Numeri, casualità e simulazioni

Senza addentrarci troppo in discussioni filosofiche sul significato di casualità, accenneremo ad alcune tra le molteplici applicazioni nelle quali intervengono i cosiddetti numeri casuali ed a come la matematica consenta di affrontare problemi legati al concetto di caso.

Numeri casuali sono utilizzati per costruire simulazioni di natura probabilistica di fenomeni fisici (reattori nucleari, traffico stradale, aereodinamica), chimico-biologici (simulazione di reazioni metaboliche) di problemi decisionali e finanziari (econometrica, previsione Dow-Jones), informatica (progettazione VLSI, rendering) o come semplice fonte di divertimento (videogiochi).

Il forte legame che esiste tra il gioco e le simulazioni probabilistiche è sottolineato dal fatto che a tali simulazioni viene generalmente dato il nome di metodi Monte Carlo (in onore del famoso casinò a Monaco).

Note storiche

L'idea di utilizzare in modo sistematico simulazioni di tipo probabilistico per risolvere un problema di natura fisica viene generalmente attribuita al matematico polacco Stanislaw Ulam (1909-1984).

Ulam fu uno dei personaggi chiave nel progetto americano per la costruzione della bomba atomica (Manhattan project) durante la seconda guerra mondiale tra il 1943 ed il 1945 a Los Alamos, New Mexico (dopo la guerra, Ulam diede contributi essenziali anche nello sviluppo della bomba a fusione di idrogeno o bomba H).

Il progetto Manhattan richiedeva infatti la risoluzione di un enorme numero di problemi incredibilmente complessi (nella sua autobiografia Ulam descrive come l'idea di utilizzare simulazioni casuali per risolveretali problemi gli sia venuta mentre giocava a carte)

Generatori di numeri casuali

Innanzitutto cosa è un numero casuale?

Le proprietà statistiche che una sequenza di numeri casuali deve possedere sono uniformità ed indipendenza. Si supponga di dividere l'intervallo [0,1] in *n* sottointervalli di uguale ampiezza.

Conseguenza della proprietà di uniformità:

- Se si eseguono *N* osservazioni di un numero casuale, il numero atteso in ogni sottointervallo è pari a *N/n*.

Conseguenza della proprietà di indipendenza:

- La probabilità di ottenere un valore in un particolare intervallo è indipendente dai valori precedentemente ottenuti

Generatori di numeri casuali

Innanzitutto cosa è un numero casuale?

Un esempio che tutti conosciamo consiste nel lancio di un dado, in effetti l'imprevedibilità del numero ottenuto come punteggio, compreso tra 1 e 6, conferisce allo stesso una forma di casualità.

L'idea stessa di utilizzare un calcolatore (quindi un'oggetto puramente deterministico e di conseguenza prevedibile), per generare un numero casuale quindi imprevedibile sembra costituire una sfida impossibile

In effetti nessun calcolatore è in grado di generare numeri puramente casuali, ma solo *numeri pseudo-casuali* o quasi-casuali ossia numeri generati da algoritmi numerici deterministici in grado di superare una serie di test statistici che conferiscono a tali numeri una apparente casualità.

Il metodo middle-square

Nei primi anni dell'era dei computer (1946) John von Neumann suggerì il famoso metodo middle-square per generare numeri pseudo casuali distribuiti in modo uniforme.

In tale distribuzione uniforme ogni possibile numero in un determinato intervallo è ugualmente probabile. Ad esempio se lanciamo un dado un certo numero di volte ognuna delle facce da 1 a 6 si presenterà circa 1/6 delle volte originando così una successione uniforme di numeri casuali compresi tra 1 e 6.

Oggi il metodo middle-square ha un'importanza solamente storica, ma nella sua semplicità evidenzia un aspetto importante nella generazione di numeri pseudo-casuali al computer, ossia la necessità di avere a disposizione molti numeri casuali ed in modo rapido.

Il metodo middle-square

Supponiamo di volere generare un numero casuale di 4 cifre, ossia un numero tra 0000 e 9999. Il metodo middle-square richiede come tutti i generatori di numeri casuali un valore iniziale, detto seme dal quale vengono generati i successivi valori.

Ad esempio a partire da 1234, elevando tale numero al quadrato abbiamo le otto cifre 01522756 delle quali teniamo solamente le quattro cifre di mezzo 5227. Da queste ripetendo il procedimento otteniamo 27321529 e quindi 3215.

