



UNIVERSITÀ DEGLI STUDI DI FIRENZE

FACOLTÀ DI INGEGNERIA

Corso di Laurea Specialistica in Ingegneria per la Tutela dell'Ambiente e del Territorio

A.A. 2006 - 2007

**Corso di Fisica Statica e dei
Fenomeni Diffusivi**

PROF. FRANCO BAGNOLI

Introduzione al metodo Monte Carlo

RELATORE

IRENE MORONI

1 - Indice generale

1 -Origini	4
2 -In cosa consiste	6
3 -Utilizzi del metodo	8
4 -Metodi Monte Carlo	10
4.1 -Un Gioco nel Principato di Monaco	10
4.2 -L'algoritmo di Toddlers e la Condition of detailed Balance	14
4.2.i -Il concetto di rigetto : l'algoritmo di Metropolis	17
4.2.ii -to Cry or to Cray !	18
4.2.iii -Una probabilità a priori	23
4.2.iv -Perturbazioni	25
4.2.v -Algoritmi di Gruppo	27
4.2.vi -Osservazioni sull'equilibrio Monte Carlo	30
5 -Esempi concreti	31
5.1 -Monte Carlo grezzo	31
5.2 -Stima dell'Attesa di un Autobus	32
5.3 -Applicazioni	34
5.4 -Applicazione didattica del metodo Monte Carlo	34
5.4.i -Lancio di due dadi	35
5.4.ii -Lancio di n monete	35
5.4.iii -Modello di Ehrenfest	36
5.4.iv -Simuliamo una passeggiata a caso: il quincunx di Galton	36
5.4.v -Il problema di Méré	37
5.4.vi -Problema dei compleanni	37
5.4.vii - Selezione darwiniana	38
5.4.viii -Il problema dell'estinzione	38

5.4.ix -Sviluppo di popolazioni	39
5.4.x -Guadagno di un giocatore	39
5.4.xi -Problema del collezionista	40
5.4.xii -Il problema delle concordanze	40
5.4.xiii -Il gioco del Craps	41
5.4.xiv -I cerini di Banach	41
5.4.xv -Il gioco dell'un-due-tre	42
5.4.xvi -Distribuzione uniforme	42
5.4.xvii -Probabilità di $(a \leq R \leq b)$ e di $(c \leq R \leq d)$ in n estrazioni di R	43
5.4.xviii -Punti a caso in un quadrato unitario	43
5.4.xix -Area della superficie sotto la retta $y=mx$	44
5.4.xx -Un metodo statistico per il calcolo di π	45
5.4.xxi -Area sotto la parabola $y=x^2$	45
5.4.xxii -Area sotto l'iperbole $y = 1/x$	46
5.5 -Modelli non parametrici	46

2 - Origini

Nel 1777 il matematico francese George Louis Leclerc de Buffon descrive nella sua opera *Essai d'Arithmetique morale* un esperimento di stima del valore tramite una "simulazione" casuale basata sul lancio di uno spillo, ottenendo un risultato di buona precisione numerica per l'epoca.

Lo stesso Enrico Fermi all'inizio degli anni '30 sosteneva di utilizzare stime ottenute con tecniche di campionamento statistico per lo studio del moto dei neutroni. Il metodo Monte Carlo fu formalizzato negli anni '40 da J. Von Neumann e S. Ulam che partecipavano al Progetto Manhattan per lo studio della dinamica delle esplosioni nucleari.

Lo stesso Von Neumann è considerato l'inventore del calcolatore elettronico. 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, sede per antonomasia dell'aleatorietà, dove, la fortuna o la rovina di coloro che per vizio, noia o altro si cimentano nel gioco, sia decretata dalla assoluta casualità che si manifesta nell'uscita di questo o quel numero.

Negli esperimenti nucleari una particella veloce colpisce in pieno il nucleo di un atomo. Questo si frantuma in molte particelle, che vanno a colpire i nuclei di altri atomi vicini, che si frantumano a loro volta, secondo una reazione a catena, nel corso della quale si libera una gran quantità d'energia. Il processo durerà fino a coinvolgere l'intero universo o s'arresterà, dopo un certo numero di reazioni? Sappiamo già che il secondo caso è quello che ci si attende, come è stato confermato sperimentalmente dai fisici. Con quale stato d'animo è stato tuttavia possibile eseguire il primo esperimento? I fisici erano certi del fatto che la reazione avrebbe avuto termine?

In realtà, l'esperimento fisico fu preceduto da una serie di simulazioni. Una volta introdotti alcuni parametri iniziali, il fenomeno fu simulato da un calcolatore per mezzo di valori casuali, trattati con metodi statistici.

Si poteva così stimare la probabilità che, dopo un certo numero di "generazioni", le particelle emesse nel corso delle reazioni a catena, cessassero di generare altre particelle. Le simulazioni dettero sufficienti garanzie e gli esperimenti reali furono eseguiti con una sufficiente dose di tranquillità.

Un primo esempio di utilizzo del metodo Monte Carlo è rappresentato dall'esperimento dell'ago di Buffon e forse il più famoso utilizzo di tale metodo è quello di Enrico Fermi, quando nel 1930 usò un metodo casuale per calcolare le proprietà del neutrone.

Sono tanti ormai i campi in cui si utilizzano metodi statistici per ottenere informazioni e stime su fenomeni legati al caso. Non occorre che i dati siano raccolti durante un esperimento reale in cui tali fenomeni avvengono. Ciò potrebbe richiedere troppo tempo e, in ogni caso, non sempre la natura è disposta a fornirci situazioni aleatorie ... a comando. I dati possono allora provenire da simulazioni fatte per mezzo di un computer, in grado di generare sequenze di numeri casuali. Esse sono quindi utilizzate per simulare per migliaia

di volte il fenomeno aleatorio, raccogliendo così rapidamente una serie di dati che, trattati con metodi statistici, forniscono stime che diventano tanto più attendibili quanto più è grande il numero delle prove fatte.

Oggi il metodo Monte Carlo trova applicazione in vari ambiti scientifici dall'analisi del rischio nella valutazione degli investimenti a molti problemi fisici in cui gli esperimenti sono molto difficili da realizzare, e richiedono molto tempo per essere riprodotti. In questi casi il metodo Montecarlo è uno strumento fondamentale per l'interpretazione di dati sperimentali e la verifica di nuove teorie.

2 - In cosa consiste

Spesso ci si trova di fronte a situazioni in cui si ha bisogno di conoscere la probabilità di un certo evento, ma le variabili che lo condizionano sono troppe e non è possibile svolgere i calcoli analitici. In tali situazioni si fa ricorso a 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 alla valutazione analitica l'osservazione empirica del fenomeno e traendo da questa le informazioni non rilevabili per via analitica. Ad esempio, la frequenza osservata di un certo evento costituisce una valutazione della probabilità di quell'evento (a patto che il campionamento sia stato simulato per un consistente numero di volte). Questa simulazione prende il nome di metodo di Montecarlo.

Il metodo di Monte Carlo consiste nel cercare la soluzione di un problema, rappresentandola quale parametro di una ipotetica popolazione e nello stimare tale parametro tramite l'esame di un campione della popolazione ottenuto mediante sequenze di numeri casuali.

Il metodo Montecarlo è una procedura numerica usata in fisica per riprodurre lo stato di un sistema. In generale questo metodo permette di generare eventi secondo un'opportuna distribuzione di probabilità, quindi può essere applicato a qualsiasi fenomeno di cui si conosca la probabilità di occorrenza.

Per esempio supponiamo che I sia il valore incognito da calcolare e che si possa interpretare quale valore medio di una variabile casuale X . Il metodo di Monte Carlo consiste, in questo caso, nello stimare I mediante il calcolo della media di un campione costruito determinando N valori di X ; ciò si ottiene tramite un procedimento che prevede l'uso di numeri casuali.

Il metodo della simulazione Montecarlo è una tecnica numerica per la trattazione di problemi caratterizzati da una sostanziale intrattabilità analitica. Inoltre, ha particolare successo per risolvere il problema della rappresentazione dei portafogli composti da strumenti finanziari ad alto contenuto opzionale o non lineare e la distribuzione dei rendimenti del portafoglio stesso.

Il metodo Montecarlo è una valida alternativa agli approcci parametrici nel caso in cui il portafoglio comprenda posizioni con andamento di prezzo non lineare. La seconda caratteristica attraente riguarda la possibilità di adattare per la generazione degli scenari random, distribuzioni di probabilità dei fattori di rischio non normali.

Rimane fondamentale, comunque, la scelta delle funzioni di valutazioni appropriate, il rischio di scegliere un modello di valutazione errato è chiamato model risk.

Il Metodo Monte Carlo fa parte della famiglia dei metodi statistici non parametrici. È utile per superare i problemi computazionali legati ai test esatti (ad esempio i metodi basati sulla distribuzione binomiale e calcolo combinatorio, che per grandi campioni generano un numero di permutazioni eccessivo).

Il metodo è usato per trarre stime attraverso simulazioni. Si basa su un algoritmo che genera una serie di numeri tra loro incorrelati, che seguono la distribuzione di probabilità

che si suppone abbia il fenomeno da indagare. L'incorrelazione tra i numeri è assicurata da un test chi quadrato.

La simulazione Monte Carlo calcola una serie di realizzazioni possibili del fenomeno in esame, con il peso proprio della probabilità di tale evenienza, cercando di esplorare in modo denso tutto lo spazio dei parametri del fenomeno. Una volta calcolato questo campione rappresentativo, la simulazione esegue delle 'misure' delle grandezze di interesse su tale campione. La simulazione Monte Carlo è ben eseguita se il valore medio di queste misure sulle realizzazioni del sistema converge al valore vero.

Da un altro punto di vista le simulazioni Monte Carlo non sono altro che una tecnica numerica per calcolare integrali.

L'algoritmo Monte Carlo è un metodo numerico che viene utilizzato per trovare le soluzioni di problemi matematici, che possono avere molte variabili e che non possono essere risolti facilmente, per esempio il calcolo integrale. L'efficienza di questo metodo aumenta rispetto agli altri metodi quando la dimensione del problema cresce.

3 - Utilizzi del metodo

Non si tratta di un metodo molto efficiente per il calcolo di π , ma illustra il concetto. Un matematico direbbe che il MC serve a calcolare integrali, in particolare integrali che sono difficili da calcolare numericamente per altra via perché per esempio l'integrando varia molto rapidamente. In fisica tante cose si esprimono tramite integrali per cui il metodo MC è di uso abbastanza comune.

Uno dei primi casi in cui il MC fu usato fu il calcolo delle reazioni nucleari a catena. Quando un neutrone urta un nucleo non si può dire con certezza se il nucleo si scinda o meno, né quanti neutroni verranno prodotti: conosciamo solo la probabilità di questi eventi. Allora simuliamo l'urto generando un numero casuale (i primi calcoli vennero fatti usando una roulette come generatore di numeri randomici) ed a seconda del numero uscito diciamo "è avvenuto questo o quell'evento". Per esempio, poniamo che si sappia che in ogni urto la probabilità di generare un neutrone è il 45%, quella di generarne due è il 30%, e quella di non generarne nessuno è il 25%. Allora estraiamo per esempio un numero fra 1 e 100 e diciamo: "se esce un numero fra 1 e 45 supponiamo che sia stato prodotto un neutrone; se esce un numero fra 46 e 75 supponiamo che ne siano stati prodotti due; e se esce un numero fra 76 e 100 supponiamo che non ne sia stato generato nessuno." La simulazione dunque procede come segue:

- si parte con un certo numero di neutroni;
- si calcola la probabilità che ciascun neutrone ne produca altri, col metodo descritto sopra;
- si calcolano le altre grandezze a cui si possa essere interessati (p.es. l'energia rilasciata in ogni scissione);
- per ognuno dei nuovi neutroni si calcola la probabilità di generarne altri, e si ripete...

Metodi di questo tipo sono tuttora usati in ingegneria nucleare, ed altri simili si usano per descrivere il moto degli elettroni nei dispositivi elettronici al fine di predirne o studiarne il funzionamento.

Altri esempi importanti riguardano il calcolo di proprietà della materia condensata (solidi e liquidi) e no (plasma etc). In un sistema che non sia allo zero assoluto atomi e molecole sono agitati in continuazione ed in modo sostanzialmente casuale (moto Browniano). Con il metodo MC si simula questa agitazione generando numeri casuali e spostando gli atomi/le molecole di conseguenza. Come esattamente vengano effettuati gli spostamenti è un argomento un po' tecnico e non posso descriverlo qui.

La simulazione consiste nello studio del comportamento di un sistema mediante la sua riproduzione in un contesto controllabile. Nella simulazione al computer si costruisce un modello matematico costituito da equazioni che descrivono le relazioni tra le componenti del sistema oggetto di studio e il loro legame con il suo funzionamento/comportamento, con l'obiettivo di effettuare esperimenti "virtuali" sul modello matematico assumendo che i risultati di tali esperimenti costituiscano una "riproduzione" sufficientemente accurata del comportamento che avrebbe il sistema.

Questo allo scopo di accrescere la comprensione del suo funzionamento, verificare (o negare) la validità di ipotesi su di esso, raccogliere informazioni per poter formulare possibili previsioni, per implementare meccanismi di controllo del sistema modellato, ecc. Oggi il metodo Montecarlo viene usato per moltissime finalità, dalla decisione di dove effettuare perforazioni petrolifere all'ottimizzazione della compattazione dei rifiuti negli impianti di trattamento. Apparentemente è semplice ma è molto potente. La base è il lancio dei dadi. Se si lancia un dado 100 volte e se ne registrano i risultati si scoprirà che i numeri presenti sulle sei facce usciranno circa un sesto di volte per uno, ma non esattamente. Questo perché i risultati sono casuali. Se si lancia il dado mille volte e poi un milione di volte, i risultati si avvicineranno sempre di più a quelli teorici. È la cosiddetta legge dei grandi numeri.

Il dado rappresenta i rischi, predicibili e distribuiti in modo eguale: ogni faccia/rischio ha un sesto di probabilità di accadere e cinque sestimi di non accadere. Ora, supponiamo che ogni dado sia un rischio di progetto e le facce rappresentino i possibili accadimenti. Esempio: un dado rappresenta il rischio che un progetto venga ritardato a causa di cambiamenti nello staff. Una faccia potrebbe portare la possibilità che il progetto ritardi sei mesi a causa di un turnover del 20%, un'altra un ritardo di due anni a causa di un turnover dell'80%. Si possono introdurre dei pesi, per cui il dado potrebbe essere sbilanciato in modo da rendere più o meno probabili certi risultati. Ogni rischio avrebbe il suo dado: sviluppo con molti errori, tagli di budget.

