

# Spunti e riflessioni per l'insegnamento del calcolo delle probabilità<sup>\*o</sup>

ALJOŠA VOLČIČ  
Dipartimento di Matematica e Informatica  
Università della Calabria  
volcic@unical.it

## ABSTRACT

*Quoting Felix Klein and Efraim Fischbein we underline the difficulties of introducing the teaching of probability at any level. We then compare two alternatives: we can either study first descriptive statistics, so the first steps in probability can be motivated by suggestive examples (as proposed by Carla Rossi) or we can begin with simple problems from classical probability approach (as proposed by Giovanni Prodi). Following this second suggestion, the basic concepts of probability are introduced through the correspondence between Fermat and Pascal. We also discuss d'Alembert's mistake, introduce the notions of conditional probability and independence, analyse Weaver's three-cards-game and an experiment by Ruma Falk. Finally we discuss some delicate aspects of independence.*

## PAROLE CHIAVE

STATISTICA DESCRITTIVA / DESCRIPTIVE STATISTICS; PROBABILITÀ CLASSICA / CLASSICAL PROBABILITY; IL PROBLEMA DELLA DIVISIONE / THE PROBLEM OF POINTS; ERRORE DI D'ALEMBERT / D'ALEMBERT'S MISTAKE; PARADOSSO DI WEAVER / WEAVER'S PARADOX; ESPERIMENTO DI RUMA / RUMA'S EXPERIMENT.

## 1. INTRODUZIONE

Questo articolo è una rielaborazione della conferenza che ho tenuto a Trieste il 23 novembre 2018 nell'ambito del Progetto *I Lincei per una nuova didattica nella scuola: una rete nazionale*. Il titolo della conferenza era "Qualche idea sull'insegnamento della probabilità tra la scuola secondaria di primo e quella di secondo grado".

Ringrazio gli organizzatori per avermi offerto l'occasione di utilizzare, nell'ambito di questo

---

\* *Title: Some suggestions and considerations concerning teaching of probability.*

<sup>o</sup> Il presente contributo trae origine da un seminario tenuto dall'autore nell'ambito del Programma di Matematica (edizione a. a. 2018-19) proposto agli insegnanti dal Polo di Trieste del Progetto "I Lincei per una nuova didattica nella scuola: una rete nazionale"; cfr. <<https://www.linceiscuola.it/trieste/>>.

tema, materiale e riflessioni che ho accumulato in anni di insegnamento di questa materia. Il grande matematico Felix Klein (1849-1925), famoso per il suo “Programma di Erlangen”<sup>1</sup>, si dedicò anche con grande impegno all’insegnamento. A partire dal 1908 pubblicò un’opera in tre volumi dal titolo *Elementarmathematik vom höheren Standpunkte aus* (Matematica elementare dal punto di vista superiore). Nell’introduzione al primo volume egli enunciò quella che chiamava la “doppia discontinuità”, che era descritta sostanzialmente come segue<sup>2</sup>:

*Una matricola di matematica si confronta all’università con problemi dei quali non si è occupato a scuola. Dopo aver finito l’università, e diventato a sua volta insegnante, ci si aspetta che insegni la matematica tradizionale con la quale non si è confrontato all’università. Pertanto, egli insegna la matematica come gli è stata insegnata alcuni anni prima e i suoi studi universitari rimangono solamente un ricordo più o meno piacevole che non influisce sul suo insegnamento.*

Le cose sono sicuramente cambiate in tutti questi anni, ma l’analisi di Klein rimane sostanzialmente corretta, poiché l’università non contribuisce a sufficienza e con continuità alla formazione dei futuri insegnanti. A questo va aggiunto che anche l’aggiornamento degli insegnanti è discontinuo e non riesce a raggiungere tutti gli interessati.

In Italia, uno dei programmi di aggiornamento degli insegnanti che ha avuto maggior diffusione a livello nazionale è stato il *Piano Nazionale Informatica*, che è nato nel 1988 ed è proseguito per molti anni. Un’informazione molto ricca in merito si trova in BAROZZI, CIARRAPICO 2003.

Ma l’aggiornamento nel settore probabilistico è un *problema nel problema*.

Il famoso psicologo (e insegnante di matematica) Efraim Fischbein (1920-1998) ha scritto che<sup>3</sup>:

*[...] è indispensabile introdurre nei corsi di matematica [...] nozioni di probabilità e statistica [...] a livelli elementari dell’insegnamento [perché] il pensiero probabilistico rappresenta un approccio intellettuale ai problemi essenzialmente diverso da quello ispirato dagli schemi deterministici. [...] Non si può attendere fino all’età, diciamo, di 18-20 anni per cominciare questo tipo di educazione. A questa età, ormai, il pensiero è completamente organizzato nelle sue strutture essenziali.*

---

<sup>1</sup> Con il cosiddetto “Programma di Erlangen” Klein proponeva una visione unificatrice della geometria. Per una traduzione italiana del testo cfr. KLEIN 1890.

<sup>2</sup> KLEIN 1908.

<sup>3</sup> FISCHBEIN 1984, pp. 37-38. Si veda anche in merito FISCHBEIN 1975.

Bisognerebbe iniziare dunque abbastanza presto, ma come si può incominciare?

Bruno De Finetti (1906-1985) ha suggerito come procedere all'inizio:

*Per spiegare con una analogia, osserviamo che c'è qualcosa di preliminare rispetto alla conoscenza della geometria e della fisica, come scienze, ed è la familiarità con alcune nozioni indispensabili per intenderne il senso, come la nozione di lunghezza, la nozione di peso, la nozione di temperatura; familiarità che non consiste nel saperne dare una definizione o spiegazione, ma nell'intendere concretamente, intuitivamente, inavvertitamente, cosa sia una lunghezza di 12 metri, un peso di 4 Kg, una temperatura di +21 °C.<sup>4</sup>*

Le attuali *Indicazioni* ministeriali (che sono ispirate alle conclusioni della Commissione Brocca che ha completato i suoi lavori nel 1992) sembrano suggerire che la statistica descrittiva possa essere una buona palestra per esercitarsi con gli "oggetti" con i quali si lavora nel calcolo delle probabilità.

Carla Rossi, che ne faceva parte, ha pubblicato nel 1999 un libro<sup>5</sup> intitolato *La matematica dell'incertezza* (sottotitolo: *Didattica della probabilità e della statistica*).

Il libro, che è molto ben scritto, è destinato agli insegnanti e non è certo inteso per essere adottato come libro di testo. Esso fornisce un modello di come si può compiere il percorso statistica descrittiva - probabilità - inferenza statistica.

Nella prefazione scrive:

*È opportuno che la trattazione della probabilità segua il tema di statistica [...] Si inizia così a parlare di probabilità avendo già a disposizione delle motivazioni e degli esempi suggestivi [...].*

Non c'è unanimità su questo punto di vista.

Anche Giovanni Prodi (1925-2010) aveva fatto parte della Commissione Brocca. Nel 1992 aveva pubblicato un libro intitolato *Metodi matematici e statistici*. In questo libro (oltre all'algebra lineare e l'analisi) compaiono la statistica descrittiva, la probabilità e la statistica inferenziale. Il libro era stato scritto per il corso di laurea in Biologia che era stato appena riformato, ma, come osservava Prodi nell'introduzione:

*[...] può essere utile anche per la preparazione matematica dei laureati in Biologia [...] che aspirano ad insegnare matematica e materie scientifiche nella Scuola Media.*

Prodi tiene a marcare le differenze tra la statistica descrittiva e la probabilità. Dice

<sup>4</sup> DE FINETTI 1965, p. 8; sull'argomento si veda anche DE FINETTI 1990, 1992.

<sup>5</sup> ROSSI 1999.

ad esempio:

[...] il valore atteso e la varianza formalmente sono definiti dalle stesse espressioni. [...] Ma si coglie la diversità [...] tenendo conto che i coefficienti  $p_i$  ... [qui] indicano delle probabilità [...] mentre nell'ambito della statistica descrittiva indicano dei pesi.<sup>6</sup>

A me sembra più condivisibile l'opinione di Prodi (che è compatibile con la breve citazione di De Finetti riportata prima) e non solamente per il diverso significato che le stesse parole assumono nei due contesti, ma anche perché la statistica descrittiva è una materia complessa, con metodi specifici e problemi che le sono propri.

