pretare la realtà, secondo uno spettro che non preveda soltanto il bianco e il nero, ma molte, più di cinquanta, sfumature di grigio.

De Finetti ha proposto una definizione e una misura della probabilità che possono essere tradotte nel linguaggio, universalmente comprensibile, delle scommesse. Un essere razionale, quando scommette, valuta il suo grado personale di fiducia nel verificarsi di un evento e in base a questo, e alla sua più o meno pronunciata avversione al rischio, decide.

D'altra parte, il filo rosso che lega la probabilità e le scommesse è stato fondamentale anche nelle indagini che le forze di polizia di tutta Europa hanno insieme condotto per portare alla luce un sistema di partite di calcio truccate in molti campionati del nostro continente. Il sospetto è sorto quando grandi quantità di denaro si andavano a concentrare su eventi altamente improbabili – e quindi ben pagati dagli scommettitori – come, ad esempio, la sconfitta del Barcellona contro una piccola squadra.

De Finetti e Kolmogorov sono considerati, penso in modo unanime, i pilastri essenziali su cui si è sviluppata, nel secolo scorso, la teoria della probabilità, fino a penetrare, come strumento irrinunciabile nelle scienze esatte e in quelle sociali.

Sia Kolmogorov che de Finetti consideravano il momento didattico della matematica come centrale. Kolmogorov si impegnò personalmente in un progetto didattico, da lui stesso curato, come docente e autore di testi scolastici, e destinato a bambini particolarmente promettenti. de Finetti continuamente sostenne la necessità di rendere la matematica intuitiva e a lungo animò un gruppo di ricerca sulla didattica della discipline matematiche. Ad una giovane insegnante che gli chiedeva consigli per una serie di lezioni televisive sulla matematica liceale – siamo negli anni '60 – de Finetti rispose: Per cercare spunti interessanti e piacevoli per le lezioni, devi *«guardarti intorno e leggere la realtà con occhi matematici»*.

Non può non colpire come e quanto il loro impatto scientifico e culturale sia stato differente nei loro paesi di origine. È evidente che la diversa predisposizione culturale, in particolare alla "pedagogia scientifica" nei due paesi abbia svolto un ruolo centrale.

⋆

Cattive abitudini, ovvero quando l'intuizione prende il posto della conoscenza

Elvira Di Nardo[1]

Ad una persona come me che mastica numeri e formule per mestiere, lascia sempre perplessa la facilità con cui opinionisti, personaggi televisivi e politici usano statistiche e percentuali per avvalorare le proprie tesi in pubblico. Come scrive Stefano Feltri in *La grande bufala dei numeri sparati dai politici* sul fatto quotidiano online (13/11/2012), per dimostrare l'urgenza delle proprie istanze, i gruppi organizzati sommergono di numeri l'opinione pubblica che "prende gran parte delle decisioni con la pancia" invece di effettuare analisi razionali. Messo di fronte a un numero, quello che si chiede allo spettatore medio è meramente un atto di fiducia. Il che, per i tempi che corrono, non è così ovvio. Anche per le menti più brillanti non c'è modo di verificare in diretta se quelle statistiche o quelle percentuali hanno fondamento, quali sono i dati su cui quei numeri si basano, in che modo sono stati raccolti, e soprattutto elaborati. Enrico Giovannini, presidente dell'Istat, spiega che «per evitare di trovarsi in balia del "diluvio di dati"» – troppe informazioni equivalgono a nessuna informazione – «dobbiamo tutti studiare di più, pretendere rigore scientifico da chi fornisce i dati».

[1] Università della Basilicata; elvira.dinardo@unibas.it.

