2.3 La legge debole dei grandi numeri

Questo argomento comprende una classe di risultati relativi al comportamento asintotico della media aritmetica delle variabili aleatorie di una successione; risultati che forniscono un'efficace giustificazione a un fenomeno che si presenta in molte situazioni reali chiamato regolarità statistica: la media aritmetica degli esiti numerici di prove ripetute di un esperimento manifesta una spiccata tendenza a stabilizzarsi all'aumentare del numero di prove.
Nell'ambito della legge dei grandi numeri si usa fare una distinzione tra legge *forte*, che implica una forma di convergenza quasi certa, e legge *debole*, che fa invece riferimento a una forma di convergenza in probabilità ed è il caso che ci interessa e che andremo ad osservare.
La formulazione più semplice della legge debole dei grandi numeri è la seguente:

Teorema 2.3.1. ***Teorema di Chebyshev*** *Siano $X_1, X_2, \dots$ variabili casuali mutuamente indipendenti ciascuna con media μ e varianza σ^2 finite. Allora se $S_n = X_1 + X_2 + \dots + X_n$, $n = \{1, 2, \dots\}$, vale*

$$lim_{n\to+\infty} P\left(|\frac{S_n}{n} - \mu| \geq \lambda\right) = 0 \quad \forall\lambda > 0$$

Questo teorema afferma che la probabilità che la media campionaria $\frac{S_n}{n}$ differisca dal suo valore probabile μ per più di ϵ tende a zero quando n tende all'infinito.
Si osserva inoltre che il teorema non richiede che la successione numerica presenti alcun carattere di convergenza, sicchè la legge rimane valida anche se tale successione diverge.

Dimostrazione. Introduciamo la seguente proposizione:

Proposizione 2.3.1. ***Disuguaglianza di Chebycev***
Valgono le seguenti disuguaglianze:

- *sia X una variabile aleatoria con previsione quadratica $P_Q(X) > 0$. Per ogni $t > 0$ vale*

$$P(|X| \geq tP_Q(X)) \leq \frac{1}{t^2};$$

- *sia X una variabile aleatoria con varianza $\sigma^2(X) > 0$, posto $\mu = P(X)$, per ogni $t > 0$ vale*

$$P(|X - \mu| \geq \sigma(X)t) \leq \frac{1}{t^2};$$

Usiamo la seconda disuguaglianza di Chebycev per dimostrare il teorema. Calcoliamo

$$P\left(\frac{S_n}{n}\right) = \frac{1}{n}(P(X_1) + ...P(X_n)) = \mu$$

$$\sigma^2\left(\frac{S_n}{n}\right) = \frac{1}{n^2}\sigma^2(S_n) = \frac{1}{n^2}\left(\sum_{i=1}^{n}\sigma^2(X_i) + \sum_{i,j=1;i\neq j}^{n} cov(X_i, Y_j)\right) = \frac{\sigma^2}{n}.$$

Per la seconda disuguaglianza di Chebycev vale

$$P\left(|\frac{S_n}{n} - \mu| \geq \frac{\sigma}{\sqrt{n}}t\right) \leq \frac{1}{t^2}.$$

Se poniamo $\lambda = \frac{\sigma}{\sqrt{n}}t$ e quindi $\frac{1}{t^2} = \frac{\sigma^2}{n\lambda^2}$, risulta che

$$P\left(|\frac{S_n}{n} - \mu| \geq \lambda\right) \leq \frac{\sigma^2}{n\lambda^2}$$

tende a 0 per $n \to +\infty$. □

E' fondamentale però notare che, in teoria, solo per un numero infinto di uscite ci sarà conincidenza fra probabilità teoriche e frequenze sperimentali; inoltre è possibile trovare sorprese controintuitive da una rigorosa applicazione della legge dei grandi numeri: lo scarto proporzionale diminuisce, ma può aumentare in valore assoluto.
Credere in una natura uniforme che tende a fornire eventi casuali con una frequenza pari a quanto previsto da un calcolo a priori delle probabilità è un errore comune e molto pericoloso per il giocatore. Troppo spesso l'esperienza temporale del giocatore non è tale da poter vedere compensati gli sbilanciamenti di una sera.