Teoremi asintotici

9.1 Successioni di variabili aleatorie

In questo capitolo vogliamo analizzare alcuni dei principali problemi asintotici, di rilevante interesse e applicabilità in statistica, coinvolgenti collezioni di un numero infinitamente grande di variabili aleatorie, ordinate in maniera da costituire delle successioni. Ricordiamo che quando prendiamo in esame più variabili aleatorie, occorre supporre che esse siano definite in uno stesso spazio di probabilità $(\Omega, \mathscr{F}, P)$. Una successione di variabili aleatorie va quindi riguardata come una successione $\{X_n(\omega);\ n = 1, 2, \ldots\}$ di funzioni misurabili da $(\Omega, \mathscr{F})$ in $(\mathbb{R}, \mathscr{B})$. Ricordiamo inoltre (v. Paragrafo 3.8) che una successione di variabili aleatorie $X_1, X_2, \ldots$, tutte definite in uno stesso spazio di probabilità, si dicono indipendenti se e solo se per ogni insieme finito di indici distinti $i_1, i_2, \ldots, i_k$ le variabili $X_{i_1}, X_{i_2}, \ldots, X_{i_k}$ sono indipendenti.

Discuteremo dapprima alcuni tipi di convergenza per successioni di variabili aleatorie e successivamente esporremo alcuni dei principali teoremi limite della probabilità.

9.2 Convergenza di variabili aleatorie

Ci proponiamo in questo paragrafo di analizzare il comportamento limite di successioni di variabili aleatorie mediante lo studio dei seguenti tipi di convergenza: *convergenza quasi certa*, *convergenza in probabilità* e *convergenza in distribuzione*.

Un criterio naturale di convergenza che emerge dalle familiari nozioni dell'analisi matematica è la *convergenza su* Ω, così definita: per ogni $\varepsilon > 0$ e per ogni $\omega \in \Omega$ esiste un intero positivo $k = k(\varepsilon, \omega)$ tale da aversi

$$|X_n(\omega) - X(\omega)| < \varepsilon \qquad \text{per ogni } n > k.$$

In altri termini, si richiede che la successione $X_n(\omega)$ converga ad $X(\omega)$ per ogni $\omega \in \Omega$.

In un contesto di tipo probabilistico risulta però eccessivo richiedere che la convergenza sussista per ogni $\omega \in \Omega$; è sufficiente che essa sussista per tutti i punti ω il cui insieme costituisce un evento di probabilità unitaria. Da ciò scaturisce quindi la seguente definizione:

Definizione 9.1 *Una successione di variabili aleatorie $X_1, X_2, \ldots$, tutte definite nello stesso spazio di probabilità $(\Omega, \mathscr{F}, P)$, converge quasi certamente ad una variabile aleatoria X, definita nello stesso spazio di probabilità, se risulta:*

$$\begin{aligned} P\Big(\lim_{n\to+\infty} X_n = X\Big) &= P\Big(\Big\{\omega \in \Omega :\ \lim_{n\to+\infty} X_n(\omega) = X(\omega)\Big\}\Big) \\ &= P\big(\{\omega \in \Omega :\ \forall\varepsilon > 0\ \ \exists k :\ \forall n \geq k\ \ |X_n(\omega) - X(\omega)| < \varepsilon\}\big) = 1. \end{aligned} \tag{9.1}$$

Tale convergenza, detta quasi certa, sarà denotata con $X_n \xrightarrow{q.c.} X$. Essa può essere così interpretata: la probabilità dell'evento $\Big\{\omega \in \Omega :\ \lim_{n\to+\infty} X_n(\omega) = X(\omega)\Big\}$ è unitaria, ossia X_n al crescere di n tende a identificarsi con X con probabilità 1. È poi possibile dimostrare il seguente teorema:

Teorema 9.1 *La successione di variabili aleatorie $X_1, X_2, \ldots$ converge quasi certamente a X se e solo se per ogni $\varepsilon > 0$ si ha*

$$\lim_{n\to+\infty} P\left(\bigcap_{k=n}^{+\infty}\Big\{\omega \in \Omega:\ |X_k(\omega) - X(\omega)| < \varepsilon\Big\}\right) = 1. \tag{9.2}$$

Introduciamo ora una ulteriore forma di convergenza, detta convergenza in probabilità, che sarà denotata con la scrittura $X_n \xrightarrow{P} X$.

