

## PAC - MAGGIO 2006

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

1. Simulare una v.a. Gaussiana  $X$  di media nulla e varianza  $\sigma^2$  usando l'algoritmo Box–Muller (per esempio  $\sigma^2 = 1$  e  $\sigma^2 = 0.5$ ). Graficare media e varianza empirica su prove ripetute.
2. Graficare il corrispondente istogramma delle frequenze nell'intervallo  $[-4\sigma, 4\sigma]$  (che dovrebbe contenere più del 99,99% dei dati).

**E 12.** Approssimazione normale (Teorema centrale del limite).

Siano  $Y_1, \dots, Y_N$  copie indipendenti di una variabile di Bernoulli di parametro  $p$ . Sia  $X_N$  la variabile

$$X_N = \frac{1}{\sqrt{p(1-p)N}} \sum_{i=1}^N (Y_i - p).$$

Si ricordi il limite di De Moivre – Laplace

$$\mathbb{P}[X_N \in (a, b)] \rightarrow \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{z^2}{2}} dz, \quad N \rightarrow \infty \quad (0.1)$$

La (0.1) stabilisce che  $X_N$  è approssimata da una normale  $N(0, 1)$  per  $N$  grande. Considerando differenti valori di  $p$ , si effettui una verifica del risultato (0.1), graficando un istogramma delle frequenze per la variabile  $X_N$  (per  $N$  fissato) e confrontandolo con l'istogramma ottenuto nell'esercitazione precedente.

### Suggerimenti

**E 10.** Poiché l'integrale Gaussiano

$$F_X(x) = \mathbb{P}[X \leq x] = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^x e^{-\frac{z^2}{2\sigma^2}} dz$$

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

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

e  $\Theta$  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  $\sigma^2$ . 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  $\Theta$  si può simulare banalmente come  $\Theta = 2\pi U$  con  $U$  uniforme in  $[0, 1)$  mentre per la variabile  $R$  si può applicare il metodo della trasformazione: Infatti

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

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  $U_1, U_2$  uniformi in  $[0, 1]$ ; si pone

$$\Theta = 2\pi U_1, \quad R = \sqrt{-2\sigma^2 \log U_2};$$

si calcola infine  $X = R \cos \Theta$ . Ripetendo  $M$  volte la simulazione di  $X$  si ottiene nel modo usuale un istogramma delle frequenze.

**E 11.** Una volta generate le bernoulliane  $\{Y_i\}_{i=1}^N$  con un  $N$  fissato (e.g.  $N = 10^3$ ) si calcola la variabile  $X_N$  come nel testo. Generiamo poi  $M$  copie indipendenti della  $X_N$ , che chiameremo  $X_N^1, \dots, X_N^M$ . Questi valori possono essere utilizzati nel modo usuale per ottenere un istogramma delle frequenze. Per  $M$  e  $N$  grandi l'istogramma approssima quello ottenuto nell'esperienza precedente.