# Convergenza e leggi dei Grandi Numeri

Appunti per il corso di CP2, a.a. 2001/2002

## 1 Convergenza

In questo paragrafo studieremo la convergenza quasi certa, in probabilità e in $L^p$, stabilendo eventuali implicazioni e mostrando esempi e controesempi.

Cominciamo dalle definizioni.

**Definizione 1.1** Sia $(\Omega, \mathcal{F}, \mathbb{P})$ uno spazio di probabilità, dov'è definita una successione $\{X_n\}_n$ di v.a. e una ulteriore v.a. $X$.

- Diremo che $\{X_n\}_n$ converge a $X$ **quasi certamente (q.c.)** se esiste $N \in \mathcal{F}$ tale che $\mathbb{P}(N) = 0$ e per ogni $\omega \notin N$,
$$\lim_{n\to\infty} X_n(\omega) = X(\omega)$$
(o, equivalentemente, se $P(\{\omega : X_n(\omega) \not\to X(\omega)\}) = 0$).
- Diremo che $\{X_n\}_n$ converge a $X$ **in probabilità** se per ogni $\delta > 0$
$$\lim_{n\to\infty} \mathbb{P}(|X_n - X| > \delta) = 0.$$
- Se $X_n \in L^p$ per ogni $n$ e $X \in L^p$, con $p \geq 1$, diremo che $\{X_n\}_n$ converge a $X$ **in** $L^p$ se[1]
$$\lim_{n\to\infty} ||X_n - X||_p = 0.$$

La convergenza quasi certa coincide quindi con la convergenza puntuale ($X_n(\omega) \stackrel{n\to\infty}{\longrightarrow} X(\omega)$) a meno di un insieme di probabilità nulla e quindi trascurabile. La convergenza in probabilità invece richiede che asintoticamente (cioè per $n \to \infty$) divenga trascurabile l'evento $\{\omega : |X_n(\omega) - X(\omega)| > \delta\}$, che corrisponde all'evento "$X_n$ dista da $X$ per più di $\delta$", per ogni $\delta > 0$. Infine, la convergenza in $L^p$ è forse quella più semplice intuitivamente. Infatti, poiché $L^p(\Omega, \mathcal{F}, \mathbb{P})$ è uno spazio di Banach con la norma $|| \cdot ||_p$, la convergenza di $X_n$ a $X$ in $L^p$ corrisponde alla convergenza a 0 della distanza (in $L^p$: $||X_n - X||_p$) tra $X_n$ e $X$.

Cominciamo a dimostrare un primo risultato sulla convergenza q.c.:

**Proposizione 1.2** *(i) $X_n \to X$ q.c. se e solo se $\mathbb{P}(|X_n - X| > \delta \ \ i.o.) = 0$, per ogni $\delta > 0$.*

*(ii) Se per ogni $\delta > 0$ la serie $\sum_n \mathbb{P}(|X_n - X| > \delta)$ converge allora $X_n \to X$ q.c.*

**Dimostrazione.** *(i)* Osserviamo anzitutto che $X_n(\omega) \to X(\omega)$ se e solo se per ogni $\delta > 0$ esiste un $n$ tale che per ogni $k \geq n$ si ha $|X_k(\omega) - X(\omega)| \leq \delta$. Ora, poiché ogni numero reale si può

[1] Ovviamente, in tal caso si ha anche $||X_n||_p \to ||X||_p$, o equivalentemente $\mathbb{E}(|X_n|^p) \to \mathbb{E}(|X|^p)$. Infatti, dalla disuguaglianza triangolare segue che $|\,||X_n||_p - ||X||_p\,| \leq ||X_n - X||_p$, quindi se $||X_n - X||_p \to 0$ per $n \to \infty$ allora $||X_n||_p \to ||X||_p$ per $n \to \infty$. In particolare, se $p = 1$ allora $\mathbb{E}(X_n) \to \mathbb{E}(X)$.

approssimare con dei razionali, non cambia nulla se il $\delta$ in questione si sceglie nei razionali $\mathbb{Q}$. Quindi, in simboli possiamo scrivere

$$\{\omega\,:\,X_n(\omega)\to X(\omega)\}=\cap_{\delta>0,\delta\in\mathbb{Q}}\cup_n\cap_{k\geq n}\{\omega\,:\,|X_k(\omega)-X(\omega)|\leq\delta\}$$

il che peraltro prova che $\{\omega\,:\,X_n(\omega)\to X(\omega)\}\in\mathcal{F}$ perché unione e intersezione numerabile di eventi ($\{\omega\,:\,|X_k(\omega)-X(\omega)|\leq\delta\}$) che appartengono a $\mathcal{F}$. Quindi,

$$\begin{aligned}\{\omega\,:\,X_n(\omega)\to X(\omega)\}^c&=\cup_{\delta>0,\delta\in\mathbb{Q}}\cap_n\cup_{k\geq n}\{\omega\,:\,|X_k(\omega)-X(\omega)|>\delta\}\\&=\cup_{\delta>0,\delta\in\mathbb{Q}}\{|X_n-X|>\delta\ \text{ i.o.}\}.\end{aligned}$$

