

POLITECNICO DI TORINO

DIPARTIMENTO DI SCIENZE MATEMATICHE G. LUIGI LAGRANGE

Corso di Laurea in Ingegneria Matematica



Tesi di laurea magistrale

**STRUMENTI QUANTITATIVI E LORO  
IMPLEMENTAZIONE SOFTWARE PER  
L'ANALISI DI PORTAFOGLI FINANZIARI,  
CON APPROFONDIMENTI SUI  
MODELLI GARCH AVANZATI**

Giorgio Bella (230721)

Fabrizio Lasaponara (230757)

ANNO ACCADEMICO 2016/2017

# Indice

<b>Introduzione</b>	<b>5</b>
<b>1 Overview sul Quantitative Risk Management</b>	<b>6</b>
<b>2 Returns</b>	<b>9</b>
2.1 Definizione . . . . .	9
2.2 Distribuzione dei return . . . . .	11
2.2.1 Distribuzione normale . . . . .	12
2.2.2 Distribuzione t-Student . . . . .	15
2.2.3 Distribuzione normale asimmetrica . . . . .	16
2.2.4 Distribuzione t-Student asimmetrica . . . . .	16
2.3 Processi stocastici e rumori bianchi . . . . .	17
2.3.1 Processi stocastici . . . . .	17
2.3.2 Processi a rumore bianco . . . . .	18
2.3.3 Testare un rumore bianco . . . . .	19
<b>3 Modelli autoregressivi</b>	<b>20</b>
3.1 Il modello AR . . . . .	20
3.1.1 Funzione di autocorrelazione di un modello AR . . . . .	22
3.2 Media Mobile (MA) . . . . .	23
3.2.1 Proprietà dei modelli MA . . . . .	24
3.3 Modelli Autoregressivi a Media Mobile (ARMA) . . . . .	25
3.3.1 Proprietà dei modelli ARMA(1,1) . . . . .	25
3.3.2 Modelli ARMA generalizzati . . . . .	27
<b>4 Misure di rischio</b>	<b>29</b>
4.1 Distribuzione di profitti e perdite (P&L) . . . . .	29
4.2 Definizione delle misure di rischio . . . . .	29
4.2.1 Volatilità . . . . .	30
4.2.2 Value-at-Risk . . . . .	30
4.2.3 Expected Shortfall . . . . .	32
4.3 Proprietà delle misure di rischio . . . . .	33
<b>5 Pianificazione del rischio di portafoglio</b>	<b>37</b>
5.1 Calcolo di portafoglio . . . . .	37
5.1.1 Media e varianza di portafoglio . . . . .	37
5.1.2 Teorema di Eulero e decomposizione del rischio . . . . .	38
5.2 Risk Budgeting . . . . .	41
5.2.1 Risultati analitici per le misure di rischio . . . . .	41

5.3	Ottimizzazione di portafoglio . . . . .	42
5.3.1	Approccio standard . . . . .	42
5.3.2	Approccio inverso . . . . .	43
5.4	Decomposizione delle misure di rischio . . . . .	43
5.4.1	Decomposizione della volatilità per un portafoglio . . . . .	44
5.4.2	Interpretazione ed uso di $MCR_i^\sigma$ . . . . .	45
5.4.3	Interpretazione di MCR con VaR o ES . . . . .	47
5.4.4	$\beta$ come misura del contributo alla $\sigma$ di portafoglio . . . . .	47
<b>6</b>	<b>Stimare le misure di rischio</b> . . . . .	<b>51</b>
6.1	Stima non parametrica delle misure di rischio . . . . .	52
6.1.1	Volatilità - stima non parametrica . . . . .	52
6.1.2	Value at risk - stima non parametrica . . . . .	53
6.1.3	Expected Shortfall - stima non parametrica . . . . .	53
6.2	Stima parametrica delle misure di rischio . . . . .	59
6.2.1	Stima di massima verosimiglianza di $\theta$ . . . . .	59
6.2.2	Calcolo della stima di max verosimiglianza . . . . .	60
6.2.3	Caso di distribuzione normale . . . . .	61
6.3	Stima semi-parametrica delle misure di rischio . . . . .	62
6.4	Stimare le misure di rischio di portafoglio . . . . .	63
6.5	Stima non parametrica delle misure di rischio di portafoglio . . . . .	65
6.5.1	Volatilità di portafoglio - stima non parametrica . . . . .	65
6.5.2	VaR ed ES di portafoglio - stima non parametrica . . . . .	66
6.6	Stima parametrica delle misure di rischio di portafoglio . . . . .	72
6.6.1	Stima nel caso di distribuzione normale multivariata . . . . .	72
6.6.2	Stima nel caso di distribuzione t di Student . . . . .	72
<b>7</b>	<b>Modelli per la volatilità</b> . . . . .	<b>74</b>
7.1	Attesa condizionata . . . . .	74
7.2	Misure di rischio condizionate vs. non condizionate . . . . .	75
7.2.1	Media, varianza e volatilità condizionata . . . . .	76
7.2.2	Misure di rischio condizionate . . . . .	76
7.2.3	Volatilità - Proprietà empiriche dei return . . . . .	76
7.3	Modelli GARCH e ARCH . . . . .	78
7.3.1	Modello ARCH(p) di Engle . . . . .	81
7.3.2	ARCH(1) - Proprietà degli errori . . . . .	82
7.3.3	ARCH(p) - Proprietà degli errori . . . . .	84
7.3.4	Modello GARCH . . . . .	84
7.3.5	Specificazione media e varianza condizionata . . . . .	86
7.3.6	GARCH-in-Mean (GARCH-M) . . . . .	87

7.3.7	Aggregazione Temporale . . . . .	87
7.3.8	Testare gli effetti di un modello ARCH . . . . .	88
7.3.9	Stimare il GARCH con MLE . . . . .	89
7.3.10	Stimare il GARCH con QMLE . . . . .	92
7.3.11	GARCH - Determinare i Lag . . . . .	92
7.3.12	GARCH - Predire la media condizionata . . . . .	93
7.3.13	EWMA come modello previsionale . . . . .	95
7.3.14	Prevedere la volatilità dei Return Multiperiodali . . . . .	96
7.3.15	Previsione del VaR . . . . .	97
<b>8</b>	<b>GARCH Univariato - Modelli Avanzati</b>	<b>104</b>
8.1	Leverage Effect . . . . .	104
8.1.1	Testare gli effetti asimmetrici sulla volatilità condizionata . . . . .	105
8.2	Modelli GARCH asimmetrici . . . . .	106
8.2.1	Modello EGARCH . . . . .	106
8.2.2	TGARCH e GJR . . . . .	107
8.2.3	PGARCH . . . . .	108
8.2.4	News Impact Curve . . . . .	108
8.2.5	Forecast per i modelli asimmetrici . . . . .	109
8.3	GARCH - errori non normali . . . . .	110
8.3.1	GARCH - errori t-Student . . . . .	110
8.3.2	Distribuzione generale degli errori . . . . .	111
8.4	Modelli GARCH a lunga memoria . . . . .	112
8.5	Valutare le predizioni della volatilità . . . . .	112
8.5.1	Statistiche tradizionali - valutazione della predizione . . . . .	113
8.5.2	Test di Diebold-Mariano per l'accuratezza della predizione . . . . .	114
8.5.3	Regressione della predizione - Mincer-Zarnowitz . . . . .	114
8.5.4	I principali problemi nel valutare le predizioni . . . . .	115
<b>9</b>	<b>Implementazione software</b>	<b>124</b>
9.1	Overview architetturale . . . . .	124
9.2	Presentation Layer - Use case . . . . .	127
9.3	Logic Layer - Use case . . . . .	129
9.4	Data Layer - Use case . . . . .	130
<b>10</b>	<b>Utilizzo della piattaforma</b>	<b>132</b>
10.1	User Experience . . . . .	132
10.2	Use case tipico . . . . .	134

10.2.1 Grafici e Analisi Tecnica . . . . .	135
10.2.2 Analisi fondamentale . . . . .	140
10.2.3 Metodi quantitativi e analisi del rischio . . . . .	144
<b>Conclusioni</b>	<b>153</b>
<b>Riferimenti</b>	<b>154</b>
Bibliografia . . . . .	154
Sitografia . . . . .	156
Software utilizzati . . . . .	157

## Introduzione

L'obiettivo del progetto di tesi è quello di mostrare gli aspetti teorici necessari alla realizzazione di una piattaforma web-based di supporto agli investimenti finanziari, attraverso la messa a disposizione di una serie di strumenti quantitativi complessi a mezzo un front-end intuitivo.

In particolare, l'elaborato è suddiviso in due parti principali:

1. capitoli teorici (dal primo all'ottavo) per introdurre i modelli esposti in piattaforma corredati da esempi;
2. capitoli tecnici e funzionali (nono e decimo) per descrivere le modalità d'implementazione della piattaforma e l'esperienza d'uso della stessa.

I capitoli teorici sono organizzati in modo da fornire in primo luogo un background su distribuzioni e processi stocastici degli oggetti principali dell'analisi quantitativa trattata: i return dei titoli negoziati sul mercato.

In secondo luogo, dopo una prima introduzione ai modelli autoregressivi AR e MA per la modellizzazione dei return, viene dedicato ampio spazio alle misure di rischio principali, al loro uso nella pianificazione del rischio di portafoglio e alle diverse modalità per la loro stima.

In fine, vengono riportati i principali modelli per la volatilità: ARCH, GARCH e GARCH asimmetrici.

I capitoli tecnici e funzionali, invece, in primo luogo dettagliano le modalità di sviluppo progettuale descrivendo l'architettura applicativa dal punto di vista logico e tecnologico, per poi descrivere l'esperienza d'uso della piattaforma in cui viene mostrato lo use case tipico di un trader in cerca di insights quantitativi.

Poiché la piattaforma è stata sviluppata a quattro mani, il presente elaborato e quello di Bella G. e Lasaponara F.- *Strumenti quantitativi e loro implementazione software per l'analisi di portafogli finanziari, con approfondimenti sui modelli GARCH multivariati*, PoliTo (2018) sono identici a meno del Capitolo ottavo e del Capitolo nono, in cui Bella ha approfondito i modelli GARCH Multivariati in luogo di quelli asimmetrici e ha descritto in modo tecnologico anziché funzionale i layer dell'architettura implementata.

# 1 Overview sul Quantitative Risk Management

Il *rischio* - secondo l'Oxford English Dictionary - è un azzardo, una possibilità che si verifichino conseguenze negative, perdite o esposizioni a eventi ignoti. Si tratta di una definizione molto generale, che nel contesto finanziario il rischio può essere definito come:

“la probabilità quantificabile di perdita o di ritorni al di sotto delle attese”.

Vi sono numerosi tipi di rischi legati al contesto finanziario:

- **Market Risk:** il rischio di cambiamento nel valore di una posizione finanziaria a causa delle oscillazioni di valore di una componente sottostante da cui dipende quella stessa posizione, come ad esempio i prezzi di azioni e obbligazioni, tassi d'interesse, prezzi di materie prime, ecc. Questo tipo di rischio sarà prevalentemente quello trattato da questo elaborato.
- **Credit Risk:** il rischio di non ricevere un pagamento concordato da un investimento in essere - come ad esempio crediti e bond - a causa del default del debitore. I modelli di *credit risk* si basano su un'estensione dei modelli di *market risk* tenendo conto del rischio di fallimento.
- **Operational Risk:** il rischio di perdite risultanti dall'inadeguatezza o errore di un processo interno, di una persona, del sistema o a causa di un evento esterno.
- **Liquidity Risk:** il rischio derivante dalla mancanza di sufficiente desiderabilità di mercato per un investimento che perciò non può essere comprato o venduto abbastanza rapidamente da prevenire o minimizzare le perdite. Questo tipo di rischio è molto alto durante le crisi finanziarie ed il rischio di mercato può notevolmente essere influenzato da quest'ultimo.
- **Model Risk:** è il rischio di usare un modello scorretto per misurare il rischio (ad esempio Black-Scholes usato in contesti in cui il prezzo non oscilla come un Browniano geometrico).
- **Systemic Risk:** è il rischio che si riferisce alla probabilità e alla magnitudo che certe conseguenze negative possano espandersi a tutto un certo sistema finanziario (ad esempio il fallimento di Lehman Bros.).

E' evidente come il concetto di rischio sia strettamente collegato a quello di aleatorietà, perciò nella trattazione faremo largo uso della teoria probabilistica e della statistica, che sono di fatto alla base del quantitative risk management. E' importante tenere presente a questo punto che c'è differenza tra misurare il rischio e gestire il rischio.

Il *Risk Measurement* riguarda la quantificazione della probabilità di perdita. Esempi di questo tipo sono ad esempio la volatilità giornaliera dei ritorni di un portafoglio o la sua massima perdita possibile su un anno.

Il *Risk Management* è, invece, la disciplina per convivere con la possibilità che eventi futuri possano causare effetti avversi. La gestione del rischio coinvolge perciò l'uso di sofisticati strumenti, come derivati su titoli (opzioni, swap, ecc) e altri tipi di prodotti per il trasferimento del rischio al mercato.

L'importanza di questa disciplina è stata messa in risalto da diversi eventi storici moderni:

- la creazione di mercati di future e opzioni;
- lo sviluppo accademico della matematica per il pricing dei derivati;
- l'avanzamento tecnologico;
- le crisi finanziarie;
- il fallimento di grandi imprese quotate;
- l'abolizione del gold standard (che rendeva fissi i tassi di cambio);
- la deregolamentazione finanziaria degli anni '90;
- la crescita esponenziale dei mercati OTC (*Over The Counter*).

In particolare, la comunità internazionale ha fatto fronte nel tempo a questo tipo di criticità con l'introduzione di standard ed accordi, come i celebri Basilea I, II e III.

Proprio nel processo di definizione di questi accordi, nascono diverse definizioni di misura del rischio al fine di definire alcune linee guida per la sua gestione. Acquisisce così nel tempo sempre più importanza la capacità di gestire il rischio, che al giorno d'oggi rappresenta addirittura parte del valore riconosciuto dagli azionisti ad un'impresa.

L'obiettivo di una gestione del rischio di tipo quantitativo è dare precise definizioni e una connotazione matematica ai concetti di rischio. Il focus è sugli output inattesi, anormali o estremi, piuttosto che quelli attesi, normali o medi che sono oggetto di molte applicazioni classiche.

I modelli di Quantitative Risk Management descrivono l'interdipendenza e la concentrazione dei rischi relativi al realizzarsi di eventi estremi prendendo spunto da diverse discipline quantitative affini quali - tra le altre - matematica finanziaria e attuariale, statistica ed econometria.

## 2 Returns

### 2.1 Definizione

Nel corso di tutta la trattazione verrà fatto uso del concetto di *return*, ossia della variazione di valore di un asset a seguito di un investimento in un certo arco di tempo. Definiamo:

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} \quad (1)$$

$$G_t = 1 + R_t = \frac{P_t}{P_{t-1}} \quad (2)$$

con  $R_t$  e  $G_t$  intesi come *Simple Net Return* e *Gross Return* rispettivamente.

Chiaramente gli archi temporali possono essere molteplici, in tal caso (k periodi) si avrebbe:

$$\begin{aligned} R_t(k) &= \frac{P_t - P_{t-k}}{P_{t-k}} = \frac{P_t}{P_{t-k}} - 1 \\ R_t(k) &= \frac{P_t}{P_{t-1}} \cdot \frac{P_{t-1}}{P_{t-2}} \cdot \dots \cdot \frac{P_{t-k+1}}{P_{t-k}} - 1 \\ &= G_t \cdot G_{t-1} \cdot \dots \cdot G_{t-k+1} - 1 \end{aligned}$$

Da qui si ricavano:

$$R_t(k) = \prod_{j=0}^{k-1} G_{t-j} - 1 \quad (3)$$

$$G_t(k) = 1 + R_t(k) = \prod_{j=0}^{k-1} G_{t-j} \quad (4)$$

Se siamo in presenza di un titolo che rilascia dividendi, modifichiamo la (1) come segue (dove  $D_t$  è l'ammontare del dividendo pagato nel periodo):

$$R_t = \frac{P_t + D_t - P_{t-1}}{P_{t-1}} = \frac{P_t - P_{t-1}}{P_{t-1}} + \frac{D_t}{P_{t-1}} \quad (5)$$

Un'ulteriore modifica si rende necessaria nel caso in cui si voglia tenere conto dell'inflazione:

$$1 + R_t^{\text{Real}} = (1 + R_t) \cdot \frac{CPI_t}{CPI_{t-1}} \quad (6)$$

Dove  $CPI_t$  rappresenta il *consumer price index*, ossia il valore di un certo paniere di beni di consumo valutato al tempo  $t$ .

E' ora opportuno introdurre il concetto di portafoglio, inteso come set di titoli posseduti da un soggetto. In questo caso parliamo di *portfolio return*. Indichiamo con  $P_{i,t}$  il prezzo del titolo  $i$  nel al tempo  $t$ . Convenzionalmente indichiamo con  $P_{p,t}$  il prezzo complessivo di tutti i titoli in portafoglio al tempo  $t$ . Si ha perciò:

$$P_{p,t} = \sum_{i=1}^n w_i P_{i,t} \quad \text{con} \quad \sum_{i=1}^n w_i = 1 \quad (7)$$

$$R_{p,t} = \sum_{i=1}^n w_i R_{i,t} \quad (8)$$

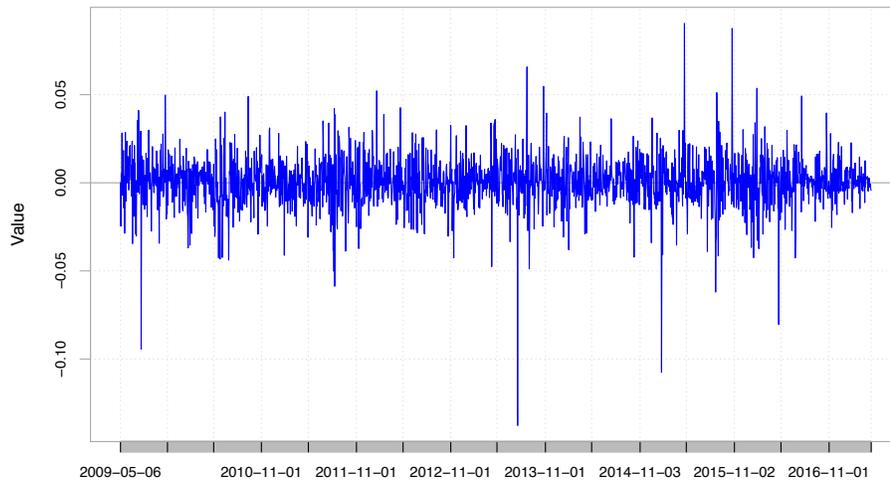


Figura 1: Esempio di plot dei log-return giornalieri per il titolo Microsoft negli ultimi 8 anni. Si noti come i log-return abbiano pressoché media nulla.

Nel contesto della misura quantitativa del rischio spesso, anziché utilizzare il concetto di *return*, si preferisce lavorare con i *log-return*:

$$\begin{aligned}
r_t &= \ln(1 + R_t) = \ln\left(\frac{P_t}{P_{t-1}}\right) \\
&= \ln(P_t) - \ln(P_{t-1}) \\
&= p_t - p_{t-1} \\
e^{r_t} &= 1 + R_t = \frac{P_t}{P_{t-1}} \\
\Rightarrow P_t &= P_{t-1}e^{r_t}
\end{aligned}$$

In caso di più periodi, partendo dalla (4) e applicando il logaritmo si ha:

$$\begin{aligned}
r_t(k) &= \ln(1 + R_t(k)) = \ln \prod_{j=0}^{k-1} (1 + R_{t-j}) \\
&= \sum_{j=0}^{k-1} r_{t-j}
\end{aligned}$$

Analogamente a come fatto nella pagine precedenti, consideriamo il caso di presenza del dividendo

$$\begin{aligned}
r_t &= \ln(1 + R_t) = \ln\left(\frac{P_t + D_t}{P_{t-1}}\right) \\
&= \ln(P_t + D_t) - \ln(P_{t-1})
\end{aligned}$$

e di inflazione

$$\begin{aligned}
r_t^{Real} &= \ln(1 + R_t^{Real}) = \ln\left(\frac{P_t}{P_{t-1}} \cdot \frac{CPI_{t-1}}{CPI_t}\right) \\
&= r_t - \pi_t
\end{aligned}$$

Utilizzando la notazione precedente, il *log-return* di portafoglio risulta essere:

$$\begin{aligned}
r_{t,p} &= \ln(1 + R_{t,p}) \\
&= \ln\left(1 + \sum_{i=1}^n w_i R_{i,t}\right)
\end{aligned}$$

## 2.2 Distribuzione dei return

I return ed i log-return - dal punto di vista statistico - sono caratterizzati da distribuzioni che presentano code pesanti, asimmetrie e assenza di correlazioni seriali. Per comprendere la forma di queste distribuzioni, è bene

considerare anzitutto alcune misure salienti. Sia  $\tilde{r}$  una variabile aleatoria di *log-return*, si ha:

$$\begin{aligned}\mu &= E[\tilde{r}] : \text{centro} \\ \sigma^2 &= \text{var}(\tilde{r}) = E[(\tilde{r} - \mu)^2] : \text{spread} \\ \text{skew}(\tilde{r}) &= E\left[\frac{(\tilde{r} - \mu)^3}{\sigma^3}\right] : \text{simmetria} \\ \text{kurt}(\tilde{r}) &= E\left[\frac{(\tilde{r} - \mu)^4}{\sigma^4}\right] : \text{spessore coda}\end{aligned}$$

Si ricordi che il coefficiente di curtosi è dato dalla formula:  $\gamma(r) = \text{kurt}(r) - 3$ .

Se  $\gamma > 0$  si ha una curva *leptocurtica* (più appuntita di una normale), se  $\gamma < 0$  si ha una curva *platicurtica* (più “piatta” di una normale) e se  $\gamma = 0$  si ha una curva *normocurtica*. Ricordiamo ora alcune tra le distribuzioni più utilizzate in questo campo.

### 2.2.1 Distribuzione normale

Ipotizzando che i ritorni si distribuiscano come una normale  $X \sim N(\mu, \sigma^2)$ , si ha:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right), \quad -\infty \leq x \leq \infty \quad (9)$$

$$\begin{aligned}E[X] &= \mu \\ \text{var}[X] &= \sigma^2 \\ \text{skew} &= 0 \\ \text{kurt} &= 3 \\ m_k &= 0 \text{ per } k \text{ dispari}\end{aligned}$$

Con  $m_k$  momento k-esimo.

Sia  $\{r_1, \dots, r_T\}$  un campione casuale di taglia  $T$ , dove gli  $r_i$  sono delle

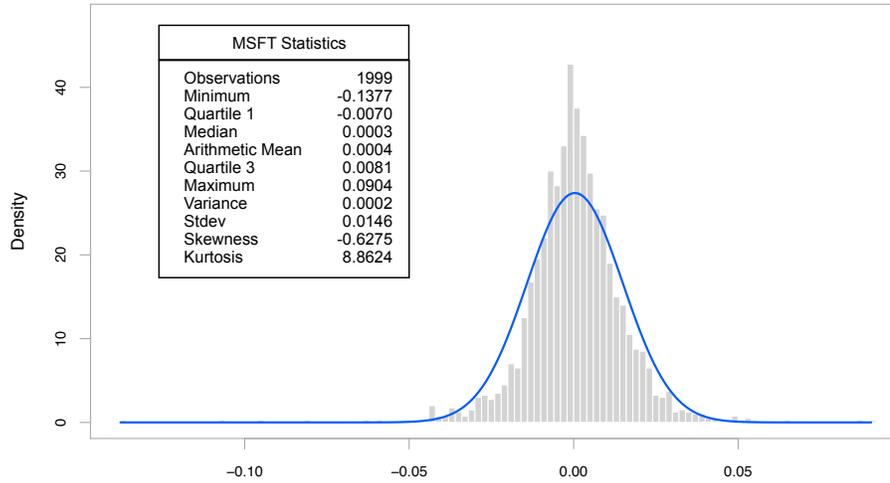


Figura 2: Fit di una normale a media nulla sui log-return giornalieri del titolo Microsoft negli ultimi 8 anni. Il valore di curtosi così elevato è indice di una distribuzione leptocurtica.

realizzazioni della v.a.  $\tilde{r}$ .

$$\hat{\mu} = \frac{1}{T} \sum_{t=1}^T r_t$$

$$\hat{m}_2 = \hat{\sigma}^2 = \frac{1}{T-1} \sum_{t=1}^T (r_t - \hat{\mu})^2$$

$$\hat{m}_k = \frac{1}{T-1} \sum_{t=1}^T (r_t - \hat{\mu})^k$$

$$\hat{\text{skew}} = \frac{\hat{m}_3}{\hat{\sigma}^3}, \quad \hat{\text{kurt}} = \frac{\hat{m}_4}{\hat{\sigma}^4}$$

Per capire se i ritorni sono veramente distribuiti come normali, possono essere eseguiti dei test.

- **QQ-plot:** grafico dei quantili empirici standardizzati vs. i quantili teorici della distribuzione specificata (si vedano la *Figura 3* e la *Figura 4* nelle pagine seguenti a titolo di esempio).

- **Jarque-Bera (JB)**: il test considera la v.a. JB, costruita in funzione dei coefficienti di asimmetria e curtosi:

$$JB = \frac{T}{6} \left( \text{skew}^2 + \frac{(\text{kurt} - 3)^2}{4} \right)$$

Sotto ipotesi di normalità, si dimostra che  $JB \stackrel{A}{\sim} \chi^2(2)$ , fatto che rende possibile la costruzione di un test d'ipotesi.

- **Kolmogorov-Smirnov (KS)**: è un test che compara la funzione di ripartizione empirica dei ritorni con quella della normale (o qualsiasi altra distribuzione che si vuole testare). In sintesi si tratta di:

- Ordinare i ritorni  $r_{(1)} \leq \dots \leq r_{(T)}$  e calcolare la funzione di ripartizione empirica  $\hat{F}_r(r_{(t)}) = t/T$
- Valutare la funzione di ripartizione normale  $\Phi\left(\frac{r_{(t)} - \hat{\mu}}{\hat{\sigma}}\right)$
- Calcolare la statistica KS:  $KS = \sup_t \left| \Phi\left(\frac{r_{(t)} - \hat{\mu}}{\hat{\sigma}}\right) - \frac{t}{T} \right|$
- Questo valore dovrà poi essere confrontato con i quantili della distribuzione KS al fine di accettare o rifiutare l'ipotesi nulla di uguaglianza della distribuzione teorica ed empirica.

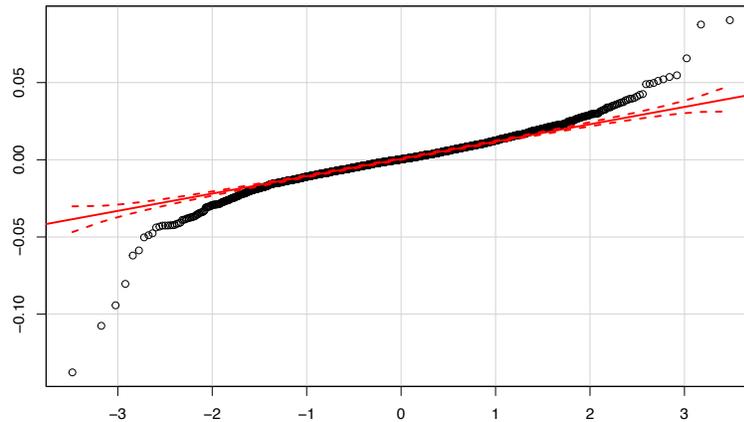


Figura 3: QQ-plot dei quantili empirici standardizzati vs. i quantili teorici della distribuzione normale. Si noti come i valori estremi non siano ben descritti dalla distribuzione considerata. Le bande tratteggiate rosse rappresentano l'intervallo di confidenza al 95%.

### 2.2.2 Distribuzione t-Student

Siano  $Z \sim N(0, 1)$ ,  $W \sim \chi^2(v)$  tali che  $Z$  e  $W$  siano indipendenti. Allora

$$X = \frac{Z}{\sqrt{W/v}} \sim t_v \quad (10)$$

dove  $t_v$  denota una distribuzione t-Student (standardizzata) con  $v$  gradi di libertà. Si noti che:

$$E[X] = 0, \quad \text{var}(X) = \frac{v}{v-2}, \quad v > 2$$
$$\text{skew} = 0, \quad \text{kurt} - 3 = \frac{6}{v-4}, \quad v > 4$$

L'esistenza dei momenti dipende di gradi di libertà  $v$ .

Se  $X \sim t_v$ , allora

$$Y = \mu + \frac{\sigma X}{\sqrt{v/(v-2)}}$$

ha momenti

$$E[Y] = \mu, \quad \text{var}(Y) = \sigma^2$$

La funzione di densità della v.a. definita nella (10) è

$$f(x; p) = \left[ \frac{\Gamma\left(\frac{v+1}{2}\right)}{(\pi v)^{1/2} \Gamma\left(\frac{v}{2}\right)} \right] \frac{1}{[1 + (x^2/v)]^{(v+1)/2}}$$

Ricordiamo la definizione della funzione gamma:

$$\Gamma(t) = \int_0^\infty x^{t-1} e^{-x} dx$$

Qualora non siano noti i gradi di libertà  $v$ , è possibile stimare questo parametro attraverso il metodo della *massima verosimiglianza*. Alternativamente è possibile utilizzare un metodo più semplice basato sulla curtosi:

$$\text{kurt} - 3 = \frac{6}{v-4} \Rightarrow v = \frac{6}{\text{kurt} - 3} + 4$$

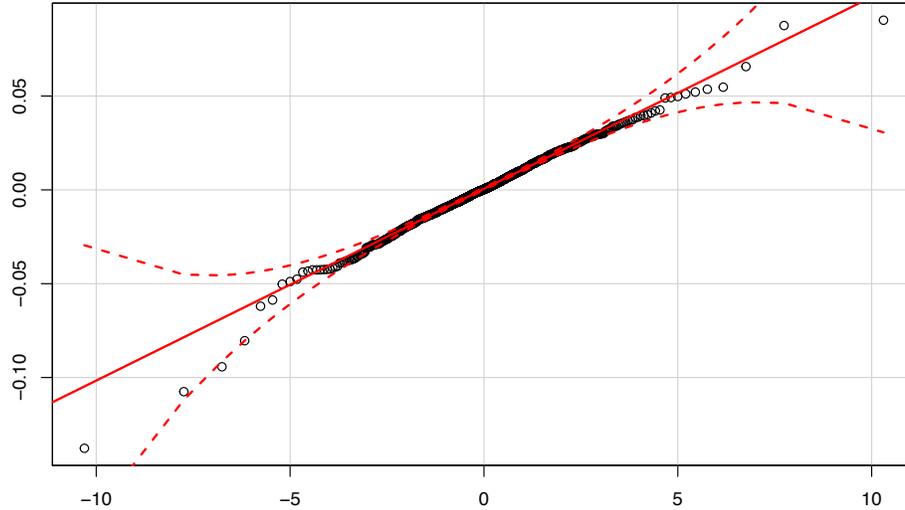


Figura 4: QQ-plot dei log-returns del titolo Microsoft ipotizzando una loro distribuzione t-student: grafico dei quantili empirici standardizzati vs. i quantili teorici della distribuzione t-student. Si noti come i valori estremi questa volta - a differenza del caso normale in *Figura 3* - risultino entro l'intervallo di confidenza al 95% rappresentato dalle barre rosse tratteggiate.

### 2.2.3 Distribuzione normale asimmetrica

Utilizzando la notazione data da Azzalini e Capitanio (2002), definiamo  $Z \sim SN(\xi, \omega, \alpha)$  come una variabile aleatoria normale asimmetrica con densità

$$f_Z(x) = 2\phi(z - \xi)\Phi(\alpha\omega^{-1}(z - \xi))$$

$$\phi(z) = (2\pi)^{-1/2} \exp\left(-\frac{z^2}{2}\right), \quad \Phi(z) = \int_{-\infty}^z \phi(x)dx$$

$\xi$  = parametro di posizione,  $-\infty < \xi < \infty$

$\omega$  = parametro di scala,  $\omega > 0$

$\alpha$  = parametro di forma (asimmetria),  $-\infty < \alpha < \infty$

Si noti che per  $\alpha = 0$ ,  $Z \sim N(\xi, \omega^2)$ ; per  $\alpha > 0$  si ha asimmetria positiva; per  $\alpha < 0$  ho asimmetria negativa.

### 2.2.4 Distribuzione t-Student asimmetrica

Utilizzando la notazione data da Azzalini e Capitanio (2002), definiamo  $Y \sim St(\xi, \omega, \alpha, \nu)$  come una variabile aleatoria t-Student asimmetrica usando la

trasformazione

$$\begin{aligned} Y &= \xi + V^{-1/2}Z \\ Z &\sim SN(\xi, \omega, \alpha) \\ V &\sim \chi^2(v)/v \end{aligned}$$

I parametri  $\xi$ ,  $\omega$  e  $\alpha$  hanno la medesima interpretazione vista per la skew-normal e  $v$  indica il numero di gradi di libertà.

### 2.3 Processi stocastici e rumori bianchi

Sempre col fine di studiare il comportamento dei *returns*, introduciamo alcune altre definizioni utili. In tutto il paragrafo considereremo un scala temporale discreta, poiché in tempi discreti avvengono le transazioni sui mercati finanziari.

#### 2.3.1 Processi stocastici

Un processo stocastico  $\{\tilde{r}_t\}$  è *stazionario in covarianza* se

$$E[\tilde{r}_t] = \mu \quad \forall t \quad (11)$$

$$\text{cov}(\tilde{r}_t, \tilde{r}_{t-j}) = E[(\tilde{r}_t - \mu)(\tilde{r}_{t-j} - \mu)] = \gamma_j \quad \forall t, j \quad (12)$$

Il parametro  $\gamma_j$  è chiamato *autocovarianza* di  $\{\tilde{r}_t\}$  di  $j$ -esimo ordine o lag  $j$ .

L'autocorrelazione di  $\{\tilde{r}_t\}$  è definita da

$$\rho_j = \frac{\text{cov}(\tilde{r}_t, \tilde{r}_{t-j})}{\sqrt{\text{var}(\tilde{r}_t) \text{var}(\tilde{r}_{t-j})}} = \frac{\text{cov}(\tilde{r}_t, \tilde{r}_{t-j})}{\text{var}(\tilde{r}_t)} = \frac{\gamma_j}{\gamma_0} \quad (13)$$

ed il grafico di  $\rho_j$  in funzione di  $j$  è chiamata *funzione di autocorrelazione* (ACF).

L'autocovarianza campionaria di lag  $j$  e l'autocorrelazione campionaria di lag  $j$  sono definite come

$$\hat{\gamma} = \frac{1}{T} \sum_{t=j+1}^T (r_t - \bar{r})(r_{t-j} - \bar{r}) \quad (14)$$

$$\hat{\rho} = \frac{\hat{\gamma}_j}{\hat{\gamma}_0} \quad (15)$$

dove  $\bar{r} = \sum_{t=1}^T r_t$  è la media semplice. La *funzione di autocorrelazione campionaria* (SACF) è il grafico di  $\hat{\rho}$  in funzione di  $j$ .

### 2.3.2 Processi a rumore bianco

Probabilmente il processo stocastico stazionario a tempo discreto più semplice in senso di serie temporale è quello composto rumori bianchi indipendenti ed identicamente distribuiti come gaussiane, *Gaussian White Noise* (GWN). Indichiamo questo processo con  $\{\tilde{r}_t\} \sim N(0, \sigma^2)$  iid  $\equiv GWN(0, \sigma^2)$ . Questo processo ha  $\mu = \gamma_j = \rho_j = 0$  ( $\forall j \neq 0$ ).

Due processi lievemente più generali sono: il processo a *rumori bianchi indipendenti* (IWN), indicato come  $\{\tilde{r}_t\} \sim IWN(0, \sigma^2)$ , e il processo a *rumori bianchi* (WN), indicato come  $\{\tilde{r}_t\} \sim WN(0, \sigma^2)$ .

Entrambi questi processi hanno media zero e varianza  $\sigma^2$ , ma il processo IWN è caratterizzato da incrementi indipendenti, mentre il processo WN ha incrementi scorrelati<sup>1</sup>.

La funzione di autocorrelazione campionaria tipicamente è mostrata con un intervallo di confidenza al 95%. Questo intorno è basato sul risultato che se  $\{\tilde{r}_t\} \sim N(0, \sigma^2)$  iid, allora

$$\hat{\rho}_j \overset{A}{\sim} N\left(0, \frac{1}{T}\right), \quad j > 0.$$

La notazione  $\hat{\rho}_j \overset{A}{\sim} N\left(0, \frac{1}{T}\right)$  significa che la distribuzione di  $\hat{\rho}_j$  è approssimata da una distribuzione normale con media 0 e varianza  $\frac{1}{T}$ . Questo fatto segue dal teorema del limite centrale:  $\sqrt{T}\hat{\rho} \overset{d}{\sim} N(0, 1)$ . L'intervallo al 95% è quindi  $\pm \frac{1.96}{\sqrt{T}}$ .

---

<sup>1</sup>Si tenga presente che variabili aleatorie scorrelate non sono necessariamente indipendenti. Nel caso particolare di due variabili aleatorie gaussiane, scorrelazione implica indipendenza.

### 2.3.3 Testare un rumore bianco

Supponiamo di voler testare l'ipotesi nulla

$$H_0 : \{\tilde{r}_t\} \sim WN(0, \sigma^2)$$

Sotto questa ipotesi, tutte le autocorrelazioni  $\rho_j$  sono uguali a zero per  $j > 0$ . Per eseguire questo test, Box e Pierce (1970) hanno introdotto la *Q-statistic*

$$Q(k) = T \sum_{j=1}^k \hat{\rho}^2 \quad (16)$$

Sotto l'ipotesi nulla, la (16) si distribuisce asintoticamente come una  $\chi^2(k)$ . In un campione a dimensione finita, la statistica-Q potrebbe però non essere ben approssimata dalla  $\chi^2(k)$ . Ljung e Box (1978) hanno suggerito la seguente modifica

$$MQ(k) = T(T+2) \sum_{j=1}^k \frac{\hat{\rho}^2}{T-j} \quad (17)$$

### 3 Modelli autoregressivi

I modelli autoregressivi vengono prevalentemente utilizzati per comprendere l'autocorrelazione dei returns al variare del lag (si veda *Paragrafo 2.3.1*).

Un modello semplice che può essere utilizzato per predire le future osservazioni è

$$r_t = \phi_0 + \phi_1 r_{t-1} + a_t$$

dove assumiamo che  $\{a_t\}$  sia un rumore bianco con media nulla e varianza  $\sigma_a^2$ . Questo modello equivale ad una regressione lineare semplice in cui  $r_t$  è la variabile esplicativa. Questo modello, nella letteratura sulle serie temporali, è noto come il modello autoregressivo (AR) di ordine 1, che indicheremo come modello AR(1).

#### 3.1 Il modello AR

Il modello AR(1) ha diverse proprietà simili a quelle proprie della regressione lineare. Esistono però alcune differenze chiave. Per i modelli AR(1) vale

$$E(r_t | r_{t-1}) = \phi_0 + \phi_1 r_{t-1} \quad (18)$$

$$Var(r_t | r_{t-1}) = Var(a_t) = \sigma_a^2 \quad (19)$$

Si tratta di una proprietà di Markov<sup>2</sup>, per cui il return  $r_t$  dipende unicamente da  $r_{t-1}$  e non da  $r_{t-i}$  per  $i > 1$ . Utilizzando una notazione formale, significa che  $\mathbb{P}(r_t | r_{t-1}, r_{t-2}, \dots) = \mathbb{P}(r_t | r_{t-1})$ . Per questa ragione spesso ci si riferisce al processo AR(1) come al *processo di Markov*.

Chiaramente ci sono modelli autoregressivi di ordine superiore al primo; una generalizzazione del modello AR(1) può essere scritta come modello AR( $p$ )

$$r_t = \phi_0 + \phi_1 r_{t-1} + \dots + \phi_p r_{t-p} + a_t \quad p \in \mathbb{N}^+$$

dove  $\{a_t\}$  è la serie di rumori bianchi mostrata nel precedente modello AR(1). Nell'AR( $p$ ), l'attesa condizionata<sup>3</sup> di  $r_t$  è determinata da una combinazione dei precedenti  $p$  valori  $r_{t-i}$  per ( $i = 1, \dots, p$ ). In modo analogo al modello AR(1), il modello AR( $p$ ) è equivalente ad una regressione lineare multipla

---

<sup>2</sup>Bharucha-Reid A. T. - *Elements of the Theory of Markov Processes and Their Applications*, McGraw-Hill (1960)

<sup>3</sup>Ulteriori approfondimenti nel *Paragrafo 7.1*

con i coefficienti di lag  $\phi_i$  utilizzati come variabili esplicative.

E' utile descrivere le proprietà caratteristiche dei modelli AR al fine delle applicazioni. Assumendo di avere un modello AR(1) debolmente stazionario, possiamo determinare delle proprietà univoche. Utilizzando la proprietà di stazionarietà debole,  $E(r_t) = \mu$ ,  $\text{Var}(r_t) = \gamma_0$  e  $\text{Cov}(r_t, r_{t-l}) = \gamma_l$ , dove  $\mu$  e  $\gamma_0$  sono costanti e  $\gamma_l$  è una funzione di  $l$  e non di  $t$ ; possiamo riscrivere l'equazione dell'AR(1) in modo da ottenere la seguente espressione

$$E(r_t) = \phi_0 + \phi_1 E(r_{t-1})$$

dove  $E(a_t) = 0$ . Sotto condizione di stazionarietà,  $E(r_t) = E(r_{t-1}) = \mu$ . Troviamo perciò l'equazione

$$\mu = \phi_0 + \phi_1 \mu$$

che può essere riscritta per ottenere

$$E(r_t) = \mu = \frac{\phi_0}{1 - \phi_1}.$$

Questo ha due implicazioni per  $r_t$ : la sua media esiste solo se  $\phi_1 \neq 1$  ed è nulla se e solo se  $\phi_0 = 0$ . Significa che per il modello AR(1),  $\phi_0$  è in relazione alla media di  $r_t$  e, in particolare,  $\phi_0 = 0$  implica che  $E(r_t) = 0$ .

Utilizzando l'equazione  $\phi_0 = (1 - \phi_1)\mu$ , il modello AR(1) può essere riscritto nel seguente modo

$$r_t - \mu = \phi_1(r_{t-1} - \mu) + a_t \tag{20}$$

Sostituendo iterativamente, la precedente equazione implica che

$$r_t - \mu = a_t + \phi_1 a_{t-1} + \phi_1^2 a_{t-2} + \dots = \sum_{i=0}^{+\infty} \phi_1^i a_{t-i}$$

Inoltre,  $r_t - \mu$  è una funzione lineare in  $a_{t-i}$  per  $i \geq 0$ . Usando questa proprietà e l'indipendenza della serie  $\{a_t\}$  otteniamo il valore atteso  $E[(r_{t-1} - \mu)a_t] = 0$ . Dall'ipotesi di stazionarietà, possiamo asserire che  $\text{Cov}(r_{t-1}, a_t) = E[(r_{t-1} - \mu)a_t] = 0$ . Questo risultato deriva dal fatto che  $r_{t-1}$  si realizza prima del tempo  $t$  e  $a_t$  non dipende dalle informazioni passate. Elevando al quadrato e calcolando il valore atteso della (20) si ottiene,

sfruttando il fatto che la covarianza tra  $r_t$  and  $a_t$  è zero

$$\text{Var}(r_t) = \phi_1^2 \text{Var}(r_{t-1}) + \sigma_a^2$$

dove  $\sigma_a^2$  è la varianza di  $a_t$ .

Sotto l'ipotesi di stazionarietà debole,  $\text{Var}(r_t) = \text{Var}(r_{t-1})$  e perciò

$$\text{Var}(r_t) = \frac{\sigma_a^2}{1 - \phi_1^2}.$$

Tuttavia, deve valere  $\phi_1^2 < 1$ . Si ha tale vincolo poiché la varianza di una variabile aleatoria deve essere sempre limitata e non negativa. In aggiunta, la stazionarietà debole del modello AR(1) implica che  $-1 < \phi_1 < 1$  è una condizione sufficiente. Se  $-1 < \phi_1 < 1$ , allora usando l'equazione (20), così come l'indipendenza della serie  $\{a_t\}$ , è possibile riscrivere il modello nella sua forma lineare e mostrare che media e varianza di  $r_t$  sono finite, soddisfacendo la condizione necessaria per l'esistenza della stazionarietà debole. Perciò, la condizione necessaria e sufficiente affinché il modello AR(1) sia debolmente stazionario è  $|\phi_1| < 1$ .

### 3.1.1 Funzione di autocorrelazione di un modello AR

Assumendo per semplicità che la media sia nulla ( $\phi_0 = 0$ ), utilizzando l'indipendenza tra  $\{a_t\}$  e  $\{r_{t-1}\}$  - come mostrato precedentemente - e calcolando il valore atteso dell'equazione (20), otteniamo

$$E[a_t(r_t - \mu)] = \phi_1 E[a_t(r_{t-1} - \mu)] + E(a_t^2) = E(a_t^2) = \sigma_a^2$$

dove  $\sigma_a^2$  è la varianza di  $a_t$ . Utilizzando questo risultato, moltiplicando l'equazione (20) per  $(r_{t-1} - \mu)$  e calcolando il suo valore atteso, possiamo ottenere l'autocovarianza a lag-1 del modello AR(1).

$$(r_t - \mu)(r_{t-1} - \mu) = \phi_1(r_{t-1} - \mu)(r_{t-1} - \mu) + a_t(r_{t-1} - \mu)$$

$$E((r_t - \mu)(r_{t-1} - \mu)) = \phi_1 E(r_{t-1} - \mu)(r_{t-1} - \mu) + E(a_t(r_{t-1} - \mu))$$

Utilizzando l'equazione sovrastante, congiuntamente al fatto che  $E[a_t(r_t - \mu)] = \sigma_a^2$ , quando il rumore bianco e l'osservazione sono nel medesimo time-frame, troviamo che l'autocovarianza a lag-1 è

$$\gamma_l = \begin{cases} \phi_1 \gamma_1 + \sigma_a^2 & \text{se } l = 0 \\ \phi_1 \gamma_{l-1} & \text{se } l \geq 1 \end{cases}$$

Assumendo che il modello AR(1) sia debolmente stazionario, è possibile trovare un'equazione che colleghi direttamente il coefficiente di correlazione  $\rho_l$  e il coefficiente  $\phi_1$ .

$$\gamma_1 = E(r_t r_{t+1}) = E(r_t(\phi_1 r_t + a_t)) = \phi_1 E(r_t^2) + E(r_t a_t) = \phi_1 \text{Var}(r_t)$$

$$\rho_1 = \frac{\text{Cov}(r_t, r_{t+1})}{\text{Var}(r_t)} = \frac{\phi_1 \text{Var}(r_t)}{\text{Var}(r_t)} = \phi_1$$

Utilizzando queste equazioni, così come la funzione di autocovarianza a lag-1, è possibile trovare l'equazione  $\gamma_l = \phi_1^l \gamma_0$ . Per definizione  $\gamma_l = \phi_1^l \text{Var}(r_t)$ , perciò  $\gamma_l = \phi_1^l \text{Var}(r_t)$ . La correlazione risulta perciò

$$\rho_l = \frac{\gamma_l}{\gamma_0} = \frac{\phi_1^l \text{Var}(r_t)}{\text{Var}(r_t)} = \phi_1^l.$$

Questo risultato indica che una serie AR(1) debolmente stazionaria decade esponenzialmente con tasso  $\phi_1$ , assumendo come valore iniziale  $\rho_0 = 1$ . Per un  $\phi_1$  positivo, il grafico della funzione di autocorrelazione (ACF) di un modello AR(1) mostra infatti un marcato decadimento esponenziale. Per un  $\phi_1$  negativo, il grafico oscilla tra valori positivi e negativi che si riducono in ampiezza come due decadimenti esponenziali alternati con tasso  $\phi_1^2$ .

