



UNIVERSITÀ DEGLI STUDI DI PARMA
Facoltà di Economia

Corso di Laurea Magistrale in Finanza e Risk Management
Metodi e Modelli per la Finanza

Distribuzioni Lévy-Pareto stabili
e loro impiego nel modellare
i rendimenti di titoli azionari americani.
Analisi dei cluster.

Tesi di Laurea Magistrale

Relatore Prof. Gino Favero

Correlatore Dott. Riccardo Donati

Laureanda Dott.sa Alice Pisani

Sessione di Laurea 21 - 25 Novembre 2011

Anno Accademico 2010 - 2011

Introduzione		5
1 Rendimenti e distribuzioni di probabilità		9
1.1 Rendimenti		9
1.1.1 Serie Storiche		9
1.1.2 Rendimento percentuale e logaritmico		10
1.2 Distribuzione Gaussiana		11
1.3 Distribuzione Lévy-Pareto stabile		13
1.3.1 Definizione di stabilità		13
1.3.2 Dalla funzione caratteristica alla funzione di densità		14
1.3.3 Parametrizzazione della legge di stabilità		15
1.3.4 Tre casi particolari di distribuzione stabile		17
1.3.5 Somma di variabili aleatorie stabili		18
1.3.6 Decadimento delle code e leptocurtosi		20
1.3.7 Momenti		24
1.3.8 Teorema del limite centrale generalizzato		26
2 Analisi empirica		29
2.1 Raccolta e Manipolazione dei dati		29
2.1.1 Distribuzione temporale		31
2.1.2 Filtro sui rendimenti nulli		32

2.2	Distribution fitting	33
2.2.1	Scelta del modello	33
2.2.2	Stima dei parametri	36
2.2.3	<i>Goodness of fit test</i>	37
2.3	<i>Clustering</i> settoriale	43
2.3.1	Classificazione settoriale	48
2.3.2	Distribuzioni α -settoriali	49
2.3.3	Test di Kruskal-Wallis	53
	Conclusioni	63
	Bibliografia	65

INTRODUZIONE

I rendimenti di attività finanziarie sono interpretabili come il risultato di un vasto numero di informazioni e decisioni individuali che giungono sul mercato in modo quasi continuo nel tempo. Per tale motivo, fin dal pioneristico lavoro di Louis Bachelier nel 1900 [1], le variazioni logaritmiche di prezzo di attività finanziarie sono state modellate utilizzando una distribuzione gaussiana. Statisticamente questa teoria è comprovata dal Teorema del limite centrale, secondo il quale il limite della somma di variabili indipendenti ed identicamente distribuite con varianza finita tende a una distribuzione normale.

L'uso del modello normale è attualmente diffuso in maniera capillare in ambito economico-finanziario, sebbene sia universalmente riconosciuto che i rendimenti presentino code spesse, o caratteristiche *leptocurtiche*, non descrivibili mediante una distribuzione di Gauss.

Mandelbrot [12] e Fama [6], in risposta alle evidenze empiriche riscontrate, hanno proposto un modello alternativo: le distribuzioni *Lévy-Pareto stabili*. Queste ultime sono particolarmente efficaci nel modellare rendimenti causati da eventi estremi quali crolli di mercato o catastrofi naturali, ovvero l'insieme degli accadimenti definiti “*cigni neri*” nell'universo gaussiano. Le code di una distribuzione definiscono le probabilità di fenomeni rari (positivi–coda destra, negativi–coda sinistra): maggiore è lo spessore delle code, maggiore è la probabilità di osservare ampie variazioni dei prezzi che generano rendimenti non usuali.

Nonostante che siano stati sviluppati numerosi modelli econometrici (ARCH, GARCH, ...) e che siano state proposte diverse leggi leptocurtiche alternative alla curva di Gauss (t di Student, iperbolica, ...) esistono almeno due buoni motivi, adeguatamente approfonditi nel seguito, per scegliere la famiglia di leggi stabili quale modello di stima dei rendimenti azionari:

- il Teorema generalizzato del limite centrale, che identifica le distribuzioni stabili quale unico limite per variabili indipendenti ed identicamente distribuite una volta abbandonata l'ipotesi di varianza finita;
- l'invarianza di scala dei rendimenti logaritmici, che trova la sua corrispondenza nella proprietà di stabilità delle distribuzioni stabili.

Il presente elaborato si propone, in prima battuta, di verificare empiricamente la validità del modello Lévy-Pareto stabile. A tal fine verranno analizzate le serie storiche dei rendimenti logaritmici giornalieri di oltre 3000 titoli azionari trattati sui due principali mercati borsistici a livello globale: il New York Stock Exchange (NYSE) e il NASDAQ.

Nel capitolo 1 a pagina 9, dopo una breve introduzione relativa alle serie storiche dei rendimenti e alla ben nota distribuzione di Gauss, si fornisce al lettore una solida base teorica sulle caratteristiche e proprietà della classe di leggi stabili. Una buona conoscenza di tali distribuzioni risulta essenziale per comprendere appieno l'analisi empirica svolta.

Il capitolo 2 a pagina 29 illustra obiettivi, metodo e risultati della ricerca svolta. Il paragrafo 2.1 a pagina 29 mostra la metodologia usata nel processo di raccolta dei dati, le fonti e le trasformazioni propedeutiche alla successiva analisi degli stessi. Nel paragrafo 2.2 a pagina 33 le serie storiche dei rendimenti vengono modellate attraverso l'uso delle distribuzioni Lévy-Pareto stabili e, comparativamente, delle distribuzioni gaussiane; gli esiti del procedimento, definito *distribution fitting*, vengono poi posti a verifica mediante l'impiego di test statistici.

Una volta verificata empiricamente la superiorità del modello stabile quale metodo per approssimare le distribuzioni dei rendimenti azionari, si propone una metodologia alternativa per la stima del parametro α , il cosiddetto *parametro di forma*. Lo

scopo è dimostrare l'esistenza di una connessione forte e univoca tra il valore del parametro e l'appartenenza settoriale della società emittente. Il riscontro statistico di tale legame, mostrato nel paragrafo 2.3 a pagina 43, permette di suddividere il campione in *cluster* α -settoriali, ovvero gruppi di titoli omogenei per valori del parametro di forma.

La ricerca sperimentale condotta, per ampiezza del campione statistico e risultati raggiunti, è nuova in letteratura e fornisce alcuni spunti per applicazioni pratiche nell'ambito del *Risk Management*. In particolare l'esistenza di correlazione tra il settore di operatività dell'azienda e il parametro α permette di costruire stime Lévy-Pareto stabili ragionevolmente affidabili anche per titoli con storia recente o di prima emissione.

CAPITOLO 1

RENDIMENTI E DISTRIBUZIONI DI PROBABILITÀ

1.1 Rendimenti

1.1.1 Serie Storiche

Per *serie storica* si intende la classificazione di un fenomeno rispetto al tempo. Se il fenomeno analizzato è l'insieme dei prezzi di transazione di una attività finanziaria si parla di *serie storica dei prezzi*. I prezzi si formano sul mercato dall'incontro tra domanda ed offerta. Per i titoli quotati su un mercato regolamentato, il numero di compravendite è molto frequente e quasi continuo nel tempo. Ciò è dovuto all'elevato numero di controparti interessate alla vendita, o all'acquisto, di determinate attività finanziarie.

È possibile campionare a intervalli prefissati di tempo ΔT il prezzo dell'ultima transazione del titolo sul mercato, ottenendo una serie storica $\{P_t | t \in \mathbb{N}\}$, ove t indica il numero d'ordine del campione. Per esempio, campionando l'ultimo prezzo disponibile in chiusura di ogni giornata di contrattazione, si ottiene una serie storica giornaliera dei prezzi dell'attività analizzata. Se, al posto dei prezzi dell'attività finanziaria, si considerano i rendimenti, è possibile ricavare, in maniera del tutto analoga, una *serie storica dei rendimenti*.

Affinché l'analisi statistica della serie sia significativa, è necessario che l'attività

finanziaria sia *liquida* sull'orizzonte temporale considerato, ovvero che il numero di transazioni tra due intervalli temporali consecutivi sia elevato.

1.1.2 Rendimento percentuale e logaritmico

Poiché lo scopo di ogni investitore è quello di trarre profitto dalle proprie azioni è naturale introdurre una misura del rendimento ottenibile dall'attività di *trading*. Siano P il prezzo, D i dividendi, t il tempo; il *rendimento percentuale* $R_{t,\Delta}$ calcolato nell'intervallo $t, t + \Delta$ è definito come:

$$R_{t,\Delta} = 100 \left(\frac{P_{t+\Delta} + D_{t+\Delta}}{P_t} - 1 \right).$$

Il rendimento percentuale è la più conosciuta forma di quantificazione dei profitti, tuttavia presenta alcuni limiti. Si pensi, per esempio, a un titolo azionario che perde il 50% e guadagna il 50%. Il rendimento complessivo non è nullo, bensì pari a -25% .

Una misura di rendimento in grado di superare tale limite è il *rendimento logaritmico* $r_{t,\Delta}$: *

$$r_{t,\Delta} = \log(1 + R_{t,\Delta}) = \log \frac{P_{t+\Delta} + D_{t+\Delta}}{P_t} = \log(P_{t+\Delta} + D_{t+\Delta}) - \log P_t. \quad (1.1)$$

Il rendimento logaritmico multiperiodale è la somma dei rendimenti uniperiodali:

$$\begin{aligned} r_{t,t+k} &= \log[1 + R_{t,t+k}] = \log[(1 + R_{t,t+1}) + \cdots + (1 + R_{t+k-1,t+k})] \\ &= r_{t,t+1} + r_{t+1,t+2} + \cdots + r_{t+k-1,t+k} = \sum_{i=1}^k r_{t+i}, \end{aligned} \quad (1.2)$$

inoltre, se R_t tende a zero, il rendimento logaritmico tende al rendimento percentuale

* $\log P_t$ è definito prezzo logaritmico, o *log-prezzo*.

(a meno di un fattore 100):

$$\begin{aligned} r_{t,\Delta} &= \log \frac{P_{t+\Delta} + D_{t+\Delta}}{P_t} = \log \left[1 + \left(\frac{P_{t+\Delta} + D_{t+\Delta}}{P_t} - 1 \right) \right] = \\ &= 1 + \left(\frac{P_{t+\Delta} + D_{t+\Delta}}{P_t} - 1 \right) + o\left(\frac{P_{t+\Delta} + D_{t+\Delta}}{P_t} - 1 \right) \\ &= \frac{1}{100} [R_{t,\Delta} + o(R_{t,\Delta})]. \end{aligned}$$

1.2 Distribuzione Gaussiana

I rendimenti finanziari possono essere considerati come il risultato di un vasto numero di informazioni e decisioni individuali che arrivano sui mercati in modo continuo nel tempo: matematicamente parlando, le variazioni dei prezzi osservate sono manifestazioni di una variabile aleatoria, alle quali può essere associata una probabilità di accadimento.

Per modellare il comportamento di una variabile casuale, e poterne studiare le caratteristiche, viene comunemente utilizzata una distribuzione di probabilità.

Una *distribuzione di probabilità* è un modello matematico che collega i valori di una variabile aleatoria alle probabilità che tali valori possano essere osservati. Formalmente, le distribuzioni di probabilità vengono espresse da una legge matematica detta funzione di densità di probabilità (PDF) definita come l'applicazione $f(x)$ non negativa e integrabile tale che:

$$\Pr\{a < X < b\} = \int_a^b f(x)dx, \quad \text{con } a, b \in \Omega.$$

Sia $X : \Omega \rightarrow \mathbb{R}$ una variabile aleatoria. La funzione:

$$\begin{aligned} F(x) &= \Pr\{X(\omega) \leq x\} \\ &= \int_{-\infty}^x f(t)dt. \end{aligned}$$

è chiamata *funzione di ripartizione* (DF) di X .

La distribuzione *normale* fu proposta inizialmente da Carl Gauss per modellare gli errori casuali di misurazione. Tale distribuzione è usualmente indicata $\mathbf{N}(\mu; \sigma^2)$

e la sua PDF è rappresentata graficamente dalla nota curva “a campana” di Gauss. Il parametro μ rappresenta la media, ed il parametro σ^2 la varianza della distribuzione. In particolare, se $\mu = 0$ e $\sigma^2 = 1$, $\mathbf{N}(0; 1)$ è detta distribuzione *normale standardizzata*.

In letteratura si parla spesso di log-normalità dei prezzi, ovvero le variazioni dei prezzi logaritmici (1.1) tendono, per intervalli di tempo sufficientemente ampi (un giorno, una settimana, ecc...), a una distribuzione di Gauss. Tale affermazione, basata sul Teorema del limite centrale (1.10), richiede che il numero di transazioni sia uniforme nel tempo e che i rendimenti siano indipendenti e identicamente distribuiti (*iid*), con varianza finita.

Una variabile aleatoria continua X si distribuisce normalmente con parametri μ e σ^2 se la sua funzione di densità è:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in \mathbb{R}.$$

La sua funzione di distribuzione è:

$$F(x) = \int_{-\infty}^x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt.$$

La Gaussiana è considerata il caso base (da qui il nome “normale”) delle distribuzioni di probabilità continue, a causa del suo ruolo nel Teorema del limite centrale. Più specificamente la somma di n variabili casuali con media e varianza finite tende a una distribuzione normale al tendere di n all’infinito.

Questo teorema, insieme con le ipotesi di rendimenti *iid*, varianza finita e uniformità temporale delle transazioni, giustifica l’uso della distribuzione normale come approssimazione della distribuzione dei rendimenti nei principali modelli finanziari, quale, per esempio, il modello di Black, Scholes e Merton.

Le distribuzioni empiriche dei rendimenti, tuttavia, differiscono dalla distribuzione gaussiana, presentando *leptocurtosi*, ovvero code più pesanti e valori centrali più frequenti di distribuzioni normali con la medesima varianza. Tale fenomeno è stato descritto per la prima volta in letteratura negli studi di Mandelbrot [12], e da allora

è stato ampiamente confermato.

Il momento quarto standardizzato (cioè indipendente da scala e locazione), o coefficiente di *curtosi*,

$$\mathbb{K}[X] = \frac{\mathbb{E}[(X - \mu)^4]}{\sigma^4},$$

con μ e σ media e deviazione standard di X , è utilizzato per misurare il grado di leptocurtosi di una data distribuzione.

Per la distribuzione normale, $\mathbb{K} = 3$; $\mathbb{K} > 3$ implica un *eccesso di curtosi* nella distribuzione, considerato come indicatore di una forma leptocurtica.

Nel paragrafo seguente verrà presentata una famiglia di distribuzioni leptocurtiche, adatte a modellare i rendimenti finanziari.

1.3 Distribuzione Lévy-Pareto stabile

1.3.1 Definizione di stabilità

Le distribuzioni *Lévy-Pareto stabili*, definite anche *stabili* o α -*stabili*, sviluppate da Paul Lévy, sono una famiglia di distribuzioni di probabilità caratterizzate da quattro parametri: α , β , γ e δ .

La caratteristica di tale famiglia di distribuzioni è la proprietà di *stabilità*: se più variabili aleatorie casuali, indipendenti e identicamente distribuite seguono una distribuzione stabile, allora anche una combinazione lineare di tali variabili seguirà la medesima legge di distribuzione. Una variabile aleatoria X si definisce *stabile* se per ogni X_1, X_2 , copie indipendenti di X è verificata la seguente uguaglianza:

$$\begin{aligned} aX_1 + bX_2 &\stackrel{d}{=} cX + d, \\ a, b, c, d &\in \mathbb{R}, \\ a, b, c &> 0, \end{aligned} \tag{1.3}$$

dove il simbolo $\stackrel{d}{=}$ indica uguaglianza in distribuzione, ovvero le variabili in esame seguono la medesima legge di probabilità. La variabile aleatoria è detta *stabile in senso stretto* se l'equazione (1.3) è verificata con costante $d = 0$; X è

simmetricamente stabile se è stabile e è simmetricamente distribuita attorno allo 0, ovvero $X \stackrel{d}{=} -X$.

Il termine “stabile” è utilizzato perché la forma di tali distribuzioni è stabile o invariante per somma del tipo (1.3). La distribuzione normale, la distribuzione di Cauchy e quella di Lévy verificano la condizione di stabilità e sono pertanto considerate casi particolari della famiglia delle distribuzioni stabili. †

Le distribuzioni appartenenti alla famiglia stabile sono inoltre *infinitamente divisibili*. Ciò significa che, se X è una variabile aleatoria con distribuzione stabile F , allora per ogni intero positivo n esistono n variabili indipendenti e identicamente distribuite X_1, X_2, \dots, X_n la cui somma è uguale a X in distribuzione.

1.3.2 Dalla funzione caratteristica alla funzione di densità

La funzione di densità di una distribuzione stabile non è generalmente esprimibile analiticamente (se non per le distribuzioni normale, di Cauchy e di Lévy). È tuttavia possibile ricavarla a partire dalla *funzione caratteristica*.

