

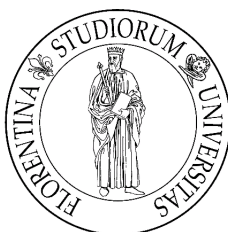
UNIVERSITÀ DEGLI STUDI DI FIRENZE

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

Anno Accademico 2008/2009

Corso di Metodi Matematici per la Finanza

INTRODUZIONE AI PROCESSI STOCASTICI



Indice

1	Prerequisiti	5
1.1	Spazi misurabili e funzioni misurabili	5
1.2	Integrale rispetto ad una misura astratta	7
1.3	Alcune disuguaglianze notevoli	9
1.4	Spazi L^p	9
1.5	Teoremi di misurabilità	10
2	Cenni di teoria della probabilità	12
2.1	Spazi di probabilità e variabili aleatorie	12
2.2	Varianza e legge immagine	13
2.3	Indipendenza di variabili aleatorie	14
2.4	Prodotto di convoluzione	17
2.5	Leggi e densità marginali	18
2.6	Funzioni caratteristiche	19
2.7	Leggi normali	21
2.8	Il Lemma di Borel-Cantelli	23
3	Processi stocastici	25
3.1	Generalità	25
3.2	Tempi d'arresto	30
3.3	Costruzione di processi stocastici	33

Premessa

Lo scopo principale di questi appunti è quello di fornire una concisa introduzione allo studio degli aspetti basilari della teoria della probabilità e dei processi stocastici.

L'esposizione degli argomenti sarà articolata come segue.

Il **capitolo 1** costituirà una parte preliminare in cui riprenderemo brevemente gli strumenti di analisi necessari alla comprensione dei capitoli successivi.

- Definiremo i concetti di σ -algebra e algebra boreliana.
- Ricorderemo cosa siano una misura su uno spazio misurabile, l'integrale rispetto ad una misura e due importanti teoremi di misurabilità.
- Introduciamo infine gli spazi L^p assieme ad alcune note disuguaglianze su tali spazi (quelle di Jensen, Hölder e Minkowski).

Nel **capitolo 2** si riprenderanno i concetti fondamentali del calcolo delle probabilità, ridefinendoli con gli strumenti più generali della teoria della misura astratta.

- Prima di tutto definiremo gli spazi di probabilità e le variabili aleatorie utilizzando le σ -algebre. Saranno ricordati l'espressione ed il significato di valore atteso, varianza e legge immagine. Si introdurrà il concetto di indipendenza di eventi, di variabili aleatorie e di leggi immagine, illustrandone alcune caratteristiche e contestualmente definendo la covarianza.
- Successivamente verrà definita la densità di probabilità mostrando come ricavare la legge immagine di somme di variabili aleatorie mediante l'operazione di convoluzione. Parleremo di leggi e densità marginali di un vettore aleatorio, riportando formule esplicite per il calcolo. Sarà introdotta la funzione caratteristica (o trasformata di Fourier) con le principali proprietà che la contraddistinguono.
- Infine analizzeremo la legge normale di probabilità, mettendone in luce le peculiarità e le più importanti e utili proprietà.

Nel **capitolo 3** analizzeremo i lineamenti della teoria dei processi stocastici, sfruttando gli strumenti introdotti nella parte precedente del lavoro.

- Per cominciare, daremo la definizione di *processo stocastico*, introducendo allo scopo particolari famiglie di σ -algebre dette *filtrazioni* e descrivendo cosa siano lo *spazio degli stati* e quello *delle traiettorie* di un processo. Faremo poi un semplice e familiare esempio di un fenomeno di natura aleatoria (il lancio di una moneta) in cui tali strutture risultino immediatamente riconoscibili.
- Mostreremo come si possano paragonare due processi stocastici diversi, in particolare cosa si intenda per *modificazione* ed *equivalenza* di processi e quand'è che due processi sono *indistinguibili*.
- Diremo cosa si intenda per *continuità* di un processo stocastico, enunciando a riguardo un importante teorema (che non dimostreremo). Da qui capiremo come si caratterizza una particolare categoria di processi, i cosiddetti *processi standard*.
- Parleremo del fondamentale concetto di *tempo d'arresto*, per familiarizzare col quale discuteremo due facili esempi, ed esporremo a riguardo una breve serie di risultati.
- Per concludere, affronteremo il problema della costruzione di un processo stocastico, a partire per esempio dalle cosiddette *distribuzioni di dimensione finita*, esponendo poi un fondamentale teorema dovuto a Kolmogorov. Nella parte finale elencheremo brevemente i principali aspetti che in generale intervengono nello studio dei processi stessi.

In tutto quanto segue eviteremo di soffermarci sugli aspetti tecnici della teoria e tenderemo invece di dare un'idea intuitiva dei concetti più importanti, anche se in alcuni punti sarà inevitabile ricorrere a qualche nozione di analisi non elementare.

Capitolo 1

Prerequisiti

In questo capitolo ci proponiamo di introdurre tutti gli strumenti necessari alla comprensione della teoria dei processi stocastici, focalizzandoci in particolare sulle basi della teoria della misura astratta.

In quanto segue si presuppone una conoscenza elementare della teoria della misura di Lebesgue e del concetto di integrale secondo Lebesgue, di cui si farà uso sistematico. Un'esposizione dettagliata di tali argomenti si può comunque trovare in vari testi e corsi di analisi [2].

1.1 Spazi misurabili e funzioni misurabili

Definizione 1.1.1. *Sia X un insieme qualsiasi. Una collezione \mathcal{M} di sottoinsiemi di X è una σ -algebra se:*

1. $X \in \mathcal{M}$,
2. $\forall A \in \mathcal{M}$ si ha che $A^c = X \setminus A \in \mathcal{M}$,
3. $\forall A_n \in \mathcal{M}, n \in \mathbb{N}$, si ha che $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{M}$.

Alcuni semplici esempi di σ -algebre sono l'insieme delle parti di un qualsiasi insieme X , o la famiglia formata dal solo X e dall'insieme vuoto \emptyset . Si noti che ogni σ -algebra contiene l'insieme vuoto e che, ponendo $A_n = \emptyset \forall n = m+1, m+2, \dots$, si ha che *le unioni finite di insiemi di una σ -algebra \mathcal{M} sono ancora contenute in \mathcal{M}* . Inoltre, anche le intersezioni numerabili di insiemi di \mathcal{M} sono ancora contenute in \mathcal{M} in quanto:

$$\bigcap_{n \in \mathbb{N}} A_n = \left(\bigcup_{n \in \mathbb{N}} A_n^c \right)^c.$$

La modalità (teorica e non costruttiva) con la quale da una qualunque famiglia di sottoinsiemi di X si passa ad una famiglia articolata come una σ -algebra è illustrata nella proposizione seguente.

Proposizione 1.1.2. *Sia \mathcal{F} una qualsiasi collezione di sottoinsiemi di un insieme X . Allora esiste una σ -algebra \mathcal{M}^* in X che è la più piccola σ -algebra contenente \mathcal{F} . \mathcal{M}^* si dice la σ -algebra generata da \mathcal{F} .*

Dimostrazione. Sia Ω la famiglia di σ -algre in X contenenti \mathcal{F} ; Ω è non vuota perché contiene almeno l'insieme delle parti di X . Consideriamo allora

$$\mathcal{M}^* = \bigcap_{\mathcal{M} \in \Omega} \mathcal{M}.$$

Si verifica facilmente che \mathcal{M}^* è ancora una σ -algebra e contiene \mathcal{F} . □

Definizione 1.1.3. *Sia (X, τ) uno spazio topologico. Si dice algebra boreliana di X (rispetto a τ), e si indica con \mathcal{B} , la σ -algebra generata da τ cioè la più piccola σ -algebra che contiene tutti gli aperti di X e i suoi elementi sono detti insiemi boreliani.*

Un esempio tipico di σ -algebra di questo tipo (fondamentale nelle applicazioni probabilistiche) è quella dei boreliani generati dagli aperti di \mathbb{R} . Esempi di insiemi che ne fanno parte sono le unioni numerabili di chiusi e le intersezioni numerabili di aperti, ma ve ne rientrano anche di più complicati (né aperti né chiusi, discreti, e altre forme ancora più strane). Nonostante la sua complessità la famiglia degli insiemi di Borel risulta enormemente meno ricca dell'insieme delle parti di \mathbb{R} , famiglia della quale è persino impossibile comprendere la cardinalità.

Definizione 1.1.4. *Sia \mathcal{M} una σ -algebra su X . La coppia (X, \mathcal{M}) si dice spazio misurabile e gli elementi di \mathcal{M} sono detti insiemi misurabili.*

Siano (X, \mathcal{M}_X) e (Y, \mathcal{M}_Y) due spazi misurabili. Una funzione $f : X \rightarrow Y$ si dice *misurabile* se

$$f^{-1}(A) \in \mathcal{M}_X \quad \forall A \in \mathcal{M}_Y,$$

cioè se $f^{-1}(A)$ è misurabile in X per ogni insieme misurabile A in Y .

Siano ora (X, τ_X) e (Y, τ_Y) spazi topologici. Si dice che $f : X \rightarrow Y$ è *boreliana* (o misurabile secondo Borel) se

$$f^{-1}(V) \in \mathcal{B}_X \quad \forall V \in \tau_Y.$$

Si noti che se $f : \mathbb{R}^n \rightarrow \mathbb{R}$ è boreliana allora è chiaramente misurabile secondo Lebesgue¹.

Per la trattazione che seguirà è inoltre opportuno ricordare che, dato uno spazio misurabile X , una funzione $s : X \rightarrow \mathbb{R}$ si dice *semplice* se è misurabile e assume solo un numero finito di valori. Detti c_1, \dots, c_n tali valori (distinti), si può scrivere:

$$s = \sum_{i=1}^n c_i \mathbb{1}_{E_i} \quad (1.1.1)$$

dove $E_i = \{x : s(x) = c_i\}, i = 1, \dots, n$. Risulta dunque che $\bigcup_{i=1}^n E_i = X$ e $E_i \cap E_j = \emptyset$ se $i \neq j$.

