



**UNIVERSITA' DEGLI STUDI DI TORINO**

**DIPARTIMENTO DI SCIENZE ECONOMICO-SOCIALI E  
MATEMATICO-STATISTICHE**

**Corso di Laurea Magistrale in Economia e Management**

**TESI DI LAUREA**

**TITOLO**

*La natura frattale dei mercati finanziari: dal moto browniano al modello  
multifrattale con un'applicazione dell'analisi R/S ad aziende europee*

Relatore: Prof. Davide Canavesio

Correlatore: Prof. Luigi Bollani

Candidato:

Simone Bianco

Anno accademico: 2019/2020



# ABSTRACT

All'interno della tesi, si espone una nuova interpretazione del comportamento dei mercati finanziari, sfruttando i risultati della geometria frattale, che si contrappone alla teoria finanziaria classica, sviluppata da Louis Bachelier, all'inizio del 900'. La teoria classica consente di modellizzare le variazioni dei prezzi attraverso la matematica dei processi stocastici. In particolare, si ipotizzano oscillazioni indipendenti e normalmente distribuite, in un cammino casuale in tempo continuo, noto come moto browniano. Questa impostazione presenta alcune criticità, se applicata alle fluttuazioni reali, in quanto la dinamica evolutiva delle turbolenze, presenti nei mercati finanziari, sembra essere inadatta ad essere interpolata in una curva normale. La tesi ripercorre gli studi di Mandelbrot, allo scopo di elaborare una nuova teoria finanziaria, tramite l'uso dei frattali, oggetti geometrici autosimili con dimensione non intera, generati attraverso algoritmi di iterazione. Il nuovo modello dei mercati si basa su tre importanti pilastri, esaminati singolarmente nella tesi, le distribuzioni stabili, la dipendenza a lungo termine e il tempo di contrattazione. In ultimo, nella parte di ricerca, si usa il test non parametrico R/S per stimare l'esponente di Hurst, che quantifica la dipendenza a lungo periodo, per le variazioni di cinque aziende europee.

Parole chiave: teoria finanziaria classica, moto browniano, processi stocastici, random walk, mercato efficiente, oggetti frattali, dimensione frattale, auto similarità, leggi di potenza, distribuzioni stabili, dipendenza a lungo periodo, multifrattali, esponente di Hurst.



# Indice

<b>INTRODUZIONE</b>	13
---------------------	----

## **CAPITOLO I. I PROCESSI STOCASTICI E IL MOTO BROWNIANO**

<i>1.1 Le variabili aleatorie e le distribuzioni di probabilità</i>	16
<i>1.1.1 I momenti e l'equazione caratteristica</i>	19
<i>1.2 La distribuzione normale e il teorema del limite centrale</i>	22
<i>1.2.1 La proprietà di stabilità per una distribuzione</i>	26
<i>1.3 Introduzione ai processi stocastici</i>	27
<i>1.3.1 I processi stocastici con tempo discreto e il random walk monodimensionale</i>	28
<i>1.3.2 I processi di Markov e la definizione di martingala</i>	29
<i>1.4 Il moto browniano e i processi stocastici con tempo continuo</i>	32
<i>1.5 Cenni di calcolo stocastico: l'integrale Ito</i>	37
<i>1.6 Cenni di calcolo stocastico: il lemma di Ito</i>	42
<i>Appendice A. Una simulazione di cammino casuale: The Drunkard's Walk</i>	45

## **CAPITOLO II. LA TEORIA FINANZIARIA CLASSICA E IL MERCATO EFFICIENTE**

<i>2.1 L. Bachelier e il nuovo modo di concepire la finanza</i>	50
<i>2.2 La teoria del mercato efficiente e i livelli di efficienza informativa</i>	53
<i>2.2.1 Le specificazioni del modello del mercato efficiente</i>	55

2.2.2 <i>Le condizioni e le evidenze empiriche</i>	56
2.3 <i>La teoria di portafoglio di Markowitz</i>	57
2.3.1 <i>Il free risk e il teorema di separazione di Tobin</i>	61
2.4 <i>Il Capital Asset Pricing Model</i>	62
2.5 <i>L'Arbitrage Pricing Theory come alternativa al CAPM</i>	64
2.6 <i>Il modello di Black e Scholes</i>	66
2.6.1 <i>Le ipotesi e la derivazione dell'equazione di Black-Scholes</i>	68
2.7 <i>I limiti della teoria finanziaria classica</i>	72
Appendice B. <i>Il moto browniano geometrico</i>	74
Appendice C. <i>Le equazioni differenziali alle derivate parziali e l'equazione di diffusione</i>	75

## **CAPITOLO III. GLI OGGETTI FRATTALI E LA DIMENSIONE FRAZIONARIA**

3.1 <i>Introduzione e definizione del concetto di frattale</i>	84
3.1.1 <i>L'insieme di Mandelbrot</i>	89
3.2 <i>La dimensione frattale e la misura della linea costiera</i>	91
3.3 <i>La dimensione di Minkowski-Bouligand</i>	95
3.4 <i>La misura e la dimensione di Hausdorff</i>	97
3.5 <i>Esempi di insiemi frattali</i>	103
3.5.1 <i>L'insieme di Cantor</i>	103
3.5.2 <i>La curva di Koch</i>	106
3.5.3 <i>Il triangolo di Sierpinski</i>	111
Appendice D. <i>I sistemi di funzioni iterate e gli attrattori</i>	113

## **CAPITOLO IV. LA NUOVA TEORIA FINANZIARIA E IL MODELLO MULTIFRATTALE DEI MERCATI**

<i>4.1 I cartoni di Brown-Bachelier e i primi esempi di grafici frattali</i>	120
<i>4.2 L'analisi dei prezzi del cotone e le leggi di potenza</i>	124
<i>4.2.1 La legge di Zipf</i>	129
<i>4.2.2 La legge di Pareto</i>	132
<i>4.3 Le distribuzioni L-stabili e le code spesse</i>	134
<i>4.4 I processi anomali e i voli di Levy</i>	140
<i>4.5 Le variazioni del cotone e l'invarianza di scala</i>	143
<i>4.6 La dipendenza a lungo termine e l'esponente di Hurst</i>	145
<i>4.7 Il moto browniano frazionario</i>	149
<i>4.8 L'effetto Noè e l'effetto Giuseppe</i>	154
<i>4.9 il modello multifrattale del rendimento delle attività</i>	156
<i>4.9.1 Il tempo di contrattazione e i cartoni multifrattali</i>	157
<i>4.9.2 Le misure e i processi multifrattali</i>	159
<i>4.9.3 Lo spettro multifrattale e l'esponente di Hölder locale</i>	165
<i>4.9.4 Definizione e proprietà dell'MMAR</i>	167
<i>Appendice E. Il metodo della massima verosimiglianza</i>	170

## **CAPITOLO V. LA DIPENDENZA A LUNGO TERMINE NEL MERCATO EUROPEO**

<i>5.1 La Rescaled Range Analysis</i>	176
<i>5.2 Il dataset utilizzato</i>	179
<i>5.3 L'applicazione dell'analisi R/S per calcolare l'esponente di Hurst</i>	182
<i>5.4 I risultati</i>	185

<i>Appendice F. L'analisi R/S con Matlab</i>	188
<b>CONCLUSIONI</b>	191
<b>BIBLIOGRAFIA</b>	195



# Indice delle figure

<i>Fig. 1</i> PDF e CDF della distribuzione normale	23
<i>Fig. 2</i> Simulazioni del random walk	29
<i>Fig. 3</i> Traiettorie di un moto browniano	36
<i>Fig. 4</i> Approssimazione di un processo integrando continuo	40
<i>Fig. 5</i> Distanza media dall'origine dell'ubriaco	47
<i>Fig. 6</i> Simulazione della posizione dell'ubriaco dopo 100 passi	47
<i>Fig. 7</i> Frontiera per due titoli al variare di $\rho$	61
<i>Fig. 8</i> Capital Market Line	62
<i>Fig. 9</i> Variazioni del Dow Jones e del moto browniano	73
<i>Fig. 10</i> Soluzione fondamentale dell'equazione di diffusione	82
<i>Fig. 11</i> Felce frattale	86
<i>Fig. 12</i> Insieme di Mandelbrot	90
<i>Fig. 13</i> Relazione tra insieme di Mandelbrot e insiemi di Julia	90
<i>Fig. 14</i> Approssimazioni della lunghezza della costa della Gran Bretagna	94
<i>Fig. 15</i> Dimensione box-counting	96
<i>Fig. 16</i> $\delta$ -ricoprimento di una spirale	100
<i>Fig. 17</i> Misura di Hausdorff	102
<i>Fig. 18</i> Insieme di Cantor	104
<i>Fig. 19</i> Fiocco di neve di Koch	107
<i>Fig. 20</i> Curva di Koch casuale	110
<i>Fig. 21</i> Triangolo di Sierpinski	112
<i>Fig. 22</i> IFS del triangolo di Sierpinski	118
<i>Fig. 23</i> IFS con diverso insieme di partenza	118

<i>Fig. 24</i> Cartone finanziario di Brown-Bachelier	121
<i>Fig. 25</i> Cartone di Brown-Bachelier con casualità	122
<i>Fig. 26</i> Generatori per costruire cartoni multifrattali	123
<i>Fig. 27</i> Variazioni dei prezzi nei cartoni multifrattali	124
<i>Fig. 28</i> Distribuzione delle parole nella versione inglese di Robinson Crusoe	139
<i>Fig. 29</i> Grafico di Pareto della distribuzione della ricchezza	133
<i>Fig. 30</i> Funzione di densità standardizzata $N(0,1)$ , $C(0,1)$ e $L(0,1)$	138
<i>Fig. 31</i> Traiettorie di un volo di Levy	142
<i>Fig. 32</i> Grafico delle variazioni del cotone	144
<i>Fig. 33</i> Cartone frattale con discontinuità	145
<i>Fig. 34</i> Serie antipersistente $H = 0,1$ (in alto), serie browniana $H = 0,5$ (in mezzo) e serie persistente $H = 0,9$ (in basso)	148
<i>Fig. 35</i> Cartone frattale anti-persistente $H < \frac{1}{2}$	155
<i>Fig. 36</i> Cartone frattale persistente $H > \frac{1}{2}$	155
<i>Fig. 37</i> Costruzione del generatore multifrattale	158
<i>Fig. 38</i> Cubo del mercato multifrattale	159
<i>Fig. 39</i> Cascata moltiplicativa della misura binomiale	161
<i>Fig. 40</i> Campo di variazione di Hurst	177
<i>Fig. 41</i> Grafici del prezzo e del rendimento logaritmico di Bayer	180
<i>Fig. 42</i> Grafici del prezzo e del rendimento logaritmico di Exor	180
<i>Fig. 43</i> Grafici del prezzo e del rendimento logaritmico di Lufthansa	181
<i>Fig. 44</i> Grafici del prezzo e del rendimento logaritmico di Philips	181
<i>Fig. 45</i> Grafici del prezzo e del rendimento logaritmico di Renault	182
<i>Fig. 46</i> Grafico R/S per il titolo Bayer	183
<i>Fig. 47</i> Grafico R/S per il titolo Exor	183

<i>Fig. 48</i> Grafico R/S per il titolo Lufthansa	184
<i>Fig. 49</i> Grafico R/S per il titolo Philips	184
<i>Fig. 50</i> Grafico R/S per il titolo Renault	185
<i>Fig. 51</i> Tabella delle stime ottenute per l'esponente di Hurst	185

*‘In economia, non potrà mai esistere una «teoria del tutto». Sono convinto, però, che ogni tentativo porti sempre più vicino a una comprensione adeguata del comportamento dei mercati.’*

***B. B. MANDELBROT.***

# INTRODUZIONE

L'interpretazione moderna della finanza nasce a inizio 900' dal lavoro di Bachelier, il primo ad utilizzare la teoria della probabilità per studiare i mercati finanziari. Le oscillazioni dei prezzi sono interpretate come un cammino casuale, in cui le aspettative future non dipendono dai valori passati e in cui le osservazioni si distribuiscono secondo una curva normale. Questa impostazione ha permesso di cambiare il metodo di indagine all'interno della finanza, favorendo la creazione di modelli matematici per il comportamento dei mercati, la cui realistica è stata messa in discussione in tempi recenti. In particolare, l'indipendenza e l'ipotesi di normalità sembrano non adattarsi al tipo di casualità presente nelle fluttuazioni finanziarie e sembra essere necessario ricorrere a nuovi strumenti. L'idea della tesi è quella di presentare una diversa interpretazione dei mercati finanziari, tramite i risultati della geometria frattale. Nel Capitolo I, si introducono gli strumenti matematici essenziali per descrivere l'evoluzione dei prezzi, i processi stocastici. Il prezzo di un titolo si può rappresentare, infatti, come una successione di variabili casuali, all'interno di un processo probabilistico, del quale si è interessati a studiarne le caratteristiche. Nel Capitolo II, si approfondisce la teoria finanziaria classica, originata dal modello di Bachelier, attraverso i suoi principali risultati. In particolare, si espone il modello del mercato efficiente, la teoria di portafoglio e il modello di Black e Scholes. Il Capitolo III permette di indagare gli oggetti frattali, gli strumenti matematici su cui si basa la nuova teoria. La geometria frattale presenta numerose proprietà controintuitive, in particolare ci si focalizza sull'interpretazione del concetto di dimensione frazionaria. L'utilizzo dei frattali in finanza si concretizza nel Capitolo IV, in cui si presentano gli sviluppi

del modello di Mandelbrot dei mercati. In particolare, esso poggia le basi su tre importanti elementi, le distribuzioni stabili, la dipendenza a lungo termine e il tempo di contrattazione. All'interno della tesi si cerca di coniugare l'aspetto intuitivo, tramite una rappresentazione grafica dei fenomeni, e l'aspetto matematico, attraverso le definizioni. Nel Capitolo V si espone, infine, la parte di ricerca, nella quale si applica l'analisi R/S alle variazioni dei titoli di cinque aziende europee, allo scopo di stimare il parametro di Hurst, che quantifica la memoria a lungo termine delle oscillazioni.

# **CAPITOLO I.**

## **I PROCESSI STOCASTICI E IL MOTO BROWNIANO**

## 1.1 Le variabili aleatorie e le distribuzioni di probabilità

L'esposizione dei concetti principali della teoria della probabilità deve necessariamente partire da ciò che, in un certo senso, ne giustifica l'esistenza, gli esperimenti aleatori. Si tratta di esperimenti in cui non si può conoscere il risultato a priori. Il singolo esito è detto evento elementare  $\omega$  e il loro insieme si dice spazio campionario  $\Omega := \{\omega_1, \omega_2, \dots, \omega_n, \dots\}$ . Si indica con evento  $E$  un sottoinsieme  $E \subseteq \Omega$  contenente uno o più eventi elementari. Per definire la probabilità è necessario introdurre il concetto di  $\sigma$ -algebra, detta anche spazio degli eventi.

*Definizione.* ( $\sigma$ -algebra). Dato uno spazio  $\Omega$ , una famiglia di eventi  $\mathcal{F} = \{E_1, E_2, \dots\}$  è detta  $\sigma$ -algebra se soddisfa le seguenti proprietà:

- i.  $\emptyset \in \mathcal{F}$ ;
- ii.  $E \in \mathcal{F} \Rightarrow \bar{E} \in \mathcal{F}$ ;
- iii.  $E_i \in \mathcal{F}, \forall i \in \mathbb{N} \Rightarrow \bigcup_{i=1}^{\infty} E_i \in \mathcal{F}$ .

Una  $\sigma$ -algebra è quindi una collezione di insiemi chiusa rispetto alle operazioni di complementazione e unione numerabile, inoltre dalle proprietà i. e ii. si osserva che  $\Omega \in \mathcal{F}$ . L'insieme di tutti i possibili sottoinsiemi di  $\Omega$  si dice insieme delle parti  $\mathcal{P}(\Omega)$ . Lo spazio ottenuto  $(\Omega, \mathcal{F})$  si dice spazio misurabile o probabilizzabile. La  $\sigma$ -algebra più utilizzata è quella che contiene tutti gli insiemi aperti di  $\Omega$ , detta  $\sigma$ -algebra o campo di Borel, denotata  $\mathcal{B}(\Omega)$ , e i suoi elementi si dicono insiemi boreliani. In particolare, in  $\mathbb{R}$  essa è la più piccola  $\sigma$ -algebra che contiene tutti gli intervalli  $(-\infty, x)$  della retta reale.  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  è un esempio di spazio misurabile.



*Definizione.* (Probabilità). Dato uno spazio misurabile  $(\Omega, \mathcal{F})$ , si definisce misura di probabilità su  $\Omega$  la funzione  $P: \mathcal{F} \rightarrow [0,1]$  che soddisfa le proprietà:

- i.  $0 \leq P(E) \leq 1, \forall E \in \mathcal{F}$ ;
- ii.  $P(\Omega) = 1$ ;
- iii. Se  $E_1, E_2, \dots \in \mathcal{F}$ ,  $E_i \cap E_j = \emptyset \forall i \neq j$ , allora vale l'assioma di  $\sigma$ -additività:

$$P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i).$$

Lo spazio  $(\Omega, \mathcal{F}, P)$ , così definito, si dice spazio di probabilità. La probabilità è quindi una funzione che assegna a ogni evento dell'esperimento aleatorio la probabilità di realizzarsi, in particolare si ha che  $P(E) = 0$  indica un evento trascurabile e  $P(E) = 1$  indica l'evento certo. Dalle definizioni precedenti si può ricavare che:

$$P(E \cup \bar{E}) = P(E) + P(\bar{E}) = P(\Omega) = 1,$$

da cui si osserva che  $P(E) = 1 - P(\bar{E}) \geq 0$ . Per sviluppare una teoria della probabilità si è interessati a definire delle funzioni che associano ad ogni evento elementare  $\omega$  un numero reale  $x \in \mathbb{R}$ .

*Definizione.* (Funzione misurabile). Dati due spazi misurabili  $(\Omega, \mathcal{F})$  e  $(A, \mathcal{S})$ , una funzione  $f: \Omega \rightarrow A$  è detta misurabile se  $\forall E \in \mathcal{S}$  vale che:

$$f^{-1}(E) := \{x \in \Omega | f(x) \in E\} \in \mathcal{F}.$$

$\sigma(f) \subseteq \mathcal{F}$  è detta  $\sigma$ -algebra generata da  $f$ . Le funzioni misurabili all'interno di uno spazio di probabilità sono dette variabili aleatorie.

*Definizione.* (Variabile aleatoria). Dato uno spazio di probabilità  $(\Omega, \mathcal{B}, P)$ , si definisce variabile aleatoria o casuale una funzione misurabile a valori reali  $X: \Omega \rightarrow \mathbb{R}$ , tale che  $\forall B_i \in \mathbb{R}$  si ha che  $X^{-1}(B_i) \in \mathcal{B}$ .

In particolare, una funzione  $X$  definita su  $\Omega$  è misurabile rispetto al campo di Borel se  $\forall x \in \mathbb{R}$  l'insieme  $\{\omega \in \Omega: X(\omega) \leq x\} \in \mathcal{B}$ . La variabile aleatoria è un modo per indurre una misura di probabilità sullo spazio di arrivo a partire da quella definita sullo spazio di partenza  $\Omega$ , ovvero  $X(\omega)$  assume un valore  $x \in \mathbb{R}$  con una probabilità  $P^*(x) = P(\omega) \in [0,1]$ <sup>1</sup>. Le variabili casuali si possono distinguere in discrete, se possono assumere solo valori interi, o continue se possono assumere qualunque valore reale. Ogni variabile aleatoria discreta è caratterizzata da una distribuzione di probabilità, detta anche funzione di massa di probabilità (PMF), che assegna ad ogni esito dell'esperimento aleatorio la rispettiva probabilità di realizzarsi. In particolare, data la variabile casuale discreta  $X(\omega): \Omega \rightarrow \mathbb{R}$ , la PMF di  $X$  associa a ogni punto dello spazio campionario la probabilità che la variabile assuma quel valore:

$$p_X(x) = P(X = x) = P(\{\omega \in \Omega: X(\omega) = x\}).$$

Tale funzione è definita in modo che  $p_X(x) \geq 0$  e  $\sum_x p_X(x) = 1$ . Una variabile aleatoria che può assumere qualunque valore in  $\mathbb{R}$  si dice continua se può essere descritta da una funzione di densità di probabilità (PDF), denotata con  $f_X(x)$  o  $p(x)$ , che soddisfa le seguenti proprietà:

- i.  $f_X(x) \geq 0, \forall x \in \mathbb{R}$ ;
- ii.  $\int_{-\infty}^{\infty} f_X(x) dx = 1$ ;
- iii.  $P(a \leq X \leq b) = \int_a^b f_X(x) dx$ .

Considerando che l'insieme  $\{\omega \in \Omega: X(\omega) \leq x\} \in \mathcal{B}$ , per una funzione  $\mathcal{B}$ -misurabile, si può definire un ulteriore modo di rappresentare la distribuzione di una variabile aleatoria, tramite la funzione di ripartizione (CDF), che rappresenta la probabilità che  $X$  assuma valori minori o uguali a  $x$ :

---

<sup>1</sup> R. RIGANTI, *Elementi di probabilità e statistica*, 2003, cap. 2.

$$F_X(x) = P(X \leq x) = \begin{cases} \sum_{x_k \leq x} P(X = x_k) = \sum_{x_k \leq x} p_k & (\text{se } X \text{ discreta}) \\ P(-\infty \leq X \leq x) = \int_{-\infty}^x f_X(t) dt & (\text{se } X \text{ continua}) \end{cases}.$$

La distribuzione cumulata di probabilità gode delle seguenti proprietà:

- i.  $0 \leq F_X(x) \leq 1$ ;
- ii.  $\lim_{x \rightarrow -\infty} F_X(x) = 0$  e  $\lim_{x \rightarrow +\infty} F_X(x) = 1$ ;
- iii.  $F_X(x) = 1 - F_X(-x)$ ;
- iv.  $f_X(x) = \frac{\partial}{\partial x} F_X(x)$ .

### 1.1.1 I momenti e l'equazione caratteristica

In alcuni casi, data una variabile aleatoria  $X$ , non si è interessati all'intera distribuzione di probabilità, ma solamente ad alcuni parametri caratterizzanti. Il più noto è il valore atteso, detto anche speranza matematica, che formalizza l'idea di valore medio di un fenomeno casuale:

$$E(X) := \begin{cases} \sum_k x_k p_k & (\text{se } X \text{ discreta}) \\ \int_{-\infty}^{\infty} x f_X(x) dx & (\text{se } X \text{ continua}) \end{cases}.$$

Il valore atteso è un parametro di posizione che indica dove è centrata la distribuzione di probabilità. È un operatore lineare in quanto, date  $X_1$  e  $X_2$  variabili aleatorie, si ha  $E(aX_1 + bX_2) = aE(X_1) + bE(X_2)$ , inoltre gode di monotonia, poiché se  $X_1 \geq X_2 \Rightarrow E(X_1) \geq E(X_2)$ . Il valore atteso, indicato  $E(X) \equiv \langle X \rangle \equiv \bar{X}$ , è finito solo se la sommatoria (o l'integrale) converge assolutamente, inoltre, dalla definizione, il valore atteso degli scarti  $X - \langle X \rangle$  è nullo. Un secondo importante parametro di interesse, in una distribuzione di probabilità, è la varianza  $\sigma_X^2$ , definita come valore atteso degli scarti al quadrato:

$$\sigma_X^2 := E[(X - \langle X \rangle)^2] = \begin{cases} \sum_{k=-\infty}^{\infty} (x_k - \bar{x})^2 p_k & (\text{se } X \text{ discreta}) \\ \int_{-\infty}^{\infty} (x - \bar{x})^2 f_X(x) dx & (\text{se } X \text{ continua}) \end{cases}.$$

La varianza è un parametro di dispersione che indica la larghezza della distribuzione, inoltre, per ottenere una dimensione confrontabile con quella della variabile aleatoria, spesso si considera la deviazione standard  $\sigma_X = \sqrt{\sigma_X^2}$ .

Dalla definizione si può dimostrare<sup>2</sup> che:

$$\sigma_X^2 = E(X^2) - E(X)^2.$$

I concetti di valore atteso e varianza possono essere generalizzati definendo il momento di ordine  $q$ , per  $q$  intero positivo, di una variabile aleatoria  $X$ , come il valore atteso della  $q$ -esima potenza:

$$M_0^q := E(X^q) = \begin{cases} \sum_{k=-\infty}^{\infty} x_k^q p_k & (\text{se } X \text{ discreta}) \\ \int_{-\infty}^{\infty} x^q f_X(x) dx & (\text{se } X \text{ continua}) \end{cases}.$$

Per  $q = 1$ ,  $M_0^1$  corrisponde alla speranza e per  $q = 2$ ,  $M_0^2$  corrisponde alla varianza, a meno di una traslazione. Infatti, i momenti  $M^q$  di ordine superiori al primo vengono definiti rispetto al valore atteso, operando la traslazione  $X' = X - E(X)$ . I momenti di una variabile aleatoria forniscono un modo alternativo alla PDF e alla CDF per identificarne la distribuzione di probabilità, in quanto possono essere racchiusi all'interno della funzione generatrice dei momenti  $M_X(t)$ :

---

<sup>2</sup>  $\sigma_X^2 = E[(X - E(X))^2] = E[X^2 - 2XE(X) + E(X)^2] = E(X^2) - 2E(X)E(X) + E(X)^2 = E(X^2) - E(X)^2.$

$$M_X(t) := E(e^{tX}) = \begin{cases} \sum_{k=-\infty}^{\infty} e^{x_k t} p_k & (\text{se } X \text{ discreta}) \\ \int_{-\infty}^{\infty} e^{xt} f_X(x) dx & (\text{se } X \text{ continua}) \end{cases}.$$

Dall'espansione di Taylor dell'esponenziale si può scrivere che  $e^x = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \dots$  e che  $e^{tx} = 1 + tx + \frac{t^2}{2!}x^2 + \frac{t^3}{3!}x^3 + \dots$ . Calcolandone il valore atteso si ha:

$$M_X(t) = E(e^{tX}) = 1 + tE(X) + \frac{t^2}{2!}E(X^2) + \frac{t^3}{3!}E(X^3) + \dots.$$

I singoli momenti della distribuzione si possono ricavare differenziando  $M_X(t)$  rispetto a  $t$  e ponendo  $t = 0$ :

$$E(X^q) = \frac{\partial^q}{\partial t^q} M_X(t) |_{t=0}.$$

La prima derivata restituisce il valore atteso<sup>3</sup> e la seconda consente di esprimere la varianza. La funzione generatrice non esiste per qualunque  $X$ , in quanto i momenti sono definiti solo se l'integrale converge in senso assoluto, cioè  $\int_{-\infty}^{\infty} |x|^q f_X(x) dx < +\infty$ . Spesso si utilizza quindi la funzione caratteristica  $\varphi_X(t)$ , che esprime in modo univoco la distribuzione e il comportamento di una variabile aleatoria ed è sempre definita, in particolare  $\varphi_X(t) = M_X(it)$ , cioè  $\varphi_X: \mathbb{R} \rightarrow \mathbb{C}$  consente di valutare la funzione dei momenti nell'asse immaginario<sup>4</sup>:

---

<sup>3</sup>  $\frac{\partial}{\partial t} M_X(t) = 0 + E(X) + tE(X^2) + t^2E(X^3) + \dots |_{t=0} = E(X)$ .

<sup>4</sup> L'equazione caratteristica rappresenta il valore atteso della funzione complessa  $g(x) = e^{itX} = \cos(tX) + i \sin(tX)$ , quindi si può scrivere che:

$$\varphi_X(t) = E[\cos(tX)] + iE[\sin(tX)].$$

Se  $X$  è discreta si ha che:

$$\varphi_X(t) = \sum_k e^{itx_k} p_k = \sum_k \cos(tx_k) p_k + i \sum_k \sin(tx_k) p_k.$$

$$\varphi_X(t) := E(e^{itX}) = \int_{-\infty}^{\infty} e^{itx} f_X(x) dx.$$

Si osserva che  $\varphi_X(0) = 1$ . Se due variabili hanno la stessa equazione caratteristica allora hanno anche la stessa funzione di densità, inoltre, date  $X_1, \dots, X_n$  variabili indipendenti, vale che  $\varphi_{X_1, \dots, X_n}(t) = \varphi_{X_1}(t) \varphi_{X_2}(t) \dots \varphi_{X_n}(t)$ <sup>5</sup>. Conoscendo l'equazione caratteristica è possibile ricavare la corrispondente densità come:

$$f_X(x) = F'_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \varphi_X(t) dt.$$

## 1.2 La distribuzione normale e il teorema del limite centrale

La distribuzione normale o gaussiana è un tipo di distribuzione di probabilità, continua e simmetrica, a forma campanulare che fornisce una rappresentazione verosimile della distribuzione di molti fenomeni, come ad esempio l'altezza di una popolazione o i punteggi dei test per valutare i QI. Prende il nome dal matematico tedesco Carl Friedrich Gauss che, nel XIX secolo, ne ha teorizzato le proprietà<sup>6</sup>. Una variabile aleatoria  $X$  ha una distribuzione normale  $N(\mu, \sigma^2)$  se la sua funzione di densità è pari a:

$$f_X(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

ove  $\mu \in \mathbb{R}$  è la media e  $\sigma^2 > 0$  è la varianza. Per  $\mu = 0$  e  $\sigma = 1$  si ha una distribuzione normale standard con densità:

---

<sup>5</sup> R. RIGANTI, *Elementi di probabilità e statistica*, 2003, cap. 3.

<sup>6</sup> Gauss lavorando a misurazioni astronomiche si è trovato a fronteggiare il problema degli errori, i quali sembravano addensarsi attorno a un valore centrale per poi disperdersi, all'aumentare della distanza da esso.

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

La funzione di ripartizione della distribuzione gaussiana è data da:

$$F_X(x) = \int_{-\infty}^x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \frac{1}{2} \operatorname{erfc}\left(\frac{\mu-x}{\sigma\sqrt{2}}\right),$$

con  $\operatorname{erfc}(x) := \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-y^2} dy$  funzione degli errori complementare<sup>7</sup>. Dalla simmetria della distribuzione, si nota che il valore medio è anche la moda e la mediana, inoltre  $P(X \geq x) = P(X \leq x) = \frac{1}{2}$ .

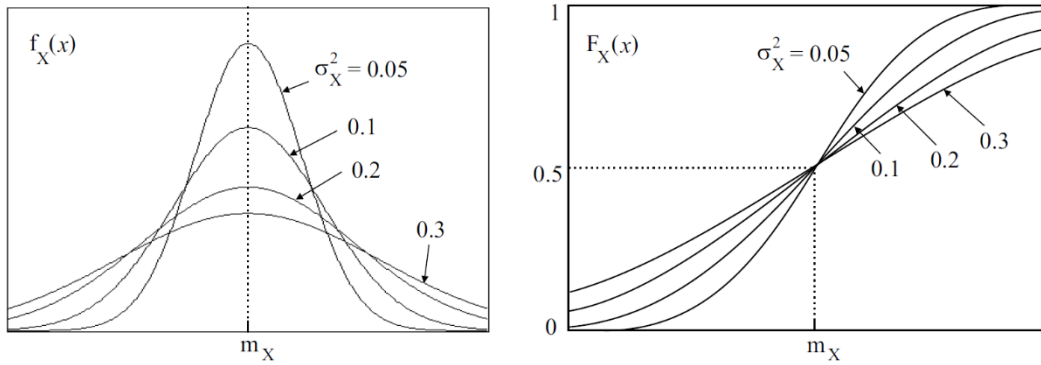


Fig. 1 PDF e CDF della distribuzione normale. [68]

La maggior parte delle osservazioni è concentrata attorno alla media quindi, in corrispondenza di tale punto la funzione di densità presenta un massimo:

$$f_X(\mu) = \frac{1}{\sigma\sqrt{2\pi}}$$

In particolare, il 68,27% dei valori si colloca nell'intervallo  $[\mu - \sigma, \mu + \sigma]$ , i cui estremi rappresentano anche i punti di flesso della curva, il 95,45% dei valori si trova nell'intervallo  $[\mu - 2\sigma, \mu + 2\sigma]$  e il 99,73% all'interno di  $[\mu - 3\sigma, \mu + 3\sigma]$ .

<sup>7</sup> Ponendo  $y = \frac{\mu-x}{\sigma\sqrt{2}}$ , da cui  $dx = -\sigma\sqrt{2}dy$ , si può scrivere  $F_X(x)$  come:

$$F_X(x) = \int_{\frac{\mu-x}{\sigma\sqrt{2}}}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{y^2} (-\sigma\sqrt{2}) dy = \frac{1}{\pi} \int_{\frac{\mu-x}{\sigma\sqrt{2}}}^{\infty} e^{y^2} dy = \frac{1}{2} \operatorname{erfc}\left(\frac{\mu-x}{\sigma\sqrt{2}}\right).$$

$3\sigma]$ <sup>8</sup>. Uno dei principali utilizzi della distribuzione normale fa riferimento a una delle proprietà più note, secondo cui:

$$N(x, \mu, \sigma) = N(z, 0, 1).$$

Il valore della funzione di ripartizione di una generica distribuzione normale coincide con il valore della funzione di ripartizione della normale standardizzata

$F_X(z) = \frac{1}{2} \operatorname{erfc}\left(\frac{-z}{\sqrt{2}}\right)$ , in cui:

$$z = \frac{x - \mu}{\sigma}$$

L'integrale non è esprimibile in termini di funzioni elementari, tuttavia attraverso le tavole della funzione di ripartizione di  $z$ <sup>9</sup> è possibile calcolare le aree di probabilità della variabile in esame. La funzione generatrice dei momenti di  $X \sim N(\mu, \sigma^2)$  è data da:

$$M_X(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = e^{\mu t + \frac{1}{2}\sigma^2 t^2}.$$

Più in generale, l'equazione caratteristica di una distribuzione normale si può scrivere come:

---

<sup>8</sup> Questi risultati sono ottenuti usando la disuguaglianza di Tchebycheff, secondo cui, se si costruisce un intervallo simmetrico di raggio  $\varepsilon = \lambda\sigma$  attorno alla media, la probabilità che la variabile casuale  $X$  assuma valori esterni all'intervallo non può superare la soglia  $1/\lambda^2$ , mentre la probabilità interna non può scendere al di sotto della soglia  $1 - 1/\lambda^2$ , ovvero:

$$\begin{aligned} \Pr(|X - \mu| \geq \lambda\sigma) &\leq 1/\lambda^2 \\ \Pr(|X - \mu| < \lambda\sigma) &\geq 1 - 1/\lambda^2. \end{aligned}$$

<sup>9</sup> Le prime tabelle della funzione di ripartizione di  $z$  furono create da W. F. Sheppard.



$$\varphi_X(t) = E(e^{itX}) = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{itx} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = e^{i\mu t - \frac{1}{2}\sigma^2 t^2}.^{10}$$

Esistono ulteriori parametri con cui può essere descritta la forma di una distribuzione, in particolare l'asimmetria, che indica quanto è simmetrica la distribuzione rispetto al valore centrale, e la curtosi, che indica lo spessore delle code, ovvero la probabilità di valori outlier. L'indice di asimmetria  $\gamma_1 = \frac{M_0^3}{(M_0^2)^{3/2}}$

è un indice adimensionale che è uguale a zero se la distribuzione è simmetrica, maggiore di zero se presenta asimmetria positiva e minore di zero nel caso di asimmetria negativa. L'indice di curtosi  $\gamma_2 = \beta_2 - 3$ , con  $\beta_2 = \frac{M_0^4}{(M_0^2)^2}$ , indica la tendenza della distribuzione ad assumere una forma a campana. Si ha mesocurtosi se  $\beta_2 = 3$ , platicurtosi se  $\beta_2 < 3$  e leptocurtosi se  $\beta_2 > 3$ <sup>11</sup>. Il largo utilizzo che ha trovato la distribuzione normale nelle applicazioni deriva principalmente dal teorema del limite centrale, che consente di approssimare distribuzioni non note ad una gaussiana.

*Teorema.* (Teorema del limite centrale). Data  $\{X_n\}$ ,  $n \in \mathbb{N}^+$ , una successione di variabili aleatorie indipendenti, con uguale  $f_{X_n}(x_n)$  e varianza finita, allora la funzione di densità della somma:

$$S_n = \sum_{i=1}^n X_i,$$

---

<sup>10</sup> Sostituendo  $y = x - \mu$  si può scrivere:

$$\varphi_X(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{it\mu} \int_{-\infty}^{\infty} e^{ity} e^{-\frac{y^2}{2\sigma^2}} dy = \frac{1}{\sigma\sqrt{2\pi}} e^{it\mu} \int_{-\infty}^{\infty} e^{-\frac{y(y-2\sigma^2 it)}{2\sigma^2}} dy,$$

sostituendo nuovamente  $u = y - it\sigma^2$ , e sfruttando le proprietà di  $f_X(x)$ , si ottiene:

$$\varphi_X(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{it\mu} \int_{-\infty}^{\infty} e^{-\frac{(u+it\sigma^2)(u-it\sigma^2)}{2\sigma^2}} du = e^{i\mu t} e^{-\frac{1}{2}\sigma^2 t^2} \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{u^2}{2\sigma^2}} du = e^{i\mu t - \frac{1}{2}\sigma^2 t^2}.$$

<sup>11</sup> L. BOLLANI, L. BOTTACIN, *Statistica per l'azienda, metodi per gli ambiti operativi e per la sfera sociale*, 2018.

converge a  $N(n\mu, n\sigma^2)$ , per  $n \rightarrow \infty$ . È possibile dimostrare tale risultato sfruttando le proprietà della funzione caratteristica.

*Dimostrazione.* Date  $X_1, \dots, X_n$  variabili aleatorie indipendenti con media  $\mu$  e varianza  $\sigma^2$ , si considera la variabile  $Z_n$ :

$$Z_n = \frac{X_1 + \dots + X_n - n\mu}{\sqrt{n\sigma^2}} = \sum_{i=1}^n \frac{X_i - \mu}{\sigma\sqrt{n}} = \sum_{i=1}^n \frac{Y_i}{\sqrt{n}}$$

dove  $Y_i = \frac{X_i - \mu}{\sigma}$ , con  $E(Y_i) = 0$  e  $\sigma_2(Y_i) = 1$ . L'equazione caratteristica  $\varphi_{Z_n}(t)$  è data da:

$$\varphi_{Z_n}(t) = \varphi_{\sum_i Y_i} \left( \frac{t}{\sqrt{n}} \right) = \prod_{i=1}^n \varphi_{Y_i} \left( \frac{t}{\sqrt{n}} \right) = \left[ \varphi_{Y_1} \left( \frac{t}{\sqrt{n}} \right) \right]^n.$$

Sfruttando l'espansione di Taylor, l'equazione caratteristica di  $Y_i$  si può scrivere come  $\varphi_{Y_i} \left( \frac{t}{\sqrt{n}} \right) = 1 - \frac{t^2}{2n} + o \left( \frac{t^2}{n} \right)$ . Sostituendo nell'espressione precedente si ottiene:

$$\varphi_{Z_n}(t) = \left[ 1 - \frac{t^2}{2n} + o \left( \frac{t^2}{n} \right) \right]^n \rightarrow \lim_{n \rightarrow \infty} \left( 1 - \frac{t^2}{2n} \right)^n = e^{-\frac{t^2}{2}} = \varphi_{N(0,1)}(t).$$

Si è quindi dimostrato che  $Z_n \sim N(0,1)$  e, di conseguenza,  $S_n = \sum_i X_i \sim N(n\mu, n\sigma^2)$ , per  $n \rightarrow \infty$ .

Questo importante teorema consente di descrivere in termini di probabilità tutti quei fenomeni che possono essere considerati come prodotto di un elevato numero di cause indipendenti<sup>12</sup>.

### 1.2.1 La proprietà di stabilità per una distribuzione

Per concludere questa parte sulla distribuzione normale, è importante introdurre una proprietà che risulterà centrale per alcune considerazioni successive, la stabilità di una distribuzione. Essa descrive in realtà un'ampia classe di

---

<sup>12</sup> R. RIGANTI, *Elementi di probabilità e statistica*, 2003, cap. 2.

distribuzioni di probabilità, di cui i più noti esempi sono la distribuzione normale, di Cauchy e di Levy, che hanno molte applicazioni in diversi tipi di modelli fisici ed economici. Se si considerano due variabili casuali gaussiane, la somma di queste due risulterà ancora una variabile normale, ovvero:

$$aX_1 + bX_2 \stackrel{d}{=} cX + d.$$

Se  $X$  è normale, allora combinazioni lineari di  $X_1$  e  $X_2$ , copie indipendenti di  $X$ , con  $a, b > 0$ , presentano la stessa legge di probabilità di  $X$ , con  $c > 0$  e  $d \in \mathbb{R}$ . In altre parole, la forma di  $X$  è preservata dall'addizione e  $X$  è detta variabile casuale stabile se questa condizione rimane soddisfatta al variare dei parametri. Si dice che  $X$  è strettamente stabile se la condizione vale al variare di  $a$  e  $b$  con  $d = 0$ . Infine, si dice simmetrica stabile una distribuzione  $X$  che sia stabile e distribuita simmetricamente rispetto a 0<sup>13</sup>:

$$X \stackrel{d}{=} -X.$$

### *1.3 Introduzione ai processi stocastici*

Un processo stocastico<sup>14</sup>, in modo intuitivo, può essere descritto come una rappresentazione di una grandezza che varia nel tempo in modo casuale, ovvero una collezione di variabili casuali nel tempo. Una definizione alternativa vede un processo stocastico come una distribuzione di probabilità su uno spazio di percorsi, che descrivono l'evoluzione nel tempo di un sistema casuale. Più precisamente, un processo stocastico è una famiglia di variabili casuali  $\{X_t, t \in [0, T] \subset \mathbb{R}\}$  definite su uno stesso spazio di probabilità  $(\Omega, \mathcal{F}, P)$ :

---

<sup>13</sup> J. P. NOLAN, *Stable distributions models for heavy tailed data*, 2018.

<sup>14</sup> L'aggettivo stocastico deriva dal greco *stochastikos* e significa 'che mira bene', 'abile nel fare congetture'.

$$X_t : \Omega \rightarrow \mathbb{R}^n.$$

Una famiglia crescente  $\mathcal{F}_t = \{\mathcal{F}_t\}_{t \in I}$ , con  $I \in \mathbb{R}_+$ ,  $\sigma$ -algebra su  $\mathcal{F}$ , si definisce filtrazione in  $(\Omega, \mathcal{F}, P)$ . In altri termini, una filtrazione rappresenta tutti gli eventi di cui si può conoscere l'esito, ovvero le informazioni contenute nel processo al tempo  $t \in T$ . Un processo  $X_t$  è adattato alla filtrazione  $\mathcal{F}_t$  se è  $\mathcal{F}_t$ -misurabile  $\forall t \in I$ <sup>15</sup>. La coppia  $\{X_t, \mathcal{F}_t\}$  si dice sequenza stocastica adattata. Considerando due processi stocastici  $\{X_t\}$  e  $\{Y_t\}$  definiti su uno stesso spazio di probabilità,  $\{Y_t\}$  è una modificazione di  $\{X_t\}$  se,  $\forall t \in T$ , le variabili aleatorie  $X$  e  $Y$  coincidono quasi certamente. Un processo stocastico si dice sommabile se, per ogni  $t$ ,  $E[|X_t|] < \infty$ . A differenza dei processi deterministici, le traiettorie delle variabili casuali in esame non possono essere stabilite a priori, tuttavia le informazioni sui percorsi possono essere dedotte dalla distribuzione di probabilità. È possibile generalizzare un processo stocastico definendo il campo casuale, in cui il parametro  $t$ , solitamente associato al tempo, non è un valore intero o reale, ma è un vettore multidimensionale di  $\mathbb{R}^n$ .

### *1.3.1 I processi stocastici con tempo discreto e il random walk monodimensionale*

Un processo stocastico si dice discreto se la variabile temporale assume valori interi positivi,  $X_0, X_1, X_2, \dots, X_t$ ,  $t \in \mathbb{N}^+$ , oppure si dice continuo se la variabile temporale può assumere qualunque valore reale positivo,  $\{X_t\}_{t \geq 0}$ . In questa parte, si prendono in considerazione i processi discreti, in particolare il processo più semplice è il random walk monodimensionale<sup>16</sup>. Siano  $Y_1, Y_2, \dots, Y_i$  variabili casuali indipendenti e identicamente distribuite tali che:

---

<sup>15</sup> S. E. SHREVE, *Stochastic calculus for finance II, continuous time models*, 2004, cap. 2.

<sup>16</sup> P. KEMPTHORNE, C. LEE, V. STRELA, J. XIA, 18.S096, *Topics in Mathematics with Applications in Finance*, Fall 2013, Massachusetts Institute of Technology: MIT.

$$Y_i := \begin{cases} +1 & P(Y_i) = p \\ -1 & P(Y_i) = 1 - p \end{cases}$$

con  $p = \frac{1}{2}$ . Sia  $X_0 = 0$  e  $X_t = \sum_{i=1}^n Y_i$ ,  $\forall n$ , allora la sequenza di variabili casuali  $\{X_0, X_1, X_2, \dots, X_t\}$  è un processo stocastico discreto chiamato cammino casuale simmetrico unidimensionale. Per  $p \neq \frac{1}{2}$ , si ha un random walk asimmetrico. Questo tipo di moto è caratterizzato da:

- i.  $E[X_t] = \sum_{i=1}^n E[Y_i] = 0$ ,  $\forall t$ ;
- ii. (incrementi indipendenti). Per  $0 = t_0 \leq t_1 \leq \dots \leq t_n$  le variabili casuali  $X_{t_{i-1}} - X_{t_i}$  sono reciprocamente indipendenti;
- iii. (stazionarietà).  $\forall h \geq 0$  e  $\forall t \geq 0$  la distribuzione di  $X_{t+h} - X_t$  è uguale alla distribuzione di  $X_t - X_{t-h}$ , ovvero:

$$\{X_{t_1}, X_{t_2}, \dots, X_{t_n}\} \stackrel{d}{=} \{X_{t_1+h}, X_{t_2+h}, \dots, X_{t_n+h}\}.$$

Se  $X_t \sim N(0, n\sigma^2)$  si ha un cammino casuale gaussiano. Considerando un lancio di una moneta in cui, partendo dall'origine, ogni lancio ha la stessa probabilità di generare un esito positivo  $+1$  o negativo  $-1$ , se ogni lancio è indipendente, il bilancio di un giocatore ha una distribuzione assimilabile a un random walk.

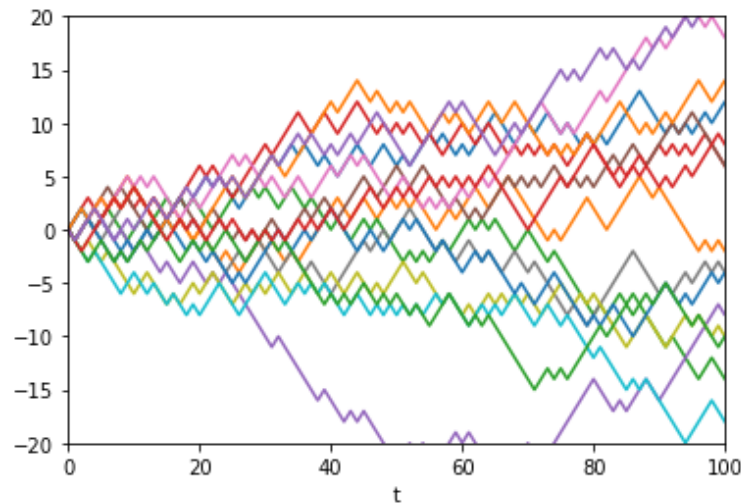


Fig. 2 Simulazioni del random walk.

### 1.3.2 I processi di Markov e la definizione di martingala

Il processo di random walk, descritto in precedenza, possiede una proprietà che caratterizza una tipologia di processi stocastici, denominati processi markoviani, in cui l'effetto dei valori passati è interamente contenuto nel valore corrente, e quest'ultimo è l'unico determinante dei valori futuri. In altri termini, la distribuzione di probabilità di  $X_{t+1}$  dipende solamente dal valore di  $X_t$  e non dall'intero set di valori  $X_0, X_1, \dots, X_t$ . Formalmente questi processi sono descritti dalla condizione di “assenza di memoria” o proprietà di Markov.

*Definizione.* (Proprietà di Markov). Un processo stocastico discreto  $X_0, X_1, X_2, \dots, X_t$  è un processo markoviano se:

$$P(X_{t+1} = s \mid X_0, X_1, \dots, X_t) = P(X_{t+1} = s \mid X_t), \quad \forall t \geq 0, \forall s.$$

Il random walk è un processo di Markov in quanto entrambe queste due probabilità sono pari a  $\frac{1}{2}$ . Un processo a tempo discreto che soddisfa la condizione di assenza di memoria è detto catena di Markov, con  $S$  insieme degli stati, finito o infinito numerabile, che il processo può assumere. Nel caso in cui l'insieme degli stati  $S$  sia un insieme finito,  $S = [m]$ , un processo di Markov può essere descritto in termini di probabilità di transizione:

$$p_{ij} = P(X_{t+1} = j \mid X_t = i), \quad i, j \in S.$$

Gli elementi  $p_{ij}$  possono essere racchiusi nella matrice di transizione  $A \in \mathbb{R}^{m \times m}$ :

$$A = \begin{pmatrix} p_{11} & \cdots & p_{1m} \\ \vdots & \ddots & \vdots \\ p_{m1} & \cdots & p_{mm} \end{pmatrix}.$$

All'interno della matrice sono contenute tutte le informazioni del processo stocastico. Gli elementi sulla diagonale principale indicano le probabilità di rimanere nello stesso stato, inoltre la somma degli elementi lungo le colonne è

uguale a 1, ovvero  $\sum_{j \in S} p_{ij} = 1$ <sup>17</sup>. Una catena di Markov è omogenea se la probabilità di transizione tra gli stati è costante nel tempo, ovvero  $\forall t \geq 1$ :

$$P(X_{t+n} = i | X_t = j) = P(X_{t-1+n} = i | X_{t-1} = j).$$

Una catena di Markov si dice irriducibile se la probabilità di transizione tra stati è sempre positiva:

$$P(X_{t+n} = i | X_t = j) > 0, \quad \forall i, j \in S, \forall t, n \in \mathbb{N}.$$

Un'ulteriore importante categoria di processi stocastici è la martingala<sup>18</sup>. Questa nozione, all'interno della teoria delle probabilità, è stata introdotta per la prima volta dal matematico francese Paul Levy.

*Definizione.* (Martingala). Data una sequenza stocastica adattata  $\{X_t, \mathcal{F}_t\}$ , un processo stocastico a tempo discreto  $\{X_0, X_1, \dots, X_t\}$ ,  $t \in T$ , è una martingala se:

$$E[|X_t|] < \infty,$$

$$E[X_{t+1} | \mathcal{F}_t] = X_t, \quad \forall t > 0,$$

con  $\mathcal{F}_t = \{X_0, \dots, X_t\}$ <sup>19</sup>.