Ora, $X_n\to X$ q.c. se e solo se il complementare di $\{\omega\,:\,X_n(\omega)\to X(\omega)\}$ ha probabilità nulla, ciè se e solo se

$$\mathbb{P}(\cup_{\delta>0,\delta\in\mathbb{Q}}\{|X_n-X|>\delta\ \text{ i.o.}\})=0,$$

il che è vero se e solo se $\mathbb{P}(|X_n-X|>\delta\ \text{ i.o.})=0$ per ogni $\delta>0$. Infatti, se $\mathbb{P}(\cup_{\delta>0,\delta\in\mathbb{Q}}\{|X_n-X|>\delta\ \text{ i.o.}\})=0$ allora evidentemente $\mathbb{P}(\{|X_n-X|>\delta\ \text{ i.o.}\})=0$ per ogni $\delta>0$, $\delta\in\mathbb{Q}$. Preso $\delta\in\mathbb{R}\setminus\mathbb{Q}$ con $\delta>0$, allora esiste un $\delta'\in\mathbb{Q}$ tale che $\delta'\in(0,\delta)$, quindi $\{|X_n-X|>\delta\ \text{ i.o.}\}\subset\{|X_n-X|>\delta'\ \text{ i.o.}\}$, da cui segue che $\mathbb{P}(|X_n-X|>\delta\ \text{ i.o.})=0$ per ogni $\delta>0$.

Viceversa, se $\mathbb{P}(|X_n-X|>\delta\ \text{ i.o.})=0$ per ogni $\delta>0$, allora $\cup_{\delta>0,\delta\in\mathbb{Q}}\{|X_n-X|>\delta\ \text{ i.o.}\}$ ha probabilità nulla, quindi $X_n\to X$ q.c.

*(ii)* Se per ogni $\delta>0$ la serie $\sum_n\mathbb{P}(|X_n-X|>\delta)$ converge allora usando il primo Lemma di Borel-Cantelli segue che $\mathbb{P}(\{|X_n-X|>\delta\ \text{ i.o.}\})=0$ e da *(i)* si ottiene $X_n\to X$ q.c.

■

Osserviamo che la proposizione precedente dà delle condizioni sulla convergenza q.c. che coinvolgono le leggi di $X_n$ e $X$ e non le loro espressioni come funzioni di $\omega$. Questo è particolarmente importante in probabilità perché spesso (quasi sempre...) non si conosce il valore che una generica v.a. assume sulle $\omega$ ma si hanno informazioni su come si comporta in probabilità, ad esempio se ne conosce la legge, oppure la media o la varianza o i momenti etc (tipicamente, solo negli esempi "mirati", come i controesempi, si sa chi sono $\Omega$, $\mathcal{F}$, $\mathbb{P}$ e quindi chi è $X(\omega)$...). A tale riguardo, si propone il seguente esercizio.

**Esercizio 1.3** *Sia* $\{Y_n\}_n$ *una successione di v.a. i.i.d.* $\text{Exp}(\lambda)$[2]. *Posto* $X_n=Y_n/\log n$, *discutere se* $X_n\to 0$ *in probabilità e/o in* $L^p$ *e/o q.c.*

**Soluzione.** Fissato $\delta>0$, studiamo $\mathbb{P}(|X_n-X|>\delta)$, dove qui $X=0$:

$$\begin{aligned}&\mathbb{P}(|X_n-X|>\delta)=\mathbb{P}(X_n>\delta)=\mathbb{P}(Y_n>\delta\log n)\\&=\int_{\delta\log n}^{\infty}\lambda\,e^{-\lambda x}dx=e^{-\lambda\delta\log n}=\frac{1}{n^{\lambda\delta}}\to 0\quad\text{ se } n\to\infty\end{aligned}$$

quindi $X_n\to 0$ in probabilità. Inoltre, una volta osservato che una $Z\sim\text{Exp}(\lambda)$ è in $L^p$ per ogni $p\geq 1$ (dimostrare!!), si ha

$$||X_n-X||_p^p=\mathbb{E}(X_n^p)=\frac{1}{(\log n)^p}\mathbb{E}(Y_n^p)=\frac{c_{\lambda,p}}{(\log n)^p}\to 0\quad\text{ se } n\to\infty$$

[2]Cioè, con densità $p(x)=\lambda e^{-\lambda x}1\!\!1_{x>0}$

avendo posto $c_{\lambda,p} = \mathbb{E}(Z^p)$, con $Z \sim \text{Exp}(\lambda)$. Quindi $X_n \to 0$ in $L^p$.

Sudiamo ora la convergenza q.c. Cominciamo a vedere se vale la condizione sufficiente del punto *(ii)* della Proposizione 1.2. Per $\delta > 0$, abbiamo visto che $\mathbb{P}(|X_n - X| > \delta) = \frac{1}{n^{\lambda\delta}}$ quindi la serie $\sum_n \mathbb{P}(|X_n - X| > \delta)$ converge se e solo se $\delta > 1/\lambda$. Ciò significa che, purtroppo, non possiamo dire se $X_n \to 0$ q.c. Però, abbiamo visto che la serie $\sum_n \mathbb{P}(|X_n - X| > \delta)$ diverge per $\delta \leq 1/\lambda$ e abbiamo a disposizione uno strumento, il secondo Lemma di Borel-Cantelli, che dà informazioni sulla probabilità dell'evento $\{|X_n - X| > \delta \ \text{ i.o.}\}$ purché però gli eventi $\{|X_n - X| > \delta\}$ siano indipendenti. Qui,

$$\{|X_n - X| > \delta\} = \{Y_n > \delta \log n\}$$

e le v.a. $Y_n$ sono indipendenti per ipotesi, quindi gli eventi $\{|X_n - X| > \delta\}$ sono indipendenti. Allora, usando BC2, si ha

$$\mathbb{P}(|X_n - X| > \delta \ \text{ i.o.}) = 1$$

per $\delta \in (0, 1/\lambda]$ e dal punto *(i)* della Proposizione 1.2 possiamo concludere che non c'è convergenza q.c.

