

Capitolo 7

Funzioni caratteristiche e generatrici

7.1. Funzioni caratteristiche

La funzione caratteristica di una v.a. è una particolare funzione a valori complessi, le cui proprietà analitiche sono molto utili per ottenere leggi di probabilità o momenti di v.a. trasformate. Nel prossimo Capitolo, la funzione caratteristica sarà anche centrale nello studio del comportamento asintotico di successioni di v.a.

Prima di introdurre la definizione di funzione caratteristica è conveniente estendere il concetto di v.a. all'insieme dei numeri complessi. Per quanto riguarda la notazione adottata in questo Capitolo, $i = \sqrt{-1}$ rappresenta come al solito l'unità immaginaria. Inoltre, dato il numero complesso $c = a + ib$, allora la parte reale e immaginaria di c sono rispettivamente denotate con $\Re(c) = a$ e $\Im(c) = b$, mentre il modulo di c è dato da $|c| = \sqrt{a^2 + b^2}$.

Definizione 7.1.1. Se (Ω, \mathcal{F}, P) è uno spazio probabilizzato, un'applicazione $X : \Omega \rightarrow \mathbb{C}$ è detta *variabile aleatoria complessa* se $\Re(X)$ e $\Im(X)$ sono variabili aleatorie. \square

Se X è una v.a. complessa, allora il relativo valore atteso è definito come

$$E[X] = E[\Re(X)] + i E[\Im(X)].$$

Affinchè il valore atteso di una v.a. complessa esista finito è sufficiente che il valore atteso

$$E[|X|] = E[\sqrt{\Re(X)^2 + \Im(X)^2}]$$

esista finito. Infatti, le proprietà del valore atteso di una v.a. complessa sono simili a quelle di una v.a. reale e in particolare si ha $|E[X]| \leq E[|X|]$ (si veda Billingsley, 1995, p.218). Inoltre, vale anche $E[cX] = cE[X]$ per $c \in \mathbb{C}$.

La definizione formale di funzione caratteristica è quindi la seguente.

Definizione 7.1.2. Data la v.a. X definita sullo spazio probabilizzato (Ω, \mathcal{F}, P) , si dice *funzione caratteristica* (f.c.) della v.a. X la funzione $\varphi_X : \mathbb{R} \rightarrow \mathbb{C}$

$$\varphi_X(t) = E[\cos(tX)] + i E[\sin(tX)] = E[e^{itX}],$$

dove $t \in \mathbb{R}$. \square

Nel linguaggio della Teoria della Misura, la f.c. è in effetti la trasformata di Fourier (a meno di una costante moltiplicativa). Tenendo presente quanto detto nella Sezione 4.1, se la v.a. X è discreta a valori su insieme numerabile S con f.p. p_X , la f.c. si riduce a

$$\varphi_X(t) = \sum_{x \in S} \cos(tx) p_X(x) + i \sum_{x \in S} \sin(tx) p_X(x) = \sum_{x \in S} e^{itx} p_X(x),$$

mentre, se la v.a. assolutamente continua X ammette d.p. f_X , la f.c. si riduce a

$$\varphi_X(t) = \int_{-\infty}^{\infty} \cos(tx) f_X(x) dx + i \int_{-\infty}^{\infty} \sin(tx) f_X(x) dx = \int_{-\infty}^{\infty} e^{itx} f_X(x) dx .$$

• **Esempio 7.1.1.** Si consideri la v.a. X distribuita con legge di Bernoulli $\mathcal{B}(1, p)$. In questo caso, la f.c. della v.a. X è data da

$$\varphi_X(t) = \sum_{x=0}^1 \cos(tx) p^x q^{1-x} + i \sum_{x=0}^1 \sin(tx) p^x q^{1-x} = q + p \cos(t) + ip \sin(t) = q + pe^{it} .$$

Se si considera più in generale la v.a. X distribuita con legge Binomiale $\mathcal{B}(n, p)$, tenendo presente l'espressione del binomio di Newton in \mathbb{C} , si ha

$$\varphi_X(t) = \sum_{x=0}^n e^{itx} \binom{n}{x} p^x q^{n-x} = \sum_{x=0}^n \binom{n}{x} (pe^{it})^x q^{n-x} = (q + pe^{it})^n . \quad \square$$

• **Esempio 7.1.2.** Si consideri la v.a. X distribuita con legge di Poisson $\mathcal{P}(\lambda)$. Tenendo presente l'espressione della serie esponenziale in \mathbb{C} , si ha

$$\varphi_X(t) = \sum_{x=0}^{\infty} e^{itx} e^{-\lambda} \frac{\lambda^x}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^{it})^x}{x!} = e^{\lambda(e^{it}-1)} . \quad \square$$

• **Esempio 7.1.3.** Si consideri la v.a. X distribuita con legge Binomiale Negativa $\mathcal{BN}(k, p)$. Tenendo presente l'espressione della serie binomiale negativa in \mathbb{C} , si ha

$$\varphi_X(t) = \sum_{x=0}^{\infty} e^{itx} \binom{k+x-1}{k-1} p^k q^x = p^k \sum_{x=0}^{\infty} \binom{k+x-1}{x} (qe^{it})^x = \left(\frac{p}{1 - qe^{it}} \right)^k . \quad \square$$

• **Esempio 7.1.4.** Si consideri la v.a. Z distribuita con legge Normale ridotta $\mathcal{N}(0, 1)$. Tenendo presente che Z è simmetrica rispetto all'origine, si ha

$$E[\sin(tZ)] = \int_{-\infty}^{\infty} \sin(tz) \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz = 0 ,$$

e quindi

$$\varphi_Z(t) = E[\cos(tZ)] = \int_{-\infty}^{\infty} \cos(tz) \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz = e^{-\frac{1}{2}t^2} .$$

Per ulteriori dettagli su questo risultato si veda l'Esercizio 7.1.1. Si osservi che in questo caso $\varphi_Z(t)$ è una funzione puramente reale. In generale, risulta immediato verificare che la f.c. è una funzione puramente reale se una v.a. è simmetrica rispetto all'origine. \square

• **Esempio 7.1.5.** Si consideri la v.a. Z distribuita con legge Esponenziale ridotta $\mathcal{G}(0, 1, 1)$. Dal momento che

$$E[\cos(tZ)] = \int_0^{\infty} \cos(tz) e^{-z} dz = \frac{1}{1+t^2}$$

e

$$E[\sin(tZ)] = \int_0^\infty \sin(tz) e^{-z} dz = \frac{t}{1+t^2},$$

allora si ha

$$\varphi_Z(t) = \frac{1}{1+t^2} + i \frac{t}{1+t^2} = (1-it)^{-1}.$$

Più in generale, se si considera la v.a. Z distribuita con legge Gamma ridotta $\mathcal{G}(0, 1, k)$, tenendo presente le espressioni delle serie esponenziale e binomiale negativa in \mathbb{C} , si ha

$$\begin{aligned} \varphi_Z(t) &= \int_0^\infty e^{itz} \frac{1}{\Gamma(k)} z^{k-1} e^{-z} dz = \frac{1}{\Gamma(k)} \int_0^\infty \sum_{n=0}^\infty \frac{(itz)^n}{n!} z^{k-1} e^{-z} dz \\ &= \frac{1}{\Gamma(k)} \sum_{n=0}^\infty \frac{(it)^n}{n!} \int_0^\infty z^{n+k-1} e^{-z} dz = \sum_{n=0}^\infty \frac{\Gamma(n+k)}{n! \Gamma(k)} (it)^n \\ &= \sum_{n=0}^\infty \binom{n+k-1}{n} (it)^n = (1-it)^{-k}. \end{aligned}$$

Ovviamente, l'espressione coincide con quella ottenuta per la legge Esponenziale ridotta se $k = 1$. \square

Di seguito vengono considerate una serie di proprietà che riguardano l'esistenza della f.c., la continuità uniforme della f.c. e l'espressione della f.c. di particolari v.a. trasformate.

Proposizione 7.1.3. *Se X è una v.a. definita sullo spazio probabilizzato (Ω, \mathcal{F}, P) con f.c. φ_X , allora $|\varphi_X(t)| \leq 1$ per ogni $t \in \mathbb{R}$ e in particolare $\varphi_X(0) = 1$.*

Dimostrazione. Tenendo presente che per ogni $x \in \mathbb{R}$ si ha

$$|e^{itx}| = |\cos(tx) + i \sin(tx)| = \sqrt{\cos^2(tx) + \sin^2(tx)} = 1,$$

allora

$$|\varphi_X(t)| = |E[e^{itX}]| \leq E[|e^{itX}|] = 1.$$

Inoltre, si ha

$$|\varphi_X(0)| = \int dP = 1. \quad \square$$

Proposizione 7.1.4. *Se X è una v.a. definita sullo spazio probabilizzato (Ω, \mathcal{F}, P) con f.c. φ_X , allora φ_X è uniformemente continua su \mathbb{R} .*

Dimostrazione. Per $h \in \mathbb{R}$ si ha

$$\begin{aligned} |\varphi_X(t+h) - \varphi_X(t)| &= |E[e^{itX}(e^{ihX} - 1)]| \leq E[|e^{itX}(e^{ihX} - 1)|] \\ &= E[|e^{itX}| \cdot |e^{ihX} - 1|] = E[|e^{ihX} - 1|], \end{aligned}$$

essendo $|e^{itX}| = 1$. Dal momento che si hanno le disuguaglianze $|e^{ihx} - 1| \leq 2$ e $|e^{ihx} - 1| \leq |hx|$ per ogni $x \in \mathbb{R}$, per $a \in]0, \infty[$ risulta

$$\begin{aligned} E[|e^{ihX} - 1|] &= E[|e^{ihX} - 1| \mathbf{1}_{[0,a]}(|X|)] + E[|e^{ihX} - 1| \mathbf{1}_{]a,\infty[}(|X|)] \\ &\leq E[|hX| \mathbf{1}_{[0,a]}(|X|)] + 2E[\mathbf{1}_{]a,\infty[}(|X|)] = |h| E[|X| \mathbf{1}_{[0,a]}(|X|)] + 2P(|X| > a) \\ &\leq a|h| + 2P(|X| > a). \end{aligned}$$

Dunque, fissato $\varepsilon > 0$, se h e a sono scelti in modo tale che $a|h| < \frac{\varepsilon}{2}$ e $P(|X| > a) < \frac{\varepsilon}{4}$, si ha

$$|\varphi_X(t+h) - \varphi_X(t)| \leq a|h| + 2P(|X| > a) < \varepsilon,$$

da cui segue la continuità uniforme, dal momento che ε non dipende da t . \square

Proposizione 7.1.5. *Se X è una v.a. definita sullo spazio probabilizzato (Ω, \mathcal{F}, P) con f.c. φ_X , allora la f.c. della v.a. trasformata $Y = a + bX$, dove $a, b \in \mathbb{R}$, è data da*

$$\varphi_Y(t) = e^{iat} \varphi_X(bt).$$

Dimostrazione. Si ha

$$\varphi_Y(t) = E[e^{it(a+bX)}] = e^{iat} E[e^{ibtX}] = e^{iat} \varphi_X(bt),$$

che è quanto si voleva dimostrare. \square

• **Esempio 7.1.6.** Si consideri la v.a. X distribuita con legge Normale $\mathcal{N}(\mu, \sigma^2)$. Tenendo presente l'Esempio 7.1.4, dal momento che $X = \mu + \sigma Z$, si ha

$$\varphi_X(t) = e^{i\mu t} \varphi_Z(\sigma t) = e^{i\mu t - \frac{1}{2}\sigma^2 t^2}.$$

In questo caso, al contrario della legge ridotta $\mathcal{N}(0, 1)$, la f.c. φ_X non è puramente reale, dal momento che la v.a. X non è simmetrica rispetto all'origine. \square

• **Esempio 7.1.7.** Si consideri la v.a. X distribuita con legge Gamma $\mathcal{G}(a, b, k)$. Tenendo presente l'Esempio 7.1.5, dal momento che $X = a + bZ$, si ha

$$\varphi_X(t) = e^{iat} \varphi_Z(bt) = e^{iat} (1 - ibt)^{-k}.$$

In particolare, se la v.a. X si distribuisce con legge Chi-quadrato χ_n^2 , dalla precedente espressione con $a = 0$, $b = 2$ e $k = \frac{n}{2}$ risulta

$$\varphi_X(t) = (1 - 2it)^{-\frac{1}{2}n}. \quad \square$$

Le seguenti Proposizioni forniscono rispettivamente un'importante relazione fra la f.c. e i momenti di una v.a., e lo sviluppo in serie della f.c. basato sui momenti in un intorno di zero.

Proposizione 7.1.6. *Se la v.a. X definita sullo spazio probabilizzato (Ω, \mathcal{F}, P) possiede momento di ordine r finito, allora la relativa f.c. φ_X è derivabile r volte e*

$$\varphi_X^{(s)}(0) = i^s E[X^s]$$

per $s = 1, \dots, r$.

Dimostrazione. Dalla definizione di derivata di funzione complessa si ha

$$\frac{d}{dt} e^{itx} = \frac{d}{dt} \cos(tx) + i \frac{d}{dt} \sin(tx) = ix e^{itx}.$$

Dunque, per ogni $t \in \mathbb{R}$, dal Teorema A.7 risulta che

$$\varphi_X^{(1)}(t) = \frac{d}{dt} E[e^{itX}] = E\left[\frac{d}{dt} e^{itX}\right] = E[ix e^{itX}].$$

La funzione $\varphi_X^{(1)}$ è definita, dal momento che, tenendo presente le ipotesi fatte, si ha

$$|E[ix e^{itX}]| \leq E[|ix e^{itX}|] = E[|i| \cdot |X| \cdot |e^{itX}|] = E[|X|] < \infty.$$

In particolare, risulta

$$\varphi_X^{(1)}(0) = iE[X].$$

In maniera analoga, si ha

$$\varphi_X^{(2)}(t) = \frac{d}{dt} E[iX e^{itX}] = E[i^2 X^2 e^{itX}],$$

che è definita, essendo per le ipotesi fatte

$$|E[i^2 X^2 e^{itX}]| \leq E[|i^2 X^2 e^{itX}|] = E[X^2 |e^{itX}|] = E[X^2] < \infty.$$

Dunque, si ha

$$\varphi_X^{(2)}(0) = i^2 E[X^2].$$

Iterando il procedimento r volte si ottiene la tesi. □

• **Esempio 7.1.8.** Si consideri la v.a. Z distribuita con legge Normale ridotta $\mathcal{N}(0, 1)$. La v.a. Z possiede momenti di ogni ordine, essendo $E[Z^r] = 0$ se r è dispari, mentre risulta

$$\begin{aligned} E[Z^r] &= \int_{-\infty}^{\infty} z^r \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz = \frac{2}{\sqrt{2\pi}} \int_0^{\infty} z^r e^{-\frac{1}{2}z^2} dz \\ &= \frac{2^{\frac{1}{2}r}}{\Gamma(\frac{1}{2})} \int_0^{\infty} y^{\frac{1}{2}(r+1)-1} e^{-y} dy = \frac{2^{\frac{1}{2}r} \Gamma(\frac{r+1}{2})}{\Gamma(\frac{1}{2})} \end{aligned}$$

se r è pari. Tenendo presente la Formula di Duplicazione per la funzione Gamma di Eulero, la precedente espressione può essere anche riscritta come

$$E[Z^r] = \frac{\Gamma(r)}{2^{\frac{1}{2}r-1} \Gamma(\frac{r}{2})} = \frac{r!}{2^{\frac{1}{2}r} (\frac{r}{2})!}.$$

In base alla Proposizione 7.1.5, la f.c. ammette dunque derivata di ogni ordine. In particolare, tenendo presente l'Esempio 7.1.4, si ha

$$\varphi_Z^{(1)}(t) = \frac{d}{dt} e^{-\frac{1}{2}t^2} = -t\varphi_Z(t),$$

da cui $\varphi_Z^{(1)}(0) = 0$, mentre

$$\varphi_Z^{(2)}(t) = \frac{d^2}{dt^2} e^{-\frac{1}{2}t^2} = (t^2 - 1)\varphi_Z(t),$$

da cui da cui $\varphi_Z^{(2)}(0) = -1$. Confermando i precedenti risultati, si ha dunque $E[Z] = i^{-1}\varphi_Z^{(1)}(0) = 0$, mentre $\text{Var}[Z] = E[Z^2] = i^{-2}\varphi_Z^{(2)}(0) = 1$. □

• **Esempio 7.1.9.** Si consideri la v.a. Z con legge di Cauchy ridotta analizzata nell'Esempio 4.1.4. Dal medesimo Esempio e tenendo presente la Proposizione 4.3.3, la v.a. Z non possiede momento di ogni ordine. In effetti, dal momento che la v.a. Z è simmetrica rispetto all'origine, la f.c. è data da

$$\varphi_Z(t) = E[\cos(tZ)] = \int_{-\infty}^{\infty} \cos(tz) \frac{1}{\pi(1+z^2)} dz = e^{-|t|},$$

che appunto non risulta derivabile in $t = 0$. Per maggiori dettagli sulla precedente espressione, si veda l'Esercizio 7.3.2. □

Proposizione 7.1.7. *Se la v.a. X definita sullo spazio probabilizzato (Ω, \mathcal{F}, P) con f.c. φ_X possiede momento di ordine r finito, allora per $t \rightarrow 0$ si ha*

$$\varphi_X(t) = \sum_{s=0}^r \frac{(it)^s}{s!} E[X^s] + o(|t|^r).$$

Dimostrazione. Si veda Billingsley (1995, p.341).

• **Esempio 7.1.10.** Si consideri la v.a. Z con legge Normale ridotta $\mathcal{N}(0, 1)$. Per $t \rightarrow 0$ si ha dunque

$$\varphi_Z(t) = 1 + itE[Z] + \frac{(it)^2}{2} E[Z^2] + o(t^2) = 1 - \frac{t^2}{2} + o(t^2). \quad \square$$

In alcune applicazioni della Teoria della Probabilità (ad esempio in molti problemi della statistica matematica) risulta piuttosto semplice ottenere la f.c. di alcune v.a. Il problema che sorge di conseguenza consiste nel risalire alla f.r. di queste v.a., se è nota la corrispondente f.c. I prossimi Teoremi consentono appunto di ottenere la f.r. di una v.a. (eventualmente la f.p. e la d.p. nel caso di v.a. discrete e di v.a. assolutamente continue) a partire dalla f.c.

Teorema 7.1.8. (Teorema di inversione di Lévy) *Si consideri la v.a. X definita sullo spazio probabilizzato (Ω, \mathcal{F}, P) con f.r. F_X e f.c. φ_X . Se $x, y \in \mathbb{R}$ sono punti di continuità di F_X tali che $x < y$, allora si ha*

$$F_X(y) - F_X(x) = \lim_{T \rightarrow \infty} \frac{1}{2\pi} \int_{-T}^T \frac{e^{itx} - e^{ity}}{it} \varphi_X(t) dt.$$

Dimostrazione. Si veda Billingsley (1995, p.346).

Teorema 7.1.9. *Si consideri la v.a. X definita sullo spazio probabilizzato (Ω, \mathcal{F}, P) con f.c. φ_X . Se $P(X = x) > 0$, allora si ha*

$$P(X = x) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T e^{-itx} \varphi_X(t) dt.$$

Dimostrazione. Si veda Billingsley (1995, p.354).

Teorema 7.1.10. *Si consideri la v.a. X definita sullo spazio probabilizzato (Ω, \mathcal{F}, P) con f.c. φ_X . Se $\int_{-\infty}^{\infty} |\varphi_X(t)| dt < \infty$, la v.a. X è assolutamente continua e ammette d.p.*

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \varphi_X(t) dt.$$

Dimostrazione. Si veda Billingsley (1995, p.347).

Si noti che dal Teorema 7.1.8 segue anche la proprietà di unicità della f.c., ovvero, date le v.a. X e Y , allora $X \stackrel{L}{=} Y$ se e solo se $\varphi_X = \varphi_Y$. La f.c. deve appunto la sua denominazione al fatto che “caratterizza” in modo univoco la corrispondente legge.

• **Esempio 7.1.11.** Si consideri la v.a. Z con legge di Cauchy ridotta (si veda l'Esempio 4.1.4). Tenendo presente l'Esempio 7.1.9, si ha

$$\int_{-\infty}^{\infty} |\varphi_X(t)| dt = \int_{-\infty}^{\infty} e^{-|t|} dt = 2 < \infty$$

e dal momento che la f.c. φ_X è simmetrica rispetto all'origine, si ottiene

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} e^{-|t|} dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \cos(tx) e^{-|t|} dt = \frac{1}{\pi} \int_0^{\infty} \cos(tx) e^{-t} dt = \frac{1}{\pi(1+x^2)},$$

come ovviamente doveva risultare. \square

7.2. Funzioni generatrici

Nel caso in cui si consideri una v.a. discreta X a valori su \mathbb{N} (o su un suo sottoinsieme), è conveniente considerare la cosiddetta funzione generatrice di probabilità piuttosto che la f.c. Formalmente, si ha la seguente definizione.

Definizione 7.2.1. Data la v.a. discreta X definita sullo spazio probabilizzato (Ω, \mathcal{F}, P) a valori su $S \subseteq \mathbb{N}$ e con f.p. p_X , si dice *funzione generatrice di probabilità* (f.g.) (o semplicemente *funzione generatrice*) della v.a. X la funzione $G_X : [0, 1] \rightarrow \mathbb{R}$ tale che

$$G_X(t) = E[t^X] = \sum_{x=0}^{\infty} t^x p_X(x),$$

con la convenzione che $0^0 = 1$. \square

Nel linguaggio della Teoria della Misura la f.g. corrisponde alla funzione generatrice di Laplace. Risulta immediato verificare che per ciascun punto di $[0, 1]$ la f.g. G_X coincide per definizione con la serie di potenze avente $(p_X(x))_{x \geq 0}$ come successione dei coefficienti. Inoltre, la serie di potenze ha raggio di convergenza almeno pari ad 1 dal momento che converge nel punto $t = 1$, essendo $G_X(1) = 1$. Inoltre, dal momento che $G_X(0) = p_X(0)$, la f.g. è definita per ogni valore dell'intervallo $[0, 1]$ ed è crescente in questo intervallo essendo $E[t^X] \leq E[s^X]$ per ogni $t < s$.

• **Esempio 7.2.1.** Se si considera la v.a. X distribuita con legge Binomiale $\mathcal{B}(n, p)$, tenendo presente l'espressione del binomio di Newton, si ha

$$G_X(t) = \sum_{x=0}^n t^x \binom{n}{x} p^x q^{n-x} = \sum_{x=0}^n \binom{n}{x} (pt)^x q^{n-x} = (q + pt)^n.$$

Evidentemente, si ha $G_X(0) = p_X(0) = q^n$ e $G_X(1) = (q + p)^n = 1$, mentre G_X è una funzione crescente. \square

• **Esempio 7.2.2.** Si consideri la v.a. X distribuita con legge di Poisson $\mathcal{P}(\lambda)$. Tenendo presente l'espressione della serie esponenziale, si ha

$$G_X(t) = \sum_{x=0}^{\infty} t^x e^{-\lambda} \frac{\lambda^x}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda t)^x}{x!} = e^{\lambda(t-1)}. \quad \square$$

• **Esempio 7.2.3.** Si consideri la v.a. X con legge Binomiale Negativa $\mathcal{BN}(k, p)$. Tenendo presente l'espressione della serie binomiale negativa, si ha

$$G_X(t) = \sum_{x=0}^{\infty} t^x \binom{k+x-1}{k-1} p^k q^x = p^k \sum_{x=0}^{\infty} \binom{k+x-1}{x} (qt)^x = \left(\frac{p}{1-qt} \right)^k. \quad \square$$

La seguente Proposizione mette in relazione la f.p. con la f.g. e giustifica in effetti la denominazione di G_X .

Proposizione 7.2.2. *Data la v.a. discreta X definita sullo spazio probabilizzato (Ω, \mathcal{F}, P) a valori su $S \subseteq \mathbb{N}$ con f.p. p_X e f.g. G_X , si ha*

$$p_X(x) = \frac{G_X^{(x)}(0)}{x!}$$

dove $x \in \mathbb{N}$.

Dimostrazione. Dal momento che una serie di potenze può essere derivata termine a termine per ogni valore nel raggio di convergenza, per $t \in [0, 1[$ si ha

$$G_X^{(1)}(t) = \sum_{n=1}^{\infty} n t^{n-1} p_X(n)$$

e

$$G_X^{(2)}(t) = \sum_{n=2}^{\infty} n(n-1) t^{n-2} p_X(n),$$

mentre in generale si ha

$$G_X^{(x)}(t) = \sum_{n=x}^{\infty} n(n-1)\dots(n-x+1) t^{n-x} p_X(n).$$

Dunque, si ottiene $G_X^{(1)}(0) = p_X(1)$ e $G_X^{(2)}(0) = 2p_X(2)$, mentre in generale risulta

$$G_X^{(x)}(0) = x! p_X(x).$$

Tenendo presente che $G_X^{(0)}(0) = p_X(0)$ si ha dunque la tesi. \square

Dalla precedente Proposizione segue anche la proprietà di unicità della f.g., ovvero, date le v.a. X e Y , allora si ha $X \stackrel{\mathcal{L}}{=} Y$ se e solo se $G_X = G_Y$. La seguente Proposizione fornisce la relazione fra la f.g. e i momenti di una v.a.

Proposizione 7.2.3. *Data la v.a. discreta X definita sullo spazio probabilizzato (Ω, \mathcal{F}, P) a valori su $S \subseteq \mathbb{N}$ con f.p. p_X e f.g. G_X , se la v.a. X possiede momento di ordine r finito, si ha*

$$G_X^{(s)}(1) = E[X(X-1)\dots(X-s+1)]$$

per $s = 1, \dots, r$.

Dimostrazione. Tenendo presente la Proposizione 7.2.2, $G_X^{(s)}$ è positiva e crescente in $[0, 1[$ e quindi ammette un limite sinistro per $t \rightarrow 1$ (che può non essere finito). Si ha

$$\lim_{t \rightarrow 1^-} G_X^{(1)}(t) = \sum_{x=0}^{\infty} x p_X(x) = E[X]$$

e

$$\lim_{t \rightarrow 1^-} G_X^{(2)}(t) = \sum_{x=0}^{\infty} x(x-1) p_X(x) = E[X(X-1)],$$

mentre in generale si ha

$$\lim_{t \rightarrow 1^-} G_X^{(s)}(t) = \sum_{x=0}^{\infty} x(x-1)\dots(x-s+1)p_X(x) = E[X(X-1)\dots(X-s+1)].$$

Dal momento che risulta $E[X(X-1)\dots(X-s+1)] \leq E[X^s] < \infty$, in quanto per ipotesi $E[X^s]$ è finito per $s = 1, \dots, r$, si ha quindi la tesi. \square

In particolare dalla precedente Proposizione si ha che

$$E[X] = G_X^{(1)}(1)$$

se $E[X]$ è finito, mentre

$$\text{Var}[X] = E[X(X-1)] + E[X] - E[X]^2 = G_X^{(2)}(1) + G_X^{(1)}(1) - G_X^{(1)}(1)^2,$$

se $E[X^2]$ è finito.

• **Esempio 7.2.4.** Se si considera la v.a. X distribuita con legge Binomiale $\mathcal{B}(n, p)$, si ha

$$G_X^{(1)}(t) = \frac{d}{dt} (q + pt)^n = np(q + pt)^{n-1}$$

e

$$G_X^{(2)}(t) = \frac{d^2}{dt^2} (q + pt)^n = n(n-1)p^2(q + pt)^{n-2}.$$

Dunque, risulta $E[X] = np$ e $E[X(X-1)] = n(n-1)p^2$, da cui è immediato verificare che $\text{Var}[X] = npq$, un risultato noto dalla Sezione 6.1. \square

• **Esempio 7.2.5.** Se si considera la v.a. X distribuita con legge di Poisson $\mathcal{P}(\lambda)$, si ha

$$G_X^{(1)}(t) = \frac{d}{dt} e^{\lambda(t-1)} = \lambda e^{\lambda(t-1)}$$

e

$$G_X^{(2)}(t) = \frac{d^2}{dt^2} e^{\lambda(t-1)} = \lambda^2 e^{\lambda(t-1)}.$$

Dunque, risulta $E[X] = \lambda$ e $E[X(X-1)] = \lambda^2$, da cui $\text{Var}[X] = \lambda$, un risultato verificato nella Sezione 6.2. \square

• **Esempio 7.2.6.** Se si considera la v.a. X distribuita con legge Binomiale Negativa $\mathcal{BN}(k, p)$, si ha

$$G_X^{(1)}(t) = \frac{d}{dt} \left(\frac{p}{1-qt} \right)^k = \frac{kq}{p} \left(\frac{p}{1-qt} \right)^{k+1}$$

e

$$G_X^{(2)}(t) = \frac{d^2}{dt^2} \left(\frac{p}{1-qt} \right)^k = \frac{k(k+1)q^2}{p^2} \left(\frac{p}{1-qt} \right)^{k+2}.$$

Dunque, risulta $E[X] = \frac{kq}{p}$ e $E[X(X-1)] = \frac{k(k+1)q^2}{p^2}$, da cui è immediato verificare che $\text{Var}[X] = \frac{kq}{p^2}$, un risultato ottenuto in precedenza nella Sezione 6.3. \square

• **Esempio 7.2.7.** (Legge Zeta o di Zipf) Si consideri la v.a. discreta X con f.p. data da

$$p_X(x) = \frac{1}{\zeta(s)} x^{-s} \mathbf{1}_{\{1,2,\dots\}}(x),$$

dove $s \in]1, \infty[$, mentre $\zeta(s) = \sum_{x=1}^{\infty} x^{-s}$ rappresenta la funzione Zeta di Riemann. La legge associata a questa v.a. è detta di Zipf in onore del linguista e filologo statunitense George Kingsley Zipf (1902-1950), che si occupò delle applicazioni probabilistiche alle analisi dei testi. Ovviamente, la funzione Zeta prende nome dal matematico tedesco Georg Friedrich Bernhard Riemann (1826-1866). Dal momento che si ha $\zeta(2) = \frac{\pi^2}{6}$, per $s = 2$ la f.p. della v.a. X risulta

$$p_X(x) = \frac{6}{\pi^2} x^{-2} \mathbf{1}_{\{1,2,\dots\}}(x).$$

Per questo valore di s , la v.a. X non possiede momenti finiti di alcun ordine, essendo

$$E[X] = \frac{6}{\pi^2} \sum_{x=1}^{\infty} x^{-1} = \infty.$$

In effetti, la f.g. risulta

$$G_X(t) = \frac{6}{\pi^2} \sum_{x=1}^{\infty} t^x x^{-2},$$

da cui, tenendo presente l'espressione della serie logaritmica, si ha

$$G_X^{(1)}(t) = \frac{6}{\pi^2 t} \sum_{x=1}^{\infty} \frac{t^x}{x} = -\frac{6}{\pi^2} \frac{\log(1-t)}{t},$$

e quindi

$$\lim_{t \rightarrow 1^-} G_X^{(1)}(t) = \infty. \quad \square$$

7.3. Funzioni caratteristiche multivariate

Viene considerata di seguito l'estensione multivariata della f.c. quando si dispone di un v.v.a. Formalmente si ha la seguente definizione.

Definizione 7.3.1. Dato il v.v.a. $X = (X_1, \dots, X_n)^T$ definito sullo spazio probabilizzato (Ω, \mathcal{F}, P) , si dice *funzione caratteristica multivariata* (f.c.m.) del v.v.a. X la funzione $\varphi_X : \mathbb{R}^n \rightarrow \mathbb{C}$

$$\varphi_X(t) = E[\cos(t^T X)] + i E[\sin(t^T X)] = E[e^{it^T X}],$$

dove $t = (t_1, \dots, t_n)^T \in \mathbb{R}^n$. □

Dalla Definizione 7.3.1 è ovvio che la f.c.m. possiede la seguente espressione alternativa

$$\varphi_X(t) = \varphi_X(t_1, \dots, t_n) = E\left[e^{\sum_{k=1}^n it_k X_k}\right] = E\left[\prod_{k=1}^n e^{it_k X_k}\right].$$

Risulta importante osservare che, se nella f.c.m. relativa al v.v.a. X si pone $t_l = 0$ per ogni $l \neq k = 1, \dots, n$, si ha

$$\varphi_X(0, \dots, t_k, \dots, 0) = E[e^{it_k X_k}] = \varphi_{X_k}(t_k),$$

ovvero si ottiene la f.c. della componente marginale X_k . In modo analogo, ponendo $t_l = 0$ per ogni $l \neq k, j = 1, \dots, n$, si ha

$$\varphi_X(0, \dots, t_k, \dots, t_j, \dots, 0) = E[e^{it_k X_k + it_j X_j}] = \varphi_{(X_k, X_j)}(t_k, t_j),$$

che costituisce la f.c.m. del v.v.a. $(X_k, X_j)^T$. In generale, azzerando opportunamente alcune componenti del vettore t , si ottengono le f.c.m. di tutte le possibili scelte di componenti marginali del v.v.a. X .

• **Esempio 7.3.1.** Si consideri il v.v.a. X distribuito con legge Normale Multivariata $\mathcal{N}_n(\mu, \Sigma)$. Tenendo presente quanto visto nella Sezione 6.11, in base al cambio di variabile $y = \Sigma^{-\frac{1}{2}}(x - \mu)$ con $x = \mu + \Sigma^{\frac{1}{2}}y$ e tenendo presente le regole di integrazione di una funzione complessa, la f.c.m. relativa al v.v.a. X è data da

$$\begin{aligned} \varphi_X(t) &= \int_{\mathbb{R}^n} e^{it^T x} \det(2\pi\Sigma)^{-\frac{1}{2}} e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)} dx = (2\pi)^{-\frac{1}{2}n} \int_{\mathbb{R}^n} e^{it^T \mu + it^T \Sigma^{\frac{1}{2}} y - \frac{1}{2} y^T y} dy \\ &= (2\pi)^{-\frac{1}{2}n} e^{i\mu^T t - \frac{1}{2} t^T \Sigma t} \int_{\mathbb{R}^n} e^{-\frac{1}{2}(y - i \Sigma^{\frac{1}{2}} t)^T (y - i \Sigma^{\frac{1}{2}} t)} dy. \end{aligned}$$

Effettuando inoltre il cambio di variabile $z = y - i \Sigma^{\frac{1}{2}} t$, con Jacobiano pari a 1, si ottiene

$$\varphi_X(t) = (2\pi)^{-\frac{1}{2}n} e^{i\mu^T t - \frac{1}{2} t^T \Sigma t} \int_{\mathbb{R}^n} e^{-\frac{1}{2} z^T z} dz = e^{i\mu^T t - \frac{1}{2} t^T \Sigma t},$$

dal momento che

$$(2\pi)^{-\frac{1}{2}n} \int_{\mathbb{R}^n} e^{-\frac{1}{2} z^T z} dz = \left(\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} u^2} du \right)^n = 1.$$

Si ha inoltre

$$\varphi_{X_k}(t_k) = \varphi_X(0, \dots, t_k, \dots, 0) = e^{i\mu_{X_k} t_k - \frac{1}{2} \sigma_{X_k}^2 t_k^2},$$

ovvero ogni componente marginale X_k si distribuisce con legge Normale $\mathcal{N}(\mu_{X_k}, \sigma_{X_k}^2)$. In modo simile, si dimostra che ogni scelta di componenti marginali del v.v.a. X si distribuisce a sua volta con legge Normale Multivariata. \square

La f.c.m. di un v.v.a. possiede proprietà analoghe rispetto alla f.c. di una v.a. Più esattamente, si hanno le seguenti Proposizioni.

Proposizione 7.3.2. Se $X = (X_1, \dots, X_n)^T$ è un v.v.a. definito sullo spazio probabilizzato (Ω, \mathcal{F}, P) con f.c.m. φ_X , allora si ha $|\varphi_X(t)| \leq 1$ per ogni $t \in \mathbb{R}^n$ e in particolare $\varphi_X(0) = 1$, dove $0 = (0, \dots, 0)^T$ rappresenta l'origine di \mathbb{R}^n , mentre $\varphi_X(t)$ è uniformemente continua su \mathbb{R}^n .

Dimostrazione. Risulta simile a quella delle Proposizioni 7.1.3. e 7.1.4. \square

Proposizione 7.3.3. Se $X = (X_1, \dots, X_n)^T$ è un v.v.a. definito sullo spazio probabilizzato (Ω, \mathcal{F}, P) con f.c.m. $\varphi_X(t_1, \dots, t_n)$, la f.c. della v.a. trasformata $Y = a + \sum_{k=1}^n b_k X_k$, dove $a \in \mathbb{R}$ e $b = (b_1, \dots, b_n)^T \in \mathbb{R}^n$, risulta

$$\varphi_Y(t) = e^{iat} \varphi_X(tb_1, \dots, tb_n) = e^{iat} \varphi_X(tb),$$

dove $t \in \mathbb{R}$.

Dimostrazione. Si ha

$$\varphi_Y(t) = E[e^{itY}] = E\left[e^{iat + \sum_{k=1}^n ib_k t X_k}\right] = e^{iat} E\left[e^{\sum_{k=1}^n ib_k t X_k}\right] = e^{iat} \varphi_X(b_1 t, \dots, b_n t) = e^{iat} \varphi_X(t b),$$

che è quanto si voleva dimostrare. \square

Come conseguenza della precedente Proposizione si ottiene che la f.c. della v.a. $Y = \sum_{k=1}^n X_k$ è data da

$$\varphi_Y(t) = \varphi_X(t, \dots, t),$$

un risultato che riveste notevole importanza pratica nel determinare la legge della v.a. Y .

• **Esempio 7.3.2.** Si consideri il v.v.a. X distribuito con legge Normale Multivariata $\mathcal{N}_n(\mu, \Sigma)$. Tenendo presente l'Esempio 7.3.1, per la Proposizione 7.3.3 la f.c. della v.a. trasformata $Y = a + \sum_{k=1}^n b_k X_k$ è data da

$$\varphi_Y(t) = e^{iat} e^{i(tb)^T \mu - \frac{1}{2}(tb)^T \Sigma (tb)} = e^{i(a+b^T \mu)t - \frac{1}{2}(b^T \Sigma b)t^2},$$

che risulta essere la f.c. di una v.a. distribuita con legge Normale $\mathcal{N}(a + b^T \mu, b^T \Sigma b)$. Dunque, ogni combinazione lineare (e quindi anche la somma) delle componenti di un v.v.a. con legge Normale Multivariata si distribuisce a sua volta con legge Normale. \square

Proposizione 7.3.4. Dato il v.v.a. $X = (X_1, \dots, X_n)^T$ definito sullo spazio probabilizzato (Ω, \mathcal{F}, P) con f.c.m. $\varphi_X(t_1, \dots, t_n)$, se $E[X_1^{r_1} \dots X_n^{r_n}]$ esiste finito, φ_X è derivabile r_1 volte rispetto a t_1, \dots, r_n volte rispetto a t_n e

$$\left. \frac{\partial^s \varphi_X(t_1, \dots, t_n)}{\partial t_1^{s_1} \dots \partial t_n^{s_n}} \right|_{t_1, \dots, t_n=0} = i^s E[X_1^{s_1} \dots X_n^{s_n}],$$

dove $s_k = 1, \dots, r_k, k = 1, \dots, n$, mentre $s = \sum_{k=1}^n s_k$.

Dimostrazione. Risulta analoga a quella della Proposizione 7.1.6. \square

Teorema 7.3.5. (Teorema di inversione multivariata) Dato il v.v.a. $X = (X_1, \dots, X_n)^T$ definito sullo spazio probabilizzato (Ω, \mathcal{F}, P) con f.r. F_X e f.c.m. φ_X , se $B =]x_1, y_1] \times \dots \times]x_n, y_n]$, con $x_k < y_k, k = 1, \dots, n$, è un insieme la cui frontiera è un insieme di continuità di F_X , allora

$$P(X \in B) = \lim_{T \rightarrow \infty} \frac{1}{(2\pi)^n} \int_{[-T, T]^n} \left(\prod_{k=1}^n \frac{e^{it_k x_k} - e^{it_k y_k}}{it} \right) \varphi_X(t) dt.$$

Dimostrazione. Si veda Billingsley (1995, p.382). \square

La seguente Proposizione fornisce un risultato di particolare rilevanza nel caso in cui le componenti del v.v.a. X sono indipendenti.

Proposizione 7.3.6. Dato il v.v.a. $X = (X_1, \dots, X_n)^T$ definito sullo spazio probabilizzato (Ω, \mathcal{F}, P) con f.c.m. $\varphi_X(t_1, \dots, t_n)$, allora si ha

$$\varphi_X(t_1, \dots, t_n) = \prod_{k=1}^n \varphi_{X_k}(t_k),$$

se e solo se X_1, \dots, X_n sono indipendenti.

Dimostrazione. Se X_1, \dots, X_n sono v.a. indipendenti, allora dalla Proposizione 4.2.5 si ottiene

$$\varphi_X(t) = \mathbb{E}\left[\prod_{k=1}^n e^{it_k X_k}\right] = \prod_{k=1}^n \mathbb{E}[e^{it_k X_k}] = \prod_{k=1}^n \varphi_{X_k}(t_k).$$

Inversamente, se $\varphi_X(t) = \prod_{k=1}^n \varphi_{X_k}(t_k)$ e ponendo $B =]x_1, y_1] \times \cdots \times]x_n, y_n]$, per il Teorema 7.3.5 si ottiene

$$P(X \in B) = \lim_{T \rightarrow \infty} \frac{1}{(2\pi)^n} \int_{[-T, T]^n} \left(\prod_{k=1}^n \frac{e^{it_k x_k} - e^{it_k y_k}}{it} \varphi_{X_k}(t_k) \right) dt_1 \dots dt_n$$

e, tenendo presente i Teoremi A.7 e A.11, si ha

$$P(X \in B) = \prod_{k=1}^n \lim_{T \rightarrow \infty} \frac{1}{2\pi} \int_{[-T, T]} \frac{e^{it x_k} - e^{it y_k}}{it} \varphi_{X_k}(t_k) dt_k = \prod_{k=1}^n P(X_k \in]x_k, y_k]),$$

che dalla Definizione 3.6.1 implica l'indipendenza delle v.a. X_1, \dots, X_n . □

Il Teorema 7.3.5 implica la proprietà di unicità della f.c.m., ovvero, dati i v.v.a. X e Y , allora si ha $X \stackrel{\mathcal{L}}{=} Y$ se e solo se $\varphi_X = \varphi_Y$. Inoltre, tenendo presente la Proposizione 7.3.3, se X_1, \dots, X_n sono v.a. indipendenti, la f.c. della v.a. trasformata $Y = a + \sum_{k=1}^n b_k X_k$, dove $a \in \mathbb{R}$ e $b = (b_1, \dots, b_n)^T \in \mathbb{R}^n$, risulta

$$\varphi_Y(t) = e^{iat} \varphi_X(b_1 t, \dots, b_n t) = e^{iat} \prod_{k=1}^n \varphi_{X_k}(b_k t).$$

In particolare, se $Y = \sum_{k=1}^n X_k$ è la somma di n v.a. indipendenti, la relativa f.c. risulta

$$\varphi_Y(t) = \prod_{k=1}^n \varphi_{X_k}(t).$$

Dunque, quando le v.a. X_1, \dots, X_n si distribuiscono con la stessa legge che coincide con quella della v.a. X , ovvero quando $\varphi_{X_k}(t) = \varphi_X(t)$ per ogni $k = 1, \dots, n$, si ottiene

$$\varphi_Y(t) = \varphi_X(t)^n.$$

In generale, si può dimostrare in modo semplice ma laborioso che se X_1, \dots, X_n sono v.v.a. indipendenti con k componenti marginali, la f.c.m. del v.v.a. trasformato $Y = \sum_{j=1}^n X_j$ risulta

$$\varphi_Y(t_1, \dots, t_k) = \prod_{j=1}^n \varphi_{X_j}(t_1, \dots, t_k).$$

• **Esempio 7.3.3.** Si consideri la v.a. Z distribuita con legge Esponenziale ridotta $\mathcal{G}(0, 1, 1)$. Dunque, dall'Esempio 7.1.5 la f.c. della v.a. Z è data da

$$\varphi_Z(t) = (1 - it)^{-1}.$$

Di conseguenza, se Z_1, \dots, Z_k sono v.a. indipendenti con la stessa legge della v.a. Z , la v.a. $Y = \sum_{j=1}^k X_j$ possiede f.c. data da

$$\varphi_Y(t) = ((1 - it)^{-1})^k = (1 - it)^{-k}.$$

Dunque, tenendo ancora presente l'Esempio 7.1.5, la v.a. Y si distribuisce con legge Gamma ridotta $\mathcal{G}(0, 1, k)$. □

• **Esempio 7.3.4.** Si consideri la v.a. X distribuita con legge Chi-quadrato χ_1^2 . Dunque, dall'Esempio 7.1.7 la f.c. di X è data da

$$\varphi_X(t) = (1 - 2it)^{-\frac{1}{2}}.$$

Di conseguenza, se X_1, \dots, X_n sono v.a. indipendenti con la stessa legge di X , la v.a. $Y = \sum_{k=1}^n X_k$ possiede f.c. data da

$$\varphi_Y(t) = ((1 - 2it)^{-\frac{1}{2}})^n = (1 - 2it)^{-\frac{1}{2}n}.$$

Dunque, tenendo presente l'Esempio 7.1.7, la v.a. Y si distribuisce con legge Chi-quadrato χ_n^2 . \square

• **Esempio 7.3.5.** Si consideri le v.a. indipendenti X_1, \dots, X_n , rispettivamente distribuite con legge Normale $\mathcal{N}(\mu_{X_k}, \sigma_{X_k}^2)$ per $k = 1, \dots, n$. Tenendo presente l'Esempio 7.1.6, la v.a. $Y = \sum_{k=1}^n X_k$ possiede f.c. data da

$$\varphi_Y(t) = \prod_{k=1}^n e^{i\mu_{X_k}t - \frac{1}{2}\sigma_{X_k}^2 t^2} = e^{i\mu_Y t - \frac{1}{2}\sigma_Y^2 t^2},$$

dove $\mu_Y = \sum_{k=1}^n \mu_{X_k}$ e $\sigma_Y^2 = \sum_{k=1}^n \sigma_{X_k}^2$. Dunque, la v.a. Y si distribuisce con legge Normale $\mathcal{N}(\mu_Y, \sigma_Y^2)$, un risultato che conferma un caso particolare di quanto ottenuto nell'Esempio 7.3.2. \square

• **Esempio 7.3.6.** Si consideri il v.v.a. X con legge Normale Multivariata $\mathcal{N}_k(\mu, \Sigma)$. Se X_1, \dots, X_n sono v.v.a. indipendenti con la stessa legge del v.v.a. X , sulla base dei risultati ottenuti nell'Esempio 7.3.1, il v.v.a. $Y = \sum_{j=1}^n X_j$ possiede f.c.m. data da

$$\varphi_Y(t) = (e^{i\mu^T t - \frac{1}{2}t^T \Sigma t})^n = e^{i(n\mu)^T t - \frac{1}{2}t^T (n\Sigma)t}.$$

Dunque, tenendo ancora presente l'Esempio 7.3.1, il v.v.a. Y si distribuisce con legge Normale Multivariata $\mathcal{N}_k(n\mu, n\Sigma)$. \square

7.4. Funzioni generatrici multivariate

Quando si dispone di un v.v.a. discreto con n componenti marginali e a valori su \mathbb{N}^n (o su un suo sottoinsieme) si può estendere il concetto di funzione generatrice. Più esattamente, si ha la seguente definizione.

Definizione 7.4.1. Dato il v.v.a. discreto $X = (X_1, \dots, X_n)^T$ definito sullo spazio probabilizzato (Ω, \mathcal{F}, P) a valori su $S \subseteq \mathbb{N}^n$ e con f.p.c. p_X , si dice *funzione generatrice multivariata di probabilità* (f.g.m.) (o semplicemente *funzione generatrice multivariata*) del v.v.a. X la funzione $G_X : [0, 1]^n \rightarrow \mathbb{R}$ tale che

$$G_X(t_1, \dots, t_n) = E\left[\prod_{k=1}^n t_k^{X_k}\right] = \sum_{(x_1, \dots, x_n) \in S} \left(\prod_{k=1}^n t_k^{x_k}\right) p_X(x_1, \dots, x_n),$$

con la convenzione che $0^0 = 1$. \square

Se nella f.g.m. relativa al v.v.a. X si pone $t_l = 1$ per ogni $l \neq k = 1, \dots, n$, si ottiene

$$G_X(1, \dots, t_k, \dots, 1) = E[t_k^{X_k}] = G_{X_k}(t_k),$$

che costituisce la f.g. della componente marginale X_k . In modo analogo, ponendo $t_l = 1$ per ogni $l \neq k, j = 1, \dots, n$, si ottiene

$$G_X(1, \dots, t_k, \dots, t_j, \dots, 1) = E[t_k^{X_k} t_j^{X_j}] = G_{(X_k, X_j)}(t_k, t_j),$$

che costituisce la f.g.m. del v.v.a. (X_k, X_j) . In generale, ponendo pari all'unità in modo opportuno alcune componenti del vettore (t_1, \dots, t_n) , si ottengono le f.g.m. di tutte le possibili scelte di componenti marginali del v.v.a. X .

• **Esempio 7.4.1.** Si consideri il v.v.a. X con legge Multinomiale $\mathcal{M}(n, p)$. Dunque, per il Teorema Multinomiale si ha

$$\begin{aligned} G_X(t_1, \dots, t_k) &= \sum_{(x_1, \dots, x_k) \in S} \left(\prod_{j=1}^k t_j^{x_j} \right) \binom{n}{x_1 \dots x_k} \prod_{j=1}^k p_j^{x_j} \\ &= \sum_{(x_1, \dots, x_k) \in S} \binom{n}{x_1 \dots x_k} \prod_{j=1}^k (p_j t_j)^{x_j} = \left(\sum_{j=1}^k p_j t_j \right)^n, \end{aligned}$$

dove $S \subset \mathbb{N}^k$ è definito nella Sezione 6.5. Inoltre, risulta

$$G_{X_j}(t_j) = G_X(1, \dots, t_j, \dots, 1) = (q_j + p_j t_j)^n,$$

dove $q_j = 1 - p_j = \sum_{l \neq j} p_l$, ovvero ogni componente marginale X_j si distribuisce con legge Binomiale $\mathcal{B}(n, p_j)$. In modo analogo, si dimostra che ogni scelta di componenti marginali del v.v.a. X si distribuisce ancora con legge Multinomiale. \square

La f.g.m. di un v.v.a. possiede proprietà analoghe alla f.g. di una v.a. In particolare, la f.g.m. è definita per ogni $t \in [0, 1]^n$ e risulta $G_X(1, \dots, 1) = 1$, mentre $G_X(0, \dots, 0) = p_X(0, \dots, 0)$. Si ha inoltre la seguente Proposizione.

Proposizione 7.4.2. Dato il v.v.a. discreto $X = (X_1, \dots, X_n)^T$ definito sullo spazio probabilitizzato (Ω, \mathcal{F}, P) a valori su $S \subseteq \mathbb{N}^n$ con f.p.c. p_X e f.g.m. G_X , si ha

$$p_X(x_1, \dots, x_n) = \frac{1}{x_1! \dots x_n!} \left. \frac{\partial^{x_1 + \dots + x_n} G_X(t_1, \dots, t_n)}{\partial t_1^{x_1} \dots \partial t_n^{x_n}} \right|_{t_1, \dots, t_n=0}$$

dove $(x_1, \dots, x_n) \in \mathbb{N}^n$.

Dimostrazione. Risulta simile a quella della Proposizione 7.2.2. \square

Dalla precedente Proposizione segue anche la proprietà di unicità della f.g.m., ovvero, dati i v.v.a. X e Y , allora si ha $X \stackrel{\mathcal{L}}{=} Y$ se e solo se $G_X = G_Y$.

Proposizione 7.4.3. Dato il v.v.a. discreto $X = (X_1, \dots, X_n)^T$ definito sullo spazio probabilitizzato (Ω, \mathcal{F}, P) a valori su $S \subseteq \mathbb{N}^n$ con f.g.m. G_X , se $E[X_1^{r_1} \dots X_n^{r_n}]$ esiste finito, G_X è derivabile r_1 volte rispetto a t_1, \dots, r_n volte rispetto a t_n e

$$\left. \frac{\partial^s G_X(t_1, \dots, t_n)}{\partial t_1^{s_1} \dots \partial t_n^{s_n}} \right|_{t_1, \dots, t_n=1} = E \left[\prod_{k=1}^n X_k (X_k - 1) \dots (X_k - s_k + 1) \right],$$

dove $s_k = 1, \dots, r_k$ e $k = 1, \dots, n$, mentre $s = \sum_{k=1}^n s_k$.

Dimostrazione. Risulta simile a quella della Proposizione 7.2.3. \square

• **Esempio 7.4.2.** Si consideri il v.v.a. X distribuito con legge Multinomiale $\mathcal{M}(n, p)$. Dunque, dalla Proposizione 7.4.3 si ha

$$E[X_j] = \frac{\partial}{\partial t_j} \left(\sum_{l=1}^k p_l t_l \right)^n \Big|_{t_1, \dots, t_k=1} = np_j,$$

per cui il vettore delle medie del v.v.a. X è dato da $\mu = np$. Inoltre, si ha

$$E[X_j(X_j - 1)] = \frac{\partial^2}{\partial t_j^2} \left(\sum_{l=1}^k p_l t_l \right)^n \Big|_{t_1, \dots, t_k=1} = n(n-1)p_j^2$$

e

$$E[X_i X_j] = \frac{\partial^2}{\partial t_i \partial t_j} \left(\sum_{l=1}^k p_l t_l \right)^n \Big|_{t_1, \dots, t_k=1} = n(n-1)p_i p_j.$$

Dunque, risulta $\text{Var}[X_j] = np_j(1-p_j)$ e $\text{Cov}[X_i, X_j] = -np_i p_j$, e quindi la matrice di varianza-covarianza è data da $\Sigma = n(\text{diag}(p) - pp^T)$. \square

Proposizione 7.4.4. Dato il v.v.a. discreto $X = (X_1, \dots, X_n)^T$ definito sullo spazio probabilitizzato (Ω, \mathcal{F}, P) a valori su $S \subseteq \mathbb{N}^n$ con f.g.m. G_X , la f.g.m. della v.a. trasformata $Y = \sum_{k=1}^n X_k$ risulta

$$G_Y(t) = G_X(t, \dots, t),$$

dove $t \in \mathbb{R}$.

Dimostrazione. Si ha

$$G_Y(t) = E[t^{\sum_{k=1}^n X_k}] = E\left[\prod_{k=1}^n t^{X_k}\right] = G_X(t, \dots, t),$$

che è quanto si voleva dimostrare. \square

La seguente Proposizione fornisce l'espressione della f.g.m. nel caso in cui le componenti del v.v.a. X risultano indipendenti.

Proposizione 7.4.5. Dato il v.v.a. discreto $X = (X_1, \dots, X_n)^T$ definito sullo spazio probabilitizzato (Ω, \mathcal{F}, P) a valori su $S \subseteq \mathbb{N}^n$ con f.g.m. G_X , allora si ha

$$G_X(t_1, \dots, t_n) = \prod_{k=1}^n G_{X_k}(t_k),$$

se e solo se X_1, \dots, X_n sono indipendenti.

Dimostrazione. Se X_1, \dots, X_n sono v.a. indipendenti, allora dalla Proposizione 4.2.5 si ottiene

$$G_X(t_1, \dots, t_n) = E\left[\prod_{k=1}^n t_k^{X_k}\right] = \prod_{k=1}^n E[t_k^{X_k}] = \prod_{k=1}^n G_{X_k}(t_k).$$

Inversamente, se $G_X(t_1, \dots, t_n) = \prod_{k=1}^n G_{X_k}(t_k)$, dalla Proposizione 7.4.2 si ha

$$\begin{aligned}
 p_X(x_1, \dots, x_n) &= \frac{1}{x_1! \dots x_n!} \left. \frac{\partial^{x_1 + \dots + x_n} \prod_{k=1}^n G_{X_k}(t_k)}{\partial t_1^{x_1} \dots \partial t_n^{x_n}} \right|_{t_1, \dots, t_n=0} \\
 &= \prod_{k=1}^n \frac{1}{x_k!} \left. \frac{\partial^{x_k} G_{X_k}(t_k)}{\partial t_k^{x_k}} \right|_{t_1, \dots, t_n=0} = \prod_{k=1}^n p_{X_k}(x_k),
 \end{aligned}$$

e il Teorema 3.6.1 implica l'indipendenza delle v.a. X_1, \dots, X_n . □

Tenendo presente la Proposizione 7.4.4, se X_1, \dots, X_n sono v.a. indipendenti, la f.g. della v.a. trasformata $Y = \sum_{k=1}^n X_k$ risulta

$$G_Y(t) = \prod_{k=1}^n G_{X_k}(t).$$

Dunque, quando le v.a. X_k possiedono la stessa legge che coincide con quella di una data v.a. X , ovvero quando $G_{X_k}(t) = G_X(t)$ per ogni $k = 1, \dots, n$, si ottiene

$$G_Y(t) = G_X(t)^n.$$

In generale, si può dimostrare in modo semplice ma laborioso che se X_1, \dots, X_n sono v.v.a. indipendenti con k componenti marginali, la f.g.m. del v.v.a. trasformato $Y = \sum_{j=1}^n X_j$ risulta

$$G_Y(t_1, \dots, t_k) = \prod_{j=1}^n G_{X_j}(t_1, \dots, t_k).$$

• **Esempio 7.4.3.** Si consideri la v.a. X distribuita con legge di Bernoulli $\mathcal{B}(1, p)$. Dunque, dall'Esempio 7.2.1 la f.g. di X è data da

$$G_X(t) = q + pt.$$

Di conseguenza, se X_1, \dots, X_n sono v.a. indipendenti con la stessa legge della v.a. X , la v.a. $Y = \sum_{k=1}^n X_k$ possiede f.g. data da

$$G_Y(t) = (q + pt)^n.$$

Dunque, tenendo ancora presente l'Esempio 7.2.1, la v.a. Y si distribuisce con legge Binomiale $\mathcal{B}(n, p)$. □

• **Esempio 7.4.4.** Si consideri la v.a. X distribuita con legge Binomiale Negativa $\mathcal{BN}(1, p)$, ovvero con legge Geometrica. Dunque, dall'Esempio 7.2.3 la f.g. della v.a. X è data da

$$G_X(t) = \frac{p}{1 - qt}.$$

Di conseguenza, se X_1, \dots, X_k sono v.a. indipendenti con la stessa legge della v.a. X , la v.a. $Y = \sum_{k=1}^k X_k$ possiede f.g. data da

$$G_Y(t) = \left(\frac{p}{1 - qt} \right)^k.$$

Dunque, tenendo ancora presente l'Esempio 7.2.3, la v.a. Y si distribuisce con legge Binomiale Negativa $\mathcal{BN}(k, p)$. □

• **Esempio 7.4.5.** Si consideri il v.v.a. X distribuito con legge Multinomiale $\mathcal{M}(1, p)$. Dunque, dall'Esempio 7.4.1 la f.g.m. del v.v.a. X è data da

$$G_X(t_1, \dots, t_k) = \sum_{j=1}^k p_j t_j.$$

Di conseguenza, se X_1, \dots, X_n sono v.v.a. indipendenti con la stessa legge del v.v.a. X , il v.v.a. $Y = \sum_{j=1}^k X_j$ possiede f.g.m. data da

$$G_Y(t_1, \dots, t_k) = \left(\sum_{j=1}^k p_j t_j \right)^n.$$

Dunque, tenendo ancora presente l'Esempio 7.4.1, il v.v.a. Y si distribuisce con legge Multinomiale $\mathcal{M}(n, p)$. \square

7.5. Leggi additive e infinitamente divisibili

In questa sezione vengono considerate due importanti famiglie di leggi che sono caratterizzate da particolari comportamenti rispetto alla somma di v.a. indipendenti. In particolare, si ha la seguente definizione.

