



**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 JPMorgan Chase & Co."**

**RELATORE:**

**CH.MO PROF. NUNZIO CAPPuccio**

**LAUREANDO: FEDERICO RUGGIU**

**MATRICOLA N. 1136067**

**ANNO ACCADEMICO 2018 - 2019**

## INDICE

### CAPITOLO 1

1.1	Concetto di Serie Storica	4
1.2	Concetto di Processo Stocastico	5
1.2.1	Momenti del Processo Stocastico	6
1.2.2	Concetto di Processo Stazionario	7
1.2.3	Concetto di Processo non Stazionario	8

### CAPITOLO 2

2.1	Funzione di Autocorrelazione e Correlogramma	10
2.2	Grafico Serie Storica ( $\Delta t = 1$ giorno, $\Delta t = 5$ giorni, $\Delta t = 30$ giorni)	11
2.3	Analisi variabile $\ln Y_t$ ( $\Delta t = 1$ giorno, $\Delta t = 5$ giorni, $\Delta t = 30$ giorni)	13
2.4	Analisi variabile $r_t$ ( $\Delta t = 1$ giorno, $\Delta t = 5$ giorni, $\Delta t = 30$ giorni)	16
2.5	Analisi variabile $r_t^2$ ( $\Delta t = 1$ giorno, $\Delta t = 5$ giorni, $\Delta t = 30$ giorni)	19
2.6	Test di Radice Unitaria	22
2.6.1	Test Dickey-Fuller	22
2.6.2	Estensioni del test DF: il test ADF (Augmented DF)	23
2.7	Applicazione Test ADF	24
2.7.1	Test ADF per la variabile $\ln(Y_t)$	24
2.7.2	Test ADF per la variabile $r_t$	25

### CAPITOLO 3

3.1	Modello Autoregressivo (AR)	27
3.1.1	Scelta della lunghezza dei ritardi	28
3.2	Modello a Media Mobile (MA)	29
3.3	Modello Autoregressivo a Media Mobile (ARMA)	30
3.4	Applicazione modello ARMA	30
3.5	Modello Autoregressivo a Eteroschedasticità Condizionata Generalizzato (GARCH)	31
3.6	Applicazione modello GARCH	32

<b>Bibliografia</b>	33
---------------------	----

## **ABSTRACT**

Il presente lavoro ha come oggetto l'analisi della serie storica del titolo JPMorgan Chase & Co. con lo scopo di analizzare l'andamento del prezzo del titolo in diversi intervalli di tempo.

Il lavoro si divide in tre capitoli.

Nel primo capitolo vengono introdotti i concetti fondamentali di serie storica, di processo stocastico e gli argomenti ad essi correlati.

Successivamente si presenta la definizione di funzione di autocorrelazione e la descrizione del correlogramma, necessarie per svolgere la preliminare analisi grafica che verrà poi ulteriormente approfondita mediante l'applicazione dei test econometrici, atti a verificare la presenza di radici unitarie nei periodi della serie storica prese in considerazione.

Infine nel terzo capitolo, alla luce dei risultati emersi dai test, si costruiranno dei modelli econometrici per spiegare l'andamento medio e la volatilità della serie nei diversi intervalli di tempo presi in esame.

## CAPITOLO 1

Nel presente elaborato viene presa in considerazione la serie storica del titolo JPMorgan Chase & Co. con lo scopo di andare ad analizzare come l'andamento del prezzo di tale titolo sia mutato rispetto a tre intervalli di tempo: 1 giorno, 5 giorni, 30 giorni. Prima di descrivere il processo di analisi in dettaglio, introduciamo di seguito diversi concetti basilari indispensabili per la comprensione dell'elaborato.

### 1.1 Concetto di Serie Storica

Kirchgässner (2012) definisce una serie storica come un insieme di osservazioni quantitative disposte in ordine cronologico. Generalmente si suppone che il tempo sia una variabile discreta, ovvero quando la variabile assume soltanto un insieme discreto di numeri reali.

Come definito da Warren M. Pearson (1919), le serie storiche sono composte da un numero finito di componenti indipendenti, non direttamente osservabili, che hanno un impatto sulla variabile oggetto di studio. In particolare, Pearson identifica quattro componenti:

- i) Trend (T), componente che misura il movimento tendenziale di lungo periodo. Tale movimento è determinato prevalentemente da fenomeni quali lo sviluppo e l'evoluzione strutturale del sistema economico, che per loro natura si manifestano gradualmente nel tempo;
- ii) Business cycle (C), componente che descrive i cambiamenti nella serie storica con periodo superiore a un anno. Descrive le oscillazioni di lungo periodo rispetto al trend generate da espansioni o contrazioni del sistema economico;
- iii) Seasonal cycle (S), componente che registra le variazioni della variabile in un periodo inferiore a un anno. Essa descrive le fluttuazioni del fenomeno che tendono a ripetersi nei medesimi periodi nel corso degli anni seguenti (Di Fonzo, Lisi, 2007).
- iv) Residual ( $\epsilon$ ), componente di disturbo che registra tutti i movimenti che non appartengono alle componenti precedenti. Questa è la componente che genera irregolarità nelle serie.

Al fine di analizzare le serie storiche per rilevare le regolarità nelle osservazioni di una variabile, per ricavare da esse delle leggi che regolano l'andamento dei dati osservati, oppure per sfruttare tutte le informazioni incluse nella variabile per prevedere meglio gli sviluppi futuri, è opportuno far riferimento a due diversi approcci di analisi.

Originariamente le serie storiche venivano studiate e analizzate mediante un approccio classico in cui si assumeva che le componenti sistematiche, trend, business cycle e seasonal cycle, non erano influenzate da disturbi stocastici, e quindi potevano essere rappresentate come funzioni deterministiche del tempo. L'impatto stocastico era considerato limitato ai soli residuals, che non contenendo nessun movimento sistematico, erano presentati come una serie di variabili casuali indipendenti, o non correlate, con media nulla e varianza costante. Come verrà presentato più avanti, tali componenti vengono definite come white noise, ovvero un processo stocastico che presenta media zero e varianza  $\sigma^2_A$ .

Tuttavia, dagli anni '70, con gli studi di George E.P Box e Gwilm M. Jenkins (1970) ha iniziato a prendere maggiore seguito un nuovo approccio che ha portato a una rivalutazione del ruolo del processo stocastico all'interno delle serie storiche. L'approccio moderno, infatti, presuppone che tutte le componenti di una serie storica siano influenzate da un processo stocastico, e pertanto la serie storica diviene la diretta realizzazione del processo stocastico stesso.

## 1.2 Concetto di Processo Stocastico

Siano dati uno spazio parametrico  $T$ , e uno spazio di probabilità  $(\Omega, F, P)$ , dove  $\Omega$  è lo spazio degli eventi,  $F$  è una  $\sigma$ -algebra di  $\Omega$ , e  $P$  è una misura di probabilità su  $(\Omega; F)$ , che viene definito spazio misurabile. L'insieme  $T$  costituisce un insieme ordinato di numeri reali  $R$ : il processo stocastico è definito discreto se l'insieme dei tempi  $T$  è un sottoinsieme di  $N$  o  $Z$ , mentre assume le caratteristiche di un processo stocastico continuo quando  $T$  è un intervallo di  $R$ . L'insieme dei parametri  $T$  può essere considerato di qualsivoglia natura, nel nostro caso però, dato che stiamo parlando di serie temporali che, come accennato precedentemente, sono rappresentazioni di processi stocastici temporali, verrà interpretato come l'insieme dei tempi del processo stocastico.

Possiamo ora definire un processo stocastico come un insieme di variabili casuali  $\{Y_t\}_{t \in T}$ , indicizzate rispetto a un parametro reale  $t \in T$ , definite su uno stesso spazio di probabilità  $(\Omega, F, P)$ . È possibile indicare un processo stocastico anche con la seguente definizione:

$$\{Y(t, \omega) : t \in T\}, \text{ con } \omega \in \Omega$$

in cui la funzione reale  $Y(t, \omega)$  è costituita rispettivamente da un istante temporale ( $t$ ) e da una eventualità ( $\omega$ ).

In particolare, dato un processo stocastico  $(Y_t)_{t \in T}$ , si ha che:

- i) la variabile aleatoria  $Y_t$ , con  $t \in T$ , è detta stato del processo all'istante  $t$ .
- ii)  $\forall \omega \in \Omega$ , la funzione  $t \rightarrow Y_t(\omega)$  è detta realizzazione o traiettoria del processo associata all'eventualità  $\omega$  e viene denotata con il simbolo  $Y(., \omega)$ , ovvero una particolare successione generata dal processo.

I processi stocastici sono modelli matematici che vengono utilizzati per descrivere la legge probabilistica, o stocastica, con cui un certo fenomeno si evolve nel tempo. Alla luce di quanto detto sinora, quindi, è possibile definire una serie storica  $Y_t$  come una successione finita di valori, ognuno dei quali è la realizzazione di una variabile aleatoria  $Y_t$  appartenente a un processo stocastico  $(Y_t)_{t \in T}$ .

### 1.2.1 Momenti del Processo Statistico

I processi stocastici, data la loro definizione, sono in grado di generare delle serie storiche di lunghezza illimitata. Poiché però non è pensabile di stimare tutte le caratteristiche dei processi stocastici mediante le serie temporali da essi generate, le loro principali proprietà vengono riassunte tramite delle grandezze caratteristiche, i momenti.

In particolare noi ci concentreremo sui momenti primi e secondi considerando le seguenti funzioni:

$$\text{media: } \mu_t = E(Y_t)$$

$$\text{varianza: } \sigma_t^2 = \text{var}(Y_t) = E[(Y_t - \mu)^2]$$

$$\text{covarianza: } \gamma(t, t-k) = \text{cov}(Y_t, Y_{t-k}) = E[(Y_t - \mu_t)(Y_{t-k} - \mu_{t-k})]$$

### 1.2.2 Concetto di Processo Stocastico Stazionario

Dagli studi di Faliva e Zoia (1999) un processo stocastico  $(Y_t)_{t \in T}$  si dice stazionario, rispetto a determinate caratteristiche o elementi del pertinente spazio parametrico, se dette caratteristiche risultano essere invarianti rispetto a qualsivoglia traslazione dell'asse dei tempi. Questa peculiare caratteristica conferisce un carattere di omogeneità, permanenza e stabilità al fenomeno rappresentato dal processo.

Il processo stocastico si dice stazionario in senso stretto (o forte) qualora la distribuzione delle v.c.  $(Y_{t_1}, Y_{t_2}, \dots, Y_{t_k})$  non sia funzione di  $(t_1, t_2, \dots, t_k)$  per ogni  $k \geq 1$ .

Formalmente:

$$(Y_{t_1}, Y_{t_2}, \dots, Y_{t_k}) = (Y_{t_1+h}, Y_{t_2+h}, \dots, Y_{t_k+h}), \forall (t_1, t_2, \dots, t_k) \text{ e } \forall h.$$

L'ipotesi di stazionarietà in senso stretto è una condizione ideale e, nella maggior parte dei casi, quasi impossibile da soddisfare. Pertanto per l'analisi delle serie storiche si preferisce fare ricorso ad una condizione meno limitativa, ossia alla stazionarietà del secondo ordine o stazionarietà in senso lato, riconducendo la verifica alle proprietà di media, varianza e covarianza.

In particolare, un processo stocastico si dice stazionario in senso lato (o debole) se sono soddisfatte le seguenti condizioni:

i)  $E(Y_t) = \mu$ ,

Il valor medio del processo è costante e pari a  $\mu$  al variare di  $t$ ;

ii)  $E(Y_t - \mu)^2 = \sigma^2 < +\infty$

Il processo ha varianza  $\sigma^2$  finita e costante al variare di  $t$ ;

iii)  $E[(Y_t - \mu_t)(Y_{t-k} - \mu_{t-k})] = \gamma(t, t-k) = \text{cov}(Y_t, Y_{t-k})$

La funzione di covarianza non dipende dai due istanti temporali  $t$  e  $t-k$  in cui viene calcolata, ma solo dal loro ritardo (lag):  $k$ .

