

Appunti di Simulazione

M. Gianfelice

Corso di modelli probabilistici per le applicazioni
Master di II livello in Matematica per le Applicazioni

a.a. 2004/2005

1 Simulazione di numeri aleatori con distribuzione discreta

1.1 Metodo della *Trasformazione Inversa*

Supponiamo che U sia un numero aleatorio (n.a.) uniformemente distribuito nell'intervallo $[0, 1]$ con densità di probabilità assolutamente continua rispetto alla misura di Lebesgue. Allora vale il seguente risultato:

Teorema 1 *Se F è una funzione di distribuzione (o ripartizione) di un n.a. a valori interi non negativi, allora la variabile aleatoria X così definita:*

$$X = k \quad \text{se e solo se} \quad F(k-1) < U \leq F(k); \quad k \in \mathbb{N}$$

ha distribuzione F .

Dimostrazione: Per la definizione di funzione di distribuzione (ripartizione) di un n.a. si rimanda per es. a [BC] pag.109. Basta osservare che

$$\mathbb{P}(F(k-1) < U \leq F(k)) = F(k) - F(k-1)$$

ma per definizione $F(x) = \sum_{i \in I(X): i \leq x} \mathbb{P}(X = i)$, dove $I(X)$ è l'insieme dei valori possibili di X . ■

Esempio 2 Distribuzione di Bernoulli. *Poiché in questo caso $F(x) = (1-p)\mathbf{1}_{[0,1)}(x) + \mathbf{1}_{[1,+\infty)}(x)$, allora per il teorema precedente $X = 1$ se e solo se*

$$\begin{aligned} F(0) < U \leq F(1) \\ 1-p < U \leq 1 \\ \mathbb{P}(1-p < U \leq 1) &= \int_{1-p}^1 du = p \end{aligned}$$

ma $1 - p < U \leq 1$ è equivalente a $0 \leq 1 - U < p$. Ovviamente anche il n.a. $V = 1 - U$ è distribuito uniformemente in $[0, 1]$, cioè U e V sono equivalenti in distribuzione, pertanto $X = \mathbf{1}_{\{U \leq p\}}$.

Esempio 3 Distribuzione geometrica. Sia X un n.a. con distribuzione geometrica di parametro p , ovvero

$$\mathbb{P}(X = i) = pq^{i-1}, \quad i \geq 1, \quad q = 1 - p.$$

X rappresenta l'istante di primo successo in una successione di prove indipendenti che seguono la legge di Bernoulli (cfr [BC] pag. 26). Allora, per $j \geq 1$,

$$F(j-1) = \sum_{i=1}^{j-1} \mathbb{P}(X = i) = 1 - \mathbb{P}(X > j-1) = 1 - p \sum_{i \geq j} q^{i-1} = 1 - q^{j-1},$$

quindi per il teorema precedente, il n.a. X assumerà il valore j se estraendo U si ha

$$1 - q^{j-1} < U \leq 1 - q^j,$$

ovvero

$$q^j \leq 1 - U < q^{j-1}.$$

Quindi posto $X = \min\{j \geq 1 : q^j \leq 1 - U\}$, poiché la funzione $\log x$ è monotona crescente, si ha

$$X = \min\left\{j \geq 1 : j \geq \left\lceil \frac{\log(1 - U)}{\log q} \right\rceil\right\}$$

dove si è usato che essendo $0 < q, U < 1$, $\log q$ e $\log(1 - U)$ sono minori di 0. Pertanto, se $\lceil x \rceil$ indica la parte intera di x , cioè il numero intero tale che $0 \leq x - \lceil x \rceil < 1$, poiché $1 - U$ è anch'esso un n.a. uniformemente distribuito in $[0, 1]$, si ottiene che

$$X = \left\lceil \frac{\log U}{\log(1 - p)} \right\rceil + 1$$

ha distribuzione geometrica di parametro p .

Esempio 4 Distribuzione di Poisson. Sia X un n.a. con distribuzione di Poisson di parametro λ . Poiché

$$p_i = \mathbb{P}(X = i) = e^{-\lambda} \frac{\lambda^i}{i!}, \quad i \in \mathbb{N}$$

vale la relazione ricorsiva

$$p_{i+1} = \frac{\lambda}{i+1} p_i,$$

che può essere usata per calcolare la probabilità di X usando il seguente

Algoritmo:

passo 1 generare il n.a. U distribuito uniformemente in $[0, 1]$;

passo 2 porre $i = 0, p = e^{-\lambda}, F(i) = p$;

passo 3 se $U < F(i)$, allora porre $X = i$ e stop. Oppure andare al passo successivo;

passo 4 porre $p = \frac{\lambda}{i+1}p, F(i) = F(i) + p$ e $i = i + 1$;

passo 5 andare al passo 3.

