



PETE MCARTHUR (CORBIS)

# Statistiche imperfette

**Charles Wheelan, Medium, Stati Uniti**

Le analisi statistiche permettono di capire molti aspetti della vita di tutti i giorni. Ma a condizione che i dati su cui si basano siano corretti e appropriati



**D**ietro ogni studio importante ci sono dati di qualità che rendono possibile l'analisi. E dietro ogni studio scadente... bè, continuate a leggere. Spesso si dice che è possibile "mentire con le statistiche". Io dico invece che alla base degli errori statistici più grossolani ci sono le bugie raccontate dai dati. In molti casi l'analisi statistica è giusta, ma i dati su cui si basano i calcoli sono falsi o inadeguati. Ecco alcuni esempi comuni di come "se entra spazzatura esce spazzatura".

**Selezione** Si racconta che Pauline Kael, storica critica cinematografica del *New Yorker*, commentò così la rielezione di Richard Nixon: "Nixon non può aver vinto, non conosco nessuno che l'ha votato". Molto probabilmente la citazione è falsa, ma è un perfetto esempio di come un campione scelto male (in questo caso, gli amici progressisti della Kael) possa offrire un quadro fuorviante di un bacino più allargato (gli elettori statunitensi). E introduce la domanda che dovremmo sempre farci: come abbiamo scelto il campione o i campioni che stiamo analizzando? Se tutti quelli che compongono la popolazione oggetto dell'analisi non hanno la stessa possibilità di finire nel campione, i risultati forniti dal campione saranno sempre discutibili. Un rito delle campagne presidenziali statunitensi è lo *straw poll* dell'Iowa, in cui i candidati alle primarie repubblicane si presentano nella cittadina di Ames nell'agosto dell'anno precedente alle presidenziali per attirare gli elettori, che pagano trenta dollari per partecipare a questo sondaggio sulle primarie. Lo *straw poll* dell'Iowa non ci dice molto su chi sarà il candidato repubblicano alle presidenziali (ha anticipato solo tre delle ultime cinque nomination repubblicane). Perché? Perché gli abitanti dell'Iowa che pagano trenta dollari allo *straw poll* sono diversi dagli altri cittadini dello stato che votano repubblicano. E i repubblicani dell'Iowa, a loro volta, sono diversi dagli elettori repubblicani nel resto del paese.

La distorsione dovuta alla selezione (*selection bias*) può manifestarsi in molti altri modi. Un sondaggio sui consumatori in un aeroporto sarà influenzato dal fatto che chi viaggia in aereo molto probabilmente è più ricco della media. Un sondaggio in una stazione di servizio sull'autostrada avrà il problema opposto. Entrambi i sondaggi saranno influenzati dal fatto che una persona disposta a rispondere a un questionario è probabilmente molto diversa da un'altra che preferisce non essere disturbata. Se chiedete a cento persone in un luogo pubblico di sottoporsi a un breve questionario e sessanta accettano di rispondere, è probabile che quelle sessanta persone siano significativamente diverse dalle quaranta che non vi hanno degnato di uno sguardo.

**Pubblicazione** I risultati positivi hanno più probabilità di essere pubblicati di quelli negativi, che possono compromettere un'analisi. Immaginiamo di avere appena concluso un rigoroso studio longitudinale in cui abbiamo stabilito che giocare ai videogiochi non aiuta a prevenire il cancro

al colon. Abbiamo seguito un campione rappresentativo di centomila statunitensi per vent'anni: l'incidenza del cancro al colon nei soggetti che giocano ai videogiochi è più o meno la stessa rispetto a quella di chi non gioca. Diamo per assodato che la nostra metodologia sia inappuntabile. Quale prestigiosa rivista medica pubblicherà i risultati della nostra ricerca? Nessuna, per due motivi. Primo, non ci sono solidi motivi scientifici per ritenere che giocare ai videogiochi abbia un impatto sul cancro al colon, perciò non si capisce perché lo studio sia stato fatto. Secondo, il fatto che qualcosa non aiuti a prevenire il cancro non è una scoperta particolarmente interessante. In fin dei conti la maggior parte delle cose non aiuta a prevenire il cancro. Le evidenze negative non sono molto seducenti, in medicina e in altri campi.