Un processo stocastico  $\{X_t\}$  caratterizzato da valore atteso condizionato nullo è detto differenza di martingala:

$$E[X_{t+1} | \mathcal{F}_t] = 0, \quad \forall t > 0.$$

Al tempo  $t$  il valore atteso della distribuzione al tempo  $t + 1$  è esattamente uguale a quello al tempo  $t$ , quindi il valore atteso degli incrementi durante il processo è nullo, per ogni  $t$ . Per questa ragione, la martingala si colloca all'interno dei giochi equi, inoltre essa è sempre un processo markoviano. Si può immaginare la martingala come un processo senza tendenza, in cui non vi è nessuna variazione attesa, e di conseguenza:

---

<sup>17</sup> S. ROSS, *A first course in probability*, 2010, cap 9.

<sup>18</sup> Il termine martingala fa riferimento a una serie di strategie utilizzate dagli scommettitori nel XVIII secolo. In particolare, la più semplice, usata nei lanci della moneta, prevedeva di raddoppiare la puntata dopo ogni lancio perso.

<sup>19</sup> R. DURRETT, *Probability: theory and examples*, 2019, cap.4.

$$P(X_{t+1}) = P(X_t).$$

Il random walk è una martingala mentre il bilancio di un giocatore di roulette non lo è, poiché ci si aspetta di perdere. Il processo  $X_t$  è detto submartingala se è soddisfatta la seguente condizione:

$$E[X_{t+1} | \mathcal{F}_t] \geq X_t, \quad \forall t > 0.$$

Infine, si dice supermartingala un processo stocastico che decresce in media, in cui:

$$E[X_{t+1} | \mathcal{F}_t] \leq X_t, \quad \forall t > 0.$$

#### *1.4 Il moto browniano e i processi stocastici con tempo continuo*

Dopo aver introdotto le caratteristiche dei processi stocastici discreti, si generalizzeranno ora i risultati precedenti definendo il caso in cui la variabile temporale sia un numero reale, ovvero considerando il processo stocastico con tempo continuo,  $\{X_t\}_{t \geq 0}$ ,  $t \in \mathbb{R}$ . La prima classe di processi continui che si considera è quella dei processi gaussiani, che rappresenta un'estensione dei vettori casuali normali.

*Definizione.* (Processo gaussiano). Un processo stocastico reale  $\{X_t\}_{t \geq 0}$ , è detto processo gaussiano se  $\forall t_n \in T$ , l'insieme  $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$  ha distribuzione normale, inoltre ogni combinazione lineare finita di  $X_t$  è normale.

Il più importante processo stocastico continuo, il cui ruolo all'interno della finanza sarà approfondito nel prossimo capitolo, è il moto browniano semplice o processo di Wiener.



*Definizione.* (Moto browniano semplice). Dato uno spazio di probabilità  $(\Omega, \mathcal{F}, P)$ , il processo stocastico  $\{B_t\}_{t \in T}$ , si chiama moto browniano semplice (SBM) se soddisfa le seguenti proprietà:

- i.  $P(B_0 = 0) = 1$ ;
- ii. (stazionarietà).  $\forall 0 \leq s < t, B(t) - B(s) \sim N(0, t - s)$ ;
- iii. (incrementi indipendenti). Se gli intervalli  $[s_i, t_i]$  non si sovrappongono, allora  $B(t_i) - B(s_i)$  sono reciprocamente indipendenti  $\forall i$ ;
- iv.  $\exists A \in \mathcal{F}$  con  $P(A) = 1$  tale che  $\forall \omega \in A$  la funzione  $t \rightarrow B_t(\omega)$  è continua.

Nel 1923 Wiener ha dimostrato l'esistenza di questo processo all'interno di uno spazio di probabilità, inoltre si può dimostrare che  $\{B_t\}_{t \in T}$  è un moto browniano semplice solo se è un processo gaussiano continuo con:

- i.  $E(B_t) = 0, E(B_t - B_s) = 0$ ;
- ii.  $\sigma^2(B_t) = t; \sigma^2(B_t - B_s) = t - s$ ;
- iii.  $Cov(B_t, B_s) = \begin{cases} s & \text{se } s \leq t \\ t & \text{se } t \leq s \end{cases} = \min(t, s)$ .

In altri termini, esiste un processo gaussiano, che inizia quasi certamente dall'origine, i cui intervalli sono normalmente distribuiti e tra i quali è garantita l'indipendenza, avente valore atteso nullo e varianza pari alla lunghezza dell'intervallo. Se  $\{B_t\}_{t \in T}$  è un SBM, allora il processo  $\{X_t\}_{t \in T}$  è un moto browniano nel caso in cui:

- i. (riflessione spaziale).  $X_t := -B_t$ ;
- ii. (traslazione temporale).  $X_t := B_{t_0+\tau} - B_{t_0}, \forall t_0 \geq 0$ ;
- iii. (riflessione temporale).  $X_t := B_{t_0-\tau} - B_{t_0}, \forall t_0 \geq 0$ ;
- iv. (riscaldamento diffusivo).  $X_t := \frac{1}{\sqrt{c}} B_{ct}, c > 0$ ;
- v. (inversione temporale).  $X_t := t B_{1/t}, t > 0, X_0 = 0$ .

Il moto browniano fa parte di una più ampia classe di processi stocastici continui, detti processi di Levy, caratterizzati da incrementi stazionari e indipendenti, inoltre l'SBM è un processo di Markov ed è un esempio di martingala a tempo continuo<sup>20</sup>. Il nome deriva dal botanico scozzese Robert Brown, che per primo, nel 1827, ne ha osservato fisicamente le caratteristiche. Guardando al microscopio il comportamento di granuli di polline in acqua, egli ha notato la presenza di due tipi di corpuscoli, uno a forma oblunga e uno a forma sferica, che erano soggetti ad uno strano tipo di movimento orizzontale lungo direzioni casuali, in ogni istante di tempo. Questo movimento non era causato né dalla forza di gravità né dalla resistenza dell'acqua e, come ha verificato egli stesso, non era un fenomeno di tipo organico. Brown non è riuscito a spiegare matematicamente questo evento e bisogna attendere una pubblicazione di Albert Einstein, del 1905, per ottenerne una descrizione accurata. La motivazione risiede nella collisione tra i corpuscoli e le molecole d'acqua, che origina spostamenti quasi impercettibili, continui in ogni direzione<sup>21</sup>. Un ulteriore modo per definire il moto browniano è come limite di un random walk discreto. In particolare, si considera  $\{Y_0, Y_1, \dots, Y_n\}$  un random walk con incrementi aventi media nulla e varianza unitaria. Sia  $Z$  una funzione lineare a tratti da  $[0,1] \rightarrow \mathbb{R}$  definita come:

$$Z\left(\frac{t}{n}\right) = Y_t, \quad \text{per } t = 0, \dots, n.$$

Per  $n \rightarrow \infty$ , la distribuzione risultante è un moto browniano. Intuitivamente, aumentando sempre di più le  $n$  suddivisioni dell'intervallo  $[0,1]$ ,  $Z$  si avvicina sempre di più a un processo di Wiener, in particolare la distribuzione  $Z(t)$

---

<sup>20</sup> F. CARAVENNA, *Moto browniano e analisti stocastica*, 2011.

<sup>21</sup> R.RENÒ, *Polline, atomi e Wall Street: l'imprevedibile viaggio del moto browniano*, 2014.

converge a  $N(0, t)^{22}$ . Partendo dal moto browniano standard si possono costruire diverse generalizzazioni, modificando le proprietà iniziali, in particolare si prende in esame il moto browniano aritmetico.

*Definizione.* (Moto browniano aritmetico). Dato uno spazio di probabilità  $(\Omega, \mathcal{F}, P)$ , il processo stocastico continuo  $\{B_t^A\}_{t \in T}$  è un moto browniano aritmetico se soddisfa le proprietà:

- i.  $P(B_0^A = 0) = 1$ ;
- ii. Gli intervalli  $B^A(t_i) - B^A(s_i)$  sono reciprocamente indipendenti per  $[s_i, t_i]$  non sovrapposti,  $\forall i$ ;
- iii.  $\forall 0 < s < t, B^A(t) - B^A(s) \sim N[\mu(t - s), \sigma^2(t - s)]$ ;
- iv.  $t \rightarrow B_t^A(\omega)$  funzione continua.

I parametri  $\mu$  e  $\sigma^2$  sono detti drift e diffusione del processo. Per un moto browniano aritmetico si ha che  $E(B_t^A) = \mu t$  e  $\sigma^2(B_t^A) = \sigma^2 t$ , inoltre può essere ottenuto tramite combinazione lineare di un SMB come:

$$B_t^A = \mu t + \sigma^2 B_t.$$

Un ulteriore modo per generalizzare il SBM consiste nell'imporre a  $B_t$  il passaggio dall'origine al tempo  $t = 0$  e  $t = T$ . Tale processo è detto ponte browniano  $B_t^B$ :

$$B_t^B = \{B_t | B_T = 0\}, \quad t \in [0, T].$$

---

<sup>22</sup> P. KEMPTHORNE, C. LEE, V. STRELA, J. XIA, 18.S096, *Topics in Mathematics with Applications in Finance*, Fall 2013, Massachusetts Institute of Technology: MIT.

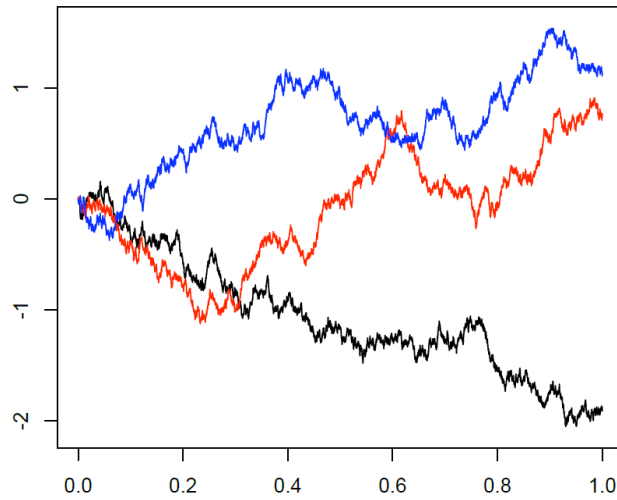


Fig. 3 Traiettorie di un moto browniano. [30]

Il moto browniano semplice è un processo stocastico continuo ma non differenziabile in ogni punto del dominio, ovvero l'espressione:

$$\frac{dB_t}{dt},$$

non ha significato e il calcolo differenziale classico non può essere usato. Per giustificare questa proprietà, che rende necessario lo sviluppo di un calcolo differenziale stocastico alternativo, si introduce il concetto di variazione quadratica di una funzione.

*Definizione.* (Variazione quadratica). Data una funzione  $f(t)$  definita nell'intervallo  $t \in [0, T]$ , considerando una partizione  $\Pi = \{t_0, t_1, \dots, t_n\}$ , con  $0 = t_0 < t_1 < \dots < t_n = T$ , si definisce variazione quadratica di  $f$  rispetto a  $T$  l'espressione:

$$[f, f](T) := \lim_{\|\Pi\| \rightarrow 0} \sum_{i=1}^{n-1} [f(t_{i+1}) - f(t_i)]^2.$$

Per funzioni continue e derivabili, si può dimostrare che la variazione quadratica è nulla e non ha dunque un particolare ruolo nel calcolo classico<sup>23</sup>. La non

---

<sup>23</sup> S. E. SHREVE, *Stochastic calculus for finance II, continuous time models*, 2004, cap. 3.

derivabilità del SBM conduce a un diverso risultato della variazione quadratica, che evidenzia l'elevata fluttuazione stocastica intrinseca a questo tipo di processo.

*Teorema.* (Variazione quadratica SBM). Sia  $\{B_t\}_{t \in T}$  un moto browniano standard in un intervallo  $[0, T]$ , si considera  $\Pi = \{t_0, t_1, \dots, t_n\}$ , con  $0 = t_0 < t_1 < \dots < t_n = T$ . Allora, con  $P = 1$ , vale la seguente condizione,  $\forall T > 0$ :

$$[B, B](T) := \lim_{\|\Pi\| \rightarrow 0} \sum_{i=0}^{n-1} [B(t_{i+1}) - B(t_i)]^2 = T.$$

*Dimostrazione.* Prendiamo in considerazione  $\sum_i [B(t_{i+1}) - B(t_i)]^2$ . Dalla condizione di stazionarietà del moto browniano,  $B(t_{i+1}) - B(t_i) \sim N(0, t_{i+1} - t_i) = N\left(0, \frac{T}{n}\right)$ , possiamo scrivere la somma precedente come  $\sum_{i=0}^{n-1} X_i^2$ , con  $X_i \sim N\left(0, \frac{T}{n}\right)$ .  $E[X_i^2] = \frac{T}{n}$ . Considerando  $Y_i$  variabili casuali i.i.d. con  $E[Y_i] = \frac{T}{n}$ , si ha:

$$\sum_i X_i^2 = \sum_i Y_i = n \left( \frac{1}{n} \sum_i Y_i \right) = T, \text{ per } n \rightarrow \infty$$

In quanto, per la legge forte dei grandi numeri  $\frac{1}{n} \sum_i Y_i \rightarrow \frac{T}{n}$ , per  $n \rightarrow \infty$ <sup>24</sup>.

In termini differenziali, la variazione quadratica del moto browniano standard si può scrivere come:

$$dBdB = (dB)^2 = dt.$$

## 1.5 Cenni di calcolo stocastico: l'integrale di Ito

---

<sup>24</sup> P. KEMPTHORNE, C. LEE, V. STRELA, J. XIA, 18.S096, *Topics in Mathematics with Applications in Finance*, Fall 2013, Massachusetts Institute of Technology: MIT.

La variazione quadratica non nulla e la non differenziabilità, in ogni punto, rende inutilizzabili, per il moto browniano, i classici metodi di calcolo differenziale e integrale. In questo paragrafo, si cerca di fornire un'idea, senza pretesa di esaustività, dall'integrazione stocastica. In particolare, dato  $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ , sia  $\{B_t\}_{t \in T}$  un moto browniano standard  $\mathcal{F}_t$ -misurabile e  $\Delta(t)$  un processo stocastico non anticipativo, fissato  $T > 0$  si è interessati a definire formalmente l'integrale stocastico:

$$\int_0^T \Delta(t) dB(t).$$

Nonostante  $B_t$  non sia differenziabile è possibile definire gli incrementi istantanei come  $B(t + \Delta t) - B(t)$ , per  $\Delta t \rightarrow 0$ . La definizione rigorosa dell'integrale precedente è stata formulata nel 1944 dal matematico giapponese Kiyoshi Ito, perciò spesso è chiamato integrale di Ito. Per dare un significato a tale espressione, il processo integrando  $\Delta(t)$  deve essere non anticipativo rispetto al moto browniano.

*Definizione.* (Processo non anticipativo). Dato  $\{B_t\}_{t \in T}$  un SBM adattato alla filtrazione  $\mathcal{F}_t$ , un processo stocastico  $\{X_t\}_{t \in T}$  si dice non anticipativo rispetto a  $B_t$  se:

- i.  $X_t$  è  $\mathcal{F}_t$ -misurabile per ogni  $t \geq 0$ ;
- ii.  $X_t$  è indipendente dagli incrementi  $B(s) - B(t)$ , con  $s > t$ .

Questa definizione implica che l'informazione al tempo  $t$  è sufficiente per valutare  $\Delta(t)$ , il quale dipende solo dai valori passati e presenti del SBM ma non dagli incrementi futuri. Per formalizzare la definizione dell'integrale di Ito si considera inizialmente il caso in cui  $\Delta(t)$  sia un processo semplice e poi si

estende ad integrandi generici come limite dell'integrale di funzioni semplici<sup>25</sup>.

Data  $\Pi = \{t_0, t_1, \dots, t_n\}$  una partizione di  $[0, T]$ ,  $0 = t_0 < t_1 < \dots < t_n = T$ , si assume che  $\Delta(t)$  sia costante in ogni sotto intervallo  $[t_i, t_{i+1})$ . Intuitivamente, per interpretare la relazione tra  $\Delta(t)$  e  $B(t)$ , si può immaginare che  $B(t)$  rappresenti il prezzo di un'azione al tempo  $t$  e  $\Delta(t_0), \Delta(t_1), \dots, \Delta(t_{n-1})$  sia il numero di azioni comprate in ogni periodo e mantenute fino a quello successivo.

Il guadagno per ogni tempo  $t$  è dato da:

$$I(t) = \Delta(t_0)[B(t) - B(0)] = \Delta(0)B(t), \quad 0 \leq t \leq t_1,$$

$$I(t) = \Delta(0)B(t_1) + \Delta(t_1)[B(t) - B(t_1)], \quad t_1 \leq t \leq t_2,$$

$$I(t) = \Delta(0)B(t_1) + \Delta(t_1)[B(t_2) - B(t_1)] + \Delta(t_2)[B(t) - B(t_2)], \quad t_3 \leq t \leq t_2.$$

In generale, per  $t_k \leq t \leq t_{k+1}$ , si ha:

$$I(t) = \sum_{i=0}^{k-1} \Delta(t_i)[B(t_{i+1}) - B(t_i)] + \Delta(t_k)[B(t) - B(t_k)].$$

Il processo  $I(t)$ , così definito, è l'integrale di Ito del processo semplice  $\Delta(t)$ :

$$I(t) = \int_0^t \Delta(u)dB(u),$$

ove scegliendo  $t = T$  si ottiene l'espressione integrale iniziale. Si può notare che  $I(t)$  non è un numero, come per l'integrale classico, ma è a sua volta un processo stocastico,  $\{I(t)\}_{t \in T}$ . A questo punto, è possibile generalizzare la costruzione dell'integrale di Ito per qualunque  $\Delta(t)$ , processo stocastico non anticipativo a quadrato integrabile<sup>26</sup>, approssimandolo tramite l'integrale di

<sup>25</sup> S. E. SHREVE, *Stochastic calculus for finance II, continuous time models*, 2004, cap. 4.

<sup>26</sup> Una funzione  $f$  misurabile si dice a quadrato integrabile, o appartenente allo spazio  $L^2(X, P)$ , se vale che  $\int |f|^2 dP < \infty$ . Se  $f \in L^2$ , si definisce la norma  $L^2$  di  $f$  come  $\|f\|_{L^2} := (\int |f|^2 dP)^{\frac{1}{2}} < \infty$ . In probabilità gli spazi  $L^2$ , detti spazi di Hilbert, sono molti utilizzati poiché, in modo intuitivo, si può dire che permettono di caratterizzare, oltre al valore atteso, la varianza di una variabile aleatoria.

processi semplici<sup>27</sup>. Definendo una partizione  $0 = t_0 < t_1 < \dots < t_n = T$ , si sceglie come approssimazione il processo semplice  $\Delta(t_i)$  in ogni  $t_i$ , costante nell'intervallo  $[t_i, t_{i+1})$ . Facendo tendere a zero la dimensione della partizione il processo semplice approssimerà sempre meglio  $\Delta(t)$ , ovvero si può scegliere una sequenza  $\Delta_n(t)$  di processi semplici tali che per  $n \rightarrow \infty$ ,  $\Delta_n(t)$  converge in media quadratica<sup>28</sup> a  $\Delta(t)$ , cioè:

$$\lim_{n \rightarrow \infty} E \int_0^T |\Delta_n(t) - \Delta(t)|^2 dt = 0.$$

Si definisce quindi l'integrale di Ito per un processo continuo  $\Delta(t)$  come:

$$I(t) = \int_0^t \Delta(u) dB(u) = \lim_{n \rightarrow \infty} \int_0^t \Delta_n(u) dB(u), \quad 0 \leq t \leq T.$$

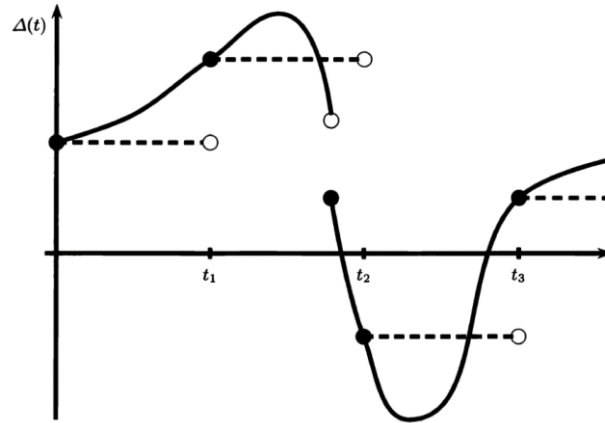


Fig. 4 Approssimazione di un processo integrando continuo. [71]

<sup>27</sup> Questo risultato è un'applicazione del teorema di convergenza dominata di Lebesgue, secondo cui, dato  $(X, \mathcal{F}, P)$ , si considera una sequenza di funzioni misurabili  $\{f_n\}$ ,  $f_n: X \rightarrow \mathbb{R}$  in modo che  $\exists f: X \rightarrow \mathbb{R}$ , tale che  $\lim_{n \rightarrow \infty} f_n(x) = f(x)$ ,  $\forall x \in X$ ,  $\forall n \in \mathbb{N}$ . Se  $\exists g$ , funzione a quadrato integrabile, ovvero  $g \in L^2(P)$ :  $\int |g|^2 dP < \infty$ , tale che  $|f_n(x)| \leq g(x)$ ,  $\forall x \in X$ ,  $\forall n \in \mathbb{N}$ , allora  $f \in L^2(P)$  e  $\lim_{n \rightarrow \infty} \int f_n dP = \int f dP$ , ovvero  $f_n \rightarrow f$  in media quadratica.

<sup>28</sup> Una sequenza di variabili casuali  $X_n$  converge in media  $r$ -esima alla variabile casuale  $X$  se  $E(|X_n|^r)$  e  $E(|X|^r)$  esistono finiti e  $\lim_{n \rightarrow \infty} E(|X_n - X|^r) = 0$ . Se  $r = 2$  si ha una convergenza quadratica media, che si può scrivere come  $X_n \xrightarrow{L^2} X$ .



Nel caso in cui  $\Delta(t)$  sia una funzione deterministica  $f(t)$ , che dipende solo dal tempo  $t \in T$ , l'integrale:

$$I(t) = \int_a^b f(t)dB(t),$$

è chiamato integrale di Wiener, ed è una variabile casuale con media nulla e varianza  $\int_a^b f(t)^2 dt$ . Inoltre, l'integrale di Wiener<sup>29</sup> si può definire solo se  $f(t)$  è a quadrato integrabile e  $B(t)$  ha incrementi indipendenti. Definendo l'integrale di Ito in termini di guadagni delle compravendite rispetto a  $B_t$ , si può intuire che  $I(t)$  non segue particolari tendenze, ovvero è una martingala, in quanto, per  $0 < s < t < T$ , vale che:

$$E(I_t | \mathcal{F}_s) = I_s.$$

Inoltre, per l'integrale di Ito, si possono dimostrare le seguenti proprietà:

- i. (continuità).  $I(t)$  è continuo e  $\mathcal{F}_t$ -misurabile per ogni  $t$ ;
- ii. (linearità). Se  $I(t) = \int_0^t \Delta(u)dB(u)$  e  $J(t) = \int_0^t \Psi(u)dB(u)$  allora  

$$I(t) \pm J(t) = \int_0^t [\Delta(u) \pm \Psi(u)] dB(u);$$
- iii. (isometria di Ito).  $E \left[ \left( \int_0^t \Delta(u)dB(u) \right)^2 \right] = E \left( \int_0^t \Delta^2(u)dB(u) \right);$
- iv. (variazione quadratica).  $[I, I](t) = \int_0^t \Delta^2(u)du.$

La variazione quadratica dell'integrale di Ito si può anche scrivere, in forma differenziale, come:

$$dI(t)dI(t) = \Delta^2(t)dB(t)dB(t) = \Delta^2(t)dt.$$

Si può notare quindi che  $[I, I](t)$  è un processo stocastico e  $I(t)$  accumula una variazione quadratica di  $\Delta^2(t)$  per ogni unità di tempo<sup>30</sup>. In ultimo, dopo aver

---

<sup>29</sup> La costruzione dell'integrale di Wiener è simile a quella dell'integrale di Ito, in quanto si può approssimare  $f(t)$  con una sequenza di funzioni semplici  $\Delta_n(t)$ , le quali convergono in media quadratica a  $f(t)$ , per  $n \rightarrow \infty$ , dal teorema di convergenza di Lebesgue.

<sup>30</sup> S. E. SHREVE, *Stochastic calculus for finance II, continuous time models*, 2004, cap. 4.

definito l'integrale stocastico, si può introdurre un ulteriore tipo di processo stocastico, noto come processo di Ito, la cui applicazione in finanza sarà esaminata successivamente.

*Definizione.* (Processo di Ito). Dato  $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ , il processo stocastico  $\{X_t\}_{t \in [0, \infty)}$  è detto processo di Ito se può essere rappresentato nella forma:

$$X_t = X_0 + \int_0^t g(u) du + \int_0^t h(u) dB(u),$$

con  $X_0 \in \mathbb{R}$  condizione iniziale,  $B(u)$  moto browniano standard,  $g$  funzione  $\mathcal{F}_t$ -misurabile per ogni  $t$  e integrabile,  $h \in L^2$  funzione a quadrato sommabile. L'integrale  $\int g du$  è detto drift o deriva ed è un processo continuo a variazione finita,  $g \in L^1$ ,  $\int |g| du < \infty$ , invece  $\int h dB(u)$  è un integrale di Ito, detto diffusione, con  $\int |h|^2 du < \infty$ . In applicazioni finanziarie, di solito,  $g(t)$  indica il rendimento di un portafoglio,  $\mu_t$ , e  $h(t)$  indica la deviazione standard,  $\sigma_t$ <sup>31</sup>. In generale, il processo di Ito può essere scritto, in forma differenziale, come:

$$dX_t = a(X_t, t)dt + b(X_t, t)dB_t.$$

## 1.6 Cenni di calcolo stocastico: il lemma di Ito

Utilizzando il moto browniano come modello, ad esempio, per le fluttuazioni dei mercati, spesso si è interessati a valutare l'espressione:

$$f(B_t),$$

stimandone le differenze su un intervallo di tempo infinitesimale. (si suppone  $f$  funzione liscia di classe  $C^\infty$ ). A causa della non differenziabilità di  $B_t$  non si può scrivere, come nel caso classico:

---

<sup>31</sup> S. PRINA, D. CARACCIOLO, *Finanza matematica*, 2013.

$$df = f'(B_t)dB_t.$$

Per capirne il motivo, si considera lo sviluppo di Taylor della funzione  $f(x)$ :

$$f(x + \Delta x) = f(x) + f'(x)\Delta x + f''(x)\frac{(\Delta x)^2}{2!} + f'''(x)\frac{(\Delta x)^3}{3!} + \dots$$

In forma differenziale, quando  $\Delta x \rightarrow 0$ , per piccoli incrementi, i termini di secondo e terzo grado possono essere trascurati e, ponendo  $df = f(x + \Delta x) - f(x)$ , si ottiene quindi:

$$df = f'(x)dx.$$

Se si scrive lo sviluppo di Taylor per un moto browniano,  $x = B_t$ , si ottiene:

$$\begin{aligned} f(B_t + \Delta B_t) - f(B_t) \\ = f'(B_t)(\Delta B_t) + f''(B_t)\frac{(\Delta B_t)^2}{2!} + f'''(B_t)\frac{(\Delta B_t)^3}{3!} + \dots \end{aligned}$$

Si nota come, a causa della variazione quadratica non nulla,  $(\Delta B)^2 = \Delta t$ , il termine di secondo grado non può essere trascurato. Scrivendo in forma differenziale, si ottiene il lemma di Ito per funzioni con una variabile:

$$df = f'(B_t)dB_t + \frac{1}{2}f''(B_t)dt.$$

Si può generalizzare l'espressione appena ottenuta, considerando una funzione con due variabili del tipo:

$$f(t, B_t).$$

Seguendo gli stessi passi del caso precedente, si considera lo sviluppo di Taylor per una generica funzione  $f(t, x)$ :

$$\begin{aligned} f(t + \Delta t, x + \Delta x) \\ = f(t, x) + \frac{\partial f}{\partial t}(t, x)\Delta t + \frac{\partial f}{\partial x}(t, x)\Delta x \\ + \frac{1}{2}\left[\frac{\partial^2 f}{\partial t^2}(x, t)(\Delta t)^2 + 2\frac{\partial^2 f}{\partial t \partial x}(x, t)\Delta t\Delta x + \frac{\partial^2 f}{\partial x^2}(x, t)(\Delta x)^2\right]. \end{aligned}$$

Lo stesso sviluppo, in forma differenziale, per un moto browniano è dato da:

$$\begin{aligned}
& f(t + dt, B_t + dB_t) - f(t, B_t) \\
&= \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dB_t \\
&+ \frac{1}{2} \left( \frac{\partial^2 f}{\partial t^2} (dt)^2 + 2 \frac{\partial^2 f}{\partial t \partial x} dt dB_t + \frac{\partial^2 f}{\partial x^2} (dB_t)^2 \right)
\end{aligned}$$

Si considera nuovamente la variazione quadratica, che impedisce di approssimare il termine quadratico, e si nota come  $(dt)^2$  e  $dt dB_t$  siano trascurabili rispetto a  $dt^{3/2}$ , in modo da ottenere una versione più generale del lemma di Ito:

$$df(t, B_t) = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dB_t + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} dt = \left( \frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} \right) dt + \frac{\partial f}{\partial x} dB_t.$$

L'esistenza del lemma di Ito rende il moto browniano un modello molto utile, poiché consente di compiere analisi differenziali di processi stocastici. In ultimo, prendiamo in esame una versione del lemma di Ito ancora più generale, che considera un processo stocastico di Ito,  $X_t$ .

*Teorema.* (Lemma di Ito). Sia  $f(t, x)$  una funzione liscia a due variabili e sia  $X_t$  un processo di Ito, che soddisfa  $dX_t = \mu_t dt + \sigma_t dB_t$ , con  $B_t$  moto browniano classico,  $\mu_t$  termine di drift e  $\sigma_t$  parametro di diffusione. Allora,

$$df(t, X_t) = \left( \frac{\partial f}{\partial t} + \mu_t \frac{\partial f}{\partial x} + \frac{1}{2} \sigma_t^2 \frac{\partial^2 f}{\partial x^2} \right) dt + \sigma_t \frac{\partial f}{\partial x} dB_t.$$

*Dimostrazione.*

$$\begin{aligned}
df(t, X_t) &= \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dX_t + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} (dX_t)^2 \\
&= \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} (\mu_t dt + \sigma_t dB_t) + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} (\mu_t dt + \sigma_t dB_t)^2
\end{aligned}$$

---

<sup>32</sup>  $dt dB_t$  può essere visto come  $dt dt^{1/2} = dt^{3/2}$ , quindi risulta anch'esso trascurabile rispetto a  $dt$ .

Dallo sviluppo del quadrato,  $(\mu_t dt + \sigma_t dB_t)^2 = (\mu_t^2 (dt)^2 + \sigma_t^2 (dB_t)^2 + 2\mu_t \sigma_t dt dB_t)$ , rimane solamente il termine  $\sigma_t^2 (dB_t)^2 = \sigma_t^2 dt$ . Raccogliendo  $dt$  si ricava<sup>33</sup>:

$$df(t, X_t) = \left( \frac{\partial f}{\partial t} + \mu_t \frac{\partial f}{\partial x} + \frac{1}{2} \sigma_t^2 \frac{\partial^2 f}{\partial x^2} \right) dt + \sigma_t \frac{\partial f}{\partial x} dB_t.$$

## *Appendice A. Una simulazione di cammino casuale: The Drunkard's Walk*

Il random walk ha molte applicazioni pratiche, oltre che teoriche. Fino ad ora, è stato analizzato enunciandone le principali proprietà e caratteristiche all'interno della più ampia categoria dei processi stocastici, sia con tempo discreto sia con tempo continuo, nella sua estensione di moto browniano. In questa appendice si adotterà un approccio diverso, sfruttando una simulazione della nota passeggiata dell'ubriaco<sup>34</sup>. Si immagina un ubriaco che parte dal centro di un campo e ogni secondo compie un passo in una direzione casuale. Attraverso questa simulazione si cercherà di studiare l'evoluzione di questo moto e la distanza dall'origine, all'aumentare del numero di passi. Per semplicità, si considera il caso in cui la persona si muove solamente lungo direzioni parallele agli assi cartesiani. Al tempo  $t = 0$ , l'ubriaco occuperà la posizione  $[0,0]$ , con possibilità di muoversi, nell'istante successivo, nelle posizioni  $[1,0]$ ,  $[0,1]$ ,  $[-1,0]$ ,  $[0,-1]$ , con la stessa probabilità,  $P = 1/4$ . Nel

---

<sup>33</sup> P. KEMPTHORNE, C. LEE, V. STRELA, J. XIA, 18.S096, *Topics in Mathematics with Applications in Finance*, Fall 2013, Massachusetts Institute of Technology: MIT.

<sup>34</sup> J.V. GUTTAG, *Introduction to computation and programming using Python with application to understanding data*, 2016, cap.14.

caso si sia mosso in posizione  $[1,0]$ , in  $t = 2$ , esso potrà muoversi nelle posizioni  $[2,0]$ ,  $[0,0]$ ,  $[1,-1]$ ,  $[1,1]$ , sempre con la stessa probabilità, e così via. Se si volesse calcolare la distanza dal centro più probabile al secondo passo, essa sarebbe di 2 unità con  $P = 0,25$ , di 0 unità con  $P = 0,25$  e di  $\sqrt{2}$  unità con  $P = 0,5$ . All'aumentare del numero di passi si può calcolare la distanza media percorsa<sup>35</sup>.

```
UsualDrunk random walk of 10 steps
Mean = 2.863
Max = 7.2 Min = 0.0
UsualDrunk random walk of 100 steps
Mean = 8.296
Max = 21.6 Min = 1.4
UsualDrunk random walk of 1000 steps
Mean = 27.297
Max = 66.3 Min = 4.2
UsualDrunk random walk of 10000 steps
Mean = 89.241
Max = 226.5 Min = 10.0
```

Dopo aver compiuto 10, 100, 1000, 10000 passi l'ubriaco si troverà ad una distanza media di 2,86, 8,29, 27,29 e 89,24. Sembra evidente una relazione positiva tra numero di passi e distanza dall'origine, come confermato nella *Fig.5*. Essa mostra come l'ubriaco si trovi, in media, all'interno di un cerchio, di centro nell'origine e di raggio pari alla distanza media, tuttavia fornisce poche informazioni su dove potrebbe trovarsi alla fine di una particolare sequenza.

---

<sup>35</sup> E. GRIMSON, J. GUTTAG, A. BELL, 6.0002, *Introduction to Computational Thinking and Data Science*, Fall 2016, Massachusetts Institute of Technology: MIT.

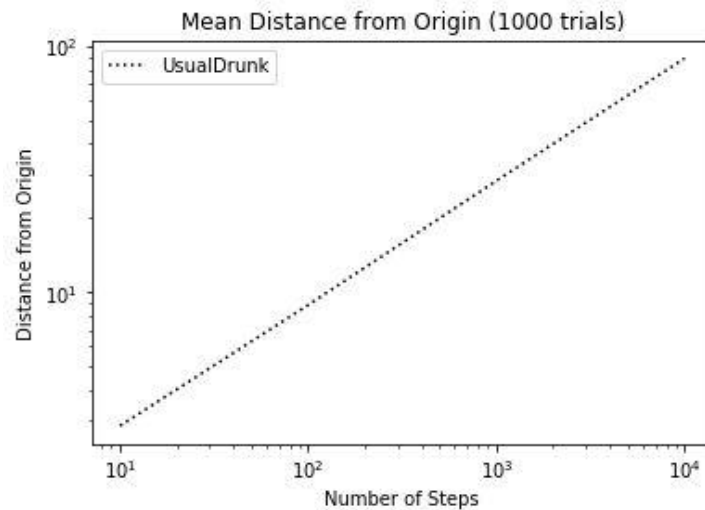


Fig. 5 Distanza media dall'origine dell'ubriaco.

Si costruisce ora un ulteriore di grafico, sul quale è raffigurato il percorso dell'ubriaco per un singolo ammontare di passi e per una singola simulazione.

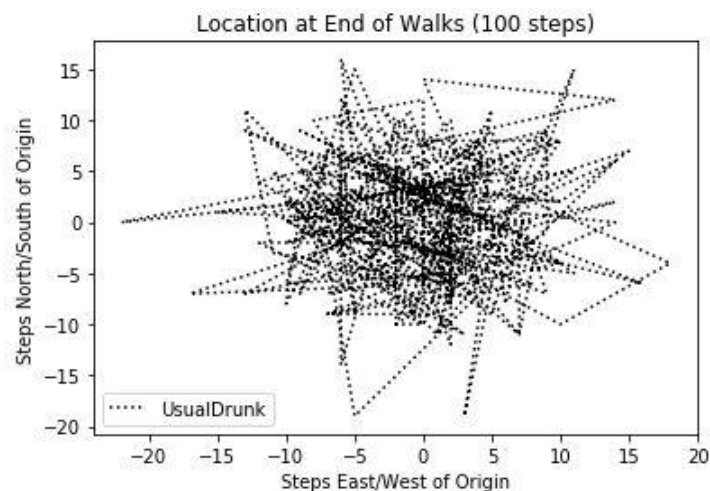


Fig. 6 Simulazione della posizione dell'ubriaco dopo 100 passi.

Il moto rappresenta una corretta raffigurazione di un random walk. Seguendo traiettorie casuali, l'ubriaco non sembra in realtà allontanarsi molto dall'origine e sembra inoltre che passi infinite volte dal punto  $[0,0]$ <sup>36</sup>.

<sup>36</sup> Si può dimostrare che la frequenza dei ritorni all'origine si abbassa all'aumentare dei passi ma, per un numero infinito di passi la probabilità di ritorno tende a 1.





## **CAPITOLO II.**

### **LA TEORIA FINANZIARIA CLASSICA E IL MERCATO EFFICIENTE**

## 2.1 L. Bachelier e il nuovo modo di concepire la finanza

Nel precedente capitolo sono stati definiti i concetti di random walk e di moto browniano, accennando a come essi possano essere applicati allo studio dei mercati finanziari. Per approfondire questa applicazione è necessario fare riferimento al lavoro di Louis Bachelier, un matematico francese che ha gettato le basi della teoria economica classica, studiando le fluttuazioni dei mercati in ottica probabilistica. Si tratta di un approccio del tutto nuovo e rivoluzionario che, anni prima della definizione formale di Wiener del moto browniano come processo stocastico, ne anticipava molte conclusioni. Nel marzo del 1900, Bachelier, ha presentato, all'università di Parigi, la sua tesi di dottorato, *Théorie de la Spéculation*, davanti a una commissione di noti matematici, tra cui Henri Poincaré. La tesi non è stata particolarmente apprezzata in quanto, nonostante fossero state riconosciute varie intuizioni originali, si occupava di un ambito che non suscitava un particolare interesse, a quel tempo, nel mondo accademico. La tesi non gli ha procurato grande notorietà e, dal lato economico, è stata ignorata fino agli anni 60', nei quali ottiene il giusto riconoscimento, diventando un punto fermo delle teorie economiche sviluppate. L'idea di partenza è che i fattori che influenzano le fluttuazioni dei mercati finanziari sono infiniti e impossibili da prevedere matematicamente. Lo studio dei mercati tramite la teoria statistica non porterà mai a conclusioni esatte, tuttavia è possibile stabilire delle leggi per le oscillazioni dei prezzi, attraverso il calcolo delle probabilità<sup>37</sup>. Bachelier si concentra principalmente sul mercato delle opzioni, un tipo di strumento finanziario che fornisce al possessore il diritto di comprare o vendere

---

<sup>37</sup> L. BACHELIER, *Théorie de la Spéculation*, 1900.

un asset, chiamato sottostante, a un prezzo prefissato in una data futura, a fronte di un premio pagato. Un'opzione call fornisce il diritto di comprare, un'opzione put il diritto di vendere<sup>38</sup>. Facendo riferimento a questo tipo di mercato si possono considerare due differenti tipi di probabilità:

- i. Una probabilità matematica, che può essere determinata a priori, studiata nella sua opera;
- ii. Una probabilità condizionata a eventi futuri, impossibile da calcolare matematicamente.

Bachelier definisce la speranza matematica, come il prodotto tra il profitto potenziale e la corrispondente probabilità, e la speranza matematica totale, come somma delle singole speranze di un gioco. Il gioco è equo se la speranza totale è nulla. Il concetto di speranza dà un'indicazione sulla vantaggiosità del gioco, tuttavia non fornisce un coefficiente che ne quantifica il valore intrinseco. Ciò porta a introdurre il vantaggio matematico di un gioco come il rapporto tra le speranze positive e le speranze totali. Esso è compreso tra 0 e 1 e vale  $\frac{1}{2}$  se il gioco è equo. Per il mercato, la speranza totale è nulla, non solo sulle operazioni correnti, in un dato istante, ma anche su quelle future. In altre parole, la speranza matematica di un'operazione che ha prodotto una variazione del prezzo è nulla a priori. L'idea è che in un'analisi ex post si possono notare le variazioni e inferire sulle cause ma, facendo riferimento al primo tipo di probabilità a priori, non c'è ragione per aspettarsi un cambiamento positivo o negativo, quindi la miglior previsione possibile è che i prezzi si mantengano costanti, in quanto riflettono già le informazioni disponibili<sup>39</sup>. Da queste considerazioni, appare chiaro come il mercato descritto da Bachelier sia equiparabile a un cammino

---

<sup>38</sup> J. BERK, P. DEMARZO, *Corporate finance*, 2013, cap.20.

<sup>39</sup> B. B. MANDELBROT, *Il disordine dei mercati. Una visione frattale di rischio, rovina e redditività*, 2005, cap. 3.

casuale. La funzione di probabilità che lo descrive si può immaginare su un grafico che mette in relazione probabilità e prezzi a un tempo dato, dove l'origine corrisponde al prezzo equo, con deviazioni indipendenti, inoltre essa è simmetrica rispetto al prezzo equo e presenta un massimo in tale punto. Si può intuire come tale funzione contenga le caratteristiche della funzione normale, in particolare la legge di probabilità ottenuta da Bachelier è del tipo:

$$p = \frac{1}{2\pi k\sqrt{t}} e^{-\frac{x^2}{4\pi k^2 t}},$$

ove  $k$  è la speranza matematica.

Giudicando le fluttuazioni indipendenti tra loro e normalmente distribuite, Bachelier ha anticipato la condizione di assenza di memoria di Markov e, tramite quella che oggi viene definita equazione Chapman-Kolmogorov<sup>40</sup>, ha identificato un'analogia con le equazioni di diffusione del calore di Fourier<sup>41</sup>, definendo una legge di radiazione di probabilità. Si suppone che, al tempo  $t$ , i prezzi  $X_{n-2}, X_{n-1}, X_n, X_{n+1}, X_{n+2}, \dots$ , differiscano di una quantità  $\Delta x$  e abbiano, rispettivamente, probabilità  $p_{n-2}, p_{n-1}, p_n, p_{n+1}, p_{n+2}$ . Conoscendo la distribuzione di probabilità al tempo  $t$  è possibile dedurre quella al tempo  $t + \Delta t$ . Se al tempo  $t$  si è realizzato  $X_n$ , al tempo  $t + \Delta t$  potranno realizzarsi  $X_{n-1}$  e  $X_{n+1}$ , entrambi con uguale probabilità  $\frac{p_n}{2}$ . Se al tempo  $t + \Delta t$  si è realizzato  $X_{n-1}$ , al tempo  $t$  si sono realizzati  $X_n$  o  $X_{n-2}$ , la probabilità di  $X_{n-1}$  sarà quindi  $\frac{(p_n + p_{n-2})}{2}$ , quella di  $X_n$ , allo stesso tempo, sarà  $\frac{(p_{n-1} + p_{n+1})}{2}$ , e così via. Questo si riassume dicendo che ogni prezzo trasmette, in un intervallo di tempo

---

<sup>40</sup> Questa equazione descrive la probabilità di transizione di un processo da uno stato  $i$  a uno  $j$  in un tempo  $t$ , come somma delle probabilità di transizione per qualche stato intermedio  $k$ :

$$p_{ij}(t_1 + t_2) = \sum_{k \in S} p_{ik}(t_1) p_{kj}(t_2) \quad \forall i, j \in S, \forall t_1, t_2 \geq 0.$$

<sup>41</sup> M. DAVIS, A. ETHERIDGE, *Louis Bachelier's theory of speculation: the origins of modern finance*, 2006.

elementare, una quantità di probabilità ai prezzi vicini, proporzionale alla differenza delle loro probabilità. Si considera ora la probabilità  $\mathcal{P}$  che il prezzo  $x$  sia situato al tempo  $t$  nell'intervallo  $[x, +\infty]$ ,  $\mathcal{P} = \int_x^{+\infty} p \, dx$ , e si cerca di valutare come essa vari, in un intervallo di tempo  $\Delta t$ . Se si denota con  $p = -\frac{\partial \mathcal{P}}{\partial x}$  la probabilità di  $x$  al tempo  $t$ , la relazione che descrive l'ammontare di probabilità che attraversa il prezzo durante un intervallo di tempo  $\Delta t$  può essere espressa come:

$$\frac{1}{c^2} \left( p - \frac{\partial p}{\partial x} - p \right) \Delta t = -\frac{1}{c^2} \frac{\partial p}{\partial x} \Delta t = \frac{1}{c^2} \frac{\partial^2 \mathcal{P}}{\partial x^2} \Delta t,$$

con  $c$  costante arbitraria. Questa probabilità può essere espressa, allo stesso modo, come  $\left( \frac{\partial \mathcal{P}}{\partial t} \right) \Delta t$ . Eguagliando le due espressioni, si ottiene l'equazione di diffusione<sup>42</sup>:

$$\frac{1}{c^2} \frac{\partial^2 \mathcal{P}}{\partial x^2} \Delta t = \left( \frac{\partial \mathcal{P}}{\partial t} \right) \Delta t \rightarrow c^2 \frac{\partial \mathcal{P}}{\partial t} - \frac{\partial^2 \mathcal{P}}{\partial x^2} = 0.$$

## 2.2 La teoria del mercato efficiente e i livelli di efficienza informativa

Il lavoro di Bachelier, come detto, è stato ignorato fino agli anni 60', quando, dopo essere stato tradotto in inglese, vari economisti si sono interessati al suo modello, testandone le assunzioni. Tra questi è necessario fare riferimento al lavoro dell'economista statunitense Eugene Fama, che inizialmente si è preoccupato di testarne la rilevanza empirica e, in seguito, ha sviluppato quella

---

<sup>42</sup> L. BACHELIER, *Théorie de la Spéculation*, 1900.

che oggi è nota come teoria dei mercati efficienti<sup>43</sup>. Partendo dal modello di Bachelier, Fama giunge a una definizione più rigorosa dei fattori che influenzano i mercati, studiandone l'efficienza. Un mercato è detto efficiente se in esso i prezzi riflettono tutta l'informazione disponibile. In particolare, dato  $\Phi_t$  l'insieme delle informazioni disponibili al tempo  $t$ , Fama distingue tre differenti declinazioni del concetto di efficienza, che costituiscono tre differenti forme di mercato efficiente:

- i. Efficienza informativa debole, nella quale  $\Phi_t$  racchiude solamente le informazioni sui valori storici dei prezzi;
- ii. Efficienza informativa semi-forte, nella quale  $\Phi_t$  racchiude, oltre alle informazioni passate, quelle pubblicamente disponibili;
- iii. Efficienza informativa forte, nella quale  $\Phi_t$  racchiude la totalità delle informazioni, sia pubbliche che private.

In termini più rigorosi, si può formulare un modello di efficienza usando come strumento matematico il valore atteso dei rendimenti, descritto come:

$$E(\tilde{p}_{j,t+1}|\Phi_t) = [1 + E(\tilde{r}_{j,t+1}|\Phi_t)]p_{j,t},$$

dove  $p_{j,t}$  è il prezzo dello strumento  $j$  al tempo  $t$ ,  $\tilde{p}_{j,t+1}$  è la stima del prezzo dello strumento  $j$  al tempo  $t + 1$  e  $\tilde{r}_{j,t+1}$  è la stima del rendimento percentuale dell'asset  $t$  al tempo  $t + 1$ <sup>44</sup>. L'idea è che il valore atteso della stima del prezzo nel periodo successivo, date le informazioni disponibili, è uguale al prezzo attuale più il valore atteso della stima del rendimento che otterrà lo strumento nel periodo successivo. L'ipotesi che l'informazione è riflessa nel prezzo è descritta dal concetto di valore atteso condizionato, inoltre nella determinazione del rendimento atteso è utilizzata tutta l'informazione contenuta in  $\Phi_t$ .

---

<sup>43</sup> E. F. FAMA, *The behavior of stock market prices*, 1965.

<sup>44</sup> Il rendimento percentuale in un singolo periodo è dato da  $\frac{(p_{j,t+1}-p_{j,t})}{p_{j,t}}$ .

L'eccesso di valore nel mercato,  $x_{j,t+1}$ , è definito come la differenza tra il prezzo effettivamente osservato e l'aspettativa:

$$x_{j,t+1} = p_{j,t+1} - E(\tilde{p}_{j,t+1}|\Phi_t).$$

Quindi, dalle assunzioni deriva che:

$$E(\tilde{x}_{j,t+1}|\Phi_t) = 0.$$

In termini di rendimento si può scrivere, equivalentemente:

$$z_{j,t+1} = r_{j,t+1} - E(\tilde{r}_{j,t+1}|\Phi_t),$$

$$E(\tilde{z}_{j,t+1}|\Phi_t) = 0.$$

La successione dei rendimenti attesi è dunque un gioco equo, in riferimento alle informazioni possedute al tempo  $t$ , e il mercato è in equilibrio con nessuna possibilità di batterlo<sup>45</sup>.

### 2.2.1 Le specificazioni del modello del mercato efficiente

Basandosi sul modello appena esposto, Fama ha proposto due ulteriori specificazioni del modello del mercato efficiente, che portano a conclusioni interessanti. La prima sfrutta la nozione di submartingala, infatti se si assume che:

$$E(\tilde{p}_{j,t+1}|\Phi_t) \geq p_{j,t}, \quad \forall t, \Phi_t,$$

oppure in altri termini,

$$E(\tilde{r}_{j,t+1}|\Phi_t) \geq 0, \quad \forall t, \Phi_t,$$

si ottiene un modello di mercato in cui l'aspettativa del prezzo nel periodo successivo, date le informazioni al tempo  $t$ , è maggiore del prezzo attuale. Questo ha una importante implicazione economica, infatti, un qualunque meccanismo di trading, con questa assunzione, risulta sub-ottimale rispetto a una semplice strategia di 'buy and hold'. La seconda specificazione del mercato

---

<sup>45</sup> E.F. FAMA, *Efficient capital markets: a review of theory and empirical work*, 1970.

efficiente considera il caso di random walk. In particolare, si aggiunge un'ulteriore condizione, di identica distribuzione delle variazioni successive, a quella di indipendenza. Si può ipotizzare che il rendimento atteso non dipende dalle informazioni al tempo  $t$ :

$$E(\tilde{r}_{j,t+1}|\Phi_t) = E(\tilde{r}_{j,t+1}).$$

Ciò implica che l'ordine dei rendimenti passati non ha influenza sulla stima di quelli futuri. Questa risulta essere un'estensione più restrittiva del modello iniziale.