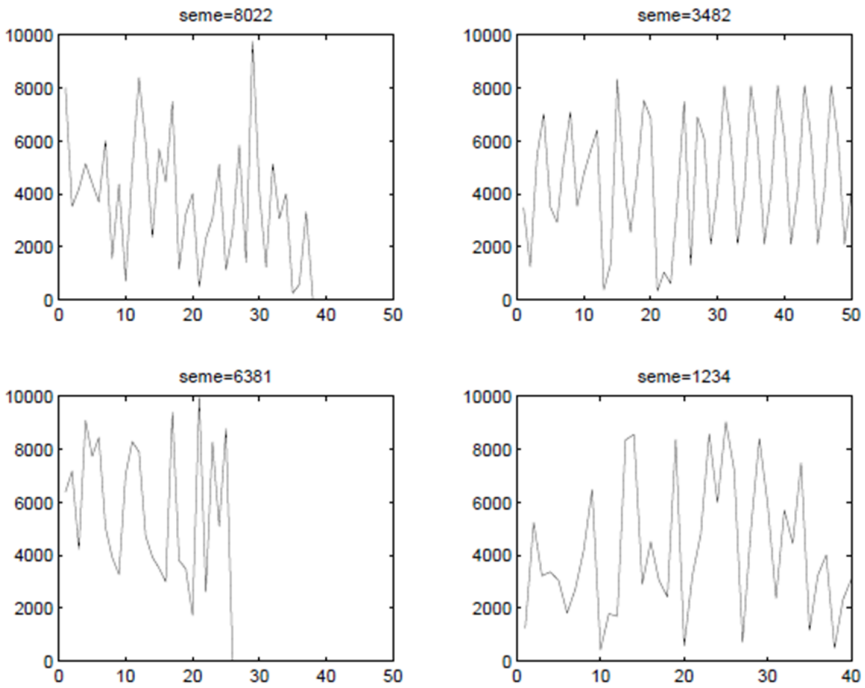
Ogni nuovo numero nella successione è determinato univocamente dal suo predecessore. La successione generata quindi non potrà essere casuale ma avrà solo il carattere di apparente casualità.

Limiti

In particolare ogni successione di numeri generati da questo algoritmo inizierà a ripetersi prima o poi. Il numero di numeri della sequenza prima che intervenga una ripetizione è detto periodo della sequenza. La lunghezza di tale periodo può essere considerata una misura della bontà del generatore di numeri pseudo-casuali.

Sfortunatamente il metodo può degenerare in sequenze con periodi molto brevi, ad esempio a partire dal valore 0 la sequenza ha sempre periodo 1, oppure partendo con 43 e numeri a 2 cifre otteniamo la sequenza 43, 84, 05, 02, 00, 00, ...

Semi e metodo middle-square



Un dado simulato

Proviamo ad utilizzare il metodo middle-square per effettuare un semplice esperimento di simulazione. Consideriamo la simulazione del lancio di un dado definendo il risultato ottenuto d come

$$d = 1 + \left[\frac{5ms}{10^4} \right]$$

dove ms e' il numero generato tramite il metodo middle-square.

Simulando 10 lanci consecutivi a partire dal seme 8022 otteniamo risultati che sembrano abbastanza realistici

5 3 3 4 3 3 4 2 3 1

Basta però aumentare il numero di lanci per giungere a risultati insoddisfacenti (la successione ha infatti periodo 38).

Generatori lineari congruenziali

Nel 1948 venne proposto un generatore di numeri casuali distribuiti uniformemente detto generatore lineare congruenziale o più brevemente LCG che a tutt'oggi è ancora utilizzato. Tale generatore venne presentato per la prima volta dal matematico D.H. Lemer esperto di teoria dei numeri.

Il metodo LCG, analogamente al metodo middle-square, ha bisogno di un seme per inizializzare la sequenza di numeri secondo la seguente regola

$$x_{n+1} = (ax_n + c) \bmod m \qquad n \geq 0$$

dove *a*, *c* ed *m* sono opportuni numeri interi costanti. La notazione "mod *m*", ossia modulo *m*, significa che ax_n+c viene diviso per *m* e x_n+1 posto uguale al resto intero della divisione. Quindi x_n+1 assume valori interi tra 0, 1, 2, . . . ,m–1.

Alcuni esempi

Per esempio le scelte a=13, c=0 (ossia generatore puramente moltiplicativo) e m=31 partendo dal valore iniziale x_0=1 otteniamo

1 13 14 27 10 6 16 22 7 29 5 3 8 11 19 30 18 17 4 21 25 15 9 24 2 26 28 23 20 12

Tale successione ha periodo 30 (ossia m–1). Tutti i numeri da 1 a 30 compaiono per poi ripetersi. Questo non dipende dalla scelta del seme iniziale se non è nullo

L'uso del seme 0 origina invece la successione costante uguale a 0 per c=0 indipendentemente da a e m. Quindi il massimo periodo di un generatore puramente moltiplicativo è m–1. Nel caso di $c \neq 0$ il massimo periodo sarà invece m, in quanto comparirà anche lo 0. Tale periodo massimo non è però raggiunto per tutte le scelte di a, c ed m, ad esempio per a=7, c=7 e m=10 partendo dal seme 7 si ha per n=8

7 6 9 0 7 6 9 0 che ha solo periodo 4.

Scelte sbagliate

Scelte sbagliate dei parametri per questo tipo di generatori hanno rovinato risultati sperimentali fin dai lontani anni 60 (generatore RANDU)

Ad esempio, un errore nella specifica dell'ANSI C ha portato alla propagazione di un terribile generatore per la rand() del C e del C++

Quando usiamo metodi che richiedono la generazione di molti numeri casuali (metodi Montecarlo, meta-euristiche, ecc.) bisogna innanzitutto controllare quale tipo di generatore è utilizzato

I difetti più comuni di una routine di generazione di numeri casuali sono:

- numeri non uniformemente distribuiti
- discretizzazione dei numeri generati
- media o varianza non corrette
- presenza di variazioni cicliche

Bontà di un generatore

Il problema della scelta dei migliori valori per *a*, *c* ed *m* è quindi il punto cruciale del metodo. Cosa si intende però per migliori?

