Università degli Studi di Roma "La Sapienza"
Anno Accademico 2008-2009

Facoltà di Scienze Matematiche Fisiche e Naturali
Master Calcolo Scientifico

Alcuni appunti per il corso di
FINANZA MATEMATICA

Giovanna Nappo

A.A. 2008/2009

versione del 5 giugno-2009

Indice

1 Cenni sul funzionamento e sulla storia dei mercati finanziari **1**
1.1 Mercati Finanziari . . . 2
Appendice: tasso di interesse a tempo continuo . . . 8
Appendice: Richiami sulle equazioni differenziali lineari . . . 9
1.2 Un esempio concreto di derivato finanziario . . . 11
1.3 Teorema dell'arbitraggio . . . 13
1.3.1 Applicazioni . . . 14
1.4 Il modello binomiale uniperiodale . . . 17
1.4.1 Modello binomiale uniperiodale con contingent claim . . . 21
1.5 Il modello binomiale multiperiodale . . . 28
1.6 Appendice: Alberi binomiali e modello CRR*** . . . 36
1.7 Appendice: Approccio soggettivista alla probabilità . . . 44

2 Il modello di Black e Scholes come limite del modello binomiale multiperiodale **47**
2.1 Il Modello Binomiale Multiperiodale . . . 47
2.1.1 Ipotesi e notazioni . . . 47
2.2 Approssimazione del Modello Binomiale Multiperiodale . . . 49
2.2.1 Il modello approssimato, a tempo continuo . . . 49
2.2.2 Dimostrazione della formula di Black e Scholes . . . 52
2.3 Il moto Browniano . . . 57
2.3.1 Approssimazione del moto browniano per t fissato . . . 57
2.3.2 Indipendenza ed omogeneità degli incrementi . . . 60
2.3.3 Definizione del moto browniano e del modello di Black e Scholes . . . 60

3 Processi aleatori a tempo continuo **62**
3.1 Processi aleatori, definizioni ed esempi . . . 62
3.2 Le traiettorie del processo di Wiener non sono a variazione limitata . . . 69
3.3 Appendice: variabili Gaussiane . . . 72

4 Processi a incrementi indipendenti e martingale **75**
4.1 Processi ad incrementi indipendenti ed omogenei . . . 75
4.2 Esempi di martingale a tempo continuo . . . 77
4.3 APPENDICE: Proprietà del valore atteso condizionale e qualche esempio . . . 83

5 Integrale stocastico: cenni **88**
5.1 Integrale stocastico per processi integrabili, a partire dai processi elementari . . . 88
5.1.1 Esempi di calcolo di integrali stocastici . . . 93
5.1.2 Integrale stocastico per processi localmente di quadrato integrabile . . . 95
5.1.3 Osservazioni su alcune proprietà degli integrali stocastici . . . 97
5.2 Calcolo stocastico e formula di Itô . . . 97
5.2.1 Moto browniano geometrico e il suo differenziale stocastico . . . 102
5.3 Equazioni differenziali stocastiche . . . 103
5.4 Il processo di Cox-Ingersoll-Ross . . . 106

6 Il modello di Black e Scholes **109**
6.1 Copertura perfetta e non arbitraggio . . . 111
6.2 L'equazione di Black e Scholes . . . 111

Bibliografia **116**

Capitolo 1

Cenni sul funzionamento e sulla storia dei mercati finanziari

Prima di tutto proviamo a definire di cosa si occupa quella che viene chiamata la teoria matematica della finanza (o matematica finanziaria)[1].

I quattro protagonisti sono

- **gli individui**, la cui attività viene descritta in termini del dilemma consumo-investimento: consumare di più ora o investire per ottenere di più dopo? L'ambivalenza del loro comportamento sia come consumatore che come investitore porta a problemi di ottimizzazione formulati in termini di economia matematica come problemi di consumo-risparmio e di decisione sulla composizione del portafoglio[2]. Nell'ambito della teoria dell'utilità il primo problema è trattato in base al postulato del comportamento razionale (Von Neumann-Morgestern) degli individui in stato di incertezza. Questa teoria porta a determinare strategie preferibili in termini di analisi qualitativa, ad esempio con il valore atteso delle funzioni di utilità. Il problema della composizione del portafoglio porta a problemi di miglior allocazione dei fondi, tenendo conto del rischio, tra i diversi beni possibili, quali proprietà, oro, titoli (o securities: buoni, azioni, opzioni, future...).
- **le corporazioni** (compagnie, ditte,...) che posseggono beni di valore (terra, fabbriche, macchinari, tecnologie,...) organizzano affari, mantengono relazioni commerciali. Per aumentare il loro capitale a volte le corporazioni emettono azioni (stocks) o buoni (bonds). I buoni sono emessi anche dai governi. Lo scopo delle corporazioni è quello di andare incontro agli interessi dei possessori delle azioni (shareholders) e dei buoni (bondholders).
- **gli intermediari** o meglio le strutture finanziare intermediatrici (banche, compagnie di investimenti, fondi pensioni, compagnie di assicurazioni...). Tra queste si possono mettere anche i mercati finanziari, che scambiano azioni opzioni, future, etc..., Tra i mercati finanziari, quelli degli Stati Uniti sono tra i più famosi:
  - NYSE: New York Stock Exchange
  - AMEX: American Stock Exchange
  - NASDAQ: NASDAQ Stock Exchange
  - NYFE: New York Future Exchange
  - CBOT: Chicago Board of Trade
- **i mercati finanziari** (di denaro, di metalli preziosi, di strumenti finanziari). In particolare nei mercati di strumenti finanziari, di solito si distingue tra
  - strumenti sottostanti o primari come
    - conti bancari
    - buoni

[1] Questa prima parte è basata principalmente sul testo di Shiryaev [17].

[2] Il termine portafoglio (portfolio) significa nei modelli base la suddivisione tra investimenti e risparmio, con eventuali altre restrizioni, come ad esempio limiti superiori od inferiori delle quantità investite. In modelli più complessi può riguardare anche le quantità utilizzate in consumo.

    * azioni
  - strumenti derivati o secondari
    * opzioni
    * contratti future
    * warrants
    * swaps
    * combinazioni
    * etc.

Notiamo che l'ingegneria finanziaria è spesso pensata come la manipolazione dei derivati per aumentare il capitale e ridurre il rischio causato dall'incertezza della situazione del mercato nel futuro.

1.1 Mercati Finanziari

1 **Denaro** Si tratta di un meccanismo che permette di commerciare cose/beni che si hanno in modo da ottenere poi cose/beni che si desiderano. Al momento attuale monete e banconote sono solo una piccola parte del denaro esistente. La maggior parte dei pagamenti viene effettuata tramite assegni o per via telematica (bancomat, carte di credito, web-banking, ...). Oltre alla funzione di mezzo di circolazione il denaro ha un ruolo importante anche come mezzo di valutazione e come mezzo di risparmio.

2 **Moneta, Cambio, Numerario** Le riserve di moneta di altri paesi, i tassi di cambio, etc. sono un'indicatore importante del benessere di una nazione, del suo sviluppo; inoltre è spesso un mezzo di pagamento per il commercio con l'estero.

Nell'ambito della storia dei cambi di valuta spesso si sono avuti accordi internazionali e unioni monetarie. Ad esempio a Bretton-Woods (New Hampshire, USA) nel 1944 si svolse una famosa conferenza durante la quale si decise il sistema di credito e di valuta del mondo occidentale, in particolare i tassi di cambio delle valute coinvolte potevano variare solo del 1% rispetto a quelli ufficiali. In quella occasione fu istituito la Fondazione Monetaria Internazionale (International Monetary Foundation, IMF). L'accordo rimase in vigore fino alla crisi petrolifera e alla crisi monetaria del 1973, che coinvolse il dollaro statunitense, il marco tedesco e lo yen giapponese. Nel 1979 furono poste le basi per l'Unione Monetaria Europea (il famoso Serpentone: venne stipulato un patto secondo il quale le variazioni dei tassi di cambio potevano variare in una fascia del $\pm 2.25\%$. Per ottenere questo risultato le banche centrali nazionali dovevano intervenire per assicurare la stabilità dei tassi di cambio. Successivamente si è arrivati alla moneta unica: a partire dalla fine del 2001, nei paesi dell'Unione Europa circola l'euro.

3 **Metalli Preziosi** Si tratta di oro, argento, platino e altri (iridio, palladio, osmio, rodio, rutenio). Hanno avuto un ruolo importante nel passato, specialmente nel $19-$*simo* secolo, ma hanno ancora un ruolo ai nostri giorni nel sistema del credito internazionale e del cambio di valute.

Un po' di storia: si può considerare che l'età dell'oro sia iniziata nel 1821, anno in cui il governo britannico proclamò la convertibilità in oro della sterlina. Poco dopo anche gli Stati Uniti fecero lo stesso con il dollaro (as good as gold). Lo standard dell'oro ebbe il suo apice tra il 1880 e il 1914, ma dopo la prima guerra mondiale non recuperò più il suo status. Le sue tracce si persero definitivamente quando Nixon nell'agosto del 1971 dichiarò formalmente la fine della convertibilità in oro del dollaro[3].

[3] In realtà dopo la crisi del 1929, e precisamente nel 1934, il governo degli USA dichiarò che un'oncia (28,35 grammi) d'oro valeva 35 dollari. Così rimase formalmente fino al 1971, anche se già da tempo era chiaro che i dollari in circolazione erano molti di più di quelli che si sarebbero potuti convertire in oro con le riserve di questo metallo prezioso in possesso degli USA. La dichiarazione di Nixon, che va ricordato anche come il presidente che gestì la fine della guerra in Vietnam, ebbe forti ripercussioni su tutta l'economia mondiale, e portò ad una svalutazione del dollaro e ad una conseguente impennata dei prezzi del petrolio, forse anche maggiore di quella che stiamo vivendo in questo periodo (autunno 2004). La svalutazione del dollaro comportò la svalutazione delle altre monete, in particolare della lira, legata al dollaro dopo che il piano Roosevelt aveva permesso all'Italia di riprendersi dopo la seconda guerra mondiale. Anche l'inflazione era impressionante: dell'ordine del 17% annuo. La crisi fu tale che furono inventate le domeniche a piedi: ma non per motivi ecologici (la parola *ecologia* non era entrata ancora nell'uso comune), bensì per cercare di risparmiare sul consumo del petrolio e di conseguenza di diminuire nel bilancio dello stato la voce dei pagamenti all'estero. La fine della convertibilità del dollaro ebbe come conseguenza una forte instabilità dei prezzi e fu uno dei motivi per cui nacque l'esigenza di avere delle coperture finanziarie contro la grande variabilità dei prezzi. Per dare ancora un'idea delle fluttuazioni si tenga presente che prima del 1971 il dollaro veniva scambiato a 600 lire, mentre nel giro di poco tempo (non so precisare al momento quanto tempo) il cambio si aggirava attorno al doppio. Comunque per essere più precisi si può considerare che l'oro passò dai 35 dollari per oncia del periodo $1934-1971$ al massimo di 570 dollari per oncia del 1980. Successivamente

4 Conto bancario Un conto bancario (bank account) è un titolo (o una security) dello stesso tipo dei buoni[4], in quanto si riduce all'obbligo da parte della banca di pagare certi interessi sulla somma che è stata messa sul conto. I conti in banca sono convenienti come misura dei prezzi di varie altre security. Si distinguono vari tipi di interesse

- **interesse semplice** a tasso r significa che se oggi ho la cifra x_0 dopo n anni avrò la stessa cifra più $n \cdot r \cdot x_0$, cioè $x_0\big(1 + n \cdot r\big)$. Sostanzialmente è quello che accade se ogni anno gli interessi vengono ritirati e non rimessi sul conto in banca.
- **interesse composto** Se invece gli interessi venissero messi sul conto si avrebbe la seguente tabella

0	1	2	$\cdots$	k	$\cdots$	n
x_0	$x_1 = x_0\big(1+r\big)$	$x_2 = x_1\big(1+r\big)$	$\cdots$	$x_k = x_{k-1}\big(1+r\big)$	$\cdots$	$x_n = x_{n-1}\big(1+r\big)$
x_0	$x_1 = x_0\big(1+r\big)$	$x_2 = x_0\big(1+r\big)^2$	$\cdots$	$x_k = x_0\big(1+r\big)^k$	$\cdots$	$x_n = x_0\big(1+r\big)^n$

  dove x_0 rappresenta il valore inizialmente ($t = 0$) depositato, (ovvero all'inizio del primo periodo cioè per $0 \leq t < 1$) x_1 rappresenta l'ammontare dalla fine del primo periodo all'inizio (escluso) del secondo periodo (cioè per $1 \leq t < 2$), mentre x_k rappresenta l'ammontare dalla fine del $k - simo$ periodo alla sua fine (cioè per $k \leq t < k + 1$).
- **interesse composto (m volte in un anno)** a tasso (nominale) r significa che gli interessi sono versati sul conto alla fine di ogni periodo di durata l'$m - sima$ parte di anno, ovvero alla fine del primo periodo si avrà $x_0(1 + r/m)$, alla fine del secondo periodo si avrà $x_0(1 + r/m)^2$, e alla fine dell'$h - simo$ periodo si avrà $x_0(1 + r/m)^h$. Se non vengono effettuati prelievi la quantità di denaro al tempo $t = N + k/m$, ovvero dopo N anni e k periodi di un $m - simo$ di anno, cioè dopo $N \cdot m + k$ periodi, si avrà sul conto la quantità

$$x_t^{(m)} = x_0(1 + r/m)^{N \cdot m+k} = x_0(1 + r/m)^{m(N+k/m)} = x_0(1 + r/m)^{m \cdot t} \tag{1.1}$$

  È interessante notare che alla fine dell'anno, cioè dopo m periodi, si ha a disposizione la cifra di $x_0(1+r/m)^m$. Si può quindi definire (e calcolare) il ***tasso semplice effettivo*** $r^*_{eff}(m)$ equivalente al tasso r composto su m periodi in un anno come quel valore $r^*_{eff}(m)$ tale che:

$$1 + r^*_{eff}(m) = \big(1 + r/m\big)^m \Leftrightarrow r^*_{eff}(m) = \big(1 + r/m\big)^m - 1$$

- **interesse composto a tempo continuo**[5] Nel caso in cui il numero di periodi per anno tende ad infinito, ovvero nel caso in cui gli interessi vengono pagati con scadenze così ravvicinate da poter essere pensate in

precipitò ai 308 dollari per oncia del 1984, per poi continuare ad oscillare tra i 300 e i 400 dollari.

Può essere interessante riportare i 10 punti con cui Nixon diede l'annuncio il 15 agosto 1971 (come riportato dai giornali italiani dell'epoca):

1 Sospensione temporanea della convertibilità in oro del dollaro, eccezion fatta per le operazioni che saranno di interesse per gli Stati Uniti.

2 Gli Stati Uniti chiederanno al Fondo Monetario Internazionale il varo di un nuovo sistema monetario internazionale e terranno in sospeso la convertibilità in oro fino a quando non si saranno trovati adeguati accordi.

3 Sarà introdotta una tassa temporanea del 10% su tutte le importazioni negli Stati Uniti.

4 Saranno congelati per tre mesi prezzi, stipendi, affitti e dividendi.

5 Sarà abrogata la sovrattassa del 7% sull'acquisto di vetture nuove nazionali o straniere.

6 Saranno anticipate al gennaio 1972 le riduzioni fiscali già previste per il gennaio 1973.

7 Sarà richiesto al Congresso di approvare un piano per l'estensione della mano d'opera, con la possibilità di riduzione delle tasse per coloro che seguiranno questa idea.

8 Ricerche e sviluppo tecnologico e industriale saranno stimolati e incoraggiati.

9 È previsto un risparmio di 4 miliardi e 700 milioni nelle spese federali, comprese alcune limitazioni negli aumenti degli stipendi degli impiegati.

10 È prevista una riduzione del 10% degli aiuti americani all'estero.

[4]Si veda più avanti una brevissima spiegazione sui buoni.

[5]In questo paragrafo consideriamo solo il caso in cui il tasso rimane costante per tutto il periodo al quale siamo interessati. Per il caso in cui il tasso varia nel tempo, in modo deterministico, si veda l'appendice a questa sezione.

tempo continuo appare naturale che al tempo t vada considerato il valore definito da

$$x_t := \lim_{m\to\infty} x_t^{(m)}.$$

Tenendo conto che[6] $x_t^{(m)} = x_{\lfloor m\cdot t\rfloor/m}^{(m)}$ dalla (1.1) si ottiene che

$$x_t = \lim_{m\to\infty} x_{\lfloor m\cdot t\rfloor/m}^{(m)} = \lim_{m\to\infty} \big(1 + r/m\big)^{m\cdot\left(\lfloor m\cdot t\rfloor/m\right)} = e^{rt}.$$

Anche in questo caso si possono definire dei tassi equivalenti: ad esempio, dato il tasso nominale a tempo continuo r_c esiste un tasso (nominale) di interesse annuale $r(m)(= r(m, r_c)$ composto in m periodi, equivalente al tasso nominale a tempo continuo r_c, ovvero tale che

$$x_1 = x_0 \cdot e^{r_c} = x_0\big(1 + r(m)/m\big)^m \Leftrightarrow r(m) = m\big(e^{r_c/m} - 1\big).$$

La formula inversa, cioè la formula che, dato il tasso di interesse $r(m)$ composto in m periodi, permette di ottenere il valore del tasso $r = r_c(r(m))$ a tempo continuo corrispondente è ovviamente

$$r = m\,\log\big(1 + r(m)/m\big).$$

Vale la pena di sottolineare il caso in cui $m = 1$, che corrisponde al tasso (effettivo) di interesse semplice $r(1) = \hat r$ (sempre su base annua), in cui le due formule diventano

$$\hat r = e^r - 1, \qquad r = \log\big(1 + \hat r\big),$$

Infine va ricordata anche la definizione di ***tasso di sconto*** $\hat q$ (su base annua), ovvero quella quantità $\hat q$ che permette di calcolare la somma B_0 che devo mettere in banca oggi, se voglio ottenere tra un anno la somma B_1, attraverso la formula

$$B_0 = B_1\big(1 - \hat q\big).$$

Tenendo conto che $B_1 = B_0\big(1 + \hat r\big)$ si ottiene che

$$\big(1 - \hat q\big)\,\big(1 + \hat r\big) = 1, \tag{1.2}$$

ovvero[7]

$$\hat q = \frac{\hat r}{1 + \hat r}, \qquad \hat r = \frac{\hat q}{1 - \hat q}$$

5 Buoni (DA RIVEDERE) I buoni (o bond) sono obbligazioni emesse da un governo, o da una corporazione, da una banca, o da un'altro ente finanziario per aumentare il proprio capitale. I buoni sono molto popolari in alcuni paesi, in particolare perché impegnano l'ente che li emette ad uno scadenzario prefissato in modo deterministico: il compratore paga inizialmente il prezzo (iniziale) del buono, e l'interesse gli viene pagato dall'ente emittente

[6]Infatti il valore della funzione $s \mapsto x_s^{(m)}$ è costante sugli intervalli di tempo del tipo $[k/m, (k+1)/m)$. Dato t si tratta di trovare $k = k(t)$ per il quale valga che $t \in [k/m, (k+1)/m)$, ovvero per il quale $k/m \le t < (k+1)/m$, chiaramente $k(t) = \lfloor m \cdot t\rfloor$, la parte intera inferiore di t. Per quel che segue è poi utile notare che

$$0 \le t - \lfloor m\cdot t\rfloor/m = \frac{m\cdot t - \lfloor m \cdot t\rfloor}{m} < 1/m$$

e che quindi

$$t = \lim_{m\to\infty} \lfloor m\cdot t\rfloor/m.$$

[7]È interessante notare che se il tasso di sconto viene aumentato di α allora il tasso di interesse corrispondente passa dal valore $\hat r(\hat q) = \frac{\hat q}{1-\hat q}$ al valore

$$\hat r(\hat q + \alpha) = \frac{\hat q + \alpha}{1 - \hat q - \alpha} = \hat r(\hat q) + \alpha\,\frac{d}{dq}\left(\frac{q}{1-q}\right)\bigg|_{q=\hat q} + o(\alpha) = \hat r(\hat q) + \alpha\,\frac{1}{(1-\hat q)^2} + o(\alpha).$$

Inoltre va sottolineato che la relazione (1.2) va pensata come una definizione del tasso di sconto. Inoltre in tutta la trattazione precedente si è considerato che il tasso di credito e quello di prestito sono gli stessi, ovvero non abbiamo considerato il segno di x_0. Ciò non è vero di solito nella realtà: la banca prevede un tasso di interesse se depositate delle somme di denaro (cioè se $x_0 > 0$), mentre prevede un tasso di interesse diverso (e più elevato) se siete creditori di somme di denaro nei confronti della banca (cioè se $x_0 > 0$). Comunque per semplicità di trattazione, di solito si ammette che i due tassi coincidano.

con scadenze regolari (in cedole), mentre il pagamento dell'intero prestito è garantito ad una scadenza prefissata (maturità). Esistono anche buoni senza cedole (zero coupon bond), tipicamente con maturità breve. Si considera quindi l'investimento in buoni un investimento senza rischio[8].

Per caratterizzare un buono servono delle caratteristiche numeriche (caso a tempo discreto, con cedole costanti, emesso al tempo $t = 0$):

valore facciale (face value) $P(T,T)$

maturità (maturity date) T

- **breve termine (short term)** da 3 mesi a 1 anno
- **medio termine (middle term)** da 2 a 10 anni
- **lungo termine (long term)** oltre 30 anni

tasso di interesse, o rendimento (nominale???) del buono (coupon yield) r_c che permette di calcolare il valore di ciascuna cedola[9]: $C_k = r_c \cdot P(T,T)$, per $k = 1, 2, ..., T$

prezzo iniziale (original price) $P(0,T)$, che è il prezzo pagato al tempo $t = 0$

valore di mercato(market value) $P(t,T)$, che è il valore del contratto al tempo $t \in (0,T)$ e che può variare in modo aleatorio (a causa di vari fattori economici: domanda/offerta e viene di solito modellato come un processo aleatorio)

rendimento corrente (current yield) $r_c(t,T) = \dfrac{r_c \cdot P(T,T)}{P(t,T)}$, che è il rapporto tra il valore di una cedola rispetto al valore di mercato del buono al tempo t, e che è importante per comparare i valori di buoni differenti.

rendimento alla maturità (yield to maturity), su base percentuale $\rho(T-t,T)$, che è definito[10] come la soluzione (unica) di

$$P(t,T) = \sum_{k=1}^{T-t} \frac{r_c \cdot P(T,T)}{(1+\rho)^k} + \frac{P(T,T)}{(1+\rho)^{T-t}}.$$

In altre parole $\rho(T-t,T)$ è definito come il tasso di rendimento interno del flusso dei pagamenti residui. Infine va sottolineato che poiché dato $P(t,T)$ si ricava $\rho(T-t,T)$ e viceversa[11], la descrizione aleatoria (la modellizzazione) del valore di mercato di un buono può essere fatta a partire dalla struttura temporale di $\rho(T-t,T)$ invece che di $P(t,T)$. La descrizione attraverso il valore di mercato P è detta diretta, mentre la descrizione attraverso ρ è detta indiretta.

6 **Azioni (Stock o Share)** Come già detto le azioni, come i buoni, sono emesse da compagnie per aumentare il capitale. Anche se esistono diversi tipi di azioni, i tipi principali sono due: azioni ordinarie (equity e common stock) e azioni preferenziali (preferred stock). Le differenze sono nel tipo di rischio e nel pagamento dei dividendi:

- chi possiede azioni ordinarie ottiene come dividendi la sua parte dei profitti della compagnia, e il loro ammontare dipende dal suo successo finanziario, mentre se la ditta fallisce perde tutto il suo investimento;

[8] Ovviamente c'è il rischio di insolvenza (o default risk), dovuto alla possibilità che l'ente che ha emesso il buono fallisca e non ottemperi l'impegno preso. Ovviamente i buoni emessi dai governi sono in genere meno esposti al rischio di credito rispetto ai buoni emessi da ditte, ma ovviamente ci sono controesempi clamorosi (si veda il caso dell'Argentina).

Questo tipo di rischio è di tipo diverso da quello dovuto alle fluttuazioni del mercato: si tratta di ***rischio di credito***, dovuto all'incertezza sulla solidità dell'ente emittente, e non di un rischio dovuto al fatto che il contratto stesso riguarda quantità aleatorie, come invece accade nel caso delle azioni o delle opzioni. Il rischio di credito è presente quando vengono comperate delle azioni: è possibile che la corporazione che le emette fallisca e, in questo caso, le azioni potrebbero perdere parzialmente o del tutto il loro valore (si vedano il caso Cirio e Parmalat).

[9] Il tasso r_c va inteso come tasso annuale se le cedole sono staccate alla fine di ogni anno, trimestrale, se vengono staccate alla fine di ogni trimestre, o mensile se vengono staccate alla fine di ogni mese, e così via

[10] È da ricordare che $T-t$ va inteso come tempo di vita residuo del buono, o tempo residuo alla maturità e da osservare che nel caso in cui fosse $P(t,T) = P(T,T)$ allora $\rho = r_c$ è l'unica soluzione di

$$1 = \sum_{k=1}^{T-t} \frac{r_c \cdot 1}{(1+\rho)^k} + \frac{1}{(1+\rho)^{T-t}}.$$

[11] Esistono buoni con interessi composti, trimestralmente, mensilmente o anche a tempo continuo, di cui viene dato il tasso nominale r_c annuo. Ovviamente di questo fatto va tenuto conto nel momento in cui si definisce la relazione che lega $P(t,T)$ e $\rho(T-t,T)$.

- chi possiede azioni preferenziali ha minor rischio di perdere tutto, i suoi dividendi sono garantiti, ma non aumentato con i profitti della compagnia.

Di solito però l'investitore, cioè colui che compra azioni, ma ciò vale anche per i bond, è attratto più che dai dividendi, dalla opportunità di fare soldi dalle fluttuazioni dei prezzi delle azioni, ovvero di comprare a un prezzo basso (prima degli altri) e vendere ad un prezzo alto (sempre prima degli altri[12])

7 Mercati dei derivati (DA RIVEDERE) Questi mercati ebbero grande espansione all'inizio degli anni settanta del ventesimo secolo. Fino al decennio precedente i cambi erano stati stabili e la volatilità del mercato era bassa. La situazione cambiò radicalmente[13] e questo causò l'interesse verso strumenti finanziari[14] che permettevano di coprirsi contro i rischi di inflazione, i cambi sfavorevoli e la grande volatilità dei mercati finanziari. Nel paragrafo 1.2 vedremo un esempio che illustra il tipo di problemi che possono essere affrontati con questo tipo di strumenti.

[12] Va tenuta presente la legge di mercato, che vale per ogni tipo di merce: se un prezzo è basso, la domanda (di comprare) aumenta, e se la domanda aumenta allora il prezzo sale, mentre, viceversa, se il prezzo è alto, l'offerta aumenta, e se l'offerta aumenta, allora il prezzo scende.

[13] Tra le motivazioni di questi cambiamenti vanno ricordate le seguenti:
i. il passaggio dal cambio fisso al cambio fluttuante, a seguito della crisi monetaria del 1973,
ii. la svalutazione del dollaro rispetto all'oro (1971), che invece dal 1934 era sempre stato scambiato a 35 dollari per oncia,
iii. la crisi mondiale del petrolio provocata dalla politica dell'OPEC, la comunità economica dei paesi produttori di petrolio, che divenne il principale *price maker* del petrolio,
iv. il declino dei mercati azionari (negli USA il declino era stato più forte che durante la crisi degli anni trenta, la Grande Depressione)

[14] I primi strumenti furono le opzioni (insieme anche ai futures) Tali titoli derivati erano presenti sui mercati non ufficiali (over the counter), ma il primo mercato ufficiale specializzato in opzioni fu il Chicago Board Option Exchange (CBOE) aperto il 26 aprile del 1973. Nel primo giorno di apertura furono trattati 911 contratti di opzioni, mentre appena un anno dopo il numero di opzioni scambiate giornalmente era di 20 000, tre anni più tardi di 100 000, e 700 000 nel 1987. Per capire le dimensioni del fenomeno va ricordato che ogni contratto riguardava 100 azioni e quindi gli scambi giornalieri riguardavano opzioni su 70 milioni di azioni, cioè poco più di un terzo dei 190 milioni di azioni scambiate giornalmente nel New York Stock Exchange (NYSE) nello stesso anno.

Il 1973 va ricordato anche per essere l'anno in cui furono pubblicati due articoli fondamentali, uno di Black e Scholes [6] e l'altro di Merton [13], che influenzarono notevolmente i metodi di prezzaggio.

**Valutazione di un flusso di cassa

Il tasso di interesse può essere utilizzato per attualizare i valori nominali di denaro e confrontare diversi flussi di cassa. Ci mettiamo per semplicità nel caso dell'interesse composto, con l'ipotesi che il tasso di interesse rimanga costante.

Ad esempio un contratto (come ad esempio nel caso dei bond, ma anche nel caso dei mutui stipulati con una banca) può prevedere il pagamento di a_0 al tempo iniziale $t=0$, e successivamente la riscossione di cifre nominali b_j al tempo $t=j$. Per poter confrontare cifre di denaro a tempi diversi, tali valori vanno attualizzati, (ad esempio al tempo $t=0$, cioè il valore attualizzato al tempo $t=0$ della cifra b_j al tempo $t=j$, è quella cifra $\beta_0^{(j)}$ (al tempo $t=0$) tale che il suo valore al tempo $t=j$, cioè $\beta_0^{(j)}(1+r)^j$, sia uguale a b_j, in sintesi

$$\beta_0^{(j)}(1+r)^j = b_j \qquad \Leftrightarrow \qquad \beta_0^{(j)} = \frac{b_j}{(1+r)^j}$$

Il valore attualizzato del flusso di cassa $(-a_0, b_1, \cdots, b_n)$ precedentemente descritto vale quindi

$$A(r) := -a_0 + \sum_{j=1^n} \beta_0^{(j)} = -a_0 + \sum_{j=1^n} \frac{b_j}{(1+r)^j}.$$

Per confrontare il flusso di cassa $(-a_0, b_1, \cdots, b_n)$ con un altro flusso di cassa $(-a_0', b_1', \cdots, b_n')$, dobbiamo quindi calcolare anche

$$A'(r) = -a_0' + \sum_{j=1^n} \frac{b_j'}{(1+r)^j},$$

e poi confrontare $A(r)$ e $A'(r)$. Il primo flusso di cassa sarà preferibile al secondo flusso se e solo se $A(r) > A'(r)$. Può accadere che il verificarsi di tale disuguaglianza dipenda dal valore di r, ossia che per alcuni valori di r il primo flusso sia preferibile al secondo e per altri valori di r accada il contrario.

Passiamo ora a parlare del rendimento interno, di un contratto che prevede un flusso di cassa $(-a_0, b_1, \cdots, b_n)$:
• se i segni di a_0 e di b_j sono tutti positivi,
e (come è naturale aspettarsi, altrimenti nessuno sottoscriverebbe un tale contratto)
• se vale la seguente disuguaglianza

$$\sum_{j=1}^n b_j > a_0,$$

allora possiamo anche definire il ***tasso di rendimento interno del flusso di cassa*** $(-a_0, b_1, \cdots, b_n)$, come quell'unico valore ρ^* tale che

$$A(\rho^*) = 0, \qquad \Leftrightarrow \qquad -a_0' + \sum_{j=1^n} \frac{b_j'}{(1+\rho^*)^j} = 0,$$

ossia quell'unico valore (costante) del tasso di interesse che rende uguali la cifra a_0 al tempo $t=0$ e il valore dei pagamenti b_j, $j=1,\ldots,n$, (ciascuno al tempo $t=j$).

L'esistenza e l'unicità di tale valore si possono ottenere facilmente: infatti, sotto le precedenti condizioni, è facile convincersi che il polinomio

$$P(x) := -a_0 + \sum_{j=1^n} b_j\, x^j$$

gode della proprietà che esiste uno e un solo $x^* \in (0,1)$ tale che $P(x^*) = 0$, e che quindi ρ^* è univocamente determinato da $x^* = \frac{1}{1+\rho^*}$, cioè $\rho^* = \frac{x^*-1}{x^*}$. E infatti

$$P(0) = -a_0 < 0, \qquad P(1) = -a_0 + \sum_{j=1^n} b_j > 0,$$

e quindi c'è almeno un $x^* \in (0,1)$ tale che $P(x^*) = 0$ e inoltre, tale x^* è unico, in quanto la funzione $P(x)$ è strettamente crescente in $(0,1)$, infatti

$$P'(x) = \sum_{j=1^n} j\, b_j\, x^{j-1} > 0, \quad \text{per } x \in (0,1).$$

Appendice: tasso di interesse a tempo continuo

Si supponga che $D(t)$ rappresenti il valore di un deposito in banca al tempo t. Il ***tasso di interesse istantaneo*** (o ***spot rate***) al tempo t sia denotato da $r(t)$: per definizione questo significa che nell'intervallo di tempo $[t, t+h]$, per ogni $h > 0$, ma abbastanza piccolo la cifra depositata passa dal valore $D(t)$ al valore

$$D(t+h) = D(t) + D(t) \cdot r(t) \cdot h = D(t)\big(1 + \cdot r(t) \cdot h\big), \tag{1.3}$$

ovvero[15] (a meno di infinitesimi) il tasso nell'intervallo considerato è proporzionale all'ampiezza h dell'intervallo considerato $(t, t+h)$, esattamente come nel caso a tasso costante, ma dipende dall'istante iniziale t. Come conseguenza della precedente uguaglianza si ottiene che (sempre a meno di infinitesimi)

$$D(t+h) - D(t) \approx D(t) \cdot r(t) \cdot h \quad \text{ovvero} \quad \frac{D(t+h) - D(t)}{h} = D(t) \cdot r(t).$$

Con un semplice passaggio al limite per h che tendo a zero[16] si ottiene che la condizione (1.3) equivale all'esistenza della derivata di $D(t)$ e al fatto che

$$\frac{d}{dt} D(t) = D(t) \cdot r(t)$$

equivale a dire che $D(t)$ è soluzione di una equazione differenziale (lineare omogenea a coefficienti non costanti) e più precisamente del problema di Cauchy

$$\begin{cases} \dot{x}(t) = \alpha(t)\, x(t), & \text{per } t \geq t_0 \\ x(t_0) = x_0 \end{cases} \tag{1.4}$$

con $\alpha(t) = r(t)$, e $x_0 = D(t_0)$.

Essendo la soluzione[17] del precedente problema (1.4) data da

$$x(t) = x_0 \exp\left\{ \int_{t_0}^{t} \alpha(s)\, ds \right\}, \tag{1.5}$$

si ottiene che, se $t_0 = 0$ e D_0 è il valore di $D(0)$,

$$D(t) = D_0 \exp\left\{ \int_0^t r(s)\, ds \right\}.$$

di conseguenza, ragionando come nel caso a tempo continuo[18], si ha che il valore attualizzato (al tempo $t = 0$) di una cifra D che verrà data al tempo t si può esprimere come

$$D \exp\left\{ - \int_0^t r(s)\, ds \right\}$$

È interessante notare che nell'ottenere l'equazione differenziale per $D(t)$ abbiamo ipotizzato implicitamente che i soldi depositati non fossero prelevati per costi o consumi, e che inoltre abbiamo ipotizzato che non venissero effettuati

[15] La formulazione esatta sarebbe

$$D(t+h) = D(t) + D(t) \cdot r(t) \cdot h + o(h) = D(t)\big(1 + \cdot r(t) \cdot h\big) + o(h).$$

[16] Si tenga presente che in realtà il limite si sta facendo per $h \to 0^+$ e quindi si ottiene solo l'esistenza della derivata destra.

[17] Il fatto che (1.5) sia soluzione del problema di Cauchy (1.4), si può verificare per calcolo diretto.

[18] Il valore attualizzato (al tempo 0) di una cifra D data al tempo t è definito come quella cifra $\bar{D} = \bar{D}(t)$ tale che

$$D = \bar{D} \exp\left\{ \int_0^t r(s)\, ds \right\},$$

da cui immediatamente

$$\bar{D}(t) = D \exp\left\{ - \int_0^t r(s)\, ds \right\}.$$

neanche ulteriori depositi dovuti a entrate (o income) di alcun tipo. Se invece prelievi e ulteriori depositi fossero ammessi, ovviamente bisognerebbe tenerne conto. E in tale caso, se $I(t)$ rappresenta il totale delle entrate fino al tempo t, e $C(t)$ il totale dei consumi effettuati fino al tempo t, allora si ha che, sempre nell'intervallo $[t, t+h]$ le entrate totali sono $I(t+h) - I(t)$, mentre le uscite sono $C(t+h) - C(t)$. Se le funzioni $I(t)$ e $C(t)$ sono derivabili con derivata $i(t)$ e $c(t)$ rispettivamente, allora[19]

$$\begin{aligned} D(t+h) =& D(t) + D(t) \cdot r(t) \cdot h + I(t+h) - I(t) - \big(C(t+h) - C(t)\big) \\ \approx& D(t) + D(t) \cdot r(t) \cdot h + i(t) \cdot h - c(t) \cdot h. \end{aligned}$$

Procedendo come nel caso precedente si ottiene che in questo caso $D(t)$ soddisfa un'altra equazione differenziale (lineare non omogenea a coefficienti non costanti)

$$\frac{d}{dt} D(t) = D(t) \cdot r(t) + i(t) - c(t),$$

e più precisamente $D(t)$ è soluzione del problema di Cauchy

$$\begin{cases} \dot{x}(t) = \alpha(t)\, x(t) + \beta(t), & \text{per } t \geq t_0 \\ x(t_0) = x_0 \end{cases} \tag{1.6}$$

con $\alpha(t) = r(t)$, $\beta(t) = i(t) - c(t)$, e $x_0 = D(t_0)$.
La soluzione del precedente problema (1.6) vale[20]

$$x(t) = \exp\left\{ \int_{t_0}^{t} \alpha(s)\, ds \right\} \left(x_0 + \int_{t_0}^{t} \beta(s)\, \exp\left\{ - \textstyle\int_s^t \alpha(u)\, du \right\} ds \right) \tag{1.7}$$

Appendice: Richiami sulle equazioni differenziali lineari

In questo paragrafo, ad uso degli studenti che non avessero ancora superato un'esame di equazioni differenziali, ricordiamo il metodo per ottenere la soluzione dei problemi di Cauchy (1.4) e (1.6). Infatti pur essendo facile verificare che le soluzioni (1.5) e (1.7) date precedentemente sono soluzioni dei rispettivi problemi (1.4) e (1.6), è interessante sapere come si arriva a tali soluzioni.

1 Problema omogeneo Il problema

$$\begin{cases} \dot{x}(t) = \alpha(t)\, x(t), & \text{per } t \geq t_0 \\ x(t_0) = x_0 \end{cases}$$

è equivalente a

$$\begin{cases} \dfrac{dx(t)}{x(t)} = \alpha(t)\, dt, & \text{per } t \geq t_0 \\ x(t_0) = x_0. \end{cases}$$

[19]Nell'equazione per il calcolo di $D(t+h)$ abbiamo trascurato l'apporto degli interessi maturati nell'intervallo $[t, t+h]$ per via delle ulteriori entrate e quelli dovuti alla banca per via delle spese effettuate. Del resto nelle ipotesi di regolarità per le funzioni $I(t)$ e $C(t)$, si ha che

$$I(t+h) - I(t) \approx i(t) \cdot h \quad \text{e} \quad C(t+h) - C(t) \approx c(t) \cdot h,$$

da cui le corrispondenti somme di denaro a credito e a debito dovuti agli interessi maturati sono rispettivamente

$$\big(I(t+h) - I(t)\big) \cdot r(t) \cdot h \approx \big(i(t) \cdot h \cdot r(t) \cdot h \quad \text{e} \quad \big(C(t+h) - C(t) \cdot r(t) \cdot h \approx \big(c(t) \cdot h \cdot r(t) \cdot h.$$

Si tratta quindi di infinitesimi dell'ordine di h^2, e quindi sono trascurabili nei passaggi successivi.

[20]Il fatto che (1.7) sia soluzione del problema di Cauchy (1.6), si può verificare per calcolo diretto.

Integrando tra t_0 e t, si ottiene

$$\begin{aligned}
\int_{t_0}^{t} \frac{dx(s)}{x(s)} &= \int_{t_0}^{t} \alpha(s)\,ds \\
&\Updownarrow \\
\int_{t_0}^{t} d\log\big(x(s)\big) = \log\big(x(t)\big) - \log\big(x(t_0)\big) &= \int_{t_0}^{t} \alpha(s)\,ds \\
\text{(tenendo conto che } x(t_0) = x_0\text{)} \qquad &\Updownarrow \\
\log\big(x(t)\big) - \log\big(x_0\big) = \log\left(\frac{x(t)}{x_0}\right) &= \int_{t_0}^{t} \alpha(s)\,ds \\
&\Updownarrow \\
\frac{x(t)}{x_0} = \exp\left\{\log\left(\tfrac{x(t)}{x_0}\right)\right\} &= \exp\left\{\int_{t_0}^{t} \alpha(s)\,ds\right\}
\end{aligned}$$

da cui

$$x(t) = x_0 \exp\left\{\int_{t_0}^{t} \alpha(s)\,ds\right\}$$

2 Problema non omogeneo - Metodo della variazione della costante

L'idea consiste nel cercare una soluzione del problema

$$\begin{cases} \dot{x}(t) = \alpha(t)\,x(t) + \beta(t), & \text{per } t \geq t_0 \\ x(t_0) = x_0 \end{cases}$$

che si possa esprimere come

$$x(t) = C(t)\,x_o(t)$$

dove

$$x_o(t) := \exp\left\{\int_{t_0}^{t} \alpha(s)\,ds\right\}$$

e $C(t)$ è una funzione da determinare.

Si noti che $x_o(t)$ è soluzione dell'equazione differenziale omogenea, infatti

$$x_o(t) := \exp\left\{\int_{t_0}^{t} \alpha(s)\,ds\right\} \Rightarrow \frac{d}{dt}x_o(t) = \alpha(t)\,x_o(t).$$

Se $x(t) = C(t)\,x_o(t)$ allora, da una parte

$$\begin{aligned}
\frac{d}{dt}x(t) &= \left(\frac{d}{dt}C(t)\right)\,x_o(t) + C(t)\,\left(\frac{d}{dt}x_o(t)\right) \\
&= \left(\frac{d}{dt}C(t)\right)\,x_o(t) + C(t)\,\alpha(t)\,x_o(t) \\
&= \left(\frac{d}{dt}C(t)\right)\,x_o(t) + \alpha(t)\,x(t),
\end{aligned}$$

mentre dall'altra

$$\frac{d}{dt}x(t) = \beta(t) + \alpha(t)\,x(t).$$

Confrontando le ultime due espressioni si ottiene che, affinché la funzione $x(t) = C(t)\,x_o(t)$ sia soluzione dell'equazione lineare non omogenea, è sufficiente che

$$\left(\frac{d}{dt}C(t)\right)\,x_o(t) = \beta(t), \quad \Leftrightarrow \quad \frac{d}{dt}C(t) = \frac{\beta(t)}{x_o(t)} = \beta(t)\,\exp\left\{-\int_{t_0}^{t} \alpha(s)\,ds\right\}$$

e che, affinché $x(t) = C(t)\, x_o(t)$ soddisfi la condizione iniziale,

$$C(t_0)\, x_o(t_0) = x_0 \quad \Leftrightarrow C(t_0) = x_0.$$

Di conseguenza, integrando tra t_0 e t si ottiene

$$C(t) - C(t_0) = \int_{t_0}^{t} \beta(s) \exp\left\{-\int_{t_0}^{s} \alpha(u)\, du\right\} \quad \Leftrightarrow \quad C(t) = x_0 + \int_{t_0}^{t} \beta(s) \exp\left\{-\int_{t_0}^{s} \alpha(u)\, du\right\},$$

e da cui infine si ottiene che

$$x(t) = C(t)\, x_o(t) = \left(x_0 + \int_{t_0}^{t} \beta(s) \exp\left\{-\int_{t_0}^{s} \alpha(u)\, du\right\}\right) \exp\left\{\int_{t_0}^{t} \alpha(u)\, du\right\},$$

ovvero

$$x(t) = \exp\left\{\int_{t_0}^{s} \alpha(u)\, du\right\} x_0 + \int_{t_0}^{t} \beta(s) \exp\left\{-\int_{s}^{t} \alpha(u)\, du\right\} ds.$$

Infine rimane da sottolineare che si può parlare della soluzione in quanto valgono le condizioni di unicità. Si rimanda ai corsi di Analisi Matematica sulle equazioni differenziali per i dettagli.

1.2 Un esempio concreto di derivato finanziario

Si consideri una compagnia italiana, la GPR, che oggi (denotato come tempo $t = 0$) ha firmato un contratto con la controparte americana BCG. Il contratto stipulato prevede che esattamente fra sei mesi (denotato come tempo $t = T$) la BCG invii alla ditta italiana 1000 computer, il contratto prevede anche che la ditta italiana GPR paghi 1000 dollari USA per ogni computer. Diamo anche come informazione[21] che il cambio euro/dollaro è $0,80$ euro per un dollaro[22].

Il problema di questo contratto, dal punto di vista della ditta italiana, sta nel rischio dovuto al cambio: nonostante si sappia che deve pagare $1\,000\,000\,\$$, non si sa quale sarà il cambio tra sei mesi, di conseguenza la GPR non sa esattamente quanto dovrà pagare in euro per i computer. Ad esempio se dovesse pagare tutto subito, cioè al tempo $t = 0$, allora dovrebbe pagare

$$1000 \times 1000\,\$ \times 0,80\, euro/\$ = 800\,000\, euro.$$

Se invece tra sei mesi, cioè al tempo $t = T$, il cambio fosse, sempre ad esempio, di $0,85$ euro per un dollaro, allora la ditta GPR dovrebbe pagare

$$1000 \times 1000\,\$ \times 0,85\, euro/\$ = 850\,000\, euro.$$

Per questo motivo la GPR deve affrontare il problema di come *coprirsi* contro il rischio dovuto al cambio. Di seguito diamo alcune strategie naturali:

1 Una strategia naïve potrebbe essere quella di comprare oggi i $1\,000\,000\,\$$ con $800\,000\, euro$, e tenere questi dollari bloccati. Il vantaggio di questo procedimento sta nel fatto che elimina completamente il rischio del cambio, ma c'è qualche grave controindicazione: prima di tutto si blocca una ingente somma di denaro per un periodo abbastanza lungo, in secondo luogo potrebbe anche darsi il caso che al tempo $t = 0$, la ditta italiana non possegga affatto il denaro necessario per effettuare questo tipo di operazione.

2 Una soluzione più sofisticata, e che non richiede alcun esborso di denaro al tempo $t = 0$, consiste nel fatto che la GPR vada sul mercato dei contratti forward per $1\,000\,000\,\$$, con scadenza a sei mesi. Un tale contratto può essere stipulato con una banca commerciale e prevede due punti

- La banca fornirà alla ditta GPR $1\,000\,000\,\$$, al tempo $t = T$.
- La ditta GPR pagherà, al tempo $t = T$, al tasso di cambio $K\, euro/\$$

[21] Nel mese di ottobre 2004, data in cui scriviamo, il cambio è solo approssimativamente giusto, i dati sono stati modificati per ottenere numeri più semplici.

[22] Questo esempio è tratto dall'introduzione del libro di Björk [5].

Il tasso di cambio K viene detto ***prezzo in avanti (forward)***, o anche tasso di cambio forward, al tempo $t = 0$ con tempo di consegna $t = T$. Nei contratti forward non ci sono costi di stipulazione[23], ed il tasso K deve essere determinato dalla domanda e dall'offerta del mercato forward.

Assumiamo che $K = 0,81$, allora la GPR sa che tra sei mesi dovrà dare alla banca 810 000 *euro*. Ed in questo modo il rischio dovuto al cambio è completamente eliminato.

Tuttavia ci sono anche qui alcune controindicazioni, dovute al fatto che un contratto forward è un contratto che **deve** essere onorato:

- Supponiamo che il tasso di cambio al tempo $t = T$ sia di $0,82\,euro/\$$.

  In questo caso la ditta risparmia la differenza tra 820 000 *euro* che avrebbe pagato senza il contratto e gli 810 000 *euro* che paga, cioè risparmia 10 000 *euro*.
- Supponiamo che il tasso di cambio al tempo $t = T$ sia di $0,79\,euro/\$$.

  Essendo la ditta costretta ad onorare il contratto, deve pagare il milione di dollari sempre 810 000 *euro*, mentre se non avesse dovuto onorare il contratto avrebbe pagato solo 790 000 *euro*, con una perdita di 20 000 *euro*

3 A questo punto è chiaro che l'ideale sarebbe un contratto che *coprisse contro eventuali tassi di cambio alti* al tempo $t = T$, ma permettesse di usufruire del vantaggio di un eventuale tasso di cambio basso. Questo tipo di contratto esiste e si chiama ***opzione call europea***, e prevede **la possibilità**, ***ma non l'obbligo***, di comprare al tempo $t = T$ (***tempo di esercizio***) un dollaro (o comunque un titolo sottostante) al prezzo di K *euro* (***prezzo di esercizio***)

Ovviamente, mentre il contratto forward non prevede costi iniziali, un'opzione deve prevedere un ***prezzo iniziale*** o ***premio*** (altrimenti non si troverebbe nessuna banca disponibile a vendere l'opzione)[24].

Un problema importante è proprio quello di come determinare tale prezzo (o problema del ***prezzaggio-pricing***), sotto opportune ipotesi sul mercato e sulla base del valore del cambio attuale. Questo sarà la motivazione principale del corso, ovviamente in contesti più generali.

Una prima risposta potrebbe essere la seguente: siamo in condizioni di incertezza, sono possibili diverse situazioni (o scenari) e per ciascuna situazione possiamo valutare la probabilità che si verifichi. Per valutare il prezzo si potrebbe calcolare il valore atteso del futuro guadagno stocastico[25]. Un problema che nasce immediatamente risiede nel fatto che i fenomeni economici e finanziari, pur essendo aleatori, non sono ripetibili, e quindi non esiste immediatamente una definizione oggettiva di probabilità, come nel caso dei fenomeni ripetibili, in cui è ragionevole prendere la frequenza come probabilità.

Un altro problema importante, che si pone chi vende l'opzione call europea è il seguente: il venditore si è impegnato a fornire un bene ad un prezzo prefissato, e ciò comporta un rischio finanziario; come fare a proteggersi (o coprirsi) dal rischio finanziario che corre all'istante $t = T$? Si tratta del problema della ***copertura-hedging***. E anche questo è un problema interessante da affrontare.

Introduciamo ora un po' di notazioni e di gergo economico.

Il termine ***opzione*** si usa tutte le volte in cui si tratta di avere la possibilità, ma non l'obbligo di comprare o vendere un titolo (o security). Il titolo da comprare o da vendere viene detto ***titolo sottostante o primario***, mentre l'opzione viene detto ***titolo derivato o secondario***. Il termine ***call*** si usa quando si tratta dell'opportunità di comprare qualcosa, se invece si tratta dell'opportunità di vendere qualcosa si parla di opzione ***put***. Il prezzo K concordato per l'acquisto o la vendita del bene viene detto ***prezzo di esercizio*** (dell'opzione) o ***prezzo di strike***.

[23] Va detto che comunque un contratto forward può richiedere delle spese durante l'intervallo $[0, T]$, ma non entriamo qui nella descrizione di questo tipo di contratti.

[24] Inoltre se non ci fosse un costo iniziale sarebbe possibile, con un capitale iniziale nullo e senza rischi, ottenere un guadagno non negativo e che può risultare addirittura strettamente positivo. La situazione vantaggiosa appena descritta è detta in termine tecnico arbitraggio. Daremo comunque una definizione formale di arbitraggio nel seguito di queste note.

[25] Si tratta di seguire la definizione soggettivista di probabilità, o meglio di valore atteso $\mathbb{E}(X)$ di una variabile aleatoria X, come quel prezzo c (certo) che si è disposti a pagare per accettare una scommessa e in cui reputiamo equivalente prendere il ruolo di scommettitore o di broker. A questo proposito si veda il paragrafo dedicato all'impostazione soggettiva delle probabilità, alla fine del capitolo con i richiami di probabilità.

L'istante finale $t = T$ viene detto ***tempo di esercizio*** (dell'opzione) o ***tempo di strike***. L'aggettivo ***europea*** si riferisce al fatto che l'opzione può essere esercitata solo alla fine del periodo $[0, T]$, si parlerebbe invece di opzione ***americana*** nel caso in cui l'opzione fosse esercitabile in un istante qualunque dell'intervallo di tempo $[0, T]$.

Il compratore del bene viene detto ***holder***, mentre il venditore ***writer***, inoltre si dice che l'holder assume una ***posizione lunga*** (***long position***), mentre il writer assume la ***posizione corta*** (***short position***).

Più in generale si considerano quelli che sono detti ***pagamenti a scadenza*** (o ***terminal pay-off***) che corrispondono all'obbligo di corrispondere una quantità aleatoria $\boldsymbol{f_T}$, che dipende dal valore del bene sottostante al tempo T (anche detti ***plain vanilla***), oppure dai valori che il bene sottostante assume durante l'intervallo di tempo $[0, T]$. Per questo motivo tali tipi di obbligazioni sono detti anche ***derivati*** (o in inglese anche ***contingent claim***, cioè affermazioni che dipendono dal valore contingente che assume appunto il sottostante).

1.3 Teorema dell'arbitraggio

Supponiamo di avere un mercato in cui si possano verificare solo un numero finito di casi possibili (detti anche ***scenari***). In altre parole supponiamo che l'evento certo sia finito: $\Omega = \{\omega_1, \omega_2, \omega_3, ..., \omega_m\}$. Supponiamo inoltre che si possano fare $d+1$ scommesse. Le scommesse sono caratterizzate[26] da $d+1$ variabili aleatorie Δ_0, Δ_1, Δ_2, ..., Δ_d nel seguente modo: se scommettiamo la quantità x sulla scommessa i, con $i = 0, 1, 2, ..., d$, e si verifica ω_j, allora ricaveremo[27] la somma $x\, \Delta_i(\omega_j)$.

Una ***strategia di scommessa*** è un vettore $\mathbf{x} = (x_0, x_1, ..., x_d)$, con x_i che indica la quantità (positiva o negativa) relativa alla scommessa $i - esima$.

Supponiamo che le strategie di scommessa $\mathbf{x} = (x_0, x_1, ..., x_d)$, con x_i scommesse, godano della ***proprietà di linearità***, ossia la strategia di scommessa $\mathbf{x}$ produce un guadagno[28] pari a

$$\sum_{i=0}^{d} x_i\, \Delta_i(\omega_j).$$

Osservazione 1.1. *In molti esempi la scommessa Δ_0 corrisponde al titolo così detto non rischioso, come ad esempio il deposito in banca o un titolo di tipo bond, e che viene poi usato come unità di misura per tutti gli altri beni. In alcuni casi si distingue rispetto a scommesse fatte in tempi diversi, e quindi per poter confrontare guadagni ottenuti in tempi diversi, tali guadagni vanno attualizzati, in molte applicazioni la scommessa Δ_0 vale identicamente 0 (si veda la parte relativa agli esempi)*

Come definizione di arbitraggio si utilizza la seguente:

Definizione 1.1 (arbitraggio forte)**.** *Data una strategia $\mathbf{x}$, si dice che quest'ultima permette un* **arbitraggio forte** *se a partire da un capitale iniziale nullo (cioè se $\sum_{i=0}^{d} x_i = 0$) si può ottenere un guadagno (sempre) strettamente positivo (ovvero $\sum_{i=0}^{d} x_i\, \Delta_i(\omega_j) > 0$) per ogni $j = 1, 2, ..., m$.*

[26] Si noti che se si indica con $r_i(j)$ il valore $\Delta_i(\omega_j)$, per $i = 0, 1, 2, ..., d$, e $j = 1, 2, ..., m$, si ottengono m vettori (colonna) $\mathbf{r}(j)$ $d+1$-dimensionali

$$\mathbf{r}(1) = \begin{pmatrix} r_0(1) \\ r_1(1) \\ \vdots \\ r_d(1) \end{pmatrix} \qquad \mathbf{r}(2) = \begin{pmatrix} r_0(2) \\ r_1(2) \\ \vdots \\ r_d(2) \end{pmatrix} \quad \cdots \quad \mathbf{r}(m) = \begin{pmatrix} r_0(m) \\ r_1(m) \\ \vdots \\ r_d(m) \end{pmatrix},$$

o equivalentemente si ottengono $d+1$ vettori (riga)

$$\mathbf{r}_0 = \begin{pmatrix} r_0(1) & r_0(2) & \cdots & r_0(m) \end{pmatrix} \qquad \mathbf{r}_1 = \begin{pmatrix} r_1(1) & r_1(2) & \cdots & r_1(m) \end{pmatrix} \quad \cdots \quad \mathbf{r}_d = \begin{pmatrix} r_d(1) & r_d(2) & \cdots & r_d(m) \end{pmatrix}.$$

[27] In Δ_j vanno considerate incorporate sia la somma ottenuta alla fine del gioco che l'eventuale somma da intascare alla fine del gioco.

[28] Con la notazione della nota precedente il guadagno è pari a la prodotto scalare

$$\mathbf{x} \cdot \mathbf{r}(j) = \sum_{i=0}^{d} x_i\, r_i(j).$$

Tuttavia va detto che c'è un'altra definizione di arbitraggio, che poi utilizzeremo in seguito, con una richiesta leggermente più debole:

Definizione 1.2 (arbitraggio debole)**.** *Data una strategia* $\mathbf{x}$*, si dice che quest'ultima permette un* **arbitraggio** *se a partire da un capitale iniziale nullo (cioè se* $\sum_{i=0}^{d} x_i = 0$*) si può ottenere un guadagno non negativo (ovvero* $\sum_{i=0}^{d} x_i\, \Delta_i(\omega_j) \geq 0$*) per ogni* $j = 1, 2, ..., m$*, ma non identicamente nullo (ovvero con almeno un* $j \in \{1, 2, ..., m\}$ *per il quale* $\sum_{i=0}^{d} x_i\, \Delta_i(\omega_j) > 0$*).*

A questo punto possiamo enunciare il risultato principale di questo capitolo.

Teorema 1.1 (Teorema dell'arbitraggio)**.** *In un mercato con le condizioni precedenti può verificarsi una delle seguenti alternative*[29]*:*
(a) *esiste una probabilità* $\tilde{\mathbb{P}}$ *su* Ω*, con* $\tilde{p}_j := \tilde{\mathbb{P}}(\omega_j) > 0$*, per ogni* $j = 1, 2, ..., m$*, ovvero caratterizzata dal vettore di probabilità* $\tilde{\mathbf{p}} := (\tilde{p}_1, \tilde{p}_2, ..., \tilde{p}_m)$*, con* $\tilde{p}_j > 0$*, per ogni* $j = 1, 2, ..., m$*, e tale che*

$$\widetilde{\mathbb{E}}(\Delta_i) = 0, \quad \textit{per ogni } i = 0, 1, 2, ..., d,$$

dove $\widetilde{\mathbb{E}}$ *indica il valore atteso rispetto alla probabilità* $\tilde{\mathbb{P}}$*, ovvero*

$$\widetilde{\mathbb{E}}(\Delta_i) = \sum_{j=1}^{m} \tilde{p}_j \Delta_i(\omega_j) = 0 \quad \textit{per ogni } i = 0, 1, 2, ..., d;$$

(b) *esiste una strategia* $\mathbf{x} = (x_0, x_1, ..., x_d)$ *tale che*

$$\sum_{i=0}^{d} x_i\, \Delta_i(\omega_j) > 0 \quad \textit{per ogni } j = 1, 2, ..., m;$$

in altre parole ci sono opportunità di arbitraggio (forte).

Osservazione 1.2. *Una (misura di) probabilità* $\tilde{\mathbb{P}}$ *che soddisfi le condizioni dell'alternativa* ***(a)*** *è detta nell'ambito finanziario* **misura martingala equivalente***. Il termine* **misura** *si riferisce al fatto che le probabilità sono dette anche misure di probabilità. Il termine* **equivalente** *corrisponde al fatto che prevedono che ogni scenario/evento elementare* ω_j *abbia probabilità positiva, infine il termine* **martingala** *si riferisce al fatto che le scommesse danno luogo a guadagni nulli, ed in un certo senso, rispetto alla probabilità* $\tilde{\mathbb{P}}$ *si tratta di scommesse eque. Una buona parte del seguito di questi appunti è dedicata alla precisazione del concetto di martingala, che verrà ripreso in seguito.*

Non diamo in queste note la dimostrazione del teorema dell'arbitraggio, perché nel seguito vedremo la dimostrazione probabilistica nel caso del modello binomiale multiperiodale. Segnaliamo tuttavia che nelle note di Baldi e Caramellino [3] se ne può trovare la dimostrazione analitico/geometrica.
Nella prossima sezione illustriamo l'uso del teorema dell'arbitraggio in un esempio relativo alle scommesse dei cavalli. Gli esempi relativi al modello di mercato più semplice si trovano nelle sezioni successive.

1.3.1 Applicazioni

Esempio 1.1. *Come primo esempio di applicazione vediamo il caso delle scommesse sui cavalli*[30]*. In una gara ippica con n cavalli ci sono ovviamente n scommesse, ciascuna caratterizzata dal dare vincente un cavallo diverso. Assumiamo*

[29] Sempre con le notazioni delle note precedenti le due alternative si possono scrivere come:
(a) esiste un vettore di probabilità $\tilde{\mathbf{p}} = (\tilde{p}_1, \tilde{p}_2, ..., \tilde{p}_m)$, con $\tilde{p}_j > 0$, per ogni $j = 1, 2, ..., m$, e tale che

$$\tilde{\mathbf{p}} \cdot \mathbf{r}_i' = \sum_{j=1}^{m} \tilde{p}_j\, r_i(j) = 0 \quad \text{per ogni } i = 0, 1, 2, ..., d;$$

(b) esiste una strategia $\mathbf{x} = (x_0, x_1, ..., x_d)$ tale che

$$\mathbf{x} \cdot \mathbf{r}(j) = \sum_{i=0}^{d} x_i\, r_i(j) > 0 \quad \text{per ogni } j = 1, 2, ..., m.$$

[30] Questo esempio è tratto dal libro di Ross [14]. Si veda anche l'esempio I.5.1 del libro di Dall'Aglio [9].

per semplicità che questa sia la sola tipologia di scommesse ammissibili. Ovviamente si ha $\Omega = \{\omega_1, \omega_2, ..., \omega_n\}$ *dove* ω_j *corrisponde alla vincita da parte del cavallo* j*, e in questo caso si ha anche che il numero delle scommesse vale* $n(= d+1)$*. Le scommesse vengono date in termini dei così detti* **odds***, ovvero si dice che sono date* 1 *a* o_i *intendendo che se si scommette sul cavallo* i *si paga immediatamente la cifra* 1*, mentre si riceve la cifra* $o_i + 1$ *se solo se vince il cavallo corrispondente. Per comodità di notazioni conviene assumere che le scommesse siano indicizzate da* 1 *ad* n *invece che da* 0 *ad* $n-1$*, in modo che nella scommessa* Δ_i *si riceva*[31] $o_i + 1$ *se e solo si verifica il caso* ω_i*. Con questa convenzione le scommesse si possano esprimere come*

$$\Delta_i(\omega_j) = \begin{cases} o_i(= o_i + 1 - 1) & \textit{se } j = i \\ -1 & \textit{se } j \neq i. \end{cases}$$

La condizione che esista una probabilità $\tilde{\mathbb{P}}$ *per la quale valga* $\tilde{\mathbb{E}}(\Delta_i) = 0$ *per ogni* $i = 1, 2, ..., n$ *diviene*

$$\begin{cases} \tilde{\mathbb{E}}(\Delta_i) = o_i\, \tilde{p}_i + (-1)\,(1 - \tilde{p}_i) = 0 & \textit{per ogni } i = 1, 2, ..., n; \\ \displaystyle\sum_{j=1}^{n} \tilde{p}_j = 1, & \textit{con } \tilde{p}_j > 0,\ \forall j = 1, 2, ..., n; \end{cases}$$

che equivale a

$$\begin{cases} \tilde{p}_i = \dfrac{1}{1 + o_i}; & \textit{per ogni } i = 1, 2, ..., n; \\ \displaystyle\sum_{j=1}^{n} \frac{1}{1 + o_j} = 1, & \dfrac{1}{1 + o_j} > 0,\ \forall j = 1, 2, ..., n. \end{cases}$$

La condizione che $\tilde{p}_i = \frac{1}{1+o_i}$ *sia strettamente positiva (e sia un numero reale) corrisponde alla ovvia condizione che la quantità che si riceve in caso di vincita per la scommessa* i*, sia strettamente positiva, ovvero*

$$o_i + 1 > 0.$$

Esempio 1.2. *Continuando il precedente Esempio 1.1 è interessante mostrare come, se la condizione* $\sum_{j=1}^{n} \frac{1}{1+o_j} = 1$ *di assenza di opportunità di arbitraggio non è soddisfatta, ovvero se accade che*

$$O := \sum_{j=1}^{n} \frac{1}{1 + o_j} \neq 1,$$

allora si può trovare esplicitamente una strategia che permette un arbitraggio, e precisamente se si punta

$$x_i = \frac{1}{1 - O} \frac{1}{1 + o_i}, \quad \textit{per } i = 1, 2, ..., n,$$

si ottiene una vincita certa di 1*. Ciò è chiaramente equivalente a mostrare che se*

$$x_i = \alpha\, \frac{1}{1 + o_i}, \quad \textit{per } i = 1, 2, ..., n,$$

allora si vince un guadagno certo $\alpha\,(1 - O)$*. Ovviamente per ottenere che il guadagno sia positivo, il segno di* α *va scelto in dipendenza del segno di* $1 - O$*.*

Infatti se si verifica la vincita del cavallo j*, ovvero se si verifica* ω_j*, allora, tenendo presente che*

$$\Delta_j(\omega_j) = (1 + o_j) - 1 = o_j, \qquad \Delta_i(\omega_j) = 0 - 1 = -1, \quad \textit{per } i \neq j,$$

[31] Ovvero in caso di vincita si riceve $1 + o_i$ e quindi, tenuto conto del pagamento iniziale di 1, il ricavo totale è di $o_i = (1 + o_i) - 1$.

si ottiene come guadagno complessivo

$$\begin{aligned}
\sum_{i=1}^{n} x_i\, \Delta_i(\omega_j) &= x_j\, \Delta_j(\omega_j) + \sum_{i\neq j}^{1,n} x_i\, \Delta_i(\omega_j) \\
&= x_j\, o_j - \sum_{\ell\neq j} x_\ell = \alpha\, \frac{1}{1+o_j}\, o_j - \sum_{\ell\neq j} \alpha \frac{1}{1+o_\ell} \\
&= \alpha \left(1 - \frac{1}{1+o_j} - \sum_{\ell\neq j} \frac{1}{1+o_\ell} \right) = \alpha \left(1 - \sum_{\ell=1}^{n} \frac{1}{1+o_\ell} \right) \\
&= \alpha\,(1-O)\,.
\end{aligned}$$

Osservazione 1.3. *Nell'esempio precedente quindi accade che un capitale iniziale*

$$\alpha \sum_{i=1}^{n} \frac{1}{1+o_j} = \alpha\, O$$

permette di effettuare le n scommesse e, alla fine, permette di ottenere al botteghino sempre

$$\alpha = \alpha \frac{1}{1+o_j}\,(1+o_j),$$

qualunque sia il cavallo vincente. Il guadagno è quindi sempre $\alpha(1-O)$.

