187. Casualità Matematica e Metodo Monte Carlo

Nicola De Nitti
nicoladenitti@gmail.com

Premessa

Il concetto di probabilità, impiegato a partire dal XVII secolo, è diventato con il passare del tempo fondamentale per diverse discipline. I primi studi che portarono a sviluppare concetti legati alla probabilità si ritrovano nel *Liber de ludo aleæ* di Girolamo Cardano (scritto nel 1526, ma pubblicato solo nel 1663) e in *Sulla scoperta dei dadi* di Galileo Galilei (pubblicato nel 1656), ma la nascita del concetto moderno di probabilità viene attribuita a Blaise Pascal (1623-1662) e Pierre de Fermat (1601-1665). Nel 1657 Christiaan Huygens (1629-1695) scrisse un *Libellus de ratiociniis in ludo aleæ*, il primo vero e proprio trattato sul calcolo delle probabilità. Nel 1713 venne pubblicato postumo *Ars conjectandi* di Jakob Bernoulli, in cui veniva dimostrato il teorema che porta il suo nome, noto anche come "legge dei grandi numeri". Successivamente, de Moivre pervenne ad una prima formulazione, in seguito generalizzata da Pierre Simon Laplace (1749-1827), del fondamentale "Teorema del limite centrale". La teoria della probabilità raggiunse così basi matematicamente solide e, dunque, il rango di nuova disciplina.

Il calcolo delle probabilità è uno strumento essenziale per la statistica, in quanto dà una risposta al problema inverso di quello della statistica inferenziale. Infatti, mentre essa cerca di determinare, tramite la conoscenza dei risultati di un esperimento (o più esperimenti), quali siano le caratteristiche della popolazione su cui tale sperimentazione è stata eseguita, nel calcolo delle probabilità, si assume che tutte le caratteristiche della popolazione siano note e si cerca di calcolare a priori la 'probabilità' che un esperimento abbia un determinato risultato.

Metodo Monte Carlo e casualità matematica

Negli ultimi anni ha assunto crescente importanza il "metodo Monte Carlo", un metodo numerico basato su procedimenti probabilistici, usato in statistica per la risoluzione di problemi di varia natura che presentano difficoltà analitiche difficilmente (o non in altro modo) superabili. Tale metodo è stato citato in un articolo di J. Dongarra e F. Sullivan, pubblicato sulla prestigiosa rivista "Computing in Science and Engineering", tra i dieci algoritmi con "la più grande influenza sullo sviluppo e la pratica della scienza e dell'ingegneria del XX secolo". Il metodo Monte Carlo fu formalizzato negli anni '40 del Novecento da John von Neumann e Stanisław Marcin Ulam, che partecipavano al Progetto Manhattan per lo studio della dinamica delle esplosioni nucleari. A quanto pare, il nome "Monte Carlo" fu coniato da Nicholas Constantine Metropolis in riferimento alla capitale del Principato di Monaco, Montecarlo, dove ha sede il celebre casinò, luogo dell'aleatorietà per antonomasia. Infatti alla base dell'algoritmo ci sono proprio ripetuti campionamenti casuali.

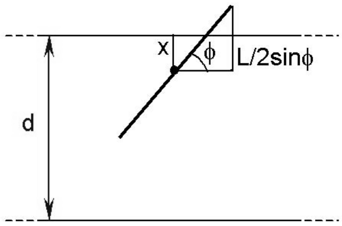
Non è tuttavia una operazione semplice dare una definizione univoca di casualità matematica. Una tra le più note proposte di caratterizzazione formale di questo concetto (cioè quali criteri debba soddisfare una sequenza di numeri per essere casuale) è il criterio di Richard von Mises: una sequenza di numeri è casuale quando sono completamente assenti regole che possano essere applicate con successo per migliorare le previsioni circa il numero successivo della serie. Tale principio è noto come "principio dell'impossibilità di un sistema di gioco" (o "assioma del disordine"). Il criterio di von Mises presenta, tuttavia, una evidente ambiguità quando è applicato a sequenze infinite, in quanto viene a mancare la possibilità di qualsiasi controllo effettivo della casualità della sequenza stessa.

Negli anni '30, Karl Popper propose un altro tipo di sequenza casuale: una sequenza finita costruita con una regola matematica. Le idee di Popper si ritrovano nei sistemi fondati su algoritmi per la costruzione delle sequenze di numeri casuali che vengono comunemente adoperate. Tuttavia, è evidente che se si conosce la legge con cui comporre una sequenza, questa non è più definibile, a rigore, casuale: ogni numero è infatti predicibile con probabilità pari al 100%.