### 3.2 Media Mobile (MA)

Il modello a media mobile (*Moving Average*, MA) si caratterizza per il fatto di prendere in considerazione la possibilità che ci sia una relazione tra la variabile  $\{r_t\}$  e i residui  $\{a_{t-i}\}$  per  $i = (1, 2, \dots, q)$ . Possiamo esprimere la media mobile di ordine q (MA(q)) nel modo seguente

$$r_t = \mu + a_t + \sum_{i=1}^q \theta_i a_{t-i}$$

### 3.2.1 Proprietà dei modelli MA

Principalmente, ci concentreremo sui modelli a media mobile MA(1) ed MA(2), dai quali le proprietà dei modelli MA(q) possono essere desunte attraverso le medesime procedure.

#### Stazionarietà

I modelli MA sono sempre debolmente stazionari in quanto sono combinazioni lineari finite di sequenze di rumori bianchi per i quali i primi due momenti sono invarianti rispetto al tempo (poiché la sequenza di rumori bianchi è iid). Per esempio, se consideriamo l'attesa e la varianza di un modello MA(1), otteniamo

$$\begin{aligned} E(r_t) &= \mu + E(a_t) + \theta_1 E(a_{t-1}) = \mu \\ \text{Var}(r_t) &= \sigma_a^2 + \theta_1^2 \sigma_a^2 = (1 + \theta_1^2) \sigma_a^2 \end{aligned}$$

poiché l'attesa della sequenza di rumori bianchi è nulla. Questa proprietà si applica facilmente a tutti i modelli MA(q). La varianza di un modello MA(q) si presenta sempre nella seguente forma

$$\text{Var}(r_t) = (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2) \sigma_a^2$$

#### Funzione di Autocorrelazione

Per semplicità, assumiamo che  $\mu = 0$  per un certo modello MA(1). Per mostrare l'autocorrelazione a lag-1, moltiplichiamo il modello per  $r_{t-1}$

$$r_{t-1}r_t = r_{t-1}a_t + \theta_1 r_{t-1}a_{t-1}$$

calcoliamone ora l'attesa e otteniamo

$$\gamma_1 = \theta_1 \sigma_a^2, \text{ and } \gamma_l = 0, \text{ per } l > 0.$$

Questo risultato, combinato al fatto che  $\text{Var}(r_t) = (1 + \theta_1^2) \sigma_a^2$ , permette di individuare i coefficienti di correlazione

$$\rho_0 = 1, \quad \rho_1 = \frac{\theta_1}{1 + \theta_1^2}, \quad \rho_l = 0 \text{ per } l > 0$$

Si noti che la funzione di autocorrelazione (ACF) per ogni lag-1 superiore ad 1 è nulla. Significa che l'ACF di un modello MA(1) è troncata al lag-1. Osservando un modello MA(2), invece, notiamo che

$$\rho_1 = \frac{\theta_1 + \theta_1 \theta_2}{1 + \theta_1^2 + \theta_2^2}, \quad \rho_2 = \frac{\theta_2}{1 + \theta_1^2 + \theta_2^2}, \quad \rho_l = 0 \text{ per } l > 2$$

L'ACF per un modello MA(2) è perciò troncata al lag-2. Questa proprietà può essere generalizzata per tutti i modelli MA, dove per un modello MA( $q$ ) l'ACF a lag- $q$  è non nulla e  $\rho_l = 0$  per  $l > q$ . Per questa ragione una serie MA( $q$ ) dipende unicamente dai suoi  $q$  valori precedenti e si tratta perciò di un modello a *memoria finita*.

La funzione di autocorrelazione nei modelli MA è utile per determinare il loro ordine. Se l'ACF di una serie temporale  $r_t$  ha  $\rho_q \neq 0$ , ma  $\rho_l = 0$  per  $l > q$ , allora  $r_t$  segue un modello MA( $q$ ).

### 3.3 Modelli Autoregressivi a Media Mobile (ARMA)

Una serie temporale  $r_t$  segue un modello ARMA(1,1) se soddisfa la seguente equazione

$$r_t - \phi_1 r_{t-1} = \phi_0 + a_t + \theta_1 a_{t-1} \quad (21)$$

dove  $\{a_t\}$  è una serie di rumori bianchi. La parte sinistra dell'equazione rappresenta la componente AR del modello, mentre la parte destra quella MA. Affinché questo modello abbia significato,  $\phi_1 \neq \theta_1$  altrimenti vi sarebbe una cancellazione diretta nell'equazione che renderebbe il processo una semplice serie di rumori bianchi con  $(1 - \phi B)r_t = (1 - \theta B)a_t$ .

#### 3.3.1 Proprietà dei modelli ARMA(1,1)

Le proprietà dei modelli ARMA(1,1) sono molto simil a quelle dei modelli AR(1) con alcune differenze dovute alla componente MA. Iniziamo considerando il valore atteso dell'equazione (21), si ottiene

$$E(r_t) - \phi_1 E(r_t) = \phi_0 + E(a_t) - \theta_1 E(a_{t-1}).$$

Dato che  $E(a_i) = 0 \forall i$ , la media  $r_t$  è quella del modello AR(1)

$$E(r_t) = \mu = \frac{\phi_0}{1 - \phi_1}$$

sotto la condizione che la serie sia debolmente stazionaria. Calcoliamo ora l'ACF di  $r_t$ , assumendo per semplicità che  $\phi_0 = 0$ . Il primo passo consiste nel determinare la varianza. A tal fine consideriamo la seguente equazione,

ottenuta esstraendo il valore atteso dell'equazione (21) moltiplicata per  $a_t$ .

Si ottiene

$$E(r_t a_t) = E(a_t^2) - \theta_1 E(a_t a_{t-1}) = E(a_t^2) = \sigma_a^2$$

Riscriviamo ora il modello ARMA come

$$r_t = \phi_1 r_{t-1} + a_t - \theta_1 a_{t-1}$$

e calcoliamo la sua varianza

$$Var(r_t) = \phi_1^2 Var(r_{t-1}) + \sigma_a^2 + \theta_1 \sigma_a^2 + 2\phi_1 \theta_1 E(r_{t-1} a_{t-1})$$

Poichè  $E(r_{t-1} a_t) = 0$ , sostituendo nell'equazione  $E(r_t a_t) = \sigma_a^2$ , otteniamo

$$Var(r_t) - \phi_1^2 Var(r_{t-1}) = (1 + 2\phi_1 \theta_1 + \theta_1^2) \sigma_a^2$$

Sotto la condizione che la serie sia debolmente stazionaria, troviamo che

$$Var(r_t) = \frac{(1 - 2\phi_1 \theta_1 + \theta_1^2) \sigma_a^2}{1 - \phi_1^2}$$

Come visto per il modello AR(1), affinchè la serie temporale sia stazionaria, è necessario che  $\phi_1^2 < 1$ . Per ottenere l'ACF dobbiamo anche assumere che  $\phi_0 = 0$ .

Moltiplicando l'equazione (21) per  $r_{t-l}$ , si ottiene la seguente espressione

$$r_{t-l} r_t - \phi_1 r_{t-l} r_{t-1} = r_{t-l} a_t - \theta_1 r_{t-l} a_{t-1}$$

Per  $l=1$ , calcolando il valore atteso, si ottiene

$$E(r_t r_{t-1}) - \phi_1 E(r_{t-1} r_{t-1}) = E(r_{t-1} a_t) - \theta_1 E(r_{t-1} a_{t-1})$$

che equivale a

$$\gamma_1 - \phi_1 \gamma_0 = -\theta_1 \sigma_a^2$$

Per  $l=2$ , invece, si ha

$$E(r_t r_{t-2}) - \phi_1 E(r_{t-1} r_{t-2}) = E(r_{t-2} a_t) - \theta_1 E(r_{t-2} a_{t-1})$$

equivalente a

$$\gamma_2 - \gamma_1 = 0$$

che è identica ad una serie AR(1). La precedente equazione può essere infatti generalizzata come

$$\gamma_l - \phi_1 \gamma_{l-1} = 0 \quad \text{for } l > 0$$

In conclusione, il risultato precedente mostra che l'ACF di un modello ARMA(1,1) stazionario è

$$\rho_1 = \phi_1 - \frac{\phi_1 \sigma_a^2}{\gamma_0}, \quad \rho_l = \phi_1 \rho_{l-1} \quad \text{for } l > 0$$

L'ACF di un modello ARMA(1,1) si comporta in modo molto simile a quella del modello AR(1), con la differenza che il decadimento esponenziale inizia al lag-2. Tuttavia, l'ACF di un modello ARMA(1,1) non è troncata a nessun lag finito.

### 3.3.2 Modelli ARMA generalizzati

Un modello ARMA(p,q) si presenta nella forma

$$r_t = \phi_0 + \sum_{i=1}^p \phi_i r_{t-i} - \sum_{i=1}^q \theta_i a_{t-i}$$

dove  $\{a_t\}$  è una serie di rumori bianchi e  $p$  e  $q$  sono interi non negativi. I modelli AR e MA sono considerati dei casi particolari dei modelli ARMA(p,q) (rispettivamente nei casi  $p = 0$  o  $q = 0$ ). Con l'operatore lag (anche noto come operatore *back-shift*<sup>4</sup>) è possibile riscrivere il modello nella seguente forma

$$(1 - \phi_1 B - \dots - \phi_p B^p) r_t = \phi_0 + (1 - \theta_1 B - \dots - \theta_q B^q) a_t \quad \text{or}$$

$$\phi_p(B)(r_t - \mu) = \theta_q(B) a_t$$

Dove  $\phi_p(B)$  e  $\theta_q(B)$  sono degli operatori lag tali che  $\phi_p(B) = 1 - \phi_1 B - \dots - \phi_p B^p$  e  $\theta_q(B) = 1 - \theta_1 B - \dots - \theta_q B^q$ . Il lato sinistro e destro non sono altro che i polinomi dell'AR e del MA, rispettivamente. E' essenziale che non ci siano fattori comuni tra i polinomi dell'AR e del MA, altrimenti l'ordine del modello potrebbe ridursi.

Come nel modello AR puro, consideriamo l'equazione caratteristica del polinomio AR. Se tutte le soluzioni del modello caratteristico sono minori di uno, allora il modello ARMA è debolmente stazionario. In questo caso, la media del modello ARMA(p,q) è  $E(r_t) = \phi_0 / (1 - \phi_1 - \dots - \phi_p)$ .

---

<sup>4</sup>Hamilton, Douglas - *Time Series Analysis*, Princeton University Press (1994)

Un modello ARMA( $p, q$ ) è detto invertibile se la serie temporale  $\{r_t\}$  può essere riscritta nella seguente forma

$$\pi(B)r_t = \sum_{i=0}^{\infty} \pi_i r_{t-i} = a_t$$

dove  $\pi(B) = \sum_{i=0}^{\infty} \pi_i B^i = \frac{\phi(B)}{\theta(B)}$  e  $\sum_{i=0}^{\infty} |\pi_i| < \infty$ , con  $\pi_0 = 1$ .

Un processo ARMA è dunque invertibile se e solo se tutte le radici del polinomio del MA giacciono all'esterno della circonferenza a raggio unitario. Una condizione equivalente è che per ogni  $\{i = 1, 2, \dots, q\}$  valga  $|\theta_i| < 1$ , dove  $\theta_i$  rappresenta i pesi della componente MA del modello ARMA.

## 4 Misure di rischio

### 4.1 Distribuzione di profitti e perdite (P&L)

Introduciamo alcune notazioni che utilizzeremo nella trattazione:

- $P_t$  = valore (prezzo) di un asset al tempo  $t$  (noto), misurato in €
- $P_{t+1}$  = valore (prezzo) di un asset al tempo  $t + 1$  (incognito), in €
- $\Pi_{t+1} = P_{t+1} - P_t$  = profitto misurato in € durante il periodo di possesso
- $L_{t+1} = -\Pi_{t+1}$  = perdita misurata in € durante il periodo di possesso

Si noti che un profitto positivo  $\Pi_{t+1} > 0 \Rightarrow$  una perdita negativa  $L_{t+1} < 0$ ; un profitto negativo  $\Pi_{t+1} < 0 \Rightarrow$  una perdita positiva  $L_{t+1} > 0$ .

Le misure di rischio tipicamente sono definite in termini di perdite. Perciò un grande valore positivo per una misura di rischio indica una grande perdita positiva.

Volendo definire P&L in funzione dei returns definiti nella (1), si ha

$$\begin{aligned}R_{t+1} &= \frac{P_{t+1} - P_t}{P_t} \\ \Pi_{t+1} &= P_{t+1} - P_t = P_t R_{t+1} \\ L_{t+1} &= -P_t R_{t+1}\end{aligned}$$

Poichè al tempo  $t$  il valore futuro di  $P_{t+1}$  è incognito,  $\Pi_{t+1}$ ,  $L_{t+1}$  e  $R_{t+1}$  sono variabili aleatorie.

Per le variabili aleatorie  $X_1 = \Pi_{t+1}$ ,  $X_2 = L_{t+1}$  e  $X_3 = R_{t+1}$ , denotiamo con  $F_{X_i}$  la relativa funzione di ripartizione e con  $f_{X_i}$  la funzione di densità. Naturalmente queste tre distribuzioni sono collegate.

Assumiamo (a meno di differenti specifiche) che  $F_{X_i}$  e  $f_{X_i}$  siano funzioni continue conosciute.

### 4.2 Definizione delle misure di rischio

Nel misurare il rischio ci si pone l'obiettivo di confrontare una varietà di asset in modo da supportare il processo decisionale in un qualche modo.

Nella gestione del rischio, l'iter tipico presuppone anzitutto la quantificazione del rischio di capitale e la valutazione della sua adeguatezza. In seguito ci si avvale di strumenti di management atti a determinare l'opportuno strumento di protezione.

In questo paragrafo ci occuperemo del primo step, ossia della quantificazione del rischio attraverso il calcolo di stime (*risk measurement*) ottenute applicando certe realizzazioni ad alcuni stimatori (*risk measure*).

Tra le misure più utilizzate, si annoverano: volatilità (vol o  $\sigma$ ), value-at-risk (VaR) ed expected shortfall (ES).

#### 4.2.1 Volatilità

La *volatilità* è una misura di rischio atta a quantificare lo scostamento dalla perdita o dal return medio. Si tratta di una misura simmetrica e positiva, il cui uso è appropriato se perdite o returns hanno distribuzioni normali. Si tenga presente che la volatilità potrebbe non esistere (potrebbe ad esempio essere un numero non finito). In formule:

Volatilità di P&L

$$\sigma_{\Pi} = \left( \mathbb{E}[(\Pi_{t+1} - \mu_{\Pi})^2] \right)^{1/2} = \sigma_L = \left( \mathbb{E}[(L_{t+1} - \mu_L)^2] \right)^{1/2} \quad (22)$$

Volatilità dei returns

$$\sigma_R = \left( \mathbb{E}[(R_{t+1} - \mu_R)^2] \right)^{1/2} \quad (23)$$

Relazione tra  $\sigma_L$  e  $\sigma_R$

$$L_{t+1} = P_t R_{t+1} \Rightarrow \sigma_L = P_t \sigma_R \quad (24)$$

#### 4.2.2 Value-at-Risk

Indichiamo con  $F_L$  la funzione di ripartizione della perdita  $L_{t+1}$  su un asset detenuto per un certo periodo di tempo.

**Definizione** Dato un certo livello di confidenza  $\alpha \in (0, 1)$ , il VaR su un

asset al livello di confidenza  $\alpha$  è dato dal più piccolo  $l$  tale che la probabilità di perdita al tempo  $t + 1$  sia inferiore a  $(1 - \alpha)$ .

$$\text{VaR}_\alpha = \inf\{l \in \mathbb{R} : \mathbb{P}(L_{t+1} > l) \leq 1 - \alpha\} \quad (25)$$

Se  $F_L$  è continua, allora il  $\text{VaR}_\alpha$  può essere definito implicitamente usando

$$\mathbb{P}(L_{t+1} \geq \text{VaR}_\alpha) = 1 - \alpha \quad (26)$$

oppure

$$F_L(\text{VaR}_\alpha) = \mathbb{P}(L_{t+1} \leq \text{VaR}_\alpha) = \alpha \quad (27)$$

Il  $\text{VaR}_\alpha$  può essere definito anche attraverso la funzione quantile  $F_L^{-1}$  (sotto l'ipotesi che la funzione di ripartizione  $F$  sia invertibile) a partire dalla (27)

$$\text{VaR}_\alpha = F_L^{-1}(\alpha) = q_\alpha^L$$

Tipicamente  $\alpha = 0.90, 0.95$  o  $0.99$ . Se  $\alpha = 0.95$  allora col 95% di confidenza (ossia con una probabilità del 5%) possiamo perdere  $\text{VaR}_{0.95}$  o di più;  $\text{VaR}_\alpha$  è un *lower bound* della possibile perdita che potrebbe accadere con una confidenza al livello  $\alpha$  (ossia con una probabilità  $1 - \alpha$ ). Alternativamente il  $\text{VaR}_\alpha$  indica il valore massimo che mi posso aspettare di perdere con una probabilità  $\alpha$ .

### Esempio

Calcolo del VaR nel caso di perdite/returns distribuiti come normali. Sia  $L_{t+1} \sim N(\mu_L, \sigma_L^2)$  dove  $\mu_L$  e  $\sigma_L$  sono note. Conosciamo perciò la sua funzione di distribuzione (9) e, integrando, quella di ripartizione.

Allora, dato un certo intervallo di confidenza  $\alpha \in (0, 1)$

$$\text{VaR}_\alpha = F_L^{-1}(\alpha, \mu_L, \sigma_L) = q_\alpha^L$$

dove  $F_L^{-1}(\cdot; \mu_L, \sigma_L)$  è la funzione quantile della distribuzione normale con media  $\mu_L$  e deviazione standard  $\sigma_L$ <sup>5</sup>.

Se  $L_{t+1} \sim N(\mu_L, \sigma_L^2)$ , allora

$$\text{VaR}_\alpha = q_\alpha^L = \mu_L + \sigma_L \cdot q_\alpha^Z$$

---

<sup>5</sup>Si noti che  $F_L^{-1}(\cdot; \mu_L, \sigma_L)$  non ha una forma chiusa ma una soluzione può essere facilmente determinata in modo numerico. Ad esempio con il software R è possibile utilizzare la funzione `qnorm()`.

dove  $q_\alpha^Z$  è il quantile  $\alpha$ -esimo di una normale standard definito come

$$F_Z^{-1}(\alpha) = \Phi^{-1}(\alpha) = q_\alpha^Z \text{ tale che } \Phi(q_\alpha^Z) = \alpha$$

dove

$$F_Z(x) = \Phi(x) = \mathbb{P}(Z \leq x) \text{ e } Z \sim N(0, 1)$$

### 4.2.3 Expected Shortfall

Denotiamo con  $F_L$  la distribuzione della perdita  $L_{t+1}$  su un asset in un certo periodo di tempo e assumiamo che  $F_L$  sia continua.

**Definizione** *L'Expected Shortfall (ES) al livello di confidenza  $\alpha$  è la perdita attesa condizionata al fatto che la perdita sia maggiore del  $\text{VaR}_\alpha$ .*

$$\text{ES}_\alpha = \mathbb{E}[L_{t+1} | L_{t+1} \geq \text{VaR}_\alpha] \quad (28)$$

In altre parole, ES è la perdita attesa nella coda superiore della distribuzione della perdita.

Se  $F_L$  non è continua, potrebbe essere impossibile calcolare un valore atteso per certi  $\alpha$ , perciò l' $\text{ES}_\alpha$  in questa circostanza è definito come

$$\text{ES}_\alpha = \frac{1}{1 - \alpha} \int_\alpha^1 \text{VaR}_u du$$

che è la media del  $\text{VaR}_u$  su tutti gli  $u$  maggiori o uguali ad  $\alpha \in (0, 1)$ .

Per calcolare  $\text{ES}_\alpha = \mathbb{E}[L_{t+1} | L_{t+1} \geq \text{VaR}_\alpha]$ , bisogna calcolare il valore atteso della distribuzione troncata della perdita

$$\begin{aligned} \text{ES}_\alpha &= \mathbb{E}[L_{t+1} | L_{t+1} \geq \text{VaR}_\alpha] \\ &= \int_{-\infty}^{+\infty} l \mathbb{P}[L_{t+1} = l | L_{t+1} \geq \text{VaR}_\alpha] dl \\ &= \int_{-\infty}^{+\infty} l \frac{\mathbb{P}[L_{t+1} \geq \text{VaR}_\alpha | L_{t+1} = l] \cdot \mathbb{P}[L_{t+1} = l]}{\mathbb{P}[L_{t+1} \geq \text{VaR}_\alpha]} dl \\ &= \int_{\text{VaR}_\alpha}^{+\infty} l \frac{\mathbb{P}[L_{t+1} = l]}{\mathbb{P}[L_{t+1} \geq \text{VaR}_\alpha]} dl \\ &= \frac{1}{1 - \alpha} \int_{\text{VaR}_\alpha}^{+\infty} l f_L(l) dl \end{aligned}$$

In termini del profitto e dei returns si ha

$$\begin{aligned} \text{ES}_\alpha &= -\mathbb{E}[\Pi_{t+1} | \Pi_{t+1} \leq -\text{VaR}_\alpha] \\ \text{ES}_\alpha &= -P_t \cdot \mathbb{E}[-R_{t+1} | -R_{t+1} \leq q_\alpha^{-R}] \end{aligned}$$

### Esempio

Calcolare ES quando perdite/returns seguono una distribuzione normale.

Sia  $L_{t+1} \sim N(\mu_L, \sigma_L^2)$  dove  $\mu_L$  e  $\sigma_L$  sono noti. Per il livello di confidenza  $\alpha$

$$\begin{aligned} \text{ES}_\alpha &= \mathbb{E}[L_{t+1} | L_{t+1} \geq \text{VaR}_\alpha] \\ &= \text{valore atteso di una distribuzione normale troncata} \\ &= \mu_L + \sigma_L \frac{\phi(q_\alpha^Z)}{1 - \alpha} \end{aligned}$$

dove  $\phi(z) = f_Z(z) = \text{pdf of } Z \sim N(0, 1)$

### 4.3 Proprietà delle misure di rischio

Indichiamo con  $\text{RM}(\cdot)$  una generica misura di rischio, ad esempio Volatilità, Value at Risk o Expected Shortfall.

**Definizione** (Misura di rischio coerente<sup>6</sup>) *Si considerino due variabili aleatorie  $X$  e  $Y$  (entrambe a supporto  $\Omega$  e spazio campionario  $V \in \mathbb{R}$ ) rappresentanti la perdita di due asset posseduti. Una funzione  $\text{RM}(\cdot) : X, Y \rightarrow \mathbb{R}$  è chiamata misura di rischio coerente se soddisfa per  $X, Y$  e una costante  $C$ :*

#### 1. Monotonicità

$$X, Y : \Omega \rightarrow V \subset \mathbb{R}, F_X(z) \geq F_Y(z) \forall z \in \Omega \Rightarrow \text{RM}(X) \geq \text{RM}(Y)$$

Se la perdita di  $X$  è sempre maggiore di quella di  $Y$ , il rischio di  $X$  deve sempre eccedere quello di  $Y$ .

---

<sup>6</sup>Artzner P., Delbaen F., Eber J. e Heath D. - *Coherent Measures of Risk*, Mathematical Finance (1999), analizza quali proprietà dovrebbero avere le misure di rischio per essere considerate sensibili e utili. Il testo identifica quattro assiomi che le misure di rischio dovrebbero soddisfare per essere considerate *coerenti*.

## 2. Subaddittività

$$X, Y, X + Y : \Omega \rightarrow V \in \mathbb{R} \Rightarrow \text{RM}(X + Y) \leq \text{RM}(X) + \text{RM}(Y)$$

Il rischio del portafoglio di  $X$  e  $Y$  non può essere peggiore che la somma dei due rischi individuali - una manifestazione del principio di diversificazione.

## 3. Positiva omogeneità

$$X : \Omega \rightarrow V \subset \mathbb{R}, c > 0 \Rightarrow \text{RM}(cX) = c \text{RM}(X)$$

Per esempio, se l'asset raddoppia di valore ( $c = 2$ ), allora il rischio raddoppia di conseguenza. Questa proprietà è però spesso violata nelle applicazioni pratiche per  $c$  molto grandi, a causa di effetti di liquidità.

## 4. Invarianza rispetto alle traslazioni

$$X : \Omega \rightarrow V \subset \mathbb{R}, c \in \mathbb{R} \Rightarrow \text{RM}(X + c) = \text{RM}(X) + c$$

Per esempio, aggiungere  $c < 0$  alla perdita è come aggiungere denaro liquido - un po' come un'assicurazione - in modo che il rischio di  $X + c$  sia minore del rischio di  $X$  per un ammontare di denaro pari a  $c$ .

Può essere mostrato<sup>7</sup> che la misura ES è coerente, mentre volatilità e VaR non lo sono. In particolare il VaR soddisfa sempre tutti i criteri a meno della subaddittività che solo talvolta è verificata.

Mostriamo due esempi di verifica della proprietà di subaddittività.

### **Esempio**

La volatilità è subadditiva.

Consideriamo un portafoglio di due titoli  $X$  e  $Y$  con ritorni  $R_x$  e  $R_y$  e pesi di portafoglio  $w_x > 0$  e  $w_y > 0$  tali che  $w_x + w_y = 1$ . Allora

$$\sigma_p = \sqrt{w_x^2 \sigma_x^2 + w_y^2 \sigma_y^2 + 2w_x w_y \sigma_{xy}} \leq w_x \sigma_x + w_y \sigma_y$$

---

<sup>7</sup> Artzner P., Delbaen F., Eber J. e Heath D. - *Coherent measures of risk*, Mathematical Finance (1999)

Ossia, la volatilità del portafoglio è minore della media ponderata delle volatilità dei singoli asset.

Per dimostrarlo, anzitutto esplicitiamo la volatilità del portafoglio

$$\sigma_p^2 = w_x^2 \sigma_x^2 + w_y^2 \sigma_y^2 + 2w_x w_y \sigma_x \sigma_y \rho_{xy} = w_x^2 \sigma_x^2 + w_y^2 \sigma_y^2 + 2w_x w_y \sigma_x \sigma_y \rho_{xy}$$

Ora aggiungiamo e sottraiamo  $2w_x w_y \sigma_x \sigma_y$

$$\begin{aligned} \sigma_p^2 &= (w_x^2 \sigma_x^2 + w_y^2 \sigma_y^2 + 2w_x w_y \sigma_x \sigma_y) - 2w_x w_y \sigma_x \sigma_y + 2w_x w_y \sigma_x \sigma_y \rho_{xy} \\ &= (w_x \sigma_x + w_y \sigma_y)^2 - 2w_x w_y \sigma_x \sigma_y (1 - \rho_{xy}) \end{aligned}$$

Poiché

$$2w_x w_y \sigma_x \sigma_y (1 - \rho_{xy}) > 0$$

Segue che

$$\sigma_p = \sqrt{(w_x \sigma_x + w_y \sigma_y)^2 - 2w_x w_y \sigma_x \sigma_y (1 - \rho_{xy})} \leq w_x \sigma_x + w_y \sigma_y$$

e di conseguenza la volatilità del portafoglio è subadditiva.

Ad ogni modo, la volatilità di portafoglio non è invariante rispetto alle traslazioni, perciò non è una misura di rischio coerente. Infatti

$$\text{var}(R + c) = \text{var}(R) \quad \forall c \in \mathbb{R}$$

### Esempio

VaR non è subadditiva.

Consideriamo due investitori che operano con derivati riferiti alla stessa impresa:

1. vendita allo scoperto<sup>8</sup> di un'opzione<sup>9</sup> call che è *out of the money*<sup>10</sup>, con un 4% di probabilità di perdere denaro;

<sup>8</sup>John C. Hull - *Opzioni futures e altri derivati*, Springer (2015): la *vendita allo scoperto*, in inglese *short selling*, è un'operazione finanziaria che consiste nella vendita di titoli non direttamente posseduti dal venditore.

<sup>9</sup>Fredric S. Mishkin, Stanley G. Eakins - *Istituzioni e mercati finanziari*, Pearson (2012): un'opzione call è un contratto che dà al compratore il diritto di acquistare uno strumento finanziario al prezzo di esercizio (strike price) entro un determinato periodo di tempo; un'opzione put è un contratto che dà al compratore il diritto di vendere uno strumento finanziario al prezzo di esercizio entro un determinato periodo di tempo.

<sup>10</sup>Fredric S. Mishkin, Stanley G. Eakins - *Istituzioni e mercati finanziari*, Pearson (2012): un'opzione si dice *out of the money* quando il prezzo dello strumento finanziario sottostante è inferiore (per la call) o superiore (per la put) allo strike price.

2. vendita allo scoperto di un'opzione *put* che è *out of the money*, con un 4% di probabilità di perdere denaro.

Nessuna delle due posizioni ha un VaR al livello del 5% poiché - per entrambe - una qualsiasi perdita è possibile solo col 4% di probabilità. Si avrebbe un VaR positivo solo se la probabilità di perdere denaro fosse maggiore del 5%.

Nonostante questo, per la posizione complessiva potrebbe esserci un VaR al livello 5%. Per mostrare questo fatto, ipotizziamo l'indipendenza delle due posizioni. Allora

$$\begin{aligned}\mathbb{P}(\text{almeno un investitore perde denaro}) &= 1 - \mathbb{P}(\text{nessuno perde denaro}) \\ &= 1 - (0.96)^2 = 0.078 > 5\%\end{aligned}$$

Perciò VaR al 5% non è subadditivo.

## 5 Pianificazione del rischio di portafoglio

La pianificazione del rischio di portafoglio - *Portfolio Risk Budgeting* - consiste della decomposizione delle misure di rischio di portafoglio nelle rispettive contribuzioni asset per asset, in modo da permettere al gestore di conoscere le fonti di rischio degli asset a fini di copertura e ripartizione ottimale del capitale.

Il risk manager ha così la possibilità di valutare il portafoglio dal punto di vista dell'asset risk.

### 5.1 Calcolo di portafoglio

Siano  $R_1, \dots, R_n$  i *simple return* - definiti in (1) - di  $n$  differenti asset e siano  $w_1, \dots, w_n$  i pesi di portafoglio, tali che  $\sum_{i=1}^n w_i = 1$ .

Definiamo perciò il *portfolio return* come

$$R_p = \mathbf{w}'\mathbf{R} = \sum_{i=1}^N w_i R_i, \quad \mathbf{w}'\mathbf{1} = 1 \quad (29)$$

dove  $\mathbf{R} = (R_1, \dots, R_N)'$ ,  $\mathbf{w} = (w_1, \dots, w_n)'$  e  $\mathbf{1} = (1, \dots, 1)'$ .

#### 5.1.1 Media e varianza di portafoglio

Sia  $\mathbf{R}$  un vettore aleatorio con

$$\begin{aligned} \mathbb{E}[\mathbf{R}] &= \boldsymbol{\mu} = (\mu_1, \dots, \mu_n)' \\ \text{var}(\mathbf{R}) &= \mathbb{E}[(\mathbf{R} - \boldsymbol{\mu})(\mathbf{R} - \boldsymbol{\mu})'] = \Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1n} \\ \sigma_{12} & \sigma_2^2 & \dots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1n} & \sigma_{2n} & \dots & \sigma_n^2 \end{pmatrix} \end{aligned}$$

Allora

$$\mu_p = \mathbf{w}'\boldsymbol{\mu}, \quad \sigma_p^2 = \text{var}(R_p) = \text{var}(\mathbf{w}'\mathbf{R}) = \mathbf{w}'\Sigma\mathbf{w}$$

#### Esempio

Decomposizione del rischio di portafoglio nel caso di due asset

$$\begin{aligned}
R_p &= w_1 R_1 + w_2 R_2 \\
\sigma_p^2 &= w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2 + 2w_1 w_2 \sigma_{12} \\
\sigma_p &= (w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2 + 2w_1 w_2 \sigma_{12})^{1/2}
\end{aligned}$$

Per ottenere una decomposizione additiva per  $\sigma_p^2$ , scriviamo

$$\begin{aligned}
\sigma_p^2 &= w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2 + 2w_1 w_2 \sigma_{12} \\
&= (w_1^2 \sigma_1^2 + w_1 w_2 \sigma_{12}) + (w_2^2 \sigma_2^2 + w_1 w_2 \sigma_{12})
\end{aligned}$$

In questo modo abbiamo diviso il contributo della covarianza  $2w_1 w_2 \sigma_{12}$  di portafoglio, tra i due asset. Definiamo

$$\begin{aligned}
w_1^2 \sigma_1^2 + w_1 w_2 \sigma_{12} &= \text{contributo dell'asset 1 sulla varianza} \\
w_2^2 \sigma_2^2 + w_1 w_2 \sigma_{12} &= \text{contributo dell'asset 2 sulla varianza}
\end{aligned}$$

Possiamo anche definire una decomposizione additiva per  $\sigma_p$

$$\sigma_p = \frac{w_1^2 \sigma_1^2 + w_1 w_2 \sigma_{12}}{\sigma_p} + \frac{w_2^2 \sigma_2^2 + w_1 w_2 \sigma_{12}}{\sigma_p}$$

dove i due addendi rappresentano i contributi dell'asset 1 e 2 alla deviazione standard del portafoglio.

### 5.1.2 Teorema di Eulero e decomposizione del rischio

Quando utilizziamo  $\sigma_p$  per misurare il rischio di portafoglio, siamo in grado di derivare facilmente una decomposizione additiva del rischio. Se, invece, il rischio viene misurato usando VaR o ES, la determinazione dei contributi di rischio individuali può non essere agevole da calcolare.

Il teorema di Eulero fornisce un metodo per la decomposizione del rischio nel caso in cui la *risk measure* di portafoglio sia una funzione omogenea dei pesi di primo grado.

Iniziamo col definire una funzione omogenea di primo grado.

**Definizione** (funzione omogenea di primo grado) *Sia  $f(w_1, \dots, w_n)$  una funzione continua e differenziabile.  $f$  è detta omogenea di primo grado se per*

ogni costante  $c > 0$ , vale  $f(c \cdot w_1, \dots, c \cdot w_n) = c \cdot f(w_1, \dots, w_n)$ .

### Esempio

- Sia  $f(w_1, w_2) = w_1 + w_2$ . Allora  
 $f(c \cdot w_1, c \cdot w_2) = c \cdot w_1 + c \cdot w_2 = c \cdot f(w_1, w_2)$
- Sia  $f(w_1, w_2) = w_1^2 + w_2^2$ . Allora  
 $f(c \cdot w_1, c \cdot w_2) = c^2 \cdot w_1^2 + c^2 \cdot w_2^2 = c^2 \cdot (w_1^2, w_2^2) \neq c \cdot f(w_1, w_2)$
- Sia  $f(w_1, w_2) = \sqrt{w_1^2 + w_2^2}$ . Allora  
 $f(c \cdot w_1, c \cdot w_2) = \sqrt{c^2 \cdot w_1^2 + c^2 \cdot w_2^2} = c \cdot \sqrt{(w_1^2, w_2^2)} = c \cdot f(w_1, w_2)$

Consideriamo ora un portafoglio di  $n$  asset con pesi  $w = (w_1, \dots, w_n)'$  e valore iniziale  $V_0$ . Indichiamo con  $\alpha \in (0, 1)$  un certo livello di confidenza.

$$\mathbf{R} = (R_1, \dots, R_n)' , \quad \mathbf{w} = (w_1, \dots, w_n)'$$

$$\mathbb{E}[\mathbf{R}] = \boldsymbol{\mu} , \quad \text{cov}(\mathbf{R}) = \Sigma , \quad \mathbf{R} \sim N(\boldsymbol{\mu}, \Sigma)$$

Avendo definito

$$R_p = R_p(\mathbf{w}) = \mathbf{w}'\mathbf{R}$$

$$\mu_p = \mu_p(\mathbf{w}) = \mathbf{w}'\boldsymbol{\mu} , \quad \sigma_p^2 = \sigma_p^2(\mathbf{w}) = \mathbf{w}'\Sigma\mathbf{w} ,$$

$$\sigma_p = \sigma_p(\mathbf{w}) = (\mathbf{w}'\Sigma\mathbf{w})^{1/2}$$

$$q_{1-\alpha}^{R_p} = q_{1-\alpha}^{R_p}(\mathbf{w}) = \mu_p(\mathbf{w}) + \sigma_p(\mathbf{w}) q_{1-\alpha}^Z$$

Le misure di rischio VaR ed ES risultano

$$\text{VaR}(\mathbf{w}) = -q_{1-\alpha}^{R_p}(\mathbf{w}) P_0$$

$$\text{ES}_\alpha(\mathbf{w}) = -P_0 \left( \mu_p(\mathbf{w}) + \sigma(\mathbf{w}) \frac{\phi(q_{1-\alpha}^Z)}{1-\alpha} \right)$$

Il return di portafoglio  $R_p(\mathbf{w})$ , il return atteso  $\mu_p(\mathbf{w})$ , la deviazione standard  $\sigma_p$ , il quantile normale  $q_{1-\alpha}^{R_p}(\mathbf{w})$ ,  $\text{VaR}_\alpha(\mathbf{w})$  e  $\text{ES}_\alpha(\mathbf{w})$ <sup>11</sup> sono funzioni

<sup>11</sup>Per il calcolo delle misure di rischio, sono stati assunti return normali, anche se è possibile dimostrare più in generale l'omogeneità lineare delle misure di rischio per altri tipi di distribuzione. Si veda S. Rachev - *Desirable Properties of an Ideal Risk Measure in Portfolio Theory*, University of California (2008).

omogenee di primo grado.

Perciò se  $\text{RM}(\mathbf{w})$  - che denota una misura di rischio - è una funzione omogenea lineare di  $\mathbf{w}$ , allora  $\text{RM}(c \cdot \mathbf{w}) = c \cdot \text{RM}(\mathbf{w}) \quad \forall c \geq 0$

**Teorema** (di Eulero<sup>12</sup>)

Sia  $f(w_1, \dots, w_n) = f(\mathbf{w})$  una funzione continua, differenziabile e omogenea di primo grado delle variabili  $w = (w_1, \dots, w_n)'$ . Allora

$$f(\mathbf{w}) = w_1 \cdot \frac{\partial f(\mathbf{w})}{\partial w_1} + w_2 \cdot \frac{\partial f(\mathbf{w})}{\partial w_2} + \dots + w_n \cdot \frac{\partial f(\mathbf{w})}{\partial w_n} = \mathbf{w}' \frac{\partial f(\mathbf{w})}{\partial \mathbf{w}} \quad (30)$$

dove

$$\frac{\partial f(\mathbf{w})}{\partial \mathbf{w}} = \begin{pmatrix} \frac{\partial f(\mathbf{w})}{\partial w_1} \\ \vdots \\ \frac{\partial f(\mathbf{w})}{\partial w_n} \end{pmatrix}$$

Verifichiamo ora il teorema in alcuni casi semplici.

#### Verifica del Teorema di Eulero - Caso 1

La funzione  $f(w_1, w_2) = w_1 + w_2 = f(\mathbf{w}) = \mathbf{w}'\mathbf{1}$  è omogenea di primo grado (si veda esempio a pagina precedente), inoltre

$$\begin{aligned} \frac{\partial f(\mathbf{w})}{\partial w_1} &= \frac{\partial f(\mathbf{w})}{\partial w_2} = 1 \\ \frac{\partial f(\mathbf{w})}{\partial \mathbf{w}} &= \begin{pmatrix} \frac{\partial f(\mathbf{w})}{\partial w_1} \\ \frac{\partial f(\mathbf{w})}{\partial w_2} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \mathbf{1} \end{aligned}$$

Per il teorema di Eulero

$$f(\mathbf{w}) = w_1 \cdot 1 + w_2 \cdot 1 = w_1 + w_2 = \mathbf{w}'\mathbf{1}$$

#### Verifica del Teorema di Eulero - Caso 2

La funzione  $f(w_1, w_2) = (w_1^2 + w_2^2)^{1/2} = f(\mathbf{w}) = (\mathbf{w}'\mathbf{w})^{1/2}$  è omogenea di primo grado (si veda esempio a pagina precedente), inoltre

$$\begin{aligned} \frac{\partial f(\mathbf{w})}{\partial w_1} &= \frac{1}{2}(w_1^2 + w_2^2)^{-1/2} 2w_1 = w_1(w_1^2 + w_2^2)^{-1/2} \\ \frac{\partial f(\mathbf{w})}{\partial w_2} &= \frac{1}{2}(w_1^2 + w_2^2)^{-1/2} 2w_2 = w_2(w_1^2 + w_2^2)^{-1/2} \end{aligned}$$

---

<sup>12</sup>Glasserman P. - *Measuring marginal risk contributions in credit portfolios*, The Journal of Computational Finance (2005) Vol. 9, No. 2, 1-41

Per il teorema di Eulero

$$\begin{aligned}
 f(\mathbf{w}) &= w_1 \cdot w_1(w_1^2 + w_2^2)^{-1/2} + w_2 \cdot w_2(w_1^2 + w_2^2)^{-1/2} \\
 &= (w_1^2 + w_2^2) (w_1^2 + w_2^2)^{-1/2} \\
 &= (w_1^2 + w_2^2)^{1/2}
 \end{aligned}$$

## 5.2 Risk Budgeting

Nella sezione precedente abbiamo mostrato come le misure di rischio introdotte siano funzioni omogenee lineari di  $\mathbf{w}$ . Grazie al teorema di Eulero otteniamo quindi

$$\begin{aligned}
 \text{RM}(\mathbf{w}) &= \sum_{i=1}^n w_i \frac{\partial \text{RM}(\mathbf{w})}{\partial w_i} \\
 &= w_1 \frac{\partial \text{RM}(\mathbf{w})}{\partial w_1} + \dots + w_n \frac{\partial \text{RM}(\mathbf{w})}{\partial w_n}
 \end{aligned}$$

Convenzionalmente chiameremo

$$\frac{\partial \text{RM}(\mathbf{w})}{\partial w_i} \longrightarrow \text{contribuzione marginale al rischio dell'asset } i \quad (31)$$

$$w_i \frac{\partial \text{RM}(\mathbf{w})}{\partial w_i} \longrightarrow \text{contribuzione al rischio dell'asset } i \quad (32)$$

$$\frac{w_i \frac{\partial \text{RM}(\mathbf{w})}{\partial w_i}}{\text{RM}(\mathbf{w})} \longrightarrow \text{contribuzione percentuale al rischio dell'asset } i \quad (33)$$

### 5.2.1 Risultati analitici per le misure di rischio

Nel caso  $\text{RM}(\mathbf{w}) = \sigma(\mathbf{w})$ , si ha

$$\begin{aligned}
 R_p &= \mathbf{w}'\mathbf{R}, \quad \text{var}(\mathbf{R}) = \Sigma \\
 \sigma(\mathbf{w}) &= (\mathbf{w}'\Sigma\mathbf{w})^{1/2} \\
 \frac{\partial \sigma(\mathbf{w})}{\partial \mathbf{w}} &= \frac{1}{\sigma(\mathbf{w})} \Sigma\mathbf{w}
 \end{aligned}$$

Dove

$$\begin{aligned}
 \Sigma\mathbf{w} &= \begin{pmatrix} \text{cov}(R_1, R_p) \\ \vdots \\ \text{cov}(R_n, R_p) \end{pmatrix} = \sigma(\mathbf{w}) \begin{pmatrix} \beta_{1,p} \\ \vdots \\ \beta_{n,p} \end{pmatrix} \\
 \beta_{i,p} &= \text{cov}(R_i, R_p) / \sigma^2(\mathbf{w})
 \end{aligned}$$

Nel caso  $\text{RM}(\mathbf{w}) = \text{VaR}_\alpha(\mathbf{w})$ ,  $\text{ES}_\alpha(\mathbf{w})$ , si ha<sup>13</sup>

$$\begin{aligned}\frac{\partial \text{VaR}_\alpha(\mathbf{w})}{\partial w_i} &= \mathbb{E}[R_i | R_p = \text{VaR}_\alpha(\mathbf{w})], \quad i = 1, \dots, n \\ \frac{\partial \text{ES}_\alpha(\mathbf{w})}{\partial w_i} &= \mathbb{E}[R_i | R_p \leq \text{VaR}_\alpha(\mathbf{w})], \quad i = 1, \dots, n\end{aligned}$$

Proviamo a comprendere questo risultato intuitivamente: partendo dal ritorno di portafoglio definito in (29), si ha

$$\begin{aligned}\text{VaR}_\alpha(\mathbf{w}) &= \mathbb{E}[R_p | R_p = \text{VaR}_\alpha] = \sum_{i=1}^n w_i \mathbb{E}[R_i | R_p = \text{VaR}_\alpha] \\ \text{ES}_\alpha(\mathbf{w}) &= \mathbb{E}[R_p | R_p \leq \text{VaR}_\alpha] = \sum_{i=1}^n w_i \mathbb{E}[R_i | R_p \leq \text{VaR}_\alpha]\end{aligned}$$

Differenziando  $\text{VaR}_\alpha(\mathbf{w})$  e  $\text{ES}_\alpha(\mathbf{w})$  rispetto a  $w_i$ , si ottiene

$$\begin{aligned}\frac{\partial \text{VaR}_\alpha(\mathbf{w})}{\partial w_i} &= \mathbb{E}[R_i | R_p = \text{VaR}_\alpha] \\ \frac{\partial \text{ES}_\alpha(\mathbf{w})}{\partial w_i} &= \mathbb{E}[R_i | R_p \leq \text{VaR}_\alpha]\end{aligned}$$

### 5.3 Ottimizzazione di portafoglio

L'ottimizzazione di portafoglio svolta in modo standard necessita di un insieme di valori attesi dei returns e delle previsioni sui rischi. Questi valori sono processati da un algoritmo di ottimizzazione che produce i pesi di portafoglio atti a massimizzare un certo rapporto rischio-rendimento, di solito soggetti a certi vincoli. Al contrario, quando si parla di *ottimizzazione inversa*, s'intende quel procedimento che, a partire dai pesi di portafoglio e le previsioni di rischio, determina quali dovrebbero essere i returns attesi al fine di soddisfare l'ottimalità.

