Università degli Studi di Roma “La Sapienza”
Anno Accademico 2007-2008

Facoltà di Scienze Matematiche Fisiche e Naturali
Corso di Laurea in Matematica

Alcuni appunti per il corso di

METODI PROBABILISTICI PER L’ECONOMIA E LA FINANZA

Giovanna Nappo

A.A. 2007/08

versione del 4 marzo 2008

Indice

Introduzione

Questi appunti sono una raccolta di vari appunti, scritti durante vari anni di insegnamento di corsi universitari sui processi aleatori. E questo spiega la ancora non completa coerenza della presentazione degli argomenti. Nella preparazione delle lezioni ho utilizzato diversi testi e i principali sono il testo di P. Baldi [1] e quello di A. N. Shiryaev [17]. Ma andrebbero citati anche altri testi, come ad esempio il testo di L. Breiman [6] e i testi di S. Ross [14] e [15]. Inoltre, in particolare per le parti dedicate agli esempi in finanza, ho utilizzato, rimaneggiandole, anche alcune parti delle tesi di laurea di due studentesse Valeria Belleudi e Stefania Latella, che ringrazio per il lavoro svolto. Questo spiega la grande discontinuità di stile delle varie parti di questi appunti.

A partire dall'anno accademico 2003/04 il corso di Metodi Probabilistici per l'Economia e la Finanza è divenuto un corso complementare della laurea triennale in Matematica, di conseguenza si è posto il problema di riuscire a presentare argomenti come il valore atteso condizionato e le martingale in modo che uno studente con un bagaglio minimo di conoscenze in probabilità possa ugualmente capire l'utilità di questi concetti e le sue applicazioni in finanza. Lo sforzo principale fatto consiste appunto in questo tentativo, ma per ora si tratta solo di un tentativo preliminare.

Nonostante ciò rimangono nel testo alcune note che erano state scritte per studenti con un bagaglio di conoscenze più elevato, quelle parti però non sono state ancora riguardate.

Notazioni: I capitoli, o le sezioni, con il segno Ⓡ contengono richiami di nozioni di base di teoria delle probabilità, mentre i capitoli, o le sezioni, con il segno † contengono argomenti di approfondimento.

Programma per l'anno accademico 2003/04: I capitoli, le sezioni, singoli teoremi o dimostrazioni con il segno ‡ non sono in programma per l'anno accademico 2003/04.

Si ricorda che per la preparazione all'esame si richiedono anche alcuni elementi di Matematica Finanziaria (che corrispondono ai capitoli 4 e 5 del testo di Ross [15]), l'enunciato del teorema dell'arbitraggio, con relative applicazioni e lo studio elementare del modello binomiale multiperiodale (Cox, Ross e Rubistein) (che corrispondono alle sezioni 6.1 e 6.2 del testo di Ross [15] e ai capitoli 1 e 2 del testo di Björk [4]). Infine si richiede la trattazione elementare del moto browniano geometrico e la formula di Black e Scholes (che corrispondono al capitolo 3 e alle sezioni 7.1 e 7.2 del testo di Ross [15]).

Programma per l'anno accademico 2004/05: Rispetto allo scorso anno manca la parte sui processi a tempo continuo. Rimangono validi i richiami dei testi di Björk [4] e di Ross [15], ma in queste note compaiono anche alcuni dei temi che prima erano affrontati solo in quei testi.

Programma per l'anno accademico 2005/06 e 2006/07: Si tratta di studiare i capitoli 1-6, con l'esclusione delle parti con il segno ‡. Si noti che è stata aggiunta la sezione 6.3 sugli alberi binomiali. Inoltre, oltre ad alcune parti di elementi di matematica finanziaria (testo base il Ross) si richiede di studiare l'approssimazione del modello di Black e Scholes (le note di questa parte sono su un file a parte).

Programma di massima, per l'anno accademico 2007/08: Si tratta di studiare i capitoli 1-7, con l'esclusione delle parti con il segno ‡ (il capitolo 7 contiene l'approssimazione del modello di Black e Scholes). Inoltre, sono necessarie alcuni di elementi di base di matematica finanziaria (testo di riferimento il Ross).
Si noti che l'appendice sulle variabili Gaussiane multidimensionali è stata spostata nel Capitolo 2, Sezione 2.5.

ATTENZIONE: I capitoli, le sezioni, singoli teoremi o dimostrazioni con il segno ‡ non sono in programma.

Per una versione aggiornata dei capitoli 8, 9, si faccia riferimento agli appunti per il corso di Dottorato *Metodi probabilistici per le equazioni alle derivate parziali* (A.A. 2007-08).

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

[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.

- strumenti sottostanti o primari come
  - conti bancari
  - buoni
  - 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

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 $0 \le t < 1$) x_1 rappresenta l'ammontare dalla fine del primo periodo all'inizio (escluso) del secondo periodo (cioè per $1 \le t < 2$), mentre x_k rappresenta l'ammontare dalla fine del $k-simo$ periodo alla sua fine (cioè per $k \le t < k+1$).

- **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

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

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 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.

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 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] $x_t^{(m)} = x^{(m)}_{\lfloor m\cdot t\rfloor/m}$ dalla (1.1) si ottiene che

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

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

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

[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.

[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

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

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 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]

---

[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.

[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

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

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;
- 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.

[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 $\rho(T-t,T)$.

[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 [5] e l'altro di Merton [11], che influenzarono notevolmente i metodi di prezzaggio.

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

[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

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

effettuati 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 $t = T$.

[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 [4].

- La ditta GPR pagherà, al tempo $t = T$, al tasso di cambio K *euro*/\$

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

[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.

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***. 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, $\ldots$, 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:

[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)

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

[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

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

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* $j = 1, 2, ..., m$*.*

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*

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

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

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

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

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

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 (si veda il capitolo 4 corrispondente).*

Non diamo in queste note la dimostrazione del teorema dell'arbitraggio, perché nel seguito vedremo la dimostrazione probabilistica nel caso del modello binomiale multiperiodale (capitoli 5 e 6). Segnaliamo tuttavia che nelle note di Baldi e Caramellino [2] 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.

---

[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

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

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 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*

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

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*

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

che equivale a

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

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

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

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

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

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* $1 - O$*.*

[30] Questo esempio è tratto dal libro di Ross [15]. Si veda anche l'esempio I.5.1 del libro di Dall'Aglio [7].

[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 $o_i = (1 + o_i) - 1$.

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

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

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). *DA RIVEDERE IN MODO PESANTE!!! 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 Z è una variabile aleatoria che può assumere solo due valori u (per UP) e d (per DOWN) (da cui il nome binomiale)

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

$$\text{al tempo t=0} \qquad \tilde{B}_0 = \frac{B_0}{B_0} = 1 \qquad \tilde{S}_0 = \frac{S_0}{B_0} = \frac{s_0}{B_0}$$

$$\text{al tempo t=1} \qquad \tilde{B}_1 = \frac{B_1}{B_1} = 1 \qquad \tilde{S}_1 = \frac{S_1}{B_1} = \frac{S_1}{B_0(1+r)} = \frac{s_0\, Z}{B_0(1+r)}$$

[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,

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

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

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

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 osservazione 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*

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

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

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

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

Dimostriamo ora direttamente (senza far uso del teorema dell'arbitraggio) che la condizione precedente, ovvero $d < 1 + r < u$, ***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[34]) mentre γS_0 è la quantità di denaro investita nel titolo rischioso (l'azione), ovvero γ è il numero di azioni comprate[35].

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***[36] 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, 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

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

• 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

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

[34]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.

[35]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.

[36]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.

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, 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***[37], 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[38] 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[39] *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).*

[37] 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).

[38] Si veda la successiva osservazione 1.6 per il motivo per cui si parla di derivati.

[39] 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.

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

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

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.*

Ovviamente nessun problema sorge nella seguente situazione: esiste (ed è 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

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

In questo caso infatti la condizione

$$\begin{aligned} &\tilde{\mathbb{E}}[\Delta_\Psi] = 0 \\ &\Updownarrow \\ &\beta\tilde{\mathbb{E}}[\Delta_0] + \sum_{i=1}^{m} \gamma^i \tilde{\mathbb{E}}[\Delta_i] = 0 \end{aligned}$$

è automaticamente soddisfatta.

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[40]*. 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 investimento o portfolio**[41].

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

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

[41] 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.

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

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

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*[42]*.*

Dimostrazione dell'affermazione precedente*: si distinguono due casi*

• *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

[42]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.

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

• *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.*[43]*. 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

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

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

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

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

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

[43] 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.

.

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

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

ovvero[44] *per* $\alpha=-1$ *(*$B_0=1$*, per semplicità)*

$$\gamma=\frac{f_1(s_0\,u)-f_1(s_0\,d)}{s_0u-s_0d}. \tag{1.26}$$

Di conseguenza la prima condizione (1.24) *permette di trovare*[45] 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*[46] *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:

[44] 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.

[45] Si osservi che

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

[46] 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 $\gamma=\gamma(-1)$.

infatti si ricava che[47]

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

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

***Nella seguente Figura 1.1 riassumiamo il procedimento con un albero:

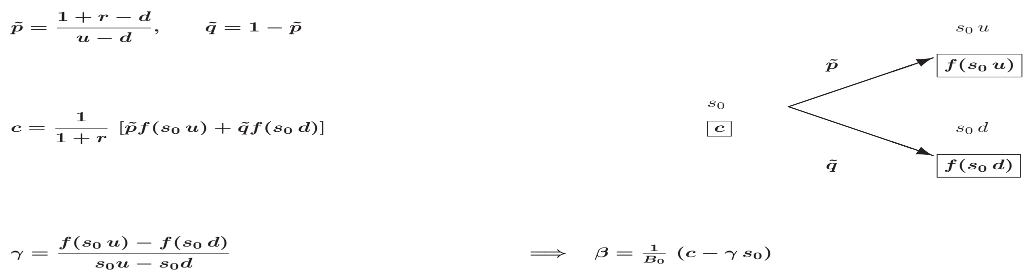

Figura 1.1: Albero binomiale per il calcolo del prezzo c e della strategia di copertura perfetta: trovati c e γ, il valore β è automaticamente determinato.

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. ***

[47] Con semplici passaggi si ottiene

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

Capitolo 2

Richiami su spazi di probabilità ®

2.1 Esempi di spazi di probabilità

Come dovrebbe essere noto uno spazio di probabilità è una terna $(\Omega, \mathcal{F}, \mathbb{P})$, dove

$\mathcal{F}$ è una ***σ-algebra***, ovvero $\mathcal{F}$ è una famiglia di sottoinsiemi di Ω, cioè $\mathcal{F}$ è un sottoinsieme di $\mathcal{P}(\Omega)$, tale che

$$\Omega \in \mathcal{F}; \tag{2.1}$$
$$\text{se } A \in \mathcal{F}, \text{ allora } A^c \in \mathcal{F}; \tag{2.2}$$
$$\text{se } A_n \in \mathcal{F}, n \in \mathbb{N}, \text{ allora } \cup_{n\in\mathbb{N}} A_n \in \mathcal{F}; \tag{2.3}$$

$\mathbb{P}$ è una ***misura di probabilità***, ovvero

$$\mathbb{P} : \mathcal{F} \mapsto [0,1]; \quad A \mapsto \mathbb{P}(A)$$

con le proprietà che

$$\mathbb{P}(\Omega) = 1; \tag{2.4}$$
$$\text{se } A_n \in \mathcal{F}, n \in \mathbb{N}, \text{ con } A_n \cap A_m = \emptyset \text{ per } n \neq m, \tag{2.5}$$
$$\text{allora } \mathbb{P}\big(\bigcup_{n\in\mathbb{N}} A_n\big) = \sum_{n\in\mathbb{N}} \mathbb{P}\big(A_n\big).$$

La σ-algebra $\mathcal{F}$ rappresenta l'informazione disponibile, ovvero gli eventi appartenenti a $\mathcal{F}$ sono gli unici eventi di cui abbiamo la possibilità di sapere se si sono verificati oppure no.

Oltre alla misura di probabilità $\mathbb{P}$, per tutti gli eventi $A \in \mathcal{F}$ con $\mathbb{P}(A) > 0$, si possono definire le ***probabilità condizionate***[1] ***all'evento*** A, che rappresentano la valutazione della probabilità nel caso in cui si verificasse l'evento A:

$$\mathbb{P}(\cdot|A)\, \mathcal{F} : \to [0,1] \tag{2.6}$$
$$E \mapsto \mathbb{P}(E|A) := \frac{\mathbb{P}(E \cap A)}{\mathbb{P}(A)} \tag{2.7}$$

Vediamo ora alcuni esempi elementari di spazi di probabilità:

[1] È facile verificare che la funzione $\mathbb{P}(\cdot|A)$ definita in (2.6) è una probabilità, cioè soddisfa gli assiomi delle probabilità. Per mettere in evidenza tale fatto va detto che Kolmogorov aveva adottato la notazione $\mathbb{P}_A(\cdot)$, ovvero $\mathbb{P}_A(E)$ invece di $\mathbb{P}(E|A)$, anche per mettere meglio in evidenza questa proprietà.

Esempio 2.1. *Qualunque sia* Ω*, la* σ*-algebra banale* $\mathcal{F} = \{\emptyset, \Omega\}$ *è una* σ*-algebra, e necessariamente* $\mathbb{P}(\Omega) = 1$ *e* $\mathbb{P}(\emptyset) = 0$*.*

Esempio 2.2. *Qualunque sia* Ω*, preso un sottoinsieme proprio* A *di* Ω *la* σ*-algebra* $\mathcal{F} = \{\emptyset, A, A^c, \Omega\}$ *è una* σ*-algebra, e necessariamente* $\mathbb{P}(\Omega) = 1$, $\mathbb{P}(\emptyset) = 0$, $\mathbb{P}(A) = p$, $\mathbb{P}(A^c) = 1 - p$, *per un* $p \in [0, 1]$*.*

Esempio 2.3. *Qualunque sia* Ω*, sia* $\{H_m,\, m = 1, 2, \ldots, N\}$ *una* **partizione finita** *di* Ω*, cioè se gli eventi sono* **incompatibili***:*

$$H_n \cap H_m = \emptyset \text{ per } n \neq m, \; n, m \in \{1, 2, \ldots, N\}$$

ed **esaustivi***:*

$$\bigcup_{m=1}^{N} H_m = \Omega,$$

allora la famiglia $\mathcal{M} = \{A = \bigcup_{m \in I} H_m, \text{ al variare di } I \subseteq \{1, 2, \ldots, N\}\}$*, (con la convenzione che* $\bigcup_{m \in \emptyset} H_m = \emptyset$*) è una* σ*-algebra. Inoltre se* $p_1, p_2, \ldots, p_N$ *sono numeri non negativi, a somma* 1*, ovvero*

$$p_m \geq 0,\, m = 1, 2, \ldots, N, \qquad \sum_{m=1}^{N} p_m = 1,$$

allora $\mathbb{P}: \mathcal{M} \mapsto [0, 1];\; A \mapsto \mathbb{P}(A)$*, con*

$$\mathbb{P}(A) = \sum_{m \in I} p_m, \qquad \text{per } A = \bigcup_{m \in I} H_m, \tag{2.8}$$

definisce una probabilità su $(\Omega, \mathcal{M})$*.*

Esempio 2.4. *Le proprietà dell'esempio precedente valgono anche nel caso di una* **partizione numerabile** $\{H_m,\, m \in \mathbb{N}\}$ *con i dovuti cambiamenti: cioè, se*

$$H_n \cap H_m = \emptyset \text{ per } n \neq m, \; n, m \in \mathbb{N},$$
$$\bigcup_{m \in \mathbb{N}} H_m = \Omega,$$

allora la famiglia

$$\mathcal{F} = \{A = \bigcup_{m \in I} H_m, \text{ al variare di } I \subseteq \mathbb{N}\},$$

(con la convenzione che $\bigcup_{m \in \emptyset} H_m = \emptyset$*), è una* σ*-algebra*[2]*.*
Inoltre se $p_1, p_2, \ldots, p_m, \ldots$ *sono numeri non negativi, somma* 1*, ovvero*

$$p_m \geq 0,\, m \in \mathbb{N}, \qquad \sum_{m \in \mathbb{N}} p_m = 1,$$

[2] La verifica è banale:

$$\Omega = \bigcup_{m \in \mathbb{N}} H_m, \text{ ovvero } I = \mathbb{N}$$

$$\text{se } A = \bigcup_{m \in I} H_m, \text{ allora } A^c = \bigcup_{m \in I^c} H_m$$

$$\text{se } A_n = \bigcup_{m \in I_n} H_m,\, n \geq 1, \text{ allora } \bigcup_{n=1}^{\infty} A_n = \bigcup_{m \in I} H_m, \text{ per } I = \cup_{n=1}^{\infty} I_n.$$

allora $\mathbb{P}:\ \mathcal{F} \mapsto [0,1];\ A \mapsto \mathbb{P}(A)$*, con*

$$\mathbb{P}(A) = \sum_{m \in I} p_m, \qquad \textit{per } A = \bigcup_{m \in I} H_m, \tag{2.9}$$

definisce una probabilità su $(\Omega, \mathcal{F})$*.*

La verifica di quest'ultima proprietà è banale[3]*.*

Elenchiamo adesso alcune ***proprietà e notazioni relative alle*** σ***-algebre***:

1 ***l'intersezione di*** σ***-algebre è una*** σ***-algebra***
Sia $\{\mathcal{G}_\alpha, \alpha \in \Lambda\}$ una famiglia di σ-algebre, allora $\mathcal{F} := \bigcap_{\alpha \in \Lambda} \mathcal{G}_\alpha$ è una σ-algebra[4].

2 ***l'unione di*** σ***-algebre non è (in generale) una*** σ***-algebra***
Basta mostrare con un controesempio che l'unione di due σ-algebre non è una σ-algebra: ad esempio se $\mathcal{G}_i = \{\emptyset, A_i, A_i^c, \Omega\}$, con $A_1 \cap A_2 \neq \emptyset, A_1, A_2$, allora $\mathcal{G}_1 \cup \mathcal{G}_2 = \{\emptyset, A_1, A_2, A_1^c, A_2^c, \Omega\}$ non è una σ-algebra.

3 ***la*** σ***-algebra generata da una collezione di eventi***
Sia $\mathcal{K}$ un sottoinsieme di $\mathcal{P}(\Omega)$, l'insieme delle parti di Ω, allora

$$\sigma(\mathcal{K}) := \bigcap_{\mathcal{G}:\mathcal{K} \subseteq \mathcal{G}}$$

è la σ-algebra[5] generata da $\mathcal{K}$.

In particolare quindi la σ-algebra $\mathcal{M}$, generata dalla partizione $\{H_m;\ m \in \mathbb{N}\}$ come nell'Esempio 2.4, coincide con $\sigma(\{H_m;\ m \in \mathbb{N}\})$, in quanto, come già visto $\mathcal{M}$ è una σ-algebra, e inoltre ogni σ-algebra che contenga $\{H_m;\ m \in \mathbb{N}\}$, deve necessariamente contenere tutte le unioni del tipo $\cup_{m \in I} H_m$.

4 ***la*** σ***-algebra generata da una collezione di*** σ***-algebre***
Nel caso in cui $\mathcal{K} = \bigcup_{\alpha \in \Lambda} \mathcal{G}_\alpha$, dove $\mathcal{G}_\alpha$ sono σ-algebre, allora si pone

$$\bigvee_{\alpha \in \Lambda} \mathcal{G}_\alpha := \sigma\big(\bigcup_{\alpha \in \Lambda} \mathcal{G}_\alpha\big).$$

In particolare se $\mathcal{M} = \sigma(\{H_m;\ m \in \mathbb{N}\})$ e $\mathcal{N} = \sigma\big(\{K_\ell;\ \ell \in \mathbb{N}\}\big)$, allora

$$\mathcal{M} \vee \mathcal{N} = \sigma(\{H_m \cap K_\ell;\ m \in \mathbb{N},\ \ell \in \mathbb{N}\}) = \big\{E = \bigcup_{(m,\ell) \in J} H_m \cap K_\ell;\ \text{con } J \subseteq \mathbb{N} \times \mathbb{N}\big\}.$$

---

[3] La funzione $\mathbb{P}: \mathcal{M} \mapsto [0,1]$ definita in (2.9) è una probabilità, infatti

$$\mathbb{P}(\Omega) = \sum_{m \in \mathbb{N}} p_m = 1,$$

$$\text{se } A_n = \bigcup_{m \in I_n} H_m \in \mathcal{M}, n \in \mathbb{N},\ \text{con } A_n \cap A_{n'} = \emptyset \text{ per } n \neq n',$$

$$\text{allora } \bigcup_{n \in \mathbb{N}} A_n = \bigcup_{m \in I} H_m \text{ con } I = \bigcup_{n \in \mathbb{N}} I_n,\ \text{e con } I_n \cap I_{n'} = \emptyset \text{ per } n \neq n',$$

$$\text{e quindi } \mathbb{P}\big(\bigcup_{n \in \mathbb{N}} A_n\big) = \mathbb{P}(A) = \sum_{\ell \in I} p_\ell = \sum_{n \in \mathbb{N}} \sum_{m \in I_n} p_m = \sum_{n \in \mathbb{N}} \mathbb{P}(A_n),$$

[4] La verifica è banale:

$\Omega \in \mathcal{F},$ in quanto $\Omega \in \mathcal{G}_\alpha,$ per ogni $\alpha \in \Lambda;$

se $A \in \mathcal{F},$ cioè se $A \in \mathcal{G}_\alpha,$ per ogni $\alpha \in \Lambda,$ allora $A^c \in \mathcal{G}_\alpha,$ per ogni $\alpha \in \Lambda,$ e quindi $A^c \in \mathcal{F};$

se $A_n \in \mathcal{F}, n \in \mathbb{N}$ cioè se $A_n \in \mathcal{G}_\alpha,$ per ogni $\alpha \in \Lambda, n \in \mathbb{N}$ allora $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{G}_\alpha,$ per ogni $\alpha \in \Lambda,$ e quindi $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{F};$

[5] Il fatto che $\bigcap_{\mathcal{G}:\mathcal{K} \subseteq \mathcal{G}}$ sia una σ-algebra, deriva dalla proprietà che l'intersezione di σ-algebre è una σ-algebra.

5 ***la σ-algebra dei Boreliani*** Nel caso in cui $\mathcal{K} = \mathcal{A}$, la famiglia degli aperti di $\mathbb{R}^k$, allora

$$\mathcal{B}(\mathbb{R}^k) := \sigma(\mathcal{A})$$

è detta σ-algebra dei boreliani, o σ-algebra di Borel, ed ogni elemento di $\mathcal{I}$ di $\mathcal{B}(\mathbb{R}^k)$ è detto ***boreliano***.

2.2 Variabili aleatorie

Definizione 2.1. *Dato uno spazio di probabilità* $(\Omega, \mathcal{F}, \mathbb{P})$[6]*, una* **variabile aleatoria reale** X *è una funzione* $\mathcal{F}$**-misurabile***, ovvero una funzione*

$$X : \Omega \mapsto \mathbb{R}; \quad \omega \mapsto X(\omega),$$

tale che la controimmagine di ogni aperto $\mathcal{O} \in \mathcal{A}$ *sia un elemento di* $\mathcal{F}$[7]*, cioè tale che*

$$X^{-1}(\mathcal{O}) := \{\omega \text{ tali che } X(\omega) \in \mathcal{O}\} \in \mathcal{F}, \text{ per ogni aperto } \mathcal{O} \in \mathcal{A}.$$

Si dice anche che X *è una* **variabile aleatoria** $\mathcal{F}$**-misurabile**.
Una definizione analoga vale nel caso di variabili aleatorie multidimensionali

$$\boldsymbol{X} : \Omega \mapsto \mathbb{R}^k; \quad \omega \mapsto \boldsymbol{X}(\omega) = \big(X_1(\omega), \dots, X_k(\omega)\big),$$

basta infatti sostituire $\mathbb{R}$ *con* $\mathbb{R}^k$*.*

Vediamo alcuni esempi di variabili aleatorie $\mathcal{F}$-misurabili, al variare della σ-algebra $\mathcal{F}$.

Esempio 2.5. *Se* $\mathcal{F} = \{\emptyset, \Omega\}$*, allora le uniche variabili aleatorie reali* X $\mathcal{F}$*-misurabili sono le costanti:*
Se $X : \Omega \mapsto \mathbb{R}$*;* $\omega \mapsto X(\omega) = c$*, allora* $X^{-1}(\mathcal{O})$ *è l'evento impossibile(=insieme vuoto* $\emptyset$*), se* $c \notin \mathcal{O}$*, oppure è l'insieme certo(=*Ω*), se* $c \in \mathcal{O}$*.*
Viceversa se $X : \Omega \mapsto \mathbb{R}$*;* $\omega \mapsto X(\omega)$ *non è costante allora* X *assume almeno due valori* c_1 *e* c_2 *distinti (cioè esistono* ω_i *tale che* $X(\omega_i) = c_i$*, per* $i = 1, 2$*, con* $c_1 \neq c_2$*). Quindi se* $c_1 \in \mathcal{O}$*, ma* $c_2 \notin \mathcal{O}$*, allora* $\omega_1 \in X^{-1}(\mathcal{O})$*, mentre* $\omega_2 \notin X^{-1}(\mathcal{O})$*, ovvero* $\emptyset \subset X^{-1}(\mathcal{O}) \subset \Omega$ *(dove le inclusioni sono in senso stretto), e quindi* X *non è* $\mathcal{F}$*-misurabile.*

Si noti che l'esempio precedente mostra anche che tutte le variabili aleatorie costanti sono misurabili rispetto a qualunque σ-algebra ($\{\emptyset, \Omega\} \subseteq \mathcal{F}$, per ogni σ-algebra $\mathcal{F}$).

Esempio 2.6. *Sia* $\{H_m, m \in \mathbb{N}\}$ *una partizione numerabile, e sia* $\mathcal{M}$ *come nell'esempio 2.4. Allora* $X : \Omega \mapsto \mathbb{R}$*;* $\omega \mapsto X(\omega)$ *è* $\mathcal{M}$*-misurabile, se e solo se esiste una successione di costanti* $\{c_m, m \in \mathbb{N}\}$[8]*, tale che*

$$X(\omega) = \sum_{m \in \mathbb{N}} c_m \mathbb{I}_{H_m}(\omega). \tag{2.10}$$

Se X *è definita come in (2.10) allora* X *è* $\mathcal{M}$*-misurabile, infatti per ogni aperto* $\mathcal{O}$*,*

$$X^{-1}(\mathcal{O}) = \bigcup_{m : c_m \in \mathcal{O}} H_m,$$

[6] In realtà basta che ci sia uno spazio ***probabilizzabile***, ovvero basta solo la coppia $(\Omega, \mathcal{F})$, mentre non è necessario specificare la misura di probabilità $\mathbb{P}$.

[7] Si noti l'analogia con la definizione di funzione continua $f : \mathbb{R}^k \mapsto \mathbb{R}^d$, come una funzione tale che le controimmagini di aperti sono aperti.

[8] Si noti che non si assume che i valori di $\{c_m\}$ siano tutti distinti, ad esempio nel caso della successione costante, cioè $c_m = c$ per ogni $m \in \mathbb{N}$, si trova una variabile aleatoria costante.

ovvero $X^{-1}(\mathcal{O}) = \bigcup_{m\in I} H_m \in \mathcal{M}$, *per* $I = \{m : c_m \in \mathcal{O}\}$.
Viceversa se X *è* $\mathcal{M}$*-misurabile, cioè, per ogni aperto* $\mathcal{O}$*, esiste un* $I \subseteq \mathbb{N}$ *tale che*

$$X^{-1}(\mathcal{O}) = \bigcup_{m\in I} H_m,$$

allora qualunque sia $c \in \mathbb{R}$*, preso* $\mathcal{O}^n$ *l'intervallo aperto* $(c - 1/n, c + 1/n)$ *si ha che*

$$X^{-1}(\{c\}) = X^{-1}\big(\bigcap_n \mathcal{O}^n\big) = \bigcap_n X^{-1}(\mathcal{O}^n) = \bigcap_n \bigcup_{m\in I^n} H_m = \bigcup_{m\in \bigcap_n I^n} H_m \in \mathcal{M},$$

Esempio 2.7. *Sia* $X : \Omega \mapsto \mathbb{R};\ \omega \mapsto X(\omega)$*, una funzione discreta, ovvero tale che l'immagine* $X(\Omega) = \{x \in \mathbb{R},$ *tali che esiste un* ω *con* $X(\omega) = x\}$ *di* X *sia un insieme numerabile (finito o infinito), cioè* $X(\Omega) = \{x_m, m \in \mathbb{N}\}$*, con* $x_n \neq x_m$ *per* $n \neq m$*. Allora*

$$X(\omega) = \sum_{m\in\mathbb{N}} x_m \mathbb{I}_{H_m}(\omega), \tag{2.11}$$

dove

$$H_m = X^{-1}(\{x_m\}) = \{\omega \text{ tali che } X(\omega) = x_m\}.$$

Si noti che $\{H_m,\, m \in \mathbb{N}\}$ *forma una partizione numerabile.*

Inoltre la funzione X *è una variabile aleatoria* $\boldsymbol{\mathcal{F}}$**-misurabile***, se e solo se*

$$H_m = X^{-1}(\{x_m\}) \in \mathcal{F}, \text{ per ogni } m \in \mathbb{N},$$

come è immediato da (2.11), osservando che, come nel caso precedente,

$$X^{-1}(\mathcal{O}) = \bigcup_{m: x_m\in\mathcal{O}} H_m.$$

Infine la variabile aleatoria X *si dice* **semplice** *o* **elementare***, se l'insieme* $X(\Omega)$ *è un insieme finito.*

Si può dimostrare che

1 se X è una variabile aleatoria $\mathcal{F}$-misurabile, allora la controimmagine $X^{-1}(\mathcal{I}) \in \mathcal{F}$, per ogni boreliano $\mathcal{I} \in \mathcal{B}(\mathbb{R})$,

2 la variabile aleatoria $\boldsymbol{X}$ è $\mathcal{F}$-misurabile, se e solo se ciascuna componente X_i è $\mathcal{F}$-misurabile[9], per ogni $i = 1, \dots, k$. In particolare $X^{-1}(\{x\}) \in \mathcal{F}$, per ogni $x \in \mathbb{R}$, in quanto $\{x\} = \bigcap_n (x - 1/n, x + 1/n)$.

Connessa con la precedente Definizione 2.1 è la seguente definizione:

Definizione 2.2. *Sia data una funzione* $\boldsymbol{X} : \Omega \mapsto \mathbb{R}^k;\quad \omega \mapsto \boldsymbol{X}(\omega) = \big(X_1(\omega), \dots, X_k(\omega)\big)$*. Si dice* $\boldsymbol{\sigma}$**-algebra generata da** $\boldsymbol{X}$*, la* σ*-algebra*

$$\sigma(\boldsymbol{X}) = \bigcap_{\mathcal{G}\in\mathcal{R}^{\boldsymbol{X}}} \mathcal{G}$$

dove $\mathcal{R}^{\boldsymbol{X}}$ *è la famiglia delle* σ*-algebre, per le quali* $\boldsymbol{X}$ *è* $\mathcal{G}$*-misurabile*[10].

Si dimostra che

3 La σ-algebra generata da $\boldsymbol{X}$, si può caratterizzare come:

$$\sigma(\boldsymbol{X}) = \{A = \boldsymbol{X}^{-1}(\mathcal{I}), \text{ per } \mathcal{I} \in \mathcal{B}(\mathbb{R}^k)\},$$

[9]Dimostriamo solo la necessità, che è immediata: basta prendere $\mathcal{O} = \underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{i-1 \text{ volte}} \times \mathcal{O}_i \times \underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{k-i \text{ volte}}$.

[10]La famiglia $\mathcal{R}^{\boldsymbol{X}}$ non è vuota, in quanto contiene almeno $\mathcal{G} = \mathcal{P}(\Omega)$, l'insieme delle parti di Ω.

4 la funzione $\boldsymbol{X}$ è $\mathcal{F}$-misurabile, se e solo se $\sigma(\boldsymbol{X}) \subseteq \mathcal{F}$,

5 le variabili aleatorie $\sigma(\boldsymbol{X})$-misurabili a valori in $\mathbb{R}^d$ sono tutte e sole le variabili aleatorie $\boldsymbol{Z}$ per le quali esiste una funzione g boreliana[11] tale che
$$Z = g(\boldsymbol{X}).$$

Esempio 2.8. *Sia X una funzione semplice, come in Esempio 2.7, allora*

$$\sigma(X) = \sigma(\{H_m,\, m \in \mathbb{N}\}) = \{A = \bigcup_{m \in I} H_m;\ I \subseteq \mathbb{N}\},$$

dove $H_m = X^{-1}(\{x_m\})$.

Inoltre tutte e sole le variabili aleatorie $\sigma(X)$-misurabili sono le funzioni

$$Z : \Omega \mapsto \mathbb{R};\ \omega \mapsto Z(\omega) := \sum_m c_m \mathbb{I}_{H_m},$$

come discende immediatamente dall'Esempio 2.6. Di conseguenza se $g : \mathbb{R} \mapsto \mathbb{R}$ tale che $g(x_m) = c_m$, per ogni $m \in \mathbb{N}$, allora

$$Z(\omega) := \sum_m c_m \mathbb{I}_{H_m} = Z(\omega) = \sum_m g(x_m)\mathbb{I}_{X^{-1}(\{x_m\})}(\omega) = \sum_m g(x_m)\mathbb{I}_{\{x_m\}}\big(X(\omega)\big) = g(X(\omega)).$$

Terminiamo questa sezione, ricordando che le operazioni di massimo, minimo, somma, prodotto, di due funzioni misurabili, danno luogo a funzioni misurabili: quindi se X ed Y sono variabili aleatorie $\mathcal{F}$-misurabili, lo sono anche $X \vee Y = \max(X,Y)$, $X \wedge Y = \min(X,Y)$, $X+Y$, XY. In particolare sono variabili aleatorie $X^+ := X \vee 0$ e $X^- := (-X) \vee 0$.

2.3 Distribuzioni di variabili aleatorie

Sia $(\Omega,\, \mathcal{F},\, \mathbb{P})$ uno spazio di probabilità e sia

$$X : \Omega \mapsto \mathbb{R}^k;\ \omega \mapsto X(\omega)$$

una variabile aleatoria a valori ******in $\mathbb{R}^k$ ******. Tramite X è possibile definire una misura di probabilità $\mathbf{P}_X$ sullo spazio misurabile $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ nel seguente modo:

$$\mathbf{P}_X\ : \mathcal{B}(\mathbb{R}^k) \mapsto [0,1] \quad \mathcal{I} \mapsto \mathbf{P}_X(\mathcal{I}) := \mathbb{P}\big(X \in \mathcal{I}\big).$$

È facile verificare che effettivamente $\mathbf{P}_X$ definisce una probabilità sui boreliani $\mathcal{B}(\mathbb{R}^k)$. La misura di probabilità così definita è detta ***misura di probabilità indotta da*** X, o ***distribuzione di*** X.

***

A volte, per indicare la misura di probabilità indotta, si usa il simbolo $\mathbb{P}X^{-1}$, che nasce dall fatto che $\mathbf{P}_X(\mathcal{I}) := \mathbb{P}\big(X \in \mathcal{I}\big) = \mathbb{P}\big(X^{-1}(\mathcal{I})\big)$. Nel seguito, a volte useremo anche il simbolo μ_X per indicare la distribuzione di probabilità di X.

***

Esempio 2.9 (una variabile aleatoria binomiale). *Sia*

$$\Omega = \{0,1\}^N = \{\omega = (\omega_1, \omega_2, \ldots, \omega_N),\ \textit{con } \omega_i \in \{0,1\},\ \textit{per } i = 1,2,\ldots,N\},$$

[11] Una funzione $g : \mathbb{R}^k \mapsto \mathbb{R}^d$, si dice boreliana se è una funzione tale che le controimmagini di aperti sono boreliani.
Ovviamente le funzioni continue sono boreliane. Sono boreliane anche le funzioni continue a tratti, o meglio ancora costanti a tratti.
Per chi non avesse familiarità con i concetti di misurabilità può pensare a queste funzioni, o a funzioni che siano limite puntuale di funzioni di uno dei due tipi precedenti.

sia

$$\mathcal{F} = \mathcal{P}(\Omega),$$

l'insieme delle parti di Ω, *sia la probabilità definita attraverso* ****la relazione* ***

$$\mathbb{P}(***\{\omega\}***) := p^{\sum_{i=1}^N \omega_i}\,(1-p)^{N-\sum_{i=1}^N \omega_i},$$

dove p *è un numero fissato con la condizione che* $p \in (0,1)$. *Sia infine* X *la variabile aleatoria definita da*

$$X(\omega) := \sum_{i=1}^N \omega_i.$$

Si vede facilmente che

- 1 *la variabile aleatoria* X *assume solo i valori* $\{0, 1, \dots N\}$,

- 2 *per* $h \in \{0, 1, \dots N\}$ *si ha*[12]

$$\mathbf{P}_X(h) := \mathbb{P}(X = h) = \binom{N}{h} p^h\,(1-p)^{N-h},$$

- 3 *per ogni boreliano* $\mathcal{I}$

$$\mathbf{P}_X(\mathcal{I}) := \mathbb{P}(X \in \mathcal{I}) = \sum_{\substack{h=0\\ h\in\mathcal{I}}}^N \binom{N}{h} p^h\,(1-p)^{N-h}$$

Definizione 2.3 (variabili aleatorie con distribuzione binomiale). *Ogni variabile aleatoria* X *che* *******per la quale*

$$\mathbf{P}_X(\mathcal{I}) := \sum_{\substack{h=0\\ h\in\mathcal{I}}}^n \binom{n}{h} p^h\,(1-p)^{n-h}$$

*******viene detta una variabile aleatoria binomiale di parametri* n *e* p *e si scrive in breve* $X \sim Bin(n,p)$.

A volte, invece di definire lo spazio di probabilità e la variabile aleatoria X ed infine trovare la distribuzione di X, si può dare direttamente la distribuzione di X. Questo è il caso delle variabili aleatorie che vengono caratterizzate solo attraverso la densità discreta o con densità (di probabilità).

Definizione 2.4 (variabili aleatorie con densità discreta). *Si dice che una variabile aleatoria elementare* X *ha densità discreta*

$$\begin{pmatrix} x_1 & x_2 & \cdots & \cdots & x_m \\ p_1 & p_2 & \cdots & \cdots & p_m \end{pmatrix}$$

dove $x_1, x_2, \cdots x_m$ *sono elementi di* $\mathbb{R}^k$ *e* $p_1, p_2, \cdots p_m$ *sono numeri reali tali che*

$$p_j \geq 0 \quad \textit{per ogni } j = 1, 2, \cdots, m, \qquad \sum_{j=1}^m p_j = 1,$$

[12] L'evento $A_h := \{X = h\}$ è rappresentato dall'insieme, di cardinalità $\binom{N}{h}$, i cui elementi $\omega = (\omega_1, \omega_2, \dots, \omega_N)$ hanno la proprietà che $\sum_{i=1}^N \omega_i = h$. La probabilità di ciascuno di questi ω vale quindi

$$\mathbb{P}(\omega) = p^{\sum_{i=1}^N \omega_i}\,(1-p)^{N-\sum_{i=1}^N \omega_i} = p^h\,(1-p)^{N-h}$$

e la probabilità dell'insieme vale

$$\mathbb{P}(X = h) = \mathbb{P}(A_h) = \sum_{\omega\in A_h} \mathbb{P}(\omega) = \sum_{\omega\in A_h} p^h\,(1-p)^{N-h} = |A_h| p^h\,(1-p)^{N-h} = \binom{N}{h} p^h\,(1-p)^{N-h}$$

se, per ogni boreliano $\mathcal{I}$, vale

$$\mathbf{P}_X(\mathcal{I}) := \mathbb{P}(X \in \mathcal{I}) = \sum_{\substack{j=1 \\ x_j \in \mathcal{I}}}^{m} p_j.$$

In particolare quindi il significato di p_j è chiaro, essendo

$$\mathbb{P}(X = x_j) = p_j.$$

La definizione è analoga nel caso di variabili aleatorie discrete, la cui distribuzione viene caratterizzata attraverso una densità discreta *******su un insieme numerabile $\{x_k,\ k \geq 1\}$*******

$$\begin{pmatrix} x_1 & x_2 & \cdots & \cdots & x_m & x_{m+1} & \cdots \\ p_1 & p_2 & \cdots & \cdots & p_m & p_{m+1} & \cdots \end{pmatrix}$$

Definizione 2.5 (variabili con densità). *Si supponga di avere una funzione $f : \mathbb{R} \mapsto \mathbb{R}$ con le proprietà:*

$$f(x) \geq 0 \quad \text{per ogni } x \in \mathbb{R}^k, \qquad \int_{******\mathbb{R}^k} f(x)\, dx = 1,$$

si dice che X ha distribuzione con **densità (di probabilità)** *f se accade che, per ogni boreliano $\mathcal{I}$,*

$$\mathbf{P}_X(\mathcal{I}) := \int_{\mathcal{I}} f(x)\, dx.$$

Esempio 2.10 (distribuzione gaussiana). *Come caso particolare si consideri il caso della variabile aleatoria unidimensionale con densità*

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

dove μ è un numero reale e σ è un numero (strettamente) positivo. Una variabile aleatoria con questa distribuzione è detta **gaussiana** *o* **normale** *di valore atteso (o valore medio) μ e varianza σ^2. Brevemente si indica $X \sim N(\mu, \sigma^2)$. Se $\mu = 0$ e $\sigma^2 = 1$ si dice che X è una variabile gaussiana o normale* **standard**.

2.4 Valori attesi

In questa sezione ricordiamo come si può definire il valore atteso per variabili aleatorie generali, a partire dalla sua definizione per variabili aleatorie semplici. Per maggiori approfondimenti si rimanda, ad esempio, al libro di Billingsley [3] o a quello di Williams [19].

Definizione 2.6 (Valore atteso per variabili semplici). *Sia X una variabile aleatoria in $(\Omega, \mathcal{F}, \mathbb{P})$, non negativa e semplice, cioè come in Esempio 2.7,*

$$X(\omega) = \sum_{m \in \mathbb{N}} x_m \mathbb{I}_{H_m}(\omega), \quad \text{con } H_m \in \mathcal{F} \text{ per ogni } m \in \mathbb{N},$$

allora si definisce

$$\mathbb{E}[X] = \sum_{m \in \mathbb{N}} x_m \mathbb{P}(H_m).$$

Osservazione 2.1. *Ogni variabile aleatoria X in $(\Omega, \mathcal{F}, \mathbb{P})$, non negativa, ammette una successione di variabili aleatorie X_n, semplici e non negative, tali che*

$$0 \leq X_n(\omega) \leq X_{n+1}(\omega), \quad \text{e tali che } \lim_{n \to \infty} X_n(\omega) = X(\omega).$$

Infatti[13] *basta prendere*

$$X_n(\omega) = \sum_{m=0}^{n2^n-1} \frac{m}{2^n} \mathbb{I}_{H_m^{(n)}}(\omega) + n\mathbb{I}_{H_{n2^n}^{(n)}}(\omega) = \sum_{m=0}^{n2^n-1} \frac{m}{2^n} \mathbf{1}_{[\frac{m}{2^n}, \frac{m+1}{2^n})}(X(\omega)) + n\mathbf{1}_{[n,\infty)}(X(\omega)), \tag{2.12}$$

dove si è posto

$$H_m^{(n)} = X^{-1}\left(\left[\frac{m}{2^n}, \frac{m+1}{2^n}\right)\right) \in \mathcal{F} \text{ per } 0 \le m \le n2^n - 1, \quad H_{n2^n}^{(n)} = X^{-1}\big([n,\infty)\big),$$

e, per $A \in \mathcal{F}$,

$$\mathbb{I}_A(\omega) = 1 \quad se \quad \omega \in A \qquad e \qquad \mathbb{I}_A(\omega) = 0 \quad se \quad \omega \notin A,$$

ed infine, per $a < b$ *numeri reali,*

$$\mathbf{1}_{[a,b)}(x) = 1 \quad se \quad x \in [a,b) \qquad e \qquad \mathbf{1}_{[a,b)}(x) = 0 \quad se \quad x \notin [a,b).$$

È infine interessante notare che, posto $\lfloor x \rfloor$ *la parte intera inferiore*[14] *di* x*, si può riscrivere nel seguente modo*

$$X_n(\omega) = \frac{\lfloor 2^n X(\omega) \rfloor}{2^n} \wedge n.$$

Definizione 2.7 (Valore atteso per variabili nonnegative). *Sia* X *una variabile aleatoria in* $(\Omega, \mathcal{F}, \mathbb{P})$*, non negativa, si definisce*

$$\mathbb{E}[X] = \lim_{n\to\infty} \mathbb{E}[X_n],$$

dove $\{X_n;\ n \in \mathbb{N}\}$ *è la successione monotona definita come in (2.12) dell'Osservazione precedente. Il limite esiste ed è monotono, per la proprietà di monotonia del valore atteso, sulle variabili aleatorie semplici. Si noti bene che tale limite può valere anche* $+\infty$*, nel qual caso si dice che la variabile* X *ha valore atteso infinito.*

Arriviamo ora alla definizione generale del valore atteso:

Definizione 2.8 (Valore atteso per variabili generali). *Sia* X *una variabile aleatoria in* $(\Omega, \mathcal{F}, \mathbb{P})$*, Siano* $X^+ := X \vee 0$ *e* $X^- := (-X) \vee 0$*, le variabili aleatorie non negative, definite alla fine della sezione precedente. Si noti che* $X = X^+ - X^-$ *e che invece* $|X| = X^+ + X^-$*. Si definisce allora, se ha senso*[15]

$$\mathbb{E}[X] = \mathbb{E}[X^+] - \mathbb{E}[X^-].$$

---

[13] La monotonia della successione delle variabili aleatorie X_n è evidente:
• se $X_n(\omega) = m/2^n$, con $m < n2^n$, allora i soli casi possibili sono

$$X_{n+1}(\omega) = (2m)/2^{n+1} = m/2^n = X_n(\omega),$$

oppure

$$X_{n+1}(\omega) = (2m+1)/2^{n+1} = m/2^n + 1/2^{n+1} > X_n(\omega);$$

• se $X_n(\omega) = n$ allora $X_{n+1}(\omega)$ può assumere un valore compreso tra n ed $n+1$.

Per la convergenza basta osservare che, qualunque sia ω, pur di prendere n sufficientemente grande e in modo che $X(\omega) < n$, si ha che

$$0 \le X(\omega) - X_n(\omega) \le 1/2^n.$$

[14] La parte intera inferiore $\lfloor x \rfloor$ di x è quel numero intero k tale che $k \le x < k+1$.

[15] Si considera che la somma $\mathbb{E}[X^+] - \mathbb{E}[X^-]$ ha senso

1 se $\mathbb{E}[X^+] < \infty$, $\mathbb{E}[X^-] < \infty$, nel qual caso $\mathbb{E}[X] \in \mathbb{R}$ e inoltre si ha anche $\mathbb{E}[|X|] = \mathbb{E}[X^+] + \mathbb{E}[X^-] < \infty$;

2 se $\mathbb{E}[X^+] < \infty$, $\mathbb{E}[X^-] = \infty$, nel qual caso $\mathbb{E}[X] = -\infty$;

3 se $\mathbb{E}[X^+] = \infty$, $\mathbb{E}[X^-] < \infty$, nel qual caso $\mathbb{E}[X] = +\infty$;

Il caso che rimane escluso è quindi il caso in cui $\mathbb{E}[X^+] = \infty$, $\mathbb{E}[X^-] = \infty$, del resto si avrebbe la forma indeterminata $\infty - \infty$.

Se invece di usare la probabilità $\mathbb{P}$ si usa la probabilità condizionata ad un evento A, ovvero $\mathbb{P}(\cdot|A)$, allora si parla di ***valore atteso di*** X ***condizionato all'evento*** A e si usa la notazione

$$\mathbb{E}[X|A].$$

Ciò significa che, nel caso di una variabile aleatoria semplice

$$X(\omega) = \sum_{m\in\mathbb{N}} x_m \mathbb{I}_{H_m}(\omega), \quad \text{con } H_m \in \mathcal{F} \text{ per ogni } m \in \mathbb{N},$$

si ha

$$\mathbb{E}[X|A] = \sum_{m\in\mathbb{N}} x_m \mathbb{P}(H_m|A).$$

2.5 Appendice: variabili Gaussiane multidimensionali

Cominciamo con il definire una variabile aleatoria gaussiana standard unidimensionale:

Definizione 2.9. *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[16]

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

$$\mathbb{E}[Z^k] = \sigma^k\mathbb{E}[Y^k]$$

[16] 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 $span(\boldsymbol{a}_{(i)}, i = 1, \cdot, k)$.

e

$$\mathbb{E}[|Z|^k] = |\sigma|^k \mathbb{E}[|Y|^k].$$

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[17] 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 $\mathbb{E}[|Y|] = \sqrt{\frac{2}{\pi}}$.

***Concludiamo questo paragrafo con un lemma che riguarda il comportamento asintotico della funzione di sopravvivenza di una gaussiana standard e del modulo di una gaussiana standard.

Lemma 2.1. ‡*Sia Y una gaussiana standard, allora, posto $f_Y(y) = \frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}$, si ha, per $x>0$,*

$$\left(x+\frac{1}{x}\right)^{-1} f_Y(x) \le \mathbb{P}(Y>x) \le \frac{1}{x}\,f_Y(x), \qquad x>0, \tag{2.13}$$

$$\left(x+\frac{1}{x}\right)^{-1} f_{|Y|}(x) \le \mathbb{P}(|Y|>x) \le \frac{1}{x}\,f_{|Y|}(x), \qquad x>0, \tag{2.14}$$

$$\mathbb{P}(|Y|>x) \le e^{-\frac{x^2}{2}}, \qquad x>0. \tag{2.15}$$

[17]Si noti che dalle ultime due relazioni sui momenti si ottiene che

$$\mathbb{E}[|Y|^m] = (m-1)!!C_{(-1)^m}, \quad \text{con } C_{+1}=1, C_{-1}=\sqrt{\frac{2}{\pi}}.$$

Dimostrazione. La disuguaglianza (2.14) discende immediatamente dalla prima disuguaglianza (2.13), la quale equivale a

$$\left(x+\frac{1}{x}\right)^{-1}\frac{1}{\sqrt{2\pi}}\,e^{-\frac{x^2}{2}}\leq \mathbb{P}(Y>x)\leq \frac{1}{x}\,\frac{1}{\sqrt{2\pi}}\,e^{-\frac{x^2}{2}},$$

e discende dalla seguente relazione

$$\left(w+\frac{1}{w}\right)^{-1}e^{-\frac{w^2}{2}}\leq \int_w^{+\infty} e^{-\frac{z^2}{2}}\,dz\leq \frac{1}{w}\,e^{-\frac{w^2}{2}},\qquad w>0. \tag{2.16}$$

La disuguaglianza destra della (2.16) discende da

$$\int_w^{+\infty} e^{-\frac{z^2}{2}}\,dz\leq \frac{1}{w}\int_w^{+\infty} z\,e^{-\frac{z^2}{2}}\,dz=\frac{1}{w}\,e^{-\frac{w^2}{2}}\quad ,$$

Inoltre

$$\frac{d}{dw}\frac{1}{w}\,e^{-\frac{w^2}{2}}=-\left(1+\frac{1}{w^2}\right)\,e^{-\frac{w^2}{2}}$$

e quindi

$$\frac{1}{w}\,e^{-\frac{w^2}{2}}=\int_w^{+\infty}\left(1+\frac{1}{z^2}\right)\,e^{-\frac{z^2}{2}}\,dz\leq\left(1+\frac{1}{w^2}\right)\int_w^{+\infty} e^{-\frac{z^2}{2}}\,dz\quad ,$$

che prova l'altra disuguaglianza nella (2.16).

Infine, per provare la disuguaglianza (2.15), basta osservare che,

$$\begin{aligned}\mathbb{P}(|Y|>x)&=2\int_x^{+\infty}\frac{1}{\sqrt{2\pi}}\,e^{-\frac{y^2}{2}}\,dy=2\,e^{-\frac{x^2}{2}}\int_x^{+\infty}\frac{1}{\sqrt{2\pi}}\,e^{-\frac{y^2-x^2}{2}}\,dy\\&=2\,e^{-\frac{x^2}{2}}\int_x^{+\infty}\frac{1}{\sqrt{2\pi}}\,e^{-\frac{(y+x)(y-x)}{2}}\,dy=2\,e^{-\frac{x^2}{2}}\int_0^{+\infty}\frac{1}{\sqrt{2\pi}}\,e^{-\frac{(z+2x)z}{2}}\,dz\qquad\text{(essendo } x>0,\text{)}\\&\leq 2\,e^{-\frac{x^2}{2}}\int_0^{+\infty}\frac{1}{\sqrt{2\pi}}\,e^{-\frac{z^2}{2}}\,dz=e^{-\frac{x^2}{2}}.\end{aligned}$$

$\square$

2.6 Appendice: Spazi di variabili aleatorie

Sia dato uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$. L'insieme delle variabili aleatorie X, che sono $\mathcal{F}$-misurabili ed integrabili[18], cioè per le quali $\mathbb{E}[|X|] < \infty$, forma uno **spazio vettoriale reale**: se $\mathbb{E}[|X_i|] < \infty$, per $i = 1, 2$, con X_i $\mathcal{F}$-misurabili, allora la variabile aleatoria $a_1X_1 + a_2X_2$ è ancora $\mathcal{F}$-misurabile e inoltre

$$\mathbb{E}[|a_1X_1 + a_2X_2|] \leq \mathbb{E}[|a_1||X_1| + |a_2||X_2|] \leq |a_1|\mathbb{E}[|X_1|] + |a_2|\mathbb{E}[|X_2|] < \infty.$$

L'insieme di tali variabili aleatorie è indicato con $\mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$.

Questo spazio differisce dall'usuale spazio $L^1(\Omega, \mathcal{F}, \mathbb{P})$ dell'analisi, solo in quanto in quest'ultimo spazio si considerano equivalenti due variabili aleatorie X ed Y se e solo se $\mathbb{P}(X = Y) = 1$. Lo spazio $L^1(\Omega, \mathcal{F}, \mathbb{P})$ è uno **spazio metrico completo e separabile** rispetto alla distanza

$$d_{L^1}(X, Y) = \mathbb{E}[|X - Y|]$$

(cioè si passa alle classi di equivalenza modulo la relazione di equivalenza[19]: $X \sim Y$ se e solo se $\mathbb{E}[|X - Y|] = 0$, perché altrimenti $d(X, Y)$ non gode della proprietà delle metriche che $d(X, Y) = 0$ se e **solo se** $X = Y$).

Anche l'insieme delle variabili aleatorie X, che sono $\mathcal{F}$-misurabili e ***quadrato integrabili***, cioè per le quali $\mathbb{E}[|X|^2] < \infty$, forma uno **spazio vettoriale reale** che si indica con $\mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$. Come prima passando alle classi di equivalenza[20] si ottiene l'usuale spazio $L^2(\Omega, \mathcal{F}, \mathbb{P})$, che inoltre è uno spazio metrico completo e separabile rispetto alla distanza

$$d^2_{L^2}(X, Y) = \mathbb{E}[|X - Y|^2]$$

(modulo passare a classi di equivalenza $X \sim Y$ se e solo se $\mathbb{E}[|X - Y|^2] = 0$). Inoltre $L^2(\Omega, \mathcal{F}, \mathbb{P})$ è **uno spazio di Hilbert**: questo significa che è possibile introdurre un **prodotto scalare**

$$< X_1, X_2 >:= \mathbb{E}[X_1X_2],$$

e definire[21] $d^2_{L^2}(X, Y) =< X - Y, X - Y >= \mathbb{E}[|X - Y|^2]$. Si noti che, per la **disuguaglianza di Cauchy**,

$$|\mathbb{E}(X_1X_2)| \leq \mathbb{E}(|X_1X_2|) \leq \mathbb{E}^{1/2}(|X_1|^2)\mathbb{E}^{1/2}(|X_2|^2),$$

e quindi $< X_1, X_2 >$ è finito se $\mathbb{E}[|X_i|^2] < \infty$, per $i = 1, 2$.

Ovviamente si ha

$$\mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P}) \subseteq \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P}),$$

ed analogamente

$$L^2(\Omega, \mathcal{F}, \mathbb{P}) \subseteq L^1(\Omega, \mathcal{F}, \mathbb{P}),$$

Infatti se X è di quadrato integrabile allora se X è integrabile:

$$\mathbb{E}[X^2] < \infty \Rightarrow \mathbb{E}[|X|] < \infty,$$

come si vede immediatamente, ad esempio, con la disuguaglianza di Cauchy:

$$\mathbb{E}[|X|] = \mathbb{E}[1|X|] \leq \big(\mathbb{E}[1^2]\big)^{1/2}\big(\mathbb{E}[X^2]\big)^{1/2} = \big(\mathbb{E}[X^2]\big)^{1/2},$$

[18] Ovvero, in termini analitici,

$$\mathbb{E}[|X|] = \int_\Omega |X(\omega)|\, \mathbb{P}(d\omega) =< \infty.$$

[19] Se $\mathbb{E}[|X|] < \infty$ e $\mathbb{E}[|Y|] < \infty$, allora $\mathbb{P}(X = Y) = 1$ se e solo se $\mathbb{E}[|X - Y|] = 0$.

[20] Se $\mathbb{E}[|X|^2] < \infty$ e $\mathbb{E}[|Y|^2] < \infty$, allora $\mathbb{P}(X = Y) = 1$ se e solo se $\mathbb{E}[|X - Y|^2] = 0$.

[21] Va ricordato che è usuale indicare $d^2_{L^2}(X, Y) =< X - Y, X - Y >= \mathbb{E}[|X - Y|^2]$ con $d^2_{L^2}(X, Y) = \|X - Y\|^2_{L^2}$, o equivalentemente

$$\|X\| \, (= \|X\|_{L^2}) := \big(\mathbb{E}[|X|^2]\big)^{1/2}.$$

oppure direttamente considerando che $|x| \leq 1 + x^2$, e per la **proprietà di monotonia** del valore atteso

$$\mathbb{E}[|X|] \leq 1 + \mathbb{E}[X^2].$$

2.6.1 Sottospazi dello spazio delle v.a. di quadrato integrabile†

Sia $\mathcal{G}$ una sotto-σ-algebra di $\mathcal{F}$, ovvero $\mathcal{G} \subseteq \mathcal{F}$: se X è $\mathcal{G}$-misurabile, allora X è anche $\mathcal{F}$-misurabile[22].

Di conseguenza[23]

$$\mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P}) \subseteq \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P}),$$

ed analogamente[24]

$$L^2(\Omega, \mathcal{G}, \mathbb{P}) \subseteq L^2(\Omega, \mathcal{F}, \mathbb{P}),$$

Nel Capitolo 3 sui valori attesi condizionali, per ogni variabile aleatoria $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ verrà definita la variabile aleatoria $\mathbb{E}[X|\mathcal{G}]$, sostanzialmente come la proiezione di X sul sottospazio $L^2(\Omega, \mathcal{G}, \mathbb{P})$, ovvero come quella variabile aleatoria[25] $\tilde{X} \in L^2(\Omega, \mathcal{G}, \mathbb{P})$ per la quale

$$\mathbb{E}[(X - \tilde{X})^2] = \min_{Z \in L^2(\Omega, \mathcal{G}, \mathbb{P})} \mathbb{E}[(X - Z)^2]. \tag{2.20}$$

In particolare la proiezione sullo spazio generato dalla variabile aleatoria costante $\mathbf{1}$, $\omega \mapsto \mathbf{1}(\omega) = 1$, cioè prendendo $\mathcal{G} = \{\emptyset, \Omega\}$ si ottiene il valore atteso.

Nel caso in cui $\mathcal{G} = \sigma(Y)$, tenendo presente che le variabili aleatorie $\sigma(Y)$-misurabili sono le variabili aleatorie $Z = g(Y)$, con g boreliana, allora il valore atteso condizionato $\mathbb{E}[X|\sigma(Y)]$, secondo la precedente definizione (2.20) (si veda anche la successiva Definizione 3.2), è quella variabile aleatoria $\tilde{X} = \tilde{g}(Y)$, tale che

$$\mathbb{E}\big[\big(X - \tilde{g}(Y)\big)^2\big] = \min_{g \text{ boreliane}} \mathbb{E}\big[\big(X - g(Y)\big)^2\big],$$

dove il minimo va preso rispetto alle funzioni g per le quali $\mathbb{E}\big[|g(Y)|^2\big]$ è finito. In tale caso la media condizionata viene indicata con $\mathbb{E}[X|Y] = \tilde{g}(Y)$, invece che con $\mathbb{E}[X|\sigma(Y)]$.

2.6.2 Regressione lineare

Siano X ed Y due variabili aleatorie di quadrato integrabile. Il problema di trovare la retta di regressione lineare di X rispetto ad Y corrisponde al problema di trovare il punto di minimo, tra tutte le funzioni affini $a + \boldsymbol{b} \cdot y$ del valore atteso del quadrato della differenza tra X, ovvero trovare α e $\boldsymbol{\beta}$ tali che

$$\mathbb{E}\big[\big(X - (\alpha + \boldsymbol{\beta} \cdot Y)\big)^2\big] = \min_{a, \boldsymbol{b}} \mathbb{E}\big[\big(X - (a + \boldsymbol{b} \cdot Y)\big)^2\big].$$

[22]Infatti se la controimmagine di ogni aperto è un insieme (evento) di $\mathcal{G}$, allora ovviamente la controimmagine di ogni aperto è un insieme (evento) di $\mathcal{F}$.

[23]Si noti che, per maggiore precisione, si dovrebbe scrivere $\mathcal{L}^2(\Omega, \mathcal{G}, \widehat{\mathbb{P}})$ oppure $L^2(\Omega, \mathcal{G}, \widehat{\mathbb{P}})$ dove $\widehat{\mathbb{P}} = \mathbb{P}|_{\mathcal{G}}$, cioè la restrizione di $\mathbb{P}$ a $\mathcal{G}$.

[24]***Bisogna però fare una precisazione:
se lo spazio $L^2(\Omega, \mathcal{F}, \mathbb{P})$ viene considerato come spazio di Hilbert, siamo passati alle classi di equivalenza delle variabili aleatorie $\mathcal{F}$-misurabili e tali che $P(X \neq X') = 0$, ossia che differiscono su un insieme di $\mathcal{N}_{\mathcal{F}}$, la classe degli insiemi $\mathcal{F}$-misurabili di $\mathbb{P}$-probabilità nulla. Invece lo spazio $L^2(\Omega, \mathcal{G}, \mathbb{P})$, [o meglio $L^2(\Omega, \mathcal{G}, \widehat{\mathbb{P}})$, dove $\widehat{\mathbb{P}}$ è la misura $\mathbb{P}$ ristretta a $\mathcal{G}$,] è lo spazio delle classi di equivalenza delle variabili aleatorie $\mathcal{G}$-misurabili e tali che $P(X \neq X') = 0$, ossia che differiscono su un insieme di $\mathcal{N}_{\mathcal{G}}$, la classe degli insiemi $\mathcal{G}$-misurabili di $\mathbb{P}$-probabilità nulla. Quindi il passaggio agli spazi di Hilbert si può fare soltanto se gli insiemi $\mathcal{N}_{\mathcal{F}}$ e $\mathcal{N}_{\mathcal{G}}$ coincidono. Ciò è sicuramente vero se $\mathcal{N}_{\mathcal{F}} \subset \mathcal{G}$.***

[25]Sarebbe più corretto dire quella classe di equivalenza di variabili aleatorie.

Si noti che anche qui sia ha la proiezione su un sottospazio di $L^2(\Omega, \mathcal{F}, \mathbb{P})$. Questa volta però si tratta di un sottospazio di dimensione 2:

$$V = \{Z = a + bY;\ a, b \in \mathbb{R}\}$$

La soluzione è data in analogia con il caso delle proiezioni sui sottospazi vettoriali in $\mathbb{R}^d$. Si denota con $\hat{X} = \alpha + \boldsymbol{\beta} \cdot Y$ la variabile aleatoria che si cerca. Allora deve essere

$$\frac{\hat{X} - \mathbb{E}(X)}{\sigma_X} = \rho_{X,Y} \frac{Y - \mathbb{E}(Y)}{\sigma_Y} \tag{2.21}$$

dove $\rho_{X,Y}$ è il ***coefficiente di correlazione***, cioè

$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_X\, \sigma_Y}$$

L'analogia sta nel fatto che $\frac{Y-\mathbb{E}(Y)}{\sigma_Y}$ rappresenta il vettore unitario con "direzione" parallela ad Y, e invece $\rho_{X,Y}$ rappresenta il coseno tra l'angolo formato tra due vettori[26]:

$$\rho_{X,Y} = \frac{< X - \mathbb{E}(X), Y - \mathbb{E}(Y) >}{\|X - \mathbb{E}(X)\|\, \|Y - \mathbb{E}(Y)\|}$$

Dalla (2.21) si possono ottenere i valori di α e di $\boldsymbol{\beta}$. In particolare

$$\hat{X} = \frac{\rho_{X,Y}}{\sigma_Y}\, \sigma_X\, (Y - \mathbb{E}(Y)) + \mathbb{E}(X),$$

da cui immediatamente, tenendo conto della definizione di coefficiente di correlazione

$$\boldsymbol{\beta} = \frac{\rho_{X,Y}}{\sigma_Y}\, \sigma_X = \frac{Cov(X,Y)}{\sigma_Y^2}$$

Vedremo nel capitolo sui valori attesi condizionali che, per variabili aleatorie (X, Y) congiuntamente gaussiane il valore atteso condizionale $\mathbb{E}[X|Y]$ di X dato Y è una funzione affine di Y, e che quindi coincide anche con la retta di regressione.

[26]Si ricordi che se $\boldsymbol{u}$ e $\boldsymbol{v}$ sono due vettori non nulli, allora

$$\cos(\widehat{\boldsymbol{u}, \boldsymbol{v}}) = \frac{\boldsymbol{u} \cdot \boldsymbol{v}}{\|\boldsymbol{u}\| \cdot \boldsymbol{v}\|}$$

e che con la notazione $\|X\| := \left(\mathbb{E}[|X|^2]\right)^{1/2}$ si ha $Var(X) = \|X - \mathbb{E}(X)\|^{1/2}$.

2.7 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 2.10 (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 2.11 (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[27]: 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[28].

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" $-p$)

[27] 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*!!!

[28] 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***

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

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[29] 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***.

[29]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***[30] ***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[31].

[30] Vale un'interpretazione analoga per le probabilità condizionate ad un evento A, prendendo $X = I_E$.
[31] Come a dire che il contratto o la scommessa non hanno validità se non si verifica la condizione A.

Capitolo 3

Valori attesi e probabilità condizionali

3.1 Definizioni

Sia $(\Omega, \mathcal{F}, \mathbb{P})$ uno spazio di probabilità e sia X una variabile aleatoria. Si supponga di avere una sotto σ-algebra $\mathcal{G} \subseteq \mathcal{F}$. Cerchiamo una variabile aleatoria $\widetilde{X}$ che sia $\mathcal{G}$-misurabile e che "in qualche senso" abbia un comportamento simile a X. Vale la pena di ricordare che una σ-algebra può essere interpretata come "informazione" disponibile, e quindi cercare una variabile $\widetilde{X}$ che sia $\mathcal{G}$-misurabile con un comportamento simile ad X, significa cercare una variabile aleatoria che, sulla base dell'informazione disponibile $\mathcal{G}$, sia simile. Un altro modo di definire questa variabile $\widetilde{X}$ consiste, più banalmente, nel richiedere che sia una variabile aleatoria $\mathcal{G}$-misurabile "vicina" ad X. Naturalmente è necessario definire il senso di vicinanza, cioè quale metrica mettere sullo spazio delle variabili aleatorie.

Diamo ora due pre-definizioni, in cui però mancano le ipotesi da fare su X e delle precisazioni, affinché risultino definizioni ben poste.

pre-Definizione 1. *Si cerca una variabile aleatoria $\widetilde{X}$, $\mathcal{G}$-misurabile, per la quale valga*

$$\mathbb{E}[X|A] = \mathbb{E}[\widetilde{X}|A], \qquad \forall A \in \mathcal{G}, \text{ con } \mathbb{P}(A) > 0. \tag{3.1}$$

dove

$$\mathbb{E}[X|A] = \frac{\mathbb{E}[XI_A]}{\mathbb{P}(A)}.$$

Si noti che (3.1) equivale a

$$\mathbb{E}[XI_A] = \mathbb{E}[\widetilde{X}I_A], \qquad \forall A \in \mathcal{G}, \text{ con } \mathbb{P}(A) > 0. \tag{3.2}$$

e che quindi la richiesta che $\mathbb{P}(A) > 0$ si può omettere.

Si noti inoltre che la precedente (3.2) per $A = \Omega$, implica che, se una tale variabile aleatoria $\widetilde{X}$ esiste, allora $\mathbb{E}[\widetilde{X}] = \mathbb{E}[X]$.

pre-Definizione 2. *Si cerca una variabile aleatoria $\widetilde{X}$, $\mathcal{G}$-misurabile, per la quale valga*

$$\mathbb{E}[(X - Z)^2] \geq \mathbb{E}[(X - \widetilde{X})^2], \qquad \forall Z \ \mathcal{G} - \text{misurabile}. \tag{3.3}$$

Prima di tutto, dobbiamo trovare sotto quali condizioni le pre-definizioni siano ben poste, cioè, in questo caso, che esista una variabile $\widetilde{X}$ per cui valga la (3.1) (o equivalentemente la (3.2)) oppure valga la (3.3), ed in che senso ne viene individuata una sola. Notiamo che intanto ci sono delle condizioni necessarie da rispettare:

chiaramente, per la pre-definizione 1, è necessario che X sia una variabile aleatoria integrabile[1], cioè $\mathbb{E}[|X|] < \infty$, mentre, per la pre-definizione 2, è necessario richiedere che X sia di quadrato integrabile [2], cioè $\mathbb{E}[|X|^2] < \infty$, ed inoltre anche la (3.3) va modificata, nel senso che è necessario richiedere che anche Z sia di quadrato integrabile, oltre che $\mathcal{G}$-misurabile[3]. Inoltre è chiaro che se $\widetilde{X}'$ è una variabile aleatoria $\mathcal{G}$-misurabile, che differisce da $\widetilde{X}$ a meno di un insieme di misura nulla, anche $\widetilde{X}'$ gode della proprietà (3.2) o (3.3) rispettivamente e quindi non si individua una sola variabile aleatoria, ma una classe di variabili aleatorie. Queste modifiche in realtà sono sufficienti a garantire che le due pre-definizioni diventino due definizioni.

Definizione 3.1 (valore atteso condizionale 1). *Sia X una variabile aleatoria integrabile, cioè $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$. Sia data una variabile aleatoria (integrabile) $\widetilde{X}$, $\mathcal{G}$-misurabile, per la quale valga*

$$\mathbb{E}[X I_A] = \mathbb{E}[\widetilde{X} I_A], \qquad \forall A \in \mathcal{G}. \tag{3.4}$$

In questo modo si individua univocamente una classe di funzioni che si indica con $\mathbb{E}[X \mid \mathcal{G}]$ e che si chiama anche **media condizionale** *(o* **condizionata***) di X data $\mathcal{G}$. Si dice inoltre che $\widetilde{X}$ è una versione di $\mathbb{E}[X \mid \mathcal{G}]$.*

Definizione 3.2 (valore atteso condizionale 2). *Sia X una variabile aleatoria di quadrato integrabile, cioè $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$. Sia data una variabile aleatoria (di quadrato integrabile) $\widetilde{X}$, $\mathcal{G}-$misurabile, per la quale valga*[4]

$$\mathbb{E}[(X - Z)^2] \geq \mathbb{E}[(X - \widetilde{X})^2], \forall Z \quad \mathcal{G} - \textit{misurabile, con } \mathbb{E}[Z^2] < \infty. \tag{3.7}$$

In questo modo si individua univocamente una classe di funzioni che si indica con $\mathbb{E}[X \mid \mathcal{G}]$ e che si chiama anche **media condizionale** *(o* **condizionata***) di X data $\mathcal{G}$. Si dice inoltre che $\widetilde{X}$ è una versione di $\mathbb{E}[X \mid \mathcal{G}]$.*

Rimandiamo la verifica che effettivamente la Definizione 3.1 e la Definizione 3.2 sono ben poste a dopo aver trattato alcuni esempi, e anticipiamo che, se X è di quadrato integrabile[5], allora le Definizioni 3.1 e 3.2 sono

[1] L'insieme delle variabili aleatorie X, che sono $\mathcal{F}$-misurabili ed integrabili, cioè per le quali $\mathbb{E}[|X|] < \infty$, forma uno spazio vettoriale reale: se $\mathbb{E}[|X_i|] < \infty$, per $i = 1, 2$, con X_i $\mathcal{F}$-misurabili, allora la variabile aleatoria $a_1 X_1 + a_2 X_2$ è ancora $\mathcal{F}$-misurabile e inoltre

$$\mathbb{E}[|a_1 X_1 + a_2 X_2|] \leq \mathbb{E}[|a_1||X_1| + |a_2||X_2|] \leq |a_1|\mathbb{E}[|X_1|] + |a_2|\mathbb{E}[|X_2|] < \infty.$$

L'insieme di tali variabili aleatorie è indicato con $L^1(\Omega, \mathcal{F}, \mathbb{P})$ che è uno spazio metrico completo e separabile rispetto alla distanza $d(X, Y) = \mathbb{E}[|X - Y|]$ (modulo passare a classi di equivalenza $X \sim Y$ se e solo se $\mathbb{E}[|X - Y|] = 0$, perché altrimenti $d(X, Y)$ non gode della proprietà delle metriche che $d(X, Y) = 0$ se e **solo se** $X = Y$).

[2] Anche in questo caso l'insieme delle variabili aleatorie X, che sono $\mathcal{F}$-misurabili e quadrato integrabili, cioè per le quali $\mathbb{E}[|X|^2] < \infty$, forma uno spazio vettoriale reale che si indica con $L^2(\Omega, \mathcal{F}, \mathbb{P})$, che inoltre è uno spazio metrico completo e separabile rispetto alla distanza $d^2_{L^2}(X, Y) = \mathbb{E}[|X - Y|^2]$ (modulo passare a classi di equivalenza $X \sim Y$ se e solo se $\mathbb{E}[|X - Y|^2] = 0$). Inoltre $L^2(\Omega, \mathcal{F}, \mathbb{P})$ è uno spazio di Hilbert. Questo significa che è possibile introdurre un prodotto scalare

$$< X_1, X_2 >:= \mathbb{E}[X_1 X_2],$$

e definire $d^2_{L^2}(X, Y) =< X - Y, X - Y >= \mathbb{E}[|X - Y|^2]$. Si noti che, per la disuguaglianza di Cauchy,

$$|\mathbb{E}(X_1 X_2)| \leq \mathbb{E}(|X_1 X_2|) \leq \mathbb{E}^{1/2}(|X_1|^2)\mathbb{E}^{1/2}(|X_2|^2),$$

e quindi $< X_1, X_2 >$ è finito se $\mathbb{E}[|X_i|^2] < \infty$, per $i = 1, 2$.

[3] † Si noti che quindi deve essere $Z \in L^2(\Omega, \mathcal{G}, \widehat{\mathbb{P}})$ dove $\widehat{\mathbb{P}} = \mathbb{P}|_{\mathcal{G}}$, cioè la restrizione di $\mathbb{P}$ a $\mathcal{G}$.

[4] La proprietà (3.7) è equivalente alla proprietà

$$\mathbb{E}[(X - \widetilde{X})^2] = \min_{Z \in L^2(\Omega, \mathcal{G}, \widehat{\mathbb{P}})} \left(\mathbb{E}[(X - Z)^2]\right), \tag{3.5}$$

ovvero

$$d^2_{L^2}(X, \widetilde{X}) = \min_{Z \in L^2(\Omega, \mathcal{G}, \widehat{\mathbb{P}})} d^2_{L^2}(X, Z). \tag{3.6}$$

(vedere le note precedenti per la definizione di $L^2(\Omega, \mathcal{G}, \widehat{\mathbb{P}})$ e di $d^2_{L^2}$)

[5] Si ricordi che se X è di quadrato integrabile allora se X è integrabile:

$$\mathbb{E}[X^2] < \infty \Rightarrow \mathbb{E}[|X|] < \infty,$$

come si vede immediatamente, ad esempio, con la disuguaglianza di Cauchy:

$$\mathbb{E}[|X|] = \mathbb{E}[1|X|] \leq \left(\mathbb{E}[1^2]\right)^{1/2}\left(\mathbb{E}[X^2]\right)^{1/2} = \left(\mathbb{E}[X^2]\right)^{1/2},$$

oppure direttamente considerando che $|x| \leq 1 + x^2$, e per la proprietà di monotonia del valore atteso

$$\mathbb{E}[|X|] \leq 1 + \mathbb{E}[X^2].$$

equivalenti, e quindi non c'è ambiguità nello scegliere una definizione o l'altra e che in seguito, riferendoci a $\mathbb{E}[X \mid \mathcal{G}]$, intenderemo riferirci alla Definizione 3.1, che valendo per variabili aleatorie integrabili, è più generale. Non c'è quindi ambiguità nella seguente definizione che viene data in analogia con $\mathbb{P}(A) = \mathbb{E}[I_A]$.

Definizione 3.3 (probabilità condizionale di un evento). *Sia $A \in \mathcal{F}$ un evento e $\mathcal{G}$ una sotto $\sigma-$algebra di $\mathcal{F}$, allora*

$$\mathbb{P}(A \mid \mathcal{G}) := \mathbb{E}[I_A \mid \mathcal{G}]$$

è detta **probabilità condizionale** *di A data $\mathcal{G}$.*

Va inoltre sottolineato ***un abuso di notazione*** per cui si identifica la classe di equivalenza $\mathbb{E}[X \mid \mathcal{G}]$ e la variabile aleatoria $\widetilde{X}$ che ne è un rappresentante.

3.2 Esempi

Esempio 3.1. *Caso in cui $\mathcal{G} = \{\Omega, \emptyset\}$.*
In questo caso $\mathbb{E}[X \mid \mathcal{G}]$ si riduce al valore medio usuale, cioè

$$\mathbb{E}[X \mid \mathcal{G}] = \mathbb{E}[X],$$

in quanto ovviamente $\mathbb{E}[XI_\Omega] = \mathbb{E}[\mathbb{E}[X]I_\Omega] = \mathbb{E}[X] \cdot 1$, mentre banalmente $\mathbb{E}[XI_\emptyset] = \mathbb{E}[\mathbb{E}[X]I_\emptyset] = \mathbb{E}[X] \cdot 0$.

Esempio 3.2. *Caso in cui $\mathcal{G} = \mathcal{M} = \sigma\{H_m,\, m \in \mathbb{N}\}$ ed $\{H_m,\, m \in \mathbb{N}\}$ è una partizione.*

Intanto ricordiamo che in questo caso le variabili aleatorie $\mathcal{G}$-misurabili sono le funzioni del tipo

$$Z(\omega) = \sum_{m=1}^{\infty} c_m \mathbb{I}_{H_m}(\omega),$$

dove c_m sono costanti reali, e gli insiemi $\mathcal{G}$-misurabili sono le unioni di sottofamiglie numerabili di elementi della partizione. Basta quindi calcolare i valori $\widetilde{c}_m$ che caratterizzano $\widetilde{X}$, imponendo la condizione[6] *che*

$$\mathbb{E}[X\mathbb{I}_{H_n}] = \mathbb{E}[\widetilde{X}\mathbb{I}_{H_n}] = \mathbb{E}[\sum_{m=1}^{\infty} \widetilde{c}_m \mathbb{I}_{H_m}\mathbb{I}_{H_n}] = \mathbb{E}[\widetilde{c}_n \mathbb{I}_{H_n}] = \widetilde{c}_n \mathbb{P}[H_n] \tag{3.8}$$

Quindi

$$\widetilde{c}_m = \frac{\mathbb{E}[XI_{H_m}]}{\mathbb{P}(H_m)} \quad \text{se } \mathbb{P}(H_m) > 0 \qquad e \qquad \widetilde{X} = \mathbb{E}[X \mid \mathcal{G}] = \sum_{n\geq 1}^{*} \frac{\mathbb{E}[X\mathbb{I}_{H_n}]}{\mathbb{P}(H_n)} \mathbb{I}_{H_n},$$

dove $\sum_{n\geq 1}^{*}$ è la somma estesa agli indici n per cui $\mathbb{P}(H_n) > 0$.
Si noti l'abuso di notazione: in realtà

$$\mathbb{E}[X \mid \mathcal{G}] = \{\xi \text{ v.a. } \mathcal{G} - \text{misurabili, t.c. } \xi = \sum_{n\geq 1}^{*} \frac{\mathbb{E}[X\mathbb{I}_{H_n}]}{\mathbb{P}(H_n)} \mathbb{I}_{H_n} + \sum_{m\geq 1}^{**} c_m \mathbb{I}_{H_m}, \text{ per } c_m \in \mathbb{R}\}$$

dove $\sum_{m\geq 1}^{**}$ è la somma estesa agli indici m per cui $\mathbb{P}(H_m) = 0$.