I simulatori che usano il metodo di Montecarlo "lanciano" tutti i dadi/rischio assieme e ne registrano i risultati combinati. Più volte si tira il dado, più la distribuzione dei possibili risultati è esatta. Se si mettono i risultati su un grafico cartesiano ne risulta una curva che sembra un formicaio, dove il punto più elevato rappresenta il risultato combinato più probabile e le terminazioni laterali i risultati sempre possibili ma meno probabili.

Una volta determinato in questo modo il profilo di rischio di un progetto, è possibile inserire nel progetto ulteriori risorse (denaro e tempo, per esempio) per mitigare i rischi che si trovano nei punti più alti della curva. Se così la distribuzione dice che il progetto ha il 50% di probabilità di andare in ritardo di sei mesi, si può decidere di prevedere ulteriori tre mesi di tempo di sviluppo per mitigare questo rischio.

Le simulazioni di Montecarlo permettono di effettuare "analisi di sensibilità", ossia lanciare un solo dado tenendo fermi gli altri su un certo risultato per vedere cosa succede quando cambia un solo rischio. Questo può consentire di scoprire che solo una piccola parte dei rischi presi in considerazione rappresenta la stragrande maggioranza delle probabilità globali di rischio di un progetto, e quindi concentrare su quelli le energie.

Le simulazioni di Montecarlo dovrebbero essere effettuate su tutti i progetti presenti a portafoglio, per classificarli in ordine di rischio. Ciò consente di generare una frontiera della massima efficienza, ossia una linea che collega la combinazione di progetti che garantisce i maggiori benefici a un livello di rischio predeterminato. La "frontiera efficiente" permette di evitare i rischi non necessari e impedisce che venga scelto un portafoglio con rischi uguali ma benefici più bassi di un altro.

5 - Metodi Monte Carlo

5.1 - UN GIOCO NEL PRINCIPATO DI MONACO

Una curiosa applicazione del metodo Monte Carlo può essere trovata in un gioco molto popolare nel Principato di Monaco, che anche i bambini praticano sulle spiagge. I bambini si riuniscono, raccolgono un bastone grande, disegnano un cerchio ed un quadrato come mostrato nell'illustrazione 1,1 e riempiono le loro tasche con dei ciottoli¹.



Illustrazione 1,1: I bambini al gioco sulle spiagge di Monaco. Passano i loro pomeriggi calcolando il π con un metodo che può essere facilmente esteso a integrali generali.

Poi stanno in piedi intorno al quadrato, chiudono i loro occhi, e lanciano i ciottoli a caso in direzione del quadrato stesso. Qualcuno tiene al corrente del numero dei ciottoli che colpiscono il quadrato, e di quanti cascano dentro il cerchio (vedi illustrazione 1,1).

Si può verificare facilmente che il rapporto fra i ciottoli nel cerchio, e quelli nel quadrato sarà $\pi/4$.

Questo passa tempo è il solo metodo per calcolare il π con precisione arbitraria senza usare strumenti di misura (il metro, bilancia) o le matematiche avanzate (la divisione, la moltiplicazione, la trigonometria). I bambini in Monaco possono passare il pomeriggio intero con questo gioco.

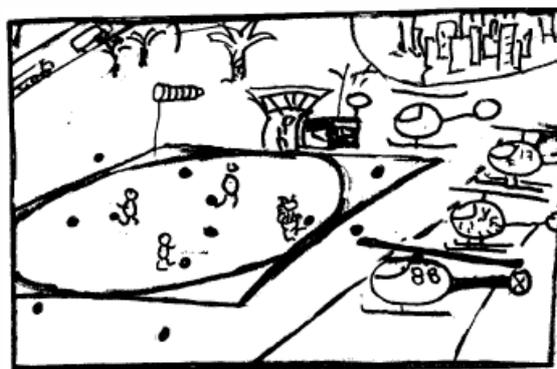


Illustrazione 1,2: Gli adulti al gioco sull'eliporto di Monte Carlo. Il loro metodo di calcolare il π è stato esteso all'investigazione di liquidi, alle sospensioni, e alla teoria di calibro di griglia.

¹ ciottoli è il calculus in latino

Anche gli adulti in Monaco praticano un gioco simile. A tarda sera, quando tutti gli elicotteri sono fermi a terra, prendono possesso dell'eliporto locale (vedi illustrazione 1,2), che offre la stessa disposizione fondamentale del gioco dei bambini. Riempiono le loro costose borsette con dei ciottoli, ma, poiché il campo è grande, giocano un po' diversamente : cominciano in un posto a caso sul campo, chiudono i loro occhi, e lanciano poi un ciottolo in una direzione a caso. Poi camminano fino a dove la pietra è atterrata, quindi tirano fuori dalla loro borsetta un altro ciottolo e lo lanciano, e così via. Si realizza che usando questo metodo, uno può spazzare uniformemente il quadrato dell'eliporto, e quindi calcolare il numero π .

Questa versione del gioco è molto simile al gioco dei bambini, ma si faccia attenzione a questa situazione : consideriamo una signora posta in un punto ψ . Ha appena lanciato un ciottolo in ψ' , che è fuori del quadrato. Che dovrebbe fare?

- Muoversi semplicemente verso ψ' .
- Posizionarsi fuori dal campo di gioco e lanciare un nuovo ciottolo, e continuare finché, per caso, il ciottolo non cade dentro i campo.
- Altro ...

Daremo una risposta corretta a questa domanda successivamente, in quanto contiene l'essenza del concetto di equilibrio dettagliato. Molti programmi che applicano Monte Carlo sono errati perché sbagliano il comportamento in questa situazione.

I due casi, il gioco dei bambini e gli adulti, sono delle illustrazioni perfette di ciò che succede negli algoritmi di Monte Carlo. In ogni caso, è interessante valutare questo integrale :

$$\int_{(x,y) \in \square} dx dy \pi(x,y) f(x,y) \tag{1,1}$$

con una densità di probabilità π , che nel nostro caso è il quadrato :

$$\pi(x,y) = \begin{cases} 1 & \text{se } |x| \leq 1 \text{ e } |y| \leq 1 \\ 0 & \text{altrimenti} \end{cases} \tag{1,2}$$

e la funzione f , che nel nostro caso è il cerchio :

$$f(x,y) = \begin{cases} 1 & \text{se } x^2 + y^2 < 1 \\ 0 & \text{altrimenti} \end{cases} \tag{1,3}$$

Sia i bambini che gli adulti riempiono il quadrato con una densità costante di ciottoli (corrispondenti a $\pi(x,y)=1$), si può dire che campionano la funzione $\pi(x,y)$ sul quadrato fondamentale. Se si riflette su questo, si realizzerà che nei due giochi il π può essere calcolato soltanto perché l'area del quadrato fondamentale è nota. Se così non fosse, ci si ridurrebbe al calcolo del rapporto fra le aree del cerchio e del quadrato, cioè, nel caso generale, il rapporto degli integrali :

$$\frac{\int_{(x,y) \in \square} dx dy \pi(x,y) f(x,y)}{\int_{(x,y) \in \square} dx dy \pi(x,y)} \tag{1,4}$$

Si capisce quindi che nei metodi Monte Carlo si possono avere due approcci :

- Il campionamento diretto (i bambini sulla spiaggia)
- Il campionamento della catena di Markov (gli adulti sull'eliporto)

Il campionamento diretto è noto come *GOLD PURE*, significa che si può chiamare una subroutine che fornisce un dato indipendente dalla sua funzione di distribuzione $\pi(x, y)$, ovvero che ogni dato è indipendente dagli altri.

Questo è esattamente quello che i bambini fanno quando prendono dalle loro tasche un nuovo ciottolo.

Spesso, non c'è modo di fare il campionamento diretto in un modo ragionevole. Quindi si ripiega sul campionamento della catena di Markov, come gli adulti sull'eliporto. Quasi tutti i sistemi fisici sono di quest'ultima classe.

Un esempio famoso, che non ci obbliga a parlare di energia, hamiltoniano ecc. e che ha occupato più di una generazione di fisici, può essere facilmente creato con una scatola da scarpe, e delle monete (l'illustrazione 1,3) : come possiamo generare, applicando un campionamento diretto, configurazioni casuali di monete tali che queste non si sovrappongano?

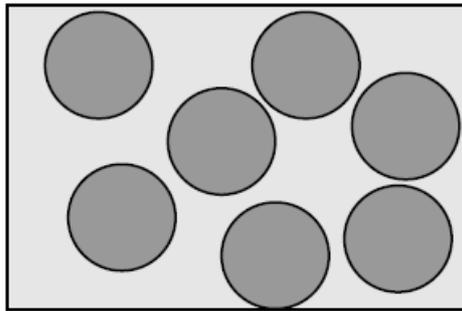


Figura 1,3 : Il problema delle Monete in una di scatola da scarpe ha occupato gli applicatori di Monte Carlo dal 1953. Questo non è un algoritmo a campionamento diretto, ed il campionamento della catena di Markov è estremamente tedioso ad alcune densità.

Immaginiamo lo spazio di configurazione $2N$ -dimensionale di N monete non-sovrapposte in una scatola da scarpe. Nessuno ha trovato una subroutine che farebbe il campionamento diretto di questo spazio di configurazione, cioè in grado di creare ogni dato con probabilità uguale e quindi indipendente dagli altri. A questo punto spesso viene proposto un metodo chiamato *Random Sequential Adsorption* : depositiamo una prima moneta in una posizione a caso, poi una seconda, ecc., (se si sovrappongono, semplicemente si tenta ancora). Ci si deve rendere conto che questo non ha niente a che fare con la ricerca di una configurazione di monete non-sovrapposte disposte a caso nella scatola da scarpe.

Il campionamento diretto è spesso impossibile, e questo rappresenta la frustrazione fondamentale dei metodi di Monte Carlo (nella fisica statistica). Invece, il gioco degli adulti può essere giocato in qualunque situazione (l'eliporto è già troppo grande per il campionamento diretto). Il campionamento della catena di Markov è stato usato in un numero innumerabile di modelli di fisica statistica; nel suddetto problema delle monete

nella scatola per scarpe, nel modello di Hubbard, ecc. ecc. Solitamente è però un sostituto molto povero, ed estremamente costoso di quello che vogliamo fare realmente.

È importante a questo punto fare attenzione a questi punti :

- Nel caso del gioco dei bambini, abbiamo bisogno soltanto di poche dozzine di ciottoli (i campioni) per avere un'idea fisica del valore di π , che sarà sufficiente nella maggior parte dei casi. Esattamente lo stesso è vero per la maggior parte dei più difficili problemi di fisica statica; poche dozzine di campioni (diretti) in grandi problemi del tipo moneta-nella-scatola, a qualunque densità, sarebbero sufficienti per risolvere controversie di vecchia data.
Similmente, poche dozzine di campioni diretti del modello di Hubbard di *fermionic correlation* ci darebbero delle conoscenze importanti nella superconduttività. Ancora, alcuni dei più grandi computer del mondo girano intorno lontani, giorno dopo giorno su problemi come quello menzionato. Tentano di costruire un ponte che colmi il divario esistente tra i miliardi di campioni delle catene di Markov e l'equivalente dei pochi lanci casuali nel gioco dei bambini.
- Solo dopo aver compreso il problema della generazione dei campioni indipendenti delle catene di Markov, potremmo cominciare a preoccuparci della convergenza lenta dei valori medi. Questo è già evidente nel gioco dei bambini, come in ogni misura nella fisica sperimentale : la precisione del valore numerico diminuisce soltanto come $1/\sqrt{N}$, dove N è il numero di misure indipendenti. Lasciatemi però sottolineare, in contrasto forse con il pensiero più diffuso, che la discussione sulla precisione assoluta di una misura non è poi così importante come potrebbe sembrare: nella valutazione della propria temperatura corporea, solitamente non è necessario valutare cinque cifre decimali significative per determinare se si ha la febbre oppure no. Nella pratica quotidiana Monte Carlo è di solito più critico nell'essere assolutamente sicuro che il programma ha dato dei campioni indipendenti¹ piuttosto che se di questi ve ne siano milioni o miliardi.
- È essenziale per capire che una catena di Markov lunga, anche se produce soltanto un numero piccolo di campioni indipendenti (e quindi un risultato molto approssimativo) è di solito estremamente sensibile a errori di programmazione (bugs), e a irregolarità varie, anche piccole, del programma.

1 che ha de-correlato dalla configurazione iniziale

5.2 - L'ALGORITMO DI TODDLERS E LA CONDITION OF DETAILED BALANCE

Dobbiamo ancora determinare ciò che dovrebbe fare la signora ferma al punto ψ nel gioco del eliporto. È una domanda difficile, e piena di conseguenze, e dunque non vogliamo dargli un consiglio sbagliato. Prima analizziamo un simile gioco discretizzato ; il noto puzzle mostrato nell'illustrazione 1,4.

4	2	8
1		7
6	3	5

Illustrazione 1,4: Questo noto puzzle è il punto di partenza per la nostra analisi teorica del metodo di Monte Carlo.

Questo puzzle ci consentirà dare un consiglio ottimale alla giocatrice ferma in ψ . Adesso lo scopo è creare un perfetto algoritmo di strapazzo che genera qualunque configurazione possibile del puzzle con la medesima probabilità. Un tale metodo, l'algoritmo dei bambini, è ben noto, ed è mostrato nell'illustrazione 1,5 insieme a uno dei suoi inventori.

Quanto visto sinora, ci consente di classificare senza esitazione questo metodo fra i metodi di campionamento diretto (come il gioco dei bambini appunto), poiché crea un caso indipendente ogni volta.

Adesso siamo però interessati all'algoritmo usato dalle persone adulte, che rispetti l'architettura del gioco.

Te cosa faresti ? Quasi certamente, terresti il puzzle nelle mani e, ad ogni passo temporale, muoveresti il quadrato vuoto in una direzione a caso. La condizione di equilibrio dettagliata che scopriremo in questo paragrafo, dimostrerà che quest'algoritmo, completamente naturale, è errato.



Illustrazione 1,5 : Questo è un algoritmo di campionamento diretto per il gioco di puzzle.

Se il quadrato vuoto è all'interno (come nell'illustrazione 1,4) allora l'algoritmo dovrebbe scegliere chiaramente una delle quattro direzioni possibili $\uparrow \rightarrow \downarrow \leftarrow$ con uguale probabilità, e muovere il quadrato vuoto.

