Ratio Math. n. 7 - 1994
125 - 133

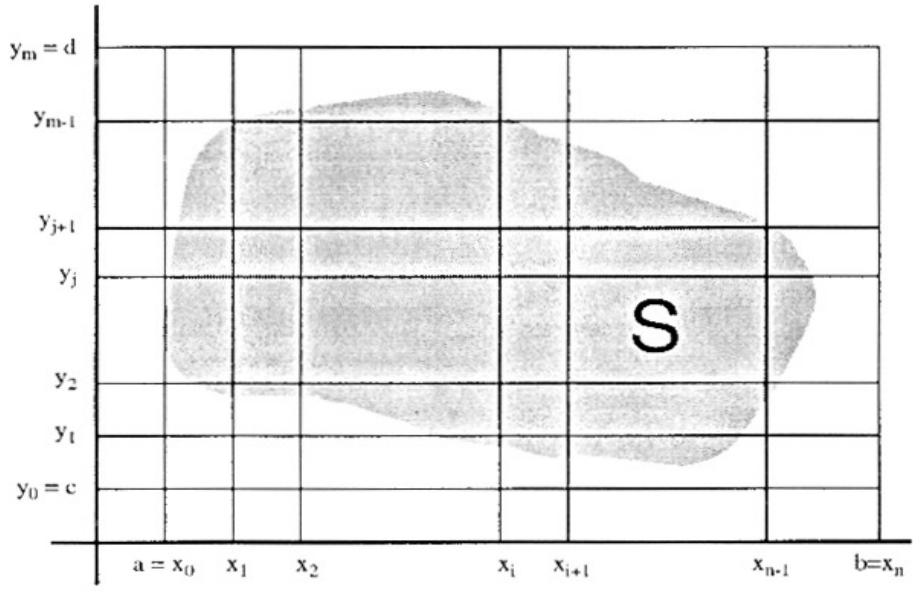
CALCOLO DI AREE E VOLUMI PER MEZZO DI SUCCESSIONI DI NUMERI A CASO*

Giuseppina Varone
Istituto Mecenate, via Marco Polo - Pescara

1. AREA DI UN INSIEME PIANO: IL PROCEDIMENTO DI PEANO-JORDAN

Sia S un insieme limitato del piano. Allora esiste un intervallo $I = [a, b] \times [c, d]$ che lo contiene.
Siano $x_0 = a < x_1 < x_2 .... < x_{n-1} < x_n = b$ punti dell'intervallo [a, b] e siano $y_0 = c < y_1 < y_2 .... < y_{m-1} < y = d$ punti dell' intervallo [c, d].

Tracciando dai punti x_i le parallele all'asse y e dai punti y_i le parallele all'asse x, l'intervallo I viene suddiviso in m· n intervallini $I_{ij} = [x_i, x_{i+1}] \times [y_j, y_{j+1}]$.



---

* Conferenza tenuta al Convegno Interregionale *"Probabilità, Statistica e Crittografia"* - Isernia, 1992

L'insieme degli intervallini I_{ij} si dice partizione dell'intervallo I. Indichiamolo con Π. La somma delle misure degli intervallini I_{ij} contenuti in S è un valore per difetto dell'area di S e si indica con $S_1(\Pi)$. La somma delle misure degli intervallini I_{ij} che hanno intersezione non vuota con S è un valore per eccesso dell'area di S e si indica con $S_2(P)$.

Indichiamo con Σ_1 l'insieme delle somme $S_1(\Pi)$ e con Σ_2 l'insieme delle somme $S_2(\Pi)$ al variare di Π nell'insieme delle partizioni dell'intervallo I.

I due insiemi Σ_1 e Σ_2 sono separati ossia si dimostra che dette Π_1 e Π_2 due partizioni di I risulta sempre $S_1(\Pi_1) \leq S_2(\Pi_2)$.

