Caso, Probabilità e Complessità

Angelo VULPIANI

Dipartimento di Fisica, Università "Sapienza", Roma, Italia
<angelo.vulpiani@roma1.infn.it>

Malaga Novembre 2015

Perchè questo seminario?

Lo scopo di una conferenza non è tanto quello di fornire soluzioni pronte dei problemi, quanto porre i problemi e dare lo stimolo per la loro risoluzione.
L. Boltzmann

La teoria della probabilità non è in fondo che il buon senso ridotto a calcolo: essa fa apprezzare con precisione ci che gli spiriti giusti sentono per una sorta di istinto, senza che essi possano, sovente, rendersene conto.
P.S. Laplace

Perché vale la pena sapere un po' probabilità

La probabilità è una scienza giovane, la sua (prei)storia inizia solo nel XVI-mo secolo nel frivolo mondo dei giochi (dadi, carte, scommesse); ha poi avuto un ruolo fondamentale nella **meccanica statistica** e più recentemente ha avuto intersezioni con il **caos deterministico** e la **complessità**.

Probabilità nella **vita di tutti i giorni**:
* gioco del lotto, poker etc;
* investimenti in borsa;
* analisi mediche.

Probabilità nelle **scienze** e nella **tecnologia**:
* chimica, ecologia;
* navicelle spaziali.

Purtroppo, benché la probabilità sia molto importante sia a livello pratico che concettuale, è (relativamente) poco studiata e quasi assente nella trattazioni divulgativa.

Opinioni (non sempre concordi) di persone importanti:

* Diceria maligna (del passato)
Tra la probabilità e la matematica c'è la stessa relazione che intercorre tra il mercato nero e l'economia.

* K. Popper:
Le stime probabilistiche **non** *sono falsificabili. E, naturalmente non sono neppure verificabili ...*
Inoltre Popper riteneva inconsistente l' uso della probabilità in ambito deterministico.

*A. Einstein aveva l'opinione esattamente opposta:
... non credo che lei abbia ragione quando sostiene la tesi che è impossibile derivare conclusioni statistiche da una teoria deterministica. Le basta pensare alla meccanica statistica.
(Da una lettera a Popper)

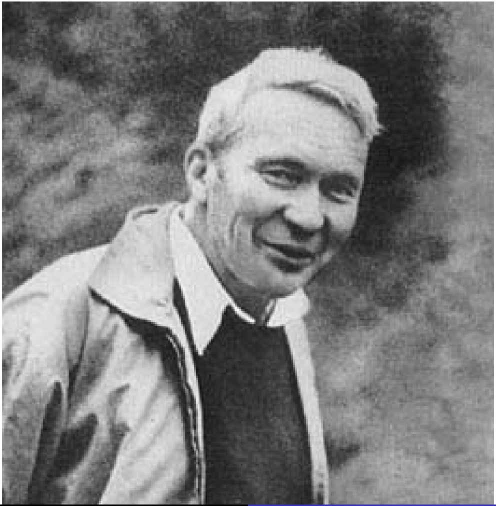
* J. Clerk Maxwell:
La vera logica di questo mondo è il calcolo delle probabilità... Questa branca della matematica che di solito viene ritenuta favorire il gioco d'azzardo, quello dei dadi e le scommesse, e quindi estremamente immorale, è la sola "matematica per uomini pratici".

* B. de Finetti

La probabilità non esiste.

* A.N.K. Kolmogorov

L'assunzione che una definita probabilità esiste per un dato evento sotto certe condizioni è un ipotesi che deve essere verificata e giustifica in ciascun caso individuale.



* A.A. Markov
Uno dei compiti più importanti della teoria delle probabilità è identificare quegli eventi la cui probabilità è vicino a zero od ad uno.

* M. Kac
La teoria della probabilità è la teoria della misura più un'anima.
(L' anima è la nozione di indipendenza.)

* J.A. Wheeler
La probabilità, come il tempo, è un concetto inventato dagli esseri umani, e gli esseri umani devono assumersi la responsabilità delle oscurità che lo circondano.

I concetti di base della probabilità sono delicati e potenzialmente MOLTO pericolosi...

