

Comportamento del prezzo delle azioni,  
il modello Black-Scholes e alcune applicazioni  
delle catene di Markov nel campo finanziario

Dario Andrei

18 settembre 2009

## Introduzione

Ogni variabile che cambia nel tempo in un modo non del tutto certo è detta seguire un *processo stocastico*. I processi stocastici possono essere categorizzati come *a tempo discreto* oppure *a tempo continuo*. Nel primo caso il valore della variabile può cambiare soltanto a punti stabiliti nel tempo, mentre, nel secondo, il valore della variabile può cambiare in qualunque momento.

Esistono molti modelli per analizzare il rischio di credito: il loro principale intento è quello di determinare le riserve e il capitale necessario a coprire il rischio di un portfolio di strumenti di credito. La qualità del credito è progressivamente denotata da uno o più indici che riassumono, su di un orizzonte specificato, una probabilità di mancato pagamento del credito, un tasso di perdita data una mancanza nel pagamento, oppure entrambi.

Le banche, i supervisori ed altre istituzioni dipendono da questi sistemi per creare delle accurate e stabili rappresentazioni dei rischi di perdite. Istituzioni finanziarie usano queste informazioni per selezionare un portfolio. Informazioni storiche sulla fluttuazione del valore di credito da un livello di qualità, o da un indice, ad un altro sono usate per giudicare molti modelli che descrivono l'evoluzione probabilistica della qualità del credito.

Il testo è organizzato come segue: nella Sezione 1 vengono presentate le catene di Markov, enunciando la proprietà che le caratterizza e descrivendo il loro funzionamento. Nella Sezione 2 sono descritti i processi di Wiener, processi stocastici vitali per enunciare il comportamento delle azioni. Nella Sezione 3 viene presentato il processo per prezzare le azioni. Nella Sezione 4 è presentato il Modello di Black-Scholes. Infine nella Sezione 5 sono riportate due applicazioni dove, impiegando le catene di Markov, è stato possibile produrre modelli per stilare graduatorie di bontà per i titoli oppure modelli per estrarre la volatilità dal prezzo di un bene.

# 1 Catene di Markov

## 1.1 La proprietà di Markov

Un processo di Markov è una tipologia di processo stocastico dove solo il valore attuale di una variabile è rilevante per predire il futuro. La storia passata della variabile e il modo con il quale il presente è legato al passato sono irrilevanti. Normalmente i prezzi delle azioni sono considerati seguire un processo Markoviano, le previsioni per il futuro non dovrebbero essere influenzate dal prezzo dalla settimana precedente, dal mese passato oppure dall'anno passato. L'unica informazione rilevante è il prezzo odierno<sup>1</sup>. Stime sul futuro non sono certe ed è necessario esprimerle in termini di distribuzioni di probabilità. La proprietà di Markov indica che la distribuzione di probabilità del prezzo ad un particolare istante di tempo non è dipendente dal cammino che il prezzo ha seguito nel passato. È conosciuta anche come *memoryless property* e può essere espressa:

$$P[X(k) = s_k | X(k-1) = s_{k-1}, \dots, X(0) = s_0] = P[X(k) = s_k | X(k-1) = s_{k-1}].$$

identificando con  $k(\geq 0)$  lo stato attuale,  $k-1$  lo stato precedente,  $\{X(k) | k = 0, 1, 2, \dots\}$  la variabile stocastica e, infine, i valori assunti da questa  $(s_0, \dots, s_k)$  sono chiamati stati. In questa formulazione sono considerate solo catene omogenee, dove la probabilità di transizione dipende unicamente dallo stato del sistema immediatamente precedente e non anche dal tempo  $t$ .

La proprietà di Markov per il prezzo delle azioni è congrua con la forma debole di efficienza di mercato: quest'ultima afferma che il prezzo presente di un'azione cattura tutta l'informazione contenuta in un prezzo passato<sup>2</sup>. Se la forma debole di efficienza di mercato non fosse vera, gli analisti potrebbero fare guadagni sopra la media tramite l'interpretazione di grafici della passata storia dei prezzi delle azioni (in realtà sono capaci di farlo).

È la concorrenza nel mercato ad assicurare che la forma debole di efficienza di mercato valga. Ci sono moltissimi investitori che controllano il mercato di azioni

---

<sup>1</sup>Le proprietà statistiche della storia del prezzo dell'azione possono essere utili per determinare le caratteristiche del processo stocastico che il prezzo dell'azione segue. In questo caso il cammino seguito dall'azione nel passato è irrilevante.

<sup>2</sup>I prezzi osservati sul mercato riflettono tutta l'informazione contenuta nella serie storica dei prezzi stessi.

attentamente. Provando a guadagnare da questo si arriva ad una situazione nella quale il prezzo di un'azione, fissato un qualsiasi tempo, riflette l'informazione dei prezzi passati. Basta pensare al caso in cui si scopra un pattern particolare nei prezzi di un'azione. Gli investitori tenterebbero di acquistare l'azione al più presto nel momento in cui il pattern viene individuato. Questo porterebbe ad un immediato incremento nel suo prezzo, e l'effetto dato dall'osservazione sarebbe eliminato, come anche i profitti dati da questo.

## 1.2 Catene di Markov a tempo discreto e continuo

Dopo aver enunciato la proprietà di Markov vengono presentati i processi stocastici che la soddisfano.

Il primo ad essere presentato è la catena di Markov a tempo discreto (o DTMC, acronimo di *discrete-time Markov chain*) dove, come già accennato, i valori assunti dalla variabile  $X(k)$  sono osservazioni fatte a esatti istanti di tempo. L'insieme dei possibili stati, chiamato anche spazio degli stati, è discreto (per motivi di semplicità in questa trattazione lo spazio degli stati è considerato finito).

Per ultimo viene presentata la catena di Markov a tempo continuo (o CTMC, acronimo di *continuous-time Markov chain*). Questa estende, sotto alcuni punti di vista, il modello discreto: mentre ogni transizione tra gli stati in una DTMC corrisponde ad un momento preciso nel tempo, in una CTMC le transizioni possono occorrere in tempo reale. Formalmente, una CTMC è una famiglia di variabili aleatorie  $\{X(t) \mid t \geq 0\}$ , dove  $X(t)$  è un'osservazione fatta nell'istante di tempo  $t$  con  $t \in \mathbb{R}^{\geq 0}$ . Lo spazio degli stati, l'insieme di tutti i possibili valori che  $X(t)$  può assumere, è discreto (sempre per motivi di semplicità, anche in questo caso lo spazio degli stati è considerato finito). La proprietà di Markov può essere riscritta:

**Definizione 1.2.1** Per ogni intero  $k \geq 0$ , sequenze di tempo  $t_0 < t_1 < \dots < t_k$  e stati  $s_0, \dots, s_k$ .

$$P[X(t_k) = s_k \mid X(t_{k-1}) = s_{k-1}, \dots, X(t_0) = s_0] = P[X(t_k) = s_k \mid X(t_{k-1}) = s_{k-1}]$$

Intuitivamente questo significa che la probabilità di compiere una transizione

verso un particolare stato dipende soltanto dallo stato corrente, non dalla storia degli stati passati.

### 1.2.1 Catena di Markov a tempo discreto

In generale, per descrivere una DTMC è sufficiente dare la probabilità di transizione da uno stato ad un altro. Formalmente, una DTMC può essere definita come:

**Definizione 1.2.2** Una DTMC è una tripla  $\mathcal{D} = (S, \bar{s}, P)$  dove:

- $S$  è un insieme finito di stati;
- $\bar{s}$  è lo stato iniziale;
- $P : S \times S \rightarrow [0, 1]$  è una matrice di probabilità di transizione, tale che:

$$\sum_{s' \in S} P(s, s') = 1$$

$$\forall s \in S.$$

Ogni elemento  $P(s, s')$  della matrice di probabilità di transizione indica la probabilità di compiere una transizione da uno stato  $s$  ad uno stato  $s'$  che, per ogni  $k \geq 0$ , è:

$$P(s, s') = P[X(k+1) = s' \mid X(k) = s].$$

La probabilità totale di uscita da uno stato deve essere pari a 1. Ad esempio, quei sistemi che non possono muoversi in un altro stato, possono essere modellati aggiungendo un arco rientrante con probabilità 1.

Per poter analizzare le DTMC è necessario aggiungere altre informazioni. Data una funzione di etichettatura  $L : S \rightarrow 2^{AP}$ , essa associa stati a insiemi di proposizioni atomiche, dall'insieme  $AP$ , che sono utilizzate per etichettare gli stati di particolare interesse. A seconda del tipo di studio che viene compiuto con la catena di Markov, è possibile inoltre annotare la DTMC con una struttura di costo oppure di premio, in pratica una funzione  $C : S \times S \rightarrow \mathbb{R}^{\geq 0}$  che mette in relazione coppie di stati ad un valore reale non negativo, dove il valore  $C(s, s')$  è il costo che deriva da una transizione da  $s$  a  $s'$ . Una struttura dei costi può essere usata per rappresentare informazioni riguardanti il sistema che

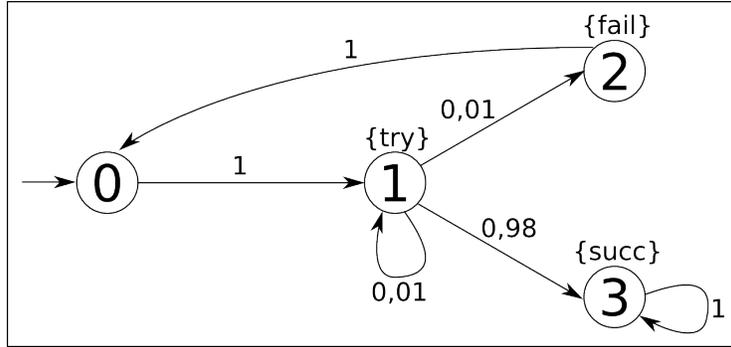


Figura 1: Una DTMC a quattro stati.

la DTMC rappresenta, come ad esempio il consumo energetico oppure il numero di compratori persi.

**Definizione 1.2.3** Una DTMC etichettata è una tripla  $(\mathcal{D}, L, C)$  dove:

- $\mathcal{D} = (S, \bar{s}, P)$  è una DTMC;
- $L : S \rightarrow 2^{AP}$  è la funzione di etichettatura;
- $C : S \times S \rightarrow \mathbb{R}^{\geq 0}$  è la struttura di costo.

La Figura 1 mostra un semplice esempio di una DTMC a quattro stati:  $S = \{0, 1, 2, 3\}$ . Gli stati sono i cerchi e le transizioni sono gli archi, etichettati con le loro probabilità associate. Lo stato iniziale è indicato con la freccia entrante. La matrice di probabilità di transizione  $P$  è:

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0,01 & 0,01 & 0,98 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Le proposizioni atomiche usate per etichettare gli stati sono prese dall'insieme  $\{try, fail, succ\}$ . Questa DTMC descrive un processo che tenta di inviare un messaggio. Dopo il passaggio nello stato 1 (dopo un intervallo di tempo), la catena entra in uno stato dal quale: con probabilità 0.01 attende un intervallo di tempo, con probabilità 0.98 invia con successo il messaggio e con probabilità

0.01 prova ad inviare il messaggio ma fallisce. Nell'ultimo caso il processo riparte. La funzione di etichettatura permette di assegnare nomi significativi agli stati della DTMC:

$$L(1) = \{try\}, L(2) = \{fail\}, L(3) = \{succ\}.$$

**Cammini e misure di probabilità:** Un'esecuzione del sistema che è stato modellato da una DTMC è rappresentata da un cammino. Formalmente un cammino  $\omega$  è una sequenza non vuota di stati  $s_0s_1s_2\dots$  dove  $s_i \in S$  e  $P(s_i, s_{i+1}) > 0 \forall i \geq 0$ . Un cammino può essere sia finito che infinito.  $\omega(i)$  è l' $i$ -esimo stato del cammino  $\omega$ ,  $|\omega|$  è la lunghezza di  $\omega$  (numero di transizioni) e, per un cammino finito  $\omega_{fin}$ ,  $last(\omega_{fin})$  è l'ultimo stato. Un cammino finito  $\omega_{fin}$  di lunghezza  $n$  è un prefisso del cammino infinito  $\omega$  se  $\omega_{fin}(i) = \omega(i)$  per  $0 \leq i \leq n$ . Gli insiemi di tutti i cammini finiti ed infiniti, che partono dallo stato  $s$ , sono indicati rispettivamente da  $Path_s^{fin}$  e  $Path_s$ . A meno che non sia esplicitamente detto, si ha a che fare con cammini infiniti.

Per capire il comportamento probabilistico della DTMC, è necessario determinare la probabilità che determinati cammini hanno di essere percorsi. Questo è possibile definendo, per ogni stato  $s \in S$ , una misura di probabilità  $Prob_s$  sui  $Path_s$ . La misura di probabilità è indotta dalla matrice di probabilità di transizione  $P$ . Per prima cosa per ogni cammino finito  $\omega_{fin} \in Path_s^{fin}$  viene data una probabilità  $P_s(\omega_{fin})$ :

$$P_s(\omega_{fin}) \stackrel{\text{def}}{=} \begin{cases} 1 & \text{se } n = 0 \\ P(\omega(0), \omega(1)) * \dots * P(\omega(n-1), \omega(n)) & \text{altrimenti} \end{cases}$$

dove  $n = |\omega_{fin}|$ .

In secondo luogo viene definito l'insieme cilindrico  $C(\omega_{fin})$ :

$$C(\omega_{fin}) \stackrel{\text{def}}{=} \{\omega \in Path_s \mid \omega_{fin} \text{ è un prefisso di } \omega\}$$

che è un insieme di infiniti cammini con prefisso  $\omega_{fin}$ .

Per quantificare la probabilità che una DTMC si comporti in un modo specificato abbiamo:

$$Prob_s(C(\omega_{fin})) = P_s(\omega_{fin}) \forall \omega_{fin} \in Path_s^{fin}.$$

Considerando ancora la DTMC in Figura 1, sono presenti cinque cammini finiti distinti di lunghezza 3 che partono dallo stato 0. La misura di probabilità degli insiemi cilindrici associati con ognuno di essi è:

- $Prob_0(C(0111)) = 1 * 0.01 * 0.01 = 0.0001$ ;
- $Prob_0(C(0112)) = 1 * 0.01 * 0.01 = 0.0001$ ;
- $Prob_0(C(0113)) = 1 * 0.01 * 0.98 = 0.0098$ ;
- $Prob_0(C(0120)) = 1 * 0.01 * 1 = 0.01$ ;
- $Prob_0(C(0133)) = 1 * 0.98 * 1 = 0.98$ .

### 1.2.2 Catena di Markov a tempo continuo

Una CTMC è definita come segue:

**Definizione 1.2.4** Una CTMC è una tripla  $(S, \bar{s}, R)$  dove:

- $S$  è un insieme finito di stati;
- $\bar{s} \in S$  è lo stato iniziale;
- $R: S \times S \rightarrow \mathbb{R}_{\geq 0}$  è la matrice dei tassi di transizione.

La matrice dei tassi di transizione  $R$  assegna un tasso ad ogni coppia di stati nella CTMC. La transizione tra  $s$  e  $s'$  può avvenire se  $R(s, s') > 0$  e questo valore è la sua probabilità di occorrere. Tipicamente, in uno stato  $s$ , c'è più di uno stato  $s'$  tale che  $R(s, s') > 0$ , questa è conosciuta come la condizione di competizione: la prima transizione a trionfare determina il prossimo stato della CTMC. Il tempo speso nello stato  $s$ , prima che avvenga qualsiasi transizione, è distribuito con un tasso  $E(s)$ :

$$E(s) \stackrel{\text{def}}{=} \sum_{s' \in S} R(s, s')$$

$E(s)$  è detto tasso di uscita dallo stato  $s$ . È possibile anche determinare la probabilità di ogni stato  $s'$  di essere il prossimo stato, indipendentemente dal tempo in cui occorrerà la transizione. Questa è definita con una DTMC.

**Definizione 1.2.5** *La catena di Markov immersa  $emb(\mathcal{C})$  di una CTMC  $\mathcal{C} = (S, \bar{s}, R)$  è una DTMC  $(S, \bar{s}, P)$  dove per  $s, s' \in S$ :*

$$P(s, s') = \begin{cases} R(s, s')/E(s) & \text{se } E(s) \neq 0 \\ 1 & \text{se } E(s) \neq 0 \text{ e } s = s' \\ 0 & \text{altrimenti (ad esempio se non ci sono archi uscenti)} \end{cases}$$

Diversamente dalle DTMC, alle CTMC è permesso non avere archi uscenti, questo tipo di stati sono detti stati assorbenti. Riassumendo: le CTMC possono rimanere in uno stato  $s$  per un intervallo di tempo con tasso  $E(s)$ .

Viene ora definita una matrice supplementare, utile per analizzare la CTMC.

**Definizione 1.2.6** *La matrice generatrice infinitesimale per CTMC  $(S, \bar{s}, R)$  è una matrice  $Q : S \times S \rightarrow \mathbb{R}$  definita come:*

$$Q(s, s') = \begin{cases} R(s, s') & \text{se } s \neq s' \\ -\sum_{s'' \neq s} R(s, s'') & \text{altrimenti} \end{cases}$$

Come per le DMTC, è possibile annotare le CTMC con informazioni addizionali. Esattamente come con le DTMC, usiamo una funzione di etichettatura che mette in relazione, da un insieme AP, stati dell'insieme delle proposizioni atomiche. In più, viene dato un costo al modello ma, per le CTMC, ci sono due tipi di costi: costo istantaneo, associato alle transizioni, e costo cumulativo, associato agli stati. Sono definiti rispettivamente come  $C$  e  $c$ . Il costo istantaneo  $C(s, s')$  viene considerato ogni volta che avviene una transizione tra uno stato  $s$  e  $s'$ . Il costo cumulativo  $c(s)$  è il tasso di costo unitario per lo stato  $s$ ; ad esempio c'è un costo  $t * c(s)$  se la CTMC rimane nello stato  $s$  per  $t$  unità di tempo.

**Definizione 1.2.7** *Una CTMC etichettata è una quadrupla  $(\mathcal{C}, L, C, c)$  dove:*

- $\mathcal{C} = (S, \bar{s}, R)$  è una CTMC;
- $L : S \rightarrow 2^{AP}$  è una funzione di etichettatura;
- $C : S \times S \rightarrow \mathbb{R}^{\geq 0}$  è la funzione di costo istantaneo;
- $c : S \rightarrow \mathbb{R}^{\geq 0}$  è la funzione di costo cumulativo.

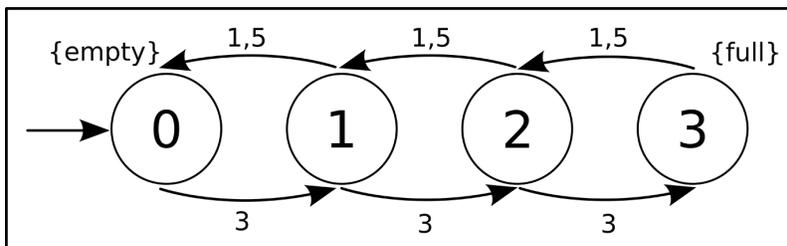


Figura 2: Una CTMC a quattro stati.

Nella Figura 2 viene riportata una semplice CTMC con quattro stati  $\{0, 1, 2, 3\}$ . La notazione grafica è identica a quella per le DTMC eccetto per le transizioni, che adesso sono etichettate con tassi e non probabilità. Il modello descritto da questa CTMC è una coda dove lo stato  $i$ -esimo indica che ci sono  $i$  posizioni della coda occupate. Inizialmente la coda è vuota ed ha una dimensione massima di 3. Con tasso 3 un posto viene occupato mentre con tasso 1.5 viene liberato.

Le matrici che descrivono, secondo quello esposto precedentemente, sono:

$$R = \begin{pmatrix} 0 & 1.5 & 0 & 0 \\ 3 & 0 & 1.5 & 0 \\ 0 & 3 & 0 & 1.5 \\ 0 & 0 & 3 & 0 \end{pmatrix}, Q = \begin{pmatrix} -1.5 & 1.5 & 0 & 0 \\ 3 & -4.5 & 1.5 & 0 \\ 0 & 3 & -4.5 & 1.5 \\ 0 & 0 & 3 & -3 \end{pmatrix},$$

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 \\ \frac{2}{3} & 0 & \frac{1}{3} & 0 \\ 0 & \frac{2}{3} & 0 & \frac{1}{3} \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

**Cammini e misure di probabilità:** Un cammino in una CTMC è una sequenza  $s_0 t_0 s_1 t_1 s_2 \dots$  dove  $R(s_i, s_{i+1}) > 0$  e  $t_i \in \mathbb{R}_{>0}$  per ogni  $i \geq 0$ . Il valore  $t_i$  rappresenta il totale del tempo speso nello stato  $s_i$ . Come con le DTMC, indichiamo con  $\omega(i)$  lo stato  $i$ -esimo del cammino  $\omega$ , in altre parole  $s_i$ . Inoltre  $time(\omega, i)$  indica il totale del tempo speso nello stato  $s_i$ , che è  $t_i$ , e  $\omega @ t$  lo stato occupato al tempo  $t$ .

