Capitolo 1

Il modello di Black e Scholes come limite del modello binomiale multiperiodale

1.1 Il Modello Binomiale Multiperiodale

Ricordiamo brevemente il Modello Binomiale Multiperiodale (o Cox-Ross-Rubinstein)

IPOTESI e NOTAZIONI:

$$u = 1 + b, \qquad d = 1 + a$$

$$d < 1 + r < u$$

o equivalentemente

$$a < r < b.$$

$$S_0 = s_0 > 0, \qquad S_{n+1} = (1 + \rho_{n+1})S_n = Z_{n+1}S_n$$

o in altre parole

$$S_0 = s_0 > 0, \qquad S_1 = (1 + \rho_1)S_0 = Z_1 S_0 \quad \ldots \quad S_n = (1 + \rho_n) \cdots (1 + \rho_2)(1 + \rho_1)S_0 = Z_n \cdots Z_2 Z_1 S_0$$

Si definisca

$$\xi_i = \mathbf{1}_{\{\rho_i = b\}} = \mathbf{1}_{\{Z_i = u\}}$$

ovvero la variabile aleatoria che vale 1 se il prezzo dell'azione sale e zero altrimenti, in modo che[1]

$$Z_i = u^{\xi_i}\, d^{1-\xi_i}.$$

Sia $H_n(\omega)$ la v.a. che conta il numero delle volte in cui il prezzo sale tra il passo 1 e il passo n, ossia

$$H_k(\omega) = \sum_{i=1}^{k} \mathbf{1}_{\{\rho_i = b\}} = \sum_{i=1}^{k} \mathbf{1}_{\{Z_i = u\}} = \sum_{i=1}^{k} \xi_i$$

[1]Il fatto che $Z_i = u^{\xi_i}\, d^{1-\xi_i}$ si verifica per ispezione:

$$Z_i = u \Leftrightarrow \xi_i = 1 \Leftrightarrow u^{\xi_i}\, d^{1-\xi_i} = u^1\, d^{1-1} = u$$

$$Z_i = d \Leftrightarrow \xi_i = 0 \Leftrightarrow u^{\xi_i}\, d^{1-\xi_i} = u^0\, d^{1-0} = d$$

in modo che

$$S_N = S_0\, u^{H_N}\, d^{N-H_N} = S_0\, \left(\frac{u}{d}\right)^{H_N}\, d^N$$
$$= S_0\,(1+b)^{H_N}\,(1+a)^{N-H_N} = S_0\, \left(\frac{1+b}{1+a}\right)^{H_N}\, (1+a)^N$$

per ogni N e per ogni pay-off terminale[2] f_N $\mathcal{F}_N$-misurabile il prezzo di esercizio può essere descritto dalla formula

$$C(f_N,\mathbb{P}) = B_0\,\widetilde{\mathbb{E}}\left[\frac{f_N}{B_N}\right] = \widetilde{\mathbb{E}}\left[\frac{f_N}{(1+r)^N}\right]. \tag{1.1}$$

dove $\widetilde{\mathbb{E}}$ è il valore atteso rispetto alla probabilità $\widetilde{\mathbb{P}}$ rispetto alla quale gli eventi $\{Z_i = u\}$ sono indipendenti e con probabilità

$$\tilde{p} = \frac{1+r-d}{u-d} = \frac{1+r-(1+a)}{(1+b)-(1+a)} = \frac{r-a}{b-a}.$$

Allora

$$C_{call}(K,\mathbb{P}) = \widetilde{\mathbb{E}}\left[\frac{(S_N-K)^+}{B_N}\right] \tag{1.2}$$
$$= S_0 \sum_{h=h_0}^{N} \binom{N}{h} \widetilde{p}^{\,h}(1-\widetilde{p})^{N-h} \left(\frac{1+b}{1+r}\right)^h \left(\frac{1+a}{1+r}\right)^{N-h}$$
$$- \frac{K}{(1+r)^N} \sum_{h=h_0}^{N} \binom{N}{h} \widetilde{p}^{\,h}(1-\widetilde{p})^{N-h} \tag{1.3}$$
$$= S_0 \sum_{h=h_0}^{N} \binom{N}{h} \left(\frac{1+b}{1+r}\,\widetilde{p}\right)^h \left(\frac{1+a}{1+r}\,(1-\widetilde{p})\right)^{N-h}$$
$$- \frac{K}{(1+r)^N} \sum_{h=h_0}^{N} \binom{N}{h} \widetilde{p}^{\,h}(1-\widetilde{p})^{N-h}. \tag{1.4}$$

1.2 Approssimazione del Modello Binomiale Multiperiodale

Si consideri ora il caso in cui gli scambi avvengono sempre più vicini nel tempo ovvero ai tempi $t_k^{(n)} = k/n$.

Considereremo il tempo continuo, ma, per n fissato, i processi che ci interessano sono costanti negli intervalli tra un tempo $t_k^{(n)} = k/n$ e l'altro.

Continueremo ad indicare con B_k ed S_k il prezzo del titolo non rischioso (conto in banca) e del titolo rischioso (l'azione) rispettivamente, anche se, per mettere in evidenza la dipendenza dal parametro n sarebbe più opportuno denotarli con $B_k^{[n]}$ e $S_k^{[n]}$.