Per la linearità delle scommesse si tratta quindi di un arbitraggio considerando che, se inizialmente ho un capitale nullo, devo prendere in prestito dalla banca la quantità di denaro $\alpha\, O$, che alla fine devo restituire (corrisponde alla scommessa non rischiosa $\Delta_0 = \alpha\, O - \alpha\, O \equiv 0$) e quindi all'inizio ho $\alpha\, O - \alpha\, O$ e alla fine dal broker ricevo sempre α, qualunque sia il cavallo vincente, ma devo restituire il capitale $\alpha\, O$ inizialmente preso dalla banca, e quindi alla fine ho in totale $\alpha(1-O)$, che è strettamente positivo a seconda del segno di $\alpha(1-O)$.

In particolare se $O > 1$ allora per ottenere un guadagno sicuramente positivo si deve avere α negativo, il che corrisponde a prendere la parte del broker, mentre se $O < 1$ allora α è positivo e quindi è lo scommettitore che vince di sicuro con questa strategia.

1.4 Il modello binomiale uniperiodale

In questa sezione vedremo il modello di mercato più semplice possibile.

Definizione 1.3 (modello binomiale uniperiodale)**.** *Consideriamo ora il* **modello binomiale uniperiodale** *ovvero il caso in cui si abbia solo due tipi di titoli B (titolo non rischioso) ed S (titolo rischioso):*

al tempo t=0	$B_0 > 0$ *(si suppone spesso $B_0 = 1$ per semplicità)*	$S_0 = s_0 > 0$
al tempo t=1	$B_1 = B_0(1 + r)$	$S_1 = S_0\, Z = s_0\, Z$

dove il montante Z è una variabile aleatoria che può assumere solo due valori u (per UP) e d (per DOWN), da cui il nome binomiale (Attenzione: non è detto che necessariamente sia abbia $d < 1$, anche se questa circostanza è possibile, per questo motivo a volte useremo anche la simbologia $u = 1 + b$ e $d = 1 + a$).

Osservazione 1.4. *La circostanza che Z possa assumere solo due valori, si può tradurre nel fatto che $\Omega = \{\omega_0, \omega_1\}$, e che $Z(\omega_0) = d$ e $Z(\omega_1) = u$. Tuttavia questa interpretazione può risultare riduttiva, e conviene più in generale non assumere che la cardinalità di Ω sia due, ma conviene invece assumere che*

$$Z(\omega) = d\, I_{A_0}(\omega) + u\, I_{A_1}(\omega), \qquad \text{dove } A_0 = A_1^c,$$

o in altre parole che

$$Z(\omega) = \begin{cases} d & \text{se } \omega \in A_0 \\ u & \text{se } \omega \in A_1, \end{cases}$$

e assumere come sigma-algebra

$$\mathcal{F} = \{\emptyset,\, A_0,\, A_1,\, \Omega\},$$

di modo che sia la sigma-algebra $\mathcal{F}$ ad essere un insieme con un numero finito di elementi e non Ω.

Cominciamo con il caso in cui le uniche scommesse ammissibili[32] sono quelle del tipo

- *comprare (o anche vendere a corto[33]) il titolo rischioso al tempo $t = 0$* e *rivenderlo (o comprare, se si era venduto a corto, al tempo successivo $t = 1$.*

- *Mettere in banca (o chiedere in prestito) al tempo $t = 0$* e *ritirare (o restituire) al tempo successivo $t = 1$.*

Ovviamente le combinazioni lineari delle due precedenti danno luogo a tutte le scommesse ammissibili.

Il titolo non rischioso, come si usa solitamente, viene utilizzato per attualizzare i valori ai vari tempi. Quindi si considera anche un mercato $\tilde{B}$, $\tilde{S}$ con i valori attualizzati come segue

al tempo t=0	$\tilde{B}_0 = \dfrac{B_0}{B_0} = 1$	$\tilde{S}_0 = \dfrac{S_0}{B_0} = \dfrac{s_0}{B_0}$
al tempo t=1	$\tilde{B}_1 = \dfrac{B_1}{B_1} = 1$	$\tilde{S}_1 = \dfrac{S_1}{B_1} = \dfrac{S_1}{B_0(1+r)} = \dfrac{s_0\, Z}{B_0(1+r)}$

[32]Ovviamente i titoli sono due, e poi sono ammesse anche le combinazioni lineari delle due relative scommesse.
[33]Vendere a corto (o short selling) significa vendere anche senza possedere l'azione. La differenza tra comprare e vendere sarà nel segno della strategia.

Allora la scommessa relativa al titolo non rischioso B

$$\Delta_0 = \tilde{B}_1 - \tilde{B}_0 = 1 - 1 = 0$$

non comporta alcun contributo, mentre l'altra scommessa, relativa al titolo S diviene

$$\Delta_1 = \tilde{S}_1 - \tilde{S}_0 = \frac{s_0\, Z}{B_0(1+r)} - \frac{s_0}{B_0}.$$

Questo modello verrà studiato prima come un'applicazione del teorema dell'arbitraggio (Esempio 1.3) e successivamente si otterranno direttamente le condizioni necessarie e sufficienti per l'assenza di opportunità di arbitraggio (Teorema 1.2)

Esempio 1.3 (il modello binomiale uniperiodale con il teorema dell'arbitraggio)**.** *Per il teorema dell'arbitraggio, la condizione di assenza di opportunità di arbitraggio è equivalente alla condizione di esistenza della misura martingala equivalente, che in questo modello si riduce alla richiesta che esista una probabilità $\tilde{\mathbb{P}}$ su Ω per la quale valga*

$$\tilde{\mathbb{E}}(\Delta_1) = 0 \quad \Leftrightarrow \quad \tilde{\mathbb{E}}(\tilde{S}_1) = \tilde{s}_0 = \frac{s_0}{B_0} \quad \Leftrightarrow \quad \tilde{\mathbb{E}}\left(\frac{s_0\, Z}{B_0(1+r)} - \frac{s_0}{B_0}\right) = 0 \quad \Leftrightarrow \quad \tilde{\mathbb{E}}\,(Z) = 1 + r,$$

e che dia probabilità positiva a tutti gli scenari o gli eventi possibili.

In altre parole, la condizione equivale all'esistenza di due numeri $\tilde{p} > 0$ e $\tilde{q} > 0$, con $\tilde{p} + \tilde{q} = 1$, che rappresentino $\tilde{p} = \tilde{\mathbb{P}}(Z = u)$ e $\tilde{q} = \tilde{\mathbb{P}}(Z = d)$, rispettivamente, e per i quali si abbia

$$u\,\tilde{\mathbb{P}}(Z = u) + d\,\tilde{\mathbb{P}}(Z = d) = 1 + r \quad \Leftrightarrow \quad u\,\tilde{\mathbb{P}}(Z = u) + d\,\big(1 - \tilde{\mathbb{P}}(Z = u)\big) = 1 + r \quad \Leftrightarrow \quad u\,\tilde{p} + d\,(1 - \tilde{p}) = 1 + r$$

da cui, immediatamente,[34]

$$\tilde{p} = \frac{1 + r - d}{u - d}, \qquad \tilde{q} = 1 - \tilde{p} = \frac{u - (1+r)}{u - d}. \tag{1.8}$$

Infine, la condizione che i due numeri $\tilde{p}$ e $\tilde{q}$ definiscano una probabilità con $\tilde{p} = \tilde{\mathbb{P}}(Z = u) > 0$ e $\tilde{q} = \tilde{\mathbb{P}}(Z = d) > 0$ è soddisfatta se e solo se[35]

$$\boxed{d < 1 + r < u.}$$

**Il resto della sezione è dedicato alla dimostrazione diretta (senza far uso del teorema dell'arbitraggio) che la condizione precedente, ovvero $d < 1 + r < u$, è necessaria e sufficiente per non avere arbitraggi.

Premettiamo, però, un'osservazione che il lettore può tranquillamente saltare in una prima lettura.

Osservazione 1.5. *Nel caso in cui Ω sia composto solo di due eventi elementari, ovvero in cui $\Omega = \{\omega_0, \omega_1\}$ e $Z(\omega_0) = d$ mentre $Z(\omega_1) = u$, ovviamente la precedente **condizione permette di individuare univocamente la probabilità (misura martingala equivalente) sull'insieme delle parti. Nel caso in cui invece la cardinalità di Ω sia strettamente maggiore di due, allora la misura martingala equivalente $\tilde{\mathbb{P}}$ è univocamente determinata solo sulla σ-algebra generata da Z, ovvero, con le notazione dell'Osservazione 1.4, su $\mathcal{F}$, ma non su σ-algebre più grandi. Ad esempio se $\Omega = \{\omega_0, \omega_1, \omega_2, \omega_3\}$, e $Z(\omega_0) = Z(\omega_2) = d$ e $Z(\omega_1) = Z(\omega_3) = u$, allora $A_0 = \{Z = d\} = \{\omega_0, \omega_2\}$ e $A_1 = \{Z = u\} = \{\omega_1, \omega_3\}$. Se invece si prendesse come σ-algebra l'insieme delle parti, allora ogni misura di probabilità $\mathbb{Q}$ sarebbe determinata da*

$$p_0, p_1, p_2, p_3, \quad p_i \geq 0, \quad \sum_{i=0}^{3} p_i = 1,$$

[34]**Usando la simbologia $d = 1 + a$ e $u = 1 + b$, le espressioni per $\tilde{p}$ e $\tilde{q}$ diventano

$$\tilde{p} = \frac{r - a}{b - a}, \qquad \tilde{q} = 1 - \tilde{p} = \frac{b - r}{b - a}.$$

[35]**Usando la simbologia $d = 1 + a$ e $u = 1 + b$, la condizione $d < 1 + r < u$ diviene $1 + a < 1 + r < 1 + b$, ossia

$$a < r < b.$$

dove $p_i = \mathbb{Q}(\{\omega_i\})$. *La condizione che* $\mathbb{Q}$ *sia una misura martingala equivalente, diventerebbe*

$$p_1 + p_3 = \tilde{p}, \qquad p_0 + p_2 = \tilde{q} = 1 - \tilde{p}, \qquad 0 \le p_1,\, p_3 \le \tilde{p}, \quad 0 \le p_0,\, p_2 \le 1 - \tilde{p},$$

e non si avrebbe più l'unicità della misura martingala equivalente.

**Passiamo ora alla dimostrazione diretta che la condizione $d < 1 + r < u$, è necessaria e sufficiente per non avere arbitraggi.

Teorema 1.2. *Per il modello di mercato binomiale uniperiodale la condizione*

$$d < 1 + r < u. \tag{1.9}$$

è **necessaria e sufficiente** *per non avere arbitraggi (in senso debole), ovvero affinché con un capitale iniziale nullo, non sia possibile non avere (con certezza) perdite, ed avere un guadagno positivo con probabilità positiva.*

Prima di iniziare la dimostrazione fissiamo le notazioni: indichiamo una strategia con

$$\pi = (\beta, \gamma)$$

ovvero βB_0 è la quantità di denaro investita nel titolo non rischioso (in banca, o in buoni del tesoro[36]) mentre γS_0 è la quantità di denaro investita nel titolo rischioso (l'azione), ovvero γ è il numero di azioni comprate[37].

Allora il capitale iniziale (o il valore della strategia al tempo $t = 0$) è

$$X_0^\pi = \beta\, B_0 + \gamma\, S_0\, (= \beta + \gamma\, s_0 \quad \text{se } B_0 = 1)$$

mentre al tempo $t = 1$ vale

$$X_1^\pi = \beta\, B_1 + \gamma\, S_1 = \beta\, B_0\,(1 + r) + \gamma\, s_0\, Z\, (= \beta\,(1 + r) + \gamma\, s_0\, Z \quad \text{se } B_0 = 1)$$

ed il suo valore attualizzato è

$$\tilde{X}_1^\pi = \frac{X_1^\pi}{B_1} = \frac{\beta\, B_1 + \gamma\, S_1}{B_1} = \beta + \gamma\, \tilde{S}_1 = \beta + \gamma\, \frac{s_0\, Z}{B_0(1 + r)}$$

Dimostrazione.
La condizione (1.9) ***è necessaria***[38] affinché non ci siano opportunità di arbitraggio:

Mostreremo infatti che, se la condizione (1.9) non è soddisfatta, allora in ciascuno dei due casi $d < u \le 1 + r$ e $1 + r \le d < u$ esistono strategie (β, γ) di arbitraggio.

Se fosse

$$d < u \le 1 + r,$$

allora,**si otterrebbe un *pasto gratis* (*free lunch*) vendendo a corto l'azione e mettendo in banca i soldi ricevuti dalla vendita.**

Infatti, formalmente, con capitale iniziale nullo

• al tempo $t = 0$ si potrebbe vendere corto una azione $(\gamma = -1)$ al prezzo s_0, e mettere i soldi ottenuti in banca

$$0 = X_0^\pi = \beta\, B_0 + \gamma\, S_0 = \beta\, B_0 - s_0 \qquad (\Rightarrow \beta = +\frac{s_0}{B_0})$$

[36] Nel caso dei buoni β rappresenta il numero di buoni ***comprati***. È importante capire che quando β è negativo significa che stiamo prendendo in prestito i soldi dalla banca, oppure che stiamo vendendo β buoni.

[37] È anche importante capire che è ammessa la ***vendita a corto***, cioè ***si può vendere oggi un titolo***, ***impegnandosi a fornirlo domani***, ***pur non avendolo ancora comprato oggi***. Ciò vale anche per i buoni, ovvero per i titoli non rischiosi e non solo per le azioni, ovvero per i titoli rischiosi.

[38] La dimostrazione formale oscura leggermente il significato:
Se $d < u \le 1 + r$, allora conviene vendere l'azione oggi e mettere in banca s_0, domani in banca ci sarà di sicuro una somma sufficiente per comprare l'azione, e c'è anche la possibilità (se l'azione non cresce troppo) di un guadagno strettamente positivo.
Se $1 + r \le d < u$, allora conviene comprare l'azione oggi, prendendo in prestito dalla banca la cifra s_0 necessaria, domani rivendendo l'azione di sicuro si otterrà una somma sufficiente per restituire i soldi alla banca, e c'è anche la possibilità (se l'azione cresce abbastanza) di un guadagno strettamente positivo.

• al tempo $t=1$ si potrebbe comprare l'azione al prezzo $s_0\,Z$ (e che risulta minore di $s_0(1+r)$, cioè il valore di quanto depositato in banca); dopo aver ritirato i soldi in banca, si compra l'azione (che si era venduta a corto) ricavando quindi

$$\beta\,B_1+\gamma\,S_1=s_0(1+r)-s_0\,Z=\begin{cases} s_0(1+r-u)\geq 0 & \text{se } Z=u,\\ s_0(1+r-d)>0 & \text{se } Z=d.\end{cases}$$

In termini del mercato attualizzato

$$\beta+\gamma\,\tilde{S}_1=s_0-\frac{s_0\,Z}{1+r}=\begin{cases} s_0\Big(1-\dfrac{u}{1+r}\Big)\geq 0 & \text{se } Z=u,\\[2ex] s_0\Big(1-\dfrac{d}{1+r}\Big)>0 & \text{se } Z=d.\end{cases}$$

Se invece fosse

$$1+r\leq d<u,$$

allora, ,**si otterrebbe un *pasto gratis* (*free lunch*) comprando l'azione, con i soldi presi a prestito i soldi necessari dalla banca.**

Infatti, formalmente, con capitale iniziale nullo,

• al tempo $t=0$ si potrebbe comprare una azione ($\gamma=1$) al prezzo s_0, prendendo i soldi in prestito dalla banca

$$0=X_0^\pi=\beta\,B_0+\gamma\,S_0=\beta\,B_0+s_0\qquad(\Rightarrow\beta=-\frac{s_0}{B_0})$$

• al tempo $t=1$ si potrebbe vendere l'azione al prezzo $s_0\,Z$ (e che risulta strettamente maggiore di $s_0(1+r)$, cioè il valore di quanto va restituito in banca); dopo aver venduto l'azione si restituiscono i soldi (che si erano presi in prestito) alla banca, ricavando quindi

$$\beta\,B_1+\gamma\,S_1=-s_0(1+r)+s_0\,Z=\begin{cases} s_0\big(u-(1+r)\big)>0 & \text{se } Z=u,\\ s_0\big(d-(1+r)\big)\geq 0 & \text{se } Z=d.\end{cases}$$

In termini del mercato attualizzato

$$\beta+\gamma\,\tilde{S}_1=-s_0+\frac{s_0\,Z}{1+r}=\begin{cases} s_0\Big(\dfrac{u}{1+r}-1\Big)>0 & \text{se } Z=u,\\[2ex] s_0\Big(\dfrac{d}{1+r}-1\Big)\geq 0 & \text{se } Z=d.\end{cases}$$

La condizione (1.9) **è sufficiente affinché non ci siano strategie di arbitraggio**:

Mostreremo infatti che, se la condizione (1.9) è soddisfatta, allora non esistono strategie (β,γ) di arbitraggio.

Dobbiamo cioè mostrare che, se vale (1.9), allora per ogni (β,γ) tale che

$$X_0^\pi=\beta\,B_0+\gamma\,s_0=0\qquad\text{e}\qquad X_1^\pi(\omega)=\beta\,B_0(1+r)+\gamma\,s_0\,Z(\omega)\geq 0\qquad\text{per ogni }\omega$$

si ha che

$$X_1^\pi(\omega)=0,\qquad\text{per ogni }\omega.$$

Infatti

$$X_0^\pi=\beta\,B_0+\gamma\,s_0=0,\qquad\Leftrightarrow\beta\,B_0=-\gamma\,s_0,$$

e di conseguenza

$$X_1^\pi(\omega)=\beta\,B_0(1+r)+\gamma\,s_0\,Z(\omega)=\begin{cases} -\gamma\,s_0\big(1+r-u\big) & \text{se } Z=u,\\ -\gamma\,s_0\big(1+r-d\big) & \text{se } Z=d,\end{cases}$$

Poiché $1 + r - u < 0$ mentre $1 + r - d > 0$, è impossibile che $-\gamma\, s_0\big(1 + r - u\big)$ e $-\gamma\, s_0\big(1 + r - d\big)$ siano entrambi maggiori o uguali a zero, tranne nel caso in cui siano entrambi uguali a zero, ovvero se $\gamma s_0 = 0$. Poiché $s_0 > 0$, ciò significa che deve essere $\gamma = 0$. Lo stesso vale per $\beta = -\dfrac{\gamma\, s_0}{B_0}$ e quindi $X_1^\pi(\omega) = \beta\, B_0(1 + r) + \gamma\, s_0\, Z(\omega) = 0$. □

1.4.1 Modello binomiale uniperiodale con contingent claim

Siamo sempre nel mercato binomiale uniperiodale della definizione 1.3, ma ora sono ammesse anche scommesse del tipo ***contingent claim***[39], ovvero il valore della scommessa dipende dal valore del titolo rischioso:

- oggi (al tempo $t = 0$) pago c e domani (al tempo $t = 1$) ricevo $f_1(S_1) = f_1(s_0\, Z) =: \Phi(Z)$.

Il valore attualizzato di questa scommessa è

$$\Delta_2 = \frac{f_1(S_1)}{B_1} - \frac{c}{B_0} = \frac{f_1(S_1)}{B_0(1+r)} - \frac{c}{B_0} = \frac{\Phi(Z)}{B_0(1+r)} - \frac{c}{B_0}.$$

Il modello binomiale uniperiodale con derivati[40] verrà studiato prima come un'applicazione del teorema dell'arbitraggio (Esempio 1.4) per determinare il prezzo del derivato in modo che non ci siano opportunità di arbitraggio. Successivamente (si veda l'esempio 1.5) troveremo di nuovo tale prezzo con un procedimento collegato con il concetto di strategia di copertura perfetta (si veda la definizione 1.4).

Esempio 1.4 (modello binomiale uniperiodale con contingent claim: il prezzo, sempre con il teorema dell'arbitraggio)**.** *Rispetto all'Esempio 1.3, abbiamo un'altra scommessa, ma lo spazio Ω è rimasto lo stesso di prima, e abbiamo anche la scommessa di prima, quindi la misura martingala equivalente o rimane la stessa di prima o non esiste.*

Il teorema dell'arbitraggio assicura l'assenza di opportunità di arbitraggio se e solo se la scommessa Δ_2 ha valore atteso nullo rispetto a tale misura (oltre alla scommessa Δ_1). Quindi l'unico prezzo che non permette arbitraggi è perciò quel valore c per il quale si abbia

$$\tilde{\mathbb{E}}(\Delta_2) = \tilde{\mathbb{E}}\left(\frac{f_1(S_1)}{B_1} - \frac{c}{B_0}\right) = 0 \quad \Leftrightarrow \quad c = B_0\, \tilde{\mathbb{E}}\left(\frac{f_1(S_1)}{B_1}\right)$$

ossia

$$\tilde{\mathbb{E}}(\Delta_2) = 0 \quad \Leftrightarrow \quad \tilde{\mathbb{E}}\left(\frac{\Phi(Z)}{1+r} - c\right) = 0 \quad \Leftrightarrow \quad c = \tilde{\mathbb{E}}\left(\frac{\Phi(Z)}{1+r}\right),$$

da cui

$$c = \frac{1}{1+r}\left[f_1(s_0\, u)\, \frac{1+r-d}{u-d} + f_1(s_0\, d)\, \frac{u-(1+r)}{u-d}\right]. \tag{1.10}$$

Quindi il teorema dell'arbitraggio[41] *ci permette di ottenere il prezzo del contingent claim $f_1(S_1)$.*

Osservazione 1.6. ***Contingent claim*** *è uno dei possibili nomi che si può dare ad un contratto che preveda che la quantità di denaro (o di beni) che ci si impegna a dare dipenda (sia contingente) dalle circostanze future. In questo esempio abbiamo supposto che la dipendenza sia stabilita in modo deterministico dall'andamento dei prezzi di un altro bene, che è detto* **(bene) sottostante** *(in questo caso il sottostante, o* **titolo primario***, è l'azione).*

[39] Per chiarire meglio il significato delle due parole
CONTINGENT: (aggettivo, formale) contingent on/upon something, depending on something else in the future in order to happen:
Outdoor arrangements are, as ever, ***contingent*** *on the weather and we have other plans in the event of rain.*
Our success is ***contingent*** *upon your support.*
CLAIM: (fra gli altri significati) a right to have something or obtain something from someone:
She has no rightful ***claim*** *to the title.*
Our neighbours have no ***claim*** *to* (= cannot say that they own) *that strip of land between our houses.*
My ex-wife has no ***claims*** *on me* (= has no right to any of my money).
[40] Si veda la successiva osservazione 1.6 per il motivo per cui si parla di derivati.

[41] Infatti solo scegliendo come prezzo $c := \tilde{\mathbb{E}}\left(\frac{\Phi(Z)}{1+r}\right)$, si ottiene che non ci sono opportunità di arbitraggio.

**Il resto di questa sezione è dedicato alla dimostrazione diretta, attraverso il concetto di copertura perfetta.
Primo di procedere in tale senso, diamo un'osservazione che riguarda l'unicità o meno della misura martingala equivalente, e che puo' essere tralasciata in una prima lettura.

Osservazione 1.7. *Oltre a derivati che sono funzioni deterministiche del sottostante, sarebbe possibile considerare anche scommesse (attualizzate) più generali, ovvero del tipo*

$$\Delta_\Psi(\omega) = \frac{\Psi(\omega)}{B_1} - \frac{c_\Psi}{B_0} = \frac{\Psi(\omega)}{B_0(1+r)} - \frac{c_\Psi}{B_0}$$

Se lo spazio degli eventi non contenesse solo due elementi, allora con il teorema dell'arbitraggio si individuerebbe in modo unico la misura martingala equivalente solo sulla σ-algebra generata da Z, cioè su $\mathcal{F} = \big\{\emptyset, \{Z = d\}, \{Z = u\}, \Omega\big\}$, ma non si individuerebbe in modo unico su tutte le possibili σ-algebre.
Per capire meglio il significato di questa osservazione riprendiamo il caso esaminato nell'Osservazione 1.5 di $\Omega = \{\omega_0, \omega_1, \omega_2, \omega_3\}$, e la σ-algebra sia l'insieme delle parti di Ω.
La richiesta che $\mathbb{Q}$ sia una misura martingala equivalente corrisponde a chiedere che

$$\mathbb{E}^{\mathbb{Q}}[\Delta_0] = 0$$
$$\mathbb{E}^{\mathbb{Q}}[\Delta_1] = 0$$
$$\mathbb{E}^{\mathbb{Q}}[\Delta_\Psi] = 0$$

La prima condizione è ovvia, la seconda condizione (come abbiamo già visto) corrisponde a

$$p_1 + p_3 = \tilde{p}, \qquad p_0 + p_2 = \tilde{q} = 1 - \tilde{p}, \qquad 0 \le p_1,\, p_3 \le \tilde{p}, \quad 0 \le p_0,\, p_2 \le 1 - \tilde{p},$$

ovvero

$$0 \le p_0 \le 1 - \tilde{p}, \quad 0 \le p_1 \le \tilde{p}, \quad p_2 = 1 - \tilde{p} - p_0, \quad p_3 = \tilde{p} - p_1, \tag{1.11}$$

che come è chiaro ha due gradi di libertà (cioè le soluzioni dipendono da $(p_0, p_1) \in [0, \tilde{p}] \times [0, 1 - \tilde{p}]$.
Infine la terza condizione diviene

$$p_0\,\Psi(\omega_0) + p_1\,\Psi(\omega_1) + p_2\,\Psi(\omega_2) + p_3\,\Psi(\omega_3) - (1+r)B_0\frac{c_\Psi}{B_0} = 0$$

ovvero tenendo conto della (1.11)

$$p_0\,\Psi(\omega_0) + p_1\,\Psi(\omega_1) + (1 - \tilde{p} - p_0)\,\Psi(\omega_2) + (\tilde{p} - p_1)\,\Psi(\omega_3) - (1+r)B_0\frac{c_\Psi}{B_0} = 0. \tag{1.12}$$

Da questa relazione si ottiene quindi che c_Ψ non è univocamente determinato dalla richiesta che non ci siano opportunità di arbitraggio, ossia, mentre nel caso in cui $\mathcal{F} = \{\emptyset, A_0, A_1, \Omega\}$ il suo prezzo è univocamente determinato dalla (unica) misura di probabilità martingala equivalente dalla relazione $c_\Psi = B_0\,\tilde{\mathbb{E}}\Big(\dfrac{\Psi}{B_1}\Big)$, nel caso in esame, in cui $\mathcal{F} = \mathcal{P}(\Omega)$, il prezzo può variare in un range che dipende, in non solo dai valori che può assumere $\Psi(\omega)$, ma anche da (p_0, p_1), senza che ci siano opportunità di arbitraggio. Tuttavia se accade che

$$\Psi(\omega_0) = \Psi(\omega_2) \quad e \quad \Psi(\omega_1) = \Psi(\omega_3), \qquad \Leftrightarrow \qquad \Psi(\omega) = \Phi(Z(\omega)),$$

per una opportuna funzione $\Phi(z)$, allora invece il suo prezzo è univocamente determinato.

Vale la pena anche di osservare che potrebbe anche darsi la situazione in cui alcuni contingent claim $\Psi_j(\omega)$, $j = 1, \cdots, \ell$ sono prezzati dal mercato e sia c_j il corrispondente prezzo di mercato. In altre parole non abbiamo il problema di fare il prezzo di questi contingent claim, ma prendiamo per buoni i prezzi di questi contingent claim. In questo caso le condizioni di assenza di opportunità di arbitraggio

$$\mathbb{E}^{\mathbb{Q}}[\Delta_0] = 0 \tag{1.13}$$
$$\mathbb{E}^{\mathbb{Q}}[\Delta_1] = 0 \tag{1.14}$$
$$\mathbb{E}^{\mathbb{Q}}[\Delta_{\Psi_j}] = 0 \qquad j = 1, \ldots \ell \tag{1.15}$$

dove

$$\Delta_{\Psi_j}(\omega) = \frac{\Psi_j(\omega)}{B_1)} - \frac{c_j}{B_0} = \frac{\Psi(\omega)}{B_0(1+r)} - \frac{c_j}{B_0}$$

divengono ulteriori condizioni sulla probabilità martingala equivalente. Ad esempio, se applicate sempre all'esempio dell'Osservazione 1.5 con $\Omega = \{\omega_0, \omega_1, \omega_2, \omega_3\}$*, permetterebbero di ricavare relazioni analoghe alla* (1.12)*, ovvero*

$$p_0\,\Psi(\omega_0) + p_1\,\Psi(\omega_1) + (1-\tilde{p}-p_0)\,\Psi(\omega_2) + (\tilde{p}-p_1)\,\Psi(\omega_3) - (1+r)B_0\frac{c_j}{B_0} = 0, \qquad j = 1, \cdots \ell. \tag{1.16}$$

Queste relazioni si possono interpretare in questo caso invece come una condizione su (p_0, p_1)*. Non solo, esse ci permettono altre considerazioni:*
(i) *se* $\ell = 1$*, allora è possibile determinare* p_1 *in funzione di* p_0 *(o viceversa);*
(ii) *se* $\ell = 2$*, allora è possibile che siano univocamente determinate sia* p_0 *che* p_1*;*
(iii) *se* $\ell \geq 3$*, allora è possibile che nessuna coppia di valori* (p_0, p_1) *soddisfi tutte le condizioni* (1.16)*, ovvero che il sistema* (1.13) (1.14) (1.15) *di assenza di opportunità di arbitraggio non abbia nessuna soluzione.*

**Consideriamo ora la seguente situazione: come nell'osservzione precedente oltre a derivati che sono funzioni deterministiche del sottostante, consideriamo anche scommesse (attualizzate) più generali, ovvero del tipo

$$\Delta_\Psi(\omega) = \frac{\Psi(\omega)}{B_1} - \frac{c_\Psi}{B_0} = \frac{\Psi(\omega)}{B_0(1+r)} - \frac{c_\Psi}{B_0};$$

supponiamo inoltre che

• esista (e sia unica) una misura di probabilità $\tilde{\mathbb{P}}$, rispetto alla quale le variabili aleatorie Δ_i hanno valore atteso nullo, per $i = 0, 1, \cdots, m$,

e

• il contingent claim Ψ è una variabile aleatoria che si può scrivere come combinazione lineare delle precedenti, ovvero esiste un vettore di numeri reali $(\beta, \gamma^1, \cdots, \gamma^m)$ per i quali

$$**\Delta_\Psi(\omega) = \beta\Delta_0 + \sum_{i=1}^m \gamma^i \Delta_i(\omega) \qquad \text{per ogni } \omega \in \Omega.$$

In questo caso infatti la condizione

$$\tilde{\mathbb{E}}[\Delta_\Psi] = 0 \quad \Leftrightarrow \quad \beta\tilde{\mathbb{E}}[\Delta_0] + \sum_{i=1}^m \gamma^i \tilde{\mathbb{E}}[\Delta_i] = 0$$

è automaticamente soddisfatta, e quindi, aggiungere al mercato tali tipi di scommesse non comporta arbitraggi. Di conseguenza il valore c_Ψ si può anche ricavare come quel valore, se esiste, per cui il contingente claim Ψ si puo' scrivere come combinazione lineare delle scommesse di base.**

Questa banale osservazione ha delle interessanti interpretazioni in Finanza come si vede nel seguente esempio. Il vettore $(\beta, \gamma^1, \cdots, \gamma^m)$ viene interpretato come una ***strategia di copertura perfetta***, e contemporaneamente permette, in alcuni casi la determinazione dei prezzi.

Esempio 1.5 (modello binomiale uniperiodale con contingent claim: la strategia di copertura perfetta)**.** *Il prezzo del derivato con contingent claim* $f_1(S_1)$ *(che brevemente chiameremo anche opzione) si può ottenere anche a partire dal concetto di copertura perfetta (si veda la successiva definizione 1.4).*

Sia c *il prezzo (ancora da determinare) dell'opzione (cioè il prezzo del derivato con contingent claim* $f_1(S_1) = f(s_0\,Z) =: \Phi(Z)$*). Si immagini di comprare* α *derivati caratterizzati dal precedente contingent claim* $\Phi(Z)$*. Nel seguito considereremo solo il caso* $\alpha = -1$*, che invece corrisponde a vendere l'opzione, e quindi prenderemo il punto di vista del venditore, ma questo non lede la generalità*[42]*. Siano inoltre* $\beta\,B_0$ *la quantità di denaro investita nel titolo non rischioso (B), e* γS_0 *la quantità di denaro investita nel titolo rischioso (S). La coppia* $\pi := (\beta, \gamma)$ *viene detta* **strategia di**

[42] È chiaro che supporre $\alpha = -1$ non comporta nessuna perdita in generalità in quanto basterebbe considerare $\beta' = \frac{\beta}{-\alpha}$ e analogamente $\gamma' = \frac{\gamma}{-\alpha}$.

investimento o portfolio[43].

Si suppone inoltre che la strategia sia **autofinanziante**, *cioè che non ci siano costi di transazione, non ci siano consumi, e neppure introiti ulteriori.*

La condizione che il capitale iniziale sia nullo diviene

$$(X_0^{\alpha,\pi} =)\alpha\, c + X_0^{\pi} = \alpha\, c + \beta\, B_0 + \gamma\, S_0 = 0 \tag{1.17}$$

nel caso in cui $\alpha = -1$ *corrisponde a chiedere che*

$$X_0^{\pi} = \beta\, B_0 + \gamma\, S_0 = c. \tag{1.18}$$

Questa formulazione corrisponde a chiedere che il prezzo di vendita c sia il capitale iniziale che viene ripartito nel mercato (B, S)*, ovvero investito in banca (o nel titolo non rischioso) e nelle azioni (o nel titolo rischioso).*
Ovviamente alla fine del periodo, ovvero al tempo $t = 1$*, il valore complessivo si esprime come*

$$(X_1^{\alpha,\pi} =)\alpha\, f_1(S_1) + X_1^{\pi} = \alpha\, f_1(S_1) + \beta\, B_1 + \gamma\, S_1,$$

mentre il valore complessivo attualizzato si esprime come

$$(\tilde{X}_1^{\alpha,\pi} =)\alpha\, \frac{f_1(S_1)}{B_0(1+r)} + \beta\, 1 + \gamma\, \tilde{S}_1 = \begin{cases} \alpha\, \dfrac{f_1(s_0\, u)}{B_0(1+r)} + \beta\, 1 + \gamma\, \dfrac{s_0\, u}{B_0(1+r)} & \text{se} \quad Z = u \\[2ex] \alpha\, \dfrac{f_1(s_0\, d)}{B_0(1+r)} + \beta\, 1 + \gamma\, \dfrac{s_0\, d}{B_0(1+r)} & \text{se} \quad Z = d. \end{cases}$$

Definizione 1.4 (copertura perfetta nel modello binomiale uniperiodale)**.** *Una* **strategia di copertura perfetta** *è una strategia* (β, γ) *per la quale il capitale finale*

$$\alpha\, f_1(S_1) + \beta\, B_0(1+r) + \gamma\, S_1 = 0$$

o, il che è lo stesso, che il capitale finale attualizzato

$$\alpha\, \frac{f_1(S_1)}{B_0(1+r)} + \beta\, 1 + \gamma\, \tilde{S}_1 = 0$$

sia nel caso $Z = u$ *che nel caso* $Z = d$*.*
Nel caso $\alpha = -1$ *ciò corrisponde a chiedere che il valore*

$$X_1^{\pi} = \beta\, B_1 + \gamma\, S_1 = f(S_1), \tag{1.19}$$

ovvero che la strategia (β, γ) *permetta di onorare esattamente il contratto: vendendo l'opzione il venditore (cioè chi ha la posizione corta) si è impegnato a fornire al compratore (cioè chi ha la posizione lunga) la quantità contingente* $f(S_1)$*, e questa strategia* (β, γ) *gli permette di ottenere esattamente la quantità di denaro necessaria.*

Affermazione: *Si supponga che esista e siano unico un valore c che assicura l'esistenza di una strategia di copertura perfetta* (β, γ)*, a partire da un capitale iniziale nullo (ossia c e* (β, γ) *sono tali che valgono le relazioni* (1.18) *e* (1.19)*).*
Allora il valore c è l'unico valore del prezzo che assicura che non ci siano opportunità di arbitraggio[44].
Dimostrazione dell'affermazione precedente*: si distinguono due casi*

[43] Secondo la terminologia del teorema dell'arbitraggio 1.1 la strategia sarebbe la terna (α, β, γ), ma come già osservato si può considerare $\alpha = -1$, e quindi solo la parte relativa a (β, γ) è interessante.

[44] Tale valore c prende anche il nome di ***prezzo di copertura***. Il suo interesse sta nel seguente fatto: colui che vende l'opzione al prezzo c di copertura riceve la somma c al tempo $t = 0$ e la investe nel mercato (B, S) secondo la strategia di copertura perfetta (β, γ), ed è sicuro in tale modo di ottemperare all'impegno preso, cioè di poter fornire al tempo $t = 1$ il contingent claim $f_1(S_1)$ al compratore. Questo è interessante anche per il compratore, in quanto sa che il venditore riuscirà ad ottenere il contingent claim.

• *se il prezzo dell'opzione della strategia fosse* $\bar{c}$*, con* $\bar{c} > c$*, allora vendendo al tempo* $t = 0$ *l'opzione al prezzo* $\bar{c}$*, si potrebbe utilizzare* c *per mettere in atto la strategia* (β, γ) *(che permette di ottenere* $f(S_1)$ *al tempo* $t = 1$*) e mettere in banca la differenza* $\bar{c} - c > 0$*. Al tempo* $t = 1$ *il venditore potrebbe ottemperare al contratto, dando, come dovuto, al compratore* $f(S_1)$*, e avere inoltre la quantità di denaro* $(1 + r)(\bar{c} - c) > 0$*. Si ha cioè un arbitraggio.*

Riassumendo

	opzione	*strategia = azioni e bond/banca*	*bond/banca*	*totale*
$t = 0$	$+\bar{c}$	$-\gamma S_0 - \beta B_0 = -c$	$\bar{c} - c$	0
$t = 1$	$-f(S_1)$	$\gamma S_1 + \beta B_1 = f(S_1)$	$(1 + r)(\bar{c} - c)$	$(1 + r)(\bar{c} - c) > 0$

• *se invece il prezzo dell'opzione della strategia fosse* $\underline{c}$*, con* $\underline{c} < c$*, allora vendendo al tempo* $t = 0$ *il portfolio* (β, γ) *al prezzo* $c = \beta\, B_0 + \gamma\, S_0 > \underline{c}$ *(impegnandosi a restituire al tempo* $t = 1$ *il corrispettivo valore, ossia* $\beta\, B_1 + \gamma\, S_1$*), comprando al tempo* $t = 0$ *l'opzione al prezzo* $\underline{c}$*, e infine, sempre al tempo* $t = 0$*, mettendo in banca la differenza* $c - \underline{c}$*, al tempo* $t = 1$ *si ottempera all'impegno preso prendendo il valore* $f(S_1)$ *tramite l'opzione: per tale valore infatti vale* $f(S_1) = \beta\, B_1 + \gamma\, S_1$*. A questo punto al tempo* $t = 1$ *in banca è rimasta la quantità* $(1 + r)(c - \underline{c}) > 0$ *strettamente positiva. Si ha cioè un arbitraggio.*

Riassumendo

	opzione	*strategia = azioni e bond/banca*	*bond/banca*	*totale*
$t = 0$	$-\underline{c}$	$\gamma S_0 + \beta B_0 = c$	$c - \underline{c}$	0
$t = 1$	$f(S_1)$	$-\gamma S_1 - \beta B_1 = -f(S_1)$	$(1 + r)(c - \underline{c})$	$(1 + r)(c - \underline{c}) > 0$

Rimane da vedere che il prezzo c *di copertura dell'affermazione esista e sia unico e che esso coincide con il prezzo ottenuto precedentemente, attraverso la formula* (1.10)*, ottenuta con il teorema dell'arbitraggio. L'idea alla base del ragionamento sta nella seguente osservazione: affinché si trovi qualcuno disposto a comprare, il prezzo deve essere tale che il venditore non possa fare un guadagno sicuro, ma anche non si troverebbe nessuno disposto a vendere se il compratore avesse un guadagno sicuro.*[45]*. Questo è possibile solo se, sempre a partire da capitale iniziale nullo* $X_0^{\alpha,\pi} = 0$*, c'è una strategia per la quale* $\tilde{X}_1^{\alpha,\pi} = 0$*, indipendentemente dal valore di* Z*.*

In particolare deve valere $X_0^{\alpha,\pi} = 0$*, cioè la relazione* (1.17)*, e la condizione di copertura, cioè* $\tilde{X}_1^{\alpha,\pi} = 0$*. Quest'ultima si traduce nelle due condizioni seguenti:*

$$\alpha\, \frac{f_1(s_0\, u)}{B_0(1+r)} + \beta\, 1 + \gamma\, \frac{s_0\, u}{B_0(1+r)} = 0 \tag{1.20}$$

$$\alpha\, \frac{f_1(s_0\, d)}{B_0(1+r)} + \beta\, 1 + \gamma\, \frac{s_0\, d}{B_0(1+r)} = 0 \tag{1.21}$$

o equivalentemente

$$\alpha\, \frac{f_1(s_0\, u)}{B_0(1+r)} + \beta\, 1 + \gamma\, \frac{s_0\, u}{B_0(1+r)} = 0 \tag{1.22}$$

$$\alpha\, \frac{f_1(s_0\, u)}{B_0(1+r)} + \beta\, 1 + \gamma\, \frac{s_0\, u}{B_0(1+r)} = \alpha\, \frac{f_1(s_0\, d)}{B_0(1+r)} + \beta\, 1 + \gamma\, \frac{s_0\, d}{B_0(1+r)} \tag{1.23}$$

[45]Va inoltre ricordato che implicitamente abbiamo supposto che il mercato sia ***liquido***, cioè che sia possibile trovare sempre qualcuno disposto a comprare e qualcuno disposto a vendere: se ci fossero opportunità di arbitraggio per il venditore, allora tutti vorrebbero vendere e non si troverebbe nessuno disposto a comprare, mentre accadrebbe il contrario se ci fossero opportunità di arbitraggio per il compratore.
.

ovvero, esplicitando β *in funzione di* α *e di* γ *tramite la* (1.22)

$$\beta = \beta(\alpha,\gamma) = -\alpha\,\frac{f_1(s_0\,u)}{B_0(1+r)} - \gamma\,\frac{s_0\,u}{B_0(1+r)} \tag{1.24}$$

$$\alpha\,\frac{f_1(s_0\,u)}{B_0(1+r)} + \gamma\,\frac{s_0\,u}{B_0(1+r)} = \alpha\,\frac{f_1(s_0\,d)}{B_0(1+r)} + \gamma\,\frac{s_0\,d}{B_0(1+r)}. \tag{1.25}$$

La seconda condizione (1.25) *permette di trovare* γ *in funzione di* α

$$\alpha\,\big(f_1(s_0\,u) - f_1(s_0\,d)\big) = \gamma\,\big(s_0\,d - s_0\,u\big) \Leftrightarrow \gamma = \gamma(\alpha) = -\alpha\frac{1}{s_0}\,\frac{f_1(s_0\,u) - f_1(s_0\,d)}{u-d},$$

ovvero[46] *per* $\alpha = -1$ *(*$B_0 = 1$*, per semplicità)*

$$\gamma = \frac{f_1(s_0\,u) - f_1(s_0\,d)}{s_0 u - s_0 d}. \tag{1.26}$$

Di conseguenza la prima condizione (1.24) *permette di trovare*[47] β*, infatti si ha*

$$\beta = \beta(\alpha) = -\alpha\,\frac{1}{B_0(1+r)}\,\frac{f_1(s_0\,d)\,u - f_1(s_0\,u)\,d}{u-d}$$

che per $\alpha = -1$ *diviene*

$$\beta = \frac{1}{B_0(1+r)}\,\frac{f_1(s_0\,d)\,u - f_1(s_0\,u)\,d}{u-d} \tag{1.27}$$

La condizione precedente (1.17) *di capitale iniziale nullo, ossia* $X_0^{\alpha,\pi} = 0$*, diviene quindi*[48] *per* $\alpha = -1$ *(si può supporre anche* $B_0 = 1$*, per semplicità)*

$$\beta\,B_0 + \gamma\,S_0 = \beta + \gamma\,s_0 = c \qquad \big(\Rightarrow \beta = c - \gamma\,s_0\big)$$

da cui si può trovare (in modo nuovo) il valore che deve avere c affinché non ci siano opportunità di arbitraggio: infatti

---

[46]Si osservi che il numero di azioni che permettono una copertura perfetta del titolo derivato, cioè il numero γ, si ottiene attraverso una sorta di ***derivata discreta*** della funzione costo $f(x)$, come funzione del prezzo dell'azione sottostante. Inoltre $\gamma(\alpha)$ dipende linearmente da α.

[47]Si osservi che

$$\begin{aligned}
\beta = \beta(\alpha,\gamma(\alpha)) &= -\alpha\,\frac{f_1(s_0\,u)}{B_0(1+r)} - \gamma(\alpha)\,\frac{s_0\,u}{B_0(1+r)}\\
&= -\alpha\,\frac{f_1(s_0\,u)}{B_0(1+r)} + \alpha\frac{1}{s_0}\,\frac{f_1(s_0\,u) - f_1(s_0\,d)}{u-d}\,\frac{s_0\,u}{B_0(1+r)}\\
&= -\alpha\,\frac{1}{B_0(1+r)}\Big[f_1(s_0\,u) - \frac{1}{s_0}\,\frac{f_1(s_0\,u) - f_1(s_0\,d)}{u-d}\,s_0\,u\Big]\\
&= -\alpha\,\frac{1}{B_0(1+r)}\Big[\frac{f_1(s_0\,u)\,(u-d)}{u-d} - \frac{f_1(s_0\,u)\,u - f_1(s_0\,d)\,u}{u-d}\Big]\\
&= -\alpha\,\frac{1}{B_0(1+r)}\,\frac{f_1(s_0\,d)\,u - f_1(s_0\,u)\,d}{u-d}
\end{aligned}$$

[48]Per α generico si ottiene lo stesso:

$$\alpha\,c + \beta(\alpha)\,B_0 + \gamma(\alpha)\,S_0 = \alpha\,c + (-\alpha)\,\beta(-1)\,B_0 + (-\alpha)\,\gamma(-1)\,s_0 = 0$$

che equivale a

$$\beta\,B_0 + \gamma\,S_0 = \beta + \gamma\,s_0 = c$$

dove per semplicità $\beta = \beta(-1)$ e $\gamma = \gamma(-1)$.

si ricava che[49]

$$
\begin{aligned}
c &= \frac{1}{1+r}\,\frac{f_1(s_0\,d)\,[u-(1+r)] + f_1(s_0\,u)\,[(1+r)-d]}{u-d}\\
&= \frac{1}{1+r}\Big(f_1(s_0\,d)\,\tilde p + f_1(s_0\,u)\,(1-\tilde p)\Big),
\end{aligned}
$$

dove $\tilde p$ *è definita in* (1.8).

Nella seguente Figura 1.1 riassumiamo il procedimento con un albero:

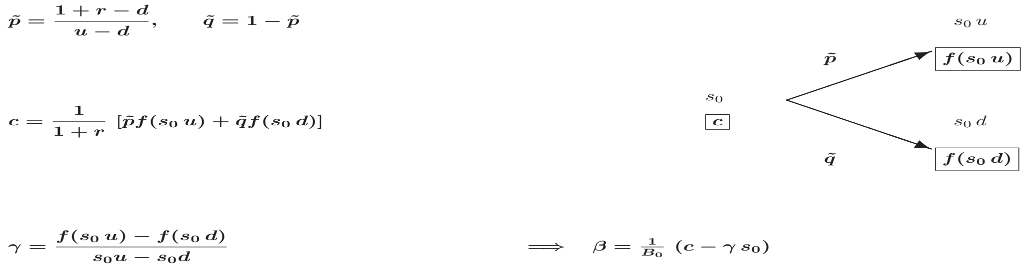

Figura 1.1: Albero binomiale per il calcolo del prezzo c e della strategia di copertura perfetta: trovati c e γ, il valore β è automaticamente determinato.

Osservazione 1.8. *La condizione di copertura perfetta è stata definita attraverso le condizioni* (1.18) *ed* (1.19)*, ossia*

$$
\begin{aligned}
c &= \beta\,B_0 + \gamma\,S_0,\\
f(S_1) &= \beta\,B_1 + \gamma\,S_1.
\end{aligned}
$$

Da queste relazioni è immediato verificare che allora, attualizzando i prezzi al tempo $t=0$,

$$
\begin{aligned}
\frac{c}{B_0} &= \beta\,\frac{B_0}{B_0} + \gamma\,\frac{S_0}{B_0},\\
\frac{f(S_1)}{B_1} &= \beta\,\frac{B_1}{B_1} + \gamma\,\frac{S_1}{B_1},
\end{aligned}
$$

da cui immediatamente, facendo la differenza tra la seconda e la prima riga si ottiene

$$
\frac{f(S_1)}{B_1} - \frac{c}{B_0} = \beta\left(\frac{B_1}{B_1} - \frac{B_0}{B_0}\right) + \gamma\left(\frac{S_1}{B_1} - \frac{S_0}{B_0}\right) \qquad \textit{ossia} \quad \Delta_2 = \beta\,\Delta_0 + \gamma\Delta_1.
$$

Abbiamo quindi ritrovato il fatto che trovare la copertura perfetta equivale ad esprimere la scommessa Δ_2*, relativa al contingent claim, come una combinazione lineare delle scommesse* Δ_0 *e* Δ_1*.*

[49]Con semplici passaggi si ottiene

$$
\begin{aligned}
c = \beta\,B_0 + \gamma\,s_0 &= \frac{1}{B_0(1+r)}\,\frac{f_1(s_0\,d)\,u - f_1(s_0\,u)\,d}{u-d}\,B_0 + \frac{1}{s_0}\,\frac{f_1(s_0\,u) - f_1(s_0\,d)}{u-d}\,s_0\\
&= \frac{1}{1+r}\,\frac{f_1(s_0\,d)\,u - f_1(s_0\,u)\,d}{u-d} + \frac{f_1(s_0\,u) - f_1(s_0\,d)}{u-d}\\
&= \frac{1}{1+r}\left[\frac{f_1(s_0\,d)\,u - f_1(s_0\,u)\,d}{u-d} + \frac{f_1(s_0\,u)(1+r) - f_1(s_0\,d)(1+r)}{u-d}\right]\\
&= \frac{1}{1+r}\,\frac{f_1(s_0\,d)\,[u-(1+r)] + f_1(s_0\,u)\,[(1+r)-d]}{u-d}
\end{aligned}
$$

1.5 Il modello binomiale multiperiodale

Il modello binomiale multiperiodale cui si fa riferimento è noto anche come il **modello di Cox Ross Rubinstein**, ed è una generalizzazione dell'esempio che abbiamo visto nella precedente Sezione 1.4.

Siamo interessati al caso di un mercato che ammette due titoli, uno non rischioso e uno rischioso, i valori dei titoli cambiano in istanti discreti. Più precisamente consideriamo un mercato in cui sia possibile effettuare transazioni in un periodo di tempo discreto caratterizzato da N date, definito sullo scadenzario:

$$\underline{t} = \{t_0, t_1, ..., t_N\} \qquad \text{con} \qquad t_0 < t_1 < \cdots < t_N,$$

dove gli istanti t_i sono equispaziati ($t_i - t_{i-1} = t_1 - t_0$).

Per quanto riguarda l'evoluzione del prezzo del titolo non-rischioso, si considera il tasso di interesse costante su ciascun intervallo di tempo, con valore

$$B_k = B_{k-1}(1+r), \qquad \text{nell'intervallo di tempo } (t_k, t_{k+1}) \qquad \forall\, k = 1...N$$

da cui, procedendo per induzione ricaviamo:

$$B_k = B_0(1+r)^k.$$

A proposito dell'evoluzione del prezzo del titolo rischioso (o stock) S, invece, in analogia con il caso uniperiodale, il valore non è deterministico, ma dipende da una variabile aleatoria Z. Nel caso multiperiodale, invece di considerare una sola variabile aleatoria binomiale, consideriamo una successione di v.a. binomiali Z_n, per $n = 1, \ldots, N$. Tali variabili aleatorie assumono i d ue valori, u oppure d, e l'evoluzione del titolo rischioso è data da

$$S_k = S_{k-1}Z_k \qquad \text{nell'intervallo di tempo } (t_k, t_{k+1}) \qquad \forall\, k = 1...N.$$

Lo spazio su cui è definita la successione Z_n è $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\geq 0}, \mathbf{P})$, dove $\mathcal{F}_n \subseteq \mathcal{F}$ è la famiglia di eventi[50] costituita da:

$$\mathcal{F}_n = \bigvee_{k=1}^{n} \left\{\emptyset, A_0^{(k)}, A_1^{(k)}, \Omega\right\} \qquad \forall\, n = 1...N,$$

dove $A_0^{(k)}$ ed $A_1^{(k)}$ è la successione di insiemi di eventi definiti in modo simile al caso uniperiodale: per $k = 1, \ldots, N$.

$$Z_k(w) = \begin{cases} d \text{ se } w \in A_0^{(k)} \\ \\ u \text{ se } w \in A_1^{(k)} \end{cases}.$$

Per evitare arbitraggi (ossia guadagni privi di rischio) deve valere, come nel caso uniperiodale, $d < 1 + r < u$.

Definendo, per $i = 1, \ldots, N$:

$$\xi_i = \mathbf{1}_{\{Z_i = u\}}$$

la variabile aleatoria che indica se il prezzo del sottostante sale (in tal caso vale uno) o scende (allora vale zero), possiamo riscrivere:

$$Z_i = u^{\xi_i} d^{1-\xi_i} = \begin{cases} u \Leftrightarrow \xi_i = 1 \\ \\ d \Leftrightarrow \xi_i = 0 \end{cases}$$

da cui $\forall\, i = 1, \ldots, N$:

$$S_i = S_{i-1}Z_i = S_{i-1}u^{\xi_i}d^{1-\xi_i}.$$

Iterando questa fattorizzazione ricaviamo la formula del prezzo:

$$S_i = S_0 Z_1 Z_2 \cdots Z_i = S_0 \prod_{k=1}^{i} Z_k = S_0 \prod_{k=1}^{i} u^{\xi_k} d^{1-\xi_k} \tag{1.28}$$

[50] Si ricorda che $\vee_\alpha \mathcal{K}_\alpha$ è la sigma-algebra generata dall'unione $\cup_\alpha \mathcal{K}_\alpha$. Per la definizione di sigma-algebra, si rimanda al capitolo sui richiami di probabilità. In questo caso la famiglia di eventi $\mathcal{F}_n$ è la famiglia degli eventi del tipo unioni finite di

$$\bigcap_{k=1}^{n} B_k, \qquad \text{al variare di } B_k \in \left\{A_0^{(k)}, A_1^{(k)}\right\}$$

più l'insieme vuoto.

Anche le variabili aleatorie ξ_i sono definite sullo spazio $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\geq 0}, \mathbf{P})$ e verificano

$$\mathbf{P}(\xi_1 = x_1,\, \xi_2 = x_2,\, \cdots, \xi_N = x_N) > 0, \quad \text{per ogni } x_i \in \{0,1\}.$$

Questa condizione si verifica se, ad esempio, rispetto alla misura $\mathbf{P}$ le variabili aleatorie Z_i sono indipendenti con $\mathbf{P}(Z_i = u) = p$ e $\mathbf{P}(Z_i = d) = 1 - p =: q$, con $0 < p < 1$. In tale caso ovviamente anche le variabili ξ_i sono indipendenti e sia ha

$$\begin{aligned} &\mathbf{P}(\xi_i = 1) = \mathbf{P}(Z_i = u) = p \\ &\mathbf{P}(\xi_i = 0) = \mathbf{P}(Z_i = d) = 1 - p =: q \quad \forall\, i = 1, ..., N. \end{aligned}$$

Possiamo considerare il numero di successi (numero di volte che il prezzo sale) fino all'istante $k - esimo$ definendo la variabile:

$$H_k = \sum_{i=1}^{k} \xi_i = \sum_{i=1}^{k} \mathbf{1}_{\{Z_i = u\}}$$

Allora la (1.28) può essere riscritta $\forall\, k = 1...N$, in questo modo:

$$S_k = S_0 u^{\sum_{i=1}^k \xi_i} d^{k - \sum_{i=1}^k \xi_i} = S_0 u^{H_k} d^{k - H_k} = S_0 \left(\frac{u}{d}\right)^{H_k} d^k \tag{1.29}$$

Attraverso il teorema dell'arbitraggio vogliamo dimostrare che esiste unica una probabilità martingala equivalente $\tilde{\mathbf{P}}$, definita sullo spazio $(\Omega, \mathcal{F})$, ossia per la quale le scommesse elementari hanno media nulla.

Come vedremo, tale probabilità è caratterizzata dal fatto che $\tilde{\mathbf{P}}$ rende le variabili aleatorie ξ_i indipendenti ed identicamente distribuite con probabilità

$$\tilde{p} = \frac{1 + r - d}{u - d}. \tag{1.30}$$

A questo scopo definiamo innanzitutto le scommesse elementari ammissibili[51] all'istante $t = i - 1$.

Possiamo fare solo scommesse del tipo[52]: $\forall\, i = 1, \ldots, N$ e $x_j \in \{0,1\}$, $j = 1, \ldots i - 1$,

$$\Delta_{i, x_1, \cdots, x_{i-1}} = \left(\frac{S_i(w)}{(1+r)^i} - \frac{S_{i-1}(w)}{(1+r)^{i-1}}\right) \underbrace{\mathbf{1}_{\{\xi_1 = x_1, \xi_2 = x_2, \cdots, \xi_{i-1} = x_{i-1}\}}}_{\text{si tiene conto di tutti gli eventi precedenti all'}i\text{-esimo}}, \tag{1.31}$$

ovvero, con notazioni simili al caso uniperiodale, posto $\widetilde{S}_i(w) = \frac{S_i(w)}{B_0 (1+r)^i}$,

$$= \left(\widetilde{S}_i(w) - \widetilde{S}_{i-1}(w)\right) \mathbf{1}_{\{\xi_1 = x_1, \xi_2 = x_2, \cdots, \xi_{i-1} = x_{i-1}\}}. \tag{1.32}$$

con la convenzione che quando $i = 1$, il termine $\mathbf{1}_{\{\xi_1 = x_1, \xi_2 = x_2, \cdots, \xi_{i-1} = x_{i-1}\}}$ va considerato identicamente uguale a 1.

Osservazione 1.9. *La restrizione alle scommesse di questo tipo è molto importante: significa che posso effettuare la scommessa del tipo compro l'azione al tempo $i - 1$ e la rivendo al tempo i solo sulla base dell'andamento dei prezzi, nel periodo di tempo fino all'istante $i - 1$, ma non posso utilizzare informazioni relative al futuro. Dal punto di vista economico equivale a richiedere che non sia possibile fare* **insider trading***:*

[51] Nel caso uniperiodale l'unica scommessa possibile è

$$\Delta_1 = \tilde{S}_1 - \tilde{S}_0 = \; = \frac{S_0 Z}{B_0(1+r)} - \frac{S_0}{B_0},$$

quindi

$$\tilde{\mathbf{E}}(\Delta_1) = 0 \Leftrightarrow \tilde{\mathbf{E}}(\tilde{S}_1) = \tilde{S}_0 = \frac{S_0}{B_0},$$

quindi deve accadere che $\tilde{\mathbf{E}}(Z) = (1 + r)$, ossia che

$$u\tilde{\mathbf{P}}(Z = u) + d\tilde{\mathbf{P}}(Z = d) = 1 + r$$

da cui

$$u\tilde{p} + d(1 - \tilde{p}) = 1 + r \Leftrightarrow \tilde{p} = \frac{1 + r - d}{u - d}$$

[52] Assumiamo per convenienza e senza perdita di generalità $B_0 = 1$

a volte i dirigenti di una azienda sanno in anticipo alcuni eventi che accadranno nel prossimo futuro (ad esempio un crack finanziario o una fusione) e sanno che tali eventi influenzeranno negativamente o positivamente sull'andamento del prezzo di mercato delle azione della loro azienda.
In tali casi i dirigenti non possono speculare sulle azioni della azienda da loro diretta, perché agirebbero in regime di insider trading.