Un esempio di processo stocastico stazionario è il cosiddetto processo di rumore bianco (white noise). Tale processo è costituito da variabili casuali non sistematiche, omoschedastiche ed incorrelate; è caratterizzato dal valore medio nullo e varianza costante, ovvero non dipendente da  $t$ :

$$E(\varepsilon_t) = 0, \forall t$$

$$E(\varepsilon_t^2) = \text{Var}(\varepsilon_t) = \sigma^2$$

$$\text{Cov}(\varepsilon_t, \varepsilon_{t-k}) = E(\varepsilon_t, \varepsilon_{t-k}) = \begin{cases} \sigma^2, & k = 0 \\ 0, & k \neq 0 \end{cases}$$

e viene indicato con la seguente notazione:

$$\varepsilon_t \sim \text{WN}(0; \sigma^2)$$

dove WN è l'acronimo del termine inglese white noise.

### 1.2.3 Concetto di Processo non Stazionario

Nel paragrafo precedente abbiamo introdotto il concetto di stazionarietà e delle sue implicazioni. Tuttavia, in economia, ma non solo, è piuttosto comune la presenza di serie originate da un processo generatore dei dati non stazionario, e ciò comporta l'impossibilità di utilizzare i dati storici per fare previsioni future attendibili.

Come viene definito, e spiegato nel corso del paragrafo, da Stock e Watson (2016) la caratteristica di "non stazionarietà" di una serie storica è data dalla presenza nella serie di un trend e dalla presenza di rotture strutturali (break).

Per trend si intende un costante movimento tendenziale di lungo periodo di una variabile nel corso del tempo. In particolare nelle serie storiche si possono osservare due tipi di trend: un trend deterministico e uno stocastico.

Un trend deterministico è definito da una funzione deterministica, o non aleatoria, del tempo.

In generale, una serie temporale  $Y_t$  costituita da tale trend viene presentata come:

$$Y_t = f(t) + \varepsilon_t$$

dove  $f(t)$  è perfettamente prevedibile, mentre la componente di errore  $\varepsilon_t$  viene rappresentata come un processo stocastico stazionario a media nulla (ad es.  $\varepsilon_t \sim WN(0, \sigma^2)$ ).

Un trend stocastico, invece, è descritto da una funzione aleatoria che varia nel tempo in maniera causale, e quindi non prevedibile. È caratterizzato dalla presenza contestuale di un trend deterministico ( $Y_t$ ) e di un trend stocastico ( $\varepsilon_t$ ).

Il modello di trend stocastico più semplice è la random walk (passeggiata aleatoria):

$$Y_t = Y_{t-1} + \varepsilon_t$$

dove  $\varepsilon_t$  è indipendentemente ed identicamente distribuita.

Se  $Y_t$  segue una passeggiata aleatoria, allora il valore di  $Y$  domani è il valore di  $Y$  oggi più un disturbo imprevedibile  $\varepsilon_t$ . Se  $Y_t$  segue il modello di passeggiata aleatoria, allora:

$$Y_{t-1} = Y_{t-2} + \varepsilon_{t-1}$$

Sostituendo ancora  $Y_{t-2}, Y_{t-3}, \dots$ , si ottiene:

$$Y_t = \varepsilon_t + \varepsilon_{t-1} + \varepsilon_{t-2} + \varepsilon_{t-3} + \dots + \varepsilon_{t-k}$$

Data questa scrittura, possiamo dedurre che la media di questo processo è

$$E(Y_t) = 0$$

Di conseguenza, poiché  $\varepsilon_t$  è serialmente incorrelato:

$$\text{var}(Y_t) = \text{var}(\varepsilon_t + \varepsilon_{t-1} + \varepsilon_{t-2} + \varepsilon_{t-3} + \dots + \varepsilon_{t-k}) = t\sigma_\varepsilon^2$$

Perciò poiché la varianza di  $Y_t$  dipende da  $t$ , precisamente aumenta al crescere di  $t$ , la sua distribuzione, come definito precedentemente, sarà detta non stazionaria.

Un seconda causa di non stazionarietà si ha quando la funzione di regressione muta nel corso del tempo, cioè quando si verificano delle rotture strutturali (break). In particolare, i coefficienti della serie storica cambino nel corso del tempo a causa di continui cambiamenti nella struttura dell'economia oppure da un'evoluzione della politica economica, e questo mutamento non fornisce una base solida e attendibile per la previsione delle serie storica presa in esame.

## CAPITOLO 2

Ricordiamo che in questo elaborato ci occuperemo dell'analisi della serie storica del prezzo del titolo JPMorgan Chase & Co. con l'obiettivo di individuare un opportuno processo stocastico per comprendere in che modo i fenomeni si evolvono nel tempo e poi formulare previsioni future. Per fare ciò è necessario verificare che la serie storica, osservata a diversi intervalli di tempo, abbia un comportamento stazionario. Per effettuare tale controllo possiamo inizialmente svolgere un'analisi grafica, attraverso il correlogramma; in seguito per confermare tali risultati ottenuti graficamente possiamo utilizzare una serie di test per il controllo della presenza o meno di radici unitarie che sono più rigorosi rispetto alla semplice analisi grafica della serie, tali test ricadono sotto il nome di test per le radici unitarie (unit root test).

### 2.1 Funzione di Autocorrelazione e Correlogramma

La funzione di autocorrelazione misura l'associazione dei valori di una serie temporale con i propri valori ritardati, e risulta particolarmente utile per indicare quali sono i valori passati più utili per prevedere i valori futuri. Più precisamente, la funzione di autocorrelazione (ACF) è data dalla seguente relazione:

$$\rho_k = \frac{\gamma(t, t-k)}{\sigma_t \sigma_{t-k}}$$

dove  $\gamma(t, t-k)$  indica la covarianza tra  $Y_t$  e  $Y_{t-k}$ :

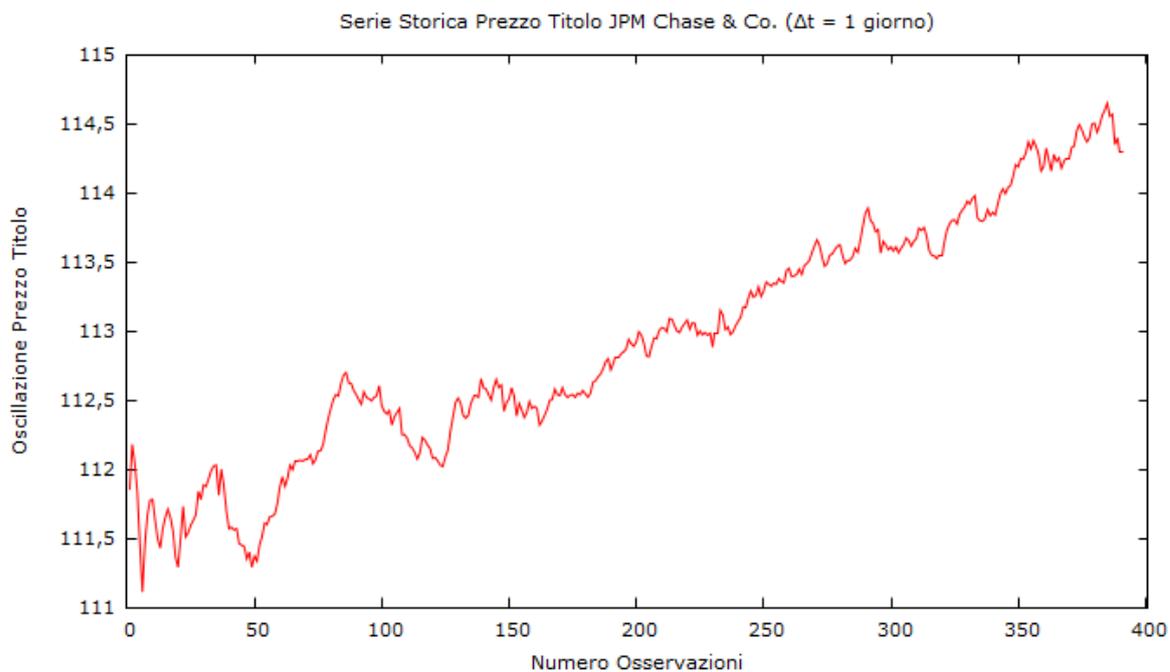
$$\gamma(t, t-k) = \text{cov}(Y_t, Y_{t-k}) = E[(Y_t - \mu_t)(Y_{t-k} - \mu_{t-k})]$$

La funzione di autocorrelazione esprime quindi la dipendenza che esiste tra i valori della serie al tempo  $t$  con i valori ritardati al tempo  $t-k$  della stessa serie storica.

Uno strumento grafico utile per analizzare la correlazione fra i diversi valori assunti dalla serie storica nel corso del tempo è il correlogramma. Esso è un grafico che mostra i valori, tra  $-1$  e  $1$ , che assume il coefficiente di correlazione  $\rho_k$  tra il valore della variabile assunto in  $t$  e i valori in  $t-1, t-2, \dots, t-k$ . Nel grafico ogni barretta verticale rappresenta il valore assunto da  $\rho_k$ , dove  $k$  rappresenta il ritardo dall'istante  $t$  con cui l'autocorrelazione è calcolata.

Utilizzando il software statistico Gretl, di seguito rappresentiamo diversi grafici per studiare l'andamento della serie storica del prezzo del titolo JPMorgan Chase & Co. rispetto ai seguenti intervalli di tempo: 1 giorno, 5 giorni, 30 giorni.

## 2.2 Grafici Serie Storica ( $\Delta t = 1$ giorno, $\Delta t = 5$ giorni, $\Delta t = 30$ giorni)



*Figura 1*

Il grafico rappresentato in *Figura 1* descrive i valori che il prezzo del titolo JPM Chase & Co. ha assunto dalle 9:30 alle 16:00 nella giornata del 17/04/2019 per un totale di 390 osservazioni.

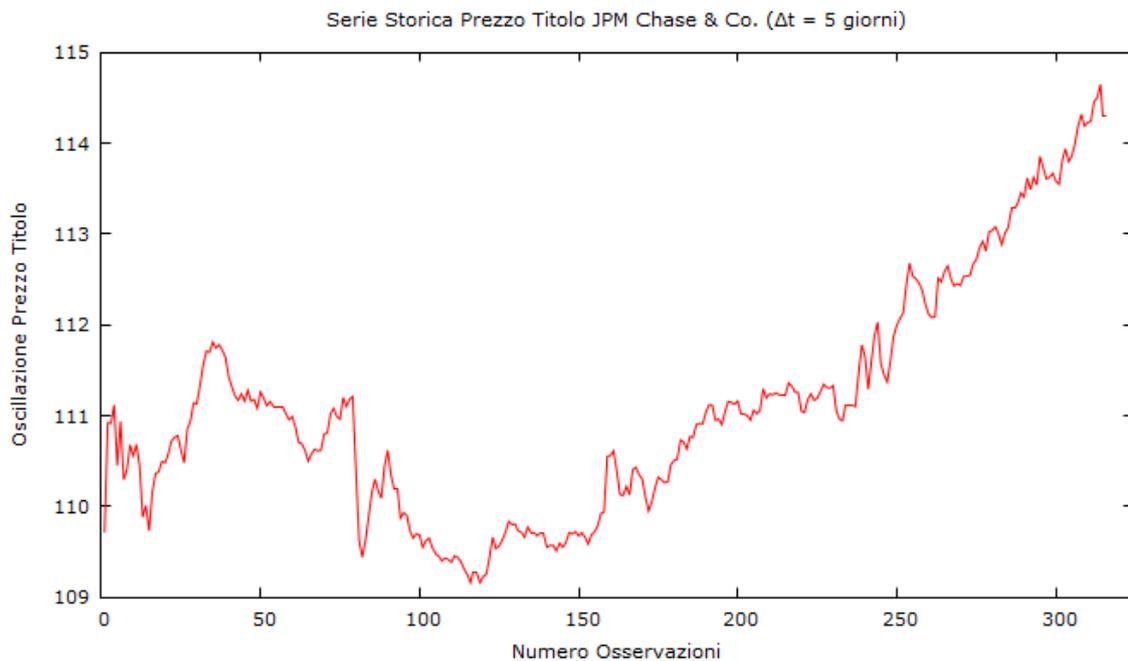


Figura 2

Il grafico rappresentato in *Figura 2* descrive i valori che il prezzo del titolo JPM Chase & Co. ha assunto dalle 9:30 alle 16:30 nelle giornate dal 12/04/2019 al 17/04/2019 per un totale di 316 osservazioni.