---

[6] È chiaro che la condizione che $\mathbb{E}[X\mathbb{I}_A] = \mathbb{E}[\widetilde{X}\mathbb{I}_A]$ per ogni $A = \bigcup_{n\in I} H_n$, implica la condizione (3.8): basta prendere $A = H_n$. Tuttavia vale anche il viceversa, in quanto

$$\begin{aligned}\mathbb{E}[X\mathbb{I}_A] &= \mathbb{E}[X \sum_{n\in I} \mathbb{I}_{H_n}] = \sum_{n\in I} \mathbb{E}[X\mathbb{I}_{H_n}] \\ &= \sum_{n\in I} \mathbb{E}[\widetilde{X}\mathbb{I}_{H_n}] = \mathbb{E}[\widetilde{X} \sum_{n\in I} \mathbb{I}_{H_n}] = \mathbb{E}[\widetilde{X}\mathbb{I}_A].\end{aligned}$$

In particolare se $X = \mathbb{I}_B$, con $B \in \mathcal{F}$, e ponendo (come è usuale)

$$\mathbb{P}(B \mid \mathcal{G})(\omega) = \mathbb{E}[\mathbb{I}_B \mid \mathcal{G}](\omega),$$

otteniamo che una versione[7] *di $\mathbb{P}(B \mid \mathcal{G})(\omega)$ è data da*

$$\mathbb{P}(B \mid \mathcal{G}) = \sum_{n\geq 1}{}^{*} \frac{\mathbb{E}[\mathbb{I}_B \mathbb{I}_{H_n}]}{\mathbb{P}(H_n)} \mathbb{I}_{H_n} = \sum_{n\geq 1}{}^{*} \frac{\mathbb{P}(B \cap H_n)}{\mathbb{P}(H_n)} \mathbb{I}_{H_n} = \sum_{n\geq 1}{}^{*} \mathbb{P}(B \mid H_n) \mathbb{I}_{H_n}. \tag{3.9}$$

Ritroviamo quindi forse più chiaramente l'idea che se si verifica H_n cambiamo la probabilità prendendo $\mathbb{P}(B \mid H_n)$ al posto di $\mathbb{P}(B)$, che consideriamo se non abbiamo alcuna informazione (corrisponde al caso in cui la σ-algebra a nostra disposizione è quella banale).
Vale la pena di considerare il caso in cui B sia un evento della σ-algebra $\mathcal{G}$ a nostra disposizione: $B \in \mathcal{G}$, o equivalentemente $B = \cup_{n\in I} H_n$, per un insieme di indici I. A parole si tratta del caso in cui B è un evento completamente osservabile, ovverosia il caso in cui B sia un evento che possiamo conoscere perfettamente. In tale caso una versione di $\mathbb{P}(B \mid \mathcal{G})(\omega)$ è proprio la funzione indicatrice di B, ovvero $\mathbb{I}_B(\omega)$. Infatti $\mathbb{P}(B \mid H_n) = 1$ per $n \in I$ e $\mathbb{P}(H_n) > 0$, mentre $\mathbb{P}(B \mid H_n) = 0$ per $n \notin I$ e $\mathbb{P}(H_n) > 0$.

È interessante osservare che, se $\mathcal{F}$ è a sua volta generato da una partizione $\{K_\ell,\ \ell \in \mathbb{N}\}$, più fine[8] di $\{H_m,\ m \in \mathbb{N}\}$, e nel caso in cui $\mathbb{P}(H_n) > 0$ per ogni $n \in \mathbb{N}$, allora (3.9) definisce una probabilità su $\mathcal{F}$:
per iniziare

$$\mathbb{P}(\Omega \mid \mathcal{G})(\omega) = \sum_{n\geq 1}{}^{*} \frac{\mathbb{E}[\mathbb{I}_\Omega \mathbb{I}_{H_n}(\omega)]}{\mathbb{P}(H_n)} \mathbb{I}_{H_n} = \sum_{n\geq 1} \frac{\mathbb{E}[\mathbb{I}_\Omega \mathbb{I}_{H_n}(\omega)]}{\mathbb{P}(H_n)} \mathbb{I}_{H_n} = \sum_{n\geq 1} \mathbb{I}_{H_n} = 1,$$

inoltre, se $B = \bigcup_{\ell\in I_B} K_\ell$, allora

$$\begin{aligned}
\mathbb{P}(B \mid \mathcal{G})(\omega) &= \sum_{n\geq 1}{}^{*} \frac{\mathbb{E}[\mathbb{I}_B \mathbb{I}_{H_n}(\omega)]}{\mathbb{P}(H_n)} \mathbb{I}_{H_n} = \sum_{n\geq 1} \frac{\mathbb{E}[\mathbb{I}_B \mathbb{I}_{H_n}(\omega)]}{\mathbb{P}(H_n)} \mathbb{I}_{H_n} = \sum_{n\geq 1} \frac{\mathbb{P}(B \cap H_n)}{\mathbb{P}(H_n)} \mathbb{I}_{H_n}(\omega) \\
&= \sum_{n\geq 1} \sum_{\ell\in I_B} \frac{\mathbb{P}(K_\ell \cap H_n)}{\mathbb{P}(H_n)} \mathbb{I}_{H_n}(\omega) = \sum_{\ell\in I_B} \sum_{n\geq 1} \frac{\mathbb{P}(K_\ell \cap H_n)}{\mathbb{P}(H_n)} \mathbb{I}_{H_n}(\omega) = \sum_{\ell\in I_B} \mathbb{P}(K_\ell \mid \mathcal{G})(\omega),
\end{aligned}$$

e da questa relazione immediatamente si ricava che, qualunque sia ω, l'applicazione $B \mapsto \mathbb{P}(B \mid \mathcal{G})(\omega)$ definisce una probabilità. Anche nel caso in cui **non** si faccia l'ipotesi che $\mathbb{P}(H_n) > 0$ per ogni $n \in \mathbb{N}$, si può trovare una versione di $\mathbb{P}(B \mid \mathcal{G})(\omega)$ in modo che per ogni ω la funzione

$$\mathbb{P}(\cdot \mid \mathcal{G})(\omega) : \mathcal{F} \mapsto [0,1];\ B \mapsto \mathbb{P}(B \mid \mathcal{G})(\omega)$$

sia una probabilità:
fissata a piacere una probabilità $\mathbb{P}^0$ su $\mathcal{F}$[9], si definisce

$$\mathbb{P}(K_\ell \mid \mathcal{G})(\omega) = \sum_{n\geq 1}{}^{*} \frac{\mathbb{P}(K_\ell \cap H_n)}{\mathbb{P}(H_n)} \mathbb{I}_{H_n}(\omega) + \sum_{n\geq 1}{}^{**} \mathbb{P}^0(K_\ell) \mathbb{I}_{H_n}(\omega) = \tag{3.10}$$

$$= \sum_{n\geq 1}{}^{*} \frac{\mathbb{P}(K_\ell \cap H_n)}{\mathbb{P}(H_n)} \mathbb{I}_{H_n}(\omega) + \mathbb{P}^0(K_\ell) \mathbb{I}_{\{\cup^{**}_{n\geq 1} H_n\}}(\omega). \tag{3.11}$$

e

$$\mathbb{P}(B \mid \mathcal{G})(\omega) := \sum_{\ell\in I_B} \mathbb{P}(K_\ell \mid \mathcal{G})(\omega). \tag{3.12}$$

---

[7] La versione di $\mathbb{P}(B \mid \mathcal{G})(\omega)$ data in (3.9) non è in generale una probabilità per ogni ω: ad esempio se $\omega \in H_m$ e $\mathbb{P}(H_m) = 0$, allora $\mathbb{P}(\Omega \mid \mathcal{G})(\omega) = 0$ invece di 1.

[8] La partizione $\{K_\ell,\ \ell \in \mathbb{N}\}$ è più fine della partizione $\{H_m,\ m \in \mathbb{N}\}$ se e solo se per ogni $m \in \mathbb{N}$ esiste un $I_m \subseteq \mathbb{N}$ tale che

$$H_m = \bigcup_{\ell\in I_m} K_\ell.$$

[9] Ad esempio fissando una successione $\{p_\ell;\ \ell \in \mathbb{N}\}$, con $p_\ell \geq 0$, per ogni $\ell \in \mathbb{N}$ e $\sum_{\ell\in\mathbb{N}} p_\ell = 1$, in modo che $\mathbb{P}^0(K_\ell) = p_\ell$.

Si vede immediatamente che la parte a destra di (3.12) è effettivamente una versione di $\mathbb{P}(B \mid \mathcal{G})$, e si vede facilmente che in questo modo, qualunque sia ω, $\mathbb{P}(\cdot \mid \mathcal{G})(\omega)$ definisce una probabilità[10].

Inoltre questa probabilità gode della proprietà che se X è $\mathcal{F}$-misurabile, cioè se

$$X = \sum_{\ell \in \mathbb{N}} c_\ell \mathbb{I}_{K_\ell},$$

allora

$$\mathbb{E}[X \mid \mathcal{G}](\omega) = \sum_{\ell \in \mathbb{N}} c_\ell \mathbb{P}(K_\ell \mid \mathcal{G})(\omega) = \int_\Omega X(\omega') d\mathbb{P}(d\omega' \mid \mathcal{G})(\omega).$$

Per σ-algebre $\mathcal{F}$ più generali del caso di σ-algebre generate da una partizione, non è detto che queste proprietà valgano (per approfondimenti vedere la Sezione 4).

Esempio 3.3. *Ritorniamo nel caso dell'Esempio precedente, quando $\mathcal{G} = \sigma(Y)$, e Y è una variabile aleatoria discreta, a valori in $\{y_m;\, m \in \mathbb{N}\}$. Infatti allora*

$$\mathcal{G} = \sigma(\{H_m := Y^{-1}(\{y_m\});\; m \in \mathbb{N}\}),$$

e di conseguenza, 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.

---

[10] Infatti in questo caso è come se avessimo definito una successione $\{p_\ell(\omega); \ell \in \mathbb{N}\}$, con $\sum_{\ell \in \mathbb{N}} p_\ell(\omega) = 1$ per ogni ω

$$\begin{cases} p_\ell(\omega) = \mathbb{P}(K_\ell | H_m) & \text{se } \omega \in H_m, \text{ con } \mathbb{P}(H_m) > 0, \\ p_\ell(\omega) = \mathbb{P}^0(K_\ell) & \text{se } \omega \in H_m, \text{ con } \mathbb{P}(H_m) = 0, \end{cases}$$

e poi avessimo definito, per $B = \bigcup_{\ell \in I}$,

$$\mathbb{P}(B|\mathcal{G})(\omega) = \sum_{\ell \in I} p_\ell(\omega).$$

Così, ad esempio,

$$\mathbb{P}(\Omega \mid \mathcal{G})(\omega) := \sum_{\ell \in \mathbb{N}} \mathbb{P}(K_\ell \mid \mathcal{G})(\omega) = \begin{cases} \mathbb{P}(\Omega | H_m) = 1 & \text{se } \omega \in H_m, \text{ con } \mathbb{P}(H_m) > 0, \\ \mathbb{P}^0(\Omega) = 1 & \text{se } \omega \in H_m, \text{ con } \mathbb{P}(H_m) = 0. \end{cases}$$

Analogamente a quanto fatto nell'esempio precedente si ottiene che in tale caso, cioè se anche X *è una variabile aleatoria discreta a valori in* $\{x_n;\, n \in \mathbb{N}\}$*, allora*

$$\mathbb{E}[X \mid Y](\omega) = \sum_{n\in\mathbb{N}} x_n \mathbb{P}(X = x_n \mid \{Y = y\})\bigg|_{y=Y(\omega)}$$

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*

$$\mathbb{E}[h(X) \mid Y](\omega) = \sum_{n\in\mathbb{N}} h(x_n) \mathbb{P}(X = x_n \mid \{Y = y\})\bigg|_{y=Y(\omega)}$$

Nel caso in cui ***non*** *si abbia che* $\mathbb{P}(Y = y_m) > 0$ *per ogni* $m \in \mathbb{N}$*, si ottiene, fissata una probabilità* $\mathbb{P}^0$*, come nell'esempio precedente,*

$$\begin{aligned}\mathbb{E}[X \mid \sigma(Y)](\omega) &= \sum_{m\in\mathbb{N}}{}^{*} \mathbb{E}[X \mid \{Y = y_m\}] \mathbb{I}_{\{Y=y_m\}}(\omega) + \sum_{m\in\mathbb{N}}{}^{**} \mathbb{E}^0[X] \mathbb{I}_{\{Y=y_m\}}(\omega)\\ &= \sum_{m\in\mathbb{N}}{}^{*} \mathbb{E}[X \mid \{Y = y_m\}] \mathbb{I}_{\{y_m\}}(Y(\omega)) + \sum_{m\in\mathbb{N}}{}^{**} \mathbb{E}^0[X] \mathbb{I}_{\{y_m\}}(Y(\omega)),\end{aligned}$$

con il solito abuso di notazione (il secondo membro è un rappresentante della classe di equivalenza $\mathbb{E}[X \mid \sigma(Y)]$*).*

Quindi posto

$$\psi(y) = \sum_{m\in\mathbb{N}}{}^{*} \mathbb{E}[X \mid \{Y = y_m\}] \mathbb{I}_{\{y_m\}}(y) + \sum_{m\in\mathbb{N}}{}^{**} \mathbb{E}^0[X] \mathbb{I}_{\{y_m\}}(y),$$

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:*

$$\mathbb{E}[X \mid Y](\omega) = \quad \mathbb{E}[X \mid \{Y = y\}]\bigg|_{y=Y(\omega)} = \psi(Y(\omega)).$$

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 3.4. *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]

$$\widetilde{X}(\omega) = \int_{\mathbb{R}} x \left(I_{\{z: f_Y(z)>0\}}(y) \frac{f_{X,Y}(x, y)}{f_Y(y)} \right)\bigg|_{y=Y(\omega)} dx + \int_{\mathbb{R}} x \; I_{\{z: f_Y(z)=0\}}(y)\big|_{y=Y(\omega)} \; f_0(x) dx,$$

[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 $\mathbb{P}\big(Y(\omega) \in \{z : f_Y(z) = 0\}\big) = \int_{\{z: f_Y(z)=0\}} f_y(z)\, dz = \int_{\{z: f_Y(z)=0\}} 0\, dz = 0$.

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*

$$\mathbb{E}[XI_A] = \mathbb{E}[XI_B(Y)] = \int_{\mathbb{R}\times\mathbb{R}^d} xI_B(y) f_{X,Y}(x,y)\ dx\ dy.$$

Cominciamo con il 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],

$$\begin{aligned}
\mathbb{E}[\widetilde{X}(\omega)I_A] &= \int_{\mathbb{R}}\int_{\mathbb{R}^d} xI_B(y)\frac{f_{X,Y}(x,y)}{f_Y(y)} f_Y(y)\ dx\ dy \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}^d} xI_B(y) f_{X,Y}(x,y)\ dx\ dy.
\end{aligned}$$

D'altra parte, nel caso generale

$$\begin{aligned}
\mathbb{E}[\widetilde{X}(\omega)I_A] &= \mathbb{E}[\widetilde{X}(\omega)I_B(Y)] \\
&= \int_{\mathbb{R}^d}\left(\int_{\mathbb{R}} xI_{\{z: f_Y(z)>0\}}(y)\frac{f_{X,Y}(x,y)}{f_Y(y)}\ dx\right) I_B(y) f_Y(y)\ dy \\
&\quad + \int_{\mathbb{R}^d}\left(\int_{\mathbb{R}} xI_{\{z: f_Y(z)=0\}}(y) f_0(x)\ dx\right) I_B(y) f_Y(y)\ dy
\end{aligned}$$

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 $B \in \mathcal{B}(\mathbb{R}^d)$

$$\int_{\mathbb{R}}\int_{\mathbb{R}^d} xI_B(y) I_{\{z: f_Y(z)>0\}}(y) f_{X,Y}(x,y)\ dx\ dy = \int_{\mathbb{R}}\int_{\mathbb{R}^d} xI_B(y) f_{X,Y}(x,y)\ dx\ dy.$$

La verifica è immediata in quanto

$$\int_{\mathbb{R}}\int_{\mathbb{R}^d} xI_B(y) I_{\{z: f_Y(z)=0\}}(y) f_{X,Y}(x,y)\ dx\ dy = 0,$$

[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

$$\int_{R^{n_1}}\left(\int_{R^{n_2}} |\psi(x,y)| dy\right) dx < \infty \ \Leftrightarrow\ \int_{R^{n_2}}\left(\int_{R^{n_1}} |\psi(x,y)| dx\right) dy < \infty \ \Leftrightarrow\ \int_{R^{n_1}\times_R^{n_2}} |\psi(x,y)| dx dy < \infty.$$

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.

infatti, se $I_{\{z: f_Y(z)=0\}}(y) = 1$, *ovvero se* $f_Y(y) = 0 \left(= \int_{\mathbb{R}} f_{X,Y}(x,y)\, dx\right)$, *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 $\mathbb{R}$.

Esempio 3.5. *Caso in cui* $\mathcal{G} = \sigma(Y)$ *e* (X,Y) *è una variabile (congiuntamente) gaussiana bidimensionale, di media nulla. Come caso particolare dell'esempio precedente si ottiene*

$$\mathbb{E}[X \mid Y] = \frac{\sigma_{X,Y}}{\sigma_Y^2} Y.$$

A sua volta questo risultato si ottiene dal caso più in generale: $(X_1, \cdots, X_n)$ *è un vettore aleatorio gaussiano di media nulla e densità congiunta*

$$f(x_1, \cdots, x_n) = c \exp\{-\sum_{i,j}^{1,n} \alpha_{i,j} x_i x_j\} = c \exp\{-\sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_{i,j} x_i x_j\} \qquad \text{con } \alpha_{i,j} = \alpha_{j,i},\ \alpha_{n,n} > 0$$

e

$$\mathbb{E}[X_n \mid X_1, \cdots, X_{n-1}] = -\sum_{i=1}^{n-1} \frac{\alpha_{i,n}}{\alpha_{n,n}} X_i. \tag{3.13}$$

soluzione[13]*: Si tratta del caso* $d = n-1$ *con* $X = X_n$ *e* $Y = (X_1, \cdots, X_{n-1})$ *e con* $f_{X,Y}(x,y) > 0$. *Per calcolare* $\mathbb{E}[X_n \mid X_1, \cdots, X_{n-1}]$ *dobbiamo innanzitutto*

$$\frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{f_{Y,X}(y,x)}{f_Y(y)} = \frac{f_{X_1,\ldots,X_{n-1},X_n}(y,x)}{f_{X_1,\ldots,X_{n-1}}(y)}, \qquad \text{con } y = (x_1 \ldots, x_{n-1}).$$

Abbiamo quindi

$$f(x_1, \cdots, x_{n-1}, x_n) = c \exp\left\{ -\sum_{i,j}^{1,n-1} \alpha_{i,j} x_i x_j - x_n \left(\sum_{j=1}^{n-1} \alpha_{n,j} x_j\right) - \left(\sum_{i=1}^{n-1} \alpha_{i,n} x_i\right) x_n - \alpha_{n,n} x_n^2 \right\}$$

Tenendo presente che $\alpha_{i,j} = \alpha_{j,i}$, *e che* $\alpha_{n,n} > 0$ *si ha che*

$$m(y) = m(x_1, \ldots, x_{n-1}) := \sum_{j=1}^{n-1} \frac{\alpha_{n,j}}{\alpha_{n,n}} x_j = \sum_{i=1}^{n-1} \frac{\alpha_{i,n}}{\alpha_{n,n}} x_i,$$

e quindi che

$$f(x_1, \cdots, x_{n-1}, x_n) = f(y, x_n) = c \exp\left\{ -\alpha_{n,n} \left[\sum_{i,j}^{1,n-1} \frac{\alpha_{i,j}}{\alpha_{n,n}} x_i x_j + 2 x_n m(x_1, \ldots, x_{n-1}) + x_n^2 \right] \right\}$$

$$= c \exp\left\{ -\alpha_{n,n} \left[\sum_{i,j}^{1,n-1} \frac{\alpha_{i,j}}{\alpha_{n,n}} x_i x_j + m^2(y) + 2 x_n m(y) + x_n^2 - m^2(y) \right] \right\}$$

$$= c(y) \exp\left\{ -\alpha_{n,n} \left[m^2(y) + 2 x_n m(y) + x_n^2 \right] \right\} = c(y) \exp\left\{ -\alpha_{n,n} \left[x_n + m(y) \right]^2 \right\}$$

[13] Pur essendo possibile utilizzare la tecnica dell'esempio precedente si consiglia lo svolgimento dei calcoli dopo l'introduzione delle distribuzioni condizionali

Di conseguenza

$$\frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{c(y)}{f_Y(y)} \exp\left\{-\alpha_{n,n}\left[x_n - (-m(y))\right]^2\right\} = K(y)\exp\left\{-\frac{\left[x_n - (-m(y))\right]^2}{2\frac{1}{2\alpha_{n,n}}}\right\}.$$

Come osservato nel caso generale dell'esempio precedente, qualunque sia y, $\frac{f_{X,Y}(x,y)}{f_Y(y)}$ *deve essere una densità di probabilità: è quindi chiaro che deve coincidere con la densità di una variabile aleatoria gaussiana*[14] $N(-m(y), \frac{1}{2\alpha_{n,n}})$*, di media* $-m(y)$ *e di varianza* $\frac{1}{2\alpha_{n,n}}$*. Il valore atteso si calcola quindi come*

$$\begin{aligned}\mathbb{E}[X_n \mid X_1, \cdots, X_{n-1}] &= \int_{\mathbb{R}} xK(y)\exp\left\{-\frac{\left[x_n - (-m(y))\right]^2}{2\frac{1}{2\alpha_{n,n}}}\right\} dx\Bigg|_{y=(X_1,\ldots,X_{n-1})} \\ &= \; -m(y)\Big|_{y=(X_1,\ldots,X_{n-1})},\end{aligned}$$

da cui si ottiene (3.13).

È infine interessante notare che essendo l'espressione del valore condizionato $\mathbb{E}[X_n \mid X_1, \cdots, X_{n-1}]$*, una funzione lineare di* $X_1, \cdots, X_{n-1}$*, si ottiene che coincide con la retta di regressione*[15] *di* X_n *rispetto a* $X_1, \cdots, X_{n-1}$*.*

Esempio 3.6. *L'Esempio 3.4 si generalizza facilmente al caso in cui X è una variabile aleatoria a valori in* $\mathbb{R}^k$*, e si vuole calcolare il valore atteso condizionale* $\mathbb{E}[h(X) \mid Y]$*, dove* $h(\cdot)$ *è una funzione misurabile, a valori reali, ripetendo tutti i passaggi con i dovuti cambiamenti:*

$$\widetilde{h(X)}(\omega) = \mathbb{E}[h(X) \mid Y](\omega) = \int_{\mathbb{R}^k} h(x)\; f_{X|Y}(x|y)\big|_{y=Y(\omega)}\; dx$$

3.3 Proprietà del valore atteso condizionale

Enunciamo ora (senza dimostrarle, per il momento), le proprietà fondamentali della media condizionale (secondo la Definizione 3.1).

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

$$\mathbb{E}[aX + bY \mid \mathcal{G}] = a\mathbb{E}[X \mid \mathcal{G}] + b\mathbb{E}[Y \mid \mathcal{G}]$$

[14] La funzione $K(y)$ di y deve inoltre necessariamente valere:

$$K(y) = \frac{1}{\sqrt{2\pi\frac{1}{2\alpha_{n,n}}}} = \sqrt{\frac{\alpha_{n,n}}{\pi}}.$$

[15] Ricordiamo che il problema di trovare la retta di regressione lineare di X rispetto ad Y corrisponde al problema di trovare il punto di minimo, tra tutte le funzioni affini $\boldsymbol{a}\cdot y + b$ del valore atteso del quadrato della differenza tra X, ovvero trovare $\tilde{\boldsymbol{a}}$ e $\tilde{b}$ tali che

$$\mathbb{E}\left[\left(X - (\tilde{\boldsymbol{a}}\cdot Y + \tilde{b})\right)^2\right] = \min_{\boldsymbol{a},b}\mathbb{E}\left[\left(X - (\boldsymbol{a}\cdot Y + b)\right)^2\right].$$

Tenendo presente che le variabili aleatorie $\sigma(Y)$-misurabili sono le variabili aleatorie $Z = g(Y)$, con g boreliana, allora il valore atteso condizionato, secondo la Definizione 3.2, è quella variabile aleatoria $\mathbb{E}[X|Y] = \tilde{X} = \tilde{g}(Y)$ tale che

$$\mathbb{E}\left[\left(X - \tilde{g}(Y)\right)^2\right] = \min_{g \text{ boreliane}}\mathbb{E}\left[\left(X - g(Y)\right)^2\right],$$

di conseguenza si ha che se $\mathbb{E}[X|Y] = \tilde{X} = \tilde{\boldsymbol{a}}\cdot Y + \tilde{b}$ è una funzione affine, allora coincide anche con la retta di regressione.

2. Monotonia

$$\mathbb{P}(X \le Y) = 1 \text{ implica } \mathbb{P}(\mathbb{E}[X \mid \mathcal{G}] \le \mathbb{E}[Y \mid \mathcal{G}]\ ) = 1$$

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

$$\phi(\mathbb{E}[X \mid \mathcal{G}]) \le \mathbb{E}[\phi(X) \mid \mathcal{G}]$$

Osservazione In particolare per $\phi(x) = x^2$ ed X in $L^2(\Omega, \mathcal{F}, \mathbb{P})$, si ottiene che

$$(\mathbb{E}[X \mid \mathcal{G}])^2 \le \mathbb{E}[X^2 \mid \mathcal{G}],$$

e quindi, passando al valore atteso, che

$$\mathbb{E}[(\mathbb{E}[X \mid \mathcal{G}])^2] \le \mathbb{E}[\mathbb{E}[X^2 \mid \mathcal{G}]] = \mathbb{E}[X^2].$$

8. Convergenza sotto il segno di media condizionale, monotona e dominata

Se $\{X_n\}_{n\ge1}$ è una successione di variabili aleatorie non negative ed integrabili, convergente con probabilità 1 ad X, monotonamente, cioè $0 \le X_n \le X_{n+1}$, 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 $\mid X_n \mid \le Y$, allora

$$\mathbb{E}[X_n \mid \mathcal{G}] \to \mathbb{E}[X \mid \mathcal{G}], \text{ in } L^1.$$

A questo punto passiamo ad osservare che la Definizione 3.1 è ben posta, in quanto esiste almeno una variabile aleatoria che verifica la (3.4).

Dimostrazione. ‡ Si definisca infatti la misura $\nu(\cdot)$ su $\mathcal{G}$ come $\nu(A) := \mathbb{E}[XI_A]$, per $A \in \mathcal{G}$. La misura $\nu(\cdot)$ risulta assolutamente continua[16] rispetto a $\widehat{\mathbb{P}} = \mathbb{P}|_{\mathcal{G}}$, in quanto se $\widehat{\mathbb{P}}$(A)=$\mathbb{P}$(A)=0, allora $\nu(A) = 0$. Esiste quindi ***la derivata di Radon-Nikodym*** di ν rispetto a $\widehat{\mathbb{P}}$, cioè una funzione $f(\omega) = \frac{d\nu}{d\widehat{\mathbb{P}}}(\omega)$, $\mathcal{G}-$ misurabile, per la quale valga, qualunque sia $A \in \mathcal{G}$

$$\nu(A) = \int_A f(\omega) d\widehat{\mathbb{P}}(\omega) *** \qquad \text{cioè} \qquad \int_A X(\omega) d\mathbb{P}(\omega) = \int_A f(\omega) d\widehat{\mathbb{P}}(\omega)$$

ovvero

$$\mathbb{E}[XI_A] = \mathbb{E}[fI_A].$$

La derivata di Radon-Nikodym è definita a meno di insiemi di $\widehat{\mathbb{P}}$-misura nulla. Infatti se $g(\omega)$ è un'altra funzione $\mathcal{G}$-misurabile per la quale valga $\mathbb{E}[XI_A] = \mathbb{E}[gI_A]$ per ogni $A \in \mathcal{G}$, allora $\mathbb{E}[fI_A] = \mathbb{E}[gI_A]$ per ogni $A \in \mathcal{G}$, ed in particolare per $A = \{\omega : f(\omega) \geq g(\omega)\}$ e per $A = \{\omega : f(\omega) < g(\omega)\}$, per cui $\mathbb{E}[|\ f - g\ |] = 0$. Basta quindi prendere come $\widetilde{X}$ la derivata di Radon-Nikodym f o qualunque altra variabile aleatoria che differisca da $\widetilde{X}$ al più in un insieme ($\mathcal{G}$-misurabile) di probabilità nulla.

□

Anche la Definizione 3.2 è ben posta, per convincersene basta considerare che lo spazio $L^2(\Omega, \mathcal{F}, \mathbb{P})$, pensato come classi di equivalenza, è uno spazio di Hilbert con la norma $\|\ X\ \|^2 = \mathbb{E}[|\ X\ |^2]$ ed $L^2(\Omega, \mathcal{G}, \widehat{\mathbb{P}})$ è un suo sottospazio chiuso. Quindi la classe di equivalenza $\mathbb{E}[X \mid \mathcal{G}]$ della Definizione 3.2 è la proiezione di X su $L^2(\Omega, \mathcal{G}, \widehat{\mathbb{P}})$. ***A questo proposito va però ricordato quanto già detto in una nota della sezione 2.6.1: per passare agli spazi di Hilbert bisogna assicurarsi che $\mathcal{G}$ contenga gli insiemi $\mathcal{F}$-misurabili e di probabilità nulla. ***

Si noti la (3.4) potrebbe essere modificata come

$$\mathbb{E}[XW] = \mathbb{E}[\widetilde{X}W], \quad \forall \text{ v.a. } W \quad \mathcal{G} - \text{misurabile e per cui } XW \text{è integrabile.} \tag{3.14}$$

Dimostrazione. ‡ **Attenzione**: *la dimostrazione si basa sulle proprietà di linearità e di monotonia dei valori attesi condizionali, che dimostreremo più in là, basandoci solo sulla Definizione 3.1 e quindi sulla (3.4)*
Ovviamente (3.14) implica (3.4). Per mostrare il viceversa, basta considerare il caso in cui $W \geq 0$. In tale caso esiste una successione $W_n \uparrow$W, con W_n funzioni elementari, cioè combinazioni lineari di funzioni indicatrici. D'altra parte è facile vedere (confronta le proprietà di linearità e di monotonia, dimostrate nella sezione successiva) che, posto

$$X = X^+ - X^-, \qquad \text{con } X^+ = X \vee 0, \text{ e con } X^- = (-X) \vee 0,$$

si ha $\widetilde{X} = \widetilde{X^+} - \widetilde{X^-}$, per la proprietà di linearità, con $\widetilde{X^+}$ e $\widetilde{X^-} \geq 0$, per la proprietà di monotonia. Inoltre si ha che

$$\mathbb{E}[XW_n] = \mathbb{E}[X^+W_n] - \mathbb{E}[X^-W_n] = \mathbb{E}[\widetilde{X^+}W_n] - \mathbb{E}[\widetilde{X^-}W_n].$$

Per la proprietà di convergenza monotona dei valori attesi, (se $Z \geq 0$ allora $0 \leq ZW_n \leq ZW_{n+1} \uparrow ZW$), si ha

[16]† Date due misure ν e μ su una σ-algebra $\mathcal{G}$, si dice che ν è risulta assolutamente rispetto a μ se e solo se per ogni $A \in \mathcal{G}$ con $\mu(A) = 0$ risulta $\nu(A) = 0$. Questa proprietà si può anche esprimere dicendo che la famiglia $\mathcal{N}^\nu = \{A : \nu(A) = 0\}$ degli insiemi di ν-misura nulla contiene la famiglia $\mathcal{N}^\mu = \{A : \mu(A) = 0\}$ degli insiemi di μ-misura nulla. La ***derivata di Radon-Nikodym*** $\frac{d\nu}{d\mu}$ è una funzione h, che sia $\mathcal{G}$-misurabile, e per la quale valga

$$\nu(A) = \int_A h(\omega)\mu(d\omega) = \int_\Omega \mathbb{I}_A h(\omega)\mu(d\omega), \quad \text{per ogni } A \in \mathcal{G}.$$

$$\mathbb{E}[X^+W_n] \uparrow \mathbb{E}[X^+W], \qquad \mathbb{E}[X^-W_n] \uparrow \mathbb{E}[X^-W],$$

$$\mathbb{E}[\widetilde{X^+}W_n] \uparrow \mathbb{E}[\widetilde{X^+}W], \qquad \mathbb{E}[\widetilde{X^-}W_n] \uparrow \mathbb{E}[\widetilde{X^-}W],$$

quindi $\mathbb{E}[XW] = \mathbb{E}[\widetilde{X}W]$, in quanto

$$\mathbb{E}[XW] \leftarrow \mathbb{E}[XW_n] = \mathbb{E}[\widetilde{X}W_n] \rightarrow \mathbb{E}[\widetilde{X^+}W] - \mathbb{E}[\widetilde{X^-}W] = \mathbb{E}[\widetilde{X}W]$$

□

3.4 Equivalenza tra le definizioni di valore atteso condizionale per variabili aleatorie di quadrato sommabile

Mostreremo ora che, se X è di quadrato sommabile, allora ogni variabile aleatoria $\widetilde{X}_2$ che soddisfa la Definizione 3.2, soddisfa anche la Definizione 3.1. Per l'implicazione inversa, sempre nel caso in cui X sia di quadrato integrabile, abbiamo bisogno di alcune delle proprietà della media condizionale secondo la Definizione 3.1, enunciate precedentemente e che dimostreremo in seguito.

1) Se $\widetilde{X}_2$ è la (o meglio un rappresentante della) media condizionale di X secondo la Definizione 3.2, allora lo è anche secondo la Definizione 3.1.

Se infatti $\widetilde{X}_2$ soddisfa la condizione (3.7) allora, qualunque sia $Z \in L^2(\Omega, \mathcal{G}, \widehat{\mathbb{P}})$ la funzione

$$\phi_2(t) := \mathbb{E}[(X - \{(1-t)\widetilde{X}_2 + tZ\})^2] = \mathbb{E}[(X - \widetilde{X}_2)^2 - 2t(X - \widetilde{X}_2)(Z - \widetilde{X}_2) + t^2(Z - \widetilde{X}_2)^2]$$

ammette un minimo in $t = 0$, e quindi

$$\phi_2'(t)|_{t=0} = \frac{d}{dt}\mathbb{E}[(X - \{(1-t)\widetilde{X}_2 + tZ\})^2]\Big|_{t=0} = -2\mathbb{E}[(X - \widetilde{X}_2)(Z - \widetilde{X}_2)] = 0$$

Quindi posto $W = Z - \widetilde{X}_2$ deve valere, qualunque sia $W \in L^2(\Omega, \mathcal{G}, \widehat{\mathbb{P}})$

$$\mathbb{E}[XW] = \mathbb{E}[\widetilde{X}_2 W].$$

Considerando che tutte le funzioni indicatrici del tipo I_A, con $A \in \mathcal{G}$, sono in $L^2(\Omega, \mathcal{G}, \widehat{\mathbb{P}})$, otteniamo che se vale la (3.7) allora vale la (3.4).

2) Se $\widetilde{X}$ è la media condizionale di X secondo la Definizione 3.1 (o meglio ne è un rappresentante), allora lo è anche secondo la Definizione 3.2.

Infatti, allora $\widetilde{X}$ è di quadrato sommabile (confrontare l'osservazione alla disuguaglianza di Jensen, proprietà 7.) e quindi la seguente funzione è finita per ogni Z di quadrato sommabile

$$\phi(t) := \mathbb{E}[(X - \{(1-t)\widetilde{X} + tZ\})^2] = \mathbb{E}[(X - \widetilde{X})^2 - 2t(X - \widetilde{X})(Z - \widetilde{X}) + t^2(Z - \widetilde{X})^2]$$

Poiché si tratta di una parabola essa ammette un minimo nel punto $\bar{t}$, in cui la derivata si annulla:

$$\phi'(\bar{t}) = -2\mathbb{E}[(X - \widetilde{X})(Z - \widetilde{X})] + 2\bar{t}\,\mathbb{E}[(Z - \widetilde{X})^2] = 0$$

Posto $W = Z - \widetilde{X}$ e tenendo conto della (3.14), si ottiene $\phi'(\bar{t}) = 2\bar{t}\,\mathbb{E}[(Z - \widetilde{X})^2] = 0$, da cui, per l'arbitrarietà di Z, segue che $\bar{t} = 0$. Quindi $\widetilde{X}$ gode della proprietà che per ogni Z di quadrato sommabile $\phi(1) \geq \phi(0)$, ovvero la (3.7).

Osservazione 3.1. *Si noti l'analogia*[17] *con il problema classico di geometria euclidea del trovare la proiezione di un vettore* $\boldsymbol{x} \in \mathbb{R}^n$ *su un sottospazio vettoriale* $V \subset \mathbb{R}^n$ *come quel vettore* $\tilde{\boldsymbol{x}} \in V$ *che gode della proprietà di minimizzare la distanza da* V, *ovvero*

$$\|\boldsymbol{x} - \tilde{\boldsymbol{x}}\|^2 = \min_{\boldsymbol{z} \in V} \|\boldsymbol{x} - \boldsymbol{z}\|^2,$$

che è notoriamente equivalente alla condizione di ortogonalità

$$< \boldsymbol{x} - \tilde{\boldsymbol{x}}, \boldsymbol{z} >= 0 \text{ per ogni } \boldsymbol{z} \in V \quad \Leftrightarrow \quad < \boldsymbol{x}, \boldsymbol{z} >=< \tilde{\boldsymbol{x}}, \boldsymbol{z} > \text{ per ogni } \boldsymbol{z} \in V.$$

3.5 Dimostrazioni delle proprietà del valore atteso condizionale

Daremo ora le dimostrazioni delle proprietà enunciate nella sezione precedente utilizzando solo la Definizione 3.1.

Si noti che ciò permette di concludere che la dimostrazione dell'equivalenza delle Definizioni 3.1 e 3.2 è autocontenuta: in realtà andrebbero prima dimostrate le proprietà 1,2, e 7; successivamente la (3.14) e infine l'equivalenza delle Definizioni 3.1 e 3.2.

1. Linearità: $\mathbb{E}[aX + bY \mid \mathcal{G}] = a\mathbb{E}[X \mid \mathcal{G}] + b\mathbb{E}[Y \mid \mathcal{G}]$

La dimostrazione è ovvia, infatti se $\tilde{X}$ e $\tilde{Y}$ sono rappresentanti di $\mathbb{E}[X \mid \mathcal{G}]$ e di $\mathbb{E}[Y \mid \mathcal{G}]$, rispettivamente, ovvero se $\mathbb{E}[XI_A] = \mathbb{E}[\widetilde{X}I_A]$ e $\mathbb{E}[YI_A] = \mathbb{E}[\widetilde{Y}I_A]$ per ogni $A \in \mathcal{G}$, allora basta verificare che $\mathbb{E}[(aX + bY)I_A] = \mathbb{E}[(a\widetilde{X} + b\widetilde{Y})I_A]$, e ciò segue immediatamente da

$$\mathbb{E}[(aX + bY)I_A] = a\mathbb{E}[XI_A] + b\mathbb{E}[YI_A] = a\mathbb{E}[\widetilde{X}I_A] + b\mathbb{E}[\widetilde{Y}I_A] = \mathbb{E}[(a\widetilde{X} + b\widetilde{Y})I_A].$$

2. Monotonia: $\mathbb{P}(X \leq Y) = 1$ implica $\mathbb{P}(\mathbb{E}[X \mid \mathcal{G}] \leq \mathbb{E}[Y \mid \mathcal{G}]\ ) = 1$

Se $\mathbb{P}(X \leq Y) = 1$, allora $\mathbb{E}[(Y - X)I_A] = \mathbb{E}[(\widetilde{Y} - \widetilde{X})I_A] \geq 0$ per ogni $A \in \mathcal{G}$ e quindi in particolare per $A = \{\widetilde{X} < \widetilde{Y}\}$, si ottiene che $\mathbb{P}(\widetilde{X} \leq \widetilde{Y}) = 1$, ovvero $\mathbb{P}(\mathbb{E}[X \mid \mathcal{G}] \leq \mathbb{E}[Y \mid \mathcal{G}]) = 1$.

3. Formula dei condizionamenti successivi: se $\mathcal{G} \subseteq \mathcal{H}$, allora $\mathbb{E}[X \mid \mathcal{G}] = \mathbb{E}[\mathbb{E}[X \mid \mathcal{H}] \mid \mathcal{G}]$

Sia $\mathcal{G} \subseteq \mathcal{H}$ e siano $\widetilde{X}$ e $\widehat{X}$ tali che $\mathbb{E}[XI_A] = \mathbb{E}[\widetilde{X}I_A]$ per ogni $A \in \mathcal{G} \subseteq \mathcal{H}$ e $\mathbb{E}[XI_B] = \mathbb{E}[\widehat{X}I_B]$ per ogni $B \in \mathcal{H}$, ed in particolare per ogni $A \in \mathcal{G}$. Allora

$$\mathbb{E}[\widehat{X}I_A] = \mathbb{E}[\widetilde{X}I_A] (= \mathbb{E}[XI_A]), \quad \text{per ogni } A \in \mathcal{G}$$

e quindi $\widetilde{X} = \mathbb{E}[\widehat{X} \mid \mathcal{G}]$. La seconda parte deriva dal fatto che per la σ-algebra banale la media e la media condizionale coincidono.

4. Fattorizzazione: e Z è $\mathcal{G}$-misurabile e ZX è integrabile allora $\mathbb{E}[ZX \mid \mathcal{G}] = Z\mathbb{E}[X \mid \mathcal{G}]$.

Infatti allora $\mathbb{E}[XZI_A] = \mathbb{E}[\widetilde{X}ZI_A]$, per ogni $A \in \mathcal{G}$ e la funzione $Z\widetilde{X}$ è ovviamente $\mathcal{G}$-misurabile.

5. Condizionamento rispetto a σ-algebre indipendenti: se X e $\mathcal{G}$ sono indipendenti allora $\mathbb{E}[X \mid \mathcal{G}] = \mathbb{E}[X]$.

Infatti allora, per ogni $A \in \mathcal{G}$

$$\mathbb{E}[XZI_A] = \mathbb{E}[X]\mathbb{E}[ZI_A] = \mathbb{E}[Z\mathbb{E}[X]I_A]$$

e la funzione $Z\mathbb{E}[X]$ è ovviamente $\mathcal{G}$-misurabile.

[17] L'analogia si vede meglio se si sostituisce $\|\boldsymbol{x} - \boldsymbol{z}\|^2$ con $d_{L^2}(X, Z) = \mathbb{E}[|X - Z|^2]$ e $< \boldsymbol{x}, \boldsymbol{z} >$ con $\mathbb{E}[XZ]$.

6. Condizionamento ridondante: se X e $\mathcal{G}$ sono indipendenti da $\mathcal{H}$ allora, $\mathbb{E}[X \mid \mathcal{G} \vee \mathcal{H}] = \mathbb{E}[X \mid \mathcal{G}]$.

La dimostrazione segue da una leggera variante del Lemma di Dynkin[18] $\mathbb{E}[X \mid \mathcal{G} \vee \mathcal{H}] = \mathbb{E}[X \mid \mathcal{G}]$.
e si basa sul fatto la famiglia $\mathcal{A}$ di eventi del tipo $C = A \cap B$, con $A \in \mathcal{G}$ e $B \in \mathcal{H}$, formano un sistema chiuso rispetto all'intersezione e generano $\mathcal{G} \vee \mathcal{H}$,per il lemma di Dynkin ***basterà quindi verificare che le due misure $\overline{\nu}(C) := \mathbb{E}[XI_C]$ e $\overline{\mu}(C) := \mathbb{E}[\widetilde{X}I_C]$ coincidono per $C = A \cap B$, $A \in \mathcal{G}$ e $B \in \mathcal{H}$ ***

$$\begin{aligned} \mathbb{E}[XI_{A\cap B}] &= \mathbb{E}[XI_A I_B] \qquad \text{(per l'indipendenza di } \sigma(X) \vee \mathcal{G} \text{ da } \mathcal{H}) \\ &= \mathbb{E}[XI_A]\mathbb{E}[I_B] = \mathbb{E}[\widetilde{X}I_A]\mathbb{E}[I_B] = \mathbb{E}[\widetilde{X}I_A I_B] = \mathbb{E}[\widetilde{X}I_{A\cap B}]. \end{aligned} \tag{3.15}$$

***e che $\mathcal{A}$ contiene Ω (essendo ovviamente $\Omega = \Omega \cap \Omega$).

Nella dimostrazione abbiamo utilizzato l'indipendenza della σ-algebra $\sigma(X) \vee \mathcal{G}$, generata da $\sigma(X)$ e $\mathcal{G}$ da $\mathcal{H}$, ossia: per ogni evento $F \in \sigma(X) \vee \mathcal{G}$ e per ogni evento $H \in \mathcal{H}$ si ha $\mathbb{P}(F \cap H) = \mathbb{P}(F)\mathbb{P}(H)$. Attenzione, la sola indipendenza di $\sigma(X)$ da $\mathcal{H}$ e di $\mathcal{G}$ da $\mathcal{H}$ non è sufficiente[19] nel passaggio (3.15). ***

7. Disuguaglianza di Jensen[20]: se ϕ è una funzione convessa, e $\phi(X)$ è integrabile, allora $\phi(\mathbb{E}[X \mid \mathcal{G}]) \leq \mathbb{E}[\phi(X) \mid \mathcal{G}]$.

Per risultati classici di analisi ϕ è l'inviluppo delle sue tangenti (o sotto tangenti) ovvero esistono una successione di rette $\phi_n(x) = \alpha_n x + \beta_n$ per cui $\phi(x) = \sup_n\{\phi_n(x)\}$, per ogni x in $\mathbb{R}$. Per le proprietà **1.** di linearità e **2.** di monotonia si ha allora che

$$\phi_n(\mathbb{E}[X \mid \mathcal{G}]) = \alpha_n \mathbb{E}[X \mid \mathcal{G}] + \beta_n = \mathbb{E}[\phi_n(X) \mid \mathcal{G}] \leq \mathbb{E}[\phi(X) \mid \mathcal{G}];$$

[18] Ricordiamo l'enunciato del Lemma di Dynkin, che tra l'altro è la base per molti risultati di unicità,***ed è noto anche come teorema dell'unicità della misura.
***
*****Lemma** [Lemma di Dynkin, Billingsley 1984 [3]] *Sia $\mathcal{A}$ una famiglia di eventi che genera la σ-algebra $\mathcal{G}$ e che è chiusa rispetto alla intersezione finita.*
Se due misure di probabilità ν e μ coincidono su $\mathcal{A}$, allora le due misure di probabilità coincidono su $\mathcal{G}$.
Di conseguenza, se $\overline{\nu}$ e $\overline{\mu}$ sono due misure non negative e finite che coincidono su $\mathcal{A} \cup \{\Omega\}$ (cioè con $\overline{\nu}(\Omega) = \overline{\mu}(\Omega)$) allora le due misure coincidono su $\mathcal{G}$.
(la dimostrazione della affermazione sulle misure non negative riconduce subito al caso delle misure di probabilità, considerando le misure di probabilità $\nu(A) := \overline{\nu}(A)/\overline{\nu}(\Omega)$ e $\nu(A) := \overline{\mu}(A)/\overline{\mu}(\Omega)$.)***

[19] ***Il fatto che la sola indipendenza di $\sigma(X)$ da $\mathcal{H}$ e di $\mathcal{G}$ da $\mathcal{H}$ non è sufficiente si può mostrare con il seguente controesempio: $X = \mathbb{I}_E$, $\mathcal{G} = \{\emptyset, G\, G^c, \Omega\}$, $\mathcal{H} = \{\emptyset, H\, H^c, \Omega\}$, con E, G ed H tali che $\mathbb{P}(E \cap H) = \mathbb{P}(E)\mathbb{P}(H)$, $\mathbb{P}(G \cap H) = \mathbb{P}(G)\mathbb{P}(H)$, ma $\mathbb{P}(E \cap G \cap H) \neq \mathbb{P}(E \cap G)\mathbb{P}(H)$ (ad esempio si prenda $\Omega = (0,1) \times (0,1)$, $E = (0, \frac{1}{2}) \times (0,1)$, $G = (0, \frac{1}{4}) \times (0,1) \cup (\frac{1}{4}, \frac{1}{2}) \times (0, \frac{1}{2})$ e $H = (\frac{1}{4}, \frac{3}{4}) \times (0,1)$).***

[20] Si noti che nel caso particolare in cui $\mathcal{G}$ è la σ-algebra banale la disuguaglianza di Jensen per i valori attesi condizionali diviene l'usuale disuguaglianza di Jensen

$$\phi(\mathbb{E}[X]) \leq \mathbb{E}[\phi(X)].$$

Infine va notato che la precedente disuguaglianza di Jensen, nel caso particolare di una variabile aleatoria X semplice, a valori in $\{x_1, \ldots, x_n\}$, con $\mathbb{P}(\{X = x_i\}) = \lambda_i$ si scrive come

$$\phi(\mathbb{E}[X]) = \phi(\sum_{i=1}^{n} \lambda_i x_i) \leq \sum_{i=1}^{n} \lambda_i \phi(x_i) = \mathbb{E}[\phi(X)], \quad \text{, con } \lambda_i \geq 0,\ \sum_{i=1}^{n} \lambda_i = 1$$

La disuguaglianza interna è equivalente alla definizione di funzione convessa. Lo è esattamente nel caso $n = 2$

$$\phi(\lambda x_1 + (1-\lambda)x_2) \leq \lambda\phi(x_1) + (1-\lambda)\phi(x_2)$$

Il caso generale si ottiene per induzione su n: Se

$$\phi(\sum_{i=1}^{n} \lambda_i x_i) \leq \sum_{i=1}^{n} \lambda_i \phi(x_i), \text{ per ogni } \lambda_i \geq 0,\ \sum_{i=1}^{n} \lambda_i = 1, \text{ e per ogni } x_i$$

allora

$$\phi(\sum_{i=1}^{n+1} \mu_i x_i) \leq \sum_{i=1}^{n+1} \mu_i \phi(x_i), \text{ per ogni } \mu_i \geq 0,\ \sum_{i=1}^{n+1} \mu_i = 1, \text{e per ogni } y_i$$

passando all'estremo superiore su n si ottiene

$$\sup_n\{\phi_n(\mathbb{E}[X \mid \mathcal{G}])\} = \phi(\mathbb{E}[X \mid \mathcal{G}]) \leq \mathbb{E}[\phi(X) \mid \mathcal{G}].$$

Si noti che nella dimostrazione vengono utilizzate solo le proprietà di linearità e di monotonia del valore atteso condizionato.

8. Convergenza sotto il segno di media condizionale, monotona e dominata

8i) Se $\{X_n\}_{n\geq 1}$ è una successione di variabili aleatorie non negative ed integrabili, convergente con probabilità 1 ad X, **monotonamente**, cioè la successione $\{X_n\}_{n\geq 1}$ ha la proprietà che $0 \leq X_n \leq X_{n+1}$, allora $\mathbb{E}[X_n \mid \mathcal{G}] \nearrow \mathbb{E}[X \mid \mathcal{G}]$, con probabilità 1.

Dalla proprietà di monotonia si ottiene che se $\{X_n\}_{n\geq 1}$ è una successione monotona allora anche la successione delle medie condizionate $\widetilde{X}_n = \mathbb{E}[X_n \mid \mathcal{G}]$ è una successione monotona, ed è quindi convergente ad una variabile aleatoria $\widetilde{Z}$, $\mathcal{G}$-misurabile. (È importante notare che per ogni n esiste un insieme $A_n \in \mathcal{G}$ di probabilità nulla nel cui complementare vale $\widetilde{X}_n \leq \widetilde{X}_{n+1}$, che queste disuguaglianze valgono contemporaneamente nel complementare di $A = \bigcup_{n\geq 1} A_n$, ed infine che A ha ancora misura nulla).
Per mostrare che la successione converge ad un rappresentante di $\mathbb{E}[X \mid \mathcal{G}]$, basta notare che $\mathbb{E}[\widetilde{X}_n I_A] = \mathbb{E}[X_n I_A] \uparrow \mathbb{E}[X I_A]$, per la convergenza monotona di $\{X_n\}$ a X (e quindi di $\{X_n I_A\}$ a $X I_A$), e che $\mathbb{E}[\widetilde{X}_n I_A] \uparrow \mathbb{E}[\widetilde{Z} I_A]$, per la convergenza monotona di $\{\widetilde{X}_n\}$ a $\widetilde{Z}$ (e quindi di $\{\widetilde{X}_n I_A\}$ a $\widetilde{Z} I_A$). Di conseguenza $\mathbb{E}[X I_A] = \mathbb{E}[\widetilde{Z} I_A], \forall A \in \mathcal{G}$.

8ii) Se $\{X_n\}_{n\geq 1}$ è una successione di variabili aleatorie integrabili, convergente con probabilità 1 ad X, **dominatamente**, cioè $\mid X_n \mid \leq Y$, per una variabile aleatoria Y integrabile, allora si ha che $\mathbb{E}[X_n \mid \mathcal{G}] \to \mathbb{E}[X \mid \mathcal{G}]$, in L^1.
La dimostrazione (‡) relativa alla convergenza dominata si basa sulla seguente osservazione: se $X_n \to X$ q.c. e $\mid X_n \mid \leq Y$ allora $X_n \to X$ in L^1 (infatti allora $\mid X_n - X \mid \leq \mid Y \mid + \mid X \mid$ e quindi $\mid X_n - X \mid \to 0$ dominatamente e perciò $\mathbb{E}[\mid X_n - X \mid] \to 0$). Di conseguenza per le proprietà di linearità e per la disuguaglianza di Jensen applicata alla funzione $\phi(x) = \mid x \mid$ e alla variabile aleatoria $X_n - X$,

$$\mid \mathbb{E}[X_n \mid \mathcal{G}] - \mathbb{E}[X \mid \mathcal{G}] \mid = \mid \mathbb{E}[X_n - X \mid \mathcal{G}] \mid \leq \mathbb{E}[|X_n - X| \mid \mathcal{G}].$$

Passando ai valori medi

$$\mathbb{E}[\mid \mathbb{E}[X_n \mid \mathcal{G}] - \mathbb{E}[X \mid \mathcal{G}] \mid] = \mathbb{E}[\mid \mathbb{E}[X_n - X \mid \mathcal{G}] \mid] \leq \mathbb{E}[\mathbb{E}[\mid X_n - X \mid \mid \mathcal{G}]] = \mathbb{E}[\mid X_n - X \mid] \to 0.$$

Osservazione. Dalla dimostrazione è chiaro che basta la convergenza $X_n \to X$ in L^1, per ottenere la convergenza delle rispettive medie condizionali in L^1. In realtà, nel caso di convergenza dominata, c'è anche la convergenza puntuale delle medie condizionali. La dimostrazione di questo fatto è rimandata a dopo aver mostrato l'esistenza di distribuzioni condizionali regolari.

3.6 Probabilità condizionali regolari ‡

Il problema è il seguente:

È possibile trovare una versione di $\mathbb{P}(A \mid \mathcal{G})$ in modo che l'applicazione

$$\mathbb{P}(\cdot \mid \mathcal{G})(\omega):\ \mathcal{F} \to [0,1]; A \to \mathbb{P}(A \mid \mathcal{G})(\omega)$$

sia una probabilità per ogni ω?

Ad un primo sguardo superficiale sembrerebbe di sì:
Ovviamente $\mathbb{P}(\Omega \mid \mathcal{G})(\omega)$=1. Dalla proprietà di monotonia si ha che $\mathbb{P}(C \mid \mathcal{G}) \in [0,1]$, per ogni evento $C \in \mathcal{F}$. Dalla proprietà di convergenza monotona, applicata alla successione $X_n = \sum_{k=1}^n I_{A_k}$, dove $\{A_n\}$ è una successione di eventi di $\mathcal{F}$, oppure di una σ-algebra $\mathcal{A} \subseteq \mathcal{F}$, si ottiene che, comunque scelta una successione di eventi di $\mathcal{A}$, disgiunti a due a due, vale

$$\mathbb{P}(\bigcup_{n\geq 1} A_n \mid \mathcal{G}) = \sum_{n\geq 1} \mathbb{P}(A_n \mid \mathcal{G})$$

nel senso che al più differiscono per un insieme di misura nulla N.

Sembrerebbe quindi tutto funzionare. Il problema sta nel fatto che l'insieme N può dipendere però, in generale, dalla particolare successione $\{A_n, n \geq 1\}$ scelta e l'unione su tutte le successioni possibili è un'unione non numerabile, quindi non è detto che sia un evento e, anche se lo fosse, non è detto che sia di probabilità nulla. È proprio questo che in generale impedisce di affermare che $A \mapsto \mathbb{P}(A \mid \mathcal{G})(\omega)$ è una probabilità.

Lo stesso tipo di problema si pone nel caso in cui invece si cerchi una versione di $\mathbb{P}(A \mid \mathcal{G})(\omega)$ per $A \in \mathcal{A} \subseteq \mathcal{F}$, mentre non c'è nessun problema del tipo precedente se $\mathcal{A}$ è un'algebra finita (e quindi una $\sigma-$algebra).
Comunque, nel caso in cui sia possibile trovare un **nucleo di misure di probabilità**, cioè una famiglia

$$\mathbb{Q}(\cdot,\cdot):\ \mathcal{A} \times \Omega \to [0,1]; (A,\omega) \mapsto \mathbb{Q}(A,\omega)$$

1) per ogni $\omega \in \Omega$, $\mathbb{Q}(\cdot,\omega)$ è una misura di probabilità su $(\Omega,\mathcal{A})$

2) per ogni $C \in \mathcal{A}$, $\mathbb{Q}(C,\cdot)$ è $\mathcal{G}$-misurabile ed è una versione di $\mathbb{P}(C \mid \mathcal{G})(\omega)$,

allora si dice che $\mathbb{Q}(\cdot,\omega)$ è una **versione regolare** di $\mathbb{P}(\cdot \mid \mathcal{G})(\omega)$, cioè **delle probabilità condizionali**.

L'interesse di tali versioni deriva dal fatto che allora, per ogni v.a. Z, $\mathcal{A}$-misurabile e integrabile, vale

$$\mathbb{E}(Z \mid \mathcal{G})(\omega) = \int_\Omega Z(\omega')\mathbb{Q}(d\omega',\omega). \tag{3}$$

Per convincersene basta capire che ciò è vero per $Z = I_C$, e quindi per ogni funzione elementare, cioè combinazione lineare di funzioni indicatrici. Quindi (3) vale per ogni v.a. non negativa integrabile, in quanto limite monotono di funzioni elementari, e quindi per ogni v.a. integrabile in quanto differenza di due v.a. non negative ed integrabili.

L'interesse per l'esistenza di una versione regolare delle probabilità condizionali sta anche nel fatto che tutte le dimostrazioni delle proprietà delle medie condizionali sarebbero immediate (in particolare la disuguaglianza di Jensen e la convergenza monotona e dominata, anche nella versione dell'osservazione relativa, con la convergenza puntuale).

Non sempre, purtroppo, si hanno versioni regolari delle probabilità condizionali. In generale dipende dalla σ-algebra $\mathcal{A}$, qui di seguito viene dimostrato che ciò è vero se $\mathcal{A} = \sigma(X)$, dove X è una variabile aleatoria a valori in $\mathbb{R}$, cioè se $C = \{X \in H\}$, $H \in \mathcal{B}(\mathbb{R})$. In realtà si può vedere che ciò è vero anche se X è una variabile aleatoria d-dimensionale, oppure se X è una variabile aleatoria a valori in uno spazio metrico S, completo e separabile (ovvero uno spazio polacco). In questi casi si ottiene anche una probabilità sullo spazio S degli stati di X e si parla di distribuzione condizionale invece che di probabilità condizionale, che invece è una misura di probabilità su Ω.

Proposizione 3.1. *Sia X una variabile aleatoria reale in $(\Omega, \mathcal{F}, \mathbb{P})$, sia $\mathcal{G}$ una sotto σ-algebra di $\mathcal{F}$, e sia $\mathcal{A} = \sigma(X) = \big\{A = \{\omega \in \Omega,\ t.c.\, X(\omega) \in H\},\ H \in \mathcal{B}(\mathbb{R})\big\}$. Allora esiste una versione regolare $\mathbb{Q}(\cdot, \cdot)$ di $\mathbb{P}(\cdot|\mathcal{G})(\omega)$.*

Dimostrazione. L'idea è molto semplice:

1. Si costruisce una funzione $\hat{F}(\cdot, \cdot) : \mathbb{R} \times \Omega \mapsto [0, 1]$ che, per ogni ω, $\hat{F}(\cdot, \omega) : \mathbb{R} \mapsto [0, 1]$ soddisfa tutte le proprietà di una funzione di ripartizione, e che è una versione della probabilità condizionale $\mathbb{P}(X \leq s \mid \mathcal{G})(\omega)$. Indicheremo tale funzione con $F(s \mid \mathcal{G})(\omega)$, per ricordare questa proprietà.

dimostrazione di 1. Si inizia considerando $F(t \mid \mathcal{G})(\omega) := \mathbb{P}(X \leq t \mid \mathcal{G})(\omega)$ per t razionale. Per la proprietà della monotonia, possiamo prendere una versione per cui, per ogni $t_1 \leq t_2$ razionali, $F(t_1 \mid \mathcal{G})(\omega) \leq F(t_2 \mid \mathcal{G})(\omega)$, in un evento Ω_0 di probabilità 1: in questo caso l'evento in cui ciò non si verifica è un'unione numerabile di eventi di probabilità nulla. Per $\omega \in \Omega_0^c$ si definisce $F(t \mid \mathcal{G})(\omega) = F_0(t)$, con F_0 una fissata funzione di distribuzione, ad esempio (sempre sui razionali). Su tale evento Ω_0 esiste, per monotonia, il limite di $F(t \mid \mathcal{G})(\omega)$ sia per $t \to +\infty$ che per $t \to -\infty$, sempre per t razionale. Tali limiti sono rispettivamente uguali ai limiti di $F(n \mid \mathcal{G})(\omega)$ e di $F(-n \mid \mathcal{G})(\omega)$, e, di nuovo a parte un insieme di probabilità nulla, coincidono rispettivamente con $1 = \mathbb{P}(X < +\infty \mid \mathcal{G})$ e $0 = \mathbb{P}(X < -\infty \mid \mathcal{G})$.(Di nuovo su tale insieme si definisca $F(t \mid \mathcal{G})(\omega) = F_0(t)$)

Per ogni valore s reale, ma non razionale, si definisce una successione $t_n \downarrow s$. Per la proprietà della convergenza monotona (applicato a $1 - I_{\{X \leq t_n\}}$) si ottiene che il limite di $F(t_n \mid \mathcal{G})(\omega)$ esiste ed è una versione di $\mathbb{P}(X \leq s \mid \mathcal{G})(\omega)$. In questo modo si è ottenuta, per ogni ω una funzione $F(s \mid \mathcal{G})(\omega)$, definita su tutti i reali, e che soddisfa tutte le proprietà di una funzione di ripartizione, come si può vedere facilmente.

2. Ad ogni ω è associata una misura $\mu_X(\cdot \mid \mathcal{G})(\omega)$ di probabilità, per cui $\mu_X((-\infty, t] \mid \mathcal{G})(\omega) = F(t \mid \mathcal{G})(\omega)$, per ogni t reale.
Come conseguenza del punto **2.**, per ogni ***$H \in \mathcal{B}(\mathbb{R})$, la misura ***$\mu_X(H \mid \mathcal{G})(\omega)$ è una versione di $\mathbb{P}(X \in H)$, oltre ad essere una misura.
La misura $\mu_X(\cdot \mid \mathcal{G})(\omega)$ è detta appunto ***la distribuzione condizionale di X data $\mathcal{G}$***.

3. Per $A \in \mathcal{A}$, con $A = \{X \in H\}$, si definisce

$$\mathbb{Q}(A, \omega) := \mu_X(H \mid \mathcal{G})(\omega),$$

e risulta la versione regolare delle probabilità cercata. $\square$

Si faccia attenzione: la distribuzione condizionale $\mu_X(\cdot \mid \mathcal{G})(\omega)$, definita nel punto **2.** della dimostrazione precedente, è una probabilità su $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, mentre $\mathbb{Q}(\cdot, \omega)$ è una misura su $(\Omega, \sigma(X))$. Comunque accade che $\mathbb{P}(\{X \in H\} \mid \mathcal{G})(\omega) = \mu_X(H \mid \mathcal{G})(\omega)$. Inoltre poiché le v.a. $\sigma(X)-$misurabili sono le variabili aleatorie del tipo $Z = f(X)$, con f boreliana, si ha che

$$\mathbb{E}(f(X) \mid \mathcal{G})(\omega) = \mu_X(f \mid \mathcal{G})(\omega),$$

dove, in generale,

$$\mu(f) := \int f(x) d\mu(x).$$

Per la verità esiste una generalizzazione a tutti gli ***spazi di Borel*** $(S, \mathcal{S})$, cioè quegli spazi misurabili $(S, \mathcal{S})$ per cui esiste un insieme $E \in \mathcal{B}(\mathbb{R})$ e una funzione biunivoca $\phi : S \to E$ che sia $(S, \mathcal{S}) - (E, \mathcal{B}(\mathbb{R})|_E)$ misurabile, e la cui inversa sia $(E, \mathcal{B}(\mathbb{R})|_E) - (S, \mathcal{S})$ misurabile.

In particolare quindi il risultato di esistenza della versione regolare è valido per $(S,\mathcal{S}) = (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ o $(S,\mathcal{S}) = (\mathbb{R}^{\mathbb{N}}, \mathcal{R}^{\mathbb{N}})$, che sono spazi di Borel (vedere ad esempio P. Billingsley [3] "Convergence of Probability measures" pag. 218 e seguenti).

Prima della dimostrazione va notato che questo fatto permette di applicare il risultato anche al caso in cui siano coinvolte le variabili aleatorie X_n, $n \geq 1$, ed X (confrontare di nuovo l'osservazione alla proprietà della convergenza dominata per i valori attesi condizionali).

La dimostrazione è basata sull'osservazione che se Y è una variabile aleatoria a valori in S, allora $X := \phi(Y)$ è una variabile aleatoria reale,e quindi esiste una distribuzione condizionale $\mu_X(\cdot \mid \mathcal{G})$. La distribuzione condizionale di Y su $(S,\mathcal{S})$ si ottiene, per $J \in \mathcal{S}$, come

$$\nu_Y(J \mid \mathcal{G})(\omega) = \mu_X(\phi(J) \mid \mathcal{G})(\omega)$$

in quanto la seconda è una versione di

$$\mathbb{P}(\{X \in \phi(J)\} \mid \mathcal{G})(\omega) \equiv \mathbb{P}(\{\phi^{-1}(X) \in J\} \mid \mathcal{G})(\omega) \equiv \mathbb{P}(\{Y \in J\} \mid \mathcal{G})(\omega).$$

Abbiamo già usato il fatto che ogni funzione $\sigma(X)-$misurabile si può esprimere come $f(X)$. Supponiamo ora che $\mathcal{G} = \sigma(Y)$ allora, necessariamente deve accadere che $\mu_X(H)(\omega)$ sia una funzione di $Y(\omega)$, ovvero che, per ogni boreliano H, esista una funzione $f(H,y)$, misurabile in y, per cui

1) $f(H, Y(\omega)) = \mu_X(H)(\omega)$

2) $f(\cdot, y)$ sia una misura di probabilità per ogni y

(questo è sicuramente vero se $y \in Y(\Omega)$, mentre se $y \notin Y(\Omega)$ basta definire $f(H,y) = \nu(H)$ per una fissata misura di probabilità ν).

Indicheremo $f(H,y)$ con la notazione più evocativa di $\mu_X(H \mid Y = y)$ o anche $\mu_{X|Y}(H \mid y)$. Analogamente indicheremo con $F_X(t \mid Y = y)$ o con $F_{X|Y}(t \mid y)$ la funzione di distribuzione condizionale di X data Y, "valutata in $Y = y$". Ovviamente tale espressione non va confusa con $\mathbb{P}(X \leq t \mid \{Y = y\})$, che tra l'altro potrebbe non avere senso nel caso in cui $\mathbb{P}(\{Y = y\}) = 0$.

3.7 Esempi ‡

Esempio 3.7 (caso dominato). *Si tratta della generalizzazione dell'Esempio 3.4. Siano X ed Y due variabili aleatorie con legge congiunta assolutamente continua rispetto ad una misura prodotto $\nu_1(dx) \times \nu_2(dy)$ cioè con legge congiunta data da*

$$\mu_{X,Y}(dx, dy) = f(x,y)\nu_1(dx) \times \nu_2(dy).$$

È facile vedere che la legge di X è assolutamente continua rispetto a $\nu_1(dx)$ e la legge di Y lo è rispetto a $\nu_2(dy)$, o meglio

$$\mu_X(dx) = \left[\int f(x,y)\nu_2(dy)\right] \nu_1(dx) = f_X(x)\nu_1(dx),$$

$$\mu_Y(dy) = \left[\int f(x,y)\nu_1(dx)\right] \nu_2(dy) = f_Y(y)\nu_2(dy).$$

In questo caso anche la legge condizionale di X data Y è assolutamente continua rispetto a $\nu_1(dx)$ e risulta

$$\mu_{X|Y}(dx|y) = \frac{f(x,y)}{f_Y(y)}\nu_1(dx),$$

per μ_Y-quasi ogni y.

Infatti è facile verificare che, per ogni funzione $g : \mathbb{R} \to \mathbb{R}$, misurabile,

$$\mathbb{E}[g(X)h(Y)] = \mathbb{E}[\phi(Y)h(Y)], \quad \text{per ogni funzione } h \text{ misurabile} \qquad (*)$$

dove

$$\phi(y) = \int g(x)\mu_{X|Y}(dx|y) = \int g(x)\frac{f(x,y)}{f_Y(y)}\nu_1(dx);$$

(ovviamente g ed h devono soddisfare ipotesi che garantiscano l'integrabilità delle v.a. $g(X)$ ed $h(Y)$)

Tutte le v.a. W che siano $\sigma(Y)-$misurabili sono del tipo $W = h(Y)$ per h boreliana, di conseguenza

$$\mathbb{E}[g(X) \mid Y] = \phi(Y) = \int g(x)\ \mu_{X|Y}(dx|y)\big|_{y=Y(\omega)} = \int g(x)\ \frac{f(x,y)}{f_Y(y)}\nu_1(dx)\bigg|_{y=Y(\omega)}$$

Un risultato analogo vale ovviamente anche per la legge di Y data X.
Casi particolari sono i casi in cui $\nu_1(dx) = dx$, e $\nu_2(dy) = dy$, cioè torniamo al caso della misura di Lebesgue esaminato nell'Esempio 3.4, oppure $\nu_1(dx) = \sum_{n=0}^{\infty}\delta_{x_n}(dx)$, e $\nu_2(dy) = \sum_{n=0}^{\infty}\delta_{y_n}(dy)$ e si trova allora che

$$\mathbb{E}[g(X) \mid Y] = \sum_n g(x_n)\ \mu_{X|Y}(\{x_n\}|y)\big|_{y=Y(\omega)} = \sum_n g(x_n)\ \mathbb{P}(X = x_n \mid Y = y)|_{y=Y(\omega)}\,.$$

Esempio 3.8 (caso non dominato). *Siano X ed Y due variabili aleatorie con legge congiunta data da*

$$\mu_{X,Y}(dx, dy) = p\, f_1(x)\, dx\, \delta_{\alpha(x)}(dy) + q\, f_2(x,y)\, dx\, dy,$$

dove $p + q = 1$, $f_1(x)$ è una densità di probabilità su $\mathbb{R}$, $f_2(x,y)$ è una densità di probabilità su $\mathbb{R}^2$, $\delta_z(dx)$ è la misura concentrata in z (ovvero $\delta_z(A) = 1$ se $z \in A$, mentre $\delta_z(A) = 0$ se $z \notin A$), α è una funzione invertibile, C^1 e con $\mid \alpha'(x) \mid$ strettamente positiva.

Le distribuzioni marginali sono equivalenti alla misura di Lebesgue, essendo

$$\mu_X(dx) = \left[p f_1(x) + q\int f_2(x,y)dy\right] dx,$$

$$\mu_Y(dy) = \left[p f_1(\alpha^{-1}(y))\frac{1}{|\alpha'(\alpha^{-1}(y))|} + q\int f_2(x,y)dx\right] dy.$$

L'ultima uguaglianza deriva da

$$\mathbb{E}[h(Y)] = p\int h(\alpha(x))f_1(x)dx + q\int h(y)dy\int f_2(x,y)dx =$$

$$= p\int h(y)f_1(\alpha^{-1}(y))\frac{1}{|\alpha'(\alpha^{-1}(y))|}dy + q\int h(y)dy\int f_2(x,y)dx = \int h(y)\mu_Y(dy)$$

Come si procede per calcolare la media condizionata di $g(X)$ data Y?

Come nell'esempio precedente si deve trovare una v.a. $\sigma(Y)$-misurabile, cioè una funzione $\phi(Y)$ per cui valga

$$\mathbb{E}[g(X)h(Y)] = \mathbb{E}[\phi(Y)h(Y)], \qquad \text{per ogni funzione } h \text{ misurabile} \qquad (*)$$

Se poi troviamo che

$$\phi(y) = \int g(x)\nu(dx; y)$$

per una misura di probabilità $\nu(\cdot; y)$, allora potremo affermare che $\nu(\cdot; y)$ è una versione regolare della distribuzione condizionale di X data $Y = y$.

Ora, qualunque sia la funzione h

$$
\begin{aligned}
\mathbb{E}[g(X)h(Y)] &= p\iint g(x)h(y)f_1(x)dx\delta_{\alpha(x)}(dy) + q\iint g(x)h(y)f_2(x,y)dxdy \\
&= p\int g(x)h(\alpha(x))f_1(x)dx + q\int\Big(\int g(x)f_2(x,y)dx\Big)h(y)dy \\
&= p\int g(\alpha^{-1}(y))f_1(\alpha^{-1}(y))\frac{1}{|\alpha'(\alpha^{-1}(y))|}h(y)dy + q\int\Big(\int g(x)f_2(x,y)dx\Big)h(y)dy \\
&= \int\Big[pg(\alpha^{-1}(y))f_1(\alpha^{-1}(y))\frac{1}{|\alpha'(\alpha^{-1}(y))|}+q\Big(\int g(x)f_2(x,y)dx\Big)\Big]\,h(y)dy,
\end{aligned}
$$

mentre invece

$$
\begin{aligned}
\mathbb{E}[\phi(Y)h(Y)] = p\int \phi(y)h(y)f_1(\alpha^{-1}(y))\frac{1}{|\alpha'(\alpha^{-1}(y))|}dy + q\int\phi(y)h(y)dy\int f_2(x,y)dx \\
= \int\phi(y)\,\Big[pf_1(\alpha^{-1}(y))\frac{1}{|\alpha'(\alpha^{-1}(y))|}+q\Big(\int f_2(x,y)dx\Big)\Big]\,h(y)dy.
\end{aligned}
$$

Quindi affinché valga l'uguaglianza (*), qualunque sia h, è necessario e sufficiente che

$$
\begin{aligned}
\Big[pg(\alpha^{-1}(y))f_1(\alpha^{-1}(y))\frac{1}{|\alpha'(\alpha^{-1}(y))|}+q\Big(\int g(x)f_2(x,y)dx\Big)\Big] = \\
= \phi(y)\,\Big[pf_1(\alpha^{-1}(y))\frac{1}{|\alpha'(\alpha^{-1}(y))|}+q\Big(\int f_2(x,y)dx\Big)\Big]
\end{aligned}
$$

ovvero che

$$
\phi(y) = \frac{pg(\alpha^{-1}(y))f_1(\alpha^{-1}(y))\frac{1}{|\alpha'(\alpha^{-1}(y))|}+q\Big(\int g(x)f_2(x,y)dx\Big)}{pf_1(\alpha^{-1}(y))\frac{1}{|\alpha'(\alpha^{-1}(y))|}+q\Big(\int f_2(x,y)dx\Big)}.
$$

Si noti ancora che

$$
g(\alpha^{-1}(y)) = \int g(x)\delta_{\alpha^{-1}(y)}(dx).
$$

Per questo motivo, qualunque sia g, si può riscrivere il numeratore come

$$
pf_1(\alpha^{-1}(y))\frac{1}{|\alpha'(\alpha^{-1}(y))|}\int g(x)\delta_{\alpha^{-1}(y)}(dx) + q\Big(\int g(x)f_2(x,y)dx\Big)
$$

e quindi la legge $\mu_{X|Y}(dx|y)$ di X, condizionata ad $Y=y$, è proporzionale a

$$
pf_1(\alpha^{-1}(y))\frac{1}{|\alpha'(\alpha^{-1}(y))|}\delta_{\alpha^{-1}(y)}(dx) + qf_2(x,y)dx.
$$

Si noti che quindi la legge di X condizionata a $Y=y$ non è assolutamente continua rispetto alla distribuzione iniziale di X, che ha invece una densità rispetto alla misura di Lebesgue.

Capitolo 4

Martingale

In questo capitolo introduciamo il concetto di martingala, che in un certo senso si può considerare una formalizzazione del concetto di gioco equo. Per definire una martingala abbiamo però prima bisogno di dare la seguente definizione.

Definizione 4.1 (Filtrazione). *In uno spazio $(\Omega, \mathcal{F})$, la famiglia $\{\mathcal{F}_t, t \geq 0\}$ si dice una* **filtrazione** *se è una famiglia crescente di σ-algebre di $\mathcal{F}$, cioè $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$, per $0 \leq s \leq t$. (La definizione ha senso anche nel caso in cui $t = n \in \mathbb{N}$, e nel caso in cui $t \in I \subseteq \mathbb{R}$, ad esempio $t \in [0, T]$).*

La $\sigma-$algebra $\mathcal{F}_t$ rappresenta l'informazione disponibile fino al tempo t, e più in generale la filtrazione $\{\mathcal{F}_t\}$ viene detta anche ***flusso di σ-algebre***, in quanto rappresenta il flusso di informazioni disponibili, al variare del tempo.

Definizione 4.2 (Martingala). *Un processo aleatorio*[1] *(X_t), definito in uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$, si dice una* **martingala** *rispetto ad una filtrazione $\{\mathcal{F}_t\}$, con $\mathcal{F}_t \subseteq \mathcal{F}$, se*

0) *X_t è $\mathcal{F}_t$-misurabile per ogni $t \geq 0$. (o più rapidamente il processo X_t è* **adattato** *ad $\mathcal{F}_t$)*

1) *X_t è integrabile per ogni $t \geq 0$, cioè $\mathbb{E}[| X_t |] < +\infty$ per ogni $t \geq 0$.*

2) *$\mathbb{E}[X_{t+s} \mid \mathcal{F}_t] = X_t$, per ogni $t, s \geq 0$*

Nel caso in cui $t = n \in \mathbb{N}$ la **2)** *può essere sostituita con la richiesta che*

$$\mathbb{E}[X_{n+1} \mid \mathcal{F}_n] = X_n, \text{ per ogni } n \in \mathbb{N}.$$

Si noti che la proprietà **0)** è sovrabbondante in quanto la **2)** implica che X_t sia $\mathcal{F}_t$-misurabile.

Ciò non è vero nella seguente definizione di submartingala (supermartingala), che in un certo senso si può considerare una formalizzazione del concetto di gioco favorevole (sfavorevole).

Definizione 4.3 (Submartingala (Supermartingala)). *Un processo aleatorio X_t si dice una* **submartingala** *(***supermartingala***) rispetto ad una filtrazione $\{\mathcal{F}_t\}$ se*

0) *X_t è $\mathcal{F}_t$-misurabile per ogni $t \geq 0$. (o più rapidamente il processo X_t è adattato ad $\mathcal{F}_t$)*

1) *X_t è integrabile per ogni $t \geq 0$, cioè $\mathbb{E}[| X_t |] < +\infty$ per ogni $t \geq 0$.*

2) *$\mathbb{E}[X_{t+s} \mid \mathcal{F}_t] \geq X_t(\mathbb{E}[X_{t+s} \mid \mathcal{F}_t] \leq X_t)$, per ogni $t, s \geq 0$,*

[1] Il lettore per il momento può pensare di considerare solo il caso tempo discreto, e sostituire la parola processo con la parola successione $\{X_n\}_n$ di variabili aleatorie. Inoltre può limitarsi a considerare il caso di una filtrazione $\{\mathcal{F}_n\}_n$ generata dal processo stesso, ovvero il caso in cui $\mathcal{F}_n$ è σ-algebra generata da $\{X_1, X_2, \cdots, X_n\}$. In altre parole

$$\mathcal{F}_n = \{A \subseteq \Omega, \text{ tali che esiste un boreliano } H_n \in \mathcal{B}(\mathbb{R}^n) \text{ per il quale } A = \{\omega \in \Omega : (X_1, X_2, \cdots, X_n) \in H_n\}\}.$$

Per la definizione formale di processo aleatorio si rimanda al Capitolo 8.