Il libro di Kolmogorov sui fondamenti della teoria delle probabilità venne originariamente pubblicato nel 1933 in tedesco, che all'epoca era la lingua più importante per la matematica. Per ovviare a possibili critiche, dopo pochi anni Kolmogorov lo fece tradurre in russo.
Quando il libro venne pubblicato in russo un ministro sollevò delle perplessità riguardo al concetto di **indipendenza** e chiese se fosse compatibile con il determinismo del materialismo dialettico alla base del marxismo-leninismo.
Kolmogorov dovette replicare immediatamente, secondo la leggenda* rispose: *Compagno ministro, immagini un remoto villaggio dove c' è una lunga siccità. I contadini disperati chiedono al pope di pregare per la pioggia; il pope invoca S. Nicola ed il giorno dopo piove. Sicuramente lei conviene che la pioggia e la preghiera sono due eventi indipendenti.*

* accreditata da persone credibili (ad esempio alcuni suoi allievi)

Follie sul gioco del lotto: i famigerati numeri ritardatari e le ricevitorie fortunate

Molti pensano (con grave danno per le loro finanze) che se un numero non è uscito per un gran numero di estrazioni allora è favorito nella prossima estrazione.
Quindi si ritiene "sensato" che se il 53 non esce da 150 estrazioni, è meglio giocarlo.

Non bisogna fare confusione tra due eventi completamenti diversi:
A) non avere l'uscita del 53 per 151 volte di seguito;
B) il 53 non esce dopo che non è uscito 150 volte.
Poiché la probabilità che il 53 esca è $1/18$, la probabilità che non esca è $1 - 1/18 = 17/18$, quindi la probabilità nel primo caso è

$$\left(1 - \frac{1}{18}\right)^{151} = \left(\frac{17}{18}\right)^{151} \simeq 0,000178 = 0,0178\% \ .$$

Nel secondo caso sapere che il 53 non è uscito 150 volte è irrilevante (in quanto le estrazioni sono indipendenti), quindi la probabilità è

$$1 - 1/18 = 17/18 \simeq 0,9444 = 94,44\%$$

due valori ben diversi!

Un' altra leggenda diffusa tra i giocatori del lotto è fare le puntate in ricevitorie dove ci sono state vincite, ecco il "ragionamento":
è ovvio: ci sono state tante vincite, quindi sarà più probabile che ce ne siano altre!
Ma questo argomento è l'opposto di quello nella "teoria" sui numeri ritardatari; per coerenza si dovrebbe giocare in qualche ricevitoria dove nessuno ha mai vinto.

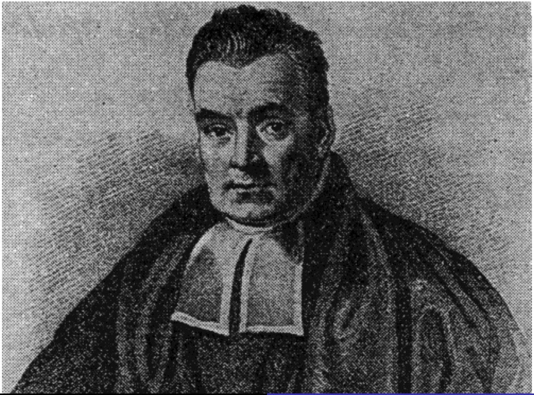
Basterebbe avere chiaro i concetti di **eventi indipendenti** e **probabilità condizionata**.

Daniel Bernoulli (1700-1782) fu il primo a usare la probabilità in medicina e biologia

La famosa formula del reverendo (non conformista) Thomas Bayes (1702-1761)

Probabilità condizionata dell'evento A sapendo che l'evento B è avvenuto

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$



Probabilità **condizionata** negli ospedali
Se nelle ASL conoscessero la formula di Bayes...

Negli anni 90 due importanti ricercatori di psicologia fecero un esperimento in cui veniva proposta ad un gruppo di medici e studenti senior di medicina dell' Università di Harvard la seguente domanda:

Una malattia ha un tasso di incidenza di $1/1000$. **Esiste un test che permette di individuarne la presenza. Questo test ha un tasso di falsi positivi del** 5%. **Un individuo si sottopone al test. L'esito è positivo. Qual è la probabilità che l'individuo sia effettivamente malato?**

Il risultato fu impressionante (in negativo):
solo il 18% rispose in modo esatto!

Per la rispondere correttamente dobbiamo usare la probabilità condizionata e la formula di Bayes:

$$\textit{Probabilità di essere malati essendo risultati positivi} =$$

$$\frac{\textit{Numero di malati positivi}}{\textit{Numero di persone risultate positive}} \simeq \frac{1}{1 + P_e/P_M} \simeq 0.02 = 2\%$$

Ove P_M è la probabilità di essere malati e P_e è la probabilità che il test sia errato.