Per una variabile aleatoria X con funzione di distribuzione $F(x)$, e funzione di densità $f(x)$, la funzione caratteristica (CF) è definita da:

$$\varphi_X(t) = \mathbb{E}[e^{itX}] = \int_{-\infty}^{+\infty} e^{itX} dF(x) = \int_{-\infty}^{+\infty} e^{itX} f(x) dx.$$

La funzione $\varphi_X(t)$ caratterizza in modo univoco la distribuzione di X , nel senso che se due distribuzioni di probabilità hanno la stessa CF la loro densità di probabilità è identica.

La CF è detta anche *trasformata di Fourier* della funzione $f(x)$. Le proprietà che contraddistinguono una funzione di densità implicano che la funzione caratteristica esista sempre per qualsiasi variabile casuale considerata. Inoltre, se è nota la funzione caratteristica, è possibile invertirla riottenendo da essa la $f(x)$ attraverso la

† Con il termine *distribuzione di Lévy* si intende indicare la particolare distribuzione con parametri $\alpha = 0.5$ e $\beta = 1$, appartenente alla classe generale delle distribuzioni Lévy Pareto stabili definite nel seguito.

trasformata inversa di Fourier:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \varphi_X(t) e^{-ixt} dt.$$

Ampliando la definizione di stabilità (1.3), una variabile aleatoria X è stabile se e solo se:

$$X \stackrel{d}{=} \gamma Z + \delta, \quad (1.4)$$

con $\gamma > 0$, $\delta \in \mathbb{R}$. La variabile aleatoria Z ha funzione caratteristica:

$$\varphi_Z(t) = \mathbb{E}[e^{itZ}] := \begin{cases} \exp(-|t|^\alpha [1 - i\beta \tan \frac{\pi\alpha}{2} \text{sign}(t)]), & \text{se } \alpha \neq 1, \\ \exp(-|t|[1 + i\beta \frac{2}{\pi} \text{sign}(t) \log|t|]), & \text{se } \alpha = 1, \end{cases} \quad (1.5)$$

con $0 < \alpha \leq 2$, $-1 \leq \beta \leq 1$. La funzione $\text{sign}(t)$ è definita come segue: ‡

$$\text{sign}(t) := \begin{cases} -1, & \text{se } t < 0, \\ 0, & \text{se } t = 0, \\ 1, & \text{se } t > 0. \end{cases}$$

1.3.3 Parametrizzazione della legge di stabilità

Come indicato nell'equazione (1.5), una distribuzione stabile viene descritta utilizzando quattro parametri: un *indice o esponente di stabilità* $\alpha \in (0, 2]$, un *parametro di asimmetria* $\beta \in [-1, 1]$, un *parametro di scala* $\gamma > 0$ e un *parametro di locazione* $\delta \in \mathbb{R}$. La distribuzione stabile sarà pertanto indicata con $\mathbf{S}(\alpha, \beta, \gamma, \delta)$. Seguendo la notazione proposta da Nolan [16] verranno utilizzati i parametri γ per la scala e δ per la locazione, al fine di evitare confusione con i simboli σ e μ , i quali saranno usati esclusivamente per *standard deviation* e media della distribuzione.

Per la famiglia di distribuzioni stabili infatti, il parametro di scala γ non rappresenta la deviazione standard (neppure nel caso normale) e il parametro di locazione δ non indica generalmente la media.

‡Nella espressione (1.5) per il caso $\alpha = 1$, $0 \cdot \log 0$ è interpretato come 0.

Un possibile problema con la parametrizzazione della distribuzioni stabili è la molteplicità di definizioni dei parametri γ e δ , i quali caratterizzano la relazione tra X e Z nella equazione (1.4). Per distinguere la parametrizzazione scelta verrà utilizzato un numero intero $k = (0, 1, 2, \dots)$. In letteratura le parametrizzazioni più comuni sono del tipo $\mathbf{S}(\alpha, \beta, \gamma, \delta; k = 0)$ e $\mathbf{S}(\alpha, \beta, \gamma, \delta; k = 1)$, figura 1.1 a pagina 19.

I parametri α , β e γ hanno lo stesso significato nelle due parametrizzazioni, mentre il parametro di locazione δ è definito diversamente. In particolare, indicando con δ il parametro nel caso $k = 1$ e δ_0 nel caso $k = 0$, la relazione fra i due parametri risulta:

$$\delta_0 := \begin{cases} \delta + \gamma\beta \tan \frac{\pi\alpha}{2}, & \text{se } \alpha \neq 1, \\ \delta + \beta^{\frac{2}{\pi}}\gamma \log \gamma, & \text{se } \alpha = 1. \end{cases}$$

Se $\beta = 0$ le due parametrizzazioni coincidono. Quando $\beta \neq 0$ i valori di δ e δ_0 nelle due parametrizzazioni differiscono per un fattore pari a $\gamma\beta \tan \frac{\pi\alpha}{2}$, il quale tende a infinito per $\alpha \rightarrow 1$.

In particolare, la moda della densità di una distribuzione $\mathbf{S}(\alpha, \beta, \gamma, \delta; 1)$ diverge a $+\infty$ (se $\text{sign}(\alpha - 1)\beta > 0$), oppure a $-\infty$ (altrimenti), quando $\alpha \rightarrow 1$.[§]

Nel corso delle analisi di seguito presentate verrà utilizzata la parametrizzazione $k = 1$, poiché, nonostante gli svantaggi illustrati sopra, gode di semplici proprietà algebriche. Inoltre, nel caso specifico dei dati empirici studiati, α risulta sempre ben lontano da 1, eludendo così eventuali problemi di instabilità numerica nella stima dei parametri. Una variabile aleatoria X è $\mathbf{S}(\alpha, \beta, \gamma, \delta; 1)$ se:

$$X \stackrel{d}{=} \begin{cases} \gamma Z + \delta, & \text{se } \alpha \neq 1, \\ \gamma Z + (\delta + \beta^{\frac{2}{\pi}}\gamma \log \gamma), & \text{se } \alpha = 1. \end{cases}$$

con $Z = Z(\alpha, \beta)$, come definita nell'equazione (1.5). La funzione caratteristica di

[§]La *moda* di una distribuzione di probabilità continua è il valore x che massimizza la funzione di densità; in modo informale, la moda coincide con il picco più alto nella curva della distribuzione.

X è:

$$\varphi_X(t) := \begin{cases} \exp(-\gamma^\alpha |t|^\alpha [1 - i\beta \tan \frac{\pi\alpha}{2} \text{sign}(t)] + i\delta t), & \text{se } \alpha \neq 1, \\ \exp(-\gamma |t| [1 + i\beta \frac{2}{\pi} \text{sign}(t) \log|t|] + i\delta t), & \text{se } \alpha = 1. \end{cases} \quad (1.6)$$

Applicando la trasformata inversa di Fourier (1.3.2) alla funzione caratteristica, così come calcolata nell'equazione (1.6), è possibile derivare la funzione di densità di una $\mathbf{S}(\alpha, \beta, \gamma, \delta; 1)$:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \varphi_X(t | \alpha, \beta, \gamma, \delta) e^{-ixt} dt.$$

Il caso $\alpha = 1$ verrà tralasciato nella trattazione che segue, perché non risulta significativo nello studio della distribuzione dei rendimenti di attività finanziarie. In particolare i log-rendimenti empirici si distribuiscono generalmente con parametro di stabilità $1 < \alpha < 2$. Le definizioni matematiche enunciate nel testo che segue saranno, pertanto, esclusivamente relative al caso $\alpha > 1$. Per approfondimenti si rimanda a [16].

1.3.4 Tre casi particolari di distribuzione stabile

Come già accennato, la funzione di densità di una distribuzione stabile non è generalmente esprimibile analiticamente, ovvero non è sempre possibile calcolare l'integrale della (1.3.3). L'equazione può essere risolta esclusivamente per alcuni valori specifici dei parametri $\alpha, \beta, \gamma, \delta$.

Esistono tre casi nei quali è possibile scrivere la PDF in forma chiusa e verificarne direttamente la stabilità: la distribuzione normale, di Cauchy e di Lévy (figura 1.2 a pagina 19).

Esempio 1. (*Distribuzione normale*). $X \sim \mathbf{N}(\mu, \sigma^2)$ se ha densità

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in \mathbb{R}.$$

La distribuzione di Gauss è stabile con parametri ($\alpha = 2, \beta = 0, \gamma = \frac{\sigma}{\sqrt{2}}, \delta = \mu$).

Più precisamente: $\mathbf{N}(\mu, \sigma^2) = \mathbf{S}(2, 0, \frac{\sigma}{\sqrt{2}}, \mu; 1)$.

Esempio 2. (*Distribuzione di Cauchy*). $X \sim \mathbf{C}(a, b)$ se ha densità

$$f(x) = \frac{1}{\pi} \cdot \frac{a}{a^2 + (x - b)^2}, \quad x \in \mathbb{R}.$$

La distribuzione di Cauchy è stabile con parametri ($\alpha = 1, \beta = 0, \gamma = a, \delta = b$).

Più precisamente: $\mathbf{C}(a, b) = \mathbf{S}(1, 0, a, b; 1)$.

Esempio 3. (*Distribuzione di Lévy*). $X \sim \mathbf{L}(a, b)$ se ha densità

$$f(x) = \sqrt{\frac{a}{2\pi}} \cdot \frac{1}{(x - b)^{\frac{3}{2}}} \cdot e^{-\frac{a}{2(x-b)}}, \quad b < x < \infty. \quad (1.7)$$

La distribuzione di Lévy è stabile con parametri ($\alpha = \frac{1}{2}, \beta = 1, \gamma = a, \delta = b$). Più precisamente $\mathbf{L}(a, b) = \mathbf{S}(\frac{1}{2}, 1, a, b; 1)$.

1.3.5 Somma di variabili aleatorie stabili

La stabilità così come illustrata nell'equazione (1.3) implica l'invarianza per addizione, cioè la somma di variabili α -stabili è anch'essa una variabile α -stabile. Affinché tale proprietà valga, è necessario che le variabili sommate abbiano lo stesso valore di α . Il valore dei parametri risultanti dalla somma dipende dalla parametrizzazione scelta. In particolare la parametrizzazione $\mathbf{S}(\alpha, \beta, \gamma, \delta; 1)$, gode delle seguenti proprietà:

1. Se $X \sim \mathbf{S}(\alpha, \beta, \gamma, \delta; 1)$, allora per ogni $a \neq 0, b \in \mathbb{R}$,
 $aX + b \sim \mathbf{S}(\alpha, [\text{sign}(a)\beta], |a|\gamma, a\delta + b; 1)$.
2. La funzione caratteristica, la funzione di densità e la funzione di distribuzione sono continue.

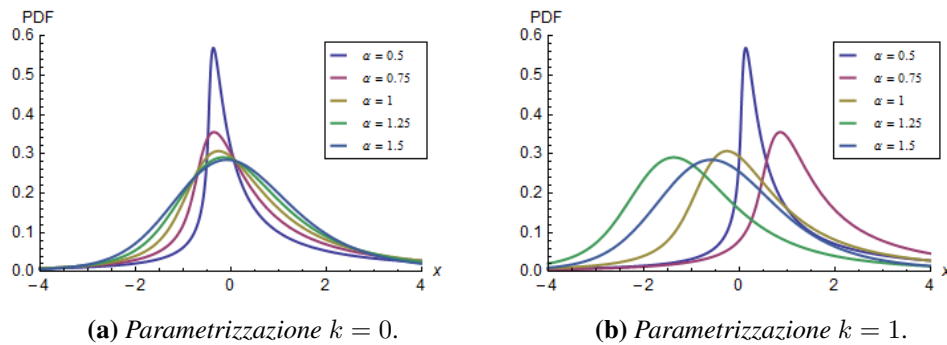


Figura 1.1. A confronto la parametrizzazione $k = 1$, e la parametrizzazione $k = 0$, con parametri $\beta = 0.5$, $\delta = 0$, $\gamma = 1$ e diversi parametri α , .

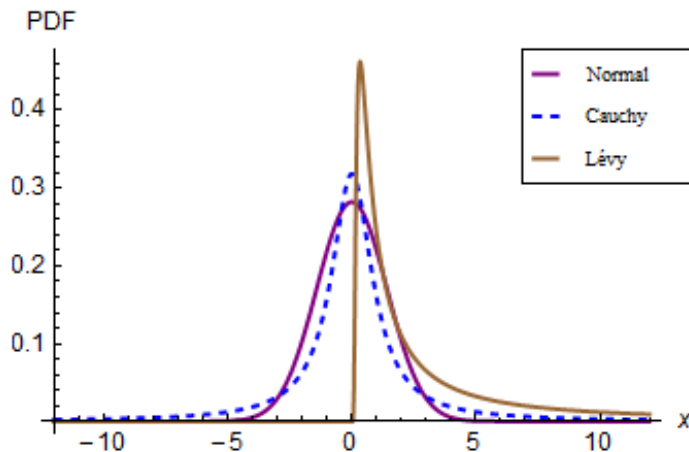


Figura 1.2. Tre casi particolari di distribuzione stabile.

Distribuzione Gaussiana $\sim \mathbf{S}(2, 0, \frac{\sigma}{\sqrt{2}}, \mu; 1)$, di Cauchy $\sim \mathbf{S}(1, 0, a, b; 1)$ e di Lévy $\sim \mathbf{S}(0.5, 1, a, b; 1)$. La distribuzione normale e la distribuzione di Cauchy sono simmetriche con curva “a campana”. La seconda ha tuttavia code più pesanti, in particolare $\Pr\{X > |3|\}$ è 100 volte superiore in una distribuzione di Cauchy rispetto a una normale. Questa è la ragione per cui le distribuzioni stabili sono definite leptocurtiche o *heavy tailed*. La distribuzione di Lévy è fortemente asimmetrica, con distribuzione concentrata esclusivamente per valori $X > 0$, e presenta code ancora più pesanti rispetto alla distribuzione di Cauchy.

3. Se $X_1 \sim \mathbf{S}(\alpha, \beta_1, \gamma_1, \delta_1; 1)$, e $X_2 \sim \mathbf{S}(\alpha, \beta_2, \gamma_2, \delta_2; 1)$ sono indipendenti, allora $X_1 + X_2 \sim \mathbf{S}(\alpha, \beta, \gamma, \delta; 1)$ con

$$\beta = \frac{\beta_1 \gamma_1^\alpha + \beta_2 \gamma_2^\alpha}{\gamma_1^\alpha + \gamma_2^\alpha}, \quad \gamma^\alpha = \gamma_1^\alpha + \gamma_2^\alpha, \quad \delta = \delta_1 + \delta_2.$$

La formula utilizzata per γ^α è la generalizzazione della regola di addizione delle varianze di variabili casuali indipendenti: $\sigma^2 = \sigma_1^2 + \sigma_2^2$.

Il punto 1 e il punto 3 affermano che il parametro di locazione di una trasformazione lineare coincide con la trasformazione lineare del parametro δ e il parametro di locazione di una somma è la somma dei parametri di locazione.

Come conseguenza delle proprietà sopra elencate (combinazioni lineari di variabili aleatorie stabili), è possibile affermare che X è strettamente stabile se e solo se $\delta = 0$. Qualsiasi grado di asimmetria (β) può dare stabilità in senso stretto, purché $\delta = 0$. Questo equivale a dire che qualsiasi distribuzione stabile può essere resa strettamente stabile mediante uno spostamento del parametro di locazione.

1.3.6 Decadimento delle code e leptocurtosi

La figura 1.3 nella pagina successiva illustra l'impatto dei parametri α e β sulla forma della distribuzione. Le distribuzioni stabili permettono di modellare asimmetria e leptocurtosi. In particolare β controlla l'asimmetria, mentre α è responsabile del decadimento delle code della distribuzione. Per tale motivo i due parametri, α e β sono anche detti *parametri di forma*.

Nel caso gaussiano, $X \sim \mathbf{N}(0, 1)$, esiste una formula ben nota per approssimare la probabilità nelle code, quando $x \rightarrow \infty$:[¶]

$$\Pr\{X > x\} \sim \frac{e^{-\frac{x^2}{2}}}{x\sqrt{2\pi}}.$$

Le code di una distribuzione normale presentano quindi un decadimento di tipo esponenziale. Nel caso non gaussiano, ovvero quando $\alpha < 2$, Lévy (1925) ha

[¶]Nell'equazione (1.3.6) e seguenti: $h(x) \sim g(x)$ quando $x \rightarrow \infty$, significa $\lim_{x \rightarrow \infty} h(x)/g(x) = 1$.