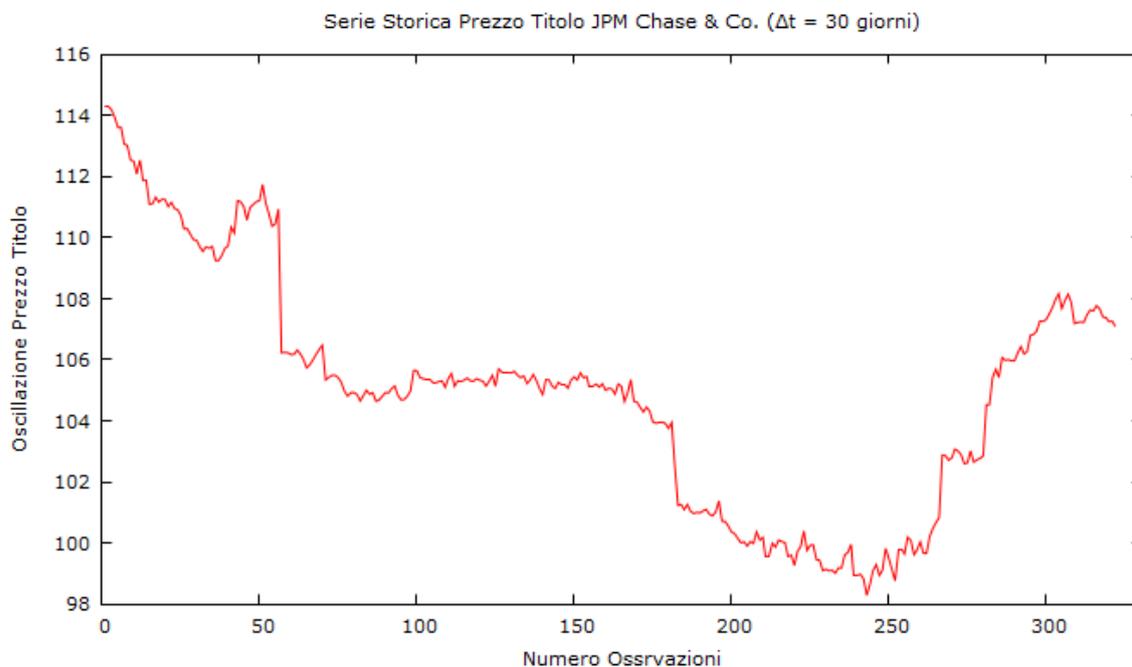


Figura 3

Il grafico rappresentato in *Figura 3* descrive i valori che il prezzo del titolo JPM Chase & Co. ha assunto dalle 10:00 alle 16:30 nelle giornate dal 18/03/2019 al 17/04/2019 per un totale di 323 osservazioni.

### 2.3 Analisi grafica variabile $\ln(Y_t)$

Per una corretta analisi di una serie storica è opportuno calcolare il logaritmo ( $\ln(Y_t)$ ).

Presentiamo di seguito le trasformazioni logaritmiche di  $Y_t$ , rispetto ai tre intervalli temporali:

$\Delta t = 1$  giorno,  $\Delta t = 5$  giorni,  $\Delta t = 30$  giorni.

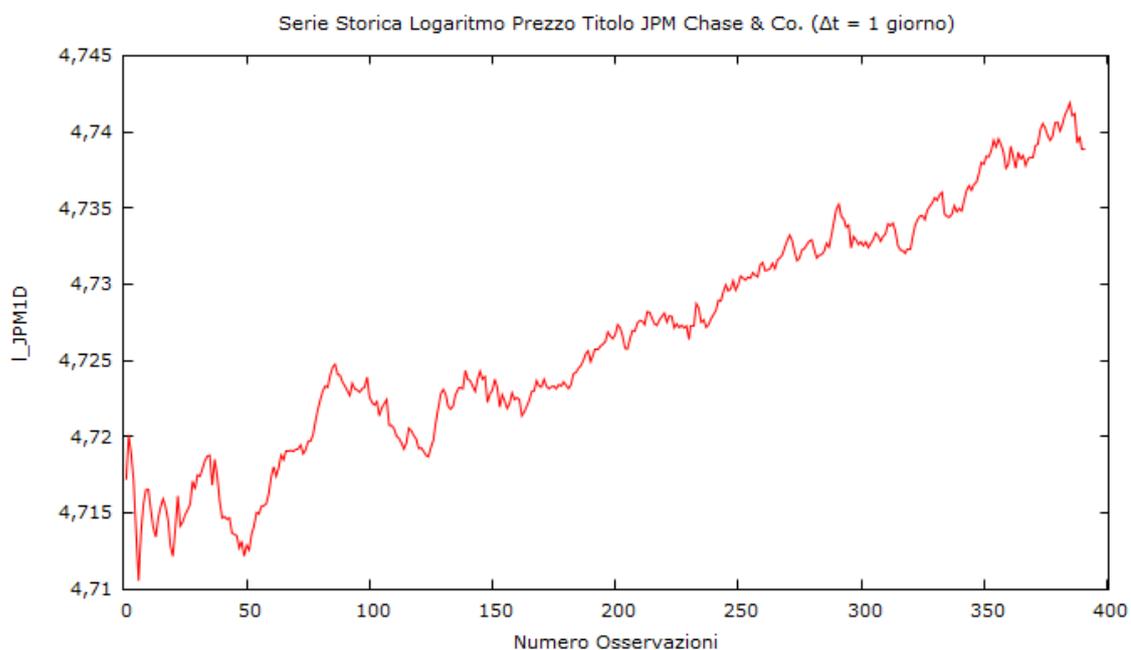


Figura 4

Il grafico rappresentato in *Figura 4* descrive il logaritmo dei valori che il prezzo del titolo JPM Chase & Co. ha assunto dalle 9:30 alle 16:00 nella giornata del 17/04/2019 per un totale di 390 osservazioni.

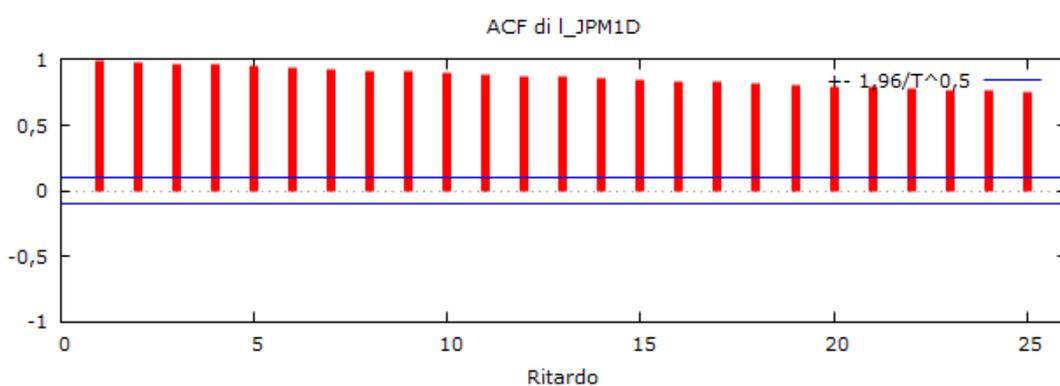


Figura 5

Il correlogramma mostrato in *Figura 5* rappresenta il coefficiente di correlazione  $\rho_k$  tra la variabile  $\ln(Y_t)$ , presentata in *Figura 4*, ed i suoi valori passati considerando un totale di 25 ritardi. Dal grafico si può notare che  $\rho_k$  è sempre positivo e decresce lentamente all'aumentare di  $k$ .

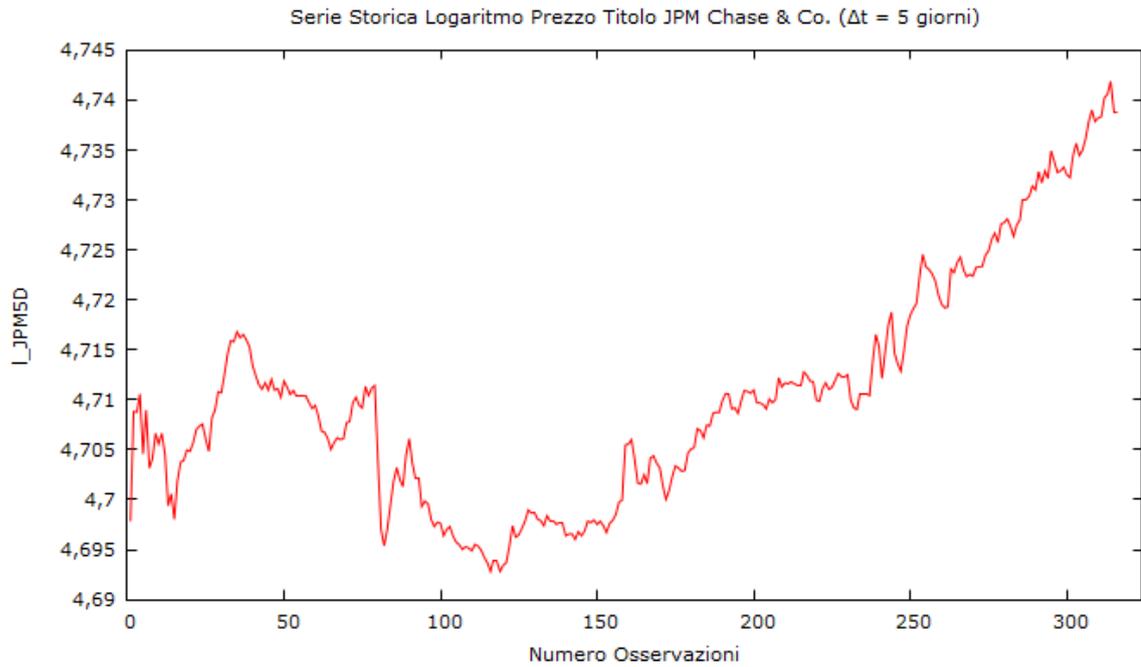


Figura 6

Il grafico rappresentato in *Figura 6* descrive il logaritmo dei valori che il prezzo del titolo JPM Chase & Co. ha assunto dalle 9:30 alle 16:30 nelle giornate dal 12/04/2019 al 17/04/2019 per un totale di 316 osservazioni.