Perché se è vero che un fiume non può risalire verso la sorgente, potrebbe anche essere vero il contrario, se da qualche parte fosse nascosta una pompa. A questo proposito, può essere illuminante un esempio che lessi qualche tempo fa sul libro scritto da Darrell Huff (1991), *How to lie with statistics*. Un ragazzo, che eseguiva interviste per un sondaggio di opinioni, disse di essersi recato alla stazione dei treni per selezionare le persone cui somministrare il questionario perché lì si potevano incontrare tipologie molto diverse. Qualcuno gli fece notare che era difficile poter intervistare mamme di bambini molto piccoli con questa scelta. E dunque il suo campione non era rappresentativo dell'intera popolazione. Ma il problema non è solo legato alle scelte strategiche degli intervistatori. I dati sono il materiale di base della conoscenza. Al giorno d'oggi ogni cosa può essere monitorata e misurata. Ma come dice Erik Brynjolfsson, direttore ed economista del MIT (Massachusetts Institute of Technology), il vero problema risiede nella capacità umana di usare, analizzare e dare senso ai dati raccolti. E in questo la teoria della probabilità può essere di grande aiuto. Laplace, scienziato del periodo napoleonico, ben noto alla comunità dei matematici, scrisse che la teoria della probabilità era null'altro che senso comune ridotto a calcoli. In effetti il rischio associato alla previsione dei terremoti è solo una percentuale. Il vero problema sta nella corretta interpretazione di quel numero, appunto secondo buon senso. Scrive Rovelli «Non sappiamo se domani ci sarà un terremoto. In questo mondo incerto, chiedere certezze assolute è una sciocchezza. Chi esibisce risposte certe è di solito il meno affidabile». Io preferisco rifarmi ad una citazione di Keynes «Gli uomini pratici che credono di essere alquanto esenti da qualsiasi influenza intellettuale, sono in genere schiavi di qualche economista defunto». In sostanza le certezze sono quasi sempre esibite per motivi economici. Che sono tra le cause del fallimento di molte previsioni. Perché nemmeno i dati, anche ammesso che siano stati raccolti senza fonti di condizionamento, possono fornire certezze. L'esempio della malattia rara non infettiva che colpisce cinque colleghi di uno stesso istituto mi ha fatto tornare in mente un altro esempio, che spesso racconto ai miei studenti. Immaginate di lanciare una moneta equa 20 volte. Quale risultato vi sembra maggiormente plausibile? TTCTCTCCTTCCCTCTCTTC oppure TTTTTTTTTTTTTTTTTTTT? In genere, lo studente concentra la propria attenzione sul primo risultato, andando a contare il numero di occorrenze di testa e di croce. E scarta immediatamente il secondo, che sembra alquanto inverosimile. E invece la probabilità che si verifichi l'uno o l'altro è sempre la stessa, anche se decisamente bassa, dell'ordine di 10^(-7). Questo numero può essere di aiuto nel chiarire la differenza tra probabilità ed imprecisione. Un computer è in grado di riprodurre l'esito del lancio di una moneta equa (a questo proposito si usa un termine che non sempre ha una accezione positiva: si parla di simulazione dell'esperimento). Tuttavia non osserverete (quasi) mai una sequenza di sole teste per il semplice motivo che un computer lavora in aritmetica finita (ossia i numeri irrazionali vengono arrotondati) e come dice Samuel Johnson (cfr. Boswell, 1986), i numeri arrotondati sono sempre falsi. E fonte di imprecisione, aggiungo io. Ciò nonostante furono i computer ad essere usati a Los Alamos per verificare che la reazione a catena innestata dallo scoppio di una bomba atomica avesse fine e non portasse alla distruzione di tutta la Terra. Questo perché l'imprecisione di un computer è quantificabile in termini statistici ed è in genere mantenuta molto bassa, quando si usano i risultati sperimentali. È anche per questo che la probabilità e la statistica sono inscindibili: non si può lavorare con l'una senza usare l'altra. Perché l'uso e la comprensione dei dati attiene più propriamente alla statistica, ma la ricerca e l'impiego di modelli che spieghino l'origine di quei dati è invece dominio di applicazione della probabilità. Il problema nasce quando l'impiego di queste due discipline porta a dei paradossi con cui la nostra mente difficilmente riesce a fare i conti. Perché la teoria della probabilità non è in grado di misurare l'imprecisione o l'incertezza che nasce dal comportamento umano. Si pensi al gioco del lotto. Le tabaccherie napoletane espongono spesso cartelli che riportano il numero di settimane di ritardo di un numero su di una ruota. I turisti in genere considerano questi cartelli puro folclore. Sorridono e li fotografano. Eppure ci sono molte persone che consultano quei cartelli prima di puntare su di un numero. È assolutamente intuitivo e razionale immaginare che se un numero porta un numero di settimane consistente di ritardo, stia per "uscire" (come si dice nel gergo). E per chi sa fare un po' di conti, usando modelli probabilistici, è facile dimostrare che la probabilità che il ritardo superi le n settimane, quando n cresce, tende a diventare nulla. Questa, in teoria, sarebbe una buona notizia per lo scommettitore medio. Peccato che questa probabilità va a zero molto lentamente, al punto che potrebbero passare anni prima che quel numero si verifichi. Non solo. La probabilità che quel numero "esca", di settimana in settimana, è sempre la stessa, perché si presuppone che ad ogni giocata, vengano ripristinate sempre le medesime condizioni iniziali. Quindi il ritardo non ha alcuna influenza sulla probabilità di occorrenza del numero. Differentemente da quanto scrive Rovelli, questo è uno di quei casi in cui lo spirito bayesiano, ossia la valutazione delle probabilità in seguito all'acquisizione di informazioni sull'esperimento, non aiuta il giocatore a desistere dal rovinarsi. La logica fuzzy nasce proprio dalla difficoltà di misurare l'imprecisione e l'incertezza del comportamento umano con i semplici strumenti probabilistici e supera questo inconveniente considerando il ragionamento preciso come caso limite. In sostanza ad un elemento di un insieme si associa il grado di appartenenza a quell'insieme, tenendo così in considerazione le fonti di imprecisione dei processi decisionali e superando la logica bivalente (A e non A) del calcolo delle probabilità.