Di nuovo, nel caso in cui $t = n \in \mathbb{N}$ la **2)** *può essere sostituita con la richiesta che*

$$\mathbb{E}[X_{n+1} \mid \mathcal{F}_n] \geq X_n (\mathbb{E}[X_{n+1} \mid \mathcal{F}_n] \leq X_n), \text{ per ogni } n \in \mathbb{N}.$$

Osservazione 4.1. *Se X_t è una martingala (o submartingala) rispetto a una filtrazione $\{\mathcal{F}_t\}$ e se $\{\mathcal{G}_t\}$ è una filtrazione per cui $\sigma(X_t) \subseteq \mathcal{G}_t \subseteq \mathcal{F}_t$, per ogni $t \geq 0$, allora X_t lo è anche rispetto alla nuova filtrazione:*

$$\mathbb{E}[X_{t+s} \mid \mathcal{G}_t] = \mathbb{E}[\mathbb{E}[X_{t+s} \mid \mathcal{F}_t] \mid \mathcal{G}_t] = \mathbb{E}[X_t \mid \mathcal{G}_t] = X_t.$$

In particolare per ogni martingala (o submartingala) si può prendere sempre la filtrazione minimale $\mathcal{F}_t^X = \sigma\big(\{X_u, u \leq t\}\big) = \sigma\big(\bigcup_{0 \leq u \leq t} \sigma(X_u)\big)$, che per la proprietà **0)** *è sempre contenuta in $\mathcal{F}_t$. Quindi, in genere* [2]*, se la filtrazione non è specificata, si deve intendere che si tratti della filtrazione naturale.*

Se invece $\{\mathcal{H}_t\}$ è una filtrazione per cui $\mathcal{H}_t$ è indipendente da $\mathcal{F}_t$ (e quindi da X_t), allora X_t è una martingala (o submartingala) anche rispetto alla filtrazione $\mathcal{G}_t := \mathcal{F}_t \vee \mathcal{H}_t$ (confrontare la proprietà **6.** *delle medie condizionali: il condizionamento ridondante).*

Osservazione 4.2. *Ogni martingala ha media costante e ogni submartingala ha media crescente (in senso lato): basta passare al valore medio nella proprietà* **3.** *delle medie condizionali (condizionamenti successivi) e tenere presente che il valore medio di Y coincide con il valore medio di $\mathbb{E}[Y \mid \mathcal{G}]$.*

Osservazione 4.3. *Data una submartingala X_t si ottiene una supermartingala considerando $-X_t$, e viceversa.*

4.1 Esempi di martingale e di submartingale

Esempio 4.1. *Sia data una variabile aleatoria Y integrabile ed una filtrazione $\{\mathcal{F}_t\}$. Si definisca $X_t := \mathbb{E}[Y \mid \mathcal{F}_t]$. Si dimostra facilmente che X_t è una martingala:*

$$\mathbb{E}[X_{t+s} \mid \mathcal{F}_t] \underset{(per\ def.)}{=} \mathbb{E}[\mathbb{E}[Y \mid \mathcal{F}_{t+s}] \mid \mathcal{F}_t] \underset{(condiz.\ successivi)}{=} \mathbb{E}[Y \mid \mathcal{F}_t]$$

L'esempio precedente aiuta a capire il nome di submartingala: sia ora X_t una submartingala, allora, fissato $v = t + s$, il processo

$$Z_t := \mathbb{E}[X_{t+s} \mid \mathcal{F}_t] \equiv \mathbb{E}[X_v \mid \mathcal{F}_t],$$

è una martingala per $t \in [0, v]$, quindi la proprietà **2)** per le submartingale diviene

$$Z_t := \mathbb{E}[X_v \mid \mathcal{F}_t] \geq X_t, \qquad \text{per } t \in [0, v],$$

ovvero che X_t sia *sotto* la martingala $Z_t := \mathbb{E}[X_v \mid \mathcal{F}_t]$, per $t \in [0, v]$.

*** Inoltre sempre in relazione all'Esempio 4.1, si può osservare che se due martingale M_t^1 ed M_t^2 coincidono al tempo T, ossia se $M_T^1 = M_T^2$, allora esse coincidono per ogni $t \leq T$, ossia $M_t^1 = M_t^2$ in $[0, T]$, in quanto per tali valori di t si ha $M_t^i = \mathbb{E}[M_T^i \mid \mathcal{F}_t]$. ***

Esempio 4.2. *Sia data una successione di variabili aleatorie Y_n indipendenti, identicamente distribuite*[3]*, integrabili e a media nulla. Si definisca*

$$S_0 = 0, \qquad S_n = \sum_{k=1}^{n} Y_k, \quad n \geq 1,$$

o equivalentemente

$$S_n = \sum_{k=1}^{n} Y_k, \quad n \geq 0,$$

[2] Attenzione: ovviamente la filtrazione potrebbe anche essere specificata anche all'inizio di una sezione, o di un capitolo.

[3] Non è necessario che le variabili aleatorie Y_n abbiano la stessa distribuzione, basta che siano integrabili e abbiano valore atteso nullo.

(con la convenzione che $\sum_{k=1}^{0} a_k = \sum_{1\leq k\leq 0} a_k = \sum_{k\in\emptyset} a_k = 0$).

Si ottiene facilmente che S_n è una martingala, rispetto[4] *a*

$$\mathcal{F}_n^S = \begin{cases} \{\emptyset, \Omega\} & se\ n = 0, \\ \sigma(\{Y_1, \cdots, Y_n\}) & se\ n > 0. \end{cases}$$

Infatti

$$\mathbb{E}[S_{n+1} \mid \mathcal{F}_n] = \mathbb{E}[S_n + Y_{n+1} \mid \mathcal{F}_n] = S_n + \mathbb{E}[Y_{n+1} \mid \mathcal{F}_n] = S_n + \mathbb{E}[Y_{n+1}] = S_n.$$

Se le variabili aleatorie Y_k non sono a media nulla, basta sostituire $Y_k - \mu$, se $\mu = \mathbb{E}[Y_1]$, ed ottenere che

$$\widetilde{S}_n = \sum_{k=1}^{n} (Y_k - \mu) = \sum_{k=1}^{n} Y_k - n\mu = S_n - n\mu$$

è una martingala[5].

Infine si osservi che, nel caso in cui $\mu > 0$, si ottiene che S_n è una submartingala, mentre, nel caso in cui $\mu < 0$, si ottiene che S_n è una supermartingala.

Esempio 4.3. *Se siamo nelle stesse ipotesi del precedente Esempio 4.2 ed inoltre le variabili aleatorie Y_k ammettono momento secondo finito, e quindi varianza σ^2, allora $M_n := S_n^2 - n\sigma^2$ è una martingala:*

$$M_{n+1} - M_n = S_{n+1}^2 - (n+1)\sigma^2 - (S_n^2 - n\sigma^2) = (S_n^2 + 2Y_{n+1}S_n + Y_{n+1}^2) - S_n^2 - \sigma^2 = 2Y_{n+1}S_n + Y_{n+1}^2 - \sigma^2$$

Passando alle medie condizionali si ottiene che

$$\begin{aligned} &\mathbb{E}[M_{n+1} - M_n \mid \mathcal{F}_n] = \mathbb{E}[2Y_{n+1}S_n + Y_{n+1}^2 - \sigma^2 \mid \mathcal{F}_n] \\ &= 2S_n\mathbb{E}[Y_{n+1} \mid \mathcal{F}_n] + \mathbb{E}[Y_{n+1}^2 \mid \mathcal{F}_n] - \sigma^2 = 2S_n \cdot 0 + \sigma^2 - \sigma^2 = 0 \end{aligned}$$

in quanto Y_{n+1} è indipendente da $\mathcal{F}_n$ e quindi i valori medi condizionali $\mathbb{E}[Y_{n+1} \mid \mathcal{F}_n]$ ed $\mathbb{E}[Y_{n+1}^2 \mid \mathcal{F}_n]$ coincidono con i rispettivi valori medi.

Se le variabili aleatorie Y_k non sono a media nulla, allora

$$M_n := (S_n - n\mu)^2 - n\sigma^2$$

è una martingala[6].

Il seguente esempio permette di generare submartingale a partire da martingale e da submartingale.

---

[4]La successione $\{Y_k\}$ si ottiene immediatamente dalla successione $\{S_k\}$:

$$Y_n = S_n - S_{n-1}, \qquad \text{per } n \geq 1,$$

di conseguenza $\mathcal{F}_n^S = \mathcal{F}_n^Y$, per $n \geq 1$.

[5]Sempre nel caso in cui le variabili aleatorie Y_k non abbiano la stessa legge si dovr1à sostituire $Y_k - \mu_k$, dove $\mu_k = \mathbb{E}[Y_k]$. In questo caso

$$\widetilde{S}_n = \sum_{k=1}^{n} (Y_k - \mu_k) = \sum_{k=1}^{n} Y_k - \sum_{k=1}^{n} \mu_k$$

[6]Di nuovo la stessa dimostrazione funziona anche nel caso in cui le variabili aleatorie Y_k non hanno la stessa legge, purché abbiano momento secondo finito e siano indipendenti. In tale caso

$$M_n := (S_n - \sum_{k=1}^{n} \mu_k)^2 - \sum_{k=1}^{n} \sigma_k^2$$

è una martingala. Si confronti questo risultato con il successivo Teorema di Decomposizione di Doob 4.1.

Esempio 4.4. *Sia X_t una martingala con $\mathbb{E}[| X_t |^{\alpha}] < +\infty$, per un $\alpha \geq 1$. Allora il processo $| X_t |^{\alpha}$ è una submartingala: basta applicare la disuguaglianza di Jensen alla funzione $|x|^{\alpha}$, che è convessa per $\alpha \geq 1$*

$$| X_t |^{\alpha} \underset{(X_t \text{ è una MG})}{=} | \mathbb{E}[X_{t+s} \mid \mathcal{F}_t] |^{\alpha} \underset{(\text{dis. Jensen per } |x|^{\alpha})}{\leq} \mathbb{E}[| X_{t+s} |^{\alpha} | \mathcal{F}_t].$$

Questo esempio si generalizza immediatamente al caso di ogni funzione convessa ϕ, purché, ovviamente, $\mathbb{E}[\phi(X_t)] < +\infty$.

$$\phi(X_t) \underset{(X_t \text{ è una MG})}{=} \phi(\mathbb{E}[X_{t+s} \mid \mathcal{F}_t]) \underset{(\text{dis. Jensen per } \phi)}{\leq} \mathbb{E}[\phi(X_{t+s}) \mid \mathcal{F}_t].$$

Per ottenere un risultato analogo nel caso in cui X_t sia una submartingala bisogna aggiungere l'ipotesi che ϕ sia una funzione crescente, in modo che, essendo

$$\begin{aligned} X_t &\underset{(X_t \text{ è una subMG})}{\leq} \mathbb{E}[X_{t+s} \mid \mathcal{F}_t], \\ &\Downarrow \\ \phi(X_t) &\underset{(\phi \text{ è crescente})}{\leq} \phi(\mathbb{E}[X_{t+s} \mid \mathcal{F}_t]) \underset{(\text{dis. Jensen per } \phi)}{\leq} \mathbb{E}[\phi(X_{t+s}) \mid \mathcal{F}_t]. \end{aligned}$$

Esempio 4.5. *Se le variabili aleatorie W_k sono indipendenti identicamente distribuite*[7]*, con $\mathbb{E}[W_1] = 1$, allora la successione di variabili aleatorie*

$$Z_0 = 1, \qquad Z_n := \prod_{k=1}^{n} W_k, \quad n \geq 1,$$

o equivalentemente

$$Z_n := \prod_{k=1}^{n} W_k, \quad n \geq 0,$$

(con la convenzione che $\prod_{k=1}^{0} a_k = 1$) definisce una martingala, rispetto [8] *a*

$$\mathcal{F}_n = \begin{cases} \{\emptyset, \Omega\} & \text{se } n = 0, \\ \sigma(\{W_1, \cdots, W_n\}) & \text{se } n > 0, \end{cases}$$

ovvero, dato che la condizione di misurabilità è ovvia, quella di integrabilità deriva dall'integrabilità di ciascuna delle W_k e dalla loro indipendenza, e infine

$$\mathbb{E}[Z_{n+1} \mid \mathcal{F}_n] = \mathbb{E}[Z_n W_{n+1} \mid \mathcal{F}_n] = Z_n \mathbb{E}[W_{n+1} \mid \mathcal{F}_n] = Z_n \mathbb{E}[W_{n+1}] = Z_n.$$

Come caso particolare si consideri la situazione dell'Esempio 4.2 con l'ulteriore ipotesi che per un $\theta \in \mathbb{R}$ valga

$$\mathbb{E}[\exp\{\theta Y_1\}] = \exp\{\psi(\theta)\} < +\infty$$

*(si noti che **non è necessario** supporre $\mathbb{E}[Y_1] = 0$). Allora ovviamente*

$$\mathbb{E}[\exp\{\theta Y_1\} \exp\{-\psi(\theta)\}] = \mathbb{E}[\exp\{\theta Y_1 - \psi(\theta)\}] = 1.$$

Come conseguenza, posto $W_k = \exp\{\theta Y_k - \psi(\theta)\}$, si ha che

$$Z_n = \exp\{\theta S_n - n\psi(\theta)\}$$

è una martingala strettamente positiva di media 1.

[7] L'ipotesi che abbiano la stessa legge è superflua, basta che le variabili aleatorie W_k abbiano valore atteso 1 e siano indipendenti.

[8] In questo caso $\mathcal{F}_n^Z$ può essere strettamente contenuta in $\mathcal{F}_n = \mathcal{F}_n^W$, infatti mentre $Z_n = g_n(W_1, \ldots, W_n)$, e quindi ogni funzione misurabile rispetto a $Z_1, \ldots, Z_n$ è funzione misurabile di $W_1, \ldots, W_n$, il viceversa in genere non è vero: se $Z_1, \ldots, Z_n$ sono tutti positivi, allora $W_k = Z_k / Z_{k-1}$, ma se $Z_n = 0$ allora $Z_{n+m} = 0$ per ogni $m \geq 0$ e quindi è impossibile ricavare i valori di W_{n+m} per $m > 0$. Se tuttavia le variabili aleatorie $W_k(\omega)$, assumono valori strettamente positivi per ogni k e per ogni ω, allora la corrispondenza tra $\{Z_k\}$ e $\{W_k\}$ è biunivoca e $\mathcal{F}_n^Z = \mathcal{F}_n^W$, per ogni $n \geq 1$.

Esempio 4.6 (*Integrale stocastico a tempo discreto***).** *Sia $\tilde{S}_n$ una martingala rispetto a $\mathcal{F}_n$, nello spazio $(\Omega, \mathcal{F}, \tilde{\mathbb{P}})$, e sia γ_n* **predicibile** *rispetto a $\mathcal{F}_n$, ovvero sia γ_n misurabile rispetto a $\mathcal{F}_{n-1}$ per ogni $n \geq 1$. Allora l'***integrale stocastico discreto**

$$(\gamma.\tilde{S})_n = I_n(\gamma) := \sum_{k=1}^{n} \gamma_k(\tilde{S}_k - \tilde{S}_{k-1}) \tag{4.1}$$

definisce una $\mathcal{F}_n$-martingala, sotto una delle due seguenti condizioni:
a) *per ogni k esiste una costante c_k tale che $\tilde{\mathbb{P}}(\{\omega \text{ tali che } \gamma_k(\omega) \leq c_k\}) = 1$*
b) *La martingala $\tilde{S}_k$ e il processo γ_k sono di quadrato integrabile, ovvero*

$$\tilde{\mathbb{E}}[|\tilde{S}_k|^2] < \infty, \text{ per ogni } k \geq 0 \quad \tilde{\mathbb{E}}[|\gamma_k|^2] < \infty, \text{ per ogni } k \geq 1$$

Verifica. Per la misurabilità basta osservare che se $k \leq n$ allora γ_k e $\tilde{S}_{k-1}$ sono $\mathcal{F}_{k-1}$-misurabile e quindi anche $\mathcal{F}_n$-misurabile, analogamente $\tilde{S}_k$ è $\mathcal{F}_k$-misurabile e quindi anche $\mathcal{F}_n$-misurabile, di conseguenza

$$I_n(\gamma) = \sum_{k=1}^{n} \gamma_k(\tilde{S}_k - \tilde{S}_{k-1})$$

è $\mathcal{F}_n$-misurabile. Per l'integrabilità si osservi che

$$|I_n(\gamma)| \leq \sum_{k=1}^{n} |\gamma_k|(|\tilde{S}_k| + |\tilde{S}_{k-1}|)$$

e che, se vale la condizione **a)***, allora*

$$\tilde{\mathbb{E}}[|\gamma_k||\tilde{S}_k|] \leq c_k \tilde{\mathbb{E}}[|\tilde{S}_k|] < \infty \qquad \tilde{\mathbb{E}}[|\gamma_k||\tilde{S}_{k-1}|] \leq c_k \tilde{\mathbb{E}}[|\tilde{S}_{k-1}|] < \infty,$$

mentre se vale la condizione **b)***, allora per la disuguaglianza di Cauchy*

$$\tilde{\mathbb{E}}[|\gamma_k||\tilde{S}_k|] \leq \tilde{\mathbb{E}}^{1/2}[|\gamma_k|^2]\tilde{\mathbb{E}}^{1/2}[|\tilde{S}_k|^2] < \infty \qquad \tilde{\mathbb{E}}[|\gamma_k||\tilde{S}_{k-1}|] \leq \tilde{\mathbb{E}}^{1/2}[|\gamma_k|^2]\tilde{\mathbb{E}}^{1/2}[|\tilde{S}_{k-1}|^2] < \infty.$$

Infine basta osservare che

$$\begin{aligned}\tilde{\mathbb{E}}[I_{n+1}(\gamma) - I_n(\gamma) \mid \mathcal{F}_n] &= \tilde{\mathbb{E}}[\gamma_{n+1}(\tilde{S}_{n+1} - \tilde{S}_n) \mid \mathcal{F}_n] \\ &\underset{(\gamma_{n+1} \text{ è } \mathcal{F}_n-mis.)}{=} \gamma_{n+1}\, \tilde{\mathbb{E}}[(\tilde{S}_{n+1} - \tilde{S}_n) \mid \mathcal{F}_n] \\ &\underset{(\tilde{S}_n \text{ è una } \mathcal{F}_n-MG)}{=} \gamma_{n+1}\, 0 = 0\end{aligned}$$

Esempio 4.7. [9] *Dato uno spazio $(\Omega, \mathcal{F})$ e su di esso una filtrazione $\{\mathcal{F}_t\}$ e due misure di probabilità $\mathbb{P}$ e $\mathbb{Q}$, con $\mathbb{P}$ assolutamente continua rispetto a $\mathbb{Q}$ (di conseguenza lo sono anche rispetto ad $\mathcal{F}_t$ per ogni t). Si definisca*[10]

[9] Questo esempio richiede la conoscenza del Teorema di Radon Nikodym, e può essere tralasciato in una prima lettura. In alternativa il lettore può considerare solo il caso a tempo discreto con $t = k \in \{1, \cdots n\}$, $\Omega = \mathbb{R}^n$, $\mathcal{F}_k = \{A = H_k \times \mathbb{R}^{n-k},$ con $H_k \in \mathcal{B}(\mathbb{R}^k)\}$, ed infine $\mathbb{P}(dx_1 \cdots dx_n) = p(x_1, \cdots, x_n)dx_1 \cdots dx_n$ e $\mathbb{Q}(dx_1 \cdots dx_n) = q(x_1, \cdots, x_n)dx_1 \cdots dx_n$ con p e q densità di probabilità. La assoluta continuità di $\mathbb{P}$ rispetto a $\mathbb{Q}$ diviene allora la condizione

$$\{(x_1, \cdots, x_n) : q(x_1, \cdots, x_n) = 0\} \subseteq \{(x_1, \cdots, x_n) : p(x_1, \cdots, x_n) = 0\},$$

e

$$\frac{d\mathbb{P}}{d\mathbb{Q}}(x_1, \cdots, x_n) = \frac{p(x_1, \cdots, x_n)}{q(x_1, \cdots, x_n)}.$$

[10] Nel caso a tempo discreto della nota precedente

$$L_k((x_1, \cdots, x_n) := \frac{p_k(x_1, \cdots, x_k)}{q_k(x_1, \cdots, x_k)},$$

dove

$$p_k(x_1, \cdots, x_k) = \int_{\mathbb{R}}^{n-k} p(x_1, \cdots, x_k, y_{k+1}, \cdots, y_n)dy_{k+1}, \cdots, dy_n$$

e

$$q_k(x_1, \cdots, x_k) = \int_{\mathbb{R}}^{n-k} q(x_1, \cdots, x_k, y_{k+1}, \cdots, y_n)dy_{k+1}, \cdots, dy_n$$

sono le densità marginali su $\mathbb{R}^k$ di p e q, rispettivamente.

la derivata di Radon Nikodym di $\mathbb{P}$ rispetto a $\mathbb{Q}$, entrambe ristrette a $\mathcal{F}_t$, cioè

$$L_t := \left.\frac{d\mathbb{P}}{d\mathbb{Q}}\right|_{\mathcal{F}_t}.$$

Il processo L_t è una $\mathcal{F}_t$-martingala nello spazio $(\Omega, \mathcal{F}, \mathbb{Q})$, anzi più in generale risulta che se X_t è un processo adattato ad $\mathcal{F}_t$, allora X_t è una $\mathcal{F}_t$-martingala nello spazio $(\Omega, \mathcal{F}, \mathbb{P})$ se e solo se X_tL_t è una $\mathcal{F}_t$-martingala nello spazio $(\Omega, \mathcal{F}, \mathbb{Q})$ (quindi il caso precedente deriva prendendo banalmente $X_t \equiv 1$).

Infatti X_t è una martingala in $(\Omega, \mathcal{F}, \mathbb{P})$, se e solo è integrabile rispetto a $\mathbb{P}$ e se per ogni $0 \leq s \leq t$ e $A \in \mathcal{F}_s$

$$\mathbb{E}^{\mathbb{P}}[I_A X_t] = \mathbb{E}^{\mathbb{P}}[I_A X_s],$$

mentre X_tL_t lo è in $(\Omega, \mathcal{F}, \mathbb{Q})$ se e solo se è integrabile rispetto a $\mathbb{Q}$ e se per ogni $0 \leq s \leq t$ e $A \in \mathcal{F}_s$

$$\mathbb{E}^{\mathbb{Q}}[I_A X_t L_t] = \mathbb{E}^{\mathbb{Q}}[I_A X_s L_s].$$

Ovviamente, essendo $I_A X_s$ una v.a. $\mathcal{F}_s$-misurabile, si ha $\mathbb{E}^{\mathbb{P}}[I_A X_s] = \mathbb{E}^{\mathbb{Q}}[I_A X_s L_s]$ e, essendo $I_A X_t$ una v.a. $\mathcal{F}_t$-misurabile, in quanto $A \in \mathcal{F}_s \subseteq \mathcal{F}_t$, si ha $\mathbb{E}^{\mathbb{P}}[I_A X_t] = \mathbb{E}^{\mathbb{Q}}[I_A X_t L_t]$. La verifica dell'integrabilità è banale.

Esempio 4.8. *Sia X_n una* **catena di Markov omogenea**[11] *con spazio degli stati finito e con matrice delle probabilità di transizione $(p_{i,j})_{i,j}$. Sia inoltre h una* **funzione armonica** *rispetto alla matrice delle probabilità di transizione $P = (p_{i,j})_{i,j}$, cioè*

$$h(i) = (Ph)(i) := \sum_j p_{i,j} h(j), \qquad \textit{per ogni } i$$

(si noti che ciò corrisponde a chiedere che h sia la soluzione di $(P - I)h = 0$).

Il processo

$$M_n^h := h(X_n)$$

è una martingala (rispetto a $\mathcal{F}_n^X = \sigma\{X_k,\ k = 0, \cdots, n\}$).
Questo risultato deriva da un caso più generale: qualunque sia f

$$M_n^f := f(X_n) - \sum_{k=0}^{n-1} (P - I) f(X_k)$$

è una martingala rispetto a $\mathcal{F}_n^X$.

Cominciamo con il caso h armonica. Basta controllare che

$$\mathbb{E}[M_{n+1}^h \mid \mathcal{F}_n^X] = \mathbb{E}[M_{n+1}^h \mid X_n, X_{n-1}, \cdots, X_0] = M_n^h,$$

ovvero che

$$\mathbb{E}[h(X_{n+1}) \mid X_n, X_{n-1}, \cdots, X_0] = h(X_n).$$

Essendo $(X_n, n \geq 0)$ una catena di Markov si ha[12]

[11] Si ricorda che la successione $\{X_n\}_n$ è una catena di Markov omogenea con matrice delle probabilità di transizione $(p_{i,j})_{i,j}$ significa che vale la proprietà di Markov, ovvero: qualunque siano n, j, i, $i_{n-1}, \cdots, i_0$

$$\mathbb{P}(X_{n+1} = j | X_n = i,\ X_{n-1} = i_{n-1}, \cdots, X_0 = i_0) = \mathbb{P}(X_{n+1} = j | X_n = i) = p_{i,j},$$

purché $\mathbb{P}(X_n = i,\ X_{n-1} = i_{n-1}, \cdots, X_0 = i_0) > 0$.

[12] Nel caso di variabili aleatorie discrete possiamo applicare i risultati sulle densità condizionali (vedere l'Esempio 3.3): posto $\mathbf{X} = (X_n, X_{n-1}, \cdots, X_0)$, sappiamo che

$$\mathbb{E}[f(X_{n+1}) \mid \mathbf{X}](\omega) = \sum_j f(j)\ \mathbb{P}(X_{n+1} = j \mid \{\mathbf{X} = \mathbf{x}\})|_{\mathbf{x} = \mathbf{X}(\omega)}.$$

$$\mathbb{E}[f(X_{n+1}) \mid X_n, X_{n-1}, \cdots, X_0] = (Pf)(X_n)$$

ed il caso $f = h$ armonica è immediato. Il caso generale deriva dall'osservare che

$$M_{n+1}^f - M_n^f = f(X_{n+1}) - \sum_{k=0}^{n}(P-I)f(X_k) - f(X_n) + \sum_{k=0}^{n-1}(P-I)f(X_k)$$

$$M_{n+1}^f - M_n^f = f(X_{n+1}) - (P-I)f(X_n) - f(X_n) = f(X_{n+1}) - (Pf)(X_n),$$

e quindi

$$\mathbb{E}[M_{n+1}^f - M_n^f \mid \mathcal{F}_n^X] = \mathbb{E}[f(X_{n+1}) - (Pf)(X_n) \mid \mathcal{F}_n^X] = \mathbb{E}[f(X_{n+1}) \mid \mathcal{F}_n^X] - (Pf)(X_n) = 0.$$

Per la proprietà di Markov $\mathbb{P}(X_{n+1} = j \mid \mathbf{X} = \mathbf{x}) = \mathbb{P}(X_{n+1} = j \mid X_n = x_n) = p_{x_n,j}$, quindi

$$\mathbb{P}(X_{n+1} = j \mid \mathbf{X}) = \mathbb{P}(X_{n+1} = j \mid X_n) = p_{X_n,j}$$

e

$$\sum_j f(j)\mathbb{P}(X_{n+1} = j \mid \mathbf{X} = \mathbf{x}) = (Pf)(x_n),$$

e perciò

$$\mathbb{E}[f(X_{n+1}) \mid X_n, X_{n-1}, \cdots, X_0] = (Pf)(X_n).$$

4.2 Decomposizione di Doob

Tenendo conto dell'Esempio 4.4, possiamo affermare che, se sono soddisfatte delle condizioni di integrabilità, il quadrato della martingala S_n dell'Esempio 4.2 è una submartingala. Nell'Esempio 4.3, si ottiene che $S_n^2 - n\mu$ è una martingala, quindi si può scrivere come la somma di due processi

$$S_n^2 = (S_n^2 - n\sigma^2) + n\mu = M_n + A_n$$

dove M_n è una martingala, ed $A_n = n\sigma^2$ è un processo deterministico crescente.

Il seguente teorema generalizza tale esempio a tutte le submartingale.

Teorema 4.1 (Decomposizione di Doob). *: Sia $(\Omega, \mathcal{F}, \mathbb{P})$ uno spazio di probabilità, e sia $(\mathcal{F}_n)_n \in \mathbb{N}$ una filtrazione con $\mathcal{F}_n \subseteq \mathcal{F}$. Data una $\mathcal{F}_n$-submartingala X_n a tempo discreto, essa si può sempre scrivere in modo unico (a meno di insiemi di misura nulla) come*

$$X_n = X_0 + M_n + A_n,$$

dove M_n è una martingala ed A_n è un processo predicibile (cioè, per ogni $n \in \mathbb{N}$, A_n è $\mathcal{F}_{n-1}$-misurabile) crescente in senso lato, con $A_0 = 0$.

Dimostrazione. Se una tale decomposizione esiste necessariamente deve accadere che $M_0 = 0$, ed inoltre, essendo $A_n = X_n - X_0 - M_n$, deve accadere che

$$A_{n+1} - A_n = X_{n+1} - X_0 - M_{n+1} - (X_n - X_0 - M_n) = X_{n+1} - X_n - (M_{n+1} - M_n).$$

Essendo $A_{n+1} - A_n$ una v.a. $\mathcal{F}_n$-misurabile, passando alla media condizionale rispetto ad $\mathcal{F}_n$, essa non cambia, per cui deve necessariamente accadere che

$$\begin{aligned} A_{n+1} - A_n &= \mathbb{E}[X_{n+1} - X_n - (M_{n+1} - M_n) \mid \mathcal{F}_n] \\ &= \mathbb{E}[X_{n+1} - X_n \mid \mathcal{F}_n] - \mathbb{E}[M_{n+1} - M_n \mid \mathcal{F}_n] \underset{(M_n \text{ è una } \mathcal{F}_n\text{-martingala})}{=} \\ &= \mathbb{E}[X_{n+1} - X_n \mid \mathcal{F}_n] = \mathbb{E}[X_{n+1} \mid \mathcal{F}_n] - X_n. \end{aligned}$$

Quindi l'unico modo per definire[13] A_{n+1}, con $A_0 = 0$, è il seguente

$$A_{n+1} := A_{n+1} - A_0 = \sum_{k=0}^{n} (A_{k+1} - A_k) = \sum_{k=0}^{n} (\mathbb{E}[X_{k+1} \mid \mathcal{F}_k] - X_k). \tag{4.2}$$

È immediato verificare[14] che con questa definizione il processo $A_n := \sum_{\ell=1}^{n} (\mathbb{E}[X_\ell \mid \mathcal{F}_{\ell-1}] - X_{\ell-1})$ è integrabile, predicibile e crescente, con $A_0 = 0$.

Si tratta ora solo di verificare che con questa definizione di $\{A_n\}_{n\geq 0}$ il processo $M_n := X_n - X_0 - A_n$ è una martingala, con $M_0 = 0$. Ma ovviamente

$$M_{n+1} - M_n = X_{n+1} - X_0 - A_{n+1} - (X_n - X_0 - A_n) = X_{n+1} - X_n - (A_{n+1} - A_n),$$

per cui, tenendo conto che per definizione $A_{n+1} - A_n = \mathbb{E}[X_{n+1} \mid \mathcal{F}_n] - X_n$, e passando alla media condizionale, si ottiene la tesi.

[13] Da cui segue l'unicità della decomposizione, in quanto A_n deve essere definito come in (4.2) e poi si dovrà definire necessariamente $M_n := X_n - X_0 - A_n$.

[14] Per l'integrabilità basta osservare che

$$|A_n| \leq \sum_{\ell=1}^{n-1} \big(|\mathbb{E}[X_\ell \mid \mathcal{F}_{\ell-1}]| + |X_{\ell-1}|\big) \underset{\text{(per dis. Jensen)}}{\leq} \sum_{\ell=1}^{n} \big(\mathbb{E}[|X_\ell| \mid \mathcal{F}_{\ell-1}] + |X_{\ell-1}|\big),$$

ed utilizzare il fatto che X_n sono tutte integrabili.

Per la predicibilità, basta osservare che, per ogni $\ell \leq n$, $\mathbb{E}[X_\ell \mid \mathcal{F}_{\ell-1}]$ ed $X_{\ell-1}$ sono $\mathcal{F}_{\ell-1}$-misurabili e quindi $\mathcal{F}_{n-1}$-misurabili.

Per la crescenza basta osservare che, per definizione

$$A_{n+1} - A_n = \mathbb{E}[X_{n+1} \mid \mathcal{F}_n] - X_n \geq 0,$$

dove la disuguaglianza vale in quanto X_n è una submartingala.

Infine il fatto che $A_0 = 0$, è vero per definizione.

Osservazione 4.4. *Ogni processo crescente, integrabile ed adattato è una submartingala, quindi si può applicare il Teorema di Doob e riscriverlo come la somma di un processo predicibile e di una martingala. Più in generale, la somma di un processo crescente, integrabile ed adattato e di una martingala è una submartingala, e quindi, per il Teorema di Doob, si può riscrivere ancora come la somma di un'altra martingala e di un processo crescente che inoltre è predicibile.*

Osservazione 4.5. *Si noti che nella dimostrazione del teorema di decomposizione di Doob, il fatto che X_n sia una submartingala è servito solo nei seguenti punti:*
(i) *ha senso calcolare la media condizionata di $\mathbb{E}[X_{n+1} - X_n \mid \mathcal{F}_n]$, in quanto X_k è integrabile per ogni k,*
(ii) *il valore atteso condizionato $\mathbb{E}[X_{n+1} - X_n \mid \mathcal{F}_n]$ coincide con $\mathbb{E}[X_{n+1} \mid \mathcal{F}_n] - X_n$ in quanto X_n è adattato alla filtrazione $\{\mathcal{F}_n\}$,*
(iii) *il processo A_n risulta crescente (in senso lato), in quanto $\mathbb{E}[X_{n+1} \mid \mathcal{F}_n] - X_n \geq 0$.*

Si vede quindi che il procedimento si applica a qualunque processo che sia integrabile ed $\mathcal{F}_n$-adattato, si ottiene però una decomposizione nella somma di una martingala e di un processo predicibile (che, in generale, non è crescente).

Sarebbe interessante rivedere l'Esempio 4.8, sotto questa luce, in fondo applicando il procedimento della decomposizione di Doob al processo $\mathcal{F}_n^X$-adattato e integrabile $f(X_n)$, si ottiene che

$$\begin{aligned}
f(X_n) &= f(X_n) - \sum_{k=0}^{n-1} \big(\mathbb{E}[f(X_{k+1}) \mid \mathcal{F}_k] - f(X_k)\big) + \sum_{k=0}^{n-1} \big(\mathbb{E}[f(X_{k+1}) \mid \mathcal{F}_k] - f(X_k)\big) \\
&= f(X_n) - \sum_{k=0}^{n-1} \big(Pf(X_k) - f(X_k)\big) + \sum_{k=0}^{n-1} \big(Pf(X_k) - f(X_k)\big) \\
&= f(X_n) - \sum_{k=0}^{n-1} \big(P - I\big) f(X_k) + \sum_{k=0}^{n-1} \big(P - I\big) f(X_k) = M_n^f + \sum_{k=0}^{n-1} \big(P - I\big) f(X_k) \\
&= f(X_0) + [M_n^f - f(X_0)] + \sum_{k=0}^{n-1} \big(P - I\big) f(X_k).
\end{aligned}$$

In effetti il processo

$$A_n^f := \sum_{k=0}^{n-1} \big(P - I\big) f(X_k)$$

è un processo $\mathcal{F}_n^X$-predicibile, in quanto chiaramente A_n^f è $\mathcal{F}_{n-1}^X$-misurabile.

Esempio 4.9 (Variazione quadratica predicibile di una martingala). *Se M_n è una martingala di quadrato integrabile, allora M_n^2 è una submartingala ***(come sappiamo dall'Esempio 4.4).****

Se $M_0 = 0$, allora la decomposizione di Doob in questo caso diviene:

$$M_n^2 = \sum_{k=1}^{n} \mathbb{E}[\Delta(M^2)_k | \mathcal{F}_{k-1}] + \sum_{k=1}^{n} \big(\Delta(M^2)_k - \mathbb{E}[\Delta(M^2)_k | \mathcal{F}_{k-1}]\big),$$

dove $\Delta(M^2)_k := M_k^2 - M_{k-1}^2$.

Così, detto **variazione quadratica predicibile**, *o* **caratteristica quadratica**, *il processo predicibile definito da*

$$< M >_n := \sum_{k=1}^{n} \mathbb{E}[\Delta(M^2)_k | \mathcal{F}_{k-1}] = \sum_{k=1}^{n} \mathbb{E}[M_k^2 - M_{k-1}^2 | \mathcal{F}_{k-1}], \tag{4.3}$$

e definita

$$m_n := \sum_{k=1}^{n} \big(\Delta(M^2)_k - \mathbb{E}[\Delta(M^2)_k | \mathcal{F}_{k-1}]\big),$$

si ha la decomposizione di Doob

$$M_n^2 = < M >_n + m_n.$$

Va menzionato il fatto che, essendo M_n una martingala,

$$< M >_k - < M >_{k-1} := \mathbb{E}[\Delta(M^2)_k|\mathcal{F}_{k-1}] = \mathbb{E}[(\Delta M_k)^2\,|\mathcal{F}_{k-1}], \tag{4.4}$$

****ossia* $\mathbb{E}[M_k^2 - M_{k-1}^2|\mathcal{F}_{k-1}] = \mathbb{E}[(M_k - M_{k-1})^2\,|\mathcal{F}_{k-1}]$, ****come si vede facilmente*[15], *e quindi*

$$< M >_n := \sum_{k=1}^{n} \mathbb{E}[\Delta(M^2)_k|\mathcal{F}_{k-1}] = \sum_{k=1}^{n} \mathbb{E}\left[(M_k - M_{k-1})^2\,|\mathcal{F}_{k-1}\right]. \tag{4.5}$$

Va infine menzionato anche il fatto che il processo

$$[M]_n := \sum_{k=1}^{n} (\Delta M_k)^2 = \sum_{k=1}^{n} (M_k - M_{k-1})^2$$

viene detto **variazione quadratica (opzionale)**. ****Si noti inoltre che* $[M]_n$, *essendo un processo adattato e crescente è una submartingala e che* $[M]_n - < M >_n$ *è una martingala, ossia* $< M >_n$ *è il processo crescente e predicibile della decomposizione di Doob, relativo alla submartingala* $[M]_n$. ***

Esempio 4.10 (Decomposizione dell'integrale stocastico a tempo discreto). ****Ci mettiamo nelle stesse ipotesi e notazioni dell'Esempio 4.9 precedente:* ***M_n *è una martingala di quadrato integrabile* ***. *Se* γ_n *è un processo predicibile, di quadrato integrabile, per l'Esempio 4.6, l'integrale stocastico a tempo discreto*

$$I_n(\gamma) := \sum_{k=1}^{n} \gamma_k\,(M_k - M_{k-1})$$

è una martingala, con

$$\Delta I_k(\gamma) = \gamma_k\,(M_k - M_{k-1}).$$

Si consideri ora il caso in cui γ_k *è limitato, ovvero esiste un* $L \in \mathbb{R}^+$ *tale che* $|\gamma_k(\omega)| \leq L$. *Si osservi che*

$$\begin{aligned} I_n^2(\gamma) &= \sum_{k=1}^{n} \gamma_k\,(M_k - M_{k-1}) \sum_{h=1}^{n} \gamma_h\,(M_h - M_{h-1}) \\ &= \sum_{k=1}^{n} \gamma_k^2\,(M_k - M_{k-1})^2 + 2\sum_{k=1}^{n-1} \gamma_k\,(M_k - M_{k-1}) \sum_{h=k+1}^{n} \gamma_h\,(M_h - M_{h-1}), \end{aligned}$$

da cui

$$\begin{aligned} \mathbb{E}\left[I_n^2(\gamma)\right] &= \mathbb{E}\left[\sum_{k=1}^{n} \gamma_k^2\,(M_k - M_{k-1})^2 + 2\sum_{k=1}^{n-1} \gamma_k\,(M_k - M_{k-1}) \sum_{h=k+1}^{n} \gamma_h\,(M_h - M_{h-1})\right] \\ &= \sum_{k=1}^{n} \mathbb{E}\left[\gamma_k^2\,(M_k - M_{k-1})^2\right] + 2\sum_{k=1}^{n-1} \sum_{h=k+1}^{n} \mathbb{E}\left[\gamma_k\,(M_k - M_{k-1})\,\gamma_h\,(M_h - M_{h-1})\right]. \end{aligned}$$

---

[15] Infatti

$$\begin{aligned} \Delta(M^2)_k &= M_k^2 - M_{k-1}^2 = (M_{k-1} + \Delta M_k)^2 - M_{k-1}^2 = M_{k-1}^2 + 2\,M_{k-1}\,\Delta M_k + (\Delta M_k)^2 - M_{k-1}^2 \\ &= 2\,M_{k-1}\,\Delta M_k + (\Delta M_k)^2, \end{aligned}$$

da cui, per la $\mathcal{F}_{k-1}$-misurabilità di M_{k-1},

$$\begin{aligned} &\mathbb{E}[\Delta(M^2)_k|\mathcal{F}_{k-1}] = \mathbb{E}[2\,M_{k-1}\,\Delta M_k|\mathcal{F}_{k-1}] + \mathbb{E}[(\Delta M_k)^2\,|\mathcal{F}_{k-1}] \\ &= 2\,M_{k-1}\,\mathbb{E}[\Delta M_k|\mathcal{F}_{k-1}] + \mathbb{E}[(\Delta M_k)^2\,|\mathcal{F}_{k-1}], \end{aligned}$$

e quindi, poiché $\mathbb{E}[\Delta M_k|\mathcal{F}_{k-1}] = 0$, in quanto M_n è una martingala,

$$\mathbb{E}[\Delta(M^2)_k|\mathcal{F}_{k-1}] = \mathbb{E}[(\Delta M_k)^2\,|\mathcal{F}_{k-1}].$$

È chiaro che se $|\gamma_k(\omega)| \le L$*, allora* $\mathbb{E}\left[I_n^2(\gamma)\right]$ *risulta finita*[16].

Ora si può vedere direttamente[17] *che*

$$\mathbb{E}\left[I_n^2(\gamma)\right] = \sum_{k=1}^{n} \mathbb{E}\left[\gamma_k^2(< M >_k - < M >_{k-1})\right]$$
$$= \sum_{k=1}^{n} \mathbb{E}\left[\gamma_k^2 \left(M_k - M_{k-1}\right)^2\right] = \mathbb{E}\left[\sum_{k=1}^{n} \gamma_k^2 \left(M_k - M_{k-1}\right)^2\right],$$

tuttavia si può procedere anche in un altro modo.

Come visto, se γ_k *è limitato, allora* $I_n(\gamma)$ *è una martingala di quadrato integrabile. Per le formule (4.3) e (4.5) dell'Esempio precedente, applicate alla martingala* $I_n(\gamma)$*, si ha*

$$I_n^2(\gamma) =< I(\gamma) >_n + \mathcal{M}_n,$$

con

$$< I(\gamma) >_n := \sum_{k=1}^{n} \mathbb{E}[I_k^2(\gamma) - I_{k-1}^2(\gamma)|\mathcal{F}_{k-1}] = \sum_{k=1}^{n} \mathbb{E}\left[\left(I_k(\gamma) - I_{k-1}(\gamma)\right)^2|\mathcal{F}_{k-1}\right],$$

ed $\mathcal{M}_n$ *una martingala a media nulla.*

Ovviamente

$$\left(I_k(\gamma) - I_{k-1}(\gamma)\right)^2 = \gamma_k^2(M_k - M_{k-1})^2,$$

da cui

$$\mathbb{E}[\left(I_k(\gamma) - I_{k-1}(\gamma)\right)^2|\mathcal{F}_{k-1}] = \mathbb{E}[\gamma_k^2\,(M_k - M_{k-1})^2|\mathcal{F}_{k-1}]$$
$$= \gamma_k^2\,\mathbb{E}[(M_k - M_{k-1})^2|\mathcal{F}_{k-1}] = \gamma_k^2\,(< M >_k - < M >_{k-1}),$$

ovvero

$$< I(\gamma) >_n = \sum_{k=1}^{n} \gamma_k^2(< M >_k - < M >_{k-1}).$$

In altre parole

$$\mathcal{M}_n := I_n^2(\gamma) - < I(\gamma) >_n$$
$$= I_n^2(\gamma) - \sum_{k=1}^{n} \gamma_k^2(< M >_k - < M >_{k-1})$$

[16] Se $|\gamma_k(\omega)| \le L$, ed M_n è quadrato integrabile, allora

$$\mathbb{E}\left[\gamma_k^2\,(M_k - M_{k-1})^2\right] \le L^2\,\mathbb{E}\left[(M_k - M_{k-1})^2\right] < \infty$$

e, per la disuguaglianza di Cauchy,

$$\mathbb{E}\left[|\gamma_k|\,|M_k - M_{k-1}|\,|\gamma_h|\,|M_h - M_{h-1}|\right] \le L^2\mathbb{E}\left[(M_k - M_{k-1})^2\right]^{1/2}\,\mathbb{E}\left[(M_h - M_{h-1})^2\right]^{1/2} < \infty.$$

[17] Si osservi che

$$\mathbb{E}\left[\gamma_k^2\,(M_k - M_{k-1})^2\right] = \mathbb{E}\left[\mathbb{E}\left[\gamma_k^2\,(M_k - M_{k-1})^2\,|\mathcal{F}_{k-1}\right]\right] = \mathbb{E}\left[\gamma_k^2\mathbb{E}\left[(M_k - M_{k-1})^2\,|\mathcal{F}_{k-1}\right]\right]$$
$$= \mathbb{E}\left[\gamma_k^2(< M >_k - < M >_{k-1})\right],$$

mentre, per $k < h$,

$$\mathbb{E}\left[\gamma_k\,(M_k - M_{k-1})\,\gamma_h\,(M_h - M_{h-1})\right] = \mathbb{E}\left[\mathbb{E}\left[\gamma_k\,(M_k - M_{k-1})\,\gamma_h\,(M_h - M_{h-1})\,|\mathcal{F}_{h-1}\right]\right]$$
$$= \mathbb{E}\left[\gamma_k\,(M_k - M_{k-1})\,\gamma_h\mathbb{E}\left[(M_h - M_{h-1})\,|\mathcal{F}_{h-1}\right]\right] = \mathbb{E}\left[\gamma_k\,(M_k - M_{k-1})\,\gamma_h\,0\right] = 0.$$

Quindi, tenendo conto dell'espressione trovata per $\mathbb{E}\left[I_n^2(\gamma)\right]$ si ottiene

$$\mathbb{E}\left[I_n^2(\gamma)\right] = \sum_{k=1}^{n} \mathbb{E}\left[\gamma_k^2(< M >_k - < M >_{k-1})\right].$$

è una martingala a media nulla. E ciò implica che

$$\mathbb{E}\left[I_n^2(\gamma)\right] = \sum_{k=1}^{n} \mathbb{E}\left[\gamma_k^2(< M >_k - < M >_{k-1})\right].$$

4.2.1 Applicazioni: verso l'integrale stocastico a tempo continuo ***

Questa sezione richiede la conoscenza dei processi a tempo continuo. Se il lettore non è familiare con tali processi può tranquillamente saltare questa parte e rimandarne la lettura.

Esempio 4.11 (Primi passi verso l'integrale stocastico a tempo continuo). *Sia $(X_t)_{t\in[0,T]}$ una $\mathcal{G}_t$-martingala di quadrato integrabile.*

*Sia inoltre $(f(s))_{s\in[0,T]}$ un **processo elementare $\mathcal{G}_t$-predicibile**, di quadrato integrabile, ovvero un processo definito da*

$$f(s,\omega) := \sum_{k=1}^{N} H_k(\omega)\mathbb{I}_{(t_{k-1},t_k]}(s),$$

per una partizione $\{t_k, k = 0, \cdots, N\}$ di $(0,T]$, con $0 = t_0 < t_1 < \cdots < t_N = T$, dove H_k sono variabili aleatorie $\mathcal{G}_{t_{k-1}}$-misurabili e di quadrato integrabile.

Si definiscano

$$\int_\alpha^\beta f(s,\omega)dX(s,\omega) := \sum_{k=1}^{N} H_k(\omega)\big(X_{\beta\wedge t_k\vee\alpha} - X_{\beta\wedge t_{k-1}\vee\alpha}\big), \tag{4.6}$$

e, nel caso $\alpha = 0$ e $\beta = t$

$$\mathcal{I}_t^X(f)(\omega) := \int_0^t f(s,\omega)dX(s,\omega) = \sum_{k=1}^{N} H_k(\omega)\big(X_{t_k\wedge t} - X_{t_{k-1}\wedge t}\big), \tag{4.7}$$

$$= \sum_{k=1}^{n} H_k(\omega)\big(X_{t_k} - X_{t_{k-1}}\big) + H_{n+1}(\omega)\big(X_t - X_{t_n}\big) \qquad \text{per } t_n \le t < t_{n+1} \tag{4.8}$$

che è indicato anche, più brevemente come

$$\mathcal{I}_t^X(f) := \int_0^t f(s)dX(s).$$

*Si ha che il processo $(\mathcal{I}_t^X(f))_{t\in[0,T]}$ è una martingala. Si osservi che, dalla (4.8), si deduce immediatamente che $\mathcal{I}_t^X(f)$ ha traiettorie continue, se $(X_t)_t$ ha traiettorie continue). Se inoltre, se le variabili H_k sono limitate, allora il processo $(\mathcal{I}_t^X(f))_{t\in[0,T]}$ è di quadrato integrabile. ***Inoltre, il processo definito da*

$$\mathcal{M}_t^f := \big(\mathcal{I}_t^X(f)\big)^2 - \left(\sum_{k=1}^{n} H_k^2(\omega)\mathbb{E}[(X_{t_k} - X_{t_{k-1}})^2|\mathcal{G}_{t_{k-1}}] + H_{n+1}^2(\omega)\mathbb{E}[(X_t - X_{t_n})^2|\mathcal{G}_{t_n}]\right), \qquad \text{per } t_n \le t < t_{n+1} \tag{4.9}$$

è una $\mathcal{G}_t$-martingala, ed in particolare si ha

$$\mathbb{E}\left[\big(\mathcal{I}_t^X(f)\big)^2\right] = \mathbb{E}\left[\sum_{k=1}^{n} H_k^2(\omega)\mathbb{E}\left[(X_{t_k} - X_{t_{k-1}})^2|\mathcal{G}_{t_{k-1}}\right] + H_{n+1}^2(\omega)\mathbb{E}[(X_t - X_{t_n})^2|\mathcal{G}_{t_n}]\right], \qquad \text{per } t_n \le t < t_{n+1} \tag{4.10}$$

Le precedenti proprietà si dimostrano utilizzando i risultati dei precedenti Esempi 4.6 e 4.10, notando che nella partizione $\{t_k, k = 0, \cdots, N\}$ che definisce $f(s)$ si può sempre supporre[18] ****che ci sia un indice h tale*

[18]Infatti se $t \notin \{t_k : k = 0, \cdots, N\}$, allora esiste un ℓ tale che $t \in (t_{\ell-1}, t_\ell)$ e quindi

$$H_\ell(\omega)\mathbb{I}_{(t_{\ell-1},t]}(s) = H_\ell(\omega)\mathbb{I}_{(t_{\ell-1},t_\ell]}(s) + H_\ell(\omega)\mathbb{I}_{(t,t_\ell]}(s).$$

che $t = t_h$. In realtà nella dimostrazione della proprietà di martingala servono due tempi $t' \leq t''$ e si deve mostrare che $\mathbb{E}\left[\mathcal{I}^X_{t''}(f)|\mathcal{G}_{t'}\right] = \mathcal{I}^X_{t'}(f)$ e che $\mathbb{E}[\mathcal{M}^f_{t''}|\mathcal{G}_{t'}] = \mathcal{M}^f_{t'}$. Basta pensare che entrambi facciano parte della partizione $\{t_k, k = 0, \cdots, N\}$.

A titolo di esempio si consideri la proprietà di martingala dell'integrale stocastico $\mathcal{I}^X_t(f)$ (tralasciando le proprietà di misurabilità e di integrabilità: si tratta di dimostrare che per ogni $t' \leq t''$

$$\mathbb{E}\left[\mathcal{I}^X_{t''}(f)|\mathcal{G}_{t'}\right] = \mathcal{I}^X_{t'}(f). \tag{4.11}$$

Se nella partizione $\{t_k, k = 0, \cdots, N\}$, si ha $t' = t_m \leq t'' = t_n$ (con $m \leq n$), posto

$$M_k = X_{t_k}, \quad \mathcal{F}_k = \mathcal{G}_{t_k}, \quad \gamma_k = f(t_k) = H_k,$$

si ottiene che $(M_n)_n$ è una martingala rispetto alla filtrazione $(\mathcal{F}_n)_n$, e, con le notazione dell'Esempio 4.10, che

$$\mathcal{I}^X_{t'}(f) = I_m(\gamma), \quad \mathcal{I}^X_{t''}(f) = I_n(\gamma)$$

e quindi la (4.11) diviene

$$\mathbb{E}\left[I_n(\gamma)|\mathcal{F}_m\right] = I_m(\gamma),$$

che è esattamente la proprietà di martingala dell'integrale stocastico a tempo discreto dell'Esempio 4.6.
Infine si noti che, la (4.8)*, si può riscrivere, per $t \in [t_k, t_{k+1}]$, come*

$$\mathcal{I}^X_t(f) = \mathcal{I}^X_{t_k}(f) + H_{k+1}\left(X_t - X_{t_k}\right),$$

e che, entrambe queste espressioni hanno come conseguenza che, se la martingala $(X_t)_t$ è una martingala a traiettorie continue, allora anche $(\mathcal{I}^X_t(f))_t$ è una martingala a traiettorie continue.

Supponiamo ora che esista un processo crescente e adattato[19] $< X >= (< X >_t)_{t\in[0,T]}$*, tale che $X^2_t- < X >_t$ è una martingala. In altre parole il processo $< X >$ è la variazione quadratica della martingala X. Allora si può dimostrare facilmente*[20] *che, per ogni $0 \leq s \leq t \leq T$,*

$$\mathbb{E}\big[(X_t - X_s)^2\big|\mathcal{G}_s\big] = \mathbb{E}\big[< X >_t - < X >_s \big|\mathcal{G}_s\big], \tag{4.12}$$

che a sua volta è la versione a tempo continuo della (4.4).

Ovviamente H_ℓ è $\mathcal{G}_t$-misurabile, in quanto Sia $\{t'_k, k = 0, \cdots, N+1\}$ la partizione ottenuta dalla partizione $\{t_k, k = 0, \cdots, N\}$ inserendo il punto t al posto $\ell + 1$ (ovvero $t'_k = t_k$ per $k \leq \ell$, $t'_{\ell+1} = t$, e $t'_k = t_{k-1}$ per i rimanenti valori di k) e sia $\{H'_k, k = 0, \cdots, N+1\}$ la famiglia di variabili aleatorie ottenuta in modo analogo dalla famiglia $\{H_k, k = 0, \cdots, N\}$ inserendo la variabile aleatoria H_ℓ al posto $\ell + 1$ (ovvero $H'_k = H_k$ per $k \leq \ell$, $H'_{\ell+1} = t$, e $H'_k = H_{k-1}$ per i rimanenti valori di k). Allora ovviamente

$$f(s, \omega) := \sum_{k=1}^{N} H_k(\omega)\mathbb{I}_{(t_{k-1},t_k]}(s) = \sum_{k=1}^{N+1} H'_k(\omega)\mathbb{I}_{(t'_{k-1},t'_k]}(s).$$

Si noti che, in entrambe le rappresentazioni, si ha che $H_k = f(t_k)$, con $f(t_k)$ che risulta $\mathcal{G}_{t_{k-1}}$ misurabile, e $H'_k = f(t'_k)$, con $f(t'_k)$ che risulta $\mathcal{G}_{t'_{k-1}}$ misurabile.

Si noti infine che anche l'integrale stocastico relativo non cambia cambiando rappresentazione, ovvero se f elementare e predicibile ammette due diverse rappresentazioni, l'integrale stocastico è sempre definito dallo stesso processo, che è come dire che la definizione è ben posta.

[19] Per essere la variazione quadratica deve avere anche la proprietà di essere predicibile, ma la definizione di predicibilità viene data dopo....????

[20] La dimostrazione della (4.12), si basa sul fatto che

$$(X_t-X_s)^2-[< X >_t - < X >_s] = X^2_t-X^2_s-2\,X_s\,(X_t-X_s)- < X >_t + < X >_s= X^2_t- < X >_t -[X^2_s- < X >_s]-2\,X_s\,(X_t-X_s).$$

Infatti, passando ai valori attesi condizionali si ha che, essendo X_s una variabile aleatoria $\mathcal{G}_s$ misurabile,

$$\mathbb{E}\big[(X_t - X_s)^2 - [< X >_t - < X >_s]\big|\mathcal{G}_s\big] = \mathbb{E}\big[X^2_t- < X >_t -[X^2_s- < X >_s]\big|\mathcal{G}_s\big] - 2\,X_s\,\mathbb{E}\big[(X_t - X_s)\big|\mathcal{G}_s\big] = 0,$$

dove l'ultima uguaglianza dipende dal fatto che $X^2_t- < X >_t$ e X_t sono martingale.

Allora le (4.9) *e la* (4.10) *si possono riscrivere, come*

$$\mathcal{M}_t^f := \left(\mathcal{I}_t^X(f)\right)^2 - \int_0^t f^2(s)\, d<X>_s, \qquad \text{è una martingala, a media nulla} \tag{4.13}$$

$$\mathbb{E}\left[\left(\mathcal{I}_t^X(f)\right)^2\right] = \mathbb{E}\left[\int_0^t f^2(s)\, d<X>_s\right], \tag{4.14}$$

tenendo conto del fatto che

$$\begin{aligned}
&\sum_{k=1}^n H_k^2(\omega)\mathbb{E}\left[(X_{t_k} - X_{t_{k-1}})^2 | \mathcal{G}_{t_{k-1}}\right] + H_{n+1}^2(\omega)\mathbb{E}\left[(X_t - X_{t_n})^2 | \mathcal{G}_{t_n}\right] \\
&= \sum_{k=1}^n H_k^2(\omega)\mathbb{E}\left[<X>_{t_k} - <X>_{t_{k-1}} | \mathcal{G}_{t_{k-1}}\right] + H_{n+1}^2(\omega)\mathbb{E}\left[<X>_t - <X>_{t_n} | \mathcal{G}_{t_n}\right] \\
&= \sum_{k=1}^n H_k^2(\omega)\ \left[<X>_{t_k} - <X>_{t_{k-1}}\right] + H_{n+1}^2(\omega)\left[<X>_t - <X>_{t_n}\right] \\
&= \int_0^t f^2(s)\, d<X>_s
\end{aligned}$$

Questa identità è la base per poi definire l'integrale stocastico rispetto ad una martingala di quadrato integrabile, che ammetta come variazione quadratica (predicibile) il processo $<X>$. *Infine, si noti che, a sua volta, la relazione* (4.13) *si può esprimere dicendo che la variazione quadratica (predicibile) di* $\mathcal{I}_t^X(f)$ *coincide con l'integrale di* f *rispetto a* $d<X>_s$, *ossia*

$$<\mathcal{I}^X(f)>_t = \int_0^t f^2(s)\, d<X>_s\ .$$

Esercizio 4.1. *Dare la dimostrazione diretta del fatto che* $\mathcal{I}_{t'}^X(f)$ *e* $\mathcal{M}_t^f$ *sono martingale.*

soluzione: *La misurabilità e l'integrabilità di* $\mathcal{I}_{t'}^X(f)$ *(se le variabili* H_k *sono di quadrato integrabile) e di* $\mathcal{M}_t^f$ *(se le variabili* H_k *sono limitate) sono banali da verificare.*

Per la proprietà di martingala di $\mathcal{I}_{t'}^X(f)$ *si osservi che, se* $t' \le t_i \le t_j \le t''$, *posto* $t'_0 = t' \le t'_0 = t_i \le \cdots \le t'_{m-1} = t'_j \le t'_m = t''$, *e posto* $H'_{k+1} = f(t'_{k+1})$, *allora* H'_{k+1} *è* $\mathcal{G}_{t_k}$*-misurabile, per* $k = 0, \cdots, m-1$, *e*

$$\begin{aligned}
\tilde{\mathbb{E}}[I_{t''}^X(f) - I_{t'}^X(f) \mid \mathcal{G}_{t'}] &= \sum_{k=0}^m \mathbb{E}[H'_{k+1}(X_{t'_{k+1}} - X_{t'_k}) \mid \mathcal{G}_{t'_0}] = \sum_{k=0}^m \mathbb{E}\big[\mathbb{E}[H'_{k+1}(X_{t'_{k+1}} - X_{t'_k}) \mid \mathcal{G}_{t'_k}] \mid \mathcal{G}_{t'_0}\big] \\
&\underset{(H'_{k+1}\ \text{è}\ \mathcal{G}_{t_k}-mis.)}{=} \sum_{k=0}^m \mathbb{E}\big[H'_{k+1}\, \mathbb{E}[(X_{t'_{k+1}} - X_{t'_k}) \mid \mathcal{G}_{t'_k}] \mid \mathcal{G}_{t'_0}\big] \underset{(X_t\ \text{è una}\ \mathcal{G}_t-MG)}{=} \sum_{k=0}^m \mathbb{E}\big[H'_{k+1}\ 0\big] = 0
\end{aligned}$$

Per quanto riguarda $\mathcal{M}_t^f$, *con la stessa tecnica usata precedentemente, ci si convince subito che basta dimostrare che* $\mathbb{E}[\mathcal{M}_{t''}^f | \mathcal{G}_{t'}] = \mathcal{M}_{t'}^f$ *per due tempi* t' *e* t'', *che sono due tempi consecutivi di una partizione, ossia, per cui*

$$\mathcal{I}_{t''}^X(f) = \mathcal{I}_{t'}^X(f) + H'\,(X_{t''} - X_{t'})\,, \quad \text{con } H' = f(t''), \text{ che è } \mathcal{G}_{t'}\text{-misurabile.}$$

A questo punto basta osservare che

$$\begin{aligned}
\mathcal{M}_{t''}^f - \mathcal{M}_{t'}^f &= \Big(\mathcal{I}_{t'}^X(f) + H'\,(X_{t''} - X_{t'})\Big)^2 - \Big(\mathcal{I}_{t'}^X(f)\Big)^2 - (H')^2\,\mathbb{E}\left[(X_{t''} - X_{t'})^2 \mid \mathcal{G}_{t'}\right] \\
&= \Big(\mathcal{I}_{t'}^X(f)\Big)^2 + \big(H'\,(X_{t''} - X_{t'})\big)^2 - 2\,\mathcal{I}_{t'}^X(f)\,H'\ (X_{t''} - X_{t'}) - \Big(\mathcal{I}_{t'}^X(f)\Big)^2 - (H')^2\,\mathbb{E}\left[(X_{t''} - X_{t'})^2 \mid \mathcal{G}_{t'}\right] \\
&= (H')^2\,(X_{t''} - X_{t'})^2 - 2\,\mathcal{I}_{t'}^X(f)\,H'\,(X_{t''} - X_{t'}) - (H')^2\,\mathbb{E}\left[(X_{t''} - X_{t'})^2 \mid \mathcal{G}_{t'}\right]
\end{aligned}$$

da cui, passando ai valori attesi condizionali, e sfruttando il fatto che H' *e* $\mathcal{I}_{t'}^X(f)$ *sono* $G_{t'}$*-misurabili, si ottiene*

$$\begin{aligned}
\mathbb{E}\left[\mathcal{M}_{t''}^f - \mathcal{M}_{t'}^f \mid \mathcal{G}_{t'}\right] &= \mathbb{E}\left[(H')^2\,(X_{t''} - X_{t'})^2 - 2\,\mathcal{I}_{t'}^X(f)\,H'\ (X_{t''} - X_{t'}) - (H')^2\,\mathbb{E}\left[(X_{t''} - X_{t'})^2 \mid \mathcal{G}_{t'}\right] \mid \mathcal{G}_{t'}\right] \\
&= (H')^2\,\mathbb{E}\left[(X_{t''} - X_{t'})^2 \mid \mathcal{G}_{t'}\right] - 2\,\mathcal{I}_{t'}^X(f)\,H'\,\mathbb{E}\left[X_{t''} - X_{t'} \mid \mathcal{G}_{t'}\right] - (H')^2\,\mathbb{E}\left[(X_{t''} - X_{t'})^2 \mid \mathcal{G}_{t'}\right] = 0
\end{aligned}$$

.

Esercizio 4.2. *Si dimostri che non è necessario fare l'ulteriore ipotesi di limitatezza di f (ossia di H_k) per avere l'integrabilità della martingala $\mathcal{M}_t^f$, ma basta supporre che*

$$\mathbb{E}\left[\int_0^T f^2(s)\,d<X>_s\right] = \mathbb{E}\left[\sum_{k=1}^N H_k^2(\omega)\ \left[<X>_{t_k} - <X>_{t_{k-1}}\right]\right] < \infty$$

soluzione: *Posto $t'_k = t_k$ per $k \le n$ e $t'_{n+1} = t$ si ha*

$$\left(\sum_{k=1}^n H_k(\omega)\big(X_{t_k} - X_{t_{k-1}}\big) + H_{n+1}(\omega)\big(X_t - X_{t_n}\big)\right)^2 = \left(\sum_{k=1}^{n+1} H_k(\omega)\big(X_{t'_k} - X_{t'_{k-1}}\big)\right)^2$$
$$= \sum_{k=1}^{n+1}\Big(H_k(\omega)\big(X_{t'_k} - X_{t'_{k-1}}\big)\Big)^2 + \sum_{k=1}^{n+1}\sum_{j=1}^{n+1} H_k(\omega)\big(X_{t'_k} - X_{t'_{k-1}}\big)\, H_j(\omega)\big(X_{t'_j} - X_{t'_{j-1}}\big)$$

Ciascun termine del tipo

$$\big(H_k(\omega)\big)^2\,\big(X_{t'_k} - X_{t'_{k-1}}\big)^2$$

ammette valore atteso (finito o no, a priori) che è dato dal limite

$$\lim_{M\nearrow\infty}\mathbb{E}\left[(H_k(\omega)\wedge M)^2\,\big(X_{t'_k} - X_{t'_{k-1}}\big)^2\right] = \lim_{M\nearrow\infty}\mathbb{E}\left[(H_k(\omega)\wedge M)^2\,\mathbb{E}\big[\big(X_{t'_k} - X_{t'_{k-1}}\big)^2|\mathcal{G}_{t'_{k-1}}\big]\right]$$
$$= \lim_{M\nearrow\infty}\mathbb{E}\left[(H_k(\omega)\wedge M)^2\,\mathbb{E}\big[<X>_{t'_k} - <X>_{t'_{k-1}}\,|\mathcal{G}_{t'_{k-1}}\big]\right] = \lim_{M\nearrow\infty}\mathbb{E}\left[(H_k(\omega)\wedge M)^2\,\big(<X>_{t'_k} - <X>_{t'_{k-1}}\big)\right]$$
$$= \mathbb{E}\left[\big(H_k(\omega)\big)^2\big(<X>_{t'_k} - <X>_{t'_{k-1}}\big)\right]$$

Per l'integrabilità dei termini del tipo $H_k(\omega)\big(X_{t'_k} - X_{t'_{k-1}}\big)\,H_j(\omega)\big(X_{t'_j} - X_{t'_{j-1}}\big)$, basta poi usare la disuguaglianza di Cauchy.

L'esempio che segue è in realtà un'anticipazione, in quanto richiede la conoscenza del processo di Wiener, che in queste note si trova nei capitoli successivi. La lettura di questo esempio va quindi rinviata e deve essere effettuata dopo aver introdotto tale processo.

Esempio 4.12. *Se nell'Esempio 4.11 si prende $X_t = W_t$, il processo di Wiener standard, e $\mathcal{G}_t = \mathcal{F}_t^W$ (oppure $\mathcal{G}_t = \mathcal{F}_t$, se si tratta di un processo di Wiener rispetto alla filtrazione $(\mathcal{F}_t)_t$) si ottiene*[21] *che*

$$\mathbb{E}[(X_{t_k} - X_{t_{k-1}})^2|\mathcal{G}_{t_{k-1}}] = \mathbb{E}[(W_{t_k} - W_{t_{k-1}})^2|\mathcal{F}^W_{t_{k-1}}] = t_k - t_{k-1}$$

e quindi che

$$\mathcal{M}_t^f := \mathcal{I}_t^2(f) - \sum_k H_k^2(\omega)(t_k\wedge t - t_{k-1}\wedge t) = \Big(\int_0^t f(s)dW(s)\Big)^2 - \int_0^t f^2(s)ds \tag{4.15}$$

è una martingala, ed infine che

$$\mathbb{E}[\mathcal{I}_t^2(f)] = \mathbb{E}\left[\Big(\int_0^t f(s)dW(s)\Big)^2\right] = \mathbb{E}\left[\int_0^t f^2(s)ds\right]. \tag{4.16}$$

Si noti che saremmo arrivati alla stessa conclusione anche nel caso di una martingala $(X_t)_{t\ge 0}$ di quadrato integrabile, con variazione quadratica (predicibile) $<X>_t = t$. Se inoltre la martingala ha traiettorie continue, allora anche l'integrale stocastico $\mathcal{I}_t^X(f)$, sempre per processi elementari, ha traiettorie continue.

[21] Si ricordi che il processo di Wiener W_t è un processo ad incrementi indipendenti, e quindi $W_{t_k} - W_{t_{k-1}}$ è una variabile aleatoria gaussiana $N(0, t_k - t_{k-1})$ indipendente dalla σ-algebra $\mathcal{F}^W_{t_{k-1}}$. Di conseguenza

$$\mathbb{E}[(W_{t_k} - W_{t_{k-1}})^2|\mathcal{F}^W_{t_{k-1}}] = \mathbb{E}[(W_{t_k} - W_{t_{k-1}})^2] = t_k - t_{k-1}$$

in quanto $\mathbb{E}[(W_{t_k} - W_{t_{k-1}})^2] = Var(W_{t_k} - W_{t_{k-1}})$.

Conclusione In questa sezione abbiamo mostrato come si può definire l'integrale stocastico di $(f(s))_{s\in[0,T]}$ rispetto ad una $\mathcal{G}_t$-martingala X_t, con $(f(s))_{s\in[0,T]}$ **processo elementare $\mathcal{G}_t$-predicibile**, ovvero[22]

$$f(s,\omega) := \sum_{k=1}^{N} H_k(\omega)\mathbb{I}_{(t_{k-1},t_k]}(s),$$

per una partizione $\{t_k, k = 0, \cdots, N\}$ di $(0, T]$, con $0 = t_0 < t_1 < \cdots < t_N = T$, dove H_k sono variabili aleatorie $\mathcal{G}_{t_{k-1}}$-misurabili:

$$\mathcal{I}_t^X(f)(\omega) := \int_0^t f(s,\omega)dX(s,\omega) = \sum_{k=1}^{N} H_k(\omega)\big(X_{t_k\wedge t} - X_{t_{k-1}\wedge t}\big),$$

Se inoltre la martingala $(X_t)_{t\geq 0}$ è di quadrato integrabile, con variazione quadratica (predicibile) $< X >_t$ (ossia con $X_t^2- < X >_t$ una martingala), allora, nell'ipotesi che H_k siano variabili aleatorie limitate, oppure (si veda l'Esercizio 4.2) nell'ipotesi che

$$\mathbb{E}\left[\int_0^T f^2(s)\, d < X >_s\right] = \mathbb{E}\left[\sum_{k=1}^{N} H_k^2(\omega)\ \big[< X >_{t_k} - < X >_{t_{k-1}}\big]\right] < \infty,$$

si ottiene che

$$\big(\mathcal{I}_t^X(f)\big)^2 - \int_0^t f^2(s)\, d < X >_s, \qquad \text{è una martingala, a media nulla,}$$

ossia

$$< \mathcal{I}^X(f) >_t = \int_0^t f^2(s)\, d < X >_s\ .$$

Infine se $(X_t)_{t\geq 0}$ è una martingala di quadrato integrabile, con variazione quadratica (predicibile) $< X >_t = t$, ed è a traiettorie continue, allora il processo $\mathcal{I}_t^X(f)$ è a traiettorie continue e

$$< \mathcal{I}^X(f) >_t = \int_0^t f^2(s)\, ds,$$

per ogni funzione elementare con $\mathbb{E}\left[\int_0^t f^2(s)\, ds\right] < \infty$.

[22]**Attenzione:** nella parte relativa all'integrale stocastico (Sezione 9.5), al posto di H_k, che è $\mathcal{F}_{t_{k-1}}$-misurabile, si scrive C_{k-1}, che è $\mathcal{F}_{t_{k-1}}$-misurabile, ossia si considerano i processi del tipo

$$\sum_{k=1}^{N} C_{k-1}(\omega)\mathbb{I}_{(t_{k-1},t_k]}(s),$$

ponendo cioè $C_{k-1} = H_k$. Ciò può ingenerare qualche confusione, ma confidiamo nell'intelligenza del lettore....

4.3 Martingale, submartingale e tempi d'arresto

Definizione 4.4. . *Sia $\{\mathcal{F}_t\}$ una filtrazione e sia $\tau : \Omega \to \mathbb{R}^+ \cup \{\infty\}$, una variabile aleatoria (la v.a. τ deve prendere valori negli indici del tempo preso in considerazione e quindi, ad esempio, in $\mathbb{N}$ se si tratta tempo discreto, ma può anche prendere il valore infinito). La v.a. τ si dice **tempo d'arresto** (o **stopping time**) rispetto ad $\{\mathcal{F}_t\}$, se per ogni t*

$$\{\tau \leq t\} \in \mathcal{F}_t$$

*Si dice inoltre che τ è un **tempo d'arresto finito** se*

$$\mathbb{P}(\tau < +\infty) = 1$$

*e che τ è un **tempo d'arresto limitato** se esiste un numero $L< +\infty$ per cui*

$$\mathbb{P}(\tau < L) = 1$$

(Si noti che a volte la filtrazione in considerazione è ovvia, e quindi non viene specificata.)

Nel caso in cui l'insieme dei tempi sia $\mathbb{N}$ è equivalente chiedere che $\{\tau = n\} \in \mathcal{F}_n$ per ogni n, come si vede facilmente[23].

Esempio 4.13. *Con le stesse notazioni dell'Esempio 4.3, e per ogni $a \in \mathbb{R}$,*

$$\tau(\omega) = \inf\{n : S_n \geq a\}, \ \textit{(con la convenzione che } \inf\{\emptyset\} = +\infty)$$

è un tempo d'arresto rispetto ad $\mathcal{F}_n^S$, in quanto per decidere se l'evento $\{\tau \leq k\}$ si è verificato, basta esaminare le prime k v.a. $S_1, \cdots, S_k$.

Invece $\sigma(\omega) = \sup\{n \leq 10 : S_n \geq a\}$, se un tale n esiste e 10 altrimenti, non è un tempo d'arresto, in quanto, ad esempio, per decidere se l'evento $\{\sigma \leq 3\}$ si è verificato, bisogna esaminare tutte le v.a. $S_1, \cdots, S_{10}$ e non solo S_1, S_2, S_3, perciò $\{\sigma \leq 3\}$ non è misurabile rispetto a $\mathcal{F}_3^S$.

Definizione 4.5. ‡ *Dato un tempo d'arresto τ, si definisce* **la $\boldsymbol{\sigma}$-algebra degli eventi fino al tempo** τ *come*

$$\mathcal{F}_\tau = \{A \in \mathcal{F}_\infty, \ \textit{per cui } A \cap \{\tau \leq t\} \in \mathcal{F}_t \ \textit{per ogni } t\}$$

dove $\mathcal{F}_\infty = \bigvee_t \mathcal{F}_t$.

Si tratta cioè degli eventi, per i quali stabilire il loro verificarsi insieme al verificarsi di $\{\tau \leq t\}$ dipende solo dall'informazione disponibile fino al tempo t.
Nel caso in cui l'insieme dei tempi sia $\mathbb{N}$ è equivalente chiedere che $A \cap \{\tau = n\} \in \mathcal{F}_n$ per ogni n, come si vede facilmente[24].

---

[23]Ovviamente

$$\{\tau = n\} = \{\tau \leq n\} \backslash \{\tau \leq n-1\},$$

quindi, se $\{\tau \leq n\} \in \mathcal{F}_n$ e $\{\tau \leq n-1\} \in \mathcal{F}_{n-1} \subseteq \mathcal{F}_n$, allora $\{\tau = n\} \in \mathcal{F}_n$ per ogni n, mentre

$$\{\tau \leq n\} = \cup_{k=0}^n \{\tau = k\}$$

e quindi se $\{\tau = k\} \in \mathcal{F}_k$ per ogni k, essendo $\mathcal{F}_k \subseteq \mathcal{F}_n$ per $k \leq n$, si ha che $\{\tau \leq n\} \in \mathcal{F}_n$ per ogni n.

[24]Infatti: se $A \cap \{\tau = n\} \in \mathcal{F}_n$ per ogni n, allora

$$A \cap \{\tau \leq n\} = \cup_{k=0}^n A \cap \{\tau = k\}, \quad A \cap \{\tau = k\} \in \mathcal{F}_k \subseteq \mathcal{F}_n, \ k \leq n,$$

e quindi $A \cap \{\tau \leq n\} \in \mathcal{F}_n$.
Se invece $A \cap \{\tau \leq n\} \in \mathcal{F}_n$ per ogni n, allora

$$A \cap \{\tau = n\} = A \cap \{\tau \leq n\} \backslash A \cap \{\tau \leq n-1\}, \quad A \cap \{\tau \leq n-1\} \in \mathcal{F}_{n-1} \subseteq \mathcal{F}_n,$$

e quindi $A \cap \{\tau = n\} \in \mathcal{F}_n$.

Esercizio 4.3. ‡ *Controllare che $\mathcal{F}_\tau$ è una σ-algebra.*
(suggerimento: se $A \in \mathcal{F}_\tau$ allora l'evento $A^c \cap \{\tau \le t\} = \{\tau \le t\} \setminus \{A \cap \{\tau \le t\}\} \in \mathcal{F}_t$)

Esempio 4.14. *Se $\mathcal{F}_n = \sigma\{X_1, \cdots, X_n\}$ e $\tau = \inf\{n \text{ t.c. } X_n \in I\}$, con $I \in \mathcal{B}(\mathbb{R})$, allora τ è un $\mathcal{F}_n$ tempo d'arresto, infatti l'evento $\{\tau \le n\} = \{\exists k \le n \text{ t.c. } X_k \in I\} \in \mathcal{F}_n$.*

4.3.1 Tempo continuo‡

***Per i tempi di uscita, nel caso a tempo continuo, le cose non sono così semplici come nel caso a tempo discreto, però qualcosa si può dire. ***

Definizione 4.6. ‡ *Sia*

$$\tau_A = \inf\{t > 0 \text{ t.c. } X_t \notin A\},$$

con la convenzione che l'estremo inferiore dell'insieme vuoto è uguale a $+\infty$.

La v. a. τ_A è detta ***tempo di prima uscita da*** *A.*

Lemma 4.2. ‡ *Se X_t è un processo a traiettorie continue e A è aperto, allora τ_A è un tempo d'arresto.*

Dimostrazione. Si tratta di notare che, essendo A^c chiuso la funzione $x \mapsto \operatorname{dist}(x, A^c)$ è continua, e di conseguenza, essendo X_t a traiettorie continue, si ha che la funzione $s \mapsto \operatorname{dist}(X_s, A^c)$ è continua. Perciò

$$\inf_{0 \le s \le t} \operatorname{dist}(X_s, A^c) = \min_{0 \le s \le t} \operatorname{dist}(X_s, A^c)$$

e quindi

$$\{\tau_A > t\} = \{\inf_{0 \le s \le t} \operatorname{dist}(X_s, A^c) > 0\} = \bigcup_n \{\inf_{0 \le s \le t} \operatorname{dist}(X_s, A^c) > \frac{1}{n}\} =$$

$$= \bigcup_n \{\inf_{\substack{0 \le s \le t \\ s \in \mathbb{Q}}} \operatorname{dist}(X_s, A^c) \ge \frac{1}{n}\} = \bigcup_n \bigcap_{\substack{0 \le s \le t \\ s \in \mathbb{Q}}} \{\operatorname{dist}(X_s, A^c) \ge \frac{1}{n}\} \in \mathcal{F}_t.$$

Si noti che se il $\inf_{0 \le s \le t} \operatorname{dist}(X_s, A^c)$ non fosse un minimo, allora potrebbero verificarsi contemporaneamente gli eventi

$$\{\tau_A > t\} \text{ e } \{\inf_{0 \le s \le t} \operatorname{dist}(X_s, A^c) = 0\},$$

e la prima delle precedenti uguaglianze non sarebbe valida.

□

Nel caso in cui l'insieme A non sia aperto non è detto che τ_A sia un tempo d'arresto. Se A è un **insieme chiuso** allora τ_A è un **tempo d'arresto in senso debole**, ovvero

$$\{\tau_A < t\} \in \mathcal{F}_t \text{ per ogni } t \ge 0.$$

Osservazione 4.6. ‡ *Affermare che τ è un tempo d'arresto debole è equivalente ad affermare che è un tempo di arresto rispetto alla filtrazione*

$$\mathcal{F}_{t^+} := \bigcap_{s > t} \mathcal{F}_s.$$

Infatti in generale, se $\{\tau < t\} \in \mathcal{F}_t$ per ogni $t \ge 0$, allora, qualunque sia $m \ge 1$

$$\{\tau \le t\} = \bigcap_{n \ge m} \{\tau < t + \frac{1}{n}\} \in \mathcal{F}_{t + \frac{1}{m}},$$

e quindi $\{\tau \le t\} \in \mathcal{F}_s$ per ogni $s > t$, ovvero [25] *$\{\tau \le t\} \in \mathcal{F}_{t^+}$ per ogni t.*

[25] Alternativamente: per ogni $s > t$ si ha $\{\tau \le t\} = \bigcap_{n \ge 1} \{\tau < t + \frac{s-t}{n}\} \in \mathcal{F}_{t+(s-t)} = \mathcal{F}_s$

Lemma 4.3. ‡ *Se X_t è un processo con traiettorie continue a destra con limiti a sinistra (* cadlag *acronimo dal francese* continue à droite limite à gauche*), la filtrazione è continua a destra (cioè $\mathcal{F}_t = \mathcal{F}_{t^+} := \bigcap_{s>t} \mathcal{F}_s$) ed F è un chiuso allora τ_F è un tempo d'arresto.*

Dimostrazione. Basta dimostrare che τ_F è un tempo d'arresto in senso debole, e difatti

$$\{\tau_F \geq t\} = \bigcap_{0 \leq s < t} \{X_s \in F\},$$

ed essendo il processo X_t a traiettorie *cadlag* ed F chiuso si ha

$$\bigcap_{0 \leq s < t} \{X_s \in F\} = \bigcap_{\substack{0 \leq s < t \\ s \in \mathbb{Q}}} \{X_s \in F\} \in \mathcal{F}_t.$$

(infatti se $X_s \in F$ per ogni $s \in \mathbb{Q} \cap [0, t)$, allora $\forall r < t$ $X_r = \lim_{\substack{s \to r \\ s \in \mathbb{Q} \cap (r,t)}} X_s \in F$)

□

4.4 Alcune proprietà dei tempi d'arresto

1) Se τ e σ sono tempi d'arresto allora $\tau \wedge \sigma$ e $\tau \vee \sigma$ sono tempi d'arresto.

