Capitolo 2

La Convergenza Stocastica

2.1 I lemmi di Borel–Cantelli

I seguenti due risultati sono, nella loro semplicità, fondamentali.

Lemma 2.1.1. (Primo lemma di Borel–Cantelli). *Siano* $(\Omega, \mathcal{F}, \mathbb{P})$ *uno spazio di probabilità e* $(A_n) \subseteq \mathcal{F}$ *una successione di insiemi misurabili tale che*

$$\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) < +\infty\,.$$

Allora

$$\mathbb{P}\left(\limsup_{n\to+\infty} A_n\right) = 0\,.$$

Dimostrazione. La successione di insiemi $(\cup_{j\geq n} A_j)_{n\in\mathbb{N}}$ è decrescente, sicché

$$\begin{aligned}\mathbb{P}\left(\limsup_{n\to+\infty} A_n\right) &= \mathbb{P}\left(\bigcap_{n\in\mathbb{N}}\bigcup_{j\geq n} A_j\right)\\ &= \lim_{n\to+\infty} \mathbb{P}\left(\bigcup_{j\geq n} A_j\right) \leq \lim_{n\to+\infty} \sum_{j\geq n} \mathbb{P}(A_j) = 0\,,\end{aligned}$$

perché $\sum_{j\geq n} \mathbb{P}(A_j)$ è il resto $(n-1)$–esimo della serie $\sum_{n\in\mathbb{N}} \mathbb{P}(A_n)$ che è, per ipotesi, convergente. □

Lemma 2.1.2. (Secondo lemma di Borel–Cantelli). *Se gli eventi* $A_n \in \mathcal{F}$ *sono indipendenti e se* $\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) = +\infty$*, allora*

$$\mathbb{P}\left(\limsup_{n\to+\infty} A_n\right) = 1\,.$$

Dimostrazione. Come sopra

$$\mathbb{P}\left(\limsup_{n\to+\infty} A_n\right) = \lim_{n\to+\infty} \mathbb{P}\left(\bigcup_{j\geq n} A_j\right)\,.$$

Poiché

$$\mathbb{P}\left(\bigcup_{j\geq n} A_j\right) = 1 - \mathbb{P}\left(\bigcap_{j\geq n} A_j^c\right) = 1 - \prod_{j\geq n} \mathbb{P}(A_j^c)\,,$$

basterà mostrare che $\prod_{j\geq n} \mathbb{P}(A_j^c) \xrightarrow[n\to+\infty]{} 0$. A tal fine, si ricordi la diseguaglianza $\ln x \leq x-1$, nella quale il segno d'eguaglianza vale se, e solo se, $x=1$. Posto, per $t \in [0,1[$, $x = 1-t$, si ha $\ln(1-t) \leq -t$, onde, per $s \geq n$,

$$\begin{aligned}\ln \prod_{j=n}^{s} \mathbb{P}(A_j^c) &= \sum_{j=n}^{s} \ln \mathbb{P}(A_j^c) \\ &= \sum_{j=n}^{s} \ln\left(1 - \mathbb{P}(A_j)\right) \leq -\sum_{j=n}^{s} \mathbb{P}(A_j) \xrightarrow[s\to+\infty]{} -\infty\,,\end{aligned}$$

onde $\prod_{j\geq n} \mathbb{P}(A_j^c) \xrightarrow[n\to+\infty]{} 0$. □

Il seguente risultato è una conseguenza immediata dei lemmi di Borel–Cantelli. Esso è il prototipo delle cosiddette *leggi zero–uno*; sono, queste, teoremi che asseriscono che la probabilità di certi insiemi può assumere solo i valori zero o uno.

Teorema 2.1.1. (Legge zero-uno di Borel). *Se (A_n) è una successione di eventi indipendenti di $\mathcal{F}$, risulta*

$$\mathbb{P}\left(\limsup_{n\to+\infty} A_n\right) = 0 \qquad \textit{oppure} \qquad \mathbb{P}\left(\limsup_{n\to+\infty} A_n\right) = 1\,,$$

secondo che risulti convergente o divergente la serie $\sum_{n\in\mathbb{N}} \mathbb{P}(A_n)$.

Ecco alcune applicazioni dei lemmi di Borel–Cantelli al gioco di testa o croce (o, se si preferisce, ad un processo di Bernoulli).

Esempio 2.1.1. Si consideri un'assegnata sequenza finita di risultati,

$$s := (x_1, \dots, x_k)$$

ove $x_i = 1$ oppure 0 $(i = 1, 2, \dots, k)$ e si considerino gli eventi

$$A_n := \left\{(x_n)_{n\in\mathbb{N}} \in \{0,1\}^{\mathbb{N}} : (x_n, x_{n+1}, \dots, x_{n+k-1}) = s\right\}$$

e sia $p \in]0,1[$ la probabilità di successo, $\mathbb{P}(x_n = 1) = 1$; $\{0,1\}^{\mathbb{N}}$ è l'insieme di tutte le successioni composte di 0 e di 1.

Per far vedere che si ha $\mathbb{P}(\limsup_{n\to+\infty} A_n) = 1$, non si può applicare il secondo lemma di Borel–Cantelli direttamente alla successione (A_n) perché gli eventi che la compongono non sono indipendenti. Si considerino, tuttavia, gli eventi

$$B_n := A_{(n-1)k+1} = \left\{(x_n) \in \{0,1\}^{\mathbb{N}} : (x_{(n-1)k+1}, x_{(n-1)k+2}, \dots, x_{nk}) = s\right\}\,.$$

Questi sono, evidentemente, indipendenti ed inoltre verificano l'inclusione

$$\limsup_{n\to+\infty} B_n \subseteq \limsup_{n\to+\infty} A_n\,.$$

Ora, $\mathbb{P}(B_n) = \mathbb{P}(B_1) = \mathbb{P}(s) > 0$ per ogni $n \in \mathbb{N}$, sicché $\sum_{n\in\mathbb{N}} \mathbb{P}(B_n) = +\infty$, ciò che implica

$$\mathbb{P}\left(\limsup_{n\to+\infty} B_n\right) = 1\,,$$

e, *a fortiori*, $\mathbb{P}(\limsup_{n\to+\infty} A_n) = 1$. Perciò, in una serie infinita di lanci di una moneta, è eguale a 1 la probabilità che un'assegnata sequenza finita di teste e di croci torni a verificarsi un numero infinito di volte.

Questo risultato complementa quello che si è già incontrato per il quale occorrerà aspettare, perché si realizzi per la prima volta la sequenza s, un tempo che è inversamente proporzionale alla probabilità di s. ■

Ancora, in un processo di Bernoulli, si ponga $Y_i(1) = 1$, $Y_i(0) = -1$ e $G_n = \sum_{i\leq n} Y_i$. Quando $G_n = 0$, si dirà che, al tempo n, il gioco è in parità, oppure, pensando alla passeggiata aleatoria, che al tempo n, il processo è nell'origine. Se $A_n := \{G_n = 0\}$, $\limsup_{n\to+\infty} A_n$ rappresenta l'evento "si ha parità infinite volte" oppure, pensando alla passeggiata aleatoria, l'evento "il processo torna infinite volte nell'origine".

Teorema 2.1.2. *Se $p \neq 1/2$, allora si ha*

$$\mathbb{P}\left(\limsup_{n\to+\infty} A_n\right) = 0\,.$$

Dimostrazione. Ora $\mathbb{P}(A_{2n+1}) = 0$, mentre

$$\mathbb{P}(A_{2n}) = \mathbb{P}(G_{2n} = 0) = \binom{2n}{n} p^n\, q^n\,,$$

onde

$$\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) = \sum_{j\in\mathbb{N}} \mathbb{P}(A_{2j}) = \sum_{j\in\mathbb{N}} \binom{2j}{j} p^j\, q^j\,.$$

Quest'ultima serie è convergente, come subito ci si assicura, usando, ad esempio, il criterio del rapporto; l'asserto è ora un'immediata conseguenza del Lemma 2.1.1. □

Rimane da esaminare il caso $p = 1/2$. Stabiliremo prima il segunete risultato ausiliario

Lemma 2.1.3. *Per ogni $\alpha \in\,]0,1[$ esiste una funzione $\varphi : \mathbb{N} \to \mathbb{N}$ tale che*

$$\mathbb{P}\left(\left|G_{\varphi(j)}\right| < j\right) \leq \alpha\,.$$

Dimostrazione. Per ogni $k \in \mathbb{Z}$, si ha

$$\lim_{n\to+\infty} \mathbb{P}(G_n = k) = 0\,;$$

perciò

$$\lim_{n\to+\infty} \sum_{|k|<j} \mathbb{P}(G_n = k) = 0\,.$$

Basta allora scegliere $\varphi(j)$ abbastanza grande da avere

$$\sum_{|k|<j} \mathbb{P}\left(G_{\varphi(j)} = k\right) \leq \alpha\,,$$

che conclude la dimostrazione. □

Teorema 2.1.3. *Se $p = 1/2$, allora si ha*

$$\mathbb{P}\left(\limsup_{n\to+\infty} A_n\right) = 1\,.$$

Dimostrazione. Come nell'esempio 2.1.1 non si può usare direttamente il secondo lemma di Borel–Cantelli perché gli eventi della successione (A_n) non sono indipendenti. Si consideri una successione strettamente crescente (n_j) estratta da quella dei numeri naturali e si supponga che sia $n_{j+1} - n_j \geq 2$. Per ogni $j \in \mathbb{N}$, sia s_j un naturale con $n_j < s_j < n_{j+1}$ e si definisca

$$B_j := \left\{\sum_{k=n_j+1}^{s_j} Y_k \leq -n_j\right\} \bigcap \left\{\sum_{k=s_j+1}^{n_{j+1}} Y_k \geq s_j\right\}\,.$$

Si noti che B_j è un insieme misurabile e che dipende dalle v.a.

$$Y_{n_j+1}, Y_{n_j+2}, \dots, Y_{n_{j+1}}\,;$$

poiché le v.a. della successione (Y_n) sono indipendenti anche gli insiemi (B_j) lo sono.

Ora, poiché ogni v.a. Y_k assume solo i valori 1 e -1, si ha

$$G_{n_j} = \sum_{k=1}^{n_j} Y_k \leq n_j \qquad \text{e} \qquad G_{s_j} \geq -s_j\,.$$

perciò, se $\Omega \in B_j$, vale tanto $G_{s_j}(\Omega) \leq 0$ quanto $G_{n_{j+1}}(\Omega) \geq 0$, sicché esiste un naturale n compreso tra $n_j + 1$ e n_{j+1} per il quale $G_n(\Omega) = 0$. Dunque

$$\limsup_{j\to+\infty} B_j \subseteq \limsup_{n\to+\infty} \{G_n = 0\}\,.$$

In virtú del secondo lemma di Borel–Cantelli, per dimostrare l'asserto, basta far vedere che è possibile scegliere i numeri n_j e s_j in modo che sia

$$\sum_{j\in\mathbb{N}} \mathbb{P}(B_j) = +\infty\,. \tag{2.1.1}$$

Si proceda ora a scegliere i numeri n_j e s_j come segue

$$n_1 = 1 \qquad\qquad s_1 = 1 + \varphi(1)\,,$$

e, per $j > 1$,

$$s_j = n_j + \varphi(n_j), \qquad n_{j+1} = s_j + \varphi(s_j)\,.$$

Ora

$$\mathbb{P}(B_j) = \mathbb{P}\left(\sum_{k=n_j+1}^{s_j} Y_k \leq -n_j\right) \mathbb{P}\left(\sum_{k=s_j+1}^{n_{j+1}} Y_k \geq s_j\right)\,,$$

e di qui, per la simmetria della passeggiata aleatoria ($p = 1/2$),

$$\mathbb{P}(B_j) = \frac{1}{4}\,\mathbb{P}\left(\sum_{k=n_j+1}^{s_j} |Y_k| \geq n_j\right) \mathbb{P}\left(\sum_{k=s_j+1}^{n_{j+1}} |Y_k| \geq s_j\right)\,;$$

ma i vettori aleatorî $(Y_{i+1}, Y_{i+2}, \dots, Y_{i+k})$ e $(Y_1, Y_2, \dots, Y_k)$ hanno la stessa legge, sicché per il Lemma 2.1.3

$$\mathbb{P}(B_j) = \frac{1}{4}\,\mathbb{P}\left(\sum_{k=1}^{\varphi(n_j)} |Y_k| \geq n_j\right)\,\mathbb{P}\left(\sum_{k=1}^{\varphi(s_j)} |Y_k| \geq s_j\right) \geq \frac{1}{4}\,(1-\alpha)^2 > 0\,,$$

ciò che prova la (2.1.1) e quindi l'asserto. □

2.2 Varî tipi di convergenza stocastica

L'approccio piú ovvio alla convergenza di una successione di v.a. $(X_n)_{n\in\mathbb{N}}$ è di stabilire che la successione converga per ogni punto $\omega \in \Omega$. Per esempio, se X è una v.a. e $X_n := X + 1/n$, la successione $\{X_n\}$ cosí definita converge a X in ogni $\omega \in \Omega$. Tuttavia, questa definizione è eccessivamente restrittiva. Si consideri infatti un processo di Bernoulli nel quale sia $p \in \,]0,1[$ la probabilità di un successo ad ogni prova; se $X_n = 1$ oppure 0 secondo che la n–esima prova abbia avuto come risultato un successo o un fallimento, e se $S_n := (1/n)\sum_{j\leq n} X_j$ indica la frequenza dei successi, la legge dei grandi numeri di Bernoulli asserisce che la probabilità che S_n si discosti da p può essere resa arbitrariamente piccola al tendere di n a $+\infty$. Però S_n non tende a p in ogni punto di $\Omega = \{0,1\}^{\mathbb{N}}$; per esempio, per la successione $(0, 0, \dots, 0, \dots)$ si ha $S_n = 0$ per ogni $n \in \mathbb{N}$, mentre per la successione $(1, 1, \dots, 1, \dots)$ è $S_n = 1$ per ogni $n \in \mathbb{N}$. è facile dare esempî di successioni per le quali la frequenza dei successi è eguale ad un preassegnato numero (razionale). È cosí giustificata l'introduzione di altri tipi di convergenza. *Tutte le v.a. che compaiono nelle definizioni e nei teoremi di questa sezione si supporranno definite sopra il medesimo spazio di probabilità* $(\Omega, \mathcal{F}, \mathbb{P})$.

Definizione 2.2.1. Si dice che la successione (X_n) converge *quasi certamente* alla v.a. X se esiste un insieme trascurabile $N \in \mathcal{F}$, $\mathbb{P}(N) = 0$, tale che per ogni $\omega \in N^c$ si abbia $X_n(\omega) \to X(\omega)$. Si scrive allora

$$X_n \xrightarrow[n\to+\infty]{q.c.} X\,.$$

Diremo questo tipo di convergenza *convergenza quasi certa.* ◇

È opportuno porre la Definizione 2.2.1 in altri termini. Si definisca, per $\varepsilon > 0$ fissato arbitrariamente, l'insieme $A_{j,\varepsilon} := \{|X_j - X| \leq \varepsilon\}$. Poiché tanto X_j quanto X sono misurabili, si ha che $A_{j,\varepsilon}$ appartiene a $\mathcal{F}$. Si considerino, inoltre, gli insiemi, che sono tutti misurabili,

$$S_{n,\varepsilon} := \bigcap_{j\geq n} A_{j,\varepsilon}$$

e

$$S_\varepsilon := \bigcup_{n\in\mathbb{N}} S_{n,\varepsilon} = \bigcup_{n\in\mathbb{N}}\bigcap_{j\geq n} A_{j,\varepsilon} = \liminf_{n\to+\infty} A_{n,\varepsilon}\,.$$

Si osservi che, se $0 < \varepsilon < \varepsilon'$, allora, per ogni $j \in \mathbb{N}$, si ha $A_{j,\varepsilon} \subseteq A_{j,\varepsilon'}$, e, di conseguenza, $S_\varepsilon \subseteq S_{\varepsilon'}$. Sia, ora, (ε_n) una successione decrescente e infinitesima con $\varepsilon_n > 0$ per ogni $n \in \mathbb{N}$ e si ponga $S := \cap_{n\in\mathbb{N}} S_{\varepsilon_n} \in \mathcal{F}$. L'insieme S non dipende dalla scelta della successione (ε_n); i punti di S sono tutti, e soli, i punti $\omega \in \Omega$ per i quali risulta $|X_n(\omega) - X(\omega)| \leq \varepsilon$, comunque si fissi $\varepsilon > 0$, per tutti gli indici n, tranne, al piú, un numero finito di essi; tale numero dipende in generale dal punto ω; si può dunque scrivere $\mathbb{P}(S) = \mathbb{P}(\lim_{n\to+\infty} X_n = X)$. Pertanto, dire che la successione (X_n) converge q.c. alla v.a. X equivale a dire che $\mathbb{P}(S) = 1$.

Poiché la scelta della successione infinitesima (ε_n) è irrilevante, si può prendere, senza perdita di generalità, $\varepsilon_n = 1/n$. In questo modo, la condizione che esprime la convergenza q.c. può essere espressa nella forma

$$\mathbb{P}\left(\bigcap_{m\in\mathbb{N}} \bigcup_{n\in\mathbb{N}} \bigcap_{j\geq n} \left\{|X_j - X| \leq \frac{1}{m}\right\}\right) = 1\,.$$

È spesso utile la seguente caratterizzazione

Teorema 2.2.1. *Per la successione (X_n) di v.a. sono equivalenti le proprietà:*

(a) *(X_n) converge q.c. alla v.a. X;*

(b) *per ogni $\delta > 0$, risulta $\lim_{n\to+\infty} \mathbb{P}(S_{n,\delta}) = 1$.*

Dimostrazione. (a) $\Longrightarrow$ (b) Si è visto che $\mathbb{P}(S) = 1$ e poiché $S \subseteq S_{\varepsilon_n}$ per ogni $n \in \mathbb{N}$, si ha $\mathbb{P}(S_{\varepsilon_n}) = 1$ per ogni $n \in \mathbb{N}$. Fissato arbitrariamente $\delta > 0$, si scelga un elemento ε_k della successione (ε_n) in modo che sia $\varepsilon_k < \delta$. Segue dalla definizione che $S_{\varepsilon_k} \subseteq S_\delta$ cosí che $\mathbb{P}(S_\delta) = 1$. Poiché $S_\delta = \cup_n S_{n,\delta} = \lim_{n\to+\infty} S_{n,\delta}$, si ha l'asserto.
(b) $\Longrightarrow$ (a) Fissato arbitrariamente $\varepsilon > 0$, si ponga $\delta = \varepsilon 2^{-k}$; perciò, esiste un numero naturale $n_k = n(\varepsilon, k)$ tale che, per ogni $n \geq n_k$, sia

$$\mathbb{P}(S_{n,\varepsilon 2^{-k}}) > 1 - \varepsilon\, 2^{-k}\,.$$

Poiché

$$S = \bigcap_{k\in\mathbb{N}} S_{\varepsilon 2^{-k}} \supset \bigcap_{k\in\mathbb{N}} S_{n_k,\,\varepsilon 2^{-k}} = \left(\bigcup_{k\in\mathbb{N}} S^c_{n_k,\,\varepsilon 2^{-k}}\right)^c,$$

si ha

$$\begin{aligned}\mathbb{P}(S) &\geq 1 - \mathbb{P}\left(\bigcup_{k\in\mathbb{N}} S^c_{n_k,\,\varepsilon 2^{-k}}\right)\\ &\geq 1 - \sum_{k\in\mathbb{N}} \mathbb{P}\left(S^c_{n_k,\,\varepsilon 2^{-k}}\right) > 1 - \sum_{k\in\mathbb{N}} \varepsilon\, 2^{-k} = 1 - \varepsilon\,,\end{aligned}$$

che dà l'asserto. $\square$

La condizione (b) dell'ultimo teorema si può scrivere in una delle forme equivalenti:

$$\lim_{n\to+\infty} \mathbb{P}\left(\bigcap_{j=n}^{\infty}\{|X_j - X| \leq \delta\}\right) = 1\,; \tag{2.2.1}$$

$$\lim_{n\to+\infty} \mathbb{P}\left(\sup_{j\geq n}|X_j - X| \leq \delta\right) = 1\,; \tag{2.2.2}$$

$$\lim_{n\to+\infty} \mathbb{P}\left(\bigcup_{j=n}^{\infty}\{|X_j - X| > \delta\}\right) = 0\,; \tag{2.2.3}$$

$$\lim_{n\to+\infty} \mathbb{P}\left(\sup_{j\geq n}|X_j - X| > \delta\right) = 0\,. \tag{2.2.4}$$

Il seguente semplice risultato illustra l'uso dei lemmi di Borel–Cantelli nello studio della convergenza quasi certa.

