

Numeri pseudocasuali

“Numeri casuali non devono essere generati con un metodo scelto a caso” D.Knuth

- Il periodo deve essere il più lungo possibile;
- la distribuzione deve essere uniforme in $[0, 1]$
 $p(x) = \text{costante}$ in $[0, 1]$;
- le correlazioni devono essere trascurabili
 $\langle x_{n+1} \cdot x_n \rangle - \langle x_{n+1} \rangle \langle x_n \rangle = 0$;
- distribuzioni uniformi:
 - metodi lineari congruenti;
 - metodi Fibonacci lagged;
 - metodi non lineari, etc
- distribuzioni non uniformi:
 - distribuzione gaussiana e metodo di Box-Müller
 - distribuzione qualsiasi;

Metodi lineari congruenti (LCM)

- $x_{n+1} = (a \cdot x_n + b) \bmod m$;
- di solito $b = 0$;
- a è primo;
- m è primo, oppure $m = 2^n$

Problema

overflow per moltiplicazione per a .

Soluzione

$m = a \cdot q + r$ con $r < q$ scrivo $x_n = z \cdot q + w$
allora

$$a \cdot x_n = a \cdot (z \cdot q + w) = a \cdot q \cdot z + a \cdot w = (a \cdot q + r) \cdot z - r \cdot z + a \cdot w$$

$$a \cdot x_n = m \cdot z - r \cdot z + a \cdot w$$

con $r \cdot z < q \cdot z < x_n < m$ e $a \cdot w < a \cdot q < m$

Quindi $a \cdot w - r \cdot z$ è minore di m ed il risultato è

$$a \cdot x_n \bmod m = a \cdot w - r \cdot z (+m) =$$

$$a \cdot (x_n \bmod m) - r \cdot (x_n/q) (+m)$$

Esempio:

$$a = 16807 \quad m = 2^{31} = 2147483648$$

$$q = 127773 \quad r = 2836$$

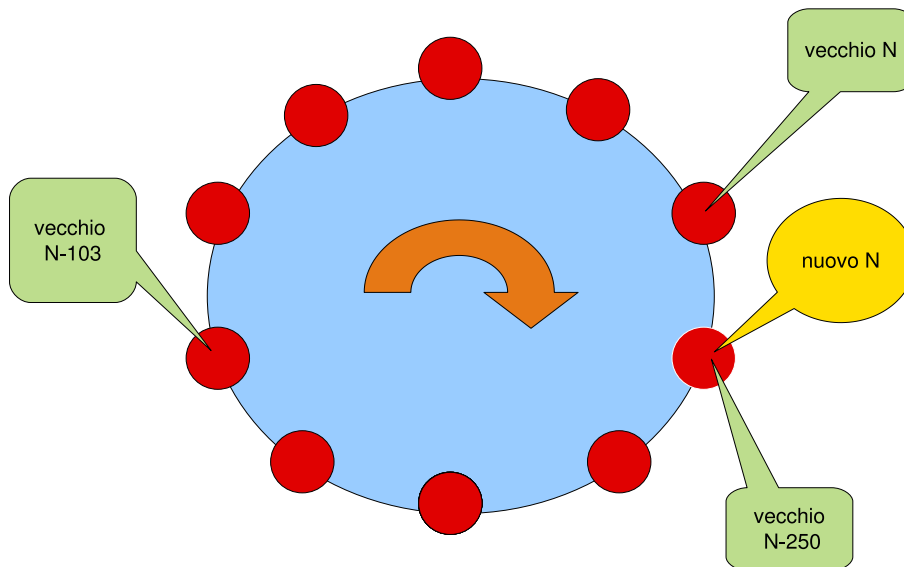
Metodi Fibonacci lagged

$$x_{n+1} = \sum_{q=0}^N a_q x_{n-q}$$

- Simili alla successione di Fibonacci, ma con più termini; di solito solo un paio di c_k sono diversi da zero;
- occorre una certa attenzione alle condizioni iniziali; si può provare a usare LCM come "starter" per i primi termini;
- hanno periodo molto lungo e scarse correlazioni;
- più successioni indipendenti con la stessa regola;

Esempio: r250

- $x_n = \sum_{j=1}^{250} a_j x_{n-j}$
- periodo $N \approx 2^{250}$
- $a_q = 0$ escluso che per $q = 103, 250$ per cui $a_q = 1$
- quindi $x_n = x_{n-103} + x_{n-250}$
- in realtà si usa *xor*: $x_n = x_{n-103} \text{ XOR } x_{n-250}$
- memorizzo gli ultimi 250 termini in uno stack circolare per non dover spostare ad ogni passo un intero vettore