Concetti chiave del metodo Monte Carlo

Con il termine di "metodo Monte Carlo" vengono in generale denominate tutte quelle tecniche che adoperano variabili aleatorie artificiali (ovvero generate con un calcolatore) per la risoluzione di problemi (quali il calcolo di quantità o la simulazione di fenomeni).
Spesso i ricercatori si trovano a fronteggiare situazioni in cui hanno bisogno di conoscere la probabilità di un determinato evento condizionato da un numero elevato di variabili che rendono molto difficoltosi i calcoli analitici. In tali situazioni, generalmente si adottano metodi di campionamento simulato (cioè si simula la situazione nella quale si vuole calcolare la probabilità di un certo evento). La simulazione stocastica si attua riproducendo il meccanismo preso in esame; sostituendo la valutazione analitica con l'osservazione empirica del fenomeno; e traendo da quest'ultima le informazioni non rilevabili per via analitica. Tuttavia, è accertato che questo non è il metodo più efficace per trovare la soluzione di un problema, in quanto la procedura del campionamento simulato porta ad un risultato sempre influenzato dall'errore statistico.
L'applicazione di questo metodo non è ristretta solamente ai problemi di natura statistica, ma include tutti i casi in cui è possibile trovare un collegamento tra il problema in esame ed il comportamento di un certo sistema aleatorio.

L'ago di Buffon



Le basi teoriche del metodo Monte Carlo risalgono al 1777, quando il matematico francese George Louis Leclerc, conte di Buffon (1707-1788), propose e risolse un curioso problema, che, per il profondo interesse che suscita, per la sua rilevanza storica, e per la semplicità dell'approccio con cui può essere affrontato, costituirà il principale oggetto d'analisi di questo articolo, insieme ad altri due esempi significativi dei concetti alla base del metodo Monte Carlo.
Questo problema, comunemente noto come "il problema dell'ago di Buffon" fa intervenire π in un contesto assolutamente originale. Buffon considerò un'area piana sulla quale erano state tracciate linee rette parallele a distanza d l'una dall'altra ed immaginò, poi, di lanciare 'a caso' sulla superficie un ago di lunghezza $L < d$. Si presentavano due eventualità: 1) l'ago incontrava una delle linee; 2) cadeva tra una linea e l'altra.
La condizione affinché l'ago incontri una retta è

$$x < \frac{L}{2}\sin\varphi ,$$

per cui per risolvere in problema è necessario calcolare la probabilità:

$$P\left(x < \frac{L}{2}\sin\varphi\right)$$

A tal fine, si può tentare un approccio geometrico, trasformando la definizione classica di probabilità in un rapporto tra aree convenientemente individuate.
I limiti geometrici del problema sono:

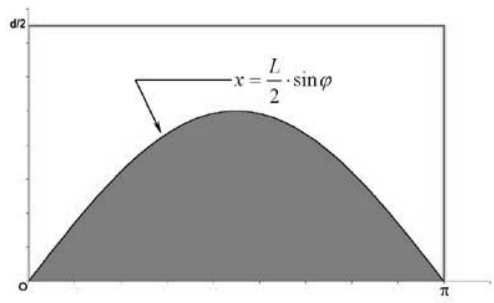
$$0 < x < \frac{d}{2} \text{ e } 0 < \varphi < \pi$$

Utilizzando un grafico (f; *x*), queste condizioni consentono di individuare un rettangolo la cui area costituisce la somma di tutte le possibili combinazioni tra l'orientazione dell'ago e la posizione del suo centro sul piano. Fra queste, le combinazioni favorevoli sono quelle individuate dai punti le cui coordinate soddisfano la condizione sopra determinata, ovvero da tutti i punti al di sotto della curva

$$x = \frac{L}{2}\sin\varphi$$

Da queste considerazioni segue:

$$P\left(x < \frac{L}{2}\sin\varphi\right) = \frac{\text{casi favorevoli}}{\text{casi possibili}} = \frac{\frac{L}{2}\cdot\int_0^{\pi}\sin\varphi\cdot d\varphi}{\pi\cdot\frac{d}{2}}$$



ovvero: $P = \frac{2L}{\pi d}$.

