PAC - PROBABILITA' AL CALCOLATORE
2002

E 1. *Variabili uniformi in* $[0,1]$.

- Simulare M variabili aleatorie $X_1, \ldots, X_M$ indipendenti e uniformemente distribuite nell'intervallo $[0,1]$.
- Calcolare la media empirica

$$S_M = \frac{1}{M} \sum_{i=1}^{M} X_i \tag{1}$$

e la varianza empirica

$$V_M = \frac{1}{M} \sum_{i=1}^{M} \left(X_i - S_M\right)^2 . \tag{2}$$

Graficarne l'andamento per diversi valori di M e osservare che per M grandi $S_M \sim \frac{1}{2}$ e $V_M \sim \frac{1}{12}$.

E 2. *Prove di Bernoulli*

- Simulare M v.a. $Y_1, \ldots, Y_M$ indipendenti, ciascuna distribuita come una variabile di Bernoulli di parametro p.
- Calcolare media e varianza empiriche (vedi (1) e (2)) nei casi $p = \frac{1}{4}, \frac{1}{2}, \frac{3}{4}$ e graficarne l'andamento al variare di M. Osservare che per M grande si ha $S_M \sim p$, $V_M \sim p(1-p)$.

E 3. *Gioco delle tre porte.*

Ci sono tre porte. Dietro una a caso delle tre c'è un premio, dietro le rimanenti due il vuoto. Scegliamo una delle porte. A questo punto il presentatore Monty Hall apre una delle due porte che rimangono dopo la nostra scelta, mostrando il vuoto. Rimangono quindi due porte chiuse una delle quali nasconde il premio. Abbiamo la scelta fra tre strategie:

- A: rimaniamo con la porta scelta all'inizio
- B: cambiamo porta
- C: tiriamo una moneta, se è testa rimaniamo, se è croce cambiamo

Simulare un numero M di prove del gioco e graficare la frequenza $f_M^{(A)}, f_M^{(B)}, f_M^{(C)}$ di vittorie con le diverse strategie A, B, C. Osservare che per M grande si ha $f_M^{(A)} \sim \frac{1}{3}$, $f_M^{(B)} \sim \frac{2}{3}$, $f_M^{(C)} \sim \frac{1}{2}$.

Note del corso di PAC, per informazioni: Pietro Caputo, `caputo@mat.uniroma3.it`.

E 4. *Distribuzione binomiale*

Siano $N \in \mathbb{N}$ e $p \in [0,1]$ due parametri e Z la v.a. con distribuzione binomiale definita dalle probabilità:

$$p_k = \mathbb{P}(Z = k) = \binom{N}{k} p^k (1-p)^{N-k}\,, \qquad k = 0, 1, 2, \dots, N \tag{3}$$

Simulare Z utilizzando N prove di Bernoulli di parametro p. Simulare quindi M versioni indipendenti $Z_1, \dots, Z_M$ della variabile binomiale Z definita da (3). Verificare la bonta' della simulazione graficando media e varianza empiriche $S_M = \frac{1}{M}\sum_{i=1}^M Z_i$ e $V_M = \frac{1}{M}\sum_{i=1}^M (Z_i - S_M)^2$ al variare di M. Si effettui l'esperimento nei casi $N = 10, 50, 500$ con $p = 0.4$ e si osservi che $S_M \sim Np$, $V_M \sim Np(1-p)$.

E 5. *Distribuzione geometrica*

Sia $p \in [0,1]$ un parametro e Z la v.a. definita da

$$p_k = \mathbb{P}(Z = k) = p(1-p)^{k-1}\,, \qquad k = 1, 2, \dots \tag{4}$$

Simulare Z come istante di primo successo in uno schema di Bernoulli (di lunghezza indefinita). Simulare quindi M versioni indipendenti $Z_1, \dots, Z_M$ e ripetere le usuali verifiche per medie e varianze empiriche nei casi $p = 0.01$ e $p = 0.5$. Osservare che $S_M \sim \frac{1}{p}$, mentre $V_M \sim \frac{1-p}{p^2}$.

E 6. *Fluttuazioni nel lancio di una moneta*

Una moneta (equa) viene lanciata N volte (per esempio $N = 10^4$). Il giocatore A ottiene un punto ogni volta che il lancio dà "testa". Il giocatore B ottiene un punto ogni volta che il lancio dà "croce".