Ovviamente è necessario che il tasso di interesse sia proporzionale all'ampiezza degli intervalli $[t_k^{(n)}, t_{k+1}^{(n)}) = [\frac{k}{n}, \frac{k+1}{n})$, ossia si abbia

$$B^{(n)}_{t_k^{(n)}} = B_k\big(= B_k^{[n]}\big) = B_0\left(1+\frac{r}{n}\right)^k$$

e che rimanga costante in tutto l'intervallo $[t_k^{(n)}, t_{k+1}^{(n)})$, ovvero che

$$B_t^{(n)} = B^{(n)}_{t_k^{(n)}} = B_0\left(1+\frac{r}{n}\right)^k \qquad \text{per } t \in [t_k^{(n)}, t_{k+1}^{(n)}) = [\tfrac{k}{n}, \tfrac{k+1}{n})$$

[2]Si ricorda che stiamo trattando obbligazioni derivate di tipo europeo, che possono essere esercitate solo al tempo finale N, o tempo si esercizio, al contrario di quelle di tipo americano, che invece possono essere esercitate in un qualunque istante tra l'inizio del contratto e il tempo di esercizio.

o in altre parole che, se $N_n(t) = \lfloor nt \rfloor$ è il numero di intervalli di ampiezza $1/n$ che si trovano nell'intervallo $[0,t]$, allora

$$B_t^{(n)} = B_0 \Big(1 + \frac{r}{n}\Big)^{N_n(t)}$$

o meglio

$$B_t^{(n)} = B^{(n)}_{\frac{\lfloor nt \rfloor}{n}} = B_0 \Big(1 + \frac{r}{n}\Big)^{\lfloor nt \rfloor}. \tag{1.5}$$

Inoltre è ragionevole pensare che i cambiamenti del prezzo si discostino di poco in un intervallo di tempo così piccolo. Più precisamente si suppone che

$$S^{(n)}_{t_k^{(n)}} = Z_k^{(n)}\, S^{(n)}_{t_{k-1}^{(n)}} = \Big(1 + \rho_k^{(n)}\Big)\, S^{(n)}_{t_{k-1}^{(n)}}$$

dove i valori ammissibili per $Z_k^{(n)}$ sono solo

$$u^{(n)} = e^{\sigma/\sqrt{n}} \qquad d^{(n)} = e^{-\sigma/\sqrt{n}}.$$

In altri termini si suppone che

$$u^{(n)} = 1 + b^{(n)} \qquad d^{(n)} = 1 + a^{(n)}.$$

dove

$$b^{(n)} = \frac{\sigma}{\sqrt{n}} + \tfrac{1}{2}\,\frac{\sigma^2}{n} + o(\tfrac{1}{n}) \qquad a^{(n)} = -\frac{\sigma}{\sqrt{n}} + \tfrac{1}{2}\,\frac{\sigma^2}{n} + o(\tfrac{1}{n})$$

e che

$$S_t^{(n)} = S^{(n)}_{t_k^{(n)}}, \qquad \text{per } t \in [t_k^{(n)}, t_{k+1}^{(n)}) = [\tfrac{k}{n}, \tfrac{k+1}{n})$$

ovvero, se come prima, $N_n(t) = \lfloor nt \rfloor$ è il numero di intervalli di ampiezza $1/n$ che si trovano nell'intervallo $[0,t]$, allora

$$S_t^{(n)} = S_{\frac{\lfloor nt \rfloor}{n}} \big(= S^{[n]}_{\frac{\lfloor nt \rfloor}{n}} \big) = S_0 \left(\frac{u^{(n)}}{d^{(n)}}\right)^{H_{N_n(t)}} \Big(d^{(n)}\Big)^{\lfloor N_n(t) \rfloor}$$

o meglio

$$S_t^{(n)} = S^{(n)}_{\frac{\lfloor nt \rfloor}{n}} = S_0 \left(\frac{u^{(n)}}{d^{(n)}}\right)^{H_{\lfloor nt \rfloor}} \Big(d^{(n)}\Big)^{\lfloor nt \rfloor}.$$

Tenendo presente che

$$\frac{u^{(n)}}{d^{(n)}} = \frac{e^{\sigma/\sqrt{n}}}{e^{-\sigma/\sqrt{n}}} = e^{2\sigma/\sqrt{n}},$$

si può anche scrivere

$$\begin{aligned} S_t^{(n)} &= S_0\, e^{2\sigma \frac{1}{\sqrt{n}} H_{\lfloor nt \rfloor}}\, e^{-\sigma \frac{1}{\sqrt{n}} \lfloor nt \rfloor} \\ &= S_0\, e^{\sigma \frac{1}{\sqrt{n}} \left(2\, H_{\lfloor nt \rfloor} - \lfloor nt \rfloor\right)}. \end{aligned} \tag{1.6}$$