Dunque, è possibile calcolare π utilizzando la formula:

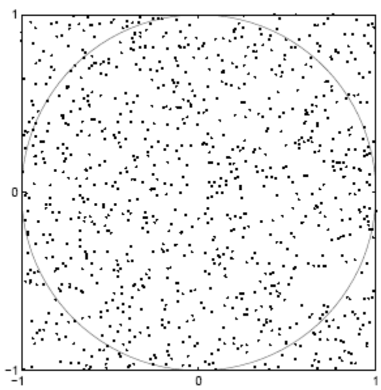
$$\pi = \frac{2L}{Pd}.$$

Dal punto di vista di un elaboratore elettronico, il procedimento seguito nella risoluzione del problema può essere così schematizzato:

a. generare una prima sequenza di numeri casuali (x) per simulare la posizione dell'ago;
b. generare una seconda sequenza di numeri casuali (φ) per simulare l'orientazione dell'ago;
c. controllare che ogni coppia di numeri casuali x e φ verifichi o meno la disuguaglianza che caratterizza il problema;
d. contare quante volte la condizione è soddisfatta (tenendo conto del numero di prove effettuate);
e. calcolare π.

Il problema e la soluzione proposta da Buffon furono, tuttavia, persi di vista finché Pierre Simon Laplace (1749-1827) ne propose una generalizzazione, così gettando le basi del metodo Monte Carlo.

Metodo Monte Carlo per la determinazione di π

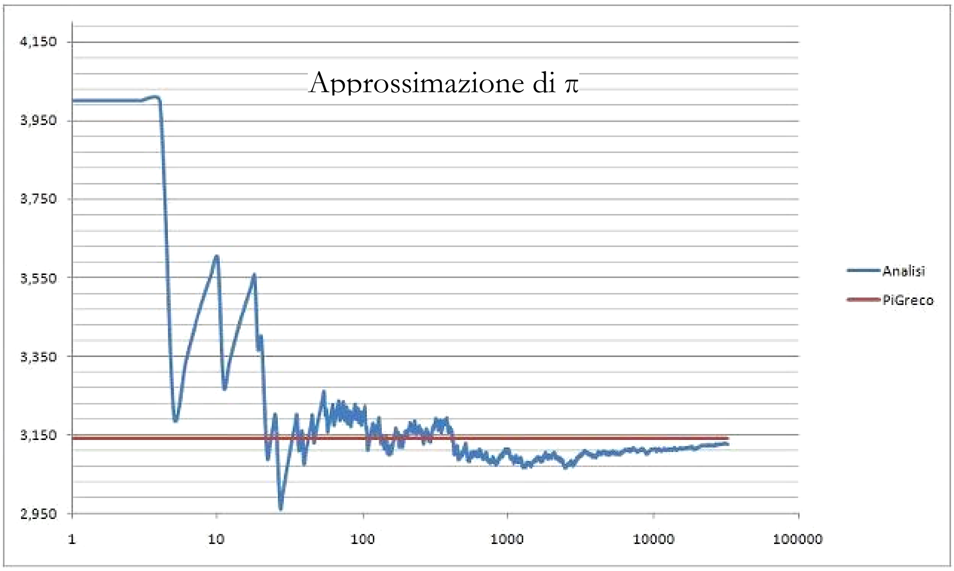


Esiste un secondo modo per determinare le cifre decimali di π adoperando il metodo Monte Carlo (cambia unicamente l'evento casuale preso in considerazione).
Si suppone di lanciare N freccette ad un bersaglio costituito da un quadrato di lato $2L$ circoscritto ad una circonferenza. Si assume che le freccette vengano lanciate casualmente all'interno del quadrato e che, quindi, colpiscano il quadrato in ogni posizione con uguale probabilità. Per analizzare questa situazione, si può stabilire un sistema cartesiano xOy avente per origine il centro della circonferenza, in cui la posizione di ogni freccetta sarà indicata da un punto individuato da una coppia 'casuale' di coordinate: $Z_i(x; y)$.
Dopo un numero consistente di lanci, la frazione di freccette che ha colpito la circonferenza (ovvero il numero di punti che hanno distanza dall'origine minore del raggio della circonferenza) sarà circa uguale al rapporto tra l'area della circonferenza e quella del quadrato.
Pertanto:

$$\frac{\pi L^2}{4L^2} = \frac{1}{4}\pi \cong \frac{N_c}{N}$$

dove N è il numero totale di freccette lanciate e N_c indica il numero di freccette cadute all'interno della circonferenza.

Si potrà dunque usare il valore $\frac{4N_c}{N}$ come approssimazione di π.



Si può notare che la stima di π migliora all'aumentare di N. Tuttavia, la convergenza è tutt'altro che veloce ed uniforme: la presenza di fluttuazioni dovute all'approccio probabilistico è, infatti, una caratteristica del metodo Monte Carlo.