- Si calcoli la frazione di tempo v in cui il giocatore A è in vantaggio sul giocatore B. Contrariamente a una concezione ingenua delle fluttuazioni si osserverá che i valori tipici di v non si avvicinano (nemmeno per N molto grande) al valor medio $\frac{1}{2}$.
- Si ripeta l'esperimanto M volte e si indichino con $v_1, v_2, \dots, v_M$ le frequenze corrispondenti. Si ponga, al solito $S_M = \frac{1}{M}\sum_{i=1}^M v_i$. Visualizzando l'andamento di S_M al crescere di M si osserverá che $S_M \sim \frac{1}{2}$, il che restituisce, almeno in parte, l'intuizione perduta.
- Siano $v_1, v_2, \dots, v_M$ come sopra e si costruisca un istogramma delle frequenze come segue. Suddividiamo l'intervallo $[0,1]$ in Δ (per es. $\Delta = 10$) parti uguali e calcoliamo la frazione di esperimenti che fornisce un valore della frequenza nei vari intervalli:

  $$f(k) = \frac{1}{M}\sum_{i=1}^M \mathbb{1}\Big(v_i \in \Big[\frac{k}{\Delta}, \frac{k+1}{\Delta}\Big)\Big)\,, \qquad k = 0, 1, 2, \dots, \Delta - 1$$

  Fissato M (per es. $M = 10^5$) si grafichi f al variare di k fra 0 e $\Delta - 1$. Si osservi la concentrazione intorno agli estremi.

E 7. *Limite di Poisson per distribuzione Binomiale*

Si considerino v.a. binomiali di parametri N e p con $N = 15, 20, 25, \ldots, 995, 1000$ e $p = \lambda/N$, per $\lambda = 1.5, 5, 8, 15$. Si calcolino le frequenze empiriche $f_{N,p}(\cdot)$ per variabili binomiali $Z_1, \ldots, Z_M$ di parametri N e p (per ognuno dei valori sopra elencati e per valori grandi di M, per es. $M = 10^5$):

$$f_{N,p}(k) = \frac{1}{M} \sum_{i=1}^{M} 1\!\!1(Z_i = k)\,, \qquad k = 0, 1, \ldots, N\,.$$

Si riportino su un unico grafico le funzioni $f_{N,p}(k)$ e $\pi_\lambda(k) := e^{-\lambda} \frac{\lambda^k}{k!}$, con $\lambda = Np$, al variare di k. Si verifichi la convergenza della distribuzione binomiale alla distribuzione di Poisson calcolando l'errore

$$\epsilon_{N,\lambda} = \frac{1}{2} \sum_{k=0}^{N} |f_{N,p}(k) - \pi_\lambda(k)|\,.$$

Si riportino su grafico i valori di $\epsilon_{N,p}$ al crescere di N per i diversi valori di λ.

E 8. *Monte Carlo per il calcolo di integrali*

- Si scriva un programma per calcolare l'integrale
$$\int_0^1 \frac{4}{1+x^2} = 4 \operatorname{arctg} x \Big|_0^1 = \pi$$
attraverso le somme parziali $S_M = \frac{1}{M} \sum_{i=1}^{M} f(X_i)$, dove $f(x) = \frac{4}{1+x^2}$ e $X_1, \ldots, X_M$ sono realizzazioni indipendenti di una v.a. uniforme in $[0, 1)$.
- Utilizzando la disuguaglianza di Chebyshev si dia una stima teorica del valore M^* tale che con probabilità maggiore di $\delta = 0.9$ la stima ottenuta con $M = M^*$ iterazioni sia precisa alla seconda cifra decimale. Si osservi che la stima ottenuta non è ottimale. Dare una stima empirica della probabilità che S_{M^*} sia invece precisa alla terza cifra decimale.
- Si calcoli l'integrale gaussiano
$$I(x) = \int_0^x \frac{\mathrm{e}^{-z^2/2}}{\sqrt{2\pi}}\, dz\,, \quad x \geq 0\,.$$
Graficare la funzione $I(x)$ per valori di x nell intervallo $[0, 10]$.

E 9. *Monte Carlo per il calcolo di aree (metodo del rigetto)*

- Si scriva un programma che utilizzi il metodo del rigetto per calcolare il valore di π come l'area di un cerchio di raggio 1 inscritto in un quadrato di lato 2.
- Si usi la disuguaglianza di Chebyshev per dare una stima teorica del valore M^* tale che con probabilità maggiore di $\delta = 0.9$ la stima ottenuta con M^* iterazioni sia precisa alla seconda cifra decimale.
- Si discuta l'efficienza dell'algoritmo rispetto all'esercitazione E 8.

E 10. *Metodo della Trasformazione per v.a. continue*

- *Esponenziale:* Simulare la v.a. continua X con densità $f(t) = \lambda\, e^{-\lambda t}$, $t \in [0, \infty)$, nei diversi casi $\lambda = 0.1, 1, 10$. Calcolare media e varianza empirica su un numero M di prove e confrontare graficamente con i valori teorici $\mathrm{E}[X] = \frac{1}{\lambda}$ e $\mathrm{Var}[X] = \frac{1}{\lambda^2}$.
- *Cauchy:* Si ripeta l'esperimento per la v.a. X con densità $f(t) = \frac{1}{\pi(1+x^2)}$, $t \in (-\infty, \infty)$.