1.2 Integrale rispetto ad una misura astratta

Sia (X, \mathcal{M}) uno spazio misurabile. Una *misura (positiva)* è una funzione $\mu : \mathcal{M} \rightarrow [0, \infty]$ numerabilmente additiva, cioè tale che

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

per ogni collezione numerabile di elementi E_i di \mathcal{M} a due a due disgiunti. Se μ è una misura su (X, \mathcal{M}) , la terna (X, \mathcal{M}, μ) si dice uno *spazio di misura*. Esempi classici di spazi di misura sono i seguenti.

- $X = \mathbb{R}^N$, $\mathcal{M} = \sigma$ -algebra degli insiemi misurabili secondo Lebesgue e $\mu =$ misura di Lebesgue.
- Siano $X = \mathbb{R}^N$, $\mathcal{M} = \sigma$ -algebra degli insiemi misurabili secondo Lebesgue e $f : \mathbb{R}^N \rightarrow [0, +\infty]$ sommabile in \mathbb{R}^N , allora

$$\mu(E) = \int_E f(x) dx, \quad E \in \mathcal{M}$$

definisce una misura.

- X insieme qualsiasi, $\mathcal{M} = \mathcal{P}(X)$, $x_0 \in X$ e

$$\delta_{x_0}(E) = \begin{cases} 1, & \text{se } x_0 \in E \\ 0, & \text{se } x_0 \notin E \end{cases}.$$

La misura definita in questo modo si dice *delta di Dirac*.

¹Infatti ogni boreliano di \mathbb{R}^n è misurabile secondo Lebesgue (il contrario non è vero).

Ricordiamo ora, senza dimostrarle, alcune importanti proprietà di uno spazio di misura.

Teorema 1.2.1. *Sia (X, \mathcal{M}, μ) uno spazio di misura. Allora:*

1. $\mu(\emptyset) = 0$ se almeno un $E \in \mathcal{M}$ ha misura finita;
2. $\mu(E_1 \cup \dots \cup E_n) = \mu(E_1) + \dots + \mu(E_n)$ se E_1, \dots, E_n sono misurabili e a due a due disgiunti;
3. se $E \subseteq F$, allora $\mu(E) \leq \mu(F)$;
4. se $E_n \in \mathcal{M}$ per $n = 1, 2, \dots$ e $E_1 \subseteq E_2 \subseteq \dots$, allora

$$\lim_{n \rightarrow \infty} \mu(E_n) = \mu\left(\bigcup_{n \in \mathbb{N}} E_n\right);$$

5. se $E_n \in \mathcal{M}$ per $n = 1, 2, \dots$ e $E_1 \supseteq E_2 \supseteq \dots$, e $\mu(E_1) < +\infty$, allora

$$\lim_{n \rightarrow \infty} \mu(E_n) = \mu\left(\bigcap_{n \in \mathbb{N}} E_n\right).$$

Se (X, \mathcal{M}, μ) è uno spazio di misura ed f una funzione misurabile, possiamo a questo punto definire l'integrale di f rispetto alla misura μ come segue:

- se s è una funzione semplice misurabile su X della forma (1.1), si definisce l'integrale su X di s rispetto alla misura μ come:

$$\int_X s \, d\mu = \sum_{i=1}^n c_i \mu(E_i);$$

- se f è una funzione misurabile non negativa, si definisce l'integrale di f su X rispetto a μ come:

$$\int_X f \, d\mu = \sup \left\{ \int_X s \, d\mu : s \text{ è semplice, } 0 \leq s \leq f \right\};$$

- una funzione f si dice sommabile rispetto μ se esiste finito l'integrale $\int_X |f| \, d\mu$ e si pone:

$$\int_X f \, d\mu = \int_X f^+ \, d\mu - \int_X f^- \, d\mu$$

dove $f^+ = \max(f, 0)$ e $f^- = \max(-f, 0)$.

Osserviamo inoltre che si definisce

$$\int_E f \, d\mu = \int_X f \mathbb{1}_E \, d\mu$$

se $E \in \mathcal{M}$.

1.3 Alcune disuguaglianze notevoli

Teorema 1.3.1. (Disuguaglianza di Jensen) *Sia μ una misura positiva su una σ -algebra \mathcal{M} sull'insieme Ω . Sia $\varphi : (a, b) \rightarrow \mathbb{R}$ convessa e sia f una funzione a valori reali, sommabile su Ω rispetto a μ e tale che valga $a < f(x) < b \quad \forall x \in \Omega$. Allora, se $\mu(\Omega) = 1$, risulta che*

$$\varphi\left(\int_{\Omega} f(x) d\mu\right) \leq \int_{\Omega} \varphi(f(x)) d\mu. \quad (1.3.1)$$

Teorema 1.3.2. (Disuguaglianza di Hölder) *Sia E un sottoinsieme misurabile di \mathbb{R}^N e siano $p, p' > 1$ tali che $\frac{1}{p} + \frac{1}{p'} = 1$. Se $|f|^p, |g|^{p'}$ sono sommabili su E , allora fg è sommabile su E e si ha:*

$$\int_E |f(x)g(x)| dx \leq \left(\int_E |f(x)|^p dx\right)^{\frac{1}{p}} \left(\int_E |g(x)|^{p'} dx\right)^{\frac{1}{p'}}.$$

Il segno di uguaglianza vale se e solo se esiste $\lambda > 0$ tale che $|g| = \lambda|f|^{p-1}$ quasi ovunque in E .

Teorema 1.3.3. (Disuguaglianza di Minkowski) *Sia $1 \leq p \leq +\infty$. Se $|f|^p$ e $|g|^p$ sono sommabili su E allora anche $|f+g|^p$ è sommabile su E e risulta che*

$$\left(\int_E |f(x) + g(x)|^p dx\right)^{\frac{1}{p}} \leq \left(\int_E |f(x)|^p dx\right)^{\frac{1}{p}} + \left(\int_E |g(x)|^p dx\right)^{\frac{1}{p}}.$$

Inoltre, il segno di uguaglianza vale se e solo se f e g sono linearmente dipendenti.

1.4 Spazi L^p

Sia p un numero tale che $0 < p < +\infty$ e sia $E \subset \mathbb{R}^N$ misurabile. Definiamo:

$$L^p(E) = \left\{ f : E \rightarrow \mathbb{R} \mid f \text{ è misurabile e } \int_E |f(x)|^p dx \leq \infty \right\} / \sim$$

dove

$$f \sim g \Leftrightarrow \text{mis}\{x \in E \mid f(x) \neq g(x)\} = 0.$$

Se inoltre $f \in L^p(E)$, si definisce

$$\|f\|_p := \left(\int_E |f(x)|^p dx\right)^{1/p}$$

dove f è un qualsiasi rappresentante della classe di equivalenza. Si dimostra facilmente, osservando quanto segue, che $\|\cdot\|_p$ è una norma su $L^p(E)$.

- $\|f\|_p \geq 0$ (ovvio) e vale l'uguale se e solo se $f = 0$ (perché $f = 0$ va inteso nel senso di $f = 0$ quasi ovunque²).
- $\|f + g\|_p \leq \|f\|_p + \|g\|_p$ per la disuguaglianza di Minkowski.
- $\|\alpha f\|_p = |\alpha| \|f\|_p$ (ovvio).

Possiamo inoltre definire lo spazio $L^\infty(E)$ come

$$L^\infty(E) := \left\{ f : E \rightarrow \mathbb{R} \mid f \text{ è misurabile e } \text{ess sup } |f(x)| < +\infty \right\}$$

con la norma

$$\|f\|_\infty = \text{ess sup } |f| := \inf \left\{ M \geq 0 \mid \text{mis} \{x \in E : |f(x)| > M\} = 0 \right\}.$$

Si può dimostrare che per ogni $1 \leq p \leq +\infty$ lo spazio $L^p(E)$ (con la norma $\|f\|_p$) è uno *spazio di Banach*, cioè uno spazio normato completo. L^2 in particolare è anche uno *spazio di Hilbert*, col prodotto scalare definito da

$$\langle f, g \rangle = \int_E fg \, dx.$$

Osserviamo infine che, dati p, q t.c. $1 \leq p, q \leq \infty$, se $\text{mis}(E) < \infty$ allora $L^q(E) \subset L^p(E)$ (per provarlo basta prendere una generica funzione $f \in L^q(E)$ e applicare la disuguaglianza di Hölder alle funzioni $|f|^p \in L^{(q/p)}(E)$ e $\mathbb{1} \in L^{(q/p)'}(E)$ per dedurre che $f \in L^p(E)$).

1.5 Teoremi di misurabilità

Concludiamo questa introduzione enunciando due teoremi frequentemente utilizzati nella teoria. Dato un insieme E e una classe di suoi sottoinsiemi \mathcal{C} , denoteremo con $\sigma(\mathcal{C})$ la σ -algebra generata da \mathcal{C} .

Teorema 1.5.1. *Sia E un insieme e \mathcal{C} una classe di parti di E stabile per l'intersezione finita. Sia poi \mathcal{H} uno spazio vettoriale di funzioni limitate tale che:*

1. *se $f_n \in \mathcal{H}$ è una successione crescente di elementi tali che $|f_n| < h \in \mathcal{H}$ per ogni $n \in \mathbb{N}$, allora:*

$$\sup_{n \in \mathbb{N}} f_n \in \mathcal{H};$$

²Ossia tranne al più su un insieme di misura nulla.

