PAC - 22 MAGGIO 2003

E 11. Algoritmo Box–Muller per v.a. Gaussiane

1. Simulare una v.a. Gaussiana X di media nulla e varianza $\sigma^2 = 0.5$ usando l'algoritmo Box–Muller. Graficare media e varianza empirica su un numero di prove M e stabilire un numero M_0 a partire dal quale la misura si stabilizza.
2. Dare una stima empirica di $F(\xi) = \mathbb{P}[X \leqslant \xi]$ e graficare per tutti i multipli di 0.2 nell'intervallo $[0, 20]$. Confrontare il risultato con il metodo Monte Carlo per il calcolo di integrali Gaussiani discusso in **E 7**.

E 12. Teorema del limite centrale

Si calcoli l'integrale Gaussiano

$$I(\xi) = \int_0^\xi \frac{\mathrm{e}^{-z^2/2}}{\sqrt{2\pi}}\, dz\,, \quad \xi \geqslant 0$$

utilizzando la formula di De Moivre–Laplace (teorema del limite centrale)

$$\mathbb{P}\left[\frac{1}{\sqrt{p(1-p)N}} \sum_{i=1}^{N} (X_i - p) \in [0, \xi]\right] \to I(\xi)\,, \quad N \to \infty\,, \tag{0.1}$$

dove $X_1, \ldots, X_N$ sono copie indipendenti di una variabile di Bernoulli di parametro p. Considerando i casi $p = 0.2$ e $p = 0.5$, e riportando i valori ottenuti su un grafico (per ξ multipli di 0.1 nell'intervallo $[0, 10]$) si confronti il risultato con le stime ottenute in **E 7** e **E 11**.

caputo@mat.uniroma3.it.

Suggerimenti

E 11. Poiché l'integrale Gaussiano

$$\mathbb{P}[X \leqslant \xi] = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\xi} e^{-\frac{t^2}{\sigma^2}}\, dt$$

non è noto analiticamente, il metodo della trasformazione non si può applicare direttamente in questo caso. L'algoritmo di Box–Muller si basa sull'osservazione che se R è v.a. continua in $[0, \infty)$ con densità

$$f_R(r) = \frac{r}{\sigma^2} e^{-\frac{r^2}{\sigma^2}}$$

e Θ è v.a. uniforme in $[0, 2\pi)$, allora $X = R\cos\Theta$ è v.a. $N(0, \sigma^2)$, cioè Gaussiana di media nulla e varianza σ^2. In effetti, tramite un passaggio a coordinate polari nel piano cartesiano si dimostra facilmente che se $Y = R\sin\Theta$, allora X e Y sono due variabili $N(0, \sigma^2)$ *indipendenti.* Ora la variabile Θ si può simulare banalmente come $\Theta = 2\pi Z$ con Z uniforme in $[0, 1)$ mentre per la variabile R si può applicare il metodo della trasformazione: Infatti

$$F_R(x) = \frac{1}{\sigma^2} \int_0^x r\, e^{-\frac{r^2}{\sigma^2}}\, dr = 1 - e^{-\frac{x^2}{\sigma^2}}\,, \quad x \geqslant 0\,.$$

e invertendo si ha $F_R^{-1}(y) = \sqrt{-2\sigma^2 \log(1-y)}$, $y \in (0, 1)$. Dunque, per simulare una Gaussiana $X = N(0, \sigma^2)$ si può utilizzare il seguente algoritmo: si generano 2 v.a. indipendenti Z_1, Z_2 uniformi in $[0, 1]$; si pone

$$\Theta = 2\pi Z_1\,, \qquad R = \sqrt{-2\sigma^2 \log Z_2}\,,$$

e si calcola infine $X = R\cos\Theta$.

Ripetendo un certo numero M di volte la simulazione di X si può stimare empiricamente la probabilità $\mathbb{P}[X \leqslant \xi]$ (come frazione dei sorteggi in cui tale evento si è verificato). Nell'esercitazione **E 7** si propone di calcolare l'integrale Gaussiano $I(\xi) = \int_0^\xi \frac{e^{-t^2/2}}{\sqrt{2\pi}}\, dt$, $\xi \geqslant 0$. Un confronto può essere fatto sulla base dell'identità $\mathbb{P}[X \leqslant \xi] = \frac{1}{2} + I(\xi/\sigma)$.

E 12. Una volta generate le $\{X_i\}_{i=1}^N$ con un N fissato (e.g. $N = 10^3$) si calcola la variabile

$$Y^{(N)} = \frac{1}{\sqrt{p(1-p)N}} \sum_{i=1}^{N} (X_i - p)$$

Per evere una stima delle probabilità in (0.1) si realizzano M (e.g. $M = 10^3$) copie indipendenti $Y_1^{(N)}, \ldots, Y_M^{(N)}$ della variabile $Y^{(N)}$ e si calcola la frequenza empirica dell'evento $\{Y^{(N)} \in [0, \xi]\}$ per i diversi valori di ξ. Combinando legge dei grandi numeri e teorema del limite centrale si ottiene dunque una stima di $I(\xi)$.