Generazione di determinazioni di modelli di probabilità$_0$

Nelle tecniche di simulazione per risolvere problemi complessi è spesso necessario generare determinazioni di modelli di probabilità assegnati, quali possono essere: binomiale, Poisson, normale, esponenziale negativa, uniforme, e quant'altro possa servire per descrivere nel modo più aderente e completo possibile la realtà che si vuole imitare.

Sorge allora il problema della generazione dei modelli di probabilità, che può essere validamente affrontato e risolto con l'ausilio del computer e l'applicazione di uno due seguenti metodi:

- Metodo della funzione di ripartizione inversa;
- Metodo del teorema centrale del limite.

Entrambi i metodi si appoggiano alla generazione di numeri pseudocasuali ottenuti con le funzioni proprie di ciascun linguaggio o foglio elettronico, o con i metodi classici, come ad esempio il metodo congruente lineare.

La scelta fra i due metodi dipende dal tipo di modello di probabilità da simulare, in particolare il secondo è usato per il modello normale, mentre il primo per tutti gli altri modelli, con i dovuti adattamenti.

Metodo della funzione di ripartizione inversa

Questo metodo si appoggia ad una delle proprietà che i generatori di numeri pseudocasuali devono possedere per essere considerati "buoni" e cioè che i numeri da essi ricavati devono seguire la distribuzione uniforme nell'intervallo [0,1). Poiché variano in [0,1) possono essere interpretati come valori di probabilità o, in modo ancora più comodo, come valori di Funzione di Ripartizione. Questa funzione, che permette di calcolare la probabilità di eventi composti del tipo $X \le x_0$, ha tre proprietà di forma, tra le quali la monotonicità non decrescente. Quest'ultima permette di ricavare la sua funzione inversa $F^{-1}(X = x_0)$ e dunque di risalire alla determinazione x_0 alla quale è associato il valore $F(X = x_0) = P(X \le x_0)$. È necessario a questo punto distinguere due situazioni:

a) la Funzione di Ripartizione del modello di probabilità da simulare è invertibile facilmente dal punto di vista matematico;
b) la Funzione di Ripartizione del modello di probabilità da simulare **non** è invertibile facilmente dal punto di vista matematico.

Fra i modelli che abbiamo studiato rientrano nel primo caso quello esponenziale negativo e quello uniforme, nel secondo tutti quelli discreti e cioè binomiale, di Poisson, e ipergeometrico, infatti dalla seguente tabella che riporta le forme analitiche delle Funzioni di Ripartizioni solo negli intervalli in cui sono diverse da 0 o da 1:

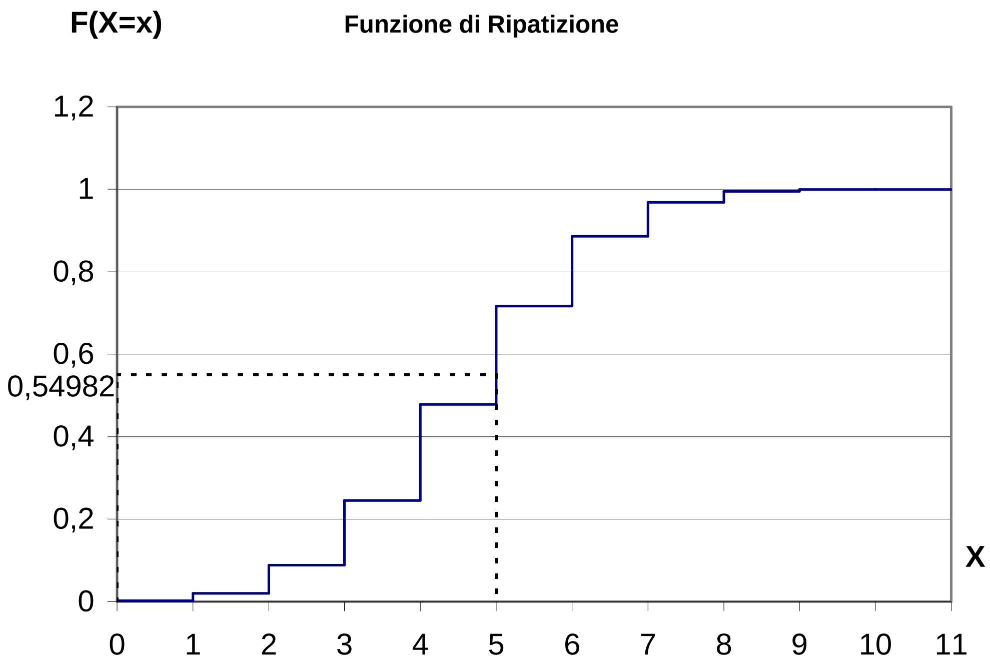
Modello	**Funzione di Ripartizione**	**Modello**	**Funzione di Ripartizione**
Esponenziale negativo	$F(X = x) = 1 - e^{-\lambda x}$	Poisson	$F(X = x) = \sum_{k=0}^{x} e^{-\lambda} \frac{\lambda^k}{k!}$
Uniforme [a;b]	$F(X = x) = \frac{x-a}{b-a}$	ipergeometrico	$F(X = x) = \sum_{i=0}^{x} \frac{\binom{K}{i}\binom{N-K}{n-i}}{\binom{N}{n}}$
binomiale	$F(X = x) = \sum_{k=0}^{x} \binom{n}{k} p^k (1-p)^{n-k}$		