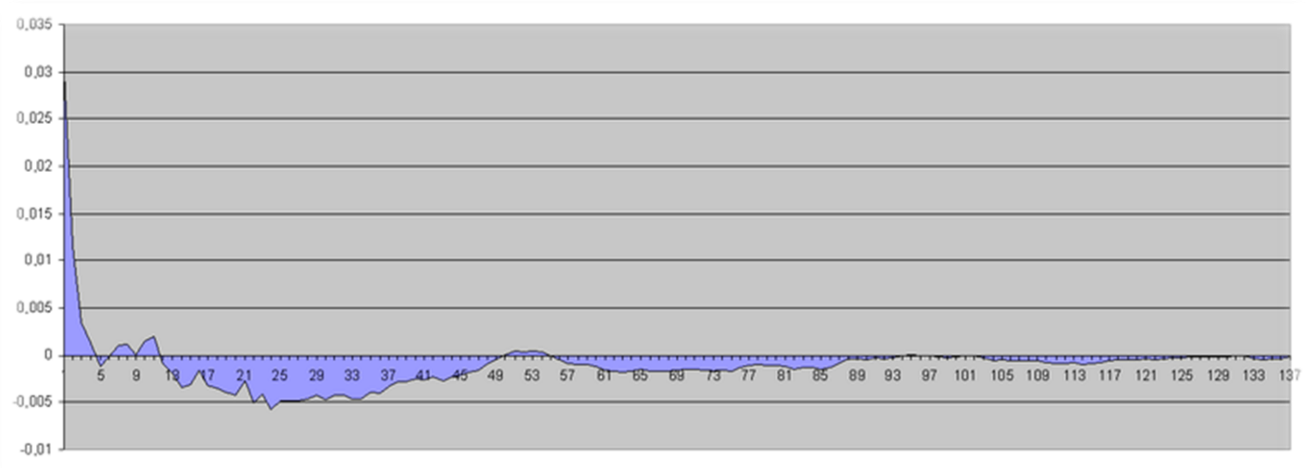


Immagine tratta da http://it.wikipedia.org/wiki/File:Errore_montecarlo_calcolo_pi_greco.PNG
Andamento percentuale dell'errore tra il π teorico e il π calcolato. Il programma ha eseguito 1370 milioni di lanci. Si noti che, all'inizio, l'errore è molto elevato, ma rapidamente tende a decrescere.

Determinazione della superficie di una figura piana

In molti casi pratici, il calcolo dell'area di una regione di piano richiede troppe variabili per darne una stima in modo rapido (è il caso del calcolo della superficie di una macchia). Anche in questo caso è possibile adoperare il metodo Monte Carlo per trovare almeno un valore approssimativo della misura cercata. Come si evince dagli esempi precedenti, la probabilità P che un punto appartenga ad una superficie piana S_1 racchiusa all'interno di un'altra superficie piana S_2 è data dal rapporto fra le aree delle due superfici.
Viceversa, nota la probabilità P, possiamo affermare che l'area di S_1 è uguale al prodotto di P per l'area di S_2.

$$S_1 = P \cdot S_2$$

Per un gran numero di prove, si può sostituire a P la frequenza relativa f_n:

$$S_1 = f_n \cdot S_2$$

Considerazioni

Sono numerosi ormai i campi in cui si utilizzano metodi statistici per ottenere informazioni e stime su fenomeni legati al caso. In questi contesti non occorre nemmeno che i dati siano raccolti durante un esperimento reale: infatti, possono anche provenire da simulazioni effettuate per mezzo di un computer attraverso l'uso di sequenze di numeri casuali. Esse sono utilizzate per simulare per migliaia di volte il fenomeno aleatorio, e per raccogliere rapidamente serie di dati che, trattati con metodi statistici, forniscono stime che diventano tanto più attendibili quanto più è grande il numero delle prove eseguite. Oggi il metodo Monte Carlo trova applicazione in vari ambiti scientifici, fra cui l'analisi di problemi fisici che richiederebbero esperimenti difficilmente realizzabili. In questi casi il metodo Monte Carlo costituisce uno strumento vitale per la verifica di nuove teorie.

Bibliografia

BECKMANN, P., *A History of Pi*, St. Martin's Press, New York, 1971.
DEL PICCOLO, A., *Appunti per una lezione.*
http://areeweb.polito.it/didattica/polymath/htmlS/argoment/APPUNTI/TESTI/Apr_03/APPUNTI.HTM
DONGARRA, J. e SULLIVAN, F., *Guest editors' introduction: the top 10 algorithms*, in "Computing in Science and Engineering", 2(1):22–23, January/February 2000.
GOULD, H. e TOBOCHNIK, J., *An Introduction to Computer Simulation Methods*, Addison-Wesley, New York, 1988.
ROBERT , C. P. e CASELLA, G., *Monte Carlo Statistical Methods*, Springer-Verlag, New York, 2004.