Come nel gioco dell'eliporto, dobbiamo analizzare ciò che succede sul confine del quadrato¹.

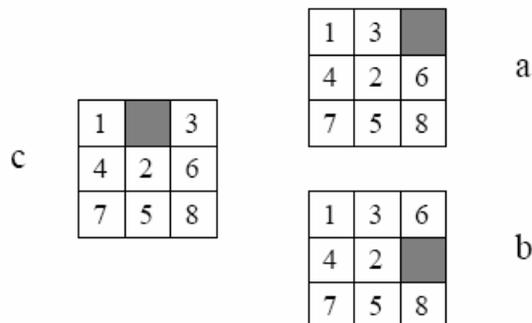


Figura 1,6 : La configurazione dell'angolo a è in contatto con le configurazioni b e c

Consideriamo la configurazione di un angolo a , il quale comunica con le configurazioni b e c , come mostrato nell'illustrazione 1,6. Se il nostro algoritmo (che deve ancora essere trovato) genera le configurazioni a , b e c con le probabilità $\pi(a)$, $\pi(b)$, $\pi(c)$, rispettivamente (e imponiamo loro di essere uguali), possiamo derivare un'equazione semplicemente legando il π alle probabilità di transizione $p(a \rightarrow b)$ ecc. Poiché a può essere generata solo da b , c , o da sé stessa, abbiamo :

$$\pi(a) = \pi(b)p(b \rightarrow a) + \pi(c)p(c \rightarrow a) + \pi(a)p(a \rightarrow a) \quad (1,5)$$

questo da :

$$\pi(a)[1 - p(a \rightarrow a)] = \pi(b)p(b \rightarrow a) + \pi(c)p(c \rightarrow a) \quad (1,6)$$

Scriviamo la condizione che ci dice che il quadrato vuoto, una volta in a , può stare in a o muoversi verso b o c :

$$1 = p(a \rightarrow a) + p(a \rightarrow b) + p(a \rightarrow c) \quad (1,7)$$

che da : $[1 - p(a \rightarrow a)] = p(a \rightarrow b) + p(a \rightarrow c)$. Questa formula, introdotta nella (1,6) comporta :

$$\pi(a) \underbrace{p(a \rightarrow b) + p(a \rightarrow c)}_{= \pi(c)p(c \rightarrow a) + \pi(b)p(b \rightarrow a)} = \pi(c)p(c \rightarrow a) + \pi(b)p(b \rightarrow a) \quad (1,8)$$

Possiamo soddisfare certamente quest'equazione se paragoniamo i termini legati separatamente:

$$\pi(a)p(a \rightarrow b) = \pi(b)p(b \rightarrow a) \dots \quad (1,9)$$

Questa equazione è la celebre Condizione di Equilibrio Dettagliato.

Osserviamola per un momento, e poi passiamo alla sua prima applicazione nel nostro puzzle.

¹ Condizioni di confine periodiche non sono consentite.

Imponiamo probabilità uguali per tutte le configurazioni accessibili, come $p(a \rightarrow b) = p(b \rightarrow a)$, ecc., dove la sola maniera semplice di collegare l'algoritmo proposto per l'interno del quadrato col confine è postulare una probabilità uguale a $1/4$ per qualunque movimento possibile. Osservando la (1.7), vediamo che dobbiamo consentire una probabilità di $1/2$ per $p(a \rightarrow a)$.

La conseguenza di quest'analisi è che per strapazzare al massimo il puzzle, dobbiamo essere immobili con una probabilità $1/2$ negli angoli, e con la probabilità $1/4$ sui bordi.

Possiamo montare tutte le altre regole nell'algoritmo seguente :

- ad ogni passo di tempo $t=0,1,\dots$:
 - Scegliere uno delle quattro direzioni possibili $\uparrow \rightarrow \downarrow \leftarrow$ con uguale probabilità.
 - Muovere il quadrato vuoto nella direzione corrispondente se possibile. Altrimenti, restare dove siamo, ma comunque il tempo va avanti.

In verità ho presentato quest'algoritmo del puzzle nella speranza che tu non creda che sia vero. Questo esempio è utile solo per per portarci a esaminare la condizione di equilibrio dettagliata, e capirla meglio.

Ricordiamo che il gioco del puzzle è, per certi aspetti, una versione discretizzata del gioco dell'eliporto, e adesso possiamo rispondere per analogia alla domanda della signora posta nella posizione ψ .

Si noti che c'è già un ciottolo in questa posizione. La signora dovrebbe adesso prendere dalla sua borsetta due pietre, collocarne una sulla pietra già a terra, e usare l'altra per tentare un nuovo lancio. Se questa dovesse cadere ancora all'esterno del perimetro, dovrebbe ammucciare a terra un'altra pietra sulle precedenti ecc.

Se si osserva l'eliporto dopo che il gioco è andato avanti per un certo tempo, si noterà una strana disposizione dei ciottoli a terra (si veda l'illustrazione 1,7).

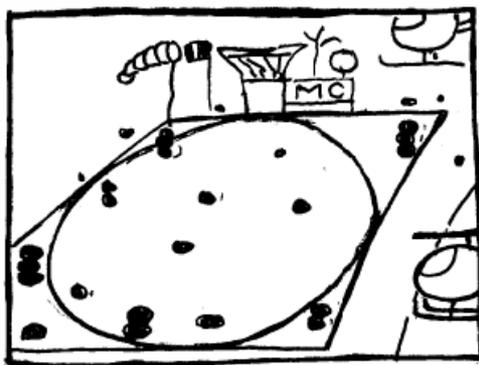


Illustrazione 1,7: Il piano di atterraggio dell'eliporto di Monte Carlo dopo che tutti i giocatori sono andati a casa. Nota i mucchi ai margini, particolarmente vicino agli angoli.

Ci sono delle pietre isolate soprattutto nell'interno, ed alcuni mucchi più o meno alti come ci si avvicina al bordo, soprattutto negli angoli. Questa è la maniera più standard per imporre una densità $\pi(x, y) = \text{costante}$.

A primo impatto questa disposizione dei ciottoli può generare incredulità. Infatti si può ritenere l'ammucchio di pietre assurdo e concludere che deve esserci un'errore di esposizione. Per dimostrare che non ho preso un abbaglio, posso soltanto mostrare le simulazioni (di signore che lanciano le pietre) con le casistiche che fanno o non fanno ammucciare i ciottoli.

5.2.1 - IL CONCETTO DI RIGETTO : L'ALGORITMO DI METROPOLIS

In effetti abbiamo visto appena il primo caso di Rigetto, che è una parola chiave dei metodi di Monte Carlo. Il concetto di rigetto è così fondamentale che vale la pena discuterne nel contesto completamente sterile di un campo d'aviazione di Monaco. I rigetti sono il metodo principale con cui Monte Carlo impone una densità corretta $\pi(x, y)$ sul quadrato, con un algoritmo (a lanci randomici) che non è particolarmente conforme con la geometria. I rigetti sono anche dissipatori (di ciottoli), e costosi (si lanciano ciottoli ma alla fine si resta fermi). Prossimamente approfondiremo decisamente la nostra comprensione dei rigetti.

Abbiamo introdotto i rigetti da un'ispezione diretta della condizione di equilibrio dettagliata. Questo trucco è stato elevato allo stato di metodo generale nel cosiddetto algoritmo di Metropolis. Quindi si è compreso che l'equazione della Condizione di Equilibrio Dettagliato (1,9) è verificata se si usa :

$$P(a \rightarrow b) = \min\left(1, \frac{\pi(b)}{\pi(a)}\right) \quad (1,10)$$

Questa equazione cosa significa ? Valutiamola nel caso dell'eliporto : stando in piedi nel punto a (che è nel quadrato, cioè $\pi(a)=1$), lanci il tuo ciottolo in una direzione a caso (verso b). Due cose possono succedere: Se b è nel quadrato ($\pi(b)=1$), l'equazione (1,10) dice di accettare il movimento con probabilità 1, se b è fuori ($\pi(b)=0$), dovresti accettarlo con probabilità 0, ovvero rigettare il movimento, e restare dove sei.

Dopo aver visto, tramite l'esempio come funziona l'algoritmo di Metropolis, potrebbe sorgere la curiosità di verificarne la correttezza nel caso generale. Per una volta, la burocrazia ci è di aiuto, infatti possiamo controllare il teorema tramite una procedura burocratica: riempiendo semplicemente il seguente modulo :

	$\pi(a) > \pi(b)$	$\pi(b) > \pi(a)$	
$P(a \rightarrow b)$			1
$\pi(a)P(a \rightarrow b)$			2
$P(a \rightarrow b)$			3
$\pi(a)P(a \rightarrow b)$			4

Riempendolo, si realizzerà per entrambe le colonne che la seconda e la quarta riga sono identiche, come stabilito dalla condizione di equilibrio dettagliata.

Questo è tutto quello che c'è da capire dell'algoritmo di Metropolis.

5.2.II - TO CRY OR TO CRAY !

L'algoritmo di Metropolis consente di trasformare una densità arbitraria di movimenti di prova (come i lanci randomici dei ciottoli) in una densità di probabilità $\pi(x, y)$.

Per cambiare la configurazione, discutiamo adesso come un modello di fisica generale, classico, statistico con un'energia alto-dimensionale arbitraria $E(x_1, x_2, \dots, x_N)$ viene simulato. In questo caso, la densità di probabilità è la misura di Boltzmann $\pi(x_1, \dots, x_N) = \exp(-\beta E)$, ed i valori di aspettativa fisici (le energie, i valori di aspettativa) sono dati dalle formule come la (1.4). Tutto quello che si deve fare (se il problema non è troppo complesso) è :

- Preparare una regola di spostamento (piuttosto arbitraria), che dovrebbe generalizzare i lanci randomici dei ciottoli del gioco di eliporto. Per esempio, si può andare da una configurazione iniziale x_1, x_2, \dots, x_N in una nuova configurazione scegliendo una dimensione arbitraria i , e facendo lo spostamento a caso $x_i \rightarrow x_i + \delta x$ con $-\epsilon < \delta x < +\epsilon$. Le variabili discrete sono trattate semplicemente.
- Avendo registrato le energie E^a e E^b con a il punto iniziale e con b il punto finale, potremmo usare l'algoritmo di Metropolis per calcolare la probabilità p , che effettivamente può andare così :

$$p(a \rightarrow b) = \min[1, \exp(-\beta(E^b - E^a))] \quad (1,11)$$

Questo è implementato usando un singolo numero casuale uniformemente distribuito $0 < \text{ran} < 1$. Muoviamo il nostro sistema verso b sotto la condizione $\text{ran} < p(a \rightarrow b)$, come mostrato in figura.

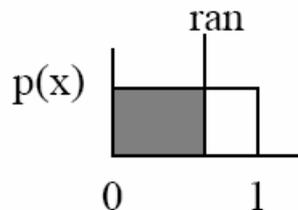


Illustrazione 1,8: La probabilità per un numero casuale uniformemente distribuito $0 < \text{ran} < 1$ sarà più piccola di $p(a \rightarrow b)$ e $\min[1, P(a \rightarrow b)]$. Un tipico programma Monte Carlo contiene almeno un confronto per ogni passo.

- Noterai che, per i sistemi continui, la libertà residua nel nostro algoritmo è l'aggiustamento del valore ϵ . La procedura "time-honored" consiste nello scegliere ϵ tale che circa la metà delle mosse vengano rigettate. La regola $\langle p(\epsilon) \rangle \sim 1/2$, come potrebbe essere chiamata, dovrebbe in generale assicurare una diffusione abbastanza veloce della particella attraverso lo spazio di configurazione. Se si vuole fare di meglio, si deve controllare la velocità della nostra diffusione di Monte Carlo, ma di solito non ne vale la pena.

Come presentato, l'algoritmo di Metropolis è estremamente potente, e molti problemi sono adeguatamente trattati con questa procedura. Il metodo è così potente che per molte persone la teoria di Monte Carlo si ferma dopo l'equazione (1,9). Tutto il resto sono dettagli di implementazione, raccolta dati, e regolazione di ϵ , che è stata appena discussa. Molti problemi, soprattutto a molte dimensioni, sfidano tuttavia questa regola.

I programmi scritti sulla linea del caso presentato sopra funzioneranno correttamente, ma, proprio per i problemi appena accennati, saranno tutti molto lenti generando campioni indipendenti.

Esempi prominenti di difficili problemi sono le transizioni di fasi nei modelli statistici generali; il modello di Hubbard, di Bosons, ed i sistemi disordinati. Il forte senso di frustrazione può essere evidenziato nel caso della sfera liquida dura, che è stato trattato nella pubblicazione del 1953 introducendo l'algoritmo di Metropolis, e che ha generato una serie di calcoli non convergenti, e infuocate controversie.

Un sistema molto utile per illustrare una difficile simulazione, mostrato nell'illustrazione 1,9, è la catena di N perle con energie z_0 e z_{N+1} fissate a zero, con z_1, \dots, z_N posizioni laterali dei punti, le variabili del sistema.

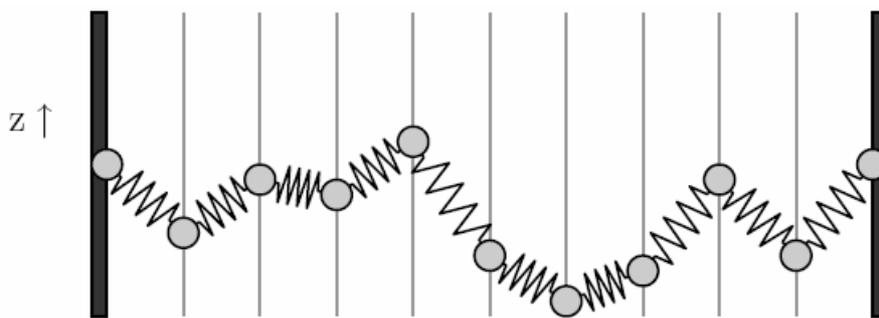


Illustrazione 1,9: La catena di perle accoppiate serve da esempio di un problema di Monte Carlo di grandi dimensioni.

$$E = \frac{(z_1 - z_0)^2}{2} + \frac{(z_2 - z_1)^2}{2} + \dots + \frac{(z_{N+1} - z_N)^2}{2} \quad (1,12)$$

Come generiamo delle distribuzioni secondo $\exp(-\beta E)$?