Tutto ciò sembrerebbe avvalorare la tesi secondo cui l'insegnamento della probabilità e della statistica nelle scuole è superfluo, perché difficilmente aiuta a chiarire le idee sui modi in cui vanno prese decisioni in condizioni di incertezza. O almeno induce a riflettere sulla complessità dei problemi con cui uno studente di scuola media superiore dovrebbe misurare le proprie conoscenze. Eppure, basterebbero già poche nozioni di base per evitare, per lo meno, alcune delle imprecisioni comuni nel linguaggio corrente. Lo stesso Rovelli scrive «Non è raro sentire un telegiornale riportare con rilievo il fatto che in una certa località la percentuale di qualcosa sia superiore alla media. *La percentuale di qualunque cosa è superiore alla media in più o meno la metà delle località (inferiore nell'altra metà)*». Questa affermazione dipende fortemente da cosa si intende per "media" prima ancora di capire questo "qualcosa" a cosa si riferisce. Perché la parola media può avere un significato molto vago. A questo altro tipo di imprecisione, molti si rifanno, in buona o cattiva fede, per avvalorare le proprie tesi. Se si parla di mediana, ossia del valore che si trova al centro di una sequenza di etichette (o numeri) ordinabili (ad esempio reddito basso, medio, alto), allora l'affermazione di Rovelli è corretta: la mediana è quel centro tale che, stacca nella sequenza ordinata lo stesso numero di elementi di sotto (a sinistra) e di sopra (a destra), e questo sempre che la sequenza abbia un numero di elementi dispari. Se si parla di moda, ossia di quel valore che maggiormente compare nella sequenza, allora non è detto che essa goda della stessa proprietà della mediana: moda e mediana godono delle stesse proprietà solo se nella "sequenza" coincidono. Se poi si parla di "media", la faccenda si complica ancora di più. Di quali medie si parla? Aritmetica, geometrica, armonica? Qui entra in gioco anche a cosa si riferisce quel "qualcosa" citato da Rovelli. Perché se si parla di reddito, non è detto che si possano calcolare indici numerici: dipende da come sono state raccolte le informazioni. È stato chiesto di dichiarare il valore numerico del proprio reddito o di apporre una crocetta alla categoria (basso, medio, alto) cui il proprio reddito appartiene? I valori medi calcolabili sono diversi nei due casi. Pertanto sono assolutamente d'accordo con Rovelli che gli studenti vadano iniziati alla statistica già nelle scuole. Imparare a districarsi nei paradossi che la statistica, ma più ancora la probabilità, aprono ai neofiti non può che giovare alla formazione di un pensiero critico e alla ricerca di conoscenza. Come scriveva Steve Lohr sul New York Times del 5 agosto 2009 nell'articolo *For today's graduate, just one word: statistics* ai computer bisogna lasciar fare quello che sanno fare meglio, ossia maneggiare ingenti quantità di dati per analizzare situazioni anomale e difficilmente modellabili per via matematica. Ma è meglio che sia la mente umana a trovare spiegazioni per queste anomalie. E per spiegare queste anomalie bisogna aver fatto esercizio di conoscenza. Per le generazioni che popoleranno il nostro futuro e che così tanto dipendono proprio dai computer, oramai più che dalla televisione, la palestra dell'improbabile può essere un ottimo labirinto dove esercitare la mente.