$Path_s$  è l'insieme di tutti i cammini infiniti che partono dallo stato  $s$ . La probabilità  $Prob_s$ , su  $Path_s$ , è definita come segue:

**Definizione 1.2.8** Se gli stati  $s_0, \dots, s_n \in S$  soddisfano  $R(s_i, s_{i+1}) > 0$  per ogni  $0 \leq i < n$  e  $I_0, \dots, I_{n-1}$  sono intervalli non vuoti contenenti valori appartenenti a  $\mathbb{R}_{\geq 0}$ , l'insieme cilindrico  $C(s_0, I_0, \dots, i_{n-1}, s_n)$  contiene tutti gli infiniti cammini  $s'_0 t_0 s'_1 t_1 s'_2 \dots$  dove  $s_i = s'_i$  per  $i \leq n$  e  $t_i \in I_i$  per  $i < n$ .

La misura di probabilità  $Prob_s$  è descritta in modo induttivo da  $Prob_s(C(s_0)) = 1$  e:

$$\begin{aligned} Prob_s(C(s_0, \dots, s_n, I_n, s_{n+1})) &= \\ &= Prob_s(C(s_0, \dots, s_n)) \cdot P(s_n, s_{n+1}) \cdot (e^{-\inf I_n \cdot E(s_n)} - e^{-\sup I_n \cdot E(s_n)}) \end{aligned}$$

dove  $P$  è la matrice di probabilità di transizione della catena di Markov immersa.

Prendendo la CTMC in Figura 2 e la sequenza di stati e intervalli  $0, [0, 2], 1$  ( $s_0 = 0, I_0 = [0, 2]$  e  $s_1 = 1$ ) e considerando la misura di probabilità dello stato iniziale emerge, per l'insieme cilindrico  $C(0, [0, 2], 1)$ :

$$Prob_0(C(0, [0, 2], 1)) = 1 \cdot P(0, 1) \cdot (e^{-0 \cdot E(0)} - e^{-2 \cdot E(0)}) = 1 \cdot 1 \cdot (e^0 - e^{-3}) = 1 - e^{-3}.$$

Questo significa che la probabilità di lasciare lo stato iniziale e passare allo stato 1 entro le prime due unità di tempo è circa 0.950213.

Prendendo una variabile che segue un processo stocastico di Markov, supponendo che questa abbia come valore iniziale 10 e che il cambiamento nel suo valore durante un anno sia  $\phi(0, 1)$ , dove  $\phi(\mu, \sigma)$  denota una distribuzione di probabilità che è normalmente distribuita con media  $\mu$  e deviazione standard  $\sigma$ , è possibile dedurre quale sia la distribuzione di probabilità del cambiamento nel suo valore durante due anni. Il cambiamento in due anni è la somma di due distribuzioni normali, ognuna delle quali ha una media pari a zero e deviazione standard 1. Dato che la variabile è Markoviana, le due distribuzioni di probabilità sono indipendenti. Sommare due distribuzioni normali indipendenti porta come risultato una distribuzione normale dove la media è la somma delle medie e la varianza è la somma delle varianze<sup>3</sup>. La media di fluttuazione durante due anni nella variabile che è presa in considerazione è quindi zero, mentre la varianza è 2. La fluttuazione nella variabile è esprimibile nella forma  $\phi(0, \sqrt{2})$ .

<sup>3</sup>da ricordare che la varianza di una distribuzione di probabilità è il quadrato della sua deviazione standard. La varianza di fluttuazione di un anno nel valore di una variabile che consideriamo ergo è 1.

Adesso lo stesso ragionamento si può riproporre considerando la fluttuazione nella variabile durante sei mesi. La varianza di fluttuazione nel suo valore durante un anno è uguale alla varianza di fluttuazione nei primi sei mesi più la varianza di fluttuazione nei seguenti sei mesi (sono assunte come equivalenti). Ne segue che la varianza di fluttuazione durante un periodo di sei mesi debba essere 0.5. Allo stesso modo, la deviazione standard di fluttuazione è  $\sqrt{0.5}$ , quindi la distribuzione di probabilità della fluttuazione nel valore di una variabile durante sei mesi è  $\phi(0, \sqrt{0.5})$ .

È possibile fare un analogo ragionamento per mostrare la fluttuazione nel valore di una variabile durante un periodo di tre mesi: questo valore è  $\phi(0, \sqrt{0.25})$ . In generale la fluttuazione durante un qualsiasi periodo di tempo di lunghezza  $T$  è  $\phi(0, \sqrt{T})$ . In particolare, la fluttuazione durante un periodo di tempo molto corto di lunghezza  $\delta t$  è  $\phi(0, \sqrt{\delta t})$ .

La presenza della radice quadrata in questi risultati può apparire strana. Questo perché, quando viene considerato un processo di Markov, la varianza delle fluttuazioni in periodi di tempo successivi è additiva. Le deviazioni standard delle fluttuazioni in periodi di tempo successivi non sono additive. La varianza della fluttuazione in una variabile nell'esempio riportato è 1 per anno, così la varianza di fluttuazione, secondo quello descritto, è pari a 2 per due anni e così via. La deviazione standard della fluttuazione per due e tre anni è rispettivamente  $\sqrt{2}$  e  $\sqrt{3}$ . Quindi, non c'è da riferirsi alla deviazione standard della variabile come 1 per anno, ma dovrebbe essere "1 per la radice quadrata degli anni". Il risultato illustra come mai l'incertezza sia spesso riportata come proporzionale alla radice quadrata del tempo.

### 1.3 Catene di Markov nascoste

Esistono anche varianti dei due modelli presentati. Come ultima osservazione sulle catene di Markov vengono riportate le catene di Markov nascoste: un modello statistico dove il sistema modellato viene assunto essere una catena di Markov a tutti gli effetti ma con stati non osservabili. Una catena di Markov nascosta può essere considerata come la più semplice rete dinamica Bayesiana<sup>4</sup>.

<sup>4</sup>Una rete dinamica Bayesiana è una rete che rappresenta una sequenza di variabili. Queste sequenze sono spesso serie temporali (ad esempio sistemi per il riconoscimento vocale, dei

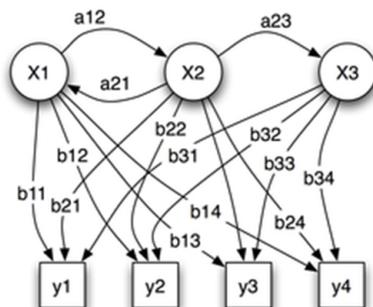


Figura 3: Esempio di una catena di Markov nascosta:

- x: stati
- y: possibili osservazioni
- a: probabilità di transizione di stato
- b: probabilità di output

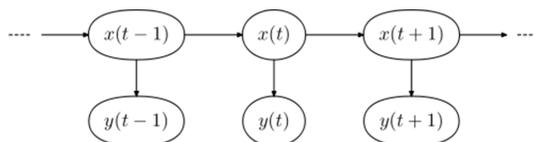


Figura 4: Diagramma dell'architettura generale di una catena di Markov nascosta.

Nei modelli di catena di Markov visti fino ad ora lo stato è direttamente visibile all'osservatore del fenomeno descritto da questa e, conseguentemente, le probabilità di transizione tra stati sono gli unici parametri. In una catena di Markov nascosta lo stato non è direttamente visibile, ma l'output dipende dallo stato visibile. Ogni stato ha una distribuzione di probabilità sui possibili output e la sequenza di output generati dalla catena nascosta dà informazioni riguardanti la sequenza di stati (l'aggettivo "nascosta" si riferisce agli stati nei quali il modello passa, non ai parametri del modello; anche se questi sono conosciuti esattamente, il modello rimane "nascosto").

Il diagramma in Figura 4 mostra l'architettura di un'istanza di una catena di Markov nascosta (ad esempio, della scrittura manuale e in bioinformatica) oppure sequenze di simboli (sequenze di proteine).

di Markov nascosta. Ogni ovale rappresenta una variabile stocastica che può assumere qualsiasi valore. La variabile casuale  $x(t)$  è lo stato nascosto al tempo  $t$  e la variabile casuale  $y(t)$  è un'osservazione al tempo  $t$  (considerando il modello riportato in Figura 3:  $x(t) \in \{x_1, x_2, x_3\}$  e  $y(t) \in \{y_1, y_2, y_3, y_4\}$ ). Le frecce nel diagramma (chiamato anche diagramma a *traliccio*<sup>5</sup>) indicano dipendenze condizionali.

Dal diagramma è chiaro che la distribuzione di probabilità condizionale della variabile nascosta  $x(t)$  al tempo  $t$ , dato il valore della variabile nascosta  $x(t-1)$ , dipende solo dal valore della variabile nascosta  $x(t-1)$ : i valori al tempo  $t-2$  e precedenti non hanno influenza (Sezione 1.1). In modo simile il valore della variabile osservata  $y(t)$  dipende soltanto dal valore della variabile nascosta  $x(t)$  (entrambe al tempo  $t$ ).

### 1.3.1 Probabilità di una sequenza osservata e uso del modello

La probabilità di osservare una sequenza

$$Y = y(0), y(1), \dots, y(L-1)$$

di lunghezza  $L$  è data da

$$P(Y) = \sum_X P(Y|X)P(X)$$

dove la sommatoria si muove su tutte le possibili sequenze di nodi nascosti

$$X = x(0), x(1), \dots, x(L-1).$$

Per la maggior parte dei problemi quotidiani, il calcolo di  $P(Y)$  tramite forza bruta è intrattabile, poiché il numero di possibili sequenze di nodi nascosti è tipicamente molto alto e scala esponenzialmente in base alla lunghezza della sequenza. Il calcolo può comunque essere velocizzato enormemente utilizzando algoritmi di programmazione dinamica<sup>6</sup>.

<sup>5</sup>Un traliccio è un grafico dove i nodi sono ordinati secondo fette verticali e ogni nodo, allo stesso tempo, è connesso ad almeno un nodo precedente e ad almeno un nodo successivo. Nel traliccio il primo e l'ultimo elemento hanno solo un nodo. I trallicci sono usati anche come encoders e decoders in teoria della comunicazione e nella crittografia.

<sup>6</sup>La programmazione dinamica è una metodologia di risoluzione di problemi complessi spezzandoli in sottoproblemi più semplici. È applicabile a problemi che esibiscono una sovrapposizione di sottoproblemi. Quando è applicabile, il metodo dinamico impiega molto meno tempo del metodo originale senza ottimizzazioni.

I problemi canonici associati alle catene di Markov nascoste sono tre:

- Dato un parametro del modello, calcolare la probabilità di una particolare sequenza in output.
- Dati i parametri di un modello e una particolare sequenza di output, trovare la sequenza di stati che in modo più probabile abbia generato la sequenza di output data.
- Data una sequenza di output o un insieme di queste sequenze, trovare l'insieme di transizioni da stati e probabilità di output.

## 2 Processi di Wiener

I processi di Wiener sono processi stocastici a tempo continuo con media zero e varianza uno; questi giocano un ruolo importante sia in matematica applicata che teorica. In matematica teorica, i processi di Wiener hanno indotto allo studio delle martingale<sup>7</sup> a tempo continuo. Sono processi chiave poiché altri processi, più complicati, possono essere descritti in termini di questi. In matematica applicata i processi di Wiener sono molto utili per dare modelli di rumore, in ingegneria elettronica per filtrare gli errori degli strumenti.

Ci sono molte applicazioni di questi processi in tutte le scienze matematiche ed una di queste, dove sono molto importanti, è la teoria della finanza, in particolare il modello di prezzaggio delle opzioni di Black-Scholes.

La variabile presente nell'esempio descritto per ultimo nella Sezione 1.2.2 segue un processo di Wiener. Formalmente una variabile  $z$  segue un processo di Wiener se soddisfa le seguenti due proprietà:

**Proprietà 2.0.1** *La fluttuazione  $\delta z$  durante un breve intervallo di tempo  $\delta t$  è*

$$\delta z = \epsilon \sqrt{\delta t} \tag{1}$$

dove  $\epsilon$  è una rappresentazione casuale da una distribuzione standardizzata normale  $\Phi(0, 1)$ .

**Proprietà 2.0.2** *Per un qualsiasi breve intervallo di tempo  $\delta t$ , si ha un relativo valore di  $\delta z$ .*

Segue dalla prima proprietà, dato che  $\delta z$  ha una distribuzione normale con:

- media di  $\delta z = 0$ ;
- deviazione standard di  $\delta z = \sqrt{\delta t}$ ;
- varianza di  $\delta z = \delta t$ .

La seconda proprietà implica che  $z$  segua un processo di Markov.

---

<sup>7</sup>Una martingala è un processo stocastico  $X_t$  indicizzato da un parametro crescente  $t$  (spesso il tempo), con la seguente proprietà: per ogni  $s \leq t$ , l'attesa di  $X_t$  condizionata rispetto ai valori di  $X_r$ ,  $r \leq s$ , è uguale ad  $X_s$ .

Considerando l'incremento nel valore di  $z$  durante un relativamente lungo intervallo di tempo  $T$ , questo può essere rappresentato da  $z(T) - z(0)$ . È possibile considerarlo come la somma di incrementi in  $z$  in  $N$  piccoli intervalli di lunghezza  $\delta t$ , dove:

$$N = \frac{T}{\delta t}$$

di conseguenza

$$z(T) - z(0) = \sum_{i=1}^N \epsilon_i \sqrt{\delta t} \quad (2)$$

dove  $\epsilon_i$  ( $i = 1, 2, \dots, N$ ) sono rappresentazioni casuali da  $\Phi(0, 1)$ . Dalla seconda proprietà dei processi di Wiener, le  $\epsilon_i$  sono indipendenti l'una dalle altre. Segue dall'equazione (2) che  $z(T) - z(0)$  è normalmente distribuita con:

- media di  $[z(T) - z(0)] = 0$ ;
- varianza di  $[z(T) - z(0)] = N\delta t = T$ ;
- deviazione standard di  $[z(T) - z(0)] = \sqrt{T}$

Questo è consistente con ciò che è stato discusso all'inizio di questa Sezione.

Volendo fare un esempio, supposto che il valore di  $z$  (variabile che segue un processo di Wiener) sia inizialmente 25 e il tempo sia misurato in anni, alla fine di un anno il valore della variabile è normalmente distribuito con una media di 25 e una deviazione standard di 1. Alla fine di cinque anni, è normalmente distribuito con una media di 25 e deviazione standard di  $\sqrt{5}$ . L'incertezza circa il valore della variabile ad un determinato periodo nel futuro, misurata dalla deviazione standard, aumenta secondo la radice quadrata di quanto lontano nel tempo vogliamo osservare.

Nel calcolo ordinario è norma procedere da piccoli intervalli al limite fino a che i piccoli intervalli si avvicinano allo zero, quindi  $\frac{\delta y}{\delta x}$  diventa  $\frac{dy}{dx}$  al limite. È possibile agire in modo simile anche quando si ha a che fare con i processi stocastici. Un processo di Wiener è il limite  $\delta t \rightarrow 0$  del processo descritto per  $z$  nell'esempio.

Le Figure 5, 6 e 7 mostrano cosa accade al cammino seguito da  $z$  all'avvicinarsi del limite  $\delta t \rightarrow 0$ . È da notare che il cammino è leggermente "dentellato", questo perché la dimensione di un movimento in  $z$  nel tempo  $\delta t$  è proporzionale a  $\sqrt{\delta t}$  e, quando  $\delta t$  è piccolo,  $\sqrt{\delta t}$  è molto più grande di  $\delta t$ . Due affascinanti

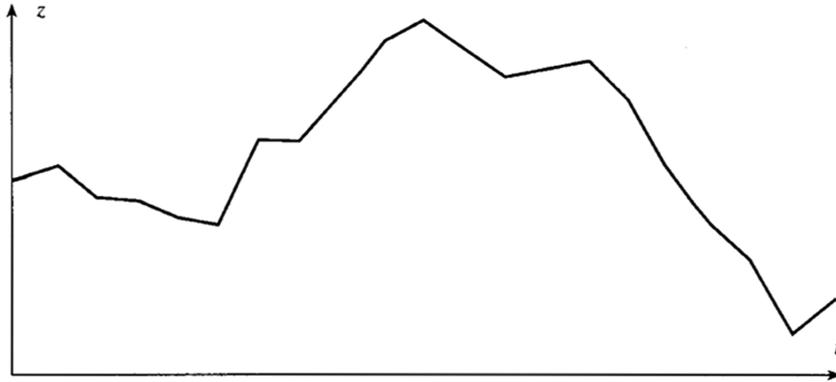


Figura 5: Valore di  $\delta t$  relativamente grande.

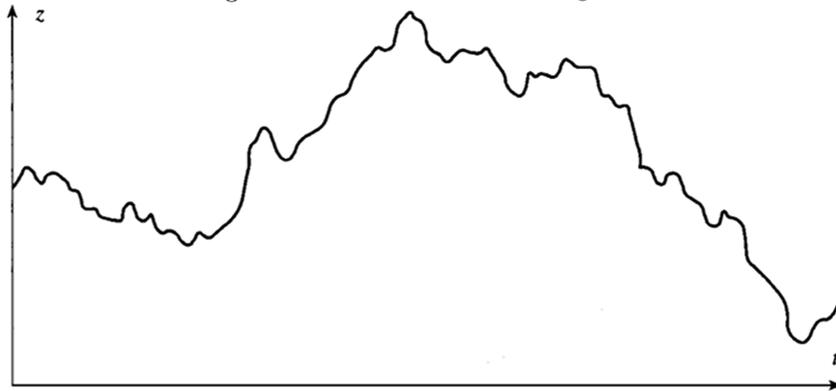


Figura 6: Valore di  $\delta t$  più piccolo.

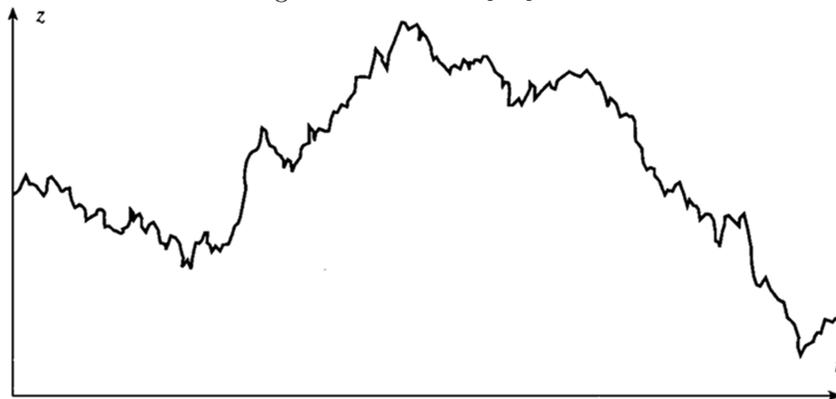


Figura 7: Il vero processo ottenuto con  $\delta t \rightarrow 0$ .

proprietà dei processi di Wiener, legate alla proprietà di  $\sqrt{\delta t}$  appena accennata, sono:

1. La lunghezza presunta del cammino seguito da  $z$  in qualsiasi intervallo di tempo è infinito.
2. È infinito il presunto numero di volte in cui  $z$  è uguale ad un qualsiasi valore fissato, in qualsiasi intervallo di tempo.

## 2.1 Processi di Wiener generalizzati

Il processo di Wiener base,  $dz$ , che è stato sviluppato fino ad adesso ha un tasso di slittamento pari a zero e un tasso di varianza pari a 1. Il tasso di slittamento pari a zero significa che il valore atteso di  $z$  ad ogni intervallo futuro di tempo è uguale al suo valore corrente. Il tasso di varianza pari a 1 significa che la varianza del cambiamento in  $z$  in un intervallo di tempo pari a  $T$  equivale a  $T$ . Un processo di Wiener generalizzato per una variabile  $x$  può essere definito in termini di  $dz$  come segue:

$$dx = a dt + b dz \quad (3)$$

dove  $a$  e  $b$  sono costanti.