### *2.2.2 Le condizioni e le evidenze empiriche*

Le prime condizioni di esistenza del modello del mercato efficiente riguardano la sua formulazione matematica. In particolare, si suppone che la condizione di equilibrio sia espressa sotto forma di valore atteso e che il parametro  $\Phi_t$  racchiuda tutte le informazioni disponibili. Inoltre, possono essere determinate delle condizioni sufficienti, ma non necessarie, per un mercato pienamente efficiente:

- i. Assenza di costi di transazione nelle operazioni;
- ii. Accessibilità di tutta l'informazione disponibile, senza costi, da parte dei partecipanti al mercato;
- iii. Concordanza sull'impatto di  $\Phi_t$  sui prezzi correnti e su quelli futuri.

Verosimilmente, queste condizioni sono raramente soddisfatte da un mercato reale, tuttavia, non sono strettamente necessarie. Ad esempio, è sufficiente che una parte degli operatori abbia accesso alle informazioni disponibili per generare potenzialmente un mercato efficiente<sup>46</sup>. Si può affermare che le prove empiriche a favore di questo modello erano inizialmente solide e che, in seguito,

---

<sup>46</sup> E.F. FAMA, *Efficient capital markets: a review of theory and empirical work*, 1970.



sono aumentate quelle sfavorevoli<sup>47</sup>. Tra quelle favorevoli, è da sottolineare il fatto che il mercato sembra effettivamente impossibile da battere. Comparando i rendimenti di analisti finanziari professionisti con i rendimenti di titoli scelti casualmente, gli analisti finanziari sembrano non ottenere sistematicamente performance migliori<sup>48</sup>. Inoltre, si ritiene ragionevole l'ipotesi di incrementi indipendenti in quanto, esaminando serie storiche di prezzi non sembra esserci una correlazione significativa tra le variazioni. Per quanto riguarda gli argomenti contrari, negli ultimi anni, la generalizzata applicabilità dell'efficienza sembra essere messa in discussione. In particolare, si cita l'effetto piccole dimensioni, secondo cui le aziende di minore dimensione sembrano ottenere rendimenti superiori alle previsioni, per lunghi periodi di tempo<sup>49</sup>, e l'eccesso di reattività del mercato, secondo cui le reazioni dei prezzi agli annunci sono eccessive e gli eventuali errori si correggono lentamente<sup>50</sup>.

### *2.3 La teoria di portafoglio di Markowitz*

Le conseguenze che il lavoro di Bachelier ha avuto sulla teoria economica classica sono state numerose. Partendo dalla normalità della distribuzione delle variazioni dei prezzi possono essere analizzati diversi fenomeni e aspetti economici, tra i quali è importante citare l'opera dell'economista statunitense Harry Markowitz, che ha rivoluzionato il modo di analizzare le scelte finanziarie, studiando la composizione di un portafoglio di titoli, attraverso il

---

<sup>47</sup> F.S. MISHKIN, S. G. EAKINS, G. FORESTIERI, *Istituzioni e mercati finanziari*, 2015, cap.6.

<sup>48</sup> M.C. JENSEN, *The performance of mutual funds in the period 1945-64*, 1968.

<sup>49</sup> M.R. REINGANUM, *The Anomalous stock market behavior of small firms in January: empirical tests for tax-loss selling effects*, 1983.

<sup>50</sup> W.F.M. DE BONDT, R. THALER, *Further evidence on investor overreaction and stock market seasonality*, 1987.

rischio e il rendimento atteso. Non è nello scopo di questa tesi compiere un'ampia digressione sul tema del rischio e del rendimento, perciò si assumerà che essi possano essere ben descritti, all'interno della teoria di portafoglio, dalla varianza e dal valore atteso, precedentemente definiti in termini di momenti della distribuzione gaussiana. L'idea di partenza è che un investitore, con un capitale iniziale, difficilmente sceglierà di investirlo interamente in un unico titolo, in quanto, spesso, sceglierà di comporre un portafoglio di titoli, partizionando il proprio capitale. Questa operazione di diversificazione consente di aumentare il rendimento diminuendo il rischio, rispetto all'acquisto di un unico titolo. Il portafoglio che consente di ottenere il massimo rendimento al minimo rischio è chiamato portafoglio efficiente e, grazie all'analisi di Markowitz, ne possono essere studiate le proprietà. Il modello prende in esame un orizzonte temporale tipicamente uniperiodale. Si considera un portafoglio formato da  $n$  titoli rischiosi, con un rendimento al tempo  $t$ ,  $R_{i,t}$ . Il rendimento ha media  $\mu_{i,t}$  e varianza  $\sigma_{i,t}^2$ , inoltre si indica con  $w_i$  la frazione di capitale investita nell' $i$ -esimo asset del portafoglio, in modo che  $\sum_{i=1}^n w_i = 1$ ,  $w_i \geq 0$ . Conoscendo queste informazioni, al tempo  $t$ , è possibile calcolare la media e la varianza (o la deviazione standard) del portafoglio:

$$E[R] = E \left[ \sum_i w_i R_i \right] = \sum_i w_i E[R_i] = \sum_i w_i \mu_i = w' \mu.$$

$$\begin{aligned}
\sigma^2[R] &= E[(R - \mu)^2] = E\left[\left(\sum_i w_i(R_i - \mu_i)\right)^2\right] \\
&= E\left[\sum_i \sum_j w_i w_j (R_i - \mu_i)(R_j - \mu_j)\right] \\
&= \sum_i \sum_j E[w_i w_j (R_i - \mu_i)(R_j - \mu_j)] \\
&= \sum_i \sum_j w_i w_j \text{Cov}(R_i, R_j) = \sum_i \sum_j w_i w_j \sigma_{i,j} = w' \Sigma w,
\end{aligned}$$

con  $w = \begin{pmatrix} w_1 \\ w_2 \\ w_i \end{pmatrix}$  e  $\Sigma = [\sigma_{i,j}] = \begin{bmatrix} \sigma_{1,1} & \cdots & \sigma_{1,n} \\ \vdots & \ddots & \vdots \\ \sigma_{n,1} & \cdots & \sigma_{n,n} \end{bmatrix}$ , matrice delle covarianze.

Spesso al posto della covarianza si utilizza una misura standardizzata, il coefficiente di correlazione tra  $R_i$  e  $R_j$ , ottenuto come  $\rho_{i,j} = \frac{\sigma_{i,j}}{\sigma_i \sigma_j}$ . Per visualizzare l'insieme dei portafogli efficienti si considera, per semplicità, il caso in cui siano disponibili sul mercato solamente due titoli rischiosi in cui investire, in modo da ottenere un portafoglio con:

$$\begin{aligned}
E[R] &= w_1 \mu_1 + w_2 \mu_2 \\
\sigma^2[R] &= w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2 + 2w_1 w_2 \rho \sigma_1 \sigma_2.
\end{aligned}$$

Con queste considerazioni, sul piano media-scarto quadratico medio, si possono disegnare, al variare di  $w_1$  e  $w_2$ , tutte le possibili combinazioni tra i due titoli<sup>51</sup>. L'insieme delle combinazioni assume forma diversa in base al coefficiente di correlazione, in particolare:

- i. Se  $\rho = 1$ , c'è perfetta correlazione positiva tra i titoli e  $\sigma^2[R] = (w_1 \sigma_1 + w_2 \sigma_2)^2$ , da cui  $\sigma[R] = w_1 \sigma_1 + w_2 \sigma_2$ . Si ottiene una retta che congiunge i due portafogli;

---

<sup>51</sup> G. WEST, *An introduction to modern portfolio theory: Markowitz, CAP-M, APT and Black-Litterman*, 2004, cap. 1.

- ii. Se  $\rho = 0$ , c'è assenza di correlazione e  $\sigma^2[R] = w_1^2\sigma_1^2 + w_2^2\sigma_2^2$ . Si ottiene un arco di iperbole;
- iii. Se  $\rho = -1$ , c'è perfetta correlazione negativa e si ha  $\sigma^2[R] = (w_1\sigma_1 - w_2\sigma_2)^2$  e  $\sigma[R] = |w_1\sigma_1 - w_2\sigma_2|$ . Si ottiene una coppia di segmenti con vertice sull'asse verticale.

Tornando al caso generale, risolvendo il problema di minimizzazione vincolata della varianza per il vettore  $w$ , tramite il metodo dei moltiplicatori di Lagrange<sup>52</sup>, è possibile calcolare il portafoglio di rischio minimo, rappresentato dal vertice dell'iperbole. A partire da questo punto, si costruisce poi l'insieme dei portafogli efficienti, o frontiera efficiente, che è costituita dalla parte superiore dell'iperbole. In base alla propria curva di indifferenza, l'investitore sceglierà un portafoglio all'interno della frontiera. Un risultato importante è che tutti i portafogli della frontiera efficiente possono essere scritti come combinazione lineare di due soli portafogli efficienti. L'investitore dunque non è tenuto a conoscere l'intera composizione di portafogli efficienti, ma è sufficiente decidere l'allocazione tra due soli fondi<sup>53</sup>. Risulta evidente come, per determinare la combinazione desiderata, devono valere due assunti<sup>54</sup>:

- i. L'investitore sceglie il portafoglio minimizzando il rischio e massimizzando il rendimento;
- ii. È possibile ottenere delle stime ragionevoli per  $\mu_i$  e  $\sigma_{i,j}$

---

<sup>52</sup> Si considerano  $f, g: A \rightarrow \mathbb{R}$ , due funzioni di classe  $C^1$  con  $A \subseteq \mathbb{R}^n$ . Dato il problema  $\max/\min f(x)$  con vincolo  $g(x) = b$ , si definisce la funzione Lagrangiana come  $\mathcal{L}(x, \lambda) = f(x) - \lambda(g(x) - b)$ . Se la funzione  $f$  ha un massimo o minimo vincolato allora,  $\exists \lambda \in \mathbb{R} \text{ t. c. } \nabla \mathcal{L}(x, \lambda) = 0$ . Questa è una condizione necessaria del primo ordine. I punti critici della Lagrangiana possono essere anche punti di sella, quindi per definire in modo esatto gli estremi vincolati di  $f(x)$  si possono applicare le condizioni sufficienti di secondo ordine, usando il teorema dell'hessiana.  $\lambda$  è denominato moltiplicatore di Lagrange e se ne inseriscono tanti quanti sono i vincoli.

<sup>53</sup> N. AMENC, V. LE SOURD, *Portfolio theory and performance analysis*, 2003, cap. 3.

<sup>54</sup> H. MARKOWITZ, *Portfolio selection*, 1952.

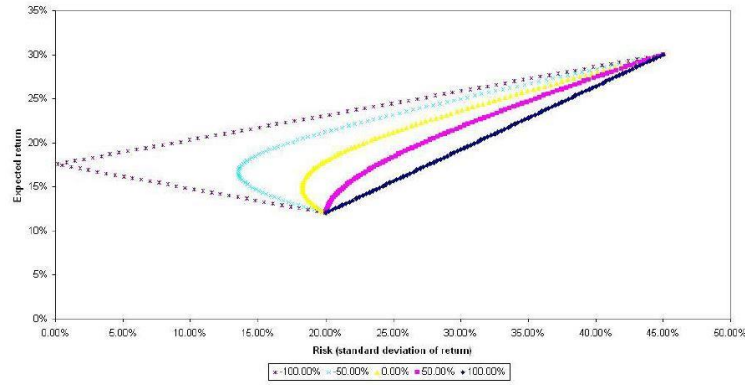


Fig. 7 Frontiera per due titoli al variare di  $\rho$ . [35]

### 2.3.1 Il free risk e il teorema di separazione di Tobin

Si può estendere la teoria di portafoglio ammettendo l'esistenza del portafoglio di rischio nullo, con  $\sigma = 0$ . La scelta per l'investitore è ora diversa, in quanto oltre ai portafogli sulla frontiera può acquistare anche quello free risk. Sia  $w$  la porzione di capitale investita in portafogli rischiosi, con rendimento  $\mu_p$  e deviazione  $\sigma_p$ , e  $1 - w$  la porzione di capitale investita nel free risk con ritorno  $r$ , allora il portafoglio risulterà avere:

$$\sigma[R] = w\sigma_p,$$

$$E[R] = (1 - w)r + w\mu_p = r + w(\mu_p - r) = r + \frac{\mu_p - r}{\sigma_p} \sigma[R].$$

Quest'ultima è l'equazione di una retta, con intercetta  $r$  e pendenza  $\frac{\mu_p - r}{\sigma_p}$  che, se calcolata nel punto di tangenza con la frontiera efficiente, chiamato Optimum Portfolio of Risky Assets (OPRA), rappresenta l'equazione della CML, ovvero tutte le combinazioni ottime tra OPRA e free risk che l'investitore sceglierà di acquistare, in base alla propria funzione di utilità<sup>55</sup>. Il portafoglio di tangenza  $(\mu_t, \sigma_t)$  è detto anche portafoglio di mercato e si può dimostrare che tutti gli investitori decideranno di comprare combinazioni lineari di soli due portafogli,

<sup>55</sup> G. WEST, *An introduction to modern portfolio theory: Markowitz, CAP-M, APT and Black-Litterman*, 2004, cap. 1.

il free risk e l'OPRA. L'introduzione del portafoglio risk free ha cambiato la natura della frontiera efficiente, in quanto, con queste condizioni, tutti gli operatori sceglieranno di allocare il capitale nel portafoglio non rischioso, e in una sola tipologia di portafoglio rischioso, quello di mercato. Questo risultato è noto come teorema di separazione di Tobin, in quanto si vengono a creare due distinte fasi di investimento. Nella prima, identica per tutti gli operatori, si definisce l'OPRA, nella seconda ogni investitore si colloca sulla CML, massimizzando la funzione di utilità individuale<sup>56</sup>. La pendenza della CML è detta indice di Sharpe e rappresenta la remunerazione per unità di rischio di ogni partecipante al mercato.

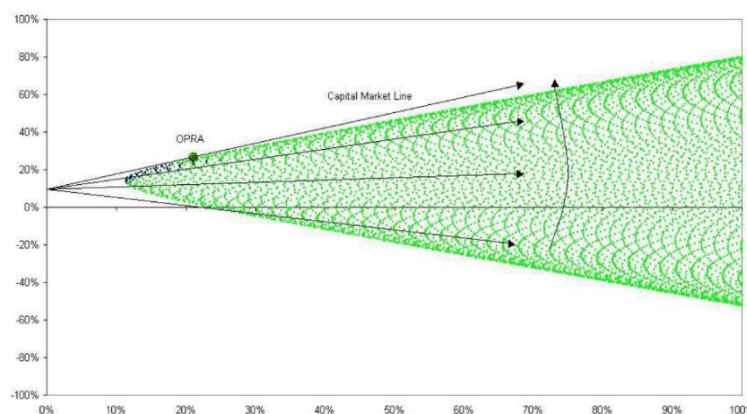


Fig. 8 Capital Market Line. [35]

## 2.4 Il Capital Asset Pricing Model

Partendo dalle considerazioni ottenute nella teoria di portafoglio in presenza di risk free, è possibile formulare un modello economico di equilibrio del prezzo dei titoli in condizioni di incertezza, noto come Capital Asset Pricing Model (CAPM), sviluppato dall'economista statunitense William Sharpe. L'obiettivo

<sup>56</sup> J. TOBIN, *Liquidity preference as behavior towards risk*, 1958.

del modello è quello di quantificare la relazione tra rendimento e rischio di ogni titolo. Dal teorema di separazione si ricava che la frontiera efficiente descritta da Markowitz si riduce, in realtà, alla CML, sulla quale è raffigurato un unico portafoglio rischioso, l'OPRA, che rappresenta l'unica scelta efficiente per gli investitori, insieme alle attività prive di rischio. Per ricavare il rendimento di un qualunque titolo rischioso è sufficiente quindi considerare solamente queste due componenti, attraverso la seguente relazione:

$$E(\tilde{R}_i) = R_f + \beta_i[E(\tilde{R}_m) - R_f],$$

con  $\tilde{R}_i$  rendimento atteso dell'asset  $i$ ,  $\tilde{R}_m$  rendimento atteso del portafoglio di mercato,  $R_f$  rendimento atteso privo di rischio e  $\beta_i$  parametro di sensibilità del rendimento rispetto a quello di mercato, definito matematicamente come:

$$\beta_i = \frac{cov(\tilde{R}_i, \tilde{R}_m)}{var(\tilde{R}_m)}.$$

La prima equazione rappresenta la formula chiave del CAPM e risulta essere applicabile sotto le seguenti condizioni<sup>57</sup>:

- i. Tutti gli investitori hanno le medesime opinioni sulle possibili variazioni del valore dei titoli alla fine del periodo;
- ii. La distribuzione di probabilità che descrive il rendimento atteso dei titoli disponibili ha una forma normale;
- iii. Gli investitori scelgono un portafoglio che massimizza la loro utilità alla fine del periodo e sono tutti avversi al rischio;
- iv. Un investitore può allocare una qualunque quantità di risorse all'interno dei titoli rischiosi e all'interno di quelli non rischiosi.

Se  $\beta_i = 0$ ,  $E(\tilde{R}_i) = R_f$ , il titolo ha rischio nullo e se  $\beta_i = 1$ ,  $E(\tilde{R}_i) = E(\tilde{R}_m)$ , il titolo ha la stessa rischio del mercato. La relazione che associa al

---

<sup>57</sup> F. BLACK, *Capital market equilibrium with restricted borrowing*, 1972.

rendimento di ciascun titolo il rispettivo  $\beta_i$  è detta Security Market Line (SML). Essa ha intercetta in  $R_f$  e inclinazione positiva, in quanto il premio per il rischio di mercato,  $E(\tilde{R}_m) - R_f$ , è sempre positivo. Tutti i titoli dovrebbero collocarsi sulla SML, che rappresenta quindi la linea del mercato azionario<sup>58</sup>. Il punto fondamentale per le applicazioni pratiche del CAPM riguarda la stima dei coefficienti  $\beta_i$ . Il metodo più comune consiste nello stimare una regressione lineare usando i dati storici. L'idea è che il rendimento atteso di un titolo è legato a quello del mercato (si usano come proxy per il portafoglio di mercato indici come S&P 500 o FTSE 100) e la quantificazione di questa relazione è rappresentata da  $\beta_i$ <sup>59</sup>. Il modello ha dunque la seguente espressione:

$$R_i = \alpha + \beta R_m + \varepsilon,$$

con,  $\alpha$  valore dell'intercetta e  $\varepsilon$  termine di errore. Se  $\beta > 1$  il titolo tende ad assumere rendimenti superiori a quelli del mercato, se invece  $\beta < 1$  il titolo tende ad avere rendimenti inferiori.

## 2.5 L'Arbitrage Pricing Theory come alternativa al CAPM

L'ipotesi che il rendimento atteso di un'attività possa essere determinato come somma tra il rendimento senza rischio e una percentuale del premio per il rischio di mercato, basata su un coefficiente  $\beta$ , può essere controversa, in quanto tutte le informazioni sulla rischiosità del titolo dovrebbero essere contenute in un unico parametro. Un'estensione del CAPM, sviluppata dall'economista americano Stephen Ross, che si propone di eliminare questa criticità, è nota

---

<sup>58</sup> G. TARDIVO, R. SCHIESARI, N. MIGLIETTA, *Finanza aziendale*, 2014, cap. 7.

<sup>59</sup> THE CORPORATE FINANCE INSTITUTE, *Investment banking manual*, 2018.



come Arbitrage Pricing Theory (APT). Le assunzioni su cui è basato il modello APT sono<sup>60</sup>:

- i. Gli investitori sono avversi al rischio e hanno come obiettivo quello di massimizzare la loro ricchezza finale;
- ii. Esiste un tasso di interesse senza rischio a cui prendere a prestito/prestare denaro;
- iii. Gli investitori concordano sul numero e sulla tipologia di fattori rilevanti nella determinazione dei prezzi;
- iv. Non esistono possibilità di arbitraggio, ovvero non è possibile ottenere profitti positivi senza rischio.

Sotto queste ipotesi, l'APT prevede che il rendimento atteso di un titolo possa essere espresso come una funzione lineare di una serie di fattori<sup>61</sup>:

$$R_i(t) = a_i + \sum_j b_{i,j} I_j(t) + \varepsilon_i(t),$$

in cui  $a_i$  rappresenta il tasso di rendimento privo di rischio,  $I_j$  sono i diversi fattori che concorrono alla determinazione del rendimento dell'attività, moltiplicati per un proprio fattore di sensibilità  $b_{i,j}$  e  $\varepsilon_i$  rappresenta il termine di errore, ovvero il rischio specifico relativo al titolo  $i$ -esimo, che può essere eliminato tramite diversificazione. In particolare, si ipotizza che  $\sigma_{\varepsilon_i}^2 < \infty \forall i$ ,  $E[\varepsilon_i \varepsilon_j] = 0$  e  $E[\varepsilon_j I_j] = 0$ . Il modello APT ideato da Ross non prevede esplicitamente il numero e il tipo di fattori che è necessario inserire per valutare correttamente il rendimento atteso. Sono state create, in seguito, molte versioni di APT considerando come  $I_j(t)$  delle variabili macroeconomiche, come PIL o inflazione, tuttavia la versione più famosa è quella elaborata da Fama e French,

---

<sup>60</sup> S. ROSS, *The arbitrage theory of capital asset pricing*, 1976.

<sup>61</sup> G. WEST, *An introduction to modern portfolio theory: Markowitz, CAP-M, APT and Black-Litterman*, 2004.

nota come modello a tre fattori<sup>62</sup>. In particolare, è possibile stimare il rendimento atteso di qualunque attività rischiosa come funzione di solamente tre fattori rilevanti:

- i. Il premio per il rischio di mercato  $R_m - R_f$ ;
- ii. La differenza di rendimento tra le imprese di piccole e grandi dimensioni, SMB (Small Minus Big);
- iii. La differenza di rendimento tra imprese con alto e basso rapporto tra valore contabile e valore di mercato, HML (High Minus Low).

Il modello a tre fattori si può esprimere come:

$$R_i - R_f = a_i + b_i[R_m - R_f] + s_iSMB + h_iHML.$$

Per  $a_i, s_i, h_i = 0$  si ottiene la formula dal CAPM per il rendimento di un titolo.

## 2.6 Il modello di Black e Scholes

La più diretta evoluzione del lavoro di Bachelier è certamente quello di Black e Scholes, due economisti americani che hanno ideato un modello per definire il valore di strumenti derivati, in particolare delle opzioni. La loro formula ha acquisito un enorme rilievo all'interno dei mercati finanziari, aumentando l'interesse verso questo particolare tipo di strumenti. In precedenza, si è definita un'opzione come un contratto che fornisce il diritto di acquistare o vendere uno specifico sottostante a un determinato prezzo in una data prestabilita, distinguendo le due principali tipologie di opzioni, di tipo call e di tipo put. Si denota con  $V$  il valore di un'opzione, che varia in funzione del prezzo attuale del sottostante  $S$  e del tempo  $t$ ,  $V = V(S, t)$ , distinguendo tra opzione call

---

<sup>62</sup> E. F. FAMA, K. R. FRENCH, *Common risk factors in the return on stocks and bonds*, 1993.

$C(S, t)$  e opzione put  $P(S, t)$ . Inoltre, si indica con  $K$  il prezzo di esercizio dell'opzione, detto strike, e con  $T$  la scadenza. Si considera come esempio un'opzione call. Alla scadenza, quando  $t = T$ , se  $S > K$  l'opzione verrà esercitata, poiché si ha la possibilità di comprare un'azione a un prezzo inferiore al valore corrente, generando un profitto positivo pari a  $S - K$ . Se invece alla scadenza si ha  $S < K$  l'opzione non sarà esercitata in quanto si potrebbe acquistare il sottostante sul mercato, pagando un prezzo inferiore. Il valore dell'opzione in questo caso è nullo. Un ragionamento opposto vale per un'opzione put. In modo compatto, si può scrivere il valore come:

$$C(S, T) = \max(S - K, 0),$$

$$P(S, T) = \max(K - S, 0).$$

Un'importante ipotesi, nel modello del mercato delle opzioni, è la condizione di assenza di arbitraggio, intesa come impossibilità di ottenere del profitto senza rischio istantaneamente, dalla quale si ricava una rilevante relazione, nota come put-call parity<sup>63</sup>. Si ipotizza di comprare un'azione, sulla quale viene acquistata una opzione put e viene venduta una opzione call, con la stessa scadenza  $T$  e lo stesso strike  $K$ . Il valore complessivo del portafoglio  $\pi$  è dato da:

$$\pi = S + P - C.$$

In modo alternativo si può scrivere, in termini di payoff:

$$\pi = S + \max(K - S, 0) - \max(S - K, 0).$$

Segue che:

$$\pi = \begin{cases} S + (K - S) - 0 = K & \text{se } S < K \\ S + 0 - (S - K) = K & \text{se } S > K \end{cases}$$

In qualunque caso, un portafoglio costruito in questo modo produce, alla scadenza, un guadagno pari a  $K$ . Ci si chiede ora quanto dovrebbe essere pagato

---

<sup>63</sup> P. WILMOTT, S. HOWISON, J. DEWYNNE, *The mathematics of financial derivatives, a student introduction*, 1995, cap.3.

un portafoglio di questo tipo. Per ottenere un valore di  $K$  al tempo  $T$  bisognerebbe depositare in una banca, al tempo  $t$  e al tasso di interesse continuo composto  $r$ , una somma  $K_0 = Ke^{-r(T-t)}$ <sup>64</sup>. Data l'assenza di arbitraggio deve valere la seguente relazione:

$$S + P - C = Ke^{-r(T-t)} \rightarrow C - P = S - Ke^{-r(T-t)}.$$

In altri termini, la differenza tra il valore di un'opzione call e un'opzione put deve essere uguale, in assenza di arbitraggio, alla differenza tra il prezzo del sottostante e il valore attuale dello strike.

### 2.6.1 Le ipotesi e la derivazione dell'equazione di Black-Scholes

L'obiettivo del modello di Black e Scholes è quello di trovare una espressione che consenta di calcolare il prezzo corretto di un'opzione in un tempo  $t < T$ . L'idea è quella di applicare un bilanciamento continuo per eliminare il rischio simile a quello precedentemente usato. L'equazione per il valore di un'opzione può essere derivata solamente se sono soddisfatte le seguenti ipotesi<sup>65</sup>:

- i. Il tasso di interesse  $r$  a breve termine è noto e costante;
- ii. Il prezzo del sottostante segue un moto browniano con una distribuzione di probabilità log-normale<sup>66</sup>(detto moto browniano geometrico);
- iii. Il sottostante non paga dividendi;

---

<sup>64</sup> Dato il tasso di interesse  $r$  si considera il periodo di tempo  $T - t$ . Se si immagina di dividere il periodo in tanti sottoperiodi  $\tau$ , al termine di ognuno sarà corrisposta una frazione dell'interesse  $r/\tau$ . Passando al limite, quando  $\tau \rightarrow \infty$ , il montante  $K$  è dato da:

$$K = \lim_{\tau \rightarrow \infty} K_0 \left(1 + \frac{r}{\tau}\right)^{\tau(T-t)} = K_0 \lim_{\tau \rightarrow \infty} \left(1 + \frac{r}{\tau}\right)^{\tau(T-t)} = K_0 e^{r(T-t)}.$$

Da questa relazione si ottiene facilmente il capitale iniziale necessario per ottenere  $K$  alla fine del periodo tramite attualizzazione:

$$K_0 = Ke^{-r(T-t)}.$$

<sup>65</sup> F. BLACK, M. SCHOLES, *The pricing of options and corporate liabilities*, 1973.

<sup>66</sup> Una variabile aleatoria  $X = e^K$  ha una distribuzione log-normale con media  $\mu$  e varianza  $\sigma^2$  se e solo se  $K = \log X \sim N(\mu, \sigma^2)$ .

- iv. L'opzione è di tipo europeo, cioè il diritto può essere esercitato solamente alla scadenza;
- v. Non ci sono costi di transazione;
- vi. È possibile scambiare qualunque frazione del prezzo di un titolo sul mercato al tasso di interesse di breve termine;
- vii. È ammessa la vendita allo scoperto.

Si suppone di avere un'opzione dal valore  $V(S, t)$  che dipende solamente dal prezzo del sottostante  $S$  e dal tempo  $t$ . Il prezzo del sottostante è un moto browniano geometrico che soddisfa l'equazione differenziale stocastica  $dS = \mu S dt + \sigma S dB_t$ . Usando il lemma di Ito si può scrivere:

$$dV(S, t) = \left( \frac{\partial V}{\partial t} + \mu S \frac{\partial V}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right) dt + \sigma S \frac{\partial V}{\partial S} dB_t.$$

Si considera ora il valore di un portafoglio  $\pi$  formato da un'opzione e da un numero  $-\Delta$  di asset sottostanti<sup>67</sup>:

$$\pi = V - \Delta S = V - \frac{\partial V}{\partial S} S$$

Studiando la variazione  $d\pi$  in un tempo infinitesimale e applicando nuovamente il lemma di Ito al valore del portafoglio si ottiene:

$$\begin{aligned} d\pi &= dV - \frac{\partial V}{\partial S} dS \\ &= \left( \frac{\partial V}{\partial t} + \mu S \frac{\partial V}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right) dt + \sigma S \frac{\partial V}{\partial S} dB_t - \frac{\partial V}{\partial S} \mu S dt \\ &\quad - \frac{\partial V}{\partial S} \sigma S dB_t = \left( \frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right) dt. \end{aligned}$$

---

<sup>67</sup> Imporre  $\Delta = \frac{\partial V}{\partial S}$  consente di eliminare la componente aleatoria nella variazione del portafoglio  $d\pi$ , rendendola completamente deterministica, in quanto i termini  $dW$  si semplificano.

Sfruttando la condizione di assenza di arbitraggio, deve valere che il rendimento di una somma  $\pi$ , depositata in una banca al tasso  $r$ , genera in un tempo  $dt$  un valore  $r\pi dt$ . In altri termini deve valere:

$$d\pi = r\pi dt = r \left( V - \frac{\partial V}{\partial S} S \right) dt.$$

Eguagliando le due espressioni ottenute per  $d\pi$  si ottiene l'equazione differenziale alle derivate parziali di Black-Scholes:

$$\left( \frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right) dt = r \left( V - \frac{\partial V}{\partial S} S \right) dt$$

$$\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} - rV = 0.$$

Questa relazione, in assenza di arbitraggio, deve essere soddisfatta dal prezzo di qualunque strumento derivato. Si può notare che l'espressione ottenuta per il delta

$$\Delta = \frac{\partial V}{\partial S},$$

rappresenta la variazione del valore dell'opzione rispetto al prezzo del sottostante, ovvero misura la sensibilità del valore rispetto al sottostante<sup>68</sup>. Il delta può essere utilizzato per costruire una strategia di 'delta-neutral'<sup>69</sup>, che consente di assumere una posizione neutrale al rischio per piccole variazioni del prezzo del titolo. L'equazione di Black-Scholes non dipende dal termine di drift  $\mu$  ma solamente dalla volatilità  $\sigma^2$ . Una misura diretta del parametro  $\sigma$  può essere complicata da identificare, perciò si assume solitamente che le

---

<sup>68</sup> P. WILMOTT, S. HOWISON, J. DEWYNNE, *The mathematics of financial derivatives, a student introduction*, 1995, cap.3.

<sup>69</sup> In una strategia di delta-neutral, il delta stabilisce la quantità di sottostante da acquistare o vendere per compensare le variazioni nel valore dell'opzione, per piccoli movimenti del prezzo del titolo. Si annulla il rischio aggiustando in ogni istante il delta. Si possono usare misure alternative più accurate, come gamma  $\Gamma$ :

$$\Gamma = \frac{\partial^2 V}{\partial S^2}.$$

informazioni relative alla volatilità siano contenute nel prezzo di mercato delle opzioni. Si assume quindi una volatilità implicita. Per risolvere esplicitamente l'equazione differenziale ottenuta è necessario imporre delle condizioni al contorno<sup>70</sup>. Prendendo come esempio un'opzione di tipo call, si assume che al tempo  $t = T$  sia noto il payoff, che sia nullo il valore dell'opzione quando il prezzo del titolo è uguale a zero e che quando il prezzo del sottostante tende a infinito anche il valore dell'opzione segue lo stesso andamento:

- i.  $t = T \rightarrow C(S, T) = \max(S - K, 0)$ ;
- ii.  $S = 0 \rightarrow C(0, t) = 0$ ;
- iii.  $S \rightarrow \infty \rightarrow C(S, t) \sim S$ .

Sfruttando queste condizioni e applicando la formula di Feynman-Kac<sup>71</sup>, si può ottenere la soluzione esplicita dell'equazione di Black-Scholes sia per un'opzione di tipo call che per una di tipo put:

$$C(S, t) = SN(d_1) - Ke^{-r(T-t)}N(d_2)$$

$$P(S, t) = Ke^{-r(T-t)}N(-d_2) - SN(-d_1),$$

ove:

$$d_1 = \frac{\ln\left(\frac{S}{K}\right) + \left(r + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}; \quad d_2 = d_1 - \sigma\sqrt{T-t}.$$

Questa formula ha ottenuto un notevole successo negli ambienti finanziari, diventando in poco tempo un fondamento della teoria classica.

---

<sup>70</sup> W. PAUL, J. BASCHNAGEL, *Stochastic processes, from physics to finance*, 2013, cap. 5.

<sup>71</sup> Questa equazione consente di trovare la soluzione di PDE paraboliche sfruttando le proprietà dei processi stocastici. Fissata una scadenza costante  $T > 0$ ,  $t \in [0, T]$ ,  $f \in C^2(\mathbb{R}^n)$   $v \in C^1(\mathbb{R}^n)$ , il rischio è descritto da  $X_t$  processo stocastico che soddisfa  $dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dB_t$ . Si definisce la funzione  $v(t, x)$  come valore atteso condizionale di un qualche payoff futuro attualizzato:

$$v(t, x) := E[f(X_T)e^{-r(T-t)}|X_t = x],$$

allora  $\forall x \in \mathbb{R}^n$ ,

- i. per  $t = T$   $v(T, x) = f(x)$
- ii. per  $0 < t < T$ ,  $v(t, x)$  è soluzione dell'equazione differenziale alle derivate parziali:

$$\frac{\partial v}{\partial t} + \mu(t, x) \frac{\partial v}{\partial x} + \frac{1}{2} \sigma^2(t, x) \frac{\partial^2 v}{\partial x^2} = 0.$$

## 2.7 I limiti della teoria finanziaria classica

Dopo aver illustrato i principali modelli che costituiscono la teoria finanziaria classica, è utile soffermarsi sulle debolezze che tali modelli presentano nel descrivere la realtà<sup>72</sup>. In primo luogo, qualunque modello, che si propone di descrivere un fenomeno reale, comporta necessariamente delle semplificazioni, in particolare, nei modelli economici decidere quali variabili trattare in modo endogeno e quali in modo esogeno è un processo cruciale. In seguito, oltre alla scelta delle variabili, un aspetto importante è rappresentato dalle ipotesi, le condizioni di esistenza sotto le quali il modello è verificato. Soffermandoci su queste ultime, una delle più comuni riguarda la razionalità degli investitori. Con razionalità ci si riferisce all'idea che ognuno massimizza la propria funzione di utilità con lo scopo di raggiungere il massimo grado di ricchezza possibile, inoltre, tutti investitori sono uguali e hanno aspettative omogenee. È noto come questa assunzione sia particolarmente irrealistica, nello specifico la finanza comportamentale si occupa di analizzare le scelte individuali irrazionali e i relativi bias<sup>73</sup>. Le variazioni dei prezzi, come visto in precedenza, si assumono

---

<sup>72</sup> B. B. MANDELBROT, *Il disordine dei mercati. Una visione frattale di rischio, rovina e redditività*, 2005, cap. 5.

<sup>73</sup> La finanza comportamentale è una branca dell'economia sviluppatasi a partire dagli anni 80', che studia, partendo da evidenze empiriche, le cause dell'irrazionalità delle scelte degli individui. Una delle teorie più note è la Prospect Theory, formulata da Kahneman e Tversky, secondo cui le scelte razionali sono sistematicamente violate da tre importanti effetti psicologici:

- i. L'effetto contesto, secondo il quale il modo in cui è formulato un problema ha un effetto diretto sulla scelta dell'individuo;
- ii. L'avversione alle perdite, secondo cui le persone forniscono maggior peso all'evitare una perdita rispetto a realizzare un guadagno;
- iii. L'effetto isolamento, secondo il quale gli individui, allo scopo di semplificare la scelta tra alternative, tendono a distinguerne le componenti, isolando probabilità consecutive invece di considerarle insieme. Ciò potrebbe creare preferenze incoerenti poiché, data la numerosità elevata delle componenti, alcune scelte probabili potrebbero essere ignorate in favore di altre.



indipendenti tra loro, stazionarie e distribuite lungo una curva normale. In realtà, le oscillazioni superiori agli intervalli predetti dalla disuguaglianza di Tchebycheff sono molto più comuni di quanto si può dedurre dalla teoria classica.

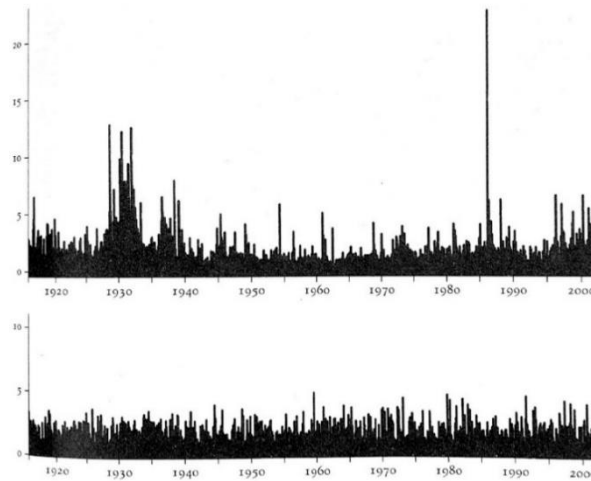


Fig. 9 Variazioni del Dow Jones e del moto browniano. [10]

La Fig. 9 mostra un confronto tra le variazioni del prezzo dell'indice Dow Jones, espresse in termini di deviazione standard, e una simulazione di un moto browniano<sup>74</sup>. Si nota che le oscillazioni del moto browniano, in basso, sono distribuite normalmente attorno alla media, con il 68% delle variazioni contenute in una deviazione standard  $\sigma$ , il 95% contenute in  $2\sigma$  e il 99% circa in  $3\sigma$ . L'andamento del Dow Jones mostra chiaramente la debolezza dell'ipotesi di normalità, in quanto è immediato notare come alcune oscillazioni siano ben

---

La teoria del prospetto distingue due diverse fasi nelle scelte individuali: la fase di editing, che consiste in un'analisi preliminare delle alternative, e la fase di valutazione dove viene selezionata l'alternativa che genera un valore maggiore. Queste considerazioni vengono riassunte nella funzione del valore  $v(x)$ , una legge matematica, nella quale la probabilità delle alternative viene ponderata per il peso individuale che ognuno attribuisce. È una funzione non lineare, concava nella zona dei guadagni ( $v''(x) < 0, x > 0$ ) e convessa nella zona delle perdite ( $v''(x) > 0, x < 0$ ), inoltre nella regione delle perdite presenta una pendenza maggiore.

D. KAHNEMAN, A. TVERSKY, *Prospect theory: an analysis of decision under risk*, 1979.

<sup>74</sup> B. B. MANDELBROT, *Il disordine dei mercati. Una visione frattale di rischio, rovina e redditività*, 2005, cap. 5.

superiori e rilevanti di quanto ipotizzato. Ad esempio, nel 1987 si è verificata una variazione pari a  $22\sigma$ , la cui probabilità è trascurabile per tabelle normali, in quanto apparentemente impossibile. Nei seguenti capitoli si espone una nuova impostazione, un nuovo modo di analizzare i mercati finanziari, attraverso quella che è nota come geometria frattale.

## *Appendice B. Il moto browniano geometrico*

Una delle ipotesi più importanti del modello di Black e Scholes è che il prezzo del sottostante di un'opzione deve seguire un moto browniano con distribuzione log-normale. Questo tipo di processo stocastico è definito come moto browniano geometrico e rappresenta una necessaria evoluzione del moto browniano classico, se applicato alla teoria finanziaria. Il moto browniano standard, ipotizzato da Bachelier, permette, infatti, ai prezzi dei titoli di assumere valori negativi, che sono per definizione irrealistici. Per sviluppare un modello coerente è quindi necessario permettere ai prezzi delle attività di assumere solamente valori positivi, considerando un processo stocastico il cui logaritmo descrive un moto browniano con distribuzione normale. Per studiare una variazione infinitesimale dei prezzi  $dS$  si considerano due tipi di contributi, uno deterministico, nel quale il termine di drift  $\mu$  rappresenta un tasso medio di crescita del prezzo, e uno casuale, nel quale si introduce il parametro di volatilità  $\sigma$ , in modo da ricavare l'equazione differenziale stocastica, che descrive un moto browniano geometrico, introdotta precedentemente, che rappresenta un caso particolare di un processo di Ito, in cui le funzioni  $a(S, t)$  e  $b(S, t)$  sono proporzionali a  $S$ :

$$dS = \mu S dt + \sigma S dB_t.$$

Considerando una funzione  $f(S, t)$  e applicando il lemma di Ito si ottiene:

$$df = \left( \frac{\partial f}{\partial t} + \mu S \frac{\partial f}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} \right) dt + \sigma S \frac{\partial f}{\partial S} dB_t.$$

Se si pone  $f(S, t) = \ln(S)$ , il lemma di Ito può essere scritto come:

$$df = \left( \mu S \frac{1}{S} - \frac{1}{2} \sigma^2 S^2 \frac{1}{S^2} \right) dt + \sigma S \frac{1}{S} dB_t = \left( \mu - \frac{1}{2} \sigma^2 \right) dt + \sigma dB_t.$$

Integrando entrambi i membri si ha:

$$\begin{aligned} \int_0^t df &= \left( \mu - \frac{1}{2} \sigma^2 \right) \int_0^t dt + \sigma \int_0^t dB_t \\ f_t - f_0 &= \left( \mu - \frac{1}{2} \sigma^2 \right) t + \sigma (B_t - B_0). \end{aligned}$$

Sostituendo, infine,  $f = \ln(S)$ , si può ricavare la soluzione analitica dell'equazione differenziale per un moto browniano geometrico<sup>75</sup>:

$$\begin{aligned} \ln(S_t) - \ln(S_0) &= \ln\left(\frac{S_t}{S_0}\right) = \ln e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma(B_t - B_0)} \\ S_t &= S_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma(B_t - B_0)}. \end{aligned}$$

La variabile casuale  $\ln(S_t)$  ha una distribuzione normale con media e varianza pari a:

$$\begin{aligned} E[\ln(S_t)] &= \left( \mu - \frac{1}{2} \sigma^2 \right) t \\ \text{Var}[\ln(S_t)] &= \sigma^2 t. \end{aligned}$$

*Appendice C. Le equazioni differenziali alle derivate parziali e l'equazione di diffusione*

---

<sup>75</sup> W. PAUL, J. BASCHNAGEL, *Stochastic processes, from physics to finance*, 2013, cap.5.

Un'equazione differenziale alle derivate parziali (PDE) è un'equazione composta dalle derivate parziali di una funzione non nota di due o più variabili indipendenti. Sono utilizzate in diversi ambiti, in quanto forniscono un modo per rappresentare l'evoluzione di sistemi dinamici nel tempo. In modo più formale, una generica PDE di ordine  $k$  ha la forma:

$$F(D^k u(x), D^{k-1} u(x), \dots, Du(x), u(x), x) = 0,$$

con  $k \in \mathbb{N}$ ,  $D^k$  operatore di derivazione di ordine  $k$ ,  $x \in U \subseteq \mathbb{R}^n$ ,  $F$  funzione nota,  $F: \mathbb{R}^{n^k} \times \mathbb{R}^{n^{k-1}} \times \dots \times \mathbb{R} \times U \rightarrow \mathbb{R}$ , e  $u$  funzione incognita,  $u: U \rightarrow \mathbb{R}$ . Una PDE si dice lineare se la funzione incognita e le sue derivate sono lineari, ovvero ha la forma:

$$\sum_{|\alpha| \leq k} a_\alpha(x) D^\alpha u = f(x),$$

con  $a_\alpha(x)$  e  $f(x)$  funzioni date. Se  $f = 0$  l'equazione si dice omogenea. Una PDE si dice invece non lineare se dipende in modo non lineare dal più alto grado di derivazione<sup>76</sup>. Risolvere una PDE vuol dire trovare un'espressione esplicita della funzione  $u$  che soddisfa l'equazione, tuttavia non esiste un modo univoco per farlo e non sempre è possibile. Spesso se la PDE dipende dal tempo si può identificare il relativo problema di Cauchy individuando delle condizioni al contorno, come ad esempio il valore della funzione in un particolare istante iniziale o finale. Tale problema si dice ben posto se esiste un'unica soluzione che dipende in modo continuo dai dati presenti. Inoltre, se la soluzione è una funzione liscia di classe  $C^k$ , per una PDE di ordine  $k$ , essa si denota come soluzione classica. La maggior parte delle volte non è possibile risolvere una PDE in senso classico ma, se si trova una soluzione unica non differenziabile  $k$  volte, essa si denota come soluzione debole. In alcuni casi non si riesce ad

---

<sup>76</sup> L.C. EVANS, *Partial differential equations*, 1998, cap. 1.

individuare una soluzione esplicita ma possono essere dedotte alcune sue proprietà e caratteristiche, infine, in altri casi, è possibile utilizzare metodi di analisi numerica. Prendendo in considerazione PDE del secondo ordine è possibile fare un'ulteriore classificazione. Data una PDE nella forma:

$$A \frac{\partial^2 f}{\partial x^2} + B \frac{\partial^2 f}{\partial x \partial y} + C \frac{\partial^2 f}{\partial y^2} + D \frac{\partial f}{\partial x} + E \frac{\partial f}{\partial y} + Ff = G,$$

si può analizzare il segno del discriminante  $B^2 - 4AC$ :

- i. Se  $B^2 - 4AC > 0$ , si dice PDE iperbolica, in quanto è possibile ottenere, per sostituzione, un'equazione nella forma  $\frac{\partial^2 f}{\partial \varepsilon^2} - \frac{\partial^2 f}{\partial \omega^2} + g = 0$ <sup>77</sup>, che è simile all'equazione di un'iperbole  $x^2 - y^2 = 1$ ;
- ii. Se  $B^2 - 4AC = 0$ , si dice PDE parabolica, in quanto si può operare una sostituzione che consente di ottenere un'equazione nella forma  $\frac{\partial^2 f}{\partial \varepsilon^2} + g = 0$ , che è simile a una parabola con equazione  $x^2 + y = 0$ ;
- iii. Se  $B^2 - 4AC < 0$ , si dice PDE ellittica, poiché si può ottenere una forma del tipo  $\frac{\partial^2 f}{\partial \varepsilon^2} + \frac{\partial^2 f}{\partial \omega^2} + g = 0$ , che assomiglia a quella di un'elisse  $x^2 + y^2 = 1$ .

L'esempio più noto di PDE parabolica di secondo ordine è l'equazione di diffusione del calore, che è stata ricavata nel capitolo precedente in modo qualitativo, dal lavoro di Bachelier, partendo dalla legge di radiazione di probabilità. Si mostrerà ora come sia possibile ricavare la stessa equazione imponendo alcune trasformazioni alla formula di Black e Scholes per il valore di un'opzione<sup>78</sup>:

$$\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} - rV = 0.$$

<sup>77</sup>  $g$  racchiude i termini di grado inferiore al secondo.

<sup>78</sup> G. MATTEI, F. LIBERATI, *Modelli di option pricing: l'equazione di Black e Scholes*, 2013.

Ponendo  $S = e^x$  e  $t = T - \frac{2\tau}{\sigma^2}$ , si ottiene una nuova funzione del valore

$V(S, t) = v(x, \tau) = v\left(\ln S, \frac{\sigma^2}{2}(T - t)\right)$ <sup>79</sup>. Dalle derivate parziali di  $V$  in funzione di  $S$  e  $t$ , si ricavano le derivate di  $v$  in funzione delle nuove coordinate:

$$\frac{\partial V}{\partial t} = \frac{\partial v}{\partial \tau} \frac{\partial \tau}{\partial t} = -\frac{\sigma^2}{2} \frac{\partial v}{\partial \tau},$$

$$\frac{\partial V}{\partial S} = \frac{\partial v}{\partial x} \frac{\partial x}{\partial S} = \frac{1}{S} \frac{\partial v}{\partial x},$$

$$\begin{aligned} \frac{\partial^2 V}{\partial S^2} &= \frac{\partial}{\partial S} \left( \frac{1}{S} \frac{\partial v}{\partial x} \right) = -\frac{1}{S^2} \frac{\partial v}{\partial x} + \frac{1}{S} \frac{\partial}{\partial S} \left( \frac{\partial v}{\partial x} \right) = -\frac{1}{S^2} \frac{\partial v}{\partial x} + \frac{1}{S} \frac{\partial}{\partial x} \left( \frac{\partial v}{\partial x} \right) \frac{\partial x}{\partial S} \\ &= -\frac{1}{S^2} \frac{\partial v}{\partial x} + \frac{1}{S^2} \frac{\partial^2 v}{\partial x^2}. \end{aligned}$$

Sostituendo queste ultime nella PDE si ricava:

$$\begin{aligned} -\frac{\sigma^2}{2} \frac{\partial v}{\partial \tau} + \frac{1}{2} \sigma^2 S^2 \left( -\frac{1}{S^2} \frac{\partial v}{\partial x} + \frac{1}{S^2} \frac{\partial^2 v}{\partial x^2} \right) + rS \frac{1}{S} \frac{\partial v}{\partial x} - rv &= 0 \\ -\frac{\sigma^2}{2} \frac{\partial v}{\partial \tau} + \left( r - \frac{1}{2} \sigma^2 \right) \frac{\partial v}{\partial x} + \frac{1}{2} \sigma^2 \frac{\partial^2 v}{\partial x^2} - rv &= 0 \\ \frac{\partial v}{\partial \tau} + \left( \frac{2r}{\sigma^2} - 1 \right) \frac{\partial v}{\partial x} + \frac{\partial^2 v}{\partial x^2} - \frac{2r}{\sigma^2} v &= 0. \end{aligned}$$

Per semplicità si pone  $k = \frac{2r}{\sigma^2}$  e  $t = \tau$ , in modo da poter riscrivere:

$$\frac{\partial v}{\partial t} + (k - 1) \frac{\partial v}{\partial x} + \frac{\partial^2 v}{\partial x^2} - kv = 0.$$

Per derivare l'equazione di diffusione è necessario imporre un'ulteriore trasformazione  $v(x, t) = e^{\alpha x + \beta t} u(x, t) = \phi u$ , da cui si ottengono le derivate parziali:

$$\frac{\partial v}{\partial t} = \beta e^{\alpha x + \beta t} u + e^{\alpha x + \beta t} \frac{\partial u}{\partial t} = \beta \phi u + \phi \frac{\partial u}{\partial t},$$

---

<sup>79</sup> In termini differenziali si può scrivere  $dx = \frac{1}{S} dS$  e  $d\tau = -\frac{\sigma^2}{2} dt$ .

$$\begin{aligned}\frac{\partial v}{\partial x} &= \alpha e^{\alpha x + \beta t} u + e^{\alpha x + \beta t} \frac{\partial u}{\partial t} = \alpha \phi u + \phi \frac{\partial u}{\partial t}, \\ \frac{\partial^2 v}{\partial x^2} &= \frac{\partial}{\partial x} \left( \alpha \phi u + \phi \frac{\partial u}{\partial t} \right) = \alpha^2 e^{\alpha x + \beta t} u + 2\alpha e^{\alpha x + \beta t} \frac{\partial u}{\partial x} + e^{\alpha x + \beta t} \frac{\partial^2 u}{\partial x^2} \\ &= \alpha^2 \phi u + 2\alpha \phi \frac{\partial u}{\partial x} + \phi \frac{\partial^2 u}{\partial x^2}.\end{aligned}$$

Sostituendo nella PDE e ponendo:

$$\begin{aligned}\alpha &= -\frac{1}{2}(k-1) = \frac{\sigma^2 - 2r}{2\sigma^2} \\ \beta &= -\frac{1}{4}(k-1)^2 = -\left(\frac{\sigma^2 + 2r}{2\sigma^2}\right)^2,\end{aligned}$$

si può ricavare, semplificando i termini, l'equazione del calore:

$$\begin{aligned}\left(\frac{\sigma^2 + 2r}{2\sigma^2}\right)^2 \phi u - \phi \frac{\partial u}{\partial t} + \left(\frac{2r - \sigma^2}{\sigma^2}\right) \left(\frac{\sigma^2 - 2r}{2\sigma^2} \phi u + \phi \frac{\partial u}{\partial x}\right) \\ + \left(\frac{\sigma^2 - 2r}{2\sigma^2}\right)^2 \phi u + \frac{\sigma^2 - 2r}{2\sigma^2} \phi \frac{\partial u}{\partial x} + \phi \frac{\partial^2 u}{\partial x^2} - \frac{2r}{\sigma^2} \phi u = 0, \\ \phi \frac{\partial u}{\partial t} - \phi \frac{\partial^2 u}{\partial x^2} = 0, \\ \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}.\end{aligned}$$

Questa equazione è stata formalizzata per la prima volta da Joseph Fourier per descrivere l'evoluzione nel tempo del flusso di calore su una superficie, ma può essere sfruttata per rappresentare numerosi fenomeni fisici, chimici e, come visto, anche economici. Essa descrive l'evoluzione della probabilità di identificare la posizione di una particella soggetta a un moto browniano. Può

essere generalizzata per una qualunque superficie a  $n$  dimensioni, tramite l'operatore di Laplace, scrivendo<sup>80</sup>:

$$\dot{u} = \alpha^2 \nabla^2 u,$$

ove  $\dot{u} = \frac{\partial u}{\partial t}$  e  $\nabla^2 = \left( \frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \dots, \frac{\partial}{\partial x_n} \right)$ . Si considera un'asta metallica uniforme perfettamente isolata lateralmente con una distribuzione di temperatura iniziale  $u(x, t)$ . Dal secondo principio della termodinamica, il calore passa da zone più calde a zone più fredde, quindi la temperatura della barra tenderà a uniformarsi, diventando costante in equilibrio, seguendo l'equazione di diffusione. È possibile trovare la soluzione fondamentale considerando il problema con la condizione iniziale  $u(x, 0) = f(x)$ , con  $(-\infty < x < \infty)$ , e la condizione di bordo  $u(x, t) \rightarrow 0$  per  $x \rightarrow \pm\infty$ . Sfruttando la trasformata di Fourier  $\mathcal{F}$ <sup>81</sup> si può

<sup>80</sup>  $\alpha^2 = \frac{K}{\sigma\rho}$ , rappresenta la diffusività termica del materiale che dipende dalla conducibilità termica  $K$ , dalla densità  $\rho$  e dalla capacità termica specifica  $\sigma$ . È una caratteristica intrinseca di ogni materiale, si misura in  $m^2/s$  nel SI e indica la velocità di propagazione nel materiale del calore.

