere predittivo dei modelli sulla volatilità, se valutati da criteri di previsione standard che utilizzano a_t^2 come misura della volatilità *ex post*, è una conseguenza inevitabile del termine ϵ_t .

2.5 Previsioni con *GARCH (1,1)*

Dato il modello *GARCH (1,1)* stimato

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

$$\sigma_t^2 = 4,56068 \cdot 10^{-6} + 0,104188 a_{t-1}^2 + 0,895332 \sigma_{t-1}^2$$

consideriamo come origine delle previsioni l'ultima sessione di borsa presente nel campione, cioè $h = 29/05/2020$, giorno in cui le osservazioni sull'innovazione e sul suo quadrato sono state rispettivamente $a_h = -0,0153025$ e $a_h^2 = 0,000234167$ e la volatilità stimata e il suo quadrato ammontavano a $\sigma_h = 0,03412$ e $\sigma_h^2 = 0,001164$.

La volatilità prevista per il periodo $h + 1$, quindi per il 01/06/2020 è $\sigma_h(1)$ e si ottiene come segue:

$$\sigma_h^2(1) = \alpha_0 + \alpha_1 a_h^2 + \beta_1 \sigma_h^2$$

$$\sigma_h^2(1) = 4,56068 \cdot 10^{-6} + 0,104188 \cdot 0,000234167 + 0,895332 \cdot 0,001164$$

$$\sigma_h^2(1) = 0,001071125$$

$$\sigma_h(1) = 0,03272805$$

La volatilità prevista per il periodo $h + 2$, quindi per il 02/06/2020 è $\sigma_h(2)$ e si ottiene come segue:

$$\sigma_h^2(2) = \alpha_0 + (\alpha_1 + \beta_1) \sigma_h^2(1)$$

$$\sigma_h^2(2) = 4,56068 \cdot 10^{-6} + (0,104188 + 0,895332) \cdot 0,001071125$$

$$\sigma_h^2(2) = 0,001075172$$

$$\sigma_h(2) = 0,03278982$$

La volatilità prevista per il periodo $h + 3$, quindi per il 03/06/2020 è $\sigma_h(3)$ e si ottiene come segue:

$$\sigma_h^2(3) = \alpha_0 + (\alpha_1 + \beta_1) \sigma_h^2(2)$$

$$\sigma_h^2(3) = 4,56068 \cdot 10^{-6} + (0,104188 + 0,895332) \cdot 0,001075172$$

$$\sigma_h^2(3) = 0,001079217$$

$$\sigma_h(3) = 0,03285144$$

2.6 Test di asimmetria: *Sign Bias Test*

Testiamo la presenza dell'effetto *leverage* nel campione in esame attraverso il *Sign Bias Test*.

Stimiamo la regressione

$$\frac{a_t^2}{\sigma_t^2} = \alpha + \beta D_{t-1} + u_t$$

$$D_{t-1} = \begin{cases} 1, & a_{t-1} < 0 \\ 0, & a_{t-1} \geq 0 \end{cases}$$

dove $\frac{a_t^2}{\sigma_t^2}$ è la serie dei residui standardizzati ottenuti a partire dal modello *GARCH (1,1)* e D_{t-i} è una variabile dummy che assume il valore 1 se $a_{t-1} < 0$ e 0 altrimenti.

	<i>Coefficiente</i>	<i>Errore std.</i>	<i>Statistica t</i>	<i>p-value</i>
α	0,924253	0,0543518	17,01	<2e-016 ***
β	0,154255	0,0780873	1,975	0,0483 **

β risulta essere positivo, con un *p-value* associato alla statistica *t* pari a 0,0483.

Con un livello di significatività del 5% possiamo quindi rifiutare l'ipotesi nulla e pronunciarsi per la presenza di asimmetria nella volatilità, mentre non è possibile affermarlo considerando una significatività dell'1%.

2.7 Specificazione modello *E-GARCH*

Specifichiamo il modello *E-GARCH* (1,1):

$$a_t = \sigma_t \epsilon_t$$

$$\ln(\sigma_t^2) = \alpha_0 + \alpha_1 \frac{|a_{t-1}| + \gamma a_{t-1}}{\sigma_{t-1}} + \beta_1 \ln(\sigma_{t-1}^2)$$

Ne stimiamo i parametri:

	<i>Coefficiente</i>	<i>Errore std.</i>	<i>z</i>	<i>p-value</i>
α_0	-0,222666	0,0412647	-5,396	6,81e-08 ***
α_1	0,148405	0,0327544	4,531	5,88e-06 ***
γ	-0,0900795	0,0136023	-6,622	3,54e-011 ***
β_1	0,985858	0,00411460	239,6	<2e-016 ***

I coefficienti stimati sono tutti positivi e significativi all'1%.