Per capire al meglio l'equazione (3), è utile considerare i due componenti del membro destro separatamente.  $a dt$  implica che  $x$  abbia un previsto tasso di slittamento di  $a$  unità di tempo. Senza il termine  $b dz$ , l'equazione diventa:

$$dx = a dt$$

che implica

$$\frac{dx}{dt} = a$$

Integrando rispetto al tempo viene:

$$x = x_0 + at$$

dove  $x_0$  è il valore di  $x$  al tempo zero. In un intervallo di tempo di lunghezza  $T$ , il valore di  $x$  aumenta di un valore  $aT$ .  $b dz$  può essere visto come del rumore per il cammino seguito da  $x$ . Il valore di questo rumore è  $b$  volte un processo di Wiener con una deviazione standard pari a  $b$ . In intervalli di tempo piccoli  $\delta t$ , il cambiamento di  $\delta x$  nel valore  $x$  è dato dalle equazioni (1) e (3):

$$\delta x = a \delta t + b \epsilon \sqrt{\delta t}$$

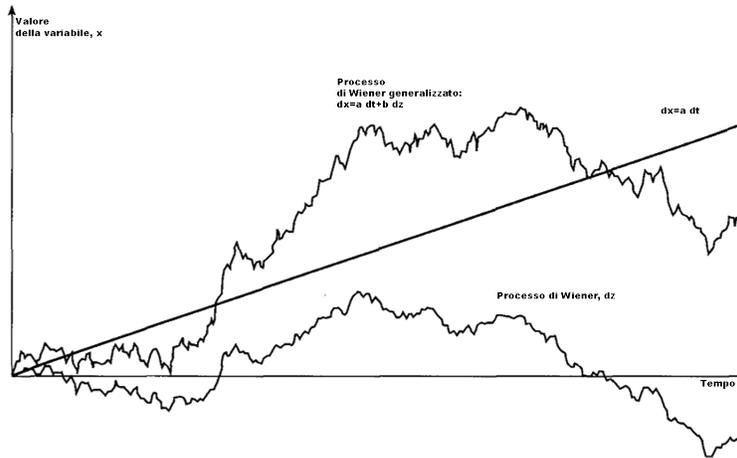


Figura 8: Processo di Wiener generalizzato:  $a = 0.3$ ,  $b=1.5$ .

dove, come prima,  $\epsilon$  è una rappresentazione casuale proveniente dalla distribuzione normale standardizzata. In più  $\delta x$  ha una distribuzione normale con:

- media di  $\delta x = a\delta t$ ;
- deviazione standard di  $\delta x = b\sqrt{\delta t}$ ;
- varianza di  $\delta x = b^2\delta t$ .

Parametri simili a quelli dati per un processo di Wiener mostrano che il cambiamento nel valore di  $x$ , in qualsiasi intervallo di tempo  $T$ , segue una distribuzione normale con parametri:

- media della fluttuazione in  $x = aT$ ;
- deviazione standard della fluttuazione in  $x = b\sqrt{T}$ ;
- varianza della fluttuazione in  $x = b^2T$ .

Inoltre, il processo di Wiener generalizzato dato nell'equazione (3) ha un previsto tasso di slittamento (in questo caso lo slittamento medio per unità di tempo) di  $a$  e un tasso di varianza (in questo caso la varianza per unità di tempo) di  $b^2$ .

Supposta una situazione dove il capitale di una compagnia, misurato in migliaia di dollari, segua un processo di Wiener generalizzato con uno slittamento

di 20 per anno e tasso di varianza di 900 per anno, inizialmente il capitale è 50. Alla fine di un anno il capitale sarà normalmente distribuito con una media di 70 e una deviazione standard di 30 ( $\sqrt{900}$ ). Alla fine di sei mesi sarà normalmente distribuito con una media di 60 e una deviazione standard di  $30\sqrt{0.5}$ . L'incertezza sul capitale in un certo istante nel futuro, come misurato dalla sua deviazione standard, aumenterà seguendo la radice quadrata di quanto lontano si voglia guardare. Il capitale può diventare negativo (è interpretato come una situazione nella quale la compagnia richieda dei prestiti).

## 2.2 Processi di Itô

Possono essere definiti un ulteriore tipo di processi stocastici: i processi di Itô. Questi sono dei processi di Wiener generalizzati nei quali i parametri  $a$  e  $b$  sono funzioni del valore della variabile chiave  $x$  e del tempo  $t$ . Algebricamente un processo di Itô può essere scritto:

$$dx = a(x, t)dt + b(x, t)dz. \quad (4)$$

Entrambi i previsti tassi di slittamento e di varianza nei processi di Itô sono suscettibili di cambiamento nel tempo. In un piccolo intervallo di tempo tra  $t$  e  $t + \delta t$ , la variabile cambia da  $x$  a  $x + \delta x$ , dove:

$$\delta x = a(x, t)\delta t + b(x, t)\epsilon\sqrt{\delta t}.$$

Questa relazione comprende una piccola approssimazione: viene assunto che il tasso di slittamento e di varianza di  $x$  rimanga costante, uguale a  $a(x, t)$  e  $b(x, t)^2$  rispettivamente, durante l'intervallo di tempo che va da  $t$  a  $t + \delta t$ .

### 3 Il processo per il prezzaggio delle azioni

Si può pensare, dopo ciò che è stato riportato, che il prezzo delle azioni (considerando le azioni che non pagano dividendi) segua un processo di Wiener generalizzato che abbia un previsto tasso di slittamento ed un tasso di varianza costanti. Il modello fallisce nel catturare un aspetto chiave dei prezzi delle azioni: la prevista percentuale di ritorno richiesta dall'investitore da un'azione è indipendente dal prezzo dell'azione. Se gli investitori richiedono un 14% annuo come previsto ritorno quando il prezzo dell'azione è a 10, allora, stando così le cose, richiederanno lo stesso un 14% annuo quando avrà prezzo 50.

L'assunzione di un costante tasso di slittamento previsto è inappropriata e necessita di essere sostituita dall'assunzione che il ritorno previsto (in questo caso lo slittamento previsto diviso il prezzo dell'azione) sia costante. Se  $S$  è il prezzo di un'azione al tempo  $t$ , il tasso di slittamento previsto in  $S$  potrebbe essere preso come  $\mu S$  per qualche parametro  $\mu$ . Questo significa che in un piccolo intervallo di tempo  $\delta t$ , l'aumento previsto in  $S$  è pari a  $\mu S \delta t$ . Il parametro  $\mu$  è il previsto tasso di ritorno sull'azione, espresso in decimali.

Se la volatilità del prezzo dell'azione è sempre zero, questo modello implica che:

$$\delta S = \mu S \delta t$$

Al limite, con  $\delta t \rightarrow 0$

$$dS = \mu S dt$$

oppure

$$\frac{dS}{S} = \mu dt$$

Integrando tra zero ed il tempo  $T$

$$S_T = S_0 e^{\mu T} \tag{5}$$

dove  $S_0$  e  $S_T$  sono il prezzo dell'azione al tempo zero ed al tempo  $T$ . L'equazione (5) mostra che, quando il tasso di varianza è zero, il prezzo dell'azione cresce ad un tasso continuamente composto di  $\mu$  per unità di tempo.

In realtà, chiaramente, un prezzo di un'azione esibisce volatilità. Un'assunzione ragionevole è che la variabilità della percentuale di ritorno in un breve periodo di tempo  $\delta t$  sia la stessa senza considerare il prezzo dell'azione. In altre

parole, un investitore è incerto della percentuale di ritorno quando il prezzo dell'azione è 50 tanto quando è 10. Questo suggerisce che la deviazione standard della fluttuazione in un breve periodo di tempo  $\delta t$  potrebbe essere proporzionale al prezzo dell'azione e conduce al modello

$$dS = \mu S dt + \sigma S dz$$

oppure

$$\frac{dS}{S} = \mu dt + \sigma dz \quad (6)$$

L'equazione (6) è il modello più largamente usato per il comportamento del prezzo dell'azione. La variabile  $\sigma$  è la volatilità del prezzo dell'azione e la variabile  $\mu$  è il suo tasso di ritorno previsto.

Si consideri un'azione che non paghi dividendi con una volatilità del 30% annua e che fornisca un ritorno previsto del 15% annuo con composizione continua. In questo caso i parametri sono  $\mu = 0.15$  e  $\sigma = 0.30$ , quindi il processo per il prezzo dell'azione è

$$\frac{dS}{S} = 0.15 dt + 0.30 dz$$

Se  $S$  è il prezzo dell'azione ad un istante di tempo e  $\delta S$  è l'incremento nel prezzo dell'azione nel successivo piccolo intervallo di tempo, quindi:

$$\frac{\delta S}{S} = 0.15 \delta t + 0.30 \epsilon \sqrt{\delta t}$$

dove  $\epsilon$  è una rappresentazione casuale da una distribuzione normale standardizzata. Considerando, per ultimo, un periodo di tempo di una settimana (0.0192 anni) ed un prezzo iniziale per l'azione di 100 risulta che  $\delta t = 0.0192$ ,  $S = 100$  e

$$\delta S = 100(0.00288 + 0.0416\epsilon)$$

oppure

$$\delta S = 0.288 + 4.16\epsilon$$

mostrando che l'incremento del prezzo è una rappresentazione casuale da una distribuzione normale con media 0.288 e deviazione standard 4.16.

### 3.1 Analisi del modello

Il modello del comportamento del prezzo dell'azione emerso è conosciuto come *moto geometrico Browniano*<sup>8</sup>. La versione a tempo discreto del modello è

$$\frac{\delta S}{S} = \mu \delta t + \sigma \epsilon \sqrt{\delta t} \quad (7)$$

oppure

$$\delta S = \mu S \delta t + \sigma S \epsilon \sqrt{\delta t} \quad (8)$$

La variabile  $\delta S$  è la fluttuazione nel prezzo dell'azione  $S$  in un piccolo intervallo di tempo  $\delta t$  e  $\epsilon$  è una rappresentazione casuale da una distribuzione normale standardizzata (una distribuzione normale con una media di zero e deviazione standard pari a 1). Il parametro  $\mu$  è il tasso di ritorno previsto per unità di tempo dall'azione ed il parametro  $\sigma$  è la volatilità del prezzo dell'azione. Entrambi questi parametri sono considerati costanti. Il membro sinistro dell'equazione (7) è il ritorno dato dall'azione in un breve periodo di tempo  $\delta t$ . Il termine  $\mu \delta t$  è il valore di ritorno previsto mentre il termine  $\sigma \sqrt{\delta t}$  è la parte stocastica del ritorno. La varianza della componente stocastica (e di conseguenza di tutto il ritorno) è  $\sigma^2 \delta t$ .

L'equazione (7) mostra che  $\frac{\delta S}{S}$  è normalmente distribuito con media  $\mu \delta t$  e deviazione standard  $\sigma \sqrt{\delta t}$ . In altre parole

$$\frac{\delta S}{S} \sim \Phi(\mu \delta t, \sigma \sqrt{\delta t}) \quad (9)$$

---

<sup>8</sup>Il moto browniano è quello osservato su singole particelle pesanti (dell'ordine del micrometro) presenti in fluidi o sospensioni fluide.

Quando un fluido si trova all'equilibrio termodinamico si potrebbe pensare che le molecole che lo compongono siano essenzialmente ferme o che comunque vibrino attorno alla loro posizione di equilibrio per effetto della temperatura. Se però si osserva il moto di un tale fluido, ad esempio disperdendovi delle particelle colorate molto leggere ed osservandone il movimento, si nota che queste sono tutt'altro che a riposo. Quello che si osserva è che ciascuna particella segue un moto assolutamente disordinato la cui natura appare essere indipendente dalla natura della particella stessa.

Questo è dovuto al fatto che la particella in questione subisce un gran numero di urti da parte delle molecole del fluido in cui è immersa. Quanto più piccole sono le particelle tanto più rapido è il moto browniano.

La matematica del moto browniano utilizzata nell'ambito della finanza differisce da quella comunemente utilizzata in ambito fisico; in finanza si utilizzano per lo più il calcolo stocastico basato sul lemma di Itô. Applicazioni numeriche nel prezzaggio dei prodotti finanziari spesso ricorrono a metodi di simulazione Monte Carlo.

### 3.1.1 Simulazioni Monte Carlo

Una simulazione di Monte Carlo di un processo stocastico è una procedura per ricavare dei campioni random dai processi, che verrà impiegata per comprendere la natura del processo del prezzo delle azioni nell'equazione (6).

Facendo una supposizione che il ritorno previsto da un'azione sia il 14% annuo e che la deviazione standard del ritorno (in questo caso la volatilità) sia del 20% annuo, allora il modello avrà  $\mu = 0.14$  e  $\sigma = 0.20$ . Con  $\delta t = 0.01$  vengono considerate fluttuazioni nel prezzo dell'azione in intervalli di tempo di lunghezza 0.01 anni (3.65 giorni). Dall'equazione (8) si nota:

$$\delta S = 0.14 \times 0.01S + 0.2\sqrt{0.01}S\epsilon$$

oppure

$$\delta S = 0.0014S + 0.02S\epsilon \tag{10}$$

Un cammino per il prezzo dell'azione può essere simulato ricavando campioni ripetutamente di  $\epsilon$  da  $\Phi(0, 1)$  e sostituendolo nell'equazione (10); la Tabella 1 mostra un insieme di risultati di questa simulazione.

Prezzo dell'azione all'inizio del periodo	Campione casuale per $\epsilon$	Fluttuazione nel prezzo dell'azione durante il periodo
20.000	0.52	0.236
20.236	1.44	0.611
20.847	-0.86	-0.329
20.518	1.46	0.628
21.146	-0.69	-0.262
20.883	-0.74	-0.280
20.603	0.21	0.115
20.719	-1.10	-0.427
20.292	0.73	0.325
20.617	1.16	0.507
21.124	2.56	1.111

Tabella 1: Simulazione del prezzo di un'azione con parametri  $\mu = 0.14$  e  $\sigma = 0.20$  durante un periodo di lunghezza 0.01 anni.

Viene preso 20 come valore iniziale del prezzo dell'azione. Per il primo periodo,  $\epsilon$  è registrato come 0.52 e, per l'equazione (10), la fluttuazione durante il primo periodo è:

$$\delta S = 0.0014 \times 20 + 0.02 \times 20 \times 0.52 = 0.236$$

Di conseguenza, all'inizio del secondo periodo, il prezzo dell'azione è 20.236. Il valore di  $\epsilon$  registrato nel seguente periodo corrisponde a 1.44. Sempre per l'equazione (10), la fluttuazione durante il secondo periodo è pari a:

$$\delta S = 0.0014 \times 20.236 + 0.02 \times 20.236 \times 1.44 = 0.611$$

Conseguentemente, all'inizio del periodo successivo, il valore dell'azione è pari a 20.847 e così via. Da notare che, siccome il processo simulato è markoviano, i campioni per  $\epsilon$  sono indipendenti l'uno dall'altro.

La Tabella 1 mostra soltanto una possibile sequenza di movimenti del prezzo dell'azione. Campioni casuali differenti possono portare a differenti movimenti nei prezzi. Qualsiasi piccolo intervallo  $\delta t$  può essere usato nella simulazione; al limite  $\delta t \rightarrow 0$  viene ottenuta una perfetta descrizione del processo stocastico. Il prezzo finale dell'azione di 21.124 nella Tabella 1 può essere considerato come un campione casuale dalla distribuzione dei prezzi dell'azione alla fine di 10 intervalli di tempo (in questo caso alla fine di un decimo di anno). Ripetendo la simulazione dei movimenti nel prezzo delle azioni, come quello fatto in Tabella 1, viene delineato un completo profilo della distribuzione di probabilità per il prezzo dell'azione alla fine di questo tempo.

### 3.2 I parametri

Il processo per il prezzo dell'azione mostrato, come accennato, coinvolge due parametri:  $\mu$  e  $\sigma$ . Il parametro  $\mu$  è il previsto ritorno continuativamente composto guadagnato dall'investitore per anno. La maggior parte degli investitori richiedono ritorni previsti più alti che inducono ad accettare rischi più alti. Segue che il valore di  $\mu$  dovrebbe dipendere dal rischio del ritorno dall'azione.<sup>9</sup> e anche dal tasso di interesse nell'economia. Più alto è il tasso di interesse, più alto sarà il ritorno previsto dall'azione.

Il parametro  $\sigma$ , la volatilità del prezzo dell'azione, è, per contrasto, importante per la determinazione del valore della maggior parte dei derivati. Tipicamente, i valori di  $\sigma$  per un'azione sono nell'intervallo  $[0.20, 0.50]$  ([20%, 50%]).

La deviazione standard della fluttuazione proporzionale nel prezzo dell'azione in un piccolo intervallo di tempo  $\delta t$  è  $\sigma\sqrt{\delta t}$ . Come approssimazione grezza, la

<sup>9</sup> $\mu$  è condizionato da un rischio che non può essere diversificato dall'investitore.

deviazione standard della fluttuazione proporzionale nel prezzo dell'azione su di un relativo lungo periodo di tempo  $T$  è  $\sigma\sqrt{T}$ . Questo significa che la volatilità può essere interpretata come la deviazione standard della fluttuazione nel prezzo dell'azione in un anno.

### 3.3 Il lemma di Itô

Il prezzo di un'opzione su un'azione è una funzione del prezzo dell'azione considerata e del tempo. Più in generale, possiamo dire che il prezzo di qualsiasi derivato sia una funzione della variabile stocastica alla base del derivato e del tempo. Un importante risultato nell'area del comportamento di funzioni di variabili stocastiche è stato scoperto nel 1951 da Kiyosi Itô e, proprio dal suo nome, è conosciuto come lemma di Itô.

Partendo dal presupposto che un valore di una variabile  $x$  segua un processo di Itô

$$dx = a(x, t)dt + b(x, t)dz \quad (11)$$

dove  $dz$  è un processo di Wiener e  $a$  e  $b$  sono funzioni di  $x$  e  $t$ . La variabile  $x$  ha un tasso di slittamento di  $a$  e un tasso di varianza di  $b^2$ . Il lemma di Itô mostra che una funzione  $G$  di  $x$  e  $t$  segue il processo

$$dG = \left( \frac{\partial G}{\partial x} a + \frac{\partial G}{\partial t} + \frac{1}{2} \frac{\partial^2 G}{\partial x^2} b^2 \right) dt + \frac{\partial G}{\partial x} b dz \quad (12)$$

dove il  $dz$  è lo stesso processo di Wiener nell'equazione (11). Di conseguenza,  $G$  segue anche un processo di Itô con un tasso di slittamento di

$$\frac{\partial G}{\partial x} a + \frac{\partial G}{\partial t} + \frac{1}{2} \frac{\partial^2 G}{\partial x^2} b^2$$

ed un tasso di varianza di

$$\left( \frac{\partial G}{\partial x} \right)^2 b^2$$

Precedentemente è stato affermato che

$$dS = \mu S dt + \sigma S dz \quad (13)$$

con  $\mu$  e  $\sigma$  costanti, è un ragionevole modello per i movimenti del prezzo dell'azione. Dal lemma di Itô si deduce che questo processo, seguito da una funzione  $G$  in  $S$  e in  $t$ , sia

$$dG = \left( \frac{\partial G}{\partial S} \mu S + \frac{\partial G}{\partial t} + \frac{1}{2} \frac{\partial^2 G}{\partial S^2} \sigma^2 S^2 \right) dt + \frac{\partial G}{\partial S} \sigma S dz \quad (14)$$

È da notare che entrambi  $S$  e  $G$  sono influenzati dalla stessa sorgente di incertezza,  $dz$ .

### 3.3.1 Applicazione ai contratti Forward

Il forward è il contratto derivato con il quale due parti si impegnano a negoziare a scadenza un'attività o una merce ad un prezzo stabilito (delivery price)<sup>10</sup>.

Per illustrare il lemma di Itô viene preso, come esempio, un contratto forward su di un'azione che non paga dividendi. Assunto che il tasso di interesse libero dai rischi sia costante ed uguale a  $r$  per tutto il tempo vale l'equazione

$$F_0 = S_0 e^{rT}$$

dove  $F_0$  è il valore del forward al tempo zero,  $S_0$  è il prezzo di mercato al tempo zero e  $T$  è il tempo di maturità del forward.

L'interesse è riposto su cosa accade al prezzo del forward quando il tempo passa; vengono definiti  $F$  e  $S$  come, rispettivamente, il prezzo del forward e il prezzo dell'azione ad un tempo  $t$ , con  $t < T$ . La relazione tra  $F$  e  $S$  è

$$F = S e^{r(T-t)} \tag{15}$$

Assunto che il processo per  $S$  è definito dall'equazione (13), viene usato il lemma di Itô per determinare il processo  $F$ .

---

<sup>10</sup>Il forward più diffuso è il Forward Rate Agreement (FRA), in cui le parti si accordano sul tasso forward di un deposito o di un prestito a termine.

Colui che acquista il FRA (buyer) acquista un deposito e paga un tasso fisso, chi vende il FRA (seller) concede un prestito ed incassa il tasso fisso. Il tasso fisso (tasso contrattuale o delivery rate) è quello contrattualmente stabilito dalle parti contraenti.