Esempio 5 Distribuzione binomiale. Sia X un n.a. con distribuzione binomiale di parametri n e p . Poiché

$$\rho_i = \mathbb{P}(X = i) = \frac{n!}{i!(n-i)!} p^i (1-p)^{n-i}, \quad i \in \mathbb{N}$$

vale la relazione ricorsiva

$$\rho_{i+1} = \frac{n-i}{i+1} \frac{p}{1-p} \rho_i$$

che può essere usata per calcolare la probabilità di X usando il seguente

Algoritmo:

passo 1 generare il n.a. U distribuito uniformemente in $[0, 1]$;

passo 2 porre $i = 0, \rho = (1-p)^n, F(i) = \rho$;

passo 3 se $U < F(i)$, allora porre $X = i$ e stop. Oppure andare al passo successivo;

passo 4 porre $\rho = \frac{p}{1-p} \frac{n-i}{i+1} \rho, F(i) = F(i) + \rho$ e $i = i + 1$;

passo 5 andare al passo 3.

1.2 Metodo del *rigetto*

Supponiamo che si abbia un metodo efficiente per simulare un n.a. Y con distribuzione $\mathbb{P}_q := \{q_j, j \in \mathbb{N}\}$. È possibile usare quest'ultima per simulare un altro n.a. X che abbia distribuzione $\mathbb{P}_p := \{p_j, j \in \mathbb{N}\}$ simulando dapprima Y ed accettandone poi i valori ottenuti come valori di X con probabilità proporzionale al rapporto tra i corrispondenti valori di \mathbb{P}_q e \mathbb{P}_p . Infatti, vale il seguente risultato:

Teorema 6 *Dati il n.a. Y con distribuzione $\mathbb{P}_q := \{q_j, j \in \mathbb{N}\}$ e la distribuzione di probabilità $\mathbb{P}_p := \{p_j, j \in \mathbb{N}\}$, sia c una costante positiva tale che*

$$\frac{p_j}{q_j} \leq c \quad \forall j : p_j \neq 0.$$

Allora se U è un n.a. uniformemente distribuito in $[0, 1]$ stocasticamente indipendente da Y , si ha

$$\mathbb{P} \left(Y = i \mid \left\{ Y : U \leq \frac{\mathbb{P}_p(Y)}{c\mathbb{P}_q(Y)} \right\} \right) = p_i, \quad i \in \mathbb{N}.$$

Dimostrazione: Poiché Y ha distribuzione \mathbb{P}_q , per definizione di probabilità subordinata (cfr [BC] pag. 16) abbiamo

$$\begin{aligned} \mathbb{P} \left(Y = i \mid \left\{ Y : U \leq \frac{\mathbb{P}_p(Y)}{c\mathbb{P}_q(Y)} \right\} \right) &= \frac{\mathbb{P} \left(\{Y = i\} \cap \left\{ Y : U \leq \frac{\mathbb{P}_p(Y)}{c\mathbb{P}_q(Y)} \right\} \right)}{\mathbb{P} \left\{ Y : U \leq \frac{\mathbb{P}_p(Y)}{c\mathbb{P}_q(Y)} \right\}} \\ &= \frac{\mathbb{P} \left(Y = i, U \leq \frac{p_i}{cq_i} \right)}{\mathbb{P} \left\{ Y : U \leq \frac{\mathbb{P}_p(Y)}{c\mathbb{P}_q(Y)} \right\}}. \end{aligned}$$

Per l'indipendenza stocastica tra U ed Y si ha

$$\mathbb{P} \left(Y = i, U \leq \frac{p_i}{cq_i} \right) = \mathbb{P}_q(Y = i) \mathbb{P} \left(U \leq \frac{p_i}{cq_i} \right) = q_i \int_0^{\frac{p_i}{cq_i}} du = \frac{p_i}{c}.$$

D'altronde per la formula delle probabilità totali (cfr [BC] pag. 17),

$$\begin{aligned} \mathbb{P} \left\{ Y : U \leq \frac{\mathbb{P}_p(Y)}{c\mathbb{P}_q(Y)} \right\} &= \sum_{j \in \mathbb{N}} \mathbb{P} \left(\left\{ Y : U \leq \frac{\mathbb{P}_p(Y)}{c\mathbb{P}_q(Y)} \right\} \mid Y = j \right) \mathbb{P}_q(Y = j) \\ &= \sum_{j \in \mathbb{N}} \mathbb{P} \left(\left\{ Y : U \leq \frac{\mathbb{P}_p(Y)}{c\mathbb{P}_q(Y)} \right\} \mid Y = j \right) q_j, \end{aligned}$$