<sup>81</sup> La serie di Fourier consente di rappresentare una funzione periodica tramite combinazione lineare di funzioni sinusoidali con frequenza crescente, in un intervallo del dominio  $[-L, L]$ . In particolare, l'idea di Fourier è stata quella di identificare delle basi trigonometriche ortonormali dello spazio di Hilbert, ovvero un insieme di vettori indipendenti ortogonali di norma unitaria. Una funzione  $f \in L^2$  a quadrato sommabile, quindi, può essere rappresentata come combinazione lineare delle basi. I coefficienti della combinazione lineare sono la proiezione della funzione rispetto alle basi  $\psi_k$  dello spazio (prodotto interno tra  $f$  e  $\psi_k$ ) e sono detti coefficienti di Fourier. Si ipotizza che all'esterno dell'intervallo in esame la funzione si ripeta con la stessa periodicità  $2L$  all'infinito.

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} [a_k \cos(\omega x) + b_k \sin(\omega x)] = \sum_{-\infty}^{\infty} c_k e^{i\omega x},$$

con  $\omega = k\pi/L$  e  $c_k = \frac{1}{2L} \langle f(x), \psi_k \rangle = \frac{1}{2L} \int_{-L}^L f(x) e^{-i\omega x} dx = \frac{1}{2L} \int_{-L}^L f(x) e^{-i\frac{k\pi}{L}x} dx$ .

È possibile generalizzare questa idea introducendo la trasformata di Fourier come limite della serie di Fourier sul dominio  $[-L, L]$  per  $L \rightarrow \infty$  (oppure per  $\Delta\omega = \pi/L \rightarrow 0$ ), permettendo di definire una funzione in  $(-\infty, \infty)$  ed escludendo la ripetizione.

$$f(x) = \lim_{\Delta\omega \rightarrow 0} \sum_{-\infty}^{\infty} \frac{\Delta\omega}{2\pi} \int_{-\pi/\Delta\omega}^{\pi/\Delta\omega} f(\xi) e^{-ik\Delta\omega\xi} d\xi e^{ik\Delta\omega x} = \int_{-\infty}^{\infty} \frac{1}{2\pi} \int_{-\infty}^{\infty} f(\xi) e^{-i\omega\xi} d\xi e^{i\omega x} d\omega.$$

Riconoscendo il coefficiente di Fourier  $\hat{f}(\omega) = \int_{-\infty}^{\infty} f(\xi) e^{-i\omega\xi} d\xi$  è possibile formulare la trasformata di Fourier  $\mathcal{F}$  e la trasformata inversa di Fourier  $\mathcal{F}^{-1}$  della funzione  $f(x)$  come:

$$\begin{aligned} \hat{f}(\omega) &= \mathcal{F}[f(x)] = \int_{-\infty}^{\infty} f(x) e^{-i\omega x} dx \\ f(x) &= \mathcal{F}^{-1}[\hat{f}(\omega)] = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\omega) e^{i\omega x} d\omega. \end{aligned}$$



trasformare la PDE in una equazione differenziale ordinaria (ODE), operando una trasformazione di coordinate spaziali  $u(x, t) \rightarrow \hat{u}(\omega, t)$ , facilmente risolvibile, per poi ritornare alle originali coordinate usando il teorema di inversione di Fourier<sup>82</sup>. Scrivendo l'equazione del calore come  $u_t = \alpha^2 u_{xx}$  si può calcolare  $\hat{u}(\omega, t) = \mathcal{F}(u)$ :

$$\begin{aligned}\mathcal{F}(u_x) &= i\omega \mathcal{F}(u) = i\omega \hat{u} \\ \mathcal{F}(u_{xx}) &= (i\omega)^2 \mathcal{F}(u) = -\omega^2 \hat{u},\end{aligned}$$

da cui si ha,

$$\frac{\partial \hat{u}}{\partial t} = -\alpha^2 \omega^2 \hat{u}.$$

Quest'ultima è una ODE lineare del primo ordine con soluzione:

$$\hat{u}(\omega, t) = \hat{f}(\omega) e^{-\alpha^2 \omega^2 t},$$

con condizione iniziale  $\hat{f}(\omega) = \mathcal{F}[f(x)] = \int_{-\infty}^{\infty} f(x) e^{-i\omega x} dx$ . Si usa ora la trasformata inversa  $\mathcal{F}^{-1}$  di Fourier per esprimere la soluzione dell'equazione del calore in  $(x, t)$ :

---

Un'importante proprietà usata nella risoluzione delle PDE riguarda la trasformata di Fourier della derivata di una funzione:

$$\begin{aligned}\mathcal{F}\left(\frac{\partial f}{\partial x}\right) &= \int_{-\infty}^{\infty} f'(x) e^{-i\omega x} dx = [f(x) e^{-i\omega x}]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} f(x) (-i\omega e^{-i\omega x}) dx = i\omega \int_{-\infty}^{\infty} f(x) e^{-i\omega x} dx, \\ \mathcal{F}\left[\frac{\partial}{\partial x} f(x)\right] &= i\omega \mathcal{F}[f(x)].\end{aligned}$$

S.B. BRUNTON, J.N. KUTZ, *Data-driven science and engineering, machine learning, dynamical systems, and control*, 2019, cap. 2.

<sup>82</sup> E. KREYSZIG, *Advanced engineering mathematics*, 2006, cap. 12.

$$\begin{aligned}
u(x, t) &= \mathcal{F}^{-1}[\hat{u}(\omega, t)] = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{f}(\omega) e^{-\alpha^2 \omega^2 t} e^{i\omega x} d\omega \\
&= \frac{1}{2\pi} \iint_{-\infty}^{+\infty} f(y) e^{-i\omega y} e^{-\alpha^2 \omega^2 t} e^{i\omega x} d\omega dy \\
&= \frac{1}{2\pi} \iint_{-\infty}^{+\infty} e^{i\omega(x-y) - \alpha^2 \omega^2 t} f(y) dy d\omega.
\end{aligned}$$

Ponendo  $\xi = x - y$  si può scrivere:

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega \xi} e^{-\alpha^2 \omega^2 t} d\omega = \mathcal{F}^{-1}(e^{-\alpha^2 \omega^2 t}) = \frac{1}{2\alpha\sqrt{\pi t}} e^{-\frac{\xi^2}{4\alpha^2 t}}.$$

Sostituendo si ottiene l'espressione generale della soluzione come:

$$u(x, t) = \frac{1}{2\alpha\sqrt{\pi t}} \int_{-\infty}^{\infty} e^{-\frac{(x-y)^2}{4\alpha^2 t}} f(y) dy.$$

Ponendo, infine,  $g(\xi, t) = \frac{1}{2\alpha\sqrt{\pi t}} e^{-\frac{\xi^2}{4\alpha^2 t}}$  la soluzione può essere scritta come convoluzione di  $g$  e  $f$ :

$$u(x, t) = \int_{-\infty}^{\infty} g(x - y, t) f(y) dy = (g * f)(x, t).$$

Si può notare come  $g(x, t)$  sia una gaussiana la cui deviazione aumenta nel tempo in modo lineare, appiattendosi fino alla temperatura di equilibrio.

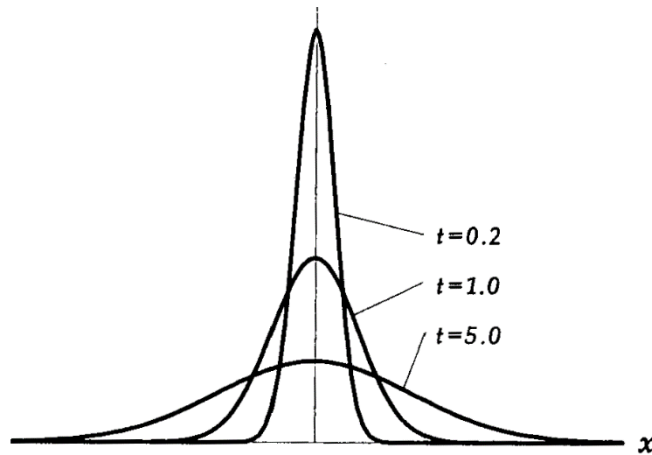


Fig. 10 Soluzione fondamentale dell'equazione di diffusione. [63]

## **CAPITOLO III.**

### **GLI OGGETTI FRATTALI E LA DIMENSIONE FRAZIONARIA**

### *3.1 Introduzione e definizione del concetto di frattale*

La teoria economica che sarà descritta nei successivi paragrafi poggia le basi sulla nozione di irregolarità, contrapposta a quella di normalità classica, che è il punto di partenza per descrivere una nuova classe di oggetti geometrici per molto tempo ignorata, gli oggetti frattali. Il termine è stato ideato dal matematico polacco Benoit Mandelbrot nel 1975 e deriva dal latino ‘fractus’, participio passato di ‘frangere’, che significa spezzato, interrotto, irregolare. Per approfondire il significato di questa espressione è necessario andare oltre la semplice irregolarità. Una delle principali proprietà dei frattali infatti è l’omotetia interna, detta anche invarianza di scala<sup>83</sup> poiché ingrandendo una parte della figura è possibile ritrovare, invariata, la stessa immagine generale. Una copia del frattale può essere identificata ingrandendo una qualsiasi parte della figura, in modo che il particolare sia uguale al generale. In altre parole, la forma del frattale rimane inalterata se si modifica la scala con la quale la si osserva. Questa idea, apparentemente caotica e disordinata, fornisce una rappresentazione innovativa e accurata della complessità, che è stata sviluppata solo recentemente. La complessità di un frattale è formalizzata all’interno della definizione di dimensione frazionaria, la cui più nota declinazione è la dimensione di Hausdorff. Questi oggetti, infatti, possono assumere una dimensione non intera, che può essere rappresentata anche da un numero irrazionale, come ad esempio  $\frac{\log 4}{\log 3} = 1,261859$ , per una curva frattale che si

---

<sup>83</sup> Una funzione  $f(x)$  ha, in generale, la proprietà di invarianza di scala se, applicando una dilatazione con fattore di scala  $\lambda$  alla variabile  $x$  vale, per un qualche esponente  $\Delta$ , che:

$$f(\lambda x) = \lambda^\Delta f(x).$$

Sotto questa condizione,  $f(x)$  risulta invariante per qualunque valore di  $\lambda$ . I frattali sono solitamente invarianti per un insieme discreto di valori di  $\lambda$ , quindi si parla formalmente di auto similarità.

esaminerà in seguito. Questa interpretazione di dimensione è una nozione fondamentale, estranea a qualunque concezione di geometria euclidea<sup>84</sup>. In altri termini, definire una linea come oggetto a 1 dimensione, un piano come oggetto a 2 dimensioni e una sfera come oggetto a 3 dimensioni non è più sufficiente, all'interno della geometria frattale. La dimensione di un oggetto non è più definita a priori ma è in funzione del grado di irregolarità della figura geometrica, in particolare una curva molto irregolare potrà avere dimensione compresa tra 1 e 2, una superficie molto irregolare potrà averne una compresa tra 2 e 3 e dei segmenti infinitesimali di retta potranno avere una dimensione compresa tra 0 e 1<sup>85</sup>. Da queste proprietà si potrebbe immaginare un frattale come un oggetto esclusivamente astratto, in realtà esistono numerosi esempi di frattali naturali, come gli alberi, le felci, i fulmini, il cavolo romano, le nuvole, il polmone umano o le linee costiere. Questi esempi possiedono tutti una struttura frattale e sono impossibili da descrivere all'interno della geometria euclidea. Considerando ad esempio una felce, ogni singola foglia è esattamente identica alla felce più grande.

---

<sup>84</sup> Con geometria euclidea si intende il sistema matematico descritto per la prima volta in modo completo da Euclide attorno al 300 a.C. negli 'Elementi'. Quest'ultimo è composto da 13 libri, riguardanti la geometria piana, i rapporti tra grandezze e la geometria solida, che fissano le basi dell'impostazione classica della geometria, studiata ancora oggi. Nel primo libro, in particolare, egli definisce cinque assiomi, senza dimostrazione, necessari per lo sviluppo logico dell'opera:

- i. Tra due punti si può tracciare una e una sola retta;
- ii. È possibile prolungare un segmento oltre i due punti indefinitamente;
- iii. Fissato un punto e una lunghezza si può definire un cerchio;
- iv. Gli angoli retti sono tra loro congruenti;
- v. Se una retta che taglia altre due rette determina dalla stessa parte due angoli interni minori di due angoli retti, prolungando le due rette all'infinito, esse si incontreranno dalla parte di questi due angoli.

L'ultimo postulato è noto anche come assioma delle rette parallele, poiché afferma indirettamente che due rette parallele non si incontrano mai, ed è il più delicato, in quanto non empiricamente osservabile. Dalla negazione dell'ultimo assioma sono stati creati diversi tipi di geometrie non euclidee, come ad esempio la geometria iperbolica e quella ellittica.

R. FITZPATRICK, *Euclid's elements of geometry*, 2008.

<sup>85</sup> B. B. MANDELBROT, *Gli oggetti frattali. Forma, caso e dimensione*, 2000, cap. 1.



Fig. 11 Felce frattale. [38]

Al momento non esiste una definizione unanime di frattale. Mandelbrot ha proposto di definire con questo termine tutti gli oggetti con una dimensione di Hausdorff non intera maggiore di quella topologica<sup>86</sup>, tuttavia si è in seguito mostrato come possono esistere curve frattali con dimensione intera, ad esempio si può dimostrare che il moto browniano in  $\mathbb{R}^n$ , con  $n \geq 2$ , ha un carattere frattale e dimensione di Hausdorff pari a 2, con probabilità unitaria<sup>87</sup>. Si possono classificare i frattali in base alle proprietà possedute dal rapporto di simmetria<sup>88</sup> dividendoli in:

---

<sup>86</sup> La dimensione topologica, detta anche dimensione di ricoprimento di Lebesgue, è uno dei diversi modi di definire una dimensione, in senso matematico, in modo che sia invariante per omeomorfismo (due spazi topologici  $X$  e  $Y$  sono omeomorfi se esiste un omeomorfismo tra loro, ovvero se esiste una funzione  $f: X \rightarrow Y$  continua e biunivoca la cui inversa  $f^{-1}: Y \rightarrow X$  è continua). Dato uno spazio topologico  $X$  un ricoprimento aperto  $\{U_i\}$  di  $X$  è una famiglia di insiemi aperti tale che  $X \subseteq \bigcup_{i=1}^n U_i$ . Un suo raffinamento è un ulteriore ricoprimento aperto  $\{V_j\}$  tale che  $U_i \subseteq \bigcup_{j=1}^n V_j$ . Si dice dimensione topologica di  $X$  il più piccolo intero  $n$  tale che ogni ricoprimento aperto  $U_i$  di  $X$  ha un raffinamento in cui ogni punto è contenuto al massimo in  $n + 1$  insiemi. Se  $n$  non esiste  $\dim X = \infty$ .  $X$  ha dimensione nulla se  $\{U_i\}$  ha un raffinamento formato da insiemi aperti disgiunti.

A. R. PEARS, *Dimension theory of general spaces*, 1975, cap.3.

<sup>87</sup> K.J FALCONER, *The geometry of fractal sets*, 1986, cap. 8.

<sup>88</sup> B. B. MANDEBROT, *Il disordine dei mercati. Una visione frattale di rischio, rovina e redditività*, 2005, cap. 7.

- i. Frattali auto-simili, se la riduzione di scala avviene con lo stesso rapporto in tutte le direzioni;
- ii. Frattali auto-affini, se la riduzione si ha maggiormente in una direzione rispetto ad un'altra;
- iii. Multifrattali, se si hanno diversi fattori di riduzione in molteplici punti distinti.

Non è totalmente corretto assegnare l'esclusiva paternità dei frattali a Mandelbrot, in quanto, all'interno della matematica, alcune di queste figure erano già state scoperte anni prima, tuttavia esse rappresentavano delle eccezioni, delle curiosità matematiche che erano più simili a paradossi che ad oggetti geometrici degni di attenzione. Erano già note, ad esempio, delle funzioni continue sul dominio ma che non ammettono una tangente in nessun punto<sup>89</sup>, così come erano note delle curve in grado di passare per tutti i punti di una superficie ricoprendola<sup>90</sup>. Uno dei meriti di Mandelbrot è stato quello di raccogliere tutti questi paradossi all'interno di una categoria ben precisa, assegnandogli un nome. Se si pensa ad una qualunque curva piana, essa può essere descritta da un'equazione matematica. Per i frattali ciò non è possibile, in quanto sono costruiti tramite un algoritmo, ovvero un metodo utilizzato per disegnare la curva. Per quelli più semplici, si parte da un iniziatore, una figura geometrica classica. In seguito, si costruisce un generatore, di solito una semplice forma geometrica, come una linea spezzata, e si applica una regola di ricorsione, sostituendo il generatore nelle varie parti dell'iniziatore,

---

<sup>89</sup> La funzione di Weierstrass, scoperta nel 1872, continua in ogni punto ma non derivabile per nessun valore di  $x$ , rappresenta il primo storico controesempio all'idea secondo cui ogni funzione continua è derivabile, di conseguenza. Essa ha un carattere di auto similarità e si può considerare un primo esempio di curva frattale.

<sup>90</sup> Il primo esempio di curva che ricopre uno spazio è la curva di Peano, scoperta nel 1890. Essa è descritta come limite di una successione di curve all'interno di un quadrato. Si può dimostrare che è una funzione continua non iniettiva e non derivabile che passa per tutti i punti del quadrato.

comprimendolo<sup>91</sup>. In base al tipo di algoritmo utilizzato per creare i frattali, essi si possono dividere in:

- i. Frattali per sostituzione, ottenuti sostituendo progressivamente il generatore al posto dell'iniziatore;
- ii. Frattali IFS (sistemi di funzioni iterate), ottenuti partendo da una figura generatrice a cui si applicano  $n$  trasformazioni affini (ad esempio rotazioni, traslazioni o omotetie) a ogni punto, che creano  $n$  punti trasformati, ai quali si applicano nuovamente le stesse trasformazioni affini, ottenendo  $n_2$  punti trasformati, fino a quando, all'ultima iterazione, si ottiene l'attrattore del sistema, ovvero un frattale auto simile tipico per ogni IFS, che si origina indipendentemente dai punti iniziali scelti;
- iii. L-Frattali (sistemi di Lindenmayer), creati a partire da una lista ordinata di elementi, detta stringa iniziale o assioma, alla quale si applicano in parallelo delle regole di riscrittura prestabilite, all'infinito. È il metodo utilizzato ad esempio per studiare i processi di crescita delle piante.

Per i frattali più complessi, si può avere un'equazione generatrice iterata al posto del generatore geometrico, in modo da poter essere classificati, in base al grado dell'equazione, in:

- i. Frattali lineari, se l'equazione generatrice contiene solo termini del primo ordine;
- ii. Frattali non lineari, se l'equazione generatrice è di ordine superiore al primo;

---

<sup>91</sup> B. B. MANDEBROT, *Il disordine dei mercati. Una visione frattale di rischio, rovina e redditività*, 2005, cap. 7.



- iii. Frattali aleatori, se l'equazione generatrice segue una legge di distribuzione particolare. Si tratta in questo caso di una figura non deterministica.

### 3.1.1 L'insieme di Mandelbrot

Il frattale più famoso, scoperto da Mandelbrot nel 1979, è senza dubbio l'insieme di Mandelbrot, che è il risultato di una rielaborazione degli studi del matematico francese Gaston Julia, sui processi di iterazione delle funzioni razionali. L'insieme di Mandelbrot è costruito sul piano complesso ed è descritto come:

$$M = \begin{cases} z_{n+1} = z_n^2 + c \\ z_0 = 0 \end{cases} \quad z, c \in \mathbb{C}.$$

La successione definita parte dall'origine e si sviluppa come:

$$n = 0 \rightarrow z_1 = c$$

$$n = 1 \rightarrow z_2 = c^2 + c$$

$$n = 2 \rightarrow z_3 = (c^2 + c)^2 + c$$

$$n = 3 \rightarrow [(c^2 + c)^2 + c]^2 + c, \dots$$

L'insieme di Mandelbrot è costituito da tutti quei punti  $c \in \mathbb{C}$  tali che la successione si mantiene limitata, ovvero non diverge a infinito<sup>92</sup>. Si può dimostrare come per  $|z_n| \geq 2$  la successione diverge, quindi l'insieme è interamente contenuto in un cerchio di raggio 2. In generale, si colorano di solito in nero i punti che appartengono all'insieme e con sfumature diverse i punti che non vi appartengono, in base alla velocità con cui la successione diverge. È importante citare anche gli insiemi di Julia  $J_c$ , che sono definiti dalla successione  $f_c(z) = z^2 + c$ , fissato il punto  $c \in \mathbb{C}$ . Alcuni punti di questa

---

<sup>92</sup> Data la funzione  $f_c: \mathbb{C} \rightarrow \mathbb{C} = z^2 + c$ , denotando con  $f_c^n(z)$  la  $n$ -esima iterazione di  $f_c(z)$  si può esprimere formalmente l'insieme di Mandelbrot come sottoinsieme di  $\mathbb{C}$ :

$$M = \{c \in \mathbb{C} : \sup |f_c^n(0)| < \infty, n \in \mathbb{N}\}.$$

successione convergono in una regione che ha una forma frattale. Esiste una importante relazione tra l'insieme di Mandelbrot e quelli di Julia<sup>93</sup>, in quanto se  $c \in M$  allora il corrispondente  $J_c$  è connesso, se  $c \notin M$  allora  $J_c$  non è connesso, infine se  $c$  appartiene alla frontiera di  $M$  il corrispondente insieme di Julia è una regione connessa senza area<sup>94</sup>. In generale, gli insiemi di Julia e di Mandelbrot non rientrano nella definizione classica di frattale poiché, aumentando la scala si ottengono figure più complicate sempre diverse che richiamano la parte generale.

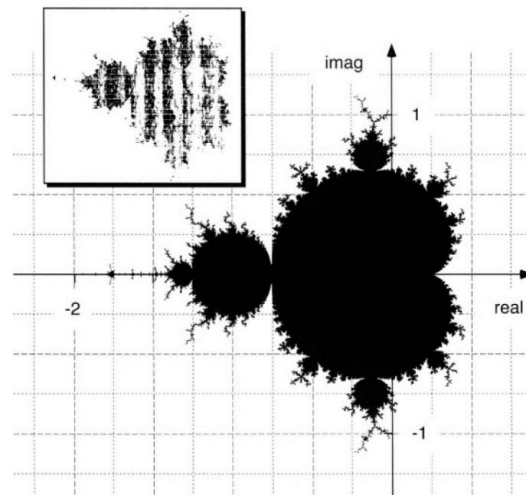


Fig. 12 Insieme di Mandelbrot. [38]

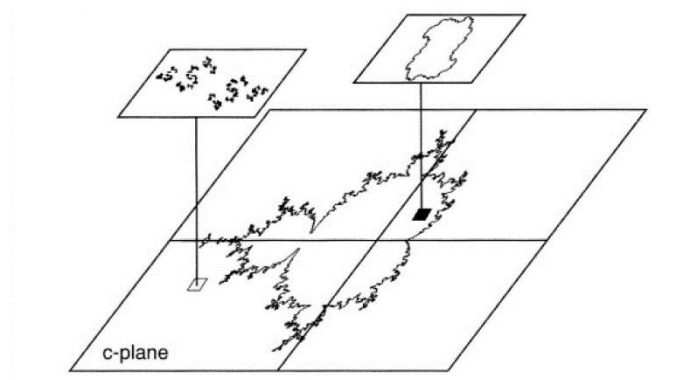


Fig. 13 Relazione tra insieme di Mandelbrot e insiemi di Julia. [38]

<sup>93</sup> Ogni punto del piano complesso denota un particolare insieme di Julia, quindi si può definire l'insieme di Mandelbrot come:

$$M = \{c \in \mathbb{C} \mid J_c \text{ è connesso}\}.$$

<sup>94</sup> H. PEITGEN, H. JÜRGENS, D. SAUPE, *Chaos and fractals. New frontiers of science*, 2004, cap. 14.

### *3.2 La dimensione frattale e la misura della linea costiera*

Una delle principali caratteristiche dei frattali è il concetto di dimensione frattale, che può assumere valori non interi. L'idea che una curva irregolare possa avere una dimensione a metà fra una retta e un piano, può generare notevole confusione ed è necessario soffermarsi su essa. La dimensione euclidea può essere considerata come una caratteristica intrinseca degli oggetti, con la quale si possiede una notevole familiarità. Risulta facile assegnare a una retta un'unica dimensione, a un piano due dimensioni e allo spazio che ci circonda tre dimensioni, tuttavia non si considera solitamente l'impatto che la scala di osservazione può avere sulla dimensione di un oggetto. Prendendo come esempio un gomito di spago idealizzato, se lo si osserva da molto lontano, dal punto di vista euclideo, esso è visto come un punto adimensionale. Se si prende il gomito in mano risulta essere un oggetto tridimensionale, se si osserva da più vicino si notano solo delle fibre monodimensionali, infine se si aumenta nuovamente la scala di osservazione si vede che le fibre sono, in realtà, filamenti a tre dimensioni<sup>95</sup>. Si può quindi considerare la dimensione come uno strumento di misura, invece che come una proprietà caratterizzante. In particolare, facendo riferimento alla dimensione frattale, essa può essere vista come uno strumento di misura del grado di irregolarità. La dimensione frattale può avere diverse formulazioni matematiche, tuttavia per capire come un oggetto possa avere una dimensione non intera è ragionevole partire dal caso più semplice, la

---

<sup>95</sup> B. B. MANDEBROT, *Il disordine dei mercati. Una visione frattale di rischio, rovina e redditività*, 2005, cap. 7.

dimensione di omotetia<sup>96</sup> o di auto-similarità. Questa dimensione si può calcolare per gli oggetti geometrici più semplici in particolare per i frattali auto simili, ovvero quelli simmetrici in cui il rapporto di dilatazione o di riduzione è costante in tutte le direzioni. La dimensione di omotetia è espressa in termini di rapporto di omotetia,  $r$ , e numero di parti simili in cui l'oggetto è diviso,  $N$ . Se si considera una retta e la si divide a metà in due parti uguali, si ha che  $r = \frac{1}{2}$  e  $N = 2$ , quindi vale la relazione:

$$\left(\frac{1}{r}\right)^1 = N.$$

L'esponente rappresenta la dimensione di omotetia, in questo caso uguale a 1. Per un quadrato vale una rappresentazione analoga, dividendo il lato a metà si ottiene  $r = \frac{1}{2}$  e  $N = 4$ , poiché si generano quattro quadrati uguali, e si può notare che:

$$\left(\frac{1}{r}\right)^2 = N.$$

Un quadrato ha quindi dimensione di similitudine pari a 2. È possibile ottenere un'espressione più generale, per ogni dimensione  $D$ :

$$\left(\frac{1}{r}\right)^D = N,$$

da cui si ricava:

$$D = \frac{\log(N)}{\log\left(\frac{1}{r}\right)}.$$

---

<sup>96</sup> Con il termine omotetia ci si riferisce ad una particolare trasformazione dello spazio euclideo di centro  $C$  che modifica la distanza di tutti i punti da  $C$  di un fattore  $k$  costante. In altre parole, ad ogni punto  $P$  fa corrispondere un punto  $P'$  in modo che  $\overline{CP'} = k\overline{CP}$ .  $C$  è anche detto punto unito della trasformazione e se  $|k| > 1$  la figura è dilatata, se  $|k| < 1$  la figura viene ridotta, se  $k = 1$  si ha un'identità e, infine, se  $k = -1$  si ottiene una simmetria rispetto al centro. Con riferimento ai frattali l'omotetia così definita perde buona parte del significato che ha per oggetti regolari, tuttavia consente di rendere intuitivo il concetto di auto similarità.

Risulta evidente che, scegliendo opportuni  $r$  e  $N$ , si può ottenere che  $D \notin \mathbb{N}$ . La formula ottenuta non è una semplice curiosità matematica, bensì ha le sue radici in un problema reale, riguardante la misurazione della lunghezza di una linea costiera. Quello che a prima vista sembra un problema semplice nasconde in sé un interessante paradosso, indagato per la prima volta nel 1961 dal ricercatore inglese Lewis Fry Richardson, il quale ha notato, esaminando i rapporti ufficiali, l'esistenza di difformità nelle stime delle lunghezze dei confini tra gli Stati<sup>97</sup>. La natura frastagliata di una costa rende complicata una misura diretta e univoca della sua lunghezza. Questa misura può essere effettuata con vari metodi, ad esempio si può immaginare di usare un compasso di apertura  $\varepsilon$  che si muove lungo la costa, in modo che ogni nuovo passo inizi dalla fine del precedente. La lunghezza approssimata  $L(\varepsilon)$  sarebbe in questo caso il prodotto tra  $\varepsilon$  e il numero di passi<sup>98</sup>. Diminuendo progressivamente  $\varepsilon$  e ripetendo il procedimento si può immaginare che il risultato converga al valore esatto della linea costiera. Il paradosso dimostrato da Richardson, noto come effetto Richardson, prevede che in realtà, diminuendo  $\varepsilon$ ,  $L(\varepsilon)$  tende ad aumentare all'infinito senza convergere. In altre parole, ad ogni nuova misurazione la lunghezza della costa risulta essere maggiore, in quanto si includono sempre più dettagli nell'approssimazione, ottenendo idealmente una curva di lunghezza infinita. È ragionevole comunque sottolineare che il concetto di omotetia, in riferimento a frattali naturali, ha senso solo all'interno di determinati limiti, infatti con scale estremamente piccole la costa cessa di essere una linea vera e propria e risulta predominante la natura atomica degli elementi che la compongono. Affinché la dimensione abbia quindi un senso fisico si opera un taglio alle piccole scale,

---

<sup>97</sup> B. B. MANDEBROT, *Il disordine dei mercati. Una visione frattale di rischio, rovina e redditività*, 2005, cap. 7.

<sup>98</sup> B. B. MANDELBROT, *The fractal geometry of the nature*, 1982, cap. 5.

ponendo un limite interno alla scala di osservazione. Con un ragionamento analogo, è possibile imporre anche un taglio alle grandi scale, ovvero un limite esterno oltre il quale non ha più senso calcolare la lunghezza della costa. Un esempio di limite esterno, per tratti di costa chiusa, può essere rappresentato dalla distanza tra le due estremità<sup>99</sup>.

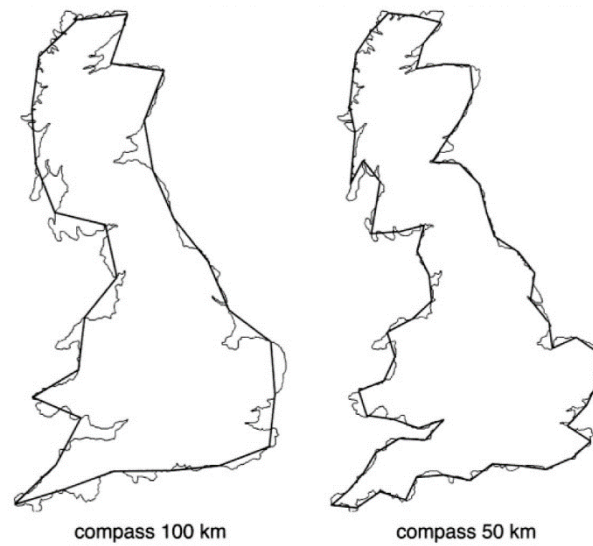


Fig. 14 Approssimazioni della lunghezza della costa della Gran Bretagna. [38]

Richardson, nei suoi scritti, esprime la lunghezza costiera  $L(\varepsilon)$  come una misura proporzionale a:

$$L(\varepsilon) \propto M\varepsilon^{1-D},$$

con  $M > 0$  costante positiva e  $D \geq 1$ . Il valore di  $D$  dipende dalla costa scelta e per Richardson rappresenta un semplice esponente senza un significato particolare. In seguito, Mandelbrot ha notato come il parametro rappresenti una effettiva dimensione della curva, che può assumere valori non interi. È possibile ottenere infatti l'espressione precedente dalla formula della dimensione di omotetia, operando le sostituzioni:

$$\begin{cases} r = \varepsilon/\varepsilon_0 \\ N = L/\varepsilon \end{cases},$$

<sup>99</sup> B. B. MANDELBROT, *Gli oggetti frattali. Forma, caso e dimensione*, 2000, cap. 2.

con  $M = \varepsilon_0^D$ . Per calcolare il valore effettivo della dimensione frattale si può scrivere:

$$\log(L) = \log(M) + (1 - D) \log(\varepsilon),$$

notando come il valore  $(1 - D)$  sia la pendenza della retta passante per il logaritmo dei punti ottenuti empiricamente. La costa della Gran Bretagna, così calcolata, ha  $D = 1,25$ , la Germania  $D = 1,15$ , il confine tra Spagna e Portogallo ha  $D = 1,14$ , la costa dell’Australia  $D = 1,13$ , infine quella del Sud Africa ha  $D = 1,02$ <sup>100</sup>. La dimensione di omotetia può essere usata per frattali con precise caratteristiche di simmetria e invarianza di scala, tuttavia essi rappresentano una sottocategoria della totalità degli oggetti frattali e, al fine di ricavare un’espressione più generale, è necessario considerare delle specificazioni diverse, per esprimere il concetto di dimensione frattale.

### *3.3 La dimensione di Minkowski-Bouligand*

Il procedimento descritto in precedenza, tramite il quale è possibile trovare il valore della dimensione frattale, generando una retta di regressione che approssima i punti ottenuti su un grafico bilogaritmico, consente di definire una diversa specificazione della dimensione frazionaria, chiamata dimensione di Minkowski-Bouligand o dimensione box-counting. La costa della Gran Bretagna non è, in realtà, un frattale auto simile, quindi la dimensione di omotetia è poco adatta, tuttavia è possibile sviluppare un metodo di misurazione di una dimensione approssimata, sfruttando un approccio simile all’uso di un compasso di apertura  $\varepsilon$ , per figure più generiche. L’idea è quella di porre la

---

<sup>100</sup> B. B. MANDELBROT, *How long is the coast of Britain? Statistical self-similarity and fractional dimension*, 1976.

figura in una griglia formata da celle quadrate uguali e contare il numero di celle in cui è contenuta almeno una parte della figura (da qui il termine dimensione box-counting). Diminuendo il lato della singola cella, in modo simile a quanto osservato per la linea costiera da Richardson, il numero di celle in cui è contenuta la figura aumenta<sup>101</sup>. Portando questo procedimento al limite, si ottiene la dimensione di Minkowski-Bouligand della figura in esame. Nel caso in cui non esista finito il limite si fa riferimento alla dimensione superiore e alla dimensione inferiore, che corrispondono rispettivamente al calcolo del limite superiore e inferiore. Dal punto di vista pratico, ciò equivale a calcolare la pendenza della retta di regressione, usando il lato della cella e il numero di celle come variabile indipendente e dipendente, prendendone il logaritmo.

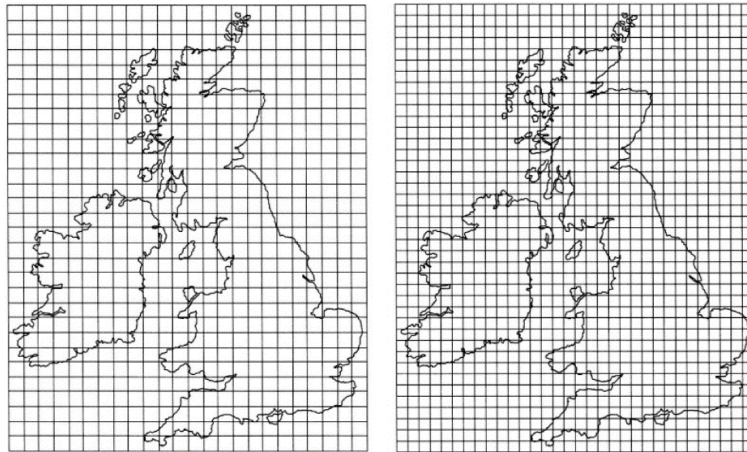


Fig. 15 Dimensione box-counting. [38]

Formalmente, dato un insieme non vuoto  $F \subseteq \mathbb{R}^n$ , si denota con  $N_\delta(F)$  il numero di celle di lato  $\delta$  che possono ricoprire  $F$ <sup>102</sup>. Si definisce dimensione inferiore e superiore di Minkowski-Bouligand di  $F$  rispettivamente:

$$\underline{\dim}_B F = \liminf_{\delta \rightarrow 0} \frac{\log N_\delta(F)}{-\log \delta}$$

<sup>101</sup> H. PEITGEN, H. JÜRGENS, D. SAUPE, *Chaos and fractals. New frontiers of science*, 2004, cap. 4.

<sup>102</sup> K. FALCONER, *Fractal geometry, mathematical foundations and applications*, 2003, cap. 3.



$$\overline{dim}_B F = \limsup_{\delta \rightarrow 0} \frac{\log N_\delta(F)}{-\log \delta}.$$

Se i due limiti esistono e convergono allo stesso valore si definisce dimensione box-counting di  $F$  l'espressione:

$$dim_B F = \lim_{\delta \rightarrow 0} \frac{\log N_\delta(F)}{-\log \delta}.$$

Esistono diversi modi per generalizzare questo tipo di dimensione, ad esempio non utilizzando delle celle quadrate come ricoprimento dell'insieme. Per insiemi regolari questa dimensione coincide con la dimensione di omotetia.

### *3.4 La misura e la dimensione di Hausdorff*

Tramite la dimensione di omotetia si è in precedenza mostrato come un concetto apparentemente intuitivo, come quello di dimensione, possa in realtà celare notevoli complicazioni, se applicato a particolari insiemi con proprietà frattali. Al fine di fornire una cornice teorica più formale al concetto di dimensione frattale, è necessario ricorrere ad ulteriori strumenti matematici, in particolare si sfrutta, in questo paragrafo, la teoria della misura, per definirne una nuova specificazione, nota come dimensione di Hausdorff, sviluppata dal matematico tedesco Felix Hausdorff nel 1918. La maggior parte degli strumenti utilizzati è stata formalizzata dal matematico russo Abram Besicovic, perciò spesso viene chiamata dimensione di Hausdorff-Besicovitch. Dall'esempio della linea costiera si nota che l'effetto Richardson rappresenta un ostacolo per definire, in modo esatto, una dimensione, tuttavia è possibile sviluppare una nozione di contenuto di una curva che non perda di significato se applica a curve di lunghezza tendente a infinito. A livello intuitivo, si può dire che un contenuto

lineare si calcola sommando  $n^1$  passi, ottenendo dall'esponente la dimensione della retta, un contenuto di un'area formata da quadrati si calcola sommando  $n^2$  lati, ottenendo la dimensione di un piano e, infine, per calcolare la lunghezza approssimata di una costa, ricoperta da cerchi di raggio  $r$ , si possono sommare  $r^D$  passi, ottenendo un contenuto approssimato di dimensione  $D$  che varia di poco al variare di  $r$ <sup>103</sup>. Per definire in modo formale e rigoroso la dimensione di Hausdorff è necessario ricavare prima la nozione di misura di Hausdorff. Con questo fine, si fornisce la definizione di misura e di misura esterna, per un dato insieme  $E$  non vuoto, che rappresentano i fondamenti matematici essenziali.

*Definizione. (Misura).* Dato uno spazio misurabile  $(E, S)$  si chiama misura su  $E$  una funzione  $\mu: S \rightarrow [0, +\infty]$  tale che:

- i.  $\mu(\emptyset) = 0$ ,
- ii. ( $\sigma$ -additività)  $\mu(\bigcup_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} \mu(A_k)$ ,  $\forall A_k \in S, A_i \cap A_j = \emptyset, i \neq j$ .

La misura più utilizzata all'interno dello spazio  $\mathbb{R}^n$  è la misura di Lebesgue<sup>104</sup>, che consente di generalizzare i concetti di area e volume per insiemi dello spazio euclideo.

<sup>103</sup> B. B. MANDELBROT, *Gli oggetti frattali. Forma, caso e dimensione*, 2000, cap. 2.

<sup>104</sup> La misura di Lebesgue consente di definire una misura per insiemi  $n$ -dimensionali dello spazio euclideo ed è stata formalizzata dal matematico francese Henri Lebesgue nel 1901. In particolare, si considerano dei rettangoli chiusi  $n$ -dimensionali  $R \subset \mathbb{R}^n$  con i bordi paralleli agli assi del tipo:

$$R = \prod_{i=1}^n [a_i, b_i], \quad b_i \geq a_i \quad \forall i.$$

Il volume dei rettangoli  $v(R)$  è dato da:

$$v(R) = \prod_{i=1}^n (b_i - a_i).$$

Per  $i = 1$  si ottiene la nozione di lunghezza,  $l(R) = b - a$ , e per  $i = 2$  quella di area del rettangolo come prodotto tra base e altezza,  $a(R) = (b_1 - a_1)(b_2 - a_2)$ . Si può definire una misura esterna che approssima un insieme di  $\mathbb{R}^n$  costruendo un ricoprimento come somma di rettangoli  $R_i$  non disgiunti e prendendone l'estremo inferiore. Dato  $E \subset \mathbb{R}^n$  e  $m^*: \mathcal{P}(\mathbb{R}^n) \rightarrow [0, +\infty]$  si definisce misura esterna di Lebesgue il valore:

$$m^*(E) = \inf \left\{ \sum_{i=1}^{\infty} v(R_i) : E \subseteq \bigcup_{i=1}^{\infty} R_i \right\}.$$

*Definizione.* (Misura esterna). Dato un insieme  $E$  non vuoto e  $\mathcal{P}(E)$  insieme delle parti, si definisce misura esterna su  $E$  la funzione  $\varphi: \mathcal{P}(E) \rightarrow [0, +\infty]$  tale che:

- i.  $\varphi(\emptyset) = 0$ ,
- ii. (monotonicità) se  $A \subset B \subset E$ , allora  $\varphi(A) \leq \varphi(B)$ ,
- iii. ( $\sigma$ -subadditività)  $\varphi(\bigcup_{k=1}^{\infty} A_k) \leq \sum_{k=1}^{\infty} \varphi(A_k)$ ,  $\forall A_k \in \mathcal{P}(E)$ .

È importante notare come  $\varphi$  non sia necessariamente una misura poiché l'insieme delle parti, in generale, non costituisce una  $\sigma$ -algebra di  $E$ , inoltre  $\varphi$  non gode della proprietà di  $\sigma$ -additività, in quanto si ammette che  $A_i \cap A_j \neq \emptyset$ ,  $i \neq j$ <sup>105</sup>. Tuttavia, sfruttando il teorema di Caratheodory, è possibile definire uno spazio di misura partendo da una misura esterna.

*Teorema.* (Teorema di Caratheodory). Sia  $\varphi: \mathcal{P}(E) \rightarrow [0, +\infty]$  una misura esterna. Un insieme  $Q \in \mathcal{P}(E)$  è detto  $\varphi$ -misurabile se si ha che:

$$\varphi(A) = \varphi(A \cap Q) + \varphi(A \cap Q^c), \forall A \in \mathcal{P}(E),$$

allora  $\mathcal{M}_{\varphi} = \{Q \subseteq E \mid Q \text{ è } \varphi\text{-misurabile}\}$  è una  $\sigma$ -algebra e di conseguenza esiste  $\mu: \mathcal{M}_{\varphi} \rightarrow [0, +\infty]$  tale che  $\mu(Q) = \varphi(Q)$ , ove  $\mu$  è una misura.

In altre parole, il teorema garantisce sempre l'esistenza di una  $\sigma$ -algebra su cui è possibile definire una misura esterna in termini di una misura. Questo è un risultato fondamentale in quanto si è in grado di descrivere un insieme partendo da una sua approssimazione, sviluppando una misura in grado di generalizzare il concetto di area.

*Definizione.* (Diametro). Dato un insieme  $E$  non vuoto si definisce diametro di  $E$ :

---

Dal teorema di Caratheodory esiste una  $\sigma$ -algebra  $\mathcal{L}(\mathbb{R}^n)$  di insiemi Lebesgue-misurabili in cui si può definire una misura di Lebesgue  $m(E) = m^*(E)$ , ove  $m: \mathcal{L}(\mathbb{R}^n) \rightarrow [0, +\infty]$ .

J. K. HUNTER, *Measure theory*, 2011.

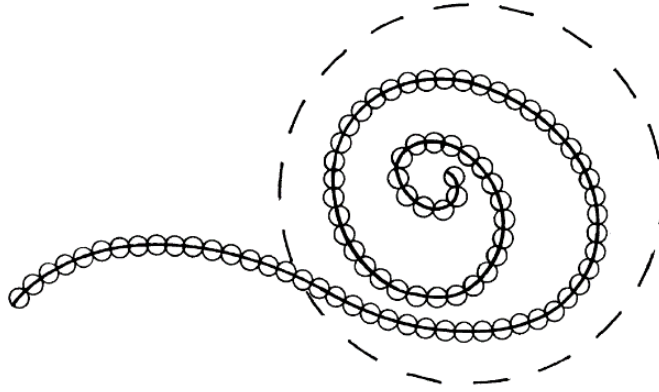
<sup>105</sup> D. L. COHN, *Measure theory*, 1980, cap. 1.

$$\text{diam}(E) = \sup\{|x - y|; x, y \in E\},$$

con  $|\cdot|$  norma euclidea.

*Definizione.* ( $\delta$ -ricoprimento). Dato un insieme  $E$  non vuoto tale che  $E \subseteq \mathbb{R}^n$  e dato un numero reale  $\delta > 0$ , si definisce  $\delta$ -ricoprimento di  $E$  una famiglia numerabile di sottoinsiemi  $\{B_k\}$  non vuoti di  $\mathbb{R}^n$  tali che:

$$\text{diam}(B_k) \leq \delta \text{ e } E \subseteq \bigcup_{k=1}^{\infty} B_k.$$



*Fig. 16*  $\delta$ -ricoprimento di una spirale. Essa è ben approssimata dalla somma dei diametri dei cerchi piccoli e sottostimata dal diametro del cerchio più grande. [32]

Il punto di partenza per ricavare la misura di Hausdorff consiste nel definire il contenuto di  $s$ -dimensionale di Hausdorff.

*Definizione.* (Contenuto di Hausdorff). Dato  $E \subseteq \mathbb{R}^n$  non vuoto,  $\delta > 0$ ,  $s \geq 0$  si definisce il contenuto  $s$ -dimensionale di Hausdorff come:

$$\mathcal{H}_{\delta}^s(E) = \inf \left\{ \sum_{k=1}^{\infty} [\text{diam}(B_k)]^s : \{B_k\} \text{ è un } \delta - \text{ricoprimento di } E \right\},$$

e si pone in seguito  $\mathcal{H}_{\delta}^s(\emptyset) = 0$ .

Se si prende  $\delta' < \delta$ , ogni  $\delta'$ -ricoprimento di  $E$  è anche un  $\delta$ -ricoprimento di  $E$ , quindi ne deriva che  $\mathcal{H}_{\delta'}^s(E) \geq \mathcal{H}_{\delta}^s(E)$ . Si ha allora che la funzione  $\delta \rightarrow \mathcal{H}_{\delta}^s(E)$

è monotona decrescente e ammette limite per  $\delta \rightarrow 0^+$ <sup>106</sup>. Intuitivamente si può pensare all'esempio della linea costiera, in cui, se si diminuisce il raggio dei cerchi che la ricoprono, il contenuto aumenta, in quanto si includono sempre più dettagli al suo interno. Da ciò risulta immediato definire il concetto di misura di Hausdorff, considerando il contenuto quando il diametro tende a zero.