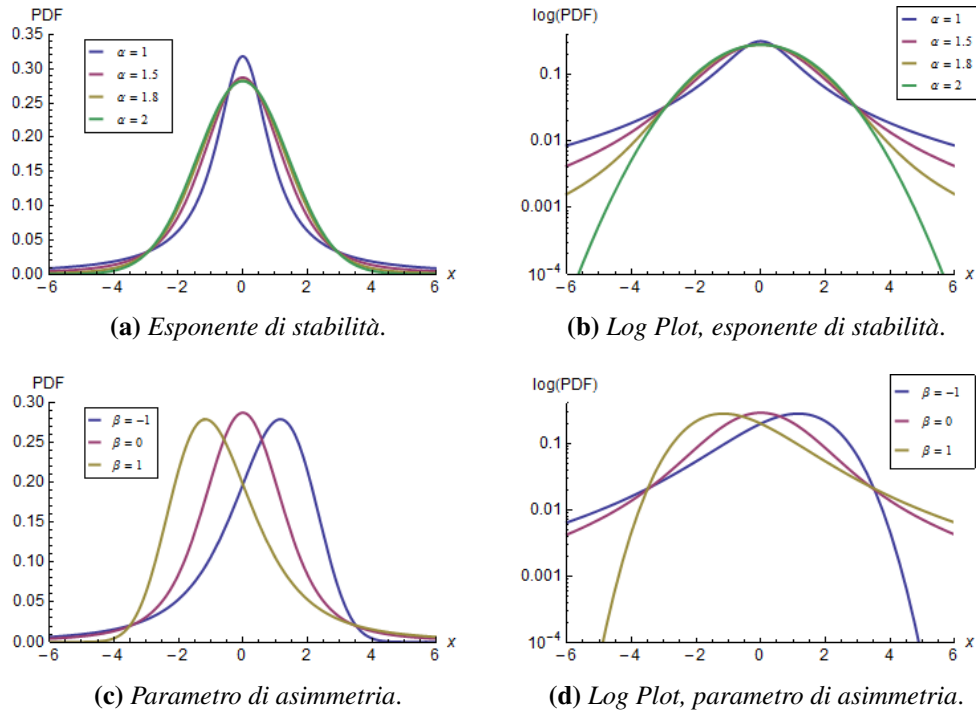


Figura 1.3. Dipendenza della distribuzione dai *parametri di forma*.

(a): funzione di densità (PDF) di una distribuzione stabile standard e simmetrica attorno allo 0 ($\gamma = 1$, $\delta = \beta = 0$) per diversi valori del parametro α . Tale parametro determina il grado di leptocurtosi della distribuzione; è possibile infatti osservare che più l'esponente è piccolo, più la distribuzione si concentra attorno al valor medio e presenta contemporaneamente maggiori probabilità nelle code. Si ricorda che per $\alpha = 2$ la distribuzione è normale, $X \sim \mathbf{N}(\mu, \sigma^2)$.

(c): distribuzione stabile standard con esponente di stabilità $\alpha = 1.2$ e parametri di asimmetria $\beta = -1$, $\beta = 0$, $\beta = 1$. Quando $\beta = 1$, la distribuzione è asimmetrica verso destra, viceversa quando $\beta = -1$, la PDF tende verso sinistra.

(b),(d): *log plot* delle densità di distribuzione sopraccitate, le curve sono generate considerando il logaritmo della funzione di densità, per meglio visualizzare il comportamento delle code. I valori indicati nel grafico (asse delle ordinate), si riferiscono ai valori originari della PDF.

dimostrato che le code di una distribuzione stabile sono asintoticamente equivalenti a una distribuzione di Pareto.

Distribuzione di Pareto

Se X è una variabile aleatoria continua con distribuzione di *Pareto* (del primo tipo), allora la probabilità che X sia maggiore di un qualsiasi numero x è data da:

$$\Pr\{X > x\} := \begin{cases} \left(\frac{x}{k}\right)^{-\alpha}, & \text{per } x \geq k, \\ 1, & \text{per } x < k, \end{cases}$$

dove $k > 0$ rappresenta il valore minimo possibile di X , e $\alpha > 0$ è il parametro di forma. La funzione di densità è

$$f(x) := \begin{cases} k^\alpha \alpha x^{-(\alpha+1)}, & \text{per } x \geq k, \\ 0, & \text{per } x < k. \end{cases}$$

Una distribuzione paretiana è una distribuzione a *legge di potenza* (o *power law*), che nella sua forma generale è descritta dalla relazione:

$$f(x) \propto L(x)x^{-\alpha}, \quad (1.8)$$

dove \propto indica proporzionale, cioè uguale a meno di un fattore moltiplicativo, e $\alpha > 1$. $L(x)$ è una funzione che “varia lentamente” (*slowly varying function*), cioè una qualsiasi funzione tale per cui:

$$\lim_{x \rightarrow \infty} \frac{L(tx)}{L(x)} = 1, \quad t \in \mathbb{R}.$$

Se $L(x)$ è una costante, allora la relazione diventa: $f(x) = Cx^{-\alpha}$.

Una proprietà importante delle leggi di potenza è l'*invarianza di scala*,

$$f(tx) = C(tx)^{-\alpha} = t^{-\alpha}f(x) \propto f(x) :$$

cambiando la scala o l'unità di misura della variabile x la distribuzione di probabilità rimane invariata a meno di una costante moltiplicativa t .

È possibile dimostrare che l'invarianza di scala è condizione necessaria e sufficiente affinché una distribuzione sia *power law*.

Solitamente le distribuzioni di Pareto vengono rappresentate su di un grafico "log-log", ovvero un piano cartesiano in cui entrambe le variabili sugli assi sono misurate in scala logaritmica. Infatti, $\log f(x) = \log(C) - \alpha \log(x)$: è una retta con intercetta $\log(C)$ e coefficiente angolare $-\alpha$. Per chiarezza interpretativa, nei grafici logaritmici presentati verranno indicati sugli assi i valori originari delle variabili (figura 1.4 a pagina 25).

Approssimazione delle code

Per approssimare il decadimento delle code nella famiglia delle leggi di probabilità stabili verrà utilizzata una $\mathbf{S}(\alpha, \beta, 1, 0; 1)$, ovvero una distribuzione *stabile standardizzata* con varianza unitaria e media nulla. Il punto di interesse, infatti, è la forma delle code della distribuzione, e non la locazione delle stesse. Quindi, se X è una variabile aleatoria standardizzata con indice di stabilità $0 < \alpha < 2$, e β parametro di asimmetria, è possibile dimostrare che per $x \rightarrow \infty$,

$$\begin{aligned} F(x) = P\{X > x\} &\sim C_\alpha(1 + \beta)x^{-\alpha}, \\ f(x \mid \alpha, \beta, 1, 0; 1) &\sim \alpha C_\alpha(1 + \beta)x^{-(\alpha+1)}, \end{aligned} \quad (1.9)$$

dove \parallel

$$C_\alpha = \left(2 \int_0^\infty x^{-\alpha} \sin x dx\right)^{-1} = \frac{1}{\pi} \sin\left(\frac{\pi\alpha}{2}\right) \Gamma(\alpha).$$

Inoltre, poiché per ogni α e β , $Z \sim \mathbf{S}(\alpha, \beta; k)$, $k = 0, 1, 2$, vale l'uguaglianza (*reflection property*):

$$Z(\alpha, -\beta) \stackrel{d}{=} -Z(\alpha, \beta),$$

è possibile affermare che, per $x \rightarrow \infty$, la coda sinistra della distribuzione sarà

\parallel Funzione Gamma di Eulero: $\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt$. Si dimostra inoltre che $\Gamma(x+1) = x!$, per $x \in \mathbb{N}$, cioè per tutti i numeri naturali (interi e non negativi).

approssimata come segue:

$$\begin{aligned} F(-x) = P\{X < -x\} &\sim C_\alpha(1 - \beta)x^{-\alpha}, \\ f(-x \mid \alpha, \beta, 1, 0; 1) &\sim \alpha C_\alpha(1 - \beta)x^{-(\alpha+1)}. \end{aligned}$$

Per ogni $\alpha < 2$ e $-1 < \beta < 1$, entrambe le code, e le densità di probabilità seguono asintoticamente una legge di potenza, così come descritta nell'equazione (1.8). Quando $\beta = -1$, la coda destra della distribuzione non segue una legge di potenza; allo stesso modo, quando $\beta = 1$, la coda sinistra non decade con legge di potenza.

Il termine distribuzioni *Pareto stabili* è utilizzato per distinguere tra il decadimento veloce (esponenziale) delle code di una distribuzione normale e il decadimento *power law* delle code di una distribuzione stabile nel caso $\alpha < 2$ (figura 1.5 nella pagina successiva).

1.3.7 Momenti

Una conseguenza della leptocurtosi, ovvero della presenza di code pesanti, è la mancanza di alcuni momenti (p). Generalmente per descrivere una distribuzione vengono utilizzati il primo e il secondo momento, vale a dire la media (μ), e la varianza (σ^2).

Tuttavia nel caso delle distribuzioni stabili è possibile dimostrare che per $0 < \alpha < 2$, $\mathbb{E}[|X|^p]$ esiste finito per $0 < p < \alpha$, mentre $\mathbb{E}[|X|^p] = +\infty$ per $p \geq \alpha$. Perciò, la varianza delle distribuzioni stabili risulta infinita (tranne che nel caso gaussiano, ovvero per $\alpha = 2$), mentre la media esiste finita per ogni $1 < \alpha \leq 2$. Per una variabile aleatoria $X \sim \mathbf{S}(\alpha, \beta, \gamma, \delta; 1)$, la media è pari a $\mu = \mathbb{E}[X] = \delta$.

È necessario sottolineare che il parametro di asimmetria β non coincide con il “classico” parametro di misura dell'asimmetria di una distribuzione (*skewness*). Quest'ultimo, definito come:

$$\zeta[X] = \frac{\mathbb{E}[(X - \mu)^3]}{\sigma^3},$$

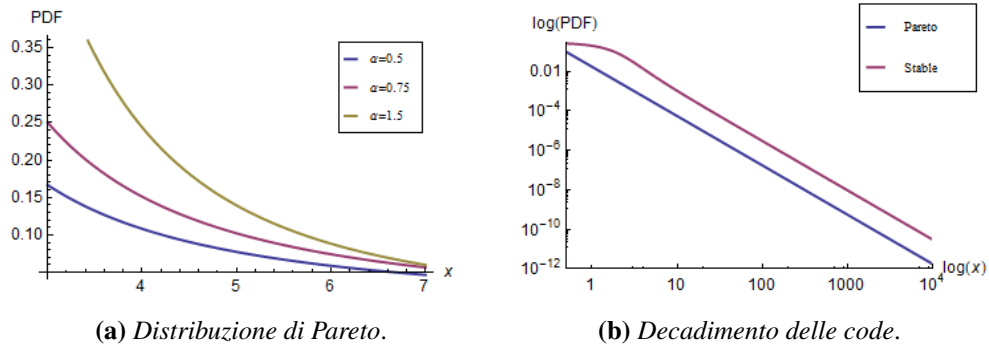


Figura 1.4. (a): distribuzione di Pareto, con $k = 3$ e parametri $\alpha = 0.5$, $\alpha = 0.75$, $\alpha = 1.5$. (b): grafico “log-log” delle densità di probabilità di una distribuzione stabile standard simmetrica e di una distribuzione di Pareto entrambe con esponente di stabilità $\alpha = 1.5$. Le distribuzioni stabili, così come le distribuzioni paretiane presentano un decadimento delle code che segue una legge di potenza. In particolare: $f(x) = k^\alpha \alpha x^{-(\alpha+1)}$, per $x \geq k$ nel caso di una distribuzione di Pareto; mentre $f(x | \alpha, 0, 1, 0) \sim \alpha C_\alpha x^{-(\alpha+1)}$ per $x \rightarrow \infty$ nel caso di una distribuzione stabile. Le code delle due distribuzioni presentano effettivamente comportamento asintotico simile e tipico di una distribuzione *power law*. In particolare, il valore della costante k nella distribuzione paretiana è stato scelto affinché fosse verificata l’uguaglianza: $k^\alpha = C_\alpha$.

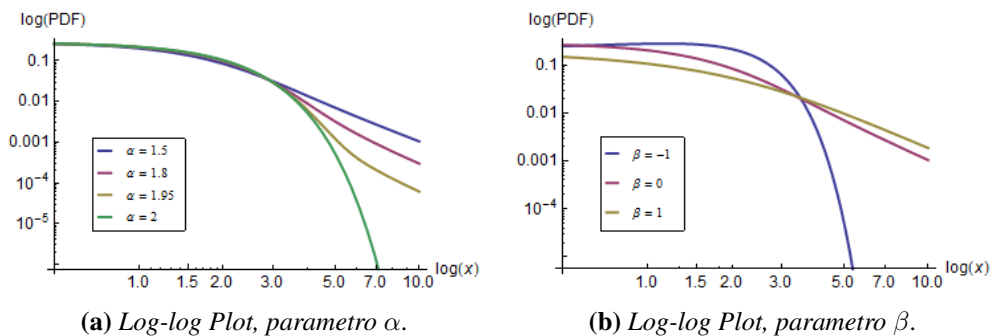


Figura 1.5. Nel *log-log plot* le curve sono generate dai logaritmi della funzione di densità e x , tuttavia sugli assi sono indicati i valori originari delle variabili (x , PDF). Nel grafico sono rappresentate le code positive (per $x > 0.5$). Per $\alpha < 2$ e $-1 < \beta < 1$, entrambe le code di una distribuzione stabile seguono asintoticamente una legge di potenza del tipo $f(x) = Cx^{-\alpha}$. Utilizzando un grafico “log-log” la relazione diventa lineare: $\log f(x) = \log(C) - \alpha \log(x)$, con intercetta $\log(C)$ e coefficiente angolare $-\alpha$. Quando $\alpha = 2$, la distribuzione è normale, $X \sim \mathbf{N}(\mu, \sigma^2)$ e il decadimento delle code è più veloce, rispetto a una *power law*, (a). Quando $\beta = -1$, la coda destra della distribuzione non segue una legge di potenza, allo stesso modo quando $\beta = 1$, la coda sinistra non decade secondo una legge di potenza, (b).

risulta indeterminato per ogni distribuzione stabile non gaussiana, poiché né il momento terzo né la varianza esistono. Parimenti anche l'indice di curtosi risulta indefinito per ogni distribuzione stabile non normale.

1.3.8 Teorema del limite centrale generalizzato

Il classico *Teorema del limite centrale* (TLC) afferma che la somma normalizzata di termini indipendenti e identicamente distribuiti con varianza finita converge a una distribuzione normale. Più precisamente, siano $X_1, X_2, X_3 \dots$, variabili aleatorie *iid*, con media μ , varianza σ^2 , e media campionaria $\bar{X} = (X_1 + \dots + X_n)/n$,

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \xrightarrow{d} Z \sim \mathbf{N}(0, 1), \quad \text{per } n \rightarrow \infty. \quad (1.10)$$

Utilizzando la notazione del Teorema generalizzato, la (1.10), può essere riscritta come

$$a_n(X_1 + \dots + X_n) - b_n \xrightarrow{d} Z \sim \mathbf{N}(0, 1), \quad \text{per } n \rightarrow \infty, \quad (1.11)$$

dove $a_n = 1/\sigma\sqrt{n}$ e $b_n = \sqrt{n}\mu/\sigma$.

Il TLC presuppone che la varianza esista e sia finita, poiché compare al denominatore di entrambe le equazioni sopra. Come già accennato, per le distribuzioni stabili il momento secondo risulta indefinito, impedendo l'uso del Teorema del limite centrale. Tuttavia una generalizzazione di tale teorema (di Gnedenko e Kolmogorov [9]), mostra che, abbandonando l'ipotesi di varianza finita, la somma di variabili *iid* con media 0 e varianza unitaria tende a una distribuzione stabile.

Teorema del limite centrale generalizzato (TLCG). *Una variabile aleatoria non degenera Z è stabile per $0 < \alpha \leq 2$, se e solo se esiste una successione di variabili aleatorie *iid* $X_1, X_2, X_3 \dots$ e costanti $a_n > 0$, $b_n \in \mathbb{R}$ tali che*

$$a_n(X_1 + \dots + X_n) - b_n \xrightarrow{d} Z. \quad (1.12)$$

Se una variabile aleatoria X soddisfa la (1.12) allora si dice che X è nel *dominio di attrazione* di Z , $DA(Z)$. Il TLCG afferma che le uniche distribuzioni non degeneri

con un dominio di attrazione sono stabili.

Il teorema caratterizza le distribuzioni nel $DA(Z)$ in termini di probabilità delle code. Ad esempio si consideri la variabile X la cui approssimazione delle code soddisfi la seguente legge di potenza, con $(C^+ + C^-) > 0$, $0 < \alpha \leq 2$, per $x \rightarrow \infty$:

$$\begin{aligned}\Pr\{X > x\} &\rightarrow C^+ x^{-\alpha}, \\ \Pr\{X < -x\} &\rightarrow C^- x^{-\alpha}.\end{aligned}$$

Allora, $\mu = \mathbb{E}[X]$ esiste finita e l'equazione analoga alla (1.11) nel caso del Teorema generalizzato risulta:

$$a_n(X_1 + \dots + X_n) - b_n \xrightarrow{d} Z \sim \mathbf{S}(\alpha, \beta, 1, 0; 1), \quad \text{per } n \rightarrow \infty,$$

con costanti $a_n = [(2\Gamma(\alpha) \sin(\frac{\pi\alpha}{2})/\pi(C^+ + C^-))]^{1/\alpha} n^{-1/\alpha}$, e $b_n = na_n\mu$. Il tasso al quale decadono le code della distribuzione determina l'indice α , mentre il peso relativo della coda sinistra e destra caratterizza il parametro di asimmetria, $\beta = (C^+ - C^-)/(C^+ + C^-)$.