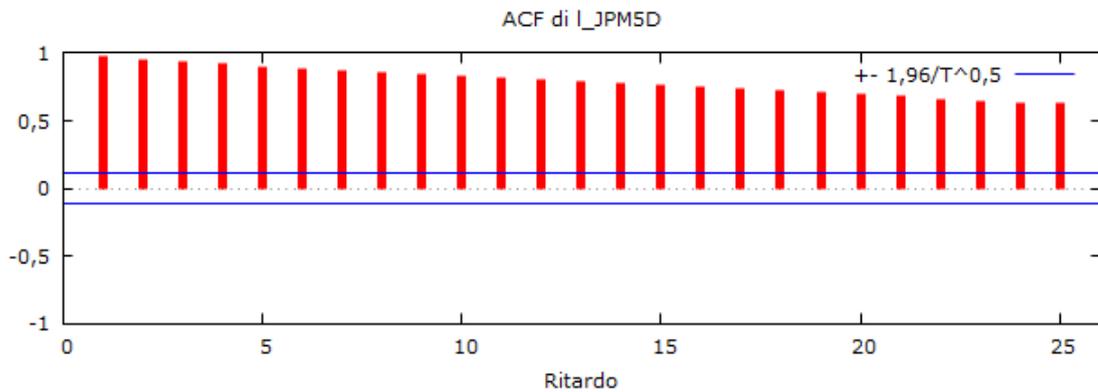


Figura 7

Il correlogramma mostrato in *Figura 7* rappresenta il coefficiente di correlazione  $\rho_k$  tra la variabile  $\ln(Y_t)$ , presentata in *Figura 6*, ed i suoi valori passati considerando un totale di 25 ritardi. Come nel caso precedente, il grafico evidenzia che  $\rho_k$  è sempre positivo e decresce lentamente all'aumentare di  $k$ .

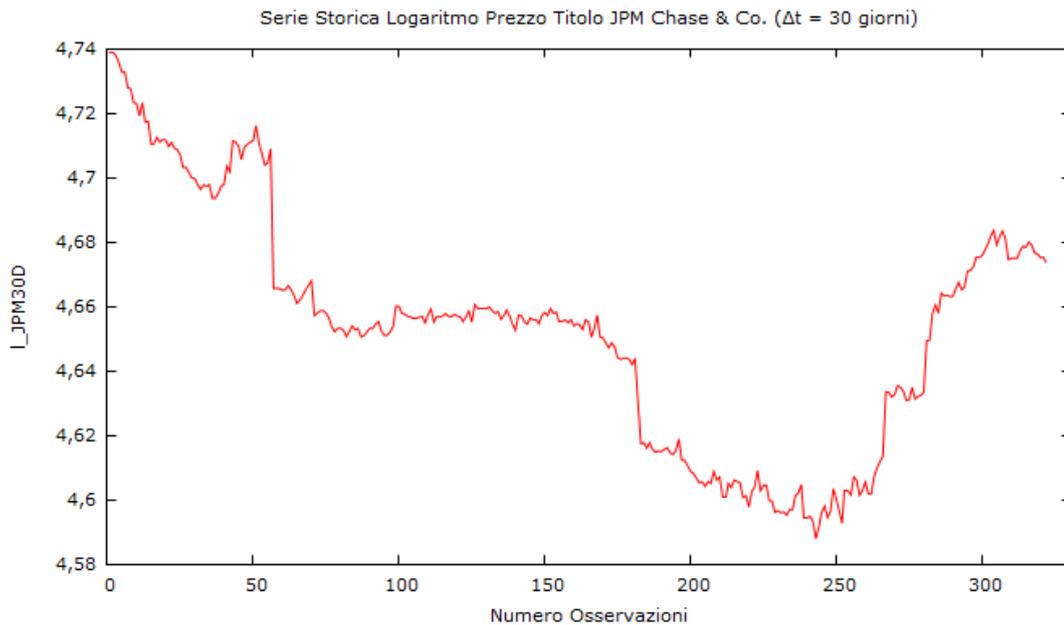


Figura 8

Il grafico rappresentato in *Figura 8* descrive il logaritmo dei valori che il prezzo del titolo JPM Chase & Co. ha assunto dalle 10:00 alle 16:30 nelle giornate dal 18/03/2019 al 17/04/2019 per un totale di 323 osservazioni.

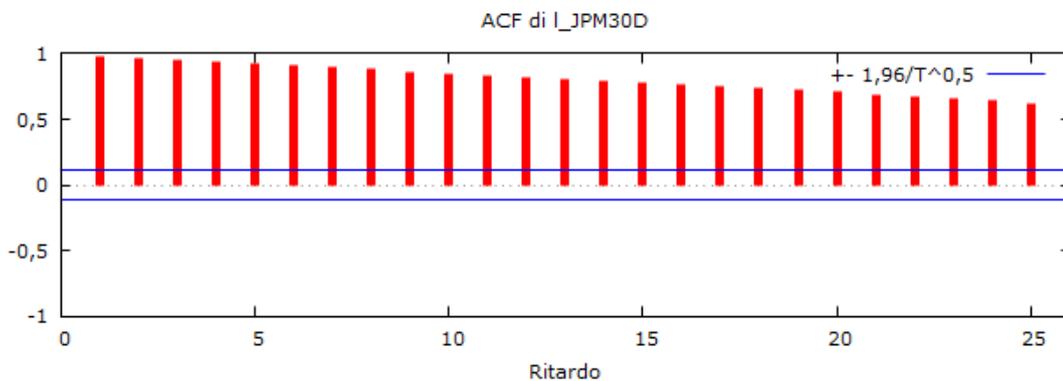


Figura 9

Il correlogramma mostrato in *Figura 9* rappresenta il coefficiente di correlazione  $\rho_k$  tra la variabile  $\ln(Y_t)$ , presentata in *Figura 8*, ed i suoi valori passati considerando un totale di 25 ritardi. Come nel caso precedente, il grafico evidenzia che  $\rho_k$  è sempre positivo e decresce lentamente all'aumentare di  $k$ .

In tutti e tre i grafici, i valori assunti dalla serie storica al tempo  $t$  sono fortemente correlati con il primo ritardo, poi tale correlazione si riduce gradualmente nel tempo, evidenziando nella serie storica la presenza di una componente tendenziale.

## 2.4 Analisi grafica variabile $r_t$ ( $r_t = \ln(Y_t) - \ln(Y_{t-1})$ )

Per analizzare le variazioni del prezzo del titolo nel tempo, introduciamo il concetto di rendimento ( $r_t$ ), definito come la differenza prima del logaritmo ( $\ln(Y_t) - \ln(Y_{t-1})$ ), dei valori della serie storica presi in considerazione. Presentiamo di seguito i rendimenti di  $Y_t$ , rispetto ai tre intervalli temporali:  $\Delta t = 1$  giorno,  $\Delta t = 5$  giorni,  $\Delta t = 30$  giorni.

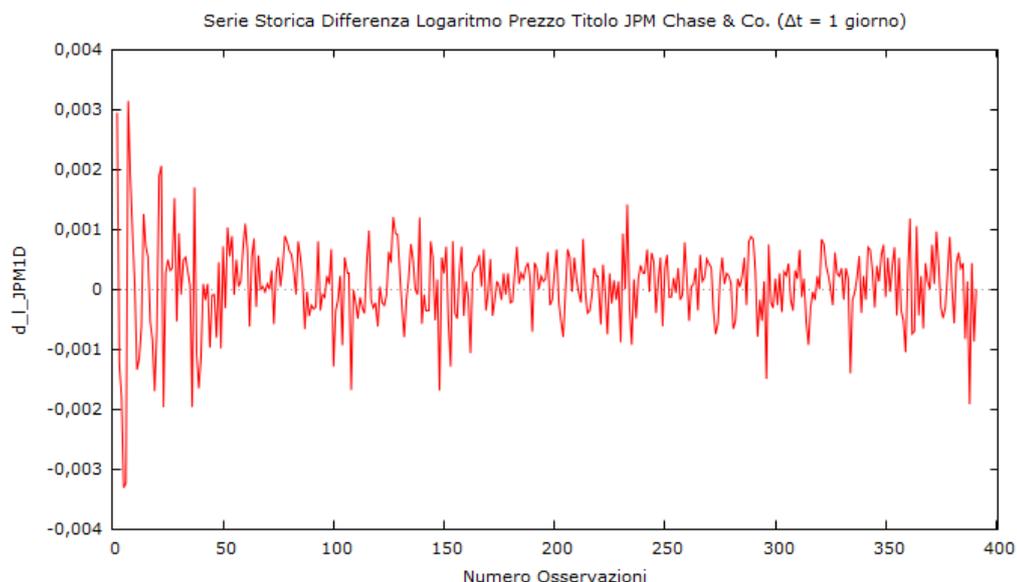


Figura 10

Il grafico rappresentato in *Figura 10* descrive la differenza prima del logaritmo dei valori che il prezzo del titolo JPM Chase & Co. ha assunto dalle 9:30 alle 16:00 nella giornata del 17/04/2019 per un totale di 390 osservazioni.

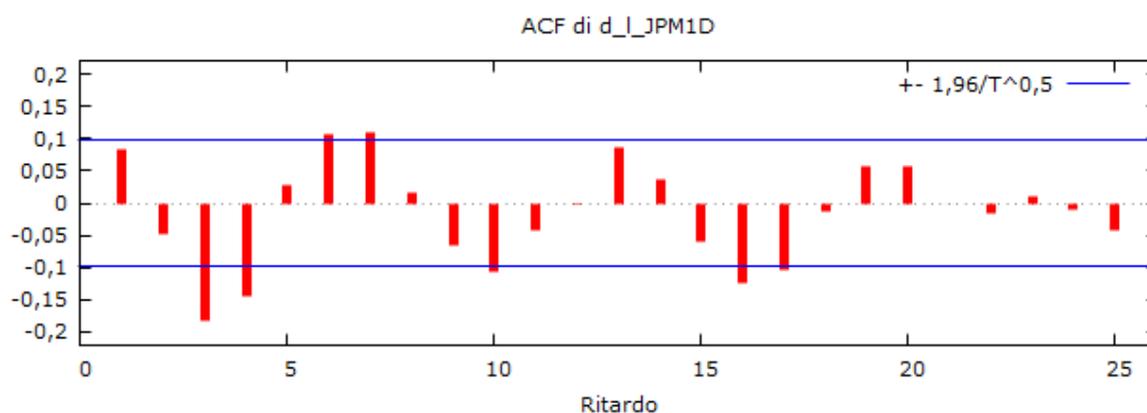


Figura 11

Il correlogramma mostrato in *Figura 11* rappresenta il coefficiente di correlazione  $\rho_k$  tra la variabile  $r_t(Y_t)$ , presentata in *Figura 10*, ed i suoi valori passati considerando un totale di 25

ritardi. In questo caso, il grafico evidenzia che i valori assunti dalla serie storica sono debolmente correlati fra di loro.

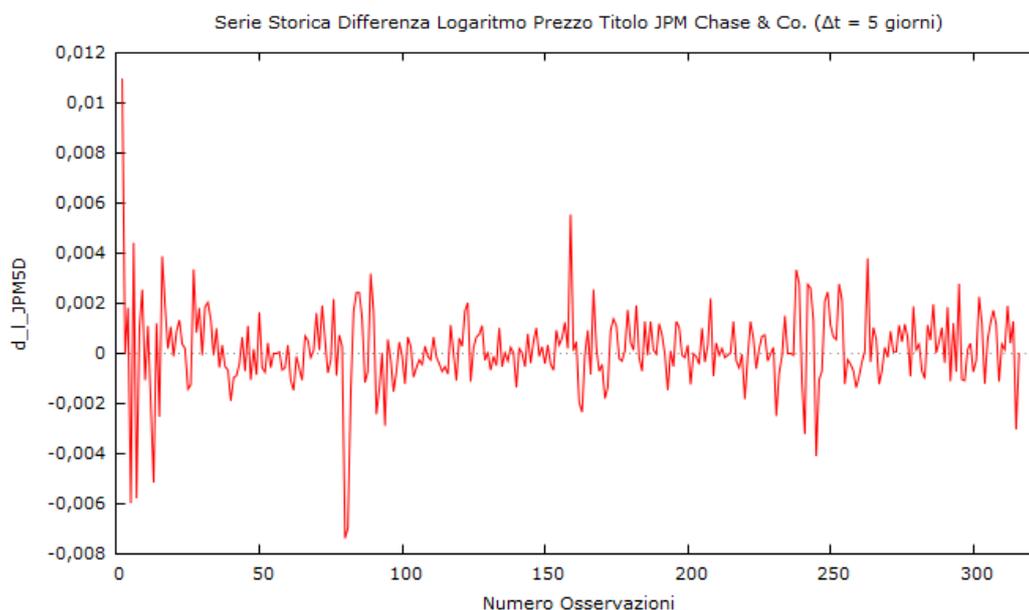


Figura 12

Il grafico rappresentato in *Figura 12* descrive la differenza prima del logaritmo dei valori che il prezzo del titolo JPM Chase & Co. ha assunto dalle 9:30 alle 16:00 nelle giornate dal 12/04/2019 al 17/04/2019 per un totale di 316 osservazioni.