Infatti, per ogni t,

$$\{\tau \vee \sigma \leq t\} = \{\tau \leq t\} \cap \{\sigma \leq t\} \in \mathcal{F}_t \quad \text{e} \quad \{\tau \wedge \sigma \leq t\} = \{\tau \leq t\} \cup \{\sigma \leq t\} \in \mathcal{F}_t$$

2) ‡ Se τ è un tempo d'arresto allora la successione $\tau_n = \frac{\lceil \tau n \rceil}{n}$ è una successione di tempi d'arresto per cui $\tau_n \to \tau$, con $\tau_n \geq \tau$ per ogni n. (qui $\lceil$x$\rceil$ denota la parte intera superiore)

Infatti, $\forall t$, posto $t_n = \frac{\lfloor nt \rfloor}{n}$, risulta

$$\{\tau_n \leq t\} = \{\lceil \tau n \rceil \leq nt\} = \{\tau \leq \frac{\lfloor nt \rfloor}{n}\} \in \mathcal{F}_{t_n} \subseteq \mathcal{F}_t,$$

in quanto, posto $k = \lceil \tau n \rceil$, cioè $k - 1 < \tau n \leq k$, e $k \leq nt$, allora $\lfloor nt \rfloor \geq k \geq \tau n$, ovvero $\tau \leq \frac{\lfloor nt \rfloor}{n}$, ed ovviamente risulta $t_n \leq t$.
Per la convergenza di τ_n a τ, basta osservare che qualunque sia $x \in \mathbb{R}$ si ha

$$\frac{\lceil nx \rceil}{n} - \frac{1}{n} = \frac{\lceil nx \rceil - 1}{n} < x \leq \frac{\lceil nx \rceil}{n}$$

Si osservi che, se di prende $\widetilde{\tau}_n = \tau_{2^n}$, allora $\widetilde{\tau}_n \searrow \tau$, cioè $\widetilde{\tau}_n$ è anche una successione monotona non crescente, e che inoltre $\{\widetilde{\tau}_n\} \subseteq D$, dove D è l'insieme dei diadici. Infine va notato che tale risultato è interessante solo nel caso di tempi d'arresto che assumono valori in un insieme non discreto.

3) ‡ La v.a. τ è $\mathcal{F}_\tau$-misurabile.

Infatti, per ogni s, $\{\tau \leq$s$\} \in \mathcal{F}_\tau$ in quanto, qualunque sia t,

$$\{\tau \leq s\} \cap \{\tau \leq t\} = \{\tau \leq s \wedge t\} \in \mathcal{F}_{s \wedge t} \subseteq \mathcal{F}_t$$

4) ‡ Se τ e σ sono tempi d'arresto e $\mathbb{P}\{\sigma \leq \tau\} = 1$, ed $\mathcal{F}_0$ contiene tutti gli eventi trascurabili (cioè gli insiemi contenuti in insiemi di probabilità nulla), allora $\mathcal{F}_\sigma \subseteq \mathcal{F}_\tau$.

Infatti se $A \in \mathcal{F}_\infty$ e $A \cap \{\sigma \le t\} \in \mathcal{F}_t$ per ogni t, allora

$$A \cap \{\tau \le t\} = (A \cap \{\tau \le t\} \cap \{\sigma \le t\}) \cup (A \cap \{\tau \le t\} \cap \{\sigma > t\}) =$$

$$= ((A \cap \{\sigma \le t\}) \cap \{\tau \le t\}) \cup C \in \mathcal{F}_t$$

in quanto $C := (A \cap \{\tau \le t\} \cap \{\sigma > t\}) \in \mathcal{F}_0 \subseteq \mathcal{F}_t$, essendo un evento trascurabile, e $\{\tau \le t\}$ e $A \cap \{\sigma \le t\}$ sono in $\mathcal{F}_t$ per ipotesi.

Si osservi che se invece $\sigma \le \tau$ certamente, allora la condizione che $\mathcal{F}_0$ contenga tutti gli eventi trascurabili non è necessaria.

Per terminare questa sezione, osserviamo ***che, se una filtrazione soddisfa le cosiddette***condizioni "abituali"***, cioè

i) $\mathcal{F}_0$ contiene tutti gli eventi trascurabili, ***

ii) la filtrazione è continua a destra (cioè $\mathcal{F}_t = \mathcal{F}_{t^+} := \bigcap_{s>t} \mathcal{F}_s$),

***allora si possono applicare sia il Lemma 2, ed ottenere così dalla **ii)** che i tempi di uscita da un chiuso sono tempi d'arresto (e non solo tempi d'arresto in senso debole), sia la proprietà 4, e ottenere così dalla **i)** che $\mathcal{F}_\sigma \subseteq \mathcal{F}_\tau$ ogni volta che $\mathbb{P}(\sigma \le \tau) = 1$.

4.5 Caso a tempo discreto

Proposizione 4.4 (Martingale arrestate). *Data una $\mathcal{F}_n$-martingala (o submartingala) X_n ed un tempo d'arresto τ, il processo*

$$Y_n := X_{n\wedge\tau}$$

è una $\mathcal{F}_n$-martingala (o submartingala).

Nota bene: si usa anche la notazione

$$X_n^\tau := X_{n\wedge\tau},$$

e il processo $(X_n^\tau)_n$ è detto **martingala arrestata al tempo $\boldsymbol{\tau}$**.

Dimostrazione. Cominciamo con il dimostrare che Y_n è $\mathcal{F}_n$-misurabile e che è integrabile. Si noti che

$$Y_n = X_\tau I_{\{\tau \le n\}} + X_n I_{\{\tau > n\}},$$

e quindi

$$Y_n = X_0 I_{\{\tau=0\}} + X_1 I_{\{\tau=1\}} + X_2 I_{\{\tau=2\}} + \cdots + X_n I_{\{\tau=n\}} + X_n I_{\{\tau>n\}}, \tag{4.17}$$
$$Y_n = X_0 I_{\{\tau=0\}} + X_1 I_{\{\tau=1\}} + X_2 I_{\{\tau=2\}} + \cdots + X_n I_{\{\tau\ge n\}} \tag{4.18}$$

e che $\{\tau = k\} \in \mathcal{F}_k \subseteq \mathcal{F}_n, \{\tau > n\} \in \mathcal{F}_n$ e che X_k sono $\mathcal{F}_n$-misurabili per $k \le n$, da cui la $\mathcal{F}_n$-misurabilità di Y_n. Per ottenere l'integrabilità basta notare che dalla (4.18)

$$| Y_n |\le| X_0 | + | X_1 | + | X_2 | + \cdots + | X_n |$$

e sfruttare l'integrabilità delle X_k.
Per ottenere il resto dobbiamo notare che dalla (4.18), con $n+1$ al posto di n, si ha

$$Y_{n+1} = X_0 I_{\{\tau=0\}} + X_1 I_{\{\tau=1\}} + X_2 I_{\{\tau=2\}} + \cdots + X_n I_{\{\tau=n\}} + X_{n+1} I_{\{\tau>n\}}, \tag{4.19}$$

e quindi, confrontando (4.17) e (4.19)

$$Y_{n+1} - Y_n = (X_{n+1} - X_n) I_{\{\tau>n\}}$$

Di conseguenza, essendo $I_{\{\tau>n\}}$ una v.a. $\mathcal{F}_n$-misurabile, ed X_n una martingala,

$$\mathbb{E}[Y_{n+1} - Y_n \mid \mathcal{F}_n] = \mathbb{E}[X_{n+1} - X_n \mid \mathcal{F}_n] I_{\{\tau>n\}} = 0.$$

Nel caso in cui X_n sia una submartingala risulta ovviamente $\mathbb{E}[Y_{n+1} - Y_n \mid \mathcal{F}_n] \geq 0$.

Proposizione 4.5. *Data una $\mathcal{F}_n$-submartingala (o martingala) X_n ed un tempo d'arresto τ, limitato quasi certamente, ovvero per cui esiste un $n \in \mathbb{N}$, tale che*

$$\mathbb{P}(1 \leq \tau \leq n) = 1$$

allora

$$\mathbb{E}[X_1] \leq \mathbb{E}[X_\tau] \leq \mathbb{E}[X_n].$$

(Nel caso di martingale $\mathbb{E}[X_1] = \mathbb{E}[X_\tau] = \mathbb{E}[X_n]$.)

Dimostrazione. Per la proposizione precedente si ha che $\{X_{\tau\wedge k}\}_k$ è una submartingala, e quindi in particolare il valore medio è crescente (in senso lato). Poiché $X_{\tau\wedge 1} = X_1$, la prima disuguaglianza segue immediatamente.

Per la seconda disuguaglianza basta mostrare che $\mathbb{E}[X_n - X_\tau] \geq 0$ e infatti

$$X_n - X_\tau = \sum_{i=1}^{n} (X_n - X_\tau) I_{\{\tau=i\}} = \sum_{i=1}^{n} (X_n - X_i) I_{\{\tau=i\}}$$

e quindi

$$\mathbb{E}[X_n - X_\tau] = \sum_{i=1}^{n} \mathbb{E}[(X_n - X_i) I_{\{\tau=i\}}] = \sum_{i=1}^{n} \mathbb{E}[\mathbb{E}[(X_n - X_i) I_{\{\tau=i\}} \mid \mathcal{F}_i]]$$

La tesi segue in quanto

$$\mathbb{E}[(X_n - X_i) I_{\{\tau=i\}} \mid \mathcal{F}_i] = I_{\{\tau=i\}} \mathbb{E}[(X_n - X_i) \mid \mathcal{F}_i] \geq 0,$$

nel caso delle submartingale ($= 0$ nel caso delle martingale).

Questa sezione termina con una versione del famoso teorema del campionamento opzionale.

Teorema 4.6 (Optional Sampling Theorem o Teorema del campionamento opzionale). ‡ *Sia X_n una submartingala. Siano σ e τ due tempi d'arresto limitati, e tali che*

$$\sigma \leq \tau,$$

allora

$$\mathbb{E}[X_\tau | \mathcal{F}_\sigma] \geq X_\sigma.$$

Dimostrazione. Sia n tale che $\sigma \leq \tau \leq n$. Allora chiaramente

$$X_\tau = X_0 I_{\{\tau=0\}} + X_1 I_{\{\tau=1\}} + X_2 I_{\{\tau=2\}} + \cdots + X_n I_{\{\tau=n\}},$$
$$X_\sigma = X_0 I_{\{\sigma=0\}} + X_1 I_{\{\sigma=1\}} + X_2 I_{\{\sigma=2\}} + \cdots + X_n I_{\{\sigma=n\}}.$$

Ciò mostra che

$$|X_\tau| \leq |X_0| + |X_1| + |X_2| + \cdots + |X_n|,$$
$$|X_\sigma| \leq |X_0| + |X_1| + |X_2| + \cdots + |X_n|$$

e quindi in particolare che X_τ è integrabile, per cui ha senso calcolare la sua media condizionale.

Inoltre X_σ è $\mathcal{F}_\sigma$-misurabile, infatti, qualunque sia x, l'evento $\{X_\sigma \leq x\} \in \mathcal{F}_\sigma$, in quanto per ogni h

$$\{X_\sigma \leq x\} \cap \{\sigma = h\} = \{X_h \leq x\} \cap \{\sigma = h\} \in \mathcal{F}_h.$$

Infine per mostrare che $\mathbb{E}[X_\tau|\mathcal{F}_\sigma] \geq X_\sigma$, basta mostrare che

$$\mathbb{E}[X_\tau \mathbb{I}_A] \geq \mathbb{E}[X_\sigma \mathbb{I}_A], \quad \forall A \in \mathcal{F}_\sigma,$$

o equivalentemente (essendo $\tau = \tau \wedge n$ e $\sigma = \sigma \wedge n$) che

$$\mathbb{E}[X_{\tau\wedge n} \mathbb{I}_A] \geq \mathbb{E}[X_{\sigma\wedge n} \mathbb{I}_A], \quad \forall A \in \mathcal{F}_\sigma.$$

Infatti

$$\mathbb{E}[X_{\tau\wedge n} \mathbb{I}_A] = \sum_{h \leq n} \mathbb{E}[X_n^\tau \mathbb{I}_A \mathbb{I}_{\{\sigma=h\}}] *** = \sum_{h \leq n} \mathbb{E}[X_n^\tau \mathbb{I}_{A\cap\{\sigma=h\}}]. ***$$

Essendo A un evento $\mathcal{F}_\sigma$-misurabile, si ha che $A \cap \{\sigma = h\}$ è un ***evento ***di $\mathcal{F}_h$; ***inoltre, su tale insieme $\tau \wedge h = h$, in quanto $h = \sigma$ e $\sigma \leq \tau$; infine, il processo $(X_n^\tau)_n$ è una submartingala. ***Di conseguenza,

$$\begin{aligned}\mathbb{E}[X_n^\tau \mathbb{I}_{A\cap\{\sigma=h\}}] &\geq \mathbb{E}[X_h^\tau \mathbb{I}_{A\cap\{\sigma=h\}}] \\ &= \mathbb{E}[X_{\tau\wedge h} \mathbb{I}_A \mathbb{I}_{\{\sigma=h\}}] = \mathbb{E}[X_h \mathbb{I}_A \mathbb{I}_{\{\sigma=h\}}].\end{aligned}$$

Per ottenere la tesi basta osservare che, sommando su $h \leq n$, si ottiene

$$\mathbb{E}[X_\sigma \mathbb{I}_A].$$

□

4.6 Applicazione: la rovina del giocatore con le martingale

1) caso simmetrico
Sia Y_k una successione di v.a. indipendenti, con

$$\mathbb{P}(Y_k = 1) = \mathbb{P}(Y_k = -1) = \frac{1}{2},$$

e sia

$$S_n = \sum_{k=1}^n Y_k.$$

Sappiamo (vedi Esempio 4.2) che S_n è una martingala, in quanto se $p = 1/2$ allora $\mathbb{E}[Y_k] = 0$. Siano a e b numeri naturali non nulli e sia

$$\tau = \tau(a,b) := \inf\{n \text{ t.c. } S_n \notin (-a,b)\} = \inf\{n \text{ t.c. } S_n = -a \text{ o } S_n = b\}.$$

La variabile aleatoria τ è finita [26] con probabilità 1, cioè $\mathbb{P}(\tau < +\infty) = 1$, e quindi il gioco finisce in un tempo finito.

Per uno dei risultati precedenti sappiamo che $S_{n\wedge\tau}$ è una martingala e che quindi

[26] Si può dimostrare direttamente, anche nel caso generale, con $\mathbb{P}(Y_h = 1) = p$, che

$$\mathbb{P}(\tau = +\infty) = \lim_{n\to\infty} \mathbb{P}(\tau > n(a+b)) = 0.$$

Infatti, si ha $\{\tau = +\infty\} = \bigcap_{n\geq 1}\{\tau > n(a+b)\}$ e $\mathbb{P}(\tau > n(a+b)) \leq \alpha^n$ per un $\alpha < 1$. La prima uguaglianza è ovvia, mentre la seconda si può vedere facilmente osservando che

$$\{\omega \text{ tali che esiste } k < n \text{ per cui } Y_{k(a+b)+1} = Y_{k(a+b)+2} = \cdots = Y_{k(a+b)+a+b} = 1\} \subseteq \{\tau \leq n(a+b)\},$$

$$\mathbb{E}[S_{n\wedge\tau}] = \mathbb{E}[S_{1\wedge\tau}] = \mathbb{E}[S_1] = 0$$

Inoltre $S_{n\wedge\tau} \to S_\tau$ per $n \to +\infty$, e $|S_{n\wedge\tau}| \leq \max(a,b)$ e quindi per il teorema della convergenza dominata

$$\mathbb{E}[S_{n\wedge\tau}] \to \mathbb{E}[S_\tau].$$

Di conseguenza

$$\mathbb{E}[S_\tau] = -a\mathbb{P}(S_\tau = -a) + b\,(1 - \mathbb{P}(S_\tau = -a)) = 0,$$

da cui immediatamente

$$\mathbb{P}(S_\tau = -a) = \frac{b}{a+b}.$$

2) caso generale
Come nell'applicazione **1)** ma con $\mathbb{P}(Y_k = 1) = p$ e $\mathbb{P}(Y_k = -1) = q = 1-p$.

Si procede[27] in modo analogo al caso precedente, ma questa volta si prende come martingala

$$Z_n = \exp\{\theta S_n - n\psi(\theta)\}$$

dove

$$\exp\{\psi(\theta)\} = \mathbb{E}[\exp\{\theta Y_1\}] = \exp\{\theta\}p + \exp\{-\theta\}q.$$

Si cerca, se esiste, θ in modo che $\psi(\theta) = 0$ ovvero, posto $\exp\{\theta\} = \alpha$, si cerca $\alpha p + \alpha^{-1}q = 1$, ovvero $\alpha^2 p - \alpha + q = 0$. Ciò è possibile solo per

$$\alpha = \frac{1 \pm \sqrt{1-4pq}}{2p} = \frac{1 \pm \sqrt{1-4p+4p^2}}{2p} = \frac{1 \pm \sqrt{1-4p+4p^2}}{2p} = \frac{1 \pm |1-2p|}{2p} = \frac{1 \pm (1-2p)}{2p}$$

ovvero per $\alpha = 1$ o $\alpha = \frac{q}{p}$ (come del resto si può vedere subito, anche direttamente). Il caso $\alpha = 1$ corrisponderebbe a $Z_n \equiv 1$ e non porterebbe ad alcun risultato, mentre $\exp\{\theta\} = \alpha = \frac{q}{p}$ corrisponde a $Z_n = \left(\frac{q}{p}\right)^{S_n}$.

Di nuovo, sempre per convergenza dominata,

$$1 \equiv \mathbb{E}[Z_{n\wedge\tau}] \to \mathbb{E}[Z_\tau] = \left(\frac{q}{p}\right)^{-a}\mathbb{P}(S_\tau = -a) + \left(\frac{q}{p}\right)^{b}[1 - \mathbb{P}(S_\tau = -a)] = 1,$$

da cui di nuovo si può ricavare, posto $\rho = \frac{q}{p}$

$$\mathbb{P}(S_\tau = -a) = \frac{1-\rho^b}{\rho^{-a} - \rho^b} = \frac{\rho^a - \rho^{b+a}}{1 - \rho^{b+a}} = \frac{\rho^{b+a} - \rho^a}{\rho^{b+a} - 1}.$$

---

e che quindi $\quad \mathbb{P}(\bigcup_{k<n} \{Y_{k(a+b)+1} = Y_{k(a+b)+2} = \cdots = Y_{k(a+b)+a+b} = 1\}) \leq \mathbb{P}(\tau \leq n(a+b))$

ovvero, passando ai complementari, e utilizzando l'indipendenza delle v.a. Y_h,

$$\mathbb{P}(\bigcap_{k<n} \{Y_{k(a+b)+1} = Y_{k(a+b)+2} = \cdots = Y_{k(a+b)+a+b} = 1\}^c) = \prod_{k<n} \mathbb{P}(\{Y_{k(a+b)+1} = \cdots = Y_{k(a+b)+a+b} = 1\}^c)$$
$$= \prod_{k<n} (1 - p^{a+b}) = (1-p^{a+b})^n = \alpha^n \geq \mathbb{P}(\tau > n(a+b)).$$

La dimostrazione è finita in quanto α^n tende a zero.
Si noti che in sostanza la precedente dimostrazione si riduce a dimostrare che $\tau \leq T(a+b)$, dove T è la variabile aleatoria geometrica di parametro $\beta = p^{a+b} = 1 - \alpha$, definita come

$$T(\omega) = k+1 \quad \Leftrightarrow \quad ***\omega \in B_k \cap \Big(\cap_{h<k} B_h^c\Big)$$

dove $B_k = \{Y_{k(a+b)+1} = \cdots = Y_{k(a+b)+a+b} = 1\}^c$. Allora la v.a. τ è finita, essendo la v.a. T finita.

[27]In questo caso $\mathbb{E}[Y_k] = 2p-1 \neq 0$. Se si prendesse la martingala a media nulla $M_n := S_n - (2p-1)n$ ci sarebbero due problemi: il primo è che la corrispondente martingala arrestata $M_{n\wedge\tau}$ non è limitata, per cui pur convergendo a $M_\tau = S_\tau - (2p-1)\tau$, **non** possiamo immediatamente dire che anche i valori attesi di $M_{n\wedge\tau}$ convergono al valore atteso di M_τ,
il secondo è che, anche se avessimo dimostrato che i valori attesi convergono, e quindi si avesse che $0 = \mathbb{E}[M_\tau] = \mathbb{E}[S_\tau - (2p-1)\tau]$, non avremmo finito, in quanto dovremmo conoscere anche il valore atteso di τ.

Si noti che

$$\mathbb{P}(S_\tau = -a) \to \frac{b}{a+b}$$

per $\rho \to 1$, cioè per $p \to \frac{1}{2}$.

***Per concludere questa sezione proponiamo un esercizio: mostrare che, per $p \neq \frac{1}{2}$, la funzione $h(x) := \left(\frac{q}{p}\right)^x$ è armonica rispetto alla matrice P delle probabilità di transizione della passeggiata aleatoria. Nel caso per $p = \frac{1}{2}$, mostrare che lo stesso vale per la funzione $h(x) = x$. ***

4.7 Disuguaglianza di Kolmogorov per submartingale non negative

Proposizione 4.7 (Disuguaglianza di Kolmogorov). *Sia X_n una submartingala non negativa allora*

(i)
$$\mathbb{P}(\max(X_1, X_2, \cdots, X_n) > \gamma) \leq \frac{\mathbb{E}[X_n]}{\gamma}$$

Sia X_n una martingala con $\mathbb{E}[|X_n|^\alpha] < +\infty, \alpha \geq 1$, allora

(ii)
$$\mathbb{P}(\max(|X_1|, |X_2|, \cdots, |X_n|) > \gamma) \leq \frac{\mathbb{E}[|X_n|^\alpha]}{\gamma^\alpha}$$

Dimostrazione. Cominciamo con il primo caso. Si definisca

$$\begin{cases} \tau := \inf\{k \text{ tali che } 1 \leq k \leq n, X_k > \gamma\}, & \text{se un tale } k \text{ esiste} \\ \tau := n, & \text{altrimenti.} \end{cases}$$

Ovviamente τ è un tempo d'arresto e $\{X_\tau > \gamma\} = \{\max(X_1, \cdots, X_n) > \gamma\}$ e quindi per la disuguaglianza di Markov

$$\mathbb{P}(\max(X_1, X_2, \cdots, X_n) > \gamma) = \mathbb{P}(X_\tau > \gamma) \leq \frac{\mathbb{E}[|X_\tau|]}{\gamma} = \frac{\mathbb{E}[X_\tau]}{\gamma}$$

Basta quindi mostrare che $\mathbb{E}[X_\tau] \leq \mathbb{E}[X_n]$, ***precedente Proposizione 4.5, ***in quanto τ è a valori in $1, 2, \cdots, n$.

Per il caso delle martingale, la tesi segue osservando che

$$\mathbb{P}(\max(|X_1|, |X_2|, \cdots, |X_n|) > \gamma) = \mathbb{P}(\max(|X_1|^\alpha, |X_2|^\alpha, \cdots, |X_n|^\alpha) > \gamma^\alpha),$$

la funzione $|x|^\alpha$, per $\alpha \geq 1$ è convessa e quindi $|\mathrm{X}_n|^\alpha$ è una submartingala non negativa (confrontare proprietà 4) e infine applicando la disuguaglianza precedente.

4.8 Convergenza di martingale

***Nell'esempio della rovina del giocatore abbiamo trovato che delle martingale limitate per le quali esisteva il limite per n che converge ad infinito: le martingale $S_{n\wedge\tau}$ nel caso simmetrico e $\left(\frac{q}{p}\right)^{S_{n\wedge\tau}}$, nel caso generale. Questo fatto è un caso particolare di un risultato più generale. ***

Proposizione 4.8. *Sia X_n una martingala uniformemente limitata in L^1, cioè tale che*

$$\sup_n \mathbb{E}[|X_n|] \leq M < +\infty.$$

Allora

$$\mathbb{P}(\{\omega \ t.c. \ \exists \ finito \lim_{n\to\infty} X_n(\omega)\}) = 1$$

Dimostrazione. ‡ Daremo la dimostrazione solo nel caso in cui valga una ipotesi più forte:

$$\sup_n \mathbb{E}[|X_n|^2] \leq M < +\infty.$$

(si noti infatti che in tale caso $\mathbb{E}[|X_n|] \leq \Big(\mathbb{E}[|X_n|^2]\Big)^{1/2} \leq M$)

Basta mostrare che la successione X_n è una successione di Cauchy con probabilità 1. Ciò significa che

$$\{\omega \text{ t.c. } \forall \epsilon > 0 \exists m \geq 1 \text{ t.c.} \forall k \geq 1 |X_{m+k}(\omega) - X_m(\omega)| \leq \epsilon\}$$

ovvero

$$\bigcap_{\epsilon>0} \bigcup_{m\geq 1} \bigcap_{k\geq 1} \{|X_{m+k} - X_m| \leq \epsilon\}$$

è un insieme misurabile di probabilità 1. La misurabilità segue osservando che è equivalente prendere ϵ razionale positivo. Per calcolare la probabilità, ricordiamo che in generale se B_h sono eventi $\mathbb{P}(\bigcap_{h\geq 1} B_h) = 1$ se e solo se $\mathbb{P}(B_h) = 1 \forall h \geq 1$.

Infatti se $\mathbb{P}(\bigcap_{h\geq 1} B_h) = 1$ allora, poiché $\bigcap_{h\geq 1} B_h \subseteq B_{\overline{h}}$, ne segue che $\mathbb{P}(B_{\overline{h}}) = 1$, comunque fissato $\overline{h} \geq 1$. Se viceversa $\mathbb{P}(B_h) = 1 \forall h \geq 1$, allora

$$\mathbb{P}(\bigcap_{h\geq 1} B_h) = 1 - \mathbb{P}(\bigcup_{h\geq 1} B_h^c), e \mathbb{P}(\bigcup_{h\geq 1} B_h^c) \leq \sum_{h\geq 1} \mathbb{P}(B_h^c) = 0.$$

Di conseguenza la tesi equivale a mostrare che, per ogni $\epsilon > 0$,

$$\mathbb{P}(\bigcup_{m\geq 1} \bigcap_{k\geq 1} \{|X_{m+k} - X_m| \leq \epsilon\}) = 1,$$

ovvero che, per ogni $\epsilon > 0$,

$$\mathbb{P}(\bigcap_{m\geq 1} \bigcup_{k\geq 1} \{|X_{m+k} - X_m| > \epsilon\}) = 0.$$

Si osservi che,$\forall \overline{m} \geq 1$,

$$\bigcap_{m\geq 1} \bigcup_{k\geq 1} \{|X_{m+k} - X_m| > \epsilon\} \subseteq \bigcup_{k\geq 1} \{|X_{\overline{m}+k} - X_{\overline{m}}| > \epsilon\}$$

e che

$$\mathbb{P}(\bigcup_{k\geq 1} \{|X_{\overline{m}+k} - X_{\overline{m}}| > \epsilon\}) = \lim_{n\to\infty} \mathbb{P}(\bigcup_{k=1}^{n} \{|X_{\overline{m}+k} - X_{\overline{m}}| > \epsilon\}).$$

Di conseguenza

$$\mathbb{P}(\bigcap_{m\geq 1} \bigcup_{k\geq 1} \{|X_{m+k} - X_m| > \epsilon\}) \leq \lim_{n\to\infty} \mathbb{P}(\bigcup_{k=1}^{n} \{|X_{\overline{m}+k} - X_{\overline{m}}| > \epsilon\}),$$

e quindi

$$\mathbb{P}(\bigcap_{m\geq 1} \bigcup_{k\geq 1} \{|X_{m+k} - X_m| > \epsilon\}) \leq \lim_{\overline{m}\to\infty} \lim_{n\to\infty} \mathbb{P}(\bigcup_{k=1}^{n} \{|X_{\overline{m}+k} - X_{\overline{m}}| > \epsilon\}).$$

Non rimane che dimostrare che

$$\lim_{\overline{m}\to\infty} \lim_{n\to\infty} \mathbb{P}(\bigcup_{k=1}^{n} \{|X_{\overline{m}+k} - X_{\overline{m}}| > \epsilon\}) = 0.$$

Si osservi ora che $\overline{X}_k := X_{\overline{m}+k} - X_{\overline{m}}$, è una martingala rispetto alla filtrazione $\mathcal{G}_k := \mathcal{F}_{\overline{m}+k}$ e che

$$\bigcup_{k=1}^{n} \{|X_{\overline{m}+k} - X_{\overline{m}}| > \epsilon\} = \{\max(|\bar{X}_1|, |\bar{X}_2|, \cdots, |\bar{X}_n|) > \epsilon\}.$$

Basta quindi applicare la disuguaglianza di Kolmogorov per $\alpha = 2$ per ottenere che

$$\mathbb{P}(\bigcup_{k=1}^{n} \{|X_{\overline{m}+k} - X_{\overline{m}}| > \epsilon\}) \leq \frac{\mathbb{E}[|X_{\overline{m}+n} - X_{\overline{m}}|^2]}{\epsilon^2} = \frac{\mathbb{E}[X_{\overline{m}+n}^2] - \mathbb{E}[X_{\overline{m}}^2]}{\epsilon^2}.$$

Infatti

$$\mathbb{E}[|X_{\overline{m}+n} - X_{\overline{m}}|^2] = \mathbb{E}[X^2_{\overline{m}+n}] + \mathbb{E}[X^2_{\overline{m}}] - 2\mathbb{E}[X_{\overline{m}+n}X_{\overline{m}}] =$$

$$= \mathbb{E}[X^2_{\overline{m}+n}] + \mathbb{E}[X^2_{\overline{m}}] - 2\mathbb{E}[\mathbb{E}[X_{\overline{m}+n}X_{\overline{m}} \mid \mathcal{F}_{\overline{m}}] =$$

$$= \mathbb{E}[X^2_{\overline{m}+n}] + \mathbb{E}[X^2_{\overline{m}}] - 2\mathbb{E}[X_{\overline{m}}\mathbb{E}[X_{\overline{m}+n} \mid \mathcal{F}_{\overline{m}}] = \mathbb{E}[X^2_{\overline{m}+n}] + \mathbb{E}[X^2_{\overline{m}}] - 2\mathbb{E}[X^2_{\overline{m}}]$$

A questo punto si noti che $\mathbb{E}[X^2_{\overline{m}}]$ è una successione convergente ad un numero $\mu \leq M$, in quanto è una successione limitata per ipotesi, e monotona non decrescente ($X^2_{\overline{m}}$ è una submartingala). Quindi

$$\lim_{\overline{m}\to\infty} \lim_{n\to\infty} \mathbb{P}(\bigcup_{k=1}^{n} \{|X_{\overline{m}+k} - X_{\overline{m}}| > \epsilon\}) \leq \lim_{\overline{m}\to\infty} \lim_{n\to\infty} \frac{\mathbb{E}[X^2_{\overline{m}+n}] - \mathbb{E}[X^2_{\overline{m}}]}{\epsilon^2} = \lim_{\overline{m}\to\infty} \frac{\mu - \mathbb{E}[X^2_{\overline{m}}]}{\epsilon^2} = 0.$$

4.9 Disuguaglianza di Doob

Proposizione 4.9 (Disuguaglianza di Doob). *Sia data una submartingala non negativa X_n, con $\mathbb{E}[(X_n)^p] < +\infty$, per un $p > 1$. Posto $X^*_n = \max_{k\leq n}(X_k)$ vale la seguente disuguaglianza:*

$$\mathbb{E}[(X^*_n)^p] \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}[(X_n)^p].$$

Nel caso di una martingala X_n, con $\mathbb{E}[|X_n|^p] < +\infty$, e posto $X^*_n = \max_{k\leq n}(|X_k|)$ vale la seguente disuguaglianza:

$$\mathbb{E}[|X^*_n|^p] \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}[|X_n|^p])$$

Dimostrazione. Si definisca

$$\begin{cases} \tau_\gamma := \inf\{k \text{ tali che } 1 \leq k \leq n, X_k > \gamma\} & \text{se un tale } k \text{ esiste,} \\ \tau_\gamma := n+1 & \text{altrimenti.} \end{cases}$$

Ovviamente τ_γ è un tempo d'arresto e l'evento $\{X^*_n > \gamma\}$ coincide con $\bigcup_{k=1}^{n}\{\tau_\gamma = k\}$.
Quindi per ogni $\beta > 0$

$$(X^*_n)^\beta = \int_0^{X^*_n} \beta\gamma^{\beta-1} d\gamma = \int_0^\infty \beta\gamma^{\beta-1} I_{\{X^*_n > \gamma\}} d\gamma = \int_0^\infty \beta\gamma^{\beta-1} \sum_{k=1}^{n} I_{\{\tau_\gamma = k\}} d\gamma$$

Passando al valore medio, per $\beta > 0$,

$$\mathbb{E}[(X^*_n)^\beta] = \sum_{k=1}^{n} \int_0^\infty \beta\gamma^{\beta-2}\mathbb{E}[\gamma I_{\{\tau_\gamma = k\}}] d\gamma \leq \sum_{k=1}^{n} \int_0^\infty \beta\gamma^{\beta-2}\mathbb{E}[X_k I_{\{\tau_\gamma = k\}}] d\gamma$$

in quanto $\gamma I_{\{\tau_\gamma = k\}} \leq X_k I_{\{\tau_\gamma = k\}}$. Inoltre, essendo X_n una submartingala $X_k \leq \mathbb{E}[X_n \mid \mathcal{F}_k]$, non negativa, ed essendo $\{\tau_\gamma = k\} \in \mathcal{F}_k$ si ha

$$\mathbb{E}[X_k I_{\{\tau_\gamma = k\}}] \leq \mathbb{E}[\mathbb{E}[X_n \mid \mathcal{F}_k] I_{\{\tau_\gamma = k\}}] = \mathbb{E}[\mathbb{E}[X_n I_{\{\tau_\gamma = k\}} \mid \mathcal{F}_k]] = \mathbb{E}[X_n I_{\{\tau_\gamma = k\}}]$$

e quindi

$$\mathbb{E}[(X^*_n)^\beta] \leq \sum_{k=1}^{n} \int_0^\infty \beta\gamma^{\beta-2}\mathbb{E}[X_n I_{\{\tau_\gamma = k\}}] d\gamma = \int_0^\infty \beta\gamma^{\beta-2}\mathbb{E}[X_n \sum_{k=1}^{n} I_{\{\tau_\gamma = k\}}] d\gamma$$

$$= \mathbb{E}[\int_0^\infty \beta\gamma^{\beta-2} X_n I_{\{X^*_n > \gamma\}} d\gamma],$$

ovvero, se $\beta - 1 > 0$,

$$\mathbb{E}[(X_n^*)^\beta] \leq \beta \mathbb{E}[X_n \int_0^\infty \gamma^{\beta-2} I_{\{X_n^* > \gamma\}} d\gamma] = \frac{\beta}{\beta - 1} \mathbb{E}[X_n (X_n^*)^{\beta-1}].$$

Inoltre, prendendo $\beta = p$, e usando la disuguaglianza di Hölder [28] si ottiene la tesi osservando che

$$\begin{aligned}\mathbb{E}[(X_n^*)^p] &\leq \frac{p}{p-1} \mathbb{E}[(X_n)^p]^{1/p} \mathbb{E}\left[\left((X_n^*)^{p-1}\right)^{p/(p-1)}\right]^{(p-1)/p} \\ &= \frac{p}{p-1} \mathbb{E}[(X_n)^p]^{1/p} \mathbb{E}[(X_n^*)^p]^{(p-1)/p}.\end{aligned}$$

Infatti questa disuguaglianza non è banale (ovvero non è del tipo $+\infty \leq +\infty$), in quanto si ha che $\mathbb{E}[(X_n^*)^p] = \mathbb{E}[\max_{k=1,\cdots,n}(X_k)^p]] \leq \mathbb{E}[\sum_{k=1}^n (X_k)^p] < +\infty$ (il caso banale in cui $\mathbb{E}[(X_n^*)^p] = 0$ si può escludere, in questo caso la tesi del teorema è banalmente vera, ed il caso è poco interessante), e non rimane che dividere ambo i membri della precedente disuguaglianza per $\mathbb{E}[(X_n^*)^p]^{(p-1)/p}$.

□

Infine va menzionato il fatto che la disuguaglianza di Doob, così come quella di Kolmogorov, si estendono al caso di submartingale X_t a tempo continuo e con traiettorie continue, considerando che

$$\max_{t \in [0,T]} |X_t| = \lim_{n \to \infty} \max_{t=(i/2^n)T, i \leq 2^n} |X_t|.$$

Per maggiori dettagli ***si veda la seguente Sezione 4.10.

4.10 Estensioni alle martingale a tempo continuo (cenni)‡***

Quasi tutti i risultati delle sezioni precedenti si estendono al caso di martingale (o submartingale) continue o *cadlag* (cioè continue a destra con limite a sinistra).

Ad esempio, la disuguaglianza di Doob, così come quella di Kolmogorov, si estendono al caso di submartingale X_t a tempo continuo e con traiettorie continue, considerando che

$$\max_{t \in [0,T]} |X_t| = \lim_{n \to \infty} \max_{t=(i/2^n)T, i \leq 2^n} |X_t|.$$

Per quanto riguarda le submartingale *cadlag* basta passare dal massimo all'estremo superiore:

$$\sup_{t \in [0,T]} |X_t| = \lim_{n \to \infty} \max_{t=(i/2^n)T, i \leq 2^n} |X_t|.$$

Per ottenere le disuguaglianze va anche osservato, che la successione $X^{*,n} = \max_{t=(i/2^n)T, i \leq 2^n} |X_t|$ è una successione monotona, e la monotonia permette di passare al limite sotto il segno di valore atteso (o di integrale rispetto a $\mathbb{P}$).

Un poco più delicato è il problema dell'estensione dei risultati che coinvolgono X_τ, quando τ è un tempo d'arresto, come ad esempio il fatto che X_τ è una variabile aleatoria, e addirittura una variabile aleatoria $\mathcal{F}_\tau$-misurabile. A questo problema dedichiamo una sezione alla fine di questo capitolo (Sezione 4.11).

Ad esempio, vale il seguente risultato:

Lemma 4.10. *Un processo stocastico X_t, con traiettorie* cadlag, *e $\mathcal{F}_t$-adattato è una martingala se e solo se*

$$\mathbb{E}[X_\tau] = \mathbb{E}[X_0], \qquad \textit{per ogni tempo d'arresto limitato } \tau. \tag{4.20}$$

[28]Si ricordi che, per $p > 1$ si ha

$$\frac{1}{p} + \frac{1}{q} = 1 \quad \Leftrightarrow \quad q = \frac{p}{p-1},$$

e che la disuguaglianza di Hölder garantisce che per tutte le variabili aleatorie

$$\mathbb{E}[|XY|] \leq \mathbb{E}[|X|^p]^{\frac{1}{p}} \mathbb{E}[|Y|^q]^{\frac{1}{q}},$$

con la convenzione che se almeno uno dei due valori attesi $\mathbb{E}[|X|^p]$ o $\mathbb{E}[|Y|^q]$ non è finito, allora il secondo membro vale infinito, e la disuguaglianza è banale. Si tratta di una generalizzazione della disuguaglianza di Cauchy, che corrisponde al caso $p = q = 2$.

Dimostrazione. La parte più semplice è quella relativa alla condizione sufficiente: se vale la (4.20), allora, prima di tutto X_t è una variabile aleatoria integrabile, per ogni t, dato che t è un tempo d'arresto rispetto a qualunque filtrazione. Inoltre per ogni $s \leq t$, e per ogni $A \in \mathcal{F}_s$, posto

$$\tau = t\,\mathbb{I}_A + s\,\mathbb{I}_{A^c},$$

allora τ è un $\mathcal{F}_t$-tempo d'arresto e quindi

$$\mathbb{E}[X_\tau] = \mathbb{E}[X_t\,\mathbb{I}_A] + \mathbb{E}[X_s\,\mathbb{I}_{A^c}] = \mathbb{E}[X_0]$$

Ovviamente anche

$$\mathbb{E}[X_s] = \mathbb{E}[X_s\,\mathbb{I}_A] + \mathbb{E}[X_s\,\mathbb{I}_{A^c}] = \mathbb{E}[X_0],$$

e quindi, per ogni $s \leq t$, e per ogni $A \in \mathcal{F}_s$, si ha

$$\mathbb{E}[X_t\,\mathbb{I}_A] = \mathbb{E}[X_s\,\mathbb{I}_A]$$

ovvero X_t è una $\mathcal{F}_t$-martingala.

Per quanto riguarda l'altra implicazione, spieghiamo solo l'idea: se τ è un tempo d'arresto limitato da L, allora anche $\tau_m = \frac{\lceil m\tau \rceil}{m}$ è una successione di tempi d'arresto limitata e a valori in $\{0, \frac{1}{m}, \cdots \frac{\lceil mL \rceil}{m}\}$. Il processo a tempo discreto $Y_k = X_{k/m}$ è una martingala a tempo discreto rispetto alla filtrazione $\mathcal{G}_k = \mathcal{F}_{k/m}$. Allora sappiamo che

$$\mathbb{E}[X_{\tau_m}] = \mathbb{E}[Y_{\lceil m\tau \rceil}] = \mathbb{E}[Y_0] = \mathbb{E}[X_0]$$

Inoltre, essendo X_t a traiettorie continue a destra si ha che la successione X_{τ_m} converge a X_τ. Per ottenere la tesi basta osservare che

$$\lim_{m\to\infty} \mathbb{E}[X_{\tau_m}] = \mathbb{E}[X_\tau],$$

in quanto la successione X_{τ_m} è una successione uniformemente integrabile. L'ultima affermazione segue dall'optional sampling theorem, che assicura che $X_{\tau_m} = \mathbb{E}[X_{L+1}|\mathcal{F}_{\tau_m}]$, e dal fatto, per ogni variabile aleatoria integrabile Z in uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$, la famiglia dei valori attesi condizionali $\{\mathbb{E}[Z|\mathcal{G}],\ \mathcal{G}$ sotto σ-algebre di $\mathcal{F}\}$ è una famiglia uniformemente integrabile.

□

L'optional sampling theorem si generalizza al caso di martingale a tempo continuo e *cadlag*. Può essere interessante però sapere che esiste una generalizzazione, anche al caso di martingale uniformemente integrabili e di tempi d'arresto generali: per tali martingale esiste $X_\infty = \lim_{t\to\infty} X_t$ e si ha che per ogni coppia di tempo d'arresto $\sigma \leq \tau$, non necessariamente finiti, si ha

$$X_\sigma = \mathbb{E}[X_\tau|\mathcal{F}_\sigma] = \mathbb{E}[X_\infty|\mathcal{F}_\sigma].$$

4.11 Processi misurabili e tempi d'arresto‡***

Lo scopo principale di questa sezione è affrontare il seguente problema: sotto quali condizioni sul processo X_t possiamo affermare che, se τ è un tempo d'arresto, allora X_τ è una variabile aleatoria, e/o una variabile aleatoria $\mathcal{F}_\tau$-misurabile?

A questo scopo ricordiamo la seguente definizione

Definizione 4.7 (Processi (congiuntamente) misurabili). *Sia* $X : [0,T] \times \Omega :\mapsto \mathbb{R}$, $(t,\omega) \mapsto X(t,\omega)\big(= X_t(\omega)\big)$. *Sia* $\mathcal{B}([0,T]) \times \mathcal{F}$ *la* σ*-algebra prodotto su* $[0,T] \times \Omega$ *, ovvero la* σ*-algebra generata da* $\{J \times A \subseteq [0,T] \times \Omega;\ J \in \mathcal{B}([0,T]), A \in \mathcal{F}\}$. *Il processo* $(X_t)_t$ *si dice* **congiuntamente misurabile***, o più semplicemente* **misurabile***, se la funzione* X *è* $\mathcal{B}([0,T]) \times \mathcal{F}$*-misurabile. La definizione è analoga per i processi definiti per tutti i tempi* $t \geq 0$*: basta sostituire* $[0,T]$ *con* $[0,\infty)$.

Cominciamo con l'osservare che se τ è una variabile aleatoria a valori in $[0,T]$ (non necessariamente un tempo d'arresto) e $(X_t)_{t\in[0,T]}$ è un processo misurabile, allora X_τ è una variabile aleatoria. Infatti basta considerare che la funzione

$$\big(\Omega,\mathcal{F}\big) \to \big([0,T]\times\Omega, \mathcal{B}([0,T])\times\mathcal{F}\big) \qquad \omega \mapsto \big(\tau(\omega),\omega\big)$$

è misurabile, e che $\omega \mapsto X_\tau(\omega) = X(\tau(\omega),\omega)$ è misurabile, in quanto composizione di due funzioni misurabili.

Nel caso in cui τ sia un tempo d'arresto (rispetto alla filtrazione $\{\mathcal{F}_t\}$), ciò dimostra solo che X_τ è una variabile aleatoria $\mathcal{F}$-misurabile, ma non ci assicura che sia una variabile aleatoria $\mathcal{F}_\tau$-misurabile. Per ottenere ciò dobbiamo assumere un'ulteriore proprietà sul processo X_t.

Definizione 4.8 (Processi $\{\mathcal{F}_t\}$-progressivamente misurabili). *Sia $X : [0,T]\times\Omega :\mapsto \mathbb{R}$, $(t,\omega) \mapsto X(t,\omega)\big(= X_t(\omega)\big)$. Sia $\mathcal{B}([0,t])\times\mathcal{F}_t$ la σ-algebra prodotto su $[0,t]\times\Omega$, ovvero la σ-algebra generata da $\{J\times A \subseteq [0,t]\times\Omega;\ J\in\mathcal{B}([0,t]), A\in\mathcal{F}_t\}$. Il processo $(X_t)_t$ si dice* **progressivamente misurabile** *se, per ogni $t\in[0,T]$, la funzione X ristretta a $[0,t]\times\Omega$ è $\mathcal{B}([0,t])\times\mathcal{F}_t$-misurabile. La definizione è analoga per i processi definiti per tutti i tempi $t\geq 0$: basta sostituire $[0,T]$ con $[0,\infty)$.*

Lemma 4.11. *Se X_t è un processo $\{\mathcal{F}_t\}$-progressivamente misurabile, e τ è un $\{\mathcal{F}_t\}$-tempo d'arresto a valori in $[0,T]$, allora la variabile aleatoria X_τ, è una variabile aleatoria $\mathcal{F}_\tau$-misurabile.*

Dimostrazione. La dimostrazione è banale, infatti basta dimostrare che, per ogni $t\in[0,T]$ la variabile aleatoria $X_\tau\,\mathbb{I}_{\{\tau\leq t\}}$ è $\mathcal{F}_t$-misurabile. Ma ciò è banalmente vero in quanto, posta $\overline{X}^t$ la funzione X ristretta a $[0,t]\times\Omega$, cioè $\overline{X}^t(s,\omega) = X(s,\omega)$, per $s\in[0,t]$ e $\omega\in\Omega$, allora $X_\tau\,\mathbb{I}_{\{\tau\leq t\}} = \overline{X}^t_{\tau\wedge t}\,\mathbb{I}_{\{\tau\leq t\}}$. La tesi segue osservando che $\tau\wedge t$ è una variabile $\mathcal{F}_t$- misurabile e $\overline{X}^t$ è un processo congiuntamente misurabile rispetto a $\mathcal{B}([0,t])\times\mathcal{F}_t$. □

Per finire basta osservare che ogni processo $\mathcal{F}_t$-adattato e *cadlag* (oppure continuo a sinistra) è progressivamente misurabile. Infatti si ha che, per ogni $s\leq t$ $X(s,\omega) = \lim_{n\to\infty} X^{(n,t)}(s,\omega)$,dove, posto $s_k^{(n,t)} = \frac{k}{2^n}\,t$, si ha

$$X^{(n,t)}(s,\omega) := X(0,\omega)\,\mathbf{1}_{\{0\}}(s) + \sum_{k=1}^{2^n} X(s_k^{(n,t)},\omega)\,\mathbf{1}_{(s_{k-1}^{(n,t)},s_k^{(n,t)}]}(s), \qquad s\leq t$$

che, chiaramente, sono processi $\mathcal{B}([0,t])\times\mathcal{F}_t$-misurabili. Ciò mostra che il processo X ristretto all'intervallo $[0,t]$ è anch'esso $\mathcal{B}([0,t])\times\mathcal{F}_t$-misurabile. Per l'arbitrarietà di t, si ottiene la progressiva misurabilità.

Il caso dei processi continui a sinistra si può trattare in modo simile, prendendo però la successione

$$X^{(n,t)}(s,\omega) := X(0,\omega)\,\mathbf{1}_{\{0\}}(s) + \sum_{k=1}^{2^n} X(s_{k-1}^{(n,t)},\omega)\,\mathbf{1}_{[s_{k-1}^{(n,t)},s_k^{(n,t)})}(s), \qquad s\leq t.$$

Tuttavia in questo caso c'è un metodo più semplice: infatti $X(t,\omega) = \lim_{n\to\infty} X(\frac{\lfloor 2^n t\rfloor}{2^n},\omega)$, in quanto $\frac{\lfloor 2^n t\rfloor}{2^n}\leq t$ e $\lim_{n\to\infty}\frac{\lfloor 2^n t\rfloor}{2^n} = t$. I processi $X^{(n)} : (t,\omega)\mapsto X(\frac{\lfloor 2^n t\rfloor}{2^n},\omega)$ sono chiaramente progressivamente misurabili, come si vede subito da:

$$X^{(n)}(t,\omega) := \sum_{k\geq 1} X(t_{k-1}^{(n)},\omega)\,\mathbf{1}_{[t_{k-1}^{(n)},t_k^{(n)})}(t),$$

dove $t_k^{(n)} = k/2^n$.

Per maggiori dettagli e ulteriori generalizzazioni si consiglia, ad esempio, il testo di P. Baldi [**?**].

Capitolo 5

Mercato *(B,S)*: investimenti, proprietà e caratteristiche

5.1 Struttura del mercato *(B,S)*

Si considera un mercato che opera sotto condizioni di incertezza, rappresentabili in termini probabilistici da uno spazio di probabilità dotato di filtrazione

$$(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\geq 0}, \mathbb{P}).$$

Come al solito il flusso $\mathcal{F} = (\mathcal{F}_n)_{n\geq 0}$ di σ-algebre può essere interpretato come il flusso di informazioni $\mathcal{F}_n$ accessibile fino all'istante n, con $n \geq 0$.
Si definisce *mercato*-(B,S) la coppia di elementi formata da $d+1$ operazioni finanziarie:

un	conto bancario o bond		B	*titolo non rischioso*
d	azioni o stocks		$S = (S^1, \dots, S^d)$	*titoli rischiosi*

in cui si assume che l'evoluzione del conto bancario sia descrivibile tramite una successione stocastica (strettamente) positiva

$$B = (B_n)_{n\geq 0}, \qquad \text{con } B_n > 0 \text{ per ogni } n,$$

dove le variabili B_n sono $\mathcal{F}_{n-1}$-misurabili per ogni $n \geq 1$, e B_0 è $\mathcal{F}_{-1} = \mathcal{F}_0$-misurabile, ovvero il processo $B = (B_n)_{n\geq 0}$ è ***predicibile***[1] rispetto alla filtrazione $(\mathcal{F}_n)$.
Anche la dinamica dei valori dell'i-esimo titolo S^i può essere descritta da una successione stocastica

$$S^i = (S^i_n)_{n\geq 0}$$

dove però le S^i_n sono variabili $\mathcal{F}_n$-misurabili per ogni $n \geq 0$, ovvero i processi $S^i = (S^i_n)_{n\geq 0}$ sono ***adattati***[2] rispetto alla filtrazione $(\mathcal{F}_n)$.
Dalle definizioni date si vede che esiste una differenza sostanziale tra un conto bancario e un'azione; ovvero, la $\mathcal{F}_{n-1}$-misurabilità di B_n indica che lo stato del conto bancario al tempo n risulta già conosciuto (si hanno tutte le informazioni) al tempo $n-1$: la successione di variabili aleatorie (B_n) è detta predicibile appunto per questa sua caratteristica. La situazione per i prezzi delle azioni è differente: le variabili S^i_n sono $\mathcal{F}_n$-misurabili, ciò significa che i loro valori attuali si possono determinare solo dopo che risultano disponibili tutte le informazioni $\mathcal{F}_n$ arrivate fino al tempo n.
Tali considerazioni permettono di capire perché si dice che un conto bancario è un'operazione non rischiosa mentre le azioni sono dei titoli rischiosi.

[1] Come al solito il lettore può pensare all'esempio in cui $\mathcal{F}_n = \sigma\{\xi_k,\, k \leq n\}$, e allora ciò significa che $B_n = b_n(\xi_0, \xi_1, \cdots, \xi_{n-1})$.
[2] Come al solito il lettore può pensare all'esempio in cui $\mathcal{F}_n = \sigma\{\xi_k,\, k \leq n\}$, e allora ciò significa che $S^i_n = s^i_n(\xi_0, \xi_1, \cdots, \xi_n)$.

Ponendo

$$r_n \quad := \quad \frac{B_n - B_{n-1}}{B_{n-1}} \quad = \quad \frac{\Delta B_n}{B_{n-1}} \qquad tasso\, di\, interesse$$

$$\rho_n^i \quad := \quad \frac{S_n^i - S_{n-1}^i}{S_{n-i}^i} \quad = \quad \frac{\Delta S_n^i}{S_{n-1}^i} \qquad rendimento\, dell'azione\, i,\ per\, i = 1, 2, \cdots, d,$$

si può scrivere

$$\Delta B_n = r_n B_{n-1}, \tag{5.1}$$

$$\Delta S_n^i = \rho_n^i S_{n-1}^i,\ per\ i = 1, 2, \cdots, d, \tag{5.2}$$

dove le r_n sono $\mathcal{F}_{n-1}$-misurabili (ovvero predicibili) e le ρ_n^i sono $\mathcal{F}_n$-misurabili (ovvero adattati).
Allora per $n \geq 1$ si ha

$$B_n = B_0 \prod_{k=1}^{n} (1 + r_k) \tag{5.3}$$

$$S_n^i = S_0^i \prod_{k=1}^{n} (1 + \rho_k^i). \tag{5.4}$$

Si noti che l'ipotesi (ragionevole) che B_n sia strettamente positivo si traduce nell'ipotesi che $r_k > -1$ per ogni k. Inoltre spesso per comodità si assume che $B_0 = 1$.

5.1.1 Strategia di investimento di un portfolio

Si consideri un investitore sul mercato (B, S) che può:

1 depositare o prendere soldi dal conto bancario

2 vendere e comprare azioni.

Si assume che un trasferimento di denaro da un'operazione ad un'altra non richieda costi di transazione e che le operazioni risultino infinitamente divisibili, cioè che l'investitore possa comprare, o vendere, qualunque porzione di azione e prelevare, o depositare, qualsiasi ammontare dal conto bancario.
Si vogliono ora introdurre alcune definizioni in merito alle operazioni, alle posizioni e alle strategie finanziarie possibili in tale mercato.

Definizione 5.1 (Portfolio o Strategia d'investimento). *Una successione stocastica predicibile*

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

dove $\beta = (\beta_n(\omega))_{n\geq 0}$ *e* $\gamma = (\gamma_n^1(\omega), \ldots, \gamma_n^d(\omega))_{n\geq 0}$ *sono predicibili, ovvero tali che* $\beta_n(\omega)$ *e* $\gamma_n^i(\omega)$ *risultano* $\mathcal{F}_{n-1}$*-misurabili per ogni* $n \geq 0$ *e* $i = 1, \ldots, d$ *è detta* **investimento di un portfolio** *(o più semplicemente* **portfolio***) sul mercato* (B, S)*.*
Per enfatizzare il dinamismo al quale risulta sottoposto l'investimento di un portfolio viene spesso usato il termine **strategia di investimento.** *(o più semplicemente* **strategia***)*

Si osservi che le variabili $\beta_n(\omega)$ e $\gamma_n^i(\omega)$ possono essere positive, nulle e anche negative; quest'ultimo caso indica che l'investitore può prendere un prestito dalla banca e vendere azioni che non possiede (***vendita a corto***, o ***vendita allo scoperto***, o ***short sale***).
Un altro punto importante da sottolineare è che la $\mathcal{F}_{n-1}$-misurabilità indica che le variabili $\beta_n(\omega)$ e $\gamma_n^i(\omega)$, che descrivono la posizione finanziaria dell'investitore al tempo n (ovvero l'ammontare presente sul conto bancario e le azioni in suo possesso), sono determinabili tramite le informazioni ottenibili fino al tempo $n - 1$ mentre ***non*** dipendono da quelle relative al tempo n: (*la posizione di domani è completamente definita dalla situazione presente oggi*).
Il tempo $n = 0$ gioca un ruolo importante; infatti la predicibilità in tale istante, formalmente equivalente alla $\mathcal{F}_{-1}$-misurabilità, corrisponde alla $\mathcal{F}_0$-misurabilità, si assume cioè $\mathcal{F}_{-1} = \mathcal{F}_0$.

Si è già assunto $n \geq 0$ verrà fatta anche l'ipotesi aggiuntiva di tempo limitato (***orizzonte finito***), ovvero $n \leq N$ da cui segue che si considereranno solo gli istanti di tempo $0 \leq n \leq N$.
Nell'indicare una generica successione $(a_n)_{n=0}^N$ si useranno, equivalentemente, le espressioni: $(a_n)_{n\geq 0}$, $(a_n)_{0\leq n\leq N}$ o addirittura (a_n).

Definizione 5.2 (Valore di un portfolio). *Il processo aleatorio* **valore associato a una strategia di investimento $\boldsymbol{\pi}$** *è la successione stocastica:*

$$X^\pi = (X_n^\pi)_{n\geq 0}$$

dove, all'istante n

$$X_n^\pi = \beta_n B_n + \sum_{i=1}^d \gamma_n^i S_n^i. \tag{5.5}$$

La definizione (5.5) significa solo che se al tempo n possediamo β_n titoli bancari (o bond), il cui valore unitario[3] è B_n e γ_n^i azioni del titolo i, che al tempo n vale[4] S_n^i, allora il valore totale del portfolio è la somma dei valori investiti nei singoli titoli.

Per evitare di appesantire le notazioni si indicherà con $\gamma_n S_n$ il prodotto scalare dei vettori $\gamma_n = (\gamma_n^1, \ldots, \gamma_n^d)$ e $S_n = (S_n^1, \ldots, S_n^d)$ per cui la (5.5) diviene

$$X_n^\pi = \beta_n B_n + \gamma_n S_n. \tag{5.6}$$

Si assume inoltre che le uniche strategie ammissibili siano le strategie autofinanzianti, ovvero se non ci sono consumi o costi di transazione, e neppure ulteriori investimenti esterni, ma ad ogni istante tutto (e solo) il valore viene semplicemente reinvestito nel mercato[5].

Definizione 5.3 (Strategia autofinanziante). *Una strategia $\pi = (\beta, \gamma)$ si dice* **autofinanziante** *(o* **self-financing***), se per ogni n*

$$\beta_{n-1} B_{n-1} + \gamma_{n-1} S_{n-1} = \beta_n B_{n-1} + \gamma_n S_{n-1}, \tag{5.7}$$

o equivalentemente

$$(\beta_n - \beta_{n-1}) B_{n-1} + (\gamma_n - \gamma_{n-1}) S_{n-1} = 0. \tag{5.8}$$

Da questa definizione si ottiene immediatamente che

$$\begin{aligned} X_n^\pi - X_{n-1}^\pi &= \beta_n B_n + \gamma_n S_n - (\beta_{n-1} B_{n-1} + \gamma_{n-1} S_{n-1}) \\ &= \beta_n B_n + \gamma_n S_n - (\beta_n B_{n-1} + \gamma_n S_{n-1}) \\ &= \beta_n (B_n - B_{n-1}) + \gamma_n (S_n - S_{n-1}) \\ &= \beta_n \Delta B_n + \gamma_n \Delta S_n. \end{aligned}$$

[3] Allora il valore globale investito in banca (o nel bond) è $\beta_n B_n$
[4] Allora il valore globale investito nel titolo i è $\gamma_n^i S_n^i$.
[5] Per capire il significato della (5.7) si può ragionare così : nell'istante $n-1$ si ha che il valore del portfolio è

$$X_{n-1}^\pi = \beta_{n-1} B_{n-1} + \gamma_{n-1} S_{n-1},$$

durante l'intervallo $(n-1, n)$ si cambia strategia e ci si ritrova con β_n titoli bancari, che però valgono ancora B_{n-1}, e con γ_n^i titoli i, che però valgono ancora S_{n-1}^i. Il fatto che la strategia sia autofinanziante si riflette nel fatto che per fare questa operazione non ho bisogno di prendere altro denaro da fuori l'investimento (ad esempio comprando azioni con il ricavato di un lavoro) e neppure di effettuare consumi (ad esempio per comprare beni, come case, automobili, etc.) o di pagare dei costi (si ipotizza che non ci siano costi di transazione), ma tutto il denaro necessario per avere il nuovo portfolio, ovvero

$$\beta_n B_{n-1} + \gamma_n S_{n-1}$$

è esattamente il valore X_{n-1}^π del portfolio all'istante $n-1$.

Dalla precedente relazione si ottiene quindi che

$$\begin{aligned} X_n^\pi - X_0^\pi &= \sum_{k=1}^{n} (X_k^\pi - X_{k-1}^\pi = \\ &= \sum_{k=1}^{n} (\beta_k \Delta B_k + \gamma_k \Delta S_k) \\ &= \sum_{k=1}^{n} \beta_k \Delta B_k + \sum_{k=1}^{n} \gamma_k \Delta S_k. \end{aligned}$$

Da quanto osservato si può concludere che il capitale guadagnato (o *capital gains*) $X_n^\pi - X_0^\pi$, tramite l'investimento di un portfolio π autofinanziante, può essere descritto da una successione $G^\pi = (G_n^\pi)_{n\geq 0}$, dove

$$G_0^\pi = 0 \qquad \text{e} \qquad G_n^\pi = \sum_{k=1}^{n} (\beta_k \Delta B_k + \gamma_k \Delta S_k), \tag{5.9}$$

quindi il valore del portfolio al tempo n è dato dall'espressione

$$X_n^\pi = X_0^\pi + G_n^\pi, \tag{5.10}$$

si arriva così alla seguente definizione alternativa ed equivalente[6]:

Definizione 5.4 (Strategia autofinanziante bis). *Un portfolio π è* **autofinanziante** *se il suo valore $X^\pi = (X_n^\pi)_{n\geq 0}$ può essere rappresentato dalla somma:*

$$X_n^\pi = X_0^\pi + \sum_{k=1}^{n} (\beta_k \Delta B_k + \gamma_k \Delta S_k) \qquad n \geq 1. \tag{5.13}$$

La classe delle strategie π autofinanzianti verrà indicata con *SF* (da ***self-financing***).

Osservazione 5.1. *Le ipotesi fondamentali che caratterizzano il mercato finanziario considerato, ovvero:*

- ▷ **assenza di costi di transazione***;*
- ▷ **titoli infinitamente divisibili***: non ci sono limiti sulle quantità minime dei titoli trattati;*
- ▷ **possibilità di vendite allo scoperto (short sale)***: è possibile vendere titoli che non si possiedono, ciò equivale a ipotizzare che è sempre consentito assumere la posizione di debitore;*

---

[6]Si osservi inoltre che prese due successioni arbitrarie $a = (a_n)_{n\geq 0}$ e $b = (b_n)_{n\geq 0}$ sussiste la *regola di differenziazione discreta*:

$$\Delta(a_n b_n) = a_n \Delta b_n + b_{n-1} \Delta a_n, \tag{5.11}$$

dove si è posto $\Delta a_n = a_n - a_{n-1}$. Applicando questo risultato alla parte destra della (5.6) si ottiene l'espressione:

$$\begin{aligned} \Delta X_n^\pi &= \Delta(\beta_n B_n) + \Delta(\gamma_n S_n) \\ &= [\beta_n \Delta B_n + \gamma_n \Delta S_n] + [B_{n-1} \Delta \beta_n + S_{n-1} \Delta \gamma_n]. \end{aligned} \tag{5.12}$$

Ciò mostra che il cambiamento del valore del portfolio ($\Delta X_n^\pi = X_n^\pi - X_{n-1}^\pi$) dipende, in generale, da due fattori:

1 dalla *variazione* dovuta al *conto bancario* e ai *prezzi delle azioni*

$$\beta_n \Delta B_n + \gamma_n \Delta S_n;$$

2 dalla *variazione* della composizione del *portfolio* (*cambio di strategia*)

$$B_{n-1} \Delta \beta_n + S_{n-1} \Delta \gamma_n.$$

Risulta, dunque, che in caso di autofinanziamento (una sorta di autarchia) il reale cambio di valore di X_n^π sia sempre dovuto alle variazioni ΔB_n e ΔS_n e non $\Delta \beta_n$ e $\Delta \gamma_n$ (si assume, cioè, che le posizioni finanziarie prese dall'investitore non siano soggette a cambiamenti). Si vede immediatamente che assumere π autofinanziante equivale al verificarsi della seguente condizione

$$B_{n-1} \Delta \beta_n + S_{n-1} \Delta \gamma_n = 0 \qquad n \geq 1.$$

Si osservi infine che con la definizione bis di strategia autofinanziante, non è importante conoscere il valore esplicito di (β_0, γ_0), ma bastano i valori di (β_n, γ_n) per $n = 1, \cdots, N$.

▷ **assenza di rischio di insolvenza (default risk)**: *si assume che i contratti di compravendita stipulati vengano sicuramente onorati;*

possono essere sintetizzate dicendo che il **mercato è privo di frizionalità** *(per maggiori dettagli si consulti, ad esempio, il testo di Moriconi [12]).*

5.1.2 Mercato scontato $(\widetilde{B}, \widetilde{S})$

Si consideri un mercato (B, S) come descritto precedentemente. È interessante notare che partendo da questo è sempre possibile[7] costruire un nuovo mercato $(\widetilde{B}, \widetilde{S})$ detto ***mercato scontato*** o ***attualizzato*** tale che

$$\widetilde{B} = (\widetilde{B}_n)_{n\geq 0}, \qquad \text{con} \qquad \widetilde{B}_n \equiv 1$$

e

$$\widetilde{S} = (\widetilde{S}_n)_{n\geq 0}, \qquad \text{con} \qquad \widetilde{S}_n = \frac{S_n}{B_n}.$$

Allora preso un portfolio $\pi = (\beta, \gamma)$ il suo ***valore scontato***

$$\widetilde{X}^\pi_n = (\widetilde{X}^\pi_n)_{n\geq 0}, \qquad \text{con} \qquad \widetilde{X}_n = \frac{X_n}{B_n}$$

risulta dato dall'espressione

$$\widetilde{X}^\pi_n = \beta_n \widetilde{B}_n + \gamma_n \widetilde{S}_n = \beta_n + \gamma_n \widetilde{S}_n \tag{5.14}$$

in quanto

$$\widetilde{X}^\pi_n = \frac{X^\pi_n}{B_n} = \frac{1}{B_n}\,(\beta_n B_n + \gamma_n S_n),$$

e nel caso in cui π è autofinanziante nel mercato (B, S) si ha che questa proprietà si trasmette al mercato $(\widetilde{B}, \widetilde{S})$, infatti la condizione (5.8) che caratterizza le strategie autofinanzianti è equivalente[8] a

$$(\beta_n - \beta_{n-1})\widetilde{B}_{n-1} + (\gamma_n - \gamma_{n-1})\widetilde{S}_{n-1} = 0 \tag{5.15}$$

In altre parole una strategia è autofinanziante nel mercato (B, S) se e solo se è autofinanziante nel mercato attualizzato $(\widetilde{B}, \widetilde{S})$.

Pertanto vale la (5.13) attualizzata dove, cioè, al posto di X, B, S si considera $\widetilde{X}$, $\widetilde{B}$, $\widetilde{S}$ ed essendo $\Delta\widetilde{B}_k \equiv 0$ segue, inoltre, che per $\pi \in SF$

$$\widetilde{X}^\pi_n = \widetilde{X}^\pi_0 + \sum_{k=1}^{n} \gamma_k \Delta\widetilde{S}_k, \tag{5.16}$$

quindi l'espressione esplicita del valore del portfolio è

$$\widetilde{X}^\pi_n = \widetilde{X}^\pi_0 + \sum_{k=1}^{n}\sum_{i=1}^{d} \gamma^i_k \Delta\widetilde{S}^i_k, \qquad \widetilde{S}^i_k = \frac{S^i_k}{B_k}. \tag{5.17}$$

Allora si conclude considerando la (5.14) e la (5.15) che il valore scontato $\widetilde{X}^\pi = \left(\widetilde{X}^\pi_n\right)_{n\geq 0} = \left(\frac{X^\pi_n}{B_n}\right)_{n\geq 0}$ con $\pi \in SF$ soddisfa la relazione

$$\Delta\widetilde{X}^\pi_n = \gamma_n \Delta\widetilde{S}_n, \quad \text{ovvero} \quad \Delta\left(\frac{X^\pi_n}{B_n}\right) = \gamma_n \Delta\left(\frac{S_n}{B_n}\right). \tag{5.18}$$

[7] Si ricordi che per ipotesi $B_n > 0$ per ogni n.

[8] Chiaramente la condizione (5.8), ovvero $(\beta_n - \beta_{n-1})B_{n-1} + (\gamma_n - \gamma_{n-1})S_{n-1} = 0$, rimane invariata se si divide per B_{n-1}, ovvero

$$(\beta_n - \beta_{n-1})\widetilde{B}_{n-1} + (\gamma_n - \gamma_{n-1})\widetilde{S}_{n-1} = \frac{1}{B_{n-1}}\,[(\beta_n - \beta_{n-1})B_{n-1} + (\gamma_n - \gamma_{n-1})S_{n-1}] = 0.$$

Da quanto illustrato risulta chiaro che nel caso si sappia a priori che $B_n > 0$ con $n \geq 1$ si può semplificare la struttura del mercato, senza ledere in generalità, assumendo $B_n \equiv 1$, invece di passare al mercato scontato.

Dopo aver dato le nozioni base di mercato (B, S) un punto fondamentale che si vuole sottolineare è che si considererà *sempre e soltanto* un mercato che non prevede opportunità di arbitraggio (non esistono *free-lunch*, cioè pasti gratis), ovvero:

Definizione 5.5. *Un mercato è detto* **privo di opportunità di arbitraggio** *(o* **arbitrage-free***), o anche* **razionale** *(o* **fair***), se non permette di ottenere guadagni sicuri, cioè se non si possono ottenere dei profitti senza sottoporsi a dei rischi*

N.B. La relazione illustrata nella (5.18), nella sua semplicità, risulterà assumere un ruolo chiave in molti calcoli relativi al concetto di arbitrage-free, ciò spiega il perché dell'introduzione del mercato scontato.

5.2 Nozione di copertura. Prezzo superiore e inferiore.

Sia $f_N = f_N(\omega)$ una funzione non negativa $\mathcal{F}_N$-misurabile detta ***obbligazione*** o ***pay-off (prezzo di pagamento) terminale***[9], in questo paragrafo vengono illustrati alcuni concetti di ***copertura*** (o ***hedge***) per il pay-off terminale.

Definizione 5.6 (Copertura superiore, inferiore e perfetta). *Un portfolio $\pi = (\beta, \gamma)$ con $\beta = (\beta_n)$, $\gamma = (\gamma_n)$, per $n = 0, 1, \ldots, N$ è detto* **copertura superiore** *per (x, f_N), con $x \geq 0$, se il valore iniziale del portfolio è x e se $X_N^\pi \geq f_N$, $\mathbb{P}$ q.c., ovvero se*

$$X_0^\pi = x \qquad e \qquad \mathbb{P}(X_N^\pi \geq f_N) = 1.$$

Un portfolio $\pi = (\beta, \gamma)$ è invece detto **copertura inferiore** *per (x, f_N),con $x \geq 0$, se il valore iniziale del portfolio è x e se $X_N^\pi \leq f_N$, $\mathbb{P}$ q.c., ovvero*

$$X_0^\pi = x \qquad e \qquad \mathbb{P}(X_N^\pi \leq f_N) = 1.$$

Infine un portfolio $\pi = (\beta, \gamma)$ è invece detto **copertura perfetta** *per (x, f_N), con $x \geq 0$, se è sia una copertura inferiore che una copertura superiore, una copertura, cioè se il valore iniziale del portfolio è x e se $X_N^\pi = f_N$, $\mathbb{P}$ q.c., ovvero*

$$X_0^\pi = x \qquad e \qquad \mathbb{P}(X_N^\pi = f_N) = 1.$$

Il concetto di ***copertura*** gioca un ruolo di fondamentale importanza in ambito finanziario in quanto rappresenta uno strumento di protezione che permette di ottenere un livello di garanzia relativamente a un certo investimento.

Tramite le successive definizioni si potranno formalizzare le azioni che si devono compiere per assicurarsi un livello di garanzia. Si indichi con

$$\begin{aligned} H^*(x, f_N; \mathbb{P}) &= \{\pi \, \textit{ammissibili}: \, X_0^\pi = x, \quad X_N^\pi \geq f_N, \quad \mathbb{P} \text{ q.c.}\} \\ &= \{\pi \, \textit{ammissibili}: \, X_0^\pi = x, \quad \mathbb{P}(X_N^\pi \geq f_N) = 1\} \end{aligned}$$

la ***classe delle coperture superiori per (x, f_N)*** e con

$$\begin{aligned} H_*(x, f_N; \mathbb{P}) &= \{\pi \, \textit{ammissibili}: \, X_0^\pi = x, \quad X_N^\pi \leq f_N, \mathbb{P} \text{ q.c.}\} \\ &= \{\pi \, \textit{ammissibili}: \, X_0^\pi = x, \quad \mathbb{P}(X_N^\pi \leq f_N) = 1\} \end{aligned}$$

la ***classe delle coperture inferiori per (x, f_N)***.

[9]Esempio tipico, nel caso $d = 1$, è il caso della opzione call con prezzo di esercizio K, per cui $f_N = (S_N - K)^+$. Altri esempi sono dati dal prezzo massimo dell'azione $f_N = \max\{S_n; n = 1, \ldots, N\}$, o ancora dal suo prezzo medio $f_N = (S_1 + \cdots + S_N)/N$.

Definizione 5.7 (Prezzo superiore e prezzo inferiore). *Sia f_N una funzione di pay-off, allora la quantità*

$$\begin{aligned} C^* = C^*(f_N;\mathbb{P}) &= \inf\{x \geq 0 :\ H^*(x, f_N;\mathbb{P}) \neq \emptyset\} \\ &= \inf\{x \geq 0 : \exists\, \pi\, \textit{ammissibile} :\ X_0^\pi = x, \quad \mathbb{P}(X_N^\pi \geq f_N) = 1\} \end{aligned}$$

è detta **prezzo superiore** *(di copertura per f_N) e*

$$\begin{aligned} C_* = C_*(f_N;\mathbb{P}) &= \sup\{x \geq 0 :\ H_*(x, f_N;\mathbb{P}) \neq \emptyset\} \\ &= \sup\{x \geq 0 :: \exists\, \pi\, \textit{ammissibile} :\ X_0^\pi = x, \quad \mathbb{P}(X_N^\pi \leq f_N) = 1\} \end{aligned}$$

è detta **prezzo inferiore** *(di copertura per f_N).*

N.B. Per convenzione si pone $C^*(f_N;\mathbb{P}) = \infty$ se $H^*(x, f_N;\mathbb{P}) = \emptyset$ per tutti gli $x \geq 0$ e $C_*(f_N;\mathbb{P}) = \infty$ se $H_*(x, f_N;\mathbb{P}) = \emptyset$ per tutti gli $x \geq 0$.

Osservazione 5.2. *Le definizioni date precedentemente assumono, implicitamente, che le strategie π considerate seguono le costrizioni imposte sulla dinamica del valore del portfolio; ad esempio se, come si fa in questo capitolo, ci si limita alle strategie autofinanzianti, allora più esplicitamente*

$$\begin{aligned} H^*(x, f_N;\mathbb{P}) &= \{\pi \in SF :\ X_0^\pi = x, \quad X_N^\pi \geq f_N \quad (\mathbb{P}\ q.c.)\}, \\ H_*(x, f_N;\mathbb{P}) &= \{\pi \in SF :\ X_0^\pi = x, \quad X_N^\pi \leq f_N \quad (\mathbb{P}\ q.c.)\}. \end{aligned}$$

Inoltre va osservato che nel mercato attualizzato $(\widetilde{B}, \widetilde{S})$, si ha

$$H^*(x, f_N;\mathbb{P}) = \widetilde{H}^*(\widetilde{x}, \widetilde{f}_N;\mathbb{P}), \quad e \quad H_*(x, f_N;\mathbb{P}) = \widetilde{H}_*(\widetilde{x}, \widetilde{f}_N;\mathbb{P}),$$

dove, $\widetilde{x} = x/B_0$ e

$$\begin{aligned} \widetilde{H}^*(\widetilde{x}, \widetilde{f}_N;\mathbb{P}) &:= \{\pi \in SF :\ \widetilde{X}_0^\pi = \widetilde{x}, \quad \widetilde{X}_N^\pi \geq \widetilde{f}_N \quad (\mathbb{P}\ q.c.)\}, \\ \widetilde{H}_*(\widetilde{x}, \widetilde{f}_N;\mathbb{P}) &:= \{\pi \in SF :\ \widetilde{X}_0^\pi = \widetilde{x}, \quad \widetilde{X}_N^\pi \leq \widetilde{f}_N \quad (\mathbb{P}\ q.c.)\}. \end{aligned}$$

Dopo aver visto formalmente i concetti di copertura, risulta interessante darne una visione pratica, cioè spiegare il loro utilizzo in ambito finanziario e il loro legame con la condizione di assenza di opportunità di arbitraggio.

Va osservato che nel seguito si dovrebbe considerare un mercato scontato $(\widetilde{B}, \widetilde{S})$ e il valore scontato $\widetilde{X}^\pi$ relativo ad una strategia π. Tuttavia, per semplicità di notazione tale mercato viene denotato come (B, S) e il valore con X^π, che è come dire che si suppone direttamente che sia $B_n = 1$, ciò permette di considerare i guadagni ad esse relativi nell'istante iniziale senza dover attualizzare i valori.

Si immagini di vendere a un prezzo x un generico contratto al quale corrisponde, all'istante N, il pagamento di un pay-off f_N. Naturalmente lo scopo è venderlo a un prezzo elevato, però si deve tener conto che il compratore lo vuole acquistare a un prezzo basso, ma vuole anche essere certo che il venditore sia in grado di fornire al tempo N l'obbligazione promessa f_N, cioè che non ci sia rischio di insolvenza (o default risk). Considerando queste posizioni opposte si pone il problema di determinare il ***più piccolo prezzo accettabile***, ovvero quel prezzo che permette al venditore di rispettare i termini del contratto (cioè di pagare l'ammontare f_N al tempo N) senza però dare al venditore un'opportunità di ottenere un arbitraggio altrimenti il compratore non avrebbe ragione di accettare la vendita.

D'altra parte acquistando un contratto lo si vuole comprare a buon mercato, ma non si può pretendere di ottenere un guadagno sicuro, senza rischi (ovvero un'opportunità di arbitraggio per il compratore), altrimenti il venditore non avrebbe interesse a cederlo.

Da quanto richiesto si ottiene, come si spiegherà tra poco, che i prezzi superiore e inferiore, $C^* = C^*(f_N;\mathbb{P})$ e $C_* = C_*(f_N;\mathbb{P})$, definiti precedentemente, soddisfano la seguente proprietà:

Gli intervalli $[0, C_*)$ ***e*** (C^*, ∞) ***sono insiemi dei valori dei prezzi che danno, rispettivamente, al compratore e al venditore opportunità di arbitraggio.***

Verifica della proprietà (C^*, ∞) *è un intervallo dei prezzi con opportunità di arbitraggio per il venditore:*

Si assuma che il contratto venga pagato un prezzo x maggiore di C^* allora il venditore può ottenere un free-lunch agendo in questo modo: egli preleva dalla somma totale x una quota y in modo tale che

$$C^* < y < x \text{ e } H^*(y, f_N; \mathbb{P}) \neq \emptyset$$

(un tale y esiste per la definizione di estremo inferiore) e la usa per costruire un portfolio $\pi^*(y)$ che assume i seguenti valori

$$X_0^{\pi^*(y)} = y \text{ e } \mathbb{P}(X_N^{\pi^*(y)} \geq f_N) = 1$$

(un tale portfolio esiste perché $H^*(y, f_N; \mathbb{P}) \neq \emptyset$).

