Marco Malvaldi e il teorema di Bayes

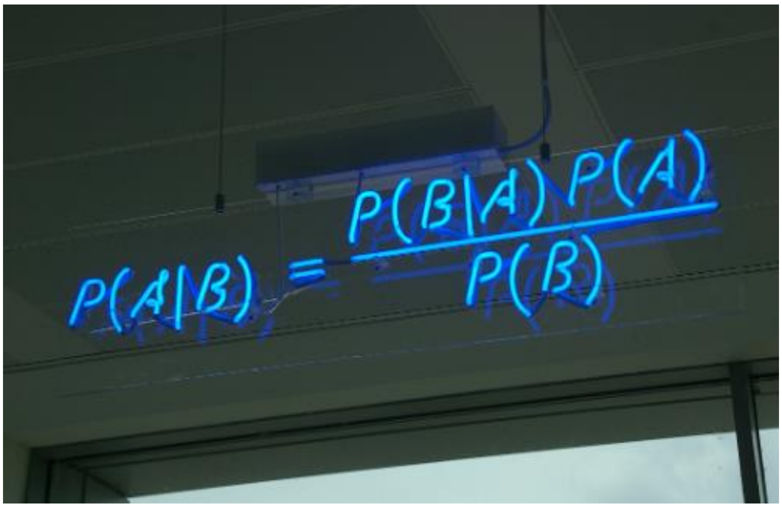

Immagine Wikipedia

Nella realtà fisica, una causa non produce un effetto,
ma una moltitudine di cause distinte contribuiscono
a produrlo senza che si abbia mezzo alcuno per
discernere il ruolo di ciascuna di esse

H. Poincaré

Errore cognitivo: sostituisco invece di imparare.

La mattina del 3 ottobre 1995, gli Stati Uniti erano in uno stato di altissima tensione. La paura di una rivolta razziale era tale che le misure antisommossa a livello nazionale erano state decise dallo stesso presidente, Bill Clinton.

Poi, arrivò l'ora x. Il giudice Lance Ito entrò in aula e dette la parola alla giuria, che lesse la sentenza.

Non colpevole.

O. J. Simpson, ex fuoriclasse di football americano e indimenticato interprete di Nordberg in *Una pallottola spuntata* a fianco di Leslie Nielsen, era stato giudicato non colpevole dell'omicidio della moglie Nicole Brown e del suo amante.

Il processo aveva un doppio spartiacque che lo rendeva pericoloso agli occhi dell'opinione pubblica: da una parte l'accusato, che era nero. Dall'altra una delle due vittime, che era una donna. Divisione di genere contro divisione di razza.

L'accusa puntò decisa sulla prima. Venne ricordato più volte che Simpson era geloso e possessivo, e in un'occasione aveva picchiato la moglie. Uno schiaffo, disse il rappresentante dell'accusa, è il preludio a un omicidio.

Fu a questo punto che il difensore di O.J. Simpson, l’avvocato Alan Dershowitz, ebbe il colpo di genio. Nella sua arringa finale ricordò che l’accusa aveva sostenuto che uno schiaffo era un prodromo dell’omicidio. Ma era davvero così? No, dichiarò Dershowitz. Dalle statistiche si evinceva che negli Stati Uniti ogni anno dai tre ai quattro milioni di donne venivano picchiate o maltrattate dai mariti e dai compagni, ma c’erano stati solo 913 casi in cui la donna era stata uccisa dal marito e 512 in cui era stata uccisa dall’amante. La probabilità che una donna picchiata venisse uccisa era quindi meno di 1 su 2500. Non c’erano dunque prove scientifiche a carico dell’asserzione dell’accusa.

Se qualcuno tra i giurati avesse conosciuto la regola di Bayes, avrebbe forse fatto notare ai propri compagni assolvituri (sic) che il discorso di Dershowitz non era esattamente onesto. La probabilità di essere uccisa, se ti picchiano, è di 1 su 2500, okay. Ma qual è la probabilità di essere uccisa a priori, a prescindere dal fatto che ti picchino o meno? È di 1 su 20.000. Dieci volte minore.