Teorema 2.2.2. *Per una successione* (X_n) *di v.a. definite sullo spazio di probabilità* $(\Omega, \mathcal{F}, \mathbb{P})$ *sono equivalenti le seguenti condizioni:*

(a) $X_n \to 0$ *q.c.;*

(b) $\forall\, \varepsilon > 0 \quad \mathbb{P}\left(\limsup_{n\to+\infty}\{|X_n| \geq \varepsilon\}\right) = 0$*;*

(c) $\forall\, \varepsilon > 0 \quad \mathbb{P}\left(\liminf_{n\to+\infty}\{|X_n| < \varepsilon\}\right) = 1$*;*

Dimostrazione. Poiché è ovvia l'equivalenza tra le condizioni (b) e (c), basta dimostrare che una di queste, per esempio la (c), è equivalente alla (a).

Ora, per il Teorema 2.2.1, $X_n \to 0$ q.c. se, e solo se, posto

$$S_{n,\varepsilon} := \bigcap_{j\geq n}\{|X_j| < \varepsilon\}\,,$$

è $\lim_{n\to+\infty}\mathbb{P}(S_{n,\varepsilon}) = 1$, per ogni $\varepsilon > 0$, che è la (c) perché la successione $(S_{n,\varepsilon})$ è crescente. □

Definizione 2.2.2. Si dice che una successione (X_n) di v.a. definite sullo stesso spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ è *di Cauchy rispetto alla convergenza* q.c. se esiste un insieme $N \in \mathcal{F}$ con $\mathbb{P}(N) = 0$ tale che, per ogni $\varepsilon > 0$ e per ogni $\omega \in N^c$, si possa determinare un naturale $n_0 = n_0(\varepsilon) \in \mathbb{N}$ con la proprietà che, per ogni scelta di n e m con $m, n \geq n_0$, si abbia $|X_n(\omega) - X_m(\omega)| < \varepsilon$. In termini piú formali, se

$$\mathbb{P}\left(\bigcap_{m\in\mathbb{N}}\bigcup_{n\in\mathbb{N}}\bigcap_{j,k\geq n}\left\{|X_j - X_k| < \frac{1}{m}\right\}\right) = 1\,,$$

allora (X_n) è una successione di Cauchy. ◇

L'ultima equazione può essere scritta in forme differenti ma equivalenti; per esempio, equivale a richiedere che, per ogni $\varepsilon > 0$, valga

$$\lim_{n\to+\infty} \mathbb{P}\left(\sup_{k\in\mathbb{N}} |X_{n+k} - X_n| \le \varepsilon\right) = 1\,. \tag{2.2.5}$$

Teorema 2.2.3. *Per una successione (X_n) di v.a. definite sullo spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ sono equivalenti le affermazioni:*

(a) *(X_n) converge q.c.*

(b) *(X_n) è di Cauchy rispetto alla convergenza quasi certa.*

Dimostrazione. (a) $\Longrightarrow$ (b) Si supponga che esista una v.a. X tale $X_n \to X$ q.c.. Vogliamo dimostrare che per ogni $\varepsilon > 0$ vale la (2.2.5). Si introducano gli insiemi

$$B^\varepsilon_{n,k} := \{|X_{n+k} - X_n| \le \varepsilon\}\,, \qquad T^\varepsilon_k := \bigcap_{n\in\mathbb{N}} B^\varepsilon_{n,k}$$
$$S^\varepsilon_n := \bigcap_{k\ge n} T^\varepsilon_k, \qquad S^\varepsilon := \bigcup_{n\in\mathbb{N}} S^\varepsilon_n\,.$$

Poiché $|X_{n+k} - X_n| \le |X_{n+k} - X| + |X_n - X|$, si ha, se $A_{n,\varepsilon} := \{|X_n - X| \le \varepsilon\}$,

$$A_{n+k,\varepsilon/2} \bigcap A_{n,\varepsilon/2} \subseteq B^\varepsilon_{n,k}\,,$$

sicché

$$S^{\varepsilon/2}_n \subseteq A_{n+k,\varepsilon/2} \bigcap A_{n,\varepsilon/2} \subseteq B^\varepsilon_{n,k}\,,$$

per ogni $k \in \mathbb{N}$. Di qui

$$S^{\varepsilon/2}_n \subseteq T^\varepsilon_k = \left\{\sup_{k\in\mathbb{N}} |X_{n+k} - X_n| \le \varepsilon\right\}\,.$$

Poiché, per il Teorema 2.2.1, si ha $\mathbb{P}\left(S^{\varepsilon/2}_n\right) \to 1$, segue anche la (2.2.5).

(b) $\Longrightarrow$ (a) Sia (X_n) una successione di Cauchy rispetto alla convergenza quasi certa, vale a dire sia vera la (2.2.5); si noti che, per ogni naturale $r \ge n$, si ha

$$\begin{aligned} T^{\varepsilon/2}_k &\subseteq \{|X_{r+k} - X_n| \le \varepsilon/2\} \bigcap \{|X_r - X_n| \le \varepsilon/2\} \\ &\subseteq \{|X_{r+k} - X_r| \le \varepsilon\}\,, \end{aligned}$$

sicché

$$T^{\varepsilon/2}_k \subseteq B^\varepsilon_{r,k}$$

per ogni $r \ge n$ e per ogni $k \in \mathbb{N}$. Quindi

$$T^{\varepsilon/2}_k \subseteq T^\varepsilon_k\,,$$

sicché $T^{\varepsilon/2}_n \subseteq S^\varepsilon_n$ e perciò

$$\lim_{n\to+\infty} \mathbb{P}\left(T^{\varepsilon/2}_n\right) \le \lim_{n\to+\infty} \mathbb{P}\left(S^\varepsilon_n\right) = \mathbb{P}\left(S^\varepsilon\right)\,.$$

Quest'ultima relazione e la (2.2.5) implicano

$$\mathbb{P}\left(S^{\varepsilon}\right)=1\,.$$

Sia ora (ε_j) un'arbitraria successione infinitesima di numeri strettamente positivi e si consideri l'insieme

$$S:=\bigcap_{j\in\mathbb{N}}S^{\varepsilon_j}\,,$$

per il quale risulta $\mathbb{P}(S)=1$. La successione $(X_n(\omega))$ soddisfà alla condizione di Cauchy per ogni $\omega\in S$. Esiste dunque una funzione Y tale che per ogni $\omega\in S$ sia $\lim_{n\to+\infty}X_n(\omega)=Y(\omega)$. Tale Y può essere estesa all'intero spazio Ω, ponendo

$$X(\omega):=\begin{cases}Y(\omega), & \omega\in S\,,\\ 0, & \omega\in S^c\,.\end{cases}$$

Poiché $\mathbb{P}(S^c)=0$, si ha che X è una v.a. e che $X_n\to X$ q.c.. □

La convergenza quasi certa non discende, in generale, da una metrica, anzi, come si dimostrerà oltre, non è neppure topologica. Vale il seguente teorema che non dimostriamo.

Teorema 2.2.4. *Sia L^0 lo spazio delle v.a. che sono quasi certamente finite su $(\Omega,\mathcal{F},\mathbb{P})$. Affinché esista una metrica su L^0 che sia compatibile con la convergenza quasi certa occorre e basta che Ω sia l'unione numerabile punti $\{\omega_j:j\in\mathbb{N}\}$ con $\mathbb{P}(\{\omega_j\})>0$ per ogni $j\in\mathbb{N}$.*

Definizione 2.2.3. Si dice che la successione (X_n) di v.a. converge *in probabilità* alla v.a. X se, per ogni $\varepsilon>0$, risulta

$$\lim_{n\to+\infty}\mathbb{P}\left(|X_n-X|>\varepsilon\right)=0\,. \tag{2.2.6}$$

Si scriverà $X_n\xrightarrow[n\to+\infty]{\mathbb{P}}X$. Chiameremo questo tipo di convergenza *convergenza in probabilità*. ◇

Vale la pena osservare che la definizione di convergenza in probabilità asserisce che, per ogni $\varepsilon>0$ e per ogni $\delta>0$, esiste un numero naturale $n_0=n_0(\varepsilon,\delta)$, tale che per ogni $n\geq n_0$,

$$\mathbb{P}\left(|X_n-X|>\varepsilon\right)\leq\delta\,.$$

Si osservi che si è già incontrato un risultato significativo nel quale compariva la convergenza in probabilità: si tratta della legge dei grandi numeri di Bernoulli.

Teorema 2.2.5. *La convergenza quasi certa implica la convergenza in probabilità.*

Dimostrazione. Se la successione (X_n) converge q.c. a X essa verifica la (2.2.4) e quindi, *a fortiori*, anche la (2.2.6). □

Si vedrà piú avanti che non è necessariamente vero il viceversa; è tuttavia notevole il seguente risultato

Teorema 2.2.6. *Da ogni successione di v.a.* (X_n) *che converga in probabilità ad una v.a.* X*, si può estrarre una sottosuccessione che converge quasi certamente allo stesso limite* X*.*

Dimostrazione. Si supponga che (X_n) converga in probabilità a X e si scelga $n(k)$ in $\mathbb{N}$ in modo che risulti $\mathbb{P}(|X_{n(k)} - X| > 2^{-k}) \leq 2^{-k}$; si sceglierà l'indice $n(k+1)$ in maniera analoga, con la sola condizione che sia $n(k+1) > n(k)$. Il primo lemma di Borel–Cantelli implica

$$\mathbb{P}\left(\limsup_{k\to +\infty}\{|X_{n(k)} - X| > 2^{-k}\}\right) = 0\,,$$

cioè

$$\mathbb{P}\left(\liminf_{k\to +\infty}\{|X_{n(k)} - X| \leq 2^{-k}\}\right) = 1\,.$$

Nella notazione di questa sezione si è appena dimostrato che, con riferimento alla successione $(X_{n(k)})$, per ogni $k \in \mathbb{N}$, si ha $\mathbb{P}\left(S_{2^{-k}}\right) = 1$, sicché

$$\mathbb{P}(S) = \mathbb{P}\left(\bigcap_{k\in\mathbb{N}} S_{2^{-k}}\right) = 1\,,$$

vale a dire che $X_{n(k)} \xrightarrow[k\to+\infty]{} X$ q.c.. □

Si vedrà oltre che la convergenza in probabilità discende da una metrica su L^0. Nel seguito useremo liberamente questo fatto.

Definizione 2.2.4. Si dice che la successione $(X_n) \subseteq L^p$ converge alla v.a. X in L^p (oppure in media di ordine p) con $p \in]0, +\infty[$ se

$$\lim_{n\to+\infty} \mathbb{E}\left(|X_n - X|^p\right) = 0\,.$$

Questo modo di convergenza sarà detto *convergenza in* L^p. ◇

Nelle convergenze delle Definizioni 2.2.1, 2.2.3 e 2.2.4, (X_n) converge a X se, e solo se, $(X_n - X)$ converge alla v.a. q.c. nulla. Nelle dimostrazioni che seguono ci si limiterà dunque, senza perdita di generalità, ad esaminare il caso in cui la v.a. limite sia (q.c.) nulla.

Teorema 2.2.7. *Se* $p < r$*, allora la convergenza in* L^r *implica quella in* L^p*.*

Dimostrazione. È conseguenza immediata della (4.5.9). □

Teorema 2.2.8. *La convergenza in* L^p *con* $p \in]0, +\infty[$ *implica quella in probabilità.*

Dimostrazione. È conseguenza immediata della (1.9.4). □

Quest'ultimo risultato ammette un reciproco parziale. Si vedrà nel seguito che questo teorema è conseguenza di un risultato piú generale (Teorema 6.4.2).

Teorema 2.2.9. *Se (X_n) converge a X in probabilità e se esiste una v.a. Y di L^p tale che $|X_n| \leq Y$, per ogni $n \in \mathbb{N}$, allora (X_n) converge a X anche in L^p.*

Dimostrazione. È, intanto, evidente che la condizione $|X_n| \leq Y$ per ogni $n \in \mathbb{N}$, con $Y \in L^p$, assicura che la successione (X_n) è in L^p. Basta far vedere che da ogni sottosuccessione di (X_n) se ne può estrarre un'altra convergente a X in L^p. Per il Teorema 2.2.6 esiste una sottosuccessione $(X_{n(k)})$ che converge a X q.c.. Poiché, evidentemente, si ha $|X| \leq Y$, risulta, per ogni $k \in \mathbb{N}$,

$$|X_{n(k)} - X|^p \leq 2^p \, Y^p \,,$$

sicché, per il teorema di convergenza dominata, si ha

$$\lim_{k \to +\infty} \mathbb{E}\left(|X_{n(k)} - X|^p\right) = 0 \,,$$

che dà l'asserto. □

Gli esempî che seguono illustrano i rapporti tra i varî modi di convergenza sin qui introdotti.

Esempio 2.2.1. (La convergenza in L^p non implica la convergenza quasi certa). Sia $\Omega = [0,1]$ dotato della misura di Lebesgue; si considerino le v.a.

$$X_{rs} := \mathbf{1}_{\left]\frac{s-1}{r}, \frac{s}{r}\right[} \qquad (r \in \mathbb{N}; s = 1, \dots, r)$$

e le si pensino ordinate secondo l'ordine lessicografico sí da avere la successione

$$(X_{1,1}, X_{2,1}, X_{2,2}, X_{3,1}, X_{3,2}, X_{3,3}, \dots) \,.$$

Questa converge in L^p per ogni $p \in \,]0, +\infty[$; infatti

$$\mathbb{E}\left(|X_{r,s}|^p\right) = \int |X_{r,s}|^p \, \mathrm{d}\,\mathbb{P} = \int X_{r,s}^p \, \mathrm{d}\,\mathbb{P} = \frac{1}{r} \,.$$

D'altra parte, se $\omega \neq 0$, $(X_{r,s}(\omega))$ è una successione reale composta da 0 e da 1, sicché essa converge se, e solo se, è definitivamente costante; però, essa assume, per ogni valore di r, esattamente una volta il valore 1 e $r-1$ volte il valore 0, sicché non converge. Perciò $(X_{r,s})$ non converge in alcun punto di $]0,1]$ e quindi non converge q.c.. Si osservi che, in virtú del Teorema 2.2.8, $(X_{r,s})$ tende in probabilità alla v.a. nulla, sicché essa fornisce anche l'esempio di una successione che converge in probabilità senza convergere quasi certamente. ■

Esempio 2.2.2. (La convergenza quasi certa non implica quella in L^p). Nello stesso spazio di probabilità dell'esempio precedente, si consideri la successione (X_n) ove

$$X_n := e^n \, \mathbf{1}_{\left[0, \frac{1}{n}\right]} \,.$$

Si controlla subito che $\lim_{n \to +\infty} X_n(\omega) = 0$ se $\omega \neq 0$, onde $X_n \to 0$ q.c.. D'altra parte, per ogni $p > 0$ è

$$\mathbb{E}\left(|X_n|^p\right) = \int X_n^p \, \mathrm{d}\,\mathbb{P} = \frac{e^{np}}{n} \to +\infty \,.$$

Si osservi che, in virtú del Teorema 2.2.7, (X_n) tende in probabilità alla v.a. nulla, sicché essa fornisce l'esempio di una succesione di v.a. che converge in probabilità ma non in L^p. ■

Nei due esempî 2.2.1 e 2.2.2 si è visto che la convergenza quasi certa non implica né è implicata dalla convergenza in L^p. Qui di seguito, diamo l'esempio di una successione che converge sia in L^p sia quasi certamente. I rapporti tra i due tipi di convergenza restano cosí del tutto chiariti.

Esempio 2.2.3. (Una successione che converge sia quasi certamente sia in L^p). Sia X_n una v.a. che assume i valori $-1/n$ e $1/n$, entrambi con probabilità $1/2$. Pertanto, $\mathbb{E}\left(|X_n|^p\right) = 1/(n^p)$ e, quindi, $X_n \to 0$ in L^p. D'altro canto, sia $A_{j,\delta} := \{|X_j| \leq \delta\}$ e si noti che, per $j < k$, si ha $|X_j| > |X_k|$, onde $A_{j,\delta} \subseteq A_{k,\delta}$ e perciò

$$S_{n,\delta} = \bigcap_{j \geq n} A_{j,\delta} = A_{n,\delta} \,.$$

Fissato arbitrariamente $\delta > 0$, risulta

$$\mathbb{P}(S_{n,\delta}) = \mathbb{P}(A_{n,\delta}) = \mathbb{P}(|X_n| \leq \delta) = 1 \,,$$

se $n \geq 1/\delta$, sicché $X_n \to 0$ q.c. in virtú del Teorema 2.2.1. ■

Esempio 2.2.4. (Se $s > r$, la convergenza in L^r non implica quella in L^s). Sia (X_n) una successione di v.a. tali che $X_n = n$ con probabilità $\mathbb{P}(X_n = n) = 1/n^s$ e $X_n = 0$ con probabilità $\mathbb{P}(X_n = 0) = 1 - 1/n^s$. Perciò, $\mathbb{E}(|X_n|^r) = n^r/n^s = 1/n^{s-r} \to 0$, sicché $X_n \to 0$ in L^r, ma $\mathbb{E}(|X_n|^s) = 1$ per ogni $n \in \mathbb{N}$, e quindi (X_n) non tende a zero in L^s. ■

La convergenza in probabilità discende da una metrica. Nel prossimo teorema si dà l'esempio di una tale metrica; poiché discende da una metrica si potranno usare i risultati noti dalla topologia, in primo luogo, si potrà parlare della *topologia* della convergenza in probabilità.

Teorema 2.2.10. *Sia dato lo spazio di probabilità* $(\Omega, \mathcal{F}, \mathbb{P})$. *La funzione* $d_{KF} : L^0 \times L^0 \to \mathbb{R}_+$ *definita da*

$$d_{KF}(X, Y) := \inf \{\varepsilon > 0 : \mathbb{P}\left(|X - Y| > \varepsilon\right) < \varepsilon\} \tag{2.2.7}$$

è una metrica su L^0, *detta* metrica di Ky Fan*; inoltre la successione di v.a.* (X_n) *converge in probabilità alla v.a.* X *se, e solo se,* $d_{KF}(X_n, X) \xrightarrow[n\to+\infty]{} 0$.

Dimostrazione. Che d_{KF} sia positiva e simmetrica è una conseguenza immediata della definizione (2.2.7). Si supponga $d_{KF}(X, Y) = 0$; ciò vuol dire, in particolare che, per ogni $n \in \mathbb{N}$ risulta

$$\mathbb{P}\left(|X - Y| > 1/n^2\right) \leq \frac{1}{n^2} \,,$$

onde, per il primo lemma di Borel–Cantelli,

$$\mathbb{P}\left(\limsup_{n\to+\infty} \left\{|X - Y| > \frac{1}{n^2}\right\}\right) = 0 \,,$$

o equivalentemente,

$$\mathbb{P}\left(\liminf_{n\to+\infty} \left\{|X - Y| \leq \frac{1}{n^2}\right\}\right) = 1 \,,$$

cioè $X = Y$ q.c., vale a dire $X = Y$ in L^0.