Un aspetto importante è la lunghezza del periodo da cui segue che m dovrà essere molto grande: per un dato *m* i valori di *a* e *c* dovranno essere tali che la successione abbia periodo massimo (se *m* è grande la differenza tra *m* e *m*–1 è irrilevante e ci si può restringere al caso di generatori moltiplicativi, *c*=0, che sono più veloci). Una delle scelte più popolari è

$$m = 2^{31} - 1,\ a = 75,\ c = 0$$

Questo garantisce un periodo di $2^{31}-2$=2.147.483.646 ossia oltre 2 miliardi di numeri pseudo-casuali (il fatto che $m=2^{31}-1$ sia un numero primo è fondamentale al fine di ottenere il massimo periodo).

Bontà di un generatore

Il problema della scelta dei migliori valori per *a*, *c* ed *m* è quindi il punto cruciale del metodo. Cosa si intende però per migliori?

Si dice che la sequenza ottenuta ha periodo pieno se il suo periodo è proprio *m*, e ciò si verifica quando sono verificate le seguenti condizioni:

- Se *m* e *c* sono primi tra loro
- Se *m* è divisibile per un numero primo *b*, per il quale deve essere divisibile anche *a*–1
- Se *m* è divisibile per 4, allora anche a–1 deve essere divisibile per 4

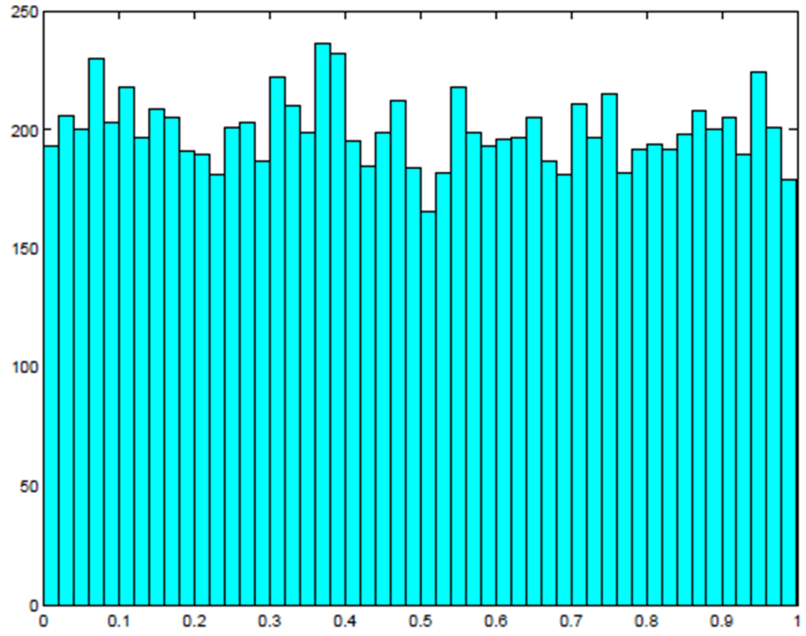
Istogrammi di frequenza

Da un punto di vista pratico molti generatori di numeri casuali restituiscono invece di x_n+1 il suo valore diviso per *m*, al fine di evitare numeri troppo grandi. Inoltre evitano di reinizializzare la sequenza, se non esplicitamente richiesto.

Generiamo una tabella di *N* valori pseudo-casuali compresi tra 0 ed 1 e verifichiamone l'uniforme distribuzione tramite l'istogramma di frequenza ottenuto suddividendo l'intervallo [0,1] in *M* sottointervalli di uguale ampiezza e calcolando quanti valori cadono in un dato intervallo. Nel nostro caso tale valore teorico dovrebbe essere pari a *N/M* (con fluttuazioni – appunto - casuali)