Perciò il modello può essere scritto come

$$\ln(\sigma_t^2) = -0,222666 + 0,148405 \frac{|a_{t-1}| + -0,0900795 a_{t-1}}{\sigma_{t-1}} + 0,985858 \cdot \ln(\sigma_{t-1}^2)$$

In particolare notiamo che il coefficiente γ è negativo e significativo, il che conferma l'ipotesi di presenza di asimmetrie a favore degli shock negativi, rispetto a quelli positivi.

2.8 Specificazione modello *T-GARCH*

Specifichiamo il modello *T-GARCH* (1,1):

$$a_t = \sigma_t \epsilon_t$$

$$\sigma_t^2 = \alpha_0 + (\alpha_1 + \gamma D_{t-1}) a_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

$$D_{t-1} = \begin{cases} 1, & a_{t-1} < 0 \\ 0, & a_{t-1} \geq 0 \end{cases}$$

Ne stimiamo i parametri:

	<i>Coefficiente</i>	<i>Errore std.</i>	<i>z</i>	<i>p-value</i>
α_0	$5,78715e-06$	$2,05250e-06$	2,820	0,0048 ***
α_1	0,0801931	0,0176560	4,542	5,57e-06 ***
γ	0,604708	0,145751	4,149	3,34e-05 ***
β_1	0,928808	0,0136316	68,14	<2e-016 ***

I coefficienti stimati sono tutti positivi e significativi all'1%.

Perciò il modello può essere scritto come

$$\sigma_t^2 = 5,78715 \cdot 10^{-6} + (0,0801931 + 0,604708N_{t-1})a_{t-1}^2 + 0,928808\sigma_{t-1}^2$$

$$D_{t-i} = \begin{cases} 1, & a_{t-1} < 0 \\ 0, & a_{t-1} \geq 0 \end{cases}$$

In particolare notiamo che il coefficiente γ è positivo e significativo, il che conferma l'ipotesi di presenza di asimmetrie a favore degli shock negativi, rispetto a quelli positivi.

2.9 Serie storica dei rendimenti mensili

Analizziamo ora la serie dei rendimenti mensili, ottenuta a partire dai prezzi di chiusura aggiustati dell'ultimo giorno del mese in cui la borsa era aperta. Il campione presenta 125 osservazione, da gennaio 2010 a maggio 2020.