E 11. *Algoritmo Box–Muller per v.a. Gaussiane*

- Simulare una v.a. Gaussiana X di media nulla e varianza $\sigma^2 = 0.1, 0.5, 1$ usando l'algoritmo Box–Muller. Graficare media e varianza empirica su un numero di prove M e stabilire un numero M_0 a partire dal quale la misura si stabilizza.
- Dare una stima empirica di $F(\xi) = \mathbb{P}[X \leq \xi]$ per diversi valori di ξ in $[0, \infty)$. Confrontare il risultato con il metodo Monte Carlo per il calcolo di integrali Gaussiani discusso in **E 8**.

E 12. *Teorema del limite centrale*

Si calcoli l'integrale Gaussiano

$$I(\xi) = \int_0^\xi \frac{\mathrm{e}^{-z^2/2}}{\sqrt{2\pi}}\, dz\,, \quad \xi \geq 0$$

utilizzando l'approssimazione del teorema del limite centrale

$$\mathbb{P}\Big(\frac{1}{\sigma\sqrt{M}} \sum_{i=1}^{M} (X_i - \mu) \in [0, \xi]\Big) \to I(\xi)\,, \qquad M \to \infty \tag{5}$$

dove $X_1, \ldots, X_M$ sono v.a. indipendenti e identicamente distribuite (i.i.d.) con media $\mu = \mathbb{E}(X_1)$ e varianza $\sigma^2 = \mathbb{E}(X_1^2) - \mu^2$, nei seguenti casi. Riportando i valori empirici per $I(\xi)$ su un grafico (come funzione di $\xi \in [0, 10]$ per esempio) si confrontino i risultati con le stime ottenute in **E 8** e **E 11**.

- *Bernoulli:* $X_1, \ldots, X_M$ sono bernoulliane di parametro $p = 0.1, 0.5$ ($\mu = p$ e $\sigma^2 = p(1-p)$).
- *Geometrica:* $X_1, \ldots, X_M$ sono geometriche di parametro $q = 0.1, 0.4$ ($\mu = \frac{1}{q}$, $\sigma^2 = \frac{1-q}{q^2}$).
- *Esponenziale:* $X_1, \ldots, X_M$ sono v.a. continue con legge esponenziale di parametro $\lambda = 0.1, 1$ ($\mu = \frac{1}{\lambda}$ e $\sigma^2 = \frac{1}{\lambda^2}$).

E 13. *Simulazione di v.a. di Poisson*

Simulare variabili aleatorie di Poisson X^λ di parametro $\lambda > 0$, utilizzando la rappresentazione in termini di v.a. di legge esponenziale. Calcolare media empirica S_M e varianza empirica V_M su un numero di prove M sufficientemente grande, al variare di λ tra 0.1 e 1. Graficare i valori ottenuti e confrontare con i valori teorici $\mathbb{E}[X^\lambda] = \mathrm{Var}[X^\lambda] = \lambda$.

E 14. *Catena di Markov: 3 colori*

Supponiamo di colorare i vertici A, B, C, D di un quadrato nel modo seguente: disponiamo di 3 colori e due vertici adiacenti non possono avere lo stesso colore. Sia $\sigma = \{\sigma(A), \sigma(B), \sigma(C), \sigma(D)\}$ la configurazione di colori sui vertici A, B, C, D e sia Ω lo spazio delle configurazioni σ realizzabili con la suddetta regola. Sia ν la misura di probabilità uniforme su Ω.

- Costruire una catena di Markov $\{\sigma_1, \sigma_2, \ldots\}$ su Ω che ha come misura invariante ν.
- Simulare l'evoluzione temporale della catena di Markov per una fissata condizione iniziale di colori σ_1. Verificare la convergenza della catena alla misura invariante ν calcolando la frequenza di un'arbitraria configurazione di riferimento $\xi \in \Omega$. Indicare il tempo necessario per poter parlare di "perdita di memoria" della catena.

- Sia Ω_0 il sottoinsieme di Ω in cui vertici opposti hanno lo stesso colore:

  $$\Omega_0 = \{\sigma \in \Omega : \ \sigma(A) = \sigma(C) \ \text{ e } \ \sigma(B) = \sigma(D)\}\,.$$

  Sia h_n il numero di volte in cui la catena si è trovata in Ω_0 fino al tempo n:

  $$h_n = \sum_{k=1}^{n} \mathbb{1}(\sigma_k \in \Omega_0)\,.$$

  Verificare graficamente la convergenza per un valore arbitrario di σ_1:

  $$\lim_{n\to\infty} \frac{h_n}{n} = \nu(\Omega_0) = \frac{1}{3}\,.$$

- Estendere l'analisi al caso di un esagono o altro poligono regolare.