Anche le nozioni che poi si ritrovano nella probabilità rischiano di rimanere in ombra perché giocano un ruolo secondario nell'affrontare argomenti impegnativi e focalizzati su problemi non probabilistici.

Esiste una vasta letteratura sull'apprendimento della casualità dei ragazzi a partire dall'età prescolare e fino all'età preuniversitaria e la varietà di approcci non è confinata alla dicotomia su quale degli argomenti vada trattato prima, la probabilità o la statistica descrittiva.

Uno studioso che, forse per primo, si era dedicato a questo tema già molti anni fa è il famoso psicologo svizzero Jean Piaget (1896-1980). Spesso viene ricordato come uno dei promotori dell'insegnamento della cosiddetta "matematica moderna", in particolare dell' "insiemistica", a tutti i livelli scolastici.

Per quanto riguarda la probabilità, il suo interesse era concentrato sulla capacità di comprensione dei bambini, a partire dall'età di quattro anni, del concetto di *casualità*. Anche se il suo interesse principale non era la didattica, i suoi studi sono senz'altro da tenere in considerazione.

La sua opera principale in questo campo è apparsa in Italia nel 1976, quando è stato pubblicato il suo libro *La genesi dell'idea di fortuito nel bambino*<sup>7</sup>.

Per quanto importanti, i suoi studi sono datati. Il volume appena citato è stato pubblicato per la prima volta in francese nel 1951 ed è basato su articoli pubblicati

<sup>6</sup> PRODI 1992, pp. 214-215. Sul tema della didattica della probabilità a livello preuniversitario si veda ad es. PRODI 1991.

<sup>7</sup> PIAGET 1976.

in riviste verso la fine degli anni '20 del Novecento.

Sarebbe interessante rifare quelle ricerche nella situazione attuale. Ci deve essere una bella differenza tra gli studenti di Neuchâtel di quasi un secolo fa (con i quali lavorava lui) e quelli, per dire, che frequentano oggi le scuole di Gorizia.

È curiosa la sua invenzione di un congegno che utilizzava per testare le competenze dei bambini. Si trattava di una specie di cassetto oscillante che conteneva delle palline inizialmente ordinate per colore, che poi si mescolavano ottenendo quello che lui chiamava “melange”. La sua intenzione era di vedere a che età i bambini incominciavano a capire che il disordine era irreversibile e, nel contempo, di far vedere loro, sperimentalmente, che l'ordine iniziale non si riformava.

Jean Piaget riteneva che le difficoltà che i ragazzi hanno con l'apprendimento del concetto di *casualità* siano dovute al fatto che lo sviluppo psicologico del bambino ripercorre, in miniatura, lo sviluppo storico del pensiero umano. E la probabilità è l'ultima nata tra le discipline matematiche.

Altri psicologi che si sono interessati alle peculiarità dell'apprendimento della probabilità sono il già citato Fischbein, Ruma Falk e la sua collaboratrice Maya Bar-Hillel. Tra gli italiani i nomi sono tanti e ricordo ancora una volta De Finetti e Prodi, che sono quelli che più sistematicamente si sono dedicati alla probabilità anche da un punto di vista psicologico.

In questo contesto è opportuno menzionare Amos Tversky (1937-1996) e Daniel Kahneman (1934-), quest'ultimo premio Nobel per l'economia nel 2002.

I due studiosi fecero numerosi e ingegnosi esperimenti sugli adulti, dimostrando che in certe situazioni i processi decisionali umani violavano sistematicamente alcuni principi di razionalità<sup>8</sup>.

Per questo motivo i loro studi risultano di grande rilevanza per l'economia matematica, nella quale l'assunzione fondamentale è che il comportamento degli “agenti” sia razionale, cioè finalizzato alla massimizzazione dell'utilità.

---

<sup>8</sup> TVERSKY, KAHNEMAN 1974.

Alcune delle “trappole” tese alle loro cavie sono degli indovinelli probabilistici molto raffinati e ben congegnati che si sono dimostrati utili per comprendere meglio le difficoltà di apprendimento della probabilità e sono quindi interessanti per chi si occupa della didattica. In (VOLČIČ 2015) si possono trovare alcune variazioni su un problema da loro proposto.

Lo schema mentale causa-effetto è molto naturale e si sviluppa abbastanza presto. Ogni effetto “deve” avere la sua causa.

Era così nelle civiltà antiche. Una tempesta si scatenava e faceva affondare una nave perché Nettuno era arrabbiato, la caccia andava bene se Diana era soddisfatta del sacrificio fatto.

L'evento casuale era considerato un segno divino. Si lanciavano gli astragali per avere suggerimenti sulle azioni da intraprendere.

La frase di Cesare “*alea iacta est*” non vuol dire che si è fatto consigliare dai dadi per prendere la sua decisione, era troppo navigato. Ma forse era destinata ai suoi soldati per sentirsi confortati dall'approvazione divina.

D'altra parte ancora oggi sopravvive nella società cosiddetta evoluta una buona dose di “pensiero magico”. Sugli aerei, ad esempio, non esistono le file 13 e 17. Eppure non risulta che i passeggeri di quelle file siano mai stati vittime di incidenti che avessero risparmiato gli altri.

A Capodanno veniamo sommersi dagli oroscopi.

*Direct Line*, una compagnia di assicurazione automobilistica online, ha pubblicato uno studio che valuta il legame statistico che nel 2011 c'è stato tra il segno zodiacale dei guidatori e gli incidenti stradali. Vi raccomando cautela se vedete per strada uno del segno del Cancro! Quel segno zodiacale è risultato il più incidentato con un tasso superiore alla media del 9,43%!

Naturalmente se credessimo in queste cose e nei numeri del lotto ritardatari potremmo interrompere qui, io la scrittura e voi la lettura, e tirare fuori un mazzo di tarocchi.

Con monete, dadi, sacchetti contenenti palline colorate e/o numerate, mazzi di carte e numeri della tombola si può proporre una serie di indovinelli istruttivi. Uno dei pregi della matematica è che gli oggetti con cui si fanno gli esperimenti sono intercambiabili e ciò contribuisce a rendere il pensiero degli studenti più flessibile. In letteratura si trovano tanti “esercizi” pronti all’uso, ma quelli semplici si possono fare anche a casa. L’importante è rendersi conto che nella probabilità, come in altre parti della matematica, gli esercizi non sono fine a se stessi, ma servono per illustrare e, alle volte, introdurre i concetti.

Ad esempio (se vogliamo seguire il suggerimento di De Finetti), si può prendere un mazzo composto da tre cuori e tre picche. Si mescola e si pesca una carta. Che cosa ci possiamo aspettare? (Uno dei due semi, con la stessa probabilità).

E se rimettiamo la carta nel mazzo e rimescoliamo, alla seconda estrazione ci possiamo attendere qualcosa di diverso? (Come prima).

E se peschiamo una seconda carta senza rimettere la carta precedentemente estratta che era, mettiamo, un cuori? (Diminuisce la probabilità che si estragga un cuori). Se continuiamo a estrarre carte dal mazzo si esauriranno prima i cuori o le picche? (Stessa probabilità).

E se aggiungiamo al mazzo una (o due, o tre) carte di cuori e rifacciamo gli stessi esperimenti? Così si dovrebbe creare l’idea del più probabile e dell’ancora più probabile.

Ma incominciamo con ordine, partendo dalla nascita “ufficiale” della probabilità.