Nel meccanismo del FRA, a scadenza, le parti liquidano la differenza tra il tasso fisso (contrattuale) ed il tasso corrente di riferimento (tasso di liquidazione o settlement rate) ma non si scambiano il sottostante seppur la base del FRA sia un prestito o un deposito. Precisamente, a scadenza, se il tasso fisso è maggiore di quello corrente di mercato, il FRA buyer versa la differenza al seller; viceversa, se il tasso fisso risulterà essere inferiore a quello di liquidazione, sarà il venditore a versare la differenza al compratore del contratto.

I FRA consentono quindi di coprire il rischio di rialzo o di ribasso dei tassi forward e possono essere acquistati o venduti per speculare sulla tendenza (al rialzo o al ribasso) dei tassi a breve termine.

Dall'equazione (15):

$$\frac{\partial F}{\partial S} = e^{r(T-t)}, \quad \frac{\partial^2 F}{\partial S^2} = 0, \quad \frac{\partial F}{\partial t} = -rSe^{r(T-t)}$$

Dall'equazione (14), il processo per  $F$  è dato da

$$dF = [e^{r(T-t)}\mu S - rSe^{r(T-t)}]dt + e^{r(T-t)}\sigma S dz$$

dove, applicando la sostituzione  $F = Se^{r(T-t)}$ , si ottiene

$$dF = (\mu - r)F dt + \sigma F dz \quad (16)$$

### 3.4 La proprietà lognormale

Viene utilizzato il lemma di Itô per derivare il processo seguito da  $\ln S$  quando  $S$  segue il processo nell'equazione (13).

Definita  $G = \ln S$ , dato che

$$\frac{\partial G}{\partial S} = \frac{1}{S}, \quad \frac{\partial^2 G}{\partial S^2} = -\frac{1}{S^2}, \quad \frac{\partial G}{\partial t} = 0$$

segue, dall'equazione (14), che il processo seguito da  $G$  è

$$dG = \left(\mu - \frac{\sigma^2}{2}\right)dt + \sigma dz \quad (17)$$

Dato che  $\mu$  e  $\sigma$  sono costanti questa equazione indica che  $G = \ln S$  segue un processo di Wiener generalizzato. Questa ha un tasso di slittamento costante pari a  $\mu - \frac{\sigma^2}{2}$  e un tasso di varianza costante pari a  $\sigma^2$ . La fluttuazione in  $\ln S$  al tempo zero ed in un istante nel futuro  $T$  è distribuita normalmente con media  $\left(\mu - \frac{\sigma^2}{2}\right)T$  e varianza  $\sigma^2 T$ .

Questo significa che

$$\ln S_T - \ln S_0 \sim \Phi\left[\left(\mu - \frac{\sigma^2}{2}\right)T, \sigma\sqrt{T}\right] \quad (18)$$

oppure

$$\ln S_T \sim \Phi\left[\ln S_0 + \left(\mu - \frac{\sigma^2}{2}\right)T, \sigma\sqrt{T}\right] \quad (19)$$

dove  $S_T$  è il prezzo dell'azione al tempo (futuro)  $T$ ,  $S_0$  è il prezzo dell'azione al tempo zero e  $\Phi(m, s)$  indica una distribuzione normale con media  $m$  e deviazione standard  $s$ .

L'equazione (19) mostra che  $\ln S_T$  è normalmente distribuito. Una variabile ha una *distribuzione lognormale* se il logaritmo naturale della variabile è normalmente distribuito. Dato il prezzo di oggi, il modello per il comportamento

del prezzo dell'azione al tempo  $T$ , che è stato descritto fino ad ora, è distribuito lognormalmente. La deviazione standard del logaritmo del prezzo dell'azione è  $\sigma\sqrt{T}$ . È proporzionale alla radice quadrata di quanto si guardi in avanti.

## 4 Il modello di Black-Scholes

Nei primi anni '70 del secolo scorso, Fischer Black, Myron Scholes e Robert Merton hanno compiuto una delle più grandi rivoluzioni nel prezzaggio delle azioni, attraverso il lavoro che viene riconosciuto sotto il nome di modello di Black-Scholes. Questo ha avuto una grande influenza nel modo con il quale viene fatto, da parte degli operatori, il commercio e la protezione delle opzioni.

### 4.1 La proprietà lognormale dei prezzi delle azioni

Il modello per il comportamento dei prezzi delle azioni usato da Black, Scholes e Merton è il modello presentato nelle sezioni precedenti. Viene assunto che le fluttuazioni della percentuale nel prezzo delle azioni in un periodo di tempo breve sono normalmente distribuite. Definito:

$\mu$ : ritorno previsto sull'azione;

$\sigma$ : volatilità del prezzo dell'azione.

La media della fluttuazione di percentuale nel tempo  $\delta t$  è  $\mu \delta t$  e la deviazione standard della stessa fluttuazione di percentuale è  $\sigma\sqrt{\delta t}$ , così che:

$$\frac{\delta S}{S} \sim \Phi(\mu \delta t, \sigma\sqrt{\delta t}) \quad (20)$$

dove  $\delta S$  è la fluttuazione nel prezzo dell'azione  $S$  al tempo  $\delta t$  e  $\Phi(m, s)$  indica una distribuzione standard con media  $m$  e deviazione standard  $s$ .

Come visto nella Sezione 3.4, il modello implica che

$$\ln S_T - \ln S_0 \sim \Phi\left[\left(\mu - \frac{\sigma^2}{2}\right)T, \sigma\sqrt{T}\right]$$

dalla quale, utilizzando le proprietà dei logaritmi:

$$\ln \frac{S_T}{S_0} \sim \Phi\left[\left(\mu - \frac{\sigma^2}{2}\right)T, \sigma\sqrt{T}\right] \quad (21)$$

e

$$\ln S_T \sim \Phi\left[\ln S_0 + \left(\mu - \frac{\sigma^2}{2}\right)T, \sigma\sqrt{T}\right] \quad (22)$$

dove  $S_T$  è il prezzo dell'azione, nel futuro, al tempo  $T$  e  $S_0$  è il prezzo dell'azione al tempo zero. L'equazione (22) mostra che  $\ln S_T$  è normalmente distribuito, quindi  $S_T$  è lognormalmente distribuito.

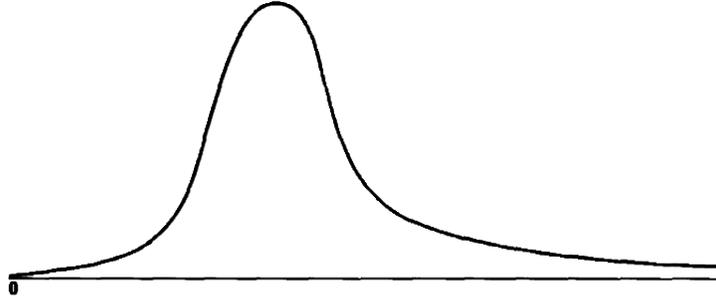


Figura 9: Distribuzione lognormale.

Come esempio: avendo un'azione con prezzo iniziale di 40, un previsto ritorno del 16% annuo ed una volatilità del 20% annuo, dall'equazione (22) si ricava che la distribuzione di probabilità del prezzo  $S_T$  dell'azione in sei mesi è data da:

$$\ln S_T \sim \Phi\left[\ln 40 + \left(0.16 - \frac{0.2^2}{2}\right) \times 0.5, 0.2\sqrt{0.5}\right]$$

$$\ln S_T \sim \Phi(3.759, 0.141)$$

C'è un 95% di probabilità che una variabile normalmente distribuita abbia un valore al di sotto di 1.96 deviazioni standard della sua media. Di conseguenza, con un intervallo di confidenza del 95%,

$$3.759 - 1.96 \times 0.141 < \ln S_T < 3.759 + 1.96 \times 0.141$$

Questo può essere scritto come

$$e^{3.759 - 1.96 \times 0.141} < S_T < e^{3.759 + 1.96 \times 0.141}$$

oppure

$$32.55 < S_T < 56.56$$

Di conseguenza c'è un 95% di probabilità che il prezzo dell'azione in sei mesi sarà situato tra 32.55 e 56.56.

Una variabile che ha una distribuzione lognormale può assumere qualunque valore tra zero ed infinito. La Figura 9 mostra la forma di una distribuzione lognormale.

Diversamente dalla distribuzione normale, è distorta in modo che la media, mediana e moda siano tutte differenti. Il valore atteso  $E(S_T)$  è dato da

$$E(S_T) = S_0 e^{\mu T} \tag{23}$$

Questo calza con la definizione di  $\mu$  come il tasso previsto di ritorno. La varianza,  $var(S_T)$ , è data da

$$var(S_T) = S_0^2 e^{2\mu T} (e^{\sigma^2 T} - 1) \quad (24)$$

Considerando un'azione che ha un prezzo corrente pari a 20, un ritorno previsto del 20% per anno e una volatilità del 40% annua, il prezzo previsto di questa azione in un anno  $E(S_T)$  e la varianza del suo prezzo  $var(S_T)$  in un anno risultano essere

$$E(S_T) = 20e^{0.2 \times 1} = 24.43$$

$$var(S_T) = 400e^{2 \times 0.2 \times 1} (e^{0.4^2 \times 1} - 1) = 103.54$$

La deviazione standard del prezzo dell'azione in un anno è  $\sqrt{103.54}$  pari a 10.18.

## 4.2 La distribuzione del tasso di ritorno

La proprietà lognormale dei prezzi dell'azione può essere usata per fornire informazioni sulla distribuzione di probabilità del tasso di ritorno composto continuativamente su di un'azione tra zero e  $T$ . Definito con  $\eta$  tasso di ritorno annuo composto continuativamente tra zero e  $T$ , segue che

$$S_T = S_0 e^{\eta T}$$

tale che

$$\eta = \frac{1}{T} \ln \frac{S_T}{S_0} \quad (25)$$

Dall'equazione (21) segue che

$$\eta \sim \Phi\left(\mu - \frac{\sigma^2}{2}, \frac{\sigma}{\sqrt{T}}\right) \quad (26)$$

Di conseguenza, tasso di ritorno annuo composto continuativamente è distribuito normalmente con media  $\mu - \frac{\sigma^2}{2}$  e deviazione standard  $\frac{\sigma}{\sqrt{T}}$ .

Considerando un'azione con un tasso previsto del 17% annuo e una volatilità del 20% annuo, la distribuzione di probabilità per il tasso attuale di ritorno raggiunto su tre anni è normale con media

$$0.17 - \frac{0.2^2}{2} = 0.15$$

oppure il 15% annuo e deviazione standard

$$\frac{0.2}{\sqrt{3}} = 0.1155$$

oppure l'11.55% annuo. Dato che c'è il 95% di possibilità che una variabile normalmente distribuita giaccia entro 1.96 deviazioni standard della sua media, c'è una confidenza del 95% che il ritorno attuale raggiunto su tre anni abbia un valore tra  $-7.6\%$  e  $+37.6\%$  per anno.

### 4.3 Ritorno previsto

Il ritorno previsto  $\mu$ , richiesto dagli investitori da un'azione, dipende dalla rischiosità dell'azione. Più grande è il rischio, più alto è il ritorno previsto. C'è una dipendenza anche rispetto al livello di tasso di interesse: più alto è il tasso più alto sarà il ritorno previsto dall'azione. È chiaro, quindi, che il valore di un'azione, quando è espresso in termini di valore dell'azione in questione, non dipende totalmente da  $\mu$ .

L'equazione (20) mostra che  $\mu\delta t$  è la percentuale prevista di fluttuazione nel prezzo dell'azione in un breve intervallo di tempo  $\delta t$ . Questo significa che  $\mu$  è il ritorno previsto in un breve intervallo di tempo  $\delta t$ .  $\mu$  è anche il previsto tasso di ritorno composto continuativamente su di un'azione in un relativamente lungo periodo di tempo.

Il tasso composto continuativamente accumulato su  $T$  anni è

$$\frac{1}{T} \ln \frac{S_T}{S_0}$$

e l'equazione (26) mostra che il valore previsto da questo è  $\mu - \frac{\sigma^2}{2}$ .

Il motivo di questa distinzione tra  $\mu$  nell'equazione (20) e il  $\mu - \frac{\sigma^2}{2}$  nell'equazione (26) è sottile ma importante. Partendo con l'equazione (23):

$$E(S_T) = S_0 e^{\mu T}$$

Usando i logaritmi abbiamo

$$\ln[E(S_T)] = \ln(S_0) + \mu T$$

A questo punto è possibile affermare

$$\ln[E(S_T)] = E[\ln(S_T)]$$

tale che  $E[\ln(S_T)] - \ln(S_0) = \mu T$  oppure  $E[\ln(\frac{S_T}{S_0})] = \mu T$ . Comunque, non è possibile fare questo poiché  $\ln$  è una funzione non lineare. Infatti  $\ln[E(S_T)] > E[\ln(S_T)]$ , così che  $E[\ln(\frac{S_T}{S_0})] < \mu T$ . Questo è compatibile con l'equazione (26).

Considerando un numero grande di piccoli periodi di lunghezza  $\delta t$ , viene definito  $S_i$  come il prezzo dell'azione alla fine dell' $i$ -esimo intervallo e  $\delta S_i$  come  $S_{i+1} - S_i$ . Sotto le assunzioni fatte sul comportamento del prezzo dell'azione, la media del ritorno su di un'azione in ogni intervallo è vicina a  $\mu$ . In altre parole,  $\mu$  è vicino alla media aritmetica del rapporto  $\frac{\delta S_i}{S_i}$ . In ogni caso, il ritorno previsto su tutto il periodo coperto dal dato, espresso con un periodo composto di  $\delta t$ , si avvicina a  $\mu - \frac{\sigma^2}{2}$  e non a  $\mu$ .<sup>11</sup>

Supposto che la seguente successione sia la sequenza di ritorni per anno di un'azione misurata usando una composizione annuale:

$$15\%, 20\%, 30\%, -20\%, 25\%$$

la media aritmetica dei ritorni, calcolata come la somma dei ritorni diviso 5, è 14%. Comunque, un investitore potrebbe guadagnare meno del 14% annuo lasciando i soldi investiti nell'azione per cinque anni. Il valore di 100 alla fine di cinque anni potrebbe essere

$$100 \times 1.15 \times 1.20 \times 1.30 \times 0.80 \times 1.25 = 179.40$$

in contrasto con il risultato nel caso del 14%, che sarebbe:

$$100 \times 1.14^5 = 192.54$$

l'effettivo ritorno medio guadagnato dall'investitore sarebbe, con composizione annuale:

$$(1.7940)^{\frac{1}{5}} - 1 = 0.124$$

oppure 12.4% annuo.

Quindi è chiaro che il termine *ritorno previsto* è ambiguo, perché si può riferire sia a  $\mu$  che a  $\mu - \frac{\sigma^2}{2}$  (in generale, se non specificato, il termine "ritorno previsto" è riferito a  $\mu$ ).

---

<sup>11</sup>Se venisse definito il *ritorno obeso* come uno più il ritorno regolare, il ritorno obeso su tutto il periodo coperto dal dato è la media geometrica del ritorno obeso in ogni piccolo intervallo di lunghezza  $\delta t$ , non la media aritmetica. La media geometrica è più piccola della media aritmetica.

## 4.4 Volatilità

La volatilità  $\sigma$  di un'azione è la misura dell'incertezza riguardo al ritorno previsto dall'azione. Le azioni, tipicamente, hanno una volatilità che varia tra il 20% ed il 50%.

Dall'equazione (26) la volatilità di un prezzo di un'azione può essere definita come la deviazione standard del ritorno fornito dall'azione in un anno, quando il ritorno è espresso usando una composizione continua.

Quando  $T$  è piccolo, l'equazione (20) mostra che  $\sigma\sqrt{T}$  è approssimativamente uguale alla deviazione standard della percentuale di fluttuazione nel prezzo dell'azione al tempo  $T$ . Supposto  $\sigma = 0.3$  oppure 30% annuo e il prezzo corrente di 50, la deviazione standard della percentuale di fluttuazione nel prezzo di un'azione in una settimana è approssimativamente

$$30 \times \sqrt{\frac{1}{52}} = 4.16\%$$

Un movimento di una deviazione standard nel prezzo dell'azione in una settimana è di conseguenza  $50 \times 0.0416$  oppure 2.08.

L'equazione (20) mostra che l'incertezza sul futuro prezzo dell'azione, misurata dalla sua deviazione standard, aumenta approssimativamente della radice quadrata di quanto lontano si sta guardando. Ad esempio, la deviazione standard del prezzo dell'azione in quattro settimane è approssimativamente il doppio della deviazione standard in una settimana.

### 4.4.1 Stimare la volatilità dai dati storici

Per stimare la volatilità del prezzo di un'azione empiricamente, il prezzo dell'azione normalmente è osservato a intervalli di tempo fissati (per esempio ogni giorno, settimana o mese).

Dati:

$n + 1$  : numero di osservazioni;

$S_i$  : prezzo dell'azione alla fine dell' $i$ -esimo intervallo ( $i = 0, 1, \dots, n$ );

$\tau$  : lunghezza degli intervalli di tempo in anni.

posto

$$u_i = \ln \left( \frac{S_i}{S_{i-1}} \right)$$

per  $i = 1, 2, \dots, n$ .

La stima  $s$  della deviazione standard dell' $u_i$ -esimo è data da

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (u_i - \bar{u})^2}$$

oppure

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n u_i^2 - \frac{1}{n(n-1)} \left( \sum_{i=1}^n u_i \right)^2}$$

dove  $\bar{u}$  è la media dell' $u_i$ -esimo.

Dall'equazione (21) la deviazione standard dell' $u_i$ -esimo è  $\sigma\sqrt{\tau}$ . La variabile  $s$  è un'approssimazione di  $\sigma\sqrt{\tau}$  e ne segue che  $\sigma$  possa essere stimata come

$$\hat{\sigma} = \frac{s}{\sqrt{\tau}}$$

L'errore di questa stima può essere approssimativamente  $\frac{\hat{\sigma}}{\sqrt{2n}}$ .

Scegliere un appropriato valore di  $n$  non è così facile; in generale più dati portano ad una maggiore accuratezza, ma  $\sigma$  cambia su tempi e dati che sono troppo vecchi e potrebbero non essere rilevanti per predire il futuro. Un compromesso che sembra funzionare in modo ragionevole è usare i prezzi di chiusura giornalieri dai più recenti 90 ai 180 giorni. Un metodo spesso usato è impostare  $n$  uguale al numero di giorni dove si vuole applicare la volatilità (ad esempio, se la stima della volatilità fosse fatta ad un valore di un'opzione di due anni, allora verrebbero usati i dati degli ultimi due anni).

La tabella 2 mostra una possibile sequenza di prezzi di azioni durante 21 giorni di mercato; in questo caso

$$\sum u_i = 0.09531 \text{ e } \sum u_i^2 = 0.00326$$

e l'approssimazione della deviazione standard sul ritorno giornaliero è pari a

$$\sqrt{\frac{0.00326}{19} - \frac{0.09531^2}{380}} = 0.01216$$

oppure 1.216%. Assumendo che in un anno ci sono 252 giorni di mercato  $\tau = \frac{1}{252}$  e il dato permette di avere una stima per la volatilità annua di  $0.01216\sqrt{252} = 0.193$  oppure 19.3%. L'errore di questa stima è di

$$\frac{0.193}{\sqrt{2 \times 20}} = 0.031$$

Giorno	Prezzo di chiusura dell'azione	Prezzo relativo $\frac{S_i}{S_{i-1}}$	ritorno giornaliero $u_i = \ln(\frac{S_i}{S_{i-1}})$
0	20.000		
1	20.10	1.00500	0.00499
2	19.90	0.99005	-0.01000
3	20.00	1.00503	0.00501
4	20.50	1.02500	0.02469
5	21.25	0.98780	-0.01227
6	20.90	1.03210	0.03159
7	20.90	1.00000	0.00000
8	20.90	1.00000	0.00000
9	20.75	0.99282	-0.00720
10	20.75	1.00000	0.00000
11	21.00	1.01205	0.01198
12	21.10	1.00476	0.00475
13	20.90	0.99052	-0.00952
14	20.90	1.00000	0.00000
15	21.25	1.01675	0.01661
16	21.40	1.00706	0.00703
17	21.40	1.00000	0.00000
18	21.25	0.99299	-0.00703
19	21.75	1.02353	0.02326
20	22.00	1.01149	0.01143

Tabella 2: Calcolo della volatilità.

oppure il 3.1%.