∎

La proposizione che segue dà delle interessanti implicazioni tra le convergenze definite nella Definizione 1.1.

**Proposizione 1.4** *(a) Se $X_n \to X$ q.c. allora $X_n \to X$ in probabilità.*

*(b) Se $X_n \to X$ in probabilità allora esiste una sottosuccessione $\{X_{n_k}\}_k$ tale che $X_{n_k} \to X$ q.c.*

*(c) Se $X_n \to X$ in $L^p$ allora $X_n \to X$ in probabilità.*

*(d) Se $X_n \to X$ q.c. ed esiste $Z \in L^p$, $Z \geq 0$ q.c. tale che $|X_n| \leq Z$ allora $X_n \to X$ in $L^p$. In particolare, $||X_n||_p \to ||X||_p$ per $n \to \infty$ e se $p = 1$, $\mathbb{E}(X_n) \to \mathbb{E}(X)$.*

**Dimostrazione.** *(a)* Se $X_n \to X$ q.c. allora, dalla Proposizione 1.2, per ogni $\delta > 0$ si ha

$$\begin{aligned} 0 &= \mathbb{P}(|X_n - X| > \delta \ \text{ i.o.}) = \mathbb{P}(\cap_n \cup_{k \geq n} \{\omega : |X_k(\omega) - X(\omega)| > \delta\}) \\ &= \lim_{n\to\infty} \mathbb{P}(\cup_{k\geq n}\{\omega : |X_k(\omega) - X(\omega)| > \delta\}) \geq \lim_{n\to\infty} \mathbb{P}(\{|X_n - X| > \delta\}) \end{aligned}$$

quindi $\lim_{n\to\infty} \mathbb{P}(\{|X_n - X| > \delta\}) = 0$, per ogni $\delta > 0$, e $X_n \to X$ in probabilità.

*(b)* Per ipotesi, per ogni $\delta > 0$ si ha che $\lim_{n\to\infty} \mathbb{P}(\{|X_n - X| > \delta\}) = 0$. Quindi, per ogni $k \geq 1$, $\lim_{n\to\infty} \mathbb{P}(|X_n - X| > 1/k) = 0$, dunque esiste un indice $n_k$ tale che

$$\mathbb{P}(|X_{n_k} - X| > 1/k) < \frac{1}{2^k}.$$

Consideriamo la sottosuccessione $\{X_{n_k}\}_k$ appena costruita e mostriamo che $X_{n_k} \to X$ q.c. per $k \to \infty$.

Fissato $\varepsilon > 0$, esiste un $k_0$ tale che $1/k < \varepsilon$ per ogni $k \geq k_0$. Allora, per $k \geq k_0$,

$$\{|X_{n_k} - X| > \varepsilon\} \subset \{|X_{n_k} - X| > 1/k\}$$

quindi $\sum_k \mathbb{P}(|X_{n_k} - X| > \varepsilon) = \sum_{k<k_0} \mathbb{P}(|X_{n_k} - X| > \varepsilon) + \sum_{k\geq k_0} \mathbb{P}(|X_{n_k} - X| > \varepsilon) \leq \sum_{k<k_0} \mathbb{P}(|X_{n_k} - X| > \varepsilon) + \sum_{k\geq k_0} 1/2^k < \infty$. Il punto *(ii)* della Proposizione 1.2 consente di affermare che $X_{n_k} \to X$ q.c. per $k \to \infty$.

*(c)* Usando la disuguaglianza di Markov,

$$\mathbb{P}(|X_n - X| > \delta) \leq \frac{\mathbb{E}(|X_n - X|^p)}{\delta^p} = \frac{||X_n - X||_p^p}{\delta^p}.$$

Quindi se $||X_n - X||_p \to 0$ allora $\mathbb{P}(|X_n - X| > \delta) \to 0$ per $n \to \infty$, per ogni $\delta > 0$, cioè $X_n \to X$ in probabilità.

*(d)* Posto $U_n = |X_n - X|^p$, basta dimostrare che $\mathbb{E}(U_n) \to 0$ per $n \to \infty$ (si noti infatti che $\mathbb{E}(U_n) = ||X_n - X||_p^p$). Per ipotesi, $U_n \to U = 0$ q.c. Inoltre, $|U_n| \leq c_p(|X_n|^p + |X|^p) \leq 2\,c_p\,Z^p$, con $Z^p \in L^1$ e $c_p$ opportuna. Allora, usando il Teorema della Convergenza Dominata, si ottiene che $\mathbb{E}(U_n) \to \mathbb{E}(U) = 0$. Ricordando infine la Nota 1, la tesi è dimostrata.

■

**Osservazione 1.5** La proposizione precedente si può riassumere come segue:

- convergenza q.c. $\overset{\not\Leftarrow}{\Rightarrow}$ convergenza in probabilità, e convergenza in probabilità $\Rightarrow$ esistenza di una sottosuccessione che converge q.c.;
- convergenza in $L^p$ $\overset{\not\Leftarrow}{\Rightarrow}$ convergenza in probabilità;
- convergenza q.c. $\overset{\not\Leftarrow}{\not\Rightarrow}$ convergenza in $L^p$, ma se la successione è dominata da una v.a. in $L^p$ allora convergenza q.c. $\Rightarrow$ convergenza in $L^p$.

Mostriamo alcuni controesempi che mostrano la validità delle $\not\Rightarrow$ e $\not\Leftarrow$ nell'Osservazione 1.5.