Definizione 7.5.1. Sia data la v.a. X definita sullo spazio probabilizzato (Ω, \mathcal{F}, P) con f.r. $F_{X, \theta}$ che dipende da un vettore di parametri $\theta \in \Theta \subseteq \mathbb{R}^m$. La v.a. X e la relativa legge si dicono *additive rispetto a θ* se, date le v.a. indipendenti X_1 e X_2 con rispettive f.r. F_{X_1, θ_1} e F_{X_2, θ_2} , la f.r. della v.a. $(X_1 + X_2)$ risulta essere $F_{X_1+X_2, \theta_1+\theta_2}$ per ogni $\theta_1, \theta_2 \in \Theta$. \square

Tenendo presente la proprietà di unicità della f.c., la Definizione 7.5.1 può essere anche data in termini della f.c. $\varphi_{X, \theta}$ della v.a. X . Evidentemente, se la v.a. X e la rispettiva legge risultano additive rispetto a θ , si verifica che

$$\varphi_{X_1, \theta_1}(t) \varphi_{X_2, \theta_2}(t) = \varphi_{X_1+X_2, \theta_1+\theta_2}(t),$$

per ogni $\theta_1, \theta_2 \in \Theta$. Ovviamente, nel caso di una v.a. discreta X a valori su $S \subseteq \mathbb{N}$, una analoga definizione può essere data in termini della corrispondente f.g. $G_{X, \theta}$, ovvero si ha

$$G_{X_1, \theta_1}(t) G_{X_2, \theta_2}(t) = G_{X_1+X_2, \theta_1+\theta_2}(t),$$

per ogni $\theta_1, \theta_2 \in \Theta$. Si osservi che il vettore di parametri θ può non comprendere l'intero insieme di parametri che caratterizzano la legge in questione, ma soltanto il sottoinsieme di parametri rispetto ai quali si vuole verificare l'additività. In questo caso, i rimanenti parametri devono invece essere costanti.

• **Esempio 7.5.1.** Si consideri la v.a. X distribuita con legge Normale $\mathcal{N}(\mu, \sigma^2)$. In questo caso, si pone $\theta = (\mu, \sigma^2)^T$ con $\Theta = \mathbb{R} \times]0, \infty[$. La legge risulta additiva rispetto al vettore di parametri θ , in quanto dall'Esempio 7.3.5 si ha

$$\varphi_{X_1, \theta_1}(t) \varphi_{X_2, \theta_2}(t) = e^{i\mu_{X_1}t - \frac{1}{2}\sigma_{X_1}^2 t^2} e^{i\mu_{X_2}t - \frac{1}{2}\sigma_{X_2}^2 t^2} = e^{i(\mu_{X_1} + \mu_{X_2})t - \frac{1}{2}(\sigma_{X_1}^2 + \sigma_{X_2}^2)t^2} = \varphi_{X_1+X_2, \theta_1+\theta_2}(t),$$

ovvero, in altri termini, la legge della somma di due v.a. indipendenti, rispettivamente con legge Normale $\mathcal{N}(\mu_{X_1}, \sigma_{X_1}^2)$ e $\mathcal{N}(\mu_{X_2}, \sigma_{X_2}^2)$, è Normale $\mathcal{N}(\mu_{X_1} + \mu_{X_2}, \sigma_{X_1}^2 + \sigma_{X_2}^2)$. \square

• **Esempio 7.5.2.** Si consideri la v.a. X distribuita con legge Gamma $\mathcal{G}(0, b, k)$. In questo caso, si pone $\theta = k$ con $\Theta =]0, \infty[$, ovvero si vuole verificare l'additività rispetto al parametro k e non al parametro b . La legge è additiva rispetto a θ , in quanto tenendo presente l'Esempio 7.1.7 si ha

$$\varphi_{X_1, \theta_1}(t) \varphi_{X_2, \theta_2}(t) = (1 - ibt)^{-k_1} (1 - ibt)^{-k_2} = (1 - ibt)^{-(k_1+k_2)} = \varphi_{X_1+X_2, \theta_1+\theta_2}(t).$$

Dunque, la legge della somma di due v.a. indipendenti, rispettivamente con legge Gamma $\mathcal{G}(0, b, k_1)$ e $\mathcal{G}(0, b, k_2)$, è Gamma $\mathcal{G}(0, b, k_1 + k_2)$. In particolare, se la legge della v.a. X è Chi-quadrato χ_n^2 , dalla precedente espressione con $b = 2$ e $k = \frac{n}{2}$, la legge risulta additiva rispetto al parametro n . Dunque, la legge della somma di due v.a. indipendenti, rispettivamente con legge Chi-quadrato $\chi_{n_1}^2$ e $\chi_{n_2}^2$, è Chi-quadrato $\chi_{n_1+n_2}^2$. Si osservi infine che l'additività non vale per il parametro di scala b . In effetti, se si considera il vettore di parametri $\theta = (b, k)^T$ con $\Theta =]0, \infty[\times]0, \infty[$, si ha

$$\varphi_{X_1, \theta_1}(t) \varphi_{X_2, \theta_2}(t) = (1 - ib_1 t)^{-k_1} (1 - ib_2 t)^{-k_2} \neq \varphi_{X_1+X_2, \theta_1+\theta_2}(t) = (1 - i(b_1 + b_2)t)^{-(k_1+k_2)}.$$

In altre parole, la legge della somma di due v.a. indipendenti, distribuite rispettivamente con legge Gamma $\mathcal{G}(0, b_1, k_1)$ e $\mathcal{G}(0, b_2, k_2)$, non è Gamma $\mathcal{G}(0, b_1 + b_2, k_1 + k_2)$. \square

• **Esempio 7.5.3.** Si consideri la v.a. X distribuita con legge Binomiale $\mathcal{B}(n, p)$. In questo caso, si pone $\theta = n$ con $\Theta = \{1, 2, \dots\}$, ovvero si vuole verificare l'additività rispetto al parametro n e non al parametro p . La legge risulta additiva rispetto a n , in quanto tenendo presente l'Esempio 7.2.1 si ha

$$G_{X_1, \theta_1}(t) G_{X_2, \theta_2}(t) = (q + pt)^{n_1} (q + pt)^{n_2} = (q + pt)^{n_1+n_2} = G_{X_1+X_2, \theta_1+\theta_2}(t).$$

Quindi, la legge della somma di due v.a. indipendenti, distribuite rispettivamente con legge Binomiale $\mathcal{B}(n_1, p)$ e $\mathcal{B}(n_2, p)$, è Binomiale $\mathcal{B}(n_1 + n_2, p)$. Risulta immediato verificare che l'additività non vale per il parametro p . Dunque, se si considera il vettore di parametri $\theta = (n, p)^T$ con $\Theta = \mathbb{Z}^+ \times]0, 1[$, la legge non risulta additiva rispetto a θ . \square

• **Esempio 7.5.4.** Si consideri la v.a. X distribuita con legge di Poisson $\mathcal{P}(\lambda)$. Si pone $\theta = \lambda$ con $\Theta =]0, \infty[$. La legge risulta additiva rispetto al parametro λ , in quanto tenendo presente l'Esempio 7.2.2 si ha

$$G_{X_1, \theta_1}(t) G_{X_2, \theta_2}(t) = e^{\lambda_1(t-1)} e^{\lambda_2(t-1)} = e^{(\lambda_1+\lambda_2)(t-1)} = G_{X_1+X_2, \theta_1+\theta_2}(t).$$

Dunque, la legge della somma di due v.a. indipendenti, distribuite rispettivamente con legge di Poisson $\mathcal{P}(\lambda_1)$ e $\mathcal{P}(\lambda_2)$, è Poisson $\mathcal{P}(\lambda_1 + \lambda_2)$. \square

• **Esempio 7.5.5.** Si consideri la v.a. X distribuita con legge Binomiale Negativa $\mathcal{BN}(k, p)$. Si pone $\theta = k$ con $\Theta = \mathbb{Z}^+$, ovvero si vuole verificare l'additività rispetto al parametro k e non al parametro p . La legge risulta additiva rispetto a θ , in quanto tenendo presente l'Esempio 7.2.3 si ha

$$G_{X_1, \theta_1}(t) G_{X_2, \theta_2}(t) = \left(\frac{p}{1 - qt} \right)^{k_1} \left(\frac{p}{1 - qt} \right)^{k_2} = \left(\frac{p}{1 - qt} \right)^{k_1+k_2} = G_{X_1+X_2, \theta_1+\theta_2}(t),$$

ovvero si ha che la legge della somma di due v.a. indipendenti, distribuite rispettivamente con legge Binomiale Negativa $\mathcal{BN}(k_1, p)$ e $\mathcal{BN}(k_2, p)$, è Binomiale Negativa $\mathcal{BN}(k_1 + k_2, p)$. In modo analogo alla legge Binomiale, risulta inoltre immediato verificare che l'additività non vale per il parametro p . Dunque, se si considera il vettore di parametri $\theta = (k, p)^T$ con $\Theta = \mathbb{Z}^+ \times]0, 1[$, la legge non è additiva rispetto a θ . \square

Definizione 7.5.2. Si consideri la v.a. X definita sullo spazio probabilizzato (Ω, \mathcal{F}, P) . La v.a. X e la rispettiva legge sono dette *infinitamente divisibili* se per ogni $n = 1, 2, \dots$ si ha $X \stackrel{\mathcal{L}}{=} \sum_{k=1}^n X_k$, dove $(X_1, \dots, X_n)^T$ è un v.v.a. indipendenti e ugualmente distribuite. \square

Tenendo presente la proprietà di unicità della f.c., la Definizione 7.5.2 può essere data in modo alternativo utilizzando la f.c. In questo caso, se la v.a. X possiede f.c. φ_X , mentre φ_{X_n} è la f.c. comune alle componenti marginali del v.v.a. $(X_1, \dots, X_n)^T$, allora la v.a. X è infinitamente divisibile se per ogni $n = 1, 2, \dots$ si ha

$$\varphi_X(t) = \varphi_{X_n}(t)^n .$$

Nel caso in cui la v.a. X è discreta a valori su $S \subseteq \mathbb{N}$, una analoga definizione può essere data in termini della f.g., nel senso che la precedente condizione è equivalente a

$$G_X(t) = G_{X_n}(t)^n .$$

Per una v.a. infinitamente divisibile vale inoltre la seguente Proposizione.

Proposizione 7.5.3. *Data la v.a. X definita sullo spazio probabilizzato (Ω, \mathcal{F}, P) , se la v.a. X è infinitamente divisibile, anche la trasformata $Y = a + bX$, dove $a, b \in \mathbb{R}$, è infinitamente divisibile.*

Dimostrazione. Tenendo presente la Proposizione 7.3.3, si ha

$$\varphi_Y(t) = e^{iat} \varphi_X(bt) = e^{iat} \varphi_{X_n}(bt)^n = (e^{\frac{iat}{n}} \varphi_{X_n}(bt))^n ,$$

da cui segue che $Y \stackrel{\mathcal{L}}{=} \sum_{k=1}^n (\frac{a}{n} + bX_k)$, ovvero la v.a. Y è infinitamente divisibile. \square

In generale, dalla precedente Proposizione, è immediato verificare che, se una legge ridotta è infinitamente divisibile, lo è anche la relativa legge con parametro di posizione e di scala. Dunque, è sufficiente verificare che legge ridotta sia infinitamente divisibile per concludere che la legge con parametro di posizione e di scala è a sua volta infinitamente divisibile.

• **Esempio 7.5.6.** Si consideri la v.a. X distribuita con legge Normale $\mathcal{N}(0, 1)$. Dal momento che per ogni $n = 1, 2, \dots$ si ha

$$\varphi_X(t) = e^{-\frac{1}{2}t^2} = \left(e^{-\frac{1}{2n}t^2} \right)^n = \varphi_{X_n}(t)^n ,$$

la legge è infinitamente divisibile. In altri termini, una v.a. X con legge Normale $\mathcal{N}(0, 1)$ è equivalente in legge alla somma di n v.a. indipendenti con legge Normale $\mathcal{N}(0, n^{-1})$ per ogni $n = 1, 2, \dots$. In base alla Proposizione 7.5.3 è inoltre sufficiente verificare che legge Normale ridotta $\mathcal{N}(0, 1)$ è infinitamente divisibile per concludere che la legge Normale $\mathcal{N}(\mu, \sigma^2)$ è a sua volta infinitamente divisibile. \square

• **Esempio 7.5.7.** Si consideri la v.a. X distribuita con legge Gamma ridotta $\mathcal{G}(0, 1, k)$. Dal momento che per ogni $n = 1, 2, \dots$ si ha

$$\varphi_X(t) = (1 - it)^{-k} = \left((1 - it)^{-\frac{k}{n}} \right)^n = \varphi_{X_n}(t)^n ,$$

la legge è infinitamente divisibile. In altri termini, una v.a. X con legge Gamma ridotta $\mathcal{G}(0, 1, k)$ è equivalente in legge alla somma di n v.a. indipendenti con legge Gamma $\mathcal{G}(0, 1, \frac{k}{n})$ per ogni $n = 1, 2, \dots$. \square

• **Esempio 7.5.8.** (Legge Stabile) Si consideri una v.a. X con f.c.

$$\varphi_X(t) = e^{-|t|^\alpha} ,$$

dove $\alpha \in]0, 2]$. La legge associata alla v.a. X è detta *legge Stabile simmetrica* di parametro α e, ovviamente, contiene come casi particolari la legge $\mathcal{N}(0, 2)$ per $\alpha = 2$ (si veda l'Esempio 7.1.6) e la legge di Cauchy ridotta per $\alpha = 1$ (si veda l'esempio 7.1.9). Dal momento che per $\alpha \in]0, 2]$ si ha

$$\int_{-\infty}^{\infty} |\varphi_X(t)| dt = \int_{-\infty}^{\infty} e^{-|t|^\alpha} dt = 2\Gamma\left(1 + \frac{1}{\alpha}\right) < \infty,$$

sulla base del Teorema 7.1.10 la v.a. X è assolutamente continua. La d.p. della v.a. X non può essere ottenuta in forma chiusa, eccetto che per i due casi particolari precedentemente citati. Inoltre, tenendo presente la Proposizione 7.1.6, $E[X]$ non è definito se $\alpha \in]0, 1]$, mentre esiste finito $E[X]$ ma non $E[X^2]$ se $\alpha \in]1, 2[$. Ovviamente, i momenti di ogni ordine esistono finiti se $\alpha = 2$. Dal momento che per ogni legge della famiglia si ha

$$\varphi_X(t) = e^{-|t|^\alpha} = (e^{-|n^{-\frac{1}{\alpha}}t|^\alpha})^n = \varphi_{X_n}(t)^n,$$

la legge Stabile simmetrica è infinitamente divisibile. In altri termini, una v.a. X con legge Stabile simmetrica di parametro α è equivalente in legge alla somma di n v.a. indipendenti del tipo $n^{-\frac{1}{\alpha}}X_k$, $k = 1, \dots, n$, dove la legge di ogni v.a. X_k è ancora Stabile simmetrica di parametro α . \square

• **Esempio 7.5.9.** Si consideri la v.a. X distribuita con legge Binomiale $\mathcal{B}(m, p)$. Dal momento che si ha

$$G_X(t) = (q + pt)^m = \left((q + pt)^{\frac{m}{n}}\right)^n,$$

la legge è infinitamente divisibile solamente se m è multiplo di n . Dunque, la legge Binomiale non è infinitamente divisibile. \square

• **Esempio 7.5.10.** Si consideri la v.a. X distribuita con legge di Poisson $\mathcal{P}(\lambda)$. Dal momento che si ha

$$G_X(t) = e^{\lambda(t-1)} = \left(e^{\frac{\lambda}{n}(t-1)}\right)^n = G_{X_n}(t)^n,$$

la legge è infinitamente divisibile. In altri termini, una v.a. X distribuita con legge di Poisson $\mathcal{P}(\lambda)$ è equivalente in legge alla somma di n v.a. indipendenti con legge di Poisson $\mathcal{P}(\frac{\lambda}{n})$ per ogni $n = 1, 2, \dots$. \square

7.6. Riferimenti bibliografici

Lukacs (1970) è una monografia sulle funzioni caratteristiche. I classici testi di Feller (1968, 1971) trattano estesamente le funzioni generatrici e caratteristiche. Un testo più recente sulle funzioni caratteristiche è Sasvári (2013). Wilf (2006) è una monografia espressamente dedicata alle funzioni generatrici. Sull'argomento si veda anche Petkovšek *et al.* (1996) e Spivey (2019). Una trattazione esauriente delle leggi infinitamente divisibili è data in Sato (1999) e Steutel e van Harn (2004). Le leggi stabili sono estesamente considerate in Uchaikin e Zolotarev (1999) e Zolotarev (1986).

7.7. Esercizi svolti

• **Nota.** Al fine di evitare notazioni ridondanti e ripetizioni, nei seguenti esercizi si assume l'esistenza di uno spazio probabilizzato (Ω, \mathcal{F}, P) , a meno che non venga specificato diversamente.

Sezione 7.1

• **Esercizio 7.1.1.** Si consideri la v.a. Z con legge Normale ridotta $\mathcal{N}(0, 1)$ e se ne determini la f.c.

Soluzione. Tenendo presente che la v.a. Z è simmetrica rispetto all'origine e che la funzione seno è dispari, dalla definizione di f.c. si ha

$$\varphi_Z(t) = E[\cos(tZ)] = \int_{-\infty}^{\infty} \cos(tz) \phi(z) dz .$$

Differenziando i termini della precedente relazione ed effettuando una integrazione per parti tenendo presente che

$$\frac{d}{dz} \phi(z) = -z\phi(z) ,$$

si ottiene

$$\frac{d}{dt} \varphi_Z(t) = - \int_{-\infty}^{\infty} \sin(tz) z\phi(z) dz = \sin(tz) \phi(z) \Big|_{-\infty}^{\infty} - t \int_{-\infty}^{\infty} \cos(tz) \phi(z) dz = -t\varphi_Z(t) .$$

Dunque, la precedente espressione fornisce una equazione differenziale con condizione iniziale data da $\varphi_Z(0) = 1$, di cui si determina facilmente la soluzione $\varphi_Z(t) = e^{-\frac{1}{2}t^2}$. \square

• **Esercizio 7.1.2.** Si consideri la v.a. assolutamente continua X con f.c.

$$\varphi_X(t) = \left(1 - \frac{|t|}{\pi}\right) \mathbf{1}_{[-\pi, \pi]}(t) .$$

Si determini la d.p. f_X .

Soluzione. Si ha

$$\int_{-\infty}^{\infty} |\varphi_X(t)| dt = \int_{-\pi}^{\pi} \left(1 - \frac{|t|}{\pi}\right) dt = \pi < \infty .$$

Dal momento che la f.c. φ_X è una funzione pari, dal Teorema di Inversione di Lévy si ottiene

$$f_X(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-itx} \left(1 - \frac{|t|}{\pi}\right) dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} \cos(tx) \left(1 - \frac{|t|}{\pi}\right) dt = \frac{1 - \cos(\pi x)}{\pi^2 x^2} .$$

La d.p. f_X è in effetti tale, in quanto $f_X(x) \geq 0$ per ogni $x \in \mathbb{R}$ ed inoltre

$$\int_{-\infty}^{\infty} f_X(x) dx = \int_{-\infty}^{\infty} \frac{1 - \cos(\pi x)}{\pi^2 x^2} dx = \int_{-\infty}^{\infty} \frac{2 \sin^2(\frac{\pi x}{2})}{\pi^2 x^2} dx = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\sin^2(x)}{x^2} dx = 1 .$$

Per una verifica probabilistica dell'integrale $\int_{-\infty}^{\infty} \left(\frac{\sin(x)}{x}\right)^2 dx = \pi$ mediante l'uso delle f.c. si veda l'Esercizio 7.3.5. \square

• **Esercizio 7.1.3.** Si considerino le v.a. indipendenti X_1 e X_2 con legge Normale ridotta $\mathcal{N}(0, 1)$ e si determini la f.c. della v.a. $Y = X_1 X_2$.

Soluzione. Tenendo presente l'espressione della f.c. di una v.a. con legge Normale, si ha

$$\varphi_Y(t) = E[e^{itX_1 X_2}] = E[E[e^{itX_1 X_2} | X_2]] = E[e^{-\frac{1}{2}t^2 X_2^2}] .$$

Dunque, risulta

$$\varphi_Y(t) = \int_{-\infty}^{\infty} e^{-\frac{1}{2}t^2 x_2^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_2^2} dx_2 = \frac{1}{\sqrt{1+t^2}} \int_{-\infty}^{\infty} \frac{\sqrt{1+t^2}}{\sqrt{2\pi}} e^{-\frac{1}{2}(1+t^2)x_2^2} dx_2 = \frac{1}{\sqrt{1+t^2}} ,$$

dove si è tenuto presente che la funzione integranda nell'ultimo integrale è la d.p. relativa ad una v.a. con legge Normale $\mathcal{N}(0, (1+t^2)^{-1})$. \square

• **Esercizio 7.1.4.** (Mistura di leggi rispetto al parametro di scala) Si considerino le v.a. indipendenti X_1 e X_2 , rispettivamente distribuite con legge Normale ridotta $\mathcal{N}(0, 1)$ e legge Esponenziale ridotta $\mathcal{G}(0, 1, 1)$. Si determini la legge della v.a. $Y = X_1\sqrt{2X_2}$.

Soluzione. Tenendo presente l'espressione della f.c. di una v.a. con legge Normale, si ha

$$\varphi_Y(t) = E[e^{itX_1\sqrt{2X_2}}] = E[E[e^{itX_1\sqrt{2X_2}} \mid X_2]] = E[e^{-t^2X_2}].$$

Dunque, risulta

$$\varphi_Y(t) = \int_0^\infty e^{-t^2x_2} e^{-x_2} dx_2 = \frac{1}{1+t^2}.$$

Tenendo inoltre presente che

$$\int_{-\infty}^\infty |\varphi_Y(t)| dt = \int_{-\infty}^\infty \frac{1}{1+t^2} dt = \pi < \infty$$

e dal momento che la f.c. φ_Y è simmetrica rispetto all'origine, la d.p. della v.a. Y è data da

$$f_Y(y) = \frac{1}{2\pi} \int_{-\infty}^\infty e^{-ity} \frac{1}{1+t^2} dt = \frac{1}{2\pi} \int_{-\infty}^\infty \cos(ty) \frac{1}{1+t^2} dt = \frac{1}{2} e^{-|y|}.$$

Si osservi che la v.a. Y si distribuisce con legge di Laplace. Inoltre, la legge può essere considerata come una mistura di leggi Normali rispetto al parametro di scala (dove il parametro di scala è aleatorio). \square

Sezione 7.2

• **Esercizio 7.2.1.** (Valori record, terza parte) Tenendo presente gli Esercizi 2.7.3 e 3.3.2, si consideri la v.a. discreta X con f.p. data da

$$p_X(x) = \frac{1}{x(x+1)} \mathbf{1}_{\{1,2,\dots\}}(x)$$

e si determini la f.g. G_X . Si calcoli inoltre $E[X]$ mediante la f.g.

Soluzione. Tenendo presente la serie logaritmica, si ha

$$\begin{aligned} G_X(t) &= \sum_{x=1}^\infty t^x \frac{1}{x(x+1)} = \sum_{x=1}^\infty t^x \left(\frac{1}{x} - \frac{1}{x+1} \right) = \sum_{x=1}^\infty \frac{t^x}{x} - \frac{1}{t} \sum_{x=1}^\infty \frac{t^{x+1}}{x+1} \\ &= -\log(1-t) - \frac{1}{t} (-t - \log(1-t)) = 1 + \frac{1-t}{t} \log(1-t). \end{aligned}$$

Inoltre, risulta

$$G_X^{(1)}(t) = -\frac{t + \log(1-t)}{t^2},$$

da cui $E[X] = \lim_{t \rightarrow 1^-} G_X^{(1)}(t) = \infty$. \square

• **Esercizio 7.2.2.** (Legge Uniforme discreta) Si consideri la v.a. discreta X con f.p. data da

$$p_X(x) = \frac{1}{N} \mathbf{1}_{\{1,\dots,N\}}(x),$$

dove $N \in \mathbb{Z}^+$, ovvero la v.a. X si distribuisce con legge Uniforme discreta su $\{1, \dots, N\}$. Si determini la f.g. G_X della v.a. X . Si calcoli inoltre $E[X]$ e $\text{Var}[X]$ mediante la f.g.

Soluzione. Tenendo presente le proprietà della serie geometrica, per $t \neq 1$ la f.g. della v.a. X può essere espressa come

$$G_X(t) = \frac{1}{N} \sum_{x=1}^N t^x = \frac{t(1-t^N)}{N(1-t)},$$

mentre $G_X(1) = 1$. Dunque, si ha

$$G_X^{(1)}(t) = \frac{1 - (N+1)t^N + Nt^{N+1}}{N(1-t)^2},$$

e quindi

$$E[X] = \lim_{t \rightarrow 1^-} G_X^{(1)}(t) = \frac{N+1}{2}.$$

Inoltre, risulta

$$G_X^{(2)}(t) = \frac{2 - N(N+1)t^{N-1} + 2(N^2-1)t^N - N(N-1)t^{N+1}}{N(1-t)^3},$$

e quindi

$$E[X(X-1)] = \lim_{t \rightarrow 1^-} G_X^{(2)}(t) = \frac{N^2-1}{3},$$

da cui

$$\text{Var}[X] = \frac{N^2-1}{12}.$$

In questo esercizio sono state implicitamente dimostrate le due relazioni $\sum_{k=1}^n k = \frac{1}{2}n(n+1)$ e $\sum_{k=1}^n k^2 = \frac{1}{6}n(n+1)(2n+1)$. Alla stessa maniera, con calcoli evidentemente sempre più laboriosi, si può ottenere le espressioni di $\sum_{k=1}^n k^s$ per $s = 3, 4, \dots$ \square

• **Esercizio 7.2.3.** (Mistura di leggi) Dato il v.v.a. $(X, Y)^T$, si assuma che la legge condizionata della v.a. X all'evento $\{Y = y\}$ sia la legge di Poisson $\mathcal{P}(y)$. Inoltre, si assuma che la v.a. Y si distribuisca con legge Esponenziale ridotta $\mathcal{G}(0, 1, 1)$. Si determini la legge della v.a. X .

Soluzione. Tenendo presente l'espressione della f.g. di una v.a. con legge di Poisson, la f.g. della v.a. X è data da

$$G_X(t) = E[E[t^X | Y]] = E[e^{Y(t-1)}] = \int_0^\infty e^{y(t-1)} e^{-y} dy = \frac{1}{2-t} = \frac{\frac{1}{2}}{1 - \frac{1}{2}t}.$$

Dunque, la v.a. X si distribuisce con legge Geometrica di parametro $p = \frac{1}{2}$. In modo alternativo, la f.p. della v.a. X può essere ottenuta come

$$p_X(x) = \int_0^\infty e^{-y} \frac{y^x}{x!} \mathbf{1}_{\mathbb{N}}(x) e^{-y} dy = \frac{1}{x!} 2^{-x-1} \mathbf{1}_{\mathbb{N}}(x) \int_0^\infty z^x e^{-z} dz = 2^{-x-1} \mathbf{1}_{\mathbb{N}}(x),$$

ovvero si ha una conferma del precedente risultato. In generale, una *mistura di leggi* viene definita come una legge con un parametro che a sua volta è una v.a. con una ulteriore legge. Nel presente caso, la legge Geometrica è una mistura di leggi di Poisson il cui parametro si distribuisce con legge Esponenziale ridotta. \square

Sezione 7.3

• **Esercizio 7.3.1.** Si considerino le v.a. indipendenti X_1 e X_2 che si distribuiscono con la stessa legge della v.a. X . Se φ_X è la f.c. della v.a. X , si verifichi che la f.c. della v.a. $Y = X_1 - X_2$ è data da $\varphi_Y(t) = |\varphi_X(t)|^2$.

Soluzione. La f.c. della v.a. Y risulta

$$\varphi_Y(t) = E[e^{it(X_1 - X_2)}] = E[e^{itX_1}]E[e^{-itX_2}] = \varphi_X(t)\varphi_X(-t).$$

Dalla definizione di f.c. si ha inoltre

$$\begin{aligned} \varphi_X(t)\varphi_X(-t) &= (E[\cos(tX)] + iE[\sin(tX)])(E[\cos(-tX)] + iE[\sin(-tX)]) && \square \\ &= (E[\cos(tX)] + iE[\sin(tX)])(E[\cos(tX)] - iE[\sin(tX)]) \\ &= E[\cos(tX)]^2 + E[\sin(tX)]^2 = |\varphi_X(t)|^2. \end{aligned}$$

• **Esercizio 7.3.2.** Si consideri la v.a. X distribuita con legge di Cauchy ridotta, ovvero la v.a. X ammette d.p. data da

$$f_X(x) = \frac{1}{\pi(1+x^2)}.$$

Si determini la f.c. φ_X della v.a. X .

Soluzione. Dal momento che la v.a. X è simmetrica rispetto all'origine, la f.c. può essere espressa come

$$\varphi_X(t) = E[\cos(tZ)] = \int_{-\infty}^{\infty} \cos(tx) \frac{1}{\pi(1+x^2)} dx.$$

Al fine di calcolare il precedente integrale, si considerino due v.a. Y_1 e Y_2 indipendenti con legge Esponenziale ridotta e la v.a. differenza $Y = Y_1 - Y_2$. Tenendo presente l'Esercizio 6.8.1, la v.a. Y si distribuisce con legge di Laplace ridotta e dunque si ha

$$f_Y(y) = \frac{1}{2} e^{-|y|}.$$

Inoltre, risulta $\varphi_{Y_1}(t) = \varphi_{Y_2}(t) = (1 - it)^{-1}$ e, sulla base dell'Esercizio 7.3.1, la f.c. della v.a. Y è data da $\varphi_Y(t) = (1 + t^2)^{-1}$. Dal momento che

$$\int_{-\infty}^{\infty} |\varphi_Y(t)| dt = \int_{-\infty}^{\infty} \frac{1}{1+t^2} dt = \pi < \infty,$$

dal Teorema di Inversione si ottiene un'ulteriore espressione per f_Y data da

$$f_Y(y) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ity} \frac{1}{1+t^2} dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} \cos(ty) \frac{1}{1+t^2} dt,$$

ovvero si ha

$$\int_{-\infty}^{\infty} \cos(ty) \frac{1}{\pi(1+t^2)} dt = e^{-|y|}.$$

Quindi, si deve concludere che $\varphi_X(t) = e^{-|t|}$. □

• **Esercizio 7.3.3.** Si consideri la v.a. X distribuita con legge di Cauchy ridotta descritta nell'Esercizio 7.3.2. Se le v.a. X_1 e X_2 sono distribuite con la stessa legge di X e sono v.a. tali che $\{X_1 = X_2\}$ q.c., si verifichi che

$$\varphi_{X_1+X_2}(t) = \varphi_{X_1}(t)\varphi_{X_2}(t).$$

Soluzione. Ovviamente, le v.a. X_1 e X_2 non sono indipendenti. Dall'Esercizio 7.3.2 la f.c. della v.a. X risulta $\varphi_X(t) = e^{-|t|}$. Dal momento che $X_1 + X_2 \stackrel{L}{=} Y$ dove $Y = 2X$, si ha

$$\varphi_{X_1+X_2}(t) = \varphi_Y(t) = \varphi_X(2t) = e^{-2|t|} = \varphi_{X_1}(t)\varphi_{X_2}(t).$$

Dunque, la f.c. della somma di due v.a. può essere identica al prodotto delle f.c. delle v.a., anche se queste non sono indipendenti. \square

• **Esercizio 7.3.4.** Si consideri la v.a. X con legge di Laplace ridotta descritta nell'Esercizio 6.8.1. Se inoltre $(X_1, \dots, X_n)^T$ è un v.v.a. con componenti indipendenti e con la stessa legge di X , si determini la legge della v.a. $Y = \sum_{k=1}^n |X_k|$.

Soluzione. Posto $Z = |X|$, dal momento che la v.a. X è simmetrica rispetto all'origine, allora risulta

$$f_Z(z) = 2f_X(z) \mathbf{1}_{[0, \infty[}(z) = e^{-z} \mathbf{1}_{[0, \infty[}(z),$$

ovvero la v.a. Z si distribuisce con legge Esponenziale ridotta. Dal momento che si ha $\varphi_Z(t) = (1 - it)^{-1}$, allora

$$\varphi_Y(t) = \varphi_Z(t)^n = (1 - it)^{-n},$$

ovvero la v.a. Y si distribuisce con legge Gamma ridotta $\mathcal{G}(0, 1, n)$. \square

• **Esercizio 7.3.5.** (Legge Triangolare, terza parte) Si consideri la v.a. X distribuita con la legge Triangolare descritta nell'Esercizio 3.7.1. Sulla base delle proprietà di questa legge e mediante l'uso della f.c. si verifichi che

$$\int_{-\infty}^{\infty} \frac{\sin^2(x)}{x^2} dx = \pi.$$

Soluzione. Tenendo presente l'Esercizio 3.7.2, si ha $X = \frac{1}{2}Y$, dove Y è una v.a. data dalla somma di due v.a. che si distribuiscono come una v.a. Z con legge Uniforme su $] -1, 1[$. Dal momento che

$$\varphi_Z(t) = \int_{-1}^1 e^{itz} \frac{1}{2} dz = \frac{1}{2} \int_{-1}^1 \cos(tz) dz = \frac{\sin(t)}{t},$$

allora si ha $\varphi_Y(t) = \frac{\sin^2(t)}{t^2}$ e quindi

$$\varphi_X(t) = \varphi_Y\left(\frac{t}{2}\right) = \frac{\sin^2\left(\frac{t}{2}\right)}{\left(\frac{t}{2}\right)^2}.$$

Inoltre, essendo

$$\int_{-\infty}^{\infty} |\varphi_X(t)| dt = \int_{-\infty}^{\infty} \frac{\sin^2\left(\frac{t}{2}\right)}{\left(\frac{t}{2}\right)^2} dt = 2 \int_{-\infty}^{\infty} \frac{\sin^2(t)}{t^2} dt < 2 \int_{-\infty}^{\infty} \min(1, t^{-2}) dt = 8 < \infty,$$

dal Teorema di Inversione si ottiene un'ulteriore espressione per f_X data da

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \frac{\sin^2\left(\frac{t}{2}\right)}{\left(\frac{t}{2}\right)^2} dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \cos(tx) \frac{\sin^2\left(\frac{t}{2}\right)}{\left(\frac{t}{2}\right)^2} dt = \frac{1}{\pi} \int_{-\infty}^{\infty} \cos(2tx) \frac{\sin^2(t)}{t^2} dt,$$

ovvero, tenendo presente l'Esercizio 3.7.1, si ha

$$(1 - |x|) \mathbf{1}_{]-1,1[}(x) = \frac{1}{\pi} \int_{-\infty}^{\infty} \cos(2tx) \frac{\sin^2(t)}{t^2} dt .$$

Ponendo $x = 0$ nella precedente espressione si ha l'integrale desiderato. \square

• **Esercizio 7.3.6.** Si considerino le v.a. indipendenti X_1 e X_2 , tali che X_1 si distribuisce con legge Uniforme su $[-1, 1]$, mentre X_2 è una v.a. discreta con f.p. data da

$$p_{X_2}(x_2) = \frac{1}{2} \mathbf{1}_{\{-1,1\}}(x_2) .$$

Si determini la legge della v.a. $Y = X_1 + X_2$.

Soluzione. La f.c. della v.a. X_1 è data da $\varphi_{X_1}(t) = \frac{\sin(t)}{t}$, mentre la f.c. della v.a. X_2 è data da $\varphi_{X_2}(t) = \cos(t)$. Dunque, la f.c. della v.a. Y risulta

$$\varphi_Y(t) = \varphi_{X_1}(t)\varphi_{X_2}(t) = \frac{\sin(t)\cos(t)}{t} = \frac{\sin(2t)}{2t} ,$$

da cui risulta $Y \stackrel{\mathcal{L}}{=} 2X_1$. Dunque, la v.a. Y si distribuisce con legge Uniforme su $[-2, 2]$. \square

• **Esercizio 7.3.7.** (Identità di Wald) Si consideri la v.a. $X = \sum_{k=1}^Y Z_k$, dove Y è una v.a. discreta a valori su \mathbb{N} tale che $E[Y] < \infty$, mentre le Z_k sono v.a. indipendenti con la medesima legge della v.a. Z , dove $E[Z] < \infty$. Inoltre le v.a. Z_k sono indipendenti dalla v.a. Y . Si verifichi che risulta $E[X] = E[Y]E[Z]$.

Soluzione. Si ha

$$\varphi_X(t) = E[e^{it\sum_{k=1}^Y Z_k}] = E[E[e^{it\sum_{k=1}^Y Z_k} | Y]] = E[\varphi_Z(t)^Y] = G_Y(\varphi_Z(t)) .$$

Dunque, dalla precedente relazione risulta

$$\varphi_X^{(1)}(t) = G_Y^{(1)}(\varphi_Z(t))\varphi_Z^{(1)}(t) ,$$

da cui

$$E[X] = i^{-1}\varphi_X^{(1)}(0) = i^{-1}G_Y^{(1)}(\varphi_Z(0))\varphi_Z^{(1)}(0) = G_Y^{(1)}(1) i^{-1}\varphi_Z^{(1)}(0) = E[Y]E[Z] .$$

L'identità prende il nome dal matematico e probabilista ungherese Abraham Wald (1902-1950), che ha introdotto le tecniche di statistica sequenziale. \square

Sezione 7.4

• **Esercizio 7.4.1.** (Legge di Poisson composta) Si consideri la v.a. Y con legge di Poisson $\mathcal{P}(\lambda)$ e la v.a. discreta Z con f.p.

$$p_Z(z) = - \frac{1}{\log(1 - q)} \frac{q^z}{z} \mathbf{1}_{\{1,2,\dots\}}(z) ,$$

dove $q \in]0,1[$. Si determini la legge relativa alla v.a. $X = \sum_{k=1}^Y Z_k$, dove le Z_k sono v.a. indipendenti con la medesima legge della v.a. Z .

Soluzione. Tenendo presente l'Esempio 3.3.1, si osservi che la v.a. Z si distribuisce con legge Logaritmica di parametro q . La f.g. della v.a. Z è quindi data da

$$G_Z(t) = \sum_{z=1}^{\infty} t^z \left(-\frac{1}{\log(1-q)} \frac{q^z}{z} \right) = -\frac{1}{\log(1-q)} \sum_{z=1}^{\infty} \frac{(qt)^z}{z} = \frac{\log(1-qt)}{\log(1-q)}.$$

Dal momento che X è la somma di v.a. indipendenti con la medesima legge e un numero aleatorio di addendi distribuito con legge di Poisson, la f.g. della v.a. X è data da

$$G_X(t) = E[t^{\sum_{k=1}^Y Z_k}] = E[E[t^{\sum_{k=1}^Y Z_k} | Y]] = E[G_Z(t)^Y] = G_Y(G_Z(t)),$$

da cui

$$G_X(t) = e^{\lambda \left(\frac{\log(1-qt)}{\log(1-q)} - 1 \right)} = \left(\frac{1-q}{1-qt} \right)^{-\frac{\lambda}{\log(1-q)}} = \left(\frac{p}{1-qt} \right)^k,$$

dove si è posto $p = 1 - q$ e $k = -\frac{\lambda}{\log(p)}$. Si deve dunque concludere che la v.a. X si distribuisce con legge Binomiale Negativa $\mathcal{BN}(k, p)$. La legge Binomiale Negativa è quindi una distribuzione di Poisson composta, ovvero è la legge della somma di v.a. indipendenti e ugualmente distribuite con legge Logaritmica e con un numero aleatorio di addendi distribuito con legge di Poisson. \square

• **Esercizio 7.4.2.** (Legge di Poisson composta, seconda parte) Si consideri la v.a. X tale che $X = \sum_{k=1}^Y Z_k$, dove la legge della v.a. Y è Poisson $\mathcal{P}(\lambda)$, mentre le Z_k sono v.a. indipendenti con legge di Bernoulli $\mathcal{B}(1, p)$. Si determini la legge della v.a. X .

Soluzione. In modo simile all'Esercizio 7.4.1, tenendo presente l'espressione della f.g. di una v.a. con legge di Bernoulli $\mathcal{B}(1, p)$, si ha

$$G_X(t) = G_Y(G_Z(t)) = G_Y(1 - p + pt) = e^{p\lambda(t-1)}.$$

Dunque, si deve concludere che la v.a. X si distribuisce con legge di Poisson $\mathcal{P}(p\lambda)$. \square

• **Esercizio 7.4.3.** (Formula di Waring) Si dimostri mediante l'uso delle funzioni generatrici la formula di Waring introdotta nell'Esercizio 2.2.1.

Soluzione. Assumendo le notazioni dell'Esercizio 2.2.1, sia $Y = \sum_{k=1}^n X_k$, dove $X_k = \mathbf{1}_{E_k}$. Evidentemente, il v.v.a. $(X_1, \dots, X_n)^T$ possiede componenti marginali che sono v.a. non indipendenti con legge di Bernoulli di parametro $P(E_k) = E[\mathbf{1}_{E_k}] = E[X_k]$. La f.g. della v.a. X_k è quindi data da

$$G_{X_k}(t) = 1 - E[X_k] + E[X_k]t = E[1 + (t-1)X_k]$$

e quindi la f.g. della v.a. Y risulta

$$G_Y(t) = E\left[\prod_{k=1}^n (1 + (t-1)X_k)\right] = E\left[1 + \sum_{k=1}^n \sum_{1 \leq j_1 < \dots < j_k \leq n} (t-1)^k X_{j_1} \dots X_{j_k}\right].$$

Tenendo presente che $E[X_{j_1} \dots X_{j_k}] = P(E_{j_1} \cap \dots \cap E_{j_k})$, si ha dunque

$$G_Y(t) = 1 + \sum_{k=1}^n (t-1)^k \sum_{1 \leq j_1 < \dots < j_k \leq n} P(E_{j_1} \cap \dots \cap E_{j_k}) = 1 + \sum_{k=1}^n (t-1)^k S_{k,n}.$$

Dal momento che

$$G_Y^{(m)}(t) = \sum_{k=m}^n \frac{k!}{(k-m)!} (t-1)^{k-m} S_{k,m},$$

per le proprietà della f.g. si ha

$$p_Y(m) = \frac{G_Y^{(m)}(0)}{m!} = \sum_{k=m}^n \binom{k}{m} (-1)^{k-m} S_{k,n}.$$

Si osservi che

$$P\left(\bigcup_{k=1}^n E_k\right) = 1 - p_Y(0) = \sum_{k=1}^n (-1)^{k+1} S_{k,n},$$

ovvero si è anche ottenuto una ulteriore dimostrazione della Formula di Inclusionione ed Esclusione. \square

Sezione 7.5

• **Esercizio 7.5.1.** Si consideri la v.a. X con legge di Cauchy descritta nell'Esercizio 7.3.2 e si verifichi che è infinitamente divisibile.

Soluzione. Tenendo presente l'Esercizio 7.3.2, per ogni $n = 1, 2, \dots$ si ha

$$\varphi_X(t) = e^{-|t|} = \left(e^{-\frac{1}{n}|t|}\right)^n = \varphi_{X_n}(t)^n,$$

dove per le proprietà della f.c. la v.a. X_n ammette d.p. data da

$$f_{X_n}(x) = \frac{n}{\pi(1 + n^2x^2)}.$$

Dunque, la legge è infinitamente divisibile. In altri termini, una v.a. X con legge di Cauchy è equivalente in legge alla somma di n v.a. indipendenti con legge di Cauchy con parametro di scala pari a n^{-1} per ogni $n = 1, 2, \dots$ \square

• **Esercizio 7.5.2.** (Legge di Poisson composta, terza parte) Si consideri la v.a. X tale che $X = \sum_{k=1}^Y Z_k$, dove la v.a. Y si distribuisce con legge di Poisson $\mathcal{P}(\lambda)$, mentre le Z_k sono v.a. indipendenti con la stessa legge della v.a. Z . Si verifichi che la v.a. X è infinitamente divisibile.

Soluzione. Tenendo presente l'espressione della f.g. della v.a. Y , la f.c. della v.a. X è data da

$$\varphi_X(t) = E[e^{it\sum_{k=1}^Y Z_k}] = E[E[e^{it\sum_{k=1}^Y Z_k} | Y]] = E[\varphi_Z(t)^Y] = G_Y(\varphi_Z(t)) = e^{\lambda(\varphi_Z(t)-1)}.$$

Dunque, la v.a. è infinitamente divisibile dal momento che per ogni $n = 1, 2, \dots$ si ha

$$\varphi_X(t) = \left(e^{\frac{\lambda}{n}(\varphi_Z(t)-1)}\right)^n = \varphi_{X_n}(t)^n,$$

dove $X_n = \sum_{k=1}^{Y_n} Z_k$ e la legge della v.a. Y_n è Poisson $\mathcal{P}(\frac{\lambda}{n})$. \square

Pagina intenzionalmente vuota

Capitolo 8

Convergenze

8.1. Convergenza in legge

Un insieme numerabile di v.a. (X_1, X_2, \dots) è detta *successione di v.a.* e viene indicata con la notazione $(X_n)_{n \geq 1}$. Se la successione viene indicizzata come (X_0, X_1, \dots) , si adotta eventualmente la notazione $(X_n)_{n \geq 0}$. Dal momento che alla successione $(X_n)_{n \geq 1}$ è associata la corrispondente successione di f.r. $(F_{X_n})_{n \geq 1}$, risulta interessante definire il concetto di convergenza per una successione di f.r. Formalmente, la cosiddetta *convergenza in legge* è data nella seguente definizione.

Definizione 8.1.1. Se $(X_n)_{n \geq 1}$ è una successione di v.a., a cui corrisponde la successione di f.r. $(F_{X_n})_{n \geq 1}$, si dice che la successione di v.a. *converge in legge* (o *in distribuzione*) alla v.a. X con f.r. F_X se per ogni x tale che $\Delta F_X(x) = 0$ si ha

$$\lim_n F_{X_n}(x) = F_X(x)$$

e si scrive $X_n \xrightarrow{\mathcal{L}} X$ per $n \rightarrow \infty$. □

In pratica, la definizione di convergenza in legge implica che la f.r. della v.a. X_n converga puntualmente a quella della v.a. X per ogni x per cui F_X è continua. Si osservi inoltre che la convergenza in legge non richiede che le v.a. della successione e la v.a. X siano definite su uno stesso spazio probabilizzato. In effetti, la condizione della Definizione 8.1.1 riguarda solamente le f.r. e introduce la cosiddetta *convergenza debole* della successione di f.r. $(F_{X_n})_{n \geq 1}$ alla f.r. F_X . Data la biunivocità esistente fra legge e f.r. di una v.a., se $(P_{X_n})_{n \geq 1}$ denota la successione di leggi corrispondente alla successione di f.r. $(F_{X_n})_{n \geq 1}$ e se P_X è la legge relativa alla f.r. F_X , la convergenza debole può essere anche definita mediante la condizione

$$\lim_n P_{X_n}(B) = P_X(B)$$

per ogni $B \in \mathcal{B}(\mathbb{R})$ la cui frontiera (ovvero l'insieme di punti di discontinuità di $\mathbf{1}_B$) ha probabilità nulla secondo P_X .

Nel caso in cui la v.a. X sia degenera e tale che $P(X = c) = 1$, allora si ha $F_X(x) = \mathbf{1}_{[c, \infty[}(x)$ e F_X è continua per $x \neq c$. Dunque, la condizione della Definizione 8.1.1 si riduce a

$$\lim_n F_{X_n}(x) = \mathbf{1}_{[c, \infty[}(x)$$

per ogni $x \neq c$. In questo caso, con un leggero abuso in notazione generalmente adottato nei testi di Teoria della Probabilità, la convergenza in legge viene denotata come $X_n \xrightarrow{\mathcal{L}} c$ per $n \rightarrow \infty$.

• **Esempio 8.1.1.** Si consideri la successione di v.a. $(X_n)_{n \geq 1}$ tale che la v.a. X_n si distribuisce con legge Beta ridotta $\mathcal{BE}(0, 1, 1 + n^{-1}, 1)$. Tenendo presente quanto detto nella Sezione 6.9, si ha

$$F_{X_n}(x) = x^{1+\frac{1}{n}} \mathbf{1}_{]0,1[}(x) + \mathbf{1}_{[1, \infty[}(x),$$

da cui

$$\lim_n F_{X_n}(x) = x \mathbf{1}_{]0,1[}(x) + \mathbf{1}_{[1,\infty[}(x).$$

Dal momento che la precedente espressione fornisce la f.r. di una v.a. X con legge Uniforme su $]0, 1[$, allora si ha $X_n \xrightarrow{\mathcal{L}} X$ per $n \rightarrow \infty$. \square

• **Esempio 8.1.2.** Si consideri la successione di v.a. $(X_n)_{n \geq 1}$ tale che la v.a. X_n si distribuisce con legge Uniforme su $[-n^{-1}, n^{-1}]$. Tenendo presente quanto detto nella Sezione 6.9, si ha

$$F_{X_n}(x) = \frac{nx + 1}{2} \mathbf{1}_{[-n^{-1}, n^{-1}]}(x) + \mathbf{1}_{]n^{-1}, \infty[}(x),$$

da cui

$$\lim_n F_{X_n}(x) = \frac{1}{2} \mathbf{1}_{\{0\}}(x) + \mathbf{1}_{]0, \infty[}(x).$$

Risulta immediato verificare che la precedente espressione non costituisce una f.r. Tuttavia, si consideri la v.a. degenere X concentrata sullo 0 e che possiede f.r. data da $F_X(x) = \mathbf{1}_{[0, \infty[}(x)$. Dal momento che $\lim_n F_{X_n}(x) = F_X(x)$ coincidono per ogni $x \neq 0$, allora si ha $X_n \xrightarrow{\mathcal{L}} 0$ per $n \rightarrow \infty$. \square

• **Esempio 8.1.3.** Sia $(Z_n)_{n \geq 1}$ una successione di v.a. indipendenti definite sullo stesso spazio probabilizzato (Ω, \mathcal{F}, P) e tale che la v.a. Z_n si distribuisce con legge Esponenziale ridotta $\mathcal{G}(0, 1, 1)$ per ogni n . Inoltre, si consideri la successione di v.a. $(Y_n)_{n \geq 1}$, dove $Y_n = \max_{1 \leq k \leq n} Z_k$. Tenendo presente i risultati della Sezione 6.8, dall'assunzione di indipendenza si ha

$$\begin{aligned} F_{Y_n}(x) &= P(\max_{1 \leq k \leq n} Z_k \leq x) = P(Z_1 \leq x, \dots, Z_n \leq x) = \prod_{k=1}^n P(Z_k \leq x) \\ &= \prod_{k=1}^n F_{Z_k}(x) = \prod_{k=1}^n (1 - e^{-x}) \mathbf{1}_{]0, \infty[}(x) = (1 - e^{-x})^n \mathbf{1}_{]0, \infty[}(x). \end{aligned}$$

Si consideri l'ulteriore successione di v.a. $(X_n)_{n \geq 1}$, con $X_n = Y_n - \log(n)$. In questo caso, si ha

$$\begin{aligned} F_{X_n}(x) &= P(X_n \leq x) = P(Y_n - \log(n) \leq x) = P(Y_n \leq x + \log(n)) \\ &= F_{Y_n}(x + \log(n)) = (1 - e^{-x - \log(n)})^n \mathbf{1}_{] - \log(n), \infty[}(x) = \left(1 - \frac{e^{-x}}{n}\right)^n \mathbf{1}_{] - \log(n), \infty[}(x), \end{aligned}$$

da cui

$$\lim_n F_{X_n}(x) = \lim_n \left(1 - \frac{e^{-x}}{n}\right)^n \mathbf{1}_{] - \log(n), \infty[}(x) = e^{-e^{-x}}.$$

La precedente espressione fornisce la f.r. di una v.a. X con legge di Gumbel ridotta (si veda l'Esempio 6.6.1) e quindi si ha che $X_n \xrightarrow{\mathcal{L}} X$ per $n \rightarrow \infty$. In altre parole, la successione delle f.r. relative ai massimi ("centrati" con un opportuno parametro di posizione) di n v.a. indipendenti con legge Esponenziale ridotta converge alla f.r. relativa ad una legge di Gumbel ridotta. \square

I prossimi Teoremi risultano fondamentali per ottenere la convergenza in legge mediante la convergenza delle f.c. o delle f.g. In effetti, dal momento che la f.c. o la f.g. determinano in maniera univoca la f.r. di una v.a., è naturale aspettarsi che la convergenza in legge si possa stabilire tramite la successione delle f.c. o delle f.g.

Teorema 8.1.2. (Teorema di Lévy) Se $(X_n)_{n \geq 1}$ è una successione di v.a. a cui corrisponde la successione di f.c. $(\varphi_{X_n})_{n \geq 1}$ e se X è una v.a. con f.c. φ_X , allora $X_n \xrightarrow{\mathcal{L}} X$ per $n \rightarrow \infty$ se e solo se

$$\lim_n \varphi_{X_n}(t) = \varphi_X(t)$$

per ogni $t \in \mathbb{R}$.

Dimostrazione. Si veda Billingsley (1995, p.349). □

Si osservi che, nel caso in cui la v.a. X sia degenere e concentrata su c , allora si ha $\varphi_X(t) = e^{ict}$, e la condizione del Teorema di Lévy si riduce alla condizione

$$\lim_n \varphi_{X_n}(t) = e^{ict}$$

per ogni $t \in \mathbb{R}$.

• **Esempio 8.1.4.** Sia $(Z_n)_{n \geq 1}$ una successione di v.a. tale che la v.a. Z_n si distribuisce con legge Chi-quadrato χ_n^2 . Tenendo presente che dalla Sezione 6.8 risulta $E[Z_n] = n$ e $\text{Var}[Z_n] = 2n$, si consideri la successione di v.a. standardizzate $(X_n)_{n \geq 1}$, dove $X_n = \frac{Z_n - n}{\sqrt{2n}}$. Dunque, dall'Esempio 7.1.7 e dalla Proposizione 7.1.5 si ha

$$\log(\varphi_{X_n}(t)) = -\sqrt{\frac{n}{2}} it - \frac{n}{2} \log\left(1 - \sqrt{\frac{2}{n}} it\right).$$

Inoltre, dalle proprietà dell'espansione in serie della funzione logaritmica in campo complesso, per ogni $c \in \mathbb{C}$ tale che $|c| < 1$ risulta

$$\log(1 + c) = c - \frac{1}{2} c^2 + o(c^2),$$

da cui

$$\log\left(1 - \sqrt{\frac{2}{n}} it\right) = -\sqrt{\frac{2}{n}} it + \frac{t^2}{n} + o(n^{-1}).$$

Dunque, si ha

$$\log(\varphi_{X_n}(t)) = -\frac{t^2}{2} + no(n^{-1}),$$

da cui

$$\lim_n \varphi_{X_n}(t) = e^{-\frac{1}{2}t^2}.$$

Risulta immediato riconoscere che la precedente espressione è la f.c. relativa a una v.a. X con legge Normale ridotta $\mathcal{N}(0, 1)$ e perciò si ottiene infine che $X_n \xrightarrow{\mathcal{L}} X$ per $n \rightarrow \infty$. □

• **Esempio 8.1.5.** Analogamente all'Esempio 8.1.4, si consideri la successione di v.a. $(Z_n)_{n \geq 1}$, tale che la v.a. Z_n si distribuisce con legge Chi-quadrato χ_n^2 . Si consideri inoltre la successione di v.a. $(X_n)_{n \geq 1}$, dove $X_n = n^{-1}Z_n$. Tenendo presente l'Esempio 7.1.7, si ottiene

$$\varphi_{X_n}(t) = \left(1 - \frac{2it}{n}\right)^{-\frac{1}{2}n}.$$

Dunque, si ha

$$\lim_n \varphi_{X_n}(t) = e^{it}.$$

Dal momento che la precedente espressione fornisce la f.c. di una v.a. degenera X concentrata sul valore 1, risulta dunque $X_n \xrightarrow{\mathcal{L}} 1$ per $n \rightarrow \infty$. \square

• **Esempio 8.1.6.** Si consideri la successione di v.a. $(Z_n)_{n \geq 1}$ tale che la v.a. Z_n si distribuisce con legge di Poisson $\mathcal{P}(\lambda_n)$. Inoltre, $(\lambda_n)_{n \geq 1}$ è una successione per cui $\lambda_n \rightarrow \infty$ per $n \rightarrow \infty$. Dalla Sezione 6.2 si ha $E[Z_n] = \lambda_n$ e $\text{Var}[Z_n] = \lambda_n$, e quindi si consideri la successione di v.a. standardizzate $(X_n)_{n \geq 1}$, dove $X_n = \frac{Z_n - \lambda_n}{\sqrt{\lambda_n}}$. Dunque, dall'Esempio 7.1.2 e dalla Proposizione 7.1.5 si ha

$$\log(\varphi_{X_n}(t)) = \lambda_n e^{\frac{1}{\sqrt{\lambda_n}}it} - \lambda_n - \sqrt{\lambda_n} it.$$

Inoltre, dalle proprietà dell'espansione in serie della funzione esponenziale in campo complesso, per ogni $c \in \mathbb{C}$ risulta

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

da cui

$$e^{\frac{1}{\sqrt{\lambda_n}}it} = 1 + \frac{it}{\sqrt{\lambda_n}} - \frac{t^2}{2\lambda_n} + o(\lambda_n^{-1}).$$

Dunque, si ha

$$\log(\varphi_{X_n}(t)) = -\frac{t^2}{2} + \lambda_n o(\lambda_n^{-1}),$$

da cui

$$\lim_n \varphi_{X_n}(t) = e^{-\frac{1}{2}t^2}.$$

La precedente espressione fornisce la f.c. relativa a una v.a. X con legge Normale ridotta $\mathcal{N}(0, 1)$ e quindi si ha infine che $X_n \xrightarrow{\mathcal{L}} X$ per $n \rightarrow \infty$. \square

Teorema 8.1.3. Se $(X_n)_{n \geq 1}$ è una successione di v.a. discrete a valori su $S_n \subseteq \mathbb{N}$ a cui corrisponde la successione di f.g. $(G_{X_n})_{n \geq 1}$ e se X è un'ulteriore v.a. discreta a valori su $S \subseteq \mathbb{N}$ con f.g. G_X , allora $X_n \xrightarrow{\mathcal{L}} X$ per $n \rightarrow \infty$ se e solo se

$$\lim_n G_{X_n}(t) = G_X(t)$$

per ogni $t \in [0, 1]$.

Dimostrazione. Si veda Gut (2005, p.240). \square

• **Esempio 8.1.7.** Si consideri la successione di v.a. $(X_n)_{n \geq 1}$ tale che la v.a. X_n si distribuisce con legge Binomiale $\mathcal{B}(n, p_n)$. Inoltre, la successione $(p_n)_{n \geq 1}$ è tale che l'ulteriore successione $(np_n)_{n \geq 1}$ converge alla costante $\lambda \in]0, \infty[$. Tenendo presente l'Esempio 7.2.1, si ha

$$G_{X_n}(t) = (1 - p_n + p_n t)^n = \left(1 + \frac{np_n(t-1)}{n}\right)^n.$$

Dunque, si ha

$$\lim_n G_{X_n}(t) = e^{\lambda(t-1)}.$$

La precedente espressione fornisce la f.g. di una v.a. X distribuita con legge di Poisson $\mathcal{P}(\lambda)$ e dunque $X_n \xrightarrow{\mathcal{L}} X$ per $n \rightarrow \infty$. Questo risultato è in effetti già noto dalla Sezione 6.2. \square

• **Esempio 8.1.8.** Si consideri la successione di v.a. $(X_n)_{n \geq 1}$ tale che la v.a. X_n si distribuisce con legge Binomiale Negativa $\mathcal{BN}(n, p_n)$. Inoltre, la successione $(p_n)_{n \geq 1}$ è tale che l'ulteriore successione $(\frac{q_n}{p_n})_{n \geq 1}$, dove $q_n = 1 - p_n$, converge alla costante $\lambda \in]0, \infty[$. Tenendo presente l'Esempio 7.2.3, si ha

$$G_{X_n}(t) = \left(\frac{p_n}{1 - q_n t} \right)^n = \left(1 + \frac{n \frac{q_n}{p_n} (1 - t)}{n} \right)^{-n}.$$

Dunque, si ha

$$\lim_n G_{X_n}(t) = e^{\lambda(t-1)}.$$

Dal momento che la precedente espressione fornisce la f.g. di una v.a. X con legge di Poisson $\mathcal{P}(\lambda)$, si ha $X_n \xrightarrow{\mathcal{L}} X$ per $n \rightarrow \infty$. Questo risultato è in effetti già noto dalla Sezione 6.3. \square

Teorema 8.1.4. (Teorema della trasformata continua) Sia $(X_n)_{n \geq 1}$ una successione di v.a. e X una v.a. tale che $X_n \xrightarrow{\mathcal{L}} X$ per $n \rightarrow \infty$. Se $g: \mathbb{R} \rightarrow \mathbb{R}$ è una funzione continua, allora si ha $g(X_n) \xrightarrow{\mathcal{L}} g(X)$ per $n \rightarrow \infty$.