CAPITOLO 2

ANALISI EMPIRICA

Di seguito verrà presentata un'analisi empirica effettuata su un ampio campione di rendimenti di titoli azionari statunitensi. I rendimenti saranno modellati mediante la distribuzione Lévy-Pareto stabile e, comparativamente, la distribuzione di Gauss. Successivamente si procederà alla ricerca di possibili *cluster* settoriali per il parametro di stabilità α , paragrafo 1.3.3 a pagina 15.

Per l'elaborazione tecnica dei dati verrà utilizzato il software Wolfram Mathematica 8.0. I dati sono stati raccolti dal sito www.yahoofinance.com.

I risultati di seguito illustrati sono stati conseguiti nell'ambito di uno stage formativo presso Redex "Risk Management and Finance" Srl, in collaborazione con il Dott. Riccardo Donati. Gli esiti raggiunti hanno supportato la verifica empirica della misura di rischio *RedES*TM sviluppata ed utilizzata da Redex. Questo capitolo è stato redatto quale relazione tecnica a conclusione dello stage.

2.1 Raccolta e Manipolazione dei dati

L'analisi empirica effettuata si basa su un campione di titoli rappresentanti una quota importante del mercato azionario statunitense. In particolare sono state utilizzate le serie storiche dei prezzi giornalieri di chiusura *adjusted* di *

*Prezzi di chiusura aggiustati per eventuali dividendi distribuiti nella giornata di transazione considerata o frazionamenti azionari.

- 1007 titoli azionari componenti il NYSE Composite Index (^NYA);
- 2600 titoli azionari componenti il NASDAQ Composite Index (^IXIC).

Il *NYSE Composite Index* è un indice azionario contenente tutte le azioni, statunitensi e estere, quotate presso il New York Stock Exchange (NYSE). Il peso relativo delle aziende nell'indice è calcolato in base al *flozzante* delle stesse, ovvero la quota di capitale azionario effettivamente circolante sul mercato, non detenuta dall'azionista di riferimento o dal gruppo di controllo di una società. L'indice raccoglie il 61% della capitalizzazione totale di tutte le aziende quotate nel mondo. L'indice è stato costituito nel dicembre 1965 con base 50; nel gennaio 2003 è stato ricalcolato e reintrodotta con base 5000. Attualmente ha raggiunto una quotazione pari a circa 7000 punti.

Il *NASDAQ Composite Index* è un indice azionario che raccoglie le azioni quotate presso il NASDAQ, mercato di riferimento per i titoli tecnologici della borsa americana e secondo mercato nel mondo per capitalizzazione, dopo il NYSE. Le componenti dell'indice sono pesate in base alla loro capitalizzazione di mercato. L'indice è stato costituito nel febbraio 1971 con base 100. Ad oggi la sua quotazione ha superato i 2600 punti.

Il numero di titoli azionari componenti i due indici riflette esattamente la numerosità delle aziende quotate presso il NYSE e il NASDAQ il 5 Luglio 2011, ovvero la data di analisi considerata.

Analizzare le componenti dei due indici sopraccitati significa disporre di un campione di titoli rappresentanti la quasi totalità del mercato azionario statunitense e una quota rilevante del mercato azionario mondiale. Questo permette di conferire validità ai risultati conseguiti, sia a livello teorico, sia per possibili applicazioni pratiche nell'ambito del *Risk Management*. Le serie storiche utilizzate coprono un intervallo temporale di oltre cinquant'anni, contando a ritroso dal 5 Luglio 2011, data di analisi. Più precisamente le prime osservazioni risalgono al 3 Marzo 1962.

Per ogni titolo, oltre alla serie storica dei prezzi, sono stati importati:

- nome dell'azienda, ad esempio *Apple Inc.*;

- *ticker*, ovvero il codice identificativo dell'azione, ad esempio *AAPL*;
- settore di appartenenza;
- capitalizzazione, cioè il valore di mercato della società;
- flottante;
- data del primo rilievo di prezzo, ovvero la data di prima quotazione del titolo nel listino di borsa considerato.

Una prima scrematura dei dati ha consentito di eliminare i titoli per i quali risulta mancante uno o più indicatori sopraelencati, riducendo il campione a 3140 titoli suddivisi in:

- 936 titoli azionari componenti il NYSE Composite Index (^NYA);
- 2204 titoli azionari componenti il NASDAQ Composite Index (^IXIC),

Per un totale di 3140 attività considerate. A partire dalle serie storiche dei prezzi sono state ottenute le serie storiche dei rendimenti logaritmici (1.1). Per semplicità i log-rendimenti saranno d'ora in avanti definiti rendimenti.

2.1.1 Distribuzione temporale

Come è logico presupporre, non tutti i titoli quotati nei listini di interesse il 5 Luglio 2001 erano presenti sul mercato anche cinquant'anni prima, data in cui risale la prima osservazione disponibile. Dal 1962, infatti, il NYSE e il NASDAQ hanno subito un notevole ampliamento, fino ad arrivare a racchiudere oltre il 70% della capitalizzazione mondiale di borsa. È necessario quindi, scegliere una data di inizio per le serie storiche dei rendimenti che sia più significativa rispetto al marzo '62, data in cui il nostro campione è ridotto a solo nove titoli quotati sul NYSE, tra cui The Coca Cola Company, Caterpillar Inc., The Walt Disney Company,

Per ogni serie storica è stato calcolato l'intervallo temporale tra il primo e l'ultimo rilievo di prezzo, ricavando una lista contenente il numero di anni di presenza sul mercato di ciascun titolo analizzato. La figura 2.1 a pagina 35 mostra la distribuzione

dei titoli rispetto al momento di quotazione delle azioni della società presso uno dei due mercati considerati, il NYSE o il NASDAQ. Sull'asse delle ordinate è indicato il numero totale di anni di permanenza sul mercato del titolo, dalla data di ingresso fino alla data di analisi, 5 Luglio 2011.

Più di 300 titoli sono stati quotati tra il 1990 e il 1991 e oltre 200 società sono entrate in borsa poco prima del 2001. Per assecondare questi boom di ingressi nei mercati regolamentati, è stata fatta la scelta di considerare due diverse finestre temporali: dieci e vent'anni. È stato inoltre ritenuto opportuno evidenziare un terzo intervallo temporale, più recente, di soli due anni. Le serie storiche dei rendimenti giornalieri saranno studiate per tre periodi di lunghezza definita come segue:

- 5 Luglio 1991 - 5 Luglio 2011, 20 anni, 759 titoli considerati;
- 5 Luglio 2001 - 5 Luglio 2011, 10 anni, 2233 titoli considerati;
- 5 Luglio 2009 - 5 Luglio 2011, 2 anni, 3050 titoli considerati.

Per ogni singolo titolo è stato calcolato il numero di sessioni in cui l'attività è stata effettivamente scambiata sul mercato e l'intervallo medio fra sessioni successive. Quest'ultimo è circa pari a 1.45 giorni e risulta coerente con una serie giornaliera dei rendimenti, poiché l'attività di *trading* sui mercati regolamentati è solitamente effettuata cinque giorni a settimana.

2.1.2 Filtro sui rendimenti nulli

Esaminando le serie temporali è stata osservata la presenza di rendimenti esattamente nulli. La probabilità che tale rendimento sia "reale" è piuttosto bassa e trascurabile. La spiegazione più plausibile è che il titolo non sia stato scambiato in tale giornata. È stato quindi costruito un filtro per eliminare i rendimenti pari a zero ottenendo tre ulteriori serie storiche (2, 10, 20 anni "filtrate").[†]

Nel corso della trattazione verranno esposti i risultati delle analisi comparativamente per le serie con e senza filtro sullo zero.

[†]Il filtro sui rendimenti nulli verrà definito per semplicità *filtro sullo zero*.

2.2 Distribution fitting

Con il termine *distribution fitting* si indica la rappresentazione analitica delle distribuzioni statistiche. Essa consiste nel trovare una funzione matematica interpolante che rappresenti opportunamente un fenomeno statistico osservato. Il problema affrontato è il seguente: si ha una serie di osservazioni di un carattere quantitativo indicate con x_1, x_2, \dots, x_n , e si vuole verificare se tali osservazioni – che costituiscono un campione di una popolazione non nota a priori – provengono da una determinata popolazione caratterizzata da una funzione di densità di frequenza della quale si conosce l'espressione in termini analitici. Indichiamo con $f(x | \Theta)$ tale funzione, che rappresenta la distribuzione del carattere, dove Θ è un vettore di parametri da stimare sulla base dei dati disponibili. Nel caso trattato la variabile di interesse è il rendimento logaritmico giornaliero e il campione dei dati a disposizione è composto dalle serie storiche individuate nel precedente paragrafo.

Nella rappresentazione analitica possiamo individuare tre fasi:

- scelta del modello;
- stima dei parametri del modello scelto;
- applicazione di test statistici per saggiare la bontà del modello (*goodness of fit*).

2.2.1 Scelta del modello

La prima fase del *fitting* altro non è che un problema di specificazione di un modello matematico adeguato alla rappresentazione dei dati raccolti. Un primo approccio esplorativo con i dati può essere quello grafico. Utilizzando il software Mathematica è stato possibile stimare la funzione di densità della frequenza dei dati. La figura 2.2 a pagina 35 mostra la densità stimata dei log-rendimenti giornalieri della Foot Locker Inc., con orizzonte temporale ventennale e senza filtro sullo zero. Esiste un modello teorico che descriva in maniera soddisfacente la distribuzione osservata? Se, per esempio, fosse possibile trovare una distribuzione avente la stessa “forma”

della PDF dei rendimenti della Foot Locker Inc., sarebbe verosimile utilizzare le proprietà conosciute del modello teorico per prevedere *ex-ante* il comportamento della variabile analizzata.

Louis Bachelier [1] fu il primo che introdusse l'idea che le variazioni dei prezzi sui mercati si muovessero seguendo una "passeggiata casuale" (*random walk*) e che scelse il moto browniano come candidato per la descrizione delle fluttuazioni di prezzo. Il modello di Bachelier applicato ai prezzi logaritmici divenne molto popolare negli anni cinquanta ed è uno degli ingredienti della famosa formula di Black, Scholes e Merton per il *pricing* delle opzioni. Questo modello implica che i log-rendimenti azionari siano variabili *iid* con distribuzione gaussiana. Tale concetto è inoltre confermato dal Teorema del limite centrale (1.10), che propone la distribuzione normale quale candidata ideale per modellare i rendimenti.

Oggi è tuttavia ampiamente riconosciuta la scarsa capacità della distribuzione di Gauss di descrivere uno dei principali fatti stilizzati dei rendimenti azionari: la leptocurtosi. Mandelbrot [13] propose una distribuzione α -stabile per lo studio delle variazioni dei log-prezzi del cotone. Il suo lavoro fu completato da Fama [7], il quale, con un'analisi dei rendimenti giornalieri delle azioni appartenenti al Dow Jones Industrial Average, confermò le teorie di Mandelbrot, arrivando ad una stima media del parametro di stabilità di circa 1.9. Le scoperte dei due studiosi hanno aperto un ampio dibattito accademico sulla scelta della distribuzione adeguata a modellare le distribuzioni di rendimenti azionari.

L'analisi statistica dei dati raccolti presentata in questo capitolo utilizza la classe di distribuzioni Lévy-Pareto stabili con esponente di stabilità $0 < \alpha < 2$ (paragrafo 1.3 a pagina 13) come modello teorico per la definizione del comportamento dei rendimenti azionari osservati. La distribuzione normale è adoperata a fini comparativi.

Alla base della scelta delle distribuzioni stabili come modello per le variazioni dei log-prezzi si trova la proprietà di stabilità (1.3), la quale afferma che la somma di due variabili aleatorie *iid* con distribuzione stabile e α fissato genera una variabile stabile con il medesimo parametro α . Tale proprietà, anche definita *invarianza di scala*, riflette il comportamento dei log-rendimenti: il rendimento calcolato con-

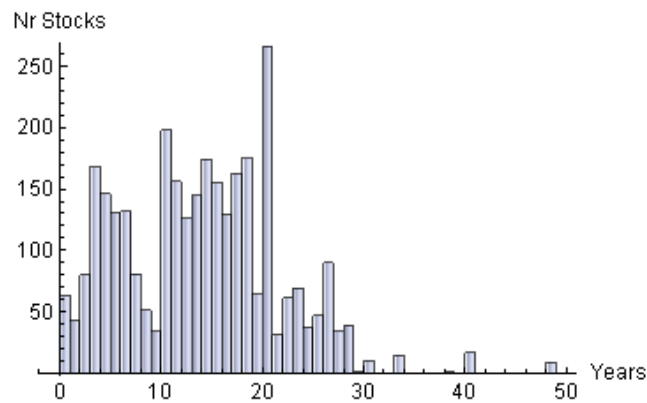
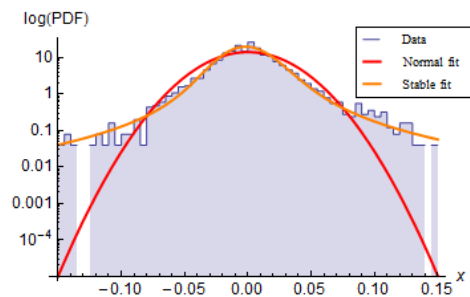
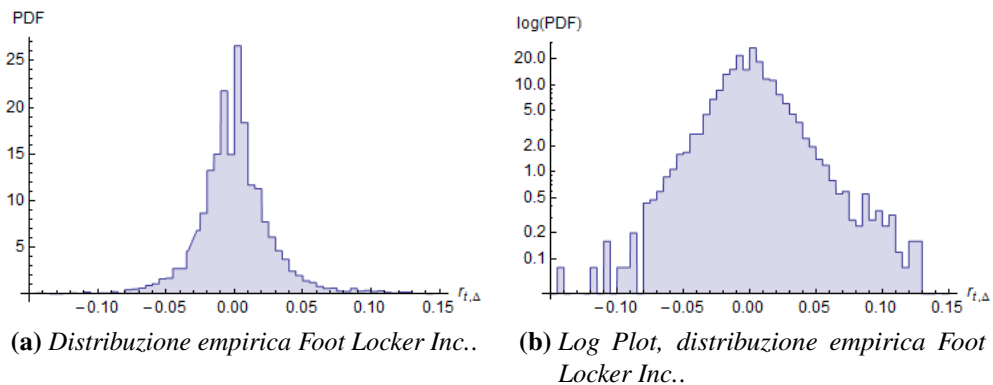


Figura 2.1. Distribuzione temporale dei titoli. Sull'asse delle ordinate è indicato il numero totale di anni di permanenza sul mercato del titolo, dalla data di ingresso, fino alla data di analisi, 5 Luglio 2011. L'altezza delle barre indica la numerosità dei titoli presenti nella classe. Il grafico può essere anche interpretato come la quantità di azioni quotate per la prima volta presso il NYSE e il NASDAQ x anni fa, contando a ritroso dalla data di analisi (l'origine nel grafico).



(c) Log Plot, Foot Locker Inc., distribution fitting.

Figura 2.2. Distribuzione empirica e stimata del titolo Foot Locker Inc., finestra temporale ventennale. I grafici (a) e (b) mostrano la distribuzione empirica dei log-rendimenti giornalieri del titolo. (c) illustra due diversi *fit* utilizzando alternativamente una distribuzione normale e una distribuzione stabile.

siderando, per esempio, un intervallo temporale mensile, è pari alla somma dei rendimenti settimanali o giornalieri componenti l'intervallo stesso (1.2). Il Teorema del limite centrale generalizzato (1.12) conferisce validità alla scelta del modello stabile, permettendo di superare l'ostacolo di varianza infinita che preclude l'uso del classico Teorema del limite centrale.

2.2.2 Stima dei parametri

Una volta scelta la funzione che meglio si adatta al fenomeno da rappresentare è necessario stimare i parametri che caratterizzano tale modello sulla base dei dati disponibili. Il metodo utilizzato è il metodo della massima verosimiglianza (ML, *maximum likelihood*). Si suppone nota $f(x | \Theta)$, ovvero la funzione di densità della distribuzione teorica che descrive il carattere quantitativo a livello di popolazione. Di tale funzione occorre stimare il vettore dei parametri Θ in base alle osservazioni campionarie x_1, x_2, \dots, x_n . Viene definita *funzione di verosimiglianza* la seguente espressione:

$$L(x_1, x_2, \dots, x_n | \Theta) = \prod_{i=1}^n f(x_i | \Theta).$$

Il metodo della massima verosimiglianza consiste nello stimare Θ in modo che venga massimizzata $L(x_1, x_2, \dots, x_n | \Theta)$, oppure, per semplificare i calcoli, il suo logaritmo:

$$L(x_1, x_2, \dots, x_n | \Theta) = \sum_{i=1}^n \log f(x_i | \Theta).$$

Nei casi più semplici lo stimatore Θ è ottenuto uguagliando a zero le derivate parziali; quando la funzione di verosimiglianza assume forme troppo complesse si usano invece i metodi di calcolo numerico.

Gli stimatori ML godono di una serie di proprietà molto utili da punto di vista statistico, tra le quali la minima varianza e l'assenza di errori sistematici. Quest'ultima caratteristica può essere chiarita pensando di estrarre un numero sufficientemente ampio di campioni casuali da una popolazione, in mancanza di un errore sistematico il valore medio dei parametri stimati risulterà uguale ai valori esatti dei parametri dell'intera popolazione. Uno stimatore con minima varianza è, invece, lo stimatore

che possiede la varianza più piccola, ovvero l'intervallo di confidenza più "stretto" di tutti gli stimatori del medesimo tipo.