Un esempio

Istogramma di frequenza di *N*=10000 valori in *M*=50 sottointervalli, *N/M*=200



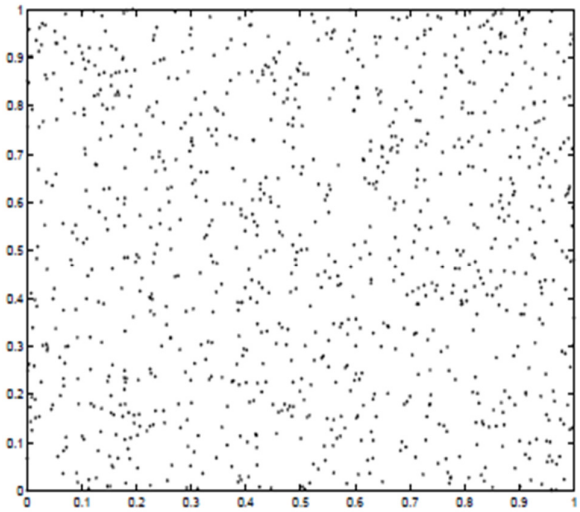
Verificare la casualità

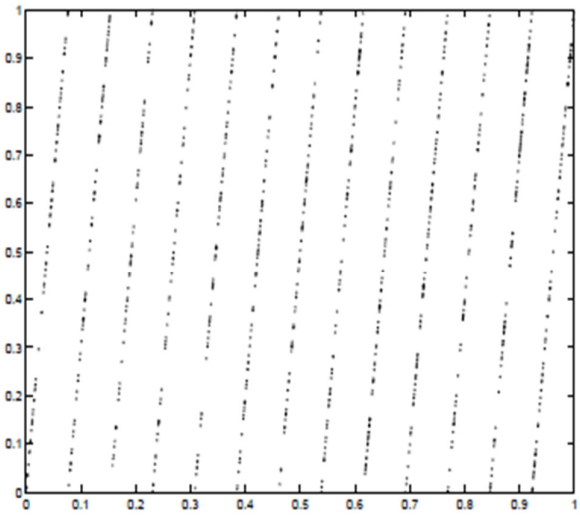
Una richiesta più importante nel valutare la bontà di un generatore uniforme di numeri pseudo-casuali è l'assenza di correlazione tra i numeri generati dall'algoritmo. In altre parole non deve emergere nessuna relazione tra x_n e x_n+1 per n>0.

Questa proprietà può essere verificata graficamente realizzando il grafico bidimensionale dei punti (x_n, x_{n+j}) per j>0. In tale grafico non dovranno comparire linee, forme o altre strutture regolari.

Correlazione

In figura riportiamo il risultato per j = 1 con 1000 punti ottenuto con il generatore LCG con scelta ottimale e con lo stesso generatore modificato scegliendo come valori m = 31, a = 13 e c = 0.





Generatori

I generatori di numeri casuali più recenti non sono basati sul metodo LCG.

Essi sono una combinazione di operazioni di spostamento di registri e manipolazione sui bit che non richiedono nessuna operazione di moltiplicazione o divisione.

Questo nuovo approccio risulta estremamente veloce ed inoltre garantisce periodi incredibilmente lunghi. Nelle ultime versioni MATLAB il periodo è 2^{1492}, che ad un milione di numeri casuali al secondo richiederebbe 10^{435} anni (circa 10^{425} volte l'età dell'universo) prima di ripetersi!. Data la coincidenza dell'esponente con la data della scoperta delle Americhe questo generatore viene comunemente chiamato il generatore di "Cristoforo Colombo".

Altri generatori

I generatori di numeri casuali più recenti non sono basati sul metodo LCG.

Essi sono una combinazione di operazioni di spostamento di registri e manipolazione sui bit che non richiedono nessuna operazione di moltiplicazione o divisione.

Mersenne twister

- Veloce e di alta qualità - permette di generare punti equidistribuiti in spazi fino a 623 dimensioni
- Periodo molto lungo ($2^{19937} - 1$) (i creatori di questo algoritmo hanno dimostrato questa proprietà)
- Utilizza i numeri primi di Mersenne (da cui il nome), e alcune delle costanti dell'algoritmo sono anch'esse numeri primi di Mersenne
- Ha passato numerosi test statistici di casualità, tra cui il test Diehard

L'ambiente C

La libreria di utilità generiche stdlib.h fornisce, tra le altre cose:

- la funzione rand() per la generazione di numeri casuali, con valori compresi tra 0 e la costante RAND_MAX, anch'essa definita in stdlib.h, con valore che può variare a seconda del sistema
- La funzione srand(), che inizializza la sequenza di numeri casuali di rand(). Richiede un parametro di input di tipo unsigned (intero senza segno)

In assenza di inizializzazione non è sempre definito cosa fa un eseguibile (può ad esempio prendere il seme iniziale dal clock del computer, ma non è sempre detto)

L'ambiente C

Per fare in modo (se lo vogliamo) che ogni esecuzione del programma generi un diverso valore del lancio del dado, dobbiamo variare il seme usato per inizializzare il generatore di numeri pseudo-casuali

- Se intendiamo che l'utente abbia il controllo del sistema, gli possiamo permettere di indicare il seme iniziale
- Se vogliamo automaticamente generare ogni volta numeri differenti possiamo utilizzare la funzione time() della libreria time.h che, invocata con parametro NULL restituisce un valore unsigned che rappresenta l'ora corrente del giorno espressa in secondi

Probabilità discreta

Consideriamo il concetto di evento discreto. Un evento E_k rappresenta un elemento di un insieme, al quale è associato un numero non negativo, minore o uguale ad uno, p_k, che rappresenta la probabilità che l'evento k si verifichi.

$$P(E_k) = p_k$$

Se un evento non può verificarsi la sua probabilità è zero. Se un evento si verifica sicuramente (evento certo) la sua probabilità è uno. Esempi:

- L'insieme costituito dalle possibili risultati del lancio di una moneta. Tale insieme è costituito dai due eventi elementari: E_1=«esce testa» ed E_2=«esce croce». A ciascuno di tali eventi viene data probabilità pari a 1/2, poiché i due eventi sono ritenuti equiprobabili

- I possibili risultati del lancio di un dado. L'evento E_k si può definire come la cifra che compare in seguiti ad un lancio, e a ciascun evento è associata la stessa probabilità 1/6.