Dimostrazione. Posto $Y_n = g(X_n)$ e $Y = g(X)$, tenendo presente il Lemma di Fatou (Teorema A.7), si ha

$$\lim_n \varphi_{Y_n}(t) = \lim_n E[e^{itY_n}] = E[\lim_n e^{itY_n}] = E[e^{itY}] = \varphi_Y(t)$$

e la tesi segue dal Teorema 8.1.2. \square

• **Esempio 8.1.9.** Analogamente all'Esempio 8.1.4, si consideri la successione di v.a. standardizzate $(X_n)_{n \geq 1}$ dove $X_n = \frac{Z_n - n}{\sqrt{2n}}$, mentre la v.a. Z_n si distribuisce con legge Chi-quadrato χ_n^2 . Nel medesimo

Esempio è stato verificato che $X_n \xrightarrow{\mathcal{L}} X$ per $n \rightarrow \infty$, dove la v.a. X si distribuisce con legge Normale ridotta $\mathcal{N}(0, 1)$. Tenendo presente quanto detto nella Sezione 6.8, se si considera la funzione continua $g: x \mapsto x^2$, allora risulta $X_n^2 \xrightarrow{\mathcal{L}} Y$ per $n \rightarrow \infty$, dove la v.a. trasformata $Y = X^2$ si distribuisce con legge Chi-quadrato χ_1^2 . \square

8.2. Convergenza in probabilità

Un secondo concetto di convergenza, detto *convergenza in probabilità*, viene introdotto nella seguente definizione.

Definizione 8.2.1. Se $(X_n)_{n \geq 1}$ è una successione di v.a. definite sullo spazio probabilizzato (Ω, \mathcal{F}, P) , allora si dice che la successione *converge in probabilità* alla v.a. X , definita sul medesimo spazio probabilizzato, se per ogni $\varepsilon > 0$ si ha

$$\lim_n P(|X_n - X| > \varepsilon) = 0$$

e si scrive $X_n \xrightarrow{P} X$ per $n \rightarrow \infty$. □

In pratica, la definizione di convergenza in probabilità richiede che la probabilità dell'evento che si realizza quando le v.a. X_n e X differiscono tenda a 0 per $n \rightarrow \infty$. Nel caso in cui la v.a. X sia degenera e tale che $P(X = c) = 1$, allora la condizione della Definizione 8.2.1 si riduce a

$$\lim_n P(|X_n - c| > \varepsilon) = 0$$

e con un abuso in notazione quasi universalmente adottato nei testi di Teoria della Probabilità, la convergenza in probabilità viene denotata con $X_n \xrightarrow{P} c$ per $n \rightarrow \infty$. Infine, si osservi che nel linguaggio della Teoria della Misura questo tipo di convergenza è in effetti la cosiddetta convergenza in misura.

• **Esempio 8.2.1.** Data la v.a. X con legge Normale ridotta $\mathcal{N}(0, 1)$, si consideri la successione di v.a. $(X_n)_{n \geq 1}$ tale che $X_n = (1 - n^{-1})X$. Si vuole verificare che la successione di v.a. converge in probabilità alla v.a. X . Dal momento che

$$\begin{aligned} P(|X_n - X| > \varepsilon) &= 1 - P(|X_n - X| \leq \varepsilon) = 1 - P(|X| \leq n\varepsilon) \\ &= 1 - \int_{-n\varepsilon}^{n\varepsilon} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz = 2 - 2\Phi(n\varepsilon), \end{aligned}$$

allora si ha

$$\lim_n P(|X_n - X| > \varepsilon) = 2 - 2 \lim_n \Phi(n\varepsilon) = 0$$

e quindi $X_n \xrightarrow{P} X$ per $n \rightarrow \infty$. □

• **Esempio 8.2.2.** Data la v.a. Z con legge ridotta di Cauchy (si veda l'Esempio 4.1.4), si consideri la successione di v.a. $(X_n)_{n \geq 1}$ tale che $X_n = n^{-1}Z$. Si vuole verificare che la successione di v.a. converge in probabilità ad una v.a. degenera concentrata su 0. In effetti, si ha

$$\begin{aligned} P(|X_n| > \varepsilon) &= 1 - P(|X_n| \leq \varepsilon) = 1 - P(|Z| \leq n\varepsilon) \\ &= 1 - \int_{-n\varepsilon}^{n\varepsilon} \frac{1}{\pi(1+z^2)} dz = 1 - \frac{2}{\pi} \arctan(n\varepsilon), \end{aligned}$$

da cui

$$\lim_n P(|X_n| > \varepsilon) = 1 - \frac{2}{\pi} \lim_n \arctan(n\varepsilon) = 0$$

e quindi $X_n \xrightarrow{P} 0$ per $n \rightarrow \infty$. □

La seguente Proposizione fornisce la condizione sufficiente per la convergenza in probabilità ad una v.a. degenera.

Proposizione 8.2.2. Sia $(X_n)_{n \geq 1}$ una successione di v.a. definite sullo spazio probabilizzato (Ω, \mathcal{F}, P) . Se esiste un $\delta > 0$ per cui si ha

$$\lim_n E[|X_n - c|^\delta] = 0,$$

allora $X_n \xrightarrow{P} c$ per $n \rightarrow \infty$.

Dimostrazione. Dalla disuguaglianza di Markov (Teorema 4.3.6), per $\varepsilon > 0$ si ha

$$P(|X_n - c| > \varepsilon) = P(|X_n - c|^\delta > \varepsilon^\delta) \leq \frac{E[|X_n - c|^\delta]}{\varepsilon^\delta}$$

e quindi dall'ipotesi fatta risulta $\lim_n P(|X_n - c| > \varepsilon) = 0$. □

Dal momento che $E[(X_n - c)^2] = \text{Var}[X_n] + (E[X_n] - c)^2$, si osservi che per $\delta = 2$ la condizione data nella Proposizione 8.2.2 è equivalente alle condizioni $\lim_n E[X_n] = c$ e $\lim_n \text{Var}[X_n] = 0$.

• **Esempio 8.2.3.** (Legge dei Grandi Numeri di Bernoulli) Si consideri la successione di v.a. indipendenti $(X_n)_{n \geq 1}$ tale che la v.a. X_n si distribuisce con legge Binomiale $\mathcal{B}(1, p)$ per ogni n . Inoltre, si consideri la successione di v.a. $(Z_n)_{n \geq 1}$, dove $Z_n = n^{-1} S_n$ e $S_n = \sum_{k=1}^n X_k$. Essendo

$$E[Z_n] = E[X_n] = p,$$

allora

$$E[(Z_n - p)^2] = \text{Var}[Z_n] = \frac{1}{n} \text{Var}[X_n] = \frac{pq}{n},$$

da cui $\lim_n E[(Z_n - p)^2] = 0$. Dunque, si può concludere che $Z_n \xrightarrow{P} p$ per $n \rightarrow \infty$. Dal momento che $n^{-1} S_n$ rappresenta la percentuale di successi in n esperimenti indipendenti dicotomici con probabilità di successo pari a p (ovvero quando si considera lo schema delle prove ripetute, vedi Esempio 1.5.2), allora il precedente risultato stabilisce che la percentuale di successi converge alla probabilità di successo nel singolo esperimento. Questa particolare convergenza in probabilità costituisce in effetti la Legge Debole dei Grandi Numeri di Jakob Bernoulli e sarà analizzata nella sua forma generale in questa Sezione. □

La condizione della Proposizione 8.2.2 è sufficiente, ma non necessaria, per la convergenza in probabilità verso una v.a. degenere. In effetti, tale convergenza si può verificare senza che $E[(X_n - c)^2]$ converga verso zero. Ad esempio, questo può accadere con successioni di v.a. che non possiedono primo o secondo momento finito.

• **Esempio 8.2.4.** Si consideri la successione di v.a. discrete $(X_n)_{n \geq 1}$, tale che la v.a. X_n è a valori sull'insieme $S = \{0, n\}$ con legge essenziale $P(X_n = 0) = 1 - n^{-1}$ e $P(X_n = n) = n^{-1}$. La successione di v.a. $(X_n)_{n \geq 1}$ converge in probabilità ad una v.a. degenere concentrata sullo 0. In effetti, per $\varepsilon \geq n$ risulta $P(|X_n| > \varepsilon) = 0$, mentre per $\varepsilon < n$ si ha

$$P(|X_n| > \varepsilon) = 1 - P(X_n = 0) = \frac{1}{n}.$$

Dunque, si ha $\lim_n P(|X_n| > \varepsilon) = 0$, da cui $X_n \xrightarrow{P} 0$ per $n \rightarrow \infty$. Tuttavia, è immediato verificare che $E[X_n^2] = n$ e quindi $\lim_n E[X_n^2] = \infty$. Inoltre, si ha $E[X_n] = 1$, da cui ovviamente segue che

$$\lim_n E[X_n] = 1,$$

ovvero la successione delle medie $(E[X_n])_{n \geq 1}$ non converge verso 0. □

• **Esempio 8.2.5.** (Legge di Pareto) Si consideri la v.a. assolutamente continua Z che ammette d.p. data da

$$f_Z(x) = \alpha x^{-\alpha-1} \mathbf{1}_{[1, \infty)}(z),$$

dove $\alpha \in]0, \infty[$. Si consideri inoltre la successione di v.a. $(X_n)_{n \geq 1}$, dove $X_n = n^{-1}Z$. Si vuole verificare che per ogni α la successione di v.a. converge in probabilità ad una v.a. degenerare concentrata sullo 0. In effetti, si ha

$$P(|X_n| > \varepsilon) = 1 - P(|X_n| \leq \varepsilon) = 1 - P(|Z| \leq n\varepsilon) = 1 - \int_1^{n\varepsilon} \alpha z^{-\alpha-1} dz = (n\varepsilon)^{-\alpha},$$

da cui

$$\lim_n P(|X_n| > \varepsilon) = \lim_n (n\varepsilon)^{-\alpha} = 0$$

e quindi $X_n \xrightarrow{P} 0$ per $n \rightarrow \infty$. Tuttavia, per $\alpha \in]0, 2]$ risulta $\lim_n E[X_n^2] = \infty$. La legge è stata introdotta dall'economista Vilfredo Pareto (1848-1923). \square

Le prossime Proposizioni riguardano il legame esistente tra la convergenza in probabilità e la convergenza in distribuzione. Più esattamente, si evidenzia che la convergenza in probabilità implica la convergenza in distribuzione, mentre la proposizione inversa è vera quando si ha la convergenza in distribuzione verso una v.a. degenerare.

Proposizione 8.2.3. *Date una v.a. X e una successione di v.a. $(X_n)_{n \geq 1}$ definite sullo stesso spazio probabilitizzato (Ω, \mathcal{F}, P) , se $X_n \xrightarrow{P} X$ per $n \rightarrow \infty$ allora si ha $X_n \xrightarrow{\mathcal{L}} X$ per $n \rightarrow \infty$.*

Dimostrazione. Se $\varepsilon > 0$, allora risulta

$$\begin{aligned} F_{X_n}(x) &= P(X_n \leq x) = P(\{X_n \leq x\} \cap \{|X_n - X| \leq \varepsilon\}) + P(\{X_n \leq x\} \cap \{|X_n - X| > \varepsilon\}) \\ &\leq P(\{X \leq x + \varepsilon\} \cap \{|X_n - X| \leq \varepsilon\}) + P(|X_n - X| > \varepsilon) \\ &\leq P(X \leq x + \varepsilon) + P(|X_n - X| > \varepsilon). \end{aligned}$$

Dalla definizione di convergenza in probabilità si ha $\lim_n P(|X_n - X| > \varepsilon) = 0$ e dunque

$$\limsup_n F_{X_n}(x) \leq F_X(x + \varepsilon).$$

In modo analogo si ottiene

$$\liminf_n F_{X_n}(x) \geq F_X(x - \varepsilon).$$

Dal momento che ε è arbitrario, per ogni x per cui F_X è continua si ha

$$F_X(x) = \lim_{\varepsilon \rightarrow 0^+} F_X(x - \varepsilon) \leq \liminf_n F_{X_n}(x) \leq \limsup_n F_{X_n}(x) \leq \lim_{\varepsilon \rightarrow 0^+} F_X(x + \varepsilon) = F_X(x),$$

da cui segue $\lim_n F_{X_n}(x) = F_X(x)$, ovvero $X_n \xrightarrow{\mathcal{L}} X$ per $n \rightarrow \infty$. \square

Proposizione 8.2.4. *Data una successione di v.a. $(X_n)_{n \geq 1}$ definite sullo spazio probabilitizzato (Ω, \mathcal{F}, P) , allora $X_n \xrightarrow{P} c$ per $n \rightarrow \infty$ se e solo se $X_n \xrightarrow{\mathcal{L}} c$ per $n \rightarrow \infty$.*

Dimostrazione. Sulla base della Proposizione 8.2.3 si deve dimostrare solo l'implicazione inversa. Se $\varepsilon > 0$, si ha

$$\begin{aligned} P(|X_n - c| > \varepsilon) &= 1 - P(|X_n - c| \leq \varepsilon) = 1 - F_{X_n}(c + \varepsilon) + F_{X_n}(c - \varepsilon) - P(X_n = c - \varepsilon) \\ &\leq 1 - F_{X_n}(c + \varepsilon) + F_{X_n}(c - \varepsilon). \end{aligned}$$

Dal momento che $X_n \xrightarrow{\mathcal{L}} c$ per $n \rightarrow \infty$, ovvero $\lim_n F_{X_n}(x) = \mathbf{1}_{[c, \infty[}(x)$ per ogni $x \neq c$, allora si ha $\lim_n F_{X_n}(c + \varepsilon) = 1$ e $\lim_n F_{X_n}(c - \varepsilon) = 0$. Dunque, risulta

$$\lim_n (1 - F_{X_n}(c + \varepsilon) + F_{X_n}(c - \varepsilon)) = 0,$$

da cui segue che $X_n \xrightarrow{P} c$ per $n \rightarrow \infty$. □

• **Esempio 8.2.6.** Analogamente all'Esempio 8.1.5, si consideri la successione di v.a. $(Z_n)_{n \geq 1}$, tale che la v.a. Z_n si distribuisce con legge Chi-quadrato χ_n^2 , e la successione di v.a. $(X_n)_{n \geq 1}$, dove $X_n = n^{-1}Z_n$. Nel medesimo esempio è stato verificato che $X_n \xrightarrow{L} 1$ per $n \rightarrow \infty$ e dunque dalla Proposizione 8.2.4 si ha anche $X_n \xrightarrow{P} 1$ per $n \rightarrow \infty$. □

Si introducono di seguito alcuni risultati sulla convergenza per successioni di trasformate di v.a. In particolare, viene dato il cosiddetto Teorema di Cramér-Slutsky, che prende nome dallo statistico matematico svedese Harald Cramér (1893-1985) e dal probabilista ed economista russo Evgeny Evgenievich Slutsky (1880-1948).

Teorema 8.2.5. (Teorema della trasformata continua) *Date una v.a. X e una successione di v.a. $(X_n)_{n \geq 1}$ definite sullo stesso spazio probabilizzato (Ω, \mathcal{F}, P) , se $X_n \xrightarrow{P} X$ per $n \rightarrow \infty$ e $g : \mathbb{R} \rightarrow \mathbb{R}$ è una funzione continua, allora si ha $g(X_n) \xrightarrow{P} g(X)$ per $n \rightarrow \infty$.*

Dimostrazione. Sia $c > 0$ una costante per cui si ha $P(|X| > c) < \frac{1}{2}\delta$ con $\delta > 0$. Inoltre, dal momento che g è continua, allora per n abbastanza elevato esiste un $\delta = \delta(\varepsilon)$ tale che

$$\begin{aligned} P(|g(X_n) - g(X)| > \varepsilon) &= P(\{|g(X_n) - g(X)| > \varepsilon\} \cap \{|X| \leq c\}) \\ &\quad + P(\{|g(X_n) - g(X)| > \varepsilon\} \cap \{|X| > c\}) \\ &\leq P(|X_n - X| > \delta) + P(|X| > c) \leq \frac{\delta}{2} + \frac{\delta}{2} = \delta, \end{aligned}$$

da cui segue che $g(X_n) \xrightarrow{P} g(X)$ per $n \rightarrow \infty$. □

Teorema 8.2.6. (Teorema di Cramér-Slutsky) *Si considerino due successioni di v.a. $(X_n)_{n \geq 1}$ e $(Y_n)_{n \geq 1}$ definite sullo stesso spazio probabilizzato (Ω, \mathcal{F}, P) e la v.a. X definita sul medesimo spazio probabilizzato. Se $X_n \xrightarrow{L} X$ e $Y_n \xrightarrow{P} c$ per $n \rightarrow \infty$, allora si ha*

$$X_n + Y_n \xrightarrow{L} X + c,$$

$$X_n - Y_n \xrightarrow{L} X - c,$$

$$X_n Y_n \xrightarrow{L} cX,$$

$$\frac{X_n}{Y_n} \xrightarrow{L} \frac{X}{c}, c \neq 0,$$

per $n \rightarrow \infty$.

Dimostrazione. Si dimostra la prima relazione. Per $\varepsilon > 0$, si ha

$$\begin{aligned} F_{X_n+Y_n}(x) &= P(X_n + Y_n \leq x) \\ &= P(\{X_n + Y_n \leq x\} \cap \{|Y_n - c| \leq \varepsilon\}) + P(\{X_n + Y_n \leq x\} \cap \{|Y_n - c| > \varepsilon\}) \\ &\leq P(\{X_n \leq x - c + \varepsilon\} \cap \{|Y_n - c| \leq \varepsilon\}) + P(|Y_n - c| > \varepsilon) \\ &\leq P(X_n \leq x - c + \varepsilon) + P(|Y_n - c| > \varepsilon). \end{aligned}$$

Dalla definizione di convergenza in probabilità si ha $\lim_n P(|Y_n - c| > \varepsilon) = 0$ e dal momento che $X_n \xrightarrow{\mathcal{L}} X$, se $(x - c + \varepsilon)$ è un punto di continuità per F_X , si ha

$$\limsup_n F_{X_n+Y_n}(x) \leq F_X(x - c + \varepsilon).$$

In modo analogo, se $(x - c - \varepsilon)$ è un punto di continuità per F_X , si ottiene

$$\liminf_n F_{X_n+Y_n}(x) \geq F_X(x - c - \varepsilon).$$

Dal momento che ε può essere scelto in modo arbitrario e che l'insieme dei punti di discontinuità di F_X è al più numerabile (Proposizione 3.2.6), per ogni $(x - c)$ per cui F_X è continua si ha

$$\lim_n F_{X_n+Y_n}(x) = F_X(x - c),$$

ovvero $X_n + Y_n \xrightarrow{\mathcal{L}} X + c$ per $n \rightarrow \infty$. Le altre relazioni si verificano in modo analogo. \square

• **Esempio 8.2.7.** Si consideri la successione di v.a. (Z_n) , dove la v.a. Z_n si distribuisce con legge t di Student con n gradi di libertà. Tenendo presente la Sezione 6.10, la v.a. Z_n può essere rappresentata come $Z_n = \frac{X_n}{\sqrt{n^{-1}Y_n}}$, dove la v.a. X_n si distribuisce con legge Normale ridotta $\mathcal{N}(0, 1)$, mentre la v.a. Y_n si distribuisce con legge Chi-quadrato χ_n^2 . Dall'Esempio 8.2.6 si ha che $n^{-1}Y_n \xrightarrow{P} 1$ per $n \rightarrow \infty$. Inoltre, dal momento che la funzione $g : x \mapsto \sqrt{x}$ è continua, dal Teorema 8.2.5 si ha $\sqrt{n^{-1}Y_n} \xrightarrow{P} 1$ per $n \rightarrow \infty$. Infine, tenendo presente che banalmente si ha $X_n \xrightarrow{\mathcal{L}} Z$, dove la v.a. Z si distribuisce con legge Normale ridotta $\mathcal{N}(0, 1)$, dal Teorema 8.2.6 si ha infine che $Z_n \xrightarrow{\mathcal{L}} Z$ per $n \rightarrow \infty$. \square

Il seguente Teorema fornisce un celebre risultato, ovvero la cosiddetta Legge Debole dei Grandi Numeri, nella sua forma più comune.

Teorema 8.2.7. (Legge Debole dei Grandi Numeri) Sia $(X_n)_{n \geq 1}$ una successione di v.a. non correlate definite sullo spazio probabilizzato (Ω, \mathcal{F}, P) , tali che per ogni n si ha $E[X_n] = \mu$ e $\text{Var}[X_n] < c < \infty$. Data l'ulteriore successione di v.a. $(Z_n)_{n \geq 1}$, dove $Z_n = n^{-1}S_n$ e $S_n = \sum_{k=1}^n X_k$, allora si ha $Z_n \xrightarrow{P} \mu$ per $n \rightarrow \infty$.

Dimostrazione. Dal momento che

$$E[Z_n] = \frac{1}{n} \sum_{k=1}^n E[X_k] = \mu$$

e, dal momento che le v.a. della successione sono non correlate, si ha

$$\text{Var}[Z_n] = \frac{1}{n^2} \sum_{k=1}^n \text{Var}[X_k] < \frac{c}{n}.$$

Inoltre, dalla disuguaglianza di Chebyshev (Teorema 4.3.7), per $\varepsilon > 0$ si ha

$$P(|Z_n - \mu| > \varepsilon) \leq \frac{\text{Var}[Z_n]}{\varepsilon^2} < \frac{c}{n\varepsilon^2}.$$

Dunque, risulta

$$\lim_n P(|Z_n - \mu| > \varepsilon) = 0,$$

ovvero $Z_n \xrightarrow{P} \mu$ per $n \rightarrow \infty$. \square

Un'ulteriore versione fondamentale della Legge Debole dei Grandi Numeri è stata ottenuta in un approccio generale dal probabilista russo Aleksandr Yakovlevich Khinchin (1894-1959) e viene data di seguito.

Teorema 8.2.8. (Legge Debole dei Grandi Numeri di Khinchin) Sia $(X_n)_{n \geq 1}$ una successione di v.a. indipendenti con la medesima legge definite sullo spazio probabilizzato (Ω, \mathcal{F}, P) , tali che per ogni n si ha $E[X_n] = \mu$. Data l'ulteriore successione di v.a. $(Z_n)_{n \geq 1}$, dove $Z_n = n^{-1}S_n$ e $S_n = \sum_{k=1}^n X_k$, allora si ha $Z_n \xrightarrow{P} \mu$ per $n \rightarrow \infty$.

Dimostrazione. Dalla Proposizione 7.1.7 si ha

$$\varphi_{X_n}(t) = 1 + i\mu t + o(t) .$$

Inoltre, dalle Proposizioni 7.3.6 e 7.1.5 si ha

$$\varphi_{Z_n}(t) = \left(1 + \frac{i\mu t}{n} + o(n^{-1}) \right)^n ,$$

da cui

$$\lim_n \varphi_{X_n}(t) = e^{i\mu t} .$$

Dunque, tenendo presente il Teorema 8.1.2 si ha $Z_n \xrightarrow{\mathcal{L}} \mu$ per $n \rightarrow \infty$, ovvero dalla Proposizione 8.2.4 si ottiene infine che $Z_n \xrightarrow{P} \mu$ per $n \rightarrow \infty$. □

8.3. Convergenza quasi certa

Un terzo concetto di convergenza, detto *convergenza quasi certa*, viene introdotto nella seguente definizione.

Definizione 8.3.1. Se $(X_n)_{n \geq 1}$ è una successione di v.a. definite sullo spazio probabilizzato (Ω, \mathcal{F}, P) , allora si dice che la successione *converge quasi certamente* alla v.a. X , definita sul medesimo spazio probabilizzato, se l'evento

$$\{\omega \in \Omega : \lim_n X_n(\omega) = X(\omega)\} = \{\lim_n X_n = X\}$$

si verifica *q.c.*, ovvero se

$$P(\lim_n X_n = X) = 1 ,$$

e si scrive $X_n \xrightarrow{q.c.} X$ per $n \rightarrow \infty$. □

Nel caso in cui la v.a. X sia degenere e tale che $P(X = c) = 1$, allora la condizione della Definizione 8.3.1 si riduce a

$$P(\lim_n X_n = c) = 1 ,$$

e con il solito abuso in notazione, la convergenza quasi certa viene denotata con $X_n \xrightarrow{q.c.} c$ per $n \rightarrow \infty$. Nel linguaggio della Teoria della Misura questo tipo di convergenza è in effetti la cosiddetta convergenza puntuale. Dal momento che la misura dello spazio fondamentale è finita, in Teoria della Probabilità la convergenza puntuale coincide con la convergenza uniforme.

Risulta inoltre interessante notare che, dalla definizione di limite, l'evento $\{\lim_n X_n = X\}$ si verifica se per ogni intero m esiste un intero n tale che per ogni $k \geq n$ si verifica l'evento

$\{|X_k - X| \leq m^{-1}\}$. Dunque, anche tenendo presente la definizione di limite superiore di una successione di eventi, si ha

$$\{\lim_n X_n = X\} = \bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} \{|X_k - X| \leq m^{-1}\} = \lim_m \limsup_n \{|X_n - X| \leq m^{-1}\},$$

ovvero, dal momento che $(\limsup_n \{|X_n - X| \leq m^{-1}\})_{m \geq 1}$ è una successione decrescente di eventi e in base al Teorema 2.2.10, la condizione della Definizione 8.3.1 è equivalente a

$$\begin{aligned} P(\lim_n X_n = X) &= P\left(\bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} \{|X_k - X| \leq m^{-1}\}\right) \\ &= \lim_m P(\limsup_n \{|X_n - X| \leq m^{-1}\}) = 1. \end{aligned}$$

La seguente Proposizione stabilisce una condizione necessaria e sufficiente affinché si verifichi la convergenza quasi certa.

Proposizione 8.3.2. *Date una v.a. X e una successione di v.a. $(X_n)_{n \geq 1}$ definite sullo stesso spazio probabilizzato (Ω, \mathcal{F}, P) , allora $X_n \xrightarrow{q.c.} X$ per $n \rightarrow \infty$ se e solo se per ogni $\varepsilon > 0$ si ha*

$$\lim_n P\left(\bigcap_{k=n}^{\infty} \{|X_k - X| \leq \varepsilon\}\right) = 1,$$

o equivalentemente che

$$\lim_n P\left(\bigcup_{k=n}^{\infty} \{|X_k - X| > \varepsilon\}\right) = 0.$$

Dimostrazione. Posto

$$F_{n,m} = \bigcap_{k=n}^{\infty} \{|X_k - X| \leq m^{-1}\},$$

si osservi che la successione di eventi $(F_{n,m})_{n \geq 1}$ è crescente per ogni intero m . Inoltre, la condizione della Definizione 8.3.1 è verificata se e solo se per ogni intero m si ha

$$P\left(\bigcup_{n=1}^{\infty} F_{n,m}\right) = 1.$$

Tenendo presente la definizione di limite di successione crescente di eventi, dal Teorema 2.2.10 si ottiene inoltre

$$P\left(\bigcup_{n=1}^{\infty} F_{n,m}\right) = P(\lim_n F_{n,m}) = \lim_n P(F_{n,m}),$$

da cui segue la prima parte. Inoltre, dalla relazione di De Morgan si ha

$$F_{n,m} = \left(\bigcup_{k=n}^{\infty} \{|X_k - X| > m^{-1}\}\right)^c,$$

da cui segue la seconda parte. □

• **Esempio 8.3.1.** Si consideri la successione di v.a. $(X_n)_{n \geq 1}$ dell'Esempio 8.2.1. Dal momento che si ha

$$P\left(\bigcup_{k=n}^{\infty} \{|X_k - X| > \varepsilon\}\right) = P\left(\bigcup_{k=n}^{\infty} \{|X| > k\varepsilon\}\right) = P(|X| > n\varepsilon) = 2 - 2\Phi(n\varepsilon)$$

allora risulta

$$\lim_n P\left(\bigcup_{k=n}^{\infty} \{|X_k - X| > \varepsilon\}\right) = 2 - 2 \lim_n \Phi(n\varepsilon) = 0$$

e quindi $X_n \xrightarrow{q.c.} X$ per $n \rightarrow \infty$. □

• **Esempio 8.3.2.** Si consideri la successione di v.a. $(X_n)_{n \geq 1}$ dell'Esempio 8.2.2. Dal momento che si ha

$$P\left(\bigcup_{k=n}^{\infty} \{|X_k| > \varepsilon\}\right) = P\left(\bigcup_{k=n}^{\infty} \{|Z| > k\varepsilon\}\right) = P(|Z| > n\varepsilon) = 1 - \frac{2}{\pi} \arctan(n\varepsilon),$$

allora risulta

$$\lim_n P\left(\bigcup_{k=n}^{\infty} \{|X_k| > \varepsilon\}\right) = 1 - \frac{2}{\pi} \lim_n \arctan(n\varepsilon) = 0$$

e quindi $X_n \xrightarrow{q.c.} 0$ per $n \rightarrow \infty$. □

Dalla Proposizione 8.3.2 si intuisce come la convergenza quasi certa implichi una condizione più forte di quella richiesta nella convergenza in probabilità. In effetti, vale la seguente Proposizione.

Proposizione 8.3.3. *Date una v.a. X e una successione di v.a. $(X_n)_{n \geq 1}$ definite sullo stesso spazio probabilizzato (Ω, \mathcal{F}, P) , se $X_n \xrightarrow{q.c.} X$ per $n \rightarrow \infty$ allora si ha $X_n \xrightarrow{P} X$ per $n \rightarrow \infty$.*

Dimostrazione. Dato $\varepsilon > 0$, per ogni n si ha

$$P(|X_n - X| > \varepsilon) \leq P\left(\bigcup_{k=n}^{\infty} \{|X_k - X| > \varepsilon\}\right).$$

Dunque, se $X_n \xrightarrow{q.c.} X$ per $n \rightarrow \infty$, dalla Proposizione 8.3.2 e dalla precedente relazione si ha $\lim_n P(|X_n - X| > \varepsilon) = 0$, ovvero $X_n \xrightarrow{P} X$ per $n \rightarrow \infty$. □

• **Esempio 8.3.3.** Si consideri la successione di v.a. discrete $(X_n)_{n \geq 1}$ dell'Esempio 8.2.4, assumendo l'indipendenza delle v.a. Evidentemente, anche in questo caso si ha $X_n \xrightarrow{P} 0$ per $n \rightarrow \infty$. Dunque, in base alla Proposizione 8.3.3, se vi è convergenza quasi certa la successione deve necessariamente convergere verso una v.a. degenerare concentrata sullo 0. Tuttavia, per ogni $\varepsilon > 0$ si ha $P(|X_n| > \varepsilon) = n^{-1}$, e dal momento che la serie armonica non converge risulta

$$P\left(\bigcup_{n=1}^{\infty} \{|X_n| > \varepsilon\}\right) = \sum_{n=1}^{\infty} P(|X_n| > \varepsilon) = \sum_{n=1}^{\infty} \frac{1}{n} = \infty.$$

Dunque, dal Lemma di Borel-Cantelli (Teorema 2.7.1) si ottiene che $P(\limsup_n |X_n| > \varepsilon) = 1$, ovvero la successione non può convergere quasi certamente a 0. □

La seguente Proposizione fornisce condizioni sufficienti per la convergenza quasi certa verso una v.a. degenerare.

Proposizione 8.3.4. *Data una successione di v.a. $(X_n)_{n \geq 1}$ definite sullo stesso spazio probabilizzato (Ω, \mathcal{F}, P) , allora $X_n \xrightarrow{q.c.} c$ per $n \rightarrow \infty$ se per ogni ε si ha*

$$\sum_{n=1}^{\infty} P(|X_n - c| > \varepsilon) < \infty.$$

Alternativamente, $X_n \xrightarrow{q.c.} c$ per $n \rightarrow \infty$ se esiste un $\delta > 0$ per cui si ha

$$\sum_{n=1}^{\infty} E[|X_n - c|^\delta] < \infty.$$

Dimostrazione. Dalla disuguaglianza di Bonferroni (Teorema 2.2.8), segue che

$$P\left(\bigcup_{k=n}^{\infty} \{|X_k - c| > \varepsilon\}\right) \leq \sum_{k=n}^{\infty} P(|X_k - c| > \varepsilon).$$

Dal momento che l'ipotesi di convergenza della serie $\sum_{n=1}^{\infty} P(|X_n - c| > \varepsilon)$ implica

$$\lim_n \sum_{k=n}^{\infty} P(|X_k - c| > \varepsilon) = 0,$$

segue la prima parte. Tenendo presente la disuguaglianza di Markov (Teorema 4.3.6), si ha inoltre

$$\sum_{k=n}^{\infty} P(|X_k - c| > \varepsilon) = \sum_{k=n}^{\infty} P(|X_k - c|^\delta > \varepsilon^\delta) \leq \frac{1}{\varepsilon^\delta} \sum_{k=n}^{\infty} E[|X_k - c|^\delta].$$

Dal momento che l'ipotesi di convergenza della serie $\sum_{n=1}^{\infty} E[|X_n - c|^\delta]$ implica

$$\lim_n \sum_{k=n}^{\infty} E[|X_k - c|^\delta] = 0,$$

allora si ha la seconda parte. □

• **Esempio 8.3.4.** In modo simile all'Esempio 8.2.3, si consideri la successione di v.a. indipendenti $(X_n)_{n \geq 1}$ tale che la v.a. X_n si distribuisce con legge Binomiale $\mathcal{B}(1, p)$ per ogni n . Inoltre, si consideri l'ulteriore successione di v.a. $(Z_n)_{n \geq 1}$, dove $Z_n = n^{-1}S_n$ e $S_n = \sum_{k=1}^n X_k$. Tenendo presente la seconda condizione della Proposizione 8.3.4, posto $\delta = 4$, dal momento che S_n si distribuisce con legge Binomiale $\mathcal{B}(n, p)$ si ha

$$E[(Z_n - p)^4] = \frac{1}{n^4} E[(S_n - np)^4] = \frac{3p^2q^2}{n^2} + \frac{pq(1 - 6pq)}{n^3},$$

dove l'ultima relazione è stata ottenuta attraverso l'applicazione laboriosa, ma ovvia, della Proposizione 7.1.6. Dunque, dalla definizione della funzione Zeta di Riemann (si veda Esempio 7.2.7), risulta

$$\sum_{n=1}^{\infty} E[(Z_n - p)^4] = 3p^2q^2\zeta(2) + pq(1 - 6pq)\zeta(3) < \infty,$$

dove $\zeta(2) = \frac{\pi^2}{6}$ e $\zeta(3) < \zeta(4) = \frac{\pi^4}{90}$. Dunque, si può concludere che $Z_n \xrightarrow{q.c.} p$ per $n \rightarrow \infty$. Questa particolare convergenza quasi certa costituisce un caso particolare della Legge Forte dei Grandi Numeri. \square

Il seguente Teorema fornisce la cosiddetta Legge Forte dei Grandi Numeri, nella forma introdotta da Andrej Kolmogorov.

Teorema 8.3.5. (Legge Forte dei Grandi Numeri di Kolmogorov) Sia $(X_n)_{n \geq 1}$ una successione di v.a. indipendenti con la medesima legge definita sullo spazio probabilizzato (Ω, \mathcal{F}, P) , tali che per ogni n si ha $E[X_n] = \mu$. Data l'ulteriore successione di v.a. $(Z_n)_{n \geq 1}$, dove $Z_n = n^{-1}S_n$ e $S_n = \sum_{k=1}^n X_k$, allora si ha $Z_n \xrightarrow{q.c.} \mu$ per $n \rightarrow \infty$.

Dimostrazione. Si veda Billingsley (1995, p.85). \square

8.4. Convergenza in media

Un ulteriore concetto di convergenza, che è usualmente detto *convergenza in media di ordine p* (o semplicemente *in media quadratica* quando $p = 2$), viene definito di seguito. La convergenza in media sarà centrale nello sviluppo della teoria dell'integrazione stocastica presentata nel Capitolo 10.

Definizione 8.4.1. Se $(X_n)_{n \geq 1}$ è una successione di v.a. definite sullo spazio probabilizzato (Ω, \mathcal{F}, P) e se $p \in]0, \infty[$, allora si dice che la successione *converge in media di ordine p* alla v.a. X , definita sul medesimo spazio probabilizzato, se si ha

$$\lim_n E[|X_n - X|^p] = 0$$

e si scrive $X_n \xrightarrow{L^p} X$ per $n \rightarrow \infty$. \square

La convergenza in media di ordine p risulta effettivamente equivalente alla convergenza in norma di ordine p considerata nella Teoria della Misura. Questo è il motivo per cui questo tipo di convergenza viene indicata con il simbolo L^p (si veda l'Appendice A). Nel caso in cui la v.a. X sia degenere e tale che $P(X = c) = 1$, allora la condizione della Definizione 8.4.1 si riduce a

$$\lim_n E[|X_n - c|^p] = 0$$

e si scrive $X_n \xrightarrow{L^p} c$ per $n \rightarrow \infty$ con il solito abuso in notazione.

• **Esempio 8.4.1.** Si consideri una successione $(X_n)_{n \geq 1}$ di v.a. indipendenti con la medesima legge e tali che per ogni n si ha $E[X_n] = \mu$ e $\text{Var}[X_n] = \sigma^2 < \infty$. Si consideri inoltre l'ulteriore successione di v.a. $(Z_n)_{n \geq 1}$ con $Z_n = n^{-1}S_n$ e $S_n = \sum_{k=1}^n X_k$. Dal momento che $E[S_n] = n\mu$ e $\text{Var}[S_n] = n\sigma^2$, allora si ha

$$\lim_n E[(Z_n - \mu)^2] = \lim_n \frac{E[(S_n - n\mu)^2]}{n^2} = \lim_n \frac{\text{Var}[S_n]}{n^2} = \lim_n \frac{\sigma^2}{n} = 0.$$

Dunque, risulta $Z_n \xrightarrow{L^2} \mu$ per $n \rightarrow \infty$. Inoltre, dal momento che si ha

$$E[|Z_n - \mu|] \leq E[(Z_n - \mu)^2]^{\frac{1}{2}}$$

per la disuguaglianza di Lyapunov (Teorema 4.3.2), risulta anche $Z_n \xrightarrow{L^1} \mu$ per $n \rightarrow \infty$. \square

La convergenza in media di ordine p implica la convergenza in probabilità. Questo risultato era già stato in effetti evidenziato nella Proposizione 8.2.2 nel caso di convergenza ad una variabile degenera e viene generalizzato nella seguente Proposizione.

Proposizione 8.4.2. *Date una v.a. X e una successione di v.a. $(X_n)_{n \geq 1}$ definite sullo stesso spazio probabilizzato (Ω, \mathcal{F}, P) , se esiste un $p \in]0, \infty[$ per cui si ha $X_n \xrightarrow{L^p} X$ per $n \rightarrow \infty$ allora risulta $X_n \xrightarrow{P} X$ per $n \rightarrow \infty$.*

Dimostrazione. Dalla disuguaglianza di Markov (Teorema 4.3.6), per $\varepsilon > 0$ si ha

$$P(|X_n - X| > \varepsilon) = P(|X_n - X|^p > \varepsilon^p) \leq \frac{E[|X_n - X|^p]}{\varepsilon^p}$$

e, dal momento che dalle ipotesi fatte si ha $\lim_n E[|X_n - X|^p] = 0$, allora risulta anche $\lim_n P(|X_n - X| > \varepsilon) = 0$, ovvero $X_n \xrightarrow{P} X$ per $n \rightarrow \infty$. \square

Per quanto riguarda la relazione fra convergenza quasi certa e convergenza in media di ordine p , in generale non si possono fare affermazioni. In effetti, ci sono casi in cui vi è convergenza quasi certa, ma non quella in media di ordine p e viceversa. I seguenti esempi evidenziano queste situazioni.

• **Esempio 8.4.2.** Si consideri una v.a. Z con legge Uniforme su $]0, 1[$ e la successione $(X_n)_{n \geq 1}$ dove $X_n = 2^n \mathbf{1}_{]0, \frac{1}{n}[}(Z)$. Evidentemente la v.a. X_n è discreta ed è definita su $\{0, 2^n\}$ con probabilità $P(X_n = 0) = 1 - n^{-1}$ e $P(X_n = 2^n) = n^{-1}$. Tenendo presente la Proposizione 8.3.2 e dal momento che per ogni $k > n$ si ha $\{|X_k| > \varepsilon\} \subset \{|X_n| > \varepsilon\}$ q.c., allora per $\varepsilon > 0$ risulta

$$\lim_n P\left(\bigcup_{k=n}^{\infty} \{|X_k| > \varepsilon\}\right) = \lim_n P(X_n > \varepsilon) = \lim_n \frac{1}{n} = 0,$$

ovvero $X_n \xrightarrow{q.c.} 0$. Tuttavia, dal momento che

$$\lim_n E[|X_n|^p] = \lim_n \frac{2^{np}}{n} = \infty,$$

non vi è convergenza in media per nessun valore di p . \square

• **Esempio 8.4.3.** Si consideri una v.a. Z con legge Uniforme su $]0, 1[$ e sia data inoltre la successione $(X_n)_{n \geq 1}$ dove $X_n = \mathbf{1}_{[j2^{-m}, (j+1)2^{-m}[}(Z)$ con $n = 2^m + j$, $j = \{0, 1, \dots, 2^m - 1\}$ e $m \in \mathbb{N}$. Dunque, $X_1 = \mathbf{1}_{]0, 1[}(Z)$, $X_2 = \mathbf{1}_{]0, \frac{1}{2}[}(Z)$, $X_3 = \mathbf{1}_{]0, \frac{1}{4}[}(Z)$, $X_4 = \mathbf{1}_{]0, \frac{1}{8}[}(Z)$, $X_5 = \mathbf{1}_{]0, \frac{1}{4}[}(Z)$, $X_6 = \mathbf{1}_{]0, \frac{3}{4}[}(Z)$, $X_7 = \mathbf{1}_{]0, \frac{3}{4}[}(Z)$ e così via. In questo caso, X_1 è una v.a. degenera tale che $P(X_1 = 1) = 1$, mentre X_n è una v.a. di Bernoulli di parametro $P(X_n = 1) = 2^{-m}$ per $n = 2^m + j > 1$. Dal momento che

$$\lim_n E[|X_n|^p] = \lim_m 2^{-mp} = 0,$$

allora $X_n \xrightarrow{L^p} 0$ per ogni valore di p . Tuttavia, dalla Proposizione 8.3.2, si ha

$$\lim_n P\left(\bigcup_{k=n}^{\infty} \{|X_k| > \varepsilon\}\right) = \lim_n P(X_1 = 1) = 1,$$

ovvero la successione non converge quasi certamente a 0. \square

La prossima Proposizione fornisce una condizione per la quale la convergenza quasi certa implica la convergenza in media di ordine p .

Proposizione 8.4.3. *Date una v.a. X e una successione di v.a. $(X_n)_{n \geq 1}$ definite sullo stesso spazio probabilizzato (Ω, \mathcal{F}, P) , se $X_n \xrightarrow{q.c.} X$ per $n \rightarrow \infty$ e $\{|X_n| \leq Y\}$ q.c. per una v.a. Y tale che $E[Y^p] < \infty$, allora si ha $X_n \xrightarrow{L^p} X$ per $n \rightarrow \infty$.*

Dimostrazione. Dalle assunzioni fatte si ha $E[|X_n|^p] \leq E[Y^p] < \infty$. Inoltre, dal momento che $X_n \xrightarrow{q.c.} X$, allora risulta anche $\{|X| \leq Y\}$ q.c. Dal momento che $\{|X_n - X| \leq |X_n| + |X|\}$ q.c., si ha anche $\{|X_n - X| \leq 2Y\}$ q.c. Quindi, per ogni $\varepsilon > 0$ si ottiene

$$\begin{aligned} E[|X_n - X|^p] &= E[|X_n - X|^p \mathbf{1}_{[0, \varepsilon]}(|X_n - X|)] + E[|X_n - X|^p \mathbf{1}_{[\varepsilon, \infty]}(|X_n - X|)] \\ &\leq \varepsilon^p + 2^p E[Y^p \mathbf{1}_{[\varepsilon, \infty]}(|X_n - X|)]. \end{aligned}$$

Tenendo ancora presente che $X_n \xrightarrow{q.c.} X$, allora $\lim_n E[Y^p \mathbf{1}_{[\varepsilon, \infty]}(|X_n - X|)] = 0$. Dunque, essendo ε arbitrario, dalla precedente relazione segue che $\lim_n E[|X_n - X|^p] = 0$. \square

8.5. Teoremi limite

Si considerano di seguito alcuni teoremi di convergenza che hanno importanti aspetti applicativi in ambiti come la statistica inferenziale. Questi classici risultati sono comunemente denominati come Teoremi Centrali del Limite (dove centrale va inteso nel senso di fondamentale) sulla base di una traduzione leggermente impropria dalla terminologia inglese, ma di uso comune nella letteratura italiana.

Teorema 8.5.1. (Teorema Centrale del Limite di Lindeberg-Lévy) *Sia $(X_n)_{n \geq 1}$ una successione di v.a. indipendenti con la medesima legge e definite sullo spazio probabilizzato (Ω, \mathcal{F}, P) , tali che per ogni n si ha $E[X_n] = \mu$ e $\text{Var}[X_n] = \sigma^2 < \infty$. Si consideri inoltre la successione di v.a. $(Z_n)_{n \geq 1}$ con*

$$Z_n = \frac{S_n - E[S_n]}{\text{Var}[S_n]} = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$

e $S_n = \sum_{k=1}^n X_k$. Se la v.a. Z si distribuisce con legge Normale $\mathcal{N}(0, 1)$, allora $Z_n \xrightarrow{\mathcal{L}} Z$ per $n \rightarrow \infty$.

Dimostrazione. Posto $Y_k = \frac{X_k - \mu}{\sigma}$, si ha $E[Y_k] = 0$ e $\text{Var}[Y_k] = 1$. Sulla base della Proposizione 7.1.6, per $t \rightarrow 0$ la f.c. della v.a. Y_k può essere espressa come

$$\varphi_{Y_k}(t) = 1 + it E[Y_k] + \frac{(it)^2}{2} E[Y_k^2] + o(t^2) = 1 - \frac{t^2}{2} + o(t^2).$$

Dal momento che

$$Z_n = \frac{1}{\sqrt{n}} \sum_{k=1}^n Y_k,$$

allora si ottiene

$$\varphi_{Z_n}(t) = \varphi_{Y_k} \left(\frac{t}{\sqrt{n}} \right)^n = \left(1 - \frac{t^2}{2n} + o(n^{-1}) \right)^n.$$

Dunque, si ha

$$\lim_n \varphi_{Z_n}(t) = e^{-\frac{1}{2}t^2},$$

ovvero dal Teorema di Lévy (Teorema 8.1.2) si ottiene la tesi. \square

Questa versione di Teorema Centrale del Limite è stata appunto introdotta da Paul Lévy e dal probabilista finlandese Jarl Waldemar Lindeberg (1876-1932). Il Teorema 8.5.1 contiene come caso particolare la versione primitiva del Teorema Centrale del Limite data da Abraham de Moivre, nella quale si assume che $(X_n)_{n \geq 1}$ sia una successione di v.a. indipendenti con la medesima legge e tale che ogni v.a. X_n si distribuisce con legge di Bernoulli $\mathcal{B}(1, p)$. Dal momento che $S_n = \sum_{k=1}^n X_k$ si distribuisce con legge Binomiale $\mathcal{B}(n, p)$ (si veda l'Esempio 7.4.3), il Teorema Centrale del Limite dato di De Moivre fornisce la convergenza in legge della successione di v.a. standardizzate $(Z_n)_{n \geq 1}$, dove $Z_n = \frac{S_n - np}{\sqrt{npq}}$, ad una v.a. con legge $\mathcal{N}(0, 1)$ per $n \rightarrow \infty$.

La seguente versione di Teorema Centrale del Limite può essere applicata a successioni di v.a. indipendenti con differenti medie e varianze. Il Teorema, oltre che da Paul Lévy e da Jarl Lindeberg, prende nome dal probabilista croato Vilibald (William) Srečko Feller (1906-1970).



Figure 8.5.1. Vilibald (William) Srečko Feller (1906-1970).

Teorema 8.5.2. (Teorema Centrale del Limite di Lindeberg-Lévy-Feller) Sia $(X_n)_{n \geq 1}$ una successione di v.a. indipendenti definite sullo spazio probabilizzato (Ω, \mathcal{F}, P) , tali che $E[X_n] = \mu_n$ e $\text{Var}[X_n] = \sigma_n^2 < \infty$ per ogni n . Posto $v_n^2 = \sum_{k=1}^n \sigma_k^2$, si consideri inoltre la successione di v.a. $(Z_n)_{n \geq 1}$ con

$$Z_n = \frac{1}{v_n} \sum_{k=1}^n (X_k - \mu_k).$$

Se per ogni $\varepsilon > 0$ si ha

$$\lim_n \frac{1}{v_n^2} \sum_{k=1}^n E[(X_k - \mu_k)^2 \mathbf{1}_{\varepsilon v_n, \infty}(|X_k - \mu_k|)] = 0,$$

allora $Z_n \xrightarrow{\mathcal{L}} Z$ per $n \rightarrow \infty$, dove la v.a. Z si distribuisce con legge Normale $\mathcal{N}(0, 1)$.

Dimostrazione. Si veda Billingsley (1995, p.359). □

La prossima versione di Teorema Centrale del Limite è dovuta a Lyapunov. Nelle medesime ipotesi del Teorema Centrale del Limite dato da Lindeberg-Lévy-Feller, questa versione richiede una condizione più forte, ma spesso più facile da verificare. Si tenga comunque presente che la condizione del Teorema Centrale del Limite dato da Lindeberg-Lévy-Feller è necessaria e sufficiente.

Teorema 8.5.3. (Teorema Centrale del Limite di Lyapunov) Sia $(X_n)_{n \geq 1}$ una successione di v.a. indipendenti definite sullo spazio probabilizzato (Ω, \mathcal{F}, P) , tali che $E[X_n] = \mu_n$ e $\text{Var}[X_n] = \sigma_n^2 < \infty$ per ogni n . Posto $v_n^2 = \sum_{k=1}^n \sigma_k^2$, si consideri inoltre la successione di v.a. $(Z_n)_{n \geq 1}$ con

$$Z_n = \frac{1}{v_n} \sum_{k=1}^n (X_k - \mu_k).$$

Se esiste un $\delta > 2$ tale che $E[|X_n|^\delta] < \infty$ per ogni n e che

$$\lim_n \frac{1}{v_n^{\frac{\delta}{2}}} \sum_{k=1}^n E[|X_k - \mu_k|^\delta] = 0,$$

allora $Z_n \xrightarrow{\mathcal{L}} Z$ per $n \rightarrow \infty$, dove la v.a. Z si distribuisce con legge Normale $\mathcal{N}(0, 1)$.

Dimostrazione. Si veda Billingsley (1995, p.362). □

• **Esempio 8.5.1.** Si consideri la successione di v.a. indipendenti $(X_n)_{n \geq 1}$ tale che la v.a. X_n si distribuisce con legge Binomiale $\mathcal{B}(1, p_n)$. Risulta dunque $v_n^2 = \sum_{k=1}^n p_k q_k$. Inoltre, dal momento che $\{|X_k - p_k| \leq 1\}$ q.c., segue che

$$E[|X_k - p_k|^3] = E[(X_k - p_k)^2 |X_k - p_k|] \leq E[(X_k - p_k)^2] = p_k q_k.$$

Dunque, considerando il Teorema 8.5.3 con $\delta = 3$, si ha

$$\frac{1}{\sqrt{v_n^3}} \sum_{k=1}^n E[|X_k - \mu_k|^3] \leq \frac{1}{\sqrt{v_n^3}} \sum_{k=1}^n p_k q_k = \frac{1}{\sqrt{v_n}}$$

e la relativa condizione è dunque soddisfatta se $\lim_n v_n = \infty$. In questo caso, se si considera la successione di v.a. $(Z_n)_{n \geq 1}$, con

$$Z_n = \frac{1}{v_n} \sum_{k=1}^n (X_k - p_k) = \frac{S_n - \sum_{k=1}^n p_k}{\sqrt{\sum_{k=1}^n p_k q_k}}$$

e $S_n = \sum_{k=1}^n X_k$, risulta $Z_n \xrightarrow{\mathcal{L}} Z$ per $n \rightarrow \infty$ dove la v.a. Z si distribuisce con legge Normale $\mathcal{N}(0, 1)$. Questo risultato costituisce in effetti una estensione del Teorema Centrale del Limite dato da De Moivre. In particolare, se esiste $c \in]0, 1[$ tale che $p_k \in]c, 1 - c[$, allora si ha anche $q_k \in]c, 1 - c[$ e da cui segue $v_n > nc^2$, ovvero la condizione di Lyapunov è soddisfatta. La condizione di Lyapunov può essere soddisfatta perfino se $\lim_n p_n = 0$. Ad esempio, se $p_n = n^{-1}$, allora si ha

$$v_n = \sum_{k=1}^n \frac{1}{k} - \sum_{k=1}^n \frac{1}{k^2}$$

e quindi la condizione è verificata dal momento che la serie armonica $\sum_{n=1}^{\infty} n^{-1}$ non converge, mentre risulta $\sum_{n=1}^{\infty} n^{-2} = \zeta(2) = \frac{\pi^2}{6}$. □

Si considera infine un Teorema di convergenza, comunemente detto metodo Delta, che è spesso di notevole utilità pratica per lo studio di successioni di v.a.

Teorema 8.5.4. (Metodo Delta) Sia $(X_n)_{n \geq 1}$ una successione di v.a. indipendenti con la medesima legge e definite sullo spazio probabilizzato (Ω, \mathcal{F}, P) . Sia inoltre $(Z_n)_{n \geq 1}$ la successione di v.a. con

$$Z_n = \sqrt{n} \frac{X_n - \theta}{\psi},$$

dove θ e ψ sono opportune costanti, tale che $Z_n \xrightarrow{\mathcal{L}} Z$ per $n \rightarrow \infty$ e dove la v.a. Z è distribuita con legge Normale $\mathcal{N}(0, 1)$. Se $g : \mathbb{R} \rightarrow \mathbb{R}$ è una funzione continua e differenziabile tale che $g'(\theta) \neq 0$, data la successione di v.a. $(Y_n)_{n \geq 1}$ con

$$Y_n = \sqrt{n} \frac{g(X_n) - g(\theta)}{|g'(\theta)|\psi},$$

si ha $Y_n \xrightarrow{\mathcal{L}} Z$ per $n \rightarrow \infty$.

Dimostrazione. Si veda Billingsley (1995). □

• **Esempio 8.5.2.** Sia $(Z_n)_{n \geq 1}$, dove $Z_n = n^{-1}S_n$, una successione di v.a. indipendenti e con la medesima legge tale che ogni v.a. S_n è distribuita con legge Binomiale $\mathcal{B}(n, p)$. Tenendo presente il Teorema Centrale del Limite di De Moivre, la successione di v.a. $(V_n)_{n \geq 1}$ con

$$V_n = \sqrt{n} \frac{Z_n - p}{\sqrt{p(1-p)}}$$

è tale che $V_n \xrightarrow{\mathcal{L}} V$ per $n \rightarrow \infty$, dove la v.a. V si distribuisce con legge Normale $\mathcal{N}(0, 1)$. Dunque, se $g: \mathbb{R} \rightarrow \mathbb{R}$ è una funzione continua e differenziabile tale che $g'(p) \neq 0$, per la successione di v.a. $(Y_n)_{n \geq 1}$ con

$$Y_n = \sqrt{n} \frac{g(Z_n) - g(p)}{|g'(p)|\sqrt{p(1-p)}},$$

risulta $Y_n \xrightarrow{\mathcal{L}} V$ per $n \rightarrow \infty$. In particolare, se si richiede che $|g'(p)|\sqrt{p(1-p)} = 1$, allora deve essere $g(p) = 2\arcsin(\sqrt{p})$, ovvero in questo caso si ha

$$Y_n = 2\sqrt{n} (\arcsin(\sqrt{Z_n}) - \arcsin(\sqrt{p})).$$

Questa particolare scelta della funzione g è detta trasformazione stabilizzatrice, in quanto permette di ottenere una costante al denominatore della precedente definizione della v.a. Y_n . Questo risultato è spesso utile nella statistica inferenziale. □

8.6. Convergenza di vettori aleatori

I concetti di convergenza possono essere estesi al caso di v.v.a. Per quanto riguarda la notazione adottata in questo capitolo, $\|\cdot\|$ denota come al solito la distanza euclidea in \mathbb{R}^k . Inoltre, se F_X rappresenta la f.r.c. di un v.v.a. X con k componenti marginali, $x = (x_1, \dots, x_k)^T$ è un punto di continuità di F_X se la frontiera dell'insieme $]-\infty, x_1] \times \dots \times]-\infty, x_k]$ ha probabilità nulla.

Definizione 8.6.1. Se $(X_n)_{n \geq 1}$ è una successione di v.v.a. (ognuno dei quali possiede k componenti marginali) a cui corrisponde la successione di f.r.c. $(F_{X_n})_{n \geq 1}$, si dice che la successione di v.v.a. converge in legge al v.v.a. X con f.r.c. F_X se per ogni punto di continuità $x \in \mathbb{R}^k$ di F_X si ha

$$\lim_n F_{X_n}(x) = F_X(x)$$

e si scrive $X_n \xrightarrow{\mathcal{L}} X$ per $n \rightarrow \infty$. Se $(X_n)_{n \geq 1}$ è una successione di v.v.a. definiti sullo spazio probabilizzato (Ω, \mathcal{F}, P) , si dice che la successione di v.v.a. converge in probabilità al v.v.a. X , definito sul medesimo spazio probabilizzato, se per ogni $\varepsilon > 0$ si ha

$$\lim_n P(\|X_n - X\| > \varepsilon) = 0$$

e si scrive $X_n \xrightarrow{P} X$ per $n \rightarrow \infty$. Se $(X_n)_{n \geq 1}$ è una successione di v.v.a. definiti sullo spazio probabilizzato (Ω, \mathcal{F}, P) , si dice che la successione di v.v.a. *converge quasi certamente* al v.v.a. X , definito sul medesimo spazio probabilizzato, se l'evento

$$\{\omega \in \Omega : \lim_n X_n(\omega) = X(\omega)\} = \{\lim_n X_n = X\}$$

si verifica *q.c.* e si scrive $X_n \xrightarrow{q.c.} X$ per $n \rightarrow \infty$. Se $(X_n)_{n \geq 1}$ è una successione di v.v.a. definiti sullo spazio probabilizzato (Ω, \mathcal{F}, P) , si dice che la successione di v.v.a. *converge in media di ordine p* al v.v.a. X , definito sul medesimo spazio probabilizzato, se

$$\lim_n E(\|X_n - X\|^p) = 0$$

per $p \in]0, \infty[$ e si scrive $X_n \xrightarrow{L^p} X$ per $n \rightarrow \infty$. □

La gran parte dei risultati ottenuti per successioni di v.a. possono essere opportunamente estesi al caso di v.v.a. Di conseguenza, vengono considerati nel seguito solo alcuni Teoremi che hanno interesse specifico per v.v.a.

Il prossimo Teorema consente di determinare la convergenza in legge di una successione di v.v.a. attraverso la convergenza in legge di combinazioni lineari degli elementi della successione ed è dovuto a Harald Cramér e allo statistico e matematico norvegese Herman Ole Andreas Wold (1908-1992).