Per stabilire la diseguaglianza triangolare, siano X, Y e Z tre v.a. e si ponga $\alpha := d_{KF}(X,Y)$ e $\beta := d_{KF}(Y,Z)$. Per ogni $\varepsilon > 0$, esistono allora due insiemi misurabili $E_\alpha(\varepsilon)$ e $F_\beta(\varepsilon)$ con $\mathbb{P}(E_\alpha(\varepsilon)) \leq \alpha + \varepsilon/2$ e $\mathbb{P}(F_\beta(\varepsilon)) \leq \beta + \varepsilon/2$, tali che $|X - Y| \leq \alpha + \varepsilon/2$ in $E^c_\alpha(\varepsilon)$ e $|Y - Z| \leq \beta + \varepsilon/2$ in $F^c_\beta(\varepsilon)$. Dunque nell'insieme

$$E^c_\alpha(\varepsilon) \bigcap F^c_\beta(\varepsilon) = \Big(E_\alpha(\varepsilon) \bigcup F_\beta(\varepsilon)\Big)^c$$

è

$$|X - Z| \leq |X - Y| + |Y - Z| \leq \alpha + \beta + \varepsilon\,.$$

Ora $\mathbb{P}(E_\alpha(\varepsilon) \cup F_\beta(\varepsilon)) \leq \mathbb{P}(E_\alpha(\varepsilon)) + \mathbb{P}(F_\beta(\varepsilon)) \leq \alpha + \beta + \varepsilon$. In conclusione,

$$d_{KF}(X,Z) \leq \alpha + \beta + \varepsilon = d_{KF}(X,Y) + d_{KF}(Y,Z) + \varepsilon\,,$$

onde l'asserto in virtú dell'arbitrarietà di ε.

Si supponga ora che $X_n \to X$ in probabilità; allora, per ogni $\varepsilon > 0$, si ha definitivamente

$$\mathbb{P}(|X_n - X| > \varepsilon) \leq \varepsilon\,;$$

perciò, $d_{KF}(X_n, X) \to 0$.

Viceversa, si supponga che sia $\lim_{n\to+\infty} d_{KF}(X_n, X) = 0$. Si scelgano ora, arbitrariamente, $\varepsilon > 0$ e $\eta > 0$ e si scelga altresí $\delta < \varepsilon \wedge \eta = \min\{\varepsilon, \eta\}$. Per n sufficientemente grande, si ha $d_{KF}(X_n, X) < \delta$, che, a sua volta, implica $\mathbb{P}(|X_n - X| > \delta) < \delta$. Poiché $\eta > \delta$, si ha

$$\mathbb{P}(|X_n - X| > \eta) \leq \mathbb{P}(|X_n - X| > \delta) < \delta < \varepsilon\,,$$

cioè $X_n \to X$ in probabilità. □

Si osservi che da questo teorema scende come conseguenza l'unicità del limite in probabilità, e quindi, *a fortiori*, di quello quasi certo. La (2.2.7) non costituisce l'unica maniera di metrizzare la convergenza in probabilità su L^0. Esiste un gran numero di tali metriche; si vedrà negli esercizî come costruirle.

Si è visto sopra che la convergenza in probabilità deriva da una metrica; in genere, essa è, però, incompatibile con l'esistenza di una norma.

Teorema 2.2.11. *Dato lo spazio di probabilità sono equivalenti le proprietà:*

(a) *la convergenza in probabilità derivi da una norma su* $L^0 = L^0(\Omega, \mathcal{F}, \mathbb{P})$;

(b) Ω *è l'unione di un numero finito di punti* $\{\omega_i\}$ *con* $\mathbb{P}(\{\omega_i\}) > 0$.

Teorema 2.2.12. *Lo spazio* $L^0 = L^0(\Omega, \mathcal{F}, \mathbb{P})$ *è completo rispetto alla convergenza in probabilità; in altre parole,* (L^0, d_{KF}) *è uno spazio metrico completo.*

Dimostrazione. Sia (X_n) una successione di Cauchy in probabilità, vale a dire che, per ogni $\varepsilon > 0$, esiste un indice $\nu \in \mathbb{N}$ tale che sia

$$d_{KF}(X_n, X_m) < \varepsilon$$

se $m, n \geq \nu$. Come nel Teorema 2.2.6 si può mostrare che è possibile estrarre da (X_n) una successione $(X_{n(k)})_{k\in\mathbb{N}}$ che sia di Cauchy rispetto alla convergenza q.c.. Esiste, dunque, una v.a. X tale che $\lim_{k\to+\infty} X_{n(k)} = X$ q.c. e quindi anche in probabilità. Per mostrare che l'intera successione (X_n) converge a X in probabilità si consideri che

$$d_{KF}(X_n, X) \leq d_{KF}(X_n, X_{n(k)}) + d_{KF}(X_{n(k)}, X) < 2\,\varepsilon\,,$$

se n e $n(k)$ sono abbastanza grandi. □

2.3 La convergenza completa

Un tipo di convergenza che riveste particolare importanza in probabilità è la *convergenza completa*. È spesso importante, si ricordi il teorema integrale di de Moivre–Laplace, considerare, data una successione (X_n) di v.a. definite sullo stesso spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$, il limite della probabilità $\mathbb{P}(X_n \in\,]a, b])$ che la v.a. X_n assuma valori nell'intervallo $]a, b]$. Poiché, ricorrendo alla f.r. F_n della v.a. X_n, tale probabilità si scrive

$$\mathbb{P}(X_n \in\,]a, b]) = F_n(b) - F_n(a)\,,$$

è evidente l'interesse dello studio del limite della successione $(F_n(x))$, eventualmente con qualche restrizione sul punto x nel quale essa è calcolata.

Definizione 2.3.1. Si dice che la successione di f.r. (F_n) *converge completamente* alla f.r. F, se, per ogni punto di continuità x di F, si ha la convergenza puntuale di $(F_n(x))$ a $F(x)$. Se $C(F)$ indica l'insieme dei punti di continuità di F la convergenza completa si esprime nella forma

$$\forall\, x \in C(F) \qquad \lim_{n\to+\infty} F_n(x) = F(x)\,;$$

ciò che si indica con $F_n \xrightarrow{c} F$. ◇

La dizione di convergenza completa non è adottata universalmente; piú usata è l'espressione *convergenza debole* che, però, in queste lezioni è riservata ad un altro tipo di convergenza.

Esistono modi equivalenti di porre la convergenza completa; alcuni si incontreranno in queste lezioni, altri sono dati negli esercizî. Lo spazio delle f.r. è indicato da $\mathcal{D}$.

Teorema 2.3.1. *Se (F_n) è una successione di f.r. di $\mathcal{D}$ e F appartiene a $\mathcal{D}$, sono equivalenti le proprietà:*

(a) $F_n \xrightarrow{c} F$;

(b) *esiste un sottoinsieme denso D di $\mathbb{R}$ tale che $\lim_{n\to+\infty} F_n(x) = F(x)$ per ogni $x \in D$.*

Dimostrazione. (a) $\Longrightarrow$ (b) Basta prendere $D = C(F)$.
(b) $\Longrightarrow$ (a) Sia F continua in x_0, $x_0 \in C(F)$. Allora, comunque si fissi $\varepsilon > 0$, esiste $\delta = \delta(\varepsilon) > 0$ tale che

$$|F(x) - F(x_0)| < \frac{\varepsilon}{2}$$

se $|x-x_0| < \delta$. Poiché D è denso in $\mathbb{R}$, esistono due punti a e b di D che appartengono all'intervallo $]x_0 - \delta, x_0 + \delta[$ e tali che sia $a < x_0 < b$. Per ipotesi, esiste $n_0 \in \mathbb{N}$ tale che per ogni $n \geq n_0$, si abbia

$$|F_n(a) - F(a)| < \frac{\varepsilon}{2} \qquad \text{e} \qquad |F_n(b) - F(b)| < \frac{\varepsilon}{2}.$$

Allora, per ogni $n \geq n_0$, si ha

$$F_n(x_0) - F(x_0) \leq F_n(b) - F(x_0) < F(b) - F(x_0) + \frac{\varepsilon}{2} < \varepsilon$$

e

$$F_n(x_0) - F(x_0) \geq F_n(a) - F(x_0) > F(a) - F(x_0) - \frac{\varepsilon}{2} > -\varepsilon,$$

onde $|F_n(x_0) - F(x_0)| < \varepsilon$, se $n \geq n_0$. □

Non si deve credere che il limite di una successione convergente di f.r. coincida (tranne, eventualmente, per un insieme numerabile) con una f.r. come mostrano gli esempî che seguono. Nel trattare della convergenza completa di f.r. è, perciò, indispensabile assicurarsi che il limite sia anch'esso una f.r..

Esempio 2.3.1. Sia $F_n = \varepsilon_n (= \mathbf{1}_{[n,\infty[})$ $(n \in \mathbb{N})$ la f.r. di una v.a. che assume q.c. il valore n; allora $F_n(x) \to 0$ per ogni $x \in \mathbb{R}$ sicché la funzione identicamente nulla non è una f.r.. ■

Esempio 2.3.2. Se, per ogni $n \in \mathbb{N}$, F_n è la f.r. della legge uniforme in $(-n, n)$

$$F_n(x) := \frac{1}{2n} \int_{-\infty}^{x} \mathbf{1}_{(-n,n)}(t)\, \mathrm{d}t = \begin{cases} 0, & x \leq -n, \\ \dfrac{(x+n)}{2n}, & x \in [-n, n], \\ 1, & x \geq n, \end{cases}$$

allora $F_n(x) \to 1/2$ per ogni $x \in \mathbb{R}$. ■

Esempio 2.3.3. Si consideri la successione di v.a. definita, per ogni $n \in \mathbb{N}$, da

$$X_n = \begin{cases} \dfrac{1}{n}, & n \text{ dispari}, \\ 0, & n \text{ pari}. \end{cases}$$

Si consideri la successione delle corrispondenti f.r. (F_n). Per ogni $x > 0$, si ha $F_n(x) = \mathbb{P}(X_n \leq x) \to 1$. Perciò $F_n \xrightarrow{c} \varepsilon_0$. Nell'origine si ha

$$F_n(0) = \mathbb{P}(X_n \leq 0) = \begin{cases} 0, & \text{se } n \text{ è dispari}, \\ 1, & \text{se } n \text{ è pari}. \end{cases}$$

In questo caso la successione $(F_n(0))$ non ammette neanche limite. ■

Il seguente è un concetto piú debole della convergenza completa quando sia applicato alle f.r..

Definizione 2.3.2. Si dice che una successione (f_n) di funzioni reali a valori reali crescenti (f_n da $\mathbb{R}$ in $\mathbb{R}$, $x' < x'' \implies f_n(x') \leq f_n(x'')$) *converge debolmente* alla funzione (pure crescente) f se, per ogni $x \in C(f)$, è

$$\lim_{n\to+\infty} f_n(x) = f(x)\,.$$

Tale modo di convergenza sarà detto *convergenza debole*. ◇

Il seguente risultato è di frequente applicazione in Probabilità. Tuttavia, esso non è probabilistico in natura, sicché tutte le volte che lo si applicherà ad un insieme di f.r. occorrerà poi accertarsi che il limite di cui esso asserisce l'esistenza sia effettivamente una f.r..

Teorema 2.3.2. (Primo teorema di Helly). *Sia (f_n) una successione di funzioni crescenti ed uniformemente limitate, cioè $|f_n| \leq K$ per ogni $n \in \mathbb{N}$; esiste una successione estratta dalla data che converge debolmente ad una funzione g crescente, con $|g| \leq K$ e che può essere presa continua a destra.*

Dimostrazione. Supporremo, senza scapito di generalità, che ogni funzione f_n sia positiva e continua a destra. Sia $\mathbb{Q}$ l'insieme dei numeri razionali che si supporrà numerato, $\mathbb{Q} = \{q_1, q_2, \dots\}$. La succesione reale $(f_n(q_1))$ è limitata; esiste perciò una successione estratta $(f_{1n}(q_1))_{n\in\mathbb{N}}$ convergente ad un limite c_1. Anche la successione di numeri reali $(f_{1n}(q_2)$ è limitata ed ha perciò una sottosuccessione che converge ad un limite c_2; la si indichi con $(f_{2n}(q_2))$. La successione di funzioni (f_{2n}) converge tanto in q_1 quanto in q_2 e si ha

$$\lim_{n\to+\infty} f_{2n}(q_1) = c_1 \qquad \text{e} \qquad \lim_{n\to+\infty} f_{2n}(q_2) = c_2\,.$$

Procedendo in questa maniera, si costruiscono successioni $(f_{kn})_{n\in\mathbb{N}}$ tutte estratte da quella data e con la proprietà che $\lim_{n\to+\infty} f_{kn}(q_j) = c_j$ per ogni $j \leq k$. Si consideri la successione "diagonale" (f_{nn}) e la funzione $\varphi : \mathbb{Q} \to \mathbb{R}$ definita da $\varphi(q_j) := c_j$. Si ha cosí $f_{nn}(q) \to \varphi(q)$ per ogni $q \in \mathbb{Q}$. Si definisca $g : \mathbb{R} \to \mathbb{R}$ mediante $g(x) := \inf\{\varphi(q) : q \in \mathbb{Q}, q > x\}$; la funzione g cosí definita è crescente e continua a destra. Infatti, sia x_0 un arbitrario numero reale; esiste allora, per ogni $\varepsilon > 0$, $q \in \mathbb{Q}$ con $q > x_0$ tale che $\varphi(q) - g(x_0) < \varepsilon$. Si consideri ora $x \in]x_0, q[$; dalla definizione di g segue che $g(x) \leq \varphi(q)$, sicché, per ogni $x \in]x_0, q[$, è $0 \leq g(x) - g(x_0) \leq \varphi(q) - g(x_0) < \varepsilon$, ciò che prova la continuità a destra di g. Sia $x \in \mathbb{R}$; allora per ogni $q \in \mathbb{Q}$ con $q > x$, si ha, per ogni $q \in \mathbb{Q}$,

$$\limsup_{n\to+\infty} f_{nn}(x) \leq \limsup_{n\to+\infty} f_{nn}(q) = \varphi(q)\,,$$

sicché $\limsup_{n\to+\infty} f_{nn}(x) \leq g(x)$. Analogamente, si ha, per ogni $q \in \mathbb{Q}$ con $q < x$,

$$\liminf_{n\to+\infty} f_{nn}(x) \geq \liminf_{n\to+\infty} f_{nn}(q) = \varphi(q)\,,$$

e, dunque, se $y < x$ e $q \in \mathbb{Q}$ è tale che $q < x$,

$$\liminf_{n\to+\infty} f_{nn}(x) \geq \sup\{\varphi(q) : q \in \mathbb{Q}, q < x\} \geq g(y)\,,$$

e $\liminf_{n\to+\infty} f_{nn}(x) \geq \ell^- g(x)$. In particolare, se g è continua in x, si ha

$$\lim_{n\to+\infty} f_{nn}(x) = g(x)\,.$$

Poiché $0 \leq f \leq K$, è pure $0 \leq g \leq K$. □

Quando il teorema di Helly si applica ad una successione di f.r., il limite non è necessariamente una f.r. perché potrebbero non essere soddisfatte le proprietà di limite a $-\infty$ e a $+\infty$.

Oltre a quelli incontrati in precedenza vi è un altro tipo di convergenza per una successione di v.a..

Definizione 2.3.3. Si dice che una successione di v.a. (X_n) *converge in legge o in distribuzione* alla v.a. X se, essendo F_n la f.r. di X_n e F la f.r. di X, la successione (F_n) converge completamente a F; tale convergenza si indica con $X_n \xrightarrow{\mathcal{L}} X$. $\diamond$

Per la convergenza in legge di una successione di v.a. non occorre che tutte le v.a. della successione siano definite sopra il medesimo spazio di probabilità. La convergenza in legge non ngode delle proprietà alle quali si è soliti pensare trattando di limiti; si vedrà che il limite in legge non è necessariamente unico, che il limite di una somma non sempre è eguale alla somma dei limiti e cosí via. Ovviamente se si considera la convergenza in legge in relazione ad altri tipi di convergenza, allora si richiederà che tutte le v.a. siano definite sullo stesso spazio di probabilità.

Teorema 2.3.3. (La convergenza in probabilità implica la convergenza in legge). *Se (X_n) è una successione di v.a. sullo spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ che converge in probabilità alla v.a. X, allora vi converge anche in legge,*

$$X_n \xrightarrow[n\to+\infty]{\mathbb{P}} X \Longrightarrow X_n \xrightarrow[n\to+\infty]{\mathcal{L}} X\,.$$

Dimostrazione. Sia $x' < x < x''$. Poiché

$$\begin{aligned}\{X \le x'\} &= \left(\{X_n \le x\} \cap \{X \le x'\}\right) \bigcup \left(\{X_n > x\} \cap \{X \le x'\}\right) \\ &\subseteq \{X_n \le x\} \bigcup \left(\{X_n > x\} \cap \{X \le x'\}\right)\,,\end{aligned}$$

si ha

$$F(x') \le F_n(x) + \mathbb{P}\left(X_n > x, X \le x'\right)\,.$$

Poiché $X_n \xrightarrow[n\to+\infty]{\mathbb{P}} X$ e $x - x' > 0$, si ha

$$\mathbb{P}\left(X_n > x, X \le x'\right) \le \mathbb{P}\left(|X_n - X| > x - x'\right) \xrightarrow[n\to+\infty]{} 0\,.$$

Perciò

$$F(x') \le \liminf_{n\to+\infty} F_n(x)\,.$$

Scambiando i ruoli di X_n e X, si ha

$$F_n(x) \le F(x'') + \mathbb{P}\left(X > x'', X_n \le x\right)$$

e

$$\limsup_{n\to+\infty} F_n(x) \le F(x'')\,.$$

Perciò

$$F(x') \le \liminf_{n\to+\infty} F_n(x) \le \limsup_{n\to+\infty} F_n(x) \le F(x'')\,. \tag{2.3.1}$$

Si supponga ora che x sia un punto di continuità per F, $x \in C(F)$, e si facciano tendere x' e x'' a x, x' crescendo e x'' decrescendo, $x' \uparrow x$, $x'' \downarrow x$, ottenendo cosídalla (2.3.1)

$$\lim_{n \to +\infty} F_n(x) = F(x) ,$$

cioè l'asserto. □

Esempio 2.3.4. (La convergenza in legge non implica quella in probabilità). Sia X una v.a. che assume i valori 0 e 1, entrambi con probabilità eguale a 1/2 e sia Y la v.a. definita da $Y := 1 - X$. Evidentemente, X e Y hanno la stessa f.r. $F = (\varepsilon_0 + \varepsilon_1)/2$. Si consideri la successione (X_n) con $X_n = Y$; essa converge in legge a X, ma poiché $|X_n - X| = |Y - X| = 1$, non vi converge in probabilità. ■

L'ultimo esempio stabilisce inoltre che il limite in legge non è unico perché la successione (X_n) lí considerata converge in legge sia a X sia a Y.