Eventi e operazioni

Alle probabilità degli eventi si associano le operazioni insiemistiche di unione e intersezione dati due eventi E_i ed E_j vale la relazione

$$P(E_i \cap E_j) = P(E_i) + P(E_j) - P(E_i \cup E_j)$$

Nel caso in cui due eventi E_i ed E_j non possano accadere entrambi simultaneamente si dicono mutuamente esclusivi (o eventi disgiunti). Allora si ha:

$$P(E_i \cup E_j) = P(E_i) + P(E_j) \qquad P(E_i \cap E_j) = 0$$

Se $P(\overline{E}_k)$ indica il complementare di un evento E_k allora vale:

$$P(\overline{E}_k) = 1 - P(E_i)$$

Infine E_i ed E_j si dicono eventi indipendenti se:

$$P(E_i \cap E_j) = P(E_i)P(E_j)$$

Variabile aleatoria o casuale

In taluni casi ad un evento è possibile associare un numero. Ad esempio, nel caso del lancio dei dadi è naturale associare all'evento E_k il numero k che rappresenta il risultato.

Un evento casuale al quale possiamo associare un valore numerico, oltre che probabilità è detto variabile aleatoria. Una variabile aleatoria discreta X è caratterizzata quindi dall'insieme di valori x_i, $i = 1, \ldots, N$ che essa può assumere, e dalle probabilità con la quale essa assume tali valori.

$$p_i = P(\{X = x_i\})$$

La funzione $f(x_i)=p_i$ è detta funzione di probabilità di X. Valgono:

$$\sum_{i=1}^{N} f(x_i) = 1 \qquad \mu = \sum_{i=1}^{N} f(x_i) x_i \qquad \sigma^2 = \sum_{i=1}^{N} f(x_i)(x_i - \mu)^2$$

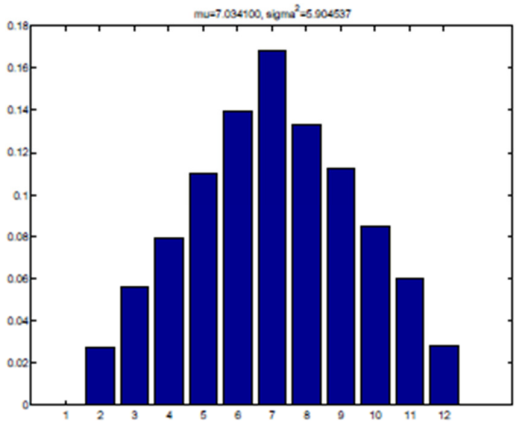
Dadi

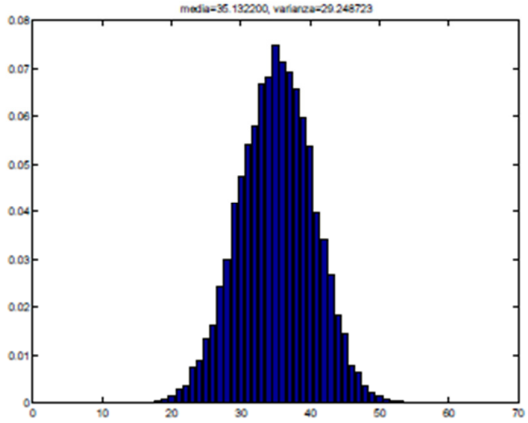
Consideriamo ora il problema del lancio simultaneo di n dadi ed indichiamo con X la variabile aleatoria che indica il punteggio ottenuto. Se n=2 non è difficile calcolare tale probabilità. Infatti X potrà assumere i valori 2, . . . , 12 e è sufficiente considerare il rapporto tra i casi favorevoli ed il numero di possibili combinazioni del punteggio dei due dadi (che si verifica subito essere 36)

$$\begin{array}{cccc}
(1,1) & (1,2) & (1,3)\ldots & (1,6) \\
(2,1) & (2,2) & (2,3)\ldots & (2,6) \\
\vdots & \vdots & \vdots & \vdots \\
(6,1) & (6,2) & (6,3)\ldots & (6,6)
\end{array}$$

Se n è molto grande il calcolo diretto risulta più complesso. Da un punto di vista probabilistico il lancio di n dadi è equivalente al lancio di un singolo dado effettuato n volte in modo indipendente.