Per ciascuna serie dei rendimenti osservata (2, 10, 20 anni; con e senza filtro sullo zero) sono state effettuate due diverse stime, considerando le due distribuzioni teoriche di maggior interesse nella caratterizzazione dei rendimenti finanziari:

- gaussiana, con funzione di densità $f(x | \mu, \sigma^2)$ e vettore dei parametri da stimare $\Theta = (\mu, \sigma^2)$.
- Lévy-Pareto stabile, con funzione di densità $\hat{f}(x | \alpha, \beta, \gamma, \delta)$ e vettore dei parametri da stimare $\Theta = (\alpha, \beta, \gamma, \delta)$. L'accento circonflesso sottolinea il fatto che, in generale, la forma esplicita della densità non è conosciuta e deve essere approssimata numericamente.

L'elaborazione dei due *fit* per la totalità del campione raccolto ha richiesto circa 5 giorni di calcolo parallelo su un *server* con 8 processori. Uno dei principali svantaggi della stima ML è infatti la lentezza del metodo, compensato tuttavia dall'altissima precisione dei risultati.

2.2.3 Goodness of fit test

La verifica della conformità dei dati osservati ad un modello teorico (*goodness of fit*) può essere effettuata con diversi test statistici. Si tratta di test talvolta definiti *test omnibus*. Essi offrono un approccio globale al problema: la conformità tra i dati campionari e la popolazione viene esaminata in un quadro complessivo che include tutte le caratteristiche del carattere oggetto di studio (media, variabilità, forma della distribuzione, ...). Tali test sono definiti *distribution free*, ossia risultano indipendenti dalla distribuzione del carattere.

Un primo test per verificare la bontà del *fit* dei dati osservati rispetto ad un modello teorico è il *Test χ^2 di Pearson*. Esso si basa sul confronto delle frequenze empiriche con quelle attese calcolate in base alla funzione di densità della frequenza impiegata. Le ipotesi sottoposte a verifica dalla statistica sono:

- H_0 : i dati sono compatibili con la distribuzione teorica specificata;

- H_1 : i dati non sono compatibili con la distribuzione teorica specificata.

Il campione è diviso in k classi di ampiezza omogenea e la statistica di riferimento è:

$$X^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i},$$

dove O_i indica la frequenza osservata e E_i la frequenza attesa per la i -esima classe considerata. Le frequenze attese sono determinata a partire dalla funzione di distribuzione del modello teorico scelto. Tale statistica si distribuisce asintoticamente secondo una variabile casuale χ^2 con $(k - n - 1)$ gradi di libertà (n è il numero di parametri che caratterizzano la distribuzione).

L'ipotesi nulla che i dati analizzati provengano da una popolazione con distribuzione assegnata è rifiutata se:

$$X^2 > \chi_{a, (k-n-1)}^2,$$

dove il secondo termine della disuguaglianza rappresenta il valore di una distribuzione χ^2 con $(k - n - 1)$ gradi di libertà e livello di significatività pari ad a .
‡

Ad ogni valore assunto da X^2 è possibile assegnare il corrispondente valore di probabilità o *p-value*. Si accetta l'ipotesi di conformità al modello se il valore di probabilità ottenuto dal test è superiore al livello di significatività prefissato. In altre parole, l' H_0 viene rifiutata se la probabilità di ottenere i dati osservati, a partire dalla distribuzione teorica scelta, è minore di a .

Uso del test χ^2 di Pearson

Il test statistico è applicato ad ogni singolo *fit*, utilizzando alternativamente le frequenze attese della distribuzione gaussiana o stabile. Al fine di rendere più chiara la visualizzazione grafica dei risultati, è stata utilizzata una trasformazione logaritmica del valore di probabilità: $g = -\log_{10}(p\text{-value})$. Si ricorda al lettore che,

‡Al fine di evitare confusione con il parametro di stabilità α , il livello di significatività dei test statistici utilizzati (test χ^2 di Pearson e test di Kruskal-Wallis) è indicato con la lettera dell'alfabeto a .

poiché $0 < (p\text{-value}) < 1$, la trasformazione effettuata rende valori di g positivi e crescenti per $p\text{-value}$ via via più vicini allo zero. Per valori di g superiori a $-\log_{10}(a)$ viene rifiutata l'ipotesi nulla di conformità dei dati con il modello teorico scelto.

Nel processo di interpretazione degli esiti del test non è stato fissato a priori un livello di significatività, cioè il procedimento di valutazione di bontà del modello selezionato non si è concluso con un rifiuto o un accoglimento dell'ipotesi nulla. I valori di g sono interpretabili come quantificazione dell'errore del *fit*:

- alti valori di g sono significativi di una scarsa capacità del modello di spiegare la distribuzione empirica dei dati;
- bassi valori di g implicano una buona aderenza del modello teorico alle frequenze osservate.

Effetto del filtro sui rendimenti nulli

Il primo uso dei risultati del Test χ^2 di Pearson è un confronto tra le stime dei parametri effettuate sulle serie storiche alle quali è stato applicato il filtro sui rendimenti nulli e le stime sulle serie originali (paragrafo 2.1.2 a pagina 32). Il filtro sullo zero è stato costruito per eliminare le variazioni di prezzo nulle. L'idea alla base di tale scelta è la convinzione che i rendimenti nulli siano generati dalla mancata transazione del titolo in tale giornata. Eliminando questi dati dovrebbe essere possibile ottenere serie storiche più aderenti alla realtà e, di conseguenza, *fit* migliori.

I risultati sono mostrati nelle figure 2.3 a pagina 41 e 2.4 a pagina 42. Si sottolinea che lo scopo dei grafici non è quello di comparare la bontà delle stime stabili e normali, bensì capire se l'applicazione del filtro sullo zero permette di ottenere *fit* migliori, trascurando il modello teorico scelto. Peraltra un'eventuale confronto tra stima stabile e normale sarebbe reso difficile dalle diverse scale, sia sull'asse delle ascisse sia su quello delle ordinate, utilizzate nelle figure sopraccitate. Indicando con $\bar{g} = \frac{1}{N} \sum_{i=1}^N g_i$ la media aritmetica dei valori assunti dalla trasformazione logaritmica del $p\text{-value}$, N la numerosità del campione e $[]_f$ la presenza del filtro sui rendimenti nulli, per il *fit* stabile si ha:

- finestra temporale 2 anni, $\bar{g} = 10.542 > \bar{g}_f = 1.944$;
- finestra temporale 10 anni, $\bar{g} = 100.705 > \bar{g}_f = 8.232$;
- finestra temporale 20 anni, $\bar{g} = 427.297 > \bar{g}_f = 22.406$.

Similarmente per la stima del modello gaussiano, i risultati medi del Test di Pearson sono:

- finestra temporale 2 anni, $\bar{g} = 19.404 > \bar{g}_f = 11.331$;
- finestra temporale 10 anni, $\bar{g} = 278.371 > \bar{g}_f = 201.579$;
- finestra temporale 20 anni, $\bar{g} = 754.798 > \bar{g}_f = 503.493$.

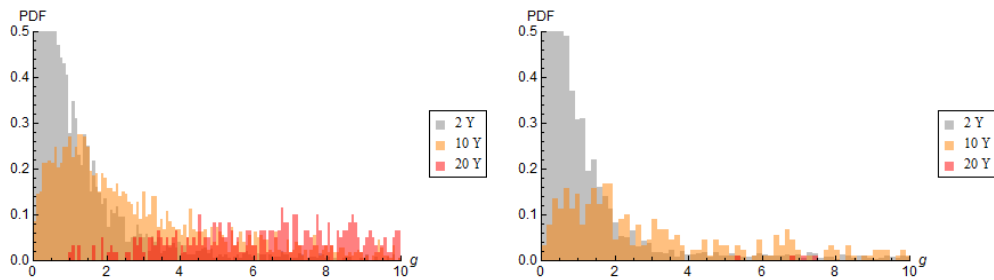
A prescindere dal tipo di modello usato per la stima parametrica, i valori di g rilevati sono maggiori sia in media sia in varianza per le serie non filtrate. Seguendo l'interpretazione data nel paragrafo precedente, è possibile affermare che l'applicazione del filtro permette di migliorare la bontà del *fit*, mantenendo l'errore (g) a livelli più contenuti.

Confronto tra fit normale e fit stabile

Nel paragrafo 2.2.1 a pagina 33 sono illustrati i vantaggi di una distribuzione stabile quale modello per descrivere i log-rendimenti azionari, rispetto al classico uso della distribuzione di Gauss. Il *goodness of fit test* di Pearson permette di capire se, a partire dai dati empirici, il modello stabile sia effettivamente più efficace nel descrivere le frequenze osservate.

I risultati, suddivisi per orizzonte temporale considerato, sono illustrati nelle figure 2.5- 2.7 a pagina 46. I valori del test per le serie senza filtro sui rendimenti nulli vengono temporaneamente tralasciati perché poco significativi per la trattazione del caso in esame.

Per ciascun titolo considerato e per ogni periodo di osservazione (2, 10, 20 anni), il *fit* stabile sembra descrivere con errore g relativamente basso la distribuzione empirica dei dati osservati. Gli esiti del test per la stima dei rendimenti mediante distribuzione normale mostrano valori di g maggiori in media e notevolmente più



(a) Presenza del filtro sui rendimenti nulli, fit stabile, Test χ^2 di Pearson.

(b) Assenza del filtro sui rendimenti nulli, fit stabile, Test χ^2 di Pearson.

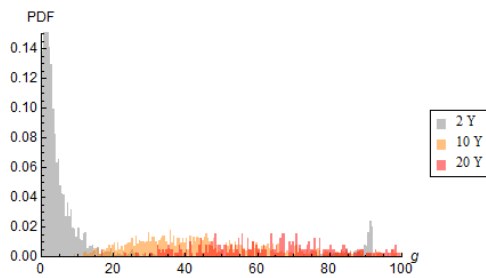
	Samples	Mean	Std. Dev.	Skewness	Kurtosis
2Y	3050	1.944	5.309	11.498	202.288
10Y	2226	8.232	25.404	8.271	89.726
20Y	755	22.406	41.851	8.193	90.119

	Samples	Mean	Std. Dev.	Skewness	Kurtosis
2Y	2937	10.542	49.206	10.347	131.119
10Y	593	100.705	206.424	3.455	17.607
20Y	707	427.297	577.381	2.718	12.238

(c) Presenza del filtro sui rendimenti nulli, statistiche descrittive.

(d) Assenza del filtro sui rendimenti nulli, statistiche descrittive.

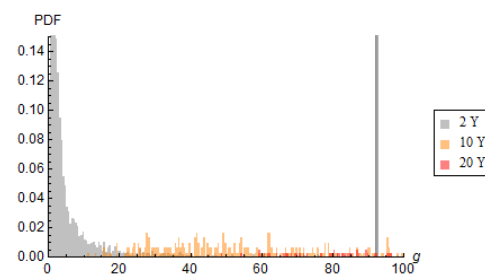
Figura 2.3. Effetto del filtro sui rendimenti nulli, Test χ^2 di Pearson sul fit stabile. A sinistra: distribuzione (a) e statistiche descrittive (c) dei risultati del test effettuato sulle serie storiche dopo l'applicazione del filtro sui rendimenti nulli. A destra: distribuzione (b) e statistiche descrittive (d) dei risultati del test effettuato sulle serie storiche originarie. $g = -\log_{10}$ è interpretabile come l'errore della stima parametrica rispetto ai dati empirici osservati. I fit effettuati sulle serie di rendimenti con applicazione del filtro sullo zero presentano valori di g più contenuti. L'applicazione del filtro sui rendimenti nulli permette di migliorare la bontà delle stime.



(a) Presenza del filtro sui rendimenti nulli, fit normale, Test χ^2 di Pearson.

	Samples	Mean	Std. Dev.	Skewness	Kurtosis
2Y	3050	11.331	23.051	2.830	9.591
10Y	2226	201.579	197.959	0.590	1.475
20Y	755	503.493	429.749	0.048	1.059

(c) Presenza del filtro sui rendimenti nulli, statistiche descrittive.



(b) Assenza del filtro sui rendimenti nulli, fit normale, Test χ^2 di Pearson.

	Samples	Mean	Std. Dev.	Skewness	Kurtosis
2Y	3050	11.331	23.051	2.830	9.591
10Y	2226	201.579	197.959	0.590	1.475
20Y	755	503.493	429.749	0.048	1.059

(d) Assenza del filtro sui rendimenti nulli, statistiche descrittive.

Figura 2.4. Effetto del filtro sui rendimenti nulli, Test χ^2 di Pearson sul fit normale. A sinistra: distribuzione (a) e statistiche descrittive (c) dei risultati del test effettuato sulle serie storiche dopo l'applicazione del filtro sui rendimenti nulli. A destra: distribuzione (b) e statistiche descrittive (d) dei risultati del test effettuato sulle serie storiche originarie. $g = -\log_{10}$ è interpretabile come l'errore della stima parametrica rispetto ai dati empirici osservati. I fit effettuati sulle serie di rendimenti con applicazione del filtro sullo zero presentano valori di g più contenuti. L'applicazione del filtro sui rendimenti nulli permette di migliorare la bontà delle stime.

	2Y		10Y		20Y	
	Filter	No Filter	Filter	No Filter	Filter	No Filter
Stable fit	1.943	10.542	8.232	100.705	22.406	427.497
Normal fit	11.331	19.404	201.579	278.371	503.493	754.798

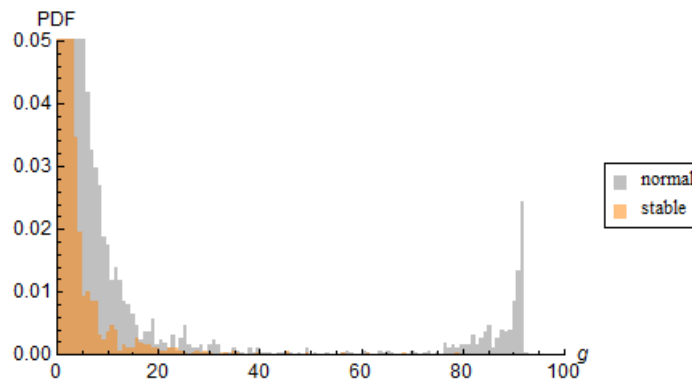
Tabella 2.1. Test χ^2 di Pearson, tabella riassuntiva delle probabilità $g = -\log_{10}$. Vengono presentati i valori di probabilità medi ottenuti dall'applicazione del test di bontà del modello applicato alle stime normali e stabili delle serie storiche dei rendimenti, con e senza l'applicazione del filtro sui rendimenti nulli. Si ricorda che il valore di probabilità g è da intendersi come errore di stima del modello; più il valore di g è alto, meno il *fit* è aderente ai dati osservati. L'uso delle distribuzioni Lévy-Pareto stabili permette di descrivere in modo efficace i dati raccolti con errori contenuti per ogni periodo temporale considerato. Il filtro sullo zero depura, invece, le serie dai rendimenti nulli "fittizi" dovuti a giornate di mancata transazione del titolo. L'uso del filtro permette di migliorare le stime a prescindere dal modello teorico scelto.

dispersi rispetto all'uso di stime stabili. Ciò significa che solo per poche serie storiche il *fit* normale si rivela adeguato tanto quanto quello stabile. In particolare si ricorda che la stima del modello può essere considerata tanto più buona quanto più il valore di probabilità è vicino allo zero. Nel caso in cui $g = 0$, la distribuzione empirica coincide con la distribuzione teorica scelta per la rappresentazione dei dati.

La tabella 2.1 riassume le analisi presentate finora mostrando i valori di probabilità medi (\bar{g}) del *goodness of fit test* condotto sui *fit* stabili e normali. È possibile affermare che le distribuzioni Lévy-Pareto stabili sono effettivamente un modello teorico efficace nel descrivere le fluttuazioni dei log-prezzi. È inoltre evidente che l'uso del filtro sui rendimenti nulli permette di migliorare ulteriormente le stime stabili.

2.3 Clustering settoriale

Noti i fatti stilizzati presentati nel paragrafo 1.3 a pagina 13, di particolare interesse è il valore assunto dal parametro di stabilità α nel processo di stima delle distribuzioni osservate mediante la classe di modelli stabili. Se i rendimenti rilevati sono distribuiti come una Lévy-Pareto stabile $\mathbf{S}(\alpha, \beta, \gamma, \delta; 1)$, allora α deve assumere valori positivi inferiori o uguali a due. Se $\alpha = 2$ la distribuzione è gaussiana con parametri (μ, σ^2) .

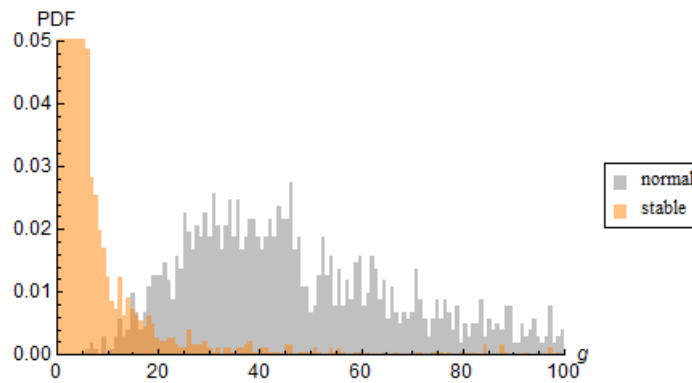


(a) Test χ^2 di Pearson, fit stabile e fit normale, 2Y.