*Definizione.* (Misura di Hausdorff): dato  $E \subseteq \mathbb{R}^n$  e  $0 \leq s < \infty$ , si definisce la misura di Hausdorff  $s$ -dimensionale di  $E$  come:

$$\mathcal{H}^s(E) = \lim_{\delta \rightarrow 0} \mathcal{H}_\delta^s(E) = \sup_{\delta > 0} \mathcal{H}_\delta^s(E).$$

$\mathcal{H}_\delta^s(E)$  e  $\mathcal{H}^s(E)$  sono entrambe misure esterne su  $\mathbb{R}^n$  e, segue dalla definizione che  $\mathcal{H}^s(E) \geq \mathcal{H}_\delta^s(E), \forall \delta > 0$ . Dal teorema di Caratheodory si ottiene che esiste una  $\sigma$ -algebra su cui il contenuto e la dimensione di Hausdorff siano misure. La misura di Hausdorff rappresenta una generalizzazione di quella di Lebesgue, utile per misurare insiemi frattali. Per definire la dimensione di Hausdorff è necessario fare un'ultima osservazione. Si può dimostrare che, dato  $E \in \mathbb{R}^n$ , si ha:

- i. Se  $\mathcal{H}^s(E) < \infty \rightarrow \mathcal{H}^t(E) = 0, con t > s$ ;
- ii. Se  $\mathcal{H}^s(E) > 0 \rightarrow \mathcal{H}^t(E) = \infty, con t < s$ .

Da questa implicazione si ricava la nozione di dimensione di Hausdorff di un insieme, in quanto esiste un unico caso critico in cui  $\mathcal{H}^s(E)$  assume un valore diverso da 0 e  $+\infty$ .

---

<sup>106</sup> G. EDGAR, *Measure, topology and fractal geometry*, 2008, cap. 6.

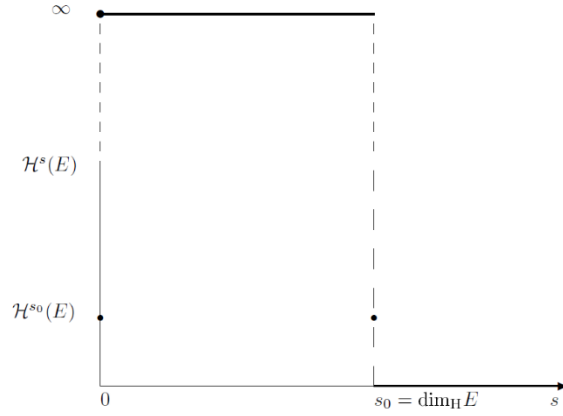


Fig. 17 Misura di Hausdorff. Il punto di discontinuità  $s_0 = \dim_H E$  rappresenta la dimensione di Hausdorff. [58]

*Definizione.* (Dimensione di Hausdorff). Dato un insieme  $E \subseteq \mathbb{R}^n$  è detta dimensione di Hausdorff di  $E$  il valore:

$$\dim_H E = \inf\{s \geq 0 | \mathcal{H}^s(E) = 0\} = \sup\{s \geq 0 | \mathcal{H}^s(E) = +\infty\}.$$

Inoltre, si ha che:

$$0 < \mathcal{H}^s(E) < +\infty \Rightarrow \dim_H E = s,$$

ovvero se esiste una misura di Hausdorff finita e diversa da zero si può facilmente calcolare il valore della dimensione di Hausdorff. Si può dimostrare che  $\dim_H(E) \leq n$  se  $E \subseteq \mathbb{R}^n$  e che  $\dim_H(E) = 0$  se l'insieme  $E$  è numerabile<sup>107</sup>. La dimensione di Hausdorff, in generale, è maggiore di quella topologica per insiemi frattali. Intuitivamente, se l'insieme  $E$  è una curva la lunghezza costituisce un buon modo di misurarne la dimensione e l'area e il volume sono nulli. In altri termini,  $\dim_H E = 2$  e  $\dim_H E = 3$  sono troppo ampi per misurare correttamente la dimensione di della curva  $E$ . Allo stesso modo, se l'insieme  $E$  è identificato dalla superficie di una sfera la dimensione 1 sarà troppo limitata e la dimensione 3 sarà troppo ampia, perciò deve necessariamente verificarsi  $\dim_H E = 2$ <sup>108</sup>. Per insiemi regolari essa coincide

<sup>107</sup> M. M. DODSON, S. KRISTENSEN, *Hausdorff dimension and diophantine approximation*, 2003.

<sup>108</sup> G. EDGAR, *Measure, topology and fractal geometry*, 2008, cap. 6.

con la dimensione di Minkowski-Bouligand, tuttavia esse possono in generale differire. Per esempio,  $\dim_B E = n$  per ogni insieme numerabile  $E \subseteq \mathbb{R}^n$ . Considerando la dimensione dell'insieme dei numeri razionali  $[0,1]$ , si ha che  $\dim_H = 0$  e  $\dim_b = 1^{109}$ . In generale, vale la seguente disuguaglianza:

$$\dim_H \leq \underline{\dim}_B \leq \overline{\dim}_B.$$

### 3.5 Esempi di insiemi frattali

In precedenza, si è visto che gli oggetti frattali sono molto vicini al senso comune, in quanto esempi di strutture frattali possono essere osservate in oggetti presenti in natura, come alberi o felci. Da un punto di vista matematico, nonostante questa familiarità intrinseca, uno sviluppo effettivo della teoria frattale è avvenuto solo in tempi recenti. Per molto tempo, infatti, gli insiemi di tipo frattale sono stati ignorati dalla matematica, riducendoli a semplici paradossi non meritevoli di attenzione. In questo paragrafo si cercherà di applicare quanto visto ad alcuni dei più noti esempi di frattali, dei quali si esaminerà la costruzione, le proprietà e le applicazioni, allo scopo di fornire esempi concreti alle definizioni enunciate. In particolare, si considera una classe di frattali deterministici, simmetrici e auto-simili, in cui i più noti esponenti sono l'insieme di Cantor, la curva di Koch e il triangolo di Sierpinski.

#### 3.5.1 L'insieme di Cantor

L'insieme di Cantor è uno dei primi esempi di paradossi matematici ricaduti all'interno della definizione di Mandelbrot di frattali. È stato teorizzato dal

---

<sup>109</sup> H. PEITGEN, H. JÜRGENS, D. SAUPE, *Chaos and fractals. New frontiers of science*, 2004, cap. 4.

matematico tedesco Georg Cantor nel 1883 come un sottoinsieme dei numeri reali. Esso può essere definito in modo ricorsivo, partendo dall'intervallo  $[0,1]$  ed eliminando il segmento centrale aperto, per ogni iterazione. Ciò che rimane nell'intervallo, dopo infinite ripetizioni, costituisce l'insieme o la polvere di Cantor. Partendo dall'insieme iniziale  $C_0 = [0,1]$  si elimina l'intervallo aperto  $\left(\frac{1}{3}, \frac{2}{3}\right)$  ottenendo  $C_1 = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right]$ . In seguito, si elimina da  $C_1$  l'insieme aperto  $\left(\frac{1}{9}, \frac{2}{9}\right) \cup \left(\frac{7}{9}, \frac{8}{9}\right)$  in modo da ottenere  $C_2 = \left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right] \cup \left[\frac{2}{3}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right]$ . Procedendo all'infinito per la  $n$ -esima iterazione si può scrivere la polvere di Cantor come:

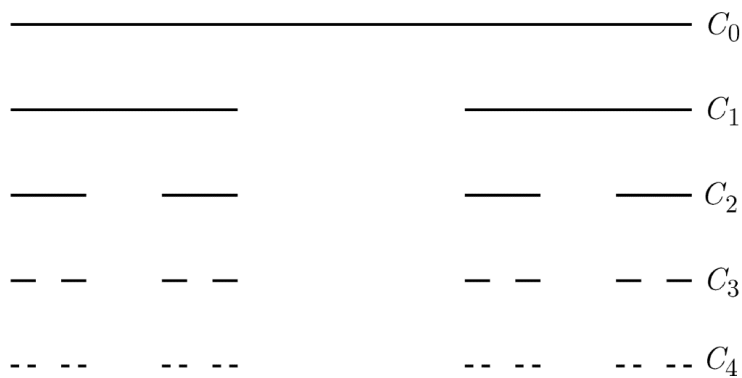
$$C := \bigcap_{n=1}^{\infty} C_n, \quad \text{con } C_n = \frac{C_{n-1}}{3} \cup \left(\frac{2}{3} + \frac{C_{n-1}}{3}\right).$$


Fig. 18 Insieme di Cantor. [32]

È possibile calcolare la misura dell'insieme notando come  $C_n$  sia l'unione di  $2^n$  intervalli disgiunti chiusi, ciascuno con lunghezza  $\left(\frac{1}{3}\right)^n$ . La lunghezza di totale di  $C_n$  è  $\left(\frac{2}{3}\right)^n$  e la misura dell'insieme di Cantor si può esprimere calcolandone il limite, per  $n \rightarrow \infty$ :

$$C = \lim_{n \rightarrow \infty} C_n = \lim_{n \rightarrow \infty} \left(\frac{2}{3}\right)^n = 0.$$



Il risultato ottenuto mostra come la polvere abbia lunghezza totale nulla, tuttavia si può notare che, rimuovendo ad ogni passo insiemi aperti, gli estremi appartengono necessariamente, per definizione, all'insieme finale di Cantor<sup>110</sup>. Ciò si può verificare calcolando la lunghezza totale rimossa,  $C_R$ . Essa può essere rappresentata tramite una serie geometrica convergente con  $|q| < 1$ <sup>111</sup>. Per  $n = 1$  si elimina 1 segmento di lunghezza  $\frac{1}{3}$ , per  $n = 2$  si eliminano 2 segmenti di lunghezza  $\frac{1}{9}$ , procedendo all'infinito si può scrivere:

$$C_R = \frac{1}{3} + \frac{2}{9} + \frac{4}{27} + \frac{8}{81} + \dots = \sum_{n=1}^{\infty} \frac{2^n}{3^{n+1}} = \frac{1}{3} \sum_{n=1}^{\infty} \left(\frac{2}{3}\right)^n = \frac{1}{3} \left( \frac{1}{1 - 2/3} \right) = 1.$$

Si osserva che la lunghezza totale eliminata è uguale all'intero intervallo di partenza  $[0,1]$ . Questa contraddizione rende inaccurato l'utilizzo della misura di Lebesgue, tuttavia l'insieme di Cantor presenta le proprietà di simmetria e invarianza di scala necessarie per calcolare la dimensione frazionaria. In particolare, ogni ingrandimento è una esatta copia dell'insieme totale e si può facilmente calcolare la dimensione di omotetia, notando che il rapporto di scala è  $r = \frac{1}{3}$  e il numero di parti è  $N(r) = 2$ . Il valore della dimensione frattale di similitudine è dato da:

$$D = \frac{\log N}{\log \left(\frac{1}{r}\right)} = \frac{\log 2}{\log 3} \cong 0,631.$$

---

<sup>110</sup> G. EDGAR, *Measure, topology and fractal geometry*, 2008, cap. 1.

<sup>111</sup> Una serie geometrica è una serie del tipo  $s_n = \sum_{n=0}^{\infty} q^n = 1 + q + q^2 + \dots + q^n$  con  $q \in \mathbb{R}$ . Il rapporto di ogni termine con quello precedente è costante ed è detto ragione della serie. Si può dimostrare che:

$$s_n = \begin{cases} \frac{1 - q^{n+1}}{1 - q} & \text{se } q \neq 1 \\ n + 1 & \text{se } q = 1 \end{cases}.$$

Inoltre, per studiare il carattere della serie è necessario calcolare il limite di  $s_n$  per  $n \rightarrow \infty$ :

$$\lim_{n \rightarrow \infty} \sum_{k=0}^n q^k = \begin{cases} \frac{1}{1 - q} & \text{se } |q| < 1 \\ +\infty & \text{se } q \geq 1 \\ \nexists & \text{se } q \leq -1 \end{cases}.$$

L'insieme di Cantor ha quindi una dimensione non intera compresa tra quella di un punto e quella di una retta. L'insieme inoltre non è numerabile e si può dimostrare che ha la stessa cardinalità dell'intervallo  $[0,1]$ , ovvero la potenza del continuo<sup>112</sup>. Il complementare è un insieme formato solamente da intervalli aperti quindi l'insieme di Cantor è un insieme chiuso, inoltre nell'intorno di ogni punto sono presenti sia punti appartenenti che non appartenenti all'insieme stesso, ne segue quindi che ogni punto di  $C$  è un punto di accumulazione<sup>113</sup>.  $C$  è un insieme perfetto, in quanto chiuso senza punti isolati, con la parte interna vuota. Infine, per citare un'applicazione concreta, si è scoperto che sulla Terra lo spettro di emissione di alcuni composti organici, ovvero l'insieme delle frequenze emesse dagli elettroni quando questi compiono una transizione tra stati energetici, assomiglia a una polvere di Cantor<sup>114</sup>.

### 3.5.2 La curva di Koch

Uno degli esempi più famosi di curva frattale è stata immaginata dal matematico svedese Helge von Koch, in un suo articolo del 1904. A causa dell'elevato grado di complicazione e delle proprietà controintuitive, la curva di Koch non ha suscitato un reale interesse, poiché giudicata come una strana curva patologica senza un valore pratico<sup>115</sup>. In realtà, la curva di Koch si colloca all'interno di quella classe di funzioni continue, come ad esempio la funzione di Weierstrass, che non ammettono una tangente in nessun punto. La curva è generata da un algoritmo ricorsivo di costruzione, simile a quello dell'insieme di Cantor, non interseca mai sé stessa e presenta evidenti proprietà di simmetria e invarianza

---

<sup>112</sup> H. PEITGEN, H. JÜRGENS, D. SAUPE, *Chaos and fractals. New frontiers of science*, 2004, cap. 2.

<sup>113</sup> Dato  $A \subset \mathbb{R}$  e  $x_0 \in \mathbb{R}$ ,  $x_0$  è un punto di accumulazione di  $A$  se:

$$\forall I(x_0) \exists x \in A: x \in I(x_0) \setminus \{x_0\}.$$

<sup>114</sup> B. B. MANDEBROT, *Il disordine dei mercati. Una visione frattale di rischio, rovina e redditività*, 2005, cap. 7.

<sup>115</sup> H. LAUWERIER, *Fractals, endlessly repeated geometrical figures*, 1991, cap. 3.

di scala, che ne fanno un ottimo esempio di frattale auto-simile. Per costruirla si parte da un segmento di lunghezza determinata e lo si divide in tre parti uguali. Il secondo passo consiste nell'eliminare il segmento centrale sostituendolo con altri due segmenti uguali disposti in modo da rappresentare due lati di un triangolo equilatero. In questo modo si creano 4 nuovi segmenti di lunghezza pari a  $\frac{1}{3}$  di quello iniziale. Per il terzo passo della costruzione si attua lo stesso procedimento, applicato ai 4 segmenti, in modo da ottenere 16 nuovi segmenti di lunghezza  $\frac{1}{9}$  rispetto all'originale. Ripetendo un numero infinito di volte questo processo si ottiene la rappresentazione della curva Koch<sup>116</sup>. Se si usa come iniziatore un triangolo equilatero, al posto di un segmento di lunghezza fissata, e si itera la costruzione precedente per ognuno dei lati, si ottiene il fiocco di neve o isola di Koch.

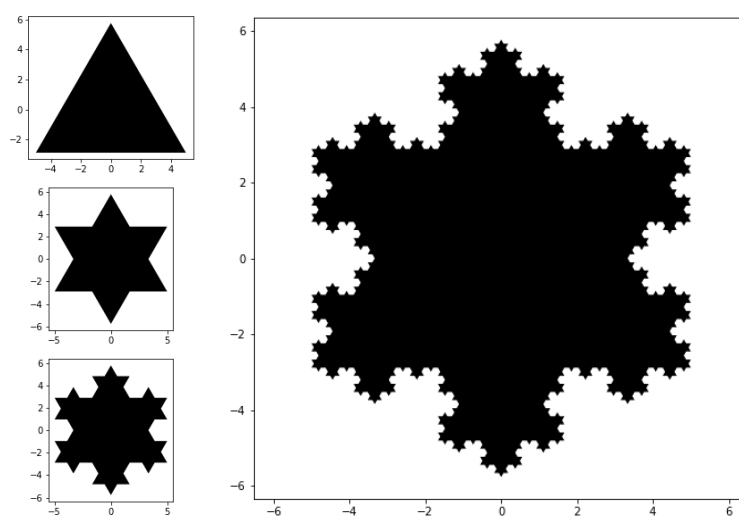


Fig. 19 Fiocco di neve di Koch.

Uno degli aspetti più interessanti di questo frattale, oltre il notevole impatto visivo che suscita la figura, riguarda il calcolo dell'area e del perimetro. Si può facilmente intuire che la lunghezza della curva aumenta, per ogni lato del

<sup>116</sup> H. PEITGEN, H. JÜRGENS, D. SAUPE, *Chaos and fractals. New frontiers of science*, 2004, cap. 2.

triangolo, di  $\frac{4}{3}$  ad ogni iterazione. Definita con  $s$  la lunghezza iniziale del lato, per la  $n$ -esima iterazione, il numero di lati ottenuti è dato da  $N_n = 3 \cdot 4^n$ , ciascuno con lunghezza  $L_n = s \left(\frac{4}{3}\right)^n$ . Si può calcolare quindi il perimetro dell' $n$ -esima iterazione come:

$$P_n = 3s \left(\frac{4}{3}\right)^n.$$

Per esprimere il perimetro totale del fiocco di neve di Koch è sufficiente calcolare il limite per  $n \rightarrow +\infty$ :

$$P = \lim_{n \rightarrow +\infty} P_n = \lim_{n \rightarrow +\infty} 3s \left(\frac{4}{3}\right)^n = +\infty.$$

Al fine di calcolare l'area della regione si ricorda in primo luogo che, se si denota con  $s$  il lato del triangolo equilatero, l'area è data da  $A_0 = \frac{\sqrt{3}}{4} s^2$ , dalle relazioni sui triangoli rettangoli. Dopo la prima iterazione si aggiungono 3 nuovi triangoli equilateri con lato pari a  $\frac{1}{3}s$ . L'area totale è data da  $A_1 = \frac{\sqrt{3}}{4} s^2 + 3 \frac{\sqrt{3}}{4} \left(\frac{s}{3}\right)^2 = \frac{\sqrt{3}}{4} s^2 \left(1 + \frac{3}{9}\right)$ . La seconda iterazione permette di creare 12 nuovi triangoli con lato uguale a  $\frac{1}{9}s$ . L'espressione dell'area è data quindi da  $A_2 = \frac{\sqrt{3}}{4} s^2 \left(1 + \frac{3}{9}\right) + 3 \cdot 4 \frac{\sqrt{3}}{4} \left(\frac{s}{3^2}\right)^2 = \frac{\sqrt{3}}{4} s^2 \left(1 + \frac{3}{9} + \frac{3 \cdot 4}{9^2}\right)$ . In modo analogo, è possibile esprimere l'area dell'isola di Koch per il terzo ciclo di iterazione come:

$$\begin{aligned} A_3 &= \frac{\sqrt{3}}{4} s^2 \left(1 + \frac{3}{9} + \frac{3 \cdot 4}{9^2}\right) + 3 \cdot 4 \cdot 4 \frac{\sqrt{3}}{4} \left(\frac{s}{3^3}\right)^2 \\ &= \frac{\sqrt{3}}{4} s^2 \left(1 + \frac{3}{9} + \frac{3 \cdot 4}{9^2} + \frac{3 \cdot 4^2}{9^3}\right). \end{aligned}$$

Si nota che l'area della  $n$ -esima iterazione è esprimibile attraverso l'uso di una serie geometria:

$$A_n = \frac{\sqrt{3}}{4} s^2 \left( 1 + \sum_{k=1}^n \frac{3 \cdot 4^{k-1}}{9^k} \right).$$

Dalla la convergenza della serie geometrica con  $|q| < 1$ , come per l'insieme di Cantor, l'area totale dall'isola di Koch è data dal limite per  $n \rightarrow +\infty$  dell'area dell' $n$ -esima iterazione:

$$\lim_{n \rightarrow +\infty} \frac{\sqrt{3}}{4} s^2 \left( 1 + \sum_{k=1}^n \frac{3}{9} \left( \frac{4}{9} \right)^{k-1} \right) = \frac{\sqrt{3}}{4} s^2 \left( 1 + \frac{\frac{3}{9}}{1 - \frac{4}{9}} \right) = \frac{\sqrt{3}}{4} s^2 \frac{8}{5} = \frac{2\sqrt{3}}{5} s^2.$$

Sostituendo  $s^2 = \frac{4A_0}{\sqrt{3}}$ , si può esprimere l'area in termini di quella iniziale<sup>117</sup>:

$$A = \frac{8}{5} A_0.$$

Appare evidente la paradossale proprietà della curva di Koch, secondo cui si è ottenuta una curva di lunghezza infinita confinata all'interno di un'area finita, in particolare il fiocco di Koch può essere inscritto in un esagono regolare convesso. In modo analogo a quanto fatto per l'insieme di Cantor è possibile calcolare la dimensione di omotetia della curva di Koch, sfruttando la proprietà di simmetria e il rapporto di scala costante. In particolare, si ha che il numero di parti è  $N(r) = 4$  e il rapporto costante di omotetia è  $r = \frac{1}{3}$ . La dimensione frattale si può calcolare come:

$$D = \frac{\log N}{\log \left( \frac{1}{r} \right)} = \frac{\log 4}{\log 3} \cong 1,2618.$$

Intuitivamente, si osserva che la curva di Koch occupa più spazio di una retta unidimensionale ma non è in grado di riempire interamente il piano bidimensionale. L'idea di una curva infinita all'interno di uno spazio finito è analoga all'analisi di Richardson sulla lunghezza della costa della Gran

---

<sup>117</sup> P. S. ADDISON, *Fractals and chaos, an illustrated course*, 1997, cap. 2.

Bretagna. La curva di Koch, infatti, potrebbe essere un buon modello di partenza per descrivere l'evolversi di una linea costiera, tuttavia presenta alcuni difetti. Nello specifico, la simmetria e il rapporto di omotetia costante di  $\frac{1}{3}$  forniscono un'eccessiva regolarità al frattale di Koch, che non può essere ritrovata in una linea costiera. Quest'ultima, infatti, è il risultato dell'azione nel tempo di innumerevoli eventi, impossibili da descrivere all'interno di una rappresentazione deterministica<sup>118</sup>. Per rendere la curva di Koch un'efficace simulazione costiera è dunque necessario introdurre un elemento casuale all'interno dell'algoritmo di costruzione. In particolare, partendo da un iniziatore e da un generatore noto, è possibile ottenere un frattale aleatorio, scegliendo l'orientamento di sostituzione del generatore nell'iniziatore con una certa distribuzione di probabilità. Ad esempio, facendo riferimento alla costruzione di Koch, si può immaginare di aggiungere ad ogni iterazione triangoli equilateri rivolti verso l'alto, con una probabilità, e triangoli rivolti verso il basso, con una probabilità complementare. La linea irregolare così ottenuta presenta un andamento più naturale, che si adatta meglio all'approssimazione di una costa<sup>119</sup>.

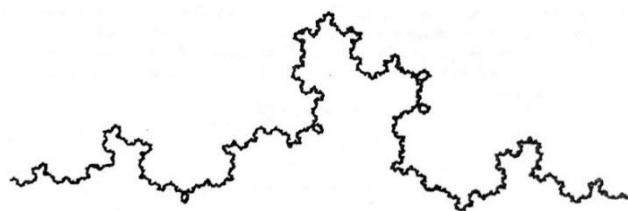


Fig. 20 Curva di Koch casuale. [10]

<sup>118</sup> B. B. MANDELBROT, *Gli oggetti frattali. Forma, caso e dimensione*, 2000, cap. 3.

<sup>119</sup> B. B. MANDELBROT, *Il disordine dei mercati. Una visione frattale di rischio, rovina e redditività*, 2005, cap. 7.

### 3.5.3 Il triangolo di Sierpinski

L'ultimo esempio di frattale che si prende in esame è il triangolo di Sierpinski. Esso è stato descritto dal punto di vista matematico per la prima volta dal matematico polacco Waclaw Sierpinski nel 1915, tuttavia forme geometriche simili sono state identificate in vari mosaici del XIII secolo. Ne sono un esempio il pavimento dei Cosmati della cattedrale di Anagni e quelli della basilica di Santa Maria in Cosmedin a Roma<sup>120</sup>. Al pari degli altri frattali analizzati, il triangolo di Sierpinski non è descritto da un'unica equazione matematica ma da un algoritmo predefinito. L'iniziatore di partenza, come nel caso della curva di Koch, è un triangolo equilatero. Il primo passo consiste nell'identificare e congiungere i punti medi dei lati, in modo da formare quattro nuovi triangoli equilateri. L'eliminazione del triangolo centrale consente di creare tre triangoli equilateri congruenti con lunghezza del lato pari alla metà di quello dell'iniziatore. La seconda fase consiste nel ripetere nuovamente la costruzione precedente per i tre triangoli, identificando e congiungendo i punti medi ed eliminando il triangolo centrale, in modo da ottenere nove nuovi triangoli equilateri congruenti. Iterando un numero infinito di volte questa costruzione, i punti rimanenti formano la figura frattale di Sierpinski. Appare evidente il carattere auto-similare del triangolo, in quanto ogni nuovo triangolo equilatero è una versione ridotta dell'originale, con un costante rapporto di scala<sup>121</sup>.

---

<sup>120</sup> E. CONVERSANO, L. L. TEDESCHINI, *Sierpinski triangles in stone, on medieval floors in Rome*, 2011.

<sup>121</sup> H. PEITGEN, H. JÜRGENS, D. SAUPE, *Chaos and fractals. New frontiers of science*, 2004, cap. 2.

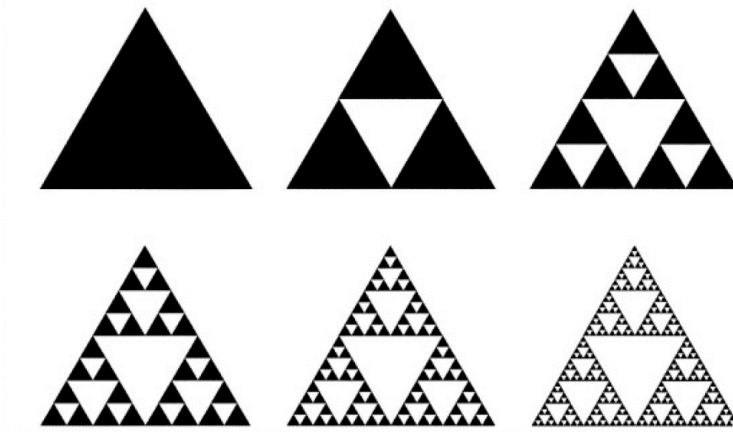


Fig. 21 Triangolo di Sierpinski. [38]

Come nel caso dell'isola di Koch, è interessante calcolare il perimetro e l'area del triangolo di Sierpinski. Se si identifica con  $A_0$  l'area del triangolo iniziale, si può notare come, dopo la prima iterazione, l'area totale, data dalla somma dell'area dei tre nuovi triangoli ottenuti, sia  $A_1 = A_0 \left(\frac{3}{4}\right)$ . Allo stesso modo, dopo la seconda iterazione, si può dire che  $A_2 = A_1 \left(\frac{3}{4}\right) = A_0 \left(\frac{3}{4}\right)^2$ . Procedendo con la costruzione, l'area della  $n$ -esima iterazione è data da:

$$A_n = A_0 \left(\frac{3}{4}\right)^n.$$

L'area finale del triangolo di Sierpinski si ottiene calcolando il limite per  $n \rightarrow +\infty$  di  $A_n$ :

$$A = \lim_{n \rightarrow +\infty} A_n = \lim_{n \rightarrow +\infty} A_0 \left(\frac{3}{4}\right)^n = 0.$$

Per calcolare il perimetro si può notare che, fissato  $l$  il lato dell'iniziatore, con perimetro iniziale  $P_0 = 3l$ , dopo la prima iterazione si formato 3 triangoli equilateri ciascuno con lato  $\frac{1}{2}l$ . Il perimetro dopo il primo passo è dato da  $P_1 = \frac{3}{2}P_0$ . Si può calcolare il perimetro della  $n$ -esima iterazione come:

$$P_n = P_0 \left(\frac{3}{2}\right)^n.$$



Per  $n \rightarrow +\infty$ , si ottiene la lunghezza totale dei lati del triangolo di Sierpinski:

$$P = \lim_{n \rightarrow +\infty} P_n = \lim_{n \rightarrow +\infty} P_0 \left(\frac{3}{2}\right)^n = +\infty.$$

Si può osservare, in modo simile a quanto visto per il fiocco di neve di Koch, una contraddizione intrinseca alla natura frattale del triangolo, il quale ha un perimetro infinito contenuto in un'area nulla. Sfruttando la simmetria e l'invarianza di scala si può calcolare la dimensione frattale, osservando che  $N(r) = 3$  e  $r = \frac{1}{2}$ . La dimensione di omotetia è data da:

$$D = \frac{\log N}{\log \left(\frac{1}{r}\right)} = \frac{\log 3}{\log 2} \cong 1,585.$$

Un metodo alternativo per costruire il triangolo di Sierpinski consiste in quello che è comunemente definito come 'chaos game'. Fissati  $v_1, v_2, v_3$ , i tre vertici del triangolo, e un punto  $p_1$  scelto in modo casuale si definisce la successione:

$$p_{n+1} = \frac{1}{2}(p_n + v_{r_n}),$$

ove  $r_n$  è un numero casuale scelto tra 1, 2 o 3. Se si disegnano tutti i punti da  $p_1$  a  $p_\infty$  si può osservare la formazione del triangolo di Sierpinski, a condizione che il punto iniziale appartenga al triangolo. In altre parole, da qualunque punto iniziale, scelto interno al triangolo di Sierpinski, se si posizionano i punti successivi calcolando la semi-distanza tra un vertice casuale e il punto precedente, appare sempre il triangolo di Sierpinski come figura finale. Si dice che il triangolo di Sierpinski è un attrattore del 'chaos game'<sup>122</sup>.

## *Appendice D. I sistemi di funzioni iterate e gli attrattori*

---

<sup>122</sup> H. PEITGEN, H. JÜRGENS, D. SAUPE, *Chaos and fractals. New frontiers of science*, 2004, cap. 6.

Fino ad ora, si è descritto il processo di costruzione dei frattali da un punto di vista geometrico, attraverso trasformazioni grafiche, allo scopo di rendere immediata la comprensione delle peculiarità di questa particolare classe di oggetti geometrici. Nella prima parte del capitolo si è introdotta la distinzione dei frattali in base al tipo di algoritmo di costruzione, tuttavia è necessario essere più specifici, per comprendere come vengono costruiti gli oggetti frattali, anche da un punto di vista algebrico. Si fa riferimento, in questo paragrafo, all'algoritmo più comune, ovvero la costruzione tramite sistemi di funzioni iterate (IFS). In generale, si può definire una funzione come una regola che trasforma un valore in un altro valore. L'idea di iterare una funzione consiste, invece, nell'applicare ripetutamente la regola in questione ad ogni output ottenuto, in modo da creare una sequenza dinamica di valori. Se le funzioni iterate sono trasformazioni affini di contrazione, grazie alla proprietà di auto-similarità, si ottiene un modo per costruire frattali, facendo uso di una rappresentazione matriciale<sup>123</sup>. Prima di approfondire gli IFS è necessario definire lo spazio nel quale si applicano queste trasformazioni, lo spazio metrico<sup>124</sup>.

*Definizione.* (Spazio metrico). Uno spazio metrico  $(X, d)$  è un insieme  $X$  su cui è definita una funzione  $d: X \times X \rightarrow \mathbb{R}$ , che misura la distanza tra due elementi  $x, y \in X$ . La funzione  $d$  è detta metrica se valgono le seguenti proprietà:

---

<sup>123</sup> M. BARNESLEY, *Fractals everywhere*, 2000, cap. 2.

<sup>124</sup> Più precisamente, lo spazio metrico in cui operano gli IFS è uno spazio metrico completo, ovvero uno spazio metrico in cui tutte le successioni di Cauchy convergono ad un elemento dello spazio. Una successione  $\{x_n\}_{n=1}^{\infty}$  di punti dello spazio metrico  $(X, d)$  è una successione di Cauchy se  $\forall \varepsilon > 0$  esiste un intero  $N > 0$  tale che  $d(x_n, x_m) < \varepsilon$ ,  $\forall n, m > N$ . In altre parole, da un certo punto in poi i punti della successione tendono ad avvicinarsi. La successione si dice convergente a un punto  $x \in X$  se  $d(x_n, x) < \varepsilon$ ,  $\forall n > N$ , ovvero  $\lim_{n \rightarrow \infty} x_n = x$ . Lo spazio metrico si dice completo se ogni  $\{x_n\}_{n=1}^{\infty}$  in  $X$  ha il limite in  $x \in X$ . Uno spazio metrico completo è detto spazio di Banach.

- i.  $d(x, y) = d(y, x) \quad \forall x, y \in X$ ;
- ii.  $0 < d < \infty \quad \forall x, y \in X, x \neq y$ ;
- iii.  $d(x, y) = 0 \quad \forall x, y \in X, x = y$ ;
- iv. (disuguaglianza triangolare)  $d(x, y) \leq d(x, z) + d(z, y) \quad \forall x, y, z \in X$ .

Lo spazio metrico più semplice è lo spazio euclideo, sul quale è definita la comune distanza tra due punti<sup>125</sup>. Questa distanza non è adatta a descrivere con esattezza un frattale ed è necessario introdurre una più generale, che fa riferimento alla distanza tra due sottoinsiemi di uno spazio metrico, nota come distanza di Hausdorff.

*Definizione.* ( $\varepsilon$ -espansione). Dato un insieme  $A \subseteq X$ , si definisce  $\varepsilon$ -espansione di  $A$ ,  $A_\varepsilon$ , l'insieme di tutti i punti del piano che distano da  $A$  non più di  $\varepsilon$ :

$$A_\varepsilon = \{x \in X \mid d(x, y) \leq \varepsilon, y \in A\}.$$

Usando questa nozione, si può definire la distanza di Hausdorff,  $h(A, B)$ , tra due insiemi  $A$  e  $B$ . L'idea è quella di inserire  $A$  in  $B_\varepsilon$  e  $B$  in  $A_\varepsilon$ , variando  $\varepsilon$ . Il più piccolo valore di  $\varepsilon$ , così definito, è la distanza di Hausdorff<sup>126</sup>.

*Definizione.* (Distanza di Hausdorff). Dato uno spazio metrico  $(X, d)$ ,  $\forall A \subseteq X$ ,  $\varepsilon > 0$ , si considera una  $\varepsilon$ -espansione di  $A$ :

$$A_\varepsilon = \{x \in X \mid d(x, y) \leq \varepsilon, y \in A\}.$$

Dati  $A, B \subseteq X$  si dice distanza di Hausdorff la quantità:

$$h(A, B) = \inf\{\varepsilon > 0 \mid A \subset B_\varepsilon \text{ e } B \subset A_\varepsilon\}.$$

Per la distanza di Hausdorff valgono le seguenti proprietà:

- i.  $h(A, B) = 0$  se  $A = B$ ;
- ii.  $h(A, B) = h(B, A)$ ;
- iii.  $h(A, B) \leq h(A, C) + h(C, B)$ .

---

<sup>125</sup> La distanza euclidea tra due punti  $P_1(x_1, y_1)$  e  $P_2(x_2, y_2)$  in uno spazio bi-dimensionale è data da:

$$d(P_1, P_2) = |(x_1, y_1) - (x_2, y_2)| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}.$$

<sup>126</sup> J. HENRIKSON, *Completeness and total boundedness of the Hausdorff metric*, 1999.

Se  $A$  e  $B$  sono punti si ricava la distanza in senso euclideo. È ora possibile introdurre la funzione di contrazione, che permette di ridurre la distanza tra due elementi di uno spazio metrico, che rappresenta l'operazione di base degli IFS.

*Definizione.* (Contrazione). Dato uno spazio metrico  $(X, d)$ , si definisce contrazione una funzione  $f: X \rightarrow X$  tale che esiste una costante  $0 < s < 1$  che soddisfa la seguente condizione:

$$d(f(x), f(y)) \leq s \cdot d(x, y) \quad \forall x, y \in X,$$

con  $s$  detto fattore di contrazione per  $f$ .

La validità dei sistemi di funzioni iterate, come metodo di costruzione di alcuni oggetti frattali, è giustificata da un fondamentale teorema, noto come teorema di Banach-Caccioppoli o teorema delle contrazioni, il quale garantisce l'esistenza e l'unicità di un punto fisso<sup>127</sup> per contrazioni in spazi di Banach.

*Teorema.* (Teorema delle contrazioni). Data  $f: X \rightarrow X$  una contrazione sullo spazio di Banach  $(X, d)$ ,  $f$  possiede un unico punto fisso  $x_f \in X$ , inoltre  $\forall x \in X$  la sequenza  $x_0, x_1, x_2, \dots$  definita da  $x_{n+1} = f(x_n)$  converge a  $x_f$ , ovvero si ha che:

$$\lim_{n \rightarrow \infty} f(x_n) = x_f.$$

Questi ultimi risultati sono stati definiti per un generico spazio di Banach  $(X, d)$ , senza specificare l'uso della distanza di Hausdorff, in quanto si può dimostrare che data  $f: X \rightarrow X$  su  $(X, d)$ , con fattore di contrazione  $s$ , allora  $f: \mathcal{H}(X) \rightarrow \mathcal{H}(X)$  è una contrazione di fattore  $s$  dello spazio  $(\mathcal{H}(X), h(d))$ , con  $h(d)$  distanza di Hausdorff<sup>128</sup>. Con questa precisazione è ora possibile definire i sistemi di funzioni iterate o IFS.

---

<sup>127</sup> Data  $f: A \rightarrow A$  un punto fisso è un elemento  $x \in A$  che coincide con la sua immagine, ovvero tale che  $x = f(x)$ .

<sup>128</sup> M. BARNESLEY, *Fractals everywhere*, 2000, cap. 3.

*Definizione.* (IFS). Un sistema di funzioni iterate è una famiglia finita di funzioni di contrazione  $f_i$  su uno spazio metrico, ciascuna con fattore di contrazione  $s_i$ . In formula:

$$\{f_i: X \rightarrow X | 1 \leq i \leq N\} \quad N \in \mathbb{N}.$$

Dato un IFS si definisce l'operatore di Hutchinson  $\Phi$  associato al sistema la funzione che associa a un sottoinsieme  $S \subset X$  l'unione delle sue contrazioni:

$$\Phi_n(S) = f_1(S) \cup f_2(S) \cup f_3(S) \dots \cup f_n(S) = \bigcup_{i=1}^n f_i(S).$$

Si può dimostrare che  $\Phi$ , a sua volta, è una contrazione rispetto alla distanza di Hausdorff, ovvero  $h(\Phi(B), \Phi(C)) \leq s \cdot h(B, C)$ ,  $\forall B, C \in \mathcal{H}(X)$ , quindi, dal teorema delle contrazioni, esiste ed è unico il punto fisso  $A \in \mathcal{H}(X)$  tale che  $A = \Phi(A)$  dato da:

$$\lim_{n \rightarrow \infty} \Phi_n(S) = A, \quad \forall S \in \mathcal{H}(X).$$

Il punto fisso  $A \in \mathcal{H}(X)$  è detto attrattore del sistema di funzioni iterate. Esso è unico per ogni IFS e si può generare partendo da qualunque sottoinsieme iniziale  $S_0 \subset X$  e applicando ripetutamente l'operatore di Hutchinson, definendo la successione  $S_{n+1} = \Phi(S_n)$ ,  $n = 0, 1, 2, \dots$ . Questa successione converge all'attrattore del sistema, che in alcuni casi ha natura frattale<sup>129</sup>. In altri termini:

$$\lim_{n \rightarrow \infty} h(S_n, A) = 0.$$

Iterando  $\Phi$ , la distanza di Hausdorff tra la successione e l'attrattore tende ad annullarsi. Gli algoritmi per creare IFS possono essere deterministici, se si applica successivamente, partendo da  $S_0$ , l'iterazione di  $\Phi$ , oppure casuali, se si assegna ad ogni  $f_i$  una probabilità  $p_i$  tale che  $\sum_i p_i = 1$ , e ad ogni iterazione si

---

<sup>129</sup> H. PEITGEN, H. JÜRGENS, D. SAUPE, *Chaos and fractals. New frontiers of science*, 2004, cap. 5.

sceglie casualmente la trasformazione di contrazione, in base alla probabilità assegnata<sup>130</sup>. Partendo da specifici IFS è possibile costruire, con questo metodo, tutti i frattali citati nel precedente capitolo. Lo studio degli IFS è molto importante in riferimento a sistemi caotici, nei quali spesso si presenta un attrattore con proprietà frattali, detto attrattore strano<sup>131</sup>.

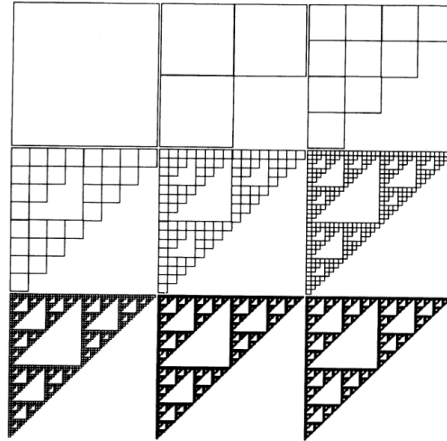


Fig. 22 IFS del triangolo di Sierpinski. [53]

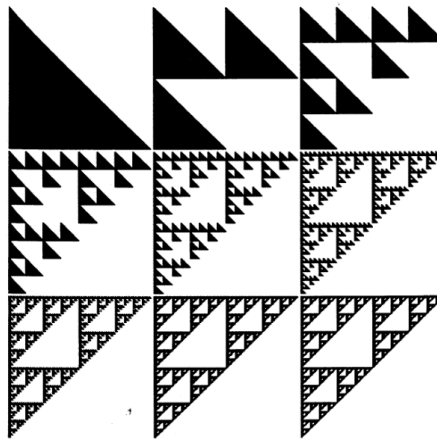


Fig. 23 IFS con diverso insieme di partenza. Al crescere di  $n$  si genera nuovamente il triangolo di Sierpinski, quindi ogni IFS ha un preciso attrattore, indipendentemente da  $S_0$ . [53]

<sup>130</sup> M. BARNSLEY, *Fractals everywhere*, 2000, cap. 3.

<sup>131</sup> H. PEITGEN, H. JÜRGENS, D. SAUPE, *Chaos and fractals. New frontiers of science*, 2004, cap. 12.

## **CAPITOLO IV.**

# **LA NUOVA TEORIA FINANZIARIA E IL MODELLO MULTIFRATTALE DEI MERCATI**

#### *4.1 La costruzione di grafici frattali e i cartoni finanziari di Brown-Bachelier*

Nel precedente capitolo sono stati presentati gli oggetti frattali, specificando le proprietà controintuitive che caratterizzano questo nuovo settore di indagine. A questo punto, si può compiere un ulteriore passo e implementare la teoria dei frattali in un contesto artificiale, come i mercati finanziari. La teorizzazione degli oggetti frattali non è avvenuta in un momento preciso ma è stato un processo, che ha riguardato buona parte della vita di Mandelbrot. Allo stesso modo, gli elementi essenziali che caratterizzano la teoria frattale dei mercati sono il risultato di modelli intermedi, elaborati dal matematico polacco nell'arco di più di dieci anni. In primo luogo, ci si può chiedere come sia possibile realizzare dei grafici frattali dei mercati e che tipo di informazioni se ne possono trarre. Per rispondere, si prende in esame un modello noto, quello di Bachelier, e si costruisce un grafico frattale, chiamato cartone finanziario, che ne rispetti le proprietà. In particolare, il modello ipotizza che il prezzo di uno strumento finanziario si muova come un moto browniano standard, con variazioni del prezzo indipendenti tra loro e distribuite secondo una curva normale, con la maggior parte di esse disposte attorno al valore medio. Il processo di costruzione dei frattali, esaminato nel precedente capitolo, prevede l'esistenza di un iniziatore, di un generatore e di una regola ricorsiva. Per realizzare il cartone di Brown-Bachelier si segue un metodo simile. Si inizia disegnando un riquadro con dimensioni unitarie (nella fig. 24 la larghezza è dilatata per chiarezza espositiva), all'interno del quale si traccia un segmento dall'angolo in basso a sinistra a quello in alto a destra. Questo segmento è l'iniziatore, detto linea di tendenza del mercato, in quanto fornisce l'indicazione che, indipendentemente



dalle oscillazioni, il mercato ha una tendenza rialzista. Per incorporare una diminuzione generale si può scegliere come linea di tendenza la diagonale opposta. A questo punto si disegna il generatore, ovvero una linea spezzata che interseca l'inziatore, formata in questo caso da tre punti di rottura, la cui posizione è in accordo con il modello di Bachelier<sup>132</sup>. La regola di ricorsione si attua sostituendo una copia del generatore ridotto in ogni parte dell'inziatore e ripetendo il processo per segmenti sempre più piccoli. Il risultato è una linea irregolare e frastagliata che ricorda un grafico finanziario.

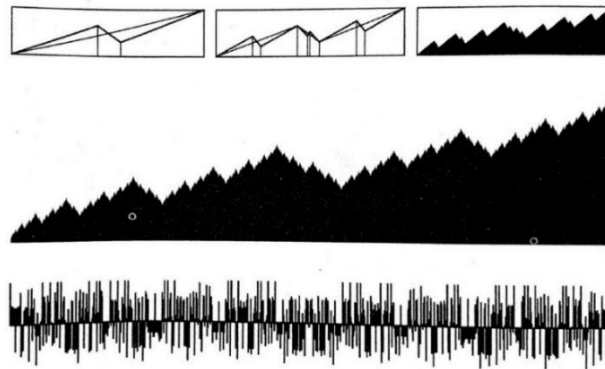


Fig. 24 Cartone finanziario di Brown-Bachelier. [10]

Come si può immaginare, la scelta del generatore determina il risultato finale del grafico. In particolare, la scelta di un generatore a zig-zag di tipo su-giù-su, come nell'esempio precedente, determina un risultato molto lontano dalla

<sup>132</sup> Il generatore è formato da tre parti, che si incontrano in punti con una specifica proprietà. L'altezza di ogni pezzo è infatti la radice quadrata della sua lunghezza. La prima parte ha altezza  $\frac{2}{3}$  e lunghezza  $\frac{4}{9}$ , la seconda  $\frac{1}{3}$  di altezza e  $\frac{1}{9}$  di lunghezza e la terza ha le stesse dimensioni della prima. In questo modo si può definire una nuova grandezza  $H$  per ogni parte pari a:

$$H = \frac{\log(\text{altezza})}{\log(\text{larghezza})}.$$

Per il primo pezzo si ottiene  $H = \frac{\log(2/3)}{\log(2/3)^2} = \frac{1}{2}$ . Questo valore è costante per l' $n$ -esimo intervallo, in questo esempio. Se si considera che  $N$  variabili indipendenti con distribuzione normale standardizzata possiedono  $\sigma = N^{\frac{1}{2}}$ , si può notare che il cartone finanziario presenta un comportamento analogo a un moto browniano standard.

B. B. MANDELBROT, *Scaling in financial prices, III: Cartoon brownian motions in multifractal time*, 2001.

quotidianità e dall'andamento reale dei prezzi. Per aumentare il grado di realismo, si può adottare un processo casuale per decidere il tipo di sostituzione. Nella figura seguente sono stati costruiti tre diversi generatori iniziali, di tipo giù-su-su, su-giù-su e su-su-giù, il cui ordine di sostituzione sulla linea di tendenza è scelto in modo casuale.

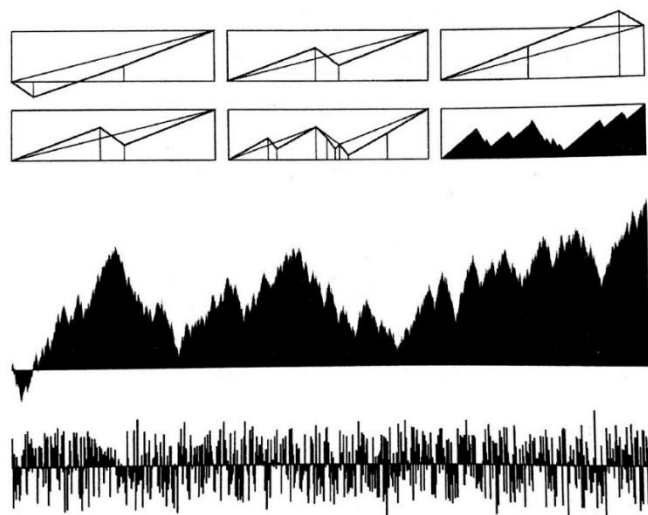


Fig. 25 Cartone di Brown-Bachelier con casualità. [10]

È possibile costruire numerosi tipi di generatori, eliminando anche il quadrato iniziale, per creare risultati complessi, sempre più simili all'andamento dei titoli. Tuttavia, nella realtà il valore del prezzo non è determinato da questo processo geometrico. Incorporare in un modello ogni singola variazione annulla l'utilità del modello stesso, perciò i cartoni si propongono di replicare statisticamente le variazioni reali, sviluppando un nuovo strumento per indagare la dinamica evolutiva delle turbolenze presenti nei mercati finanziari<sup>133</sup>. Modificando la scala di osservazione di un grafico di un titolo si osserva un andamento simile a quello originario, che può presentare fattori di scala differenti in varie parti. Nel capitolo precedente, gli oggetti frattali con questa caratteristica sono stati

<sup>133</sup> B. B. MANDEBROT, *Il disordine dei mercati. Una visione frattale di rischio, rovina e redditività*, 2005, cap. 6.

definiti multifrattali. Il concetto di multifrattalità è un elemento che fa parte delle ultime specificazioni del modello di Mandelbrot in esame, tuttavia è utile introdurlo ora in modo intuitivo, facendo riferimento ai cartoni finanziari. La volatilità limitata del modello di Brown-Bachelier è rispecchiata nel grafico dalla proporzione costante tra prezzo e tempo, ovvero tra le lunghezze dei lati del riquadro iniziale. L'uso dei multifrattali consente di generare un modello in cui il prezzo possa variare in modo non costante, in base all'attività del mercato. Dal punto di vista grafico l'idea è quella di allungare o ridurre l'asse temporale, mantenendo invariato quello del prezzo, in modo da dilatare o contrarre il generatore. Nel primo pezzo si avrà un'alta attività del mercato, nell'intervallo di tempo, e nel secondo una velocità più bassa. Procedendo con il processo di iterazione, il multifrattale che si ottiene è caratterizzato da una maggiore volatilità, positivamente correlata con la contrazione iniziale del generatore<sup>134</sup>. Il modello multifrattale presenta un maggior grado di complessità, tuttavia questa intuizione grafica permette di intravedere le potenzialità del modello.