si intuisce che solo per i primi due modelli è agevole dal punto di vista matematico ricavare le funzioni inverse, che diventano rispettivamente:

Modello	Funzione di Ripartizione inversa
Esponenziale negativo	$F^{-1}(X = x) = -\frac{1}{\lambda}\ln(1-y)$
Uniforme [a;b]	$F^{-1}(X = x) = (b-a)y + a$

essendo y = F(X = x) cioè, per quanto spiegato in precedenza, un numero pseudocasuale in [0;1).

Per i modelli discreti bisogna passare per via grafica, cioè disegnare il grafico della Funzione di Ripartizione del modello teorico e proiettare su di esso il valore pseudocasuale, in modo da individuare la determinazione che può averlo generato.



Il grafico precedente rappresenta la Funzione di Ripartizione di una v.c. Bi(10;0,46). Su di esso è messo in evidenza il numero pseudocasuale 0,54982, scelto come valore della F(X) e perciò è riportato sull'asse delle ordinate. Questo valore viene proiettato sul grafico della Funzione di Ripartizione fino ad intersecarlo e dal punto d'intersezione si traccia la proiezione sull'asse delle ascisse, individuando così la determinazione da associare al numero pseudocasuale, cioè x=5. Con l'ausilio del calcolatore il procedimento grafico può essere sostituito da una procedura che, se ripetuta n volte, permette di generare una quantità n anche elevata di determinazioni. I passi della procedura sono i seguenti:

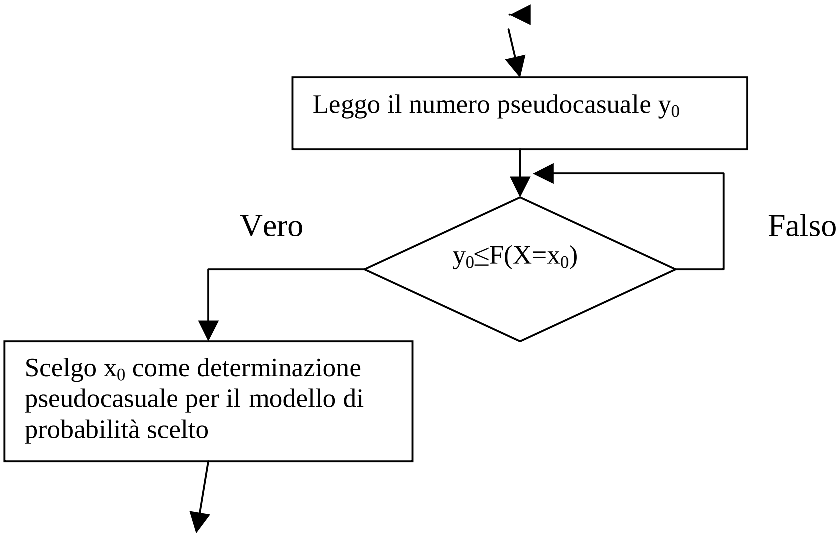
passo 1: scelto il modello e fissato/i il/i suo/i parametro/i, si costruisce la funzione di ripartizione;