ma

$$\begin{aligned} \mathbb{P} \left(\left\{ Y : U \leq \frac{\mathbb{P}_p(Y)}{c\mathbb{P}_q(Y)} \right\} \mid Y = j \right) &= \frac{\mathbb{P} \left(\{Y = j\} \cap \left\{ Y : U \leq \frac{\mathbb{P}_p(Y)}{c\mathbb{P}_q(Y)} \right\} \right)}{\mathbb{P}_q(Y = j)} \\ &= \frac{\mathbb{P} \left(Y = j, U \leq \frac{p_j}{cq_j} \right)}{\mathbb{P}_q(Y = j)} = \mathbb{P} \left(U \leq \frac{p_j}{cq_j} \right) = \frac{p_j}{cq_j}, \end{aligned}$$

quindi

$$\mathbb{P} \left\{ Y : U \leq \frac{\mathbb{P}_p(Y)}{c\mathbb{P}_q(Y)} \right\} = \sum_{j \in \mathbb{N}} \frac{p_j}{c} = c^{-1},$$

da cui segue la tesi. ■

Pertanto possiamo simulare un n.a. X con distribuzione di probabilità $\mathbb{P}\{X = j\} = p_j, j \in \mathbb{N}$, per mezzo del seguente

Algoritmo:

passo 1 Simulare il valore di Y avente distribuzione di probabilità \mathbb{P}_q .

passo 2 Generare un n.a. U con distribuzione uniforme in $[0, 1]$.

passo 3 Se $U < \frac{\mathbb{P}_p(Y)}{c\mathbb{P}_q(Y)}$ porre $X = Y$ e stop. Altrimenti tornare al passo 1.

2 Simulazione di numeri aleatori con distribuzione assolutamente continua

2.1 Metodo della *Trasformazione Inversa*

In quanto segue supporremo, se non diversamente specificato, tutti i n.a. aventi distribuzione assolutamente continua (a.c.) rispetto alla misura di Lebesgue della retta reale (cfr [BC] pag. 41).

Lemma 7 *Siano X e Y due n.a. tali che X abbia distribuzione (a.c.) con densità ϕ e $Y = f(X)$, dove $f : (a, b) \subseteq \mathbb{R} \rightarrow \mathbb{R}$ sia $C^1((a, b); \mathbb{R})$ e monotona. Allora la densità di probabilità di Y è $\varphi(y) = \frac{\phi(f^{-1}(y))}{|f'(f^{-1}(y))|}$.*

Dimostrazione: Supponiamo che f sia crescente, allora per $f(a) < y < f(b)$

$$\mathbb{P}(Y \leq y) = \mathbb{P}(f(X) \leq y) = \mathbb{P}(X \leq f^{-1}(y)) = \int_{-\infty}^{f^{-1}(y)} dx \phi(x).$$

Posto $z = f(x)$, $dz = f'(x) dx$, ovvero $dx = \frac{dz}{f'(f^{-1}(z))}$. Quindi,

$$\mathbb{P}(X \leq f^{-1}(y)) = \int_{f(a)}^y dz \frac{\phi(f^{-1}(z))}{f'(f^{-1}(z))},$$

da cui segue $\varphi(y) = \frac{d}{dy} \mathbb{P}(Y \leq y) = \frac{\phi(f^{-1}(y))}{f'(f^{-1}(y))}$. Analogamente, se f è decrescente,

$$\mathbb{P}(Y \leq y) = \mathbb{P}(f(X) \leq y) = \mathbb{P}(X \geq f^{-1}(y)) = \int_{f^{-1}(y)}^{+\infty} dx \phi(x).$$

Ponendo di nuovo $z = f(x)$ si ha

$$\mathbb{P}(X \geq f^{-1}(y)) = \int_y^{f(b)} dz \frac{\phi(f^{-1}(z))}{f'(f^{-1}(z))},$$

da cui segue $\varphi(y) = \frac{d}{dy} \mathbb{P}(Y \leq y) = -\frac{\phi(f^{-1}(y))}{f'(f^{-1}(y))}$, ma poiché f è decrescente, $f'(x) \leq 0, \forall x \in (a, b)$. ■

Osservazione 8 Se l'insieme di definizione di f è l'unione disgiunta di intervalli $\bigcup_{k=1}^n (a_k, b_k) \subseteq \mathbb{R}$ in cui f è monotona e di classe C^1 allora se f_k denota la restrizione di f all'intervallo (a_k, b_k) , $k = 1, \dots, n$, si ha $\varphi(y) = \sum_{k=1}^n \frac{\phi(f_k^{-1}(y))}{|f'_k(f_k^{-1}(y))|} \mathbf{1}_{(f_k(a_k) \wedge f_k(b_k), f_k(a_k) \vee f_k(b_k))}(y)$.

Teorema 9 Sia U un n.a. con distribuzione uniforme in $[0, 1]$. Se F è una funzione di distribuzione allora il n.a. $X = F^{-1}(U)$ ha funzione di distribuzione F .