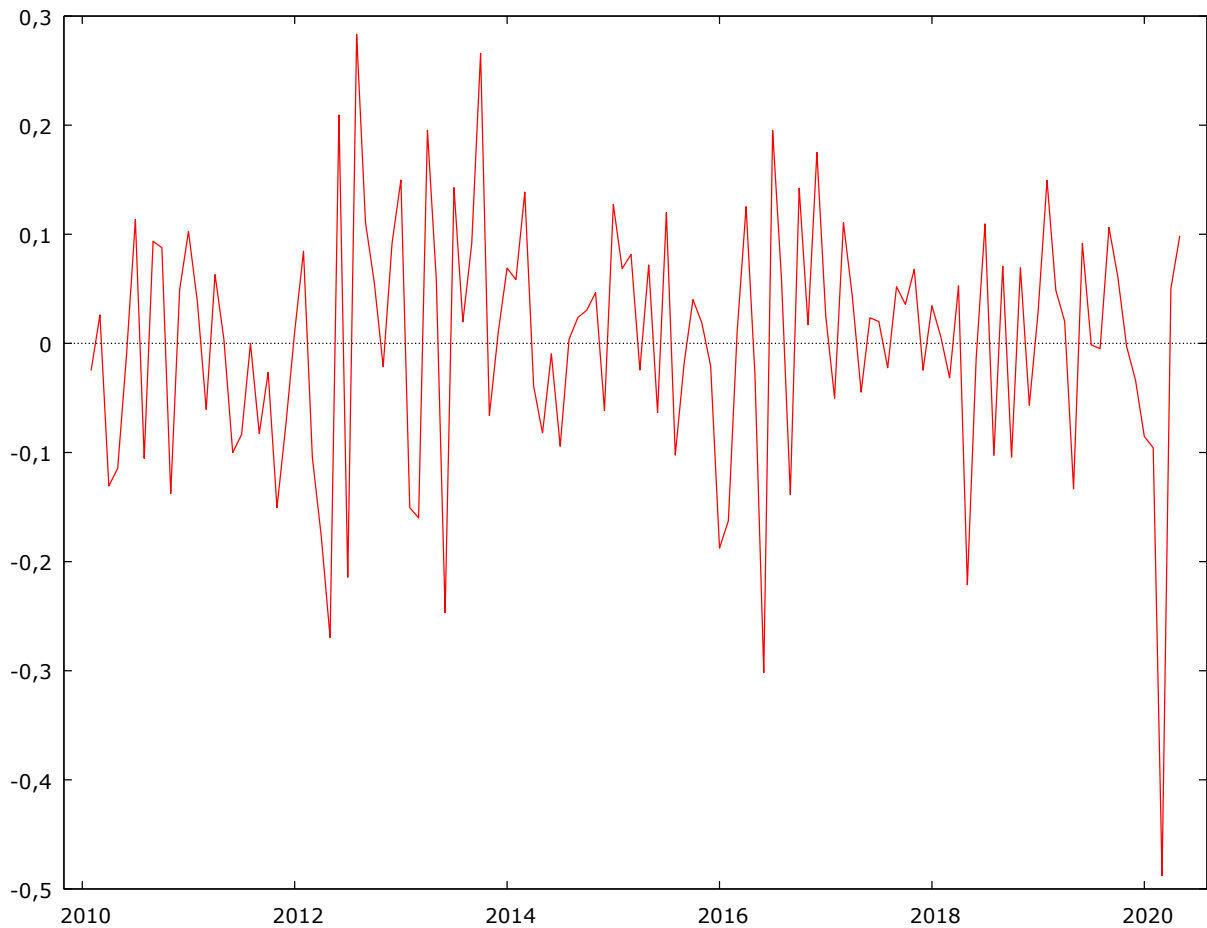


Figura 12: Grafico della serie storica dei rendimenti percentuali mensili del titolo Mediobanca.

Osservando il correlogramma dei rendimenti mensili si nota una correlazione praticamente nulla con ciascuno dei ritardi, il che ci porta a dire che la serie sembra essere stazionaria.

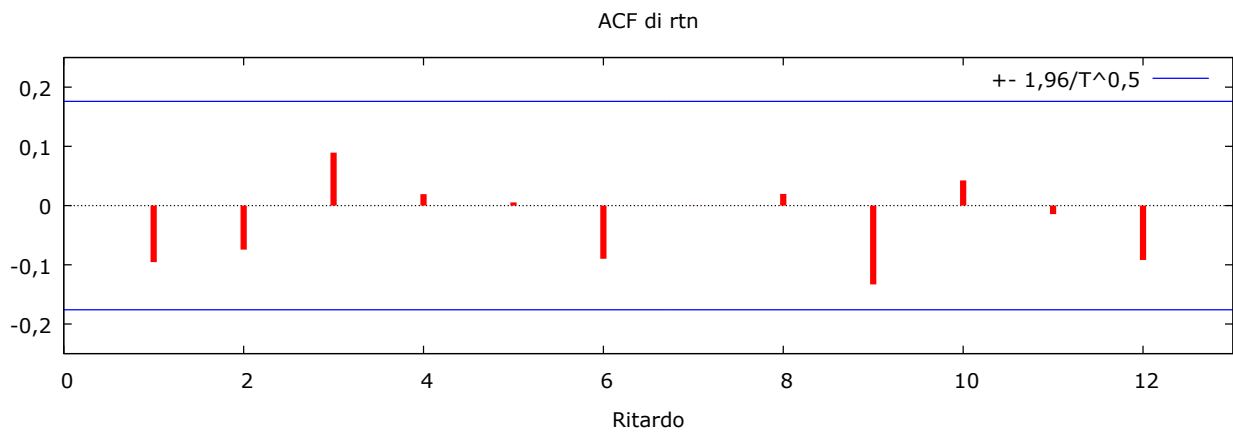


Figura 13: Correlogramma con funzione di autocorrelazione (ACF) della serie storica dei rendimenti percentuali mensili. Vengono considerati i primi 12 ritardi.

Tuttavia nemmeno i quadrati e i valori assoluti dei rendimenti mensili presentano una correlazione significativa con i relativi ritardi e ciò ci impedisce di procedere con la costruzione di modelli per lo studio della volatilità.