passo 2: si genera il numero pseudocasuale U(0;1), che indichiamo con y_0;

passo 3: si confronta il numero y_0 con tutti i valori di F(X); il confronto prosegue finché y_0 è maggiore di F(X). Quando si trova il primo valore di Funzione di Ripartizione tale che $y_0 \leq F(X=x_0)$, allora si legge la corrispondente determinazione x_0 che viene assunta come determinazione pseudocasuale del modello di probabilità scelto;

passo 4: se si vuole generare una sequenza di n determinazioni si ritorna al passo 2.

Il diagramma che segue descrive i passi 2 e 3.



Metodo del teorema centrale del limite

Questo metodo è utilizzato per generare determinazioni di v.c. normali, poiché il teorema centrale sancisce la convergenza di v.c. trasformate lineari proprio al modello normale. In particolare si sfrutta la velocità della convergenza alla distribuzione normale della trasformata lineare mediante la formula:

$S_n = \sum_{i=1}^{n} U_i$ quando le v.c. U_i si distribuiscono in modo uniforme, e quindi si applica il seguente teorema centrale:

> *"sia data una successione di v.c. $U_i \sim U(0;1)$, tra loro stocasticamente indipendenti, con i = 1,2,..12, con media $M(U_i)$ = ½ e $Var(U_i)$ = 1/12, allora la v.c. trasformata $S_{12} \rightarrow N(1/2;1)$"*

Questo teorema indica che per ottenere **una determinazione** di v.c. N(1/2;1) sono necessari 12 numeri pseudocasuali nell'intervallo [0;1). Ovviamente se si vuole ottenere una determinazione di N(0;1) bisognerà sottrarre la media ½ alla determinazione costruita, cioè se con U_i indichiamo il generico numero pseudocasuale in [0;1) la formula per costruire **una determinazione z_0 di N(0;1) è**:

$$z_0 = \left(\sum_{i=1}^{12} U_i \right) - \frac{1}{2}$$

Per ottenere una determinazione x_0 di $N(\mu;\sigma^2)$ con $\mu \neq 0$ e $\sigma^2 \neq 1$ sarà necessario prima costruire una determinazione z_0 come descritto appena descritto e poi applicare la formula di standardizzazione inversa:

$$x_0 = \left[\left(\sum_{i=1}^{12} U_i \right) - \frac{1}{2} \right] * \sigma + \mu$$

Esempi

Modello esponenziale negativo

$\lambda = 2$	
n°pseudocasuale	det. Esp neg(λ = 2)
0,651481309	0,527031709
0,700748017	0,603234655
0,257606206	0,148937729
0,161877607	0,088295568
0,011303742	0,005684057
0,225378287	0,12769024

Modello uniforme

a = -4	
b = 7	
n°pseudocasuale	det. U(-4;7)
0,484652421	1,33117663
0,700748017	3,708228192
0,257606206	-1,166331732
0,161877607	-2,219346325
0,011303742	-3,875658837
0,937935527	6,317290802

Modello binomiale

n = 5			
p = 0,34			
k	F(X=k)	n°pseudocasuale	det Bi(5;0,34)
0	0,125233	0,651481309	2
1	0,447804	0,862096547	3
2	0,780149	0,971895981	4
3	0,951357	0,011303742	0
4	0,995456	0,998763158	5
5	1	0,225378287	1

Modello di Poisson

$\lambda = 2{,}7$			
k	F(X=k)	n°pseudocasuale	det di P (2,7)
0	0,06720551	0,651481309	3
1	0,2486604	0,700748017	3
2	0,49362449	0,956043748	6
3	0,71409218	0,161877607	1
4	0,86290786	0,011303742	0
5	0,94326833	0,865425512	5
6	0,97943055	0,439452831	2