Teorema 2.3.4. *Per una successione di v.a. (X_n) sono equivalenti le condizioni:*

(a) (X_n) *converge in probabilità alla v.a. $X = a$ q.c.;*

(b) (X_n) *converge in legge alla v.a. $X = a$ q.c..*

Dimostrazione. L'implicazione (a) $\Longrightarrow$ (b) è contenuta nel Teorema 2.3.3.
(b) $\Longrightarrow$ (a) Se $X \to a$ in legge, si ha $\lim_{n \to +\infty} F_n(x) = \varepsilon_a(x)$ per ogni $x \neq a$ e quindi, per ogni $\varepsilon > 0$,

$$\mathbb{P}(X_n - a < -\varepsilon) = \mathbb{P}(X_n < a - \varepsilon) \leq F_n(a - \varepsilon) ,$$
$$\mathbb{P}(X_n - a > \varepsilon) = \mathbb{P}(X_n > a + \varepsilon) = 1 - F_n(a + \varepsilon) ,$$

relazioni dalle quali scende immediatamente $\mathbb{P}(|X_n - a| > \varepsilon) \to 0$. □

I risultati sulle relazioni tra i varî tipi di convergenza sono riassunti nella Figura 3.1, nella quale le frecce indicano le implicazioni.

È notevole e spesso utile il seguente

Teorema 2.3.5. (Skorohod). *Siano F_n $(n \in \mathbb{N})$ e F f.r. e si supponga che (F_n) converga completamente a F. Esistono allora uno spazio di probabilità $(\Omega, \mathcal{F}, \mu)$ e v.a. X_n $(n \in \mathbb{N})$ e X su $(\Omega, \mathcal{F}, \mu)$ tali che*

(a) *per ogni $n \in \mathbb{N}$, $F_{X_n} = F_n$ e $F_X = F$;*

(b) $\lim_{n \to +\infty} X_n = X$ *μ–q.c..*

Dimostrazione. Si ponga $\Omega =]0, 1[$, $\mathcal{F} = \mathcal{B}(]0, 1[)$ e sia μ la restrizione della misura di Lebesgue a(i boreliani di) $]0, 1[$. Per $\omega \in]0, 1[$, si ponga

$$X_n(\omega) := \inf\{x \in \mathbb{R} : \omega \leq F_n(x)\} \quad \text{e} \quad X(\omega) := \inf\{x \in \mathbb{R} : \omega \leq F(x)\} .$$

Risulta, allora, $\omega \leq F_n(x)$ se, e solo se, $X_n(\omega) \leq x$ e perciò

$$\mu(\{\omega : X_n(\omega) \leq x\}) = \mu(\{\omega : \omega \leq F_n(x)\}) = F_n(x) ,$$

sicché F_n è la f.r. della v.a. X_n; similmente si dimostra che F è la f.r. di X. Fissato $\omega \in]0,1[$ e dato $\varepsilon > 0$, si scelga un punto x di continuità per F, $x \in C(F)$ in modo che $X(\omega) - \varepsilon < x < X(\omega)$. Allora, è anche $F(x) < \omega$. Ora, $F_n(x) \to F(x)$ implica, per n sufficientemente grande, che sia $F_n(x) < \omega$, onde

$$X(\omega) - \varepsilon < x < X_n(\omega)\,;$$

perciò $\liminf_{n\to+\infty} X_n(\omega) \geq X(\omega)$. Se $\omega' > \omega$, si scelga $x' \in C(F)$ tale che $X(\omega') < x' < X(\omega') + \varepsilon$; di qui $\omega < \omega' < F(x')$ e, perciò, per n sufficientemente grande, è $\omega < F_n(x')$, da cui scende

$$X_n(\omega) < x' < X(\omega') + \varepsilon\,.$$

Quindi, per $\omega' > \omega$, si ha $\limsup_{n\to+\infty} X_n(\omega) \leq X(\omega')$. Dunque, $X_n(\omega) \to X(\omega)$ se X è continua in ω. È però facile vedere che X è crescente in $]0,1[$ sicché ha, al piú, un'infinità numerabile di punti di discontinuità; inoltre, l'insieme dei punti di discontinuità di X ha misura nulla rispetto alla misura di Lebesgue. □

La dimostrazione appena data del teorema di Skorohod usa in maniera essenziale l'ordinamento dei reali. Benché il teorema sia ancora valido in condizioni di maggiore generalità, la dimostrazione è, in quei casi, assai piú complessa.

2.4 Le convergenze vaga e stretta

Definizione 2.4.1. Siano F_n e F $\quad (n \in \mathbb{N})$ f.r. e siano $\mathbb{P}_n$ e $\mathbb{P}$ le misure di probabilità di Stieltjes su $(\mathbb{R}, \mathcal{B})$ generate da F_n e da F, $\mathbb{P}_n := \mu_{F_n}$ e $\mathbb{P} := \mu_F$. Se $F_n \xrightarrow[n\to+\infty]{\mathrm{c}} F$, si dirà che $\mathbb{P}_n$ converge completamente a $\mathbb{P}$ ($\mathbb{P}_n \xrightarrow[n\to+\infty]{\mathrm{c}} \mathbb{P}$). ◇

Teorema 2.4.1. *Sia $\varphi : \mathbb{R} \to \mathbb{R}$ misurabile e si supponga che l'insieme $D(\varphi)$ dei punti di discontinuità di φ sia misurabile, $D(\varphi) \in \mathcal{B}$.*

Se $\mathbb{P}\,(D(\varphi)) = 0$ e $\mathbb{P}_n \xrightarrow[n\to+\infty]{c} \mathbb{P}$, allora

$$\mathbb{P}_n \circ \varphi^{-1} \xrightarrow[n\to+\infty]{c} \mathbb{P} \circ \varphi^{-1}\,.$$

Dimostrazione. Siano $(\Omega, \mathcal{F}, \mu)$, X_n $(n \in \mathbb{N})$ e X come nel teorema di Skorohod. Si osservi che la legge della v.a. $\varphi \circ X$ è $\mathbb{P} \circ \varphi^{-1}$; infatti, per ogni boreliano A, è

$$\mu\left((\varphi \circ X) \in A\right) = \mu\left(X \in \varphi^{-1}(A)\right) = \mathbb{P}\left(\varphi^{-1}(A)\right)\,.$$

Similmente $\varphi \circ X_n^{-1}$ ha legge $\mathbb{P}_n \circ \varphi^{-1}$.

Se $X(\omega)$ non appartiene a $D(\varphi)$, da $X_n(\omega) \to X(\omega)$ scende

$$\varphi[X_n(\omega)] \xrightarrow[n\to+\infty]{} \varphi[X(\omega)]$$

e, poiché

$$\mu(\{\omega : X(\omega) \in D(\varphi)\}) = \mathbb{P}\,(D(\varphi)) = 0\,,$$

si ha $\varphi \circ X_n \xrightarrow[n\to+\infty]{} \varphi \circ X$ μ–q.c.. Di conseguenza, si ha anche $\varphi \circ X_n \xrightarrow[n\to+\infty]{} \varphi \circ X$ in legge. Allora, la convergenza in legge di $\varphi \circ X_n$ a $\varphi \circ X$ equivale a dire che $\mathbb{P}_n \circ \varphi^{-1} \xrightarrow[n\to+\infty]{} \mathbb{P} \circ \varphi^{-1}$. □

Il risultato che segue è essenziale per stabilire l'unicità del limite nelle convergenze che saranno introdotte nella successiva definizione 2.4.2.

Teorema 2.4.2. *Se due misure finite μ e ν su $(\mathbb{R}, \mathcal{B})$ sono tali che $\int f \,\mathrm{d}\mu = \int f \,\mathrm{d}\nu$ per ogni funzione continua e con il supporto compatto, $f \in C_c = C_c(\mathbb{R})$, allora esse sono eguali, $\mu = \nu$.*

Dimostrazione. Poiché la famiglia degli intervalli chiusi genera la tribú di Borel ed è stabile rispetto all'intersezione, basta far vedere che $\mu([a,b]) = \nu([a,b])$ per ogni coppia di numeri reali a, b con $a < b$. Esiste una successione $(f_n) \subseteq C_c$ tale che $f_n \to \mathbf{1}_{[a,b]}$ q.o.; una tale successione è, per esempio, data da

$$f_n(x) := \begin{cases} 0, & x \notin \left[a - \dfrac{1}{n}, b + \dfrac{1}{n}\right] , \\ n\left(x - a + \dfrac{1}{n}\right), & x \in [a - 1/n, a] , \\ 1, & x \in [a, b] , \\ n\left(b + \dfrac{1}{n} - x\right), & x \in \left[b, b + \dfrac{1}{n}\right] . \end{cases}$$

Poiché, per ogni $n \in \mathbb{N}$, $0 \le f_n \le f_1$ e poiché f_1 è integrabile sia rispetto a μ sia rispetto a ν, il teorema di convergenza dominata applicato due volte dà

$$\begin{aligned} \mu([a,b]) &= \int \mathbf{1}_{[a,b]} \,\mathrm{d}\mu = \lim_{n\to+\infty} \int f_n \,\mathrm{d}\mu \\ &= \lim_{n\to+\infty} \int f_n \,\mathrm{d}\nu = \int \mathbf{1}_{[a,b]} \,\mathrm{d}\nu = \nu([a,b]) , \end{aligned}$$

che stabilisce l'eguaglianza delle misure μ e ν. □

Definizione 2.4.2. Siano $\mu_n \quad (n \in \mathbb{N})$ e μ misure su $(\mathbb{R}, \mathcal{B})$ finite; si dice che (μ_n) *converge vagamente* a μ, e si scrive $\mu_n \xrightarrow[n\to+\infty]{v} \mu$, se per ogni funzione f continua e con il supporto compatto, $f \in C_c(\mathbb{R})$, si ha

$$\lim_{n\to+\infty} \int f \,\mathrm{d}\mu_n = \int f \,\mathrm{d}\mu .$$

Si dice invece che (μ_n) *converge strettamente* a μ, e si scrive $\mu_n \xrightarrow[n\to+\infty]{s} \mu$, se, per ogni funzione f continua e limitata, $f \in C_b(\mathbb{R})$, si ha

$$\lim_{n\to+\infty} \int f \,\mathrm{d}\mu_n = \int f \,\mathrm{d}\mu .$$

Tali modi di convergenza si diranno *convergenza vaga* o *stretta*, rispettivamente. ◇

Poiché $C_c \subseteq C_b$, la convergenza stretta implica la convergenza vaga. Il Teorema 2.4.2 assicura l'unicità del limite vago (e, di conseguenza, di quello stretto).

Teorema 2.4.3. *Per una successione (μ_n) di misure finite su $(\mathbb{R}, \mathcal{B})$ sono equivalenti le proprietà:*

(a) (μ_n) *converge strettamente a* μ*;*

(b) (μ_n) *converge vagamente a* μ *ed inoltre si ha* $\lim_{n\to+\infty} \mu_n(\mathbb{R}) = \mu(\mathbb{R})$.

Dimostrazione. (a) $\Longrightarrow$ (b) Basta dimostrare il secondo asserto di (b). Poiché la funzione identicamente eguale a 1 è continua e limitata, $1 \in C_b$, è

$$\mu_n(\mathbb{R}) = \int 1 \,\mathrm{d}\mu_n \to \int 1 \,\mathrm{d}\mu = \mu(\mathbb{R}) \,.$$

(b) $\Longrightarrow$ (a) Alla luce della definizione di integrale non è restrittivo dare la dimostrazione per le sole funzioni positive limitate superiormente da 1: $f \in C_b^+$, $0 \le f \le 1$. Si può trovare una successione $(f_r) \subseteq C_c$ tale che $f_r \uparrow f$. Per esempio, si consideri la funzione $g_r \in C_c$ definita da

$$g_r(x) := \begin{cases} 0, & x \notin [-(r+1), r+1] \,, \\ x+r+1, & x \in [-(r+1), -r] \,, \\ 1, & x \in [-r, r] \,, \\ r+1-x, & x \in [r, r+1] \,. \end{cases}$$

La funzione $f_r := g_r \wedge f$ soddisfà al requisito di sopra. Allora, per ogni $r \in \mathbb{N}$,

$$\int f_r \,\mathrm{d}\mu = \lim_{n\to+\infty} \int f_r \,\mathrm{d}\mu_n = \liminf_{n\to+\infty} \int f_r \,\mathrm{d}\mu_n \le \liminf_{n\to+\infty} \int f \,\mathrm{d}\mu_n \,,$$

onde

$$\int f \,\mathrm{d}\mu = \sup_{r\in\mathbb{N}} \int f_r \,\mathrm{d}\mu \le \liminf_{n\to+\infty} \int f \,\mathrm{d}\mu_n \,. \tag{2.4.1}$$

Applicando quest'ultima diseguaglianza alla funzione $1-f$, che pure è continua ed assume valori in $[0,1]$, si ottiene

$$\int (1-f) \,\mathrm{d}\mu \le \liminf_{n\to+\infty} \int (1-f) \,\mathrm{d}\mu_n \,.$$

D'altro canto, ancora per la (2.4.1)

$$\begin{aligned} \int f \,\mathrm{d}\mu = \mu(\mathbb{R}) - \int (1-f) \,\mathrm{d}\mu &= \limsup_{n\to+\infty} \mu_n(\mathbb{R}) - \int (1-f) \,\mathrm{d}\mu \\ &\ge \limsup_{n\to+\infty} \mu_n(\mathbb{R}) - \liminf_{n\to+\infty} \int (1-f) \,\mathrm{d}\mu_n \\ &\ge \limsup_{n\to+\infty} \left(\mu_n(\mathbb{R}) - \int (1-f) \,\mathrm{d}\mu_n \right) = \limsup_{n\to+\infty} \int f \,\mathrm{d}\mu_n \,, \end{aligned}$$

che insieme alla (2.4.1) fornisce l'asserto. □

Cosí, per le misure di probabilità, la convergenza vaga e quella stretta coincidono.

Se $A \in \mathcal{B}$ si dirà *frontiera* di A l'insieme $\partial A := \overline{A} - A^\circ$ (ove $\overline{A}$ e A° sono rispettivamente la chiusura e l'interno di A). Se $\mathbb{P}$ è una probabilità su $(\mathbb{R}, \mathcal{B})$ si dirà che $A \in \mathcal{B}$ è un *insieme di* $(\mathbb{P})$*–continuità* se $\mathbb{P}(\partial A) = 0$. Il teorema che segue dà una formulazione di convergenza completa che non dipende dalle f.r..

Teorema 2.4.4. *Siano* $\mathbb{P}_n \quad (n \in \mathbb{N})$ *e* $\mathbb{P}$ *misure di probabilità su* $(\mathbb{R}, \mathcal{B})$ *e siano* F_n *e* F *le corrispondenti f.r.. Sono allora equivalenti le condizioni:*

(a) $F_n \xrightarrow[n\to+\infty]{c} F$;

(b) $(\mathbb{P}_n)$ *converge strettamente a* $\mathbb{P}$;

(c) $\lim_{n\to+\infty} \mathbb{P}_n(A) = \mathbb{P}(A)$ *per ogni insieme di* $\mathbb{P}$*–continuità.*

Dimostrazione. (a) $\Longrightarrow$ (b) In virtú del teorema di Skorohod, esistono uno spazio di probabilità $(\Omega, \mathcal{F}, \mu)$ e v.a. X_n e X definite su di esso, tali che le leggi di X_n e di X siano, rispettivamente, $\mathbb{P}_n$ e $\mathbb{P}$. Sia, inoltre, $\varphi : \mathbb{R} \to \mathbb{R}$ una funzione misurabile, limitata e tale che l'insieme $D(\varphi)$ dei suoi punti di discontinuità sia misurabile e di probabilità nulla, $\mathbb{P}\,(D(\varphi)) = 0$. Poiché $\mu\,(X \in D(\varphi)) = (\mu \circ X^{-1})(D(\varphi)) = \mathbb{P}(D(\varphi)) = 0$, si ha $\varphi \circ X_n \xrightarrow[n\to+\infty]{} \varphi \circ X$ μ–q.c.. Il teorema di convergenza dominata dà

$$\int_{\mathbb{R}} \varphi \, \mathrm{d}\,\mathbb{P}_n = \int_{\mathbb{R}} \varphi \, \mathrm{d}(\mu \circ X_n^{-1}) = \int_{\Omega} \varphi \circ X_n \, \mathrm{d}\mu \xrightarrow[n\to+\infty]{} \int_{\Omega} \varphi \circ X \, \mathrm{d}\mu = \int_{\mathbb{R}} \varphi \, \mathrm{d}\,\mathbb{P} \,.$$

Ciò mostra, in particolare, che (a) implica (b), poiché, se φ, oltre che essere limitata, è anche continua, è $D(\varphi) = \emptyset$ e dunque, *a fortiori*, $\mathbb{P}(D(\varphi)) = 0$.
(a) $\Longrightarrow$ (c) Preso $\varphi = \mathbf{1}_A$ con $A \in \mathcal{B}$ insieme di $\mathbb{P}$–continuità, risulta $D(\varphi) = \partial A$, sicché quanto visto sopra dà l'asserto.
(c) $\Longrightarrow$ (a) Considerato l'insieme $]-\infty, x]$, la sua frontiera è costituita dal singoletto $\{x\}$, sicché esso è di continuità per $\mathbb{P}$ se, e solo se, $\mathbb{P}(\{x\}) = 0$, vale a dire se, e solo se, $x \in C(F)$. Perciò, se tale condizione è verificata, è

$$F_n(x) = \mathbb{P}_n(]-\infty, x]) \xrightarrow[n\to+\infty]{} \mathbb{P}(]-\infty, x]) = F(x) \,.$$

Per completare la dimostrazione basta far vedere che (b) $\Longrightarrow$ (a). Si supponga allora che $(\mathbb{P}_n)$ converga strettamente a $\mathbb{P}$. Non è difficile stabilire che l'insieme degli atomi di $\mathbb{P}$, cioè l'insieme $A \subseteq \mathbb{R}$ di punti con probabilità strettamente positiva, è, al piú numerabile: se

$$A := \{x \in \mathbb{R} : \mathbb{P}(\{x\}) > 0\} \,,$$

è $\operatorname{card}(A) \leq \aleph_0$. Si prenda $D = A^c$; evidentemente, $\overline{D} = \mathbb{R}$. Siano a e b punti di D con $a < b$ e sia $f = \mathbf{1}_{]a,b]}$. Dato $\varepsilon > 0$, esiste $\delta = \delta(\varepsilon) > 0$ tale che $a + \delta < b - \delta$ e $\mathbb{P}(I) < \varepsilon$, ove

$$I :=]a - \delta, a + \delta[\bigcup]b - \delta, b + \delta[\;.$$

Si definiscano due funzioni $f_1, f_2 : \mathbb{R} \to [0, 1]$ di C_c mediante

$$f_1(x) := \begin{cases} 0\,, & x \in]-\infty, a] \cup [b, +\infty[\;, \\ \dfrac{x-a}{\delta}\,, & x \in]a, a+\delta] \;, \\ 1\,, & x \in]a+\delta, b-\delta] \;, \\ \dfrac{b-x}{\delta}\,, & x \in [b-\delta, b] \;. \end{cases}$$

e

$$
f_2(x) := \begin{cases} 0\,, & x \in \,]-\infty, a-\delta] \cup [b+\delta, +\infty[\;, \\ \dfrac{x-a+\delta}{\delta}\,, & x \in \,]a-\delta, a]\;, \\ 1\,, & x \in \,]a, b]\;, \\ \dfrac{b+\delta-x}{\delta}\,, & x \in [b, b+\delta]\;. \end{cases}
$$

Evidentemente, f_1 e f_2 sono in C_c e si ha $0 \le f_1 \le f \le f_2 \le f_1 + 1$. Allora

$$
\int f_1 \, d\mathbb{P} \le \int f \, d\mathbb{P} \le \int f_2 \, d\mathbb{P} \qquad \text{e} \qquad \int f_1 \, d\mathbb{P}_n \le \int f \, d\mathbb{P}_n \le \int f_2 \, d\mathbb{P}_n
$$

per ogni $n \in \mathbb{N}$. Poiché $\lim_{n\to+\infty} \int f_j \, d\mathbb{P}_n = \int f_j \, d\mathbb{P}$ $(j = 1, 2)$ si ha, per n sufficientemente grande,

$$
\left| \int f \, d\mathbb{P}_n - \int f \, d\mathbb{P} \right| \le \int f_2 \, d\mathbb{P} - \int f_1 \, d\mathbb{P} + \varepsilon < P(I) + \varepsilon < 2\varepsilon\,,
$$

sicché

$$
\lim_{n\to+\infty} (F_n(b) - F_n(a)) = \lim_{n\to+\infty} \int f \, d\mathbb{P}_n = \int f \, d\mathbb{P} = F(b) - F(a)\,. \tag{2.4.2}
$$

Poiché $D = A^c = C(F)$, la (2.4.2) vale per ogni coppia a, b di punti di $C(F)$ con $a < b$. Per mostrare che $\lim_{n\to+\infty} F_n(x) = F(x)$ per ogni $x \in C(F)$ si scelgano a e b in $C(F)$ in modo da avere $a < x < b$, $F(a) < \varepsilon$ e $F(b) > 1 - \varepsilon$. Allora,

$$
|F_n(x) - F(x)| \le |F_n(x) - F_n(a) - F(x) + F(a)| + F_n(a) + F(a)\,,
$$

mentre dalla (2.4.2), con $b = x$, scende

$$
\limsup_{n\to+\infty} |F_n(x) - F(x)| \le F(a) + \limsup_{n\to+\infty} F_n(a)\,. \tag{2.4.3}
$$

Ora,

$$
\limsup_{n\to+\infty} F_n(a) \le \limsup_{n\to+\infty} (1 - F_n(b) + F_n(a)) \le 1 - F(b) + F(a) < 2\varepsilon\,,
$$

onde, sostituendo nella (2.4.3), si ha $\limsup_{n\to+\infty} |F_n(x) - F(x)| = 0$. □

L'ultimo teorema consente di formulare in maniera differente, ma equivalente, la convergenza in legge di v.a.: una successione (X_n) di v.a. definite sopra il medesimo spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ converge in legge alla v.a. X se, e solo se, per ogni funzione f di $C_b(\mathbb{R})$, o di $C_c(\mathbb{R})$, si ha

$$
\mathbb{E}(f \circ X_n) \xrightarrow[n\to+\infty]{} \mathbb{E}(f \circ X)\,.
$$

Infatti, per il teorema del cambio di variabile (3.8.2), si ha

$$
\mathbb{E}(f \circ X_n) = \int_\Omega f \circ X_n \, d\mu = \int_\mathbb{R} f \, d(\mu\, X_n^{-1}) = \int_\mathbb{R} f \, d\mathbb{P}_n\,,
$$

dove $\mathbb{P}_n$ è la legge di X_n. Analogamente si procede per $\mathbb{E}(f \circ X)$.