#### 5.3.1 Approccio standard

Supponiamo che l'obiettivo sia quello di costituire un portafoglio volto a massimizzare il valore atteso del rapporto ritorno-rischio (il cosiddetto *shar-*

<sup>13</sup>come mostrato - sotto ipotesi di normalità - da Gouriéroux (2000) e Scalliet (2002)

*pe ratio*, indicatore comunissimo in finanza per valutare le strategie di trading):

$$\max_{\mathbf{w}} \frac{\mu_p(\mathbf{w})}{\text{RM}(\mathbf{w})}$$

$$\mu_p(\mathbf{w}) = \mathbf{w}'\boldsymbol{\mu}$$

Dove  $\text{RM}(\mathbf{w})$  è una generica misura di rischio omogenea lineare.

Le condizioni necessarie e sufficienti di primo ordine per l'ottimalità sono

$$0 = \frac{\partial}{\partial w_i} \left( \frac{\mu_p(\mathbf{w})}{\text{RM}(\mathbf{w})} \right) = \frac{1}{\text{RM}(\mathbf{w})} \frac{\partial \mu_p(\mathbf{w})}{\partial w_i} - \frac{\mu_p(\mathbf{w})}{\text{RM}(\mathbf{w})^2} \frac{\partial \text{RM}(\mathbf{w})}{\partial w_i} \quad \forall i \quad (34)$$

### 5.3.2 Approccio inverso

L'ottimizzazione inversa utilizza la condizione (34) con pesi di portafoglio fissati per determinare i ritorni attesi ottimali. Questi ultimi sono chiamati *ritorni impliciti* e soddisfano

$$\mu_i^{impl}(\mathbf{w}) = \frac{\mu_p(\mathbf{w})}{\text{RM}(\mathbf{w})} \frac{\partial \text{RM}(\mathbf{w})}{\partial w_i}$$

Il ritorno implicito  $i$ -esimo risulta essere proporzionale al contributo marginale al rischio, con una costante di proporzionalità pari allo sharpe ratio. Perciò, dato un generico shape ratio di portafoglio,  $\mu_i^{impl}(\mathbf{w})$  è grande se  $\frac{\partial \text{RM}(\mathbf{w})}{\partial w_i}$  è grande.

Se il return attuale o quello previsto per l'asset  $i$ -esimo è inferiore al suo ritorno implicito, allora è necessario diminuire l'ammontare posseduto di quell'asset. Al contrario, se il ritorno atteso fosse maggiore di quello implicito, sarebbe necessario aumentare la quota posseduta.

## 5.4 Decomposizione delle misure di rischio

Denotiamo con  $\text{RM}(\mathbf{w})$  una misura di rischio omogenea lineare di primo grado per un certo portafoglio con pesi definiti dal vettore  $\mathbf{w}$ . In questo caso l'applicazione del teorema di Eulero si traduce nella (30), ossia

$$\text{RM}_p(\mathbf{w}) = w_1 \frac{\partial \text{RM}_p(\mathbf{w})}{\partial w_1} + w_2 \frac{\partial \text{RM}_p(\mathbf{w})}{\partial w_2} + \dots + w_n \frac{\partial \text{RM}_p(\mathbf{w})}{\partial w_n} \quad (35)$$

$$= \sum_{i=1}^n w_i \frac{\partial \text{RM}_p(\mathbf{w})}{\partial w_i} = \mathbf{w}' \frac{\partial \text{RM}_p(\mathbf{w})}{\partial \mathbf{w}} \quad (36)$$

Assegnamo una notazione per definire i rapporti (31), (32) e (33):

$$\text{MCR}_i^{\text{RM}} = \frac{\partial \text{RM}(\mathbf{w})}{\partial w_i}; \quad \text{CR}_i^{\text{RM}} = w_i \frac{\partial \text{RM}(\mathbf{w})}{\partial w_i}; \quad \text{PCR}_i^{\text{RM}} = \frac{w_i \frac{\partial \text{RM}(\mathbf{w})}{\partial w_i}}{\text{RM}(\mathbf{w})}.$$

Riscriviamo quindi la (35) con la notazione appena introdotta

$$\begin{aligned} \text{RM}_p(\mathbf{w}) &= w_1 \cdot \text{MCR}_1^{\text{RM}} + w_2 \cdot \text{MCR}_2^{\text{RM}} + \dots + w_n \cdot \text{MCR}_n^{\text{RM}} \\ &= \text{CR}_1^{\text{RM}} + \text{CR}_2^{\text{RM}} + \dots + \text{CR}_n^{\text{RM}} \\ 1 &= \frac{\text{CR}_1^{\text{RM}}}{\text{RM}_p(\mathbf{w})} + \dots + \frac{\text{CR}_n^{\text{RM}}}{\text{RM}_p(\mathbf{w})} = \text{PCR}_1^{\text{RM}} + \dots + \text{PCR}_n^{\text{RM}} \end{aligned}$$

#### 5.4.1 Decomposizione della volatilità per un portafoglio

$$\text{RM}_p(\mathbf{w}) = \sigma_p(\mathbf{w}) = (\mathbf{w}'\Sigma\mathbf{w})^{1/2} = \mathbf{w}' \frac{\partial \sigma_p(\mathbf{w})}{\partial \mathbf{w}} \quad (37)$$

Dove

$$\begin{aligned} \frac{\partial \sigma_p(\mathbf{w})}{\partial \mathbf{w}} &= \frac{\partial (\mathbf{w}'\Sigma\mathbf{w})^{1/2}}{\partial \mathbf{w}} = \frac{1}{2} (\mathbf{w}'\Sigma\mathbf{w})^{-1/2} 2\Sigma\mathbf{w} \\ &= \frac{\Sigma\mathbf{w}}{(\mathbf{w}'\Sigma\mathbf{w})^{1/2}} = \frac{\Sigma\mathbf{w}}{\sigma_p(\mathbf{w})} \end{aligned} \quad (38)$$

Per ottenere  $\frac{\Sigma\mathbf{w}}{\sigma_p(\mathbf{w})}$  è sufficiente calcolare le derivate parziali  $\frac{\partial \sigma_p(\mathbf{w})}{\partial w_i}$  al variare di  $i$ , che risultano essere le righe della matrice incognita.

#### Esempio

Decomposizione della misura di rischio volatilità per un portafoglio di due soli asset

$$\begin{aligned} \sigma_p(\mathbf{w}) &= (\mathbf{w}'\Sigma\mathbf{w})^{1/2} = (w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2\sigma_{12})^{1/2} \\ \Sigma\mathbf{w} &= \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} w_1\sigma_1^2 + w_2\sigma_{12} \\ w_2\sigma_2^2 + w_1\sigma_{12} \end{pmatrix} \\ \frac{\Sigma\mathbf{w}}{\sigma_p(\mathbf{w})} &= \begin{pmatrix} (w_1\sigma_1^2 + w_2\sigma_{12})/\sigma_p(\mathbf{w}) \\ (w_2\sigma_2^2 + w_1\sigma_{12})/\sigma_p(\mathbf{w}) \end{pmatrix} \end{aligned}$$

Si ha perciò

$$\begin{aligned} \text{MCR}_1^\sigma &= (w_1\sigma_1^2 + w_2\sigma_{12})/\sigma_p(\mathbf{w}) \\ \text{MCR}_2^\sigma &= (w_2\sigma_2^2 + w_1\sigma_{12})/\sigma_p(\mathbf{w}) \\ \text{CR}_1^\sigma &= w_1 \cdot (w_1\sigma_1^2 + w_2\sigma_{12})/\sigma_p(\mathbf{w}) \end{aligned}$$

$$\text{PCR}_1^\sigma = \text{CR}_1^\sigma / \sigma_p(\mathbf{w}) = (w_1^2 \sigma_1^2 + w_1 w_2 \sigma_{12}) / \sigma_p^2(\mathbf{w})$$

$$\text{PCR}_2^\sigma = \text{CR}_2^\sigma / \sigma_p(\mathbf{w}) = (w_2^2 \sigma_2^2 + w_1 w_2 \sigma_{12}) / \sigma_p^2(\mathbf{w})$$

#### 5.4.2 Interpretazione ed uso di $\text{MCR}_i^\sigma$

A partire dalla definizione della *Contribuzione Marginale al Rischio* della componente *i-esima*

$$\text{MCR}_i^\sigma = \frac{\partial \sigma_p(\mathbf{w})}{\partial w_i}$$

può essere utile ricorrere ad un'approssimazione al fine di semplificare i calcoli:

$$\text{MCR}_i^\sigma = \frac{\partial \sigma_p(\mathbf{w})}{\partial w_i} \approx \frac{\Delta \sigma_p}{\Delta w_i},$$

segue perciò che  $\Delta \sigma_p \approx \text{MCR}_i^\sigma \cdot \Delta w_i$ .

In un portafoglio con  $n$  asset si ha

$$w_1 + w_2 + \dots + w_n = 1$$

e quindi l'incremento o il decremento di  $w_i$  comporta la necessità di decrementare o incrementare l'allocazione su uno o più altri asset. Perciò la formula

$$\Delta \sigma_p \approx \text{MCR}_i^\sigma \cdot \Delta w_i$$

ignora questo effetto di ri-allocazione.

Se l'aumento dell'allocazione relativo all'asset  $i$  è compensato dalla riduzione dell'allocazione sull'asset  $j$ , allora

$$\Delta w_j = -\Delta w_i$$

ed il cambiamento nella volatilità di portafoglio è approssimativamente

$$\begin{aligned} \Delta \sigma_p &\approx \text{MCR}_i^\sigma \cdot \Delta w_i + \text{MCR}_j^\sigma \cdot \Delta w_j \\ &= \text{MCR}_i^\sigma \cdot \Delta w_i - \text{MCR}_j^\sigma \cdot \Delta w_i \\ &= (\text{MCR}_i^\sigma - \text{MCR}_j^\sigma) \cdot \Delta w_i \end{aligned}$$

### Esempio

Decomposizione rischio di portafoglio a due titoli rispetto alla misura di rischio volatilità. I due asset sono caratterizzati dai seguenti parametri:

$\mu_1$	$\mu_2$	$\sigma_1^2$	$\sigma_2^2$	$\sigma_1$	$\sigma_2$	$\sigma_{12}$	$\rho_{12}$
0.175	0.055	0.067	0.013	0.258	0.115	-0.004875	-0.164

Immaginiamo due diverse possibili configurazioni dei pesi (supponendo siano ammesse vendite allo scoperto):

- $w_1 = w_2 = 0.5$
- $w_1 = 1.5$  e  $w_2 = -0.5$

La decomposizione per le due configurazioni risulta

	$\sigma_i$	$w_i$	$MCR_i^\sigma$	$CR_i^\sigma$	$PCR_i^\sigma$
$\sigma_p = 0.1323$					
Asset 1	0.258	0.5	0.23310	0.11655	0.8807
Asset 2	0.115	0.5	0.03158	0.01579	0.1193
$\sigma_p = 0.4005$					
Asset 1	0.258	1.5	0.25540	0.38310	0.95663
Asset 2	0.115	-0.5	-0.03474	0.01737	0.04337

Variando il peso  $w_1$  del primo portafoglio da 0.5 a 0.6,  $w_2$  passa da 0.5 a 0.4. Perciò

$$\begin{aligned}\Delta\sigma_p &\approx (MCR_1^\sigma - MCR_2^\sigma) \cdot \Delta w_i \\ &= (0.23310 - 0.03158)(0.1) \\ &= 0.02015\end{aligned}$$

A riguardo del secondo portafoglio, aumentando  $w_1$  da 1.5 a 1.6,  $w_2$  decresce da -0.5 a -0.6. Perciò

$$\begin{aligned}\Delta\sigma_p &\approx (MCR_1^\sigma - MCR_2^\sigma) \cdot \Delta w_i \\ &= (0.25540 - 0.03474)(0.1) \\ &= 0.02901\end{aligned}$$

### Nota

La riallocazione non deve per forza essere effettuata su un singolo asset: ipotizziamo una suddivisione su tutti gli altri asset  $j \neq i$  in modo che

$$\Delta w_j = -\alpha_j \Delta w_i \quad \text{t.c.} \quad \sum_{j \neq i} \alpha_j = 1$$

Allora

$$\sum_{j \neq i} \Delta w_j = - \sum_{j \neq i} \alpha_j \Delta w_i = -\Delta w_i \sum_{j \neq i} \alpha_j = -\Delta w_i$$

e

$$\begin{aligned} \Delta \sigma_p &\approx \text{MCR}_i^\sigma \cdot \Delta w_i + \sum_{j \neq i} \text{MCR}_j^\sigma \cdot \Delta w_j \\ &= \left[ \text{MCR}_i^\sigma \cdot \left( - \sum_{j \neq i} \alpha_j \right) \cdot \text{MCR}_j^\sigma \right] \Delta w_i \end{aligned}$$

Usando la notazione matriciale il risultato può essere scritto come

$$\begin{aligned} \Delta \sigma_p &\approx (\text{MCR}^{\sigma'} \alpha) \Delta w_i \\ \text{MCR}^\alpha &= (\text{MCR}_1^\sigma, \dots, \text{MCR}_n^\sigma)' \\ \alpha &= (-\alpha_1, \dots, -\alpha_{i-1}, 1, -\alpha_{i+1}, \dots, -\alpha_n)' \end{aligned}$$

#### 5.4.3 Interpretazione di MCR con VaR o ES

Analogamente a come si è mostrato per la volatilità, l'MCR per le misure di rischio VaR e ES consente di descrivere la variazione della misura di rischio di portafoglio al variare dell'allocazione sugli asset. In particolare si ha

- $\Delta w_j = -\Delta w_i$

$$\begin{aligned} \Delta \text{VaR}_\alpha(\mathbf{w}) &= \left( \text{MCR}_i^{\text{VaR}_\alpha} - \text{MCR}_j^{\text{VaR}_\alpha} \right) \cdot \Delta w_i \\ \Delta \text{ES}_\alpha(\mathbf{w}) &= \left( \text{MCR}_i^{\text{ES}_\alpha} - \text{MCR}_j^{\text{ES}_\alpha} \right) \cdot \Delta w_i \end{aligned}$$

- $\Delta w_j = -\alpha_j \Delta w_i$  t.c.  $\sum_{j \neq i} \alpha_j = 1$

$$\begin{aligned} \Delta \text{VaR}_\alpha(\mathbf{w}) &\approx \left( \text{MCR}_i^{\text{VaR}_\alpha'} \alpha \right) \cdot \Delta w_i \\ \Delta \text{ES}_\alpha(\mathbf{w}) &\approx \left( \text{MCR}_i^{\text{ES}_\alpha'} \alpha \right) \cdot \Delta w_i \end{aligned}$$

#### 5.4.4 $\beta$ come misura del contributo alla $\sigma$ di portafoglio

Per un portafoglio di  $n$  asset con return

$$R_p(\mathbf{w}) = w_1 R_1 + \dots + w_n R_n = \mathbf{w}' \mathbf{R}$$

usando la (37) e la (38) si ricava

$$\text{MCR}_i^\sigma = \frac{\partial \sigma_p(\mathbf{w})}{\partial w_i} = \text{i-esima riga di } \frac{\Sigma \mathbf{w}}{\sigma_p(\mathbf{w})}$$

**Definizione** La beta di un asset  $i$  rispetto ad un certo portafoglio è definita come

$$\beta_i = \frac{\text{cov}(R_i, R_p(\mathbf{w}))}{\text{var}(R_p(\mathbf{w}))} = \frac{\text{cov}(R_i, R_p(\mathbf{w}))}{\sigma_p^2(\mathbf{w})}$$

Ne consegue che  $\beta_i$  misura la contribuzione dello specifico asset alla misura  $\sigma_p(\mathbf{w})$ :

$$\begin{aligned}\text{MCR}_i^\sigma &= \frac{\partial \sigma_p(\mathbf{w})}{\partial w_i} = \beta_i \sigma_p(\mathbf{w}) \\ \text{CR}_i^\sigma &= w_i \beta_i \sigma_p(\mathbf{w}) \\ \text{PCR}_i^\sigma &= w_i \beta_i\end{aligned}$$

#### Note

Per costruzione, la beta di un portafoglio è 1:

$$\beta_p = \frac{\text{cov}(R_p(\mathbf{w}), R_p(\mathbf{w}))}{\text{var}(R_p(\mathbf{w}))} = \frac{\text{var}(R_p(\mathbf{w}))}{\text{var}(R_p(\mathbf{w}))} = 1$$

Quando  $\beta_i = 1$ :

$$\begin{aligned}\text{MCR}_i^\sigma &= \sigma_p(\mathbf{w}) \\ \text{CR}_i^\sigma &= w_i \sigma_p(\mathbf{w}) \\ \text{PCR}_i^\sigma &= w_i\end{aligned}$$

Quando  $\beta_i > 1$ :

$$\begin{aligned}\text{MCR}_i^\sigma &> \sigma_p(\mathbf{w}) \\ \text{CR}_i^\sigma &> w_i \sigma_p(\mathbf{w}) \\ \text{PCR}_i^\sigma &> w_i\end{aligned}$$

Quando  $\beta_i < 1$ :

$$\begin{aligned}\text{MCR}_i^\sigma &< \sigma_p(\mathbf{w}) \\ \text{CR}_i^\sigma &< w_i \sigma_p(\mathbf{w}) \\ \text{PCR}_i^\sigma &< w_i\end{aligned}$$

### Esempio

Riprendiamo l'esempio del *Paragrafo 5.4.2*, aggiungendo il calcolo dei  $\beta_i$

	$\sigma_i$	$w_i$	$\beta_i$	$\text{MCR}_i^\sigma$	$\text{CR}_i^\sigma$	$\text{PCR}_i^\sigma$
$\sigma_p = 0.1323$						
Asset 1	0.258	0.5	1.76	0.23310	0.11655	0.8807
Asset 2	0.115	0.5	0.24	0.03158	0.01579	0.1193
$\sigma_p = 0.4005$						
Asset 1	0.258	1.5	0.64	0.25540	0.38310	0.95663
Asset 2	0.115	-0.5	-0.09	-0.03474	0.01737	0.04337

Si nota che per il portafoglio ad uguali pesi, l'asset 1 accresce il rischio ( $\beta > 1$ ), l'asset 2 lo riduce ( $\beta < 1$ ); per il portafoglio con una posizione di vendita allo scoperto si ha addirittura  $\beta_2 < 0$ .

Richiamiamo

$$\text{MCR}_i^\sigma = \frac{\partial \sigma_p(\mathbf{w})}{\partial w_i} = \text{i-esima riga di } \frac{\Sigma \mathbf{w}}{\sigma_p(\mathbf{w})} = \frac{\text{cov}(R_i, R_p(\mathbf{w}))}{\sigma_p(\mathbf{w})}.$$

Usando

$$\begin{aligned} \rho_{i,p} &= \text{corr}(R_i, R_p(\mathbf{w})) = \frac{\text{cov}(R_i, R_p(\mathbf{w}))}{\sigma_i \sigma_p(\mathbf{w})} \\ \Rightarrow \text{cov}(R_i, R_p(\mathbf{w})) &= \rho_{i,p} \sigma_i \sigma_p(\mathbf{w}) \end{aligned}$$

si ottiene

$$\text{MCR}_i^\sigma = \frac{\rho_{i,p} \sigma_i \sigma_p(\mathbf{w})}{\sigma_p(\mathbf{w})} = \rho_{i,p} \sigma_i.$$

Perciò

$$\text{CR}_i^\sigma = w_i \cdot \text{MCR}_i^\sigma = w_i \cdot \sigma_i \cdot \rho_{i,p}$$

$$\text{CR}_i^\sigma = \text{allocazione} \cdot \text{rischio singolo asset} \cdot \text{correlazione con il portafoglio}$$

**Note**

- $w_i \cdot \sigma_i$  = contribuzione al rischio del singolo asset (ignora effetti di correlazione con altri titoli)
- $\text{CR}_i^\sigma = w_i \cdot \sigma_i$  solo quando  $\rho_{i,p} = 1$

- se  $\rho_{i,p} \neq 1$  allora  $CR_i^\sigma < w_i \cdot \sigma_i$

### Esempio

Riprendiamo nuovamente l'esempio del *Paragrafo 5.4.2*, aggiungendo il calcolo dei  $\rho_i$

	$\sigma_i$	$w_i$	$\rho_i$	$MCR_i^\sigma$	$CR_i^\sigma$	$PCR_i^\sigma$
$\sigma_p = 0.1323$						
Asset 1	0.258	0.5	0.90	0.23310	0.11655	0.8807
Asset 2	0.115	0.5	0.27	0.03158	0.01579	0.1193
$\sigma_p = 0.4005$						
Asset 1	0.258	1.5	0.99	0.25540	0.38310	0.95663
Asset 2	0.115	-0.5	-0.30	-0.03474	0.01737	0.04337

Si nota che per il portafoglio ad uguali pesi, entrambi gli asset sono positivamente correlati con il portafoglio; per il portafoglio con una posizione di vendita allo scoperto si ha che l'asset 2 è negativamente correlato con il portafoglio.

I risultati ottenuti su  $\beta$  e  $\rho$  per la volatilità possono essere estesi anche alla misura ES<sup>14</sup>.

<sup>14</sup>Goldberg, Hayes, Menchero e Mitra - *Extreme Risk Management*, MSCI Barra Research Paper (2009)

## 6 Stimare le misure di rischio

Nella disciplina del Risk Management risulta di primaria importanza quantificare le misure di rischio, attraverso stime ricavabili dai dati storici dei return.

Chiamiamo  $R$  il return di un asset o di un portafoglio con funzione di ripartizione  $F_R$  e funzione di densità  $f_R$ . Siano  $\text{RM} = \sigma$ ,  $\text{VaR}_\alpha$  e  $\text{ES}_\alpha$  le misure di rischio di nostro interesse. Dato un certo livello di confidenza  $\alpha$  e un valore iniziale di investimento  $P_0$ , le nostre misure di rischio varranno:

$$\sigma_L = P_0 \left( \mathbb{E}[(R - \mu_R)^2] \right)^{\frac{1}{2}} = P_0 \sigma_R \quad (39)$$

$$\text{VaR}_\alpha = -P_0 q_{1-\alpha}^R = -P_0 F_R^{-1}(1 - \alpha) \quad (40)$$

$$\text{ES}_\alpha = -P_0 \mathbb{E}[R | R \leq q_{1-\alpha}^R] = \frac{P_0}{1 - \alpha} \int_{-\infty}^{q_{1-\alpha}^R} x f_R(x) dx \quad (41)$$

Dunque siamo interessati a stimare  $\sigma_R$ ,  $q_{1-\alpha}^R$  e  $\mathbb{E}[R | R \leq q_{1-\alpha}^R]$  attraverso il campione osservato di return  $\{R_1 = r_1, \dots, R_T = r_t\}$

Gli approcci possibili per effettuare la stima sono molteplici:

- **non parametrico:** vengono utilizzate distribuzioni empiriche per stimare le misure di rischio;
- **parametrico:** viene dapprima specificata la distribuzione (parametrica) dei return, si stimano poi i parametri della distribuzione e infine si darà una stima che non sarà nient'altro che una funzione dei parametri stimati;
- **semi-parametrico:** un approccio ibrido modellizzato per la prima volta da Cornish-Fisher<sup>15</sup>

Le stime delle misure di rischio sono soggette al cosiddetto *errore di stima*. Il modo più semplice per il calcolo degli errori standard e gli intervalli di confidenza è attraverso la tecnica di *bootstrapping error estimation*<sup>16</sup>. Nelle applicazioni pratiche spesso non si tiene conto di questi errori che si riscontrano in sede di stima.

---

<sup>15</sup>Cornish E. A., Fisher Ronald A. - *Moments and Cumulants in the Specification of Distributions*, Revue de l'Institut International de Statistique (1938)

<sup>16</sup>Efron B., Tibshirani R. - *An Introduction to the Bootstrap*, Boca Raton, FL: Chapman & Hall (1993)

Prima di vedere gli approcci di stima nello specifico è necessario fare alcune assunzioni:

- i return  $R_t \sim iid$  per  $t = 1, \dots, T$
- $\mathbb{E}[R_t] = \mu_R < \infty$
- $var(R_t) = \sigma_R^2 < \infty$
- $cov(R_t, R_s) = 0$  per ogni  $t \neq s$

## 6.1 Stima non parametrica delle misure di rischio

L'idea che sta alla base di questo tipo di approccio è supporre la funzione di ripartizione dei return  $F_R$  incognita e stimarla usando la distribuzione empirica

$$\begin{aligned}\hat{F}_R(r) &= \frac{\# \text{ di return minori o uguali a } r}{T} \\ &= \frac{\sum_{t=1}^T \mathbb{I}_{R_t \leq r}}{T}\end{aligned}$$

dove  $\mathbb{I}_{R_t < r} = 1$  se  $R_t \leq r$ ; 0 altrimenti.

Man mano che il campione dei return diventa più grande, la distribuzione campionaria o empirica tende a quella reale:  $\hat{F}_R(r) \rightarrow F_R(r)$  quando  $T \rightarrow \infty$  per la legge forte dei grandi numeri.

### 6.1.1 Volatilità - stima non parametrica

Per la volatilità abbiamo bisogno di uno stimatore della deviazione standard  $\sigma_R$ :

$$\hat{\sigma}_R = \left( \frac{1}{T-1} \sum_{t=1}^T (R_t - \hat{\mu})^2 \right)^{\frac{1}{2}}, \quad \hat{\mu}_R = \frac{1}{T} \sum_{t=1}^T R_t \quad (42)$$

Lo stimatore gode di alcune proprietà:

$$\begin{aligned}\mathbb{E}[\hat{\sigma}_R] &\neq \sigma_R \\ \hat{\sigma}_R &\rightarrow \sigma_R \text{ per } T \rightarrow \infty\end{aligned}$$

Si può dimostrare che  $\hat{\sigma}_R$  converge asintoticamente ad una normale<sup>17</sup>:

$$\hat{\sigma}_R \sim^A N\left(\sigma_R, \frac{\sigma_R^2}{2T}\right).$$

Di conseguenza lo Standard Error risulta:

$$SE(\hat{\sigma}_R) = \frac{\sigma_R}{\sqrt{2T}}$$

$$\hat{SE}(\hat{\sigma}_R) = \frac{\hat{\sigma}_R}{\sqrt{2T}}$$

### 6.1.2 Value at risk - stima non parametrica

Per stimare il Value at Risk abbiamo bisogno innanzitutto di una stimatore del quantile:

$$\hat{q}_{1-\alpha}^R = \text{quantile empirico di livello } 1 - \alpha$$

$$= r_t \text{ tale che } 100 \cdot (1 - \alpha)\% \text{ dei dati è inferiore a } r_t$$

$$\widehat{\text{VaR}}_{\alpha}^{HS} = -P_0 \cdot \hat{q}_{1-\alpha}^R, \text{ dove } HS = \text{'historical simulation'} \quad (43)$$

Come fatto nell'esempio precedente, andiamo a ispezionare le proprietà dello stimatore:

$$\hat{q}_{1-\alpha}^R \rightarrow q_{1-\alpha}^R \text{ per } T \rightarrow \infty$$

$$\hat{q}_{1-\alpha}^R \sim^A N\left(q_{1-\alpha}^R, \frac{\alpha(1-\alpha)}{T \cdot f_R(q_{1-\alpha}^R)^2}\right)$$

$$SE(\hat{q}_{1-\alpha}^R) = \frac{\sqrt{\alpha(1-\alpha)}}{\sqrt{T \cdot f_R(q_{1-\alpha}^R)^2}}$$

Dunque per stimare l'errore è necessario prima stimare  $f_R$ .

### 6.1.3 Expected Shortfall - stima non parametrica

Per stimare l'ES partiamo dal valore atteso condizionato definito in (41)

$$\hat{\mathbb{E}}[R | R \leq q_{1-\alpha}^R] = \frac{1}{B_{1-\alpha}} \sum_{t=1}^T R_t \cdot \mathbb{I}_{R_t \leq r}$$

<sup>17</sup>David Ruppert e David S. Matteson - *Statistics and Data Analysis for Financial Engineering*, Springer (2015)

dove  $\Pi_{R_t < r} = 1$  se  $R_t \leq r$ ; 0 altrimenti.

$$B_{1-\alpha} = \sum_{t=1}^T \Pi_{R_t < r} = \# \text{ di return } \leq \hat{q}_{1-\alpha}^R$$

$$\hat{E}S_{\alpha}^{HS} = -P_0 \cdot \hat{\mathbb{E}}[R|R \leq q_{1-\alpha}^R]$$

Anche in questo caso esplicitiamo le proprietà:

$$\begin{aligned} \hat{\mathbb{E}}[R|R \leq q_{1-\alpha}^R] &\rightarrow \mathbb{E}[R|R \leq q_{1-\alpha}^R] \text{ per } T \rightarrow \infty \\ \hat{\mathbb{E}}[R|R \leq q_{1-\alpha}^R] &\sim^A N\left(\mathbb{E}[R|R \leq q_{1-\alpha}^R], \frac{\sigma_R^2}{B_{1-\alpha}}\right) \end{aligned}$$

## Esempio

Al fine di comprendere gli argomenti trattati nel paragrafo, segue un esempio in cui vengono mostrate le modalità in cui ricavare le misure di rischio non parametriche per titoli realmente negoziati sui mercati finanziari.

Consideriamo, ad esempio, la timeseries dei prezzi relativi al titolo Microsoft negoziato sul NYSE. In particolare, ogni punto corrisponde al prezzo di chiusura<sup>18</sup> giornaliero.



Figura 5: Grafico relativo ai prezzi di chiusura giornalieri del titolo Microsoft negoziato sul NYSE tra il 30 nov 2005 e il 1 nov 2017

Poichè i modelli si basano sui log-returns, calcoliamoli e rappresentiamoli su un grafico.

<sup>18</sup>Il prezzo di chiusura di uno strumento finanziario quotato su un indice di borsa è il prezzo a cui è stata eseguita l'ultima transazione durante l'orario di apertura del mercato. Questo valore non dipende dai prezzi degli scambi eseguiti durante la seduta after-hour.

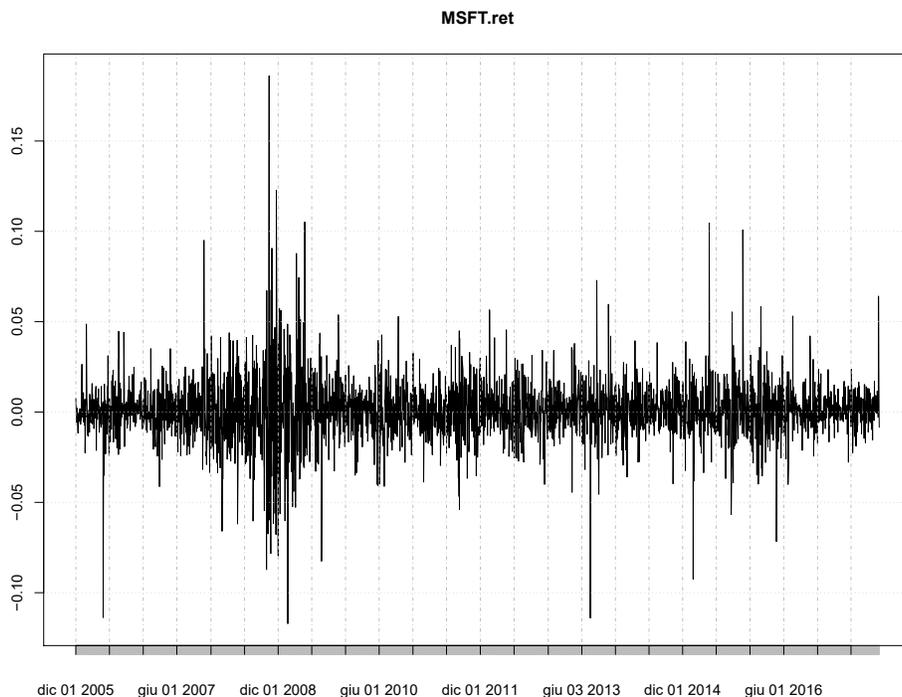


Figura 6: Grafico relativo ai log-returns giornalieri del titolo Microsoft negoziato sul NYSE tra il 30 nov 2005 e il 1 nov 2017

A partire dal campione (il vettore dei log-returns) procediamo al calcolo delle misure di rischio. Ipotizziamo una distribuzione normale dei log-returns.

Usando la (42) stimiamo la volatilità:  $\sigma = 0,01698$

Usando la (43) stimiamo il Value at Risk:  $\text{VaR}_{5\%} = -0,02379$

Usando la (44) stimiamo l'Expected Shortfall:  $\text{ES}_{5\%} = -0,03865$

Rappresentiamo sulla *Figura 7* queste due ultime misure di rischio.

La retta blu ( $\text{VaR}_{5\%}$ ) rappresenta il quantile della distribuzione empirica tale da racchiudere il 5% dei log-returns al di sotto di essa. Investendo perciò oggi in una azione Microsoft, la peggior perdita che potremmo riscontrare con probabilità del 95% è:  $e^{\text{VaR}_{5\%}}\$ = e^{-0,02379}\$ = 0,976491\$$ .

Nel 5% dei casi la perdita risulta superiore a 0,976491\$; in questa circostanza il valore atteso dei log-retrurn è rappresentato dall'Expected Shortfall (retta verde nel grafico). Si tratta infatti del log-return atteso condizionato al fatto di avere valori inferiori al VaR.

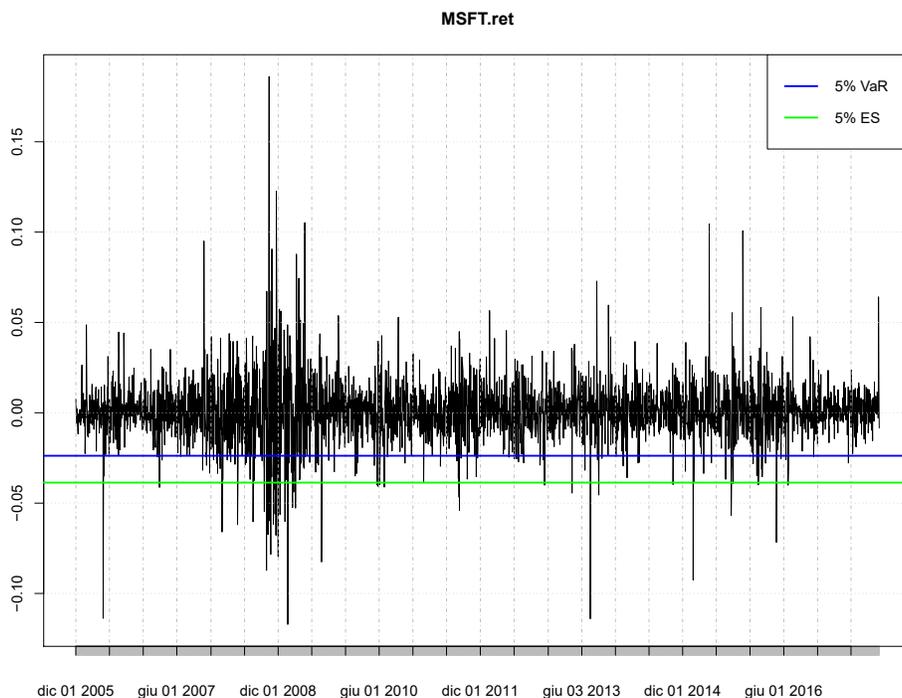


Figura 7: Grafico relativo ai log-returns giornalieri del titolo Microsoft negoziato sul NYSE tra il 30 nov 2005 e il 1 nov 2017 con indicazione del  $\text{VaR}_{5\%}$  e dell' $\text{ES}_{5\%}$

Per quanto riguarda il  $\text{VaR}_{\alpha}$ , poichè rappresenta il valore del quantile  $\alpha$ -esimo, è d'interesse capire come varia lo Standard Error al variare di  $\alpha$ . Se l'ipotesi di normalità sui log-return è corretta, ci aspettiamo di ottenere un grafico simmetrico rispetto alla retta  $x = 0,5$ , e riscontrare un minimo alla medesima ascissa.

Nella *Figura 8* vengono riportati gli Standard Error rilevati nel caso preso in considerazione al variare di  $\alpha$ . Si noti come il valore assoluto della derivata prima assuma i valori più alti negli intervalli (0%; 5%) e (95%; 100%), giustificando perciò la scelta del livello del 5% per il calcolo del VaR.

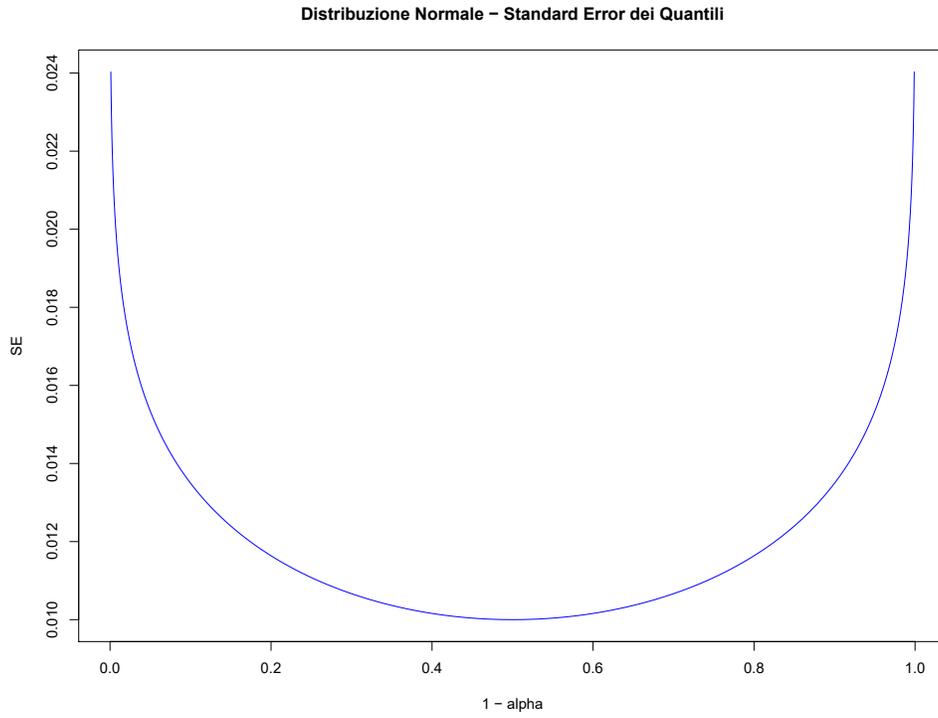


Figura 8: Standard Error al variare degli  $(1 - \alpha)$  quantili nel caso di return normali

## 6.2 Stima parametrica delle misure di rischio

Chiamiamo  $F_R = F_R(r; \theta)$  e  $f_r = f_r(r; \theta)$  funzione di ripartizione e funzione di densità parametriche che dipendono da  $p$  parametri ignoti  $\theta = (\theta_1, \dots, \theta_p)'$ . L'obiettivo è stimare  $\theta$  a partire dai dati osservati (tipicamente usando la funzione di verosimiglianza) ottenendo così delle stime delle misure di rischio.

### 6.2.1 Stima di massima verosimiglianza di $\theta$

Sia  $R_1, \dots, R_n$  un campione casuale di  $n$  variabili aleatorie iid con densità  $f_R(r_i; \theta)$ , dove  $\theta$  è un vettore di  $p$  parametri. La distribuzione congiunta del campione  $r = (r_1, \dots, r_n)'$  è, per l'indipendenza, equivalente al prodotto delle densità marginali

$$f_R(r; \theta) = f_R(r_1; \theta), \dots, f_R(r_n; \theta) = \prod_{i=1}^n f_R(r_i; \theta)$$

La *funzione di verosimiglianza* è definita come la funzione di densità congiunta al variare del parametro  $\theta$  quando il campione si realizza:

$$L_R(\theta|r) = f_R(r; \theta) = \prod_{i=1}^n f_R(r_i; \theta)$$

Lo stimatore di massima verosimiglianza (che indicheremo con  $\hat{\theta}_{mle}$ ) è il valore di  $\theta$  che massimizza  $L_R(\theta|r)$ :

$$\hat{\theta}_{mle} = \arg \max_{\theta} L_R(\theta|r)$$

Usualmente, è molto più semplice massimizzare la funzione di verosimiglianza logaritmica (log-likelihood). Infatti, poichè il logaritmo è una funzione monotona, vale:

$$\hat{\theta}_{mle} = \arg \max_{\theta} \ln(L_R(\theta|r))$$

In particolare è più agevole lavorare perché la log-likelihood è una sommatoria:

$$\ln(L_R(\theta|r)) = \ln\left(\prod_{i=1}^n f_R(r_i; \theta)\right) = \sum_{i=1}^n \ln(f_R(r_i; \theta))$$

Tipicamente possiamo trovare una stima di massima verosimiglianza derivando  $L_R(\theta|r)$  e risolvendo le condizioni del primo ordine

$$\frac{\partial \ln(L_R(\theta|r))}{\partial \theta} = 0 \tag{44}$$

Fino a quando  $\theta$  è un vettore  $p \times 1$  le condizioni di primo ordine definiscono  $p$  -potenzialmente non lineari- equazioni in  $p$  incogniti valori:

$$\frac{\partial \ln(L_R(\hat{\theta}_{mle}|r))}{\partial \theta} = \begin{pmatrix} \frac{\partial \ln(L_R(\hat{\theta}_{mle}|r))}{\partial \theta_1} \\ \vdots \\ \frac{\partial \ln(L_R(\hat{\theta}_{mle}|r))}{\partial \theta_p} \end{pmatrix} \quad (45)$$

Per semplicità chiameremo *vettore score* il vettore di derivate della funzione di verosimiglianza logaritmica:

$$S_R(\theta|r) = \frac{\partial \ln L_R(\theta|r)}{\partial \theta}$$

Per definizione quindi la stima di massima verosimiglianza (MLE) soddisfa

$$S_R(\hat{\theta}_{mle}|r) = 0$$

La matrice  $p \cdot p$  delle derivate al secondo ordine della verosimiglianza logaritmica è chiamata *Hessiano*

$$H_R(\theta|r) = \frac{\partial^2 \ln L_R(\theta|r)}{\partial \theta \partial \theta'} = \begin{pmatrix} \frac{\partial^2 \ln(L_R(\theta|r))}{\partial \theta_1^2} & \cdots & \frac{\partial^2 \ln(L_R(\theta|r))}{\partial \theta_1 \theta_p} \\ \vdots & \ddots & \vdots \\ \frac{\partial \ln(L_R(\hat{\theta}_{mle}|r))}{\theta_p \theta_1} & \cdots & \frac{\partial \ln(L_R(\hat{\theta}_{mle}|r))}{\theta_p^2} \end{pmatrix} \quad (46)$$

La *matrice di informazione* è definita come l'opposto del valore atteso dell'Hessiana

$$I_R(\theta|r) = -\mathbb{E}[H_R(\theta|r)]$$

### 6.2.2 Calcolo della stima di max verosimiglianza

Naturalmente solo nel caso di alcune distribuzioni semplici dei return ( es. distribuzione normale) è sempre possibile ottenere una soluzione in forma chiusa per  $\hat{\theta}_{mle}$ , ma in generale tale risultato non è garantito. Nei casi più complessi viene, tipicamente utilizzato un metodo numerico iterativo.<sup>19</sup>

$$\begin{aligned} \hat{\sigma}_{mle,n+1} &= \hat{\sigma}_{mle,n} - H(\hat{\sigma}_{mle,n+1}|r)^{-1} S(\hat{\sigma}_{mle,n}|r) \\ \hat{\sigma}_{mle,0} &= \sigma_0 = \text{valore iniziale} \end{aligned}$$

<sup>19</sup>Solitamente viene utilizzato il metodo iterativo di Newton-Raphson - anche conosciuto col nome di *metodo delle tangenti* - che si basa sull'espansione di Taylor a secondo ordine della funzione di massima verosimiglianza.

e l'iterazione si ferma quando  $S(\hat{\sigma}_{mle,n}|r) \approx 0$

Enunciamo di seguito alcune proprietà asintotiche dello stimatore di massima verosimiglianza sotto condizioni di regolarità:

$$\begin{aligned}\hat{\theta}_{mle} &\xrightarrow{p} \theta \\ \hat{\theta}_{mle} &\sim N\left(\theta, \frac{1}{T}H(\hat{\theta}_{mle}|r)^{-1}\right) \\ SE(\hat{\theta}_{i,mle}) &= \text{elemento diagonale } i\text{-esimo di } \frac{1}{H}(\hat{\theta}_{mle}|r)^{-1}\end{aligned}$$

Scriviamo ora la stima parametrica del Value at risk e dell'Expected Shortfall per esteso:

$$\begin{aligned}\hat{\theta} &= \hat{\theta}_{mle} \\ F_R(r, \hat{\theta}) &= \text{stima parametrica di } F_R \\ F_R^{-1}(r, \hat{\theta}) &= \text{stima parametrica di } F_R^{-1} \\ f_r(r, \hat{\theta}) &= \text{stima parametrica di } f_R\end{aligned}$$

Allora

$$\begin{aligned}\widehat{\text{VaR}}_{\alpha}^{par} &= -V_0 \cdot F_R^{-1}(1 - \alpha; \hat{\theta}) \\ \widehat{ES}^{par} &= \frac{-V_0}{1 - \alpha} \int_{-\infty}^{q_{1-\alpha}^R(\hat{\theta})} x f_R(x; \hat{\theta}) dx\end{aligned}$$

La scrittura *par* indica una distribuzione parametrica.