## 2. CARTEGGIO FERMAT-PASCAL

Un’opinione largamente condivisa è che il calcolo delle probabilità sia nato nel 1654 con una lettera che il famoso matematico (che di mestiere faceva il magistrato) Pierre de Fermat (1601-1665) scrisse all’altrettanto famoso matematico (oltre che fisico, filosofo e teologo) Blaise Pascal (1623-1662). In essa Fermat risolveva anche alcuni altri problemi riguardanti dei giochi con i dadi, ma lo

scritto è soprattutto interessante per la soluzione dell'antico problema della divisione della posta. Pascal, nella lettera di risposta, concordava con la dimostrazione di Fermat e ne proponeva una variante. Il carteggio tra i due scienziati si può leggere nella traduzione inglese indicata in nota<sup>9</sup>.

Ci tengo a segnalare un divertente “falso” costruito dal famoso probabilista ungherese Alfréd Rényi (1921–1970), il quale ha confezionato quattro lettere apocrife di Fermat a Pascal che “interpolano” parte della corrispondenza andata perduta. L'accuratezza della ambientazione storica, il linguaggio e la fedeltà del testo al pensiero di Fermat rendono quest'opera un piccolo gioiello. La prima traduzione italiana apparve sulla rivista *Induzioni* e successivamente venne riprodotta in (RÉNYI, PASCAL 1991) con una tiratura, credo, bassa. Meriterebbe una terza ristampa e una più larga diffusione.

### 3. LA DIVISIONE DELLA POSTA

Fermat e Pascal si occuparono del problema della divisione della posta in un gioco improvvisamente interrotto, problema che in termini più semplici e attuali possiamo formulare come segue:

*Supponiamo che due giocatori, Anna e Bruno, giochino a testa e croce. L'accordo è che la posta di 100 euro che hanno messo in gioco vada a chi per primo ottiene sei vittorie. Ma nel momento in cui Anna conduce per 5:3 la partita deve essere interrotta e, per circostanze che ignoriamo, non può essere ripresa più tardi. Come va divisa la posta? Certamente Anna deve ricevere di più, ma quanto?*

Fermat e Pascal ebbero la felice idea di formulare il problema in termini probabilistici. Nella Tabella 1 sono riportate in colonna tutte le possibili continuazioni della partita. *Si noti che sono riportate anche le continuazioni “inutili”* (sono quelle delle prime sei colonne della tabella).

Nel calcio, quando si va ai rigori, si fa diversamente. Se la squadra A ne ha segnati cinque e la B tre, e toccherebbe alla B calciare l'ultimo rigore, questo, per regolamento, non viene tirato. Qui invece è conveniente pensare a tutte le otto possibilità che si

<sup>9</sup> Cfr. Siti web <<https://www.york.ac.uk/depts/maths/histstat/pascal.pdf>>.

possono presentare:

A	A	A	A	B	B	B	B
A	A	B	B	A	A	B	B
A	B	A	B	A	B	A	B

Tabella 1. Le possibili continuazioni della partita interrotta.

Delle otto possibilità, per simmetria tutte ugualmente probabili, le prime sette sono favorevoli ad Anna. Solo nell'ultima, Bruno, dopo aver pareggiato, si aggiudica anche il punto decisivo. Quindi a Bruno spetta solo  $1/8$  della posta (12.50 euro), gli altri  $7/8$ , cioè 87.50 euro, vanno ad Anna.

Il calcolo delle probabilità inizia quindi con un grande successo, risolvendo in modo molto semplice un problema posto nel '400 e con il quale si erano inutilmente cimentati, tra gli altri, Luca Pacioli, Niccolò Tartaglia e Gerolamo Cardano.

Ho voluto iniziare la parte più tecnica con questo esempio (che è molto conosciuto) perché introduce con naturalezza gli elementi fondamentali con i quali più tardi Pierre-Simon Laplace (1749–1827) ha posto le basi della probabilità<sup>10</sup>.

#### 4. FORMULAZIONE CLASSICA (DI LAPLACE)

In termini moderni, ma seguendo l'impostazione classica risalente a Laplace, i problemi probabilistici più semplici si possono rappresentare con un insieme  $\Omega$ , chiamato *spazio di probabilità*, (evento certo), costituito da un numero finito di  $n$  elementi (che si chiamano *eventi elementari*) ai quali è naturale assegnare la stessa probabilità  $\frac{1}{n}$ .

I sottoinsiemi di  $\Omega$  si chiamano *eventi*. Ciascuno di essi è composto da zero, uno o più eventi elementari (anche tutti).

La probabilità di  $E$  è *il rapporto tra il numero  $k$  degli eventi elementari favorevoli a  $E$  e il*

<sup>10</sup> LAPLACE 1812, 1814.

numero degli eventi possibili (gli elementi di  $\Omega$ ):

$$P(E) = \frac{k}{n}$$

Nel nostro esempio, gli eventi elementari favorevoli ad Anna sono sette, i casi possibili sono otto e quindi:

$$P(A) = \frac{7}{8}$$

Un altro bell'esempio, adattato dal libro di Giovanni Prodi, fa vedere come un problema di soluzione non immediata può essere reso semplice se si considera lo spazio di probabilità "giusto":

*Mescoliamo un mazzo di 52 carte "francesi". Due giocatori pescano a sorte alternativamente una carta. Vince il gioco colui che pesca l'asso di cuori. Chi è favorito nel gioco?  
Forse si è tentati di dire che è favorito colui che pesca per primo, ma non è vero: il gioco è equo, come si dice.*

Come nel problema della divisione, conviene pensare che i due giocatori continuino a pescare fino alla fine (anche dopo che l'asso di cuori è uscito) e che si distribuiscano 26 carte a testa. Diventa evidente allora che la probabilità che l'asso di cuori stia da una o dall'altra parte è la stessa.

E che succede se togliamo dal mazzo una carta, diversa dall'asso di cuori, in modo che le carte siano 51? In questo caso chi pesca per primo riceve 26 carte, l'altro 25. Più probabile che l'asso di cuori stia nel gruppo più numeroso, quindi il gioco è favorevole a chi comincia.

## 5. GIOCHI CON I DADI

I dadi offrono molte possibilità di fare degli esercizi interessanti. Ne possiamo lanciare uno o due o tre, ...

La prima pubblicazione scientifica che parla di un problema probabilistico precede di qualche decennio il carteggio tra Fermat e Pascal. Si tratta di un breve trattato scritto da Galileo Galilei (1564–1642) nel primo periodo della sua permanenza in

Toscana e si intitola *Sopra le scoperte de i dadi*<sup>11</sup>. Il Granduca di Toscana gli aveva chiesto come mai, lanciando tre dadi, il punteggio di 10 e 11 punti è più frequente di quello di 9 e 12, rispettivamente.

Incominciamo a vedere che cosa succede se lanciamo due dadi.

## 5.1 DUE DADI

Problemi connessi con il lancio di due dadi sono stati frequentemente utilizzati nei test di ammissione a varie facoltà universitarie, ma non è questo a renderli interessanti.

Se lanciamo due volte un dado, (o due dadi contemporaneamente, che è lo stesso) i risultati possibili sono le coppie ordinate di interi  $(m, n)$ , con  $1 \leq m \leq 6$ , e  $1 \leq n \leq 6$ , eventi che riportiamo nella Tabella 2.

(1,6)	(2,6)	(3,6)	(4,6)	(5,6)	(6,6)
(1,5)	(2,5)	(3,5)	(4,5)	(5,5)	(6,5)
(1,4)	(2,4)	(3,4)	(4,4)	(5,4)	(6,4)
(1,3)	(2,3)	(3,3)	(4,3)	(5,3)	(6,3)
(1,2)	(2,2)	(3,2)	(4,2)	(5,2)	(6,2)
(1,1)	(2,1)	(3,1)	(4,1)	(5,1)	(6,1)

Tabella 2. Risultati possibili nel lancio di due dadi.

Il lancio di due dadi ci porta a uscire in modo naturale dallo schema, un po' ristretto, della probabilità classica (ovvero quella di Laplace). Infatti nei giochi, tradizionali e non, capita spesso di lanciare due dadi e di prestare attenzione non tanto alla coppia ordinata che si ottiene, ma al punteggio totale. Esso è rappresentato nella Tabella 3.

<sup>11</sup> GALILEI 1596.