Si può arivare al risultato anche **senza formule**: poiché la probabilità di essere malati è 10^{-3}, su 100000 soggetti circa 100 sono malati e 99900 sani; il test sbaglia nel 5% dei casi quindi si avranno circa 4995 soggetti sani che risultano positivi e circa 95 malati che risultano positivi. La probabilità di essere malati risultando positivi è $95/(95 + 4995) \simeq 2\%$.

Notare che per un test di laboratorio non è tanto importante la probabilità di errore del test (P_e) quanto il rapporto P_e/P_M: tanto più una malattia è rara, tanto più il test deve essere accurato, altrimenti il risultato non è significativo.

Il Problema Determinismo/Probabilità

Poincaré si domandava (retoricamente?):
Perché molte persone trovano del tutto naturale pregare per avere la pioggia o il sereno, mentre giudicherebbero ridicolo invocare un'eclisse con la preghiera?

Se osserviamo il mondo che ci circonda notiamo che esistono fenomeni regolari e prevedibili, come il susseguirsi del giorno e della notte, l'alternarsi delle stagioni e le eclissi che sono calcolate dagli astronomi con grande anticipo e precisione.

Ci sono però anche fenomeni che non sembrano affatto seguire leggi precise come quelle che valgono per le eclissi o i corpi che cadono.

Quando abbiamo a che fare con giochi come i dadi, la roulette, il lotto e così via, invece di parlare di leggi usiamo termini come caso, aleatorietà, probabilità.

UN APPROCCIO PRAGMATICO

Assumere che esistano due tipi di leggi completamente diverse:
* **derministiche**
* **probabilistiche**

Sistemi deterministici

Esiste una regola di evoluzione, non necessariamente esplicita, che dato lo stato al tempo iniziale $\mathbf{x}(0)$, determina univocamente lo stato ad ogni tempo futuro $\mathbf{x}(t)$.
* Le leggi di Newton della meccanica sono di questo tipo, il caso paradigmatico è la caduta di un grave sotto l'azione della gravità: lasciandolo cadere da un'altezza h tocca terra dopo un tempo

$$t = \sqrt{\frac{2h}{g}}$$

Ancora sistemi deterministici (a tempi discreti)

Regole ricorsive: $\mathbf{x}(t+1) = \mathbf{f}(\mathbf{x}(t))$, in questo caso il determinismo è ovvio:

$$\mathbf{x}(0) \to \mathbf{x}(1) = \mathbf{f}(\mathbf{x}(0)) \to \mathbf{x}(2) = \mathbf{f}(\mathbf{x}(1)) \ \ etc.$$

Parlando di determinismo non si può non citare **Laplace**:
Dobbiamo dunque considerare lo stato presente dell' universo come effetto del suo stato anteriore e come causa del suo stato futuro. Un' intelligenza che, per un dato istante, conoscesse tutte le forze di cui è animata la natura e le posizioni rispettiva degli esseri che la compongono, se per di più fosse abbastanza profonda per sottomettere questi dati all' analisi, abbraccerebbe nella stessa formula i movimenti dei più grandi corpi dell'universo e dell'atomo più leggero: nulla sarebbe incerto per essa e l'avvenire, come il passato, sarebbe presente ai suoi occhi.

Anche se sembra strano il manifesto del determinismo è nell' introduzione del **Saggio Filosofico sulle Probabilità**.

Il marchese Pierre Simon Laplace (1749 - 1827)

Molto spesso Laplace è considerato il gran patron del determinismo, ma ha dato grandi contributi anche alla teoria delle probabilità.

Ovviamente il super matematico di Laplace non esiste; solo in poche situazioni la regola per l'evoluzione è nota esplicitamente, questi sono i casi facili (quelli degli esercizi dei primi anni universitari).

Si potrebbe sperare che un supercomputer possa aiutarci a creare un super matematico (o almeno una sua buona approssimazione) capace di fare previsioni con la precisione desiderata.
Negli anni cinquanta anche il grande John von Neumann ne era convinto, pensava sarebbe stato possibile prevedere (ed addirittura controllare) il clima, ma si sbagliava....
non aveva considerato il caos!