**Esempio 1.6** [convergenza in probabilità, convergenza in $L^p \not\Rightarrow$ convergenza q.c.]
Sia $(\Omega, \mathcal{F}, \mathbb{P}) = ([0,1], \mathcal{B}[0,1], \mathrm{Leb})$ e fissati $m \geq 1$ e $k = 0, 1, 2, \ldots, 2^m - 1$, sia

$$Y_{k,2^m}(\omega) = 1\!\!1_{[k/2^m,(k+1)/2^m]}(\omega).$$

Si ponga ora $X_1 = 1$ e per $n \geq 1$, $X_n = Y_{k,2^m}$, dove $m \geq 1$ e $k \leq 2^m - 1$ sono tali che $n = k + 2^m$.

Mostriamo che $X_n \to 0$ in $L^p$:

$$\mathbb{E}(|X_n|^p) = \mathbb{E}(1\!\!1_{[k/2^m,(k+1)/2^m]}) = \mathbb{P}([k/2^m, (k+1)/2^m]) = \mathrm{Leb}([k/2^m, (k+1)/2^m]) = 1/2^m$$

quindi

$$\lim_{n\to\infty} \mathbb{E}(|X_n|^p) = \lim_{m\to\infty} 1/2^m = 0$$

e, in particolare (parte *(c)* della Proposizione 1.4), $X_n \to 0$ in probabilità[3]. Ma $X_n$ non converge a 0 q.c. Infatti, per ogni $\omega \in \Omega = [0,1]$ e per ogni $m \geq 1$ esiste un $k^* = k^*(\omega, m)$ tale che

[3] Si può anche verificare direttamente. Infatti, se $\delta \geq 1$ allora $\{|X_n| > \delta\} = \emptyset$; se $0 < \delta < 1$, $\{|X_n| > \delta\} = \{1\!\!1_{[k/2^m,(k+1)/2^m]} > \delta\} = \{1\!\!1_{[k/2^m,(k+1)/2^m]} = 1\} = [k/2^m, (k+1)/2^m]$. Quindi, per ogni $\delta > 0$, $\mathbb{P}(|X_n| > \delta) \leq 1/2^m$, che va a 0 per $m \to \infty$, cioè se $n \to \infty$.

$\omega \in [k/2^m, (k+1)/2^m]$. Consideriamo la successione (numerica) $\xi_m = Y_{k^*+2^m}(\omega)$: ovviamente $\xi_m = 1$ per ogni $m$, quindi $\xi_m \to 1$ per $m \to \infty$. Ora, poiché $X_n \leq 1$ per ogni $n$, si ha $\limsup_{n\to\infty} X_n(\omega) \leq 1$. Ma per ogni $\omega$, esiste una sottosuccessione $\xi_m$ di $X_n(\omega)$ tale che $\xi_m \to 1$ per $m \to \infty$. Ciò significa che, per ogni $\omega$ fissato,

$$\limsup_{n\to\infty} X_n(\omega) = 1$$

e quindi $X_n$ non può convergere a 0 q.c.

**Esempio 1.7** [convergenza q.c., convergenza in probabilità$\not\Rightarrow$ convergenza in $L^p$]
Sia ancora $(\Omega, \mathcal{F}, \mathbb{P}) = ([0,1], \mathcal{B}[0,1], \mathrm{Leb})$. Fissato $n \geq 1$, sia

$$X_n(\omega) = n^\alpha \, 1\!\!1_{(0,1/n)}(\omega),$$

dove $\alpha \in \mathbb{R}$ è fissato. Ora, poiché per ogni $\omega \in [0,1]$ esiste $n_0$ tale che per ogni $n > n_0$ si ha $\omega > 1/n$, evidentemente $X_n(\omega) \to 0$ per $n \to \infty$, quindi $X_n \to 0$ q.c. e dunque in probabilità. Ma non è detto che $X_n \to 0$ in $L^p$. Infatti,

$$\mathbb{E}(|X_n|^p) = n^{\alpha p}\, \mathbb{P}((0,1/n)) = n^{\alpha p}\, \mathrm{Leb}((0,1/n)) = n^{\alpha p - 1}.$$

Dunque, $X_n \in L^p$ per ogni $p$ e se si sceglie $\alpha p \geq 1$, cioè $\alpha \geq 1/p$, allora $\mathbb{E}(|X_n|^p)$ non converge a 0, cioè $X_n$ non converge a 0 in $L^p$.

Proponiamo infine il seguente esercizio, che in qualche modo anticipa il contenuto del prossimo paragrafo.

**Esercizio 1.8** *Sia $\{X_n\}_n$ una successione di v.a. i.i.d., con media (comune) $\mu$ e varianza (comune) $\sigma^2$. Posto $S_n = \sum_{k=1}^n X_k$, dimostrare che $S_n/n \to \mu$ in probabilità.*

**Soluzione.** Fissiamo $\delta > 0$ e studiamo $\mathbb{P}(|S_n/n - \mu| > \delta)$. Osserviamo che $\mathbb{E}(S_n/n) = \mu$, quindi stimiamo la probabilità sopra scritta con la disuguaglianza di Chebycev[4]:

$$\mathbb{P}(|S_n/n - \mu| > \delta) \leq \frac{\mathrm{Var}(S_n/n)}{\delta^2}.$$

Ora, ricordando che le $X_k$ sono indipendenti e usando le proprietà della varianza, abbiamo

$$\mathrm{Var}(S_n/n) = \frac{1}{n^2}\mathrm{Var}(\sum_{k=1}^n X_k) = \frac{1}{n^2}\sum_{k=1}^n \mathrm{Var}(X_k) = \frac{\sigma^2}{n},$$

quindi

$$\mathbb{P}(|S_n/n - \mu| > \delta) \leq \frac{\sigma^2}{n\delta^2} \to 0 \quad \text{se } n \to \infty.$$

Osserviamo infine che in realtà abbiamo dimostrato che $S_n/n \to \mu$ in $L^2$, e quindi in probabilità.