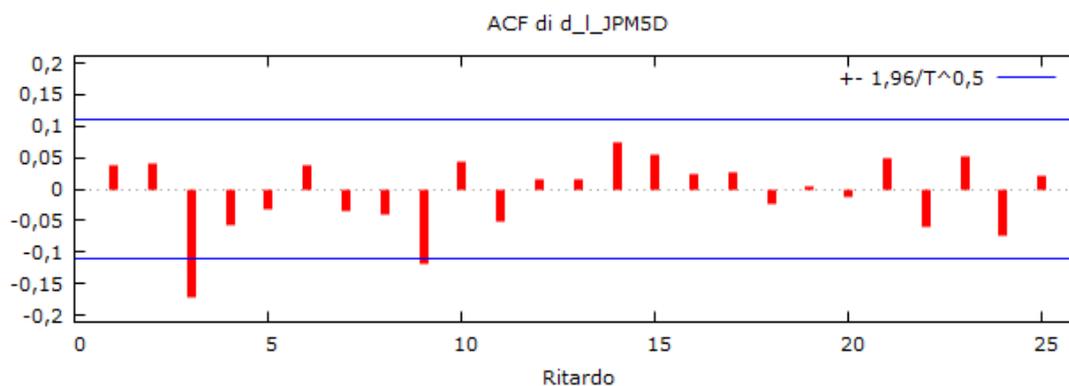


Figura 13

Il correlogramma mostrato in *Figura 12* rappresenta il coefficiente di correlazione  $\rho_k$  tra la variabile  $r_t(Y_t)$ , presentata in *Figura 11*, ed i suoi valori passati considerando un totale di 25 ritardi. Come nel caso precedente, il grafico evidenzia che i valori assunti dalla serie storica sono debolmente correlati fra di loro.

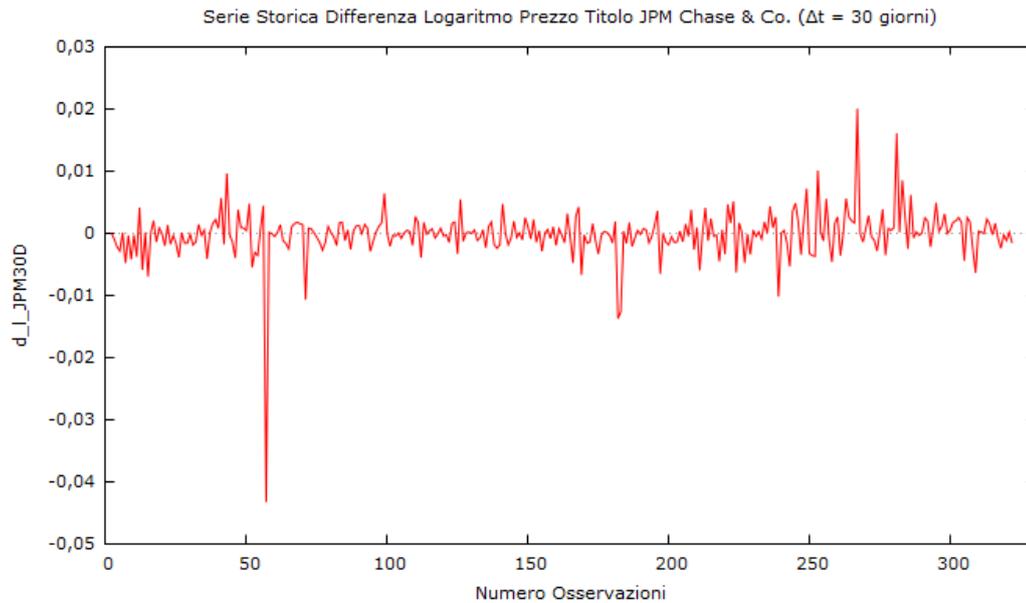


Figura 14

Il grafico rappresentato in *Figura 14* descrive la differenza prima del logaritmo dei valori che il prezzo del titolo JPM Chase & Co. ha assunto dalle 10:00 alle 16:30 nelle giornate dal 18/03/2019 al 17/04/2019 per un totale di 323 osservazioni.

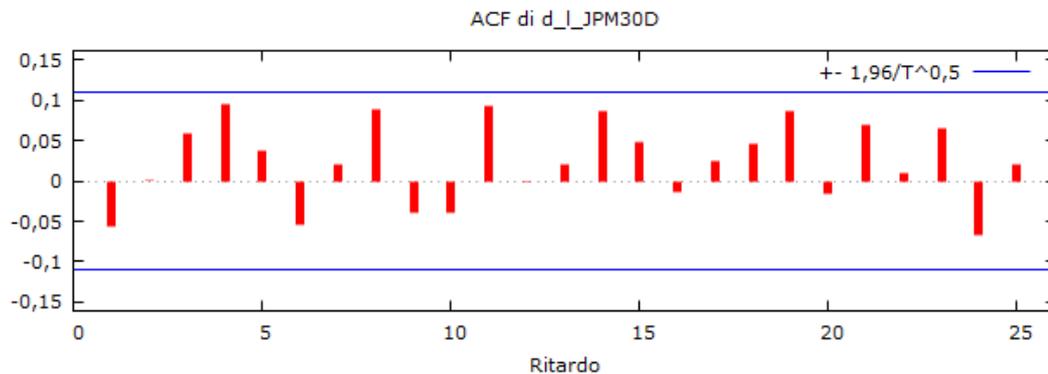


Figura 15

Il correlogramma mostrato in *Figura 15* rappresenta il coefficiente di correlazione  $\rho_k$  tra la variabile  $r_t(Y_t)$ , presentata in *Figura 14*, ed i suoi valori passati considerando un totale di 25 ritardi. Come nei casi precedente, anche questa volta il grafico evidenzia che i valori assunti dalla serie storica sono debolmente correlati fra di loro.

## 2.5 Analisi grafica variabile $r_t^2$ ( $r_t^2 = (\ln(Y_t) - \ln(Y_{t-1}))^2$ )

Per rappresentare l'andamento della volatilità all'interno della serie storica presentiamo di seguito i quadrati dei rendimenti  $r_t^2 = (\ln(Y_t) - \ln(Y_{t-1}))^2$  di  $Y_t$ , rispetto ai tre intervalli temporali:  $\Delta t = 1$  giorno,  $\Delta t = 5$  giorni,  $\Delta t = 30$  giorni.

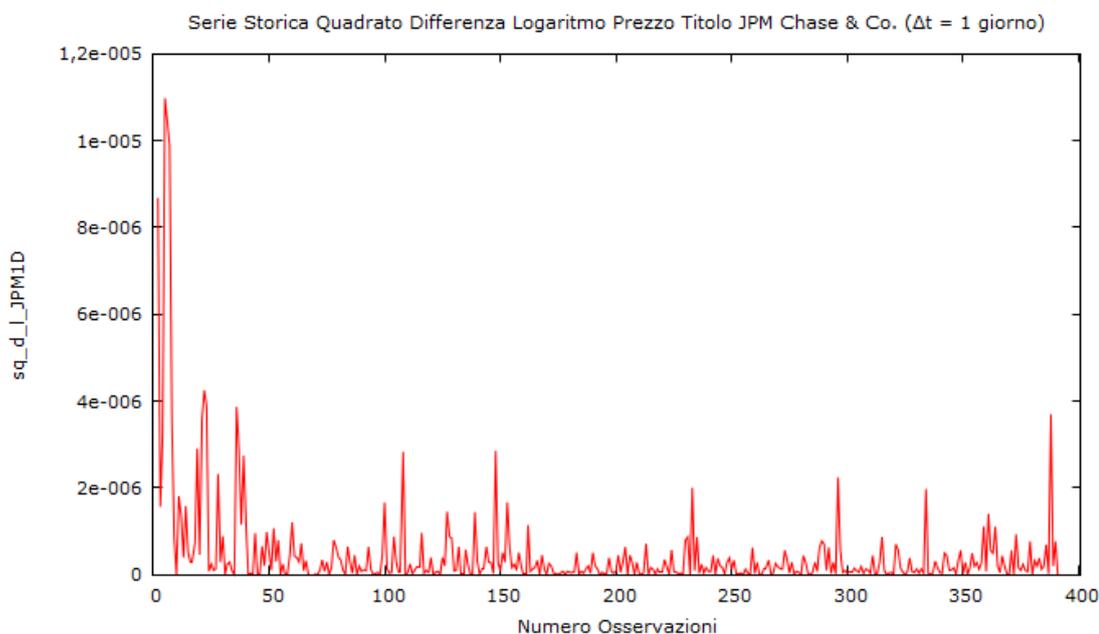


Figura 16

Il grafico rappresentato in *Figura 16* descrive il quadrato della differenza prima del logaritmo dei valori che il prezzo del titolo JPM Chase & Co. ha assunto dalle 9:30 alle 16:00 nella giornata del 17/04/2019 per un totale di 390 osservazioni.

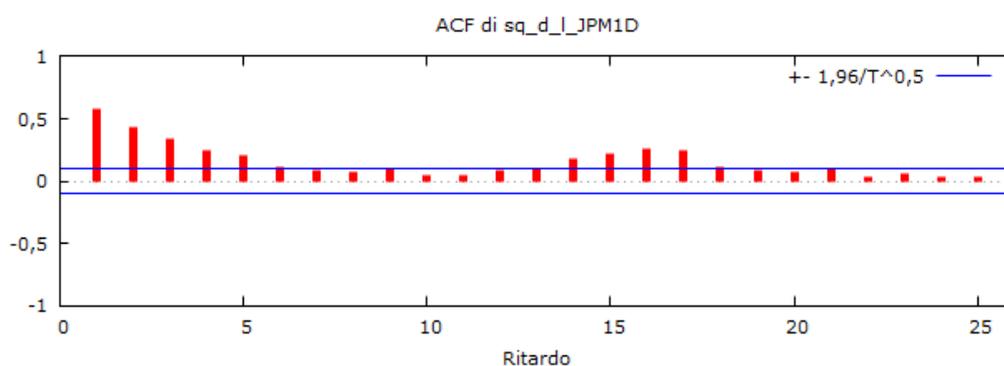


Figura 17

Il correlogramma mostrato in *Figura 17* rappresenta il coefficiente di correlazione  $\rho_k$  tra la variabile  $r_t^2(Y_t)$ , presentata in *Figura 16*, ed i suoi valori passati considerando un totale di 25 ritardi. In questo caso, il grafico evidenzia qualche grado di correlazione con i valori passati.