### 6.2.3 Caso di distribuzione normale

Calcoliamo ora la stima parametrica nel caso semplice di distribuzione normale dei return. Innanzitutto scriviamo le stime dei parametri della distribuzione normale. Sia  $R_t \sim iid N(\mu_R, \sigma_R^2)$ . Il vettore dei parametri è dunque  $\theta = (\mu_R, \sigma_R^2)'$ . Chiamiamo  $\mathbf{r} = (r_1, \dots, r_T)$  il campione osservato. Le stime dei parametri sono:

$$\begin{aligned}\hat{\mu}_{R,mle} &= \frac{1}{T} \sum_{t=1}^T r_t \\ \hat{\sigma}_{R,mle}^2 &= \frac{1}{T} \sum_{t=1}^T (r_t - \hat{\mu}_{R,mle})^2\end{aligned}$$

Ora esplicitiamo il valore di  $F_R^{-1}(1 - \alpha; \hat{\theta}_{mle})$  necessario al calcolo del Var e dell'ES:

$$\begin{aligned}
F_R^{-1}(1 - \alpha; \hat{\theta}_{mle}) &= q_{1-\alpha}^R(\hat{\theta}_{mle}) = \hat{\mu}_{R,mle} + \hat{\sigma}_{R,mle} \cdot q_{1-\alpha}^Z \\
q_{1-\alpha}^Z &= \text{quantile di livello } 1 - \alpha \text{ di } Z \sim N(0, 1) \\
\mathbb{E}[R | R \leq q_{1-\alpha}^R(\hat{\theta}_{mle})] &= - \left( \hat{\mu}_{R,mle} + \hat{\sigma}_{R,mle} \cdot \frac{\phi(q_{1-\alpha}^Z)}{1 - \alpha} \right) \\
\hat{\text{VaR}}_\alpha^{norm} &= -V_0 \left( \hat{\mu}_{R,mle} + \hat{\sigma}_{R,mle} \cdot q_{1-\alpha}^Z \right) \\
\hat{\text{ES}}_\alpha^{norm} &= -V_0 \left( \hat{\mu}_{R,mle} + \hat{\sigma}_{R,mle} \cdot \frac{\phi(q_{1-\alpha}^Z)}{1 - \alpha} \right)
\end{aligned}$$

L'accuratezza delle stime si calcola utilizzando il teorema del limite centrale, che in questo caso da il seguente risultato:

$$\begin{pmatrix} \hat{\mu}_{R,mle} \\ \hat{\sigma}_{R,mle} \end{pmatrix} \sim N \left( \begin{pmatrix} \mu_R \\ \sigma_R \end{pmatrix}, \begin{pmatrix} \frac{\sigma_R^2}{T} & 0 \\ 0 & \frac{\sigma_R^2}{2T} \end{pmatrix} \right)$$

Dunque è semplice far vedere che

$$SE(q_{1-\alpha}^R(\hat{\theta}_{mle})) = \frac{\sigma_R}{\sqrt{T}} \left( 1 + \frac{1}{2}(q_{1-\alpha}^Z)^2 \right)^{\frac{1}{2}}$$

Per un determinato valore di  $\sigma_R$  e di  $T$ ,  $SE(q_{1-\alpha}^R(\hat{\theta}_{mle}))$  cresce non linearmente all'aumentare di  $\alpha$ .

### 6.3 Stima semi-parametrica delle misure di rischio

Come suggerisce il nome, con questo tipo di approccio ibrido alcune parti della funzione di ripartizione  $F_R$  sono trattate in modo parametrico, mentre altre no. Accenniamo di seguito ad un esempio che dà l'idea di come viene usato questo metodo per stimare il VaR.

Approssimiamo la funzione di ripartizione  $F_Z$  di  $Z = (R - \mu_R)/\sigma_R$  usando i termini dell'espansione di Edgeworth<sup>20</sup> per la funzione di ripartizione normale e invertiamo poi l'espansione per ottenere l'approssimazione di

<sup>20</sup>Peter Hall - *The Bootstrap and Edgeworth Expansion*, Springer (2013)

Cornish-Fisher della funzione quantile  $F_{Z,CF}^{-1}$  di  $Z$ :

$$F_{Z,CF}^{-1}(1 - \alpha) = z_{1-\alpha} + \frac{1}{6}(z_{1-\alpha}^2 - 1) \cdot skew + \frac{1}{24}(z_{1-\alpha}^3 - 3z_{1-\alpha}) \cdot ekurt - \frac{1}{36}(2z_{1-\alpha}^3 - 5z_{1-\alpha}) \cdot skew$$

$$z_{1-\alpha} = \Phi^{-1}(1 - \alpha) = \text{quantile di } N(0, 1)$$

Ricordiamo che in questo caso  $skew = \mathbb{E}[Z^3]$ , mentre  $ekurt = \mathbb{E}[Z^4] - 3$ .

L'approssimazione della funzione quantile di Cornish-Fisher risulta essere:

$$\hat{q}_{1-\alpha}^{CF} = \hat{\mu}_R + \hat{\sigma}_R \cdot \hat{F}_{Z,CF}^{-1}(1 - \alpha)$$

Segue che:

$$\text{VaR}_\alpha^{CF} = -V_0 \cdot q_{1-\alpha}^{CF}$$

#### 6.4 Stimare le misure di rischio di portafoglio

Sia  $\mathbf{R} = (R_1, R_2, \dots, R_n)'$  un vettore di *returns* su  $n$  asset e  $\mathbf{w} = (w_1, w_2, \dots, w_n)'$  denoti i pesi di portafoglio tali che  $\sum_{i=1}^n w_i = 1$ .

Per le modellazioni che seguiranno sono necessarie le seguenti ipotesi:

- $\mathbf{R}_t = (R_{1t}, \dots, R_{nt})'$  è iid al variare del tempo  $t$  con funzione di ripartizione congiunta  $F_{\mathbf{R}}$  e funzione di densità congiunta  $f_{\mathbf{R}}$
- $\mathbb{E}[\mathbf{R}_t] = \mu = (\mu_1, \dots, \mu_n)' \forall t$

$$\bullet \text{ var}(\mathbf{R}_t) = \mathbb{E}[(\mathbf{R}_t - \mu)(\mathbf{R}_t - \mu)'] = \Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1n} \\ \sigma_{12} & \sigma_2^2 & \dots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1n} & \sigma_{2n} & \dots & \sigma_n^2 \end{pmatrix}$$

A questo punto possiamo definire la variabile aleatoria *portfolio return*  $R_{pt}$ , definita in funzione dei *return*  $\mathbf{R}_t$  e dei pesi  $\mathbf{w}$ :

$$R_{pt} = \mathbf{w}'\mathbf{R}_t = \sum_{i=1}^N w_i R_{it}$$

$R_{pt} \sim iid F_{\mathbf{R}_p}$  che dipende dalla distribuzione congiunta  $F_{\mathbf{R}}$

Il suo valor medio e varianza risultano essere:

$$\mu_p = \mathbf{w}'\boldsymbol{\mu} \quad \sigma_p^2 = \mathbf{w}'\boldsymbol{\Sigma}\mathbf{w}$$

Per quanto riguarda le misure di rischio di portafoglio, calcoleremo - come in precedenza - la Volatilità, il Value at Risk e l'Expected Shortfall. Dato un intervallo di confidenza  $\alpha$  e un investimento iniziale  $V_0$ , le misure di rischio della perdita sono:

$$\begin{aligned} \sigma_L &= V_0 \left( \mathbb{E}[(R_p - \mu_p)^2] \right)^{1/2} = V_0 \sigma_p \\ \text{VaR}_\alpha &= -V_0 q_{1-\alpha}^{R_p} = -V_0 F_{R_p}^{-1}(1 - \alpha) \\ \text{ES}_\alpha &= -V_0 \mathbb{E}[R_p | R_p \leq q_{1-\alpha}^{R_p}] = \frac{-V_0}{1 - \alpha} \int_{-\infty}^{q_{1-\alpha}^{R_p}} x f_{R_p}(x) dx \end{aligned}$$

Poiché  $R_p = \mathbf{w}'\mathbf{R}$ , le misure di rischio indicate sono tutte funzioni di  $\mathbf{w}$ .

Sia  $\text{RM}(\mathbf{w})$  una delle misure di rischio (volatilità, value at risk, expected shortfall) in funzione dei pesi di portafoglio  $\mathbf{w}$ .

Ricordiamo i rapporti (31), (32) e (33):

$$\text{MCR}_i^{\text{RM}} = \frac{\partial \text{RM}(\mathbf{w})}{\partial w_i}; \quad \text{CR}_i^{\text{RM}} = w_i \frac{\partial \text{RM}(\mathbf{w})}{\partial w_i}; \quad \text{PCR}_i^{\text{RM}} = \frac{w_i \frac{\partial \text{RM}(\mathbf{w})}{\partial w_i}}{\text{RM}(\mathbf{w})}.$$

Poiché  $\text{CR}_i^{\text{RM}}$  e  $\text{PCR}_i^{\text{RM}}$  sono facilmente ricavabili da  $\text{MCR}_i^{\text{RM}}$ , verrà esPLICITATO solo quest'ultimo al variare delle misure di rischio:

$$\text{MCR}_i^\sigma = \frac{\partial \sigma(\mathbf{w})}{\partial w_i} = \left[ \frac{1}{\sigma(\mathbf{w})} \boldsymbol{\Sigma} \mathbf{w} \right]_i \quad (47)$$

$$\text{MCR}_i^{\text{VaR}} = \frac{\partial \text{VaR}_\alpha(\mathbf{w})}{\partial w_i} = \mathbb{E}[r_i | R_p = \text{VaR}_\alpha(\mathbf{w})] \quad (48)$$

$$\text{MCR}_i^{\text{ES}} = \frac{\partial \text{ES}_\alpha(\mathbf{w})}{\partial w_i} = \mathbb{E}[r_i | R_p \leq \text{VaR}_\alpha(\mathbf{w})] \quad (49)$$

Come fatto precedentemente nel caso di una singola componente in portafoglio, siamo interessati alle stime delle misure di rischio

$$\sigma_p; \quad q_{1-\alpha}^{R_p}; \quad \mathbb{E}[R_p | R_p \leq q_{1-\alpha}^{R_p}] \quad \text{e} \quad \text{MCR}_i^{\text{RM}}$$

per un vettore di osservazioni campionarie  $\{\mathbf{R}_1 = \mathbf{r}_1, \dots, \mathbf{R}_T = \mathbf{r}_T\}$

## 6.5 Stima non parametrica delle misure di rischio di portafoglio

Nella sezione 6.1 abbiamo visto come stimare in modo non parametrico le misure di rischio di un singolo asset. Ora applichiamo lo stesso approccio nel caso di un portafoglio contenente più strumenti finanziari.

### 6.5.1 Volatilità di portafoglio - stima non parametrica

Presentiamo due metodi differenti per stimare la *volatilità di portafoglio*.

1. Per prima cosa, prendiamo in considerazione una time series di *portfolio return*  $R_{pt} = \mathbf{w}'\mathbf{R}_t$  e calcoliamo la deviazione standard:

$$\hat{\sigma}_p = \left( \frac{1}{T-1} \sum_{t=1}^T (R_{pt} - \hat{\mu})^2 \right)^{\frac{1}{2}}, \quad \hat{\mu}_p = \frac{1}{T} \sum_{t=1}^T R_{pt} \quad (50)$$

2. Utilizziamo la formula  $\sigma_p^2 = \mathbf{w}'\Sigma\mathbf{w}$  vista in precedenza. Calcoliamo dunque la matrice di covarianza campionaria:

$$\hat{\Sigma} = \frac{1}{T-1} \sum_{t=1}^T (\mathbf{R}_t - \hat{\mu})(\mathbf{R}_t - \hat{\mu})' \quad (51)$$

$$\hat{\mu} = \frac{1}{T} \sum_{t=1}^T \mathbf{R}_t \quad (52)$$

Infine sono in grado di trovare  $\hat{\sigma}_p(\mathbf{w}) = \left( \mathbf{w}'\hat{\Sigma}\mathbf{w} \right)^{\frac{1}{2}}$

Per stimare il contributo marginale al rischio di portafoglio dell'asset  $i$ -esimo è sufficiente riprendere la definizione (47) da cui

$$\text{MCR}^\sigma = \frac{\partial \sigma_p(\mathbf{w})}{\partial w} = \frac{1}{\sigma_p(\mathbf{w})} \Sigma \mathbf{w} = (\mathbf{w}'\Sigma\mathbf{w})^{\frac{1}{2}} \Sigma \mathbf{w}$$

e sostituire la matrice di covarianza campionaria:

$$\text{M}\hat{\text{C}}\text{R}^\sigma = \frac{\partial \hat{\sigma}_p(\mathbf{w})}{\partial w} = (\mathbf{w}'\hat{\Sigma}\mathbf{w})^{\frac{1}{2}} \hat{\Sigma} \mathbf{w} \quad (53)$$

### 6.5.2 VaR ed ES di portafoglio - stima non parametrica

$$\widehat{\text{VaR}}_\alpha^{\text{HS}} = -V_0 \cdot \hat{q}_{1-\alpha}^{R_p}, \quad \text{HS} = \text{simulazione storica} \quad (54)$$

$$\hat{q}_{1-\alpha}^R = \text{quantile empirico di } R_{p,t} \quad (55)$$

$$\widehat{\text{ES}}_\alpha^{\text{HS}} = -V_0 \cdot \hat{\mathbb{E}} \left[ R_p | R_p \leq \hat{q}_{1-\alpha}^R \right] \quad (56)$$

$$\hat{\mathbb{E}}[R_p | R_p \leq \hat{q}_{1-\alpha}^R] = \frac{1}{B_{1-\alpha}} \sum_{t=1}^T R_{pt} \cdot 1 \left\{ R_{pt} \leq \hat{q}_{1-\alpha}^R \right\} \quad (57)$$

Indichiamo con  $\text{RM}(\mathbf{w})$  la misura di rischio di portafoglio  $\text{VaR}_\alpha$  o  $\text{ES}_w$ .

Assumiamo che il vettore  $n \times 1$  dei returns  $\mathbf{R}_t$  sia *iid* al variare del tempo  $t$  ma non facciamo assunzioni sulla distribuzione:

$$\begin{aligned} \{\mathbf{R}_1, \dots, \mathbf{R}_T\} &= \text{campione iid osservato} \\ R_{p,t} &= \mathbf{w}'\mathbf{R}_t \end{aligned}$$

Stimiamo i contributi marginali al rischio utilizzando la simulazione storica<sup>21</sup>

$$\hat{\mathbb{E}}^{\text{HS}}[R_{it} | R_{p,t} = \text{VaR}_\alpha] = \frac{1}{m} \sum_{t=1}^T R_{it} \cdot 1 \left\{ \widehat{\text{VaR}}_\alpha^{\text{HS}} - \epsilon \leq R_{p,t} \leq \widehat{\text{VaR}}_\alpha^{\text{HS}} + \epsilon \right\}$$

$$\hat{\mathbb{E}}^{\text{HS}}[R_{it} | R_{p,t} \leq \text{VaR}_\alpha] = \frac{1}{B_{1-\alpha}} \sum_{t=1}^T R_{it} \cdot 1 \left\{ R_{p,t} \leq \widehat{\text{VaR}}_\alpha^{\text{HS}} \right\} \quad (58)$$

In questo caso  $\text{VaR}_\alpha = q_{1-\alpha}^R$  e  $\widehat{\text{VaR}}_\alpha^{\text{HS}} = \hat{q}_{1-\alpha}^R$  è il quantile empirico  $1 - \alpha$  dei return.

Come si evince dall'espressione, lo stimatore di fatto calcola una media locale dei return nel vicinato di  $R_{p,t} = \text{VaR}_\alpha$ .

---

<sup>21</sup>Si tratta di un caso particolare dello stimatore non parametrico "kernel", dove la funzione peso è il kernel rettangolare. Si veda in proposito See Yamai Y. e T. Yoshida - *Comparative Analyses of Expected Shortfall and Value-at-Risk*, Institute for Monetary and Economic Studies, Bank of Japan (2002).

## Esempio

Al fine di comprendere gli argomenti trattati nel paragrafo, segue un esempio in cui vengono mostrate le modalità in cui ricavare le misure di rischio non parametriche per titoli realmente negoziati sui mercati finanziari.

Consideriamo i seguenti strumenti: MSFT (azione Microsoft negoziata sul NYSE) e un ETF<sup>22</sup> avente come sottostante il Nasdaq (indice tecnologico del NYSE).

Rappresentiamo in *Figura 9* uno scatterplot con i log-return empirici dei due strumenti nell'intervallo di tempo 30 nov 2005 e 1 nov 2017.

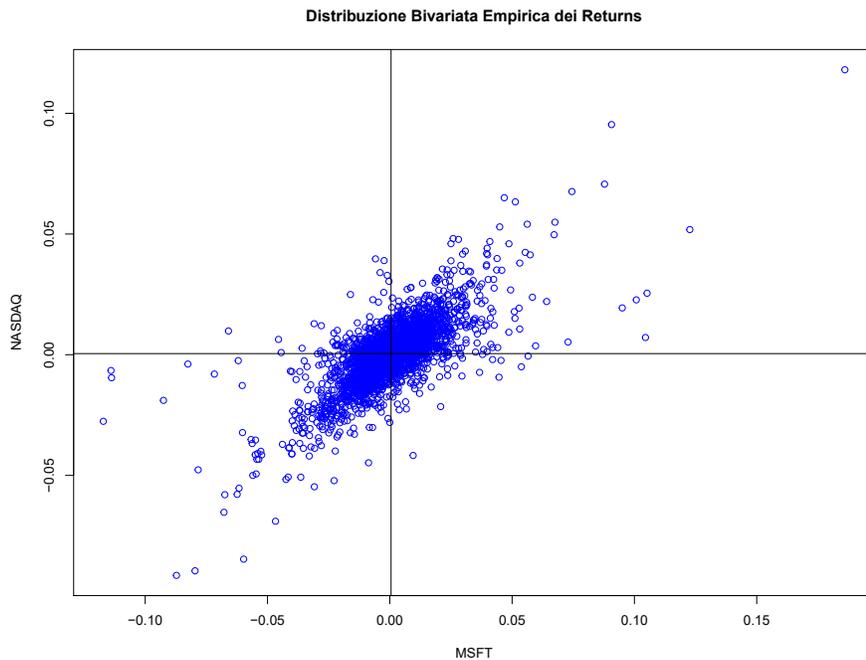


Figura 9: Scatterplot relativo ai log-returns giornalieri degli strumenti MSFT ed un ETF sul Nasdaq tra il 30 nov 2005 e il 1 nov 2017

---

<sup>22</sup>ETF è l'acronimo di Exchange Traded Fund, un termine con il quale si identifica una particolare tipologia di fondo d'investimento o Sicav che è negoziato in Borsa come un'azione ed ha come unico obiettivo d'investimento quello di replicare l'indice al quale si riferisce (benchmark) attraverso una gestione totalmente passiva.

Nel grafico possiamo notare una certa correlazione positiva tra i due set di log-return; questo fenomeno è motivato dalla presenza stessa di Microsoft all'interno del set di titoli che costituiscono l'indice Nasdaq, il quale è caratterizzato da un paniere di titoli afferenti al comparto IT che tendenzialmente reagiscono al medesimo modo a valle di eventi macroeconomici.

Costruiamo ora il nostro portafoglio equilibrando le quote dei due titoli. La distribuzione dei log-return del portafoglio così costituito risulta come da *Figura 10*.

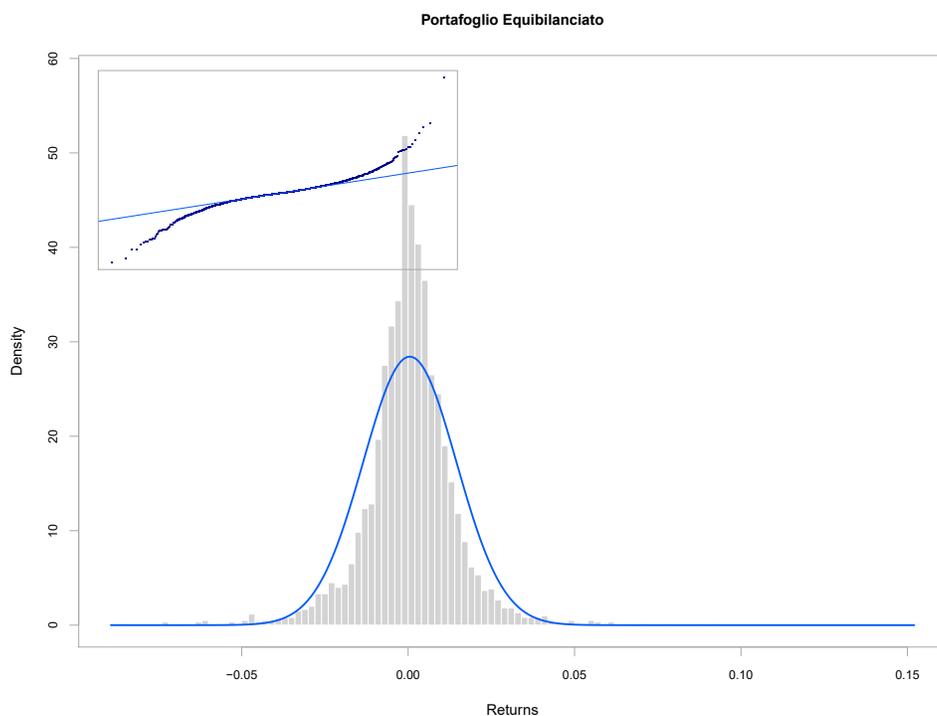


Figura 10: Distribuzione dei log-returns del portafoglio costituito dagli strumenti MSFT ed un ETF sul Nasdaq tra il 30 nov 2005 e il 1 nov 2017. In alto a sinistra il relativo qq-plot.

Possiamo notare come la distribuzione normale approssimi apprezzabilmente bene quella empirica a meno delle code (si veda qq-plot riportato nella medesima figura in alto a sinistra).

Le misure di rischio di portafoglio Value at Risk ed Expected Shortfall

calcolate usando, rispettivamente, le formule (54) e (56) risultano:

$$\text{VaR}_{5\%} = -0,02174 \quad \text{ES}_{5\%} = -0,03349$$

Per comprendere se l'ipotesi di distribuzione normale bivariata delle delle due componenti del portafoglio sia corretta, stimiamo i parametri di quest'ultima a partire dai campioni e proviamo a rappresentare la normale bivariata così descritta sullo stesso grafico di quella empirica (*Figura 11*).

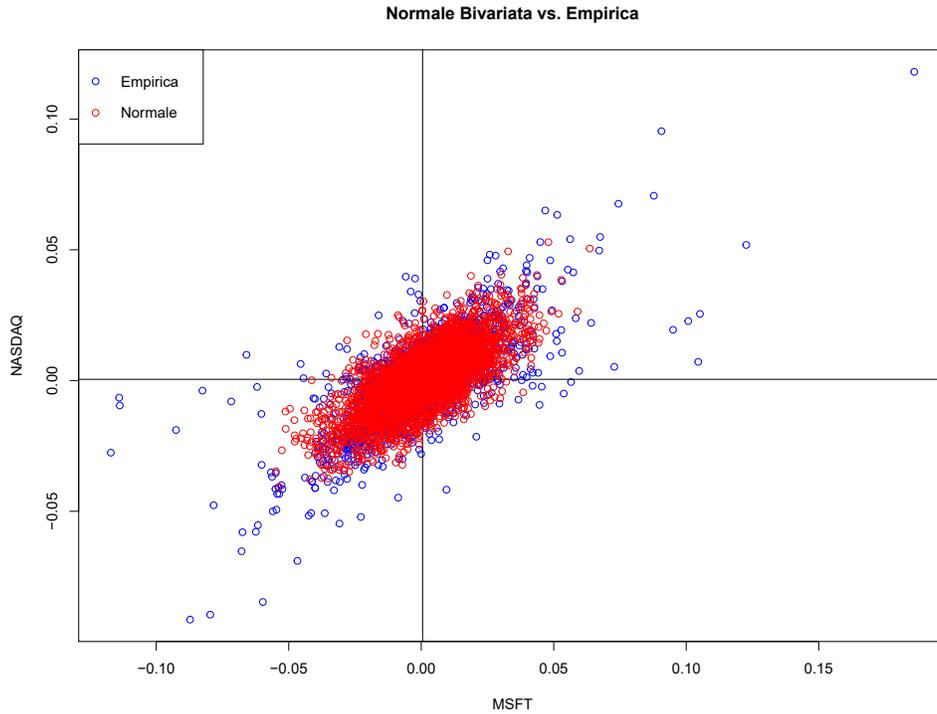


Figura 11: Distribuzione dei log-returns del portafoglio costituito dagli strumenti MSFT ed un ETF sul Nasdaq tra il 30 nov 2005 e il 1 nov 2017. In alto a sinistra il relativo qq-plot.

Considerata la normale bivariata  $N(\mu, \Sigma)$  I parametri stimati a partire dal campione risultano:

$$\hat{\mu} = \begin{pmatrix} 0,0005140 \\ 0,0004539 \end{pmatrix}$$

$$\hat{\Sigma} = \begin{pmatrix} 0,0002882 & 0,0001627 \\ 0,0001627 & 0,0001738 \end{pmatrix}$$

La distribuzione bivariata coincide con quella empirica molto finemente nella parte centrale, mentre le code di quella empirica si discostano sensibilmente, come rilevato già nel qq-plot in *Figura 10*.

Per concludere l'esempio, possiamo - utilizzando i dati sin qui considerati - rappresentare la contribuzione marginale di uno dei due titoli (ad esempio Microsoft) alle misure di rischio. In particolare calcoleremo e rappresenteremo (*Figura 12*) la contribuzione marginale all'Expected Shortfall di Microsoft ( $\text{MCEs}_{5\%}^{MSFT}$ ) usando la (58), ossia il valore atteso dei return di MSFT quando i return di portafoglio sono inferiori al  $\text{VaR}_{5\%}$ .

In *Figura 12* nel primo grafico sono evidenziati con dei mark rossi i return di portafoglio al di sotto del  $\text{VaR}_{5\%}$ ; considerando gli specifici return di MSFT associati a questi punti (mark rossi nel secondo grafico di *Figura 12*) e calcolando il valore medio di questi ultimi, si ottiene la contribuzione marginale di Microsoft all'Expected Shortfall di portafoglio.

L'Expected Shortfall del portafoglio equilibrato formato da MSFT e Nasdaq risulta:

$$\text{ES}_{5\%}^{\text{ptf}} = 0,05635$$

Il numero di return di portafoglio che, nell'intervallo di tempo considerato, risultano inferiori al  $\text{VaR}_{5\%}$  di portafoglio è:

$$(\# \text{ return di ptf} < \text{VaR}_{5\%}) = 149$$

A partire dalla (58), indichiamo la contribuzione marginale al rischio di Microsoft (retta rossa rappresentata nel secondo grafico di *Figura 12*) e quella del Nasdaq:

$$\text{MCEs}_{5\%}^{MSFT} = -0,06378 \quad \text{MCEs}_{5\%}^{Nasdaq} = -0,04891$$

A questo punto, utilizzando la (33), calcoliamo le due contribuzioni percentuali all'Expected Shortfall di portafoglio:

$$\text{PCES}_{5\%}^{MSFT} = 56,6\% \quad \text{PCES}_{5\%}^{Nasdaq} = 43,4\%$$

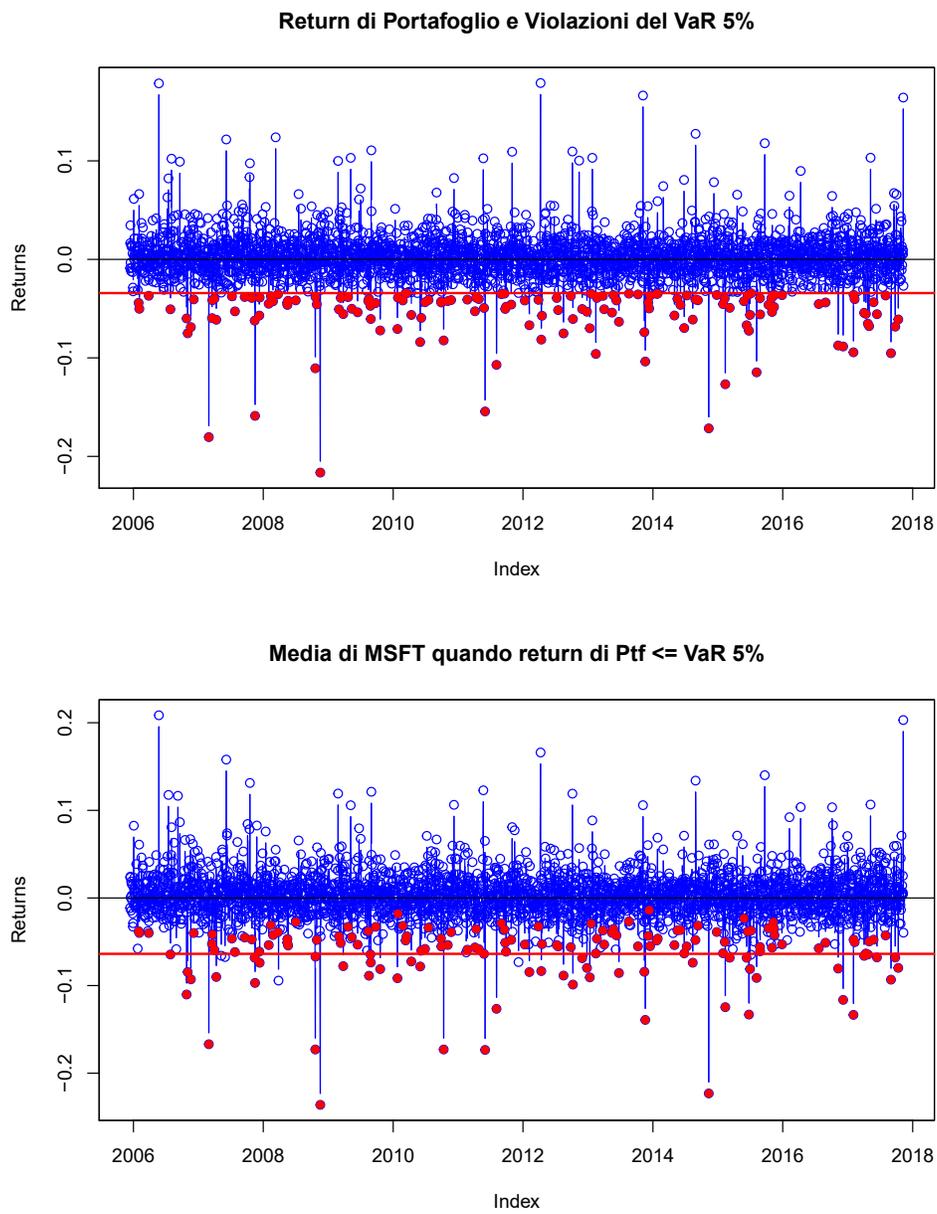


Figura 12: Rappresentazione delle violazioni dei return di portafoglio oltre  $VaR_{5\%}$  (primo grafico) e contribuzione marginale del titolo MSFT all'ES di portafoglio (retta rossa nel secondo grafico).

## 6.6 Stima parametrica delle misure di rischio di portafoglio

Consideriamo due casi rilevanti: quello in cui i return son distribuiti come una normale multivariate e quello in cui sono distribuiti come una t di Student multivariata.

### 6.6.1 Stima nel caso di distribuzione normale multivariata

$$\begin{aligned} \mathbf{R} &\sim N(\mu, \Sigma), \quad \theta = (\mu, \text{vech}(\Sigma))' \\ \Rightarrow R_p &\sim N(\mu_p, \sigma_p^2), \quad \mu_p = \mathbf{w}'\mu \quad e \quad \sigma_p^2 = \mathbf{w}'\Sigma\mathbf{w} \end{aligned}$$

Dove con  $\text{vech}(\Sigma)$  s'intende *vettorizzazione*, ossia quella trasformazione lineare che, impilando le colonne di  $\Sigma$ , trasforma la matrice in un vettore colonna.

Gli stimatori di massima verosimiglianza per  $\mu$  e  $\Sigma$  sono:

$$\begin{aligned} f_{\mathbf{R}}(\mathbf{r}) &= (2\pi)^{-n/2} \det(\Sigma)^{-1/2} \exp\left(-\frac{1}{2}(\mathbf{r} - \mu)' \Sigma^{-1}(\mathbf{r} - \mu)\right) \\ \hat{\mu}_{mle} &= \frac{1}{T} \sum_{t=1}^T \mathbf{R}_t \\ \hat{\Sigma}_{mle} &= \frac{1}{T} \sum_{t=1}^T (\mathbf{R}_t - \hat{\mu})(\mathbf{R}_t - \hat{\mu})' \end{aligned}$$

Le stime delle misure di rischio di portafoglio sono perciò:

$$\begin{aligned} \hat{\sigma}_{p,mle} &= (\mathbf{w}'\hat{\Sigma}_{mle}\mathbf{w})^{1/2} \\ \hat{q}_{1-\alpha}^{R_p}(\hat{\theta}_{mle}) &= \hat{\mu}_{p,mle} + \hat{\sigma}_{p,mle} \times q_{1-\alpha}^Z = \mathbf{w}'\hat{\mu}_{mle} + \hat{\sigma}_{p,mle} \times q_{1-\alpha}^Z \\ \mathbb{E}[R_p | R_p \leq q_{1-\alpha}^{R_p}(\hat{\sigma}_{p,mle})] &= - \left( \hat{\mu}_{p,mle} + \hat{\sigma}_{p,mle} \times \frac{\phi(q_{1-\alpha}^Z)}{1-\alpha} \right) \end{aligned}$$

### 6.6.2 Stima nel caso di distribuzione t di Student

Un vettore aleatorio  $\mathbf{Y}$  di dimensione  $n \times 1$  distribuito come una t di Student multivariata avente per media il vettore  $\mu$ , matrice di scala  $\Sigma$  e gradi di

libertà  $v$  può essere definito come

$$\begin{aligned}\mathbf{Y} &= \mu + \sqrt{\frac{v}{W}}\mathbf{Z} \\ \mathbf{Z} &\sim N(0, \Sigma) \\ W &\sim \chi_v^2 \text{ indipendente da } \mathbf{Z}\end{aligned}$$

Segue perciò

$$\begin{aligned}\mathbb{E}[\mathbf{T}] &= \mu \\ \text{var}(\mathbf{Y}) &= \frac{v}{(v-2)}\Sigma \neq \Sigma\end{aligned}$$

Considerando l'intero portafoglio si ha:

$$\begin{aligned}\mathbf{R} &\sim t(\mu, \Sigma, v) \text{ e } R_p = \mathbf{w}'\mathbf{R} \\ \Rightarrow R_p &\sim t(\mu_p, \sigma_p^2, v), \\ \mu_p &= \mathbf{w}'\mu, \quad \sigma_p^2 = \mathbf{w}'\Sigma\mathbf{w} \\ \text{var}(R_p) &= \frac{v}{v-2}\sigma_p^2\end{aligned}$$

## 7 Modelli per la volatilità

### 7.1 Attesa condizionata

Sia  $\{Y_t\}$  una sequenza di variabili aleatorie e sia  $\{I_t\}$  una sequenza di set di informazioni con  $I_t \in I$  per ogni  $t$  e  $I$  l'insieme universo di tutti i possibili set informativi. Seguono alcuni esempi di set informativi:

$$I_t = \{Y_1, Y_2, \dots, Y_t\} = \text{storia passata di } Y_t$$
$$I_t = \{(Y_s, Z_s)_{s=1}^t\}, \quad \{Z_t\} = \text{variabili ausiliarie}$$

Seguono alcune definizioni e proposizioni<sup>23</sup> che verranno utilizzate nel corso della trattazione:

**Definizione 1** (Attesa condizionata) *Sia  $Y_t$  una variabile aleatoria con funzione di densità condizionata  $f(y_t|I_s)$ , dove  $I_s$  è un set di informazioni con  $s < t$ . Allora*

$$\mu_{t|s} = \mathbb{E}[Y_t|I_s] = \int_{-\infty}^{\infty} y_t f(y_t|I_s) dy_t$$
$$\sigma_{t|s}^2 = \text{var}(Y_t|I_s) = \int_{-\infty}^{\infty} (y_t - \mathbb{E}[Y_t|I_s])^2 f(y_t|I_s) dy_t$$

**Nota:** siano  $X$  e  $Y$  due variabili aleatorie e  $I$  denoti un set di informazioni. Se  $X \in I$ , allora

$$\mathbb{E}[XY|I] = X\mathbb{E}[Y|I]$$

**Proposizione 1** *Legge dell'attesa iterata*

Siano  $I_1$  e  $I_2$  due set d'informazioni tali che  $I_1 \subseteq I_2$  e sia  $Y$  una variabile aleatoria tali che siano definiti  $\mathbb{E}[Y|I_1]$  e  $\mathbb{E}[Y|I_2]$ . Allora

$$\mathbb{E}[Y|I_1] = \mathbb{E}[\mathbb{E}[Y|I_2]|I_1] \quad (\text{l'insieme più piccolo vince})$$

Se  $I_1 = \emptyset$  (insieme nullo), allora

$$\mathbb{E}[Y|I_1] = \mathbb{E}[Y] \quad (\text{attesa non condizionata})$$
$$\mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[Y|I_2]|I_1] = \mathbb{E}[\mathbb{E}[Y|I_2]]$$

---

<sup>23</sup>David Williams - *Probability with Martingales*, Cambridge Mathematical Textbooks (1991)

**Definizione 2** (Differenza di martingala) *Sia  $Y_t$  una variabile aleatoria e  $I_t$  denoti un set d'informazioni. La coppia  $(Y_t, I_t)$  è una differenza di martingala (Martingale Difference Sequence) se*

1.  $I_t \in I_{t+1}$  (sequenza crescente di set d'informazioni - una filtrazione)
2.  $Y_t \in I_t$  ( $Y_t$  è adattato a  $I_t$ ; e.g.  $Y_t$  è un evento in  $I_t$ )
3.  $\mathbb{E}[Y_t | I_{t-1}] = 0$

## 7.2 Misure di rischio condizionate vs. non condizionate

Sia  $R_{t+1}$  il return di un asset tra i tempi  $t$  e  $t+1$  (composto nel continuo o nel discreto).

**Definizione 3** (Modellazione non condizionata) *La modellazione non condizionata di  $R_{t+1}$  è basata sulla distribuzione marginale non condizionata di  $R_{t+1}$ . Ovvero, le misure di rischio sono calcolate a partire dalla distribuzione marginale  $F_R$ .*

$$R_{t+1} \sim F_R, \quad \mathbb{E}[R_{t+1}] = \mu, \quad \text{var}(R_{t+1}) = \sigma^2$$

Definiamo

$$Z_{t+1} = \frac{R_{t+1} - \mu}{\sigma}, \quad Z_{t+1} \sim F_Z, \quad \mathbb{E}[Z_{t+1}] = 0, \quad \text{var}(Z_{t+1}) = 1$$

in modo tale che

$$R_{t+1} = \mu + \sigma \cdot Z_{t+1}$$

$I_t$  denoti l'informazione nota al tempo  $t$ . Per esempio  $I_t = \{R_t, R_{t-1}, \dots, R_0\}$  oppure  $I_t = \{(R_t, X_t), (R_{t-1}, X_{t-1}), \dots, (R_0, X_0)\}$

**Definizione 4** (Modellazione condizionata) *La modellazione condizionata di  $R_{t+1}$  è basata sulla distribuzione marginale condizionata di  $R_{t+1}$  dato  $I_t$ . Ovvero, le misure di rischio sono calcolate a partire dalla distribuzione marginale  $F_{R|I_t}$ .*

$$R_{t+1} \sim F_{R|I_t}, \quad \mathbb{E}[R_{t+1} | I_t] = \mu_{t+1|t}, \quad \text{var}(R_{t+1} | I_t) = \sigma_{t+1|t}^2$$

Definiamo

$$Z_{t+1} = \frac{R_{t+1} - \mu_{t+1|t}}{\sigma_{t+1|t}}, \quad Z_{t+1} \sim F_Z, \quad \mathbb{E}[Z_{t+1}] = 0, \quad \text{var}(Z_{t+1}) = 1$$

in modo tale che

$$R_{t+1} = \mu_{t+1|t} + \sigma_{t+1|t} \cdot Z_{t+1}$$

### 7.2.1 Media, varianza e volatilità condizionata

$$\begin{aligned}\mathbb{E}[R_{t+1}|I_t] &= \mu_{t+1|t} = \text{media condizionata} \\ \text{var}(R_{t+1}|I_t) &= \sigma_{t+1|t}^2 = \text{varianza condizionata} \\ \sigma_{t+1|t} &= \text{volatilità condizionata}\end{aligned}$$

Al variare di  $I_t$  nel tempo, variano  $\mu_{t+1|t}$  e  $\sigma_{t+1|t}$ .

Per i return giornalieri della maggior parte degli asset, tipicamente si assume  $\mu_{t|t-1} = \mu \approx 0$ . Ipotizzare, invece,  $\sigma_{t+1|t} = \sigma$  è da evitare. Un'evidenza empirica di questa considerazione verrà fornita nell'esempio del prossimo capitolo in *Figura 13*.

### 7.2.2 Misure di rischio condizionate

$$\begin{aligned}\sigma_{t+1|t} &= \text{volatilità condizionata} \\ q_\alpha^{R_{t+1|t}} &= \alpha \text{ quantile condizionato} \\ \mathbb{E}[R_{t+1}|R_{t+1}] &\leq q_\alpha^{R_{t+1|t}} = \text{Shortfall condizionato}\end{aligned}$$

### 7.2.3 Volatilità - Proprietà empiriche dei return

Con l'obiettivo di modellizzare il comportamento dei returns, è necessario tenere in considerazione una serie di proprietà derivanti dalle evidenze sperimentali riscontrate sulle timeseries dei principali strumenti quotati:

- Code pesanti

Le distribuzioni dei log-return sono tipicamente leptocurtiche; l'eccessiva curtosi diminuisce con l'aggregazione dei dati

- Clustering di volatilità  
Grandi variazioni seguono a grandi variazioni; piccole variazioni seguono a piccole variazioni
  
- Effetto leva  
I cambiamenti nel prezzo sono spesso correlati negativamente con le variazioni della volatilità<sup>24</sup>
  
- Periodi con pochi scambi  
La volatilità si riduce nei periodi in cui i mercati sono chiusi più frequentemente
  
- Eventi attesi o programmati  
La diffusione di informazioni attese (e.g. pubblicazione di bilanci, indicatori macroeconomici, etc.) è associata ad una volatilità ex-ante molto alta
  
- Volatilità ed autocorrelazione  
E' osservabile una relazione inversa tra volatilità e autocorrelazione degli indici di mercato
  
- Movimenti congiunti di volatilità  
Fattori comuni spiegano la volatilità di più serie di return (e.g. scandalo Diselgate che ha ripercussione su tutto il settore automotive)

---

<sup>24</sup>Come illustrato da Black F. - *The pricing of commodity contracts*, Journal of Financial Economics, 3, 167-179 (1976)

### 7.3 Modelli GARCH e ARCH

I modelli GARCH sono stati sviluppati per tenere conto delle proprietà empiriche sulla volatilità degli asset returns mostrate nel *Paragrafo 7.2.3*.

Sebbene i returns appaiano stazionari e poco autocorrelati, se si considerano i returns quadrati o in modulo, questi ultimi risultano decisamente autocorrelati. In particolare, la volatilità dei returns pare essere clusterizzabile: grandi variazioni di prezzo (positive o negative) sono seguite da altrettanto sensibili variazioni di prezzo.

Si noti il fenomeno di clustering della volatilità appena descritto nella rappresentazione seguente in *Figura 13*.

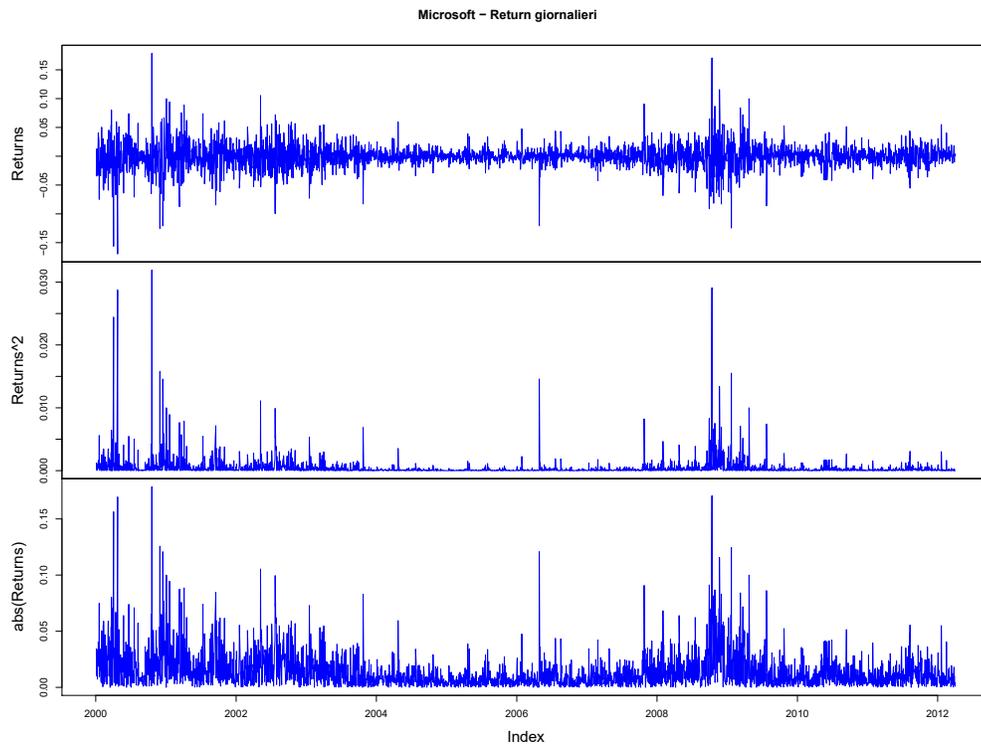


Figura 13: Distribuzione dei returns, dei returns quadrati e in modulo del titolo Microsoft tra gennaio 2000 e aprile 2012.

Provando, invece, a rappresentare le autocorrelazioni (si veda *Paragrafo 2.3.1*) tra i returns semplici, quadrati ed in modulo per il titolo Microsoft e l'indice S&P500 tra gennaio 2000 e aprile 2012 (si veda *Figura 14*) si evidenzia una forte correlazione seriale per i returns quadrati e, ancora più accentuata, per quelli in modulo.