Regole stocastiche (probabilistiche)

Lo stato al tempo iniziale $\mathbf{x}(0)$ non determina univocamente lo stato ad un tempo futuro $\mathbf{x}(t)$.
ESEMPIO: puntando sistematicamente 1 euro sull'uscita di un numero alla roulette, se $x(0)$ è il capitale iniziale allora dopo t giocate il capitale il tasca (o il debito) sarà

$$x(t) = x(0) + \delta_1 + \delta_2 + ... + \delta_t$$

ove $\delta_j = 35$ se è uscito il numero (con probabilità $1/37$), oppure $\delta_j = -1$ se non è uscito (con probabilità $36/37$). Le $\{\delta_j\}$ sono variabili aleatorie indipendenti (se la roulette non è truccata).

Polvere sotto il tappeto?

È onesto sperare di cavarsela semplicemente assumendo che esistono due tipi di situazioni, quelle regolate da leggi certe (deterministiche), e quelle che seguono leggi aleatorie? Si potrebbe infatti notare che i dadi, e le palline delle roulette, obbediscono alle leggi della meccanica di Newton, proprio come i sassi che cadono ed i corpi celesti.

Il famoso dialogo di Democrito tra i Sensi e la Ragione potrebbe essere riproposto per confrontare due posizioni diverse ed opposte sul caso.

Dice il **Determinista**:
Solo in apparenza c'è il caso, solo in apparenza il numero 27 alla roulette esce per motivi aleatori; in realtà esistono solo le leggi deterministiche della meccanica di Newton, il 27 è uscito perché la pallina è stata lanciata con una certa velocità ed un certo angolo, non poteva fare altrimenti.

Risponde il **Probabilista**:
Sei un ingenuo! tu credi che da leggi certe non possa che venire l'ordine; sei in errore: il determinismo non esclude il caso, tuttavia l'aleatorio può portare alla certezza.

Come superare questa dicotomia apparentemente inconciliabile?

Il padre del caos: Henri Poincaré (1854 - 1912)

Caos deterministico

Se pure accadesse che le leggi della natura non avessero più alcun segreto per noi, anche in questo caso potremmo conoscere la situazione iniziale solo approssimativamente .... può accadere che piccole differenze nelle condizioni iniziali ne producano di grandissime nei fenomeni finali. Un piccolo errore nelle prime produce un errore enorme nei secondi. La previsione diventa impossibile e si ha un fenomeno fortuito.
(H. Poincaré)

Ci sono sistemi deterministiche che, nonostante la natura che sembrerebbe suggerire un comportamento regolare, mostrano un' evoluzione temporale è piuttosto irregolare (come ci si aspetta nei processi stocastici).
Può esserci una forte dipendenza dalla condizione iniziale: piccole differenze al tempo iniziali vengono amplificate in modo esponenziale.
Questo è l'effetto farfalla: una farfalla che batte le ali in Brasile può provocare un tornado in Texas dopo un paio di settimane.

Se il sistema è caotico un' incertezza inizialmente molto piccola cresce esponenzialmente nel tempo:

$$|\delta \mathbf{x}(t)| \sim e^{\lambda t} |\delta \mathbf{x}(0)| \ ,$$

ove λ è chiamato *esponente di Lyapunov*. Il sistema può essere predetto con una tolleranza Δ solo fino al tempo di predicibilità T_p che dipende (poco) dall'incertezza iniziale e dall' esponente di Lyapunov $\lambda > 0$

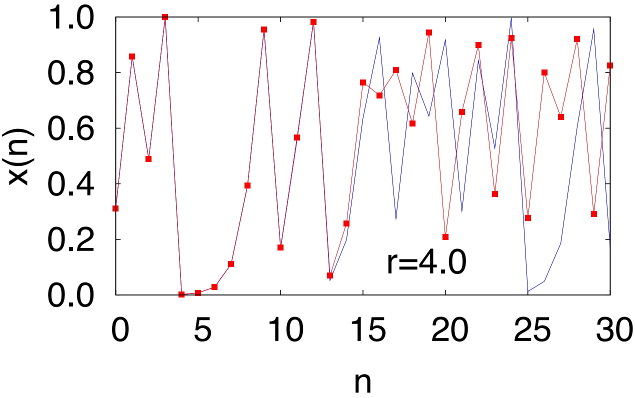
$$T_p \sim \frac{1}{\lambda} \ln \Big(\frac{\Delta}{|\delta \mathbf{x}(0)|} \Big) \ .$$

Poiché la funzione logaritmo cresce molto lentamente $(\ln(5) = 1.609.., \ln(10) = 2.302.., \ln(20) = 2.995.., \ln(40) = 3.688..)$ il tempo di predicibilità è determinato sostanzialmente dall' esponente di Lyapunov e poco dallo sforzo (ed il costo) per determinare la condizioni iniziale con grande precisione (cioè diminuire $|\delta \mathbf{x}(0)|$).