Note, suggerimenti e algoritmi

Numeri pseudo–random. La simulazione di (sequenze di) numeri aleatori tramite calcolatore è un argomento molto studiato e con notevoli applicazioni. Per chi fosse interessato al tema sono dedicati ampi spazi sul web, si veda per es. il sito [2]. Per gli scopi di questo corso ci accontenteremo delle funzioni standard incluse nei comuni linguaggi di programmazione. Nel linguaggio C (l'unico a cui faremo esplicito riferimento in seguito), ad esempio si puo' utilizzare la funzione `rand()` (che fornisce un intero pseudo-random tra 0 e un valore massimo `RAND_MAX`) oppure `drand48()` (che fornisce un numero reale pseudo-random tra 0 e 1). In entrambi i casi le funzioni vanno inizializzate con un opportuno *seme*. Nel primo caso il comando è `srand(seme)` mentre nel secondo si usa `srand48(seme)`. Qui *seme* è un numero arbitrario che va cambiato ad ogni nuovo giro del programma. Per comodità si puo' usare il timer interno della macchina e porre il *seme* uguale a `(unsigned)time(NULL)` (in questo caso è necessario il comando `#include <time.h>` all'inizio del programma). Esempio di uso:

```
double r;
srand48((unsigned)time(NULL));
r = drand48();
```

fornisce una variabile r in doppia–precisione "distribuita uniformemente" in $[0, 1)$.

Grafici. Per i grafici si fa riferimento al programma GNUPLOT. Si veda la pagina web [3] per un'introduzione all'uso.

E 1. Un modo semplice di simulare le M variabili $X_1, \ldots, X_M$ è quello di inserire la riga `r=drand48();` all'interno di un ciclo `for (i=0; i < M; i++)`.

E 2. Per esempio

$$Y_i = 1\!\!1(X_i \le p) = \begin{cases} 1 & \text{se } X_i \le p \\ 0 & \text{se } X_i > p \end{cases}$$

se X_i sono uniformi in $[0, 1)$ (come in E 1).

E 3. Simulazione di una prova:

- il premio è X: v.a. uniforme in $\{0, 1, 2\}$
- la nostra prima scelta è Y: v.a. uniforme in $\{0, 1, 2\}$
- la porta aperta da Monty Hall puo' essere definita per esempio scegliendo una porta a caso Z (v.a. uniforme in $\{0, 1, 2\}$) finché questa è diversa da X e da Y (mediante un semplice ciclo `do...while`)
- Le vincite secondo le tre strategie sono: $A = 1\!\!1_{\{X=Y\}}$, $B = 1\!\!1_{\{X=3-Y-Z\}}$ (notare che $3 - Y - Z$ è la porta che si ottiene cambiando) e infine

  $$C = \alpha 1\!\!1_{\{X=Y\}} + (1 - \alpha) 1\!\!1_{\{X=3-Y-Z\}},$$

  dove α è v.a. di Bernoulli (indipendente) di parametro $p = \frac{1}{2}$.

Simulazione di M prove e calcolo frequenze:

- Chiamiamo X_i, Y_i, Z_i, $i = 1, \ldots, M$ le variabili sopra definite, scelte in maniera indipendente ad ogni prova.
- A_i, B_i, C_i denotano le corrispondenti vincite relative alla la prova i–esima
- Calcoliamo le frequenze

$$f_M^{(A)} = \frac{1}{M} \sum_{i=1}^{M} A_i\,, \quad f_M^{(B)} = \frac{1}{M} \sum_{i=1}^{M} B_i\,, \quad f_M^{(C)} = \frac{1}{M} \sum_{i=1}^{M} C_i$$

E 4. Con riferimento a E 2 basterà porre $Z = \sum_{i=1}^{N} Y_i$. Questo fornisce una singola realizzazione della binomiale di parametri N, p. Ripetendo M volte l'operazione si ottiene la sequenza $Z_1, Z_2, \dots, Z_M$.

E 5. La v.a. aleatoria Z puo' essere rappresentata come

$$Z = \min\{m \in \mathbb{N} : Y_m = 1\}$$

se $Y_1, Y_2, \dots$ sono v.a. di Bernoulli. Poiché il numero di prove di Bernoulli è indefinito qui si deve usare un ciclo del tipo

`do ... while`

che viene interrotto al primo successo.

E 6. Siano $X_1, \dots, X_N$ bernoulliane indipendenti di parametro $p = \frac{1}{2}$ (i risultati degli N lanci di una moneta). Ponendo $Y_i = 2X_i - 1$ si ha $+1$ se vince A e -1 se vince B. Al tempo $k = 1, 2, \dots, N$ il bilancio tra i due giocatori è dato da

$$Z_k = \sum_{i=1}^{k} Y_i = \#(\text{vincite } A) - \#(\text{vincite } B)$$

e diremo che A è in vantaggio se $Z_k > 0$. Quindi v si puo' calcolare mediante la formula

$$v = \frac{1}{N} \sum_{k=1}^{N} 1\!\!1(Z_k > 0) + \frac{1}{2N} \sum_{k=1}^{N} 1\!\!1(Z_k = 0) .$$

Qui il secondo termine è stato aggiunto per motivi di simmetria (altrimenti non è esatto dire che $S_M \to \frac{1}{2}$). Ripetendo M volte l'esperimento si ottengono i valori di S_M e dell'istogramma $f(k)$. Per il calcolo di $f(k)$ si deve fare attenzione a non escludere gli estremi $v = 0, v = 1$. L'espressione data nel testo per esempio include $v = 0$ ma esclude $v = 1$. Si consiglia quindi di correggerla aggiungendo il valore $\frac{1}{M} \sum_{i=1}^{M} 1\!\!1(v_i = 1)$ alla variabile $f(\Delta - 1)$.