Dimostrazione: Dal Lemma precedente, dato che F è una funzione monotona crescente, differenziabile sul suo insieme di definizione esclusi, segue che

$$\mathbb{P}(X \leq x) = \mathbb{P}(F^{-1}(U) \leq x) = \mathbb{P}(U \leq F(x)) = \int_0^{F(x)} du = F(x).$$

■

Esempio 10 distribuzione esponenziale. La funzione di distribuzione esponenziale di parametro λ è $F(t) = (1 - e^{-\lambda t}) \mathbf{1}_{[0, +\infty)}(t)$ (cfr [BC] pag.45). Pertanto $F^{-1}(s) = -\frac{1}{\lambda} \log(1 - s) \mathbf{1}_{(0, 1]}(s)$, allora $X = -\frac{1}{\lambda} \log(1 - U)$ ha distribuzione esponenziale. Ma poiché $V = 1 - U$ è uniformemente distribuita in $[0, 1]$, X è equivalente in distribuzione a $Y = -\frac{1}{\lambda} \log U$.

2.2 Metodo del rigetto

2.2.1 Primo metodo

Vogliamo simulare un n.a. X con densità di probabilità f . Supponiamo di essere in grado di simulare un n.a. Y con distribuzione g . È possibile allora simulare il n.a. X con densità di probabilità f , simulando dapprima Y ed accettandone poi i valori ottenuti come valori di X con probabilità proporzionale al rapporto tra i corrispondenti valori di $f(Y)$ e $g(Y)$.

Teorema 11 Dati il n.a. Y con densità di probabilità g e la densità di probabilità f entrambe definite su $\mathcal{D} \subseteq \mathbb{R}$, sia c una costante positiva tale che

$$\frac{f(y)}{g(y)} \leq c \quad \forall y \in \mathcal{D}.$$

Allora se U è un n.a. uniformemente distribuito in $[0, 1]$ stocasticamente indipendente da Y , si ha

$$\mathbb{P}\left(Y \leq x \mid \left\{Y : U \leq \frac{f(Y)}{cg(Y)}\right\}\right) = \int_{-\infty}^x dy f(y), \quad x \in \mathcal{D}.$$

Dimostrazione: Poiché Y ha densità di probabilità g otteniamo

$$\mathbb{P}\left(Y \leq x \mid \left\{Y : U \leq \frac{f(Y)}{cg(Y)}\right\}\right) = \frac{\int_{-\infty}^x dy \mathbb{P}\left(U \leq \frac{f(Y)}{cg(Y)} \mid Y = y\right) g(y)}{\int_{-\infty}^{+\infty} dy \mathbb{P}\left(U \leq \frac{f(Y)}{cg(Y)} \mid Y = y\right) g(y)}.$$

D'altronde poiché U e Y sono stocasticamente indipendenti si ha che

$$\begin{aligned} \mathbb{P}\left(U \leq \frac{f(Y)}{cg(Y)} \mid Y = y\right) &= \frac{\mathbb{P}\left(U \leq \frac{f(Y)}{cg(Y)}, Y = y\right)}{\mathbb{P}(Y = y)} \\ &= \mathbb{P}\left(U \leq \frac{f(y)}{cg(y)}\right) = \int_0^{\frac{f(y)}{cg(y)}} du = \frac{f(y)}{cg(y)}, \end{aligned}$$

che sostituita nella precedente relazione da la tesi. ■

Pertanto possiamo simulare un n.a. X con densità di probabilità f , per mezzo del seguente

Algoritmo:

passo 1 Simulare il valore di Y avente densità di probabilità g .

passo 2 Generare un n.a. U con distribuzione uniforme in $[0, 1]$.

passo 3 Se $U \leq \frac{f(Y)}{cg(Y)}$ porre $X = Y$ e stop. Altrimenti tornare al passo 1.

2.2.2 Secondo metodo

Il vantaggio di questo metodo sta nel fatto che utilizza soltanto n.a. con distribuzione uniforme.

Teorema 12 *Sia f una densità di probabilità tale che:*

1. $f(x) = 0$ se $x \leq 0$;
2. $f(x) \leq 1 \wedge \frac{1}{x^2}$ se $x > 0$.