Risultati riguardanti la convergenza degli integrali si hanno anche nel caso della convergenza debole anziché completa; essi torneranno utili nel seguito.

Lemma 2.4.1. *Sia $(\varphi_n)_{n\in\mathbb{Z}_+}$ una successione di funzioni φ_n da $\mathbb{R}$ in $\mathbb{R}$ ($n \in \mathbb{Z}_+$) crescenti che converge debolmente a φ. Si ponga, per $a \in]0, +\infty[$,*

$$\begin{aligned}\delta\varphi &:= \varphi(+\infty) - \varphi(-\infty)\,,\\ \delta\varphi_n &:= \varphi_n(+\infty) - \varphi_n(-\infty)\,,\\ \delta\varphi_n(a) &:= \varphi_n(a) - \varphi_n(-a)\,.\end{aligned}$$

Allora

$$\delta\varphi \leq \liminf_{n\to+\infty} \delta\varphi_n\,. \tag{2.4.4}$$

Inoltre, se $\delta\varphi_n(a)$ converge uniformemente a $\delta\varphi_n$ al tendere di a a $+\infty$, vale a dire se

$$\lim_{a\to+\infty} \sup_{n\in\mathbb{N}} (\delta\varphi_n - \delta\varphi_n(a)) = 0\,,$$

si ha

$$\lim_{n\to+\infty} \varphi_n(\pm\infty) = \varphi(\pm\infty)\,.$$

Dimostrazione. Nelle diseguaglianze $\varphi_n(-\infty) \leq \varphi_n(x) \leq \varphi_n(+\infty)$ si prenda x in $C(\varphi)$ e si passi al limite per $n \to \infty$ per ottenere

$$\limsup_{n\to+\infty} \varphi_n(-\infty) \leq \varphi(x) \leq \limsup_{n\to+\infty} \varphi_n(+\infty)$$

e

$$\liminf_{n\to+\infty} \varphi_n(-\infty) \leq \varphi(x) \leq \liminf_{n\to+\infty} \varphi_n(+\infty)$$

che insieme danno la (2.4.4).

Fissato, arbitrariamente, $\varepsilon > 0$, si scelga $\alpha = \alpha(\varepsilon) > 0$ in modo che

$$\delta\varphi_n - \delta\varphi_n(a) < \varepsilon$$

per ogni $n \in \mathbb{N}$ se $a \geq \alpha$. Allora, se tanto a quanto $-a$ sono punti di continuità per φ, è

$$\limsup_{n\to+\infty} \delta\varphi_n \leq \delta\varphi(a) + \varepsilon < +\infty\,.$$

In virtú della (2.4.4), ciò implica $\delta\varphi < +\infty$; poiché ε è arbitrario, risulta

$$\limsup_{n\to+\infty} \delta\varphi_n \leq \delta\varphi\,.$$

Questa diseguaglianza e la (2.4.4) danno

$$\lim_{n\to+\infty} \delta\varphi_n = \delta\varphi\,.$$

Ora, le diseguaglianze di sopra danno

$$\begin{aligned}\limsup_{n\to+\infty} \varphi_n(+\infty) &= \limsup_{n\to+\infty} (\delta\varphi_n + \varphi_n(-\infty))\\ &\leq \delta\varphi + \limsup_{n\to+\infty} \varphi_n(-\infty) \leq \varphi(+\infty) \leq \liminf_{n\to+\infty} \varphi_n(+\infty)\end{aligned}$$

cioè $\lim_{n\to+\infty} \varphi_n(+\infty) = \varphi(+\infty)$. Di conseguenza si ha anche

$$\lim_{n\to+\infty} \varphi_n(-\infty) = \varphi(-\infty)\,,$$

con il che la dimostrazione è conclusa. □

Teorema 2.4.5. (Secondo teorema di Helly). *Sia (F_n) una successione uniformemente limitata di funzioni crescenti che converge debolmente a F. Se a e b sono punti di continuità per F $(a, b \in C(F))$, per ogni funzione continua $g \in C([a, b])$ è*

$$\lim_{n\to+\infty} \int_a^b g \,\mathrm{d}F_n = \int_a^b g \,\mathrm{d}F \,.$$

Dimostrazione. Dall'uniforme continuità di g in $[a, b]$ scende che, per ogni $\varepsilon > 0$, esistono punti

$$x_0, x_1, \dots, x_m$$

con $a_0 = x_0 < x_1 < \dots < x_m = b$ tali che

$$|g(x) - g(x_j)| < \varepsilon$$

per ogni $x \in [x_j, x_{j+1}]$ $\quad (j = 0, 1, \dots, m - 1)$. Si definisca

$$g_\varepsilon := g(x_0)\, \mathbf{1}_{[x_0,x_1]} + \sum_{j=1}^{m-1} g(x_j)\, \mathbf{1}_{]x_j,x_{j+1}]} \,.$$

Non è, inoltre, restrittivo supporre che $x_0, x_1, \dots, x_m$ siano punti di continuità per F. Per ogni $x \in [a, b]$, si ha $|g(x) - g_\varepsilon(x)| < \varepsilon$. Si consideri la norma della convergenza uniforme $\|g\| := \max\{|g(x)| : x \in [a, b]\}$; si può, allora, determinare $n_0 = n_0(\varepsilon) \in \mathbb{N}$ in modo che per ogni $n \geq n_0$ sia

$$|F_n(x_j) - F(x_j)| \leq \frac{\varepsilon}{\|g\| m} \qquad (j = 0, 1, \dots, m)\,.$$

Allora,

$$\begin{aligned}\left|\int_a^b g \,\mathrm{d}F - \int_a^b g \,\mathrm{d}F_n\right| &\leq \left|\int_a^b g \,\mathrm{d}F - \int_a^b g_\varepsilon \,\mathrm{d}F\right| \\ &\quad + \left|\int_a^b g_\varepsilon \,\mathrm{d}F - \int_a^b g_\varepsilon \,\mathrm{d}F_n\right| + \left|\int_a^b g_\varepsilon \,\mathrm{d}F_n - \int_a^b g \,\mathrm{d}F_n\right| .\end{aligned}$$

Si supponga che sia $|F| \leq K$ e $|F_n| \leq K$ per ogni $n \in \mathbb{N}$. Allora,

$$\left|\int_a^b g \,\mathrm{d}F - \int_a^b g_\varepsilon \,\mathrm{d}F\right| \leq \varepsilon\, \mu_F(]a, b]) \leq 2\, \varepsilon\, K \,,$$

mentre

$$\left|\int_a^b g_\varepsilon \,\mathrm{d}F_n - \int_a^b g \,\mathrm{d}F_n\right| \leq \varepsilon\, \mu_{F_n}(]a, b]) \leq 2\, \varepsilon\, K \,.$$

Inoltre, per $n \geq n_0$, si ha

$$\begin{aligned}
&\left|\int_a^b g_\varepsilon \,\mathrm{d}F - \int_a^b g_\varepsilon \,\mathrm{d}F_n\right| \\
&\quad = \left|\sum_{j=0}^{m-1} g(x_j)\,(F(x_{j+1}) - F(x_j)) - \sum_{j=0}^{m-1} g(x_j)\,(F_n(x_{j+1}) - F_n(x_j))\right| \\
&\quad = \left|\sum_{j=0}^{m-1} g(x_j)\,(F(x_{j+1}) - F_n(x_{j+1})) - \sum_{j=0}^{m-1} g(x_j)\,(F(x_j) - F_n(x_j))\right| \\
&\quad \leq 2\|g\|\, m \frac{\varepsilon}{\|g\|\, m} = 2\,\varepsilon\,,
\end{aligned}$$

sicché, per $n \geq n_0$, si ha

$$\left|\int_a^b g \,\mathrm{d}F - \int_a^b g \,\mathrm{d}F_n\right| \leq (4\,K + 2)\,\varepsilon\,,$$

che dà l'asserto. □

Teorema 2.4.6. *Se (F_n) è una successione uniformemente limitata di funzioni crescenti che converge debolmente a F e se $g \in C(\mathbb{R})$ è una funzione continua tale che $\lim_{|x|\to+\infty} g(x) = 0$, allora*

$$\lim_{n\to+\infty} \int_{\mathbb{R}} g \,\mathrm{d}F_n = \int_{\mathbb{R}} g \,\mathrm{d}F\,.$$

Dimostrazione. È ovvio che anche la funzione F è crescente e limitata. Si può supporre che sia $g \geq 0$ (altrimenti basta dimostrare il teorema separatamente per g^+ e per g^-). Allora se $a \in C(F)$ e $x \in \overline{\mathbb{R}}$, si definisca

$$\varphi_n(x) := \int_a^x g \,\mathrm{d}F_n \qquad \text{e} \qquad \varphi(x) := \int_a^x g \,\mathrm{d}F\,.$$

Per il Teorema 2.4.5, $\varphi_n(x)$ converge a $\varphi(x)$ per ogni $x \in C(F)$. Si scelga, ora, $a > 0$ sufficientemente grande perché sia $g(y) < \varepsilon$, per ogni y con $|y| > a$. Allora,

$$\begin{aligned}
\delta\varphi_n - \delta\varphi_n(y) &= \int_a^{+\infty} g \,\mathrm{d}F_n - \int_a^{-\infty} g \,\mathrm{d}F_n - \int_a^y g \,\mathrm{d}F_n + \int_a^{-y} g \,\mathrm{d}F_n \\
&= \int_y^{+\infty} g \,\mathrm{d}F_n + \int_{-\infty}^{-y} g \,\mathrm{d}F_n\,,
\end{aligned}$$

sicché, per ogni $n \in \mathbb{N}$, risulta

$$\delta\varphi_n - \delta\varphi_n(y) \leq 2\varepsilon K \qquad (\text{se } |F_n| \leq K)\,,$$

cioè

$$\limsup_{n\to+\infty}\,(\delta\varphi_n - \delta\varphi_n(y)) = 0\,.$$

L'asserto segue ore dal Lemma 2.4.1. □

2.5 Metriche

Esaminiamo ora il problema dell'esistenza di una metrica per la convergenza completa. Una tale metrica è quella di Lévy.

Si indichi con $\mathcal{D}$ l'insieme delle f.r., detto anche spazio delle f.r.; si noti che $\mathcal{D}$ non è uno spazio vettoriale; per vederlo basta osservare, per esempio, che la somma di due f.r. non è una f.r.. A $\mathcal{D}$ si può dare la struttura di spazio metrico introducendo la *metrica di Lévy.* Si denoterà con $(F, G; h)$ ove $h > 0$ la condizione

$$\forall\, x \in \mathbb{R} \qquad F(x-h) - h \leq G(x) \leq F(x+h) + h\,.$$

Teorema 2.5.1. *Sia $d_L : \mathcal{D} \times \mathcal{D} \to \mathbb{R}$ definita da*

$$d_L(F,G) := \inf\{h > 0 : \textit{valgono sia } (F,G;h) \textit{sia } (G,F;h)\}\,.$$

Allora $(\mathcal{D}, d_L)$ è uno spazio metrico e d_L una metrica detta metrica di Lévy.

Si osservi che per ogni coppia $F, G \in \mathcal{D}$ è $d_L(F,G) \leq 1$; infatti, per ogni $x \in \mathbb{R}$ si ha $F(x-1) - 1 \leq 0 \leq G(x) \leq 1 \leq F(x+1) + 1$. Alla dimostrazione è opportuno premettere il seguente

Lemma 2.5.1. *Se $d_L(F,G) = \alpha > 0$ valgono sia $(F,G;\alpha)$ sia $(G,F;\alpha)$.*

Dimostrazione. Sia $t > 0$ e e si scelga y in modo che $x < y$; scende dalla definizione di d_L che

$$F(y-\alpha-t) - \alpha - t \leq G(y) \leq F(y+\alpha+t) + \alpha + t\,. \tag{2.5.1}$$

Nel limite $t \downarrow 0$, la (2.5.1) dà, per la continuità a destra di F,

$$\ell^- F(y-\alpha) - \alpha \leq G(y) \leq F(y+\alpha) + \alpha\,,$$

per ogni $y \in \mathbb{R}$, onde, per $y \downarrow x$, si ha $F(x-\alpha) - \alpha \leq G(x) \leq F(x+\alpha) + \alpha$ cioè la $(F,G;\alpha)$. Analogamente si procede per la $(G,F;\alpha)$. □

Dimostrazione del Teorema 2.5.1. Le condizioni $d_L(F,F) = 0$ per ogni $F \in \mathcal{D}$ e $d_L(F,G) = d_L(G,F)$ per ogni coppia $F, G \in \mathcal{D}$ sono di banale dimostrazione. Si supponga ora che $d_L(F,G) = 0$. Segue dalla definizione di d_L che $\ell^- F(x) \leq G(x)$ e $\ell^- G(x) \leq F(x)$ per ogni $x \in \mathbb{R}$, sicché F e G coincidono tranne, al piú, nell'insieme al piú numerabile dei loro punti di discontinuità. Per la continuità a destra delle f.r. si ha $F = G$.

Rimane da dimostrare la diseguaglianza triangolare

$$d_L(F,H) \leq d_L(F,G) + d_L(G,H)$$

per ogni scelta delle f.r. F, G, H. Posto $\alpha = d_L(F,G)$ e $\beta = d_L(G,H)$, non vi è alcunché da dimostrare se $d_L(F,H) = 0$ oppure se $\alpha + \beta \geq 1$. Si può, perciò, supporre che $d_L(F,H) > 0$, che $\alpha + \beta < 1$, e che $\alpha > 0$, $\beta > 0$. In virtú del Lemma 2.5.1 si ha, per ogni $x \in \mathbb{R}$,

$$\begin{aligned} F(x-\alpha-\beta) - \alpha - \beta &\leq G(x-\beta) - \beta \\ &\leq H(x) \leq G(x+\beta) + \beta \leq F(x+\alpha+\beta) + \alpha + \beta \end{aligned}$$

sicché vale $(F,H;\alpha+\beta)$.

Similmente, si deduce che vale $(H,F;\alpha+\beta)$; perciò, è $d_L(F,H) \leq \alpha + \beta$. □

Per il calcolo della metrica di Lévy è spesso utile la seguente osservazione. Se si completa, come in Figura 5.1, il grafico delle f.r. F e G tracciando segmenti verticali che congiungano i limiti a sinistra e a destra negli eventuali punti di discontinuità, si considerino i segmenti $P_c Q_c$ intercettati tra i due grafici, completati come appena detto, dalle rette parallele alla bisettrice del secondo e quarto quadrante $x + y = c$. Allora

$$d_L(F, G) = \sup \left\{ \frac{\overline{P_c Q_c}}{\sqrt{2}} : c \in \mathbb{R} \right\} .$$

La metrica di Lévy metrizza la convergenza completa.

Teorema 2.5.2. *Per una successione (F_n) di f.r. di $\mathcal{D}$ sono equivalenti le affermazioni:*

(a) *(F_n) converge completamente a $F \in \mathcal{D}$;*

(b) $\lim_{n \to +\infty} d_L(F_n, F) = 0$.

