PAC - 27 MARZO 2008

E 4. Distribuzione binomiale

Siano $N \in \mathbb{N}$ e $p \in [0,1]$ due parametri e Z la v.a. con distribuzione binomiale definita dalle probabilità:

$$\mathbb{P}(Z = k) = \binom{N}{k} p^k (1-p)^{N-k}\,, \qquad k = 0, 1, 2, \ldots, N \tag{0.1}$$

Si effettui l'esperimento seguente per diversi valori di N e p. Per esempio si considerino i casi $(N, p) = (10, 0.2)\,;\ (50, 0.4)\,;\ (500, 0.1)$.

- Simulare Z utilizzando N prove di Bernoulli di parametro p (vedi esercitazione E2).
- Simulare n copie indipendenti $Z_1, \ldots, Z_n$ della variabile binomiale Z definita dalla (0.1) e rappresentare i dati tramite un istogramma di frequenze. Si prenda per esempio $n = 10^5$ o 10^6.
- Graficare media e varianza empiriche m_n e v_n, al variare di n. Per esempio, se $n = 10^5$, si grafichi la tabella di valori (k, m_{k*m}, v_{k*m}), dove $m = 10^3$ e $k = 1, 2, \ldots, 100$. Si osservi che

$$m_n \sim Np\,, \qquad v_n \sim Np(1-p)\,. \tag{0.2}$$

Suggerimenti

Siano N e p fissati. Siano $X_1, \ldots, X_N$ i risultati della simulazione di N bernoulliane indipendenti di parametro p. Ponendo $Z = X_1 + \cdots + X_N$ otteniamo una simulazione della variabile binomiale di parametri N e p.

Ripetendo l'operazione descritta sopra per un numero n di volte, otteniamo una simulazione di n copie indipendenti della variabile Z. Chiamiamo $Z_1, \ldots, Z_n$ i valori ottenuti. Un istogramma di frequenze puo' essere ottenuto come segue. Consideriamo gli $N+1$ valori che una singola Z puo' raggiungere, cioè $\{0, 1, \ldots, N\}$. A ognuno di essi associamo una colonna la cui altezza è data dal numero complessivo delle variabili che assumono quel valore. Precisamente: l'istogramma è costituito dalle altezze $h_0, h_1, \ldots, h_N$, dove

$$h_k = \#\{i = 1, \ldots, n\,:\, Z_i = k\}\,, \quad k = 0, 1, \ldots, N\,.$$

Per ottenere un grafico normalizzato, ossia un'approssimazione della distribuzione binomiale (0.1), si puo' graficare la tabella $(k, h_k/n)$, $k = 0, 1, \ldots, N$.

`www.mat.uniroma3.it/users/caputo`.