Osservazione 1.10. *Un'altra osservazione interessante riguarda invece la globalità delle scommesse ammissibili che sono le combinazioni lineari di quelle elementari, ossia del tipo*

$$y_1\Delta_1 + \sum_{i=2}^{N} \sum_{x_1,x_2,\cdots,x_{i-1}} y_i(x_1,x_2,\cdots,x_{i-1})\Delta_{i,x_1,\cdots,x_{i-1}}.$$

È interessante notare che tali scommesse si possono riscrivere come segue:

$$\begin{aligned}
&= y_1\left(\frac{S_0Z_1}{B_0(1+r)} - \frac{S_0}{B_0}\right)\\
&+\sum_{i=2}^{N} \sum_{x_1,x_2,\cdots,x_{i-1}} y_i(x_1,x_2,\cdots,x_{i-1})\frac{S_{i-1}}{B_0(1+r)^{i-1}}\left(\frac{u^{\xi_i}d^{1-\xi_i}}{1+r}-1\right)\mathbf{1}_{\{\xi_1=x_1,\xi_2=x_2,\cdots,\xi_{i-1}=x_{i-1}\}}\\
&= y_1\left(\frac{S_0Z_1}{B_0(1+r)} - \frac{S_0}{B_0}\right)\\
&+\sum_{i=2}^{N} \frac{S_{i-1}}{B_0(1+r)^{i-1}}\left(\frac{u^{\xi_i}d^{1-\xi_i}}{1+r}-1\right)\sum_{x_1,x_2,\cdots,x_{i-1}} y_i(x_1,x_2,\cdots,x_{i-1})\mathbf{1}_{\{\xi_1=x_1,\xi_2=x_2,\cdots,\xi_{i-1}=x_{i-1}\}}\\
&= y_1\left(\frac{S_0Z_1}{B_0(1+r)} - \frac{S_0}{B_0}\right) + \sum_{i=2}^{N} \frac{S_{i-1}}{B_0(1+r)^{i-1}}\left(\frac{u^{\xi_i}d^{1-\xi_i}}{1+r}-1\right) y_i(\xi_1,\xi_2,\cdots,\xi_{i-1})
\end{aligned}$$

ovvero, tenendo conto che $\widetilde{S}_{i-1} = \frac{S_{i-1}}{B_0(1+r)^{i-1}}$ *e che* $\widetilde{S}_i = \frac{S_{i-1}\,u^{\xi_i}d^{1-\xi_i}}{B_0(1+r)^i}$,

$$= y_1\left(\widetilde{S}_1 - \widetilde{S}_0\right) + \sum_{i=2}^{N}\left(\widetilde{S}_i - \widetilde{S}_{i-1}\right) y_i(\xi_1,\xi_2,\cdots,\xi_{i-1}).$$

Posto $\gamma_1 := y_1$, *e*

$$\gamma_i(\omega) := y_i(\xi_1(\omega),\xi_2(\omega),\cdots,\xi_{i-1}(\omega)), \quad i \geq 2$$

il processo stocastico (o successione di variabili aleatorie) $(\gamma_i)_i$ *si dice predicibile, perché, per ogni* i, γ_i *è funzione delle variabili aleatorie* ξ_j *per* $j \leq i-1$, *che, al tempo* i *saranno tutte note (osservate). Con tale notazione si ottiene che le scommesse ammissibili sono del tipo*

$$\sum_{i=1}^{N} \gamma_i\left(\widetilde{S}_i - \widetilde{S}_{i-1}\right),$$

e quindi se il capitale iniziale al tempo 0 *è* X_0 *(e quindi* $\widetilde{X}_0 = \frac{X_0}{B_0}$*) e la strategia seguita è la precedente, allora, al tempo* N, *il capitale finale, scontato, diventa*

$$\widetilde{X}_N = \widetilde{X}_0 + \sum_{i=1}^{N} \gamma_i\left(\widetilde{S}_i - \widetilde{S}_{i-1}\right).$$

Vedremo ne seguito che questo tipo di scommesse si possono considerare integrali stocastici a tempo discreto.

Vogliamo dimostrare, per il teorema dell'arbitraggio, che $\exists!\,\tilde{\mathbf{P}}$ per cui:

$$\tilde{\mathbf{E}}(\Delta_{i,x_1,\cdots,x_{i-1}}) = 0 \tag{1.33}$$

Possiamo riscrivere la (1.31) in questo modo:

$$\Delta_{i,x_1,\cdots,x_{i-1}} = \frac{S_{i-1}}{(1+r)^{i-1}}\left(\frac{u^{\xi_i}d^{1-\xi_i}}{1+r}-1\right)\mathbf{1}_{\{\xi_1=x_1,\xi_2=x_2,\cdots,\xi_{i-1}=x_{i-1}\}}$$

che equivale, considerando la prima uguaglianza nella (1.29) ad avere:

$$\begin{aligned}\Delta_{i,x_1,\cdots,x_{i-1}} &= \frac{S_0 u^{\sum_{k=1}^{i-1}\xi_k} d^{i-1-\sum_{k=1}^{i-1}\xi_k}}{(1+r)^{i-1}}\mathbf{1}_{\{\xi_1=x_1,\xi_2=x_2,\cdots,\xi_{i-1}=x_{i-1}\}}\left(\frac{u^{\xi_i}d^{1-\xi_i}}{1+r}-1\right)\\ &= \frac{S_0 u^{\sum_{k=1}^{i-1}x_k} d^{i-1-\sum_{k=1}^{i-1}x_k}}{(1+r)^{i-1}}\mathbf{1}_{\{\xi_1=x_1,\xi_2=x_2,\cdots,\xi_{i-1}=x_{i-1}\}}\left(\frac{u^{\xi_i}d^{1-\xi_i}}{1+r}-1\right)\end{aligned}$$

Riscritta in questo modo è evidente che conosciamo il valore di $\frac{S_{i-1}}{(1+r)^{i-1}}$, ed è diverso da zero, quindi richiedere la (1.33) è equivalente a richiedere che:

$$\tilde{\mathbf{E}}\left[\left(\frac{u^{\xi_i}d^{1-\xi_i}}{1+r}-1\right)\mathbf{1}_{\{\xi_1=x_1,\xi_2=x_2,\cdots,\xi_{i-1}=x_{i-1}\}}\right] = 0$$

Svolgendo i calcoli[53] ricaviamo

$$\tilde{\mathbf{P}}\left(\xi_i = 1|\xi_1 = x_1,\cdots,\xi_{i-1} = x_{i-1}\right) = \frac{1+r-d}{u-d}.$$

[53]Si osservi che

$$\begin{aligned}&\tilde{\mathbf{E}}\left[\left(\frac{u^{\xi_i}d^{1-\xi_i}}{1+r}-1\right)\mathbf{1}_{\{\xi_1=x_1,\xi_2=x_2,\cdots,\xi_{i-1}=x_{i-1}\}}\right] = \tilde{\mathbf{E}}\left[(\mathbf{1}_{\{\xi_i=1\}}+\mathbf{1}_{\{\xi_i=0\}})\left(\frac{u^{\xi_i}d^{1-\xi_i}}{1+r}-1\right)\mathbf{1}_{\{\xi_1=\mathbf{x_1},\xi_2=\mathbf{x_2},\cdots,\xi_{i-1}=\mathbf{x_{i-1}}\}}\right]\\ &= \tilde{\mathbf{E}}\left[\left(\frac{u}{1+r}-1\right)\mathbf{1}_{\{\xi_1=x_1,\xi_2=x_2,\cdots,\xi_{i-1}=x_{i-1},\xi_i=1\}} + \left(\frac{d}{1+r}-1\right)\mathbf{1}_{\{\xi_1=x_1,\xi_2=x_2,\cdots,\xi_{i-1}=x_{i-1},\xi_i=0\}}\right]\\ &= \left(\frac{u}{1+r}-1\right)\tilde{\mathbf{P}}\left(\xi_1=x_1,\xi_2=x_2,\cdots,\xi_{i-1}=x_{i-1},\xi_i=1\right) + \left(\frac{d}{1+r}-1\right)\tilde{\mathbf{P}}\left(\xi_1=x_1,\xi_2=x_2,\cdots,\xi_{i-1}=x_{i-1},\xi_i=0\right)\\ &= \left(\frac{u}{1+r}-1\right)\tilde{\mathbf{P}}\left(\xi_1=x_1,\cdots,\xi_{i-1}=x_{i-1}\right)\tilde{\mathbf{P}}\left(\xi_i=1|\xi_1=x_1,\cdots,\xi_{i-1}=x_{i-1}\right)\\ &+ \left(\frac{d}{1+r}-1\right)\tilde{\mathbf{P}}\left(\xi_1=x_1,\cdots,\xi_{i-1}=x_{i-1}\right)\tilde{\mathbf{P}}\left(\xi_i=0|\xi_1=x_1,\cdots,\xi_{i-1}=x_{i-1}\right).\end{aligned}$$

Di conseguenza si ottiene:

$$\tilde{\mathbf{E}}\left(\Delta_{i,x_1,\cdots,x_{i-1}}\right) = 0$$

se e solo se

$$\begin{aligned}&\left(\frac{u}{1+r}-1\right)\tilde{\mathbf{P}}(\xi_i=1|\xi_1=x_1,\cdots,\xi_{i-1}=x_{i-1})+\\ &+\left(\frac{d}{1+r}-1\right)(1-\tilde{\mathbf{P}}\left(\xi_i=1|\xi_1=x_1,\cdots,\xi_{i-1}=x_{i-1}\right)) = 0\end{aligned}$$

e raccogliendo a fattor comune si ottiene

$$\tilde{\mathbf{P}}\left(\xi_i=1|\xi_1=x_1,\cdots,\xi_{i-1}=x_{i-1}\right)\left[\left(\frac{u}{1+r}-1\right)-\left(\frac{d}{1+r}-1\right)\right]+\left(\frac{d}{1+r}-1\right)=0$$

da cui

$$\tilde{\mathbf{P}}\left(\xi_i=1|\xi_1=x_1,\cdots,\xi_{i-1}=x_{i-1}\right)\left(\frac{u-d}{1+r}\right) = \frac{1+r-d}{1+r}$$

quindi

$$\tilde{\mathbf{P}}\left(\xi_i=1|\xi_1=x_1,\cdots,\xi_{i-1}=x_{i-1}\right) = \left(\frac{1+r-d}{u-d}\right) := \tilde{p}$$

Per dimostrare effettivamente che le variabili aleatorie ξ_i sono indipendenti ed identicamente distribuite con probabilità $\tilde{p} = \dfrac{1+r-d}{u-d}$ sullo spazio $(\Omega, F, \tilde{\mathbf{P}})$, dobbiamo far vedere che $\forall\, i = 1...N$

$$\tilde{\mathbf{P}}\left(\xi_i = 1\right) = \tilde{\mathbf{P}}\left(\xi_i = 1 | \xi_1 = x_1, \cdots, \xi_{i-1} = x_{i-1}\right),$$

ovvero che il verificarsi dell'evento $\{\xi_i = 1\}$ non è influenzato dai valori assunti dalla v.a. ξ_k fino all'istante $i - esimo$. Infatti, con semplici passaggi, possiamo vedere che:

$$\begin{aligned}
\tilde{\mathbf{P}}(\xi_i = 1) &= \sum_{x_1 \cdots x_{i-1}} \tilde{\mathbf{P}}(\xi_i = 1 | \xi_1 = x_1 \cdots \xi_{i-1} = x_{i-1}) \tilde{\mathbf{P}}(\xi_1 = x_1 \cdots \xi_{i-1} = x_{i-1}) \\
&= \tilde{p} \sum_{x_1 \cdots x_{i-1}} \tilde{\mathbf{P}}(\xi_1 = x_1 \cdots \xi_{i-1} = x_{i-1}) = \tilde{p}.
\end{aligned}$$

L'indipendenza delle variabili ξ_i è poi ovvia: infatti l'indipendenza dipende dal fatto che se vale

$$\tilde{\mathbf{P}}\left(\xi_i = 1\right) = \tilde{\mathbf{P}}\left(\xi_i = 1 | \xi_1 = x_1, \cdots, \xi_{i-1} = x_{i-1}\right),$$

allora, essendo ξ_i una variabile aleatoria che assume solo due valori, si ha anche

$$\tilde{\mathbf{P}}\left(\xi_i = 0\right) = \tilde{\mathbf{P}}\left(\xi_i = 0 | \xi_1 = x_1, \cdots, \xi_{i-1} = x_{i-1}\right),$$

ossia, qualunque siano $x_1,\, x_2,\, \cdots,\, x_i, \cdots, x_N \in \{0, 1\}$

$$\tilde{\mathbf{P}}\left(\xi_i = x_i\right) = \tilde{\mathbf{P}}\left(\xi_i = 0 | \xi_1 = x_1, \cdots, \xi_{i-1} = x_{i-1}\right).$$

L'indipendenza, cioè il fatto che, comunque siano scelti $x_1,\, x_2,\, \cdots, x_N \in \{0, 1\}$,

$$\tilde{\mathbf{P}}\left(\xi_1 = x_1, \cdots, \xi_{N-1} = x_{N-1},\ \xi_N = x_N\right) = \prod_{j=1}^{N} \tilde{\mathbf{P}}\left(\xi_j = x_j\right),$$

segue immediatamente per induzione, dalla formula delle probabilità composte:

$$\begin{aligned}
\tilde{\mathbf{P}}\left(\xi_1 = x_1, \cdots, \xi_{i-1} = x_{i-1},\ \xi_i = x_i\right) &= \tilde{\mathbf{P}}\left(\xi_i = x_i | \xi_1 = x_1, \cdots, \xi_{i-1} = x_{i-1}\right) \tilde{\mathbf{P}}\left(\xi_1 = x_1, \cdots, \xi_{i-1} = x_{i-1}\right) \\
&= \tilde{\mathbf{P}}\left(\xi_i = x_i\right) \tilde{\mathbf{P}}\left(\xi_1 = x_1, \cdots, \xi_{i-1} = x_{i-1}\right) \\
&= \tilde{\mathbf{P}}\left(\xi_i = x_i\right)\, \tilde{\mathbf{P}}\left(\xi_{i-1} = x_{i-1} | \xi_1 = x_1, \cdots, \xi_{i-2} = x_{i-2}\right)\, \tilde{\mathbf{P}}\left(\xi_1 = x_1, \cdots, \xi_{i-2} = x_{i-2}\right) \\
&= \tilde{\mathbf{P}}\left(\xi_i = x_i\right)\, \tilde{\mathbf{P}}\left(\xi_{i-1} = x_{i-1}\right)\, \tilde{\mathbf{P}}\left(\xi_1 = x_1, \cdots, \xi_{i-2} = x_{i-2}\right) = \cdots = \prod_{j=1}^{i} \tilde{\mathbf{P}}\left(\xi_j = x_j\right).
\end{aligned}$$

Quindi la misura di probabilità $\tilde{\mathbf{P}}$ rende ξ_i variabili aleatorie indipendenti, identicamente distribuite, con:

$$\tilde{\mathbf{P}}(\xi_i = 1) = \tilde{p} = \frac{1+r-d}{u-d}$$

e di conseguenza, rispetto alla misura di probabilità $\tilde{\mathbf{P}}$, si ha che **

$$H_N = \sum_{i=1}^{N} \xi_i \sim Bin(N, \tilde{p}), \quad \Leftrightarrow \quad \tilde{\mathbf{P}}\left(H_N = h\right) = \binom{N}{h} \tilde{p}^h (1-\tilde{p})^{N-h}, \quad h = 0, \cdots, N$$

Dopo aver dimostrato che effettivamente il prezzo del sottostante puó essere scritto attraverso la (1.29) con H_N variabile aleatoria con distribuzione binomiale di parametri N e $\tilde{\mathbf{P}}$, possiamo trovare un'espressione per il prezzo di un contratto derivato con pay-off terminale f_N.

Il prezzo di esercizio è dato da:

$$C(f_N, \mathbf{P}) = B_0\, \tilde{\mathbf{E}}\left[\frac{f_N}{B_N}\right] = \frac{1}{(1+r)^N} \tilde{\mathbf{E}}\left[f_N\right]$$

dove $\frac{1}{(1+r)^N}$ rappresenta il fattore di sconto relativo al periodo $[t_0, t_N]$ e $\tilde{\mathbf{E}}[f_N]$ è l'aspettazione calcolata rispetto alla probabilità $\tilde{\mathbf{P}}$ del pay-off f_N.

In particolare per un'opzione call di tipo europeo con prezzo di esercizio K e tempo di esercizio N si ottiene:

$$C_{call}(K, \mathbf{P}) = \frac{1}{(1+r)^N} \tilde{\mathbf{E}}\left[(S_N - K)^+\right]$$

Visto che possiamo scrivere il valore del sottostante all'istante t_N come in (1.29) ed il valore atteso $\tilde{\mathbf{E}}[(S_N - K)^+]$ dipende solo dalla distribuzione di H_N, come si vede dall'espressione:

$$\tilde{\mathbf{E}}\left[(S_N - K)^+\right] = \sum_{h=0}^{N} \tilde{\mathbf{P}}(H_N = h) \left(S_0 u^h d^{N-h} - K **\right)^+,$$

se definiamo[54]:

$$x_0 := \min\left\{x \in \mathbb{R} : \; S_0 d^{N-x} u^x - K > 0\right\} **$$

ovvero

$$S_0 d^{N-x_0} u^{x_0} - K = 0 \qquad \Leftrightarrow \qquad \left(\frac{u}{d}\right)^{x_0} = \frac{K}{S_0\, d^N} \qquad \Leftrightarrow \qquad x_0 = \frac{\ln\left(\frac{K}{S_0 d^N}\right)}{\ln\left(\frac{u}{d}\right)} ****$$

e

$$h_0 := \min\left\{j \in N : S_0\, d^{N-j}\, u^j - K > 0\right\} **$$

$\forall\, h \geq x_0$ (o equivalentemente $\forall\, h \geq h_0$) ovviamente si ha $S_N > K$, quindi:

$$\begin{aligned}
C_{call}(K, \mathbf{P}) &= \frac{\tilde{\mathbf{E}}[(S_N - K)^+]}{(1+r)^N} = \frac{\tilde{\mathbf{E}}\left[(S_0\, u^{H_N}\, d^{N-H_N} - K)^+\right]}{(1+r)^N} \\
&= \sum_{0 \leq h \leq N} \tilde{\mathbf{P}}(H_N = h) \frac{\left(S_0\, u^h\, d^{N-h} - K\right)^+}{(1+r)^N} = \sum_{h_0 \leq h \leq N} \tilde{\mathbf{P}}(H_N = h) \frac{S_0\, u^h\, d^{N-h} - K}{(1+r)^N} \\
&= S_0 \sum_{h_0 \leq h \leq N} \binom{N}{h} \tilde{p}^h (1-\tilde{p})^{N-h} \left(\frac{u}{1+r}\right)^h \left(\frac{d}{1+r}\right)^{N-h} \\
&- \frac{K}{(1+r)^N} \sum_{h_0 \leq h \leq N} \binom{N}{h} \tilde{p}^{\,h} (1-\tilde{p})^{N-h} \\
& S_0 \sum_{h_0 \leq h \leq N} \binom{N}{h} \left(\frac{u}{1+r}\tilde{p}\right)^h \left(\frac{d}{1+r}(1-\tilde{p})\right)^{N-h} - \frac{K}{(1+r)^N} \sum_{h_0 \leq h \leq N} \binom{N}{h} \tilde{p}^{\,h} (1-\tilde{p})^{N-h}
\end{aligned}$$

Equivalentemente si può eseguire la somma sull'insieme dei valori h interi e tali che $h > x_0$. In altre parole

$$C_{call}(K, \mathbf{P}) = S_0\, \overline{F}_{N,p^*}(x_0) - \frac{K}{(1+r)^N} \overline{F}_{N,\tilde{p}}(x_0) \tag{1.34}$$

dove $\tilde{p} \in (0,1)$ è definita come in (1.30),

$$p^* = \frac{u}{1+r}\tilde{p}, \qquad \text{con} \quad p^* \in (0,1), \;\; \text{(come è facile verificare)},$$

---

[54] **Usando la simbologia $d = 1 + a$ e $u = 1 + b$,

$$x_0 := \min\left\{x \in \mathbb{R} : \; S_0(1+a)^{N-x}(1+b)^x - K > 0\right\},$$

ovvero

$$S_0(1+a)^{N-x_0}(1+b)^{x_0} - K = 0 \qquad \Leftrightarrow \qquad x_0 = \ln\left(\frac{K}{S_0(1+a)^N}\right) \ln\left(\frac{1+b}{1+a}\right).$$

Per cui

$$h_0 := \min\left\{j \in N : S_0(1+a)^{N-j}(1+b)^j - K > 0\right\}.$$

$$F_{n,p}(x) = \sum_{0 \le k \le x} \binom{n}{h} p^h \, (1-p)^{n-k} {**}$$

è la funzione di distribuzione di una variabile binomiale e infine

$$\overline{F}_{n,p}(x) = 1 - F_{n,p}(x) = \sum_{k>x} \binom{n}{h} p^h \, (1-p)^{n-k} {**}$$

è la funzione di sopravvivenza di una binomiale calcolata nel punto x.

La (1.34) definisce quindi il prezzo di un contratto di opzione call nel caso in cui consideriamo un modello di mercato binomiale multiperiodale a tempo discreto (in cui cioè siano possibili transazioni solo nelle date definite dallo scadenzario $\underline{\mathbf{t}}$).

Osservazione 1.11. *Possiamo facilmente estendere i risultati ottenuti per il contratto di opzione call al caso di un contratto di opzione put, infatti dall'uguaglianza:*

$$(K - S_N)^+ = (S_N - K)^+ - S_N + K$$

segue che il prezzo per un'opzione put è dato da:

$$\begin{aligned} C_{put}(K, \mathbf{P}) \;&=\; \tfrac{1}{(1+r)^N} \, \tilde{\mathbf{E}} \left[(K - S_N)^+ \right] = \\ &= C_{call}(K, \mathbf{P}) - \; \tilde{\mathbf{E}} \left[\tfrac{S_N}{(1+r)^N} \right] + \tfrac{K}{(1+r)^N} \end{aligned}$$

ed essendo[55] $\tilde{\mathbf{E}} \left[\dfrac{S_N}{(1+r)^N} \right] = S_0$ *si ottiene la seguente relazione:*

$$C_{put}(K, \mathbf{P}) = C_{call}(K, \mathbf{P}) - S_0 + \frac{K}{(1+r)^N}$$

che viene detta formula di parità per le opzioni call-put.

Terminiamo questo paragrafo con il dimostrare che il calcolo del prezzo per le call europee permette di effettuare il calcolo del prezzo per contratti con contingent claim del tipo $f(S_N)$ mediante una semplice operazione di integrazione, per f in un'ampia classe di funzioni.

Osservazione 1.12. *Sia* $f = f(x)$, *con* $x \geqslant 0$, *una funzione non negativa, sia* $f_N = f(S_N)$ *il pay-off e sia, come al solito,* $C(f_N, \mathbb{P}) = B_0 \, \widetilde{\mathbb{E}} \Big[\frac{f(S_N)}{B_N} \Big]$ *il prezzo razionale corrispondente. È possibile determinare il valore del prezzo di un'opzione generica di questo tipo usando il prezzo razionale di un'opzione call.*
Si assuma f *derivabile con derivata*

$$f'(x) = f'(0) + \int_{(0,x]} \mu(dy),$$

[55] Il fatto che $\tilde{\mathbf{E}} \left[\dfrac{S_N}{(1+r)^N} \right] = S_0$ si può dedurre facilmente osservando che

$$\frac{S_N}{B_0 \, (1+r)^N} - \frac{S_0}{B_0} = \frac{1}{B_0} \sum_{k=1}^{N} \left(\frac{S_k}{(1+r)^k} - \frac{S_{k-1}}{(1+r)^{k-1}} \right)$$

e che per le scommesse $\frac{S_k}{(1+r)^k} - \frac{S_{k-1}}{(1+r)^{k-1}}$ hanno media nulla sotto $\widetilde{\mathbb{P}}$, in quanto

$$\frac{S_k}{(1+r)^k} - \frac{S_{k-1}}{(1+r)^{k-1}} = \sum_{x_1, \cdots, x_{k-1}} \Delta_{k, x_1, \cdots, x_{k-1}} .$$

dove $\mu = \mu(dy)$ è una misura finita[56]*, non necessariamente positiva, su $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$. Si noti che se f è derivabile due volte si ha la precedente rappresentazione con $\mu(dy) = f''(y)dy$.*
Allora è chiaro che

$$f(x) = f(0) + xf'(0) + \int_{(0,x]} (x-y)^+ \mu(dy) = f(0) + xf'(0) + \int_{(0,\infty)} (x-y)^+ \mu(dy), \tag{1.35}$$

quindi ponendo $x = S_N$ e cambiando notazione nell'integrale

$$f_N = f(S_N) = f(0) + S_N f'(0) + \int_{(0,\infty)} (S_N - y)^+ \mu(dy) \quad (\mathbb{P}\ q.c.).$$

Se ora si attualizzano i valori

$$\widetilde{f}_N = \frac{f(S_N)}{B_N} = \frac{f(0)}{B_N} + \frac{S_N}{B_N} f'(0) + \int_{(0,\infty)} \frac{(S_N - y)^+}{B_N} \mu(dy) \quad (\mathbb{P}\ q.c.).$$

e si considera la media rispetto alla misura $\widetilde{\mathbb{P}}$, essendo $\widetilde{\mathbb{E}}\Big[\frac{S_N}{B_N}\Big] = \widetilde{\mathbb{E}}\Big[\frac{S_0}{B_0}\Big] = \frac{S_0}{B_0}$, e tenendo conto che $B_N = B_0(1+r)^N$ è deterministico, si ottiene

$$\widetilde{\mathbb{E}}\Big[\frac{f_N}{B_N}\Big] = \frac{f(0)}{B_N} + \frac{S_0}{B_0} f'(0) + \int_{(0,\infty)} \widetilde{\mathbb{E}}\left[\frac{(S_N - y)^+}{B_N}\right] \mu(dy),$$

da cui segue che

$$C(f_N, \mathbb{P}) = B_0 \widetilde{\mathbb{E}}\Big[\frac{f_N}{B_N}\Big] = \frac{f(0)}{(1+r)^N} + S_0 f'(0) + \int_{(0,\infty)} C_{call}(y, \mathbb{P}) \mu(dy). \tag{1.36}$$

Si osservi che se $f_N = f(S_N) = (S_N - K_*)^+$, $K_* > 0$, allora $f(0) = f'(0) = 0$, $\mu(dy)$ è concentrata nel punto K_*, cioè $\mu_*(dy) = \delta_{\{K_*\}}(dy)$, e ciò implica, come deve essere, $C(f_N, \mathbb{P}) = C_{call}(K_*, \mathbb{P})$.

---

[56] Se il lettore trova difficoltà a comprendere questa espressione può limitarsi al caso di funzioni derivabili due volte, con derivate seconde continue, e sostituire a $\mu(dy)$ l'espressione $f''(y)dy$.

Per ottenere poi la formula (1.35) basta considerare che allora

$$f'(z) = f'(0) + \int_0^z f''(y)dy,$$

e quindi

$$f(x) = f(0) + \int_0^x f'(z)dz = f(0) + \int_0^x \left(f'(0) + \int_0^z f''(y)dy\right) dz = f(0) + xf'(0) + \int_0^x \left(\int_0^z f''(y)dy\right) dz$$

e scambiando l'ordine integrali per cui $0 < z \leq x$ e $0 < y < z$ diviene $0 < y < x$ e $y < z \leq x$

$$\begin{aligned} &= f(0) + xf'(0) + \int_0^x \left(\int_y^x f''(y)dz\right) dy = f(0) + xf'(0) + \int_0^x f''(y) \left(\int_y^x dz\right) dy \\ &= f(0) + xf'(0) + \int_0^x f''(y)\,(x-y)\,dy \end{aligned}$$

Tuttavia, ad esempio, rientrano nella categoria sopra citata anche funzioni del tipo

$$f(x) = (x - K_1)^+ - (x - K_2)^+, \qquad 0 < K_1 < K_2.$$

Ovviamente per tale funzione non c'è bisogno di trovare quale sia la misura μ per ottenere il prezzo (basta usare la linearità del valore atteso per capire che basta fare la differenza tra il prezzo della call con prezzo di strike K_1 e quello della call con prezzo di strike K_2. Comunque, posto $g_i(x) = (x - K_i)^+$ si ha $g_i'(x) = 0$ per $x < K_1$ e $g_i'(x) = 1$ per $x > K_1$, ossia $g_i'(x)$ coincide con la funzione di Heaviside $H(x - K_i)$, che a sua volta corrisponde all'integrale della misura di Dirac $\delta_{K_i}(dx)$, che e' una misura concentrata in K_i con peso 1 in K_i.

1.6 Appendice: Alberi binomiali e modello CRR***

In questa Appendice vogliamo discutere la relazione tra i risultati ottenuti in questo capitolo, e la costruzione degli alberi binomiali, così come si trova nei testi standard come il testo di Hull [10] o di Luenberger [12], ad esempio, ma anche nel testo di Björk [5]. Il metodo si applica nel caso in cui l'opzione sia di tipo europeo e *plain vanilla*, cioè il caso in cui il terminal pay-off $f_N = f(S_N)$, con f funzione deterministica.

Per iniziare l'albero permette di scrivere l'evoluzione del prezzo dell'azione in modo che l'albero rappresenta lo spazio dei possibili eventi (si veda la Figura 1.2).

La martingala[57] M_n è data dal valore atteso condizionato, sotto la misura martingala equivalente $\widetilde{\mathbb{P}}$, del terminal pay-off attualizzato

$$\tilde{f}_N = \frac{f(S_N)}{B_N},$$

rispetto a $\mathcal{F}_n$, ossia da

$$M_N = \frac{f(S_N)}{B_N} \quad \text{e} \quad M_n = \widetilde{\mathbb{E}}\left[M_N \middle| \mathcal{F}_n\right].$$

Quindi, nel caso delle opzioni plain vanilla, il valore di M_N si ottiene semplicemente calcolando la funzione f negli $N+1$ valori possibili per S_N e dividendo per $B_N = B_0(1+r)^N$ (si veda la Figura 1.4). Essendo B_N un valore deterministico, si ha che

$$M_N = \hat{g}_N(S_N),$$

dove **$\hat{g}_N$ è una funzione deterministica **($\hat{g}_N(s) = \frac{f(s)}{B_0(1+r)^N}$). Sempre nel caso delle opzioni plain vanilla, anche M_n gode di una proprietà analoga, ovvero della proprietà che

$$M_n = \hat{g}_n(S_n),$$

per un'opportuna funzione deterministica $\hat{g}_n$. Questa proprietà si dimostra facilmente per induzione: per $n = N-1$

[57]**Per capire questa parte bisogna conoscere il concetto di martingala e di valore atteso condizionale: il lettore che non lo conosce ancora nesuno di questi concetti può procedere come segue:nel seguito $\mathcal{F}_n$ è la sigma algebra generata da $Z_1, \cdots, Z_n$ (o equivalentemente da $S_1, \cdots, S_n$ o ancora da $\xi_1, \cdots, xi_n$). Dire che una variabile aleatoria X_n è $\mathcal{F}_n$ misurabile, vuole dire che esiste una funzione $g_n(z_1, \cdots, z_n)$ tale che $X_n(\omega) = g_n(Z_1(\omega), \cdots, Z_n(\omega))$. Data una variabile aleatoria X

$$\mathbb{E}[X|\mathcal{F}_n]$$

è un modo rapido per scrivere

$$\mathbb{E}[X|Z_1, \cdots, Z_n]$$

Si tratta di una variabile aleatoria che è una funzione di $Z_1, \cdots, Z_n$, ossia

$$\mathbb{E}[X|Z_1, \cdots, Z_n] = \psi(Z_1, \cdots, Z_n)$$

con

$$\psi(z_1, \cdots, z_n) = \sum_i x_i \mathbb{P}(X = x_i | \{Z_1 = z_1, \cdots, Z_n = z_n\})$$

se le variabili $X, Z_1, \cdots, Z_n$ sono variabili discrete, mentre,

$$\psi(z_1, \cdots, z_n) = \int_{-\infty}^{+\infty} x \frac{p_{X,Z_1\cdots,Z_n}(x, z_1, \cdots, z_n)}{p_{Z_1\cdots,Z_n}(z_1, \cdots, z_n)}\, dx$$

se $X, Z_1, \cdots, Z_n$ sono variabili aleatorie che ammettono densità congiunta $p_{X,Z_1\cdots,Z_n}(x, z_1, \cdots, z_n)$. Una martingala $(M_n)_{n\geq 0}$ è allora successione di variabili aleatorie per cui per ogni $n \geq 1$ M_n una funzione ψ_n^M di $Z_1, \cdots Z_n$ che gode della proprietà di avere valore atteso finito e per la quale vale

$$\mathbb{E}[M_{n+1}|\mathcal{F}_n] = M_n$$

cioè

$$\mathbb{E}[\psi_{n+1}^M(Z_1, \cdots, Z_n, Z_{n+1})|Z_1, \cdots Z_n] = \psi_n^M(Z_1, \cdots, Z_n)$$

si ottiene immediatamente

$$\begin{aligned}
M_{N-1} &= \widetilde{\mathbb{E}}\left[M_N \middle| \mathcal{F}_{N-1}\right] = \widetilde{\mathbb{E}}\left[\frac{f(S_N)}{B_N} \middle| \mathcal{F}_{N-1}\right] \\
&= \widetilde{\mathbb{E}}\left[\frac{f(S_{N-1}\cdot u)}{B_N}\cdot \mathbf{1}_{\{Z_N=u\}} + \frac{f(S_{N-1}\cdot d)}{B_N}\cdot \mathbf{1}_{\{Z_N=d\}} \middle| \mathcal{F}_{N-1}\right] \\
&= \frac{f(S_{N-1}\cdot u)}{B_N}\cdot \widetilde{\mathbb{E}}\left[\mathbf{1}_{\{Z_N=u\}} \middle| \mathcal{F}_{N-1}\right] + \frac{f(S_{N-1}\cdot d)}{B_N}\cdot \widetilde{\mathbb{E}}\left[\mathbf{1}_{\{Z_N=d\}} \middle| \mathcal{F}_{N-1}\right] \\
&= \frac{f(S_{N-1}\cdot u)}{B_N}\cdot \widetilde{p} + \frac{f(S_{N-1}\cdot d)}{B_N}\cdot \widetilde{q} =: \hat{g}_{N-1}(S_{N-1}).
\end{aligned}$$

Analogamente se $M_n = \hat{g}_n(S_n)$, allora

$$\begin{aligned}
M_{n-1} &= \widetilde{\mathbb{E}}\left[M_n \middle| \mathcal{F}_{n-1}\right] = \widetilde{\mathbb{E}}\left[\hat{g}_n(S_n) \middle| \mathcal{F}_{n-1}\right] \\
&= \widetilde{\mathbb{E}}\left[\hat{g}_n(S_{n-1}\cdot u)\cdot \mathbf{1}_{\{Z_n=u\}} + \hat{g}_n(S_{n-1}\cdot d)\cdot \mathbf{1}_{\{Z_n=d\}} \middle| \mathcal{F}_{n-1}\right] \\
&= \hat{g}_n(S_{n-1}\cdot u)\cdot \widetilde{\mathbb{E}}\left[\mathbf{1}_{\{Z_n=u\}} \middle| \mathcal{F}_{n-1}\right] + \hat{g}_n(S_{n-1}\cdot d)\cdot \widetilde{\mathbb{E}}\left[\mathbf{1}_{\{Z_n=d\}} \middle| \mathcal{F}_{n-1}\right] \\
&= \hat{g}_n(S_{n-1}\cdot u)\cdot \widetilde{p} + \hat{g}_n(S_{n-1}\cdot d)\cdot \widetilde{q} =: \hat{g}_{n-1}(S_{n-1}).
\end{aligned}$$

Riassumendo abbiamo visto che quindi, nel caso delle opzioni plain vanilla, per ottenere la martingala M_n basta

1) calcolare M_N alla fine dell'albero binomiale,**tramite $M_N = \widehat{g}_N(S_N) = \dfrac{f_N(S_N)}{B_0(1+r)^N}$
2)**noto $\widehat{g}_n$, calcolare $M_{n-1} = \hat{g}_{n-1}(S_{n-1}) = \hat{g}_n(S_{n-1}\cdot u)\cdot \widetilde{p} + \hat{g}_n(S_{n-1}\cdot d)\cdot \widetilde{q}$.

Ovviamente queste operazioni corrispondono alle operazioni deterministiche[58]**che corrispondono a calcolare la funzione $\hat{g}_n(s_n)$, per $= 0, 1, ..., n$, (con $n = N, N-1, \ldots, 1, 0$), e dove s_n sono i possibili valori che può assumere la variabile aleatoria S_n, ossia per $s_n \in \{s_0\, u^k\, d^{n-k},$ con $k = 0, 1, ..., n\}$, attraverso i seguenti passi:

[58]Queste operazioni corrispondono anche a calcolare la funzione $\hat{g}_n(s_0\, u^k\, d^{n-k})$ per $k = 0, 1, ..., n$, (con $n = N, N-1, \ldots, 1, 0$) attraverso i passi

(1) calcolare $\hat{g}_N(s_0\, u^k\, d^{N-k}) = f(s_0\, u^k\, d^{N-k})/B_N$, per $k = 0, 1, ...N$,
(2) calcolare, per $n \leq N$,

$$\begin{aligned}
\hat{g}_{n-1}(s_0\, u^k\, d^{n-1-k}) &= \hat{g}_n(s_0\, u^k\, d^{n-1-k}\cdot u)\cdot \widetilde{p} + \hat{g}_n(s_0\, u^k\, d^{n-1-k}\cdot d)\cdot \widetilde{q} \\
&= \hat{g}_n(s_0\, u^{k+1}\, d^{n-k+1})\cdot \widetilde{p} + \hat{g}_n(s_0\, u^k\, d^{n-k})\cdot \widetilde{q}
\end{aligned}$$

Si osservi che se si indica con

$$\tilde{V}_n(k) := \hat{g}_n(s_0\, u^k\, d^{n-k}),$$

allora la precedente formula ricorsiva diviene

(1) calcolare $\tilde{V}_N(k) = f(s_0\, u^k\, d^{N-k})/B_N$, per $k = 0, 1, ...N$,
(2) calcolare $\tilde{V}_n(k)$, per $k = 0, 1, ...N$, attraverso la formula ricorsiva

$$\tilde{V}_{n-1}(k) = \tilde{V}_n(k+1)\cdot \widetilde{p} + \tilde{V}_n(k)\cdot \widetilde{q}.$$

Inoltre, tenendo conto del fatto che

$$S_n = s_0\, u^{H_n}\, d^{n-H_n}$$

si ottiene che

$$M_n = \hat{g}_n(s_0\, u^{H_n}\, d^{n-H_n}) = \tilde{V}_n(H_n).$$

La funzione

$$V_n(k) = B_n\, \tilde{V}_n(k) = B_n\, \hat{g}_n(s_0\, u^k\, d^{n-k}) = B_0\,(1+r)^n\, \hat{g}_n(s_0\, u^k\, d^{n-k}),$$

anche ammette una formula ricorsiva, che risulta

$$V_{n-1}(k) = \frac{1}{1+r}\left(V_n(k+1)\cdot \widetilde{p} + V_n(k)\cdot \widetilde{q}\right),$$

1) porre

$$\hat{g}_N(s_N) := \frac{f(s_N)}{B_N},$$

per tutti i valori che può assumere S_N, ossia per $s_N \in \{s_0\, u^k\, d^{N-k},\ \text{con}\ k = 0, 1, ..., N\}$;

2) calcolare $\hat{g}_{n-1}(s_{n-1})$, attraverso la formula ricorsiva

$$\hat{g}_{n-1}(s_{n-1}) = \hat{g}_n(s_{n-1} \cdot u) \cdot \widetilde{p} + \hat{g}_n(s_{n-1} \cdot d) \cdot \widetilde{q},$$

per tutti i valori che può assumere S_{n-1}, ossia per $s_{n-1} \in \{s_0\, u^k\, d^{n-1-k},\ \text{con}\ k = 0, 1, ..., n-1\}$;

ATTENZIONE INIZIO CAMBIAMENTO

Come già sappiamo, l'interesse della rappresentazione della martingala M_n è dovuto ai seguenti fatti:

(I) Calcolo del prezzo dell'opzione: Vedremo come si possa arrivare ad una formula ricorsiva per il calcolo del prezzo, non attualizzato. Anche in questo caso c'è una formula ricorsiva (all'indietro) per la funzione $c(n, x) = B_n \hat{g}_n(x)$, che permette di calcolare il valore del portafoglio al tempo n come $c(n, S_n)$, ed il prezzo al tempo zero come $c(0, s_0)$.
(II) Calcolo della strategia di copertura perfetta: si ottiene l'espressione esplicita della strategia autofinanziante di copertura perfetta data dal cosidetto Delta-hedging ossia

$$\tilde{\gamma}_n(\omega) = \frac{c(n, S_{n-1}(\omega)\, u) - c(n, S_{n-1}(\omega)\, d)}{S_{n-1}(\omega)u - S_{n-1}(\omega)d},$$

(si ricordi che $\tilde{\beta}_n$ si ottiene immediatamente dalla condizione che la strategia sia autofianziante).
Il nome di Delta-hedging, viene dal fatto che la strategia (o portafoglio) si può riscrivere

$$\tilde{\gamma}_n(\omega) = \frac{\Delta c(n)}{\Delta S(n)}$$

dove $\Delta c(n)$ è la differenza tra i due possibili valori del portafoglio al tempo n, e $\Delta S(n)$ è la differenza tra i due possibili prezzi del titolo rischioso al tempo n, entrambe le differenze sono calcolate al tempo $n - 1$, quando sappiamo che il prezzo del titolo rischioso vale S_{n-1}.

(I) Calcolo del prezzo dell'opzione
Se $\tilde{\pi} = \big(\tilde{\beta}_n, \tilde{\gamma}_n\big)_n$ è una strategia di copertura perfetta, allora, sotto la misura martingala equivalente $\widetilde{\mathbb{P}}$, il processo valore attualizzato $\tilde{X}^{\tilde{\pi}}_n = X^{\tilde{\pi}}_n / B_n$ gode delle proprietà che

$$\tilde{X}^{\tilde{\pi}}_N = \frac{f(S_N)}{B_N}, \qquad \tilde{X}^{\tilde{\pi}}_n = \sum_{k=1}^{n} \tilde{\gamma}_k \big(\tilde{S}_k - \tilde{S}_{k-1}\big), \quad 1 \le n \le N. \qquad \textit{cambiamento}$$

per cui è una martingala (sempre sotto $\widetilde{\mathbb{P}}$). Poiché entrambe M_N e $\tilde{X}^{\tilde{\pi}}_N$ coincidono con il terminal pay-off attualizzato, ovvero

$$M_N = \tilde{X}^{\tilde{\pi}}_N = \frac{f(S_N)}{B_N},$$

come si vede subito tenendo conto del fatto che

$$\begin{aligned} V_{n-1}(k) &= B_0\,(1+r)^{n-1}\,\tilde{V}_{n-1}(k) = \frac{1}{1+r} B_0\,(1+r)^n\,\tilde{V}_{n-1}(k) \\ &= \frac{1}{1+r} B_0\,(1+r)^n\,\Big(\tilde{V}_n(k+1) \cdot \widetilde{p} + \tilde{V}_n(k) \cdot \widetilde{q}\Big) \\ &= \frac{1}{1+r}\,\left(V_n(k+1) \cdot \widetilde{p} + V_n(k) \cdot \widetilde{q}\right). \end{aligned}$$

Abbiamo quindi ottenuto la formula ricorsiva che viene presentata nel libro di Björk [5].

necessariamente deve essere (se due martingale coincidono, con probabilità 1, nell'istante N, allora coincidono, con probabilità 1, in tutti i tempi precedenti $n \leq N$)

$$M_n = \tilde{X}_n^{\tilde{\pi}}, \qquad \text{per ogni} \quad n = 0, 1, \ldots, N.$$

Da ciò discende anche che si può ottenere una rappresentanzione per il valore della strategia di copertura, non attualizzato, cioè

$$X_n^{\tilde{\pi}} = B_n\, \tilde{X}_n^{\tilde{\pi}} = B_n\, M_n = B_n\, \widetilde{\mathbb{E}}\left[\frac{f_N}{B_N}\Big|\mathcal{F}_n\right] = B_n\, g_n(\rho_1, \cdots, \rho_n).$$

Nel caso delle opzioni plain vanilla, quindi, la funzione valore diviene

$$X_n^{\tilde{\pi}} = B_n\, \hat{g}_n(S_n) = c(n, S_n), \tag{1.37}$$

dove

$$c(n, x) = c_N(n, x) := B_0\, (1 + r)^n \hat{g}_n(x). \tag{1.38}$$

Nella precedente definizione abbiamo messo in evidenza la dipendenza dal tempo di esercizio N, in quanto anche la funzione $\hat{g}_n(x)$ dipende da tale valore[59].

Ovviamente anche $c_N(n, x)$ si può ottenere con una ***formula ricorsiva***:

1) porre

$$c_N(N, s_N) := f(s_N) \left(= B_N\, \frac{f(s_N)}{B_N}\right),$$

per tutti i valori che può assumere S_N, ossia per $s_N \in \{s_0\, u^k\, d^{N-k},$ con $k = 0, 1, ..., N\}$;

2) calcolare $c_N(n-1, s_{n-1})$, attraverso la formula ricorsiva

$$c_N(n-1, s_{n-1}) = \frac{1}{1+r}\, \big(c_N(n, s_{n-1} \cdot u) \cdot \widetilde{p} + c_N(n, s_{n-1} \cdot d) \cdot \widetilde{q}\big),$$

per tutti i valori che può assumere S_{n-1}, ossia per $s_{n-1} \in \{s_0\, u^k\, d^{n-1-k},$ con $k = 0, 1, ..., n-1\}$;

Questa osservazione ha anche l'importante conseguenza che il prezzo dell'opzione si ottiene come[60]

$$x_0 = B_0\, \tilde{X}_0^{\tilde{\pi}} = c_N(0, s_0).$$

(II) Calcolo della strategia di copertura perfetta

La strategia di copertura perfetta si ottiene immediatamente dalla $\tilde{S}-$rappresentazione di $M_n = \widetilde{\mathbb{E}}\big[\tilde{f}_N \big|\mathcal{F}_n\big]$, ossia nel momento in cui si abbia che

$$M_n = \sum_{k=1}^{n} \tilde{\gamma}_k \big(\tilde{S}_k - \tilde{S}_{k-1}\big),$$

in quanto, appunto il valore attualizzato $\tilde{X}_n^{\tilde{\pi}}$ coincide con M_n, se appunto la strategia scelta è data dall'investire in γ_k azioni nell'intervallo $(k-1, k)$ e nel mettere in banca il rimanente, ossia la differenza tra $\tilde{X}_{k-1}^{\tilde{\pi}} - \tilde{\gamma}_k\, \tilde{S}_{k-1}$, infatti $\tilde{X}_{k-1}^{\tilde{\pi}} = \tilde{\beta}_{k-1} + \tilde{\gamma}_{k-1}\, \tilde{S}_{k-1}$ per definizione, ma deve essere anche $\tilde{X}_{k-1}^{\tilde{\pi}} = \tilde{\beta}_k + \tilde{\gamma}_k\, \tilde{S}_{k-1}$, se la strategia $\tilde{\pi} = (\tilde{\beta}_n, \tilde{\gamma}_n)_n$ è autofinanziante.

[59] Abbiamo indicato tale funzione con $c(n, x)$, ma ovviamente il valore del portafoglio di copertura perfetta dipende anche dai parametri r, a e b (o se si preferisce da r, u e d).

[60] Nelle notazioni del Björk [5] ciò corrisponde a

$$x_0 = B_0\, \tilde{X}_0^{\tilde{\pi}} = B_0\, \tilde{V}(H_0) = B_0\, \tilde{V}(0) = V(0).$$

Ricordiamo brevemente come si ottiene il numero $\tilde{\gamma}_k$ di azioni da comprare nel caso generale: una volta data la funzione $g_n(\rho_1, \cdots, \rho_n)$, che individua[61] M_n, allora la condizione che M_n sia una martingala, ossia

$$M_{n-1} = \widetilde{\mathbb{E}}\big[M_n \big| \mathcal{F}_{n-1}\big]$$

diviene la condizione[62]

$$g_{n-1}(\rho_1(\omega), \cdots, \rho_{n-1}(\omega)) = \tilde{p}\, g_n(\rho_1(\omega), \cdots, \rho_{n-1}(\omega), b) + \tilde{q}\, g_n(\rho_1(\omega), \cdots, \rho_{n-1}(\omega), a).$$

Inoltre, come abbiamo visto, la condizione che M_n sia una martingala equivale ad affermare che

$$\gamma_n'(\omega) := \frac{g_n(\rho_1(\omega), \cdots, \rho_{n-1}(\omega), b) - g_{n-1}(\rho_1(\omega), \cdots, \rho_{n-1}(\omega))}{b-r} \tag{1.39}$$

coincide con l'analoga espressione con a al posto di b, ovvero che vale anche l'uguaglianza

$$\gamma_n'(\omega) = \frac{g_n(\rho_1(\omega), \cdots, \rho_{n-1}(\omega), a) - g_{n-1}(\rho_1(\omega), \cdots, \rho_{n-1}(\omega))}{a-r}. \tag{1.40}$$

Inoltre, come abbiamo già visto, nel caso generale abbiamo

$$\begin{aligned} M_n(\omega) - M_{n-1}(\omega) =& \mathbf{1}_{\{\rho_n = a\}}(\omega) \left[g_n(\rho_1(\omega), \cdots, \rho_{n-1}(\omega), a) - g_{n-1}(\rho_1(\omega), \cdots, \rho_{n-1}(\omega))\right] \\ &+ \mathbf{1}_{\{\rho_n = b\}}(\omega) \left[g_n(\rho_1(\omega), \cdots, \rho_{n-1}(\omega), b) - g_{n-1}(\rho_1(\omega), \cdots, \rho_{n-1}(\omega))\right] \\ =& \mathbf{1}_{\{\rho_n = a\}}(\omega)\, \gamma_n'(\omega) \big(a - r\big) + \mathbf{1}_{\{\rho_n = b\}}(\omega)\, \gamma_n'(\omega) \big(b - r\big) \\ =& \gamma_n'(\omega) \big(\rho_n - r\big) \end{aligned}$$

ed infine, tenendo conto che $\tilde{S}_n - \tilde{S}_{n-1} = \frac{(1+\rho_n)S_{n-1}}{(1+r)B_{n-1}} - \frac{S_{n-1}}{B_{n-1}} = \frac{S_{n-1}}{B_{n-1}} \left[\frac{1+\rho_n}{1+r} - 1\right] = \frac{S_{n-1}}{B_{n-1}} \frac{\rho_n - r}{1+r} = \frac{S_{n-1}\,(\rho_n - r)}{B_{n-1}(1+r)}$, si ha

$$M_n(\omega) - M_{n-1}(\omega) = \gamma_n'(\omega)\, \frac{B_{n-1}(1+r)}{S_{n-1}(\omega)} \big(\tilde{S}_n(\omega) - \tilde{S}_{n-1}(\omega)\big) = B_n\, \frac{\gamma_n'(\omega)}{S_{n-1}(\omega)} \big(\tilde{S}_n(\omega) - \tilde{S}_{n-1}(\omega)\big),$$

da cui

$$\tilde{\gamma}_n(\omega) = B_n\, \frac{\gamma_n'(\omega)}{S_{n-1}(\omega)}. \tag{1.41}$$

Nel caso delle opzioni plain vanilla, tenendo conto che $g_n(\rho_1(\omega), \cdots, \rho_{n-1}(\omega), \rho_n(\omega)) = \hat{g}_n(S_n(\omega))$, la condizione che M_n sia una martingala sotto $\widetilde{\mathbb{P}}$ diviene, come abbiamo visto all'inizio del paragrafo in modo diretto,

$$\hat{g}_{n-1}(S_{n-1}) = \tilde{p}\, \hat{g}_n(S_{n-1}\, u) + \tilde{q} \hat{g}_n(S_{n-1}\, d), \tag{1.42}$$

mentre la condizione che (1.39) coincide con (1.40) diviene

$$\gamma_n'(\omega) = \frac{\hat{g}_n(S_{n-1}(\omega)\, u) - \hat{g}_{n-1}(S_{n-1}(\omega))}{b-r} = \frac{\hat{g}_n(S_{n-1}(\omega)\, d) - \hat{g}_{n-1}(S_{n-1}(\omega))}{a-r}. \tag{1.43}$$

Inserendo in quest'ultima relazione (1.43) la precedente espressione (1.42) per $\hat{g}_{n-1}(S_{n-1})$, e tenendo conto che

[61] Come abbiamo già visto prima, nel caso delle opzioni plain vanilla sappiamo che la funzione $g_n(\rho_1, \cdots, \rho_n)$ risulta essere funzione deterministica di S_n ossia $M_n = \hat{g}_n(S_n)$.

[62] Ricordiamo che i punti chiave sono scrivere $M_n = \mathbf{1}_{\{\rho_n = a\}}\, g_n(\rho_1, \cdots, \rho_{n-1}, a) + \mathbf{1}_{\{\rho_n = b\}}\, g_n(\rho_1, \cdots, \rho_{n-1}, b)$, la linearità del valore atteso condizionato, ed il fatto che $\widetilde{\mathbb{E}}\big[\mathbf{1}_{\{\rho_n = x\}} \big| \mathcal{F}_n\big] = \widetilde{\mathbb{P}}(\{\rho_n = x\} \big| \mathcal{F}_n) = \widetilde{\mathbb{P}}(\{\rho_n = x\})$ vale $\tilde{q}$ per $x = a$ e $\tilde{p}$ per $x = b$.

$\tilde{q} = 1 - \tilde{p} = \frac{b-r}{b-a}$ e che $b - a = 1 + b - (1 + a) = u - d$, si ottiene

$$\begin{aligned}
\gamma'_n(\omega) &:= \frac{\hat{g}_n(S_{n-1}(\omega)\,u) - [\tilde{p}\,\hat{g}_n(S_{n-1}(\omega)\,u) + \tilde{q}\hat{g}_n(S_{n-1}(\omega)\,d)]}{b-r} \\
&= \frac{\hat{g}_n(S_{n-1}(\omega)\,u) - \tilde{p}\,\hat{g}_n(S_{n-1}(\omega)\,u) - \tilde{q}\hat{g}_n(S_{n-1}(\omega)\,d)}{b-r} \\
&= \frac{\tilde{q}\,\hat{g}_n(S_{n-1}(\omega)\,u) - \tilde{q}\hat{g}_n(S_{n-1}(\omega)\,d)}{b-r} \\
&= \frac{\tilde{q}}{b-r}\left[\hat{g}_n(S_{n-1}(\omega)\,u) - \hat{g}_n(S_{n-1}(\omega)\,d)\right] \\
&= \frac{\frac{b-r}{b-a}}{b-r}\left[\hat{g}_n(S_{n-1}(\omega)\,u) - \hat{g}_n(S_{n-1}(\omega)\,d)\right] \\
&= \frac{\hat{g}_n(S_{n-1}(\omega)\,u) - \hat{g}_n(S_{n-1}(\omega)\,d)}{b-a} \\
&= \frac{\hat{g}_n(S_{n-1}(\omega)\,u) - \hat{g}_n(S_{n-1}(\omega)\,d)}{u-d}.
\end{aligned}$$

Per ottenere $\tilde{\gamma}_n$, non rimane che utilizzare questa uguaglianza in (1.41) e osservare che, sempre per le opzioni plain vanilla, la strategia di copertura perfetta si ottiene da una derivata discreta, ossia

$$\begin{aligned}
\tilde{\gamma}_n(\omega) = B_n\,\frac{\frac{\hat{g}_n(S_{n-1}(\omega)\,u) - \hat{g}_n(S_{n-1}(\omega)\,d)}{u-d}}{S_{n-1}(\omega)} &= B_n\,\frac{\hat{g}_n(S_{n-1}(\omega)\,u) - \hat{g}_n(S_{n-1}(\omega)\,d)}{S_{n-1}(\omega)u - S_{n-1}(\omega)d} \\
= \frac{B_n\,\hat{g}_n(S_{n-1}(\omega)\,u) - B_n\,\hat{g}_n(S_{n-1}(\omega)\,d)}{S_{n-1}(\omega)u - S_{n-1}(\omega)d}
\end{aligned}$$

Ricordando la (1.37) e la definizione (1.38) della funzione $c(n,x)$, si ha dunque che in questo caso la strategia di copertura perfetta diviene una sorta di derivata discreta rispetto al parametro x della funzione $c(n,x)$, in quanto

$$\tilde{\gamma}_n(\omega) = \frac{c(n, S_{n-1}(\omega)\,u) - c(n, S_{n-1}(\omega)\,d)}{S_{n-1}(\omega)u - S_{n-1}(\omega)d}.$$

FIGURA 1:
come si trovano i valori di S_n per $n = 0, 1, ..., N$

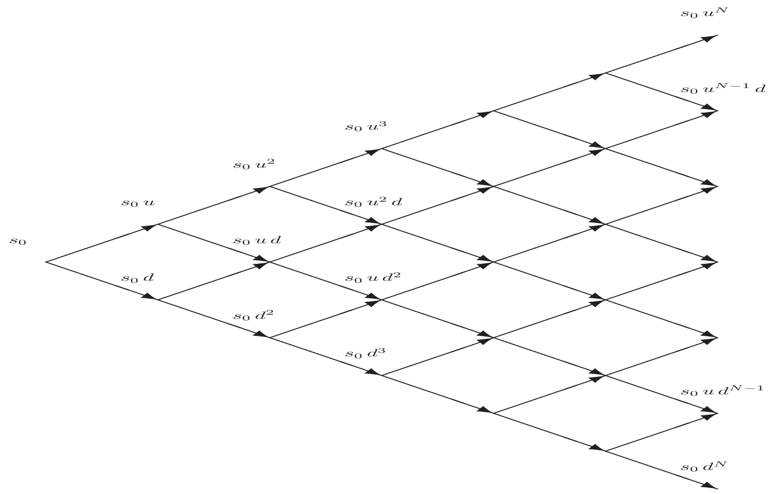

Figura 1.2: Albero binomiale per il calcolo di S_n per $n = 0, 1, ..., N$. Lo spazio Ω viene rappresentato come l'insieme dei cammini possibili.

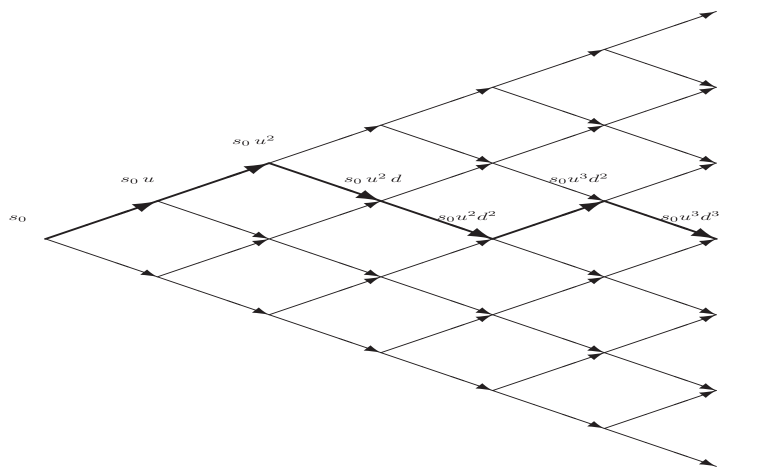

Figura 1.3: Albero binomiale per il calcolo di S_n per $n = 0, 1, ..., N$. Esempio di cammino: in grassetto si vede il cammino relativo all'*evento* $\{Z_1 = u, Z_2 = u, Z_3 = d, Z_4 = d, Z_5 = u, Z_6 = d\}$

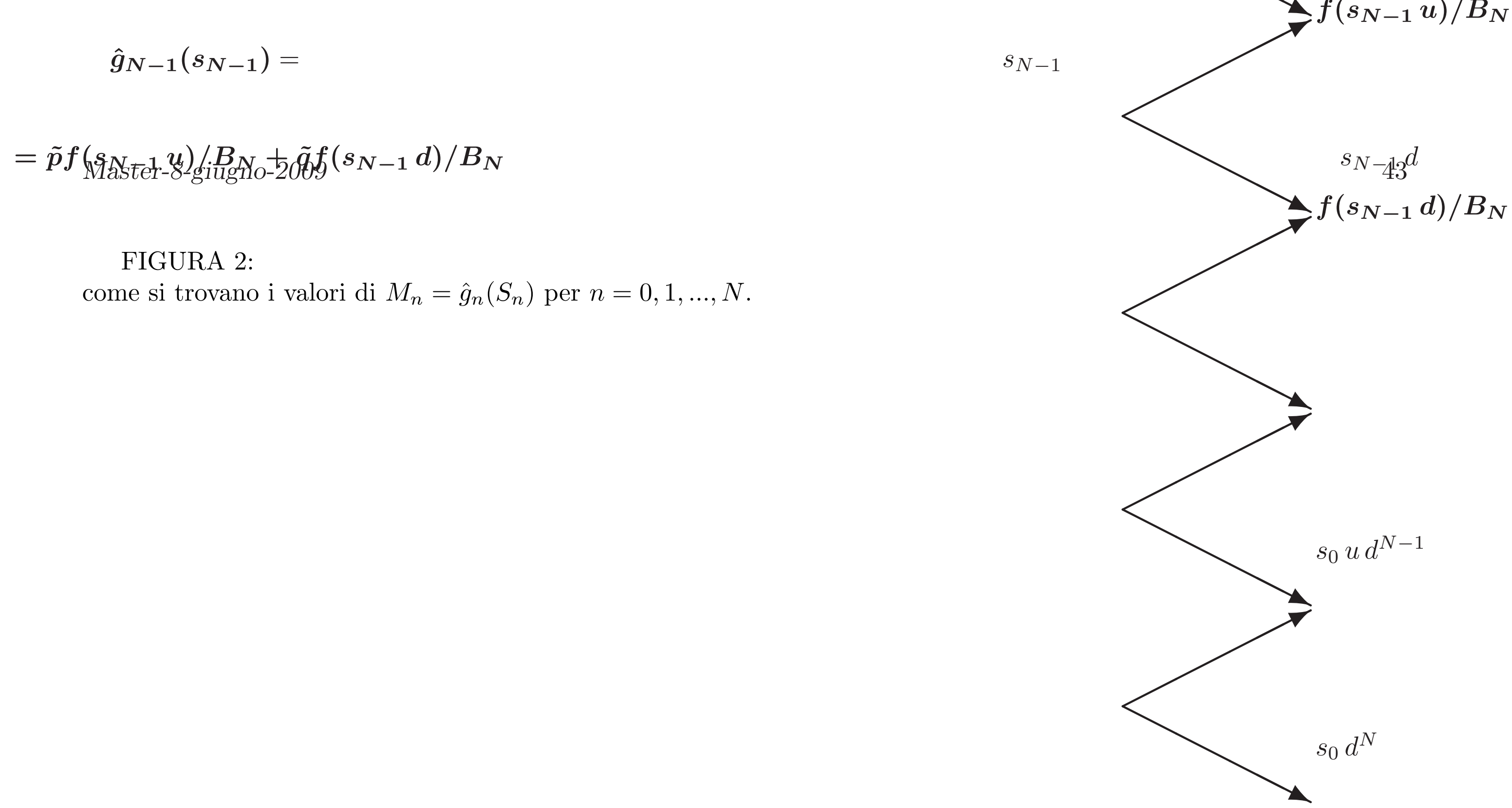

FIGURA 2:
come si trovano i valori di $M_n = \hat{g}_n(S_n)$ per $n = 0, 1, ..., N$.

Figura 1.4: Albero binomiale per il calcolo di $\hat{g}_{N-1}$ e della strategia di copertura perfetta $\tilde{\gamma}_N$

In particolare la strategia di copertura perfetta per $n = N$ ed $S_{N-1} = s_{N-1}$

$$\tilde{\gamma}_N = \frac{B_N\, f(s_{N-1}\, u)/B_N - B_N\, f(s_{N-1}\, d)/B_N}{s_{N-1}u - s_{N-1}d} = \frac{f(s_{N-1}\, u) - f(s_{N-1}\, d)}{s_{N-1}u - s_{N-1}d},$$

mentre per n generico ed $S_{n-1} = s_{n-1}$, in un nodo generico $s_{n-1} = s_0\, u^k\, d^{n-1-k}$, si ha

$$\boldsymbol{\tilde{\gamma}_n = \frac{B_n\, \hat{g}_n(s_{n-1}\, u) - B_n\, \hat{g}_n(s_{n-1}\, d)}{s_{n-1}u - s_{n-1}d} = \frac{c(n, s_{n-1}\, u) - c(n, s_{n-1}\, d)}{s_{n-1}u - s_{n-1}d}}.$$

Figura 1.5: Albero binomiale per il calcolo della strategia di copertura perfetta $\tilde{\gamma}_n$.