Sappiamo che rispetto alla misura martingala equivalente $\widetilde{\mathbb{P}}^{(n)}$ rispetto alla quale gli eventi $\{Z_i^{(n)} = u\}$ sono indipendenti e con probabilità

$$\tilde{p}^{(n)} = \frac{1 + \frac{r}{n} - d^{(n)}}{u^{(n)} - d^{(n)}} = \frac{\frac{r}{n} - a^{(n)}}{b^{(n)} - a^{(n)}} = \frac{\frac{r}{n} - (-\frac{\sigma}{\sqrt{n}} + \frac{1}{2}\,\frac{\sigma^2}{n} + o(\frac{1}{n}))}{\frac{\sigma}{\sqrt{n}} + \frac{1}{2}\,\frac{\sigma^2}{n} + o(\frac{1}{n}) - (-\frac{\sigma}{\sqrt{n}} + \frac{1}{2}\,\frac{\sigma^2}{n} + o(\frac{1}{n}))},$$

o meglio

$$\tilde{p}^{(n)} = \frac{\frac{\sigma}{\sqrt{n}} + \frac{1}{n}(r - \frac{1}{2}\,\sigma^2) + o(\frac{1}{n})}{2\,\frac{\sigma}{\sqrt{n}} + \frac{1}{2}\,\frac{\sigma^2}{n} - \frac{1}{2}\,\frac{\sigma^2}{n} + o(\frac{1}{n})} = \frac{\frac{\sigma}{\sqrt{n}} + \frac{1}{n}(r - \frac{1}{2}\,\sigma^2) + o(\frac{1}{n})}{2\,\frac{\sigma}{\sqrt{n}} + o(\frac{1}{n})} = \frac{1}{2} + \frac{1}{2\sqrt{n}}\,\frac{r - \frac{1}{2}\,\sigma^2}{\sigma} + o(\tfrac{1}{\sqrt{n}}).$$

Equivalentemente le variabili aleatorie ξ_i sono indipendenti e di valore atteso

$$\widetilde{\mathbb{E}}^{(n)}[\xi_i] = \tilde{p}^{(n)} = \frac{1}{2} + \frac{1}{2\sqrt{n}} \, \frac{r - \frac{1}{2}\,\sigma^2}{\sigma} + o(\tfrac{1}{\sqrt{n}}).$$

Di conseguenza il prezzo di un derivato con maturità (tempo di esercizio) T e con pay-off terminale (contingent claim)

$$f(S_T^{(n)})$$

dovrà necessariamente avere come prezzo

$$C^{(n)}(f, \mathbb{P}) = \widetilde{\mathbb{E}}^{(n)}\left[\frac{f(S_T^{(n)})}{B_T^{(n)}}\right] \tag{1.7}$$

in particolare per l'opzione call

$$C^{(n)}_{call}(K, \mathbb{P}) = \widetilde{\mathbb{E}}^{(n)}\left[\frac{(S_T^{(n)} - K)^+}{B_T^{(n)}}\right] = \widetilde{\mathbb{E}}^{(n)}\left[\frac{(S_{N_n(T)} - K)^+}{\left(1 + \frac{r}{n}\right)^{N_n(T)}}\right] \tag{1.8}$$

Abbiamo quindi un'espressione del prezzo, ma il problema a questo punto diviene complesso dal punto di vista numerico, almeno per n grande: il denominatore non comporta problemi in quanto si può approssimare con e^{rT}, mentre lo stesso non si può dire del numeratore. Ci viene in aiuto il Teorema Centrale del Limite. Si noti infatti che, qualunque sia $t > 0$

$$\log S_t^{(n)} = \log S_0 + H_{N_n(t)}\, \log\left(\frac{u^{(n)}}{d^{(n)}}\right) + N_n(t)\, \log\left(d^{(n)}\right) = \log S_0 + \sum_{i=1}^{N_n(t)} \xi_i\, \log\left(\frac{u^{(n)}}{d^{(n)}}\right) + N_n(t)\, \log\left(d^{(n)}\right),$$

e che $H_{N_n(t)}$, rispetto alla misura martingala equivalente $\widetilde{\mathbb{P}}^{(n)}$, è la somma di variabili aleatorie indipendenti tutte con la stessa distribuzione, che per $t > 0$ il numero $N_n(t) = \lfloor nt \rfloor$ converge all'infinito. Grazie al Teorema Centrale del Limite si ha che quindi $H_{N_n(t)}$ ha una distribuzione ***approssimativamente*** gaussiana, di valore atteso

$$\widetilde{\mathbb{E}}^{(n)}[H_{N_n(t)}] = N_n(t)\, \tilde{p}^{(n)}$$

e varianza

$$\widetilde{Var}^{(n)}(H_{N_n(t)}) = N_n(t)\, \tilde{p}^{(n)}(1 - \tilde{p}^{(n)}).$$

Per lo stesso motivo[3], sempre rispetto alla misura martingala equivalente $\widetilde{\mathbb{P}}^{(n)}$, anche il logaritmo di $S_t^{(n)}$ è approssimata da una variabile aleatoria gaussiana di valore atteso

$$\widetilde{\mathbb{E}}^{(n)}\left[\log\left(S_t^{(n)}\right)\right] = \log S_0 + N_n(t)\, \tilde{p}^{(n)} \log\left(\frac{u^{(n)}}{d^{(n)}}\right) + N_n(t)\, \log\left(d^{(n)}\right)$$

e varianza

$$\widetilde{Var}^{(n)}\left(\log\left(S_t^{(n)}\right)\right) = N_n(t)\, \tilde{p}^{(n)}(1 - \tilde{p}^{(n)})\left(\log\left(\frac{u^{(n)}}{d^{(n)}}\right)\right)^2.$$

Ricordando che

$$\frac{u^{(n)}}{d^{(n)}} = \frac{e^{\sigma/\sqrt{n}}}{e^{-\sigma/\sqrt{n}}} = e^{2\sigma/\sqrt{n}}, \qquad d^{(n)} = e^{-\sigma/\sqrt{n}},$$

si ha

$$\log\left(\frac{u^{(n)}}{d^{(n)}}\right) = 2\,\frac{\sigma}{\sqrt{n}} \qquad \log\left(d^{(n)}\right) = -\frac{\sigma}{\sqrt{n}},$$