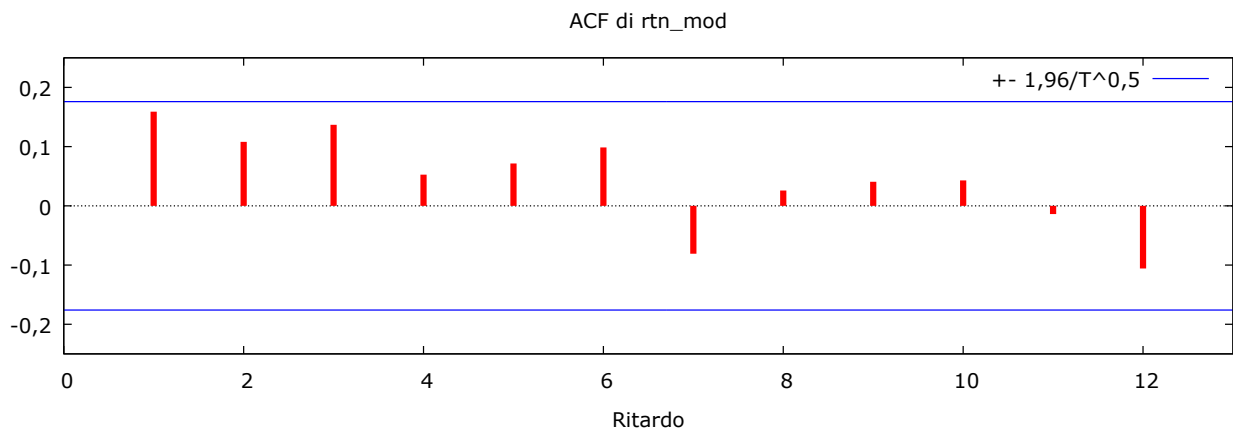


Figura 14: Correlogramma con funzione di autocorrelazione (ACF) della serie storica dei rendimenti percentuali mensili in valore assoluto. Vengono considerati i primi 12 ritardi.

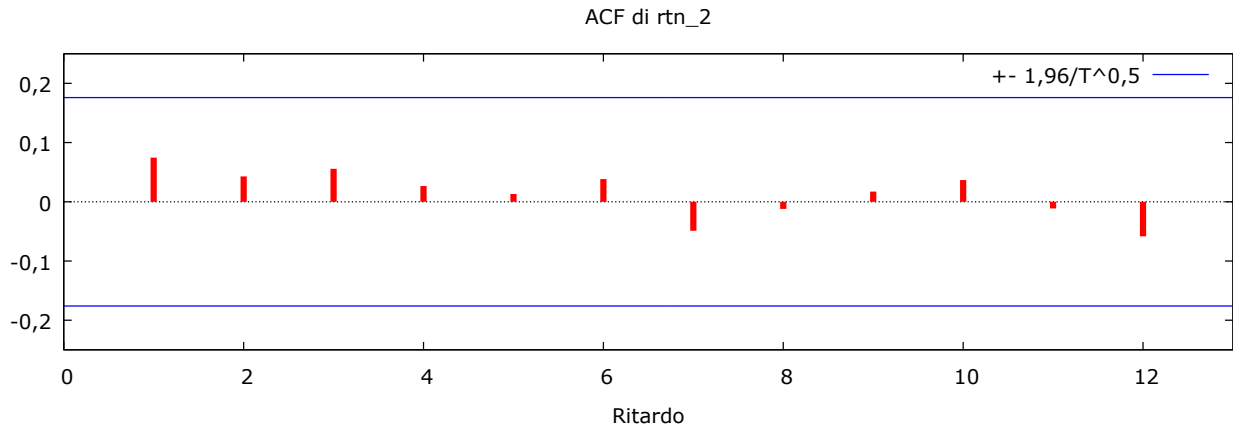


Figura 15: Correlogramma con funzione di autocorrelazione (ACF) della serie storica dei quadrati dei rendimenti percentuali mensili. Vengono considerati i primi 12 ritardi.

Questo risultato tuttavia non ci sorprende.

Osserviamo infatti cosa succede alla varianza condizionata di un processo ARCH se la serie storica può essere osservata in intervalli di tempo più ampi rispetto alla frequenza del processo di generazione dei dati (Kirchgässner e Wolters 2007).