1.7 Appendice: Approccio soggettivista alla probabilità

In questo paragrafo vogliamo illustrare come con l'approccio soggettivista, si possa arrivare agli assiomi della probabilità (con la additività semplice) attraverso una opportuna definizione di valore atteso e di probabilità come prezzo, e attraverso opportune regole di linearità e di coerenza.

Daremo due definizioni che sostanzialmente si equivalgono, una con un linguaggio più neutrale e generale ed un'altra con un linguaggio più economico.

Definizione 1.5 (valore atteso)**.** *Il valore atteso $\mathbb{E}(X)$ di una variabile aleatoria X è quel valore certo c che sono disposto a scambiare con il valore aleatorio dato proprio da X, nel senso che per me è indifferente ricevere c o X. La probabilità $\mathbb{P}(A)$ di un evento A è definito come il valore atteso della variabile aleatoria indicatrice di A, ovvero la variabile aleatoria $X = I_A$ che vale 1 se si verifica A e vale 0 altrimenti.*

La seconda definizione vede X come valore aleatorio che si ottiene come esito di una scommessa, mentre c come prezzo da pagare per prendere parte alla scommessa (o al gioco).

Definizione 1.6 (valore atteso come prezzo)**.** *Il valore atteso di X è quel valore certo c che si è disposti a pagare per effettuare la scommessa in cui si ottiene il valore aleatorio X, con l'accordo che si è disposti indifferentemente a prendere il ruolo sia dello scommettitore che del banco.*

Regola di linearità[63]: Siano X ed Y due variabili aleatorie allora

i $\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)$

ii $\mathbb{E}(\alpha\, X + \beta\, Y) = \alpha\, \mathbb{E}(X) + \beta\, \mathbb{E}(Y)$, per ogni α e β costanti reali

La regola di coerenza seguente è basata sul fatto alla fine della scommessa ricevo X e pago c e quindi alla fine ho $X - c$, se ho il ruolo dello scommettitore, mentre ricevo c e pago X se ho il ruolo del banco (broker).

Regola di coerenza: Non è possibile che $X - c$ sia certamente positivo o certamente negativo[64].

La giustificazione della precedente regola di coerenza sta nel fatto che se $X - c$ ha segno costante, non si troverebbe nessuno disposto a prendere il ruolo dello scommettitore, se $X - c$ fosse certamente negativo, e analogamente non si troverebbe nessuno disposto a prendere il ruolo del banco se invece $X - c$ fosse certamente positivo.

A titolo di esempio mostriamo come dalla regola di coerenza si ottenga che se X è una funzione indicatrice di un evento A, ovvero X assume solo i valori 0 ed 1 (e allora $A = \{X = 1\}$) e, posto per definizione $\mathbb{P}(A) = \mathbb{E}(X)$, allora necessariamente deve valere $\mathbb{P}(A) \in [0, 1]$.

La seguente tavola illustra i ricavi possibili se $\mathbb{P}(A) = p$ è il prezzo da pagare per la scommessa in cui si vince 1 se si verifica A (e quindi si ottiene globalmente $1 - p$), mentre se si verifica A^c non si vince nulla (e quindi si ottiene $-p$)

[63]In realtà, per ottenere regola di linearità **[i]**, basterebbe la regola di coerenza (data immediatamente dopo la regola di linearità) e aggiungere l'ipotesi che sia sempre possibile trovare qualcuno disposto a scommettere su ciascuna singola scommessa:
se il prezzo di due scommesse insieme fosse maggiore della somma dei prezzi delle due scommesse separatamente, allora converrebbe accettare (comprare) due singole scommesse e prendere la posizione del banco (vendere) la scommessa relativa alla somma

se invece il prezzo della somma fosse minore della somma dei prezzi, allora converrebbe vendere le due scommesse separatamente e invece accettare la scommessa. Per capire meglio si veda la seguente tabella: se non ci possono essere arbitraggi, allora necessariamente deve accadere che $c - c_1 - c_2 = 0$

tipo di scommessa	solo la 1	solo la 2	1 e 2, ma con il ruolo opposto	le precedenti insieme
vincita	X_1	X_2	$-X = -(X_1 + X_2)$	$X_1 + X_2 - X = 0$
pagamento	$-c_1$	$-c_2$	$+c$	$c_1 - c_2 + c$
ricavo	$X_1 - c_1$	$X_2 - c_2$	$c - X$	$X_1 - c_1 + X_2 - c_2 + c - X = c - c_1 - c_2$

Per la regola di linearità **[ii]** si può invece pensare all'assenza di sconti, nel senso che se si scommette αX si paga esattamente α volte il prezzo della scommessa X, nessun *prendi tre paghi due*!!!

[64]Interpretando $X - c$ come il guadagno di chi ha pagato c per ottenere in cambio la cifra aleatoria X, e quindi $c - X$ come il guadagno del venditore, la Regola di coerenza afferma che non può esserci un arbitraggio (forte) ne' per il compratore, ne' per il venditore.

evento	A	A^c
pagamento	$-p$	$-p$
vincita	1	0
ricavo	$1-p$	$-p$

Per la regola di coerenza si ottiene che ***non può essere***

$$\Big(1-p<0 \quad \text{e} \quad -p<0\Big) \quad \text{oppure} \quad \Big(1-p>0 \quad \text{e} \quad -p>0\Big),$$

Equivalentemente ***non può essere***

$$\Big(1<p \quad \text{e} \quad 0<p\Big) \quad \text{oppure} \quad \Big(1>p \quad \text{e} \quad 0>p\Big),$$

cioè ***non può essere***

$$1<p \quad \text{oppure} \quad 0>p,$$

e quindi in altre parole ***deve essere necessariamente***

$$0 \le p = \mathbb{P}(A) \le 1,$$

che corrisponde alla richiesta degli assiomi che la probabilità sia a valori in $[0,1]$.

Nel caso particolare in cui $A=\Omega$ è l'evento certo la tabella si precedente si riduce a

evento	Ω
pagamento	$-p$
vincita	1
ricavo	$1-p$

Non potendo essere ne' $1-p<0$ ne' $1-p>0$, ovvero non potendo essere ne' $1<p$ ne' $1>p$, deve necessariamente essere $p=\mathbb{P}(\Omega)=1$, che è un altro degli assiomi delle probabilità.

Se inoltre ho n scommesse relative ad n eventi A_1, A_2, ..., A_n incompatibili (cioè $A_i \cap A_j = \emptyset$ per $i \neq j$) allora la scommessa su $A=\bigcup_{i=1}^n A_i$ equivale[65] alla somma delle singole n scommesse su A_i, in quanto grazie all'incompatibilità degli A_i in entrambe le scommesse ricevo 1 se e solo se si verifica uno degli A_i, mentre altrimenti ottengo 0. Quindi la linearità dei prezzi, più il fatto implicito che se due scommesse danno luogo alla stessa vincita, allora devono avere lo stesso prezzo, si ottiene che

$$\mathbb{P}(A)\,\Big(=\mathbb{P}\Big(\cup_{i=1}^n A_i\Big)\Big) = \sum_{i=1}^n \mathbb{P}(A_i),$$

che corrisponde all'assioma dell'additività finita.

Riassumendo abbiamo mostrato come si possano riottenere gli assiomi della probabilità (con l'esclusione della $\sigma-$additività) con la definizione di valore atteso come prezzo di scommesse aleatorie, con la ***regola di coerenza*** che corrisponde all'***assenza*** di ***opportunità di arbitraggio***.

[65]Ovvero se $A=\bigcup_{i=1}^n A_i$ e se $A_i \cap A_j = \emptyset$ per $i \neq j$, allora

$$I_A = \sum_{i=1}^n I_{A_i}.$$

Va notato che l'ipotesi di linearità corrisponde all'assenza di costi di transazione: quando ci sono costi di transazione allora comprare all'ingrosso (ovvero fare un'unica scommessa su $A = \cup_{i=1}^{n} A_i$) è in genere più conveniente che comprare al dettaglio (ovvero fare n scommesse separate su ciascun A_i, per $i = 1, \ldots, n$).

Terminiamo queste osservazioni sulle probabilità soggettive dando l'***interpretazione del valore atteso condizionato***[66] ***ad un evento*** A di una variabile aleatoria X come il valore certo c_A che si è disposti a scambiare con X, tenendo presente che lo scambio avviene ***solo se si verifica*** l'eventualità rappresentata dall'evento A, ovvero con l'intesa che se l'evento A non si verifica, allora non viene effettuato alcuno scambio[67].

[66] Vale un'interpretazione analoga per le probabilità condizionate ad un evento A, prendendo $X = I_E$.

[67] Come a dire che il contratto o la scommessa non hanno validità se non si verifica la condizione A.

Capitolo 2

Il modello di Black e Scholes come limite del modello binomiale multiperiodale

2.1 Il Modello Binomiale Multiperiodale

Ricordiamo brevemente il Modello Binomiale Multiperiodale (o Cox-Ross-Rubinstein)

2.1.1 Ipotesi e notazioni

Il tasso di interesse è costante e vale r, mentre il prezzo dell'azione è dato da

$$S_0 = s_0 > 0, \qquad S_{n+1} = (1 + \rho_{n+1})S_n = Z_{n+1}S_n$$

o in altre parole

$$S_0 = s_0 > 0, \qquad S_1 = (1 + \rho_1)S_0 = Z_1S_0 \quad \dots \quad S_n = (1 + \rho_n) \cdots (1 + \rho_2)(1 + \rho_1)S_0 = Z_n \cdots Z_2Z_1S_0$$

dove le variabili aleatorie ρ_i possono assumere solo il valori a e b, con la condizione

$$a < r < b,$$

o equivalentemente le variabili aleatorie $Z_i = 1 + \rho_i$ possono assumere solo i valori

$$u = 1 + b, \qquad d = 1 + a,$$

con la condizione

$$d < 1 + r < u.$$

Si definisca

$$\xi_i = \mathbf{1}_{\{\rho_i=b\}} = \mathbf{1}_{\{Z_i=u\}}$$

ovvero la variabile aleatoria che vale 1 se il prezzo dell'azione sale e zero altrimenti, in modo che[1]

$$Z_i = u^{\xi_i}\, d^{1-\xi_i}.$$

[1]Il fatto che $Z_i = u^{\xi_i}\, d^{1-\xi_i}$ si verifica per ispezione:

$$Z_i = u \Leftrightarrow \xi_i = 1 \Leftrightarrow u^{\xi_i}\, d^{1-\xi_i} = u^1\, d^{1-1} = u$$
$$Z_i = d \Leftrightarrow \xi_i = 0 \Leftrightarrow u^{\xi_i}\, d^{1-\xi_i} = u^0\, d^{1-0} = d$$

Sia $H_n(\omega)$ la v.a. che conta il numero delle volte in cui il prezzo sale tra il passo 1 e il passo n, ossia

$$H_k(\omega) = \sum_{i=1}^{k} \mathbf{1}_{\{\rho_i = b\}} = \sum_{i=1}^{k} \mathbf{1}_{\{Z_i = u\}} = \sum_{i=1}^{k} \xi_i$$

in modo che

$$\begin{aligned} S_k N &= S_0\, u^{H_k}\, d^{k-H_k} = S_0\, \left(\frac{u}{d}\right)^{H_k}\, d^k \\ &= S_0\, (1+b)^{H_N}\, (1+a)^{k-H_k} = S_0\, \left(\frac{1+b}{1+a}\right)^{H_k}\, (1+a)^k \end{aligned}$$

per ogni $k = 0, 1, ..., N$ e per ogni pay-off terminale[2] f_N (che deve essere $\mathcal{F}_N$-misurabile) il prezzo di esercizio può essere descritto dalla formula

$$C(f_N, \mathbb{P}) = C_N(f_N, \mathbb{P}) = B_0\, \widetilde{\mathbb{E}}\left[\frac{f_N}{B_N}\right] = \widetilde{\mathbb{E}}\left[\frac{f_N}{(1+r)^N}\right]. \tag{2.1}$$

dove $\widetilde{\mathbb{E}}$ è il valore atteso rispetto alla probabilità $\widetilde{\mathbb{P}}$ rispetto alla quale gli eventi $\{Z_i = u\}$ sono indipendenti e tutti con probabilità

$$\tilde{p} = \frac{1+r-d}{u-d} = \frac{1+r-(1+a)}{(1+b)-(1+a)} = \frac{r-a}{b-a}.$$

Allora, per l'opzione call europea con prezzo di esercizio K e tempo di esercizio N,

$$C_{call}(K, \mathbb{P}) = C_{call,N}(K, \mathbb{P}) \widetilde{\mathbb{E}}\left[\frac{(S_N - K)^+}{B_N}\right] \tag{2.2}$$

e quindi, tenendo conto che H_N ha distribuzione binomiale $Bin(N, \tilde{p})$ rispetto a $\widetilde{\mathbb{P}}$

$$\begin{aligned} C_{call}(K, \mathbb{P}) &= S_0 \sum_{x_0 < h \le N} \binom{N}{h} \widetilde{p}^{\,h} (1-\widetilde{p})^{N-h} \left(\frac{1+b}{1+r}\right)^h \left(\frac{1+a}{1+r}\right)^{N-h} \\ &\quad - \frac{K}{(1+r)^N} \sum_{x_0 < h \le N} \binom{N}{h} \widetilde{p}^{\,h} (1-\widetilde{p})^{N-h} \qquad (2.3) \\ &= S_0 \sum_{h=h_0}^{N} \binom{N}{h} \left(\frac{1+b}{1+r}\,\widetilde{p}\right)^h \left(\frac{1+a}{1+r}\,(1-\widetilde{p})\right)^{N-h} \\ &\quad - \frac{K}{(1+r)^N} \sum_{h=h_0}^{N} \binom{N}{h} \widetilde{p}^{\,h} (1-\widetilde{p})^{N-h}. \\ &= s_0\, \overline{F}_{N, \widetilde{p}\,\frac{u}{1+r}}(x_0) - \frac{K}{(1+r)^N} \overline{F}_{N,\widetilde{p}}(x_0) ** \qquad (2.4) \end{aligned}$$

**dove $\overline{F}_{m,p}(x)$ è la funzione di sopravvivenza di una variabile aleatoria binomiale di parametri m e p, e dove

$$x_0 = \frac{\ln\left(\frac{K}{S_0(1+a)^N}\right)}{\ln\left(\frac{1+b}{1+a}\right)} ****$$

o equivalentemente

$$h_0 = \min\{j \,\epsilon\, \mathbb{N} : S_0(1+a)^{N-j}(1+b)^j - K > 0\}$$

[2]Si ricorda che stiamo trattando obbligazioni derivate di tipo europeo, che possono essere esercitate solo al tempo finale N, o tempo si esercizio, al contrario di quelle di tipo americano, che invece possono essere esercitate in un qualunque istante tra l'inizio del contratto e il tempo di esercizio.

2.2 Approssimazione del Modello Binomiale Multiperiodale

2.2.1 Il modello approssimato, a tempo continuo

Consideriamo ora il caso in cui gli scambi avvengono sempre più vicini nel tempo ovvero ai tempi $t_k^{(n)} = k/n$.

Consideriamo il tempo continuo, ma, per n fissato, i processi che ci interessano sono costanti negli intervalli tra un tempo $t_k^{(n)} = k/n$ e l'altro. Inoltre i parametri del modello, ossia r, u e d, dipenderanno da n, in modo da specificare e da tenere conto del fatto che si tratta di intervalli di ampiezza $1/n$.

Continuiamo ad indicare con B_k ed S_k il prezzo del titolo non rischioso (conto in banca) e del titolo rischioso (l'azione) rispettivamente, anche se, per mettere in evidenza la dipendenza dal parametro n sarebbe più opportuno denotarli con $B_k^{[n]}$ e $S_k^{[n]}$.

Supponiamo

$$B_t^{(n)} = B^{(n)}_{\frac{\lfloor nt \rfloor}{n}} = B_{\lfloor nt \rfloor} = B_0 \left(1 + r(n)\right)^{\lfloor nt \rfloor} \tag{2.5}$$

$$S_t^{(n)} = S^{(n)}_{\frac{\lfloor nt \rfloor}{n}} = S_{\lfloor nt \rfloor} = s_0 \left(\frac{u^{(n)}}{d^{(n)}}\right)^{H_{\lfloor nt \rfloor}} \left(d^{(n)}\right)^{\lfloor nt \rfloor} \tag{2.6}$$

dove il primo segno di uguaglianza sia in (2.5) che in (2.6), garantisce il fatto che i due processi $B_t^{(n)}$ ed $S_t^{(n)}$ sono costanti sugli intervalli di ampiezza $1/n$, mentre nell'ultima uguaglianza in (2.5) si sceglie

$$r^{(n)} = \frac{r}{n}, \tag{2.7}$$

e dove infine nell'ultima uguaglianza[3] in (2.6) si sceglie

$$u^{(n)} = e^{\sigma/\sqrt{n}} \qquad d^{(n)} = e^{-\sigma/\sqrt{n}}, \tag{2.8}$$

per una costante $\sigma > 0$.

Infine anche il valore del tempo di esercizio dipende da n in modo che

$$N^{(n)} = N_n(T), \tag{2.9}$$

dove $N_n(T)$ è il numero di intervalli di ampiezza $1/n$ contenuti in $[0, T]$, ovvero $N_n(T) = \lfloor nT \rfloor$.

Il risultato principale di questa sezione è riassunto nel seguente teorema, che corrisponde ad ottenere la formula di Black e Scholes come limite.

Teorema 2.1. *Posto $C_{call}^{(n)}(K, \mathbb{P})$ il prezzo della call europea nel modello binomiale dato da* (2.5) *e* (2.6)*, con i parametri definiti come in* (2.7) *e* (2.8)*, prezzo di esercizio K e tempo di esercizio $N_n(T)$, si ottiene che*

$$\lim_{n\to\infty} C_{call}^{(n)}(K, \mathbb{P}) = s_0\, \Phi\left(-\zeta + \sigma\sqrt{T}\right) - K e^{-rT} \Phi(-\zeta), \tag{2.10}$$

dove Φ è la funzione di ripartizione della legge $\mathcal{N}(0,1)$,

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-z^2/2} dz \left(= \frac{1}{\sqrt{2\pi}} \int_{-x}^{+\infty} e^{-z^2/2} dz = 1 - \Phi(-x) \right), \quad x \in \mathbb{R}$$

e ζ dipende da K, T, r e σ nel seguente modo:

$$\zeta = \frac{1}{\sigma\sqrt{T}} \left(\log\left(\tfrac{K}{s_0}\right) - \left(r - \frac{\sigma^2}{2}\right) T \right).$$

[3]Si noti che di nuovo, per non appesantire la notazione, non abbiamo utilizzato la notazione $H^{[n]}_{\lfloor nt \rfloor}$, che sarebbe stata più precisa.

Osservazione 2.1. *Tenendo conto che* $-r\,T = \log(e^{-rT}) = -\log(e^{rT})$*, con semplici passaggi si ottiene*

$$\zeta = -\left(\tfrac{1}{\sigma\sqrt{T}}\log\left(\tfrac{s_0\,e^{rT}}{K}\right) - \tfrac{\sigma\sqrt{T}}{2}\right), ** \qquad -\zeta + \sigma\sqrt{T} = \tfrac{1}{\sigma\sqrt{T}}\log\left(\tfrac{s_0\,e^{rT}}{K}\right) + \tfrac{\sigma\sqrt{T}}{2}, **$$

da cui, poiché, come già ricordato $\Phi(x) = 1 - \Phi(-x)$*, il limite in* (2.10) *si riscrive come*

$$s_0\,\Phi\left(\tfrac{1}{\sigma\sqrt{T}}\log\left(\tfrac{s_0\,e^{rT}}{K}\right) + \tfrac{\sigma\sqrt{T}}{2}\right) - Ke^{-rT}\Phi(\tfrac{1}{\sigma\sqrt{T}}\log\left(\tfrac{s_0\,e^{rT}}{K}\right) - \tfrac{\sigma\sqrt{T}}{2}) \tag{2.11}$$

$$= s_0\left[\Phi\left(\tfrac{1}{\sigma\sqrt{T}}\log\left(\tfrac{s_0\,e^{rT}}{K}\right) + \tfrac{\sigma\sqrt{T}}{2}\right) - \tfrac{K}{s_0\,e^{rT}}\,\Phi(\tfrac{1}{\sigma\sqrt{T}}\log\left(\tfrac{s_0\,e^{rT}}{K}\right) - \tfrac{\sigma\sqrt{T}}{2})\right] \tag{2.12}$$

La forma (2.12)*, mette in evidenza come il prezzo si scriva come il prezzo iniziale* s_0 *della opzione, moltiplicato per un fattore che dipende solo dal rapporto*[4] $\frac{K}{s_0\,e^{rT}}$*, fra prezzo di esercizio* K *e prezzo forward della opzione stessa, (cioè* $s_0\,e^{rT}$*, il valore di* s_0 *attualizzato al tempo* T*).*

Prima di tutto commentiamo la scelta del riscalamento, ossia la scelta dei due processi (2.5) e (2.6). Ovviamente la scelta di $r^{(n)}$ in (2.7) corrisponde alla necessità che il tasso di interesse sia proporzionale all'ampiezza degli intervalli $[t_k^{(n)}, t_{k+1}^{(n)}) = [\frac{k}{n}, \frac{k+1}{n})$, ossia si abbia

$$B_{t_k^{(n)}}^{(n)} = B_k\left(= B_k^{[n]}\right) = B_0\left(1 + \frac{r}{n}\right)^k$$

e che rimanga costante in tutto l'intervallo $[t_k^{(n)}, t_{k+1}^{(n)})$, ovvero che

$$B_t^{(n)} = B_{t_k^{(n)}}^{(n)} = B_0\left(1 + \frac{r}{n}\right)^k \qquad \text{per } t \in [t_k^{(n)}, t_{k+1}^{(n)}) = [\tfrac{k}{n}, \tfrac{k+1}{n})$$

o in altre parole che, se, analogamente a $N_n(T)$,

$$N_n(t) = \lfloor nt \rfloor \tag{2.13}$$

è il numero di intervalli di ampiezza $1/n$ che si trovano nell'intervallo $[0, t]$, allora

$$B_t^{(n)} = B_0\left(1 + \frac{r}{n}\right)^{N_n(t)}$$

o meglio

$$B_t^{(n)} = B_{\frac{\lfloor nt \rfloor}{n}}^{(n)} = B_0\left(1 + \frac{r}{n}\right)^{\lfloor nt \rfloor}. \tag{2.14}$$

Inoltre è ragionevole pensare che i cambiamenti del prezzo si discostino di poco in un intervallo di tempo così piccolo, ed in effetti si suppone che

$$S_{t_k^{(n)}}^{(n)} = Z_k^{(n)}\,S_{t_{k-1}^{(n)}}^{(n)} = \left(1 + \rho_k^{(n)}\right)\,S_{t_{k-1}^{(n)}}^{(n)}$$

dove i valori ammissibili per $Z_k^{(n)}$ sono solo $u^{(n)} = e^{\sigma/\sqrt{n}}$ e $d^{(n)} = e^{-\sigma/\sqrt{n}}$, ed in effetti le successioni $u^{(n)}$ e $d^{(n)}$ convergono entrambe ad 1, ovvero, più precisamente,

$$u^{(n)} \underset{n\to\infty}{\searrow} 1 \qquad d^{(n)} \underset{n\to\infty}{\nearrow} 1,$$

Si noti inoltre che, tenendo conto che

$$e^x = 1 + x + \frac{1}{2}\,x^2 + o(x^2),$$

la definizione (2.8) corrisponde a

$$u^{(n)} = 1 + (\sigma/\sqrt{n}) + \frac{1}{2}\,(\sigma^2/n) + o(1/n) \qquad d^{(n)} = 1 - (\sigma/\sqrt{n}) + \frac{1}{2}\,(\sigma^2/n) + o(1/n), \tag{2.15}$$

[4]Ricordiamo che l'opzione si dice alla pari o *at the money* se $K = s_0\,e^{rT}$, *out of the money* se $K < s_0\,e^{rT}$, e *in the money* se invece $K > s_0\,e^{rT}$.

e quindi la condizione di completezza e di assenza di arbitraggio, ossia

$$d^{(n)} < 1 + r^{(n)} < u^{(n)} \quad \Leftrightarrow \quad a^{(n)} < r^{(n)} < b^{(n)}$$

diviene

$$-(\sigma/\sqrt{n}) + \frac{1}{2}\,(\sigma^2/n) + o(1/n) < r/n < (\sigma/\sqrt{n}) + \frac{1}{2}\,(\sigma^2/n) + o(1/n), \tag{2.16}$$

che è chiaramente soddisfatta per n sufficientemente grande.

In altri termini si suppone che

$$u^{(n)} = 1 + b^{(n)} \qquad d^{(n)} = 1 + a^{(n)}.$$

dove

$$b^{(n)} = \frac{\sigma}{\sqrt{n}} + \tfrac{1}{2}\,\frac{\sigma^2}{n} + o(\tfrac{1}{n}) \qquad a^{(n)} = -\frac{\sigma}{\sqrt{n}} + \tfrac{1}{2}\,\frac{\sigma^2}{n} + o(\tfrac{1}{n})$$

e che

$$S_t^{(n)} = S_{t_k^{(n)}}^{(n)}, \qquad \text{per } t \in [t_k^{(n)}, t_{k+1}^{(n)}) = [\tfrac{k}{n}, \tfrac{k+1}{n})$$

ovvero, se come prima, $N_n(t) = \lfloor nt \rfloor$ è il numero di intervalli di ampiezza $1/n$ che si trovano nell'intervallo $[0, t]$, allora, tenendo presente che

$$\frac{u^{(n)}}{d^{(n)}} = \frac{e^{\sigma/\sqrt{n}}}{e^{-\sigma/\sqrt{n}}} = e^{2\sigma/\sqrt{n}},$$

si può anche scrivere

$$\begin{aligned}
S_t^{(n)} &= S_{\frac{\lfloor nt \rfloor}{n}} \left(= S^{[n]}_{\frac{\lfloor nt \rfloor}{n}} \right) = S_0 \left(\frac{u^{(n)}}{d^{(n)}}\right)^{H_{N_n(t)}} \left(d^{(n)}\right)^{\lfloor N_n(t) \rfloor} \\
&= S_0 \left(\frac{u^{(n)}}{d^{(n)}}\right)^{H_{\lfloor nt \rfloor}} \left(d^{(n)}\right)^{\lfloor nt \rfloor} \\
&= S_0\, e^{2\sigma \frac{1}{\sqrt{n}} H_{\lfloor nt \rfloor}}\, e^{-\sigma \frac{1}{\sqrt{n}} \lfloor nt \rfloor} \\
&= S_0\, e^{\sigma \frac{1}{\sqrt{n}} \left(2\, H_{\lfloor nt \rfloor} - \lfloor nt \rfloor\right)}.
\end{aligned} \tag{2.17}$$

Passiamo ora alla dimostrazione del Teorema 2.1, ossia della formula di Black e Scholes.

2.2.2 Dimostrazione della formula di Black e Scholes

Grazie al fatto che (almeno per n sufficientemente grande) il modello di mercato considerato è completo e privo di opportunità di arbitraggio, sappiamo che la misura martingala equivalente $\widetilde{\mathbb{P}}^{(n)}$ esiste ed è unica, e che, rispetto alla misura martingala equivalente $\widetilde{\mathbb{P}}^{(n)}$, gli eventi $\{Z_i^{(n)} = u\}$ sono indipendenti e con probabilità

$$\tilde{p}^{(n)} = \frac{1+\frac{r}{n} - d^{(n)}}{u^{(n)} - d^{(n)}} = \frac{\frac{r}{n} - a^{(n)}}{b^{(n)} - a^{(n)}} = \frac{\frac{r}{n} - (-\frac{\sigma}{\sqrt{n}} + \frac{1}{2}\frac{\sigma^2}{n} + o(\frac{1}{n}))}{\frac{\sigma}{\sqrt{n}} + \frac{1}{2}\frac{\sigma^2}{n} + o(\frac{1}{n}) - (-\frac{\sigma}{\sqrt{n}} + \frac{1}{2}\frac{\sigma^2}{n} + o(\frac{1}{n}))},$$

o meglio

$$\begin{aligned}\tilde{p}^{(n)} &= \frac{\frac{\sigma}{\sqrt{n}} + \frac{1}{n}(r - \frac{1}{2}\sigma^2) + o(\frac{1}{n})}{2\frac{\sigma}{\sqrt{n}} + \frac{1}{2}\frac{\sigma^2}{n} - \frac{1}{2}\frac{\sigma^2}{n} + o(\frac{1}{n})} = \frac{\frac{\sigma}{\sqrt{n}} + \frac{1}{n}(r - \frac{1}{2}\sigma^2) + o(\frac{1}{n})}{2\frac{\sigma}{\sqrt{n}} + o(\frac{1}{n})}\\ &= \frac{1}{2} + \frac{1}{2\sqrt{n}}\,\frac{r - \frac{1}{2}\sigma^2}{\sigma} + o(\tfrac{1}{\sqrt{n}}) = \frac{1}{2} + \frac{1}{2\sqrt{n}}\,\theta_n,\end{aligned}$$

dove $\theta_n = (r - \frac{1}{2}\sigma^2)/\sigma + o(1)$ converge a $(r - \frac{1}{2}\sigma^2)/\sigma$ per n che tende all'infinito.

Equivalentemente le variabili aleatorie ξ_i sono indipendenti e di valore atteso

$$\widetilde{\mathbb{E}}^{(n)}[\xi_i] = \tilde{p}^{(n)} = \frac{1}{2} + \frac{1}{2\sqrt{n}}\,\frac{r - \frac{1}{2}\sigma^2}{\sigma} + o(\tfrac{1}{\sqrt{n}}).$$

Di conseguenza il prezzo di un derivato con maturità (tempo di esercizio) T e con pay-off terminale (contingent claim)

$$f(S_T^{(n)})$$

dovrà necessariamente avere come prezzo

$$C^{(n)}(f, \mathbb{P}) = \widetilde{\mathbb{E}}^{(n)}\left[\frac{f(S_T^{(n)})}{B_T^{(n)}}\right] \tag{2.18}$$

in particolare per l'opzione call

$$C_{call}^{(n)}(K, \mathbb{P}) = \widetilde{\mathbb{E}}^{(n)}\left[\frac{(S_T^{(n)} - K)^+}{B_T^{(n)}}\right] = \widetilde{\mathbb{E}}^{(n)}\left[\frac{(S_{N_n(T)} - K)^+}{\left(1+\frac{r}{n}\right)^{N_n(T)}}\right] \tag{2.19}$$

Abbiamo quindi un'espressione del prezzo, ma il problema a questo punto diviene complesso dal punto di vista numerico, almeno per n grande: il denominatore non comporta problemi in quanto si può approssimare con e^{rT}, mentre lo stesso non si può dire del numeratore[5].

[5]**La dimostrazione della formula di Black e Scholes potrebbe tuttavia essere ottenuta direttamente dalla formula del prezzo della call per il modello binomiale multiperiodale: Infatti sappiamo che

$$C_{call}^{(n)}(K, \mathbb{P}) = s_0\, \overline{F}_{N_n(T), \tilde{p}^{(n)}\,\frac{u^{(n)}}{1+\frac{r}{n}}}(\zeta_n) - \frac{K}{(1+\frac{r}{n})^{N_n(T)}}\,\overline{F}_{N_n(T), \tilde{p}^{(n)}}(\zeta_n)$$

dove $\overline{F}_{m,p}(x)$ è la funzione di sopravvivenza di una variabile aleatoria binomiale di parametri m e p, e dove

$$\zeta_n = \frac{\ln\left(\frac{K}{S_0 (d^{(n)})^{N_n(T)}}\right)}{\ln\left(\frac{u^{(n)}}{d^{(n)}}\right)} = \frac{\sqrt{n}}{2\sigma}\ln\left(\frac{K}{S_0}\right) + \frac{\sqrt{n}}{2\sigma}\,\frac{\sigma}{\sqrt{n}}\lfloor nT\rfloor = \frac{\sqrt{n}}{2\sigma}\ln\left(\frac{K}{S_0}\right) + \frac{\lfloor nT\rfloor}{2}.$$

Tenendo conto che il teorema centrale del limite assicura che, per m grande (e p non troppo piccolo) si ha

$$\overline{F}_{m,p}(x) \simeq \Phi\Big(\frac{x - mp}{\sqrt{m\,p(1-p)}}\Big),$$

si ottiene che il prezzo della call può essere approssimato con

$$C_{call}^{(n)}(K, \mathbb{P}) \simeq s_0\, \overline{\Phi}\Bigg(\frac{\zeta_n - \lfloor nT\rfloor\, \tilde{p}^{(n)}\, \frac{u^{(n)}}{1+\frac{r}{n}}}{\sqrt{\lfloor nT\rfloor \tilde{p}^{(n)}\,\frac{u^{(n)}}{1+\frac{r}{n}}(1 - \tilde{p}^{(n)})\,\frac{u^{(n)}}{1+\frac{r}{n}}}}\Bigg) - \frac{K}{(1+\frac{r}{n})^{\lfloor nT\rfloor}}\,\overline{\Phi}\Bigg(\frac{\zeta_n - \lfloor nT\rfloor\, \tilde{p}^{(n)}}{\sqrt{\lfloor nT\rfloor \tilde{p}^{(n)}(1 - \tilde{p}^{(n)})}}\Bigg).$$

Per risolvere questo problema ci viene in aiuto il Teorema Centrale del Limite. Si noti infatti che, qualunque sia $t > 0$

$$\log S_t^{(n)} = \log S_0 + H_{N_n(t)} \log\left(\frac{u^{(n)}}{d^{(n)}}\right) + N_n(t) \log\left(d^{(n)}\right) = \log S_0 + \sum_{i=1}^{N_n(t)} \xi_i \log\left(\frac{u^{(n)}}{d^{(n)}}\right) + N_n(t) \log\left(d^{(n)}\right),$$

e che $H_{N_n(t)}$, rispetto alla misura martingala equivalente $\widetilde{\mathbb{P}}^{(n)}$, è la somma di variabili aleatorie indipendenti tutte con la stessa distribuzione, che per $t > 0$ il numero $N_n(t) = \lfloor nt \rfloor$ converge all'infinito. Grazie al Teorema Centrale del Limite si ha che quindi $H_{N_n(t)}$ ha una distribuzione ***approssimativamente*** gaussiana, di valore atteso

$$\widetilde{\mathbb{E}}^{(n)}[H_{N_n(t)}] = N_n(t)\,\tilde{p}^{(n)}$$

e varianza

$$\widetilde{Var}^{(n)}(H_{N_n(t)}) = N_n(t)\,\tilde{p}^{(n)}(1 - \tilde{p}^{(n)}).$$

Per lo stesso motivo[6], sempre rispetto alla misura martingala equivalente $\widetilde{\mathbb{P}}^{(n)}$, anche il logaritmo di $S_t^{(n)}$ è approssimata da una variabile aleatoria gaussiana di valore atteso

$$\widetilde{\mathbb{E}}^{(n)}\left[\log\left(S_t^{(n)}\right)\right] = \log S_0 + N_n(t)\,\tilde{p}^{(n)} \log\left(\frac{u^{(n)}}{d^{(n)}}\right) + N_n(t) \log\left(d^{(n)}\right)$$

e varianza

$$\widetilde{Var}^{(n)}\left(\log\left(S_t^{(n)}\right)\right) = N_n(t)\,\tilde{p}^{(n)}(1 - \tilde{p}^{(n)})\left(\log\left(\frac{u^{(n)}}{d^{(n)}}\right)\right)^2.$$

Ricordando che

$$\frac{u^{(n)}}{d^{(n)}} = \frac{e^{\sigma/\sqrt{n}}}{e^{-\sigma/\sqrt{n}}} = e^{2\sigma/\sqrt{n}}, \qquad d^{(n)} = e^{-\sigma/\sqrt{n}},$$

si ha

$$\log\left(\frac{u^{(n)}}{d^{(n)}}\right) = 2\,\frac{\sigma}{\sqrt{n}} \qquad \log\left(d^{(n)}\right) = -\frac{\sigma}{\sqrt{n}},$$

Per ottenere la formula di Black e Scholes basta poi osservare che $(1 + \frac{r}{n})^{\lfloor nT \rfloor}$ converge a e^{rT}, che

$$\begin{aligned}
\frac{\zeta_n - \lfloor nT \rfloor\,\widetilde{p}^{(n)}}{\sqrt{\lfloor nT \rfloor \widetilde{p}^{(n)}(1 - \widetilde{p}^{(n)})}} &= \frac{\frac{\sqrt{n}}{2\sigma} \ln\left(\frac{K}{S_0}\right) + \frac{\lfloor nT \rfloor}{2} - \lfloor nT \rfloor\left(\frac{1}{2} + \frac{\theta_n}{2\sqrt{n}}\right)}{\sqrt{\lfloor nT \rfloor\left(\frac{1}{2} + \frac{\theta_n}{2\sqrt{n}}\right)\left(1 - \left(\frac{1}{2} + \frac{\theta_n}{2\sqrt{n}}\right)\right)}} \\
&= \ln\left(\frac{K}{S_0}\right) \frac{\frac{\sqrt{n}}{2\sigma}}{\sqrt{\lfloor nT \rfloor\left(\frac{1}{4} - \frac{\theta_n^2}{4}\right)}} - \frac{\lfloor nT \rfloor\,\frac{\theta_n}{2\sqrt{n}}}{\sqrt{\lfloor nT \rfloor\left(\frac{1}{4} - \frac{\theta_n^2}{4}\right)}} \\
&\simeq \ln\left(\frac{K}{S_0}\right) \frac{1}{\sigma\sqrt{T}} - \sqrt{T} \lim_n \theta_n = \ln\left(\frac{K}{S_0}\right) \frac{1}{\sigma\sqrt{T}} - \sqrt{T}\frac{r - \frac{1}{2}\sigma^2}{\sigma} \\
&= \ln\left(\frac{K}{S_0}\right) \frac{1}{\sigma\sqrt{T}} - \frac{rT}{\sigma\sqrt{T}} + \frac{\sigma\sqrt{T}}{2} = \frac{\ln\left(\frac{K}{S_0\,e^{rT}}\right)}{\sigma\sqrt{T}} + \frac{\sigma\sqrt{T}}{2},
\end{aligned}$$

e infine che

$$\frac{\zeta_n - \lfloor nT \rfloor\,\widetilde{p}^{(n)}\,\frac{u^{(n)}}{1+\frac{r}{n}}}{\sqrt{\lfloor nT \rfloor \widetilde{p}^{(n)}\,\frac{u^{(n)}}{1+\frac{r}{n}}(1 - \widetilde{p}^{(n)})\,\frac{u^{(n)}}{1+\frac{r}{n}}}} \simeq \frac{\zeta_n - \lfloor nT \rfloor\,\widetilde{p}^{(n)}\,u^{(n)}}{\sqrt{\lfloor nT \rfloor \widetilde{p}^{(n)}\,u^{(n)}(1 - \widetilde{p}^{(n)})\,u^{(n)}}} = \frac{\zeta_n - \lfloor nT \rfloor\left(\frac{1}{2} + \frac{\theta_n}{2\sqrt{n}}\right)\left(1 + \frac{\sigma}{\sqrt{n}} + o(\frac{1}{\sqrt{n}})\right)}{\sqrt{\lfloor nT \rfloor \widetilde{p}^{(n)}\,(1 - \widetilde{p}^{(n)})}\,\sqrt{u^{(n)}}}$$

quindi definendo $\widetilde{\theta}_n$ in modo che $\left(\frac{1}{2} + \frac{\theta_n}{2\sqrt{n}}\right)\left(1 + \frac{\sigma}{\sqrt{n}} + o(\frac{1}{\sqrt{n}})\right) = \frac{1}{2} + \frac{\widetilde{\theta}_n}{2\sqrt{n}}$ si ha che $\widetilde{\theta}_n := \theta_n + 2\sqrt{n}(u^{(n)} - 1) + o(1) \simeq \theta_n + \sigma$, in modo che $\lim_n \widetilde{\theta}_n = \frac{r - \frac{\sigma^2}{2}}{\sigma} + \sigma = \frac{r + \frac{\sigma^2}{2}}{\sigma}$, e tenendo conto che $u^{(n)}$ converge ad uno, si ottiene, con conti del tutto analoghi ai precedenti che

$$= \frac{\zeta_n - \lfloor nT \rfloor\left(\frac{1}{2} + \frac{\widetilde{\theta}_n}{2\sqrt{n}}\right)}{\sqrt{\lfloor nT \rfloor \widetilde{p}^{(n)}\,(1 - \widetilde{p}^{(n)})}}\,\frac{1}{\sqrt{u^{(n)}}} \simeq \frac{\ln\left(\frac{K}{S_0\,e^{rT}}\right)}{\sigma\sqrt{T}} - \frac{\sigma\sqrt{T}}{2}.$$

[6]Si ricordi che se Z ha distribuzione $N(\alpha, \beta^2)$, ovvero distribuzione gaussiana di valore atteso α e varianza β^2, allora anche $W = a + bZ$ ha distribuzione gaussiana, ma di valore atteso $\mathbb{E}[W] = \mathbb{E}[a + bZ] = a + b\alpha$, e varianza $Var(W) = Var(a + bZ) = Var(bZ) = b^2 Var(Z) = b^2\beta^2$.

si ottiene che il valore atteso del logaritmo di $S_t^{(n)}$, vale

$$\begin{aligned}
\widetilde{\mathbb{E}}^{(n)}\left[\log\left(S_t^{(n)}\right)\right] &= \log S_0 + \lfloor nt \rfloor \left(\frac{1}{2} + \frac{1}{2\sqrt{n}}\,\frac{r-\frac{1}{2}\sigma^2}{\sigma} + o(\tfrac{1}{\sqrt{n}})\right) 2\,\frac{\sigma}{\sqrt{n}} + \lfloor nt \rfloor \left(-\frac{\sigma}{\sqrt{n}}\right)\\
&= \log S_0 + \lfloor nt \rfloor \left(\frac{1}{2}\,2\,\frac{\sigma}{\sqrt{n}} + \frac{1}{2\sqrt{n}}\,\frac{r-\frac{1}{2}\sigma^2}{\sigma}\,2\,\frac{\sigma}{\sqrt{n}} + o(\tfrac{1}{\sqrt{n}})\,2\,\frac{\sigma}{\sqrt{n}} - \frac{\sigma}{\sqrt{n}}\right)\\
&= \log S_0 + \lfloor nt \rfloor \left(\frac{1}{n}\,\left(r-\tfrac{1}{2}\sigma^2\right) + o(\tfrac{1}{\sqrt{n}})\,2\,\frac{\sigma}{\sqrt{n}}\right)\\
&= \log S_0 + \frac{\lfloor nt \rfloor}{n}\,\left(r-\tfrac{1}{2}\sigma^2\right) \approx \log S_0 + t\left(r-\tfrac{1}{2}\sigma^2\right).
\end{aligned}$$

Analogamente la varianza vale

$$\begin{aligned}
\widetilde{Var}^{(n)}\left(\log\left(S_t^{(n)}\right)\right) &= N_n(t)\,\tilde{p}^{(n)}(1-\tilde{p}^{(n)})\left(\log\left(\frac{u^{(n)}}{d^{(n)}}\right)\right)^2\\
&= \lfloor nt \rfloor \left(\frac{1}{2} + \frac{1}{2\sqrt{n}}\,\frac{r-\frac{1}{2}\sigma^2}{\sigma} + o(\tfrac{1}{\sqrt{n}})\right)\left(1 - \left(\frac{1}{2} + \frac{1}{2\sqrt{n}}\,\frac{r-\frac{1}{2}\sigma^2}{\sigma} + o(\tfrac{1}{\sqrt{n}})\right)\right)\left(\frac{2\sigma}{\sqrt{n}}\right)^2\\
&= \lfloor nt \rfloor \left(\frac{1}{2} + \frac{1}{2\sqrt{n}}\,\theta_n\right)\left(\frac{1}{2} - \frac{1}{2\sqrt{n}}\,\theta_n\right)\frac{4\sigma^2}{n} = \lfloor nt \rfloor \left(\frac{1}{4} - \frac{1}{4\,n}\,\theta_n^2\right)\frac{4\sigma^2}{n}\\
&\approx \lfloor nt \rfloor\,\frac{1}{4}\,\frac{4\sigma^2}{n} \approx t\sigma^2
\end{aligned}$$

Di conseguenza, se W_t è una variabile aleatoria gaussiana di valore atteso 0 e varianza t

$$\log(S_t^{(n)}) \overset{distr}{\to} \log S_0 + t\left(r-\tfrac{1}{2}\sigma^2\right) + \sigma\,W_t. \tag{2.20}$$

Osservazione 2.2. *Va notato che il precedente risultato si potrebbe ottenere anche direttamente, dimostrando che la funzione caratteristica della variabile aleatoria* $\log S_t^{(n)}$ *converge alla funzione caratteristica di* $\log(s_0) + r\,t - \frac{1}{2}\sigma^2\,t + \sigma\,W_t$*, ossia che*

$$\lim_{n\to\infty}\widetilde{\mathbb{E}}^{(n)}\left[\exp\{i\,u\log\left(S_t^{(n)}\right)\}\right] = e^{iu\left(\log(s_0\,e^{rt}) - \frac{1}{2}\sigma^2\,t\right) + \frac{1}{2}\sigma^2\,t\,u^2}.$$

In modo ancora più semplice si noti che dalla (2.17)*, si ottiene*

$$S_t^{(n)} = S_0\,e^{\sigma\frac{\sqrt{\lfloor nt \rfloor}}{\sqrt{nt}}\,\frac{1}{\sqrt{\lfloor nt \rfloor}}\left(2\,H_{\lfloor nt \rfloor} - \lfloor nt \rfloor\right)}.$$

Ovviamente $\frac{\sqrt{\lfloor nt \rfloor}}{\sqrt{nt}}$ *converge ad* 1*, e per ottenere la convergenza basterebbe dimostrare direttamente che la funzione caratteristica della variabile aleatoria*

$$\frac{1}{\sqrt{\lfloor nt \rfloor}}\left(2\,H_{\lfloor nt \rfloor} - \lfloor nt \rfloor\right)$$

converge alla funzione caratteristica di $\frac{1}{\sigma}\left(r\,t - \frac{1}{2}\sigma^2\,t + \sigma\,W_t\right) = \frac{r-\frac{1}{2}\sigma^2}{\sigma}\,t + W_t$*, **ossia che*

$$\lim_{n\to\infty}\widetilde{\mathbb{E}}^{(n)}\left[\exp\left\{i\,u\log\left(S_t^{(n)}\right)\right\}\right] = e^{iu\left(r\,t - \frac{1}{2}\sigma^2\,t\right) + \frac{1}{2}\sigma^2\,t\,u^2}.$$

La dimostrazione sarebbe sostanzialmente la stessa di quella del teorema centrale del limite, in quanto possiamo scrivere

$$2H_k - k = \sum_{j=1}^{k}(2\,\xi_j - 1)\Big(= 2\,H_k^{[n]} - k) = \sum_{j=1}^{k}(2\,\xi_j^{[n]} - 1)\Big)$$

come somma di variabili aleatorie indipendenti ed identicamente distribuite. L'unica differenza risiede nel fatto che le variabili aleatorie $2\,\xi_j - 1\big(= 2\,\xi_j^{[n]} - 1\big)$*, hanno distribuzione che dipende da* n*: esattamente* $2\,\xi_j^{[n]} - 1$ *assumono i* *valori +1 e −1 con probabilità $\tilde{p}^{(n)}$ e $1 - \tilde{p}^{(n)}$, rispettivamente*[7]*. Il calcolo del limite della funzione caratteristica in questo caso richiede quindi un minimo di attenzione in più.*

Va infine notato che, nel caso in cui invece si avesse più semplicemente il caso di variabili aleatorie $X_j := 2\,\xi_j - 1$, indipendenti e che assumono i valori +1 e −1 con probabilità 1/2 (e quindi a valore atteso nullo), il teorema centrale del limite ci darebbe direttamente che

$$\frac{1}{\sqrt{\lfloor nt \rfloor}}\big(2\,H_{\lfloor nt \rfloor} - \lfloor nt \rfloor\big) = \frac{1}{\sqrt{\lfloor nt \rfloor}}\sum_{j=1}^{\lfloor nt \rfloor}(2\,\xi_j - 1)$$

converge in distribuzione ad una variabile aleatoria gaussiana standard e quindi

$$\frac{1}{\sqrt{n}}\big(2\,H_{\lfloor nt \rfloor} - \lfloor nt \rfloor\big) = \frac{1}{\sqrt{n}}\sum_{j=1}^{\lfloor nt \rfloor}(2\,\xi_j - 1)$$

converge ad una variabile aleatoria gaussiana di valore atteso nullo e varianza t.

Alternativamente, nel caso in cui invece la probabilità che $2\,\xi_j - 1 = +1$ fosse ancora $\tilde{p}^{(n)}(\neq 1/2)$ si potrebbe sottrarre ed aggiungere il valore atteso di $2\,\xi_j - 1$, ossia $2\,\tilde{p}^{(n)} - 1$ nella somma, ottenendo

$$\frac{1}{\sqrt{n}}\sum_{j=1}^{\lfloor nt \rfloor}(2\,\xi_j - 1) = \frac{1}{\sqrt{n}}\sum_{j=1}^{\lfloor nt \rfloor}\big(2\,\xi_j - 1 - (2\,\tilde{p}^{(n)} - 1)\big) + \frac{1}{\sqrt{n}}\sum_{j=1}^{\lfloor nt \rfloor}(2\,\tilde{p}^{(n)} - 1).$$

La prima somma converge allora ad una variabile aleatoria gaussiana a valore medio nullo, mentre si potrebbe dimostrare che la seconda somma converge a $rt - \frac{1}{2}\sigma^2\,t$. Questa osservazione è alla base dell'idea che permette di costruire il moto browniano standard. Ad esso è dedicata la sezione successiva.

Prima di proseguire nella dimostrazione del Teorema 2.1, dobbiamo ricordare che il simbolo $X_n \overset{distr}{\to}_{n\to\infty} X$ significa che la successione di variabili aleatorie X_n converge in distribuzione (o in legge) alla variabile aleatoria X. Non è necessario che le variabili aleatorie X_n ed X vivano sullo stesso spazio di probabilità, ma si può supporre che, per ogni n, la variabile aleatoria X_n sia definita nello spazio di probabilità $(\Omega^n, \mathcal{F}^n, \mathbf{P^n})$ e X sia definita nello spazio di probabilità $(\Omega, \mathcal{F}, \mathbf{P})$. La convergenza in distribuzione significa la convergenza di $F_n(x) := \mathbf{P^n}(X_n \leq x)$ a $F(x) := \mathbf{P}(X \leq x)$ per ogni x tale che $F(x) = F(x^-)$, e che ciò è equivalente al fatto che[8] per ogni funzione continua e limitata f

$$\lim_{n\to\infty} \mathbf{E^n}[f(X_n)] = \mathbf{E}[f(X)].$$

In altre parole che si può approssimare $\mathbf{E^n}[f(X_n)]$ con $\mathbf{E}[f(X)]$. Dopo questo richiamo possiamo tornare alla nostra dimostrazione.

A questo punto il prezzo di una opzione europea si può calcolare come

$$C^{(n)}(f, \mathbb{P}) = \widetilde{\mathbb{E}}^{(n)}\left[\frac{f(S_T^{(n)})}{B_T^{(n)}}\right] \approx \mathbf{E}\left[\frac{1}{e^{rT}}\,f(S_0 e^{(r - \frac{1}{2}\,\sigma^2)T + \sigma\,W_T})\right] \tag{2.21}$$

ed in particolare per l'opzione call

$$C_{call}^{(n)}(K, \mathbb{P}) = \widetilde{\mathbb{E}}^{(n)}\left[\frac{(S_T^{(n)} - K)^+}{\big(1 + \frac{r}{n}\big)^{N_n(T)}}\right] \approx \mathbf{E}\left[\frac{1}{e^{rT}}\,(S_0 e^{(r - \frac{1}{2}\,\sigma^2)T + \sigma\,W_T} - K)^+\right], \tag{2.22}$$

dove l'unica variabile aleatoria è W_T (che sotto $\mathbf{P}$ ha distribuzione gaussiana di media 0 e varianza T), in quanto $S_0 = s_0$ è il valore iniziale del prezzo dell'azione.

[7] Si osservi che $2\,x - 1 = 2 \cdot 1 - 1 = 1$ se $x = +1$ e che $2\,x - 1 = 2 \cdot 0 - 1 = -1$ se $x = 0$.

[8] Di solito nei corsi di base di probabilità si usa l'equivalenza con la convergenza per $f(x) = \exp i\,t\,x = \cos(tx) + i\,\sin(tx)$, e che corrisponde alla scelta di $f(x) = \cos(tx)$ o $f(x) = \sin(tx)$.

Abbiamo quindi visto come il prezzo di una opzione call europea si possa ottenere dalla formula

$$C_{call} = C(s_0, K, T, r, \sigma) = \mathbb{E}\left[e^{-rT}\,(s_0 e^{(r-\frac{1}{2}\,\sigma^2)T+\sigma\, W_T} - K)^+\right]$$

dove si è messo in evidenza la dipendenza dai parametri del modello: S_0 prezzo iniziale della azione, K prezzo di esercizio i di strike dell'opzione, T tempo di maturità o di strike dell'opzione, r tasso nominale di interesse composto in modo continuo, ed infine il parametro σ, che è detto volatilità.

È importante sottolineare che, al contrario di tutti gli altri parametri, che sono noti e direttamente osservabili, il valore della volatilità non è direttamente osservabile, ma deve essere stimato. Un problema interessante riguarda proprio la stima statistica della volatilità.

Si osservi che fino ad ora abbiamo addirittura trovato una formula che permette di calcolare in modo approssimato tutti i tipi di opzioni plain vanilla di tipo europeo, con la condizione che la funzione f sia continua[9]

Per terminare la dimostrazione rimane solamente da calcolare esplicitamente il valore limite del prezzo della call europea.

A questo proposito notiamo che, per ogni t, la legge della variabile aleatoria $S_t^{(n)}$ rispetto a $\mathbb{P}^{(n)}$ è approssimativamente la stessa[10] della variabile aleatoria $s_0 \exp\{(r-\sigma^2/2)t + \sigma\sqrt{t}Z\}$, dove Z è una variabile aleatoria gaussiana standard $N(0,1)$.

Quindi, per $s_0 = x$, si tratta di calcolare

$$\begin{aligned} C_0(x) &= e^{-rT}\mathbf{E}\left[(S_T-K)^+\right] \\ &= e^{-rT}\mathbf{E}\left[\left(xe^{(r-\frac{\sigma^2}{2})T+\sigma\sqrt{T}Z} - K\right)^+\right]. \end{aligned}$$

La speranza matematica a destra vale

$$e^{-rT}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\left(xe^{(r-\frac{\sigma^2}{2})T+\sigma\sqrt{T}z} - K\right)^+ e^{-z^2/2}dz.$$

L'integrando si annulla per $z \leq \zeta$, dove

$$\zeta = \frac{1}{\sigma\sqrt{T}}\left(\log\left(\tfrac{K}{x}\right) - \left(r - \frac{\sigma^2}{2}\right)T\right),$$

e quindi

$$C_0(x) = \frac{e^{-rT}}{\sqrt{2\pi}}\int_{\zeta}^{+\infty}\left(xe^{(r-\frac{\sigma^2}{2})T+\sigma\sqrt{T}z} - K\right)e^{-z^2/2}dz.$$

Ricordando che per la funzione di distribuzione di una gaussiana standard si ha $1-\Phi(x) = \Phi(-x)$, si ottiene allora

$$\begin{aligned} C_0(x) &= \frac{x}{\sqrt{2\pi}}\int_{\zeta}^{+\infty} e^{-rT}\,e^{(r-\frac{\sigma^2}{2})T+\sigma\sqrt{T}z}e^{-z^2/2}dz - \frac{K\,e^{-rT}}{\sqrt{2\pi}}\int_{\zeta}^{+\infty} e^{-z^2/2}dz \\ &= \frac{x}{\sqrt{2\pi}}\int_{\zeta}^{+\infty} e^{-\frac{1}{2}(z-\sigma\sqrt{T})^2}dz - Ke^{-rT}\Phi(-\zeta) \\ &= \frac{x}{\sqrt{2\pi}}\int_{\zeta-\sigma\sqrt{T}}^{+\infty} e^{-z^2/2}dz - Ke^{-rT}\Phi(-\zeta) \\ &= x\Phi\left(-\zeta+\sigma\sqrt{T}\right) - Ke^{-rT}\Phi(-\zeta), \end{aligned}$$

[9] In realtà la condizione che f sia una funzione continua è superflua.

[10] L'uguaglianza in legge si vede dall'espressione approssimata del logaritmo del prezzo (2.20), e tenendo conto del fatto che la legge della variabile aleatoria $\widetilde{W}_t$ è $N(0,t)$ e che anche la variabile aleatoria $\sqrt{t}Z$ ha legge $N(0,t)$, se Z ha legge $N(0,1)$. È importante sottolineare che ovviamente ciò vale solo come variabili aleatorie e non vale come processi (ossia il processo $(W_t)_{t\geq 0}$ non coincide affatto con il processo $(\sqrt{t}\,Z)_{t\geq 0}$: infatti le traiettorie tipiche del processo $(W_t)_{t\geq 0}$ non sono molto regolari ne' prevedibili, e sono quindi molto diverse dalle traiettorie di $(\sqrt{t}\,Z)_{t\geq 0}$, che invece sono molto regolari e prevedibili.

la così detta *formula di Black-Scholes* (2.10).

Mediante la formula di Black-Scholes, possiamo quindi ricavare il prezzo equo di un'opzione call europea. La formula di Black-Scholes ha il pregio di essere semplice e di dipendere da tre parametri: r, μ e σ. L'unico parametro difficile da stimare è σ^2, cioè la volatilità[11].

Va inoltre notato che poiché a tempo discreto vale l'interpretazione di $c_N(n,f) = c_{N-n}(0,f)$, analogamente si ottiene che, se entrassimo nel mercato al tempo t, e se il prezzo della azione sottostante al tempo t fosse noto ed uguale ad S_t, allora il prezzo della opzione, che indichiamo con C_t, sarebbe calcolato approssimativamente con la formula (2.10) in cui, però, al posto di s_0 andrebbe messo S_t e al posto di T andrebbe messo $T-t$:

$$C_t = S_t \left[\Phi\left(\tfrac{1}{\sigma\sqrt{T-t}} \log\left(\tfrac{x\, e^{r(T-t)}}{K}\right) + \tfrac{\sigma\sqrt{T-t}}{2}\right) - \tfrac{K}{x\, e^{r(T-t)}}\, \Phi\left(\tfrac{1}{\sigma\sqrt{T-t}} \log\left(\tfrac{x\, e^{r(T-t)}}{K}\right) - \tfrac{\sigma\sqrt{T-t}}{2}\right)\right]\Big|_{x=S_t}.$$

Grazie alla formula di parità (si veda ad esempio il libro di S. Ross [14]), si ricava immediatamente anche il prezzo equo di una put europea P_t è dato da

$$P_t = C_t - S_t + Ke^{-r(T-t)},$$

sempre se si vuole comprare l'opzione al tempo t ed il prezzo della azione vale S_t.

Per ulteriori approfondimenti sai consiglia di consultare il libro di P. Baldi [2], o quello di D. Lamberton e B. Lapeyre [11], oppure di J.M. Steele [19] .

Terminiamo con l'osservare che in analogia a quanto detto in Osservazione 1.12, una volta calcolato il prezzo delle call europee nel modello di Black e Scholes, è possibile ottenere il prezzo delle opzioni con terminal payoff $f(S_T)$, ossia funzione deterministica del valore del prezzo del sottostante al tempo T, nel caso in cui, $f(x)$ ammetta una rappresentazione del tipo (1.35), ossia (la riportiamo per comodità del lettore)

$$f(x) = f(0) + xf'(0) + \int_{(0,x]} (x-y)^+ \mu(dy) = f(0) + xf'(0) + \int_{(0,\infty)} (x-y)^+ \mu(dy).$$

Coem ricordato in precedenza in una nota esempi di funzioni di questo tipo sono le funzioni derivabili due volte, con derivate seconde continue (si tratta di sostituire a $\mu(dy)$ l'espressione $f^{''}(y)dy$) o combinazioni di funzioni call, come ad esempio,

$$f(x) = (x-K_1)^+ - (x-K_2)^+, \qquad K_1 \neq K_2,$$

per le quali non c'è bisogno nemmeno di individuare la misura $\mu(dy)$.

2.3 Il moto Browniano

2.3.1 Approssimazione del moto browniano per t fissato

Nella derivazione precedente del prezzo abbiamo incontrato il processo del logaritmo dei prezzi, che a parte il contributo dovuto al prezzo iniziale, si esprime come

$$\sum_{i=1}^{N_n(t)} \xi_i \log\left(\frac{u^{(n)}}{d^{(n)}}\right) + N_n(t) \log\left(d^{(n)}\right) = 2\sigma \frac{1}{\sqrt{n}} \sum_{i=1}^{N_n(t)} \xi_i - N_n(t)\, \sigma \frac{1}{\sqrt{n}}$$

e che si può ulteriormente riscrivere come

$$\sigma \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{N_n(t)} (2\xi_i - 1)\right) = \sigma \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} (2\xi_i - 1)\right).$$

[11]La volatilità è un parametro che gioca un ruolo importante nelle applicazioni. Per questo motivo, negli ultimi anni, è stato molto studiato, in statistica, il problema di stimare il coefficiente di diffusione, a partire dall'osservazione di una traiettoria.

Si definisca

$$W_t^{(n)} := \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} \left[(2\xi_i - 1) - \widetilde{\mathbb{E}}^{(n)}[2\xi_i - 1]\right] \tag{2.23}$$

di modo che

$$\log\left(S_t^{(n)}\right) = \log\left(S_0\right) + \sigma W_t^{(n)} + \sigma\,\frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} \widetilde{\mathbb{E}}^{(n)}[2\xi_i - 1] \tag{2.24}$$

Con calcoli analoghi a quelli della sezione precedente, si può vedere che il valore atteso

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} \widetilde{\mathbb{E}}^{(n)}[2\xi_i - 1] = \frac{1}{\sqrt{n}} \lfloor nt \rfloor \,(2\tilde{p}^{(n)} - 1) = \frac{1}{\sqrt{n}} \lfloor nt \rfloor \left(1 + \frac{1}{\sqrt{n}}\,\frac{r - \frac{1}{2}\sigma^2}{\sigma} + o\big(\tfrac{1}{\sqrt{n}}\big) - 1\right) \to_{n\to\infty} \frac{r - \frac{1}{2}\sigma^2}{\sigma}\,t.$$

Ovviamente il processo $W_t^{(n)}$, vale zero all'istante iniziale, ovvero

$$W_0^{(n)} = 0,$$

ed inoltre

$$\widetilde{\mathbb{E}}^{(n)}\left[W_t^{(n)}\right] = 0.$$

Per $t > 0$, con gli stessi calcoli della sezione precedente, si può vedere che il processo $W_t^{(n)}$ converge (per n che tende all'infinito) alla legge gaussiana di valore atteso nullo e varianza che tende a t.

