Numeri pseudocasuali

- Il periodo deve essere il più lungo possibile;
- la distribuzione deve essere uniforme in $[0, 1]$
  $p(x)$=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

$$\begin{aligned} a \cdot x_n &= a \cdot (z \cdot q + w) = a \cdot q \cdot z + a \cdot w \\ &= (a \cdot q + r) \cdot z - r \cdot z + a \cdot w \\ a \cdot x_n &= m \cdot z - r \cdot z + a \cdot w \end{aligned}$$

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 risulato è
$a \cdot x_n \bmod m = a \cdot w - r \cdot z (+m) =$
$a \cdot (x_n \bmod m) - r \cdot (x_n/q)\,(+m)$

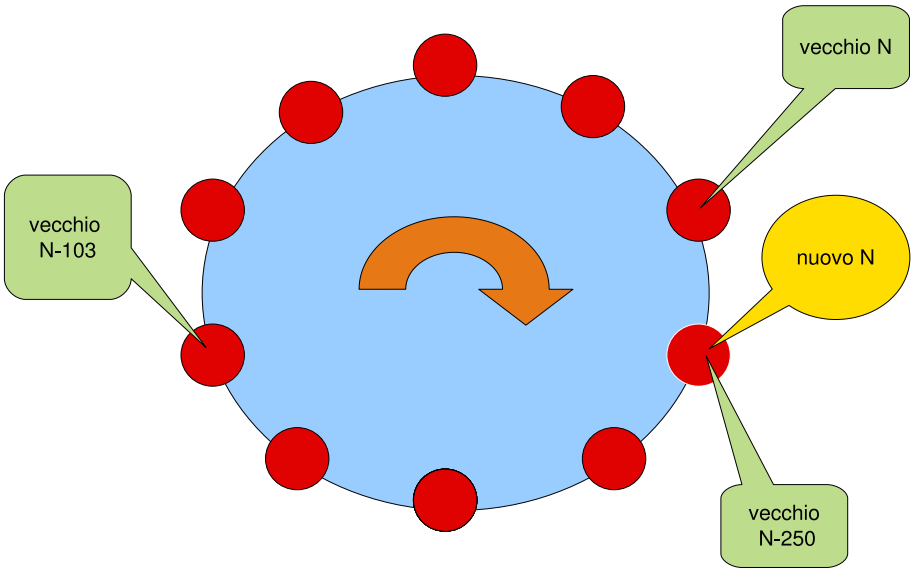
Esempio:
$a = 16807 \quad m = 2147483647 \quad q = 127773 \quad r = 2836$

Metodi Fibonacci lagged

- Simili alla successione di Fibonacci, ma con più termini;
- usano LCM come "starter" per i primi termini;
- hanno periodo molto lungo e scarse correlazioni;
- più successioni indipendenti con la stessa regola. Occorre una certa attenzione alle condizioni iniziali;

Esempio: r250

- $x_n = \sum_{j=1}^{250} a_q x_{n-q}$
- 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: $x_n = x_{n-103}\ xor\ x_{n-250}$
- 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 hanno distribuzione uniforme p_x in $(0,1)$ cerco la distribuzione p_y di y_1 e y_2 definiti da

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

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

Quindi $x_1 = \exp\left(-\frac{1}{2}(y_1^2+y_2^2)\right)$ e

$$p_y(y_1)\,p_y(y_2)\,dy_1\,dy_2 = p_x(x_1)\,p_x(x_2)\,dx_1\,dx_2 = dx_1\,dx_2$$

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

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

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

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

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

Gaussiana con media e varianza qualsiasi

- *voglio una distribuzione di probabilità della forma*

$$\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-(x-\mu)^2/2\sigma^2\right)$$

- *per ottenere la stessa distribuzione con media μ basta aggiungere μ ai numeri random generati*
- *per ottenere una varianza σ^2 basta moltiplicare il numero random per σ. Se*

$$p_x(x) = \frac{1}{\sqrt{2\pi}}\exp\left(-x^2/2\right)$$

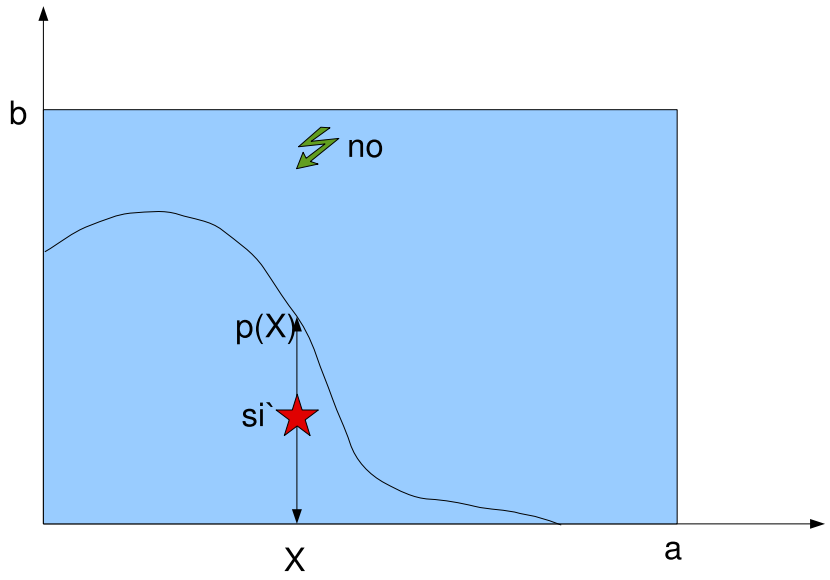
allora, definendo $y = \sigma \cdot x$ ottengo

$$p_y(y) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-y^2/2\sigma^2\right)$$