∎

[4]**Disuguaglianza di Chebycev.** Sia $Z$ una v.a. con media $m$ e varianza $\rho^2$. Allora, $\mathbb{P}(|Z-m| \geq \delta) \leq \rho^2/\delta^2$. Infatti, per la disuguaglianza di Markov, $\mathbb{P}(|Z-m| \geq \delta) \leq \mathbb{E}(|Z-m|^2)/\delta^2$ e ricordando che $\rho^2 = \mathrm{Var}(Z) = \mathbb{E}((Z-m)^2)$, la disuguaglianza di Chebycev segue immediatamente.

## 2 Legge dei Grandi Numeri

Con il nome di "legge dei Grandi Numeri" si intende lo studio della convergenza della media empirica $S_n/n$ alla media teorica $\mu$ come nell'Esercizio 1.8, dov'è stata dimostrata una versione classica della Legge "Debole" dei Grandi Numeri, così detta perché la convergenza è in probabilità. Più in generale, data una successione $\{X_n\}_n$ di v.a. con medie $\mu_n = \mathbb{E}(X_n)$, studieremo la convergenza a 0 di

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

Parleremo di Legge "Forte" dei Grandi Numeri se la convergenza in questione è quasi certa.

**Teorema 2.1 [Legge forte di Rajchmann]** *Se le v.a. $X_n$ sono a due a due non correlate e se esiste $L > 0$ tale che $sup_n \mathrm{Var}(X_n) \leq L$ allora $S_n/n \to 0$ q.c.*

**Dimostrazione.** Senza perdere in generalità, possiamo supporre $\mu_n = 0$ per ogni $n$ (altrimenti lavoriamo con le v.a. $\tilde{X}_n = X_n - \mu_n$). Poiché le $X_n$ sono a due a due non correlate, si ha $\mathrm{Var}(\sum_{k=1}^n X_k) = \sum_{k=1}^n \mathrm{Var}(X_k)$, quindi come nell'Esercizio 1.8 si ha

$$\mathbb{P}(|S_n/n| > \delta) \leq \frac{\mathrm{Var}(S_n/n)}{\delta^2} = \frac{\sum_{k=1}^n \mathrm{Var}(X_k)}{n^2\delta^2} \leq \frac{L}{n\delta} \tag{1}$$

perché le varianze sono uniformemente limitate. Quindi ancora $S_n/n \to 0$ in $L^2$ e in probabilità. Osserviamo però che la stima appena ottenuta non consente di stabilire che $\sum_n \mathbb{P}(|S_n/n| > \delta) < \infty$ e quindi (Proposizione 1.2) la convergenza q.c. Per tale ragione, consideriamo la sottosuccessione $\{S_{n^2}/n^2\}_n$ di $\{S_n/n\}_n$: da (1) otteniamo, per $\delta > 0$,

$$\mathbb{P}(|S_{n^2}/n^2| > \delta) \leq \frac{L}{n^2\delta}.$$

Quindi $\sum_n \mathbb{P}(|S_{n^2}/n^2| > \delta) < \infty$ e dalla Proposizione 1.2 si ha

$$\lim_{n\to\infty} \frac{1}{n^2} S_{n^2} = 0 \quad \text{ q.c.} \tag{2}$$

Ora, fissato $n$ poniamo

$$D_n = \sup_{n^2 \leq k < (n+1)^2} |S_k - S_{n^2}|.$$

Per ogni $k \geq 1$

$$|S_k| = |S_k - S_{n^2} + S_{n^2}| \leq |S_k - S_{n^2}| + |S_{n^2}|,$$

quindi se $n$ è scelto tale che $n^2 \leq k < (n+1)^2$ si ha

$$\left|\frac{1}{k} S_k\right| \leq \frac{1}{k}\Big(\sup_{n^2 \leq k < (n+1)^2} |S_k - S_{n^2}| + |S_{n^2}|\Big) = \frac{1}{k}\Big(D_n + |S_{n^2}|\Big) \leq \frac{1}{n^2} D_n + \left|\frac{1}{n^2} S_{n^2}\right|.$$

Da (2), $|S_{n^2}/n^2| \to 0$ q.c., quindi basta dimostrare che $D_n/n^2 \to 0$ q.c. Infatti, fissato $\delta > 0$, usando la disuguaglianza di Markov,

$$\mathbb{P}(D_n/n^2 > \delta) = \mathbb{P}(D_n > n^2\delta) \leq \frac{\mathbb{E}(D_n^2)}{n^4\delta^2}. \tag{3}$$

Stimiamo $\mathbb{E}(D_n^2)$:

$$D_n^2 = \sup_{n^2 \le k < (n+1)^2} |S_k - S_{n^2}|^2 \le \sum_{k=n^2}^{(n+1)^2-1} |S_k - S_{n^2}|^2 = \sum_{k=n^2}^{(n+1)^2-1} \Big| \sum_{j=n^2+1}^{k} X_j \Big|^2.$$