Osservazione 2.3. *Se si vuole ottenere una dimostrazione più precisa della convergenza in distribuzione di $W_t^{(n)}$ a W_t, si può ripetere lo stesso tipo di dimostrazione che si usa per il Teorema Centrale del Limite: tenendo conto del fatto che, definite le variabili aleatorie*

$$X_j := 2\xi_j - 1 \quad \textit{ed} \quad \widetilde{X}_j^{(n)} := \frac{1}{\sqrt{n}}\left[X_j - \widetilde{\mathbb{E}}^{(n)}[X_j]\right],$$

si può riscrivere (2.23) *come*

$$\begin{aligned} W_t^{(n)} &= \frac{1}{\sqrt{n}} \sum_{j=1}^{\lfloor nt \rfloor} \left[X_j - \widetilde{\mathbb{E}}^{(n)}[X_j]\right] \\ &= \sum_{j=1}^{\lfloor nt \rfloor} \frac{1}{\sqrt{n}}\left[X_j - \widetilde{\mathbb{E}}^{(n)}[X_j]\right] = \sum_{j=1}^{\lfloor nt \rfloor} \widetilde{X}_j^{(n)} \end{aligned}$$

e quindi, dato che

$$\exp\{i\,u\,W_t^{(n)}\} = \prod_{j=1}^{\lfloor nt \rfloor} \exp\{i\,u\,\widetilde{X}_j^{(n)}\}$$

e che le variabili aleatorie $\widetilde{X}_j^{(n)}$, $j \geq 1$ sono indipendenti, si ottiene che

$$\widetilde{\mathbb{E}}^{(n)}\left[\exp\{i\,u\,W_t^{(n)}\}\right] = \prod_{j=1}^{\lfloor nt \rfloor} \widetilde{\mathbb{E}}^{(n)}\left[\exp\{i\,u\,\widetilde{X}_j^{(n)}\}\right]. \tag{2.25}$$

Le variabili aleatorie $\widetilde{X}_j^{(n)}$ hanno media nulla e varianza

$$\begin{aligned} \widetilde{Var}^{(n)}\left[\widetilde{X}_j^{(n)}\right] &= \widetilde{Var}^{(n)}\left[\frac{1}{\sqrt{n}}\,2\,\xi_j\right] = \left(\frac{1}{\sqrt{n}}\,2\right)^2 \tilde{p}^{(n)}(1 - \tilde{p}^{(n)}) \\ &= \frac{4}{n}\left(\frac{1}{2} + \frac{1}{2\sqrt{n}}\,\theta_n\right)\left(\frac{1}{2} - \frac{1}{2\sqrt{n}}\,\theta_n\right) \end{aligned}$$

dove $\theta_n = (r - \frac{1}{2}\sigma^2)/\sigma + o(1)$, *di modo che*

$$\widetilde{Var}^{(n)}[\widetilde{X}_j^{(n)}] = \frac{1}{n} + o(\frac{1}{n}).$$

Per ottenere la convergenza in distribuzione non rimane che utilizzare tale fatto per ottenere l'espressione approssimata della funzione caratteristica di $\widetilde{X}_j^{(n)}$*, sostituire tale espressione in* (2.25)*, e procedere esattamente come nella dimostrazione del teorema centrale del limite.*

Dimostrazione della convergenza

Iniziamo con il calcolare la funzione caratteristica di $\widetilde{X}_j^{(n)}$, ossia

$$\widetilde{\varphi}^{(n)}(u) := \widetilde{\mathbb{E}}^{(n)}\big[\exp\{i\,u\,\widetilde{X}_j^{(n)}\}\big] = \widetilde{\mathbb{E}}^{(n)}\Big[\exp\big\{i\,\frac{u}{\sqrt{n}}\,[X_j - \widetilde{\mathbb{E}}^{(n)}[X_j]]\big\}\Big]$$

Tenendo presente che X_j assume solo i valori $+1$ e -1 con probabilità $\tilde{p}^{(n)} = \frac{1}{2} + \frac{\theta_n}{2\sqrt{n}}$ e $1 - \tilde{p}^{(n)} = \frac{1}{2} - \frac{\theta_n}{2\sqrt{n}}$, rispettivamente, e quindi $\widetilde{\mathbb{E}}^{(n)}[X_j] = \frac{\theta_n}{\sqrt{n}}$ si ottiene

$$\begin{aligned}
\widetilde{\varphi}^{(n)}(u) &= (\frac{1}{2} + \frac{\theta_n}{2\sqrt{n}})\exp\big\{i\,\frac{u}{\sqrt{n}}\,[1 - \frac{\theta_n}{\sqrt{n}}]\big\} + (\frac{1}{2} - \frac{\theta_n}{2\sqrt{n}})\exp\big\{i\,\frac{u}{\sqrt{n}}\,[-1 - \frac{\theta_n}{\sqrt{n}}]\big\} \\
&= \frac{\exp\{i\,\frac{u}{\sqrt{n}}\} + \exp\{-i\,\frac{u}{\sqrt{n}}\}}{2}\,\exp\{-i\,\tfrac{u\,\theta_n}{n}\} + \frac{\theta_n}{\sqrt{n}}\,\frac{\exp\{i\,\frac{u}{\sqrt{n}}\} - \exp\{-i\,\frac{u}{\sqrt{n}}\}}{2}\,\exp\{-i\,\tfrac{u\,\theta_n}{n}\} \\
&= \exp\{-i\,\tfrac{u\,\theta_n}{n}\}\,\Big[\frac{1}{2}(1 + \frac{iu}{\sqrt{n}}) - \frac{u^2}{2n} + o(\tfrac{u^2}{n}) + 1 - \frac{iu}{\sqrt{n}}) - \frac{u^2}{2n} + o(\tfrac{u^2}{n})) \\
&\qquad\qquad + \frac{\theta_n}{\sqrt{n}}\,\frac{1}{2}(1 + \frac{iu}{\sqrt{n}}) - \frac{u^2}{2n} + o(\tfrac{u^2}{n}) - 1 + \frac{iu}{\sqrt{n}}) + \frac{u^2}{2n} + o(\tfrac{u^2}{n}))\Big] \\
&= \exp\{-i\,\tfrac{u\,\theta_n}{n}\}\,\Big[1 - \frac{u^2}{2n} + o(\tfrac{u^2}{n}) + \frac{\theta_n}{\sqrt{n}}\,(\frac{iu}{\sqrt{n}} + o(\tfrac{u^2}{n}))\Big] \\
&= \exp\{-i\,\tfrac{u\,\theta_n}{n}\}\,\Big[1 - \frac{u^2}{2n} + \frac{iu\theta_n}{n} + o(\tfrac{u^2}{n})\Big].
\end{aligned}$$

Di conseguenza, da (2.25), la funzione caratteristica di $W_t^{(n)}$, vale

$$\widetilde{\mathbb{E}}^{(n)}\big[\exp\{i\,u\,W_t^{(n)}\}\big] = \Big(\exp\{-i\,\tfrac{u\,\theta_n}{n}\}\,\Big[1 - \frac{u^2}{2n} + \frac{iu\theta_n}{n} + o(\tfrac{u^2}{n})\Big]\Big)^{\lfloor nt\rfloor}$$

e converge, per n che tende ad infinito

$$\begin{aligned}
\lim_{n\to\infty}\widetilde{\mathbb{E}}^{(n)}\big[\exp\{i\,u\,W_t^{(n)}\}\big] &= \lim_{n\to\infty}\Big(\exp\{-i\,\tfrac{u\,\theta_n}{n}\}\,\Big[1 - \frac{u^2}{2n} + \frac{iu\theta_n}{n} + o(\tfrac{u^2}{n})\Big]\Big)^{\lfloor nt\rfloor} \\
&= \exp\{-i\,u\,\theta t\}\,\exp\{-\frac{u^2 t}{2} + iu\theta t\} \\
&= \exp\{-\frac{u^2 t}{2}\},
\end{aligned}$$

che è appunto la funzione caratteristica di una variabile gaussiana di media nulla e varianza t.

FINE DIMOSTRAZIONE

2.3.2 Indipendenza ed omogeneità degli incrementi

Ancora, se si considerano i tempi $0 \leq t_1 < t_2$, allora l'incremento $W^{(n)}_{t_2} - W^{(n)}_{t_1}$ ha la stessa legge di $W^{(n)}_{t_2-t_1}$, in quanto $W^{(n)}_{t_2} - W^{(n)}_{t_1}$ è funzione deterministica delle variabili aleatorie ξ_j, per j tale che $t^{(n)}_j = j/k \in (t_1, t_2]$, e la stessa funzione determina nello stesso modo $W^{(n)}_{t_2-t_1}$ a partire dalle variabili aleatorie ξ_i, per i tale che $t^{(n)}_i = i/n \in (0, t_2 - t_1]$, e di conseguenza la distribuzione dell'incremento $W^{(n)}_{t_2} - W^{(n)}_{t_1}$ converge ad una distribuzione gaussiana di valore atteso 0 e varianza $t_2 - t_1$, come si vede facilmente[12].

Se invece si considerano i tempi $0 \leq t_1 < t_2 < \ldots < t_{m-1} < t_m$ allora gli incrementi

$$W^{(n)}_{t_1} - W^{(n)}_0, \quad W^{(n)}_{t_2} - W^{(n)}_{t_1}, \quad \ldots \quad W^{(n)}_{t_m} - W^{(n)}_{t_{m-1}}$$

sono funzioni deterministiche delle variabili aleatorie

$$\{\xi_{j_1} : \tfrac{j_1}{n} \in (0, t_1]\}, \quad \{\xi_{j_2} : \tfrac{j_2}{n} \in (t_1, t_2]\} \quad \ldots \quad \{\xi_{j_m} : \tfrac{j_m}{n} \in (t_{m-1}, t_m]\}$$

che sono indipendenti, e di conseguenza anche gli incrementi di $W^{(n)}$ sono indipendenti. Come ulteriore conseguenza questa proprietà si mantiene al tendere di n all'infinito.

2.3.3 Definizione del moto browniano e del modello di Black e Scholes

Le osservazioni precedenti portano naturalmente alla seguente definizione del processo di Wiener standard o moto browniano.

Definizione 2.1 (moto browniano)**.** *Si chiama moto browniano un processo W_t per $t \in \mathbb{R}^+$ un processo tale che*

1 $W_0 = 0$,

2 se $0 \leq t_1 < t_2 < \ldots < t_{m-1} < t_m$ *allora gli incrementi*

$$W_{t_1} - W_0, \quad W_{t_2} - W_{t_1}, \quad \ldots \quad W_{t_m} - W_{t_{m-1}}$$

sono indipendenti,

3 se $0 \leq t_1 < t_2$ *ed* $s > 0$ *allora gli incrementi*

$$W_{t_2} - W_{t_1} \quad e \quad W_{t_2+s} - W_{t_1+s} \sim N(0, t_2 - t_1),$$

ovvero hanno la stessa distribuzione gaussiana di valore atteso 0 *e varianza* $t_2 - t_1$. *In altre parole si dice anche che gli incrementi sono omogenei.*

Di solito oltre alle tre precedenti proprietà si aggiunge anche la proprietà che le traiettorie sono continue, ossia che per ogni ω *la funzione* $t \mapsto W_t(\omega)$ *è una funzione continua.*

---

[12]Infatti, utilizzando le stesse notazioni dell'Osservazione 2.3, si ha

$$W^{(n)}_{t_2} - W^{(n)}_{t_1} = \sum_{j=\lfloor nt_1 \rfloor+1}^{\lfloor nt_2 \rfloor} \widetilde{X}^{(n)}_j,$$

come si vede subito osservando

$$W^{(n)}_{t_2} = \sum_{j=1}^{\lfloor nt_2 \rfloor} \widetilde{X}^{(n)}_j = \sum_{j=1}^{\lfloor nt_1 \rfloor} \widetilde{X}^{(n)}_j + \sum_{j=\lfloor nt_1 \rfloor+1}^{\lfloor nt_2 \rfloor} \widetilde{X}^{(n)}_j = W^{(n)}_{t_1} + (W^{(n)}_{t_2} - W^{(n)}_{t_1}).$$

Questa scomposizione, insieme al fatto che $W^{(n)}_0 = 0$ permette anche di vedere che $W^{(n)}_{t_1} (= W^{(n)}_{t_1} - W^{(n)}_0 = W^{(n)}_{t_1} - 0)$ e $W^{(n)}_{t_2} - W^{(n)}_{t_1}$ sono indipendenti. L'indipendenza di questi incrementi permette di affermare che, oltre alla convergenza in distribuzione di $W^{(n)}_{t_1}$ ad una variabile aleatoria gaussiana di valore medio nullo e varianza t_1 e la convergenza di $W^{(n)}_{t_2} - W^{(n)}_{t_1}$ ad una variabile aleatoria gaussiana di valore medio nullo e varianza $t_2 - t_1$, si ha anche la convergenza congiunta di $\big(W^{(n)}_{t_1}, W^{(n)}_{t_2} - W^{(n)}_{t_1}\big)$ ad una variabile aleatoria bidimensionale gaussiana di valori medi nulli, indipendenti e di varianza rispettivamente t_1 e $t_2 - t_1$. Questo fatto implica anche la convergenza della distribuzione congiunta di $\big(W^{(n)}_{t_1}, W^{(n)}_{t_2}\big)$. Questa osservazione si estende facilmente al caso in cui si considerino le variabili aleatorie $\big(W^{(n)}_{t_1}, W^{(n)}_{t_2}, \cdots W^{(n)}_{t_m}\big)$, con $0 \leq t_1 < t_2 < \ldots < t_{m-1} < t_m$.

Inoltre quanto abbiamo visto nella sezione precedente permette di dare una definizione di un processo S_t, che, alla luce di (2.24), si può interpretare come il limite, rispetto alle misure martingala equivalenti, dei processi

$$S_t^{(n)} = \exp\left\{ \log\left(S_t^{(n)}\right)\right\} = S_0 \, \exp\left\{\sigma W_t^{(n)} + \sigma \, \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} \widetilde{\mathbb{E}}^{(n)}[2\xi_i - 1]\right\},$$

da cui si definisce

$$S_t = S_0 \, \exp\left\{\sigma W_t + \sigma \, \frac{r - \frac{1}{2}\sigma^2}{\sigma} \, t\right\} = S_0 \, \exp\left\{\sigma W_t + (r - \tfrac{1}{2}\sigma^2)\, t\right\},$$

in un opportuno spazio di probabilità $(\Omega, \mathcal{F}, \widetilde{\mathbb{P}})$, rispetto al quale il processo attualizzato dei prezzi, ossia

$$\widetilde{S}_t = \frac{S_t}{B_t} = \frac{S_t}{B_0 \, e^{rt}} = \frac{S_0}{B_0} \, \exp\left\{\sigma W_t + - \tfrac{1}{2}\sigma^2 \, t\right\},$$

risulta una martingala.

Osservazione 2.4 (Il modello di Black e Scholes). *In realtà il modello di Black e Scholes viene di solito presentato nel seguente modo. Si assume che nel mercato ci siano due titoli, uno non rischioso*

$$B_t := B_0 \, e^{rt},$$

ed uno rischioso

$$S_t = S_0 \, \exp\left\{\mu \, t + \sigma \, \mathbb{W}_t - \tfrac{1}{2}\sigma^2 \, t\right\},$$

dove il processo stocastico $\mathbb{W}_t$ è un moto browniano standard in uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$. Si dimostra che esiste una misura $\widetilde{\mathbb{P}}$, equivalente a $\mathbb{P}$, e sotto la quale il processo attualizzato si scrive come

$$\widetilde{S}_t = \frac{S_t}{B_t} = \frac{S_t}{B_0 \, e^{rt}} = \frac{S_0}{B_0} \, \exp\left\{\sigma W_t - \tfrac{1}{2}\sigma^2 \, t\right\}, \tag{2.26}$$

e dove il processo W_t è un moto browniano nello spazio di probabilità $(\Omega, \mathcal{F}, \widetilde{\mathbb{P}})$. Non dimostriamo questo fatto, ma osserviamo solo che, essendo

$$\widetilde{S}_t = \frac{S_t}{B_t} = \frac{S_t}{B_0 \, e^{rt}} = \frac{S_0}{B_0} \, \exp\left\{(\mu - r)\, t + \sigma \, \mathbb{W}_t - \tfrac{1}{2}\sigma^2 \, t\right\}, \tag{2.27}$$

il processo W_t deve necessariamente soddisfare

$$\sigma W_t = (\mu - r)\, t + \sigma \, \mathbb{W}_t,$$

ovvero il processo W_t deve essere necessariamente definito come

$$W_t := \mathbb{W}_t - \frac{\mu - r}{\sigma} \, t,$$

e quindi la misura di probabilità $\widetilde{\mathbb{P}}$ deve essere necessariamente definita come la misura che rende il processo $\mathbb{W}_t - \frac{\mu - r}{\sigma} \, t$ un moto browniano.

Capitolo 3

Processi aleatori a tempo continuo

3.1 Processi aleatori, definizioni ed esempi

Esistono diversi modi di definire un processo stocastico. Il primo consiste nel considerare semplicemente una famiglia di variabili aleatorie.

Definizione 3.1 (Processo stocastico come famiglia di v.a.). *Dato uno spazio* $(\Omega, \mathcal{F}, \mathbb{P})$ *ed un insieme* I *di indici*[1], *allora una famiglia di variabili aleatorie* $\{X_t : \Omega \mapsto \mathbb{R}(o\,\mathbb{R}^d),\, per\, t \in I\}$ *è detta* **processo stocastico**, *e* $\mathbb{R}$ *(o* $\mathbb{R}^d$*) è detto lo* **spazio degli stati** *del processo.*

In alcuni casi si vuole mettere in evidenza l'evoluzione rispetto al tempo e si preferisce dare una definizione di processo stocastico come funzione aleatoria.

Definizione 3.2 (Processo stocastico come funzione aleatoria). *Dato uno spazio* $(\Omega, \mathcal{F}, \mathbb{P})$ *ed un insieme* I *di indici, si definisce* **processo stocastico (come funzione aleatoria)** *la funzione (misurabile)*[2]

$$X : \Omega \mapsto \mathbb{R}^I; \omega \mapsto \big(t \mapsto X(t, \omega)\big).$$

Una definizione analoga vale nel caso di processi con spazio degli stati $\mathbb{R}^d$*.*

Altre volte si è interessati anche ad alcune proprietà delle traiettorie $t \mapsto X(t, \omega)$, quali ad esempio la continuità. Viene allora naturale dare la definizione di processo a traiettorie continue.

Definizione 3.3 (Processo stocastico come funzione aleatoria continua). *Dato uno spazio* $(\Omega, \mathcal{F}, \mathbb{P})$ *ed un insieme* I *di indici, si definisce* **processo stocastico (come funzione aleatoria continua)** *la funzione (misurabile)*[3]

$$X : \Omega \mapsto C(I; \mathbb{R}); \omega \mapsto \big(t \mapsto X(t, \omega)\big),$$

[1] Tipicamente l'insieme degli indici è $I = [0, \infty)$, o $I = [0, T]$, o ancora $I = \mathbb{N}$ (e in questo caso si parla più propriamente di successioni aleatorie), ma è possibile anche che l'insieme degli indici sia multidimensionale, ad esempio $I = \mathbb{R}^2$ (e in questo caso si parla più propriamente di campi aleatori).

[2] Per rendere la definizione completa andrebbe precisata la σ-algebra che si mette sullo spazio di tutte le funzioni $\mathbb{R}^I$. Di solito si tratta in realtà di richiedere almeno che tutte le funzioni

$$\Omega \mapsto \mathbb{R}; \omega \mapsto X(t, \omega)$$

siano variabili aleatorie $\mathcal{F}$-misurabili. La σ-algebra considerata su $\mathbb{R}^I$ è di solito $\mathcal{R}^I$, la σ-algebra del Teorema di Kolmogorov 3.1 e la nota 3.1 corrispondente (per maggiori dettagli si veda, ad esempio, Billingsley [4]) .

[3] Nel caso $I = [0, T]$ lo spazio $C(I; \mathbb{R})$ è uno spazio metrico con la metrica uniforme:

$$d(x(\cdot), y(\cdot)) := \sup_{t \in [0,T]} |x(t) - y(t)|,$$

e la σ-algebra su $C(I; \mathbb{R})$ è quella generata dagli aperti. Se invece $I = [0, \infty)$ si può, ad esempio, usare la metrica

$$d(x(\cdot), y(\cdot)) = \sum_{N=1}^{\infty} \frac{1}{2^N} \sup_{t \in [0,N]} |x(t) - y(t)| \wedge 1,$$

e di nuovo la σ-algebra su $C([0, \infty); \mathbb{R})$ è quella generata dagli aperti.

dove $C(I;\mathbb{R})$ è lo spazio metrico delle traiettorie continue. Una definizione analoga vale nel caso di processi con spazio degli stati $\mathbb{R}^d$.

Dato un processo si definiscono le funzioni di distribuzione finito-dimensionali.

Definizione 3.4 (Famiglia delle funzioni di distribuzione finito-dimensionali)**.** *Dato un processo stocastico (secondo una delle precedenti definizioni) la famiglia delle funzioni di distribuzione $F_{t_1,\cdots,t_n}$, definite per $n \geq 1$, e $t_1, \cdots t_n$, con $t_i \in I$, come le funzioni di distribuzione congiunte delle variabili $(X_{t_1}, \cdots, X_{t_n})$, ovvero*

$$F_{t_1,\cdots,t_n}(x_1, \cdots, x_n) = \mathbb{P}(X_{t_1} \leq x_1, \cdots, X_{t_n} \leq x_n),$$

è detta famiglia delle funzioni di distribuzione finito-dimensionali del processo $X = (X_t)_{t\in I}$.

Va detto che, così come una variabile aleatoria viene spesso individuata attraverso la sua distribuzione, senza specificare quale sia lo spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ sul quale è definita, spesso anche un processo viene individuato attraverso le sue distribuzioni finito-dimensionali. Si pone quindi il problema di individuare quali sono le famiglie di distribuzioni che sono effettivamente famiglie di distribuzioni finito-dimensionali di un processo.

Si individuano facilmente due condizioni necessarie (**(C1)** e **(C2)**), dette ***condizioni di consistenza***:

(C1) Sia $k > 1$, sia π una permutazione di $\{1, \cdots, k\}$, si ponga

$$\Phi_\pi : \mathbb{R}^k \to \mathbb{R}^k, (x_1, \cdots, x_k) \to \Phi_\pi(x_1, \cdots, x_k) := (x_{\pi_1}, \cdots, x_{\pi_k}).$$

Si richiede che per ogni $k > 1$, π, $t_1, \cdots, t_k$ e $(x_1, \cdots, x_k)$ valga

$$F_{t_1,\cdots,t_k}(x_1, \cdots, x_k) = F_{t_{\pi_1},\cdots,t_{\pi_k}}(\Phi_\pi(x_1, \cdots, x_k)) = F_{t_{\pi_1},\cdots,t_{\pi_k}}(x_{\pi_1}, \cdots, x_{\pi_k}). \tag{3.1}$$

La precedente condizione **(C1)** è chiaramente necessaria, infatti

$$\begin{aligned}&F_{t_1,\cdots,t_k}(x_1, \cdots, x_k) = \mathbb{P}\{(X_{t_1} \leq x_1, \cdots, X_{t_k} \leq x_k)\}\\ &= \mathbb{P}\{X_{t_{\pi_1}} \leq x_{\pi_1}, \cdots, X_{t_{\pi_k}} \leq x_{\pi_k}\} = F_{t_{\pi_1},\cdots,t_{\pi_k}}(x_{\pi_1}, \cdots, x_{\pi_k}).\end{aligned}$$

(C2) Si richiede che per ogni $k \geq 1$, $t_1, \cdots, t_k, t_{k+1}$ e $(x_1, \cdots, x_k)$ valga

$$\lim_{x_{k+1}\to\infty} F_{t_1,\cdots,t_k,t_{k+1}}(x_1, \cdots, x_k, x_{k+1}) = F_{t_1,\cdots,t_k}(x_1, \cdots, x_k). \tag{3.2}$$

La precedente condizione **(C1)** è chiaramente necessaria, infatti

$$\begin{aligned}&\lim_{x_{k+1}\to\infty} F_{t_1,\cdots,t_k,t_{k+1}}(x_1, \cdots, x_k, x_{k+1})\\ &= \lim_{x_{k+1}\to\infty} \mathbb{P}(X_{t_1} \leq x_1, \cdots, X_{t_k} \leq x_k, X_{t_{k+1}} \leq x_{k+1})\\ &= \mathbb{P}(X_{t_1} \leq x_1, \cdots, X_{t_k} \leq x_k) = F_{t_1,\cdots,t_k}(x_1, \cdots, x_k).\end{aligned}$$

È interessante notare che la prima condizione potrebbe essere automaticamente soddisfatta dando le funzioni di distribuzione solo per $t_1 < t_2 < \cdots, < t_k$ (e definendole negli altri casi in modo che la condizione **(C1)** sia soddisfatta). Allora, però, la condizione **(C2)**, va modificata:

(C2′) Sia $k \geq 1$, siano $t_1 < t_2 < \cdots < t_k < t_{k+1}$, allora

$$\lim_{x\to\infty} F_{t_1,\cdots,t_k,t_{k+1}}(x_1, \cdots, x_k, x) = F_{t_1,\cdots,t_k}(x_1, \cdots, x_k), \tag{3.3}$$

$$\begin{aligned}&\lim_{x\to\infty} F_{t_1,\cdots,t_{i-1},t_i,t_{i+1},\cdots,t_{k+1}}(x_1, \cdots, x_{i-1}, x, x_{i+1}, \cdots, x_{k+1})\\ &= F_{t_1,\cdots,t_{i-1},t_{i+1},\cdots,t_{k+1}}(x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_{k+1}), \quad \text{per } i = 1, \cdots, k.\end{aligned} \tag{3.4}$$

Il seguente teorema garantisce che le precedenti condizioni necessarie, sono anche sufficienti.

Teorema 3.1 (di Kolmogorov). *Sia data una famiglia $F_{t_1,\cdots,t_k}$ di funzioni di distribuzione finito-dimensionali consistente, cioè che verifica le condizioni di consistenza* **(C1)** *e* **(C2)** *(ovvero (3.1) e (3.2)), allora esiste uno spazio di probabilità ed un processo aleatorio che ammette $F_{t_1,\cdots,t_k}$ come funzioni di distribuzione finito-dimensionali.*[4]

La tesi rimane valida se valgono le condizioni equivalenti **(C1)** *e* **(C2′)***(ovvero (3.1), (3.3) e (3.3).*

Vediamo subito delle applicazioni del precedente teorema, mentre per la dimostrazione rimandiamo al Billingsley [4].

Esempio 3.1 (Processi a coordinate indipendenti). *Data una famiglia di funzioni di distribuzione $\{F_t, t \in I\}$ ad un tempo, si consideri la famiglia delle funzioni di distribuzione finito-dimensionali $F_{t_1,\cdots,t_k}(x_1,\cdots,x_k) := F_{t_1}(x_1)\times\cdots\times F_{t_k}(x_k)$. Tale famiglia è chiaramente una famiglia consistente e quindi esiste un processo con tali funzioni di distribuzione finito-dimensionali.*
Si noti che non si richiede che I sia numerabile, e che questo esempio, nel caso numerabile, garantisce l'esistenza di successioni di v.a. indipendenti[5] *.*

Esempio 3.2 (Processo di Poisson). *Un processo di Poisson di parametro $\lambda(>0)$ è un processo $(N_t)_{t\geq 0}$, a valori in $0,1,2,\ldots$, con $N_0 = 0$, a incrementi indipendenti, e con distribuzione di Poisson. Il parametro dell'incremento sull'intervallo $(s,t]$ è dato dal prodotto $\lambda(t-s)$. Questo significa che per ogni $k \geq 1$, e $t_0 = 0 \leq t_1 \leq t_2 \cdots \leq t_k$, e per ogni $n_1,\cdots,n_k \in \{0,1,2, cdots\}$:*
INDIPENDENZA DEGLI INCREMENTI:

$$\mathbb{P}(N_{t_1}-N_{t_0}=n_1,\,N_{t_2}-N_{t_1}=n_2,\cdots,N_{t_k}-N_{t_{k-1}}=n_k)=\mathbb{P}(N_{t_1}-N_{t_0}=n_1)\,\mathbb{P}(N_{t_2}-N_{t_1}=n_2)\cdots\mathbb{P}(N_{t_k}-N_{t_{k-1}}=n_k)$$

DISTRIBUZIONE DI POISSON:

$$\mathbb{P}(N_t - N_s = n) = \frac{\big(\lambda(t-s)\big)^n}{n!}\,e^{-\lambda(t-s)}.$$

Dalle precedenti relazioni si ottiene che, se $0 \leq \ell_1 \leq \ell_2 \leq \cdots \leq \ell_k$

$$\begin{aligned}&\mathbb{P}(N_{t_1}=\ell_1,\,N_{t_2}=\ell_2,\cdots N_{t_k}=\ell_k)\\&=\frac{\big(\lambda(t_1-t_0)\big)^{\ell_1}}{\ell_1!}\,e^{-\lambda(t_1-t_0)}\,\frac{\big(\lambda(t_2-t_1)\big)^{\ell_2-\ell_1}}{(\ell_2-\ell_1)!}\,e^{-\lambda(t_k-t_{k-1})}\cdots\frac{\big(\lambda(t_k-t_{k-1})\big)^{\ell_k-\ell_{k-1}}}{(\ell_k-\ell_{k-1})!}\,e^{-\lambda(t_k-t_{k-1})}.\end{aligned}$$

In questo modo abbiamo individuato univocamente le distribuzioni finito dimensionali. È facile vedere che quest'ultime formano una famiglia consistente di distribuzioni finito dimensionali.

Tuttavia è interessante osservare che è possibile costruire un processo di Poisson anche in un altro modo: data una successione di variabili aletorie $\{U_n\}_{n\geq 0}$ indipendenti, identicamente distribuite, tutte con legge esponenziale di parametro $\lambda(>0)$ (ossia U_n sono variabili aleatorie a valori positivie con $\mathbb{P}(U_n > u) = e^{-\lambda u}$ per $u > 0$), si definiscono

$$T_0 = 0, \quad T_1 = U_1, \quad ,T_n = T_{n-1} + U_n (= U_1 + U_2 + \cdots + U_n),$$

e si pone

$$N_t = n \quad \Leftrightarrow T_n \leq t < T_{n+1}.$$

Si può dimostrare che questo processo ha proprio le distribuzioni finito dimensionali precedenti, ossia definisce un Processo di Poisson di parametro λ.

Esempio 3.3 (Processi gaussiani). *Siano date una funzione $m : I \to \mathbb{R}$, $t \to m(t)$ e una funzione, detta* **funzione di correlazione***, $K : I \times I \to \mathbb{R}$, $(t,s) \to K(t,s)$* **definita non negativa***, cioè tale che comunque scelti $n \geq 1$, $t_1,\cdots,t_n \in I$, $\eta_1,\cdots,\eta_n \in \mathbb{C}$ valga*

$$\sum_{h,k}^{1,n} \eta_h \overline{\eta}_k K(t_h,t_k) \geq 0.$$

[4]Inoltre è sempre possibile prendere come spazio di probabilità lo spazio canonico $\mathbb{R}^I$, come σ-algebra la σ-algebra generata dai cilindri, ovvero dagli insiemi del tipo

$$\mathcal{C} = \{x(\cdot) \in \mathbb{R}^I \mid (x(t_1), x(t_2),\cdots,x(t_k)) \in H\}, \text{ dove } H \in \mathcal{B}(\mathbb{R}^k)$$

ed infine come processo X_t il processo canonico $X_t(x(\cdot)) = x(t)$.

[5]L'esistenza di successioni indipendenti è di solito sottintesa nei teoremi fondamentali del *Calcolo delle Probabilità*, come ad esempio la Legge dei grandi numeri, o il Teorema centrale del limite.

Si noti che quindi necessariamente $K(t,s) = K(s,t)$ e $K(t,t) \geq 0$ e che la condizione che la funzione di correlazione sia definita non negativa corrisponde alla richiesta che la matrice $\Gamma = (\Gamma_{i,j} := K(t_i,t_j))_{i,j=1,\cdots,n}$ sia definita non negativa (ovvero positiva in senso lato) qualunque siano t_j, per $j = 1,\cdots,n$.

Sia $F_{t_1,\cdots,t_k}$ la funzione di distribuzione congiunta gaussiana di media $(m(t_1),\cdots,m(t_k))$ e matrice di covarianza Γ definita da $\Gamma_{i,j} = K(t_i,t_j)$ per $i,j = 1,\cdots,k$, e sia $f_{t_1,\cdots,t_k}$ la sua densità.

Per convincersi dell'esistenza di un processo con tali distribuzioni finito-dimensionali si consideri il caso in cui $K(t,s)$ sia strettamente definita positiva e si noti che se $(Y_1,\cdots,Y_k)$ è un vettore aleatorio con componenti indipendenti e ciascuna con distribuzione normale $N(0,1)$, allora il vettore definito da $Z = (Z_1,\cdots,Z_k) = A(Y_1,\cdots,Y_k) + (m(t_1),\cdots,m(t_k)) = AY + m$, dove $A = \Gamma^{1/2}$ (cioè $\Gamma = A^tA = AA^t$), è un vettore con la distribuzione cercata[6].

Per la consistenza della famiglia $F_{t_1,\cdots,t_k}$ così definita, si noti che l'operatore Φ_π di (3.1) è una trasformazione lineare e che trasformazioni lineari di vettori gaussiani sono ancora gaussiani. Inoltre, indicando ancora con Φ_π la matrice associata all'operatore di (3.1), allora $\Phi_\pi Z = \Phi_\pi AY + \Phi_\pi m$, segue una legge gaussiana di media $\Phi_\pi m = (m(t_{\pi_1}),\cdots,m(t_{\pi_k}))$ e con matrice di covarianza $(\Phi_\pi A)(\Phi_\pi A)^t = \Phi_\pi AA^t\Phi_\pi^t = \Phi_\pi \Gamma \Phi_\pi^t = (\Gamma_{\pi_i,\pi_j})_{i,j=1,\cdots,k}$. Da queste osservazioni si deduce immediatamente che per la densità vale

$$f_{t_1,\cdots,t_k}(x_1,\cdots,x_k) = f_{t_{\pi_1},\cdots,t_{\pi_k}}(\Phi_\pi(x_1,\cdots,x_k)) = f_{t_{\pi_1},\cdots,t_{\pi_k}}(x_{\pi_1},\cdots,x_{\pi_k}),$$

ovvero la (3.1). Per quanto riguarda la (3.2), basta ricordare che ogni sottovettore di un vettore gaussiano è ancora un vettore gaussiano.

Esempio 3.4 (Processo di Wiener standard). *Come caso particolare dell'Esempio precedente possiamo stabilire l'esistenza del* **processo di Wiener standard**, *detto anche* **moto browniano**, *cioè del processo gaussiano con $m(t) = 0$ e $K(t,s) = t \wedge s$.*

Bisogna ovviamente controllare che la funzione $K(\cdot,\cdot)$ sia definita non negativa. Ciò può essere fatto direttamente, ma con una certa fatica[7].

[6] Si veda l'Appendice 3.3

[7] Dati $t_1 < t_2 < \cdots < t_k$ la matrice $(t_i \wedge t_j)_{i,j}$ si può scrivere come

$$\begin{pmatrix} t_1 & t_1 & \cdots & t_1 & t_1 \\ t_1 & t_2 & \cdots & t_2 & t_2 \\ \vdots & & \ddots & & \vdots \\ t_1 & t_2 & \cdot & t_{k-1} & t_{k-1} \\ t_1 & t_2 & \cdot & t_{k-1} & t_k \end{pmatrix}$$

Si può dimostrare che il determinante di questa matrice è $t_1(t_2 - t_1)(t_3 - t_2)\cdots(t_k - t_{k-1}) > 0$, e ciò dimostra subito il fatto che $K(t,s) = t \wedge s$ è definita positiva. Il primo passo per dimostrare tale identità è osservare che

$$det \begin{pmatrix} t_1 & t_1 & \cdots & t_1 & t_1 \\ t_1 & t_2 & \cdots & t_2 & t_2 \\ \vdots & & \ddots & & \vdots \\ t_1 & t_2 & \cdot & t_{k-1} & t_{k-1} \\ t_1 & t_2 & \cdot & t_{k-1} & t_k \end{pmatrix} = det \begin{pmatrix} t_1 & 0 & \cdots & 0 & 0 \\ t_1 & t_2 - t_1 & \cdots & t_2 - t_1 & t_2 - t_1 \\ \vdots & & \ddots & & \vdots \\ t_1 & t_2 - t_1 & \cdot & t_{k-1} - t_1 & t_{k-1} - t_1 \\ t_1 & t_2 - t_1 & \cdot & t_{k-1} - t_1 & t_k - t_1 \end{pmatrix}$$

da cui

$$det \begin{pmatrix} t_1 & t_1 & \cdots & t_1 & t_1 \\ t_1 & t_2 & \cdots & t_2 & t_2 \\ \vdots & & \ddots & & \vdots \\ t_1 & t_2 & \cdot & t_{k-1} & t_{k-1} \\ t_1 & t_2 & \cdot & t_{k-1} & t_k \end{pmatrix} = t_1\, det \begin{pmatrix} t_2 - t_1 & \cdots & t_2 - t_1 & t_2 - t_1 \\ \vdots & \ddots & & \vdots \\ t_2 - t_1 & \cdot & t_{k-1} - t_1 & t_{k-1} - t_1 \\ t_2 - t_1 & \cdot & t_{k-1} - t_1 & t_k - t_1 \end{pmatrix},$$

poi basta procedere per induzione, notando che $t_i - t_1 - (t_2 - t_1) = t_i - t_2$.

Un metodo decisamente più probabilistico[8] *è il seguente: consiste nell'osservare che la funzione di correlazione*

$$K(t,s) := Cov(N_t, N_s)$$

di un processo N_t di Poisson standard (cioè con $\lambda = 1$) è proprio $t \wedge s$, ciò è sufficiente[9] *a garantire la sua non negatività. Infatti*

$$\begin{aligned}
Cov(N_t, N_s) &= Cov(N_{t\wedge s}, N_{t\vee s}) \\
&= \mathbb{E}[N_{t\wedge s}(N_{t\wedge s} + N_{t\vee s} - N_{t\wedge s})] - \mathbb{E}[N_{t\wedge s}]\mathbb{E}[N_{t\wedge s} + N_{t\vee s} - N_{t\wedge s}] \\
&= \mathbb{E}[N_{t\wedge s}N_{t\wedge s}] + \mathbb{E}[N_{t\wedge s}(N_{t\vee s} - N_{t\wedge s})] - \mathbb{E}[N_{t\wedge s}](\mathbb{E}[N_{t\wedge s}] + \mathbb{E}[N_{t\vee s} - N_{t\wedge s}]) \\
&= Var(N_{t\wedge s}) + \mathbb{E}[N_{t\wedge s}(N_{t\vee s} - N_{t\wedge s})] - \mathbb{E}[N_{t\wedge s}]\mathbb{E}[N_{t\vee s} - N_{t\wedge s}] = Var(N_{t\wedge s}) \\
&= t \wedge s
\end{aligned}$$

Si noti che la proprietà del processo di Wiener di avere la stessa funzione di correlazione del processo di Poisson standard, implica che gli incrementi sono non correlati. Poiché gli incrementi del processo di Wiener $(W_t)_{t\geq 0}$ hanno distribuzione gaussiana, sono quindi anche indipendenti. Inoltre per ogni $s < t$, l'incremento $W_t - W_s$ deve essere una variabile aleatoria gaussiana (in quanto differenza di due v.a. congiuntamente gaussiane), deve avere media nulla e varianza[10] *uguale alla varianza di $N_t - N_s$, ovvero uguale a $t - s$. Infine deve essere $\mathbb{P}(W_0 = 0) = 1$, in quanto la media deve essere nulla e la varianza deve essere uguale a $K(0,0) = 0$.*

Questa proprietà di indipendenza degli incrementi è fondamentale per ottenere la densità congiunta di $(W_{t_1}, \cdots, W_{t_n})$. Infatti si può pensare che il vettore $(W_{t_1}, \cdots, W_{t_n})$ si ottiene dal vettore degli incrementi

$$(\Delta W_{t_1}, \Delta W_{t_2}, \cdots, \Delta W_{t_n}) = (W_{t_1} - W_0, W_{t_2} - W_{t_1} \cdots, W_{t_n} - W_{t_{n-1}})$$

con la seguente trasformazione lineare

$$W_{t_h} = W_{t_h} - W_0 = \sum_{i=1}^{h}(W_{t_i} - W_{t_{i-1}}) = \sum_{i=1}^{h} \Delta W_{t_i},$$

(dove si è posto $t_0 = 0$) che corrisponde ad una matrice B la cui inversa B^{-1} è la matrice la cui diagonale ha tutti gli elementi uguali ad 1, la cui sottodiagonale ha gli elementi tutti uguali a -1 e tutti i rimanenti elementi uguali a 0, in quanto banalmente

$$\Delta W_{t_h} = W_{t_h} - W_{t_{h-1}}.$$

[8] Tuttavia questo metodo presuppone la conoscenza del processo di Poisson, che in queste note viene definito come processo ad incrementi indipendenti in Sezione 4.1.

[9] La matrice di covarianza di un vettore aleatorio è sempre definita non negativa, e di conseguenza la funzione di correlazione $K(s,t) = Cov(X_s, X_t)$ di un processo X_t qualsiasi è sempre definita non negativa:

$$\begin{aligned}
\sum_{h,k}^{1,n} \eta_h \overline{\eta}_k K(t_h, t_k) &= \sum_{h,k}^{1,n} \eta_h \overline{\eta}_k \mathbb{E}\left[(X_{t_h} - \mathbb{E}[X_{t_h}])(X_{t_k} - \mathbb{E}[X_{t_k}])\right] \\
&= \mathbb{E}\left[\sum_{h,k}^{1,n} \eta_h \overline{\eta}_k (X_{t_h} - \mathbb{E}[X_{t_h}])(X_{t_k} - \mathbb{E}[X_{t_k}])\right] \\
&= \mathbb{E}\left[\sum_{h=1}^{n} \eta_h (X_{t_h} - \mathbb{E}[X_{t_h}]) \sum_{k=1}^{n} \overline{\eta}_k (X_{t_k} - \mathbb{E}[X_{t_k}])\right] \\
&= \mathbb{E}\left[\left|\sum_{h=1}^{n} \eta_h (X_{t_h} - \mathbb{E}[X_{t_h}])\right|^2\right] \geq 0.
\end{aligned}$$

[10] Il lettore che non conosce il processo di Poisson, può procedere anche nel seguente semplicissimo modo

$$\begin{aligned}
Var(W_t - W_s) &= Cov(W_t - W_s, W_t - W_s) = Cov(W_t, W_t) - Cov(W_t, W_s) - Cov(W_s, W_t) + Cov(W_s, W_s) \\
&= K(t,t) - 2K(s,t) + K(s,s) = t - 2s \wedge t + s = t - s.
\end{aligned}$$

Si noti che il procedimento permette di calcolare la varianza degli incrementi per un qualunque processo, purché sia nota la funzione di covarianza $K(s,t)$.

Da questa osservazione, dalla formula di trasformazione della densità e dall'osservazione che gli incrementi sono indipendenti, ovvero che

$$f_{\Delta W_{t_1},\cdots,\Delta W_{t_n}}(y_1,\cdots,y_n) = f_{\Delta W_{t_1}}(y_1)\cdots f_{\Delta W_{t_n}}(y_n) = \prod_{h=1}^{n} g_{t_h-t_{h-1}}(y_h),$$

dove

$$g_s(y) = \frac{1}{\sqrt{2\pi s}}\exp\{-\frac{y^2}{2s}\}$$

si ottiene la densità congiunta di $(W_{t_1},\cdots,W_{t_n})$

$$f_{W_{t_1},\cdots,W_{t_n}}(x_1,x_2,\cdots,x_n) = f_{\Delta W_{t_1},\cdots,\Delta W_{t_n}}(x_1,x_2-x_1\cdots,x_n-x_{n-1}) \tag{3.5}$$

$$= \prod_{h=1}^{n} g_{t_h-t_{h-1}}(x_h-x_{h-1}) \tag{3.6}$$

$$= \frac{1}{\sqrt{2\pi t_1}}e^{-\frac{x_1^2}{2t_1}}\frac{1}{\sqrt{2\pi(t_2-t_1)}}e^{-\frac{(x_2-x_1)^2}{2(t_2-t_1)}}\cdots\frac{1}{\sqrt{2\pi(t_n-t_{n-1})}}e^{-\frac{(x_n-x_{n-1})^2}{2(t_n-t_{n-1})}} \tag{3.7}$$

La precedente formula 3.5 si ottiene come caso particolare del seguente problema generale: calcolare la distribuzione di una trasformazione di una variabile aleatoria, ossia, ad esempio, se Z è una variabile aleatoria con distribuzione $\mathbf{P}_Y$ ed $Z=\varphi(Y)$, allora la distribuzione $\mathbf{P}_Z$ di Z è data da

$$\mathbf{P}_Z(B) = \mathbf{P}_Y\big(\varphi^{-1}(B)\big),$$

come è immediato verificare. Infatti

$$\mathbf{P}_Z(B) = \mathbb{P}(Z\in B) = \mathbb{P}(\varphi(Y)\in B) = \mathbb{P}\big(Y\in\varphi^{-1}(B)\big) = \mathbf{P}_Y\big(\varphi^{-1}(B)\big),$$

Nel caso di variabili aleatorie multivariate, e per funzioni φ sufficientemente regolari, si possono ottenere formule esplicite, utilizzando noti risultati di analisi: ad esempio, se Y ammette densità f_Y e φ è invertibile[11] e con derivate continue, allora anche Z ammette densità e si ha

$$f_Z(z) = f_Y(\varphi^{-1}(z))\left|det\left(\frac{\partial\varphi^{-1}(z)}{\partial z}\right)\right| = f_Y(\varphi^{-1}(z))\frac{1}{\left|det\left(\frac{\partial\varphi(y)}{\partial y}\right)\right|_{y=\varphi^{-1}(z)}}.$$

Particolarmente semplice è il caso di trasformazioni lineari (o affini) in cui lo Jacobiano è il determinante della matrice. Ad esempio se $Z=AY$, con A invertibile, allora $\varphi^{-1}(z)=A^{-1}z$ e la formula precedente diviene

$$f_Z(z) = f_Y(A^{-1}(z))\frac{1}{|det(A)|}.$$

Esempio 3.5. *Un esempio di trasformazione che incontreremo spesso nel seguito è il caso in cui* $Y=(Y_1,Y_2,\cdots,Y_m)$ *e*

$$\begin{aligned} &Z_1 = Y_1, && && z_1 = y_1,\\ &Z_2 = Y_1+Y_2, && \textit{ossia } z=\varphi(y)=A\,y, \textit{ con} && z_2 = y_1+y_2,\\ &\cdots && && \cdots\\ &Z_m = Y_1+Y_2+\cdots+Y_m, && && z_m = y_1+y_2+\cdots+y_m. \end{aligned}$$

[11] In realtà basta che esista un aperto $\mathcal{O}$, tale che la densità $f_Y(y)=0$ per $y\notin A$ e tale che φ sia invertibile da $\mathcal{O}$ a $\varphi(\mathcal{O}$,

Allora la matrice A è la matrice triangolare $A = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & 1 & 1 & \cdots & 1 \end{bmatrix}$ *con determinante uguale ad* 1.

La trasformazione inversa è

$$\begin{aligned} y_1 &= z_1, \\ y_2 &= z_2 - z_1, \\ &\cdots \\ y_m &= z_m - z_{m-1}, \end{aligned} \qquad \textit{ossia } y = \varphi^{-1}(z) = A^{-1}y \textit{ dove } A^{-1} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix}$$

per cui, se Y ammette densità di probabilità,

$$f_Z(z_1, z_2, \cdots, z_m) = f_Y(z_1, z_2 - z_1, \cdots, z_m - z_{m-1}).$$

Il caso $m = 2$ è particolarmente interessante in quanto permette di ricavare la densità della somma di due variabili aleatorie, semplicemente calcolando la densità marginale di $Z_2 = Y_1 + Y_2$: per $z \in \mathbb{R}$

$$f_{Y_1+Y_2}(x)\Big(= f_{Z_2}(x)\Big) = \int_{\mathbb{R}} f_{Z_1,Z_2}(x, x')\, dx' = \int_{\mathbb{R}} f_{Y_1,Y_2}(x, x' - x)\, dx'.$$

3.2 Le traiettorie del processo di Wiener non sono a variazione limitata

In questa sezione si dimostra una proprietà che non permette di definire nel modo usuale (Lebesgue-Stieltjes) l'integrale rispetto al processo di Wiener. Nonostante si abbia che le traiettorie della versione continua del processo di Wiener siano hölderiane, tuttavia esse non possono essere molto regolari, in quanto ad esempio sono anche a variazione non limitata con probabilità 1 su ogni intervallo $[s,t]$. Infatti, per ogni versione continua[12] di W_u, se esistesse un intervallo $[s,t]$ ed un insieme A con $\mathbb{P}(A) > 0$, su cui la variazione[13] di W_u, cioè se

$$\begin{aligned} V(\omega) &= V(s,t,W;\omega) \\ &:= \sup\Big\{ \sum_{k=0}^{n-1} |W_{t_{k+1}}(\omega) - W_{t_k}(\omega)|, \text{ al variare delle partizioni } \pi : t_0 = s \le t_1 \le \cdots \le t_n = t \Big\} \end{aligned}$$

fosse finita per ogni $\omega \in A$, allora si avrebbe che, se $|\pi| := \max_k |t_{k+1} - t_k| \to 0$,

$$\begin{aligned} S_\pi &:= \sum_{k=0}^{n-1} |W_{t_{k+1}} - W_{t_k}|^2 \le \sum_{k=0}^{n-1} \left(\sup_h |W_{t_{h+1}} - W_{t_h}| \right) |W_{t_{k+1}} - W_{t_k}| \\ &= \left(\sup_h |W_{t_{h+1}} - W_{t_h}| \right) \sum_{k=0}^{n-1} |W_{t_{k+1}} - W_{t_k}| \le \left(\sup_{u,v\in[s,t], |u-v| \le |\pi|} |W_u - W_v| \right) V(\omega) \to 0. \end{aligned}$$

Questo fatto è in contraddizione con la seguente proprietà del moto browniano (proprietà che è valida anche per versioni non continue del processo di Wiener):

Lemma 3.2. *Si definisca*

$$S_\pi := \sum_{k=0}^{n-1} |W_{t_{k+1}} - W_{t_k}|^2,$$

allora S_π converge in media quadratica a $t-s$, ovvero $S_\pi \xrightarrow{L^2} (t-s)$, al tendere a zero di $|\pi| := \max_k |t_{k+1} - t_k|$.

Prima di passare alla dimostrazione notiamo che il significato di questo Lemma può essere espresso in modo euristico, dicendo che $\big(\Delta W_{t_k} big)^2 = \big(W_{t_k} - W_{t_{k-1}}\big)^2 \simeq t_k - t_{k-1}$ e in forma infinitesimale che $(dW_t)^2 = dt$.

Dimostrazione. Si tratta di mostrare che $\mathbb{E}[(S_\pi - (t-s))^2] \to 0$ al tendere a zero di $|\pi| := \max_k |t_{k+1} - t_k|$, e infatti

$$\mathbb{E}\left[(S_\pi - (t-s))^2\right] = 2 \sum_{k=0}^{n-1} |t_{k+1} - t_k|^2,$$

[12] Se il processo W_u ha le traiettorie continue, allora su ogni intervallo chiuso e limitato $[s,t]$, esse sono uniformemente continue, e quindi, per ogni ω

$$\sup_{u,v\in[s,t], |u-v|\le\delta} |W_u(\omega) - W_v(\omega)| \to 0 \qquad \text{per } \delta \to 0.$$

[13] Ricordiamo che se f è una funzione definita su un intervallo $[a,b]$ la variazione di f su $[a,b]$ è appunto definita come

$$\sup\{\sum_{k=0}^{n} |f(x_{k+1}) - f(x_k)|\},$$

dove l'estremo superiore è preso su tutte le partizioni $\{x_0, x_1, \cdots, x_n\}$ di $[a,b]$ tali che $x_0 = a \le x_1 \le \cdots \le x_n = b$.

come si vede facilmente[14]. Di conseguenza si ottiene che

$$
\begin{aligned}
\mathbb{E}\left[(S_\pi-(t-s))^2\right] &\le 2\sum_{k=0}^{n-1}|t_{k+1}-t_k|(\max_h|t_{h+1}-t_h|)\\
&=2(\max_h|t_{h+1}-t_h|)\sum_{k=0}^{n-1}|t_{k+1}-t_k| = 2|\pi|(t-s)\to 0.
\end{aligned}
$$

La conseguenza di questo risultato è che le traiettorie di un processo di Wiener non sono a variazione limitata[15], o meglio[16],

$$
\mathbb{P}\left(V(s,t,W;\omega) := \sup_\pi \sum_{k=0}^{n-1}|W_{t_{k+1}}-W_{t_k}| = \infty\right)=1.
$$

**Infatti per la disuguaglianza di Chebyshev il precedente Lemma implica che S_π converge in probabilità a $t-s$, mentre da quanto abbiamo visto sull'insieme degli ω in cui $V(\omega)$ è finito, necessariamente S_π converge a zero. **

Questo fatto in particolare pone dei problemi nel tentare di definire $\int_0^t f(s,\omega)dW_s(\omega)$, in quanto non si può adottare la classica definizione di integrale di Lebesgue-Stieltjes, che avrebbe senso solo se W_s avesse traiettorie a variazione finita. La definizione che si dà è quindi diversa e si parla in questo caso di integrale stocastico, inoltre per ottenerla si ha bisogno del concetto di martingala (Tuttavia cercheremo di dare la definizione anche senza questo concetto, ma sfrutteremo la proprietà di indipendenza degli incrementi del processo di Wiener.**

---

[14] Basta osservare che

$$
\begin{aligned}
&\mathbb{E}[(S_\pi-(t-s))^2] = \mathbb{E}[S_\pi^2+(t-s))^2-2S_\pi(t-s)]\\
&=\mathbb{E}\left[\Big(\sum_{k=0}^{n-1}|W_{t_{k+1}}-W_{t_k}|^2\Big)^2+(t-s)^2-2\Big(\sum_{k=0}^{n-1}|W_{t_{k+1}}-W_{t_k}|^2\Big)(t-s)\right]\\
&=\mathbb{E}\left[\sum_{k=0}^{n-1}|W_{t_{k+1}}-W_{t_k}|^2\sum_{h=0}^{n-1}|W_{t_{h+1}}-W_{t_h}|^2+(t-s)^2-2\Big(\sum_{k=0}^{n-1}|W_{t_{k+1}}-W_{t_k}|^2\Big)(t-s)\right]\\
&=\sum_{k=0}^{n-1}\mathbb{E}\left[|W_{t_{k+1}}-W_{t_k}|^4\right]+\sum_{k=0}^{n-1}\sum_{\substack{h=0\\h\neq k}}^{n-1}\mathbb{E}\left[|W_{t_{k+1}}-W_{t_k}|^2|W_{t_{h+1}}-W_{t_h}|^2\right]\ +\\
&\quad+(t-s)^2-2(t-s)\sum_{k=0}^{n-1}\mathbb{E}\left[|W_{t_{k+1}}-W_{t_k}|^2\right]\\
&=\sum_{k=0}^{n-1}3|t_{k+1}-t_k|^2+\sum_{k=0}^{n-1}\sum_{\substack{h=0\\h\neq k}}^{n-1}|t_{k+1}-t_k||t_{h+1}-t_h|+(t-s)^2-2(t-s)\sum_{k=0}^{n-1}|t_{k+1}-t_k|
\end{aligned}
$$

Di conseguenza si ottiene che

$$
\begin{aligned}
&\mathbb{E}\left[(S_\pi-(t-s))^2\right]\\
&=\sum_{k=0}^{n-1}2|t_{k+1}-t_k|^2+\sum_{k=0}^{n-1}\sum_{h=0}^{n-1}|t_{k+1}-t_k||t_{h+1}-t_h|+(t-s)^2-2(t-s)(t-s)\\
&=\sum_{k=0}^{n-1}2|t_{k+1}-t_k|^2+(t-s)^2+(t-s)^2-2(t-s)(t-s)=2\sum_{k=0}^{n-1}|t_{k+1}-t_k|^2
\end{aligned}
$$

[15] In realtà con la stessa tecnica del lemma si dimostra anche che il processo di Wiener non ha traiettorie di Hölder con esponente $\gamma>1/2$: se così fosse allora, posto $L_\gamma(\omega)$ la costante di Hölder, si avrebbe necessariamente

$$
S_\pi(\omega)\le\sum_{k=0}^{n-1}L_\gamma^2(\omega)|t_{k+1}-t_k|^{2\gamma}\le\sum_{k=0}^{n-1}L_\gamma^2(\omega)|t_{k+1}-t_k|\delta^{2\gamma-1}=L_\gamma^2(\omega)\delta^{2\gamma-1}(t-s)\to 0,\ q.c.
$$

[16] **Si noti che, per $\pi: t_0=s\le t_1\le\cdots\le t_n=t$ posto $V_\pi(\omega):=\sum_{k=0}^{n-1}|W_{t_{k+1}}(\omega)-W_{t_k}(\omega)|$, per la continuità delle traiettorie $t\mapsto W_t(\omega)$, si ottiene che $V(\omega)=V(s,t,W;\omega)=\sup\{V_{\pi^m};\ \pi=\{t_0=s,\ t_k=s+\frac{k}{2^m}(t-s),\ k=1,\cdots,2^m\}\ m\in\mathbb{N}\}$, e che la successione V_{π^m} è una successione monotona non decrescente, quindi $V(\omega)=\lim_{m\to\infty}V_{\pi^m}(\omega)$, per ogni ω.

**A questo proposito facciamo notare che non si può definire l'integrale stocastico come si definisce l'usuale integrale di Riemann-Stieltjes.

Esempio 3.6 (Come definire $\int_a^b 2W_s\,dW_s$?)**.** *Prima di tutto ricordiamo alcune formule che sono banalmente valide: sia $\{b_k\}$ una successione di numeri reali, allora*

$$
\begin{aligned}
2 \sum_{k_{min} \le k \le k_{max}-1} b_k(b_{k+1}-b_k) &= -\sum_k [b_{k+1}-b_k]^2 + [b_{k_{max}}^2 - b_{k_{min}}^2],\\
2 \sum_{k_{min} \le k \le k_{max}-1} b_{k+1}(b_{k+1}-b_k) &= \sum_k [b_{k+1}-b_k]^2 + [b_{k_{max}}^2 - b_{k_{min}}^2],
\end{aligned}
$$

la cui dimostrazione è molto semplice[17]*.*

Posto $b_k = W_{t_k}$ con $t_{k_{min}} = a$ e $t_{k_{max}} = b$, le precedenti formule, insieme al Lemma 3.2, danno rispettivamente i seguenti limiti (in L^2):

$$
\lim_{n\to\infty} 2 \sum_{k_{min} \le k \le k_{max}-1} W_{t_k}(W_{t_{k+1}} - W_{t_k}) = -\lim_{n\to\infty} \sum_k [W_{t_{k+1}} - W_{t_k}]^2 + W_b^2 - W_a^2 = -b - a + W_b^2 - W_a^2
$$

$$
\lim_{n\to\infty} 2 \sum_{k_{min} \le k \le k_{max}-1} W_{t_{k+1}}(W_{t_{k+1}} - W_{t_k}) = \lim_{n\to\infty} \sum_k [W_{t_{k+1}} - W_{t_k}]^2 + W_b^2 - W_a^2 = b - a + W_b^2 - W_a^2.
$$

In conclusione, come del resto c'era da aspettarsi, anche se le somme di Riemann

$$
2 \sum_{k_{min} \le k \le k_{max}-1} W_{t_k^*}(W_{t_{k+1}} - W_{t_k}) \qquad (\text{con } t_k^* \in [t_k, t_{k+1}])
$$

convergono per alcune scelte di t_k^*, ma i limiti dipendono dalla scelta del punto t_k^*.

---

[17] Per ogni successione $\{b_k\}$ si ha

$$
\begin{aligned}
2 \sum_{k_{min} \le k \le k_{max}-1} b_k(b_{k+1}-b_k) &= \sum_k [2b_{k+1}b_k - 2b_k^2] = \sum_k [2b_{k+1}b_k - 2b_k^2 - b_{k+1}^2 + b_{k+1}^2]\\
&= \sum_k [(2b_{k+1}b_k - b_k^2 - b_{k+1}^2) + b_{k+1}^2 - b_k^2] = -\sum_k [b_{k+1}-b_k]^2 + \sum_k [b_{k+1}^2 - b_k^2]\\
&= -\sum_k [b_{k+1}-b_k]^2 + [b_{k_{max}}^2 - b_{k_{min}}^2].
\end{aligned}
$$

Invece

$$
\begin{aligned}
2 \sum_{k_{min} \le k \le k_{max}-1} b_{k+1}(b_{k+1}-b_k) &= \sum_k [2b_{k+1}^2 - 2b_{k+1}b_k] = \sum_k [2b_{k+1}^2 + b_k^2 - b_k^2 - 2b_{k+1}b_k]\\
{**}\ &= \sum_k [(b_{k+1}^2 - 2b_{k+1}b_k + b_k^2) + b_{k+1}^2 - b_k^2] = \sum_k [b_{k+1}-b_k]^2 + \sum_k [b_{k+1}^2 - b_k^2]\\
&= \sum_k [b_{k+1}-b_k]^2 + [b_{k_{max}}^2 - b_{k_{min}}^2].
\end{aligned}
$$

3.3 Appendice: variabili Gaussiane

Cominciamo con il definire una variabile aleatoria gaussiana standard unidimensionale:

Definizione 3.5. *Si dice che una variabile aleatoria reale Z è* ***gaussiana*** *di valore atteso μ e varianza σ^2, se ammette densità*

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\}.$$

In questo caso si usa la notazione $Z \sim N(\mu, \sigma^2)$. Se $\mu = 0$ e $\sigma^2 = 1$ allora si dice che Z segue una ***legge normale o gaussiana standard****.*

Caso $n-$dimensionale: iniziamo con il caso di un vettore (colonna) aleatorio

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \\ \cdots \\ Y_k \\ \cdots \\ Y_n \end{pmatrix}$$

a componenti indipendenti e tutte gaussiane standard, ovvero il caso in cui

$$\begin{aligned} f_Y(\boldsymbol{y}) &= \prod_{i=1}^n f_{Y_i}(y_i) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}y_i^2\right\} \\ &= \frac{1}{(\sqrt{2\pi})^n} \exp\left\{-\frac{1}{2}\sum_{i=1}^n y_i^2\right\} = \frac{1}{(2\pi)^{n/2}} \exp\left\{-\frac{1}{2}\boldsymbol{y}'\boldsymbol{y}\right\}. \end{aligned}$$

dove l'apice indica l'operazione di trasposizione, ovvero $\boldsymbol{y}'$ è il vettore riga $(y_1, y_2, \cdots, y_n)$.

È immediato verificare che $\mathbb{E}(Y_i) = 0$, $Var(Y_i) = 1$ e che $Cov(Y_i, Y_j) = 0$, per $i \neq j$.