Si può descrivere la stessa azione dicendo che il venditore *investe* la quota y nel mercato (B, S) in accordo con la strategia $\pi^*(y) = (\beta_n^*(y), \gamma_n^*(y))_{0 \leq n \leq N}$.

Il valore[10] di $\pi^*(y)$ al tempo N è $X_N^{\pi^*(y)}$, il guadagno totale delle due transazioni (vendere il contratto e comprare il portfolio $\pi^*(y)$) risulta[11] quindi

$$(x - f_N) + (X_N^{\pi^*(y)} - y) = (x - y) + (X_N^{\pi^*(y)} - f_N) \geq x - y > 0.$$

Allora, c'è un *guadagno netto senza rischio* (arbitraggio) del venditore di almeno $x - y > 0$.

Verifica della proprietà $[0, C_*)$ *è un intervallo dei prezzi con opportunità di arbitraggio per il compratore:*

Si consideri ora l'opportunità di arbitraggio per il compratore. Se il compratore acquista, per un prezzo $x < C_*$, un contratto che preveda all'istante N il pagamento di un ammontare f_N, allora egli può ottenere un free-lunch, come segue: prima di tutto il compratore sceglie y tale che

$$x < y < C_* \text{ e } H_*(y, f_N; \mathbb{P}) \neq \emptyset$$

(un tale y esiste per definizione di estremo superiore). Sia ora $\pi_*(y)$ un portfolio che appartiene a $H_*(y, f_N; \mathbb{P})$, ovvero il cui valore iniziale è y e il cui valore al tempo N risulti $X_N^{\pi_*(y)} \leq f_N$, in formule

$$X_0^{\pi_*(y)} = y \text{ e } \mathbb{P}(X_N^{\pi_*(y)} \leq f_N) = 1.$$

Il compratore agisce come segue: al tempo $n = 0$ investe la quota $-y$ in accordo con la strategia $\overline{\pi}(-y) = -\pi_*(y)$, dove

$$\pi_*(y) = (\beta_{*n}(y), \gamma_{*n}(y))_{0 \leq n \leq N}.$$

Allora

$$\overline{\pi}(-y) = (-\beta_{*n}(y), -\gamma_{*n}(y))_{0 \leq n \leq N},$$

in modo tale che

$$-y = -\beta_{*0}(y) B_0 - \gamma_{*0}(y) S_0$$

e il valore di $\overline{\pi}(-y)$ sia

$$X_N^{\overline{\pi}(-y)} = X_N^{-\pi_*(y)} = -X_N^{\pi_*(y)},$$

[10] Si ricordi che abbiamo assunto per semplicità che $B_n = 1$ per ogni $n = 0, \ldots, N$: altrimenti si dovrebbero cambiare le quantità $X_N^{\pi^*(y)}$ con $\widetilde{X}_N^{\pi^*(y)} = X_N^{\pi^*(y)}/B_N$, ed f_N con $\widetilde{f}_N = f_N/B_N$.

[11] Al tempo 0 il venditore riceve x dal compratore e al tempo N gli deve restituire f_N, inoltre tiene per sé il guadagno $X_N^{\pi^*(y)} - y$, ottenuto investendo y con la strategia $\pi^*(y)$.

si ottiene, quindi, che il guadagno totale dato dalle due transazioni (ossia comprare il contratto e investire $-y$, con la strategia $\overline{\pi}(-y)$)[12] è

$$(f_N - x) + (X_N^{\overline{\pi}(-y)} - (-y)) = (f_N - X_N^{\pi_*(y)}) + (y - x) \geq y - x > 0.$$

Si vede, infine, che essendo questo il guadagno netto strettamente positivo del compratore relativo all'acquisto di un contratto di prezzo $x < C^*$ implica un arbitraggio.

Osservazione 5.3. *Nella discussione precedente è stato considerato un investimento con un ammontare negativo $-y$, praticamente ciò significa che esiste la possibilità di trovare uno speculatore che mette a disposizione la quota y al tempo $n = 0$ al patto di ricevere $X_N^{\pi_*(y)}$ al tempo $n = N$; si noti che quest'ultimo valore può essere superiore o inferiore a y data l'aleatorietà dei prezzi.*

Dalla serie di considerazioni fatte si evince che l'unica possibilità per non avere opportunità di arbitraggio è porsi nell'ipotesi che $C_* \leq C^*$ e prendere come prezzo del contratto un valore $x \in [C_*, C^*]$; per questo motivo si dà a $[C_*, C^*]$ il nome di ***intervallo dei prezzi accettabili (o razionali)***.

In altre parole: se non ci sono opportunità di arbitraggio (ne' per il compratore, ne' per il venditore) deve necessariamente valere

$$C_* = C_*(f_N, \mathbb{P}) \leq C^*(f_N, \mathbb{P}) = C^* \tag{5.19}$$

La trattazione precedente dei prezzi razionali può essere schematizzata come segue

arbitraggio compratore *arbitraggio venditore*

0 C_* C^*

arbitrage-free

Si osservi, infine, che operare con prezzi appartenenti al range di quelli accettabili esclude soltanto il caso di profitti sicuri, ma non la possibilità di profitti, e che un profitto sicuro è una situazione logicamente e economicamente assurda anche perché un guadagno può essere visto come una *compensazione per il rischio.*

5.2.1 Mercato completo e incompleto

Si consideri ora il caso particolare in cui esista, per un fissato valore x e una funzione di pay-off f_N, una copertura perfetta π per (x, f_N), cioè una strategia tale che

$$X_0^\pi = x \text{ e } \mathbb{P}(X_N^\pi = f_N) = 1.$$

Assumere che $X_N^\pi = f_N$, $\mathbb{P}$ q.c., significa prendere una copertura π in grado di ***replicare*** la f_N richiesta.
Per molte ragioni risulta auspicabile che ogni obbligazione si possa replicare per qualche valore di $x = \overline{x}$; il motivo principale è che se questo accade l'intersezione tra le classi $H^*(x, f_N; \mathbb{P})$ e $H_*(x, f_N; \mathbb{P})$ è diversa dall'insieme vuoto in quanto contiene la strategia π che dà la replicabilità. Dalla definizione dei prezzi di copertura segue allora

$$C^*(f_N; \mathbb{P}) \leq \overline{x} \leq C_*(f_N; \mathbb{P})$$

ovvero, in questo caso, se inoltre non ci sono opportunità di arbitraggio e quindi deve valere (5.19), allora l'intervallo dei prezzi accettabili si riduce a un ***unico prezzo***

$$C(f_N; \mathbb{P}) = C^*(f_N; \mathbb{P}) = C_*(f_N; \mathbb{P}) = \overline{x},$$

[12] Per il contratto, al tempo zero deve pagare x e al tempo N riceve f_N, mentre per l'investimento al tempo 0 deve pagare $(-y)$ e al tempo N riceve $X_N^{\overline{\pi}(-y)} = -X_N^{\pi_*(y)}$.

detto ***prezzo razionale*** (o ***fair price***) per la richiesta f_N (è il prezzo che risulta accettabile sia per il venditore che per il compratore: ogni deviazione da questo comporterebbe la possibilità di ottenere guadagni senza rischi). Il caso trattato, vista la sua importanza, presenta una terminologia specifica.

Definizione 5.8 (Replicabilità). *Una obbligazione f_N si dice* **replicabile***, se esiste un x, per il quale esiste una copertura perfetta π di (x, f_N), ovvero esiste un portfolio π tale che*

$$X_0^\pi = x, \;\; e \; \mathbb{P}(X_N^\pi = f_N) = 1.$$

Definizione 5.9 (Mercato N-perfetto). *Un mercato (B, S) è detto* N-perfetto *o* perfetto rispetto il tempo N*, se* ogni *funzione f_N $\mathcal{F}_N$-misurabile può essere replicata, cioè, per un qualche x esiste una copertura perfetta π di (x, f_N), altrimenti il mercato viene detto* N-imperfetto*.*

La condizione che un mercato (B, S) sia perfetto è molto forte e impone delle costrizioni severe sulla struttura del mercato. Fortunatamente, non è necessario, in molti casi, che la copertura perfetta esista per tutte le funzioni f_N $\mathcal{F}_N$-misurabili a volte è sufficiente considerare solo funzioni limitate oppure funzioni con opportune condizioni di integrabilità o di misurabilità: è questo il caso di un mercato completo.

Definizione 5.10 (Mercato N-completo). *Un mercato (B, S) si dice* **N-completo** *o* **completo rispetto il tempo** *N, se ogni funzione f_N di pay-off limitata e $\mathcal{F}_N$-misurabile è replicabile.*

Osservazione 5.4. *Si noti che nel caso in cui $\mathcal{F}_N$ è una σ-algebra finita, allora tutte le variabili aleatorie $\mathcal{F}_N$-misurabili sono variabili aleatorie discrete e che assumono un numero finito di valori, e quindi tutte le variabili aleatorie $\mathcal{F}_N$-misurabili sono anche limitate. In questo caso non c'è differenza tra le due condizioni di mercato N-perfetto e mercato N-completo.*
Inoltre conviene osservare in questo caso, tutte le variabili aleatorie $\mathcal{F}_N$-misurabili sono anche ovviamente integrabili, essendo limitate. Quest'ultima proprietà sarà utile nel seguito per semplificare alcune dimostrazioni.

Determinare se e quando un mercato è perfetto o completo è uno dei punti di maggiore interesse per la matematica finanziaria rispondere a queste domande risulta estremamente difficile nel caso generale, ma, considerando delle ipotesi aggiuntive sulla struttura del mercato, si può arrivare a una soluzione esaustiva del problema. È questo il caso di un mercato arbitrage-free dove la completezza risulta strettamente connessa all'esistenza e all'unicità di una misura di probabilità: ***la misura martingala equivalente***, che viene trattata nelle sezioni successive e che è l'oggetto principale dei così detti teoremi dell'Asset Pricing *APT1* e *APT2* (vedere le sezioni successive).

5.3 Mercato senza opportunità di arbitraggio

In poche parole, come preannunciato, dire che un mercato è ***senza opportunità di arbitraggio*** (o ***arbitrage-free***) significa che il mercato è razionale: non si ottengono profitti senza rischiare. La Definizione 5.5 di arbitrage-free è piuttosto intuitiva la si vuole formalizzare illustrandone i legami con gli elementi finanziari introdotti.
Si fissi un $N \geq 1$ si è interessati al valore X_N^π di una strategia $\pi \in SF$ all'istante terminale.

Definizione 5.11 (Opportunità di arbitraggio). *Si dice che una strategia autofinanziante π permette un'***opportunità di arbitraggio** *(al tempo N) se, per un capitale iniziale*

$$X_0^\pi = 0$$

si ha

$$\mathbb{P}(X_N^\pi \geq 0) = 1$$

e $X_N^\pi > 0$ con probabilità $\mathbb{P}$ positiva, cioè

$$\mathbb{P}(X_N^\pi > 0) > 0$$

o, equivalentemente[13]

$$E(X_N^\pi) > 0.$$

[13] In generale una variabile aleatoria Z con $\mathbb{P}(Z \geq 0) = 1$ si ha:
se $\mathbb{P}(Z > 0) > 0$ allora $\mathbb{E}[Z] > 0$;
se $\mathbb{E}[Z] > 0$ allora necessariamente deve essere $\mathbb{P}(Z > 0) > 0$ (se fosse $\mathbb{P}(Z > 0) = 0$ allora sarebbe $\mathbb{P}(Z = 0) = 1$ e $\mathbb{E}[Z] = 0$)

Si indichi con SF_{arb} la classe delle strategie autofinanzianti con opportunità di arbitraggio. Se $\pi \in SF_{arb}$ e $X_0^\pi = 0$, allora

$$\mathbb{P}(X_N^\pi \geq 0) = 1 \quad \Longrightarrow \quad \mathbb{P}(X_N^\pi > 0) > 0.$$

Definizione 5.12 (Assenza di opportunità di arbitraggio). *Si dice che* non esiste opportunità di arbitraggio *su un mercato* (B, S) *o che il mercato è* arbitrage-free *se* $SF_{arb} = \emptyset$. *In altre parole, se il capitale iniziale* X_0^π *di una strategia* π *è zero e al tempo* N *il guadagno è non negativo* $\mathbb{P}$ *q.c., allora il guadagno è nullo* $\mathbb{P}$ *q.c., ovvero in formule, per ogni strategia autofinanziante*

$$X_0^\pi = 0 \ e \ \mathbb{P}(X_N^\pi \geq 0) = 1 \quad \Longrightarrow \quad \mathbb{P}(X_N^\pi = 0) = 1.$$

Si noti che nelle definizioni date sopra si considerano eventi del tipo $\{X_N^\pi > 0\}$, $\{X_N^\pi \geq 0\}$, o $\{X_N^\pi = 0\}$, che sono, rispettivamente, gli stessi[14] di $\{\widetilde{X}_N^\pi > 0\}$, $\{\widetilde{X}_N^\pi \geq 0\}$, o $\{\widetilde{X}_N^\pi = 0\}$ dove $\widetilde{X}_N^\pi = \frac{X_N^\pi}{B_N}$. Una considerazione analoga vale per gli eventi del tipo $\{X_N^\pi \geq f_N\}$, $\{X_N^\pi \leq f_N\}$ e $\{X_N^\pi = f_N\}$, che sono uguali agli eventi $\{\widetilde{X}_N^\pi \geq \widetilde{f}_N\}$, $\{\widetilde{X}_N^\pi \leq \widetilde{f}_N\}$ e $\{\widetilde{X}_N^\pi = \widetilde{f}_N\}$, come già osservato in precedenza. Ciò spiega perché la discussione relativa all'assenza o, alla presenza, di arbitraggio su un mercato (B, S), può essere ristretta al mercato $(\widetilde{B}, \widetilde{S})$ con $\widetilde{B}_n \equiv 1$ e $\widetilde{S}_n = \frac{S_n}{B_n}$. Nel seguito quindi si deve intendere che le ipotesi sono riferite al mercato scontato. Tuttavia, per semplificare le notazioni (sempre nell'ipotesi che valga $B_n > 0$ per ogni $n = 0, \ldots, N$) le ipotesi saranno scritte per il mercato (B, S)[15].

5.4 Primo e Secondo Teorema Fondamentale

In questo paragrafo si vogliono illustrare i legami esistenti tra le proprietà di assenza di arbitraggio e di completezza in un mercato finanziario e la nozione di misura martingala[16] equivalente. Si consideri ora un mercato (B, S) strutturato come descritto nel primo paragrafo, in quest'ambito viene data la seguente definizione

Definizione 5.13. *Una misura di probabilità* $\widetilde{\mathbb{P}}$ *è una* **misura martingala equivalente** *alla misura* $\mathbb{P}$ *se la successione d-dimensionale scontata*

$$\widetilde{S} = \left(\widetilde{S}_n := \frac{S_n}{B_n} \right)_{n \geq 0}$$

è una $(\widetilde{\mathbb{P}}, \mathcal{F}_n)$*-martingala, cioè, esplicitamente, presa* $\widetilde{\mathbb{E}}$ *come la media rispetto alla* $\widetilde{\mathbb{P}}$ *si ha che per ogni* $i = 1, \ldots, d$

$$\widetilde{\mathbb{E}} \left| \widetilde{S}_n^i \right| = \widetilde{\mathbb{E}} \left| \frac{S_n^i}{B_n} \right| < \infty \tag{5.20}$$

con $n = 0, 1, \ldots, N$ *e*

$$\widetilde{\mathbb{E}} \left(\widetilde{S}_n^i \middle| \mathcal{F}_{n-1} \right) = \widetilde{S}_{n-1}^i \quad ovvero \ \widetilde{\mathbb{E}} \left(\frac{S_n^i}{B_n} \middle| \mathcal{F}_{n-1} \right) = \frac{S_{n-1}^i}{B_{n-1}} \qquad (\widetilde{\mathbb{P}} \ q.c.) \tag{5.21}$$

per $n = 1, \ldots, N$.

Introdotti questi elementi si può enunciare il seguente teorema che data la sua importanza è chiamato ***Primo Teorema fondamentale dell'Asset Pricing*** (o ***First Fundamental Asset Pricing Theorem***), e verrà richiamato nel seguito come ***APT1***.

[14] Si ricordi che si ipotizza che $B_N > 0$, e quindi $\{\omega \in \Omega : X_N^\pi(\omega) > 0\} = \{\omega \in \Omega : \frac{X_N^\pi(\omega)}{B_N(\omega)} > 0\}$

[15] Scrivere le assunzioni per il mercato (B, S), invece che per il mercato $(\widetilde{B}, \widetilde{S})$, si può esprimere anche dicendo che si può assumere senza ledere la generalità che $B_n \equiv 1$, per ogni $n = 0, \ldots, N$.

[16] Ricordiamo la definizione di martingala a tempo discreto per comodità del lettore Un processo $X = (X_n)_{n \geq 0}$ definito su uno spazio di probabilità filtrato $(\Omega, \mathcal{F}, (\mathcal{F}_n), \mathbb{P})$ è una *martingala* rispetto $(\mathcal{F}_n)$ sotto $\mathbb{P}$, $(\mathbb{P}, \mathcal{F}_n)$*-martingala*, se:

0) X_n è $\mathcal{F}_n$-misurabile $\forall\, n$ $\qquad$ (X_n è $\mathcal{F}_n$-adattato);

1) X_n è integrabile $\qquad$ ($\mathbb{E}|X_n| < \infty \quad \forall\, n$);

2) $\mathbb{E}(X_n | \mathcal{F}_{n-1}) = X_{n-1}$ $\qquad$ (condizione martingala).

Teorema 5.1 (APT1). *Un mercato finanziario (B, S) con $N < \infty$ e $d < \infty$ definito su uno spazio di probabilità filtrato in cui $\mathcal{F}_0 = \{\emptyset, \Omega\}$ e $\mathcal{F}_N = \mathcal{F}$ è arbitrage-free se e solo se esiste almeno una misura martingala $\widetilde{\mathbb{P}}$ equivalente a $\mathbb{P}$.*

Si è già detto che l'assunzione di assenza di arbitraggio ha un chiaro significato economico: questa proprietà rende il mercato razionale; è interessante osservare come tale caratteristica puramente finanziaria sia in realtà strettamente connessa a modelli probabilistici quali risultano essere le martingale, considerando ciò appare chiara l'importanza del teorema e il perché dell'attributo fondamentale!
Pur avendo elogiato il valore del teorema ***APT1*** bisogna sottolineare che le ipotesi su cui poggia risultano abbastanza restrittive, infatti si sta immaginando di operare su un mercato con orizzonte finito ($N < \infty$) e numero di azioni limitato ($d < \infty$). Queste assunzioni pur limitando notevolmente l'applicazione del teorema risultano, purtroppo, indispensabili: si possono, infatti, fornire dei ***controesempi*** che mostrano la non validità del teorema nel caso di $d = \infty$ o $N = \infty$.

Esempio 5.1. *Questo esempio, dovuto a W. Schachermayer [16], mostra che se $d = \infty$ (e $N = 1$) allora esiste un mercato arbitrage-free senza misura martingala equivalente, così che la parte necessaria del teorema risulta non verificata per un numero infinito di azioni.*
Sia $\Omega = \{1, 2, \ldots\}$, sia $\mathcal{F}_0 = \{\emptyset, \Omega\}$, sia $\mathcal{F} = \mathcal{F}_1$ la σ-algebra generata dall'insieme delle parti di Ω e sia $\mathbb{P} = \sum_{k=1}^{\infty} 2^{-k}\delta_k$, cioè, $\mathbb{P}\{k\} = 2^{-k}$.
Si definisca la successione dei prezzi $S = (S_n^i)$ per $i = 1, 2, \ldots$ e $n = 0, 1$ come segue:

$$S_0^i(\omega) = 1 \quad \forall\, i,\, \omega \qquad\qquad \textit{mentre} \qquad\qquad \Delta S_1^i(\omega)(= S_1^i - S_0^i) := \begin{cases} 1, & \omega = i \\ -1, & \omega = i+1 \\ 0 & \textit{altrimenti} \end{cases}$$

e si consideri il mercato (B, S) corrispondente alla successione $S = (S_n^i)$, per $n = 0, 1$, e con $B_0 = B_1 = 1$.
Si inizi con il vedere che è **arbitrage-free**.

Il valore al tempo 0 è

$$X_0^\pi = c_0 + \sum_{i=1}^{\infty} c_i;$$

ovvero $c_0(= \beta_0)$ rappresenta la quantità di denaro investita nel conto in banca (o nel bond) e le $c_i(= \gamma_0^i)$ con $i = 1, 2, \ldots$ sono le quantità iniziali delle azioni (si assume che $\sum |c_i| < \infty$), che inizialmente valgono $S_0^i = 1$.

Il valore $X_1^\pi(\omega)$ del generico portfolio π può essere allora rappresentato con la somma

$$X_1^\pi = c_0 + \sum_{i=1}^{\infty} c_i S_1^i == X_0^\pi + c_0 + \sum_{i=1}^{\infty} c_i S_1^i - X_0^\pi = X_0^\pi + \sum_{i=1}^{\infty} c_i \Delta S_1^i.$$

Se $X_0^\pi = 0$ (cioè $c_0 + \sum_{i=1}^{\infty} c_i = 0$) allora dalla condizione $\mathbb{P}(X_1^\pi \geq 0) = 1$ si ottiene[17]

$$***X_1^\pi(1) = c_1 - c_0*** \geq 0;\ X_1^\pi(2) = c_2 - c_1 \geq 0;\ \ldots\ ; X_1^\pi(k) = c_k - c_{k-1} \geq 0;\ \ldots$$

ovvero si ottiene che la $\{c_i\}$ è una successione crescente in senso lato. D'altra parte la successione $\{c_i\}$ è assolutamente convergente, e quindi si devono avere tutte le c_i uguali a zero. Da ciò segue che $X_1^\pi = 0\,(\mathbb{P}\,q.c.)$ e, quindi, ricordando la Definizione 5.12 si ottiene l'assenza di opportunità di arbitraggio.

Tuttavia **non esiste nessuna misura martingala equivalente***: se esistesse una misura $\widetilde{\mathbb{P}} \sim \mathbb{P}$ tale che S è una martingala rispetto a $\widetilde{\mathbb{P}}$ si dovrebbe avere (si veda la (5.21)) che per ogni $i = 1, 2, \ldots$*

$$\widetilde{\mathbb{E}}[\Delta S_1^i | \mathcal{F}_0] = \widetilde{\mathbb{E}}[\Delta S_1^i] = 0.$$

[17]Se $\omega = 1$ allora $\Delta S_1^1(\omega)\Delta S_1^1(1) = 1$, mentre $\Delta S_1^i(1) = 0$ per ogni altro valore di $i > 1$, da cui $X_1^\pi(1) = c_1$.
Se $\omega = k \geq 2$, allora $\Delta S_k^1(\omega) = \Delta S_k^1(k) = 1$ e $\Delta S_{k-1}^1(\omega) = \Delta S_{k-1}^1(k) = -1$, mentre $\Delta S_1^i(\omega) = \Delta S_1^i(k) = 0$ per ogni altro valore di i, da cui $X_1^\pi(k) = c_k - c_{k-1}$.

Esplicitando tale condizione si ottiene

$$\begin{aligned}\sum_{\omega\in\Omega} \Delta S_1^i(\omega)\widetilde{\mathbb{P}}(\omega) &= \sum_{\omega=1}^{\infty} \Delta S_1^i(\omega)\widetilde{\mathbb{P}}(\omega)\\ &= \Delta S_1^i(i)\widetilde{\mathbb{P}}(i) + \Delta S_1^i(i+1)\widetilde{\mathbb{P}}(i+1) = \widetilde{\mathbb{P}}(i) - \widetilde{\mathbb{P}}(i+1) = 0,\end{aligned}$$

cioè $\widetilde{\mathbb{P}}(i) = \widetilde{\mathbb{P}}(i+1)$, *per* $i = 1, 2, \ldots$; *ciò è chiaramente impossibile per una misura di probabilità* σ*-additiva (si avrebbe, infatti,* $\sum_{\omega=1}^{\infty} \widetilde{\mathbb{P}}(\omega) = \infty$ *oppure* $= 0$*).*

Esempio 5.2. *Questo controesempio, che è il classico* **Paradosso di S. Pietroburgo***, mostra che nel caso* $N = \infty$ *l'esistenza di una misura martingala non assicura l'assenza di opportunità di arbitraggio, ovvero la parte sufficiente del teorema risulta non verificata nel caso di orizzonte infinito.*
Sia $\xi = (\xi_n)_{n\geq 0}$ *una successione di variabili aleatorie i.i.d. su uno spazio di probabilità* $(\Omega, \mathcal{F}, \mathbb{P})$ *tali che* $\mathbb{P}(\xi_n = 1) = \mathbb{P}(\xi_n = -1) = \frac{1}{2}$. *Se si considera il mercato* (B, S) *tale che* $S_0 = 0$, $S_n = \xi_1 + \ldots + \xi_n$ *e* $B_n \equiv 1$, *si vede subito che* $\mathbb{P}$ **è essa stessa una misura martingala equivalente** *in quanto* $\widetilde{S}_n = S_n$ *è una martingala (confrontare l'Esempio 4.2) Tuttavia se si considera la strategia* π *autofinanziante, con*

$$\gamma_k = \begin{cases} 2^{k-1} & se \quad \xi_1 = \ldots = \xi_{k-1} = -1,\\ 0 & altrimenti,\end{cases}$$

e con capitale iniziale X_0^π *nullo, si ottiene che il guadagno è*

$$X_n^\pi = \sum_{1\leq k\leq n} \gamma_k \Delta S_k = \sum_{1\leq k\leq n} \gamma_k \xi_k.$$

Questa strategia permette arbitraggio*, anzi per la precisione, questa strategia assicura una vincita sicura di* 1. *Prima di dimostrare questa proprietà, che permette di affermare che il Paradosso di S. Pietroburgo fornisce un controesempio al Teorema* **APT1***, risulta interessante osservare che* X_n^π *può essere visto come il guadagno associato a un giocatore d'azzardo che sta giocando contro un avversario di pari livello (da qui la simmetria), che l'esito del gioco è descritto dalle variabili aleatorie* ξ_k *(egli vince se* $\xi_k = 1$ *e perde se* $\xi_k = -1$*) e infine che la strategia del giocatore consiste nel raddoppiare la puntata dopo una perdita e smettere di giocare dopo la prima vincita.*
Chiaramente se $\xi_1 = \ldots = \xi_k = -1$ *il giocatore è in netta perdita e il guadagno in questo caso assume il valore*

$$X_k^\pi = -\sum_{i=1}^{k} 2^{i-1} = -(2^k - 1),$$

tuttavia se all'istante successivo $k+1$ *si ha una vincita, cioè se* $\xi_{k+1} = 1$, *il guadagno diventa*

$$X_{k+1}^\pi = X_k^\pi + 2^k = -(2^k - 1) + 2^k = 1;$$

quindi la strategia scelta ammette un tempo di arresto (aleatorio) τ, *cui è associato un guadagno positivo.*
Si può definire τ *come*

$$\tau = \inf\{k : \ X_k^\pi = 1\}$$

e poiché $\mathbb{P}(\tau = k) = \left(\frac{1}{2}\right)^k$, *risulta* $\mathbb{P}(\tau < \infty) = 1$, *e quindi essendo* $\mathbb{P}(X_\tau^\pi = 1) = 1$ *si ha*

$$X_N^\pi = X_\infty^\pi = \sum_{k\geq 1} \gamma_k \xi_k = \sum_{1\leq k\leq \tau} \gamma_k \xi_k = X_\tau^\pi = 1$$

(e quindi anche $\mathbb{E}X_\tau^\pi = 1$*) sebbene il capitale iniziale* X_0^π *sia nullo.*
Allora nel mercato (B, S) *considerato esiste un'opportunità di arbitraggio: esiste un portfolio* π *tale che* $X_0^\pi = 0$ *e* $\mathbb{E}X_\tau^\pi = 1 > 0$ *per un qualche istante* τ.

N.B. *Ipotizzare una strategia che raddoppi la posta dopo ogni perdita significa considerare, implicitamente, che egli possa prendere dei prestiti senza alcun limite (oppure che il giocatore sia immensamente ricco)*[18]*; la possibilità di prendere in prestito una cifra di qualunque entità (così come l'eventualità che il giocatore sia immensamente ricco) risulta altamente improbabile,* nel mondo reale. *Per questo motivo si capisce che tra le considerazioni da fare, in relazione a un mercato, vi è anche quella di considerare, fra le strategie ammissibili, solo quelle economicamente ragionevoli.*

Tornando al caso generale ***si denoti con $\mathcal{P}(\mathbb{P})$ l'insieme delle misure martingale equivalenti a $\mathbb{P}$

$$\mathcal{P}(\mathbb{P}) = \{\widetilde{\mathbb{P}} \sim \mathbb{P} : \; \widetilde{S} \text{ è una } (\widetilde{\mathbb{P}}, \mathcal{F}_n) - \text{martingala}\},$$

***ossia l'insieme di tutte le misure martingala $\widetilde{\mathbb{P}}$ equivalenti a $\mathbb{P}$ e per le quali la successione dei prezzi scontati $\widetilde{S} = \left(\widetilde{S}_n = \frac{S_n}{B_n}\right)_{n \geq 0}$ è una $(\widetilde{\mathbb{P}}, \mathcal{F}_n)$-martingala.

***A questo proposito va ricordato che

Definizione 5.14 (misure equivalenti). *Due misure μ e ν sullo spazio misurabile $(\Omega, \mathcal{F})$ sono equivalenti se e solo se*

$$\{A \in \mathcal{F} : \; \mu(A) = 0\} = \{A \in \mathcal{F} : \; \nu(A) = 0\}.$$

Di conseguenza, una misura di probabilità $\widetilde{\mathbb{P}}$ è equivalente alla misura di probabilità $\mathbb{P}$ su $(\Omega, \mathcal{F})$ se e solo se

$$\{A \in \mathcal{F} : \; \mathbb{P}(A) = 0\} = \{A \in \mathcal{F} : \; \widetilde{\mathbb{P}}(A) = 0\}.$$

Trattandosi di misure di probabilità la precedente condizione equivale a

$$\{B \in \mathcal{F} : \; \mathbb{P}(B) = 1\} = \{B \in \mathcal{F} : \; \widetilde{\mathbb{P}}(B) = 1\},$$

come si vede subito passando ai complementari.***

Definito l'insieme $\mathcal{P}(\mathbb{P})$ si può introdurre il seguente risultato che per la sua importanza viene detto ***Secondo teorema fondamentale dell'Asset Pricing*** (o ***Second Fundamental Asset Pricing Theorem***) e denotato in seguito con ***APT2***.

Teorema 5.2 (APT2). *Un mercato finanziario (B, S) arbitrage-free con $N < \infty$ e $d < \infty$ definito su uno spazio di probabilità filtrato in cui $\mathcal{F}_0 = \{\emptyset, \Omega\}$ e $\mathcal{F}_N = \mathcal{F}$ è completo se e solo se l'insieme $\mathcal{P}(\mathbb{P})$ delle misure martingala equivalenti contiene un singolo elemento.*

Allora si può osservare che mentre l'assenza di opportunità di arbitraggio implica

$$\mathcal{P}(\mathbb{P}) \neq \emptyset$$

la completezza di un mercato arbitrage-free può essere scritta come

$$|\mathcal{P}(\mathbb{P})| = 1.$$

[18]***Il prestito (o la perdita, prima della vincita) vale , in valore assoluto $L_\tau = \sum_{k=0}^{\tau-1} 2^k = 2^\tau - 1$. Il suo valore atteso vale dunque

$$\mathbb{E}[L_\tau] = \sum_{n=1}^{\infty} (2^n - 1)\, \mathbb{P}(\tau = n) = \sum_{n=1}^{\infty} (2^n - 1)\, \frac{1}{2^n} = \infty.***$$

5.4.1 Sufficienza del Teorema APT1

In questa sezione si dimostra che l'esistenza di una misura di probabilità $\widetilde{\mathbb{P}}$ equivalente a $\mathbb{P}$ tale che il valore

$$\widetilde{S} = \left(\widetilde{S}_n = \frac{S_n}{B_n}\right)_{0\leq n\leq N}$$

sia una $(\widetilde{\mathbb{P}}, \mathcal{F}_n)$-martingala implica l'assenza di opportunità di arbitraggio per il mercato (B, S).
Si ricorda che siamo nell'ipotesi $B_n > 0$ per $n \geq 0$ per cui si può assumere, senza ledere in generalità, che $B_n \equiv 1$, quindi, usando la formula (5.18) relativa al mercato scontato, si ha

$$X_n^\pi = X_0^\pi + G_n^\pi \qquad e \qquad G_n^\pi = \sum_{k=1}^{n} \gamma_k \Delta S_k \tag{5.22}$$

dove $S = (S_n)$ è una $(\widetilde{\mathbb{P}}, \mathcal{F}_n)$-martingala.
Per provare quanto richiesto si deve mostrare che, presa una qualunque strategia autofinanziante $\pi \in SF$ tale che $X_0^\pi = 0$ e $X_N^\pi \geq 0$, $\mathbb{P}$ q.c. o, equivalentemente, $\widetilde{\mathbb{P}}$ q.c., cioè

$$\mathbb{P}(X_N^\pi = G_N^\pi = \sum_{k=1}^{N} \gamma_k \Delta S_k \geq 0) = 1 \Leftrightarrow \widetilde{\mathbb{P}}(X_N^\pi = G_N^\pi = \sum_{k=1}^{N} \gamma_k \Delta S_k \geq 0) = 1 \tag{5.23}$$

si ha $X_N^\pi = G_N^\pi = 0$ $\mathbb{P}$ q.c. o, equivalentemente, $\widetilde{\mathbb{P}}$ q.c.. ovvero

$$\mathbb{P}(X_N^\pi = G_N^\pi = 0) = 1 \Leftrightarrow \widetilde{\mathbb{P}}(X_N^\pi = G_N^\pi = 0) = 1.$$

***Dimostrazione della sufficienza del teorema* APT1.**
Si comincia qui la dimostrazione nel caso in cui $\mathcal{F}_N = \mathcal{F}$ sia una famiglia finita, ciò evita tutti i problemi di integrabilità, in quanto in uno spazio di probabilità con un numero finito di eventi, tutte le variabili aleatorie sono limitate e quindi integrabili (ovvero ammettono valore atteso finito).

Dall'espressione del guadagno $(G_n^\pi)_{n=0}^N$ risulta chiaro che questa successione è un integrale stocastico a tempo discreto[19] e quindi (vedere l'Esempio 4.6) è una $(\widetilde{\mathbb{P}}, \mathcal{F}_n)$-martingala, purché siano soddisfatte le condizioni di integrabilità[20]. Inoltre, poiché si assume $X_0^\pi = 0$, si ha $X_n^\pi = G_n^\pi$, per ogni n, e quindi si ha che $(X_n^\pi)_n$ è una martingala. Quindi, ricordando che dalla condizione di martingala discende che la successione considerata ha media costante, si ottiene la seguente serie di implicazioni

$$X_0^\pi = 0 \quad \Longrightarrow \quad \widetilde{\mathbb{E}}[X_0^\pi] = 0 \quad \Longrightarrow \quad \widetilde{\mathbb{E}}[X_N^\pi] = 0$$

ed avendo assunto per ipotesi $\mathbb{P}(X_N^\pi \geq 0) = 1$ ovvero $\widetilde{\mathbb{P}}(X_N^\pi \geq 0) = 1$ si deve avere $X_N^\pi = 0 (= G_N^\pi)$ $\widetilde{\mathbb{P}}$ q.c. e quindi anche $\mathbb{P}$ q.c., ovvero la tesi.

In ciò che segue viene accennato a come si può ottenere la dimostrazione anche nel caso generale, ovvero senza l'ipotesi che $\mathcal{F}$ sia una famiglia finita.

Per poter arrivare a dimostrare che, sotto le ipotesi fatte, $G_N^\pi = 0$, occorre introdurre delle definizioni e dei teoremi su nuovi elementi probabilistici: le martingale locali.

Definizione 5.15. ‡ *Una successione stocastica $X = (X_n)$ è una $(\mathbb{P}, \mathcal{F}_n)$-**martingala locale** se esiste una successione di tempi di Markov $(\tau_k)_{k\geq 1}$ (cioè, di variabili aleatorie che soddisfano la condizione $\{\omega : \tau_k \leq n\} \in \mathcal{F}_n, \quad n \geq 1$) tali che $\mathbb{P}(\tau_k \leq \tau_{k+1}) = 1$, $\mathbb{P}(\tau_k \uparrow \infty,$ per $k \to \infty) = 1$ e per ogni k il processo arrestato al tempo τ_k*

$$X^{\tau_k} := (X_{\tau_k \wedge n})_{n\geq 0}$$

è una $(\mathbb{P}, \mathcal{F}_n)$-martingala.

[19] In alcuni testi viene denotato come una trasformazione di martingala rispetto a $\widetilde{\mathbb{P}}$
[20] Come si è già chiarito si sono fatte le ipotesi che garantiscano l'integrabilità di qualsiasi variabile aleatoria

Teorema 5.3. ‡ *Sia il processo* $(M_n)_{n\geq 0}$ *una* $(\mathbb{P}, \mathcal{F}_n)$*-martingala e* Y_n *un processo predicibile, allora la trasformazione di martingala* $X = Y \cdot M$*, cioè la successione stocastica definita come integrale stocastico a tempo discreto*

$$X_n = Y_0 M_0 + \sum_{k=1}^{n} Y_k (M_k - M_{k-1}) \tag{5.24}$$

è una martingala locale.

N.B. Per approfondimenti del concetto di integrale stocastico e dei legami tra questo e le martingale si rimanda alla lettura del libro di P. Baldi [1].

Lemma 5.4. ‡ *1) Se il processo* $(X_n)_{n\geq 0}$ *è una* $(\mathbb{P}, \mathcal{F}_n)$*-martingala locale tale che* $\mathbb{E}X_0 < \infty$ *e risulta verificata una fra le due condizioni seguenti*

$$\mathbb{E}X_n^- < \infty, \qquad n \geq 0$$
$$\mathbb{E}X_n^+ < \infty, \qquad n \geq 0$$

allora $X = (X_n)_{n\geq 0}$ *è una* $(\mathbb{P}, \mathcal{F}_n)$*-martingala.*
2) Sia $X = (X_n)_{0\leq n\leq N}$ *una* $(\mathbb{P}, \mathcal{F}_n)$*-martingala locale e si assuma che* $N < \infty$*,* $\mathbb{E}X_0 < \infty$ *e sia* $\mathbb{E}X_N^- < \infty$ *oppure* $\mathbb{E}X_N^+ < \infty$*. Allora le ipotesi del punto 1) sono verificate per ogni* $n \leq N$ *e* $X = (X_n)_{0\leq n\leq N}$ *è una* $(\mathbb{P}, \mathcal{F}_n)$*-martingala.*

Introdotte queste nozioni si può procedere nella dimostrazione della sufficienza del teorema *APT1*.
Dall'espressione del guadagno $(G_n^\pi)_{n=0}^N$ risulta chiaro che questa successione è una trasformazione di martingala rispetto a $\widetilde{\mathbb{P}}$ e quindi per Teorema 5.3 è una martingala locale. Inoltre, poiché si assume $G_0^\pi = 0$ e $\mathbb{P}(G_N^\pi \geq 0) = 1$, risultano verificate le ipotesi di integrabilità della parte 2) del Lemma 5.4 da cui (G_n^π) è una $(\widetilde{\mathbb{P}}, \mathcal{F}_n)$-martingala. Infine ricordando che dalla condizione di martingala discende che la successione considerata ha media costante si ottiene la seguente serie di implicazioni

$$G_0^\pi = 0 \quad \Longrightarrow \quad \widetilde{\mathbb{E}}G_0^\pi = 0 \quad \Longrightarrow \quad \widetilde{\mathbb{E}}G_N^\pi = 0$$

ed avendo assunto per ipotesi $\mathbb{P}(G_N^\pi \geq 0) = 1$ si deve avere $G_N^\pi = 0 = X_N^\pi$ ($\mathbb{P}$ q.c. e $\widetilde{\mathbb{P}}$ q.c.) ovvero quanto desiderato.

Osservazione 5.5. ‡ *Come illustrato, la sufficienza del teorema* APT1 *risulta banalmente dimostrata se* G_n^π *è una* $(\widetilde{\mathbb{P}}, \mathcal{F}_n)$*-martingala. Tale proprietà è strettamente connessa all'integrabilità della successione* $(\gamma_n S_n)$*; infatti, dall'ipotesi che* (S_n) *è una* $(\widetilde{\mathbb{P}}, \mathcal{F}_n)$*-martingala e data la* $\mathcal{F}_{n-1}$*-misurabilità delle* (γ_n)*, si ottiene la condizione di martingala (si veda il punto 2 della Definizione 4.2) per il guadagno, risulta cioè:*

$$\widetilde{\mathbb{E}}(G_n^\pi - G_{n-1}^\pi | \mathcal{F}_{n-1}) = \widetilde{\mathbb{E}}(\gamma_n (S_n - S_{n-1}) | \mathcal{F}_{n-1}) = \gamma_n \widetilde{\mathbb{E}}(S_n - S_{n-1} | \mathcal{F}_{n-1}) = 0.$$

Quindi per avere verificate tutte le condizioni di martingala serve l'integrabilità della successione $(\gamma_n S_n)$ *rispetto a* $\widetilde{\mathbb{P}}$*, questo è, ad esempio, il caso in cui le* (γ_n) *sono uniformemente limitate.*
Assumere la successione $(\gamma_n S_n)$ *integrabile, pur semplificando notevolmente la dimostrazione del teorema, lederebbe la sua generalità, è per tale motivo che si preferisce ricorrere al concetto di martingala locale che permette di apprezzare l'*efficacia *dell'*APT1.

5.4.2 Necessità del Teorema APT1: trasformazione condizionale di Esscher † ‡

Si deve dimostrare che l'assenza di opportunità di arbitraggio significa l'esistenza di una misura di probabilità $\widetilde{\mathbb{P}} \sim \mathbb{P}$ in $(\Omega, \mathcal{F})$ tale che la successione scontata $S = (S_n)_{0\leq n\leq N}$ è una $(\widetilde{\mathbb{P}}, \mathcal{F}_n)$-martingala.
Esistono diverse dimostrazioni rigorose di questo risultato le quali sfruttano, in un modo o in un altro, dei risultati relativi all'analisi funzionale. Nessuna di queste suggerisce, però, la costruzione esplicita delle misure martingala equivalenti: non compare mai l'esplicita descrizione di tutte le misure martingala $\widetilde{\mathbb{P}}$ equivalenti all'originale misura $\mathbb{P}$.
Per questo motivo si è cercata una dimostrazione alternativa che potesse portare a una costruzione della misura martingala equivalente: l'idea da utilizzare, come ha illustrato per primo L.C.G. Rogers nel 1994 [13], è quella di sfruttare la *trasformazione condizionale di Esscher*, si veda il Lemma 5.6, tale trasformazione è una

generalizzazione della (5.35).

Per spiegare l'idea base si considera inizialmente un modello a un singolo passo ($N = 1$) dove, per semplicità, si assume: $d = 1, \quad B_0 = B_1 = 1$ e $\mathcal{F}_0 = \{\emptyset, \Omega\}$. Si ipotizza inoltre che $\mathbb{P}(S_1 \neq S_0) > 0$ altrimenti si avrebbe un mercato banale in cui si può prendere come misura martingala equivalente la misura originale $\mathbb{P}$.
Considerato che ogni portfolio π è una coppia di numeri (β, γ), se $X_0^\pi = 0$ allora le sole coppie ammissibili sono quelle per cui si ha $\beta + \gamma S_0 = 0$.
Assumere che il mercato sia arbitrage-free significa dire che devono verificarsi (sotto l'ipotesi di mercato non banale) le condizioni :

$$\mathbb{P}(\Delta S_1 > 0) > 0 \qquad \text{e} \qquad \mathbb{P}(\Delta S_1 < 0) > 0. \tag{5.25}$$

Si vuole dedurre dalla (5.25) che esiste una misura $\widetilde{\mathbb{P}} \sim \mathbb{P}$ tale che

1) $\widetilde{\mathbb{E}}|\Delta S_1| < \infty$;

2) $\widetilde{\mathbb{E}}\Delta S_1 = 0$;

per poter giungere a questa conclusione si applicano i risultati ottenuti dal seguente lemma.

Lemma 5.5. *Sia X una variabile aleatoria reale con distribuzione di probabilità su $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ tale che*

$$\mathbb{P}(X > 0) > 0 \qquad e \qquad \mathbb{P}(X < 0) > 0. \tag{5.26}$$

Allora esiste una misura $\widetilde{\mathbb{P}} \sim \mathbb{P}$ tale che

$$\widetilde{\mathbb{E}}\, e^{aX} < \infty \tag{5.27}$$

per ogni $a \in \mathbb{R}$; in particolare,

$$\widetilde{\mathbb{E}}|X| < \infty. \tag{5.28}$$

Inoltre, $\widetilde{\mathbb{P}}$ ha la seguente proprietà:

$$\widetilde{\mathbb{E}}X = 0. \tag{5.29}$$

Dimostrazione. Data la misura $\mathbb{P}$, si costruisce a partire da questa la misura di probabilità equivalente

$$\mathbb{Q}(dx) = ce^{-x^2}\mathbb{P}(dx), \qquad x \in \mathbb{R},$$

dove c è la costante di normalizzazione, cioè

$$c^{-1} = \mathbb{E}e^{-X^2}.$$

Considerando la funzione

$$\varphi(a) = \mathbb{E}_{\mathbb{Q}}e^{aX} = \frac{\mathbb{E}(e^{aX}e^{-X^2})}{\mathbb{E}(e^{-X^2})}, \quad a \in \mathbb{R}, \tag{5.30}$$

dalla costruzione di $\mathbb{Q}$ segue che $0 < \varphi(a) < \infty$ per ogni $a \in \mathbb{R}$ (si verificherà ciò nelle Osservazione 5.6); quindi ponendo

$$Z_a(x) = \frac{e^{ax}}{\varphi(a)}, \tag{5.31}$$

si ha $Z_a(x) > 0$ e $\mathbb{E}_{\mathbb{Q}}Z_a(X) = 1$. Risulta possibile, dunque, definire per ogni $a \in \mathbb{R}$ la misura di probabilità

$$\widetilde{\mathbb{P}}_a(dx) = Z_a(x)\mathbb{Q}(dx) \tag{5.32}$$

tale che $\widetilde{\mathbb{P}}_a \sim \mathbb{Q} \sim \mathbb{P}$.
Dalla definizione di $\widetilde{\mathbb{P}}_a$ segue che, per ogni scelta di a, assumendo $\widetilde{\mathbb{P}} = \widetilde{\mathbb{P}}_a$ vale la (5.27), infatti

$$\mathbb{E}_{\widetilde{\mathbb{P}}_a}e^{aX} = \mathbb{E}_{\mathbb{Q}}\left(\frac{e^{2aX}}{\varphi(a)}\right) = \frac{\varphi(2a)}{\varphi(a)} < \infty. \tag{5.33}$$

Si può notare, inoltre, che la funzione $\varphi = \varphi(a)$ definita $\forall\, a \in \mathbb{R}$ è strettamente convessa poiché

$$\varphi^{''}(a) = \frac{\mathbb{E}(X^2 e^{aX} e^{-X^2})}{\mathbb{E}(e^{-X^2})} > 0;$$

per cui se si pone

$$\varphi_* = \inf\{\varphi(a) : \ a \in \mathbb{R}\}$$

vi sono solo due casi possibili

1) esiste a_* tale che $\varphi(a_*) = \varphi_*$;

2) non esiste a_* tale che $\varphi(a_*) = \varphi_*$.

Nel primo caso risulta, come ovvio, $\varphi^{'}(a_*) = 0$ da cui

$$\mathbb{E}_{\widetilde{\mathbb{P}}_{a^*}} X = \mathbb{E}_{\mathbb{Q}}\left(\frac{Xe^{a_* X}}{\varphi(a_*)}\right) = \frac{\varphi^{'}(a_*)}{\varphi(a_*)} = 0; \tag{5.34}$$

allora in virtù della (5.33) e della (5.34) si ha che prendendo $\widetilde{\mathbb{P}} = \widetilde{\mathbb{P}}_{a^*}$ il lemma è dimostrato.
Il secondo caso è irrealizzabile data l'ipotesi (5.26). Infatti, se un tale a_* non esiste, si può prendere una successione $\{a_n\}$ per cui valgano le relazioni

$$\varphi_* < \varphi(a_n) \qquad \text{e} \qquad \varphi(a_n) \searrow \varphi_*,$$

inoltre la successione $\{a_n\}$ deve tendere a $+\infty$ o a $-\infty$ poiché, altrimenti, si potrebbe scegliere una sottosuccessione convergente in modo tale che il valore di minimo venga assunto in un punto finito e ciò contraddice l'assunzione **2)**.
Sia $u_n = \frac{a_n}{|a_n|}$ e sia $u = \lim u_n$, ne segue $u = \pm 1$. Allora considerando l'ipotesi (5.26) e l'equivalenza tra le misure $\mathbb{P}$ e $\mathbb{Q}$ si ha

$$\mathbb{Q}(uX > 0) > 0,$$

per cui esiste $\delta > 0$ tale che

$$\mathbb{Q}(uX > \delta) > 0;$$

definendo $\varepsilon := \mathbb{Q}(uX > \delta)$ e scegliendo δ in modo tale che risulti un punto di continuità per $\mathbb{Q}$, cioè

$$\mathbb{Q}(uX = \delta) = 0$$

si ottiene

$$\mathbb{Q}(a_n X > |a_n|\delta) = \mathbb{Q}(u_n X > \delta) \to \varepsilon \qquad \text{per} \qquad n \to \infty.$$

Quindi per n sufficientemente grande

$$\begin{aligned} \varphi(a_n) &= \mathbb{E}_{\mathbb{Q}}(e^{a_n X}) \geq \mathbb{E}_{\mathbb{Q}}(e^{a_n X} \mathbb{I}_{(a_n X > \delta |a_n|)}) \\ &\geq e^{\delta |a_n|} \mathbb{E}_{\mathbb{Q}}(\mathbb{I}_{(a_n X > \delta |a_n|)}) = e^{\delta |a_n|} \mathbb{Q}(u_n X > \delta) \end{aligned}$$

da cui definitivamente

$$\varphi(a_n) \geq \frac{1}{2} \varepsilon e^{\delta |a_n|} \longrightarrow \infty.$$

Si è giunti a un assurdo poiché, come visto, $\varphi(a_n) \searrow \varphi_*$ e $\varphi_* < \infty$. $\square$

Si osservi che la dimostrazione di tale lemma porta a una costruzione della misura $\widetilde{\mathbb{P}}_a$ basata sulla trasformazione

$$x \mapsto \frac{e^{ax}}{\varphi(a)} \tag{5.35}$$

detta *trasformazione di Esscher*.

Osservazione 5.6. *Nella precedente dimostrazione si è tralasciato di verificare l'asserzione* $0 < \varphi(a) < \infty$ *per ogni* $a \in \mathbb{R}$*; la prima disuguaglianza è ovvia, la seconda è diretta conseguenza della semplice relazione algebrica:*

$$e^{-x^2}e^{ax} = e^{-\left(x-\frac{a}{2}\right)^2}e^{\frac{a^2}{4}},$$

infatti applicandola alla definizione di $\varphi(a)$ *si ottiene*

$$\varphi(a) = \frac{\mathbb{E}(e^{-X^2}e^{aX})}{\mathbb{E}(e^{-X^2})} = e^{\frac{a^2}{4}}\frac{\mathbb{E}e^{-\left(X-\frac{a}{2}\right)^2}}{\mathbb{E}(e^{-X^2})} \le \frac{e^{\frac{a^2}{4}}}{\mathbb{E}(e^{-X^2})} < +\infty.$$

Dalla limitatezza delle $\varphi(a) = \mathbb{E}_{\mathbb{Q}}(e^{aX})$ *seguono alcune considerazioni che sono state adoperate nel corso della dimostrazione:*

- $\varphi_* < \infty$ *(*φ_* *è l'estremo inferiore di funzioni limitate)*
- $\mathbb{E}_{\mathbb{Q}}(Xe^{aX}) = \mathbb{E}_{\mathbb{Q}}\left(\frac{d}{da}(e^{aX})\right) = \frac{d}{da}\mathbb{E}_{\mathbb{Q}}(e^{aX}) = \frac{d}{da}\varphi(a)$.

Dimostrato il lemma è facile vederne la generalizzazione al caso vettoriale quando, cioè, si considera al posto di X una successione di variabili aleatorie $(X_0, X_1, \dots, X_N)$ tali che X_n è $\mathcal{F}_n$-misurabile per $0 \le n \le N$ con $\mathcal{F}_0 = \{\emptyset, \Omega\}$ e $\mathcal{F}_N = \mathcal{F}$.
Per procedere nella trattazione del caso generale si deve tener presente il concetto di differenza di martingala.

Definizione 5.16. *Un processo* $X = (X_n)_{n\ge 1}$ *con* $\mathbb{E}|X_n| < \infty$ *è una* $(\mathbb{P}, \mathcal{F}_n)$-differenza di martingala *se*

$$\mathbb{E}(X_n|\mathcal{F}_{n-1}) = 0 \qquad (\mathbb{P}\ q.c.), \qquad\qquad n \ge 1.$$

Lemma 5.6. *Si ipotizzi che*

$$\mathbb{P}(X_n > 0|\mathcal{F}_{n-1}) > 0 \qquad e \qquad \mathbb{P}(X_n < 0|\mathcal{F}_{n-1}) > 0 \tag{5.36}$$

per $1 \le n \le N$*. Allora esiste una misura di probabilità* $\widetilde{\mathbb{P}} \sim \mathbb{P}$ *nello spazio* $(\Omega, \mathcal{F})$ *tale che la successione* $(X_0, X_1, \dots, X_N)$ *è una* $(\widetilde{\mathbb{P}}, \mathcal{F}_n)$*-differenza di martingala.*

Dimostrazione. La dimostrazione di questo lemma ripercorre pari passo quella del precedente. Data $\mathbb{P}$ si considera la misura di probabilità $\mathbb{Q}$ tale che

$$\mathbb{Q}(d\omega) = c\, exp\left\{-\sum_{i=0}^{N} X_i^2(\omega)\right\}\mathbb{P}(d\omega) \tag{5.37}$$

e $\mathbb{E}_{\mathbb{Q}} exp\left\{-\sum_{i=0}^{N} a_i X_i\right\}$ è finito. Per costruire la misura $\widetilde{\mathbb{P}}$ richiesta si considerano le funzioni

$$\varphi_n(a;\omega) = \mathbb{E}_{\mathbb{Q}}(e^{aX_n}|\mathcal{F}_{n-1})(\omega)$$

che fissato ω risultano strettamente convesse in a. Si può mostrare procedendo come fatto per il Lemma 5.5 che esiste un unico punto (finito) $a_n^*(\omega)$ tale che il più piccolo valore $\inf_a \varphi_n(a;\omega)$ viene assunto in questo punto, ovvero:

$$\inf_a \varphi_n(a;\omega) = \varphi_n(a_n^*(\omega);\omega).$$

La funzione $\inf_a \varphi_n(a;\omega)$ è $\mathcal{F}_{n-1}$-misurabile da ciò segue che anche $a_n^*(\omega)$ è $\mathcal{F}_{n-1}$-misurabile.
Ora si definisce ricorsivamente una successione $Z_0, Z_1(\omega), \dots, Z_N(\omega)$ ponendo

$$Z_0 = 1 \quad \text{e} \quad Z_n(\omega) = Z_{n-1}(\omega)\frac{exp\{a_n^*(\omega)X_n(\omega)\}}{\mathbb{E}_{\mathbb{Q}}(exp\{a_n^*\, X_n\}|\mathcal{F}_{n-1})(\omega)}$$

per $n \ge 1$. Chiaramente, le variabili $Z_n(\omega)$ sono $\mathcal{F}_n$-misurabili e formano una $(\mathbb{Q}, \mathcal{F}_n)$-martingala

$$\mathbb{E}_{\mathbb{Q}}(Z_n|\mathcal{F}_{n-1}) = Z_{n-1} \qquad (\mathbb{Q}\text{ q.c. e }\mathbb{P}\text{ q.c.}).$$

Quanto mostrato permette di definire la $\widetilde{\mathbb{P}}$ richiesta come:

$$\widetilde{\mathbb{P}}(d\omega) = Z_N(\omega)\mathbb{P}(d\omega),$$

ricordando la dimostrazione del Lemma 5.5 si vede che $\widetilde{\mathbb{E}}|X_n| < \infty$, $0 \le n \le N$,

$$\widetilde{\mathbb{E}}(X_n|\mathcal{F}_{n-1}) = 0 \qquad 1 \le n \le N$$

e inoltre $\widetilde{\mathbb{E}}X_0$=0. Quindi la successione $(X_0, X_1, \ldots, X_N)$ è una differenza di martingala rispetto la $\widetilde{\mathbb{P}}$ ciò prova il lemma. □

Si evince immediatamente che nel caso $d = 1$, la *necessità* dell'esistenza di una misura martingala $\widetilde{\mathbb{P}} \sim \mathbb{P}$ (in un mercato arbitrage-free) è conseguenza del Lemma 5.6: basta prendere come successione di variabili aleatorie $X_0 = S_0, X_1 = \Delta S_1, \ldots, X_N = \Delta S_N$.
Considerando un mercato privo di opportunità di arbitraggio si può assumere senza ledere la generalità che

$$\mathbb{P}(\Delta S_n > 0|\mathcal{F}_{n-1}) > 0 \qquad \text{e} \qquad \mathbb{P}(\Delta S_n < 0|\mathcal{F}_{n-1}) > 0 \tag{5.38}$$

per ogni $n = 1, \ldots, N$; infatti se esiste un istante n per cui $\mathbb{P}(\Delta S_n = 0) = 1$ questo può essere trascurato in quanto, preso un generico portfolio π autofinanziante, l'istante n non apporta contributi al valore X_N^π. Allora la (5.38) risulta vera per tutti gli $n \le N$ da cui si conclude, come annunciato, che applicando il Lemma 5.6 a $X_0 = S_0$ e $X_n = \Delta S_n$ per $1 \le n \le N$ esiste una misura martingala equivalente (si veda la Definizione 5.13).

Per il caso generale, $d \ge 1$, la dimostrazione della necessità è concettualmente la stessa del caso $d = 1$, si deve però tener conto della generalizzazione del Lemma 5.6 al caso di valori vettoriali.

Lemma 5.7. *Sia $(X_0, X_1, \ldots, X_N)$ una successione di d-vettori $\mathcal{F}_n$-misurabili*

$$X_n = \begin{pmatrix} X_n^1 \\ \cdot \\ \cdot \\ \cdot \\ X_n^d \end{pmatrix}, \quad 0 \le n \le N$$

definita su uno spazio di probabilità filtrato $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{0\le n\le N}, \mathbb{P})$ con $\mathcal{F}_0 = \{\emptyset, \Omega\}$ e $\mathcal{F}_N = \mathcal{F}$. Si assuma inoltre che

$$\gamma_n = \begin{pmatrix} \gamma_n^1 \\ \cdot \\ \cdot \\ \cdot \\ \gamma_n^d \end{pmatrix}$$

è un vettore di variabili vettoriali, diverse da zero, $\mathcal{F}_{n-1}$-misurabili avente componenti limitate ($|\gamma_n^i(\omega)| \le c < \infty, \quad \omega \in \Omega$) tale che

$$\mathbb{P}((\gamma_n, X_n) > 0|\mathcal{F}_{n-1}) > 0 \qquad e \qquad \mathbb{P}((\gamma_n, X_n) > 0|\mathcal{F}_{n-1}) > 0 \qquad (\mathbb{P}\ q.c.)$$

dove (γ_n, X_n) è il prodotto scalare.
Allora esiste una misura di probabilità $\widetilde{\mathbb{P}}$ equivalente a $\mathbb{P}$ su $(\Omega, \mathcal{F})$ tale che la successione $(X_0, X_1, \ldots, X_N)$ è una differenza di martingala d-dimensionale rispetto a $\widetilde{\mathbb{P}}$, cioè: $\widetilde{\mathbb{E}}|X_n| < \infty$, $\widetilde{\mathbb{E}}X_0 = 0$ e $\widetilde{\mathbb{E}}(X_n|\mathcal{F}_{n-1}) = 0 \quad 1 \le n \le N$.

Risulta interessante notare che la costruzione della misura martingala basata sulla trasformazione condizionale di Esscher dà solo *una* misura particolare, sebbene la classe di misure martingala equivalente all'originale possa contenerne altre. A titolo informativo si osservi che esiste un procedimento rigoroso basato sulle *trasformazioni di Girsanov* che può essere usato nella costruzione di una famiglia di misure $\widetilde{\mathbb{P}}$ (per un maggior approfondimento di questo argomento si veda il capitolo V dello Shiryaev [17]).

5.5 Completezza e S-rappresentabilità

In questo paragrafo si vuole illustrare l'equivalenza tra la completezza e la proprietà di S-rappresentabilità di martingale locali.

Definizione 5.17. *Sia $(\Omega, \mathcal{F}, (\mathcal{F}_n), \mathbb{Q})$ uno spazio di probabilità filtrato con*

$$\text{una } (\mathbb{Q}, \mathcal{F}_n)\text{-martingala} \quad d \text{ dimensionale } S = (S_n)$$

e

$$\text{una } (\mathbb{Q}, \mathcal{F}_n)\text{-martingala locale} \quad \text{unidimensionale } M = (M_n).$$

Allora diciamo che la $(\mathbb{Q}, \mathcal{F}_n)$-martingala locale M ammette una **S-rappresentazione** *rispetto a $\mathbb{Q}$, o una* **rappresentazione in termini della $(\mathbb{Q}, \mathcal{F}_n)$-martingala** *$S$, se esiste una successione predicibile $\gamma = (\gamma_n)$, dove $\gamma_n = (\gamma_n^1, \dots, \gamma_n^d)$, tale che*

$$M_n = M_0 + \sum_{k=1}^{n} \gamma_k \Delta S_k = M_0 + \sum_{k=1}^{n} \left(\sum_{j=1}^{d} \gamma_k^j [S_k^j - S_{k-1}^j] \right) \qquad (\mathbb{Q} \text{ q.c.}) \tag{5.39}$$

per ogni $n \geq 1$, cioè M è l'integrale stocastico di una successione predicibile γ_n, ovvero M è una trasformazione di martingala ottenuta dalla $(\mathbb{Q}, \mathcal{F}_n)$-martingala S attraverso l'integrazione di una successione predicibile γ_n.

Lemma 5.8. *Sia (B, S) un mercato arbitrage-free con orizzonte finito N e $B_n \equiv 1$ per $n \leq N$. Allora questo mercato è completo se e solo se esiste una misura $\widetilde{\mathbb{P}} \in \mathcal{P}(\mathbb{P})$ tale che ogni $(\widetilde{\mathbb{P}}, \mathcal{F}_n)$-martingala limitata $M = (M_n)$ (con $|M_n(\omega)| \leq C \quad n \leq N \quad \omega \in \Omega$) ammette una S-rappresentazione rispetto a $\widetilde{\mathbb{P}}$.*

Dimostrazione. ($\Longrightarrow$: ***la completezza implica la S-rappresentabilità***) Si assuma che il mercato (arbitrage-free) sia completo. Presa una misura arbitraria $\widetilde{\mathbb{P}}$ appartenente a $\mathcal{P}(\mathbb{P})$ si consideri una $(\widetilde{\mathbb{P}}, \mathcal{F}_n)$-martingala $M = (M_n)$ con $|M_n(\omega)| \leq C$ per $n \leq N$ e $\omega \in \Omega$. Si può scegliere come funzione di pay-off

$$f_N = M_N,$$

dall'ipotesi di completezza segue l'esistenza di un portfolio π autofinanziante e di un capitale iniziale x tale che

$$X_n^\pi = x + \sum_{k=1}^{n} \gamma_k \Delta S_k$$

e $X_N^\pi = f_N = M_N$ ($\mathbb{P}$ q.c. e $\widetilde{\mathbb{P}}$ q.c.), ovvero

$$\mathbb{P}(X_N^\pi = f_N = M_N) = 1 \Leftrightarrow \widetilde{\mathbb{P}}(X_N^\pi = f_N = M_N) = 1.$$

Dalla limitatezza dell'obbligazione $|f_N| = |M_N| \leq C$ segue [21] che $X^\pi = (X_n^\pi)_{n \leq N}$ è una $(\widetilde{\mathbb{P}}, \mathcal{F}_n)$-martingala. Le martingale X^π ed M presentano la stessa funzione terminale, il pay-off f_N, e quindi la condizione di martingala porta alle seguenti uguaglianze

$$X_n^\pi = \widetilde{\mathbb{E}}(X_N^\pi | \mathcal{F}_n) = \widetilde{\mathbb{E}}(M_N | \mathcal{F}_n) = M_n \qquad \widetilde{\mathbb{P}} \text{ q.c.}$$

e da ciò si può concludere che

$$\widetilde{\mathbb{P}}(X_n^\pi = M_n) = 1 \Leftrightarrow \mathbb{P}(X_n^\pi = M_n) = 1$$

per ogni $0 \leq n \leq N$. La $(\widetilde{\mathbb{P}}, \mathcal{F}_n)$-martingala M ammette, quindi, una S-rappresentazione.

($\Longleftarrow$: ***la S-rappresentabilità implica la completezza***) Si ipotizzi che esista una misura $\widetilde{\mathbb{P}} \in \mathcal{P}(\mathbb{P})$, tale che ogni $(\widetilde{\mathbb{P}}, \mathcal{F}_n)$-martingala limitata ammette un S-rappresentazione e sia $f_N = f_N(\omega)$ una funzione limitata

[21] Nel caso in cui $\mathcal{F}$ sia una famiglia finita, non ci sono problemi di integrabilità e risulta immediatamente che $(X_n^\pi)_{n \leq N}$ è una martingala.
Nel caso generale si ha che $\widetilde{\mathbb{E}} X_N^\pi < \infty$ ed essendo $X_0^\pi = x$ risultano verificate le ipotesi del Lemma 5.4, quindi si ottiene ugualmente che$(X_n^\pi)_{n \leq N}$ è una martingala.

$\mathcal{F}_N$-misurabile ($|f_N| \leq C < \infty \quad \mathbb{P}$ q.c.) si vuole ottenere che esiste una strategia π autofinanziante e un capitale iniziale x tale che $X_N^\pi = f_N$ ($\mathbb{P}$ q.c.).
Si consideri la $(\widetilde{\mathbb{P}}, \mathcal{F}_n)$-martingala

$$M = (M_n)_{n \leq N}, \quad \text{definita da} \quad M_n := \widetilde{\mathbb{E}}[f_N | \mathcal{F}_n],$$

essendo $|f_N| \leq C$ si ha che M è una martingala limitata e per ipotesi ammette una S-rappresentazione (5.39). Inoltre, ricordando che, nelle stesse ipotesi dei teoremi dell'Asset Pricing, si ha $\mathcal{F}_0 = \{\emptyset, \Omega\}$, risulta

$$M_0 = \widetilde{\mathbb{E}}[f_N | \mathcal{F}_0] = \widetilde{\mathbb{E}}[f_N].$$

Esiste dunque una successione di variabili aleatorie predicibili $\gamma_k^j,\ j = 1, \ldots, d,\ k \leq N$ tale che

$$M_n = M_0 + \sum_{k=1}^n \gamma_k \Delta S_k \left(= M_0 + \sum_{k=1}^n \sum_{j=1}^d \gamma_k^j [S_k^j - S_{k-1}^j] \right),$$

attraverso queste variabili si costruisce un portfolio $\pi^* = (\beta^*, \gamma^*)$ dove

$$\gamma_n^* = \gamma_n, \tag{5.40}$$
$$\beta_n^* = \beta_{n-1}^* - (\gamma_n - \gamma_{n-1}) S_{n-1}, \tag{5.41}$$

in quanto si cerca una strategia autofinanziante, e basta allora ricordare la condizione (5.8)[22].

Si osserva che le β_n^* risultano predicibili data la $\mathcal{F}_{n-1}$-misurabilità delle γ_n. Inoltre π^* è un portfolio autofinanziante di valore[23]

$$X_n^{\pi*} = \beta_n^* + \gamma_n S_n \left(= \beta_n^* + \sum_{j=1}^d \gamma_n^j S_n^j \right) = M_n.$$

In particolare si vede che per l'istante terminale N si ha $X_N^{\pi*} = M_N = f_N$ ($\widetilde{\mathbb{P}}$ e $\mathbb{P}$ q.c.) da cui segue la completezza del mercato (B, S).

Da quest'ultima uguaglianza si ottiene anche la definizione alternativa

$$\beta_n^* = M_n - \gamma_n S_n. \tag{5.42}$$

$\square$

N.B. Se non si assume che $B_n \equiv 1, n \leq N$, allora tutti i risultati ottenuti risultano, comunque, validi se si considera al posto della $(\widetilde{\mathbb{P}}, \mathcal{F}_n)$-martingala $S = (S_n)_{n \leq N}$ la $(\widetilde{\mathbb{P}}, \mathcal{F}_n)$-martingala $\widetilde{S} = \left(\widetilde{S}_n = \frac{S_n}{B_n}\right)_{n \leq N}$, e al posto

[22] Si noti che sarebbe necessario anche che $\beta_0^* + \gamma_0^* S_0 = M_0$, con β_0^* e γ_0^* v.a. $\mathcal{F}{-}1 = \mathcal{F}_0 = \{\emptyset, \Omega\}$-misurabili, ovvero numeri certi. Si noti inoltre che per $n = 1$ si ha $\beta_1^* = \beta_0^* - (\gamma_1^* - \gamma_0^*) S_0 = \beta_0^* + \gamma_0^* S_0 - \gamma_1^* S_0 = M_0 - \gamma_1^* S_0$ e che quest'ultima espressione permette di definire β_1 anche senza conoscere (β_0^*, γ_0^*).
In realtà, noto $(\gamma_n)_{n \leq N}$ si potrebbe definire $\beta_n := M_n - \gamma_n S_n$, andrebbe mostrato che β_n risulta una v.a. $\mathcal{F}_{n-1}$-misurabile, e infine si potrebbe procede cominciando dal "fondo": $\beta_N = M_N - \gamma_N S_N = f_N - \gamma_N S_N$ e poi, usando la (5.41) si ottiene $\beta_{n-1}^* = \beta_n^* + (\gamma_n - \gamma_{n-1}) S_{n-1}$.

[23] L'uguaglianza si vede subito considerando che, per la definizione di β_n^* si ha

$$\begin{aligned} M_n - M_{n-1} = \gamma_n (S_n - S_{n-1}) &= \gamma_n S_n - \gamma_n S_{n-1} + \gamma_{n-1} S_{n-1} - \gamma_{n-1} S_{n-1} + (\beta_n^* - \beta_{n-1}^*) - (\beta_n^* - \beta_{n-1}^*) \\ &= \gamma_n S_n - \gamma_{n-1} S_{n-1} + (\beta_n^* - \beta_{n-1}^*), \end{aligned}$$

da cui

$$\begin{aligned} M_n - M_0 &= \gamma_n S_n - \gamma_0 S_0 + (\beta_n^* - \beta_0^*) \\ &= \gamma_n S_n + \beta_n^* - (\gamma_0 S_0 + \beta_0^*) = \gamma_n S_n + \beta_n^* - M_0. \end{aligned}$$

Si noti infine che quest'ultima uguaglianza permette di dare la definizione alternativa di β_n^* come

$$\beta_n^* := M_n - \gamma_n S_n,$$

che è quella originariamente usata da Shiryaev in [17].

di f_N si considera il valore attualizzato $\widetilde{f}_N = \frac{f_N}{B_N}$. In particolare si noti che allora serve la $\widetilde{S}$-rappresentabilità, e che allora

$$M_0 = \widetilde{\mathbb{E}}(\widetilde{f}_N|\mathcal{F}_0) = \widetilde{\mathbb{E}}(\frac{f_N}{B_N}),$$

essendo $\mathcal{F}_0$ la σ-algebra banale.

Allora $M_0 = \frac{x_0}{B_0}$, dove x_0 si può interpretare come il capitale iniziale che permette, attraverso la strategia $\pi*$, descritta nella dimostrazione della parte sufficiente del lemma, di ottenere una copertura perfetta per f_N.

RIASSUMENDO - per ottenere quindi il prezzo si può procedere nel seguente modo:

1 si mostra l'esistenza di una misura martingala equivalente $\widetilde{\mathbb{P}}$; da cui per la necessità dell'APT1 si ottiene che non ci sono opportunità di arbitraggio e quindi si ha

$$C_* \le C^*;$$

2 si mostra che vale la $\widetilde{S}$-rappresentabilità, e quindi che il mercato $(\widetilde{B}_n, \widetilde{S}_n)$ è completo, da cui deve valere che, se x_0 è il capitale iniziale che permette una copertura perfetta per f_N, allora

$$C^* \le x_0 \le C_*;$$

3 dai punti precedenti si ottiene che allora l'unico prezzo che non permette arbitraggi è proprio x_0

4 dalla dimostrazione del lemma si ha inoltre che

$$x_0 = B_0 M_0 = B_0 \widetilde{\mathbb{E}}[\frac{f_N}{B_N}]$$

e la strategia di copertura si ottiene prendendo

$$\gamma_k^* = \gamma_k$$

della rappresentazione della martingala $\widetilde{\mathbb{E}}[f_N|\mathcal{F}_n]$ e che

$$\beta_0^* = M_0 - \gamma_0 S_0, \qquad \beta_n^* = \beta_{n-1}^* - (\gamma_n - \gamma_{n-1}) S_{n-1},$$

Il precedente programma verrà realizzato nel seguente capitolo in cui si tratta il modello binomiale multiperiodale.

Si noti comunque che procedendo come nella dimostrazione del lemma (parte: la completezza implica la $\widetilde{S}$-rappresentabilità) si può ottenere la dimostrazione che la completezza implica l'unicità della misura martingala equivalente, se si è in presenza di un mercato senza opportunità di arbitraggio.

Dimostrazione della sufficienza dell'APT2: se il mercato è completo, allora, qualunque sia $A \in \mathcal{F}_N = \mathcal{F}$, per $\widetilde{f}_N = \mathbb{I}_A$ si può trovare un capitale iniziale $\widetilde{x}_A$ e una strategia π_A^*, che permettono una copertura perfetta di $\widetilde{f}_N$. Allora necessariamente deve essere

$$\widetilde{x}_A = \widetilde{\mathbb{E}}(\mathbb{I}_A) = \widetilde{\mathbb{P}}(A). \tag{5.43}$$

Ragionando come nei punti **1**, **2** e **3**, e tenendo presente che per ipotesi il mercato è privo di opportunità di arbitraggio, si ottiene che

$$\widetilde{x}_A = C_*(\mathbb{I}_A, \mathbb{P}) = C^*(\mathbb{I}_A, \mathbb{P}), \tag{5.44}$$

da cui l'unicità della misura martingala equivalente[24].

[24] Da (5.44) si ottiene che il capitale iniziale che permette una copertura perfetta è univocamente definito come $C_*(\mathbb{I}_A, \mathbb{P}) = C^*(\mathbb{I}_A, \mathbb{P})$.
Se esistesse una seconda misura martingala equivalente $\widetilde{\mathbb{P}}'$, allora per (5.43), applicata a questa nuova misura, si avrebbe che

$$\widetilde{x}'_A = \widetilde{\mathbb{E}}'(\mathbb{I}_A) = \widetilde{\mathbb{P}}'(A),$$

e contemporaneamente $\widetilde{x}'_A = C_*(\mathbb{I}_A, \mathbb{P}) = C^*(\mathbb{I}_A, \mathbb{P}) = \widetilde{x}_A$, da cui

$$\widetilde{\mathbb{P}}'(A) = \widetilde{\mathbb{P}}(A),$$

ovvero l'unicità.

Capitolo 6

Il modello di Cox Ross Rubinstein *(CRR-model)*

In questo capitolo si vuole illustrare l'applicazione di alcuni risultati teorici trattati, ovvero considerando un modello del quale si conoscono le caratteristiche peculiari si vogliono trarre delle conclusioni in merito al mercato su cui esso opera (arbitrage-free, completo, incompleto) e ai prezzi di copertura per opzioni Europee.

6.1 Caratteristiche del modello

Il modello che si andrà a considerare è il ***Cox Ross Rubinstein*** (CRR-*model*) detto anche ***modello binomiale multiperiodale***: si prende un mercato (B,S) che risulta formato da due operazioni finanziarie:

1 un conto bancario $B = (B_n)$

2 un'azione $S = (S_n)$

per le quali si ha (ricordando le formule (5.1) e (5.2))

$$\triangle B_n = r_n B_{n-1};$$

$$\triangle S_n = \rho_n S_{n-1}.$$

Si ipotizza che il tasso di interesse[1] sia costante $r_n = r$ e che la successione di variabili aleatorie indipendenti $\rho = (\rho_n)$ possa prendere solo due valori[2] a e b tali che

$$-1 < a < r < b. \tag{6.1}$$

Inoltre si assume che la successione $\rho = (\rho_n)$ definita sullo spazio di probabilità $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\geqslant 0}, \mathbb{P})$ sia $\mathcal{F}_n$-misurabile per ogni n e abbia la proprietà:

$$\mathbb{P}(\rho_n = b) = p \qquad \text{e} \qquad \mathbb{P}(\rho_n = a) = q \tag{6.2}$$

con $p + q = 1$ e $0 < p < 1$.
Si può osservare che tutta l'aleatorietàà del modello risulta data dalle variabili ρ_n, quindi si può assumere come spazio dei risultati elementari lo spazio $\Omega = \Omega_n = \{a,b\}^N$ di successioni finite $x = (x_1, x_2, \dots, x_N)$ tali che $x_n = a$ o $x_n = b$ con $n \leqslant N$. Allora $\rho_n(x) = x_n$ e la misura di probabilità $\mathbb{P}_N$ sui corrispondenti insiemi di Borel risulta completamente definita dalle distribuzioni finito dimensionali $\mathbb{P}_n$ dove $n \leqslant N$: se $\nu_b(x_1, \dots, x_n) = \sum_{i=1}^n \mathbb{I}_b(x_i)$ è il numero di componenti $x_i = b$ per $i \leqslant n$ allora

$$\mathbb{P}_n(x_1, \dots, x_n) = p^{\nu_b(x_1,\dots,x_n)} q^{n-\nu_b(x_1,\dots,x_n)}. \tag{6.3}$$

[1]Nel testo di T. Björk [4] si pone $R = r$, mentre nel testo di S. Ross [15] si usa la stessa notazione.

[2]Nei testi di S. Ross [15] e di T. Björk [4] si pone $u = 1 + b$ e $d = 1 + a$, o equivalentemente si ha qui $b = u - 1$ e $a = d - 1$. Inoltre la condizione (6.1) seguente corrisponde alla condizione $0 < d < 1 + r < u$ e $0 < d < 1 + R < u$, rispettivamente.

Da tale considerazione segue che $\mathbb{P}_n$ è uguale[3] a un prodotto diretto di n misure di tipo Q dove con Q si indica la misura caratterizzata da $Q(\{b\}) = p$ e $Q(\{a\}) = q$.

Nei prossimi paragrafi si mostrerà che il modello CRR è arbitrage-free e completo. Per i Teoremi *APT1* e *APT2* (si vedano i Teoremi 5.1 e 5.2) ciò significa, rispettivamente, che per ogni $n \geqslant 1$ esiste ed è unica la misura martingala $\widetilde{\mathbb{P}}_n$ equivalente a $\mathbb{P}_n$, si otterrà che $\widetilde{\mathbb{P}}_n$ presenta la seguente struttura:

$$\widetilde{\mathbb{P}}_n(x_1, \dots, x_n) = \widetilde{p}^{\,\nu_b(x_1,\dots,x_n)}\, \widetilde{q}^{\,n-\nu_b(x_1,\dots,x_n)} \tag{6.4}$$

dove

$$\widetilde{p} = \frac{r-a}{b-a} \qquad \text{e} \qquad \widetilde{q} = \frac{b-r}{b-a}. \tag{6.5}$$

Si può osservare che dalla (6.4) segue che anche $\widetilde{\mathbb{P}}_n$, così come $\mathbb{P}_n$, presenta la struttura di un prodotto diretto di misure $\widetilde{Q}$ dove $\widetilde{Q}(\{b\}) = \widetilde{p}$ e $\widetilde{Q}(\{a\}) = \widetilde{q}$.

6.1.1 CRR è arbitrage-free e completo

Definita la struttura del modello si vuole dimostrare che il mercato su cui opera è senza opportunità di arbitraggio. Per realizzare tale scopo, come ricordato sopra, basta dimostrare l'esistenza di una misura di probabilità equivalente alla misura di probabilità definita dalla (6.3) rispetto alla quale la successione $\widetilde{S}_n = \frac{S_n}{B_n}$ è una martingala.

Si cerca quindi di individuare quali sono le condizioni che le misure $\tilde{\mathbb{P}}$ martingala equivalenti (a $\mathbb{P}$) devono soddisfare. Si ricorda che $\widetilde{S}_n$ è una $(\widetilde{\mathbb{P}}, \mathcal{F}_n)$-martingala se risulta $\mathcal{F}_n$-adattato, integrabile e gode della proprietà $\widetilde{\mathbb{E}}(\widetilde{S}_n|\mathcal{F}_{n-1}) = \widetilde{S}_{n-1}$.