Un algoritmo è scrivibile molto facilmente. Secondo la ricetta data sopra, potresti semplicemente scegliere una perla a caso da un piccolo mucchio e spostarla. Un programma che implementa questo algoritmo può essere scritto in 10 linee. Tramite il programma si può tentare anche di ottimizzare il valore di ϵ (che uscirà tale che il cambiamento medio di energia in un solo movimento è $\beta |E^b - E^a| - 1$).

Quello che si scoprirà è che il programma funziona, ma è estremamente lento. È così lento che potrai scegliere solo se piangere oppure usare il super-computer Cray ("to Cry or to Cray"), ed in entrambi i casi sarebbe ridicolo.

Analizziamo perché un programma di questo tipo è così lento. La ragione è la seguente :

- La regola $\langle p(\epsilon) \rangle \sim 1/2$ stabilisce la misura del passo, che è necessariamente molto piccolo.
- La distribuzione della perla media, detta $z_{N/2}$, come espresso in unità di ϵ è molto grande, poiché la catena intera avrà un'estensione laterale dell'ordine di \sqrt{N} . È assolutamente essenziale realizzare che la distribuzione non può mai essere intrinsecamente larga, ma soltanto nelle unità di misura del passo imposto (che è una proprietà dell'algoritmo).
- È molto difficile il campionamento di una grande distribuzione con un passo piccolo, per capirsi, se l'eliporto è molto più grande della tua capacità di lancio, sarai costretto a portare una borsa di ciottoli molto grande per avere un'idea del valore numerico di π (si veda l'illustrazione 1,10).

La difficoltà di campionare una distribuzione grande è la ragione fondamentale perché le simulazioni possono avere difficoltà a convergere.

A questo punto, incontro spesso questa osservazione : perché non posso muovere indipendentemente diverse perle collegate, così da realizzare un grande movimento? Questa strategia è inutile. Nel gioco degli adulti, corrisponde a tentare, per ogni seconda pietra, lanciarla in una direzione a caso, raggiungerla, ovunque essa sia andata a finire (potrebbe essere necessario arrampicarsi sopra il recinto), raccoglierla e lanciarla nuovamente.

È chiaro che non si guadagna niente: abbiamo avuto già una probabilità di rigetto nel momento del primo lancio, legata alla probabilità che il lancio portasse il ciottolo al di fuori del quadrato. Non è difficile immaginare che a questa si deve sommare la probabilità che il secondo lancio ti porti fuori del quadrato, e considerato che è sufficiente che uno dei due lanci ti porti fuori del perimetro perché vi sia rigetto, ne deriva che la probabilità di avere un rigetto sarebbe proibitiva.

L'aumento della probabilità di rigetto sarà molto maggiore che il guadagno del passo.

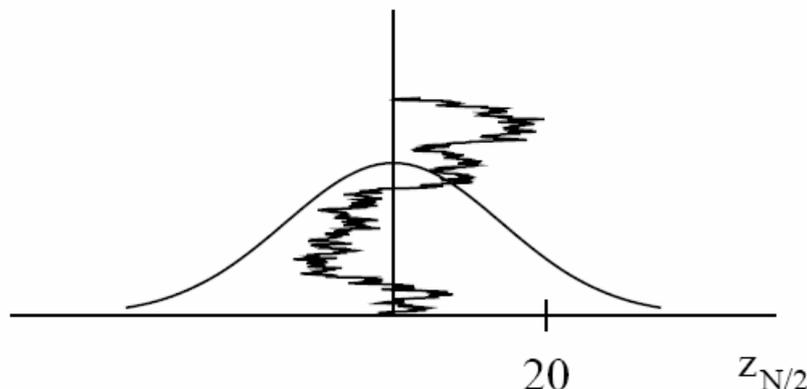


Illustrazione 1,10: Il problema della cordicella elastica con $N = 40$ è un problema abbastanza difficile perché la distribuzione di $z_{N/2}$ è grande. Il percorso di simulazione mostrato nella figura corrisponde a 400,000 passi.

Come configurato, il problema delle perle è difficile perché i numerosi gradi di libertà x_1, \dots, x_N sono fortemente accoppiati. Gli ϵ movimenti randomici sono una maniera estremamente povera di trattare tali sistemi. Senza una strategia migliore per i movimenti tentati, il programma non converge molto rapidamente e produce solo una manciata di campioni indipendenti.

Nel problema delle perle accoppiate, ci sono essenzialmente due maniere per migliorare la situazione.

La prima consiste nell'usare una trasformazione di base, in questo caso andando semplicemente nello spazio di Fourier (accelerazione di Fourier). Questo evidentemente riduce i gradi di libertà. Si potrebbero identificare i modi di Fourier che hanno una probabilità di essere sollecitati.

Un esercizio di valore insuperato è esaminare l'energia che contiene un piccolo accoppiamento anisotropo addizionale di perle e trattarlo con entrambi gli algoritmi. Fermi, Pasta ed Ulam, nel 1945, furono i primi a simulare il problema della catena anisotropa accoppiata su un computer (con un algoritmo deterministico) e a scoprire l'estremamente lenta esecuzione.

La trasformazione di base è un metodo specifico per consentire i grandi movimenti. Di solito, è comunque impossibile trovare tale trasformazione.

Un metodo più generale di trattare i problemi difficili interagenti consiste nell'isolare i sotto problemi che possono essere più o meno semplicemente campionati e risolti esattamente. Le informazioni a priori ottenute da questo lavoro analitico possono essere poi usate per proporre esattamente quei (grandi) movimenti che sono compatibili col sistema. I movimenti proposti sono poi resi corretti per mezzo di un algoritmo di Monte Carlo generalizzato. Un rigetto di Metropolis modificato, corregge la densità "progettata" nella vera probabilità stazionaria.

Motiveremo dapprima questo metodo molto importante nel caso dell'esempio delle catene accoppiate, quindi daremo la teoria generale, e presenteremo una domanda molto importante per generare i modelli.

Per capire realmente ciò che accade nel problema delle perle accoppiate, torniamo a vedere la figura 1,9, ed analizziamo solo una parte del gioco : il movimento di una specifica perla che identifichiamo con i , con $i-1$ ed $i+1$ (per il momento) immobilizzate ad alcuni valori z_{i-1} e a z_{i+1} . È facile fare questa simulazione e vedere che i risultati ottenuti sono dati come nell'illustrazione 1,11.

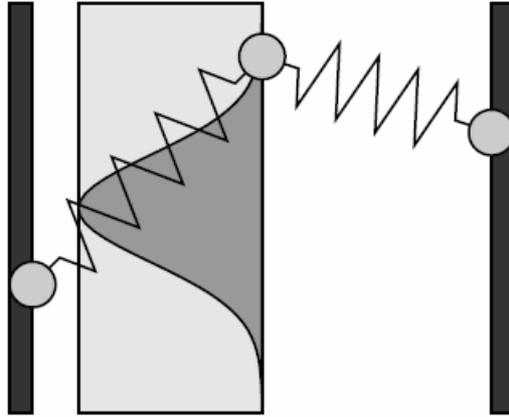


Illustrazione 1,11: Qui analizziamo il caso della simulazione del problema delle perle accoppiate, con z_{i-1} e z_{i+1} immobilizzate. Il grande tasso di rigetto è dovuto al fatto che la distribuzione di probabilità proposta per lo z_i (grigio chiaro) è molto differente dalla distribuzione accettata (il Gaussiano, grigio scuro).

Sei vede che la funzione di distribuzione $P(z_i)$, (con z_{i-1} e z_{i+1} ferme) è un Gaussiano concentrato intorno a $z_i = (z_{i-1} + z_{i+1})/2$.

Notare comunque che ci sono infatti due distribuzioni: quella accettata e quella rigettata. È l'algoritmo di Metropolis che, attraverso i rigetti, modifica la proposta distribuzione piatta nella corretta Gaussiana.

Adesso vediamo che i rigetti, visti inizialmente come una noia, giocano invece il ruolo molto costruttivo di modificare i movimenti proposti nella densità di probabilità corretta.

C'è un'intera letteratura di ricerca sull'uso dei metodi di rigetto nel campionamento di distribuzioni 1-dimensionali ...

Un argomento che lasceremo immediatamente perché siamo più interessati ai casi alto-dimensionali.

5.2.III - UNA PROBABILITÀ A PRIORI

Estendiamo adesso le formule della condizione di equilibrio dettagliata e dell'algorithmo di Metropolis, tenendo conto delle possibili scelte "a priori" dei movimenti, che sono descritte da una probabilità a priori $A(a \rightarrow b)$ di tentare il movimento. Nel gioco dell'eliporto, questa probabilità era semplicemente :

$$A(a \rightarrow b) = \begin{cases} 1 & \text{if } |a-b| < \epsilon \\ 0 & \text{altrimenti} \end{cases} \quad (1,13)$$

Con ϵ il range del lancio, del quale non abbiamo sinora notato la presenza.

Nell'esempio delle perle, la probabilità di scegliere una perla i , e di muoverla di una piccola quantità $-\epsilon < \delta < \epsilon$ era indipendente sia da i sia dalla posizione attuale z_i .

Adesso, rivisitiamo l'equazione di equilibrio dettagliata, ed esplicitiamo l'algorithmo: la probabilità $p(a \rightarrow b)$ è divisa in due parti separate :

$$p(a \rightarrow b) = A(a \rightarrow b) P(a \rightarrow b) \quad (1,14)$$

Dove $P(a \rightarrow b)$ è la probabilità di accettazione del movimento proposto con $A(a \rightarrow b)$.

Cos'è questa probabilità di rigetto ?

Questa è molto facile da ottenere dall'equazione di pieno equilibrio

$$\pi(a) A(a \rightarrow b) P(a \rightarrow b) = \pi(b) A(b \rightarrow a) P(b \rightarrow a) \quad (1,15)$$

Si può vedere, che per qualunque probabilità a priori, cioè per qualunque algorithmo di Monte Carlo potremmo trovare il tasso di accettazione che è necessario per portare questa probabilità conformemente alla nostra inalterata condizione di equilibrio dettagliata. Come prima, possiamo usare un algorithmo di Metropolis per ottenere (un possibile) tasso di accettazione corretto

$$\frac{P(a \rightarrow b)}{P(b \rightarrow a)} = \frac{\pi(b)}{A(a \rightarrow b)} \frac{A(b \rightarrow a)}{\pi(a)} \quad (1,16)$$

che risulta

$$P(a \rightarrow b) = \min \left(1, \frac{\pi(b)}{A(a \rightarrow b)} \frac{A(b \rightarrow a)}{\pi(a)} \right) \quad (1,17)$$

Evidentemente, questa formula si riduce alla prescrizione originale di Metropolis se introduciamo una probabilità a priori piatta (come nell'equazione (1,13)). Siccome sta in piedi, nell'equazione (1,17) dichiara la libertà fondamentale algoritmica che abbiamo configurando i nostri algoritmi di Monte Carlo : Qualunque pregiudizio possibile $A(a \rightarrow b)$ può essere fatto in un valido algorithmo di Monte Carlo poiché possiamo correggerlo sempre col tasso di accettazione corrispondente $P(a \rightarrow b)$. Certo, soltanto una probabilità scelta molto attentamente sarà utilizzabile, o anche superiore alla scelta semplice e popolare nell'equazione (1,10).

L'inclusione di una probabilità a priori generale è matematicamente innocua, ma genera un cambiamento profondo nell'organizzazione pratica della nostra simulazione.

Per valutare la probabilità di accettazione, $P(a \rightarrow b)$, nell'equazione (1,17) proponiamo non solo il movimento verso b , ma necessitiamo anche delle valutazioni esplicite sia di $A(a \rightarrow b)$ che del movimento di ritorno $A(b \rightarrow a)$.

Cosa è cambiato nell'equazione (1,17) in riferimento alla precedente l'equazione (1,10) ?

Prima, avevamo necessariamente una grande probabilità di rigetto quando partivamo da un punto a con l'alta probabilità (grande $\pi(a)$) verso un punto b con una probabilità bassa (piccolo $\pi(b)$). L'algoritmo di Metropolis ingenuo potrebbe produrre la corretta densità di probabilità soltanto installando i rigetti. Adesso possiamo neutralizzarli, scegliendo semplicemente una probabilità a priori che è anche molto più piccola. Infatti, puoi vedere facilmente che c'è una scelta ottima: noi potremmo essere in grado di usare una probabilità a priori $A(a \rightarrow b)$ con densità di probabilità $\pi(b)$ ed $A(b \rightarrow a)$ con $\pi(a)$. In quel caso, il rapporto espresso nell'equazione (1,17) sarà sempre uguale a 1, e non ci sarà mai un rigetto. Certo però c'è da dire che siamo tornati al campionamento diretto, da dove ci eravamo allontanati in quanto ritenuto troppo difficoltoso.

La discussione non è circolare, come potrebbe apparire, perché può sempre essere applicata solo a una parte del sistema.

Per capire questo punto, è meglio ritornare all'esempio delle perle. Sappiamo che la distribuzione di probabilità $\pi(z_i)$ per z_{i-1} e z_{i+1} immobili è

$$\pi(z_i | z_{i-1}, z_{i+1}) \sim \exp[-\beta(z_i - \bar{z}_i)^2] \quad (1,18)$$

Con $\bar{z}_i = (z_{i-1} + z_{i+1})/2$.

Possiamo usare adesso esattamente questa formula come una probabilità a priori $A(z_i | z_{i-1}, z_{i+1})$ e generare un algoritmo senza rigetti, che muove la perla i a ogni passo, con le sue vicine. Per programmare questa regola, abbiamo bisogno dei numeri casuali Gaussiani. Fino ad ora, comunque, il beneficio della nostra operazione è essenzialmente inesistente¹.

È adesso possibile estendere la formula per lo z_i a z_{i-1} e z_{i+1} immobili, sino a una più grande finestra.

Integrando tra z_{i-1} e z_{i+1} in $\pi(z_{i-2}, \dots, z_{i+2})$, troviamo che

$$\pi(z_i | z_{i-1}, z_{i+1}) \sim \exp[-2\beta(z_i - \bar{z}_i)^2] \quad (1,19)$$

Dove, adesso, $\bar{z}_i = (z_{i-2} + z_{i+2})/2$. Avendo determinato z_i da un campionamento dell'equazione (1,19), possiamo trovare in seguito dei valori per z_{i-1} e z_{i+1} usando l'equazione (1,18). Il risultato netto di questo è che possiamo aggiornare z_{i-1} e z_{i+1} simultaneamente. Questa procedura viene chiamata "costruzione di Lévy" e genera grandi movimenti con una misura di finestra arbitraria, senza rigetti.