Questa analisi assume che l'azione non paghi dividendi, ma può essere adattata anche alle azioni che pagano dividendi. Il ritorno  $u_i$  durante l'intervallo di tempo che comprende un giorno di dividendo passato è dato da

$$u_i = \ln \frac{S_i + D}{S_{i-1}}$$

dove  $D$  è il totale del dividendo. Il ritorno negli altri periodi di tempo rimane

$$u_i = \ln \frac{S_i}{S_{i-1}}$$

## 4.5 Concetti fondamentali sull'equazione differenziale di Black-Scholes-Merton

L'equazione differenziale di Black-Scholes-Merton è un'equazione che deve essere soddisfatta dal prezzo di qualsiasi derivato dipendente da un'azione che non paghi dividendi.

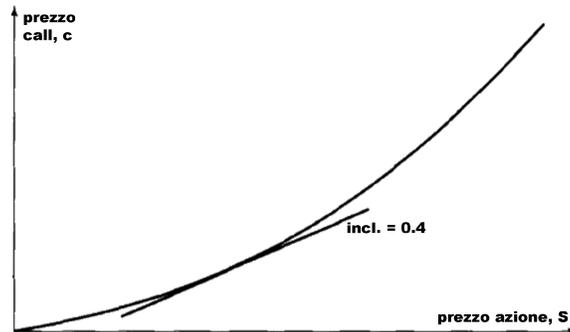


Figura 10: Relazione tra  $c$  e  $S$ .

Gli argomenti di questa equazione richiedono di preparare un portfolio senza rischi<sup>12</sup> fatto da una posizione nei derivati e una posizione delle azioni. In assenza di opportunità di scambio, il ritorno dal portfolio deve essere  $r$ , il tasso di interesse senza rischi.

La ragione di un portfolio senza rischi può essere stabilita dal fatto che il prezzo dell'azione ed il prezzo del derivato sono entrambi affetti dalla stessa fonte di incertezza: il movimento del prezzo delle azioni. In un breve periodo, il prezzo del derivato è perfettamente correlato con il prezzo dell'azione in considerazione. Quando viene creato un appropriato portfolio di azioni e derivati, il guadagno o la perdita dalla posizione dell'azione bilancia sempre il guadagno o la perdita dalla posizione del derivato tale che il valore totale del portfolio, alla fine del breve periodo, è conosciuto con certezza.

Per esempio, supposto che, a un particolare momento, la relazione tra una piccola fluttuazione nel prezzo dell'azione  $\delta S$  e la conseguente piccola fluttuazione nel prezzo di un'opzione call Europea  $\delta c$  è data da

$$\delta c = 0.4\delta S$$

Questo significa che l'inclinazione della retta rappresentante la relazione tra  $c$  e  $S$  è 0.4 come mostrato in Figura 10.

<sup>12</sup>è un tasso di interesse che può essere ottenuto investendo in strumenti finanziari senza il rischio di perdita. In ogni caso lo strumento finanziario può portare altri tipi di rischi come ad esempio il rischio di mercato (il rischio di fluttuazioni nei tassi interesse), rischio di liquidità (il rischio di non avere la possibilità di vendere lo strumento per soldi).

**Assunzioni:** per derivare l'equazione di Black-Scholes-Merton sono le seguenti:

- Il prezzo dell'azione segue il processo descritto nelle sezioni precedenti con i parametri  $\sigma$  e  $\mu$  costanti;
- È consentito lo short selling<sup>13</sup>;
- Non ci sono tasse o costi di transazione;
- Non ci sono dividendi durante la vita del derivato;
- Non ci sono opportunità di scambi senza rischio;
- Il trading è continuo;
- Il tasso di interesse privo di rischio  $r$  è costante ed è lo stesso.

Come visto in precedenza alcune di queste assunzioni possono essere rilassate; ad esempio  $\mu$  e  $\sigma$  possono essere in funzione (conosciuta) del tempo.

## 4.6 Derivazione dell'equazione differenziale di Black-Scholes-Merton

Viene considerato il processo del prezzo delle azioni discusso nella Sezione 3

$$dS = \mu S dt + \sigma S dz \quad (27)$$

Supposto che  $f$  sia il prezzo dell'opzione call o di un altro derivato determinato da  $S$ , la variabile  $f$  deve essere una qualche funzione di  $S$  e  $t$ . Per questa ragione, dall'equazione (14),

$$df = \left( \frac{\partial f}{\partial S} \mu S + \frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial S^2} \sigma^2 S^2 \right) dt + \frac{\partial f}{\partial S} \sigma S dz \quad (28)$$

Le versioni discrete delle equazioni (27) e (28) sono

$$\delta S = \mu S \delta t + \sigma S \delta z \quad (29)$$

---

<sup>13</sup>È la pratica di vendere risorse economiche che sono state prese da una terza parte con l'intenzione di comprare la stessa risorsa successivamente alla data di ritorno a colui che le ha prestate. Fare short selling è l'inverso di fare long selling dove l'investitore ha i suoi profitti dall'aumento del prezzo della risorsa economica.

e

$$\delta f = \left( \frac{\partial f}{\partial S} \mu S + \frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial S^2} \sigma^2 S^2 \right) \delta t + \frac{\partial f}{\partial S} \sigma S \delta z \quad (30)$$

dove  $\delta S$  e  $\delta f$  sono le fluttuazioni in  $S$  e  $f$  in un piccolo intervallo di tempo  $\delta t$ . Il  $\delta t (= \epsilon \sqrt{\delta t})$  nelle equazioni (29) e (30) è lo stesso. Segue che, scegliendo un portfolio di azioni e derivati, il processo di Wiener può essere eliminato (i processi di Wiener sottostanti  $f$  e  $S$  sono gli stesso, come discusso in Sezione 3.3). Il portfolio appropriato è:

-1: derivati;  
 $+\frac{\partial f}{\partial S}$ : azioni.

Il possessore del portfolio è in posizione short su di un derivato e in posizione long sulla quantità  $\frac{\partial f}{\partial S}$  di azioni. Definito  $\Pi$  come il valore del portfolio, dalla definizione

$$\Pi = -f + \frac{\partial f}{\partial S} S \quad (31)$$

la fluttuazione  $\delta \Pi$  nel valore del portfolio nel periodo di tempo  $\delta t$  è data da

$$\delta \Pi = -\delta f + \frac{\partial f}{\partial S} \delta S \quad (32)$$

Sostituendo le equazioni (29) e (30) nell'equazione (32) produce:

$$\delta \Pi = \left( -\frac{\partial f}{\partial t} - \frac{1}{2} \frac{\partial^2 f}{\partial S^2} \sigma^2 S^2 \right) \delta t \quad (33)$$

Dato che questa equazione non coinvolge  $\delta z$ , il portfolio deve essere senza rischi durante il periodo di tempo  $\delta t$ . Le assunzioni riportate in fondo alla Sezione precedente implicano che il portfolio debba ottenere istantaneamente lo stesso tasso di ritorno come tutti gli altri derivati liberi da rischio a corto termine. Se è ottenuto di più del ritorno, coloro che commerciano potrebbero produrre profitti senza rischio prendendo denaro in prestito per comprare il portfolio; se guadagnassero meno, potrebbero produrre un profitto senza rischio creando un portfolio short e comprando derivati privi di rischio. Segue che

$$\delta \Pi = r \Pi \delta t$$

dove  $r$  è il tasso di interesse libero da rischi. Sostituendo dall'equazione (31) e (33) si ottiene

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

tale che

$$\frac{\partial f}{\partial t} + rS \frac{\partial f}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} = rf \quad (34)$$

L'equazione (34) è l'equazione differenziale di Black-Scholes-Merton. Ha molteplici soluzioni, corrispondenti a tutti i differenti derivati che possono essere definiti dalla variabile  $S$ . Il particolare derivato che è ottenuto quando l'equazione è risolta dipende dalle condizioni al contorno che sono state usate. Queste specificano i valori del derivato nei limiti dei possibili valori di  $S$  e  $t$ . Nel caso delle opzioni europee di tipo call, la condizione al contorno chiave è

$$f = \max(S - K, 0) \text{ quando } t = T$$

Nel caso delle opzioni europee di tipo put invece

$$f = \max(K - S, 0) \text{ quando } t = T$$

È importante sottolineare che il portfolio usato nell'equazione (34) non è senza permanentemente rischi, ma è senza rischio in un infinitesimale breve intervallo di tempo. Come  $S$  e  $t$  cambiano, anche  $\frac{\partial f}{\partial S}$  cambia. Per mantenere il portfolio senza rischi è di conseguenza necessario modificare le proporzioni tra derivati e azioni nel portfolio stesso.

Un contratto forward su delle azioni che non paga dividendi è un derivato dipendente dall'azione e soddisfa l'equazione (34). Il valore di un contratto forward  $f$  ad un tempo generico  $t$  è dato in termini del prezzo dell'azione  $S$  a questo tempo da

$$f = S - Ke^{-r(T-t)}$$

dove  $K$  è il prezzo di consegna. Questo significa che

$$\frac{\partial f}{\partial t} = -rKe^{-r(T-t)}, \quad \frac{\partial f}{\partial S} = 1, \quad \frac{\partial^2 f}{\partial S^2} = 0$$

questi vengono sostituiti nella parte sinistra dell'equazione (34) così da ottenere

$$-rKe^{-r(T-t)} + rS$$

Questo è uguale a  $rf$ , e quindi l'equazione (34) è soddisfatta.

#### 4.6.1 Il prezzo dei derivati commerciabili

Qualsiasi funzione  $f(S, t)$  che è soluzione dell'equazione differenziale (34) è il prezzo teorico di un derivato che potrebbe essere commerciato. Se un derivato con quel prezzo esistesse, non dovrebbe creare alcun tipo di opportunità

di commercio. Per contro, se non esistesse alcuna funzione  $f(S, t)$  che soddisfi l'equazione differenziale (34), essa non potrebbe essere il prezzo di un derivato senza creare opportunità di scambio per chi commercia.

Per illustrare il punto si può considerare la funzione  $e^S$ . Questo non soddisfa l'equazione differenziale (34). Di conseguenza non è una candidata per essere un prezzo di un derivato dipendente dal prezzo dell'azione. Se esistesse uno strumento finanziario che ha sempre avuto quel prezzo, potrebbe essere un'opportunità di scambio. Come secondo esempio si consideri la funzione

$$\frac{e^{(\sigma^2 - 2r)(T-t)}}{S}$$

Questa soddisfa l'equazione differenziale e, da come detto precedentemente, è il prezzo di qualche derivato commerciabile (è il prezzo di un derivato che paga  $\frac{1}{S_T}$  al tempo  $T$ ).

## 4.7 Valutazione risk-neutral

È interessante fare una valutazione risk-neutral<sup>14</sup>. Questa valutazione emerge da una proprietà chiave dell'equazione di Black-Scholes-Merton (34), la proprietà è che l'equazione non interessa nessuna variabile che è affetta dalle preferenze di rischio degli investitori. Le variabili che appaiono nell'equazione sono il prezzo corrente dell'azione, il tempo, la volatilità del prezzo dell'azione e il tasso di interesse senza rischi; tutte queste sono indipendenti dalle preferenze di rischio.

L'equazione di Black-Scholes-Merton non dovrebbe essere indipendente dalle preferenze di rischio se implica il ritorno previsto  $\mu$  dell'azione. Questo perché il valore di  $\mu$  dipende dalle preferenze di rischio: più alto è il livello di avversione al rischio dell'investitore, più alto  $\mu$  dovrà essere per qualsiasi azione. Se le preferenze di rischio non interessano l'equazione allora non possono interessare la sua soluzione. Qualunque insieme di preferenze di rischio può, di conseguenza, essere usato quando viene valutata  $f$ . In particolare, è possibile fare la semplice assunzione che tutti gli investitori sono insensibili al rischio.

In altre parole, dove gli investitori sono risk-neutral, il ritorno previsto su tutti i derivati è il tasso di interesse privo di rischi  $r$ . La ragione è che gli investi-

<sup>14</sup>Significa che è completamente insensibile al rischio incluso negli investimenti ed è solo interessato al ritorno previsto.

tori insensibili al rischio non vogliono un premio per motivarli a prendere rischi. È anche vero che il valore attuale di qualsiasi totale di soldi, trasferibili all'interno e all'esterno del commercio in uno scenario di insensibilità al rischio, può essere ottenuto dalla sottrazione del suo valore previsto al tasso senza rischio, rendendo l'analisi dei derivati più semplice.

Prendendo un derivato che garantisca un profitto ad un tempo fissato, può essere valutato usando la valutazione risk-neutral secondo la procedura seguente:

- Assumere che il prezzo di ritorno previsto per la risorsa sottostante sia il tasso di interesse privo di rischio  $r$ , cioè assumere  $\mu = r$ ;
- Calcolare il profitto previsto dall'opzione e dal suo tasso di maturità;
- Sottrarre il profitto previsto al tasso di interesse senza rischio.

È importante sottolineare che la valutazione risk-neutral (o l'assunzione che tutti gli investitori siano insensibili al rischio) è semplicemente un metodo che non rispetta la realtà per ottenere soluzioni dell'equazione di Black-Scholes. Le soluzioni ottenute sono sempre valide per tutti gli scenari, non solo in quelli dove gli investitori sono insensibili al rischio. Quando lo scenario passa da risk-neutral a ostile al rischio succedono due cose: cambiano il tasso di crescita previsto nel prezzo dell'azione e il tasso di sottrazione che è usato per i profitti dai derivati.

#### 4.7.1 Applicazione ai contratti forward su di un'azione

Nell'esempio sui contratti forward riportato in fondo alla Sezione 4.6 è stato verificato che la formula di prezzo soddisfa l'equazione di Black-Scholes; adesso viene derivata la formula facendo una valutazione risk-neutral, assumendo che i tassi d'interesse siano costanti e uguali a  $r$ .

Viene preso un contratto forward che matura al tempo  $T$  con un prezzo di consegna  $K$ . Il valore del contratto al momento della sua maturità è

$$S_T - K$$

dove  $S_T$  è il prezzo dell'azione al tempo  $T$ . Secondo la valutazione risk-neutral, il valore del contratto al tempo zero è il suo valore previsto al tempo  $T$  in uno scenario insensibile al rischio, ripulito dal tasso di interesse senza rischio.

Indicando con  $f$  il valore del contratto forward al tempo zero

$$f = e^{-rT} \hat{E}(S_T - K) \quad (35)$$

dove  $\hat{E}$  indica il valore previsto in uno scenario insensibile al rischio. Dato che  $K$  è una costante, l'equazione (35) diventa

$$f = e^{-rT} \hat{E}(S_T) - Ke^{-rT} \quad (36)$$

Il tasso  $\mu$  di crescita previsto per il prezzo dell'azione diventa  $r$  in una valutazione risk-neutral. In più dall'equazione (23)

$$\hat{E}(S_T) = S_0 e^{rT} \quad (37)$$

Sostituendo l'equazione (37) nell'equazione (36) abbiamo

$$f = S_0 - Ke^{-rT} \quad (38)$$

## 4.8 Le formule di prezzaggio di Black-Scholes

Le formule di Black-Scholes per i prezzi al tempo zero delle opzioni call Europee su di un'azione che non paga i dividendi e per le put Europee, sempre su di un'azione che non paga dividendi, sono:

$$c = S_0 N(d_1) - Ke^{-rT} N(d_2) \quad (39)$$

e

$$p = Ke^{-rT} N(-d_2) - S_0 N(-d_1) \quad (40)$$

dove

$$d_1 = \frac{\ln \frac{S_0}{K} + (r + \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}$$

$$d_2 = \frac{\ln(\frac{S_0}{K}) + (r - \frac{\sigma^2}{2})T}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T}$$

La funzione  $N(x)$  è la funzione cumulativa di distribuzione di probabilità per una distribuzione normale standardizzata. In altre parole, è la probabilità che una variabile con una distribuzione normale standard,  $\Phi(0, 1)$ , sia meno di  $x$ , come mostrato in Figura 11.

Le altre variabili rappresentano rispettivamente:

- $c$ , il prezzo dell'opzione Europea call;
- $p$ , il prezzo dell'opzione Europea put;

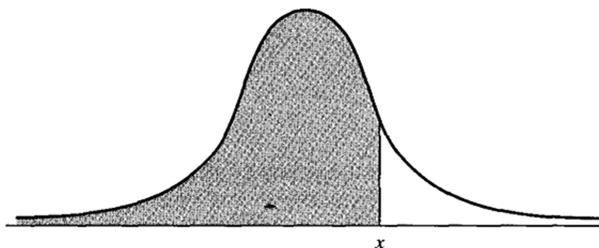


Figura 11: L'area scura rappresenta la funzione  $N(x)$ .

- $S_0$ , il prezzo dell'azione al tempo zero;
- $K$ , il prezzo al quale l'opzione può essere esercitata (strike price);
- $r$ , il tasso composto di interesse senza rischio;
- $\sigma$ , la volatilità del prezzo dell'azione;
- $T$  è il tempo di esercizio dell'opzione.

In modo di derivare le formule di Black-Scholes è quello di risolvere l'equazione differenziale (34) sotto le condizioni al contorno presentate nella Sezione 4.6<sup>15</sup>.

Un altro approccio è quello di usare una valutazione risk-neutral. Considerando un'opzione Europea call, il valore previsto dell'opzione, al momento della sua esercitazione, è

$$\hat{E}[\max(S_T - K, 0)]$$

dove, come prima, con  $\hat{E}$  si esprime il valore previsto in uno scenario insensibile al rischio. Nella valutazione risk-neutral l'opzione Europea call  $c$  ha un valore pari a

$$c = e^{-rT} \hat{E}[\max(S_T - K, 0)] \quad (41)$$

<sup>15</sup>L'equazione differenziale fornisce i prezzi della call e della put al generico tempo  $t$ . Per esempio, il prezzo della call che soddisfa l'equazione differenziale è  $c = SN(d_1) - Ke^{-r(T-t)}N(d_2)$ , dove

$$d_1 = \frac{\ln \frac{S}{K} + (r + \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}$$

e  $d_2 = d_1 - \sigma\sqrt{T-t}$ .

Per fornire un'interpretazione dei termini nell'equazione (39) è possibile notare che può essere riscritta

$$c = e^{-rT} [S_0 N(d_1) e^{rT} - K N(d_2)] \quad (42)$$

L'espressione  $N(d_2)$  è la probabilità che l'opzione possa essere esercitata in uno scenario insensibile al rischio,  $KN(d_2)$  è il prezzo al quale può essere esercitata per la probabilità che lo strike price venga pagato. L'espressione  $S_0 N(d_1) e^{rT}$  è il valore previsto di una variabile che equivalga a  $S_T$ , se  $S_T > K$  o zero, altrimenti in uno scenario insensibile al rischio.

#### 4.8.1 Proprietà delle formule di Black-Scholes

Le formule di Black-Scholes hanno le giuste proprietà generali considerando quello che succede quando uno dei parametri assume valori estremi.

Quando il prezzo dell'azione  $S_0$  diventa molto grande un'opzione call quasi sicuramente sarà esercitata, diventando molto simile ad un contratto forward con tempo di consegna  $K$ . Il prezzo previsto per la call sarà

$$S_0 - Ke^{-rT}$$

Questo è infatti il prezzo della call dato dall'equazione (39) perché, quando  $S_0$  diventa un valore molto grande, entrambi  $d_1$  e  $d_2$  diventano grandi ed i valori di  $N(d_1)$  e  $N(d_2)$  si avvicinano a 1. Quando il prezzo dell'azione diventa molto grande, il prezzo dell'opzione Europea put  $p$  si avvicinerà a zero. Questo è consistente con l'equazione (40) poiché  $N(-d_1)$  e  $N(-d_2)$  sono entrambi vicini a zero.