2Y	Mean	Std. Dev.	Skewness	Kurtosis
Stable fit	1.944	5.309	11.498	202.288
Normal fit	11.331	23.051	2.830	9.591

(b) Statistiche descrittive, 2Y.

Figura 2.5. Bontà del *fit* normale e stabile a confronto sulla finestra temporale di due anni. I valori di probabilità ($g = -\log_{10}$) del Test di Pearson per la stima stabile dei parametri della distribuzione sono inferiori rispetto ai valori del *fit* gaussiano. Il modello stabile è in grado di fornire una descrizione più accurata delle frequenze osservate rispetto ad un modello normale.



(a) Test χ^2 di Pearson, fit stabile e fit normale, 10Y.

10Y	Mean	Std. Dev.	Skewness	Kurtosis
Stable fit	8.232	25.404	8.271	89.726
Normal fit	201.579	197.959	0.590	1.475

(b) Statistiche descrittive, 10Y.

Figura 2.6. Bontà del *fit* normale e stabile a confronto sulla finestra temporale di dieci anni. I valori di probabilità ($g = -\log_{10}$) del Test di Pearson per la stima stabile dei parametri della distribuzione sono inferiori rispetto ai valori del *fit* gaussiano. Il modello stabile è in grado di fornire una descrizione più accurata delle frequenze osservate rispetto ad un modello normale.

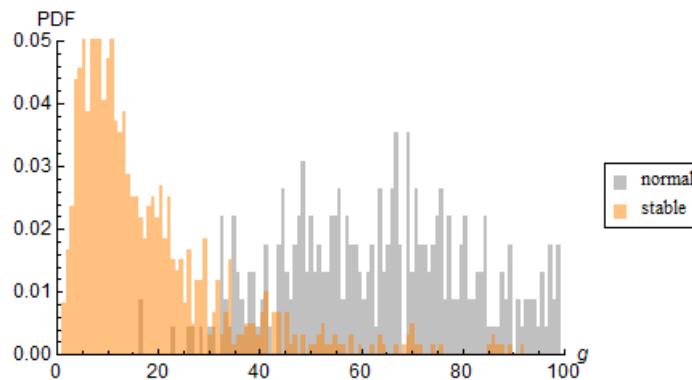
Il parametro di stabilità determina la forma della distribuzione definendo lo spessore delle code e il valore assunto dalla moda. Per $\alpha < 2$ la distribuzione è detta leptocurtica con code più spesse e valori centrali più frequenti rispetto ad una distribuzione normale. Più il parametro è prossimo al valore minimo ($\alpha = 0$), più aumenta lo spessore delle code, figura 1.3 a pagina 21, cioè eventi cosiddetti estremi (sia positivi sia negativi) risultano via via più frequenti. In generale, una distribuzione gaussiana considera rendimenti molto elevati o molto bassi come poco plausibili. In realtà, come verrà dimostrato in seguito, le distribuzioni empiriche mostrano, per orizzonti temporali sufficientemente ampi, un numero considerevole di rendimenti “non normali”.

L’analisi dell’indice di stabilità α è effettuata in tre *step* successivi:

- stima dei parametri $(\alpha, \beta, \gamma, \delta)$ del modello stabile sulle serie storiche campionarie (paragrafo 2.2.2 a pagina 36).
- Estrapolazione del parametro di stabilità α da ciascuno dei *fit* effettuati; di fatto per ogni orizzonte temporale (2, 10, 20 anni) e per ogni tipologia di serie storica (con e senza filtro) si ottengono tanti valori di α quanti sono i titoli del gruppo considerato. Per esempio nel gruppo “20 anni–con filtro” sono presenti 755 serie storiche, per le quali sono calcolati altrettanti *fit* stabili. Dalle stime stabili vengono quindi raccolti 755 valori di α .
- Calcolo della distribuzione empirica di α a partire dalle frequenze rilevate all’interno del gruppo considerato.

Le PDF empiriche di α così ottenute sono mostrate nella figura 2.8 a pagina 47; il parametro di stabilità sembra assumere valori coerenti con l’ipotesi di distribuzione stabile, mediamente compresi tra 1.5 e 1.8. Le statistiche descrittive della densità di probabilità di α sono presentate nella tabella 2.2 nella pagina seguente.

Ottenere una stima del parametro di stabilità statisticamente significativa e precisa non sempre è fattibile. Il processo di *fitting* mediante il metodo di massima verosimiglianza richiede calcolatori molto potenti; inoltre può capitare che un titolo



(a) Test χ^2 di Pearson, fit stabile e fit normale, 20Y.

20Y	Mean	Std. Dev.	Skewness	Kurtosis
Stable fit	22.406	41.851	8.193	90.119
Normal fit	503.493	429.749	0.048	1.059

(b) Statistiche descrittive, 20Y.

Figura 2.7. Bontà del *fit* normale e stabile a confronto sulla finestra temporale di vent'anni. I valori di probabilità ($g = -\log_{10}$) del Test di Pearson per la stima stabile dei parametri della distribuzione sono inferiori rispetto ai valori del *fit* gaussiano. Il modello stabile è in grado di fornire una descrizione più accurata delle frequenze osservate rispetto ad un modello normale.

		Samples	Mean	Std. Dev.	Skewness	Kurtosis
2Y	Filter	3050	1.742	0.150	-2.501	19.120
	No Filter	2937	1.708	0.196	-2.788	17.183
10Y	Filter	2226	1.561	0.128	-1.145	5.112
	No Filter	593	1.509	0.149	-1.474	6.414
20Y	Filter	755	1.657	0.094	-0.942	4.417
	No Filter	707	1.594	0.110	-1.026	4.227

Tabella 2.2. Statistiche descrittive della distribuzione empirica del parametro di stabilità α . I valori di α derivano dai *fit* stabili effettuati sul campione complessivo di dati a disposizione. La media dei valori oscilla tra 1.5 e 1.8 a seconda della tipologia della serie storica considerata. La deviazione standard è piuttosto contenuta. I valori assunti dal parametro sono coerenti con una distribuzione Lévy-Pareto stabile con $0 < \alpha < 2$.

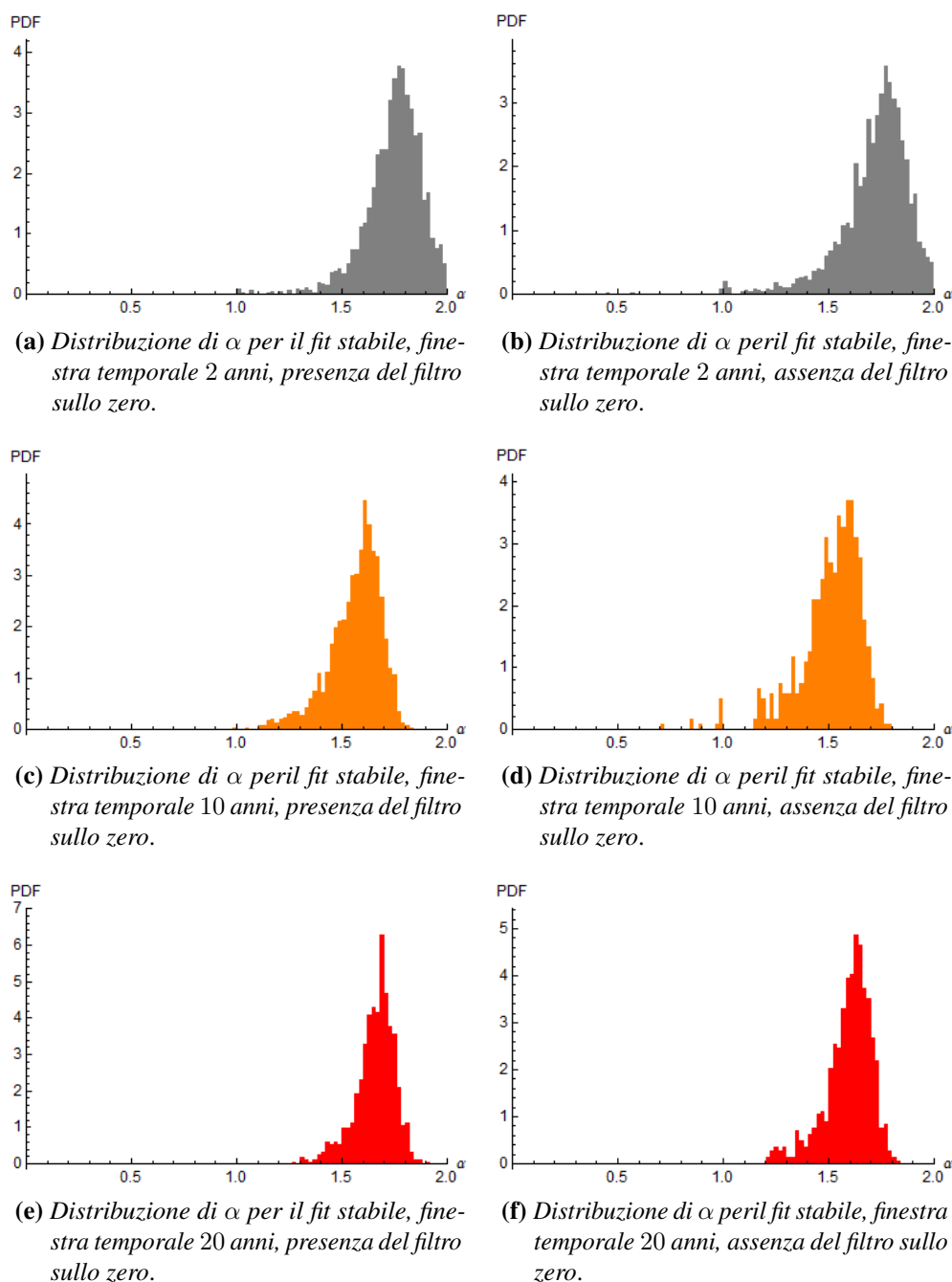


Figura 2.8. Distribuzione empirica del parametro di stabilità α derivante dai *fit* stabili effettuati sulle serie storiche 2, 10, 20 anni, con e senza filtro sui rendimenti nulli. α (insieme a β) determina la forma della distribuzione definendo lo spessore delle code e il valore della moda. Per una distribuzione Lévy-Pareto stabile si ha: $0 < \alpha \leq 2$. Si ricorda che nel caso limite $\alpha = 2$, la distribuzione è gaussiana. I valori assunti dal parametro nei diversi *fit* sono coerenti con l'ipotesi di distribuzione stabile, poiché sono in media compresi tra 1.5 e 1.8.

disponga di una serie storica di dati limitata, ad esempio nel caso di una IPO (*Initial Public Offering*).

È possibile stabilire a priori il valore di α da assegnare a un titolo azionario? Utilizzando il vasto campione di dati a disposizione è probabile che esista un parametro velocemente reperibile o noto con un legame forte e univoco con α ? L'esistenza di un indicatore con tali caratteristiche permetterebbe di stabilire *ex-ante* il valore del parametro di stabilità del titolo considerato, eliminando la necessità di stime storiche.

L'indicatore scelto per la verifica empirica delle ipotesi poste sopra è l'appartenenza settoriale della società emittente.

2.3.1 Classificazione settoriale

Il sito www.yahoofinance.com, sorgente dei dati analizzati, utilizza una classificazione industriale, che suddivide il mercato in macro-settori (primo livello) e settori (secondo livello). YahooFinance (YF) individua 9 macro-settori, divisi al loro interno in 215 settori totali.[§] Lo scopo di tale tassonomia industriale è la classificazione delle imprese in base alla principale fonte di reddito, il cosiddetto *core business*.

Per chiarezza d'esposizione, nei paragrafi a seguire, verrà utilizzata esclusivamente la classificazione YF di primo livello. Tale scelta, pur diminuendo la precisione dei risultati ottenuti, non compromette la validità del metodo statistico utilizzato. A ciascun macro-settore YF assegna un codice compreso tra 100 e 900:

- 100, *Basic Materials*;
- 200, *Conglomerates*;
- 300, *Consumer Goods*;
- 400, *Financial*;

[§]A partire dalla classificazione YF è stata ottenuta la classificazione ICB (*Industry Classification Benchmark*) dei titoli, sviluppata dal Dow Jones e dal FTSE e utilizzata globalmente. I risultati delle analisi di *clustering* settoriale rispetto a tale tassonomia industriale non saranno qui pubblicati poiché del tutto coerenti con gli esiti derivanti dall'uso della classificazione YF.

- 500, *Healthcare*;
- 600, *Industrial Goods*;
- 700, *Services*;
- 800, *Technology*;
- 900, *Utilities*.

I grafici a torta della figura 2.9 nella pagina successiva mostrano come è variata la suddivisione settoriale del campione nei tre orizzonti temporali considerati: 2, 10, 20 anni. Comparando l'orizzonte temporale più breve, 2 anni, con quello più ampio, 20 anni, è possibile notare l'aumento di titoli tecnologici e finanziari nel mercato, avvenuto a discapito delle società produttrici di beni di consumo e *utilities*. Tale cambiamento è avvenuto non solo per il campione di dati analizzato, bensì a livello globale nel mercato: si pensi, per esempio, al boom di ingressi nei mercati regolamentati da parte di titoli tecnologici (e finanziari) durante la cosiddetta “*new economy*”. Questa corrispondenza fornisce, ancora una volta, la conferma che il campione scelto è rappresentativo della popolazione di riferimento: i rendimenti della totalità dei titoli azionari trattati mondialmente.

2.3.2 Distribuzioni α -settoriali

In questo paragrafo si indaga l'esistenza di un eventuale legame tra i valori assunti dal parametro di stabilità α e l'appartenenza settoriale del titolo. ¶

I diversi valori di α derivanti dai *fit* stabili calcolati sulle singole serie storiche sono raggruppati per classificazione industriale assegnata all'attività finanziaria (YF). A partire dalle frequenze di α osservate nelle classi YF sono calcolate le distribuzioni empiriche del parametro di stabilità, una per ciascuno dei 9 macro-settori, definite distribuzioni α -settoriali.

Le figure 2.10- 2.12 a pagina 58 illustrano graficamente i risultati ottenuti. Sull'asse delle ascisse è indicato il settore industriale, mentre sull'asse delle ordinate è

¶ Sono utilizzati esclusivamente i valori di α derivanti da *fit* calcolati sulle serie storiche dei rendimenti con filtro sui rendimenti nulli. Per chiarimenti si rimanda al paragrafo 2.2.3 a pagina 39.

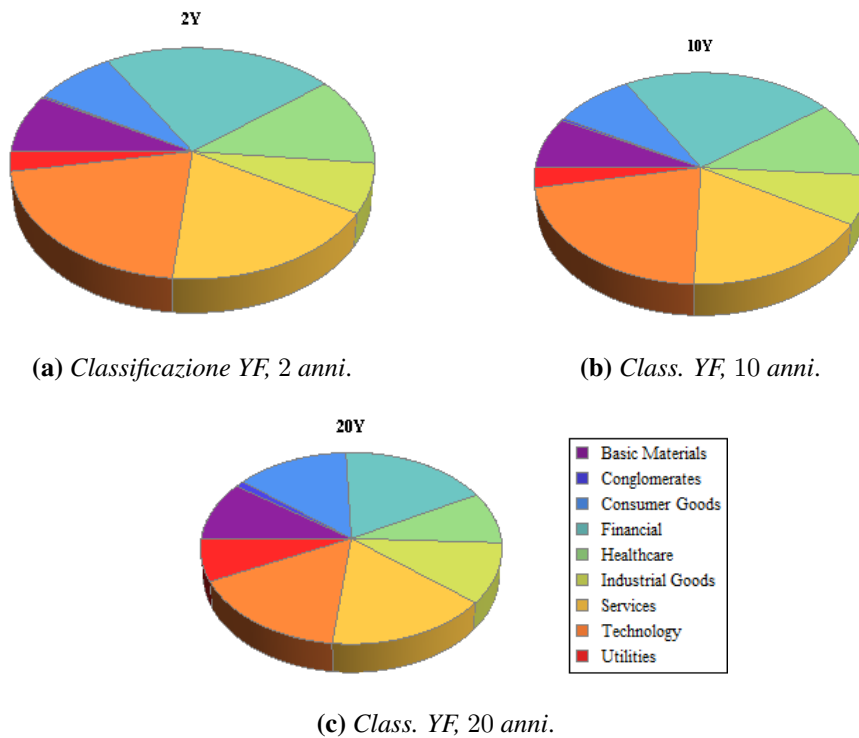


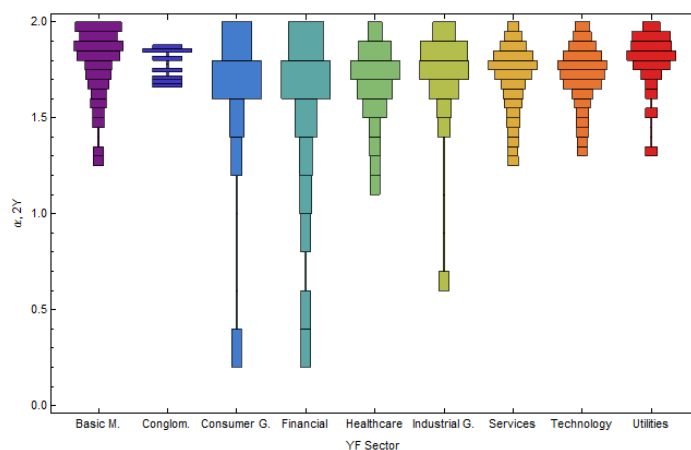
Figura 2.9. Classificazione settoriale del campione di titoli considerato mediante la tassonomia industriale fornita dal sito www.yahoofinance.com (YF). Le diverse dimensioni delle “torte” vogliono ricordare al lettore che la numerosità del campione analizzato varia a seconda della finestra temporale considerata: 3050 titoli (a), 2226 titoli (b), 755 titoli (c). Comparando l’orizzonte temporale più breve, 2 anni, con quello più ampio, 20 anni, è possibile notare l’aumento di titoli tecnologici e finanziari nel mercato, avvenuto a discapito delle società produttrici di beni di consumo e *utilities*.

mostrato α . La larghezza delle barre orizzontali indica la numerosità dei titoli per i quali il calcolo della stima stabile ha reso valori di α uguali. Di fatto le figure citate sono la rappresentazione grafica delle PDF del parametro di stabilità per settore; diversamente dai grafici usuali, le densità sono mostrate in verticale.