Definizione 9.2 *Date una successione di variabili aleatorie $X_1, X_2, \ldots$ ed una variabile aleatoria X, tutte definite nello stesso spazio di probabilità $(\Omega, \mathscr{F}, P)$, la successione $X_1, X_2, \ldots$ si dice convergere in probabilità a X se per ogni $\varepsilon > 0$ risulta:*

$$\lim_{n\to+\infty} P(|X_n - X| < \varepsilon) = \lim_{n\to+\infty} P\big(\{\omega \in \Omega:\ |X_n(\omega) - X(\omega)| < \varepsilon\}\big) = 1. \tag{9.3}$$

Teorema 9.2 *Date una successione di variabili aleatorie $X_1, X_2, \ldots$ ed una variabile aleatoria X, tutte definite nello stesso spazio di probabilità $(\Omega, \mathscr{F}, P)$, si ha:*

$$X_n \xrightarrow{q.c.} X \quad\Longrightarrow\quad X_n \xrightarrow{P} X.$$

Un altro tipo di convergenza è la convergenza in distribuzione, che definiamo qui di seguito.

Definizione 9.3 *Date una successione di variabili aleatorie $X_1, X_2, \ldots$ con rispettive funzioni di distribuzione $F_{X_1}(x), F_{X_2}(x), \ldots$ ed una variabile aleatoria X con funzione di distribuzione $F_X(x)$, la successione $X_1, X_2, \ldots$ converge in distribuzione a X se*

$$\lim_{n\to+\infty} F_{X_n}(x) = F_X(x) \tag{9.4}$$

in tutti i punti $x \in \mathbb{R}$ in cui $F_X(x)$ è continua.

La scrittura $X_n \xrightarrow{d} X$ denota la convergenza in distribuzione ad X della successione $X_1, X_2, \ldots$.

Teorema 9.3 *Date una successione di variabili aleatorie $X_1, X_2, \ldots$ ed una variabile aleatoria X tutte definite nello stesso spazio di probabilità $(\Omega, \mathscr{F}, P)$, sussiste la seguente implicazione:*

$$X_n \xrightarrow{P} X \quad\Longrightarrow\quad X_n \xrightarrow{d} X.$$

Inoltre, se X è una variabile aleatoria degenere si ha:

$$X_n \xrightarrow{d} X \quad\Longrightarrow\quad X_n \xrightarrow{P} X.$$

9.3 Teorema centrale di convergenza

Vogliamo ora introdurre uno dei più importanti risultati della teoria della probabilità, noto quale *teorema centrale di convergenza* o *teorema centrale del limite*, che fornisce una semplice ed utile approssimazione alla distribuzione della somma di variabili aleatorie indipendenti, evidenziando al contempo la grande importanza della distribuzione normale.

Teorema 9.8 (Teorema centrale di convergenza) *Sia $X_1, X_2, \ldots$ una successione di variabili aleatorie, definite nello stesso spazio di probabilità, indipendenti e identicamente distribuite con valore medio μ finito e varianza σ^2 finita e positiva. Posto per ogni intero n*

positivo $Y_n = X_1 + X_2 + \ldots + X_n$, *per ogni* $x \in \mathbb{R}$ *risulta:*

$$\lim_{n\to+\infty} P\left(\frac{Y_n - n\,\mu}{\sigma\,\sqrt{n}} \le x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2}\,dy = \Phi(x), \tag{9.7}$$

ossia la successione delle variabili aleatorie standardizzate

$$Z_n = \frac{Y_n - E(Y_n)}{\sqrt{\mathrm{Var}(Y_n)}} = \frac{Y_n - n\,\mu}{\sqrt{n}\,\sigma}, \qquad n = 1, 2, \ldots$$

converge in distribuzione alla variabile aleatoria normale standard.

Il Teorema 9.8 mostra inoltre che sottraendo a $Y_n = X_1 + X_2 + \ldots + X_n$ la sua media $n\,\mu$ e dividendo la differenza per la deviazione standard di Y_n, ossia per $\sigma\,\sqrt{n}$, si ottiene una variabile aleatoria standardizzata Z_n la cui funzione di distribuzione $P(Z_n \le x)$ è per n sufficientemente grande approssimativamente normale standard:

$$P\left(\frac{Y_n - n\,\mu}{\sigma\,\sqrt{n}} \le x\right) \simeq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2}\,dy = \Phi(x) \qquad (x \in \mathbb{R}). \tag{9.12}$$