Sia ora A una matrice non singolare e sia $\boldsymbol{m}$ un vettore (colonna). Definiamo ora $Z = AY + \boldsymbol{m}$ e cerchiamo la sua densità. Sappiamo dai risultati generali che se Y ammette densità e $Z = \varphi(Y)$ con φ invertibile e con derivate continue, allora anche Z ammette densità:

$$f_Z(z) = f_Y(\varphi^{-1}(z)) \left| det\left(\frac{\partial\varphi^{-1}(z)}{\partial z}\right)\right| = f_Y(\varphi^{-1}(z)) \frac{1}{\left|det\left(\frac{\partial\varphi(y)}{\partial y}\right)\right|_{y=\varphi^{-1}(z)}}$$

di conseguenza, poiché nel nostro caso $\varphi(y) = Ay + m$ e $\varphi^{-1}(z) = A^{-1}(z-m)$

$$f_Z(z) = \frac{1}{\sqrt{2\pi}^n} \exp\left\{-\frac{1}{2}\left(A^{-1}(z-m)\right)' A^{-1}(z-m)\right\} \frac{1}{|det(A)|}.$$

Essendo

$$\begin{aligned} &(A^{-1}(z-m))'A^{-1}(z-m) = (z-m)'(A^{-1})'A^{-1}(z-m) \\ &= (z-m)'(A')^{-1}A^{-1}(z-m) = (z-m)'(AA')^{-1}(z-m) \end{aligned}$$

si ottiene

$$f_Z(z) = \frac{1}{(2\pi)^{n/2}} \frac{1}{|det(A)|} \exp\left\{-\frac{1}{2}(z-m)'\left(AA'\right)^{-1}(z-m)\right\}.$$

La precedente espressione si basa sulle seguenti proprietà:

(i) $\boxed{(A')^{-1} = (A^{-1})'}$

in quanto

$$A'z = w \Leftrightarrow z = (A')^{-1}w$$

e inoltre

$$A'z = w \Leftrightarrow (z'A)' = w \Leftrightarrow z'A = w' \Leftrightarrow z' = w'A^{-1}$$

$$\Leftrightarrow z = \left(w'A^{-1}\right)' \Leftrightarrow z = \left(A^{-1}\right)' w.$$

(ii) $\boxed{(AA')^{-1} = (A')^{-1} A^{-1}}$

in quanto

$$(AA')^{-1} z = w \Leftrightarrow z = AA'w \Leftrightarrow A^{-1}z = A'w \Leftrightarrow (A')^{-1} A^{-1}z = w.$$

È interessante notare che sia il vettore $\boldsymbol{m}$ che la matrice $AA' = A'A$ hanno una interpretazione probabilistica:

$$\mathbb{E}(Z_i) = \mathbb{E}(\sum_{k=1}^{n} a_{i,k}Y_k) + m_i = \sum_{k=1}^{n} a_{i,k}\mathbb{E}(Y_k) + m_i = m_i$$

$$Cov(Z_i, Z_j) = \mathbb{E}[(Z_i - m_i)(Z_j - m_j)] = \mathbb{E}[\sum_{k=1}^{n} a_{i,k}Y_k \sum_{h=1}^{n} a_{j,h}Y_h] = \sum_{k=1}^{n}\sum_{h=1}^{n} a_{i,k}a_{j,h}\mathbb{E}[Y_kY_h]$$

e quindi

$$Cov(Z_i, Z_j) = \sum_{k=1}^{n} a_{i,k}a_{j,k}\mathbb{E}[Y_kY_k] + \sum_{k=1}^{n}\sum_{h\neq k}^{1,n} a_{i,k}a_{j,h}\mathbb{E}[Y_kY_h] = \sum_{k=1}^{n} a_{i,k}a_{j,k} = (AA')_{i,j}$$

Si osservi che se $\boldsymbol{Z} = (Z_1, ..., Z_n)$ è un vettore gaussiano allora $(Z_1...Z_k)$ e $(Z_{k+1}, ..., Z_n)$ sono indipendenti, se e solo se $Cov(Z_i, Z_h) = 0$ per ogni $i = 1, \cdots, k$ e $h = k+1, \cdots, n$. In tale caso allora è ovvio che il vettore $(Z_1...Z_k)$ è un vettore gaussiano[18]

Terminiamo questo paragrafo con il ricordare quanto valgono i ***momenti di una variabile aleatoria gaussiana***. Sia Z una variabile aleatoria $\mathcal{N}(0, \sigma^2)$. Per quanto visto prima possiamo considerare $Z = \sigma Y$ con Y una variabile aleatoria $\mathcal{N}(0,1)$. Da questa osservazione segue subito che

$$\mathbb{E}[Z^k] = \sigma^k\mathbb{E}[Y^k]$$

e

$$\mathbb{E}[|Z|^k] = |\sigma|^k\mathbb{E}[|Y|^k].$$

[18]Per ottenere lo stesso risultato nel caso generale, ovvero che se $\boldsymbol{Z} = (Z_1, ..., Z_n)$ è un vettore gaussiano allora $(Z_1...Z_k)$ è un vettore gaussiano, si può procedere nel seguente modo. Innanzitutto basta considerare il caso in cui i valori attesi sono nulli senza ledere in generalità. Inoltre si può pensare che $\boldsymbol{Z} = A\boldsymbol{Y}$. Se la matrice $A' = (a'_{ij})$ è definita in modo che $a'_{ij} = a_{ij}$ qualunque siano $i = 1, ...k$ e $j = 1, ..., n$, e il vettore aleatorio $\boldsymbol{Z}'$ è definito da

$$\boldsymbol{Z}' = A'\boldsymbol{Y},$$

allora, chiaramente,

$$Z'_i = (A'\boldsymbol{Y})_i = Z_i = (A\boldsymbol{Y})_i, \quad \text{per } i = 1, ...k.$$

Se inoltre a'_{hj} per $h = k+1, ...n$ e $j = 1, ..., n$ sono presi in modo che il vettore $(Z'_1, \cdots, Z'_k) = (Z_1, \cdot, Z_k)$ sia indipendente dal vettore $(Z'_{k+1}, \cdots, Z'_n)$, ovvero in modo che

$$0 = E[Z_iZ'_h] = Cov(Z_i, Z'_h) = \sum_{\ell=1}^{n} a_{i,\ell}a'_{h,\ell}$$

per $i = 1, ...k$ e $h = k+1, ...n$, allora si ottiene il risultato voluto.
Nel caso in cui la matrice A sia non singolare ciò è sempre possibile perché i vettori $\boldsymbol{a}_{(i)} = (a_{i1}, a_{i2}, \cdots, a_{in})$ sono linearmente indipendenti e quindi basta trovare $n-k$ vettori $\boldsymbol{a}'_{(h)} = (a'_{i1}, a'_{h2}, \cdots, a'_{hn})$ ortogonali allo spazio vettoriale k-dimensionale $span(\boldsymbol{a}_{(i)}, i = 1, \cdot, k)$.

Vale poi la pena di ricordare che

$$\mathbb{E}[Y^{2k+1}] = 0,$$

$$\mathbb{E}[Y^{2k}] = (2k-1)!! = (2k-1)(2k-3)\cdots 5\cdot 3\cdot 1,$$

mentre[19] infine

$$\mathbb{E}[|Y|^{2k+1}] = \sqrt{\frac{2}{\pi}}(2k)!! = \sqrt{\frac{2}{\pi}}(2k)(2k-2)\cdots 4\cdot 2 = \sqrt{\frac{2}{\pi}}2^k k!.$$

Prima di dimostrare queste tre uguaglianze si osservi che le ultime due si possono scrivere in modo sintetico come

$$\mathbb{E}[|Y|^n] = C_{((-1)^n)}\,(n-1)!! \qquad C_{(+1)} = 1 \qquad C_{(-1)} = \sqrt{\frac{2}{\pi}}.$$

La prima relazione è banale, per ragioni di simmetria, e permette di ricavare la seconda osservando che

$$\mathbb{E}[\mathrm{e}^{uY}] = \mathrm{e}^{\frac{u^2}{2}} = \sum_{h=0}^{\infty}\frac{1}{h!}\frac{u^2}{2}^h = \sum_{h=0}^{\infty}\frac{1}{h!}\frac{u^{2h}}{2^h}.$$

e d'altra parte, essendo appunto ovviamente $\mathbb{E}[Y^{2k+1}] = 0$,

$$\mathbb{E}[\mathrm{e}^{uY}] = \mathbb{E}[\sum_{k=0}^{\infty}\frac{1}{k!}u^k Y^k] = \sum_{h=0}^{\infty}\frac{1}{(2h)!}u^{2h}\mathbb{E}[Y^{2h}]$$

si deve necessariamente avere che i coefficienti delle due serie devono coincidere:

$$\frac{1}{h!}\frac{1}{2^h} = \frac{1}{(2h)!}\mathbb{E}[Y^{2h}],$$

ovvero

$$\begin{aligned}\mathbb{E}[Y^{2h}] &= \frac{(2h)!}{h!2^h} = \frac{2h(2h-1)(2h-2)(2h-3)\cdots 3\cdot 2\cdot 1}{h(h-1)\cdots 3\cdot 2\cdot 1\cdot 2^h}\\ &= \frac{(2h)!!(2h-1)!!}{h!2^h} = \frac{2^h h!\cdot(2h-1)!!}{2^h h!} = (2h-1)!!.\end{aligned}$$

Infine la terza si ricava per integrazione per parti e calcolando a mano che $\mathbb{E}[|Y|] = \sqrt{\frac{2}{\pi}}$.

[19] Si noti che dalle ultime due relazioni sui momenti si ottiene che

$$\mathbb{E}[|Y|^m] = (m-1)!!C_{(-1)^m}, \quad \text{con } C_{+1} = 1, C_{-1} = \sqrt{\frac{2}{\pi}}.$$

Capitolo 4

Processi a incrementi indipendenti e martingale

4.1 Processi ad incrementi indipendenti ed omogenei

Il procedimento usato per ottenere le distribuzioni finito dimensionali del processo di Wiener, si può estendere ad una classe più generale:

Definizione 4.1 (Processi ad incrementi indipendenti ed omogenei). *Un processo* $(X_t)_{t\geq 0}$ *si dice* **ad incrementi indipendenti ed omogenei** *se*

(0) $X_0 = 0$;

(1) *per* $0 = t_0 < t_1 < \cdots t_n$ *gli incrementi* $\Delta X_{t_i} = X_{t_i} - X_{t_{i-1}}$ *sono variabili aleatorie indipendenti;*

(2) *gli incrementi* $\Delta X_{t_i} = X_{t_i} - X_{t_{i-1}}$ *sono variabili aleatorie la cui distribuzione dipende solo dall'ampiezza dell'intervallo* $(t_i - t_{i-1})$*;*

Per fissare le idee e capire meglio la condizione **(2)**, si consideri la famiglia $(F_u)_{u\geq 0}$ di funzioni di distribuzione dipendente da un parametro, per cui $X_{t+u} - X_t \sim F_u$, qualunque siano t ed u in $[0, \infty)$.

La famiglia $(F_u)_{u\geq 0}$ non può essere presa a piacere, ma deve soddisfare la seguente condizione necessaria[1]:

$$F_u * F_v = F_{u+v}, \qquad \text{per ogni } u, v \geq 0, \tag{4.1}$$

dove $*$ corrisponde alla convoluzione.

Infatti ciò corrisponde alla condizione che

$$X_u = X_u - X_0 \sim F_u, \quad X_{u+v} - X_u \sim F_v, \quad X_{u+v} \sim F_{u+v},$$

e d'altra parte

$$X_{u+v} = (X_{u+v} - X_u) + (X_u - X_0) \sim F_v * F_u.$$

In realtà questa condizione risulta anche sufficiente, come si può verificare facilmente. La dimostrazione è riportata per completezza, ma non è necessaria per gli studenti del Master.

Infatti le funzioni $F_{t_1,\cdots,t_k}(x_1, \cdots, x_k)$ di distribuzione finito dimensionale risultano definite[2], per $0 < t_1 < \cdots < t_k$, come, la funzione di distribuzione ottenuta da quella degli incrementi, cioè

$$\mathbb{P}\left(X_{t_1} - X_0 \leq z_1, X_{t_2} - X_{t_1} \leq z_2, \cdots X_{t_k} - X_{t_{k-1}} \leq z_k\right) = F_{t_1}(z_1) F_{t_2 - t_1}(z_2) \cdots F_{t_k - t_{k-1}}(z_k),$$

[1] Inoltre è necessario che $F_0(x) = 0$ per $x < 0$ ed $F_0(x) = 1$ per $x \geq 0$, ovvero che l'*incremento* $X_t - X_t$ sia concentrato nello 0.

[2] Nel caso $0 = t_0 < t_1 < \cdots < t_k$ si ottiene immediatamente che

$$F_{t_0,t_1,\cdots,t_k}(x_0, x_1, \cdots, x_k) = 0, \qquad \text{per } x_0 < 0,$$

$$F_{t_0,t_1,\cdots,t_k}(x_0, x_1, \cdots, x_k) = F_{t_1,\cdots,t_k}(x_1, \cdots, x_k), \qquad \text{per } x_0 \geq 0.$$

attraverso la trasformazione[3] $x_1 = z_1$, $x_2 = z_1 + z_2, \cdots, x_k = z_1 + \cdots + z_k$. La condizione (4.1) implica immediatamente che la condizione di consistenza di Kolmogorov **(C2′)** sia soddisfatta.

Per rendere piú concreta la verifica, si consideri, ad esempio, il caso con densità, ovvero il caso in cui

$$F_u(x) = \int_{-\infty}^{x} q_u(y)dy.$$

Procedendo come per il processo di Wiener, si ottiene che, per $0 < t_1 < \cdots < t_k$

$$F_{t_1,\cdots,t_k}(x_1,\cdots,x_k) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_k} q_{t_1}(y_1)\cdots q_{t_k - t_{k-1}}(y_k - y_{k-1})dy_1 \cdots dy_k.$$

Per controllare la condizione di consistenza **(C2′)**, si prendano $k \geq 1$ e $t_1 < t_2 < \cdots < t_k < t_{k+1}$, allora la 3.3) è verificata:

$$\begin{aligned}
&\lim_{x\to\infty} F_{t_1,\cdots,t_k,t_{k+1}}(x_1,\cdots,x_k,x)\\
&= \lim_{x\to\infty} \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_k} \int_{-\infty}^{x} q_{t_1}(y_1)\cdots q_{t_k - t_{k-1}}(y_k - y_{k-1}) q_{t_{k+1}-t_k}(y - y_k) dy_1 \cdots dy_k\, dy\\
&= \lim_{x\to\infty} \int_{-\infty}^{x_1} q_{t_1}(y_1)\, dy_1 \cdots \int_{-\infty}^{x_k} q_{t_k - t_{k-1}}(y_k - y_{k-1})\, dy_k \int_{-\infty}^{x} \cdots q_{t_{k+1}-t_k}(y - y_k)\, dy\\
&= \lim_{x\to\infty} \int_{-\infty}^{x_1} q_{t_1}(y_1)\, dy_1 \cdots \int_{-\infty}^{x_k} q_{t_k - t_{k-1}}(y_k - y_{k-1})\, dy_k \int_{-\infty}^{x-y_k} \cdots q_{t_{k+1}-t_k}(y')\, dy'\\
&= \int_{-\infty}^{x_1} q_{t_1}(y_1)\, dy_1 \cdots \int_{-\infty}^{x_k} q_{t_k - t_{k-1}}(y_k - y_{k-1})\, dy_k = F_{t_1,\cdots,t_k}(x_1,\cdots,x_k),
\end{aligned}$$

e la (3.4) anche:

$$\begin{aligned}
&\lim_{x\to\infty} F_{t_1,\cdots,t_{i-1},t_i,t_{i+1},\cdots,t_{k+1}}(x_1,\cdots,x_{i-1},x,x_{i+1},\cdots,x_{k+1})\\
&= \lim_{x\to\infty} \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_{i-1}} \int_{-\infty}^{x} \int_{-\infty}^{x_{i+1}} \cdots \int_{-\infty}^{x_{k+1}} q_{t_1}(y_1)\cdots q_{t_{i-1}-t_{i-2}}(y_{i-1} - y_{i-2})\\
&\qquad q_{t_i - t_{i-1}}(y_i - y_{i-1}) q_{t_{i+1}-t_i}(y_{i+1} - y_i) \cdots q_{t_{k+1}-t_k}(y_{k+1} - y_k) dy_1 \cdots dy_{i-1}\, dy_i\, dy_{i+1} \cdots dy_{k+1}\\
&= \lim_{x\to\infty} \int_{-\infty}^{x_1} q_{t_1}(y_1)\, dy_1 \cdots \int_{-\infty}^{x_{i-1}} q_{t_{i-1}-t_{i-2}}(y_{i-1} - y_{i-2})\, dy_{i-1}\\
&\qquad \int_{-\infty}^{x_{i+1}} \left(\int_{-\infty}^{x} q_{t_i - t_{i-1}}(y_i - y_{i-1})\, q_{t_{i+1}-t_i}(y_{i+1} - y_i)\, dy_i \right) dy_{i+1} \cdots\\
&\qquad\qquad \cdots \int_{-\infty}^{x_{k+1}} q_{t_{k+1}-t_k}(y_{k+1} - y_k)\, dy_{k+1}\\
&= \int_{-\infty}^{x_1} q_{t_1}(y_1)\, dy_1 \cdots \int_{-\infty}^{x_{i-1}} q_{t_{i-1}-t_{i-2}}(y_{i-1} - y_{i-2})\, dy_{i-1}\\
&\qquad \int_{-\infty}^{x_{i+1}} q_{t_{i+1}-t_i}(y_{i+1} - y_i)\, dy_{i+1} \cdots \int_{-\infty}^{x_{k+1}} q_{t_{k+1}-t_k}(y_{k+1} - y_k)\, dy_{k+1}\\
&= F_{t_1,\cdots,t_{i-1},t_{i+1},\cdots,t_{k+1}}(x_1,\cdots,x_{i-1},x_{i+1},\cdots,x_{k+1}), \quad \text{per } i = 1,\cdots,k.
\end{aligned}$$

Nella penultima uguaglianza si è tenuto conto del fatto che la condizione $F_u * F_v = F_{u+v}$ implica che, posto $y = y_i - y_{i-1}$, in modo che $y_{i+1} - y_i = (y_{i+1} - y_{i-1}) - y$, si abbia

$$\begin{aligned}
&\lim_{x\to\infty} \int_{-\infty}^{x} q_{t_i - t_{i-1}}(y_i - y_{i-1})\, q_{t_{i+1}-t_i}(y_{i+1} - y_i)\, dy_i\\
&= \lim_{x\to\infty} \int_{-\infty}^{x - y_{i-1}} q_{t_i - t_{i-1}}(y) q_{t_{i+1}-t_i}(y_{i+1} - y_i - y)\, dy\\
&= \int_{-\infty}^{\infty} q_{t_i - t_{i-1}}(y) q_{t_{i+1}-t_i}(y_{i+1} - y_i - y)\, dy\\
&= q_{t_{i+1}-t_i} * q_{t_i - t_{i-1}}(y_{i+1} - y_{i-1}) = q_{t_{i+1}-t_{i-1}}(y_{i+1} - y_{i-1}).
\end{aligned}$$

Nel caso in cui F_v sia discreta il discorso si ripete identico, mettendo le densità discrete al posto delle densità di probabilità e le somme al posto degli integrali.

Come esempi di famiglie ad un parametro di funzioni di distribuzione, oltre al caso della famiglia gaussiana $F_u \sim N(0, u)$, che dà luogo al processo di Wiener standard, si possono considerare

1 il ***processo di Wiener con drift*** (o deriva) μ e ***coefficiente di diffusione*** σ^2, ovvero

$$F_u \sim N(\mu u, \sigma^2 u);$$

[3] Si tratta solo di notare che se $Z_i := X_{t_i} - X_{t_{i-1}}$ allora $(X_{t_1}, X_{t_2}, \cdots, X_{t_k}) = (Z_1, Z_1 + Z_2, \cdots, Z_1 + \cdots + Z_k)$

2 il ***processo di Cauchy***, ovvero il caso in cui

$$F_u \sim Cauchy(u), \text{ ovvero } q_u(x) = \frac{u}{\pi}\frac{1}{u^2+x^2};$$

3 il ***processo di Poisson*** di ***parametro*** λ, ovvero il caso in cui

$$F_u \sim Poisson(\lambda u), \text{ ovvero } p_u(k) = F_u(k) - F_u(k-1) = \frac{(\lambda u)^k}{k!}\exp(-\lambda u), \quad k \in \mathbb{N};$$

4 i ***processi di Poisson composti***, ovvero i processi $(X_t)_{t\geq 0}$ ottenuti per mezzo di un processo di Poisson $(N_t)_{t\geq 0}$ ed una successione di variabili aleatorie indipendenti ed identicamente distribuite $\{\xi_n\}_{n\in\mathbb{N}}$, (tutte indipendenti dal processo di Poisson) tramite la seguente regola

$$X_t = 0, \text{ se } N_t = 0; \qquad X_t = \sum_{k=1}^{n} \xi_k, \text{ se } N_t = n.$$

Terminiamo questa sezione ricordando anche la definizione di processi ad incrementi indipendenti rispetto ad una filtrazione $(\mathcal{F}_t)_{t\geq 0}$.

Definizione 4.2 (Processi ad incrementi indipendenti ed omogenei rispetto ad una filtrazione)**.** *Un processo* $(X_t)_{t\geq 0}$ *si dice* **ad incrementi indipendenti ed omogenei rispetto alla filtrazione** $(\mathcal{F}_t)_{t\geq 0}$**, con** $(\mathcal{F}_t)_{t\geq 0} \supseteq \mathcal{F}_t^X = \sigma\{X_u;\ u \in [0,t]\}$ *se*

(0) $X_0 = 0$*;*

(1) *per* $s, t \geq 0$ *gli incrementi* $X_{t+s} - X_t$ *sono variabili aleatorie indipendenti da* $\mathcal{F}_t$*;*

(2) *gli incrementi* $X_{t+s} - X_t$ *sono variabili aleatorie la cui distribuzione dipende solo da* s*.*

Si vede facilmente che questa definizione implica l'altra considerando che $X_{t_n} - X_{t_{n-1}}$ è indipendente da $\mathcal{F}_{t_{n-1}} \supseteq \sigma\{X_{t_i} - X_{t_{i-1}};\ i = 1, \cdots, n-1\}$. Si può anche vedere che la prima definizione implica la seconda con $\mathcal{F}_t = \mathcal{F}_t^X = \sigma\{X_u;\ u \in [0,t]\} = \sigma\{X_u - X_v;\ u, v \in [0,t]\}$.

4.2 Esempi di martingale a tempo continuo

Dimostreremo ora che se X_t è un processo ad incrementi indipendenti (e omogenei), rispetto ad una filtrazione $(\mathcal{F}_t)_{t\geq 0}$, con $X_0 = 0$, integrabile e con media nulla allora gode della seguente proprietà:

1) per ogni $t \geq 0$, la variabile aleatoria X_t è $\mathcal{F}_t$ misurabile (infatti $\mathcal{F}_t \supseteq \mathcal{F}_t^X$, e X_t è $\mathcal{F}_t^X$-misurabile)
2) per ogni $t \geq 0$, la variabile aleatoria X_t è integrabile (X_t è integrabile, per ipotesi)
3) per ogni $t, s \geq 0$, vale la seguente relazione:

$$\mathbb{E}\big[X_{t+s}\,\mathbb{I}_A\big] = \mathbb{E}\big[X_t\,\mathbb{I}_A\big], \qquad \text{per ogni } A \in \mathcal{F}_t.$$

La verifica della proprietà **3)** è immediata, non appena si nota che equivale a chiedere

$$\mathbb{E}\big[(X_{t+s} - X_t)\,\mathbb{I}_A\big] = 0, \qquad \text{per ogni } A \in \mathcal{F}_t,$$

e che, poiché l'incremento $X_{t+s} - X_t$ è indipendente da $\mathcal{F}_t$ e $A \in \mathcal{F}_t$, si ha

$$\mathbb{E}\big[(X_{t+s} - X_t)\,\mathbb{I}_A\big] = \mathbb{E}\big[X_{t+s} - X_t\big]\,\mathbb{E}\big[\mathbb{I}_A\big] = 0, \qquad \text{per ogni } A \in \mathcal{F}_t,$$

in quanto $\mathbb{E}\big[X_{t+s} - X_t\big] = \mathbb{E}\big[X_{t+s}\big] - \mathbb{E}\big[X_t\big] = 0$.

Definizione 4.3 (MARTINGALA)**.** *Più in generale un processo* $(X_t)_{t\geq 0}$ *(non necessariamente un processo a incrementi indipendenti, e non necessariamente a media nulla) che goda delle precedenti tre proprietà è detto essere una* ***martingala rispetto alla filtrazione*** $(\mathcal{F}_t))_{t\geq 0}$.

Inoltre la proprietà **3)** per una martingala viene di solito formulata nel seguente modo:
3) per ogni t, $s \geq 0$, vale la seguente relazione:

$$\mathbb{E}\big[X_{t+s}|F_t\big] = X_t,$$

che si legge il valore atteso condizionale di X_{t+s} dato $\mathcal{F}_t$ coincide con X_t.

Il motivo deriva dalla definizione di valore atteso condizionale di una variabile aleatoria rispetto ad una sigma-algebra:

Definizione 4.4 (VALORE ATTESO CONDIZIONALE)**.** *Sia $(\Omega, \mathcal{F}, \mathbb{P})$ uno spazio di probabilità. Sia X una variabile aleatoria $\mathcal{F}$ misurabile e integrabile. Sia infine $\mathcal{G}$ una sigma-algebra contenuta in $\mathcal{F}$.*

Sia infine $\widetilde{X}$ una variabile aleatoria $\mathcal{G}$-misurabile e tale che

$$\mathbb{E}\big[X\,\mathbb{I}_A\big] = \mathbb{E}\big[\widetilde{X}\,\mathbb{I}_A\big] \qquad \textit{per ogni } A \in \mathcal{G}.$$

Allora $\widetilde{X}$ è una versione del valore atteso condizionale di X dato $\mathcal{G}$.

La classe di tutte le variabili aleatorie che verificano la precedente proprietà viene detta valore atteso di X dato $\mathcal{G}$ e viene indicata con

$$\mathbb{E}\big[X|\mathcal{G}\big].$$

Ovviamente ogni altra variabile aleatoria $\widetilde{X}'$, che sia $\mathcal{G}$ misurabile e che differisca da $\widetilde{X}$, su un insieme ($\mathcal{G}$ misurabile) di misura nulla, è una versione del valore atteso condizionale di X dato $\mathcal{G}$. E vale anche il viceversa, cioè se $\widetilde{X}$ e $\widetilde{X}'$ sono due versioni del valore atteso condizionale di X dato $\mathcal{G}$, allora $\widetilde{X}$ e $\widetilde{X}'$ differiscono su un insieme ($\mathcal{G}$ misurabile) di misura nulla.[4]. Alcune proprietà del valore atteso condizionale sono riportate in appendice.

La definizione di Martingala si estende anche al caso di processi del tipo $(X_t)_{t\in I}$, con I un intervallo (ad esempio $[0,T]$ o un sottoinsieme discreto di $\mathbb{R}$ (ad esempio $\mathbb{N}$, ed in tal caso si parla di successioni aleatorie, o nel caso più generale di $I = \{t_k\}_{k\in\mathbb{N}}$, di processi a tempo discreto). La definizione di martingala va però leggermente cambiata (in quanto non è ovvio che se $s, t \in I$ allora $s + t \in I$). Inoltre si può dare anche la definizione di submartingala e di supermartingala

Definizione 4.5 (Martingale, Submartingale e Supermartingale)**.** *Data una filtrazione $(\mathcal{F}_t)_{t\in I}$ (cioè per ogni $t_1, t_2 \in I$ con $t_1 \leq t_2$ si ha $\mathcal{F}_{t_1} \subseteq \mathcal{F}_{t_2}$), un processo $(X_t)_{t\in I}$ si dice **martingala** rispetto a $(\mathcal{F}_t)_{t\in I}$ se verifica le seguenti proprietà:*
***1)** per ogni $t \in I$, la variabile aleatoria X_t è $\mathcal{F}_t$ misurabile, ossia il processo $(X_t)_{t\in I}$ è $\mathcal{F}_t$-adattato;*

[4]Rimarrebbe da controllare che la definizione di valore atteso condizionale è ben posta, ossia esistenza e unicità. Per l'unicità non ci sono problemi perché, come si fa usualemente per gli spazi L^p, si è usata una classe di equivalenza. Per l'esistenza invece bisognerebbe usare il Teorema di Radon-Nikodym. Qui ci limitiamo a ricordare il fatto che se la variabile aleatoria X è di quadrato integrabile, allora si può dare una seconda definizione di valore atteso condizionale, sempre come classe di equivalenza delle variabili aleatorie $\widetilde{X}_2$, che siano $\mathcal{G}$-misurabili, di quadrato integrabile e tali che

$$\mathbb{E}\big[(X-\widetilde{X}_2)^2\big] \leq \mathbb{E}\big[(X-Z)^2\big] \qquad \text{per ogni variabile aleatoria } Z\text{, che sia } \mathcal{G}\text{-misurabile e di quadrato integrabile.}$$

Considerando lo spazio delle variabili aleatorie $\mathcal{F}$ misurabili di quadrato integrabile come uno spazio di Hilbert con il prodotto scalare dato da $< X, Y >= \mathbb{E}\big[XY\big]$, la seconda definizione di valore atteso condizionale risulta quindi equivalente alla definizione di proiezione ortogonale del *vettore* X nel sottospazio lineare (chiuso) delle variabili aleatorie $\mathbb{G}$ misurabili e di quadrato integrabile, ossia $L^2(\Omega, \mathcal{G}, \mathbb{P})$.
La condizione che caratterizza tali proiezioni diviene quindi
per ogni Z in $L^2(\Omega, \mathcal{G}, \mathbb{P})$ $< X - \widetilde{X}_2, Z >= 0$, ossia

$$\mathbb{E}\big[Z(X-\widetilde{X}_2)\big] = 0 \quad \Leftrightarrow \mathbb{E}\big[X\,Z\big] = \mathbb{E}\big[\widetilde{X}_2\,Z\big] \quad \text{per ogni } Z \text{ in } L^2(\Omega, \mathcal{G}, \mathbb{P}).$$

Rimane solo da osservare che la condizione della definizione iniziale di valore atteso condizionale, ossia

$$\mathbb{E}\big[X\,\mathbb{I}_A\big] = \mathbb{E}\big[\widetilde{X}\,\mathbb{I}_A\big] \qquad \text{per ogni } A \in \mathcal{G}.$$

si estende facilmente a variabili aleatorie W che sono combinazioni lineari di funzioni indicatrici di insiemi $\mathcal{G}$ misurabili, ossia

$$\mathbb{E}\big[X\,W\big] = \mathbb{E}\big[\widetilde{X}\,W\big] \qquad \text{per ogni } W = \sum_{k=1}^{n} c_k\,\mathbb{I}_{A_k} \text{ con } A_k \in \mathcal{G},$$

e poi per variabili aleatorie più generali per approssimazione, ossia

$$\mathbb{E}\big[X\,W\big] = \mathbb{E}\big[\widetilde{X}\,W\big] \qquad \text{per ogni variabile aleatoria } W \text{ che sia } \mathcal{G} \text{ misurabile e per cui } \mathbb{E}\big[|X\,W|\big] < \infty.$$

In particolare, se X è di quadrato integrabile, la precedente relazione vale per ogni variabile aleatoria $\mathcal{G}$ misurabile e di quadrato integrabile.

2) *per ogni $t \in I$, la variabile aleatoria X_t è integrabile, ossia $\mathbb{E}\big[|X_t|\big] < \infty$;*
3) *per ogni t_1, $t_2 \in I$, con $t_1 \le t_2$, vale la seguente relazione:*

$$\mathbb{E}\big[X_{t_2}\,\mathbb{I}_A\big] = \mathbb{E}\big[X_{t_1}\,\mathbb{I}_A\big], \qquad \textit{per ogni } A \in \mathcal{F}_{t_1},$$

ossia,

$$\mathbb{E}\big[X_{t_2}|\mathcal{F}_{t_1}\big] = X_{t_1}.$$

Inoltre un processo $(X_t)_{t\in I}$ si dice ***submartingala (supermartingala)*** *rispetto a $(\mathcal{F}_t)_{t\in I}$ se verifica le proprietà* ***1)***, ***2)*** *e la proprietà:*
3') *per ogni t_1, $t_2 \in I$, con $t_1 \le t_2$,*

$$\mathbb{E}\big[X_{t_2}|\mathcal{F}_{t_1}\big] \ge X_{t_1} \qquad \textit{(per le submartingale)},$$

$$\Big(\mathbb{E}\big[X_{t_2}|\mathcal{F}_{t_1}\big] \le X_{t_1} \quad \textit{(per le supermartingale)}\Big).$$

Diamo, senza dimostrazione, alcuni esempi di martingale

Esempio 4.1. *Sia data una variabile aleatoria Y integrabile ed una filtrazione $\{\mathcal{F}_t\}$. Allora il processo definito da*

$$X_t := \mathbb{E}[Y \mid \mathcal{F}_t],$$

è una martingala.

Si dimostra facilmente che X_t è una martingala utilizzando le proprietà del valore atteso condizionale (si vedano le proprietà del valore atteso condizionale in appendice):

$$\mathbb{E}[X_{t+s} \mid \mathcal{F}_t] \underset{(per\ def.)}{=} \mathbb{E}[\mathbb{E}[Y \mid \mathcal{F}_{t+s}] \mid \mathcal{F}_t] \underset{(condiz.\ successivi)}{=} \mathbb{E}[Y \mid \mathcal{F}_t]$$

Esempio 4.2. *Sia X_t una martingala con $\mathbb{E}[\mid X_t \mid^\alpha] < +\infty$, per un $\alpha \ge 1$. Allora il processo $\mid X_t \mid^\alpha$ è una submartingala. Questo esempio si generalizza immediatamente al caso di ogni funzione convessa ϕ, purché, ovviamente, $\mathbb{E}[\phi(X_t)] < +\infty$.*

Per la dimostrazione basta applicare la disuguaglianza di Jensen (si vedano le proprietà del valore atteso condizionale in appendice) alla funzione $|x|^\alpha$, che è convessa per $\alpha \ge 1$

$$\mid X_t \mid^\alpha \underset{(X_t\ \text{è una}\ MG)}{=} \mid \mathbb{E}[X_{t+s} \mid \mathcal{F}_t] \mid^\alpha \underset{(dis.\ Jensen\ per\ |x|^\alpha)}{\le} \mathbb{E}[\mid X_{t+s} \mid^\alpha \mid \mathcal{F}_t].$$

Nel caso generale di una funzione convessa i passaggi sono identici:

$$\phi(X_t) \underset{(X_t\ \text{è una}\ MG)}{=} \phi(\mathbb{E}[X_{t+s} \mid \mathcal{F}_t]) \underset{(dis.\ Jensen\ per\ \phi)}{\le} \mathbb{E}[\phi(X_{t+s}) \mid \mathcal{F}_t].$$

Per ottenere un risultato analogo nel caso in cui X_t sia una submartingala bisogna aggiungere l'ipotesi che ϕ sia una funzione crescente: ossia se X_t è una submartingala e ϕ una funzione convessa e crescente, con $\mathbb{E}[\phi(X_t)] < +\infty$, allora il precesso $\phi(X_t)$ è una submartingala.

La dimostrazione è simile alla procedente: si tratta in più di notare che

$$X_t \underset{(X_t\ \text{è una}\ subMG)}{\le} \mathbb{E}[X_{t+s} \mid \mathcal{F}_t],$$
$$\Downarrow$$
$$\phi(X_t) \underset{(\phi\ \text{è crescente})}{\le} \phi(\mathbb{E}[X_{t+s} \mid \mathcal{F}_t]) \underset{(dis.\ Jensen\ per\ \phi)}{\le} \mathbb{E}[\phi(X_{t+s}) \mid \mathcal{F}_t],$$

e si giunge alla tesi.

Torniamo ai processi ad incrementi indipendenti: Ci si potrebbe chiedere, ma esistono processi a incrementi indipendenti e omogenei, a valore atteso nullo? e che quindi sono martingale?
La risposta è positiva, infatti è facile mostrare che se X_t è un processo ad incrementi indipendenti (e omogenei), integrabile e con $X_0 = 0$, allora, sotto semplici ipotesi di regolarità,

$$\mathbb{E}[X_t] = \mu t \quad \text{per } t \geq 0, \tag{4.2}$$

basterà allora considerare il processo $X_t - \mu\, t$, che è ancora un processo ad incrementi indipendenti e omogenei. Inoltre vedremo come ottenere altre martingale, a partire da un processo ad incrementi indipendenti e omogenei.

Per inziare la dimostrazione di (4.2), è immediato verificare che $\mathbb{E}[X_t] = \mu t$ per t nei razionali:

$$\mathbb{E}[X_1] = \sum_{k=1}^{n} \mathbb{E}[X_{k/n} - X_{(k-1)/n}] = n\mathbb{E}[X_{1/n}]$$

da cui $\mathbb{E}[X_{1/n}] = \frac{1}{n}\mathbb{E}[X_1]$ e analogamente $\mathbb{E}[X_{m/n}] = \sum_{k=1}^{m} \mathbb{E}[X_{k/n} - X_{(k-1)/n}] = \frac{m}{n}\mathbb{E}[X_1]$.
Per ottenere che ciò valga anche per ogni t reale, si deve notare che comunque $\mathbb{E}[X_{t+s}] = \mathbb{E}[X_t] + \mathbb{E}[X_s]$ e aggiungere una piccola ulteriore **ipotesi di regolarità**: la $\mathbb{E}[X_t]$ è una funzione continua in t (o continua a destra). Con questa ipotesi si ottiene immediatamente la tesi per continuità.

Il processo $X_t - \mathbb{E}[X_t] = X_t - \mu t$ è allora un processo ad incrementi indipendenti (ed omogenei), a media nulla e quindi è una martingala.

Se ancora X_t ammette momento secondo finito, allora, con una dimostrazione simile[5] si ha che $Var(X_t) = \sigma^2 t$, purché si possa affermare a priori che $Var(X_t)$ è una funzione continua. Di nuovo similmente al caso a tempo discreto, accade che $(X_t - \mathbb{E}[X_t])^2 - Var(X_t) = (X_t - \mu t)^2 - \sigma^2 t$ è una martingala (a media nulla) .

Infine è possibile mostrare[6] che, sotto opportune ipotesi di regolarità (continuità in probabilità), se $\mathbb{E}[\exp\{\theta\, X_t\}] < +\infty$, allora

$$\mathbb{E}[\exp\{\theta\, X_t\}] = \exp\{K(\theta)t\}, \qquad \text{dove}\, \mathbb{E}[\exp\{\theta\, X_1\}] = \exp\{K(\theta)\}$$

e che quindi

$$Z_t := \exp\{\theta\, X_t - K(\theta)t\} \tag{4.3}$$

è una martingala[7] a media 1.
Tutte le proprietà precedenti valgono anche per i processi $Y_t = Y_0 + X_t$, con dato iniziale Y_0 indipendente da $\{X_t\}$

[5]Si tratta di osservare che la varianza della somma degli incrementi è la somma delle varianze degli incrementi e quindi si procede come nel caso del valore atteso, sostituendo Var a $\mathbb{E}$.

[6]In questo caso si sfrutta il fatto che

$$\exp\{\theta X_1\} = \prod_{k=1}^{n} \exp\{\theta \left(X_{k/n} - X_{(k-1)/n}\right)\},$$

da cui, passando al valore atteso, per l'indipendenza degli incrementi e per l'omogeneità

$$\mathbb{E}[\exp\{\theta X_1\}] = \prod_{k=1}^{n} \mathbb{E}[\exp\{\theta \left(X_{k/n} - X_{(k-1)/n}\right)\}] = \left(\mathbb{E}[\exp\{\theta \left(X_{1/n} - X_0\right)\}]\right)^n .$$

Posto $\exp\{K(\theta)\} := \mathbb{E}[\exp\{\theta X_1\}]$ si ottiene dunque che

$$\mathbb{E}[\exp\{\theta \left(X_{k/n} - X_{(k-1)/n}\right)\}] = \mathbb{E}[\exp\{\theta \left(X_{1/n}\right)\}] = \exp\{K(\theta)(1/n)\},$$

da cui ancora la tesi per ogni t razionale.

[7]Di nuovo misurabilità e integrabilità sono banali. Osservando che

$$\begin{aligned} Z_{t+s} &= \exp\{\theta\, X_{t+s} - K(\theta)(t+s)\} \\ &= \exp\{\theta\, X_t - K(\theta)t\} \exp\{\theta(X_{t+s} - X_t) - K(\theta)s\} \\ &= Z_t \exp\{\theta(X_{t+s} - X_t) - K(\theta)s\} \end{aligned}$$

(tranne per i valori medi). Bisogna però che Y_0 sia $\mathcal{F}_0$-misurabile[8], e soddisfi alcuni requisiti di integrabilità. Ad esempio

$$\widetilde{Z}_t := \mathbb{E}[\exp\{\theta(Y_t - \mu t) - K(\theta)t\}] \tag{4.4}$$

è ancora una martingala a media costante uguale a $\mathbb{E}[\widetilde{Z}_0] = \mathbb{E}[\exp\{\theta Y_0\}]$, purché ovviamente il valore medio di $\exp\{\theta Y_0\}$ sia finito.

Esempio 4.3 (Decomposizione di Doob e martingala esponenziale per il processo di Wiener)**.** *Come applicazione si consideri che il processo di Wiener standard o moto browniano W_t è una martingala, anche $M_t = W_t^2 - t$ e infine, per ogni θ reale*

$$Z_t^\theta := \exp\{\theta W_t - \frac{1}{2}\theta^2 t\} \tag{4.5}$$

è una martingala[9] *a media* 1*, che viene detta* **martingala esponenziale**.

Si noti che W_t^2 è una submartingala (in quanto quadrato di una martingala) e che quindi si può decomporre nella somma di una martingala (la martingala M_t) e di un processo crescente (il processo deterministico t), cioè $W_t^2 = M_t + t$. Si tratta di un caso particolare della decomposizione di Doob a tempo continuo.

Esempio 4.4 (Decomposizione di Doob e martingala esponenziale per il processo di Poisson)**.** *Anche il processo di Poisson N_t di parametro λ, essendo un processo crescente è una submartingala, e si può decomporre nella somma di una martingala $M_t := N_t - \lambda t$ e di un processo crescente $A_t = \lambda t$. Si tratta anche qui di un caso particolare della decomposizione di Doob a tempo continuo.*

Si osservi che in generale data una v.a. Z non negativa e a media 1 in uno spazio $(\Omega, \mathcal{F}, \mathbb{P})$,

$$\mathbb{Q}(C) := \mathbb{E}^{\mathbb{P}}[I_C Z], \qquad C \in \mathcal{F}, \tag{4.6}$$

---

si ottiene subito, che per ogni $A \in \mathcal{F}_t$

$$\begin{aligned}
\mathbb{E}[(Z_{t+s} - Z_t)\mathbb{I}_A] &= \mathbb{E}\left[(\exp\{\theta(X_{t+s} - X_t) - K(\theta)s\} - 1)\, Z_t\, \mathbb{I}_A\right] \\
&\qquad \text{(per la } \mathcal{F}_t\text{-misurabilità di } Z_t\,\mathbb{I}_A\text{) e per l'indipendenza degli incrementi da } \mathcal{F}_t\text{)} \\
&= \mathbb{E}\left[(\exp\{\theta(X_{t+s} - X_t) - K(\theta)s\} - 1)\right] \mathbb{E}\left[Z_t\, \mathbb{I}_A\right] \\
&\qquad \text{(e infine, per l'omogeneità degli incrementi)} \\
&= \mathbb{E}\left[(\exp\{\theta\, X_s - K(\theta)s\} - 1)\right] \mathbb{E}\left[Z_t\, \mathbb{I}_A\right] = (1-1)\, \mathbb{E}\left[Z_t\, \mathbb{I}_A\right] = 0
\end{aligned}$$

In alternativa, con il linguaggio dei valori attesi condizionali, le precedenti relazioni divengono

$$\begin{aligned}
\mathbb{E}[Z_{t+s} - Z_t|\mathcal{F}_t] &= \mathbb{E}\left[Z_t\,(\exp\{\theta(X_{t+s} - X_t) - K(\theta)s\} - 1)\,|\mathcal{F}_t\right] \\
&\qquad \text{(per la misurabilità di } Z_t\text{)} \\
&= Z_t\mathbb{E}\left[(\exp\{\theta(X_{t+s} - X_t) - K(\theta)s\} - 1)\,|\mathcal{F}_t\right] \\
&\qquad \text{(per l'indipendenza degli incrementi da } \mathcal{F}_t\text{)} \\
&= Z_t\mathbb{E}\left[(\exp\{\theta(X_{t+s} - X_t) - K(\theta)s\} - 1)\right] \\
&\qquad \text{(per l'omogeneità degli incrementi)} \\
&= Z_t\mathbb{E}\left[(\exp\{\theta\, X_s - K(\theta)s\} - 1)\right] = Z_t\,(1-1) = 0
\end{aligned}$$

[8]† In realtà la richiesta che Y_0 sia $\mathcal{F}_0$-misurabile non è strettamente necessaria, se vale la condizione di indipendenza di tra il processo (X_t) e la variabile aleatoria Y_0: nel caso in cui $\mathcal{F}_t = \mathcal{F}_t^X$ si potrebbe, in alternativa, cambiare la filtrazione e prendere la filtrazione definita da $\mathcal{F}_t \vee \sigma\{Y_0\}$. In questo caso infatti il processo $Y_0 + X_t$ viene automaticamente adattato alla nuova filtrazione. Inoltre $\{X_t\}$ è ancora una martingala rispetto alla nuova filtrazione, come si può vedere facilmente usando la proprietà dei condizionamenti ridondanti: infatti la σ-algebra $\mathcal{H} = \sigma\{Y_0\}$ è indipendente da $X = X_{t+s}$ e da $\mathcal{G} = \mathcal{F}_t$.

$$\mathbb{E}[X|\mathcal{G} \vee \mathcal{H}] = \mathbb{E}[X|\mathcal{G}] \quad \Longleftrightarrow \quad \mathbb{E}[X_{t+s}|\mathcal{F}_t \vee \sigma\{Y_0\}] = \mathbb{E}[X_{t+s}|\mathcal{F}_t] = X_t;$$

lo stesso discorso vale nel caso in cui si assuma Y_0 indipendente da $\mathcal{F}_t \supseteq F_t^X$ per ogni t (e quindi anche dal processo $\{X_t\}$.

[9]Basta ricordare che $K(\theta)$ è definito dal fatto che $\exp\{K(\theta)\} = \mathbb{E}[\exp\{\theta(X_1 - \mu 1)\}]$. In questo caso quindi $\exp\{K(\theta)\} = \mathbb{E}[\exp\{\theta W_1\}] = \exp\{\frac{1}{2}\theta^2\}$, da cui $K(\theta) = \frac{1}{2}\theta^2$.

definisce una nuova misura di probabilità[10] $\mathbb{Q}$ su $(\Omega, \mathcal{F})$.

Esercizio 4.1 (Un caso particolare del Teorema di Girsanov)**.** *Posto*

$$Z = Z_T^\theta \Big(= \exp\{\theta W_T - \frac{1}{2}\theta^2 T\}\Big)$$

dove Z_t^θ è la martingala esponenziale (4.5) relativa al processo di Wiener, e $\mathcal{F} = \mathcal{F}_T$ nell'espressione precedente (4.6), si trovi

1) la derivata di Radon-Nikodym

$$h_t = \frac{d\mathbb{Q}}{d\mathbb{P}}\Big|_{\mathcal{F}_t},$$

ovvero quella variabile aleatoria** $h_t(\omega)$**, $\mathcal{F}_t$-misurabile, tale che

$$\mathbb{Q}(A) = \int_A h_t(\omega)\mathbb{P}(d\omega) = \mathbb{E}^{\mathbb{P}}\big[h_t\, \mathbb{I}_A\big], \qquad \textit{\textbf{per ogni}}\ A \in \mathcal{F}_t;$$

soluzione: $h_t = \frac{d\mathbb{Q}}{d\mathbb{P}}\big|_{\mathcal{F}_t} = Z_t^\theta$.

suggerimento: *qualunque sia $A \in \mathcal{F}_T$ si ha che $\mathbb{Q}(A) = \mathbb{E}^{\mathbb{P}}\big[\mathbb{I}_A Z_T^\theta\big]$ e quindi anche per ogni $A \in \mathcal{F}_t$ $(t \leq T)$ si ha*

$$\begin{aligned}\mathbb{Q}(A) &= \mathbb{E}^{\mathbb{P}}\big[\mathbb{I}_A Z_T^\theta\big] = \mathbb{E}^{\mathbb{P}}\big[\mathbb{I}_A\, Z_t^\theta\, \frac{Z_T^\theta}{Z_t^\theta}\big] = \mathbb{E}^{\mathbb{P}}\big[\mathbb{I}_A\, Z_t^\theta\, \exp\{\theta W_T - \tfrac{1}{2}\theta^2 T\}\, \exp\{-\theta W_t + \tfrac{1}{2}\theta^2 t\}\big] \\ &= \mathbb{E}^{\mathbb{P}}\big[\mathbb{I}_A\, Z_t^\theta\, \exp\{\theta(W_T - W_t) - \tfrac{1}{2}\theta^2(T-t)\}\big] = \mathbb{E}^{\mathbb{P}}\big[\mathbb{I}_A\, Z_t^\theta\big]\, \mathbb{E}^{\mathbb{P}}\big[\exp\{\theta(W_T - W_t) - \tfrac{1}{2}\theta^2(T-t)\}\big]\end{aligned}$$

dove l'ultima uguaglianza vale in quanto, se $A \in \mathcal{F}_t$ allora $\mathbb{I}_A\, Z_t^\theta \in \mathcal{F}_t$ e quindi è indipendente dall'incremento $W_T - W_t$. Di conseguenza, tenendo conto che abbiamo che

$$\mathbb{Q}(A) = \mathbb{E}^{\mathbb{P}}\big[\mathbb{I}_A\, Z_T^\theta\big] = \mathbb{E}^{\mathbb{P}}\big[\mathbb{I}_A\, Z_t^\theta\big] \qquad \textit{per ogni } A \in \mathcal{F}_t.$$

Osservazione, abbiamo quindi ritrovato che il processo $(Z_t^\theta)_{t\in[0,T]}$ è una martingala rispetto ala filtrazione $(\mathcal{F}_t)_{t\in[0,T]}$

2) la legge di** W_t **rispetto a** $\mathbb{Q}$**; ***soluzione:*** *la legge di W_t è $N(\theta t, t)$*

suggerimento: *basta capire che $\mathbb{E}^{\mathbb{Q}}[g(W_t)]$ si calcola equivalentemente come*

$$\begin{aligned}&\mathbb{E}^{\mathbb{Q}}[g(W_t)] = \mathbb{E}^{\mathbb{P}}[g(W_t) Z_t^\theta] = \mathbb{E}^{\mathbb{P}}[g(W_t)\exp\{\theta W_t - \frac{1}{2}\theta^2 t\}] \\ &= \int_{\mathbb{R}} g(w)\exp\{\theta w - \frac{1}{2}\theta^2 t\}\exp\{-\frac{1}{2t}w^2\}dw = \int_{\mathbb{R}} g(w)\exp\{-\frac{1}{2t}(w^2 - 2w\theta t + \theta^2 t^2)\}dw \\ &= \int_{\mathbb{R}} g(w)\exp\{-\frac{1}{2t}(w - \theta t)^2\}dw\end{aligned}$$

3) le distribuzioni finito dimensionali di** $(W_t, t \geq 0)$ **rispetto a** $\mathbb{Q}$**.

soluzione: *il processo $\{W_t\}$ diviene sotto $\mathbb{Q}$ un processo di Wiener con drift θ e coefficiente di diffusione* 1.

[10] È ovvio che $\mathbb{Q}(C) \geq 0$, essendo Z non negativa, e che $\mathbb{Q}(\Omega) = 1$, in quanto $\mathbb{Q}(\Omega) = \mathbb{E}^{\mathbb{P}}[I_\Omega Z] = \mathbb{E}^{\mathbb{P}} Z] = 1$. La σ-additività segue dalle proprietà di σ-additività di $\mathbb{P}$ e dal fatto che

$$I_{\{\cup_n A_n\}} = \sum_n I_{A_n}, \qquad \text{se gli insiemi } A_n \text{ sono disgiunti a due a due.}$$

suggerimento: *si tratta di capire che* $\mathbb{E}^{\mathbb{Q}}[g_1(W_{t_1})g_2(W_{t_2} - W_{t_1}) \cdots g_n(W_{t_n} - W_{t_n-1})]$ *si calcola come*

$$
\begin{aligned}
&\mathbb{E}^{\mathbb{P}}[g_1(W_{t_1})g_2(W_{t_2} - W_{t_1}) \cdots g_n(W_{t_n} - W_{t_n-1})Z^{\theta}_{t_n}] \\
&= \mathbb{E}^{\mathbb{P}}[g_1(W_{t_1})g_2(W_{t_2} - W_{t_1}) \cdots g_n(W_{t_n} - W_{t_n-1}) \prod_{k=1}^{n} \frac{Z^{\theta}_{t_k}}{Z^{\theta}_{t_{k-1}}}] \\
&= \mathbb{E}^{\mathbb{P}}[g_1(W_{t_1})g_2(W_{t_2} - W_{t_1}) \cdots g_n(W_{t_n} - W_{t_{n-1}}) \prod_{k=1}^{n} \exp\{\theta(W_{t_k} - W_{t_{k-1}}) - \frac{1}{2}\theta^2(t_k - t_{k-1})\}] \\
&= \mathbb{E}^{\mathbb{P}}[\prod_{k=1}^{n} g_k(W_{t_k} - W_{t_{k-1}}) \exp\{\theta(W_{t_k} - W_{t_{k-1}}) - \frac{1}{2}\theta^2(t_k - t_{k-1})\}] \\
&= \prod_{k=1}^{n} \mathbb{E}^{\mathbb{P}}[g_k(W_{t_k} - W_{t_{k-1}}) \exp\{\theta(W_{t_k} - W_{t_{k-1}}) - \frac{1}{2}\theta^2(t_k - t_{k-1})\}] \\
&= \prod_{i=1}^{n} \int_{\mathbb{R}} g(w_i - w_{i-1}) \exp\{-\frac{(w_i - w_{i-1} - \theta(t_i - t_{i-1}))^2}{2(t_i - t_{i-1})}\} dw_i.
\end{aligned}
$$

dove per convenzione si è posto $t_0 = 0$ *e* $w_0 = 0$*.*

4.3 APPENDICE: Proprietà del valore atteso condizionale e qualche esempio

Enunciamo ora (senza dimostrarle), le proprietà fondamentali della media condizionale.

Siano X e Y variabili aleatorie integrabili in $(\Omega, \mathcal{F}, \mathbb{P})$ e siano $\mathcal{G}$ e $\mathcal{H}$ sotto σ-algebre di $\mathcal{F}$, allora valgono le seguenti proprietà:

1. Linearità

$$\mathbb{E}[aX + bY \mid \mathcal{G}] = a\mathbb{E}[X \mid \mathcal{G}] + b\mathbb{E}[Y \mid \mathcal{G}]$$

2. Monotonia

$$\mathbb{P}(X \leq Y) = 1 \text{ implica } \mathbb{P}(\mathbb{E}[X \mid \mathcal{G}] \leq \mathbb{E}[Y \mid \mathcal{G}]\ ) = 1$$

3. Formula dei condizionamenti successivi (caso particolare $\mathcal{G} = \{\Omega, \emptyset\}$)

$$\text{se } \mathcal{G} \subseteq \mathcal{H}, \text{ allora } \mathbb{E}[X \mid \mathcal{G}] = \mathbb{E}[\mathbb{E}[X \mid \mathcal{H}] \mid \mathcal{G}]$$

quindi, in particolare,

$$\text{se } \mathcal{G} = \{\Omega, \emptyset\}, \text{ allora } \mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X \mid \mathcal{H}]]$$

4. Fattorizzazione

Se Z è $\mathcal{G}$-misurabile e ZX è integrabile allora

$$\mathbb{E}[ZX \mid \mathcal{G}] = Z\mathbb{E}[X \mid \mathcal{G}]$$

5. Condizionamento rispetto a σ-algebre indipendenti

Se X e $\mathcal{G}$ sono indipendenti allora

$$\mathbb{E}[X \mid \mathcal{G}] = \mathbb{E}[X]$$

6. Condizionamento ridondante, cioè rispetto ad allargamenti indipendenti di σ-algebre

Se X e $\mathcal{G}$ sono indipendenti da $\mathcal{H}$, nel senso che $\sigma(X) \vee \mathcal{G}$ è indipendente da $\mathcal{H}$, allora

$$\mathbb{E}[X \mid \mathcal{G} \vee \mathcal{H}] = \mathbb{E}[X \mid \mathcal{G}]$$

7. Disuguaglianza di Jensen per funzioni convesse (caso particolare $\phi(x) = x^2$)

Se ϕ è una funzione convessa, e $\phi(X)$ è integrabile, allora

$$\phi(\mathbb{E}[X \mid \mathcal{G}]) \leq \mathbb{E}[\phi(X) \mid \mathcal{G}]$$

Osservazione In particolare per $\phi(x) = x^2$ ed X in $L^2(\Omega, \mathcal{F}, \mathbb{P})$, si ottiene che

$$(\mathbb{E}[X \mid \mathcal{G}])^2 \leq \mathbb{E}[X^2 \mid \mathcal{G}],$$

e quindi, passando al valore atteso, che

$$\mathbb{E}[(\mathbb{E}[X \mid \mathcal{G}])^2] \leq \mathbb{E}[\mathbb{E}[X^2 \mid \mathcal{G}]] = \mathbb{E}[X^2].$$

8. Convergenza sotto il segno di media condizionale, monotona e dominata

Se $\{X_n\}_{n\geq 1}$ è una successione di variabili aleatorie non negative ed integrabili, convergente con probabilità 1 ad X, monotonamente, cioè $0 \leq X_n \leq X_{n+1}$, allora

$$\mathbb{E}[X_n \mid \mathcal{G}] \uparrow \mathbb{E}[X \mid \mathcal{G}], \text{ con probabilità 1},$$

Se invece la successione converge ad X dominatamente, cioè $\mid X_n \mid \leq Y$, allora

$$\mathbb{E}[X_n \mid \mathcal{G}] \to \mathbb{E}[X \mid \mathcal{G}], \text{ in } L^1.$$

Terminiamo con un paio di esempi di calcolo del valore atteso condizionale

Esempio 4.5. *Consideriamo il caso in cui $\mathcal{G} = \sigma(Y)$, e Y è una variabile aleatoria discreta, a valori in $\{y_m;\, m \in \mathbb{N}\}$.*

Se $\mathbb{P}(Y = y_m) > 0$ per ogni $m \in \mathbb{N}$, si ha

$$\mathbb{E}[X \mid \sigma(Y)](\omega) = \sum_{m\in\mathbb{N}} \mathbb{E}[X \mid \{Y = y_m\}]\mathbb{I}_{\{Y=y_m\}}(\omega) = \sum_{m\in\mathbb{N}} \mathbb{E}[X \mid \{Y = y_m\}]\mathbb{I}_{\{y_m\}}(Y(\omega)),$$

con il solito abuso di notazione (il secondo membro è un rappresentante della classe di equivalenza $\mathbb{E}[X \mid \sigma(Y)]$).

Quindi posto

$$\psi(y) = \sum_{m\in\mathbb{N}} \mathbb{E}[X \mid \{Y = y_m\}]\mathbb{I}_{\{y_m\}}(y),$$

e indicato con

$$\mathbb{E}[X \mid \{Y = y\}] = \psi(y),$$

cioè la funzione, che vale $\mathbb{E}[X \mid \{Y = y\}]$*, se* $y \in \{y_m;\, m \in \mathbb{N}\}$*, e zero altrimenti, si ha:*

$$\mathbb{E}[X \mid \sigma(Y)](\omega) = \quad \mathbb{E}[X \mid \{Y = y\}]\Big|_{y=Y(\omega)} = \psi(Y(\omega)).$$

Ciò giustifica il fatto che si usa scrivere

$$\mathbb{E}[X \mid \sigma(Y)] = \mathbb{E}[X \mid Y],$$

e ci fa ritrovare il concetto elementare di valore atteso condizionato di una variabile aleatoria discreta X *rispetto a una variabile aleatoria discreta* Y*.*

In tale caso, cioè quando anche X *è una variabile aleatoria discreta a valori in* $\{x_n;\, n \in \mathbb{N}\}$*, allora*

$$\mathbb{E}[X \mid Y](\omega) = \sum_{n\in\mathbb{N}} x_n \mathbb{P}(X = x_n \mid \{Y = y\})\Big|_{y=Y(\omega)}$$

Infine notiamo che è facile ripetere quanto sopra nel caso in cui al posto di X *ci sia una variabile aleatoria* $Z = h(X)$*, integrabile, e ottenere che*

$$\mathbb{E}[h(X) \mid Y](\omega) = \sum_{n\in\mathbb{N}} h(x_n) \mathbb{P}(X = x_n \mid \{Y = y\})\Big|_{y=Y(\omega)}$$

Nel caso in cui ***non*** *si abbia che* $\mathbb{P}(Y = y_m) > 0$ *per ogni* $m \in \mathbb{N}$*, si può ottenere un rappresentante del valore atteso condizionale come segue: fissata una probabilità* $\mathbb{P}^0$*, e posto* $\sum_m^*$ *la somma sugli indici* m *tali che* $\mathbb{P}(Y = y_m) > 0$ *e* $\sum_m^{**}$ *la somma sui rimanenti indici (ossia sulgi* m *tali che* $\mathbb{P}(Y = y_m) = 0$*)*

$$\begin{aligned}\mathbb{E}[X \mid \sigma(Y)](\omega) &= \sum_{m\in\mathbb{N}}{}^{*} \mathbb{E}[X \mid \{Y = y_m\}]\mathbb{I}_{\{Y=y_m\}}(\omega) + \sum_{m\in\mathbb{N}}{}^{**} \mathbb{E}^0[X]\mathbb{I}_{\{Y=y_m\}}(\omega)\\ &= \sum_{m\in\mathbb{N}}{}^{*} \mathbb{E}[X \mid \{Y = y_m\}]\mathbb{I}_{\{y_m\}}(Y(\omega)) + \sum_{m\in\mathbb{N}}{}^{**} \mathbb{E}^0[X]\mathbb{I}_{\{y_m\}}(Y(\omega)),\end{aligned}$$

con un abuso di notazione usuale (il secondo membro è un rappresentante della classe di equivalenza $\mathbb{E}[X \mid \sigma(Y)]$*).*

Quindi posto

$$\psi(y) = \sum_{m\in\mathbb{N}}{}^{*} \mathbb{E}[X \mid \{Y = y_m\}]\mathbb{I}_{\{y_m\}}(y) + \sum_{m\in\mathbb{N}}{}^{**} \mathbb{E}^0[X]\mathbb{I}_{\{y_m\}}(y),$$

e indicato con

$$\mathbb{E}[X \mid \{Y = y\}] = \psi(y),$$

cioè la funzione, che vale $\mathbb{E}[X \mid \{Y = y\}]$*, se* $y \in \{y_m;\, m \in \mathbb{N}\}$ *e se* $\mathbb{P}(\{Y = y\}) > 0$*, e vale* $\mathbb{E}^0[X]$ *altrimenti, si ha:*

$$\mathbb{E}[X \mid Y](\omega) = \quad \mathbb{E}[X \mid \{Y = y\}]\Big|_{y=Y(\omega)} = \psi(Y(\omega)).$$

Più in generale si ha anche che posto

$$\psi_h(y) = \sum_{m\in\mathbb{N}}{}^{*} \mathbb{E}[h(X) \mid \{Y = y_m\}]\mathbb{I}_{\{y_m\}}(y) + \sum_{m\in\mathbb{N}}{}^{**} \mathbb{E}^0[h(X)]\mathbb{I}_{\{y_m\}}(y),$$

si ottiene che

$$\mathbb{E}[h(X) \mid Y](\omega) = \psi_h(Y(\omega)).$$

Esempio 4.6. *Caso in cui* $\mathcal{G} = \sigma(Y)$*, con* Y *una variabile aleatoria a valori in* $\mathbb{R}^d$*, e* (X, Y) *ammette densità di probabilità congiunta* $f_{X,Y}(x, y)$*. Allora, posto*

$$f_{X|Y}(x|y) = I_{\{z: f_Y(z)>0\}}(y)\frac{f_{X,Y}(x, y)}{f_Y(y)} + I_{\{z: f_Y(z)=0\}}(y) f_0(x)$$

dove $f_0(x)$ *è una qualunque densità di probabilità prefissata, si ha*

$$\widetilde{X}(\omega) = \mathbb{E}[X \mid Y](\omega) = \int_{\mathbb{R}} x\, f_{X|Y}(x|y)\Big|_{y=Y(\omega)}\, dx$$

ossia[11]

$$\widetilde{X}(\omega) = \int_{\mathbb{R}} x \left(I_{\{z: f_Y(z)>0\}}(y) \frac{f_{X,Y}(x,y)}{f_Y(y)} \right)\Big|_{y=Y(\omega)} dx + \int_{\mathbb{R}} x \; I_{\{z: f_Y(z)=0\}}(y)\big|_{y=Y(\omega)} \; f_0(x) dx,$$

dove $\mathbb{E}[X \mid Y]$ *è una abbreviazione per* $\mathbb{E}[X \mid \sigma(Y)]$.