2. \mathcal{H} contiene la funzione identità;
3. \mathcal{H} contiene le funzioni indicatrici degli elementi di \mathcal{C} :

$$\mathbf{1}_c \in \mathcal{H} \quad \forall c \in \mathcal{C}.$$

Allora \mathcal{H} contiene tutte le funzioni $\sigma(\mathcal{C})$ -misurabili e limitate.

Teorema 1.5.2. Sia \mathcal{H} uno spazio vettoriale di funzioni reali limitate e definite su un insieme E tale che:

1. la funzione identità sta in \mathcal{H} ;
2. se $f_n \in \mathcal{H}$ è una successione convergente uniformemente, allora

$$\lim_{n \rightarrow +\infty} f_n = f \in \mathcal{H};$$

3. se $f_n \in \mathcal{H}$ è successione crescente di funzioni positive equilimitate³, allora

$$\lim_{n \rightarrow +\infty} f_n \in \mathcal{H}.$$

Allora \mathcal{H} contiene tutte le funzioni $\sigma(\mathcal{L})$ -misurabili limitate, dove \mathcal{L} è un sottospazio di \mathcal{H} stabile per la moltiplicazione.

³Una successione di funzioni f_n definite su un insieme E si dice *equilimitata* se $\exists c \geq 0$ tale che $|f_n(x)| \leq c \forall n \in \mathbb{N}, \forall x \in E$.

Capitolo 2

Cenni di teoria della probabilità

In questo capitolo si riprendono i concetti base della teoria della probabilità. Senza farne una trattazione troppo approfondita, si sono ridefiniti i più importanti strumenti in uso con la teoria dell'integrazione rispetto alle misure astratte. A chiusura del capitolo si presentano la distribuzione normale di v.a. unidimensionali e di vettori aleatori m -dimensionali.

2.1 Spazi di probabilità e variabili aleatorie

Definizione 2.1.1. Si dice spazio di probabilità una terna (Ω, \mathcal{F}, P) , dove:

- Ω è un insieme;
- \mathcal{F} è una σ -algebra di parti di Ω ;
- P è una misura positiva su \mathcal{F} tale che $P(\Omega) = 1$.

Gli elementi di \mathcal{F} sono detti *eventi* e P una (*misura di*) *probabilità*.

Definizione 2.1.2. Sia (Ω, \mathcal{F}, P) uno spazio di probabilità e (E, \mathcal{A}) uno spazio misurabile, si dice *variabile aleatoria* (abbreviato v.a.) una funzione $X : \Omega \rightarrow E$ misurabile, cioè tale che $X^{-1}(A) \in \mathcal{F}$ per ogni $A \in \mathcal{A}$. Nel caso $E = \mathbb{R}$ e $\mathcal{A} = \mathcal{B}(\mathbb{R})$ (boreliani di \mathbb{R}) si dice *variabile aleatoria reale* (abbreviato v.a.r.).

Consideriamo una v.a.r. X sullo spazio di probabilità (Ω, \mathcal{F}, P) . Se X è integrabile chiameremo *valore atteso* di X il valore

$$E(X) = \int X dP.$$

Se X è m -dimensionale, cioè a valori in $(\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m))$, e le singole componenti X_i sono integrabili allora si pone

$$E(X) = (E(X_1), E(X_2), \dots, E(X_m)).$$

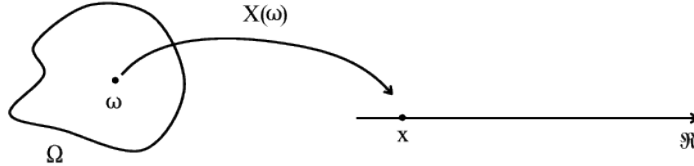


Figura 2.1: Rappresentazione intuitiva di una variabile aleatoria reale.

2.2 Varianza e legge immagine

Se la v.a.r. X sta in L^2 possiamo definire la *varianza* di X come la quantità

$$Var(X) = E[(X - E(X))^2] = E(X^2) - E(X)^2.$$

La varianza misura quanto la variabile aleatoria assume valori lontani dalla sua media $M = E(X)$. Più precisamente, se X è una v.a.r. in L^2 allora per ogni $\alpha > 0$

$$P\{\omega : |X(\omega) - M| \geq \alpha\} \leq \frac{Var(X)}{\alpha^2}.$$

questo risultato è noto come *disuguaglianza di Chebychev*.

Introduciamo ora il concetto di *legge immagine* considerando una v.a. X a valori nello spazio misurabile (E, \mathcal{A}) e la funzione Q definita su \mathcal{A} come segue:

$$Q(A) = P(X^{-1}(A)) \quad \text{con } A \in \mathcal{A}.$$

Si può facilmente verificare che così definita questa funzione è a sua volta una misura di probabilità. Si dice allora che Q è la *legge immagine di P tramite X* (o più semplicemente la legge di X , una volta fissato lo spazio (Ω, \mathcal{F}, P)). Da ora in poi utilizzeremo le notazioni $X \sim \mu$ e $X \sim Y$ per indicare rispettivamente che X ha legge μ e che X e Y hanno la stessa legge.

Proposizione 2.2.1. *Sia $f : (E, \mathcal{A}) \rightarrow (\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m))$ una funzione misurabile, $X : (\Omega, \mathcal{F}, P) \rightarrow (E, \mathcal{A})$ una v.a. e sia Q la legge di X . Allora f è Q -integrabile se e solo se $f \circ X$ è P -integrabile e si ha*

$$\int_E f(x) Q(dx) = \int_\Omega f \circ X(\omega) dP(\omega).$$

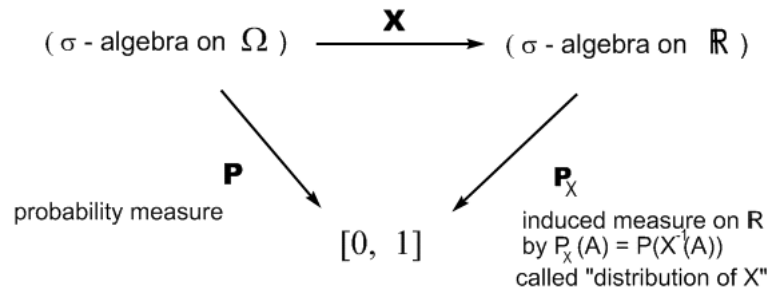


Figura 2.2: Misura di probabilità, v.a. reale e legge immagine (schema).

In particolare perciò, se X è una v.a.r. e μ la sua legge, valgono le seguenti relazioni:

$$E(X) = \int x \mu(dx),$$

$$E(|X|^\alpha) = \int |x|^\alpha \mu(dx).$$

La quantità $E(|X|^\alpha) = \int |x|^\alpha dx$, quando esiste finita, si chiama *momento di ordine α di μ* . Dunque si deduce immediatamente che una v.a.r. X sta in L^p se e solo se ha un momento assoluto di ordine p finito.

Per i momenti vale una versione generalizzata della disuguaglianza di Chebychev, nella forma

$$P\{|X| \geq \alpha\} \leq \frac{E(|X|^\beta)}{\alpha^\beta} \quad \forall \alpha, \beta > 0.$$

2.3 Indipendenza di variabili aleatorie

Informalmente, due eventi si dicono *indipendenti* se non si influenzano l'uno con l'altro, cioè se il verificarsi del primo non modifica le probabilità che si verifichi il secondo.

Un esempio è il lancio di due dadi: il fatto che nel primo dado sia uscito il 6 non influisce affatto sull'uscita del 5 nel secondo lancio.

Viceversa in un'estrazione senza reinserimento si hanno eventi dipendenti. Consideriamo la probabilità che alla prima estrazione della tombola esca 90 e alla seconda esca 89. Chiaramente il primo evento influenza il secondo, poiché dopo la prima estrazione è stato tolto un numero dal sacco della tombola.

Diamo ora la definizione matematica di v.a. indipendenti, quindi associate ad eventi indipendenti.

Definizione 2.3.1. *Siano $X_i : (\Omega, \mathcal{F}, P) \rightarrow (E_i, \mathcal{E}_i)$ delle variabili aleatorie per $i = 1, \dots, m$. Esse si dicono indipendenti se per ogni m -upla di possibili risultati $A'_1 \in \mathcal{E}_1, \dots, A'_m \in \mathcal{E}_m$ si ha che:*

$$P[X_1 \in A'_1, \dots, X_m \in A'_m] = P[X_1 \in A'_1] \cdots P[X_m \in A'_m]$$

In realtà questa definizione viene spesso utilizzata al contrario. Invece che verificare l'uguaglianza per dimostrare che due o più v.a. sono indipendenti, si sfrutta la relazione per semplificare il calcolo della probabilità di un'evento composto quando si suppone che gli eventi semplici che lo costituiscono siano indipendenti.

Per esempio, nell'esperimento con i dadi di prima, supponiamo che il lancio di un dado sia indipendente dal lancio dell'altro. Invece di calcolare la probabilità dell'evento composto “esce un 6 ed un 5 nel lancio successivo di due dadi”, possiamo suddividerlo in due sotto-eventi indipendenti e sfruttare la relazione data dalla definizione:

$$\begin{aligned} & P[\text{“esce un 6 ed un 5 nel lancio successivo di due dadi”}] \\ &= P[\text{“esce 6 nel primo dado”, “esce 5 nel secondo dado”}] \\ &= P[\text{“esce 6 nel primo dado”}] \cdot P[\text{“esce 5 nel secondo dado”}] \\ &= \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}. \end{aligned}$$

Si può vedere invece che un ragionamento analogo non vale se vogliamo risolvere il secondo problema, poiché in tal caso gli eventi non sono indipendenti:

$$\begin{aligned} & P[\text{“si estragga alla tombola il 90 e poi il 89”}] \\ &\neq P[\text{“si estrae il 90”, “si estrae il 89”}] \\ &= P[\text{“si estrae il 90”}] \cdot P[\text{“si estrae il 89”}] \\ &= \frac{1}{90} \cdot \frac{1}{90} = \frac{1}{90^2}. \end{aligned}$$

Analogamente a quanto fatto per le v.a. possiamo definire gli eventi A_1, \dots, A_m indipendenti se sono indipendenti le variabili aleatorie indicatrici $\mathbb{1}_{A_1}, \dots, \mathbb{1}_{A_m}$. Quindi se e solo se vale la relazione

$$P[A_{i_1}, \dots, A_{i_s}] = P[A_{i_1}] \cdots P[A_{i_s}]$$

per ogni scelta di indici $i_1, \dots, i_s \leq m$.