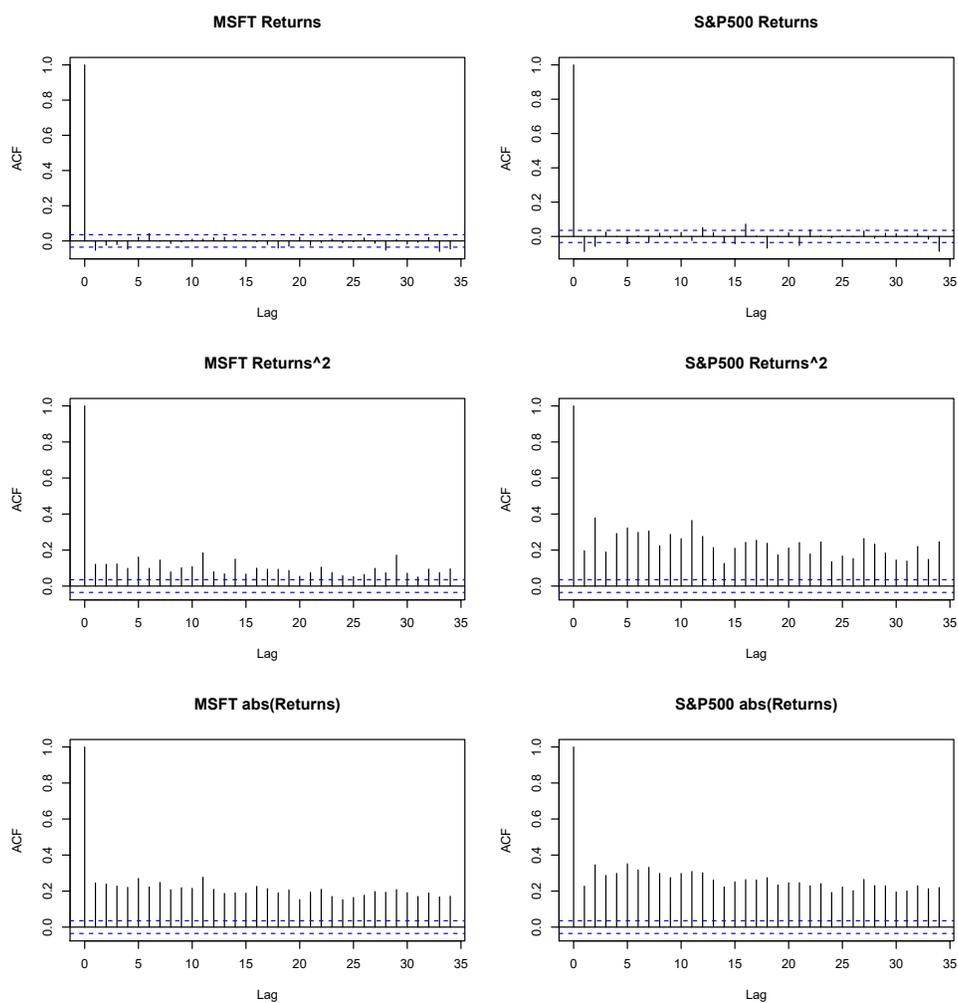


Figura 14: Autocorrelazione dei returns semplici, quadrati e in modulo del titolo Microsoft e dell'indice S&P500 tra gennaio 2000 e aprile 2012.

Questa autocorrelazione tra i returns quadrati e in modulo venne inizialmente modellata da Engle<sup>25</sup> con il modello ARCH (*AutoRegressive Conditional Heteroskedasticity*) e, in seguito da Bollerslev<sup>26</sup> con il modello *Generalized ARCH* (GARCH).

Quando si hanno a disposizione dati finanziari completi e intraday<sup>27</sup> la volatilità viene trattata secondo quelli che vengono chiamati *Financial markets on the intraday evolution of volatility models*<sup>28</sup> che non tratteremo in questa sede.

Volendo considerare una profondità storica significativa, i dati sui returns intraday risultano essere difficilmente reperibili. I modelli ARMA e GARCH, infatti, rispondo proprio alla specifica necessità di lavorare con timeseries di return con una risoluzione più bassa, tipicamente giornaliera.

Tali modelli sono utili a descrivere la varianza condizionata al variare del tempo, modellando - almeno parzialmente - anche il fenomeno delle code pesanti delle distribuzioni dei return.

Oltretutto, il fatto che i return tendano ad essere correlati negativamente con i cambiamenti di volatilità (come visto nel *Paragrafo 7.2.3*) rende i modelli GARCH particolarmente adatti a descriverli. Infatti, questi modelli hanno la prerogativa di poter modellare una grande varietà di proprietà derivanti dalle evidenze empiriche sul comportamento degli asset finanziari.

I più recenti studi, ricerche e pubblicazioni nell'ambito di questi modelli trovano fondamento nei lavori di Bollerslev, Engle e Nelson (1994), Palm (1996) e Li, Ling e McAleer (2002).

Nei successivi paragrafi del capitolo tratteremo i principali modelli autoregressivi per la volatilità, la cui struttura può essere modellata come segue:

---

<sup>25</sup>Robert F. Engle - *Econometrica*, Volume 50, Issue 4 (1982)

<sup>26</sup>Bollerslev T. - *Generalized autoregressive conditional heteroskedasticity*, Journal of Econometrics 31 (1986)

<sup>27</sup>Prezzi intraday: insieme di tutti i prezzi di scambio di un certo titolo all'interno di una singola giornata di contrattazione

<sup>28</sup>Eric Jondeau, Ser-Huang Poon e Michael Rockinger - *Financial Modeling Under Non-Gaussian Distributions*, Springer (2000) cap.3

$$r_t = \mu_t(\theta) + \epsilon_t \quad (59)$$

$$\epsilon_t = \sigma_t(\theta)z_t \quad (60)$$

con  $\theta$  set dei parametri del modello e dove

$$\begin{aligned} \mu_t(\theta) &= \mathbb{E}[r_t|I_{t-1}] \\ \sigma_t^2 &= \mathbb{E}[(r_t - \mu_t(\sigma))^2|I_{t-1}] \end{aligned}$$

Nella (59) abbiamo decomposto il return  $x_t$  in una media condizionata  $\mu_t(\theta)$  e un termine residuale  $\epsilon_t$ . La media condizionata può essere descritta da un processo ARMA( $p, q$ ) e/o avere delle stagionalità.  $I_t$  è il set d'informazioni disponibili al tempo  $t$ : può includere i return, i residui passati/presenti o, in generale, ogni altra variabile conosciuta al tempo  $t$ . In accordo con la (60), il termine residuale  $\epsilon_t$  ha una volatilità condizionata al set d'informazioni disponibile al tempo  $t - 1$  - indicata con  $\sigma_t$  - che può variare nel tempo. Assumeremo che la variabile  $z_t$  segua una qualche distribuzione con media 0 e varianza 1. Sebbene per semplicità ipotizzeremo inizialmente  $z_t$  distribuita normalmente, questa ipotesi non è necessaria.

Un modello di volatilità è un modello che descrive l'evoluzione di  $\sigma_t^2(\theta)$ . Esistono principalmente due modelli per descrivere le dinamiche di volatilità:

1. Nella prima categoria, la volatilità è descritta in modo deterministico come una funzione di un certo insieme di variabili. Questa categoria include i modelli (G)ARCH;
2. Nella seconda categoria, la volatilità è descritta in modo aleatorio attraverso l'uso di funzioni stocastiche. Questa categoria include i modelli di volatilità stocastica.

Il focus del capitolo sarà sulla prima categoria, sebbene in coda verranno riportati dei cenni ai modelli stocastici per la volatilità.

### 7.3.1 Modello ARCH( $p$ ) di Engle

*Intuizione:* usare un modello autoregressivo per descrivere la dipendenza temporale nella volatilità condizionata degli asset returns.

Il modello ARCH(p) per  $r_t = \ln P_t - \ln P_{t-1}$  è

$$\begin{aligned} r_t &= \mathbb{E}[r_t|I_{t-1}] + \epsilon_t, \quad \epsilon|I_{t-1} \sim iid(0, \sigma_t^2) \\ \sigma_t^2 &= a_0 + a_1\epsilon_{t-1}^2 + \dots + a_p\epsilon_{t-p}^2, \quad a_0 > 0, a_i \geq 0 \\ &= a_0 + a(L)\epsilon_t^2, \quad a(L) = \sum_{j=1}^p a_j L^j \end{aligned}$$

Ricordando che  $Lr_t = r_{t-1}$ ,  $L^2r_t = r_{t-2}$ , etc.

Come già accennato nella (60), è possibile descrivere il termine residuale anche come:

$$\begin{aligned} \epsilon_t &= z_t \sigma_t \\ z_t &\sim iid(0, 1) \\ \sigma_t^2 &= a_0 + a(L)\epsilon_t^2 \end{aligned}$$

Si consideri che la variabile  $z_t$  non dev'essere necessariamente normale, potrebbe - ad esempio - avere una distribuzione con code pesanti come la  $t$  di Student. Chiaramente, nel caso in cui la  $z_t$  non fosse normale, sarebbe necessario provvedere alla sua standardizzazione al fine di ottenere  $var(z_t) = \mathbb{E}[z_t^2] = 1$ .

### 7.3.2 ARCH(1) - Proprietà degli errori

Nel caso particolare di un modello ARCH(1), esplicitiamo in formule gli errori ed evidenziamo alcune loro proprietà:

$$\begin{aligned} \epsilon_t &= r_t - \mathbb{E}[r_t|I_{t-1}] \\ \epsilon_t &= z_t \sigma_t, \quad z_t \sim iid(0, 1) \\ \sigma_t^2 &= a_0 + a_1\epsilon_{t-1}^2, \quad a_0 > 0, a_1 \geq 0 \end{aligned}$$

- Il processo  $\{\epsilon_t\}$  è stazionario in covarianza<sup>29</sup> se  $0 \leq a_1 < 1$ .
- Il processo  $\{\epsilon_t, I_{t-1}\}$  è una *Martingale Difference Sequence* (si veda *Paragrafo 7.1*) con errori eteroschedastici (se  $a_1 \neq 0$ )

$$\begin{aligned} \mathbb{E}[\epsilon_t|I_{t-1}] &= \mathbb{E}[z_t \sigma_t|I_{t-1}] = \sigma_t \mathbb{E}[z_t|I_{t-1}] = 0 \\ var(\epsilon_t|I_{t-1}) &= \mathbb{E}[\epsilon_t^2|I_{t-1}] = \sigma_t^2 \mathbb{E}[z_t^2|I_{t-1}] = \sigma_t^2 \\ \mathbb{E}[\epsilon_t^m|I_{t-1}] &= 0 \text{ per } m \text{ dispari} \end{aligned}$$

<sup>29</sup>Eric Jondeau, Ser-Huang Poon e Michael Rockinger - *Financial Modeling Under Non-Gaussian Distributions*, Springer (2000) cap. 4.3

- $\epsilon_t$  ha media nulla e varianza non condizionata per tutti i  $t$

$$\begin{aligned}\mathbb{E}[\epsilon_t] &= \mathbb{E}[\mathbb{E}[z_t \sigma_t | I_{t-1}]] \\ &= \mathbb{E}[\sigma_t \mathbb{E}[z_t | I_{t-1}]] = 0 \\ \text{var}(\epsilon_t) &= \mathbb{E}[\epsilon_t^2] = \mathbb{E}[\mathbb{E}[z_t^2 \sigma_t^2 | I_{t-1}]] \\ &= \mathbb{E}[\sigma_t^2 \mathbb{E}[z_t^2 | I_{t-1}]] = \mathbb{E}[\sigma_t^2]\end{aligned}$$

Sotto condizione di stazionarietà e poichè  $\mathbb{E}[\epsilon_t^2] = \mathbb{E}[\sigma_t^2]$

$$\begin{aligned}\mathbb{E}[\sigma_t^2] &= \mathbb{E}[a_0 + a_1 \epsilon_{t-1}^2] = a_0 + a_1 \mathbb{E}[\epsilon_{t-1}^2] \\ &= a_0 + a_1 \mathbb{E}[\epsilon_t^2] \text{ per stazionarietà} \\ &= a_0 + a_1 \mathbb{E}[\sigma_t^2] \text{ poichè } \mathbb{E}[\epsilon_t^2] = \mathbb{E}[\sigma_t^2] \\ \Rightarrow \mathbb{E}[\sigma_t^2] &= \bar{\sigma}^2 = \frac{a_0}{1 - a_1} > 0 \text{ posto che } 0 \leq a_1 < 1\end{aligned}$$

- $\epsilon_t$  è un processo non correlato:  $\mathbb{E}[\epsilon_t \epsilon_{t-j}] = 0$  per  $j = 1, 2, \dots$

$$\begin{aligned}\mathbb{E}[\epsilon_t \epsilon_{t-1}] &= \mathbb{E}[\mathbb{E}[z_t \sigma_t z_{t-1} \sigma_{t-1} | I_{t-1}]] \\ &= \mathbb{E}[z_{t-1} \sigma_{t-1} \sigma_t \mathbb{E}[z_t | I_{t-1}]] = 0\end{aligned}$$

- $\epsilon_t$  è leptocurtica

$$\begin{aligned}\mathbb{E}[\epsilon_t^4] &= \mathbb{E}[\sigma_t^4 \mathbb{E}[z_t^4 | I_{t-1}]] = \mathbb{E}[\sigma_t^4] \cdot 3 \\ &\geq \left(\mathbb{E}[\sigma_t^2]\right)^2 \cdot 3 \text{ per la disuguaglianza Jensen}^{30} \\ &= \left(\mathbb{E}[\epsilon_t^2]\right)^2 \cdot 3 \\ \Rightarrow \frac{\mathbb{E}[\epsilon_t^4]}{\left(\mathbb{E}[\epsilon_t^2]\right)^2} &> 3\end{aligned}$$

Si ha quindi che

$$\text{kurt}(\epsilon_t) > 3 = \text{kurt}(\text{normale})$$

- $\sigma_t^2$  è una variabile aleatoria autocorrelata

$$\begin{aligned}\sigma_t^2 &= a_0 + a_1 \epsilon_{t-1}^2 \\ \mathbb{E}[\sigma_t^2] &= \frac{a_0}{1 - a_1} = \bar{\sigma}^2\end{aligned}$$

---

<sup>30</sup>Jensen J. L. W. V. - *Sur les fonctions convexes et les inégalités entre les valeurs moyennes*, Acta Mathematica (1906)

Usando  $a_0 = (1 - a_1)\bar{\sigma}^2$ ,  $\sigma_t^2$  può essere scritta come

$$\begin{aligned}\sigma_t^2 &= (1 - a_1)\bar{\sigma}^2 + a_1\epsilon_{t-1}^2 \\ \Rightarrow \sigma_t^2 - \bar{\sigma}^2 &= a_1(\epsilon_{t-1}^2 - \bar{\sigma}^2)\end{aligned}$$

- Il processo  $\{\epsilon_t^2\}$  può essere rappresentato in modo stazionario con un processo AR(1)

$$\begin{aligned}\sigma_t^2 + \epsilon_t^2 &= a_0 + a_1\epsilon_{t-1}^2 + \epsilon_t^2 \\ \Rightarrow \epsilon_t^2 &= a_0 + a_1\epsilon_{t-1}^2 + (\epsilon_t^2 - \sigma_t^2)\end{aligned}$$

dove  $(\epsilon_t^2 - \sigma_t^2) = v_t$  è una *Martingale Difference Sequence* (si veda *Paragrafo 7.1*) ereroschedastica.

### 7.3.3 ARCH(p) - Proprietà degli errori

Nel caso generale di un modello ARCH(p), esplicitiamo in formule gli errori ed evidenziamo alcune loro proprietà:

$$\begin{aligned}\epsilon_t &= z_t\sigma_t \\ z_t &\sim iid(0, 1) \\ \sigma_t^2 &= a_0 + a(L)\epsilon_t^2, \quad a(L) = a_0L + \dots + a_pL^p, \quad a_0 > 0, \quad a_i \geq 0\end{aligned}$$

- $\{\epsilon_t\}$  è stazionario in covarianza se  $0 \leq a_1 + \dots + a_p < 1$  ( $0 \leq a(1) < 1$ )
- $\mathbb{E}[\sigma_t^2] = \bar{\sigma}^2 = \frac{a_0}{1 - a_1 - \dots - a_p} = \frac{a_0}{a(1)}$ ,  $a(1) > 0$
- $\sigma_t^2$  è una variabile aleatoria autocorrelata  $\sigma_t^2 - \bar{\sigma}^2 = a(L)(\epsilon_t^2 - \bar{\sigma}^2)$
- $\epsilon_t^2$  può essere rappresentata in modo stazionario con un processo AR(p)  
 $\epsilon_t^2 = a_0 + a(L)\epsilon_t^2 + (\epsilon_t^2 - \sigma_t^2)$

### 7.3.4 Modello GARCH

Fino ad ora abbiamo presentato il modello ARCH, il quale si presenta come un modello AR per la volatilità (si veda *Paragrafo 3.1*). In questa sezione illustreremo il modello GARCH che, invece, può essere brevemente descritto come un modello ARMA (si veda *Paragrafo 3.3*) per la volatilità.

Il modello GARCH( $p, q$ ) è così definito:

$$\begin{aligned}\epsilon_t &= z_t \sigma_t, \quad z_t \sim iid(0, 1) \\ \sigma_t^2 &= a_0 + a(L)\epsilon_t^2 + b(L)\sigma_t^2, \quad a_0 > 0 \\ a(L) &= a_1 L + \dots + a_p L^p, \quad a_i \geq 0 \\ b(L) &= b_1 L + \dots + b_q L^q, \quad b_j \geq 0\end{aligned}$$

**Nota 1** Per identificare  $b_j$  è necessario avere almeno un coefficiente  $a_i > 0$ .

**Nota 2** Il modello GARCH(1,1) è molto spesso il miglior modello<sup>31</sup> da utilizzare per le applicazioni pratiche.

Il modello che viene più comunemente utilizzato è il GARCH(1,1):

$$\begin{aligned}\sigma_t^2 &= a_0 + a_1 \epsilon_{t-1}^2 + b_1 \sigma_{t-1}^2 \\ a_0 &> 0, \quad a_1 \geq 0, \quad b_1 \geq 0\end{aligned}$$

Ricordiamo alcune proprietà principali:

Condizione di stazionarietà:  $a_1 + b_1 < 1$

$$\text{ARCH}(\infty): a_i = a_1 b_1^{i-1}$$

$$\text{ARMA}(1,1): \epsilon_t^2 = a_0 + (a_1 + b_1)\epsilon_{t-1}^2 + \mu_t - b_1 \mu_{t-1}$$

$$\mu_t = \epsilon_t^2 - \mathbb{E}_{t-1}(\epsilon_t^2)$$

$$\text{Varianza non condizionata: } \bar{\sigma}^2 = \frac{a_0}{1 - a_1 - b_1}$$

Ci soffermiamo ora sulle proprietà specifiche del modello GARCH:

- (i) Il GARCH( $p, q$ ) è equivalente al modello ARCH( $\infty$ ). Se  $1 - b(z) = 0$  ha tutte le radici che giacciono all'esterno della circonferenza goniometrica, allora

$$\begin{aligned}\sigma_t^2 &= \frac{a_0}{1 - b(1)} + \frac{a(L)}{1 - b(L)} \epsilon_t^2 \\ &= a_0^* + \delta(L)\epsilon_t^2, \quad \delta(L) = \sum_{k=0}^{\infty} \delta_k L^k\end{aligned}$$

---

<sup>31</sup>Questa osservazione viene trattata in dettaglio da Peter R. Hansen, Asger Lunde - *A Forecast Comparison of Volatility Models: does anything beat a GARCH(1,1)?*, Journal of Applied Econometrics (2005)

- (ii)  $\epsilon_t$  è una *Martingale Difference Sequence* stazionaria ed ergodica con varianza finita, sotto l'ipotesi che  $a(1) + b(1) < 1$

$$\begin{aligned}\mathbb{E}[\epsilon_t] &= 0 \\ \text{var}(\epsilon_t) = \mathbb{E}[\epsilon_t^2] &= \frac{a_0}{1 - a(1) - b(1)} = \mathbb{E}[\sigma_t^2] = \bar{\sigma}^2 \\ \epsilon_t^2 &\sim \text{ARMA}(m, q), \quad m = \max(p, q)\end{aligned}$$

### 7.3.5 Specificazione media e varianza condizionata

$\mathbb{E}[r_t|I_{t-1}]$  è tipicamente specificata come una costante o, possibilmente, come un processo ARMA di ordine inferiore, al fine di descrivere l'autocorrelazione causata dalle condizioni micro-strutturali di mercato (e.g. rimbalzi denaro-lettera<sup>32</sup>) o dagli effetti derivanti dall'assenza di scambi.

Se durante il periodo di campionamento sono avvenuti eventi estremi o inusuali sul mercato (e.g. scoppio di una bolla speculativa, pubblicazione di notizie penalmente rilevanti per le imprese, decisioni delle Banche Centrali, etc.), allora certe variabili dummy<sup>33</sup> associate a questi eventi vengono spesso aggiunte alla specifica della media condizionata al fine di isolare gli effetti di tali eventi. In generale, questo tipo di specificazione della media condizionata è nella forma:

$$\mathbb{E}[r_t|I_{t-1}] = c + \sum_{i=1}^r \phi_i r_{t-i} + \sum_{j=1}^s \theta_j \epsilon_{t-j} + \sum_{l=0}^L \beta'_l x_{t-l} + \epsilon_t,$$

dove  $x_t$  è un vettore  $k \times 1$  di variabili esplicative esogene.

Le variabili esplicative esogene possono essere aggiunte alla formula della varianza condizionata:

$$\sigma_t^2 = a_0 + \sum_{i=1}^p a_i \epsilon_{t-i}^2 + \sum_{j=1}^q b_j \sigma_{t-j}^2 + \sum_{k=1}^K \delta'_k \mathbf{z}_{t-k},$$

---

<sup>32</sup>Gli importi denaro e lettera sono, rispettivamente, il prezzo unitario a cui il miglior offerente è disposto ad acquistare un determinato asset e il prezzo unitario più basso a cui un proprietario è disposto a cedere il medesimo asset. Nel breve periodo le logiche di fluttuazione denaro-lettera sono governati dai modelli microeconomici. P. Krugman e R. Wells - *Microeconomia*, Zanichelli (2006).

<sup>33</sup>Anche conosciuta come *variabile di comodo*, è una variabile che può assumere unicamente i valori 0 e 1, a secondo che sia soddisfatta o meno una data condizione.

dove  $\mathbf{z}_t$  è un vettore  $m \times 1$  di variabili e  $\delta$  è un vettore  $m \times 1$  di coefficienti positivi.

Le variabili esogene generalmente utilizzate come supporto nella previsione della volatilità sono: il volume degli scambi, i tassi di interesse, gli annunci di notizie macroeconomiche, la volatilità implicita e realizzata dei prezzi delle opzioni, i return overnight e la volatilità realizzata nell'after hour.

### 7.3.6 GARCH-in-Mean (GARCH-M)

Le moderne teorie finanziarie suggeriscono che la volatilità dovrebbe dipendere dai *premi di rischio*<sup>34</sup> degli asset. Il modello GARCH-M permette di considerare una volatilità variabile nel tempo dipendente dai return attesi:

$$\begin{aligned} r_t &= c + \alpha g(\sigma_t) + \epsilon_t \\ \epsilon_t &\sim \text{GARCH} \\ g(\sigma_t) &= \begin{cases} \sigma_t \\ \sigma_t^2 \\ \ln(\sigma_t^2) \end{cases} \end{aligned}$$

La scelta di  $g(\sigma_t)$  dipende dal tipo di esposizione al rischio a cui l'asset è sottoposto. Condizionano perciò la scelta fattori quali il tipo di asset (e.g. obbligazioni, azioni, derivati, etc.), il livello di liquidità, la solvibilità della controparte, etc.

### 7.3.7 Aggregazione Temporale

Il fenomeno di clustering di volatilità e il comportamento non-gaussiano dei return finanziari è tipicamente evidente nei dati settimanali, giornalieri o intraday. La persistenza della volatilità condizionata tende a crescere all'aumentare della frequenza di campionamento.

Per i modelli GARCH non esiste un principio di aggregazione semplice che colleghi i parametri del modello ad una certa frequenza di campionamento

---

<sup>34</sup>Il *premio di rischio* è la differenza tra il valore atteso di una lotteria (variabile casuale) rischiosa e l'ammontare certo (detto *certo equivalente*) che un individuo sarebbe disposto ad accettare al posto della lotteria rischiosa. E. Pitacco - *Elementi di Matematica delle Assicurazioni*, Edizioni LINT (2002).

a quelli specificati per una diversa frequenza. Questo accade perché i modelli GARCH implicano che il processo descritto dai residui quadrati segua un processo di tipo ARMA con le innovazioni di tipo *Martingale Difference Sequence*, che non sono indipendenti rispetto alle aggregazioni temporali.

Nelle applicazioni pratiche è necessario perciò un tuning manuale dei parametri, sebbene questa strategia non fornisca una soluzione generalmente applicabile per le previsioni di volatilità. Per esempio, è stato mostrato<sup>35</sup> che un modello GARCH che fitta i return giornalieri dell'S&P 500 produce previsioni migliori sulla volatilità settimanale e mensile rispetto ad un GARCH costruito direttamente sui return settimanali e mensili.

### 7.3.8 Testare gli effetti di un modello ARCH

Consideriamo il caso in cui si voglia effettuare il test di ipotesi:

$$H_0 : (\text{NO ARCH}) \quad a_1 = a_2 = \dots = a_p = 0$$

$$H_1 : (\text{ARCH}) \quad \text{almeno un } a_i \neq 0$$

Engle ha derivato un semplice test LM che può essere sintetizzato nei seguenti step:

- Step 1: calcolare i residui quadrati  $\epsilon_t$  dall'equazione della media nella regressione
- Step 2: stimare la regressione ausiliaria

$$\hat{\epsilon}_t^2 = a_0 + a_1 \hat{\epsilon}_{t-1}^2 + \dots + a_p \hat{\epsilon}_{t-p}^2 + \text{errore}_t$$

- Step 3: costruire la statistica di test LM

$$LM_{\text{ARCH}} = T \cdot R_{\text{AUX}}^2$$

dove  $T$  = campione proveniente dalla regressione ausiliaria ed  $R_{\text{AUX}}^2$  è l'R-quadro non centrato della regressione ausiliaria. Sotto l'ipotesi nulla che tutti i parametri  $a_i$  siano congruenti a 0, è possibile dimostrare che la statistica  $LM_{\text{ARCH}}$  si distribuisce come un  $\chi^2(p)$ :

$$H_0 : (\text{NO ARCH}) \quad \longrightarrow \quad LM_{\text{ARCH}} \sim \chi^2(p)$$

---

<sup>35</sup>M. Martens - *Measuring and Forecasting S&P 500 Index-Futures Volatility Using High-Frequency Data*, The Journal of Futures Markets (2002).

### 7.3.9 Stimare il GARCH con MLE

Considerando il modello di stima

$$\begin{aligned} r_t &= \mathbb{E}[r_t|I_{t-1}] + \epsilon_t = \mathbf{x}'_t \beta + \epsilon_t \\ \epsilon_t &= z_t \sigma_t, \quad z_t \sim iidN(0, 1) \\ \sigma_t^2 &= a_0 + a(L)\epsilon_t^2 + b(L)\sigma_t^2 \end{aligned}$$

è possibile ottenere i parametri  $\beta$  della regressione e quelli del GARCH.

$\gamma = (a_0, a_1, \dots, a_p, b_1, \dots, b_q)'$  può essere stimata separatamente poiché la matrice d'informazione per  $\theta = (\beta', \gamma)'$  è diagonale a blocchi.

In particolare,

- Step 1: stimare  $\beta$  col metodo dei minimi quadrati ignorando gli errori dell'ARCH e costruire i residui  $\hat{\epsilon}_t = r_t - \mathbf{x}'_t \hat{\beta}$
- Step 2: stimare il processo ARCH per i residui  $\hat{\epsilon}_t$  attraverso la massima verosimiglianza (MLE)

Si noti che la matrice d'informazione potrebbe non essere diagonale a blocchi se la funzione di densità di probabilità di  $z_t$  non è simmetrica o se  $\beta$  e  $\gamma$  non sono prive di variazioni (e.g. modelli GARCH-M).

Sotto ipotesi di normalità, determiniamo ora la funzione di verosimiglianza per il GARCH.

Assumiamo  $\mathbb{E}[r_t|I_{t-1}] = 0$  e che  $\theta = (a_0, a_1, \dots, a_p, b_1, \dots, b_q)'$  sia il set di parametri da stimare. Poiché  $\epsilon_t = z_t \sigma_t$

$$\begin{aligned} f(\epsilon_t|I_{t-1}; \theta) &= f(z_t) \left| \frac{\partial z_t}{\partial \epsilon_t} \right| = f\left(\frac{\epsilon_t}{\sigma_t}\right) \left| \frac{1}{\sigma_t} \right| \\ &= (2\pi\sigma_t^2)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma_t^2}\epsilon_t^2\right\} \end{aligned}$$

Per un campione di taglia  $T$ , la decomposizione dell'errore di previsione restituisce

$$\begin{aligned} f(\epsilon_T, \epsilon_{T-1}, \dots, \epsilon_1, \theta) &= \left( \prod_{t=m+1}^T f(\epsilon_t|I_{t-1}; \theta) \right) \cdot f(\epsilon_1, \dots, \epsilon_m; \theta) \\ &= \left( \prod_{t=m+1}^T (2\pi\sigma_t^2)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma_t^2}\epsilon_t^2\right\} \right) \cdot f(\epsilon_1, \dots, \epsilon_m; \theta) \end{aligned}$$

Si tenga presente che:

- $\sigma_t^2 = a_0 + a(L)\epsilon_t^2 + b(L)\sigma_t^2$  va valutata ricorsivamente dato  $\theta$  ed i valori iniziali  $\sigma_t^2$  ed  $\epsilon_t^2$ . Per esempio, considerando il GARCH(1, 1)

$$\sigma_1^2 = a_0 + a_1\epsilon_0^2 + b_1\sigma_0^2,$$

una volta specificati i valori iniziali  $\epsilon_0^2$  e  $\sigma_0^2$ , tutti gli altri valori di  $\sigma_t^2$  al variare del tempo  $t$  possono essere calcolati ricorsivamente.

- La funzione di log-verosimiglianza è

$$-\frac{(T-m)}{2} \ln(2\pi) - \sum_{t=m+1}^T \left[ \frac{1}{2} \ln(\sigma_t^2) + \frac{1}{2} \frac{\epsilon_t^2}{\sigma_t^2} \right] + \ln(f(\epsilon_1, \dots, \epsilon_m; \gamma))$$

La densità marginale per i valori iniziali  $f(\epsilon_1, \dots, \epsilon_m; \theta)$  non è esprimibile in forma chiusa, perciò non è possibile ottenere una stima di massima verosimiglianza esatta. Operativamente, i valori iniziali  $\epsilon_1, \dots, \epsilon_m$  sono fissati a 0 e la densità marginale  $f(\epsilon_1, \dots, \epsilon_m; \theta)$  viene ignorata. Si è così ottenuta una stima di massima verosimiglianza condizionata.

Seguono alcune indicazioni operative di cui solitamente ci si avvale nel calcolo delle stime:

- Utilizzando la log-verosimiglianza è necessario scegliere i parametri  $a_i$  ( $i = 0, \dots, p$ ) e  $b_j$  ( $j = 0, \dots, q$ ) ed inizializzare  $\epsilon_t^2$  e  $\sigma_t^2$
- I parametri sulla varianza condizionata vengono solitamente fissati a 0, mentre  $a_0$  viene posto uguale alla varianza incondizionata di  $r_t$ . Per i valori iniziali di  $\sigma_t^2$  una scelta comune è

$$\sigma_t^2 = \epsilon_t^2 = \frac{1}{T} \sum_{s=m+1}^T r_s^2, \quad t \leq m$$

- Il metodo della *variance targeting*<sup>36</sup> (puntamento della varianza) può essere utilizzato per eliminare la stima di  $a_0$  e migliorare la stabilità numerica. A titolo di esempio applichiamo tale metodo nel GARCH(1, 1)

$$\begin{aligned} \sigma_t^2 &= a_0 + a_1\epsilon_{t-1}^2 + b_1\sigma_{t-1}^2 \\ \bar{\sigma}^2 &= \frac{a_0}{1 - a_1 - b_1} \Rightarrow a_0 = \bar{\sigma}^2(1 - a_1 - b_1) \end{aligned}$$

<sup>36</sup>John C. Hull - *Risk management e istituzioni finanziari*, Pearson (2008) cap. 5

$$\begin{aligned}\Rightarrow \sigma_t^2 &= \bar{\sigma}^2(1 - a_1 - b_1) + a_1\epsilon_{t-1}^2 + b_1\sigma_{t-1}^2 \\ &= \bar{\sigma}^2 + a_1(\epsilon_{t-1}^2 - \bar{\sigma}^2) + b_1(\sigma_{t-1}^2 - \bar{\sigma}^2)\end{aligned}$$

Nella stima,  $\bar{\Sigma}^2$  è fissato al valore della varianza dei return semplici e non è un parametro da stimare.

- Dopo aver inizializzato la log-verosimiglianza, la funzione può essere massimizzata utilizzando tecniche di ottimizzazione numeriche. Il metodo più comunemente utilizzato si basa sull'iterazione di Newton-Raphson<sup>37</sup> nella forma

$$\hat{\theta}_{n+1} = \hat{\theta}_n - \lambda_n \mathbf{H}(\hat{\theta}_n)^{-1} s(\hat{\theta}_n)$$

- Per i modelli GARCH, spesso viene adoperato l'algoritmo di ottimizzazione numerica BHHH<sup>38</sup>. Questo algoritmo approssima la matrice hessiana usando solo l'informazione della derivata prima

$$-\mathbf{H}(\theta) \approx \mathbf{B}(\theta) = \sum_{t=1}^T \frac{\partial l_t}{\partial \theta} \frac{\partial l_t}{\partial \theta'}$$

- Sotto condizioni opportune di regolarità, le stime di massima verosimiglianza sono consistenti, distribuite asintoticamente come normali e una stima della matrice di covarianza asintotica delle stime di massima verosimiglianza è costruibile a partire dalla stima della matrice hessiana finale ottenuta con l'algoritmo di ottimizzazione numerica scelto.

**Osservazione:** sulla base dell'esperienza d'uso, la stima del GARCH disponibile sui software commerciali tipicamente viene implementata lasciando come variabili da scegliere i valori di partenza e la tipologia di algoritmo, criteri di convergenza e spesso anche l'uso del calcolo differenziale in modo analitico oppure numerico. Tutte queste variabili influenzano la stima numerica dei parametri del modello GARCH. Inoltre, la funzione di verosimiglianza - o log-verosimiglianza - del GARCH non ha sempre un comportamento regolare: specialmente nei modelli molto complessi e con molti parametri non è

<sup>37</sup>Conosciuto più comunemente come *metodo delle tangenti*, si basa sull'espansione di Taylor a secondo ordine della funzione di massima verosimiglianza.

<sup>38</sup>Il Berndt-Hall-Hall-Hausman (BHHH) è un algoritmo di ottimizzazione numerica simile a quello di Gauss-Newton, maggiori dettagli nel paper di Berndt E., Hall B., Hall R., Hausman J. - *Estimation and Inference in Nonlinear Structural Models*, Annals of Economic and Social Measurement (1974).

garantito che si riesca a trovare il massimo globale di tale funzione (e.g. la scelta sconsigliata di valori di partenza dell'algoritmo possono condurre a problemi di convergenza).

### 7.3.10 Stimare il GARCH con QMLE

Nella sezione precedente abbiamo ipotizzato che la distribuzione condizionata dell'errore  $\epsilon_t$  sia normale. Tale assunzione non è da considerarsi sempre appropriata.

In ogni caso, anche quando l'assunzione è inappropriata, l'utilizzo delle *quasi-maximum likelihood estimates* (QMLEs - stima di quasi-massima verosimiglianza), le quali sono consistenti e asintoticamente normalmente distribuite, garantisce che la media e varianza condizionata siano correttamente specificate.

Una matrice di covarianza asintotica per le *quasi-maximum likelihood estimates*, la quale è robusta all'ipotesi di non-normalità è stimata utilizzando:

$$H(\hat{\theta}_{QML})^{-1}B(\hat{\theta}_{QML})H(\hat{\theta}_{QML})^{-1},$$

dove  $\hat{\theta}_{QML}$  denota la stima di quasi-massima verosimiglianza di  $\theta$  e in letteratura viene spesso chiamata *sandwich estimator*<sup>39</sup>.

### 7.3.11 GARCH - Determinare i Lag

Per la determinazione dei lag si occorre a vari criteri di selezione. I più utilizzati sono AIC e BIC.

Il criterio d'informazione di Akaike<sup>40</sup> ( Akaike's Information Criteria - AIC), è un metodo per la valutazione e il confronto tra modelli statistici che fornisce una misura della qualità della stima di un modello statistico tenendo conto sia della bontà di adattamento che della complessità del modello. Tendenzialmente è opportuno scegliere il modello con il valore AIC più basso. In generale il modello è

$$AIC = 2k - 2\ln(L),$$

---

<sup>39</sup>Per approfondimenti sulle tematiche che vedono coinvolta la robustezza sotto ipotesi di non-normalità si veda: David A. Freedman - *On the So-Called "Huber Sandwich Estimator" and "Robust Standard Errors"*, The American Statistician (2005)

<sup>40</sup>Il criterio è stato presentato alla comunità matematica per la prima volta proprio da Akaike Hirotugu - *A new look at the statistical model identification*, IEEE (1974)

dove  $k$  è il numero di parametri del modello statistico e  $L$  è il valore massimizzato della funzione di verosimiglianza del modello.

Il secondo criterio per la selezione di modelliche abbiamo citato, ovvero il Criterio di Informazione Bayesiano<sup>41</sup> (Bayesian Information Criteria - BIC), è allo stesso modo un criterio che introduce un termine di penalizzazione per il numero di parametri del modello, evitando così il fenomeno di overfitting del modello. Nel BIC la penalizzazione risulta addirittura più forte rispetto all'AIC.

Generalmente è definito come:

$$AIC = -2\ln(L) + k \ln(n),$$

dove  $k$  è il numero di parametri del modello statistico,  $L$  è il valore massimizzato della funzione di verosimiglianza del modello stimato e  $n$  il numero di osservazioni.

Utilizzando tali criteri di selezione, per i modelli GARCH( $p, q$ ) usualmente vengono selezionati  $p, q \leq 2$ . Come già evidenziato è difficile trovare casi in cui non si debba scegliere il GARCH(1, 1).

I modelli GARCH( $p, q$ ) di piccolo ordine (ovvero  $p, q \leq 2$ ) sono generalmente da preferirsi rispetto ai modelli ARCH( $p$ ) di ordine maggiore per varie motivazioni (e.g. migliore stabilità numerica della stima, complessità del modello inferiore, etc.).

### 7.3.12 GARCH - Predire la media condizionata

Assumiamo il modello sia così specificato:

$$\begin{aligned} r_t &= \mathbb{E}[r_t | I_{t-1}] + \epsilon_t = \mu + \epsilon_t \\ \epsilon_t &= z_t \sigma_t, \quad z_t \sim iid N(0, 1) \\ \sigma_t^2 &= a_0 + a_1 \epsilon_{t-1}^2 + b_1 \sigma_{t-1}^2 \end{aligned}$$

Supponiamo di essere interessati a prevedere i valori futuri di  $r_{T+h}$  basandoci sul set di informazioni  $I_t$ . In questo caso, l'errore medio quadratico minimo dato dalla previsione ad  $h$  tempi successivi  $r_{T+h}$  è:

$$\mathbb{E}[r_{T+h} | I_T] = \mu \quad e \quad \epsilon_{T+h} = r_{T+h} - \mu,$$

---

<sup>41</sup>Il BIC è stato sviluppato da Gideon E. Schwarz - *Estimating the dimension of a model*, Annals of Statistics (1978)

che non dipende dai parametri del GARCH.

Per produrre degli intervalli di confidenza per gli errori è necessario calcolare la varianza condizionata di degli errori per cui abbiamo fatto la previsione.

$$\text{var}(r_{T+h}|I_T) = \mathbb{E}[\sigma_{T+h}^2|I_T],$$

che analogamente non dipende dai parametri del GARCH. Dunque gli intervalli di confidenza dipenderanno direttamente dalla previsione della volatilità  $\mathbb{E}[\sigma_{T+h}^2|I_T]$ .

Vediamo come predire la volatilità partendo da un caso semplice e poi cerchiamo di generalizzare. Consideriamo il modello GARCH(1, 1):

$$\sigma_t^2 = a_0 + a_1\epsilon_{t-1}^2 + b_1\sigma_{t-1}^2$$

con  $t = 1, 2, \dots, T$ . Il miglior predittore lineare di  $\sigma_{t+1}^2$  utilizzando il set di informazioni al tempo  $t$  è:

$$\begin{aligned}\mathbb{E}[\sigma_{T+1}^2|I_T] &= a_0 + a_1\mathbb{E}[\sigma_T^2|I_T] + b_1\mathbb{E}[\sigma_T^2|I_T] \\ &= a_0 + a_1\epsilon_{T-1}^2 + b_1\sigma_{T-1}^2\end{aligned}$$

Usando il fatto che  $\mathbb{E}[\epsilon_{T+1}^2|I_T] = \mathbb{E}[\sigma_{T+1}^2|I_T]$

$$\begin{aligned}\mathbb{E}[\sigma_{T+2}^2|I_T] &= a_0 + a_1\mathbb{E}[\sigma_{T+1}^2|I_T] + b_1\mathbb{E}[\sigma_{T+1}^2|I_T] \\ &= a_0 + (a_1 + b_1)\mathbb{E}[\sigma_{T+1}^2|I_T]\end{aligned}$$

In generale per  $k \geq 2$

$$\begin{aligned}\mathbb{E}[\sigma_{T+2}^2|I_T] &= a_0 + (a_1 + b_1)\mathbb{E}[\sigma_{T+k-1}^2|I_T] \\ &= a_0 \sum_{i=0}^{k-1} (a_1 + b_1)^i + (a_1 + b_1)^{k-1} (a_1\epsilon_t^2 + b_1\sigma_T^2)\mathbb{E}[\sigma_{T+1}^2|I_T]\end{aligned}$$

**Nota:** Se  $|a_1 + b_1| < 1$ , allora per  $k \rightarrow \infty$

$$\mathbb{E}[\sigma_{T+k}^2|I_T] \rightarrow \mathbb{E}[\sigma_T^2] = \frac{a_0}{1 - a_1 - b_1}$$

Esiste una rappresentazione alternativa dell'equazione previsionale. Partiamo dalla forma con media aggiustata

$$\sigma_{T+1}^2 - \bar{\sigma}^2 = a_1 (\epsilon_T^2 - \bar{\sigma}^2) + b_1 (\sigma_T^2 - \bar{\sigma}^2)$$

dove  $\bar{\sigma}^2 = a_0/(1 - a_1 - b_1)$  è la varianza non condizionata. Con una sostituzione ricorsiva otteniamo

$$\mathbb{E}_t \left[ \sigma_{T+k}^2 \right] - \bar{\sigma}^2 = (a_1 + b_1)^{k-1} \left( \mathbb{E} \left[ \sigma_{T+1}^2 \right] - \bar{\sigma}^2 \right).$$

### 7.3.13 EWMA come modello previsionale

L'algoritmo previsionale del modello GARCH(1,1) è fortemente collegato alla media mobile esponenziale pesata (EWMA - *Exponentially Weighted Moving Average*) dei precedenti valori di  $\epsilon_t^2$ . Questo tipo di modello previsionale è stato proposto per la prima volta dal gruppo di metriche di rischio in J.P. Morgan.

Il modello previsionale EWMA di  $\sigma_{T+1}^2$  ha la forma

$$\sigma_{T+1,EWMA}^2 = (1 - \lambda) \sum_{s=0}^{\infty} \lambda^s \epsilon_{T-s}^2 = \sum_{s=0}^{\infty} w_s \epsilon_{T-s}^2, \quad \lambda \in (0, 1)$$

$$w_s = (1 - \lambda) \lambda^s, \quad \sum_{s=0}^{\infty} w_s = 1$$

Come si osserva, il modello previsionale per  $\sigma_{T+1}^2$  si basa su pesi geometricamente decrescenti determinati da  $\lambda$ , in particolare

$$\lambda \approx 1 \Rightarrow \text{pesi decrescenti molto lentamente}$$

$$\lambda \approx 0 \Rightarrow \text{pesi decrescenti molto rapidamente}$$

Definiamo come *half-life* (mezza vita) del modello previsionale EWMA il numero di periodi necessari a dimezzarne i pesi

$$\text{half-life} = \frac{\ln(0.5)}{\ln(\lambda)}$$

In particolare, per dati a granularità giornaliera, J.P. Morgan suggerisce l'uso di  $\lambda = 0.94$  per disporre della corretta sensibilità previsionale a breve termine

$$\lambda = 0.94 \Rightarrow \text{half-life} = \frac{\ln(0.5)}{\ln(0.94)} = 11.2 \text{ giorni}$$

Attraverso sostituzioni ricorsive, il modello previsionale EWMA può essere espresso nella forma

$$\sigma_{T+1,EWMA}^2 = (1 - \lambda)\epsilon_T^2 + \lambda\sigma_{T,EWMA}^2$$

che è nella forma del modello GARCH(1, 1) con  $a_0 = 0$ ,  $a_1 = 1 - \lambda$  e  $b_1 = \lambda$ . Questo modello viene chiamato *integrated GARCH* - IGARCH(1, 1).

In questo modello  $a_1 + b_1 = 1 - \lambda + \lambda = 1$  dunque  $\epsilon_t$  non è stazionario in covarianza e così  $\mathbb{E}[\epsilon_t^2] = \mathbb{E}[\sigma_t^2]$ . Inoltre l' EWMA forecast si comporta verosimilmente come un random walk e ciò significa che la miglior predizione è il valore corrente.

$$\begin{aligned}\sigma_{T+1,EMWA}^2 &= (1 - \lambda)\epsilon_T^2 + \lambda\sigma_{T,EMWA}^2 \\ \sigma_{T+1,EMWA}^2 &= (1 - \lambda)\epsilon_{T+1}^2 + \lambda\sigma_{T+1,EMWA}^2 \\ \sigma_{T+1,EMWA}^2 &= (1 - \lambda)\sigma_{T+1}^2 + \lambda\sigma_{T+1,EMWA}^2 = \sigma_{T+1,EMWA}^2\end{aligned}$$

### 7.3.14 Prevedere la volatilità dei Return Multiperiodali

Sia  $r_t = \ln P_t - \ln P_{t-1}$ . E' possibile fare delle previsioni utilizzando il modello GARCH per la volatilità giornaliera a differenti giorni di distanza di tempo. Per la gestione dei rischi e per il pricing delle opzioni, a partire dalla volatilità intesa come processo stocastico, sono necessarie previsioni dei return multiperiodali. Con la composizione continua dei return, i return negli  $h$ -giorni tra  $T$  e  $T + h$  sono semplicemente la somma dei return giornalieri:

$$r_{T+h}(h) = \sum_{j=1}^h r_{T+j}$$

Assumendo che i return siano scorrelati, la varianza condizionata associata al return nel giorno  $h$  è:

$$\begin{aligned}var(r_{T+h}(h)|I_t) &= \sigma_T^2(h) = \sum_{j=1}^h var(r_{T+j}|I_t) \\ &= \mathbb{E}[\sigma_{T+1}^2|I_T] + \dots + \mathbb{E}[\sigma_{T+h}^2|I_T]\end{aligned}$$

Se i return hanno varianza costante  $\bar{\sigma}^2$ , allora

$$\sigma_T^2(h) = h\bar{\sigma}^2 \quad \text{e} \quad \sigma_T(h) = \sqrt{h}\bar{\sigma}$$

Questa regola - chiamata *square root time* - indica che la volatilità ad  $h$ -giorni di distanza si riscalda con la radice  $\sqrt{h}$ . In questo specifico caso determina il rapporto  $\frac{\sigma_T^2(h)}{h}$  rimanga costante.