* Il comportamento caotico non è un fatto eccezionale: lo si incontra un po' ovunque (ecologia, astronomia, geofisica, ottica etc)
* Il caos può essere presente anche in sistemi apparentemente innocenti (E. Lorenz, M. Hénon, B.V. Chirikov, ...)
Ad esempio nella mappa logistica

$$x(n+1) = 4x(n)(1 - x(n)) \ ,$$

questo è un modello (iper)semplificato, ma non banale, per la dinamica delle popolazioni.
* L' andamento temporale $x(0), x(1), x(2), ...$ appare "erratico".
* Un piccolo errore raddoppia ad ogni passo, $\delta x(n) \sim 2\delta x(n-1)$; dopo pochi passi l'errore è enorme anche se $\delta x(0)$ è molto piccolo.



Due traiettorie della mappa logistica con condizioni iniziali molto vicine $|x(0) - x'(0)| = 4 \times 10^{-6}$, notare:

* il comportamento irregolare;

* come solo dopo $16 - 17$ iterazioni le due traiettorie diventino completamente diverse.

La domanda
* *cosa succede nel futuro se conosco* $\mathbf{x}(0)$*?*
è (in qualche modo) senza senso anche in ambito deterministico, infatti dopo T_p non si riesce a seguire l' evoluzione.

Il fatto che dagli stessi antecedenti seguano le stesse conseguenze è una dottrina metafisica.
[...] Ma non è molto utile nel mondo in cui viviamo, ove non si verificano mai gli stessi antecedenti e nulla accade identico a se stesso due volte.
[...] L' assioma della fisica che ha, in un certo senso, la stessa natura è "che da antecedenti simili seguono conseguenze simili".
(J. Clerk Maxwell)

Allora?
La domanda "giusta" è:
* *cosa succede nel futuro se conosco* $\mathbf{x}(0)$ *con una data incertezza?*
Questo è un problema di probabilità.

Probabilità in un mondo deterministico: non è un ossimoro.

Se un sistema è caotico allora in qualche modo si comporta come un processo stocastico.
Immaginiamo di avere un'incertezza al tempo $t = 0$, questo corrisponde ad avere una densità di probabilità $p_0(\mathbf{x})$, ad esempio se la condizione iniziale $\mathbf{x}(0)$ è nota con precisione ϵ allora $p_0(\mathbf{x})$ è diversa da zero nella regione $|\mathbf{x} - \mathbf{x}(0)| < \epsilon$.
È possibile mostrare che esiste un'equazione che permette di determinare la densità di probabilità $p_t(\mathbf{x})$ al tempo t, proprio come nei processi stocastici.
Ad esempio per una mappa unidimensionale $x(t+1) = f(x(t))$ si ha la regola:

$$p_{t+1}(x) = \int p_t(y)\delta[x - f(y)]dy$$

ove $\delta[\ ]$ è la delta di Dirac.

Inoltre, sotto ipotesi abbastanza generali, accade che dopo un tempo sufficientemente lungo $p_t(\mathbf{x})$ tende ad una densità $p_\infty(\mathbf{x})$ indipendente dalla densità iniziale.
Nel caso della mappa logistica si ha

$$\lim_{t\to\infty} p_t(x) = p_\infty(x) = \frac{1}{\pi\sqrt{x(1-x)}} \ .$$

Abbiamo quindi che la densità $p_\infty(\mathbf{x})$ è una proprietà intrinseca del sistema ed ha un carattere oggettivo.
I risultati sono indipendenti dalla grandezza dell' incertezza, per quanto piccola sia (purché non nulla), si ha sempre la stessa densità di probabilità asintotica.

Ludwig BOLTZMANN

Sulla connessione tra probabilità e sistemi deterministici avevano ragione Maxwell, Boltzmann ed Einstein; ed aveva torto Popper!

Anche in un contesto puramente deterministico ha senso parlare di probabilità (che non ha un carattere soggettivo).

Comportamento deterministico in ambito probabilistico: non è un ossimoro!

