GENERAZIONI DI NUMERI CASUALI

Un generatore di numeri casuali è un programma la cui sequenza di output è la simulazione del comportamento di una sequenza di variabili casuali indipendenti. Tuttavia bisogna osservare che il comportamento del programma è fondamentalmente deterministico; a parità di seme iniziale la sequenza di output è la stessa.

Una sequenza di numeri pseudo-casuali deve soddisfare, al minimo, le seguenti proprietà statistiche:

1. distribuzione casuale degli elementi della sequenza secondo una funzione di distribuzione predefinita f(x): di solito si richiede una distribuzione uniforme su un intervallo specificato, cioè f(x) = 1 / (xmax − xmin) nell'intervallo [xmin,xmax] e f(x) = 0 fuori da tale intervallo.
2. indipendenza tra elementi successivi della sequenza: se la funzione di distribuzione per un singolo elemento è f(x), la funzione di distribuzione per le coppie x, y di elementi successivi deve essere f(x,y)= f(x)·f(y)

METODI DI GENERAZIONE DI NUMERI CASUALI

Esistono diverse classi di generatori di numeri pseudo-casuali, che si differenziano per il tipo di algoritmo usato. Spesso producono una sequenza di numeri interi uniformemente distribuiti tra 0 e un certo valore massimo, oppure di numeri reali tra 0 e 1. Prima di essere usato, un generatore deve essere inizializzato, cioè bisogna assegnare un opportuno valore a un parametro numerico, o gruppo di parametri, che viene chiamato seme. Ogni volta che si usa lo stesso seme, si otterrà sempre la stessa identica sequenza. Il periodo di un generatore non può quindi superare il numero dei possibili valori del seme.

Metodo della Trasformazione

Una variabile casuale r sia distribuita in modo uniforme. Si vuole determinare una funzione x(r) la quale segua una distribuzione f(x).

Sia p(r) la funzione densità di probabilità della variabile r. La probabilità di avere un valore tra r e r + dr, p(r)dr, deve essere uguale alla probabilità di avere un valore tra x(r) e x(r) + dx(r) cioè

$$p(r)\,dr = P(x)dx$$

Integrando:

$$\int_{-\infty}^{r} p(r)\,dr = \int_{-\infty}^{x} P(x)\,dx$$

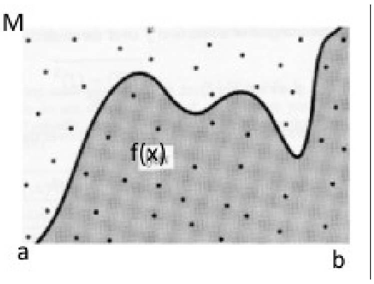
Se la distribuzione è uniforme, allora p(r) = 1 per $0 \le r \le 1$:

$$\int_{0}^{r} 1\,dr = \int_{-\infty}^{x} P(x)\,dx \rightarrow r = \int_{-\infty}^{x} P(x)\,dx$$

Se si riesce a risolvere analicamente la relazione precedente, si ottiene x(r). Spesso il metodo necessita di una integrazione numerica.

Rejection Method (Metodo dell'accettazione-rigetto)

Quando non si può applicare il metodo della trasformazione , allora si può utilizzare il Rejection Method.



1. Si vogliono generare numeri casuali compresi tra a e b distribuiti secondo una densità di probabilità P(x).
2. Supponiamo M è massimo valore della funzione P(x) in [a, b].
3. Generiamo due numeri casuali r_1 e r_2 uniformemente distribuiti tra [0, 1] e poniamo : $x_i = a + (b-a) \cdot r_1$ (la x_i sarà sempre compresa tra a e b) e per questo x_i prendiamo $y_i = M \cdot r_2$ (la y_i non sarà mai più grande di M)

4. Se $y_i < f(x_i)$ il punto viene accettato.
5. Se $y_i > f(x_i)$ il punto viene rigettato e si ritorna al punto 3. I punti accettati sono distribuiti come P(x) per costruzione.

In questo metodo se la P(x) fosse stretta avrebbe bassa efficienza (e quindi alto costo computazionale) perchè la gran parte dei punti generati starebbe fuori della funzione P(x) e quindi verrebbero rigettati. L'efficienza può migliorare se si stabiliscono gli estremi a e b in modo appropriato.

Metodo dell'utilizzo del Teorema del Limite Centrale

Una variabile casuale r sia distribuita in modo uniforme nell'intervallo [0,1] . Il valore di aspettazione è $\langle r \rangle = ½$ e la varianza $\sigma^2 = 1/12$.

Secondo il TCL se una variabile casuale è la somma di molte variabili casuali indipendenti ciascuna di media m_j e varianza σ^2_j :

$$x=\sum_{j=1}^{N} a_j r_j$$

la distribuzione della x tende ad essere distribuzione normale indipendentemente dalle distribuzione di r_j. Il valore aspettato della x sarà $m_x=\sum_{j=1}^{N} a_j m_j$ e la varianza $\sigma_x^2=\sum_{j=1}^{N} a_j^2 \sigma_j^2$

Se si sommano 12 valori casuali della nostra r_j:

$$x=\sum_{j=1}^{12} r_j$$

si ottiene una variabile gaussiana con valore aspettato 6 e varianza 1.

Metodo di Box-Muller.

Il metodo di Box-Muller (George Edward Pelham Box e Mervin Edgar Muller, 1958) è un metodo per generare coppie di numeri casuali indipendenti e distribuiti normalmente con media nulla e varianza uno.

Metodo:

1. Si estraggono due numeri r_1, r_2 dalla una distribuzione uniforme sull’intervallo (0,1].
2. Si può dimostrare che se costruiamo due variabili casuali z_1 e z_2:

$$z_1=\sqrt{-2\ln r_1}\cdot\cos(2\pi r_2) \quad \text{e} \quad z_2=\sqrt{-2\ln r_1}\cdot\sin(2\pi r_2)$$

Allora z_1 e z_2 sono variabili casualie indipendenti con distribuzione normale e di deviazione standard unitaria.

Esiste anche una versione, detta forma polare del metodo, che campiona due numeri su un intervallo differente ([−1,+1]) e permette di ricavare due numeri distribuiti normalmente senza l’uso delle funzioni seno e coseno. La forma polare differisce da quella base in quanto è anche una tecnica di rigetto. Vengono scartati alcuni numeri casuali, ma l’algoritmo è più veloce perché meno oneroso da valutare numericamente.