Allora,

$$\mathbb{E}(D_n^2) \le \sum_{k=n^2}^{(n+1)^2-1} \mathbb{E}\Big( \Big| \sum_{j=n^2+1}^{k} X_j \Big|^2 \Big) = \sum_{k=n^2}^{(n+1)^2-1} \mathbb{E}\Big( \sum_{j,\ell=n^2+1}^{k} X_j \, X_\ell \Big)$$
$$= \sum_{k=n^2}^{(n+1)^2-1} \Big( \sum_{j=n^2+1}^{k} \mathbb{E}(X_j^2) + \sum_{j,\ell=n^2+1, j\neq\ell}^{k} \mathbb{E}(X_j \, X_\ell) \Big).$$

Poiché $\mathbb{E}(X_n) = 0$, si ha che $\mathbb{E}(X_j^2) = \mathrm{Var}(X_j) \le L$ per ogni $j$ e $\mathbb{E}(X_j \, X_\ell) = \mathrm{Cov}(X_j, X_\ell) = 0$ per $j \neq \ell$, da cui segue che[5]

$$\mathbb{E}(D_n^2) \le \sum_{k=n^2}^{(n+1)^2-1} \sum_{j=n^2+1}^{k} L = L \sum_{k=n^2}^{(n+1)^2-1} (k - n^2)$$
$$\le L \sum_{k=n^2}^{(n+1)^2-1} \Big( (n+1)^2 - 1 - n^2 \Big) \le L \Big( (n+1)^2 - n^2 \Big) \Big( (n+1)^2 - 1 - n^2 \Big) = 2Ln(2n+1).$$

Da (3), si ottiene

$$\mathbb{P}(D_n/n^2 > \delta) \le \frac{2L\, n\, (2n+1)}{n^4 \delta^2} = O\Big(\frac{1}{n^2}\Big),$$

quindi $\sum_n \mathbb{P}(D_n/n^2 > \delta) < \infty$ e, usando la Proposizione 1.2, possiamo concludere che $D_n/n^2 \to 0$ q.c.

∎

Come immediata conseguenza, possiamo rafforzare quanto dimostrato nell'Esercizio 1.8:

**Corollario 2.2** *Se $X_1, X_2, \ldots$ sono v.a. i.i.d. con media $\mu$ e varianza $\sigma^2$ allora*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} X_k = \mu \qquad q.c.$$

**Dimostrazione.** Poiché sono verificate le ipotesi del Teorema 2.1, si ha

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} (X_k - \mu) = 0 \qquad \text{q.c.}$$

e poiché $\frac{1}{n} \sum_{k=1}^{n} \mu = \mu$, la tesi segue immediatamente.

∎

Proponiamo il seguente

[5] Ricordiamo che $\sum_{j=n_1}^{n_2} 1 = n_2 - n_1 + 1$.

**Esercizio 2.3** *Sia $\{X_n\}_n$ una successione di v.a. i.i.d. tali che*

$$\mathbb{P}(X_i > x) = \begin{cases} x^{-\lambda} & se\ x > 1 \\ 1 & se\ x \leq 1 \end{cases}$$

*dove $\lambda > 1$.*

*a) Calcolare media e varianza delle $X_i$, se esistono.*

*b) Posto $Y_i = \log X_i$, determinare la legge di $X_i$.*

*c) Studiare la convergenza q.c. della successione $\{(X_1 X_2 \cdots X_n)^{1/n}\}_n$.*

**Soluzione.** a) La f.d. comune è $F(x) = 1 - \mathbb{P}(X_i > x)$ e per q.o. $x$, $F'(x) = \lambda\, x^{-(\lambda+1)}\, 1\!\!1_{x>1} =: f(x)$. Ora, $\int_{\mathbb{R}} f(x)dx = 1$ quindi $f$ è la densità di probabilità delle $X_i$. Inoltre, si verifica facilmente che $xf(x) \in L^1$ se e solo se $\lambda > 1$ e $x^2 f(x) \in L^2$ se e solo se $\lambda > 2$ e in tal caso,

$$\mathbb{E}(X_i) = \int_{\mathbb{R}} x\, f(x)\, dx = \frac{\lambda}{\lambda - 1} \qquad \mathbb{E}(X_i^2) = \int_{\mathbb{R}} x^2\, f(x)\, dx = \frac{\lambda}{\lambda - 2},$$

quindi se $\lambda > 2$ allora esiste

$$\mathrm{Var}(X_i) = \frac{\lambda}{\lambda - 2} - \Big(\frac{\lambda}{\lambda - 1}\Big)^2 = \frac{\lambda}{(\lambda - 1)^2(\lambda - 2)}.$$

b) Intanto, osserviamo che poiché $X_i \geq 1 > 0$ q.c., le $Y_i$ sono ben poste. Inoltre, è evidente che le $Y_i$ rimangono i.i.d. Poi, detta $G$ la f.d. delle $Y_i$, si ha

$$G(y) = \mathbb{P}(Y_i \leq y) = \mathbb{P}(X_i \leq e^y) = F(e^y)$$

quindi, per q.o. $y$,

$$g(y) := G'(y) = F'(e^y)\, e^y = f(e^y) e^y = \lambda\, x^{-(\lambda+1)}\, 1\!\!1_{x>1}\Big|_{x=e^y} e^y = \lambda\, e^{-\lambda y}\, 1\!\!1_{y>0}$$

che è la densità di una $\mathrm{Exp}(\lambda)$, quindi $Y_i \sim \mathrm{Exp}(\lambda)$.

c) Osserviamo che

$$\log(X_1 X_2 \cdots X_n)^{1/n} = \frac{1}{n} \sum_{k=1}^n \log X_k = \frac{1}{n} \sum_{k=1}^n Y_k$$

e, per ogni $\lambda > 0$ le $Y_k$ sono i.i.d. di media $\mu = 1/\lambda$ e varianza finita. Allora, dal Corollario 2.2 deduciamo che

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^n Y_k = \frac{1}{\lambda} \qquad \text{q.c.}$$

e quindi

$$\lim_{n\to\infty} (X_1 X_2 \cdots X_n)^{1/n} = \lim_{n\to\infty} \exp\Big(\frac{1}{n} \sum_{k=1}^n Y_k\Big) = e^{1/\lambda} \qquad \text{q.c.}$$

perché $x \mapsto e^x$ è una funzione continua.