Teorema 8.6.2. (Teorema di Cramér-Wold) *Data una successione di v.v.a. $(X_n)_{n \geq 1}$ (ognuno dei quali possiede k componenti marginali), allora $X_n \xrightarrow{\mathcal{L}} X$ per $n \rightarrow \infty$ se e solo se $c^T X_n \xrightarrow{\mathcal{L}} c^T X$ per ogni $c \in \mathbb{R}^k$.*

Dimostrazione. Si veda Billingsley (1995, p.383). □

• **Esempio 8.6.1.** Si consideri la successione di v.v.a. $(Z_n)_{n \geq 1}$ tale che ogni v.v.a. Z_n è distribuito con legge Normale Multivariata $\mathcal{N}_k(\mu, \Sigma)$ e l'ulteriore successione di v.v.a. $(X_n)_{n \geq 1}$ dove $X_n = n^{-1} Z_n$. Dall'Esempio 7.3.2 è noto che la v.a. $c^T Z_n$ si distribuisce con legge Normale $\mathcal{N}(c^T \mu, c^T \Sigma c)$. Dunque, la v.a. $c^T X_n$ è tale che

$$E[c^T X_n] = \frac{1}{n} c^T \mu$$

e

$$\text{Var}[c^T X_n] = \frac{1}{n^2} c^T \Sigma c.$$

Di conseguenza, dalla Proposizione 8.2.2 segue che $c^T X_n \xrightarrow{P} 0$ per $n \rightarrow \infty$ e quindi dalla Proposizione 8.2.4 si ottiene che $c^T X_n \xrightarrow{\mathcal{L}} 0$ per $n \rightarrow \infty$. Dunque, dal Teorema 8.6.2 risulta infine che $X_n \xrightarrow{\mathcal{L}} 0$ per $n \rightarrow \infty$, ovvero la successione di v.v.a. converge in legge ad un v.v.a. degenere concentrato sul vettore a componenti nulle. □

Teorema 8.6.3. *Data una successione di v.v.a. $(X_n)_{n \geq 1}$ (ognuno dei quali possiede k componenti marginali) e se il v.v.a. X è degenere e tale che $P(X = c) = 1$ con $c \in \mathbb{R}^k$, allora $X_n \xrightarrow{P} X$ per $n \rightarrow \infty$ se e solo se ogni componente marginale del v.v.a. X_n converge in probabilità alla rispettiva componente marginale del v.v.a. X . Analogamente, $X_n \xrightarrow{q.c.} X$ per $n \rightarrow \infty$ se e solo se ogni componente marginale del v.v.a. X_n converge quasi certamente alla rispettiva componente marginale del v.v.a. X .*

Dimostrazione. Si veda Billingsley (1995, p.378). □

• **Esempio 8.6.2.** Si consideri la successione di v.v.a. $(Z_n)_{n \geq 1}$ tale che il v.v.a. Z_n è distribuito con legge Multinomiale $\mathcal{M}(n, p)$ dove $p = (p_1, \dots, p_k)^T$ e l'ulteriore successione di v.v.a. $(X_n)_{n \geq 1}$ dove $X_n = n^{-1}Z_n$. Dall'Esempio 7.4.1 è noto che la j -esima componente marginale del v.v.a. Z_n si distribuisce con legge Binomiale $\mathcal{B}(n, p_j)$ e dunque dall'Esempio 8.3.4 si ottiene che la j -esima componente marginale del v.v.a. X_n converge quasi certamente ad una v.a. degenera concentrata su p_j . Quindi, dal Teorema 8.6.3 si ha infine che $X_n \xrightarrow{q.c.} p$ per $n \rightarrow \infty$, ovvero la successione di v.v.a. converge quasi certamente al v.v.a. degenera concentrato sul vettore p . \square

Teorema 8.6.4. (Teorema Centrale Multivariato del Limite) Sia $(X_n)_{n \geq 1}$ una successione di v.v.a. indipendenti con la medesima legge (ognuno dei quali possiede k componenti marginali) e definiti sullo spazio probabilizzato (Ω, \mathcal{F}, P) , tali che per ogni n si ha $E[X_n] = \mu$ e $\text{Var}[X_n] = \Sigma$ dove $\det(\Sigma) < \infty$. Si consideri inoltre la successione di v.v.a. $(Z_n)_{n \geq 1}$ con

$$Z_n = \frac{1}{\sqrt{n}}(S_n - n\mu)$$

e $S_n = \sum_{k=1}^n X_k$. Se il v.v.a. Z è distribuito con legge Normale Multivariata $\mathcal{N}_k(0, \Sigma)$, allora $Z_n \xrightarrow{\mathcal{L}} Z$ per $n \rightarrow \infty$.

Dimostrazione. Si veda Billingsley (1995, p.385). \square

• **Esempio 8.6.3.** Si consideri la successione di v.v.a. $(X_n)_{n \geq 1}$ tale che ogni v.v.a. Z_n è distribuito con legge Multinomiale $\mathcal{M}(1, p)$. Inoltre, dall'Esempio 7.4.5 risulta che la legge del v.v.a. S_n è Multinomiale $\mathcal{M}(n, p)$. Dunque, il Teorema Centrale Multivariato del Limite fornisce la convergenza in legge della successione di v.v.a. $(Z_n)_{n \geq 1}$, dove $Z_n = n^{-\frac{1}{2}}(S_n - np)$, ad un v.v.a. con legge $\mathcal{N}_k(0, \text{diag}(p) - pp^T)$ per $n \rightarrow \infty$. \square

Teorema 8.6.5. (Metodo Delta Multivariato) Sia $(X_n)_{n \geq 1}$ una successione di v.v.a. indipendenti con la medesima legge (ognuno dei quali possiede k componenti marginali) e definiti sullo spazio probabilizzato (Ω, \mathcal{F}, P) . Sia inoltre $(Z_n)_{n \geq 1}$ la successione di v.v.a. con $Z_n = \sqrt{n}(X_n - \theta)$, tale che $Z_n \xrightarrow{\mathcal{L}} Z$ per $n \rightarrow \infty$, dove il v.v.a. Z è distribuito con legge Normale Multivariata $\mathcal{N}_k(0, \Psi)$, mentre $\theta \in \mathbb{R}^k$ e Ψ è una matrice simmetrica definita positiva. Se $g: \mathbb{R}^k \rightarrow \mathbb{R}^h$ è una funzione continua e differenziabile tale che $A = \frac{\partial}{\partial x} g(x)|_{x=\theta}$ è una matrice non nulla, data la successione di v.v.a. $(Y_n)_{n \geq 1}$ con $Y_n = \sqrt{n}(g(X_n) - g(\theta))$, si ha $Y_n \xrightarrow{\mathcal{L}} Y$ per $n \rightarrow \infty$, dove il v.v.a. Z è distribuito con legge Normale Multivariata $\mathcal{N}_h(0, A^T \Psi A)$.

Dimostrazione. Si veda Billingsley (1995, p.385). \square

8.7. Riferimenti bibliografici

I testi di Ferguson (1996), Serfling (1980) e van der Vaart (1998) considerano estesamente gli argomenti relativi alle convergenze probabilistiche, con speciale attenzione alle applicazioni statistiche. Il testo di Petrov (1995) è espressamente dedicato ai teoremi limite. I testi con approccio alla probabilità basato sulla Teoria della Misura hanno ampie parti dedicati alle convergenze e ai teoremi limite, come ad esempio Ash e Doléans-Dade (2000), Billingsley (1995), Gut (2005), Resnick (2014) e Shiryaev (2019).

8.8. Esercizi svolti

• **Nota.** Al fine di evitare notazioni ridondanti e ripetizioni, nei seguenti esercizi si assume l'esistenza di uno spazio probabilizzato (Ω, \mathcal{F}, P) , a meno che non venga specificato diversamente.

Sezione 8.1

• **Esercizio 8.1.1.** Si consideri la successione di v.a. $(X_n)_{n \geq 1}$ tale che la v.a. X_n si distribuisce con legge Normale $\mathcal{N}(\mu_n, \sigma_n^2)$. Sia inoltre $\lim_n \mu_n = \mu$ e $\lim_n \sigma_n^2 = \sigma^2$. Si determini la legge della v.a. X tale che $X_n \xrightarrow{\mathcal{L}} X$ per $n \rightarrow \infty$.

Soluzione. La f.c. della v.a. X_n è data da

$$\varphi_{X_n}(t) = e^{i\mu_n t - \frac{1}{2}\sigma_n^2 t^2},$$

e quindi

$$\lim_n \varphi_{X_n}(t) = e^{i\mu t - \frac{1}{2}\sigma^2 t^2}.$$

Dunque, per il Teorema di Lévy la v.a. X si distribuisce con legge Normale $\mathcal{N}(\mu, \sigma^2)$. □

• **Esercizio 8.1.2.** Si consideri la successione di v.a. indipendenti $(X_n)_{n \geq 1}$ tale che la f.p. della v.a. X_n è data da

$$p_{X_n}(x) = \frac{1}{2} \mathbf{1}_{\{-2^{-n}, 2^{-n}\}}(x).$$

Inoltre, si consideri l'ulteriore successione di v.a. $(S_n)_{n \geq 1}$, dove $S_n = \sum_{k=1}^n X_k$. Se la v.a. X si distribuisce con legge Uniforme su $] -1, 1[$, si verifichi che $S_n \xrightarrow{\mathcal{L}} X$ per $n \rightarrow \infty$.

Soluzione. Dal momento che la v.a. X_n è simmetrica rispetto all'origine, la f.c. è data da

$$\varphi_{X_n}(t) = E[\cos(tX_n)] = \cos(2^{-n}t)$$

e quindi

$$\varphi_{S_n}(t) = \prod_{k=1}^n \cos(2^{-k}t).$$

Evidentemente, per $t = 0$ risulta $\varphi_{S_n}(0) = 1$. Inoltre, tenendo presente la formula di duplicazione $\sin(2\theta) = 2\sin(\theta)\cos(\theta)$, per $t \neq 0$ si ha

$$\varphi_{S_n}(t) = \prod_{k=1}^n \frac{2 \sin(2^{-k}t) \cos(2^{-k}t)}{2 \sin(2^{-k}t)} = \prod_{k=1}^n \frac{\sin(2^{-k+1}t)}{2 \sin(2^{-k}t)} = \frac{\sin(t)}{2^n \sin(2^{-n}t)},$$

dal momento che il prodotto è telescopico. Quindi, risulta

$$\lim_n \varphi_{S_n}(t) = \lim_n \frac{\sin(t)}{t} \frac{2^{-n}t}{\sin(2^{-n}t)} = \frac{\sin(t)}{t}.$$

Inoltre, la v.a. X è simmetrica rispetto all'origine e la corrispondente f.c. è data da

$$\varphi_X(t) = E[\cos(tX)] = \int_{-1}^1 \cos(tx) \frac{1}{2} dx = \frac{\sin(t)}{t}.$$

Dunque, per il Teorema di Lévy si ha $S_n \xrightarrow{\mathcal{L}} X$ per $n \rightarrow \infty$. □

• **Esercizio 8.1.3.** Si consideri la successione $(X_n)_{n \geq 1}$ di v.a. indipendenti e con la medesima legge, tale che la f.p. della v.a. discreta X_n è data da

$$p_{X_n}(x) = \frac{1}{10} \mathbf{1}_{\{0,1,\dots,9\}}(x),$$

ovvero la v.a. X_n è distribuita con legge Uniforme discreta su $\{0, 1, \dots, 9\}$. Inoltre, si consideri la successione di v.a. $(S_n)_{n \geq 1}$, dove $S_n = \sum_{k=1}^n 10^{-k} X_k$. Se la v.a. X è distribuita con legge Uniforme su $]0, 1[$, si verifichi che $S_n \xrightarrow{\mathcal{L}} X$ per $n \rightarrow \infty$.

Soluzione. Dal momento che per ogni $a \in \mathbb{C}$ e tale che $a \neq 1$ si ha

$$\sum_{k=0}^n a^k = \frac{1 - a^{n+1}}{1 - a},$$

per $t \neq 0$ la f.c. della v.a. X_n è data da

$$\varphi_{X_n}(t) = \sum_{x=0}^9 \frac{1}{10} e^{itx} = \frac{1}{10} \frac{1 - e^{10it}}{1 - e^{it}},$$

mentre $\varphi_{X_n}(0) = 1$. Quindi, risulta

$$\varphi_{S_n}(t) = \prod_{k=1}^n \varphi_{X_n}(10^{-k}t) = \prod_{k=1}^n \frac{1}{10} \frac{1 - e^{10^{-k+1}it}}{1 - e^{10^{-k}it}} = \frac{1}{10^n} \frac{1 - e^{it}}{1 - e^{10^{-n}it}},$$

dal momento che il prodotto è telescopico. Dalla precedente espressione si ha

$$\lim_n \varphi_{S_n}(t) = \lim_n \frac{e^{it} - 1}{it} \frac{10^{-n}it}{e^{10^{-n}it} - 1} = \frac{e^{it} - 1}{it}.$$

Inoltre, la f.c. della v.a. X è data da

$$\varphi_X(t) = \int_0^1 \cos(tx) dx + i \int_0^1 \sin(tx) dx = \frac{e^{it} - 1}{it}.$$

Dunque, per il Teorema di Lévy risulta $S_n \xrightarrow{\mathcal{L}} X$ per $n \rightarrow \infty$. □

• **Esercizio 8.1.4.** (Legge di Fréchet) Si consideri una successione $(Y_n)_{n \geq 1}$ di v.a. indipendenti tale che la v.a. Y_n è distribuita con legge di Pareto ridotta con parametro di forma pari ad α , ovvero la f.r. di Y_n è data da

$$F_{Y_n}(y) = (1 - y^{-\alpha}) \mathbf{1}_{]1, \infty[}(y)$$

con $\alpha > 0$ (vedi Esempio 8.2.5). Data la successione di v.a. $(X_n)_{n \geq 1}$, dove $X_n = n^{-\frac{1}{\alpha}} \max_{1 \leq k \leq n} Y_k$, si determini la legge della v.a. X tale che $X_n \xrightarrow{\mathcal{L}} X$ per $n \rightarrow \infty$.

Soluzione. Considerata la successione $(Z_n)_{n \geq 1}$, dove $Z_n = \max_{1 \leq k \leq n} Y_k$, dall'assunzione di indipendenza si ha

$$F_{Z_n}(z) = P(Y_1 \leq z, \dots, Y_n \leq z) = \prod_{k=1}^n F_{Y_k}(z) = (1 - z^{-\alpha})^n \mathbf{1}_{]1, \infty[}(z).$$

Dunque, risulta

$$F_{X_n}(x) = P(X_n \leq x) = P(Z_n \leq n^{\frac{1}{\alpha}} x) = \left(1 - \frac{x^{-\alpha}}{n}\right)^n \mathbf{1}_{]n^{-\frac{1}{\alpha}}, \infty[}(x),$$

da cui

$$\lim_n F_{X_n}(x) = \lim_n \left(1 - \frac{x^{-\alpha}}{n}\right)^n \mathbf{1}_{]n^{-\frac{1}{\alpha}}, \infty[}(x) = e^{-x^{-\alpha}} \mathbf{1}_{]0, \infty[}(x).$$

La precedente espressione fornisce la f.r. di una v.a. X con legge di Fréchet ridotta con parametro di forma pari ad α . Si deve concludere che $X_n \xrightarrow{\mathcal{L}} X$ per $n \rightarrow \infty$. La legge prende il nome dal matematico francese Maurice René Fréchet (1878-1973). \square

• **Esercizio 8.1.5.** (Legge Zeta, seconda parte) Si consideri la v.a. X con legge Zeta (si veda l'Esercizio 4.1.2 e l'Esempio 7.2.7). Tenendo presente la formula di Eulero

$$\zeta(s) = \prod_{p \in \mathbb{P}} \frac{1}{1 - p^{-s}},$$

dove \mathbb{P} rappresenta l'insieme dei numeri primi, e considerata la successione di v.a. indipendenti $(X_p)_{p \in \mathbb{P}}$, dove la v.a. X_p è distribuita con legge Geometrica di parametro $(1 - p^{-s})$, si verifichi che

$$X = \prod_{p \in \mathbb{P}} p^{X_p}.$$

Soluzione. Posto $Y = -\log(X)$ e $Y_p = -\log(p)X_p$, la relazione da verificare è equivalente all'ulteriore relazione

$$Y = \sum_{p \in \mathbb{P}} Y_p.$$

Si osservi che la f.c. della v.a. Y è data da

$$\varphi_Y(t) = E[e^{-it \log(X)}] = E[X^{-it}] = \sum_{x=1}^{\infty} x^{-it} \frac{1}{\zeta(s)} x^{-s} = \frac{1}{\zeta(s)} \sum_{x=1}^{\infty} x^{-(s+it)} = \frac{\zeta(s+it)}{\zeta(s)}.$$

Tenendo presente l'espressione della f.c. di una v.a. con legge Geometrica, la f.c. della v.a. Y_p risulta

$$\varphi_{Y_p}(t) = E[e^{-it \log(p)X_p}] = \frac{1 - p^{-s}}{1 - p^{-s} e^{-it \log(p)}} = \frac{1 - p^{-s}}{1 - p^{-(s+it)}}.$$

Si consideri inoltre la successione di v.a. $(S_n)_{n \geq 1}$, dove $S_n = \sum_{p \leq n, p \in \mathbb{P}} Y_p$. Dal momento che le v.a. Y_p sono indipendenti, in quanto trasformate di v.a. indipendenti, la f.c. della v.a. S_n risulta

$$\varphi_{S_n}(t) = \prod_{p \leq n, p \in \mathbb{P}} \varphi_{Y_p}(t) = \prod_{p \leq n, p \in \mathbb{P}} \frac{1 - p^{-s}}{1 - p^{-(s+it)}}.$$

Dunque, essendo $\frac{1}{\zeta(s)} = \prod_{p \in \mathbb{P}} (1 - p^{-s})$ per la formula di Eulero, si ha

$$\lim_n \varphi_{S_n}(t) = \prod_{p \in \mathbb{P}} \frac{1 - p^{-s}}{1 - p^{-(s+it)}} = \frac{\prod_{p \in \mathbb{P}} (1 - p^{-s})}{\prod_{p \in \mathbb{P}} (1 - p^{-(s+it)})} = \frac{\zeta(s+it)}{\zeta(s)}$$

e per il Teorema di Lévy si ottiene la relazione richiesta. \square

• **Esercizio 8.1.6.** Si consideri la successione di v.a. indipendenti $(X_n)_{n \geq 1}$, tale che la v.a. X_n si distribuisce con legge di Cauchy e relativa f.r.

$$F_{X_n}(x) = \frac{1}{2} + \frac{1}{\pi} \arctan(x).$$

Inoltre, si consideri la successione di v.a. $(Y_n)_{n \geq 1}$, dove

$$Y_n = \frac{1}{n} \max_{1 \leq k \leq n} X_k.$$

Si determini la legge della v.a. Y tale che $Y_n \xrightarrow{\mathcal{L}} Y$ per $n \rightarrow \infty$.

Soluzione. La f.r. di Y_n è data da

$$\begin{aligned} F_{Y_n}(x) &= P(\max_{1 \leq k \leq n} X_k \leq nx) = P(X_1 \leq nx, \dots, X_n \leq nx) = \prod_{k=1}^n P(X_k \leq nx) \\ &= \prod_{k=1}^n F_{X_k}(nx) = \prod_{k=1}^n \left(\frac{1}{2} + \frac{1}{\pi} \arctan(nx) \right) = \left(\frac{1}{2} + \frac{1}{\pi} \arctan(nx) \right)^n. \end{aligned}$$

Per $x < 0$ si ha

$$\lim_n \left(\frac{1}{2} + \frac{1}{\pi} \arctan(nx) \right)^n = 0,$$

mentre, tenendo presente che $\arctan(x) = \frac{\pi}{2} - \arctan(\frac{1}{x})$ e che $\arctan(x) = x + O(x^3)$, per $x \geq 0$ si ha

$$\lim_n \left(\frac{1}{2} + \frac{1}{\pi} \arctan(nx) \right)^n = \lim_n \left(1 - \frac{1}{\pi nx} + O\left(\frac{1}{n^3}\right) \right)^n = e^{-\frac{1}{\pi x}}.$$

Dunque, si ha

$$F_Y(x) = \lim_n F_{Y_n}(x) = e^{-\frac{1}{\pi x}} \mathbf{1}_{[0, \infty[}(x).$$

Se si considera la v.a. trasformata $Z = 1/Y$, la corrispondente f.r. è data da

$$F_Z(x) = P\left(\frac{1}{Y} \leq x\right) = P\left(Y \geq \frac{1}{x}\right) = (1 - e^{-\frac{\pi}{x}}) \mathbf{1}_{[0, \infty[}(x).$$

Si deve quindi concludere che $Y_n \xrightarrow{\mathcal{L}} 1/Z$ per $n \rightarrow \infty$, dove la v.a. Z è distribuita con legge esponenziale con parametro di posizione nullo e parametro di scala pari a π . \square

• **Esercizio 8.1.7.** (Mistura di leggi) Si consideri il v.v.a. $(X_1, X_2)^T$ tale che la legge condizionata della v.a. X_2 all'evento $\{X_1 = x_1\}$ è la legge Normale $\mathcal{N}(0, x_1)$, mentre la v.a. X_1 è distribuita con legge di Poisson $\mathcal{P}(\lambda)$. Si verifichi che $X_2/\sqrt{\lambda} \xrightarrow{\mathcal{L}} Z$ per $\lambda \rightarrow \infty$, dove la v.a. Z si distribuisce con legge Normale $\mathcal{N}(0, 1)$.

Soluzione. La f.c. della v.a. X_2 è data da

$$\varphi_{X_2}(t) = \mathbb{E}[e^{itX_2}] = \mathbb{E}[\mathbb{E}[e^{itX_2} | X_1]] = \mathbb{E}[e^{-\frac{1}{2}t^2 X_1}] = e^{\lambda(e^{-\frac{1}{2}t^2} - 1)}.$$

Quindi, la f.c. della v.a. $Y_\lambda = X_2/\sqrt{\lambda}$ risulta

$$\varphi_{Y_\lambda}(t) = e^{\lambda(e^{-\frac{1}{2}t^2} - 1)}.$$

Infine, si ha

$$\lim_{\lambda \rightarrow \infty} \varphi_{Y_\lambda}(t) = e^{-\frac{1}{2}t^2}.$$

La precedente espressione fornisce la f.c. di una v.a. distribuita con legge Normale $\mathcal{N}(0, 1)$ e dunque $X_\lambda \xrightarrow{\mathcal{L}} Z$ per $\lambda \rightarrow \infty$. \square

• **Esercizio 8.1.8.** Si consideri il v.v.a. discreto con componenti indipendenti e ugualmente distribuite (X_{n1}, \dots, X_{nn}) , tale che la f.p. della v.a. X_{nk} è data da

$$p_{X_{nk}}(x) = \left(1 - \frac{1}{n} - \frac{1}{n^2}\right) \mathbf{1}_{\{0\}}(x) + \frac{1}{n} \mathbf{1}_{\{1\}}(x) + \frac{1}{n^2} \mathbf{1}_{\{2\}}(x),$$

con $k = 1, \dots, n$ e $n = 2, 3, \dots$. Si consideri inoltre la successione di v.a. $(S_n)_{n \geq 2}$, dove $S_n = \sum_{k=1}^n X_{nk}$. Se la v.a. X è distribuita con legge di Poisson $\mathcal{P}(1)$, si verifichi che $S_n \xrightarrow{\mathcal{L}} X$ per $n \rightarrow \infty$.

Soluzione. La f.g. della v.a. X_{nk} è data da

$$G_{X_{nk}}(t) = 1 + \frac{t-1}{n} - \frac{t^2-1}{n^2},$$

per cui la f.g. di S_n risulta

$$G_{S_n}(t) = \left(1 + \frac{t-1}{n} - \frac{t^2-1}{n^2}\right)^n.$$

Dunque, si ha

$$\lim_n G_{S_n}(t) = e^{t-1}.$$

La precedente espressione fornisce la f.g. di una v.a. X distribuita con legge di Poisson $\mathcal{P}(1)$ e dunque $S_n \xrightarrow{\mathcal{L}} X$ per $n \rightarrow \infty$. \square

• **Esercizio 8.1.9.** (Legge di Poisson composta) Si consideri la successione di v.a. indipendenti $(Z_n)_{n \geq 1}$, tale che la v.a. Z_n si distribuisce con legge di Poisson $\mathcal{P}(\lambda)$. Si consideri inoltre la v.a. Y , indipendente dalle v.a. della successione, che si distribuisce con legge di Poisson $\mathcal{P}(\mu)$. Posto $X_{\lambda\mu} = \sum_{n=1}^Y Z_n$, si determini la legge della v.a. V per cui si ha $X_{\lambda\mu} \xrightarrow{\mathcal{L}} V$ se $\lambda \rightarrow 0$ e $\mu \rightarrow \infty$ in modo che $\lambda\mu \rightarrow \gamma > 0$.

Soluzione. Tenendo presente l'espressione della f.g. delle v.a. Z_n e Y , la f.g. della v.a. $X_{\lambda\mu}$ è data da

$$G_{X_{\lambda\mu}}(t) = E[t^{\sum_{n=1}^Y Z_n}] = E[E[t^{\sum_{n=1}^Y Z_n} | Y]] = E[e^{\lambda Y(t-1)}] = e^{\mu(e^{\lambda(t-1)} - 1)}.$$

Dunque, dal momento che $e^x = 1 + x + O(x^2)$, risulta

$$\lim_{\lambda \rightarrow 0, \mu \rightarrow \infty, \lambda\mu \rightarrow \gamma} G_{X_{\lambda\mu}}(t) = \lim_{\lambda \rightarrow 0, \mu \rightarrow \infty, \lambda\mu \rightarrow \gamma} e^{\mu(\lambda(t-1) + O(\lambda^2))} = e^{\gamma(t-1)}.$$

La precedente espressione fornisce la f.g. di una v.a. distribuita con legge di Poisson $\mathcal{P}(\gamma)$, per cui $X_{\lambda\mu} \xrightarrow{\mathcal{L}} V$ se $\lambda \rightarrow 0$ e $\mu \rightarrow \infty$ con $\lambda\mu \rightarrow \gamma$. \square

Sezione 8.2

• **Esercizio 8.2.1.** Si consideri la successione $(X_n)_{n \geq 1}$ di v.a. indipendenti con legge Uniforme su $]0, 1[$ e la successione $(Y_n)_{n \geq 1}$ tale che

$$Y_n = \left(\prod_{k=1}^n X_k \right)^{\frac{1}{n}}.$$

Si verifichi che $Y_n \xrightarrow{P} e^{-1}$ per $n \rightarrow \infty$.

Soluzione. Si consideri la successione di v.a. $(Z_n)_{n \geq 1}$ tale che

$$Z_n = \log(Y_n) = \frac{1}{n} \sum_{k=1}^n \log(X_k).$$

Si ha

$$E[\log(X_n)] = \int_0^1 \log(x) dx = -1.$$

Quindi, sulla base della Legge Debole dei Grandi Numeri di Khinchin, si ha $Z_n \xrightarrow{P} -1$ per $n \rightarrow \infty$.

Dal momento che la funzione esponenziale è continua, risulta infine $Y_n = e^{Z_n} \xrightarrow{P} e^{-1}$ per $n \rightarrow \infty$. \square

• **Esercizio 8.2.2.** Si consideri la successione di v.a. indipendenti $(X_n)_{n \geq 1}$ tale che la f.p. della v.a. X_n è data da

$$p_{X_n}(x) = \frac{1}{2} \mathbf{1}_{\{-n^a, n^a\}}(x).$$

Inoltre, si consideri l'ulteriore successione di v.a. $(Z_n)_{n \geq 1}$ con $Z_n = n^{-1} \sum_{k=1}^n X_k$. Si verifichi che $Z_n \xrightarrow{P} 0$ per $n \rightarrow \infty$ se $a < \frac{1}{2}$.

Soluzione. Si ha $E[X_n] = 0$ e $\text{Var}[X_n] = n^{2a}$, da cui $E[Z_n] = 0$ e

$$\text{Var}[Z_n] = \frac{1}{n^2} \sum_{k=1}^n k^{2a}.$$

Tenendo presente che $\sum_{k=1}^n k^c = O(n^{c+1})$ se $c > -1$, allora risulta anche $\text{Var}[Z_n] = O(n^{2a-1})$. Inoltre, dalla disuguaglianza di Chebyshev, per $\varepsilon > 0$ si ha

$$P(|Z_n| > \varepsilon) \leq \frac{\text{Var}[Z_n]}{\varepsilon^2}.$$

Dunque, si ha $Z_n \xrightarrow{P} 0$ per $n \rightarrow \infty$ se $\lim_n \text{Var}[Z_n] = 0$, ovvero deve sussistere $a < \frac{1}{2}$. \square

• **Esercizio 8.2.3.** Si consideri la successione di v.a. indipendenti $(X_n)_{n \geq 1}$, tale che la v.a. X_n si distribuisce con legge Uniforme con f.r.

$$F_{X_n}(x) = x \mathbf{1}_{]0,1[}(x) + \mathbf{1}_{[1,\infty[}(x).$$

Inoltre, si consideri la successione di v.a. $(Y_n)_{n \geq 1}$, dove $Y_n = \max_{1 \leq k \leq n} X_k$, e la successione di v.a. $(Z_n)_{n \geq 1}$, dove $Z_n = \min_{1 \leq k \leq n} X_k$. Si verifichi che $Y_n \xrightarrow{P} 1$ e $Z_n \xrightarrow{P} 0$ per $n \rightarrow \infty$.

Soluzione. La f.r. di Y_n è data da

$$\begin{aligned} F_{Y_n}(x) &= P(\max_{1 \leq k \leq n} X_k \leq x) = P(X_1 \leq x, \dots, X_n \leq x) = \prod_{k=1}^n P(X_k \leq x) \\ &= \prod_{k=1}^n F_{X_k}(x) = \prod_{k=1}^n (x \mathbf{1}_{[0,1[}(x) + \mathbf{1}_{[1,\infty[}(x)) = x^n \mathbf{1}_{[0,1[}(x) + \mathbf{1}_{[1,\infty[}(x). \end{aligned}$$

Dunque, si ha

$$\lim_n F_{Y_n}(x) = \lim_n (x^n \mathbf{1}_{[0,1[}(x) + \mathbf{1}_{[1,\infty[}(x)) = \mathbf{1}_{[1,\infty[}(x),$$

da cui $Y_n \xrightarrow{\mathcal{L}} 1$. Inoltre, la f.r. di Z_n è data da

$$\begin{aligned} F_{Z_n}(x) &= P(\min_{1 \leq k \leq n} X_k \leq x) = 1 - P(X_1 \geq x, \dots, X_n \geq x) = 1 - \prod_{k=1}^n P(X_k \geq x) \\ &= 1 - \prod_{k=1}^n (1 - F_{X_k}(x)) = 1 - \prod_{k=1}^n (\mathbf{1}_{]-\infty,0[}(x) + (1-x) \mathbf{1}_{[0,1[}(x)) \\ &= (1 - (1-x)^n) \mathbf{1}_{[0,1[}(x) + \mathbf{1}_{[1,\infty[}(x). \end{aligned}$$

Dunque, si ha

$$\lim_n F_{Z_n}(x) = \lim_n ((1 - (1-x)^n) \mathbf{1}_{[0,1[}(x) + \mathbf{1}_{[1,\infty[}(x)) = \mathbf{1}_{[0,\infty[}(x),$$

da cui $Z_n \xrightarrow{\mathcal{L}} 0$. Dal momento che le due successioni di v.a. convergono in legge a due v.a. degeneri, risulta anche $Y_n \xrightarrow{P} 1$ e $Z_n \xrightarrow{P} 0$ per $n \rightarrow \infty$. \square

• **Esercizio 8.2.4.** Si consideri la successione $(X_n)_{n \geq 1}$ di v.a. indipendenti e ugualmente distribuite con la medesima legge della v.a. X , tale che $E[X] = 0$ e $\text{Var}[X] = 1$. Inoltre, sia Y una ulteriore v.a. indipendente che si distribuisce con legge di Poisson $\mathcal{P}(\lambda)$. Posto

$$S_\lambda = \frac{1}{\sqrt{Y}} \sum_{n=1}^Y X_n,$$

si verifichi che $S_\lambda \xrightarrow{\mathcal{L}} Z$ per $\lambda \rightarrow \infty$, dove Z è una v.a. che si distribuisce con legge Normale $\mathcal{N}(0, 1)$.

Soluzione. Si osservi che $\frac{1}{\lambda} Y \xrightarrow{P} 1$ per $\lambda \rightarrow \infty$ e quindi, sulla base del Teorema della trasformata continua, risulta $\sqrt{\frac{\lambda}{Y}} \xrightarrow{P} 1$ per $\lambda \rightarrow \infty$. Posto

$$T_\lambda = \frac{1}{\sqrt{\lambda}} \sum_{n=1}^Y X_n,$$

si ha $S_\lambda = \sqrt{\frac{\lambda}{Y}} T_\lambda \xrightarrow{\mathcal{L}} Z$ e dunque, sulla base del il Teorema di Cramér-Slutsky, si ottiene che $T_\lambda \xrightarrow{\mathcal{L}} Z$ per $\lambda \rightarrow \infty$. Si assuma che φ_X sia la f.c. della v.a. X e si noti che, per le proprietà della f.c., risulta $\varphi_X(t) = 1 - \frac{1}{2}t^2 + o(t^2)$ per $t \rightarrow 0$. Tenendo presente l'espressione della f.g. della v.a. Y , la f.c. della v.a. T_λ è data da

$$\begin{aligned} \varphi_{T_\lambda}(t) &= E[e^{it \frac{1}{\sqrt{\lambda}} \sum_{n=1}^Y X_n}] = E[E[e^{it \frac{1}{\sqrt{\lambda}} \sum_{n=1}^Y X_n} \mid Y]] = E[\varphi_X(\lambda^{-\frac{1}{2}} t)^Y] \\ &= E[(1 - \frac{1}{2\lambda} t^2 + o(\lambda^{-1}))^Y] = e^{\lambda(-\frac{1}{2\lambda} t^2 + o(\lambda^{-1}))}. \end{aligned}$$

Dunque, risulta

$$\lim_{\lambda \rightarrow \infty} \varphi_{T_\lambda}(t) = \lim_{\lambda \rightarrow \infty} e^{-\frac{1}{2}t^2 + \lambda o(\lambda^{-1})} = e^{-\frac{1}{2}t^2}.$$

La precedente espressione fornisce la f.c. di una v.a. distribuita con legge Normale $\mathcal{N}(0, 1)$, per cui $T_\lambda \xrightarrow{\mathcal{L}} Z$ per $\lambda \rightarrow \infty$. \square

Sezione 8.3

• **Esercizio 8.3.1.** Si consideri la successione $(X_n)_{n \geq 1}$ di v.a. indipendenti con legge Uniforme su $[0, 1]$ e l'ulteriore successione di v.a. $(Z_n)_{n \geq 1}$ dove $Z_n = \max(X_1, \dots, X_n)$. Si verifichi che $Z_n \xrightarrow{q.c.} 1$ per $n \rightarrow \infty$.

Soluzione. La f.r. della v.a. Z_n è data da

$$F_{Z_n}(z) = \prod_{k=1}^n F_{X_k}(z) = z^n \mathbf{1}_{[0,1[}(z) + \mathbf{1}_{[1,\infty[}(z).$$

Tenendo presente che se si verifica l'evento $\{Z_n > z\}$ si verificano anche tutti gli eventi $\{Z_k > z\}$ per $k \geq n$, se $\varepsilon \in]0, 1]$ risulta

$$\begin{aligned} \lim_n P\left(\bigcap_{k=n}^{\infty} \{|Z_k - 1| \leq \varepsilon\}\right) &= \lim_n P\left(\bigcap_{k=n}^{\infty} \{Z_k > 1 - \varepsilon\}\right) = \lim_n P(Z_n > 1 - \varepsilon) \\ &= \lim_n \min(1, 1 - (1 - \varepsilon)^n) = 1, \end{aligned}$$

ovvero $Z_n \xrightarrow{q.c.} 1$ per $n \rightarrow \infty$. \square

• **Esercizio 8.3.2.** Si consideri la successione $(X_n)_{n \geq 1}$ di v.a. indipendenti tale che la v.a. X_n è distribuita con legge Normale $\mathcal{N}(n^{-1}, n^{-1})$. Si verifichi che $X_n \xrightarrow{q.c.} 0$ per $n \rightarrow \infty$.

Soluzione. Dal momento che $X_n \stackrel{\mathcal{L}}{=} n^{-\frac{1}{2}}Z + n^{-1}$, dove la v.a. Z si distribuisce con legge Normale $\mathcal{N}(0, 1)$, allora si ha

$$\begin{aligned} E[|X_n|^4] &= \frac{1}{n^4} E[(n^{\frac{1}{2}}Z + 1)^4] = \frac{1}{n^4} E[n^2 Z^4 + 4n^{\frac{3}{2}} Z^3 + 6n Z^2 + 4n^{\frac{1}{2}} Z + 1] \\ &= \frac{1}{n^2} E[Z^4] + \frac{6}{n^3} E[Z^2] + \frac{1}{n^4} = \frac{3}{n^2} + \frac{6}{n^3} + \frac{1}{n^4} \end{aligned}$$

da cui

$$\sum_{n=1}^{\infty} E[|X_n|^4] = 3\zeta(2) + 6\zeta(3) + \zeta(4) < \infty.$$

Dunque, si deve concludere che $X_n \xrightarrow{q.c.} 0$ per $n \rightarrow \infty$. \square

Sezione 8.4

• **Esercizio 8.4.1.** Data la v.a. assolutamente continua Z che ammette d.p.

$$f_Z(z) = \frac{\log(2)}{z \log^2(z)} \mathbf{1}_{[2,\infty[}(z),$$

si consideri la successione di v.a. $(X_n)_{n \geq 1}$ tale che $X_n = n^{-1}Z$. Si verifichi che la successione di v.a. converge in probabilità ad una v.a. degenera concentrata su 0, ma non converge in media.

Soluzione. La f.r. della v.a. Z è data da

$$F_Z(z) = 1 - \frac{\log(2)}{\log(z)} \mathbf{1}_{[2, \infty[}(z),$$

da cui si ottiene

$$\lim_n P(|X_n| > \varepsilon) = \lim_n P(Z > n\varepsilon) = \lim_n \frac{\log(2)}{\log(n\varepsilon)} = 0,$$

ovvero $X_n \xrightarrow{P} 0$ per $n \rightarrow \infty$. Tuttavia, per ogni $p > 0$ si ha

$$E[|X_n|^p] = \frac{1}{n^p} E[Z^p] = \frac{\log(2)}{n^p} \int_2^\infty \frac{z^{p-1}}{\log^2(z)} = \infty$$

e quindi la successione non converge in media ad una v.a. degenerata concentrata su 0. \square

• **Esercizio 8.4.2.** Si consideri la successione di v.a. indipendenti $(X_n)_{n \geq 1}$ tale che X_n è una v.a. con f.p. data da

$$p_{X_n}(x) = (1 - n^{-a}) \mathbf{1}_{\{0\}}(x) + n^{-a} \mathbf{1}_{\{n\}}(x),$$

dove $a > 0$. Si verifichi che $X_n \xrightarrow{L^p} 0$ per $n \rightarrow \infty$ se $a > p$. Si commenti inoltre la convergenza quasi certa e in probabilità della successione $(X_n)_{n \geq 1}$.

Soluzione. Si osservi che $E[|X_n|^p] = n^{p-a}$. Quindi, se $a > p$ si ha

$$\lim_n E[|X_n|^p] = \lim_n n^{p-a} = 0,$$

ovvero $X_n \xrightarrow{L^p} 0$ per $n \rightarrow \infty$. Per quanto riguarda la convergenza quasi certa, se $a > 1$ si ha

$$\sum_{n=1}^\infty P(|X_n| > \varepsilon) = \sum_{n=1}^\infty n^{-a} = \zeta(a) < \infty,$$

ovvero $X_n \xrightarrow{q.c.} 0$ per $n \rightarrow \infty$. Dunque, quando $a = 2$ si ottiene un esempio di una successione di v.a. che converge quasi certamente, ma non converge in media quadratica. Infine, se $a > 0$ si ha

$$\lim_n P(|X_n| > \varepsilon) = \lim_n n^{-a} = 0,$$

ovvero $X_n \xrightarrow{P} 0$ per $n \rightarrow \infty$. Dunque, quando $a = \frac{1}{2}$ si ottiene un esempio di una successione di v.a. che converge in probabilità, ma non converge quasi certamente e in media quadratica. \square

Sezione 8.5

• **Esercizio 8.5.1.** Si consideri la successione $(X_n)_{n \geq 1}$ di v.a. indipendenti discrete, tale che la f.p. della v.a. X_n è data da

$$p_{X_n}(x) = \frac{1}{2} \mathbf{1}_{\{-an, an\}}(x),$$

dove $a > 0$. Si consideri inoltre la successione di v.a. $(S_n)_{n \geq 1}$, dove $S_n = \sum_{k=1}^n X_k$. Si verifichi la condizione del Teorema Centrale del Limite di Lindeberg-Lévy-Feller.

Soluzione. Si ha $E[X_n] = 0$ e $\text{Var}[X_n] = a^2 n^2$. Dunque, risulta

$$v_n^2 = \sum_{k=1}^n a^2 k^2 = \frac{a^2 n(n+1)(2n+1)}{6}.$$

Si noti che per ogni $\varepsilon > 0$ esiste un valore N_ε (che non dipende da a) tale che per $n > N_\varepsilon$ si ha $\varepsilon v_n > an$, in quanto questa disuguaglianza può essere espressa come

$$\frac{(n+1)(2n+1)}{6n} > \frac{1}{\varepsilon^2}.$$

Quindi, si ha

$$\lim_n \sum_{k=1}^n E[X_k^2 \mathbf{1}_{\varepsilon v_n, \infty}(|X_k|)] < \infty.$$

Inoltre, risulta $\lim_n v_n^2 = \infty$ per cui

$$\lim_n \frac{1}{v_n^2} \sum_{k=1}^n E[X_k^2 \mathbf{1}_{\varepsilon v_n, \infty}(|X_k|)] = 0.$$

Dunque, la condizione del Teorema Centrale del Limite di Lindeberg-Lévy-Feller è verificata per ogni a . Pertanto, se si considera la successione di v.a. $(Z_n)_{n \geq 1}$ con $Z_n = \frac{1}{v_n} S_n$, allora $Z_n \xrightarrow{\mathcal{L}} Z$ per $n \rightarrow \infty$, dove la v.a. Z si distribuisce con legge Normale $\mathcal{N}(0, 1)$. Si osservi che la condizione del Teorema Centrale del Limite di Lyapunov è in questo caso più facile da verificare. In effetti, scelto $\delta = 3$, si ha $E[|X_n|^3] = a^3 n^3$, per cui risulta

$$\sum_{k=1}^n E[|X_k|^3] = \sum_{k=1}^n a^3 k^3 = \frac{a^3 n^2 (n+1)^2}{4}.$$

Dunque, per ogni $a \in \mathbb{R}^+$ si ha

$$\lim_n \frac{1}{\sqrt{v_n^3}} \sum_{k=1}^n E[|X_k|^3] = 0. \quad \square$$

• **Esercizio 8.5.2.** (Formula di Stirling) Si dimostri la formula di Stirling

$$n! \sim \sqrt{2\pi} n^{n+\frac{1}{2}} e^{-n}$$

mediante l'uso della Teorema Centrale del Limite di Lindeberg-Lévy.

Soluzione. Si consideri una successione $(X_n)_{n \geq 1}$ di v.a. indipendenti e ugualmente distribuite con legge Esponenziale ridotta. Evidentemente, si ha $E[X_n] = \text{Var}[X_n] = 1$. Posto $S_n = \sum_{k=1}^n X_k$, si consideri inoltre la successione di v.a. $(Z_n)_{n \geq 1}$ dove

$$Z_n = \frac{S_n - n}{\sqrt{n}}.$$

Quindi, dal Teorema Centrale del Limite di Lindeberg-Lévy risulta $Z_n \xrightarrow{\mathcal{L}} Z$, dove la v.a. Z si distribuisce con legge Normale $\mathcal{N}(0, 1)$. Dal momento che $E[Z_n^2] = 1$ per ogni n , allora risulta $\lim_n E[Z_n^2] = E[Z^2] = 1$ e questo implica che anche i momenti assoluti di ordine inferiore convergono, ovvero si ha

$$\lim_n E[|Z_n|] = E[|Z|] = 2 \int_0^\infty z \phi(z) dz = \sqrt{\frac{2}{\pi}}.$$

Inoltre, tenendo presente che la v.a. S_n è distribuita con legge Gamma $\mathcal{G}(0, 1, n)$, segue che

$$\begin{aligned} E[|Z_n|] &= \int_0^\infty \left| \frac{z-n}{\sqrt{n}} \right| \frac{1}{\Gamma(n)} z^{n-1} e^{-z} dz \\ &= \frac{n^{n+\frac{1}{2}}}{(n-1)!} \int_0^\infty |u-1| u^{n-1} e^{-nu} du = \frac{2n^{n+\frac{1}{2}} e^{-n}}{n!}. \end{aligned}$$

Dunque, si deve concludere che

$$\lim_n \frac{E[|Z_n|]}{E[|Z|]} = \lim_n \frac{\sqrt{2\pi} n^{n+\frac{1}{2}} e^{-n}}{n!} = 1,$$

che fornisce una verifica probabilistica della formula di Stirling. La formula prende il nome dal matematico scozzese James Stirling (1692-1770). \square

• **Esercizio 8.5.3.** Si consideri la successione $(X_n)_{n \geq 1}$ di v.a. indipendenti e con la medesima legge, tale che $\mu = E[X_n]$ e $\sigma^2 = \text{Var}[X_n] < \infty$. Si verifichi che la condizione del Teorema Centrale del Limite di Lindeberg-Lévy-Feller è soddisfatta per la successione considerata.

Soluzione. Senza perdita di generalità si può assumere $\mu = 0$. Si ha

$$v_n^2 = \sum_{k=1}^n \sigma^2 = n\sigma^2,$$

mentre

$$\sum_{k=1}^n E[X_k^2 \mathbf{1}_{\varepsilon v_n, \infty}(|X_k|)] = n E[X_1^2 \mathbf{1}_{\varepsilon \sqrt{n}\sigma, \infty}(|X_1|)].$$

Dunque, si ottiene infine

$$\lim_n \frac{1}{v_n^2} \sum_{k=1}^n E[X_k^2 \mathbf{1}_{\varepsilon v_n, \infty}(|X_k|)] = \lim_n \frac{1}{\sigma^2} E[X_1^2 \mathbf{1}_{\varepsilon \sqrt{n}\sigma, \infty}(|X_1|)] = 0,$$

ovvero la condizione del Teorema Centrale del Limite di Lindeberg-Lévy-Feller è soddisfatta. Quindi, si noti che quando le condizioni del Teorema Centrale del Limite di Lindeberg-Lévy sono soddisfatte, anche quella del Teorema Centrale del Limite di Lindeberg-Lévy-Feller è soddisfatta. \square

• **Esercizio 8.5.4.** Si consideri la successione $(X_n)_{n \geq 1}$ di v.a. indipendenti e con la medesima legge, tale che la d.p. della v.a. X_n è data da

$$f_{X_n}(x) = \frac{1}{c|x|^3 \log^2(|x|)} \mathbf{1}_{]2, \infty[}(|x|),$$

dove

$$c = 2 \int_2^\infty \frac{1}{x^3 \log^2(x)} dx = 4 \text{Ei}(-\log(4)) + \frac{1}{\log(4)},$$

mentre

$$\text{Ei}(z) = \int_{-\infty}^z \frac{e^u}{u} du$$

rappresenta la funzione integrale esponenziale. Si verifichi che la condizione del Teorema Centrale del Limite di Lyapunov non è soddisfatta per la successione considerata.

Soluzione. Si ha $\mu = E[X_n] = 0$ e

$$\sigma^2 = \text{Var}[X_n] = \frac{2}{c} \int_2^{\infty} \frac{1}{x \log^2(x)} dx = \frac{2}{c \log(2)} < \infty.$$

Inoltre, per $\delta > 2$ risulta

$$E[|X_n|^\delta] = \frac{2}{c} \int_2^{\infty} \frac{x^{\delta-3}}{\log^2(x)} dx = \infty$$

e quindi la condizione del Teorema Centrale del Limite di Lyapunov non può essere soddisfatta. Tuttavia, si noti che $(X_n)_{n \geq 1}$ è una successione di v.a. indipendenti con la medesima legge e tale che $\text{Var}[X_n] < \infty$. Dunque, sulla base dell'Esercizio 8.5.3, la condizione del Teorema Centrale del Limite di Lindeberg-Lévy-Feller è invece soddisfatta. \square

• **Esercizio 8.5.5.** Si dimostri che

$$\lim_n e^{-n} \sum_{k=0}^n \frac{n^k}{k!} = \frac{1}{2}$$

applicando il Teorema Centrale del Limite ad una opportuna successione di v.a. indipendenti distribuite con legge di Poisson.

Soluzione. Si consideri la successione di v.a. indipendenti $(X_n)_{n \geq 1}$, tale che la v.a. X_n è distribuita con legge di Poisson $\mathcal{P}(1)$. Inoltre, si abbia la successione di v.a. $(S_n)_{n \geq 1}$, dove $S_n = \sum_{k=1}^n X_k$. Per le proprietà della legge di Poisson, la v.a. S_n è distribuita con legge di Poisson $\mathcal{P}(n)$. Dunque, si ha

$$P(S_n \leq n) = e^{-n} \sum_{k=0}^n \frac{n^k}{k!}.$$

Applicando il Teorema Centrale del Limite risulta inoltre

$$\lim_n P(S_n \leq n) = \lim_n P\left(\frac{S_n - n}{\sqrt{n}} \leq 0\right) = \Phi(0) = \frac{1}{2},$$

da cui segue la relazione desiderata. \square

• **Esercizio 8.5.6.** Si consideri la successione $(X_n)_{n \geq 1}$ di v.v.a. indipendenti tali che $X_n = (Y_n, Z_n)^T$ dove $Y_n = Y + \frac{1}{n}$ e $Z_n = 1 - Y_n$, mentre Y è una v.a. distribuita con legge di Bernoulli $\mathcal{B}(1, \frac{1}{2})$. Sia inoltre Z una ulteriore v.a. distribuita con legge di Bernoulli $\mathcal{B}(1, \frac{1}{2})$ indipendente dalla v.a. Y . Si verifichi che $Y_n \xrightarrow{\mathcal{L}} Y$, $Z_n \xrightarrow{\mathcal{L}} Z$ per $n \rightarrow \infty$, anche se la successione di v.a. $(Y_n + Z_n)_{n \geq 1}$ non converge in legge al v.v.a. $(Y + Z)$, ovvero $(X_n)_{n \geq 1}$ non converge in legge al v.v.a. $X = (Y, Z)^T$.

Soluzione. Dal momento che la f.c. della v.a. Y_n è data da

$$\varphi_{Y_n}(t) = e^{i \frac{t}{n}} \varphi_Y(t),$$

si ha

$$\lim_n \varphi_{Y_n}(t) = \varphi_Y(t),$$

ovvero $Y_n \xrightarrow{\mathcal{L}} Y$. Dunque, risulta anche $Z_n \xrightarrow{\mathcal{L}} 1 - Y \stackrel{\mathcal{L}}{=} Z$. Inoltre, essendo $Y_n + Z_n = 1$, si ha $Y_n + Z_n \xrightarrow{\mathcal{L}} 1$. Tuttavia, dal momento che Y e Z sono v.a. indipendenti, si osservi che la v.a. $(Y + Z)$ è distribuita con legge Binomiale $\mathcal{B}(2, \frac{1}{2})$, ovvero $(X_n)_{n \geq 1}$ non converge in legge al v.v.a. X . \square

Sezione 8.6

• **Esercizio 8.6.1.** Si consideri la successione $(X_n)_{n \geq 1}$ di v.a. indipendenti e con la medesima legge della v.a. X tale che $E[X^4] < \infty$. Si consideri inoltre la successione di v.v.a. $(S_n)_{n \geq 1}$ tale che $S_n = (S_{n,1}, S_{n,2})^T = (\sum_{k=1}^n X_k, \sum_{k=1}^n X_k^2)^T$ e la successione di v.v.a. $(Z_n)_{n \geq 1}$ tale che

$$Z_n = \frac{1}{\sqrt{n}} (S_n - nE[S_n]).$$

Si verifichi che $Z_n \xrightarrow{\mathcal{L}} Z$ per $n \rightarrow \infty$ e si determini la legge relativa al v.v.a. Z .

Soluzione. Si assuma che $E[X] = \mu$ e $E[X^r] = \mu_r$ per $r = 2, 3, 4$, mentre $\text{Var}[X] = \sigma^2$. Inoltre, si consideri la successione $(Y_n)_{n \geq 1}$ tale che $Y_n = (Y_{n,1}, Y_{n,2})^T = (X_n, X_n^2)^T$. Dunque, risulta immediato verificare che $\nu = E[Y_n] = (\mu, \mu_2)^T$. Inoltre, si ha $\text{Var}[Y_{n,1}] = \sigma^2$ e $\text{Var}[Y_{n,2}] = \mu_4 - \mu_2^2$, mentre

$$\text{Cov}[Y_{n,1}, Y_{n,2}] = \mu_3 - \mu\mu_2.$$

Quindi, si ha

$$\Sigma = \text{Var}[Y_n] = \begin{pmatrix} \sigma^2 & \mu_3 - \mu\mu_2 \\ \mu_3 - \mu\mu_2 & \mu_4 - \mu_2^2 \end{pmatrix}.$$

Dal momento che $S_n = \sum_{k=1}^n Y_k$, allora risulta $E[S_n] = n\nu$ e $\text{Var}[S_n] = n\Sigma$. Per il Teorema Centrale Multivariato del Limite si ha dunque che il v.v.a. $Z = \frac{1}{\sqrt{n}}(S_n - n\nu)$ è distribuito con legge Normale Multivariata $\mathcal{N}_2(0, \Sigma)$. In un contesto di Statistica inferenziale, si osservi che le componenti del v.v.a. $U_n = (U_{n,1}, U_{n,2})^T = n^{-1}S_n$ sono date rispettivamente dalla media campionaria e dal secondo momento campionario. Dunque, il risultato ottenuto implica che $\sqrt{n}(U_n - \nu) \xrightarrow{\mathcal{L}} Z$, ovvero si è ottenuta la distribuzione per “grandi campioni” dei primi due momenti campionari $U_{n,1}$ e $U_{n,2}$. \square

• **Esercizio 8.6.2.** Si consideri la successione $(X_n)_{n \geq 1}$ di v.a. indipendenti e con la medesima legge della v.a. X con $E[X^4] < \infty$. Inoltre, sia $(V_n)_{n \geq 1}$ una ulteriore successione di v.a. tale che

$$V_n = \frac{1}{n} \sum_{k=1}^n X_k^2 - \left(\frac{1}{n} \sum_{k=1}^n X_k \right)^2.$$

Considerata la successione di v.a. $(Z_n)_{n \geq 1}$ con $Z_n = \sqrt{n}(V_n - \nu)$, si determini la legge relativa alla v.a. Z tale che $Z_n \xrightarrow{\mathcal{L}} Z$ per $n \rightarrow \infty$ e il valore di ν .

Soluzione. Si adotti le notazioni dell'Esercizio 8.6.1. Si osservi che

$$V_n = \frac{1}{n} \sum_{k=1}^n (X_k - \mu)^2 - \left(\frac{1}{n} \sum_{k=1}^n (X_k - \mu) \right)^2$$

e quindi per semplicità si può assumere $\mu = 0$, da cui $\mu_2 = \sigma^2$. Inoltre, risulta $V_n = g(U_n)$, dove $g(u_1, u_2) = u_2 - u_1^2$, per cui

$$A = \left. \frac{\partial}{\partial x} g(x) \right|_{x=\nu} = \begin{pmatrix} -2\mu \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Quindi, si ha $A^T \Sigma A = \mu_4 - \sigma^4$ e $\nu = g(\nu) = \sigma^2$. Sulla base del metodo Delta Multivariato, si deve dunque concludere che $Z_n = \sqrt{n}(V_n - \sigma^2) \xrightarrow{\mathcal{L}} Z$, dove la v.a. Z è distribuita con legge Normale $\mathcal{N}(0, \mu_4 - \sigma^4)$. In un contesto di Statistica inferenziale, si osservi che la v.a. V_n è in effetti la varianza campionaria, di cui si è ottenuta la distribuzione per “grandi campioni”. \square

Capitolo 9

Processi aleatori

9.1. Concetti preliminari

Molti esperimenti o fenomeni aleatori danno luogo ad un insieme di realizzazioni all'evolversi di un "parametro", come ad esempio il tempo. In questo caso, l'interesse si concentra sull'analisi di collezioni di v.a. indicizzate su un determinato insieme, che nella Teoria della Probabilità sono dette processi aleatori. Formalmente, si ha la seguente definizione.

Definizione 9.1.1. Se $\mathbb{T} \subseteq [0, \infty[$ è un insieme di indici, una collezione di v.a. $X = (X_t)_{t \in \mathbb{T}}$ definita sullo spazio probabilizzato (Ω, \mathcal{F}, P) è detta *processo aleatorio* (p.a.) se l'applicazione $X : \mathbb{T} \times \Omega \rightarrow \mathbb{R}$, dove $X(t, \omega) = X_t(\omega)$, è misurabile rispetto alla σ -algebra $\mathcal{B}(\mathbb{T}) \otimes \mathcal{F}$. Per un dato $t \in \mathbb{T}$, la v.a. X_t è detta *stato* del p.a. a tempo t . Al contrario, per un dato $\omega \in \Omega$, l'applicazione $t \mapsto X_t(\omega)$ è detta *traiettoria* del p.a. associata con ω . \square

Anche se in generale si possono considerare spazi molto generali per \mathbb{T} , si noti che nella precedente Definizione l'insieme \mathbb{T} è stato assimilato per semplicità ad un insieme di tempi. Nel seguito si assumerà principalmente che $\mathbb{T} = \mathbb{N}$ o che $\mathbb{T} = [0, \infty[$ (o eventualmente un loro sottoinsieme). Più esattamente, il p.a. sarà detto *a tempo discreto* se $\text{card}(\mathbb{T})$ è finita o numerabile o *a tempo continuo* se $\text{card}(\mathbb{T})$ non è numerabile. Inoltre, quando $\mathbb{T} = \mathbb{N}$, il p.a. si riduce ad una successione di v.a. e si enfatizzerà questo fatto adottando l'indice n piuttosto che l'indice t .

• **Esempio 9.1.1.** Si consideri una successione di v.a. $(Z_n)_{n \geq 1}$ a componenti indipendenti distribuiti con legge di Bernoulli di parametro p e si definisca il p.a. a tempo discreto $X = (X_n)_{n \geq 0}$ dove $X_n = \sum_{k=1}^n Z_k$ e $X_0 = 0$ q.c. In questo caso, si ha ovviamente $\mathbb{T} = \mathbb{N}$. Inoltre, per quanto visto nella Sezione 6.1, per un determinato $n \geq 1$ lo stato X_n è una v.a. con legge Binomiale $\mathcal{B}(n, p)$, mentre una traiettoria è una successione non decrescente a valori su \mathbb{N} tale che gli incrementi fra due termini successivi sono nulli o pari all'unità. In un ambito di gioco d'azzardo, piuttosto che la successione originale $(Z_n)_{n \geq 1}$, si considera invece la successione $(Y_n)_{n \geq 1}$, con $Y_n = 2Z_n - 1$, ovvero si ottiene il p.a. a tempo discreto $X = (X_n)_{n \geq 0}$ dove $X_n = \sum_{k=1}^n Y_k = 2 \sum_{k=1}^n Z_k - n$. In questo caso, il p.a. X viene assimilato ad una successione di giocate in cui si vince un'unità con probabilità p e si perde un'unità con probabilità $(1 - p)$ ad ogni giocata. Lo stato X_n risulta dunque la v.a. che rappresenta la somma vinta (o perduta) a tempo n , mentre una traiettoria è la successione delle somme vinte (o perdute). Con un termine pittoresco, questo p.a. è detto passeggiata aleatoria e sarà analizzato in dettaglio nella Sezione 9.3. \square

• **Esempio 9.1.2.** Sia $\mathbb{T} = [0, T]$ e sia $0 = t_0 < t_1 < \dots < t_n = T$ una partizione di $[0, T]$. Inoltre, si consideri il v.v.a. $Z = (Z_0, \dots, Z_{n-1})^T$ e si definisca il p.a. a tempo continuo $X = (X_t)_{t \in [0, T]}$ dove

$$X_t = \sum_{k=0}^{n-1} Z_k \mathbf{1}_{[t_k, t_{k+1}[}(t).$$

Per un determinato $t \in [t_k, t_{k+1}[$, lo stato X_t è una v.a. con la stessa legge della v.a. Z_k , mentre una traiettoria è una funzione semplice. Di conseguenza, questo p.a. è detto semplice. Questo tipo di p.a. sarà fondamentale nel calcolo stocastico introdotto nel Capitolo 10. \square

Definizione 9.1.2. Si consideri uno spazio probabilizzabile (Ω, \mathcal{F}) . Se $\mathbb{T} \subseteq [0, \infty[$, una *filtrazione* è una collezione di σ -algebre $(\mathcal{F}_t)_{t \in \mathbb{T}}$ per cui si ha $\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$ per ogni $s, t \in \mathbb{T}$ tali che $s \leq t$. Inoltre, dato il p.a. $X = (X_t)_{t \in \mathbb{T}}$ definito sullo spazio probabilizzato (Ω, \mathcal{F}, P) , sia $\mathcal{F}_t = \sigma(\{X_s : s \leq t\})$ la più piccola σ -algebra che contiene gli eventi del tipo $\{X_s \in B\}$ con $B \in \mathcal{B}(\mathbb{R})$ e $s \leq t$. In questo caso, la corrispondente filtrazione viene detta *filtrazione naturale* per il p.a. X . \square

Per comodità, si assume usualmente che $\inf(\{t : t \in \mathbb{T}\}) = 0$ e $\mathcal{F}_0 = \{\emptyset, \Omega\}$. Inoltre, uno spazio di probabilità (Ω, \mathcal{F}, P) a cui è associata la filtrazione $(\mathcal{F}_t)_{t \in \mathbb{T}}$ è detto *spazio di probabilità con filtrazione*. Una filtrazione può essere interpretata come l'incremento di informazione nel p.a. che si ottiene al passare del tempo. In effetti, \mathcal{F}_t può essere vista come l'informazione disponibile ad un osservatore del p.a. fino a tempo t .