L'integrabilità in questo caso non comporta alcun problema poiché $\widetilde{S}_n = \frac{S_n}{B_n}$ assume un numero finito di valori e in particolare risulta uniformemente limitata [4]. Inoltre essendo in generale:

$$\frac{S_n}{B_n} = \frac{S_{n-1}}{B_{n-1}} \frac{1+\rho_n}{1+r_n} \tag{6.6}$$

e considerata la $\mathcal{F}_{n-1}$-misurabilità di $\widetilde{S}_{n-1}$ e di $1 + r_n$ risulta chiaro[5] che la proprietà di martingala rispetto a $\tilde{\mathbb{P}}$ diviene

$$\begin{aligned}
\widetilde{\mathbb{E}}\left[\frac{S_n}{B_n}\Big|\mathcal{F}_{n-1}\right] = \frac{S_{n-1}}{B_{n-1}} \qquad &\Leftrightarrow \qquad \widetilde{\mathbb{E}}\left[\frac{1+\rho_n}{1+r_n}\Big|\mathcal{F}_{n-1}\right] = 1 \\
&\Leftrightarrow \qquad \widetilde{\mathbb{E}}\left[\rho_n|\mathcal{F}_{n-1}\right] = r_n \qquad\qquad (i) \\
&\qquad\qquad\qquad \Downarrow \\
&\qquad\qquad \widetilde{\mathbb{E}}[\rho_n] = \widetilde{\mathbb{E}}[r_n]. \qquad\qquad (ii)
\end{aligned}$$

[3] Il punto importante qui è che $\mathbb{P}_n(x_1, \dots, x_n) > 0$ per ogni $(x_1, \dots, x_n)$

[4] In questo caso possiamo anche trovare una maggiorazione superiore esplicita: infatti

$$\frac{S_n}{B_n} = \frac{S_0}{B_0} \prod_{k=1}^{n} \frac{1+\rho_k}{1+r_k} \leqslant \frac{S_0}{B_0} \left(\frac{1+b}{1+r}\right)^n.$$

[5] Si osservi che

$$\begin{aligned}
\Delta\widetilde{S}_n &= \frac{S_n}{B_n} - \frac{S_{n-1}}{B_{n-1}} = \frac{S_{n-1}}{B_{n-1}}\Big(\frac{1+\rho_n}{1+r_n} - 1\Big) \\
&= \frac{S_{n-1}}{B_{n-1}}\Big(\frac{1+\rho_n-(1+r_n)}{1+r_n}\Big) = \frac{S_{n-1}}{B_{n-1}} \frac{1}{1+r_n}\left[\rho_n - r_n\right]
\end{aligned}$$

e quindi

$$\begin{aligned}
\widetilde{\mathbb{E}}\left[\Delta\widetilde{S}_n\Big|\mathcal{F}_{n-1}\right] &= \widetilde{\mathbb{E}}\left[\frac{S_n}{B_n} - \frac{S_{n-1}}{B_{n-1}}\Big|\mathcal{F}_{n-1}\right] = \widetilde{\mathbb{E}}\left[\frac{S_{n-1}}{B_{n-1}}\big(\frac{1+\rho_n}{1+r_n} - 1\big)\Big|\mathcal{F}_{n-1}\right] \\
&= \frac{S_{n-1}}{B_{n-1}}\widetilde{\mathbb{E}}\left[\big(\frac{1+\rho_n-(1+r_n)}{1+r_n}\big)\Big|\mathcal{F}_{n-1}\right] = \frac{S_{n-1}}{B_{n-1}}\frac{1}{1+r_n}\widetilde{\mathbb{E}}\left[\rho_n - r_n\Big|\mathcal{F}_{n-1}\right]
\end{aligned}$$

Quindi, nel caso del modello CRR considerato, in cui r_n è costante e uguale a r, e ρ_n può assumere solo i due valori a e b, si ottengono le seguenti identità[6]

$$\widetilde{\mathbb{E}}\left(\rho_n|\mathcal{F}_{n-1}\right)=r, \quad \Leftrightarrow \quad b\widetilde{\mathbb{P}}\left(\rho_n=b|\mathcal{F}_{n-1}\right)+a\widetilde{\mathbb{P}}\left(\rho_n=a|\mathcal{F}_{n-1}\right)=r, \tag{6.7}$$

$$\widetilde{\mathbb{E}}(\rho_n)=r, \quad \Leftrightarrow \quad b\widetilde{\mathbb{P}}\left(\rho_n=b\right)+a\widetilde{\mathbb{P}}\left(\rho_n=a\right)=r, \tag{6.8}$$

ovvero,

$$b\widetilde{\mathbb{P}}\left(\rho_n=b|\mathcal{F}_{n-1}\right)+a\left(1-\widetilde{\mathbb{P}}\left(\rho_n=b|\mathcal{F}_{n-1}\right)\right)=r \tag{6.9}$$

$$b\widetilde{\mathbb{P}}\left(\rho_n=b\right)+a\left(1-\widetilde{\mathbb{P}}\left(\rho_n=b\right)\right)=r, \tag{6.10}$$

Si osservi che la (6.9) e la (6.10) implicano

$$\widetilde{\mathbb{P}}(\rho_n=b|\mathcal{F}_{n-1}) = \widetilde{\mathbb{P}}(\rho_n=b) = \widetilde{p} = \tfrac{r-a}{b-a} \tag{6.11}$$

$$\widetilde{\mathbb{P}}(\rho_n=a|\mathcal{F}_{n-1}) = \widetilde{\mathbb{P}}(\rho_n=a) = \widetilde{q} = \tfrac{b-r}{b-a}. \tag{6.12}$$

Da questa osservazione, e dal fatto che $\mathcal{F}_{n-1}=\sigma\{\rho_1,\cdots,\rho_{n-1}\}$ si ottiene facilmente che le variabili aleatorie $\rho_1,\cdots,\rho_N$ sono, rispetto a $\widetilde{\mathbb{P}}$, indipendenti ed identicamente distribuite[7]. Segue che *il modello CRR è un*

[6] Si noti che la condizione (i) caratterizza le misure martingala equivalenti mentre la condizione (ii) è solo una condizione necessaria affinché $\widetilde{\mathbb{P}}$ sia una misura martingala equivalente. Inoltre, grazie al fatto che ρ_n assume solo due valori, la (i) permette di individuare la distribuzione condizionata di ρ_n data $\mathcal{F}_{n-1}$, in funzione di $r_n=r_n(\omega)$. Per lo stesso motivo la (ii) permette di individuare la distribuzione di ρ_n (entrambe rispetto a $\widetilde{\mathbb{P}}$). Infine, il fatto che $r_n(\omega)$ sia una costante r implica che $r_n=r$ e quindi che la distribuzione condizionata di ρ coincida con la distribuzione stessa.

[7] ‡ Un modo più elegante per pervenire a questo stesso risultato e verificare l'indipendenza delle ρ_n si ottiene illustrando la costruzione di $\widetilde{\mathbb{P}}$.
Dalle premesse fatte si evince che si può costruire la misura $\widetilde{\mathbb{P}}\sim\mathbb{P}$ attraverso dei passi ben precisi: si prende $\mathbb{P}_n=\mathbb{P}_N|_{\mathcal{F}_n}$, dove $\mathcal{F}_n=\sigma(\rho_1,\dots,\rho_n)$, si considera una $(\mathbb{P},\mathcal{F}_n)$-martingala Z_n di media 1 e si definiscono $\widetilde{\mathbb{P}}_1,\widetilde{\mathbb{P}}_2,\dots,\widetilde{\mathbb{P}}_N$ tramite la formula

$$\widetilde{\mathbb{P}}_n(x_1,\dots,x_n)=Z_n(x_1,\dots,x_n)\mathbb{P}_n(x_1,\dots,x_n),$$

quindi, dalla formula di Bayes segue che la condizione (6.7) può essere espressa come

$$\mathbb{E}_n\left(\rho_n\frac{Z_n}{Z_{n-1}}\Big|\mathcal{F}_{n-1}\right)=r. \tag{6.13}$$

Per $n=1$, tenendo conto che $\mathcal{F}_0=\{\emptyset,\Omega\}$, che $Z_0=1$ e considerando la relazione (6.13) risulta

$$p\,bZ_1(b)+q\,aZ_1(a)=r, \tag{6.14}$$

tale equazione unita alla condizione di normalizzazione

$$pZ_1(b)+qZ_1(a)=1, \tag{6.15}$$

permette di ottenere

$$Z_1(b)=\frac{r-a}{b-a}\frac{1}{p} \qquad \text{e} \qquad Z_1(a)=\frac{b-r}{b-a}\frac{1}{q}.$$

Ponendo

$$\widetilde{p}=\frac{r-a}{b-a} \qquad \text{e} \qquad \widetilde{q}=\frac{b-r}{b-a}$$

si arriva all'identità

$$\widetilde{\mathbb{P}}_1(b)=Z_1(b)\mathbb{P}_1(b)=\widetilde{p}$$
$$\widetilde{\mathbb{P}}_1(a)=Z_1(a)\mathbb{P}_1(a)=\widetilde{q}.$$

Si osservi che $\widetilde{p}+\widetilde{q}=1$ e che, per la (6.1), $\widetilde{p},\widetilde{q}>0$ da cui segue $\widetilde{\mathbb{P}}_1$ è una probabilità equivalente a $\mathbb{P}_1$. Per determinare $\widetilde{\mathbb{P}}_2$ si usa nuovamente la (6.13) e l'indipendenza tra ρ_1 e ρ_2 rispetto a $\mathbb{P}_2$ arrivando al seguente risultato

$$p\,b\frac{Z_2(b,b)}{Z_1(b)}+q\,a\frac{Z_2(b,a)}{Z_1(b)}=r, \tag{6.16}$$

una condizione ulteriore sui valori di $Z_2(b,b)$ e $Z_2(b,a)$ è dovuta alla proprietà di martingala di Z_n

$$\mathbb{E}_2(Z_2(\rho_1,\rho_2)|\rho_1=b)=Z_1(b),$$

che conduce all'uguaglianza

$$b\frac{Z_2(b,b)}{Z_1(b)}+q\frac{Z_2(b,a)}{Z_1(b)}=1, \tag{6.17}$$

mercato privo di opportunità di arbitraggio; analizzando i passi che conducono alla costruzione di $\widetilde{\mathbb{P}}$ si evince che quest'ultima è l'*unica* misura martingala equivalente da ciò discende la *completezza*, per il Teorema APT2. Tuttavia la dimostrazione che l'unicità della misura martingala equivalente implica la completezza non è stata data in queste note, mentre è stata data la dimostrazione che la $\widetilde{S}$-rappresentabilità implica la completezza. Si procederà quindi ora alla dimostrazione della $\widetilde{S}$-rappresentabilità per il modello CRR, in modo che la dimostrazione della completezza per questo modello sia autocontenuta in queste note.

6.1.2 $\widetilde{S}$-rappresentabilità

Nel Lemma 5.8 viene stabilita l'equivalenza tra la proprietà di completezza del mercato (B,S) e la S-rappresentabilità, sotto l'ipotesi che $B_n = 1$, ovvero l'equivalenza della proprietà di completezza del mercato $(\widetilde{B},\widetilde{S})$ e della $\widetilde{S}$-rappresentabilità. È interessante osservare che in questo caso è proprio l'unicità della misura martingala a permettere la $\widetilde{S}$-rappresentazione delle martingale limitate e quindi la completezza.
N.B. Si giunge così, per il modello in questione, a una dimostrazione diretta della necessità del teorema *APT2*.

Sia $M = (M_n)_{n\geqslant 0}$ una $(\widetilde{\mathbb{P}},\mathcal{F}_n)$-martingala e siano le $g_n = g_n(x_1,\dots,x_n)$ funzioni tali che

$$M_n(\omega) = g_n(\rho_1(\omega),\dots,\rho_n(\omega)), \tag{6.18}$$

ovvero[8]

$$M_n(\omega) = \mathbb{I}_{\{\rho_n(\omega)=b\}} g_n(\rho_1(\omega),\dots,\rho_{n-1}(\omega),b) + \mathbb{I}_{\{\rho_n(\omega)=a\}} g_n(\rho_1(\omega),\dots,\rho_{n-1}(\omega),a).$$

La condizione $\widetilde{\mathbb{E}}(M_n|\mathcal{F}_{n-1}) = M_{n-1}$ diviene quindi

$$\begin{aligned}&\widetilde{\mathbb{P}}(\rho_n = b|\mathcal{F}_{n-1})\, g_n(\rho_1(\omega),\dots,\rho_{n-1}(\omega),b) + \widetilde{\mathbb{P}}(\rho_n = a|\mathcal{F}_{n-1})\, g_n(\rho_1(\omega),\dots,\rho_{n-1}(\omega),a)\\ &= g_{n-1}(\rho_1(\omega),\dots,\rho_{n-1}(\omega)).\end{aligned}$$

Tenendo conto[9] delle (6.11) e (6.12) segue che la condizione che M è una $(\widetilde{\mathbb{P}},\mathcal{F}_n)$-martingala diviene

$$\begin{aligned}&\widetilde{p}\, g_n(\rho_1(\omega),\dots,\rho_{n-1}(\omega),b) + \widetilde{q}\, g_n(\rho_1(\omega),\dots,\rho_{n-1}(\omega),a)\\ &= g_{n-1}(\rho_1(\omega),\dots,\rho_{n-1}(\omega))\\ &= \widetilde{p}\, g_{n-1}(\rho_1(\omega),\dots,\rho_{n-1}(\omega)) + \widetilde{q}\, g_{n-1}(\rho_1(\omega),\dots,\rho_{n-1}(\omega))\end{aligned}$$

o equivalentemente

$$\begin{aligned}&\frac{g_n(\rho_1(\omega),\dots,\rho_{n-1}(\omega),b) - g_{n-1}(\rho_1(\omega),\dots,\rho_{n-1}(\omega))}{\widetilde{q}}\\ &= \frac{g_{n-1}(\rho_1(\omega),\dots,\rho_{n-1}(\omega)) - g_n(\rho_1(\omega),\dots,\rho_{n-1}(\omega),a)}{\widetilde{p}}.\end{aligned}$$

---

confrontando la (6.16) e la (6.17) rispettivamente con la (6.14) e la (6.15) si vede che

$$\frac{Z_2(b,b)}{Z_1(b)} = \frac{r-a}{b-a}\frac{1}{p} = \frac{\widetilde{p}}{p} \qquad \text{e} \qquad \frac{Z_2(b,a)}{Z_1(b)} = \frac{b-r}{b-a}\frac{1}{q} = \frac{\widetilde{q}}{q}.$$

In modo del tutto simile si può ottenere

$$\frac{Z_2(a,b)}{Z_1(a)} = \frac{\widetilde{p}}{p} \qquad \text{e} \qquad \frac{Z_2(a,a)}{Z_1(a)} = \frac{\widetilde{q}}{q};$$

quindi

$$\widetilde{\mathbb{P}}_2(a,a) = Z_2(a,a)q^2 = Z_1(a)\frac{\widetilde{q}}{q}\, q^2 = \widetilde{q}^{\,2}$$

e analogamente

$$\widetilde{\mathbb{P}}_2(a,b) = \widetilde{q}\,\widetilde{p}, \qquad \widetilde{\mathbb{P}}_2(b,a) = \widetilde{p}\,\widetilde{q}, \qquad \widetilde{\mathbb{P}}_2(b,b) = \widetilde{p}^{\,2}.$$

Le variabili aleatorie ρ_1,ρ_2 risultano, dunque, i.i.d. rispetto la misura $\widetilde{\mathbb{P}}_2$; inoltre $\widetilde{\mathbb{P}}_2(\rho_i = b) = \widetilde{p}$ e $\widetilde{\mathbb{P}}_2(\rho_i = a) = \widetilde{q}$ per $i = 1,2$.
L'ultimo passo per costruire una misura martingala $\widetilde{\mathbb{P}} \sim \mathbb{P}$ consiste nell'iterare il procedimento descritto per le misure $\widetilde{\mathbb{P}}_3,\dots,\widetilde{\mathbb{P}}_N$ ottenendo così le $\widetilde{\mathbb{P}}_n$ definite dalla formula (6.4) e infine ponendo $\widetilde{\mathbb{P}} = \widetilde{\mathbb{P}}_N$.

[8] Si osservi che l'espressione di M_n come

$$M_n(\omega) = \mathbb{I}_{\{\rho_n(\omega)=b\}} g_n(\rho_1(\omega),\dots,\rho_{n-1}(\omega),b) + \mathbb{I}_{\{\rho_n(\omega)=a\}} g_n(\rho_1(\omega),\dots,\rho_{n-1}(\omega),a).$$

corrisponde alla possibilità di rappresentare i valori che M_n assume su di un albero.

[9] Si ricordi che le condizioni (6.11) e (6.12) implicano che, rispetto a $\widetilde{\mathbb{P}}$ le ρ_n sono indipendenti ed identicamente distribuite.

In virtù dell'espressione(6.5) di $\widetilde{p}$ e $\widetilde{q}$ si definisce

$$\begin{aligned}\gamma'_n(\omega) &:= \frac{g_n(\rho_1(\omega),\ldots,\rho_{n-1}(\omega),b) - g_{n-1}(\rho_1(\omega),\ldots,\rho_{n-1}(\omega))}{b-r} \\ &= \frac{g_{n-1}(\rho_1(\omega),\ldots,\rho_{n-1}(\omega)) - g_n(\rho_1(\omega),\ldots,\rho_{n-1}(\omega),a)}{r-a}\end{aligned}$$

da cui segue

$$\begin{aligned}\Delta M_n(\omega) &= M_n(\omega) - M_{n-1}(\omega) \\ &= \mathbb{I}_{\{\rho_n(\omega)=b\}}[g_n(\rho_1(\omega),\ldots,\rho_{n-1}(\omega),b) - g_{n-1}(\rho_1(\omega),\ldots,\rho_{n-1}(\omega))] + \\ &+ \mathbb{I}_{\{\rho_n(\omega)=a\}}[g_n(\rho_1(\omega),\ldots,\rho_{n-1}(\omega),a) - g_{n-1}(\rho_1(\omega),\ldots,\rho_{n-1}(\omega))] \\ &= \mathbb{I}_{\{\rho_n(\omega)=b\}}(b-r)\gamma'_n(\omega) + \mathbb{I}_{\{\rho_n(\omega)=a\}}(a-r)\gamma'_n(\omega) \\ &= \sum_{x=\{a,b\}} \mathbb{I}_{\{\rho_n(\omega)=x\}}(x-r)\gamma'_n(\omega) \\ &= (\rho_n(\omega)-r)\gamma'_n(\omega).\end{aligned}$$

La $(\widetilde{\mathbb{P}}, \mathcal{F}_n)$-martingala $M = (M_n)_{n\geqslant 0}$ ammette, dunque, la rappresentazione

$$M_n(\omega) = M_0(\omega) + \sum_{k=1}^{n}(\rho_k(\omega)-r)\gamma'_k(\omega).$$

Infine, considerando la (6.6) si ha l'uguaglianza[10]

$$\rho_n(\omega) - r = (1+r)\frac{B_{n-1}}{S_{n-1}}\,\Delta\widetilde{S}_n,$$

da cui deriva la $\widetilde{S}$-rappresentabilità della martingala considerata, cioè

$$M_n(\omega) = M_0(\omega) + \sum_{k=1}^{n}\widetilde{\gamma}_k(\omega)\Delta\widetilde{S}_n, \tag{6.19}$$

dove si è assunto

$$\widetilde{\gamma}_n(\omega) = \gamma'_n(\omega)(1+r)\frac{B_{n-1}}{S_{n-1}} \tag{6.20}$$

$$= \frac{g_n(\rho_1(\omega),\ldots,\rho_{n-1}(\omega),b) - g_{n-1}(\rho_1(\omega),\ldots,\rho_{n-1}(\omega))}{b-r}(1+r)\frac{B_{n-1}}{S_{n-1}}. \tag{6.21}$$

Si osservi che poiché le $\gamma'_n(\omega)$ risultano $\mathcal{F}_{n-1}$-misurabili (questa proprietà è diretta conseguenza della loro definizione) la successione $\widetilde{\gamma}_n(\omega)$ è predicibile .

6.2 Prezzi di copertura per opzioni Europee

Si considerino le opzioni Europee di maturità $N < \infty$ con pay-off f_N dipendenti in generale da tutte le variabili $S_0, S_1, \ldots, S_N$ o, equivalentemente, da S_0 e $\rho_1, \ldots, \rho_N$. Tenendo conto del fatto che S_0 è una costante nota, in realtà il pay-off f_N è funzione solo di $\rho_1, \ldots, \rho_N$.

[10] La (6.6) implica, nel caso generale

$$\begin{aligned}\Delta\widetilde{S}_n &= \frac{S_n}{B_n} - \frac{S_{n-1}}{B_{n-1}} = \frac{S_{n-1}}{B_{n-1}}\Big(\frac{1+\rho_n}{1+r_n} - 1\Big) \\ &= \frac{S_{n-1}}{B_{n-1}}\Big(\frac{1+\rho_n-(1+r_n)}{1+r_n}\Big) = \frac{S_{n-1}}{B_{n-1}}\frac{1}{1+r_n}\left[\rho_n - r_n\right],\end{aligned}$$

da cui

$$[\rho_n - r_n] = (1+r_n)\frac{B_{n-1}}{S_{n-1}}\Delta\widetilde{S}_n.$$

Come visto nel capitolo II se il mercato considerato è senza opportunità di arbitraggio e completo (come risulta essere il mercato binomiale (B,S) del modello CRR) il *prezzo di copertura* (o premio) per l'acquisto dell'opzione, cioè

$$C(f_N,\mathbb{P}) = \inf\{x \geqslant 0 : \exists\, \pi \quad t.c. \quad X_0^\pi = x \quad e \quad X_N^\pi = f_N \quad \mathbb{P}\, q.c.\} \tag{6.22}$$

dove con $X^\pi = (X_n^\pi)_{0\leqslant n\leqslant N}$ si indica il valore della strategia autofinanziante $\pi = (\beta,\gamma)$, può essere determinato dalla identità:

$$C(f_N,\mathbb{P}) = B_0\, \widetilde{\mathbb{E}}\left(\frac{f_N}{B_N}\right). \tag{6.23}$$

Per il modello in oggetto essendo $B_N = B_0(1+r)^N$ si ottiene:

$$C(f_N,\mathbb{P}) = \widetilde{\mathbb{E}}\left(\frac{f_N}{(1+r)^N}\right), \tag{6.24}$$

tale risultato permette di rispondere completamente al problema di determinare un prezzo razionale per il contratto di un'opzione con pay-off f_N. Si ricorda che il venditore prendendo il premio $C(f_N,\mathbb{P})$ dal compratore può dotarsi di un portfolio $\widetilde{\pi} = (\widetilde{\beta},\widetilde{\gamma})$ che replica il pay-off f_N all'istante N, cioè

$$X_N^{\widetilde{\pi}} = f_N.$$

Come menzionato nella dimostrazione del Lemma 5.8 il modo standard per determinare il portfolio $\widetilde{\pi}$ consiste nel considerare, inizialmente, la $(\widetilde{\mathbb{P}},\mathcal{F}_n)$-martingala $M = (M_n)_{n\leqslant N}$ definita da

$$M_n = \widetilde{\mathbb{E}}\left[\frac{f_N}{B_N}\Big|\mathcal{F}_n\right].$$

Si noti che la martingala si può costruire a "ritroso" partendo da N in quanto $M_N = f_N = g_N(\rho_1,\cdots,\rho_N)$ e poi utilizzando il fatto che

$$M_{n-1} = \widetilde{\mathbb{E}}\left[M_n|\mathcal{F}_{n-1}\right] = \widetilde{p} g_n(\rho_1,\cdots,\rho_{n-1},b) + \widetilde{q} g_n(\rho_1,\cdots,\rho_{n-1},a),$$

come visto nel paragrafo precedente[11].

---

[11]Si noti l'analogia con la costruzione ad albero binomiale (vedere Björk [4] capitolo 2 paragrafo 2) per le opzioni europee con $f_N = \Phi(S_N)$, ed in particolare la Proposizione 2.24: si tenga presente che se $H_n(\omega)$, è la v.a. che conta il numero delle volte in cui il prezzo sale tra il passo 1 e il passo n, in [4] la funzione valore $V_n(k)$ coincide con $X_n^{\widetilde{\pi}}(\omega)$ quando $H_n = k$, mentre M_n coincide con il valore attualizzato $\widetilde{X}_n^{\widetilde{\pi}}(\omega)$.
Vale la pena di notare che se nella Proposizione 2.24 si mette r al posto di R, $1+a$ al posto di d, $1+b$ al posto di u, $\tilde{p}$ al posto di q_u, $\tilde{q}$ al posto di q_d, n al posto di t, N al posto di T, ed infine si divide per $(1+r)^n$ (o $(1+r)^N$) si ottiene:

$$\begin{cases} \frac{V_n(k)}{(1+r)^n} = \tilde{p}\,\frac{V_{n+1}(k+1)}{(1+r)^{n+1}} + \tilde{q}\,\frac{V_{n+1}(k)}{(1+r)^{n+1}} \\ \\ \frac{V_N(k)}{(1+r)^N} = \Phi(S_0(1+b)^k(1+a)^{N-k}). \end{cases}$$

Si definisca $g_n(\rho_1,\cdots,\rho_n) = \frac{V_n(H_n)}{(1+r)^n}$, di modo che, notando che se $\rho_{n+1} = b$ allora $H_{n+1} = H_n + 1$, mentre se $\rho_{n+1} = a$ allora $H_{n+1} = H_n$, si ha

$$g_{n+1}(\rho_1,\cdots,\rho_n,b) = \frac{V_{n+1}(H_n+1)}{(1+r)^{n+1}}, \qquad g_n(\rho_1,\cdots,\rho_n,a) = \frac{V_{n+1}(H_n)}{(1+r)^{n+1}}.$$

Di conseguenza, sostituendo a k la v.a. H_n (o H_N), si ottiene che il precedente sistema diviene

$$\begin{cases} g_n(\rho_1,\cdots,\rho_n) = \tilde{p}\, g_{n+1}(\rho_1,\cdots,\rho_n,b) + \tilde{q}\, g_{n+1}(\rho_1,\cdots,\rho_n,a) \\ \\ g_N(\rho_1,\cdots,\rho_N) = \Phi(S_0(1+b)^{H_N}(1+a)^{N-H_N}) = \Phi(S_N), \end{cases}$$

che corrisponde, se nostro modello si considera $f_N = \Phi(S_N)$, ala rappresentazione cercata di $M_n = g_n(\rho_1,\cdots,\rho_n)$, o equivalentemente corrisponde al sistema

$$\begin{cases} M_n = \widetilde{\mathbb{E}}\left[M_{n+1}|\mathcal{F}_n\right] \\ M_N = f_N. \end{cases}$$

Essendo M $\widetilde{S}$-rappresentabile esiste una successione predicibile $\tilde{\gamma} = (\tilde{\gamma}_n)_{n \leqslant N}$, si vedano[12] le (6.20) e (6.21), tale che la martingala risulta data da:

$$M_n = M_0 + \sum_{k=1}^{n} \tilde{\gamma}_k \Delta \left(\frac{S_k}{B_k} \right) \qquad n \leqslant N. \tag{6.25}$$

Si può riassumere quanto detto in questa sezione, tenendo conto anche dei risultati ottenuti nella Sezione 6.1.2, enunciando il seguente teorema:

Teorema 6.1. *Dato il modello CRR*

1 *per ogni N e per ogni pay-off f_N $\mathcal{F}_N$-misurabile il prezzo di esercizio può essere descritto dalla formula*

$$C(f_N, \mathbb{P}) = B_0 \, \widetilde{\mathbb{E}} \left[\frac{f_N}{B_N} \right] = \widetilde{\mathbb{E}} \left[\frac{f_N}{(1+r)^N} \right]. \tag{6.26}$$

2 *esiste una copertura perfetta e autofinanziante $\widetilde{\pi} = (\widetilde{\beta}, \widetilde{\gamma})$ dal valore $X^{\widetilde{\pi}} = (X^{\widetilde{\pi}}_n)_{n \leqslant N}$ tale che*

$$X^{\widetilde{\pi}}_0 = C(f_N; \mathbb{P}) \qquad\qquad X^{\widetilde{\pi}}_N = f_N$$

$$e \qquad X^{\widetilde{\pi}}_n = B_n \, \widetilde{\mathbb{E}} \left[\frac{f_N}{B_0 (1+r)^N} \Big| \mathcal{F}_n \right].$$

3 *le componenti $\widetilde{\beta} = (\widetilde{\beta}_n)_{n \leqslant N}$ e $\widetilde{\gamma} = (\widetilde{\gamma}_n)_{n \leqslant N}$ della copertura $\widetilde{\pi}$ soddisfano la relazione*[13]

$$\widetilde{\beta}_n = M_n - \widetilde{\gamma}_n \frac{S_n}{B_n} \tag{6.27}$$

dove $\widetilde{\gamma}_n$ con $n \leqslant N$ è determinata attraverso la $\frac{S}{B}$-rappresentazione (6.25) della $(\widetilde{\mathbb{P}}, \mathcal{F}_n)$-martingala $M = (M_n)_{n \leqslant N}$ definita come

$$M_n = \widetilde{\mathbb{E}} \left[\frac{f_N}{B_N} \Big| \mathcal{F}_n \right].$$

L'espressione di $\widetilde{\gamma}_n$ si ottiene attraverso la formula (6.21).

*** Per la costruzione esplicita della strategia $\tilde{\pi}$ e per le connessioni con gli alberi binomiali si veda l'Appendice 6.3, alla fine di questo capitolo. ***

Concludiamo questa sezione dando una seconda rappresentazione per la martingala M. Prendendo (come in (5.42))

$$\tilde{\beta}_n = M_n - \tilde{\gamma}_n \frac{S_n}{B_n}$$

[12] La formula (6.21), assicura che, se la martingala $M_n = \widetilde{\mathbb{E}} \left(\frac{f_N}{B_N} \Big| \mathcal{F}_n \right)$ si può scrivere attraverso una funzione g_n come in (6.18), allora il processo $\tilde{\gamma}_n$ definito da

$$\tilde{\gamma}_n(\omega) = \frac{g_n(\rho_1(\omega), \ldots, \rho_{n-1}(\omega), b) - g_{n-1}(\rho_1(\omega), \ldots, \rho_{n-1}(\omega))}{b - r} (1+r) \frac{B_{n-1}}{S_{n-1}},$$

è quello che permette di rappresentare M_n come integrale stocastico discreto come in (6.19). Sarebbe interessante controllare se questa strategia corrisponde a quella proposta nella Proposizione 2.24 di Björk, tenendo presente quanto detto nella nota precedente.

[13] Se si interpreta M_n come il valore attualizzato del portafoglio al tempo n allora la relazione (6.27) diviene ovvia se riscritta come

$$M_n = \widetilde{\beta}_n + \widetilde{\gamma}_n \frac{S_n}{B_n}.$$

si ottiene una copertura autofinanziante $\tilde{\pi} = (\tilde{\beta}, \tilde{\gamma})$ di valore

$$X_n^\pi = \tilde{\beta}_n B_n + \tilde{\gamma}_n S_n = B_n \widetilde{\mathbb{E}} \left(\frac{f_N}{B_N} \Big| \mathcal{F}_n \right)$$

tale che inizialmente

$$X_0^\pi = C(f_N, \mathbb{P})$$

e all'istante N goda della proprietà di copertura perfetta.

Considerando inoltre la relazione

$$\Delta \left(\frac{S_n}{B_n} \right) = \frac{S_{n-1}(\rho_n - r)}{B_n}$$

e sostituendo questo risultato nella (6.25) si ottiene:

$$M_n = M_0 + \sum_{k=1}^n \tilde{\gamma}_k \frac{S_{k-1}}{B_k}(\rho_k - r) = M_0 + \sum_{k=1}^n \alpha_k (\rho_k - r) \tag{6.28}$$

avendo definito α_k in modo che valga la condizione $\tilde{\gamma}_k = \alpha_k \frac{B_k}{S_{k-1}}$.

Prendendo, infine, la successione $\delta = (\delta_n)$ delle variabili aleatorie

$$\delta_n = \frac{\rho_n - a}{b - a}$$

risulta immediato che

$$\rho_n = \begin{cases} b \\ a \end{cases} \qquad \Longleftrightarrow \qquad \delta_n = \begin{cases} 1 \\ 0 \end{cases}$$

per cui $\delta_n = 1$ se e solo se il prezzo sale al passo n e inoltre $\mathcal{F}_n = \sigma(\rho_1, \ldots, \rho_n) = \sigma(\delta_1, \ldots, \delta_n)$.
Dalla relazione:

$$\delta_n - \widetilde{p} = \frac{\rho_n - r}{b - a}$$

appare evidente che oltre alla (6.25) e alla (6.28) si ha anche una rappresentazione di M in funzione della successione di variabili aleatorie $\delta = \{\delta_n\}$, ovvero:

$$M_n = M_0 + \sum_{k=1}^n \widetilde{\alpha}'_k \Delta m_k^{(\delta)} \tag{6.29}$$

dove la successione $m^{(\delta)} = \left(m_n^{(\delta)} \right)_{n \leqslant N}$ di variabili $m_n^{(\delta)} = \sum_{k=1}^n (\delta_k - \widetilde{p}) = H_n - n\widetilde{p}$ è una $(\widetilde{\mathbb{P}}, \mathcal{F}_n)$-martingala e $\widetilde{\alpha}'_n = (b-a)\widetilde{\alpha}_n$.
Si osservi che la scelta di esprimere M in termini della successione δ risulta utile nel calcolo esplicito del valore del premio, in quanto la variabile aleatoria $H_n = \sum_{k=1}^n \delta_k$, che conta il numero delle volte in cui il prezzo sale tra il passo 1 e il passo n, ha legge binomiale $Bin(n, \tilde{p})$ rispetto alla misura martingala equivalente.

6.2.1 Calcolo del prezzo di copertura per l'opzione call

Per un'opzione call standard (si veda la premessa del capitolo II) la funzione di pay-off f_N risulta pari a:

$$f_N = (S_N - K)^+ \tag{6.30}$$

dove N indica il tempo di maturità e K il prezzo di esercizio. Applicando i risultati generali (descritti nel paragrafo precedente) al caso considerato si ha che la (6.24) diviene:

$$C(f_N, \mathbb{P}) = \widetilde{\mathbb{E}} \left(\frac{(S_N - K)^+}{(1+r)^N} \right). \tag{6.31}$$

Per sottolineare la dipendenza del prezzo di copertura dell'opzione call dal prezzo di strike viene introdotta la notazione $C_{call}(K, \mathbb{P}) = C(f_N, \mathbb{P})$.
Prendendo $H = H_N$ pari al numero di volte in cui l'azione è aumentata del fattore $(1+b)$ nel periodo di tempo che va da 1 a N, cioè, richiamando le notazioni del precedente paragrafo, $H = \sum_{k=1}^{N} \delta_k$ segue che la sua distribuzione è una binomiale di parametri N e $\widetilde{p}$, ovvero:

$$H \sim B(N, \widetilde{p}) \qquad \text{sotto} \quad \widetilde{\mathbb{P}}.$$

Si può assumere, dunque, che all'istante N l'azione presenti il seguente valore

$$S_N = S_0(1+a)^{N-H}(1+b)^H \tag{6.32}$$

andandolo a sostituire nella (6.31) e esplicitando il valore della media per la distribuzione in esame segue che il prezzo di acquisto risulta

$$\begin{aligned} C_{call}(K, \mathbb{P}) &= \frac{1}{(1+r)^N} \sum_{h=0}^{N} \binom{N}{h} \widetilde{p}^{\,h} (1-\widetilde{p})^{N-h} \cdot \\ &\cdot \quad \big(S_0(1+a)^{N-h}(1+b)^h - K\big)^+, \end{aligned} \tag{6.33}$$

se si prende h_0 come il più piccolo intero per cui è soddisfatta la disuguaglianza $S_0(1+a)^{N-h}(1+b)^h > K$ si può riscrivere la (6.33) in funzione di questo

$$\begin{aligned} C_{call}(K, \mathbb{P}) = S_0 &\sum_{h=h_0}^{N} \binom{N}{h} \widetilde{p}^{\,h} (1-\widetilde{p})^{N-h} \left(\frac{1+a}{1+r}\right)^{N-h} \left(\frac{1+b}{1+r}\right)^{h} \\ &- \frac{K}{(1+r)^N} \sum_{h=h_0}^{N} \binom{N}{h} \widetilde{p}^{\,h} (1-\widetilde{p})^{N-h} \\ = S_0 &\sum_{h=h_0}^{N} \binom{N}{h} \left(\frac{1+b}{1+r}\,\widetilde{p}\right)^{h} \left(\frac{1+a}{1+r}\,(1-\widetilde{p})\right)^{N-h} \\ &- \frac{K}{(1+r)^N} \sum_{h=h_0}^{N} \binom{N}{h} \widetilde{p}^{\,h} (1-\widetilde{p})^{N-h}. \end{aligned} \tag{6.34}$$

Si osserva che presa la funzione di sopravvivenza[14] di una binomiale calcolata nel punto x ovvero

$$\overline{B(N,p)}(x) = \sum_{h>x}^{N} \binom{N}{h} p^{\,h} (1-p)^{N-h} \tag{6.35}$$

si ottiene

$$C_{call}(K, \mathbb{P}) = S_0 \overline{B(N, p^*)}(x_0) - \frac{K}{(1+r)^N} \overline{B(N, \widetilde{p})}(x_0) \tag{6.36}$$

dove si è posto[15]

$$p^* = \frac{1+b}{1+r}\,\widetilde{p}. \tag{6.37}$$

[14]Si ricordi che la funzione di sopravvivenza di una variabile aleatoria X, con funzione di distribuzione $F(x) = \mathbb{P}(X \le x)$ è definita da $\overline{F}(x) = 1 - F(x) = \mathbb{P}(X > x)$. Nel caso di una variabile aleatoria binomiale di parametri n e θ, se si indica con

$$B(n,\theta)(x) = \mathbb{P}(X \le x) = \sum_{i=0}^{\lfloor x \rfloor} \binom{n}{i} \theta^{\,h} (1-\theta)^{n-i}$$

si ha che la funzione di sopravvivenza vale

$$\overline{B(n,\theta)}(x) = \mathbb{P}(X > x) = \sum_{i=1+\lfloor x \rfloor} \binom{n}{i} \theta^{\,h} (1-\theta)^{n-i}$$

[15]Si noti che ricordando che il valore di $\widetilde{p} = \frac{r-a}{b-a} \left(= \frac{1+r-d}{u-d}\right)$ si ottiene che

$$p^* = \frac{1+b}{1+r}\,\widetilde{p} = \frac{1+b}{1+r}\,\frac{r-a}{b-a}.$$

Osservazione 6.1. *Il valore x_0 deve soddisfare la condizione*

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

In termini di h_0 è equivalente richiedere che

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

per cui risolvendo si arriva a

$$h_0 = 1 + \left\lfloor \frac{\ln\left(\frac{K}{S_0(1+a)^N}\right)}{\ln\left(\frac{1+b}{1+a}\right)} \right\rfloor . \tag{6.38}$$

I risultati ottenuti portano ad enunciare il seguente teorema.

Teorema 6.2. *Il prezzo razionale per l'opzione Europea standard di tipo call con obbligazione (pay-off) terminale $f_N = (S_N - K)^+$ è pari a*

$$C_{call}(K, \mathbb{P}) = S_0 \overline{B(N, p^*)}(x_0) - \frac{K}{(1+r)^N}\overline{B(N, \widetilde{p})}(x_0)$$

dove $\overline{B(N,p)}$ è definita tramite la (6.35), p^* attraverso la (6.37) e si è assunto h_0 come nella (6.38).

Osservazione 6.2. *Si osservi che i risultati ottenuti nel caso delle opzioni call sono facilmente estendibili a quelle put (dove $f_N = (K - S_N)^+$), infatti dall'identità*

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

segue che il prezzo razionale di un'opzione put può essere definito dalla formula

$$\begin{aligned} C_{put}(K, \mathbb{P}) &= \widetilde{\mathbb{E}}\left(\frac{(K - S_N)^+}{(1+r)^N}\right) \\ &= C_{call}(K, \mathbb{P}) - \widetilde{\mathbb{E}}\left(\frac{S_N}{(1+r)^N}\right) + \frac{K}{(1+r)^N} \end{aligned}$$

ed essendo $\widetilde{\mathbb{E}}\left(\frac{S_N}{(1+r)^N}\right) = S_0$ si ottiene la seguente relazione

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

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

Da questa espressione è facile vedere che

$$1 - p^* = \frac{1+a}{1+r}(1 - \widetilde{p}) = \frac{1+a}{1+r}\,\frac{b-r}{b-a}.$$

Infatti

$$\begin{aligned} 1 - p^* &= 1 - \frac{1+b}{1+r}\widetilde{p} = 1 - \frac{1+b}{1+r}\,\frac{r-a}{b-a} \\ &= \frac{(1+r)(b-a) - (1+b)(r-a)}{(1+r)(b-a)} = \frac{b - a + br - ar - r + a - br + ba}{(1+r)(b-a)} \\ &= \frac{b - r + ba - ar}{(1+r)(b-a)} = \frac{(b-r)(1+a)}{(1+r)(b-a)}. \end{aligned}$$

Osservazione 6.3. *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, $C(f_N, \mathbb{P}) = B_0\, \widetilde{\mathbb{E}} \frac{f(S_N)}{B_N}$ 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 $f'(x) = f'(0) + \int_{(0,x]} \mu(dy)$, dove $\mu = \mu(dy)$ è una misura finita[16]*, 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

$$f(x) = f(0) + xf'(0) + \int_{(0,x]} (x-y)^+ \mu(dy) = f(0) + xf'(0) + \int_{(0,\infty)} (x-y)^+ \mu(dy), \tag{6.39}$$

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 $\widetilde{\mathbb{E}} \frac{S_N}{B_N} = \widetilde{\mathbb{E}} \frac{S_0}{B_0} = \frac{S_0}{B_0}$, e tenendo conto che $B_N = B_0(1+r)^N$ è deterministico, si ottiene

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

da cui segue, per la (6.31), che

$$C(f_N, \mathbb{P}) = \frac{f(0)}{(1+r)^N} + S_0 f'(0) + \int_{(0,\infty)} C_{call}(y, \mathbb{P}) \mu(dy). \tag{6.40}$$

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

---

[16] 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 (6.39) 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 $y < z \leq x$

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

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.

6.3 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 [8] o di Luenberger [10], ad esempio, ma anche nel testo di Björk [4]. 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 6.1).

La martingala 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 6.3). Essendo B_N un valore deterministico, si ha che

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

dove $\hat{g}$ è una funzione deterministica. 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 $n = N-1$ 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,
2) 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[17]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 $s_n \in \{s_0\, u^k\, d^{n-k},\ \text{con } k = 0, 1, ..., n\}$, attraverso i seguenti passi:

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 $s_{n-1} \in \{s_0\, u^k\, d^{n-1-k},\ \text{con } k = 0, 1, ..., n-1\}$;

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

(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

[17]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

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

come si vede subito tenendo conto del fatto che

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

Abbiamo quindi ottenuto la formula ricorsiva che viene presentata nel libro di Björk [4].

processo valore attualizzato $\tilde{X}_n^{\tilde{\pi}} = X_n^{\tilde{\pi}}/B_n$ gode delle proprietà che

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

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

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

necessariamente deve essere (si rammentino le osservazioni all'Esempio 4.1)

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

Da ciò discende anche che 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

$$X_n^{\tilde{\pi}} = B_n\, \hat{g}_n(S_n) = c(n, S_n), \tag{6.41}$$

dove

$$c(n, x) = c_N(n, x) := B_0\, (1 + r)^n \hat{g}_n(x). \tag{6.42}$$

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[18]. 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 $s_N \in \{s_0\, u^k\, d^{N-k},\ \text{con}\ k = 0, 1, ..., N\}$;

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 $s_{n-1} \in \{s_0\, u^k\, d^{n-1-k},\ \text{con}\ k = 0, 1, ..., n-1\}$;

Questa osservazione ha anche l'importante conseguenza che il prezzo dell'opzione si ottiene come[19]

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

[18] 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).

[19] Nelle notazioni del Björk [4] ciò corrisponde a

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

(II) Calcolo della strategia di copertura perfetta
La strategia di copertura perfetta si ottiene immediatamente dalla $\tilde{S}-$rappresentazione di $M_n = \widetilde{\mathbb{E}}\big[\tilde{f}_N|\mathcal{F}_n\big]$, 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.

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[20] M_n, allora la condizione che M_n sia una martingala, ossia

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

diviene la condizione[21]

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

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

$$\gamma_n'(\omega) := \frac{g_n(\rho_1(\omega),\cdots,\rho_{n-1}(\omega),b) - g_{n-1}(\rho_1(\omega),\cdots,\rho_{n-1}(\omega))}{b-r} \tag{6.43}$$

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

$$\gamma_n'(\omega) = \frac{g_n(\rho_1(\omega),\cdots,\rho_{n-1}(\omega),a) - g_{n-1}(\rho_1(\omega),\cdots,\rho_{n-1}(\omega))}{a-r}. \tag{6.44}$$

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

$$\tilde{\gamma}_n(\omega) = B_n\,\frac{\gamma_n'(\omega)}{S_{n-1}(\omega)}. \tag{6.45}$$

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,

$$\hat{g}_{n-1}(S_{n-1}) = \tilde{p}\,\hat{g}_n(S_{n-1}\,u) + \tilde{q}\hat{g}_n(S_{n-1}\,d), \tag{6.46}$$

[20] 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 $M_n = \hat{g}_n(S_n)$.

[21] 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 $\widetilde{\mathbb{E}}\big[\mathbf{1}_{\{\rho_n=x\}}|\mathcal{F}_n\big] = \widetilde{\mathbb{P}}\big(\{\rho_n=x\}|\mathcal{F}_n\big) = \widetilde{\mathbb{P}}\big(\{\rho_n=x\}\big)$ vale $\tilde{q}$ per $x=a$ e $\tilde{p}$ per $x=b$.

mentre la condizione che (6.43) coincide con (6.44) diviene

$$\gamma_n'(\omega) = \frac{\hat{g}_n(S_{n-1}(\omega)\,u) - \hat{g}_{n-1}(S_{n-1}(\omega))}{b-r} = \frac{\hat{g}_n(S_{n-1}(\omega)\,d) - \hat{g}_{n-1}(S_{n-1}(\omega))}{a-r}. \tag{6.47}$$

Inserendo in quest'ultima relazione (6.47) la precedente espressione (6.46) per $\hat{g}_{n-1}(S_{n-1})$, e tenendo conto che $\tilde{q} = 1 - \tilde{p} = \frac{b-r}{b-a}$ 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 (6.45)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 (6.41) e la definizione (6.42) 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 $n = 0, 1, ..., N$

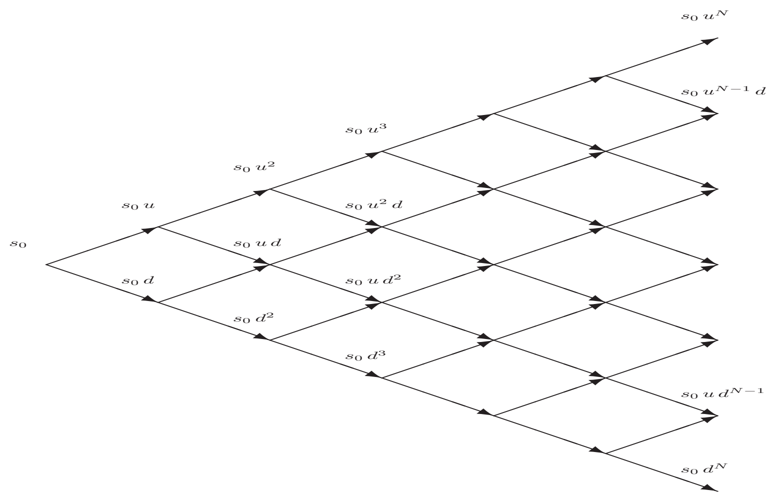

Figura 6.1: Albero binomiale per il calcolo di S_n per $n = 0, 1, ..., N$. Lo spazio Ω viene rappresentato come l'insieme dei cammini possibili.

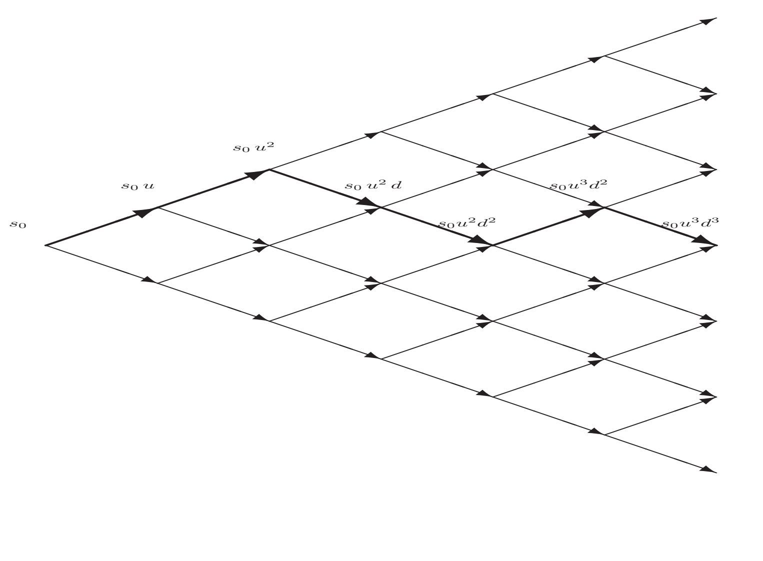

Figura 6.2: Albero binomiale per il calcolo di S_n per $n = 0, 1, ..., N$. Esempio di cammino: in grassetto si vede il cammino relativo all'*evento* $\{Z_1 = u, Z_2 = u, Z_3 = d, Z_4 = d, Z_5 = u, Z_6 = d\}$

$\hat{\boldsymbol{g}}_{\boldsymbol{N-1}}(\boldsymbol{s}_{\boldsymbol{N-1}}) =$

$= \tilde{\boldsymbol{p}} \boldsymbol{f}(\boldsymbol{s}_{\boldsymbol{N-1}}\, \boldsymbol{u})/\boldsymbol{B}_{\boldsymbol{N}} + \tilde{\boldsymbol{q}} \boldsymbol{f}(\boldsymbol{s}_{\boldsymbol{N-1}}\, \boldsymbol{d})/\boldsymbol{B}_{\boldsymbol{N}}$

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

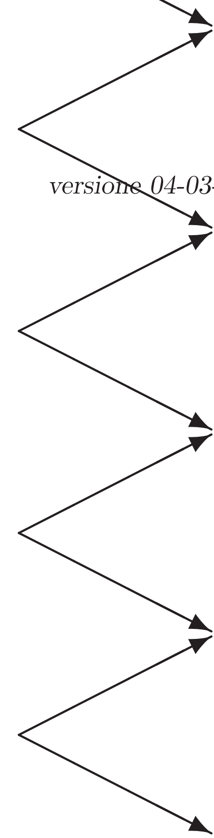

Figura 6.3: Albero binomiale per il calcolo di $\hat{g}_{N-1}$ e della strategia di copertura perfetta $\tilde{\gamma}_N$

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

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

$\hat{\boldsymbol{g}}_{\boldsymbol{n-1}}(\boldsymbol{s}_{\boldsymbol{n-1}}) =$

$= \tilde{\boldsymbol{p}} \hat{\boldsymbol{g}}_{\boldsymbol{n}}(\boldsymbol{s}_{\boldsymbol{n-1}}\, \boldsymbol{u}) + \tilde{\boldsymbol{q}} \hat{\boldsymbol{g}}_{\boldsymbol{n}}(\boldsymbol{s}_{\boldsymbol{n-1}}\, \boldsymbol{d})$

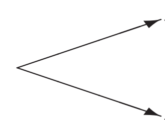

Figura 6.4: Albero binomiale per il calcolo della strategia di copertura perfetta $\tilde{\gamma}_n$.

Capitolo 7

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

7.1 Il Modello Binomiale Multiperiodale

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

7.1.1 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

$$S_0 = s_0 > 0, \qquad S_1 = (1+\rho_1)S_0 = Z_1 S_0 \quad \dots \quad S_n = (1+\rho_n)\cdots(1+\rho_2)(1+\rho_1)S_0 = Z_n \cdots Z_2 Z_1 S_0$$

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

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

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

$$C(f_N, \mathbb{P}) = C_N(f_N, \mathbb{P}) = B_0\, \widetilde{\mathbb{E}}\left[\frac{f_N}{B_N}\right] = \widetilde{\mathbb{E}}\left[\frac{f_N}{(1+r)^N}\right]. \tag{7.1}$$

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,

$$C_{call}(K, \mathbb{P}) = C_{call,N}(K, \mathbb{P})\widetilde{\mathbb{E}}\left[\frac{(S_N - K)^+}{B_N}\right] \tag{7.2}$$

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

$$\begin{aligned} C_{call}(K, \mathbb{P}) &= S_0 \sum_{x_0 < h \le N} \binom{N}{h} \widetilde{p}^{\,h}(1-\widetilde{p})^{N-h} \left(\frac{1+b}{1+r}\right)^h \left(\frac{1+a}{1+r}\right)^{N-h} \\ &\quad - \frac{K}{(1+r)^N} \sum_{x_0 < h \le N} \binom{N}{h} \widetilde{p}^{\,h}(1-\widetilde{p})^{N-h} \qquad (7.3) \\ &= S_0 \sum_{h=h_0}^{N} \binom{N}{h} \left(\frac{1+b}{1+r}\,\widetilde{p}\right)^h \left(\frac{1+a}{1+r}\,(1-\widetilde{p})\right)^{N-h} \\ &\quad - \frac{K}{(1+r)^N} \sum_{h=h_0}^{N} \binom{N}{h} \widetilde{p}^{\,h}(1-\widetilde{p})^{N-h}. \end{aligned}$$

$$\tag{7.4}$$

dove

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

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.

7.2 Approssimazione del Modello Binomiale Multiperiodale

7.2.1 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

$$B_t^{(n)} = B_{\frac{\lfloor nt \rfloor}{n}}^{(n)} = B_{\lfloor nt \rfloor} = B_0 \left(1 + r(n)\right)^{\lfloor nt \rfloor} \tag{7.5}$$

$$S_t^{(n)} = S_{\frac{\lfloor nt \rfloor}{n}}^{(n)} = S_{\lfloor nt \rfloor} = s_0 \left(\frac{u^{(n)}}{d^{(n)}}\right)^{H_{\lfloor nt \rfloor}} \left(d^{(n)}\right)^{\lfloor nt \rfloor} \tag{7.6}$$

dove il primo segno di uguaglianza sia in (7.5) che in (7.6), 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 (7.5) si sceglie

$$r^{(n)} = \frac{r}{n}, \tag{7.7}$$

e dove infine nell'ultima uguaglianza[3] in (7.6) si sceglie

$$u^{(n)} = e^{\sigma/\sqrt{n}} \qquad d^{(n)} = e^{-\sigma/\sqrt{n}}, \tag{7.8}$$

per una costante $\sigma > 0$.

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

$$N^{(n)} = N_n(T), \tag{7.9}$$

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 7.1. *Posto $C_{call}^{(n)}(K, \mathbb{P})$ il prezzo della call europea nel modello binomiale dato da* (7.5) *e* (7.6)*, con i parametri definiti come in* (7.7) *e* (7.8)*, prezzo di esercizio K e tempo di esercizio $N_n(T)$, si ottiene che*

$$\lim_{n\to\infty} C_{call}^{(n)}(K, \mathbb{P}) = s_0\, \Phi\left(-\zeta + \sigma\sqrt{T}\right) - K e^{-rT}\, \Phi(-\zeta), \tag{7.10}$$

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

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

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 $H_{\lfloor nt \rfloor}^{[n]}$, che sarebbe stata più precisa.

Osservazione 7.1. *Tenendo conto che* $-r\,T = \log(e^{-rT}) = -\log(e^{rT})$*, con semplici passaggi si ottiene*

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

da cui, poiché, come già ricordato $\Phi(x) = 1 - \Phi(-x)$*, il limite in* (7.10) *si riscrive come*

$$s_0\,\Phi\left(\tfrac{1}{\sigma\sqrt{T}}\log\left(\tfrac{s_0\,e^{rT}}{K}\right) + \tfrac{\sigma\sqrt{T}}{2}\right) - Ke^{-rT}\Phi(\tfrac{1}{\sigma\sqrt{T}}\log\left(\tfrac{s_0\,e^{rT}}{K}\right) - \tfrac{\sigma\sqrt{T}}{2}) \tag{7.11}$$

$$=s_0\,\left[\Phi\left(\tfrac{1}{\sigma\sqrt{T}}\log\left(\tfrac{s_0\,e^{rT}}{K}\right) + \tfrac{\sigma\sqrt{T}}{2}\right) - \tfrac{K}{s_0\,e^{rT}}\,\Phi(\tfrac{1}{\sigma\sqrt{T}}\log\left(\tfrac{s_0\,e^{rT}}{K}\right) - \tfrac{\sigma\sqrt{T}}{2})\right] \tag{7.12}$$

La forma (7.12)*, 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 (7.5) e (7.6). Ovviamente la scelta di $r^{(n)}$ in (7.7) 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

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

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

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

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

$$N_n(t) = \lfloor nt \rfloor \tag{7.13}$$

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

$$B_t^{(n)} = B^{(n)}_{\frac{\lfloor nt \rfloor}{n}} = B_0\left(1 + \frac{r}{n}\right)^{\lfloor nt \rfloor}. \tag{7.14}$$

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

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

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

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

[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 $K > s_0\,e^{rT}$.

la definizione (7.8) corrisponde a

$$u^{(n)} = 1 + (\sigma/\sqrt{n}) + \frac{1}{2}\,(\sigma^2/n) + o(1/n) \qquad d^{(n)} = 1 - (\sigma/\sqrt{n}) + \frac{1}{2}\,(\sigma^2/n) + o(1/n), \tag{7.15}$$

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

$$-(\sigma/\sqrt{n}) + \frac{1}{2}\,(\sigma^2/n) + o(1/n) < r/n < (\sigma/\sqrt{n}) + \frac{1}{2}\,(\sigma^2/n) + o(1/n), \tag{7.16}$$

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

$$S^{(n)}_t = S^{(n)}_{t^{(n)}_k}, \qquad \text{per } t \in [t^{(n)}_k, t^{(n)}_{k+1}) = [\tfrac{k}{n}, \tfrac{k+1}{n})$$

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

$$\begin{aligned}
S^{(n)}_t &= S_{\frac{\lfloor nt \rfloor}{n}}\big(= S^{[n]}_{\frac{\lfloor nt \rfloor}{n}}\big) = S_0 \left(\frac{u^{(n)}}{d^{(n)}}\right)^{H_{N_n(t)}} \left(d^{(n)}\right)^{\lfloor N_n(t) \rfloor} \\
&= S_0 \left(\frac{u^{(n)}}{d^{(n)}}\right)^{H_{\lfloor nt \rfloor}} \left(d^{(n)}\right)^{\lfloor nt \rfloor} \\
&= S_0\, e^{2\sigma \frac{1}{\sqrt{n}} H_{\lfloor nt \rfloor}}\, e^{-\sigma \frac{1}{\sqrt{n}} \lfloor nt \rfloor} \\
&= S_0\, e^{\sigma \frac{1}{\sqrt{n}} \left(2\, H_{\lfloor nt \rfloor} - \lfloor nt \rfloor\right)}.
\end{aligned} \tag{7.17}$$

Passiamo ora alla dimostrazione del Teorema 7.1, ossia della formula di Black e Scholes.

7.2.2 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 $\{Z^{(n)}_i = u\}$ 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

$$C^{(n)}(f,\mathbb{P}) = \widetilde{\mathbb{E}}^{(n)}\left[\frac{f(S_T^{(n)})}{B_T^{(n)}}\right] \tag{7.18}$$

in particolare per l'opzione call

$$C_{call}^{(n)}(K,\mathbb{P}) = \widetilde{\mathbb{E}}^{(n)}\left[\frac{(S_T^{(n)} - K)^+}{B_T^{(n)}}\right] = \widetilde{\mathbb{E}}^{(n)}\left[\frac{(S_{N_n(T)} - K)^+}{\left(1 + \frac{r}{n}\right)^{N_n(T)}}\right] \tag{7.19}$$

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.

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[5], 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

$$\log\left(\frac{u^{(n)}}{d^{(n)}}\right) = 2\,\frac{\sigma}{\sqrt{n}} \qquad \log\left(d^{(n)}\right) = -\frac{\sigma}{\sqrt{n}},$$

[5]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

$$\begin{aligned}
\widetilde{Var}^{(n)}\left(\log\left(S_t^{(n)}\right)\right) &= N_n(t)\,\tilde{p}^{(n)}(1 - \tilde{p}^{(n)})\left(\log\left(\frac{u^{(n)}}{d^{(n)}}\right)\right)^2\\
&= \lfloor nt \rfloor \left(\frac{1}{2} + \frac{1}{2\sqrt{n}}\,\frac{r - \frac{1}{2}\,\sigma^2}{\sigma} + o(\tfrac{1}{\sqrt{n}})\right)\left(1 - \left(\frac{1}{2} + \frac{1}{2\sqrt{n}}\,\frac{r - \frac{1}{2}\,\sigma^2}{\sigma} + o(\tfrac{1}{\sqrt{n}})\right)\right)\left(\frac{2\sigma}{\sqrt{n}}\right)^2\\
* &= \lfloor nt \rfloor \left(\frac{1}{2} + \frac{1}{2\sqrt{n}}\,\theta_n\right)\left(\frac{1}{2} - \frac{1}{2\sqrt{n}}\,\theta_n\right)\frac{4\sigma^2}{n} = \lfloor nt \rfloor \left(\frac{1}{4} - \frac{1}{4\,n^2}\,\theta_n^2\right)\frac{4\sigma^2}{n} *\\
&\approx \lfloor nt \rfloor\,\frac{1}{4}\,\frac{4\sigma^2}{n} \approx t\sigma^2
\end{aligned}$$

Di conseguenza, se W_t è una variabile aleatoria gaussiana di valore atteso 0 e varianza t

$$\log(S_t^{(n)}) \overset{distr}{\rightarrow} \log S_0 + t\left(r - \tfrac{1}{2}\,\sigma^2\right) + \sigma\,W_t. \tag{7.20}$$

Osservazione 7.2. *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 (7.17)*, 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 $r\,t - \frac{1}{2}\,\sigma^2\,t + \sigma\,W_t$*, ossia che*

$$\lim_{n\to\infty} \widetilde{\mathbb{E}}^{(n)}\left[\exp\{i\,u\log\left(S_t^{(n)}\right)\}\right] = e^{iu\left(r\,t) - \frac{1}{2}\,\sigma^2\,t\right) + \frac{1}{2}\,\sigma^2\,t\,u^2}.$$

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[6]*. 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 e' dedicata la sezione successiva.

Prima di proseguire nella dimostrazione del Teorema 7.1, 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[7] 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

$$C^{(n)}(f, \mathbb{P}) = \widetilde{\mathbb{E}}^{(n)}\left[\frac{f(S_T^{(n)})}{B_T^{(n)}}\right] \approx \mathbf{E}\left[\frac{1}{e^{rT}}\, f(S_0 e^{(r - \frac{1}{2}\,\sigma^2)T + \sigma\,W_T})\right] \tag{7.21}$$

ed in particolare per l'opzione call

$$C_{call}^{(n)}(K, \mathbb{P}) = \widetilde{\mathbb{E}}^{(n)}\left[\frac{(S_T^{(n)} - K)^+}{\big(1 + \frac{r}{n}\big)^{N_n(T)}}\right] \approx \mathbf{E}\left[\frac{1}{e^{rT}}\,(S_0 e^{(r - \frac{1}{2}\,\sigma^2)T + \sigma\,W_T} - K)^+\right], \tag{7.22}$$

[6]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 $x = 0$.

[7]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 $f(x) = \sin(tx)$.

dove l'unica variabile aleatoria è W_T, in quanto $S_0 = s_0$ è il valore iniziale del prezzo dell'azione.

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[8]

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[9] 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 $\Phi(x) = \Phi(-x)$, si ottiene allora

$$\begin{aligned} C_0(x) &= \frac{x}{\sqrt{2\pi}}\int_{\zeta}^{+\infty} e^{-rT}\,e^{(r-\frac{\sigma^2}{2})T+\sigma\sqrt{T}z}e^{-z^2/2}dz - \frac{K\,e^{-rT}}{\sqrt{2\pi}}\int_{\zeta}^{+\infty} e^{-z^2/2}dz \\ &= \frac{x}{\sqrt{2\pi}}\int_{\zeta}^{+\infty} e^{-\frac{1}{2}(z-\sigma\sqrt{T})^2}dz - Ke^{-rT}\Phi(-\zeta) \\ &= \frac{x}{\sqrt{2\pi}}\int_{\zeta-\sigma\sqrt{T}}^{+\infty} e^{-z^2/2}dz - Ke^{-rT}\Phi(-\zeta) \\ &= x\Phi\left(-\zeta + \sigma\sqrt{T}\right) - Ke^{-rT}\Phi(-\zeta), \end{aligned}$$

[8]In realtà la condizione che f sia una funzione continua è superflua.

[9]L'uguaglianza in legge si vede dall'espressione approssimata del logaritmo del prezzo (7.20), 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* (7.10).

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[10].

Va inoltre notato che poiché a tempo discreto vale l'interpretazione[11] 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 (7.10) in cui, però, al posto di s_0 andrebbe messo S_t e al posto di T andrebbe messo $T-t$:

$$C_t = S_t \left[\Phi\left(\tfrac{1}{\sigma\sqrt{T-t}} \log\left(\tfrac{x\, e^{r(T-t)}}{K}\right) + \tfrac{\sigma\sqrt{T-t}}{2} \right) - \tfrac{K}{x\, e^{r(T-t)}}\, \Phi\big(\tfrac{1}{\sigma\sqrt{T-t}} \log\left(\tfrac{x\, e^{r(T-t)}}{K}\right) - \tfrac{\sigma\sqrt{T-t}}{2} \big) \right] \Big|_{x=S_t}.$$

Grazie alla formula di parità (si veda ad esempio il libro di S. Ross [15]), 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. Alternativamente, per ottenere la formula di parità si può ripetere il ragionamento dell'Osservazione 6.2, e più in generale, per ottenere il prezzo di opzioni con terminal pay-off che dipende solo dal valore del sottostante al tempo T, si può ripetere il metodo dell'Osservazione 6.3.

Per ulteriori approfondimenti sai consiglia di consultare il libro di P. Baldi [1], o quello di D. Lamberton e B. Lapeyre [9], oppure di J.M. Steele [18] .

[10] 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.

[11] Va pero' tenuto presente che nella formula $c_N(n,f) = c_{N-n}(0,f)$, mentre nella prima (ossia $c_N(n,f)$) va considerato s_0 uguale al prezzo S_0 del sottostante al tempo 0, nella seconda (ossia $c_{N-n}(0,f)$) il valore s_0 va sostituito con S_n, cioè il prezzo del sottostante al tempo n.

7.3 Il moto Browniano

7.3.1 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

$$\sigma\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{N_n(t)}(2\xi_i-1)\right) = \sigma\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{\lfloor nt\rfloor}(2\xi_i-1)\right).$$

Si definisca

$$W_t^{(n)} := \frac{1}{\sqrt{n}}\sum_{i=1}^{\lfloor nt\rfloor}\left[(2\xi_i-1)-\widetilde{\mathbb{E}}^{(n)}[2\xi_i-1]\right] \tag{7.23}$$

di modo che

$$\log\left(S_t^{(n)}\right) = \log\left(S_0\right) + \sigma W_t^{(n)} + \sigma\frac{1}{\sqrt{n}}\sum_{i=1}^{\lfloor nt\rfloor}\widetilde{\mathbb{E}}^{(n)}[2\xi_i-1] \tag{7.24}$$

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 7.3. * *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*

$$X_j := 2\xi_j - 1 \quad ed \quad \widetilde{X}_j^{(n)} := \frac{1}{\sqrt{n}}\left[X_j - \widetilde{\mathbb{E}}^{(n)}[X_j]\right],$$

si può riscrivere (7.23) *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*

$$\widetilde{\mathbb{E}}^{(n)}\big[\exp\{i\,u\,W_t^{(n)}\}\big] = \prod_{j=1}^{\lfloor nt\rfloor} \widetilde{\mathbb{E}}^{(n)}\big[\exp\{i\,u\,\widetilde{X}_j^{(n)}\}\big]. \tag{7.25}$$

Le variabili aleatorie $\widetilde{X}_j^{(n)}$ *hanno media nulla e varianza*

$$\begin{aligned}\widetilde{Var}^{(n)}\big[\widetilde{X}_j^{(n)}\big] &= \widetilde{Var}^{(n)}\big[\frac{1}{\sqrt{n}}\,2\,\xi_j\big] = \left(\frac{1}{\sqrt{n}}\,2\right)^2 \tilde{p}^{(n)}(1-\tilde{p}^{(n)})\\ &= \frac{4}{n}\left(\frac{1}{2} + \frac{1}{2\sqrt{n}}\,\theta_n\right)\left(\frac{1}{2} - \frac{1}{2\sqrt{n}}\,\theta_n\right)\end{aligned}$$

dove $\theta_n = (r - \frac{1}{2}\,\sigma^2)/\sigma + o(1)$*, di modo che*

$$\widetilde{Var}^{(n)}\big[\widetilde{X}_j^{(n)}\big] = \frac{1}{n} + o\big(\frac{1}{n}\big).$$

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* (7.25)*, e procedere esattamente come nella dimostrazione del teorema centrale del limite.*

*

7.3.2 Indipendenza ed omogeneità degli incrementi

Ancora, se si considerano i tempi $0 \leq t_1 < t_2$, allora l'incremento $W_{t_2}^{(n)} - W_{t_1}^{(n)}$ ha la stessa legge di $W_{t_2-t_1}^{(n)}$, in quanto $W_{t_2}^{(n)} - W_{t_1}^{(n)}$ è funzione deterministica delle variabili aleatorie ξ_j, per j tale che $t_j^{(n)} = j/k \in (t_1, t_2]$, e la stessa funzione determina nello stesso modo $W_{t_2-t_1}^{(n)}$ a partire dalle variabili aleatorie ξ_i, per i tale che $t_i^{(n)} = i/n \in (0, t_2 - t_1]$, e di conseguenza la distribuzione dell'incremento $W_{t_2}^{(n)} - W_{t_1}^{(n)}$ 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

$$W_{t_1}^{(n)} - W_0^{(n)}, \quad W_{t_2}^{(n)} - W_{t_1}^{(n)}, \quad \ldots \quad W_{t_m}^{(n)} - W_{t_{m-1}}^{(n)}$$

[12]* Infatti, utilizzando le stesse notazioni dell'Osservazione 7.3, si ha

$$W_{t_2}^{(n)} - W_{t_1}^{(n)} = \sum_{j=\lfloor nt_1\rfloor+1}^{\lfloor nt_2\rfloor} \widetilde{X}_j^{(n)},$$

come si vede subito osservando

$$W_{t_2}^{(n)} = \sum_{j=1}^{\lfloor nt_2\rfloor} \widetilde{X}_j^{(n)} = \sum_{j=1}^{\lfloor nt_1\rfloor} \widetilde{X}_j^{(n)} + \sum_{j=\lfloor nt_1\rfloor+1}^{\lfloor nt_2\rfloor} \widetilde{X}_j^{(n)} = W_{t_1}^{(n)} + (W_{t_2}^{(n)} - W_{t_1}^{(n)}).$$

Questa scomposizione, insieme al fatto che $W_0^{(n)} = 0$ permette anche di vedere che $W_{t_1}^{(n)}\big(= W_{t_1}^{(n)} - W_0^{(n)} = W_{t_1}^{(n)} - 0\big)$ e $W_{t_2}^{(n)} - W_{t_1}^{(n)}$ sono indipendenti. L'indipendenza di questi incrementi permette di affermare che, oltre alla convergenza in distribuzione di $W_{t_1}^{(n)}$ ad una variabile aleatoria gaussiana di valore medio nullo e varianza t_1 e la convergenza di $W_{t_2}^{(n)} - W_{t_1}^{(n)}$ ad una variabile aleatoria gaussiana di valore medio nullo e varianza $t_2 - t_1$, si ha anche la convergenza congiunta di $\big(W_{t_1}^{(n)}, W_{t_2}^{(n)} - W_{t_1}^{(n)}\big)$ 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 $\big(W_{t_1}^{(n)}, W_{t_2}^{(n)}\big)$. Questa osservazione si estende facilmente al caso in cui si considerino le variabili aleatorie $\big(W_{t_1}^{(n)}, W_{t_2}^{(n)}, \cdots W_{t_m}^{(n)}\big)$, con $0 \leq t_1 < t_2 < \ldots < t_{m-1} < t_m$. *

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.

7.3.3 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 7.1 (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.*

Inoltre quanto abbiamo visto nella sezione precedente permette di dare una definizione di un processo S_t, che, alla luce di (7.24), 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 7.4 (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*

$$\widetilde{S}_t = \frac{S_t}{B_t} = \frac{S_t}{B_0\, e^{rt}} = \frac{S_0}{B_0} \exp\left\{\sigma W_t - \tfrac{1}{2}\sigma^2\, t\right\}, \tag{7.26}$$

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*

$$\widetilde{S}_t = \frac{S_t}{B_t} = \frac{S_t}{B_0\, e^{rt}} = \frac{S_0}{B_0} \exp\left\{(\mu - r)\, t + \sigma\, \mathbb{W}_t - \tfrac{1}{2}\sigma^2\, t\right\}, \tag{7.27}$$

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 8

Processi aleatori a tempo continuo

8.1 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 8.1 (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 8.2 (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 8.3 (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*

---

[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 8.1 e la nota 8.1 corrispondente, per maggiori dettagli si veda l'Appendice 8.8 .

(misurabile)[3]

$$X : \Omega \mapsto C(I;\mathbb{R}); \omega \mapsto \big(t \mapsto X(t,\omega)\big),$$

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 8.4 (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

$$F_{t_1,\cdots,t_k}(x_1, \cdots, x_k) = F_{t_{\pi_1},\cdots,t_{\pi_k}}(\Phi_\pi(x_1, \cdots, x_k)) = F_{t_{\pi_1},\cdots,t_{\pi_k}}(x_{\pi_1}, \cdots, x_{\pi_k}). \tag{8.1}$$

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

$$\lim_{x_{k+1}\to\infty} F_{t_1,\cdots,t_k,t_{k+1}}(x_1, \cdots, x_k, x_{k+1}) = F_{t_1,\cdots,t_k}(x_1, \cdots, x_k). \tag{8.2}$$

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:

[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.

(C2′) Sia $k \geq 1$, siano $t_1 < t_2 < \cdots < t_k < t_{k+1}$, allora

$$\lim_{x\to\infty} F_{t_1,\cdots,t_k,t_{k+1}}(x_1,\cdots,x_k,x) = F_{t_1,\cdots,t_k}(x_1,\cdots,x_k), \tag{8.3}$$

$$\begin{aligned}&\lim_{x\to\infty} F_{t_1,\cdots,t_{i-1},t_i,t_{i+1},\cdots,t_{k+1}}(x_1,\cdots,x_{i-1},x,x_{i+1},\cdots,x_{k+1})\\ &= F_{t_1,\cdots,t_{i-1},t_{i+1},\cdots,t_{k+1}}(x_1,\cdots,x_{i-1},x_{i+1},\cdots,x_{k+1}), \quad \text{per } i = 1,\cdots,k.\end{aligned} \tag{8.4}$$

Il seguente teorema garantisce che le precedenti condizioni necessarie, sono anche sufficienti.

Teorema 8.1 (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 (8.1) e (8.2)), 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 (8.1), (8.3) e (8.3).*

Vediamo subito delle applicazioni del precedente teorema, mentre per la dimostrazione rimandiamo all'Appendice 8.8.

Esempio 8.1 (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 8.2 (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*

$$\sum_{h,k}^{1,n} \eta_h \overline{\eta}_k K(t_h,t_k) \geq 0.$$

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].

---

[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.

[6]Si veda l'Appendice 2.5

Per la consistenza della famiglia $F_{t_1,\cdots,t_k}$ così definita, si noti che l'operatore Φ_π di (8.1) 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 (8.1), 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

$$f_{t_1,\cdots,t_k}(x_1,\cdots,x_k) = f_{t_{\pi_1},\cdots,t_{\pi_k}}(\Phi_\pi(x_1,\cdots,x_k)),$$

ovvero la (8.1). Per quanto riguarda la (8.2), basta ricordare che ogni sottovettore di un vettore gaussiano è ancora un vettore gaussiano.

Esempio 8.3 (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].
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

$$\begin{aligned}
Cov(N_t, N_s) &= Cov(N_{t\wedge s}, N_{t\vee s}) \\
&= \mathbb{E}[N_{t\wedge s}(N_{t\wedge s} + N_{t\vee s} - N_{t\wedge s})] - \mathbb{E}[N_{t\wedge s}]\mathbb{E}[N_{t\wedge s} + N_{t\vee s} - N_{t\wedge s}] \\
&= \mathbb{E}[N_{t\wedge s}N_{t\wedge s}] + \mathbb{E}[N_{t\wedge s}(N_{t\vee s} - N_{t\wedge s})] - \mathbb{E}[N_{t\wedge s}](\mathbb{E}[N_{t\wedge s}] + \mathbb{E}[N_{t\vee s} - N_{t\wedge s}]) \\
&= Var(N_{t\wedge s}) + \mathbb{E}[N_{t\wedge s}(N_{t\vee s} - N_{t\wedge s})] - \mathbb{E}[N_{t\wedge s}]\mathbb{E}[N_{t\vee s} - N_{t\wedge s}] = Var(N_{t\wedge s}) \\
&= t \wedge s
\end{aligned}$$

[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 $t_i - t_1 - (t_2 - t_1) = t_i - t_2$.

[8] Tuttavia questo metodo presuppone la conoscenza del processo di Poisson, che in queste note viene definito come processo ad incrementi indipendenti in Sezione 8.5.

[9] La matrice di covarianza di un vettore aleatorio è sempre definita non negativa, e di conseguenza la funzione di correlazione

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*

$$\Delta W_{t_h} = W_{t_h} - W_{t_{h-1}}.$$

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 $(W_{t_1}, \cdots, W_{t_n})$

$$f_{W_{t_1}, \cdots, W_{t_n}}(x_1, x_2, \cdots, x_n) = f_{\Delta W_{t_1}, \cdots, \Delta W_{t_n}}(x_1, x_2 - x_1 \cdots, x_n - x_{n-1}) \tag{8.5}$$

$$= \prod_{h=1}^{n} g_{t_h - t_{h-1}}(x_h - x_{h-1}) \tag{8.6}$$

$$= \frac{1}{\sqrt{2\pi t_1}} e^{-\frac{x_1^2}{2t_1}} \frac{1}{\sqrt{2\pi(t_2 - t_1)}} e^{-\frac{(x_2 - x_1)^2}{2(t_2 - t_1)}} \cdots \frac{1}{\sqrt{2\pi(t_n - t_{n-1})}} e^{-\frac{(x_n - x_{n-1})^2}{2(t_n - t_{n-1})}} \tag{8.7}$$

$K(s,t) = Cov(X_s, X_t)$ 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 $K(s,t)$.

8.2 Osservazione sulla definizione di un processo solo attraverso le sue distribuzioni finito dimensionali

La descrizione di un processo attraverso la famiglia delle distribuzioni finito dimensionali non è sufficiente a stabilire le proprietà delle sue traiettorie. In questa sezione e nella successiva ci si pone il problema di spiegare meglio questa affermazione, senza alcuna pretesa di essere esaurienti. Il primo concetto che ci serve è l'equivalenza stocastica per due processi aleatori, che, come si vedrà immediatamente dopo la definizione, implica che i due processi hanno le stesse distribuzioni finito dimensionali.

Definizione 8.5 (Equivalenza stocastica). *Due processi aleatori* $(X_t, t \in I)$ *e* $(Y_t, t \in I)$ *sullo stesso spazio di probabilità* $(\Omega, \mathcal{F}, \mathbb{P})$ *si dicono* **stocasticamente equivalenti** *se*

$$\mathbb{P}(X_t(w) = Y_t(\omega)) = 1 \text{ per ogni } t \in I.$$

(Si dice anche che $(Y_t, t \in I)$ *è una* **versione** *di* $(X_t, t \in I)$*)*

Lemma 8.2. *Due processi stocasticamente equivalenti hanno le stesse distribuzioni finito-dimensionali.*

Dimostrazione. Basta osservare che

$$\begin{aligned}
&\mathbb{P}\big(\{(X_{t_1}, ..., X_{t_k}) \in H\}\big) \\
&= \mathbb{P}\Big(\{(X_{t_1}, ..., X_{t_k}) \in H\} \cap \bigcap_{i=1}^{k}\{X_{t_i} = Y_{t_i}\}\Big) + \mathbb{P}\Big(\{(X_{t_1}, ..., X_{t_k}) \in H\} \cap \Big(\bigcap_{i=1}^{k}\{X_{t_i} = Y_{t_i}\}\Big)^c\Big) \\
&= \mathbb{P}\Big(\{(Y_{t_1}, ..., Y_{t_k}) \in H\} \cap \bigcap_{i=1}^{k}\{X_{t_i} = Y_{t_i}\}\Big) = \mathbb{P}\Big(\{(Y_{t_1}, ..., Y_{t_k}) \in H\}\Big)
\end{aligned}$$

in quanto

$$\mathbb{P}\Big(\Big(\bigcap_{i=1}^{k}\{X_{t_i} = Y_{t_i}\}\Big)^c\Big) = \mathbb{P}\Big(\bigcup_{i=1}^{k}\{X_{t_i} = Y_{t_i}\}^c\Big) \le \sum_{i=1}^{k}\mathbb{P}\big(\{X_{t_i} = Y_{t_i}\}^c\big) = 0.$$

□

Esempio 8.4. *Come primo esempio di processo si definisca, in uno spazio di probabilità* $(\Omega, \mathcal{F}, \mathbb{P})$,

$$N_t(\omega) := \mathbb{I}_{\{T \le t\}}(\omega) = \mathbb{I}_{[T(\omega),\infty)}(t) \begin{cases} = 0 & \text{se } T(\omega) > t, \\ = 1 & \text{se } T(\omega) \le t \end{cases}, \quad t \ge 0,$$

dove T *è una variabile aleatoria a valori in* $(0, \infty)$.
Si noti che, qualunque sia $\omega \in \Omega$ *le traiettorie di questo processo, cioè le funzioni*

$$t \mapsto \mathbb{I}_{\{T \le t\}}(\omega)$$

sono funzioni crescenti (in senso lato) rispetto a t.