$\blacksquare$

Enunciamo ora la più celebre Legge dei Grandi Numeri, dove le ipotesi di esistenza dei momenti sono più deboli ma si suppone che le v.a. siano i.i.d.:

**Teorema 2.4** **[Legge forte di Kolmogorov]** *Sia $\{X_n\}_n$ una successione di v.a. i.i.d.*

*(i) Se $X_i \in L^1$, allora detta $\mu = \mathbb{E}(X_i)$ si ha*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n} X_k = \mu \qquad q.c.$$

*(ii) Se $X_i \notin L^1$ allora almeno una delle due variabili terminali*

$$\liminf_{n} \frac{1}{n} \sum_{k=0}^{n} X_k \qquad\qquad \limsup_{n} \frac{1}{n} \sum_{k=0}^{n} X_k$$

*è q.c. infinita.*

Quindi, nel caso di v.a. i.i.d., possiamo dire che: *esiste $\mu = \mathbb{E}(X_i)$ se e solo se esiste $X$ v.a. finita*[6] *tale che $\frac{1}{n}\sum_{k=1}^{n} X_k$ converge q.c. a $X$, e in tal caso necessariamente dev'essere $X = \mu$ q.c.*

Infatti, se esiste $\mu = \mathbb{E}(X_i)$ allora la tesi segue da (i) del Teorema 2.4. Viceversa, se esiste una v.a. $X$ finita che è limite q.c. di $\frac{1}{n}\sum_{k=1}^{n} X_k$, allora da (ii) del Teorema 2.4 segue che $\mu = \mathbb{E}(X_i) < \infty$ perché, se così non fosse, si avrebbe $\mathbb{P}(|X| = +\infty) > 0$.

# 3 Il Teorema di Weierstrass

In questo paragrafo mostreremo con tecniche probabilistiche il

**Teorema 3.1** **[Weierstrass]** *I polinomi sono densi nello spazio delle funzioni continue su* $[0, 1]$ *dotato della norma del* sup.

In altre parole, presa $f \in C([0,1])$ allora esiste una successione di polinomi $\{P_n\}_n$ tali che

$$\lim_{n\to\infty} \sup_{x\in[0,1]} |P_n(x) - f(x)| = 0. \tag{4}$$

Inoltre, daremo a $P_n$ un'espressione esplicita. Cominciamo infatti a dimostrare il seguente risultato, conseguenza della Legge dei Grandi Numeri:

**Proposizione 3.2** *Sia $f \in C([0,1])$ e, per $x \in [0,1]$ e $n \geq 1$, sia*

$$P_n(x) = \sum_{k=0}^{n} \binom{n}{k} f\Big(\frac{k}{n}\Big) x^k (1-x)^{n-k}. \tag{5}$$

*Allora, $\lim_{n\to\infty} P_n(x) = f(x)$.*

[6] Con "$X$ è una v.a. finita" intendiamo $\mathbb{P}(|X| < +\infty) = 1$, o equivalentemente $\mathbb{P}(|X| = +\infty) = 0$.

**Dimostrazione.** Fissiamo $f \in C([0,1])$ e $x_0 \in [0,1]$. Sia $\{Z_n\}_n$ una successione di v.a. i.i.d., con $Z_k \sim \mathrm{Be}(x_0)$. Posto $\bar{Z}_n = \frac{1}{n}\sum_{k=1}^n Z_k$ allora la Legge Forte dei Grandi Numeri stabilita nel Corollario 2.2 assicura che

$$\lim_{n\to\infty} \bar{Z}_n = \mathbb{E}(Z_1) = x_0 \quad \text{q.c.}$$

Ora, poiché $f$ è continua in $x_0$, si ha anche

$$\lim_{n\to\infty} f(\bar{Z}_n) = f(x_0) \quad \text{q.c.}$$

Inoltre, essendo $f$ continua sul compatto $[0,1]$, è anche limitata, quindi usando la Proposizione 1.4 (parte (d)), $f(\bar{Z}_n)$ converge a $f(x_0)$ anche in $L^1$, quindi in particolare

$$\lim_{n\to\infty} \mathbb{E}(f(\bar{Z}_n)) = \mathbb{E}(f(x_0)) = f(x_0).$$

Mostriamo infine che $\mathbb{E}(f(\bar{Z}_n)) = P_n(x_0)$. Infatti, detta $V_n = \sum_{k=0}^n Z_k$ con $Z_1, \dots, Z_n$ i.i.d. bernoulliane di parametro $x_0$, allora $V_n \sim \mathrm{Bi}(n, x_0)$:

$$\mathbb{P}(V_n = k) = \binom{n}{k} x_0^k (1-x_0)^{n-k} \qquad k = 0, 1, \dots, n$$

quindi

$$\mathbb{E}(f(\bar{Z}_n)) = \mathbb{E}\Big(f\Big(\frac{V_n}{n}\Big)\Big) = \sum_{k=0}^n f\Big(\frac{k}{n}\Big)\mathbb{P}(V_n = k) = \sum_{k=0}^n f\Big(\frac{k}{n}\Big) \binom{n}{k} x_0^k (1-x_0)^{n-k} = P_n(x_0)$$

e la tesi è dimostrata.