Infine, se abbiamo una famiglia di v.a. $(X_i)_{i \in I}$ (anche infinita), esse si dicono indipendenti se e solo se sono tutte indipendenti tra di loro. Cioè se sono indipendenti le X_{i_1}, \dots, X_{i_m} per ogni scelta di indici i_1, \dots, i_m in I .

Se lavoriamo con variabili aleatorie indipendenti si possono utilizzare diverse importanti proprietà che semplificano spesso i calcoli. Intanto vediamo un criterio necessario e sufficiente perché delle v.a. siano indipendenti.

Proposizione 2.3.2. *Siano X_i , per $i = 1, \dots, m$, delle v.a. definite come in precedenza. Poniamo poi $X = (X_1, \dots, X_m)$ definita su $E = E_1 \times \dots \times E_m$ ed anche $\mathcal{E} = \mathcal{E}_1 \times \dots \times \mathcal{E}_m$ con la legge prodotto $\mu = \mu_1 \otimes \dots \otimes \mu_m$.*

Allora le v.a. X_1, \dots, X_m sono indipendenti se e solo se la legge di X su (E, \mathcal{E}) è la legge prodotto μ .

Grazie a questa proposizione è possibile caratterizzare la legge di una v.a. m -dimensionale a partire dalle distribuzioni delle sue componenti.

Dimostrazione (traccia). Indichiamo con ν la legge di X e prendiamo l'insieme $I = \{A \subset \mathcal{E} : A = A_1 \times \dots \times A_m, \text{ con } A_i \in \mathcal{E}_i\}$, che è un sottoinsieme dell'insieme delle parti di \mathcal{E} (si dice talvolta una *famiglia di parti* di \mathcal{E}).

Notiamo che considerando intersezioni finite di elementi di I otteniamo ancora elementi di questo insieme (si dice allora che I è *stabile per l'intersezione finita*), e che I genera per definizione la σ -algebra \mathcal{E} .

La definizione di indipendenza ci dice allora che le distribuzioni ν e $\mu = \mu_1 \otimes \dots \otimes \mu_m$ coincidono sull'insieme I .

Basta a questo punto utilizzare il *Criterio di Caratheodory*¹ per generalizzare il ragionamento e affermare che ν e μ coincidono su tutto \mathcal{E} . \square

Vorremmo ora trovare una quantità, un parametro, che ci permetta di determinare la dipendenza delle variabili aleatorie. Per fare questo osserviamo che vale un'interessante proprietà per il valore atteso del prodotto di due v.a. reali indipendenti.

Proposizione 2.3.3. *Se X e Y sono due v.a. reali indipendenti ed integrabili, allora anche XY è integrabile e vale che:*

$$E(XY) = E(X)E(Y).$$

¹**Teorema** (Criterio di Caratheodory): *Sia (E, \mathcal{E}) uno spazio misurabile e μ_1, μ_2 due misure finite su questo spazio. Sia I una famiglia di parti di \mathcal{E} stabile per intersezione finita e che genera \mathcal{E} . Allora se μ_1 e μ_2 coincidono su I , esse coincidono su tutto \mathcal{E} .*

La quantità che ci interessa definire dovrà valere zero se due v.a. sono indipendenti. Possiamo allora sfruttare questa relazione tramite la seguente definizione.

Definizione 2.3.4. *Si dice covarianza tra due v.a. il numero reale*

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y).$$

Si dice che due v.a. X e Y non sono correlate se $\text{Cov}(X, Y) = 0$.

È possibile poi generalizzare questa quantità anche nel caso di una v.a. m -dimensionale costruendo un'opportuna matrice di covarianza che ci darà informazioni sulla correlazione tra le varie componenti di $X = (X_1, \dots, X_m)$. Questa matrice $\Gamma \in \mathcal{M}(m \times m)$ sarà definita da:

$$\Gamma_{ij} = \text{Cov}(X_i, X_j).$$

Possiamo allora notare che, se le componenti sono tra loro a due a due non correlate, Γ sarà una matrice diagonale.

Se in particolare X è una v.a.r. di legge μ si ha

$$\Gamma_{i,j} = \int x_i x_j \mu(dx) - \int x_i \mu(dx) \int x_j \mu(dx).$$

In ogni caso osserviamo che, mentre due v.a. reali indipendenti non sono correlate, il viceversa non è vero in generale.

2.4 Prodotto di convoluzione

Definiamo ora la densità di una variabile aleatoria.

Definizione 2.4.1. *Data una legge di probabilità μ su \mathbb{R}^m , si dice che questa ammette una densità (rispetto alla misura di Lebesgue) se esiste una funzione $f(x)$ positiva e integrabile tale che:*

$$\mu(A) = \int_A f(x) dx$$

per ogni $A \in \mathcal{B}(\mathbb{R}^m)$, insieme dei boreliani di \mathbb{R}^m (che è un sottoinsieme dei misurabili secondo Lebesgue).

La densità è un concetto meno generale della legge, ma ci permette di caratterizzare la somma di variabili aleatorie.

Consideriamo due v.a. m -dimensionali indipendenti X e Y con le loro rispettive leggi μ e ν . Allora la legge della v.a. $X + Y$ si dice *prodotto di convoluzione* di μ e ν e si indica con $\mu \star \nu$.

Possiamo determinare la distribuzione dei prodotti di convoluzione utilizzando la seguente proposizione.

Proposizione 2.4.2. *Siano μ e ν misure di probabilità su \mathbb{R}^m con densità $f(x)$ e $g(x)$ rispettivamente. Allora il prodotto di convoluzione $\mu \star \nu$ ha densità:*

$$h(x) = f \star g(x) = \int_{\mathbb{R}^m} f(y)g(x - y)dy.$$

Dimostrazione. Prendiamo $A \in \mathcal{B}(\mathbb{R}^m)$, allora, per la regola di integrazione rispetto ad una legge immagine:

$$\begin{aligned} (\mu \star \nu)(A) &= \int_A (\mu \star \nu)(dz) \\ &= \int \mathbf{1}_A(x + y) \mu(dx) \nu(dy) \\ &= \int \mathbf{1}_A(x + y) f(x)g(y) dx dy \\ &= \int_A \left[\int_{\mathbb{R}^m} f(y)g(x - y) dy \right] dx. \end{aligned}$$

□

Quindi per ricavare la densità di una somma di v.a. basta calcolare il normale prodotto di convoluzione tra le densità degli addendi.

2.5 Leggi e densità marginali

Finora abbiamo analizzato come si può combinare le v.a. tra di loro e abbinarle inserendole in un vettore aleatorio (che diventa una nuova v.a. con una propria legge di probabilità). Nel caso di v.a. indipendenti abbiamo visto come ricavare questa legge dalle leggi delle componenti del vettore. Vediamo ora come fare il passaggio inverso e recuperare le informazioni che caratterizzano una v.a. che è componente di una variabile m -dimensionale.

Definizione 2.5.1. *Sia μ una legge di probabilità su \mathbb{R}^m . Consideriamo poi la funzione di proiezione sulla i -esima coordinata*

$$\pi_i : \mathbb{R}^m \rightarrow \mathbb{R}.$$

Si dice *i-esima legge marginale di μ* la composizione

$$\pi_i(\mu) = \mu_i.$$

Quindi se abbiamo una v.a. $X = (X_1, \dots, X_m)$ di legge μ , la legge marginale i -esima μ_i è proprio la legge di X_i . Dalla precedente definizione possiamo trovare una formula esplicita che permette di ricavare immediatamente le leggi marginali:

$$\mu_i(A) = \int \mathbf{1}_A(x_i) \mu(dx_1, \dots, dx_m).$$

In particolare è, se μ ammette una densità f , anche le μ_i ammettono delle densità f_i , date dalla formula²:

$$f_i(x) = \int f(y_1, \dots, y_{i-1}, x, y_{i+1}, \dots, y_m) dy_1 \dots dy_{i-1} dy_{i+1} \dots dy_m.$$

Infine notiamo che se esiste la densità $f(x)$ allora necessariamente esistono anche le densità $f_i(x)$ per le componenti X_i del vettore aleatorio.

Abbiamo quindi trovato delle formule per ricavare le densità e le leggi marginali da quella generale. Fare il passaggio inverso non è generalmente possibile per via delle correlazioni che ci possono essere tra le varie componenti del vettore aleatorio.

Naturalmente se le v.a. X_i sono tra loro indipendenti questo si può fare, per quanto abbiamo visto.

2.6 Funzioni caratteristiche

Uno strumento utile allo studio delle leggi e delle distribuzioni di v.a. è la funzione caratteristica, meglio conosciuta come *Trasformata di Fourier*.

Definizione 2.6.1. Sia X una v.a. m -dimensionale con legge μ . Allora si chiama funzione caratteristica di μ l'applicazione

$$\hat{\mu}(y) = E[e^{iy \cdot X}]$$

con $y \in \mathbb{R}^m$ (con l'espressione $y \cdot X$ abbiamo indicato il prodotto scalare su \mathbb{R}^m dei vettori X e y).

²L'esistenza dell'integrale e la misurabilità delle f_i sono garantite in questo caso dal Teorema di Fubini-Tonelli.