Si puo' osservare che l'istogramma f approssima (al crescere del parametro Δ) la densità di probabilità

$$\frac{1}{\pi\sqrt{x(1-x)}} , \qquad x \in [0, 1] .$$

Per una discussione dettagliata di questo e altri aspetti della cosiddetta *legge dell'arcoseno* rimandiamo a [1].

E 7. Per i diversi valori di λ si suggerisce di calcolare la stringa $\pi_\lambda(\cdot)$ utilizzando le relazioni ricorsive

$$\pi_\lambda(0) = e^{-\lambda} , \qquad \pi_\lambda(k) = \frac{\lambda}{k} \pi_\lambda(k-1) .$$

Per ogni valore di N e $p = \lambda/N$ le variabili $Z_1, \dots, Z_M$ (diciamo per $M = 10^5$) si possono ottenere tramite il programma in E 4. E' importante scegliere M grande altrimenti le frequenze sono un'approssimazione troppo povera della distribuzione binomiale e non garantiscono che $\epsilon_{N,\lambda} \to 0$ al crescere di N (a questo proposito si puo' considerare il caso $M = 10^3$ e osservare che $\epsilon_{N,\lambda} \not\to 0$).

E 8. Ricordiamo che se X è v.a. uniforme in $[0,1)$, data una funzione continua $f:[0,1]\to\mathbb{R}$ si ha $\mathbb{E}f(X)=\int_0^1 f(x)\,dx$. Per la legge dei grandi numeri tale valore può essere approssimato dalla media empirica $S_M=\frac{1}{M}\sum_{i=1}^M f(X_i)$ dove $X_1,\dots,X_M$ sono M realizzazioni indipendenti della variabile X. Scegliendo $f(x)=\frac{4}{1+x^2}$ e simulando le variabili X_i otteniamo dunque una stima di $\pi=3.14159265358\dots$.

La disuguaglianza di Chebyshev permette di stimare

$$\mathbb{P}\big[|S_M-\pi|\geq a\big]\ \leq\ \frac{\mathbb{E}\big[(f(X)-\pi)^2\big]}{a^2M}\,,\qquad a>0\,.$$

Osserviamo che $|f(x)-\pi|\ \leq\ \max f\ -\ \min f\ =\ 2$ e quindi $\mathbb{E}\big[(f(X)-\pi)^2\big]\ \leq\ 4$. Inserendo $a=0.01$ e richiedendo che la probabilità di avere $|S_M-\pi|\geq a$ sia minore di $0.1=1-\delta$ si ottiene $M^*=4\times10^5$. Ripetendo la stima S_{M^*} per N volte (per esempio $N=100$) si osserverà che $|S_{M^*}-\pi|<0.01$ molto piú che nove volte su dieci. Si osserverà invece che all'incirca sette volte su dieci $|S_{M^*}-\pi|<a'$ con $a'=0.001$.

Per il calcolo dell'integrale $I(x)$ si utilizza la formula generale

$$\int_a^b f(y)\,dy=(b-a)\int_0^1 f(a+(b-a)y)\,dy\,,\qquad -\infty<a<b<\infty\,.$$

Poniamo $a=0$, $b=x$, $f(y)=\frac{e^{-y^2/2}}{\sqrt{2\pi}}$. Se Z è v.a. uniforme in $[0,1)$ si ha dunque $I(x)=x\,\mathbb{E}f(xZ)$. Dalla legge dei grandi numeri segue che se $Z_1,\dots,Z_M$ sono realizzazioni indipendenti di Z allora

$$W_M=\frac{x}{M}\sum_{i=1}^M f(xZ_i)$$

approssima l'integrale $I(x)$ al crescere di M. Il grafico di $I(x)$ (fissando per esempio $M=10^4$ iterazioni) al crescere di x da 0 a 10 mostra la rapida convergenza $I(x)\to I(\infty)=\frac{1}{2}$.

E 9. Sia Q il quadrato di lato 2 centrato nell'origine del piano cartesiano. Sia C il cerchio di raggio 1 con centro nell'origine. Siano u e v due v.a. (indipendenti) uniformi in $[0,1)$ e si ponga $x=-1+2u$, $y=-1+2v$. Osserviamo che il punto $P=(x,y)$ di coordinate x,y è distribuito uniformemente in Q. Definiamo $Z=\mathbb{1}(P\in C)=\mathbb{1}(x^2+y^2\leq1)$. Abbiamo

$$\mathbb{P}(P\in C)=\mathbb{E}Z=\frac{\text{area di } C}{\text{area di } Q}=\frac{\pi}{4}$$

Per la legge dei grandi numeri possiamo quindi stimare π tramite

$$S_M=\frac{4}{M}\sum_{i=1}^M Z_i$$

dove $Z_1,\dots,Z_M$ sono realizzazioni indipendenti di Z.