• **Esempio 9.1.3.** Si consideri di nuovo il p.a. X dell'Esempio 9.1.1. Si osservi che ogni traiettoria del p.a. può essere posta in corrispondenza biunivoca con un punto dell'intervallo $[0, 1]$, dal momento che ogni determinazione della successione $(Z_n)_{n \geq 1}$ può essere vista come l'insieme delle cifre nella rappresentazione binaria di un punto in $[0, 1]$. In questo caso, si può considerare la filtrazione $(\mathcal{F}_n)_{n \geq 0}$ con $\mathcal{F}_0 = \{\emptyset, \Omega\}$ e dove \mathcal{F}_n è la σ -algebra generata dagli intervalli del tipo $[2^{-n}(k-1), 2^{-n}k[$ con $k = 1, \dots, 2^n$. Evidentemente, la partizione binaria di $[0, 1]$ che genera \mathcal{F}_n diventa sempre più “fine” all'aumentare di n e la relativa σ -algebra diventa contemporaneamente più “ricca”. \square

Definizione 9.1.3. Sia $X = (X_t)_{t \in \mathbb{T}}$ un p.a. definito sullo spazio di probabilità (Ω, \mathcal{F}, P) con filtrazione $(\mathcal{F}_t)_{t \in \mathbb{T}}$. Il p.a. X è detto *adattato* se X_t è misurabile rispetto a \mathcal{F}_t per ogni $t \in \mathbb{T}$. Inoltre, il p.a. X è detto *progressivamente misurabile* se l'applicazione $(s, \omega) \mapsto X_s(\omega)$ è misurabile rispetto alla σ -algebra $\mathcal{B}([0, t]) \otimes \mathcal{F}_t$ per ogni $t \in \mathbb{T}$. \square

Dalla Definizione 9.1.3 è evidente che un p.a. X è adattato se la filtrazione è naturale.

Definizione 9.1.4. Sia $X = (X_t)_{t \in \mathbb{T}}$ un p.a. definito sullo spazio di probabilità (Ω, \mathcal{F}, P) . Il p.a. X è detto *continuo* se le corrispondenti traiettorie sono continue *q.c.* Inoltre, il p.a. è detto *continuo a destra* se le corrispondenti traiettorie sono continue a destra *q.c.* \square

• **Esempio 9.1.4.** Si consideri di nuovo il p.a. semplice X introdotto nell'Esempio 9.1.2. È immediato verificare che le traiettorie del p.a. X non sono continue *q.c.* In effetti, risulta $\{X_{t_k} = Z_k\}$ *q.c.*, mentre per $\varepsilon > 0$ si ha $\{\lim_{\varepsilon \rightarrow 0^+} X_{t_k - \varepsilon} = Z_{k-1}\}$ *q.c.*, dove $k = 1, \dots, n$. Dunque, il p.a. X non è continuo *q.c.* Il p.a. X è continuo a destra *q.c.*, dal momento che $\{\lim_{\varepsilon \rightarrow 0^+} X_{t_k + \varepsilon} = Z_k\}$ *q.c.* \square

Definizione 9.1.5. Si consideri i p.a. $X = (X_t)_{t \in \mathbb{T}}$ e $Y = (Y_t)_{t \in \mathbb{T}}$ definiti sullo spazio di probabilità (Ω, \mathcal{F}, P) . Si dice che il p.a. Y è una *modificazione* del p.a. X se per ogni $t \in \mathbb{T}$ si ha

$$P(\{\omega \in \Omega : X_t(\omega) = Y_t(\omega)\}) = 1.$$

Inoltre, si dice che i p.a. X e Y sono *indistinguibili* se si ha

$$P(\{\omega \in \Omega : X_t(\omega) = Y_t(\omega), \forall t \in \mathbb{T}\}) = 1. \quad \square$$

Risulta evidente dalla precedente Definizione che se i p.a. X e Y sono indistinguibili, allora il p.a. X è anche una modificazione del p.a. Y , mentre non è vero il contrario. Sulla base di queste classi di

equivalenza fra p.a., si può selezionare il p.a. più opportuno ai fini teorici all'interno di una classe, come ad esempio un p.a. continuo.

• **Esempio 9.1.5.** Si consideri i p.a. $X = (X_t)_{t \in [0, \infty[}$ e $Y = (Y_t)_{t \in [0, \infty[}$ definiti sullo spazio di probabilità (Ω, \mathcal{F}, P) , dove $\Omega = [0, \infty[$ e $\mathcal{F} = \mathcal{B}([0, \infty[)$. Inoltre, i p.a. X e Y sono tali che $X_t = 0$, mentre $Y_t = \mathbf{1}_{\{t = \omega\}}(t)$. Dunque, i p.a. X e Y non sono indistinguibili, in quanto $X_t(\omega) \neq Y_t(\omega)$ per $t = \omega$. Tuttavia, il p.a. Y è una modificazione del p.a. X , in quanto

$$P(\{\omega \in \Omega : X_t(\omega) \neq Y_t(\omega)\}) = 0.$$

e dunque

$$P(\{\omega \in \Omega : X_t(\omega) = Y_t(\omega)\}) = 1 - P(\{\omega \in \Omega : X_t(\omega) \neq Y_t(\omega)\}) = 1. \quad \square$$

9.2. Martingale

Le cosiddette “martingale” costituiscono una vasta classe di p.a. con proprietà notevoli. Il termine “martingala” ha origine nel gioco d'azzardo, anche se la sua etimologia non è del tutto chiara. La prima apparizione del termine in ambito probabilistico è dovuta al matematico Jean Ville (1910-1989) e lo studio di questi p.a. è stato profondamente influenzato dal probabilista Joseph Leo Doob (1910-2004), che ne ha costruito le basi teoriche.



Figura 9.2.1. Joseph Leo Doob (1910-2004).

Definizione 9.2.1. Se $\mathbb{T} \subseteq [0, \infty[$, sia $M = (M_t)_{t \in \mathbb{T}}$ un p.a. adattato definito sullo spazio di probabilità (Ω, \mathcal{F}, P) con filtrazione $(\mathcal{F}_t)_{t \in \mathbb{T}}$ e tale che $E[|M_t|] < \infty$ per ogni $t \in \mathbb{T}$. Si dice che M è una *martingala* se per ogni $0 \leq s < t$ si ha

$$\{E[M_t | \mathcal{F}_s] = M_s\} \text{ q.c. ,}$$

mentre si dice che M è una *super-martingala* se

$$\{E[M_t | \mathcal{F}_s] \leq M_s\} \text{ q.c. ,}$$

e infine si dice che M è una *sub-martingala* se

$$\{E[M_t | \mathcal{F}_s] \geq M_s\} \text{ q.c. .} \quad \square$$

È evidente che la proprietà di martingala dipende dalla filtrazione e dalla misura di probabilità P che sono state considerate. Inoltre, dalla Definizione 9.2.1, per ogni t si ha

$$\{E[M_t | \mathcal{F}_0] = M_0\} \text{ q.c. ,}$$

ovvero, una martingala può essere considerato in effetti un p.a. che rimane “costante in media”. Analogamente, una super-martingala è un p.a. che “decrese in media” e una sub-martingala è un p.a. che “cresce in media”.

Risulta importante sottolineare che ogni martingala ammette sempre una modificazione continua a destra che è unica, a meno di p.a. indistinguibili (si veda Pascucci, 2011, p.115). Dal momento che questa condizione non è restrittiva, nel seguito verrà considerata tacitamente la versione continua a destra di ogni martingala.

• **Esempio 9.2.1.** Si consideri una successione di v.a. $(X_n)_{n \geq 1}$ a componenti indipendenti tale che $E[X_n] = c < \infty$ per ogni n . Si definisca inoltre il p.a. a tempo discreto $M = (M_n)_{n \geq 0}$ tale che

$$M_n = \sum_{k=1}^n X_k$$

con $\{M_0 = 0\}$ q.c. e sia $(\mathcal{F}_n)_{n \geq 0}$ la filtrazione naturale, ovvero si ha $\mathcal{F}_n = \sigma(\{X_1, \dots, X_n\})$. Tenendo presente le assunzioni fatte sul p.a. M , per ogni n si ha

$$E[|M_n|] = E\left[\left|\sum_{k=1}^n X_k\right|\right] \leq E\left[\sum_{k=1}^n |X_k|\right] = \sum_{k=1}^n E[|X_k|] < \infty .$$

Inoltre, dal momento che $\{E[M_n | \mathcal{F}_n] = M_n\}$ q.c. sulla base della Proposizione 5.2.1 e che la σ -algebra generata da X_n è indipendente dalla σ -algebra \mathcal{F}_{n-1} , si ha q.c.

$$\begin{aligned} E[M_n | \mathcal{F}_{n-1}] &= E[M_{n-1} + X_n | \mathcal{F}_{n-1}] = E[M_{n-1} | \mathcal{F}_{n-1}] + E[X_n | \mathcal{F}_{n-1}] \\ &= M_{n-1} + E[X_n] = M_{n-1} + c . \end{aligned}$$

Procedendo in modo iterativo, per $m \in \mathbb{N}$ con $m < n$, dalla precedente espressione si ha q.c.

$$E[M_n | \mathcal{F}_m] = M_m + (n - m)c .$$

Quindi, il p.a. M è una martingala se $c = 0$, una super-martingala se $c < 0$ e una sub-martingala se $c > 0$. \square

• **Esempio 9.2.2.** (Strategia di gioco) Si consideri una successione $(Y_n)_{n \geq 1}$ di v.a. indipendenti con la medesima legge, tali che $P(Y_n = 1) = p$ e $P(Y_n = -1) = q$, dove $q = 1 - p$, e una successione $(U_n)_{n \geq 1}$ di v.a. tale che $U_n = g_n(Y_1, \dots, Y_{n-1})$ con $P(U_n \geq 0) = 1$ e $E[U_n] < \infty$ per ogni $n \in \mathbb{N}$. Si definisca il p.a. a tempo discreto $M = (M_n)_{n \geq 0}$ tale che

$$M_n = \sum_{k=1}^n U_k Y_k$$

con $\{M_0 = 0\}$ q.c. e sia $(\mathcal{F}_n)_{n \geq 0}$ la filtrazione naturale, ovvero si assume che $\mathcal{F}_n = \sigma(\{Y_1, \dots, Y_n\})$. Dalle assunzioni fatte si noti che la v.a. U_n è misurabile rispetto a \mathcal{F}_{n-1} . Questo p.a. può essere assimilato ad un insieme di giocate ad un gioco d'azzardo con esito dicotomico, in cui si vince con probabilità p e si perde con probabilità q , e dove si scommette una quantità di denaro pari a U_n all' n -esima giocata. In questo caso, M_n rappresenta la somma vinta (o perduta) dal giocatore all' n -esima giocata. Nella terminologia del gioco d'azzardo, la successione di v.a. $(U_n)_{n \geq 1}$ è detta *strategia di gioco*. Risulta evidente il motivo per cui la v.a. U_n dipende solo dal v.v.a. $(Y_1, \dots, Y_{n-1})^T$. In effetti, il giocatore può prendere la decisione sulla somma da scommettere all' n -esima giocata solo basandosi

sugli esiti delle $(n - 1)$ giocate precedenti. Tenendo presente che le v.a. U_n e Y_n sono indipendenti e che $E[|Y_n|] = 1$, per ogni n si ha

$$E[|M_n|] = E\left[\left|\sum_{k=0}^n U_k Y_k\right|\right] \leq E\left[\sum_{k=0}^n |U_k Y_k|\right] = \sum_{k=0}^n E[|U_k|]E[|Y_k|] = \sum_{k=0}^n E[|U_k|] < \infty .$$

Inoltre, sulla base della Proposizione 5.2.1, dal momento che la σ -algebra generata da Y_n è indipendente dalla σ -algebra \mathcal{F}_{n-1} e che U_n è misurabile rispetto a \mathcal{F}_{n-1} , si ha *q.c.*

$$\begin{aligned} E[M_n | \mathcal{F}_{n-1}] &= E[M_{n-1} + U_n Y_n | \mathcal{F}_{n-1}] = E[M_{n-1} | \mathcal{F}_{n-1}] + E[U_n Y_n | \mathcal{F}_{n-1}] \\ &= M_{n-1} + U_n E[Y_n | \mathcal{F}_{n-1}] = M_{n-1} + U_n E[Y_n] = M_{n-1} + U_n(2p - 1) . \end{aligned}$$

Quindi, a prescindere dalla strategia di gioco adottata dal giocatore, il p.a. M è una martingala se $p = \frac{1}{2}$, una super-martingala se $p < \frac{1}{2}$ e una sub-martingala se $p > \frac{1}{2}$. Di conseguenza, per qualsiasi strategia di gioco considerata, un gioco equo rimane equo, un gioco sfavorevole rimane sfavorevole e un gioco favorevole rimane favorevole. Questo esempio mostra che è impossibile determinare una strategia di gioco per cui un gioco d'azzardo sfavorevole può essere trasformato in un gioco favorevole. \square

Di seguito vengono dati due esempi di importanti martingale. Ulteriori esempi di martingale, super-martingale e sub-martingale verranno considerati nell'analisi delle passeggiate aleatorie nella Sezione 9.3 e del moto Browniano nella Sezione 9.4.

• **Esempio 9.2.3.** (Martingala di Doob) Si consideri una v.a. X con $E[|X|] < \infty$ e si consideri il p.a. $M = (M_t)_{t \in \mathbb{T}}$ tale che

$$M_t = E[X | \mathcal{F}_t] ,$$

dove $(\mathcal{F}_t)_{t \in \mathbb{T}}$ è una filtrazione. Tenendo presente la disuguaglianza di Jensen condizionata (Teorema 5.2.3), per ogni $s < t$ si ha

$$E[|M_t|] = E[|E[X | \mathcal{F}_t]|] \leq E[E[|X| | \mathcal{F}_t]] = E[|X|] < \infty .$$

Inoltre, sulla base del Teorema 5.2.6 si ha $\{E[E[X | \mathcal{F}_t] | \mathcal{F}_s] = E[X | \mathcal{F}_s]\}$ *q.c.* se $\mathcal{F}_s \subset \mathcal{F}_t$, da cui si ha *q.c.*

$$E[M_t | \mathcal{F}_s] = E[E[X | \mathcal{F}_t] | \mathcal{F}_s] = E[X | \mathcal{F}_s] = M_s .$$

Quindi, il p.a. M è una martingala, anche detta *martingala di Doob*. Questo esempio evidenzia una procedura pratica per costruire una martingala. \square

• **Esempio 9.2.4.** (Urna di Pólya) Sia data un'urna che contiene r palline rosse e b palline blu. Si consideri inoltre un esperimento aleatorio nel quale si estrae una pallina dall'urna e, dopo averne verificato il colore, si reimmette la pallina nell'urna con un'ulteriore pallina dello stesso colore. La procedura viene successivamente ripetuta con le stesse modalità. Questo esperimento aleatorio è detto *schema dell'urna di Pólya* dal momento che è stato proposto dal matematico George Pólya (1887-1985). Se X_n è la v.a. che rappresenta il numero di palline rosse nell'urna alla n -esima ripetizione dell'esperimento, si consideri la successione di v.a. $(X_n)_{n \geq 1}$. Si osservi che si può scrivere $X_n = r + \sum_{k=1}^n Z_k$, dove ogni v.a. Z_k è tale che $Z_k = 1$ se la pallina estratta alla k -esima estrazione è rossa e $Z_k = 0$ altrimenti. Dalla costruzione dell'esperimento aleatorio risulta immediato verificare che

$$E[Z_k | X_{k-1} = x] = P(Z_k = 1 | X_{k-1} = x) = \frac{x}{k - 1 + r + b} ,$$

dove $x = r, r + 1, \dots, r + k - 1$. Si consideri dunque il p.a. $M = (M_n)_{n \geq 0}$ tale che $M_n = \frac{X_n}{n+r+b}$ e $M_0 = \frac{r}{r+b}$ con $(\mathcal{F}_n)_{n \geq 0}$ filtrazione naturale. Per ogni n si ha

$$\mathbb{E}[|M_n|] = \mathbb{E}[M_n] = \frac{1}{n+r+b} \mathbb{E}[X_n] = \frac{1}{n+r+b} \left(r + \sum_{k=1}^n \mathbb{E}[Z_k] \right) \leq \frac{n+r}{n+r+b} < 1.$$

Inoltre, tenendo presente la Proposizione 5.2.1 e che $\mathbb{E}[Z_n | \mathcal{F}_{n-1}] = \mathbb{E}[Z_n | X_{n-1}]$, si ha *q.c.*

$$\begin{aligned} \mathbb{E}[M_n | \mathcal{F}_{n-1}] &= \frac{1}{n+r+b} (\mathbb{E}[X_{n-1} | \mathcal{F}_{n-1}] + \mathbb{E}[Z_n | \mathcal{F}_{n-1}]) \\ &= \frac{1}{n+r+b} \left(X_{n-1} + \frac{X_{n-1}}{n-1+r+b} \right) = \frac{1}{n-1+r+b} X_{n-1} = M_{n-1}. \end{aligned}$$

Procedendo in modo iterativo, per $m \in \mathbb{N}$ con $m < n$, si ha dunque *q.c.*

$$\mathbb{E}[X_n | \mathcal{F}_m] = \frac{1}{m+r+b} X_m = M_m,$$

ovvero dalla Definizione 9.2.1 il p.a. M è una martingala. \square

Di seguito viene introdotta la definizione di tempo di arresto. Anche se il tempo di arresto può essere considerato in generale per un qualsiasi p.a., questo concetto risulta particolarmente importante nell'ambito delle martingale.

Definizione 9.2.2. Si consideri lo spazio di probabilità (Ω, \mathcal{F}, P) . Se $(\mathcal{F}_n)_{n \geq 0}$ è una filtrazione, un *tempo di arresto discreto* è un'applicazione $\tau : \Omega \rightarrow \mathbb{N} \cup \{\infty\}$, tale che l'evento $\{\tau \leq n\} \in \mathcal{F}_n$ per ogni $n \in \mathbb{N}$. Inoltre, se $(\mathcal{F}_t)_{t \in [0, \infty[}$ è una filtrazione, un *tempo di arresto continuo* è un'applicazione $\tau : \Omega \rightarrow [0, \infty[\cup \{\infty\}$, tale che l'evento $\{\tau \leq t\} \in \mathcal{F}_t$ per ogni $t \in [0, \infty[$. \square

Intuitivamente, si può pensare ad un tempo d'arresto come al momento in cui si prende una decisione relativa ad un fenomeno aleatorio nel tempo (per esempio, la decisione di raddoppiare la scommessa ad un certo istante in una serie di giocate ad un gioco d'azzardo). I vincoli sugli eventi, ovvero $\{\tau \leq n\} \in \mathcal{F}_n$ e $\{\tau \leq t\} \in \mathcal{F}_t$, si basano sul fatto che tale decisione deve dipendere solo dalle informazioni disponibili fino a quel momento. Si osservi infine che un tempo di arresto non è in effetti una v.a. in senso proprio in quanto potrebbe sussistere che $P(\tau < \infty) < 1$, ovvero il tempo di arresto potrebbe non essere finito.

• **Esempio 9.2.5.** Si consideri di nuovo il p.a. $M = (M_n)_{n \geq 0}$ introdotto nell'Esempio 9.2.2 e sia dato il tempo di arresto

$$\tau = \min(\{n \in \mathbb{N} : Y_n = 1\}).$$

Nell'interpretazione in termini di gioco d'azzardo, il tempo di arresto τ può essere visto come l'istante in cui il giocatore ottiene la prima vincita. Si noti che τ è in effetti un tempo di arresto, dal momento che \mathcal{F}_n è la σ -algebra generata dal v.v.a. $(Y_1, \dots, Y_n)^T$ e

$$\{\tau = n\} = \{Y_1 = -1, \dots, Y_{n-1} = -1, Y_n = 1\} \in \mathcal{F}_n,$$

per cui

$$\{\tau \leq n\} = \bigcup_{k=1}^n \{\tau = k\} \in \mathcal{F}_n.$$

Per quanto visto nella Sezione 6.3, la legge della v.a. $(\tau - 1)$ è Geometrica di parametro p , ovvero la relativa f.p. è data da

$$p_\tau(x) = pq^{x-1} \mathbf{1}_{\{1,2,\dots\}}(x).$$

Dunque, in questo caso si ha $P(\tau < \infty) = 1$. □

Definizione 9.2.3. Si consideri lo spazio di probabilità (Ω, \mathcal{F}, P) e sia $(\mathcal{F}_t)_{t \in \mathbb{T}}$ una filtrazione. In questo caso

$$\mathcal{F}_\tau = \{E \in \mathcal{F} : E \cap \{\tau < t\}, t \in \mathbb{T}\}$$

è detta σ -algebra associata al tempo di arresto τ . Inoltre, se $M = (M_t)_{t \in \mathbb{T}}$ è una martingala e τ è un tempo di arresto, allora posto $M_\tau = M(\tau(\omega), \omega)$, il p.a. $M^\tau = (M_{\tau \wedge t})_{t \in \mathbb{T}}$ è detta *martingala arrestata al tempo di arresto* τ .

Il prossimo Teorema riveste un'importanza notevole, dal momento che evidenzia che una martingala arrestata conserva ancora le caratteristiche di martingala.

Teorema 9.2.4. (Teorema del campionamento opzionale di Doob) Sia $M = (M_t)_{t \in \mathbb{T}}$ una martingala e τ_1 e τ_2 tempi di arresto tali che $\{\tau_1 \leq \tau_2 < T\}$ q.c. dove $T > 0$. Si ha q.c.

$$E[M_{\tau_2} | \mathcal{F}_{\tau_1}] = M_{\tau_1}.$$

In particolare, il p.a. $M^\tau = (M_{\tau \wedge t})_{t \in \mathbb{T}}$ è una martingala. Inoltre, se M è una super-martingala, si ha q.c.

$$E[M_{\tau_2} | \mathcal{F}_{\tau_1}] \leq M_{\tau_1}$$

e il p.a. M^τ è una super-martingala, mentre se M è una sub-martingala, si ha q.c.

$$E[M_{\tau_2} | \mathcal{F}_{\tau_1}] \geq M_{\tau_1}$$

e il p.a. M^τ è una sub-martingala.

Dimostrazione. Si veda Revuz e Yor (2005, p.123). □

• **Esempio 9.2.6.** (Strategia di gioco con martingala) Si consideri il p.a. $M = (M_n)_{n \geq 0}$ introdotto nell'Esempio 9.2.2 e si assuma inoltre una successione di v.a. $(U_n)_{n \geq 1}$ tale che

$$U_n = 2^{n-1} \prod_{k=1}^{n-1} \mathbf{1}_{\{-1\}}(Y_k).$$

Si consideri inoltre il tempo d'arresto τ introdotto nell'Esempio 9.2.5 e la martingala arrestata $M^\tau = (M_{\tau \wedge n})_{n \geq 0}$. Dal Teorema 9.2.4, si ha dunque che il p.a. M^τ è una martingala se $p = \frac{1}{2}$, una super-martingala se $p < \frac{1}{2}$ o una sub-martingala se $p > \frac{1}{2}$. Nell'interpretazione in un ambito di gioco d'azzardo, la strategia di gioco basata sulla successione di v.a. $(U_n)_{n \geq 1}$ è comunemente detta *martingala* (con una terminologia ovviamente fuorviante). In pratica, seguendo questa strategia, il giocatore scommette un'unità alla prima giocata. Nel caso in cui perda, il giocatore raddoppia la scommessa e continua in questo modo fino a quando non ottiene la prima vincita e conclude il gioco. Si noti che

$$\begin{aligned} M_{\tau \wedge n} &= \mathbf{1}_{\{1, \dots, n\}}(\tau) - \sum_{k=1}^n U_k \mathbf{1}_{\{n+1, n+2, \dots\}}(\tau) = \mathbf{1}_{\{1, \dots, n\}}(\tau) - \sum_{k=1}^n 2^{k-1} \mathbf{1}_{\{n+1, n+2, \dots\}}(\tau) \\ &= \mathbf{1}_{\{1, \dots, n\}}(\tau) - (2^n - 1) \mathbf{1}_{\{n+1, n+2, \dots\}}(\tau). \end{aligned}$$

In pratica, $M_{\tau \wedge n} = M_\tau = 1$ se $\tau = n$ (ovvero il giocatore guadagna un'unità alla prima vincita concludendo il gioco), mentre $M_{\tau \wedge n} = M_n = -(2^n - 1)$ se $\tau > n$ (ovvero il giocatore perde $2^n - 1$ unità all' n -esima giocata precedente la prima vincita). Dunque, poichè $P(\tau < \infty) = 1$, si ha anche $P(M_\tau = 1) = 1$. Dal momento che questo risultato vale per ogni p , la martingala sembra apparentemente una strategia di gioco vincente (il giocatore termina il gioco con un'unità vinta). Tuttavia, si deve tenere presente che questa strategia presuppone una disponibilità illimitata di denaro e di tempo da parte del giocatore (ovviamente un'assunzione irrealistica). Inoltre, tenendo presente l'Esempio 9.2.5, si ha

$$E[M_{\tau-1}] = \sum_{n=1}^{\infty} M_{n-1} P(\tau = n) = - \sum_{n=1}^{\infty} (2^{n-1} - 1) p q^{n-1} = -p \sum_{n=0}^{\infty} (2q)^n + 1.$$

Dunque, risulta $E[M_{\tau-1}] = -\infty$ per $q \geq \frac{1}{2}$, ovvero per $p \leq \frac{1}{2}$. Quindi, in questo caso la perdita attesa nella giocata prima della vincita finale è infinita. Sorprendentemente, questo risultato si ha anche per un gioco equo, ovvero anche nel caso in cui $p = \frac{1}{2}$. \square

Teorema 9.2.5. (Disuguaglianze di Doob) Sia $M = (M_n)_{t \in \mathbb{T}}$ una martingala. Per ogni $T > 0$ e $\lambda > 0$ vale la disuguaglianza

$$P\left(\sup_{t \in [0, T]} |M_t| \geq \lambda\right) \leq \frac{1}{\lambda} E[M_T \vee 0] \leq \frac{1}{\lambda} E[|M_T|].$$

Inoltre, per $p > 1$ e $q = \frac{p}{p-1}$ vale la disuguaglianza

$$E\left[\sup_{t \in [0, T]} |M_t|^p\right] \geq q^p E[|M_T|^p].$$

Dimostrazione. Si veda Pascucci (2011, p.113). \square

9.3. Passeggiate aleatorie

La passeggiata aleatoria è uno dei p.a. più semplici (anche se fondamentali) nella teoria della probabilità. Più esattamente, la passeggiata aleatoria è un processo a tempo discreto che viene formalmente definito come segue.

Definizione 9.3.1. Sia (Ω, \mathcal{F}, P) con $(\mathcal{F}_n)_{n \geq 0}$ uno spazio di probabilità con filtrazione. Sia inoltre $(Y_n)_{n \geq 1}$ una successione di v.a. indipendenti con la medesima legge, tali che $P(Y_n = 1) = p$ e $P(Y_n = -1) = q$ con $q = 1 - p$. Si dice *passeggiata aleatoria* il p.a. $X = (X_n)_{n \geq 0}$ dove $X_n = \sum_{k=1}^n Y_k$ e $\{X_0 = 0\}$ q.c. \square

Evidentemente, vista l'indipendenza delle componenti della successione $(Y_n)_{n \geq 1}$, si può scegliere la filtrazione naturale $(\mathcal{F}_n)_{n \geq 0}$ tale che $\mathcal{F}_n = \sigma(\{Y_1, \dots, Y_n\})$. I tempi $n \in \mathbb{N}$ sono anche detti *passi* della passeggiata aleatoria. Inoltre, nel caso particolare in cui $p = \frac{1}{2}$, la passeggiata aleatoria è detta *simmetrica*. La seguente Proposizione fornisce la legge del generico stato della passeggiata aleatoria.

Proposizione 9.3.2. Sia $X = (X_n)_{n \geq 0}$ una passeggiata aleatoria. Per ogni $n \in \mathbb{N}$ la f.p. dello stato X_n è data da

$$p_{X_n}(x) = \binom{n}{\frac{1}{2}(n+x)} p^{\frac{1}{2}(n+x)} q^{\frac{1}{2}(n-x)} \mathbf{1}_{\{-n, 2-n, \dots, n-2, n\}}(x).$$

Inoltre, $E[X_n] = n(p - q)$ e $\text{Var}[X_n] = 4npq$.

Dimostrazione. Si osservi che sussiste $X_n = 2V_n - n$, dove $V_n = \sum_{k=1}^n Z_k$ e $(Z_n)_{n \geq 1}$ è una successione di v.a. indipendenti con legge di Bernoulli di parametro p . Per quanto visto nella Sezione 6.1, la legge della v.a. V_n è Binomiale $\mathcal{B}(n, p)$. Dunque, dal momento che la v.a. X_n è una trasformata lineare della v.a. V_n è immediato ottenerne la relativa f.p. Inoltre, per le proprietà del valore atteso si ha

$$E[X_n] = 2E[V_n] - n = 2np - n = n(p - q)$$

e

$$\text{Var}[X_n] = 4\text{Var}[V_n] = 4npq. \quad \square$$

• **Esempio 9.3.1.** Si consideri il gioco d'azzardo della roulette introdotto nell'Esempio 2.6.3. Si supponga inoltre un giocatore ostinato (con disponibilità illimitate) che scommette un'unità sul rosso ad ogni giocata. La successione delle somme vinte (o perdute) dal giocatore è una passeggiata aleatoria con $p = \frac{18}{37}$ se la roulette è di tipo Monte Carlo (evitando di considerare per semplicità le regole basate su *la partage* o *en prison* quando si presenta lo zero). Sulla base della Proposizione 9.3.2, la probabilità di essere in attivo dopo $n = 20$ giocate è data da $\sum_{k=1}^{10} p_{X_{20}}(2k) = 0.365$, quella di essere in pareggio è data da $p_{X_{20}}(0) = 0.175$ e quella di essere in perdita è data da $\sum_{k=-1}^{-10} p_{X_{20}}(2k) = 0.460$. Inoltre, le medesime quantità dopo $n = 100$ giocate sono rispettivamente $\sum_{k=1}^{50} p_{X_{100}}(2k) = 0.355$, $p_{X_{100}}(0) = 0.077$ e $\sum_{k=-50}^{-1} p_{X_{100}}(2k) = 0.568$. \square

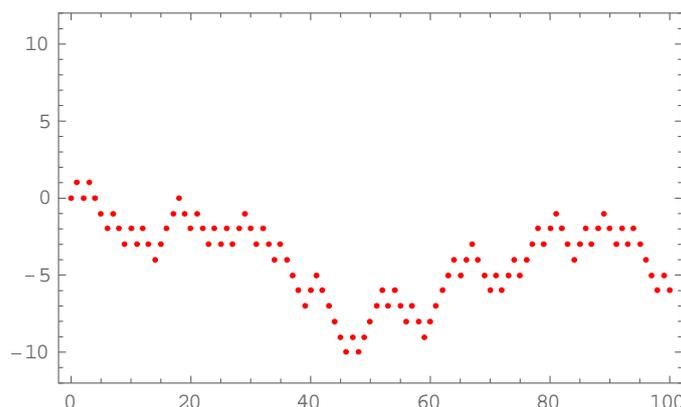


Figura 9.3.1. Una traiettoria della passeggiata aleatoria con $p = \frac{18}{37}$ e $n = 100$ passi.

La seguente Proposizione evidenzia che la passeggiata aleatoria è una martingala, una sub-martingala o una super-martingala a secondo dei valori assunti da p .

Proposizione 9.3.3. *La passeggiata aleatoria $X = (X_n)_{n \geq 0}$ è una martingala se $p = \frac{1}{2}$, una super-martingala se $p < \frac{1}{2}$ e una sub-martingala se $p > \frac{1}{2}$.*

Dimostrazione. Sulla base della Definizione 9.3.1, dal momento che $E[|Y_k|] = 1$ si ha

$$E[|X_n|] = E\left[\left|\sum_{k=1}^n Y_k\right|\right] \leq \sum_{k=1}^n E[|Y_k|] = n < \infty.$$

Inoltre, tenendo presente che $\{E[X_n | \mathcal{F}_n] = X_n\}$ q.c. sulla base della Proposizione 5.2.1 e che la σ -algebra generata da Y_n è indipendente dalla σ -algebra \mathcal{F}_{n-1} , si ha q.c.

$$E[X_n | \mathcal{F}_{n-1}] = E[X_{n-1} | \mathcal{F}_{n-1}] + E[Y_n | \mathcal{F}_{n-1}] = X_{n-1} + E[Y_n] = X_{n-1} + 2p - 1.$$

In modo analogo alla precedente relazione, si ha q.c.

$$\begin{aligned} E[X_n | \mathcal{F}_{n-2}] &= E[X_{n-2} | \mathcal{F}_{n-2}] + E[Y_{n-1} | \mathcal{F}_{n-2}] + E[Y_n | \mathcal{F}_{n-2}] \\ &= X_{n-2} + E[Y_{n-1}] + E[Y_n] = X_{n-2} + 2(2p - 1). \end{aligned}$$

Procedendo in modo iterativo, per $m \in \mathbb{N}$ con $m < n$, si ha dunque *q.c.*

$$E[X_n | \mathcal{F}_m] = X_m + (n - m)(2p - 1).$$

Dalla precedente espressione e dalla Definizione 9.2.1 segue il risultato. \square

• **Esempio 9.3.2.** Si consideri il gioco d'azzardo della testa e croce introdotto nell'Esempio 1.1.1 e si supponga di nuovo un giocatore ostinato (con disponibilità illimitate) che scommette un'unità sulla croce ad ogni giocata. Se la moneta è bilanciata, ovvero $p = \frac{1}{2}$, allora la successione delle somme vinte (o perdute) dal giocatore è una passeggiata aleatoria che è anche una martingala. La testa e croce è dunque un gioco equo. Nel caso della roulette considerato nell'Esempio 9.3.1, la passeggiata aleatoria è invece una super-martingala, dal momento che $p = \frac{18}{37} < \frac{1}{2}$. La roulette non è dunque un gioco d'azzardo vantaggioso. \square

Risulta interessante ottenere alcune proprietà della traiettoria di una passeggiata aleatoria. In questo ambito, è conveniente introdurre la seguente definizione.

Definizione 9.3.4. Sia $X = (X_n)_{n \geq 0}$ una passeggiata aleatoria. Posto $n = 2k$ con $k \in \mathbb{N}$, la quantità $u_n = P(X_n = 0)$ dove

$$u_{2k} = P(X_{2k} = 0) = \binom{2k}{k} (pq)^k,$$

mentre $u_{2k+1} = P(X_{2k+1} = 0) = 0$ è detta *probabilità di ritorno all'origine* in n passi. \square

Dunque, u_n non è nulla solo se n è pari. La seguente Proposizione fornisce alcune caratteristiche della successione delle probabilità di ritorno all'origine.

Proposizione 9.3.5. La successione $u = (u_n)_{n \geq 0}$ ammette funzione generatrice data da

$$G_u(t) = \sum_{n=0}^{\infty} t^n u_n = \frac{1}{\sqrt{1 - 4pqt^2}}.$$

Inoltre, posto $n = 2k$ con $k \in \mathbb{N}$, per $k \rightarrow \infty$ si ha

$$u_{2k} \sim \frac{1}{\sqrt{\pi k}} (4pq)^k.$$

Dimostrazione. Dal momento che sussiste la serie

$$\sum_{k=0}^{\infty} \binom{2k}{k} x^k = \frac{1}{\sqrt{1 - 4x}},$$

la funzione generatrice della successione u è data da

$$G_u(t) = \sum_{n=0}^{\infty} t^n u_n = \sum_{k=0}^{\infty} t^{2k} u_{2k} = \sum_{k=0}^{\infty} \binom{2k}{k} (pqt^2)^k = \frac{1}{\sqrt{1 - 4pqt^2}}.$$

Inoltre, per $k \rightarrow \infty$ l'approssimazione di Stirling applicata al coefficiente binomiale nell'espressione di u_{2k} fornisce

$$\binom{2k}{k} \sim \frac{4^k}{\sqrt{\pi k}},$$

da cui si ottiene la seconda parte. □

Dalla Proposizione 9.3.5, per $p = \frac{1}{2}$ si ottiene in particolare che

$$u_{2k} \sim \frac{1}{\sqrt{\pi k}}.$$

Dunque, in modo controintuitivo, la probabilità di ritorno all'origine per una passeggiata aleatoria simmetrica tende ad annullarsi all'aumentare dei passi.

La seguente definizione introduce un particolare tempo di arresto nell'ambito delle passeggiate aleatorie.

Definizione 9.3.6. Sia $X = (X_n)_{n \geq 0}$ una passeggiata aleatoria. Il tempo di arresto

$$\tau_k = \min(\{n > 0 : X_n = k\}),$$

dove $k \in \mathbb{Z}$, è detto *tempo del primo passaggio per k* . In particolare,

$$\tau_0 = \min(\{n > 0 : X_n = 0\})$$

è detto *tempo del primo ritorno all'origine*. □

Evidentemente, il tempo del primo ritorno all'origine τ_0 può prendere valori solo sui numeri pari. La seguente Proposizione fornisce la distribuzione di τ_0 e le relative caratteristiche.

Proposizione 9.3.7. Sia $X = (X_n)_{n \geq 0}$ una passeggiata aleatoria. La funzione generatrice della successione $v = (v_n)_{n \geq 1}$ con $v_n = P(\tau_0 = n)$ è data da

$$G_v(t) = \sum_{n=1}^{\infty} t^n v_n = 1 - \sqrt{1 - 4pqt^2}.$$

Posto $n = 2k$ con $k \in \{1, 2, \dots\}$, si ha

$$v_{2k} = P(\tau_0 = 2k) = \binom{2k}{k} \frac{(pq)^k}{2k-1},$$

mentre $v_{2k-1} = P(\tau_0 = 2k-1) = 0$.

Dimostrazione. Gli eventi $\{\tau_0 = k\}$, $k = 1, 2, \dots, n$, sono incompatibili e dunque risulta

$$P(X_n = 0) = \sum_{k=1}^n P(X_n = 0 \mid \tau_0 = k) P(\tau_0 = k).$$

Inoltre, si ha $P(X_n = 0 \mid \tau_0 = k) = P(X_{n-k} = 0)$ dal momento che, se si verifica $\{\tau_0 = k\}$, allora $\sum_{j=k}^n Y_j$ può essere considerato come il passo di ordine $(n-k)$ di una nuova passeggiata aleatoria.

Dunque, dalla precedente espressione segue

$$u_n = \sum_{k=1}^n u_{n-k} v_k.$$

Quindi, si ha

$$G_u(t) = \sum_{n=0}^{\infty} t^n u_n = \sum_{n=0}^{\infty} \sum_{k=1}^n t^n u_{n-k} v_k .$$

Tenendo presente la definizione di convoluzione di funzioni generatrici e poichè $u_0 = 1$, dalla precedente espressione si ottiene

$$G_u(t) = 1 + G_u(t)G_v(t) ,$$

ovvero

$$G_v(t) = 1 - \frac{1}{G_u(t)}$$

e la prima parte segue dalla Proposizione 9.3.5. Inoltre, considerando l'espressione della serie binomiale, la funzione generatrice della successione v può essere scritta come

$$G_v(t) = 1 - \sum_{k=0}^{\infty} \binom{\frac{1}{2}}{k} (-4pqt^2)^k = \sum_{k=1}^{\infty} (-1)^{k-1} \binom{\frac{1}{2}}{k} (4pq)^k t^{2k} ,$$

da cui segue la seconda parte, tenendo presente l'identità

$$(-1)^{k-1} \binom{\frac{1}{2}}{k} 2^{2k} = \binom{2k}{k} \frac{1}{2k-1}$$

applicata all'espressione di v_{2k} . □

Si osservi che dalle Proposizioni 9.3.5 e 9.3.7 risulta in modo notevole

$$v_{2k} = \frac{1}{2k-1} u_{2k} ,$$

mentre per $k \rightarrow \infty$ si ha

$$v_{2k} \sim \frac{1}{\sqrt{4\pi k^3}} (4pq)^k .$$

Inoltre, si deve evidenziare che G_v non è in generale una funzione generatrice delle probabilità. In effetti, la successione v non costituisce in generale una legge essenziale, dal momento che dalla Proposizione 9.3.7 si ha

$$P(\tau_0 < \infty) = \sum_{n=1}^{\infty} v_n = G_v(1) = 1 - \sqrt{1 - 4pq} = 1 - \sqrt{(p-q)^2} = 1 - |p-q| \leq 1 .$$

• **Esempio 9.3.3.** Si consideri di nuovo il gioco d'azzardo della roulette. In questo caso, sulla base della Proposizione 9.3.7, la probabilità del tempo di ritorno all'origine dopo $n = 20$ giocate è data da $v_{20} = 0.009$. Inoltre, la probabilità che il tempo di ritorno all'origine sia finito risulta $P(\tau_0 < \infty) = \frac{36}{37}$. Dunque, in modo sorprendente per il giocatore ingenuo (anche se con disponibilità illimitate), la probabilità che il tempo di ritorno all'origine non sia finito è dato da $1 - P(\tau_0 < \infty) = \frac{1}{37}$. □

Dalla Proposizione 9.3.7 è evidente che $P(\tau_0 < \infty) = 1$ esclusivamente per una passeggiata aleatoria simmetrica. Dunque, τ_0 è una v.a. propria solamente per $p = \frac{1}{2}$. In questo caso, si ha il seguente risultato.

Proposizione 9.3.8. Sia $X = (X_n)_{n \geq 0}$ una passeggiata aleatoria simmetrica. La f.p. della v.a. τ_0 è data da

$$p_{\tau_0}(x) = \binom{x}{\frac{1}{2}x} \frac{2^{-x}}{x-1} \mathbf{1}_{\{2,4,\dots\}}(x)$$

e $E[\tau_0] = \infty$. Inoltre, posto $n = 2k$ con $k \in \mathbb{N}$ e data la successione $z = (z_n)_{n \geq 0}$ dove

$$z_{2k} = P(\tau_0 > 2k)$$

e $z_{2k+1} = 0$, si ha $z = u$.

Dimostrazione. Per $p = \frac{1}{2}$ la v.a. τ_0 possiede f.g. data da

$$G_{\tau_0}(t) = 1 - \sqrt{1-t^2}$$

e la prima parte segue immediatamente dalla Proposizione 9.3.7. Inoltre, si ha

$$G_{\tau_0}^{(1)}(t) = \frac{t}{\sqrt{1-t^2}},$$

e dunque

$$E[\tau_0] = \lim_{t \rightarrow 1^-} G_{\tau_0}^{(1)}(t) = \infty.$$

Per quanto riguarda l'ultima parte, poichè $z_{2k} = E[\mathbf{1}_{\{1,2,\dots\}}(\tau_0 - 2k)]$, la funzione generatrice della successione z è data da

$$\begin{aligned} G_z(t) &= \sum_{n=0}^{\infty} t^n z_n = \sum_{k=0}^{\infty} t^{2k} E[\mathbf{1}_{\{1,2,\dots\}}(\tau_0 - 2k)] = E\left[\sum_{k=0}^{\infty} t^{2k} \mathbf{1}_{\{1,2,\dots\}}(\tau_0 - 2k)\right] \\ &= E\left[\sum_{k=0}^{\frac{1}{2}\tau_0-1} t^{2k}\right] = \frac{1}{1-t^2} E[1-t^{\tau_0}] = \frac{1}{1-t^2} (1-G_{\tau_0}(t)). \end{aligned}$$

Tenendo presente l'espressione di $G_{\tau_0}(t)$ e dalla Proposizione 9.3.5, si ha

$$G_z(t) = \frac{1}{\sqrt{1-t^2}} = G_u(t),$$

da cui segue la tesi. □

Dalla Proposizione 9.3.8 risulta quindi che, pur essendo $P(\tau_0 < \infty) = 1$ per $p = \frac{1}{2}$, si ha comunque $E[\tau_0] = \infty$. Inoltre, dalla medesima Proposizione segue anche una relazione notevole fra le successioni u e z per una passeggiata aleatoria simmetrica. In effetti, da questa relazione risulta che la probabilità di non ritornare all'origine in $2k$ passi è pari alla probabilità di ritornare all'origine in $2k$ passi quando $p = \frac{1}{2}$, ovvero $P(\tau_0 > 2k) = P(X_{2k} = 0)$.

• **Esempio 9.3.4.** Si consideri il gioco d'azzardo della testa e croce analizzato nell'Esempio 9.3.2. In questo caso, la probabilità del tempo di ritorno all'origine dopo $n = 20$ giocate è data da $P(X_{20} = 0) = 0.176$, per cui risulta anche $P(\tau_0 > 20) = 0.176$. □

Il prossimo Teorema fornisce un risultato classico per la passeggiata aleatoria simmetrica, ovvero la cosiddetta *legge dell'arcoseno*. Se si considera i primi $2n$ passi di una passeggiata aleatoria, questa legge fornisce la distribuzione dell'ultimo tempo di ritorno all'origine.

Teorema 9.3.9. (Legge dell'arcoseno) Sia $X = (X_n)_{n \geq 0}$ una passeggiata aleatoria simmetrica. Si consideri la v.a. $\varrho_n = \max(\{l \in \{1, \dots, n\} : X_l = 0\})$, detta ultimo tempo di ritorno all'origine in n passi. Posto $n = 2k$ con $k \in \mathbb{N}$, se $\alpha_{2l,2k} = P(\varrho_{2k} = 2l)$, si ha

$$\alpha_{2l,2k} = u_{2l}u_{2k-2l} = \binom{2l}{l} \binom{2k-2l}{k-l} 2^{-2k},$$

per $l = 0, \dots, k$, ovvero la f.p. della v.a. ϱ_n è data da

$$p_{\varrho_n}(x) = \binom{x}{\frac{1}{2}x} \binom{n-x}{\frac{1}{2}(n-x)} 2^{-n} \mathbf{1}_{\{0,2,\dots,n\}}(x).$$

Inoltre, risulta

$$\lim_n P(n^{-1}\varrho_n \leq x) = \frac{2}{\pi} \arcsin(\sqrt{x}) \mathbf{1}_{[0,1]}(x) + \mathbf{1}_{]1,\infty[}(x),$$

ovvero $n^{-1}\varrho_n \xrightarrow{\mathcal{L}} Z$ per $n \rightarrow \infty$, dove Z è una v.a. con legge Beta di tipo $\mathcal{BE}(0, 1, \frac{1}{2}, \frac{1}{2})$ che ammette d.p. data da

$$f_Z(x) = \frac{1}{\pi \sqrt{x(1-x)}} \mathbf{1}_{[0,1]}(x).$$

Dimostrazione. Per quanto riguarda la prima parte, tenendo presente che se si verifica $\{X_{2l} = 0\}$, allora i successivi passi possono essere considerati come una nuova passeggiata aleatoria, dalla Proposizione 9.3.8 si ha

$$\begin{aligned} \alpha_{2l,2k} &= P(X_{2l} = 0, X_{2l+2} \neq 0, \dots, X_{2k} \neq 0) = P(X_{2l} = 0)P(X_{2l+2} \neq 0, \dots, X_{2k} \neq 0 \mid X_{2l} = 0) \\ &= P(X_{2l} = 0)P(X_{2l+2} \neq 0, \dots, X_{2k} \neq 0) = P(X_{2l} = 0)P(\tau_0 > 2k - 2l) = u_{2l}u_{2k-2l}. \end{aligned}$$

Per quanto riguarda la seconda parte, dalla Proposizione 9.3.5 si ha

$$\alpha_{2l,2k} \sim \frac{1}{\pi \sqrt{l(k-l)}}.$$

Dunque, posto $x = \frac{l}{k}$ per $l = 0, \dots, k$, si ha

$$P(n^{-1}\varrho_n \leq x) \sim \sum_{j=0}^{kx} \frac{1}{\pi \sqrt{j(k-j)}},$$

da cui, tenendo presente la definizione dell'integrale di Riemann, si ottiene

$$\lim_n P(n^{-1}\varrho_n \leq x) = \lim_k \sum_{j=0}^{kx} \frac{1}{\pi \sqrt{j(k-j)}} = \int_0^x \frac{1}{\pi \sqrt{u(1-u)}} du.$$

La seconda parte segue dunque dalla definizione di convergenza in legge e dai risultati della Sezione 6.9. \square

Si osservi che $\alpha_{2l,2k} = \alpha_{2k-2l,2k}$. Inoltre, si deve sottolineare che ϱ_n non è un tempo di arresto. La legge dell'arcoseno fornisce un risultato sulla legge della v.a. ϱ_n che può risultare controintuitiva. In effetti, la f.p. della v.a. ϱ_n mostra la cosiddetta “forma ad U” (si veda la Figura 9.3.2), così come la d.p. della v.a. limite della successione di v.a. $(n^{-1}\varrho_n)_{n \geq 1}$ mostra la medesima morfologia. In effetti, anche se $E[\varrho_n] = \frac{1}{2}n$, la probabilità del tempo dell'ultimo ritorno all'origine in n passi è molto più elevata per valori prossimi a 0 e a n piuttosto che a $\frac{1}{2}n$. Dal momento che la v.a. limite può essere

interpretata come la proporzione di tempo dall'ultimo ritorno all'origine, la legge dell'arccoseno è ancora più sorprendente per $n \rightarrow \infty$.

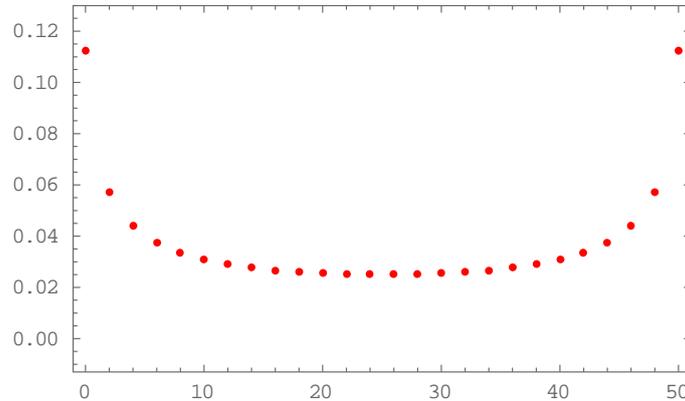


Figura 9.3.2. Funzione di probabilità per la legge dell'arccoseno con $n = 50$.

Per concludere, viene considerato un risultato classico nella Teoria della Probabilità, ovvero la cosiddetta *rovina del giocatore d'azzardo*. In questo problema, un giocatore ostinato con capitale iniziale pari ad a unità fronteggia in giocate ripetute un secondo giocatore con capitale iniziale pari a b unità. Il gioco continua fino a quando il primo giocatore esaurisce il proprio capitale o riesce a vincere il capitale del secondo giocatore. In questo caso, l'obiettivo è quello di determinare la probabilità di rovina, che viene ottenuta nel seguente risultato.

Teorema 9.3.10. (La rovina del giocatore d'azzardo) Sia $X = (X_n)_{n \geq 0}$ una passeggiata aleatoria. Se $a, b \in \mathbb{N}$, si consideri il tempo di arresto $\tau = \min(\tau_{-a}, \tau_b)$. Si ha $P(\tau < \infty) = 1$. Inoltre, per $p \neq \frac{1}{2}$

$$P(X_\tau = -a) = 1 - P(X_\tau = b) = \frac{1 - \rho^b}{\rho^{-a} - \rho^b},$$

dove $\rho = \frac{q}{p}$, mentre se $p = \frac{1}{2}$

$$P(X_\tau = -a) = 1 - P(X_\tau = b) = \frac{b}{a + b}.$$

Dimostrazione. Per quanto riguarda la prima parte, si ponga $c = a + b$ e si definisca il tempo di arresto

$$Z = \min(\{n \in \{1, 2, \dots\} : Y_{(n-1)c+1} = 1, \dots, Y_{(n-1)c+c} = 1\}).$$

Il tempo di arresto Z è quindi dato dal primo passo della passeggiata dopo il quale si verificano c vincite del primo giocatore. Dal momento che per n fissato si ha

$$P(Y_{(n-1)c+1} = 1, \dots, Y_{(n-1)c+c} = 1) = p^c,$$

la legge della v.a. Z è Geometrica di parametro p^c e dunque $E[Z] < \infty$. Inoltre, $\{\tau < cZ\}$ q.c. dal momento che τ si deve necessariamente realizzare nell'insieme $\{(n-1)c+1, \dots, (n-1)c+c\}$ se $Z = n$. Dunque, risulta $E[\tau] < cE[Z] < \infty$, da cui si ha anche $P(\tau < \infty) = 1$. Per quanto riguarda la seconda parte per $p \neq \frac{1}{2}$, si noti che $M = (M_n)_{n \geq 0}$ con $M_n = \rho^{X_n}$ è una martingala. In effetti, dal momento che $E[\rho^{Y_k}] = 1$ per $k = 1, \dots, n$, si ha

$$E[|M_n|] = E\left[\prod_{k=1}^n \rho^{Y_k}\right] = \prod_{k=1}^n E[\rho^{Y_k}] = 1 < \infty,$$

mentre risulta *q.c.*

$$E[M_n | \mathcal{F}_{n-1}] = E[\rho^{Y_n} M_{n-1} | \mathcal{F}_{n-1}] = E[\rho^{Y_n}] E[M_{n-1} | \mathcal{F}_{n-1}] = M_{n-1}$$

e dunque, in modo iterativo, per $m \in \mathbb{N}$ con $m < n$, si ha *q.c.*

$$E[X_n | \mathcal{F}_m] = X_m.$$

Quindi, dal Teorema 9.2.4 si ha che

$$E[M_\tau] = E[M_0] = 1 < \infty.$$

Dal momento che i due eventi

$$E = \{X_\tau = -a\} = \{M_\tau = \rho^{-a}\} = \{\tau_{-a} < \tau_b\}$$

e

$$F = \{X_\tau = b\} = \{M_\tau = \rho^b\} = \{\tau_{-a} > \tau_b\},$$

sono incompatibili ed esaustivi, si ha

$$E[M_\tau] = E[M_\tau | E]P(E) + E[M_\tau | F]P(F) = \rho^{-a}P(X_\tau = -a) + \rho^bP(X_\tau = b) = 1$$

e la tesi segue tenendo presente che $P(X_\tau = -a) + P(X_\tau = b) = 1$. Nel caso che $p = \frac{1}{2}$, dalla Proposizione 9.3.3 segue che X è una martingala. Dunque, dal Teorema 9.2.4 si ha che

$$E[X_\tau] = E[X_0] = 0 < \infty.$$

In modo simile a quanto visto in precedenza, si ha

$$E[X_\tau] = E[X_\tau | E]P(E) + E[X_\tau | F]P(F) = (-a)P(X_\tau = -a) + bP(X_\tau = b) = 0,$$

da cui segue la seconda parte. \square

Dal precedente Teorema, si evidenzia che il gioco termina *q.c.* dal momento che $P(\tau < \infty) = 1$ a prescindere dai valori di a e b e di p . Inoltre, nel caso che il secondo giocatore sia infinitamente ricco, ovvero se $b \rightarrow \infty$, allora si ha

$$P(X_\tau = -a) = 1,$$

se $\rho \geq 1$, mentre

$$P(X_\tau = -a) = \rho^a,$$

se $\rho < 1$. Dunque, sorprendentemente, la rovina del primo giocatore contro un giocatore infinitamente ricco avviene *q.c.* anche se il gioco è equo.

• **Esempio 9.3.5.** Si consideri di nuovo il gioco d'azzardo della roulette. Come è stato visto nell'Esempio 9.3.2, la roulette non è un gioco equo essendo $p = \frac{18}{37}$ e inoltre la casa da gioco dispone solitamente di capitali molto più elevati dei giocatori, ovvero b è molto più grande di a . Quindi, $P(X_\tau = -a)$ è prossimo ad uno nei casi usuali. Tuttavia, supponendo che $b = a$, ovvero che la casa da gioco disponga dello stesso capitale del giocatore ostinato, per $a = 50$ si ha $P(X_\tau = -50) = 0.937$ e per $a = 100$ si ha $P(X_\tau = -100) = 0.996$. Inoltre, supponendo che $b = \frac{a}{2}$, ovvero che la casa da gioco disponga della metà del capitale rispetto al giocatore ostinato, per $a = 50$ si ha $P(X_\tau = -50) = 0.754$ e per $a = 100$ si ha $P(X_\tau = -100) = 0.933$. Dunque, anche se il gioco della roulette è solo leggermente sfavorevole, questo fatto conferisce comunque un vantaggio determinante alla casa da gioco. \square

9.4. Moto Browniano

Il moto Browniano è uno dei p.a. fondamentali nella Teoria della Probabilità. Questo p.a. prende nome dal botanico Robert Brown (1773-1858), che per primo aveva notato il moto erratico delle particelle nei fluidi. Il probabilista Norbert Wiener (1894-1964) ha determinato numerose proprietà del relativo modello matematico e per questo motivo il p.a. è anche detto processo di Wiener.



Figura 9.4.1. Norbert Wiener (1894-1964).

Definizione 9.4.1. Sia (Ω, \mathcal{F}, P) con $(\mathcal{F}_t)_{t \in [0, \infty[}$ uno spazio di probabilità con filtrazione. Il *moto Browniano* o *processo di Wiener* è un p.a. $B = (B_t)_{t \in [0, \infty[}$ tale che

- i) $P(B_0 = 0) = 1$;
- ii) per ogni $0 \leq s < t$, la legge della v.a. $B_t - B_s$ è Normale $\mathcal{N}(0, t - s)$;
- iii) il p.a. B ha incrementi indipendenti, ovvero per ogni $0 \leq t_1 < \dots < t_n$ le v.a. $B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_n} - B_{t_{n-1}}$ sono indipendenti;
- iv) il p.a. B ha traiettorie continue *q.c.* □

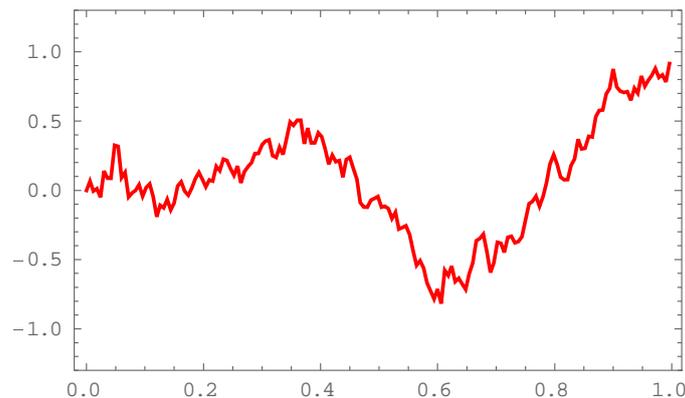


Figura 9.4.2. Una traiettoria del moto Browniano.

Per quanto riguarda la costruzione del moto Browniano, si può consultare Kuo (2006, p.23), dove a tale fine sono analizzati tre principali metodi. Il moto Browniano può essere anche ottenuto come limite in legge di una versione generale di passeggiata aleatoria “scalata” i cui passi tendono a zero. Più esattamente, si ha il seguente Teorema che prende il nome dal matematico Monroe David Donsker (1924-1991).

Teorema 9.4.2. (Teorema di Donsker) Sia (Ω, \mathcal{F}, P) con $(\mathcal{F}_t)_{t \in [0, \infty[}$ uno spazio di probabilità con filtrazione. Sia inoltre $(Y_n)_{n \geq 1}$ una successione di v.a. indipendenti con la medesima legge, tali che $E[Y_n] = 0$ e $\text{Var}[Y_n] = 1$ e sia $X_n = \sum_{k=1}^n Y_k$ e $X_0 = 0$ q.c. Se $Z_n = (Z_{n,t})_{t \in [0,1]}$ è una passeggiata aleatoria scalata tale che

$$Z_{n,t} = \frac{1}{\sqrt{n}} (X_{[nt]} + (nt - [nt])Y_{[nt]+1}),$$

dove $[x]$ rappresenta il più grande intero minore o uguale a x , allora $Z_n \xrightarrow{\mathcal{L}} B$ per $n \rightarrow \infty$, dove $B = (B_t)_{t \in [0,1]}$.