Siano inoltre U_1 e U_2 due n.a. stocasticamente indipendenti ed uniformemente distribuiti in $[0, 1]$. Se definiamo il n.a. $R = \frac{U_2}{U_1}$ e l'evento

$$E := \left\{U_1 \leq \sqrt{f\left(\frac{U_2}{U_1}\right)}\right\}, \text{ allora}$$

$$\mathbb{P}(R \leq x \mid E) = \int_0^x dy f(y).$$

Dimostrazione: Sia $T := \{(u_1, u_2) \in [0, 1]^2 : u_1 \leq \sqrt{f\left(\frac{u_2}{u_1}\right)}, u_2 \leq xu_1\}$, allora

$$\mathbb{P}(E \cap \{R \leq x\}) = \int \int_T du_1 du_2.$$

Utilizzando il cambio di variabili

$$\begin{cases} t = u_1 \\ s = \frac{u_2}{u_1} \end{cases} \implies \begin{cases} u_1 = t \\ u_2 = ts \end{cases}, \quad \left| \det \frac{\partial(u_1, u_2)}{\partial(s, t)} \right| = t,$$

otteniamo

$$\mathbb{P}(E \cap \{R \leq x\}) = \int_0^x ds \int_0^{\sqrt{f(s)}} t dt = \frac{1}{2} \int_0^x ds f(s).$$

Siccome $\mathbb{P}(E) = \lim_{x \rightarrow +\infty} \mathbb{P}(E \cap \{R \leq x\}) = \frac{1}{2}$, dalla definizione di probabilità subordinata si ha la tesi. ■

Pertanto al fine di generare un n.a. X con distribuzione f verificante le ipotesi del teorema precedente è possibile far uso del seguente

Algoritmo:

passo 1 Generare due valori u_1, u_2 del n.a. U con distribuzione uniforme in $[0, 1]$.

passo 2 Se $u_1 \leq \sqrt{f\left(\frac{u_2}{u_1}\right)}$, porre $X = \frac{u_1}{u_2}$ e stop. Altrimenti tornare al passo 1.

2.2.3 Simulazione di n.a. con distribuzione gaussiana

Oltre ai metodi sopra descritti, per generare n.a. con distribuzione gaussiana, è possibile utilizzare la rappresentazione in coordinate polari della densità congiunta di una coppia di n.a. (X, Y) stocasticamente indipendenti con distribuzione normale standard (cfr [BC] pag. 46).

Primo metodo (Metodo di *Box-Muller*) Sia (X, Y) un vettore di componenti due n.a. stocasticamente indipendenti identicamente distribuiti con distribuzione normale standard. La loro densità di probabilità congiunta (cfr [BC] pag. 51) risulta

$$f(x, y) = f_X(x) f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} = \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}}$$

Allora definendo i n.a.

$$\begin{aligned} R &:= X^2 + Y^2, \\ \Theta &:= \arctan\left(\frac{y}{x}\right) \end{aligned}$$

ed utilizzando le coordinate polari

$$\begin{cases} x = r \cos \theta \\ y = r \sin \theta \end{cases} \implies \left| \frac{\partial(x, y)}{\partial(r, \theta)} \right| = r$$

abbiamo che la loro densità di probabilità congiunta è

$$g(r, \theta) = f(x(r, \theta), y(r, \theta)) r = \frac{1}{2\pi} e^{-\frac{r^2}{2}} r = g_R(r) g_\Theta(\theta)$$

pertanto R e Θ sono stocasticamente indipendenti. Inoltre Θ è uniformemente distribuito in $[0, 2\pi]$. Ponendo $D := R^2$ otteniamo da g_R che la densità di probabilità relativa al n.a. D è $g_D(d) = \frac{1}{2} e^{-\frac{d}{2}}$ ovvero è una distribuzione esponenziale di parametro $\frac{1}{2}$. Chiaramente D è stocasticamente indipendente da Θ . Allora facendo riferimento all'Esempio 10 è possibile generare il vettore gaussiano (X, Y) generando prima il vettore (D, Θ) per poi utilizzare le coordinate polari, ovvero

Algoritmo:

passo 1 Generare due valori u_1 e u_2 del n.a. U uniformemente distribuito in $[0, 1]$.

passo 2 Porre $D = -2 \log u_1$ (cfr Esempio 10) e $\Theta = 2\pi u_2$.

passo 3 Porre

$$\begin{cases} X = R \cos \Theta = \sqrt{-2 \log u_1} \cos(2\pi u_2) \\ Y = R \sin \Theta = \sqrt{-2 \log u_1} \sin(2\pi u_2) \end{cases} \quad (1)$$

Secondo metodo Sfortunatamente il metodo precedentemente descritto per generare un vettore normale non è molto efficiente, poiché prevede il calcolo di funzioni trigonometriche. Si può allora usare il seguente metodo che prevede il calcolo indiretto del seno e del coseno di un angolo aleatorio. Infatti, se U_1 e U_2 sono due n.a. uniformemente distribuiti in $[0, 1]$, V_1 e V_2 tali che

$$\begin{cases} V_1 := 2U_1 - 1 \\ V_2 := 2U_2 - 1 \end{cases}$$

sono n.a. uniformemente distribuiti in $[-1, 1]$, pertanto il vettore aleatorio (V_1, V_2) è uniformemente distribuito in $[-1, 1]^2$. Inoltre, se (ρ, ϕ) denotano due n.a. indipendenti tali che $T := \rho^2$ e $2\pi\phi$ sono n.a. uniformemente

distribuiti in $[0, 1]$, posto $\rho^2 = V_1^2 + V_2^2$ ed $E := \{(V_1, V_2) : \rho^2(V_1, V_2) \leq 1\}$, subordinatamente all'evento E , è possibile generare la coppia di n.a.