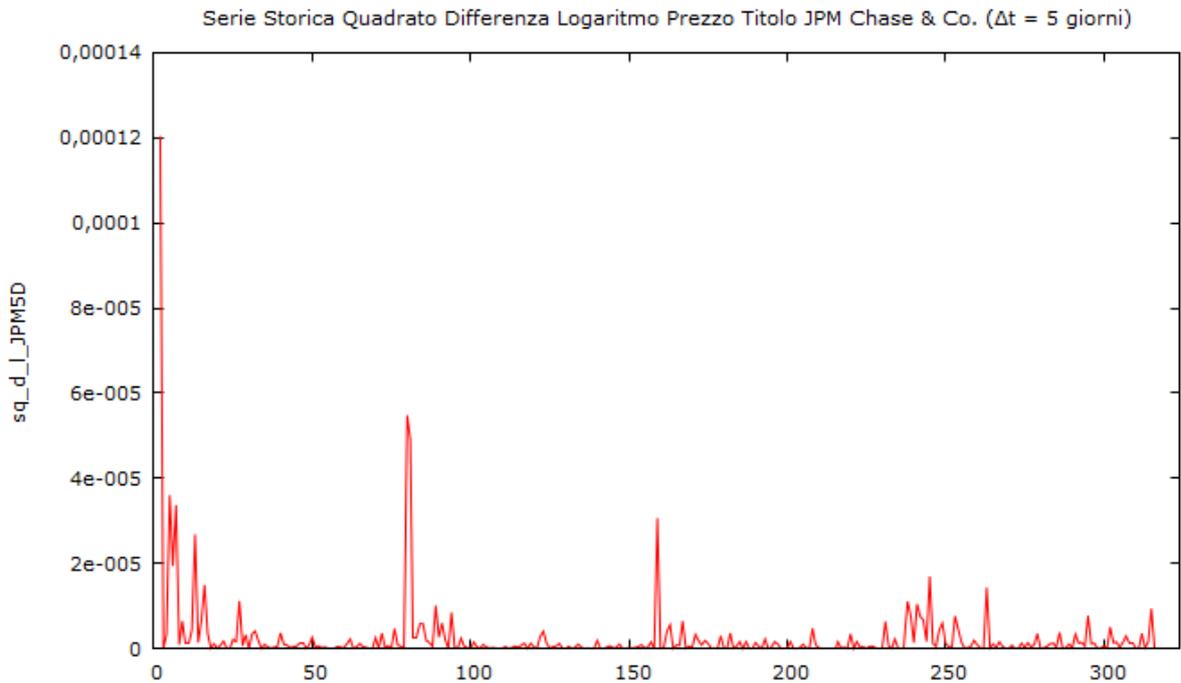


Figura 18

Il grafico rappresentato in *Figura 18* descrive il quadrato della differenza prima del logaritmo dei valori che il prezzo del titolo JPM Chase & Co. ha assunto dalle 9:30 alle 16:00 nelle giornate dal 12/04/2019 al 17/04/2019 per un totale di 316 osservazioni.

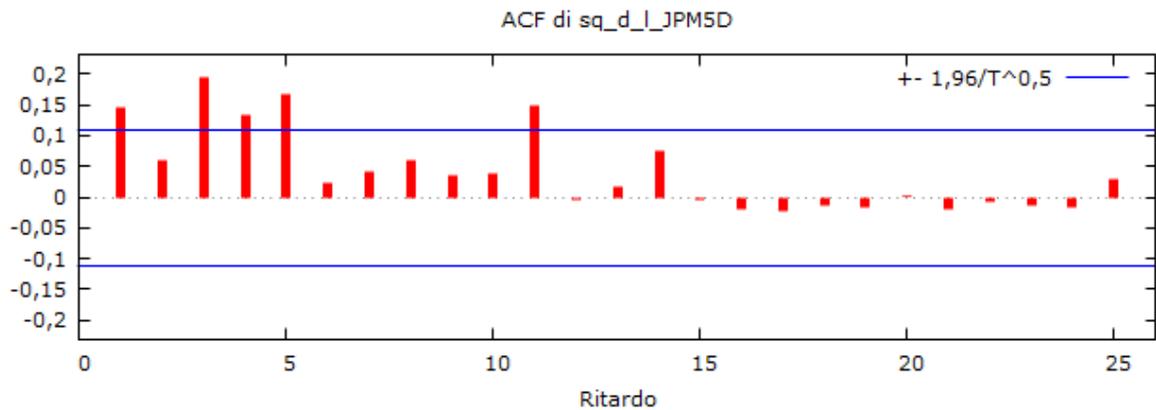


Figura 19

Il correlogramma mostrato in *Figura 19* rappresenta il coefficiente di correlazione  $\rho_k$  tra la variabile  $r_t^2(Y_t)$ , presentata in *Figura 18*, ed i suoi valori passati considerando un totale di 25 ritardi. Come nel caso precedente, il grafico evidenzia qualche grado di correlazione con i valori passati, tale correlazione però diventa più debole all'aumentare dei ritardi.

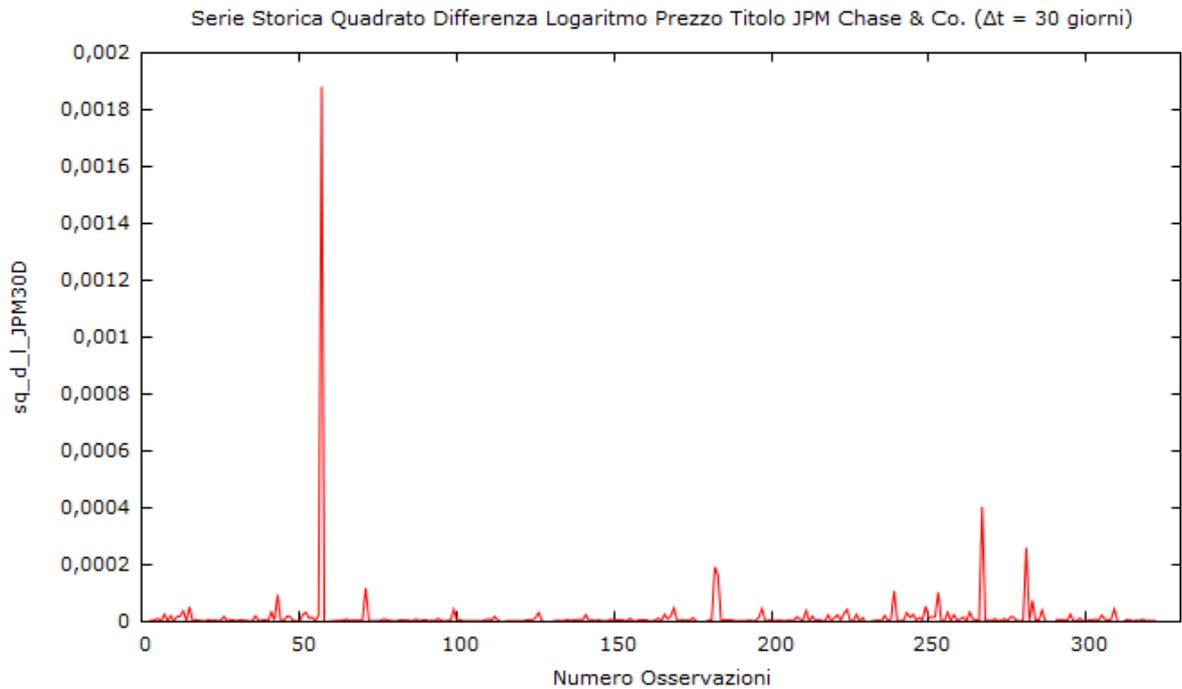


Figura 20

Il grafico rappresentato in *Figura 20* descrive il quadrato della differenza prima del logaritmo dei valori che il prezzo del titolo JPM Chase & Co. ha assunto dalle 10:00 alle 16:30 nelle giornate dal 18/03/2019 al 17/04/2019 per un totale di 323 osservazioni.

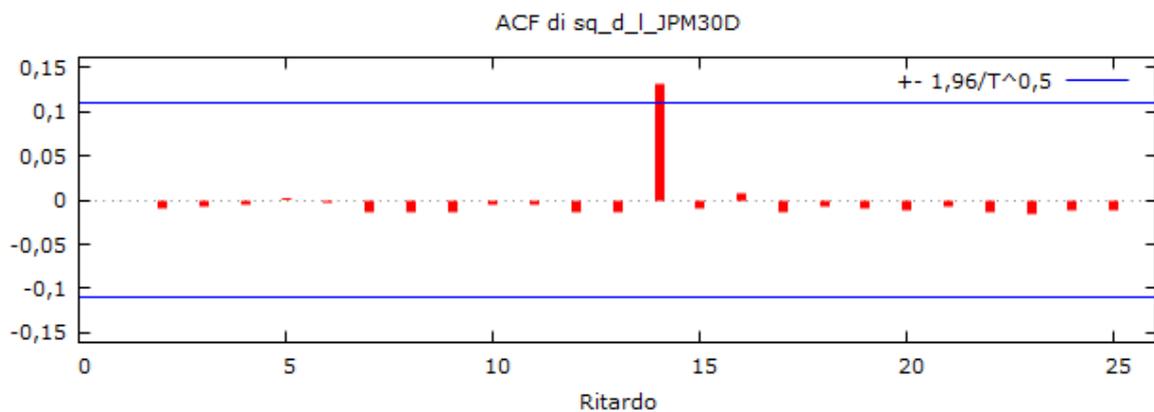


Figura 21

Il correlogramma mostrato in *Figura 21* rappresenta il coefficiente di correlazione  $\rho_k$  tra la variabile  $r_t^2(Y_t)$ , presentata in *Figura 20*, ed i suoi valori passati considerando un totale di 25 ritardi. In questo caso, il grafico evidenzia una correlazione molto debole fra i vari ritardi, con un'unica eccezione al ritardo  $k = 14$ , in cui si registra un'elevata correlazione.

## 2.6 Test di Radice Unitaria

Dalla precedente analisi grafica effettuata tramite il correlogramma si è messo in evidenza un diverso grado di correlazione tra le tre variabili prese in esame ( $\ln(Y_t)$ ,  $r_t$ ,  $r_t^2$ ) con i rispettivi valori ritardati ai differenti intervalli  $i$  tempo ( $\Delta t = 1$  giorno,  $\Delta t = 5$  giorni,  $\Delta t = 30$  giorni). Per verificare i risultati emersi dai correlogrammi delle variabili  $\ln(Y_t)$  e  $r_t$ , verrà svolta un'ulteriore analisi utilizzando il test di radice unitaria DF.

### 2.6.1 Test DF (Dickey-Fuller)

Il test DF (D. Dickey e W. Fuller, 1979) è utilizzato per verificare all'interno serie temporale la presenza di un trend stocastico. Partendo dal modello autoregressivo:

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

dove  $\beta_0$  è chiamata deriva, il test fa riferimento al seguente sistema di ipotesi:

$$\begin{cases} H_0 : \beta_1 = 1 \\ H_1 : \beta_1 < 1 \end{cases}$$

In cui, l'ipotesi nulla  $H_0$  presuppone la presenza di radice unitaria, e quindi la serie storica presenta un comportamento stocastico; mentre l'ipotesi alternativa  $H_1$  invece stabilisce che il modello autoregressivo sia stazionario.

È possibile anche verificare il test sottraendo  $Y_{t-1}$  da entrambi i lati del modello, ottenendo:

$$Y_t - Y_{t-1} = \beta_0 + (\beta_1 - 1)Y_{t-1} + \varepsilon_t$$

indicando con  $\Delta Y_t$  la differenza tra  $Y_t - Y_{t-1}$  e con  $\alpha$  la differenza  $\beta_1 - 1$ , otterremo:

$$\Delta Y_t = \beta_0 + \alpha Y_{t-1} + \varepsilon_t$$

e il nuovo sistema di ipotesi potrà essere riscritto come:

$$\begin{cases} H_0 : \alpha = 0 \\ H_1 : \alpha < 0 \end{cases}$$

La statistica test per la verifica dell'ipotesi nulla  $H_0$ , è detta statistica di Dickey-Fuller:

$$DF = \frac{\hat{\alpha}}{SE(\hat{\alpha})}$$

e viene calcolata utilizzando gli errori standard SE.