È interessante notare che, ampliando l'orizzonte temporale considerato, le PDF di α sembrano “accorciarsi” (scala sull'asse delle ordinate); cioè i valori del parametro di stabilità sono meno variabili, più bassi e concentrati attorno al valor medio per i *fit* effettuati sulle serie storiche ventennali rispetto a quelle biennali o decennali.

Nel grafico a 2 anni, α assume valori prossimi a due avvicinandosi a una distribuzione normale dei dati; è tuttavia visibile una marcata variabilità nei dati con parametri di stabilità anche vicini al valore minimo $\alpha = 0$ (Financial e Consumer Goods). Nel grafico decennale, dove gli eventi estremi non contemplati da una distribuzione gaussiana iniziano a manifestarsi, il parametro di stabilità mostra valori minori, mediamente compresi tra 1.5 e 1.7. Tuttavia eventi molto rari (M&A, LBO), che hanno probabilmente interessato solo una piccola parte del campione considerato, causano una forte dispersione dei dati. Questa problematica può facilmente essere superata ampliando l'orizzonte temporale. Considerando 20 anni di storia è plausibile pensare di ottenere stime migliori di α : le serie storiche ventennali dei log-rendimenti incorporano un numero sufficiente di shock ed eventi estremi che portano il parametro di stabilità ad assestarsi in una fascia ristretta intorno al valore $\alpha = 1.6$.

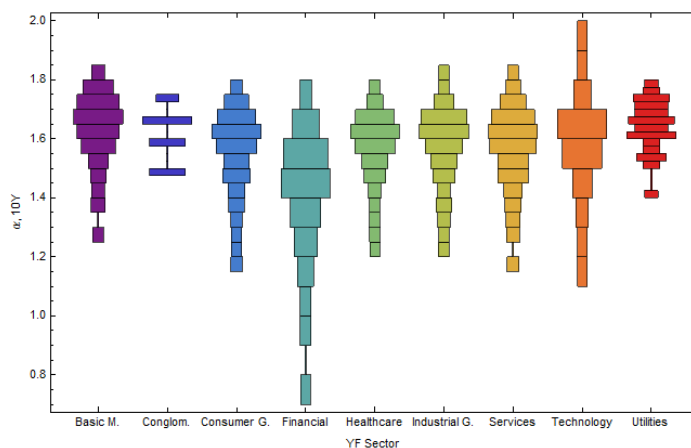
Un'ulteriore particolarità emergente dai grafici è che alcuni settori sembrano esibire comportamenti mediamente simili nella distribuzione di α , mentre altri, tendono a distaccarsi presentando caratteristiche diverse. Un veloce esame delle statistiche descrittive relative alle distribuzioni α -settoriali conferma l'idea che sia possibile formare dei gruppi nei quali porre le PDF ritenute simili. Questo processo, ovvero la suddivisione di un campione, tramite una metrica assegnata, in sottoinsiemi accomunati da una particolare caratteristica è definito *clustering*.

(a) *Classificazione YF, 2 anni.*

2Y	Basic M.	Conglom.	Consumer G.	Financial	Healthcare	Industrial G.	Services	Technology	Utilities
Mean	1.831	1.775	1.740	1.694	1.723	1.769	1.744	1.750	1.829
Median	1.853	1.788	1.752	1.734	1.743	1.785	1.764	1.756	1.832
Std. Dev.	0.123	0.080	0.140	0.210	0.120	0.136	0.117	0.107	0.097

(b) *Statistiche descrittive, 2Y.*

Figura 2.10. Distribuzione del parametro di stabilità per settore industriale. La classificazione settoriale adottata è quella del sito Yahoo Finance, sorgente dei dati. I valori di α utilizzati derivano dai *fit* stabili effettuati sulle serie storiche con orizzonte temporale biennale.

(a) *Classificazione YF, 10 anni.*

10Y	Basic M.	Conglom.	Consumer G.	Financial	Healthcare	Industrial G.	Services	Technology	Utilities
Mean	1.646	1.603	1.574	1.448	1.594	1.586	1.576	1.592	1.649
Median	1.662	1.614	1.597	1.467	1.607	1.610	1.594	1.605	1.659
Std. Dev.	0.096	0.090	0.112	0.143	0.086	0.108	0.114	0.090	0.074

(b) *Statistiche descrittive, 10Y.*

Figura 2.11. Distribuzione del parametro di stabilità per settore industriale. La classificazione settoriale adottata è quella del sito Yahoo Finance, sorgente dei dati. I valori di α utilizzati derivano dai *fit* stabili effettuati sulle serie storiche con orizzonte temporale decennale.

2.3.3 Test di Kruskal-Wallis

Il processo di *clustering* settoriale per il parametro α sarà effettuato avvalendosi del test statistico di uguaglianza delle mediane elaborato da William Kruskal e W. Allen Wallis, dai quali prende il nome. Tale paragrafo si propone di illustrare le caratteristiche del test e spiegarne le modalità di applicazione nell'ambito dell'indagine condotta.

Il *Test di Kruskal-Wallis* (KW) è un metodo non parametrico per verificare se diversi gruppi campionari provengono dalla medesima popolazione di riferimento o da popolazioni diverse con uguale mediana. Nella statistica parametrica i modelli matematici non necessitano di ipotesi a priori sulle caratteristiche della popolazione; in particolare, non si assume l'ipotesi che i dati provengono da una distribuzione normale. Questa particolarità si rivela fondamentale per l'impiego del test nell'analisi delle distribuzioni α -settoriali, poiché non esistono dati a disposizione per assumere a priori una distribuzione gaussiana della popolazione. Le ipotesi sottoposte a verifica dalla statistica sono:

- H_0 : i parametri di locazione dei gruppi sono coincidenti;
- H_1 : almeno un gruppo ha parametro di locazione differente.

I dati sono sostituiti dal loro *rango*, ovvero sono posti in ordine crescente a prescindere dal gruppo di appartenenza e a ogni dato ordinato è assegnato un numero intero ($r = 1, 2, \dots$), detto appunto rango. La statistica è calcolata come:

$$K = (N - 1) \frac{\sum_{i=1}^k n_i (\bar{r}_i - \bar{r})^2}{\sum_{i=1}^k \sum_{j=1}^{n_i} (r_{ij} - \bar{r})^2};$$

dove

- n_i è il numero di osservazioni nel gruppo i ;
- r_{ij} è il rango della j -esima osservazione nel gruppo i ;
- N è il numero totale delle osservazioni;
- $\bar{r}_{ij} = \frac{1}{n_i} \sum_{j=1}^{n_i} r_{ij}$ è la media dei ranghi nel gruppo i ;

- $\bar{r} = \frac{1}{2}(N + 1)$ è la media complessiva dei ranghi.

Tale statistica si distribuisce asintoticamente secondo una variabile χ^2 con $(k - 1)$ gradi di libertà. L'ipotesi nulla di uguaglianza delle mediane è rifiutata se

$$K \geq \chi_{a, (k-1)}^2,$$

dove il secondo termine della disuguaglianza rappresenta il valore di una distribuzione $\chi_{(k-1)}^2$ con livello di significatività pari ad a .

Ad ogni valore assunto da K è assegnato il corrispondente valore di probabilità o *p-value*. Si accetta l'ipotesi di uguaglianza delle mediane se il valore di probabilità ottenuto dal test è superiore al livello di significatività prefissato. In altre parole, l' H_0 viene rifiutata se la probabilità che due (o più) distribuzioni abbiano mediana identica è minore di a . Accettare l'ipotesi nulla significa affermare che le distribuzioni considerate sono di fatto uguali o provengono dalla medesima popolazione.

Uso del Test di KW – *Clustering*

I gruppi campionari oggetto di analisi sono le 9 distribuzioni α -settoriali, figure 2.10- 2.12 a pagina 58. Nella presentazione dei risultati quest'ultime saranno indicate semplicemente usando il nome del settore (Basic Materials, Industrial Goods, ...). Il test di KW è applicato ad ogni coppia di distribuzioni, per un totale di 45 coppie. Ovviamente, i gruppi contenenti le distribuzioni appartenenti al medesimo settore mostrano valori della statistica $K = 0$ (Basic Materials e Basic Materials, ...).

Come per il Test χ_2 di Pearson, è utilizzata una trasformazione logaritmica del valore di probabilità: $g = -\log_{10}(p\text{-value})$. La trasformazione rende valori di g positivi e crescenti per *p-value* via via più piccoli. Per valori di g superiori a $-\log_{10}(a)$ viene rifiutata l'ipotesi nulla di mediane coincidenti.

Nel processo di interpretazione dei risultati del test non è fissato a priori un livello di significatività, cioè la verifica statistica di uguaglianza delle mediane nelle coppie di distribuzioni non porta a un rifiuto o un accoglimento dell'ipotesi nulla. I valori di g ottenuti sono usati quali metrica nella procedura di *clustering*.

Il processo di *clustering* avviene, in genere, mediante l'impiego di una funzione di distanza ν (ad esempio la funzione di distanza Euclidea), che considera le coppie di elementi come simili quando il valore di ν è basso. ^{||} La funzione di distanza utilizzata è la trasformazione logaritmica del *p-value*: $\nu = g_{j,i}$, con $j, i = 1, 2, \dots, 9$, coppie di distribuzioni α -settoriali. In questo senso i valori di g sono interpretabili come la distanza esistente tra le mediane delle due distribuzioni j e i :

- alti valori di $g_{j,i}$ indicano valori mediani lontani (diversi) per la coppia di distribuzioni j, i ;
- bassi valori di $g_{j,i}$ indicano valori mediani vicini o coincidenti per la coppia di distribuzioni j, i .

All'interno di ciascun *cluster* sono state posizionate le distribuzioni per le quali il valore di g è minimo, ovvero è massimizzata la probabilità che provengano dalla stessa popolazione di riferimento. È stata fatta la scelta di dividere il campione in 4 gruppi, allo scopo di ottenere sottoinsiemi contenenti ciascuno 2, 3 distribuzioni (per un totale di 9 partizioni).

Le tabelle 2.13- 2.15 a pagina 61 mostrano i risultati del test di KW per le tre finestre temporali 2, 10, 20 anni. I colori indicano i gruppi che minimizzano i valori di probabilità. I valori nulli di g (derivanti dall'applicazione del test su coppie contenenti settori coincidenti) non sono riportati.

È interessante sottolineare che, nonostante il livello di significatività α non sia stato assegnato a priori, per ogni *cluster* e ciascun orizzonte temporale è possibile accettare l' H_0 di uguaglianza delle mediane con $\alpha > 0.5$, cioè la probabilità che le distribuzioni del gruppo siano statisticamente uguali supera il 50%.

Come si può notare dalle tabelle aumentando o diminuendo la lunghezza del periodo di osservazione cambia la composizione dei *cluster*. Tale problematica persiste anche nel caso di uso della classificazione settoriale YF di secondo livello (215 settori). Per applicazioni pratiche dei risultati è necessario scegliere un singolo orizzonte temporale, tralasciando i restanti due, e ponendo ipotesi di omogeneità

^{||}La funzione di distanza Euclidea misura la distanza fra due punti, ovvero la lunghezza del segmento con estremi i due punti considerati.

nel tempo del *core business* aziendale. Se, per esempio, si decide di considerare il *clustering* ventennale si potrebbe obiettare che dal 1991 ad oggi i settori hanno subito profondi cambiamenti e la classificazione di allora non rispecchia quella attuale. A compensare lo svantaggio illustrato si richiamano le considerazioni esposte nel paragrafo 2.3.2 a pagina 49 sulla superiorità della stima del parametro di stabilità sull'orizzonte temporale di 20 anni. Shock di mercato ed eventi estremi, responsabili della leptocurtosi dei rendimenti finanziari, sono difficilmente osservabili in intervalli ristretti (2, 10 anni), impedendo stime di α sufficientemente affidabili.

Applicazione pratica

Per illustrare una possibile applicazione pratica dei risultati raggiunti, verrà utilizzata la serie storica dei log-rendimenti di Foot-Locker Inc. (FL), 2.16 a pagina 62.

Il lettore si immagini di avere a disposizione solamente la serie storica biennale del titolo FL, riquadro (a).

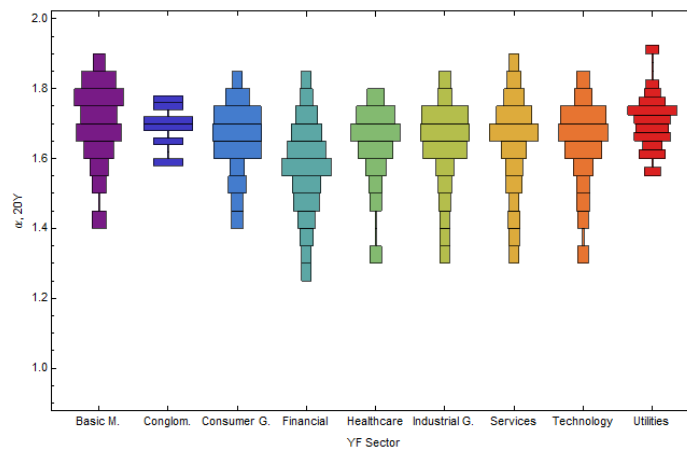
Il problema è il seguente: calcolare la migliore stima Lévy-Pareto stabile in grado di modellare la distribuzione empirica effettiva dei rendimenti. Il *fit* stabile determinato sui rendimenti degli ultimi 2 anni è $\mathbf{S}(\alpha = 1.866, \beta = -0.109, \gamma = 0.15, \delta = 0.002; 1)$ mostrato in (b), linea arancione. Una stima di questo tipo risulta tuttavia poco significativa: le code decadono molto velocemente poiché nel periodo considerato non si sono verificati eventi rari o di forte impatto sui rendimenti, tali da poter formare le cosiddette code spesse.

Come è risolvibile questo errore di valutazione? Gli studi compiuti suggeriscono di sostituire il parametro di stabilità, ottenuto dal *fit* biennale, con il valore di α del *cluster* di riferimento. Utilizzando i risultati del processo di *clustering* sull'orizzonte temporale di 20 anni, 2.15 a pagina 61, è possibile collocare il titolo nel secondo gruppo, Cluster 2. Il valore del parametro di stabilità $\alpha = 1.866$ è quindi sostituito da $\bar{\alpha}_C = 1.647$, cioè la media aritmetica dei *mean-value* di α nel *cluster* di appartenenza. Il *fit* $\mathbf{S}(\bar{\alpha}_C)$ con parametri β, γ, δ invariati è mostrato nel riquadro (b) linea rossa.

Per verificare i benefici della correzione applicata, entrambe le stime, $\mathbf{S}(\bar{\alpha}_C)$ e $\mathbf{S}(\alpha)$, sono comparate con la distribuzione dei rendimenti con periodo di osservazione

ventennale, (c). È immediato notare che $\mathbf{S}(\bar{\alpha}_C)$, sebbene prenda in considerazione esclusivamente le serie storiche con intervallo 2 anni, permetta di spiegare in maniera puntuale la densità di probabilità dei rendimenti degli ultimi 20 anni, misurando precisamente l'influsso di accadimenti anomali sulle code della distribuzione, (d).

Questo esempio pratico dovrebbe chiarire al lettore il possibile uso dei risultati raggiunti. Il *clustering* α -settoriale può diventare un'utile strumento nella correzione di stime stabili laddove siano carenti i dati storici a disposizione.