Questa funzione esiste sempre e dalla definizione precedente si ricava facilmente la sua espressione esplicita:

$$\hat{\mu}(y) = \int e^{iy \cdot x} \mu(dx).$$

Inoltre, se $\hat{\mu}$ è integrabile allora μ ammette una densità f data dalla formula

$$f(x) = \frac{1}{2\pi} \int e^{-iy \cdot x} \hat{\mu}(y) dy.$$

Di seguito analizzeremo le proprietà che contraddistinguono le funzioni caratteristiche.

Regolarità - Si può dimostrare che la funzione caratteristica è uniformemente continua su tutto \mathbb{R}^m . In più si ricava facilmente dalla scrittura esplicita che

$$\hat{\mu}(0) = 1 \quad \text{e} \quad |\hat{\mu}(y)| \leq 1 \quad \forall y \in \mathbb{R}^m.$$

Riportiamo ora (senza dimostrarle) due interessanti proprietà riguardanti la combinazione di variabili aleatorie. Siano quindi X e Y due v.a. con distribuzioni μ e ν rispettivamente. Allora :

1. la legge di $X+Y$ è, come abbiamo detto, $\mu \star \nu$ e la funzione caratteristica diventa semplicemente

$$\widehat{\mu \star \nu}(y) = \hat{\mu}(y) \hat{\nu}(y);$$

2. se abbiamo che $\hat{\mu}(y) = \hat{\nu}(y)$ per ogni $y \in \mathbb{R}^m$, allora possiamo dire che le due leggi coincidono, ovvero $\mu = \nu$.

Derivazione - Adesso passiamo ad analizzare il comportamento delle derivate delle funzioni caratteristiche, quando è possibile calcolarle.

Proposizione 2.6.2. *Sia μ una legge di probabilità con momento primo finito. Allora $\hat{\mu}$ è derivabile e vale*

$$\frac{\partial \hat{\mu}}{\partial y_j}(y) = i \int x_j e^{iy \cdot x} \mu dx.$$

Inoltre, se μ ha anche momento di ordine due finito, allora $\hat{\mu}$ è derivabile due volte e vale

$$\frac{\partial^2 \hat{\mu}}{\partial y_j \partial y_k}(y) = - \int x_j x_k e^{iy \cdot x} \mu dx.$$

Possiamo facilmente intuire come si iteri il ragionamento per la derivazione ad ordini più elevati.

Osservazione 2.6.3. *Dalla precedente proposizione possiamo vedere che la matrice Hessiana di $\hat{\mu}$ calcolata nell'origine ha espressione*

$$\frac{\partial^2 \hat{\mu}}{\partial y_j \partial y_k}(0) = - \int x_j x_k \mu dx$$

quindi è proprio la matrice di covarianza col segno opposto, quando la v.a. X è centrata.

Proprietà di indipendenza - Per concludere passiamo ad enunciare due importanti proposizioni riguardanti l'indipendenza di v.a. e le distribuzioni marginali delle stesse.

Proposizione 2.6.4. *Sia $X = (X_1, \dots, X_m)$ un vettore aleatorio con legge μ e siano le sue componenti X_1, \dots, X_m delle v.a. reali con le rispettive leggi di probabilità μ_1, \dots, μ_m . Esse sono indipendenti se e solo se vale*

$$\hat{\mu}(y_1, \dots, y_m) = \hat{\mu}_1(y_1) \cdots \hat{\mu}_m(y_m).$$

Proposizione 2.6.5. *Sia μ_i la i -esima distribuzione marginale di μ . Allora*

$$\hat{\mu}_i(y_i) = \hat{\mu}(0, \dots, 0, y_i, 0, \dots, 0).$$

2.7 Leggi normali

Tra le varie leggi di probabilità ce ne sono alcune particolarmente notevoli, sia perché ricorrono numerose volte nelle applicazioni, sia perché hanno proprietà che le rendono più facili da maneggiare.

Tra queste la più nota è la *legge normale*, la cui densità è conosciuta anche come funzione di Gauss. Questa distribuzione modella l'andamento probabilistico dei fenomeni casuali, come gli errori commessi nelle misure fisiche. Iniziamo col definire la legge normale nel caso unidimensionale.

Definizione 2.7.1. *Si dice che una legge di probabilità μ definita su \mathbb{R} è normale con media $a \in \mathbb{R}$ e varianza $\sigma > 0$ se ha per densità la funzione*

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

In tal caso si scrive $\mu \sim N(a, \sigma)$.

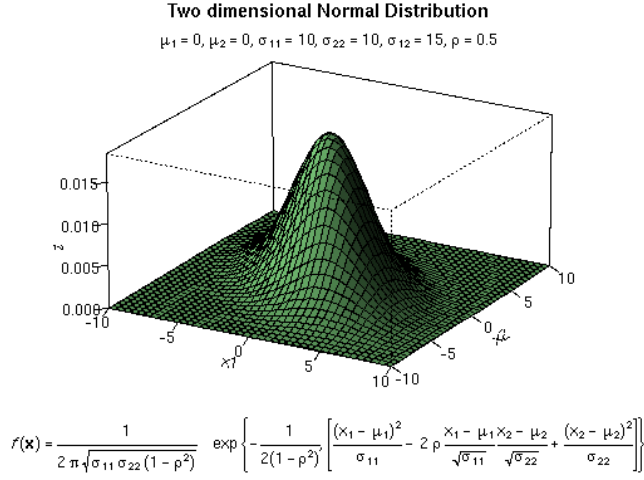


Figura 2.3: Grafico della legge normale in 2 dimensioni, disegnato con l'ausilio del software *R*.

Usando la funzione caratteristica $\hat{\mu}(y) = e^{iya}e^{-\frac{\sigma y^2}{2}}$ si può verificare che la definizione è ben posta, cioè che una distribuzione di questo tipo ha effettivamente media a e varianza σ .

Adesso andiamo a osservare un'interessante proprietà della legge normale. Siano X e Y due v.a. indipendenti con associate due leggi normali del tipo $\mu \sim N(a, \sigma)$ e $\nu \sim N(b, \tau)$ rispettivamente. Allora, per la proprietà (1) della funzione caratteristica vista in precedenza, per la v.a. $X + Y$ di legge $\mu \star \nu$ vale:

$$\widehat{\mu \star \nu}(y) = \hat{\mu}(y)\hat{\nu}(y) = e^{iy \cdot (a+b)}e^{-\frac{(\sigma+\tau)y^2}{2}}.$$

Per la proprietà (2), $\mu \star \nu$ è una legge a distribuzione normale, ed in particolare vale che

$$\mu \star \nu \sim N(a + b, \sigma + \tau).$$

Dunque, se X e Y sono v.a.r. indipendenti e normali, anche $X + Y$ è una v.a.r. normale.

In analogia con quanto fatto in precedenza possiamo ora generalizzare il concetto di legge normale al caso m -dimensionale.

Definizione 2.7.2. Una legge di probabilità μ definita su \mathbb{R}^m si dice normale di media $a \in \mathbb{R}^m$ e matrice di covarianza $\Gamma \in \mathcal{M}(m \times m, \mathbb{R})$ definita positiva se ha per densità

$$f(x) = \frac{1}{\sqrt{(2\pi)^m \det \Gamma}} e^{-\frac{1}{2}(\Gamma^{-1}(x-a)) \cdot (x-a)}$$

e si scriverà sempre $\mu \sim N(a, \Gamma)$.

Come prima, si può provare che effettivamente μ ha media a e matrice di covarianza Γ .

Si dimostra anche che:

$$\mu \sim N(a, \Gamma), \nu \sim N(b, \Lambda) \quad \Rightarrow \quad \mu \star \nu \sim N(a + b, \Gamma + \Lambda)$$

cioè che la somma di v.a. reali indipendenti e normali è una v.a. normale. Inoltre la sua media è la somma delle medie e la matrice di covarianza è data dalla somma delle matrici di covarianza delle due variabili.

Nel caso pluridimensionale possiamo poi vedere come si comportano le leggi marginali. Se μ_i è la distribuzione marginale i -esima di μ , allora sarà

$$\hat{\mu}_i(y) = e^{i y a_i} e^{-\frac{1}{2} \Gamma_{ii} y^2},$$

quindi anche la legge marginale è normale e vale $\mu_i \sim N(a_i, \Gamma_{ii})$.

In particolare notiamo che se Γ è diagonale si ha che

$$\hat{\mu}(y_1, \dots, y_m) = \hat{\mu}_1(y_1) \cdots \hat{\mu}_m(y_m)$$

e quindi per la proposizione (2.6.4) le leggi marginali sono indipendenti. Allora, per quanto noto sulla matrice di covarianza, possiamo enunciare la seguente proposizione.

Proposizione 2.7.3. *Sia $X = (X_1, \dots, X_m)$ una v.a. reale e normale. Allora le v.a. marginali X_1, \dots, X_m sono indipendenti se e solo se la matrice di covarianza è diagonale.*

Per concludere, osserviamo che si può costruire una legge normale a partire da un vettore media e una matrice di covarianza opportuna. Infatti, se Γ è una matrice simmetrica $m \times m$ semidefinita positiva, allora la funzione caratteristica

$$\hat{\mu}(y) = e^{i a \cdot y} e^{-\frac{1}{2} \Gamma y \cdot y}$$

definisce una legge di probabilità $\mu \sim N(a, \Gamma)$.

2.8 Il Lemma di Borel-Cantelli

In questa sezione riportiamo per completezza un risultato, il *Lemma di Borel-Cantelli*, che non risulta strettamente di interesse per quanto seguirà, ma che vale la pena introdurre per la sua notevole importanza (in particolare esso si rivela essenziale per dimostrare la *legge forte dei grandi numeri*).