Metodo di Box-Müller

Voglio una distribuzione gaussiana: se x_1 e x_2 sono distribuite uniformemente in $(0, 1)$ definisco

$$y_1 = \sqrt{-2 \ln x_1} \sin(2\pi x_2)$$

$$y_2 = \sqrt{-2 \ln x_1} \cos(2\pi x_2)$$

Quindi $x_1 = e^{-(y_1^2 + y_2^2)/2}$

e $p(y_1) p(y_2) dy_1 dy_2 = p(x_1) p(x_2) dx_1 dx_2 = dx_1 dx_2$

$$p(y_1) p(y_2) \frac{\partial(y_1, y_2)}{\partial(x_1, x_2)} = 1$$

$$\frac{\partial(y_1, y_2)}{\partial(x_1, x_2)} = \begin{vmatrix} \frac{-1}{x_1 \sqrt{-2 \ln x_1}} \sin(2\pi x_2) & 2\pi \sqrt{-2 \ln x_1} \cos(2\pi x_2) \\ \frac{-1}{x_1 \sqrt{-2 \ln x_1}} \cos(2\pi x_2) & -2\pi \sqrt{-2 \ln x_1} \sin(2\pi x_2) \end{vmatrix}$$

$$\frac{\partial(y_1, y_2)}{\partial(x_1, x_2)} = \frac{2\pi}{x_1}$$

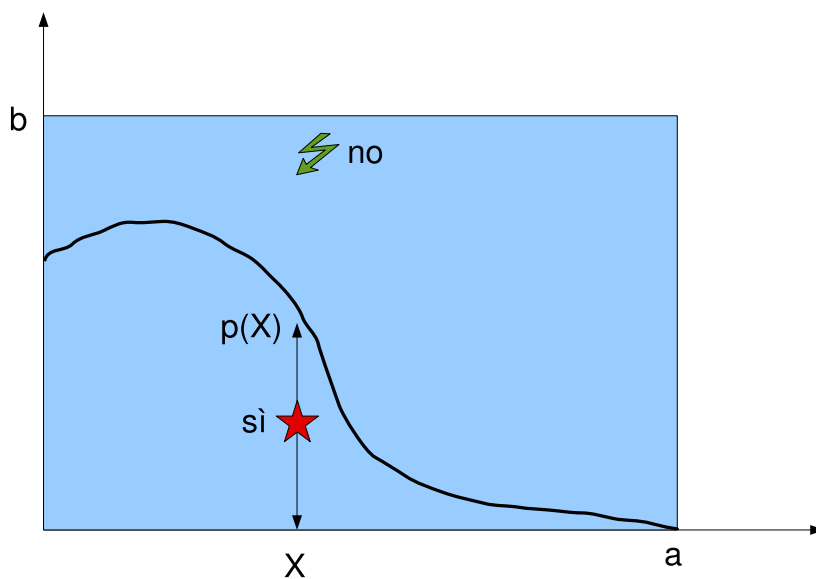
Percio' $p(y_1) p(y_2) = x_1 = \frac{1}{2\pi} e^{-(y_1^2 + y_2^2)/2}$

$$p(y_1) = \frac{1}{\sqrt{2\pi}} e^{-y_1^2/2} \quad p(y_2) = \frac{1}{\sqrt{2\pi}} e^{-y_2^2/2}$$

Distribuzioni qualsiasi

Se voglio ottenere una distribuzione con probabilità $p(x)$ arbitraria limitata superiormente.

- suppongo che mi interessi $0 \leq x \leq a$
- suppongo che $p(x) \leq b \quad \forall x$ in $[0, a]$
- scelgo un numero a caso x con distribuzione uniforme nel range $[0, a]$
- scelgo un secondo numero a caso y con $y \in [0, b]$
- se $y < p(x)$ accetto il numero, altrimenti procedo con altri due numeri