Se questa è inferiore al valore critico della distribuzione di Dickey-Fuller sarà rifiutata  $H_0$ .

<b>Regressore deterministico</b>	<b>10%</b>	<b>5%</b>	<b>1%</b>
<b>Intercetta</b>	-2.57	-2.86	-3.43
<b>Intercetta e Trend</b>	-3.12	-3.41	-3.96

Tabella 1

### 2.6.2 Estensioni del test DF: il test ADF (Augmented DF)

Il modello autoregressivo di primo ordine, su cui si basa il test DF appena descritto, non consente di calcolare l'intera correlazione seriale presente in  $Y_t$ , perciò si farà riferimento a test DF basati su modelli autoregressivi di ordine superiore al primo, chiamati test ADF (Augmented Dickey-Fuller).

Il test ADF si basa sul seguente modello:

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

dove  $p$  indica il numero di ritardi presi in considerazione.

Si può studiare anche l'ipotesi  $H_1$  in cui la serie sia stazionaria attorno a un trend deterministico  $\mu_t$ , in questo caso la regressione ADF sarà:

$$\Delta Y_t = \beta_0 + \mu_t + \alpha Y_{t-1} + \gamma_1 \Delta Y_{t-1} + \gamma_2 \Delta Y_{t-2} + \dots + \gamma_p \Delta Y_{t-p} + \varepsilon_t$$

Come nel precedente modello di autoregressione di primo ordine, anche il test ADF verifica l'ipotesi nulla  $H_0 : \alpha = 0$  contro l'alternativa unilaterale  $H_1 : \alpha < 0$ , utilizzando la medesima statistica test DF e gli stessi valori critici.

## 2.7 Applicazione Test ADF

Nei paragrafi seguenti, tramite il software Gretl, verrà svolta un'analisi rispetto alle tre variabili prese in esame per verificare la presenza di radice unitaria applicando il test ADF.

### 2.7.1 Test ADF per la variabile $\ln(Y_t)$

Test Dickey-Fuller aumentato per  $\ln_{JPM1D}$   
test all'indietro da 25 ritardi, criterio AIC  
Ampiezza campionaria 390  
Ipotesi nulla di radice unitaria:  $a = 1$

Con costante e trend  
inclusi 0 ritardi di  $(1-L)\ln_{JPM1D}$   
Modello:  $(1-L)y = b_0 + (a-1)y(-1) + e$   
Valore stimato di  $(a - 1)$ : -0,0026716  
Statistica test:  $\tau_c(1) = -0,58314$   
p-value 0,8711  
Coefficiente di autocorrelazione del prim'ordine per  $e$ : 0,086

Test Dickey-Fuller aumentato per  $\ln_{JPM5D}$   
test all'indietro da 25 ritardi, criterio AIC  
Ampiezza campionaria 312  
Ipotesi nulla di radice unitaria:  $a = 1$

Con costante e trend  
inclusi 3 ritardi di  $(1-L)\ln_{JPM5D}$   
Modello:  $(1-L)y = b_0 + b_1*t + (a-1)y(-1) + \dots + e$   
Valore stimato di  $(a - 1)$ : -0,0104596  
Statistica test:  $\tau_{ct}(1) = -1,06368$   
p-value asintotico 0,9334  
Coefficiente di autocorrelazione del prim'ordine per  $e$ : 0,001  
differenze ritardate:  $F(3, 306) = 4,061 [0,0075]$

Test Dickey-Fuller aumentato per  $\ln_{JPM30D}$   
test all'indietro da 25 ritardi, criterio AIC  
Ampiezza campionaria 321  
Ipotesi nulla di radice unitaria:  $a = 1$

Con costante e trend  
inclusi 0 ritardi di  $(1-L)\ln_{JPM30D}$   
Modello:  $(1-L)y = b_0 + b_1*t + (a-1)y(-1) + e$   
Valore stimato di  $(a - 1)$ : -0,00724129  
Statistica test:  $\tau_{ct}(1) = -0,951329$   
p-value 0,9476  
Coefficiente di autocorrelazione del prim'ordine per  $e$ : -0,074

Dai Test Dickey-Fuller aumentato per  $l\_JPM$ , effettuati rispetto ai tre intervalli temporali, risulta che l'ipotesi nulla  $H_0$  viene accettata, in quanto la statistica DF, in ogni test, risulta essere meno negativa del il valore critico al 10% uguale a -312, presente in *Tabella 1*. Tale risultato viene ulteriormente confermato dai valori assunti dal p-value che risultano essere tutti superiori al valore soglia 0.05.

Si può quindi constatare con certezza la presenza di radice unitaria, la presenza di non stazionarietà, e pertanto le serie non potranno essere studiate per fare alcun tipo di previsione.

## 2.7.2 Test ADF per la variabile $r_t$ ( $r_t = \ln(Y_t) - \ln(Y_{t-1})$ )

Test Dickey-Fuller aumentato per  $d\_l\_JPM1D$   
 test all'indietro da 25 ritardi, criterio AIC  
 Ampiezza campionaria 388  
 Ipotesi nulla di radice unitaria:  $a = 1$

Test senza costante  
 incluso un ritardo di  $(1-L)d\_l\_JPM1D$   
 Modello:  $(1-L)y = (a-1)*y(-1) + \dots + e$   
 Valore stimato di  $(a - 1)$ : -0,933889  
 Statistica test:  $\tau_{nc}(1) = -13,8761$   
 p-value asintotico  $1,402e-029$   
 Coefficiente di autocorrelazione del prim'ordine per  $e$ : -0,023

Test Dickey-Fuller aumentato per  $d\_l\_JPM5D$   
 test all'indietro da 25 ritardi, criterio AIC  
 Ampiezza campionaria 312  
 Ipotesi nulla di radice unitaria:  $a = 1$

Test senza costante  
 inclusi 2 ritardi di  $(1-L)d\_l\_JPM5D$   
 Modello:  $(1-L)y = (a-1)*y(-1) + \dots + e$   
 Valore stimato di  $(a - 1)$ : -1,08317  
 Statistica test:  $\tau_{nc}(1) = -12,0327$   
 p-value asintotico  $6,408e-025$   
 Coefficiente di autocorrelazione del prim'ordine per  $e$ : 0,002  
 differenze ritardate:  $F(2, 309) = 5,694 [0,0037]$

Test Dickey-Fuller aumentato per  $d\_l\_JPM30D$   
 test all'indietro da 25 ritardi, criterio AIC  
 Ampiezza campionaria 320  
 Ipotesi nulla di radice unitaria:  $a = 1$

Test senza costante  
 inclusi 0 ritardi di  $(1-L)d\_l\_JPM30D$   
 Modello:  $(1-L)y = (a-1)*y(-1) + e$   
 Valore stimato di  $(a - 1)$ : -1,05297  
 Statistica test:  $\tau_{nc}(1) = -18,8282$   
 p-value  $1,314e-036$   
 Coefficiente di autocorrelazione del prim'ordine per  $e$ : 0,000

Analizzando invece i risultati emersi dal Test Dickey-Fuller aumentato per  $d_1\_JPM$  effettuato rispetto ai tre intervalli temporali, l'ipotesi nulla  $H_0$  viene rifiutata in quanto la statistica test DF, in ogni test, risulta essere minore del valore critico al 10% uguale a -2.57 presente in *Tabella 1*. Ancora una volta il test viene confermato dai valori assunti dal p-value che risultano essere inferiori al valore soglia di 0.05. Da tali risultati possiamo concludere che i test confermano l'assenza di radici unitarie e, per contro, le serie storiche prese in esame assumono un comportamento stazionario. Tale caratteristica di stazionarietà consentirà quindi costruire un modello econometrico con lo scopo di studiare e predire i valori futuri delle serie storiche stesse.

## CAPITOLO 3

Come si è già visto in precedenza lo scopo dell'analisi delle serie storiche è quella di risalire ai processi stocastici che si suppone le abbiano generate. Questo consiste nell'identificare e stimare modelli statistici che garantiscano un accettabile grado di approssimazione della realtà in esame. A tale proposito in questo capitolo verrà presentata una classe di modelli statistici ampiamente utilizzati in campo econometrico e molto importanti nei processi di modellizzazione e rappresentazione delle serie storiche, i modelli autoregressivi a media mobile (ARMA). La classe dei modelli ARMA comprende sia i modelli autoregressivi (AR) che i modelli media mobile (MA).

### 3.1 Modelli autoregressivi

Un modello autoregressivo è un processo che permette di spiegare e prevedere i valori futuri di una serie storica mettendo in relazione una variabile temporale con i suoi valori passati, in altre parole esegue la regressione della variabile  $Y_t$  rispetto ai suoi valori ritardati.

I modelli autoregressivi possono essere del primo ordine e vengono indicati con AR(1), dove 1 indica che è del primo ordine, e seguono la seguente scrittura:

$$Y_t = \beta_0 + \Phi_1 Y_{t-1} + \varepsilon_t$$

dove  $\varepsilon_t$  è un processo white noise  $WN(0; \sigma^2)$ . Da tale scrittura si nota che il modello AR(1) utilizza solamente  $Y_{t-1}$ , variabile indipendente, per prevedere  $Y_t$ , variabile dipendente, non considerando altri valori passati che possono fornire informazioni potenzialmente utili per la previsione di  $Y_t$ .

Per tale motivo hanno assunto particolare importanza i modelli autoregressivi di ordine  $p$ , dove  $p$  è un numero intero non negativo, definiti AR( $p$ ). Questi modelli regrediscono la variabile  $Y_t$  considerando congiuntamente i valori passati  $Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}$ , fino al ritardo  $p$ :

$$Y_t = \beta_0 + \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + \dots + \Phi_p Y_{t-p} + \varepsilon_t$$

dove i parametri  $\Phi_1, \Phi_2, \Phi_p$  rappresentano i coefficienti del modello di regressione lineare di  $Y_t$ . I processi di tipo AR rispondono quindi al tentativo di spiegare il presente in funzione del passato, fino ad una certa distanza  $p$ .

Introducendo il concetto dell'operatore di ritardo, indicato con  $L$ , che presenta la proprietà di trasformare una variabile nel suo ritardo e applicandolo al processo  $Y_t \sim \text{AR}(p)$ , esso diventa:

$$Y_t = \beta_0 + \Phi_1 L Y_t + \Phi_2 L^2 Y_t + \dots + \Phi_p L^p Y_t + \varepsilon_t$$

dove  $LY_t = Y_{t-1}$ ,  $L^2 Y_t = Y_{t-2}$ , ...

Raccogliendo il termine  $Y_t$  si ottiene:

$$(1 - \Phi_1 L + \Phi_2 L^2 + \dots + \Phi_p L^p) Y_t = \beta_0 + \varepsilon_t$$

Ponendo:

$$(1 - \Phi_1 L + \Phi_2 L^2 + \dots + \Phi_p L^p) = \Phi(L)$$

Possiamo quindi riscrivere il processo  $\text{AR}(p)$  come:

$$\Phi(L) Y_t = \beta_0 + \varepsilon_t$$

### 3.1.1 Scelta della lunghezza dei ritardi

Per scegliere il numero di ritardi da considerare all'interno di un'autoregressione, è opportuno fare riferimento in particolare a due criteri di informazione: il criterio BIC e il criterio AIC.

Il criterio di informazione Bayesiano (Bayesian Information Criterion, BIC):

$$\text{BIC}(p) = \ln \left( \frac{\text{SSR}(p)}{T} \right) + (p + 1) \frac{\ln}{T}$$

dove  $\text{SSR}(p)$  è la somma dei quadrati dei residui della stima  $\text{AR}(p)$ .