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

Tabella 3. Punteggi totali possibili nel lancio di due dadi.

Come è ovvio, il punteggio totale va da 2 a 12 e chi abbia anche una modesta esperienza sa che non tutti i punteggi sono ugualmente probabili. Basta dare un'occhiata alla tabella precedente per capire che le probabilità sono quelle riassunte nella Tabella 4:

2	3	4	5	6	7	8	9	10	11	12
$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Tabella 4. Probabilità dei punteggi totali nel lancio di due dadi.

Una cosa importante è far capire che lanciare lo stesso dado due volte, o lanciare contemporaneamente due dadi di colore diverso, o lanciare due dadi tra loro indistinguibili, dà risultati identici.

Con un dado si può rappresentare anche il lancio di una moneta: se esce un numero pari, diciamo che abbiamo ottenuto Testa, altrimenti che abbiamo ottenuto Croce.

## 5.2 LA RISPOSTA DI GALILEO AL GRANDUCA

Lanciando tre dadi si ottengono 9 e, rispettivamente, 10 punti, con le seguenti combinazioni dei singoli risultati (Galileo le chiamava “triplicità”):

9		10	
1+2+6	6	1+3+6	6
1+3+5	6	1+4+5	6
1+4+4	3	2+2+6	3
2+2+5	3	2+3+5	6
2+3+4	6	2+4+4	3
3+3+3	1	3+3+4	3

Tabella 5. Numero di combinazioni dei risultati nel lancio di tre dadi con punteggio totale 9 e 10.

Essendo per il 9 e per il 10 il numero delle *triplicità* lo stesso, il Granduca pensava che i due risultati dovessero avere la stessa probabilità.

Galileo fece notare che, considerando i risultati che possono apparire sui tre distinti dadi, 1+2+6, per dire, conta sei volte, 2+2+5 tre volte, mentre 3+3+3 conta una volta sola! Dunque le sei “triplicità” del 9 danno, con le loro permutazioni, 25 casi favorevoli (si veda la seconda colonna della Tabella 5), mentre le altrettante “triplicità” del 10 danno, con le loro permutazioni, 27 casi favorevoli (quarta colonna). Questo spiega perché il 10 e l’11 sono punteggi più probabili del 9 e del 12, e perché gli altri punteggi 8 e 13, 7 e 14, 6 e 15, 5 e 16, 4 e 17 e infine 3 e 18 sono via via meno probabili.

Certo che il Granduca e i suoi compari dovevano avere una bella esperienza con questo gioco, detto il “gioco della zara” (citato da Dante Alighieri nella *Divina Commedia*, precisamente nel *Purgatorio*)! La differenza tra la probabilità di 9 e di 10 è pari a  $\frac{2}{216}$  e quindi meno di  $\frac{1}{100}$  ( $6^3$  è il numero dei casi possibili, 2 sono i casi favorevoli in più per il 10).

Nel suo trattato Galileo fece il conto esatto di tutte le combinazioni dei valori con cui si ottengono i vari punteggi. Il passo in più che Galileo non fece è quello di definire la probabilità dei vari punteggi.

Lo fecero, nel loro carteggio, Fermat e Pascal rispondendo anche ad alcune domande poste dal nobiluomo e scrittore Antoine Gombaud, cavaliere de Méré (1607–1684), al quale interessava il confronto della probabilità di due eventi:

*Indichiamo con  $E_1$  l'evento che esca almeno un 6 su 4 lanci di un dado e con  $E_2$  l'evento che, lanciando due dadi, esca almeno un doppio 6 su 24 lanci. Quale dei due eventi è più probabile?*

Poiché i due eventi corrispondono a due spazi di probabilità diversi, i due studiosi non poterono risolvere il problema (come nel problema della divisione) limitandosi a confrontare il numero dei casi favorevoli, ma dovettero rapportare questi numeri con il numero dei *casi possibili* che, nei due problemi, è diverso.

In questa parte del carteggio compare per la prima volta la definizione formale di probabilità *come rapporto tra casi favorevoli e casi possibili*, aprendo la strada a Laplace. Il problema di de Méré si trova descritto in tanti testi e qui non ne parleremo.

## 6. L'ERRORE DI D'ALEMBERT

Parliamo ora del “famigerato” errore di Jean-Baptiste Le Rond – detto - d'Alembert (1717-1783), noto e stimato scienziato. Nell'enciclopedia di Diderot e d'Alembert<sup>12</sup> (e in un successivo articolo) egli sostenne che, lanciando due monete indistinguibili, ci sono tre eventi elementari aventi la stessa probabilità: due teste, due croci oppure una testa e una croce.

Egli fu tratto in inganno dal fatto che se due dadi (o due monete) sono uguali, noi non siamo in grado di distinguere tra i risultati (3, 4) e (4, 3), né tra (T, C) e (C, T). I due eventi ci appaiono diversi solamente se i dadi sono colorati (o le monete sono di tipo diverso, da un euro e da due euro, per dire). In fondo, è lo stesso errore che fece il Granduca che non distingueva tra  $1+2+6$  e  $2+1+6$ .

Ma anche se operiamo con due dadi dello stesso colore che noi non siamo in grado di distinguere, abbiamo a che fare con due eventi distinti. Ce ne convinciamo pensando al lancio di due dadi colorati osservato da un daltonico e da uno che non lo è: la natura dell'esperimento probabilistico non cambia.

<sup>12</sup> D'ALEMBERT 1751.

È un errore nel quale si cade di frequente.

Negli anni '20 il fisico indiano Satyendranath Bose (1894–1974), in una lezione riguardante il comportamento di certe particelle subnucleari, voleva far vedere che la teoria probabilistica non era compatibile con i risultati sperimentali. Ma, nel fare i conti alla lavagna, commise lo stesso errore di d'Alembert e i risultati dei suoi calcoli risultarono invece compatibili con gli esperimenti. Si accorse così che certe particelle (chiamate oggi *bosoni*) obbedivano a regole probabilistiche peculiari.

Non creduto all'inizio, ne parlò con Albert Einstein (1879-1955) e insieme scrissero un articolo.

## 7. L'INTUIZIONE DI MENDEL

Gregor Mendel (1822–1884), considerato il padre della genetica, era un monaco austriaco. Studiò prima nella piccola, ma antica, Università di Olomouc e poi a Vienna, dove ebbe tra i professori il noto fisico Christian Andreas Doppler (1803-1853).

Entrò in convento a poco più di vent'anni ed ebbe la fortuna di trovare lì un ambiente favorevole allo studio. Nel prendersi cura dell'orto del convento ci mise tutto il suo spirito di osservazione, la sua curiosità e la sua preparazione scientifica. In particolare si interessò alla trasmissione di alcune caratteristiche genetiche nella sua coltivazione di piselli. Ci limiteremo qui a riportare solo una piccola parte delle sue scoperte.

Una delle caratteristiche genetiche che studiò fu il colore dei semi dei piselli, che potevano essere verdi o gialli.

Autoinseminando piselli verdi ottenne sempre discendenza “verde”. Dunque, nessuna sorpresa.

Quando invece fece la stessa l'esperienza con i piselli gialli, scoprì che c'erano in effetti due tipi di piselli gialli: alcuni davano nelle generazioni successive solamente discendenza gialla. Altri piselli gialli invece davano luogo a una discendenza mista.

I piselli verdi e i “gialli del primo tipo” (gli omozigoti) non meritavano dunque altre indagini.

Concentrò allora la sua attenzione sui piselli gialli “del secondo tipo” (gli eterozigoti). La strumentazione scientifica dell'epoca non consentiva di distinguere gli omozigoti dagli eterozigoti. Oggi questo è possibile.

Ripetendo più volte l'esperimento, osservò che dai piselli gialli eterozigoti nascevano nelle generazioni successive tutti e tre i tipi: quelli verdi (omozigoti), i gialli omozigoti e i gialli eterozigoti. Notò anche che tra questi discendenti la percentuale dei tre tipi era circa del 25%, 25% e 50%, rispettivamente.