Uno dei compiti più importanti della teoria delle probabilità è identificare quegli eventi la cui probabilità è vicino a zero od ad uno.
(A.A. Markov)

Tutto il valore epistemologico della teoria delle probabilità è basato sul questo: i fenomeni aleatori, considerati nella loro azione collettiva a grande scala, generano una regolarità non aleatoria.
(B.V. Gnedenko e A.N. Kolmogorov)

*** Legge dei grandi numeri**:

La media "empirica" di N variabili casuali indipendenti nel limite di N grande è "vicina" alla media:

$$\frac{1}{N}(x_1 + x_2 + ... + x_{N-1} + x_N)" \to " \langle x \rangle$$

*** Limite centrale**:

La somma (opportunamente normalizzata) di un grande numero di variabili casuali indipendenti è distribuita approssimativamente come una variabile casuale normale.

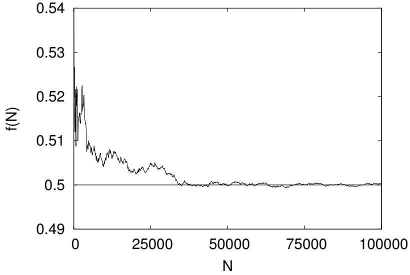
$x_1, x_2, .., x_N$ variabili indipendenti con media $\langle x \rangle$ e varianza σ^2 se $N \gg 1$ la densità di probabilità $p(z)$ della variabile

$$z = \frac{1}{\sigma\sqrt{N}} \sum_{j=1}^{N} (x_j - \langle x \rangle)$$

è ben approssimata dalla curva gaussiana $\frac{1}{\sqrt{2\pi}} e^{-z^2/2}$.

Un comportamento praticamente certo in ambito probabilistico: gioco d'azzardo (in assenza di abilità)

N lanci di moneta non truccata: frequenza $f(N) = N^*/N$ vs. N (N^* è il numero di volte che si ha testa).



I gestori dei casinò vincono sempre, infatti le regole del gioco sono a loro favore (a volte di poco), il singolo giocatore a volte vince, ma per la legge dei grandi numeri il banco alla lunga non può perdere.

George Luis Leclerc conte di Buffon (1707 - 1788)

Il suo problema dell'ago è all'origine del metodo Monte Carlo: con N lanci di un ago di lunghezza L su un pavimento listellato (d è la larghezza dei listelli) è possibile calcolare π; se f_N è la frazione di volte in cui l'ago interseca due listelli, per $N \gg 1$ dalla legge dei grandi numeri: $\pi \simeq 2\frac{L}{f_N d}$.

Una sorta di conclusione

Anche in un sistemi deterministici non è da escludere la probabilità, che (anche se di tipo epistemico) non è soggettiva se:
* Si hanno tanti gradi di libertà (meccanica statistica)
* È presente il caos (anche con pochi gradi di libertà)

Anche in presenza di legge intrinsecamente stocastiche si possono avere comportamenti certi (determinismo statistico) se sono coinvolte tante variabili (teoremi limite)

* **ESEMPIO** È impossible prevedere quando un singolo atomo di un materiale radioattivo decadrà, tuttavia il numero di atomi $N(t)$ non decaduti segue una legge ben precisa

$$N(t) = N(0)e^{-t/\tau} \ ,$$

ove τ è un tempo caratteristico del tipo di materiale.

Qualche Referenza

* G. Boffetta e A. Vulpiani *Probabilità in Fisica* (Springer Italia, 2012)
* D. Costantini *I fondamenti storico- filosofici delle discipline statistico-probabilistiche* (Bollati Boringhieri, 2004)
* B. de Finetti *L'invenzione della verità* (Raffaele Cortina Editore, 2006)
* A. Guerraggio 15 *grandi idee matematiche* (Bruno Mondadori, 2013)
* M. Kac *Gli enigmi del caso* (Bollati Boringhieri, 1996)
* E. Noël (Editore) *Aggiornamenti sull'idea di "caso"* (Bollati Boringhieri, 1992)
* H. Poincaré *Geometria e caso* (Bollati Boringhieri, 1995)
* K. Pomian (Editore) *Sul determinismo* (Il Saggiatore, 1991)
* D. Ruelle *Caso e caos* (Bollati Boringhieri, 1992)
* A. Vulpiani *Determinismo e caos* (Nuova Italia Scientifica 1994, ristampa Carocci 2004)
* A. Vulpiani *Caso, probabilità e complessità* (Ediesse, 2014)