[3]Si ricordi che se Z ha distribuzione $N(\alpha, \beta^2)$, ovvero distribuzione gaussiana di valore atteso α e varianza β^2, allora anche $W = a + bZ$ ha distribuzione gaussiana, ma di valore atteso $\mathbb{E}[W] = \mathbb{E}[a + bZ] = a + b\alpha$, e varianza $Var(W) = Var(a + bZ) = Var(bZ) = b^2 Var(Z) = b^2\beta^2$.

si ottiene che il valore atteso del logaritmo di $S_t^{(n)}$, vale

$$\begin{aligned}
\widetilde{\mathbb{E}}^{(n)}\left[\log\left(S_t^{(n)}\right)\right] &= \log S_0 + \lfloor nt \rfloor \left(\frac{1}{2} + \frac{1}{2\sqrt{n}}\,\frac{r-\frac{1}{2}\sigma^2}{\sigma} + o(\tfrac{1}{\sqrt{n}})\right) 2\,\frac{\sigma}{\sqrt{n}} + \lfloor nt \rfloor \left(-\frac{\sigma}{\sqrt{n}}\right)\\
&= \log S_0 + \lfloor nt \rfloor \left(\frac{1}{2}\,2\,\frac{\sigma}{\sqrt{n}} + \frac{1}{2\sqrt{n}}\,\frac{r-\frac{1}{2}\sigma^2}{\sigma}\,2\,\frac{\sigma}{\sqrt{n}} + o(\tfrac{1}{\sqrt{n}})\,2\,\frac{\sigma}{\sqrt{n}} - \frac{\sigma}{\sqrt{n}}\right)\\
&= \log S_0 + \lfloor nt \rfloor \left(\frac{1}{n}\,\left(r-\tfrac{1}{2}\sigma^2\right) + o(\tfrac{1}{\sqrt{n}})\,2\,\frac{\sigma}{\sqrt{n}}\right)\\
&= \log S_0 + \frac{\lfloor nt \rfloor}{n}\,\left(r-\tfrac{1}{2}\sigma^2\right) \approx \log S_0 + t\left(r-\tfrac{1}{2}\sigma^2\right).
\end{aligned}$$

Analogamente la varianza vale

$$\begin{aligned}
\widetilde{Var}^{(n)}\left(\log\left(S_t^{(n)}\right)\right) &= N_n(t)\,\tilde{p}^{(n)}(1-\tilde{p}^{(n)})\left(\log\left(\frac{u^{(n)}}{d^{(n)}}\right)\right)^2\\
&= \lfloor nt \rfloor \left(\frac{1}{2} + \frac{1}{2\sqrt{n}}\,\frac{r-\frac{1}{2}\sigma^2}{\sigma} + o(\tfrac{1}{\sqrt{n}})\right)\left(1-\left(\frac{1}{2} + \frac{1}{2\sqrt{n}}\,\frac{r-\frac{1}{2}\sigma^2}{\sigma} + o(\tfrac{1}{\sqrt{n}})\right)\right)\left(\frac{2\sigma}{\sqrt{n}}\right)^2\\
&\approx \lfloor nt \rfloor\,\frac{1}{2}\,\frac{1}{2}\,\frac{4\sigma^2}{n} \approx t\sigma^2
\end{aligned}$$

Di conseguenza, se W_t è una variabile aleatoria gaussiana di valore atteso 0 e varianza t

$$\log(S_t^{(n)}) \overset{distr}{\to} \log S_0 + t\left(r-\tfrac{1}{2}\sigma^2\right) + \sigma\,W_t.$$

Ricordiamo che il simbolo $X_n \overset{distr}{\to}_{n\to\infty} X$ significa che per ogni funzione continua e limitata f

$$\lim_{n\to\infty} \mathbb{E}[f(X_n)] = \mathbb{E}[f(X)],$$

o in altre parole che si può approssimare $\mathbb{E}[f(X_n)]$ con $\mathbb{E}[f(X)]$.

A questo punto il prezzo di una opzione europea si può calcolare come

$$C^{(n)}(f,\mathbb{P}) = \widetilde{\mathbb{E}}^{(n)}\left[\frac{f(S_T^{(n)})}{B_T^{(n)}}\right] \approx \mathbb{E}\left[\frac{1}{e^{rT}}\,f(S_0 e^{(r-\frac{1}{2}\sigma^2)T+\sigma W_T})\right] \tag{1.9}$$

ed in particolare per l'opzione call

$$C_{call}^{(n)}(K,\mathbb{P}) = \widetilde{\mathbb{E}}^{(n)}\left[\frac{(S_T^{(n)}-K)^+}{\left(1+\frac{r}{n}\right)^{N_n(T)}}\right] \approx \mathbb{E}\left[\frac{1}{e^{rT}}\,(S_0 e^{(r-\frac{1}{2}\sigma^2)T+\sigma W_T}-K)^+\right], \tag{1.10}$$

dove l'unica variabile aleatoria è W_T, in quanto S_0 è il valore iniziale del prezzo dell'azione.