Nel primo termine la somma dei quadrati dei residui ( $\text{SSR}$ ) decresce quando un ritardo viene aggiunto al modello. Al contrario, il secondo termine cresce quando si aggiunge un ritardo. Questo criterio bilancia i due termini in maniera tale che il numero di ritardi che minimizza il BIC sia uno stimatore consistente dell'ordine dell'autoregressione.

Il criterio di informazione di Akaike (Akaike Information Criterion, AIC):

$$\text{AIC}(p) = \ln \left( \frac{\text{SSR}(p)}{T} \right) + (p + 1) \frac{2}{T}$$

Lo stimatore AIC di  $p$ , a differenza del BIC, comporta la sostituzione di  $\ln(T)$  con 2. Tale criterio è molto diffuso in quanto si teme che il BIC porti ad un modello con troppo pochi ritardi.

### 3.2 Modelli a media mobile

Introduciamo adesso un'altra classe di modelli utili per modellare una serie storica: i modelli a media mobile (moving average in inglese, da cui la sigla MA). Questi modelli possono essere definiti come dei processi che generano una serie storica partendo dalla somma ponderata dei termini  $\varepsilon_t$ , dove  $\varepsilon_t$  è un processo white noise  $WN(0; \sigma^2)$ .

Come i modelli autoregressivi, anche i modelli a media mobile possono avere diversi ordini.

In particolare un processo a media mobile si dirà del primo ordine, MA(1), se:

$$Y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1}$$

dove  $\mu$  e  $\theta_1$  sono dei parametri costanti.

Dalla precedente equazione si può vedere che il modello MA(1) considera solamente i due valori più recenti di  $\varepsilon$ .

Un modello moving average di ordine  $q$  MA( $q$ ), con  $q \in \mathbb{N}$ , è invece un modello del tipo:

$$Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}$$

Dalla precedente equazione si può notare che i modelli di tipo MA rappresentano un tentativo di spiegare il presente come risultante da una successione incontrollata di impulsi casuali, statisticamente riassunti nel white noise  $\varepsilon_t$ .

Utilizzando la notazione dell'operatore di ritardo  $L$  e applicandola al processo  $Y_t \sim \text{MA}(p)$ , esso diventa:

$$Y_t = \varepsilon_t + \theta_1 L \varepsilon_t + \theta_2 L^2 \varepsilon_t + \dots + \theta_q L^q \varepsilon_t$$

Raccogliamo il termine  $\varepsilon_t$ :

$$Y_t = \varepsilon_t (1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q)$$

Ponendo:

$$\theta(L) = (1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q)$$

Possiamo quindi riscrivere il processo MA( $p$ ) come:

$$Y_t = \theta(L) \varepsilon_t$$

### 3.3 Modelli autoregressivi a media mobile

Combinando le due classi di modelli appena viste, AR(q) e MA(q), possiamo definire un modello autoregressivo a media mobile di ordine (p, q) (abbreviato in ARMA(p, q)) un processo del tipo:

$$Y_t = \beta_0 + \Phi_1LY_t + \Phi_2L^2Y_t + \dots + \Phi_pL^pY_t + \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \theta_q\varepsilon_{t-q}$$

Per comodità di rappresentazione introduciamo gli operatori di ritardo  $\beta(L)$  e  $\theta(L)$  in modo da poter riscrivere la precedente equazione come:

$$\Phi(L)Y_t = \beta_0 + \theta(L)\varepsilon_t$$

Da tale modello si può quindi pensare alla serie storica  $Y_t$  come risultante dei suoi p valori passati (modello AR) con l'interazione di un processo WN  $\varepsilon_t$  a sua volta connesso con i suoi q valori passati (modello MA). I modelli ARMA, dunque, altro non sono che combinazioni dei modelli autoregressivi e a media mobile, infatti se poniamo:

- $p = 0 \rightarrow \text{ARMA}(p, q) \equiv \text{MA}(q)$
- $q = 0 \rightarrow \text{ARMA}(p, q) \equiv \text{AR}(p)$

### 3.4 Applicazione modello ARMA

Attraverso i test di radice unitaria ADF, svolti nel capito precedente, abbiamo potuto constatare che solamente la variabile  $r_t$ , rispetto ai tre intervalli di tempo, presenta un comportamento stazionario. Questa importante caratteristica di stazionarietà ci permette di calcolare il modello ARMA al fine di studiare e spiegare l'andamento e il comportamento della serie storica stessa. Di seguito verranno quindi presentati i modelli ARMA per la variabile  $r_t$  rispetto ai tre intervalli temporali:  $\Delta t = 1$  giorno,  $\Delta t = 5$  giorni,  $\Delta t = 30$  giorni.

Modello 1: ARMA, usando le osservazioni 2-391 (T = 390)  
Variabile dipendente: d\_1\_JPM1D  
Errori standard basati sull'Hessiana

	Coefficiente	Errore Std.	p-value
const	5,62665e-05	3,75853e-05	0,1344
phi_1	-0,0587937	0,294536	0,8418
theta_1	0,151427	0,288812	0,6001

Modello 2: ARMA, usando le osservazioni 2-316 (T = 315)  
 Variabile dipendente: d\_1\_JPM5D  
 Errori standard basati sull'Hessiana

	<b>Coefficiente</b>	<b>Errore Std.</b>	<b>p-value</b>
<b>const</b>	0,000131840	9,49170e-05	0,1648
<b>phi_1</b>	0,143365	0,410922	0,7272
<b>theta_1</b>	-0,0985697	0,409125	0,8096

Modello 3: ARMA, usando le osservazioni 2-322 (T = 321)  
 Variabile dipendente: d\_1\_JPM30D  
 Errori standard basati sull'Hessiana

	<b>Coefficiente</b>	<b>Errore Std.</b>	<b>p-value</b>
<b>const</b>	-0,000202777	0,000210780	0,3360
<b>phi_1</b>	-0,0434410	0,555846	0,9377
<b>theta_1</b>	-0,0120621	0,554390	0,9826

Considerando i dati ottenuti, possiamo notare che tutti i modelli analizzati presentano i coefficienti  $\beta_0$ ,  $\Phi$ ,  $\theta$  non significativi, in quanto il valore assunto dal p-value, per ogni coefficiente, è superiore al valore soglia di 0.05. Da tale risultato possiamo trarre come conclusione che i modelli risultano essere non validi per spiegare l'andamento medio della serie nel corso del tempo e, di conseguenza, inadeguati a fornire previsioni future certe e attendibili.

### 3.5 Modelli GARCH

Nel paragrafo 2.5 è stato rappresentato il grafico dei quadrati dei rendimenti della serie storica, rispetto ai tre intervalli di tempo, che rappresenta l'andamento della variabilità della serie storica. Nei grafici che rappresentano la variabile  $r_t^2$  si può notare la presenza di picchi di volatilità (clustering) che segnalano quindi la presenza di eteroschedasticità, ovvero la presenza di una varianza caratterizzata da un comportamento non costante, che cambia al variare delle diverse osservazioni effettuate nel tempo.

Questi elementi conducono alla necessità di configurare un'ulteriore analisi introducendo il modello GARCH (Generalized Autoregressive Conditional Heteroskedasticity model), introdotto nel 1986 da Bollerslev come una famiglia generalizzata di modelli ARCH (proposti da Engle nel 1982).

Questo modello introduce la definizione di un modello autoregressivo sui quadrati dei residui della serie, e si propone di rappresentare l'evoluzione della variabilità nel tempo di una qualsiasi serie storica. Nel modello GARCH (p,q) la varianza condizionata al tempo t viene definita da:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \dots + \alpha_p \varepsilon_{t-p}^2 + \beta_1 \sigma_{t-1}^2 + \dots + \beta_q \sigma_{t-q}^2$$

dove i coefficienti  $\alpha_i$  e  $\beta_j$  sono entrambi maggiori di zero, i termini  $\varepsilon_i$ , definiti come componente ARCH di lunghezza p, sono una sequenza di variabili causali con media 0 e varianza unitaria, mentre i termini  $\sigma_q$  vengono definiti come la componente GARCH di lunghezza q.

### 3.6 Applicazione modello GARCH

Di seguito verranno presentati i modelli GARCH per la variabile  $r_t^2$ , rispetto ai tre intervalli temporali  $\Delta t = 1$  giorno,  $\Delta t = 5$  giorni,  $\Delta t = 30$  giorni, per provare a spiegare l'andamento della volatilità della serie storica nel tempo.

Modello 1: GARCH, usando le osservazioni 2-391 (T = 390)  
 Variabile dipendente: d\_1\_JPM1D  
 Errori standard basati sull'Hessiana

	Coefficiente	Errore Std.	p-value
<b>alpha(0)</b>	1,80637e-08	8,23745e-09	0,0283
<b>alpha(1)</b>	0,0921655	0,0286542	0,0013
<b>beta(1)</b>	0,854848	0,0408765	<0,0001

Modello 2: GARCH, usando le osservazioni 2-316 (T = 315)  
 Variabile dipendente: d\_1\_JPM5D  
 Errori standard basati sull'Hessiana

	Coefficiente	Errore Std.	p-value
<b>alpha(0)</b>	2,18400e-07	6,50142e-08	0,0008
<b>alpha(1)</b>	0,109892	0,0355198	0,0020
<b>beta(1)</b>	0,775145	0,0486398	<0,0001

Considerando i dati ottenuti, possiamo notare che tutti i modelli analizzati presentano i coefficienti  $\alpha_0$ ,  $\alpha_1$ ,  $\beta_1$  significativi, in quanto il valore assunto dal p-value, per ogni coefficiente, è inferiore al valore soglia di 0.05. Da tale risultato possiamo trarre come conclusione che i modelli risultano essere validi per spiegare la volatilità della serie nel corso del tempo e di conseguenza adeguati a fornire previsioni future certe e attendibili.

Numero Parole: 6586

## **Bibliografia**

- AMENDOLA B., D'ALESSANDRO M., 2017. *Processi Stocastici e Serie Storiche*. Relazione Finale, Università degli Studi di Napoli Federico II, Facoltà di Scienze Politiche
- BORRA S., DI CACCIO A., 2014. *Statistica: metodologie per le scienze economiche e sociali*. McGraw-Hill, 2014. p. 190
- BOX G. E. P., JENKINS G. M., 1970. *Time Series Analysis: Forecasting and Control*. San Francisco: Holden Day
- COTTREL A., LUCCHETTI R. J., 2019. *GNU Regression, Econometric and Time Series Library Gretl*.
- DI FONZO T., LISI F., 2007. *Serie storiche economiche: Analisi statistiche e applicazioni*. Carocci. p.32
- FALIVA M., ZOIA M.G., 1999. *Econometria delle serie storiche: profilo di un corso introduttivo*. Torino: Giappichelli Editore. p. 3
- KIRCHGASSNER G., WOLTERS J., HASSLER U., 2012. *Introduction to modern time series analysis*. Springer Science & Business Media. p. 1
- MASSAROTTO G., 2003. *Analisi delle Serie Temporal*i. Università Studi di Padova, Facoltà di Scienze Statistiche
- PAVAN S. 2013. *Gli effetti del valore iniziale sui test di radice unitaria*. Relazione Finale, Università degli Studi di Padova, Facoltà di Statistica
- PEARSONS W. M., 1919. *Indices of General Business Conditions*. Harvard Univ. Committee on Economic research.
- STOCK J., WATSON M. M., 2016. *Introduzione all'econometria*. Milano, Torino: Pearson
- VASILE G., 2011. *Processi Stazionari e Analisi delle Serie Storiche*. Relazione Finale, Università degli Studi di Bologna, Facoltà di scienze matematiche