Dimostrazione. (a) $\Longrightarrow$ (b) Fissato in maniera arbitraria $\varepsilon > 0$, si scelgano due punti a e b di continuità per F in modo tale che sia

$$F(a) < \frac{\varepsilon}{2} \qquad \text{e} \qquad F(b) < 1 - \frac{\varepsilon}{2} . \tag{2.5.2}$$

Per $i = 1, \dots, m$, si scelgano m punti di continuità per F nell'intervallo $]a, b[$, in modo che sia $a = a_0 < a_1 < \dots < a_m < a_{m+1} = b$ e $a_{j+1} - a_j < \varepsilon$ $(j = 0, 1, \dots, m)$. Esiste allora un naturale $N = N(\varepsilon)$ tale che sia

$$|F_n(a_j) - F(a_j)| \le \frac{\varepsilon}{2} \tag{2.5.3}$$

per ogni $n \ge N$ e per ogni $j = 0, 1, \dots, m + 1$. Per dimostrare che $d_L(F_n, F)$ tende a 0, basta far vedere che, per ogni $x \in \mathbb{R}$, valgono definitivamente le diseguaglianze

$$F(x - \varepsilon) - \varepsilon \le F_n(x) \le F(x + \varepsilon) + \varepsilon . \tag{2.5.4}$$

Si distinguano tre casi:
(i) $x \in [a_{j-1}, a_j]$. Si prenda $n \ge N$ e si usi la (2.5.3) per ottenere

$$F_n(x) \le F_n(a_j) \le F(a_j) + \frac{\varepsilon}{2} \le F(x + \varepsilon) + \frac{\varepsilon}{2} \le F(x + \varepsilon) + \varepsilon ,$$

e

$$F_n(x) \ge F_n(a_{j-1}) \ge F(a_{j-1}) - \frac{\varepsilon}{2} \ge F(x - \varepsilon) - \varepsilon ,$$

sicché la (2.5.4) è, in questo caso, provata.
(ii) $x < a$. La prima delle (2.5.2) e la (2.5.3) danno, per $n \ge N$,

$$F_n(x) \le F_n(a) \le F(a) + \frac{\varepsilon}{2} \le \varepsilon \le F(x + \varepsilon) + \varepsilon ,$$

e

$$F_n(x) \ge 0 \ge F(a) - \frac{\varepsilon}{2} \ge F(x - \varepsilon) - \varepsilon .$$

(iii) $x > b$. La seconda delle (2.5.2) e la (2.5.3) danno, per $n \ge N$,

$$F_n(x) \le 1 \le F(b) + \frac{\varepsilon}{2} \le F(x + \varepsilon) + \varepsilon ,$$

e

$$F_n(x) \geq F_n(b) \geq F(b) - \frac{\varepsilon}{2} \geq 1 - \varepsilon \geq F(x-\varepsilon) - \varepsilon \,.$$

(b) $\Longrightarrow$ (a) Sia x_0 un punto di continuità per F e si fissi $\varepsilon > 0$ in maniera arbitraria. Esiste allora $\delta = \delta(\varepsilon) > 0$ tale che $|x_0 - x| < \delta$ implichi

$$|F(x) - F(x_0)| < \varepsilon \,. \tag{2.5.5}$$

Posto $\eta < \min\{\varepsilon, \delta\}$, si scelga $N = N(\eta) = N(\varepsilon, \delta) \in \mathbb{N}$ in modo che si abbia $d_L(F_n, F) < \eta$ per ogni $n \geq N$. Scende allora dalla definizione di distanza di Lévy e dalla (2.5.5) che

$$F_n(x_0) \geq F(x_0 - \eta) - \eta \geq F(x_0) - 2\,\varepsilon \,,$$

e che

$$F_n(x_0) \leq F(x_0 + \eta) + \eta \leq F(x_0) + 2\,\varepsilon \,,$$

onde $|F_n(x_0) - F(x_0)| \leq 2\,\varepsilon$ se $n \geq N$. □

Teorema 2.5.3. *Lo spazio metrico $(\mathcal{D}, d_L)$ è completo.*

Dimostrazione. Sia $(F_n) \subseteq \mathcal{D}$ una successione di Cauchy; per il teorema di Helly si può estrarre dalla data una successione $(F_{n(j)})_{j\in\mathbb{N}}$ che converge debolmente alla funzione $G : \mathbb{R} \to [0, 1]$, crescente e continua a destra. Si supponga che, fissato ad arbitrio $\varepsilon > 0$, risulti $d_L(F_n, F_m) < \varepsilon$ se $m, n \geq N = N(\varepsilon)$. Si scelga ora x_m in modo che $F_m(x_m) < \varepsilon$ e si scelga $n(j)$ maggiore di N; perciò $d_L(F_{n(j)}, F_m) < \varepsilon$. Se G è continua nel punto x e se $x < x_m - \varepsilon$, allora è

$$F_{n(j)}(x) \leq F_{n(j)}(x_m - \varepsilon) \leq F_m(x_m) + \varepsilon \leq 2\varepsilon \,,$$

onde $G(x) = \lim_{j\to+\infty} F_{n(j)}(x) \leq 2\varepsilon$; perciò $\lim_{x\to-\infty} G(x) = 0$. Si dimostra in maniera analoga che $\lim_{x\to+\infty} G(x) = 1$, sicché G è effettivamente una f.r. e la successione $(F_{n(j)})_{j\in\mathbb{N}}$ converge completamente a G, cioè $d_L(F_{n(j)}, G) \to 0$. Per dimostrare l'asserto, basta mostrare che tutta la successione (F_n) converge completamente a G; a tal fine, si usi la diseguaglianza triangolare

$$d_L(F_n, G) \leq d_L(F_n, F_{n(j)}) + d_L(F_{n(j)}, G) \,,$$

cioè l'asserto. □

Vale la seguente importante diseguaglianza

Teorema 2.5.4. *Siano X e Y due v.a. definite sullo stesso spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ e siano F e G, rispettivamente, le loro f.r.. Per le metriche di Lévy e di Ky Fan vale la seguente diseguaglianza*

$$d_L(F, G) \leq d_{KF}(X, Y) \,. \tag{2.5.6}$$

Dimostrazione. Sia $\varepsilon > d_{KF}(X,Y)$; posto $A_\varepsilon := \{|X-Y| \le \varepsilon\}$, si ha

$$\mathbb{P}(A_\varepsilon) > 1-\varepsilon\,.$$

Pertanto

$$\begin{aligned}\{X \le x\} \bigcap A_\varepsilon \subseteq \{Y < x+\varepsilon\} \bigcap A_\varepsilon\,,\\ \{Y \le y\} \bigcap A_\varepsilon \subseteq \{X < y+\varepsilon\} \bigcap A_\varepsilon\,.\end{aligned}$$

Passando alle probabilità, dalla prima di queste inclusioni si ha

$$\mathbb{P}\left(\{X \le x\} \bigcap A_\varepsilon\right) \le \mathbb{P}\left(\{Y < x+\varepsilon\} \bigcap A_\varepsilon\right)\,,$$

vale a dire

$$\begin{aligned}&F(x) + \mathbb{P}(A_\varepsilon) - \mathbb{P}\left(\{X \le x\} \bigcup A_\varepsilon\right)\\ &\qquad\qquad \le G(x+\varepsilon) + \mathbb{P}(A_\varepsilon) - \mathbb{P}\left(\{Y < x+\varepsilon\} \bigcup A_\varepsilon\right)\,,\end{aligned}$$

onde

$$\begin{aligned}F(x) - G(x+\varepsilon) &\le \mathbb{P}\left(\{X \le x\} \bigcup A_\varepsilon\right) - \mathbb{P}\left(\{Y < x+\varepsilon\} \bigcup A_\varepsilon\right)\\ &\le 1 - \mathbb{P}(A_\varepsilon) \le 1-(1-\varepsilon) = \varepsilon\,,\end{aligned}$$

sicché

$$F(x) \le G(x+\varepsilon) + \varepsilon\,.$$

Similmente, dalla seconda inclusione, ponendo $y = x - \varepsilon$, si ottiene

$$\mathbb{P}\left(\{Y \le x-\varepsilon\} \cap A_\varepsilon\right) \le \mathbb{P}\left(\{X < x\} \cap A_\varepsilon\right)\,,$$

vale a dire

$$G(x-\varepsilon) + \mathbb{P}\left(A_\varepsilon\right) - \mathbb{P}\left(\{Y \le x-\varepsilon\} \cup A_\varepsilon\right) \le F(x) + \mathbb{P}\left(A_\varepsilon\right) - \mathbb{P}\left(\{X < x\} \cup A_\varepsilon\right)\,,$$

onde, come sopra,

$$\begin{aligned}G(x-\varepsilon) - F(x) &\le \mathbb{P}\left(\{Y \le x-\varepsilon\} \bigcup A_\varepsilon\right) - \mathbb{P}\left(\{X < x\} \bigcup A_\varepsilon\right)\\ &\le 1 - \mathbb{P}(A_\varepsilon) \le 1-(1-\varepsilon) = \varepsilon\,,\end{aligned}$$

e

$$G(x-\varepsilon) - \varepsilon \le F(x)\,;$$

le due diseguaglianze ottenute bastano per concludere che $d_L(F,G) \le \varepsilon$. Si faccia ora tendere ε a $d_{KF}(X,Y)$ decrescendo per avere l'asserto. □

Basta applicare la (2.5.6) per ottenere per altra via il Teorema 2.3.3.

2.6 Altri tipi di convergenza per v.a.

I tipi di convergenza per v.a. considerati sin qui non esauriscono quelli introdotti, studiati ed usati nella letteratura scientifica; qui di seguito, accenneremo brevemente ad altri tipi di convergenza per v.a..

Definizione 2.6.1. Si dice che una successione (X_n) di v.a. definite nello spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ *converge completamente* alla v.a. X se

$$\lim_{n\to+\infty} \sum_{k=n}^{\infty} \mathbb{P}\left(|X_n - X| > \varepsilon\right) = 0\,,$$

per ogni $\varepsilon > 0$. ◇

La convergenza completa di v.a. è un concetto differente dalla convergenza completa di f.r. che si è studiata nella sezione 3.

La dimostrazione della seguente caratterizzazione della convergenza completa è rinviata agli esercizî.

Teorema 2.6.1. *Sia (X_n) una successione di v.a. definite sullo spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$. Sono allora equivalenti le seguenti affermazioni:*

(a) *la successione (X_n) converge completamente a* 0*;*

(b) *ogni successione (Y_n) di v.a., definite su un qualsiasi spazio di probabilità $(\Omega', \mathcal{A}, \mu)$ e tali che, per ogni $n \in \mathbb{N}$, Y_n abbia la stessa legge di X_n, converge q.c. rispetto alla misura di probabilità μ.*

Definizione 2.6.2. Si dice che la successione (X_n) di v.a. definite sullo spazio $(\Omega, \mathcal{F}, \mathbb{P})$ *converge quasi certamente uniformemente* alla v.a. X se esiste un insieme $N \in \mathcal{F}$ con $\mathbb{P}(N) = 0$ tale che, per ogni $\omega \in N^c$, sia

$$\lim_{n\to+\infty} X_n(\omega) = X(\omega) \qquad \text{uniformemente in } \omega\,.$$

Si dice che (X_n) *converge quasi uniformemente* alla v.a. X se, per ogni $\varepsilon > 0$, esiste $N_\varepsilon \in \mathcal{F}$ con $\mathbb{P}(N_\varepsilon) < \varepsilon$ tale che (X_n) converga a X uniformemente in N_ε^c. ◇

Il seguente risultato è classico, dimostra che la convergenza quasi uniforme non è un nuovo concetto e fornisce una nuova caratterizzazione della convergenza quasi certa. La dimostrazione fa uso del teorema di Egorov non dimostrato in queste lezioni.

Teorema 2.6.2. *In una spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ la convergenza quasi certa e quella quasi uniforme sono equivalenti.*

Diamo il teorema che chiarisce i rapporti tra i tipi di convergenza considerati; la dimostrazione è lasciata al lettore come esercizio.

Teorema 2.6.3. *Sia (X_n) una successione di v.a. definite sullo spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$. Si considerino le affermazioni:*

(a) *(X_n) converge q.c. uniformememnte;*

(b) (X_n) *converge completamente;*

(c) (X_n) *converge quasi certamente, o, ciò che è lo stesso, quasi uniformemente.*

Allora, valgono le implicazioni (a) $\Longrightarrow$ (b) $\Longrightarrow$ (c), *mentre non vale alcuna delle implicazioni inverse.*

2.7 Note al Capitolo 2

Abbiamo avuto modo di usare tacitamente il seguente risultato, per il quale si veda, ad esempio, (Dixmier, 1981).

Teorema 2.7.1. *Sia* (Ω, d) *uno spazio metrico e sia* (x_n) *una successione di elementi di* Ω. *Se da ogni successione* $(x_{n(k)})$ *estratta da* (x_n) *se ne può estrarre un'altra* $(x_{k(j)})$ *convergente a* x, *allora tutta la successione* (x_n) *converge a* x.

Dimostrazione. Nelle condizioni enunciate si supponga, per assurdo, che (x_n) non converga a x. Ciò vuol dire che, per ogni $k \in \mathbb{N}$, esiste un elemento $x_{n(k)}$ della successione data tale che $d(x_{n(k)}, x) > 1/k$. Ma allora nessuna estratta da $(x_{n(k)})$ potrebbe convergere a x. □

Sezione 2.1 Esistono numerose generalizzazioni dei lemmi di Borel–Cantelli. Si veda, per esempio, (Petrov, 2004) e la bibliografia lí citata.

Due esempî notevoli di leggi $0-1$ di uso corrente nel calcolo delle probabilità sono quelle di Kolmogorov che si incontrerà piú avanti (si veda la Sezione 6.7.2) e di Hewitt & Savage che si può leggere in molti libri di Probabilità, per esempio in (Bauer, 1996).

Sezione 2.2 Le relazioni tra i varî tipi di convergenza dati in questa sezione valgono in uno spazio di probabilità, oppure, piú in generale in uno spazio di misura finita; le relazioni cambiano se si considera uno spazio di misura non necessariamente finita. Per questo si veda un buon libro di teoria della misura, per esempio (Kingman & Taylor, 1966).

Nel seguito quando si parlerà dell'integrabilità uniforme si vedrà che il Teorema 2.2.8 è un caso particolare di un risultato piú generale.

La convergenza in misura fu introdotta da F. Riesz (1909).

Sezione 2.3 Il Teorema di Helly 2.3.2 fu introdotto da Helly (1921). Esso ha un'estensione al caso multidimensionale; se si dice *crescente* una funzione $F : \mathbb{R}^d \to [0,1]$ che soddisfà alla diseguaglianza (d) (4.8.6), il teorema può essere generalizzato nella seguente maniera.

Teorema 2.7.2. *Da ogni successione* (F_n) *di f.r.* d*–dimensionali se ne può estrarre un'altra*

$$\left(F_{n(k)}\right)_{k\in\mathbb{N}}$$

che converge debolmente ad una funzione $F : \mathbb{R}^d \to [0,1]$ *crescente. La funzione limite non è necessariamente una f.r..*

Per la dimostrazione, oltreché per la condizione necessaria e sufficiente affinché la funzione limite F sia una f.r., si veda in (Feller, 1971) il Teorema 6.1 del Capitolo VIII.

La dimostrazione del teorema di Skorokhod si può trovare, nel caso $\mathbb{R}^n$, in (Billingsley, 1979), in generale in (Skorokhod, 1965), in (Rao, 1987) o in (Rogers & Williams, 1994); (Letta & Pratelli, 1997) ne hanno dato una dimostrazione generale riconducedosi al caso di $\mathbb{R}$.

Sezione 2.4 Le convergenze stretta e vaga furono introdotte da Alexandrov in tre lunghi articoli del tempo di guerra, (Alexandrov, 1940, 1941 e 1943). La forma definitiva è dovuta a Prokhorov (1956). In un ambito piú generale, vale a dire per la convergenza stretta negli spazî metrici, i due riferimenti tradizionali sono (Billingsley, 1968) e (Parthasarathy, 1967). Un approccio differente fu introdotto da Dudley (1966 e 1967); su questo è basato il libro (Pollard, 1984). Un approccio ancora differente, particolarmente utile nelle applicazioni statistiche dovuto a Hoffman–Jorgensen è adottato in (van der Vaart & Wellner, 1996).

Sezione 2.5 Ky Fan (1944) introdusse la metrica che ora porta il suo nome.

La convergenza quasi certa è, in generale, incompatibile con l'esistenza di una norma, si veda (Dugué, 1955); è compatibile se, e solo se, Ω è finito, si vedano (Marczewski, 1955), (Thomasian, 1956 e 1957). Si consultino questi ultimi riferimenti anche per il Teorema 2.2.11.

Per l'esistenza di numerose metriche su L^0 che metrizzano la convergenza in probabilità si consulti (Lukacs, 1975).

La metrica di Lévy fu introdotta da Paul Lévy nell'appendice al libro di Fréchet del 1936, ma si veda anche (Lévy, 1937).

Se si considerano le f.r. di Δ anziché quelle di $\mathcal{D}$, la metrica di Lévy non è piú adeguata a metrizzare la topologia della convergenza debole, che è quella rilevante in Δ; si dice che una successione (F_n) di f.r. di δ *converge debolmente* alla f.r. F se per ogni $x \in \mathbb{R}$ che sia punto di continuità per F, $x \in C(F)$, si ha

$$\lim_{n\to+\infty} F_n(x) = F(x) .$$

La metrica su Δ per la convergenza debole è una modifica di quella di Lévy e fu introdotta da Sibley (1971), si veda anche (Schweizer & Sklar, 1983). Altre metriche topologicamente equivalenti a quella di Sibley si possono trovare in (Sempi, 1982), (Taylor, 1985), (Sempi, 1986), (Pascali & Sempi, 1997).

Nei teoremi di questa sezione abbiamo parlato di metriche che metrizzano certi tipi di convergenza. Ricordiamo però che in uno spazio che soddisfaccia al primo assioma di numerabilità, vale a dire tale che ogni punto abbia una base di intorni numerabile, conoscere la convergenza delle successioni equivale a conoscerne la topologia (si veda, ad esempio, (Kelley, 1955)); perciò, avremmo potuto parlare egualmente bene della topologia della convergenza in probabilità in L^0 e della topologia della convergenza completa in $\mathcal{D}$ o in Δ.

Sezione 2.6 Per questa sezione ho seguito il libro (Lukacs, 1975). La convergenza completa per le v.a. fu introdotta da Hsu & Robbins (1947). Per il teorema di Egorov si veda, per esempio, (Dunford & Schwartz, 1958).

2.8 Esercizî sul Capitolo 2

2.1. Sia X una v.a. distribuita uniformemente su $(0,1)$. Se

$$A_n := \{X < 1/n\}$$

risulta $\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) = +\infty$ e $\mathbb{P}\left(\limsup_{n\to+\infty} A_n\right) = 0$ senza che ciò contraddica al secondo lemma di Borel–Cantelli.

2.2. La condizione che sia convergente la serie $\sum_{n\in\mathbb{N}} \mathbb{P}(A_n)$ è solo sufficiente, ma non necessaria affinché si abbia $\mathbb{P}(\limsup_{n\to+\infty} A_n) = 0$. Si costruiscano uno spazio di probabilità e, in questo, una successione di insiemi misurabili (A_n) tale che $\mathbb{P}(\limsup_{n\to+\infty} A_n) = 0$ e $\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) = +\infty$.

2.3. Si lancia una moneta equilibrata ($p = 1/2$) un numero infinito di volte. Qual è la probabilità di ottenere infinite volte due teste consecutivamente.

2.4. Sia (A_n) una successione di eventi indipendenti con

$$\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) = +\infty\,,$$

sia $\mathbf{1}_{A_n}$ la funzione indicatrice di A_n e si ponga $S_n := \sum_{j=1}^n \mathbf{1}_{A_j}$. Si mostri il seguente risultato, piú forte del primo lemma di Bore–Cantelli:

$$\lim_{n\to+\infty} \frac{S_n}{\mathbb{E}(S_n)} = 1 \qquad \text{q.c.}$$

(Suggerimento: si scelga una sottosuccessione $(n(j))$ tale che $\mathbb{E}(S_{n(j)}) \simeq j^2$ e si mostri prima il risultato per questa sottosuccessione, estendendolo poi all'intera successione).

2.5. Il secondo lemma di Borel–Cantelli vale se l'ipotesi che gli eventi della successione (A_n) siano indipendenti è sostituita da quella che siano *a due a due* indipendenti.

2.6. (a) Si mostri che, se in uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$, la successione di eventi (A_n) soddisfà alle due condizioni

$$\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) = +\infty$$

$$\liminf_{n\to+\infty} \frac{\sum_{j,k=1}^n \mathbb{P}\left(A_j \cap A_k\right)}{\left\{\sum_{j=1}^n \mathbb{P}(A_j)\right\}^2} = 1\,,$$

allora $\mathbb{P}\left(\limsup_{n\to+\infty} A_n\right) = 1$.
(b) Si deduca da (a) la conclusione dell'esercizio 2.6.

2.7. Nello spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ sia (X_n) una successione di v.a. indipendenti ed isonome, tutte di legge esponenziale di parametro 1, $X_n \sim \Gamma(1,1)$. Si ponga, per $\alpha > 0$,

$$A_n := \{X_n > \alpha \ln n\} \qquad \text{e} \qquad U := \limsup_{n\to+\infty} \frac{X_n}{\ln n}\,.$$

(a) Si calcoli la probabilità

$$\mathbb{P}\left(\limsup_{n\to+\infty} A_n\right)\;;$$

(b) si mostri che $\mathbb{P}(U = 1) = 1$.