¹ Un algoritmo con una probabilità a priori è chiamato algoritmo di "heatbath". È popolare nei modelli di rotazione, ma essenzialmente identico al metodo di Metropolis

5.2.IV - PERTURBAZIONI

Da questo semplice esempio del problema delle perle accoppiate, si possono apprendere rapidamente tutte le sottigliezze del metodo Monte Carlo.

Si vede che possiamo produrre un algoritmo perfetto, perché la probabilità a priori $A(z_i|z_{i-1}, z_{i+1})$ potrebbe essere scelta uguale alla probabilità statica $\pi(z_{i-1}, \dots, z_{i+1})$ azzerando il tasso di rigetto.

Questo, comunque, è solo un caso felice. Il potenziale enorme della probabilità a priori risiede nel fatto che l'equazione (1,17) si deteriora (di solito) molto graziosamente quando A e $\pi(z)$ differiscono. Un modo per capire questo punto consiste nel rivisitare l'algoritmo per il problema della catena accoppiata, aggiungendo un piccolo termine di perturbazione all'energia, come

$$E_1 = \gamma \sum_{i=1}^N f(z_i) \quad (1,20)$$

il quale è supposto essere relativamente meno importante dell'energia elastica.

È interessante vedere la maniera in cui il metodo si adatta se consideriamo il sopra descritto algoritmo della costruzione di Lévy¹.

Possiamo adesso affermare che :

- L'energia di ogni configurazione è adesso $\tilde{E}(z_1, \dots, z_N) = E_0 + E_1$, dove $E_0(a)$ è il termine dato dall'equazione (1,12), che è in qualche maniera neutralizzata dalla probabilità a priori $A(a \rightarrow b) = \exp[-\beta E_0(b)]$. Si può scrivere adesso il tasso di Accettazione di Metropolis del processo dall'equazione (1,17). Il risultato è

$$P(a \rightarrow b) = \min \left(1, \frac{\exp[-\beta E_1(b)]}{\exp[-\beta E_1(a)]} \right) \quad (1,21)$$

Questo è esattamente l'equazione dell'algoritmo di Metropolis ingenuo, (1,10), esclusivamente col termine di energia recentemente aggiunto.

- Implementando la probabilità a priori con $\gamma=0$, l'algoritmo gira con la probabilità di accettazione 1, indipendente dalla misura dell'intervallo $2l+1$. Se si include il secondo termine E_1 , si dovrà ottimizzare ancora questa misura. Certamente troverai che la finestra ottima corrisponde a una misura tipica di $\beta|(E_1^b - E_1^a)| \sim 1$.
- Con questo partizionamento del sistema in una parte che possiamo risolvere esattamente, e in un termine addizionale, scoprirai che l'algoritmo di Monte Carlo ha l'apparenza di un metodo di perturbazione. Certamente sarà sempre esatto. Tipicamente ha tutte le probabilità di essere veloce se E_1 è più piccolo di E_0 . Una principale differenza con i metodi di perturbazione è che farà sempre il campionamento della distribuzione piena perturbata $\pi(z_1, \dots, z_N) = \exp[-\beta \tilde{E}]$.

1 in Monte Carlo Quantum, si introduce un piccolo accoppiamento tra diverse cordicelle.

Uno può spendere un tempo sostanzialmente infinito meditando sulle probabilità a priori, e sulle somiglianze e differenze con la teoria della perturbazione.

In questo punto è dove il metodo Monte Carlo mente.

Questo è quello che uno deve capire prima di avventurarsi in massa nell'inseguire potenti algoritmi futuri per i problemi che oggi sono ritenuti essere irraggiungibili. Programmare un algoritmo semplice per gli oscillatori accoppiati sarà un'introduzione eccellente nell'argomento. In letteratura il problema accoppiato è mostrato come una delle applicazioni più diffuse dei metodi Monte Carlo Quantum, e dove alcune delle limitazioni dello stesso approccio vengono evidenziate.

In tutti i casi, dovremmo capire che un grande tasso di rigetti è sempre indicativo del fatto che la struttura dei movimenti proposti è inadatta alla densità di probabilità del modello alla temperatura data. Il beneficio della soluzione di questo problema, ammesso di riuscire a trovarla, è di solito colossale : fare movimenti piccoli con tasso di rigetto trascurabile ci consente spesso di fare movimenti più grandi, ed esplorare più rapidamente le regioni importanti di tutto lo spazio di configurazione.

Per concludere questo capitolo, vedremo un altro esempio : i metodi di gruppo celebrati nei sistemi di griglia, che sono stati introdotti dieci anni fa da Swendsen e da Wang e da Wolff.

Qui ancora, troviamo i due ingredienti essenziali degli algoritmi lenti : movimenti necessariamente piccoli, ed una grande distribuzione delle quantità fisiche. Usando il concetto di probabilità a priori, possiamo capire molto facilmente questi metodi.

4.2.v - ALGORITMI DI GRUPPO

Discuteremo questi metodi in una maniera che rivela chiaramente l'aspetto di "ingegneria" della probabilità a priori, dove si tenta di far fronte al problema di una grande distribuzione. Prima di fare questo, discutiamo, come prima, il caso generale e l'origine fisica della convergenza lenta. Come sappiamo, il modello di Ising è definito come un sistema di rotazioni su una griglia con un'energia:

$$E = - \sum_{\langle i, j \rangle} S_i S_j \quad (1,22)$$

Dove le rotazioni possono intraprendere dei valori di $S_i = \pm 1$, e dove la somma investe un paio di punti di griglia vicini.

Un algoritmo semplice è il seguente: si sceglie una rotazione, e tenta di ruotarlo a caso. La probabilità a priori è piatta, poiché la rotazione da eseguire è scelta arbitrariamente.

Usando questo schema, si potrà recuperare molto facilmente la transizione di fase tra una fase paramagnetica ed una ferromagnetica, che in due dimensioni ha luogo a una temperatura di $\beta = \log(1 + \sqrt{2})/2$.

Si nota anche che il programma è in aumento lento intorno alla temperatura critica. Questo è interpretato di solito come l'effetto della divergenza della lunghezza di correlazione quando ci avviciniamo al punto critico. Nei nostri termini, capiremo ugualmente bene questo rallentamento: il nostro semplice algoritmo cambia la magnetizzazione al massimo di un valore 2 a ogni passo, poiché l'algoritmo compie soltanto delle singole rotazioni. Questo valore discreto sostituisce ϵ nel nostro precedente esempio delle perle accoppiate.

Se adesso stampiamo degli istogrammi della magnetizzazione totale del sistema (con unità di misura $\Delta m = 2!$), vediamo ancora che questa distribuzione diventa "grande" quando ci avviciniamo a β_c . (si veda l'illustrazione 1,12). Chiaramente, la magnetizzazione totale ha un'ampia distribuzione, che è estremamente difficile da campionare con un algoritmo di rotazione singola per colpo.

Per apprezzare gli algoritmi di gruppo, si deve capire due cose:

- Come nel problema accoppiato, non possiamo semplicemente compiere diverse rotazioni simultaneamente (si veda la discussione a pagina 20), vogliamo colpire grandi gruppi, ma certo non possiamo semplicemente saldare insieme tutte le rotazioni di un segno che sono collegate l'un l'altra, perché queste non potrebbero mai essere prese separatamente.
- Se non possiamo fissare solidamente le rotazioni dello stesso segno con probabilità 1, dobbiamo scegliere le rotazioni adiacenti allineate con una certa probabilità p . Questa probabilità p è il parametro di costruzione della nostra probabilità a priori A .

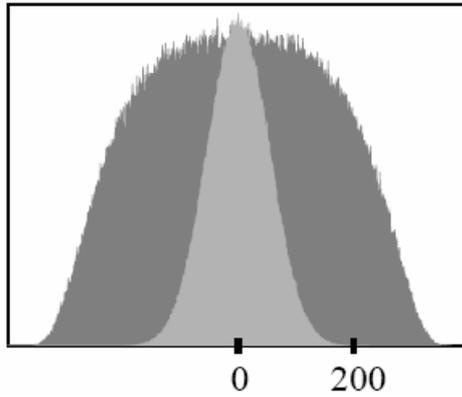


Illustrazione 1,12: Gli istogrammi della magnetizzazione nel modello di Ising 20×20 alle temperature 10% e 45% al di sopra del punto critico. Sull'asse x è stampato $m/\Delta m$, dove $\Delta m=2$ è il cambiamento della magnetizzazione che può essere ottenuto in un solo passo di Monte Carlo.

L'algorithm girerà per un valore arbitrario di p ($p = 0$ corrisponde all'algorithm a singolo spin-flip), ma il p che minimizza i rigetti sarà ottimale.

Abbiamo visto che l'algorithm di gruppo inizia con l'idea di scegliere un punto di partenza arbitrario, ed aggiunge sorte di collegamenti con probabilità p . Arriviamo quindi al primo esempio non banale della valutazione di una probabilità a priori. Consideriamo l'illustrazione 1,13. Immaginiamo di cominciare da un "+" ruotando nell'area grigia di configurazione a) ed aggiungiamo le rotazioni.

Cosa sono le probabilità a priori $A(a \rightarrow b)$ e la probabilità inversa $A(b \rightarrow a)$, e cosa sono i pesi di Boltzmann $\pi(a)$ e $\pi(b)$?

$A(a \rightarrow b)$ è data da un termine riguardante il collegamento interno "+ +", $A_{int}(a \rightarrow b)$, che vediamo difficilmente, e che neanche tentiamo di valutare, è una parte che riguarda il confine del gruppo. Questo confine è generato da due tipi di collegamenti, come riassumiamo sotto :

int	ext	number
+	-	c_1
+	+	c_2

$$E|_{\partial C} = -c_2 + c_1 \quad (1,23)$$

(Nell'esempio dell'illustrazione 1,13, abbiamo $c_1=10$ e $c_2=14$). La probabilità a priori è $A(a \rightarrow b) = A_{int} \times (1-p)^{c_2}$. Per valutare il peso di Boltzmann, ci asteniamo anche da valutare i termini che non cambiano tra a) e b).

Chiaramente abbiamo bisogno solo dell'energia di confine, che è data dall'equazione (1,23). Segue che $\pi(a) \sim \exp[-\beta(c_1 - c_2)]$.

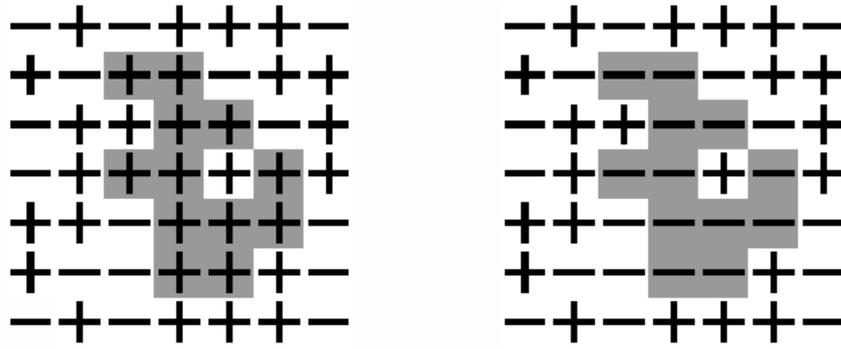


Illustrazione 1,13: il processo di costruzione dell' algoritmo di Swendsen-Wang. Cominciare da un punto iniziale +, includiamo gli altri punti + con probabilità p (sinistro). L'intero gruppo (l'area grigia) è quindi invertito (flipped). Nel testo, calcoliamo la probabilità di fermare la costruzione del gruppo grigio, e del movimento inverso. Questo produce le nostre probabilità a priori.

Consideriamo adesso il movimento inverso. Nel gruppo b), i collegamenti attraverso il confine sono come segue :

int	ext	number
-	-	c_1
-	+	c_2

$$E|_{\partial C} = -c_1 + c_2 \quad (1,24)$$

La probabilità a priori di costruire questo gruppo è composta ancora da una parte interna, che è esattamente la stessa come per il gruppo in a), ed una parte di confine, che è cambiata : $A(a \rightarrow b) = A_{int} \times (1-p)^{c_1}$. Similarmente, troviamo che $\pi(a) \sim \exp[-\beta(c_2 - c_1)]$. È adesso sufficiente mettere tutto nella formula dell'equilibrio dettagliato

$$e^{-\beta|c_1 - c_2|} (1-p)^{c_2} P(a \rightarrow b) = e^{-\beta|c_2 - c_1|} (1-p)^{c_1} P(b \rightarrow a) \quad (1,25)$$

che risulta nella probabilità di accettazione:

$$P(a \rightarrow b) = \min \left(1, \frac{e^{-\beta|c_2 - c_1|} (1-p)^{c_1}}{e^{-\beta|c_1 - c_2|} (1-p)^{c_2}} \right) \quad (1,26)$$

Il punto più importante di quest'equazione non è che può essere semplificata, come vedremo tra un minuto, ma che è perfettamente esplicita, e potrebbe essere valutata senza problemi : una volta che il suo gruppo è costruito, potremmo valutare semplicemente c_1 e c_2 e calcolare la probabilità di rigetto da quest'equazione.

Da un'ispezione più stretta della formula (1,26), vediamo che, per $1-p = \exp[-2\beta]$, la probabilità di accettazione è sempre 1. Questo è il valore "magico" implementato negli algoritmi di Swendsen-Wang e di Wolff.

L'algoritmo risultante è molto semplice e segue esattamente la descrizione data sopra : scegliamo una rotazione a caso, e aggiungiamo collegamenti tra punti vicini con probabilità p , la costruzione si ferma quando nessuno dei collegamenti che attraversano il gruppo in costruzione deve essere aggiunto.

Stupisce vedere come questo algoritmo passi il punto di Curie senza alcun rallentamento percettibile.