Se i return sono descritti come un modello GARCH, allora la regola descritta potrebbe non essere corretta. Aggiungendo il modello di previsione GARCH(1,1) per  $\mathbb{E}[\sigma_{T+1}^2|I_T], \dots, \mathbb{E}[\sigma_{T+h}^2|I_T]$  nella  $\text{var}(r_{T+h}(h)|I_t)$  si ha il seguente esito:

$$\sigma_T^2(h) = h\bar{\sigma}^2 + (\mathbb{E}[\sigma_{T+1}^2] - \bar{\sigma}^2) \left[ \frac{1 - (a_1 + b_1)^h}{1 - (a_1 + b_1)} \right]$$

Infatti, per il GARCH(1,1) la regola *square root of time* è valida solo quando  $\mathbb{E}[\sigma_{T+1}^2] = \bar{\sigma}^2$ . Che  $\sigma_T^2(h)$  sia più grande o più piccolo di  $h\bar{\sigma}^2$  dipende dal fatto che  $\mathbb{E}[\sigma_{T+1}^2]$  sia maggiore o minore rispetto a  $\bar{\sigma}^2$ .

### 7.3.15 Previsione del VaR

#### Previsione del VaR non condizionato

Siano  $r_t$  i return giornalieri composti nel continuo su di un asseto o portafoglio e sia  $\alpha$  il livello di confidenza. Allora il value at Risk non condizionato ad un giorno di distanza,  $\text{VaR}_\alpha$ , è tipicamente definito come il valore negativo del quantile  $(1 - \alpha)$  della distribuzione non condizionata dei return giornalieri. In formule:

$$\text{VaR}_\alpha = -q_{1-\alpha}^R = -F_r^{-1}(1 - \alpha)$$

dove  $F_r$  è la funzione di ripartizione di  $r_t$ .

#### Esempio

Sia  $r_t \sim iid N(\mu, \sigma^2)$ . Allora:

$$q^\alpha = \mu + \sigma X q_{1-\alpha}^z, \quad q_{1-\alpha}^z = \phi^{-1}(1 - \alpha) \quad e \quad \widehat{\text{VaR}}_\alpha = -(\hat{\mu} + \hat{\sigma} x q_{1-\alpha}^z).$$

#### Previsione del VaR condizionato

Assumiamo ora che  $r_t$  segua un processo GARCH:

$$\begin{aligned} r_t &= \mu + \sigma_t z_t \\ \sigma_t^2 &\sim GARCH(p, q) \\ z_t &\sim N(0, 1) \end{aligned}$$

Allora il Value at Risk condizionato a 1 giorno,  $\text{VaR}_{\alpha,t}$  è

$$\text{VaR}_{\alpha,t} = -q_{1-\alpha,t}^r = -(\mu + \sigma_t q_{1-\alpha}^z)$$

Nota: il  $\text{VaR}_{\alpha,t}$  varia al variare del tempo poiché  $\sigma_t$  dipende dal tempo. Il Value at Risk non condizionato è costante nel tempo.

La previsione del VaR è:

$$\widehat{\text{VaR}}_{\alpha,t} = -(\hat{\mu} + \hat{\sigma}_t q_{1-\alpha}^z), \quad \hat{\sigma}_t = \text{previsione GARCH della volatilità}$$

Per un processo GARCH, il Value at Risk condizionato a h-giorni,  $\text{VaR}_{\alpha,t}$  è:

$$\text{VaR}_{\alpha,t} = -(h\mu + \sigma_T(h) q_{1-\alpha}^z)$$

dove  $\sigma_T^2(h) = \mathbb{E}[\sigma_{T+1}^2 | I_T] + \dots + \mathbb{E}[\sigma_{T+h}^2 | I_T]$  e  $\mathbb{E}[\sigma_{T+1}^2 | I_T], \dots, \mathbb{E}[\sigma_{T+h}^2 | I_T]$  sono gli h-step previsti della varianza condizionata.

La stima del Value at Risk è:

$$\widehat{\text{VaR}}_{\alpha,t}^h = -(h\hat{\mu} + \sigma_T(h) q_{1-\alpha}^z)$$

## Esempio

Consideriamo il titolo Apple quotato sul NYSE tra il gennaio 2005 ed il dicembre 2017 al fine di fittare un modello GARCH con lo scopo di determinare una stima sulla volatilità.



Figura 15: Grafico dei prezzi di scambio del titolo Apple, quotato sul NYSE, tra il gennaio 2005 ed il dicembre 2017.

Come prima cosa calcoliamo i log-return (*Figura 16*) e calcoliamo i punti della funzione di autocorrelazione dei return semplici, quadrati ed in modulo.

In *Figura 17* possiamo notare come l'autocorrelazione dei return quadrati e in modulo sia significativa solo per i return a distanza 1 (lag-1). Questa evidenza suggerisce l'uso di un modello GARCH di tipo (1,1). Eseguiamo dunque il fit in macchina attraverso l'uso di un software statistico (e.g. R, SAS, Matlab, etc.) e osserviamo gli esiti nelle pagine a seguire.

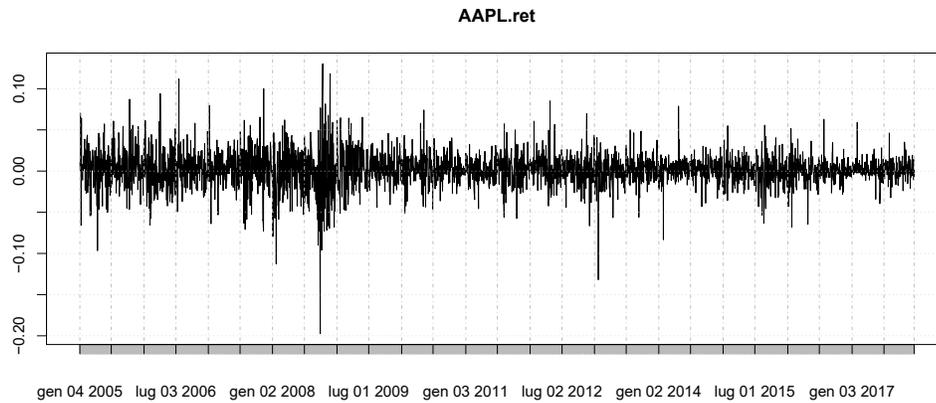


Figura 16: Grafico dei log-returns del titolo Apple tra il gennaio 2005 ed il dicembre 2017.

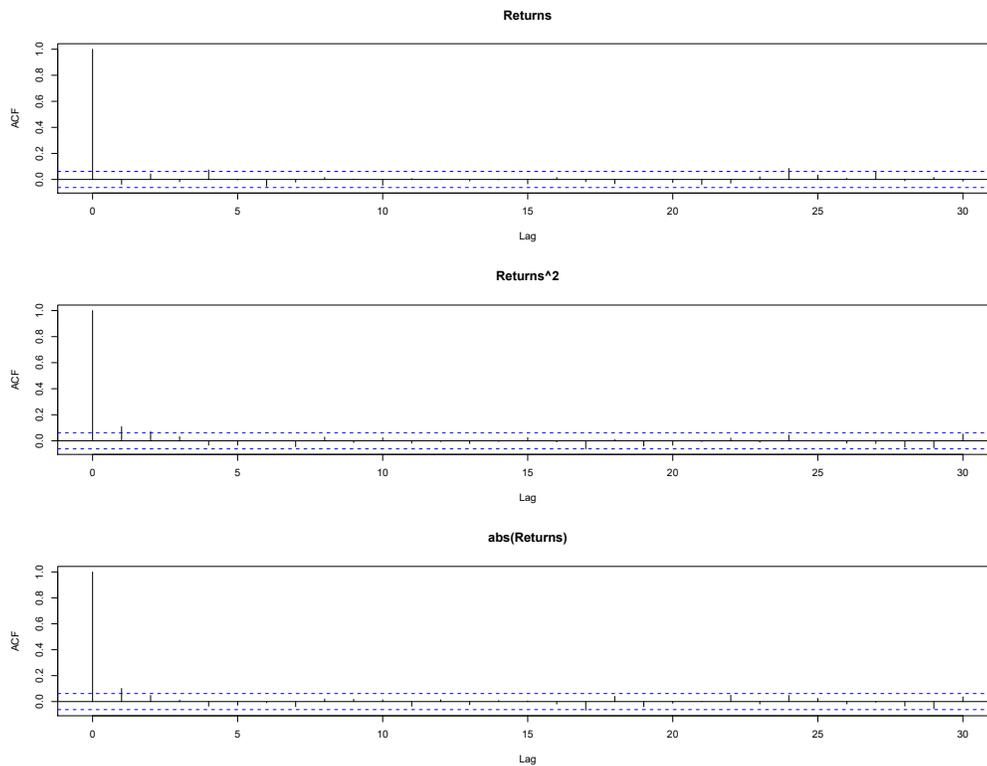


Figura 17: Grafico dell'Auto Correlation Function dei return semplici, al quadrato e in modulo del titolo Apple tra il gennaio 2005 ed il dicembre 2017.



Utilizzando i parametri derivanti dalla computazione del software, è possibile calcolare la volatilità condizionata per qualunque tempo (si veda *Figura 18*).

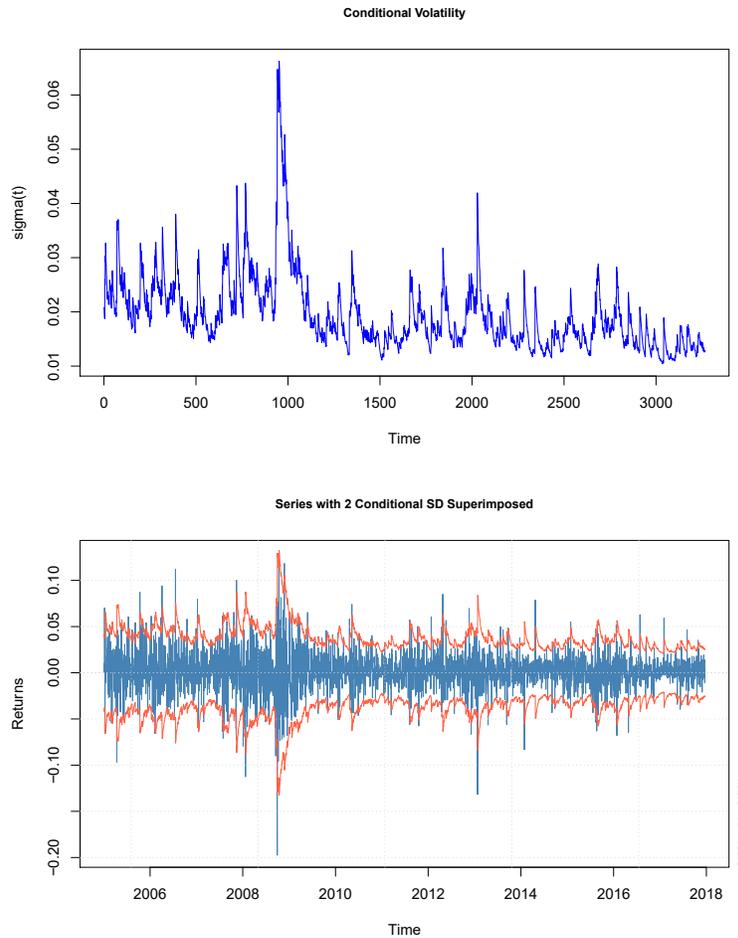


Figura 18: Grafico della volatilità condizionata del titolo Apple tra il gennaio 2005 ed il dicembre 2017 e sua rappresentazione doppia (quantile 95%) in sovrapposizione ai return.

Utilizzando i medesimi parametri stimati, è possibile calcolare anche la volatilità condizionata dei return futuri (*Figura 19*).

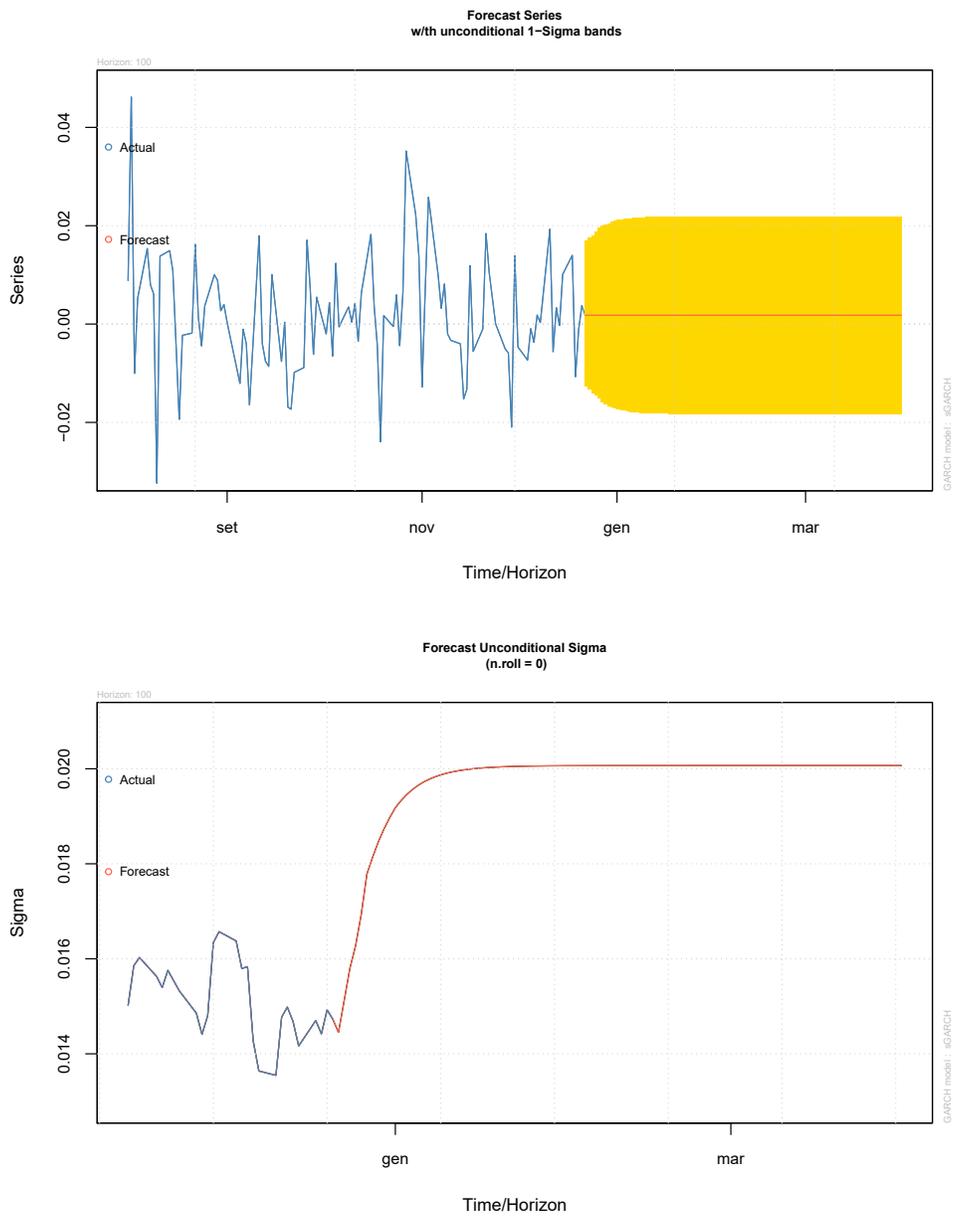


Figura 19: Grafico della stima della volatilità condizionata e dei returns dei 3 mesi successivi al periodo d'osservazione.

## 8 GARCH Univariato - Modelli Avanzati

I fenomeni di *Volatility clustering* e *leptocurtosi* (si veda *Capitolo 7.2.3*) sono tipici delle serie temporali di natura finanziaria. Un altro elemento ricorrente in cui ci si imbatte quando si analizzano i returns è il cosiddetto *leverage effect*, il quale si verifica quando le variazioni dei prezzi e/o i returns sono negativamente correlate alle variazioni della volatilità.

Osservazioni di questo tipo, sulle *time series* finanziarie, hanno condotto all'utilizzo di una vasta gamma di modelli a varianza condizionata per predire e stimare la volatilità.

Come visto nel *Paragrafo 7.3*, Engle propose nel 1982 un modello la cui varianza variava nel tempo e presentava eteroschedasticità condizionata autoregressiva, ovvero il modello ARCH. Evidenze empiriche basate sul suo lavoro mostrarono che gli ARCH di ordine elevato rappresentano bene il comportamento della varianza condizionata.

Il modello ARCH generalizzato (GARCH), introdotto da Bollerslev nel 1986 completò il lavoro di Engle poiché si basava su un modello ARCH infinito, ma riduceva il numero di parametri stimati da infinito a due. Sia il modello ARCH che il GARCH riescono a spiegare il volatility clustering e la leptocurtosi, ma poiché le loro distribuzioni sono simmetriche, falliscono nel catturare il leverage effect.

### 8.1 Leverage Effect

F. Black in 1976 fu la prima persona che prestò attenzione alla tematica del leverage effect. Notò che la volatilità cresce man mano che i prezzi diminuiscono, mentre tende a decrescere quando i prezzi aumentano. Propose la seguente spiegazione teorica a questo strano fenomeno.

La diminuzione percentuale del prezzo di un asset è identica alla diminuzione del valore di mercato dell'equity dell'impresa<sup>42</sup>. Il controvalore di mercato del debito emesso dall'azienda rimane invariato, ma il rapporto debito su equity varia positivamente. In questo modo l'impresa aumenta l'effetto leva e automaticamente diviene più esposta al rischio di default. Naturalmente

---

<sup>42</sup>Il *market value* dell'equity di una impresa si calcola moltiplicando il numero di quote di quell'impresa sul mercato per il loro prezzo di mercato

questo rischio viene trasferito direttamente agli azionisti <sup>43</sup> e solamente in parte ai creditori. L'equity della stessa azienda diviene più rischiosa e questo fattore determina l'innalzamento della volatilità dei prezzi di mercato.

Al contrario, quando i prezzi tendono ad alzarsi, il livello di indebitamento (ovvero la proporzione debito su equity) diminuisce. Risulta quindi una decrescita del rischio e di conseguenza una minor volatilità. Segue la relazione di causa ed effetto:

*la variazione dei prezzi o return affligge la volatilità futura (varianza)*

Indubbiamente, il vantaggio di utilizzare l'ipotesi dell'effetto leva è che essa spiega perfettamente la presenza dell'asimmetria della varianza nelle serie dei return finanziari. Nel modello GARCH che abbiamo trattato sino ad ora, la varianza condizionata dipende dai residui quadrati  $\epsilon_{t-i}^2$  e dunque il segno dei residui o di qualsiasi shock non ha alcun effetto sulla volatilità condizionata.

Come abbiamo visto, però, nella realtà questa ipotesi di simmetria tra le variazioni positive e negative viene infranta. Vedremo nei prossimi paragrafi come tener conto di questo importante fattore.

### 8.1.1 Testare gli effetti asimmetrici sulla volatilità condizionata

Una semplice metodologia per rivelare il possibile effetto leva asimmetrico è la correlazione campionaria tra  $r_t^2$  e  $r_{t-1}$ . Un valore negativo di questa correlazione dà evidenza di un potenziale effetto leva.

Un altro metodo deriva dal risultato della stima del seguente test di regressione:

$$\hat{\epsilon}_t^2 = \beta_0 + \beta_1 \hat{\omega}_{t-1} + \epsilon_t$$

dove  $\hat{\omega}_{t-1}$  è una variabile costruita a partire da  $\hat{\epsilon}_{t-1}$  e il suo relativo segno. Un significativo valore di  $\beta_1$  indica l'evidenza dell'effetto di asimmetria sulla volatilità condizionata. Sia  $S_{t-1}^-$  una variabile dummy equivalente ad 1 quando  $\hat{\epsilon}_{t-1}$  è negativa e 0 altrimenti. Engle e Ng considerano tre test per l'asimmetria. Imponendo:

- $\hat{\omega}_{t-1} = S_{t-1}^-$  dà il Sign Bias test;

---

<sup>43</sup>Gli azionisti vengono anche chiamati *residual claimants* in quanto sono gerarchicamente l'ultima categoria ad esercitare i diritti di credito. Per questo motivo sono i più soggetti al trasferimento dei rischi rispetto agli obbligazionisti.

- $\hat{\omega}_{t-1} = S_{t-1}^- \hat{\epsilon}_{t-1}$  dà l'effetto negativo del Sign Bias test;
- $\hat{\omega}_{t-1} = S_{t-1}^-$  dà l'effetto positivo del Sign Bias test;

## 8.2 Modelli GARCH asimmetrici

Pagan e Schwert<sup>44</sup> e Engle e Ng<sup>45</sup> hanno definito il nuovo concetto di *curva di impatto delle notizie*, la quale si riferisce al legame tra gli shock dei return passati sulla volatilità corrente. Questa curva ci fornisce indicazioni su come le nuove informazioni sono incorporate nella stima della volatilità. Nel modello GARCH, questa curva è una funzione quadratica centrata in  $\epsilon_{t-1} = 0$ . Per i modelli asimmetrici, la curva è disegnata in modo da crescere diversamente nelle due diverse direzioni.

### 8.2.1 Modello EGARCH

Il modello GARCH esponenziale (EGARCH) fu introdotto da Nelson nel 1991<sup>46</sup> per migliorare aspetti del modello standard:

1. i parametri  $\alpha$  e  $\beta$  sono calcolati in modo da garantire la positività della varianza del processo;
2. evidenze empiriche suggeriscono una risposta asimmetrica della volatilità agli shock dei prezzi.

Nel modello EGARCH, invece,  $\sigma_t^2$  dipende sia dalla ampiezza che dal segno della variazione dei return. Definiamo  $h_t = \ln(\sigma_t^2)$  e  $\epsilon_t = \sigma_t z_t$  dove gli  $z_t$  sono una serie di variabili aleatorie indipendenti e identicamente distribuite con media  $\mu = 0$  e varianza  $\sigma^2 = 1$  con  $t = 1, 2, \dots, n$ . Il modello si può dunque definire in questo modo:

$$h_t = a_0 + \sum_{i=1}^p a_i \frac{|\epsilon_{t-i}| + \gamma_i \epsilon_{t-i}}{\sigma_{t-i}} + \sum_{j=1}^q b_j h_{t-j}$$

La varianza risulta dunque sempre positiva in quanto  $\sigma_t^2 = \exp(h_t)$ . Inoltre l'effetto totale degli shock di prezzo positivi su  $\epsilon_{t-i}$  equivale a  $(1 + \gamma_i)|\epsilon_{t-i}|$ .

<sup>44</sup>A. Pagan e W. Schwert - *Alternative models for stock volatility*, Journal Of Econometrics (1990)

<sup>45</sup>R. Engle e V. Ng - *Measuring and testing the impact of news on volatility*, Journal of Finance (1993)

<sup>46</sup>Nelson, D. B. - *Conditional Heteroskedasticity in Asset Returns: A New Approach*, Econometrica. 59 (1991)

L'effetto di quelli negativi è, invece  $(1 + \gamma_i)|\epsilon_{t-i}|$ . L'effetto leva implica, invece, che  $\gamma_i < 0$ . L' EGARCH risulta stazionario in covarianza se:

$$b(1) = \text{persistenza} = \sum_{j=1}^q b_j < 1$$

**Osservazione per implementazioni software:** Nei package dei principali strumenti software di calcolo, quali R<sup>47</sup>, Matlab, Octave e Python il modello EGARCH viene specificato in modo leggermente differente:

$$h_t = a_0 + \sum_{i=1}^p (a_i z_{t-i} + \gamma_i (|z_{t-i} - \mathbb{E}[|z_{t-i}|]) + \sum_{j=1}^q b_j h_{t-j}$$

In questo caso  $a_i$  cattura l'effetto del segno: l'effetto leva impone che  $a_i < 0$ . Per quanto riguarda  $\gamma_i$ , all'aumentare del suo valore aumenta l'effetto leva. Naturalmente questo implica che  $\gamma_i > 0$ .

### 8.2.2 TGARCH e GJR

Il modello *threshold* GARCH (TGARCH) proposto da Zakoian<sup>48</sup> nel 1991 e il modello GJR<sup>49</sup> elaborato da Glosten, Jagannathan e Runkle nel 1993, definiscono la varianza condizionata come una funzione lineare a tratti:

$$\sigma_t^2 = a_0 + \sum_{i=1}^p a_i \epsilon_{t-i}^2 + \sum_{i=1}^p \gamma_i S_{t-i} \epsilon_{t-i}^2 + \sum_{j=1}^q b_j \sigma_{t-j}^2$$

$$S_{t-i} = \begin{cases} 1 & \text{se } \epsilon_{t-i} < 0 \\ 0 & \text{se } \epsilon_{t-i} \geq 0 \end{cases}$$

Quando  $\epsilon_{t-i}$  è positivo, l'effetto totale è  $a_i \epsilon_{t-i}^2$ , mentre quando negativo è  $(a_i + \gamma_i) \epsilon_{t-i}^2$ . Inoltre l'effetto leva implica che  $\gamma > 0$ .

<sup>47</sup>Per la documentazione ufficiale del package sull'*univariate GARCH* utilizzata nella redazione di questo elaborato si faccia riferimento a Alexios Ghalanos - *rugarch package*, <https://cran.r-project.org/web/packages/rugarch/rugarch.pdf>

<sup>48</sup>Jean-Michele Zakoian - *Threshold heteroskedastic models*, Journal of Economic Dynamics and Control (1991)

<sup>49</sup>Lawrence R. Glosten, Ravi Jagannathan, David E. Runkle - *On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks*, Journal of Finance (1993)

In questo modello la stazionarietà si ha quando:

$$\sum_{i=1}^p (a_i + \frac{\gamma}{2}) + \sum_{j=1}^q \beta_j < 1$$

### 8.2.3 PGARCH

Il modello studiato da Ding, Granger e Engle è definito semplicemente elevando a potenza il modello GARCH standard. Quindi per  $d > 0$  possiamo esplicitarlo come segue:

$$\sigma_t^d = a_0 + \sum_{i=1}^p a_i (|\epsilon_{t-i}| - \gamma_i \epsilon_{t-i})^d + \sum_{j=1}^q b_j \sigma_{t-j}^d$$

L'effetto leva, come negli altri modelli, implica che  $\gamma_i > 0$ . Se  $\epsilon_{t-i} < 0$  e  $\gamma_i > 0$  allora  $-\gamma_i \epsilon_{t-i} > 0$  e in questo modo i return negativi aumentano con  $\sigma_t^d$ . Al variare di  $d$  abbiamo un comportamento diverso del modello:

- se  $d = 2$  abbiamo un GARCH standard con l'aggiunta dell'effetto leva;
- se  $d = 1$  si ha un modello per  $\sigma_t$  ed è più robusto agli outlier rispetto al modello con  $d = 2$ ;
- $d$  può essere fissato oppure stimato attraverso la massima verosimiglianza.

Le condizioni di stazionarietà in covarianza sono piuttosto complesse e dipendono da  $b_j$  e  $\gamma_j$ <sup>50</sup>.

### 8.2.4 News Impact Curve

Engle e Ng proposero l'utilizzo di una funzione, chiamata *news impact curve* (NIC) per valutare l'asimmetria di un modello GARCH.

La NIC misura quanta informazione viene incorporata nella volatilità. Rappresenta, infatti, la relazione tra la varianza condizionata al tempo  $t$ , il

---

<sup>50</sup>Per la documentazione ufficiale del package sull'*univariate GARCH* utilizzata nella redazione di questo elaborato si faccia riferimento a Alexios Ghalanos - *rugarch package*, <https://cran.r-project.org/web/packages/rugarch/rugarch.pdf>

termine di errore al  $t - 1$ , mantenendo costante il set di informazioni dato al tempo  $t - 2$ . Questa curva si può facilmente ricavare per tutti i tipi di GARCH.

### Esempio

Nel modello GARCH(1,1) abbiamo:

$$NIC(\epsilon_t | h_t = h) = \alpha_0 + \beta_1 h + \alpha_1 \epsilon_t^2 = A + \alpha_1 \epsilon_t^2$$

che è una funzione quadratica di  $\epsilon_t^2$  e quindi simmetrica attorno a 0 (con intercetta  $A = \alpha_0 + \beta_1 h$ ). Come abbiamo visto in precedenza, però, questa ipotesi di simmetria sembrerebbe essere violata e ciò deriva proprio dal fatto che la curva NIC empirica risulta asimmetrica. Per i modelli EGARCH, TGARCH e PGARCH la curva è, invece, asimmetrica attorno al valore zero.

### 8.2.5 Forecast per i modelli asimmetrici

In questa sezione vediamo un esempio in cui si fanno delle predizioni a partire da un modello GARCH asimmetrico. Consideriamo un modello TGARCH(1,1) al tempo  $T$ :

$$\begin{aligned} \epsilon_T &= z_t \sigma_t \quad \text{dove } z_t \sim N(0, 1) \\ \sigma_T^2 &= a_0 + a_1 \epsilon_{T-1}^2 + b_1 \sigma_{T-1}^2 \\ S_T &= \begin{cases} 1 & \text{se } \epsilon_T < 0 \\ 0 & \text{se } \epsilon_T \geq 0 \end{cases} \end{aligned}$$

Assumiamo che  $\epsilon_t$  abbia una distribuzione simmetrica attorno a 0. La predizione della volatilità in  $T + 1$  basata sul set di informazioni al tempo  $T$  è:

$$\mathbb{E}[\sigma_{T+1}^2 | I_T] = a_0 + a_1 \epsilon_T^2 + \gamma_1 S_T \epsilon_T^2 + b_1 \sigma_T^2,$$

dove si assume che si conoscono i valori  $\epsilon_T^2$ ,  $S_T$  e  $\sigma_T^2$ . Quindi, la predizione del TGARCH(1,1) sarà differente da quella del GARCH(1,1)<sup>51</sup> se  $S_T = 1$  ( $\epsilon_T < 0$ ).

La predizione al tempo  $T + 2$  è:

$$\begin{aligned} \mathbb{E}[\sigma_{T+2}^2 | I_T] &= a_0 + a_1 \mathbb{E}[\epsilon_{T+1}^2 | I_T] + \gamma_1 \mathbb{E}[S_{T+1} \epsilon_{T+1}^2 | I_T] + b_1 \mathbb{E}[\sigma_{T+1}^2 | I_T] \\ &= a_0 + \left( \frac{\gamma_1}{2} + a_1 + b_1 \right) \mathbb{E}[\sigma_{T+1}^2 | I_T] \end{aligned}$$

<sup>51</sup>per ulteriori approfondimenti si veda il *Paragrafo 7.3.12*

Se  $S_{T+1}$  è indipendente da  $\epsilon_{T+1}^2$  e  $z_t$  ha una distribuzione simmetrica attorno a 0 :

$$\mathbb{E}[S_{T+1}\epsilon_{T+1}^2|I_T] = \mathbb{E}[S_{T+1}|I_T]\mathbb{E}[\epsilon_{T+1}^2|I_T] = \frac{1}{2}\mathbb{E}[\sigma_{T+1}^2|I_T]$$

Si può notare che l'impatto dell'effetto leva è presente anche quando  $S_T = 0$ .

Sostituendo in modo ricorsivo possiamo ottenere la predizione al tempo  $T + h$ :

$$\mathbb{E}[\sigma_{T+h}^2|I_T] = a_0 + \left(\frac{\gamma_1}{2} + a_1 + b_1\right)^{h-1} \mathbb{E}[\sigma_{T+1}^2|I_T] ,$$

la quale è molto simile a quella per il GARCH(1, 1).

La formula della media invertita è:

$$\mathbb{E}[\sigma_{T+h}^2|I_T] - \bar{\sigma}^2 = \left(\frac{\gamma_1}{2} + a_1 + b_1\right)^{h-1} \left(\mathbb{E}[\sigma_{T+1}^2|I_T] - \bar{\sigma}^2\right) ,$$

dove  $\bar{\sigma}^2 = \frac{a_0}{1 - \frac{\gamma_1}{2} - a_1 - b_1}$  è detta *long run variance*, ovvero la varianza che si ottiene

### 8.3 GARCH - errori non normali

Sino ad ora abbiamo trattato il modello GARCH con errori normali,  $\epsilon_t = \sigma_t z_t$  e  $z_t \sim iid N(0, 1)$ . Spesso la stima degli errori standardizzata  $\hat{z}_t = \frac{\hat{\epsilon}_t}{\hat{\sigma}}$  di un modello GARCH con errori standard presenta code delle code pesanti e/o asimmetriche. Questo suggerisce l'utilizzo una distribuzione standardizzata con code pesanti e/o asimmetriche per  $z_t$ , invece della classica normale  $N(0, 1)$ .

La più comune distribuzione di questo tipo è la *skewed-t*.

**Osservazione per implementazioni software:** nel package di R (*rugarch*), utilizzato nell'elaborazione di questa tesi, la distribuzione asimmetrica con code pesanti implementata è la *generalized hyperbolic skew Student distribution*.

#### 8.3.1 GARCH - errori t-Student

Vediamo il primo e più tipico di casi in cui si propone una distribuzione alternativa degli errori distribuiti normalmente. Sia  $\mu_t$  una variabile aleatoria

t-Student con  $\nu$  gradi di libertà e parametro di scala  $s_t$ . Allora:

$$f(\mu_t) = \frac{\Gamma[(\nu + 1)/2] s_t^{-1/2}}{(\pi\nu)^{1/2}\Gamma(\nu/2)[1 + \mu_t^2/(s_t\nu)]^{(\nu+1)/2}}$$

$$\text{var}(\mu_t) = \frac{s_t\nu}{\nu - 2}, \quad \nu > 2.$$

Se  $\epsilon_t$  nel modello GARCH è una t-Student con  $\mathbb{E}[\epsilon_t^2 | I_{t-1}] = \sigma_t^2$  allora imponendo

$$s_t = \frac{\sigma_t^2(\nu - 2)}{\nu}$$

creiamo la distribuzione t-Student standardizzata per  $z_t$ .

### 8.3.2 Distribuzione generale degli errori

Nelson suggerì di utilizzare una distribuzione generale degli errori ( *generalized error distribution* - GED ) con il parametro  $\nu > 0$ . la funzione di densità della variabile aleatoria  $\mu_t$ , distribuita come una GED con parametro  $\nu$  è:

$$f(\mu_t) = \frac{\nu \exp[-(1/2)|\mu_t/\alpha|^\nu]}{\alpha \cdot 2^{(\nu+1)/\nu}\Gamma(1/\nu)}$$

dove

$$\alpha = \left[ \frac{2^{-2/\nu}\Gamma(1/\nu)}{\Gamma(3/\nu)} \right]$$

Con  $\nu = 2$  abbiamo una distribuzione normale. Se  $0 < \nu < 2$  si ha una distribuzione con le code pesanti, mentre quando  $\nu > 2$  le code saranno più sottili rispetto a quelle della normale.

Il caso particolare  $\nu = 1$

$$f(\mu_t) = \frac{1}{\sqrt{2}} e^{-\text{sqrt}2|\mu_t|}$$

Per introdurre l'asimmetria è necessario utilizzare la distribuzione t-Student asimmetrica. Ci sono molte definizioni di questa distribuzione (e.g. Azzalini e Capitanio, Fernandez e Steel, etc..). Tutte le versioni hanno:

- il parametro  $\nu > 0$  indicativo i gradi di libertà, che controlla lo spessore delle code in relazione alla normale.
- il parametro  $\xi$  di asimmetria. Se  $\xi < 0$  si ha un'asimmetria negativa (ovvero una coda lunga a sinistra), mentre per  $\xi > 0$  si esattamente la situazione opposta.

## 8.4 Modelli GARCH a lunga memoria

Se i return seguono un modello GARCH( $p, q$ ), allora l'autocorrelazione dei return quadrati o in valore assoluto dovrebbe decadere esponenzialmente. La funzione di autocorrelazione SACF di  $r_t^2$  e  $|r_t|$  spesso, però, decade molto più lentamente. Questa evidenza viene chiamata *long memory* (lunga memoria).

In modo più formale, un processo stazionario possiede la proprietà di lunga memoria se la sua funzione di autocorrelazione si comporta nel seguente modo:

$$\rho(k) \rightarrow C_p k^{2d-1} \text{ per } k \rightarrow \infty,$$

dove  $C_p$  è una costante positiva, e  $d \in \mathbb{R}$  tale che  $0 < d < \frac{1}{2}$ .

Quindi la funzione di autocorrelazione di un processo a lunga memoria decade lentamente, ovvero con velocità iperbolica.

La proprietà di lunga memoria del modello può essere costruita nell'equazione della varianza condizionata in una varietà di modi. Citiamo due esempi di modelli con questa caratteristica:

- modello di Bollerslev FIGARCH <sup>52</sup>
- modello a due componenti di Engle <sup>53</sup>

La stima dei modelli GARCH a lunga memoria e raramente vengono utilizzati nelle applicazioni pratiche.

## 8.5 Valutare le predizioni della volatilità

I modelli GARCH vengono valutati e confrontati tra di loro per l'abilità nel fare predizioni *out of sample*. Questa capacità può essere misurata:

1. metriche tradizionali per calcolare l'errore della predizione come MSE
2. specifiche considerazioni economiche, quali ad esempio, le violazioni del Value-at-Risk, l'accuratezza del pricing di opzioni o le performance di portafogli di investimento

---

<sup>52</sup>Richard T.Bailliea, Tim Bollerslevb, Hans Ole Mikkelsen - *Fractionally integrated generalized autoregressive conditional heteroskedasticity*, Journal of Econometrics (1996)

<sup>53</sup>Robert Engle - *GARCH 101: The Use of ARCH/GARCH Models in Applied Econometrics*, Journal of Economic Perspective (2001)

Le predizioni *out of sample* da utilizzare nei modelli per fare delle comparazioni sono tipicamente calcolate attraverso uno di questi due metodi:

- *Recursive Forecasts*: si parte da un campione iniziale di dati da  $t = 1, \dots, T$  utilizzati per stimare il modello. Gli  $h$  passi successivi vengono predetti partendo dal tempo  $T$ . Allo step successivo il campione cresce di un elemento ed il modello viene nuovamente stimato; gli  $h$  passi successivi sono prodotti partendo ora dal tempo  $T + 1$ .
- *Rolling Forecasts*: si parte da un campione iniziale di dati da  $t = 1, \dots, T$  utilizzati per individuare una finestra temporale di ampiezza  $T$ , per stimare il modello. Gli  $h$  passi successivi vengono predetti partendo dal tempo  $T$ . Allo step successivo la finestra temporale si muove avanti di un passo ed il modello viene nuovamente stimato usando i dati da  $t = 2, \dots, T + 1$ ; gli  $h$  passi successivi sono prodotti partendo ora dal tempo  $T + 1$ .

### 8.5.1 Statistiche tradizionali - valutazione della predizione

Sia  $\mathbb{E}_i[\sigma_{T+h}^2|I_T]$  la predizione ampia  $h$  passi di  $\sigma_{T+h}^2$  al tempo  $T$  derivante da un modello GARCH  $i$  utilizzando sia il metodo *recursive* che *rolling*.

Definiamo il corrispondente errore di predizione come

$$e_{i,T+h|T} = \mathbb{E}_i[\sigma_{T+h}^2|I_T] - \sigma_{T+h}^2.$$

Le statistiche che più comunemente vengono utilizzate per valutare la bontà della predizione sono:

$$\begin{aligned} MSE_i &= \frac{1}{N} \sum_{j=T+1}^{T+N} e_{i,j+h|j}^2, \\ MAE_i &= \frac{1}{N} \sum_{j=T+1}^{T+N} |e_{i,j+h|j}|, \\ MAPE_i &= \frac{1}{N} \sum_{j=T+1}^{T+N} \frac{|e_{i,j+h|j}|}{\sigma_{j+h}}. \end{aligned}$$

Il modello che produce il più piccolo valore di queste statistiche viene giudicato come il migliore. Certamente le statistiche di valutazione della predizione sono variabili aleatorie e si dovrebbe utilizzare una procedura statistica formale per determinare se un modello ha o meno una superiore performance predittiva.

### 8.5.2 Test di Diebold-Mariano per l'accuratezza della predizione

Siano  $\{e_{1,j+h|j}\}_{T+1}^{T+N}$  e  $\{e_{2,j+h|j}\}_{T+1}^{T+N}$  gli errori della predizione di due modelli GARCH differenti. L'accuratezza per ognuna delle predizioni è misurata in base ad una particolare funzione di perdita  $L(e_{i,j+h|T})$ ,  $i = 1, 2..$ . Scegliamo per esempio una funzione di penalità quadratica:

$$L(e_{1,j+h|T}) = (e_{1,j+h|T})^2$$

Il test di Diebold-Mariano<sup>54</sup> (DM) è basato sulla differenza di queste due funzioni di perdita:

$$d_{T+h} = L(e_{1,j+h|T}) - L(e_{2,j+h|T}).$$

L'ipotesi nulla di accuratezza equivalente è  $H_0 : \mathbb{E}[d_{T+h}] = 0$ .

La statistica DM è:

$$S = \frac{\bar{d}}{(\hat{avar}(\bar{d}))^{1/2}}, \quad \bar{d} = N^{-1} \sum_{j=T+1}^{T+N} d_{j+h}$$

Diebold-Mariano indicano di utilizzare il test di Newey-West<sup>55</sup> per  $\hat{avar}(\bar{d})$  poiché il campione della differenza delle funzioni di perdita  $\{d_{j+h}\}_{T+1}^{T+N}$  sono correlati per  $h > 1$ .

Sotto l'ipotesi nulla di accuratezza predittiva vale

$$S \sim N(0, 1)$$

Dunque la statistica DM può essere utilizzata per testare se una data statistica di valutazione della predizione (e.g.  $MSE_1$ ) per un modello è statisticamente differente da quella di un altro modello preso in considerazione (e.g.  $MSE_2$ ).

### 8.5.3 Regressione della predizione - Mincer-Zarnowitz

Le predizioni sono spesso giudicate utilizzando le regressioni:

$$\sigma_{T+h}^2 = \alpha + \beta \mathbb{E}[\sigma_{T+h}^2 | I_T] + e_{i,T+h}$$

<sup>54</sup>R. Mariano e Diebold - *Testing forecast accuracy*, Blackwell Publishing (2002)

<sup>55</sup>Newey, Whitney K; West, Kenneth D - *A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix*, Econometrica (1987)

La predizione non condizionata avrebbe  $\alpha = 0$  e  $\beta = 1$ .

Per predizioni molto accurate si hanno alti valori di  $R^2$ . Le regressioni della predizione subisce i problemi derivanti dagli errori di variabile, nel momento in cui i parametri stimati con il GARCH sono usati per calcolare  $\mathbb{E}_i[\sigma_{T+h}|I_T]$ . Ciò crea un minor condizionamento nella stima di  $\beta$  e come risultato si ha che l'indicatore sul quale viene focalizzata maggiormente l'attenzione è l' $R^2$

#### 8.5.4 I principali problemi nel valutare le predizioni

Uno dei problemi principali nelle valutazioni e confronti delle predizioni sui modelli di volatilità è che la volatilità ad h-passi più avanti nel tempo  $\sigma_{T+h}^2$  non è direttamente osservabile.

Tipicamente,  $\epsilon_{T+h}^2$  viene utilizzato al posto di  $\sigma_{T+h}^2$  poiché:

$$\mathbb{E}[\epsilon_{T+h}^2|I_T] = \mathbb{E}[z_{T+h}^2 \sigma_{T+h}^2|I_T] = \mathbb{E}[\sigma_{T+h}^2|I_T]$$

$\epsilon_{T+h}^2$  è però molto più *rumoroso* rispetto a  $\sigma_{T+h}^2$ :

$$\text{var}(\epsilon_{T+h}^2) = \mathbb{E}[\epsilon_{T+h}^4](k-1)$$

dove  $k$  è il momento quarto di  $z_t$  e questo causa problemi per le interpretazioni sulla metrica di valutazione delle predizioni.

In molti paper accademici è stata valutata l'accuratezza delle predizioni dei modelli GARCH utilizzando  $\epsilon_{T+h}^2$  in sostituzione di  $\sigma_{T+h}^2$ <sup>56</sup>.

In linea di massima le evidenze sulle valutazioni del forecast sono:

1. le valutazioni sulle statistiche tendono ad essere grandi;
2. le regressioni tendono ad esser poco condizionate;
3. la regressione  $R^2$  tende ad essere molto bassa (tipicamente sotto 0,1).

In generale, i modelli GARCH asimmetrici tendono ad avere le peggiori statistiche di valutazione, ma più in generale i modelli GARCH non fanno predizioni molto accurate.

---

<sup>56</sup>Poon nel 2005 elaborò un paper onnicomprensivo sul tema in questione

Andersen e Bollerslev<sup>57</sup> diedero una spiegazione a questo fenomeno di scarsa performance nelle predizioni del GARCH quando  $\epsilon_{T+h}^2$  viene utilizzato al posto di  $\sigma_{T+h}^2$ . Per il GARCH(1, 1) nei quali  $z_t$  ha una curtosi finita  $k$ , mostrarono che l' $R^2$  della regressione del forecast con  $h = 1$  è equivalente a:

$$R^2 = \frac{a_1^2}{1 - b_1^2 - 2a_1b_1},$$

ed è limitato superiormente da  $1/k$ . Assumendo  $z_t \sim N(0, 1)$ , si ha un *upper bound* di  $1/3$ . Con distribuzioni a code pesanti gli upper bound diventano più piccoli. Dunque ci aspettiamo valori molto bassi di  $R^2$  anche se il modello reale è davvero il GARCH(1, 1).