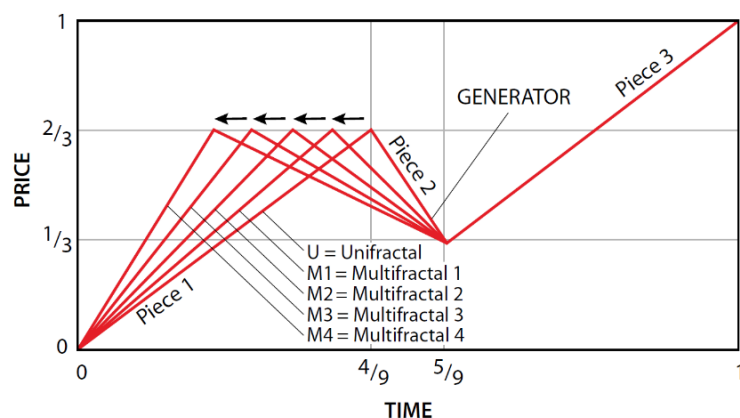


Fig. 26 Diversi generatori per costruire cartoni multifrattali. [6]

<sup>134</sup> B. B. MANDELBROT, *A multifractal walk down Wall Street*, 1999.

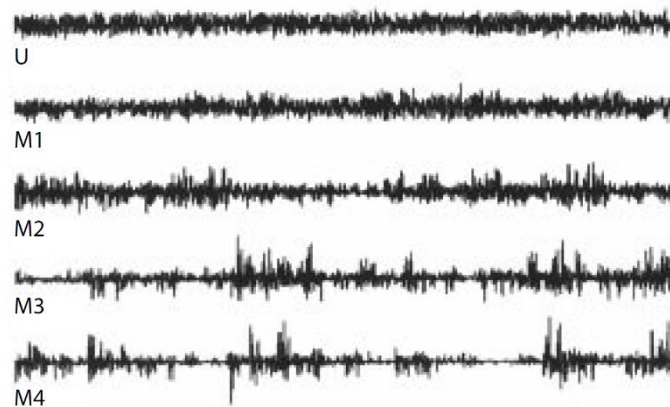


Fig. 27 Variazioni dei prezzi nei cartoni multifrattali. [6]

## 4.2 L'analisi dei prezzi del cotone e le leggi di potenza

Lo sviluppo di una teoria frattale dei mercati finanziari, come detto, è stato un processo lungo e complicato che ha richiesto numerose formulazioni intermedie, tuttavia è possibile identificare con esattezza l'episodio da cui trae origine l'intero modello. Nel 1961, invitato ad Harvard per intervenire ad una conferenza, Mandelbrot ha notato, nell'ufficio del professore Hendrik Houthakker, su una lavagna, lo stesso grafico che avrebbe dovuto presentare alla conferenza, riguardante la distribuzione dei redditi. In realtà, il grafico raffigurava le variazioni dei prezzi del cotone, le quali non erano coerenti con il modello di Bachelier. Da questo episodio è nato l'interesse del matematico polacco per i mercati finanziari e per il cotone, poiché, a differenza di altre attività, era possibile ottenere registrazioni affidabili e di lungo periodo<sup>135</sup>. La somiglianza tra la distribuzione dei redditi e le variazioni del cotone rappresenta un fenomeno singolare, che ha consentito di sviluppare un primo modello dei

<sup>135</sup> B. B. MANDEBROT, *Il disordine dei mercati. Una visione frattale di rischio, rovina e redditività*, 2005, cap. 8.

mercati, sfruttando il concetto di legge di potenza. In verità, si è già utilizzata una legge potenza, nel precedente capitolo, quando si è introdotta la dimensione di omotetia  $D$ , tuttavia è necessario ora soffermarsi su questo tipo di distribuzioni, che rappresentano il primo sostanziale distacco dal modello classico. In generale, una legge di potenza è una funzione che permette ad una variabile di evolvere secondo una potenza di un'altra variabile. Le leggi di potenza hanno suscitato un notevole interesse per le proprietà matematiche di cui godono e per le applicazioni in numerosi fenomeni naturali e artificiali che sembrano avere. In particolare, si può descrivere con una legge di potenza la magnitudo dei terremoti, le eruzioni vulcaniche o la dimensione dei meteoriti, per citare dei fenomeni naturali, la frequenza delle parole, la dimensione dei file del computer o la popolazione delle città, per citare dei fenomeni artificiali<sup>136</sup>. La distribuzione di probabilità di un fenomeno  $x$  che segue una legge di potenza è definita come:

$$p(x) = Cx^{-\alpha},$$

con  $C$  costante di normalizzazione e  $\alpha$  parametro di scala della distribuzione. Si può distinguere tra leggi di potenza continue, se assumono valori nei numeri reali, o discrete, se la quantità può assumere solo valori interi<sup>137</sup>. Si può notare che per  $x \rightarrow 0$  la distribuzione diverge, quindi si definisce solitamente un valore minimo  $x_{min}$  in cui è rispettata la legge di potenza e se ne studia il

---

<sup>136</sup> S. A. FRANK, *The common patterns of nature*, 2009.

<sup>137</sup> Nel caso di variabili discrete un modo di riscrivere la distribuzione di probabilità è  $p_k = Ck^{-\alpha}$ , con  $k$  che assume valori interi e  $\alpha$  esponente di scala. Per  $k = 0$  la funzione diverge quindi si considera solitamente il dominio  $k \geq 1$ , inoltre la costante di normalizzazione  $C$ , per leggi discrete, è data dalla condizione  $\sum_{k=1}^{\infty} p_k = C \sum_{k=1}^{\infty} k^{-\alpha} = C\zeta(\alpha) = 1$ , con  $\zeta(\alpha) = \sum_{k=1}^{\infty} \frac{1}{k^\alpha}$  funzione zeta di Riemann, da cui si ha  $C = \frac{1}{\zeta(\alpha)}$ . Si ottiene quindi che  $p_k = \frac{k^{-\alpha}}{\zeta(\alpha)}$  e, nel caso in cui la legge di potenza valga per valori  $k \geq k_{min}$ , si può scrivere l'espressione più generale della densità come  $p_k = \frac{k^{-\alpha}}{\zeta(\alpha, k_{min})}$ , con  $\zeta(\alpha, k_{min}) = \sum_{k=k_{min}}^{\infty} \frac{1}{k^\alpha}$ . Molti dei risultati ottenuti in questo paragrafo possono essere generalizzati anche per leggi di potenza discrete, tuttavia ciò porta a una maggiore complessità matematica. Distribuzioni discrete di questo tipo sono dette anche distribuzioni zeta.

comportamento in  $[x_{min}, +\infty]$ <sup>138</sup>. La costante di normalizzazione  $C$  ha lo scopo di ridurre una distribuzione generica in una funzione di densità di probabilità con area totale unitaria. In questo caso, per una variabile continua, si ha:

$$\int_{x_{min}}^{\infty} p(x) dx = C \int_{x_{min}}^{\infty} x^{-\alpha} dx = \frac{C}{1-\alpha} [x^{-\alpha+1}]_{x_{min}}^{\infty} = 1.$$

L'espressione ottenuta non diverge solo per valori  $\alpha > 1$  e si può ricavare il valore di  $C$  dato da:

$$C = (\alpha - 1)x_{min}^{\alpha-1}.$$

La funzione di densità di probabilità della legge di potenza si può scrivere come:

$$p(x) = \frac{\alpha - 1}{x_{min}} \left( \frac{x}{x_{min}} \right)^{-\alpha}.$$

I fenomeni non seguono di solito una legge di potenza nell'intero dominio, ma in un intervallo di esso, ad esempio la distribuzione della magnitudo dei terremoti ha un andamento sempre meno paragonabile ad una legge di potenza, man mano che si avvicina al limite dell'energia totale presente nella crosta terrestre. Per elevati valori di  $x$  la distribuzione diverge dalla legge di potenza convergendo velocemente a zero, quindi, per tenere conto di questo elemento, spesso si aggiunge un cutoff esponenziale alla funzione di probabilità, che prevale sulla potenza, per  $x$  grandi:

$$p(x) \propto x^{-\alpha} e^{-\lambda x}.$$

Una importante proprietà delle leggi di potenza può essere ricavata studiando i momenti della distribuzione, in quanto essi non risultano definiti per qualunque valore di  $\alpha$ <sup>139</sup>. In particolare, esaminando il caso di una variabile continua, il valore atteso è dato da:

---

<sup>138</sup> A. CLAUSET, C. R. SHALIZI, M. E. J. NEWMAN, *Power law distributions in empirical data*, 2009.

<sup>139</sup> M. E. J. NEWMAN, *Power laws, Pareto distributions and Zipf's law*, 2006.

$$\langle x \rangle = \int_{x_{min}}^{\infty} xp(x) dx = C \int_{x_{min}}^{\infty} x^{1-\alpha} dx = \frac{C}{2-\alpha} [x^{2-\alpha}]_{x_{min}}^{\infty}.$$

Per  $\alpha \leq 2$  l'espressione diverge e il valore atteso non è definito, invece per  $\alpha > 2$  si ottiene, sostituendo  $C$ :

$$\langle x \rangle = \frac{\alpha - 1}{\alpha - 2} x_{min}.$$

In modo analogo, si può calcolare il secondo momento come:

$$\langle x^2 \rangle = \int_{x_{min}}^{\infty} x^2 p(x) dx = C \int_{x_{min}}^{\infty} x^{2-\alpha} dx = \frac{C}{3-\alpha} [x^{3-\alpha}]_{x_{min}}^{\infty}.$$

Questa volta, per  $\alpha \leq 3$  esso assume un valore infinito, mentre per  $\alpha > 3$  si può ricavare un valore finito:

$$\langle x^2 \rangle = \frac{\alpha - 1}{\alpha - 3} x_{min}^2.$$

In generale, è possibile scrivere l'espressione del momento di ordine  $m$  di una legge di potenza, che non diverge per  $m < \alpha - 1$ , come:

$$\langle x^m \rangle = \int_{x_{min}}^{\infty} x^m p(x) dx = \frac{\alpha - 1}{\alpha - 1 - m} x_{min}^m.$$

La proprietà principale delle leggi di potenza, che rappresenta un'analogia con gli oggetti frattali, è l'invarianza di scala. Così come un frattale auto simile restituisce la stessa rappresentazione ingrandendo una qualunque delle sue parti, una legge di potenza raffigura lo stesso andamento per qualunque scala si scelga di adottare. In altre parole, modificando l'orizzonte di osservazione di un fenomeno, che segue una legge di potenza nel suo dominio, la funzione di probabilità non cambia forma, ovvero una legge di potenza non ammette una scala caratteristica del fenomeno. L'invarianza di scala è una condizione necessaria e sufficiente per definire una legge di potenza. In termini più formali,

data una distribuzione  $p(x)$  se si cambia la scala di misura di  $x$  di un fattore  $b$ , la distribuzione rimane invariata a meno di una costante moltiplicativa  $f(b)$ , ovvero vale che:

$$p(bx) = f(b)p(x).$$

Data  $p(x) = Cx^{-\alpha}$ , moltiplicando la variabile indipendente per un fattore  $b$ , la condizione di invarianza di scala risulta essere soddisfatta<sup>140</sup>:

$$p(bx) = C(bx)^{-\alpha} = b^{-\alpha}Cx^{-\alpha} = b^{-\alpha}p(x).$$

In alternativa, per esprimere l'invarianza di scala, si può notare che il rapporto tra valori doppi della variabile indipendente è costante per qualunque  $x$ , ovvero:

$$\frac{p(x)}{p(2x)} = k, \quad \forall x.$$

Per una legge di potenza si ottiene<sup>141</sup>:

$$\frac{p(x)}{p(2x)} = \frac{Cx^{-\alpha}}{C(2x)^{-\alpha}} = 2^{\alpha}.$$

Graficamente, le leggi di potenza si rappresentano su un grafico log-log, in quanto si ha una relazione del tipo:

$$\log p(x) = \log C - \alpha \log x.$$

La relazione lineare è una condizione necessaria, ma non sufficiente, per essere una legge di potenza, inoltre è conveniente rappresentare sul grafico la distribuzione cumulata di probabilità  $P(x)$ , in quanto essa è definita per ogni valore di  $x$ , in modo da non perdere informazioni. Si può dimostrare che  $P(x)$  è una legge di potenza con esponente più piccolo. In particolare, data una legge di potenza nella forma  $p(x) = Cx^{-\alpha}$ , la funzione di ripartizione (CDF) si può esprimere come:

---

<sup>140</sup> M. E. J. NEWMAN, *Power laws, Pareto distributions and Zipf's law*, 2006.

<sup>141</sup> S. A. FRANK, *The common patterns of nature*, 2009.



$$P(x) = P(X > x) = \frac{C}{\alpha - 1} x^{-(\alpha-1)}.^{142}$$

Una volta definita la legge di potenza, l'obiettivo è quello di sviluppare una stima del parametro di scala  $\alpha$ , che è unico per ogni distribuzione. Una prima metodologia consiste nel calcolare la pendenza della relazione lineare del grafico bilogaritmico, in modo analogo a quanto fatto in precedenza per calcolare la dimensione frattale della linea costiera. In alcuni casi, questo metodo porta a stimare un valore distorto dell'esponente reale<sup>143</sup>, quindi si può utilizzare un diverso approccio, noto come metodo della massima verosimiglianza (MLE), che è approfondito nell'Appendice E. Le leggi di potenza, come detto, hanno ampia applicazione e sono fondamentali per descrivere un elevato numero di fenomeni, in particolare si prendono in esame, di seguito, i due più noti esempi di distribuzioni di potenza, la legge di Zipf e la legge di Pareto.

#### 4.2.1 La legge di Zipf

Il primo caso di legge di potenza approfondito riguarda la distribuzione delle parole nel linguaggio. In particolare, George Kingsley Zipf, un linguista americano professore ad Harvard, si è occupato di studiare la frequenza delle parole nelle frasi, formulando quella che è nota come legge di Zipf. L'idea di base è quella di contare in un testo quante volte compare ciascuna parola. In seguito, si assegna un grado ad ogni frequenza in modo decrescente, ovvero si parte assegnando il grado uno alla parola che appare più volte, il grado due alla seconda parola più comune e così via. Disegnando su un grafico le frequenze e

---

<sup>142</sup> Si può calcolare la distribuzione cumulata di probabilità come:

$$P(x) = P(X > x) = \int_x^\infty p(y) dy = C \int_x^\infty y^{-\alpha} dy = \frac{C}{1-\alpha} [y^{1-\alpha}]_x^\infty = \frac{C}{\alpha-1} x^{-(\alpha-1)}.$$

<sup>143</sup> A. JAMES, M. J. PLANK, *On fitting power laws to ecological data*, 2007.

i gradi si ottiene una forma particolare, che diminuisce in modo irregolare, e presenta una maggiore pendenza per i gradi più bassi<sup>144</sup>. Si può dedurre che nella maggior parte dei testi e dei discorsi siano presenti parole semplici con un'elevata frequenza, che costituiscono la struttura principale del discorso, e parole con più alto contenuto informativo, che si ripetono di meno. Costruendo un grafico log-log si riscontra il tipico andamento lineare delle leggi di potenza.

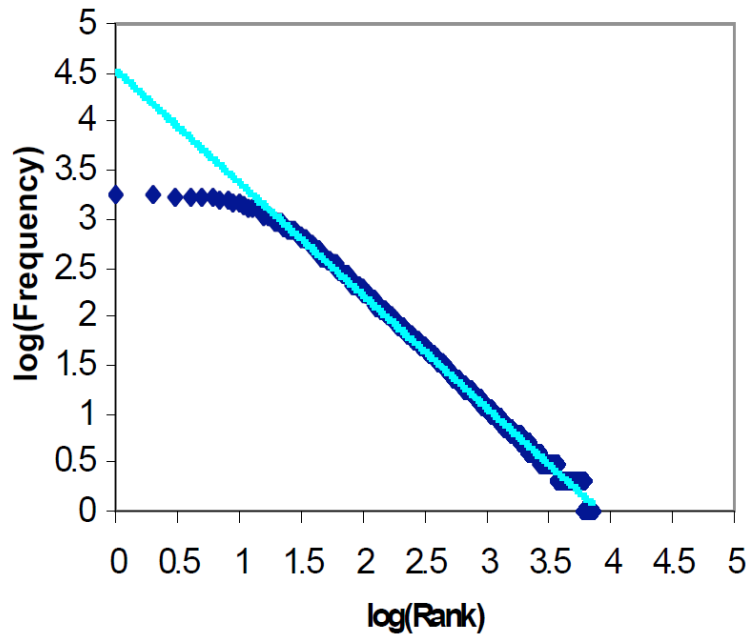


Fig. 28 Distribuzione delle parole nella versione inglese di Robinson Crusoe. [16]

Sfruttando queste informazioni, Zipf è riuscito ad elaborare una legge di potenza discreta, per la frequenza delle parole, del tipo:

$$f(x) = Cx^{-\alpha},$$

ove  $x$  è il grado,  $\alpha$  è l'esponente della distribuzione e  $f(x)$  rappresenta la frequenza delle parole. La costante di normalizzazione è pari a  $C = \frac{1}{\sum_{n=1}^N \frac{1}{n^\alpha}}$

$\frac{1}{H_{n,\alpha}}$ , ove  $N$  è il numero degli elementi e  $H_{n,\alpha} = \sum_{n=1}^N \frac{1}{n^\alpha}$  è il numero armonico generalizzato di ordine  $N$ . La legge di Zipf si può quindi scrivere come:

<sup>144</sup> B. B. MANDEBROT, *Il disordine dei mercati. Una visione frattale di rischio, rovina e redditività*, 2005, cap. 8.

$$f(x) = \frac{1}{x^\alpha H_{n,\alpha}}.$$

Zipf ha stimato un valore di  $C$  costante e pari a  $\frac{1}{10}$  e di  $\alpha$  pari a 1. L'esponente  $\alpha$  rappresenta la ricchezza del vocabolario, un valore maggiore indica un vocabolario più ricco e viceversa. Il contributo principale di Zipf consiste nell'essere stato il primo a ideare un'applicazione delle leggi di potenza all'interno delle scienze sociali, in grado di connettere diverse discipline. L'idea del linguista americano è che le distribuzioni di potenza rappresentino un principio universale, che nasce in qualunque tipo di comportamento, e ha sviluppato una legge matematica che costituisce una reale alternativa alla distribuzione normale<sup>145</sup>. In generale, la legge di Zipf ha mostrato una migliore adattabilità nelle zone centrali delle distribuzioni rispetto alle code, inoltre, egli ha probabilmente sottostimato il reale valore del parametro di scala. Mandelbrot ha approfondito questo tema, ricavando una versione più generale della formula originaria di Zipf, che consente di adattare meglio le frequenze delle parole teoriche a quelle reali attraverso un parametro di spostamento  $\beta$ , nota come legge di Zipf-Mandelbrot<sup>146</sup>:

$$f(x) = C(\beta + x)^{-\alpha}.$$

Sostituendo il valore della costante di normalizzazione la legge si può riscrivere, tramite il numero armonico generalizzato, come:

$$f(x) = \frac{1}{H_{n,\alpha,\beta}(\beta + x)^\alpha},$$

---

<sup>145</sup> B. MANARIS, L. PELLICORO, G. POTHERING, H. HODGES, *Investigating Esperanto's statistical proportions relative to other languages using neural networks and Zipf's law*, 2006.

<sup>146</sup> S. T. PIANTADOSI, *Zipf's word frequency law in natural language: a critical review and future directions*, 2014.

con  $H_{n,\alpha,\beta} = \sum_{n=1}^N \frac{1}{(n+\beta)^\alpha}$ . Per  $N$  finito e  $\beta = 0$  si ottiene la legge di Zipf normale, per  $N$  infinito e  $\beta = 0$  la legge diventa una distribuzione zeta.

#### 4.2.2 La legge di Pareto

Una delle caratteristiche delle leggi di potenza è certamente quella di poter essere applicate nel descrivere fenomeni riguardanti discipline apparentemente molto lontane. In ambito economico, una delle distribuzioni di potenza più importanti deriva dall'analisi della ricchezza degli individui, condotta dall'economista italiano Vilfredo Pareto. L'importanza dell'opera di Pareto consiste nell'aver trasformato l'economia da scienza principalmente sociale a settore di ricerca scientifica e matematica. In particolare, egli si è occupato di studiare come viene distribuita la ricchezza all'interno della società, analizzando dati di redditi e ricchezze, nei diversi Paesi, in diversi orizzonti temporali. In seguito, notando come i dati siano inadatti a essere descritti tramite la curva normale, Pareto è arrivato a formulare un'equazione matematica per descrivere la distribuzione della ricchezza tra classi di individui, nota come legge di Pareto:

$$P(X > x) = \begin{cases} \left(\frac{x}{m}\right)^{-\alpha} & \text{se } x \geq m, \\ 1 & \text{se } x < m \end{cases}$$

con  $m$  reddito minimo,  $\alpha$  esponente di scala e  $x$  ricchezza posseduta dalle persone. L'equazione rappresenta una legge di potenza continua molto simile alla legge di Zipf. Per la distribuzione della ricchezza si usa la funzione di probabilità cumulata, in quanto si è interessati a trovare la percentuale di individui con ricchezza superiore a un dato valore minimo  $m$ . Pareto ha stimato un valore dell'esponente  $\alpha = \frac{3}{2}$ , superiore a quello unitario identificato da Zipf, che indica come la maggior parte della ricchezza sia detenuta da poche persone.

Questa conclusione si può dedurre graficamente, in quanto si osserva una distribuzione fortemente asimmetrica, con la maggior parte delle persone posizionate nella parte bassa della curva.

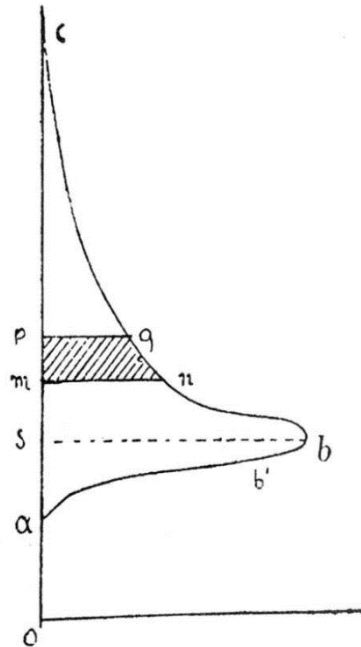


Fig. 29 Grafico di Pareto della distribuzione della ricchezza. [10]

L'esponente  $\alpha$  rappresenta, in questo caso, il grado di disuguaglianza presente nella società, al crescere di  $\alpha$  una quantità sempre maggiore di reddito viene posseduta da sempre meno persone<sup>147</sup>. In verità, Pareto ha sottostimato il valore reale di  $\alpha$ , probabilmente a causa della limitatezza dei dati disponibili. In altre parole, nella realtà i milionari sono più rari di quanto emerge dall'analisi di Pareto. Una conseguenza di questa legge di potenza è il principio di Pareto, detto anche regola 80/20. Si tratta di calcolare quanta percentuale della ricchezza totale è detenuta nelle mani della parte più benestante della società. Per qualunque legge di potenza con  $\alpha > 1$ , la mediana  $x_{1/2}$ , ovvero il valore che divide a metà la distribuzione, esiste ed è sempre definito. Si può

<sup>147</sup> B. B. MANDEBROT, *Il disordine dei mercati. Una visione frattale di rischio, rovina e redditività*, 2005, cap. 8.

immaginare che  $x_{1/2}$  divida la popolazione in una classe ricca e in una classe povera e si può calcolare la frazione della ricchezza totale posseduta dalla classe ricca. In generale, se si considera la frazione di individui con ricchezza personale superiore a  $x$  si può definire la frazione della ricchezza totale  $W(x)$  posseduta da questa porzione della popolazione<sup>148</sup>. Si può dimostrare che al diminuire del valore di  $\alpha$  aumenta lo squilibrio sociale. In particolare, si è stimato negli Stati Uniti un valore  $\alpha = 2,1$  che corrisponde ad una situazione in cui circa l'80% della ricchezza totale è detenuta dal 20% della popolazione<sup>149</sup>.

### 4.3 Le distribuzioni $L$ -stabili e le code spesse

L'importanza delle leggi di potenza nella descrizione dei fenomeni non può essere realmente compresa senza fare riferimento a quella che è probabilmente la principale peculiarità di questa categoria di distribuzioni di probabilità, il comportamento delle code. In precedenza, si è visto come, per la distribuzione normale, la maggior parte delle osservazioni siano concentrate attorno alla media, con circa il 99% di esse racchiuse in un intorno del valore centrale di  $3\sigma$ . Questo rende la probabilità di outlier molto piccola e in alcuni casi trascurabile, inoltre la persistenza dei dati attorno alla media rende le nuove osservazioni incapaci di spostarla su un diverso valore. Graficamente, la curva normale tende a zero molto velocemente, quando  $x \rightarrow \pm\infty$ , seguendo un decadimento di tipo esponenziale. Nelle leggi di potenza, invece, le code sono spesse e i valori

---

<sup>148</sup> Si può esprimere la frazione della ricchezza totale posseduta nelle mani delle persone con ricchezza superiore a  $x$  come:

$$W(x) = \frac{\int_x^\infty yp(y) dy}{\int_{x_{min}}^\infty yp(y) dy} = \frac{\frac{C}{\alpha-2}x^{2-\alpha}}{\frac{C}{\alpha-2}x_{min}^{2-\alpha}} = \left(\frac{x}{x_{min}}\right)^{-\alpha+2}.$$

<sup>149</sup> M. E. J. NEWMAN, *Power laws, Pareto distributions and Zipf's law*, 2006.

anomali sono molto più probabili. In particolare, una maggior area di probabilità è contenuta nelle code rispetto alla gaussiana. Per comprendere intuitivamente l'idea di code spesse è utile citare la metafora dell'arciere bendato, usata da Mandelbrot in riferimento a un tipo di distribuzione caratterizzata da questa proprietà, nota come distribuzione di Cauchy. Si immagini un arciere bendato davanti a un bersaglio posizionato su un muro di lunghezza infinita. L'arciere tira in modo casuale in tutte le direzioni. Se i lanci seguissero una distribuzione normale, la maggior parte di essi non sarebbe molto distante dal bersaglio, consentendo il calcolo della media e della varianza, e si stabilizzerebbero, dopo un sufficiente numero di tiri, su questi valori. Questa descrizione è evidentemente irrealistica, i tiri dell'arciere possono essere molto lontani dal bersaglio, a volte anche totalmente paralleli al muro, e non si stabilizzeranno mai attorno a un valore medio, in quanto un tiro particolarmente distante può spostare di molto la media ottenuta nei tiri precedenti, con la conseguenza che gli errori dell'arciere sono caratterizzati da valore atteso e varianza infinita. Questo comportamento è rappresentato nella distribuzione di Cauchy, che consente di descrivere un tipo di casualità che Mandelbrot definisce 'selvaggia', in contrapposizione a quella 'lieve' originata in un mondo gaussiano<sup>150</sup>. I tipi di distribuzioni di probabilità citati (Gauss, Zipf, Pareto, Cauchy) rappresentano casi particolari di una famiglia più ampia di distribuzioni, studiata dal matematico francese Paul Levy, nota come distribuzioni Levy  $\alpha$ -stabili o L-stabili. Si è visto che una variabile  $X$  è stabile se per ogni  $X_1, X_2$ , copie indipendenti di  $X$ , una combinazione lineare di tali variabili mantiene la stessa distribuzione di  $X$ , per parametri opportunamente scelti. Un modo equivalente,

---

<sup>150</sup> B. B. MANDEBROT, *Il disordine dei mercati. Una visione frattale di rischio, rovina e redditività*, 2005, cap. 2.

per formalizzare tale proprietà, consente che una variabile casuale  $X$  sia stabile se è solo se per ogni  $n > 1$  esistono costanti  $c_n > 0$  e  $d_n \in \mathbb{R}$  tali che:

$$X_1 + X_2 + \dots + X_n \stackrel{d}{=} c_n X + d_n,$$

con  $X_1, \dots, X_n$  copie indipendenti di  $X$ , se  $d_n = 0$  si parla di stabilità stretta<sup>151</sup>.

Si può dire che le distribuzioni L-stabili sono una classe di distribuzioni che ammettono code spesse e che soddisfano la proprietà di stabilità. Le distribuzioni L-stabili sono state per molto tempo ignorate a causa della difficoltà nell'applicazione concreta, in particolare non si può ricavare una forma esplicita della funzione di densità e della funzione di ripartizione, tranne per casi particolari, tuttavia esse possono essere interamente specificate tramite la funzione caratteristica. In modo più formale, la distribuzione della variabile casuale  $X$  è L-stabile se la sua funzione caratteristica  $\varphi_X$  si può esprimere, a meno di una traslazione e di una variazione di scala, come:

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

Si può notare che, a differenza del caso normale, le distribuzioni L-stabili sono descritte da quattro diversi parametri e vengono indicate spesso come  $S(\alpha, \beta, \gamma, \delta)$ , inoltre per  $\beta = 0$  la distribuzione è perfettamente simmetrica e la funzione caratteristica si semplifica diventando:

$$\varphi_X(t) = e^{-\gamma|t|^\alpha}.$$

Modificando questi parametri è possibile ottenere, come casi particolari, le distribuzioni studiate nei precedenti capitoli, per cui è importante analizzarli singolarmente:

---

<sup>151</sup> J.P. NOLAN, *Stable distributions models for heavy tailed data*, 2018.



- i.  $\alpha \in (0,2]$ : è il parametro di stabilità della distribuzione, detto anche esponente caratteristico, ed è lo stesso descritto nelle leggi di potenza. Misura la concentrazione della probabilità nelle code, più diminuisce più le code sono spesse;
- ii.  $\beta \in [-1,1]$ : è l'indice di asimmetria della distribuzione, misura quanto è sbilanciata la distribuzione;
- iii.  $\gamma > 0$ : è il parametro di scala, indica quanto è larga o stretta la distribuzione;
- iv.  $\delta \in \mathbb{R}$ : è il parametro di posizionamento, indica dove è situata la distribuzione rispetto al dominio.

$\alpha$  e  $\beta$  determinano la forma della distribuzione e sono quindi detti parametri di forma. Per  $\alpha < 2$  la varianza della distribuzione non è definita e per  $\alpha \leq 1$  non lo è la media, inoltre tutte le distribuzioni con  $\alpha < 2$  sono anche dette Pareto-stabili in quanto la decadenza delle code segue una distribuzione di Pareto<sup>152</sup>. Ponendo  $\alpha = 2$ ,  $\beta = 0$ ,  $\gamma = \frac{\sigma}{\sqrt{2}}$  e  $\delta = \mu$ , la variabile L-stabile  $X$  diventa una distribuzione normale,  $X \sim S\left(2, 0, \frac{\sigma}{\sqrt{2}}, \mu\right) = N(\mu, \sigma^2)$ , e si può calcolare la funzione di densità come:

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty.$$

Ponendo  $\alpha = 1$  e  $\beta = 0$  la variabile  $X$  assume la forma della distribuzione di Cauchy,  $X \sim S(1, 0, \gamma, \delta) = C(\gamma, \delta)$ , la cui densità è data da:

$$f_X(x) = \frac{1}{\pi} \frac{\gamma}{\gamma^2 + (x - \delta)^2}, \quad -\infty < x < \infty.$$

---

<sup>152</sup> B. B. MANDELBROT, *The variation of certain speculative prices*, 1963.

Infine, se si pone  $\alpha = \frac{1}{2}$  e  $\beta = 1$  la variabile  $X$  si dispone secondo una distribuzione di Levy,  $X \sim S\left(\frac{1}{2}, 1, \gamma, \delta\right) = L(\gamma, \delta)$ , descritta come:

$$f_X(x) = \sqrt{\frac{\gamma}{2\pi}} \frac{1}{(x - \delta)^{\frac{3}{2}}} e^{-\frac{\gamma}{2(x-\delta)}}, \quad \delta < x < \infty.$$

Questi tre esempi sono le uniche distribuzioni L-stabili per le quali è possibile ricavare una forma esplicita della funzione di densità, inoltre la distribuzione normale e quella di Cauchy sono perfettamente simmetriche, mentre quella di Levy è molto asimmetrica, verso il semiasse  $x > 0$ . La distribuzione di Levy presenta code più spesse sia della distribuzione di Cauchy che di quella gaussiana<sup>153</sup>.

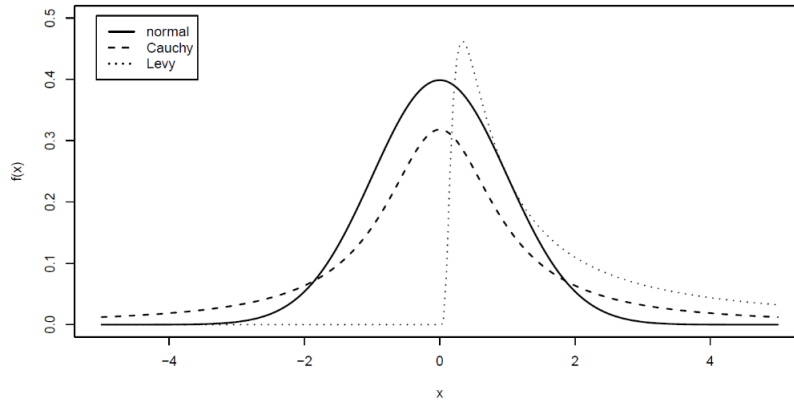


Fig. 30 Funzione di densità standardizzata  $N(0,1)$ ,  $C(0,1)$  e  $L(0,1)$ . [44]

Si può dimostrare che, data  $X \sim S(\alpha, \beta, \gamma, \delta)$  variabile casuale stabile con  $0 < \alpha < 2$ , per  $x \rightarrow \infty$  vale che:

$$F_X(x) = P(X > x) \sim \gamma^\alpha C_\alpha (1 + \beta) x^{-\alpha},$$

$$f_X(x) \sim \alpha \gamma^\alpha C_\alpha (1 + \beta) x^{-(\alpha+1)}.$$

La proprietà di riflessione afferma che,  $\forall \alpha, \beta$ ,  $X \sim S(\alpha, \beta)$ , vale:

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

<sup>153</sup> J. P. NOLAN, *Stable distributions models for heavy tailed data*, 2018.

quindi, per  $x \rightarrow \infty$ , la distribuzione cumulata della coda sinistra,  $F_X(-x) = P(X < -x)$ , e la funzione di densità,  $f_X(-x)$ , presentano lo stesso andamento asintotico della coda destra. Una conseguenza delle code spesse è che alcuni momenti della distribuzione risultano indefiniti, un modo per superare questo limite è considerare i momenti assoluti frazionari, definiti come:

$$\langle |x|^p \rangle = \int_{-\infty}^{\infty} |x|^p f_X(x) dx, \quad \text{con } p \in \mathbb{R}.$$

In particolare, nel caso di distribuzioni L-stabili con  $0 < \alpha < 2$ ,  $\langle |x|^p \rangle < \infty$  per  $0 < p < \alpha$  e  $\langle |x|^p \rangle = +\infty$  per  $p \geq \alpha$ . Un aspetto importante delle distribuzioni L-stabili riguarda il teorema del limite centrale, secondo cui la somma standardizzata di variabili aleatorie indipendenti e identicamente distribuite converge a una distribuzione gaussiana. L'applicabilità di questo noto teorema è ristretta solamente al caso di distribuzioni con varianza finita, tuttavia si può dimostrare una versione generalizzata.

*Teorema.* (Teorema generalizzato del limite centrale). Una variabile aleatoria  $Z$  è stabile con  $0 < \alpha \leq 2$  se e solo se esiste una successione di variabili aleatorie  $X_1, X_2, X_n$  tali che, date le costanti  $a_n > 0$  e  $b_n \in \mathbb{R}$ , vale la condizione:

$$a_n(X_1 + \dots + X_n) - b_n \stackrel{d}{=} Z.$$

Abbandonata l'ipotesi di varianza finita, l'unica distribuzione a cui può tendere una somma di variabili aleatorie identicamente distribuite è quella stabile, ovvero una variabile  $X$  che soddisfa la condizione appartiene al dominio di attrazione di  $Z$ , detto  $DA(Z)$ , che corrisponde a tutte le variabili casuali per cui  $Z$  è un attrattore della distribuzione<sup>154</sup>. Le uniche distribuzioni dotate di dominio di attrazione sono quelle L-stabili. Per applicazioni economiche, Mandelbrot ha

---

<sup>154</sup> J. P. NOLAN, *Stable distributions models for heavy tailed data*, 2018.

osservato che le distribuzioni che meglio si adattano sono quelle con  $1 < \alpha < 2$ , chiamate distribuzioni Levy-Pareto stabili.

#### *4.4 I processi anomali e i voli di Levy*

Nel primo capitolo si è analizzato il moto browniano da un punto di vista stocastico, tuttavia questa espressione è originariamente riferita al movimento dei granuli di polline. Il primo a formalizzare matematicamente questa dinamica irregolare è stato Albert Einstein, il quale nel 1905 ha ricavato l'equazione di Fourier per la diffusione del moto browniano:

$$\frac{\partial}{\partial t} p(x, t) = K \frac{\partial^2}{\partial x^2} p(x, t),$$

con  $K$  coefficiente di diffusione e  $p(x, t)$  densità di probabilità. Come ricavato nell'Appendice C, la soluzione generale della PDE è una gaussiana la cui varianza decade nel tempo in modo lineare, ovvero il momento secondo è dato da:

$$\langle x^2 \rangle = \int_{-\infty}^{\infty} x^2 p(x, t) dx = 2Kt.$$

Questa ipotesi di linearità è stata messa in discussione nel 1926 da Lewis Fry Richardson che, studiando fenomeni di diffusione in presenza di turbolenza, ha stimato una dipendenza cubica della varianza:

$$\langle x^2 \rangle \propto t^3.$$

Un processo di diffusione non lineare la cui varianza segue una legge di potenza è detto processo anomalo di diffusione, espresso come:

$$\langle x^2 \rangle \propto t^\alpha,$$

con  $\alpha$  esponente di diffusione della legge di potenza. Se  $0 < \alpha < 1$  si ha un processo di subdiffusione, se  $\alpha > 1$  il processo è detto di superdiffusione<sup>155</sup>. Fenomeni di diffusione anomala sono stati identificati, ad esempio, nella diffusione delle proteine all'interno cellule, nel flusso di particolari fluidi, nella diffusione della luce in particolari materiali e nelle dinamiche di migrazione. Il comportamento asintotico a una legge di potenza comporta numerose conseguenze, in particolare la varianza diverge, le distribuzioni presentano code spesse e il teorema del limite centrale non può essere applicato, nella versione classica. La distribuzione gaussiana risulta inefficace per descrivere l'andamento anomalo di questi processi ed è dunque necessario identificare una versione più generale del moto browniano. Una possibilità è utilizzare le distribuzioni L-stabili per modellizzare un diverso tipo di random walk continuo, noto come volo di Levy. I voli di Levy sono dei cammini isotropici con direzione casuale, in cui la distribuzione di probabilità dei passi è descritta dalla distribuzione di Levy. Le code spesse rendono la lunghezza degli incrementi incostante, con un'alta probabilità di passi ampi, inoltre, essendo infinita la varianza di ogni incremento, lo è anche quella di  $n$  incrementi, fornendo al volo una caratteristica frattale di invarianza di scala<sup>156</sup>. Nei voli di Levy si prende in esame solamente il punto iniziale e finale, con la conseguenza che il tempo è costante sia per grandi che per piccoli incrementi, perciò essi si distinguono dai cammini di Levy, nei quali è introdotto anche il concetto di velocità di percorrenza.

---

<sup>155</sup> R. METZLER, J. H. JEON, A. G. CHERSTVY, E. BARKAI, *Anomalous diffusion models and their properties: non-stationarity, non-ergodicity, and ageing at the centenary of single particle tracking*, 2014.

<sup>156</sup> J. P. BOUCHAUD, *Anomalous diffusion in disordered media: statistical mechanisms, models and physical applications*, 1990.

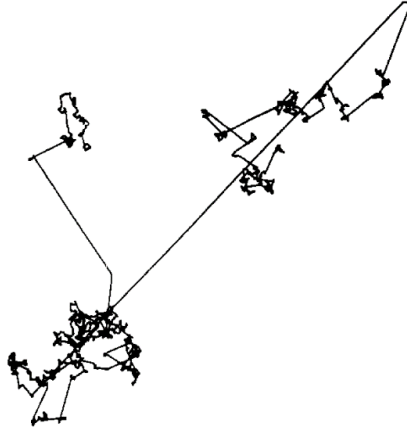


Fig. 31 Traiettoria di un volo di Levy. [43]

Supponiamo che la sequenza di passi  $\{x_i\}$  segua una distribuzione L-stabile simmetrica,  $L(a, x)$ , con funzione caratteristica  $\varphi_X(t) = e^{-a|t|^\alpha}$ . Dopo  $n$  incrementi indipendenti e identicamente distribuiti la posizione è data da  $x_N = \sum_{i=1}^N x_i$ , con equazione caratteristica  $\varphi_N(t) = e^{-aN|t|^\alpha}$ . Si può esprimere la funzione di densità di probabilità calcolando la trasformata inversa di Fourier della funzione caratteristica  $\mathcal{F}^{-1}(\varphi_N)$ :

$$p(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{itx} e^{-aN|t|^\alpha} dt.$$

Sostituendo  $t = VN^{-\frac{1}{\alpha}}$  e  $dt = dVN^{-\frac{1}{\alpha}}$ , si ottiene:

$$p(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{iV\left(\frac{x}{N^{\frac{1}{\alpha}}}\right)} e^{-aN|V|^\alpha N^{-1}} \frac{dV}{N^{\frac{1}{\alpha}}} = \frac{1}{N^{\frac{1}{\alpha}}} \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{iV\left(\frac{x}{N^{\frac{1}{\alpha}}}\right)} e^{-a|V|^\alpha} dV.$$

L'espressione può quindi essere scritta come:

$$p(x) = \frac{1}{N^{\frac{1}{\alpha}}} L\left(a, \frac{x}{N^{\frac{1}{\alpha}}}\right).$$

La funzione di densità di  $n$  incrementi è anch'essa una distribuzione L-stabile con varianza  $\langle x^2 \rangle \sim N^{\frac{2}{\alpha}}$ . Per  $\alpha < 2$  non si ha più la relazione lineare gaussiana, quindi il volo di Levy, così descritto, è un processo anomalo di

superdiffusione<sup>157</sup>. L'idea di varianza infinita può non essere riscontrata in applicazioni pratiche, anche se esse non si accordano con il modello normale. Una delle soluzioni consiste nel definire i voli troncati di Levy attraverso l'introduzione di un cutoff, che rende finita la varianza. In particolare, per elevati valori, il processo converge lentamente a una gaussiana, ma è approssimabile con una legge di potenza, per valori più bassi. Il teorema del limite centrale rimane valido e la convergenza al caso normale può essere tanto lenta da non venire osservata in applicazioni sperimentali<sup>158</sup>. I voli di Levy sono un modo per generalizzare il moto browniano, con incrementi non gaussiani.

#### *4.5 Le variazioni del cotone e l'invarianza di scala*

Le leggi di potenza rappresentano un nuovo modo di interpretare i fenomeni, eliminando l'ipotesi di normalità che caratterizza la teoria finanziaria classica, e le distribuzioni L-stabili forniscono una nuova gamma di strumenti matematici per modellizzare tali fenomeni. Quanto analizzato in precedenza consente di ritornare sull'esempio iniziale, delle variazioni anomale del cotone. In particolare, si è ora in grado di cogliere le ragioni della somiglianza tra il grafico dei redditi e il digramma delle variazioni dei prezzi, sulla lavagna di Houthakker, in quanto, le due distribuzioni rientrano all'interno della classe delle leggi L-stabili. Per confermare l'esistenza di una legge di potenza per il cotone, si può costruire un grafico bilogaritmico nel quale si raffigurano sulle ordinate le frequenze e sulle ascisse l'ampiezza delle variazioni.

---

<sup>157</sup> M. BAZANT, 18.366, *Random Walks and Diffusion*, Fall 2006, Massachusetts Institute of Technology: MIT.

<sup>158</sup> A. V. CHECHKIN, R. METZLER, J. KLAFTER, V. Y. GONCHAR, *Introduction to the theory of Levy flights*, 2008.

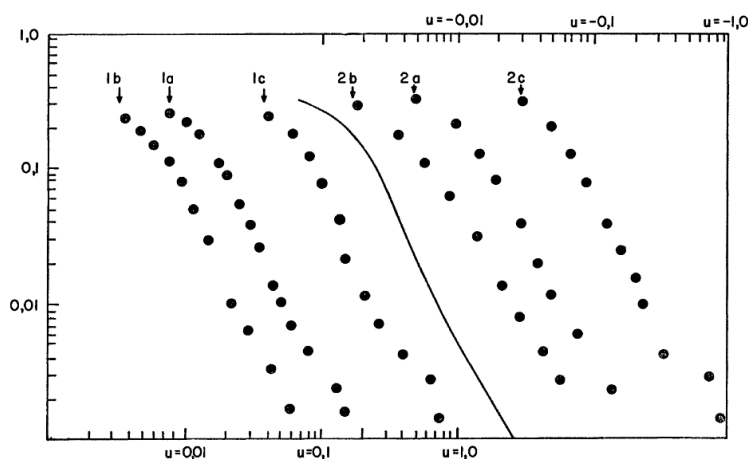


Fig. 32 Grafico delle variazioni del cotone. [15]

Nella figura le fluttuazioni sono raggruppate per periodo, in modo da evidenziare le scale temporali. *a* indica le variazioni giornaliere avvenute tra il 1900 e il 1945, *b* raffigura le variazioni del cotone tra il 1944 e il 1958 e, infine, *c* indica le fluttuazioni giornaliere tra il 1888 e il 1940. Per chiarezza espositiva sono stati separati gli incrementi positivi (*2a, 2b, 2c*) da quelli negativi (*1a, 1b, 1c*), inoltre la linea continua rappresenta una generica distribuzione L-stabile con  $\alpha = 1,7$ ,  $\beta = 0$  e  $\delta = 0$ . In verità, si può notare che le variazioni negative sono più numerose di quelle positive e ciò suggerisce che la distribuzione non è perfettamente simmetrica ma è caratterizzata da un valore di  $\beta$  lievemente negativo<sup>159</sup>. Risulta evidente l'andamento lineare delle variazioni che identifica una legge di potenza, inoltre i diversi periodi sono caratterizzati dalla stessa pendenza, ovvero la forma della distribuzione del cotone non varia in base alla scala temporale. In altre parole, Mandelbrot ha identificato uno schema frattale all'interno dei mercati finanziari, attraverso l'invarianza di scala delle fluttuazioni rispetto al tempo. All'inizio del capitolo si è mostrato come sia possibile realizzare un grafico frattale, che incorpori al suo interno le ipotesi del moto browniano. In modo analogo, è possibile

<sup>159</sup> B. B. MANDELBROT, *The variation of certain speculative prices*, 1963.



costruire un grafico finanziario frattale che segue una legge di potenza con code spesse, tramite l'introduzione di una discontinuità. L'iniziatore è costituito, come nel grafico browniano, dalla linea di tendenza che congiunge i due vertici opposti del rettangolo. Il generatore è formato, in questo caso, da cinque intervalli uguali inclinati, a cui si aggiungono due discontinuità verticali, in modo da formare, in seguito al processo di iterazione, una figura più complessa, caratterizzata da invarianza di scala, e nella quale sono possibili variazioni di elevata entità, simili a quelle riscontrate nell'analisi dei prezzi del cotone<sup>160</sup>.

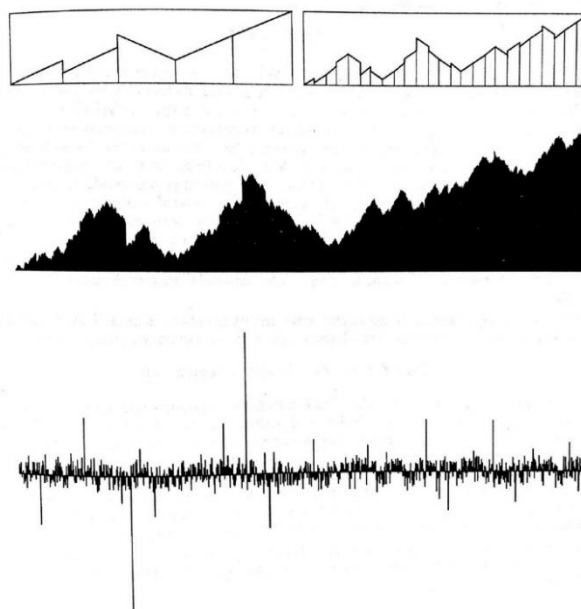


Fig. 33 Cartone frattale con discontinuità. [10]

#### 4.6 La dipendenza a lungo termine e l'esponente di Hurst

Per introdurre la seconda specificazione del modello di Mandelbrot dei mercati finanziari si può fare riferimento alle proprietà, più volte citate, del moto

<sup>160</sup> B. B. MANDEBROT, *Il disordine dei mercati. Una visione frattale di rischio, rovina e redditività*, 2005, cap. 8.

browniano. L'analisi sulle leggi di potenza e sulle distribuzioni stabili ha come risultato quello di invalidare l'ipotesi di normalità delle variazioni dei prezzi, mostrando come esse siano molto più imprevedibili di quanto mostri la curva a campana. In questa sezione, il focus è invece posto sulla contraddizione della seconda proprietà del modello di Bachelier, ovvero l'indipendenza markoviana delle oscillazioni. La formulazione di questo secondo modello si può far risalire al 1963, quando Mandelbrot ha iniziato a interessarsi al lavoro dell'idrologo britannico Harold Edwin Hurst, noto in quel periodo per aver risolto un problema apparentemente molto distante dal mondo dell'economia. Nel 1906, Hurst viene mandato in Egitto, nell'ambito di un progetto noto come 'century storage', per stimare la volatilità delle piene del Nilo, allo scopo di accumulare più acqua possibile per la siccità. Costruire un modello per il controllo delle variazioni annuali del fiume era un problema di notevole complessità, a causa dell'ampio range di variazione registrato negli anni, con una portata che poteva variare tra i 151 miliardi di  $m^3$ , nel 1878-79, e i 42 miliardi di  $m^3$ , registrati nel 1913-14. Gli ingegneri avevano sviluppato varie ipotesi per l'altezza delle dighe, assumendo che le variazioni del livello della piena potessero accordarsi al modello di Bachelier, con deviazione pari a  $n^{\frac{1}{2}}$ <sup>161</sup>. Studiando le registrazioni passate, Hurst ha notato che non è solo l'entità delle variazioni ad essere rilevante, ma lo è anche la loro sequenza. In particolare, i valori si adattavano al modello normale solo se considerati singolarmente, ignorando quindi la sequenza di successione<sup>162</sup>. Introducendo l'esistenza di una dipendenza significativa tra i livelli dei vari anni, il range di variazione cresce più

---

<sup>161</sup> Per un moto browniano classico  $B(t)$  la deviazione standard dell'incremento  $B(t + T) - B(t)$ , con  $T > 0$ , è pari a  $T^{\frac{1}{2}}$ .