4.2.VI - OSSERVAZIONI SULL'EQUILIBRIO MONTE CARLO

Arriviamo alla fine dell'introduzione ai metodi di equilibrio Monte Carlo. Spero di aver dato un'introduzione comprensibile alla via che il metodo rappresenta in molti contesti di fisica statistica:

- L'approccio (equilibrato) di Monte Carlo è un metodo di integrazione che converge lentamente, ma certamente. Eccetto in pochi casi, uno è sempre obbligato a campionare la densità fissa (la misura di Boltzmann, la matrice nel caso di densità quantum) da un approccio con catena di Markov. La questione critica è il tempo di correlazione, che può diventare astronomico. Nelle applicazioni usuali, si è contenti quando si ottengono un numero anche molto piccolo di campioni veramente indipendenti, ma numero spaventosamente grande di calcoli non capita mai nella decorrelazione di una catena di Markov dalla condizione iniziale.
- I metodi regolari lavorano bene, ma hanno alcuni svantaggi importanti. Come presentato, la condizione che il tasso di rigetto deve essere abbastanza piccolo, tipicamente dell'ordine di $1/2$, ci riduce a movimenti molto piccoli.
- Il tasso di accettazione ha delle conseguenze importanti per la velocità del programma, ma un tasso di accettazione piccolo è in particolare un indicatore che i movimenti proposti sono inadeguati.
- Sarebbe errato dare l'impressione che la situazione può essere sempre migliorata - a volte uno è semplicemente obbligato a fare calcoli molto grandi. In molti casi comunque, una scelta giudiziosa della probabilità a priori consente di ottenere tassi di accettazione molto buoni, anche per grandi movimenti. Questo lavoro è importante, e conduce frequentemente a un aumento esponenziale di efficienza.

5 - Esempi concreti

5.1 - MONTE CARLO GREZZO

Occupiamoci ora di come sia possibile stimare una quantità incognita attraverso il metodo statistico di Monte Carlo.

Nel nostro caso si tratterà della stima di un integrale. Va detto che molto spesso altri metodi numerici di risoluzione sono più consigliabili ed efficaci; ciò nondimeno, per integrali multidimensionali il Monte Carlo può risultare competitivo o, addirittura, la scelta più opportuna quando il confine che delimita il dominio di integrazione è complicato, l'integrando non esibisce picchi concentrati in regioni molto ristrette e, infine, non si pretende un'accuratezza molto elevata.

Cominciamo con un semplice integrale:

$$I = \int_a^b f(x) dx$$

dove la $f(x)$ è una funzione continua sull'intervallo di integrazione $[a,b]$.

Riscriviamo la precedente come:

$$I = \int_a^b f(x) \frac{p(x)}{p(x)} dx$$

dove $p(x)$ è una qualsiasi funzione densità di probabilità definita in $[a,b]$.

Sia X una variabile casuale con densità di probabilità $p(x)$ e sia G la variabile casuale definita da:

$$G = \frac{f(x)}{p(x)}$$

Poiché, per la definizione di valore atteso, si ha:

$$\bar{X} = \int_a^b x p(x) dx$$

nel nostro caso ne consegue che:

$$\bar{G} = \int_a^b \frac{f(x)}{p(x)} p(x) dx = I$$

La valutazione di un integrale diventa, pertanto, un problema risolvibile in chiave statistica, stimandone il valore mediante l'esame di un opportuno campione.

Detta g una generica realizzazione della variabile aleatoria G , la stima dell'integrale I diventa:

$$I = \bar{g} = \frac{1}{N} \sum_{i=1}^N g_i = \frac{1}{N} \sum_{i=1}^N \frac{f(x_i)}{p(x_i)}$$

Siffatto metodo, in cui un integrale è rappresentato come un valore medio, è detto Monte Carlo Grezzo.

5.2 - STIMA DELL'ATTESA DI UN AUTOBUS

Per scegliere a caso un elemento da un universo descritto dalla funzione di densità $f(x)$, si procede nel modo seguente:

- si traccia il grafico della funzione cumulativa ;
- si genera un numero casuale compreso tra 0 e 1 ;
- si riporta il numero così trovato sull'asse y e lo si proietta orizzontalmente sulla curva $y = F(x)$;
- il valore di x così trovato sulla curva si assume come uno dei valori del campione di x .

A questo punto, le procedure di elaborazione dei dati così ottenuti variano a seconda del tipo di simulazione.

Problema: si calcoli il tempo medio di attesa di un utente alla fermata dell'autobus, sapendo che:

- gli autobus passano ogni 15 minuti;
- gli arrivi degli autobus sono distribuiti normalmente (secondo una distribuzione gaussiana) con scarto quadratico medio di 3 minuti;
- gli utenti arrivano alla fermata secondo una distribuzione di Poisson con frequenza media di 4 utenti all'ora e i posti vuoti sugli autobus sono pure distribuiti secondo una poissoniana con media 3/2;
- non sono ammessi viaggiatori in piedi.

	Numero di posti liberi sull'autobus	Tempi di arrivo degli autobus (min.)	Assegnazione dei passeggeri sull'autobus	Tempi di arrivo degli utenti (min.)	Intervalli tra gli arrivi degli utenti (min.)	Tempi di attesa degli utenti (min.)
1	3	17	autobus 1	8	8	9
2	2	28	autobus 1	13	5	4
3	1	46	autobus 2	18	4	10
4	1	58	autobus 2	22	5	6
5	1	75	autobus 4	49	27	9
6	2	96	autobus 5	58	9	17
7	2	111	autobus 6	71	13	25
8	0	122	autobus 6	91	20	5
9	1	130	autobus 11	133	42	23
10	0	148	autobus 11	139	6	17
11	2	156	autobus 12	154	15	31
12	3	185	autobus 13	189	35	3
13	3	192	autobus 14	207	18	2
14	1	209	autobus 15	216	9	10
15	2	226	autobus 15	223	7	3
16	1	241	autobus 17	247	24	5
17	3	252	autobus 23	341	94	7
18	2	267	autobus 24	350	9	9
19	4	286	autobus 25	352	2	20
20	2	304	autobus 25	369	17	3

	Numero di posti liberi sull'autobus	Tempi di arrivo degli autobus (min.)	Assegnazione dei passeggeri sull'autobus	Tempi di arrivo degli utenti (min.)	Intervalli tra gli arrivi degli utenti (min.)	Tempi di attesa degli utenti (min.)
21	2	313	autobus 27	401	32	3
22	0	332		402	1	-
23	3	348		403	1	-
24	1	359		416	13	-
25	3	372		425	9	-
26	1	388				
27	1	404				

La situazione è stata osservata per un periodo di circa 8 ore. Se ne è ricavata questa tabella. Le prime due colonne si riferiscono al numero di posti liberi sugli autobus ed ai rispettivi tempi di arrivo e sono state ricavate col metodo di Montecarlo dalle distribuzioni indicate nel problema.

La quinta colonna indica gli intervalli fra successivi arrivi di utenti alla fermata ed è stata ricavata col metodo di Montecarlo dalla distribuzione esponenziale e^{-4t} . Il tempo è espresso in minuti, con riferimento al momento in cui ha avuto inizio il processo di campionamento simulato. Una volta noti i tempi di arrivo degli autobus e degli utenti, l'assegnazione dei passeggeri agli autobus diventa automatica. Si applica l'ovvio criterio che i passeggeri salgono sull'autobus se sono presenti alla fermata nel momento in cui arriva l'autobus e se vi sono posti liberi sullo stesso. Avendo assegnato i passeggeri agli autobus, i tempi di attesa (sesta colonna) degli utenti si ottengono sottraendo il tempo di arrivo di ogni passeggero dal tempo di arrivo del suo autobus. Nella tabella sono stati indicati i dati relativi a 21 passeggeri, per un totale di 221 minuti di attesa. Pertanto il tempo medio di attesa di un passeggero è di 10,5 minuti.

Il metodo di Montecarlo presenta vari vantaggi rispetto alla semplice osservazione ed elencazione degli arrivi, dei tempi di attesa, delle lunghezze delle code, ecc. Innanzitutto, il campionamento simulato può produrre dati relativi a simulazioni di eventi per tempi anche molto lunghi (mesi o anni) in pochi minuti, specialmente se si utilizza un calcolatore. Un altro vantaggio è rappresentato dal fatto che è possibile modificare a piacimento i parametri che determinano l'andamento della simulazione: ad es., nella simulazione si può aggiungere un ulteriore centro di servizio per vedere come evolve la situazione, senza doverlo installare nella realtà, eliminando così l'impiego di tempo e risorse che si sarebbe dovuto affrontare nel caso la modifica fosse stata compiuta materialmente.

5.3 - APPLICAZIONI

Supponiamo per esempio che un amico ci chieda qual è la probabilità di estrarre da un mazzo di carte napoletane una carta maggiore di 4 e contemporaneamente di lanciare un dado e ottenere un numero minore di 2. Noi sappiamo soltanto che ogni carta ha la stessa probabilità di essere estratta ($1/40$) e così ogni faccia del dado ($1/6$), ma non abbiamo nessuna nozione di teoria delle probabilità e non sappiamo rispondere al quesito.

In questo caso possiamo usare il metodo Montecarlo: con un generatore di numeri casuali possiamo simulare l'evento "estrazione + lancio" un numero molto elevato di volte e calcolare così la probabilità richiesta.

Tuttavia in questo caso semplice l'applicazione del metodo Montecarlo è uno spreco di energia: è possibile fare un semplice calcolo per rispondere al quesito, oppure ripetere l'esperimento "estrazione + lancio" un numero di volte grande quanto vogliamo per calcolare "sperimentalmente" la probabilità richiesta.

5.4 - APPLICAZIONE DIDATTICA DEL METODO MONTE CARLO

E' il momento di passare ora all'esame di una serie di proposte e di suggerimenti concreti sul modo di applicare i concetti fin qui sviluppati, nella didattica della matematica dell'incerto.

A tal fine è stato realizzato, e allegato alla presente relazione (Mcarlo.zip), un pacchetto di 22 programmi, che serviranno a simulare fenomeni aleatori di vario genere, dei quali illustriamo la natura e diamo qualche suggerimento sul loro uso didattico.

Si partirà da problemi dei quali è già nota la soluzione, per lo più di tipo combinatorio (lancio di monete, dadi, ecc.) Nel corso della simulazione si fisserà il numero di prove che si desidera effettuare. Sarà bene iniziare con un piccolo numero di prove, per poi passare ad un numero sempre più grande, in modo da verificare la tendenza dei dati a stabilizzarsi intorno a determinati valori, che sono poi quelli previsti dalla teoria.

Si affronteranno poi problemi di difficile soluzione, almeno per il grado di conoscenze degli alunni (in alcuni casi, anche rispetto al grado generale di conoscenze). In tal modo si potrà apprezzare l'uso del computer che, per mezzo di prove ripetute, ci offrirà almeno una misura approssimata della probabilità che ha il fenomeno di verificarsi. In alcuni esempi il risultato approssimato appare in tempo reale durante l'esperimento stesso, permettendo di verificare direttamente la progressiva stabilizzazione delle sue cifre, all'aumentare del numero di prove. In tal modo l'allievo sarà stimolato a dare egli stesso, nel corso dell'esperimento, una sua personale valutazione circa il valore che si otterrà al termine. Il raffronto dei valori stimati con quelli teorici, previsti dal calcolo delle probabilità (che in genere sarà comunicato alla fine), sarà molto più efficace di ogni discorso, circa la natura dei fenomeni aleatori e le stime sulla loro probabilità di verificarsi.

Il lavoro terminerà con alcuni esempi in cui o non è nota la soluzione dei problemi proposti, oppure si conosce la procedura per arrivare ai risultati, ma essa è talmente

complessa da renderli praticamente inaccessibili. Si resterà talvolta un po' sorpresi del grado di difficoltà presentato da alcuni problemi, vista la semplicità della loro enunciazione. In ogni caso il computer ci fornirà la soluzione, o meglio, un suo valore stimato, ma attendibile. Siamo consapevoli del fatto che si tratta di una soluzione approssimata, ma sappiamo ormai, grazie alle esperienze precedenti, che la sua approssimazione rispetto al valore cercato dipenderà solo dal numero di prove fatte.

I pochi esempi proposti saranno sufficienti a far intuire le numerose possibilità di applicazione dei metodi proposti.

5.4.I - LANCIO DI DUE DADI

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Si lanciano due dadi e si vuole determinare la probabilità dei vari esiti, costituiti dalle possibili somme che si possono ottenere. Il problema è ben noto e la sua soluzione è semplicissima. Basta costruire una tabella a doppia entrata con i 36 esiti delle somme dei due dadi. Le somme vanno da un minimo di 2 ad un massimo di 12. Le probabilità dei vari esiti sono date dai rapporti tra i rispettivi casi favorevoli (1, 2, 3, 4, 5, 6, 5, 4, 3, 2, 1) e il totale dei 36 casi possibili. Il computer serve come primo approccio all'uso dei metodi statistici, in quanto permette di determinare facilmente le *frequenze relative*, al variare del numero di lanci.

Una volta introdotto il numero di lanci, il programma procede a calcolare le **frequenze relative** (rapporto tra il numero dei casi in cui si verifica un esito rispetto al numero totale di prove effettuate), fornendo al termine una tabella con i risultati ed i raffronti rispetto ai relativi **valori teorici**. E' consigliabile eseguire più volte il programma, partendo da un piccolo numero di lanci.

Si eseguono poi prove con un numero di lanci sempre maggiore. La tabella, che compare ogni volta, fornirà gli opportuni raffronti fra le frequenze relative ottenute con il computer e le rispettive probabilità.

5.4.II - LANCIO DI N MONETE

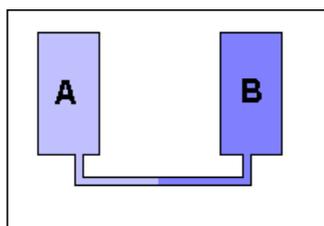


Si lanciano n monete e poi si conta il numero di teste. I possibili esiti vanno da un minimo di zero teste, fino ad un massimo di n . Per consentire la visualizzazione dei risultati sullo schermo, il numero n non deve superare 10. Le probabilità di ogni esito si possono calcolare per mezzo della combinatoria. I casi favorevoli sono dati dai coefficienti binomiali

$\binom{n}{k}$ con n numero monete lanciate e k numero di teste. I casi possibili sono 2^n .

Il computer provvederà ad effettuare i lanci, trovando le frequenze relative dei vari esiti e le rispettive probabilità.

5.4.III - MODELLO DI EHRENFEST

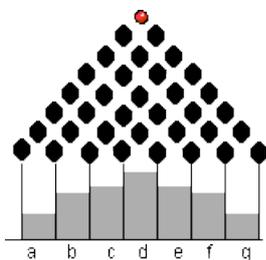


Nel 1907 i coniugi **Paul e Tatiana Ehrenfest** pubblicarono su "*Physik Zeitschrift*" un articolo, proponendo il seguente modello probabilistico per spiegare il fenomeno della diffusione dei gas. Si collegano tra loro due recipienti *A* e *B*, contenenti molecole di gas a differente pressione.