Un semplice conto mostra che $\mathbb{E}\big[(4Z-\pi)^2\big]\ =\ \pi(4-\pi)\leq 3$ e usando la disuguaglianza di Chebyshev come sopra otteniamo

$$\mathbb{P}\big[|S_M-\pi|\geq a\big]\ \leq\ \frac{\mathbb{E}\big[(4Z-\pi)^2\big]}{a^2M}\ \leq\ \frac{3}{a^2M}\,.$$

Se $\delta=0.9$ e $a=0.01$ abbiamo $M^*=3\times10^5$. Anche qui è facile verificare che la stima trovata non è ottimale. Per questo valore di M però all'incirca solo tre volte su dieci si ha precisione alla terza cifra decimale. In particolare, un confronto con i risultati del precedente programma mostra che ai fini di un calcolo di π il metodo del rigetto è tra i due il meno efficiente.

E 10. Il metodo della trasformazione può essere brevemente illustrato come segue. Sia X una v.a. continua con densità $f(t)$, $t \in [x_1, x_2] \subset \mathbb{R}$. La funzione di distribuzione associata è $F(x) = \mathbb{P}[X \leq x] = \int_{x_1}^{x} f(t)dt$. Pertanto si ha

$$\mathbb{P}\big[X \in [a,b]\big] = \int_a^b f(t)dt = F(b) - F(a)\,, \quad x_1 \leq a \leq b \leq x_2\,. \tag{6}$$

Se Y è v.a. uniforme in $[0,1]$ allora la (6) equivale a $\mathbb{P}\big[Y \in [F(a), F(b)]\big]$. In particolare, se $F : [x_1, x_2] \to [0,1]$ è strettamente crescente allora abbiamo $\{X \in [a,b]\} \iff \{F(X) \in [F(a), F(b)]\}$ e la (6) stabilisce che la v.a. $Y = F(X)$ è distribuita uniformemente in $[0,1]$, qualunque sia la distribuzione di X. Invertendo F otteniamo che X è distribuita come $F^{-1}(Y)$. Quindi se si dispone di un'espressione esplicita per F^{-1} il metodo della trasformazione permette di simulare X calcolando $F^{-1}(Y)$ dove Y è uniforme in $[0,1]$. Nei casi proposti:

- se $f(t) = \lambda e^{-\lambda t}$, $t \geq 0$ allora $F(x) = \int_0^x f(t)dt = 1 - e^{-\lambda x}$, $x \geq 0$. Quindi $F^{-1}(y) = -\frac{1}{\lambda}\log(1-y)$, $y \in (0,1)$.
- se $f(t) = \frac{1}{\pi(1+x^2)}$, $t \in \mathbb{R}$ si ha $F(x) = \int_{-\infty}^{x} f(t)dt = \frac{1}{2} + \frac{1}{\pi}\mathrm{arctg}\, x$. Invertendo: $F^{-1}(y) = \mathrm{tg}\,[\pi(y - \frac{1}{2})]$, $y \in (0,1)$.

E' noto che la v.a. aleatoria di Cauchy ha media nulla (per simmetria) ma varianza infinita.

E 11. Poiché l'integrale Gaussiano $\mathbb{P}[X \leq \xi] = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\xi} e^{-\frac{t^2}{\sigma^2}} dt$ non è noto analiticamente, il metodo della trasformazione non si può applicare direttamente in questo caso. L'algoritmo di Box–Muller si basa sull'osservazione che se R è v.a. continua in $[0,\infty)$ con densità $f_R(r) = \frac{r}{\sigma^2} e^{-\frac{r^2}{\sigma^2}}$ e Θ è v.a. uniforme in $[0, 2\pi)$, allora $X = R\sin\Theta$ è v.a. Gaussiana di media nulla e varianza σ^2. Quest'ultimo fatto si può verificare facilmente tramite un passaggio a coordinate polari nel piano cartesiano. (Lo stesso vale per la variabile $X' = R\cos\Theta$.) Ora la variabile Θ si può simulare banalmente come $\Theta = 2\pi Y$ con Y uniforme in $[0,1)$ mentre per la variabile R si può applicare il metodo della trasformazione: Infatti

$$F(x) = \frac{1}{\sigma^2} \int_0^x r\, e^{-\frac{r^2}{\sigma^2}} dr = 1 - e^{-\frac{x^2}{\sigma^2}}\,, \quad x \geq 0\,.$$

e invertendo si ha $F^{-1}(y) = \sqrt{-2\sigma^2 \log(1-y)}$, $y \in (0,1)$.