Possiamo notare (e senza dubbio questa fu un'osservazione che fece anche lui) che era lo stesso rapporto che si ottiene lanciando due monete (con le facce opposte colorate in verde V e giallo G): in media, un quarto delle volte si ottiene VV, un quarto delle volte GG e nella metà dei casi un risultato misto, VG o GV.

Questo, e il fatto che i piselli gialli eterozigoti conservavano in qualche modo una “memoria” della caratteristica verde, gli fece verosimilmente pensare che a ogni caratteristica genetica erano associati due elementi (che lui chiamò *alleli*) e che una delle due caratteristiche era dominante (un'altra sua intuizione), mentre l'altra rimaneva presente nel patrimonio genetico dell'individuo, ma non visibile (*recessiva*, la chiamava lui). Il pisello VV appare ovviamente verde, quello GG ovviamente giallo, ma GV (o VG) appaiono gialli, e questo è meno ovvio. È interessante osservare che anche le strumentazioni di oggi non riescono a dire se un individuo GV(=VG) ha preso l'allele G (o V) dalla madre o dal padre. Gli alleli (anche se intrinsecamente indistinguibili) si comportano quindi come le monete, non come le particelle studiate da Bose!

Ebbe tantissime conferme di questa sua intuizione. Si stima che per i suoi esperimenti usò nel corso della sua vita 28.000 piante di piselli.

È importante notare che non esisteva allora una strumentazione scientifica in grado di confermare queste sue intuizioni: egli costruì un modello, basato in parte sulla probabilità e in parte su congetture genetiche, che funzionava e che spiegava anche altri esperimenti che lui fece (e che qui tralasciamo).

Uno degli scopi di questa breve descrizione del lavoro di Mendel è di far notare nuovamente che gli stessi ragionamenti si possono fare con oggetti diversi: gli alleli si comportano come le monete, i dadi, le carte da gioco, le palline colorate e così via. L'insegnamento della probabilità, più di altre parti della matematica, ha il vantaggio di cambiare facilmente scenario e di far emergere in situazioni apparentemente diverse la struttura matematica sottostante.

## 8. ADDITIVITÀ

Dalla definizione data da Laplace segue immediatamente la proprietà di *additività* della probabilità:

Se  $A$  e  $B$  sono due eventi incompatibili (cioè se  $A \cap B = \emptyset$ ), allora

$$P(A \cup B) = P(A) + P(B)$$

Da questa proprietà segue anche l'additività della probabilità per tre o più eventi a due a due disgiunti: se  $A$ ,  $B$  e  $C$ , ad esempio, sono tre eventi e se  $A \cap B = \emptyset$ ,  $A \cap C = \emptyset$  e  $C \cap B = \emptyset$ , allora

$$P(A \cup B \cup C) = P(A) + P(B) + P(C)$$

I diagrammi di Venn sono un modo molto efficace per rappresentare gli eventi e le operazioni tra eventi come sottoinsiemi di una nuvoletta che rappresenta l'evento certo  $\Omega$ . Ad  $\Omega$  assegniamo area 1 e immaginiamo che  $P(A)$  rappresenti l'area del sottoinsieme  $A$ .

Con questa interpretazione è immediata la comprensione dell'importante formula che dà l'area di  $A \cup B$  nel caso in cui  $A$  e  $B$  non sono incompatibili, cioè quando  $A \cap B \neq \emptyset$ :

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

La dimostrazione formale non è difficile ma, se l'uditorio non è sufficientemente maturo, la si può aggirare con un diagramma di Venn, nel quale si vede che l'area di  $A \cap B$  è contata due volte.

## 9. PROBABILITÀ CONDIZIONATA

La probabilità di un evento è sempre condizionata dalle circostanze in cui l'esperimento aleatorio viene eseguito. Se estraiamo una pallina bianca da un'urna che contiene palline bianche e nere, la probabilità di estrarne una bianca dipende dal numero delle une e delle altre, ma può dipendere anche da altre circostanze: ad esempio, se le palline bianche sono più pesanti (e se le palline vengono estratte da sopra), la probabilità diminuisce. Ma ci può essere anche qualche altra circostanza che influisce sulla nostra valutazione di probabilità, ad esempio potrebbe succedere di sapere che si è verificato un altro evento, legato allo stesso esperimento. È di questo che parleremo adesso.

Fermat e Pascal non se ne sono resi conto (come quello che parlava in prosa ma non lo sapeva), ma il problema della divisione della posta si può vedere come un problema di *probabilità condizionata*. Anzi, è il primo problema di probabilità condizionata che la storia conosca, anche se i due studiosi hanno risolto il problema con un metodo diretto.

Anna ha, a priori, probabilità  $\frac{1}{2}$  di aggiudicarsi la posta (evento  $E$ ). Ma noi sappiamo che nelle prime otto partite ha accumulato un vantaggio di 5 : 3 (evento  $F$ ). Il problema si può dunque riformulare anche in questo modo: qual è la probabilità che si verifichi l'evento  $E$  sapendo che si è verificato l'evento  $F$ ? Il modo più veloce per risolvere il problema è comunque quello che conosciamo e non conviene cercare altre strade.

Come vedremo più avanti, i concetti di indipendenza e di probabilità condizionata sono nati parecchi decenni dopo il famoso carteggio.

Vediamo un altro esempio.

Sulla ruota della roulette i numeri sono in disordine e colorati. Lo zero è verde, mentre gli altri 36 numeri sono per la metà rossi e per la metà neri. Non tutte le ruote sono uguali. In internet mi è capitato di trovarne una sulla quale, dei diciotto numeri rossi, dieci erano dispari e otto erano pari.

Supponiamo che io stia giocando su una roulette di questo tipo e che abbia puntato sul dispari. A priori la mia probabilità di vittoria è

$$P(D) = \frac{18}{37} \sim 0.4864\dots$$

Sono lontano dal tavolo e non vedo la ruota, ma sento il mio vicino, che è in una posizione migliore, esclamare: è rosso!

Tenendo conto di questa informazione, la mia probabilità di vincere è aumentata, perché i casi possibili sono ora 18, ma quelli favorevoli sono 10 e sono quindi cresciuti, in percentuale:

$$P(D | R) = \frac{10}{18} = 0.\bar{5}$$

Abbiamo introdotto il simbolo con cui si denota la probabilità condizionata:  $P(A | B)$  si legge «probabilità di  $A$  sapendo che è avvenuto  $B$ » o, semplicemente, «probabilità di  $A$  condizionata a  $B$ ».

Il risultato di Fermat e Pascal si può riscrivere ( $E$  ed  $F$  sono quelli di prima) usando questa notazione, come

$$P(E | F) = \frac{7}{8}$$

Facciamo un altro esempio.

Sto giocando a poker (con 32 carte) e ho ricevuto una mano abbastanza promettente:  $A \heartsuit, K \heartsuit, 10 \heartsuit, 8 \heartsuit, 7 \clubsuit$ .

In questa fase del gioco posso scartare la carta (o le carte) che non mi servono e pescare nel mazzo, sperando di rafforzare il mio gioco.

Anche chi sa poco delle regole del gioco, sa che il “flush” (colore, in italiano) è un gioco molto forte.

Mi conviene quindi scartare il 7 di fiori, sperando di pescare un qualunque altro cuori. Quattro ne ho io, ne rimangono quattro. I casi possibili sono le altre 27 carte, quelle che non ho in mano. La probabilità che io peschi un cuori è dunque:

$$P(C) = \frac{4}{27} = 0.\overline{148}$$

Ho pescato ora la nuova carta. I veri giocatori non sono impazienti. Per mantenere una certa suspense apro la carta lentamente, scoprendone solo un piccolo angolino. Quanto riesco a vedere mi fa capire che la nuova carta è un sette rosso, 7R: potrebbe essere il  $7\heartsuit$  o il  $7\spadesuit$ , con uguale probabilità.