2.8. (a) Per una successione (X_n) di v.a. nello spazio $(\Omega, \mathcal{F}, \mathbb{P})$ si definisca

$$X_n \xrightarrow[n\to+\infty]{} +\infty \qquad \text{q.c.};$$

(b) si mostri che $X_n \to +\infty$ q.c. se, e solo se, per ogni $M > 0$, si ha

$$\mathbb{P}\left(\limsup_{n\to+\infty}\{X_n < M\}\right) = 0\,.$$

2.9. Si dia l'esempio di uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ e di una successione (X_n) di v.a. definite in tale spazio, per le quali $\mathbb{E}(X_n) \to 0$, mentre non esiste alcuna successione $\{n(k) : k \in \mathbb{N}\}$ per la quale $X_{n(k)} \to 0$ in probabilità.

2.10. Nello spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ una v.a. $X : \Omega \to \overline{\mathbb{R}}$ si dice *limitata in probabilità* oppure *stocasticamente limitata* se, per ogni $\varepsilon > 0$ esiste un numero reale $M(\varepsilon) > 0$ tale che

$$\mathbb{P}\left(|X| \leq M(\varepsilon)\right) \geq 1 - \varepsilon\,.$$

Si dimostri che sono equivalenti le affermazioni:

(a) X è limitata in probabilità;

(b) X è q.c. finita.

2.11. Si dimostri il Teorema 2.2.5 in maniera diversa da quella proposta nel testo.

2.12. (a) Si studii la convergenza della successione di f.r. (F_n), ove F_n è la f.r. della legge $N(0, n)$.

(b) Si studii la convergenza della successione (F_{σ_n}), ove F_{σ_n} è la f.r. della legge $N(0, \sigma_n^2)$ e $\{\sigma_n\}$ è una successione infinitesima con $\sigma_n > 0$ per ogni $n \in \mathbb{R}$.

2.13. Si studii la convergenza completa della successione di f.r. (F_n), ove F_n è la f.r. della legge di Poisson $\mathcal{P}(\lambda_n)$, nei due casi:

(a) $\lambda_n \xrightarrow[n\to+\infty]{} 0$;

(b) $\lambda_n \xrightarrow[n\to+\infty]{} +\infty$.

2.14. Siano $F_n \quad (n \in \mathbb{N})$ e F f.r. di v.a. che assumono valori

$$x_1 < x_2 < \cdots < x_j < \ldots$$

con probabilità date da $\mathbb{P}(X_n = x_j) = p_j^{(n)}$ e $\mathbb{P}(X = x_j) = p_j$. Allora $F_n \xrightarrow[n\to+\infty]{c} F$ se, e solo se, per ogni $j \in \mathbb{N}$, risulta $\lim_{n\to+\infty} p_j^{(n)} = p_j$.

2.15. Può una successione (F_n) ciascuna delle quali è assolutamente continua, convergere completamente a una f.r. che non è assolutamente continua?

2.16. (Teorema di Scheffé) Siano F_n e F f.r. assolutamente continue con densità, rispetto alla misura di Lebesgue, f_n e f, rispettivamente. Se $f_n \to f$ q.o. rispetto alla misura di Lebesgue riesce $F_n \xrightarrow[n\to+\infty]{c} F$. Vale lo stesso risultato se si suppone che $f_n \to f$ nel senso della misura di Lebesgue?

Si vedano a questo proposito (Scheffé, 1947) e (Sempi, 1989).

2.17. Una successione di f.r. assolutamente continue può convergere completamente anche se non converge q.o.la successione delle densità di probabilità. Non vale perciò il reciproco del teorema di Scheffé.

2.18. Sia (X_n) una successione di v.a. indipendenti definite sullo stesso spazio $(\Omega, \mathcal{F}, \mathbb{P})$, ognuna delle quali è distribuita uniformemente nell'intervallo $(0,1)$. Si studii la convergenza della successione (V_n) ove

$$V_n := \vee_{j=1}^n X_j = \max\{X_1, X_2, \ldots, X_n\}\,.$$

2.19. Sia (λ_n) una successione di numeri reali strettamente positivi tale che

$$\lim_{n\to+\infty} \lambda_n = +\infty\,.$$

Nello spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ sia (X_n) una successione di v.a. indipendenti tale che, per ogni $n \in \mathbb{N}$, X_n abbia legge esponenziale di parametro λ_n, $X_n \sim \Gamma(1, \lambda_n)$. Si studii l'eventuale convergenza di (X_n).

2.20. Sia (X_n) una successione di v.a. indipendenti definite sullo stesso spazio $(\Omega, \mathcal{F}, \mathbb{P})$, aventi la stessa legge esponenziale di parametro λ, $X_n \sim \Gamma(1, \lambda)$. Si studii l'eventuale convergenza in legge e in probabilità della successione (T_n), ove

$$T_n := \frac{1}{\ln n} \vee_{j=1}^n X_j = \frac{1}{\ln n} \max\{X_1, X_2, \ldots, X_n\}\,.$$

2.21. Per ogni $n \in \mathbb{N}$, si consideri la legge μ_n uniforme sui punti

$$0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-1}{n}\,.$$

(a) si mostri che (μ_n) converge completamente e si determini la misura di probabilità limite μ;

(b) si trovi un insieme misurabile A per il quale non vale la relazione

$$\lim_{n\to+\infty} \mu_n(A) = \mu(A)\,.$$

2.22. Siano μ e μ_n, per ogni $n \in \mathbb{N}$, misure di probabilità su $(\mathbb{R}, \mathcal{B})$. Sono equivalenti le affermazioni:

(a) $\mu_n \xrightarrow[n\to+\infty]{s} \mu$;

(b) $\limsup_{n\to+\infty} \mu_n(C) \le \mu(C)$ per ogni insieme chiuso C;

(c) $\liminf_{n\to+\infty} \mu_n(B) \ge \mu(B)$ per ogni insieme aperto B;

(d) per ogni funzione $f : \mathbb{R} \to \mathbb{R}$ lipschitziana, limitata e positiva, si ha
$$\liminf_{n\to+\infty} \int f \, \mathrm{d}\mu_n \ge \int f \, \mathrm{d}\mu \,;$$

(e) per ogni funzione $f : \mathbb{R} \to \mathbb{R}$ *semicontinua inferiormente* e limitata inferiormente, si ha
$$\liminf_{n\to+\infty} \int f \, \mathrm{d}\mu_n \ge \int f \, \mathrm{d}\mu \,;$$

(f) per ogni funzione $f : \mathbb{R} \to \mathbb{R}$ *semicontinua superiormente* e limitata superiormente, si ha
$$\limsup_{n\to+\infty} \int f \, \mathrm{d}\mu_n \le \int f \, \mathrm{d}\mu \,.$$

Naturalmente, si possono corrispondentemente esprimere formulazione equivalenti della convergenza in legge di una successione (X_n) di v.a. alla v.a. X.

Ricordiamo che una funzione $f : \mathbb{R} \to \mathbb{R}$ si dice *semicontinua inferiormente* se, per ogni x_0 soddisfà alla condizione
$$\liminf_{x\to x_0} f(x) \ge f(x_0)$$
o, equivalentemente, se per ogni $c \in \mathbb{R}$ gli *insiemi di livello* $\{f \le c\}$ sono chiusi, o, ancora se è chiuso l'*epigrafico* di f,
$$\{(x, y) : f(x) \le y\} \,.$$
Si dice che f è *semicontinua superiormente* se $-f$ è semicontinua inferiormente.

2.23. Ogni successione (X_n) di v.a. indipendenti definite sullo stesso spazio di probabilità ha probabilità eguale a zero o a uno di convergere quasi certamente. In particolare, se esse sono isonome (e non costanti) è nulla la probabilità che (X_n) converga.

2.24. Se (X_n) è una successione di v.a. sullo spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ per la quale esiste $\delta > 0$ tale che sia convergente la serie
$$\sum_{n\in\mathbb{N}} \mathbb{E}\left(|X_n|^\delta\right) \,,$$
allora $X_n \to 0$ q.c..

2.25. Sia X una v.a. su $(\Omega, \mathcal{F}, \mathbb{P})$ con legge uniforme su $(0,1)$; per ogni $n \in \mathbb{N}$ sia X_n la v.a. definita sullo stesso spazio da

$$X_n := n\,\mathbf{1}_{\{X \leq 1/n\}}\,.$$

Si studii la convergenza della successione (X_n).

2.26. Se $X_n \to X$ in L^p con $p \geq 1$ allora $\mathbb{E}(|X_n|^p) \to \mathbb{E}(|X|^p)$.

2.27. Siano (X_n) e (Y_n) due successioni di v.a. che convergono in legge a X e a Y, rispettivamente. Si mostri, con opportuni esempî, che, in generale, né $X_n + Y_n \xrightarrow[n\to+\infty]{\mathcal{L}} X + Y$ né $X_n Y_n \xrightarrow[n\to+\infty]{\mathcal{L}} XY$.

2.28. Siano (X_n), (Y_n) e (Z_n) successioni di v.a. definite sullo stesso spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$;

(a) (Teorema di Slutsky) se $X_n - Y_n \xrightarrow[n\to+\infty]{\mathbb{P}} 0$ e $Y_n \xrightarrow[n\to+\infty]{\mathcal{L}} X$, allora

$$X_n \xrightarrow[n\to+\infty]{\mathcal{L}} X\,;$$

(b) se $Y_n \xrightarrow[n\to+\infty]{\mathbb{P}} 0$ e $X_n \xrightarrow[n\to+\infty]{\mathcal{L}} X$, allora $X_n\,Y_n \xrightarrow[n\to+\infty]{\mathbb{P}} 0$;

(c) $X_n \xrightarrow[n\to+\infty]{\mathcal{L}} X$ e $Y_n \xrightarrow[n\to+\infty]{\mathbb{P}} c$, allora $X_n + Y_n \xrightarrow[n\to+\infty]{\mathcal{L}} X + c$;

(d) se $X_n \xrightarrow[n\to+\infty]{\mathcal{L}} X$, $Y_n \xrightarrow[n\to+\infty]{\mathbb{P}} a$ e $Z_n \xrightarrow[n\to+\infty]{\mathbb{P}} c$, allora

$$X_n\,Y_n + Z_n \xrightarrow[n\to+\infty]{\mathcal{L}} aX + c\,.$$

2.29. Se $X_n \to X$ in legge, allora $\mathbb{E}(|X|) \leq \liminf_{n\to+\infty} \mathbb{E}(|X_n|)$.

2.30. Se $X_n \to X$ in probabilità, allora $\varphi \circ X_n \to \varphi \circ X$ in probabilità per ogni funzione continua $\varphi : \mathbb{R} \to \mathbb{R}$. Il risultato può non essere vero se φ è solo misurabile.

2.31. Se $X_n \to X$ in legge e se $\varphi : \mathbb{R} \to \mathbb{R}$ è continua, è vero che $\varphi \circ X_n$ converge in legge a $\varphi \circ X$?

2.32. Se (X_n) è una successione di v.a. definite sullo stesso spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$, dominata da una v.a. integrabile Y, cioè $|X_n| \leq Y$, per ogni $n \in \mathbb{N}$, $Y \in L^0$, e se $X_n \to X$ in probabilità, allora

$$\int X_n \,\mathrm{d}\,\mathbb{P} \to \int X \,\mathrm{d}\,\mathbb{P}\,.$$

2.33. Siano Ω un insieme al piú numerabile, cioè $\mathrm{card}(\Omega) \leq \aleph_0$, sia $\mathcal{F} \subseteq \mathcal{P}(\Omega)$ una tribú, e sia P una probabilità su $\mathcal{F}$. Per una successione (X_n) di v.a. sono equivalenti le proprietà:

(a) (X_n) converge q.c.;

(b) (X_n) converge in probabilità.

2.34. Dato uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ si introducano le funzioni *essenzialmente limitate*, vale a dire, le funzioni $f : \Omega \to \mathbb{R}$ per le quali esiste una costante $k \geq 0$ tale che $\mathbb{P}(|f| > k) = 0$ e si definisca *l'estremo superiore essenziale* di f mediante

$$\text{ess sup} f := \inf\{k \geq 0 : \mathbb{P}(|f| > k) = 0\} .$$

Si pone

$$\|f\|_\infty := \text{ess sup} f .$$

Sia $\mathcal{L}^\infty$ l'insieme delle funzioni essenzialmente limitate. Posto

$$L^\infty := \mathcal{L}^\infty / \simeq_{\mathbb{P}} ,$$

ove, al solito, $\simeq_P$ è la relazione d'eguaglianza quasi certa, si mostri che

(a) se f è in L^∞ con $\|f\|_\infty = c$, allora $|f| \leq c$ q.c.;

(b) L^∞ è uno spazio vettoriale normato da $\|\cdot\|_\infty$ (anzi è anche completo, come si mostra nei corsi d'analisi matematica, sicché è uno spazio di Banach;

(c) se f è in L^∞ allora essa appartiene anche a L^p per ogni $p \in [1, +\infty[$, sicché vale l'inclusione $L^\infty \subseteq L^p$;

(d) se la successione di v.a. (X_n) di L^∞ converge alla v.a. X di L^∞, allora converge allo stesso limite in L^p per ogni $p \in [1, +\infty[$;

(e) se la successione di v.a. (X_n) di L^∞ converge alla v.a. X di L^∞, allora converge allo stesso limite anche in probabilità;

(f) il viceversa di (e) non è necessariamente vero; si dia l'esempio di uno spazio di probabilità e, su di questo, di una successione di v.a. $(X_n)_{n\in\mathbb{N}}$, ciascuna delle quali è limitata, tale che (X_n) converga in probabilità, ma non in L^∞.

2.35. (Teorema di Pólya) Si mostri che se $F_n \xrightarrow[n\to+\infty]{c} F$ e se la f.r. F è continua, allora F_n converge uniformemente a F in $\mathbb{R}$.

2.36. Se $f \in C_b(\mathbb{R})$ e se $F_{n,\theta}$ è una f.r. avente media θ e varianza $\sigma_n^2(\theta)$ tale che $\lim_{n\to+\infty} \sigma_n^2(\theta) = 0$ per ogni $\theta \in \mathbb{R}$, allora, ponendo

$$\mathbb{E}_{n,\theta}(f) := \int_{\mathbb{R}} f(x)\, \mathrm{d}F_{n,\theta}(x)$$

si ha $\lim_{n\to+\infty} \mathbb{E}_{n,\theta}(f) = f(\theta)$.

2.37. Nello spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ si consideri la famiglia

$$\mathcal{M}(P) := \left\{ f \in L^1 : f > 0 \text{ q.c.}, \mathbb{E}(f) = \int f \,\mathrm{d}\,\mathbb{P} = 1 \right\}$$

delle densità delle misure di probabilità che sono equivalenti a P. (Due misure si dicono *equivalenti* quando ognuna di esse è assolutamente continua rispetto all'altra). Si dimostri che in $\mathcal{M}(P)$ la convergenza in probabilità e quella in L^1 coincidono.

2.38. Si calcoli il limite, al tendere di n a $+\infty$, della successione (F_n), ove, per ogni $n \in \mathbb{N}$, F_n è la f.r. della legge di Student a n gradi di libertà.

2.39. Sia (X_n) una successione di v.a. indipendenti definite sullo stesso spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ e tutte con distribuzione uniforme nell'intervallo $(0, 1)$. Per ogni $n \in \mathbb{N}$, si definisca $Y_n := \prod_{j=1}^n X_j$.

(a) Si determini la legge di Y_n;

(b) si studii la convergenza di (Y_n).

2.40. Sia (E_n) una successione di eventi indipendenti sullo stesso spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$; posto $p_n := \mathbb{P}(E_n)$ e $X_n := \mathbf{1}_{E_n}$, si diano condizioni necessarie e sufficienti sulla successione (p_n) affinché si abbia la convergenza $X_n \to 0$:

(a) in probabilità;

(b) quasi certamente.

2.41. Per ogni successione infinitesima (α_n) di reali strettamente positivi esiste una successione (X_n) di v.a. sullo stesso spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ che converge a zero in probabilità senza che $(\alpha_n X_n)$ converga a zero q.c..

2.42. Sia $(X_n)_{n\in\mathbb{N}}$ una successione di v.a. indipendenti ed isonome di L^1. Per ogni $k \in \mathbb{N}$, sia $\mathcal{F}_k$ la tribú generata dalle v.a. $X_1, \dots, X_k$, cioè

$$\mathcal{F}_k := \mathcal{F}(X_1, \dots, X_k) .$$

Sia $N : \Omega \to \overline{\mathbb{Z}}_+$ una v.a. che assume valori interi positivi, tale che, per ogni $k \in \mathbb{N}$, risulti $\{N \leq k\} \in \mathcal{F}_k$ e tale che $\mathbb{E}(N) < +\infty$. Al solito, si ponga

$$S_n := \sum_{j=1}^n X_j \qquad \text{e} \qquad S_0 := 0 .$$

Allora vale l'*equazione di Wald* che costituisce una generalizzazione della proprietà d'additività delle speranze

$$\mathbb{E}(S_N) = \mathbb{E}(X_1)\,\mathbb{E}(N)$$

ove S_N è la v.a. definita da

$$S_N(\omega) := \sum_{n\in\mathbb{N}} S_n(\omega)\,\mathbf{1}_{\{N=n\}}(\omega) .$$

Si noti che S_N non è stata definita sull'evento trascurabile $\{N = +\infty\}$.

2.43. Siano (X_n) una successione di v.a. e (F_n) la corrispondente successione di f.r. e si supponga che (F_n) converga debolmente ad una funzione crescente F. Allora F è una f.r. se, e solo se, $\{X_n\}$ è stocasticamente limitata (si veda l'esercizio (2.10)).

2.44. Sia (X_n) una successione di v.a. indipendenti e si ponga, per semplicità, $F_n := F_{X_n}$. Delle due condizioni:

(a) $\forall \varepsilon > 0 \qquad \sum_{n\in\mathbb{N}} \{1 - F_n(\varepsilon) + F_n(-\varepsilon)\} < +\infty$,

(b) $X_n \to 0$ q.c.,

la prima implica la seconda; inoltre, se le v.a. della successione sono indipendenti, esse sono equivalenti.

2.45. Se $\mathbb{P}$ e $\mathbb{Q}$ sono due misure di probabilità sullo stesso spazio misurabile $(\Omega, \mathcal{F})$ e $\mathbb{Q}$ è assolutamente continua rispetto a $\mathbb{P}$, $\mathbb{Q} \ll \mathbb{P}$, la convergenza in probabilità rispetto a P implica quella rispetto a $\mathbb{Q}$. Perciò se $\mathbb{P}$ e $\mathbb{Q}$ sono equivalenti, vale a dire se ciascuna di esse è assolutamente continua rispetto all'altra, $\mathbb{Q} \ll \mathbb{P}$ e $\mathbb{P} \ll \mathbb{Q}$, le due convergenze sono equivalenti.

Se invece $\mathbb{Q} \ll \mathbb{P}$, ma $\mathbb{P}$ non è assolutamente continua rispetto a $\mathbb{Q}$, può accadere che $X_n \to 0$ in probabilità $\mathbb{Q}$, senza che X_n tenda a zero in probabilità $\mathbb{P}$.

2.46. Sia (F_n) una successione di f.r. e sia $x \in \mathbb{R}$. Le due proprietà seguenti si equivalgono:

(a) $F_n \xrightarrow[n\to+\infty]{c} \varepsilon_x$;

(b) se $\mathbb{P}_n := \mu_{F_n}$ è la misura di Borel–Stieltjes indotta da F_n, si ha

$$\mathbb{P}_n(A) \xrightarrow[n\to+\infty]{} 0$$

per ogni boreliano A tale che $x \notin \overline{A}$.

2.47. Sia $(\mathbb{P}_n)$ una successione di misure di probabilità su $(\mathbb{R}, \mathcal{B})$ che converge completamente alla misura di probabilità $\mathbb{P}$; allora per ogni $\varepsilon > 0$ esiste un compatto K di $\mathbb{R}$ tale che $\mathbb{P}(K) > 1 - \varepsilon$ e $\mathbb{P}_n(K) > 1 - \varepsilon$ per ogni $n \in \mathbb{N}$.