---

<sup>57</sup>T. Andersen e T. Bollerslev - *Answering the Skeptics: Yes, Standard Volatility Models do Provide Accurate Forecast*, International Economic (1998)

## Esempio

Mostriamo un esempio in cui si cerca di utilizzare gli elementi costitutivi del *Capitolo 8*. Come al solito il punto di partenza per i nostri esempi è una serie storica finanziaria. In questo caso scegliamo il titolo Intesa Sanpaolo quotato sul listino del mercato italiano FTSEMIB. Riportiamo di seguito il Grafico.

### Intesa Sanpaolo



Figura 20: Grafico dell'andamento azionario del titolo Intesa Sanpaolo quotato sull'indice di borsa FTSEMIB.

Partiamo dalla definizione di un classico modello GARCH(1,1) come visto nel *Paragrafo 7.3.12*:

$$\begin{aligned}r_t &= \mathbb{E}[r_t|I_{t-1}] + \epsilon_t = \mu + \epsilon_t \\ \epsilon_t &= z_t \sigma_t, \quad z_t \sim iid N(0, 1) \\ \sigma_t^2 &= a_0 + a_1 \epsilon_{t-1}^2 + b_1 \sigma_{t-1}^2\end{aligned}$$

Dopo aver definito il modello lo proviamo a fittare sulla nostra serie storica.

Otteniamo così il modello con i seguenti parametri ottimali:

$$r_t = \mathbb{E}[r_t|I_{t-1}] + \epsilon_t = 0.2573 + \epsilon_t$$

$$\epsilon_t = z_t \sigma_t, \quad z_t \sim iid N(0, 1)$$

$$\sigma_t^2 = 0.000012 + 0.161781 \cdot \epsilon_{t-1}^2 + 0.837219 \cdot \sigma_{t-1}^2$$

Vediamo ora se sul modello è più opportuno utilizzare lag più elevati di 1. Da quanto rilevato dal test di Ljung-Box, però, la serie storica non sembra evidenziare alcuna correlazione seriale: A questo punto verifichiamo se

```

Weighted Ljung-Box Test on Standardized Residuals
-----
                                statistic p-value
Lag[1]                          0.3087  0.5785
Lag[2*(p+q)+(p+q)-1][5]        1.0141  1.0000
Lag[4*(p+q)+(p+q)-1][9]        1.8089  0.9912
d.o.f=2
H0 : No serial correlation

```

Figura 21: Output del test di Ljung-Box

è corretto aver considerato un modello simmetrico per la serie in questione. Nel *Paragrafo 8.1.1* abbiamo identificato un metodo, che attraverso una semplice correlazione testa gli effetti asimmetrici. In questo specifico caso in

```

Sign Bias Test
-----
                                t-value   prob sig
Sign Bias                       0.2924  0.77001
Negative Sign Bias               2.3836  0.01721 **
Positive Sign Bias               0.1107  0.91186
Joint Effect                     8.5658  0.03566 **

```

Figura 22: Esito del test di Engle e Ng per testare l'asimmetria come definito nel *Paragrafo 8.1.1*

due casi si rifiuta l'ipotesi nulla ed è dunque possibile utilizzare un modello asimmetrico.

Prendiamo in considerazione tre dei modelli asimmetrici definiti in questo capitolo. Utilizzeremo il modelli EGARCH, GJR GARCH e APARCH. A titolo di esempio esaminiamo sempre i modelli con i parametri  $p = q = 1$ .

Partiamo dalla definizione del modello asimmetrico EGARCH definito nel

Paragrafo 8.2.1 e fittiamo il modello per trovare anche in questo caso i parametri ottimali.

$$r_t = \mathbb{E}[r_t | I_{t-1}] + \epsilon_t = 0.2627 + \epsilon_t$$

$$\epsilon_t = z_t \sigma_t, \quad z_t \sim iid N(0, 1)$$

$$h_t = a_o + 0.2275 \frac{|\epsilon_{t-1}| + 0.1825 \epsilon_{t-1}}{\sigma_{t-1}} + 0.2997 h_{t-1}$$

$$\sigma_t^2 = \exp(h_t)$$

Definiamo ora gli altri due modelli in modo analogo a quanto fatto per l'EGARCH. Quando troviamo i parametri degli tutti modelli siamo pronti

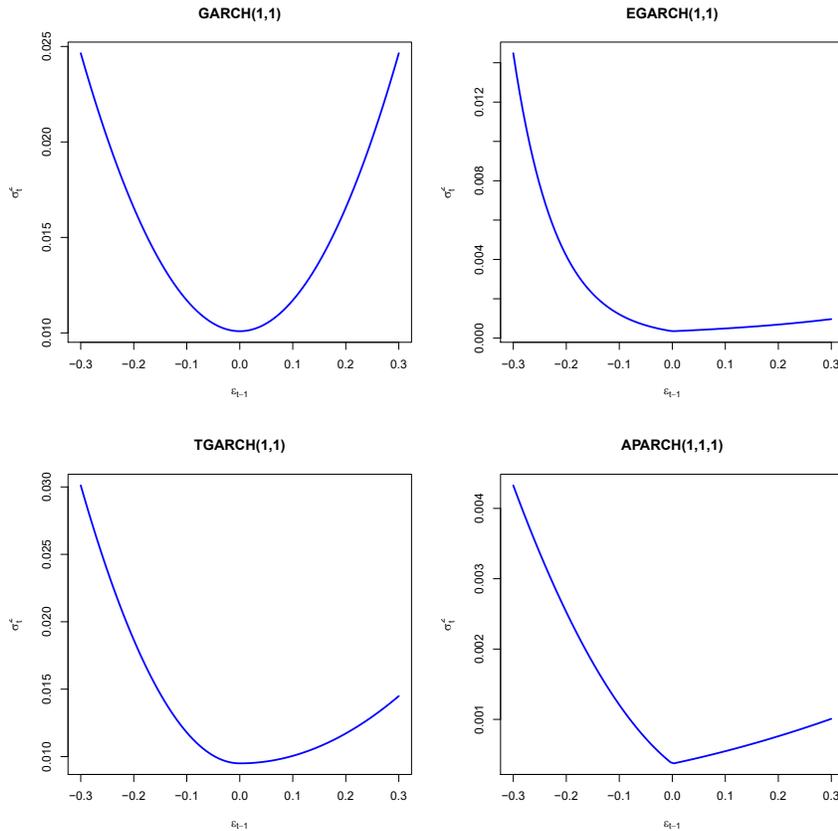


Figura 23: Grafico delle 4 *News Impact Curve* relative i modelli GARCH specificati nell'esempio

a fare qualche confronto. Nel *Paragrafo 8.2.4* abbiamo parlato della *News Impact Curve*: vediamo in *Figura 24* come cambia a seconda dei vari modelli.

E' importante ricordare che utilizzando i software di calcolo non sempre i modelli già implementati sono definiti con la notazione o con le caratteristiche cui facciamo riferimento negli elaborati è necessario dunque dare la specificazione del modello per ottenere risultati coerenti con quanto ci si aspetta.

Calcoliamo infine, alcuni degli information criteria di cui abbiamo parlato nel *Paragrafo 7.3.11* in modo da fare una valutazione.

	garch11	egarch11	gjr_garch11	aparch11.1
Akaike	-4.633	-5.217	-4.662	-5.222
Bayes	-4.621	-5.207	-4.648	-5.212
Shibata	-4.633	-5.217	-4.662	-5.222
Hannan-Quinn	-4.629	-5.213	-4.657	-5.218

Figura 24: Output del calcolo di vari Information criteria per la valutazione dei modelli

Notiamo che in questo caso il modello APARCH(1,1) sembra il migliore.

Esaminiamo gli errori standard, partendo dal caso di distribuzione normale e considerando in seguito gli elementi trattati nella *Sezione 8.3*. Utilizzando il modello GARCH(1,1) con distribuzione condizionata normale, andiamo ad analizzare il qqplot per confrontare la distribuzione degli errori empirica con quella reale.

Nella *Figura 25* si nota il classico fenomeno di code pesanti già analizzato in precedenza. Proviamo ad utilizzare una distribuzione non normale e verifichiamo se questo fenomeno diminuisce. Prendiamo in considerazione una distribuzione T - Student: in *Figura 26* possiamo infatti osservare che gli errori sembrano fittare meglio con questo tipo di distribuzione.

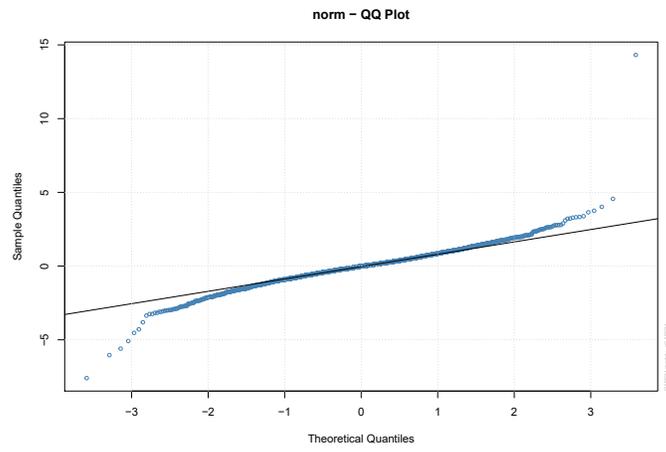


Figura 25: Output di R relativo ai parametri stimati nel modello definito in *Figura 21*

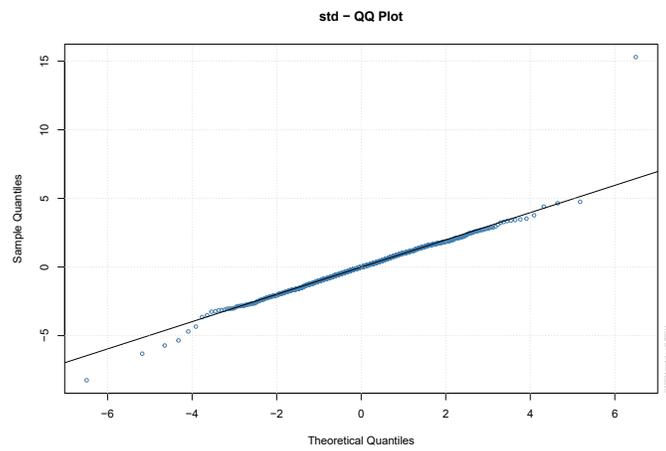


Figura 26: Output di R relativo ai parametri stimati nel modello definito in *Figura 21*

Anche il fenomeno di leptocurtosi sembra essere descritto meglio dalla distribuzione non normale. Vediamo in *Figura 27* come la linea gialla (densità della distribuzione T-Student) approssimi meglio la densità empirica degli errori standard rispetto a quella normale (linea blu).

Infine considerando distribuzioni condizionata di tipo T-Student, ridefi-

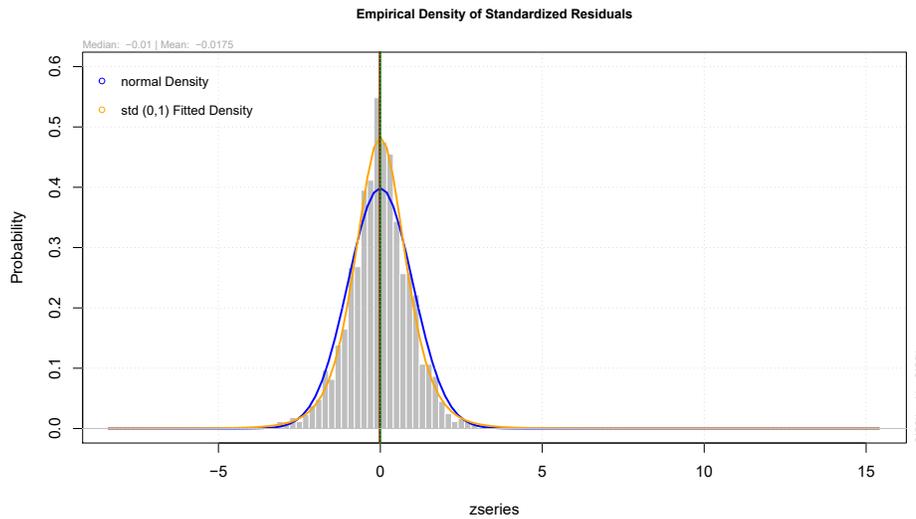


Figura 27: Grafico della densità empirica degli errori standard con le due distribuzioni reali sovrainpresse

niamo i modelli GARCH visti prima e proviamo ad effettuare il forecast della volatilità a seconda della diversa specificazione del modello. Come mostrato in *Figura 28* i diversi modelli conducono a predizioni differenti e sono necessari dei metodi per valutarle.

Nella *Sezione 8.5.1* abbiamo definito tre statistiche utili all'individuazione del miglior modello ad effettuare la predizione. Tra tutti i modelli valutiamo solamente il GARCH e APARCH e confrontiamoli: si veda la *Figura 29* dove il modello migliore in tutte e tre le statistiche risulta il modello APARCH con errori distribuiti come una T-Student.

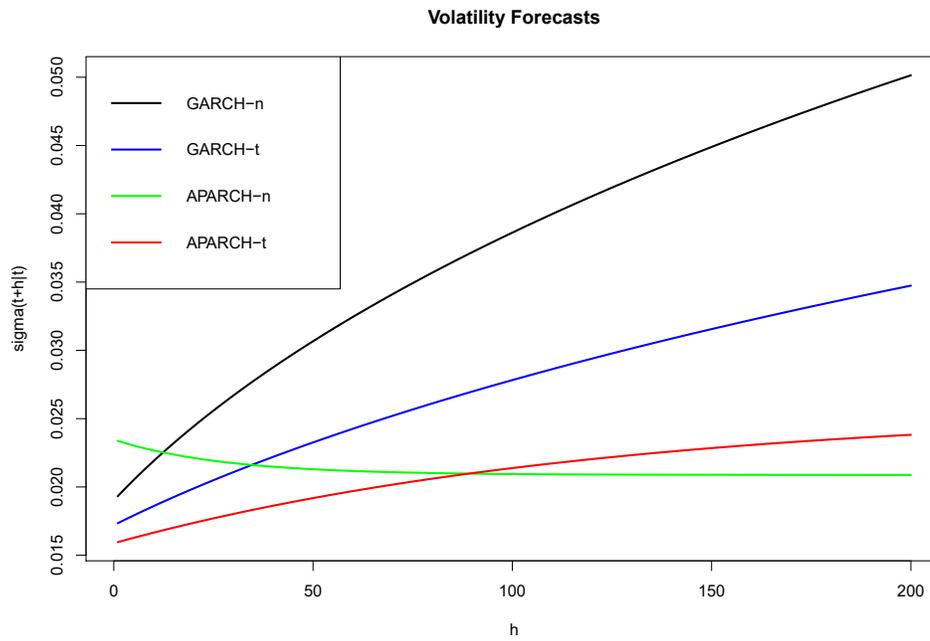


Figura 28: Grafico delle predizioni della volatilità per i vari modelli rappresentati

	<code>garch11</code>	<code>garch11.t</code>	<code>aparch11.1</code>	<code>aparch11.t.1</code>
MSE	0.000148041549465759	0.000117741401225885	0.000117991158430702	0.000116154169003905
MAE	0.00823754278010237	0.00805208447713568	0.00783087056439369	0.00762820692398096
DAC	0.52	0.5	0.49	0.48

Figura 29: Output del software R riguardante le statistiche per valutare le predizioni della volatilità

## 9 Implementazione software

La trattazione teorica del risk management, nonché quella dei principali modelli quantitativi per l'analisi delle serie temporali, è stata propedeutica alla realizzazione di una piattaforma web-based. L'obiettivo è quello di rendere fruibili agli investitori di profilo non specialistico un set di strumenti, metodologie e segnali nella forma più intuitiva possibile.

Le modalità di sviluppo progettuale sono fondate su tre cardini principali:

1. **Dinamicità e modularità:** adozione di soluzioni dinamiche e modulari in grado di permettere l'integrazione di nuove componenti, in modo semplice, veloce e governato. E' così possibile garantire una tempestiva continua evoluzione della piattaforma rispetto alla frontiera della ricerca e dell'innovazione tecnologica.
2. **Real time:** l'applicazione è fortemente orientata ai servizi web e garantisce elaborazioni scalabili on-the-fly<sup>58</sup>.
3. **Centralità dell'utente:** seamless user experience sui diversi canali (e.g. sito web desktop, mobile, etc.) e semplicità d'utilizzo e interpretazione degli strumenti e dei loro output.

### 9.1 Overview architetturale

L'architettura della piattaforma si basa sui principi *multi-tier*<sup>59</sup> e, in particolare, su una struttura a tre livelli. Questa scelta consente una maggiore semplicità di progettazione e implementazione, una maggiore manutenibilità, nonché scalabilità. I tre livelli adottati sono:

1. **Presentation Layer:** è lo strato software dedicato all'esposizione dei contenuti all'utente. Qui viene definita la grafica e le modalità d'interazione con l'utente e può differire a seconda del canale (e.g. sito web desktop, mobile, etc.). Questo strato non contiene alcuna logica di calcolo.
2. **Logic Layer:** è lo strato software dedicato al calcolo e all'implementazione delle logiche.
3. **Data Layer:** è lo strato dove vengono memorizzati i dati persistenti<sup>60</sup>.

---

<sup>58</sup>Eseguite nel momento in cui l'utente le richiede.

<sup>59</sup>Architettura nella quale vengono disaccoppiati i principali livelli logici software.

<sup>60</sup>Informazioni con finalità non d'utilizzo temporaneo (dati volatili)

Segue una rappresentazione schematica dell'architettura descritta:

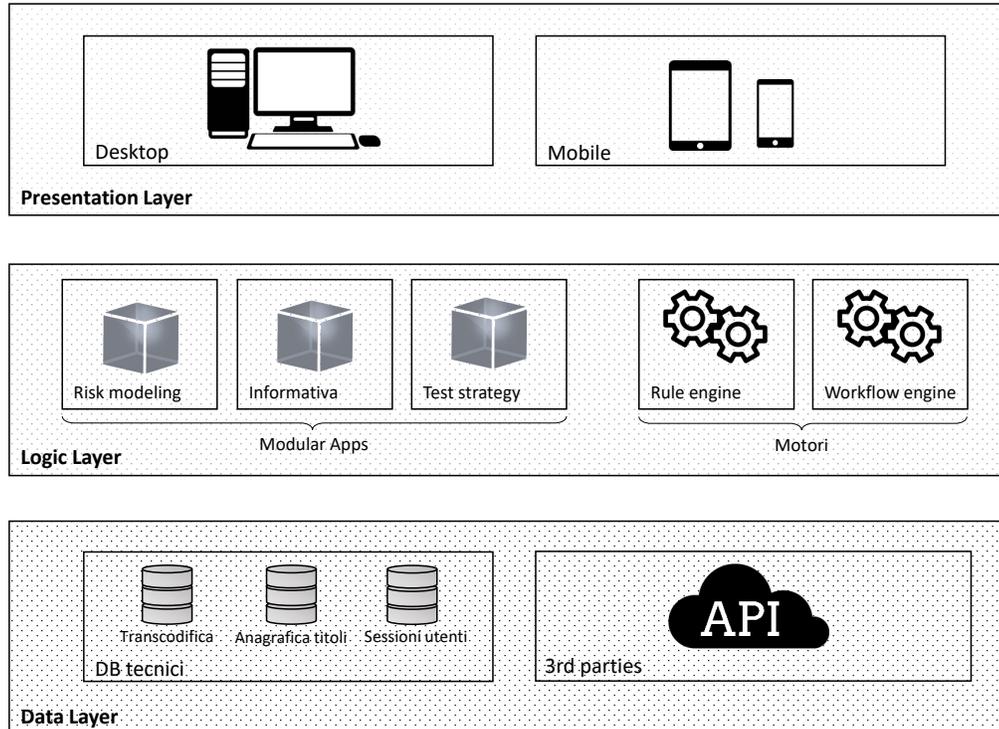


Figura 30: Rappresentazione delle componenti architettoniche divise nei tre tier: *Presentation Layer*, *Logic Layer* e *Data Layer*.

Il *Presentation Layer* è il luogo logico nel quale vengono gestite le componenti di visualizzazione dei contenuti e di interazione con l'utente. In questo tier sono stati utilizzati principalmente i linguaggi HTML5 e CSS per la componente grafica e Javascript per la gestione degli eventi. E' presente in questo tier anche una componente dedicata alle chiamate dei servizi offerti dallo strato logico.

Il *Logic Layer* è il luogo in cui vengono resi disponibili i servizi (applicazioni) richiamabili dal presentation layer. E' costituito da applicazioni per i calcoli statistici (principalmente realizzati nel linguaggio R), applicazioni per l'estrazione/manipolazione dati e motori di schedulazione dei processi e definizione delle regole (entrambi in Apache Tomcat).

Il *Data Layer* è il luogo dove vengono allocati i dati tecnici necessari al funzionamento della piattaforma (su DB open source di tipo MySQL) ed è il logo d'ingresso dei dati provenienti da servizi terzi (e.g. Google Finance, Yahoo Finance, Investing.com, etc.), richiamati tramite delle API<sup>61</sup>.

Segue la rappresentazione delle componenti architetture mostrate in *Figura 30* dal punto di vista tecnologico:

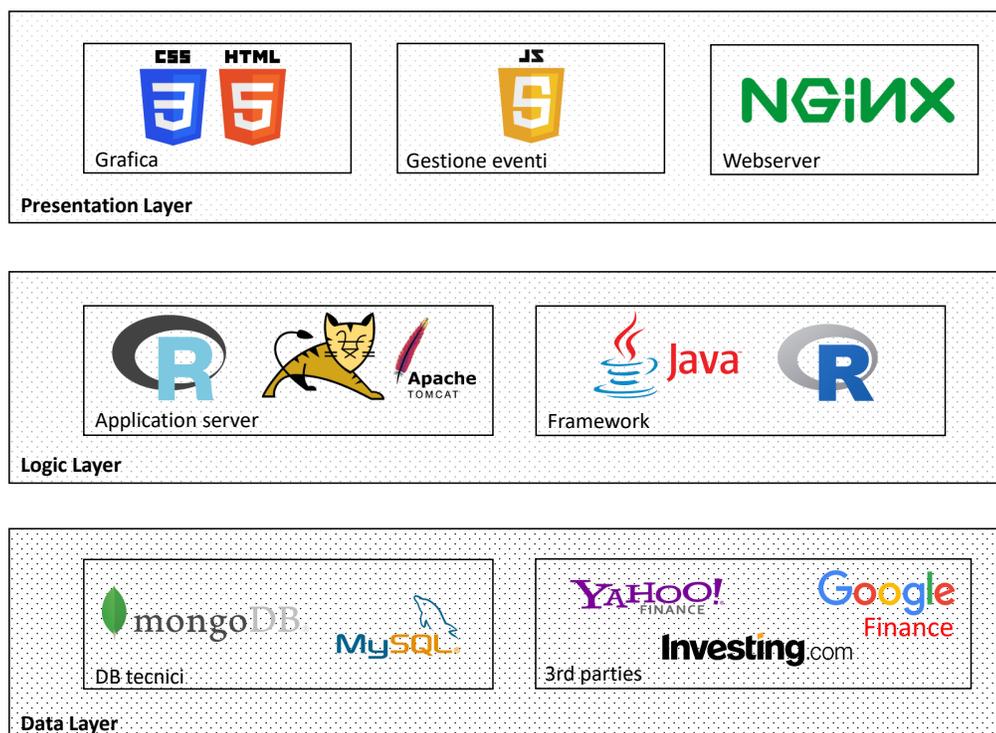


Figura 31: Rappresentazione delle componenti architetture dal punto di vista tecnologico divise nei tre tier: *Presentation Layer*, *Logic Layer* e *Data Layer*.

Verrà ora mostrata l'implementazione dell'architettura decritta all'interno della piattaforma.

<sup>61</sup>Con Application Programming Interface (API, in italiano interfaccia di programmazione di un'applicazione), si indica ogni insieme di procedure disponibili al programmatore, di solito raggruppate a formare un set di strumenti specifici per l'espletamento di un determinato compito all'interno di un certo programma.

## 9.2 Presentation Layer - Use case

Il *Presentation Layer* si basa sulla logica della *single page application*, ossia una sola pagina fissa (template HTML) viene dinamicamente popolata con del contenuto richiesto dall'utente.

In *Figura 32* viene mostrata l'home page della piattaforma sviluppata. L'interfaccia è suddivisa in moduli: una sidebar fissa nella parte sinistra e una pagina di contenuto dotata di diversi tab sulla destra. L'home-page corrisponde al tab *Chart & Prediction*, dove viene mostrato il grafico del titolo selezionato in sidebar e delle prediction.

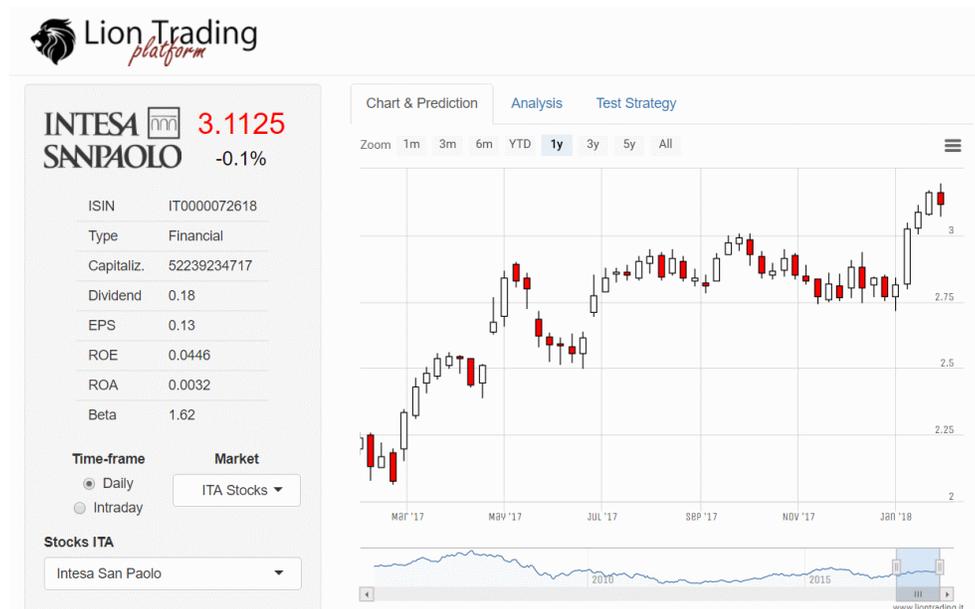


Figura 32: Home-page della piattaforma sviluppata.

Nel momento in cui si passa da un tab all'altro, il click dell'utente fa sì che venga recuperato il template specifico per il tab richiesto (pagina HTML nella quale viene definita la posizione e la grafica dei contenuti) che viene popolato con i contenuti forniti da specifici servizi del *logic layer*.

Ad esempio, cliccando sul tab *Analysis* (*Figura 33*), che contiene i principali indicatori di bilancio, la piattaforma recupera il template dello specifico tab e lo popola con i dati di bilancio. Tali dati sono ottenuti da alcuni servizi

REST<sup>62</sup> presenti sul *logic layer* attraverso la ricezione di un file JSON<sup>63</sup>.

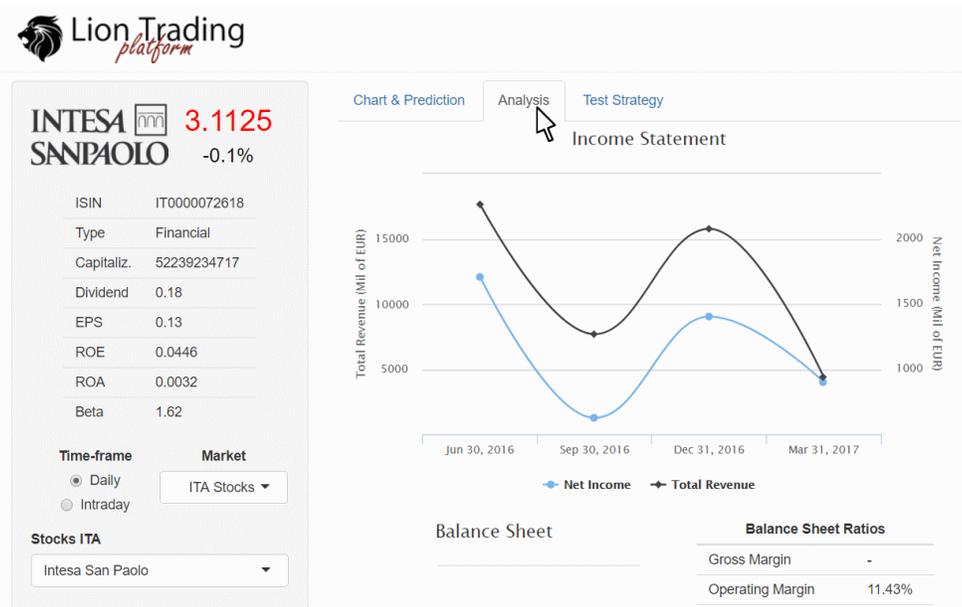


Figura 33: Tab “Analysis” della piattaforma sviluppata.

L’utente ha evidenza del fatto che si tratti un applicativo di tipo *single-page* in quanto la navigazione dei contenuti non provoca mai il reload di tutta la pagina: i contenuti vengono dinamicamente aggiornati. Questa caratteristica migliora notevolmente la user experience ed è nativamente multiplatforma.

L’approccio utilizzato prende il nome di MVC (Model View Controller) ed ha permesso una suddivisione completa tra la componente logica (su cui sono disponibili i servizi REST) e quella di presentazione.

<sup>62</sup>Un servizio REST - acronimo di REpresentational State Transfer - è un’interfaccia che mette in comunicazione i diversi layer software sfruttando il protocollo di comunicazione HTTP.

<sup>63</sup>Il JSON - acronimo di JavaScript Object Notation - è un formato adatto all’interscambio di dati fra applicazioni client-server.

### 9.3 Logic Layer - Use case

Analizziamo ora il caso d'uso di un servizio REST implementato sul *Logic Layer* della piattaforma.

A titolo d'esempio, consideriamo il modo il cui la piattaforma comunica la *Prediction* all'utente: si tratta di un'indicazione che deriva dal fit di un moto browniano asimmetrico sulla serie temporale. A seconda della probabilità che si ottenga un return positivo/negativo viene mostrato il messaggio “buy” o “sell”.

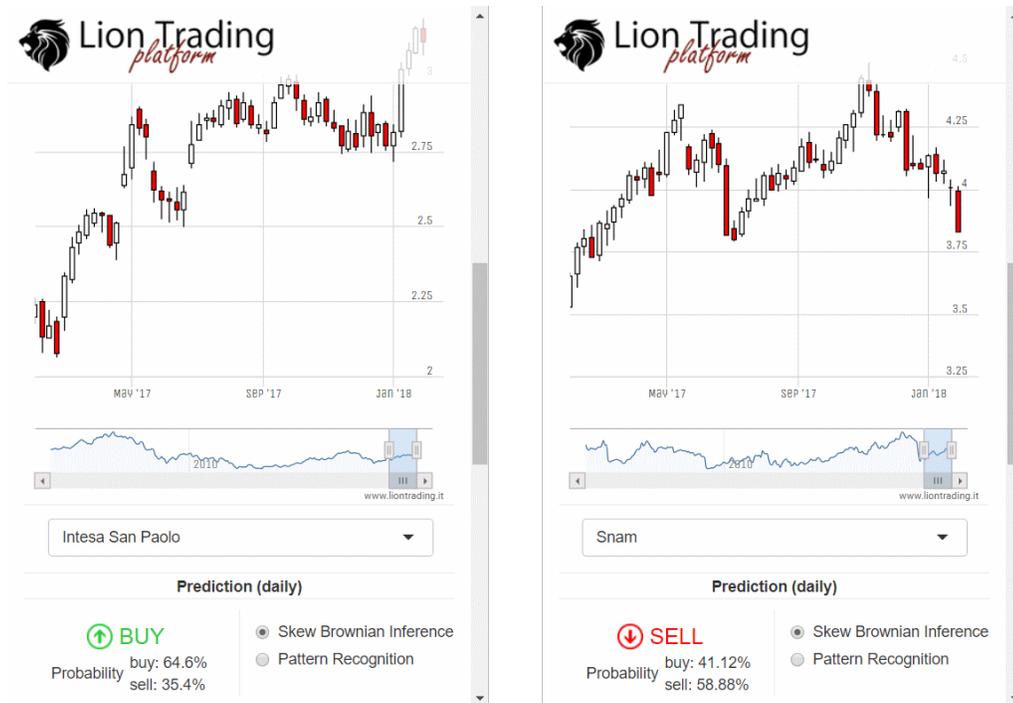


Figura 34: Esempio di output di un servizio REST sul front-end nella versione mobile dell'applicazione.

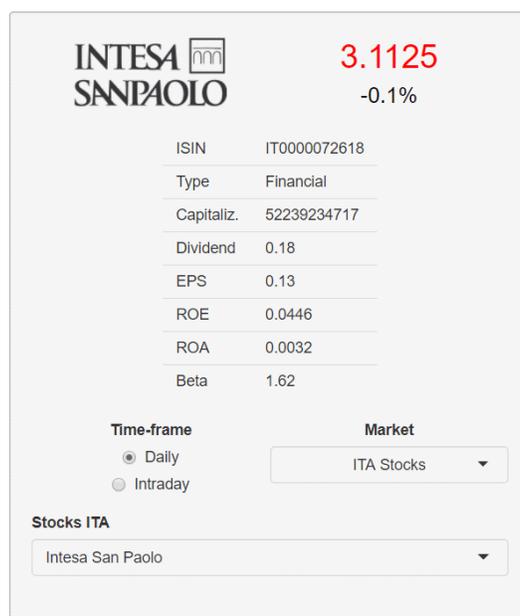
Nel momento in cui l'utente seleziona un titolo (e.g. “Intesa Sanpaolo” e “Snam” in *Figura 34*), viene effettuata una chiamata ad un servizio REST del *Logic Layer* reso disponibile da un applicativo Java. La chiamata include tutte le informazioni necessarie alla computazione: intervallo di tempo, titolo, mercato. Con tali informazioni, l'applicativo è in grado di recupe-

rare da altri servizi la serie temporale richiesta e passarla a degli script R (che vengono eseguiti su un ambiente R Server dedicato installato sulla medesima macchina) che effettuano il calcolo statistico per la stima dello Skew Brownian. Una volta che R ha terminato il Job<sup>64</sup>, l'applicativo Java si occupa di formattare l'output e renderlo fruibile da parte del front-end.

#### 9.4 Data Layer - Use case

Il *Data Layer* è composto da i database con i dati proprietari dell'applicativo e quelli di terze parti accessibili via API. Vediamo un caso d'uso in cui sono coinvolte queste componenti.

A titolo d'esempio, consideriamo la sidebar della piattaforma. In essa è possibile selezionare: mercato, titolo e time-frame per la rappresentazione e la computazione delle varie misure (si veda *Figura 35*).



The image shows a sidebar for the stock 'Intesa San Paolo'. At the top left is the logo 'INTESA SANPAOLO'. To the right of the logo, the current price is '3.1125' in red, with a change of '-0.1%'. Below this is a table of financial metrics:

ISIN	IT0000072618
Type	Financial
Capitaliz.	52239234717
Dividend	0.18
EPS	0.13
ROE	0.0446
ROA	0.0032
Beta	1.62

Below the table are two sections: 'Time-frame' with radio buttons for 'Daily' (selected) and 'Intraday'; and 'Market' with a dropdown menu showing 'ITA Stocks'. At the bottom, there is a section 'Stocks ITA' with a dropdown menu showing 'Intesa San Paolo'.

Figura 35: Sidebar della piattaforma realizzata.

<sup>64</sup>In informatica, per *Job* s'intende l'esecuzione di un processo che può essere schedato o eseguito su richiesta.

Nella sidebar è presente nella parte superiore il prezzo di ultimo scambio del titolo selezionato. Questa informazione, una volta effettuata la selezione da parte dell'utente (e.g. "Mercato Italiano" - "Intesa Sanpaolo" - "Daily"), viene esposta dal front-end secondo le logiche descritte nei paragrafi precedenti.

In particolare, il *Logic Layer* per reperire il dato e fornirlo al *Presentation Layer*, si rivolge al *Data Layer*. Il prezzo di ultimo scambio è un'informazione non disponibile all'interno delle basi dati dell'applicativo e quindi deve essere reperita da terze parti. Tale dato è ottenibile da *Google Finance* ed è richiamabile attraverso specifici URL<sup>65</sup> che vengono assemblati dall'applicativo sfruttando gli standard definiti nelle tabelle di transcodifica<sup>66</sup> disponibili nelle basi dati proprietarie.

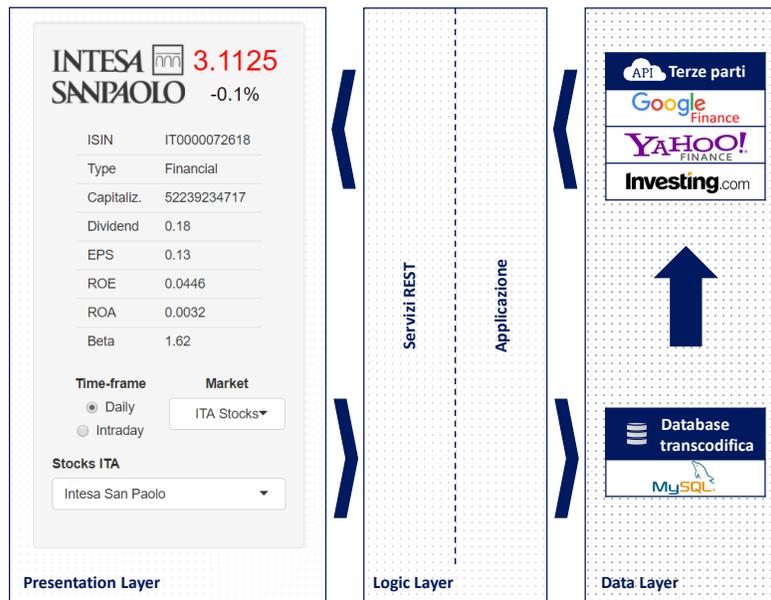


Figura 36: Sidebar della piattaforma realizzata.

<sup>65</sup>L'Uniform Resource Locator (in acronimo URL) è una sequenza di caratteri che identifica univocamente l'indirizzo di una risorsa in in network, tipicamente presente su un host server come ad esempio un documento, un immagine, un video o un servizio, rendendolo accessibile ad un client.

<sup>66</sup>La transcodifica, in informatica, è quell'operazione con la quale una certa informazione viene trasformata utilizzando specifiche regole di codifica.

## 10 Utilizzo della piattaforma

Questo capitolo finale verterà sugli aspetti funzionali della piattaforma implementata, indicando le modalità con le quali, un utente che si affaccia al mondo del trading, debba interagire con essa al fine di fare delle scelte di investimento consapevoli ed accurate.

Dopo una breve overview sull'interfaccia nel suo complesso e le modalità d'interazione con l'utente, verranno approfonditi i vari servizi offerti descrivendo in particolare il modo in cui possono fornire degli insight per la negoziazione.

### 10.1 User Experience

All'apertura della piattaforma, l'interfaccia si presenta all'utente come mostrato di seguito in *Figura 37*:

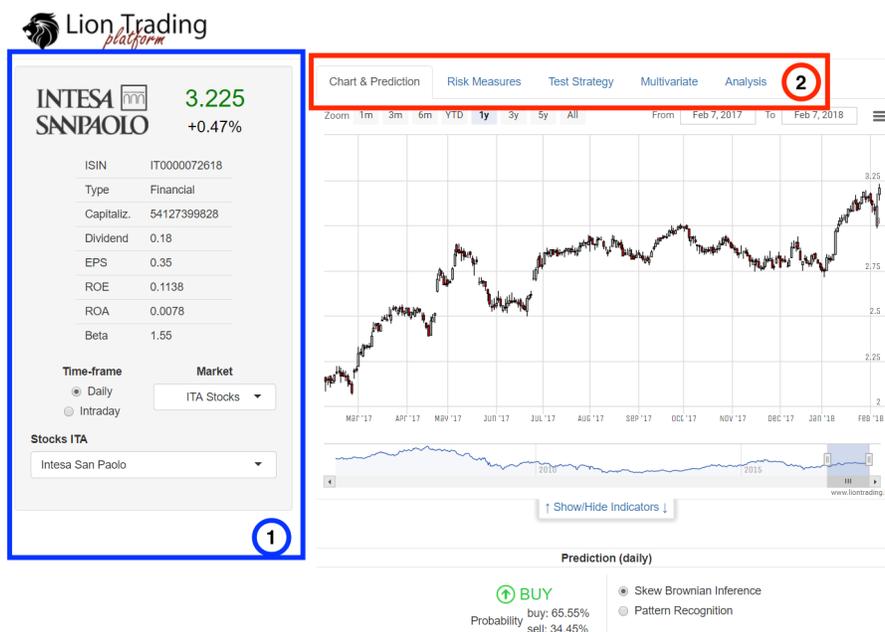


Figura 37: Principali elementi dell'interfaccia della piattaforma.

Il layout della pagina é composto da tre elementi, due dei quali statici e uno dinamico. Il primo elemento (rettangolo in blu in *Figura 37*) è una sidebar

laterale in cui sono presentate delle informazioni relative al titolo selezionato. Il secondo elemento della pagina (rettangolo rosso in *Figura 37*) è un menù che consente la navigazione tra le diverse pagine della piattaforma per l'accesso a specifiche funzionalità. Il terzo elemento è il *body*<sup>67</sup> della pagina, che cambia dinamicamente a seconda della valorizzazione dei due elementi precedenti.

Riportiamo in *Figura 38* il dettaglio della sidebar laterale.

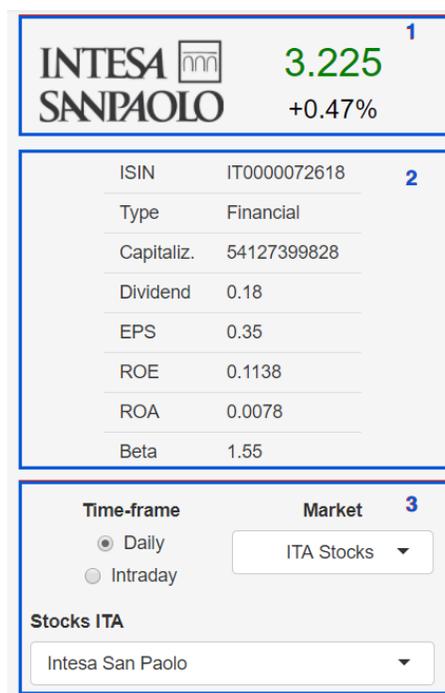


Figura 38: Sidebar della piattaforma realizzata.

La componente è suddivisa in 3 parti:

1. in alto vi è il logo del titolo selezionato, il prezzo a cui è stato effettuato l'ultimo scambio (colorato in verde se la variazione rispetto all'ultimo prezzo di chiusura della giornata precedente è positiva, in rosso altrimenti) e al di sotto la relativa variazione percentuale;

<sup>67</sup>In informatica, parlando di front-end, per *body* s'intende la porzione dell'interfaccia in cui viene mostrato il contenuto principale.

2. al centro ci sono alcune informazioni salienti (e.g. capitalizzazione, importo ultimo dividendo, etc.) sul del titolo selezionato e i più comuni indicatori utilizzati (e.g. EPS, ROA, ROE, etc.);
3. in basso è presente la modalità d'interazione con l'utente: un form in cui è possibile selezionare il time-frame<sup>68</sup>, il mercato di riferimento e infine il titolo.

Riportiamo in *Figura 39* il dettaglio del menù che consente di spostarsi<sup>69</sup> all'interno dell'applicazione tra le varie funzionalità.



Figura 39: Barra del menù della piattaforma

Il menù consente di accedere alle seguenti funzionalità:

- *Chart & Prediction*: homepage dell'applicazione, consente di fruire di alcuni strumenti di analisi tecnica e di consultare i principali forecast quantitativi;
- *Risk Measures*: sezione dedicata alle misure di rischio;
- *Test Strategy*: sezione per il test della prediction eseguita sotto ipotesi di omoschedasticità;
- *Multivariate*: sezione dedicato ai fit ed ai forecast multivariati;
- *Analysis*: sezione dedicata all'analisi fondamentale di bilancio.

## 10.2 Use case tipico

Verrà ora mostrato un dettagliato esempio di come un investitore possa utilizzare la piattaforma come strumento di supporto all'attività di trading. Ripercorriamo perciò passo passo i vari step che l'investitore dovrebbe seguire per fare un'analisi completa prima di procedere alla negoziazione di uno strumento finanziario.

<sup>68</sup>Il livello di risoluzione con il quale vuole analizzare la serie temporale (e.g. giornaliero, al minuto, etc).

<sup>69</sup>L'uso del termine *spostarsi* ha una accezione puramente funzionale in quanto il click su una delle voci del menù provoca l'aggiornamento dinamico del body e non il caricamento fisico di una nuova pagina.

### 10.2.1 Grafici e Analisi Tecnica

Il primo elemento che può fornire una serie di informazioni utili è sicuramente il grafico (in *Figura 40* quello visibile dalla sezione *Chart & Prediction*). E' possibile risalire a moltissimi fattori di analisi interpretando semplicemente il grafico del titolo scelto per valutare la bontà di un investimento.



Figura 40: Grafico a candele presente nella sezione *Chart & Prediction*.

Questa primissima analisi viene effettuata utilizzando grafici di tipo a candela, che sono molto ricorrenti nell'ambito finanziario. Questa rappresentazione consente di fornire per ogni tempo una pluralità d'informazioni: prezzo d'apertura, di chiusura, escursione min e max.

Il grafico in *Figura 40* rappresenta l'andamento dell'ultimo anno del titolo Intesa Sanpaolo (azienda quotata sull'indice italiano FTSEMIB). Sull'asse delle ascisse il tempo, mentre su quello delle ordinate il prezzo unitario di scambio di un'azione.

La selezione in sidebar del time frame fa variare dinamicamente il grafico. Ogni candela, infatti, assume valori diversi a seconda del time frame scelto. Nell'esempio è stato scelto un time frame giornaliero, dunque ogni candela corrisponde ad una giornata nel quale ci sono stati degli scambi sul mercato.

Le candele bianche nel grafico segnalano una variazione positiva, mentre quelle rosse (o nere) una variazione negativa. In *Figura 41* un diagramma esplicativo della rappresentazione *a candela*.