Notiamo inoltre che per ogni coppia di reali α, β, con $\alpha < \beta$, dalla (9.7) segue anche

$$\lim_{n\to+\infty} P\left(\alpha < \frac{Y_n - n\,\mu}{\sigma\,\sqrt{n}} \le \beta\right) = \Phi(\beta) - \Phi(\alpha), \tag{9.13}$$

così che, per n sufficientemente grande dalla (9.13) si ha:

$$P\left(\alpha < \frac{Y_n - n\,\mu}{\sigma\,\sqrt{n}} \le \beta\right) \simeq \Phi(\beta) - \Phi(\alpha) \qquad (\alpha, \beta \in \mathbb{R}, \alpha < \beta). \tag{9.14}$$

Va menzionato che la bontà delle approssimazioni (9.12) e (9.14) dipende da n, dal tipo di distribuzione delle variabili $X_1, X_2, \ldots, X_n$ e dai valori degli estremi x, α, β. Si mostra poi che per n grande l'approssimazione migliora al crescere di n. Nelle applicazioni spesso si verifica che essa è già soddisfacente per $n \ge 30$.

Ponendo $w = n\,\mu + x\,\sigma\,\sqrt{n}$ nella (9.12), $a = n\,\mu + \alpha\,\sigma\,\sqrt{n}$ e $b = n\,\mu + \beta\,\sigma\,\sqrt{n}$ nella (9.14) si ricava:

$$P(Y_n \le w) \simeq \Phi\left(\frac{w - n\,\mu}{\sigma\,\sqrt{n}}\right) \qquad (w \in \mathbb{R}), \tag{9.15}$$

$$P(a < Y_n \le b) \simeq \Phi\left(\frac{b - n\,\mu}{\sigma\,\sqrt{n}}\right) - \Phi\left(\frac{a - n\,\mu}{\sigma\,\sqrt{n}}\right) \qquad (a, b \in \mathbb{R}, a < b). \tag{9.16}$$

Per n grande la distribuzione di $Y_n = X_1 + X_2 + \ldots + X_n$ è dunque approssimativamente normale con valore medio $n\,\mu$ e varianza $n\,\sigma^2$. Si noti che nelle (9.15) e (9.16) le approssimazioni dipendono soltanto dal valore medio μ e dalla varianza σ^2 di ciascuna delle variabili $X_1, X_2, \ldots, X_n$.

Esempio 9.9 Supponiamo che i tempi, in fissate unità arbitrarie, di esecuzione di n programmi su di un elaboratore elettronico possano essere riguardati come variabili aleatorie $X_1, X_2, \ldots, X_n$ indipendenti e identicamente distribuite aventi distribuzione esponenziale di parametro $\lambda = 2$. Nell'ipotesi che $n = 40$, determiniamo un'approssimazione della probabilità che il tempo totale di esecuzione degli n programmi, indicato con $Y_n = X_1 + X_2 + \ldots + X_n$, sia compreso tra 15 e 30.

Per la Proposizione 4.16 risulta $Y_n \sim \mathcal{E}(n, \lambda)$, con $n = 40$ e $\lambda = 2$ di modo che, ricordando la (4.39), si ha:

$$P(15 < Y_{40} \leq 30) = \int_{15}^{30} \frac{2^{40}}{39!}\, x^{39}\, e^{-2x}\, dx = 0.9512. \tag{9.17}$$

Facendo ora uso della (9.16), e ricordando che per una distribuzione esponenziale di parametro $\lambda = 2$ risulta $\mu = \sigma = 1/\lambda = 0.5$, si ottiene:

$$\begin{aligned} P(15 < Y_{40} \leq 30) &\simeq \Phi\left(\frac{30 - 40 \cdot 0.5}{0.5\,\sqrt{40}}\right) - \Phi\left(\frac{15 - 40 \cdot 0.5}{0.5\,\sqrt{40}}\right) = \Phi(3.16) - \Phi(-1.58) \\ &= \Phi(3.16) - [1 - \Phi(1.58)] = 0.9992 - (1 - 0.9429) = 0.9421, \end{aligned}$$

dove si è fatto uso della relazione $\Phi(x) = 1 - \Phi(-x)$ e della Tabella C.1 dell'Appendice. Si noti che il valore approssimato 0.9421 appena ricavato si discosta per meno del 2% dal valore indicato nella (9.17). $\diamondsuit$