Se T *ammette densità di probabilità, ad esempio se è una variabile aleatoria esponenziale, allora il processo definito da*

$$M_t(\omega) := N_t(\omega) + f(t + T(\omega)),$$

dove

$$\begin{cases} f(s) = 1 & \text{per } s \in \mathbb{Q} \\ f(s) = 0 & \text{per } s \in \mathbb{R}\backslash\mathbb{Q}, \end{cases}$$

è stocasticamente equivalente ad N_t*, e quindi ha le stesse distribuzioni finito-dimensionali di* N_t*, ma evidentemente non ha traiettorie crescenti (in senso lato).*

Per controllare che $M_t(\omega)$ ed $N_t(\omega)$ sono stocasticamente equivalenti, si osservi che, $M_t(\omega) \neq N_t(\omega)$ se e solo se $f(t+T(\omega)) = 1$ ovvero se e solo se $t + T(\omega) \in \mathbb{Q}$ ed inoltre

$$\mathbb{P}(t + T(\omega) \in \mathbb{Q}) = \sum_{r\in\mathbb{Q}} \mathbb{P}(T(\omega) = r - t) = 0$$

per ogni t, in quanto T è una variabile aleatoria continua e si ha quindi che

$$\mathbb{P}(M_t(\omega) = N_t(\omega)) = 1 \textit{ per ogni } t \geq 0.$$

8.3 Esistenza di una versione continua: criterio di Chensov-Kolmogorov.

Il seguente criterio sufficiente fornisce condizioni che assicurano l'esistenza di una versione hölderiana, non solo continua. Ricordiamo che una funzione $f(x)$ è **hölderiana** di esponente γ se per ogni x esistono un $\delta(x) > 0$ e un $L_\gamma(x)$ tali che, per ogni y per il quale $|y - x| \leq \delta(x)$, si abbia

$$|f(x) - f(y)| \leq L_\gamma(x)|x - y|^\gamma.$$

Nel caso in cui $\delta(x)$ e $L_\gamma(x)$ possano essere presi in modo indipendente da x, per $x \in I$, si dice che f è **uniformemente hölderiana** nell'insieme I.

Proposizione 8.3 (Criterio di Chensov-Kolmogorov). *Sia X_t un processo per cui esistono α, β e C strettamente positivi, per cui*

$$\mathbb{E}[|X_t - X_s|^\beta] \leq C|t - s|^{1+\alpha}.$$

Allora esiste una versione $\widetilde{X}_t$ di X_t, a traiettorie continue. Inoltre le traiettorie sono uniformemente hölderiane di esponente γ, per ogni $\gamma < \frac{\alpha}{\beta}$, in ogni intervallo limitato.

Esempio 8.5 (Applicazione al processo di Wiener).

Il processo W_t di Wiener standard, o moto Browniano, ammette sempre una versione continua, anzi hölderiana per ogni $\gamma < \frac{1}{2}$. Infatti, per $s \leq t$,

$$\mathbb{E}[|W_t - W_s|^k] = \mathbb{E}[|W_{t-s}|^k] = C(k)|t - s|^{k/2}.$$

Il criterio di Chensov-Kolmogorov si può quindi applicare per $k \geq 3$ (così $\alpha = (k/2) - 1 > 0$) e ci garantisce che esiste una versione di W_t le cui traiettorie sono hölderiane di esponente γ per $\gamma < \frac{(k/2)-1}{k} = \frac{1}{2} - \frac{1}{k}$, e quindi per ogni $\gamma < \frac{1}{2}$. È possibile dimostrare che, a parte un eventuale insieme di misura nulla, le traiettorie non sono hölderiane di esponente $\frac{1}{2}$, e tantomeno di esponente maggiore di $\frac{1}{2}$.

8.4 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[11]

[11] 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

$$\sup_{u,v\in[s,t],|u-v|\leq\delta} |W_u(\omega) - W_v(\omega)| \to 0 \qquad \text{per } \delta \to 0.$$

di W_u, se esistesse un intervallo $[s,t]$ ed un insieme A con $\mathbb{P}(A) > 0$, su cui la variazione[12] 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 (valida anche per versioni non continue):

Lemma 8.4. *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 $|\pi| := \max_k |t_{k+1} - t_k|$.

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

$$\begin{aligned}
&\mathbb{E}[(S_\pi - (t-s))^2] = \mathbb{E}[S_\pi^2 + (t-s))^2 - 2S_\pi(t-s)]\\
&= \mathbb{E}\left[\Big(\sum_{k=0}^{n-1}|W_{t_{k+1}} - W_{t_k}|^2\Big)^2 + (t-s)^2 - 2\Big(\sum_{k=0}^{n-1}|W_{t_{k+1}} - W_{t_k}|^2\Big)(t-s)\right]\\
&= \mathbb{E}\left[\sum_{k=0}^{n-1}|W_{t_{k+1}} - W_{t_k}|^2\sum_{h=0}^{n-1}|W_{t_{h+1}} - W_{t_h}|^2 + (t-s)^2 - 2\Big(\sum_{k=0}^{n-1}|W_{t_{k+1}} - W_{t_k}|^2\Big)(t-s)\right]\\
&= \sum_{k=0}^{n-1}\mathbb{E}\left[|W_{t_{k+1}} - W_{t_k}|^4\right] + \sum_{k=0}^{n-1}\sum_{\substack{h=0\\h\ne k}}^{n-1}\mathbb{E}\left[|W_{t_{k+1}} - W_{t_k}|^2|W_{t_{h+1}} - W_{t_h}|^2\right] +\\
&\qquad + (t-s)^2 - 2(t-s)\sum_{k=0}^{n-1}\mathbb{E}\left[|W_{t_{k+1}} - W_{t_k}|^2\right]\\
&= \sum_{k=0}^{n-1}3|t_{k+1} - t_k|^2 + \sum_{k=0}^{n-1}\sum_{\substack{h=0\\h\ne k}}^{n-1}|t_{k+1} - t_k||t_{h+1} - t_h| + (t-s)^2 - 2(t-s)\sum_{k=0}^{n-1}|t_{k+1} - t_k|
\end{aligned}$$

[12]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 $x_0 = a \le x_1 \le \cdots \le x_n = b$.

Di conseguenza si ottiene che

$$
\begin{aligned}
&\mathbb{E}\left[(S_\pi - (t-s))^2\right] \\
&= \sum_{k=0}^{n-1} 2|t_{k+1}-t_k|^2 + \sum_{k=0}^{n-1}\sum_{h=0}^{n-1} |t_{k+1}-t_k||t_{h+1}-t_h| + (t-s)^2 - 2(t-s)(t-s) \\
&\leq 2\sum_{k=0}^{n-1} |t_{k+1}-t_k|(\max_h |t_{h+1}-t_h|) + (t-s)^2 + (t-s)^2 - 2(t-s)(t-s) \\
&= 2(\max_h |t_{h+1}-t_h|)\sum_{k=0}^{n-1} |t_{k+1}-t_k| = 2(|\pi|(t-s) \to 0.
\end{aligned}
$$

La conseguenza di questo risultato è che le traiettorie di un processo di Wiener non sono a variazione limitata[13], o meglio,

$$
\mathbb{P}\left(V(s,t,W;\omega) := \sup_\pi \sum_{k=0}^{n-1} |W_{t_{k+1}} - W_{t_k}| = \infty \right) = 1.
$$

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 (vedere la Sezione 9.5).

[13]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 necessariamente si avrebbe

$$
S_\pi(\omega) \leq \sum_{k=0}^{n-1} L^2(\omega)|t_{k+1}-t_k|^{2\gamma} \leq \sum_{k=0}^{n-1} L^2(\omega)|t_{k+1}-t_k|\delta^{2\gamma-1} = L^2(\omega)\delta^{2\gamma-1}(t-s) \to 0, \; q.c.
$$

8.5 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 8.6 (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[14]:

$$F_u * F_v = F_{u+v}, \quad \text{per ogni } u, v \geq 0, \tag{8.8}$$

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. Infatti le funzioni $F_{t_1,\cdots,t_k}(x_1, \cdots, x_k)$ di distribuzione finito dimensionale risultano definite[15], per $0 < t_1 < \cdots < t_k$, come, la funzione di distribuzione ottenuta da quella degli incrementi, cioè

$$\mathbb{P}\left(X_{t_1} - X_0 \leq z_1,\, X_{t_2} - X_{t_1} \leq z_2, \cdots X_{t_k} - X_{t_{k-1}} \leq z_k\right) = F_{t_1}(z_1) F_{t_2 - t_1}(z_2) \cdots F_{t_k - t_{k-1}}(z_k),$$

attraverso la trasformazione[16] $x_1 = z_1$, $x_2 = z_1 + z_2, \cdots, x_k = z_1 + \cdots + z_k$. La condizione (8.8) 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 $0 < t_1 < \cdots < t_k$

$$F_{t_1,\cdots,t_k}(x_1, \cdots, x_k) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_k} q_{t_1}(y_1) \cdots q_{t_k - t_{k-1}}(y_k - y_{k-1}) dy_1 \cdots dy_k.$$

[14] 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.

[15] Nel caso $0 = t_0 < t_1 < \cdots < t_k$ si ottiene immediatamente che

$$F_{t_0,t_1,\cdots,t_k}(x_0, x_1, \cdots, x_k) = 0, \qquad \text{per } x_0 < 0,$$

$$F_{t_0,t_1,\cdots,t_k}(x_0, x_1, \cdots, x_k) = F_{t_1,\cdots,t_k}(x_1, \cdots, x_k), \qquad \text{per } x_0 \geq 0.$$

[16] Si tratta solo di notare che se $Z_i := X_{t_i} - X_{t_{i-1}}$ allora $(X_{t_1}, X_{t_2}, \cdots, X_{t_k}) = (Z_1, Z_1 + Z_2, \cdots, Z_1 + \cdots + Z_k)$

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 8.3) 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 (8.4) 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

$$F_u \sim N(\mu u, \sigma^2 u);$$

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 8.7 (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 $\mathcal{F}_t = \mathcal{F}_t^X = \sigma\{X_u;\ u \in [0,t]\} = \sigma\{X_u - X_v;\ u, v \in [0,t]\}$.

8.6 Esempi di martingale a tempo continuo

In modo molto simile a quanto fatto a tempo discreto per le somme di v.a. indipendenti, si può mostrare[17] che se X_t è un processo ad incrementi indipendenti e omogenei, rispetto ad una filtrazione $(\mathcal{F}_t)_{t\geq 0}$, con $X_0 = 0$, e con media nulla allora X_t è una martingala, purché sia integrabile.

Inoltre è facile mostrare che se X_t è un processo ad incrementi indipendenti (e omogenei), integrabile e con $X_0 = 0$, allora $\mathbb{E}[X_t] = \mu t$ 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

[17] La condizione di misurabilità dipende dal fatto che $\mathcal{F}_t \supseteq \mathcal{F}_t^X$. La condizione di integrabilità è verificata per ipotesi. Infine

$$\mathbb{E}[X_{t+s} - X_t|\mathcal{F}_t] = \mathbb{E}[X_{t+s} - X_t] = \mathbb{E}[X_{t+s}] - \mathbb{E}[X_t] = 0 - 0 = 0,$$

dove nella prima uguaglianza si usa l'indipendenza di $X_{t+s} - X_t$ da $\mathcal{F}_t$, e nell'ultima si usa il fatto che la media di X_u è nulla. Si noti che l'omogeneità degli incrementi qui non è necessaria.

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[18] 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[19] che, sotto opportune ipotesi di regolarità (continuità in probabilità), se $\mathbb{E}[\exp\{\theta(X_t - \mu t)\}] < +\infty$, allora

$$\mathbb{E}[\exp\{\theta(X_t - \mu t)\}] = \exp\{K(\theta)t\}$$

e che quindi

$$Z_t := \exp\{\theta(X_t - \mu t) - K(\theta)t\} \tag{8.9}$$

è una martingala[20] 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 $\{X_t\}$ (tranne per i valori medi). Bisogna però che Y_0 sia $\mathcal{F}_0$-misurabile[21], e soddisfi alcuni requisiti di

---

[18]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 $\mathbb{E}$.

[19]In questo caso, posto $X'_t = X_t - \mu t$ 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.

[20]Di nuovo misurabilità e integrabilità sono banali. Osservando che

$$\begin{aligned} Z_{t+s} &= \exp\{\theta(X_{t+s} - \mu(t+s)) - K(\theta)(t+s)\} \\ &= \exp\{\theta(X_t - \mu t) - K(\theta)t\} \exp\{\theta(X_{t+s} - X_t - \mu s) - K(\theta)s\} \\ &= Z_t \exp\{\theta(X_{t+s} - X_t - \mu s) - K(\theta)s\} \end{aligned}$$

si ottiene subito

$$\begin{aligned} \mathbb{E}[Z_{t+s} - Z_t | \mathcal{F}_t] &= \mathbb{E}\left[Z_t \left(\exp\{\theta(X_{t+s} - X_t - \mu s) - K(\theta)s\} - 1\right) | \mathcal{F}_t\right] \\ &\qquad \text{(per la misurabilità di } Z_t) \\ &= Z_t \mathbb{E}\left[\left(\exp\{\theta(X_{t+s} - X_t - \mu s) - K(\theta)s\} - 1\right) | \mathcal{F}_t\right] \\ &\qquad \text{(per l'indipendenza degli incrementi da } \mathcal{F}_t) \\ &= Z_t \mathbb{E}\left[\left(\exp\{\theta(X_{t+s} - X_t - \mu s) - K(\theta)s\} - 1\right)\right] \\ &\qquad \text{(per l'omogeneità degli incrementi)} \\ &= Z_t \mathbb{E}\left[\left(\exp\{\theta(X_s - \mu s) - K(\theta)s\} - 1\right)\right] = Z_t (1 - 1) = 0 \end{aligned}$$

[21]† 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 $\{X_t\}$.

integrabilità. Ad esempio

$$\widetilde{Z}_t := \mathbb{E}[\exp\{\theta(Y_t - \mu t) - K(\theta)t\}] \tag{8.10}$$

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 8.6 (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*

$$Z_t^\theta := \exp\{\theta W_t - \frac{1}{2}\theta^2 t\} \tag{8.11}$$

è una martingala[22] *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 8.7 (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 $(\Omega, \mathcal{F}, \mathbb{P})$,

$$\mathbb{Q}(C) := \mathbb{E}^{\mathbb{P}}[I_C Z], \qquad C \in \mathcal{F}, \tag{8.12}$$

definisce una nuova misura di probabilità.[23] $\mathbb{Q}$ su $(\Omega, \mathcal{F})$.

Esercizio 8.1 (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 (8.11) relativa al processo di Wiener, e $\mathcal{F} = \mathcal{F}_T$ nell'espressione precedente (8.12), si trovi

1) la derivata di Radon-Nikodym

$$h_t = \frac{d\mathbb{Q}}{d\mathbb{P}}\bigg|_{\mathcal{F}_t},$$

ovvero[24] *quella variabile aleatoria $h_t(\omega)$, $\mathcal{F}_t$-misurabile, tale che*

$$\mathbb{Q}(A) = \int_A h_t(\omega)\mathbb{P}(d\omega), \quad \textit{per ogni } A \in \mathcal{F}_t;$$

suggerimento: *si veda l'esempio riguardante le martingale e le derivate di Radon-Nikodym.*

[22] 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 $K(\theta) = \frac{1}{2}\theta^2$.

[23] 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

$$I_{\{\cup_n A_n\}} = \sum_n I_{A_n}, \qquad \text{se gli insiemi } A_n \text{ sono disgiunti a due a due.}$$

[24] Si ricordi la nota 3.3.

soluzione: $\frac{d\mathbb{Q}}{d\mathbb{P}}\Big|_{\mathcal{F}_t} = Z_t^\theta$

2) la legge di W_t rispetto a $\mathbb{Q}$,

suggerimento: *basta capire che $\mathbb{E}^{\mathbb{Q}}[g(W_t)]$ si calcola equivalentemente come*

$$\begin{aligned}
&\mathbb{E}^{\mathbb{Q}}[g(W_t)] = \mathbb{E}^{\mathbb{P}}[g(W_t)Z_t^\theta] = \mathbb{E}^{\mathbb{P}}[g(W_t)\exp\{\theta W_t - \frac{1}{2}\theta^2 t\}] \\
&= \int_{\mathbb{R}} g(w)\exp\{\theta w - \frac{1}{2}\theta^2 t\}\exp\{-\frac{1}{2t}w^2\}dw = \int_{\mathbb{R}} g(w)\exp\{-\frac{1}{2t}(w^2 - 2w\theta t + \theta^2 t^2)\}dw \\
&= \int_{\mathbb{R}} g(w)\exp\{-\frac{1}{2t}(w-\theta t)^2\}dw
\end{aligned}$$

soluzione: *la legge di W_t è $N(\theta t, t)$*

3) le distribuzioni finito dimensionali di $(W_t, t \geq 0)$ rispetto a $\mathbb{Q}$.

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*

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}}[g_1(W_{t_1})g_2(W_{t_2} - W_{t_1})\cdots g_n(W_{t_n} - W_{t_n-1})Z_{t_n}^\theta] \\
&= \mathbb{E}^{\mathbb{P}}[g_1(W_{t_1})g_2(W_{t_2} - W_{t_1})\cdots g_n(W_{t_n} - W_{t_n-1})\prod_{k=1}^n \frac{Z_{t_k}^\theta}{Z_{t_{k-1}}^\theta}] \\
&= \mathbb{E}^{\mathbb{P}}[g_1(W_{t_1})g_2(W_{t_2} - W_{t_1})\cdots g_n(W_{t_n} - W_{t_{n-1}})\prod_{k=1}^n \exp\{\theta(W_{t_k} - W_{t_{k-1}}) - \frac{1}{2}\theta^2(t_k - t_{k-1})\}] \\
&= \mathbb{E}^{\mathbb{P}}[\prod_{k=1}^n g_k(W_{t_k} - W_{t_{k-1}})\exp\{\theta(W_{t_k} - W_{t_{k-1}}) - \frac{1}{2}\theta^2(t_k - t_{k-1})\}] \\
&= \prod_{k=1}^n \mathbb{E}^{\mathbb{P}}[g_k(W_{t_k} - W_{t_{k-1}})\exp\{\theta(W_{t_k} - W_{t_{k-1}}) - \frac{1}{2}\theta^2(t_k - t_{k-1})\}] \\
&= \prod_{i=1}^n \int_{\mathbb{R}} g(w_i - w_{i-1})\exp\{-\frac{(w_i - w_{i-1} - \theta(t_i - t_{i-1}))^2}{2(t_i - t_{i-1})}\}dw_i.
\end{aligned}$$

dove per convenzione si è posto $t_0 = 0$ e $w_0 = 0$.

soluzione: *il processo $\{W_t\}$ diviene sotto $\mathbb{Q}$ un processo di Wiener con drift θ e coefficiente di diffusione 1.*

8.7 Processi di Markov regolari

I processi di Markov regolari sono processi costruiti attraverso una ***famiglia di probabilità di transizione regolari*** $P(s,t,x,A)$, e attraverso una misura di probabilità μ_0, detta ***(misura delle) probabilità iniziali***.

Definizione 8.8 (Probabilità di transizione regolari). *Una famiglia $P(s,t,x,A)$*

$$P : (\mathbb{R}^+ \times \mathbb{R}^+)_+ \times \mathbb{R} \times \mathcal{B}(\mathbb{R}) \to [0,1], (s,t,x,A) \to P(s,t,x,A),$$

dove

$$(\mathbb{R}^+ \times \mathbb{R}^+)_+ = \{(s,t) \in \mathbb{R}^+ \times \mathbb{R}^+, \text{ tali che } s \leq t\},$$

che soddisfa le seguenti proprietà:

(i) *per ogni* $0 \leq s \leq t$ *ed* $x \in \mathbb{R}$,

$$P(s,t,x,\cdot) : \mathcal{B}(\mathbb{R}) \to [0,1], A \to P(s,t,x,A)$$

è una misura di probabilità,

(ii) *per ogni* $0 \leq s \leq t$ *ed* $A \in \mathcal{B}(\mathbb{R})$,

$$P(s,t,\cdot,A) : \mathbb{R} \to [0,1], x \to P(s,t,x,A)$$

è una funzione misurabile,

(iii) *la famiglia* $P(\cdot,\cdot,\cdot,\cdot)$ *soddisfa l'***equazione di Chapman-Kolmogorov***, cioè per ogni* $0 \leq r \leq s \leq t$*, per ogni* $x \in \mathbb{R}$ *e per ogni* $A \in \mathcal{B}(\mathbb{R})$ *vale*

$$P(r,t,x,A) = \int_{\mathbb{R}} P(r,s,x,dy)P(s,t,y,A),$$

viene detta **famiglia di probabilità di transizione regolari**.
Si noti che quindi necessariamente $P(t,t,x,A) = \delta_x(A)$*, cioè vale* 1 *se* $x \in A$ *e vale* 0 *altrimenti e che le proprietà (i) e (ii) permettono di dare senso all'integrale in (iii).*

Osservazione 8.1. *E' interessante notare che l'equazione di Chapman-Kolmogorov, nel caso in cui* $P(s,t,x,\cdot)$ *ammette densità, ovvero*

$$P(s,t,x,A) = \int_A p(s,t,x,y)\, dy,$$

diviene

$$\int_A p(r,t,x,z)\, dz = \int_{\mathbb{R}} p(r,s,x,y)\, dy \int_A p(s,t,y,z)\, dz = \int_A \left(\int_{\mathbb{R}} p(r,s,x,y)p(s,t,y,z)\, dy \right) dz,$$

ovvero

$$p(r,t,x,z) = \int_{\mathbb{R}} p(r,s,x,y)p(s,t,y,z)\, dy.$$

Attraverso le probabilità di transizioni regolari e la probabilità iniziale, si può definire una famiglia di funzioni di distribuzione finito dimensionale nel seguente modo: si definiscano, per $0 \leq t_1 \leq \cdots \leq t_k$, e per $(z_0, z_1, \cdots, z_k) \in \mathbb{R}^{k+1}$,

$$\begin{aligned} F_{0,t_1,\cdots,t_k}(z_0,z_1,\cdots,z_k) \quad := \quad & \int_{-\infty}^{z_0} \int_{-\infty}^{z_1} \cdots \int_{-\infty}^{z_k} \mu_0(dx_0)P(0,t_1,x_0,dy_1)P(t_1,t_2,y_1,dy_2) \cdots \\ & \cdots P(t_{k-2},t_{k-1},y_{k-2},dy_{k-1})P(t_{k-1},t_k,y_{k-1},dy_k), \end{aligned}$$

ed

$$F_{t_1,\cdots,t_k}(z_1,\cdots,z_k) := \lim_{z\to\infty} F_{0,t_1,\cdots,t_k}(z,z_1,\cdots,z_k),$$

Se invece non vale $0 \leq t_1 \leq \cdots \leq t_k$, si definiscono attraverso un'opportuna permutazione π per la quale $0 \leq t_{\pi_1} \leq \cdots \leq t_{\pi_k}$ ed in modo che valga la condizione di consistenza **(C1)**.

La proprietà (iii), ovvero l'equazione di Chapman-Kolmogorov, permette di verificare immediatamente la condizione di consistenza **(C2′)**[25]. La famiglia di distribuzioni finito-dimensionali così ottenute soddisfa quindi le proprietà di consistenza del teorema di Kolmogorov, e pertanto esiste un processo con queste distribuzioni finito-dimensionali. Un processo le cui distribuzioni finito-dimensionali si possono ottenere attraverso una famiglia di probabilità di transizione regolare e una probabilità iniziale come sopra, viene detto ***processo di Markov (o processo markoviano) regolare***.

[25] Nel caso con densità la verifica è simile a quella del caso dei processi ad incrementi indipendenti ed omogenei: si prendano

Esempio 8.8. *Il processo di Wiener ed il processo di Poisson sono processi markoviani regolari in questo senso con* $\mu_0=\delta_{\{0\}}$ *per entrambi i processi e*

$$P(s,t,x,dy) = \frac{1}{\sqrt{2\pi(t-s)}}\exp\{-\frac{(y-x)^2}{2(t-s)}\}dy$$

per il processo di Wiener, mentre

$$P(s,t,n,\{m\}) = \frac{(\lambda(t-s))^{n-m}}{(n-m)!}\exp\{-\lambda(t-s)\},\ 0\le m\le n$$

per il processo di Poisson.

Esempio 8.9 (Processi di Markov regolari omogenei e processi a incrementi indipendenti e omogenei). *In entrambi i casi dell'esempio precedente* $P(s,t,x,A)$ *dipende solo dalla differenza* $t-s$. *Più in generale ogni volta che* $P(s,t,x,A) = P(s+h,t+h,x,A) = P(0,t-s,x,A)$ *per ogni* s,t,h,x,A, *si parla di* **processo di Markov omogeneo (nel tempo)**. *Inoltre se per ogni* s,t,h,x,z,y,

$$P(s,t,x,(-\infty,z]) = P(s+h,t+h,x+y,(-\infty,z+y]), \tag{8.13}$$

allora in realtà si tratta di esempi di **processi a incrementi indipendenti ed omogenei (nello spazio)** *(con* $X_0 = 0$, *se* μ_0 *è la misura concentrata in* 0*). Nel caso generale la famiglia delle distribuzioni finito-dimensionali così ottenuta coincide con la famiglia delle distribuzioni finito-dimensionali di un processo* $(X'_t)_{t\ge 0}$, *con*

$$X'_t = Y_0 + X_t,$$

$k\ge 1$ e $0<t_1<t_2<\cdots,<t_k<t_{k+1}$, allora

$$\begin{aligned}
&\lim_{x\to\infty} F_{0,t_1,\cdots,t_{i-1},t_i,t_{i+1},\cdots,t_{k+1}}(x_0,x_1,\cdots,x_{i-1},x,x_{i+1},\cdots,x_{k+1})\\
&= \lim_{x\to\infty}\int_{-\infty}^{x_0}\int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_{i-1}}\int_{-\infty}^{x}\int_{-\infty}^{x_{i+1}}\cdots\int_{-\infty}^{x_{k+1}}\mu_0(dy_0)p(0,t_1,y_0,y_1)\cdots p(t_{i-2},t_{i-1},y_{i-2},y_{i-1})\\
&\qquad p(t_{i-1},t_i,y_{i-1},y_i)p(t_i,t_{i+1},y_i,y_{i+1})\cdots p(t_k,t_{k+1},y_k,y_{k+1})dy_1\cdots dy_{i-1}\,dy_i\,dy_{i+1}\cdots dy_{k+1}\\
&= \lim_{x\to\infty}\int_{-\infty}^{x_0}\mu_0(dy_0)\int_{-\infty}^{x_1}p(0,t_1,y_0,y_1)\,dy_1\cdots\int_{-\infty}^{x_{i-1}}p(t_{i-2},t_{i-1},y_{i-2},y_{i-1})\,dy_{i-1}\\
&\qquad\int_{-\infty}^{x_{i+1}}\left(\int_{-\infty}^{x}p(t_{i-1},t_i,y_{i-1},y_i)\,p(t_i,t_{i+1},y_i,y_{i+1})\,dy_i\right)\,dy_{i+1}\cdots\int_{-\infty}^{x_{k+1}}p(t_k,t_{k+1},y_k,y_{k+1})\,dy_{k+1} =\\
&= \int_{-\infty}^{x_0}\mu_0(dy_0)\int_{-\infty}^{x_1}p(0,t_1,y_0,y_1)\,dy_1\cdots\int_{-\infty}^{x_{i-1}}p(t_{i-2},t_{i-1},y_{i-2},y_{i-1})\,dy_{i-1}\\
&\qquad\int_{-\infty}^{x_{i+1}}\left(\int_{-\infty}^{\infty}p(t_{i-1},t_i,y_{i-1},y_i)\,p(t_i,t_{i+1},y_i,y_{i+1})\,dy_i\right)\,dy_{i+1}\cdots\int_{-\infty}^{x_{k+1}}p(t_k,t_{k+1},y_k,y_{k+1})\,dy_{k+1}\\
&= \int_{-\infty}^{x_0}\mu_0(dy_0)\int_{-\infty}^{x_1}p(0,t_1,y_0,y_1)\,dy_1\cdots\int_{-\infty}^{x_{i-1}}p(t_{i-2},t_{i-1},y_{i-2},y_{i-1})\,dy_{i-1}\\
&\qquad\int_{-\infty}^{x_{i+1}}p(t_{i-1},t_{i+1},y_{i-1},y_{i+1})\,dy_{i+1}\cdots\int_{-\infty}^{x_{k+1}}p(t_k,t_{k+1},y_k,y_{k+1})\,dy_{k+1}\\
&= F_{0,t_1,\cdots,t_{i-1},t_{i+1},\cdots,t_{k+1}}(x_0,x_1,\cdots,x_{i-1},x_{i+1},\cdots,x_{k+1}),\quad \text{per } i=1,\cdots,k.
\end{aligned}$$

Nella penultima uguaglianza si è tenuto conto dell'equazione di Chapman-Kolmogorov. Il caso $i=0$ è verificato per definizione. Infine

$$\begin{aligned}
&\lim_{x\to\infty} F_{0,t_1,\cdots,t_k,t_{k+1}}(x_0,x_1,\cdots,x_k,x)\\
&= \lim_{x\to\infty}\int_{-\infty}^{x_0}\int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_k}\int_{-\infty}^{x}\mu_0(dy_0)p(0,t_1,y_0,y_1)\cdots p(t_{k-1},t_k,y_{k-1},y_k)p(t_k,t_{k+1},y_k,y)\,dy_0\,dy_1\cdots dy_k\,dy\\
&= \lim_{x\to\infty}\int_{-\infty}^{x_0}\mu_0(dy_0)\int_{-\infty}^{x_1}p(0,t_1,y_0,y_1)\,dy_1\cdots\int_{-\infty}^{x_k}p(t_{k-1},t_k,y_{k-1},y_k)\,dy_k\int_{-\infty}^{x}\cdots p(t_k,t_{k+1},y_k,y)\,dy\\
&= \int_{-\infty}^{x_0}\mu_0(dy_0)\int_{-\infty}^{x_1}p(0,t_1,y_0,y_1)\,dy_1\cdots\int_{-\infty}^{x_k}p(t_{k-1},t_k,y_{k-1},y_k)\,dy_k\int_{-\infty}^{\infty}\cdots p(t_k,t_{k+1},y_k,y)\,dy\\
&= \int_{-\infty}^{x_0}\mu_0(dy_0)\int_{-\infty}^{x_1}p(0,t_1,y_0,y_1)\,dy_1\cdots\int_{-\infty}^{x_k}p(t_{k-1},t_k,y_{k-1},y_k)\,dy_k = F_{0,t_1,\cdots,t_k}(x_0,x_1,\cdots,x_k),
\end{aligned}$$

dove $(X_t)_{t\geq 0}$ è un processo ad incrementi indipendenti ed omogenei con $X_0 = 0$, ed Y_0 è una variabile aleatoria con funzione di distribuzione $F_0(y) = \mu_0((-\infty, y])$, indipendente dal processo $(X_t)_{t\geq 0}$.

Per verificare le precedenti affermazioni si osservi che, prendendo $h = -s$ ed $y = -x$ nella formula (8.13) si ottiene

$$P(s,t,x,(-\infty,z]) = P(0,t-s,0,(-\infty,z-x]),$$

allora basta porre

$$F_u(z) = P(0,u,0,(-\infty,z]).$$

Sempre nel caso con densità, ovvero nel caso in cui $F_u(z) = \int_{-\infty}^{z} q_u(y)dy$, l'equazione di Chapman-Kolmogorov diviene

$$q_{t-r}(z-x) = \int_{\mathbb{R}} q_{s-r}(y-x)q_{t-s}(z-y)\,dy$$

ovvero, ponendo $s - r = u$, $t - s = v$, $\zeta = z - x$ ed $\eta = y - x$,

$$q_{u+v}(\zeta) = \int_{\mathbb{R}} q_u(\eta)q_v(\zeta-\eta)\,d\eta, \quad \textit{ovvero} \quad q_{u+v} = q_u * q_v,$$

che è esattamente la condizione di compatibilità già incontrata per i processi ad incrementi indipendenti ed omogenei, e ciò dimostra che la famiglia markoviana di distribuzioni finito-dimensionali definita con $\mu_0 = \delta_0$, e quella definita per i processi ad incrementi indipendenti ed omogenei, sono le stesse, ovvero che $(X_t)_{t\geq 0}$ è un processo di quest'ultimo tipo.

Infine per mostrare che le distribuzioni finito-dimensionali del processo $(X'_t)_{t\geq 0}$ sono quelle della famiglia markoviana, osserviamo che, nel caso in cui $\mu_0(dy) = p_0(y)dy$, si ottiene

$$\begin{aligned}
&\mathbb{P}(X'_0 \leq z_0, X'_{t_1} \leq z_1, \cdots, X'_{t_k} \leq z_k)\\
&= \mathbb{P}(Y_0 \leq z_0, Y_0 + X_{t_1} \leq z_1, \cdots, Y_0 + X_{t_k} \leq z_k)\\
&= \int_{-\infty}^{z_0} \mathbb{P}(Y_0 \in dx_0)\mathbb{P}(Y_0 + X_{t_1} \leq z_1, \cdots, Y_0 + X_{t_k} \leq z_k | Y_0 = x_0)\\
&= \int_{-\infty}^{z_0} \mathbb{P}(Y_0 \in dx_0)\mathbb{P}(X_{t_1} \leq z_1 - x_0, \cdots, X_{t_k} \leq z_k - x_0 | Y_0 = x_0)\\
&\qquad\qquad\qquad\qquad\qquad\qquad \textit{(per l'indipendenza di } (X_t)_{t\geq 0} \textit{ ed } Y_0\textit{)}\\
&= \int_{-\infty}^{z_0} \mathbb{P}(Y_0 \in dx_0)\mathbb{P}(X_{t_1} \leq z_1 - x_0, \cdots, X_{t_k} \leq z_k - x_0)\\
&= \int_{-\infty}^{z_0} p_{Y_0}(x_0)dx_0 \int_{-\infty}^{z_1-x_0} \cdots \int_{-\infty}^{z_k-x_0} p_{X_{t_1},\cdots,X_{t_k}}(y'_1,\cdots,y'_k)\,dy'_1\cdots,\,dy'_k\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \textit{(posto } y_i = y'_i + x_0\textit{)}\\
&= \int_{-\infty}^{z_0} p_{Y_0}(x_0)dx_0 \int_{-\infty}^{z_1} \cdots \int_{-\infty}^{z_k} p_{X_{t_1},\cdots,X_{t_k}}(y_1 - x_0,\cdots,y_k - x_0)\,dy_1\cdots,\,dy_k\\
&= \int_{-\infty}^{z_0}\int_{-\infty}^{z_1} \cdots \int_{-\infty}^{z_k} p_0(x_0)\,dx_0\,q_{t_1}(y_1-x_0)\,dy_1\,q_{t_2-t_1}(y_2-y_1)\,dy_2\cdots\\
&\qquad\qquad \cdots q_{t_{k-1},t_{k-2}}(y_{k-1}-y_{k-2})\,dy_{k-1}\,q_{t_k-t_{k-1}}(y_k-y_{k-1})\,dy_k,\\
&=: F_{0,t_1,\cdots,t_k}(z_0,z_1,\cdots,z_k).
\end{aligned}$$

8.7.1 Processo di Orstein-Ulhenbeck

Si consideri la famiglia delle probabilità di transizione con densità, definita da

$$P(s,t,x,dy) = \frac{1}{\sqrt{2\pi\sigma^2\frac{1-\mathrm{e}^{-2\lambda(t-s)}}{2\lambda}}}\exp\{-\frac{1}{2}\frac{(y-\mathrm{e}^{-\lambda(t-s)}x)^2}{\sigma^2\frac{1-\mathrm{e}^{-2\lambda(t-s)}}{2\lambda}}\}dy.$$

Si può dimostrare che $P(s,t,x,dy)$ definisce una famiglia di probabilità di transizione regolare (gaussiana). Inoltre si potrebbe dimostrare che, se Y_0 è una variabile aleatoria con distribuzione μ_0, indipendente da un processo di Wiener standard $(W_t)_{t\geq 0}$, allora la famiglia di distribuzioni finito dimensionali ottenute tramite $P(s,t,x,dy)$ e la distribuzione iniziale μ_0, ha le stesse distribuzioni finito dimensionali del processo

$$X_t := \mathrm{e}^{-\lambda t} Y_0 + \sigma W_t - \int_0^t \sigma\lambda \mathrm{e}^{-\lambda(t-s)} W_s ds.$$

Si noti che l'integrale ha senso purché si prenda la versione a traiettorie continue del processo di Wiener standard[26].

È inoltre interessante notare che, se $Y_0 = x$, cioè se Y_0 è una variabile aleatoria degenere, allora $X_t := \mathrm{e}^{-\lambda t}x + \sigma W_t - \int_0^t \sigma\lambda \mathrm{e}^{-\lambda(t-s)} W_s ds$ è un processo gaussiano[27], di valore atteso $m(t) = \mathrm{e}^{-\lambda t}x$ e funzione di covarianza

$$K(s,t) = \mathbb{E}\left[\left(\sigma W_t - \int_0^t \sigma\lambda \mathrm{e}^{-\lambda(t-u)} W_u du\right)\left(\sigma W_s - \int_0^s \sigma\lambda \mathrm{e}^{-\lambda(s-u)} W_u du\right)\right].$$

Con un po' di calcoli[28] si può controllare che

$$K(s,t) = \frac{\sigma^2}{2\lambda}(\mathrm{e}^{-\lambda|t-s|} - \mathrm{e}^{-\lambda(t+s)}),$$

ed in particolare quindi

$$Var(X_t) = K(t,t) = \frac{\sigma^2}{2\lambda}(1 - \mathrm{e}^{-2\lambda\, t})$$

---

[26]Vale la pena ricordare che di solito questo processo viene introdotto dopo aver parlato dell'integrale stocastico rispetto al processo di Wiener. Riscrivendo

$$\sigma W_t - \int_0^t \sigma\lambda \mathrm{e}^{-\lambda(t-s)} W_s ds = \sigma \mathrm{e}^{-\lambda t}\left(\mathrm{e}^{\lambda t} W_t - \int_0^t \lambda \mathrm{e}^{\lambda s} W_s ds\right),$$

e interpretando la formula tra parentesi come un'integrazione per parti

$$\mathrm{e}^{\lambda t} W_t - \int_0^t W_s d\mathrm{e}^{\lambda s} =: \int_0^t \mathrm{e}^{\lambda s} dW_s,$$

si può riscrivere

$$X_t := \mathrm{e}^{-\lambda t}\left(Y_0 + \sigma \int_0^t \mathrm{e}^{\lambda s} dW_s\right),$$

o in forma differenziale (stocastica)

$$dX_t = -\lambda X_t dt + \sigma dW_t.$$

Per poter dare un senso più preciso a questa espressione, tuttavia andrebbe prima visto il significato dell'integrale stocastico.
In effetti la ***prima definizione di integrale stocastico data da Wiener*** è stata la seguente: per ogni funzione deterministica $h(t) \in C^1$, ovvero con derivata prima continua, Wiener definiva

$$\int_\alpha^\beta h(s) dW_s := W_\beta h(\beta) - W_\alpha h(\alpha) - \int_\alpha^\beta W_s h'(s) ds.$$

Questa variabile aleatoria risulta gaussiana (vedere la nota successiva) di valore medio nullo e di varianza $\int_\alpha^\beta h^2(s)ds$. Se $h(t) \notin C^1$, ma è misurabile e con $\int_\alpha^\beta h^2(s)ds$ finito, allora Wiener definiva

$$\int_\alpha^\beta h(s) dW_s := \lim_{n\to\infty} \int_\alpha^\beta h_n(s) dW_s,$$

dove il limite è inteso in media quadratica, e dove $\{h_n\}_n$ è una successione di funzioni C^1 con la proprietà che

$$\int_\alpha^\beta \big(h_n(s) - h(s)\big)^2 ds \to 0.$$

[27]Il fatto che sia gaussiano dipende dal fatto che l'integrale $\int_0^t \mathrm{e}^{-\lambda(t-s)} W_s ds = \mathrm{e}^{-\lambda t}\int_0^t \mathrm{e}^{\lambda s} W_s ds$ si può ottenere come limite delle somme di Riemann

$$\mathrm{e}^{-\lambda t} \sum_{k=1}^n \mathrm{e}^{\lambda s_k} W_{s_k}(s_k - s_{k-1}),$$

che sono variabili aleatorie gaussiane, e tenendo conto del fatto che il limite (in distribuzione) di variabili aleatorie gaussiane è ancora una variabile aleatoria gaussiana, purché valore atteso e varianza convergano.

[28]Questo conto risulta più agevole dopo aver introdotto l'integrale stocastico di Ito (vedere la Sezione 9.5.

Nel caso in cui $\lambda > 0$, quando t va all'infinito, allora chiaramente $\mathbb{E}[X_t] = m(t) = \mathrm{e}^{-2\lambda t}x$ converge a zero e la varianza $Var(X_t) = K(t,t)$ converge a $\frac{\sigma^2}{2\lambda}$. Di conseguenza la legge unidimensionale del processo $(X_t)_{t\geq 0}$ converge in distribuzione ad una legge gaussiana $N(0, \frac{\sigma^2}{2\lambda})$.

Questa legge gode di una interessante proprietà: se la legge iniziale μ_0 è appunto una legge gaussiana $N(0, \frac{\sigma^2}{2\lambda})$, cioè $\mu_0(dx) = \frac{1}{\sqrt{2\pi \frac{\sigma^2}{2\lambda}}} \exp\{-\frac{x^2}{2\frac{\sigma^2}{2\lambda}}\} = \sqrt{\frac{\lambda}{\pi\sigma^2}} \exp\{-\frac{\lambda}{\sigma^2}x^2\}$, allora la legge di X_t è ancora una legge gaussiana $N(0, \frac{\sigma^2}{2\lambda})$, ovvero:

$$
\begin{aligned}
\mathbb{P}(X_t \in A) &= \int_{\mathbb{R}} \mu_0(dx) \int_A p(0,t,x,y)dy \\
&= \int_{\mathbb{R}} \sqrt{\frac{\lambda}{\pi\sigma^2}} \exp\{-\frac{\lambda}{\sigma^2}x^2\}dx \int_A \frac{1}{\sqrt{2\pi\sigma^2 \frac{1-\mathrm{e}^{-2\lambda t}}{2\lambda}}} \exp\{-\frac{1}{2}\frac{(y-\mathrm{e}^{-\lambda t}x)^2}{\sigma^2 \frac{1-\mathrm{e}^{-2\lambda t}}{2\lambda}}\}dy \\
&= \int_A \left(\int_{\mathbb{R}} \sqrt{\frac{\lambda}{\pi\sigma^2}} \exp\{-\frac{\lambda}{\sigma^2}x^2\} \sqrt{\frac{\lambda}{\pi\sigma^2(1-\mathrm{e}^{-2\lambda t})}} \exp\{-\frac{\lambda(y-\mathrm{e}^{-\lambda t}x)^2}{\sigma^2(1-\mathrm{e}^{-2\lambda t})}\}dx \right) dy \\
&= \int_A \sqrt{\frac{\lambda}{\pi\sigma^2}} \exp\{-\frac{\lambda}{\sigma^2}y^2\}\, dy, .
\end{aligned}
$$

per ogni A boreliano[29].

Questo esempio porta a dare la definizione di distribuzione stazionaria.

Definizione 8.9 (Distribuzione stazionaria). *Sia μ una distribuzione tale che*

$$\int_{\mathbb{R}} \mu(dx) P(s,t,x,A) = \mu(A)$$

allora μ viene detta **distribuzione stazionaria** *o* **invariante** *del processo di Markov regolare determinato dalle probabilità di transizione $P(s,t,x,A)$.*

Si può dimostrare che se $\mu_0 = \mu$ è una distribuzione stazionaria, allora le distribuzioni finito-dimensionali godono della proprietà di stazionarietà

Definizione 8.10 (Processi stazionari). *Se per ogni $h > 0$, k, $(t_1, \cdots, t_k)$ e $(z_1, \cdots, z_k)$ la famiglia di distribuzioni finito-dimensionali $F_{t_1,\cdots,t_k}(z_1,\cdots,z_k)$ gode della proprietà che*

$$F_{t_1,\cdots,t_k}(z_0, z_1, \cdots, z_k) = F_{t_1+h,\cdots,t_k+h}(z_0, z_1, \cdots, z_k)$$

allora $F_{0,t_1,\cdots,t_k}$ è detta una famiglia di distribuzioni finito-dimensionali stazionarie, e il processo con tale famiglia di distribuzioni finito-dimensionali è detto un **processo stazionario***.*

[29] L'ultima uguaglianza deriva dal fatto che, la densità di X_t è data dall'espressione tra parentesi, ovvero da

$$
\begin{aligned}
&\sqrt{\frac{\lambda}{\pi\sigma^2}} \int_{\mathbb{R}} \frac{\sqrt{\lambda}}{\sqrt{\pi\sigma^2(1-\mathrm{e}^{-2\lambda t})}} \exp\{-(\frac{\lambda(y-\mathrm{e}^{-\lambda t}x)^2}{\sigma^2(1-\mathrm{e}^{-2\lambda t})} + \frac{\lambda x^2(1-\mathrm{e}^{-2\lambda t})}{\sigma^2(1-\mathrm{e}^{-2\lambda t})})\}dx \\
&= \sqrt{\frac{\lambda}{\pi\sigma^2}} \int_{\mathbb{R}} \frac{\sqrt{\lambda}}{\sqrt{\pi\sigma^2(1-\mathrm{e}^{-2\lambda t})}} \exp\{-\frac{\lambda}{\sigma^2(1-\mathrm{e}^{-2\lambda t})}((y-\mathrm{e}^{-\lambda t}x)^2 + x^2(1-\mathrm{e}^{-2\lambda t}))\}dx \\
&= \sqrt{\frac{\lambda}{\pi\sigma^2}} \int_{\mathbb{R}} \frac{\sqrt{\lambda}}{\sqrt{\pi\sigma^2(1-\mathrm{e}^{-2\lambda t})}} \exp\{-\frac{\lambda(y^2 - 2y\mathrm{e}^{-\lambda t}x + \mathrm{e}^{-2\lambda t}x^2 + x^2 - x^2\mathrm{e}^{-2\lambda t})}{\sigma^2(1-\mathrm{e}^{-2\lambda t})}\}dx \\
&= \sqrt{\frac{\lambda}{\pi\sigma^2}} \int_{\mathbb{R}} \frac{\sqrt{\lambda}}{\sqrt{\pi\sigma^2(1-\mathrm{e}^{-2\lambda t})}} \exp\{-\frac{\lambda(x^2 - 2xy\mathrm{e}^{-\lambda t} + y^2)}{\sigma^2(1-\mathrm{e}^{-2\lambda t})}\}dx \\
&= \sqrt{\frac{\lambda}{\pi\sigma^2}} \int_{\mathbb{R}} \frac{\sqrt{\lambda}}{\sqrt{\pi\sigma^2(1-\mathrm{e}^{-2\lambda t})}} \exp\{-\frac{\lambda(x^2 - 2xy\mathrm{e}^{-\lambda t} + y^2\mathrm{e}^{-2\lambda t})}{\sigma^2(1-\mathrm{e}^{-2\lambda t})}\} \exp\{-\frac{\lambda(-y^2\mathrm{e}^{-2\lambda t} + y^2)}{\sigma^2(1-\mathrm{e}^{-2\lambda t})}\}dx \\
&= \sqrt{\frac{\lambda}{\pi\sigma^2}} \exp\{-\frac{\lambda}{\sigma^2}y^2\} \int_{\mathbb{R}} \frac{\sqrt{\lambda}}{\sqrt{\pi\sigma^2(1-\mathrm{e}^{-2\lambda t})}} \exp\{-\frac{\lambda(x-y\mathrm{e}^{-\lambda t})^2}{\sigma^2(1-\mathrm{e}^{-2\lambda t})}\}dx = \sqrt{\frac{\lambda}{\pi\sigma^2}} \exp\{-\frac{\lambda}{\sigma^2}y^2\}.
\end{aligned}
$$

8.7.2 Moto browniano geometrico e modello di Black-Scholes

Il prossimo esempio è di fondamentale importanza nell'ambito dei modelli finanziari. Il ***modello proposto da Black e Scholes*** è un modello a tempo continuo con una azione rischiosa (una azione di prezzo S_t all'istante t) e una azione non rischiosa (di prezzo B_t all'istante t). Si suppone che l'evoluzione di B_t sia descritta dall'equazione differenziale seguente

$$dB_t = rB_t dt,$$

dove r è una costante positiva. Questo significa che il tasso di interesse è costante e uguale a r. Se $B_0 = 1$, allora ovviamente $B_t = e^{rt}$, per $t \geq 0$.
Si suppone che il prezzo dell'azione sia descritto[30] dal processo

$$S_t = S_0 \exp\left\{\mu t - \frac{\sigma^2}{2} t + \sigma W_t\right\}.$$

dove μ e σ sono due costanti e $(W_t)_t$ è un moto browniano standard, ed S_0 è una variabile aleatoria indipendente da $(W_t)_t$.

Tale processo è detto ***moto browniano geometrico***.[31]

Il modello è studiato sull'intervallo $[0, T]$, dove T è la data di scadenza dell'opzione da studiare. In particolare, risulta che la legge di S_t è una legge log-normale, vale a dire che il suo logaritmo segue una legge normale.
Il processo $(S_t)_t$ è una trasformazione biunivoca di un processo di Markov, infatti

$$Y_t = \log(S_t) = \log(S_0) + \sigma W_t + (\mu - \frac{\sigma^2}{2})t$$

è un processo ad incrementi indipendenti ed omogenei, con

$$Y_0 = \log(S_0),$$

e con

$$X_t := \sigma W_t + (\mu - \frac{\sigma^2}{2})t$$

un processo di Wiener, non standard, con coefficiente di drift $\mu' = \mu - \frac{\sigma^2}{2}$ e coefficiente di diffusione σ^2. Tenendo conto di questo fatto si può affermare immediatamente che il processo $(S_t)_t$ verifica le seguenti proprietà:

- continuità delle traiettorie;
- se $u \leq t$, S_t/S_u è indipendente da $\mathcal{F}_u^S = \sigma(S_v, v \leq u)$;
- se $u \leq t$, la legge di $(S_t - S_u)/S_u$ è identica a quella di $(S_{t-u} - S_0)/S_0$.

La prima proprietà è banale. Per verificare le altre due proprietà basta osservare che

$$S_t/S_u = \exp\{Y_t - Y_u\} = \exp\{\sigma(W_t - W_u) + (\mu - \frac{\sigma^2}{2})(t - u)\},$$

---

[30] Di solito si introduce il processo del prezzo S_t come la soluzione dell'equazione differenziale stocastica

$$dX_t = X_t(\mu dt + \sigma dW_t), \quad X_0 = S_0,$$

che si risolve esplicitamente:

$$X_t = S_0 \exp\left\{\mu t - \frac{\sigma^2}{2} t + \sigma W_t\right\}.$$

Di nuovo per capire bene il senso dell'equazione differenziale stocastica precedente andrebbe prima chiarito il significato dell'integrale stocastico di Ito e del differenziale stocastico (vedere la Sezione 9.5.

[31] Questo processo si può ottenere anche come limite del processo dei prezzi per il modello di Cox, Ross e Rubistein. Per un approccio elementare si veda il testo di S. Ross [15].

che $\mathcal{F}_u = \mathcal{F}_u^S = \mathcal{F}_u^W \vee \sigma(S_0)$ e che $(W_t - W_u)$ è indipendente da $\mathcal{F}_u^W$ e da $\sigma(S_0)$ ed ha la stessa legge di $(W_{t-u} - W_0) = W_{t-u}$.

Inoltre il processo $(S_t)_t$ risulta esso stesso un processo di Markov regolare: per $x_i > 0$, $i = 0, 1, \cdots, k$

$$F_{S_0,S_1,\cdots,S_k}(x_0, x_1, \cdots, x_k) = \mathbb{P}(S_0 \leq x_0, S_{t_1} \leq x_1, \cdots, S_{t_k} \leq x_k)$$

$$= \mathbb{P}(Y_0 \leq \log(x_0),\, Y_{t_1} \leq \log(x_1), \cdots, Y_{t_k} \leq \log(x_k))$$

$$= \int_{-\infty}^{\log(x_0)} \int_{-\infty}^{\log(x_1)} \cdots \int_{-\infty}^{\log(x_k)} p_{Y_0,Y_{t_1},\cdots,Y_{t_k}}(y_0, y_1, \cdots, y_k)\, dy_0\, dy_1 \cdots dy_k$$

$$= \int_{-\infty}^{\log(x_0)} \int_{-\infty}^{\log(x_1)} \cdots \int_{-\infty}^{\log(x_k)} p_{Y_0}(y_0) \prod_{i=1}^{k} \frac{1}{\sqrt{2\pi(t_i - t_{i-1})\sigma^2}} \mathrm{e}^{-\frac{(y_i - y_{i-1} - \mu'(t_i - t_{i-1}))^2}{2(t_i - t_{i-1})\sigma^2}} \, dy_0\, dy_1 \cdots dy_k$$

di conseguenza la densità congiunta di $(S_0, S_{t_1}, \cdots, S_{t_k})$ si ottiene derivando rispetto a x_0, x_1, $\cdots$, x_k la precedente espressione:

$$p_{S_0,S_{t_1},\cdots,S_{t_k}}(x_0, x_1, \cdots, x_k) = \frac{\partial^n}{\partial x_0 \partial x_1 \cdots \partial x_k} F_{S_0,S_1,\cdots,S_k}(x_0, x_1, \cdots, x_k)$$

$$= p_{Y_0,Y_{t_1},\cdots,Y_{t_k}}(\log(x_0), \log(x_1), \cdots, \log(x_k)) \frac{1}{x_0} \frac{1}{x_1} \cdots \frac{1}{x_k}$$

$$= \frac{p_{Y_0}(\log(x_0))}{x_0} \prod_{i=1}^{k} \frac{1}{x_i \sqrt{2\pi(t_i - t_{i-1})\sigma^2}} \mathrm{e}^{-\frac{(\log(x_i) - \log(x_{i-1}) - \mu'(t_i - t_{i-1}))^2}{2(t_i - t_{i-1})\sigma^2}}.$$

La densità risulta evidentemente nulla se una delle x_i non è strettamente positiva.

Da ciò risulta evidente che le probabilità di transizione ammettono la seguente densità

$$p(s, t, x, y) = \frac{1}{y\sqrt{2\pi(t-s)\sigma^2}} \exp\left\{ -\frac{\big(\log(y) - \log(x) - \mu'(t-s)\big)^2}{2(t-s)\sigma^2} \right\}.$$

E infine facile vedere che queste densità di probabilità di transizione verificano l'equazione di Chapman-Kolmogorov in quanto ci si riporta immediatamente attraverso un cambio di variable al caso delle densità di probabilità di transizione del processo Y_t.

La formula di Black-Scholes permette di calcolare in modo esplicito il prezzo di un'opzione call europea. Si indichi con $C_0(x)$ il prezzo dell'opzione emessa al tempo 0 con la condizione $S_0 = x$. Si ponga

$$Z_t^\theta = \exp\left\{ -\frac{1}{2}\theta^2 t + \theta W_t \right\}$$

che è la martingala esponenziale (già definita in (8.11))con $\mathbb{E}[Z_t^\theta] = 1$.

Se si prende $\theta = \frac{r-\mu}{\sigma}$ per il Teorema di Girsanov (vedere l'Esercizio 8.1) il processo scontato dei prezzi $(\tilde{S}_t = \frac{S_t}{B_t} = \mathrm{e}^{-rt} S_t)_t$ è una martingala[32] rispetto alla misura

$$\mathbb{Q}(A) = \mathbb{Q}^\theta(A) = \mathbb{E}^{\mathbb{P}}[\mathbb{I}_A Z_T^\theta], \quad A \in \mathcal{F}_T,$$

[32] Posto

$$\widetilde{W}_t = \frac{\mu - r}{\sigma} t + W_t,$$

il processo

$$\tilde{S}_t = x \exp\{(\mu - r - \frac{\sigma^2}{2})t + \sigma W_t\}$$

si può riscrivere anche come

$$\tilde{S}_t = x \exp\{-\frac{\sigma^2}{2} t + \sigma \widetilde{W}_t\}$$

Rispetto alla misura di probabilità $\mathbb{Q}$ il processo W_t risulta un processo di Wiener con coefficiente di drift θ, cioè equivalentemente

$$W_t = (W_t - \theta t) + \theta t = B_t + \theta t,$$

ovvero con $d\mathbb{Q} = Z_T^\theta d\mathbb{P}$. Estendendo la definizione data nel caso a tempo discreto, si ha quindi che la misura $\mathbb{Q}$ è una ***misura martingala equivalente*** a $\mathbb{P}$.

Inoltre si ha che, per ogni t, la legge della variabile aleatoria $S_t = \mathrm{e}^{rt}\widetilde{S}_t$ rispetto a $\mathbb{Q}$ è la stessa[33] della variabile aleatoria $x\exp\{(r - \sigma^2/2)t + \sigma\sqrt{t}Z\}$, dove Z è una variabile aleatoria gaussiana standard $N(0,1)$.

Quindi

$$\begin{aligned} C_0(x) &= e^{-rT}\mathbb{E}^{\mathbb{Q}}\left[(S_T - K)^+\right] \\ &= e^{-rT}\mathbb{E}^{\mathbb{Q}}\left[\left(xe^{(r-\frac{\sigma^2}{2})T+\sigma\sqrt{T}Z} - K\right)^+\right]. \end{aligned}$$

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

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

e quindi

$$C_0(x) = \frac{e^{-rT}}{\sqrt{2\pi}}\int_{\zeta}^{+\infty}\left(xe^{(r-\frac{\sigma^2}{2})T+\sigma\sqrt{T}z} - K\right)e^{-z^2/2}dz.$$

Se si indica con Φ la funzione di ripartizione della legge $\mathcal{N}(0,1)$,

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

allora

$$\begin{aligned} C_0(x) &= \frac{x}{\sqrt{2\pi}}\int_{\zeta}^{+\infty} e^{-\frac{1}{2}(z-\sigma\sqrt{T})^2}dz - Ke^{-rT}\Phi(-\zeta) \\ &= \frac{x}{\sqrt{2\pi}}\int_{\zeta-\sigma\sqrt{T}}^{+\infty} e^{-z^2/2}dz - Ke^{-rT}\Phi(-\zeta) \\ &= x\Phi\left(-\zeta + \sigma\sqrt{T}\right) - Ke^{-rT}\Phi(-\zeta). \end{aligned}$$

In conclusione si ottiene la *formula di Black-Scholes*

$$C_0(x) = x\Phi\left(-\zeta + \sigma\sqrt{T}\right) - Ke^{-rT}\Phi(-\zeta).$$

---

con B_t un processo di Wiener standard rispetto a $\mathbb{Q} = \mathbb{Q}^\theta$. Di conseguenza

$$\widetilde{W}_t = \frac{\mu - r}{\sigma}t + W_t = \frac{\mu - r}{\sigma}t + \theta t + B_t.$$

Prendendo quindi

$$\theta = -\frac{\mu - r}{\sigma} = \frac{r - \mu}{\sigma},$$

si ottiene che $\widetilde{W}_t = B_t$ è un processo di Wiener standard rispetto a $\mathbb{Q} = \mathbb{Q}^\theta$.
A questo punto si osserva che, proprio per questo motivo,

$$\tilde{S}_t = x\exp\{-\frac{\sigma^2}{2}t + \sigma\widetilde{W}_t\}$$

è rispetto alla misura $\mathbb{Q}$ una martingala esponenziale ovvero

$$\tilde{S}_t = x\widetilde{Z}_t^\sigma = x\exp\{-\frac{\sigma^2}{2}t + \sigma\widetilde{W}_t\}, \tag{8.14}$$

si tratta di applicare il risultato dell'Esercizio 8.1 con $\mathbb{Q}$ al posto di $\mathbb{P}$, $\widetilde{W}_t$ al posto di W_t ed infine σ al posto di θ.

[33] L'uguaglianza in legge si vede dall'espressione del prezzo scontato (8.14), e tenendo conto del fatto che la legge della variabile aleatoria $\widetilde{W}_t$ rispetto a $\mathbb{Q}$ è $N(0,t)$ e che anche la legge di $\sqrt{t}Z$ 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.

Per ulteriori approfondimenti consultare il libro di P. Baldi [1], quello di D. Lamberton e B. Lapeyre [9], oppure di J.M. Steele [18] .
Mediante la formula di Black-Scholes, possiamo ricavare il prezzo equo di un'opzione call europea. Tuttavia, grazie alla formula di parità (si veda ad esempio il libro di S. Ross [15]), si ricava immediatamente anche il prezzo equo di una put europea P_t, dato da

$$P_t = C_t - S_t + Ke^{-r(T-t)}.$$

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 la volatilità σ^2[34]

Infine va notata l'analogia della formula di Black e Scholes con il corrispondente risultato per il modello di Cox, Ross e Rubistein (CRR). A tale proposito si veda, sempre nel testo di S. Ross[35] [15], come sia possibile ottenere tale formula in modo elementare come limite del prezzo del modello CRR, mandando il numero di passi all'infinito, con opportuni riscalamenti nel modello stesso.

8.8 Appendice: dimostrazione del Teorema di esistenza di Kolmogorov †‡

Un problema noto nel caso unidimensionale è il seguente: data una distribuzione di probabilità μ su $\mathbb{R}$, esiste uno spazio di probabilità $(\Omega,\mathcal{F},\mathbb{P})$ ed una variabile aleatoria X che ammette come distribuzione μ?

Ci sono due possibili risposte "classiche" a tale quesito: la prima consiste nel prendere lo spazio canonico $\Omega = \mathbb{R}$, $\mathcal{F} = \mathcal{B}(\mathbb{R})$, X uguale all'identità, cioè $X(x) = x$, e infine $\mathbb{P} = \mu$; la seconda (Teorema di rappresentazione di Skorohod) consiste nel prendere $\Omega = (0, 1)$, $\mathcal{F} = \mathcal{B}(0, 1)$, $\mathbb{P}$ la misura di Lebesgue ristretta a (0,1) ed $X(u) = F^{-1}(u) := \inf\{x \text{ tali che } u \le F(x)\}$, dove $F(x) := \mu\big((-\infty, x]\big)$ è la funzione di distribuzione ed F^{-1} è la ***funzione inversa generalizzata*** [36] della funzione di distribuzione F.

Anche nel caso dei processi c'è qualcosa di simile. Il problema si esprime nel seguente modo: data una famiglia di distribuzioni finito-dimensionali $\mu_{t_1,\cdots,t_k}$, al variare di $k \ge 1$ e di $t_1, \cdots, t_k$ in I, esiste uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ ed un processo aleatorio $(X_t, t \ge 0)$ che ammette tali distribuzioni finito-dimensionali?

Lo spazio canonico, analogo ad $\mathbb{R}$, è $\mathbb{R}^I = \{x(\cdot) :\ I \to \mathbb{R}\}$, ed il processo canonico è il processo coordinata per il quale $X_t(x(\cdot)) := x(t)$. La scelta della σ-algebra viene fatta in modo che il processo canonico sia misurabile. Quindi necessariamente deve contenere gli insiemi del tipo

$$\{x(\cdot) \in \mathbb{R}^I \mid x(t) \in A\}, \text{ per } A \in \mathcal{B}(\mathbb{R}),$$

e delle intersezioni finite di insiemi di questo tipo, i rettangoli di base finita

$$\{x(\cdot) \in \mathbb{R}^I \mid x(t_1) \in A_1, \cdots, x(t_k) \in A_k,\}, \text{ per } A_h \in \mathcal{B}(\mathbb{R}) \text{ e } h = 1, \cdots, k.$$

Necessariamente, quindi, deve contenere anche i cilindri (o insiemi finito-dimensionali), ovvero degli insiemi del tipo

$$\mathcal{C} = \{x(\cdot) \in \mathbb{R}^I \mid (x(t_1), x(t_2), \cdots, x(t_k)) \in H\}, \text{ dove } H \in \mathcal{B}(\mathbb{R}^k) \tag{8.15}$$

Si sceglie quindi la σ-algebra $\mathcal{R}^I$ generata dall'algebra $\mathcal{R}^I_0$ dei cilindri, al variare di $k \ge 1$, degli indici $t_1, \cdots, t_k$ in I e di H in $\mathcal{B}(\mathbb{R}^k)$[37].

Infine come misura di probabilità $\mathbb{P}$, chiaramente, si vuole prendere una misura di probabilità su $\mathcal{R}^I$ per la quale valga

[34] 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.

[35] Il Ross del modello CRR e l'autore di [15] sono due persone diverse.

[36] Se non si è familiari con il teorema di Skorohod basta considerare solo il caso in cui la funzione di distribuzione $F(\cdot)$ sia strettamente crescente e continua e quindi invertibile.

[37] Per la dimostrazione del fatto che $\mathcal{R}^I_0$ sia un'algebra vedere la dimostrazione del seguente teorema.

$$\mathbb{P}(\mathcal{C}) = \mu_{t_1,\cdots,t_k}(H),$$

per ogni cilindro $\mathcal{C}$ di $\mathcal{R}_0^I$, definito come in (8.15).
A questo punto si pone il problema: esiste una tale probabilità $\mathbb{P}$? e più precisamente

(i) la definizione di $\mathbb{P}$ è ben posta su $\mathcal{R}_0^I$? (ii) $\mathbb{P}$ si può estendere a tutto $\mathcal{R}^I$?

La risposta è affermativa sotto alcune semplici condizioni di consistenza ed è il contenuto del Teorema di esistenza di Kolmogorov. Come si intuisce dal discorso precedente l'ingrediente essenziale della dimostrazione di tale teorema è il procedimento di estensione di una misura definita su un'algebra alla σ-algebra da essa generata (Teorema di Caratheodory).

Le ***condizioni di consistenza*** sono le seguenti:

Sia $k > 1$ e sia π una permutazione di $\{1, \cdots, k\}$ e sia

$$\Phi_\pi : \mathbb{R}^k \to \mathbb{R}^k, (x_1, \cdots, x_k) \to \Phi_\pi(x_1, \cdots, x_k) := (x_{\pi_1}, \cdots, x_{\pi_k}).$$

Chiaramente, essendo Φ_π biunivoca, $(x_1, \cdots, x_k) \in H$ se e solo se $\Phi_\pi(x_1, \cdots, x_k) \in \Phi_\pi(H)$, per $H \in \mathcal{B}(\mathbb{R}^k)$, e quindi

$$\mu_{t_1,\cdots,t_k}(H) = \mu_{t_{\pi_1},\cdots,t_{\pi_k}}(\Phi_\pi(H)), \tag{8.16}$$

infatti

$$\begin{aligned}\mu_{t_1,\cdots,t_k}(H) = \mathbb{P}\{(X_{t_1}, \cdots, X_{t_k}) \in H\} = \mathbb{P}\{\Phi_\pi(X_{t_1}, \cdots, X_{t_k}) \in \Phi_\pi(H)\} = \\ = \mathbb{P}\{(X_{t_{\pi_1}}, \cdots, X_{t_{\pi_k}}) \in \Phi_\pi(H)\} = \mu_{t_{\pi_1},\cdots,t_{\pi_k}}(\Phi_\pi(H)),\end{aligned}$$

inoltre

$$\mu_{t_1,\cdots,t_k,t_{k+1}}(H \times \mathbb{R}) = \mu_{t_1,\cdots,t_k}(H). \tag{8.17}$$

È immediato constatare che le condizioni di consistenza (8.16) e (8.17) si possono riscrivere nel seguente modo:
Siano $k \geq h \geq 1$, sia π una permutazione di $\{1, \cdots, k\}$ e sia $\Psi_{\pi,h} : \mathbb{R}^k \to \mathbb{R}^h$, $(x_1, \cdots, x_k) \to \Psi_{\pi,h}(x_1, \cdots, x_k) := (x_{\pi_1}, \cdots, x_{\pi_h})$, allora, per ogni $(t_1, \cdots, t_k) \in I^k$ e $H \in \mathcal{B}(\mathbb{R}^h)$

$$\mu_{t_1,\cdots,t_k}(\Psi_{\pi,h}^{-1}(H)) = \mu_{t_{\pi_1},\cdots,t_{\pi_h}}(H), \tag{8.18}$$

infatti, nel caso in cui H sia un rettangolo cioè $H = A_1 \times \cdots \times A_h$, posto π^{-1} la permutazione inversa di π, pensata quindi come funzione biunivoca di $\{1, \cdots, k\}$ in sé, e posto $A_m = \mathbb{R}$ per $m = h+1, \cdots, k$, la relazione (8.18) diviene

$$\begin{aligned}
&\mathbb{P}((X_{t_{\pi_1}}, \cdots, X_{t_{\pi_h}}) \in H) = \mathbb{P}(X_{t_{\pi_1}} \in A_1, \cdots, X_{t_{\pi_h}} \in A_h) = \\
&= \mathbb{P}(X_{t_{\pi_1}} \in A_1, \cdots, X_{t_{\pi_h}} \in A_h, X_{t_{\pi_{h+1}}} \in A_{h+1}, \cdots, X_{t_{\pi_k}} \in A_k) = \\
&= \mathbb{P}(X_{t_{\pi_1}} \in A_{\pi^{-1}(\pi_1)}, \cdots, X_{t_{\pi_h}} \in A_{\pi^{-1}(\pi_h)}, X_{t_{\pi_{h+1}}} \in A_{\pi^{-1}(\pi_{h+1})}, \cdots, X_{t_{\pi_k}} \in A_{\pi^{-1}(\pi_k)}) = \\
&\qquad\qquad\qquad\text{(riordinando opportunamente le condizioni richieste)} \\
&= \mathbb{P}(X_{t_1} \in A_{\pi^{-1}(1)}, \cdots, X_{t_h} \in A_{\pi^{-1}(h)}, X_{t_{h+1}} \in A_{\pi^{-1}(h+1)}, \cdots, X_{t_k} \in A_{\pi^{-1}(k)}) = \\
&= \mathbb{P}((X_{t_1}, \cdots, X_{t_k}) \in \Psi_{\pi,h}^{-1}(H)),
\end{aligned}$$

e se vale per i rettangoli, poi vale per ogni cilindro.

Prima di enunciare e dimostrare il teorema di Kolmogorov, menzioniamo il fatto che esiste un'altra possibilità, più simile a quanto fatto nel teorema di rappresentazione di Skorohod, nel caso in cui I è numerabile, cioè $I = \mathbb{N}$, ed illustreremo questo caso in seguito. In tale caso lo spazio canonico è di nuovo (0,1) con la $\sigma-$algebra dei boreliani $\mathcal{B}(0,1)$ e la misura di Lebesgue ristretta a $(0,1)$, e quindi, a partire dalle distribuzioni finito-dimensionali non deve essere definita la misura di probabilità, bensì il processo stesso. Ingrediente essenziale per la dimostrazione è il teorema di esistenza delle versioni regolari delle distribuzioni condizionali di una variabile aleatoria, dato una variabile aleatoria multidimensionale.

Cominciamo con l'enunciare il Teorema di Kolmogorov[38]

Teorema 8.5 (di Kolmogorov). *Sia data una famiglia $\mu_{t_1,\cdots,t_k}$ di distribuzioni finito-dimensionali consistente, cioè che verifica le condizioni di consistenza (8.16) e (8.17), allora esiste uno spazio di probabilità ed un processo aleatorio che ammette $\mu_{t_1,\cdots,t_k}$ come distribuzioni finito-dimensionali. Inoltre è sempre possibile prendere come spazio di probabilità lo spazio canonico $\mathbb{R}^I$ e come processo il processo canonico $X_t(x(\cdot)) = x(t)$.*

Dimostrazione: Cominciamo con l'ultima parte del teorema. Sia $(\Omega, \mathcal{F}, \mathbb{Q})$ uno spazio di probabilità ed $(Y_t, t \in I)$ un processo su tale spazio che ammette come distribuzioni finito-dimensionali $\mu_{t_1,\cdots,t_k}$. Si definisca la funzione $\xi : \Omega \to \mathbb{R}^I, \omega \to \xi(\omega)(\cdot) = Y(\cdot, \omega)$, cioè $\xi(\omega)$ è la funzione $t \to \xi(\omega)(t) = Y_t(\omega)$.

Risulta che ξ è una funzione $\mathcal{F}/\mathcal{R}^I$ misurabile: infatti per ogni cilindro $\mathcal{C}$ come in (8.15), l'insieme

$$\xi^{-1}(\mathcal{C}) = \{\omega \text{ tali che } (Y_{t_1}, \cdots, Y_{t_k}) \in H\} \in \mathcal{F},$$

e ciò è sufficiente a dimostrare la misurabilità di ξ.
A questo punto ponendo

$$\mathbb{P}(\mathcal{A}) = \mathbb{Q}(\xi^{-1}(\mathcal{A})), \text{ per } \mathcal{A} \in \mathcal{R}^I,$$

si ottiene la rappresentazione canonica.

Continuiamo dando lo **schema della dimostrazione**

punto 1) $\mathcal{R}_0^I$ è un'algebra.

punto 2) La definizione $\mathbb{P}$ su $\mathcal{R}_0^I$ come $\mathbb{P}(\mathcal{C}) := \mu_{t_1,\cdots,t_k}(H)$, dove $\mathcal{C}$ è il cilindro definito in (8.15), è ben posta e risulta $\mathbb{P}$ una misura di probabilità finitamente additiva su $\mathcal{R}_0^I$.

punto 3) La misura $\mathbb{P}$ è numerabilmente additiva su $\mathcal{R}_0^I$, e quindi si può estendere in modo univoco ad una misura alla σ-algebra $\mathcal{R}^I$ generata da $\mathcal{R}_0^I$.

Sia il punto 1) che il punto 2) sono basati sulla osservazione che un cilindro $\mathcal{C}$ può avere più di una rappresentazione:

$$\mathcal{C} = \{x(\cdot) \in \mathbb{R}^I \mid (x(t_1), x(t_2), \cdots, x(t_k)) \in H\}, \text{ con } H \in \mathcal{B}(\mathbb{R}^k),$$

oppure

$$\mathcal{C} = \{x(\cdot) \in \mathbb{R}^I \mid (x(s_1), x(s_2), \cdots, x(s_m)) \in J\}, \text{ con } J \in \mathcal{B}(\mathbb{R}^m).$$

Notazione: per k e $t_1, \cdots, t_k$ prefissati, indicheremo con $\mathcal{R}^{\{t_1,\cdots,t_k\}}$ la famiglia dei cilindri del tipo precedente al variare di $H \in \mathcal{B}(\mathbb{R}^k)$.
Supponiamo che $k \le m$, e che $\{t_1, \cdots, t_k\} \subseteq \{s_1, \cdots s_m\}$, allora esiste una permutazione σ di $\{1, \cdots, m\}$ per cui $t_i = s_{\sigma_i}$, $i = 1, \cdots, k$, per cui $\Psi_{\sigma,k}(J) = H$ e per cui $J = \Psi_{\pi,h}^{-1}(H)$ o, equivalentemente $\Phi_\sigma(J) = H \times \mathbb{R}^{m-k}$.
Si noti allora che dalla (8.18) si ottiene che

$$\mu_{s_1,\cdots,s_m}(J) = \mu_{s_1,\cdots,s_m}(\Psi_{\sigma,k}^{-1}(H)) = \mu_{s_{\sigma_1},\cdots,s_{\sigma_k}}(H) = \mu_{t_1,\cdots,t_k}(H). \tag{8.19}$$

[38]Le condizioni di questa formulazione del Teorema di Kolmogorov leggermente più forti di quelle richieste nell'enunciato del Teorema 8.1, ma sono equivalenti.

Inoltre, due cilindri $\mathcal{C}$ e $\mathcal{C}'$, con indici di tempo $\{t_1, \cdots, t_k\}$ e $\{t'_1, \cdots, t'_{k'}\}$ rispettivamente, si possono sempre rappresentare come cilindri con indici comuni $\{s_1, \cdots s_m\} = \{t_1, \cdots, t_k\} \cup \{t'_1, \cdots, t'_{k'}\}$.

Più precisamente se

$$\mathcal{C} = \{x(\cdot) \in \mathbb{R}^I \mid (x(t_1), x(t_2), \cdots, x(t_k)) \in H\} \in \mathcal{R}^{\{t_1,\cdots,t_k\}}$$

e

$$\mathcal{C}' = \{x(\cdot) \in \mathbb{R}^I \mid (x(t'_1), x(t'_2), \cdots, x(t'_{k'})) \in H'\} \in \mathcal{R}^{\{t'_1,\cdots,t'_{k'}\}},$$

allora, posto

$$\{s_1, \cdots, s_m\} = \{t_1, \cdots, t_k\} \cup \{t'_1, \cdots, t'_{k'}\},$$

si può supporre, senza ledere in generalità (si tratta eventualmente di ricorrere a permutazioni opportune), che

$$\{s_1, \cdots, s_h, s_{h+1}, \cdots, s_k\} = \{t_1, \cdots, t_k\}, \{s_{h+1}, \cdots, s_k\} = \{t_1, \cdots, t_k\} \cap \{t'_1, \cdots, t'_{k'}\}$$

ed infine che

$$\{s_{h+1}, \cdots, s_m\} = \{t'_1, \cdots, t'_{k'}\}.$$

Infatti, per due opportune permutazioni π di $\{1, \cdots, k\}$ e π' di $\{1, \cdots, k'\}$, si può riscrivere $(t_{\pi_1}, \cdots, t_{\pi_k}) = (s_1, \cdots, s_k)$ e $(t'_{\pi_1}, \cdots, t'_{\pi_{k'}}) = (s_{h+1}, \cdots, s_m)$ e quindi

$$\begin{aligned} \mathcal{C} &= \{x(\cdot) \in \mathbb{R}^I \mid (x(s_1), x(s_2), \cdots, x(s_k)) \in \Phi_\pi(H)\} \\ &= \{x(\cdot) \in \mathbb{R}^I \mid (x(s_1), \cdots, x(s_m)) \in \Phi_\pi(H) \times \mathbb{R}^{m-k}\} \in \mathcal{R}^{\{s_1,\cdots,s_m\}} \end{aligned} \tag{8.20}$$

e

$$\begin{aligned} \mathcal{C}' &= \{x(\cdot) \in \mathbb{R}^I \mid (x(s_{h+1}), x(s_{h+2}), \cdots, x(s_m)) \in \Phi_{\pi'}(H')\} \\ &= \{x(\cdot) \in \mathbb{R}^I \mid (x(s_1), \cdots, x(s_m)) \in \mathbb{R}^h \times \Phi_{\pi'}(H')\} \in \mathcal{R}^{\{s_1,\cdots,s_m\}} \end{aligned} \tag{8.21}$$

Dimostrazione in dettaglio dei punti 1), 2) e 3).

punto 1)

È immediato capire che, fissato $k \geq 1$ e $\{t_1, \cdots, t_k\}$, la famiglia $\mathcal{R}^{\{t_1,\cdots,t_k\}}$ dei cilindri del tipo

$$\mathcal{C} = \{x(\cdot) \in \mathbb{R}^I \mid (x(t_1), x(t_2), \cdots, x(t_k)) \in H\},$$

con $H \in \mathcal{B}(\mathbb{R}^k)$, forma un'algebra (anche se in realtà si tratta di una $\sigma-$ algebra):

Per ottenere $\mathbb{R}^I$, basta prendere $H = \mathbb{R}^k$.

Per ottenere $\mathcal{C}^c$, basta prendere H^c.

Per ottenere $\mathcal{C} \cup \mathcal{C}'$, con $\mathcal{C}' = \{x(\cdot) \in \mathbb{R}^I \mid (x(t_1), x(t_2), \cdots, x(t_k)) \in H'\}$, basta prendere $H \cup H'$.

Dalle osservazioni precedenti sappiamo che comunque presi due cilindri essi si possono esprimere come cilindri con indici comuni. Si ottiene che quindi $\mathcal{R}^I_0$ è un'algebra.

punto 2)

Le osservazioni iniziali e le proprietà di consistenza permettono di ottenere immediatamente che la definizione è ben posta. Infatti se i cilindri $\mathcal{C}$ e $\mathcal{C}'$ di (8.20) e (8.21) sono uguali allora necessariamente $\Phi_\pi(H) \times \mathbb{R}^{m-k} = \mathbb{R}^h \times \Phi_{\pi'}(H') = J$, e allora per la (8.19) 5

$$\mu_{s_1,\cdots,s_m}(J) = \mu_{t_1,\cdots,t_k}(H) = \mu_{t'_1,\cdots,t'_{k'}}(H').$$

La finita additività dipende dal fatto che ciascuna $\mu_{t_1,\cdots,t_k}$ è una misura di probabilità su $\mathcal{R}^{\{t_1,\cdots,t_k\}}$ e che, dato un numero finito di cilindri di $\mathcal{R}^I_0$, si può sempre pensare che tutti i cilindri siano appartenenti ad un'algebra del tipo $\mathcal{R}^{\{t_1,\cdots,t_k\}}$.

punto 3)

Per provare la σ-additività è sufficiente mostrare la continuità, e per questo, a sua volta, è sufficiente mostrare che se $\mathcal{A}_n \in \mathcal{R}_0^I$ ed $\mathcal{A}_n \downarrow \emptyset$ allora $\mathbb{P}(\mathcal{A}_n) \downarrow 0$. La prova procede per assurdo. Esista, per assurdo, un $\epsilon > 0$ tale che $\mathbb{P}(\mathcal{A}_n) \geq \epsilon$ per ogni n. Se mostriamo che allora $\bigcap_n \mathcal{A}_n$ non può essere l'insieme vuoto la dimostrazione è completa.