<sup>162</sup> B. B. MANDEBROT, *Il disordine dei mercati. Una visione frattale di rischio, rovina e redditività*, 2005, cap. 9.

velocemente di quanto previsto dal modello di Bachelier e l'altezza delle dighe avrebbe dovuto essere maggiore. Per descrivere questo effetto Hurst ha elaborato una formula che consente di calcolare il campo di variazione  $R$  come:

$$\log\left(\frac{R}{\sigma}\right) = H \log\left(\frac{N}{2}\right),$$

con  $\sigma$  deviazione standard,  $N$  numero di anni e  $H$  coefficiente di Hurst. Eliminando i logaritmi, si può scrivere:

$$R = \sigma \left(\frac{N}{2}\right)^H.$$

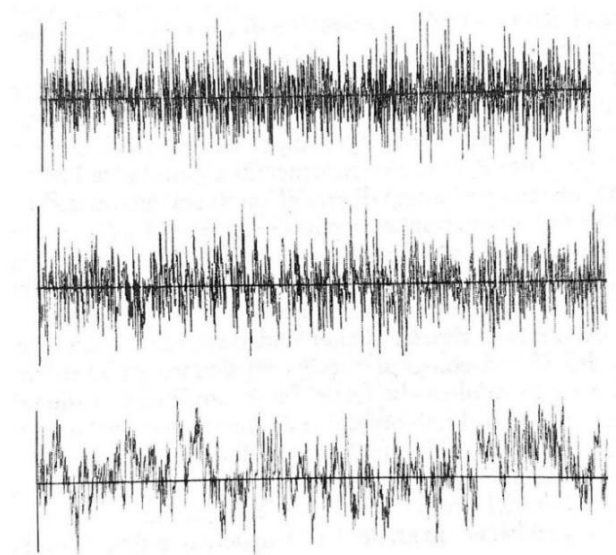
L'esponente della legge di potenza è il parametro fondamentale che quantifica il grado di dipendenza tra le registrazioni dei vari anni ed è stato stimato da Hurst pari a 0,73. La formula di Hurst si può applicare anche ad altri fenomeni, come ad esempio l'andamento delle precipitazioni annuali di New York, la crescita degli anelli nei tronchi degli alberi o la stratificazione dell'argilla sul letto dei fiumi. Il parametro  $H$  può assumere valori reali nell'intervallo  $(0,1)$  ed è stato usato da Mandelbrot per sviluppare una generalizzazione del moto browniano, nota come moto browniano frazionario, che permette di studiare fenomeni caratterizzati da memoria a lungo termine. In termini più formali, si assegna la proprietà di dipendenza a lungo termine a un processo stazionario non markoviano con varianza finita, in cui la funzione di autocorrelazione diverge seguendo una legge di potenza<sup>163</sup>. In particolare, il valore che l'esponente di Hurst assume determina il tipo di persistenza che presentano le variazioni della serie:

---

<sup>163</sup> T. GRAVES, R. B. GRAMACY, N. W. WATKINS, C. L. E. FRANZKE, *A brief history of long memory: Hurst, Mandelbrot and the road to ARFIMA*, 2016.

- i.  $H \in \left(0, \frac{1}{2}\right)$ : serie antipersistente, i valori non si allontanano molto da quelli iniziali, dopo un incremento è più probabile assistere a un decremento, la serie diventa più volatile più  $H$  è vicino a 0;
- ii.  $H = \frac{1}{2}$ : le osservazioni sono indipendenti, ciascun cambiamento non influenza o è influenzato dalle altre variazioni, è il caso del moto browniano standard;
- iii.  $H \in \left(\frac{1}{2}, 1\right)$ : serie persistente, le osservazioni tendono a seguire lo stesso andamento in modo da creare dei cicli finiti non periodici. Se si è registrato un incremento c'è una maggiore probabilità di registrarne un altro nella variazione successiva, questa probabilità è tanto maggiore più  $H$  tende a 1.

Sono stati stimati valori di  $H$  persistenti per diverse aziende americane, come per esempio 0,75 per Apple, 0,73 per Xerox o 0,72 per le variazioni di IBM.



*Fig. 34* Serie antipersistente  $H = 0,1$  (in alto), serie browniana  $H = 0,5$  (in mezzo) e serie persistente  $H = 0,9$  (in basso). [10]

## 4.7 Il moto browniano frazionario

Dopo aver spiegato l'importanza degli studi svolti da Hurst e le implicazioni della dipendenza a lungo termine, sfruttando il parametro  $H$  si può costruire un processo stocastico gaussiano che incorpori questa proprietà. In realtà, nel 1941, in seguito a studi su fenomeni di turbolenza, il matematico russo Andrej Kolmogorov è stato il primo a definire una teoria dei processi gaussiani continui a incrementi stazionari, che godevano della proprietà di auto-similarità, denominandoli spirali di Wiener. Mandelbrot fa uso di questi strumenti per definire il moto browniano frazionario, nello studio delle serie storiche.

*Definizione.* (Processo auto-simile). Un processo stocastico continuo  $X = \{X_t\}_{t \geq 0}$  è detto auto-simile, e si indica con 'H-ss', se per ogni  $a > 0$  esiste  $b > 0$  tale che:

$$\Phi(X_{at}; t \geq 0) = \Phi(bX_t; t \geq 0),$$

ovvero il processo stocastico presenta un'invarianza nella scala delle grandezze, in modo che un cambiamento nella scala temporale ( $t \rightarrow at$ ) abbia lo stesso effetto di uno che avviene nella scala degli stati ( $x \rightarrow bx$ ). Inoltre, se  $b = a^H$  si ottiene che:

$$\Phi(X_{at}; t \geq 0) = \Phi(a^H X_t; t \geq 0).$$

$X$  è, in questo caso, un processo auto-simile con coefficiente di Hurst  $H$ <sup>164</sup>.

*Definizione.* (Processo a incrementi stazionari). Un processo stocastico  $X = \{X_t\}_{t \geq 0}$  è detto a incrementi stazionari, e si scrive con 'si', se per ogni  $h > 0$  vale che:

$$\Phi(X_{t+h} - X_h; t \geq 0) = \Phi(X_t - X_0; t \geq 0).$$

---

<sup>164</sup> T. DIEKER, *Simulation of fractional brownian motion*, 2004, cap.1.

Nel caso queste due proprietà coesistessero, si può definire una nuova classe di processi stocastici.

*Definizione.* (Processi H-sssi). Dato  $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$  uno spazio di probabilità con filtrazione  $\mathcal{F}_t$ , un processo stocastico  $X = \{X_t\}_{t \geq 0}$ ,  $\mathcal{F}_t$ -misurabile, è detto processo H-sssi se è auto-simile con indice  $H$  ed ha incrementi stazionari.

A questo punto risulta facile definire il moto browniano frazionario come un processo H-sssi.

*Definizione.* (Moto browniano frazionario). Dato  $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$  uno spazio di probabilità con filtrazione  $\mathcal{F}_t$ , un processo stocastico continuo normale H-sssi  $B^H = (B_t^H)_{t \in I}$ ,  $I \subset \mathbb{R}$ , con  $0 < H < 1$ , adattato alla filtrazione  $\mathcal{F}_t$ , si dice moto browniano frazionario.

Il processo, così definito, gode di interessanti proprietà matematiche che lo caratterizzano<sup>165</sup>:

- i.  $B_t^H = 0$ , il processo inizia da un'origine;
- ii. Se  $H \neq 1$ ,  $\forall t \geq 0$  vale  $E(B_t^H) = 0$ ;
- iii.  $B^H(-t) = -B^H(t)$ , poiché per la stazionarietà degli incrementi si ha  

$$B^H(-t) = B^H(-t) - B^H(0) = B^H(0) - B^H(t) = -B^H(t)$$
- iv. La varianza è data da  $E(B_t^{2H}) = |t|^{2H} \sigma^2$ , in quanto, sfruttando la proprietà di auto-similarità, si può scrivere che  $E(B_t^{2H}) = E(B_{|t|sgn(t)}^{2H}) = |t|^{2H} E(B_1^H) = |t|^{2H} \sigma^2$ ,<sup>166</sup>
- v.  $E[(B_t^H - B_s^H)^2] = E(B_{t-s}^{2H}) = |t - s|^{2H} \sigma^2$ ;

<sup>165</sup> J. BERAN, *Statistics for long-memory processes*, 1994, cap. 2.

<sup>166</sup>  $sgn(x) := \begin{cases} 1 & \text{se } x > 0 \\ 0 & \text{se } x = 0 \\ -1 & \text{se } x < 0 \end{cases}$ , è una funzione a tratti che estrae il segno di un numero reale. Ogni numero reale si può scrivere come  $x = |x|sgn(x)$ .

vi. La funzione di autocovarianza<sup>167</sup> è data da:

$$\begin{aligned}\Gamma_{t,s}^H &= E(B_t^H B_s^H) = \frac{1}{2} [E(B_t^{2H}) + E(B_s^{2H}) - E((B_t^H - B_s^H)^2)] = \\ &= \frac{\sigma^2}{2} (|t|^{2H} + |s|^{2H} - |t - s|^{2H}).\end{aligned}$$

Se  $\sigma^2 = 1$  il processo è detto standard. Per un valore di  $H = \frac{1}{2}$  si ottiene il moto browniano classico, caratterizzato da incrementi normali indipendenti, e la funzione di autocovarianza diventa:

$$\Gamma_{t,s}^{\frac{1}{2}} = \frac{1}{2} (|t| + |s| - |t - s|) = \min(|s|, |t|), \quad \forall t, s \geq 0.$$

Il caso  $H = 1$  solitamente non viene considerato in quanto origina una semplice retta, inoltre, da queste proprietà, è facile dimostrare che  $B^H \sim N(0, \sigma^2 |t|^{2H})$  è un processo auto-simile con indice  $H$ :

$$\begin{aligned}E(B_{at}^H B_{as}^H) &= \frac{1}{2} [(at)^{2H} + (as)^{2H} - |at - as|^{2H}] \\ &= \frac{1}{2} a^{2H} (t^{2H} + s^{2H} - |t - s|^{2H}) = a^{2H} E(B_t^H B_s^H) \\ &= E[(a^{2H} B_t^H)(a^{2H} B_s^H)].\end{aligned}$$

Il moto browniano frazionario non è differenziabile nel suo dominio, tuttavia le sue traiettorie godono localmente di Hölder continuità<sup>168</sup>. Se si considera un moto browniano frazionario  $B^H(t)$ , si può dimostrare che esiste una relazione, tra la dimensione di Hausdorff e la dimensione di Minkowski-Bouligand del suo grafico e l'esponente di Hurst. In particolare, dato:

$$G_B = \{(t, Y) \in \mathbb{R}^2 | Y = B^H(t), t \in [0, +\infty]\},$$

---

<sup>167</sup> L'espressione della funzione di autocovarianza si può derivare ricordando la formula della varianza della somma di due variabili aleatorie:

$$\text{var}(X_1 + X_2) = \text{var}(X_1) + \text{var}(X_2) - 2\text{Cov}(X_1 X_2).$$

<sup>168</sup> Una funzione  $f: \mathbb{R} \rightarrow \mathbb{R}$  si dice Hölder continua di ordine  $\alpha$ , o  $\alpha$ -höllderiana, se esiste una costante  $M > 0$  tale che:

$$|f(t) - f(s)| \leq M|t - s|^\alpha.$$

La proprietà rappresenta una generalizzazione della condizione di Lipschitz di crescita limitata, che si ottiene per  $\alpha = 1$ .

vale che:

$$\dim_H(G_B) = \dim_B(G_B) = 2 - H.$$

Essendo  $H \in (0,1)$  la dimensione frattale del grafico  $G_B$  è limitata, ovvero:

$$1 < \dim_H(G_B) = \dim_B(G_B) < 2.$$

In vari settori applicativi, molto spesso, è utile fare riferimento a una versione discreta di moto browniano frazionario, studiandone il processo degli incrementi. In particolare, se si considera la quantità:

$$h_t = B_t^H - B_{t-1}^H,$$

con  $B_t^H$  moto browniano frazionario, la sequenza stazionaria  $\{h_t\}_{t \geq 1}$  è detta rumore gaussiano frazionario con indice di Hurst,  $H \in (0,1)$ . Esso è caratterizzato, per ogni  $t$ , da  $E(h_t) = 0$  e  $E(h_t^2) = 1$ , inoltre si può derivare la funzione di autocovarianza tra due elementi della sequenza come:

$$\begin{aligned} \Gamma_k^H &= Cov(h_t h_{t+k}) = Cov(h_1 h_{1+k}) = Cov(B_1^H - B_0^H, B_{k+1}^H - B_k^H) \\ &= E(B_1^H, (B_{k+1}^H - B_k^H)) = E(B_1^H)E(B_{k+1}^H - B_k^H) \\ &= E(B_{k+1}^H B_1^H) - E(B_k^H B_1^H) = \Gamma_{k+1,1}^H - \Gamma_{k,1}^H \\ &= \frac{1}{2} \sigma^2 [(k+1)^{2H} + 1^{2H} - k^{2H} - 1^{2H} - k^{2H} + (k-1)^{2H}] \\ &= \frac{\sigma^2}{2} [(k+1)^{2H} - 2k^{2H} + (k-1)^{2H}].^{169} \end{aligned}$$

Per  $H = \frac{1}{2}$ , se  $t \neq 0$ , si ha  $\Gamma_k^H = 0$ , il processo  $h_t$  è completamente casuale ed è chiamato rumore bianco. Per  $\frac{1}{2} < H < 1$ , la funzione di autocovarianza è positiva,  $\Gamma_k^H > 0$ , e il processo è detto rumore nero, infine, per  $0 < H < \frac{1}{2}$ ,  $\Gamma_k^H < 0$ , la sequenza  $\{h_t\}_{t \geq 1}$  è detta rumore rosa. Sfruttando il rumore frazionario e il comportamento della funzione di autocorrelazione, è possibile definire

---

<sup>169</sup> Si può ottenere la formula della correlazione dividendo per  $\sigma^2$ :

$$Corr(h_t h_{t+k}) = \frac{1}{2} [(k+1)^{2H} - 2k^{2H} + (k-1)^{2H}]$$



matematicamente il concetto di dipendenza a lungo termine, come misura del decadimento della correlazione tra due osservazioni del processo.

*Definizione.* (Dipendenza a lungo termine). Dato un processo stocastico stazionario  $\{X_k\}_{k \geq 0}$ , esso gode di memoria o dipendenza a lungo termine se esiste un numero reale  $\alpha$  e una costante  $c > 0$  tale che:

$$\lim_{k \rightarrow \infty} \frac{\Gamma_k}{ck^{-\alpha}} = 1.$$

In altre parole, se la correlazione tra le osservazioni di un processo decade in modo asintotico a una legge di potenza di esponente  $\alpha$ , il processo è caratterizzato da una dipendenza a lungo termine:

$$\Gamma_k = \text{Corr}(X_t X_{t+k}) \sim ck^{-\alpha}.$$

Si può dimostrare che per il rumore gaussiano frazionario vale che<sup>170</sup>:

$$\Gamma_k^H \sim H(2H - 1)k^{2H-2},$$

ovvero si ha:

$$\lim_{k \rightarrow \infty} \frac{\Gamma_k^H}{H(2H - 1)k^{2H-2}} = 1.$$

Si nota che se  $H \in (\frac{1}{2}, 1)$ , per  $k \rightarrow \infty$ , la correlazione tende a zero così lentamente che la serie della funzione di autocorrelazione diverge<sup>171</sup>:

$$\sum_{k=1}^{\infty} \Gamma_k^H = \infty.$$

---

<sup>170</sup> L'equazione di autocorrelazione si può scrivere, raccogliendo  $k^{2H}$ , come  $\Gamma_k^H = \frac{1}{2}k^{2H}g(k^{-1})$ , con  $g(x) = (1+x)^{2H} - 2 + (1-x)^{2H}$ . L'espansione di Taylor nell'origine di  $g(x)$  è data da  $g(x) = g(0) + g'(0)(x-0) - \frac{1}{2}g''(0)(x-0)^2 = (2H-2H)x + \frac{1}{2}[2H(2H-1) + 2H(2H-1)]x^2 = 2H(2H-1)x^2$ . Per  $k \rightarrow \infty$ , si ottiene  $\Gamma_k^H \sim H(2H-1)k^{2H-2}$ .

<sup>171</sup> Si osserva che la serie può essere vista come una serie armonica generalizzata  $\sum_n \frac{1}{n^\alpha}$ :

$$\sum_k H(2H-1)k^{2H-2} = \sum_k H(2H-1) \frac{1}{k^{2-2H}}.$$

Se  $H > \frac{1}{2}$ ,  $2-2H < 1$ , la serie diverge positivamente, se  $H < \frac{1}{2}$ ,  $2-2H > 1$ , la serie converge.

Se  $H < \frac{1}{2}$ , le osservazioni successive presentano una memoria a breve termine, in quanto la serie converge<sup>172</sup>:

$$\sum_{k=1}^{\infty} |\Gamma_k^H| < \infty.$$

Si è quindi definito un nuovo processo stocastico gaussiano stazionario autosimile con varianza finita e indice di Hurst  $H$ , caratterizzato da dipendenza a lungo termine tra le osservazioni, per  $H > \frac{1}{2}$ .

#### 4.8 L'effetto Noè e l'effetto Giuseppe

La dipendenza a lungo termine rappresenta, come visto, un elemento fondamentale nella descrizione dei mercati finanziari. Per darne una rappresentazione più intuitiva si può, nuovamente, fare riferimento ai cartoni finanziari. Per costruire il cartone di Bachelier è necessario che l'altezza del generatore sia pari alla radice quadrata della lunghezza, ovvero bisogna elevare a  $\frac{1}{2}$ . Immaginando che tale parametro corrisponda all'esponente di Hurst, si possono costruire dei cartoni, che incorporino il concetto di memoria di lungo termine, variandone il valore.

---

<sup>172</sup> F. BIAGINI, Y. HU, B. ØKSENDAL, T. ZHANG, *Stochastic calculus for fractional brownian motion and applications*, 2008, cap. 1.

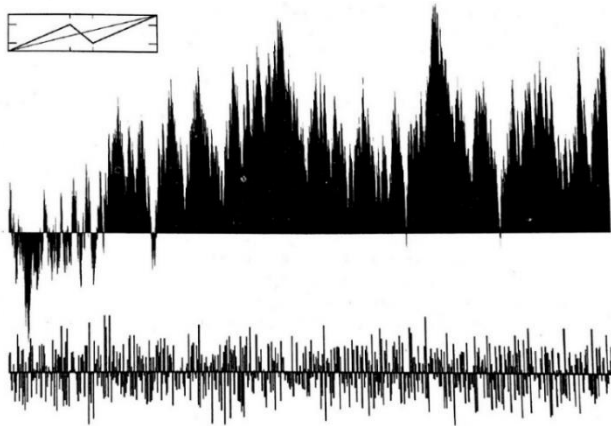


Fig. 35 Cartone frattale anti-persistente  $H < \frac{1}{2}$ . [10]

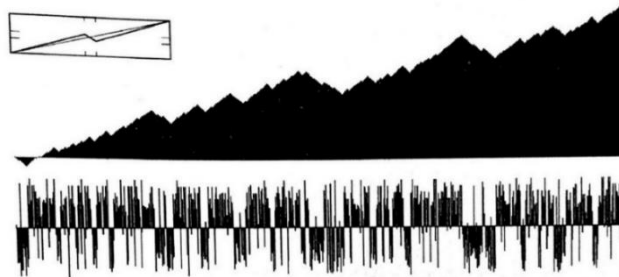


Fig. 36 Cartone frattale persistente  $H > \frac{1}{2}$ . [10]

La prima figura mostra il caso di un generatore con  $H$  inferiore a quello di Bachelier, in cui la serie ha un andamento antipersistente, con ampie oscillazioni. Il secondo cartone, invece, mostra il caso opposto, persistente<sup>173</sup>. Si possono ora riassumere le due forme di casualità ‘selvaggia’ che caratterizzano i mercati, le code spesse e la dipendenza a lungo termine. Due tratti diversi, ma legati tra loro, riassunti da due diversi parametri,  $\alpha$  e  $H$ ,

<sup>173</sup> B. B. MANDEBROT, *Il disordine dei mercati. Una visione frattale di rischio, rovina e redditività*, 2005, cap. 9.

denominati da Mandelbrot rispettivamente effetto Noè ed effetto Giuseppe<sup>174</sup>. L'episodio biblico di Noè racchiude bene l'idea di discontinuità e variazioni estese e si può immaginare una similitudine tra il diluvio e le turbolenze finanziarie. Il secondo effetto deve il nome ad un racconto della Genesi in cui Giuseppe, uno schiavo ebreo, profetizza al faraone, con successo, l'imminente arrivo di sette anni di prosperità seguiti da sette anni di carestia. La similitudine, in questo caso, riguarda l'esistenza di pseudo-cicli nei mercati finanziari, causati dalla memoria delle oscillazioni. In generale, l'effetto Giuseppe dipende dall'ordine delle variazioni e l'effetto Noè dipende dall'ampiezza. Questi due effetti sono collegati e possono presentarsi insieme, ad esempio nelle bolle finanziarie. Si può immaginare la bolla come una pseudo-tendenza che si interrompe con una forte discontinuità, cioè come un effetto Noè generato da un effetto Giuseppe. Tenere conto di queste due evidenze permette di avvicinarsi sempre di più a una migliore comprensione dei mercati finanziari<sup>175</sup>.

#### *4.9 Il modello multifrattale dei rendimenti delle attività*

I progressi ottenuti da Mandelbrot nello studio dei mercati finanziari sono sempre stati originati da eventi o indagini particolari. L'effetto Noè ha avuto origine dall'osservazione dei diagrammi delle variazioni del cotone e l'effetto Giuseppe è stato formalizzato in seguito alle scoperte di Hurst sulle piene del Nilo. L'ultima rappresentazione delle variazioni dei prezzi trae origine dallo studio dell'intermittenza della turbolenza nelle raffiche di vento e consente di

---

<sup>174</sup> B. B. MANDELBROT, *The fractal geometry of the nature*, 1982, cap. 27.

<sup>175</sup> B. B. MANDELBROT, *Il disordine dei mercati. Una visione frattale di rischio, rovina e redditività*, 2005, cap. 10.

combinare, in un unico modello, i due effetti precedenti. In particolare, gli strumenti matematici utilizzati sono una generalizzazione dei frattali già introdotta nei capitoli precedenti, i multifrattali. Un frattale è un oggetto geometrico, le cui parti presentano una particolare invarianza a riduzioni effettuate tramite un fattore di scala, i multifrattali conservano la stessa proprietà, ma ammettono diversi fattori di scala. I multifrattali premettono di realizzare un modello generalizzato e realistico delle oscillazioni nei mercati, che prende il nome di Modello Multifrattale dei Rendimenti delle Attività (MMAR) ed è analizzato, nei successivi paragrafi, dapprima a livello intuitivo, tramite l'utilizzo dei cartoni finanziari, e successivamente da un punto di vista più matematico e formale.

#### *4.9.1 Il tempo di contrattazione e i cartoni multifrattali*

La principale variazione rispetto ai modelli precedenti, che giustifica l'utilizzo del termine multifrattale, è la nozione di tempo di contrattazione. L'idea si basa sul fatto che l'attività all'interno dei mercati non è costante, ma attraversa fasi più veloci e fasi più lente. Questo suggerisce che il tempo all'interno di essi non sia uniforme e lineare in un senso fisico, ma sia caratterizzato da una forma di relatività. In altre parole, il tempo non scorre sempre in modo lineare ma varia in base all'attività che caratterizza il mercato, intesa in modo intuitivo come volume delle contrattazioni. Si assiste ad una dilatazione temporale durante le fasi di elevata attività e ad una contrazione del tempo durante le fasi di volume ridotto<sup>176</sup>. Questa nuova forma di tempo non lineare è chiamata tempo di contrattazione ed è costruita tramite un processo matematico detto cascata moltiplicativa. All'inizio del capitolo si è accennato a come realizzare un generatore multifrattale, contraendo l'asse temporale e lasciando invariato

---

<sup>176</sup> B. B. MANDELBROT, *Survey of multifractality in finance*, 1999.

l'asse del prezzo. Questa idea può essere approfondita analizzando la costruzione del generatore multifrattale, come composizione di due diversi generatori.

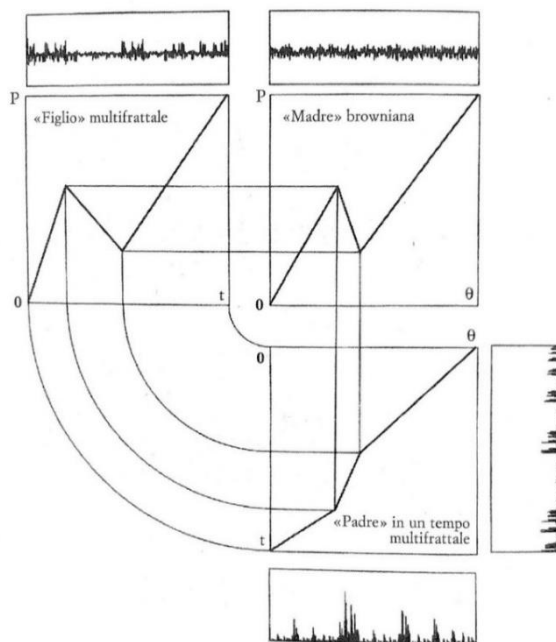


Fig. 37 Costruzione del generatore multifrattale. [10]

Denotando gli assi con  $t$  per il tempo classico in senso fisico,  $P$  per il prezzo e  $\theta$  per il tempo di contrattazione, si può immaginare di avere un generatore ‘madre’ browniano nel tempo classico e un generatore ‘padre’, che trasforma il tempo del generatore ‘madre’ secondo una nuova scala temporale caratterizzata, come si può osservare dai grafici delle variazioni in basso a destra, da elevata irregolarità. Il generatore ‘figlio’, che prende il tempo di contrattazione dal padre e le informazioni dei prezzi dalla madre, ha una natura multifrattale e, tramite interpolazione, permette di realizzare grafici finanziari verosimili<sup>177</sup>. Spesso si inserisce nel processo di iterazione una componente casuale, notando come le tre parti del generatore siano componibili secondo sei permutazioni

<sup>177</sup> B. B. MANDEBROT, *Il disordine dei mercati. Una visione frattale di rischio, rovina e redditività*, 2005, cap. 11.

diverse. Selezionandone una casualmente, ad ogni ripetizione, si può aumentare il grado di realismo del grafico finale.

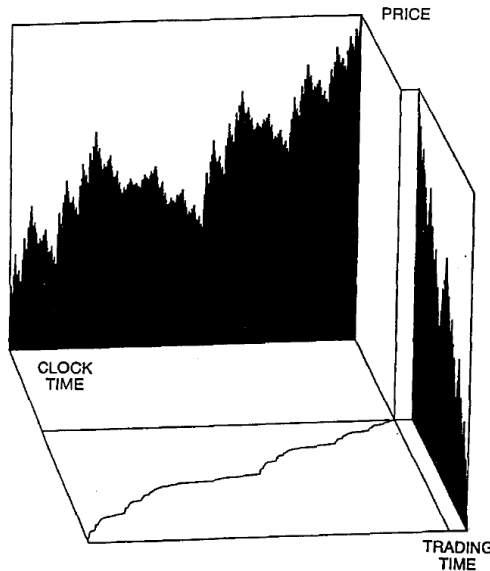


Fig. 38 Cubo del mercato multifrattale. [13]

Dal metodo di costruzione del generatore si può dedurre che il prezzo è una funzione del tempo di contrattazione che a sua volta è una funzione del tempo classico, secondo un processo composto.

#### *4.9.2 Le misure e i processi multifrattali*

Per approfondire il modello multifrattale è necessario introdurre alcuni argomenti preliminari, allo scopo di fornire un'adeguata cornice teorica al concetto di multifrattale. Essi possono essere definiti facendo uso di diverse costruzioni matematiche, in particolare si sfrutteranno le proprietà di scala dei momenti, per una definizione tramite il comportamento globale, e gli esponenti di Hölder, per analizzarne il comportamento locale. Il primo approccio permette di definire la multifrattalità per le misure ed estenderla ai processi stocastici, tramite la relazione di scala. La multifrattalità della misura è originata da un particolare processo di costruzione, noto come cascata moltiplicativa. Intuitivamente, si può immaginare, da una mappa geografica, di definire una

misura  $\mu$  della quantità di acqua nel sottosuolo. Ad ogni sottoinsieme  $S$  si attribuisce dunque una quantità  $\mu(S)$ . Se la distribuzione dell'acqua non è omogenea, dividendo in due parti la mappa,  $S_1$  e  $S_2$ , si ha che  $\mu(S_1) \neq \mu(S_2)$ . Continuando il processo di divisione, la misura continua a differire, in quanto alcune zone contengono più acqua di altre, e questa irregolarità si mantiene modificando la dimensione dei sottoinsiemi. La misura  $\mu$ , originata da questo processo moltiplicativo, è detta multifrattale<sup>178</sup>. Il più semplice esempio di multifrattale è la misura binomiale nell'intervallo chiuso  $[0,1]$ . Il processo che la origina ha come iniziatore l'intervallo stesso e come generatore un parametro  $0 < m_0 < 1$ , chiamato moltiplicatore, che identifica la frazione della probabilità (massa) che viene divisa ad ogni iterazione. Al passo  $k = 0$ , si definisce una misura di probabilità uniforme  $\mu_0$  in modo che:

$$f_0(t) = \begin{cases} \mu_0([0,1]) = 1 & \text{se } t \in [0,1] \\ 0 & \text{se } t \notin [0,1] \end{cases}$$

In  $k = 1$ , la misura  $\mu_1$  ripartisce la probabilità  $m_0$  in  $\left[0, \frac{1}{2}\right]$  e  $m_1$  in  $\left[\frac{1}{2}, 1\right]$ , in modo che  $m_0 + m_1 = 1$ , ottenendo:

$$f_1(t) = \begin{cases} \mu_1\left(\left[0, \frac{1}{2}\right]\right) = m_0 & \text{se } t \in \left[0, \frac{1}{2}\right] \\ \mu_1\left(\left[\frac{1}{2}, 1\right]\right) = m_1 & \text{se } t \in \left[\frac{1}{2}, 1\right] \\ 0 & \text{se } t \notin [0,1] \end{cases}$$

Per l'iterazione  $k = 2$ , in modo analogo, l'intervallo è nuovamente diviso, ripartendo la massa con lo stesso schema:

---

<sup>178</sup> C. J. G. EVERTSZ, B. B. MANDELBROT, *Multifractal measures*, 1992.



$$f_2(t) = \begin{cases} \mu_2\left(\left[0, \frac{1}{4}\right]\right) = m_0 m_0 & \text{se } t \in \left[0, \frac{1}{4}\right] \\ \mu_2\left(\left[\frac{1}{4}, \frac{1}{2}\right]\right) = m_0 m_1 & \text{se } t \in \left[\frac{1}{4}, \frac{1}{2}\right] \\ \mu_2\left(\left[\frac{1}{2}, \frac{3}{4}\right]\right) = m_1 m_0 & \text{se } t \in \left[\frac{1}{2}, \frac{3}{4}\right] \\ \mu_2\left(\left[\frac{3}{4}, 1\right]\right) = m_1 m_1 & \text{se } t \in \left[\frac{3}{4}, 1\right] \end{cases}$$

Si osserva che  $m_0^2 + 2m_0m_1 + m_1^2 = (m_0 + m_1)^2 = 1$ . Se si considera definita, al passo  $k + 1$ , la misura  $\mu_k$ , si può costruire  $\mu_{k+1}$  considerando l'intervallo  $[t, t + 2^{-k}]$ , ove  $t$  è un numero binario in base  $b = 2$  nella forma  $t = (0\eta_1\eta_2\cdots\eta_k)_2 = \sum_{i=1}^k \eta_i 2^{-i}$ , e, in seguito, dividendo la frazione  $m_0$  e  $m_1$  della massa contenuta in  $\mu_k([t, t + 2^{-k}])$  nei sotto-intervalli  $[t, t + 2^{-k-1}]$  e  $[t + 2^{-k-1}, t + 2^{-k}]$ . La ripetizione di questo processo genera una sequenza infinita di misure  $\{\mu_k\}$ . La misura binomiale  $\mu$  è definita come limite di tale sequenza, in modo che:

$$\{\mu_k\} \xrightarrow{d} \mu.$$

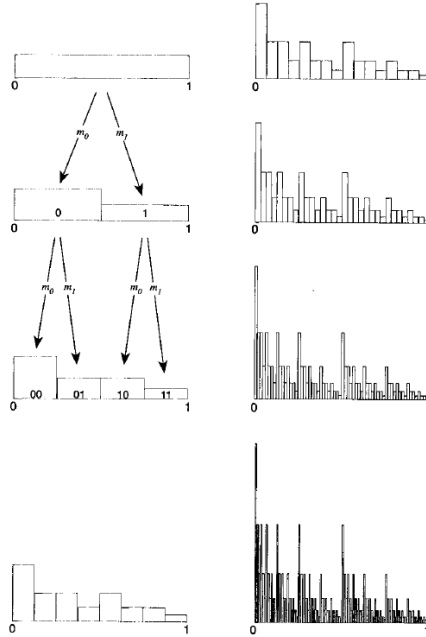


Fig. 39 Cascata moltiplicativa della misura binomiale. [17]

Da questa costruzione, si possono sviluppare numerose estensioni. Dividendo l'intervallo iniziale in  $b > 2$  parti, in modo da formare  $0 \leq \beta \leq b - 1$  sotto-intervalli, nei quali si distribuisce una frazione della massa totale  $m_0, \dots, m_{b-1}$ , si può definire una classe di misure dette multinomiali. Se la frazione della massa distribuita non è omogenea ad ogni iterazione, i moltiplicatori di ogni sotto-intervallo sono una variabile casuale discreta  $M_\beta$ , che assume valori  $m_0, m_1, \dots, m_{b-1}$ , con probabilità,  $p_0, p_1, \dots, p_{b-1}$ . È possibile generalizzare la costruzione precedente permettendo che  $M_\beta$  non sia necessariamente discreta, ma possa essere una generica variabile casuale indipendente e identicamente distribuita, definendo così le misure moltiplicative<sup>179</sup>. Come in precedenza, si parte dall'intervallo  $[0,1]$  con massa unitaria e si divide in  $\beta$  sotto-intervalli distribuendo la massa casuale  $M_\beta$ . Partendo da  $t = (0\eta_1\eta_2\ldots\eta_k)_b = \sum \eta_i b^{-i}$  e considerando la lunghezza  $\Delta t = t + b^{-k} - t = b^{-k}$ , la misura dell'intervallo  $[t, t + \Delta t]$  è data da:

$$\mu(\Delta t) = M(\eta_1)M(\eta_1\eta_2) \dots M(\eta_1 \dots \eta_k),$$

da cui segue  $[\mu(\Delta t)]^q = M(\eta_1)^q M(\eta_1\eta_2)^q \dots M(\eta_1 \dots \eta_k)^q$ , per ogni  $q \geq 0$ . Essendo i moltiplicatori indipendenti ad ogni passo della cascata, considerando il valore atteso, i momenti della misura possono essere scritti come:

$$E[\mu(\Delta t)^q] = [E(M^q)]^k. \text{ }^{180}$$

---

<sup>179</sup> B. B. MANDELBROT, A. FISHER, L. CALVET, *A multifractal model of asset returns*, 1997.

<sup>180</sup>

$$\begin{aligned} E[\mu(\Delta t)^q] &= E[M(\eta_1)^q M(\eta_1\eta_2)^q \dots M(\eta_1 \dots \eta_k)^q] = \underbrace{E(M^q)E(M^q) \dots E(M^q)}_{k \text{ volte}} \\ &= [E(M^q)]^k. \end{aligned}$$

Infine, sostituendo  $\tau(q) = -\log_b[E(M^q)] - 1$  si ottiene la relazione di scala dei momenti<sup>181</sup>:

$$E[\mu(\Delta t)^q] = \Delta t^{\tau(q)+1}.$$

In tutte le misure definite si è sempre assunta la conservazione della massa durante ogni iterazione della cascata moltiplicativa, ovvero vale  $\sum M_\beta = 1$ . Le misure di questo tipo sono dette conservative o microcanoniche. Si può definire una classe di misure in cui la massa si conserva solamente ‘in media’ in ogni fase del processo, in modo che  $E(\sum M_\beta) = 1$ . Queste misure sono dette canoniche e sono importanti all’interno del modello multifrattale dei mercati. La massa totale  $\Omega$  è, in questo caso, una variabile casuale, indipendente da  $M_\beta$ , e la misura dell’intervallo in base  $b$  si può scrivere come:

$$\mu(\Delta t) = \Omega(\eta_1 \dots \eta_k) M(\eta_1) M(\eta_1 \eta_2) \dots M(\eta_1 \dots \eta_k).$$

Tale misura soddisfa nuovamente la relazione di scala:

$$E[\mu(\Delta t)^q] = E(\Omega^q) [E(M^q)]^k = c(q) \Delta t^{\tau(q)+1},$$

ponendo  $E(\Omega^q) = c(q)$ . Sfruttando la relazione di scala dei momenti, è possibile estendere l’idea di multifrattalità dalle misure ai processi stocastici.

*Definizione.* (Processi multifrattali). Dato uno spazio di probabilità  $(\Omega, \mathcal{F}, P)$ , un processo stocastico  $\{X(t)\}$  con incrementi stazionari è detto multifrattale se  $X(0) = 0$  e soddisfa la relazione:

$$E(|X(t)|^q) = c(q) t^{\tau(q)+1},$$

---

<sup>181</sup>  $[E(M^q)]^k = \left(\frac{1}{b}\right)^{\log_1[E(M^q)]^k} = (b^{-1})^{-k \log_b[E(M^q)]} = (b^{-k})^{-\log_b[E(M^q)]} = \Delta t^{-\log_b[E(M^q)]-1+1} = \Delta t^{\tau(q)+1}.$

per ogni  $t \in T \subset \mathbb{R}$ ,  $q \in Q \subset \mathbb{R}$  e  $c, \tau: Q \rightarrow \mathbb{R}$ .  $\tau(q)$  è detta funzione di scala del processo, è concava ed ha come intercetta  $\tau(0) = -1$ <sup>182</sup>.

È possibile mostrare come i processi autosimili siano un caso particolare di processi multifrattali. Un processo stocastico autosimile  $\{X_t | t \geq 0\}$  con indice  $H$  soddisfa la condizione  $\{X_{at}\} \sim \{a^H X_t\}$ . Ponendo  $at = 1$ , si può scrivere:

$$\{X_1\} \sim \left\{ \left( \frac{1}{t} \right)^H X_t \right\} \rightarrow \{t^H X_1\} \sim \{X_t\}.$$

Inoltre, segue che:

$$\{|X_t|\} \sim \{t^H |X_1|\},$$

e che:

$$\{|X_t|^q\} \sim \{t^{Hq} |X_1|^q\}.$$

Calcolandone il valore atteso, si può scrivere:

$$E[|X_t|^q] = t^{Hq} E[|X_1|^q].$$

Sostituendo  $\tau(q) = Hq - 1$  e  $c(q) = E[|X_1|^q]$  si ottiene la relazione di scala. I processi autosimili sono un caso speciale di processi multifrattali, caratterizzati da una funzione di scala lineare, con pendenza  $H$ . Essendo determinata da un unico parametro essi sono chiamati unifrattali. Una funzione  $\tau(q)$  non lineare, concava, caratterizza invece i processi multifrattali fin qui esaminati. Si può dimostrare che la multifrattalità si può avere solo in intervalli di tempo limitati. In altre parole, un processo definito in un intervallo illimitato può essere multifrattale solo considerandolo in una partizione dell'intervallo temporale. Questo cambiamento nelle proprietà di scala di un processo si definisce *crossover*<sup>183</sup>.

---

<sup>182</sup>  $\tau(q) = -\log_b[E(M^q)] - 1 = -\log_b(\sum_{j=1}^{b-1} Pr\{M_j = m_j\} m_j^q) - 1$ . Ponendo  $q = 0$ , si può scrivere,  $\tau(0) = -\log_b(\sum_{j=1}^{b-1} Pr\{M_j = m_j\} m_j^0) - 1 = -\log_b(\sum_{j=1}^{b-1} Pr\{M_j = m_j\}) - 1 = -\log_b(1) - 1 = -1$ . Per semplicità si ipotizza  $\{M_j\}$  variabile discreta, ma tale risultato si può ricavare anche nel caso di generica variabile continua.

<sup>183</sup> B. B. MANDELBROT, A. FISHER, L. CALVET, *A multifractal model of asset returns*, 1997.

### 4.9.3 Lo spettro multifrattale e l'esponente di Hölder locale

Dopo aver ricavato la relazione di scala sfruttando il comportamento ‘globale’ dei momenti, si cerca ora di analizzare le proprietà ‘locali’ dei multifrattali. A tal fine, si introduce il concetto, dall’analisi reale, di esponente di Hölder locale.

*Definizione.* (Esponente di Hölder locale). Data una funzione  $g: A \subseteq \mathbb{R} \rightarrow \mathbb{R}$  definita in un intorno di un dato punto  $t_0$ , il valore

$$\alpha(t_0) := \sup\{\beta \geq 0: |g(t_0 + h) - g(t_0)| = O(|h|^\beta)\},$$

è detto esponente di Hölder locale di  $g$  in  $t_0$ , per  $h \rightarrow 0$ <sup>184</sup>. Si può notare che  $\alpha(t_0)$  è positivo se  $g$  è limitata attorno a  $t_0$ , inoltre  $\alpha(t_0) = 0$  in un punto di discontinuità e  $\alpha(t_0) = 1$  in un punto differenziabile. L’esponente locale di Hölder si può anche definire, nel caso  $g(t)$  sia una funzione continua, tramite il logaritmo, come:

$$\alpha(t_0) = \lim_{h \rightarrow 0} \sup \frac{\log|g(t_0 + h) - g(t_0)|}{\log|h|}.$$

Per funzioni elementari l’esponente di Hölder assume valori interi. Nel caso del moto browniano classico si ha  $\alpha(t_0) = \frac{1}{2}$  e per il moto browniano frazionario tale esponente coincide con l’indice  $H$ . È importante notare che i processi stocastici classici possiedono un unico esponente, identificato dalla pendenza della funzione lineare di scala. I multifrattali sono invece caratterizzati da una molteplicità di esponenti locali di Hölder, la cui distribuzione viene solitamente espressa tramite una funzione  $f(\alpha)$ , chiamata spettro multifrattale<sup>185</sup>. Dato l’intervallo chiuso  $[0, T] \subset \mathbb{R}$  è possibile calcolare la probabilità che un punto dell’insieme abbia un dato esponente di Hölder. L’idea è quella di dividere il

---

<sup>184</sup> Data  $g(x) \neq 0$ , si scrive che  $f(x) = O(g(x))$  in un intorno di  $x_0$  se:

$$\lim_{x \rightarrow x_0} \sup \left| \frac{f(x)}{g(x)} \right| < +\infty.$$

<sup>185</sup> L. CALVET, A. FISHER, B. B. MANDELBROT, *Large deviations and the distribution of price changes*, 1997.

dominio  $k$  volte, in  $b^k$  sotto intervalli di uguale lunghezza e di calcolare ad ogni passo, per il punto  $t = t_i \in [0, T]$ , la quantità  $|g(t_i + b^{-k}T) - g(t_i)|$ . L'esponente di Hölder è dato da:

$$\alpha_k(t_i) := \frac{\log|g(t_i + b^{-k}T) - g(t_i)|}{\log(b^{-k})}.$$

Modificando la lunghezza dell'intervallo  $\Delta t = b^{-k}$  si ottengono diversi valori  $\alpha(t_i)$ . Dividendo il range di valori degli esponenti in piccoli intervalli disgiunti  $(\alpha_j, \alpha_j + \Delta\alpha]$ , si denota con  $N_k(\alpha_j)$  il numero di esponenti locali di Hölder contenuti in ogni intervallo. Per  $k \rightarrow \infty$ , il rapporto  $\frac{N_k(\alpha_j)}{b^k}$  converge alla probabilità che un punto casuale  $t_i \in [0, T]$  abbia esponente locale di Hölder pari a  $\alpha(t_i)$ . Questo metodo non consente realmente di distinguere i processi unifrattali da quelli multifrattali, in quanto solitamente esiste un valore  $\alpha_0$  che prevale sugli altri. Per definire i multifrattali, occorre un'ulteriore considerazione. Data la funzione  $g: A \subseteq [0, T] \rightarrow \mathbb{R}$  si definisce:

$$f(\alpha) := \lim_{k \rightarrow \infty} \frac{\ln N_k(\alpha)}{\ln b^k}.$$

Se tale limite esiste ed è positivo, la funzione  $g(t)$  si chiama multifrattale e la funzione  $f(\alpha)$  è lo spettro multifrattale di  $g$ . Questa costruzione permette che  $\alpha(t)$  non sia costante in  $t$ , ovvero il processo ammette differenti comportamenti di scala locali, misurati dai diversi esponenti di Hölder. Lo spettro multifrattale può essere interpretato come la dimensione di Hausdorff di un insieme di punti caratterizzati dall'esponente locale  $\alpha$ . Per ogni  $\alpha \geq 0$ , si definisce  $T_\alpha$  l'insieme dei punti con esponente di Hölder  $\alpha$ . Ogni insieme  $T_\alpha$  ha dimensione frattale  $\dim_H(T_\alpha)$ , tale che  $0 < \dim_H(T_\alpha) < 1$ . La dimensione frattale coincide con lo spettro, in modo che:

$$f(\alpha) = \dim_H(T_\alpha).$$

#### 4.9.4 Definizione e proprietà dell'MMAR

Il modello multifrattale del rendimento delle attività (MMAR) interpreta il prezzo di un'attività come un processo multifrattale composto, con code spesse, dipendenza da lungo termine e invarianza di scala. In particolare, i rendimenti logaritmici seguono un moto browniano frazionario in un tempo di contrattazione multifrattale. L'MMAR rappresenta un'evoluzione rispetto ai modelli precedenti, in quanto ammette code spesse, senza implicare un secondo momento divergente, e memoria a lungo termine nel valore assoluto dei rendimenti, senza escludere la presenza di rumore bianco<sup>186</sup>.

*Definizione.* (Processo composto). Dato il processo stocastico  $\{Z(t)\}$  e  $\theta(t)$  una funzione crescente di  $t$ , il processo

$$X(t) := Z[\theta(t)],$$

è detto processo composto.  $t$  rappresenta il tempo classico dell'orologio e  $\theta(t)$  è il tempo di contrattazione che deforma  $t$ . In particolare,  $\theta(t)$  è la distribuzione cumulata di probabilità di una misura multifrattale. Dato il prezzo di un asset finanziario  $\{P(t)\}_{t \in [0, T]}$ , si definisce il processo stocastico  $X(t)$ , che descrive il rendimento logaritmico, come:

$$X(t) := \ln P(t) - \ln P(0).$$

L'MMAR è caratterizzato da tre importanti ipotesi:

- i.  $\{X(t)\}_{t \in [0, T]}$  è un processo composto,

$$X(t) := B_H[\theta(t)],$$

con  $B_H$  moto browniano frazionario con indice  $H$  e  $\theta(t)$  tempo di contrattazione;

---

<sup>186</sup> B. B. MANDELBROT, A. FISHER, L. CALVET, *A multifractal model of asset returns*, 1997.

- ii. Il processo stocastico  $\theta(t)$  è una distribuzione di probabilità cumulata di una misura multifrattale definita in  $[0, T]$ , è continuo ed ha incrementi crescenti e stazionari;
- iii.  $\{B_H(t)\}$  e  $\{\theta(t)\}$  sono indipendenti.

Si può notare che  $\theta(0) = 0$  in quanto  $X(0) = 0$ . Tramite la nozione di processo composto, il tempo di contrattazione trasferisce la multifrattalità ai rendimenti e consente di modellare la variabilità, senza modificare la direzione degli incrementi e la loro correlazione. Per questo motivo, si può immaginare l'esistenza di una relazione tra le funzioni di scala di  $X(t)$  e di  $\theta(t)$ . Il processo stocastico in esame è valutato nel dominio  $[0, T]$ , in quanto la multifrattalità esiste solo per intervalli limitati.

*Teorema.* Sotto queste tre ipotesi iniziali, il processo  $X(t)$  è un multifrattale con incrementi stazionari e con funzione di scala:

$$\tau_X(q) = \tau_\theta(Hq).$$

*Dimostrazione.* Dato  $X(t) = B_H[\theta(t)]$ , il momento  $q$ -esimo si può scrivere, usando la legge delle aspettative iterate<sup>187</sup>, come:

$$E[|X(t)|^q] = E[E(|X(t)|^q | \theta(t) = u)].$$

Dato che il moto browniano frazionario  $\{B_H(t)\}$  e il tempo di contrattazione  $\{\theta(t)\}$  sono due processi indipendenti, usando la proprietà di auto-similarità, si può scrivere:

$$E(|X(t)|^q | \theta(t) = u) = E(|B_H(u)|^q | \theta(t) = u) = \theta(t)^{Hq} E[|B_H(1)|^q].$$

Da cui,

$$E[|X(t)|^q] = E[\theta(t)^{Hq} E[|B_H(1)|^q]] = E[\theta(t)^{Hq}] E[|B_H(1)|^q].$$

Dato che il processo multifrattale  $\theta(t)$  soddisfa la relazione di scala:

---

<sup>187</sup> La legge delle aspettative iterate o dell'aspettativa totale afferma che, date due variabili casuali  $X$  e  $Y$ , definite su uno stesso spazio di probabilità, vale la relazione:

$$E[E(X|Y)] = E[X].$$



$$E[|\theta(t)|^q] = c_\theta(q)t^{\tau_\theta(q)+1},$$

vale che:

$$E[|X(t)|^q] = E[|B_H(1)|^q]c_\theta(Hq)t^{\tau_\theta(Hq)+1}.$$

Sostituendo  $c_X(q) = E[|B_H(1)|^q]$  e  $\tau_X(q) = \tau_\theta(Hq)$ , si trova che  $X(t)$  soddisfa la relazione multifrattale:

$$E[|X(t)|^q] = c_X(q)t^{\tau_X(q)+1}.$$

Si può notare che  $E[|X(t)|^q]$  dipende da  $E[|\theta(t)|^{Hq}]$ , quindi esso è finito solo se il processo  $\theta(t)$  ha momento finito, di ordine  $Hq$ . In altre parole, il tempo di contrattazione determina i momenti dei rendimenti logaritmici. Dalla seconda ipotesi, si può ricavare un'ulteriore proprietà dell'MMAR, poiché il tipo di misura multifrattale, da cui dipende  $\theta(t)$ , determina le caratteristiche dei momenti. Una misura microcanonica determina un processo composto  $X(t)$  con momenti finiti e code sottili. Una misura canonica, in cui la massa totale dipende dalla variabile casuale  $\Omega$ , genera code spesse e momenti divergenti. In particolare, in quest'ultimo caso, esiste un esponente critico  $q_{crit}(\theta)$ , tale che il momento  $E[|\theta(t)|^q]$  risulta finito per  $0 \leq q < q_{crit}(\theta)$  e infinito per  $q \geq q_{crit}(\theta)$ . L'MMAR è sufficientemente flessibile per consentire entrambe le tipologie di comportamento. Per  $H = \frac{1}{2}$ ,  $\{B_H(t)\}$  è un moto browniano classico, con momenti finiti, incrementi stazionari e funzione di autocovarianza nulla. Per  $H \neq \frac{1}{2}$ ,  $\{B_H(t)\}$  è un moto browniano frazionario, con funzione di autocovarianza di  $X(t)$ , per ogni  $t \geq \Delta t$ , data da:

$$\Gamma_X(t) = c_\theta(2H) \frac{\sigma^2}{2} [(t + \Delta t)^m + (t - \Delta t)^m - 2t^m],$$

ove  $m = \tau_\theta(2H) + 1$ . Se  $H > \frac{1}{2}$ ,  $\Gamma_X(t) > 0$ , se  $H < \frac{1}{2}$ ,  $\Gamma_X(t) < 0$ . La memoria a lungo termine ha un'interpretazione delicata, in quanto il processo è definito

nell'intervallo  $[0, T]$ . Per l'MMAR, si definisce memoria a lungo termine il più lungo ciclo apparente che ha la stessa lunghezza dell'intervallo del dominio<sup>188</sup>. Il processo  $X(t)$  ha uno spettro multifrattale  $f_X(\alpha)$  pari a:

$$f_X(\alpha) = f_\theta\left(\frac{\alpha}{H}\right).$$

In altre parole, si può dimostrare che l'insieme di punti dove  $X(t)$  ha esponente di Hölder locale  $\alpha$  è lo stesso insieme in cui  $\theta(t)$  ha esponente pari a  $\frac{\alpha}{H}$ , quindi essi hanno la stessa dimensione frattale e lo stesso spettro<sup>189</sup>. Si può nuovamente notare che  $\theta(t)$  contiene una moltitudine di esponenti di Hölder e trasmette a  $X(t)$  la multifrattalità. Il principale problema dell'MMAR riguarda il fatto che i metodi statistici classici non possono essere applicati, tuttavia sono stati ideati nuovi strumenti per testarne l'applicabilità<sup>190</sup>.