Quando la volatilità  $\sigma$  si avvicina a zero, siccome l'azione è virtualmente priva di rischio, il suo prezzo crescerà a tasso  $r$  verso  $S_0 e^{rT}$  al tempo  $T$  e il ritorno dall'opzione call è

$$\max(S_0 e^{rT}, 0)$$

sottraendo a tasso  $r$ , allora il valore della call oggi è

$$e^{-rT} \max(S_0 e^{rT} - K, 0) = \max(S_0 - Ke^{-rT}, 0)$$

Per mostrare che è consistente con l'equazione (39), va considerato prima il caso dove  $S_0 > Ke^{-rT}$ . Questo implica che  $\ln \frac{S_0}{K} + rT > 0$ . Siccome  $\sigma$  tende a zero,

$d_1$  e  $d_2$  tendono a  $+\infty$ , così che  $N(d_1)$  e  $N(d_2)$  tendono a 1 e l'equazione (39) diventa

$$c = S_0 - Ke^{-rT}$$

Quando  $S_0 < Ke^{-rT}$ , segue che  $\ln \frac{S_0}{K} + rT < 0$ . Siccome  $\sigma$  tende a zero,  $d_1$  e  $d_2$  tendono a  $-\infty$ , così che  $N(d_1)$  e  $N(d_2)$  tendono a zero e l'equazione (39) fornisce un prezzo per la call pari a zero. Il prezzo della call è di conseguenza sempre  $\max(S_0 - Ke^{-rT}, 0)$  nel caso di  $\sigma$  tendente a zero. In modo del tutto simile si potrebbe vedere come il prezzo delle put sia  $\max(Ke^{-rT} - S_0, 0)$

#### 4.9 La funzione di distribuzione cumulativa normale

L'unico vero problema delle equazioni (39) e (40) è nel calcolare la funzione di distribuzione normale  $N$ . Un'approssimazione polinomiale che dà un'accuratezza di sei cifre decimali è

$$N(x) = \begin{cases} 1 - N'(x)(a_1k + a_2k^2 + a_3k^3 + a_4k^4 + a_5k^5) & \text{se } x \geq 0 \\ 1 - N(-x), & \text{se } x < 0 \end{cases}$$

dove

$$k = \frac{1}{1 + \gamma x}, \quad \gamma = 0.2316419, \quad a_1 = 0.319381530, \quad a_2 = -0.356563782,$$

$$a_3 = 1.781477937, \quad a_4 = -1.821255978, \quad a_5 = 1.330274429$$

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

Il prezzo di un'azione a sei mesi dalla fine di un'opzione è 42, il prezzo di esercizio dell'opzione è 40, il tasso di interesse libero da rischio è del 10% annuo e la volatilità è del 20% annuo. Questo significa che  $S_0 = 42$ ,  $K = 40$ ,  $r = 0.1$ ,  $\sigma = 0.2$ ,  $T = 0.5$

$$d_1 = \frac{\ln \frac{42}{40} + (0.1 + \frac{0.2^2}{2}) \times 0.5}{0.2\sqrt{0.5}} = 0.7693$$

$$d_2 = \frac{\ln \frac{42}{40} + (0.1 - \frac{0.2^2}{2}) \times 0.5}{0.2\sqrt{0.5}} = 0.6278$$

e

$$Ke^{-rT} = 40e^{-0.05} = 38.049$$

Di conseguenza, se un'opzione è Europea call, il suo valore  $c$  è dato da

$$c = 42N(0.7693) - 38.049N(0.6278)$$

Se l'opzione è invece di tipo Europea put, il suo valore  $p$  è dato da

$$p = 38.049N(-0.6278) - 42N(-0.7693)$$

Usando l'approssimazione polinomiale accennata prima si ricava:

$$N(0.7693) = 0.7791, N(-0.7693) = 0.2209, N(0.6278) = 0.7349, N(-0.6278) = 0.2651$$

tale che

$$c = 4.76, p = 0.81$$

Il prezzo dell'azione dovrà salire di 2.76 per il compratore della call al pareggio, allo stesso modo il prezzo dell'azione dovrà scendere di 2.81 per il compratore della put al pareggio.

## 5 Modelli che sfruttano le catene di Markov

La semplicità sia nell'uso che nell'attuazione delle catene di Markov sono aspetti che attraggono l'intenzione di cercare di descrivere i fenomeni da analizzare con questi formalismi.

### 5.1 Stilare graduatorie di bontà per i titoli

Le catene di Markov possono essere d'aiuto alle aziende di analisi delle risorse proposte dal mercato per stimare un rank tra i prodotti. Vengono considerate le catene di Markov a tempo omogeneo<sup>16</sup>. Questa è un'assunzione forte: ad esempio può essere pensato che queste probabilità di transizione dipendano dalle condizioni macroeconomiche, oppure da condizioni specifiche per un settore industriale.

Nel caso in cui, ad esempio,  $X_t$  identificasse la categoria di rischio di una singola risorsa, se avessimo un valore pari a  $X_t = (0, 0, 1, 0, \dots, 0)$  starebbe a significare che la risorsa è nel terzo gruppo di rischio e quindi  $X_{t+1} = (p_{31}, p_{32}, \dots, p_{3k})$  sarebbe la distribuzione di probabilità per un bene al tempo  $t + 1$  dato che si trova nello stato 3 al tempo  $t$ .

In alternativa, nel caso in cui  $X_t$  sia lei stessa la distribuzione di una risorsa nelle categorie di rischio al tempo  $t$ , allora  $X_{t+1}$  è la corrispondente distribuzione al tempo  $t + 1$ . Lo stesso metodo probabilistico si applica sia ad un singolo bene che ad un intero portfolio.

La storia non importa per determinare lo stato attuale e questa affermazione, nel contesto in analisi, è equivalente a dire che dato un rank per la risorsa interessata, la probabilità che la risorsa si sia mossa ad un livello, o che questa si muova, è indipendente dalle sue passate posizioni. Questa proprietà può sembrare andar contro al modo di vedere il rank attuale come portatore di informazioni riguardo i possibili "slanci". Condizionata dall'attuale rank, una futura caduta del titolo è presumibilmente molto più probabile se la risorsa ha subito

---

<sup>16</sup>Questa caratteristica non è richiesta dalla specifica generale Markoviana, ma è un'assunzione caratterizzante in pratica (permette di rendere possibili molte applicazioni empiriche). Omogeneità temporale significa che la matrice di transizione  $P$ , il cui elemento  $i, j$ -esimo è la probabilità che dallo stato  $i$  passi allo stato  $j$ , è costante nel tempo.

una caduta nel periodo precedente piuttosto che se avesse avuto un aumento o nessun cambiamento nel rank.

Per riportare un esempio, viene considerata una semplice catena con due stati (riferendosi a risorse poco rischiose e molto rischiose). Definendo la matrice di transizione per il singolo periodo come:

$$P = \begin{bmatrix} p_{11} & 1 - p_{11} \\ 1 - p_{22} & p_{22} \end{bmatrix}$$

$n_{ij}$  è il numero di movimenti, in un campione di dimensione  $N$ , avvenuti dallo stato  $i$  allo stato  $j$ . La funzione di rischio logaritmico su di un insieme di dati raccolti con il modello di Markov è data da:

$$\ln L(P|\text{dati di un 1-periodo}) = n_{11} \ln p_{11} + n_{12} \ln(1 - p_{11}) + n_{21} \ln(1 - p_{22}) + n_{22} \ln p_{22}$$

La massimizzazione di questa funzione rispetto ai parametri sconosciuti  $p_{11}$  e  $p_{22}$  in questo caso fornisce buone soluzioni. Le stime sono date da:

$$\hat{p}_{11} = \frac{n_{11}}{n_{11} + n_{12}}, \quad \hat{p}_{22} = \frac{n_{22}}{n_{22} + n_{21}}$$

Queste stime e le loro ovvie generalizzazioni per il caso a  $K$ -stati sono appropriate quando osserviamo transizioni *one-step*.

Supposto di avere dati su due periodi di transizione è intenzione fare una stima di  $p_{11}$  e  $p_{22}$ , le probabilità per un singolo periodo in una catena di Markov a tempo omogeneo. Prima di tutto vengono identificate con  $p_{11}(2)$  e  $p_{22}(2)$  le probabilità sulle transizioni tra due periodi: sono funzioni della sottostante transizione per un singolo periodo attraverso la relazione  $P(2) = P^2$ . Specialmente:

$$P(2) = \begin{bmatrix} p_{11}(2) & 1 - p_{11}(2) \\ 1 - p_{22}(2) & p_{22}(2) \end{bmatrix} = \begin{bmatrix} p_{11}^2 + (1 - p_{11})(1 - p_{22}) & 1 - \dots \\ 1 - \dots & p_{22}^2 + (1 - p_{11})(1 - p_{22}) \end{bmatrix}$$

$n_{11}(2)$  è il numero di osservazioni che rimangono nello stato 1 per due periodi,  $n_{12}(2)$  è il numero di quelle che si muovono dallo stato 1 al 2, etc. La funzione di rischio logaritmico per due periodi di transizione è pari:

$$\begin{aligned} \ln L(P|\text{dati di un 2-periodo}) &= \\ &= n_{11}(2) \ln p_{11}(2) + n_{12}(2) \ln(1 - p_{11}(2)) + n_{21}(2) \ln(1 - p_{22}(2)) + n_{22}(2) \ln p_{22}(2) \end{aligned}$$

dove  $p_{ij}(2)$  è definito in termini di  $p_{ij}$ . Quando i dati per uno e per due periodi sono disponibili, la stima totale è la somma delle stime, ovvero:

$$\ln(P|\text{dati di 1\&2 - periodi}) = n_{11} \ln p_{11} + n_{12} \ln(1 - p_{11}) + n_{21} \ln(1 - p_{22}) + n_{22} \ln p_{22} + \\ + n_{11}(2) \ln p_{11}(2) + n_{12}(2) \ln(1 - p_{11}(2)) + n_{21} \ln(1 - p_{22}(2)) + n_{22}(2) \ln p_{22}(2)$$

che dovrebbe essere massimizzata rispetto a  $p_{ij}$ . L'estensione a più di due periodi e a più di due stati è immediata e segue lo schema descritto. Mentre a prima vista la mancanza di dati per il periodo intercorso può sembrare strana, è un aspetto comune per i dati sul rank di risorse. Per esempio, questi dati possono essere distribuiti solo annualmente, mentre gli operatori per creare ranking vorrebbero che il processo markoviano avesse transizioni mensili o trimestrali. In alternativa gli operatori potrebbero pensare che il processo markoviano abbia una frequenza annuale, ma sono disponibili per l'analisi solo i dati cumulativi di transizione nel corso di più anni.

I parametri base sono le matrici di transizioni per un singolo periodo  $P_1, P_2$  fino a  $P_T$  dove  $T$  è il più lungo intervallo di tempo dove è stata fatta un'osservazione (non corrisponde direttamente al modo in cui i dati sono osservati). Tipicamente sono osservate transizioni di uno, due, fino a massimo di quindici periodi. Queste sono molto differenti dalle transizioni di un periodo osservate al tempo 1, 2, fino al 15. Quando si osservano molte transizioni di un unico periodo, i metodi precedenti possono essere usati per parametri specifici al periodo, e può essere eseguito un test del " $\chi^2$ "<sup>17</sup> per verificare l'uguaglianza dei parametri tra i periodi.

---

<sup>17</sup>Questo test è un metodo di verifica d'ipotesi usato in statistica che utilizza la variabile casuale  $\chi^2$  per affermare se l'ipotesi è probabilisticamente compatibile con i dati. A seconda delle ipotesi di partenza usate per costruire i test, questi ultimi vengono distinti in più categorie.

I risultati ottenuti nei campioni non sempre concordano con i risultati teorici attesi, al contrario è raro che questo si verifichi. Ad esempio, in linea teorica su 100 lanci di moneta è atteso come risultato 50 teste e 50 croci, è raro che questi risultati siano ottenuti esattamente, ma non si deve comunque concludere che la moneta sia truccata.

Lo scopo del test  $\chi^2$  è quello di conoscere se le frequenze osservate differiscono

Nella situazione presentata, quando le statistiche consistono in informazioni su periodi multipli, il problema è semplificare riparametrizzando per probabilità di transizione. Certamente le probabilità dipendono dai parametri  $P_1, \dots, P_T$  e questa dipendenza complica leggermente le cose, poiché il dato della transizione al passo  $t$ -esimo  $n_{ij}(t)$  dipende dalla matrice di transizione al passo  $t$ -esimo, data da  $P_1 P_2 \dots P_{t-1}$ . In ogni caso, in questo scenario, non sono necessarie stime separate di  $P_t$ ; ciò che è necessario è massimizzare il valore del rischio logaritmico rispetto ai parametri  $P(1) = P_1, P(2) = P_1 P_2, P(3) = P_1 P_2 P_3, \dots$  tramite la formula:

$$\hat{p} = \frac{n_{ij}(t)}{\sum_{k=1}^K n_{i,k}(t)}$$

Il valore massimizzato per il rischio logaritmico è lo stesso sia rispetto ai parametri  $P(t)$  che  $P_t$ .

Nell'applicazione al ranking, lo stato peggiore è l'*absorbing state* (lo stato senza transizioni uscenti). Per convenzione il primo stato della catena è considerato il migliore mentre lo stato  $K$ -esimo il peggiore. Alcune tra le più importanti aziende di supporto agli investitori, che sfruttano questo modello, si avvalgono di questa convenzione. Tra queste ci sono, ad esempio, *Moody's* <http://www.moody.com>, *Standard & Poor* <http://www2.standardandpoors.com/portal/site/sp/en/eu> e *Fitch* <http://www.fitchratings.it/>.

Nella matrice  $P$  la riga  $K$ -esima è costituita da  $(0, 0, \dots, 1)$ . Inoltre, una matrice delle transizioni  $K \times K$  ha  $(K - 1)^2$  parametri, una riga è nota ed una colonna è limitata dal fatto che la somma delle righe è pari a 1. Ci sono altre restrizioni come ad esempio  $p_{i,j} \in [0, 1]$  (valori normalizzati) che permettono una riparametrizzazione più conveniente prima della massimizzazione. Il numero di parametri, nelle applicazioni pratiche, non è troppo grande e questo non aumenta con il numero di periodi disponibili.

---

significativamente dalle frequenze teoriche.

Se  $\chi^2 = 0$ , le frequenze osservate coincidono con quelle teoriche. Se invece  $\chi^2 > 0$ , esse differiscono. Più grande è il valore di  $\chi^2$ , più grande è la discrepanza tra le frequenze osservate e quelle teoriche. Le circostanze in cui il  $\chi^2$  è troppo vicino allo zero sono da guardare con sospetto, poiché è raro che le frequenze osservate concordino troppo bene con le frequenze teoriche.

### 5.1.1 Il test del rapporto di rischio

È possibile, adesso, provare che i dati di transizione sono generati da una catena di Markov a tempo omogeneo. Per eseguire il test, viene utilizzato un rapporto di rischio statistico che confronti il rischio ristretto calcolato generato da una catena di Markov a tempo omogeneo con un rischio alternativo calcolato senza imporre le restrizioni di omogeneità. La statistica è calcolata come  $-2 \ln LR$  dove

$$\ln LR = \max_P \ln L(P|\text{tutti i dati}) - \max_{P(1), \dots, P(T)} \ln L(P(1), \dots, P(T)|\text{tutti i dati})$$

Questo asintoticamente è il test del  $\chi^2$  dove il grado di libertà  $q$  è il numero di restrizioni imposte. Supposto di avere  $K$  stati e  $[1 \dots T]$  periodi di transizione, allora il modello ristretto ha  $(K - 1)^2$  parametri, mentre quello alternativo ne ha  $T(K - 1)^2$ . Quindi  $q = (T - 1)(K - 1)^2$ .

Dato che esiste una forma chiusa, è possibile fare il calcolo del valore alternativo massimizzando il rischio logaritmico. Il calcolo del valore massimo di rischio logaritmico per la versione ristretta è più difficile quando sono disponibili transizioni di più periodi. È utile riparametrizzare trasformando, con una funzione Logit<sup>18</sup>, nei nuovi parametri

$$\alpha_{ij} = \frac{\exp(p_{ij})}{\sum_{j=1}^K \exp(p_{ij})}$$

e massimizzare rispetto ai  $(K - 1)^2$  parametri “liberi”  $\alpha_{ij}$ . Questo tipo di massimizzazione viene fornita da molti pacchetti software. Valori convenienti per partire sono forniti dalla trasformazione Logit delle frazioni che compongono il campione di dati del primo periodo di transizione. Infine, è possibile escludere le transizioni con probabilità zero, poiché nessuna transizione, per poco probabile che sia, sarà veramente impossibile.

Per transizioni su di un lungo periodo deve essere impiegato un altro modello. In pratica questo modello riesce a fare stime fino a cinque anni in avanti, quindi,

<sup>18</sup>Il Logit è una funzione che si applica a valori compresi nell'intervallo (0,1), tipicamente valori rappresentanti probabilità. Viene definito come  $\text{Logit}(p) = \ln\left(\frac{p}{1-p}\right) = \ln(p) - \ln(1-p)$ .

In finanza questa tenta di determinare l'intensità o la magnitudine delle intenzioni del cliente a comprare, e traduce questa in una misura dell'attuale comportamento sugli acquisti.

L'analisi Logit assume che una necessità insoddisfatta nel mercato sia già stata trovata, e che un prodotto sia stato creato per coprire quel bisogno. Lo scopo dell'analisi Logit è quantificare le potenziali vendite del prodotto in questione.

per tutte quelle risorse che richiedono un'analisi in un intervallo di tempo minore, questo modello è adeguato; in casi in cui si debbano compiere analisi annuali questo modello deve essere sostituito con altri.

## 5.2 Un modello per estrarre la volatilità dal prezzo

La volatilità, come già è stato ampiamente discusso, misura l'ampiezza delle fluttuazioni nel ritorno ed è una delle quantità centrali in finanza. L'assunzione che questa quantità sia una costante indipendente dal tempo è al centro dei modelli classici, ma questa ipotesi non è sufficiente a spiegare l'eterogeneità comportamentale dei mercati. La volatilità deve essere trattata come una quantità casuale con una propria dinamica.

La volatilità è la responsabile del clustering osservabile nel cambiamento nei prezzi: grandi fluttuazioni sono comunemente seguite da altre grandi fluttuazioni (funziona, in modo analogo, per le piccole fluttuazioni). Un'altra caratteristica è che, in netto contrasto con le variazioni di prezzo che mostrano correlazioni trascurabili, la correlazione automatica rimane significativa per ritardi temporali più grandi di un anno.

L'ostacolo principale di un'appropriata analisi della volatilità è che questa non sia direttamente osservabile. La volatilità fornisce informazioni importanti per gli operatori economici, ma non è molto chiaro in che modo si possa fare una stima del processo che descrive questa grandezza. Gli investitori usano molti surrogati per dedurre il livello attuale di un prodotto, i metodi utilizzati più comunemente sono:

- renderlo equivalente al valore assoluto delle variazioni nel guadagno;
- assumere una proporzionalità tra volatilità e volume di mercato;
- utilizzare le informazioni contenute nei prezzi delle opzioni e ottenere una sorta di "volatilità implicita".

Qui viene presentata una procedura alternativa per stimare la volatilità dalle dinamiche del prezzo, che sono rappresentate da un processo di diffusione a due dimensioni: una dimensione per il prezzo e la seconda per la volatilità descritta da una catena di Markov nascosta. Viene posta attenzione su di un

particolare modello stocastico per la volatilità capace di eludere le difficoltà sia matematiche che computazionali: il modello esponenziale di Ornstein-Uhlenbeck per la volatilità. Come indica il nome, il modello assume che il logaritmo della volatilità segua un processo di Ornstein-Uhlenbeck<sup>19</sup>.

### 5.2.1 Il modello stocastico per la volatilità

Il moto geometrico Browniano (GBM) è il modello più largamente usato in finanza. In questo esempio il prezzo della risorsa  $S(t)$  è descritto tramite l'equazione di Langevin<sup>20</sup>. L'equazione di Langevin è equivalente all'equazione

---

<sup>19</sup>Quest'ultimo è un processo Gaussiano che ha varianza limitata ed ammette una distribuzione stazionaria di probabilità, in contrasto con il processo di Wiener; la differenza tra i due è il loro "termine di slittamento". Per il processo di Wiener questo slittamento è costante, mentre per il processo di Ornstein-Uhlenbeck dipende dal valore corrente del processo: se il valore attuale del processo è inferiore alla media (a lungo termine), lo slittamento sarà positivo; se il valore attuale del processo è superiore (sempre a lungo termine) alla media, lo slittamento sarà negativo. In altre parole, la media si comporta come un livello di equilibrio per il processo.

<sup>20</sup>Nel campo della fisica statistica, un'equazione di Langevin è un'equazione differenziale stocastica che descrive moto browniano in un potenziale.