Consideriamo un processo ARCH di ordine 1 ed esprimiamo σ_t^2 scomponendo ricorsivamente i suoi ritardi.

$$\begin{aligned} \sigma_t^2 &= \alpha_0 + \alpha_1 \sigma_{t-1}^2 + v_t \\ \sigma_t^2 &= \alpha_0 + \alpha_1(\alpha_0 + \alpha_1 \sigma_{t-2}^2 + v_{t-1}) + v_t \\ \sigma_t^2 &= \alpha_0 + \alpha_1(\alpha_0 + \alpha_1(\alpha_0 + \alpha_1 \sigma_{t-3}^2 + v_{t-2}) + v_{t-1}) + v_t \end{aligned}$$

L' m -esimo passo della ricorsione può essere dunque scritto come

$$\sigma_t^2 = \alpha_0 \sum_{j=0}^{m-1} \alpha_1^j + \alpha_1^m \sigma_{t-m}^2 + \sum_{j=0}^{m-1} \alpha_1^j v_{t-j}$$

Il valore atteso di σ_t^2 ricavato a partire dalle informazioni disponibili al tempo $t - 1$ è

$$E[\sigma_t^2 | I_{t-1}] = \alpha_0 + \alpha_1 \sigma_{t-1}^2$$

Ora calcoliamo il valore atteso di σ_t^2 supponendo che non tutti i suoi ritardi siano noti, ma che si possano ottenere delle osservazioni solamente ogni 2, 3, ..., m periodi.

$$E[\sigma_t^2 | I_{t-2}] = \alpha_0(1 + \alpha_1) + \alpha_1^2 \sigma_{t-2}^2$$

$$E[\sigma_t^2 | I_{t-3}] = \alpha_0 \frac{1 - \alpha_1^3}{1 - \alpha_1} + \alpha_1^2 \sigma_{t-3}^2$$

$$E[\sigma_t^2 | I_{t-m}] = \alpha_0 \frac{1 - \alpha_1^m}{1 - \alpha_1} + \alpha_1^m \sigma_{t-m}^2$$

Essendo per ipotesi del modello $0 < \alpha_1 < 1$, notiamo che il valore atteso di σ_t^2 cala al crescere del numero di periodi che è necessario aspettare prima di ottenere un'osservazione utile, cioè l'effetto *ARCH* diventa sempre più debole all'aumentare della lunghezza degli intervalli di osservazione.

Per valori molto grandi di m , σ_t^2 tende addirittura ad una costante.

$$\lim_{m \rightarrow \infty} E[\sigma_t^2 | I_{t-m}] = \frac{\alpha_0}{1 - \alpha_1}$$

BIBLIOGRAFIA

- Andersen, Torben G., e Tim Bollerslev. 1998. «Answering the Skeptics: Yes, Standard Volatility Models do Provide Accurate Forecasts». *International Economic Review* 39 (4): 885–905.
- Black, Fischer, e Myron Scholes. 1973. «The Pricing of Options and Corporate Liabilities». *Journal of Political Economy* 81 (3): 637–54.
- Black. 1976. «Studies in Stock Price Volatility Changes». *Proceedings of the 1976 Business Meeting of the Business and Economics Statistics Section*. American Statistical Association: 177-181.
- Bollerslev, Tim. 1986. «Generalized Autoregressive Conditional Heteroskedasticity». *Journal of Econometrics* 31 (3): 307–27.
- Dickey, David A., e Wayne A. Fuller. 1979. «Distribution of the Estimators for Autoregressive Time Series With a Unit Root». *Journal of the American Statistical Association* 74 (366): 427–31.
- Engle, Robert F. 1982. «Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation». *Econometrica* 50 (4): 987–1007.
- Engle, Robert F., e Victor K. Ng. 1993. «Measuring and Testing the Impact of News on Volatility». *The Journal of Finance* 48 (5): 1749–78.
- Glosten, Lawrence R., Ravi Jagannathan, e David E. Runkle. 1993. «On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks». *The Journal of Finance* 48 (5): 1779–1801.
- Kirchgässner, Gebhard, e Jürgen Wolters. 2007. *Introduction to Modern Time Series Analysis*. Berlin, New York: Springer.
- Ljung, G. M., e G. E. P. Box. 1978. «On a Measure of Lack of Fit in Time Series Models». *Biometrika* 65 (2): 297–303.
- Malkiel, Burton G. 2003. «The Efficient Market Hypothesis and Its Critics». *The Journal of Economic Perspectives* 17 (1): 59–82.
- Nelson, Daniel B. 1991. «Conditional Heteroskedasticity in Asset Returns: A New Approach». *Econometrica* 59 (2): 347–70.
- Stock, James H., e Mark W. Watson. 2016. *Introduzione all'econometria*. 4° ed. Milano, Torino: Pearson Italia S.p.a.
- Tsay, Ruey S. 2013. *An Introduction to Analysis of Financial Data with R*. Hoboken, New Jersey: Wiley.
- Zakoian, Jean-Michel. 1994. «Threshold Heteroskedastic Models». *Journal of Economic Dynamics and Control* 18 (5): 931–55.

SITOGRAFIA

<https://www.borsaitaliana.it/>

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

<https://www.mediobanca.com/>