Allora detti σ_1 e σ_2 rispettivamente l'estremo superiore di Σ_1 e l'estremo inferiore di Σ_2 risulta $\sigma_1 \leq \sigma_2$. I numeri σ_1 e σ_2 si dicono rispettivamente misura interna (m_i) e misura esterna (m_e) dell'insieme S. Se risulta $\sigma_1 = \sigma_2$ allora l'insieme S si dice misurabile secondo Peano-Jordan e il numero $\sigma_1 = \sigma_2$ si dice misura dell'insieme S secondo Peano-Jordan e si indica con m(S).

2. AREA DI UN INSIEME PIANO PER MEZZO DI NUMERI A CASO

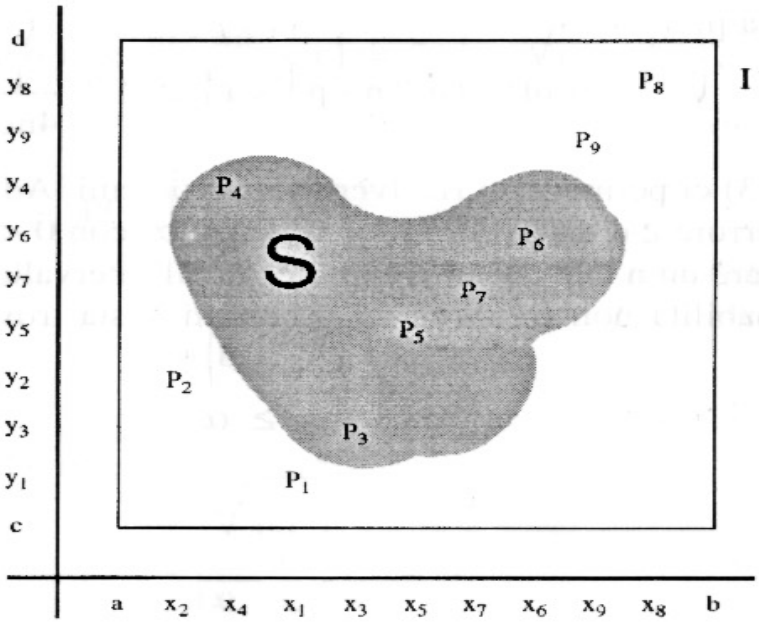
Sia S un insieme limitato del piano misurabile secondo Peano-Jordan e sia I = [a, b] x [c, d] un intervallo contenente S.

Generiamo n numeri a caso nell'intervallo [a, b] ed n numeri a caso nell'intervallo [c, d]. Siano $x_1, x_2, x_3, \ldots, x_n$ i numeri appartenenti all'intervallo [a, b] e $y_1, y_2, y_3, \ldots, y_n$ i numeri appartenenti all'intervallo [c, d].

I punti $P_i = (x_i, y_i)$ possono appartenere o no all'insieme S.

Sia m il numero di tali punti appartenenti ad S.

Per n abbastanza grande il rapporto tra l'area di S e l'area di I è approssimativamente dato dal numero m/n e generalmente l'approssimazione migliora al crescere di n.



Indichiamo con p la probabilità che un punto preso a caso nell'intervallo I cada nell'insieme S. Allora

$$p = \frac{\text{Area S}}{\text{Area I}}$$

Per la legge di Bernoulli il numero m di volte che, in n prove, un punto p preso a caso in I cade in S è una variabile casuale con media np e varianza npq, con q = (1 - p).

La frequenza m/n è una variabile casuale con media p e varianza pq/n.

Dal Teorema di Bienaymé-Tchebyceff si ha che se X è una variabile casuale con media μ e varianza σ^2 allora

$$\text{(2.1)} \qquad \text{prob} \left\{ |X - \mu| < \varepsilon \right\} \geq 1 - \frac{\sigma^2}{\varepsilon^2}$$

Segue che

$$\text{(2.2)} \qquad \text{prob} \left\{ |m/n - p| < \varepsilon \right\} \geq 1 - \frac{pq}{n\varepsilon^2} .$$

La funzione y = pq = p(1 - p) ha il massimo per p = 1/2 ed in tale punto assume il valore 1/4. Di conseguenza risulta che, qua-

lunque sia p,

$$\text{(2.3)} \qquad \text{prob}\left\{ |m/n - p| < \varepsilon \right\} \geq 1 - \frac{1}{4n\varepsilon^2} .$$