Le prime equazioni di Langevin ad essere studiate sono state quelle a potenziale costante, in modo che l'accelerazione browniana  $\mathbf{a}$  di una particella di massa  $m$  fosse espressa come la somma di una forza viscosa, che è proporzionale alla velocità  $\mathbf{v}$  della particella (legge di Stokes), un termine di rumore  $\boldsymbol{\eta}(t)$  rappresentate l'effetto di una serie continua di collisioni con gli atomi del fluido e  $\mathbf{F}(x)$ , che è la forza di interazione causata dalle interazioni intramolecolari e intermolecolari.

usata nella teoria della diffusione di  $\hat{\text{Ito}}^{21}$ . Il valore che assume  $S(t)$  è pari a

$$\frac{dS(t)}{S(t)} = \mu dt + \sigma dW_1(t) \quad (43)$$

dove  $\sigma$  è la volatilità, assunta essere costante,  $\mu$  è un qualche slittamento deterministico che indica un eventuale trend nel mercato e  $W_1(t)$  è il processo di Wiener. Data la definizione

$$X(t) = \ln\left[\frac{S(t+t_0)}{S(t_0)}\right] - \left\langle \ln\left[\frac{S(t+t_0)}{S(t_0)}\right] \right\rangle \quad (44)$$

del ritorno a media zero, dove il simbolo  $\langle \dots \rangle$  indica il valore medio e  $t_0$  è il tempo iniziale, che normalmente è esattamente zero. In termini di  $X(t)$  il GBM può essere riscritto come

$$dX(t) = \sigma dW_1(t) \quad (45)$$

In ogni caso, come è stato presentato, la volatilità non è costante e nemmeno in funzione del tempo, ma un valore casuale; dati empirici convincenti hanno messo in evidenza la dubbia correttezza delle affermazioni precedenti. La volatilità  $\sigma$  è assunta come una data funzione di un processo casuale  $Y(t)$ :

$$\sigma(t) = f[Y(t)] \quad (46)$$

La maggior parte dei modelli assumono che la volatilità stocastica (SV) sia anche un processo di diffusione che può essere o non essere correlato con il

---

<sup>21</sup>Una diffusione di  $\hat{\text{Ito}}$  (a tempo omogeneo) in uno spazio Euclideo a  $n$ -dimensioni  $R^n$  è un processo  $X : [0, +\infty) \times \Omega \rightarrow R^n$  definito su di uno spazio di probabilità  $(\Omega, \Sigma, P)$  che soddisfi l'equazione differenziale della forma:

$$dX_t = b(X_t) dt + \sigma(X_t) dB_t$$

dove  $B$  è un moto Browniano a  $m$ -dimensioni e  $b : R^2 \rightarrow R^2$  e  $\sigma : R^2 \rightarrow R^{n \times m}$  soddisfa la continuità di Lipschitz

$$|b(x) - b(y)| + |\sigma(x) - \sigma(y)| \leq C|x - y|$$

(una condizione di morbidezza per le funzioni che è più forte della continuità regolare. Intuitivamente, una funzione che soddisfa questa condizione è limitata in quanto può cambiare velocemente, una linea che congiunge due punti qualsiasi del grafico di questa funzione non avrà mai una pendenza più ripida di un certo numero, chiamato costante di Lipschitz della funzione) per qualche costante  $C$  e qualsiasi  $x$  e  $y$  in  $R^n$ . Questa condizione garantisce l'esistenza di un'unica soluzione  $X$  all'equazione differenziale stocastica data sopra. Il vettore  $\mathbf{b}$  è detto **coefficiente di slittamento** di  $X$ ; l'elemento della matrice  $\sigma$  è detto **coefficiente di diffusione** di  $X$ .

prezzo. La principale differenza tra i vari modelli è solo la parametrizzazione di questo schema. Una notazione generale per il ritorno a media zero  $X(t)$ , definito dall'equazione (44), è descritta da un insieme di equazioni differenziali

$$dX(t) = f[Y(t)]dW_1(t) \quad (47)$$

$$dY(t) = -g[Y(t)]dt + h[Y(t)]dW_2(t) \quad (48)$$

dove  $dX = \frac{dS}{S} - \langle \frac{dS}{S} \rangle$  e  $f$ ,  $g$  e  $h$  sono funzioni date di  $Y(t)$ . Come mostrato nell'equazione (46),  $f[Y(t)]$  corrisponde all'ampiezza delle fluttuazioni del ritorno. In ogni caso, poiché  $f(x)$  viene scelta monotona crescente, non è fuorviante pensare  $Y$  come una misura della volatilità. In più, per non creare confusione,  $Y(t)$  verrà chiamata anch'essa "volatilità". D'altra parte, la funzione  $g[Y(t)]$  descrive una forza inversa che porta la volatilità verso una sorta di "livello normale". Questa forza inversa porta il processo di volatilità in un regime stazionario per orizzonti temporali lunghi e il livello normale è in relazione al valore medio della volatilità al limite. Infine, il processo subordinato  $Y(t)$  può avere una diffusione non costante definita in termini della funzione  $h[Y(t)]$  che è chiamata la "volatilità della volatilità". Le funzioni  $g$  e  $h$  descrivono pienamente il processo di volatilità. La dinamica risultante è confrontabile con quella descritta da una particella intrappolata in un potenziale a spirale  $V(y)$  la cui forza associata è  $-g(y)$ , dove  $g(y) = V'(y)$ .

In ciò che segue verrà utilizzato, per la maggior parte, un particolare modello di SV: il modello esponenziale di Ornstein-Uhlenbeck (expOU) che segue dalle sostituzioni

$$f(x) = me^x, \quad g(x) = \alpha x, \quad h(x) = k$$

quindi

$$dX(t) = me^{Y(t)}dW_1(t) \quad (49)$$

$$dY(t) = -\alpha Y(t)dt + kdW_2(t) \quad (50)$$

Il processo  $Y(t)$ , in questo modello, è precisamente il logaritmo della volatilità, chiamata "log-volatilità". È da ricordare, dalla Sezione precedente dove è stata trattata la volatilità, che la distribuzione stazionaria del processo  $Y(t)$  è una Gaussiana con media zero e varianza  $\beta$ :

$$p(y) = \frac{1}{\sqrt{2\phi\beta}} \exp \frac{-y^2}{2\beta} \quad \text{dove } \beta \equiv \frac{k^2}{2\alpha} \quad (51)$$

### 5.2.2 La stima della volatilità

Presi i valori  $\mathbf{X}$  e  $\mathbf{Y}$ , rappresentanti una realizzazione simultanea delle variabili  $X(\tau)$  e  $Y(\tau)$  nell'intervallo di tempo  $t - s \leq \tau \leq t$ . Omettendo tutti i termini indipendenti da  $\mathbf{Y}$ , la densità di probabilità (rischio) della realizzazione è data approssimativamente da:

$$\ln \mathbb{P}(\mathbf{X}, \mathbf{Y}) \simeq -\frac{1}{2} \int_{t-s}^t \left[ \frac{\dot{X}(\tau)}{f[Y(\tau)]} \right]^2 d\tau - \frac{1}{2} \int_{t-s}^t \left[ \frac{\dot{Y} + g[Y(t)]}{h[Y(t)]} \right]^2 d\tau + \dots \quad (52)$$

È da notare innanzitutto che quest'equazione deve essere intesa nel senso di funzioni generalizzate, in quanto il processo di Wiener è derivabile soltanto in questo senso e, quindi,  $\dot{X}(t)$  e  $\dot{Y}(t)$  non esistono come funzioni ordinarie (l'equazione è solo un'espressione simbolica). Tuttavia la formula rimane valida quando l'integrale e le derivate in esso sono discretizzati a piccoli passi temporali. Nell'equazione (52) il primo addendo misura le fluttuazioni del ritorno a media zero rispetto alla volatilità e il loro contributo al rischio di un dato ritorno: più alto è questo contributo più basse saranno le relative fluttuazioni. Allo stesso modo, il secondo addendo misura le fluttuazioni del processo che descrive la volatilità  $Y(t)$  rispetto alla volatilità della volatilità  $h(Y)$ , sebbene in questo caso queste fluttuazioni sono calibrate con la forza respingente media  $-g(Y)$ . Esattamente come prima, più basse saranno le fluttuazioni, più alto sarà il loro contributo nell'indice logaritmico di rischio.

Mentre le equazioni (47) e (48) rappresentano un modello comune per il ritorno e la volatilità, i dati del mercato azionario includono solo il campionamento del processo di guadagno  $X$ . Il processo  $Y$  e la volatilità  $f(Y)$  devono essere dedotti indirettamente attraverso l'equazione (52). Infatti, la densità di probabilità condizionata  $\mathbf{Y}$  della realizzazione del processo nascosto  $Y$ , dato un ritorno osservato  $\mathbf{X}$ , è:

$$\ln \mathbb{P}(\mathbf{Y}|\mathbf{X}) = \ln \mathbb{P}(\mathbf{X}, \mathbf{Y}) - \ln \mathbb{P}(\mathbf{X}) \quad (53)$$

Massimizzando questa equazione rispetto a  $\mathbf{Y}$ , è possibile trovare un cammino del rischio massimo del processo (nascosto) della volatilità. Dato che il secondo addendo dell'equazione (53) è indipendente da  $\mathbf{Y}$ , è possibile trascurare  $\mathbb{P}(\mathbf{X})$  nel processo di massimizzazione e, quindi, la massimizzazione di  $\ln \mathbb{P}(\mathbf{Y}|\mathbf{X})$  fornisce lo stesso risultato che massimizzare  $\ln \mathbb{P}(\mathbf{X}, \mathbf{Y})$ . È da notare che, data la specifica del modello stocastico di volatilità, l'unico parametro libero è  $s$ : la

durata del dato sul passato ritorno da prendere in considerazione. La quantità massima del rischio, quindi è pari a:

$$\hat{\mathbf{Y}} = \operatorname{argmax}_{\mathbf{Y}} \ln \mathbb{P}(\mathbf{Y}|\mathbf{X}) = \operatorname{argmax}_{\mathbf{Y}} \ln \mathbb{P}(\mathbf{X}, \mathbf{Y}) \quad (54)$$

Problemi di massimizzazione nel contesto delle catene di Markov nascoste vengono chiamati “decodificazioni”. Quando lo spazio degli stati (il numero di possibili valori di  $X$  e  $Y$ ) è finito, l’ottimizzazione si può fare attraverso l’algoritmo di Viterbi<sup>22</sup>, mentre, nel caso del continuo, c’è stato un successo limitato.

Il principale obiettivo è quello di creare un metodo capace di filtrare il rumore Wieneriano  $dW_1(t)$  fuori dall’equazione (47) ed inoltre ottenere una stima affidabile  $\hat{Y}(t)$  del processo della volatilità nascosto  $Y(t)$ . Il metodo ha sostanzialmente cinque passi:

1. simulare un campione di cammino casuale del processo di Wiener,  $dW_1(\tau)$  per  $t - s \leq \tau \leq t$ ;
2. generare un surrogato di  $\mathbf{Y}$  da

$$\hat{Y}_s(\tau) = f^{-1} \left( \left| \frac{dX(\tau)}{d\hat{W}_1(\tau)} \right| \right) \quad (55)$$

dove  $t - s \leq \tau \leq t$  ( $f(x)$  deve essere invertibile il che implica  $f(x)$  deve essere scelta monotona);

---

<sup>22</sup>È un algoritmo generalmente utilizzato per trovare la migliore sequenza di stati (detta *Viterbi path*) in una sequenza di eventi osservati in un processo markoviano. L’algoritmo sceglie, basandosi su un processo markoviano, il percorso che è più vicino alla sequenza di simboli ricevuti all’interno del *traliccio*, ovvero del campo di tutte le possibilità.

Il criterio di scelta tra le possibilità può essere:

- la distanza minima di Hamming rispetto alla sequenza ricevuta;
- la distanza euclidea tra i segnali.

Una volta scelto il criterio è applicabile la stessa legge di decodifica. Ad ogni passo, l’algoritmo elimina i percorsi meno probabili fino a rimanere con un solo superstite. L’algoritmo è tanto più prestante quanto il numero di passi è alto. Ovviamente maggiore è il numero di passi e maggiore è la lentezza nella decodifica e maggiore è il dispendio di risorse. La complessità di calcolo del decodificatore si può immaginare calcolando che per un codice con  $i$  stati e  $t$  passi di osservazione, si hanno  $2^{i(t-1)}$  cammini possibili. Ad ogni passo vi sono  $2^i$  cammini che raggiungono ogni singolo stato. Di tutti i cammini uno solo sarà quello a distanza minima fino a quel passo.

3. Sostituire  $\hat{Y}_s$  e  $\mathbf{X}$  nell'equazione (52) per calcolare il rischio logaritmico;
4. Iterare 1-3 per  $I$  passi, mantenere la più alta realizzazione del rischio causale e assumere questo come la stima propria  $\hat{Y}(t)$ ;
5. La stima del processo nascosto al tempo  $t$  è quindi  $\hat{Y}(t)$ . La stima della volatilità è data da

$$\hat{\sigma}(t) = f[\hat{Y}(t)] \quad (56)$$

In finanza, la volatilità è spesso indentificata con il valore assoluto delle variazioni nel ritorno. Infatti, come prima approssimazione, è possibile sostituire nell'equazione (47) il termine di rumore con il suo valore atteso e scrivere

$$\sigma(t) \approx \frac{|dX(t)|}{\langle |dW_1(t)| \rangle} \quad (57)$$

L'equazione (47) può essere quindi pensata come una prima approssimazione verso la stima della volatilità. Il metodo basato sulla massimizzazione dell'equazione (53) porta questa stima due passi in avanti. Il primo passo viene fatto sostituendo la media  $\langle |dW_1(t)| \rangle$  con un cammino simulato, mentre il successivo passo sostituisce il rumore Wieneriano con una simulazione casuale ed esegue il metodo per calcolare il massimo rischio descritto dai punti 1-5 della precedente lista. In pratica in questo modo vengono separati i ritorni osservati  $dX(t)$  in due sorgenti:  $\sigma(t)$  e  $dW_1(t)$ . Per fare questo, c'è da considerare una forma particolare di volatilità stocastica, in secondo luogo viene assunto che i due rumori Wieneriani  $dW_1(t)$  e  $dW_2(t)$ , che appaiono nelle equazioni (47) e (48), non siano correlati. Inoltre deve essere fatta l'assunzione che  $\sigma(t)$  sia approssimativamente costante durante tutto il tempo nel quale vengono considerate numericamente  $\dot{Y}(t)$  e  $\dot{X}(t)$  che appaiono nell'equazione (52). È da notare che se  $h(x) = 0$ , la volatilità della volatilità uguale a zero, la soluzione stazionaria dell'equazione (48) è  $Y(t) \equiv 0$ . Di conseguenza il modello utilizza un processo di Wiener nel quale la volatilità è costante.

Per provare le performance del criterio espresso, viene fatta una simulazione del processo expOU, usando le equazioni (49) e (50) con valori realistici. La relazione tra il valore simulato della volatilità logaritmica  $Y(t)$  e la sua stima  $\hat{Y}(t)$  è data dalla Figura 12. Le due quantità giacciono tra le barre di errore, è

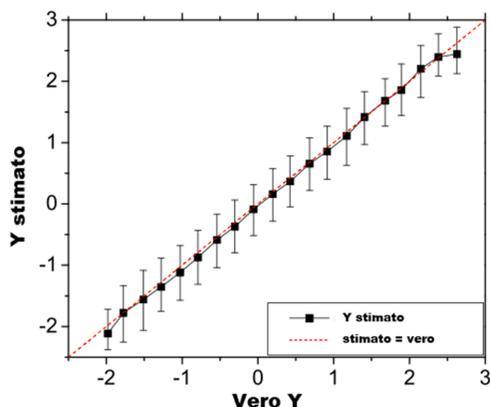


Figura 12: Valore stimato della volatilità logaritmica  $\hat{Y}$  come una funzione del valore effettivo  $Y$  dato da  $2 \times 10^5$  simulazioni del modello expOU. La ricostruzione usa gli ultimi  $s = 10$  valori del ritorno e  $I = 10^5$  iterazioni. Le barre dell'errore rappresentano i quantili 25% e 75% della stima della distribuzione della volatilità.

possibile affermare che:

$$Y(t) \approx \hat{Y}(t)$$

In ciò che segue verranno sempre usati  $s = 10$  giorni di dati precedenti e  $I = 10^5$  iterazioni per massimizzare. Il passo di tempo per la discretizzazione è  $\Delta t = 1$  giorni. Per l'ottimizzazione casuale, il numero necessario di iterazioni  $I$  cresce molto velocemente, quasi esponenzialmente, con  $\frac{s}{\Delta t}$ . I valori di  $s$  e  $\Delta t$  citati qui sono scelti per mantenere la computazione fattibile, ad esempio, per il valore  $I$  un incremento delle iterazioni a  $10^6$  non porta a miglioramenti degni di nota. Le stime generate con questi parametri hanno insignificanti distorsioni sistematiche e possono essere distinte tra periodi di bassa e alta volatilità.

Una verifica supplementare della buona performance della stima è mostrata in Figura 13, dove viene dato il cammino di prova di  $Y$  per una singola realizzazione del modello expOU insieme alla stima  $\hat{Y}$ . Come possibile vedere la stima segue la vera volatilità logaritmica  $Y(t)$  in modo ravvicinato.

Un altro possibile approccio è quello di valutare la performace di previsione del modello. È possibile, per esempio, usare l'errore medio del ritorno assoluto

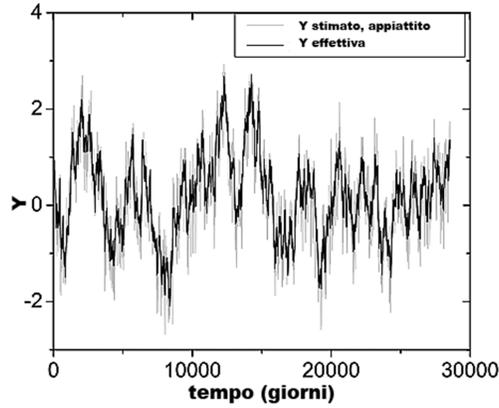


Figura 13: Volatilità stimata e attuale per un tipico cammino del metodo expOU. I valori stimati sono appiattiti dalla media dei cinque vicini per ridurre il rumore.

$|dX(t+h)|$  per un dato cammino Monte Carlo simulato:

$$E(h) = \left\langle \left| |d\hat{X}(t+h)| - |dX(t+h)| \right| \right\rangle \quad (58)$$

dove  $|d\hat{X}(t+h)|$  è la stima per il ritorno assoluto  $|dX(t+h)|$  data una stima  $\hat{Y}(t)$  mentre  $h$  è l'orizzonte temporale (numero di giorni) su cui fare la previsione. Per questo tipo di stime è possibile impiegare media condizionale per le simulazioni Monte Carlo. Di conseguenza per il modello expOU risulta

$$|d\hat{X}(t+h)| = M[|dW_1(t+h)|] m \exp[\hat{Y}(t) \exp(-\alpha h)] \quad (59)$$

dove  $M[|dW_1(t+h)|]$  è la media del valore assoluto dell'incremento di Wiener. La Figura 14 confronta cinque metodi per predire  $|dX(t+h)|$  usando solo informazioni accessibili prima del tempo  $t$  (per la descrizione viene fatto riferimento al colore della linea).

-AZZURRA La media degli ultimi cinque ritorni assoluti. Questo metodo ha un potere di previsione per  $h < 150$  giorni.

-ROSA La media degli ultimi quindici ritorni assoluti. I periodi più lunghi per fare la media porta un maggiore miglioramento per le previsioni a breve termine, ma non per quelle a lungo termine.

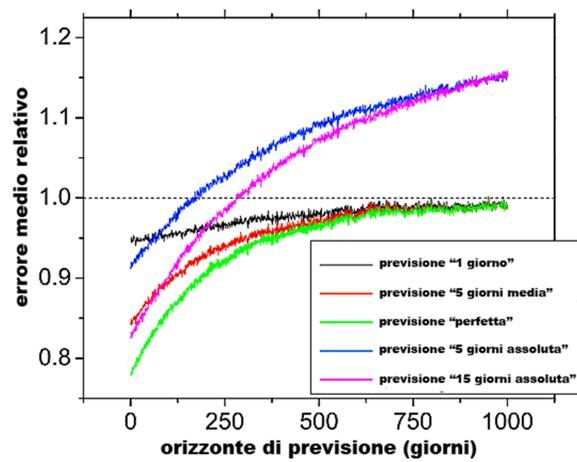


Figura 14: Previsione dell'errore medio  $E(h)$  del ritorno assoluto calcolato in accordo con l'equazione (59) tramite cinque metodi. Gli errori sono stati normalizzati con il più basso livello di errore raggiungibile con l'assunzione di un volatilità indipendente dal tempo (la linea orizzontale tratteggiata). Il risultato è stato mediato su 15000 realizzazioni indipendenti del processo expOU.

- VERDE L'equazione (59) con  $\hat{Y}(t) = Y(t)$ , che è il vero valore per la corrente simulazione del modello expOU. Questa diminuisce in modo sostanziale l'errore di previsione, che persiste fino a  $h \sim 500$  giorni. Questa curva è anche il limite inferiore raggiungibile per gli errori da qualsiasi serie temporale. I dati sono perfettamente distribuiti per la expOU ed i parametri e  $Y(t)$  sono conosciuti "perfettamente". Questo non accade mai per i dati reali, quindi ci si aspetta che questa abbia performaces cattive.
- NERO L'equazione (59) con  $\hat{Y}(t)$  stimato tramite procedure di ricostruzione. Il modello non si comporta benissimo, a causa di rumore nella procedura.
- ROSSO L'equazione (59) con  $\hat{Y}(t)$  stimato tramite procedure di ricostruzione, mediato sugli ultimi cinque giorni. L'uso della media diminuisce il rumore e l'accuratezza della previsione è incrementata nel breve termine.