Lancio di *n* dadi



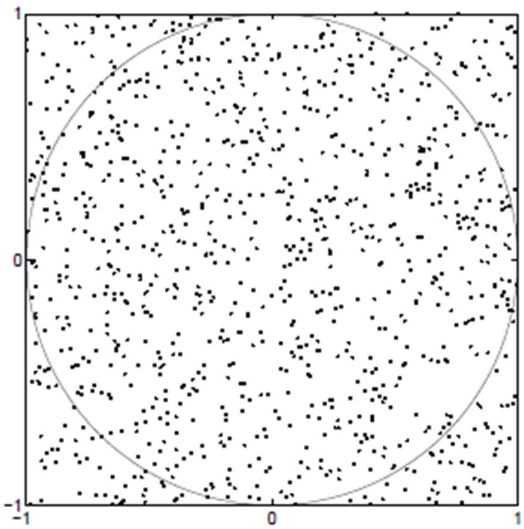


Nel caso di due dadi (figura a sinistra) il punteggio più probabile è 7 (probabilità 6/36=1/6) e quelli meno probabili 2 e 12 (probabilità 1/36).

All'aumentare del numero dei dadi (destra) le frequenze tendono ad essere distribuite secondo una campana gaussiana (la distribuzione cosiddetta normale).

Bersagli, aree e π

Supponiamo di lanciare *N* freccette ad un bersaglio formato da un quadrato di lato *L* contenente una circonferenza. Assumiamo che le freccette vengano lanciate casualmente all'interno del quadrato e che quindi colpiscano il quadrato in ogni posizione con uguale probabilità. In Figura riportiamo il bersaglio per *L*=2 e la posizione dei primi 1000 lanci.

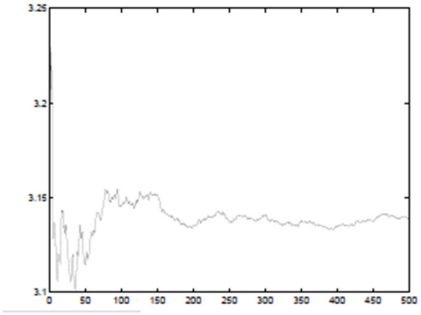


Stima di π

Dopo molti lanci la frazione di freccette che ha colpito la circonferenza sarà circa uguale al rapporto tra l'area della circonferenza e quella del quadrato. Quindi

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

dove N è il numero totale di freccette lanciate e N_c indica il numero di freccette cadute all'interno della circonferenza. Potremo quindi usare il valore $4N_c/N$ come approssimazione di π.



Calcolo di aree e volumi

La stima di π migliora all'aumentare di N ma che la convergenza è tutt'altro che veloce e uniforme. La presenza di fluttuazioni dovute all'approccio probabilistico è infatti una caratteristica dei metodi Monte Carlo.

Si noti inoltre che il metodo Monte Carlo precedente avrebbe potuto essere utilizzato per calcolare l'area della circonferenza. Infatti indicando con A_c tale area si ha

$$A_c \approx L^2 \frac{N_c}{N}$$

In modo analogo si può procedere per il calcolo di aree di altre figure piane. La generalizzazione al calcolo di volumi nello spazio è immediata. Indicato con L il lato del cubo che contiene la figura di cui si vuole misurare il volume V avremo

$$V \approx L^3 \frac{N_c}{N}$$

La generalizzazione al calcolo di volumi inscrivibili in parallelepipedi è altrettanto semplice (basta fare le proporzioni…)

Diffusione

Considerate il seguente esperimento (immaginario):

>> Calate una paratia a dividere in due questa stanza e pompate tutta l'aria in una sola metà (ovvero fate il vuoto nell'altra). Cosa succede se poi rimuovete la paratia?

L'aria diffonde fino a ridistribuirsi (in media) in egual misura nei due settori ...

>> E' possibile viceversa che da una situazione in cui c'è aria in entrambe le metà della stanza si passi a una in cui tutta l'aria viene a concentrarsi da una parte sola?

Questo è in sostanza il problema della **reversibilità/irreversibilità** dei processi in natura!

Possiamo maneggiare (e simulare) un modello per il problema in esame!

Diffusione

Prendete due urne e disponetevi N biglie numerate (ovvero n in una urna e N-n nell'altra).

Estraete un numero fra 1 e N e cercate la biglia numerata come il numero estratto: prendetela e spostatela di urna.

Iterate la procedura ...

Dopo aver iterato per un po', ha senso chiedersi qual è il numero medio di biglie in ciascuna delle due urne? In altri termini, c'è una situazione di *equilibrio*?

Il problema può essere **formalizzato** in termini matematici, **affrontato analiticamente**, ma anche simulato al calcolatore

Diffusione

In molti problemi di tipo probabilistico si incontrano urne contenenti palline di diversi colori. Tale semplice situazione infatti può servire a realizzare semplici modelli di situazioni fisiche estremamente complesse. Come esempio consideriamo il caso della diffusione di un gas (il modello di diffusione basato sul processo di estrazione e scambio da due urne è dovuto al fisico austriaco Paul Ehrenfest (1880-1933)).