Abbiamo quindi visto come il prezzo di una opzione call europea si possa ottenere dalla formula

$$C_{call} = C(S_0,K,T,r,\sigma) = \mathbb{E}\left[e^{-rT}\,(S_0 e^{(r-\frac{1}{2}\sigma^2)T+\sigma W_T}-K)^+\right]$$

dove si è messo in evidenza la dipendenza dai parametri del modello: S_0 prezzo iniziale della azione, K prezzo di esercizio i di strike dell'opzione, T tempo di maturità o di strike dell'opzione, r tasso nominale di interesse composto in modo continuo, ed infine il parametro σ, che è detto volatilità.

È importante sottolineare che, al contrario di tutti gli altri parametri, che sono noti e direttamente osservabili, il valore della volatilità non è direttamente osservabile, ma deve essere stimato. Uno dei problemi più interessanti riguarda proprio la stima della volatilità.

1.3 Il moto Browniano

Nella derivazione precedente del prezzo abbiamo incontrato il processo del logaritmo dei prezzi, che a parte il contributo dovuto al prezzo iniziale, si esprime come

$$\sum_{i=1}^{N_n(t)} \xi_i \log\left(\frac{u^{(n)}}{d^{(n)}}\right) + N_n(t) \log\left(d^{(n)}\right) = 2\sigma \frac{1}{\sqrt{n}} \sum_{i=1}^{N_n(t)} \xi_i - N_n(t)\, \sigma \frac{1}{\sqrt{n}}$$

e che si può ulteriormente riscrivere come

$$\sigma \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{N_n(t)} (2\xi_i - 1) \right) = \sigma \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} (2\xi_i - 1) \right).$$

Si definisca

$$W_t^{(n)} := \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} \left[(2\xi_i - 1) - \widetilde{\mathbb{E}}^{(n)}[2\xi_i - 1] \right] \tag{1.11}$$

di modo che

$$\log\left(S_t^{(n)}\right) = \log\left(S_0\right) + \sigma W_t^{(n)} + \sigma \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} \widetilde{\mathbb{E}}^{(n)}[2\xi_i - 1]$$

Con calcoli analoghi a quelli della sezione precedente, si può vedere che il valore atteso

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} \widetilde{\mathbb{E}}^{(n)}[2\xi_i - 1] = \frac{1}{\sqrt{n}} \lfloor nt \rfloor \,(2\tilde{p}^{(n)} - 1) = \frac{1}{\sqrt{n}} \lfloor nt \rfloor \left(1 + \frac{1}{\sqrt{n}} \frac{r - \frac{1}{2}\sigma^2}{\sigma} + o\big(\tfrac{1}{\sqrt{n}}\big) - 1 \right) \to_{n\to\infty} \frac{r - \frac{1}{2}\sigma^2}{\sigma}\, t.$$

Ovviamente il processo $W_t^{(n)}$, vale zero all'istante iniziale, ovvero

$$W_0^{(n)} = 0,$$

ed inoltre

$$\widetilde{\mathbb{E}}^{(n)}\left[W_t^{(n)}\right] = 0.$$

Per $t > 0$, con gli stessi calcoli della sezione precedente, si può vedere che il processo $W_t^{(n)}$ converge (per n che tende all'infinito) alla legge gaussiana di valore atteso nullo e varianza che tende a t.

Ancora, se si considerano i tempi $0 \le t_1 < t_2$, allora l'incremento $W_{t_2}^{(n)} - W_{t_1}^{(n)}$ ha la stessa legge di $W_{t_2 - t_1}^{(n)}$, in quanto $W_{t_2}^{(n)} - W_{t_1}^{(n)}$ è funzione deterministica delle variabili aleatorie ξ_j, per j tale che $t_j^{(n)} = j/k \in (t_1, t_2]$, e la stessa funzione determina nello stesso modo $W_{t_2-t_1}^{(n)}$ a partire dalle variabili aleatorie ξ_i, per i tale che $t_i^{(n)} = i/n \in (0, t_2 - t_1]$, e di conseguenza la distribuzione dell'incremento $W_{t_2}^{(n)} - W_{t_1}^{(n)}$ converge ad una distribuzione gaussiana di valore atteso 0 e varianza $t_2 - t_1$.

Se invece si considerano i tempi $0 \le t_1 < t_2 < \ldots < t_{m-1} < t_m$ allora gli incrementi

$$W_{t_1}^{(n)} - W_0^{(n)}, \quad W_{t_2}^{(n)} - W_{t_1}^{(n)}, \quad \ldots \quad W_{t_m}^{(n)} - W_{t_{m-1}}^{(n)}$$

sono funzioni deterministiche delle variabili aleatorie

$$\{\xi_{j_1} : \tfrac{j_1}{n} \in (0, t_1]\}, \quad \{\xi_{j_2} : \tfrac{j_2}{n} \in (t_1, t_2]\} \quad \ldots \quad \{\xi_{j_m} : \tfrac{j_m}{n} \in (t_{m-1}, t_m]\}$$

che sono indipendenti, e di conseguenza anche gli incrementi di $W^{(n)}$ sono indipendenti. Come ulteriore conseguenza questa proprietà si mantiene al tendere di n all'infinito.

Queste osservazioni portano naturalmente alla seguente definizione del processo di Wiener standard o moto browniano.