Per  $h \sim 500$  giorni tutti i modelli basati sul modello expOU perdono qualsiasi informazione inclusa in  $\hat{Y}(t)$ , e convergono tutti allo stesso livello di errore.

### 5.2.3 Applicazione ai dati del mercato delle azioni

In questa Sezione viene presentata un'applicazione del metodo sui dati del mercato delle azioni: viene preso l'indice Dow Jones (Dow Jones Industrial Average, detto anche DJIA) nell'arco di tempo 01/01/1900 - 26/05/2006, per un totale di 29038 giorni. Per poter lavorare su ritorni a media zero, il valore medio viene sottratto dal dato reale. Volumi di scambio per l'indice sono disponibili solo dal 01/04/1993 al 26/05/2006, per un totale di 3375 giorni.

Per la procedura di stima presentata è necessario assumere che venga utilizzato un modello teorico per la volatilità. Fatto questo, il successivo passo è stimare i parametri che sono interessati dal modello scelto. Per il modello expOU, le equazioni (49)-(50), questi parametri sono  $m$ ,  $k$  e  $\alpha$ .

Prima di procedere oltre è da notare che gli incrementi di tempo nei dati reali hanno una dimensione finita poiché il mercato lavora sempre in tempi discreti (per dati giornalieri l'incremento è pari a 1 giorno). Di conseguenza, in pratica, la variazione (infinitesimale) nel ritorno  $dX(t) = X(t + dt) - X(t)$  corrisponde ad un incremento (finito) nel ritorno  $\Delta X(t) = X(t + \Delta t) - X(t)$  dove  $\Delta t$  è

l'intervallo di tempo tra due “tick” consecutivi. Anche il differenziale di Wiener  $dW(t)$  corrisponde all'incremento

$$\Delta W(t) \approx \epsilon(t)\sqrt{\Delta t} \quad (60)$$

dove  $\epsilon(t)$  è un processo Gaussiano con media zero e varianza unitaria. Nel caso presentato gli intervalli temporali hanno larghezza fissata uguale a  $\Delta t = 1$  giorno. Per il parametro  $m$  è possibile fare una stima utilizzando la formula

$$\ln m \approx \frac{\gamma + \ln 2}{2} + \left\langle \ln \frac{|\Delta X|}{\sqrt{\Delta t}} \right\rangle \quad (61)$$

dove  $\gamma = 0.5772\dots$  è la costante di Eulero.

D'altra parte, se il modello expOU è appropriato allora la stima empirica di  $\hat{Y}(t)$  della volatilità nascosta  $Y(t)$  dovrebbe essere un processo Gaussiano con una distribuzione stazionaria a media zero e varianza data da  $\beta = \frac{k^2}{2\alpha}$ . Come visto nella Figura 15, se viene preso  $\beta = 0.61 \pm 0.05$  la distribuzione di  $\hat{Y}(t)$  è Gaussiana e coincide con la distribuzione teorica di  $Y(t)$  data dall'equazione (51). L'assunzione di una distribuzione Gaussiana per la stima fatta è valida per tutto l'intervallo di valori che i parametri posso assumere.

Per specificare completamente il modello deve essere recuperato il parametro  $\alpha$ . Nel caso presentato i parametri hanno i seguenti valori:

$$m = (7.5 \pm 0.5) \times 10^{-3} \text{ giorni}^{-\frac{1}{2}},$$

$$\alpha = (1.82 \pm 0.03) \times 10^{-3} \text{ giorni}^{-1}$$

$$k = (4.7 \pm 0.3) \times 10^{-2} \text{ giorni}^{-\frac{1}{2}}$$

Gli errori sono determinati in base alla Figura 15.

Le distribuzioni della volatilità logaritmica sono confrontate in base alla Figura 15 per i quattro casi. La procedura di semplificazione è determinata da:

$$\hat{Y}_{decon}(t) = \ln \frac{1}{m} \left| \frac{dX(t)}{d\hat{W}_1(t)} \right| \quad (62)$$

dove  $W_1(t)$  è la simulazione del processo di Wiener. Questa procedura di semplificazione ha una distribuzione Gaussiana ma con una varianza più grande

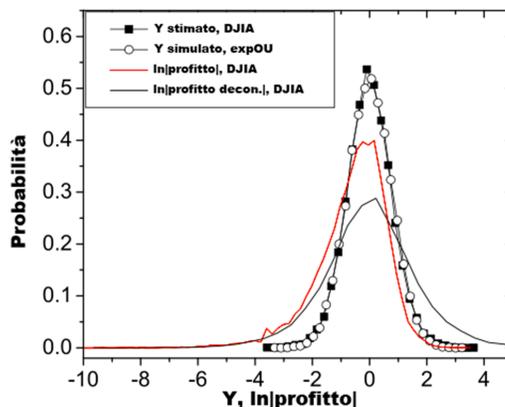


Figura 15: Correlazione delle stime della volatilità logaritmica per il Dow Jones. I quadrati neri rappresentano il metodo per il massimo rischio applicato ai dati empirici. I cerchi vuoti rappresentano la distribuzione dei cammini rappresentativi per la volatilità logaritmica, assumendo che  $Y(t)$  segua il modello expOU. Sono anche rappresentate due distribuzioni empiriche di due stime: la linea rossa è ottenuta con l’equazione (57) e la linea nera tramite l’equazione (62).

che quella suggerita nella Figura 15. Infine la stima per la volatilità logaritmica  $\ln \frac{\sigma}{m} \approx \ln |dX|$ , che è data dall’equazione (57) mostra una distribuzione non Gaussiana e parziale. Questo suggerisce che  $\hat{Y}_{decon}$  è un “modello nullo” in contrasto con  $\hat{Y}$ ; entrambe le quantità sono generate dividendo  $dX(t)$  con l’incremento di una realizzazione del processo di Wiener e, successivamente, prendendo il logaritmo del valore assoluto di questo rapporto. La differenza sta nel fatto che  $\hat{Y}$  prende realizzazioni che soddisfano un requisito di massimo rischio, mentre  $\hat{Y}_{decon}$  prende una realizzazione di un processo di Wiener che è puramente casuale.

Per mostrare che la tecnica presentata è abbastanza potente da filtrare il rumore  $dW_1$  al di fuori del profitto viene data una valutazione visiva basata su esperimenti. Nella Figura 16 è mostrata una valutazione su di un intervallo di 1000 giorni. È possibile osservare che il livello di rumore per  $\hat{Y}$  è decisamente minore che in  $\hat{Y}_{decon}$ , come anche visto in Figura 15

Per mostrare che questa relazione è responsabile della soppressione di grandi

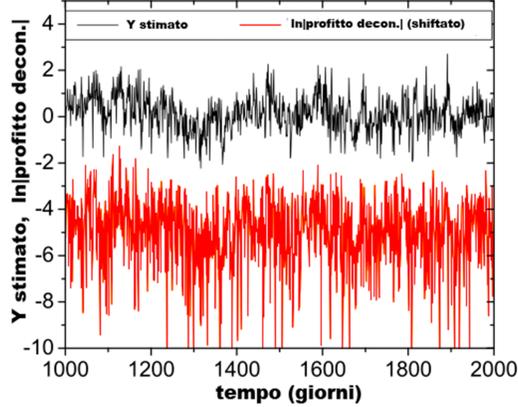


Figura 16: Valutazione tra la stima  $\hat{Y}(t)$  e  $\hat{Y}_{decon}(t)$  per un periodo di 1000 giorni del Dow Jones. Queste curve non sono livellate per mostrare la sostanziale riduzione del livello di rumore e l'asimmetria in  $\hat{Y}(t)$ .

fluttuazioni nel rapporto, viene mostrato un secondo risultato: in Figura 17 è possibile vedere una comparazione tra il logaritmo delle variazioni del profitto,  $\ln|dX|$ , e la volatilità stimata  $\hat{Y}$ .

La natura nascosta del processo di volatilità è stata trattata sotto molti aspetti, uno di questi suggerisce che, invece della volatilità, una buona stima potrebbe essere la radice quadrata del volume commerciato giornalmente che è:

$$M[\sigma(t)] \propto V(t)^\alpha \tag{63}$$

dove  $\alpha = 0.5$  e  $M[.]$  denota la media. Nella Figura 16 è mostrata l'evidenza che fa valere l'assunzione. Ancora la prima stima per la volatilità è  $|dX(t)|$ .

Ancora nella stessa figura è presentata la regressione tra la stima massima di rischio  $\hat{Y}(t)$  e  $\ln V(t)$  che appare essere meno ricco di rumore rispetto alla regressione tra  $\ln|dX(t)|$  e  $\ln V(t)$ , in accordo con la più piccola varianza di  $\hat{Y}(t)$  rispetto a  $\ln|dX(t)|$ . Comunque, l'esponente  $\alpha \approx 0.55$  è in entrambe le regressioni. L'equazione (63) non implica che la volatilità sia proporzionale al volume, solo che "il suo tipico valore è" (in questo caso la media). Le fluttuazioni intorno al comportamento medio attribuibili ai cambiamenti nella liquidità possono assumere un ruolo chiave nel processo.

Dalle equazioni (49)-(50) viene notato che per il modello expOU semplici

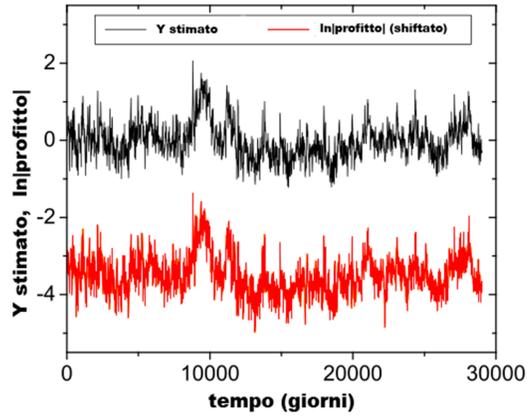


Figura 17: La stima  $\hat{Y}(t)$  ed il logaritmo delle variazioni nel profitto per l'intero Dow Jones.

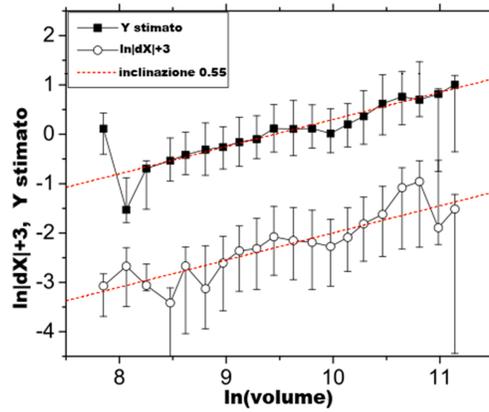


Figura 18: Logaritmo del profitto assoluto e stimata volatilità  $\hat{Y}$  come funzione del volume giornaliero. Giorni con volume simile sono scartati per una visibilità migliore. I simboli rappresentano le medie, e le barre degli errori il 25%-75% (quantili scartati).

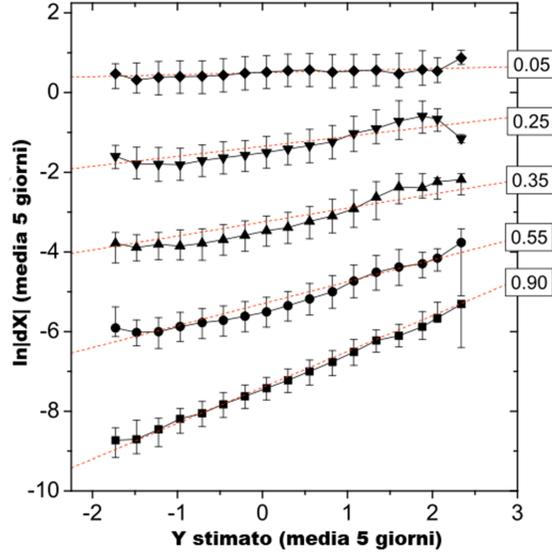


Figura 19: Proporzionalità tra la volatilità stimata e il logaritmo delle variazioni del profitto assoluto. In ordine di diminuire il rumore, sono impiegate medie a cinque giorni. I numeri a destra indicano l'inclinazione della corrispondente retta di regressione. Gli intervalli di tempo sono (dal basso verso l'alto):  $h = 0$  giorni, 5 giorni, 20 giorni, 100 giorni. I giorni con profitti simili sono stati scartati per una migliore visibilità. I simboli rappresentano le medie e le barre degli errori il 25%-75% (quantili scartati).

relazioni possono essere date tra  $\ln |dX(t)|$  e  $Y(t)$ :

$$\ln |dX(t)| = \ln(m|dW_1(t)|) + Y(t)$$

In più, la media condizionale di  $\ln |dX(t)|$  data  $Y(t)$  è:

$$M \left[ \ln |dX(t)| \middle| Y(t) \right] = \text{const.} + Y(t) \quad (64)$$

Questa relazione implica che in qualche modo sia possibile predire variazioni nel profitto se conosciuto il valore corrente della volatilità logaritmica  $Y(t)$ . Viene testata l'equazione (64) con dati reali e con  $Y(t)$  sostituita dalla stima  $\hat{Y}(t)$ . Come mostrato dalla curva più in basso della Figura 19, l'inclinazione della regressione lineare tra  $M[\ln |dX(t)| | Y]$  e  $\hat{Y}(t)$  non è uguale a 1, come doveva essere implicata dall'equazione (64), ma 0.9 che suggerisce un forte potere di predizione.

Il passo minimo temporale del dato empirico è 1 giorno. L'equazione (64) implica la predizione del profitto di domani basato sulla volatilità ed il profitto odierno.

## Appendice

### Prova della formula di Black-Scholes-Merton

Viene provato il risultato di Black-Scholes [1] dando prima un altro risultato chiave.

**Risultato chiave:** se  $V$  è distribuito lognormalmente e la deviazione standard di  $\ln V$  è  $s$ , allora

$$E[\max(V - K, 0)] = E(V)N(d_1) - KN(d_2) \quad (65)$$

dove

$$d_1 = \frac{\ln \frac{E(V)}{K} + \frac{s^2}{2}}{s}$$
$$d_2 = \frac{\ln \frac{E(V)}{K} - \frac{s^2}{2}}{s}$$

ed  $E$  è il valore previsto.

**Prova risultato chiave:**  $g(V)$  è la funzione di densità di probabilità di  $V$ . Segue che

$$E[\max(V - K, 0)] = \int_K^\infty (V - K)g(V)dV \quad (66)$$

La variabile  $\ln V$  è normalmente distribuita con deviazione standard  $s$ . Dalle proprietà della distribuzione lognormale, la media di  $\ln V$  è  $m$ , dove

$$m = \ln E(V) - \frac{s^2}{2} \quad (67)$$

Definita una nuova variabile  $Q$ , dove

$$Q = \frac{\ln V - m}{s} \quad (68)$$

Questa variabile è normalmente distribuita con media zero e deviazione standard pari a 1. Viene chiamata  $h(Q)$  la funzione di densità di  $Q$  tale che

$$h(Q) = \frac{1}{\sqrt{2\pi}} e^{-\frac{Q^2}{2}}$$

Usando l'equazione (68) per convertire l'espressione del membro destro dell'equazione (66) da un integrale su  $V$  ad un integrale su  $Q$ , viene

$$E[\max(V - K, 0)] = \int_{\frac{\ln K - m}{s}}^\infty (e^{Qs+m} - K)h(Q)dQ$$

oppure

$$E[\max(V - K, 0)] = \int_{\frac{\ln K - m}{s}}^{\infty} e^{Qs+m} h(Q) dQ - K \int_{\frac{\ln K - m}{s}}^{\infty} h(Q) dQ \quad (69)$$

Adesso

$$\begin{aligned} e^{Qs+m} h(Q) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{Q^2 + 2Qs + 2m}{2}} \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{[(Q-s)^2 + 2m + s^2]}{2}} \\ &= \frac{e^{m + \frac{s^2}{2}}}{\sqrt{2\pi}} e^{-\frac{[(Q-s)^2]}{2}} \\ &= e^{m + \frac{s^2}{2}} h(Q - s) \end{aligned}$$

Questo significa che l'equazione (69) diventa

$$E[\max(V - K, 0)] = e^{m + \frac{s^2}{2}} \int_{\frac{\ln K - m}{s}}^{\infty} h(Q) dQ \quad (70)$$

Se venisse definita  $N(x)$  come la probabilità che una variabile con media zero e deviazione standard 1 sia più piccola di  $x$ , il primo integrale nell'equazione (70) sarebbe

$$1 - N\left[\frac{\ln K - m}{s} - s\right]$$

oppure

$$N\left[\frac{-\ln K + m}{s} + s\right]$$

Applicando la sostituzione dall'equazione (67) viene

$$N\left(\frac{\ln\left[\frac{E(V)}{K}\right] + \frac{s^2}{2}}{s}\right) = N(d_1)$$

In modo simile il secondo integrale nell'equazione (70) è  $N(d_2)$ . Per questa ragione l'equazione (70) diventa

$$E[\max(V - K, 0)] = e^{m + \frac{s^2}{2}} N(d_1) - KN(d_2)$$

Applicando la sostituzione dalla formula (67) si ottiene il risultato chiave.

**Il risultato di Black-Scholes:** si considera un'opzione call che non paga dividendi e che maturi al tempo  $T$ . Il prezzo di esercizio è  $K$ , il tasso d'interesse libero da rischi è  $r$ , il prezzo attuale dell'azione è  $S_0$  e la volatilità è  $\sigma$ . Come mostrato dall'equazione (41), il prezzo  $c$  di una call, è dato da

$$c = e^{-rT} \hat{E}[\max(S_T - K, 0)] \quad (71)$$

dove  $S_T$  è il prezzo dell'azione al tempo  $T$  ed  $\hat{E}$  denota ciò che ci aspettiamo in uno scenario risk-neutral. Sotto il processo stocastico di Black-Scholes,  $S_T$  è

lognormale. Anche dalle equazioni (22) e (23)  $\hat{E}(S_T) = S_0 e^{rT}$  e la deviazione standard di  $\ln S_T$  è  $\sigma\sqrt{T}$ . Dal risultato chiave appena provato l'equazione (71) implica che

$$c = e^{-rT} [S_0 e^{rT} N(d_1) - KN(d_2)]$$

oppure

$$c = S_0 N(d_1) - Ke^{-rT} N(d_2)$$

dove

$$\begin{aligned} d_1 &= \frac{\ln \left[ \frac{\hat{E}(S_T)}{K} \right] + \sigma^2 \frac{T}{2}}{\sigma\sqrt{T}} \\ &= \frac{\ln \frac{S_0}{K} + (r + \frac{\sigma^2}{2})T}{\sigma\sqrt{T}} \end{aligned}$$

e

$$\begin{aligned} d_2 &= \frac{\ln \left[ \frac{\hat{E}(S_T)}{K} \right] - \sigma^2 \frac{T}{2}}{\sigma\sqrt{T}} \\ &= \frac{\ln \frac{S_0}{K} + (r - \frac{\sigma^2}{2})T}{\sigma\sqrt{T}} \end{aligned}$$

## Riferimenti bibliografici

- [1] JOHN C. HULL. *Options, Futures & other Derivatives - Fifth Edition* - Prentice Hall, 2006.
- [2] P. WILMOTT, S. HOWISON, J. DEWYNNE. *The Mathematics Of Financial Derivatives* - Cambridge University Press, 1995.
- [3] T.W. ANDERSON, L.A. GOODMAN. *Statistical Inference About Markov Chains* - Ann. Math. Statist. 28, 89-110, 1957.
- [4] LEA V. CARTY, JEROME S. FONS. *Measuring Changes in Corporate Credit Quality* - Moody's Special Report, Novembre 1993.
- [5] A. BANGIA, F. DIEBOLD, A. KRONIMUS, C. SHAGEN, T. SCHUERMANN. *Ratings Migration and the Business Cycle, with Applications to Credit Portfolio Stress Testing* - Journal of Banking and Finance 26, 445-474.
- [6] D. BEN-AVRAHAM, S. HAVLIN. *Diffusion and Reactions in Fractal and Disordered System* - Cambridge University Press, 2000.
- [7] C. W. GARDINER. *Handbook of Stochastic Methods* - Springer-Verlag, 1997.
- [8] T. K. MOON, W. C. STIRLING. *Mathematical Methods and Algorithms for Signal Processing* - Prentice Hall, 2000.
- [9] J.-P. FOUQUE, G. PAPANICOLAOU, K. R. SIRCAR. *Derivatives in Financial Markets with Stochastic Volatility* - Cambridge University Press, 2000.
- [10] J.MASOLIVER, J. PERELLÓ. - Quantitative Finance 6, 423, 2006.