Sullo spazio di probabilità (Ω, \mathcal{F}, P) consideriamo una successione $\{A_n\}_{n \in \mathbb{N}}$ di eventi. Poniamo

$$\limsup A_n = \bigcap_{n=1}^{\infty} \bigcup_{k \geq n} A_k = A.$$

Non è difficile vedere che l'insieme appena definito può essere interpretato nel modo seguente:

$$A = \{\omega \in \Omega \mid \omega \in A_n \text{ per infiniti indici } n\}.$$

Proposizione 2.8.1. (Lemma di Borel-Cantelli) *Con le notazioni appena definite, valgono i due fatti seguenti.*

(a) *Se si ha che*

$$\sum_{n=1}^{\infty} P(A_n) < +\infty$$

allora $P(A) = 0$.

(b) *Se invece*

$$\sum_{n=1}^{\infty} P(A_n) = +\infty$$

e per di più gli eventi A_n sono a due a due indipendenti, allora $P(A) = 1$.

In altre parole, se una successione di eventi ha probabilità sommabili, quasi certamente se ne verifica al più un numero finito; se al contrario ha probabilità non sommabili e gli eventi risultano indipendenti allora quasi certamente se ne verifica un numero infinito. In particolare, si deduce subito che effettuando un'infinità di prove indipendenti qualsiasi evento con probabilità positiva si verifica infinite volte.

A questa legge sono state associate svariate interpretazioni e applicazioni apparentemente paradossali, un esempio per tutti il cosiddetto *teorema della scimmia instancabile* (secondo il quale una scimmia che preme a caso su una tastiera per un tempo infinitamente lungo riuscirà quasi certamente a comporre un qualsiasi testo prefissato) o altri ancora. Naturalmente simili affermazioni non hanno significato da un punto di vista operativo; risulta però interessante il fatto che in senso matematico esse non risultino affatto contraddittorie, riferendosi ad eventi irrealizzabili in pratica ma a cui di fatto è possibile associare una probabilità, diciamo ϵ , misurabile e strettamente maggiore di zero.

Capitolo 3

Processi stocastici

Nel presente capitolo daremo una sintetica introduzione al concetto di *processo stocastico*, fondamentale in numerose applicazioni quando si voglia modellizzare l'evoluzione di situazioni o fenomeni di natura non deterministica (come esempi classici di questo tipo ricordiamo senz'altro il *moto browniano* e i processi di diffusione in generale). Parleremo poi del concetto di *tempo d'arresto*, mentre nella sezione conclusiva tenteremo di dare un'idea di come si proceda formalmente alla costruzione dei processi stessi, analizzando alcuni semplici esempi allo scopo di familiarizzare con i concetti introdotti.

3.1 Generalità

Definizione 3.1.1. *Si dice processo stocastico un oggetto della forma*

$$X = (\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in T}, (X_t)_{t \in T}, P) \quad (3.1.1)$$

dove

- (a) T (insieme dei tempi) è un sottoinsieme di \mathbb{R}^+ .
- (b) \mathcal{F} è una σ -algebra di parti di Ω .
- (c) P è una legge (misura) di probabilità sullo spazio misurabile (Ω, \mathcal{F}) .
- (d) $(\mathcal{F}_t)_{t \in T}$ è una filtrazione, ovvero una famiglia di sotto- σ -algebre di \mathcal{F} , crescente in t .
- (e) $(X_t)_{t \in T}$ è una famiglia di variabili aleatorie su (Ω, \mathcal{F}) a valori nello spazio misurabile $(E, \mathcal{B}(E))$, costituito da uno spazio topologico E (detto spazio degli stati) con la σ -algebra boreliana, e tale che X_t risulta \mathcal{F}_t -misurabile

per ogni istante t . Quest'ultimo fatto si esprime anche dicendo che $(X_t)_t$ è adattato alla filtrazione $(\mathcal{F}_t)_t$.

Nel caso in cui le \mathcal{F}_t non siano esplicitamente indicate, si sottintende l'applicazione di una filtrazione naturale $\mathcal{F}_t = \sigma(X_s, s \leq t)$.

Il processo stocastico è detto *a parametro discreto* se T è discreto, mentre è detto *a parametro continuo* qualora T sia continuo.

Osservazione 3.1.2. Notiamo che mediante l'applicazione $\omega \rightarrow \{t \rightarrow X_t(\omega)\}$ si può sempre supporre che $\Omega \subset E^T$. Per questo motivo, Ω viene anche detto spazio delle traiettorie (costituisce in un certo senso l'insieme delle possibili realizzazioni del processo).

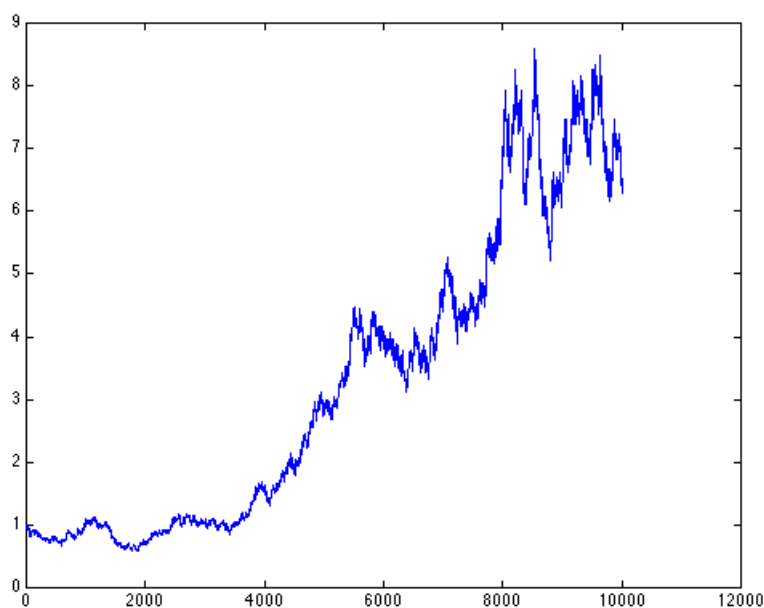


Figura 3.1: Un esempio basilare di processo stocastico: il moto browniano unidimensionale, simulato in *Matlab*.

Nei casi che ci interesseranno, $(E, \mathcal{B}(E))$ coinciderà sempre con lo spazio $(\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m))$ e T con un intervallo o una semiretta di \mathbb{R}^+ , cosa che sottintenderemo d'ora in avanti. In generale però (se ad esempio volessimo analizzare processi a parametro discreto) tale insieme potrebbe anche essere \mathbb{N} o una sua parte, come accade nell'esempio riportato di seguito.

Esempio 3.1.3. Consideriamo una partita di 30 lanci di una moneta, nella quale le regole del gioco stabiliscono che:

- quando esce testa, si guadagna un punto;
- quando esce croce, se ne perde uno.

Una buona scelta di un processo stocastico per studiare questo fenomeno può essere la seguente.

- (a) $T = \{1, 2, \dots, 30\}$.
- (b) Posto $C = \{-1, 1\}$, prendiamo $\Omega = \{-1, 1\}^{30}$. \mathcal{F} sarà l'insieme delle parti di Ω .
- (c) Detta μ la distribuzione uniforme su C , cioè $\mu(\{1\}) = \mu(\{-1\}) = \frac{1}{2}$, P sarà la legge prodotto $\mu \otimes \dots \otimes \mu$ (30 volte) su Ω .
- (d) Gli eventi $\omega \in \Omega$ (ovvero i possibili risultati di una partita) sono della forma $\omega = \{\omega_1, \dots, \omega_{30}\}$. Si pone allora $X_n(\omega) = \omega_1 + \dots + \omega_n$ (v.a. che rappresenta il punteggio ottenuto dopo n lanci).
- (e) \mathcal{F}_n è la σ -algebra $\sigma(X_1, \dots, X_n)$, cioè quella generata dagli eventi della forma $A_1 \times \dots \times A_n \times C \times \dots \times C$, dove $A_i \subset C$.

Da un punto di vista intuitivo, si può interpretare la σ -algebra \mathcal{F}_t come l'insieme degli eventi di cui al tempo t si può dire se sono verificati o meno. Nel presente caso, la forma assunta dalla σ -algebra \mathcal{F}_t esprime il fatto che al tempo n si possono conoscere i risultati dei primi n lanci, ma non quelli dei successivi.

Si ha in sostanza un processo stocastico caratterizzato da una sequenza di v.a. X_1, \dots, X_{30} e adattato alla filtrazione $\mathcal{F}_1 \subset \dots \subset \mathcal{F}_{30}$, nel quale, per costruzione, ciascuna variabile X_n diventa un valore certo non appena siano noti i primi n risultati. In altri termini, con il livello di informazione rappresentato da \mathcal{F}_n la variabile X_n risulta completamente determinata, dunque \mathcal{F}_n -misurabile, come vuole la definizione (3.1.1).

Vediamo adesso qualche ulteriore caratteristica dei processi stocastici, ad esempio come si definisce rigorosamente il fatto che processi diversi possano avere delle proprietà in qualche senso simili.

Definizione 3.1.4. Dati due processi stocastici $X = (\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in T}, (X_t)_{t \in T}, P)$ e $X' = (\Omega', \mathcal{F}', (\mathcal{F}'_t)_{t \in T}, (X'_t)_{t \in T}, P')$, si dice che essi:

- sono equivalenti se per ogni sequenza di istanti temporali t_1, \dots, t_n in T le v.a. $(X_{t_1}, \dots, X_{t_n})$ e $(X'_{t_1}, \dots, X'_{t_n})$ hanno la stessa legge;

- sono una modificazione (o una versione) l'uno dell'altro se $(\Omega, \mathcal{F}, P) = (\Omega', \mathcal{F}', P')$ e se $\forall t \in T$ si ha $X_t = X'_t$ quasi certamente¹ (q.c.) rispetto a P ;
- sono indistinguibili se $\Omega = \Omega'$, $\mathcal{F} = \mathcal{F}'$, $P = P'$ e inoltre si ha che $P(X_t = X'_t \ \forall t) = 1$.