La conseguenza è che si tende a distorcere la ricerca che vediamo (o che non vediamo). Mettiamo che un altro ricercatore faccia uno studio longitudinale diverso e scopra che la gente che passa un sacco di tempo a giocare ai videogiochi è effettivamente meno soggetta al cancro al colon. Questo sì che è interessante. È esattamente il tipo di scoperta che cattura l'attenzione delle riviste di medicina, della stampa popolare, dei blogger e delle aziende di videogiochi (che metterebbero sui loro prodotti una bella etichetta per spiegare che i videogiochi fanno bene alla salute). Nel giro di qualche tempo le mamme iperprotettive di tutto il mondo comincerebbero a "proteggere" i figli dal cancro strappandogli i libri dalle mani e costringendoli a giocare ai videogiochi.

Ovviamente, un concetto ricorrente in statistica è che le cose insolite ogni tanto succedono, per una mera questione di probabilità. Su cento studi è probabile che ce ne sia uno che dà risultati totalmente assurdi: per esempio, una correlazione statistica tra i videogiochi e l'incidenza del cancro al colon. Il problema è che i 99 studi che non rilevano alcun collegamento tra i videogiochi e il cancro al colon non saranno pubblicati perché non sono interessanti. L'unico che evidenzia un collegamento statistico, invece, sarà pubblicato e avrà grande risonanza. Questa distorsione non nasce dagli studi in sé ma dalla parzialità delle informazioni che arrivano al grande pubblico. Consultando la letteratura scientifica sui videogiochi e il cancro si troverebbe un solo studio, e quell'unico studio dice che giocare ai videogiochi aiuta a prevenire il cancro. In realtà, 99 studi su 100 non hanno evidenziato alcun collegamento.

**Ricordo** La memoria è affascinante, anche se non sempre è una fonte affidabile di dati. Il nostro impulso naturale è quello di leggere il presente come una logica conseguenza di quello che è successo nel passato. Il problema è che i nostri ricordi si rivelano “sistematicamente fragili” quando proviamo a spiegare un esito particolarmente positivo o negativo nel presente. Prendiamo per esempio uno studio sul rapporto tra la dieta e il cancro. Nel 1993 un ricercatore di Harvard ha compilato un *data set* (una raccolta di dati su soggetti diversi o sullo stesso soggetto in tempi diversi) dall’osservazione di due gruppi di donne della stessa età: il primo gruppo era composto da donne con un tumore al seno, il secondo da volontarie a cui non era stata diagnosticata la malattia. Le partecipanti di entrambi i gruppi erano state interrogate sulle loro precedenti abitudini alimentari. I risultati erano chiari: in percentuali significativamente più alte, le donne malate di cancro avevano seguito diete ricche di grassi quando erano più giovani.

In realtà, però, lo studio non voleva analizzare come la dieta incide sulla probabilità di ammalarsi di cancro. Il vero obiettivo era verificare come la consapevolezza di avere un tumore influenza i ricordi delle donne sulle loro abitudini alimentari passate. Tutte le donne dello studio si erano già sottoposte a un test sull’alimentazione anni prima che ad alcune di loro fosse diagnosticato il cancro. Ebbene, le donne che avevano il cancro ricordavano una dieta molto più ricca di grassi di quella che avevano seguito davvero. Le partecipanti sane invece no.

Ecco come il New York Times ha descritto la natura insidiosa di questa distorsione: “La diagnosi del cancro al seno non ha cambiato solo il presente e il futuro delle donne, ha alterato anche il passato. Le donne malate di cancro hanno (inconsciamente) deciso che la causa probabile della loro malattia fosse una dieta ricca di grassi e (inconsciamente) si sono ricordate una dieta ricca di grassi. Si tratta di un processo tristemente familiare per chi conosce questa malattia e sa quanto sia stigmatizzata: le donne dello studio, come migliaia di altre prima di loro, hanno cercato una causa nei loro ricordi e hanno finito per evocarla nella loro memoria”.

La distorsione del ricordo (*recall bias*) è uno dei motivi per cui gli studi longitudinali sono spesso preferiti a quelli trasversali. In uno studio longitudinale i dati sono raccolti in modo diacronico: si chiede a un volontario cosa pensa della scuola quando ha

cinque anni; poi, tredici anni dopo, si chiama lo stesso soggetto e si verifica se ha lasciato le scuole superiori. In uno studio trasversale, in cui tutti i dati sono raccolti in un momento preciso, dovremmo chiedere al diciottenne che ha lasciato le superiori cosa pensava della scuola quando aveva cinque anni. Il dato è per sua natura meno affidabile.