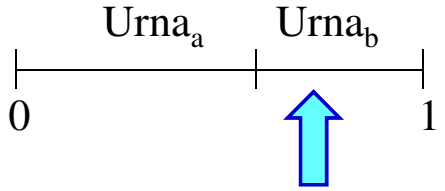
Si considerino due urne, indichiamole con *A* e *B*, ognuna contenete n palline. Al tempo t=0 nell'urna A abbiamo solo palline nere, mentre nell'urna *B* abbiamo solo palline bianche (possiamo immaginare che le diverse palline rappresentino molecole di gas di tipo diverso separate all'istante iniziale da una membrana). Al tempo t=1 (inteso come numero di iterazioni) una pallina è estratta a caso da ogni urna e posta nell'altra (in pratica corrisponde alla rottura della membrana iniziale). Il procedimento viene ripetuto ai tempi successivi t= 2, 3, ... Ad ogni istante in ogni urna avremo sempre *n* palline. Il problema può essere affrontato analiticamente ma conduce a calcoli lunghi e laboriosi; può essere conveniente utilizzare una simulazione di tipo Monte Carlo.

Diffusione

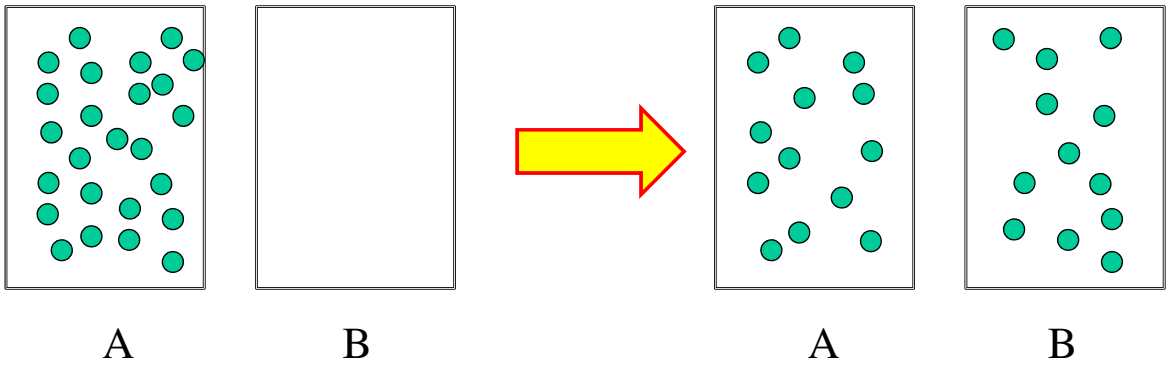
In molti problemi di tipo probabilistico si incontrano urne contenenti palline di diversi colori. Tale semplice situazione infatti può servire a realizzare semplici modelli di situazioni fisiche estremamente complesse. Come esempio consideriamo il caso della diffusione di un gas (il modello di diffusione basato sul processo di estrazione e scambio da due urne è dovuto al fisico austriaco Paul Ehrenfest (1880-1933)).

Per la simulazione possiamo quindi calcolare la probabilità che la pallina selezionata per il passaggio appartenga ad una delle due urne (possiamo stimare queste probabilità come il rapporto del numero di palline presenti in ciascuna urna, diviso il numero totale di palline). Si estrae quindi un numero casuale $\in [0,1]$: il punto di «caduta» indica l'urna da scegliere.

Dopo ogni estrazione ed il conseguente passaggio di una pallina dall'urna prescelta all'altra, le probabilità di estrazione devono essere ricalcolate.



Diffusione – situazione 1

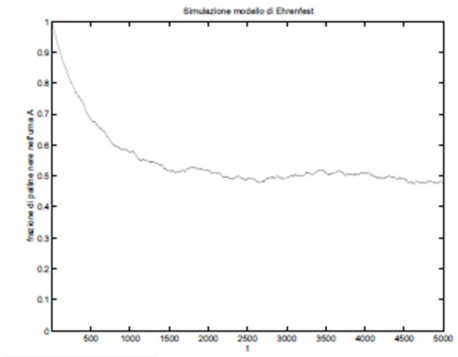


Probabilità di avere una pallina in A: palline_in_A/palline_totali

Probabilità di avere una pallina in B: palline_in_B/palline_totali

Simulazione Monte Carlo

Il risultato della simulazione è riportato in figura (frazione di palline nere nell'urna A) nel caso di 1000 palline ed un tempo finale pari a 5000. Il sistema evolve allo stato nel quale ci sono circa lo stesso numero di palline bianche e nere in ciascun'urna. In particolare osserviamo che il sistema raggiunge questo stato ma le fluttuazioni attorno a tale stato non diminuiscono nel tempo. Per diminuire tali fluttuazioni possiamo aumentare il numero di palline in ciascun urna oppure eseguire diverse simulazioni e calcolare delle medie dei risultati ottenuti.