Per la verifica è intanto importante notare che $\sigma(Y) = \{A = Y^{-1}(B), \text{ per } B \in \mathcal{B}(\mathbb{R})\}$, *quindi* $I_A(\omega) = I_B(Y(\omega))$ *e*

$$\mathbb{E}[XI_A] = \mathbb{E}[XI_B(Y)] = \int_{\mathbb{R}\times\mathbb{R}^d} xI_B(y) f_{X,Y}(x,y) \; dx \; dy.$$

Diamo ora la dimostrazione nel caso in cui $f_{X,Y}(x,y) > 0$ *per ogni* $(x,y) \in \mathbb{R}\times\mathbb{R}^d$, *così anche* $f_Y(y) > 0$ *per ogni* $y \in \mathbb{R}^d$, *e quindi*

$$\begin{aligned} \mathbb{E}[\widetilde{X}(\omega) I_A] &= \mathbb{E}[\widetilde{X}(\omega) I_B(Y)] \\ &= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}} x \frac{f_{X,Y}(x,y)}{f_Y(y)} \; dx \right) I_B(y) f_Y(y) \; dy \end{aligned}$$

ovvero, per il Teorema di Fubini[12],

$$\begin{aligned} \mathbb{E}[\widetilde{X}(\omega) I_A] &= \int_{\mathbb{R}} \int_{\mathbb{R}^d} xI_B(y) \frac{f_{X,Y}(x,y)}{f_Y(y)} f_Y(y) \; dx \; dy \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}^d} xI_B(y) f_{X,Y}(x,y) \; dx \; dy. \end{aligned}$$

Nel caso generale

$$\begin{aligned} \mathbb{E}[\widetilde{X}(\omega) I_A] &= \mathbb{E}[\widetilde{X}(\omega) I_B(Y)] \\ &= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}} x I_{\{z: f_Y(z)>0\}}(y) \frac{f_{X,Y}(x,y)}{f_Y(y)} \; dx \right) I_B(y) f_Y(y) \; dy \\ &\quad + \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}} x I_{\{z: f_Y(z)=0\}}(y) f_0(x) \; dx \right) I_B(y) f_Y(y) \; dy \end{aligned}$$

ovvero, per il Teorema di Fubini,

$$\begin{aligned} &= \int_{\mathbb{R}} \int_{\mathbb{R}^d} xI_B(y) I_{\{z: f_Y(z)>0\}}(y) \frac{f_{X,Y}(x,y)}{f_Y(y)} f_Y(y) \; dx \; dy \\ &\quad + \int_{\mathbb{R}} \int_{\mathbb{R}^d} xI_B(y) I_{\{z: f_Y(z)=0\}}(y) f_0(x) f_Y(y) \; dx \; dy \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}^d} xI_B(y) I_{\{z: f_Y(z)>0\}}(y) \frac{f_{X,Y}(x,y)}{f_Y(y)} f_Y(y) \; dx \; dy \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}^d} xI_B(y) I_{\{z: f_Y(z)>0\}}(y) f_{X,Y}(x,y) \; dx \; dy \end{aligned}$$

Si tratta quindi solo di controllare che, qualunque sia $B \in \mathcal{B}(\mathbb{R}^d)$

$$\int_{\mathbb{R}} \int_{\mathbb{R}^d} xI_B(y) I_{\{z: f_Y(z)>0\}}(y) f_{X,Y}(x,y) \; dx \; dy = \int_{\mathbb{R}} \int_{\mathbb{R}^d} xI_B(y) f_{X,Y}(x,y) \; dx \; dy.$$

[11]‡ In realtà per poter scrivere la formula esplicita per $\widetilde{X}(\omega)$ è necessario prendere $f_0(x)$ in modo che $\int_{\mathbb{R}} |x| \; f_0(x) dx < \infty$. Inoltre un altro rappresentante per il valore atteso condizionato è $\int_{\mathbb{R}} x \left(I_{\{z: f_Y(z)>0\}}(y) \frac{f_{X,Y}(x,y)}{f_Y(y)} \right)\Big|_{y=Y(\omega)} dx$ in quanto $\mathbb{P}\big(\; I_{\{z: f_Y(z)=0\}}(y)\big|_{y=Y(\omega)} = 0\big) = 1$. Quest'ultima uguaglianza dipende dal fatto che $I_{\{z: f_Y(z)=0\}}(y)\big|_{y=Y(\omega)} = 0 \Leftrightarrow Y(\omega) \notin \{z : f_Y(z) = 0$ e per il suo complementare si ha $\mathbb{P}\big(Y(\omega) \in \{z : f_Y(z) = 0\}\big) = \int_{\{z: f_Y(z)=0\}} f_y(z) \, dz = \int_{\{z: f_Y(z)=0\}} 0 \, dz = 0$.

[12]Una versione del Teorema di Fubini è la seguente: se $\psi : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \mapsto \mathbb{R}$; $(x,y) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \mapsto \psi(x,y) \in \mathbb{R}$ è una funzione boreliana, allora le seguenti condizioni sono equivalenti

$$\int_{R^{n_1}} \left(\int_{R^{n_2}} |\psi(x,y)| dy \right) dx < \infty \;\Leftrightarrow\; \int_{R^{n_2}} \left(\int_{R^{n_1}} |\psi(x,y)| dx \right) dy < \infty \;\Leftrightarrow\; \int_{R^{n_1} \times {}^{n_2}_{R}} |\psi(x,y)| dx dy < \infty.$$

Inoltre se vale una delle precedenti condizioni vale, allora tutti i valori dei precedenti integrali coincidono.
Il Teorema di Fubini è quindi usato per scambiare l'ordine degli integrali.

La verifica è immediata in quanto

$$\int_{\mathbb{R}}\int_{\mathbb{R}^d} x I_B(y) I_{\{z: f_Y(z)=0\}}(y) f_{X,Y}(x,y)\ dx\ dy = 0,$$

infatti, se $I_{\{z: f_Y(z)=0\}}(y) = 1$*, ovvero se* $f_Y(y) = 0\, \big(= \int_{\mathbb{R}} f_{X,Y}(x,y)\ dx\big)$*, allora l'insieme* $\{x : f_{X,Y}(x,y) > 0\}$ *ha misura di Lebesgue nulla, e quindi, per tali* y

$$\int_{\mathbb{R}^d} x I_B(y) f_{X,Y}(x,y)\ dx = 0.$$

Si osservi che anche in questo caso, se $f_{X,Y}(x,y) > 0$ per ogni (x,y), allora $\frac{f_{X,Y}(x,y)}{f_Y(y)}$ è una densità di probabilità in x, qualunque sia y, e che

$$\mathbb{P}(X \in C \mid Y) = \mathbb{E}[\mathbb{I}_C(X) \mid \sigma(Y)] = \int_C \frac{f_{X,Y}(x,y)}{f_Y(y)} dx$$

definisce una probabilità sui boreliani di $\mathbb{R}$.

Capitolo 5

Integrale stocastico: cenni

Sia $W = (W_t)_t$ un moto browniano continuo standard. L'obiettivo è di dare un significato ad espressioni del tipo

$$\int_0^T X_s(\omega) dW_s(\omega) \tag{5.1}$$

dove l'integrando $(X_s)_{0 \le s \le T}$ è un processo che gode di proprietà da precisare. Ricordiamo che se una funzione g è a variazione finita, allora è possibile definire l'integrale

$$\int_0^T h(t) dg(t)$$

per ogni funzione h misurabile e limitata[1].

Come abbiamo visto precedentemente nella sezione 3.2, le traiettorie del moto browniano non sono a variazione finita, quindi l'operazione (5.1) non si può definire traiettoria per traiettoria.
La (5.1) si chiamerà *integrale stocastico*. Questo tipo di calcolo è fondamentale per la costruzione e lo studio di nuovi processi.

In questa sezione $(\mathcal{F}_t)$ denota una filtrazione rispetto alla quale il processo di Wiener è una martingala. In particolare si può considerare $\mathcal{F}_t = \mathcal{F}_t^W$.

5.1 Integrale stocastico per processi integrabili, a partire dai processi elementari

Consideriamo le due seguenti classi di processi

Definizione 5.1 (Processi elementari predicibili)**.** *I processi del tipo*

$$X_t(\omega) = \sum_{i=0}^{n-1} C_i(\omega) I_{(t_i, t_{i+1}]}(t), \tag{5.2}$$

[1]† Se g è a variazione finita continua a destra, allora si può scrivere $g = g^+ - g^-$ con g^+ ed g^- funzioni crescenti in senso lato, non negative continue a destra e limitate. Allora l'integrale di Lebesgue-Stieltjes è definito da

$$\int_0^T h(t) dg(t) := \int_0^T h(t) \mu_g(dt),$$

dove

$$\mu_g = \mu_{g^+} - \mu_{g^-},$$

e $\mu_{g^\pm}$ sono le misure su $[0, T]$ univocamente definite da

$$\mu_{g^\pm}((a, b]) = g^\pm(b) - g^\pm(a).$$

Nel caso in cui g sia continua, e anche h lo sia allora, l'integrale di Lebesgue-Stieltjes coincide con l'integrale di Riemann-Stieltjes, ovvero con

$$\lim_{n\to\infty} \sum_k h(t_k^*)[g(t_k) - g(t_{k-1})],$$

dove $\{t_k\}_k$ è una partizione di $[0, T]$ e $t_k^* \in [t_{k-1}, t_k]$.

dove $\alpha = t_0 < t_1 < ... < t_n = \beta$, e le C_i sono variabili aleatorie reali $\mathcal{F}_{t_i}$-misurabili vengono detti **processi elementari** **$(\mathcal{F}_t)$-predicibili**, *o, più semplicemente,* **processi elementari predicibili** *(se non ci sono dubbi sulla filtrazione).*

**Di solito $\mathcal{F}_t = \mathcal{F}_t^W = \sigma\{W(u),\, u \leq t\}$. Esempi di variabili aleatorie $\mathcal{F}_t^W$-misurabili sono

$$C_n(\omega) = \psi_n(W_{u_1}, W_{u_2}, \cdots, W_{u_n}) \quad \text{con } u_1 \leq u_2 \leq \cdots \leq u_n \leq t.$$

Si noti che non si lede in generalità a supporre che

$$C_n(\omega) = \overline{\psi}_n(W_{u_1}, W_{u_2} - W_{u_1}, \cdots, W_{u_n} - W_{u_{n-1}}) \quad \text{con } u_1 \leq u_2 \leq \cdots \leq u_n \leq t. \tag{5.3}$$

Altri esempi si possono costruire come limite di variabili aleatorie $C_n(\omega)$ del tipo precedente: ad esempio

$$\sup_{u\in[0,t]} |W_u| = \lim_{n\to\infty} \max |W_{t_1^n}|, |W_{t_2^n}|, \cdots, |W_{t_{2^n}^n}| \quad \text{dove } t_k^n = \frac{k}{2^n}\, t;$$

oppure, per una funzione continua g,

$$\int_0^t g(W_u)du = \lim_{n\to\infty} \sum_{k\leq nt} g(W_{\frac{k}{n}})\, \frac{1}{n}.$$

**

Nel seguito la ***famiglia dei processi elementari predicibili*** viene denotata con $\mathcal{E}([\alpha, \beta])$.

Definizione 5.2. *I processi del tipo*

$$X_t(\omega) = \sum_{i=0}^{n-1} C_i(\omega) I_{[t_i,t_{i+1})}(t), \tag{5.4}$$

dove $\alpha = t_0 < t_1 < ... < t_n = \beta$, e le C_i sono variabili aleatorie reali $\mathcal{F}_{t_i}$-misurabili vengono chiamati **processi elementari $(\mathcal{F}_t)$-opzionali** *o, più semplicemente,* **processi elementari opzionali** *(se non ci sono dubbi sulla filtrazione).*

Nel seguito la ***famiglia dei processi elementari opzionali*** viene denotata con $\mathcal{E}_{op}([\alpha, \beta])$.

L'integrale stocastico rispetto al processo di Wiener per tali processi viene definito come

$$\mathcal{I}^W_{\alpha,\beta}(X) = \int_{(\alpha,\beta]} X_s\, dW_s = \sum_{i=1}^{n} C_{i-1}(\omega)\big(W_{t_i} - W_{t_{i-1}}\big) \left(= \sum_{i=0}^{n-1} C_i(\omega)\big(W_{t_{i+1}} - W_{t_i}\big)\right).$$

Si noti che si tratta di variabili aleatorie a media nulla: infatti le variabili aleatorie C_i e $\big(W_{t_{i+1}} - W_{t_i}\big)$ sono indipendenti (si consideri la forma (5.3) per C_i) e perciò

$$\mathbb{E}\Big[\mathcal{I}^W_{\alpha,\beta}(X)\Big] = \sum_{i=0}^{n-1} \mathbb{E}\big[C_i(\omega)\big(W_{t_{i+1}} - W_{t_i}\big)\big] = \sum_{i=0}^{n-1} \mathbb{E}\big[C_i(\omega)\big]\mathbb{E}\big[\big(W_{t_{i+1}} - W_{t_i}\big)\big] = 0$$

e con varianza uguale a

$$\begin{aligned}\mathbb{E}\Big[\big(\mathcal{I}^W_{\alpha,\beta}(X)\big)^2\Big] &= \sum_{i=0}^{n-1} \mathbb{E}\big[C_i^2(\omega)\big(W_{t_{i+1}} - W_{t_i}\big)^2\big] + \sum_{\substack{i=0\\ i\neq j}}^{n-1}\sum_{j=0}^{n-1} \mathbb{E}\big[C_i(\omega)\,\big(W_{t_{i+1}} - W_{t_i}\big)\, C_j(\omega)\,\big(W_{t_{j+1}} - W_{t_j}\big)\big]\\ &= \sum_{i=0}^{n-1} \mathbb{E}\big[C_i^2(\omega)\big]\big(t_{i+1} - t_i\big)\end{aligned}$$

in quanto, le variabili aleatorie C_i^2 e $\big(W_{t_{i+1}} - W_{t_i}\big)^2$ sono indipendenti (si consideri la forma (5.3) per C_i), e inoltre, prendendo ad esempio $i < j$, si ha che $C_i(\omega)\,\big(W_{t_{i+1}} - W_{t_i}\big) C_j(\omega)$ sono $\mathcal{F}^W_{t_j}$ misurabili (dipendono dagli incrementi di $W_{t_k} - W_{t_{k-1}}$, per $k \leq j$) e sono quindi indipendenti da $\big(W_{t_{j+1}} - W_{t_j}\big)\big]$, e perciò

$$\mathbb{E}\big[C_i(\omega)\,\big(W_{t_{i+1}} - W_{t_i}\big)\, C_j(\omega)\,\big(W_{t_{j+1}} - W_{t_j}\big)\big] = \mathbb{E}\big[C_i(\omega)\,\big(W_{t_{i+1}} - W_{t_i}\big)\, C_j(\omega)\big]\, \mathbb{E}\big[\big(W_{t_{j+1}} - W_{t_j}\big)\big] = 0$$

È importante individuare almeno una classe ampia di processi per i quali ha senso calcolare l'integrale stocastico: se X è un processo adattato e X ha traiettorie continue, ed è un processo per il quale vale

$$\mathbb{E}\left[\int_\alpha^\beta |X_s|^2 ds\right] < \infty. \tag{5.5}$$

è possibile mostrare che, se $\{t_i^n\}_{i=0,\cdots,m_n}$ è una partizione di $[\alpha,\beta]$ con $\max_i |t_i^n - t_{i-1}^n| \to 0$, allora la successione di processi $\{X^n\}$ definiti da

$$X_t^n(\omega) = \sum_{i=1}^{m_n} X_{t_{i-1}^n}(\omega) I_{(t_{i-1}^n, t_i^n]}(t),$$

è una successione di processi elementari e predicibili in $\mathbb{L}^2\big([\alpha,\beta]\times\Omega\big)$ che converge ad X, nel senso che

$$\lim_{n\to\infty} \mathbb{E}\left[\int_\alpha^\beta |X_t - X_t^n|^2 dt\right] = 0,$$

e per cui esiste il limite in media quadratica di $\mathcal{I}_{\alpha,\beta}^W(X^n) = \int_{(\alpha,\beta]} X_s^n\, dW_s$. Tale limite è denotato come $\mathcal{I}_{\alpha,\beta}^W(X) = \int_{(\alpha,\beta]} X_s\, dW_s$. Quindi si ha che

$$\lim_{n\to\infty} \mathbb{E}\left[\left|\mathcal{I}_{\alpha,\beta}^W(X^n) - \mathcal{I}_{\alpha,\beta}^W(X)\right|^2\right] = 0,$$

ossia che

$$\mathbb{E}\left[\left|\sum_{i=0}^{m_n-1} X_{t_i^n}\big(W_{t_{i+1}^n} - W_{t_i^n}\big) - \int_{(\alpha,\beta]} X_s\, dW_s\right|^2\right]$$

La seguente parte corrisponde ad un approfondimento, ma non è necessaria per gli studenti del Master.
I processi elementari predicibili (e anche quelli opzionali) sono esempi di processi (congiuntamente) misurabili:

Definizione 5.3 (Processi (congiuntamente) misurabili)**.** *Sia $X : [\alpha,\beta]\times\Omega :\mapsto \mathbb{R}$, $(t,\omega)\mapsto X(t,\omega)\big(= X_t(\omega)\big)$. Sia $\mathcal{B}([\alpha,\beta])\times\mathcal{F}$ la σ-algebra prodotto su $[\alpha,\beta]\times\Omega$, ovvero la σ-algebra generata da $\{J\times A\subseteq[\alpha,\beta]\times\Omega;\ J\in\mathcal{B}([\alpha,\beta]), A\in\mathcal{F}\}$. Il processo $(X_t)_t$ si dice* **(congiuntamente) misurabile** *se la funzione X è $\mathcal{B}([\alpha,\beta])\times\mathcal{F}$-misurabile.*

In realtà si tratta di processi progressivamente misurabili e continui a sinistra (continui a destra nel caso dei processi opzionali).
Inoltre va richiesto che le variabili aleatorie coinvolte nella definizione del processo predicibile elementare siano di quadrato sommabile (o addirittura limitati). Con questa richiesta (ovvero se C_i sono anche di quadrato integrabile) il processo elementare predicibile (o opzionale) X è un elemento del seguente spazio di processi, che è anche uno spazio vettoriale, metrico e completo.[2]

Definizione 5.4. *Lo spazio $\mathbb{L}^2\big([\alpha,\beta]\times\Omega\big) = \mathbb{L}^2\big([\alpha,\beta]\times\Omega, \mathcal{B}([\alpha,\beta])\times\mathcal{F}, \lambda\times\mathbb{P}\big)$, dove λ indica la misura di Lebesgue, è lo spazio delle classi di equivalenza dei processi (congiuntamente) misurabili per i quali*

$$\mathbb{E}\left[\int_\alpha^\beta |X_s|^2 ds\right] < \infty. \tag{5.6}$$

Parlando di classi di equivalenza si intende che identifichiamo due processi X e X' se

$$\mathbb{E}\left[\int_\alpha^\beta |X_s - X_s'|^2 ds\right] = 0.$$

Si può definire quindi la chiusura di

$$\mathcal{E}^2([\alpha,\beta]) = \{X\in\mathcal{E},\ \text{con } C_i \text{ di quadrato sommabile}\}$$

in tale spazio metrico, rispetto alla metrica

$$d(X,X') := \left(\mathbb{E}\left[\int_\alpha^\beta |X_s - X_s'|^2 ds\right]\right)^{\frac{1}{2}}.$$

Denoteremo tale chiusura come

$$\overline{\mathcal{E}^2}([\alpha,\beta]).$$

[2]† Il fatto che $\mathbb{L}^2\big([\alpha,\beta]\times\Omega\big)$ sia uno spazio vettoriale è ovvio, in quanto se (5.6) vale per due processi X_1 e X_2, allora vale anche per la sua combinazione lineare $Y = a_1X_1 + a_2X_2$.
Inoltre $\mathbb{L}^2\big([\alpha,\beta]\times\Omega\big)$ è uno spazio metrico completo con la metrica

$$d(X,X') := \left(\mathbb{E}\left[\int_\alpha^\beta |X_s - X_s'|^2 ds\right]\right)^{\frac{1}{2}}.$$

In realtà si tratta di uno spazio di Hilbert.

È importante identificare lo spazio $\overline{\mathcal{E}^2}([\alpha,\beta])$, perché è quello su cui (come vedremo) sappiamo definire l'integrale stocastico[3].

È altrettanto importante individuare almeno una classe ampia di processi che vi appartengono: è possibile mostrare che se X è un processo adattato e X ha traiettorie continue, ed è un processo in $\mathbb{L}^2\big([\alpha,\beta]\times\Omega\big)$, ovvero se vale (5.6), allora X appartiene a $\overline{\mathcal{E}^2}([\alpha,\beta])$: infatti se $\{t_i^n\}_{i=0,\cdots,m_n}$ è una partizione di $[\alpha,\beta]$ con $\max_i |t_i^n - t_{i-1}^n| \to 0$, allora la successione di processi $\{X^n\}$ definiti da

$$X_t^n(\omega) = \sum_{i=1}^{m_n} X_{t_{i-1}^n}(\omega) I_{(t_{i-1}^n, t_i^n]}(t),$$

è una successione di processi elementari e predicibili in $\mathbb{L}^2\big([\alpha,\beta]\times\Omega\big)$ che converge ad X nella metrica d, ovvero che

$$\lim_{n\to\infty} \mathbb{E}\left[\int_\alpha^\beta |X_t - X_t^n|^2 dt\right] = 0.$$

Se $X \in \overline{\mathcal{E}^2}([\alpha,\beta])$ allora, per definizione, esiste una successione di processi elementari e predicibili che soddisfa la precedente relazione. Di conseguenza la successione X_n è una successione di Cauchy in $\mathbb{L}^2\big([\alpha,\beta]\times\Omega\big)$, ovvero

$$\lim_{n,m\to\infty} \mathbb{E}\left[\int_\alpha^\beta |X_t^m - X_t^n|^2 dt\right] = 0.$$

Sappiamo inoltre che, per $\alpha = 0$ e $\beta = T$, per ogni processo elementare $\{X_n\}$, il processo $\mathcal{I}^W(X^n) = (\mathcal{I}_t^W(X^n))_{t\in[0,T]}$ definisce una martingala a traiettorie continue . Ovviamente anche la differenza $\mathcal{I}_t^W(X^n) - \mathcal{I}_t^W(X^m)$ è una martingala. Inoltre, chiaramente, per l'integrale stocastico dei processi elementari valgono le proprietà di linearità, e quindi

$$\mathcal{I}_t^W(X^n) - \mathcal{I}_t^W(X^m) = \mathcal{I}_t^W(X^n - X^m).$$

Dall'Esempio **??** sappiamo che, per ogni $t \in [0,T]$,

$$\mathbb{E}\left[\left|\mathcal{I}_t^W(X^n - X^m)\right|^2\right] = \mathbb{E}\left[\int_0^t |X_t^m - X_t^n|^2 dt\right]. \tag{5.7}$$

Inoltre applicando la disuguaglianza di Doob per $p = 2$, estesa alle martingale continue, alla martingala continua $\mathcal{I}^W(X^n - X^m)$, sappiamo che

$$\mathbb{E}\left[\max_{t\in[0,T]} \left|\mathcal{I}_t^W(X^n - X^m)\right|^2\right] \le 4\mathbb{E}\left[\left|\mathcal{I}_T^W(X^n - X^m)\right|^2\right]. \tag{5.8}$$

D'altra parte, ovviamente, si ha anche che

$$\left|\mathcal{I}_T^W(X^n - X^m)\right|^2 \le \max_{t\in[0,T]} \left|\mathcal{I}_t^W(X^n - X^m)\right|^2.$$

Passando ai valori attesi, nella precedente disuguaglianza, e tenendo conto della (5.8), si ha

$$\mathbb{E}\left[\left|\mathcal{I}_T^W(X^n - X^m)\right|^2\right] \le \mathbb{E}\left[\max_{t\in[0,T]} \left|\mathcal{I}_t^W(X^n - X^m)\right|^2\right] \le 4\mathbb{E}\left[\left|\mathcal{I}_T^W(X^n - X^m)\right|^2\right].$$

Utilizzato la relazione (5.7) per $t = T$, si ottiene che

$$\mathbb{E}\left[\int_0^T |X_t^m - X_t^n|^2 dt\right] \le \mathbb{E}\left[\max_{t\in[0,T]} \left|\mathcal{I}_t^W(X^n - X^m)\right|^2\right] \le 4\mathbb{E}\left[\int_0^T |X_t^m - X_t^n|^2 dt\right] \tag{5.9}$$

In altre parole, se i processi elementari X^n sono una successione di Cauchy in $\mathbb{L}^2\big([0,T]\times\Omega\big)$, allora la successione di martingale a traiettorie continue $\mathcal{I}^W(X^n) = (\mathcal{I}_t^W(X^n))_{t\in[0,T]}$, risulta una successione di Cauchy[4] rispetto alla metrica

$$\tilde{d}(M, M') := \left(\mathbb{E}\left[\max_{t\in[0,T]} |M_t - M_t'|^2\right]\right)^{\frac{1}{2}};$$

in quanto la (5.9) si può riscrivere come

$$d^2(X^n, X^m) \le \tilde{d}^2(\mathcal{I}^W(X^n), \mathcal{I}^W(X^m)) \le 4d^2(X^n, X^m)$$

Viceversa se i processi elementari X^n in $\mathbb{L}^2\big([0,T]\times\Omega\big)$ sono tali che la successione $\mathcal{I}^W(X^n) = (\mathcal{I}_t^W(X^n))_{t\in[0,T]}$ è una successione di Cauchy rispetto alla metrica $\tilde{d}$ allora la successione di processi X^n risulta di Cauchy in $\mathbb{L}^2\big([0,T]\times\Omega\big)$ rispetto alla metrica d.

---

[3] La classe **$\overline{\mathcal{E}^2}([\alpha,\beta])$ che abbiamo definito comprende tutti i processi $(\mathcal{F}_t)$-predicibili di $\mathbb{L}^2\big([\alpha,\beta]\times\Omega\big)$. Ricordiamo qui la definizione di processi $(\mathcal{F}_t)$-predicibili: i processi predicibili sono i processi da $[\alpha,\beta]\times\Omega \to \mathbb{R}$ e misurabili rispetto alla σ-algebra su $[\alpha,\beta]\times\Omega$, generata dai processi elementari e predicibili, o equivalentemente dai processi $(\mathcal{F}_t)$-adattati e continui a sinistra. Più precisamente si definisce $\mathcal{P} = \mathcal{P}\big((\mathcal{F}_t)_t\big)$ la σ-algebra su $[\alpha,\beta]\times\Omega$ generata dai processi elementari predicibili. Lo spazio $\overline{\mathcal{E}^2}([\alpha,\beta])$ è allora lo spazio $\mathbb{L}^2\big([\alpha,\beta]\times\Omega, \mathcal{P}, \lambda\times\mathbb{P}\|_{\mathcal{P}}\big)$. Tale spazio è un sottospazio chiuso di $\mathbb{L}^2\big([\alpha,\beta]\times\Omega,\big)$ in quanto i processi elementari predicibili sono congiuntamente misurabili e quindi $\mathcal{P}$ è contenuta nella σ- algebra prodotto. Va anche detto che, nel caso del processo di Wiener, si potrebbe estendere l'integrale anche ai processi opzionali, la cui definizione è simile ai processi predicibili, ma a partire dai processi elementari opzionali.

[4] In altre parola, stiamo considerando $\mathcal{I}^W(X^n) = (\mathcal{I}_t^W(X^n))_{t\in[0,T]}$ come un elemento dello spazio metrico completo $L^2(\Omega; C[0,T], \mathbb{P})$. Ovviamente identifichiamo processi che sono indistinguibili, ossia che differiscono su un insieme di probabilità nulla.

Da questa relazione si può dedurre l'esistenza di un processo limite, in questa metrica $\tilde{d}$, della successione di processi $\mathcal{I}^W(X_n)$. Il processo limite[5] viene chiamato integrale stocastico di X e denotato anche esso con

$$\mathcal{I}_t^W(X) = \int_0^t X_s dW_s.$$

Si osservi che la convergenza considerata implica la convergenza in probabilità di

$$\sup_{t\imath[0,T]} \left| \mathcal{I}_t^W(X_n) - \mathcal{I}_t^W(X) \right|$$

Inoltre il processo limite $\mathcal{I}_t^W(X)$ è a sua volta una martingala di quadrato integrabile[6], e che la sua variazione quadratica è il processo $\int_0^t X_s ds$**[7].

Ribadiamo che la definizione viene (in queste note) data quindi solo per i processi $X \in \overline{\mathcal{E}^2}([\alpha,\beta])$.

È chiaro che l'integrale stocastico $\int_0^t X_s dW_s$ è lineare in X. Inoltre $\int_0^t X_s dW_s$ è una martingala a traiettorie continue. Nel caso dei processi elementari ciò discende in modo simile a quanto visto all'inizio del capitolo per i valori attesi:

Lemma 5.1. *Se X è un processo elementare predicibile, come nella (5.4), con C_i di quadrato integrabile, allora*

$$\mathbb{E}\left(\int_\alpha^\beta X_t dW_t \middle| \mathcal{F}_\alpha \right) = 0$$

$$\mathbb{E}\left[\left(\int_\alpha^\beta X_t dW_t \right)^2 \middle| \mathcal{F}_\alpha \right] = \mathbb{E}\left[\int_\alpha^\beta X_t{}^2 \middle| \mathcal{F}_\alpha \right].$$

In particolare, l'integrale stocastico di un processo elementare di $\mathcal{E}^2$ è una variabile aleatoria centrata di quadrato integrabile e

$$\mathbb{E}\left[\left(\int_\alpha^\beta X_t dW_t \right)^2 \right] = \mathbb{E}\left[\int_\alpha^\beta X_t{}^2 dt \right].$$

Le proprietà di linearità, di martingala, di continuità e le proprietà precedenti sui valori medi si ottengono per i processi più generali, con un passaggio al limite (**vedere le note 6 e 7).

---

[5]Si può mostrare che il limite non dipende dalla particolare successione $\{X^n\}$ di processi elementari, scelta per approssimare il processo X.

[6] La convergenza di $\mathcal{I}_t^W(X_n)$ a $\mathcal{I}_t^W(X)$ nella metrica $\tilde{d}$ implica la convergenza **a zero di

$$\mathbb{E}\left[\left| \mathcal{I}_t^W(X_n) - \mathcal{I}_t^W(X) \right| \right], \qquad \forall t \in [0,T], **$$

ed implica l'esistenza di una sottosuccessione $\mathcal{I}_t^W(X_{n_k})$ convergente quasi certamente ed in norma uniforme. Inoltre, se una successione di $\mathcal{G}_t$-martingale $M_t^{(k)}$ converge, per k che tende ad infinito, a un processo M_t, allora il processo limite M_t è una $\mathcal{G}_t$-martingala: infatti, per ipotesi, qualunque siano $k \geq 1$, $s \leq t$ e $A \in \mathcal{G}_s$, si ha

$$\mathbb{E}[M_t^{(k)} \mathbb{I}_A] = \mathbb{E}[M_s^{(k)} \mathbb{I}_A].$$

Basta allora passare al limite per k che tende ad infinito, per ottenere che

$$\mathbb{E}[M_t \mathbb{I}_A] = \mathbb{E}[M_s \mathbb{I}_A],$$

in quanto, ad esempio

$$\left| \mathbb{E}[M_t^{(k)} \mathbb{I}_A] - \mathbb{E}[M_t \mathbb{I}_A] \right| \leq \mathbb{E}\left[\left| M_t^{(k)} - M_t \right| \mathbb{I}_A \right] \leq \mathbb{E}\left[\left| M_t^{(k)} - M_t \right| \right].$$

[7]**La convergenza di di $\mathcal{I}_t^W(X_n)$ a $\mathcal{I}_t^W(X)$ nella metrica $\tilde{d}$ implica la convergenza a zero di

$$\mathbb{E}\left[\left| \mathcal{I}_t^W(X_n) - \mathcal{I}_t^W(X) \right|^2 \right], \qquad \forall t \in [0,T],$$

oltre ovviamente all'ipotesi di convergenza a zero di

$$\mathbb{E}\left[\int_0^t |X_s - X_s^n|^2 ds \right].$$

La prima convergenza a zero implica la convergenza a zero di

$$\mathbb{E}\left[\left| \left(\mathcal{I}_t^W(X_n) \right)^2 - \left(\mathcal{I}_t^W(X) \right)^2 \right| \right], \qquad \forall t \in [0,T],$$

mentre la seconda implica la convergenza a zero di

$$\mathbb{E}\left[\left| \int_0^t |X_s|^2 ds - \int_0^t |X_s^n|^2 ds \right| \right].$$

Quindi la successione di martingale $\left(\mathcal{I}_t^W(X_n) \right)^2 - \int_0^t |X_s^n|^2 ds$ converge in L^1 al processo $\left(\mathcal{I}_t^W(X) \right)^2 - \int_0^t |X_s|^2 ds$, che risulta quindi una martingala (si procede come nella precedente nota 6).

5.1.1 Esempi di calcolo di integrali stocastici

Procediamo con qualche esempio di calcolo di integrale stocastico.

Esempio 5.1 (integrale stocastico di $X_t = W_t$). *In questo esempio proviamo a calcolare*

$$\int_\alpha^\beta 2W_s dW_s.$$

Essendo $X_t = 2W_t$ un processo a traiettorie continue, con

$$\mathbb{E}\left[\int_\alpha^\beta |2W_s|^2 ds\right] = \int_\alpha^\beta \mathbb{E}\left[|2W_s|^2\right] ds = \int_\alpha^\beta 4s ds < \infty$$

possiamo utilizzare come processo approssimante[8]

$$X_t^n = \sum_{i=1}^{m_n} W_{t_{i-1}^n}(\omega) I_{(t_{i-1}^n, t_i^n]}(t),$$

dove si considera la partizione $t_0^n = \alpha < t_1^n < \cdots < t_{m_n}^n = \beta$. Così

$$\int_\alpha^\beta 2W_s dW_s = \lim_n \sum_{i=1}^{m_n} W_{t_{i-1}^n}(W_{t_i^n} - W_{t_{i-1}^n}).$$

Considerando che

$$\begin{aligned}
2\sum_k b_k(b_{k+1} - b_k) &= \sum_k [2b_{k+1}b_k - 2b_k^2] = \sum_k [2b_{k+1}b_k - 2b_k^2 - b_{k+1}^2 + b_{k+1}^2] \\
&= \sum_k [(2b_{k+1}b_k - b_k^2 - b_{k+1}^2) + b_{k+1}^2 - b_k^2] = -\sum_k [b_{k+1} - b_k]^2 + \sum_k [b_{k+1}^2 - b_k^2] \\
&= -\sum_k [b_{k+1} - b_k]^2 + [b_{k_{max}}^2 - b_{k_{min}}^2]
\end{aligned}$$

ponendo $b_k = W_{t_{i-1}^n}$ si ottiene

$$\int_\alpha^\beta 2W_s dW_s = \lim_n -\sum_{i=1}^{m_n} (W_{t_i^n} - W_{t_{i-1}^n})^2 + W_\beta^2 - W_\alpha^2.$$

da cui, ricordando il Lemma 3.2, si ottiene che

$$\int_\alpha^\beta 2W_s dW_s = -(\beta - \alpha) + W_\beta^2 - W_\alpha^2. \tag{5.10}$$

C'è un altro caso in cui si può calcolare l'integrale stocastico: Se $f \in \mathbb{L}^2([0,T], \lambda)$ è una funzione deterministica, allora $f \in \overline{\mathcal{E}^2}([0,T])$ e quindi ha senso l'integrale stocastico di Itô, che in questo caso coincide con l'integrale stocastico di Wiener, e gode allora di una importante proprietà.

Proposizione 5.2. *Se f è una funzione deterministica, con $f \in \mathbb{L}^2([0,T])$, allora il processo*

$$\int_0^t f(s) dW_s$$

è gaussiano, di media zero e varianza $\int_0^t f^2(s) ds$.

[8] **Questo tipo di approssimazione si ottiene per ogni processo continuo a sinistra e adattato.

Dimostrazione. L'idea della dimostrazione è la seguente, ed è sostanzialmente l'idea seguita da Wiener, nel definire l'integrale stocastico per processi deterministici.

Sia f è una funzione deterministica e C^1, allora si osserva che vale la seguente formula di integrazione per parti (questa era la definizione originaria di Wiener)

$$\int_0^T f(s)dW_s := f(T)\,W_T - f(0)\,W_0 - \int_0^T W_s\,f'(s)\,ds. \tag{5.11}$$

Per dimostrare questa formula di integrazione per parti basta osservare che, se si considera la partizione $\pi:\ t_0^n = 0 < t_1^n < \cdots < t_{2^n}^n = T$, con $t_i^n = \frac{i}{2^n}\,T$, allora, il processo deterministico $f(s)$ si può approssimare con

$$f(s) = \lim_{n\to\infty} \sum_{i=1}^{2^n} f(t_{i-1}^n)\,\mathbf{1}_{(t_{i-1}^n, t_i^n]}(s)$$

$$\begin{aligned}
\int_0^T f(s)dW_s &= \lim_{n\to\infty} \sum_{i=1}^{2^n} f(t_{i-1}^n)\,[W_{t_i^n} - W_{t_{i-1}^n}] \qquad \text{(aggiungendo e togliendo i termini } f(t_i^n)\,W_{t_i^n})\\
&= \lim_{n\to\infty} \sum_{i=1}^{2^n} \Big([f(t_i^n)\,W_{t_i^n} - f(t_{i-1}^n)\,W_{t_{i-1}^n}] + [f(t_{i-1}^n) - f(t_i^n)]\,W_{t_i^n} \Big)\\
&= [f(T)\,W_T - f(0)\,W_0] + \\
&= \frac{1}{2}\,[f(T)\,W_T - f(0)\,W_0] + \sum_{i=1}^{2^n} \big(f'(t_i^n)[t_{i-1}^n - t_i^n] + |t_{i-1}^n - t_i^n|\,O_i(1) \big)\ W_{t_i^n}
\end{aligned}$$

dove $O_i(1)$ è infinitesimo uniformemente in i, e quindi, essendo $f'(t)\ W_t(\omega)$ continue e quindi limitate in $[0,T]$ si ottiene che

$$\int_0^T f(s)dW_s = f(T)\,W_T - f(0)\,W_0 - \lim_{n\to\infty} \sum_{i=1}^{2^n} f'(t_i^n)\,W_{t_i^n}\,(t_i^n - t_{i-1}^n),$$

e quindi la (5.11).

Si osservi che, se invece di considerare l'intervallo $[0,T]$ si fosse considerato l'intervallo $[\alpha,\beta]$, avremmo ottenuto

$$\int_\alpha^\beta f(s)dW_s = f(\beta)\,W_\beta - f(\alpha)\,W_\alpha - \lim_{|\pi|\to 0} \sum_{i=1}^{2^n} W_{t_i^n} f'(t_i^n)\,(t_i^n - t_{i-1}^n) \tag{5.12}$$

Dalla (5.12) si ottiene immediatamente che l'integrale stocastico è una variabile aleatoria gaussiana, in quanto limite di variabili aleatorie gaussiane[9]. Ovviamente si tratta di variabili gaussiane a media nulla ed è facile vedere che la varianza di $f(\beta)\,W_\beta - f(\alpha)\,W_\alpha - \sum_{i=1}^{2^n} W_{t_i^n} f'(t_i^n)\,(t_{i-1}^n - t_{i-1}^n)$ converge al valore $\int_\alpha^\beta f^2(s)ds$. Per il caso generale di funzioni f, basta considerare che lo spazio $C^1([0,T])$ è denso in $\mathbb{L}^2([0,T])$. □

Osservazione 5.1. *La proposizione precedente implica in particolare che*

$$\mathbb{E}\left[\exp\left(\int_0^t f(s)dW_s\right)\right] = \exp\left(\frac{1}{2}\int_0^t f^2(s)ds\right).$$

Infatti si riconosce a sinistra la media dell'esponenziale di una variabile aleatoria gaussiana centrata Z, con varianza σ_Z^2, cioè la sua trasformata di Laplace $\mathcal{L}(\theta) := \mathbb{E}[\exp(\theta Z)]$ calcolata in $\theta = 1$, che è uguale all'esponenziale della sua varianza divisa per 2 (ovvero $\mathcal{L}(\theta) = \exp\{\theta^2\sigma_Z^2/2\}$.

[9] Ciò si può dimostrare, ad esempio, usando il fatto che il limite di funzioni caratteristiche di tipo gaussiano converge se e solo se convergono valore atteso e varianza. Ovviamente si considera gaussiano ache il caso degenere di variabili con varianza nulla.

Terminiamo la sezione con la seguente osservzione, l'integrale stocastico si può estendere a una classe più ampia dei processi che abbiamo visto, questa classe è denotata con $M^2([0,T])$. Se la filtrazione $(\mathcal{F}_t)_{t\geq 0} = (\mathcal{F}_t^W)_{t\geq 0}$ è quella naturale completata[10] del moto browniano W, allora ogni variabile aleatoria $Z \in \mathbb{L}^2(\Omega, \mathcal{F}, \mathbf{P})$, con $\mathcal{F} = \mathcal{F}_T$, si può rappresentare nella forma

$$Z = c + \int_0^T X_s dW_s,$$

con $X \in M^2([0,T])$ e $c \in \mathbb{R}$.

Breve approfondimento su questo punto

Definizione 5.5. *Sia $M^2([\alpha,\beta])$ il sottospazio di $\mathbb{L}^2\big([\alpha,\beta]\times\Omega\big)$ definito come lo spazio delle classi d'equivalenza dei* **processi $\mathcal{F}_t^W$-progressivamente misurabili**[11] *tali che*

$$\mathbb{E}\left[\int_\alpha^\beta |X_s|^2 ds\right] < \infty.$$

Osservazione 5.2. *Si può definire l'integrale stocastico anche per il processi elementari opzionali e si può dimostrare che l'integrale stocastico è una isometria tra $M^2([0,T])$ (il sottospazio di $\mathbb{L}^2\big([0,T]\times\Omega\big)$ dei processi progressivamente misurabili) e lo spazio delle v.a. $\mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbf{P})$.*

Alla luce della precedente osservazione ci si può chiedere se tutte le variabili aleatorie di $\mathbb{L}^2(\Omega, \mathcal{F}, \mathbf{P})$ si possono ottenere come integrale stocastico di un processo di $M^2([0,T])$. Ciò non può essere dal momento che ogni integrale stocastico in $M^2([0,T])$ ha media nulla. Tuttavia si può vedere che, se la filtrazione $(\mathcal{F}_t)_{t\geq 0} = (\mathcal{F}_t^W)_{t\geq 0}$ è quella naturale completata del moto browniano W, allora ogni variabile aleatoria $Z \in \mathbb{L}^2(\Omega, \mathcal{F}, \mathbf{P})$, con $\mathcal{F} = \mathcal{F}_T$, si può rappresentare nella forma

$$Z = c + \int_0^T X_s dW_s,$$

con $X \in M^2([0,T])$ e $c \in \mathbb{R}$ e, naturalmente, con $\mathbb{E}(Z) = c$.

Prendendo $Z = M_T$, questa proprietà è equivalente alla proprietà che ogni martingala[12] $(M_t)_{t\in[0,T]}$ di quadrato integrabile si possa rappresentare tramite un'integrale stocastico, come

$$M_t = \mathbb{E}[M_T|\mathcal{F}_t] = c + \int_0^t X_s dW_s$$

La seguente sezione contiene degli approfondimenti non necessari per gli studenti del master

5.1.2 Integrale stocastico per processi localmente di quadrato integrabile

In questa sezione affrontiamo il problema della definizione dell'integrale stocastico anche per una classe più ampia di processi[13]. La seguente proposizione afferma che, se l'integrando è continuo, allora l'integrale stocastico è il limite di particolari[14] somme di Riemann, in analogia con l'integrale di Lebesgue.

Proposizione 5.3. *Se X è un processo dello spazio $\Lambda^2[\alpha,\beta]$, ovvero è progressivamente misurabile e*

$$\mathbb{P}\left(\int_\alpha^\beta X_t^2 dt < \infty\right) = 1,$$

allora si può dare un senso all'integrale stocastico $\int_\alpha^\beta X_t dW_t$, come limite in probabilità di integrali di processi elementari opzionali. Se inoltre X è un processo continuo, allora per ogni successione $(\pi_n)_n$ di partizioni $\alpha = t_0^n < t_1^n < \ldots < t_{m_n}^n = \beta$ con $|\pi_n| \to 0$ si ha

$$\sum_{k=1}^{m_n} X(t_{k-1}^n)(W_{t_k^n} - W_{t_{k-1}^n}) \underset{n\to+\infty}{\longrightarrow} \int_\alpha^\beta X(t) dW_t \qquad \textit{in probabilità.}$$

La dimostrazione di tale risultato è basata su due risultati:

[10] Per ottenere il completamento della filtrazione si aggiungono gli tutti gli insiemi di probabilità nulla e i loro sottoinsiemi.

[11] Più in generale, data una filtrazione $(\mathcal{F}_t)$, i processi $\mathcal{F}_t$-progressivamente misurabili sono i processi $X : \Omega \times [0,\infty) \mapsto \mathbb{R}$ tali che qualunque sia t la restrizione di X a funzione $\Omega \times [0,t]$ è misurabile rispetto alla σ-algebra prodotto $\mathcal{F}_t \times \mathcal{B}([0,t])$.

[12] Si tratta sempre del caso in cui la filtrazione 'e la filtrazione generata da W, completata e resa continua a destra.

[13] Questa parte non è stata ancora rivista e usa un approccio diverso da quello visto a lezione: si consiglia di vedere invece gli appunti del Prof. Bertini[**?**].

[14] Una somma di Riemann è una somma del tipo

$$\sum_{k=1}^{m_n} X(t_{k-1}^{n*})(W_{t_k^n} - W_{t_{k-1}^n})$$

con $t_{k-1}^{n*} \in [t_{k-1}^n, t_k^n]$ La differenza nel caso dell'integrale stocastico di Itô, sta nel fatto che va sempre scelto $t_{k-1}^{n*} = t_{k-1}^n$.

Lemma 5.4 (Baldi [?], Lemma???). *Se $X \in \Lambda^2([\alpha,\beta])$, allora*

1 esiste una successione $(Y^{(n)})_n$ di processi continui appartenenti a $\Lambda^2([\alpha,\beta])$ e tali che

$$\lim_{n\to\infty} \int_\alpha^\beta |X_t - Y_t^{(n)}|^2 dt = 0 \quad q.c.$$

2 esiste una successione $(X^{(n)})_n$ di processi elementari appartenenti a $\Lambda^2([\alpha,\beta])$ e tali che

$$\lim_{n\to\infty} \int_\alpha^\beta |X_t - X_t^{(n)}|^2 dt = 0 \quad q.c.$$

Lemma 5.5 (Baldi [?], lemma ???). *Per ogni processo elementare $X \in \Lambda^2([\alpha,\beta])$ e per ogni $\varepsilon > 0$, $\rho > 0$*

$$\mathbf{P}\left(\left|\int_\alpha^\beta X_t dW_t\right| > \varepsilon\right) \leq \mathbf{P}\left(\int_\alpha^\beta X_t^2 dt > \rho\right) + \frac{\rho}{\varepsilon^2}.$$

Dimostrazione. In realtà si può dimostrare un risultato più forte:

$$\mathbf{P}\left(\sup_{b\in[\alpha,\beta]}\left|\int_\alpha^b X_t dW_t\right| > \varepsilon\right) \leq \mathbf{P}\left(\int_\alpha^\beta X_t^2 dt > \rho\right) + \frac{\rho}{\varepsilon^2}. \tag{5.13}$$

La dimostrazione è basata sul fatto che, posto $X_t^\rho = X_t \, \mathbf{1}_{[0,\rho]}\big(\int_\alpha^t X_s^2 ds\big)$,

$$\begin{aligned}
\mathbf{P}\left(\sup_{b\in[\alpha,\beta]}\left|\int_\alpha^b X_t dW_t\right| > \varepsilon\right) &= \mathbf{P}\left(\sup_{b\in[\alpha,\beta]}\left|\int_\alpha^b X_t dW_t\right| > \varepsilon, \int_\alpha^\beta X_t^2 dt > \rho\right) + \mathbf{P}\left(\sup_{b\in[\alpha,\beta]}\left|\int_\alpha^b X_t dW_t\right| > \varepsilon, \int_\alpha^\beta X_t^2 dt \leq \rho\right) \\
&\leq \mathbf{P}\left(\int_\alpha^\beta X_t^2 dt > \rho\right) + \mathbf{P}\left(\sup_{b\in[\alpha,\beta]}\left|\int_\alpha^b X_t^\rho dW_t\right| > \varepsilon, \int_\alpha^\beta X_t^2 dt \leq \rho\right)
\end{aligned}$$

dove si è usato il fatto che, se $\int_\alpha^\beta X_t^2 dt \leq \rho$ allora $X_t = X_t^\rho$,

$$\begin{aligned}
&\leq \mathbf{P}\left(\int_\alpha^\beta X_t^2 dt > \rho\right) + \mathbf{P}\left(\sup_{b\in[\alpha,\beta]}\left|\int_\alpha^b X_t^\rho dW_t\right| > \varepsilon\right) \\
&\leq \mathbf{P}\left(\int_\alpha^\beta X_t^2 dt > \rho\right) + \frac{\mathbf{E}\left(\left|\int_\alpha^\beta X_t^\rho dW_t\right|^2\right)}{\varepsilon^2} \leq \mathbf{P}\left(\int_\alpha^\beta X_t^2 dt > \rho\right) + \frac{\rho}{\varepsilon^2},
\end{aligned}$$

dove si è usata la disuguaglianza di Kolmogorov e il fatto che X_t^ρ è di quadrato integrabile, e quindi

$$\mathbf{E}\left(\left|\int_\alpha^\beta X_t^\rho dW_t\right|^2\right) = \mathbf{E}\left(\int_\alpha^\beta (X_t^\rho)^2 \, d_t\right) \leq \rho$$

□

Definiamo l'integrale stocastico per ogni processo $X \in \Lambda^2([\alpha,\beta])$. Per il Lemma 5.4 esiste una successione di processi elementari $(X_n)_n \subset \Lambda^2([\alpha,\beta])$ tale che

$$\int_\alpha^\beta |X_n(t) - X(t)|^2 dt \longrightarrow_{n\to\infty} 0 \qquad \text{in probabilità.}$$

Quindi, per ogni $\eta > 0$, $\delta > 0$ esiste n_0 tale che per $n, m > n_0$

$$\mathbf{P}\left(\int_\alpha^\beta |X_n(t) - X(t)|^2 dt > \delta\right) < \eta$$

e, per il Lemma 5.5, per ogni $\rho > 0$

$$\mathbf{P}\left(\left|\int_\alpha^\beta X_n(t) dW_t - \int_\alpha^\beta X_m(t) dW_t\right| > \varepsilon\right) \leq \mathbf{P}\left(\int_\alpha^\beta |X_n(t) - X_m(t)|^2 dt > \rho\right) + \frac{\rho}{\varepsilon^2}.$$

Scegliamo, ora, prima ρ abbastanza piccolo perché sia $\frac{\rho}{\varepsilon^2} < \frac{\varepsilon}{2}$ e poi n_0 abbastanza grande perché per $n, m > n_0$ il primo termine a destra sia strettamente minore di $\frac{\varepsilon}{2}$. La successione

$$\left(\int_\alpha^\beta X_n(t) dW_t\right)_n$$

è di Cauchy, e dunque convergente, in probabilità. Indichiamo con

$$\int_\alpha^\beta X_t dW_t$$

il suo limite. Esso si chiama *integrale stocastico* di X rispetto a W.

5.1.3 Osservazioni su alcune proprietà degli integrali stocastici

In questa sezione si usano alcune notazioni introdotte nelle sezioni di approfondimento.

Osservazione 5.3. *Sappiamo che, per ogni a e $b \in \mathbb{R}$*

$$\int_\alpha^\beta [aX_t + bY_t] dW_t = a \int_\alpha^\beta X_t dW_t + b \int_\alpha^\beta Y_t dW_t,$$

se X e Y sono processi per cui ha senso l'integrale stocastico, (ad esempio se $X, Y \in \overline{\mathcal{E}}^2([\alpha,\beta])$, o se $X, Y \in \Lambda^2([\alpha,\beta])$). Ma se $A = A(\omega)$ è una variabile aleatoria (per semplicità prendiamo $b = 0$), non possiamo dire altrettanto, o meglio dipende dalle proprietà di misurabilità di A. Infatti, mentre $A \int_\alpha^\beta X_t dW_t$ si può scrivere sempre, sappiamo dare un significato a $\int_\alpha^\beta A(\omega) X_t dW_t$ solo se $AX \in \overline{\mathcal{E}}^2([\alpha,\beta])$ (o se $AX \in \Lambda^2([\alpha,\beta])$). In particolare AX_t deve essere progressivamente misurabile, di conseguenza è necessario che AX_t sia $\mathcal{F}_\alpha$-misurabile. Affinché ciò si verifichi è sufficiente che A sia $\mathcal{F}_\alpha$-misurabile. Per le condizioni di integrabilità è sufficiente che A sia una v.a. limitata.

Dunque possiamo enunciare il seguente teorema.

Teorema 5.6. *Sia A una variabile aleatoria $\mathcal{F}_\alpha$-misurabile limitata; allora per $X \in \overline{\mathcal{E}}^2([\alpha,\beta])$ (o per $X \in \Lambda^2([\alpha,\beta])$)*

$$\int_\alpha^\beta AX(t) dW_t = A \int_\alpha^\beta X(t) dW_t.$$

5.2 Calcolo stocastico e formula di Itô

Sia X un processo tale che per ogni $0 \leq t_1 < t_2 \leq T$

$$X_{t_2} - X_{t_1} = \int_{t_1}^{t_2} F_t dt + \int_{t_1}^{t_2} G_t dW_t,$$

dove F_t e G_t soddisfano opportune condizioni.[15]

Diremo allora che X ammette il ***differenziale stocastico***

$$dX_t = F_t dt + G_t dW_t.$$

Un processo che ammette differenziale stocastico si chiama anche un ***processo di Itô***.

Come primo esempio banale, prendiamo $X_t = W_t$ allora ovviamente $dW_t = dW_t$ e quindi $F_t = F_t^W = 0$ e $G_t = G_t^W = 1$. Lo stesso vale anche per $X_t = x + W_t$, con x costante reale, visto che nella definizione del differenziale, la costante x scompare.

Come secondo esempio consideriamo $X_t = W_t^2$ (o anche $X_t = x + W_t^2$). Ricordiamo che per la (5.10) si ha che qualunque siano $0 \leq t_1 < t_2 \leq T$, si ha

$$X_{t_2} - X_{t_1} = W_{t_2}^2 - W_{t_1}^2 = \int_{t_1}^{t_2} 2W_s dW_s + (t_2 - t_1),$$

ovvero in forma differenziale

$$dX_t = dW_t^2 = 2W_t dW_t + dt,$$

che è come dire che $F_t = F_t^{W^2} = 1$ e $G_t = G_t^{W^2} = 2W_t$. Si noti la differenza con il calcolo del differenziale usuale.[16]

Osservazione 5.4. *Il differenziale stocastico, se esiste, è unico, nel senso che se X ammette una rappresentazione del tipo precedente, allora F e G sono determinati univocamente*[17].

[15] Per quanto riguarda le proprietà di misurabilità, si richiede che sia F_T che G_t siano $\mathcal{F}_t^W$-progressivamente misurabili. Inoltre si richiede che abbiano le opportune proprietà di integrabilità per cui abbia senso fare gli integrali.

[16] Il motivo sta nel fatto che il processo di Wiener non ha traiettorie a variazione limitata o meglio possiamo interpretare il Lemma 3.2 come $(W_{t+\delta} - W_t)^2 = O(\delta)$.

[17] A meno di identificazione di processi.

Definiamo[18]

$$< X >_t = \int_0^t {G_s}^2 ds.$$

In maniera simile, se Y è un altro processo di Itô, avente differenziale stocastico

$$dY_t = H_t dt + K_t dW_t,$$

porremo

$$< X, Y >_t = \int_0^t G_s K_s ds.$$

Proposizione 5.7 (Differenziale del prodotto)**.** *Se X_i, $i = 1, 2$, sono due processi aventi differenziale stocastico*

$$dX_i(t) = F_i(t)dt + G_i(t)dW_t,$$

allora

$$\begin{aligned} d(X_1(t)X_2(t)) &= X_1(t)dX_2(t) + X_2(t)dX_1(t) + G_1(t)G_2(t)dt \\ &= X_1(t)dX_2(t) + X_2(t)dX_1(t) + d < X_1, X_2 >_t . \end{aligned}$$

Ovvero, se $t_1 < t_2$,

$$\begin{aligned} X_1(t_2)X_2(t_2) - X_1(t_1)X_2(t_1) &= \int_{t_1}^{t_2} [X_1(t)F_2(t) + X_2(t)F_1(t)]dt \\ &+ \int_{t_1}^{t_2} [X_1(t)G_2(t) + X_2(t)G_1(t)]dW_t \\ &+ \int_{t_1}^{t_2} G_1(t)G_2(t)dt. \end{aligned}$$

Dimostrazione. Daremo la dimostrazione solo nel caso in cui $F_i(t)$ e $G_i(t)$ sono processi aleatori costanti nell'intervallo di estremi t_1 e t_2, ossia $F_i(t) = F_i$ e $G_i(t) = G_i$, dove F_i e G_i sono variabili aleatorie $\mathcal{F}_{t_1}$-misurabili. Infatti, una volta ottenuta la validità della formula per il differenziale stocastico del prodotto in questo caso, si ottiene immediatamente la validità per il caso in cui $F_i(t)$ e $G_i(t)$ sono processi aleatori elementari (in quanto l'integrale stocastico è lineare). Ed infine il caso generale si ottiene passando al limite, considerando successioni di processi elementari $F_i^{(n)}(t)$ e $G_i^{(n)}(t)$ che convergono a $F_i(t)$ e $G_i(t)$, nel senso che

$$\mathbb{P}\left(\int_0^T |F_i^{(n)}(t) - F_i(t)|\, dt \to 0 \right) = 1$$

$$\mathbb{P}\left(\int_0^T |G_i^{(n)}(t) - G_i(t)|^2\, dt \to 0 \right) = 1.$$

Basta infine considerare che i processi $X_i^{(n)}(t)$, definiti con differenziale stocastico $dX_i^{(n)}(t) = F_i^{(n)}(t)dt + G_i^{(n)}(t)dW_t$, convergono ai processi $X_i(t)$ nel senso che converge in probabilità a zero $\sup_{t\in[0,T]} |X_i^{(n)}(t) - X(t)|$.

Torniamo al caso dei processi con $F_i(t) = F_i$ e $G_i(t) = G_i$. Per semplicità di notazione prenderemo $t_1 = 0$ e $t_2 = t$, di modo che

$$X_i(t) = X_i(0) + F_i\, t + G_i\, W_t,$$

con $X_i(0)$, F_i e G_i variabili aleatorie $\mathcal{F}_0$-misurabili. Si tratta allora di controllare

$$X_1(t)\, X_2(t) - X_1(0)\, X_2(0) := \big(X_1(0) + F_1\, t + G_1\, W_t\big)\, \big(X_2(0) + F_2\, t + G_2\, W_t\big) - X_1(0)\, X_2(0),$$

[18] Il processo $< X >$ non è altro che il processo crescente associato alla martingala $\int_0^t G_s dW_s$ che compare nella sua definizione.

coincide con

$$
\begin{aligned}
&\int_0^t X_1(u)[F_2\,du + G_2\,dW_u] + \int_0^t X_2(u)[F_1\,du + G_1\,dW_u] + G_1\,G_2\,t \\
&= \int_0^t [X_1(0) + F_1\,u + G_1\,W_u]\,[F_2\,du + G_2\,dW_u] + \int_0^t [X_2(0) + F_2\,u + G_2\,W_u]\,[F_1\,du + G_1\,dW_u] + G_1\,G_2\,t
\end{aligned}
$$

Considerando che le variabili aleatorie $X_i(0)$, F_i e G_i sono $\mathcal{F}_0$-misurabili, e quindi possono essere portate fuori dall'integrale stocastico, è facile convincersi (con un poco di pazienza) che la dimostrazione si basa sul fatto che

$$
t^2 = 2\int_0^t s\,ds, \qquad t\,W_t = \int_0^t s\,dW_s + \int_0^t W_s\,ds, \qquad \text{e infine che } W_t^2 = 2\int_0^t W_s\,dW_s + t.
$$

È interessante notare che nella seconda uguaglianza si è usata la formula di integrazione per parti (5.11) (nella dimostrazione della Proposizione 5.2) nel caso in cui l'integrando è una funzione deterministica e C^1. Va infine notato che per ottenere la formula di integrazione per parti si usa solo il fatto che il processo di Wiener ammette traiettorie continue. Invece l'ultima uguaglianza (si veda l'Esempio 5.1) vale in quanto la variazione quadratica del processo di Wiener tende a t (Lemma 3.2).

□

Siano f una funzione regolare di (t, x) e X un processo di Itô, con differenziale stocastico (5.2). Ci chiediamo quale sia il differenziale stocastico del processo $(Y_t = f(t, X_t))_t$. La risposta è fornita dalla formula di Itô.

Teorema 5.8 (Formula di Itô)**.** *Sia*

$$
dX_t = F_t dt + G_t dW_t,
$$

e sia $f(t, x)$ una funzione continua, con derivate parziali prime f_t ed f_x continue, e con la derivata parziale seconda f_{xx} continua, allora

$$
\begin{aligned}
dY_t = df(t, X_t) &= f_t(t, X_t)dt + f_x(t, X_t)dX_t + \frac{1}{2} f_{xx}(t, X_t)G_t^2 dt \\
&= f_t(t, X_t)dt + f_x(t, X_t)F_t dt + f_x(t, X_t)G_t dW_t + \frac{1}{2} f_{xx}(t, X_t)G_t^2 dt.
\end{aligned}
$$

In particolare per $X_t = W_t$ (ovvero per $F_t = 0$ e $G_t = 1$) si ha

$$
df(t, W_t) = f_t(t, W_t)dt + f_x(t, W_t)dW_t + \frac{1}{2} f_{xx}(t, W_t)dt.
$$

In realtà il Teorema 5.8 sopra enunciato non è completo in quanto mancano le ipotesi che permettono di assicurare che abbia senso, ad esempio, l'integrale stocastico di $f_x(t, X_t)G_t$ rispetto a dW_t. Un'ipotesi sufficiente è che le derivate siano limitate, tuttavia non è un'ipotesi necessaria.

Uno schema della dimostrazione si ottiene scrivendo per ogni partizione di $[t', t'']$

$$
f(t'', X_{t''}) - f(t', X_{t'}) = \sum_k \big[f(t_k, X_{t_k}) - f(t_{k-1}, X_{t_{k-1}})\big]
$$

e utilizzando la formula di Taylor[19]

$$
f(t, x) = f(t_0, x_0) + f_t(t_0, x_0)(t - t_0) + f_x(t_0, x_0)(x - x_0) + \frac{1}{2} f_{xx}(t_0, x_0)(x - x_0)^2 + o(t - t_0) + o\big((x - x_0)^2\big)
$$

e tenendo conto del fatto che

$$
\begin{aligned}
(X_{t_k} - X_{t_{k-1}})^2 &\simeq \big(F_{t_{k-1}}(t_k - t_{k-1}) + G_{t_{k-1}}(W_{t_k} - W_{t_{k-1}})\big)^2 \\
&= F_{t_{k-1}}^2(t_k - t_{k-1})^2 + 2F_{t_{k-1}}G_{t_{k-1}}(t_k - t_{k-1})(W_{t_k} - W_{t_{k-1}}) + G_{t_{k-1}}^2(W_{t_k} - W_{t_{k-1}})^2 \\
&= O\big((t_k - t_{k-1})^2\big) + o\big((t_k - t_{k-1})\big) + G_{t_{k-1}}^2(W_{t_k} - W_{t_{k-1}})^2
\end{aligned}
$$

[19]Veramente mancano i termini di ordine $O\big((t - t_0)(x - x_0)\big)$. Tuttavia successivamente, se si prende $t = t_k$, $t_0 = t_{k-1}$, $x = W_{t_k}$ e $x_0 = W_{t_{k-1}}$, si ha che $|(t - t_0)(x - x_0)| = |t_k - t_{k-1}| \cdot |W_{t_k} - W_{t_{k-1}}|$, che è un infinitesimo di ordine superiore a $|t_k - t_{k-1}|$ in quanto il processo di Wiener è a traiettorie continue. Questo è un punto fondamentale, perché mostra come per ottenere la formula di Itô serva la continuità delle traiettorie, oltre al fatto che il processo di Wiener sia una martingala con variazione quadratica uguale a t.

con $(W_{t_k} - W_{t_{k-1}})^2$ che si comporta come $(t_k - t_{k-1})$ (si riveda la sezione in cui si dimostra che il processo di Wiener non ha traiettorie a variazione limitata)

Esempio 5.2. *Sia $f(x)$ una funzione di classe C^1 e sia $F(x) = \int_0^x f(y)\,dy$ allora la formula di Itô, applicata alla funzione $F(x)$ nel caso in cui $X_t = W_t$ ci assicura che*

$$F(W_t) - F(W_s) = \int_s^t f(W_u)\,dW_u + \frac{1}{2}\int_s^t f'(W_u)\,du.$$

L'interesse di questa formula sta nel fatto che possiamo calcolare l'integrale stocastico $\int_s^t f(W_u)\,dW_u$ usando solo l'integrale ordinario, in quanto possiamo riscrive l'uguaglianza precedente come segue

$$\int_s^t f(W_u)\,dW_u = F(W_t) - F(W_s) - \frac{1}{2}\int_s^t f'(W_u)\,du.$$

Questa parte contiene un approfondimento non necessario per gli studenti del corso di Master

Un altro schema di dimostrazione della formula di Itô si basa sull'idea di dimostrare, attraverso la formula del differenziale del prodotto, che la formula di Itô vale per ogni $f(t,x) = x^n$, e quindi, per linearità, per i polinomi. Poi per ogni funzione del tipo $f(t,x) = a(t)\,p(x)$ con $a(\cdot)$ una funzione C^1 e $p(\cdot)$ un polinomio, e quindi per ogni funzione $f(t,x)$ che si possa ottenere come limite di funzioni del tipo $f_n(t,x) = \sum_{k=0}^{m_n} a_k(t)\,p_k(x)$. **Tali funzioni si possono ottenere, grazie al Teorema di Stone Weierstrass e i limiti sono uniformi sui compatti.