La (2.3) ci permette di risolvere vari problemi. Ad esempio, fissato un errore ε e un livello di confidenza α, con $0 < \alpha < 1$ si può determinare un n tale che m/n appartenga all'intervallo $(p - \varepsilon, p + \varepsilon)$ con probabilità non inferiore ad α. Infatti basta trovare un n tale che:

$$1 - \frac{1}{4n\varepsilon^2} \geq \alpha$$

ossia

$$\text{(2.4)} \qquad n \geq \frac{1}{4\varepsilon^2 (1 - \alpha)} .$$

Ad esempio, per $\alpha = 90\%$ ed $\varepsilon = 0{,}1$ si ottiene $n \geq 250$.

Le valutazioni ottenute in base alle (2.3) e (2.4) sono in generale pessimistiche, nel senso che si ottiene un errore minore di ε al livello di confidenza α per valori di n in genere molto minori al secondo membro della (2.4).

Una valutazione più precisa si ottiene tenendo conto che m è una variabile binomiale con media $n \cdot p$ e varianza $n \cdot p \cdot q$, per cui per n abbastanza grande la variabile casuale ha distribuzione normale centrata ridotta.

Indicando con $\phi(x)$ la funzione di ripartizione della variabile casuale normale centrata ridotta risulta allora per ogni $k > 0$

$$\text{prob}\left(-k < \frac{m - np}{\sqrt{n\,p\,q}} < k \right) = \phi(k) - \phi(-k) = 2\phi(k) - 1 .$$

Segue

$$\text{prob}\left(-k < \frac{m - np}{n} \cdot \frac{n}{\sqrt{npq}} < k \right) = 2\phi(k) - 1$$

Posto $k \cdot \sqrt{pq/n} = \varepsilon$ si ha

$$\text{prob}\left(-\varepsilon < \frac{m - np}{n} < \varepsilon\right) = 2\phi\left(\varepsilon \cdot \sqrt{n/pq}\right) - 1.$$

Ponendo

$$\text{prob}\left(-\varepsilon < \frac{m - np}{n} < \varepsilon\right) \geq \alpha,$$

Si ottiene

$$2\phi\left(\varepsilon\sqrt{\frac{n}{pq}}\right) - 1 \geq \alpha$$

da cui

$$\phi\left(\varepsilon\sqrt{\frac{n}{pq}}\right) \geq \frac{1+\alpha}{2},$$

$$\varepsilon\sqrt{\frac{n}{pq}} \geq \phi^{-1}\left(\frac{1+\alpha}{2}\right)$$

(2.5) $$n \geq \frac{pq}{\varepsilon^2}\left[\phi^{-1}\left(\frac{1+\alpha}{2}\right)\right]^2.$$

Essendo $pq \leq 1/4$ la (2.5) è verificata per ogni valore di p per

(2.6) $$n \geq \frac{1}{4\varepsilon^2}\left[\phi^{-1}\left(\frac{1+\alpha}{2}\right)\right]^2.$$

Ad esempio per $\alpha = 0{,}90$ ed $\varepsilon = 0{,}1$ si ottiene

$$\phi^{-1}\left(\frac{1+\alpha}{2}\right) = \phi^{-1}(0{,}95) = 1{,}65$$

e quindi dalla (2.6) si ottiene:

$$n \geq 25 \cdot [1{,}65]^2 = 68.$$

3. METODI PER GENERARE NUMERI A CASO

Uno dei metodi più utilizzati per generare numeri a caso in un intervallo [a, b] è il seguente:

(1) si fissa un intero positivo m "abbastanza grande" (nelle applicazioni è opportuno che sia almeno dell'ordine di 10^8) e si fissano due numeri interi opportuni h e k appartenenti all'insieme $\{0, 1, 2, \ldots, m - 1\}$

(2) si considera la formula

$$x_{r+1} = (hx_r + k) \bmod m, \qquad (3.1)$$

che, assegnato x_r appartenente all'insieme $\{0, 1, 2, \ldots, m - 1\}$, permette di calcolare x_{r+1} come resto della divisione di $hx_r + k$ per m.