Dimostrazione. Si veda Capasso e Bakstein (2015, p.498). \square

In particolare, il Teorema di Donsker vale per la passeggiata aleatoria simmetrica, per cui si ha $P(Y_n = 1) = P(Y_n = -1) = \frac{1}{2}$. Il Teorema di Donsker può essere considerato come la versione funzionale del Teorema Centrale del Limite (Teorema 8.4.1). Molte proprietà del moto Browniano possono essere quindi ottenute nel “limite” da quelle della passeggiata aleatoria simmetrica (si veda Bhattacharya e Waymire, 2016).

Di seguito vengono considerate una serie di Proposizioni che forniscono alcune delle principali proprietà del moto Browniano.

Proposizione 9.4.3. Sia dato il moto Browniano $B = (B_t)_{t \in [0, \infty[}$. Per ogni $t > 0$, la v.a. B_t si distribuisce con legge Normale $\mathcal{N}(0, t)$. Inoltre, per ogni $s, t \geq 0$ si ha $\text{Cov}[B_s, B_t] = \min(s, t)$.

Dimostrazione. Dall'assunzione i) della Definizione 9.4.1 si ottiene che $B_t \stackrel{\mathcal{L}}{=} B_t - B_0$ e dunque la prima parte segue dalle assunzioni i) e ii) della Definizione 9.4.1. Per quanto riguarda la seconda parte, banalmente si ha che $B_t \stackrel{\mathcal{L}}{=} (B_t - B_s) + B_s$ per $s < t$. Tenendo presente l'assunzione iii) della Definizione 9.4.1 si ha inoltre $E[(B_t - B_s)B_s] = 0$ e dunque

$$\text{Cov}[B_s, B_t] = E[B_s B_t] = E[(B_t - B_s)B_s + B_s^2] = E[B_s^2] = s.$$

La seconda parte segue immediatamente scambiando gli indici s e t . \square

La seguente Proposizione mostra che il moto Browniano è un p.a. Gaussiano, ovvero un p.a. per cui la legge di ogni scelta finita di stati è Normale Multivariata.

Proposizione 9.4.4. Sia dato il moto Browniano $B = (B_t)_{t \in [0, \infty[}$. Se $t_1 < \dots < t_n$, allora la legge del v.v.a. $(B_{t_1}, \dots, B_{t_n})^T$ è Normale Multivariata $\mathcal{N}_n(0, \Sigma)$ con $\Sigma = (\min(t_k, t_j))$.

Dimostrazione. Innanzitutto, si osservi che per un vettore di costanti $v = (v_1, \dots, v_n)^T$ risulta

$$v_1 B_{t_1} + \dots + v_n B_{t_n} = (v_1 + \dots + v_n) B_{t_1} + (v_2 + \dots + v_n)(B_{t_2} - B_{t_1}) + \dots + v_n (B_{t_n} - B_{t_{n-1}})$$

e che le v.a. $B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_n} - B_{t_{n-1}}$ sono indipendenti sulla base della assunzione iii) della Definizione 9.4.1. Dunque, la f.c.m. del v.v.a. $(B_{t_1}, \dots, B_{t_n})$ è data da

$$\varphi(v) = E\left[\prod_{k=1}^n e^{i v_k B_{t_k}}\right] = E[e^{i(v_1 + \dots + v_n) B_{t_1}}] E[e^{i(v_2 + \dots + v_n)(B_{t_2} - B_{t_1})}] \times \dots \times E[e^{i v_n (B_{t_n} - B_{t_{n-1}})}].$$

Inoltre, per la Proposizione 9.4.3 la v.a. B_{t_1} è distribuita con legge Normale $\mathcal{N}(0, t_1)$, mentre dalla assunzione ii) della Definizione 9.4.1 la v.a. $B_{t_k} - B_{t_{k-1}}$ è distribuita con legge Normale $\mathcal{N}(0, t_k - t_{k-1})$ per $k = 2, \dots, n$. Dunque, tenendo presente l'Esempio 7.1.6, risulta

$$\begin{aligned} \varphi(v) &= e^{-\frac{1}{2}t_1(v_1+\dots+v_n)^2} e^{-\frac{1}{2}(t_2-t_1)(v_2+\dots+v_n)^2} \times \dots \times e^{-\frac{1}{2}(t_n-t_{n-1})v_n^2} \\ &= e^{-\frac{1}{2}(\sum_{k=1}^n t_k v_k^2 + 2\sum_{k=1}^n \sum_{j=k+1}^n t_k v_k v_j)} = e^{-\frac{1}{2}v^T \Sigma v} \end{aligned}$$

e quindi, sulla base dell'Esempio 7.3.1, segue la tesi. \square

La seguente Proposizione mostra che il moto Browniano è in effetti una martingala.

Proposizione 9.4.5. *Sia dato il moto Browniano $B = (B_t)_{t \in [0, \infty[}$. Il p.a. B è una martingala.*

Dimostrazione. Sulla base della Proposizione 9.4.3 e dalla disuguaglianza di Lyapunov (Teorema 4.3.2), per ogni t si ha

$$E[|B_t|] \leq \sqrt{E[B_t^2]} = \sqrt{t} < \infty.$$

Inoltre, tenendo presente la dimostrazione della Proposizione 9.4.3 e le assunzioni *ii*) e *iii*) della Definizione 9.4.1, risulta

$$E[B_t | \mathcal{F}_s] = E[B_t - B_s + B_s | \mathcal{F}_s] = E[B_t - B_s | \mathcal{F}_s] + E[B_s | \mathcal{F}_s] = E[B_t - B_s] + B_s = B_s,$$

da cui segue la tesi. \square

Le seguenti Proposizioni evidenziano che se il moto Browniano viene riflesso, o traslato, o scalato, o invertito rispetto al tempo, si ottiene di nuovo un moto Browniano.

Proposizione 9.4.6. *Sia dato il moto Browniano $B = (B_t)_{t \in [0, \infty[}$. Posto $X_t = -B_t$ il p.a. $X = (X_t)_{t \in [0, \infty[}$ è un moto Browniano.*

Dimostrazione. Innanzitutto, si ha $P(X_0 = 0) = P(-B_0 = 0) = 1$. Inoltre, per ogni $0 \leq s < t$, sussiste

$$X_t - X_s = -(B_t - B_s) \stackrel{\mathcal{L}}{=} B_t - B_s$$

dal momento che la v.a. $(B_t - B_s)$ è simmetrica rispetto all'origine. Dunque, dall'assunzione *ii*) della Definizione 9.4.1, la legge della v.a. $(X_t - X_s)$ è Normale $\mathcal{N}(0, t - s)$. Dalla Proposizione 3.7.3 si ha che trasformate di v.a. indipendenti sono indipendenti, per cui le v.a. $X_t - X_s = -(B_t - B_s)$ e $X_s = -B_s$ risultano indipendenti sulla base dell'assunzione *ii*) della Definizione 9.4.1. In modo analogo, per ogni $0 \leq t_1 < \dots < t_n$ le v.a. $X_{t_1}, X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}}$ sono indipendenti. Infine, è evidente che il p.a. X ha traiettorie continue *q.c.* Dunque, dalla Definizione 9.4.1 il p.a. X è un moto Browniano. \square

Proposizione 9.4.7. *Sia dato il moto Browniano $B = (B_t)_{t \in [0, \infty[}$. Se $u \geq 0$ e posto $X_t = B_{t+u} - B_u$ il p.a. $X = (X_t)_{t \in [0, \infty[}$ è un moto Browniano.*

Dimostrazione. Risulta evidente che $P(X_0 = 0) = 1$. Inoltre, per ogni $0 \leq s < t$, la v.a. $X_t - X_s = B_{t+u} - B_{s+u}$ si distribuisce con legge Normale $\mathcal{N}(0, t - s)$ sulla base dell'assunzione *ii*) della Definizione 9.4.1. Tenendo presente la medesima assunzione, è immediato verificare che le v.a. $X_{t_1}, X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}}$ sono indipendenti. Infine, è evidente che il p.a. X ha traiettorie continue *q.c.* Dunque, dalla Definizione 9.4.1 il p.a. X è un moto Browniano. \square

Proposizione 9.4.8. *Sia dato il moto Browniano $B = (B_t)_{t \in [0, \infty[}$. Se $c \neq 0$ e posto $X_t = cB_{t/c^2}$ il p.a. $X = (X_t)_{t \in [0, \infty[}$ è un moto Browniano.*

Dimostrazione. Si ha $P(X_0 = 0) = P(cB_0 = 0) = 1$. Inoltre, per ogni $0 \leq s < t$, risulta $X_t - X_s = c(B_{t/c^2} - B_{s/c^2})$. Dunque, dall'assunzione *ii*) della Definizione 9.4.1 e dalla Proposizione 9.4.3, si ottiene

$$E[X_t - X_s] = c(E[B_{t/c^2}] - E[B_{s/c^2}]) = 0$$

e

$$\text{Var}[X_t - X_s] = c^2(\text{Var}[B_{t/c^2}] + \text{Var}[B_{s/c^2}] - 2\text{Cov}[B_{t/c^2}, B_{s/c^2}]) = t - s.$$

Tenendo presente l'Esempio 7.3.2, la v.a. $(X_t - X_s)$ si distribuisce con legge Normale $\mathcal{N}(0, t - s)$. Dal momento che trasformate di v.a. indipendenti sono indipendenti (si veda la Proposizione 3.7.3), le v.a. $X_t - X_s = c(B_{t/c^2} - B_{s/c^2})$ e $X_s = cB_{s/c^2}$ risultano indipendenti sulla base dell'assunzione ii) della Definizione 9.4.1. In modo analogo, le v.a. $X_{t_1}, X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}}$ risultano indipendenti per ogni $0 \leq t_1 < \dots < t_n$. Infine, è evidente che il p.a. X ha traiettorie continue q.c. Dunque, dalla Definizione 9.4.1, il p.a. X è un moto Browniano. \square

Proposizione 9.4.9. *Sia dato il moto Browniano $B = (B_t)_{t \in [0, \infty[}$. Posto $X_t = tB_{1/t}$ se $t > 0$ e $X_0 = 0$ q.c. il p.a. $X = (X_t)_{t \in [0, \infty[}$ è un moto Browniano.*

Dimostrazione. Evidentemente si ha $P(X_0 = 0) = 1$ dalle assunzioni fatte. Dalla Proposizione 9.4.3, la v.a. $B_{1/t}$ è distribuita con legge Normale $\mathcal{N}(0, \frac{1}{t})$ e dunque la v.a. X_t è distribuita con legge Normale $\mathcal{N}(0, t)$, mentre per ogni $0 \leq s < t$ si ha

$$E[X_t - X_s] = tE[B_{1/t}] - sE[B_{1/s}] = 0$$

e

$$\text{Var}[X_t - X_s] = t^2 \text{Var}[B_{1/t}] + s^2 \text{Var}[B_{1/s}] - 2ts \text{Cov}[B_{1/t}, B_{1/s}] = t - s.$$

Dal momento il v.v.a. $(B_{1/t}, B_{1/s})$ è distribuito con legge Normale Multivariata sulla base della Proposizione 9.4.4 e poichè $X_t - X_s = tB_{1/t} - sB_{1/s}$ è una combinazione lineare delle componenti di questo v.v.a., si ha dunque che la v.a. $(X_t - X_s)$ è distribuita con legge Normale $\mathcal{N}(0, t - s)$. Inoltre, tenendo presente che $E[B_{1/t}B_{1/s}] = \min(\frac{1}{t}, \frac{1}{s}) = \frac{1}{t}$, risulta

$$\text{Cov}[X_t - X_s, X_s] = E[(X_t - X_s)X_s] = E[X_t X_s] - E[X_s^2] = stE[B_{1/t}B_{1/s}] - s^2E[B_{1/s}^2] = 0.$$

Il v.v.a. $(X_s, X_t - X_s)$ si distribuisce con legge Normale Multivariata essendo una trasformata affine del v.v.a. $(B_{1/t}, B_{1/s})$ ed inoltre le sue componenti sono indipendenti tenendo presente la precedente relazione. In modo analogo, le v.a. $X_{t_1}, X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}}$ risultano indipendenti per ogni $0 \leq t_1 < \dots < t_n$. Dal momento che il moto Browniano è continuo q.c., risulta

$$P(\lim_{t \rightarrow 0^+} X_t = 0) = P(\lim_{t \rightarrow 0^+} tB_{1/t} = 0) = 1$$

e quindi il p.a. X ha traiettorie continue q.c. Dunque, dalla Definizione 9.4.1, il p.a. X è un moto Browniano. \square

I prossimi Teoremi evidenziano che le traiettorie del moto Browniano sono molto “irregolari”. In effetti, anche se le traiettorie di questo p.a. sono continue q.c. per definizione, queste sono non differenziabili q.o. Inoltre, la variazione di un moto Browniano è infinita q.c., mentre la rispettiva variazione quadratica è finita (per una discussione approfondita di queste quantità si veda Klebaner, 2005)

Teorema 9.4.10. *Sia dato il moto Browniano $B = (B_t)_{t \in [0, \infty[}$. Allora, q.c. le traiettorie del moto Browniano non sono differenziabili q.o.*

Dimostrazione. Si veda Capasso e Bakstein (2015, p.151). \square

Teorema 9.4.11. *Sia dato un moto Browniano $B = (B_t)_{t \in [0, T]}$. Si consideri inoltre la partizione $0 = t_{0,n} < t_{1,n} < \dots < t_{n,n} = T$ e le successioni di v.a. $(V_n)_{n \geq 1}$ e $(V_{n,2})_{n \geq 1}$, dove*

$$V_n = \sum_{k=0}^{n-1} |B_{t_{k+1,n}} - B_{t_{k,n}}|$$

e

$$V_{n,2} = \sum_{k=0}^{n-1} (B_{t_{k+1,n}} - B_{t_{k,n}})^2.$$

Allora, se $\lim_n \max(t_{k+1,n} - t_{k,n}) = 0$, si ha

$$P(\lim_n V_n = \infty) = 1$$

e

$$\lim_n E[(V_{n,2} - T)^2] = 0.$$

Dimostrazione. Si veda Klebaner (2005, p.63). □

La seguente definizione introduce il cosiddetto tempo di primo passaggio del moto Browniano.

Definizione 9.4.12. Sia $B = (B_t)_{t \in [0, \infty[}$ il moto Browniano. Il tempo di arresto

$$\tau_a = \inf\{t \in]0, \infty[: B_t = a\},$$

dove $a \in \mathbb{R}$, è detto *tempo del primo passaggio per a* . □

Teorema 9.4.13. *Sia $B = (B_t)_{t \in [0, \infty[}$ il moto Browniano. La v.a. τ_a ammette d.p. data da*

$$f_{\tau_a}(x) = \frac{|a|}{\sqrt{2\pi x^3}} e^{-\frac{a^2}{2x}} \mathbf{1}_{]0, \infty[}(x).$$

Inoltre, si ha $E[\tau_a] = \infty$.

Dimostrazione. Si veda Klebaner (2005, p.75). □

Dal precedente Teorema è immediato verificare che il tempo del primo passaggio del moto Browniano possiede in effetti la legge di Lévy introdotta nella sua versione ridotta nell'Esempio 4.1.3. Il prossimo Teorema contiene un risultato classico, ovvero la legge dell'arcoseno per il moto Browniano che fornisce la distribuzione dell'ultimo tempo di ritorno all'origine.

Teorema 9.4.14. (Legge dell'arcoseno per il moto Browniano) *Sia $B = (B_t)_{t \in [0, 1]}$ il moto Browniano su $[0, 1]$ e si consideri la v.a. $\tau = \sup\{t \in [0, 1] : B_t = 0\}$. La v.a. τ ammette d.p. data da*

$$f_{\tau}(x) = \frac{1}{\pi \sqrt{x(1-x)}} \mathbf{1}_{[0, 1]}(x).$$

ovvero τ è una v.a. con legge Beta di tipo $\mathcal{BE}(0, 1, \frac{1}{2}, \frac{1}{2})$.

Dimostrazione. Si veda Klebaner (2005, p.75). □

Il prossimo Teorema fornisce un risultato che permette di ottenere notevoli semplificazioni nel calcolo delle leggi connesse al massimo del moto Browniano e al tempo di primo passaggio.

Teorema 9.4.15. (Principio di riflessione) Sia $B = (B_t)_{t \in [0,1]}$ il moto Browniano su $[0, 1]$ e si consideri la v.a. $B_t^* = \sup_{0 \leq s \leq t} B_s$. Fra le v.a. B_t^* , τ_a e B_t sussiste la seguente relazione

$$P(B_t^* \geq a) = P(\tau_a \leq t) = 2P(B_t \geq a),$$

dove $a > 0$.

Dimostrazione. Si veda Baldi (2017, p.70). □

9.5. Riferimenti bibliografici

I testi di Bass (2011), Bhattacharya e Waymire (2016), Koralov e Sinai (2007), Knill (2009), Letta (2016) e Mishura e Shevchenko (2017) sono espressamente focalizzati sulla teoria dei processi stocastici. Ampie parti dei testi avanzati di Ash e Doléans-Dade (2000), Billingsley (1995), Çinlar (2010), Dudley (2004), Kallenberg (2021) e Stroock (2013) sono dedicati ai processi stocastici. Per un approccio introduttivo a questo argomento si possono consultare i testi di Brzeźniak e Zastawniak (2002), Durrett (2012), Resnick (2005) e Tijms (2003). I testi di Capasso e Bakstein (2021) e Revuz e Yor (2005) considerano in modo esaustivo i processi stocastici a tempo continuo. I processi a tempo discreto, con particolare enfasi sulle martingale, sono considerati in Brzeźniak e Zastawniak (2002), Pascucci (2011) e Williams (1992). In particolar modo le applicazioni delle martingale al gioco d'azzardo sono analizzate in Ethier (2010). Per una introduzione alle passeggiate aleatorie si può consultare il testo di Klafter e Sokolov (2011), mentre è fondamentale il capitolo dedicato a questo argomento in Feller (1968). Il moto Browniano è considerato in dettaglio nei testi di Bhattacharya e Waymire (2021), Mörters e Peres (2010) e Schilling e Partzsch (2021).

9.6. Esercizi svolti

• **Nota.** Al fine di evitare notazioni ridondanti e ripetizioni, nei seguenti esercizi si assume l'esistenza di uno spazio probabilizzato (Ω, \mathcal{F}, P) e di una filtrazione $(\mathcal{F}_t)_{t \in \mathbb{T}}$, a meno che non venga specificato diversamente.

Sezione 9.2

• **Esercizio 9.2.1.** Si consideri una successione di v.a. $(X_n)_{n \geq 1}$ a componenti indipendenti tale che $E[X_n] = 0$ e $E[X_n^2] = c^2 < \infty$ per ogni n . Inoltre, si definisca il p.a. a tempo discreto $M = (M_n)_{n \geq 0}$ tale che

$$M_n = S_n^2 - nc^2,$$

dove $S_n = \sum_{k=1}^n X_k$, mentre $\{M_0 = 0\}$ q.c. e $(\mathcal{F}_n)_{n \geq 0}$ è la filtrazione naturale, ovvero $\mathcal{F}_n = \sigma(\{X_1, \dots, X_n\})$. Si verifichi che M è una martingala.

Soluzione. Tenendo presente le assunzioni fatte sul p.a. M , per ogni n si ha

$$E[|M_n|] \leq E[S_n^2] + nc^2 = 2nc^2 < \infty.$$

Inoltre, dal momento che $\{E[M_n | \mathcal{F}_n] = M_n\}$ q.c. e che la σ -algebra generata da X_n è indipendente dalla σ -algebra \mathcal{F}_{n-1} , si ha q.c.

$$\begin{aligned} E[M_n | \mathcal{F}_{n-1}] &= E[S_n^2 - nc^2 | \mathcal{F}_{n-1}] = E[M_{n-1} + X_n^2 - c^2 + 2X_n S_{n-1} | \mathcal{F}_{n-1}] \\ &= E[M_{n-1} | \mathcal{F}_{n-1}] + E[X_n^2 - c^2] + 2S_{n-1}E[X_n] = M_{n-1}. \end{aligned}$$

Inoltre, per la Proprietà della torre (Proposizione 5.2.6), si ha *q.c.*

$$E[E[M_n | \mathcal{F}_{n-1}] | \mathcal{F}_{n-2}] = E[M_n | \mathcal{F}_{n-2}]$$

e quindi dalle precedenti relazioni sussiste *q.c.*

$$E[M_n | \mathcal{F}_{n-2}] = E[M_{n-1} | \mathcal{F}_{n-2}] = M_{n-2}.$$

Procedendo in modo iterativo, si ha infine $\{E[M_n | \mathcal{F}_m] = M_m\}$ *q.c.* per $m \in \mathbb{N}$ con $m < n$. Dunque, il p.a. M è una martingala. \square

• **Esercizio 9.2.2.** Si consideri una successione di v.a. $(X_n)_{n \geq 1}$ a componenti indipendenti tale che $E[X_n] = 1$ per ogni n e il p.a. $M = (M_n)_{n \geq 0}$ tale che

$$M_n = \prod_{k=1}^n X_k,$$

dove $\{M_0 = 1\}$ *q.c.* e $(\mathcal{F}_n)_{n \geq 0}$ è la filtrazione naturale. Si verifichi che M è una martingala.

Soluzione. Tenendo presente le assunzioni fatte sul p.a. M , per ogni n si ha $E[|M_n|] < \infty$. Inoltre, si ha *q.c.*

$$E[M_n | \mathcal{F}_{n-1}] = E[X_n M_{n-1} | \mathcal{F}_{n-1}] = M_{n-1} E[X_n] = M_{n-1}.$$

In modo simile all'Esercizio 9.2.1, si ha dunque $\{E[M_n | \mathcal{F}_m] = M_m\}$ *q.c.* per $m \in \mathbb{N}$ con $m < n$, ovvero il p.a. M è una martingala. \square

• **Esercizio 9.2.3.** (Martingala di Wald) Si consideri una successione di v.a. $(X_n)_{n \geq 1}$ a componenti indipendenti e con la medesima legge, tale che $\psi(t) = E[e^{tX_n}] < \infty$ per ogni $t \in \mathbb{R}$ e per ogni n . Inoltre, si definisca il p.a. a tempo discreto $M = (M_n)_{n \geq 0}$ tale che

$$M_n = \frac{e^{tS_n}}{\psi(t)^n},$$

dove $S_n = \sum_{k=1}^n X_k$, mentre $\{M_0 = 1\}$ *q.c.* e $(\mathcal{F}_n)_{n \geq 0}$ è la filtrazione naturale. Si verifichi che M è una martingala.

Soluzione. Si osservi che risulta

$$M_n = \prod_{k=1}^n \frac{e^{tX_k}}{\psi(t)}$$

e $E[\frac{e^{tX_k}}{\psi(t)}] = 1$. Dunque, sulla base dell'Esercizio 9.2.2, il p.a. M è una martingala. \square

• **Esercizio 9.2.4.** Si consideri una successione di v.a. $(X_n)_{n \geq 1}$ con $E[|X_n|] < \infty$ per ogni n . Inoltre, si definisca il p.a. a tempo discreto $M = (M_n)_{n \geq 0}$ tale che

$$M_n = \sum_{k=1}^n (X_k - E[X_k | \mathcal{F}_{k-1}]),$$

dove $\{M_0 = 0\}$ *q.c.* e $(\mathcal{F}_n)_{n \geq 0}$ è la filtrazione naturale. Si verifichi che M è una martingala.

Soluzione. Tenendo presente le assunzioni fatte sul p.a. M , per ogni n si ha

$$E[|M_n|] \leq \sum_{k=1}^n E[|X_k| + |E[X_k | \mathcal{F}_{k-1}]|] \leq \sum_{k=1}^n E[|X_k| + E[|X_k| | \mathcal{F}_{k-1}]] = 2n E[|X_1|] < \infty.$$

Inoltre, tenendo presente che $E[X_k | \mathcal{F}_{n-1}] = X_k$ per $k \leq n-1$ in quanto $\mathcal{F}_k \subset \mathcal{F}_{n-1}$, si ha *q.c.*

$$\begin{aligned} E[M_n | \mathcal{F}_{n-1}] &= E[M_{n-1} + X_n - E[X_n | \mathcal{F}_{n-1}] | \mathcal{F}_{n-1}] \\ &= E[M_{n-1} | \mathcal{F}_{n-1}] + E[X_n - E[X_n | \mathcal{F}_{n-1}] | \mathcal{F}_{n-1}] \\ &= M_{n-1} + E[X_n | \mathcal{F}_{n-1}] - E[X_n | \mathcal{F}_{n-1}] = M_{n-1}. \end{aligned}$$

In modo simile all'Esercizio 9.2.1, si ha dunque $\{E[M_n | \mathcal{F}_m] = M_m\}$ *q.c.* per $m \in \mathbb{N}$ con $m < n$, ovvero il p.a. M è una martingala. Quindi, sulla base di questo risultato, si può costruire una martingala partendo da una successione qualsiasi di v.a. con media finita. \square

Sezione 9.3

• **Esercizio 9.3.1.** Data una traiettoria $(x_n)_{n \geq 0}$ di una passeggiata aleatoria $X = (X_n)_{n \geq 0}$, dove $x_n \in \{-n, 2-n, \dots, n-2, n\}$ per $n \in \mathbb{N}$, si consideri la cardinalità dei percorsi (x_m, \dots, x_{m+n}) della passeggiata dal passo m al passo $(m+n)$, dove $m \in \mathbb{N}$. Posto $x = x_{m+n} - x_m$, si verifichi che questa cardinalità è pari a

$$N_n(x) = \binom{n}{\frac{1}{2}(n-x)}.$$

Soluzione. Data una realizzazione $(y_n)_{n \geq 1}$ della successione di v.a. $(Y_n)_{n \geq 1}$, si ponga

$$r = \sum_{k=m}^{m+n} \mathbf{1}_{\{1\}}(y_k)$$

e

$$s = \sum_{k=m}^{m+n} \mathbf{1}_{\{-1\}}(y_k).$$

Evidentemente, r e s rappresentano rispettivamente le cardinalità dei valori 1 e -1 nel vettore (y_m, \dots, y_{m+n}) . Dal momento che $n = r + s$ e $x = r - s$, allora si ha $r = \frac{1}{2}(n+x)$ e $s = \frac{1}{2}(n-x)$. Inoltre, vi sono

$$\binom{r+s}{s} = \binom{n}{\frac{1}{2}(n-x)}$$

possibili combinazioni dei valori 1 e -1 nel vettore (y_m, \dots, y_{m+n}) , da cui il risultato richiesto. Si osservi che la cardinalità dei percorsi (x_m, \dots, x_{m+n}) dal passo m al passo $(m+n)$ coincide con la cardinalità dei percorsi (x_0, \dots, x_n) dal passo 0 al passo n . \square

• **Esercizio 9.3.2.** (Principio di riflessione) Assumendo le notazioni dell'Esercizio 9.3.1 e posto $x_m, x_{m+n} > 0$, si verifichi che la cardinalità dei percorsi (x_m, \dots, x_{m+n}) tali che $x_k = 0$ per qualche $k = m+1, \dots, m+n-1$ è pari alla cardinalità dei percorsi del tipo (z_m, \dots, z_{m+n}) con $z_m = -x_m$ e $z_{m+n} = x_{m+n}$.

Soluzione. Considerato un dato percorso (x_m, \dots, x_{m+n}) , sia k il minimo indice per cui la passeggiata ritorna all'origine, ovvero $k = \min\{j : x_j = 0, j = m, \dots, m+n\}$. Dunque, a questo percorso può essere associato il percorso $(-x_m, \dots, -x_{k-1}, x_k, \dots, x_{m+n})$. Evidentemente, quest'ultimo percorso costituisce la "riflessione" del percorso originale per valori negativi della

passaggiata. Sulla base di questa osservazione, esiste dunque una corrispondenza biunivoca fra l'insieme dei percorsi (x_m, \dots, x_{m+n}) e l'insieme dei percorsi (z_m, \dots, z_{m+n}) , che hanno perciò la stessa cardinalità. Tenendo presente l'Esercizio 9.3.1, questa cardinalità è data da $N_n(x)$ dove in questo caso $x = x_m + x_{m+n}$. \square

• **Esercizio 9.3.3.** (Problema del ballottaggio) Considerata un'elezione con due candidati che ricevono rispettivamente a e b voti con $a > b$, si determini la probabilità p che il candidato con più voti risulti strettamente in vantaggio durante lo spoglio.

Soluzione. Si assuma che $n = a + b$ e $x = a - b$. Tenendo presente le notazioni dell'Esercizio 9.3.1, il problema può essere assimilato ad un percorso (x_0, \dots, x_n) di una passeggiata aleatoria, dove $x_n = x$. In questo caso, x_k rappresenta la differenza di voti fra i due candidati alla k -esima iterazione dello spoglio, dove $k = 0, 1, \dots, n$. Si osservi che, se il candidato con più voti è sempre in vantaggio nello spoglio, allora deve risultare $x_1 = 1$. Sulla base degli Esercizi 9.3.1 e 9.3.2, la cardinalità dell'insieme dei percorsi (x_1, \dots, x_n) è data da $N_{n-1}(x - 1)$, mentre la cardinalità di questi percorsi tali che $x_k = 0$ per qualche k è data da $N_{n-1}(x + 1)$. Dal momento che la cardinalità dell'insieme dei percorsi (x_0, \dots, x_n) è data da $N_n(x)$, la probabilità richiesta risulta

$$p = \frac{N_{n-1}(x - 1) - N_{n-1}(x + 1)}{N_n(x)} = \frac{\binom{n-1}{\frac{1}{2}(n+x-2)} - \binom{n-1}{\frac{1}{2}(n+x)}}{\binom{n}{\frac{1}{2}(n+x)}} = \frac{x}{n} = \frac{a - b}{a + b},$$

un risultato intuibile, sebbene abbastanza complicato da verificare con metodi diversi da quello proposto. \square

Sezione 9.4

• **Esercizio 9.4.1.** Dato il moto Browniano $B = (B_t)_{t \in [0, \infty[}$ e assumendo che $s \leq t$, si calcoli il valore atteso $E[B_s^2 B_t^2]$.

Soluzione. Si osservi che

$$E[B_s^2 B_t^2] = E[B_s^2 (B_t - B_s + B_s)^2] = E[B_s^2 (B_t - B_s)^2 + 2B_s^3 (B_t - B_s) + B_s^4].$$

Inoltre, se la v.a. Z è distribuita con legge Normale $\mathcal{N}(0, 1)$, si ha $B_s \stackrel{\mathcal{L}}{=} \sqrt{s}Z$ e quindi

$$E[B_s^4] = s^2 E[Z^4] = 3s^2.$$

Dunque, per le proprietà del moto Browniano, dalle precedenti espressioni risulta

$$E[B_s^2 B_t^2] = E[B_s^2]E[(B_t - B_s)^2] + 2E[B_s^3]E[B_t - B_s] + E[B_s^4] = st + 2s^2,$$

dal momento che $E[B_s^2] = s$, $E[(B_t - B_s)^2] = t - s$ e $E[B_t - B_s] = 0$. \square

• **Esercizio 9.4.2.** Dato il moto Browniano $B = (B_t)_{t \in [0, \infty[}$ e assumendo che $s \leq t$, si verifichi che il v.v.a. $(B_s, B_t)^\top$ ammette d.p.c. data da

$$f_{(B_s, B_t)}(x_1, x_2) = \frac{1}{\sqrt{2\pi s}} e^{-\frac{1}{2s}x_1^2} \frac{1}{\sqrt{2\pi(t-s)}} e^{-\frac{1}{2(t-s)}(x_2 - x_1)^2}.$$

Soluzione. Tenendo presente che la matrice di varianza-covarianza del v.v.a. $(B_s, B_t)^\top$ è data da

$$\Sigma = \begin{pmatrix} s & s \\ s & t \end{pmatrix},$$

si ha

$$\Sigma^{-1} = \frac{1}{s(t-s)} \begin{pmatrix} t & -s \\ -s & s \end{pmatrix}.$$

Dunque, posto $x = (x_1, x_2)^T$ si ha

$$x^T \Sigma^{-1} x = \frac{1}{s(t-s)} (tx_1^2 - 2sx_1x_2 + sx_2^2) = \frac{1}{s(t-s)} ((t-s)x_1^2 + s(x_2 - x_1)^2).$$

Il v.v.a. $(B_s, B_t)^T$ è distribuito con legge Normale bivariata $\mathcal{N}_2(0, \Sigma)$ e quindi si ottiene infine

$$f_{(B_s, B_t)}(x_1, x_2) = \det(2\pi\Sigma)^{-\frac{1}{2}} e^{-\frac{1}{2}x^T \Sigma^{-1} x} = \frac{1}{\sqrt{2\pi s}} e^{-\frac{1}{2s}x_1^2} \frac{1}{\sqrt{2\pi(t-s)}} e^{-\frac{1}{2(t-s)}(x_2 - x_1)^2}.$$

Si osservi che considerando la trasformata $(Y_1, Y_2)^T = g(B_s, B_t)$ dove $g(x_1, x_2) = (x_1, x_2 - x_1)^T$ e tale che $|J(g^{-1}(y))| = 1$, si ha

$$f_{(Y_1, Y_2)}(y_1, y_2) = \frac{1}{\sqrt{2\pi s}} e^{-\frac{1}{2s}y_1^2} \frac{1}{\sqrt{2\pi(t-s)}} e^{-\frac{1}{2(t-s)}y_2^2},$$

ovvero B_s e $(B_t - B_s)$ sono v.a. indipendenti con rispettive leggi $\mathcal{N}(0, s)$ e $\mathcal{N}(0, t-s)$, come evidentemente deve risultare per le proprietà del moto Browniano. \square

• **Esercizio 9.4.3.** Dato il moto Browniano $B = (B_t)_{t \in [0, \infty[}$, si calcoli i valori attesi $E[\mathbf{1}_{]-\infty, a]}(B_t)]$ e $E[B_t \mathbf{1}_{]-\infty, a]}(B_t)]$, dove $a \in \mathbb{R}$, e se ne determini il limite per $t \rightarrow \infty$.

Soluzione. Tenendo presente che $B_t \stackrel{\mathcal{L}}{=} \sqrt{t}Z$ dove Z è una v.a. con legge Normale $\mathcal{N}(0, 1)$, si ha

$$E[\mathbf{1}_{]-\infty, a]}(B_t)] = P(B_t \leq a) = P\left(Z \leq \frac{a}{\sqrt{t}}\right) = \Phi\left(\frac{a}{\sqrt{t}}\right)$$

e quindi, in modo controintuitivo, per ogni a si ha

$$\lim_{t \rightarrow \infty} E[\mathbf{1}_{]-\infty, a]}(B_t)] = \Phi(0) = \frac{1}{2}.$$

Inoltre, risulta

$$E[B_t \mathbf{1}_{]-\infty, a]}(B_t)] = E[\sqrt{t}Z \mathbf{1}_{]-\infty, a]}(\sqrt{t}Z)] = \sqrt{t} \int_{-\infty}^{\frac{a}{\sqrt{t}}} x \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx = -\sqrt{\frac{t}{2\pi}} e^{-\frac{1}{2}a^2}$$

e quindi, di nuovo in modo controintuitivo, per ogni a si ha

$$\lim_{t \rightarrow \infty} E[B_t \mathbf{1}_{]-\infty, a]}(B_t)] = -\infty. \quad \square$$

• **Esercizio 9.4.4.** Dato il moto Browniano $B = (B_t)_{t \in [0, \infty[}$, si consideri la v.a. $B_t^* = \sup_{0 \leq s \leq t} B_s$ e si determini la d.p. $f_{B_t^*}$. Si ottenga inoltre il valore atteso $E[B_t^*]$.

Soluzione. Dal momento che per le proprietà del moto Browniano si ha $\{B_0 = 0\}$ q.c., allora risulta $\{B_t^* \geq 0\}$ q.c. Inoltre, assumendo che $a > 0$, per il Principio di riflessione si ha

$$P(B_t^* \geq a) = 2P(B_t \geq a) = P(|B_t| \geq a).$$

Quindi, risulta $B_t^* \stackrel{\mathcal{L}}{=} |B_t|$, da cui

$$f_{B_t^*}(x) = 2f_{B_t}(x) \mathbf{1}_{]0, \infty[}(x) = \sqrt{\frac{2}{\pi t}} e^{-\frac{1}{2t}x^2} \mathbf{1}_{]0, \infty[}(x).$$

Infine, dal precedente risultato si ottiene

$$E[B_t^*] = \int_0^\infty x \sqrt{\frac{2}{\pi t}} e^{-\frac{1}{2t}x^2} dx = \sqrt{\frac{2t}{\pi}}. \quad \square$$

• **Esercizio 9.4.5.** (Tempo di occupazione) Dato il moto Browniano $B = (B_t)_{t \in [0, \infty[}$, si consideri l'insieme misurabile $A \subset \mathbb{R}$ con misura di Lebesgue $\lambda(A) > 0$ e l'insieme aleatorio

$$S_A = \{\omega \in \Omega : B_t(\omega) \in A, t \in [0, \infty[\}.$$

Si determini il valore atteso $E[\lambda(S_A)]$.

Soluzione. Si osservi che S_A è in effetti l'insieme dei valori di t per cui il moto Browniano visita l'insieme A e per questo motivo S_A viene anche definito come tempo di occupazione dell'insieme A . Ovviamente, per ogni traiettoria $B_t(\omega)$ del moto Browniano, S_A si realizza in un insieme dato di valori. Dunque, si ha che

$$\lambda(S_A) = \int_0^\infty \mathbf{1}_A(B_t) dt$$

è una v.a. che rappresenta la misura di Lebesgue dell'insieme aleatorio S_A . Sulla base di questa espressione e del Teorema di Fubini, risulta

$$E[\lambda(S_A)] = E[\int_0^\infty \mathbf{1}_A(B_t) dt] = \int_0^\infty E[\mathbf{1}_A(B_t)] dt = \int_0^\infty P(B_t \in A) dt.$$

Tenendo presente che la v.a. B_t si distribuisce con legge Normale $\mathcal{N}(0, t)$ e applicando di nuovo il Teorema di Fubini, si ha

$$E[\lambda(S_A)] = \int_0^\infty \left(\int_A \frac{1}{\sqrt{2\pi t}} e^{-\frac{1}{2t}x^2} dx \right) dt = \int_A \int_0^\infty \frac{1}{\sqrt{2\pi t}} e^{-\frac{1}{2t}x^2} dt dx.$$

Tuttavia, risulta

$$\int_0^\infty \frac{1}{\sqrt{2\pi t}} e^{-\frac{1}{2t}x^2} dt = \frac{|x|}{2\sqrt{\pi}} \int_0^\infty \frac{e^{-y}}{\sqrt{y^3}} dy = \infty$$

e dunque $E[\lambda(S_A)] = \infty$, ovvero il moto Browniano visita un qualsiasi insieme misurabile A infinitamente spesso. Risulta notevole osservare che, con una dimostrazione simile alla precedente, per il moto Browniano multivariato in \mathbb{R}^d , dove $d > 2$, e per un insieme misurabile $A \subset \mathbb{R}^d$, si ha

$$E[\lambda^d(S_A)] = \frac{\Gamma(\frac{1}{2}d - 1)}{2\pi^{\frac{1}{2}d}} \int_A \|x\|^{2-d} dx.$$

In questo caso, l'insieme aleatorio S_A viene definito opportunamente per il moto Browniano multivariato. Dunque, per $d > 2$ si può avere $E[\lambda^d(S_A)] < \infty$ per un insieme A , ovvero il moto Browniano può non visitare infinitamente spesso l'insieme A . \square

Pagina intenzionalmente vuota

Capitolo 10

Calcolo stocastico

10.1. Introduzione all'integrazione stocastica

Il problema dell'integrazione stocastica può essere fatto risalire a Louis Bachelier (1870-1946). Se X_t rappresenta il prezzo di un bene al variare del tempo, nella sua tesi di laurea *Théorie de la spéculation* (1900), Louis Bachelier propose di modellare gli incrementi infinitesimi del prezzo in modo proporzionale agli incrementi infinitesimi del moto Browniano, ovvero (con una simbologia moderna) introdusse la relazione euristica $dX_t = \sigma dB_t$ dove $\sigma \in]0, \infty[$. In questo caso, la precedente relazione porta intuitivamente a concludere che $X_t = X_0 + \sigma B_t$. Tuttavia, anche se questo modello può essere discretamente valido per descrivere la dinamica del prezzo nel breve periodo, in generale non è realistico. In effetti, tenendo presenti le proprietà del moto Browniano, il prezzo X_t può assumere valori negativi con probabilità crescente all'aumentare del tempo.



Figura 9.2.1. Kiyosi Itô (1915-2008).

L'inconveniente del modello potrebbe essere aggirato supponendo che l'incremento infinitesimo del prezzo, rapportato al prezzo stesso, sia proporzionale agli incrementi infinitesimi del moto Browniano considerando la relazione euristica $dX_t/X_t = \sigma dB_t$, ovvero $dX_t = \sigma X_t dB_t$. Anche se questa ultima espressione potrebbe ricordare una equazione differenziale ordinaria, evidentemente presenta una difficoltà teorica dal momento che il moto Browniano non è differenziabile *q.o.* e quindi non ha significato matematico. Il problema è stato risolto da Kiyosi Itô (1915-2008), che ha dato un'interpretazione rigorosa alla precedente relazione con una formulazione in termini di integrale del tipo $X_t = X_0 + \sigma \int_0^t X_s dB_s$. Nel secondo termine di quest'espressione compare appunto il cosiddetto *integrale stocastico di Itô*.

Di seguito vengono introdotte le basi per la costruzione dell'integrale stocastico di Itô partendo da una classe elementare di processi, al fine di ottenere la definizione generale nella prossima sezione.

Definizione 10.1.1. Si consideri uno spazio di probabilità (Ω, \mathcal{F}, P) con filtrazione $(\mathcal{F}_t)_{t \in [0, T]}$ dove $T > 0$. Sia inoltre $0 = t_0 < t_1 < \dots < t_n = T$ una partizione di $[0, T]$ e sia $Z = (Z_0, \dots, Z_{n-1})^T$ un v.v.a. tale che ogni v.a. Z_k è misurabile rispetto a \mathcal{F}_{t_k} con $E[Z_k^2] < \infty$ per $k = 0, \dots, n-1$. Il p.a. $X = (X_t)_{t \in [0, T]}$ è detto *semplice* se

$$X_t = \sum_{k=0}^{n-1} Z_k \mathbf{1}_{[t_k, t_{k+1}[}(t).$$

La classe dei p.a. semplici è denotata con \mathcal{S}_T^2 . □

Si osservi che la misurabilità delle v.a. Z_k rispetto alle σ -algebre \mathcal{F}_{t_k} implica che il p.a. semplice X sia adattato alla filtrazione $(\mathcal{F}_t)_{t \in [0, \infty[}$. Inoltre, l'assunzione $E[Z_k^2] < \infty$ implica che $E[X_t^2] < \infty$ per ogni $t \in [0, T]$. Infine, \mathcal{S}_T^2 è uno spazio vettoriale, dal momento che è immediato verificare che $aX + bY \in \mathcal{S}_T^2$ per ogni $X, Y \in \mathcal{S}_T^2$ e $a, b \in \mathbb{R}$.

Definizione 10.1.2. Si consideri uno spazio di probabilità (Ω, \mathcal{F}, P) con filtrazione $(\mathcal{F}_t)_{t \in [0, \infty[}$ e il moto Browniano $B = (B_t)_{t \in [0, \infty[}$ adattato alla filtrazione. Dato il p.a. $X \in \mathcal{S}_T^2$, l'*integrale stocastico di Itô* di X è definito come

$$I(X) = \sum_{k=0}^{n-1} Z_k (B_{t_{k+1}} - B_{t_k})$$

e si scrive

$$I(X) = \int_0^T X_t dB_t. \quad \square$$

Dalla Definizione 10.1.2 è evidente che l'integrale di Itô di un p.a. semplice è una v.a. Inoltre, risulta opportuno sottolineare la differenza teorica fra l'integrale di Itô e l'integrale di Lebesgue di un p.a. semplice. Più esattamente, l'integrale del p.a. semplice X è definito come

$$\int_0^T X_t dt = \sum_{k=0}^{n-1} Z_k (t_{k+1} - t_k).$$

Dunque, anche l'integrale del p.a. è ovviamente una v.a. Tuttavia, l'integrale del p.a. è basato su incrementi deterministici, al contrario dell'integrale di Itô, dove gli incrementi sono proporzionali agli incrementi di un moto Browniano. In generale, si noti che l'integrale di un p.a. (le cui traiettorie sono integrabili *q.c.*) è una v.a., dal momento che ad ogni realizzazione del processo viene associato un integrale di Lebesgue.

Si osservi inoltre che per semplicità di notazione si pone anche

$$\int_a^b X_t dB_t = \int_0^T \mathbf{1}_{[a, b[}(t) X_t dB_t,$$

dove $a, b \in [0, T]$ e $a < b$.

Il seguente Teorema evidenzia che l'integrale stocastico di Itô è una v.a. con media nulla e secondo momento finito quando il p.a. è semplice.

Teorema 10.1.3. Se $X \in \mathcal{S}_T^2$, l'integrale stocastico di Itô è una v.a. tale che $E[I(X)] = 0$ e

$$\text{Var}[I(X)] = \mathbb{E}[I(X)^2] = \int_0^T \mathbb{E}[X_t^2] dt = \mathbb{E}[\int_0^T X_t^2 dt] < \infty.$$

Dimostrazione. Per semplicità di notazione si assuma che $\Delta_k = B_{t_{k+1}} - B_{t_k}$. Si noti che la v.a. Z_k è misurabile rispetto a \mathcal{F}_{t_k} e che l'incremento Δ_k è indipendente dalla σ -algebra \mathcal{F}_{t_k} per l'assunzione *iii*) della Definizione 9.4.1. Dunque, tenendo presente la Proposizione 5.2.1 e il Teorema 5.2.4 si ha

$$\mathbb{E}[Z_k \Delta_k] = \mathbb{E}[\mathbb{E}[Z_k \Delta_k \mid \mathcal{F}_{t_k}]] = \mathbb{E}[Z_k \mathbb{E}[\Delta_k \mid \mathcal{F}_{t_k}]] = \mathbb{E}[Z_k \mathbb{E}[\Delta_k]] = 0,$$

in quanto sussiste $\mathbb{E}[\Delta_k] = 0$ sulla base dell'assunzione *ii*) della Definizione 9.4.1. Quindi, risulta

$$\mathbb{E}[I(X)] = \sum_{k=0}^{n-1} \mathbb{E}[Z_k \Delta_k] = 0.$$

Inoltre, si ha

$$I(X)^2 = \sum_{k=0}^{n-1} \sum_{j=0}^{n-1} Z_k Z_j \Delta_k \Delta_j.$$

Tenendo presente i precedenti commenti, assumendo $k < j$, si ha

$$\mathbb{E}[Z_k Z_j \Delta_k \Delta_j] = \mathbb{E}[\mathbb{E}[Z_k Z_j \Delta_k \Delta_j \mid \mathcal{F}_{t_j}]] = \mathbb{E}[Z_k Z_j \Delta_k \mathbb{E}[\Delta_j \mid \mathcal{F}_{t_j}]] = \mathbb{E}[Z_k Z_j \Delta_k \mathbb{E}[\Delta_j]] = 0,$$

mentre per $k = j$, si ha

$$\mathbb{E}[Z_k^2 \Delta_k^2] = \mathbb{E}[\mathbb{E}[Z_k^2 \Delta_k^2 \mid \mathcal{F}_{t_k}]] = \mathbb{E}[Z_k^2 \mathbb{E}[\Delta_k^2 \mid \mathcal{F}_{t_k}]] = \mathbb{E}[Z_k^2 \mathbb{E}[\Delta_k^2]] = \mathbb{E}[Z_k^2] (t_{k+1} - t_k),$$

dal momento che $\mathbb{E}[\Delta_k^2] = t_{k+1} - t_k$ dall'assunzione *ii*) della Definizione 9.4.1. Dunque, risulta

$$\mathbb{E}[I(X)^2] = \sum_{k=0}^{n-1} \mathbb{E}[Z_k^2] (t_{k+1} - t_k).$$

Inoltre, tenendo presente che $\mathbf{1}_{[t_k, t_{k+1}[}(t) \mathbf{1}_{[t_j, t_{j+1}[}(t) = 0$ per $k \neq j$, si ha

$$\begin{aligned} \int_0^T X_t^2 dt &= \int_0^T \sum_{k=0}^{n-1} \sum_{j=0}^{n-1} Z_k Z_j \mathbf{1}_{[t_k, t_{k+1}[}(t) \mathbf{1}_{[t_j, t_{j+1}[}(t) dt \\ &= \int_0^T \sum_{k=0}^{n-1} Z_k^2 \mathbf{1}_{[t_k, t_{k+1}[}(t) dt = \sum_{k=0}^{n-1} Z_k^2 (t_{k+1} - t_k), \end{aligned}$$

da cui risulta

$$\mathbb{E}[\int_0^T X_t^2 dt] = \sum_{k=0}^{n-1} \mathbb{E}[Z_k^2] (t_{k+1} - t_k).$$

Inoltre, per il Teorema di Fubini (Teorema A.11), si ha

$$\mathbb{E}[\int_0^T X_t^2 dt] = \int_0^T \mathbb{E}[X_t^2] dt.$$

La tesi segue dalle precedenti relazioni. □

• **Esempio 10.1.1.** Si assuma il p.a. semplice $X = (X_t)_{t \in [0,5]}$, dove

$$X_t = Z_0 \mathbf{1}_{[0,1[}(t) + Z_1 \mathbf{1}_{[1,2[}(t) + Z_2 \mathbf{1}_{[2,5]}(t),$$

e $Z = (Z_0, Z_1, Z_2)^T$ è un v.v.a. degenere con $P(Z_0 = 1, Z_1 = 2, Z_2 = 1) = 1$. In questo caso, l'integrale stocastico di Itô è dato da

$$I(X) = Z_0(B_1 - B_0) + Z_1(B_2 - B_1) + Z_2(B_5 - B_2) \stackrel{L}{=} B_1 + 2(B_2 - B_1) + (B_5 - B_2).$$

Dunque, dall'assunzione *iii* della Definizione 9.4.1, si ha che $I(X)$ è una v.a. con legge Normale $\mathcal{N}(0, 8)$. In effetti, dal Teorema 10.1.3 risulta $E[I(X)] = 0$ e

$$\text{Var}[I(X)] = E[Z_0^2] + E[Z_1^2] + 3E[Z_2^2] = 8. \quad \square$$

• **Esempio 10.1.2.** Si assuma il p.a. semplice $X = (X_t)_{t \in [0,1]}$, dove

$$X_t = Z_0 \mathbf{1}_{[0,1]}(t),$$

mentre Z_0 è una v.a. con legge di Bernoulli di parametro $\frac{1}{2}$. In questo caso, l'integrale stocastico di Itô è dato da

$$I(X) = Z_0(B_1 - B_0) \stackrel{L}{=} Z_0 B_1.$$

Dunque, la f.r. della v.a. $I(X)$ è data da

$$\begin{aligned} F_{I(X)}(x) &= P(I(X) \leq x) = P(\{I(X) \leq x\} \cap \{Z_0 = 0\}) + P(\{I(X) \leq x\} \cap \{Z_0 = 1\}) \\ &= P(I(X) \leq x \mid Z_0 = 0)P(Z_0 = 0) + P(I(X) \leq x \mid Z_0 = 1)P(Z_0 = 1) \\ &= \frac{1}{2} \mathbf{1}_{[0, \infty[}(x) + \frac{1}{2} \Phi(x). \end{aligned}$$

Dunque, $I(X)$ è una v.a. mista, dal momento che la corrispondente f.r. è una combinazione convessa di una f.r. di una v.a. degenere nello zero e una v.a. con legge Normale ridotta. Inoltre, risulta $E[I(X)] = 0$ e

$$\text{Var}[I(X)] = E[Z_0^2] = \frac{1}{2}. \quad \square$$

I seguenti risultati forniscono rispettivamente la proprietà di linearità l'integrale stocastico di Itô e la covarianza fra due integrali stocastici di Itô nel caso di p.a. semplici.

Teorema 10.1.4. Dati i p.a. $X, Y \in \mathcal{S}_T^2$, per ogni $a, b \in \mathbb{R}$ si ha

$$I(aX + bY) = aI(X) + bI(Y).$$

Dimostrazione. Si consideri la partizione $0 = t_0 < t_1 < \dots < t_n = T$ di $[0, T]$, in modo tale che

$$X_t = \sum_{k=0}^{n-1} Z_k \mathbf{1}_{[t_k, t_{k+1}[}(t)$$

e

$$Y_t = \sum_{k=0}^{n-1} U_k \mathbf{1}_{[t_k, t_{k+1}[}(t).$$

dove le v.a. Z_k e U_k sono misurabili rispetto a \mathcal{F}_{t_k} con $E[Z_k^2] < \infty$ e $E[U_k^2] < \infty$ per $k = 0, \dots, n-1$. Se le due partizioni per i p.a. X e Y sono differenti, è sempre possibile considerare una partizione comune considerando un raffinamento delle due partizioni originali. Dunque, si ha

$$aX_t + bY_t = \sum_{k=0}^{n-1} (aZ_k + bU_k) \mathbf{1}_{[t_k, t_{k+1}[}(t).$$

Assumendo le notazioni della dimostrazione del Teorema 10.1.3, risulta

$$I(aX + bY) = \sum_{k=0}^{n-1} (aZ_k + bU_k) \Delta_k = a \sum_{k=0}^{n-1} Z_k \Delta_k + b \sum_{k=0}^{n-1} U_k \Delta_k = aI(X) + bI(Y),$$

da cui segue la tesi. □

Proposizione 10.1.5. *Dati i p.a. $X, Y \in \mathcal{S}_T^2$, si ha*

$$\text{Cov}[I(X), I(Y)] = E[I(X)I(Y)] = \int_0^T E[X_t Y_t] dt = E[\int_0^T X_t Y_t dt].$$

Dimostrazione. Si tenga presente l'identità

$$ab = \frac{1}{2} (a + b)^2 - \frac{1}{2} a^2 - \frac{1}{2} b^2,$$

dove $a, b \in \mathbb{R}$. Dai Teoremi 10.1.3 e 10.1.4, si ha

$$\begin{aligned} E[I(X)I(Y)] &= \frac{1}{2} E[(I(X) + I(Y))^2] - \frac{1}{2} E[I(X)^2] - \frac{1}{2} E[I(Y)^2] \\ &= \frac{1}{2} E[\int_0^T (X_t + Y_t)^2 dt] - \frac{1}{2} E[\int_0^T X_t^2 dt] - \frac{1}{2} E[\int_0^T Y_t^2 dt] = E[\int_0^T X_t Y_t dt]. \end{aligned}$$

L'identità

$$E[\int_0^T X_t Y_t dt] = \int_0^T E[X_t Y_t] dt$$

segue dal Teorema di Fubini (Teorema A.11). □

10.2. Integrale di Itô

La seguente classe di p.a. è centrale nella definizione generale dell'integrale stocastico di Itô. In effetti, gli elementi di questa classe saranno le “funzioni integrande” opportune per cui l'integrale stocastico è definito.

Definizione 10.2.1. Si consideri uno spazio di probabilità (Ω, \mathcal{F}, P) con filtrazione $(\mathcal{F}_t)_{t \in [0, \infty[}$. La classe \mathcal{M}_T^2 è costituita dai p.a. $X = (X_t)_{t \in [0, T]}$ adattati alla filtrazione e tali che $E[\int_0^T X_t^2 dt] < \infty$. □

• **Esempio 10.2.1.** Si consideri il p.a. $X = (B_t)_{t \in [0, T]}$. Tenendo presente la Definizione 9.4.1 si ha

$$E[\int_0^T B_t^2 dt] = \int_0^T E[B_t^2] dt = \int_0^T t dt = \frac{T^2}{2} < \infty,$$

ovvero $X \in \mathcal{M}_T^2$. Si consideri inoltre il p.a. $X = (B_t^2)_{t \in [0, T]}$. Tenendo presente che in base alla Proposizione 9.4.3 si ha $E[B_t^4] = 3t^2$ (si veda anche l'Esempio 7.1.8), allora risulta

$$E[\int_0^T B_t^4 dt] = \int_0^T E[B_t^4] dt = \int_0^T 3t^2 dt = T^3 < \infty,$$

ovvero si ha anche $X \in \mathcal{M}_T^2$. □

È immediato verificare che $\mathcal{S}_T^2 \subset \mathcal{M}_T^2$. Inoltre, la seguente Proposizione fornisce una semplice condizione per verificare se un p.a. appartiene alla classe \mathcal{M}_T^2 , mentre il Teorema successivo evidenzia che la classe \mathcal{S}_T^2 è densa nella classe \mathcal{M}_T^2 .

Proposizione 10.2.2. *Se il p.a. $X = (X_t)_{t \in [0, T]}$ è adattato alla filtrazione e limitato, ovvero se esiste $K > 0$ tale che $|X_t(\omega)| < K$ per ogni $(t, \omega) \in [0, T] \times \Omega$, allora $X \in \mathcal{M}_T^2$.*

Dimostrazione. Dalle assunzioni si ha

$$E[\int_0^T X_t^2 dt] \leq E[\int_0^T K^2 dt] = TK^2,$$

e sulla base della Definizione 10.2.1 segue la tesi. □

Teorema 10.2.3. *Per ogni p.a. $X = (X_t)_{t \in [0, T]}$ tale che $X \in \mathcal{M}_T^2$ esiste una successione di p.a. $(X_n)_{n \geq 1}$, dove $X_n = (X_{t,n})_{t \in [0, T]}$ con $X_n \in \mathcal{S}_T^2$, che approssima il p.a. X , ovvero*

$$\lim_n E[\int_0^T (X_t - X_{t,n})^2 dt] = 0.$$

Dimostrazione. Si veda Da Prato (2014). □

Sulla base dei precedenti risultati, si può dunque introdurre la versione generale dell'integrale stocastico di Itô di un p.a. definito su \mathcal{M}_T^2 come limite in media quadratica di una successione di integrali stocastici relativi ad una successione di p.a. semplici.

Definizione 10.2.4. Si consideri uno spazio di probabilità (Ω, \mathcal{F}, P) con filtrazione $(\mathcal{F}_t)_{t \in [0, \infty[}$ e un moto Browniano $B = (B_t)_{t \in [0, \infty[}$ adattato alla filtrazione. Dato il p.a. $X \in \mathcal{M}_T^2$, si dice che $I(X)$ è l'integrale stocastico di Itô del p.a. X se

$$\lim_n E[(I(X) - I(X_n))^2] = 0$$

per ogni successione $(X_n)_{n \geq 1}$ di p.a. tali che $X_n \in \mathcal{S}_T^2$ e si scrive

$$I(X) = \int_0^T X_t dB_t. \quad \square$$

Evidentemente, la precedente Definizione implica che l'integrale stocastico di Itô può essere definito in termini di convergenza in media quadratica, ovvero risulta

$$I(X_n) \xrightarrow{L^2} I(X)$$

per $n \rightarrow \infty$. Inoltre, in modo equivalente a quanto fatto per i p.a. semplici, per comodità di notazione si pone anche

$$\int_a^b X_t dB_t = \int_0^T \mathbf{1}_{[a, b](t)} X_t dB_t.$$

Nel caso particolare in cui $X = (X_t)_{t \in [0, T]}$ è un processo deterministico, ovvero se $X_t = g(t)$ q.c. con $g : \mathbb{R} \rightarrow \mathbb{R}$, l'integrale di Itô si riduce a

$$I(X) = \int_0^T g(t) dB_t,$$

che è anche detto *integrale di Wiener*.

Dal seguente Teorema si evince che l'integrale stocastico è ben definito, nel senso che non dipende dalla scelta della successione di p.a. semplici.

Teorema 10.2.5. *Per ogni p.a. $X \in \mathcal{M}_T^2$, l'integrale stocastico di Itô $I(X)$ non dipende dalla successione di p.a. semplici $(X_n)_{n \geq 1}$ che approssimano il p.a. X .*

Dimostrazione. Si consideri due successioni di p.a. semplici $(X_n)_{n \geq 1}$ e $(Y_n)_{n \geq 1}$ che approssimano X in \mathcal{M}_T^2 , ovvero tali che $\lim_n E[\int_0^T (X_t - X_{t,n})^2 dt] = 0$ e $\lim_n E[\int_0^T (X_t - Y_{t,n})^2 dt] = 0$. Tenendo presente la disuguaglianza

$$(a + b)^2 \leq 2a^2 + 2b^2,$$

dove $a, b \in \mathbb{R}$, si ha

$$\begin{aligned} E[\int_0^T (X_{t,n} - Y_{t,n})^2 dt] &= E[\int_0^T ((X_{t,n} - X_t) + (X_t - Y_{t,n}))^2 dt] \\ &\leq 2E[\int_0^T (X_t - X_{t,n})^2 dt] + 2E[\int_0^T (X_t - Y_{t,n})^2 dt] \end{aligned}$$

e dunque

$$\lim_n E[\int_0^T (X_{t,n} - Y_{t,n})^2 dt] = 0.$$

Inoltre, dai Teoremi 10.1.3 e 10.1.4 si ha

$$E[(I(X_n) - I(Y_n))^2] = E[I(X_n - Y_n)^2] = E[\int_0^T (X_{t,n} - Y_{t,n})^2 dt],$$

da cui, sulla base del precedente risultato, si ha

$$\lim_n E[(I(X_n) - I(Y_n))^2] = 0.$$

La tesi segue sulla base delle assunzioni fatte e della Definizione 10.2.4. □

In modo parallelo per quanto visto per i p.a. semplici si hanno i seguenti Teoremi.