Definizione 1.1 (moto browniano). *Si chiama moto browniano un processo* W_t *per* $t \in \mathbb{R}^+$ *un processo tale che*

1 $W_0 = 0$,

2 *se* $0 \leq t_1 < t_2 < \ldots < t_{m-1} < t_m$ *allora gli incrementi*

$$W_{t_1} - W_0, \quad W_{t_2} - W_{t_1}, \quad \ldots \quad W_{t_m} - W_{t_{m-1}}$$

sono indipendenti,

3 *se* $0 \leq t_1 < t_2$ *ed* $s > 0$ *allora gli incrementi*

$$W_{t_2} - W_{t_1} \quad e \quad W_{t_2+s} - W_{t_1+s} \sim N(0, t_2 - t_1),$$

ovvero hanno la stessa distribuzione gaussiana di valore atteso 0 *e varianza* $t_2 - t_1$*. In altre parole si dice anche che gli incrementi sono omogenei.*

Di solito oltre alle tre precedenti proprietà si aggiunge anche la proprietà che le traiettorie sono continue, ossia che per ogni ω *la funzione* $t \mapsto W_t(\omega)$ *è una funzione continua.*

Capitolo 2

Approssimazione per variabili aleatorie

2.1 Legge dei Grandi Numeri, Teorema Centrale del Limite e Approssimazione Normale

Iniziamo ricordando due definizioni di indipendenza per variabili aleatorie.

Definizione 2.1 (indipendenza di m variabili aleatorie. 1). *Siano X_1, X_2, ..., X_m m variabili aleatorie definite tutte sullo stesso spazio di probabilità $(\Omega, \mathcal{F}, P)$. Esse si dicono*[1] *completamente (o globalmente)* **indipendenti tra loro** *se comunque scelti J_1, J_2, ..., J_m, insiemi misurabili (boreliani) di $\mathbb{R}$, si ha:*

$$P(X_1 \in J_1,\, X_2 \in J_2, \ldots, X_m \in J_m) = P(X_1 \in J_1) \cdot P(X_2 \in J_2) \cdot \ldots \cdot P(X_m \in J_m). \tag{2.1}$$

Definizione 2.2 (indipendenza di m variabili aleatorie. 2). *Siano X_1, X_2, ..., X_m m variabili aleatorie definite tutte sullo stesso spazio di probabilità $(\Omega, \mathcal{F}, P)$. Esse si dicono completamente (o globalmente)* **indipendenti tra loro** *se comunque scelte m funzioni misurabili f_1, f_2, ..., f_m, con $\mathbb{E}\big[|f_i(X_i)|\big]$ finito, si ha:*

$$\mathbb{E}\big[f_1(X_1) \cdot f_2(X_2) \cdot \ldots \cdot f_m(X_m)\big] = \mathbb{E}\big[f_1(X_1)\big] \cdot \mathbb{E}\big[f_2(X_2)\big] \cdot \ldots \cdot \mathbb{E}\big[f_m(X_m)\big]. \tag{2.2}$$

Tali definizioni sono equivalenti, anche se non ne daremo la dimostrazione. Facciamo solo notare che la (2.1) è esattamente la (2.2) per $f_i(x) = \mathbf{1}_{J_i}(x)$.

Ricordiamo che una successione di variabili aleatorie $\{X_i,\ i \geq 1\}$ si dice che è una successione di variabili aleatorie indipendenti a due a due, se comunque presi i e j, le due variabili aleatorie X_i ed X_j sono indipendenti.

Una successione di variabili aleatorie $\{X_i,\ i \geq 1\}$ si dice che è una successione di variabili aleatorie completamente (o globalmente) indipendenti, se per ogni $m \geq 2$, e comunque presi $i_1, i_2, \ldots, i_m$, le variabili aleatorie $X_{i_1}, X_{i_2}, \ldots, X_{i_m}$ sono completamente (o globalmente) indipendenti.

Proposizione **(Legge Debole dei Grandi Numeri)** Sia $\{X_i,\ i \geq 1\}$ una successione di v.a. indipendenti a due a due ed identicamente distribuite[2], per le quali esistano finiti valore atteso e varianza. Posto

$$\mathbb{E}(X_i) = \mu, \qquad Var(X_i) = \sigma^2,$$

$$S_n = \sum_{i=1}^{n} X_i \qquad Y_n = \frac{S_n}{n} = \frac{1}{n} \sum_{i=1}^{n} X_i,$$

si ha, qualunque sia $\varepsilon > 0$

$$\lim_{n\to\infty} P\left(\left|\frac{S_n}{n} - \mu\right| > \varepsilon\right) = \lim_{n\to\infty} P\left(|Y_n - \mu| > \varepsilon\right) = 0$$

[1]A volte il termine completamente può essere trascurato, e si può parlare semplicemente di variabili aleatorie indipendenti tra loro.

[2]Poiché le variabili aleatorie X_n hanno tutte la stessa distribuzione, si ha che se esistono finiti valore atteso e varianza di X_1, allora esistono finiti valore atteso e varianza di X_i e coincidono con quelli di X_1.