Esempio 9.10 Supponiamo che $X_1, X_2, \ldots$ sia una successione di variabili aleatorie indipendenti con $X_i \sim \mathcal{U}(-1/2, 1/2)$ per $i = 1, 2, \ldots$. Come mostrato nel Paragrafo 5.8.3 risulta $E(X_i) = 0$ e $\mathrm{Var}(X_i) = 1/12$. Posto $Y_n = X_1 + X_2 + \ldots + X_n$, dalla (9.12) segue:

$$P\Big(\frac{Y_n}{\sqrt{n/12}} \leq x\Big) \simeq \Phi(x) \qquad (x \in \mathbb{R}).$$

Se, in particolare, si sceglie $n = 12$, ossia se si considera la somma di 12 variabili aleatorie indipendenti e uniformi in $(-1/2, 1/2)$, si ha:

$$P(Y_{12} \leq x) \simeq \Phi(x) \qquad (x \in \mathbb{R}).$$

Il teorema centrale di convergenza mostra quindi che la funzione di distribuzione della somma di 12 variabili aleatorie indipendenti e uniformi in $(-1/2, 1/2)$ è approssimativamente normale standard. È qui opportuno notare che la simmetria della densità delle X_i concorre a far sì che con sole 12 variabili si ottenga una buona approssimazione alla normale standard. Questo risultato trova applicazione nella costruzione di un semplice algoritmo per la simulazione su elaboratori elettronici di numeri casuali di una variabile a distribuzione normale standard. A tal proposito, va menzionato che gli elaboratori posseggono routine atte a generare numeri casuali tratti da una variabile che con ottima approssimazione è $U \sim \mathcal{U}(0, 1)$. D'altronde, tramite la trasformazione $V = U - 1/2$ è possibile generare numeri tratti da una variabile $V \sim \mathcal{U}(-1/2, 1/2)$. Pertanto, se $U_1, U_2, \ldots, U_{12}$ sono indipendenti con $U_i \sim \mathcal{U}(0, 1)$, allora risulta $Y_{12} = U_1 + U_2 + \ldots + U_{12} - 6 = V_1 + V_2 + \ldots + V_{12}$, con $V_i \sim \mathcal{U}(-1/2, 1/2)$. Per il teorema centrale di convergenza segue che numeri generati da una variabile normale standard possono essere approssimativamente ottenuti generando numeri a partire dalla variabile Y_{12}. $\diamondsuit$

9.4 Cenni alle leggi dei grandi numeri

Le leggi dei grandi numeri rivestono ruolo fondamentale nel calcolo delle probabilità e in statistica. Esamineremo in questo paragrafo innanzitutto tre diverse versioni della cosiddetta *legge debole* dei grandi numeri dovute rispettivamente a Markov, Chebyshev e Khintchin.

Teorema 9.9 (Legge debole dei grandi numeri di Markov) *Sia $X_1, X_2, \ldots$ una successione di variabili aleatorie indipendenti aventi valori medi finiti e varianze $\sigma_1^2, \sigma_2^2, \ldots$ finite, e sia $Y_n = Y_! + Y_2 + \ldots + X_n$ per $n = 1, 2, \ldots$. Se risulta*

$$\lim_{n \to +\infty} \frac{\sigma_1^2 + \ldots + \sigma_n^2}{n^2} = 0, \tag{9.24}$$

allora per ogni $\varepsilon > 0$ si ha:

$$\lim_{n \to +\infty} P\Big\{\Big|\frac{Y_n - E(Y_n)}{n}\Big| < \varepsilon\Big\} = 1, \tag{9.25}$$

ovvero:

$$\frac{Y_n - E(Y_n)}{n} \xrightarrow{P} 0.$$

Teorema 9.10 (Legge debole dei grandi numeri di Chebyshev) *Sia $X_1, X_2, \ldots$ una successione di variabili aleatorie indipendenti aventi valori medi finiti e varianze $\sigma_1^2, \sigma_2^2, \ldots$ uniformemente limitate e sia $Y_n = X_1 + X_2 + \ldots + X_n$. Allora, per ogni $\varepsilon > 0$*

$$\frac{Y_n - E(Y_n)}{n} \xrightarrow{P} 0.$$

Teorema 9.11 (Legge debole dei grandi numeri di Khintchin) *Sia $X_1, X_2, \ldots$ una successione di variabili aleatorie indipendenti e identicamente distribuite, dotate di valore medio μ finito. Allora per ogni $\varepsilon > 0$ si ha:*

$$\lim_{n \to +\infty} P\big\{\big|\overline{X}_n - \mu\big| < \varepsilon\big\} = 1, \tag{9.26}$$

ossia $\overline{X}_n \xrightarrow{P} \mu$.