$$\begin{cases} W_1 := \cos \phi = \frac{V_1}{\rho} = \frac{V_1}{\sqrt{V_1^2 + V_2^2}} \\ W_2 := \sin \phi = \frac{V_2}{\rho} = \frac{V_2}{\sqrt{V_1^2 + V_2^2}} \end{cases}$$

e dunque dalla (1) il vettore normale (X, Y) di componenti

$$\begin{cases} X = R \cos \Theta = \sqrt{-2 \log \rho^2} \cos \phi = V_1 \sqrt{\frac{-2 \log(V_1^2 + V_2^2)}{(V_1^2 + V_2^2)}} \\ Y = R \sin \Theta = \sqrt{-2 \log \rho^2} \sin \phi = V_2 \sqrt{\frac{-2 \log(V_1^2 + V_2^2)}{(V_1^2 + V_2^2)}} \end{cases}.$$

Quindi, per generare il vettore di n.a. stocasticamente indipendenti (X, Y) con distribuzione normale, si può far uso del seguente

Algoritmo:

passo 1 Generare due valori u_1 e u_2 del n.a. U uniformemente distribuito in $[0, 1]$.

passo 2 Porre $v_1 = 2u_1 - 1$, $v_2 = 2u_2 - 1$.

passo 3 Se $v_1^2 + v_2^2 > 1$, tornare al passo 1. Altrimenti porre

$$\begin{cases} X = v_1 \sqrt{\frac{-2 \log(v_1^2 + v_2^2)}{(v_1^2 + v_2^2)}} \\ Y = v_2 \sqrt{\frac{-2 \log(v_1^2 + v_2^2)}{(v_1^2 + v_2^2)}} \end{cases}.$$

Notiamo che, applicando questo metodo, la probabilità di rigettare i valori di V_1 e V_2 generati è $1 - \frac{\mathbb{P}(E)}{\mathbb{P}([-1, 1]^2)} = 1 - \frac{\pi}{4}$.

3 Simulazione di Catene di Markov omogenee a tempo discreto

Sia $\{X_n\}_{n \in \mathbb{N}}$ una successione di n.a. a valori in un insieme finito \mathcal{S} . Questa è detta *catena di Markov omogenea* con *spazio degli stati* \mathcal{S} se, $\forall n \in \mathbb{N}$, si ha

$$\mathbb{P}(X_0 = x_0, X_1 = x_1, \dots, X_n = x_n) = \rho_{x_0} p_{x_0, x_1} \cdots p_{x_{n-1}, x_n},$$

dove,

1. $\rho_{x_0} := \mathbb{P}(X_0 = x_0)$, $\forall x_0 \in \mathcal{S}$ è detta *distribuzione iniziale* della catena;

2. la matrice P tale che $(P)_{x,y} = p_{x,y}$, $x, y \in \mathcal{S}$ è detta *matrice di transizione* della catena ed ha le seguenti proprietà:

- (a) P è una matrice quadrata $|\mathcal{S}| \times |\mathcal{S}|$;
- (b) $0 \leq p_{x,y} \leq 1$;
- (c) $\sum_{y \in \mathcal{S}} p_{x,y} = 1$, $\forall x \in \mathcal{S}$.

Notiamo che da queste proprietà, applicando la definizione di probabilità subordinata alla formula precedente, segue che

$$\mathbb{P}(X_n = x_n | X_{n-1} = x_{n-1}, \dots, X_0 = x_0) = \mathbb{P}(X_n = x_n | X_{n-1} = x_{n-1}) = p_{x_{n-1}, x_n}.$$

Inoltre, se la catena è *irriducibile*, ovvero $\exists m > 0$ tale che $\min_{x,y \in \mathcal{S}} (P^m)_{x,y} > 0$, ed *aperiodica*, cioè $\forall x \in \mathcal{S}$, $MCD\{n \geq 0 : (P^n)_{x,x} > 0\} = 1$, vale il seguente risultato:

Teorema 13 (Teorema Ergodico) *Sia $\{X_i\}_{i \in \mathbb{N}}$ una catena di Markov omogenea irriducibile ed aperiodica con insieme degli stati finito. Allora esistono due costanti positive c e δ , con $\delta < 1$, tali che, $\forall x \in \mathcal{S}$ ed $n > 0$, si ha*

$$\left| (P^n)_{x,y} - \pi_y \right| \leq c\delta^n,$$

dove $(\pi_x)_{x \in \mathcal{S}}$ è la distribuzione stazionaria o invariante della catena, ovvero quella distribuzione di probabilità su \mathcal{S} che verifica il sistema:

$$\begin{cases} \pi_y = \sum_{x \in \mathcal{S}} \pi_x P_{x,y}, & y \in \mathcal{S} \\ \sum_{x \in \mathcal{S}} \pi_x = 1 \end{cases}.$$

Per i dettagli della dimostrazione di quest'ultime affermazioni e per ulteriori ragguagli sulla teoria delle catene di Markov si rimanda a [BC] pag. 65.