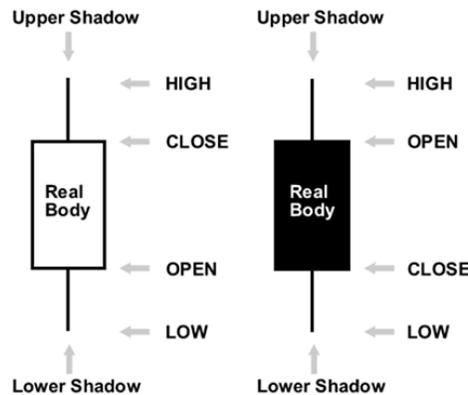


Figura 41: Il significato della rappresentazione *a candela*

Il *real body* mostra la differenza tra prezzo di apertura e di chiusura. L'*upper shadow* e la *lower shadow* ci indicano quali sono stati i valori massimi e minimi raggiunti durante un determinato periodo di scambi sul mercato (e.g. un giorno, un minuto, etc.).

L'utilizzo di un grafico a candele in luogo di quello a linea è una scelta comune di moltissimi trader in quanto consente di desumere una serie di informazioni ulteriori sulle condizioni del mercato. Ad esempio, se in un grafico ci sono una serie di candele con un *real body* molto grande e *shadows* molto piccole questo potrebbe indicare poca liquidità durante il periodo preso in considerazione.

### Pattern recognition

Come già accennato, l'uso delle candele consente ai trader di desumere certe informazioni ulteriori legate al contesto di mercato rispetto all'uso di grafici a linea. In particolare, sono comunemente utilizzati i riferimenti di alcune tipiche conformazioni grafiche come suggerimento per individuare un trend<sup>70</sup> o un'inversione di trend. Solitamente questo tipo di attività è effettuata in

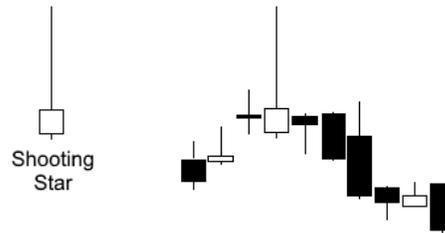
<sup>70</sup>Per *trend* s'intende la tendenza rialzista (bullish) o ribassista (bearish) del mercato.

maniera manuale e prende il nome di *analisi tecnica*.

Nella piattaforma è stato implementato un algoritmo di machine learning (supervised) per identificare questi pattern noti e fornire un'indicazione di trend all'utente.

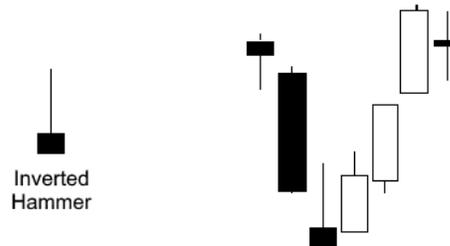
Al fine di far riconoscere alla macchina i pattern, è stato necessario classificare manualmente una serie di configurazioni grafiche basandosi sulle definizioni classiche adottate in analisi tecnica fondate sull'esperienza empirica dei trader. Sono state classificate oltre 50 figure, tra le quali ricordiamo lo *shooting star*, l'*hammer* e l'*inverted hammer* per i quali segue un breve approfondimento:

- Lo *shooting star* presenta una upper shadow molto pronunciata, almeno doppia rispetto al real body della candela e può avere anche una piccola lower shadow. È un segnale ribassista (bearish).



Il colore della candela è indifferente; quando si verifica durante un trend rialzista è un possibile segnale di inversione di tendenza, è sempre necessario attendere ulteriore conferma che validi un'apertura di posizione ribassista considerando come valido livello di stop il massimo della candela. Un'ombra più lunga viene interpretata come un maggiore potenziale di inversione.

- L'*inverted hammer* è analogo alla shooting star; in questa formazione si ha un'ombra superiore molto estesa, almeno doppia rispetto al corpo della candela.



Si trova in posizione di minimo e segnala che le iniziative dei rialzisti iniziano ad avere successo, ma non in modo sufficiente perché il mercato si assesti livelli più alti. L'eventuale conferma nella candela seguente può generare un segnale di acquisto; il colore della candela anche qui è indifferente.

- L'*hammer* segnala pressione vendite ancora sostenuta, ma non sufficiente; il colore è indifferente, ma deve presentare un'ombra ribassista almeno doppia rispetto al corpo della candela.



Eventuale conferma di un possibile cambiamento di trend verrà data solo dalla candela seguente in poi.

E' importante tenere presente che l'individuazione di questi pattern senza determinate ipotesi di partenza (e.g. andamento della media mobile) non riescono ad effettuare una precisa *detection* di trend rialzista o ribassista in quanto principalmente sono semplici segnali di cambiamento di trend.

Tutte queste considerazioni vengono svolte in autonomia dalla piattaforma che, una volta selezionato il titolo, mostra il primo pattern identificato dall'algorithm a partire dall'ultima candela. Il dato è visibile (come in *Figura 42*) sotto il grafico principale della sezione *Chart & Prediction*. Oltre a specificare il nome del pattern individuato, viene indicata la tipologia: Bearish nel caso in cui è associato ad un segnale di vendita, Bullish di acquisto e Neutral se il pattern non fornisce indicazioni.



Figura 42: Sezione della piattaforma dedicata alla Pattern recognition

## Gli indicatori

Oltre alla *pattern recognition*, l'analisi tecnica può essere integrata con l'uso di una serie d'indicatori sovrapposti al grafico a candela.

La piattaforma, infatti, rende disponibile nella sezione *Chart & Prediction* uno strumento per aggiungere al grafico questi indicatori sovrapposti (e.g. medie mobili, bande di Bollinger, MACD, etc.) con l'obiettivo di allargare il dominio di analisi dell'utente rispetto al solo studio delle candele e delle conformazioni grafiche riconosciute.

The image shows a user interface for selecting technical indicators. It features three indicator slots, each with a checked checkbox and a numerical value of 20. Indicator I is set to 'Bollinger Bands'. Indicator II is set to 'Mov Avg (arith)', and a dropdown menu is open for it, showing options: 'Mov Avg (arith)', 'Mov Avg (exp)', 'Bollinger Bands', 'MACD', and 'RSI'. Indicator III is set to 'RSI'. At the bottom, there is a button labeled 'Show/Hide Indicators' with up and down arrows.

Figura 43: Form d'interazione con l'utente per la scelta degli indicatori sovrapposti.

Come rappresentato in *Figura 43*, l'utente può inserire scegliere tra una serie di indicatori sovrapposti. In particolare, si possono inserire: Moving Average, RSI, MACD e Bollinger Bands. In *Figura 44* vengono mostrati a titolo d'esempio tre di questi indicatori:

- *Moving Average (arith)*: si tratta della media mobile semplice, viene calcolata puntualmente considerando la media algebrica dei dati di un determinato periodo di tempo precedente.
- *RSI*: il Relative Strength Index (indice di forza relativa) è un indicatore di momentum<sup>71</sup> calcolato come

$$RSI = 100 \frac{U}{U + D}$$

dove  $U$  e  $D$  sono, rispettivamente, le medie delle differenze di chiusure al rialzo e al ribasso in un determinato intervallo di tempo. Per come è costruito, l'indicatore può variare da 0 (condizione di ipervenduto,

<sup>71</sup>Il *momentum* in finanza indica la forza del mercato misurando il tasso di variazione dei prezzi rispetto ai loro livelli effettivi.

ci si aspetta un trend rialzista) ad 1 (condizione di ipercomprato, ci si aspetta un trend ribassista).

- *Bollinger Bands*: sono la rappresentazione della volatilità di un titolo in termini di deviazione standard.



Figura 44: Grafico a candele con indicatori sovrapposti: Moving Average in blu, RSI in giallo e Bollinger Bands in blu.

## 10.2.2 Analisi fondamentale

Sebbene i sostenitori dell'analisi tecnica ritengano che il prezzo incorpori già tutte le informazioni economiche necessarie, può essere opportuno - soprattutto se si ha un orizzonte temporale d'investimento molto ampio - considerare i razionali economici di base del titolo che si vuole negoziare osservando i bilanci d'impresa e, in particolare, alcuni indicatori calcolabili a partire dalle informazioni ivi contenute. Questo tipo d'indagine prende il nome di *analisi fondamentale*.

La piattaforma rende disponibili la maggior parte degli indicatori e delle misure considerate dall'analisi fondamentale nella sezione *Analysis* accessibile dal menù superiore, come mostrato in *Figura 45*.

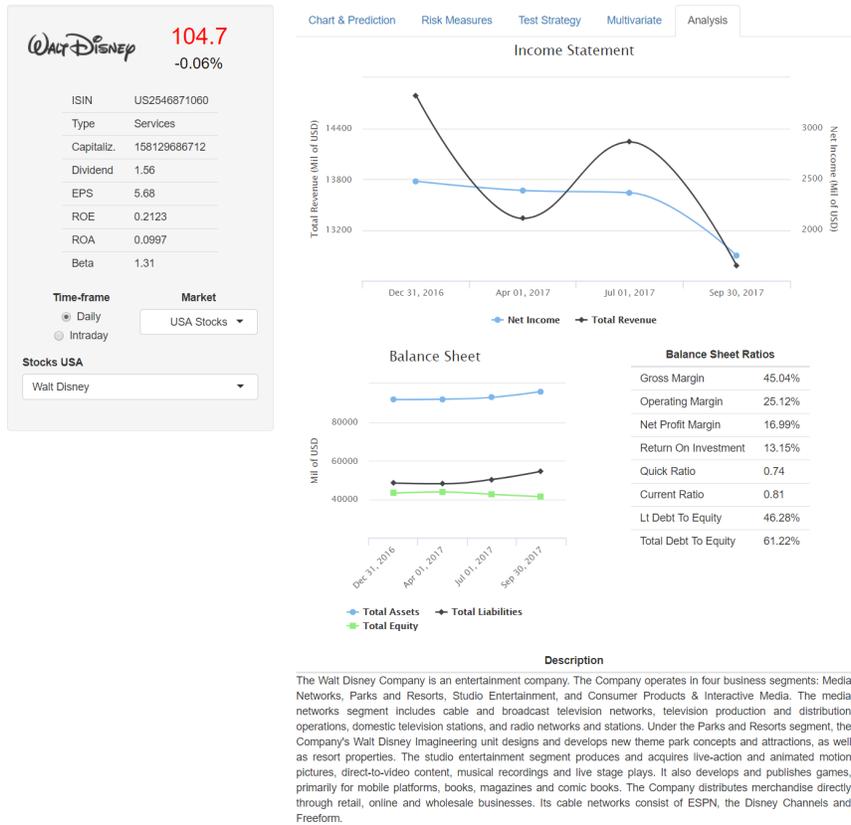


Figura 45: Screenshot della piattaforma nella sezione *Analysis*

Al fine di comprendere i contenuti della sezione e la funzione degli indicatori esposti è bene perciò anzitutto richiamare alcuni concetti basilari di ragioneria.

### Cenni di principi contabili e principali indicatori

I principi contabili internazionali (IAS) prescrivono per le società quotate l'obbligo a redigere pubblicamente un bilancio composto necessariamente da tre documenti: lo stato patrimoniale, il conto economico ed il rendiconto finanziario. In questa sede considereremo solo SP e CE, mentre non verrà descritto il rendiconto finanziario poiché completamente determinabile

a partire dai primi due e quindi ridondante per il nostro tipo di indagine.

### Stato Patrimoniale

Lo stato patrimoniale è quel documento che ha lo scopo di “fotografare” la situazione patrimoniale di un’impresa al momento della chiusura del periodo d’esercizio. In particolare è organizzato in due sezioni contrapposte, a sinistra vi è l’ *attivo* e a destra il *passivo*. Il primo raggruppa l’ammontare complessivo degli investimenti in essere in un dato momento (si tratta di macchinari, immobili, marchi, cespiti di ogni genere, crediti, ecc.), mentre il secondo raggruppa le fonti che hanno reso possibile gli investimenti (patrimonio netto o equity, finanziamenti, utile non distribuito, ecc.).

ATTIVO (ASSETS)		PASSIVO
IMPIEGHI	←	FONTI (EQUITY)
INVESTIMENTI	←	FINANZIAMENTI (LIABILITIES)

Un importante indicatore facilmente ricavabile direttamente dallo stato patrimoniale è il *debt/equity ratio*, che esprime il rapporto tra il totale dei debiti su cui si pagano gli interessi e il patrimonio netto dell’azienda (ossia i capitali provenienti dai soggetti che possiedono l’azienda). Viene utilizzato per verificare il grado di dipendenza dell’impresa da fonti finanziarie esterne: quanto più alto è il suo valore tanto meno equilibrata è giudicata la struttura finanziaria.

### Conto Economico

Il conto economico è quel documento che raggruppa i costi e i ricavi di competenza del periodo amministrativo, mostrando il risultato economico della gestione: viene misurato l’incremento o il decremento che il capitale netto aziendale ha subito per effetto della gestione di esercizio.

Operativamente il CE si presenta come una sommatoria di addendi a diverso segno intervallata da dei totali parziali:

CONTO ECONOMICO
+ Vendite (fatturato)
- Costi produzione
<b>= Margine operativo lordo (a)</b>
- Ammortamenti
- Accantonamenti
<b>= Margine operativo netto (b)</b>
+/- Altri ricavi e spese
<b>= Risultato ante imposte (c)</b>
- Imposte
<b>= Risultato netto (utile/perdita) (d)</b>

Lo schema precedente deve chiaramente essere considerato come un'estrema approssimazione, ma consente di mettere in evidenza alcuni parziali salienti: il *marginale operativo lordo* evidenzia il reddito di un'azienda basato solo sulla sua gestione operativa, quindi consente di comprendere il risultato del core business dell'impresa; il *marginale operativo netto* è ottenuto dal precedente considerando anche l'impatto di ammortamenti e accantonamenti; il *risultato netto* - se positivo - mostra l'ammontare effettivo che può essere distribuito agli azionisti della società (ossia coloro che detengono delle quote di capitale sociale) oppure reinvestito nella stessa durante l'esercizio successivo.

A partire dal CE - tra i tanti - è possibile direttamente calcolare: l'*operating margin*, che esprime il rapporto tra il margine operativo netto ed il fatturato e consente perciò di capire che proporzione dell'incasso viene utilizzata per pagare i costi variabili di produzione (materiali, personale, ecc.); il *net profit margin*, che esprime il rapporto tra risultato netto e fatturato e consente di capire quanta parte del venduto si trasforma in concreto profitto (o perdita); l'*EPS*, ossia il rapporto tra il risultato netto (privato dei dividendi privilegiati) ed il numero delle azioni in circolazione.

Altri due indicatori che vengono esposti in piattaforma sono il *ROE* ed il *ROA*, che sono calcolati come rapporti incrociati tra voci del CE e dello SP: il primo è ottenuto come rapporto tra il risultato netto (da SP) ed il patrimonio netto (da CE) e quindi costituisce un indice di

redditività del capitale proprio; il secondo è ottenuto come rapporto tra il risultato netto (da CE) ed il totale attivo (da SP) e quindi dà indicazione della redditività di un'impresa prescindendo dall'organizzazione delle fonti di finanziamento permettendo efficaci raffronti tra aziende che operano nello stesso settore.

Unitamente agli indicatori economico-patrimoniali visti fin qui, vengono esposti in piattaforma anche: il *rendimento dell'ultimo dividendo* rispetto al prezzo attuale; il *1 Year return*, ossia l'incremento o il decremento percentuale del prezzo delle azioni dell'impresa nell'ultimo anno; il livello di *capitalizzazione*, ottenuto come prodotto tra il numero di azioni in circolazione ed il loro valore unitario attuale; la media dei *volumi scambiati* sul mercato negli ultimi 3 mesi; il *beta*, ovvero la variazione che un titolo storicamente assume rispetto alle variazioni del mercato; il *settore* in cui opera l'azienda; il paese in cui è quotata.

### 10.2.3 Metodi quantitativi e analisi del rischio

Dopo lo studio qualitativo basato sull'analisi tecnica e fondamentale, verranno approfonditi ora gli aspetti più prettamente quantitativi, legati alla modellazione delle serie temporali ed il calcolo dei forecast.

#### Le misure di rischio

Uno dei fattori fondamentali per stabilire se investire in uno strumento piuttosto che un altro è il rischio connesso all'investimento. Le misure di rischio descritte nella parte iniziale dell'elaborato possono aiutare l'investitore a confrontare la sua propensione al rischio con quello connesso all'investimento e ad effettuare delle scelte, nonché avere dei riferimenti per impostare *stop loss* e *take profit*<sup>72</sup> dei titoli negoziati.

Questo tipo d'informazioni è disponibile sulla piattaforma nel tab *Risk Measures* accessibile dal menù superiore. La pagina si presenta come in *Figura 46*) e presenta:

---

<sup>72</sup>Dopo aver acquistato o venduto allo scoperto un titolo, la maggior parte di broker consente d'inserire delle soglie di profitto e perdita oltre le quali chiudere automaticamente la posizione.

- Nella parte superiore un grafico a candele e due linee sovrappresse: in blu il Value at Risk e in rosso l'Expected Shortfall. Questi due valori, calcolati partendo dai return e rappresentati a livello di prezzo, danno un'idea ben precisa di quali siano le soglie massime entro le quali ci si può aspettare la caduta del prezzo del titolo con un certo livello di confidenza.

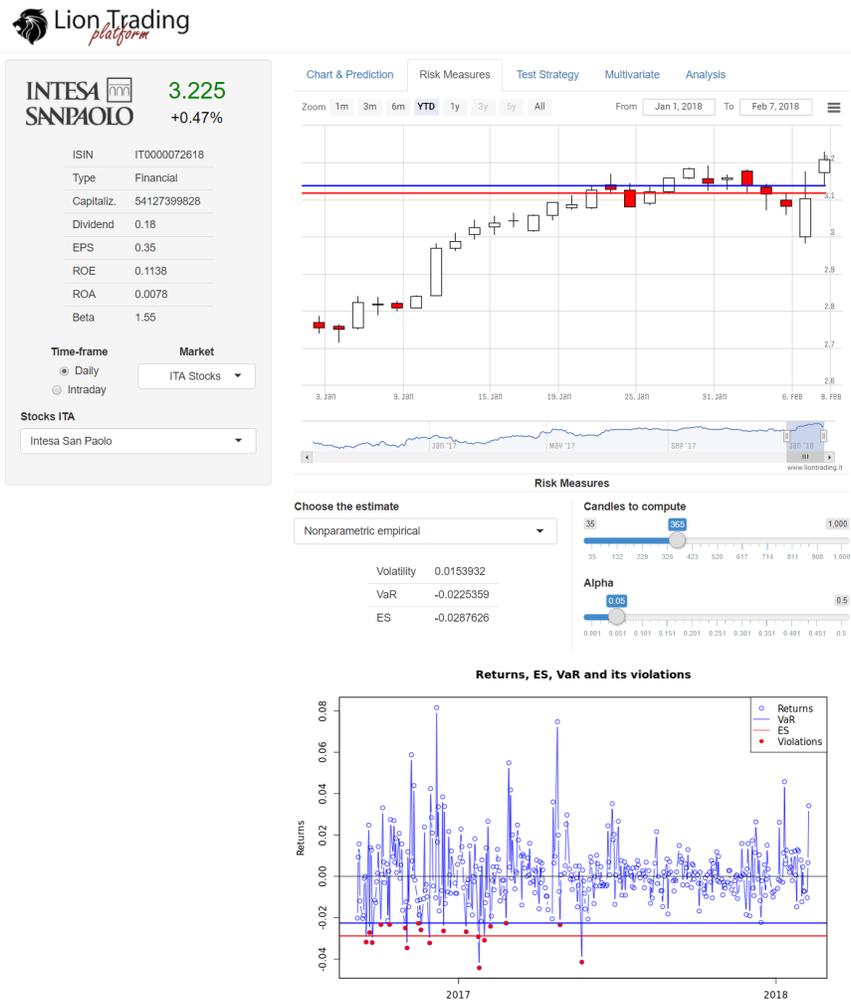


Figura 46: Sezione dedicata alle misure di rischio

- Nella parte centrale sono presenti le componenti d'interazione con l'utente e una tabella riportante il valore delle misure di rischio. E' possibile scegliere il metodo (e.g. parametrico, non parametrico) con il quale calcolare il Value at Risk, la volatilità e l'Expected ShortFall nonché l'intervallo di tempo sul quale si è interessati a calcolare le misure di rischio e l' $\alpha$  associato al calcolo dei quantili (che di fatto indicano la propensione al rischio dell'utente).
- Nella parte inferiore il grafico dei return con in sovrapposizione le rette corrispondenti al VaR e all'ES. Appaiono evidenziati i return che hanno violato il VaR.

Questi indicatori e le loro rappresentazioni costituiscono sostanzialmente l'implementazione dei contenuti descritti nel *Capitolo 4* e *6*.

Compresa l'entità del rischio è ora possibile utilizzare modelli predittivi basati sulla modellazione delle serie temporali per comprendere se acquistare o vendere allo scoperto un determinato titolo.

### Previsione sotto ip. $\text{return} \sim \text{i.i.d.}$

Nella sezione *Chart & Prediction*, sotto il grafico principale, è possibile visualizzare la previsione effettuata fittando un moto browniano asimmetrico sugli ultimi 40 return. Nel caso in cui la probabilità di un return positivo sia maggiore del 65% o minore del 35% verrà mostrata, rispettivamente, l'indicazione *Buy* o *Sell*.

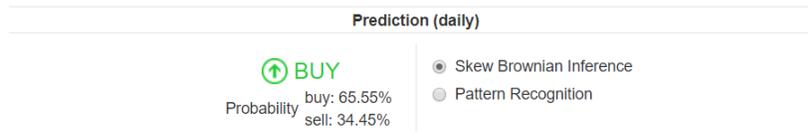


Figura 47: Previsione basata sul forecast di un moto browniano asimmetrico.

Questo tipo di modello si basa sull'assunto che i return presenti nell'ultima finestra temporale (è possibile scegliere un time frame sia daily che intraday) possano essere considerati indipendenti e identicamente distribuiti. Chiaramente si tratta di un'ipotesi molto forte e che si rivela particolarmente efficace solo su certi titoli. Per comprendere se tale

assunzione è valida su un certo titolo o meno è presente in piattaforma la sezione *Test Strategy*, nella quale è possibile simulare il rendimento che si sarebbe ottenuto su un intervallo di tempo definibile dall'utente se si fosse acquistato e venduto l'asset seguendo unicamente gli insight forniti dal moto browniano asimmetrico.

In *Figura 48* la sezione *Test Strategy*, sulla destra è possibile scegliere l'ampiezza dell'intervallo di tempo sui cui effettuare il test, sulla sinistra vengono mostrati i risultati della simulazione, mentre nella parte inferiore è presente un grafico che mostra i momenti in cui è stata effettuata un'operazione. Nel caso di un net profit maggiore di zero, può avere senso considerare le previsioni del moto browniano all'interno dell'analisi complessiva.

Questo previsione sintetizza perciò le considerazioni descritte nel *Capitolo 2* a riguardo delle distribuzioni dei returns, dato che la distribuzione normale asimmetrica che viene fittata dall'applicazione include come caso particolare quello della normale simmetrica.

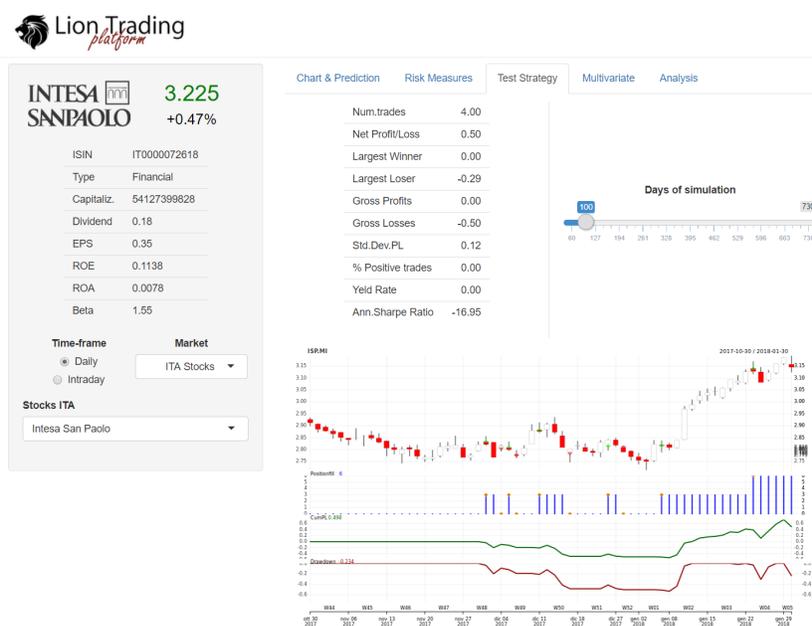


Figura 48: Sezione dedicata al test del moto browniano asimmetrico

## Previsione sotto ip. return autocorrelati

Poiché spesso non è possibile ipotizzare che i return siano i.i.d., la piattaforma mette a disposizione dell'utente una serie di strumenti classici per l'analisi dei return sotto ipotesi di autocorrelazione.

In particolare, nella parte inferiore della sezione *Chart & Prediction* è possibile fittare i modelli ARMA, GARCH e GARCH asimmetrico ed osservare i relativi forecast, come mostrato nel riquadro evidenziato in *Figura 49*.

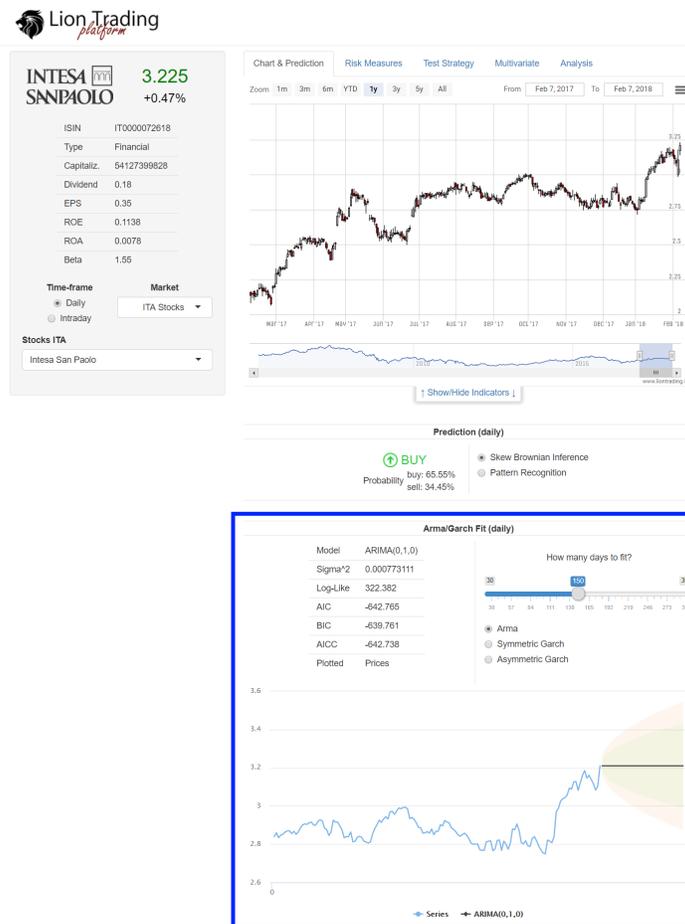


Figura 49: Sezione *Chart & Prediction* della piattaforma; la parte inferiore - evidenziata in blu - è dedicata ai modelli ARMA, GARCH e GARCH asimmetrico.

Attraverso un *radio button* è possibile scegliere il modello da fittare e utilizzare una barra di selezione per impostare il numero di giorni su cui effettuare la modellazione. A differenza del modello basato sul moto browniano asimmetrico, in questa circostanza non vengono utilizzati i return intraday ma solo quelli daily dato che con i primi la modellazione risulta meno efficace.

In *Figura 49* è possibile osservare il forecast di un modello ARMA, i cui parametri sono stati scelti automaticamente in base ai valori di AIC, BIC e AICC. La spezzata azzurra rappresenta i prezzi storici, la retta in nero il forecast, mentre in giallo chiaro e in giallo scuro sono mostrati, rispettivamente, gli intervalli di confidenza al 95% e all'80%.

Per quanto riguarda il GARCH (rif. *Figura 50*), in azzurro vengono mostrati i log-return, in blu la previsione di questi ultimi, in grigio due volte la deviazione standard ed in rosso il suo forecast.

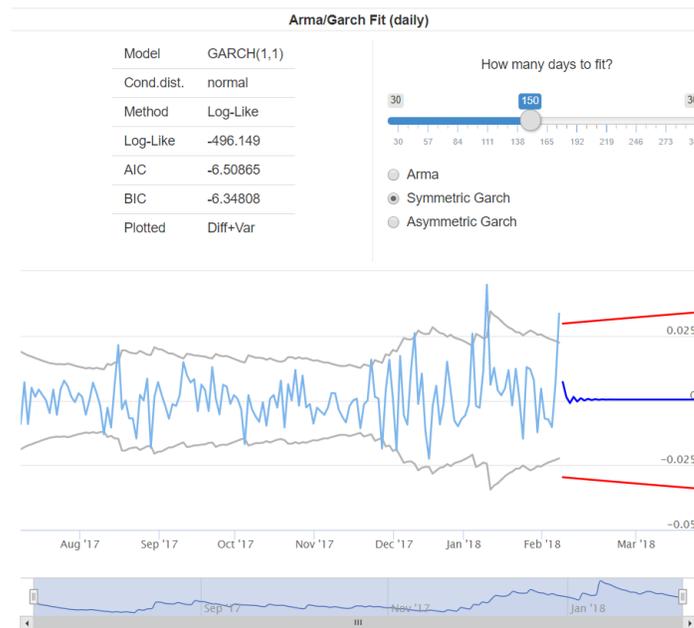


Figura 50: Fit e forecast di un modello GARCH sugli ultimi 150 return.

La scelta di utilizzare il GARCH asimmetrico rispetto a quello simme-

trico dipende dalle evidenze empiriche riscontrate: se i cambiamenti di prezzo negativi corrispondono ad una maggiore volatilità rispetto alle variazioni positive, il modello asimmetrico è da preferirsi.

I modelli ARMA e GARCH fittati dall'applicazione in questa sezione costituiscono di fatto l'implementazione applicativa dei contenuti descritti nel *Capitolo 3*, nel *Capitolo 7* e nel *Capitolo 8* dell'elaborato.

### Approccio multivariato

Come ultimo strumento per l'analisi, la piattaforma rende disponibile la sezione *Multivariate*, dove è possibile fruire di un'analisi congiunta su due titoli selezionati dall'utente.

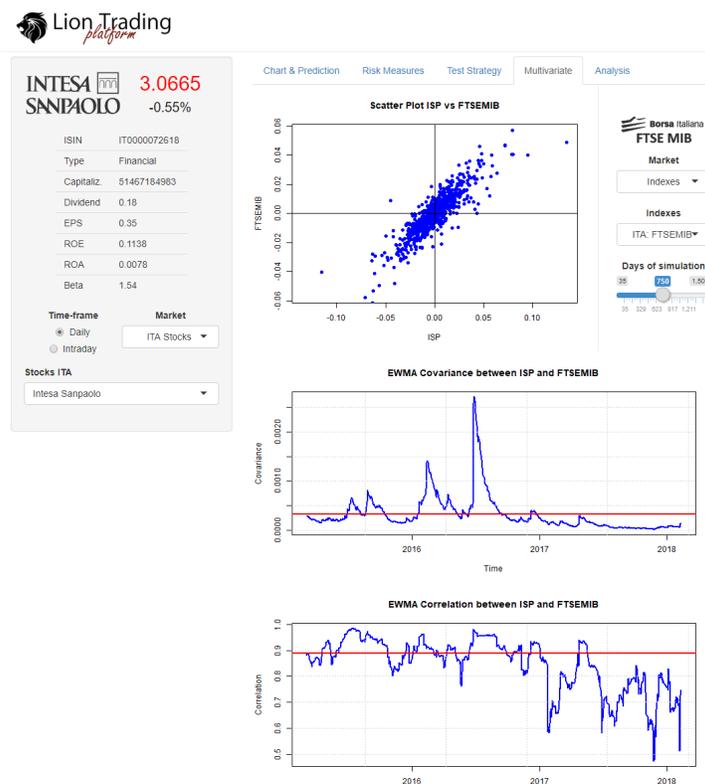
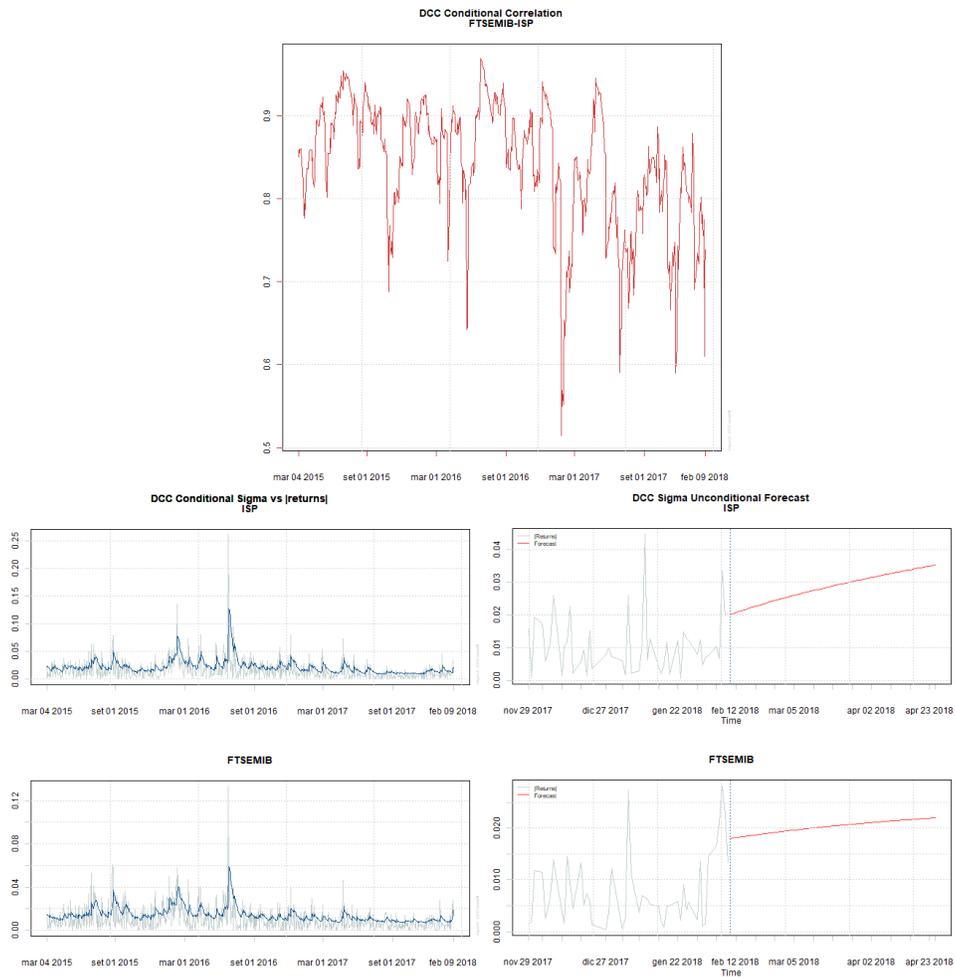


Figura 51: Sezione *Multivariate* della piattaforma; parte superiore della pagina.

Nella parte superiore destra è possibile scegliere il secondo titolo con

cui effettuare l'analisi congiunta nonché l'intervallo di tempo da utilizzare. Nella prima parte della pagina (*Figura 51*) è presente uno scatter plot dei return dei due titoli, il grafico della covarianza EWMA e quello della correlazione EWMA. Nella seconda parte della pagina (*Figura 52*) viene fittato un modello DCC: è rappresentata la correlazione condizionata DCC dei due titoli, la  $\sigma$  DCC condizionata fittata, il suo forecast e un grafico col forecast dei return di un portafoglio equilibrato formato dai due titoli e le relative bande del VaR con  $\alpha = 2, 5\%$ .



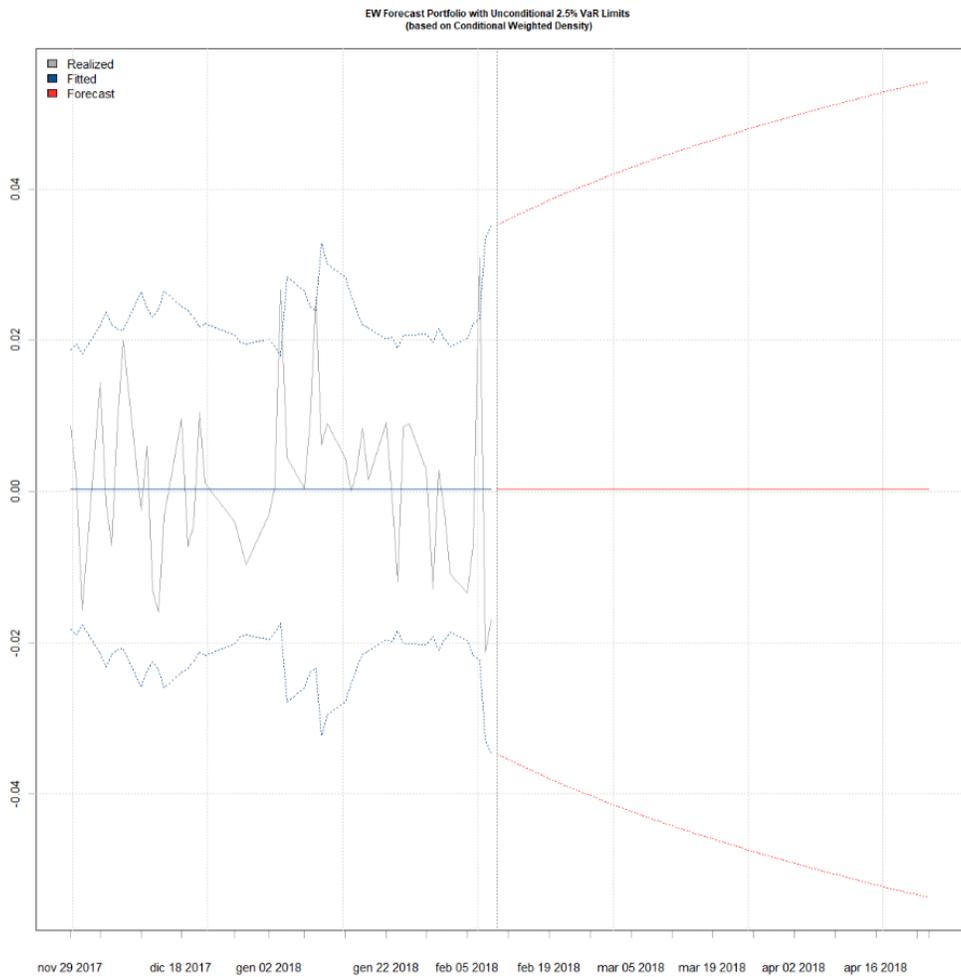


Figura 52: Sezione *Multivariate* della piattaforma; parte inferiore della pagina.

Questa sezione della piattaforma altro non è che l'implementazione dei modelli trattati nel *Capitolo 8* dell'elaborato Bella G. e Lasaponara F.- *Strumenti quantitativi e loro implementazione software per l'analisi di portafogli finanziari, con approfondimenti sui modelli GARCH multivariati*, PoliTo (2018).

## Conclusioni

Sebbene i capitoli teorici non siano sufficienti ad esaurire in modo compiuto il panorama della letteratura prodotta sul tema, sono stati mostrati quegli elementi che riteniamo essere imprescindibili per approcciare dal punto di vista quantitativo l'analisi delle serie temporali finanziarie. In particolare, l'aver optato per uno sviluppo fondato, tra gli altri, sul principio della modularità consentirà agevoli integrazioni future rispetto alla frontiera degli avanzamenti accademici.

Complessivamente, nonostante l'obiettivo del progetto fosse quello di agevolare la fruizione di strumenti quantitativi complessi per un pubblico eterogeneo e non particolarmente specializzato, nel corso della stesura dell'elaborato e durante lo sviluppo e test delle varie componenti applicative, è emersa l'evidenza che non si possa prescindere da una certa knowledge base per l'utilizzo profittevole dello strumento realizzato. Infatti, i modelli trattati e implementati coprono uno spettro molto ampio della letteratura a riguardo delle time series finanziarie che hanno come comune denominatore il tentativo di descrivere matematicamente evidenze puramente empiriche. Per questa ragione, poiché non è possibile identificare un modello univoco che si adatti efficacemente a tutti i casi, è richiesta da parte dell'utente dell'esperienza al fine di trarre le conclusioni corrette a partire dagli output forniti.

In ogni caso, riteniamo che la piattaforma implementata costituisca uno strumento di base insostituibile e l'obiettivo nel prossimo futuro è estenderne le funzionalità, anzitutto integrando la possibilità di costruire un portafoglio virtuale sul quale la piattaforma possa fornire consigli di acquisto/vendita di asset al fine di minimizzare il rischio e, in secondo luogo, aggiungendo una componente di back testing personalizzata in cui l'utente abbia la possibilità di definire autonomamente delle regole di ingresso/uscita dal mercato in funzione dei modelli disponibili sull'applicativo e, eventualmente, avere la possibilità di negoziare tali strategie con altri utenti.

## Riferimenti

### Bibliografia

- T. Andersen e T. Bollerslev - *Answering the Skeptics: Yes, Standard Volatility Models do Provide Accurate Forecast*, International Economic Review, vol.39 n.4 p.885 (1998)
- Artzner P., Delbaen F., Eber J. e Heath D. - *Coherent Measures of Risk*, Mathematical Finance, vol.9 issue 3 p.203-228 (1999)
- Berndt E., Hall B., Hall R., Hausman J. - *Estimation and Inference in Nonlinear Structural Models*, Annals of Economic and Social Measurement, vol.3 n.4 p.653-665 (1974)
- Bharucha-Reid A. T. - *Elements of the Theory of Markov Processes and Their Applications*, McGraw-Hill (1960)
- Black F. - *The pricing of commodity contracts*, Journal of Financial Economics, vol.3 issue 1-2 p.167-179 (1976)
- Bollerslev T. - *Generalized autoregressive conditional heteroskedasticity*, Journal of Econometrics, vol.31 p.307-327 (1986)
- Cornish E. A., Fisher Ronald A. - *Moments and Cumulants in the Specification of Distributions*, Revue de l'Institut International de Statistique, p.307-320 (1938)
- Jon Danielsson - *Financial Risk Forecasting*, Wiley (2011)
- R. Engle e V. Ng - *Measuring and testing the impact of news on volatility*, Journal of Finance, vol.48 n.5 p.1749-1778 (1993)
- Efron B., Tibshirani R. - *An Introduction to the Bootstrap*, Boca Raton, FL: Chapman & Hall (1993)
- Robert F. Engle - *Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation*, Econometrica, vol.50 issue 4 p.987-1007 (1982)
- David A. Freedman - *On the So-Called "Huber Sandwich Estimator" and "Robust Standard Errors"*, The American Statistician, vol.60 n.4 p.299-302 (2005)

- Glasserman P. - *Measuring marginal risk contributions in credit portfolios*, The Journal of Computational Finance, vol.9 p.1-41 (2005)
- Goldberg, Hayes, Menchero e Mitra - *Extreme Risk Management*, MSCI Barra Research, Paper n. 2009-4 (2009)
- Peter Hall - *The Bootstrap and Edgeworth Expansion*, Springer (2013)
- Peter R. Hansen, Asger Lunde - *A Forecast Comparison of Volatility Models: does anything beat a GARCH(1,1)?*, Journal of Applied Econometrics, vol.20 issue 7 p.873-889 (2005)
- Akaike Hirotugu - *A new look at the statistical model identification*, IEEE, vol.19 issue 6 (1974)
- John C. Hull - *Opzioni futures e altri derivati*, Springer (2015)
- John C. Hull - *Risk management e istituzioni finanziari*, Pearson (2008)
- Jensen J. L. W. V. - *Sur les fonctions convexes et les inégalités entre les valeurs moyennes*, Acta Mathematica, vol.30 p.175-193 (1906)
- Eric Jondeau, Ser-Huang Poon e Michael Rockinger - *Financial Modeling Under Non-Gaussian Distributions*, Springer (2000)
- P. Krugman e R. Wells - *Microeconomia*, Zanichelli (2006)
- R. Mariano e Diebold - *Testing forecast accuracy*, Companion to Economic Forecasting, p.284-298 (2002)
- M. Martens - *Measuring and Forecasting S&P 500 Index-Futures Volatility Using High-Frequency Data*, The Journal of Futures Markets, vol.22 n.6 p.497-518 (2002)
- Fredric S. Mishkin, Stanley G. Eakins - *Istituzioni e mercati finanziari*, Pearson (2012)
- Nelson, D. B. - *Conditional Heteroskedasticity in Asset Returns: A New Approach*, Econometrica, vol.59 n.2 p.347-370 (1991)

- Newey, Whitney K; West, Kenneth D - *A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix*, *Econometrica*, vol.55, n.3 p.703-708 (1987)
- A. Pagan e W. Schwert - *Alternative models for conditional stock volatility*, *Journal Of Econometrics*, n.45 p.267-290 (1990)
- E. Pitacco - *Elementi di Matematica delle Assicurazioni*, Edizioni LINT (2002)
- S. Rachev - *Desirable Properties of an Ideal Risk Measure in Portfolio Theory*, *International Journal of Theoretical and Applied Finance*, n.11 p.19-54 (2008)
- David Ruppert e David S. Matteson - *Statistics and Data Analysis for Financial Engineering*, Springer (2015)
- Gideon E. Schwarz - *Estimating the dimension of a model*, *The Annals of Statistics*, vol.6 n.2 p.461-464(1978)
- David Williams - *Probability with Martingales*, Cambridge Mathematical Textbooks (1991)
- Desheng Dash Wu - *Quantitative Financial Risk Management*, Springer (2011)
- See Yamai Y. e T. Yoshiba - *Comparative Analyses of Expected Shortfall and Value-at-Risk*, *Institute for Monetary and Economic Studies*, vol.20 issue 1 p.87-121 (2002)
- Jean-Michele Zakoian - *Threshold heteroskedastic models*, *Journal of Economic Dynamics and Control*, n.18, p.931-955 (1991)

## Sitografia

- Investing ([www.investing.com](http://www.investing.com)) - Fonte dati di bilancio
- Google Finance ([www.google.com/finance](http://www.google.com/finance)) - Fonte serie temporali prezzi titoli quotati
- Borsa Italiana ([www.borsaitaliana.it](http://www.borsaitaliana.it)) - Fonte delle definizioni di ambito finance

## **Software utilizzati**

- R 3.4.2 - Computazione statistica e rendering grafico
- R Server 3.4.2 - Engine della piattaforma
- Angular - Framework Javascript per il front-end della piattaforma
- NGINX - Http Server della piattaforma
- Apache Tomcat - Application server della piattaforma, ambiente di esecuzione degli applicativi Java EE 8
- MySQL - Database relazionale della piattaforma
- MongoDB - Database non relazionale della piattaforma