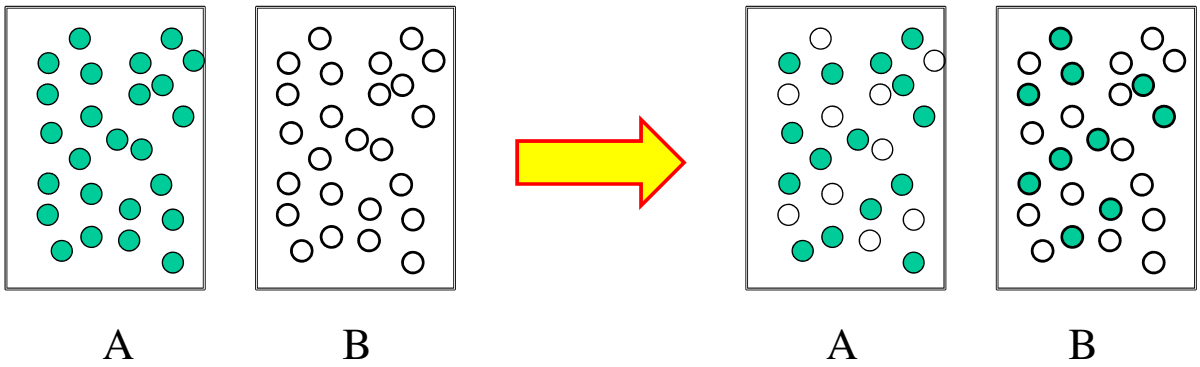
Diffusione – situazione 2

Probabilità di avere una pallina verde in A: palline_verdi_in_A/palline_totali

Probabilità di avere una pallina bianca in A: palline_bianche_in_A/palline_totali

Probabilità di avere una pallina verde in B: palline_verdi_in_B/palline_totali

Probabilità di avere una pallina bianca in B: palline_bianche_in_B/palline_totali



Moti casuali

I moti casuali sono il classico esempio di problemi facili da comprendere da un punto di vista probabilistico ma difficili (se non impossibili) da risolvere senza l'ausilio di un calcolatore.

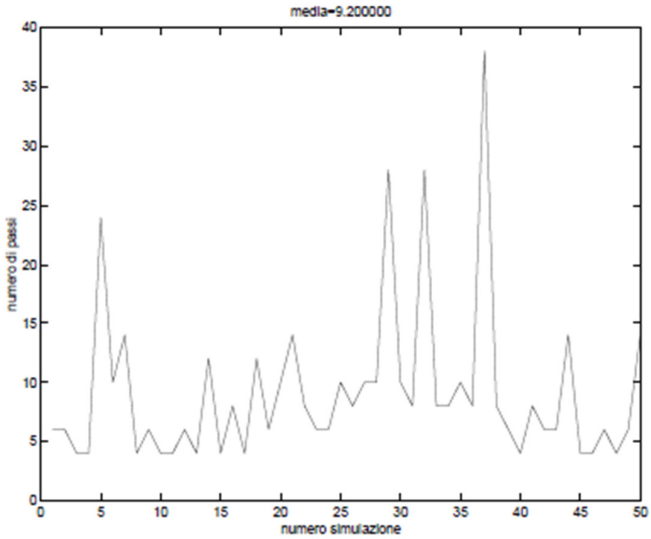
Il caso più semplice è il seguente. Si consideri l'asse delle x suddiviso in nodi equispaziati in 0,±1,±2,…. Una particella è posizionata in x=0 a t=0. Ad ogni istante la particella si sposta di un unità a destra o a sinistra in modo equiprobabile.

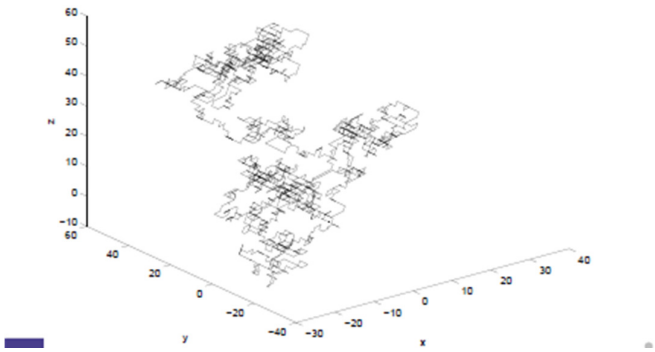
La probabilità che la particella raggiunga il punto $x=n$ nel caso non vi sia alcun limite al numero di spostamenti è uguale ad 1. Infatti, indicata con $p_i=p(\{x=i\})$ la probabilità che la particella raggiunga la posizione x=i, abbiamo $p_i=1/2(p_{i+1}+p_{i-1})$, poiché la posizione x=i può essere raggiunta solo da x=i+1 o x=i−1 in modo equiprobabile.

Quindi il grafico di p_i rispetto ad i è una linea retta. Tale retta dovrà essere orizzontale, poiché altrimenti potremmo avere probabilità minori di zero o maggiori di uno. Infine essendo p_0=1, dato che tale posizione è occupata all'istante iniziale, avremo p_i=1 per ogni valore di i (e quindi il valore atteso del numero di passi necessari a raggiungere la barriera è dato da n^2)

Simulazione Monte Carlo

Il numero di passi prima che la particella arrivi in x=3 per 50 simulazioni è riportato in figura



Moto Browniano

Lo stesso problema può essere simulato in più dimensioni, realizzando un esempio di moto browniano (un esempio è dato dal moto di piccole particelle sospese nell'acqua dovuto all'azione delle molecole di acqua. Tale moto che fornisce l'evidenza fisica dell'esistenza delle molecole e quindi dell'atomo, fu per primo riportato dal botanico scozzese Robert Brown (1773-1858) e quindi studiato matematicamente da Albert Einstein (1879-1955)).