Il dato statistico che Dershowitz avrebbe citato, se invece di essere avvocato difensore fosse stato pubblico ministero, sarebbe stato molto diverso. È vero che la probabilità che una vittima di violenze venga uccisa è di 1 su 2500: molto bassa. Ma quando succede, l’assassino è quasi sempre il compagno. Quando una donna picchiata dal compagno viene uccisa, la probabilità che sia stato lui (la stessa persona che la picchiava) a ucciderla rasenta il 90 per cento. Dershowitz, nel mostrare una probabilità condizionale, aveva mostrato come condizione iniziale il fatto che Nicole Brown era stata picchiata; in realtà avrebbe dovuto mostrare come presupposto il fatto che era stata uccisa, e che da viva era stata picchiata.

Quando si presenta una probabilità condizionale, la condizione di partenza deve essere data dall’insieme di tutti i fatti certi. È successo che Nicole Brown, una donna che veniva picchiata, è stata uccisa. La condizione è quindi questa: picchiata e uccisa. Questo bisognerebbe fare, se si vuole essere sinceri.

Ma Dershowitz, come fece notare lui stesso, non aveva giurato di dire tutta la verità. «Il giuramento che si fa in tribunale di dire tutta la verità vale solo per i testimoni. Avvocati, pubblici ministeri e giudici non giurano, e neanche potrebbero.»

E O. J. Simpson era stato assolto.

Prendo spunto da questo racconto pubblicato ne “*Le due teste del tiranno*”, scritto da Marco Malvaldi e edito da Rizzoli nel 2017, per ricordare il Teorema di Bayes (si legge *bèiʃ*) e alcune sue applicazioni.

Parto da un esempio semplice: se lanciamo un dado, la probabilità che esca un numero dispari è 3/6, cioè ½. La probabilità che esca un numero minore di 4 è di nuovo ½, perché come eventi favorevoli abbiamo che esca 1, oppure 2, oppure 3. La probabilità che esca un numero minore di 4 sapendo che è uscito un numero dispari è però maggiore, 2/3, perché di possibilità favorevoli abbiamo che esca 1 oppure 3, rispetto alle tre possibili, che sia uscito 1 o 3 o 5. Si chiama probabilità condizionale, si indica con P(A/B), ed è la probabilità che si verifichi l'evento A, sapendo che si è verificato B. Il suo valore si ottiene dal rapporto:

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

Applicando questa formula (valida solo se $P(B)\neq 0$) all'esempio precedente, abbiamo che la probabilità che nel lancio di un dado esca un numero minore di 4, sapendo che è uscito un numero dispari, si calcola effettuando il rapporto tra la probabilità di ottenere un numero che sia contemporaneamente minore di 4 e un numero dispari, fratto la probabilità che sia un numero dispari, quindi 2/6 diviso 3/6, cioè 2/3 come avevamo calcolato in precedenza.

Un altro esempio classico: in una famiglia ci sono due figli e si sa che almeno uno è maschio. Qual è la probabilità che entrambi i figli siano maschi? L'insieme dei casi possibili è {MM, MF, FM, FF} e, come è possibile vedere subito, la probabilità richiesta è 1/3, perché ci sono tre coppie con un maschio e una sola con due maschi. Allo stesso numero si arriva applicando la formula precedente, dove $P(A \cap B)$ è la probabilità di avere due maschi, cioè 1/4 e P(B) è la probabilità di avere almeno maschio, cioè ¾. Dal loro rapporto si ottiene la stessa probabilità precedentemente calcolata, 1/3.

Il quesito che segue è stato proposto all'esame di Stato per la maturità scientifica del 2010:

per la ricorrenza della festa della mamma, la sig.ra Luisa organizza una cena a casa sua, con le sue amiche che hanno almeno una figlia femmina. La sig.ra Anna è una delle invitate e perciò ha almeno una figlia femmina. Durante la cena, la sig.ra Anna dichiara di avere esattamente due figli. Si chiede: *qual è la probabilità che anche l'altro figlio della sig.ra Anna sia femmina? Si argomenti la risposta.*

Ovviamente il procedimento è lo stesso dell'esempio precedente, con due femmine anziché due maschi.

La formula vista prima si può generalizzare e ciò che si ottiene è il Teorema di Bayes.

Poiché P(A/B)= $\frac{P(A\cap B)}{P(B)}$ ma anche, per simmetria, P(B/A)= $\frac{P(A\cap B)}{P(A)}$, ricavando $P(A \cap B)$ da quest'ultima e sostituendola nella precedente, si ricava la formula di Bayes:

$$P(A/B)= \frac{P(B/A)P(A)}{P(B)}.$$

In pratica, se sappiamo come un evento A influisce su un evento B, possiamo anche determinare come l'evento B influisce sull'evento A.