Ripetendo un certo numero M_0 di volte la simulazione di X si può stimare empiricamente la probabilità $\mathbb{P}[X \leq \xi]$ (come frazione dei sorteggi in cui tale evento si è verificato). Nell'esercitazione **E 8** si propone di calcolare l'integrale Gaussiano $I(\xi) = \int_0^{\xi} \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt$, $\xi \geq 0$. Un confronto può essere fatto nel caso $\sigma^2 = 1$, sulla base dell'identità $\mathbb{P}[X \leq \xi] = \frac{1}{2} + I(\xi)$.

E 12. Per la simulazione delle X_i si rimanda alle esperienze precedenti. Una volta generate le $\{X_i\}_{i=1}^{M}$ con un M fissato (e.g. $M = 10^3$) si calcola la variabile

$$Y^{(M)} = \frac{1}{\sigma\sqrt{M}} \sum_{i=1}^{M} (X_i - \mu)$$

Per evere una stima delle probabilità in (5) si realizzano N (e.g. $N = 10^3$) copie indipendenti $Y_1^{(M)}, \ldots, Y_N^{(M)}$ della variabile $Y^{(M)}$ e si calcola la frequenza empirica dell'evento $\{Y^{(M)} \in [0,\xi]\}$

per i diversi valori di ξ. Combinando legge dei grandi numeri e teorema del limite centrale si ottiene dunque una stima di $I(\xi)$.

E 13. Ricordiamo che se $X_1, X_2, \ldots$ sono realizzazioni indipendenti di v.a. esponenziali di parametro 1, allora la rappresentazione

$$\begin{aligned} X^\lambda &= \inf\{k \geq 0 : \quad X_1 + \ldots X_{k+1} > \lambda\} \\ &= \sum_{\ell=0}^{\infty} 1\!\!1(X_1 + \ldots X_\ell \leq \lambda) \end{aligned} \tag{7}$$

fornisce una v.a. di Poisson di parametro λ:

$$\mathbb{P}(X^\lambda = k) = \mathrm{e}^{-\lambda}\,\frac{\lambda^k}{k!}\,, \qquad k = 0, 1, 2, \ldots$$

Per la simulazione delle esponenziali X_i si utilizzi la rappresentazione $X = -\log Y$, con Y uniforme in $(0,1)$ (si veda **E10**). Il programma puó essere quindi realizzato tramite un ciclo `do ... while` da arrestare non appena si realizza l'evento $X_1 + \ldots X_\ell > \lambda$.

E 13. *Spazio degli stati.* Identifichiamo i tre colori (giallo, rosso, nero) con i numeri $(0,1,2)$. Sia $\widetilde{\Omega} = \{0,1,2\}^4$. Lo spazio delle configurazioni (o stati) allora è

$$\Omega = \{\sigma \in \widetilde{\Omega} : \ \sigma(A) \neq \sigma(B)\,,\ \sigma(B) \neq \sigma(C)\,,\ \sigma(C) \neq \sigma(D)\,,\ \sigma(D) \neq \sigma(A)\}$$

Un semplice conto mostra che $|\Omega| = 18$. Infatti, fissati i colori di A e B rimangono solo 3 modi diversi di colorare C e D. Poiché i colori di A e B possono essere scelti in 6 modi diversi si ottengono 18 configurazioni. La misura uniforme su Ω è data da $\nu(\sigma) = 1/18$ per ogni $\sigma \in \Omega$.

Probabilità di transizione. Il meccanismo della transizione da una configurazione $\sigma \in \Omega$ a una nuova $\eta \in \Omega$ può essere definito come segue. Scegliamo a caso uno dei 4 vertici. Supponiamo di aver scelto C. Se i due vertici adiacenti B e D hanno lo stesso colore, cioè $\sigma(B) = \sigma(D)$ allora ci sono due possibilità per colorare il vertice scelto C. Tiriamo una moneta (equa) e scegliamo una delle due. In questo caso $\eta(C) = \sigma(C)$ con prob. $1/2$ e $\eta(C) = 3 - \sigma(C) - \sigma(B)$ con prob. $1/2$[1]. Se invece $\sigma(B) \neq \sigma(D)$ non resta che una possibilità per colorare C, quindi $\eta(C) = \sigma(C)$ e la configurazione non cambia. Osserviamo che con questa scelta la nuova configurazione η è o uguale a σ oppure differisce da σ in un solo vertice.

Volendo essere più precisi possiamo scrivere la probabilità di transizione nel modo seguente. Chiamiamo $\Gamma_A(\sigma)$ l'insieme di configurazioni che differiscono da σ solo per il colore nel vertice A:

$$\Gamma_A(\sigma) = \{\eta \in \Omega : \ \eta(A) \neq \sigma(A)\,,\ \eta(B) = \sigma(B)\,,\ \eta(C) = \sigma(C)\,,\ \eta(D) = \sigma(D)\}$$

e analogo per $\Gamma_B(\sigma)$ ecc. Siano σ, η due configurazioni tali che $\eta \in \Gamma_C(\sigma)$. Allora poniamo

$$p(\sigma,\eta) = \frac{1}{4}\begin{cases} 0 & \sigma(B) \neq \sigma(D) \\ \frac{1}{2} & \sigma(B) = \sigma(D)\,,\ \eta(C) = \sigma(C) \\ \frac{1}{2} & \sigma(B) = \sigma(D)\,,\ \eta(C) = 3 - \sigma(C) - \sigma(B) \end{cases} \qquad \big(\eta \in \Gamma_C(\sigma)\big)$$