La probabilità di avere in mano un flush e di aggiudicarsi il piatto è considerevolmente aumentata. Era del 15% scarso, ora è del 50%:

$$P(F | 7R) = 0.5$$

La cosa mi fa piacere (ma prima di puntare devo sciogliere il dubbio aprendo completamente le carte che ho in mano). Non ogni carta rossa avrebbe acceso le mie speranze: se avessi visto l'angolino di un 10 rosso, non mi sarebbe restato che aspettare il prossimo giro, perché il  $10\heartsuit$  è già tra le mie carte e quindi la carta pescata sarebbe stato sicuramente il  $10\spadesuit$ !

Il concetto di probabilità condizionata gioca un ruolo fondamentale nel calcolo delle probabilità. La formula che ora introdurremo è spesso molto comoda.

Siamo nel caso classico e  $A$  e  $B$  sono due sottoinsiemi di  $\Omega$ . Supponiamo che  $B \neq \emptyset$ .

La probabilità condizionata  $P(A|B)$  si calcola con la solita regola del rapporto tra casi favorevoli e casi possibili. I casi possibili si sono ora ridotti ai casi elementari che compongono  $B$ . Invece i casi favorevoli sono i casi elementari che compongono  $A \cap B$ . Indicando con  $|B|$  il numero degli elementi dell'insieme  $B$ , ecc., si ottiene l'importante formula:

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

che permette di calcolare la probabilità condizionata da valori della probabilità degli eventi di  $\Omega$ .

## 10. IL GIOCO DELLE TRE CARTE DI WEAVER

Il gioco è stato inventato negli anni '50 da Warren Weaver (1894-1978). Lui stesso però nel suo bel libro *Lady Luck*<sup>13</sup>, che è del 1963, riconosce che la medesima idea era venuta, molto prima, a Joseph Bertrand (1822-1900) con il suo *paradosso delle tre scatole*. Ma la scenografia di Weaver è molto più intrigante e ha il vantaggio di essere facilmente realizzabile in classe.

Si gioca in due, con tre carte un po' particolari: una (RR) è rossa su entrambi i lati, una (RB) ha un lato rosso e uno blu e la terza (BB) è blu su entrambi i lati.

Le tre carte vengono mescolate dentro un sacchetto.

Naturalmente si suppone che chi mescola lo faccia onestamente: potremmo anche chiedere di farlo a una terza persona che non partecipa al gioco.

La carta scelta viene poggiata sul tavolo, non facendo vedere la faccia di sotto che rimane coperta.

Si scommette sul colore della faccia nascosta e tocca al "banco" indicare il colore.

Il banco paga alla pari: se indovina il colore della faccia coperta vince un euro; se perde, perde un euro.

Le tre carte hanno la stessa probabilità di essere estratte e ciascuna delle due facce ha la stessa probabilità di essere quella visibile.

Voi giochereste contro il banco con queste regole?

Supponiamo che sia visibile una faccia blu. Qual è la probabilità che anche l'altra faccia sia blu?

Un ragionamento possibile è il seguente:

Le tre carte hanno la stessa probabilità, e l'altra faccia è blu se e solo se è stata pescata la carta BB. Ma la probabilità di BB è  $\frac{1}{3}$ .

Sembrerebbe dunque che al banco convenga puntare sul rosso.

Ma la formulazione del problema ci porta fuori strada perché richiama la nostra

---

<sup>13</sup> WEAVER 1982.

attenzione sulle carte piuttosto che sulle facce.

Noi sappiamo qual è il colore della faccia visibile ed è su questa informazione che dobbiamo concentrare la nostra attenzione. Indicheremo questo evento con  $B_1$ .

Questo è l'evento *condizionante*. Poiché su sei facce tre sono blu  $P(B_1) = \frac{1}{2}$ .

Espresso in termini di probabilità condizionata, noi vogliamo sapere quanto vale  $P(B_2 | B_1)$ .

Per la formula che abbiamo dimostrato prima,

$$P(B_2 | B_1) = \frac{P(B_2 \cap B_1)}{P(B_1)} = \frac{P(BB)}{\frac{1}{2}} = \frac{\frac{1}{3}}{\frac{1}{2}} = \frac{2}{3}$$

Quindi il gioco è fortemente favorevole al banco, al quale conviene puntare sul colore visibile.

La “soluzione” precedente confonde  $P(B_2 | B_1)$  con  $P(B_2 \cap B_1)$ . Weaver scrive ironicamente nel suo libro:

*Quarant'anni fa, quando i giovani laureati dovevano lavorare per mantenersi agli studi, l'autore era solito insegnare questo particolare problema ai suoi colleghi a prezzi ragionevoli, usando il metodo sperimentale.*

Una “soluzione lampo” del problema, che mi sembra molto convincente, è la seguente: Delle tre carte, due hanno le facce dello stesso colore e una ha le facce di colore diverso. La probabilità di scegliere una carta “monocolore” è  $\frac{2}{3}$ . Pertanto la strategia vincente è quella di puntare sullo stesso colore della faccia visibile e questa garantisce una probabilità di vittoria pari a  $\frac{2}{3}$ .

## 11. UN ALTRO ESPERIMENTO

L'esempio che segue è ripreso da un esperimento descritto in un articolo di Ruma Falk<sup>14</sup>. Lo scopo di questo test era di dimostrare che anche in un contesto

<sup>14</sup> FALK 1986.

probabilistico, i ragazzi continuano a usare uno schema del tipo causa-effetto.

Un'urna contiene due palline bianche e due nere. Da essa viene estratta una pallina che risulta bianca. Essa non viene rimessa nell'urna.

Poi lo sperimentatore estrae una seconda pallina e, senza farcela vedere, ci chiede qual è la probabilità che anche la seconda pallina estratta sia bianca.

Tutti sanno rispondere a questa domanda. Dopo la prima estrazione sono rimaste nell'urna una pallina bianca e due nere. La probabilità di estrarre una pallina bianca

dall'urna che ha questa composizione è  $P(B_2 | B_1) = \frac{1}{3}$ .

Fin qui tutto facile. Ma Ruma Falk continua con l'esempio.

Lo sperimentatore ripete l'esperimento. Dalla stessa urna estrae una prima pallina e non ce la mostra, né la rimette nell'urna. Poi estrae una seconda pallina. Ce la fa vedere: è bianca.

La domanda ora è: sapendo che la seconda pallina è bianca, qual è la probabilità che la prima fosse bianca? In altre parole, quanto vale  $P(B_1 | B_2)$ ?

Questa domanda ha confuso i ragazzi coinvolti nell'esperimento. L'intenzione della Falk era proprio questa: mettere in discussione l'idea che solamente un evento precedente può condizionare la probabilità di un evento successivo.

I ragazzi hanno ragionato in maniera causale, non probabilistica. Che la probabilità dell'uscita di una pallina bianca nella seconda estrazione "dipenda" da un evento precedente è sembrata loro ragionevole, ma il viceversa li ha lasciati interdetti.

La risposta preferita è stata  $\frac{1}{2}$ . Di fatto costoro hanno ignorato l'informazione ricevuta.

Alcuni sono arrivati ad affermare che la domanda era priva di senso.

La Falk argomenta così: la pallina scelta alla seconda estrazione non poteva essere estratta alla prima e quindi di fatto è come se la prima estrazione fosse avvenuta da

un'urna contenente una sola pallina bianca. Ne segue che  $P(B_1 | B_2) = \frac{1}{3}$  ( $= P(B_2 | B_1)$ ).

Propongo una soluzione più digeribile: lo sperimentatore estrae la prima pallina e se la mette nella tasca sinistra. La seconda pallina estratta se la mette nella tasca destra. Mostrare una delle due palline condiziona nella stessa maniera la probabilità del colore dell'altra.

## 12. INDIPENDENZA

La probabilità condizionata ci permette di introdurre un concetto molto importante, quello di *indipendenza* tra eventi.

Negli esempi precedenti l'evento condizionante cambia la nostra valutazione della probabilità dell'evento osservato.

Però potrebbe anche capitare che

$$P(A | B) = P(A)$$

In questo caso si dice che i due eventi sono *indipendenti*. Questo succede senz'altro nei casi banali, se cioè  $A = \emptyset$  oppure se  $B = \Omega$ . Perciò non ci interesseremo di questi due casi.