**Sopravvissuto** Mettiamo che in una scuola superiore i voti di un gruppo di studenti crescano di anno in anno. Il secondo anno i voti sono più alti rispetto al primo, il terzo anno continuano a salire e il quarto sono ancora più alti. Diamo per scontato che non ci sia nessun imbroglio né un uso “creativo” delle statistiche descrittive. Ogni anno questi studenti fanno meglio dell’anno precedente secondo tutti i parametri: media, mediana, percentuale di promossi. A questo punto cosa fate? a) Proponete il preside della scuola come preside dell’anno oppure b) chiedete di vedere altri dati?

## Le cose insolite ogni tanto succedono, per una mera questione di probabilità

Io dico b. Sento puzza di distorsione di sopravvivenza (*survivorship bias*), che entra in scena quando alcuni o molti fattori sono eliminati dal campione, modificando la composizione dei fattori rimasti e influenzando i risultati dell’analisi. Mettiamo che il preside sia un incapace. Gli studenti non imparano niente e ogni anno la metà abbandona gli studi. In questo caso è molto probabile che i voti complessivi degli studenti salgano, anche se a livello individuale nessuno è migliorato. Se ipotizziamo che gli studenti peggiori (quelli con i voti più bassi) siano quelli con la maggiore probabilità di abbandono, allora la media dei voti dei superstiti aumenta costantemente a mano a mano che aumentano gli abbandoni (se in una stanza ci sono individui di diversa altezza, far uscire quelli più bassi fa salire l’altezza media, anche se nessuno è diventato più alto).

**L'utilizzatore sano** Le persone che assumono regolarmente vitamine hanno più probabilità di essere sane, perché sono il tipo di persone che assume regolarmente vitamine! Che le vitamine abbiano un effetto o meno è un problema a parte. Facciamo un esperimento. Mettiamo che il ministero

della salute raccomandi a tutti i genitori di far indossare ai loro figli un pigiama viola prima di metterli a letto, perché secondo una teoria questo aiuta a stimolare lo sviluppo del cervello. Vent’anni dopo, uno studio longitudinale conferma che il fatto di aver indossato il pigiama viola è fortemente correlato con il successo nella vita. Scopriamo, per esempio, che il 98 per cento delle matricole di Harvard ha indossato un pigiama viola da bambino (e alcuni continuano a indossarlo), contro appena il 3 per cento dei detenuti del sistema carcerario del Massachusetts. Ovviamente il pigiama viola non c’entra niente: c’entra il fatto di aver avuto dei genitori che mettono ai figli il pigiama viola. Anche quando si cerca di tenere sotto controllo fattori come l’educazione in casa, restano comunque differenze non osservabili tra i genitori che considerano importante mettere il pigiama viola ai figli e quelli a cui invece non importa niente. Come spiega Gary Tauber, esperto di salute del New York Times: “In parole povere, il problema è che le persone che si preoccupano scrupolosamente di fare attività salutari (per esempio prendere un farmaco nei modi consigliati o seguire una dieta equilibrata) sono fondamentalmente diverse dalle altre”. Questo effetto può inquinare qualsiasi studio che cerchi di misurare la reale efficacia di attività considerate salutari, come fare esercizio o mangiare il cavolo riccio. Crediamo di confrontare gli effetti di due diete: “cavolo riccio” contro “non cavolo riccio”. In realtà, se il gruppo sperimentale e il gruppo di controllo non sono selezionati secondo criteri casuali, stiamo confrontando le diete di due diversi tipi di persone. Abbiamo un gruppo sperimentale diverso dal gruppo di controllo per due aspetti, invece che per uno solo.

Se la statistica è un lavoro investigativo, i dati sono gli indizi. Mia moglie ha insegnato per un anno in una scuola superiore nelle campagne del New Hampshire. Uno dei suoi alunni è stato arrestato per furto con scasso in una ferramenta. La polizia è riuscita a risolvere il caso perché 1) aveva appena nevicato e c’erano delle tracce sulla neve che portavano dalla ferramenta alla casa dell’alunno; 2) in casa sua sono stati ritrovati gli attrezzi rubati. Avere buoni indizi aiuta. Come avere buoni dati. Ma prima bisogna procurarseli, ed è molto più difficile di quel che sembra. ♦ *fas*

### L'AUTORE

**Charles Wheelan** è un giornalista statunitense. Questo articolo è un estratto del suo ultimo libro, *Naked statistics*.