Dimostrazione. Basta osservare che

$$E(Y_n) = E\left(\tfrac{1}{n}\,(X_1 + \cdots + X_n)\right) = \tfrac{1}{n}\,(E(X_1) + \cdots + E(X_n)) = \tfrac{1}{n}\,n\mu = \mu$$

e

$$Var(Y_n) = Var\left(\tfrac{1}{n}\,(X_1 + \cdots + X_n)\right) = \tfrac{1}{n^2} Var\,(X_1 + \cdots + X_n)$$

essendo le X_i indipendenti a due a due e quindi non correlate, la varianza della somma è la somma delle varianze

$$Var\,(X_1 + \cdots + X_n) = Var(X_1) + \cdots + Var(X_n)$$

da cui

$$Var(Y_n) = \tfrac{1}{n^2}\,(Var(X_1) + \cdots + Var(X_n)) = \tfrac{1}{n^2}\,n\sigma^2 = \tfrac{1}{n}\,\sigma^2.$$

Quindi, dalla disuguaglianza di Chebyshev, si ha

$$0 \le P\,(|Y_n - \mu| > \varepsilon) \le \frac{1}{n}\,\frac{\sigma^2}{\varepsilon^2},$$

e basta mandare n all'infinito ed usare il Teorema del confronto per le successioni numeriche:

$$0 \le \lim_{n\to\infty} P\,(|Y_n - \mu| > \varepsilon) \le \lim_{n\to\infty} \frac{1}{n}\,\frac{\sigma^2}{\varepsilon^2} = 0.$$

Osservazione Dalle varie definizioni di indipendenza, appare immediato che se $\{X_i,\ i \ge 1\}$ è una successione di variabili aleatorie completamente indipendenti, allora sono anche indipendenti a due a due, e allora la Legge Debole dei Grandi Numeri continua a valere. Sotto questa ulteriore ipotesi vale anche il così detto Teorema centrale del limite.

***Proposizione* (Teorema Centrale del Limite)** Sia $\{X_i,\ i \ge 1\}$ una successione di v.a. indipendenti ed identicamente distribuite, per le quali esistano finiti valore atteso e varianza. Posto $\mathbb{E}(X_i) = \mu$ e $Var(X_i) = \sigma^2$, si assuma che $\sigma^2 > 0$. Allora indicando con S_n^* variabile aleatoria standardizzata di S_n, si ha

$$S_n^* = \frac{S_n - \mathbb{E}(S_n)}{\sqrt{Var(S_n)}} = \frac{S_n - n\mu}{\sqrt{n\sigma^2}}, \tag{2.3}$$

e, indicando con $F_{S_n^*}(x)$ la funzione di distribuzione di S_n^*, si ha

$$\lim_{n\to\infty} F_{S_n^*}(x) = \lim_{n\to\infty} P\,(S_n^* \le x) = \Phi(x), \tag{2.4}$$

dove Φ è la funzione di distribuzione di una variabile aleatoria Gaussiana standard: in altre parole

$$\lim_{n\to\infty} P\left(\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \le x\right) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\,\pi}}\, e^{-\frac{y^2}{2}}\, dy. \tag{2.5}$$

Inoltre il limite è uniforme per $x \in \mathbb{R}$, ovvero

$$\lim_{n\to\infty} \sup_{x\in\mathbb{R}} \left| P\left(\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \le x\right) - \int_{-\infty}^{x} \frac{1}{\sqrt{2\,\pi}}\, e^{-\frac{y^2}{2}}\, dy \right| = 0. \tag{2.6}$$

Non diamo la dimostrazione di questo risultato, ma notiamo solo che la (2.3) si dimostra tenendo conto che $\mathbb{E}(S_n) = \sum_{i=1}^n \mathbb{E}(X_i) = n\mu$ e che per la completa indipendenza dalle variabili aleatorie X_i, si ha[3]

$$Var(S_n) = Var(\sum_{i=1}^n X_i) = \sum_{i=1}^n Var(X_i) = n\sigma^2.$$

[3]Come già osservato nell'***Osservazione 1***

La precedente relazione sarebbe valida anche nel caso in cui le variabili aleatorie fossero solo indipendenti a due a due(o addirittura solo non correlate), ma sottolineiamo il fatto che, mentre la Legge Debole dei Grandi Numeri, vale sotto l'ipotesi di indipendenza a due a due, e non è necessario supporre $\sigma^2 > 0$, invece ***per il Teorema Centrale del Limite, serve la condizione di completa indipendenza*** e ovviamente ***è necessario supporre*** $\sigma^2 > 0$, altrimenti non si potrebbe nemmeno formulare la tesi.

Va inoltre ricordato che la convergenza delle funzioni di distribuzione è equivalente alla proprietà, che per ogni funzione f continua e limitata,

$$\lim_{n\to\infty} E[f(S*_n)] = E[f(Y)] = \int_{-\infty}^{+\infty} f(y) \tfrac{1}{\sqrt{2\,\pi}}\, dy,$$

dove Y è una variabile aleatoria gaussiana standard.

Anche di questa proprietà non diamo la dimostrazione, ma osserviamo solo che la convergenza delle funzioni di distribuzione corrisponde alla convergenza di

$$\lim_{n\to\infty} E[h(S*_n)] = E[h(Y)],$$

per ogni funzione $h(y) = \mathbf{1}_{(-\infty,x]}(y)$, cioè della funzione $h(y)$ che vale 1 per $y \leq x$ e vale 0 altrimenti.