2.48. (Ancora sul problema dei momenti) Siano X e Y v.a. a valori in $[0,1]$ e si supponga che, per ogni $n \in \mathbb{Z}_+$, si abbia

$$\mathbb{E}\left(X^n\right) = \mathbb{E}\left(Y^n\right) .$$

Si mostri che:

(a) $\mathbb{E}\left(p \circ X\right) = \mathbb{E}\left(p \circ Y\right)$ per ogni polinomio p;

(b) $\mathbb{E}\left(f \circ X\right) = \mathbb{E}\left(f \circ X\right)$ per ogni funzione continua $f : [0,1] \to \mathbb{R}$;

(c) X e Y hanno f.r. eguali.

2.49. Sia (F_n) una successione di f.r. con

$$F_n(x) = 0 \quad \text{per } x < 0 \text{ e } \quad F_n(1) = 1 \,,$$

per ogni $n \in \mathbb{N}$. Si supponga che, per ogni $k \in \mathbb{Z}_+$, esista il limite

$$\alpha_k := \lim_{n\to+\infty} \int_{[0,1]} x^k \, \mathrm{d}F_n(x) \,.$$

Allora (F_n) converge completamente ad una f.r. F tale che

$$\int_{[0,1]} x^k \, \mathrm{d}F(x) = \alpha_k \qquad (k \in \mathbb{Z}_+) \,.$$

2.50. Sullo spazio $\mathcal{D}$ delle f.r. si considera talvolta la distanza della convergenza uniforme, detta, in questo ambito, *metrica di Kolmogorov*, definita da

$$d_K(F,G) := \sup\{|F(x) - G(x)| : x \in \mathbb{R}\} \qquad (F, G \in \mathcal{D}) .$$

Si calcoli la distanza di Kolmogorov $d_K(F,G)$ se $a < b$ e

$$F(x) = \begin{cases} 0\,, & x < 0\,, \\ 1 - \exp(-ax)\,, & x \geq 0\,; \end{cases}$$

o

$$G(x) = \begin{cases} 0\,, & x < 0\,, \\ 1 - \exp(-bx)\,, & x \geq 0\,; \end{cases}$$

2.51. Le due condizioni $(F, G; h)$ e $(G, F; h)$ che intervengono nella definizione della metrica di Lévy sono equivalenti. Inoltre, se ε_a è la f.r. della v.a. $X = a$ q.c., allora

$$d_L\,(\varepsilon_a, \varepsilon_b) = \min\{1, |b - a|\} \leq |b - a|\,.$$

2.52. Siano X e Y due v.a. discrete con

$$\mathbb{P}(X = 0) = \mathbb{P}(X = 1) = 1/2 \qquad \text{e} \qquad \mathbb{P}(Y = 1) = 1/4, \mathbb{P}(Y = 2) = 3/4\,;$$

siano F e G sono le rispettive f.r. si calcolino $d_L(F,G)$ e $d_K(F,G)$.

2.53. Sia F la f.r. della distribuzione uniforme su $(-1, 1)$ e G la f.r. della distribuzione uniforme su $(0, 1)$. Si calcolino $d_L(F,G)$ e $d_K(F,G)$.

2.54. Si calcolino $d_L(F,G)$ e $d_K(F,G)$ se F e G sono date da

$$F(x) = x\,\mathbf{1}_{[0,1[}(x) + \mathbf{1}_{[1,+\infty[}(x)\,,$$
$$G(x) = x^2\,\mathbf{1}_{[0,1[}(x) + \mathbf{1}_{[1,+\infty[}(x)\,.$$

2.55. (a) Per ogni coppia di f.r. F e G si dimostri che

$$d_L(F,G) \leq d_K(F,G)\,.$$

Si dia l'esempio di una successione $(F_n) \subseteq \mathcal{D}$ e di una f.r. F tali che

$$d_L(F_n, F) \to 0\,,$$

mentre la successione $(d_K(F_n, F))$ non tende a zero.
(b) Se la f.r. G è assolutamente continua con densità g, si ponga

$$A := \sup\{g(x) : x \in \mathbb{R}\}\,,$$

in caso contrario si ponga $A := +\infty$; allora

$$d_K(F,G) \leq\ (A + 1)\,\mathrm{d}_L(F,G)\,.$$

2.56. (a) Le f.r. assolutamente continue sono dense in $\mathcal{D}$, rispetto alla topologia della metrica di Lévy. In questa topologia, anche le f.r. discrete sono dense in $\mathcal{D}$.
(b) Se F_1, F_2, G_1, G_2 sono f.r., vale la diseguaglianza

$$d_L\left(F_1 * F_2, G_1 * G_2\right) \leq d_L(F_1, G_1) + d_L(F_2, G_2)\,.$$

(c) Si dimostri che se (F_n) e (G_n) sono due successioni di f.r. che convergono completamente alle f.r. F e G, rispettivamente, allora $(F_n * G_n)$ converge completamente a $F * G$.

2.57. Si mostri che se $d_{KF}(X, Y) = \alpha > 0$, allora

$$\mathbb{P}\left(|X - Y| > \alpha\right) \leq \alpha\,.$$

2.58. La convergenza in probabilità per v.a. arbitrarie può essere ridotta a quella per v.a. uniformemente limitate mediante la trasformazione

$$X' := \arctan X\,.$$

Si mostri che $X_n \to X$ in probabilità se, e solo se $d_0(X_n, X) \to 0$, ove

$$d_0(X, Y) := \mathbb{E}\left(|\arctan X - \arctan Y|\right)\,.$$

2.59. (a) Quella di Ky Fan non è l'unica metrica su L^0 che metrizzi la convergenza in probabilità; di questa stessa proprietà gode anche la metrica definita da

$$d(X, Y) := \mathbb{E}\left(\frac{|X - Y|}{1 + |X - Y|}\right)\,.$$

Si usi tale metrica per dimostrare che

(b) che se (X_n) tende a X in L^p $(p \geq 1)$ vi converge anche in probabilità;

(c) che se (X_n) tende a X q.c. vi converge anche in probabilità.

2.60. Sia $f : \mathbb{R}_+ \to \mathbb{R}_+$ una funzione continua, strettamente crescente, limitata, tale che $f(0) = 0$ e *subadditiva*, vale a dire tale che

$$\forall\, x, y \in \mathbb{R}_+ \quad f(x + y) \leq f(x) + f(y)\,.$$

Allora

$$d_f(X, Y) := \mathbb{E}\left[f\left(|X - Y|\right)\right]$$

definisce una metrica su $L^0 = L^0(\Omega, \mathcal{F}, \mathbb{P})$ la cui topologia è quella della convergenza in probabilità.

Esempî di funzioni con le proprietà richieste sono

$$f_1(x) := 1 - e^{-x}\,, \tag{a}$$

$$f_2(x) := \frac{kx}{1 + kx}\,, \tag{b}$$

$$f_3(x) := \tanh x \tag{c}$$

Si noti che la metrica dell'Esercizio 2.59 è un caso particolare di questo; basta considerare la funzione f_2 con $k = 1$.

2.61. Si dimostri direttamente l'implicazione (c) $\Longrightarrow$ (a) del Teorema 2.3.1.

2.62. Sia $\varphi : \mathbb{R} \to \mathbb{R}$ una funzione uniformemente continua e limitata e si ponga

$$\|\varphi\| := \sup\{|\varphi(x)| : x \in \mathbb{R}\}.$$

Se $\varepsilon > 0$ e δ sono tali che

$$|\varphi(s) - \varphi(t)| < \varepsilon \quad \text{ogni qual volta} \quad |s - t| < \delta,$$

allora, quali che siano le v.a. X e Y in $(\Omega, \mathcal{F}, \mathbb{P})$ si ha

$$|\mathbb{E}(\varphi \circ X) - \mathbb{E}(\varphi \circ (X + Y))| \le \varepsilon + 2\,\|\varphi\|\,\mathbb{P}(|Y| \ge \delta).$$

2.63. Una successione di leggi di probabilità (μ_n) sullo spazio misurabile $(\mathbb{R}, \mathcal{B})$ converge strettamente alla misura di probabilità μ se, e solo se, per ogni funzione $\varphi : \mathbb{R} \to \mathbb{R}$ uniformemente continua e limitata si ha

$$\lim_{n\to+\infty} \int_{\mathbb{R}} \varphi \,\mathrm{d}\mu_n = \int_{\mathbb{R}} \varphi \,\mathrm{d}\mu.$$

2.64. Siano X_n $(n \in \mathbb{N})$, X, Y v.a. q.c. finite definite in $(\Omega, \mathcal{F}, \mathbb{P})$. Se

$$X_n + \alpha\, Y \xrightarrow[n\to+\infty]{\mathcal{L}} X + \alpha\, Y$$

per ogni $\alpha > 0$, allora

$$X_n \xrightarrow[n\to+\infty]{\mathcal{L}} X.$$

2.65. Nello spazio $(\Omega, \mathcal{F}, \mathbb{P})$ sia (A_n) una successione di eventi. Allora

(a) vale la diseguaglianza

$$\mathbb{P}\left(\bigcup_{j=1}^{n} A_j\right) \ge \frac{E^2\left(\sum_{j=1}^{n} \mathbf{1}_{A_n}\right)}{\mathbb{E}\left[\left(\sum_{j=1}^{n} \mathbf{1}_{A_n}\right)^2\right]};$$

(b) se $\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) = +\infty$ e se esiste una costante $c > 0$ tale che si abbia, per ogni $j < k$,

$$\mathbb{P}\left(A_j \bigcap A_k\right) \le c\,\mathbb{P}(A_j)\,\mathbb{P}(A_{k-j}),$$

allora $\mathbb{P}(\limsup_{n\to+\infty} A_n) > 0$.

2.66. Nello spazio $(\Omega, \mathcal{F}, \mathbb{P})$ sia (A_n) una successione di eventi tale che

$$\lim_{n\to+\infty} \mathbb{P}(A_n) = 0 \qquad \text{e} \qquad \sum_{n\in\mathbb{N}} \mathbb{P}(A_n \setminus A_{n+1}) < +\infty;$$

allora $\mathbb{P}(\limsup_{n\to+\infty} A_n) = 0$.

2.67. Sia (μ_n) una successione di misure di probabilità su $(\mathbb{R}, \mathcal{B})$ che converge strettamente alla misura di probabilità μ. Se (f_n) è una successione di funzioni continue e limitate che converge uniformemente a f, si ha

$$\lim_{n\to+\infty} \int_{\mathbb{R}} f_n \, \mathrm{d}\mu_n = \int_{\mathbb{R}} f \, \mathrm{d}\mu \, .$$

2.68. Si mostri che se (X_n) converge in legge a X rispetto alla misura di probabilità P non è detto che vi converga rispetto ad ogni misura di probabilità Q equivalente a P.

2.69. Siano date una successione (X_n) di v.a. ed una v.a. X definite su $(\Omega, \mathcal{F}, \mathbb{P})$. Se per ogni boreliano A che sia di continuità per P_X si ha

$$\lim_{n\to+\infty} \mathbb{P}\left(\{X_n \in A\} \Delta \{X \in A\}\right) = 0 \, ,$$

allora (X_n) converge in legge a X.

2.70. Si consideri l'insieme $[0,1]$ munito della misura di Lebesgue (ristretta ai boreliani di $[0,1]$). Data la successione (X_n) di v.a. definite da $X_n(x) := x$ per $x \in [0,1]$ e per $n \in \mathbb{N}$ e la v.a.

$$X(x) := \begin{cases} 1 \, , & x = 0 \, , \\ x + \dfrac{1}{2}, & x \in \left]0, \dfrac{1}{2}\right[\, , \\ x - \dfrac{1}{2}, & x \in \left[\dfrac{1}{2}, 1\right] . \end{cases}$$

Si dica se (X_n) converge a X in legge rispetto alla misura di probabilità Q che ha densità $f(x) := 2x$ rispetto alla misura di Lebesgue.

2.71. Sia (x_n) una successione di numeri reali convergente a x,

$$\lim_{n\to+\infty} x_n = x \in \mathbb{R}$$

e si supponga che, per ogni $n \in \mathbb{N}$, sia $x_n \neq x$. Si consideri la successione di misure di probabilità (μ_n) con $\mu_n = \delta_{x_n}$, e la probabilità $\mu = \delta_x$, le misure di Dirac concentrate in x_n e in x rispettivamente. Si mostri che (μ_n) converge strettamente a μ e si dia l'esempio di un boreliano A tale che $\mu_n(A)$ non converga a $\mu(A)$.

2.72. Sia (X_n) una successione di v.a. positive di L^1. Se è convergente la serie numerica

$$\sum_{n\in\mathbb{N}} \mathbb{E}(X_n) \, ,$$

allora converge q.c. la serie

$$\sum_{n\in\mathbb{N}} X_n \, .$$

2.73. Nello spazio $(\Omega, \mathcal{F}, \mathbb{P})$ siano (X_n) e (Y_n) due successioni di v.a. tali che

$$\sum_{n\in\mathbb{N}} \mathbb{P}(X_n \neq Y_n) < +\infty \, .$$

Allora, se (a_n) è una successione crescente di numeri reali con $a_n \to +\infty$, si ha

(a) la serie $\sum_{n\in\mathbb{N}}(X_n - Y_n)$ converge q.c.;

$$\frac{1}{a_n}\sum_{j=1}^{n}(X_j - Y_j) \xrightarrow[n\to+\infty]{} 0 \quad \text{q.c.}; \tag{b}$$

(c) se $\frac{1}{a_n}\sum_{j=1}^{n} X_j$ converge q.c., allora anche $\frac{1}{a_n}\sum_{j=1}^{n} Y_j$ converge q.c. allo stesso limite.

2.74. Si dimostri il Teorema 2.6.1.

2.75. Si dimostri il Teorema 2.6.2.

2.76. Si dimostri il Teorema 2.6.3.

2.77. La convergenza completa implica la convergenza quasi certa; viceversa, se le v.a. della successione (X_n) sono indipendenti, la convergenza completa coincide con la convergenza quasi certa.

2.78. Per ogni $n \in \mathbb{N}$ sia

$$Z_n = X_n + Y_n\,,$$

dove, per ogni $n \in \mathbb{N}$, Y_n appartiene a L^2; si supponga, inoltre che sia

$$\mathbb{E}(Y_n) \xrightarrow[n\to+\infty]{} 0 \qquad \text{e} \qquad V(Y_n) \xrightarrow[n\to+\infty]{} 0\,.$$

Se (X_n) converge in legge a una v.a. X, allora anche (Z_n) converge in legge a X.
(Questo risultato è dovuto a Cramér (1946)).

2.79. Per ogni $n \in \mathbb{N}$ sia

$$Z_n = X_n + Y_n$$

dove, per ogni $n \in \mathbb{N}$, Y_n appartiene a L^2; si supponga, inoltre che sia

$$\mathbb{E}(Y_n) \xrightarrow[n\to+\infty]{} 0 \qquad \text{e} \qquad V(Y_n) \xrightarrow[n\to+\infty]{} 0\,.$$

Per ogni $n \in \mathbb{N}$ sia U_n una v.a. indipendente da X_n. Se (X_n) e (U_n) convergono in legge a v.a. di f.r. rispettivamente eguali a F e a H, allora, per ogni coppia di numeri reali x e y, si ha

$$\lim_{n\to+\infty}\mathbb{P}\left(Z_n \le x, U_n \le y\right) = F(x)\,H(y)\,.$$

(Per questo esercizio, si veda (Rényi, 1953)).

2.80. Nello spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ sia data una sucessione (X_n) di v.a. isonome e di L^1. Se

$$X_n^* := \max\{|X_j| : j = 1, 2, \dots, n\}$$

allora è

$$\frac{X_n^*}{n} \xrightarrow[n\to+\infty]{} 0$$

tanto q.c. quanto in L^1.

2.81. Date una misura di probabilità μ sullo spazio misurabile $(\Omega, \mathcal{F})$ si definisce la *variazione* di μ mediante

$$\|\mu\| := \sup\{\mu(A) : A \in \mathcal{F}\} .$$

Se ν è un'altra misura di probabilità sullo stesso spazio misurabile, resta cosídefinita la *distanza in variazione* tra μ e ν:

$$\|\mu - \nu\| := \sup \{|\mu(A) - \nu(A)| : A \in \mathcal{F}\} .$$

Si mostri che, date μ e ν,

(a) vale
$$\|\mu - \nu\| := \sup \{\mu(A) - \nu(A) : A \in \mathcal{F}\} ;$$

(b) esiste una misura λ su $(\Omega, \mathcal{F})$ rispetto alla quale μ e ν sono assolutamente conitnue e, dunque, ammettono densità f e g, $\mu = f \cdot \lambda$ e $\nu = g \cdot \lambda$;

(c) esiste un insieme $B \in \mathcal{F}$ tale che $\|\mu - \nu\| = \mu(B) - \nu(B)$; inoltre, per le due densità di (b) vale
$$\|\mu - \nu\| = \int (f - g)^+ \, \mathrm{d}\lambda = \frac{1}{2} \int |f - g| \, \mathrm{d}\lambda ;$$

(d) esistono due misure di probabilità σ e τ su $(\Omega, \mathcal{F})$ tali che, per ogni $A \in \mathcal{F}$ sia
$$\mu(A) - \nu(A) = \|\mu - \nu\| \, (\sigma(A) - \tau(A)) ;$$

(e) se $f : \Omega \to \mathbb{R}$ è limitata e la sua *oscillazione* è definita mediante
$$o(f) := \sup\{|f(\Omega) - f(\Omega')| : \Omega, \Omega' \in \Omega\} ,$$
allora vale la diseguaglianza
$$\left| \int f \, \mathrm{d}\mu - \int f \, \mathrm{d}\nu \right| \leq o(f) \, \|\mu - \nu\| .$$

2.82. Sia (μ_n) una successione di misure finite su $(\mathbb{R}, \mathcal{B})$ che converge strettamente (o, rispettivamente, vagamente) a μ. Se $\varphi : \mathbb{R} \to \mathbb{R}_+$ è continua e limitata, oppure, rispettivamente, continua e con il supporto compatto, allora, per le misure definite da $\nu_n := \varphi \cdot \mu_n$ $(n \in \mathbb{N})$ e da $\nu := \varphi \cdot \mu$ si ha che la successione (ν_n) converge strettamente (o, rispettivamente, vagamente) a ν.

2.83. Una famiglia $\mathcal{H}$ di misure di probabilità sullo spazio misurabile $(\mathbb{R}, \mathcal{B})$ si dice *tesa* (*tight* in inglese) se, per ogni $\varepsilon > 0$ esiste un compatto K_ε con la proprietà che, per ogni misura $\mu \in \mathcal{H}$, si abbia

$$\mu\,(K_\varepsilon) > 1 - \varepsilon .$$

Nello spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ sia (X_n) una successione di v.a. limitata in L^2, vale a dire tale che

$$\sup_{n \in \mathbb{N}} \|X_n\|_2^2 = \sup_{n \in \mathbb{N}} \mathbb{E}\left(X_n^2\right) < +\infty .$$

Allora la successione (μ_n) delle leggi delle v.a. di (X_n) è tesa.

2.84. Nello spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ sia (X_n) una successione di v.a. che converge q.c. allav.a. X che è q.c. finita. Allora la v.a. $Y := \sup_{n\in\mathbb{N}} |X_n|$ è q.c. finita.

2.85. Nello spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ sia (X_n) una successione di v.a. normali e centrate, $X_n \sim N(0, \sigma_n^2)$ $n \in \mathbb{N}$; se (X_n) converge in probabilità, essa converge anche in L^2.