Teorema 10.2.6. *Se $X \in \mathcal{M}_T^2$, l'integrale stocastico di Itô è una v.a. tale che $E[I(X)] = 0$ e*

$$\text{Var}[I(X)] = E[I(X)^2] = \int_0^T E[X_t^2] dt = E[\int_0^T X_t^2 dt] < \infty.$$

Dimostrazione. Si consideri una successione di p.a. semplici $(X_n)_{n \geq 1}$ che approssima il p.a. X in \mathcal{M}_T^2 . Per quanto riguarda la prima parte, tenendo presente che $E[I(X_n)] = 0$ dal Teorema 10.1.3, sulla base della disuguaglianza di Jensen (Teorema 4.2.6) si ha

$$E[I(X)]^2 = E[I(X) - I(X_n)]^2 = E[I(X - X_n)]^2 \leq E[I(X - X_n)^2] = E[(I(X) - I(X_n))^2].$$

Quindi, per $n \rightarrow \infty$ segue che $E[I(X)] = 0$ dalla Definizione 10.2.4. Per quanto riguarda la seconda parte, dal momento che \mathcal{M}_T^2 è uno spazio vettoriale, si ha

$$\left| \sqrt{E[\int_0^T X_t^2 dt]} - \sqrt{E[\int_0^T X_{t,n}^2 dt]} \right| \leq \sqrt{E[\int_0^T (X_t - X_{t,n})^2 dt]}.$$

La precedente relazione implica che

$$\lim_n E[\int_0^T X_{t,n}^2 dt] = E[\int_0^T X_t^2 dt],$$

essendo $\lim_n E[\int_0^T (X_t - X_{t,n})^2 dt] = 0$ sulla base del Teorema 10.2.3. In modo simile, la Definizione 10.2.4 implica che

$$\lim_n E[(\int_0^T X_{t,n} dB_t)^2] = E[(\int_0^T X_t dB_t)^2],$$

ovvero

$$\lim_n E[I(X_n)^2] = E[I(X)^2].$$

Inoltre, dal momento che $X_n \in \mathcal{S}_T^2$, dal Teorema 10.1.3 si ha

$$E[(\int_0^T X_{t,n} dB_t)^2] = E[\int_0^T X_{t,n}^2 dt]$$

e quindi i due limiti coincidono. □

Teorema 10.2.7. *Dati i p.a. $X, Y \in \mathcal{M}_T^2$, per ogni $a, b \in \mathbb{R}$ si ha*

$$I(aX + bY) = aI(X) + bI(Y).$$

Dimostrazione. Si consideri due successioni di p.a. semplici $(X_n)_{n \geq 1}$ e $(Y_n)_{n \geq 1}$ che approssimano i p.a. X e Y in \mathcal{M}_T^2 , ovvero tali che $\lim_n E[\int_0^T (X_t - X_{t,n})^2 dt] = 0$ e $\lim_n E[\int_0^T (Y_t - Y_{t,n})^2 dt] = 0$. Tenendo presente la disuguaglianza evidenziata nella dimostrazione del Teorema 10.2.5, si ha

$$\begin{aligned} E[\int_0^T (aX_t + bY_t - aX_{t,n} - bY_{t,n})^2 dt] &= E[\int_0^T (a(X_t - X_{t,n}) + b(Y_t - Y_{t,n}))^2 dt] \\ &\leq 2a^2 E[\int_0^T (X_t - X_{t,n})^2 dt] + 2b^2 E[\int_0^T (Y_t - Y_{t,n})^2 dt]. \end{aligned}$$

Dunque, risulta

$$\lim_n E[\int_0^T (aX_t + bY_t - aX_{t,n} - bY_{t,n})^2 dt] = 0,$$

ovvero la successione di p.a. semplici $(aX_n + bY_n)_{n \geq 1}$ approssima il p.a. $aX + bY$ in \mathcal{M}_T^2 . Dunque, dalla Definizione 10.2.4 si ha

$$\lim_n E[(I(aX + bY) - I(aX_n + bY_n))^2] = 0,$$

ovvero, tenendo presente il Teorema 10.1.4, risulta anche

$$\lim_n E[(I(aX + bY) - aI(X_n) - bI(Y_n))^2] = 0.$$

Inoltre, si ha

$$\begin{aligned} E[(aI(X) + bI(Y) - aI(X_n) - bI(Y_n))^2] &= E[(a(I(X) - I(X_n)) + b(I(Y) - I(Y_n)))^2] \\ &\leq 2a^2 E[(I(X) - I(X_n))^2] + 2b^2 E[(I(Y) - I(Y_n))^2]. \end{aligned}$$

Poichè risulta $\lim_n E[(I(X) - I(X_n))^2] = 0$ e $\lim_n E[(I(Y) - I(Y_n))^2] = 0$ dalla Definizione 10.2.4, si ha infine che

$$\lim_n E[(aI(X) + bI(Y) - aI(X_n) - bI(Y_n))^2] = 0.$$

Confrontando i precedenti limiti si ottiene la tesi. □

Proposizione 10.2.8. *Dati i p.a. $X, Y \in \mathcal{M}_T^2$, si ha*

$$\text{Cov}[I(X), I(Y)] = E[I(X)I(Y)] = \int_0^T E[X_t Y_t] dt = E[\int_0^T X_t Y_t dt].$$

Dimostrazione. È analoga a quella della Proposizione 10.1.5. □

• **Esempio 10.2.2.** Se $X = (X_t)_{[0,T]}$ è un p.a. per cui risulta che $X_t = \mathbf{1}_{[0,T]}(t)$, si vuole verificare che

$$I(X) = \int_0^T dB_t = B_T.$$

Evidentemente, X è un p.a. semplice e dunque $X \in \mathcal{M}_T^2$. Inoltre, si può banalmente scegliere la successione $(X_n)_{n \geq 1}$ con $X_n = X$ e quindi sussiste

$$I(X_n) = \sum_{k=0}^{n-1} (B_{t_{k+1,n}} - B_{t_{k,n}}) = B_T$$

dal momento che la somma è telescopica. Quindi, si ha

$$I(X_n) \xrightarrow{L^2} B_T$$

e dalla Definizione 10.2.4 risulta

$$I(X) = B_T.$$

Dunque, in questo caso l'integrale stocastico di Itô è una v.a. con legge Normale $\mathcal{N}(0, T)$. In modo coerente, dal Teorema 10.2.6 risulta $E[I(X)] = 0$ e $\text{Var}[I(X)] = T$. □

• **Esempio 10.2.3.** Se $X = (B_t)_{[0,T]}$, si vuole verificare che

$$I(X) = \int_0^T B_t dB_t = \frac{1}{2} (B_T^2 - T).$$

Nell'Esempio 10.2.1 è stato verificato che $X \in \mathcal{M}_T^2$. Dunque, posto $0 = t_{0,n} < t_{1,n} < \dots < t_{n,n} = T$, si consideri la successione di p.a. semplici $(X_n)_{n \geq 1}$, dove $X_n = (X_{t,n})_{t \in [0,T]}$ e

$$X_{t,n} = \sum_{k=0}^{n-1} B_{t_{k,n}} \mathbf{1}_{[t_{k,n}, t_{k+1,n}[}(t).$$

Inoltre, dal momento che l'integrale stocastico non dipende dalla scelta della successione, si ponga $t_{k,n} = \frac{kT}{n}$. Tenendo presente l'identità

$$a(b-a) = \frac{1}{2} (b^2 - a^2) - \frac{1}{2} (b-a)^2,$$

dove $a, b \in \mathbb{R}$, si ha

$$\begin{aligned} I(X_n) &= \sum_{k=0}^{n-1} B_{t_{k,n}} (B_{t_{k+1,n}} - B_{t_{k,n}}) = \frac{1}{2} \sum_{k=0}^{n-1} (B_{t_{k+1,n}}^2 - B_{t_{k,n}}^2) - \frac{1}{2} \sum_{k=0}^{n-1} (B_{t_{k+1,n}} - B_{t_{k,n}})^2 \\ &= \frac{1}{2} B_T^2 - \frac{1}{2} \sum_{k=0}^{n-1} (B_{t_{k+1,n}} - B_{t_{k,n}})^2. \end{aligned}$$

Dunque, tenendo presente il Teorema 9.4.11 si ha

$$\frac{1}{2} B_T^2 - \frac{1}{2} \sum_{k=0}^{n-1} (B_{t_{k+1,n}} - B_{t_{k,n}})^2 \xrightarrow{L^2} \frac{1}{2} (B_T^2 - T)$$

per $n \rightarrow \infty$. Dalla Definizione 10.2.4 si ottiene infine il risultato richiesto. In questo caso, l'integrale di Itô è una trasformata lineare del quadrato di una v.a. con legge Normale $\mathcal{N}(0, T)$, ovvero $I(X)$ si distribuisce con legge Gamma $\mathcal{G}(-\frac{1}{2}T, T, \frac{1}{2})$. Inoltre, dal Teorema 10.2.6 si ha che $E[I(X)] = 0$ e $\text{Var}[I(X)] = \frac{T^2}{2}$. Si osservi infine che il risultato ottenuto nell'integrazione stocastica è simile a quello che si verifica in un ambito di integrazione con funzioni deterministiche, eccetto la presenza di un secondo termine. Questo termine è dovuto al fatto che la variazione di un moto Browniano non è nulla (vedi Teorema 9.4.11). \square

Dai due precedenti esempi, è evidente che non è opportuno calcolare l'integrale di Itô mediante la Definizione 10.2.4. Tecniche pratiche di integrazione saranno presentate negli Esempi 10.3.5-10.3.8 sulla base della Formula di Itô introdotta nel prossimo paragrafo. Tuttavia, nel caso di un processo deterministico, si può ottenere un risultato generale. In particolare, il seguente Teorema fornisce la legge di un integrale di Wiener.

Teorema 10.2.9. *Sia $X = (X_t)_{t \in [0, T]}$ un p.a. tale che $X_t = g(t)$ q.c. con $g: \mathbb{R} \rightarrow \mathbb{R}$. Se si ha $\int_0^T g(t)^2 dt < \infty$, allora l'integrale di Wiener $I(X) = \int_0^T g(t) dB_t$ è una v.a. con legge Normale $\mathcal{N}(0, \int_0^T g(t)^2 dt)$.*

Dimostrazione. Tenendo presente il Teorema 10.2.5, si consideri la successione di p.a. semplici $(X_n)_{n \geq 1}$, dove $X_n = (X_{t,n})_{t \in [0, T]}$ con

$$X_{t,n} = \sum_{k=0}^{n-1} g(t_k) \mathbf{1}_{[t_k, t_{k+1}[}(t),$$

che approssima il p.a. X . Dalla Definizione 10.1.2 si ha

$$I(X_n) = \sum_{k=0}^{n-1} g(t_k) (B_{t_{k+1}} - B_{t_k}).$$

Gli incrementi di un moto Browniano sono indipendenti per l'assunzione *iii*) della Definizione 9.4.1 e la legge di ogni incremento è Normale per l'assunzione *ii*) della Definizione 9.4.1. Dunque, tenendo presente l'Esempio 7.3.5, la v.a. $I(X_n)$ si distribuisce con legge Normale tale che $E[I(X_n)] = 0$ e

$$\text{Var}[I(X_n)] = \sum_{k=0}^{n-1} g(t_k)^2 (t_{k+1} - t_k).$$

Inoltre, dalla definizione dell'integrale di Riemann, si ha

$$\lim_n \text{Var}[I(X_n)] = \int_0^T g(t)^2 dt.$$

Dal momento che sulla base della Definizione 10.2.4 si ha $I(X_n) \xrightarrow{L^2} I(X)$ per $n \rightarrow \infty$, allora $I(X)$ si distribuisce necessariamente con legge Normale $\mathcal{N}(0, \int_0^T g(t)^2 dt)$ (si veda l'Esercizio 8.1.1). \square

• **Esempio 10.2.4.** Se $X = (X_t)_{t \in [0, T]}$ è tale che $X_t = t$, allora si ha il seguente integrale di Wiener

$$I(X) = \int_0^T t dB_t .$$

Sulla base del Teorema 10.2.9, la v.a. $I(X)$ si distribuisce dunque con legge Normale $\mathcal{N}(0, \frac{T^3}{3})$. \square

Seguendo un percorso analogo a quello che si considera nella teoria dell'integrazione ordinaria, di seguito viene dato il concetto di processo integrale.

Definizione 10.2.10. Si consideri uno spazio di probabilità (Ω, \mathcal{F}, P) con filtrazione $(\mathcal{F}_t)_{t \in [0, \infty[}$ e un moto Browniano $B = (B_t)_{t \in [0, \infty[}$ adattato alla filtrazione. Dato il p.a. $X \in \mathcal{M}_T^2$, si dice *processo integrale* il p.a. $M = (M_t)_{t \in [0, T]}$ tale che

$$M_t = \int_0^t X_s dB_s . \quad \square$$

Il seguente Teorema considera le principali caratteristiche del processo integrale.

Teorema 10.2.11. *Se $M = (M_t)_{t \in [0, T]}$ è un processo integrale, allora M è unico q.c. e $M \in \mathcal{M}_T^2$. Inoltre, M è una martingala rispetto alla filtrazione $(\mathcal{F}_t)_{t \in [0, \infty[}$ con traiettorie continue q.c.*

Dimostrazione. Si veda Da Prato (2014). \square

• **Esempio 10.2.5.** Dall'Esempio 10.2.3, si consideri il processo integrale $M = (M_t)_{t \in [0, T]}$ tale che

$$M_t = \int_0^t B_s dB_s = \frac{1}{2} (B_t^2 - t) .$$

Tenendo presente l'Esempio 10.2.1, si ha

$$\mathbb{E}[\int_0^T M_t^2 dt] \leq \frac{1}{4} \mathbb{E}[\int_0^T (B_t^4 + t^2) dt] = \frac{1}{4} \int_0^T \mathbb{E}[B_t^4] dt + \frac{1}{4} \int_0^T t^2 dt = \frac{T^3}{3} < \infty ,$$

e dunque risulta $M \in \mathcal{M}_T^2$ a conferma della prima parte del Teorema 10.2.11. Inoltre, si ha

$$\mathbb{E}[|M_t|] \leq \frac{1}{2} \mathbb{E}[B_t^2] + \frac{t}{2} = t < \infty .$$

Tenendo presente che il moto Browniano è una martingala (Proposizione 9.4.5), per ogni $s < t$ risulta

$$\begin{aligned} \mathbb{E}[B_t^2 | \mathcal{F}_s] &= \mathbb{E}[(B_t - B_s + B_s)^2 | \mathcal{F}_s] \\ &= \mathbb{E}[(B_t - B_s)^2 | \mathcal{F}_s] + 2 \mathbb{E}[(B_t - B_s)B_s | \mathcal{F}_s] + \mathbb{E}[B_s^2 | \mathcal{F}_s] \\ &= \mathbb{E}[(B_t - B_s)^2] + 2B_s \mathbb{E}[B_t - B_s] + B_s^2 = t - s + B_s^2 , \end{aligned}$$

ovvero

$$\mathbb{E}[B_t^2 - t | \mathcal{F}_s] = B_s^2 - s ,$$

da cui infine si ha

$$\mathbb{E}[M_t | \mathcal{F}_s] = M_s .$$

Dunque, a conferma della seconda parte del Teorema 10.2.11, si ha che il p.a. M è una martingala. \square

10.3. Formula di Itô

Di seguito viene considerata un'ampia classe di processi aleatori, detti processi di Itô, che generalizza la famiglia dei processi integrali presentata nella precedente sezione. Inoltre, viene introdotta la cosiddetta formula di Itô, che permette di determinare la struttura di una trasformata di un processo di Itô e che risulta fondamentale per il calcolo differenziale stocastico.

Definizione 10.3.1. Si consideri uno spazio di probabilità (Ω, \mathcal{F}, P) con filtrazione $(\mathcal{F}_t)_{t \in [0, \infty[}$. Un p.a. $X = (X_t)_{t \in [0, T]}$ è detto *processo di Itô* o *processo di diffusione* se si ha

$$X_t = X_0 + \int_0^t \alpha_s ds + \int_0^t \beta_s dB_s,$$

dove $\alpha = (\alpha_t)_{t \in [0, T]}$ e $\beta = (\beta_t)_{t \in [0, T]}$ sono p.a. tali che $\alpha, \beta \in \mathcal{M}_T^2$. □

Quando si considera il processo di Itô, è uso comune adottare la seguente notazione alternativa

$$dX_t = \alpha_t dt + \beta_t dB_t,$$

e denominare dX_t come *differenziale stocastico*. La precedente espressione viene detta anche *notazione differenziale di Itô*. Tuttavia, si deve evidenziare che il differenziale stocastico non ha un significato matematico ben definito e dovrebbe sempre essere inteso nel senso rigoroso della Definizione 10.3.1. La notazione differenziale di Itô va dunque vista solamente come un modo efficiente di scrivere il processo di Itô.

• **Esempio 10.3.1.** Considerato il p.a. $X = (B_t)_{[0, T]}$ e tenendo presente l'Esempio 10.2.2, si ha

$$B_t = \int_0^t dB_s.$$

Dunque, il moto Browniano è un processo di Itô, dal momento che è immediato verificare che $\alpha, \beta \in \mathcal{M}_T^2$, essendo $\alpha_t = 0$ e $\beta_t = 1$. □

• **Esempio 10.3.2.** Considerato il p.a. $X = (B_t^2)_{[0, T]}$ e tenendo presente l'Esempio 10.2.3 si ha

$$B_t^2 = \int_0^t ds + 2 \int_0^t B_s dB_s$$

Dunque, X è un processo di Itô, dal momento che $\alpha, \beta \in \mathcal{M}_T^2$, essendo $\alpha_t = 1$ e $\beta_t = B_t$ (vedi anche Esempio 10.2.1). Inoltre, in notazione differenziale si può scrivere

$$dB_t^2 = dt + 2B_t dB_t. \quad \square$$

La seguente Proposizione fornisce le caratteristiche principali del processo di Itô.

Proposizione 10.3.2. Se il p.a. $X = (X_t)_{t \in [0, T]}$ è un processo di Itô, allora X ha traiettorie continue q.c. Inoltre, X è una martingala se il p.a. $\alpha = (\alpha_t)_{t \in [0, T]}$ è tale che $\alpha_t = 0$.

Dimostrazione. Segue immediatamente dalle assunzioni e dal Teorema 10.2.10. □

Il seguente Teorema considera la celebrata formula di Itô, che fornisce il effetti la rappresentazione integrale di una trasformata del processo di Itô.

Teorema 10.3.3. (Formula di Itô per trasformate univariate) *Se il p.a. X è un processo di Itô e $g : \mathbb{R} \rightarrow \mathbb{R}$ è una funzione con derivata seconda continua, allora*

$$g(X_t) = g(X_0) + \int_0^t g'(X_s) \alpha_s ds + \int_0^t g'(X_s) \beta_s dB_s + \frac{1}{2} \int_0^t g''(X_s) \beta_s^2 ds,$$

ovvero in notazione differenziale di Itô si ha

$$dg(X_t) = g'(X_t) \alpha_t dt + g'(X_t) \beta_t dB_t + \frac{1}{2} g''(X_t) \beta_t^2 dt = g'(X_t) dX_t + \frac{1}{2} g''(X_t) \beta_t^2 dt.$$

Dimostrazione. Si veda Da Prato (2014). □

È importante notare che nel precedente Teorema la funzione g non può dipendere da t . Di seguito vengono date alcune esemplificazioni di questa importante formula.

• **Esempio 10.3.3.** Si consideri il moto Browniano $X = (B_t)_{[0,T]}$ e la funzione $g(x) = x^m$ con $m \in \{1, 2, \dots\}$. Dal momento che è stato verificato che X è un processo di Itô nell'Esempio 10.3.1, dal Teorema 10.3.3 si ha

$$B_t^m = m \int_0^t B_s^{m-1} dB_s + \frac{m(m-1)}{2} \int_0^t B_s^{m-2} ds,$$

ovvero in notazione differenziale si ha

$$dB_t^m = m B_t^{m-1} dB_t + \frac{m(m-1)}{2} B_t^{m-2} dt.$$

In effetti, questa relazione generalizza quella ottenuta nell'Esempio 10.3.2 nel caso particolare $m = 2$ □

• **Esempio 10.3.4.** Dato il moto Browniano $(B_t)_{[0,T]}$, si consideri il p.a. $X = (X_t)_{[0,T]}$ tale che

$$dX_t = \frac{1}{2} X_t dt + X_t dB_t.$$

Se si suppone che $X \in \mathcal{M}_T^2$, il p.a. X è un processo di Itô con $\alpha = (\frac{1}{2} X_t)_{t \in [0,T]}$ e $\beta = (X_t)_{t \in [0,T]}$. Inoltre, dal Teorema 10.3.3 si ottiene

$$de^{B_t} = e^{B_t} dB_t + \frac{1}{2} e^{B_t} dt.$$

Si deve concludere dunque che $X_t = e^{B_t}$, ovvero si è determinata la struttura del p.a. X che soddisfa alla relazione differenziale stocastica richiesta. In effetti, risulta $X \in \mathcal{M}_T^2$ dal momento che si ha $E[e^{B_t}] = e^{\frac{1}{2}t}$ tenendo presente la Proposizione 9.4.3, e quindi

$$E[\int_0^T e^{B_t} dt] = \int_0^T E[e^{B_t}] dt = 2(e^{\frac{1}{2}T} - 1) < \infty. \quad \square$$

Il seguente Teorema estende la formula di Itô nella situazione in cui si considera due processi di Itô e una loro trasformata. La formula può essere ancora generalizzata al caso di più processi di Itô (per maggiori dettagli, si veda Da Prato, 2014, p.121).

Teorema 10.3.4. (Formula di Itô per trasformate bivariante) *Se i p.a. X e Y sono processi di Itô tali che*

$$dX_t = \alpha_{X,t} dt + \beta_{X,t} dB_t$$

e

$$dY_t = \alpha_{Y,t} dt + \beta_{Y,t} dB_t,$$

mentre $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ è una funzione con derivate parziali seconde continue, allora

$$\begin{aligned} g(X_t, Y_t) &= g(X_0, Y_0) + \int_0^t \frac{\partial g(x, y)}{\partial x} \Big|_{(X_s, Y_s)} \alpha_{X,s} ds + \int_0^t \frac{\partial g(x, y)}{\partial y} \Big|_{(X_s, Y_s)} \alpha_{Y,s} ds \\ &+ \int_0^t \frac{\partial g(x, y)}{\partial x} \Big|_{(X_s, Y_s)} \beta_{X,s} dB_s + \int_0^t \frac{\partial g(x, y)}{\partial y} \Big|_{(X_s, Y_s)} \beta_{Y,s} dB_s \\ &+ \frac{1}{2} \int_0^t \frac{\partial^2 g(x, y)}{\partial x^2} \Big|_{(X_s, Y_s)} \beta_{X,s}^2 ds + \frac{1}{2} \int_0^t \frac{\partial^2 g(x, y)}{\partial y^2} \Big|_{(X_s, Y_s)} \beta_{Y,s}^2 ds \\ &+ \int_0^t \frac{\partial^2 g(x, y)}{\partial x \partial y} \Big|_{(X_s, Y_s)} \beta_{X,s} \beta_{Y,s} ds. \end{aligned}$$

Equivalentemente, in notazione differenziale stocastica si ha

$$\begin{aligned} dg(X_t, Y_t) &= \frac{\partial g(x, y)}{\partial x} \Big|_{(X_t, Y_t)} \alpha_{X,t} dt + \frac{\partial g(x, y)}{\partial y} \Big|_{(X_t, Y_t)} \alpha_{Y,t} dt \\ &+ \frac{\partial g(x, y)}{\partial x} \Big|_{(X_t, Y_t)} \beta_{X,t} dB_t + \frac{\partial g(x, y)}{\partial y} \Big|_{(X_t, Y_t)} \beta_{Y,t} dB_t \\ &+ \frac{1}{2} \frac{\partial^2 g(x, y)}{\partial x^2} \Big|_{(X_t, Y_t)} \beta_{X,t}^2 dt + \frac{1}{2} \frac{\partial^2 g(x, y)}{\partial y^2} \Big|_{(X_t, Y_t)} \beta_{Y,t}^2 dt \\ &+ \frac{\partial^2 g(x, y)}{\partial x \partial y} \Big|_{(X_t, Y_t)} \beta_{X,t} \beta_{Y,t} dt, \end{aligned}$$

ovvero

$$\begin{aligned} dg(X_t, Y_t) &= \frac{\partial g(x, y)}{\partial x} \Big|_{(X_t, Y_t)} dX_t + \frac{\partial g(x, y)}{\partial y} \Big|_{(X_t, Y_t)} dY_t \\ &+ \frac{1}{2} \frac{\partial^2 g(x, y)}{\partial x^2} \Big|_{(X_t, Y_t)} \beta_{X,t}^2 dt + \frac{1}{2} \frac{\partial^2 g(x, y)}{\partial y^2} \Big|_{(X_t, Y_t)} \beta_{Y,t}^2 dt \\ &+ \frac{\partial^2 g(x, y)}{\partial x \partial y} \Big|_{(X_t, Y_t)} \beta_{X,t} \beta_{Y,t} dt. \end{aligned}$$

Dimostrazione. Si veda Da Prato (2014, p.121). □

• **Esempio 10.3.5.** (Formula ridotta di Itô) Si consideri i p.a. di Itô X e Y , dove $dY_t = dt$. Evidentemente, Y è in effetti un processo deterministico. In questo caso, essendo $\alpha_{Y,t} = 1$ e $\beta_{Y,t} = 0$, la formula di Itô del Teorema 10.3.4 si riduce a

$$dg(X_t, t) = \frac{\partial g(x, y)}{\partial x} \Big|_{(X_t, t)} dX_t + \frac{\partial g(x, y)}{\partial y} \Big|_{(X_t, t)} dt + \frac{1}{2} \frac{\partial^2 g(x, y)}{\partial x^2} \Big|_{(X_t, t)} \beta_{X,t}^2 dt,$$

che è anche detta *formula ridotta di Itô*. □

• **Esempio 10.3.6.** (Moto Browniano geometrico) Al fine di esemplificare la formula ridotta di Itô, si supponga che X sia il moto Browniano, ovvero $dX_t = dB_t$, e si consideri inoltre la funzione $g(x, y) = e^{\sigma x + (\mu - \frac{1}{2}\sigma^2)y}$ con $\mu \in \mathbb{R}$ e $\sigma \in]0, \infty[$. Dunque, si ottiene

$$dg(B_t, t) = \sigma g(B_t, t) dB_t + \mu g(B_t, t) dt,$$

ovvero, se si pone $Z = (Z_t)_{[0, T]}$ con $Z_t = e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma B_t}$, si è determinato quindi il p.a. che soddisfa alla relazione differenziale stocastica $dZ_t = \mu Z_t dt + \sigma Z_t dB_t$. Il p.a. Z è detto *moto Browniano geometrico* e riveste grande importanza nelle applicazioni finanziarie. Inoltre, seguendo una terminologia comunemente adottata, μ è detto *parametro di tendenza*, mentre σ è detto *parametro di volatilità*. Il cosiddetto moto Browniano geometrico ridotto si ha per $\mu = 0$ e $\sigma = 1$. Si osservi infine che per $\mu = 0$ il p.a. X si riduce a $dX_t = \sigma X_t dB_t$, ovvero si ha $X_t = X_0 + \sigma \int_0^t X_s dB_s$ in termini di integrale di Itô. Quindi, si è ottenuta la risposta formale al problema introdotto all'inizio della Sezione 10.1 e che ha condotto a considerare l'integrale di Itô. \square

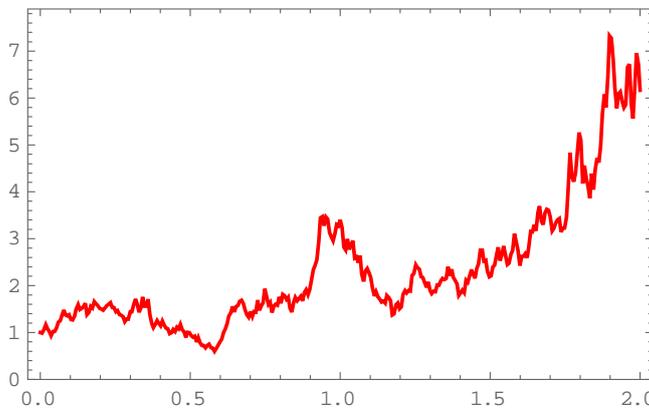


Figura 10.3.1. Una traiettoria del moto Browniano geometrico con $\mu = 1$ e $\sigma = 1$.

• **Esempio 10.3.7.** (Integrazione stocastica per sostituzione) Nella formula ridotta di Itô dell'Esempio 10.3.5 si consideri il p.a. X tale che $dX_t = dB_t$. In questo caso, la formula può essere espressa come

$$dg(B_t, t) = \frac{\partial g(x, y)}{\partial x} \Big|_{(B_t, t)} dB_t + \left(\frac{\partial g(x, y)}{\partial y} + \frac{1}{2} \frac{\partial^2 g(x, y)}{\partial x^2} \right) \Big|_{(B_t, t)} dt.$$

Si assuma che g sia una soluzione della cosiddetta equazione del calore, data dalla seguente equazione differenziale

$$\frac{\partial g(x, y)}{\partial y} + \frac{1}{2} \frac{\partial^2 g(x, y)}{\partial x^2} = 0.$$

In questo caso, la precedente formula ridotta di Itô si semplifica ulteriormente come

$$dg(B_t, t) = \frac{\partial g(x, y)}{\partial x} \Big|_{(B_t, t)} dB_t,$$

ovvero

$$\int_0^t \frac{\partial g(x, y)}{\partial x} \Big|_{(B_s, s)} dB_s = g(B_t, t),$$

che fornisce un'utile espressione per l'*integrazione stocastica per sostituzione*. Nel caso particolare dell'integrale di Itô considerato nell'Esempio 10.2.3, una conveniente soluzione dell'equazione del calore è data da $g(x, y) = \frac{1}{2}(x^2 - y)$, da cui si ottiene immediatamente

$$\int_0^t B_s dB_s = \frac{1}{2} (B_t^2 - t).$$

Come ulteriore caso particolare, se si vuole ottenere l'integrale di Itô del moto Browniano geometrico ridotto $X = (e^{-\frac{t}{2} + B_t})_{t \in [0, T]}$ considerato nell'Esempio 10.3.6, una conveniente soluzione dell'equazione del calore è data da $g(x, y) = e^{x - \frac{y}{2}}$, da cui si ottiene immediatamente

$$\int_0^t e^{B_s - \frac{s}{2}} dB_s = e^{-\frac{t}{2} + B_t}.$$

Questo risultato conferma effettivamente che il moto Browniano geometrico ridotto soddisfa alla relazione differenziale stocastica $dX_t = X_t dB_t$. Per ulteriori dettagli sulle soluzioni dell'equazione del calore si veda Calin (2015, p.157). \square

• **Esempio 10.3.8.** (Integrazione stocastica del prodotto) Si consideri i processi di Itô X e Y e la funzione $g(x, y) = xy$. Calcolando le derivate parziali e sostituendo opportunamente, dal Teorema 10.3.4 si ha dunque

$$\begin{aligned} d(X_t Y_t) &= Y_t \alpha_{X,t} dt + X_t \alpha_{Y,t} dt + Y_t \beta_{X,t} dB_t + X_t \beta_{Y,t} dB_t + \beta_{X,t} \beta_{Y,t} dt \\ &= Y_t dX_t + X_t dY_t + \beta_{X,t} \beta_{Y,t} dt. \end{aligned}$$

Si osservi infine che il risultato ottenuto nell'integrazione stocastica è simile a quello che si ha in un ambito di integrazione con funzioni deterministiche, eccetto la presenza di un terzo termine. Questo termine è dovuto al fatto in generale i processi di Itô X e Y non sono indipendenti, in quanto sono trasformate dello stesso moto Browniano. \square

• **Esempio 10.3.9.** (Integrazione stocastica per parti) Nell'Esempio 10.3.8 si consideri il caso particolare per cui i p.a. X e Y sono tali che $dX_t = h(t)dt$ e $dY_t = g(B_t)dt$ dove $h : \mathbb{R} \rightarrow \mathbb{R}$ è una funzione derivabile e $g : \mathbb{R} \rightarrow \mathbb{R}$ è una funzione con derivata seconda continua. In questo caso, tenendo presente la Formula di Itô (Teorema 10.3.3) si ha

$$d(h(t)g(B_t)) = h'(t)g(B_t)dt + h(t)dg(B_t) = h'(t)g(B_t)dt + h(t)g'(B_t)dB_t + \frac{1}{2} h(t)g''(B_t)dt,$$

ovvero

$$\int_0^t h(s)g'(B_s) dB_s = h(s)g(B_s)|_0^t - \int_0^t h'(s)g(B_s) ds - \frac{1}{2} \int_0^t h(s)g''(B_s) ds.$$

Questa utile espressione è detta formula di *integrazione stocastica per parti*. Nel caso particolare dell'integrale di Itô considerato nell'Esempio 10.2.3, posto $h(t) = 1$ e $g(x) = \frac{1}{2}x^2$, si ha

$$\int_0^t B_s dB_s = \frac{1}{2} B_t^2 - \frac{1}{2} \int_0^t ds = \frac{1}{2} (B_t^2 - t),$$

un risultato che è già stato verificato nel medesimo Esempio. Come ulteriore caso particolare, posto $h(t) = t$ e $g(x) = x$, si ottiene

$$\int_0^t s dB_s = tB_t - \int_0^t B_s ds,$$

che fornisce un interessante relazione fra l'integrale di Wiener e l'integrale del moto Browniano. Per ulteriori esemplificazioni del calcolo pratico dell'integrale di Itô si veda Calin (2015). \square

10.4. Equazioni differenziali stocastiche

Il calcolo differenziale classico è stato sviluppato principalmente al fine di porre in termini di equazioni differenziali i principi fondamentali che governano i fenomeni fisici nel tempo. Il calcolo differenziale stocastico è stato introdotto con simili scopi nel caso di sistemi che si evolvono con una componente aleatoria. Evidentemente, questo tipo di modelli fornisce una descrizione più efficiente dei fenomeni reali.

La seguente definizione introduce una classe di equazioni differenziali che ha importanza fondamentale nelle applicazioni.

Definizione 10.4.1. Si consideri uno spazio di probabilità (Ω, \mathcal{F}, P) con filtrazione $(\mathcal{F}_t)_{t \in [0, \infty[}$. Siano inoltre a e b due funzioni tali che $a : \mathbb{R}^2 \rightarrow \mathbb{R}$ e $b : \mathbb{R}^2 \rightarrow \mathbb{R}$. Un processo di Itô $X = (X_t)_{t \in [0, T]}$ è detto *soluzione dell'equazione differenziale stocastica*

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

con *condizione iniziale* $X_0 = C$, se C è una v.a. misurabile rispetto alla σ -algebra \mathcal{F}_0 , i p.a. $a = (a(X_t, t))_{t \in [0, T]}$ e $b = (b(X_t, t))_{t \in [0, T]}$ sono tali che $a, b \in \mathcal{M}_T^2$ e risulta

$$X_t = C + \int_0^t a(X_s, s) ds + \int_0^t b(X_s, s) dB_s,$$

per ogni t . \square

Il seguente Teorema fornisce le condizioni per l'esistenza e l'unicità di una soluzione di un'equazione differenziale stocastica, che in questo caso viene detta *soluzione forte*.

Teorema 10.4.2. Si consideri uno spazio di probabilità (Ω, \mathcal{F}, P) con filtrazione $(\mathcal{F}_t)_{t \in [0, \infty[}$. Siano inoltre a e b due funzioni $a : \mathbb{R}^2 \rightarrow \mathbb{R}$ e $b : \mathbb{R}^2 \rightarrow \mathbb{R}$ tali che esiste $L > 0$ per cui

$$|a(x, t) - a(y, t)| + |b(x, t) - b(y, t)| \leq L|x - y|,$$

per ogni $x, y \in \mathbb{R}$ e $t \in [0, T]$, e esiste $K > 0$ per cui

$$|a(x, t)| + |b(x, t)| \leq K(1 + |x|),$$

per ogni $x \in \mathbb{R}$ e $t \in [0, T]$. Inoltre, sia C una v.a. misurabile rispetto alla σ -algebra \mathcal{F}_0 , tale che $E[C^2] < \infty$. In questo caso, l'equazione differenziale stocastica

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

con *condizione iniziale* $X_0 = C$, possiede un'unica soluzione X con traiettorie continue e tale che $E[\int_0^T X_t^2 dt] < \infty$.

Dimostrazione. Si veda Da Prato (2014, p.134). \square

• **Esempio 10.4.1.** (Modello di Black e Scholes) Si consideri l'equazione differenziale stocastica

$$dX_t = \mu_t X_t dt + \sigma_t X_t dB_t,$$

con condizione iniziale $X_0 = C$, dove $\mu : \mathbb{R} \rightarrow \mathbb{R}$ e $\sigma : \mathbb{R} \rightarrow \mathbb{R}$ sono funzioni continue in $[0, T]$. Dunque, in questo caso si ha $a(x, t) = \mu_t x$ e $b(x, t) = \sigma_t x$. La prima condizione del Teorema 10.4.2 risulta soddisfatta, in quanto

$$|a(x, t) - a(y, t)| + |b(x, t) - b(y, t)| = (|\mu_t| + |\sigma_t|)|x - y| \leq L|x - y|,$$

dal momento che esiste $L > 0$ per cui $|\mu_t| + |\sigma_t| \leq L$ essendo μ e σ funzioni continue in $[0, T]$. Anche la seconda condizione del Teorema 10.4.2 risulta soddisfatta in modo simile, in quanto

$$|a(x, t)| + |b(x, t)| = (|\mu_t| + |\sigma_t|)|x| \leq K(1 + |x|).$$

Dunque, esiste una soluzione forte. Si consideri inoltre i processi di Itô U e V , tali che $dU_t = \sigma_t dB_t$ e $dV_t = dt$ e la trasformata

$$g(u, v) = c e^{u + \int_0^v (\mu_s - \frac{1}{2}\sigma_s^2) ds},$$

dove c è una determinazione della v.a. C . Si tenga presente che in questo caso si può applicare la formula ridotta di Itô dell'Esempio 10.3.5, data la struttura dei p.a. U e V . Dunque, sostituendo opportunamente in questa formula, si ottiene

$$dg(U_t, t) = \mu_t g(U_t, t) dt + \sigma_t g(U_t, t) dB_t.$$

Data l'equivalenza di questa espressione con l'equazione differenziale stocastica iniziale, la soluzione $X = (X_t)_{t \in [0, T]}$ è tale che

$$X_t = g(U_t, t) = C e^{\int_0^t (\mu_s - \frac{1}{2}\sigma_s^2) ds + \int_0^t \sigma_s dB_s}.$$

Questo p.a. è detto *modello di Black e Scholes*, dal momento che è stato introdotto dal matematico Fischer Sheffey Black (1938-1995) e dall'economista Myron Samuel Scholes (1941-). Il modello è in effetti una generalizzazione del moto Browniano geometrico introdotto nell'Esempio 10.3.6, che si ottiene per $\mu_t = \mu$ e $\sigma_t = \sigma$ funzioni costanti. Il modello di Black and Scholes riveste un'importanza fondamentale nella matematica finanziaria. \square

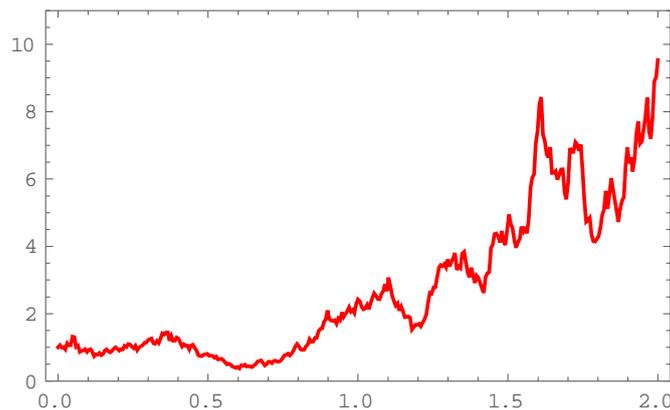


Figura 10.4.1. Una traiettoria del modello di Black e Scholes con $\mu_t = t$ e $\sigma_t = 1$.

• **Esempio 10.4.2.** (Processo di Ornstein e Uhlenbeck) Si consideri l'equazione differenziale stocastica

$$dX_t = -\mu X_t dt + \sigma dB_t,$$

con condizione iniziale $X_0 = C$, dove $\mu, \sigma \in]0, \infty[$. Dunque, in questo caso si ha $a(x, t) = -\mu x$ e $b(x, t) = \sigma$. La prima condizione del Teorema 10.4.2 risulta soddisfatta, in quanto

$$|a(x, t) - a(y, t)| + |b(x, t) - b(y, t)| = |\mu||x - y| .$$

Anche la seconda condizione del Teorema 10.4.2 risulta soddisfatta, in quanto

$$|a(x, t)| + |b(x, t)| = |\mu||x| + |\sigma| .$$

Quindi, esiste una soluzione forte. Tenendo presente la formula di integrazione stocastica del prodotto introdotta nell'Esempio 10.3.8, con il p.a. Y tale che $Y_t = e^{\mu t}$, si ha

$$d(X_t e^{\mu t}) = e^{\mu t} dX_t + \mu e^{\mu t} X_t dt = \sigma e^{\mu t} dB_t .$$

Dalla precedente relazione risulta dunque

$$e^{\mu t} X_t = C + \sigma \int_0^t e^{\mu s} dB_s ,$$

ovvero, la soluzione $X = (X_t)_{t \in [0, T]}$ dell'equazione differenziale stocastica è tale che

$$X_t = C e^{-\mu t} + \sigma e^{-\mu t} \int_0^t e^{\mu s} dB_s .$$

Questo p.a. è detto *processo di Ornstein e Uhlenbeck*, dal momento che è stato introdotto dai fisici Leonard Salomon Ornstein (1880-1941) e George Eugene Uhlenbeck (1900-1988). Anche questo modello ha grande rilevanza nella matematica finanziaria. \square

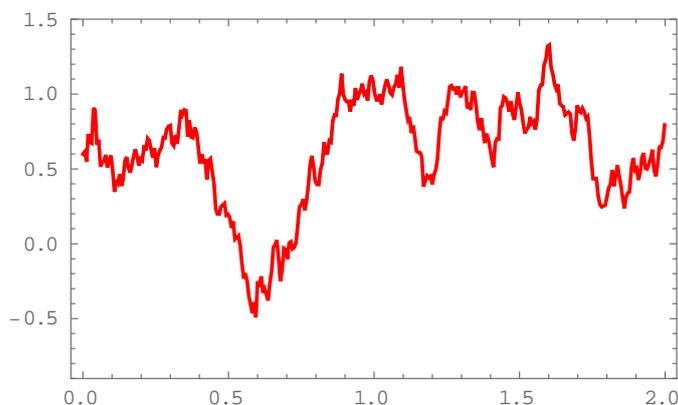


Figura 10.4.2. Una traiettoria del processo di Ornstein e Uhlenbeck con $\mu = 1$ e $\sigma = 1$.

• **Esempio 10.4.3.** (Ponte Browniano) Si consideri l'equazione differenziale stocastica

$$dX_t = \frac{1}{1-t} (c - X_t) dt + dB_t ,$$

con condizione iniziale $X_0 = 0$ e $c \in \mathbb{R}$, dove $t \in [0, 1[$. Dunque, in questo caso si ha $a(x, t) = \frac{c-x}{1-t}$ e $b(x, t) = 1$. La prima condizione del Teorema 10.4.2 risulta soddisfatta, in quanto

$$|a(x, t) - a(y, t)| + |b(x, t) - b(y, t)| = \frac{1}{1-t} |x - y| .$$

Anche la seconda condizione del Teorema 10.4.2 risulta soddisfatta, in quanto

$$|a(x, t)| + |b(x, t)| = \frac{1}{1-t} |c - x| + 1 .$$

Dunque, esiste una soluzione forte. Si consideri l'ulteriore processo di Itô $Y = (Y_t)_{t \in [0,1]}$, tale che $Y_t = \frac{1}{1-t} dt$. Evidentemente, Y è un processo deterministico. Dalla formula di integrazione stocastica del prodotto dell'Esempio 10.3.8, si ha

$$d(X_t Y_t) = \frac{1}{1-t} dX_t + \frac{1}{(1-t)^2} X_t dt = \frac{c}{(1-t)^2} dt + \frac{1}{1-t} dB_t,$$

ovvero

$$\frac{1}{1-t} X_t = X_0 + c \int_0^t \frac{1}{(1-s)^2} ds + \int_0^t \frac{1}{1-s} dB_s.$$

Dunque, la soluzione $X = (X_t)_{t \in [0,1]}$ dell'equazione differenziale stocastica è tale che

$$X_t = ct + (1-t) \int_0^t \frac{1}{1-s} dB_s.$$

Questo p.a. è detto in modo pittoresco *ponte Browniano*, dal momento che è equivalente ad un moto Browniano le cui traiettorie si originano in $(0,0)$ e terminano in $(1,c)$. \square

La seguente Proposizione fornisce un metodo per ottenere la soluzione di un particolare tipo di equazione differenziale stocastica. La tecnica è anche detta *metodo del fattore d'integrazione*.

Proposizione 10.4.3. *Si consideri uno spazio di probabilità (Ω, \mathcal{F}, P) con filtrazione $(\mathcal{F}_t)_{t \in [0, \infty[}$. Sia dato un processo di Itô $X = (X_t)_{t \in [0, T]}$, che è soluzione dell'equazione differenziale stocastica*

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

dove $b : \mathbb{R} \rightarrow \mathbb{R}$ è una funzione continua. Sia dato inoltre il p.a. $Y = (Y_t)_{t \in [0, T]}$ con

$$Y_t = e^{\frac{1}{2} \int_0^t b(s)^2 ds - \int_0^t b(s) dB_s}.$$

Si ha

$$d(X_t Y_t) = a(X_t, t) Y_t dt.$$

Dimostrazione. Se $Z = (Z_t)_{t \in [0, T]}$ è il processo di Itô tale che

$$dZ_t = \frac{1}{2} b(t)^2 dt - b(t) dB_t$$

e tenendo presente che $Y_t = e^{Z_t}$, dalla formula di Itô (Teorema 10.3.3) si ha

$$dY_t = e^{Z_t} dZ_t + \frac{1}{2} e^{Z_t} b(t)^2 dt = b(t)^2 Y_t dt - b(t) Y_t dB_t.$$

Inoltre, dalla formula di integrazione stocastica del prodotto introdotta nell'Esempio 10.3.8 si ottiene

$$\begin{aligned} d(X_t Y_t) &= Y_t(a(X_t, t) dt + b(t) X_t dB_t) + X_t(b(t)^2 Y_t dt - b(t) Y_t dB_t) - b(t)^2 X_t Y_t dt \\ &= a(X_t, t) Y_t dt, \end{aligned}$$

da cui segue la tesi. \square

Si osservi che il risultato fornito nella Proposizione 10.4.3 permette di ricondurre l'equazione differenziale stocastica originale ad una equazione differenziale ordinaria, che in pratica può essere risolta con l'utilizzo di metodi standard. Nelle applicazioni usuali, molte equazioni differenziali

stocastiche sono esprimibili nella forma considerata nella Proposizione 10.4.3 e dunque questo risultato fornisce un metodo di soluzione molto efficiente.

• **Esempio 10.4.4.** Si consideri l'equazione differenziale stocastica relativa al modello di Black e Scholes dell'Esempio 10.4.1. In questo caso, applicando la Proposizione 10.4.3 si ha

$$Y_t = e^{\frac{1}{2}\int_0^t \sigma_s^2 ds - \int_0^t \sigma_s dB_s},$$

mentre

$$d(X_t Y_t) = \mu_t X_t Y_t dt.$$

Dunque, tenendo presente che $\{Y_0 = 1\}$ q.c., si ha

$$X_t Y_t = C e^{\int_0^t \mu_s ds},$$

ovvero

$$X_t = C e^{\int_0^t (\mu_s - \frac{1}{2}\sigma_s^2) ds + \int_0^t \sigma_s dB_s},$$

che conferma il risultato ottenuto nell'Esempio 10.4.1. □

• **Esempio 10.4.5.** Si consideri l'equazione differenziale stocastica

$$dX_t = \mu dt + \sigma X_t dB_t,$$

con condizione iniziale $X_0 = C$ e dove $\mu, \sigma \in \mathbb{R}$. Dunque, in questo caso si ha $a(x, t) = \mu$ e $b(x, t) = \sigma x$. La prima condizione del Teorema 10.4.2 risulta soddisfatta, in quanto

$$|a(x, t) - a(y, t)| + |b(x, t) - b(y, t)| = |\sigma||x - y|.$$

Anche la seconda condizione del Teorema 10.4.2 risulta soddisfatta, in quanto

$$|a(x, t)| + |b(x, t)| = |\mu| + |\sigma x|.$$

Quindi, esiste una soluzione forte. In questo caso, applicando la Proposizione 10.4.3 si ha

$$Y_t = e^{\frac{1}{2}\sigma^2 t - \sigma B_t},$$

mentre

$$d(X_t Y_t) = \mu Y_t dt,$$

ovvero, tenendo presente che $\{Y_0 = 1\}$ q.c., risulta

$$X_t Y_t = C + \mu \int_0^t Y_s ds.$$

Dunque, la soluzione X dell'equazione differenziale stocastica è tale che

$$X_t = e^{-\frac{1}{2}\sigma^2 t + \sigma B_t} \left(C + \mu \int_0^t e^{\frac{1}{2}\sigma^2 s - \sigma B_s} ds \right). \quad \square$$

Un ulteriore metodo di soluzione è basato sulle cosiddette *equazioni stocastiche esatte*. Si consideri l'equazione differenziale stocastica della Definizione 10.4.1 e si voglia determinare una soluzione forte del tipo $X_t = g(B_t, t)$. In questo caso, l'equazione differenziale stocastica si riduce a

$$dg(B_t, t) = a(g(B_t, t), t) dt + b(g(B_t, t), t) dB_t.$$

Inoltre, tenendo presente la formula ridotta di Itô (Esempio 10.3.5), l'equazione differenziale stocastica è esatta se esiste una funzione g per cui si ha

$$a(g(x, y), y) = \frac{\partial g(x, y)}{\partial y} + \frac{1}{2} \frac{\partial^2 g(x, y)}{\partial x^2}$$

e

$$b(g(x, y), y) = \frac{\partial g(x, y)}{\partial x}.$$

Si osservi che in questo caso sussiste la relazione

$$\frac{\partial a(g(x, y), y)}{\partial x} = \frac{\partial b(g(x, y), y)}{\partial y} + \frac{1}{2} \frac{\partial^2 b(g(x, y), y)}{\partial x^2}.$$

In effetti, dalla formula ridotta di Itô, per questa scelta della funzione g si ha

$$dX_t = dg(B_t, t),$$

ovvero

$$X_t = g(B_t, t).$$

• **Esempio 10.4.5.** Si consideri il caso specifico dell'equazione differenziale stocastica

$$dX_t = \mu X_t dt + \sigma X_t dB_t.$$

Evidentemente, le condizioni del Teorema 10.4.2 sono soddisfatte e quindi esiste una soluzione forte. In questo caso, essendo $b(g(x, y), y) = \sigma g(x, y)$, si ha

$$\sigma g(x, y) = \frac{\partial g(x, y)}{\partial x},$$

da cui

$$g(x, y) = e^{\sigma x + h(y)},$$

dove h è una funzione da determinare. Inoltre, essendo $a(g(x, y), y) = \mu g(x, y)$ risulta

$$\mu e^{\sigma x + h(y)} = e^{\sigma x + h(y)} h'(y) + \frac{\sigma^2}{2} e^{\sigma x + h(y)},$$

da cui

$$h(x) = \left(\mu - \frac{1}{2} \sigma^2\right)t + c$$

con $c \in \mathbb{R}$. Dunque, si ottiene

$$g(x, y) = e^{(\mu - \frac{1}{2} \sigma^2)t + \sigma x + c}.$$

Dal momento che $g(0, 0) = e^c$, si deve quindi concludere che

$$X_t = X_0 e^{(\mu - \frac{1}{2} \sigma^2)t + \sigma B_t},$$

ovvero si è ottenuto il moto Browniano geometrico dell'Esempio 10.3.6. □

• **Esempio 10.4.6.** Si consideri l'equazione differenziale stocastica

$$dX_t = a(X_t) dt + b(X_t) dB_t$$

e si voglia determinare l'esistenza di una soluzione forte del tipo $X_t = g(B_t)$. In questo caso, la precedente equazione differenziale stocastica si riduce a

$$dg(B_t) = a(g(B_t)) dt + b(g(B_t)) dB_t .$$

Dalla formula di Itô del Teorema 10.3.3 si ha inoltre

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

Dunque, l'equazione differenziale stocastica è esatta se esiste una funzione g per cui si ha

$$a(g(x)) = \frac{1}{2} g''(x)$$

e

$$b(g(x)) = g'(x) ,$$

ovvero sussiste la relazione

$$a(g(x)) = \frac{1}{2} b'(g(x))g'(x) .$$

In effetti, per questa scelta risulta

$$dX_t = dg(B_t) ,$$

ovvero

$$X_t = g(B_t) .$$

Si consideri il caso specifico dell'equazione differenziale stocastica

$$dX_t = \frac{1}{2} X_t dt + \sqrt{1 + X_t^2} dB_t .$$

Le condizioni del Teorema 10.4.2 sono facilmente verificate e quindi esiste una soluzione forte. Sulla base della precedente relazione si ottiene la seguente equazione differenziale

$$g'(x) = \sqrt{1 + g(x)^2} ,$$

la cui soluzione è data da

$$g(x) = \sinh(x + c) ,$$

con $c \in \mathbb{R}$. Dunque, si deve concludere che

$$X_t = \sinh(B_t + \operatorname{arcsinh}(X_0)) .$$

□

10.5. Riferimenti bibliografici

I testi di Applebaum (2009), Baldi (2017), Chung e Williams (1990), Karatzas e Shreve (1991), Klebaner (2005), Kuo (2006), Le Gall (2016), Medvegyev (2007), Øksendal (2003), Pascucci (2011), Protter (2003) e Shreve (2004) sono dedicati al calcolo stocastico. Testi avanzati su questo argomento sono Borodin (2017), Capasso e Bakstein (2015), Cohen e Elliott (2015), Da Prato (2014) e Revuz e

Yor (2005). Per un approccio introduttivo al calcolo stocastico si possono consultare i testi di Arguin (2022), Brzeźniak e Zastawniak (2002), Calin (2015), Choe (2016) e Wiersema (2008).

10.6. Esercizi svolti

• **Nota.** Al fine di evitare notazioni ridondanti e ripetizioni, nei seguenti esercizi si assume l'esistenza di uno spazio probabilizzato (Ω, \mathcal{F}, P) e di una filtrazione $(\mathcal{F}_t)_{t \in \mathbb{T}}$, a meno che non venga specificato diversamente.

Sezione 10.1

• **Esercizio 10.1.1.** (Moto Browniano integrato) Se $B = (B_t)_{t \in \mathbb{R}^+}$ è un moto Browniano, si consideri il p.a. $X = (X_t)_{t \in \mathbb{R}^+}$ dove

$$X_t = \int_0^t B_s ds.$$

Si determini la legge relativa alla v.a. X_t . Il p.a. X è detto moto Browniano integrato.

Soluzione. Dal momento che B ha traiettorie continue *q.c.*, il p.a. X è definito. Inoltre, si consideri la successione di v.a. $(X_{t,n})_{n \geq 1}$ tale che

$$X_{t,n} = \frac{t}{n} \sum_{k=1}^n B_{kt/n}$$

e si osservi che $\lim_n X_{t,n} = X_t$ *q.c.* per le proprietà dell'integrale di Riemann. Mediante una semplice manipolazione algebrica $X_{t,n}$ può essere espresso come

$$X_{t,n} = \frac{t}{n} \sum_{k=1}^n (n - k + 1)(B_{kt/n} - B_{(k-1)t/n}).$$

Per le proprietà del moto Browniano gli incrementi $(B_{kt/n} - B_{(k-1)t/n})$ sono indipendenti con legge Normale $\mathcal{N}(0, \frac{t}{n})$. Quindi, per le proprietà della legge Normale, la v.a. $X_{t,n}$ è distribuita con legge Normale tale che $E[X_{t,n}] = 0$ e

$$\text{Var}[X_{t,n}] = \frac{t^3}{n^3} \sum_{k=1}^n (n - k + 1)^2 = \frac{t^3}{n^3} \sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)t^3}{6n^3}.$$

Dal momento che $\lim_n E[X_{t,n}] = 0$ e $\lim_n \text{Var}[X_{t,n}] = \frac{t^3}{3}$, sulla base dell'Esercizio 8.1.1 la v.a. X_t si distribuisce con legge Normale $\mathcal{N}(0, \frac{t^3}{3})$. \square

Sezione 10.2

• **Esercizio 10.2.1.** Se $B = (B_t)_{t \in \mathbb{R}^+}$ è un moto Browniano, si consideri il p.a. $X = (X_t)_{t \in [0, T]}$ dove $X_t = \int_0^t e^{sT} dB_s$. Se

$$I(X) = \int_0^T X_t dB_t,$$

si calcoli $E[I(X)]$ e $\text{Var}[I(X)]$.

Soluzione. Dal momento che X_t è un integrale di Wiener, allora si ha $E[X_t] = 0$ e

$$E[X_t^2] = \text{Var}[X_t] = \int_0^t e^{2s} ds = \frac{1}{2} (e^{2t} - 1).$$

Inoltre, risulta

$$E[\int_0^T X_t^2 dt] = \int_0^T E[X_t^2] dt = \frac{1}{4} (e^{2T} - 2T - 1) < \infty$$

e dunque $X \in \mathcal{M}_T^2$. Quindi, si ottiene che $E[I(X)] = 0$ e

$$\text{Var}[I(X)] = \int_0^T E[X_t^2] dt = \frac{1}{4} (e^{2T} - 2T - 1). \quad \square$$

• **Esercizio 10.2.2.** Se $B = (B_t)_{t \in \mathbb{R}^+}$ è un moto Browniano, si determini $E[B_s \int_0^T B_t dB_t]$ dove $s \leq T$.

Soluzione. Considerato il p.a. degenere $X = (X_t)_{t \in [0, T]}$ dove $X_t = \mathbf{1}_{[0, s]}(t)$, si osservi che risulta $X \in \mathcal{M}_T^2$ e si ponga

$$I(X) = \int_0^T \mathbf{1}_{[0, s]}(t) dB_t = B_s.$$

Inoltre, se $Y = (Y_t)_{t \in [0, T]}$ dove $Y_t = B_t$, ovviamente si ha $Y \in \mathcal{M}_T^2$ e si assuma che

$$I(Y) = \int_0^T B_t dB_t.$$

Dunque, si deve concludere che

$$E[B_s \int_0^T B_t dB_t] = E[I(X)I(Y)] = \int_0^T E[X_t Y_t] dt = \int_0^T E[\mathbf{1}_{[0, s]}(t) B_t] dB_t = 0. \quad \square$$

Sezione 10.3

• **Esercizio 10.3.1.** Se $B = (B_t)_{t \in \mathbb{R}^+}$ è un moto Browniano, si consideri il ponte Browniano $X = (X_t)_{t \in [0, 1]}$ tale che

$$X_t = (1 - t) \int_0^t \frac{1}{1 - s} dB_s$$

e si calcoli dX_t .