In generale:

Siano B_1 , B_2 ... B_n eventi incompatibili, cioè tali che $P(Bi \cap Bj) = 0$, *la cui unione sia l'insieme di tutti gli eventi possibili S. Sia A un evento arbitrario di S tale che P(A)≠0. Allora*

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

Appaiono interessanti le applicazioni del teorema in ambito economico o medico.

I due esempi che seguono sono pubblicati su *Calcolo delle probabilità, esercizi svolti e quesiti per il CdS in Economia e Finanza*, di G. Sanfilippo, del Dipartimento di Scienze Statistiche e Matematiche dell'Università degli Studi di Palermo. (http://www1.unipa.it/giuseppe.sanfilippo/pub/sigad/compiti/RaccoltadiEsercizidiProbabilita.pdf)

Primo esempio. In una ditta che vende dispositivi di un certo tipo il 60 % proviene da una fabbrica A, il 30 % da una fabbrica B e il 10 % da C. Le percentuali di lampadine difettose prodotte da A, B, C sono rispettivamente il 2%, il 4 % e il 5 %. Calcolare la probabilità α che un dispositivo venduto dalla ditta e risultato difettoso sia stato prodotto da C.

Soluzione. Indicando con D l'evento *il dispositivo è difettoso*, si ha P(D|A) = 0.02, P(D|B) = 0.04, P(D|C)=0.05, con P(A)=0.6, P(B)=0.3, P(C) = 0.1.

Allora α = P(C|D) = P(D|C)P(C)/(P(D|A)P(A)+P(D|B)P(B)+P(D|C)P(C)) =

0.05×0.1:(0.02×0.6+0.04×0.3+0.05×0.1) = 5 /29.

Secondo esempio. Una ditta A produttrice di autovetture riceve da quattro fornitori A1, A2, A3, A4 le pastiglie dei freni da installare sulle auto prodotte rispettivamente nelle seguenti percentuali: 65%, 20%, 10%, 5%. Sapendo che i quattro fornitori producono le pastiglie con una difettosità dichiarata rispettivamente del 2%, 2.5%, 4%, 10%, calcolare la probabilità p_d che la ditta A riceve una pastiglia difettosa. Inoltre, scegliendo a caso una pastiglia tra quelle ricevute ed avendo osservato che e difettosa calcolare la probabilità β che essa proviene dal fornitore A2. (Indicare con Bi l'evento "la pastiglia proviene da Ai", i = 1, 2, 3, 4, e con D l'evento "la ditta A riceve una pastiglia difettosa").

Soluzione. Utilizzando la formula di decomposizione si ottiene p_d = P(D) = P(D|B1)P(B1) + P(D|B2)P(B2) + P(D|B3)P(B3) + P(D|B4)P(B4) = 0.027. Inoltre per il teorema di Bayes si ha β = P(B2|D) = P(D|B2)P(B2)/ P(D) = 0.185.

Il seguente esercizio è imvece pubblicato in *Fondamenti di Analisi dei Segnali Biomedici con esercitazioni in Matlab*, di Luigi Landini, Edizioni Pisa University Press, 2014 (http://www.iet.unipi.it/n.vanello/statistica/Es_prob_cond_bayes.pdf).

Si indichi con m l'ipotesi "il paziente ha l'ulcera gastrica" e consideriamo un test diagnostico con le seguenti caratteristiche
- il test fornisce un risultato positivo nel 90% dei casi se il paziente ha l'ulcera gastrica
- il test risulta positivo nel 1% dei casi se il paziente non ha l'ulcera gastrica

La prima condizione può essere scritta

$$p(t_p|m) = 0.9$$

ovvero la probabilità che il test risulti positivo se il soggetto è malato è del 90%.

La seconda condizione può essere scritta

$$p(t_p|s) = 0.01$$

ovvero la probabilità che il test risulti positivo se il soggetto è sano è del 1%.

Sapendo che il medico ritiene che esista una probabilità del 60% che il soggetto sia malato, si calcoli

- la probabilità che il test risulti positivo
- la probabilità che il soggetto sia malato se il test è positivo

Il primo punto può essere risolto ricorrendo al teorema della probabilità totali, considerando che gli eventi soggetto sano (s) o malato (m) sono mutuamente esclusivi.

$$p(t_p) = p(t_p|m)p(m) + p(t_p|s)p(s) = 0.9 * 0.6 + 0.01 * (1 - 0.6) = 0.544$$

Il secondo punto si può risolvere facendo ricorso al Teorema di Bayes

$$p(m|t_p) = \frac{p(t_p|m)p(m)}{p(t_p)} = \frac{0.9 * 0.6}{0.544} = 0.9926$$

Lo studente calcoli:
- la probabilità che il test applicato al soggetto in esame sia negativo
- la probabilità che il soggetto sia malato se il test sia negativo

Come ultimo esempio propongo il successivo, tratto da *Elementi di Probabilità, Statistica ed Elementi Stocastici*, di F. Flandoli (http://users.dma.unipi.it/~flandoli/dispense_Flandoli_2011_versione1.pdf).