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 $y \epsilon [0, b]$
- se $y < p(x)$ accetto il numero, altrimenti procedo con altri due numeri



Applicazioni

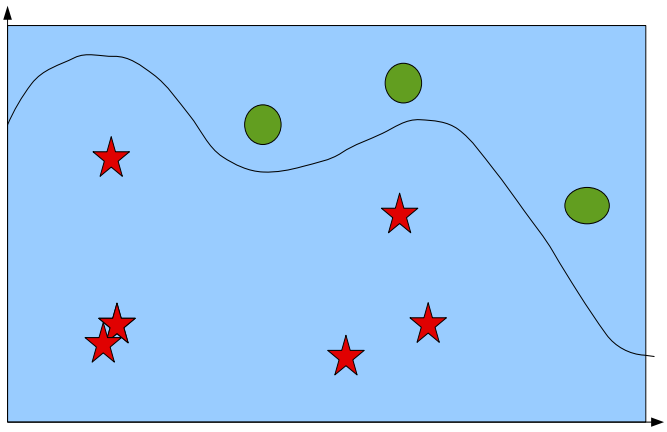
Integrazione con numeri pseudocasuali

Data $f(x)$ voglio calcolare

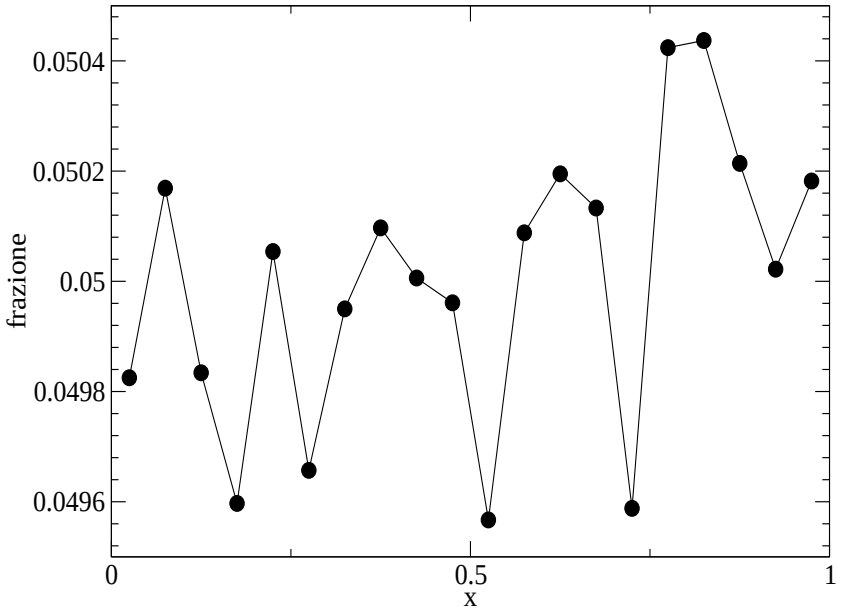
$$\int_a^b f(x)\,dx$$

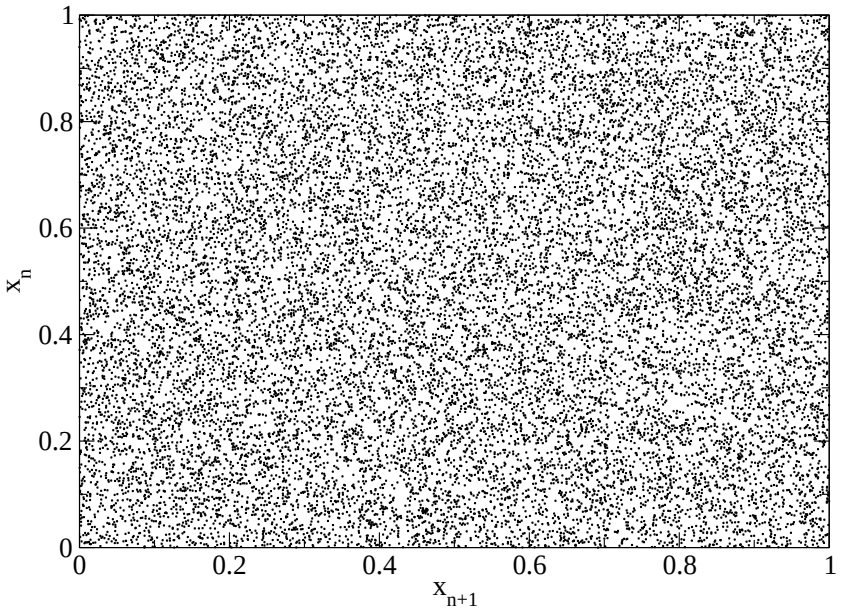
dove so che $f(x)$ è compresa tra zero e f_{max}.

- 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 $O(1/\sqrt{N})$
- In N dimensioni la sfera è data da $x_1^2 + \cdots + x_n^2 \leq 1$. Trovare il volume.

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 ϵ costante e ξ_1 e ξ_2 gaussiane;
- dopo N passi verifico che $x^2 + y^2 \sim N\epsilon^2$;

Teorema del limite centrale

- Considero le variabili $\xi_1 \dots \xi_N$;
- guardo come è distribuita la media $(\xi_1 + \xi_2 + \cdots + \xi_N)/N$ al variare di N;
- per N grande devo trovare una distribuzione gaussiana;