Il precedente Teorema 5.8 ammette diverse generalizzazioni, come ad esempio la seguente.

Teorema 5.9 (Formula di Itô multidimensionale)**.** *Siano X_i, $i = 1,...,m$, processi che ammettono differenziale stocastico*

$$dX_i(t) = F_i(t)dt + G_i(t)dW_t \qquad i = 1,...,m$$

e sia $f : \mathbb{R}^+ \times \mathbb{R}^m \to \mathbb{R}$ una funzione continua in (t,x), derivabile con continuità una volta in t e due volte in x. Allora posto $X_t = (X_1(t),...,X_m(t))$, il processo $(f(t,X_t))_t$ ammette differenziale stocastico

$$\begin{aligned} df(t,X_t) &= \left(f_t(t,X_t) + \sum_{i=1}^m f_{x_i}(t,X_t)F_i(t) + \frac{1}{2}\sum_{i,j=1}^m f_{x_i x_j}(t,X_t)G_i(t)G_j(t) \right) dt \\ &+ \sum_{i=1}^m f_{x_i}(t,X_t)G_i(t)dW_t \\ &= \sum_{i=1}^m f_{x_i}(t,X_t)dX_i(t) + \left(f_t(t,X_t) + \frac{1}{2}\sum_{i,j=1}^m f_{x_i x_j}(t,X_t)G_i(t)G_j(t) \right) dt. \end{aligned}$$

In presenza di un processo di Wiener a dimensione d, ossia in presenza di d processi di Wiener indipendenti $W_t^{(1)}$, ... ,$W_t^{(d)}$ si può considerare anche il caso in cui i processi X_i, $i = 1,...,m$ ammettono differenziale stocastico definito come segue

$$dX_i(t) = F_i(t)dt + \sum_{k=1}^d G_i^{(k)}(t)dW_t^{(k)} \qquad i = 1,...,m$$

Allora il processo $(f(t,X_t))_t$ ammette differenziale stocastico

$$df(t,X_t) = \left(f_t(t,X_t) + \sum_{i=1}^{m} f_{x_i}(t,X_t)F_i(t) + \frac{1}{2}\sum_{i,j=1}^{m} f_{x_i x_j}(t,X_t)\sum_{k=1}^{d} G_i^{(k)}(t)G_j^{(k)}(t) \right) dt + \sum_{i=1}^{m} f_{x_i}(t,X_t)G_i^{(k)}(t)dW_t^{(k)} \tag{5.14}$$

$$= {**}\sum_{i=1}^{m} f_{x_i}(t,X_t)dX_i(t) + \left(f_t(t,X_t) + \frac{1}{2}\sum_{i,j=1}^{m} f_{x_i x_j}(t,X_t)\sum_{k=1}^{d} G_i^{(k)}(t)G_j^{(k)}(t) \right) dt \tag{5.15}$$

$$= {**}\sum_{i=1}^{m} f_{x_i}(t,X_t)dX_i(t) + \left(f_t(t,X_t) + \frac{1}{2}\sum_{i,j=1}^{m} f_{x_i x_j}(t,X_t) {**} (G(t)'G(t))_{i,j} \right) dt. \tag{5.16}$$

dove abbiamo indicato con $G(t)'$ la matrice trasposta della matrice che al posto di indice (k,i) ha il valore $G_i^{(k)}(t)$.

Dimostrazione. Per la prima formula, nel caso di un unico processo di Wiener, una dimostrazione formale si basa sul considerare prima il caso dei polinomi, dimostrando la formula per induzione e utilizzando la formula del differenziale del prodotto. Per ottenere la formula nel caso generale si tratta di considerare ***(i)*** il fatto che i polinomi approssimano uniformemente sui compatti le funzioni regolari, insieme alle loro derivate, ***(ii)*** il fatto che ci si può ridurre al caso in cui X_t rimane in un compatto, per ogni t, introducendo opportuni tempi d'arresto $\tau_M = \inf\{s\, t.c.\, |X_s| \geq M\}$.

Per la seconda formula il procedimento è lo stesso, ma bisogna utilizzare una formula per il differenziale del prodotto nel caso in cui ci siano diversi processi di Wiener indipendenti: **se

$$dX_i(t) = F_i(t)dt + \sum_{k=1}^{d} G_i^{(k)}(t)dW_t^{(k)}, \qquad dX_j(t) = F_j(t)dt + \sum_{k=1}^{d} G_j^{(k)}(t)dW_t^{(k)},$$

allora

$$d(X_i\, X_j)(t) = X_i(t)dX_j(t) + X_j(t)dX_i(t) + \sum_{k=1}^{d} G_i^{(k)}(t)G_j^{(k)}(t)\, dt. \tag{5.17}$$

Si tratta di una generalizzazione della Proposizione 5.7. Anche in questo caso la dimostrazione è simile a quella nel caso unidimensionale, ma bisogna utilizzare il fatto che, **poiché per $h \neq k$ i processi di Wiener $(W_t^{(h)})$ e $(W_t^{(k)})$ sono indipendenti, allora

$$2\, W_t^{(h)}\, W_t^{(k)} = \int_0^t W_s^{(h)}\, dW_s^{(k)} + \int_0^t W_s^{(k)}\, dW_s^{(h)}.$$

□

Esercizio 5.1. *Dati due processi di Wiener $W_t^{(1)}$ e $W_t^{(2)}$ indipendenti, dimostrare che*

$$\tilde W_t^{(1)} = \frac{1}{\sqrt{2}}[W_t^{(1)} + W_t^{(2)}], \qquad \tilde W_t^{(2)} = \frac{1}{\sqrt{2}}[W_t^{(1)} - W_t^{(2)}],$$

sono processi di Wiener indipendenti. Dimostrare che

$$2\, W_t^{(1)}\, W_t^{(2)} = \int_0^t W_s^{(1)}\, dW_s^{(2)} + \int_0^t W_s^{(2)}\, dW_s^{(1)}, {**}$$

(suggerimento: si consideri $\int_0^t \tilde W_s^{(1)}\, d\tilde W_s^{(1)}$ e lo si calcoli utilizzando sia il fatto che è un processo di Wiener, sia riscrivendolo come somma di integrali stocastici)

Osservazione 5.5. *Se indichiamo con f' il gradiente di f rispetto a x. La formula di Itô si può scrivere in maniera più compatta:*

$$df(t,X_t) = f_t(t,X_t)dt + f'(t,X_t)dX(t) + \frac{1}{2}\sum_{i,j=1}^{m} f_{x_i x_j}(t,X_t)d < X_i, X_j >_t,$$

***dove $d < X_i, X_j >_t = \sum_{k=1}^{d} G_i^{(k)}(t)G_j^{(k)}(t)\, dt$ (che compare nella formula (5.17) del differenziale del prodotto di $X_i(t)$ con $X_j(t)$).*

Il differenziale stocastico si comporta, dunque, in maniera diversa da quello ordinario per la presenza dell'ultimo termine a destra.

5.2.1 Moto browniano geometrico e il suo differenziale stocastico

Ricordiamo che un ***moto browniano geometrico*** è un processo $(S_t)_t$ definito da

$$S_t = s_0 \exp\left\{\sigma W_t + \left(\mu - \frac{1}{2}\sigma^2\right) t\right\}$$

dove μ e σ sono costanti, con $\sigma \neq 0$.
Definiamo

$$f(t,x) = s_0 \exp\left\{\sigma x + \left(\mu - \frac{1}{2}\sigma^2\right) t\right\}$$

così

$$S_t = f(t, W_t).$$

Allora

$$f_t(t,x) = \left(\mu - \frac{1}{2}\sigma^2\right) f(t,x), \qquad f_x(t,x) = \sigma f(t,x), \qquad f_{xx}(t,x) = \sigma^2 f(t,x).$$

Per la formula di Itô, applicata a $X_t = W_t$ (così $F_t = F_t^W = 0$ e $G_t = G_t^W = 1$**)

$$\begin{aligned} dS_t = df(t,W_t) &= f_t(t,W_t)dt + f_x(t,W_t)dW_t + \frac{1}{2} f_{xx}(t,W_t)dt \\ &= \left(\mu - \frac{1}{2}\sigma^2\right) f(t,W_t)dt + \sigma f(t,W_t)dW_t + \frac{1}{2}\sigma^2 f(t,W_t)dt \\ &= \mu S_t dt + \sigma S_t dW_t. \end{aligned}$$

Così il moto browniano geometrico in forma differenziale è dato da

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{5.18}$$

e nella sua forma integrale

$$S_t = s_0 + \int_0^t \mu S_u du + \int_0^t \sigma S_u dW_u.$$

In particolare quindi

$$\begin{aligned} \mathbb{E}[S_t] &= s_0 + \mathbb{E}[\int_0^t \mu S_u du] + \mathbb{E}[\int_0^t \sigma S_u dW_u] \\ &= s_0 + \int_0^t \mu \mathbb{E}[S_u] du, \end{aligned}$$

da cui

$$\mathbb{E}[S_t] = s_0 \exp\{\mu t\}.$$

Riutilizzando ancora la formula di Itô per $X_t = S_t$ (così $F_t = F_t^S = \mu S_t$ e $G_t = G_t^S = \sigma S_t$) per $f(t,x) = x^2$ (per cui $f_t = 0$, $f_x = 2x$ ed $f_{xx} = 2$) si ottiene [20]

$$\begin{aligned} dS_t^2 &= 2S_t dS_t + \frac{1}{2} 2\sigma^2 S_t^2 dt \\ &= 2\mu S_t^2 dt + 2\sigma S_t^2 dW_t + \sigma^2 S_t^2 dt \\ &= (2\mu + \sigma^2) S_t^2 dt + 2\sigma S_t^2 dW_t. \end{aligned}$$

[20] Per ottenere il differenziale di S_t^2 si può considerare che

$$\begin{aligned} df(t,S_t) &= f_t(t,S_t)dt + f_x(t,S_t)dS_t + \frac{1}{2} f_{xx}(t,S_t)(G^S)_t^2 dt \\ &= 2S_t\left(\mu S_t dt + \sigma S_t dW_t\right) + \sigma^2 S_t^2 dt \end{aligned}$$

oppure si può considerare che $S_t^2 = \tilde{f}(t, W_t)$ con

$$\tilde{f}(t,x) = s_0^2 \exp\left\{2\sigma x + \left(2\mu - \sigma^2\right) t\right\}$$

e applicare la formula di Itô al processo $X_t = W_t$. Si consiglia il lettore di controllare che si ottiene lo stesso risultato con questo metodo.

Integrando quest'ultima espressione, e prendendo i valori attesi (tralasciando il problema di mostrare che l'integrale stocastico è effettivamente una martingala a media nulla) si ottiene che

$$\begin{aligned}\mathbb{E}[S_t^2] &= s_0^2 + \mathbb{E}[\int_0^t (2\mu+\sigma^2)S_u^2 du] + 2\sigma\mathbb{E}[\int_0^t S_u^2 dW_u] \\ &= s_0^2 + \int_0^t (2\mu+\sigma^2)\mathbb{E}[S_u^2]du,\end{aligned}$$

da cui

$$\mathbb{E}[S_t^2] = s_0^2 \exp\{(2\mu+\sigma^2)t\}$$

e quindi si può ricavare facilmente la varianza.[21] Il moto browniano geometrico è un processo da prendere in considerazione per modellizzare l'evoluzione di quantità che devono sempre rimanere positive. Anche per questo motivo, è un modello usato per descrivere l'evoluzione dei prezzi nei mercati finanziari. Nel modello di Black-Scholes, come già visto, il titolo rischioso segue un moto browniano geometrico.

5.3 Equazioni differenziali stocastiche

Introduciamo in questa sezione la nozione di equazione differenziale stocastica.
Siano $b(x,t) = (b_i(x,t))_{1\le i\le m}$ e $\sigma(x,t) = (\sigma_{ij}(x,t))_{1\le i\le m, 1\le j\le d}$ funzioni misurabili definite su $\mathbb{R}^m \times [0,T]$ a valori in $\mathbb{R}^m$ e in $M(m,d)$ rispettivamente, dove $M(m,d)$ indica l'insieme delle matrici $m\times d$.

Definizione 5.6. *Diremo che il processo $(\xi_t)_{t\in[u,T]}$ è soluzione dell'equazione differenziale stocastica*

$$\begin{cases} d\xi_t = b(\xi_t,t)dt + \sigma(\xi_t,t)dW_t \\ \xi_u = x \qquad x\in\mathbb{R}^m \end{cases} \tag{5.19}$$

se $(W_t)_t$ è un moto browniano d-dimensionale standard; per ogni $t\in[u,T]$ si ha

$$\xi_t = x + \int_u^t b(\xi_s,s)ds + \int_u^t \sigma(\xi_s,s)dW_s.$$

Osservazione 5.6. *Nella definizione precedente, si richiede implicitamente che $s\to b(\xi_s,s)$ e $s\to\sigma(\xi_s,s)$ siano processi per cui abbia senso fare i rispettivi integrali, deterministico e stocastico, insieme alle condizioni di misurabilità progressiva.*

Infine, ricordiamo che σ^2 viene chiamato ***coefficiente di diffusione*** e b viene chiamato ***drift*** (o ***deriva***).

Definizione 5.7. *Si dice che il sistema (5.19) ha* **soluzione forte** *se per ogni $(\Omega,\mathcal{F},\mathcal{F}_t,\mathbf{P})$ e per ogni moto browniano standard $W=(W_t)_t$, esiste un processo ξ tale che $(\xi_t)_t$ sia soluzione di (5.19).*

Con questa definizione, e tendendo conto della relazione (5.18), possiamo dire che il moto browniano geometrico è soluzione dell'equazione differenziale (con coefficienti $b(t,x)=\mu\, x$ e $\sigma(t,x)=\sigma\, x$)

$$d\xi_t = \mu\xi_t dt + \sigma\xi_t dW_t, \quad \xi_0 = s_0.$$

Per le equazioni differenziali stocastiche vale il seguente teorema di esistenza e unicità:

Teorema 5.10. *Supponiamo che i coefficienti $b(t,x)$ e $\sigma(t,x)$, oltre alla condizione di congiunta misurabilità, soddisfino la seguente condizione di Lipschitz globale:*

$$|b(x,t)-b(\bar{x},t)|^2 + |\sigma(x,t)-\sigma(\bar{x},t)|^2 \le K|x-\bar{x}|^2, \tag{5.20}$$

[21] Ovviamente la $Var(S_t)$ si può ricavare anche a direttamente, si consiglia il lettore di eseguire questo calcolo e controllare che i due metodi portano allo stesso risultato, come deve essere.

e la seguente condizione di crescita

$$|b(x,t)|^2 + |\sigma(x,t)|^2 \le K\left(1 + |x|^2\right), \tag{5.21}$$

per una costante $K > 0$.

Allora l'equazione differenziale stocastica

$$dX_t = b(X_t,t)dt + \sigma(X_t,t)dW_t, \quad t \ge s, \qquad X_s = x \tag{5.22}$$

ammette una soluzione che è un processo progressivamente misurabile, con $\mathbb{E}\big[\int_s^T |X_u|^2 du\big] < \infty$, per ogni $T \ge s$. Inoltre esiste una sola soluzione con queste caratteristiche, nel senso che, se $\bar{X}_t$ è un'altra soluzione, allora $\mathbb{P}\big(X_t = \bar{X}_t, \text{ per ogni } t\big) = 1$.

A titolo di esempio vediamo come si può dimostrare che il processo di Orstein-Ulhenbeck è soluzione dell'equazione differenziale

$$d\xi_t = -\lambda \xi_t dt + \sigma dW_t, \quad \xi_0 = x.$$

L'idea è quella di ripetere il metodo della variazione delle costanti per le equazioni differenziali ordinarie[22]

Si cerca quindi la soluzione del tipo $X_t = \eta_t \exp\{-\lambda t\}$, dove η_t è un processo da determinare che ammetta differenziale stocastico

$$d\eta_t = F_t dt + G_t dW_t, \quad \textit{ovvero} \quad \eta_t = x + \int_0^t F_s ds + \int_0^t G_s dW_s.$$

Il processo $\exp\{-\lambda t\}$ è deterministico, di conseguenza,

$$d(\exp\{-\lambda t\}) = -\lambda \exp\{-\lambda t\} dt$$

per la formula del prodotto dei differenziali stocastici[23] si ottiene che

$$\begin{aligned} dX_t = d(\exp\{-\lambda t\}\eta_t) &= \exp\{-\lambda t\} d\eta_t + \eta_t d(\exp\{-\lambda t\}) \\ &= \exp\{-\lambda t\}(F_t dt + G_t dW_t) - \lambda \exp\{-\lambda t\}\eta_t dt \\ &= \exp\{-\lambda t\} F_t dt + \exp\{-\lambda t\} G_t dW_t - \lambda X_t dt. \end{aligned}$$

Da ciò si ricava immediatamente che deve essere $F_t = 0$ ed $\exp\{-\lambda t\} G_t = \sigma$, ovvero

$$G_t = \sigma \exp\{\lambda t\}.$$

---

[22] Ricordiamo che per risolvere l'equazione differenziale ordinaria

$$\dot{y}_t = -\lambda y_t + v_t$$

si può considerare prima l'equazione

$$\dot{y}_t = -\lambda y_t,$$

la cui soluzione è data da

$$y_t^0 = C \exp\{-\lambda t\}.$$

Poi si cerca la soluzione y_t del tipo $y_t = C_t \exp\{-\lambda t\}$, da cui, essendo

$$\frac{d}{dt} y_t = \frac{d}{dt}\Big(C_t \exp\{-\lambda t\}\Big) = -\lambda \exp\{-\lambda t\} C_t + \dot{C}_t \exp\{-\lambda t\} = -\lambda y_t + \dot{C}_t \exp\{-\lambda t\},$$

si ottiene che deve essere necessariamente

$$\dot{C}_t \exp\{-\lambda t\} = v_t.$$

ovvero

$$C_t = C_0 + \int_0^t \exp\{\lambda s\} v_s \, ds.$$

In conclusione la soluzione è data da

$$y_t = \exp\{-\lambda t\}\left(C_0 + \int_0^t \exp\{\lambda s\} v_s \, ds\right).$$

[23] Il motivo per cui la tecnica della variazone delle costanti funziona anche nel caso stocastico sta nel fatto che, essendo la funzione $e^{-\lambda t}$ deterministica, che ha come differenziale $de^{-\lambda t} = -\lambda e^{-\lambda t} dt + 0 \, dW_t$, la formula del differenziale stocastico del prodotto è la stessa che nel caso dei differenziali usuali

$$d(\eta_t e^{-\lambda t}) = \eta_t de^{-\lambda t} + e^{-\lambda t} d\eta_t + G_t \cdot 0 \, dt.$$

Da quest'ultima espressione si ricava immediatamente che

$$\eta_t = x + \int_0^t \sigma \exp\{\lambda s\} dW_s,$$

e quindi

$$X_t = \exp\{-\lambda t\}\eta_t = \exp\{-\lambda t\} \left(x + \int_0^t \sigma \exp\{\lambda s\} dW_s \right),$$

ovvero il processo di Orstein-Ulhenbeck.

Una volta ottenuta la soluzione è facile verificare che è effettivamente una soluzione, riapplicando la formula del differenziale stocastico del prodotto.[24]

Come applicazione delle proprietà dell'integrale stocastico si vede quindi subito come si possono calcolare valore medio e varianza del processo X_t di Orstein-Ulhenbeck. Infatti

$$\mathbb{E}[X_t] = \exp\{-\lambda t\}x + \exp\{-\lambda t\}\mathbb{E}\left[\int_0^t \sigma \exp\{\lambda s\} dW_s \right] = \exp\{-\lambda t\}x,$$

dove l'ultima uguaglianza vale in quanto l'integrale stocastico ha valore atteso nullo.

Da ciò si può ricavare anche la varianza, in quanto

$$(X_t - \mathbb{E}[X_t])^2 = (X_t - \exp\{-\lambda t\}x)^2 = \exp\{-2\lambda t\} \left(\int_0^t \sigma \exp\{\lambda s\} dW_s \right)^2,$$

e quindi, passando ai valori attesi,

$$\begin{aligned} Var(X_t) &= \exp\{-2\lambda t\}\mathbb{E}\left[\left(\int_0^t \sigma \exp\{\lambda s\} dW_s \right)^2 \right] \\ &= \mathbb{E}\left[\exp\{-2\lambda t\} \int_0^t \sigma^2 \exp\{2\lambda s\} ds \right] = \frac{\sigma^2}{2\lambda} \left(1 - \exp\{-2\lambda t\}\right). \end{aligned}$$

Anche la covarianza si può ricavare facilmente osservando che

$$\begin{aligned} &(X_t - \mathbb{E}[X_t])\,(X_{t'} - \mathbb{E}[X_{t'}]) \\ &= \exp\{-\lambda t\} \left(\int_0^t \sigma \exp\{\lambda s\} dW_s \right) \exp\{-\lambda t'\} \left(\int_0^{t'} \sigma \exp\{\lambda s\} dW_s \right) \\ &= \exp\{-\lambda (t + t')\} M_t M_{t'}, \end{aligned}$$

dove $M_t = \int_0^t \sigma \exp\{\lambda s\} dW_s$. Poiché, per ogni $\mathcal{G}_t$-martingala di quadrato integrabile, se $t \leq t'$, si ha

$$\mathbb{E}[M_t M_{t'}] = \mathbb{E}[M_t \mathbb{E}[M_{t'}|\mathcal{G}_t]] = \mathbb{E}[M_t M_t] = \mathbb{E}[M_t^2].$$

si ottiene che, per $t \leq t'$,

$$Cov(X_t, X_{t'}) = \exp\{-\lambda (t + t')\}\mathbb{E}\left[\int_0^t \sigma^2 \exp\{2\lambda s\} ds \right].$$

---

[24] Ovvero si tratta di ripercorrere i passi precedenti al contrario:

$$d(\exp\{-\lambda t\}) = -\lambda \exp\{-\lambda t\} dt,$$

$$d\left(x + \int_0^t \sigma \exp\{\lambda s\} dW_s \right) = 0 \cdot dt + \sigma \exp\{\lambda t\} dW_t.$$

quindi se

$$X_t = \exp\{-\lambda t\} \left(x + \int_0^t \sigma \exp\{\lambda s\} dW_s \right),$$

allora

$$dX_t = \left(x + \int_0^t \sigma \exp\{\lambda s\} dW_s \right) (-\lambda \exp\{-\lambda t\} dt) + \exp\{-\lambda t\}\sigma \exp\{\lambda t\} dW_t = -\lambda X_t + \sigma dW_t.$$

5.4 Il processo di Cox-Ingersoll-Ross

In questa sezione analizzeremo una particolare equazione differenziale stocastica (EDS) che svolge un ruolo fondamentale in finanza: nel modello di Cox-Ingersoll-Ross viene usata per descrivere la dinamica dei tassi d'interesse. L'EDS è la seguente:

$$dX_t = (a - bX_t)dt + \sigma\sqrt{X_t}dW_t \tag{5.23}$$

dove σ e a sono coefficienti non negativi mentre $b \in \mathbb{R}$. In questo caso non possiamo applicare il Teorema 5.10 di esistenza ed unicità in quanto la funzione radice quadrata, definita solo su $\mathbb{R}_+$, non è lipschtziana, infatti, presenta un punto di singolarità in 0. Ciò nonostante, possiamo dimostrare l'esistenza di una soluzione unica che sotto opportune condizioni è anche strettamente positiva. A tale scopo procederemo come segue: osserveremo che le dinamiche della forma (5.23) possono essere ottenute attraverso una semplice trasformazione di un opportuno processo $n-$dimensionale. In questo modo otterremo esplicitamente una soluzione debole[25] per particolari valori di a e σ.
In realtà vale un risultato più forte e cioè l're e l'unicità di una soluzione forte (vedremo solo l'idea del motivo)**.

Sia $(W_1, W_2, ..., W_n)$ un vettore $n-$dimensionale di moti browniani indipendenti e siano, per $i = 1, ..n$, $X_i = (X_i(t))_{t\geq 0}$ processi soluzione dell'EDS

$$dX_i(t) = -\frac{1}{2}bX_i(t)dt + \frac{1}{2}\sigma dW_i(t)dt$$

con b e σ due costanti. Come sappiamo che i processi $(X_i(t))_{t\geq 0}$, al variare di $i \in \{1, ..., n\}$ con $n \geq 2$, sono processi di Ornstein Ulhenbeck unidimensionali. Poniamo $r(t) = X_1^2(t) + ... + X_n^2(t)$ e consideriamo il processo $B = (B(t))_{t\geq 0}$, definito da:

$$B(0) = 0 \qquad dB(t) = \sum_{i=1}^{n} \frac{X_i(t)dW_i(t)}{\sqrt{r(t)}},$$

cioè:

$$B(t) = \sum_{i=1}^{n} B_i(t) \text{ con } B_i(t) = \int_0^t \frac{X_i(s,\omega)dW_i(s)}{\sqrt{r(s)}}.$$

Il processo B così definito è un moto browniano (ossia un proceso di Wiener standard).

Per dimostrare che B è un moto browniano abbiamo bisogno del Teorema di caratterizzazione di Levy, che riportiamo senza dimostrazione.

Teorema 5.11 (Teorema di caratterizzazione di Levy)**.** *Sia $(\Omega, \mathcal{F}, P)$ uno spazio di probabilità su cui è definito un processo stocastico $X = (X_t)_{t\geq 0}$. Tale processo risulta essere un moto browniano se:*

1 *il processo X è a traiettorie continue;*

2 *il processo X è una martingala rispetto alla sua filtrazione naturale;*

3 *$X_t^2 - t$ è una martingala rispetto alla sua filtrazione naturale.*

Nel nostro caso si tratta di verificare le seguenti condizioni:

1 Il processo B è a traiettorie continue per come è definito;

2 Il processo B è una martingala, infatti, risulta che $B(t) = \sum_{i=1}^n B_i(t)$ dove i processi $B_i(t)$ sono martingale. Per verificare questa porprietà basta notare che, condizione sufficiente affinché un processo $I_i(t)$, $t \in [0,T]$ definito da

$$I_i(t) = \int_0^t \Psi_i(s,\omega)dW_i(s)$$

[25] Ricordiamo che per una equazione differenziale stocastica

$$dX_t = \mu(t, X_t)\,dt + \sigma(t, X_t)\,dW_t$$

ha soluzione è forte se (comunque) dato uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ e un processo di Wiener W_t su tale spazio si trova un processo X_t che soddisfa la precedente equazione rispetto a quel processo di Wiener assegnato, mentre si dice che ha soluzione debole se esiste uno uno spazio di probabilità $(\Omega', \mathcal{F}', \mathbb{P}')$, un processo di Wiener W_t' su tale spazio ed un processo X_t' per cui vale

$$dX_t' = \mu(t, X_t')\,dt + \sigma(t, X_t')\,dW_t'.$$

La differenza tra i due concetti di soluzione si capisce meglio dopo aver visto questo esempio. In questo caso si costruisce un processo soluzione, con un processo di Wiener opportuno, non con un processo assegnato a priori.

con Ψ_i progressivamente misurabile[26], sia una martingala è che valga

$$\mathbb{E}\big(\int_0^T |\Psi_i(s,\omega)|^2 ds\big) < \infty. \tag{5.24}$$

Nel nostro caso, si ha $\Psi_i(s,\omega) = \frac{X_i(s,\omega)}{\sqrt{r(s)}}$ e poiché $|\Psi_i(s,\omega)|^2 = \frac{X_k^2}{\sum_{k=1}^n X_i^2} < 1$, la condizione (5.24) è verificata, perciò $B_i(t)$ è una martingala. Di conseguenza $B(t) = \sum_{i=1}^n B_i(t)$ è una martingala in quanto somma di martingale.

3 Il processo $B^2(t) - t$ è una martingala in quanto, per ogni $i = 1, \cdots, n$

$$B_i^2(t) - \int_0^t \frac{X_i^2}{\sum_{k=1}^n X_k^2} ds$$

è una martingala, inoltre anche $B_i(t)B_j(t)$ è una martingala[27]per $i \neq j$, e infine poiché

$$\begin{aligned}
&\sum_{i=1}^n [B_i^2(t) - \int_0^t \frac{X_i^2}{\sum_{k=1}^n X_k^2} ds] + \sum_{i\neq j} B_i(t)B_j(t)\\
&= \sum_{i=1}^n B_i^2(t) + \sum_{i\neq j} B_i(t)B_j(t) - \sum_{i=1}^n \int_0^t \frac{X_i^2}{\sum_{k=1}^n X_k^2} ds]\\
&= B^2(t) - t,
\end{aligned}$$

segue che $B^2(t) - t$ è una martingala in quanto somma di martingale.

Abbiamo provato, dunque, che il processo B è un moto browniano.

Calcoliamo adesso il differenziale stocastico di

$$r(t) = f\big(X_1(t), \dots, X_n(t)\big) = X_1^2(t) + ... + X_n^2(t).$$

Dalla formula di Ito (5.15), con $m = d = n$, tenendo conto che $G_i^{(k)}(t) = \sigma/2$ per $k = i$ e $G_i^{(k)}(t) = 0$ per $k \neq i$, e che

$$f_{x_i}(x_1, \dots, x_n) = 2\, x_i, \quad f_{x_i,x_j}(x_1, \dots, x_n) = 2,\ i = j \quad f_{x_i,x_j}(x_1, \dots, x_n) = 0,\ i \neq j$$

abbiamo che

$$dr(t) = \sum_{i=1}^n (2X_i(t)dX_i(t) + \frac{\sigma^2}{4} dt).$$

Ricordando che $dX_i(t) = -(b/2)\, X_i(t)dt + (\sigma/2)\, dW_i(t)$ e sostituendo abbiamo che:

$$dr(t) = \bigg(\frac{n^2\sigma^2}{4} - b\, r(t)\bigg)dt + \sigma\sqrt{r(t)}dB(t).$$

Infatti

$$dr(t) = \sum_{i=1}^n f_{x_i}(X_t)dX_i(t) + \frac{1}{2}\sum_{i,j=1}^n f_{x_i x_j}(t, X_t) \sum_{k=1}^n G_i^{(k)}(t) G_j^{(k)}(t)\, dt$$

---

[26]Un processo $(\xi_t)_{t\in[0,T]}$ dove (se $T = +\infty$ allora si considera la semiretta $[0, +\infty)$) è detto progressivamente misurabile se $\forall u \in [0, T]$ l'applicazione $(t, \omega) \to \xi_t(\omega)$ è misurabile da $([0, u] \times \Omega, \mathcal{B}([0, u]) \times \mathcal{F}_u)$ in $(E, \mathcal{B}(E))$, avendo indicato con E lo spazio degli stati e con $\mathcal{B}(E)$ la $\sigma-$algebra boreliana su E.

[27]Osserviamo che $B^2(t) = \sum_{i=1}^n B_i^2(t) + \sum_{i\neq j} B_i(t)B_j(t)$. In generale, se $f_i(t)$ è un opportuno processo tale che l'integrale stocastico $\int_0^t f_i(s)dW_i(s)$ abbia senso, per $i = 1, \cdots, n$, e considerando che (W_i, W_j) è un processo di Wiener bidimensionale, si ha che

$$\bigg(\int_0^t f_i(s)dW_i(s)\bigg)\bigg(\int_0^t f_j(s)dW_j(s)\bigg) - \int_0^t f_i(s)f_j(s)d < W_i, W_j >_s$$

è una martingala. Nel nostro caso, $f_i(t) = \frac{X_i(t)}{\sqrt{r(t)}}$ e quindi, per $i \neq j$

$$B_i(t)B_j(t) - \int_0^t \frac{X_i(s)}{\sqrt{r(s)}}\frac{X_j(s)}{\sqrt{r(s)}} d < W_i, W_j >_s = B_i(t)B_j(t)$$

in quanto $\int_0^t \frac{X_i(s)}{\sqrt{r(s)}}\frac{X_j(s)}{\sqrt{r(s)}} d < W_i(s), W_j(s) >= 0$ essendo i moti browniani indipendenti. Perciò, per quanto detto prima, $B_i(t)B_j(t)$ è una martingala al variare di t e di i e j, sotto la condizione $i \neq j$.

(nel nostro caso $f_{x_i x_j} = 0$ per $i \neq j$ e $G_i^k = 0$ per $k \neq i$)

$$
\begin{aligned}
&= \sum_{i=1}^{n} 2\, X_i(t)\, dX_i(t) + \frac{1}{2} \sum_{i=1}^{n} 2\, \frac{\sigma^2}{4}\, dt = \sum_{i=1}^{n} 2\, X_i(t) \left(-\frac{b}{2}\, X_i(t)\, dt + \frac{\sigma}{2}\, dW_i(t) \right) + \frac{n\,\sigma^2}{4}\, dt \\
&= \frac{n\,\sigma^2}{4}\, dt - b \sum_{i=1}^{n} X_i^2(t)\, dt + \sigma \sum_{i=1}^{n} X_i(t)\, dW_i(t) = \frac{n\,\sigma^2}{4}\, dt - b\, r(t)\, dt + \sigma\, \sqrt{r(t)} \sum_{i=1}^{n} \frac{X_i(t)}{\sqrt{r(t)}}\, dW_i(t) \\
&= \left(\frac{n\sigma^2}{4} - b\, r(t) \right) dt + \sigma \sqrt{r(t)}\, dB(t).
\end{aligned}
$$

In conclusione, per opportuni valori di a e σ, più precisamente, per $a = \frac{n\sigma^2}{4}$, abbiamo ottenuto una soluzione debole $r(t) = X_1^2(t) + ... + X_n^2(t)$ dell'equazione (5.23) che è anche strettamente positiva.

Come già anticipato, vale un risultato più forte e cioè che se $a \geq \frac{\sigma^2}{2}$ esiste ed è unica la soluzione globale (cioè per ogni $t \geq 0$) in senso forte di tale equazione, e inoltre, la soluzione è anche strettamente positiva. **Per capire il motivo di tale risultato, basta considerare che, finché il valore della soluzione si mantiene lontano dallo 0, l'equazione considerata è una EDS con coefficienti lipschitziani (cioè $\sigma\, \sqrt{x}$ è una funzione di Lipschitz in ogni intervallo del tipo $[r_1, \infty)$, per ogni $r_1 > 0$) e quindi si ha esistenza e unicità della soluzione forte, fino a che la soluzione si mantiene lontana dallo 0. Quando invece la soluzione si avvicina allo 0, il coefficiente di diffusione tende a diventare nullo, e se i coefficiente $a > 0$ del drift (cioè in $a - b\, x$) è sufficientemente grande, allora la soluzione tende ad alontanarsi dallo zero, e il rumore dovuto alla diffusione non riesce a far oltrepassare lo 0, e la soluzione rimane sempre strettamente positiva e quinid esiste per tutti i tempi.**

Capitolo 6

Il modello di Black e Scholes

All'inizio degli anni $'70$, Fischer Black, Myron Scholes e Robert Merton danno un fondamentale contributo alla teoria di valutazione delle opzioni, sviluppando il modello Black e Scholes (presentato nel loro articolo del 1973 [7]). Questo modello ha un'enorme influenza sul modo di determinare il valore di una opzione e di una qualsiasi attività derivata. La sua importanza fu riconosciuta con l'assegnazione del premio Nobel per l'economia a Myron Scholes e Robert Merton. Purtroppo, Fischer Black muore nel 1995, altrimenti il premio sarebbe assegnato, senza dubbio, anche a lui. In questo capitolo vedremo come si ricava la dimostrazione della famosa formula per il calcolo del prezzo delle opzioni nel modello di Black e Scholes. È bene ricordare che esistono diverse strade per ottenere l'equazione di valutazione di Black e Scholes; abbiamo scelto quella che più si avvicina al loro procedimento originario.
Considereremo una opzione *call* scritta su S_t e con prezzo di esercizio K a tempo di esercizio T. Per determinare il premio facciamo le seguenti ipotesi di partenza:

1 L'azione evolve continuamente seguendo un moto browniano geometrico, con parametri μ e σ costanti, e non paga dividendi durante la vita dell'opzione. È possibile vendere allo scoperto azioni ed utilizzare completamente i proventi.

2 Il tasso a cui si può prendere a prestito ed investire è lo stesso ed è una costante uguale ad r durante la vita dell'opzione.

3 Il mercato in cui si opera è privo di attriti: non vi sono costi di transazione, né tasse.

4 Il valore dell'opzione dipende solo dalla stessa fonte di incertezza da cui dipende l'azione (il processo W_t) e dal tempo.

Siano:

- t l'epoca corrente,
- T l'epoca finale di scadenza dell'opzione,
- S_t il prezzo corrente dell'azione, descritto dal *moto browniano geometrico* (**??**), cioè dal processo

  $$dS_t = \mu S_t dt + \sigma S_t dW_t, \qquad S_0 = x$$

  la cui soluzione è data da

  $$S_t = x\, e^{\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W_t} = S_0\, e^{\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W_t} \tag{6.1}$$

  dove μ e σ sono due costanti e $(W_t)_t$ è un moto browniano standard.

- B_t il prezzo di un titolo non rischioso (conto bancario o titolo obbligazionario) descritto da

  $$dB_t = rB_t dt, \qquad B_0 = 1$$

con soluzione

$$B_t = B_0 e^{rt} = e^{rt} \tag{6.2}$$

dove il tasso di interesse r è una costante positiva.

Nel continuo le variazioni percentuali di prezzo possono esprimersi come $\ln(\frac{S_T}{S_t})$, la distribuzione dei rendimenti è dunque log-normale in quanto $\frac{S_T}{S_t} = e^{\mu - \frac{1}{2}\sigma^2(T-t)+\sigma(W_T - W_t)}$ è una variabile aleatoria con distribuzione log-normale.

Una opzione *call europea* con prezzo di esercizio K e tempo di maturità T ha funzione di payoff f_T pari a

$$f_T = \max(S_T - K, 0) = (S_T - K)^+$$

per cui il prezzo al tempo T è pari a $Y_T = (S_T - K)^+$.
Il problema che si pongono Black e Scholes è di determinare il prezzo dell'opzione call al tempo t, per $t \in [0, T]$; ciò a cui arrivano è la validità della seguente relazione per determinare il prezzo Y_t di una call europea al tempo t

$$Y_t = S_t \Phi\big(\widetilde{z}(T-t, S_t)\big) - e^{-r(T-t)}\,\Phi\Big(\widetilde{z}(T-t, S_t) - \sigma\sqrt{T-t}\Big)^{**}$$

dove Φ è la funzione di distribuzione di una gaussiana standard e

$$\widetilde{z} = \widetilde{z}(\tau, x) = \frac{\ln(\frac{x}{K}) + (r + \frac{\sigma^2}{2})\tau}{\sigma\sqrt{\tau}} = \frac{\ln(\frac{x}{K\,e^{r\,\tau}})}{\sigma\sqrt{\tau}} + \frac{\sigma\sqrt{\tau}}{2}.^{**} \tag{6.3}$$

Iniziamo con l'analizzare un risultato di natura più generale per spiegare il legame tra concetto di copertura perfetta e condizione di assenza di arbitraggio (si veda il Teorema 6.1).
Ricordiamo che $(a, b) = (a_t, b_t)_{0 \le t \le T}$ è una **strategia**, **se i processi $(a_t)_{0\le t\le T}$ e $(b_t)_{0\le t\le T}$ sono predicibili[1] e il valore del portfolio al tempo t è dato da

$$V_t^{(a,b)} = a_t S_t + b_t B_t \qquad \text{per ogni } t \in [0, T],$$

inoltre $(a, b) = (a_t, b_t)_{0\le t\le T}$ è una **strategia di copertura perfetta**, se oltre ad essere una strategia vale anche **

$$V_T = (S_T - K)^+.$$

Infine ricordiamo che (a_t, b_t) è **autofinanziante**, se e soltanto se per ogni t

$$V_t^{(a,b)} = a_t S_t + b_t B_t = a_0 S_0 + b_0 B_0 + \int_0^t a_\tau dS_\tau + \int_0^t b_\tau dB_\tau. \tag{6.4}$$

Ciò significa che nuovi acquisti di azioni saranno finanziati da vendite di titoli non rischiosi così come i proventi di vendite di azioni saranno reinvestiti in titoli non rischiosi.
Il seguente Teorema 6.1 ci garantisce che $Y_t = V_t^{(a,b)}$ per una tale strategia (a, b).

A questo punto dobbiamo dimostrare che il prezzo di una call europea è dato da $Y_t = C(t, S_t)$, dove

$$C(t, x) = x\Phi\big(\widetilde{z}(T-t)\big) - e^{-r(T-t)}K\,\Phi\big(\widetilde{z}(T-t) - \sigma\sqrt{T-t}\big) \tag{6.5}$$

La dimostrazione è strutturata nel seguente modo:

1 Se si ipotizza che $Y_t = C(t, S_t)$, con $C(t, x)$ sufficientemente regolare, allora da una parte $dY_t = dC(t, S_t)$, e dall'altra $dY_t = dV_t^{(a,b)}$. Il confronto tra i due diversi modi di esprimere dY_t porta a mostrare che $C(t, x)$ soddisfa l'equazione di Black e Scholes $\big($si veda la successiva (6.11)$\big)$; inoltre otteniamo che la strategia di copertura perfetta (a_t, b_t) è univocamente determinata (si veda il Teorema 6.2).

[1] **Il concetto di processo predicibile rispetto ad una filtrazione $(\mathcal{F}_t)_t$ è una generalizzazione del concetto di processo $\mathcal{F}_t$-adattato e continuo da sinistra, parlando in modo lasco, si può dire che un processo è predicibile rispetto ad una filtrazione $(\mathcal{F}_t)_t$, se è limite di processi $\mathcal{F}_t$-adattati e continui da sinistra.

2 La funzione $C(t,x) = x\,\Phi\big(\widetilde{z}(T-t)\big) - e^{-r(T-t)}K\Phi\big(\widetilde{z}(T-t) - \sigma\sqrt{T-t}\big)$ è soluzione dell'equazione di Black e Scholes (6.11) con la condizione finale $C(T,x) = (x-K)^+$ (questa verifica richiede un po' di pazienza e viene tralasciata).**

3 La strategia (a_t, b_t) definita dalle (6.12) e (6.13) è effettivamente una strategia di copertura perfetta, il cui valore è pari a $a_tS_t + b_tB_t = C(t,S_t)$.
Quindi, tenendo conto del Teorema 6.1, $C(t,S_t)$ è il prezzo dell'opzione in assenza di opportunità di arbitraggio.

6.1 Copertura perfetta e non arbitraggio

Enunciamo e dimostriamo l'anticipato Teorema dell'arbitraggio che mostra come l'assunzione di assenza di opportunità di arbitraggio (non si possono ottenere profitti senza sottoporsi a dei rischi) gioca un ruolo di fondamentale importanza nel problema del prezzaggio di un'opzione.

Teorema 6.1 (Teorema dell'arbitraggio)**.** *Sia (a_t, b_t) una strategia autofinanziante, che replichi il payoff terminale dell'opzione europea al tempo T:*

$$a_TS_T + b_TB_T = f_T, **$$

*dove f_T è una variabile aleatoria che dipende solo dal passato (cioè $\mathcal{F}_T$-misurabile).***

Se il mercato è privo di opportunità di arbitraggio, per ogni $t \in [0,T)$, allora il prezzo dell'opzione al tempo t è

$$Y_t = a_tS_t + b_tB_t. \tag{6.6}$$

ed in particolare deve valere anche

$$Y_T = a_TS_T + b_TB_T = f_T, **$$

Dimostrazione
Se

$$a_tS_t + b_tB_t < Y_t$$

allora il venditore, al tempo t, venderà l'opzione al prezzo Y_t costruendo, così, un portfolio $a_tS_t + b_tB_t$ e continuando con la strategia $(a_u, b_u)_{t\le u\le T}$; al tempo T venderà il portfolio al prezzo $a_TS_T + b_TB_T = Y_T$.
L'investimento del guadagno iniziale $\Delta_t = Y_t - a_tS_t - b_tB_t > 0$ nell'azione non rischiosa implica un guadagno pari a $e^{r(T-t)}\Delta_t$ al tempo T.
Quindi c'è opportunità di arbitraggio.

Se invece

$$a_tS_t + b_tB_t > Y_t$$

allora comprando un'opzione al tempo t e vendendo il portfolio al prezzo $a_tS_t + b_tB_t$ si otterrà al tempo T (senza alcun rischio) il guadagno pari a $\Big(a_tS_t + b_tB_t - Y_t\Big)e^{r(T-t)}$. Questo implica che c'è ancora arbitraggio.

Osservazione 6.1. *Tale teorema è valido in generale, per ogni payoff terminale $Y_T = f_T$; non è, inoltre, necessario che si tratti di una opzione europea, anche se in questo capitolo vogliamo analizzare il caso specifico di una call europea. È bene osservare che nel Teorema 6.1 non si è utilizzata la forma specifica dell'evoluzione di S_t, e che l'ipotesi di tasso r_t costante non è fondamentale: si tratta solo di sostituire $e^{r(T-t)}$ con $\frac{B_T}{B_t}$. Il Teorema 6.1 è quindi valido anche in mercati (B,S) di tipo più generale.*

6.2 L'equazione di Black e Scholes

**Iniziamo con una rapida rassegna dei modelli precursori il modello di Black e Scholes: tale modello ed in particolare **la formula di Black e Scholes è preceduta da numerosi modelli per la valutazione delle opzioni, fra questi ricordiamo l'approccio di Bachelier che è riportato nella sua tesi di dottorato discussa all'università Sorbonne di Parigi, Bachelier (1900) [1]. Tale modello (lineare) prevede che il prezzo del bene sottostante l'opzione segua un processo del tipo

$$S_t = S_0 + \mu t + \sigma W t \tag{6.7}$$

Da tale processo segue una formula di valutazione per un'opzione call europea che scritta con la solita terminologia, ossia indicando con C il prezzo dell'opzione, con $S = S_t$ il prezzo dell'azione al tempo corrente t e con $\tau = T - t$ la differenza tra l'epoca finale di scadenza dell'opzione e l'epoca corrente, diviene

$$C = (S - K)\Phi\Big(\frac{S-K}{\sigma\sqrt{\tau}}\Big) + \sigma\sqrt{\tau}\Phi'\Big(\frac{S-K}{\sigma\sqrt{\tau}}\Big) \tag{6.8}$$

L'inconveniente più grande è dovuto al fatto che l'equazione che descrive la dinamica del prezzo (6.7) consente al prezzo del bene sottostante di assumere valori negativi. Oltre mezzo secolo dopo la dissertazione di Bachelier, Sprenkle (1961) [18] propone un modello basato sulle ipotesi di neutralità al rischio e di distribuzione log-normale per i prezzi del bene sottostante. Tale modello, riportato nella successiva equazione (6.9), manifesta evidenti analogie con la formula di Black e Scholes

$$C = Se^{\rho\tau}\Phi(g_1) - (1-z)K\Phi(g_2) \tag{6.9}$$

dove:

- $g_1 = \frac{\ln(\frac{S}{K}) + (\rho + \frac{\sigma^2}{2})\tau}{\sigma\sqrt{\tau}}$,
- $g_2 = \frac{\ln(\frac{S}{K}) + (\rho - \frac{\sigma^2}{2})\tau}{\sigma\sqrt{\tau}}$,
- ρ è il tasso di crescita medio del prezzo del bene sottostante,
- z è il grado di avversione al rischio.

Boness (1964) [8] migliora il modello di Sprenkle introducendo il valore temporale della moneta; infatti, il valore finale del payoff viene attualizzato al tasso ρ. Introducendo alcune ipotesi molto vicine a quelle utilizzate da Black e Scholes, Boness propone la seguente formula di valutazione

$$C = S\Phi(g_1) - Ke^{-\rho\tau}\Phi(g_2). \tag{6.10}$$

dove ρ è il tasso di interesse atteso per un'azione.
È di particolare importanza il contributo di Samuelson (1965)[16], il quale estende il modello (6.10) permettendo all'opzione di avere una diversa forma di rischio rispetto al bene sottostante

$$C = e^{(\rho-w)\tau}S\Phi(g_1) - Ke^{-w\tau}\Phi(g_2)$$

dove

- w rappresenta il tasso di crescita del prezzo dell'opzione;
- ρ la crescita media del prezzo del bene sottostante.

È bene osservare che tutte le formule precedenti prevedono la stima non del tutto agevole di uno o più parametri e dipendono dalle preferenze dell'investitore verso il rischio o dal tasso di crescita del bene sottostante l'opzione (o dell'opzione stessa). Continuando con l'elenco dei contributi scientifici che anticipano il modello Black e Scholes vale la pena ricordare l'articolo Samuelson e Merton (1969) [15] in cui è proposta una formula di valutazione che dipende dalla funzione di utilità dell'investitore e quello di Torp e Kassouf (1971) [20] dove si studia il rapporto di copertura per la costruzione di un portfolio (protetto dal rischio di prezzo) senza riuscire a stabilire che il tasso di rendimento di tale portfolio deve coincidere con il tasso di rendimento delle attività non rischiose.

Continuiamo la nostra analisi sul modello di Black e Scholes e passiamo a considerare il seguente

Teorema 6.2. *Consideriamo le ipotesi del precedente Teorema 6.1 e assumiamo che*

$$Y_t = C(t, S_t)$$

dove $C(t,x)$, con $x \in (0,\infty)$ e $t \geq 0$, è regolare, ovvero le derivate parziali

$$C_t = \frac{\partial C}{\partial t}, \quad C_x = \frac{\partial C}{\partial x}, \quad C_{xx} = \frac{\partial^2 C}{\partial x^2},$$

sono continue.
Allora $C(t,x)$ soddisfa l'equazione di Black e Scholes

$$C_t(t,x) + rC_x(t,x)x - rC(t,x) + \frac{1}{2}C_{xx}(t,x)\sigma^2 x^2 = 0, \tag{6.11}$$

e una strategia di copertura perfetta è data da

$$a_t = C_x(t, S_t) \tag{6.12}$$

$$b_t = \frac{C(t,S_t) - C_x(t,S_t)S_t}{B_t} \tag{6.13}$$

Osservazione 6.2. ***Si noti l'analogia tra la relazione $a_t = C_x(t, S_t)$ e la relazione trovata precedentemente per il modello binomiale multiperiodale in cui avevamo ottenuto che*

$$\tilde{\gamma}_n(\omega) = \frac{c(n, S_{n-1}(\omega)\, u) - c(n, S_{n-1}(\omega)\, d)}{S_{n-1}(\omega)u - S_{n-1}(\omega)d}.$$

Inoltre si noti che la relazione $b_t = \frac{C(t,S_t)-C_x(t,S_t)S_t}{B_t}$ si può leggere dicendo che una volta comprati $a_t = C_x(t, S_t)$ titoli rischiosi al prezzo S_t, cioè una volta investito $C_x(t, S_t)S_t$ nel titolo rischioso, il resto del nostro capitale, cioè $C(t, S_t) - C_x(t, S_t)S_t$ viene investito nel titolo non rischioso.

Dimostrazione del Teorema 6.2 L'idea della dimostrazione è la seguente:
partendo dal fatto che il valore della call è funzione del tempo e del prezzo dell'opzione, ossia $Y_t = C(t, S_t)$, applicando il lemma di Ito per il calcolo di $dY_t = dC(t, S_t)$ e sfruttando il Teorema 6.1, $dY_t = a_t dS_t + b_t dB_t$, otteniamo l'equazione di Black e Scholes.
Utilizzando la formula di Ito[2]

$$\begin{aligned} dY_t &= C_t(t, S_t)dt + C_x(t, S_t)dS_t + \frac{1}{2}C_{xx}(t, S_t)d < S >_t \\ &= C_t(t, S_t)dt + C_x(t, S_t)(\mu S_t dt + \sigma S_t dW_t) + \frac{1}{2}C_{xx}(t, S_t)\sigma^2 S_t^2 dt \\ &= C_t(t, S_t)dt + C_x(t, S_t)\mu S_t dt + C_x(t, S_t)\sigma S_t dW_t + \frac{1}{2}C_{xx}(t, S_t)\sigma^2 S_t^2 dt \end{aligned}$$

cioè

$$dY_t = [C_t(t, S_t) + C_x(t, S_t)\mu S_t + \frac{1}{2}C_{xx}(t, S_t)\sigma^2 S_t^2]dt + C_x(t, S_t)\sigma S_t dW_t. \tag{6.14}$$

Dalla (6.14) segue che Y_t è ancora un processo di Ito

$$dY(t) = \mu_Y^*(t)dt + \sigma_Y^*(t)dW_t \tag{6.15}$$

dove

$$\begin{cases} \sigma_Y^*(t) &= C_x(t, S_t)\sigma S_t \\ \mu_Y^*(t) &= C_t(t, S_t) + C_x(t, S_t)\mu S_t + \frac{1}{2}C_{xx}(t, S_t)\sigma^2 S_t^2 \end{cases} \tag{6.16}$$

D'altra parte, poiché abbiamo supposto l'esistenza di una strategia autofinanziante e di copertura perfetta (a_t, b_t) sappiamo che, per il Teorema 6.1

$$Y_t = a_t S_t + b_t B_t$$

allora

$$\begin{aligned} dY_t &= a_t dS_t + b_t dB_t \\ &= a_t(\mu S_t dt + \sigma S_t dW_t) + b_t(rB_t dt) \\ &= (a_t \mu S_t + b_t r B_t)dt + a_t \sigma S_t dW_t. \end{aligned} \tag{6.17}$$

[2] S_t è un processo con differenziale stocastico dato da

$$dS_t = \mu S_t dt + \sigma S_t dW_t$$

almeno formalmente, otteniamo

$$(dS_t)^2 = \mu^2 S_t^2 (dt)^2 + 2\mu\sigma S_t^2 dt dW_t + \sigma^2 S_t^2 (dW_t)^2.$$

I termini in $(dt)^2$ e in $(dt)(dW_t)$ sono trascurabili rispetto a quello in dt, essendo $(dt)^2 = o(dt)$ e $(dt)(dW_t) = o(dt)$, inoltre $(dW_t)^2 = dt$ (nel senso del Lemma 3.2). Dalla definizione di differenziale stocastico abbiamo che

$$< S >_t = \int_0^t \sigma^2 S_t^2 ds$$

e quindi che

$$d < S >_t = \sigma^2 S_t^2 dt.$$

Anche dalla (6.17), deduciamo che Y_t è un processo di Ito, cioè

$$dY(t) = \mu_Y(t)dt + \sigma_Y(t)dW_t \tag{6.18}$$

dove

$$\begin{cases} \sigma_Y(t) &= a_t \sigma S_t \\ \\ \mu_Y(t) &= a_t \mu S_t + b_t r B_t \end{cases} \tag{6.19}$$

L'unicità nella rappresentazione dell'integrale di Ito in forma differenziale, ovvero l'uguaglianza delle (6.15) e (6.18) implica che

$$\sigma_Y(t) = \sigma^*_Y(t) \quad \text{ e } \quad \mu_Y(t) = \mu^*_Y(t).$$

Il confronto delle (6.16) e (6.19) ci consente di ottenere la strategia di copertura perfetta (a_t, b_t).
Più precisamente, dalla $\sigma_Y(t) = \sigma^*_Y(t)$ otteniamo

$$a_t \sigma S_t = C_x(t, S_t)\sigma S_t$$

e quindi, essendo $(S_t > 0)$

$$a_t = C_x(t, S_t).$$

La derivata $C_x(t, S_t)$ viene indicata con $\Delta := C_x(t, S_t)$ e chiamata di conseguenza lettera greca (o semplicemente greca) ***delta***[3]. Passiamo alla determinazione di b_t.
Essendo $Y_t = a_t S_t + b_t B_t$

$$\begin{aligned} b_t &= \frac{1}{B_t}[C(t, S_t) - a_t S_t] \\ &= \frac{1}{B_t}[C(t, S_t) - C_x(t, S_t)S_t]. \end{aligned}$$

Abbiamo così trovato che la strategia di copertura perfetta (a_t, b_t) soddisfa la (6.12) e la (6.13), cioè

$$\begin{aligned} a_t &= C_x(t, S_t) \\ b_t &= \frac{C(t, S_t) - C_x(t, S_t)S_t}{B_t}. \end{aligned}$$

Ritorniamo all'uguaglianza $\mu_Y(t) = \mu^*_Y(t)$, dalla quale otteniamo

$$a_t \mu S_t + b_t r B_t = C_t(t, S_t) + C_x(t, S_t)\mu S_t + \frac{1}{2}C_{xx}(t, S_t)\sigma^2 S_t^2.$$

Sostituendo in quest'ultima i valori espliciti di (a_t, b_t) abbiamo

$$C_x(t, S_t)\mu S_t + \frac{C(t, S_t) - C_x(t, S_t)S_t}{B_t} B_t r = C_t(t, S_t) + C_x(t, S_t)\mu S_t + \frac{1}{2}C_{xx}(t, S_t)\sigma^2 S_t^2$$

[3]**Il calcolo esplicito di Δ è possibile e si ottiene

$$\Delta = \frac{\partial C}{\partial x}(t, x) = \Phi\big(\widetilde{z}(T - t, x)\big)$$

dove Φ è come al solito la distribuzione della gaussiana standard e $\widetilde{z}$ è definito dalla (6.3), ovvero

$$\widetilde{z} = \widetilde{z}(\tau, x) = \frac{\ln(\frac{x}{K}) + (r + \frac{\sigma^2}{2})\tau}{\sigma\sqrt{\tau}}.$$

Si noti che quindi Δ (e quindi a_t) è sempre positivo.

che è ancora pari a

$$C_x \mu S_t - rC_x S_t + Cr = C_t + C_x \mu S_t + \frac{1}{2} C_{xx} \sigma^2 S_t^2$$

ovvero

$$C_t(t, S_t) + rC_x(t, S_t)S_t - rC(t, S_t) + \frac{1}{2} C_{xx}(t, S_t)\sigma^2 S_t^2 = 0. \tag{6.20}$$

**Di conseguenza, se $C(t,x)$ soddisfa l'equazione di Black e Scholes annunciata, allora l'equazione (6.20) è valida; d'altra parte bisogna assicurarsi che dal fatto che C soddisfa l'equazione (6.20) si riesca ad ottenere che allora C soddisfa l'equazione di Black e Scholes.

A questo scopo si consideri che l'opzione, per ogni $t \in [0,T]$, può assumere valori nell'intervallo $(0, \infty)$ e se fosse per assurdo che

$$\begin{aligned} B = \{ \ (t,x) \mid C_t(t,x) + rC_x(t,x)x - rC(t,x) + \tfrac{1}{2} C_{xx}(t,x)\sigma^2 x^2 \neq 0 \ \} \\ \supseteq (t_1, t_2) \times (\alpha, \beta,) \end{aligned}$$

allora avremmo che

$$\mathbf{P}\Big((t, S_t) \in B\Big) \geq \mathbf{P}\Big((t, S_t) \in (t_1, t_2) \times (\alpha, \beta,)\Big) > 0$$

ovvero che

$$\mathbf{P}\Big(C_t(t, S_t) + rC_x(t, S_t)S_t - rC(t, S_t) + \frac{1}{2} C_{xx}(t, S_t)\sigma^2 S_t^2 \neq 0\Big) > 0.$$

in contraddizione con la (6.20).

Osservazione 6.3. *Implicitamente abbiamo supposto che $C(T, S_T) = (S_T - K)^+$ e che quindi $C(t,x)$ deve soddisfare anche la condizione $C(T,x) = (x-K)^+$.*
È chiaro che nel caso $f_T = f(T, S_T)$ abbiamo invece che $C(T,x) = f(T,x)$.

Andiamo a determinare la soluzione dell'equazione di Black e Scholes, ovvero il valore di una opzione *call europea con prezzo di esercizio* K e tempo di maturità T, all'epoca $t \in [0,T)$, la cui funzione di payoff, come definito precedentemente, è pari a

$$Y_T = (S_T - K)^+.$$

È importante osservare che l'equazione (6.11) non dipende dal parametro μ, ovvero dal tasso di rendimento istantaneo atteso del bene sottostante l'azione.
Ciò porta a pensare che la soluzione deve essere del tipo

$$\widetilde{\mathbf{E}}[(S_T - K)^+ e^{-r(T-t)} \mid S_t = x] \tag{6.21}$$

dove $\widetilde{\mathbf{E}}$ è il valore atteso nel caso in cui sia valida l'ipotesi di neutralità al rischio, e questa è proprio la strada seguita da Black e Scholes[4].

[4]**In questi appunti abbiamo ottenuto la (6.21) con un procedimento limite a partire dal modello binomiale multiperiodale (o CRR).

Bibliografia

[1] Bachelier, L. Théorie de la spéculation. *Anna. Sci. École Norm. Sup. (3) 17* (1900), 21–86.

[2] Baldi, P. *Equazione differenziali stocastiche e applicazioni.* Quaderni dell'Unione Matematica Italiana, 28. Bologna: Pitagora Editrice. VIII, 309 p. (seconda edizione), 2000.

[3] Baldi, P., and Caramellino, L. Appunti del corso: Probabilità e Finanza. Dip. Mat. - Univ. di Roma Tor Vergata - http://www.mat.uniroma2.it/~processi/did_0304/pf.htm, 2004.

[4] Billingsley, P. *Probability and measure*, third ed. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1995. A Wiley-Interscience Publication.

[5] Björk, T. *Arbitrage theory in continuous time.* Oxford University Press, Oxford, 1998.

[6] Black, F., and Scholes, M. The pricing of options and corporates liabilities. *Journal of Political Economy 81* (1973), 637–659.

[7] Black, F., and Scholes, M. The pricing of options and corporates liabilities. *Journal of Political Economy 81* (1973), 637–659.

[8] Boness, A. Elements of a theory of stock option value. *Journal of Political Economy 72* (1964), 163–175.

[9] Dall'Aglio, G. *Calcolo delle Probabilità.* Zanichelli, Bologna, 2000. II edizione.

[10] Hull, C. J. *Opzioni, futures e altri derivati.* Il Sole 24 Ore, Milano, 2003. III edizione.

[11] Lamberton, D., and Lapeyre, B. *Introduction to stochastic calculus applied to finance.* Chapman & Hall, London, 1996. Translated from the 1991 French original by Nicolas Rabeau and François Mantion.

[12] Luenberger, D. C. *Investment Science.* Oxford University Press, Oxford, 1998.

[13] Merton, R. C. Theory of rational option pricing. *Bell J. Econom. and Management Sci. 4* (1973), 141–183.

[14] Ross, S. M. *An introduction to mathematical finance.* Cambridge University Press, Cambridge, 1999. Options and other topics.

[15] Samuelson, P., and Merton, R. C. A complete model of warrant prices that maximises the utility. *Industrian Management Review 10* (1969), 17–46.

[16] Samuelson, P. A. Rational theory of warrant pricing. *Industrial Management Review 6* (1965), 13–31.

[17] Shiryaev, A. N. *Essentials of stochastic finance. Facts, models, theory,*, vol. 3 of *Advanced Series on Statistical Science & Applied Probability.* World Scientific Publishing Co. Inc., River Edge, NJ, 1999. Translated from the Russian manuscript by N. Kruzhilin.

[18] Sprenkle, C. Warrant prices as indicatores of expectations. *Yole Economic Essay 1* (1961), 179–232.

[19] Steele, J. M. *Stochastic calculus and financial applications*, vol. 45 of *Applications of Mathematics (New York).* Springer-Verlag, New York, 2001.

[20] Torp, E., and Kassouf. Beat the market. *Random House, New York* (1967).