(3) a partire da un numero assegnato x_0, detto seme, si calcolano n numeri $x_1, x_2, \ldots, x_{n-1}, x_n$.

(4) si divide ogni x_i per m ottenendo dei numeri $y_i = x_i / m$, che vengono considerati come numeri a caso nell'intervallo [0, 1];

(5) dalla formula

$$z_i = a + y_i (b - a) \qquad (3.2)$$

si ottengono numeri a caso appartenti all'intervallo [a, b].

Le proprietà matematiche e statistiche delle successioni ottenute a partire dal generatore (3.1) variano notevolmente al variare di h, k ed x_0.

Per vari motivi, le scelte più opportune di m sono $m = 10^r$, $m = 2^r$ ed m numero primo.

Si dimostra che le successioni ottenute sono periodiche con periodo δ non superiore ad m ed il massimo periodo possibile $\delta = m$ si ottiene se e solo se sono soddisfatte le seguenti condizioni

(a) M.C.D. (h, m) = 1
(b) p primo divisore di m $\Rightarrow$ h mod p = 1
(c) 4 divisore di m $\Rightarrow$ h mod 4 = 1
(d) M.C.D. (k, m) = 1.

Spesso è conveniente porre k = 0 ottenendo il generatore congruenziale moltiplicativo

(3.3) $$x_{r+1} = (hx_r) \bmod m.$$

In questo modo si ha il periodo massimo per m numero primo e sotto le condizioni

(a) h ed x_0 sono entrambi primi con m,

(b) h è una radice primitiva mod m, ossia $h^r \bmod m \neq 1$

per ogni $r \in \{1, 2, ..., m - 2\}$.

ESEMPIO

Consideriamo la (3.3) con m numero primo, ad esempio m = 41. Per il Teorema di Legendre (cfr.[2] pag. 72) si ha

$$\begin{pmatrix}\text{h radice primitiva}\\ \text{mod } 41\end{pmatrix} \Leftrightarrow \begin{pmatrix}h^{20} \not\equiv 1 \bmod 41\\ h^{8} \not\equiv 1 \bmod 41\end{pmatrix}.$$

I numeri h tali che h è primo con m e $h^{20} \not\equiv 1 \bmod 41$ sono i numeri che non sono quadrati mod 41 ossia tali che non esiste nessun intero positivo tale che $b^2 = h \bmod 41$.

Poiché, rispetto al modulo 41,
$1^2 = 1$, $2^2 = 4$, $3^2 = 9$, $4^2 = 16$, $5^2 = 25$, $6^2 = 36$, $7^2 = 8$, $8^2 = 23$, $9^2 = 40$, $10^2 = 18$, $11^2 = 39$, $12^2 = 21$, $13^2 = 5$, $14^2 = 32$, $15^2 = 20$, $16^2 = 10$, $17^2 = 2$, $18^2 = 37$, $19^2 = 33$, $20^2 = 31$ e, per ogni intero positivo s tale che $20 < s \leq 40$ risulta $s^2 = (41\text{-}s)^2$, segue che $h^{20} \not\equiv 1 \bmod 41$ per h uguale ad uno dei seguenti numeri:
3 6 7 11 12 13 14 15 17 19 22 24 26 27 28 29 30 34 35 38.

Di questi numeri soddisfano l'uguaglianza $h^8 = 1 \pmod{41}$ 3, 14, 27, 38.

I rimanenti 16 numeri sono radici primitive modulo 41 in accordo con la teoria che dice che il numero di radici primitive modulo 41 è $\phi(40) = 40 \cdot 1/2 \cdot 4/5 = 16$, dove con $\phi(m)$ si denota l'indicatore di Gauss.

Consideriamo, per fissare le idee, il caso h = 7.