Analogamente definiamo $p(\sigma,\eta)$ per $\eta \in \Gamma_A(\sigma)$, $\Gamma_B(\sigma)$, $\Gamma_D(\sigma)$. Se η e σ hanno colori diversi in più di un vertice allora poniamo $p(\sigma,\eta) = 0$. Infine se $\eta = \sigma$ poniamo

$$p(\sigma,\sigma) = 1 - \sum_{\eta \in \Omega:\ \eta \neq \sigma} p(\sigma,\eta)$$

[1] Stiamo usando il fatto che dati due numeri differenti x, y in $\{0,1,2\}$ il terzo numero z in $\{0,1,2\}$ differente da entrambi è dato da $3 - x - y$.

Si può verificare per esempio che $p(\sigma,\sigma) = \frac{1}{2}$ se $\sigma \in \Omega_0$.

Misura invariante. Verifichiamo che la misura uniforme ν è invariante per la catena, cioè che per ogni $\sigma \in \Omega$ si ha

$$\nu(\sigma) = \sum_{\eta\in\Omega} \nu(\eta)p(\eta,\sigma)$$

Poiché ν è uniforme basterà quindi mostrare che $\sum_{\eta\in\Omega} p(\eta,\sigma) = 1$. Notiamo che per costruzione si ha

$$\sum_{\eta\in\Omega} p(\sigma,\eta) = 1$$

Quindi l'identità cercata segue una volta che mostriamo che $p(\cdot,\cdot)$ è simmetrica: $p(\sigma,\eta) = p(\eta,\sigma)$, $\sigma,\eta \in \Omega$. Ma questo segue immediatamente dalla definizione data: se per es. $\eta \in \Gamma_A(\sigma)$ allora $\sigma \in \Gamma_A(\eta)$ e $p(\sigma,\eta) = p(\eta,\sigma) = \frac{1}{8}\,1\!\!1(\sigma(B) = \sigma(D))$.

Ergodicità. Sia $p^n(\sigma,\eta)$ la matrice di transizione su n–passi, definita per ricorrenza dalle equazioni:

$$p^{n+1}(\sigma,\eta) = \sum_{\xi\in\Omega} p(\sigma,\xi)p^n(\xi,\eta)\,, \quad p^1(\sigma,\eta) = p(\sigma,\eta)\,.$$

E' possibile verificare che la catena è *regolare*: esiste un intero m tale che $p^m(\sigma,\eta) > 0$ per ogni $\sigma,\eta \in \Omega$. Da un teorema di Markov segue che ν è l'unica misura invariante, e soddisfa

$$\lim_{n\to\infty} p^n(\sigma,\eta) = \nu(\eta)\,, \qquad \sigma,\eta \in \Omega \tag{8}$$

Notare l'indipendenza del limite dalla condizione iniziale σ. La convergenza (8) può essere verificata come segue: fissato il valore di σ lasciamo evolvere M "storie" indipendenti della catena fino al tempo n (tutte con inizio in σ). Chiamiamo $\eta^1,\dots,\eta^M$ le configurazioni finali. Scegliamo un $\xi \in \Omega$ qualunque, per es. $\xi(A) = 0, \xi(B) = 1, \xi(C) = 0, \xi(D) = 1$. Contando le volte che $\eta^i = \xi$ e dividendo per M si ottiene una stima $\bar{p}_n(\sigma,\xi)$ per $p^n(\sigma,\xi)$. Fissiamo M molto grande (per es. $M = 10^5$) e grafichiamo la frequenza $\bar{p}_n(\sigma,\xi)$ per valori crescenti di n ($n = 1,2,3,\dots$). Si osserverà che $\bar{p}_n(\sigma,\xi)$ approssima bene il valore limite $\nu(\xi) = 1/18$ a partire da un certo n_0. Diremo che n_0 è il tempo necessario per avere "perdita di memoria".

Se $\sigma_1,\sigma_2,\dots$ denotano gli stati successivi di una *singola* "storia" della catena si può inoltre dimostrare che per ogni funzione $\varphi : \Omega \to \mathbb{R}$ si ha la convergenza:

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \varphi(\sigma_k) = \mathbb{E}_\nu(\varphi) = \sum_{\sigma\in\Omega} \nu(\sigma)\varphi(\sigma)$$

Il caso $\varphi(\sigma) = 1\!\!1(\sigma \in \Omega_0)$ fornisce il limite di h_n/n proposto nel testo.

References

[1] W. Feller, *An introduction to probability theory and its applications.* vol. 2, Wiley, 1971

[2] `http://random.mat.sbg.ac.at/links`

[3] `http://www.duke.edu/~hpgavin/gnuplot.html`