Osservazione 3.1.5. *Dalle definizioni precedenti segue subito che se due processi sono una modificazione l'uno dell'altro allora in particolare essi sono equivalenti. In modo simile si osserva che due processi indistinguibili sono anche una modificazione l'uno dell'altro, ma il viceversa è falso, come mostra l'esempio seguente.*

Esempio 3.1.6. *Consideriamo due processi definiti dagli spazi di probabilità dati da $\Omega = \Omega' = [0, 1]$, $\mathcal{F} = \mathcal{F}' = \mathcal{B}([0, 1])$, $P = P' =$ misura di Lebesgue e dalle v.a. $X_t(\omega) = \mathbf{1}_{\{\omega\}}(t)$ e $X'_t(\omega) = 0 \ \forall t \in T, \omega \in \Omega$. Risulta immediato verificare che il secondo costituisce una modificazione del primo, ma che i due processi non sono indistinguibili.*

Forniamo di seguito alcuni risultati più tecnici sui processi stocastici che risultano utili per il prosieguo della teoria. Per semplicità, denoteremo d'ora in avanti un processo con il nome della variabile che lo caratterizza, diciamo X , tenendone presente la definizione formale data in precedenza.

Definizione 3.1.7. *Un processo stocastico X si dice continuo (risp. q.c. continuo) se $\forall \omega \in \Omega$ (risp. per quasi ogni $\omega \in \Omega$) l'applicazione $(t, \omega) \rightarrow X_t(\omega)$ risulta continua rispetto a t . Analogamente si darà la definizione di processo continuo a destra e di processo q.c. continuo a destra.*

Definizione 3.1.8. *Diremo che un processo X è:*

- misurabile se l'applicazione $(t, \omega) \rightarrow X_t(\omega)$ è misurabile dall'insieme $(T \times \Omega, \mathcal{B}(T) \times \mathcal{F})$ in $(E, \mathcal{B}(E))$;
- progressivamente misurabile (o semplicemente progressivo) se $\forall u \in T$ l'applicazione $(t, \omega) \rightarrow X_t(\omega)$ è misurabile da $([0, u] \times \Omega, \mathcal{B}([0, u]) \times \mathcal{F}_u)$ in $(E, \mathcal{B}(E))$.

In particolare un processo progressivo risulta, evidentemente, misurabile.

Proposizione 3.1.9. *Se X è continuo a destra, allora in particolare è anche progressivamente misurabile.*

¹Ovvero quasi ovunque rispetto alla misura di probabilità P .

Dimostrazione. Fissato $u \in T$, e preso $k = 0, \dots, 2^n - 1$, poniamo

$$X_s^{(n)} = \begin{cases} \frac{X_{k+1}}{2^n}, & \text{se } s \in [\frac{k}{2^n}u, \frac{k+1}{2^n}u[\\ X_u, & \text{se } s = u \end{cases}$$

Evidentemente, si ha che $X_s^{(n)} \rightarrow X_s$ per ogni $s \leq u$. Inoltre, $X^{(n)}$ è progressivo: se infatti $\Gamma \in \mathcal{B}(E)$ si ha

$$\{(s, \omega) | s \leq u, X_s^{(n)}(\omega) \in \Gamma\} = \bigcup_{k < 2^n} ([\frac{ku}{2^n}, \frac{(k+1)u}{2^n}[\times \{\frac{X_{k+1}}{2^n} \in \Gamma\}) \cup (\{u\} \times \{X_u \in \Gamma\})$$

che è un elemento di $\mathcal{B}([0, u]) \times \mathcal{F}_u$. L'applicazione $(s, \omega) \rightarrow X_s(\omega)$ risulta dunque $\mathcal{B}([0, u]) \times \mathcal{F}_u$ -misurabile in quanto limite di funzioni $\mathcal{B}([0, u]) \times \mathcal{F}_u$ -misurabili, da cui la tesi. \square

Il prossimo risultato riveste una notevole importanza teorica, ma ne omettiamo la dimostrazione, la quale in ogni caso risulta strettamente tecnica e non aggiunge granchè dal punto di vista concettuale. (È comunque possibile trovarla esposta per intero in [1]).

Teorema 3.1.10. (di continuità) *Sia D un aperto di \mathbb{R}^m e $(X_y)_{y \in D}$, una famiglia di v.a. d -dimensionali su (Ω, \mathcal{F}, P) tale che esistano $\alpha > 0$, $\beta > 0$, $c > 0$ per cui vale*

$$E[|X_y - X_z|^\beta] \leq c|y - z|^{m+\alpha}.$$

Allora esiste una famiglia $(\tilde{X}_y)_y$ di v.a. a valori in \mathbb{R}^d tale che

$$X_y = \tilde{X}_y \text{ q.c. per ogni } y$$

e tale che per ogni $\omega \in \Omega$ l'applicazione $y \rightarrow \tilde{X}_y(\omega)$ è continua ed anzi hölderiana² di esponente γ per ogni $\gamma < \frac{\alpha}{\beta}$ su ogni compatto di D .

Corollario 3.1.11. *Sia X un processo a valori in \mathbb{R}^d tale che esistano $\alpha > 0$, $\beta > 0$, $c > 0$ per cui vale*

$$E[|X_t - X_s|^\beta] \leq c|t - s|^{1+\alpha}.$$

Allora esiste una modificazione Y di X che è continua. Per di più per ogni $\gamma < \frac{\alpha}{\beta}$ le traiettorie di Y sono hölderiane di esponente γ su ogni intervallo limitato.

²Ricordiamo che una funzione f di variabile reale si dice *hölderiana* con esponente di Hölder $\alpha \in [0, 1]$ se $\forall x, y$ vale $|f(x) - f(y)| \leq C|x - y|^\alpha$ per qualche $C \geq 0$.

Per concludere il presente paragrafo, parliamo un po' più in dettaglio del ruolo rivestito dalle filtrazioni.

Preso un generico processo stocastico come definito dalla (3.1.1), poniamo $\mathcal{F}_{t+} = \bigcap_{\epsilon > 0} \mathcal{F}_{t+\epsilon}$. Chiaramente si ha $\mathcal{F}_{t+} \supseteq \mathcal{F}_t$. Diremo che la filtrazione è *continua a destra* se $\mathcal{F}_{t+} = \mathcal{F}_t \forall t \in T$.

Definizione 3.1.12. *Un processo stocastico si dice standard se:*

- (a) *la filtrazione $(\mathcal{F}_t)_t$ è continua a destra;*
- (b) *per ogni $t \in T$, si ha che \mathcal{F}_t contiene tutti gli eventi di \mathcal{F} trascurabili (ovvero di misura nulla rispetto a P).*

Osserviamo come l'ipotesi che il processo sia standard riguardi unicamente la filtrazione del processo e non, ad esempio, le v.a. X_t .

Lavorare con un processo standard risulta spesso essenziale per evitare alcune difficoltà tecniche, come quelle che insorgono nell'esempio che segue.

Esempio 3.1.13. *Sia X un processo e $(Y_t)_t$ una famiglia di v.a. tali che $X_t = Y_t$ q.c. rispetto a P per ogni $t \in T$. In generale, questo non implica che $Y = (\Omega, \mathcal{F}, (\mathcal{F}_t)_t, (Y_t)_t, P)$ sia anch'esso un processo stocastico. Infatti, un evento trascurabile N_t su cui X_t e Y_t differiscono potrebbe non appartenere ad \mathcal{F}_t e quindi Y_t non essere \mathcal{F}_t -misurabile. Tale inconveniente invece non sussiste se chiediamo che X sia un processo standard.*

Notiamo inoltre che considerare processi standard implica che ogni processo q. c. continuo viene ad avere una modificazione anch'essa continua.

Anche il fatto che la filtrazione $(\mathcal{F}_t)_t$ sia continua a destra è un'ipotesi tecnica che nello sviluppo della teoria si rivela inevitabile, e per questo motivo molte volte la prima mossa da fare è cercare di provare che si può supporre che i processi a cui ci si interessa godano delle proprietà (a) e (b).

3.2 Tempi d'arresto

Introduciamo ora il fondamentale concetto di tempo d'arresto, mostrando poi, mediante alcuni risultati, l'importanza che esso riveste nello studio dei processi.

Definizione 3.2.1. *Sia $(\mathcal{F}_t)_{t \in T}$ una filtrazione. Una v.a. $\tau : \Omega \rightarrow T \cup \{+\infty\}$ si dice un tempo d'arresto se, posto $\{\tau \leq t\} = \{\omega \in \Omega | \tau(\omega) \leq t\}$, si ha che $\{\tau \leq t\} \in \mathcal{F}_t \forall t \in T$. Si pone inoltre*

$$\mathcal{F}_\tau = \{A \in \mathcal{F}_\infty | A \cap \{\tau \leq t\} \in \mathcal{F}_t \forall t \in T\}$$

dove si è definito $\mathcal{F}_\infty = \bigcup_{t \in T} \mathcal{F}_t$.

Esempio 3.2.2. Per fissare le idee, riprendiamo il processo definito nell'esempio 3.1.3. Imporre che $\{\tau \leq n\} \in \mathcal{F}_n$ non significa altro che richiedere che $\tau(\omega) \leq n$ dipenda solo dalle prime n coordinate. In questo caso, per esempio, è un tempo d'arresto (o semplicemente t.a. per brevità) la v.a.

$$\tau(\omega) = \inf\{n | \omega_1 + \dots + \omega_n \geq 13\}$$

mentre non lo è la v.a.

$$\sigma(\omega) = \sup\{n | \omega_1 + \dots + \omega_n \geq 13\}$$

poiché, fissato \bar{n} per sapere se $\sigma(\omega) \leq \bar{n}$ bisogna conoscere anche le coordinate di ω di indice $> \bar{n}$.