Si suppone che, ad ogni istante, una molecola scelta a caso passi da un recipiente all'altro. Naturalmente è più probabile che la molecola sia scelta tra quelle del recipiente che ne ha di più. Così facendo però, diminuiscono le molecole nel primo recipiente e quindi si riduce la probabilità che esse passino nel secondo recipiente. In quest'ultimo si verifica il fenomeno opposto. Si stabilisce dopo un po' un equilibrio dinamico in cui i due recipienti hanno la stessa pressione e contengono all'incirca lo stesso numero di molecole, mentre continuano a transitare dall'uno all'altro.

La simulazione al computer prevede due urne (recipienti). L'operatore assegna a ciascuna di esse un certo numero di biglie (molecole di gas). Il programma estrae una biglia a caso. La biglia estratta, passa dalla propria urna all'altra. Si continua così, fino a quando il numero di biglie nelle urne è uguale, cosa che si verifica invariabilmente, anche quando le quantità di biglie, inizialmente assegnate alle due urne, erano sensibilmente diverse tra loro.

5.4.IV - SIMULIAMO UNA PASSEGGIATA A CASO: IL QUINCUNX DI GALTON



Si deve a **John Galton** un modello per la simulazione di una passeggiata a caso. Una biglia cade lungo un piano inclinato, sul quale sono posti ostacoli che la obbligano a deviare a destra o a sinistra, in modo del tutto casuale. Continuando a scendere, la biglia supererà tutti gli ostacoli e finirà per cadere in una delle celle poste al termine dei vari percorsi. Naturalmente non tutte le celle hanno la stessa

probabilità di essere raggiunte dalla biglia che rotola. Il problema ha una stretta analogia con quello del lancio di n monete. Basta riflettere sul fatto che la biglia, nel seguire i vari percorsi, si comporta come se ogni volta lanciasse una moneta: testa, per andare a destra e croce, per andare a sinistra. Per calcolare le probabilità di passare attraverso i varchi esistenti tra gli ostacoli, si possono quindi utilizzare i *coefficienti binomiali* già visti nel lancio delle monete, oppure calcolare le successive righe del triangolo di Pascal. Sapendo che la somma dei termini di ogni riga vale 2^n , sarà facile calcolare le probabilità e raffrontarle con le frequenze relative.

Il programma realizzato al computer realizza il movimento delle palline, che cadono dall'alto e, passando attraverso le varie porte, arrivano in una delle celle poste in basso, mentre un grafico ad istogrammi evidenzia le frequenze delle biglie cadute in ciascuna di esse. Al termine sarà possibile visualizzare la solita tabella con il raffronto tra i dati statistici e quelli teorici.

5.4.v - IL PROBLEMA DI MÉRÉ

E' un problema classico nella storia della Matematica. Fu proprio **Pascal** (contemporaneamente a **Pierre de Fermat**) che, fornendo la risposta a tale problema propostogli dal Cavaliere di Méré, diede origine al calcolo delle probabilità. Il quesito posto era il seguente: "Come mai l'uscita del 6, lanciando 4 dadi, non ha la stessa probabilità dell'uscita di una coppia di 6, lanciando per 24 volte due dadi?". Il Cavaliere di Méré riteneva che l'uscita del 6 su un dado (probabilità 1/6) andasse moltiplicata per 4, lanciando 4, dadi ed assegnava perciò il valore 2/3 alla probabilità dell'evento. Allo stesso modo, riteneva che la probabilità di uscita di una coppia di 6, lanciando due dadi (1/36), andasse moltiplicata per 24, se si lanciavano per 24 volte i due dadi. Si otteneva ancora 2/3. I dati ricavati dall'esperienza tuttavia non erano in accordo con i supposti valori teorici. Il primo caso si presentava, infatti, con frequenza leggermente maggiore del secondo¹.

I calcoli del Cavaliere erano basati su alcuni ragionamenti errati. Prima di affrontare il problema, vogliamo provare a verificare se i dati statistici confermano il sospetto che il secondo tipo di gioco sia meno favorevole del primo. E' ciò che il computer s'incaricherà di fare. Eseguendo il programma per un gran numero di volte, si conferma che l'uscita di una coppia di sei, con 24 coppie di dadi è un po' meno favorevole dell'uscita del sei, con 4 dadi. In effetti, l'uscita di almeno un 6 nel lancio di 4 dadi ha probabilità $1 - (5/6)^4$ che vale circa 0,5177, mentre l'uscita di almeno una coppia di 6 lanciando per 24 volte due dadi ha invece probabilità di $1 - (35/36)^{24}$ circa 0.4914.

5.4.vi - PROBLEMA DEI COMPLEANNI

Se si scopre che, in un gruppo di una trentina di persone, due di esse festeggiano il compleanno nello stesso giorno si resta sempre un po' sorpresi. Se vi fossero 366 persone (29 febbraio a parte), tale fatto sarebbe certo e non desterebbe alcuna meraviglia. E' tuttavia sorprendente sapere che bastano solo 22 persone, affinché l'evento di due persone con lo stesso compleanno abbia una probabilità superiore al 50 %. Con 40 persone o con 56, le probabilità di coincidenza arrivano rispettivamente a più del 90 % e più del 99 %. Tali valori si determinano calcolando, per n persone il valore dell'espressione:

$$1 - \left(\frac{364}{365}\right) \cdot \left(\frac{363}{365}\right) \cdot \dots \cdot \left(\frac{365-n+1}{365}\right)$$

Il programma si limita ad intervistare persone, fino a quando non ne trova due con lo stesso compleanno, indicando poi il numero di persone intervistate, che raramente supera la trentina. Facendo ripetere più volte agli allievi il programma, si può ricavare il valore medio di persone da intervistare per avere una coincidenza nei compleanni.

¹ - E' curioso osservare che il Cavaliere di Méré, precorrendo i tempi, si avvaleva in pratica del metodo Monte Carlo per confrontare i suoi dati teorici (sbagliati) con le frequenze relative derivanti da un *gran numero di* prove da lui eseguite sui due eventi descritti nel problema. Probabilmente era propenso a dare maggiore fiducia al metodo sperimentale. La domanda formulata al matematico Pascal, gli serviva soltanto per capire il motivo della discordanza tra teoria e prassi.

5.4.VII - SELEZIONE DARWINIANA

Molte specie di animali (pesci, anfibi, insetti, ecc.) mettono al mondo un altissimo numero di figli. La maggior parte di essi è destinata ad essere eliminata. La fame, le malattie, i predatori ed altri fattori agiscono a caso sulla popolazione, eliminando gli individui in eccesso. Darwin presumeva che gli individui eliminati avessero qualche caratteristica che determinava sin dall'inizio una minore probabilità di sopravvivenza.

Il modello molto semplificato da noi proposto prevede un numero iniziale di 1000 individui che si ridurranno a 50. Gli individui sono divisi in quattro gruppi, con diverse probabilità di sopravvivenza. Le probabilità andranno assegnate prima di iniziare l'esperimento. Non occorre che le probabilità differiscano di molto tra loro. Anzi è interessante il caso in cui esse sono piuttosto vicine, ad esempio, 0.54 - 0.51 - 0.50 - 0.48 . Inizia quindi il processo di selezione, al termine del quale si vedrà quanti sono gli individui sopravvissuti in ciascuno dei quattro gruppi iniziali. Il maggior numero di sopravvissuti apparterranno quasi certamente al gruppo più favorito. Si verificheranno casi di totale estinzione dei ceppi di individui meno favoriti, pur in presenza di piccolissime differenze tra i vari gruppi. Se le probabilità di sopravvivenza iniziali sono sensibilmente diverse tra i vari gruppi, è quasi certo che sopravviveranno alla fine solo individui del gruppo più favorito. E' bene verificare tutto ciò ripetendo più volte l'esperimento e variando ogni volta le probabilità di sopravvivenza.

5.4.VIII - IL PROBLEMA DELL'ESTINZIONE

Un individuo di una certa specie vivente (ma anche una particella dell'atomo in un processo di fissione) può generare altri individui, ma può anche non riprodursi per niente. Supponiamo di conoscere le probabilità p_0 p_1 p_2 di generare 0, 1 o 2 discendenti. Qual è la probabilità che in un futuro più o meno prossimo si arrivi all'estinzione della discendenza?

Il problema è piuttosto complesso. Fu **Galton** a trovarne la soluzione. Se si suppone che un individuo possa generare al massimo due discendenti, l'equazione risolutiva sarà:

$$p_0 + p_1 x + p_2 x^2 = x$$

Se però il numero di massimo di figli aumenta fino ad un valore n , si avrà un'equazione completa di grado n , la cui risoluzione diventa molto più ardua. Non resta che usare allora il metodo Monte Carlo.

Fu proprio per questo motivo che a Los Alamos, su suggerimento di Stanislaw Ulam, gli scienziati che studiavano i processi di diffusione dei protoni, iniziarono ad usare tale metodo.

Nella nostra simulazione fisseremo a 2 il numero massimo di figli, per raffrontare i dati ottenuti dall'equazione di secondo grado con quelli ottenuti statisticamente, attraverso un congruo numero di prove.

5.4.IX - SVILUPPO DI POPOLAZIONI

Il problema sembra semplice: qual è la probabilità che, dopo un certo numero g di generazioni, una popolazione, sviluppatasi a partire da un solo individuo, sia costituita esattamente da k individui? Per un numero un po' grande di generazioni il problema raggiunge un tale livello di complessità, da non consentire altra possibilità che quella di cercare di ottenere il dato richiesto per mezzo di prove ripetute.

I parametri da introdurre per la simulazione sono:

$P_0 P_1 P_2$ probabilità di generare rispettivamente 0,1,2 figli.

g numero di generazioni da considerare (consigliabile un numero non troppo grande)

d numero di discendenti che si vogliono riscontrare alla g -esima generazione.

Si procede quindi alla simulazione del processo di sviluppo. Il risultato finale sarà la frequenza relativa dei casi in cui sono stati riscontrati esattamente d discendenti alla g -ma generazione, rispetto al numero di prove fatte. Il dato ottenuto non è raffrontabile con il dato teorico, che ... è praticamente impossibile da calcolare!

5.4.X - GUADAGNO DI UN GIOCATORE

Un giocatore ha una certa probabilità p di vincere ad un determinato gioco. Se effettua un gran numero di partite, potrà trovarsi in vincita, oppure in perdita (raramente, alla pari). Se la probabilità di vincita è pari a $1/2$, i due eventi (vincita o perdita) potranno verificarsi indifferentemente. Ma se p è anche di poco superiore o inferiore a $1/2$, allora andrà incontro rispettivamente a grandi vincite o a grandi perdite. Esse saranno tanto più grandi, quanto maggiore è il numero di partite giocate. Tutti i giochi d'azzardo danno al Banco un sia pur lieve vantaggio rispetto al giocatore. Questi è quindi destinato quasi certamente a perdere. La perdita sarà addirittura certa se deciderà di giocare un gran numero di partite. La durata del gioco dipenderà in ogni caso dall'importo del capitale iniziale, che potrà assicurargli un certo numero di partite, prima di trovarsi con le tasche vuote. Il programma è piuttosto semplice. Dopo aver introdotto il valore p della probabilità di vincita, s'inizia a giocare e sullo schermo appare il grafico del guadagno (o della perdita):

E' interessante fare più prove, assegnando inizialmente a p il valore 0,5. Si prova poi ad aumentare o a diminuire di poco la probabilità di vincita, ad esempio 0,52 e 0,48. Mentre nel primo caso il guadagno oscilla in genere intorno all'asse delle ascisse, negli altri due casi tenderà ad impennarsi rispettivamente verso l'alto (guadagno) e verso il basso (perdita).

5.4.XI - PROBLEMA DEL COLLEZIONISTA

Un bambino decide di fare una collezione di figurine. All'inizio tutte le figurine trovano posto nelle rispettive caselle dell'album. Ben presto però, il bambino scopre, nel sistemare le figurine, che alcune caselle dell'album risultano già occupate. Continuando a comprare figurine, capiteranno sempre più *doppioni* mentre le poche figurine mancanti sembreranno proprio non voler uscire.

Quante figurine dovrà acquistare in media ogni bimbo per completare il suo album? Il problema è stato già risolto. Per m caselle da riempire occorrono:

$$m \cdot \left(1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{m} \right)$$

La formula è valida nel solo caso in cui le figurine siano messe nelle bustine in modo casuale. Valori troppo diversi farebbero supporre che il fabbricante abbia reso alcune figurine *più rare* di altre, per vendere una maggiore quantità di bustine. La simulazione dimostra che il trucco non è necessario. I bambini saranno costretti ad acquistare un bel numero di figurine, se vorranno completare la collezione. Proviamo a simulare una collezione, da ripetere per n volte, e vediamo quale valore medio si otterrà. Il risultato sarà, come al solito, vicino al dato teorico, se si sceglierà di fare molte prove. Si potrà sottolineare che la formula risolutiva non è molto facile da scoprire e che il metodo Monte Carlo, oltre che fornirci un risultato statistico, ci permette di valutare il grado di attendibilità di eventuali soluzioni proposte.

5.4.XII - IL PROBLEMA DELLE CONCORDANZE

Supponiamo di avere un pacchetto di m cartellini numerati su una delle facce. Disponiamoli con la faccia nascosta e mescoliamoli accuratamente. Poi li scopriamo uno la volta, mentre contemporaneamente contiamo ad alta voce:

1, 2, 3, ..., m

Quanto siamo disposti a scommettere che riusciremo a scoprire tutti i cartellini senza che *nessuno di essi coincida con il numero che viene pronunciato*? Prima di fare la scommessa sarebbe bene sapere che probabilità di vittoria abbiamo. E' possibile dimostrare che, per un congruo numero di cartellini, tale probabilità è uguale al reciproco del numero naturale e o numero di Nepero 2.718281..., $p = 1/e = 0.367879...$

Tale risultato è sorprendente. Chi avrebbe mai pensato che dietro un giochetto con i cartellini numerati si celasse uno dei più noti numeri irrazionali? Per convincersi che non si tratta del frutto della mera fantasia di qualche matematico, non resta che cercare conferma con un buon numero di prove.

5.4.XIII - IL GIOCO DEL CRAPS

Il Craps è un gioco di dadi abbastanza diffuso. Le regole sono un po' complesse.

C'è un primo lancio di due dadi, durante il quale il Giocatore vince se esce il 7 oppure l'11 e perde se esce il 2 o il 3 o il 12;

Se non vince né il Giocatore né il Banco, allora il numero uscito viene detto *Punto* e il gioco continua.