## *Appendice E. Il metodo della massima verosimiglianza*

In precedenza, si è visto come le leggi di potenza possano rappresentare un'alternativa al modello normale, per descrivere l'evoluzione di un'ampia categoria di fenomeni, tra cui le oscillazioni dei mercati finanziari. A tal fine, è necessario stimare il valore del parametro di scala della distribuzione. Nel capitolo precedente, per calcolare la dimensione frattale della lunghezza costiera si è costruito il grafico bilogarithmico del numero di parti al variare della scala di riduzione e si è calcolata la pendenza della retta di regressione che

---

<sup>188</sup> B. B. MANDELBROT, A. FISHER, L. CALVET, *A multifractal model of asset returns*, 1997.

<sup>189</sup> CALVET, A. FISHER, B. B. MANDELBROT, *Large deviations and the distribution of price changes*, 1997.

<sup>190</sup> È possibile trovare test empirici dell'MMAR, condotti sul comportamento dei dati del tasso di cambio in A. FISHER, L. CALVET, B. B. MANDELBROT, *Multifractality of Deutschemark / US Dollar exchange rates*, 1997.

meglio approssima la distribuzione dei punti. La regressione lineare non offre un modo approfondito per stimare la grandezza dell'errore commesso nell'approssimare la distribuzione del fenomeno<sup>191</sup>. Per ottenere una stima migliore del parametro di scala, una possibilità è usare il metodo della massima verosimiglianza (MLE), ideato dal biologo britannico Ronald Fisher, nel 1922. L'idea è di calcolare il parametro della distribuzione che rende più probabili i dati campionari di partenza. Si può immaginare che, se la distribuzione di probabilità fornisce un risultato, fissato un valore del parametro, la funzione di verosimiglianza, su cui si basa la costruzione dello stimatore MLE, è una funzione del parametro, dato un particolare risultato. Massimizzando la funzione di verosimiglianza si ottiene il valore che rende più probabili i dati osservati. Tale massimo si chiama stima della massima verosimiglianza ed è, nel caso delle leggi di potenza, l'esponente di scala  $\alpha$ . Se la funzione di verosimiglianza è derivabile, è possibile applicare le condizioni di primo ordine, dal teorema di Fermat, in modo esplicito, in altri casi possono essere usati metodi numerici, per funzioni più complesse. Si considera la distribuzione di una legge di potenza continua:

$$p(x) = Cx^{-\alpha} = \frac{\alpha - 1}{x_{min}} \left( \frac{x}{x_{min}} \right)^{-\alpha}.$$

Dato un campione di  $n$  osservazioni  $x_i \geq x_{min}$ , si cerca di stimare il valore di  $\alpha$ , che rende più verosimile ottenere tali osservazioni, dalla legge di potenza. Se le osservazioni sono indipendenti tra loro, la probabilità di ottenerle si può scrivere in termini di prodotto come:

$$p(x|\alpha) = p(x_1|\alpha)p(x_2|\alpha) \dots p(x_i|\alpha) = \prod_{i=1}^n p(x_i|\alpha) = \prod_{i=1}^n \frac{\alpha - 1}{x_{min}} \left( \frac{x_i}{x_{min}} \right)^{-\alpha}.$$

---

<sup>191</sup> A. JAMES, M. J. PLANK, *On fitting power laws to ecological data*, 2007.

Per trovare il valore di  $\alpha$ , date le  $n$  osservazioni, si è interessati a calcolare la probabilità,  $p(\alpha|x)$ . Sfruttando la legge di Bayes della probabilità condizionata, si ha che  $p(\alpha|x) = p(x|\alpha) \frac{p(\alpha)}{p(x)}$  e, assumendo costanti  $p(\alpha)$  e  $p(x)$ , si può scrivere<sup>192</sup>:

$$p(\alpha|x) \propto p(x|\alpha).$$

L'espressione precedente è dunque la funzione di verosimiglianza, per una legge di potenza continua. Spesso è più utile prenderne il logaritmo, in quanto si ottengono somme di logaritmi al posto di prodotti, più facili da massimizzare. Denotando con  $\mathcal{L}$  il logaritmo della funzione di verosimiglianza si ottiene:

$$\begin{aligned} \mathcal{L} = \ln p(\alpha|x) &= \ln \prod_{i=1}^n \frac{\alpha - 1}{x_{min}} \left( \frac{x_i}{x_{min}} \right)^{-\alpha} \\ &= \sum_{i=1}^n \left[ \ln(\alpha - 1) - \ln x_{min} - \alpha \ln \frac{x_i}{x_{min}} \right] \\ &= n \ln(\alpha - 1) - n \ln x_{min} - \alpha \sum_{i=1}^n \ln \frac{x_i}{x_{min}}. \end{aligned}$$

Per calcolare il valore di  $\alpha$ , che rende massima la verosimiglianza, è sufficiente porre  $\frac{\partial \mathcal{L}}{\partial \alpha} = 0$ , da cui si ha:

$$\frac{n}{\alpha - 1} - \sum_{i=1}^n \ln \frac{x_i}{x_{min}} = 0.$$

Ricavando da questa espressione  $\alpha$ , si ottiene una formula esplicita dell'MLE per la stima del parametro di scala di una legge di potenza:

$$\hat{\alpha} = 1 + n \left[ \sum_{i=1}^n \ln \frac{x_i}{x_{min}} \right]^{-1}.$$

---

<sup>192</sup> M. E. J. NEWMAN, *Power laws, Pareto distributions and Zipf's law*, 2006.

La stima così ottenuta è consistente, ovvero al crescere di  $n$  il valore  $\hat{\alpha}$  converge al valore reale.



## **CAPITOLO V.**

### **LA DIPENDENZA A LUNGO TERMINE NEL MERCATO EUROPEO**

## 5.1 *La Rescaled Range Analysis*

Nei test statistici classici, si assume solitamente che le osservazioni empiriche siano identicamente e indipendentemente distribuite e si usa il teorema del limite centrale per approssimare ad una normale la distribuzione reale dei dati, ipotizzando una sufficiente numerosità campionaria. Gli studi di Mandelbrot sui mercati finanziari hanno evidenziato un evidente limite di tale approccio, se utilizzato per studiare le oscillazioni dei prezzi, in quanto i processi finanziari sembrano essere caratterizzati da complesse dinamiche non deterministiche. È necessario formulare una nuova metodologia statistica non parametrica, che non assuma a priori un qualche tipo di distribuzione iniziale dei dati. Un esempio di test non parametrico, chiamato R/S analisi, è stato sviluppato dall'idrologo britannico Edwin Hurst, con lo scopo di prevedere le piene del Nilo, e consente di distinguere, in generale, tra serie storiche casuali e frattali. Come descritto nel precedente capitolo, i modelli classici per studiare la variazione delle piene del Nilo si sono dimostrati inadeguati. Essi ipotizzavano indipendenti tra loro le variazioni, in un processo casuale gaussiano, la cui deviazione degli incrementi  $R$  segue, nel tempo, la regola della radice quadrata:

$$R = T^{\frac{1}{2}}.$$

Hurst ha rilevato l'esistenza di cicli di lunghezza non periodica, causati da una dipendenza statistica di lungo periodo tra le osservazioni e, al fine di stabilire l'altezza necessaria di un'ipotetica diga, ha ideato un nuovo modo per calcolare il campo di variazione. In particolare, per intervalli di tempo di durata e inizio variabile, è necessario eliminare la tendenza in atto, prima di stimare la variazione. Hurst ha notato che il campo di variazione purificato dalla tendenza



cresce più velocemente di quello browniano, seguendo una legge di potenza di coefficiente  $H$ <sup>193</sup>.

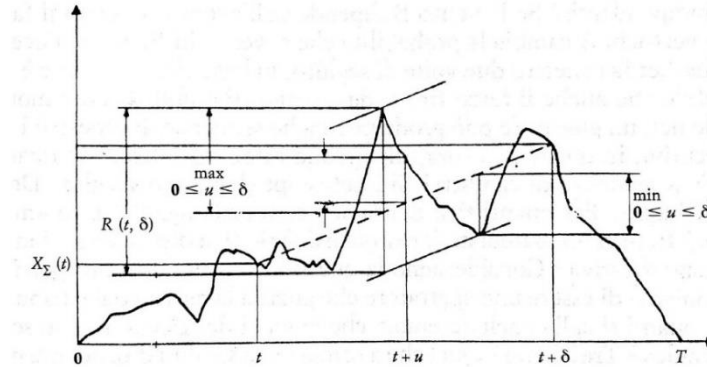


Fig. 40 Campo di variazione di Hurst. [10]

Il test R/S può essere applicato in diversi contesti, in particolare può essere usato per stimare il parametro  $H$  di serie storiche finanziarie. Si considera una serie di lunghezza  $N$  e la si divide in  $A$  sotto intervalli di lunghezza  $n$ , in modo che  $An = N$ . Il valore  $R/S$  è calcolato per ogni  $n$  tramite i seguenti passaggi:

- i. Data una serie finita di lunghezza  $n$  con valori osservati  $X_1, X_2, \dots, X_n$ , si calcola il valore medio:

$$\mu = \frac{1}{n} \sum_{i=1}^n X_i;$$

- ii. Si calcola la serie normalizzata, avente media nulla, come differenza tra le osservazioni e il valore medio:

$$Y_t = X_t - \mu, \quad t \in [1, n];$$

- iii. Si costruisce la serie cumulata normalizzata  $Z_t$ :

$$Z_t = \sum_{i=1}^n Y_i, \quad t \in [1, n];$$

<sup>193</sup> B. B. MANDEBROT, *Il disordine dei mercati. Una visione frattale di rischio, rovina e redditività*, 2005, cap. 9.

iv. Si esprime il campo di variazione  $R_n$  come:

$$R_n = \max(Z_1, Z_2, \dots, Z_n) - \min(Z_1, Z_2, \dots, Z_n);$$

v. Si calcola la deviazione standard  $S_n$ :

$$S_n = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2};$$

vi. Il valore  $R/S$  è dato, per ogni intervallo  $n$ , da:

$$\frac{R_n}{S_n} = \frac{\max(Z_1, Z_2, \dots, Z_n) - \min(Z_1, Z_2, \dots, Z_n)}{\sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2}}.$$

La serie  $(R/S)_n$  segue una legge di potenza nella forma:

$$\frac{R_n}{S_n} = an^H,$$

con  $a$  costante. In altre parole, il valore  $R/S$  varia all'aumentare degli incrementi di tempo  $n$ , in base a una legge di potenza con esponente di scala  $H$ <sup>194</sup>. Il coefficiente di Hurst può essere stimato calcolando la pendenza della retta di regressione, in un grafico bilogarithmico:

$$\log\left(\frac{R_n}{S_n}\right) = c + H \log(n),$$

con  $c = \log(a)$ .  $H$  permette di calcolare il grado di persistenza delle fluttuazioni della serie finanziaria, che si distingue in:

- i.  $H > \frac{1}{2}$ , serie persistente;
- ii.  $H = \frac{1}{2}$ , serie perfettamente casuale (moto browniano standard);
- iii.  $H < \frac{1}{2}$ , serie antipersistente.

---

<sup>194</sup> E. E. PETERS, *Fractal market analysis, applying chaos theory to investment and economics*, 1994.

Nell'applicazione empirica dei paragrafi successivi si utilizzerà la versione classica del test R/S ideata da Hurst, tuttavia è importante dire che, nel 1989, l'economista Andrew W. Lo ne ha proposta una versione modificata, utilizzando un diverso stimatore per la deviazione standard  $S$ , che permette di distinguere tra fenomeni di dipendenza a lungo e a breve periodo, in modo da rendere il test invariante per processi con memoria a breve termine<sup>195</sup>.

## 5.2 *Il dataset utilizzato*

L'idea di questo capitolo è quella di applicare il metodo R/S, appena descritto, alle variazioni dei prezzi dei titoli azionari di alcune aziende europee, allo scopo di quantificare il coefficiente di Hurst di ogni serie e di determinare, di conseguenza, il grado di dipendenza da lungo termine presente nei dati. In particolare, si è scelto di considerare un orizzonte temporale di 10 anni, dal 01-01-2010 al 01-01-2020, per le azioni di cinque aziende quotate nel mercato europeo, Bayer, Exor, Lufthansa, Philips e Renault, ottenute dal sito di Yahoo Finance, considerando le variazioni giornaliere. Bayer è uno dei leader mondiali nella produzione e commercializzazione di prodotti farmaceutici, fondato nel 1863, con sede a Leverkusen. I principali settori di interesse sono quelli dei prodotti sanitari, chimici e biotecnologici. Nel dataset utilizzato sono registrate 2536 osservazioni del prezzo giornaliero del titolo, con una media di 68,99 e una deviazione standard di 23,19.

---

<sup>195</sup> A. W. LO, *Long-term memory in stock market prices*, 1989.

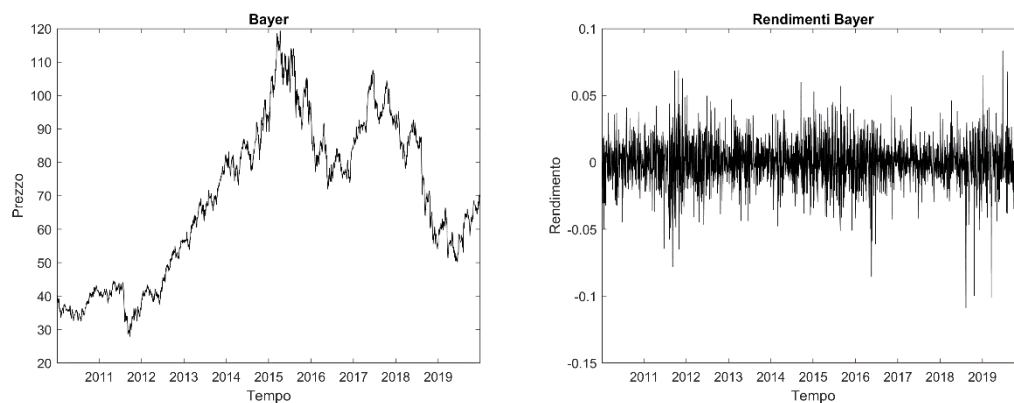


Fig. 41 Grafici del prezzo e del rendimento logaritmico di Bayer.

Exor è una holding finanziaria, con sede in olanda, fondata nel 1927 e controllata dalla famiglia Agnelli. La società si occupa di effettuare investimenti in partecipazioni a medio e lungo termine in altre società, principalmente europee ed americane. Tra le principali partecipazioni si ricordano quelle in FCA, The Economist Group e Ferrari. All'interno del campione sono state rilevate 2538 osservazioni del prezzo del titolo, con una media di 33,78 e una deviazione di 16,20.

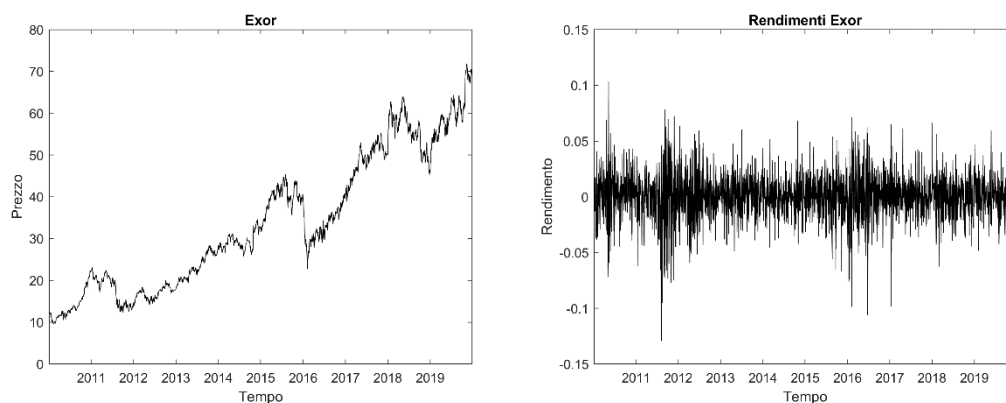


Fig. 42 Grafici del prezzo e del rendimento logaritmico di Exor.

Lufthansa è la principale compagnia aerea tedesca, fondata a Berlino nel 1953, e fornisce servizi di trasporto per merci e passeggeri in tutto il mondo. Nel periodo in esame si sono considerate 2536 osservazioni, con un prezzo medio di 13,88 e una deviazione standard di 4,70.

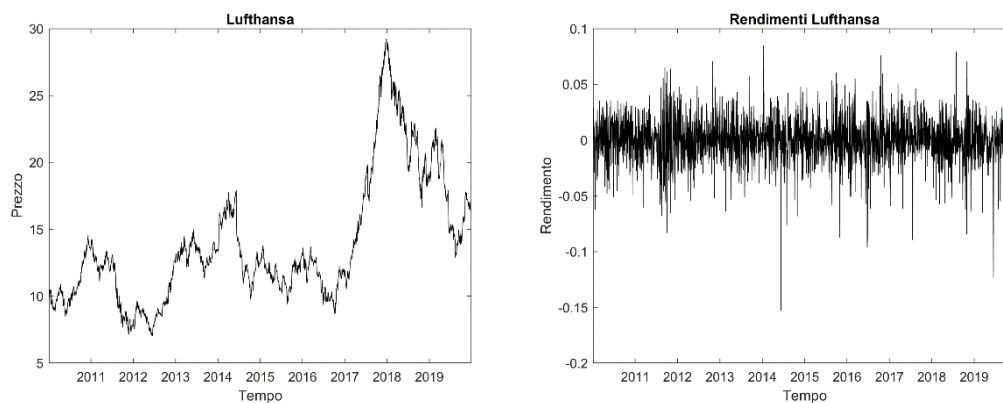


Fig. 43 Grafici del prezzo e del rendimento logaritmico di Lufthansa.

Philips è una azienda quotata olandese fondata ad Eindhoven nel 1891 che è stata per anni uno dei principali marchi nel settore elettronico, per poi specializzarsi, dal 2011, all'interno del settore delle nuove tecnologie per la salute. Le variazioni del prezzo del titolo considerate all'interno del campione sono 2557, con una media di 21,49 e uno scarto quadratico di 8,09.

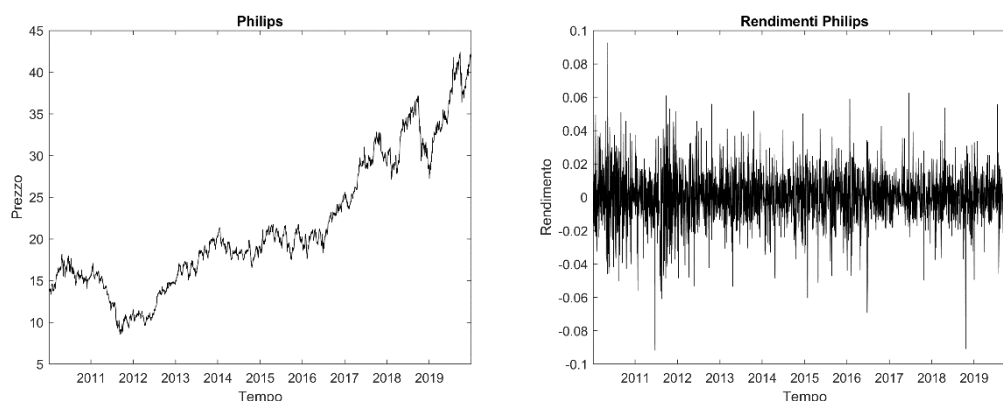


Fig. 44 Grafici del prezzo e del rendimento logaritmico di Philips.

In ultimo, Renault è una importante compagnia automobilistica francese, fondata nel 1898, con un fatturato superiore a 50 miliardi, attiva in diversi segmenti del settore. La serie storica del titolo considerata è composta da 2556 osservazioni, con un prezzo medio di 48,02 e una deviazione standard di 18,08.

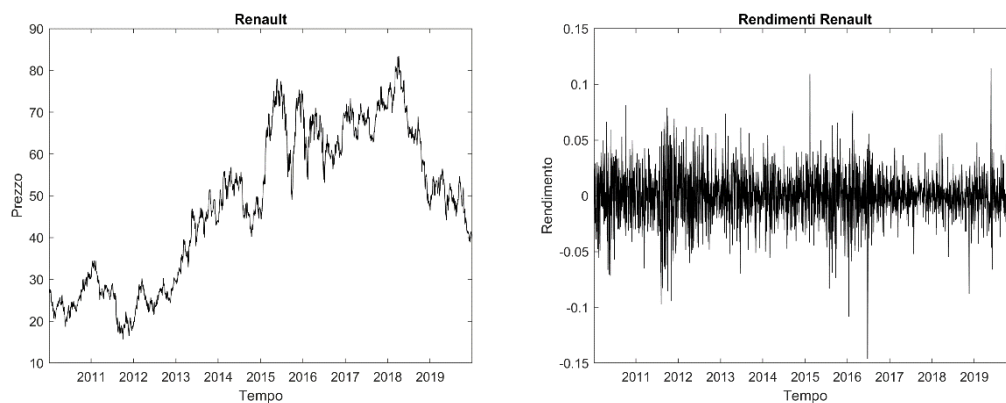


Fig. 45 Grafici del prezzo e del rendimento logaritmico di Renault.

### *5.3 L'applicazione dell'analisi R/S per calcolare l'esponente di Hurst*

Per calcolare il parametro di Hurst di ogni serie storica si è, in primo luogo, costruita la serie dei rendimenti giornalieri logaritmici con il software Matlab, in seguito si è calcolata la pendenza della retta di regressione nel grafico bilogaritmico con, sull'asse delle ordinate, la serie cumulata standardizzata  $(R/S)_n$  e, sull'asse delle ascisse, l'intervallo temporale  $n$ . Tale pendenza fornisce una stima dell'esponente di Hurst, che quantifica il grado di persistenza di una serie storica finanziaria. Al fine di rendere più significativi i risultati si è applicato il metodo descritto, al dataset in esame, utilizzando tre diversi software, Python, Matlab e Gretl. Di seguito si mostrano i grafici R/S dei cinque titoli europei scelti, realizzati con Python 3, tramite il modulo 'hurst 0.0.5', con indicato il valore stimato del parametro  $H$  e dell'intercetta  $c$ .

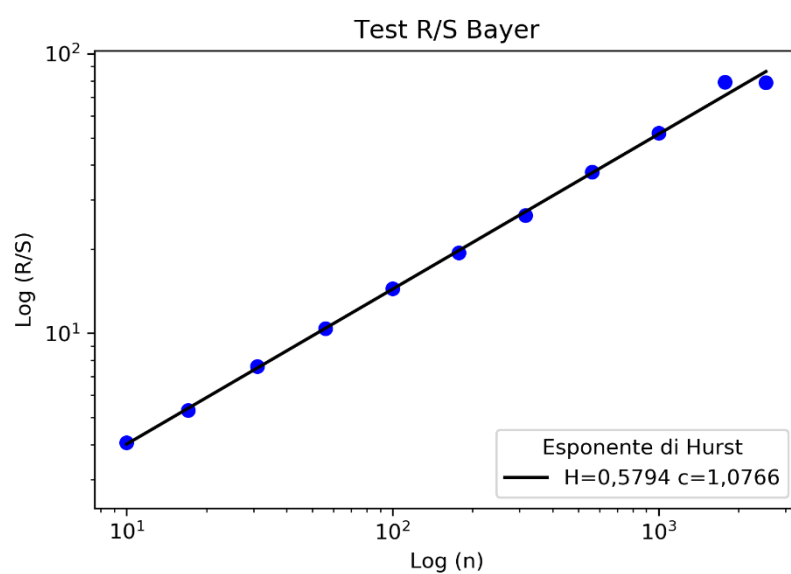


Fig. 46 Grafico R/S per il titolo Bayer.

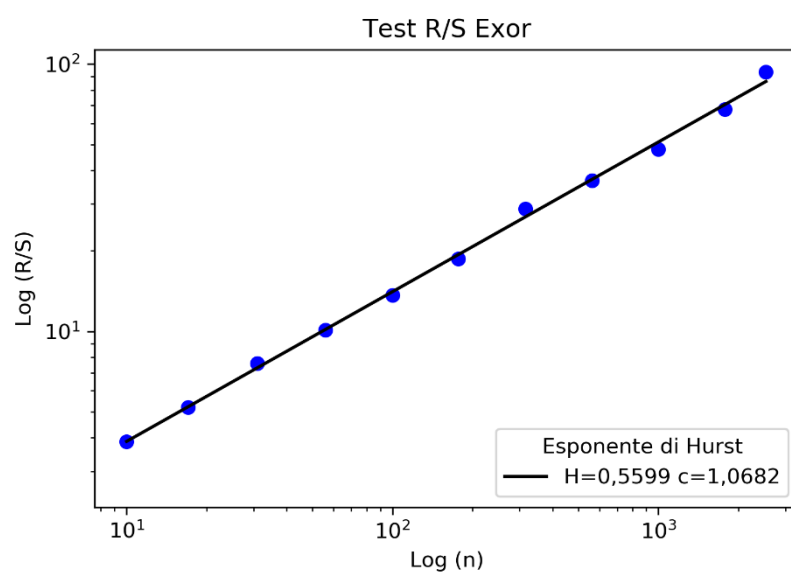


Fig. 47 Grafico R/S per il titolo Exor.

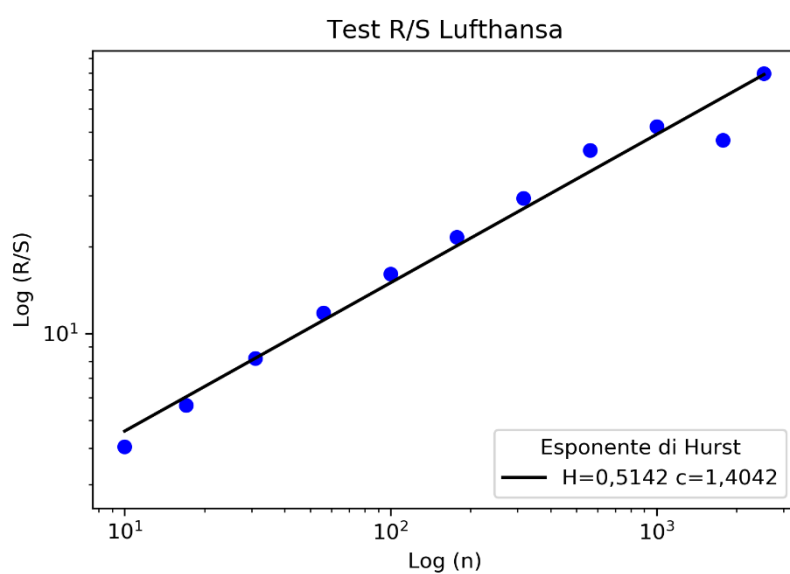


Fig. 48 Grafico R/S per il titolo Lufthansa.

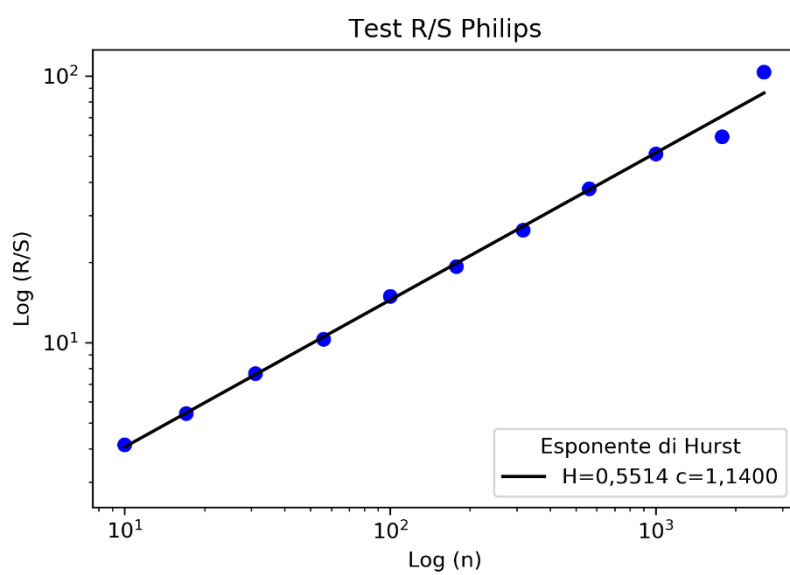


Fig. 49 Grafico R/S per il titolo Philips.



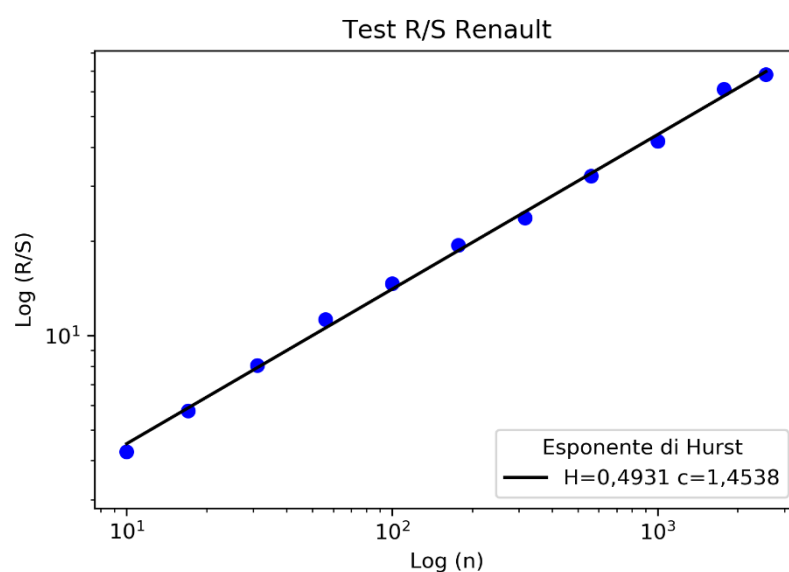


Fig. 50 Grafico R/S per il titolo Renault.

## 5.4 I risultati

Le stime del coefficiente di Hurst con i diversi software, per le cinque serie storiche, sono riportate nella tabella successiva.

Risultati analisi R/S					
	Bayer	Exor	Lufthansa	Philips	Renault
Esponente di Hurst (Python)	0,57	0,55	0,51	0,55	0,49
Esponente di Hurst (Matlab)	0,50	0,51	0,60	0,50	0,52
Esponente di Hurst (Gretl)	0,55	0,50	0,56	0,52	0,52
Errore standard	0,015	0,017	0,018	0,008	0,013

Fig. 51 Tabella delle stime ottenute per l'esponente di Hurst.

Per il primo titolo considerato, Bayer, si può notare un valore dell'esponente di Hurst compreso nell'intervallo  $[0,50, 0,57]$ , che attribuisce alla serie una memoria a lungo termine, essendo  $H > 0,5$ . In particolare, sembra essere presente una dipendenza da lungo periodo tra le osservazioni del titolo, nel periodo considerato, che rende le variazioni persistenti. Il grado di persistenza riscontrato è tuttavia contenuto e differisce molto, in verità, nelle diverse stime. Dal punto di vista aziendale, Bayer ha ottenuto, nei dieci anni in esame, una posizione di leader all'interno del settore farmaceutico, grazie anche a importanti acquisizioni concluse. Questa stabilità aziendale potrebbe comportare una minore rischiosità del titolo e a un effetto Noè poco presente, inoltre sembra realistica una presenza dell'effetto Giuseppe, che lega le variazioni presenti a quelle passate. Il coefficiente di Hurst stimato per il titolo Exor, negli ultimi dieci anni, varia all'interno dell'intervallo  $[0,50, 0,55]$ . La stima ottenuta denota una persistenza delle fluttuazioni del titolo superiore a quella del moto browniano, in quanto si ha  $H > 0,5$ . La persistenza riscontrata potrebbe indicare che il prezzo del titolo riflette le iniziative di investimento della società nei periodi precedenti. Il titolo Lufthansa presenta un parametro di Hurst contenuto nell'intervallo  $[0,51, 0,60]$ , che denota la presenza di una memoria di lungo termine superiore a quella dei casi Bayer ed Exor. Tale memoria permette l'esistenza di cicli non periodici che compromettono l'indipendenza tra le variazioni del titolo, come invece ipotizzato nel modello di Bachelier. La persistenza delle variazioni rende maggiormente probabile, per il titolo Lufthansa, l'esistenza di variazioni giornaliere dello stesso segno. È importante notare che per questo titolo si è stimato un errore standard maggiore rispetto agli altri casi, che potrebbe indicare l'esistenza di un effetto Noè nelle oscillazioni. Per quanto riguarda il caso di Philips, dalle analisi emerge un

valore stimato di  $H$  appartenente all'intervallo  $[0,50, 0,55]$ . Tale valore,  $H > 0,5$ , denota, anche in questo caso, una dipendenza da lungo termine, nell'orizzonte temporale considerato, delle fluttuazioni giornaliere del titolo, di minore entità rispetto al caso Lufthansa. L'azienda olandese è stata protagonista, nel 2011, di un importante cambio del settore di applicazione, scorporando le attività legate al settore elettronico e specializzandosi in quello delle tecnologie per la salute. L'incertezza relativa al cambio della ragione sociale potrebbe essere una causa della minore persistenza, inoltre l'esistenza di un'autocorrelazione positiva di lungo periodo potrebbe anche indicare che le variazioni più recenti riflettono anche quelle precedenti al cambiamento aziendale. Per il titolo Renault è stato stimato, con i tre differenti software, un esponente di Hurst compreso nel range  $[0,49, 0,52]$ . Dalle stime sembra emergere un valore  $H \sim 0,5$  che avvicina il processo delle variazioni giornaliere del titolo a quelle, identicamente e indipendentemente distribuite, del moto browniano standard. Il caso Renault sembra essere caratterizzato da variazioni casuali con uguale probabilità, paragonabili a un cammino casuale. In conclusione, le analisi condotte per aziende europee sembrano evidenziare un discostamento, nella maggior parte dei casi, dal modello classico, denotando la non realistica dell'ipotesi di variazioni indipendenti e normalmente distribuite. I fenomeni di persistenza riscontrati, seppur non elevati, sono coerenti con le osservazioni di Mandelbrot, secondo cui i modelli finanziari classici risultano essere limitati e inadeguati, in quanto in essi è sottostimata la reale rischiosità presente nelle oscillazioni. L'analisi R/S fornisce un metodo per quantificare la presenza di un nuovo fenomeno all'interno dei mercati e suggerisce una diversa direzione di indagine finanziaria, per avvicinare i modelli teorici alla realtà.

## *Appendice F. L'analisi R/S con Matlab*

Si riporta di seguito lo script utilizzato per stimare l'esponente di Hurst, tramite il test R/S, con il software Matlab.

```
function H = RS(sequence,isplot)
%
% 'RS' estimate the hurst parameter of a given sequence with R/S method.
%
% Inputs:
%   sequence: the input sequence for estimate
%   isplot: whether display the plot. without a plot if isplot equal to 0
% Outputs:
%   H: the estimated hurst coeffeient of the input sequence

% Author: Chu Chen
% Version 1.0, 03/10/2008
% chen-chu@163.com
%

if nargin == 1
    isplot = 0;
end

N = length(sequence);
dlarge = floor(N/5);
dsmall = max(10,log10(N)^2);
D = floor(logspace(log10(dsmall),log10(dlarge),50));
D = unique(D);
n = length(D);
x = zeros(1,n);
y = zeros(1,n);

R = cell(1,n);
S = cell(1,n);
for i = 1:n
    d = D(i);
    m = floor(N/d);
    R{i} = zeros(1,m);
    S{i} = zeros(1,m);
    matrix_sequence = reshape(sequence(1:d*m),d,m);
```

```

Z1 = cumsum(matrix_sequence);
Z2 = cumsum(repmat(mean(matrix_sequence),d,1));
R{i} = (max(Z1-Z2)-min(Z1-Z2));
S{i} = std(matrix_sequence);

if min(R{i})==0 || min(S{i}) ==0
    continue;
end

x(i) = log10(d);
y(i) = mean(log10(R{i}./S{i}));
end

% fit a line with middle part of sequence
index = x~=0;
x = x(index);
y = y(index);
n2 = length(x);
cut_min = ceil(3*n2/10);
cut_max = floor(9*n2/10);

X = x(cut_min:cut_max);
Y = y(cut_min:cut_max);
p1 = polyfit(X,Y,1);
Yfit = polyval(p1,X);
H = (Yfit(end)-Yfit(1))/(X(end)-X(1));

if isplot ~= 0
    figure,hold on;
    bound = ceil(log10(N));
    axis([0 bound 0 0.75*bound]);

    temp = (1:n).*index;
    index = temp(index);
    for i = 1:n2
        plot(x(i),log10(R{index(i)}./S{index(i)}),'b.');
```

end

```

    x = linspace(0,bound,10);
    y1 = 0.5*x;
    y2 = x;
    h1 = plot(x,y1,'b--','LineWidth',2);
    h2 = plot(x,y2,'b-','LineWidth',2);
    plot(X,Yfit,'r-','LineWidth',3);
    legend([h1,h2],'slope 1/2','slope')
```

```
    xlabel('log10(blocks of size m)'),ylabel('log10(R/S)'),title('R/S Method');  
end
```

# CONCLUSIONI

L'obiettivo di questo lavoro è stato quello di evidenziare le criticità, presenti nella teoria finanziaria classica, e di indagare un nuovo settore di ricerca, che tenta di risolverle. I mercati finanziari mostrano al loro interno dinamiche complesse, difficili da esprimere in termini deterministici. La teoria classica poggia le basi sul lavoro di Bachelier, che ha descritto l'evoluzione dei prezzi in termini probabilistici, tramite l'idea di random walk. Queste considerazioni sono state sviluppate all'interno dei processi stocastici, in cui il più noto esempio è il moto browniano, che permettono di teorizzare le fluttuazioni finanziarie, in termini di successioni di variabili casuali, indipendenti e normalmente distribuite, nel tempo. Nella prima parte della tesi, sono stati descritti i modelli finanziari classici, la cui validità è stata messa in discussione da numerose osservazioni empiriche. La curva normale si è dimostrata inadatta a spiegare il comportamento dei prezzi e l'irregolarità sembra rappresentare un tratto distintivo dei mercati finanziari. Questa irregolarità può essere studiata, sfruttando una nuova classe di oggetti geometrici, i frattali. Per riassumere l'intuizione dell'utilizzo dei frattali in economia, si può fare un parallelismo con il mondo naturale. Descrivere i mercati finanziari, tramite un cammino casuale classico, equivale a descrivere gli oggetti e i fenomeni naturali, tramite figure euclidee. Così come una montagna non è ben approssimata da un triangolo, le oscillazioni finanziarie non possono essere descritte da una distribuzione normale. La complessità dei frattali è riassunta all'interno delle loro proprietà. In particolare, l'invarianza di scala permette di ritrovare, ingrandendo la figura, lo stesso pattern generale e la dimensione frazionaria fornisce un modo per quantificare il grado di irregolarità della figura. La natura frattale dei mercati

finanziari è stata esaminata attraverso gli sviluppi del lavoro di Mandelbrot, tramite i modelli intermedi, fino alla multifrattalità. Partendo dalle variazioni del cotone, si è visto come la classe di distribuzioni L-stabili sia una valida alternativa alla gaussiana, in quanto, le code spesse, tipiche delle leggi di potenza, permettono di modellizzare meglio le turbolenze dei mercati. In seguito, si è messa in discussione l'ipotesi di indipendenza delle variazioni, teorizzando la dipendenza a lungo termine, che ha consentito di sviluppare una generalizzazione del moto browniano, nota come moto browniano frazionario, tramite il parametro di Hurst. Nell'ultima parte, tutte le considerazioni precedenti sono state riunite all'interno del modello multifrattale, introducendo il tempo di contrattazione, come alternativa al tempo classico lineare. L'idea dietro questo modello è la natura relativistica non lineare del tempo, che scorre diversamente in base all'attività del mercato. Matematicamente, questa tendenza è introdotta da oggetti frattali con diversi esponenti di scala, detti multifrattali. Nell'ultimo capitolo si è applicata l'analisi R/S nel mercato europeo. Il test permette di stimare l'esponente di Hurst, che quantifica il grado di persistenza di una serie storica, per studiare la dipendenza a lungo termine. L'analisi è stata condotta per i prezzi dei titoli, nel periodo 2010-2020, di cinque aziende europee, Bayer, Exor, Lufthansa, Philips e Renault. Il parametro  $H$  è stato stimato con tre diversi software, Python, Matlab e Gretl, al fine di confrontare i risultati e conferire maggiore significatività all'analisi. Gli esiti mostrano fenomeni di memoria a lungo termine, per le variazioni di Bayer, con  $H \in [0,50, 0,57]$ , di Exor, con  $H \in [0,50, 0,55]$ , di Lufthansa, con  $H \in [0,51, 0,60]$  e di Philips, con  $H \in [0,50, 0,55]$ . I valori, anche se non elevati, si ritengono sufficienti per confermare l'esistenza di una tendenza di lungo periodo, che caratterizza i prezzi dei titoli. In particolare, le variazioni non sono



equiprobabili, come nel moto browniano classico, ma sono caratterizzate da una autocorrelazione, che riflette, nei valori presenti, gli eventi passati. Questo ambito di ricerca è, in realtà, ancora da esplorare, infatti, è importante sottolineare che il comportamento dei mercati finanziari è ancora lontano dall'essere interamente compreso e i frattali suggeriscono una nuova direzione di analisi finanziaria, che sembra essere in grado di descrivere la rischiosità reale, meglio dei modelli classici, nonostante l'impossibilità, in economia, di una «teoria del tutto».



# BIBLIOGRAFIA

- [1] A. CLAUSET, C. R. SHALIZI, M. E. J. NEWMAN, *Power law distributions in empirical data*, 2009.
- [2] A. JAMES, M. J. PLANK, *On fitting power laws to ecological data*, 2007
- [3] A. R. PEARS, *Dimension theory of general spaces*, 1975.
- [4] A. V. CHECHKIN, R. METZLER, J. KLAFTER, V. Y. GONCHAR, *Introduction to the theory of Levy flights*, 2008.
- [5] A. W. LO, *Long-term memory in stock market prices*, 1989.
- [6] B. B. MANDELBROT, *A multifractal walk down Wall Street*, 1999.
- [7] B. B. MANDELBROT, A. FISHER, L. CALVET, *A multifractal model of asset returns*, 1997.
- [8] B. B. MANDELBROT, *Gli oggetti frattali. Forma, caso e dimensione*, 2000
- [9] B. B. MANDELBROT, *How long is the coast of Britain? Statistical self-similarity and fractional dimension*, 1976.
- [10] B. B. MANDELBROT, *Il disordine dei mercati. Una visione frattale di rischio, rovina e redditività*, 2005.
- [11] B. B. MANDELBROT, J. W. VAN NESS, *Fractional brownian motion, fractional noises and applications*, 1968.
- [12] B. B. MANDELBROT, *Scaling in financial prices, III: Cartoon brownian motions in multifractal time*, 2001.
- [13] B. B. MANDELBROT, *Survey of multifractality in finance*, 1999.

- [14] B. B. MANDELBROT, *The fractal geometry of the nature*, 1982.
- [15] B. B. MANDELBROT, *The variation of certain speculative prices*, 1963.
- [16] B. MANARIS, L. PELLICORO, G. POTHERING, H. HODGES, *Investigating Esperanto's statistical proportions relative to other languages using neural networks and Zipf's law*, 2006.
- [17] C. J. G. EVERTSZ, B. B. MANDELBROT, *Multifractal measures*, 1992.
- [18] D. KAHNEMAN, A. TVERSKY, *Prospect theory: an analysis of decision under risk*, 1979.
- [19] D. L. COHN, *Measure theory*, 1980.
- [20] E. CONVERSANO, L. L. TEDESCHINI, *Sierpinski triangles in stone, on medieval floors in Rome*, 2011.
- [21] E. E. PETERS, *Fractal market analysis, applying chaos theory to investment and economics*, 1994.
- [22] E. F. FAMA, *Efficient capital markets: a review of theory and empirical work*, 1970.
- [23] E. F. FAMA, K. R. FRENCH, *Common risk factors in the return on stocks and bonds*, 1993.
- [24] E. F. FAMA, *The behavior of stock market prices*, 1965.
- [25] E. GRIMSON, J. GUTTAG, A. BELL, 6.0002, *Introduction to Computational Thinking and Data Science*, Fall 2016, Massachusetts Institute of Technology: MIT.
- [26] E. KREYSZIG, *Advanced engineering mathematics*, 2006.
- [27] F. BIAGINI, Y. HU, B. ØKSENDAL, T. ZHANG, *Stochastic calculus for fractional brownian motion and applications*, 2008.

- [28] F. BLACK, *Capital market equilibrium with restricted borrowing*, 1972.
- [29] F. BLACK, M. SCHOLES, *The pricing of options and corporate liabilities*, 1973.
- [30] F. CARAVENNA, *Moto browniano e analisi stocastica*, 2011.
- [31] F. S. MISHKIN, S. G. EAKINS, G. FORESTIERI, *Istituzioni e mercati finanziari*, 2015.
- [32] G. EDGAR, *Measure, topology and fractal geometry*, 2008.
- [33] G. MATTEI, F. LIBERATI, *Modelli di option pricing: l'equazione di Black e Scholes*, 2013.
- [34] G. TARDIVO, R. SCHIESARI, N. MIGLIETTA, *Finanza aziendale*, 2014.
- [35] G. WEST, *An introduction to modern portfolio theory: Markowitz, CAP-M, APT and Black-Litterman*, 2004.
- [36] H. LAUWERIER, *Fractals, endlessly repeated geometrical figures*, 1991.
- [37] H. MARKOWITZ, *Portfolio selection*, 1952.
- [38] H. PEITGEN, H. JÜRGENS, D. SAUPE, *Chaos and fractals. New frontiers of science*, 2004.
- [39] J. BERAN, *Statistics for long-memory processes*, 1994.
- [40] J. BERK, P. DEMARZO, *Corporate finance*, 2013.
- [41] J. HENRIKSON, *Completeness and total boundedness of the Hausdorff metric*, 1999.
- [42] J. K. HUNTER, *Measure theory*, 2011.
- [43] J. P. BOUCHAUD, *Anomalous diffusion in disordered media: statistical mechanisms, models and physical applications*, 1990.

- [44] J. P. NOLAN, *Stable distributions models for heavy tailed data*, 2018.
- [45] J. TOBIN, *Liquidity preference as behavior towards risk*, 1958.
- [46] J. V. GUTTAG, *Introduction to computation and programming using Python with application to understanding data*, 2016.
- [47] K. J. FALCONER, *Fractal geometry, mathematical foundations and applications*, 2003.
- [48] K. J. FALCONER, *The geometry of fractal sets*, 1986.
- [49] L. BACHELIER, *Théorie de la Spéculation*, 1900.
- [50] L. BOLLANI, L. BOTTACIN, *Statistica per l'azienda, metodi per gli ambiti operativi e per la sfera sociale*, 2018.
- [51] L. C. EVANS, *Partial differential equations*, 1998.
- [52] L. CALVET, A. FISHER, B. B. MANDELBROT, *Large deviations and the distribution of price changes*, 1997.
- [53] M. BARNSLEY, *Fractals everywhere*, 2000.
- [54] M. BAZANT, 18.366, *Random Walks and Diffusion*, Fall 2006, Massachusetts Institute of Technology: MIT.
- [55] M. C. JENSEN, *The performance of mutual funds in the period 1945-64*, 1968.
- [56] M. DAVIS, A. ETHERIDGE, *Louis Bachelier's theory of speculation: the origins of modern finance*, 2006.
- [57] M. E. J. NEWMAN, *Power laws, Pareto distributions and Zipf's law*, 2006.

- [58] M. M. DODSON, S. KRISTENSEN, *Hausdorff dimension and diophantine approximation*, 2003.
- [59] M. R. REINGANUM, *The Anomalous stock market behavior of small firms in January: empirical tests for tax-loss selling effects*, 1983.
- [60] N. AMENC, V. LE SOURD, *Portfolio theory and performance analysis*, 2003.
- [61] P. KEMPTHORNE, C. LEE, V. STRELA, J. XIA, 18.S096, *Topics in Mathematics with Applications in Finance*, Fall 2013, Massachusetts Institute of Technology: MIT.
- [62] P. S. ADDISON, *Fractals and chaos, an illustrated course*, 1997.
- [63] P. WILMOTT, S. HOWISON, J. DEWYNNE, *The mathematics of financial derivatives, a student introduction*, 1995.
- [64] R. DURRETT, *Probability: theory and examples*, 2019.
- [65] R. FITZPATRICK, *Euclid's elements of geometry*, 2008.
- [66] R. METZLER, J. H. JEON, A. G. CHERSTVY, E. BARKAI, *Anomalous diffusion models and their properties: non-stationarity, non-ergodicity, and ageing at the centenary of single particle tracking*, 2014.
- [67] R. RENÒ, *Polline, atomi, Wall Street: l'imprevedibile viaggio del moto browniano*, 2014.
- [68] R. RIGANTI, *Elementi di probabilità e statistica*, 2003.
- [69] S. A. FRANK, *The common patterns of nature*, 2009.
- [70] S. B. BRUNTON, J. N. KUTZ, *Data-driven science and engineering, machine learning, dynamical systems, and control*, 2019.

- [71] S. E. SHREVE, *Stochastic calculus for finance II, continuous time models*, 2004.
- [72] S. PRINA, D. CARACCILOLO, *Finanza matematica*, 2013.
- [73] S. ROSS, *A first course in probability*, 2010.
- [74] S. ROSS, *The arbitrage theory of capital asset pricing*, 1976.
- [75] S. T. PIANTADOSI, *Zipf's word frequency law in natural language: a critical review and future directions*, 2014.
- [76] T. DIEKER, *Simulation of fractional brownian motion*, 2004.
- [77] T. GRAVES, R. B. GRAMACY, N. W. WATKINS, C. L. E. FRANZKE, *A brief history of long memory: Hurst, Mandelbrot and the road to ARFIMA*, 2016.
- [78] THE CORPORATE FINANCE INSTITUTE, *Investment banking manual*, 2018.
- [79] W. F. M. DE BONDT, R. THALER, *Further evidence on investor overreaction and stock market seasonality*, 1987.
- [80] W. PAUL, J. BASCHNAGEL, *Stochastic processes, from physics to finance*, 2013.