2.2 Approssimazione normale

Come abbiamo visto nella dimostrazione della Legge dei Grandi Numeri, la disuguaglianza di Chebyshev permette di trovare delle limitazioni inferiori alle probabilità del tipo

$$P\left(\left|\frac{S_n}{n} - \mu\right| \leq \varepsilon\right)$$

che appunto permettono di dedurre la legge dei grandi numeri. Tuttavia se si conoscesse la funzione di distribuzione $F_{S_n}(x)$ della variabile aleatoria S_n, tale probabilità si potrebbe calcolare esattamente come

$$\begin{aligned} P\left(\left|\frac{S_n}{n} - \mu\right| \leq \varepsilon\right) &= P\left(n(\mu-\varepsilon) \leq S_n \leq n(\mu+\varepsilon)\right) = F_{S_n}\left(n(\mu-\varepsilon)\right) - F_{S_n}\Big(-\left(n(\mu-\varepsilon)\right)^-\Big) \\ &= F_{S_n}\left(n(\mu-\varepsilon)\right) - F_{S_n}\left(-n(\mu-\varepsilon)\right) + P\big(\{S_n = -n(\mu-\varepsilon)\}\big) \end{aligned}$$

Appare quindi chiaro che calcolare la distribuzione della somma di variabili aleatorie $S_n = X_1+X_2+\ldots+X_n$ sia un problema interessante è, oltre che di per sé, anche per le connessioni con la legge dei grandi numeri e delle relazioni tra media aritmetica e valore atteso.

Alla luce del Teorema Centrale del Limite (o anche Teorema del Limite Centrale), si può dimostrare il seguente risultato.

Proposizione **(approssimazione normale)** Se le variabili aleatorie X_i, per $i\ =, 1, 2\ldots, n$ sono (globalmente o completamente) indipendenti, hanno la stessa distribuzione, ammettono valore atteso finito μ, varianza finita σ^2 e non nulla, allora $\mathbb{E}(S_n) = n\mu$, $Var(S_n) = n\sigma^2 > 0$, e

$$F_{S_n}(x) = P(S_n \leq x) = P\left(\frac{S_n - n\mu}{\sqrt{n\,\sigma^2}} \leq \frac{x-n\mu}{\sqrt{n\,\sigma^2}}\right) \simeq \Phi\left(\frac{x-n\mu}{\sqrt{n\,\sigma^2}}\right), \tag{2.7}$$

dove $\Phi(x)$ è la funzione di distribuzione di una variabile aleatoria gaussiana standard $N(0,1)$.

Dimostrazione Come anticipato la dimostrazione della precedente affermazione si basa sul risultato basilare che svolge un ruolo "centrale" nel Calcolo delle Probabilità il Teorema Centrale del Limite.

Fondamentale per dimostrare l'approssimazione (2.7) della funzione di distribuzione della somma S_n è il fatto che la convergenza sia uniforme[4]: infatti, posto

$$E_n(x) = F_{S_n^*}(x) - \Phi(x), \text{ e } x_n = \tfrac{x-n\mu}{\sqrt{n\,\sigma^2}},$$

si ha

$$F_{S_n}(x) = P(S_n \leq x) = P\left(\tfrac{S_n-n\mu}{\sqrt{n\,\sigma^2}} \leq \tfrac{x-n\mu}{\sqrt{n\,\sigma^2}}\right) = F_{S_n^*}(x_n) = \Phi(x_n) + E_n(x_n),$$

per cui

$$|F_{S_n}(x) - \Phi(x_n)| = |E_n(x_n)| \leq \sup_{x\in\mathbb{R}} |E_n(x)|.$$

Basta solo osservare che (2.5) garantisce che $\sup_{x\in\mathbb{R}} |E_n(x)|$ converge a zero[5] per n che tende all'infinito.

[4]Si osservi che in generale le condizioni che

$$\lim_{n\to\infty} f_n(x) = f(x) \qquad \lim_{n\to\infty} x_n = x$$

non implicano che

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

Basta pensare al seguente **controesempio**:

$$\begin{cases} f_n(x) = 0 & x < \frac{1}{n}, \\ f_n(x) = 1 & x \geq \frac{1}{n} \end{cases} \qquad \begin{cases} f(x) = 0 & x \leq 0, \\ f(x) = 1 & x > 0 \end{cases}$$

Chiaramente se $x \leq 0$ allora $f_n(x) = 0$ e quindi $\lim_{n\to\infty} f_n(x) = f(x) = 0$, analogamente, se $x > 0$, allora per $n > \frac{1}{x}$ si ha $f_n(x) = 1$, e quindi $\lim_{n\to\infty} f_n(x) = f(x) = 1$. Inoltre, posto $x_n = \frac{1}{n}$, si ha $\lim_{n\to\infty} x_n = 0$, tuttavia ovviamente $f_n(x_n) = f_n(\frac{1}{n}) = 1$ che non converge ad $f(x) = f(0) = 0$.

[5]Pur essendo assolutamente al di fuori dell'ambito di un corso elementare di probabilità, vale la pena di ricordare che esistono delle maggiorazioni per $\sup_{x\in\mathbb{R}} |E_n(x)|$, nel caso in cui si supponga che il valore atteso $\mathbb{E}(|X|^3)$ esista e sia finito. In particolare è stato dimostrato che

$$\sup_{x\in\mathbb{R}} |E_n(x)| \leq \frac{C}{\sqrt{n}} \frac{\mathbb{E}(|X|^3)}{\sigma^3},$$

con C costante. Il valore di C non è noto esattamente ma è noto che $0.4097 \leq C \leq 0.7975$, in particolare quindi vale

$$\sup_{x\in\mathbb{R}} |E_n(x)| \leq \frac{1}{\sqrt{n}} \frac{\mathbb{E}(|X|^3)}{\sigma^3},$$

I primi a fornire maggiorazioni in questa direzione sono stati Berry ed Eessen all'inizio degli anni 40 dello scorso XX secolo.