Soluzione. Applicando la formula di integrazione stocastica per il prodotto, ponendo $U_t = \int_0^t (1 - s)^{-1} dB_s$ e $V_t = 1 - t$, dal momento che

$$dU_t = \frac{1}{1 - t} dB_t,$$

si ha

$$d(U_t V_t) = V_t dU_t + U_t dV_t = dB_t - U_t dt,$$

ovvero

$$dX_t = -\frac{1}{1 - t} X_t dt + dB_t.$$

Ovviamente, la precedente espressione rappresenta l'equazione differenziale stocastica che definisce il ponte Browniano. Applicando la formula ridotta di Itô con $g(u, v) = u(1 - v)$, si ottiene lo stesso risultato, ovvero

$$dg(U_t, t) = \frac{\partial g(u, v)}{\partial u} \Big|_{(U_t, t)} dU_t + \frac{\partial g(u, v)}{\partial v} \Big|_{(U_t, t)} dt = (1 - t) dU_t - U_t dt = dB_t - U_t dt. \quad \square$$

• **Esercizio 10.3.2.** Se $B = (B_t)_{t \in \mathbb{R}^+}$ è un moto Browniano, si consideri il p.a. $X = (X_t)_{t \in \mathbb{R}^+}$ tale che

$$X_t = tB_t - \int_0^t B_s ds$$

e si calcoli dX_t . Si verifichi inoltre che il p.a. X è una martingala.

Soluzione. Si ha

$$dX_t = d(tB_t) - B_t dt.$$

e, applicando la formula di integrazione stocastica per il prodotto, risulta

$$dX_t = t dB_t + B_t dt - B_t dt = t dB_t.$$

Dal momento che $\{X_0 = 0\}$ q.c., dalla precedente espressione si ha anche

$$X_t = \int_0^t s dB_s.$$

Dunque, la v.a. X_t è in effetti un integrale di Wiener distribuito con legge Normale $\mathcal{N}(0, \frac{t^3}{3})$. Questo risultato permette di concludere che il p.a. considerato è una martingala. In effetti, il procedimento considerato è generale, nel senso che se il differenziale stocastico di un p.a. non dipende dalla variazione (ovvero, in termini informali, non contiene l'addendo in dt) allora il p.a. è un integrale di Itô ed è necessariamente una martingala. \square

• **Esercizio 10.3.3.** (Processo di Bessel in \mathbb{R}^2) Se $B = (B_t)_{t \in \mathbb{R}^+}$ e $W = (W_t)_{t \in \mathbb{R}^+}$ sono due moti Browniani indipendenti, si consideri il p.a. $X = (X_t)_{t \in \mathbb{R}^+}$ tale che

$$X_t = \sqrt{B_t^2 + W_t^2}$$

e si calcoli dX_t . Il p.a. X è comunemente detto processo di Bessel in \mathbb{R}^2 .

Soluzione. Se $g: \mathbb{R}^2 \rightarrow \mathbb{R}$ è una funzione con derivate parziali seconde continue, la generalizzazione della formula di Itô per trasformate bivariate al caso di due moti Browniani indipendenti fornisce l'espressione

$$dg(B_t, W_t) = \frac{\partial g(u, v)}{\partial u} \Big|_{(B_t, W_t)} dB_t + \frac{\partial g(u, v)}{\partial v} \Big|_{(B_t, W_t)} dW_t + \frac{1}{2} \left(\frac{\partial^2 g(u, v)}{\partial u^2} \Big|_{(B_t, W_t)} + \frac{\partial^2 g(u, v)}{\partial v^2} \Big|_{(B_t, W_t)} \right) dt.$$

Applicando la precedente formula di Itô con $g(u, v) = \sqrt{u^2 + v^2}$, risulta

$$dX_t = \frac{B_t}{X_t} dB_t + \frac{W_t}{X_t} dW_t + \frac{1}{2} \left(\frac{W_t^2}{X_t^3} + \frac{B_t^2}{X_t^3} \right) dt = \frac{B_t}{X_t} dB_t + \frac{W_t}{X_t} dW_t + \frac{1}{2X_t} dt. \quad \square$$

• **Esercizio 10.3.4.** Se $B = (B_t)_{t \in \mathbb{R}^+}$ è un moto Browniano, si considerino i p.a. $X = (X_t)_{t \in \mathbb{R}^+}$ e $Y = (Y_t)_{t \in \mathbb{R}^+}$ dove $X_t = \int_0^t s dB_s$ e $Y_t = \int_0^t B_s ds$. Si determini il coefficiente di correlazione ρ_{X_t, Y_t} .

Soluzione. Dall'Esercizio 10.3.2 si ha che $X_t + Y_t = tB_t$ e quindi

$$\text{Var}[X_t] + \text{Var}[Y_t] + 2 \text{Cov}[X_t, Y_t] = t^2 \text{Var}[B_t].$$

La v.a. X_t è un integrale di Wiener ed è una v.a. con legge Normale $\mathcal{N}(0, \frac{t^3}{3})$. Inoltre, tenendo presente l'Esercizio 10.1.1, il p.a. Y è un moto Browniano integrato e quindi Y_t è una v.a. con legge Normale $\mathcal{N}(0, \frac{t^3}{3})$. Dalla precedente relazione si ha dunque infine che $\text{Cov}[X_t, Y_t] = \frac{t^3}{6}$ e quindi $\rho_{X_t, Y_t} = \frac{1}{2}$. \square

• **Esercizio 10.3.5.** Se $B = (B_t)_{t \in \mathbb{R}^+}$ è un moto Browniano, si consideri il p.a. $X = (X_t)_{t \in \mathbb{R}^+}$ dove

$$X_t = \int_0^t e^{\frac{1}{2}s} \cos(B_s) dB_s.$$

Si determini un'espressione esplicita per il precedente integrale stocastico.

Soluzione. Tenendo presente la formula per l'integrazione stocastica per parti con $h(t) = e^{\frac{1}{2}t}$ e $g(x) = \sin(x)$ si ha

$$\begin{aligned} \int_0^t e^{\frac{1}{2}s} \cos(B_s) dB_s &= e^{\frac{1}{2}s} \sin(B_s) \Big|_0^t - \frac{1}{2} \int_0^t e^{\frac{1}{2}s} \sin(B_s) ds + \frac{1}{2} \int_0^t e^{\frac{1}{2}s} \sin(B_s) ds \\ &= e^{\frac{1}{2}t} \sin(B_t). \end{aligned} \quad \square$$

• **Esercizio 10.3.6.** (Generalizzazione del moto Browniano integrato) Se $B = (B_t)_{t \in \mathbb{R}^+}$ è un moto Browniano, si consideri il p.a. $X = (X_t)_{t \in \mathbb{R}^+}$ con

$$X_t = \int_0^t f(s) B_s ds,$$

dove f è una funzione integrabile. Si determini la legge relativa alla v.a. X_t .

Soluzione. Posto $h(t) = \int_0^t f(s) ds$ e $g(s) = s$, manipolando in modo opportuno la formula per l'integrazione stocastica per parti, si ottiene

$$\int_0^t f(s) B_s ds = h(t) B_t - \int_0^t h(s) dB_s = \int_0^t (h(t) - h(s)) dB_s.$$

Dal momento che la generalizzazione del moto Browniano integrato è espressa come integrale di Wiener, la v.a. X_t è distribuita con legge Normale $\mathcal{N}(0, \int_0^t (h(t) - h(s))^2 ds)$. Nel caso in cui $f(s) = 1$, si ha il moto Browniano integrato, per cui la legge della v.a. X_t è Normale $\mathcal{N}(0, \frac{t^3}{3})$. Si noti che questo risultato è ottenuto in modo più elegante e immediato rispetto all'Esercizio 10.1.1. \square

Sezione 10.4

• **Esercizio 10.4.1.** (Processo di Ornstein e Uhlenbeck generalizzato) Si consideri l'equazione differenziale stocastica

$$dX_t = -\mu(t)X_t dt + \sigma(t) dB_t,$$

con condizione iniziale $X_0 = C$, dove $\mu : \mathbb{R} \rightarrow \mathbb{R}$ e $\sigma : \mathbb{R} \rightarrow \mathbb{R}^+$ sono funzioni localmente limitate, e se ne determini la soluzione.

Soluzione. Si ha $a(x, t) = -\mu(t)x$ e $b(x, t) = \sigma(t)$, per cui

$$|a(x, t) - a(y, t)| + |b(x, t) - b(y, t)| = |\mu(t)||x - y|$$

e

$$|a(x, t)| + |b(x, t)| = |\mu(t)||x| + |\sigma(t)|.$$

Dal momento che μ e σ sono localmente limitate, le condizioni per l'esistenza e l'unicità dell'equazione differenziale stocastica sono soddisfatte ed esiste una soluzione forte. Assumendo che $A(t) = \int_0^t \mu(s) ds$ e tenendo presente la formula di integrazione stocastica per il prodotto, si ha

$$d(X_t e^{A(t)}) = e^{A(t)} dX_t + \mu(t) e^{A(t)} X_t dt = \sigma(t) e^{A(t)} dB_t.$$

Dalla precedente relazione ed essendo $e^{A(0)} = 1$, risulta dunque

$$e^{A(t)} X_t = C + \int_0^t \sigma(s) e^{A(s)} dB_s,$$

ovvero l'equazione differenziale stocastica ha soluzione $X = (X_t)_{t \in \mathbb{R}^+}$ con

$$X_t = C e^{-A(t)} + e^{-A(t)} \int_0^t \sigma(s) e^{A(s)} dB_s.$$

Si osservi che $\int_0^t \sigma(s) e^{A(s)} dB_s$ è un integrale di Wiener e quindi la legge della v.a. X_t condizionata all'evento $\{C = c\}$ è Normale $\mathcal{N}(c e^{-A(t)}, e^{-2A(t)} \int_0^t \sigma(s)^2 e^{2A(s)} ds)$. \square

• **Esercizio 10.4.2.** Si consideri l'equazione differenziale stocastica

$$dX_t = -\mu X_t dt + \sigma dB_t,$$

con condizione iniziale $X_0 = C$, dove $\mu, \sigma \in]0, \infty[$, che ha per soluzione il processo di Ornstein e Uhlenbeck. Si determini la legge della v.a. Y tale che $X_t \xrightarrow{\mathcal{L}} Y$ per $t \rightarrow \infty$.

Soluzione. Dall'Esercizio 10.4.1 risulta che la soluzione dell'equazione differenziale stocastica considerata è data da $X = (X_t)_{t \in \mathbb{R}^+}$ dove

$$X_t = C e^{-\mu t} + \sigma e^{-\mu t} \int_0^t e^{\mu s} dB_s.$$

e la legge della v.a. X_t condizionata all'evento $\{C = c\}$ è Normale $\mathcal{N}(c e^{-\mu t}, \frac{\sigma^2}{2\mu}(1 - e^{-2\mu t}))$. Dal momento che $\mu \in]0, \infty[$, si ha

$$\lim_{t \rightarrow \infty} \mathbb{E}[X_t] = 0$$

e

$$\lim_{t \rightarrow \infty} \text{Var}[X_t] = \frac{\sigma^2}{2\mu}.$$

Quindi, in base all'Esercizio 8.1.1, la legge della v.a. Y è Normale $\mathcal{N}(0, \frac{\sigma^2}{2\mu})$. Si osservi che la legge limite non dipende dalla condizione iniziale $X_0 = C$. Inoltre, si assuma che X_0 sia una v.a. indipendente dal moto Browniano e tale che $X_0 \stackrel{\mathcal{L}}{=} Y$. In questo caso, poichè $\int_0^t e^{\mu s} dB_s$ è una v.a. con legge Normale $\mathcal{N}(0, \frac{1}{2\mu} e^{2\mu t})$, allora è immediato verificare che la legge della v.a. X_t è ancora Normale $\mathcal{N}(0, \frac{\sigma^2}{2\mu})$. \square

• **Esercizio 10.4.3.** Assumendo che $t \in [0, 1[$, si consideri l'equazione differenziale stocastica

$$dX_t = -\frac{1}{1-t} X_t dt + dB_t,$$

con condizione iniziale $X_0 = C$, e se ne determini la soluzione.

Soluzione. Tenendo presente l'Esercizio 10.4.1, si ha $\mu(t) = \frac{1}{1-t}$ e $\sigma(t) = 1$, per cui

$$A(t) = \int_0^t \frac{1}{1-s} ds = -\log(1-t).$$

L'equazione differenziale stocastica ha soluzione $X = (X_t)_{t \in [0,1[}$ con

$$X_t = C(1-t) + (1-t) \int_0^t \frac{1}{1-s} dB_s$$

e la legge della v.a. X_t condizionata all'evento $\{C = c\}$ è Normale $\mathcal{N}(c(1-t), t(1-t))$. Inoltre, se $\{C = 0\}$ q.c., il p.a. X è un ponte Browniano le cui traiettorie si originano in $(0,0)$ e terminano in $(1,0)$. \square

• **Esercizio 10.4.4.** Si consideri l'equazione differenziale stocastica

$$dX_t = -\frac{1}{2} X_t dt + \sqrt{1-X_t^2} dB_t.$$

con condizione iniziale $X_0 = C$, e se ne determini la soluzione.

Soluzione. Si ha $a(x,t) = -\frac{1}{2}x$ e $b(x,t) = \sqrt{1-x^2}$, per cui

$$|a(x,t) - a(y,t)| + |b(x,t) - b(y,t)| = \frac{1}{2}|x-y| + |\sqrt{1-x^2} - \sqrt{1-y^2}| \leq 2|x-y|$$

e

$$|a(x,t)| + |b(x,t)| = \frac{1}{2}|x| + \sqrt{1-x^2} \leq 2(1+|x|).$$

Dunque le condizioni per l'esistenza e l'unicità dell'equazione differenziale stocastica sono soddisfatte ed esiste una soluzione forte. L'equazione differenziale stocastica è esatta e può essere risolta considerando l'equazione differenziale

$$g'(x) = \sqrt{1-g(x)^2},$$

la cui soluzione è data da

$$g(x) = \sin(x+c),$$

con $c \in \mathbb{R}$. Dunque, si deve concludere che la soluzione dell'equazione differenziale stocastica è data da $X = (X_t)_{t \in \mathbb{R}^+}$ con

$$X_t = \sin(B_t + \arcsin(C)).$$

In effetti, applicando la formula ridotta di Itô si ottiene

$$d\sin(B_t) = \cos(B_t) dB_t - \frac{1}{2} \sin(B_t) dt = \sqrt{1-\sin^2(B_t)} dB_t - \frac{1}{2} \sin(B_t) dt. \quad \square$$

• **Esercizio 10.4.5.** Si consideri l'equazione differenziale stocastica

$$dX_t = \frac{1}{X_t} dt + X_t dB_t.$$

con condizione iniziale $X_0 = C$, e se ne determini la soluzione.

Soluzione. L'equazione differenziale stocastica può essere risolta considerando il metodo del fattore d'integrazione. In questo caso si ha

$$Y_t = e^{\frac{1}{2}t - B_t},$$

mentre

$$d(X_t Y_t) = \frac{Y_t}{X_t} dt,$$

ovvero

$$X_t Y_t d(X_t Y_t) = Y_t^2 dt.$$

Tenendo presente che $X_0 Y_0 = C$, la precedente equazione differenziale ha come soluzione

$$\frac{1}{2} X_t^2 Y_t^2 = \frac{1}{2} C^2 + \int_0^t Y_s^2 ds,$$

ovvero l'equazione differenziale stocastica ha soluzione $X = (X_t)_{t \in \mathbb{R}^+}$ con

$$X_t = e^{B_t - \frac{1}{2}t} \left(C^2 + 2 \int_0^t e^{s - 2B_s} ds \right)^{\frac{1}{2}}. \quad \square$$

• **Esercizio 10.4.6.** Si consideri l'equazione differenziale stocastica

$$dX_t = \left(\sqrt{1 + X_t^2} + \frac{X_t}{2} \right) dt + \sqrt{1 + X_t^2} dB_t.$$

con condizione iniziale $X_0 = C$, e se ne determini la soluzione.

Soluzione. Si può verificare che in modo simile all'Esercizio 10.4.4 le condizioni per l'esistenza e l'unicità dell'equazione differenziale stocastica sono soddisfatte. Dunque, esiste una soluzione forte. L'equazione differenziale stocastica è esatta e può essere risolta considerando le relazioni

$$\sqrt{1 + g(x, y)^2} + \frac{g(x, y)}{2} = \frac{\partial g(x, y)}{\partial y} + \frac{1}{2} \frac{\partial^2 g(x, y)}{\partial x^2}$$

e

$$\sqrt{1 + g(x, y)^2} = \frac{\partial g(x, y)}{\partial x},$$

ovvero risolvendo l'equazione differenziale

$$g(x, y) = -2 \frac{\partial g(x, y)}{\partial x} + 2 \frac{\partial g(x, y)}{\partial y} + \frac{1}{2} \frac{\partial^2 g(x, y)}{\partial x^2},$$

la cui soluzione è data da

$$g(x, y) = \sinh(x + y + c),$$

con $c \in \mathbb{R}$. Dunque, si deve concludere che l'equazione differenziale stocastica ha soluzione $X = (X_t)_{t \in \mathbb{R}^+}$ con

$$X_t = \sinh(B_t + t + \operatorname{arcsinh}(C)) .$$

In effetti, applicando la formula ridotta di Itô e tenendo presente che $\cosh(x) = \sqrt{1 + \sinh^2(x)}$, si ha

$$\begin{aligned} d\sinh(B_t + t) &= \cosh(B_t + t) dB_t + \cosh(B_t + t) dt + \frac{1}{2} \sinh(B_t + t) dt && \square \\ &= \left(\sqrt{1 + \sinh^2(B_t + t)} + \frac{\sinh(B_t + t)}{2} \right) dt + \sqrt{1 + \sinh^2(B_t + t)} dB_t . \end{aligned}$$

Pagina intenzionalmente vuota

Appendice A

Elementi di integrazione

Sia (Ω, \mathcal{F}) uno spazio misurabile. Un'applicazione $g : \Omega \rightarrow \mathbb{R}$ è *misurabile* se

$$g^{-1}(B) = \{\omega \in \Omega : g(\omega) \in B\} \in \mathcal{F}$$

per ogni $B \in \mathcal{B}(\mathbb{R})$. In questo caso, g è anche detta funzione Boreliana su (Ω, \mathcal{F}) . Per le funzioni misurabili valgono i seguenti risultati.

Teorema A.1. *Se (Ω, \mathcal{F}) è uno spazio misurabile, allora:*

- i) g è una funzione misurabile se e solo se $g^{-1}(]-\infty, x]) \in \mathcal{F}$ per ogni $x \in \mathbb{R}$;
- ii) se g e h sono funzioni misurabili, la funzione prodotto gh è misurabile (assumendo che sia ben definita in base alla convenzione $0 \cdot \infty = 0$ e $\infty \cdot 0 = 0$);
- iii) se g e h sono funzioni misurabili e se $a, b \in \mathbb{R}$, la combinazione lineare $ag + bh$ è una funzione misurabile (assumendo che sia ben definita);
- iv) se $(g_n)_{n \geq 1}$ è una successione di funzioni misurabili, allora $\sup_n g_n$, $\inf_n g_n$, $\limsup_n g_n$ e $\liminf_n g_n$ sono funzioni misurabili.

Dimostrazione. Vedi Billingsley (1995). □

Risulta immediato dimostrare che la funzione indicatrice $\mathbf{1}_A$ dell'insieme A è una funzione misurabile se $A \in \mathcal{F}$. Una *funzione semplice* è una combinazione lineare di funzioni indicatrici, ovvero

$$\varphi = \sum_{k=1}^n a_k \mathbf{1}_{A_k},$$

dove $A_1, \dots, A_n \in \mathcal{F}$, mentre $a_1, \dots, a_n \in \mathbb{R}$. Dal Teorema A.1 si ha dunque che una funzione semplice è misurabile.

Teorema A.2. *Se (Ω, \mathcal{F}) è uno spazio misurabile e g è una funzione misurabile, allora esiste una successione di funzioni semplici $(\varphi_n)_{n \geq 1}$ tali che $0 \leq \varphi_1 \leq \varphi_2 \leq \dots \leq g$ e $\lim_n \varphi_n = g$.*

Dimostrazione. Vedi Billingsley (1995). □

Una successione di funzioni semplici come quella considerata nel precedente Teorema può essere sempre costruita scegliendo φ_n come segue

$$\varphi_n = \sum_{k=1}^{n2^n} \frac{k-1}{2^n} \mathbf{1}_{[(k-1)2^{-n}, k2^{-n}[}(g) + n \mathbf{1}_{[n, \infty[}(g).$$

La definizione di integrale di una funzione misurabile rispetto ad una misura procede per vari passi, partendo dalla definizione di integrale di una funzione semplice non negativa.

Definizione A.3. Se $(\Omega, \mathcal{F}, \mu)$ è uno spazio misurato e φ è una funzione semplice non negativa, ovvero $a_1, \dots, a_n \geq 0$, l'integrale di φ è dato da

$$\int \varphi d\mu = \sum_{k=1}^n a_k \mu(A_k),$$

con la convenzione che $0 \cdot \infty = 0$. □

Si noti che l'integrale è sempre ben definito, anche se può risultare $\int \varphi d\mu = \infty$. Inoltre, anche se φ non ha una rappresentazione univoca, nel senso che differenti scelte di A_1, \dots, A_n e a_1, \dots, a_n possono dare luogo ad una stessa funzione semplice φ , tuttavia è possibile dimostrare che tutte le differenti rappresentazioni producono lo stesso valore per $\int \varphi d\mu$ (si veda Billingsley, 1995). Tenendo presente il Teorema A.2, è quindi immediato dare la definizione di integrale per una funzione misurabile non negativa.

Definizione A.4. Sia $(\Omega, \mathcal{F}, \mu)$ uno spazio misurato e g una funzione misurabile non negativa. Se $(\varphi_n)_{n \geq 1}$ è una successione di funzioni semplici non negative tali che $0 \leq \varphi_1 \leq \varphi_2 \leq \dots \leq g$, l'integrale di g è dato da

$$\int g d\mu = \sup_n \int \varphi_n d\mu. \quad \square$$

Si noti che l'integrale non dipende dalla scelta di una particolare successione di funzioni elementari. Si definisca inoltre come parte positiva di g la funzione $g^+ = \max(g, 0)$ e come parte negativa di g la funzione $g^- = -\min(g, 0)$. Le funzioni g^+ e g^- sono misurabili e non negative, mentre risulta $g = g^+ - g^-$ e $|g| = g^+ + g^-$. In questo caso, si ha la seguente definizione generale di integrale per una funzione misurabile.

Definizione A.5. Sia $(\Omega, \mathcal{F}, \mu)$ uno spazio misurato e g una funzione misurabile. L'integrale $\int g d\mu$ esiste se e solo se almeno uno degli integrali $\int g^+ d\mu$ o $\int g^- d\mu$ è finito e in questo caso si pone

$$\int g d\mu = \int g^+ d\mu - \int g^- d\mu.$$

Se entrambi gli integrali $\int g^+ d\mu$ e $\int g^- d\mu$ sono finiti, la funzione g è detta integrabile. Se entrambi gli integrali non sono finiti (ovvero, si ha la forma del tipo $+\infty - \infty$), allora l'integrale non è definito. □

L'insieme di tutte le funzioni integrabili rispetto a μ viene indicato con $L^1(\mu)$. In generale, l'insieme delle funzioni per cui $\int |g|^p d\mu < \infty$, dove $p > 0$, viene indicato con $L^p(\mu)$. Inoltre, l'integrale

$$\|g\|_p = \int |g|^p d\mu$$

è detta norma di ordine p di g , mentre l'integrale

$$\langle g, h \rangle = \int gh d\mu$$

è detto prodotto interno. Inoltre, se $A \in \mathcal{F}$ allora l'integrale di $g \in L^1(\mu)$ su A è definito come

$$\int_A g d\mu = \int \mathbf{1}_A g d\mu.$$

Per enfatizzare l'argomento di g , per il precedente integrale si adotta talvolta anche la notazione

$$\int_A g d\mu = \int_A g(\omega) d\mu(\omega).$$

Nel caso particolare in cui $\Omega = \mathbb{R}$ e $\mathcal{F} = \mathcal{B}(\mathbb{R})$, mentre $\mu = \lambda$ è la misura di Lebesgue, allora l'integrale di g su A (detto integrale di Lebesgue su A) viene scritto semplicemente come

$$\int_A g d\mu = \int_A g(x) dx.$$

Qualora $A = [a, b]$ (dove gli estremi dell'intervallo sono eventualmente non finiti), allora l'integrale di Lebesgue si scrive

$$\int_A g d\mu = \int_a^b g(x) dx.$$

La precedente scrittura può apparire ambigua, dal momento che non distingue l'intervallo $[a, b]$ dall'intervallo $]a, b[$. Tuttavia, poichè $\lambda(\{a\}) = \lambda(\{b\}) = 0$, allora gli integrali estesi ad uno qualsiasi degli intervalli $]a, b[$, $]a, b]$, $[a, b[$ e $[a, b]$ coincidono. Infine, si può dimostrare che l'integrale $\int_a^b g(x) dx$ coincide con quello di Riemann, quando questo è definito (vedi Billingsley, 1995). Tuttavia, vi sono funzioni per cui l'integrale di Lebesgue è definito, mentre l'integrale di Riemann non è definito.

In modo simile, quando $\Omega = \mathbb{R}^n$ e $\mathcal{F} = \mathcal{B}(\mathbb{R}^n)$, mentre $\mu = \lambda^n$ è la misura di Lebesgue in \mathbb{R}^n , allora l'integrale di Lebesgue di g su A (detto integrale di Lebesgue su A) viene scritto come

$$\int_A g d\mu = \int_A g(x_1, \dots, x_n) dx_1 \dots dx_n.$$

Qualora $A = [a_1, b_1] \times \dots \times [a_n, b_n]$, allora l'integrale di Lebesgue si scrive

$$\int_A g d\mu = \int_{a_1}^{b_1} \dots \int_{a_n}^{b_n} g(x_1, \dots, x_n) dx_1 \dots dx_n.$$

Di nuovo, si può dimostrare che questo integrale coincide con quello di Riemann, quando questo è definito. Di seguito vengono considerate alcune proprietà dell'integrale.

Teorema A.6. *Sia $(\Omega, \mathcal{F}, \mu)$ uno spazio misurato e $g, h \in L^1(\mu)$:*

i) se $a, b \in \mathbb{R}$, allora

$$\int (ag + bh) d\mu = a \int g d\mu + b \int h d\mu;$$

ii) se $g \leq h$ q.o., allora

$$\int g d\mu \leq \int h d\mu;$$

iii) se $g \geq 0$ q.o. e $\int g d\mu = 0$, allora $g = 0$ q.o.

Dimostrazione. Vedi Billingsley (1995). □

Dal Teorema A.6 segue che due funzioni $g, h \in L^1(\mu)$ equivalenti q.o. rispetto a μ , ovvero $g \neq h$ solo per un insieme di misura nulla, hanno lo stesso integrale. Inoltre, dal Teorema A.6 è immediato ottenere la disuguaglianza

$$\left| \int g d\mu \right| \leq \int |g| d\mu,$$

per cui si ha che una funzione misurabile g è integrabile se e solo se $|g|$ è integrabile. Inoltre, una funzione $g \in L^1(\mu)$ è finita q.o. Il prossimo Teorema fornisce le condizioni per cui le operazioni di limite e di integrazione possono essere scambiate. La prima parte del Teorema è legata al nome del matematico francese Pierre Joseph Louis Fatou (1878-1929), mentre la seconda parte del Teorema è legata al nome del matematico italiano Beppo Levi (1875-1961).

Teorema A.7. Sia $(\Omega, \mathcal{F}, \mu)$ uno spazio misurato e $(g_n)_{n \geq 1}$ una successione di funzioni di $L^1(\mu)$:
i) (Lemma di Fatou) se $g_n \geq 0$ per ogni n , allora

$$\int \liminf_n g_n d\mu = \liminf_n \int g_n d\mu;$$

ii) (Teorema della convergenza dominata di Beppo Levi) Se $\lim_n g_n = g$ q.o. ed esiste una funzione $h \in L^1(\mu)$ tale che $|g_n| \leq h$ q.o., allora $g \in L^1(\mu)$ e

$$\int g d\mu = \int \lim_n g_n d\mu = \lim_n \int g_n d\mu;$$

Dimostrazione. Vedi Billingsley (1995). □

Sia $(\Omega, \mathcal{F}, \mu)$ uno spazio misurato e g una funzione misurabile da (Ω, \mathcal{F}) allo spazio misurabile (Λ, \mathcal{G}) . In questo caso, la misura immagine η di μ mediante g è definita come

$$\eta = \mu(g^{-1}(A))$$

per ogni $A \in \mathcal{G}$. La misura immagine η è anche detta misura indotta da g . Il seguente Teorema fornisce un'espressione per l'integrale rispetto alla misura immagine.

Teorema A.8. Sia $(\Omega, \mathcal{F}, \mu)$ uno spazio misurato e sia η la misura immagine di μ mediante g . Se h è una funzione Boreliana e la funzione $h \circ g = h(g)$ è integrabile, allora

$$\int h d\eta = \int (h \circ g) d\mu.$$

Dimostrazione. Vedi Billingsley (1995). □

Se $\Omega = \mathbb{R}^n$ e $\mu = \lambda^n$ è la misura di Lebesgue in \mathbb{R}^n , si supponga che $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$. Sia inoltre g un diffeomorfismo, ovvero un'applicazione biunivoca di un insieme aperto \mathbb{R}^n su un insieme aperto di \mathbb{R}^n differenziabile con continuità e tale che la sua inversa sia differenziabile con continuità. In questo caso, il Teorema A.8 si riduce a

$$\int_B h(y_1, \dots, y_n) dy_1 \dots dy_n = \int_{g^{-1}(B)} h(g(x_1, \dots, x_n)) |J(g(x_1, \dots, x_n))| dx_1 \dots dx_n,$$

dove $J(g(x_1, \dots, x_n))$ è lo jacobiano relativo a g .

Sia $(\Omega, \mathcal{F}, \nu)$ uno spazio misurato e f una funzione Boreliana non negativa su (Ω, \mathcal{F}) . In questo caso, l'applicazione $\eta : \mathcal{F} \rightarrow \mathbb{R}$, data da

$$\eta(A) = \int_A f d\nu$$

per ogni $A \in \mathcal{F}$, è una misura. In questo contesto, la funzione Boreliana f viene detta *densità*, mentre η è detta *misura definita* dalla densità f .

Teorema A.9. (Teorema di Radon-Nikodym) *Siano η e ν due misure definite sullo spazio misurabile (Ω, \mathcal{F}) . Se η è assolutamente continua rispetto a ν , ovvero $\nu(A) = 0$ implica che $\eta(A) = 0$, allora esiste una funzione Boreliana non negativa f tale che η è la misura definita dalla densità f . Inoltre, f è unica q.o. rispetto a ν , ovvero se h è una funzione Boreliana non negativa tale che $\eta(A) = \int_A h d\nu$ per ogni $A \in \mathcal{F}$, allora $f = h$ q.o.*

Dimostrazione. Vedi Billingsley (1995). □

In pratica, il Teorema A.9, che prende nome dal matematico austriaco Johann Karl August Radon (1887-1956) e dal matematico polacco Otto Marcin Nikodym (1887-1974), fornisce l'esistenza di una classe di funzioni non negative f che coincidono q.o. rispetto a ν e per questo motivo f viene definita come una versione della densità. Si noti che il Teorema di Radon-Nikodym fornisce sia un modo di costruire una misura, ma anche un metodo per calcolare la misura di ogni $A \in \mathcal{F}$. In effetti, se ν è una misura di cui sono ben note le proprietà (quale la misura di Lebesgue) allora, una volta determinata la densità f , il valore della misura può essere ottenuto per integrazione. In generale, ogni integrazione rispetto a η può essere ricondotta ad un'integrazione rispetto a ν , come si ottiene dal seguente Teorema.

Teorema A.10. *Sia $(\Omega, \mathcal{F}, \nu)$ uno spazio misurato e sia η la misura definita dalla densità f . Se h è una funzione Boreliana e la funzione hf è integrabile, allora*

$$\int h d\eta = \int hf d\nu.$$

Dimostrazione. Vedi Billingsley (1995). □

Il seguente Teorema evidenzia come un calcolare un integrale rispetto ad una misura prodotto per mezzo di una integrazione iterata.

Teorema A.11. (Teorema di Fubini) *Siano $(\Omega_1, \mathcal{F}_1, \mu_1)$ e $(\Omega_2, \mathcal{F}_2, \mu_2)$ due spazi misurati e sia g una funzione misurabile su $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2)$. Se $g \geq 0$ allora la funzione*

$$h(\omega_2) = \int g(\omega_1, \omega_2) d\mu_1(\omega_1)$$

esiste q.o. rispetto a μ_2 ed è misurabile su Ω_2 , mentre

$$\int g d(\mu_1 \otimes \mu_2) = \int \left(\int g(\omega_1, \omega_2) d\mu_1(\omega_1) \right) d\mu_2(\omega_2).$$

Dimostrazione. Vedi Billingsley (1995). □

Il Teorema può essere esteso a n spazi misurati ed è legato al nome del matematico italiano Guido Fubini (1879-1943), anche se più correttamente dovrebbe essere denominato Teorema di Fubini-Tonelli, dal momento che è stato esteso e perfezionato dal matematico italiano Leonida Tonelli (1885-1946).

Riferimenti bibliografici

Il classico testo di riferimento per la Teoria della Misura è Halmos (1974). Ulteriori testi sulla Teoria della Misura e integrazione sono Ambrosio *et al.* (2011), Cohn (2013), Rao (2004), Schilling (2005), Taylor (2006), Wheeden e Zygmund (2015). Un testo esaustivo sull'argomento è Bogachev (2007, volume I e II), mentre testi introduttivi sono Capiński e Kopp (2007), Kubrusly (2015) e Ovchinnikov (2013).

Appendice B

Notazioni e abbreviazioni

Capitolo 1

Ω	spazio fondamentale
\emptyset	evento impossibile
$\{\omega\}$	evento elementare
E	evento
$\text{card}(E)$	cardinalità di un evento
\mathcal{E}	classe di eventi
$\mathcal{P}(\Omega)$	insieme delle parti
$(E_i)_{i \in I}$	classe di eventi I insieme di indici
$(E_i)_{i=1}^n$	classe di eventi $I = \{1, \dots, n\}$
$(E_n)_{n \geq 1}$	successione di eventi $I = \{1, 2, \dots\}$
$(E_n)_{n \geq 0}$	successione di eventi $I = \{0, 1, \dots\}$
E^c	evento opposto
$\bigcup_{i \in I} E_i$	evento unione
$E_1 \cup E_2$	evento unione $I = \{1, 2\}$
$\bigcup_{i=1}^n E_i$	evento unione $I = \{1, \dots, n\}$
$\bigcup_{n=1}^{\infty} E_n$	evento unione $I = \{1, 2, \dots\}$
$\bigcup_{n=0}^{\infty} E_n$	evento unione $I = \{0, 1, \dots\}$
$\bigcap_{i \in I} E_i$	evento intersezione
$E_1 \cap E_2$	evento intersezione $I = \{1, 2\}$
$\bigcap_{i=1}^n E_i$	evento intersezione $I = \{1, \dots, n\}$
$\bigcap_{n=1}^{\infty} E_n$	evento intersezione $I = \{1, 2, \dots\}$
$\bigcap_{n=0}^{\infty} E_n$	evento intersezione $I = \{0, 1, \dots\}$
$E_1 \setminus E_2$	evento differenza
$\liminf_n E_n$	limite inferiore di successione di eventi
$\limsup_n E_n$	limite superiore di successione di eventi
$\lim_n E_n$	limite di successione di eventi
$(\mathcal{E}_i)_{i \in I}$	classe di classi di eventi
$(\mathcal{E}_n)_{n \geq 1}$	successione di classi di eventi $I = \{1, 2, \dots\}$
$(\mathcal{E}_n)_{n \geq 0}$	successione di classi di eventi $I = \{0, 1, \dots\}$
$\bigcup_{i \in I} \mathcal{E}_i$	unione di classi di eventi
$\bigcup_{n=1}^{\infty} \mathcal{E}_n$	unione di classi di eventi $I = \{1, 2, \dots\}$
$\bigcup_{n=0}^{\infty} \mathcal{E}_n$	unione di classi di eventi $I = \{0, 1, \dots\}$
$\bigcap_{i \in I} \mathcal{E}_i$	intersezione di classi di eventi
$\bigcap_{n=1}^{\infty} \mathcal{E}_n$	intersezione di classi di eventi $I = \{1, 2, \dots\}$
$\bigcap_{n=0}^{\infty} \mathcal{E}_n$	intersezione di classi di eventi $I = \{0, 1, \dots\}$
\mathcal{F}	σ -algebra
$\sigma(\mathcal{E})$	σ -algebra generata dalla classe \mathcal{E}
(Ω, \mathcal{F})	spazio probabilizzabile

$\mathbb{N} = \{0, 1, 2, \dots\}$	insieme degli interi non negativi
$\mathbb{Z} = \{0, \pm 1, \pm 2, \dots\}$	insieme dei numeri interi
$\mathbb{Z}^+ = \{1, 2, \dots\}$	insieme degli interi positivi
$\mathbb{Q} = \{m/n : m, n \in \mathbb{Z}, n \neq 0\}$	insieme dei numeri razionali
\mathbb{R}	insieme dei numeri reali
\mathbb{R}^+	insieme dei numeri reali positivi
\mathbb{R}^-	insieme dei numeri reali negativi
\mathbb{C}	insieme dei numeri complessi
$[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}$	intervallo chiuso
$]a, b[= \{x \in \mathbb{R} : a < x < b\}$	intervallo aperto
$]a, b] = \{x \in \mathbb{R} : a < x \leq b\}$	intervallo aperto a sinistra e chiuso a destra
$[a, b[= \{x \in \mathbb{R} : a \leq x < b\}$	intervallo chiuso a sinistra e aperto a destra
$\mathcal{B}(\mathbb{R})$	σ -algebra di Borel
$(\mathbb{R}, \mathcal{B}(\mathbb{R}))$	spazio misurabile
$\Omega_1 \times \dots \times \Omega_k$	spazio fondamentale prodotto di k spazi
$E_1 \times \dots \times E_k$	evento rettangolare di k eventi
$\mathcal{F}_1 \otimes \dots \otimes \mathcal{F}_k$	σ -algebra prodotto di k spazi
\mathbb{R}^k	spazio Euclideo
$\mathcal{B}(\mathbb{R}^k)$	σ -algebra di Borel su \mathbb{R}^k
$\prod_{i \in I} \Omega_i$	spazio fondamentale prodotto I insieme indici
$\prod_{i \in I} E_i$	evento rettangolare
$\otimes_{i \in I} \mathcal{F}_i$	σ -algebra prodotto

Capitolo 2

$\mu(\cdot)$	misura
$(\Omega, \mathcal{F}, \mu)$	spazio misurato
$P(\cdot)$	misura di probabilità
(Ω, \mathcal{F}, P)	spazio probabilizzato
$\mathbf{1}_A(\cdot)$	funzione indicatrice dell'insieme A
$\lambda(\cdot)$	misura di Lebesgue su \mathbb{R}
$(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$	spazio misurato
$P(\cdot E_0)$	probabilità condizionata
$P_1 \otimes \dots \otimes P_k$	probabilità prodotto
$\lambda^k(\cdot)$	misura di Lebesgue su \mathbb{R}^k
$(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k), \lambda^k)$	spazio misurato prodotto
$\otimes_{i \in I} P_i$	probabilità prodotto

Capitolo 3

X	variabile aleatoria (v.a.)
$P_X(\cdot)$	legge (o distribuzione) della v.a. X
$(\mathbb{R}, \mathcal{B}(\mathbb{R}), P_X)$	spazio probabilizzato
$\stackrel{\mathcal{L}}{=}$	uguaglianza in legge
$F_X(\cdot)$	funzione di ripartizione della v.a. X (f.r.)
$p_X(\cdot)$	funzione di probabilità della v.a. X (f.p.)
$[\cdot]$	funzione parte intera
$f_X(\cdot)$	densità di probabilità della v.a. X (d.p.)
$(X_1, \dots, X_k)^\top$	vettore di variabili aleatorie (v.v.a.)
X_i	i -esima componente marginale del v.v.a.

$F_X(\cdot)$	funzione di ripartizione del v.v.a. X (f.r.c.)
$F_{X_i}(\cdot)$	funzione di ripartizione della v.a. X_i (f.r.m.)
$p_X(\cdot)$	funzione di probabilità del v.v.a. X (f.p.c.)
$p_{X_i}(\cdot)$	funzione di probabilità della v.a. X_i (f.p.m.)
$f_X(\cdot)$	densità di probabilità del v.v.a. X (d.p.c.)
$f_{X_i}(\cdot)$	densità di probabilità della v.a. X_i (d.p.m.)

Capitolo 4

$E[X]$	valore atteso della v.a. X
μ_X	valore atteso della v.a. X
X^+	$\max(X, 0)$ parte positiva della v.a. X
X^-	$-\min(X, 0)$ parte negativa della v.a. X
$\mu_{X,r}$	momento di ordine r della v.a. X
$\text{Var}[X]$	varianza della v.a. X
σ_X^2	varianza della v.a. X
$\text{Cov}[X_1, X_2]$	covarianza tra le v.a. X_1 e X_2
σ_{X_1, X_2}	covarianza tra le v.a. X_1 e X_2
ρ_{X_1, X_2}	coefficiente di correlazione tra le v.a. X_1 e X_2
$\text{sgn}(\cdot)$	$\mathbf{1}_{]0, \infty[}(\cdot) - \mathbf{1}_{]-\infty, 0[}(\cdot)$ funzione segno
Σ_X	matrice di varianza-covarianza
$\det(\Sigma_X)$	varianza generalizzata

Capitolo 5

$E[X E_0]$	valore atteso condizionato all'evento E_0
$E[X \mathcal{F}_0]$	valore atteso condizionato alla σ -algebra \mathcal{F}_0
\mathcal{F}_X	σ -algebra indotta dalla v.a. X
$E[X_2 X_1]$	valore atteso condizionato di X_2 a X_1
$P_{X_2 X_1=x_1}(\cdot)$	legge condizionata di X_2 all'evento $\{X_1 = x_1\}$
$p_{X_2 X_1=x_1}$	f.p. condizionata di X_2 all'evento $\{X_1 = x_1\}$
$f_{X_2 X_1=x_1}$	d.p. condizionata di X_2 all'evento $\{X_1 = x_1\}$

Capitolo 6

$\mathcal{B}(n, p)$	legge Binomiale di parametri n e p
$\mathcal{P}(\lambda)$	legge di Poisson di parametro λ
$\mathcal{BN}(k, p)$	legge Binomiale Negativa di parametri k e p
$\mathcal{I}(n, D, N)$	legge Ipergeometrica di parametri n , D e N
$\mathcal{M}(n, p)$	legge Multinomiale di parametri n e p
$\mathcal{N}(\mu, \sigma^2)$	legge Normale di parametri μ e σ
$\phi(\cdot)$	d.p. della v.a. Z con legge Normale ridotta
$\Phi(\cdot)$	f.r. della v.a. Z con legge Normale ridotta
$\mathcal{G}(a, b, k)$	legge Gamma di parametri a , b e k
$\Gamma(k)$	funzione Gamma di Eulero
$\gamma(k, z)$	funzione Gamma incompleta
χ_n^2	legge Chi-quadrato con n gradi di libertà
$\mathcal{BE}(a, b, \alpha, \beta)$	legge Beta di parametri a , b , α e β
$B(\alpha, \beta)$	funzione Beta di Eulero
$I_z(\alpha, \beta)$	funzione Beta incompleta regolarizzata

F_{n_1, n_2}	legge di Snedecor con n_1 e n_2 gradi di libertà
t_n	legge t di Student con n gradi di libertà
$\mathcal{N}_k(\mu_X, \Sigma_X)$	legge Normale Multivariata
$\mathcal{D}(\alpha)$	legge di Dirichlet di parametro α

Capitolo 7

i	$\sqrt{-1}$ unità immaginaria
$\Re(\cdot)$	parte reale
$\Im(\cdot)$	parte immaginaria
$ \cdot $	modulo
$\varphi_X(\cdot)$	funzione caratteristica della v.a. X (f.c.)
$G_X(\cdot)$	funzione generatrice della v.a. X (f.g.)
$\zeta(s)$	funzione zeta di Riemann
$\varphi_X(\cdot)$	funzione caratteristica multivariata (f.c.m.)
$G_X(\cdot)$	funzione generatrice multivariata (f.g.m.)

Capitolo 8

$(X_n)_{n \geq 1}$	successione di v.a.
$(X_n)_{n \geq 0}$	successione di v.a.
$\xrightarrow{\mathcal{L}}$	convergenza in legge
\xrightarrow{P}	convergenza in probabilità
$\xrightarrow{q.c.}$	convergenza quasi certa
$\xrightarrow{L^p}$	convergenza in media di ordine p

Capitolo 9

$(X_n)_{t \in \mathbb{T}}$	processo aleatorio (p.a.)
$(\mathcal{F}_t)_{t \in \mathbb{T}}$	filtrazione
$(M_t)_{t \in \mathbb{T}}$	martingala
τ	tempo di arresto
\mathcal{F}_τ	σ -algebra associata al tempo di arresto τ
$(M_{\tau \wedge t})_{t \in \mathbb{T}}$	martingala arrestata al tempo di arresto τ
$a \vee b$	$\max(a, b)$
$a \wedge b$	$\min(a, b)$
$(B_t)_{t \in [0, \infty[}$	moto Browniano
τ_a	tempo di primo passaggio per a

Capitolo 10

$I(X)$	integrale stocastico di Itô del p.a. X
\mathcal{S}_T^2	classe dei p.a. semplici
\mathcal{M}_T^2	classe dei p.a. integrabili
dX_t	differenziale stocastico del p.a. X

Bibliografia

- Ambrosio, L., Da Prato, G. e Mennucci, A. (2011) *Introduction to Measure Theory and Integration*, Edizioni della Normale, Scuola Normale Superiore, Pisa.
- Applebaum, D. (2009) *Lévy Processes and Stochastic Calculus*, Cambridge University Press, Cambridge.
- Arguin, L.P. (2022) *A First Course in Stochastic Calculus*, American Mathematical Society, Providence.
- Ash, R.B. (2008) *Basic Probability Theory*, Dover, New York.
- Ash, R.B. e Doléans-Dade, C.A. (2000) *Probability and Measure Theory*, Academic Press, New York.
- Athreya, K.B. e Lahiri, S.N. (2006) *Measure Theory and Probability Theory*, Springer, New York.
- Baldi, P. (2017) *Stochastic Calculus*, Springer, New York.
- Bass, R.F. (2011) *Stochastic Processes*, Cambridge University Press, Cambridge.
- Bhattacharya, R.N. e Waymire, E.C. (2016) *Stochastic Processes with Applications*, seconda edizione, Society for Industrial and Applied Mathematics, Philadelphia.
- Bhattacharya, R.N. e Waymire, E.C. (2021) *Random Walk, Brownian Motion, and Martingales*, Springer Nature, Switzerland.
- Billingsley, P. (1995) *Probability and Measure*, terza edizione, Wiley, New York.
- Blom, G., Holst, L. e Sandell, D. (1994) *Problems and Snapshots from the World of Probability*, Springer, New York.
- Bogachev, V.I. (2007) *Measure Theory*, volume I, Springer, Berlin.
- Bogachev, V.I. (2007) *Measure Theory*, volume II, Springer, Berlin.
- Borodin, A.N. (2017) *Stochastic Processes*, Birkhäuser, Basel.
- Borovkov, A.A. (2013) *Probability Theory*, quinta edizione, Springer, London.
- Boucheron, S., Lugosi, G. e Massart, P. (2013) *Concentration Inequalities*, Oxford University Press, Oxford.
- Brémaud, P. (2020) *Probability Theory and Stochastic Processes*, Springer, New York.
- Brzeźniak, Z. e Zastawniak, T. (2002) *Basic Stochastic Processes*, Springer, London.
- Calin, O. (2015) *An Informal Introduction to Stochastic Calculus with Applications*, World Scientific Publishing, Singapore.
- Capasso, V. e Bakstein, D. (2021) *An Introduction to Continuous-Time Stochastic Processes*, quarta edizione, Birkhäuser, New York.
- Capiński, M. e Kopp, P. (2007) *Measure, Integral and Probability*, seconda edizione, Springer, New York.
- Charalambides, C.A. (2005) *Combinatorial Methods in Discrete Distribution*, Wiley, New York.
- Chaumont, L. e Yor, M. (2012) *Exercises in Probability*, seconda edizione, Cambridge University Press, Cambridge.
- Choe, G.H. (2016) *Stochastic Analysis for Finance with Simulations*, Springer, New York.
- Chung, K.L. (2001) *A Course in Probability Theory*, terza edizione, Academic Press, San Diego.
- Chung, K. e Williams, R. (1990) *Introduction to Stochastic Integration*, seconda edizione, Birkhäuser, Boston.
- Çinlar, E. (2010) *Probability and Stochastics*, Springer, New York.
- Cohen, S.N. e Elliott, R.J. (2015) *Stochastic Calculus and Applications*, seconda edizione, Birkhäuser, New York.

- Cohn, D.L. (2013) *Measure Theory*, seconda edizione, Birkhäuser, New York.
- Da Prato, G. (2014) *Introduction to Stochastic Analysis and Malliavin Calculus*, terza edizione, Edizioni della Normale, Scuola Normale Superiore, Pisa.
- Devroye, L. (1986) *Non-uniform Random Variate Generation*, Springer, New York.
- Dorogovtsev, A.Y., Silvestrov, D.S., Skorokhod, A.V. and Yadrenko, M.I. (1997) *Probability Theory: Collection of Problems*, Translations of Mathematical Monographs **163**, American Mathematical Society, Providence.
- Dubins, L.E. e Savage, L.J. (1965) *How to Gamble If You Must*, McGraw-Hill, New York.
- Dudley, R.M. (2004) *Real Analysis and Probability*, Cambridge University Press, Cambridge.
- Durrett, R. (2012) *Essentials of Stochastic Processes*, seconda edizione, Springer, New York.
- Erickson, M. (2010) *Pearls of Discrete Mathematics*, CRC Press, Boca Raton.
- Ethier, S.N. (2010) *The Doctrine of Chances*, Springer, Berlin.
- Feller, W. (1968) *An Introduction to Probability Theory and its Applications*, volume I, terza edizione, Wiley, New York.
- Feller, W. (1971) *An Introduction to Probability Theory and its Applications*, volume II, seconda edizione, Wiley, New York.
- Ferguson, T.S. (1996) *A Course in Large-sample Theory*, Chapman & Hall, New York.
- de Finetti, B. (1970) *Theory of Probability*, volume 1, Wiley, New York.
- de Finetti, B. (1970) *Theory of Probability*, volume 2, Wiley, New York.
- Foata, D. e Fuchs, A. (1998) *Calcul des Probabilités*, seconda edizione, Dunod, Paris.
- Forbes, C., Evans, M., Hastings, N. e Peacock, B. (2011) *Statistical Distributions*, quarta edizione, Wiley, New York.
- Gordon, H. (1997) *Discrete Probability*, Springer, New York.
- Gorroochurn, P. (2012) *Classic Problems of Probability*, Wiley, New York.
- Graham, R.L., Knuth, D.E. e Patashnik, O. (1994) *Concrete Mathematics*, seconda edizione, Addison-Wesley, Reading.
- Grimmett, G.R. e Stirzaker, D.R. (2001) *One Thousand Exercises in Probability*, Oxford University Press, Oxford.
- Gut, A. (2005) *Probability: a Graduate Course*, seconda edizione, Springer, New York.
- Hald, A. (1998) *A History of Mathematical Statistics from 1750 to 1930*, Wiley New York.
- Hald, A. (2003) *History of Probability and Statistics and Their Applications before 1750*, Wiley, New York.
- Halmos, P.R. (1974) *Measure Theory*, Springer, New York.
- Hamming, R.W. (1991) *The Art of Probability*, Addison-Wesley, Reading.
- Isaac, R. (1995) *The Pleasures of Probability*, Springer, New York.
- Jacod, J. e Protter, P. (2004) *Probability Essentials*, seconda edizione, Springer, New York.
- Johnson, N., Kemp, A. e Kotz, S. (2005) *Univariate Discrete Distributions*, terza edizione, New York, Wiley.
- Johnson, N. e Kotz, S. (1977) *Urn Models and Their Applications*, New York, Wiley.
- Johnson, N., Kotz, S. e Balakrishnan, N. (1994) *Continuous Univariate Distributions*, volume 1, seconda edizione, New York, Wiley.
- Johnson, N., Kotz, S. e Balakrishnan, N. (1995) *Continuous Univariate Distributions*, volume 2, seconda edizione, New York, Wiley.
- Johnson, N., Kotz, S. e Balakrishnan, N. (1997) *Discrete Multivariate Distributions*, New York, Wiley.
- Kallenberg, O. (2021) *Foundations of Modern Probability*, terza edizione, Springer Nature, Switzerland.
- Karatzas, I. e Shreve, S.E. (1991) *Brownian Motion and Stochastic Calculus*, seconda edizione, Springer, New York.
- Khoshnevisan, D. (2007) *Probability*, American Mathematical Society, Providence.
- Klafter, J. e Sokolov, I.M. (2011) *First Steps in Random Walks*, Oxford University Press, Oxford.

- Klebaner, F. (2005) *Introduction to Stochastic Calculus with Applications*, seconda edizione, World Scientific Publishing, Singapore.
- Klenke, A. (2020) *Probability Theory*, Springer Nature, Switzerland.
- Knill, O. (2009) *Probability and Stochastic Processes with Applications*, Overseas Press, New Dehli.
- Kolmogorov, A.N. (1933) *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Springer, Berlin.
Traduzione inglese: Kolmogorov, A.N. (1956) *Foundations of the Theory of Probability*, Chelsea, New York.
- Koralov, L.B. e Sinai, Y.G. (2007) *Theory of Probability and Random Processes*, seconda edizione, Springer, New York.
- Kotz, S., Balakrishnan, N. e Johnson, N.L. (2000) *Continuous Multivariate Distributions*, volume 1, seconda edizione, New York, Wiley.
- Kubrusly, C.S. (2015) *Essentials of Measure Theory*, Springer, New York.
- Kuo, H. (2006) *Introduction to Stochastic Integration*, Springer, New York.
- Le Gall, J.F. (2016) *Brownian Motion, Martingales, and Stochastic Calculus*, Springer, Berlin.
- Le Gall, J.F. (2022) *Measure Theory, Probability, and Stochastic Processes*, Springer Nature, Switzerland.
- Leadbetter, R., Cambanis, S. e Pipiras, V. (2014) *A Basic Course in Measure and Probability Theory for Applications*, Cambridge University Press, Cambridge.
- Lehmann, E.L. e Casella, G. (1998) *The Theory of Point Estimation*, seconda edizione, Springer, New York.
- Letta, G. (1993) *Probabilità Elementare*, Zanichelli, Bologna.
- Letta, G. (2016) *Elementi di Teoria dei Processi Stocastici*, Unione Matematica Italiana, Bologna.
- Lin, Z. e Bai, Z. (2010) *Probability Inequalities*, Springer, New York.
- Lovász, L., Pelikán, J. e Vesztergombi, K. (2003) *Discrete Mathematics: Elementary and Beyond*, Springer, New York.
- Lukacs, E. (1970) *Characteristic Functions*, seconda edizione, Griffin, London.
- Medvedev, P. (2007) *Stochastic Integration Theory*, Oxford University Press, Oxford.
- Mishura, Y. e Shevchenko, G. (2017) *Theoretical and Statistical Aspects of Stochastic Processes*, Elsevier, Amsterdam.
- Mörters, P. e Peres, Y. (2010) *Brownian Motion*, Cambridge University Press, Cambridge.
- Mosteller, F. (1987) *Fifty Challenging Problems in Probability with Solutions*, seconda edizione, Dover Publications, New York.
- Olofsson, P. (2015) *Probabilities - The Little Numbers that Rule Our Lives*, seconda edizione, Wiley, New York.
- Ore, Ø. (1953) *Cardano: The Gambling Scholar*, Princeton University Press, Princeton.
- Ovchinnikov, S. (2013) *Measure, Integral, Derivative*, Springer, New York.
- Øksendal, B. (2003) *Stochastic Differential Equations*, sesta edizione, Springer, Berlin.
- Pascucci, A. (2011) *PDE and Martingale Methods in Option Pricing*, Bocconi & Springer Series, Springer Italia, Milan.
- Petkovšek, M., Wilf, H.S. e Zeilberger, D. (1996) *A = B*, Peters, Wellesley.
- Petkovic, M.S. (2009) *Famous Puzzles of Great Mathematicians*, American Mathematical Society, Providence.
- Petrov, V.V. (1995) *Limit Theorems of Probability Theory*, Clarendon Press, Oxford.
- von Plato, J. (1994) *Creating Modern Probability*, Cambridge University Press, Cambridge.
- Protter, P. (2003) *Stochastic Integration and Differential Equations*, seconda edizione, Springer, New York.
- Rao, M.M. (2004) *Measure Theory and Integration*, seconda edizione, Dekker, New York.
- Rao, M.M. (2005) *Conditional Measures and Applications*, seconda edizione, Chapman & Hall, New York.
- Rao, M.M. e Swift, R.J. (2006) *Probability Theory with Applications*, seconda edizione, Springer, New York.
- Resnick, S.I. (2005) *Adventures in Stochastic Processes*, Birkhäuser, Boston.

- Resnick, S.I. (2014) *A Probability Path*, Birkhäuser, Boston.
- Revuz, D. e Yor, M. (2005) *Continuous Martingales and Brownian Motion*, terza edizione, Springer, New York.
- Romano, J.P. e Siegel, A.F. (1986) *Counterexamples in Probability and Statistics*, Wadsworth, London.
- Rosenthal, J. (2006) *A First Look at Rigorous Probability Theory*, seconda edizione, World Scientific Publishing, Singapore.
- Roussas, G.G. (2014) *An Introduction to Measure-Theoretic Probability*, Second edition, Academic Press, Oxford.
- Sasvári, Z. (2013) *Multivariate Characteristic and Correlation Functions*, De Gruyter, Berlin.
- Sato, K. (1999) *Lévy Processes and Infinitely Divisible Distributions*, Cambridge University Press, Cambridge.
- Schilling, R. (2005) *Measures, Integrals and Martingales*, Cambridge University Press, Cambridge.
- Schilling, R.L. e Partzsch, L. (2021) *Brownian Motion*, terza edizione, De Gruyter, Berlin.
- Serfling R.J. (1980) *Approximation Theorems of Statistics*, Wiley, New York.
- Shiryayev, A.N. (2016) *Probability-1*, terza edizione, Springer, New York.
- Shiryayev, A.N. (2019) *Probability-2*, terza edizione, Springer, New York.
- Shreve, S. (2004) *Stochastic Calculus for Finance II: Continuous-Time Models*, Springer, New York.
- Spivey, M.Z. (2019) *The Art of Proving Binomial Identities*, CRC Press, Boca Raton.
- Steutel, F.W. e van Harn, K. (2004) *Infinite Divisibility of Probability Distributions on the Real Line*, Dekker, New York.
- Steyer, R. e Nagel, W. (2017) *Probability and Conditional Expectation*, Wiley, Chichester.
- Stroock, D.W. (2013) *Mathematics of Probability*, American Mathematical Society, Providence.
- Székel, G.J. (1986) *Paradoxes in Probability Theory and Mathematical Statistics*, Kluwer, Dordrecht.
- Taylor, M.E. (2006) *Measure Theory and Integration*, American Mathematical Society, Providence.
- Tijms, H.C. (2003) *A First Course in Stochastic Models*, Wiley, Chichester.
- Tijms, H.C. (2012) *Understanding Probability*, terza edizione, Cambridge University Press, Cambridge.
- Uchaikin, V.V. e Zolotarev, V.M. (1999) *Chance and Stability*, De Gruyter, Utrecht.
- van der Vaart, A.W. (1998) *Asymptotic Statistics*, Cambridge University Press, Cambridge.
- Venkatesh, S.S. (2012) *The Theory of Probability: Explorations and Applications*, Cambridge University Press, Cambridge.
- Walsh, J.B. (2012) *Knowing the Odds: An Introduction to Probability*, American Mathematical Society, Providence.
- Wheeden, R.L. e Zygmund, A. (2015) *Measure and Integral*, seconda edizione, CRC Press, Boca Raton.
- Whittle, P. (2000) *Probability via Expectation*, quarta edizione, Springer, New York.
- Wiersema, U. (2008) *Brownian Motion Calculus*, Wiley, Hoboken.
- Wilf, H.S. (2006) *Generatingfunctionology*, Peters, Wellesley.
- Williams, D. (1992) *Probability with Martingales*, Cambridge University Press, Cambridge.
- Wise, G.L. e Hall, E.B. (1993) *Counterexamples in Probability and Real Analysis*, Oxford University Press, Oxford.
- Zolotarev, V.M. (1986) *One-dimensional Stable Distributions*, American Mathematical Society, Providence.