3.1 Simulazione mediante n.a. uniformemente distribuiti in $[0, 1]$

Poiché \mathcal{S} è finito, senza perdita di generalità, se $|\mathcal{S}| = n$, possiamo considerare $\mathcal{S} = \{1, \dots, n\}$. Supponendo di conoscere la matrice di transizione P , se la catena al tempo 0 si trova nello stato i , è possibile simulare la probabilità che la catena transisca in un passo nello stato j , mediante l'estrazione di un n.a. U uniformemente distribuito in $[0, 1]$. Infatti, fissando sull'intervallo $[0, 1]$ i punti $p_{i,1} + \mathbf{1}_{\{2, \dots, n\}}(j) \sum_{k=1}^j p_{i,k}$, $j = 1, \dots, n$, che non sono più di n dato che $p_{i,j}$ può essere nullo per qualche j , estraendo U , se

$$\begin{aligned} 0 \leq U < p_{i,1}, & \quad j = 1, \\ \sum_{k=1}^{j-1} p_{i,k} \leq U < \sum_{k=1}^j p_{i,k}, & \quad j = 2, \dots, n, \end{aligned}$$

si ha che la probabilità che la catena si sposti in un passo dallo stato i allo stato j è

$$\mathbb{P}(0 \leq U < p_{i,1}) = p_{i,1}, \quad j = 1,$$

$$\mathbb{P}\left(\sum_{k=0}^{j-1} p_{i,k} \leq U < \sum_{k=0}^j p_{i,k}\right) = p_{i,j}, \quad j = 2, \dots, n.$$

Quanto sopra può essere riscritto in forma algoritmica:

Algoritmo:

passo 1 Porre $m = 0$, $X_0 = i$.

passo 2 Generare il n.a. U uniformemente distribuito in $[0, 1]$.

passo 3 Se $0 \leq U < p_{i,1}$, porre $X_{m+1} = 1$ e andare al passo 5. Altrimenti andare al passo successivo.

passo 4 Porre $\lambda_{i,l} = \sum_{k=0}^{l-1} p_{i,k}$ e $j = \max\{l : \lambda_{i,l} \leq U\}$, $X_{m+1} = j$.

passo 5 Porre $m = m + 1$, $i = j$ e andare al passo 2.

3.2 Metodo Monte Carlo dinamico

Supponiamo di voler simulare una distribuzione di probabilità $(\pi_i)_{i=1,\dots,n}$ su $\{1, \dots, n\}$, allora è possibile costruire una catena di Markov omogenea che abbia come distribuzione stazionaria $(\pi_i)_{i=1,\dots,n}$. Chiaramente l'algoritmo che descrive una catena di Markov avente come distribuzione invariante $(\pi_i)_{i=1,\dots,n}$ è tanto più efficiente quanto più velocemente la catena converge a $(\pi_i)_{i=1,\dots,n}$.

Diamo ora un criterio per la costruzione di una matrice di transizione P tale che $\forall i = 1, \dots, n$, $\pi_i = \sum_{j=1}^n \pi_j P_{i,j}$ (cfr [P] pag. 97).

1. Sia Q una matrice di ordine n tale che:

- (a) $\forall i = 1, \dots, n$, $\sum_{j=1,\dots,n} Q_{i,j} = 1$;
- (b) se $Q_{i,j} > 0$ allora $Q_{j,i} > 0$.

2. Sia A una matrice di ordine n tale che $0 < A_{i,j} \leq 1$, detta *matrice d'accettazione*.

Allora definiamo $\mathcal{A}_i := \{j = 1, \dots, n : Q_{i,j} > 0\}$ e

$$P_{i,j} = \mathbb{P}(X_{n+1} = j | X_n = i) := \begin{cases} Q_{i,j} A_{i,j} & \text{se } j \in \mathcal{A}_i \setminus \{i\} \\ Q_{i,i} + \sum_{k \in \mathcal{A}_i \setminus \{i\}} (1 - A_{i,k}) Q_{i,k} & \text{se } j = i \\ 0 & \text{altrimenti} \end{cases} \quad (2)$$

ovvero, accetteremo con probabilità $A_{i,j}$ che la transizione dallo stato i allo stato j avvenga con probabilità $Q_{i,j}$, oppure la rigetteremo con probabilità $1 - A_{i,j}$.