Applicazioni

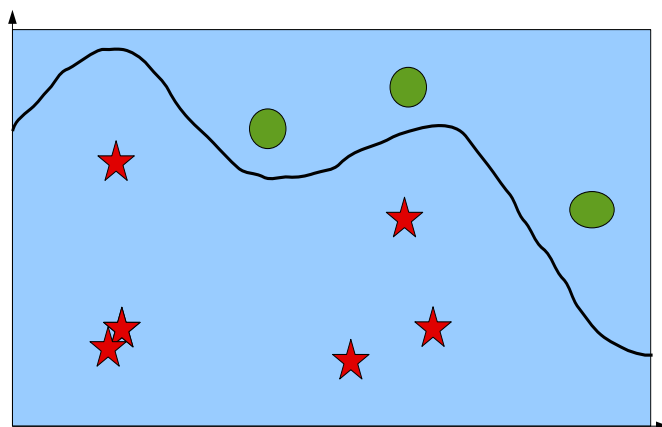
Integrazione con numeri pseudocasuali

Data $f(x)$ voglio calcolare

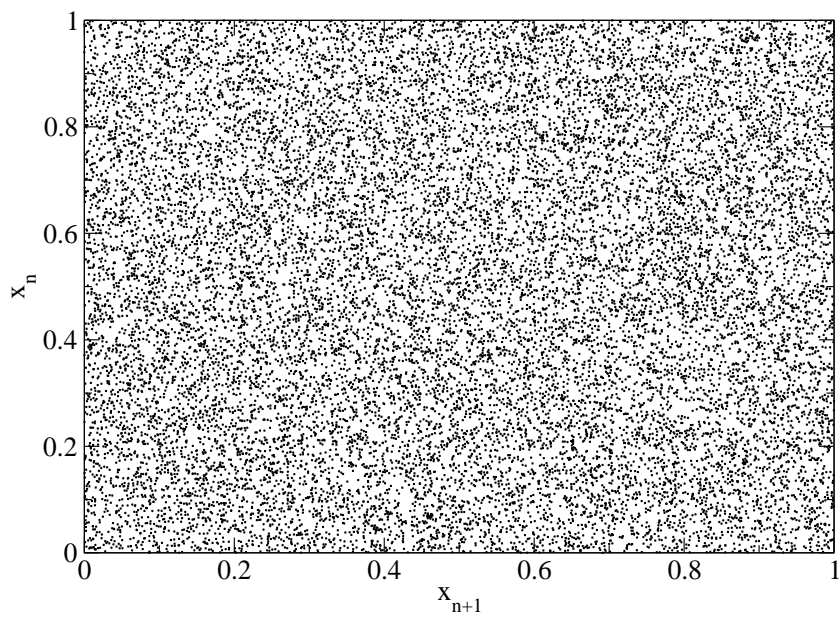
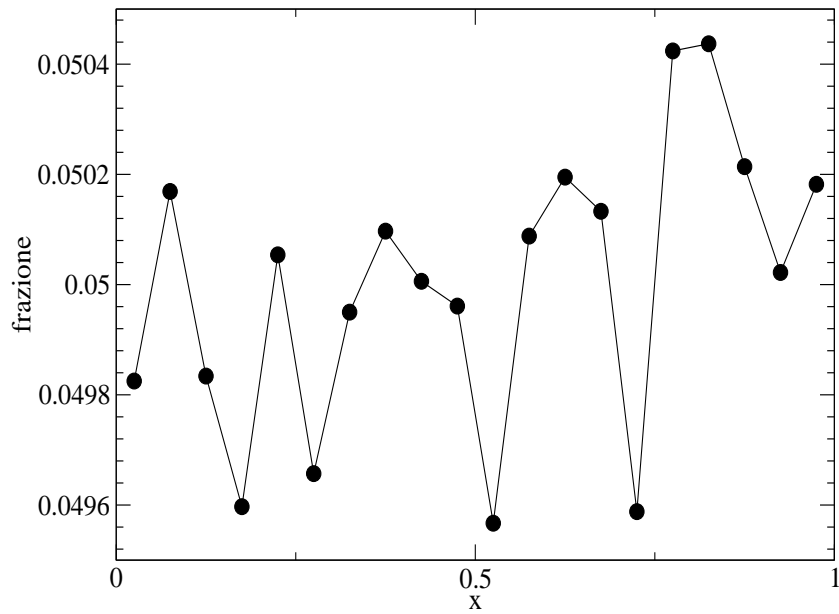
$$\int_a^b f(x) dx$$

dove so che $f(x)$ è compresa tra zero e f_{max} .

- genero coppie di numeri che corrispondono a un punto nel rettangolo che ha lati tra a e b e tra zero e f_{max} .
- calcolo la percentuale P di punti che cadono sotto la curva $y = f(x)$
- l'integrale vale $P \cdot (b - a) \cdot f_{max}$



Test di uniformità e correlazioni



Problemi

Volume di una sfera unitaria

- $x^2 + y^2 + z^2 \leq 1$;
- considero solo un ottante con $x > 0, y > 0, z > 0$;
- genero una terna di numeri casuali x, y e z ;
- guardo se $x^2 + y^2 + z^2 \leq 1$, nel qual caso accetto la terna;
- ripeto per un gran numero di terne;
- alla fine divido le terne accettate per quelle totali;
- moltiplico per 8;
- nota: l'errore è $O(1/\sqrt{N})$

Moto Browniano

- prendo $x = y = 0$ inizialmente;
- noto x al tempo t , al tempo successivo x' è dato da $x' = x + \epsilon\xi_1$ e $y' = y + \epsilon\xi_2$ con $-1 \leq \xi_{1,2} \leq 1$;
- dopo N passi verifico che $x^2 + y^2 \sim N\epsilon^2$;

Teorema del limite centrale

- considero le somme $\xi_1 \dots \xi_N$;
- guardo come è distribuita la media al variare di N ;
- per N grande devo trovare una distribuzione gaussiana;