Appendice al Cap. 11

La trasformazione di Box-Muller

Dato un vettore $\mathbf{x}$ di variabili casuali $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, le cui componenti abbiano una funzione densità di probabilità $P(x)$, la corrispondente densità di probabilità $D(y)$ delle componenti del vettore[1] $\mathbf{y} = \mathbf{y}(\mathbf{x})$ è data da

$$D(\mathbf{y}) = P(\mathbf{x}) \left|\det J\right| , \tag{11.1}$$

dove J è una matrice $n \times n$ i cui elementi sono $J_{ij} = \partial x_i / \partial y_i$.

Con questa premessa cercheremo ora di determinare una trasformazione $\mathbf{y} = \mathbf{y}(\mathbf{x})$ che trasformi un vettore di variabili casuali x_i distribuite uniformemente, in un vettore di altrettante variabili la cui distribuzione sia quella di Gauss, con varianza unitaria e media nulla. Con un computer, infatti, possiamo generare facilmente le x_i, distribuite uniformemente tra due estremi a e b. La funzione densità di probabilità di ciascuna x_i è

$$P(x) = \text{cost} = \frac{1}{b-a} , \tag{11.2}$$

e possiamo sempre scegliere $a = 0$ e $b = 1$, in modo tale che cost $= 1$.

Volendo estrarre un solo numero y potremmo pensare di limitarci al caso $n = 1$. Dall'equazione (11.1), per la scelta fatta di a e b,

$$D(y) \equiv \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} = P(x) \left| \frac{\partial x}{\partial y} \right| = \left| \frac{\partial x}{\partial y} \right| .$$

Sfortunatamente, non esistendo un'espressione esplicita per la primitiva della funzione di Gauss, in questo modo è impossibile scrivere la funzione $x = x(y)$, e di conseguenza è impossibile ottenere la sua inversa $y = y(x)$.

Possiamo quindi provare con $n = 2$. In questo caso si ha che

$$D(y_1, y_2) \equiv \frac{1}{2\pi} e^{-\frac{y_1^2}{2}} e^{-\frac{y_2^2}{2}} = P(x_1, x_2) \left|\det J\right| = \left| \frac{\partial x_1}{\partial y_1} \frac{\partial x_2}{\partial y_2} - \frac{\partial x_2}{\partial y_1} \frac{\partial x_1}{\partial y_2} \right| . \tag{11.3}$$

[1] Con leggero abuso di notazione indichiamo con y sia la variabile casuale che la sua dipendenza funzionale da x.

Dobbiamo dunque determinare le funzioni $x_1 = x_1(y_1, y_2)$ e $x_2 = x_2(y_1, y_2)$. Possiamo arbitrariamente scegliere una delle due e determinare l'altra. Una scelta semplice può essere

$$x_2 = e^{-\frac{y_1^2+y_2^2}{2}} . \tag{11.4}$$

Sostituendo le derivate di x_2 rispetto a y_1 e y_2 nell'equazione (11.3) si ottiene l'equazione

$$\left| y_1 \frac{\partial x_1}{\partial y_2} - y_2 \frac{\partial x_1}{\partial y_1} \right| = \frac{1}{2\pi} . \tag{11.5}$$

Il fatto che la differenza dei termini a primo membro dell'equazione appena trovata è una costante, induce a pensare che x_1 sia una funzione del rapporto y_1/y_2, $x_1 = F(y_1/y_2)$, da cui si ottiene che

$$\frac{\partial x_1}{\partial y_1} = F'\left(\frac{y_1}{y_2}\right)\frac{1}{y_2} \quad \text{e} \quad \frac{\partial x_1}{\partial y_2} = -F'\left(\frac{y_1}{y_2}\right)\frac{y_1}{y_2^2} .$$

L'equazione (11.5) si riscrive dunque

$$\left| F'\left(\frac{y_1}{y_2}\right)\left(\left(\frac{y_1}{y_2}\right)^2 + 1\right)\right| = \frac{1}{2\pi} ,$$

da cui si ottiene

$$\left| F'\left(\frac{y_1}{y_2}\right)\right| = \frac{1}{2\pi}\frac{1}{1+\left(\frac{y_1}{y_2}\right)^2} ,$$

la cui soluzione è

$$F\left(\frac{y_1}{y_2}\right) = A + \frac{1}{2\pi}\arctan\frac{y_1}{y_2} , \tag{11.6}$$

perciò dalle equazioni (11.6) e (11.4), in definitiva, scegliendo $A = 0$, avremo che

$$\frac{y_1}{y_2} = \tan 2\pi x_1 \quad \text{e} \quad y_1^2 + y_2^2 = -2\ln x_2 . \tag{11.7}$$

Dalle equazioni (11.7) è facile ottenere la trasformazione desiderata (osservando che il segno -1 residuo è un fattore di fase irrilevante) :

$$\begin{aligned} y_1 &= \sqrt{-2\ln x_2}\,\sin 2\pi x_1 \\ y_2 &= \sqrt{-2\ln x_2}\,\cos 2\pi x_1 , \end{aligned}$$

che, partendo da due variabili x_1, x_2 uniformemente distribuite tra 0 e 1, genera due variabili y_1 e y_2 distribuite come gaussiane di media nulla e varianza unitaria.