Bibliografia

J. Boswell (1986), *The Life of Samuel Johnson*, New York, Penguin.
H. Darrell (1991), *How to lie with statistics*, New York, Penguin, New Ed.

*

Le competenze statistiche che aiutano a scegliere e a vivere meglio

Alessandro Rosina

A partire da quale età la probabilità di morte diventa superiore al 2%? Immaginoche per molti lettori di questo articolo la risposta non sia chiara e immediata. Figuriamoci se la domanda fosse posta a tutta la popolazione italiana. Facendo riferimento all'ultima tavola di mortalità Istat (www.demo.istat.it) il rischio di morte supera il 2% a 74 anni (mentre era pari a 65 anni a metà anni Settanta). Ma la questione di fondo è: quanto alto deve essere il rischio per cominciare a preoccuparsi? Un valore pari a 0,5indica che arrivati ad un dato compleanno la probabilità di non compiere quello successivo è di una su due, quindi elevatissima. La probabilità di morte riferita alla popolazione italiana supera tale valore a partire dai 105 anni, quindi, fortunatamente, molto in là nel tempo. Il valore di 0,02 (appunto il 2%), corrisponde invece a un caso su cinquanta. Basso, ma non del tutto rassicurante. Insomma, non da sottovalutare. Ma chi arriva a 74 anni vive l'arrivo del compleanno successivo con la stessa preoccupazione che avrebbe una persona che deve prendere un aereo con una possibilità su cinquanta di precipitare?

In un articolo recentemente pubblicato sul «Sole 24 Ore» (20 gennaio 2013), Carlo Rovelli considera proprio tale soglia per ragionare sulla carenza di competenze statistiche e probabilistiche adeguate nei cittadini comuni. Una carenza che non aiuta a valutare i rischi e a fare scelte consapevoli in un mondo incerto: gestire «conoscenze parziali è più facile se abbiamo idee chiare su probabilità e statistica. Questo significa, per esempio, comprendere che una probabilità del 2%, cioè di uno su cinquanta, che ci sia un terremoto all'Aquila la prossima settimana significa che è decisamente più probabile che il terremoto non avvenga, ma il rischio è lo stesso altissimo, e quindi richiede precauzioni». Purtroppo invece l'uomo comune, non adeguatamente attrezzato con i concetti della statistica, è solo in grado di comprendere la differenza tra «ci sarà il terremoto» e «non c'è pericolo: non ci sarà un terremoto». Supponiamo ad esempio che ad un sindaco venga comunicata una probabilità pari a 0,1 che si verifichi un sisma nella sua zona. Per la maggioranza della popolazione quel 0,1 significa ben poco, eppure indica un rischio elevato. Se il sindaco decide di dare l'allarme, ordi-