Di nuovo, risulta utile a livello intuitivo immaginarsi \mathcal{F}_τ come la σ -algebra degli eventi di cui al tempo τ si può dire se sono o no verificati.

I tempi d'arresto sono uno strumento fondamentale in numerosi contesti: basti pensare ai problemi di *optimal stopping* nella teoria dei controlli (dove si presenta la questione di decidere in maniera ottimale se fermare o meno un certo processo in base allo stato presente e agli eventi passati), alla teoria delle *Martingale* in probabilità, ai tempi di collisione in un moto browniano o ad altre situazioni ancora. Facciamo qualche semplice esempio di come sfruttando le definizioni appena date si possano costruire delle *regole di fermata*.

Esempio 3.2.3. Supponiamo di voler creare una strategia per un gioco d'azzardo, diciamo a puntata costante, in cui abbiamo a disposizione una certa somma iniziale, in particolare di voler decidere quando interrompere la nostra partita.

- Effettuare una ed una sola giocata corrisponde ad un tempo d'arresto $\tau = 1$ e costituisce banalmente una regola di fermata.
- Giocare finchè o non si perde tutta la somma iniziale o si è effettuato un numero massimo fissato di puntate è anch'essa una strategia di arresto valida.
- Giocare finchè non si è raggiunto il massimo guadagno possibile non è una regola valida perchè risulta impossibile, ad un certo istante, sapere se tale criterio è verificato o no: per sapere se interrompere il gioco dovremmo essere a conoscenza anche degli esiti di tutte le partite future.

Concludiamo il paragrafo con alcuni ulteriori risultati. Del secondo, forniamo la dimostrazione, mentre per quanto riguarda il terzo rimandiamo sempre a [1].

Proposizione 3.2.4. Consideriamo due t.a. σ e τ . Si ha che:

- (a) τ è sempre \mathcal{F}_t -misurabile;
- (b) $\min(\sigma, \tau)$ e $\max(\sigma, \tau)$ sono t.a.;
- (c) se $\sigma \leq \tau$ allora $\mathcal{F}_\sigma \subseteq \mathcal{F}_\tau$.

Proposizione 3.2.5. Se X è un processo misurabile e $\sigma : \Omega \rightarrow \mathbb{R}^+$ una v.a. allora anche $X_\sigma : \omega \rightarrow X_{\sigma(\omega)}(\omega)$ è una v.a.. Inoltre, se τ è un t.a. e X è progressivamente misurabile, allora X_τ risulta \mathcal{F}_τ -misurabile.

Dimostrazione. La prima asserzione è immediata, perché X_σ si ottiene componendo le applicazioni misurabili $\omega \rightarrow (\omega, \sigma(\omega))$ e $(\omega, t) \rightarrow X_t(\omega)$. Se poi X è progressivo e $\Omega_t = \{\tau \leq t\}$, allora X_τ è misurabile da $(\Omega_t, \mathcal{F}_t)$ in $(E, \mathcal{B}(E))$ come composizione delle due applicazioni misurabili

$$\omega \rightarrow (\omega, \tau(\omega)) \text{ da } (\Omega_t, \mathcal{F}_t) \text{ in } (\Omega_t \times [0, t], \mathcal{F}_t \times \mathcal{B}([0, t]))$$

e

$$(\omega, t) \rightarrow X_t(\omega) \text{ da } (\Omega_t \times [0, t], \mathcal{F}_t \times \mathcal{B}([0, t])) \text{ in } (E, \mathcal{B}(E)).$$

Dunque, se $\Gamma \in \mathcal{B}(E)$, si ha che

$$\{\tau \leq t\} \cap \{X_\tau \in \Gamma\} = \{\omega \in \Omega_t, X_\tau(\omega) \in \Gamma\} \in \mathcal{F}_t$$

e quindi $\{X_\tau \in \Gamma\} \in \mathcal{F}_\tau$. □

Se $A \in \mathcal{B}(E)$, definiamo il *tempo di uscita* da A come

$$\tau_A = \inf\{t \in T \mid X_t \notin A\}.$$

Proposizione 3.2.6. Se E è uno spazio metrico e X un processo stocastico continuo, allora:

- (a) se A è aperto, τ_A è un tempo d'arresto.
- (b) se F è chiuso e la filtrazione $(\mathcal{F}_t)_t$ è continua a destra, τ_F è un tempo d'arresto.

3.3 Costruzione di processi stocastici

Diamo per finire alcuni brevi cenni ai risultati che riguardano la costruzione di un processo stocastico.

Definizione 3.3.1. *Sia $X = (\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in T}, (X_t)_{t \in T}, P)$ un processo a valori nello spazio topologico $(E, \mathcal{B}(E))$, e sia $\pi = (t_1, \dots, t_n)$ una n -upla di elementi di T con $t_1 < \dots < t_n$. Allora su E^n si può considerare la legge di probabilità μ_π , immagine di P tramite l'applicazione*

$$X_\pi = (X_{t_1}, \dots, X_{t_n}) : \Omega \rightarrow E^n.$$

Al variare di π , le μ_π si chiamano distribuzioni di dimensione finita di P .

La conoscenza delle distribuzioni di dimensione finita caratterizza P nel senso seguente.

Proposizione 3.3.2. *Se due processi $X = (\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in T}, (X_t)_{t \in T}, P)$ e $X' = (\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in T}, (X_t)_{t \in T}, Q)$ hanno le stesse distribuzioni di dimensione finita, allora P e Q coincidono sulla σ -algebra $\sigma(X_t, t \in T)$.*

Dimostrazione. Se X e X' hanno le stesse distribuzioni di dimensione finita, allora P e Q coincidono sugli insiemi del tipo

$$X_{t_1}^{-1}(\Gamma_1) \cap \dots \cap X_{t_n}^{-1}(\Gamma_n)$$

dove $\Gamma_1, \dots, \Gamma_n \in \mathcal{B}(E)$. Poiché questi insiemi costituiscono una classe stabile per l'intersezione finita, per il criterio di Caratheodory P e Q coincidono sulla σ -algebra da essi generata, che è appunto $\sigma(X_t, t \in T)$, il che conclude la dimostrazione. \square

Un problema che ricorre spesso è l'inverso di quello appena introdotto: assegnato uno spazio topologico E e l'insieme dei tempi T , costruire un processo stocastico $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in T}, (X_t)_{t \in T}, P)$ avente E come spazio degli stati e di assegnate distribuzioni di dimensione finita $(\mu_\pi)_{\pi \in \Pi}$.

Osservazione 3.3.3. *Anzitutto, è chiaro che le μ_π devono soddisfare la condizione di consistenza definita come segue.*

- Sia $\pi = (t_1, \dots, t_n)$, $\pi' = (t_1, \dots, t_{i-1}, t_{i+1}, \dots, t_n)$.
- Sia poi $p_i : E^n \rightarrow E^{n-1}$ l'applicazione definita da $(x_1, \dots, x_n) \rightarrow (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$.
- Allora l'immagine di μ_π tramite p_i è $\mu_{\pi'}$.

Questo perché, come si vede immediatamente, si ha $X_{\pi'} = p_i \circ X_\pi$.

La condizione precedente è anche sufficiente se lo spazio topologico E è sufficientemente regolare, come mostra il teorema seguente.

Teorema 3.3.4. (Kolmogorov) *Sia E uno spazio metrico σ -compatto³, e sia $(\mu_\pi)_{\pi \in \Pi}$ un sistema di distribuzioni di dimensione finita su E che verifichi la condizione di consistenza. Poniamo $\Omega = E^T$, e definiamo poi $X_t(\omega) = \omega(t)$, $\mathcal{F}_t = \sigma(X_s, T \ni s \leq t)$, $\mathcal{F} = \sigma(X_t, t \in T)$. Esiste allora su (Ω, \mathcal{F}) una legge di probabilità P tale che le $(\mu_\pi)_{\pi \in \Pi}$ sono le distribuzioni di dimensione finita del processo $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in T}, (X_t)_{t \in T}, P)$.*

La P definita dal teorema 3.3.4 è evidentemente unica per la proposizione 3.3.2.

Tirando le somme, nel presente capitolo abbiamo già incontrato alcuni dei problemi più importanti che intervengono nello studio dei processi stocastici:

- (a) la costruzione di processi che soddisfino particolari proprietà (riconducibili alle distribuzioni a dimensione finita);
- (b) la regolarità delle traiettorie;
- (c) la determinazione della probabilità P del processo.

Osserviamo che P è invariante per il passaggio da un processo ad una sua modificazione, a differenza della regolarità delle traiettorie, come mostra anche l'esempio 3.1.6.

I capitoli della teoria dei processi stocastici successivi a questa introduzione riguardano lo studio di classi particolari di processi, per esempio quelli di diffusione. Per trattare tali argomenti risulta necessario procurarsi delle tecniche (in particolar modo integrali stocastici) che insieme al teorema di Kolmogorov permettano prima di costruire i processi e poi di provarne l'esistenza di versioni regolari. La determinazione della probabilità P risulta essere (tranne in casi particolari) il problema più difficile, ma è possibile mostrare che in molte situazioni la probabilità di alcuni eventi si può ricondurre allo studio di problemi alle derivate parziali.

³Uno spazio topologico si dice σ -compatto se si può esprimere come unione di un'infinità numerabile di sottoinsiemi compatti.



Figura 3.2: Processi stocastici come il *random walk* intervengono anche in situazioni molto complesse come ad esempio in questa simulazione di un fenomeno di *aggregazione a diffusione limitata* di particelle.

Bibliografia

- [1] P. BALDI, *Equazioni differenziali stocastiche e applicazioni*, Quaderni dell'Unione Matematica Italiana, 2000.
- [2] R. MAGNANINI, *Istituzioni di Analisi Superiore I (dispense del corso)*, 2008.