Poiché per rappresentare un cilindro si può sempre aumentare il numero degli indici, cioè degli istanti di tempo coinvolti nella sua definizione, possiamo supporre che esista una successione $\{t_n, n \geq 1\}$ e una successione di boreliani $H_n \in \mathcal{B}(\mathbb{R}^n)$ per cui

$$\mathcal{A}_n = \{x \in \mathbb{R}^I, (x(t_1), \cdots, x(t_n)) \in H_n\}$$

(eventualmente ripetendo qualche $\mathcal{A}_n$[39])

Ovviamente $\mathbb{P}(\mathcal{A}_n) = \mu_{t_1,\cdots,t_n}(H_n)$, ed essendo $\mu_{t_1,\cdots,t_n}$ una ***misura internamente regolare*** [40] è possibile trovare un compatto $K_n \subseteq H_n$ per cui

$$\mu_{t_1,\cdots,t_n}(H_n \setminus K_n) \leq \frac{\epsilon}{2^{n+1}},$$

di modo che se $\mathcal{B}_n = \{x \in \mathbb{R}^I, (x(t_1), \cdots, x(t_n)) \in K_n\}$ allora

$$\mathbb{P}(\mathcal{A}_n \setminus \mathcal{B}_n) \leq \frac{\epsilon}{2^{n+1}}.$$

Posto $\mathcal{C}_n = \bigcap_{m=1}^n \mathcal{B}_m$, di modo che

$$\mathcal{C}_n = \{x \in \mathbb{R}^I, (x(t_1), \cdots, x(t_n)) \in \Gamma_n\}, \text{ con } \Gamma_n = \bigcap_{m=1}^n (K_m \times \mathbb{R}^{n-m}) \subseteq K_n \text{ compatto,}$$

allora

$$\mathcal{C}_n \subseteq \mathcal{B}_n \subseteq \mathcal{A}_n \text{ e } \mathbb{P}(\mathcal{A}_n \setminus \mathcal{C}_n) < \frac{\epsilon}{2}.$$

Infatti

$$\mathbb{P}(\mathcal{A}_n \setminus \mathcal{C}_n) = \mathbb{P}(\mathcal{A}_n \cap \Big(\bigcap_{m=1}^n \mathcal{B}_m\Big)^c) = \mathbb{P}(\bigcup_{m=1}^n \mathcal{A}_n \cap \mathcal{B}_m^c) \leq$$

$$\leq \sum_{m=1}^n \mathbb{P}(\mathcal{A}_m \cap \mathcal{B}_m^c) = \sum_{m=1}^n \mathbb{P}(\mathcal{A}_m \setminus \mathcal{B}_m) \leq \sum_{m=1}^n \frac{\epsilon}{2^{m+1}} < \frac{\epsilon}{2}.$$

Di conseguenza $\mathbb{P}(\mathcal{C}_n) \geq \frac{\epsilon}{2} > 0$ (essendo $\mathbb{P}(\mathcal{A}_n) \geq \epsilon$ e $\mathbb{P}(\mathcal{A}_n \setminus \mathcal{C}_n) < \frac{\epsilon}{2}$). Quindi $\mathcal{C}_n$ non è vuoto e lo stesso vale per $\Gamma_n \subseteq K_n$.

Si scelga per ogni n un punto appartenente al cilindro $\mathcal{C}_n$, ovvero una funzione, $x^{(n)}(\cdot) \in \mathcal{C}_n$. Se$n \geq k$ allora $x^{(n)}(\cdot) \in \mathcal{C}_n \subseteq \mathcal{C}_k \subseteq \mathcal{B}_k$ e quindi il punto $(x^{(n)}(t_1), \cdots, x^{(n)}(t_k)) \in K_k \subseteq J_k \times J_k \times \cdots \times J_k$ (k volte, con J_k

[39] In generale ciascun $\mathcal{A}_n$ coinvolgerà gli istanti $t_1, \cdots, t_{m_n}$; in tale caso si dovrà ripetere $\mathbb{R}^I$ per $m_1 - 1$ volte, con $H_n = \mathbb{R}_n$ per $n = 1, \cdots, m_1 - 1, \mathcal{A}_1$ per $m_2 - m_1$ volte e così via.

[40] Ricordiamo che, se S è uno spazio metrico, si dice che una misura μ sui boreliani di S è ***internamente regolare*** se per ogni insieme misurabile B si ha che $\mu(B) = \sup\{\mu(J), J \subseteq B, J \text{ compatto}\}$, e quindi esiste una successione di compatti J_h contenuti in B per cui la misura di B è il limite delle misure di J_h.

Comunque in $\mathbb{R}^n$ ciò significa che per ogni boreliano H è possibile trovare una successione monotona di iper-rettangoli chiusi e limitati $\{K_m\}_{m\geq 1}$ e convergente ad H. Per la proprietà di continuità delle probabilità vale allora $\mu(K_m) \uparrow \mu(H)$.

Questo è un punto cruciale nella dimostrazione e che vale per tutte le misure di probabilità su uno spazio metrico localmente compatto. Anzi ciò vale addirittura per spazi di Hausdorff sempre localmente compatto (confrontare, ad esempio, il Teorema di rappresentazione di Riesz). Questa osservazione permette di estendere il teorema di Kolmogorov non solo al caso di processi a valori reali, ma anche al caso di processi a valori in spazi metrici localmente compatti, ed in particolare a $\mathbb{R}^d$.

un intervallo limitato). Di conseguenza, ciascuna successione del tipo $\{x^{(1)}(t_k), x^{(2)}(t_k), \cdots, x^{(n)}(t_k), \cdots\}$ è contenuta in J_k, e quindi è limitata.

In realtà la successione $\{x^{(1)}(t_k), x^{(2)}(t_k), \cdots, x^{(n)}(t_k), \cdots\}$ è contenuta in J_k solo definitivamente, anzi $x^{(n)}(t_k) \in J_k$ per $n \geq k$, comunque è una successione che ammette almeno una sottosuccessione convergente. Anche questo è un punto da tenere presente nel caso in cui si voglia fare una generalizzazione a processi a valori in spazi più generali di $\mathbb{R}$ o $\mathbb{R}^d$.

Con il metodo diagonale si può scegliere una successione n_h crescente in modo che, qualunque sia $k \geq 1$, la successione $\{x^{(n_h)}(t_k)\}_{h\geq 1}$ sia convergente ad un punto y_k. Si noti che per ogni $k \geq 1$ il punto $(y_1, \cdots, y_k) \in K_k$, in quanto definitivamente $(x^{(n)}(t_1), \cdots, x^{(n)}(t_k)) \in K_k$.

Richiamo sul metodo diagonale:
Sia data una successione a due indici

$$\begin{array}{cccc} x_{1,1}, & x_{1,2}, & \cdots & \cdots \\ x_{2,1}, & x_{2,2}, & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ x_{n,1}, & x_{n,2}, & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{array}$$

per cui, qualunque sia n*, la successione* $\{x_{n,h}\}_{h\geq 1}$ *è limitata (o relativamente compatta). Allora esiste una successione* $\{h_m\}_{m\geq 1}$ *di interi per cui, qualunque sia* n*, la successione* $\{x_{n,h_m}\}_{m\geq 1}$ *è convergente.*
Si considera la sottosuccessione convergente $\{x_{1,h_{1,m}}\}_{m\geq 1}$ *ad* y_1*. Poi si considera la successione* $\{x_{2,h_{1,m}}\}_{m\geq 1}$*, che essendo limitata ammette una sottosuccessione convergente* $\{x_{2,h_{2,m}}\}_{m\geq 1}$ *ad* y_2*. Si noti che quindi anche* $\{x_{1,h_{2,m}}\}_{m\geq 1}$ *è ancora una successione convergente ad* y_1 *e che anche ogni sua sottosuccessione è convergente ad* y_1*. Si continua in questo modo fino ad ottenere che* $\{x_{n,h_{n,m}}\}_{m\geq 1}$ *è una successione convergente ad* y_n *e* $\{h_{n,m}\}_{m\geq 1}$ *è una sottosuccessione di* $\{h_{n-1,m}\}_{m\geq 1}$*, ed anche ogni sottosuccessione di* $\{x_{n,h_{n,m}}\}_{m\geq 1}$ *converge ad* y_n*. A questo punto abbiamo una nuova configurazione*

$$\begin{array}{cccc} x_{1,h_{1,1}}, & x_{1,h_{1,2}}, & \cdots & \to y_1 \\ x_{2,h_{2,1}}, & x_{2,h_{2,2}}, & \cdots & \to y_2 \\ \cdots & \cdots & \cdots & \cdots \\ x_{n,h_{n,1}}, & x_{n,h_{n,n}}, & \cdots & \to y_n \end{array}$$

Basta prendere $h_m = h_{m,m}$*, infatti, qualunque sia* n*, la successione* $\{x_{n,h_{m,m}}\}_{m\geq 1}$ *è (definitivamente) una sottosuccessione di* $\{x_{n,h_{n,m}}\}_{m\geq 1}$ *e quindi converge ad* y_n*, per* m *che tende ad infinito.*

Perciò, ogni funzione $x(\cdot)$ di $\mathbb{R}^I$ tale che $x(t_h) = y_h$, per ogni $h \geq 1$, appartiene ad ogni $\mathcal{B}_k$, $k \geq 1$, e quindi ad ogni $\mathcal{A}_k$, $k \geq 1$, e quindi $\bigcap_n \mathcal{A}_n$ risulta un insieme non vuoto.

$\square$

Si noti che l'argomento della dimostrazione è lo stesso della dimostrazione del teorema di Tikhonov: il prodotto infinito di compatti è un compatto. Una dimostrazione simile appare anche nella dimostrazione che l'intersezione di una successione di compatti non vuoti è non vuota.

8.8.1 Caso a tempo discreto: metodo diretto

Passiamo ora ad illustrare il preannunciato metodo costruttivo di dimostrazione del teorema di Kolmogorov, ma prima di tutto, proviamo ad ottenere una variabile aleatoria bidimensionale con funzione di distribuzione $F(x, y)$ data. Notiamo che si può riscrivere

$$F(x, y) = \int_{-\infty}^{x} F_{Y|X}(y \mid z) dF_X(z),$$

dove $F_{Y|X}(y \mid x)|_{x=X(\omega)}$ è una versione regolare delle probabilità $\mathbb{P}(Y \leq y|\sigma(X))$.

Daremo per il momento per scontato che tale scomposizione (o meglio disintegrazione) sia sempre possibile.

Si possono inoltre definire le inverse generalizzate $\Gamma_X(\cdot)$ di $F_X(\cdot)$ e $\Gamma_{Y|X}(\cdot \mid x)$ di $F_{Y|X}(\cdot \mid x)$, qualunque sia x.

Siano ora U e V due v.a. uniformi in $(0, 1)$ ed indipendenti, si definiscano

$$\widetilde{X}(\omega) = \Gamma_X(U(\omega))$$

ed

$$\widetilde{Y}(\omega) = \Gamma_{Y|X}(V(\omega) \mid \widetilde{X}(\omega)) = \Gamma_{Y|X}(V(\omega) \mid \Gamma_X(U(\omega))).$$

È facile verificare che

$$\mathbb{P}(\widetilde{X}(\omega) \leq x, \widetilde{Y}(\omega) \leq y) = F(x, y).$$

Infatti

$$\begin{aligned}
\mathbb{P}(\widetilde{X}(\omega) \leq x, \widetilde{Y}(\omega) \leq y) &= \mathbb{P}(\Gamma_X(U(\omega)) \leq x, \Gamma_{Y|X}(V(\omega) \mid \Gamma_X(U(\omega)) \leq y) = \\
&= \mathbb{P}\{U(\omega) \leq F_X(x), V(\omega) \leq F_{Y|X}(y \mid \Gamma_X(U(\omega))\} = \\
&= \int_{(0, F_X(x)]} F_{Y|X}(y \mid \Gamma_X(u)) du = \\
&\qquad\qquad \text{(applicando il cambio di variabile } z = \Gamma_X(u)) \\
&\qquad\qquad \text{(e considerando che } u \leq F_X(x) \Leftrightarrow \Gamma_X(u) = z \leq x) \\
&= \int_{-\infty}^{x} F_{Y|X}(y \mid z) dF_X(z) = F(x, y).
\end{aligned}$$

È facile generalizzare al caso in dimensione n e costruire una variabile aleatoria n-dimensionale $(X_1, \cdots, X_n)$ con funzione di distribuzione data, a partire da n variabili aleatorie uniformi in $(0, 1)$ ed indipendenti.

A questo punto è chiaro come estendere il procedimento al caso di una successione di v.a. $\{X_n\}_{n\geq 1}$ con distribuzioni finito-dimensionali $\mu_{1,\cdots,n}$ assegnate (e relativa funzione di ripartizione $F(x_1, x_2, \cdots, x_n)$) in modo che

$$\mu_{1,\cdots,n+1}(H \times \mathbb{R}) = \mu_{1,\cdots,n}(H),$$

pur di avere a disposizione una successione di v.a. uniformi in $(0, 1)$ ed indipendenti.

***Infine, va notato che sostanzialmente le condizioni di compatibilità servono solo a garantire che le distribuzioni finito-dimensionali della successione aleatoria così ottenuta, e relative a tempi non consecutivi siano quelle volute.

L'affermazione che una successione di variabili aleatorie $\{X_n\,,\ n \in \mathbb{N}\}$ è una successione di v.a. indipendenti con $\mu_{X_n} = \mu_n$, è un'affermazione che riguarda le distribuzioni finito dimensionali del processo $\{X_n\,,\ n \in \mathbb{N}\}$. L'esistenza di una tale successione si potrebbe quindi dedurre dal teorema di rappresentazione di Kolmogorov, o magari da un risultato *ad hoc* la cui prova fosse la semplificazione del procedimento usato nel dimostrare tale teorema. Tuttavia l'esistenza di una tale successione tuttavia si può dedurre direttamente, pur di dare per scontato che esiste la misura di Lebesgue su $(0, 1)$. Infatti su $(0, 1)$ si possono definire delle variabili aleatorie indipendenti ed identicamente distribuite, a valori nell'insieme $\{0, 1\}$, e che assumono il valore 0 con probabilità 1/2 (lo stesso vale per il valore 1). A partire da questa successione di variabili aleatorie si può costruire una successione di variabili aleatorie $\{U_j\,,\ j \in \mathbb{N}\}$ indipendenti ed uniformi in $(0, 1)$, come descritto qui di seguito. Infine, posto $F_n(x) = \mu_n\big((-\infty, x]\big)$, la successione cercata è data dalla successione delle v.a. $F_n^{-1}(U_n)$.

Lemma 8.6 (Successioni di v.a. indipendenti uniformi in $(0,1)$: esistenza). *Nello spazio $\Omega = (0,1)$ con la misura di Lebesgue sui boreliani, è possibile avere una successione di v.a. uniformi in $(0,1)$ ed indipendenti.*

***

Per costruire tale successione si ricordi che scrivendo $\omega \in (0,1)$ in forma diadica $\omega = \sum_1^\infty W_i(\omega)\frac{1}{2^i}$, le v.a. W_i risultano indipendenti e $\mathbb{P}(W_i = 0) = \mathbb{P}(W_i = 1) = \frac{1}{2}$. La successione U_n di v.a. uniformi ed indipendenti si può costruire, a partire dalle v.a. $\{W_i\}$, riordinandole in modo che formino una sequenza a doppio indice $\{\widetilde{W}_{i,n}\}$ così da poter definire

$$U_n(\omega) = \sum_{i=0}^{\infty} \widetilde{W}_{i,n}(\omega)\,\frac{1}{2^i}.$$

Ad esempio si può prendere ***$\widetilde{W}_{i,n}$ $= W_{2^{i-1}(2n+1)}$, che corrisponde a riordinare la successione $\{W_i\}$ in questo modo:

$$\begin{array}{cccccc}
W_1 & W_3 & W_5 & W_7 & W_9 & \cdots \\
W_2 & W_6 & W_{10} & W_{14} & \cdots & \\
W_4 & W_{12} & W_{20} & W_{28} & \cdots & \\
W_8 & W_{24} & W_{40} & \cdots & & \\
\vdots & \vdots & \vdots & & &
\end{array}$$

ottenendo da $\{W_i\}$ infinite sottosuccessioni (corrispondenti alle colonne di questa matrice) in modo tale che nessuna W_i venga tralasciata né ripetuta.
***

***Ribadiamo che questa costruzione è riportata affinché ***sia chiaro che l'affermazione che esiste una successione di v.a. indipendenti ed uniformi in $(0,1)$, non dipende dal Teorema di Kolmogorov dimostrato precedentemente.

8.8.2 Osservazione su $\mathcal{R}^I$ ‡

La σ-algebra $\mathcal{R}^I$ coincide con la famiglia degli insiemi del tipo

$$\mathcal{A} = \{x(\cdot) \mid \{x(t_k)\}_k \in \mathcal{N}\}, \tag{8.22}$$

al variare delle successioni $S = \{t_k\}_k \subseteq I$ ed $\mathcal{N} \in \mathcal{R}^{\mathbb{N}}$.

Anzi, più in generale, dato un processo, cioè una famiglia di variabili aleatorie $(Y_t, t \in I)$ in uno spazio $(\Omega, \mathcal{F}, \mathbb{P})$, e posto $\mathcal{F}_S = \sigma\{Y_s, \text{ per } s \in S\}$, valgono le seguenti affermazioni:

(i) Se $A \in \mathcal{F}_I, \omega \in A$ e $Y_t(\omega) = Y_t(\omega')$ per ogni $t \in I$, allora $\omega' \in A$.

(ii) Se $A \in \mathcal{F}_I$, allora esiste un $S \subseteq$I e numerabile per cui $A \in \mathcal{F}_S$, ovvero $\mathcal{F}_I = \bigcup_S \mathcal{F}_S$, dove l'unione è fatta su tutti i sottoinsiemi S numerabili di I.

Entrambe le due proprietà si basano sul fatto che, posto $\xi : \Omega \to \mathbb{R}^I, \omega \to \xi(\omega)(\cdot)$, dove $\xi(\omega)(t) = Y_t(\omega)$ per ogni $t \in I$, allora

$$\mathcal{F}_I = \{\xi^{-1}(M), \quad M \in \mathcal{R}^I\}.$$

Infatti allora $Y_t(\omega) = Y_t(\omega')$ per ogni $t \in I$ equivale a $\xi(\omega) = \xi(\omega')$ e quindi $\omega \in A = \xi^{-1}(M)$, cioè $\xi(\omega) \in M$ implica $\xi(\omega') \in M$ e quindi $\omega' \in A$. Infatti ogni $\sigma -$ algebra rispetto alla quale ξ è misurabile deve contenere le controimmagini dei cilindri e quindi deve contenere anche $\{\xi^{-1}(M), M \in \mathcal{R}^S\}$ per ogni S numerabile, ed inoltre la famiglia $\bigcup_S \mathcal{F}_S$ è una σ-algebra, come è facile verificare.

L'unico punto in cui è necessaria un po' di attenzione è il caso di unioni numerabili di eventi $E_n \in \bigcup_S \mathcal{F}_S$. Sia S_n tale che $E_n \in \mathcal{F}_{S_n}$. Poiché $\mathcal{F}_{S_n} \subseteq \mathcal{F}_{\cup_m S_m}$ ed $\mathcal{F}_{\cup_m S_m}$ è una $\sigma-$algebra, ovviamente $\bigcup_n E_n \in \mathcal{F}_{\cup_m S_m} \subseteq \bigcup_S \mathcal{F}_S$.

In particolare $\mathcal{R}^I = \bigcup_S \mathcal{R}^S$, dove l'unione è fatta su tutti i sottoinsiemi S numerabili di I.

8.8.3 Problemi con lo spazio canonico

Come visto nella precedente sezione $\mathcal{R}^I = \bigcup_S \mathcal{R}^S$, dove l'unione è fatta su tutti i sottoinsiemi S numerabili di I. Consideriamo il caso in cui I è un insieme continuo e per fissare le idee $I = [0, T]$ oppure $[0, \infty)$. Allora non ha senso considerare la probabilità che le traiettorie abbiano particolari proprietà tipo siano crescenti, o continue, in quanto tali insiemi non sono misurabili nello spazio canonico non essendo esprimibili come in (8.22): non è possibile trovare una successione di tempi (ovvero un sottoinsieme numerabile di tempi) e condizioni relative solo a tali istanti per determinare se una funzione è continua oppure no. In altre parole: comunque sia data una successione di tempi esistono funzioni crescenti (o continue) su tale successione di tempi, ma non su tutto I.

Tipicamente, al massimo si può sperare di ottenere un processo $(X'_t, t \in I)$, che abbia le proprietà richieste, e che sia ***stocasticamente equivalente a*** $(X_t, t \in I)$, ovvero una ***versione del processo*** $(X_t, t \in I)$[41] ed ha quindi le stesse distribuzioni finito-dimensionali di $(X_t, t \in I)$

Per proprietà tipo la continuità si può al massimo sperare di procedere nel seguente modo: si dimostra che, se ristretto ad un opportuno insieme numerabile di tempi S, il processo X_t risulta a traiettorie continue, tranne a parte un eventuale insieme di probabilità nulla. Si definisce poi un nuovo processo X'_t definito come limite delle traiettorie ristrette ad S. Si spera poi di dimostrare che il processo così ottenuto abbia le stesse distribuzioni finito-dimensionali del processo X_t.

Problemi di questo tipo sono connessi al problema della separabilità.Gli interessati possono consultare il libro di Billingsley, Probability and measures [3].

A titolo di esempio vediamo come si può procedere con il processo di Poisson, con $I = [0, 1]$. Il Teorema di Kolmogorov assicura che esiste una misura di probabilità $\mathbb{P}$ su $(\mathbb{R}^I, \mathcal{R}^I)$, in modo che il processo $X_t(x(\cdot)) = x(t)$ abbia le distribuzioni finito-dimensionali uguali a quelle del processo di Poisson. Ovviamente non possiamo dire nulla sulle traiettorie, ad esempio non possiamo affermare che siano non decrescenti, costanti a tratti e con salti unitari. Proviamo a mostrare come procedere per ottenere una versione X'_t con traiettorie non decrescenti. Consideriamo l'insieme D dei diadici. La condizione che $X_t|_D$ sia a traiettorie non decrescenti e a valori in $\mathbb{N}$ diviene $X_t|_D \in \mathcal{I}$, dove

$$\mathcal{I} = \bigcap_n \bigcap_{0 \le i < 2^n} \{x(\cdot) \text{ tali che } x(\frac{i}{2^n}) \in \mathbb{N} \text{ e } x(\frac{i}{2^n}) \le x(\frac{i+1}{2^n})\}.$$

Inoltre

$$\mathbb{P}(\bigcap_{0 \le i < 2^n} \{x(\cdot) \text{ tali che } x(\frac{i}{2^n}) \in \mathbb{N} \text{ e } x(\frac{i}{2^n}) \le x(\frac{i+1}{2^n})\}) = 1 \quad \text{per ogni } n,$$

e quindi vale anche che $(X_t, t \in D)$ è a traiettorie non decrescenti con probabilità 1.
Si noti inoltre che

$$\lim_{s \downarrow t, s \in D} x(s)$$

esiste per ogni $x(\cdot)$ appartenente a $\mathcal{I}$.

---

[41] Si ricordi che , dato un processo $(X_t, t \in I)$. Un processo $(X'_t, t \in I)$ si dice ***stocasticamente equivalente a*** $(X_t, t \in I)$ se

$$\mathbb{P}(X'_t \neq X_t) = 0 \text{ per ogni } t \ge 0.$$

In tale caso si dice anche X'_t è una ***versione*** di X_t.

Si definisca ora

$$\begin{cases} X'_t(x(\cdot)) := X_t(x(\cdot)) & \text{se } x(\cdot) \in \mathcal{I} \text{ e } t \in D \\ X'_t(x(\cdot)) := \lim\limits_{s \downarrow t, s \in D} X_s(x(\cdot)) = \lim\limits_{s \downarrow t, s \in D} x(s), & \text{se } x(\cdot) \in \mathcal{I} \text{ e } t \notin D \\ X'_t(x(\cdot)) := 0 & \text{se } x(\cdot) \notin \mathcal{I} \text{ e per ogni } t \in (0,1) \end{cases}$$

Inoltre ovviamente $\{X'_t = X_t\}$ è tutto $\mathbb{R}^I$, per $t \in D$, ed è, se $t \notin D$,

$$\bigcap_{\epsilon \in \mathbb{Q}^+} \bigcup_{\delta \in \mathbb{Q}^+} \bigcap_{\substack{t \le s \le t+\delta \\ s \in D}} \{x(\cdot) \text{ tali che } \mid x(s) - x(t) \mid < \epsilon\}$$

e, se intersecato con $\mathcal{I}$, coincide con

$$\bigcup_{\delta \in \mathbb{Q}^+} \bigcap_{\substack{t \le s \le t+\delta \\ s \in D}} \{x(\cdot) \text{ tali che } \mid x(s) - x(t) \mid = 0\}.$$

Infine [42]

$$\mathbb{P}(\bigcap_{\substack{t \le s \le t+\delta \\ s \in D}} \{x(\cdot) \text{ tali che } \mid x(s) - x(t) \mid = 0\}) \ge \exp\{-\lambda\delta\} \to 1$$

Di conseguenza $\mathbb{P}(X'_t \ne X_t) = 0$ per ogni $t \ge 0$, ovvero il processo X'_t è una versione del processo X_t, ed ha quindi le stesse distribuzioni finito-dimensionali di X_t e gode della proprietà di monotonia, continuità a destra delle traiettorie e di essere a valori in $\mathbb{N}$.

8.9 Appendice: dimostrazione del criterio di Chensov-Kolmogorov †‡

Per comodità del lettore riportiamo in appendice sia la definizione di funzione hölderiana che l'enunciato del criterio (Proposizione 8.3).

Una funzione $f(x)$ è **hölderiana** di esponente γ se per ogni x esistono un $\delta(x) > 0$ e un $L_\gamma(x)$ tali che, per ogni y per il quale $|y - x| \le \delta(x)$, si abbia

$$|f(x) - f(y)| \le L_\gamma(x)|x - y|^\gamma.$$

Nel caso in cui $\delta(x)$ e $L_\gamma(x)$ possano essere presi in modo indipendente da x, per $x \in I$, si dice che f è **uniformemente hölderiana** nell'insieme I.

Proposizione 8.3 (Criterio di Chensov-Kolmogorov) *Sia X_t un processo per cui esistono α, β e C strettamente positivi, per cui*

$$\mathbb{E}[|X_t - X_s|^\beta] \le C|t - s|^{1+\alpha}.$$

Allora esiste una versione $\widetilde{X}_t$ di X_t, a traiettorie continue. Inoltre le traiettorie sono uniformemente hölderiane di esponente γ, per ogni $\gamma < \frac{\alpha}{\beta}$, in ogni intervallo limitato.

---

[42] Si osservi che

$$\mathbb{P}(\bigcap_{\substack{t \le s \le t+\delta \\ s \in D}} \{x(\cdot) \text{ tali che } \mid x(s) - x(t) \mid = 0\}) = \lim_{n \to \infty} \mathbb{P}(\bigcap_{\substack{t \le s \le t+\delta \\ s \in D_n}} \{x(\cdot) \text{ tali che } \mid x(s) - x(t) \mid = 0\})$$

e che

$$* * * \mathbb{P}(\bigcap_{\substack{t \le s \le t+\delta \\ s \in D_n}} \{x(\cdot) \text{ tali che } \mid x(s) - x(t) \mid = 0\}) \ge * * * \mathbb{P}(x(t) = x(i/2^n) = x(t+\delta), \text{ per ogni } i \text{ tale che } t \le i/2^n \le t+\delta) = \exp\{-\lambda\delta\}.$$

Dimostrazione. Per semplicità consideriamo il caso dell'intervallo $[0,1]$. Sia come al solito D l'insieme dei diadici e $D_n = \{s = \frac{k}{2^n}$, per $0 \le k \le 2^n\}$. Si definisca Γ_n

$$\Gamma_n = \{\omega \text{ t.c. } \max_{k \le 2^n - 1} \left| X_{k/2^n} - X_{(k+1)/2^n} \right| > 2^{-n\gamma}\}$$

Se mostriamo che

$$\sum_{n=1}^{\infty} \mathbb{P}(\Gamma_n) < +\infty,$$

allora, per il Lemma di Borel-Cantelli,

$$\mathbb{P}(\cap_{m \ge 1} \cup_{n \ge m} \Gamma_n) = 0$$

ovvero

$$\mathbb{P}(\cup_{m \ge 1} \cap_{n \ge m} \cap \Gamma_n^c) = 1.$$

Ciò significa che esiste un evento $\mathcal{N}$, con $\mathbb{P}(\mathcal{N}) = 0$, e che per ogni $\omega \in \mathcal{N}^c$ esiste un $m = m(\omega)$ tale che per ogni $n \ge m = m(\omega)$

$$\left| X_{k/2^n} - X_{(k+1)/2^n} \right| \le 2^{-n\gamma}.$$

Mostreremo ora che per $\omega \in \mathcal{N}^c$ le traiettorie sono uniformemente hölderiane su D con

$$\delta = \delta(\omega) = 1/2^{m(\omega)} \text{ ed } L_\gamma = \Big(\frac{2}{2^\gamma - 1} + 1 \Big) 2^\gamma,$$

e quindi sono hölderiane su D per quasi ogni ω.

Sia quindi $\omega \in \mathcal{N}^c$.

Cominciamo con l'osservare che se

$$r \in [j/2^m, (j+1)/2^m), \text{ con } m \ge m(\omega), \text{ ed} r \in D,$$

allora

$$r = j/2^m + \sum_{h=m+1}^{m+i} \alpha_h \frac{1}{2^h},$$

per un $i \ge 0$, con $\alpha_h \in \{0,1\}$. Posto

$$r_0 = j/2^m \text{ ed } r_s = j/2^m + \sum_{h=m+1}^{m+s} \alpha_h \frac{1}{2^h},$$

si ha che $r = r_i$ e che

$$\left| X_{j/2^m} - X_r \right| \le \sum_{s=1}^{i} \left| X_{r_s} - X_{r_{s-1}} \right| \le \sum_{h=m+1}^{m+i} 2^{-h\gamma} \le 2^{-m\gamma} \sum_{l=1}^{i} 2^{-l\gamma} \le 2^{-m\gamma} \frac{1}{2^\gamma - 1}.$$

Se invece $|r - u| \le 1/2^m$ con $r, u \in D$, allora sono possibili due casi:

(i) esiste j **per cui** $r, u \in [j/2^m, (j+1)/2^m)$, e allora

$$|X_u - X_r| \le \left| X_{j/2^m} - X_u \right| + \left| X_{j/2^m} - X_r \right| \le \frac{2}{2^\gamma - 1} 2^{-m\gamma}$$

(ii) esiste j **per cui** $r \in [j/2^m, (j+1)/2^m)$ **ed** $u \in [(j+1)/2^m, (j+2)/2^m)$, e allora

$$|X_u - X_r| \le \left| X_{(j+1)/2^m} - X_u \right| + \left| X_{j/2^m} - X_{(j+1)/2^m} \right| + \left| X_{j/2^m} - X_r \right| \le \left(\frac{2}{2^\gamma - 1} + 1 \right) 2^{-m\gamma}$$

Quindi, per quasi ogni ω, si ha definitivamente, per ogni $r, u \in D$, con $|r-u| \leq 1/2^m$,

$$|X_u - X_r| \leq M_\gamma 2^{-m\gamma},$$

con $M_\gamma = \frac{2}{2^\gamma - 1} + 1$. In particolare se $|r-u| \leq 1/2^m(\omega)$,ed $m \geq m(\omega)$ è tale che $1/2^{m+1} \leq |r-u| \leq 1/2^m$, allora

$$|X_u - X_r| \leq L_\gamma 2^{-(m+1)\gamma} \leq L_\gamma |r-u|^\gamma.$$

Avremo quindi, come preannunciato, che le traiettorie sono uniformemente hölderiane su D con $\delta = \delta(\omega) = 1/2^{m(\omega)}$ ed $L_\gamma = \left(\frac{2}{2^\gamma - 1} + 1\right) 2^\gamma$, per quasi ogni ω. [1]

Si definisce, sempre per $\omega \in \mathcal{N}^c$, $\widetilde{X}_t(\omega)$ come il limite di $X_{r_n}(\omega)$ per una successione r_n di diadici convergente a t, e per $\omega \in \mathcal{N}$ si può definire, ad esempio, $\widetilde{X}_t \equiv 0$. È facile vedere che la proprietà di hölderianità si conserva. Il fatto che si tratta di una versione di X_t, deriva dal fatto che per l'ipotesi fatta X_t è continuo in L^β e quindi è continuo in probabilità e quindi contemporaneamente si ha

$$X_{r_n} \stackrel{Pr}{\to} X_t \text{ e } X_{r_n} \stackrel{q.c.}{\to} \widetilde{X}_t,$$

da cui subito $\mathbb{P}(X_t = \widetilde{X}_t) = 1$.

A questo punto rimane solo da controllare che $\sum_{n=1}^\infty \mathbb{P}(\Gamma_n) < +\infty$, e infatti:

$$\sum_{n=1}^\infty \mathbb{P}(\Gamma_n) \leq \sum_{n=1}^\infty \sum_{k=0}^{2^n-1} \mathbb{P}(\left|X_{k/2^n} - X_{(k+1)/2^n}\right| > 2^{-n\gamma}) \leq \sum_{n=1}^\infty \sum_{k=0}^{2^n-1} \mathbb{E}[\left|X_{k/2^n} - X_{(k+1)/2^n}\right|^\beta]/2^{-n\gamma\beta} \leq$$

$$\leq \sum_{n=1}^\infty \sum_{k=0}^{2^n-1} (1/2^n)^{1+\alpha} 2^{n\gamma\beta}) = \sum_{n=1}^\infty 2^n (1/2^n)^{1+\alpha} 2^{n\gamma\beta}) = \sum_{n=1}^\infty 2^{-n(\alpha-\gamma\beta)} < +\infty$$

purché $\alpha - \gamma\beta > 0$, e ciò è vero per come abbiamo preso γ. (Si noti che è fondamentale nella prova che $1+\alpha > 1$, in quanto altrimenti non ci sarebbe speranza di ottenere nessun tipo di convergenza ovvero la maggiorazione di $\mathbb{P}(\Gamma_n)$ non avrebbe nessun significato in quanto risulterebbe un numero maggiore di 1)

□

[1] Si noti che questa parte della dimostrazione è puramente deterministica: nel senso che se una funzione $x(\cdot)$ definita sui diadici di $[0,1]$, gode della proprietà che esiste un m tale che per ogni $n \geq m$ e per $k < 2^n$ si ha $|x(k/2^n) - x((k+1)/2^n)| \leq 2^{-n\gamma}$, allora $x(\cdot)$ è hölderiana sui diadici di $[0,1]$

Capitolo 9

Proprietà del moto browniano

Abbiamo visto che il moto browniano è un particolare processo gaussiano a incrementi indipendenti e omogenei con media nulla e varianza t. Abbiamo anche visto che per la proprietà delle v.a. gaussiane per cui non correlazione ed indipendenza sono equivalenti, un modo alternativo di definire è attraverso la funzione di correlazione $Cov(W_t, W_s) = t \wedge s$. Inoltre sappiamo che W_t è una martingala rispetto alla filtrazione naturale $\mathcal{F}_t^W$.

9.1 Trasformazioni del moto browniano ‡

Nei seguenti casi, dato un moto browniano W_t, si costruisce un altro processo che risulta ancora un moto browniano:

1) Per ogni $s \geq 0, \widetilde{W}_t := W_{s+t} - W_s$ è un moto browniano ed è una martingala rispetto alla filtrazione $\mathcal{G}_t := \mathcal{F}_{s+t}^W$.

2) Per ogni $c \in \mathbb{R}$, $W_t^{(c)} := cW_{c^2 t}$ è un moto browniano ed è una martingala rispetto alla filtrazione $\mathcal{G}_t^{(c)} := \mathcal{F}_{c^2 t}^W$ (in particolare per $c = -1$ si ha che $-W_t$ è ancora un moto browniano rispetto a $\mathcal{F}_t^W$).

3) Il processo definito da $\widehat{W}_t := tW_{1/t}$, per $t > 0$, e $Z_0 := 0$, è ancora un moto browniano, rispetto alla sua filtrazione naturale.

Si supponga ora di prendere una versione continua di W_t.

4) Per ogni tempo d'arresto limitato τ il processo $\widetilde{W}_t^{(\tau)} := W_{\tau+t} - W_\tau$ è ancora un moto Browniano rispetto alla filtrazione $\mathcal{F}_{\tau+t}^W$.

Le proprietà 1), 2) e 3) sono di immediata verifica. La proprietà 4), invece, non è immediata, e corrisponde a quella proprietà che, nei processi di Markov, viene detta Proprietà di Markov Forte.

9.2 Proprietà di Markov forte per il processo di Wiener‡

In questa sezione si suppone di prendere una versione continua di W_t.

Proposizione 9.1. *Sia τ un tempo d'arresto finito con probabilità 1. Allora il processo $Y_t := W_{t+\tau} - W_\tau$ è un processo di Wiener standard ed è indipendente da $\mathcal{F}_\tau$.*

Come conseguenza, per ogni funzionale Φ, che sia $\mathcal{R}^{[0,\infty)}$-misurabile delle traiettorie

$$\mathbb{E}[\Phi(W_s, s \geq \tau) \mid \mathcal{F}_\tau] = \mathbb{E}[\Phi(W_s, s \geq \tau) \mid W_\tau] = \mathbb{E}[\Phi(w + Y_u, u \geq 0)] \mid_{w=W_\tau} .$$

In questo senso la proprietà 4), ovvero la Proposizione 9.1, è una generalizzazione della proprietà di Markov, che invece di valere solo per tempi deterministici t vale anche per tempi d'arresto τ.

Più precisamente

Definizione 9.1 (Proprietà di Markov). *La* **proprietà di Markov** *(***rispetto ad una filtrazione** $\mathcal{F}_t \supseteq \mathcal{F}_t^X$*) per un generico processo* X_t *è la seguente:*

$$\mathbb{P}(X_{t+s} \in A \mid \mathcal{F}_t) = \mathbb{P}(X_{t+s} \in A \mid X_t) \qquad \textit{per ogni } s, t \geq 0.$$

Si osservi che, nel caso dei processi di Markov regolari

$$\mathbb{P}(X_{t+s} \in A \mid X_t) = p(t, t+s, x, A) \mid_{x=X_t},$$

dove $P(u, v, x, \cdot)$ sono le probabilità di transizione regolari.

Definizione 9.2 (Proprietà di Markov forte). *La* **proprietà di Markov forte**[1] **rispetto ad una filtrazione** $\mathcal{F}_t \supseteq \mathcal{F}_t^X$*) è invece*

$$\mathbb{P}(X_{\tau+s} \in A \mid \mathcal{F}_\tau) = \mathbb{P}(X_{\tau+s} \in A \mid X_\tau) \qquad \textit{per ogni tempo d'arresto finito } \tau \textit{ ed } s \geq 0.$$

Si noti che è necessario che τ sia finito affinché abbia senso X_τ ed è necessario che X_τ sia una variabile aleatoria, e che sia $\mathcal{F}_\tau$-misurabile.

Questa proprietà di misurabilità è sempre verificata per processi X_t *cadlag*, in realtà potrebbe bastare un poco meno e precisamente la progressiva misurabilità)

Dimostrazione. (della Proposizione 9.1, ovvero della proprietà di Markov forte per il processo di Wiener standard)

La tesi equivale a richiedere che per ogni $k \geq 1, 0 \leq t_1 \leq \cdots \leq t_k, A_1, \cdots, A_k$ boreliani di $\mathbb{R}$, e per ogni $C \in \mathcal{F}_\tau$

$$\mathbb{P}(Y_{t_1} \in A_1, \cdots, Y_{t_k} \in A_k, C) = \mathbb{P}(W_{t_1} \in A_1, \cdots, W_{t_k} \in A_k)\mathbb{P}(C)$$

I CASO: τ assume al più un'infinità di valori $\{s_h\}_{h\geq 1}$.

$$\begin{aligned}
\mathbb{P}(Y_{t_1} \in A_1, \cdots, Y_{t_k} \in A_k, C) &= \sum_{h=1}^{\infty} \mathbb{P}(Y_{t_1} \in A_1, \cdots, Y_{t_k} \in A_k, C, \tau = s_h) \\
&= \sum_{h=1}^{\infty} \mathbb{P}(W_{t_1+s_h} - W_{s_h} \in A_1, \cdots, W_{t_k+s_h} - W_{s_h} \in A_k, C, \tau = s_h) \\
&\qquad (C \cap \{\tau = s_h\} \in \mathcal{F}_{s_h}, \text{ poiché } C \in \mathcal{F}_\tau \text{ e } W_t \text{ ha incrementi indipendenti}) \\
&= \sum_{h=1}^{\infty} \mathbb{P}(W_{t_1+s_h} - W_{s_h} \in A_1, \cdots, W_{t_k+s_h} - W_{s_h} \in A_k)\mathbb{P}(C, \tau = s_h) \\
&\qquad (W \text{ ha incrementi omogenei}) \\
&= \sum_{h=1}^{\infty} \mathbb{P}(W_{t_1} \in A_1, \cdots, W_{t_k} \in A_k)\mathbb{P}(C, \tau = s_h) = \mathbb{P}(W_{t_1} \in A_1, \cdots, W_{t_k} \in A_k)\mathbb{P}(C)
\end{aligned}$$

[1] Si faccia attenzione che la proprietà di Markov forte non è sempre verificata per un generico processo di Markov.

II CASO: τ generico.

Per individuare la distribuzione congiunta in realtà basta verificare che per ogni $k \geq 1$, per ogni funzione continua e limitata $\Phi(y_1, \cdots, y_k)$ e per ogni $C \in \mathcal{F}_\tau$

$$\mathbb{E}[\Phi(Y_{t_1}, \cdots, Y_{t_k}) I_C] = \mathbb{E}[\Phi(W_{t_1}, \cdots, W_{t_k})]\mathbb{P}(C)$$

in quanto le funzioni continue determinano univocamente la legge di una variabile aleatoria. Per il Caso I, presa $\{\tau_m\}$ una successione[2] di tempi d'arresto, tale che $\tau_m \geq \tau$ e $\tau_m \to \tau$

$$\mathbb{E}[\Phi(W_{t_1+\tau_m} - W_{\tau_m}, \cdots, W_{t_k+\tau_m} - W_{\tau_m}) I_C] = \mathbb{E}[\Phi(W_{t_1}, \cdots, W_{t_k})]\mathbb{P}(C)$$

Essendo le traiettorie di W_t continue (ma basterebbe che fossero continue a destra), possiamo affermare che $(W_{t_{h+\tau_m}} - W_{\tau_m}) \to \ (W_{t_{h+\tau}} - W_\tau) = Y_{t_h}$ e quindi

$$\Phi(W_{t_1+\tau_m} - W_{\tau_m}, \cdots, W_{t_k+\tau_m} - W_{\tau_m}) \to \Phi(Y_{t_1}, \cdots, Y_{t_k}).$$

Essendo Φ limitata si può passare al limite sotto il segno di media.

Osservazione 9.1. *È evidente che questa dimostrazione (e quindi la proprietà di Markov forte) rimane valida se al posto del moto browniano si mette un qualsiasi processo ad incrementi indipendenti ed omogenei, con traiettorie* cadlag.

9.3 Principio di riflessione ‡

Sotto questo nome vanno diverse proprietà ma la più nota è la seguente. Per $a \geq 0$, sia

$$\tau_a = \inf\{t > 0 \text{ t.c. } W_t \geq a\},$$

cioè il tempo di prima uscita da $(-\infty, a)$, che è un tempo d'arresto se come al solito prendiamo la versione continua di W_t. Allora

$$\mathbb{P}(\tau_a \leq t) = 2\mathbb{P}(W_t \geq a). \tag{9.1}$$

Dalla precedente eguaglianza si può ricavare che allora

$$\mathbb{P}(\tau_a \leq t) = 2\mathbb{P}(W_t \geq a) = \frac{2}{\sqrt{2\pi t}} \int_a^\infty \exp\{-\frac{x^2}{2t}\}dx = \sqrt{\frac{2}{\pi}} \int_{a/\sqrt{t}}^\infty \exp\{-y^2/2\}dy.$$

Per $a < 0$ si definisce $\tau_a = \inf\{t > 0 \text{ t.c. } W_t \leq a\}$.

Notando che $\tau_a = \inf\{t > 0 \text{ t.c. } -W_t \geq -a\}$ e che $-W_t$ è ancora un moto browniano si ottiene che per ogni $a \in \mathbb{R}$

$$\mathbb{P}(\tau_a \leq t) = 2\mathbb{P}(W_t \geq |a|) = \frac{2}{\sqrt{2\pi t}} \int_{|a|}^\infty \exp\{-\frac{x^2}{2t}\}dx = \sqrt{\frac{2}{\pi}} \int_{|a|/\sqrt{t}}^\infty \exp\{-y^2/2\}dy.$$

e quindi la sua densità è

$$g_{\tau_a}(t) = \sqrt{\frac{1}{2\pi}} \frac{|a|}{t^{3/2}} \exp\{-a^2/2t\}.$$

[2]Per la proprietà **2)** dei tempi di arresto (Sezione 4.4), per ogni tempo d'arresto esiste sempre una successione con tali proprietà.

e il suo valore atteso è infinito:

$$\mathbb{E}[\tau_a] = \int_0^\infty t\sqrt{\frac{1}{2\pi}}\frac{|a|}{t^{3/2}}\exp\{-a^2/2t\}\ dt = \int_0^\infty \sqrt{\frac{1}{2\pi}}\frac{|a|}{t^{1/2}}\exp\{-a^2/2t\}\ dt = \infty$$

$$\begin{aligned}\int_0^\infty \sqrt{\frac{1}{2\pi}}\frac{|a|}{t^{1/2}}\exp\{-a^2/2t\}\ dt &\geq \int_1^\infty \sqrt{\frac{1}{2\pi}}\frac{|a|}{t^{1/2}}\exp\{-a^2/2t\}\ dt\\ &\geq \int_1^\infty \sqrt{\frac{1}{2\pi}}\frac{|a|}{t^{1/2}}\exp\{-a^2/2\}\ dt = \infty\end{aligned}$$

Inoltre questa proprietà permette di calcolare ***la distribuzione del massimo di un moto browniano*** (per $a \geq 0$):

$$\mathbb{P}(\sup_{s\leq t}(W_s) \geq a) = \mathbb{P}(\tau_a \leq t) = 2\mathbb{P}(W_t \geq a).$$

Dimostrazione intuitiva del principio di riflessione (9.1)

$$\begin{aligned}&\mathbb{P}(W_t \geq a) = \mathbb{P}(W_t \geq a \mid \tau_a \leq t)\mathbb{P}(\tau_a \leq t) + \mathbb{P}(W_t \geq a \mid \tau_a > t)\mathbb{P}(\tau_a > t)\\ &= \mathbb{P}(W_t \geq a \mid \tau_a \leq t)\mathbb{P}(\tau_a \leq t) = \frac{1}{2}\mathbb{P}(\tau_a \leq t),\end{aligned}$$

in quanto ovviamente

$$\mathbb{P}(W_t \geq a \mid \tau_a > t) = 0 \text{ e } \mathbb{P}(W_t \geq a \mid \tau_a \leq t) = \mathbb{P}(W_t \leq a \mid \tau_a \leq t) = \frac{1}{2}.$$

Nell'ultima uguaglianza è implicito l'uso del fatto che l'informazione contenuta nell'evento $\{\tau_a \leq t\}$ è riassumibile nel fatto che $\{W_{\tau_a} = a\}$ e che $W_{\tau_a+s} - W_{\tau_a}$ è ancora un moto browniano. In realtà formalmente c'è qualche problema in quanto pur essendo $\{\tau_a \leq t\} \in \mathcal{F}_\tau$, e quindi

$$\begin{aligned}\mathbb{P}(W_t \geq a \mid \tau_a \leq t) &= \mathbb{P}(W_t \geq W_{\tau_a} \mid \tau_a \leq t) = \mathbb{P}(W_t - W_{\tau_a} \geq 0 \mid \tau_a \leq t) =\\ &= \mathbb{P}(Y_{t-\tau_a} \geq 0 \mid \tau_a \leq t) = \mathbb{P}(Y_{t-\tau_a} \geq 0)\end{aligned}$$

essendoci di mezzo τ_a non è chiarissimo, anche se intuitivo che $\mathbb{P}(Y_{t-\tau_a} \geq 0) = \frac{1}{2}$.

Dimostrazione formale del principio di riflessione (9.1) Faremo vedere che le trasformate di Laplace rispetto a t, di $\mathbb{P}(W_t \geq a)$ e di $\frac{1}{2}\mathbb{P}(\tau_a \leq t)$ sono uguali, utilizzando il fatto che $W_{\tau_a} = a$, per la continuità delle traiettorie di W_t, e che $W_{\tau+s} - W_\tau$ è indipendente da $\mathcal{F}_\tau$, e quindi da τ: qualunque sia $\alpha \geq 0$,

$$\begin{aligned}&\int_0^\infty e^{-\alpha t}\mathbb{P}(W_t \geq a)dt && \text{(è necessario usare la congiunta misurabilità}\\ & && \text{in } (t,\omega) \text{ di } W_t\text{, per scambiare gli integrali)}\\ &= \mathbb{E}[\int_0^\infty e^{-\alpha t} I_{[a,+\infty)}(W_t)dt]\\ & && \text{(tenendo conto che } W_t < a \text{ per } t < \tau_a)\\ &= \mathbb{E}[\int_{\tau_a}^\infty e^{-\alpha t} I_{[a,+\infty)}(W_t)dt] = \mathbb{E}[\int_{\tau_a}^\infty e^{-\alpha\tau_a} e^{-\alpha(t-\tau_a)} I_{[0,+\infty)}(W_t - W_{\tau_a})dt]\\ & && (W_t - W_{\tau_a} = Y_{t-\tau_a} \text{ e cambio di variabile})\\ &= \mathbb{E}[e^{-\alpha\tau_a}\int_0^\infty e^{-\alpha s} I_{[0,+\infty)}(Y_s)ds]\\ & && (\{Y_t, t \geq 0\} \text{ e } \tau_a \text{ sono indipendenti})\\ &= \mathbb{E}[e^{-\alpha\tau_a}]\mathbb{E}[\int_0^\infty e^{-\alpha s} I_{[0,+\infty)}(Y_s)ds] = \mathbb{E}[e^{-\alpha\tau_a}]\int_0^\infty e^{-\alpha s}\mathbb{P}(Y_s \geq 0)ds]\\ &= \frac{1}{2\alpha}\mathbb{E}[e^{-\alpha\tau_a}].\end{aligned}$$

In modo analogo

$$\int_0^\infty e^{-\alpha t}\mathbb{P}(\tau_a \le t)dt = \mathbb{E}[\int_0^\infty e^{-\alpha t} I_{\{\tau_a \le t\}}dt] = \mathbb{E}[\int_{\tau_a}^\infty e^{-\alpha t}dt] = \frac{1}{\alpha}\mathbb{E}[e^{-\alpha\tau_a}].$$

$\square$

In alcuni testi viene dato come principio di riflessione la seguente proprietà (la cui dimostrazione è simile alla dimostrazione della proprietà di Markov forte):

Principio di riflessione: seconda formulazione
Per ogni tempo d'arresto τ finito, il processo "riflesso"

$$\begin{cases} Y_t = W_t & \text{per } t \le \tau, \\ Y_t - Y_\tau = -(W_t - W_\tau) & \text{per } t > \tau, \end{cases}$$

è ancora un moto browniano.

9.4 Tempi di uscita da una striscia ‡

Anche nel caso del moto Browniano si possono ottenere delle applicazioni simili a quelle della rovina del giocatore. Più in generale si può considerare il processo di Wiener con coefficiente di drift μ e coefficiente di diffusione σ^2, cioè

$$X_t := \sigma W_t + \mu t,$$

e il tempo τ di prima uscita da una striscia

$$\tau \equiv \tau(-a,b) := \inf\{t > 0 : X_t \notin (-a,b)\}, a, b > 0,$$

e cercare di calcolare la probabilità $p \equiv p(x,a,b) := \mathbb{P}_x(X_\tau = -a)$, e la trasformata di Laplace $\mathbb{E}_x[\exp\{-\alpha\tau\}]$, per $x \in (-a,b)$.

Sappiamo che, essendo $W_t = (X_t - \mu t)/\sigma$, il processo

$$Z_t := \exp\{\theta(X_t - \mu t)/\sigma - \theta^2 t/2\}$$

è una martingala con

$$\mathbb{E}_x[Z_t] = \exp\{\frac{\theta}{\sigma}x\}.$$

Anche $Z_{t\wedge\tau}$, è una martingala che inoltre converge a Z_τ. La convergenza è limitata:

$$Z_{t\wedge\tau} \le \exp\{|\theta|\max(a,b)/\sigma\}, \text{ se } \frac{\theta\mu}{\sigma} + \frac{\theta^2}{2} \ge 0$$

e quindi anche

$$\mathbb{E}_x[\exp\big\{\frac{\theta}{\sigma}X_\tau\big\}\exp\big\{-\frac{\theta}{2}(2\frac{\mu}{\sigma}+\theta)\tau\big\}] = \exp\big\{\frac{\theta}{\sigma}x\big\}.$$

Scegliendo θ in modo che $2\frac{\mu}{\sigma} + \theta = 0$, ovvero $\theta{=}-2\frac{\mu}{\sigma}$, la precedente uguaglianza diviene

$$\mathbb{E}_x[\exp\{\frac{\theta}{\sigma}X_\tau\}] = \mathbb{P}_x[X_\tau = -a]\exp\{-\frac{\theta}{\sigma}a\} + \mathbb{P}_x[X_\tau = b]\exp\{\frac{\theta}{\sigma}b\} = \exp\{\frac{\theta}{\sigma}x\},$$

da cui subito, tenendo conto del valore di θ,

$$p\exp\{2\frac{\mu}{\sigma^2}a\} + (1-p)\exp\{-2\frac{\mu}{\sigma^2}b\} = \exp\{-2\frac{\mu}{\sigma^2}x\},$$

e quindi

$$p = \frac{\exp\{-2\frac{\mu}{\sigma^2}x\} - \exp\{-2\frac{\mu}{\sigma^2}b\}}{\exp\{2\frac{\mu}{\sigma^2}a\{-\exp\{-2\frac{\mu}{\sigma^2}b\}}.$$

Per ogni $\alpha > 0$, si pongano ora θ_α^+ e θ_α^- i due valori per cui $\alpha := \frac{\theta}{2}(2\frac{\mu}{\sigma} + \theta)$, cioè le due soluzioni di $\theta^2 + 2\frac{\mu}{\sigma}\theta - 2\alpha = 0$, ovvero

$$\theta_\alpha^+, \theta_\alpha^- := -\frac{\mu}{\sigma} \pm \sqrt{\left(\frac{\mu}{\sigma}\right)^2 + 2\alpha}.$$

Come prima si ottiene che

$$\mathbb{E}_x[\exp\{\frac{\theta_\alpha^+}{\sigma}X_\tau\}\exp\{-\alpha\tau\}] = \exp\{\frac{\theta_\alpha^+}{\sigma}x\},\ \mathbb{E}_x[\exp\{\frac{\theta_\alpha^-}{\sigma}X_\tau\}\exp\{-\alpha\tau\}] = \exp\{\frac{\theta_\alpha^-}{\sigma}x\}.$$

Considerando che per $\theta = \theta_\alpha^+, \theta_\alpha^-$

$$\mathbb{E}_x[\exp\{\frac{\theta}{\sigma}X_\tau\}\exp\{-\alpha\tau\}] = \exp\{-\frac{\theta}{\sigma}a\}\mathbb{E}_x[I_{\{X_\tau=-a\}}\exp\{-\alpha\tau\}] + \exp\{\frac{\theta}{\sigma}b\}\mathbb{E}_x[I_{\{X_\tau=b\}}\exp\{-\alpha\tau\}]$$

si ottiene immediatamente il seguente sistema lineare nelle incognite

$$\begin{cases} y_{-a} & := & \mathbb{E}_x[I_{\{X_\tau=-a\}}\exp\{-\alpha\tau\}] \\ \\ y_b & := & \mathbb{E}_x[I_{\{X_\tau=b\}}\exp\{-\alpha\tau\}] \end{cases}$$

$$\begin{cases} \exp\{-\frac{\theta_\alpha^+}{\sigma}a\}y_{-a} + \exp\{\frac{\theta_\alpha^+}{\sigma}b\}y_b & = & \exp\{\frac{\theta_\alpha^+}{\sigma}x\} \\ \\ \exp\{-\frac{\theta_\alpha^-}{\sigma}a\}y_{-a} + \exp\{\frac{\theta_\alpha^-}{\sigma}b\}y_b & = & \exp\{\frac{\theta_\alpha^-}{\sigma}x\} \end{cases}$$

di cui si ricava la soluzione, ovvero

$$\begin{cases} y_{-a} & = & \frac{\exp\{\nu_\alpha^+x\}\exp\{\nu_\alpha^-b\} - \exp\{\nu_\alpha^+b\}\exp\{\nu_\alpha^-x\}}{\exp\{-\nu_\alpha^+a\}\exp\{\nu_\alpha^-b\} - \exp\{\nu_\alpha^+b\}\exp\{-\nu_\alpha^-a\}} \\ \\ y_b & = & \frac{\exp\{-\nu_\alpha^+a\}\exp\{\nu_\alpha^-x\} - \exp\{\nu_\alpha^+x\}\exp\{-\nu_\alpha^-a\}}{\exp\{-\nu_\alpha^+a\}\exp\{\nu_\alpha^-b\} - \exp\{\nu_\alpha^+b\}\exp\{-\nu_\alpha^-a\}}, \end{cases}$$

avendo posto

$$\nu_\alpha^\pm := \frac{\theta_\alpha^\pm}{\sigma} = -\frac{\mu}{\sigma^2} \pm \sqrt{\left(\frac{\mu}{\sigma^2}\right)^2 + 2\frac{\alpha}{\sigma^2}}.$$

Basta poi considerare la somma

$$\mathbb{E}_x[\exp\{-\alpha\tau\}] = y_{-a} + y_b =$$

$$= \frac{\exp\{\nu_\alpha^+x\}\exp\{\nu_\alpha^-b\} - \exp\{\nu_\alpha^+b\}\exp\{\nu_\alpha^-x\} + \exp\{-\nu_\alpha^+a\}\exp\{\nu_\alpha^-x\} - \exp\{\nu_\alpha^+x\}\exp\{-\nu_\alpha^-a\}}{\exp\{-\nu_\alpha^+a\}\exp\{\nu_\alpha^-b\} - \exp\{\nu_\alpha^+b\}\exp\{-\nu_\alpha^-a\}}.$$

Esercizio 9.1. *Nel caso $x = 0$, $\mu > 0$ mandare a ad infinito in modo da ottenere la trasformata di Laplace di $\tau := \tau(-\infty, b)$*

soluzione:
Si noti che, essendo $\alpha, \mu > 0$, si ha $\nu_\alpha^+ > 0$ e $\nu_\alpha^- < 0$. Di conseguenza $\exp\{-\nu_\alpha^-a\} \to \infty$, mentre $\exp\{-\nu_\alpha^+a\} \to 0$, e dalla precedente espressione per $\mathbb{E}_0[\exp\{-\alpha\tau(-a,b)\}]$ si ricava, per $a \to \infty$

$$\mathbb{E}_0[\exp\{-\alpha\tau(-a,b)\}] \to \mathbb{E}_0[\exp\{-\alpha\tau(-\infty,b)\}] = \frac{1}{\exp\{\nu_\alpha^+b\}} = \exp\left[\left(\frac{\mu}{\sigma^2} - \sqrt{\left(\frac{\mu}{\sigma^2}\right)^2 + 2\frac{\alpha}{\sigma^2}}\right)b\right].$$

Esercizio 9.2. *Si consideri il caso del moto browniano o Wiener standard, cioè con μ=0 e σ=1, e con $x = 0$ ed $a = b$.*

soluzione:

$$\mathbb{E}_0[\exp\{-\alpha\tau(-a,a)\}] = \frac{\exp\{\nu_\alpha^- a\} - \exp\{\nu_\alpha^+ a\} + \exp\{-\nu_\alpha^+ a\} - \exp\{-\nu_\alpha^- a\}}{\exp\{-\nu_\alpha^+ a\}\exp\{\nu_\alpha^- a\} - \exp\{\nu_\alpha^+ a\}\exp\{-\nu_\alpha^- a\}},$$

e quindi, poiché in questo caso $\nu_\alpha^- = -\nu_\alpha^+ = -\sqrt{2\alpha}$,

$$\begin{aligned}
\mathbb{E}_0[\exp\{-\alpha\tau(-a,a)\}] &= \frac{\exp\{-\sqrt{2\alpha}a\} - \exp\{\sqrt{2\alpha}a\} + \exp\{-\sqrt{2\alpha}a\} - \exp\{\sqrt{2\alpha}a\}}{\exp\{-\sqrt{2\alpha}a\}\exp\{-\sqrt{2\alpha}a\} - \exp\{\sqrt{2\alpha}a\}\exp\{\sqrt{2\alpha}a\}} \\
&= 2\frac{\exp\{\sqrt{2\alpha}a\} - \exp\{-\sqrt{2\alpha}a\}}{\exp\{2\sqrt{2\alpha}a\} - \exp\{-2\sqrt{2\alpha}a\}} \\
&= 2\frac{\sinh(\sqrt{2\alpha}a)}{\sinh(2\sqrt{2\alpha}a)} = 2\frac{\sinh(\sqrt{2\alpha}a)}{2\sinh(\sqrt{2\alpha}a)\cosh(\sqrt{2\alpha}a)} \\
&= \frac{1}{\cosh(\sqrt{2\alpha}a)}
\end{aligned}$$

9.5 Integrale stocastico: cenni

Sia $W = (W_t)_t$ un moto browniano continuo standard. L'obiettivo è di dare un significato ad espressioni del tipo

$$\int_0^T X_s(\omega) dW_s(\omega) \tag{9.2}$$

dove l'integrando $(X_s)_{0 \leq s \leq T}$ 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[3].

Come abbiamo visto precedentemente nella sezione 8.4 le traiettorie del moto browniano non sono a variazione finita, quindi l'operazione (9.2) non si può definire traiettoria per traiettoria.
La (9.2) 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 $\mathcal{F}_t = \mathcal{F}_t^W$.

9.5.1 Processi elementari

Consideriamo le due seguenti classi di processi

Definizione 9.3 (Processi elementari predicibili). *I processi del tipo*

$$X_t(\omega) = \sum_{i=0}^{n-1} C_i(\omega) I_{(t_i, t_{i+1}]}(t), \tag{9.3}$$

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 predicibili**.

Nel seguito la ***famiglia dei processi elementari predicibili*** viene denotata con $\mathcal{E}([\alpha, \beta])$.

Definizione 9.4. *I processi del tipo*

$$X_t(\omega) = \sum_{i=0}^{n-1} C_i(\omega) I_{[t_i, t_{i+1})}(t), \tag{9.4}$$

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 opzionali**.

---

[3]† 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 $t_k^* \in [t_{k-1}, t_k]$.

Nel seguito la ***famiglia dei processi elementari opzionali*** viene denotata con $\mathcal{E}_{op}([\alpha,\beta])$.

Si noti che i processi considerati negli Esempi 4.11 e 4.12, per $\alpha = 0$ e $\beta = T$, nel definire l'integrale stocastico, sono esempi di processi elementari predicibili. Ricordiamo che l'integrale stocastico rispetto al processo di Wiener per tali processi viene definito come

$$\mathcal{I}_t^W(X) == \sum_{i=1}^{n} C_{i-1}(\omega) W_{t_i} - W_{t_{i-1}}) \left(= \sum_{i=0}^{n-1} C_i(\omega) W_{t_{i+1}} - W_{t_i}) \right).$$

Inoltre i processi elementari predicibili (e anche quelli opzionali) sono esempi di processi (congiuntamente) misurabili:

Definizione 9.5 (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.*

Negli Esempi 4.11 e 4.12 veniva anche richiesto che le variabili aleatorie coinvolte nella definizione del processo predicibile elementare fossero 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.[4]

Definizione 9.6. *Lo spazio* $\mathbb{L}^2\big([\alpha,\beta] \times \Omega\big) = \mathbb{L}^2\big([\alpha,\beta] \times \Omega, \mathcal{B}([\alpha,\beta]) \times \mathcal{F}, \lambda \times \mathbb{P}\big)$ *è lo spazio delle classi di equivalenza dei processi (congiuntamente) misurabili per i quali*

$$\mathbb{E}\left[\int_\alpha^\beta |X_s|^2 ds\right] < \infty. \tag{9.5}$$

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

$$\overline{\mathcal{E}^2}([\alpha,\beta]).$$

È importante identificare lo spazio $\overline{\mathcal{E}^2}([\alpha,\beta])$, perché è quello su cui (come vedremo) sappiamo definire l'integrale stocastico. È possibile mostrare che se X ha traiettorie continue, ed è un processo in $\mathbb{L}^2\big([\alpha,\beta]\times\Omega\big)$, ovvero se vale

[4]† Il fatto che $\mathbb{L}^2\big([\alpha,\beta] \times \Omega\big)$ sia uno spazio vettoriale è ovvio, in quanto se (9.5) 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.

(9.5), 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 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 esiste una successione di processi 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 ogni per $\alpha = 0$ e $\beta = T$, e $t \in [0, T]$, il processo $\mathcal{I}^W(X^n) = (\mathcal{I}_t^W(X^n))_{t\in[0,T]}$ definisce una martingala. 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 4.12 sappiamo che, per ogni $t \in [0, T]$,

$$\mathbb{E}\left[\left|\mathcal{I}_t^W(X^n - X^m)\right|^2\right] = \mathbb{E}\left[\int_0^t |X_t^m - X_t^n|^2 dt\right]. \tag{9.6}$$

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

$$\mathbb{E}\left[\max_{t\in[0,T]} \left|\mathcal{I}_t^W(X^n - X^m)\right|^2\right] \le 4\mathbb{E}\left[\left|\mathcal{I}_T^W(X^n - X^m)\right|^2\right]. \tag{9.7}$$

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 (9.7), 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 (9.6) per $t = T$, si ottiene che

$$\mathbb{E}\left[\int_0^T |X_t^m - X_t^n|^2 dt\right] \le \mathbb{E}\left[\max_{t\in[0,T]} \left|\mathcal{I}_t^W(X^n - X^m)\right|^2\right] \le 4\mathbb{E}\left[\int_0^T |X_t^m - X_t^n|^2 dt\right] \tag{9.8}$$

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 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 (9.8) 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 $\mathcal{I}^W(X^n) = (\mathcal{I}^W_t(X^n))_{t\in[0,T]}$ 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.

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

$$\mathcal{I}^W_t(X) = \int_0^t X_s dW_s.$$

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 dal risultato esaminato nell'Esempio 4.12:

Lemma 9.2. *Se X è un processo elementare predicibile, come nella (9.4), con C_i di quadrato integrabile, allora*

$$\mathbb{E}\left(\int_\alpha^\beta X_t dW_t \Big| \mathcal{F}_\alpha\right) = 0$$

$$\mathbb{E}\left[\left(\int_\alpha^\beta X_t dW_t\right)^2 \Big| \mathcal{F}_\alpha\right] = \mathbb{E}\left[\int_\alpha^\beta {X_t}^2 \Big| \mathcal{F}_\alpha\right].$$

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.

Procediamo con qualche esempio di calcolo di 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 Ito, che in questo caso coincide con l'integrale stocastico di Wiener, e gode allora di una importante proprietà.

Proposizione 9.3. *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 $\int_0^t f^2(s)ds$.

Osservazione 9.2. *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, 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 $\mathcal{L}(\theta) = \exp\{\theta^2\sigma^2(Z)/2\}$.

Diamo ora un esempio di calcolo di un integrale stocastico.

[5]Si può mostrare che il limite non dipende dalla particolare successione $\{X^n\}$ di processi elementari, scelta per approssimare il processo X.

Esempio 9.1 (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

$$X_t^n = \sum_{i=1}^{m_n} W_{t_{i-1}^n}(\omega) I_{(t_{i-1}^n, t_i^n]}(t),$$

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 8.4, si ottiene che

$$\int_\alpha^\beta 2W_s dW_s = -(\beta - \alpha) + W_\beta^2 - W_\alpha^2. \tag{9.9}$$

Introduciamo ora il sottospazio $M^2([\alpha, \beta])$ di $\mathbb{L}^2\big([\alpha, \beta] \times \Omega\big)$

Definizione 9.7. *Sia $M^2([\alpha, \beta])$ lo spazio delle classi d'equivalenza dei* **processi $\mathcal{F}_t^W$-progressivamente misurabili**[6] *tali che*

$$\mathbb{E}\left[\int_\alpha^\beta |X_s|^2 ds\right] < \infty.$$

Osservazione 9.3. ‡ *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. $\mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbf{P})$, 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 $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ 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*

$$Z = c + \int_0^T X_s dW_s,$$

[6]‡ 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 $\mathcal{F}_t \times \mathcal{B}([0, t])$.

con $X \in M^2([0,T])$ *e* $c \in \mathbb{R}$. *Naturalmente* $\mathbb{E}(Z) = c$.

Prendendo $Z = M_T$, *questa proprietà è equivalente alla proprietà che ogni martingala*[7] $(M_t)_{t\in[0,T]}$ *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 proposizione afferma che l'integrale stocastico può essere definito anche per una classe più ampia di processi, ed in particolare, se l'integrando è continuo, allora l'integrale è il limite di particolari[8] somme di Riemann, in analogia con l'integrale di Lebesgue.

Proposizione 9.4. *Se* X *è un processo dello spazio* $\Lambda^2[\alpha,\beta]$, *ovvero è progressivamente misurabile per il quale*

$$\mathbb{P}\left(\int_\alpha^\beta X_t^2 dt < \infty\right) = 1$$

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* $\alpha = t_0^n < t_1^n < ... < t_{m_n}^n = \beta$ *con* $|\pi_n| \to 0$ *si ha*

$$\sum_{k=1}^{m_n} X(t_{k-1}^n)(W_{t_k^n} - W_{t_{k-1}^n}) \longrightarrow_{n\to+\infty} \int_\alpha^\beta X(t)dW_t \text{ in probabilità.}$$

Osservazione 9.4. ‡ *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 9.5. ‡ *Sia* A *una variabile aleatoria* $\mathcal{F}_\alpha$*-misurabile limitata; allora per* $X \in \overline{\mathcal{E}}^2([\alpha,\beta])$ *(o per* $X \in \Lambda^2([\alpha,\beta])$*)*

$$\int_\alpha^\beta AX(t)dW_t = A\int_\alpha^\beta X(t)dW_t.$$

9.6 Calcolo stocastico e formula di Itô

Sia X un processo tale che per ogni $0 \le t_1 < t_2 \le T$

$$X_{t_2} - X_{t_1} = \int_{t_1}^{t_2} F_t dt + \int_{t_1}^{t_2} G_t dW_t,$$

[7] Si tratta sempre del caso in cui la filtrazione 'e la filtrazione generata da W, completata e resa continua a destra.

[8] 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 Ito, sta nel fatto che va sempre scelto $t_{k-1}^{n*} = t_{k-1}^n$.

dove F_t e G_t soddisfano opportune condizioni.[9]

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 (9.9) 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.[10]

Osservazione 9.5. *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*[11].

Definiamo

$$< X >_t = \int_0^t G_s{}^2 ds.$$

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. 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 9.6 (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 $t_1 < t_2$,

$$\begin{aligned} X_1(t_2)X_2(t_2) - X_1(t_1)X_2(t_1) &= \int_{t_1}^{t_2} [X_1(t)F_2(t) + X_2(t)F_1(t)]dt \\ &+ \int_{t_1}^{t_2} [X_1(t)G_2(t) + X_2(t)G_1(t)]dW_t \\ &+ \int_{t_1}^{t_2} G_1(t)G_2(t)dt. \end{aligned}$$

[9] 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.

[10] Il motivo sta nel fatto che il processo di Wiener non ha traiettorie a variazione limitata o meglio possiamo interpretare il Lemma 8.4 come $(W_{t+\delta} - W_t)^2 = O(\delta)$.

[11] A meno di identificazione di processi. (DA CHIARIRE MEGLIO)

Siano f una funzione regolare di (t,x) e X un processo di Itô, con differenziale stocastico (9.6). 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 9.7 (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 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. Lo 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[12]

$$f(t,x) = f(t_0,x_0) + f_t(t_0,x_0)(t-t_0) + f_x(t_0,x_0)(x-x_0) + \frac{1}{2} f_{xx}(t_0,x_0)(x-x_0)^2 + o(t-t_0) + o\big((x-x_0)^2\big)$$

e tenendo conto del fatto che

$$\begin{aligned} (X_{t_k} - X_{t_{k-1}})^2 &\simeq \big(F_{t_{k-1}}(t_k - t_{k-1}) + G_{t_{k-1}}(W_{t_k} - W_{t_{k-1}})\big)^2 \\ &= F_{t_{k-1}}^2 (t_k - t_{k-1})^2 + 2F_{t_{k-1}} G_{t_{k-1}} (t_k - t_{k-1})(W_{t_k} - W_{t_{k-1}}) + G_{t_{k-1}}^2 (W_{t_k} - W_{t_{k-1}})^2 \\ &= O\big((t_k - t_{k-1})^2\big) + o\big((t_k - t_{k-1})\big) + G_{t_{k-1}}^2 (W_{t_k} - W_{t_{k-1}})^2 \end{aligned}$$

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)

Il precedente teorema ammette diverse generalizzazioni, come ad esempio la seguente.

Teorema 9.8 (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

$$\begin{aligned} df(t,X_t) \; = \; & \left(f_t(t,X_t) + \sum_{i=1}^m f_{x_i}(t,X_t) F_i(t) + \frac{1}{2} \sum_{i,j=1}^m f_{x_i x_j}(t,X_t) G_i(t) G_j(t) \right) dt \\ & + \sum_{i=1}^m f_{x_i}(t,X_t) G_i(t) dW_t \\ = \; & \left(f_t(t,X_t) + \sum_{i=1}^m f_{x_i}(t,X_t) dX_i(t) + \frac{1}{2} \sum_{i,j=1}^m f_{x_i x_j}(t,X_t) G_i(t) G_j(t) \right) dt. \end{aligned}$$

[12] 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 un infinitesimo di ordine superiore a $|t_k - t_{k-1}|$.

Osservazione 9.6. *Se indichiamo con $f^{'}$ il gradiente di f rispetto a x. La formula di Itô si può scrivere in maniera più compatta:*

$$df(t, X_t) = f_t(t, X_t)dt + f^{'}(t, X_t)dX(t) + \frac{1}{2}\sum_{i,j=1}^{m} f_{x_i x_j}(t, X_t)d < X_i, X_j >_t .$$

Il differenziale stocastico si comporta, dunque, in maniera diversa da quello ordinario per la presenza dell'ultimo termine a destra.

9.6.1 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

$$\begin{aligned} f_t(t, x) &= \left(\mu - \frac{1}{2}\sigma^2\right) f(t, x),\\ f_x(t, x) &= \sigma f(t, x),\\ f_{xx}(t, x) &= \sigma^2 f(t, x). \end{aligned}$$

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

$$dS_t = \mu S_t dt + \sigma S_t dW_t$$

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 Ito 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 [13]

$$\begin{aligned} dS_t^2 &= 2S_t dS_t + \frac{1}{2} 2\sigma^2 S_t^2 dt \\ &= 2\mu S_t^2 dt + 2\sigma S_t^2 dW_t + \sigma^2 S_t^2 dt \\ &= (2\mu + \sigma^2) S_t^2 dt + 2\sigma S_t^2 dW_t. \end{aligned}$$

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.[14] 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.

9.7 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 9.8. *Diremo che il processo $(\xi_t)_{t\in[u,T]}$ è soluzione dell'equazione differenziale stocastica*

$$\begin{cases} d\xi_t = b(\xi_t,t)dt + \sigma(\xi_t,t)dW_t \\ \xi_u = x \qquad x \in \mathbb{R}^m \end{cases} \tag{9.10}$$

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 9.7. *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***).

---

[13] Per ottenere il differenziale di S_t^2 si può considerare che

$$\begin{aligned} df(t,S_t) &= f_t(t,S_t)dt + f_x(t,S_t)dS_t + \frac{1}{2} f_{xx}(t,S_t)(G^S)_t^2 dt \\ &= 2S_t(\mu S_t dt + \sigma S_t dW_t) + \sigma^2 S_t^2 dt \end{aligned}$$

oppure si può considerare che $S_t^2 = \tilde{f}(t,W_t)$ con

$$\tilde{f}(t,x) = s_0^2 \exp\{2\sigma x + (2\mu - \sigma^2)\, t\}$$

e applicare la formula di Ito al processo $X_t = W_t$. Si consiglia il lettore di controllare che si ottiene lo stesso risultato con questo metodo.

[14] 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.

Definizione 9.9. *Si dice che il sistema (9.10) 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 (9.10).*

Con questa definizione possiamo dire che il moto browniano geometrico è soluzione dell'equazione differenziale

$$d\xi_t = \mu \xi_t dt + \sigma \xi_t dW_t, \quad \xi_0 = s_0.$$

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[15]

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 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

$$G_t = \sigma \exp\{\lambda t\}.$$

Da quest'ultima espressione si ricava immediatamente che

$$\eta_t = x + \int_0^t \sigma \exp\{\lambda s\} dW_s,$$

e quindi

$$X_t = \exp\{-\lambda t\}\eta_t = \exp\{-\lambda t\} \left(x + \int_0^t \sigma \exp\{\lambda s\} dW_s \right),$$

[15] 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

$$y_t = \exp\{-\lambda t\} \left(C_0 + \int_0^t \exp\{\lambda s\} v_s \, ds \right)$$

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.[16]

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

$$\begin{aligned} &(X_t - \mathbb{E}[X_t])\left(X_{t'} - \mathbb{E}[X_{t'}]\right) \\ &= \exp\{-\lambda t\}\left(\int_0^t \sigma \exp\{\lambda s\} dW_s\right)\exp\{-\lambda t'\}\left(\int_0^{t'} \sigma \exp\{\lambda s\} dW_s\right) \\ &= \exp\{-\lambda(t+t')\}M_t M_{t'}, \end{aligned}$$

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 $t \leq t'$,

$$Cov(X_t, X_{t'}) = \exp\{-\lambda(t+t')\}\mathbb{E}\left[\int_0^t \sigma^2 \exp\{2\lambda s\} ds\right].$$

[16]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

$$dX_t = \left(x + \int_0^t \sigma \exp\{\lambda s\} dW_s\right)(-\lambda \exp\{-\lambda t\}dt) + \exp\{-\lambda t\}\sigma \exp\{\lambda t\} dW_t = -\lambda X_t + \sigma dW_t.$$

Bibliografia

[1] BALDI, P. *Equazione differenziali stocastiche e applicazioni.* Quaderni dell'Unione Matematica Italiana, 28. Bologna: Pitagora Editrice. VIII, 309 p. (seconda edizione), 2000.

[2] 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.

[3] 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.

[4] BJÖRK, T. *Arbitrage theory in continuous time.* Oxford University Press, Oxford, 1998.

[5] BLACK, F., AND SCHOLES, M. The pricing of options and corporates liabilities. *Journal of Political Economy 81* (1973), 637–659.

[6] BREIMAN, L. *Probability*, vol. 7 of *Classics in Applied Mathematics.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992. Corrected reprint of the 1968 original.

[7] DALL'AGLIO, G. *Calcolo delle Probabilità.* Zanichelli, Bologna, 2000. II edizione.

[8] HULL, C. J. *Opzioni, futures e altri derivati.* Il Sole 24 Ore, Milano, 2003. III edizione.

[9] 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.

[10] LUENBERGER, D. C. *Investment Science.* Oxford University Press, Oxford, 1998.

[11] MERTON, R. C. Theory of rational option pricing. *Bell J. Econom. and Management Sci. 4* (1973), 141–183.

[12] MORICONI, F. *Matematica finanziaria.* Edizioni il Mulino, Bologna, 1995.

[13] ROGERS, L. C. G. Equivalent martingale measures and no-arbitrage. *Stochastics Stochastics Rep. 51*, 1-2 (1994), 41–49.

[14] ROSS, S. M. *Stochastic processes*, second ed. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, 1996.

[15] ROSS, S. M. *An introduction to mathematical finance.* Cambridge University Press, Cambridge, 1999. Options and other topics.

[16] SCHACHERMAYER, W. A Hilbert space proof of the fundamental theorem of asset pricing in finite discrete time. *Insurance Math. Econom. 11*, 4 (1992), 249–257.

[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] STEELE, J. M. *Stochastic calculus and financial applications*, vol. 45 of *Applications of Mathematics (New York).* Springer-Verlag, New York, 2001.

[19] WILLIAMS, D. *Probability with martingales.* Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.