Allora si ha il generatore congruenziale di massimo periodo, 40,

(3.4) $$x_{r+1} = 7x_r \pmod{41}.$$

Consideriamo ad esempio il problema di calcolare l'area del triangolo di vertici (0, 0), (1, 1), (0, 2) che è un sottoinsieme dell'intervallo [0, 2] x [0,1].

Ponendo $x_0 = 1$, si ottengono, successivamente, i seguenti numeri

1, 7, 8, 15, 23, 38, 20, 17, 37, 13, 9, 22, 31.

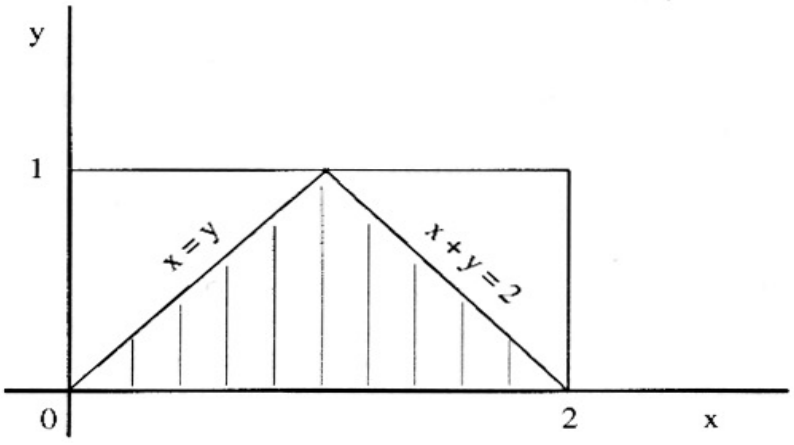
Dividendo per 41, si ottiene

0,024; 0,171; 0,195; 0,366; 0,561; 0,927; 0,458; 0,414; 0,902; 0,317; 0,220; 0,537; 0,756

Considerando, nell'intervallo I i punti

$$P_i = \left(2\,\frac{x_i}{41}\,,\,\frac{x_{i+1}}{41}\right) \text{ si ottengono}$$

$P_0 = (0{,}048;\ 0{,}171)$, $P_1 = (0{,}342;\ 0{,}195)$, $P_2 = (0{,}390;\ 0{,}366)$,
$P_3 = (0{,}732;\ 0{,}561)$, $P_4 = (1{,}122;\ 0{,}927)$, $P_5 = (1{,}854;\ 0{,}488)$,
$P_6 = (0{,}976;\ 0{,}414)$, $P_7 = (0{,}828;\ 0{,}902)$, $P_8 = (1{,}804;\ 0{,}317)$,
$P_9 = (0{,}634;\ 0{,}220)$, $P_{10} = (0{,}440;\ 0{,}537)$, $P_{11} = (1{,}074;\ 0{,}756)$.



Si vede che sono interni al triangolo i punti P_1, P_2, P_3, P_6, P_9, P_{11}, mentre sono esterni al triangolo i punti P_0, P_4, P_7, P_8, P_{10}.

Il rapporto $\dfrac{m}{n} = \dfrac{\text{punti interni al triangolo}}{\text{numero totale di punti}}$ è esattamente uguale

ad 1/2, rapporto fra l'area del triangolo e l'area del rettangolo.

BIBLIOGRAFIA

[1] G. DALL'AGLIO, *Calcolo delle probabilità* (1987), Zanichelli, Bologna.

[2] G. DI BIASE e A. MATURO, *Sulla verifica di casualità di successioni numeriche in [0, 1] da un punto di vista bayesiano: esempi e considerazioni critiche* (1993), Atti del Convegno Nazionale "Matematica Moderna e Insegnamento: se ne può riparlare?", Ed. Luciani, Roma, pp. 249-256

[3] A. MATURO, *Numeri Pseudocasuali* (1989), Libreria dell'Università, Pescara.

[4] R. SCOZZAFAVA, *Introduzione alla Probabilità e alla Statistica* (1984), Veschi Roma.

[5] G. VARONE, *Generazione di variabili pseudo-probabilistiche per le tecniche di simulazione*, Tesi di Laurea. (1991).