Lemma 14 *La matrice P è tale che $\forall i = 1, \dots, n$, $\sum_{j=1, \dots, n} P_{i,j} = 1$.*

Dimostrazione: $\forall i = 1, \dots, n$ si ha

$$\begin{aligned} \sum_{j=1, \dots, n} P_{i,j} &= P_{i,i} + \sum_{j \in \mathcal{A}_i \setminus \{i\}} P_{i,j} = Q_{i,i} + \sum_{k \in \mathcal{A}_i \setminus \{i\}} (1 - A_{i,k}) Q_{i,k} + \sum_{j \in \mathcal{A}_i \setminus \{i\}} Q_{i,j} A_{i,j} \\ &= \sum_{j \in \mathcal{A}_i} Q_{i,j} = 1. \end{aligned}$$

■

Lemma 15 *La catena di Markov associata alla matrice P è irriducibile ed aperiodica.*

Dimostrazione: Segue dal fatto che $Q_{i,j} > 0$ implica $Q_{j,i} > 0$. ■

Proposizione 16 *Data la distribuzione di probabilità $(\pi_i)_{i=1, \dots, n}$, se gli elementi di matrice della matrice di accettazione A verificano la condizione*

$$\frac{A_{i,j}}{A_{j,i}} = \frac{\pi_j Q_{j,i}}{\pi_i Q_{i,j}} \quad \forall i = 1, \dots, n; j \in \mathcal{A}_i \setminus \{i\}, \quad (3)$$

il vettore $(\pi_i)_{i=1, \dots, n}$ è la distribuzione stazionaria per la matrice di transizione P .

Dimostrazione: È sufficiente verificare che P soddisfi la condizione di bilancio dettagliato, ovvero

$$\pi_i P_{i,j} = \pi_j P_{j,i} \quad \forall i, j = 1, \dots, n \quad (4)$$

che ovviamente implica l'invarianza della distribuzione $(\pi_i)_{i=1, \dots, n}$ per la catena. Infatti, sommando su j , per il Lemma 14, dalla precedente relazione otteniamo

$$\sum_{j=1}^n \pi_j P_{j,i} = \sum_{j=1}^n \pi_i P_{i,j} = \pi_i \sum_{j=1}^n P_{i,j} = \pi_i \quad \forall i = 1, \dots, n.$$

Infatti, dalla (3), $\forall i = 1, \dots, n$, $j \in \mathcal{A}_i \setminus \{i\}$, si ha

$$\pi_i P_{i,j} = \pi_i Q_{i,j} A_{i,j} = \pi_j Q_{j,i} A_{j,i} = \pi_j P_{j,i}.$$

■

Diamo ora un criterio per costruire la matrice d'accettazione A .

Lemma 17 Sia $\phi : \bar{\mathbb{R}} \rightarrow [0, 1]$ una funzione tale che

$$\frac{\phi(x)}{\phi\left(\frac{1}{x}\right)} = x \quad x \in \bar{\mathbb{R}}^+.$$

Allora, se $\forall i = 1, \dots, n, j \in \mathcal{A}_i \setminus \{i\}$,

$$A_{i,j} = \phi\left(\frac{\pi_j Q_{j,i}}{\pi_i Q_{i,j}}\right),$$

la matrice P definita nella (2) verifica la condizione di bilancio dettagliato (4).

Dimostrazione: Dalle ipotesi su ϕ ed A otteniamo che

$$A_{i,j} = \phi\left(\frac{\pi_j Q_{j,i}}{\pi_i Q_{i,j}}\right) = \frac{\pi_j Q_{j,i}}{\pi_i Q_{i,j}} \phi\left(\frac{\pi_i Q_{i,j}}{\pi_j Q_{j,i}}\right) = \frac{\pi_j Q_{j,i}}{\pi_i Q_{i,j}} A_{j,i}$$

da cui segue (3) e quindi, per quanto esposto nella dimostrazione della proposizione precedente, la tesi. ■

Resta da dimostrare l'esistenza di una funzione ϕ che verifichi le ipotesi di quest'ultimo lemma. Per esempio si può porre $\phi(x) = 1 \wedge x$, oppure $\phi(x) = \frac{x}{1+x}$.

Riferimenti bibliografici

- [BC] F. Biagini, M. Campanino *Appunti di calcolo delle probabilità e statistica matematica* (<http://www.dm.unibo.it/~biagini/book.pdf>).
- [I] M. Isopi *Appunti sintetici sulla simulazione* (<http://www.mat.uniroma1.it/people/bassan/simulazione.pdf>).
- [P] D. Petritis *Introduction à la modélisation et la simulation stochastiques* (<http://name.math.univ-rennes1.fr/dimitri.petritis/ps/mc.pdf>).
- [R] S. Ross *A course in Simulation* Macmillan Publishing Company, New York (1990).