∎

La proposizione precedente garantisce che il polinomio $P_n$ converge a $f$ puntualmente. Dimostriamo ora che la convergenza è uniforme, come stabilito dal Teorema di Weierstrass.

**Dimostrazione del Teorema 3.1.** Presa $f \in C([0,1])$, sia $\{P_n\}_n$ la successione di polinomi definita tramite (5) e dimostriamo che vale la (4).

Nel corso della dimostrazione della Proposizione 3.2 abbiamo visto che

$$|P_n(x) - f(x)| = |\mathbb{E}(f(\bar{Z}_n^x)) - f(x)|$$

dove con $\bar{Z}_n^x$ indichiamo la v.a. costruita precedentemente, a partire da v.a. $Z_1^x, \dots, Z_n^x$ i.i.d. bernoulliane di parametro $x \in [0,1]$. Quindi,

$$|P_n(x) - f(x)| \le \mathbb{E}(|f(\bar{Z}_n^x) - f(x)|) = \mathbb{E}(|f(\bar{Z}_n^x) - f(x)|\, 1\!\!1_{|\bar{Z}_n^x - x| < \delta}) + \mathbb{E}(|f(\bar{Z}_n^x) - f(x)|\, 1\!\!1_{|\bar{Z}_n^x - x| \ge \delta})$$

dove $\delta$ denota una costante positiva. Posto $\alpha_n(x;\delta) = \mathbb{E}(|f(\bar{Z}_n^x) - f(x)|\, 1\!\!1_{|\bar{Z}_n^x - x| < \delta})$ e $\beta_n(x;\delta) = \mathbb{E}(|f(\bar{Z}_n^x) - f(x)|\, 1\!\!1_{|\bar{Z}_n^x - x| \ge \delta})$, possiamo scrivere

$$|P_n(x) - f(x)| \le \alpha_n(x;\delta) + \beta_n(x;\delta)$$

Studiamo le quantità $\alpha_n$ e $\beta_n$. Poiché $f$ è continua su $[0,1]$, è anche uniformemente continua: preso un arbitrario $\varepsilon > 0$ esiste un $\delta^*$ tale che

$$\sup_{z,y\,:\,|z-y|<\delta^*} |f(z) - f(y)| < \varepsilon.$$

Scelto $\delta = \delta^*$, per ogni $\omega$ e $x$ tali che $|\bar{Z}_n^x(\omega) - x| < \delta^*$ si ha che $|f(\bar{Z}_n^x) - f(x)| < \varepsilon$, quindi

$$\sup_{x\in[0,1]} |f(\bar{Z}_n^x) - f(x)|\, 1\!\!1_{|\bar{Z}_n^x - x| < \delta^*} < \varepsilon$$

e

$$\sup_{x\in[0,1]} \alpha_n(x;\delta^*) \le \varepsilon.$$

Possiamo allora scrivere

$$\sup_{x\in[0,1]} |P_n(x) - f(x)| \le \sup_{x\in[0,1]} \alpha_n(x;\delta^*) + \sup_{x\in[0,1]} \beta_n(x;\delta^*) \le \varepsilon + \sup_{x\in[0,1]} \beta_n(x;\delta^*). \tag{6}$$

Ora, $f$ è limitata, e sia $M$ tale che $|f(x)| \le M$ per ogni $x$. Quindi $|f(\bar{Z}_n^x) - f(x)| \le 2M$ e

$$\begin{aligned}\beta_n(x;\delta^*) &= \mathbb{E}(|f(\bar{Z}_n^x) - f(x)|\, 1\!\!1_{|\bar{Z}_n^x - x| \ge \delta^*}) \\ &\le 2M\mathbb{E}(1\!\!1_{|\bar{Z}_n^x - x| \ge \delta^*}) = 2M\mathbb{P}(|\bar{Z}_n^x - x| \ge \delta^*) \le 2M\frac{\mathrm{Var}(\bar{Z}_n^x)}{\delta^{*2}}\end{aligned}$$

dove si è usata la disuguaglianza di Chebycev. Ma[7]

$$\mathrm{Var}(\bar{Z}_n^x) = \mathrm{Var}\Big(\frac{1}{n}\sum_{k=1}^n Z_k^x\Big) = \frac{1}{n^2}\, n\, x(1-x)$$

e poiché $x(1-x) \le 1/4$ per ogni $x \in [0,1]$, si ha

$$\sup_{x\in[0,1]} \beta_n(x;\delta^*) \le \frac{M}{2n\delta^{*2}}.$$

Da (6) segue che

$$\sup_{x\in[0,1]} |P_n(x) - f(x)| \le \varepsilon + \frac{M}{2n\delta^{*2}}.$$

Passando al limite per $n \to \infty$, si ottiene

$$\lim_{n\to\infty} \sup_{x\in[0,1]} |P_n(x) - f(x)| \le \varepsilon$$

e data l'arbitrarietà di $\varepsilon$, si può concludere che

$$\lim_{n\to\infty} \sup_{x\in[0,1]} |P_n(x) - f(x)| = 0,$$

da cui la tesi.

$\blacksquare$

[7] Ricordiamo che se $Z \sim \mathrm{Be}(p)$, $\mathrm{Var}(Z) = p(1-p)$.