Si sa a priori che lo 0.2% della popolazione soffre di una certa malattia dopo i 50 anni. Quella malattia non è ovvia da diagnosticare. Se la malattia è presente, una certa analisi la evidenzia nel 90% dei casi. Se non è presente, l'analisi produce un falso positivo nel 15% dei casi. Un medico esegue l'analisi a un paziente, che risulta positivo. Il medico che decisione prende? (intendiamo: è più propenso a credere che il paziente abbia o non abbia la malattia?). Soluzione: indichiamo con C_1 l'evento: ha la malattia, con A l'evento: risulta positivo all'analisi; conosciamo: $P(C_1) = 0.002$, $P(C_2) = 0.998$, $P(A/C_1) = 0.9$, $P(A/C_2) = 0.15$, quindi calcoliamo

$P(A/C_1)\, P(C_1) = 0.9 \cdot 0.002 = 0.0018$

$P(A/C_2)\, P(C_2) = 0.15 \cdot 0.998 = 0.1497.$

La conclusione è che il medico è ancora più propenso a credere che il paziente sia sano. Quell'analisi è poco discriminante. Non si deve però pensare che l'analisi non sia servita a niente. Ora, per la prossima analisi, si parte da una probabilità a priori diversa: il paziente cade in una categoria di persone che ha probabilità 0:0018/(0.0018+0.1497) = 0.01 di essere ammalata, 0.1497/(0.0018+0.1497) = 0.99 di essere sana (proporzioni ben diverse da quelle iniziali).

Il teorema di Bayes, che potrebbe sembrare un enunciato di matematica moderna, fu in realtà il risultato degli studi del matematico e reverendo britannico Thomas Bayes, che lo formulò intorno al 1740 e che fu pubblicato postumo, nel 1763. Quello citato da Malvaldi non è stato l'unico caso giudiziario famoso in cui si sia fatto ricorso al calcolo delle probabilità. Un altro procedimento penale fu quello che scosse la Francia di fine Ottocento, un caso che «*portò alla superficie quei rigurgiti razzisti e antisemiti di cui tutta l'Europa, e non soltanto la Germania, era inquinata*» (Indro Montanelli, cfr. https://it.wikipedia.org/wiki/Affare_Dreyfus) e che coinvolse il capitano francese Alfred Dreyfus. Il trentacinquenne Dreyfus, ebreo, fu accusato di alto tradimento e rinchiuso nella colonia penale dell'Isola del Diavolo, in Guyana. Dodici anni dopo venne riconosciuto che si era trattato di un errore giudiziario. (per approfondimenti http://www.raistoria.rai.it/articoli-programma-puntate/laffare-dreyfus/24436/default.aspx; https://sites.google.com/site/bayeslegal/legal-cases-relevant-to-bayes/dreyfus). Anche Alan Turing ricorse al teorema di Bayes; gli servì per ridurre drasticamente il numero di settaggi della macchina Enigma e di conseguenza il tempo necessario per la decifrazione del codice segreto tedesco. Attualmente il teorema viene applicato a problemi di logica induttiva e assume un ruolo centrale in tutti le inferenze probabilistiche.

Per ulteriori spunti si può leggere *The Theory That Would Not Die: How Bayes' Rule Cracked the Enigma Code, Hunted Down Russian Submarines, and Emerged Triumphant from Two Centuries of Controversy,* di Sharon Bertsch McGrayne , *A History of Bayes' Theorem* (https://www.lesswrong.com/posts/RTt59BtFLqQbsSiqd/a-history-of-bayes-theorem) e *La matematica, il ragionamento dell'incerto e la discrezionalità*, di Hykel Hosni e Stefano Marmi (http://homepage.sns.it/hosni/papers/discrezionalita.pdf), infine le slides di E. Milotti dell'Università di Trieste all'indirizzo https://wwwusers.ts.infn.it/~milotti/Didattica/Bayes/Bayes.html.