Introduciamo due definizioni che ci saranno utili in seguito.

Diciamo che due eventi  $A$  e  $B$  sono *positivamente correlati*, se

$$P(A | B) > P(A)$$

cioè se, sapendo che  $B$  si è verificato, la probabilità di  $A$  aumenta.

Ovviamente diremo che due eventi sono *negativamente correlati*, se

$$P(A | B) < P(A),$$

cioè se, sapendo che  $B$  si è verificato, la probabilità di  $A$  diminuisce.

Vediamo qualche esempio.

Prendiamo il solito mazzo di 52 carte francesi. Gli eventi siano  $A$ =«peschiamo un asso» e  $C$ =«peschiamo un cuori». Si ha naturalmente  $P(A) = \frac{4}{52} = \frac{1}{13} = P(A | C)$ . Quindi

$A$  e  $C$  sono indipendenti.

Il modo più appropriato per definire l'indipendenza è la formula  $P(A | B) = P(A)$ ,

perché essa descrive in maniera significativa la relazione che c'è tra  $A$  e  $B$ . Inoltre vi si legge anche la correlazione positiva e negativa.

Però questa formula ha un difetto:  $A$  e  $B$  vi compaiono in maniera asimmetrica, giocano un ruolo diverso. Conviene allora ricordare che

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

e sostituire questa identità nella definizione. Otteniamo allora

$$\frac{P(A \cap B)}{P(B)} = P(A),$$

ovvero

$$P(A \cap B) = P(A)P(B)$$

Questa formula si chiama la *formula moltiplicativa* ed è utile quando si vuole verificare se due eventi sono indipendenti oppure no.

Questa formula è equivalente alla definizione, se  $P(B) \neq 0$ .

Sconsiglio di utilizzare la formula moltiplicativa (con l'aggiunta della condizione  $P(B) \neq 0$ ) come definizione di indipendenza (anche se qualcuno lo fa). La definizione che abbiamo dato noi indica molto più esplicitamente il significato di indipendenza. D'altra parte la formula moltiplicativa può essere comoda per verificare l'indipendenza.

La formula moltiplicativa è evidente nel caso in cui si tratti di eventi di un insieme prodotto, come nell'esempio seguente.

Se lanciamo due dadi, prendiamo gli eventi  $A = \text{«con il primo dado ottengo un numero alto, 4, 5 o 6»}$  e  $B = \text{«con il secondo dado ottengo un numero intermedio, cioè 3 o 4»}$ .

Nella Tabella 6 gli eventi elementari che compongono  $A \cap B$  sono indicati con le parentesi quadre e ce ne sono  $2 \times 3 = 6$ .

Ne segue immediatamente che  $P(A \cap B) = \frac{6}{36} = P(A)P(B) = \frac{3}{6} \cdot \frac{2}{6}$  e i due eventi sono indipendenti, come era facile capire.

(1,6)	(2,6)	(3,6)	(4,6)	(5,6)	(6,6)
(1,5)	(2,5)	(3,5)	(4,5)	(5,5)	(6,5)
(1,4)	(2,4)	(3,4)	[4,4]	[5,4]	[6,4]
(1,3)	(2,3)	(3,3)	[4,3]	[5,3]	[6,3]
(1,2)	(2,2)	(3,2)	(4,2)	(5,2)	(6,2)
(1,1)	(2,1)	(3,1)	(4,1)	(5,1)	(6,1)

Tabella 6. Gli eventi elementari nel lancio di due dadi.

Può essere comodo rappresentare l'insieme  $\Omega$  delle 52 carte francesi come un insieme prodotto:

$$\Omega = \Omega_1 \times \Omega_2 = \{2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A\} \times \{\heartsuit, \diamondsuit, \clubsuit, \spadesuit\}$$

e costruire vari esempi di coppie di eventi indipendenti che formano dei “rettangoli”, riscoprendo così la vecchia regola che l'area di un rettangolo è “base per altezza”.

Ma l'indipendenza non è sempre così evidente. Vediamo un altro esempio. Riconsideriamo il lancio di due dadi e prendiamo i due seguenti eventi:  $A$ =«con il primo dado si ottiene 4»;  $B$ =«il punteggio totale è 7».

I due eventi sono indipendenti. Infatti:  $P(A) = \frac{1}{6}$  e anche  $P(B) = \frac{1}{6}$ ; poiché  $A \cap B = \{(4, 3)\}$

risulta  $P(A \cap B) = \frac{1}{36} = P(A)P(B)$ .

È interessante notare come l'esistenza di eventi indipendenti richiede che  $|\Omega|$  non sia un numero primo.

Sia infatti  $|A| = a$ ,  $|B| = b$  e  $|A \cap B| = c$ . Se  $A$  e  $B$  sono indipendenti, dopo qualche semplificazione otteniamo  $ab = c|\Omega|$ . Ne segue che  $c < \min(a, b)$ . Ma poiché  $|\Omega| = \frac{ab}{c}$ , segue che  $|\Omega|$  non è primo.

Quest'altro esempio, più sofisticato, è tratto da un libro<sup>15</sup> di Jordan Stoyanov.

Lo spazio di probabilità consiste in tre lanci di una moneta. È l'insieme delle otto sequenze che abbiamo visto nel problema della divisione. Solo che questa volta non supponiamo che la moneta sia regolare. Poniamo  $P(C) = p$ , con  $0 < p < 1$ , e  $P(T) = q = (1 - p)$ .

Sia  $A = \text{«escono almeno due croci»}$  e  $B = \text{«i tre risultati sono uguali»}$ .

Poiché  $A = \{(C,C,C), (C,C,T), (C,T,C), (T,C,C)\}$  e  $B = \{(T,T,T), (C,C,C)\}$ , si ha:

$$P(A) = p^3 + 3p^2(1-p) \text{ e } P(B) = p^3 + (1-p)^3.$$

Inoltre  $A \cap B = \{(C,C,C)\}$  e quindi  $P(A \cap B) = p^3$

Ne segue che  $P(A \cap B) = P(A)P(B)$  se e solo se

$$(p^3 + 3p^2(1-p))(p^3 + (1-p)^3) = p^3$$

Rielaborando l'equazione si ottiene

$$p^2(1 - 4p + 5p^2 - 2p^3) = p^2(1-p)^2(1-2p) = 0$$

Si vede quindi che il polinomio di quinto grado ha due radici doppie,  $p = 0$  e  $p = 1$ .

La quinta radice, quella interessante, è  $p = \frac{1}{2}$ .

Quindi  $A$  e  $B$  sono indipendenti se e solo se la moneta è regolare.

Guardando il segno del polinomio che esprime  $P(A \cap B) - P(A)P(B)$  possiamo osservare

che  $A$  e  $B$  sono positivamente correlati se  $p \in \left]0, \frac{1}{2}\right[$  e negativamente correlati se

$$p \in \left] \frac{1}{2}, 1 \right[.$$

### 13. L'INDIPENDENZA, IL COMMENSALE SILENZIOSO

Abbiamo introdotto il concetto di *indipendenza* solo verso la fine di questa presentazione. Questa scelta corrisponde alla storia dello sviluppo della probabilità. Ma, in effetti, l'indipendenza era presente implicitamente sin dall'inizio. Galileo dava

<sup>15</sup> STOYANOV 2017.

per scontato che i risultati che si ottengono con i tre dadi siano tra loro indipendenti, e così anche per Fermat e Pascal i risultati nei lanci che rimanevano da giocare non dipendevano dai risultati precedenti. Il problema neanche se lo ponevano.

Solamente l'emergere di problemi più complessi rese naturale e, anzi, inevitabile introdurre il concetto di dipendenza e indipendenza.

In matematica succede spesso che una proprietà importante rimanga a lungo inosservata e latente.

Per fare un esempio classico: nel 1821 Augustin-Louis Cauchy (1789-1857) pubblicò una “dimostrazione” del fatto che la somma di una serie convergente di funzioni continue è continua.