(a) *Classificazione YF, 20 anni.*

20Y	Basic M.	Conglom.	Consumer G.	Financial	Healthcare	Industrial G.	Services	Technology	Utilities
Mean	1.712	1.692	1.663	1.584	1.659	1.656	1.665	1.666	1.701
Median	1.721	1.697	1.672	1.590	1.676	1.670	1.684	1.677	1.706
Std. Dev.	0.090	0.056	0.078	0.101	0.0754	0.092	0.096	0.080	0.064

(b) *Statistiche descrittive, 20Y.*

Figura 2.12. Distribuzione del parametro di stabilità per settore industriale. La classificazione settoriale adottata è quella del sito Yahoo Finance, sorgente dei dati. I valori di α utilizzati derivano dai *fit* stabili effettuati sulle serie storiche con orizzonte temporale ventennale.

2Y	Utilities	Technology	Services	Industrial G.	Healthcare	Financial	Consumer G.	Conglomerates
Basic Materials	0.478	24.949	24.718	8.787	29.944	30.958	18.982	1.184
Conglomerates	1.160	0.281	0.264	0.045	0.685	0.611	0.318	
Consumer G.	11.865	0.169	0.256	2.495	1.493	1.717		
Financial	13.766	3.873	3.466	6.480	0.175			
Healthcare	16.448	2.982	2.881	6.361				
Industrial G.	5.229	2.720	2.471					
Services	12.284	0.126						
Technology	12.736							

(a) Test di Kruskal-Wallis, finestra temporale 2 anni.

2Y	Sector	Samples	Mean	Median	Std. Dev.	Skewness	Kurtosis
Cluster 1	Basic Materials	251	1.832	1.853	0.123	-1.190	5.252
	Utilities	84	1.829	1.832	0.097	-2.263	12.070
Cluster 2	Conglomerates	8	1.775	1.788	0.079	-0.221	1.395
	Consumer G.	245	1.740	1.752	0.140	-4.129	39.592
	Industrial G.	212	1.769	1.785	0.136	-3.031	25.162
Cluster 3	Services	550	1.744	1.764	0.117	-1.015	4.423
	Technology	630	1.750	1.756	0.107	-0.564	3.866
Cluster 4	Financial	690	1.694	1.734	0.210	-2.497	14.218
	Healthcare	380	1.723	1.743	0.120	-0.723	4.279

(b) Clustering, finestra temporale 2 anni.

Figura 2.13. Clustering settoriale per il parametro α mediante l'uso del Test di Kruskal-Wallis di uguaglianza delle mediane. Finestra temporale 2 anni. In (a) sono presentati i valori di probabilità $g = -\log_{10}(p\text{-value})$ derivanti dall'applicazione del test, per ciascuna delle coppie possibili di distribuzioni α -settoriali. Il processo di *clustering* avviene tramite l'uso di g quale funzione di distanza. Alti valori di g indicano distribuzioni α -settoriali con mediane lontane (diverse); bassi valori di g indicano distribuzioni con mediane vicine (simili). I colori evidenziano i gruppi che minimizzano g , presentati con le relative statistiche descrittive nel riquadro (b)

10Y	Utilities	Technology	Services	Industrial G.	Healthcare	Financial	Consumer G.	Conglomerates
Basic Materials	0.065	11.871	12.402	7.510	9.725	63.669	10.333	0.818
Conglomerates	0.990	0.109	0.223	0.110	0.081	2.901	0.248	
Consumer G.	6.847	0.646	0.048	0.560	0.790	30.415		
Financial	30.988	75.524	47.418	32.533	55.199			
Healthcare	6.261	0.158	0.858	0.002				
Industrial G.	5.086	0.095	0.554					
Services	7.150	0.751						
Technology	6.876							

(a) Test di Kruskal-Wallis, finestra temporale 10 anni.

10Y	Sector	Samples	Mean	Median	Std. Dev.	Skewness	Kurtosis
Cluster 1	Basic Materials	174	1.646	1.662	0.096	-0.892	3.993
	Utilities	68	1.649	1.659	0.074	-0.586	3.107
Cluster 2	Conglomerates	8	1.603	1.614	0.090	-0.175	1.891
	Healthcare	258	1.594	1.606	0.086	-1.099	5.164
	Industrial G.	167	1.586	1.610	0.108	-1.054	4.094
	Technology	475	1.592	1.605	0.090	-0.686	4.671
Cluster 3	Consumer G.	193	1.574	1.597	0.112	-1.111	4.616
	Services	378	1.576	1.594	0.114	-0.951	4.089
Cluster 4	Financial	505	1.445	1.467	0.143	-0.842	4.223

(b) Clustering, finestra temporale 10 anni.

Figura 2.14. Clustering settoriale per il parametro α mediante l'uso del Test di Kruskal-Wallis di uguaglianza delle mediane. Finestra temporale 10 anni. In (a) sono presentati i valori di probabilità $g = -\log_{10}(p\text{-value})$ derivanti dall'applicazione del test, per ciascuna delle coppie possibili di distribuzioni α -settoriali. Il processo di clustering avviene tramite l'uso di g quale funzione di distanza. Alti valori di g indicano distribuzioni α -settoriali con mediane lontane (diverse); bassi valori di g indicano distribuzioni con mediane vicine (simili). I colori evidenziano i gruppi che minimizzano g , presentati con le relative statistiche descrittive nel riquadro (b)

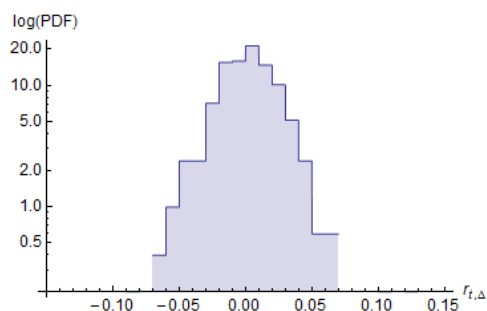
20Y	Utilities	Technology	Services	Industrial G.	Healthcare	Financial	Consumer G.	Conglomerates
Basic Materials	0.338	4.398	3.435	4.185	4.425	16.144	4.652	0.257
Conglomerates	0.146	0.526	0.313	0.637	0.699	2.497	0.510	
Consumer G.	2.241	0.088	0.327	0.092	0.168	10.459		
Financial	12.620	12.742	11.759	7.935	8.443			
Healthcare	2.685	0.279	0.511	0.020				
Industrial G.	2.251	0.257	0.429					
Services	1.521	0.180						
Technology	2.024							

(a) Test di Kruskal-Wallis, finestra temporale 20 anni.

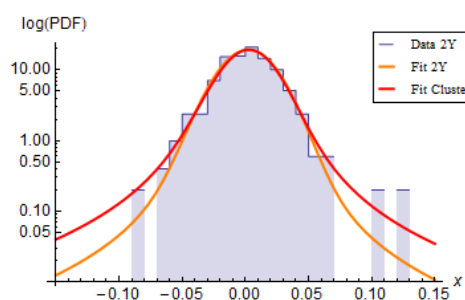
20Y	Sector	Samples	Mean	Median	Std. Dev.	Skewness	Kurtosis
Cluster 1	Basic Materials	75	1.712	1.721	0.090	-0.956	4.300
	Utilities	54	1.701	1.706	0.064	0.255	3.774
	Conglomerates	8	1.692	1.697	0.056	-0.554	2.605
Cluster 2	Consumer G.	101	1.663	1.672	0.078	-0.826	3.835
	Services	120	1.665	1.684	0.096	-1.190	4.930
	Technology	121	1.666	1.677	0.080	-1.261	6.266
Cluster 3	Healthcare	66	1.659	1.676	0.075	-1.724	8.538
	Industrial G.	77	1.656	1.670	0.092	-1.320	5.622
Cluster 4	Financial	133	1.584	1.590	0.101	-0.356	3.129

(b) Clustering, finestra temporale 20 anni.

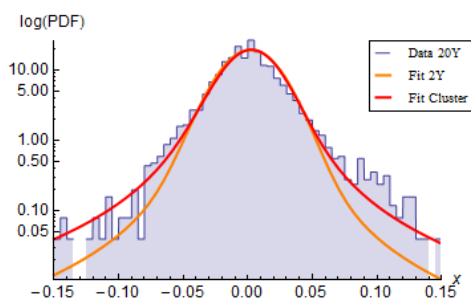
Figura 2.15. Clustering settoriale per il parametro α mediante l'uso del Test di Kruskal-Wallis di uguaglianza delle mediane. Finestra temporale 20 anni. In (a) sono presentati i valori di probabilità $g = -\log_{10}(p\text{-value})$ derivanti dall'applicazione del test, per ciascuna delle coppie possibili di distribuzioni α -settoriali. Il processo di clustering avviene tramite l'uso di g quale funzione di distanza. Alti valori di g indicano distribuzioni α -settoriali con mediane lontane (diverse); bassi valori di g indicano distribuzioni con mediane vicine (simili). I colori evidenziano i gruppi che minimizzano g , presentati con le relative statistiche descrittive nel riquadro (b)



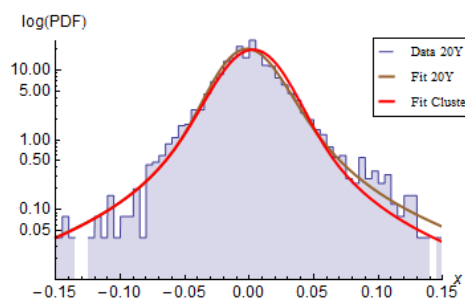
(a) Log Plot, distribuzione empirica Foot Locker Inc., 2 anni.



(b) Log Plot, Foot Locker Inc., distribution fitting, 2 anni.



(c) Log Plot, Foot Locker Inc. distribution fitting, 20 anni.



(d) Log Plot, Foot Locker Inc. distribution fitting, 20 anni.

Figura 2.16. Uso del *clustering* α settoriale per correggere errori di stima del parametro di stabilità. In (b) sono illustrati in arancione $\mathbf{S}(\alpha = 1.866, \beta = -0.109, \gamma = 0.15, \delta = 0.002; 1)$ e in rosso $\mathbf{S}(\bar{\alpha}_C = 1.647, \beta = -0.109, \gamma = 0.15, \delta = 0.002; 1)$ determinati a partire dalle serie storiche biennali di Foot-Locker Inc. (FL). $\bar{\alpha}_C$ è la media aritmetica dei *mean-value* di α nel *cluster* (ventennale) di appartenenza del titolo. In (c), (d) le stime vengono comparate con la distribuzione a 20 anni dei rendimenti e il relativo *fit* stabile (linea marrone). $\mathbf{S}(\bar{\alpha}_C)$, nonostante sia calcolata a partire dalla distribuzione biennale di FL, spiega in modo efficiente la serie storica, anche sul periodo di osservazione di 20 anni, includendo eventi estremi e shock di mercato non ancora manifestatisi nell'intervallo di più breve periodo.

CONCLUSIONI

Al termine delle analisi presentate in questa tesi, effettuate su un vasto campione di titoli azionari statunitensi, è possibile riassumere brevemente i risultati raggiunti.

In primo luogo, è stato verificato che i rendimenti giornalieri logaritmici di titoli azionari sono modellabili tramite una distribuzione Lévy-Pareto stabile con esponente α positivo e minore di 2. Tale ipotesi è stata comprovata mediante l'uso del Test χ^2 di Pearson, che ha mostrato la superiorità descrittiva della classe di leggi stabili rispetto alla tradizionale distribuzione gaussiana, a prescindere dal periodo di osservazione scelto (2, 10, 20 anni). Il processo di *distribution fitting* dei rendimenti è stato poi ulteriormente migliorato grazie all'applicazione di un filtro sulle serie storiche oggetto di analisi che ha permesso di eliminare i rendimenti nulli "fittizi", derivanti, più che da variazioni di prezzo nulle, da giornate di non trattazione del titolo. Per le stime, stabili e normali, è stato utilizzato il metodo della massima verosimiglianza.

Accertata la bontà del modello stabile, il *focus* della ricerca è stato spostato sul parametro α , responsabile della leptocurtosi della distribuzione. L'idea era di trovare un carattere del titolo esaminato noto a priori che permettesse, senza necessità di stime storiche, di collegare ad esso uno specifico valore di α . Quale indicatore è stata scelta l'appartenenza settoriale del titolo. Sono stati incrociati i valori assunti da α nei singoli *fit* stabili effettuati, con il numero rappresentativo della classificazione industriale, ottenendo delle distribuzioni α -settoriali. Il campione è stato suddiviso

in *cluster*, utilizzando quale metrica di suddivisione il Test di Kruskal-Wallis di uguaglianza delle mediane; a ogni gruppo è stato possibile collegare un valore medio del parametro di stabilità ($\bar{\alpha}_C$). È stato dimostrato, mediante un esempio, che l'uso di $\bar{\alpha}_C$ permette di cogliere gran parte della rischiosità del titolo osservato, cioè la quasi totalità degli eventi estremi ed anomali verificabili in vent'anni di trattazione dell'azione, pur disponendo di serie storiche di brevissimo periodo (2 anni).

La ricerca effettuata non ha trovato, ad oggi, corrispondenze in letteratura e si presta pertanto a spunti, approfondimenti e interessanti applicazioni pratiche nell'ambito dell'*Asset Management*. Si ricorda tuttavia al lettore che il livello di approfondimento degli esiti conseguiti è da considerarsi non esaustivo.

BIBLIOGRAFIA

- [1] L. Bachelier (1900), *Theory of Speculation*, in P. Cootner (1964), *The Random Character of Stock Market Prices*, Cambridge University, MIT Press, MA.
- [2] S. Borak, W. Härdle, R. Weron (2005), *Stable Distributions*, SFB 649 “Economic Risk”, Wiwi Fakultät, Humboldt-Universität zu Berlin, <http://edoc.hu-berlin.de/series/sfb-649-papers/2005-8/PDF/8>.
- [3] R. Cont, M. Potters, J. Bouchard (1997), *Scaling in stock market data: stable laws and beyond*, SSRN: <http://ssrn.com/abstract=40555>.
- [4] E. Castagnoli, M. Cigola, L. Peccati (2009), *Probability. A brief introduction*, Egea, Milano.
- [5] R. Donati (2005), *Studio e sviluppo di un modello di spin finalizzato all'asset management*, Tesi di Laurea, Università degli studi di Padova.
- [6] E.F. Fama (1963), *Mandelbrot and the stable Paretian hypothesis*, Journal of Business, n. 36.
- [7] E.F. Fama (1965), *The behavior of stock market prices*, Journal of Business, n. 38.
- [8] H. Fofack, J.P. Nolan (1998), *Tail Behavior, Modes and Other Characteristics of Stable Distributions*, Math/Stat Department, American University, academic2.american.edu/~jpnolan.

- [9] B.V. Gnedenko, A.N. Kokmogorov (1954) *Limit distributions for sums of independent random variables*, Addison-Wesley Pub. Co., Cambridge, MA.
- [10] M. Haas, C. Pigorsch (2009) *Financial Economics, Fat tailed Distributions*, Encyclopedia of Complexity and Systems Science, vol. 4, Springer-Verlag, Heidelberg.
- [11] M. Loreti (2006), *Teoria degli Errori e Fondamenti di Statistica–Introduzione alla Fisica Sperimentale*, Dipartimento di Fisica, Università degli studi di Padova, <http://wwwcdf.pd.infn.it/labo/INDEX.html>.
- [12] B. Mandelbrot (1963), *New methods in statistical economics*, Journal of Political Economy, n. 71.
- [13] B. Mandelbrot (1963), *The variation of certain sperculative prices*, Journal of Business, n. 36.
- [14] B. Mandelbrot (1967), *The variation of some other sperculative prices*, Journal of Business, n. 40.
- [15] J.P. Nolan (2005), *Modeling financial data with stable distributions*, Math/Stat Department, American University, academic2.american.edu/~jpnolan.
- [16] J.P. Nolan (2012), *Stable Distributions-Models for Heavy Tailed Data*, Birkhauser, Boston. Libro non terminato, academic2.american.edu/~jpnolan.
- [17] S.T. Rachev (2003), *Handbook of heavy tailed distributions in finance–Libro 1 di Handbooks in finance*, Elsevier, Amsterdam.

RINGRAZIAMENTI

Ringrazio in primo luogo chi ha effettivamente contribuito alla stesura di questa tesi: il Professore Gino Favero (relatore) e il Dott. Riccardo Donati (correlatore) senza l'aiuto dei quali queste sudate 67 pagine non sarebbero mai state scritte.

Ringrazio Wikipedia, e tutte le fonti di informazione libera online per la quantità e la qualità di conoscenze fornitemi.

Ringrazio (in qualità di capoclasse) i pochi e ultimi *FRIM-quants* con i quali ho condiviso lezioni, esami e pause pranzo negli scorsi due anni, sperando che la nuova tradizione delle cene di classe sia duratura nel tempo.

Ringrazio i professori che mi hanno saputo trasmettere passione per l'economia, per il sapere e per lo studio. Ringrazio anche quelli che non ne sono stati proprio capaci perché, ancora di più, mi hanno fatto apprezzare i primi.

Ringrazio la mia famiglia, che pur mal comprendendo tutto ciò che rientra nell'ambito finanziario, mi ha permesso di compiere gli studi scelti sostenendomi sempre (se non altro economicamente!).

Per ultimo, ma non meno importante, ringrazio Danilo per quanto ha sempre creduto in me e per l'orgoglio che vedo nei suoi occhi quando raggiungo un piccolo traguardo.

Alice Pisani