Il Giocatore vince, se esce il Punto e perde, se esce il 7.

Nella prima fase, il Giocatore ha qualche chance in più rispetto al banco; infatti, l'uscita del 7 o dell'11 ha una probabilità pari a 8/36, mentre quella del 2 o del 3 o del 12 è soltanto 4/36. E' tuttavia più probabile che non vinca nessuno dei due con una probabilità di 24/36. Quindi, nella maggior parte dei casi, si passa alla seconda fase del gioco.

E' ora il Banco ad avere più probabilità, da un massimo di 6/9 ad un minimo di 6/11.

Valutando tutte le possibilità del gioco sin dall'inizio, con un grafo ad albero (piuttosto elaborato), la situazione è la seguente: il Banco ha 251/495 (<~ 0.507) probabilità di vittoria contro 244/495 (~ 0.493) del Giocatore.

Quasi tutti i giochi d'azzardo si basano su piccolissime differenze nelle probabilità di vittoria, ovviamente a favore del banco, in modo da allettare i giocatori. Il Grafico del guadagno visto in precedenza ci assicura che il Giocatore è destinato a perdere, soprattutto se giocherà a lungo, allettato da qualche probabile vincita iniziale.

5.4.XIV - I CERINI DI BANACH

Successe proprio al matematico **Banach** di riflettere su una situazione derivata dalla sua abitudine di portare sempre con sé due scatole di cerini. Quando aveva bisogno di un cerino, lo prendeva da una delle scatole. Capitava, ad un certo momento, che la scatola da lui scelta si rivelava vuota. Era perciò costretto a prendere il cerino dalla seconda scatola se, fortunatamente, ne conteneva ancora qualcuno. Ecco nato il problema: quanti cerini restano in una delle scatole, quando l'altra viene trovata vuota¹? La soluzione è piuttosto complessa e si conosce attualmente una formula che fornisce un valore leggermente approssimato, per un numero c di cerini:

$$\frac{(2c+1)}{\sqrt{\pi c}} - 1$$

1 - Attenzione, il problema cambia se si suppone di andare a contare i fiammiferi dell'altra scatola, subito dopo aver acceso l'ultimo fiammifero di una scatola. Nel nostro caso si suppone che la scatola vuota sia rimessa in tasca insieme all'altra e, solo quando si cercherà di cavare da essa ancora un fiammifero, ci si accorgerà che è vuota. La formula si riferisce a tale situazione.

5.4.xv - IL GIOCO DELL'UN-DUE-TRE

Si tratta ancora di un gioco sulle coincidenze, sul quale si basa un solitario con le carte. Si prende un mazzo di 40 carte napoletane. Dopo averle ben mescolate si scoprono una la volta e contemporaneamente si conta:

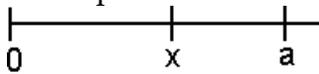
1, 2, 3, 1, 2, 3,...

Se non si verificano coincidenze tra la carta scoperta e il numero che viene detto, si guadagna un punto. Quanti punti si ottengono mediamente in n partite? Si tratta di un problema aperto, di cui non si conosce a tutt'oggi la risposta. Non resta quindi che affidarci al metodo Monte Carlo, facendo un congruo numero di simulazioni.

Questa volta non abbiamo un valore teorico della probabilità con il quale confrontare quello stimato mediante il metodo Monte Carlo. Il calcolo delle probabilità ci assicura che bastano circa 38500 prove, per avere un valore della frequenza relativa le cui prime due cifre decimali si possono ritenere *esatte*, con una probabilità del 95%.

5.4.xvi - DISTRIBUZIONE UNIFORME

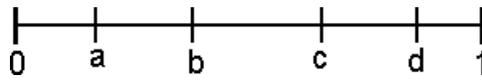
Finora abbiamo operato con fenomeni di tipo *discreto*, in cui i singoli elementi dell'insieme di eventi sono *distinguibili* l'uno dall'altro, come le biglie di un'urna, e in numero *finito*. Passiamo ora ad occuparci di eventi in cui lo spazio campione è costituito da *numeri reali* o *coppie di numeri reali*. Per i numeri reali, così come per i punti di una retta o di un piano, non è possibile *contare* gli elementi contenuti in un intervallo. La combinatoria, alla quale è spesso possibile far ricorso per il calcolo delle probabilità negli insiemi discreti, non può esserci più d'aiuto. Occorre pertanto far uso di altri concetti e di altre tecniche di calcolo.

 Esaminiamo una variabile uniforme nell'intervallo $[0, a]$, che indicheremo con aR . R è un *numero random* compreso tra 0 ad 1, quindi il prodotto aR potrà assumere solo valori compresi tra 0 ed a . Se x è un numero compreso tra 0 ed a , la probabilità che un numero scelto a caso sia compreso nell'intervallo $[0, x]$ è data dal rapporto x/a . Si può dare un'interpretazione geometrica (vedi figura a lato) al rapporto x/a : esso corrisponde al rapporto tra le lunghezze tra due segmenti, uno di lunghezza x ed uno di lunghezza a .

Scegliamo perciò due intervalli $[0, a]$ e $[0, x]$ con $x < a$ e prendiamo a caso dei numeri nell'intervallo $[0, a]$.

Qual è la probabilità che essi appartengano all'intervallo $[0, x]$? Come già detto, tale probabilità corrisponde al rapporto tra le lunghezze dei rispettivi segmenti. Il dato sarà facilmente confermato dalla simulazione.

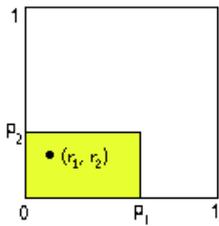
5.4.XVII - PROBABILITÀ DI $(a \leq R \leq b)$ E DI $(c \leq R \leq d)$ IN n ESTRAZIONI DI R



Se prendiamo due intervalli $[a,b]$ e $[c,d]$ della stessa lunghezza all'interno dell'intervallo $[0,1]$, la probabilità che un numero a caso appartenga all'intervallo $[a,b]$ oppure che appartenga all'intervallo $[c,d]$ è esattamente la stessa. Valgono, infatti, le considerazioni sulla distribuzione uniforme, fatte nell'esempio precedente.

Proviamo a verificarlo, con il computer. Per un numero di prove abbastanza grande le due frequenze relative tenderanno a coincidere.

5.4.XVIII - PUNTI A CASO IN UN QUADRATO UNITARIO



Finora abbiamo scelto un numero a caso aR . Ciò corrisponde, da un punto di vista geometrico a scegliere un punto a caso all'interno di un segmento di lunghezza a . Proviamo ora a scegliere ora due numeri a caso, r_1 e r_2 nell'intervallo $[0,1]$. La coppia (r_1, r_2) individua un punto a caso all'interno di un quadrato unitario. Se ora fissiamo due numeri p_1

e p_2 , ciascuno compreso nell'intervallo $[0,1]$, la probabilità che il punto di coordinate (r_1, r_2) cada nel rettangolo di lati p_1 e p_2 risulta uguale al prodotto $p_1 p_2$, cioè all'area del rettangolo.

Facciamo una simulazione, per verificare quanto è stato detto.

5.4.XIX - AREA DELLA SUPERFICIE SOTTO LA RETTA $Y=MX$



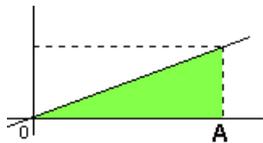
L'esempio precedente permette di affermare che la probabilità p che un punto appartenga ad una superficie piana S_1 racchiusa all'interno di un'altra superficie piana S_2 è data dal rapporto fra le corrispondenti aree. Viceversa, nota la probabilità p , possiamo affermare che l'area di S_1 è uguale al 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 $\frac{f}{n}$.

$$S_1 = \left(\frac{f}{n}\right) \cdot S_2$$

Su quest'ultima considerazione che si basa il metodo Monte Carlo per il calcolo di aree.

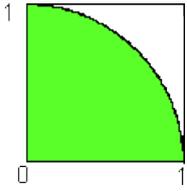


Per cominciare, consideriamo il caso molto semplice dell'area della superficie compresa tra la retta $y=mx$ e l'asse delle ascisse, fino ad un punto $A=(a,0)$ sull'asse delle ascisse. Il calcolo di tale area è molto facile: è la metà dell'area del rettangolo avente per base il segmento a che va dall'origine al punto A e per altezza l'ordinata ma della funzione, quindi è pari alla metà del prodotto $a \cdot ma$, cioè: $\frac{ma^2}{2}$

L'area stimata con il metodo Monte Carlo, è molto vicina a quella calcolata geometricamente. Certamente, è facile trovare l'area del triangolo, anche senza ricorrere al metodo Monte Carlo. L'esempio è stato scelto solo per un primo approccio con la stima di aree per mezzo di metodi statistici.

Il metodo funziona ovviamente anche per stimare aree di funzioni più complesse, per le quali si richiede il calcolo integrale. Alcune funzioni non sono facilmente integrabili, perciò il calcolo delle loro aree potrebbe essere molto difficile, o in taluni casi addirittura impossibile. A questo punto è meglio una stima, sia pure un po' approssimata ... che nulla! La situazione è simile a quella che si ha quando occorre trovare il volume di un sasso. In mancanza di formule geometriche, si usa il rapporto tra peso e peso specifico. Si ottiene una misura che non è forse quella esatta, ma è comunque una sua buona approssimazione.

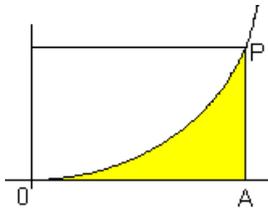
5.4.xx - UN METODO STATISTICO PER IL CALCOLO DI π



Con centro in O e raggio 1 tracciamo un arco di circonferenza nel quadrato unitario. La probabilità P che un punto scelto a caso nel quadrato appartenga al settore è, come abbiamo visto, pari all'area del settore, cioè $\pi/4$. Moltiplicando per 4 si ottiene quindi π . Anche la frequenza relativa (numero di punti presi a caso e che appartengono al settore, diviso per il numero di prove ripetute), moltiplicata per 4, darà un valore approssimato di π .

Il programma fa apparire sullo schermo sia la figura sia i punti scelti a caso. L'uso della grafica rallenta alquanto l'esecuzione, anche se risulta di un certo effetto. Togliendo la grafica il programma diventa molto più veloce ed è possibile fare un maggior numero di prove in tempi accettabili, per migliorare l'approssimazione dei risultati.

5.4.xxi - AREA SOTTO LA PARABOLA $y=x^2$



Gli alunni di scuola media o del biennio delle superiori hanno sufficienti conoscenze di geometria, per calcolare direttamente l'area del triangolo sotto la retta o quella di un settore circolare di 90° . Se si passa invece ad aree di superfici più complesse, le loro conoscenze si rivelano insufficienti.

Quando si ricorre al metodo statistico, non vi è invece alcuna differenza, dal punto di vista concettuale, tra l'area del triangolo sotto la retta e quella della superficie sotto la parabola o sotto l'iperbole.

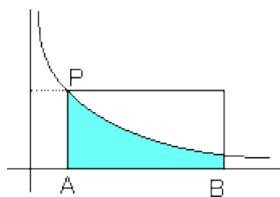
Consideriamo ad esempio la parabola definita dalla funzione $y=x^2$

Proviamo a trovare l'area delle superficie compresa tra la curva e l'asse delle ascisse, fino ad un punto $A=(a,0)$, così come abbiamo fatto nell'esempio 22. Occorrerà tracciare un rettangolo che ha per base il segmento OA di lunghezza a e per altezza la misura a^2 , corrispondente all'ordinata di P. L'area del rettangolo sarà data quindi dal prodotto $a \cdot a^2$, cioè a^3 .

L'area della superficie sotto la parabola inscritta nel rettangolo è uguale al prodotto della probabilità p che un punto scelto a caso all'interno del rettangolo appartenga alla superficie sotto la parabola, moltiplicata per l'area del rettangolo. Sostituendo la probabilità con la frequenza relativa, si troverà, per un gran numero di prove, un valore approssimato dell'area. Per un gran numero di prove, la frequenza relativa, tende al valore $1/3$. Perciò è facile arrivare a concludere che l'area della superficie sotto la parabola è $1/3$ dell'area del rettangolo circoscritto.

Una volta ottenuta l'area sotto la parabola, si può trovare quella dentro il segmento parabolico, che si trova per differenza tra l'area del rettangolo ad esso circoscritto e quella della superficie sotto la parabola. Si ottiene così la formula per trovare l'area di un segmento parabolico, che è $2/3$ dell'area del rettangolo circoscritto.

5.4.XXII - AREA SOTTO L'IPERBOLE $y = 1/x$



Occorre definire questa volta un intervallo $[a, b]$ sull'asse delle ascisse, che individua una superficie sotto l'iperbole, nel tratto che va dal punto $A = (a, 0)$ al punto $B = (b, 0)$.

Sappiamo, dal calcolo integrale, che l'area di tale superficie è pari a $\log(b) - \log(a)$, cioè $\log(b/a)$. Se $a=1$, l'area diventa semplicemente $\log(b)$.

Tutto ciò ovviamente può essere presentato ad allievi che non sanno ancora nulla di calcolo integrale, facendo semplicemente uso del metodo statistico. Solo in seguito, avendo ottenuto una sufficiente dimestichezza con le aree delle superfici sotto l'iperbole, si potrà cogliere qualche interessante spunto per introdurre elementi non banali del calcolo dei logaritmi.

Per stimare l'area col metodo Monte Carlo, si traccia un rettangolo che ha per base il segmento lungo $b-a$ sull'asse delle ascisse e per altezza l'ordinata $(1/a)$ del punto P di ascissa a .

L'area del rettangolo sarà quindi: $\frac{b-a}{a}$

Tale area, moltiplicata per la frequenza relativa $\frac{f}{n}$ dei punti che appartengono alla superficie sotto l'iperbole, rispetto al numero di prove, darà il valore approssimato della relativa area.

5.5 - MODELLI NON PARAMETRICI

Gli approcci non parametrici tentano di superare queste problematiche, esse sono la simulazione storica e la simulazione Montecarlo.

Simulazione storica

Le variazioni del valore del portafoglio sono ottenute applicando allo stesso le variazioni storiche dei fattori di rischio/prezzi titoli. I valori così ottenuti opportunamente ordinati costituiscono la distribuzione dei rendimenti di portafoglio in un'ottica "what if". L'obiettivo è simulare la perdita potenziale del portafoglio detenuto in base alle condizioni di rischio passate.