Ma nel 1826 Niels Henrik Abel (1802-1829) scrisse che ci doveva essere un errore nella dimostrazione, presentando due controesempi relativi a *serie di Fourier*. Solo Karl Weierstrass (1815-1897), alcuni decenni dopo, introdusse il concetto, evidentemente non compreso da Cauchy e neanche da Abel, di convergenza *uniforme* di una serie, proprietà che assicura la continuità della somma di una serie convergente di funzioni continue.

Per quanto riguarda il calcolo delle probabilità, il primo a rendersi conto dell'importanza dell'indipendenza (e del suo contrario, la dipendenza) fu Abraham de Moivre (1667–1754), il quale scrisse, più di sessant'anni dopo il carteggio Fermat-Pascal, nel suo trattato “The Doctrine of Chances”<sup>16</sup>:

*Due eventi sono indipendenti quando non c'è tra loro alcuna connessione e il verificarsi di uno dei due non favorisce né sfavorisce il verificarsi dell'altro.*

*Due eventi sono dipendenti, quando sono connessi tra loro in modo tale che la probabilità del verificarsi di uno dei due è modificata dal verificarsi dell'altro.*

E continuava così<sup>17</sup>:

*[...] la probabilità che si verifichino due eventi dipendenti è il prodotto della probabilità di uno dei due e della probabilità che si verifichi l'altro [...] e la stessa regola si applica per un numero qualunque di eventi assegnati.*

<sup>16</sup> DE MOIVRE 1756, p. 6 (traduzione dell'autore).

<sup>17</sup> DE MOIVRE 1756, p. 7 (traduzione dell'autore).

La dipendenza è espressa meglio, l'indipendenza invece è espressa accostando due frasi, delle quali la prima è molto vaga. Che cosa vuol dire esattamente che *tra i due eventi non c'è alcuna connessione*?

Riprendiamo l'esempio di Stoyanov.

Se lanciamo una moneta tre volte, c'è connessione tra i due eventi  $A$ =«escono almeno due croci» e  $B$ =«i tre risultati sono uguali»? Se non avessimo fatto i conti, non sapremmo come rispondere. Usando la definizione “moderna”, abbiamo visto che essi sono indipendenti se e solo se la moneta è regolare. Ma, nella formulazione verbale, la connessione tra  $A$  e  $B$  sembrerebbe la stessa per ogni  $p$ .

Ma c'è dell'altro: de Moivre non si pose la domanda se dall'essere

$$P(A \cap B) = P(A)P(B)$$

si può dedurre che il verificarsi di  $A$  non favorisce né sfavorisce il verificarsi di  $B$ .

In altre parole, la formula moltiplicativa per lui sembrerebbe essere solo un metodo di calcolo della probabilità di  $A \cap B$  se i due eventi sono indipendenti.

È plausibile immaginare che la formula moltiplicativa sia stata suggerita a de Moivre da quanto si verifica per gli spazi prodotto.

Infine, c'è un altro punto delicato che sfuggiva a de Moivre. La dipendenza di tre eventi (per dire), va espressa con quattro condizioni (STOYANOV 2017, esempi 3.2 e 3.3).

In termini “moderni”,  $A$ ,  $B$  e  $C$  si dicono indipendenti se e solo se oltre all'equazione:

$$P(A \cap B \cap C) = P(A)P(B)P(C),$$

presa in considerazione da de Moivre, sono verificate anche le tre equazioni:

$$P(A \cap B) = P(A)P(B), \quad P(A \cap C) = P(A)P(C) \quad \text{e} \quad P(C \cap B) = P(C)P(B).$$

Fu comunque per merito di de Moivre che i matematici incominciarono a prestare attenzione all'indipendenza, perfezionando nei decenni successivi quanto nella sua formulazione c'era ancora di incompiuto.

Noi, che siamo dei nani seduti sulle spalle di giganti come de Moivre, riusciamo a vedere un po' più lontano.

Concludo questo articolo con le referenze che ho usato e che possono essere utili per chi vuol fare ulteriori approfondimenti.

## BIBLIOGRAFIA

D'ALEMBERT J.

1751, *Croix ou pile*, in J. LE ROND D'ALEMBERT, D. DIDEROT (a cura di), «Encyclopédie ou Dictionnaire raisonné des sciences, des arts et des métiers», Tome IV (I edizione), Parigi, André le Breton, Michel-Antoine David, Laurent Durand, Claude Briasson, pp. 512-513.

BAROZZI G. C., CIARRAPICO L.

2003, «Il piano nazionale per l'informatica», *Boll. UMI*, ser. 8, Vol. 6-A, *La Matematica nella Società e nella Cultura*, n. 3, pp. 441-461.

DE FINETTI B.

1964, «Insegnamento di materie scientifiche nella scuola media unica e preparazione degli insegnanti», *Periodico di Matematiche*, 42, n. 1-2, pp. 72-114.

1965, *La probabilità: guida nel pensare e nell'agire*, Quaderno N. 11, Trento, Istituto Universitario di Scienze Sociali di Trento.

1990, *Theory of Probability. A critical introductory treatment*, Chichester, Wiley Interscience Publishers.

FALK R.

1986, *Conditional Probabilities: Insights and Difficulties*, in «*Proceedings of the Second International Conference on Teaching Statistics*», The Second International Committee on Teaching Statistics, Victoria, pp. 292-297.

FISCHBEIN E.

1975, *The intuitive sources of probabilistic thinking in children*, Dordrecht, Reidel.

1984, *L'insegnamento della probabilità nella scuola elementare*, in G. PRODI (a cura di), «Processi cognitivi e apprendimento della matematica nella scuola elementare 35-49», Brescia, Editrice La Scuola.

GALILEI G.

1596, *Sopra le scoperte de i dadi*, in A. FAVARO (a cura di) «Le Opere di Galileo Galilei, Edizione Nazionale», Tipografia La Barbera 1897, Firenze (1718), Vol. VIII, pp. 591-594.

KLEIN F.

1890, «Considerazioni comparative intorno a ricerche geometriche recenti» (trad. it. a cura di Gino Fano), *Annali Mat. Pura Appl.* Serie 2, vol. 17, pp. 307-343.

1908, *Elementarmathematik vom höheren Standpunkte aus*, Vol. 1, Leipzig, Teubner.

LAPLACE P. S.

1812, *Théorie analytique des probabilités*, Paris, Courcier.

1814, *Essai philosophique des probabilités*, Paris, Courcier.

DE MOIVRE A.

1756, *The Doctrine of Chances, or: A Method of Calculating the Probabilities of Events in Play. The third edition*, Londra, Millar (I edizione 1718).

PIAGET J.

1976, *La genesi dell'idea di fortuito nel bambino*, Roma, Newton Compton.

PRODI G.

1991, *Didattica della probabilità nella scuola media*, Relazione nell'ambito del Convegno "Temi dell'insegnamento della matematica nella scuola media ed aspetti interdisciplinari" (Montecatini, 11-14 aprile 1991).

1992, *Metodi matematici e statistici*, Milano, McGraw-Hill.

RÉNYI A.

1991, *Pascal, Lettere sulla probabilità*, "Millelire", Roma, Stampa Alternativa (ristampa).

ROSSI C.

1999, *La matematica dell'incertezza*, Bologna, Zanichelli.

STOYANOV J.

2017, *Counterexamples in probability*, New York, Dover (ristampa).

TVERSKY A., KAHNEMAN D.

1974, «Judgment under Uncertainty: Heuristics and Biases», *Science*, New Series, vol. 185, n. 4157 (Sep. 27, 1974), pp. 1124-1131.

VOLČIČ A.

2015, «Taxi verdi e taxi blu», *Archimede*, n. 3.

WEAVER W.

1982, *Lady Luck*, New York, Dover (ristampa).

## SITI WEB

LEE P. M.

*Fermat and Pascal on probability*,

<<https://www.york.ac.uk/depts/maths/histstat/pascal.pdf>>, sito consultato il 10.12.2019.