PAOLO GARBOLINO

Come ragionare bene, probabilmente

Il Teorema di Bayes è un potente strumento per aiutarci a non commettere errori quando dobbiamo ragionare in condizioni di incertezza, anche senza bisogno di eseguire calcoli numerici, perché focalizza l'attenzione sui due elementi centrali del ragionamento probabile a fronte di una data evidenza, il peso dell'informazione che abbiamo prima di valutare quella evidenza e il peso dell'informazione che quella data prova ci fornisce. Distinguere fra questi due elementi permette anche di evitare alcune fallacie del ragionamento intuitivo, come l'euristica della rappresentatività e la fallacia dell'accusatore.

How to think well, probably

Bayes' Theorem is a powerful tool to help us avoid errors when reasoning under uncertainty, even without the need to perform numerical calculations, because it focuses attention on the two central elements of probable reasoning in the face of given evidence: the weight of the information we have before evaluating that evidence and the weight of the information that the new evidence provides us. Distinguishing between these two elements also allows us to avoid some fallacies of intuitive reasoning, such as the representativeness heuristic and the prosecutor's fallacy.

Non è questo il luogo per ritornare sull'annoso dibattito filosofico se le regole del calcolo matematico delle probabilità siano le regole logicamente corrette da usare quando dobbiamo eseguire inferenze induttive, quando cioè abbiamo a che fare con ipotesi solamente plausibili e con evidenze incerte, e dunque non possiamo usare le regole della logica deduttiva. Sulle ragioni per sostenere questa posizione, rimando a quanto ho scritto in *Probabilità e logica della* prova[1] ma citerò qua in proposito un passaggio di un libro molto bello, *Pensieri lenti e veloci* di Daniel Kahneman, e molto utile per ciò che riguarda le fallacie del ragionamento probabilistico. Mi piace citarlo perché Kahneman, con Amos Tversky e Paul Slovic, fu l'autore, più di quarant'anni fa, di celebri e importanti ricerche sulle fallacie del ragionamento probabilistico intuitivo[2], che vennero molto utilizzate, a suo tempo, per dare fondamenti teorici, basati sulla psicologia sperimentale, alla critica della tesi che il Teorema di Bayes fosse la regola fondamentale per il ragionamento in condizioni di incertezza. Possiamo riassumere l'argomentazione di fondo di quei critici come segue: il ragionamento bayesiano non è intuitivo e non corrisponde a come, di fatto, ci comportiamo quando dobbiamo prendere delle decisioni in condizioni di incertezza; inoltre, i calcoli probabilistici richiedono una quantità di risorse cognitive che l'uomo comune non possiede o non ha il tempo di impiegare;

[1] GARBOLINO, *Probabilità e logica della prova*, Milano, 2014, 85 ss.

[2] *Judgment under uncertainty: Heuristics and Biases*, a cura di Kahneman-Slovic-Tversky, Cambridge, 1982. Nella letteratura italiana, è sempre valido GIROTTO, *Il ragionamento*, Bologna, 1994, 109 ss.

dunque, dobbiamo formulare delle regole più intuitive e che meglio si adattino alla scarsità delle nostre risorse cognitive.
Per quanto riguarda la prima obiezione, il risultato del lungo dibattito fu che, certamente, il ragionamento bayesiano in molti casi non descrive come ragioniamo quando pensiamo "velocemente", ma è quello normativamente corretto quando vogliamo, e dobbiamo, pensare "lentamente", e questa conclusione è efficacemente riassunta così dallo stesso Kahneman: "La legge di Bayes stabilisce che le credenze precedenti (le probabilità *a priori*) vanno combinate con la 'diagnosticità' delle prove, ovvero con il grado in cui le prove favoriscono l'ipotesi rispetto all'alternativa. (...) Vi sono due idee da tenere a mente riguardo alla logica bayesiana e a come tendiamo a sconvolgerla. La prima è che le probabilità *a priori* contano anche in presenza di prove sul caso in questione: è un dato che spesso non è intuitivamente ovvio. La seconda è che l'impressione che abbiamo intuitivamente della diagnosticità delle prove è spesso esagerata. (...) Ecco come si possono riassumere i punti essenziali della rigorosa logica bayesiana:
– ancorate il vostro giudizio della probabilità di un risultato a una probabilità *a priori* plausibile;
– mettete in discussione la diagnosticità delle vostre prove.
Sono entrambe idee semplici. Fu per me uno shock quando mi resi conto che non mi avevano mai insegnato ad applicarle, e ancora oggi trovo innaturale farlo"[3].
La seconda obiezione, quella per cui le inferenze probabilistiche sono cognitivamente troppo impegnative, è stata superata con lo sviluppo, dagli anni 90 in poi, di programmi che eseguono tali inferenze e che girano su normali *personal computer*, le cosiddette Reti bayesiane (*Bayesian Networks*). Le Reti bayesiane sono, forse, il canto del cigno di quella che gli informatici chiamano *Goofai* (*Good old fashioned artificial intelligence*) perché non sono delle reti neurali (il nome non inganni), le quali costituiscono invece l'architettura di base della nuova IA, l'intelligenza artificiale generativa. L'architettura delle Reti bayesiane è simbolica, è costituita da grafi matematici i cui elementi, detti 'nodi", descrivono le variabili rilevanti del problema, variabili che devono essere decise dall'agente umano, il quale deve anche disegnare il grafo, che

[3] KAHNEMAN, *Pensieri lenti e veloci*, Milano, 2012, 170. Purtroppo, ancora oggi, queste regole sono poco insegnate.

modella le relazioni (probabilistiche e/o deduttive) che esistono fra le variabili, e deve assegnare delle probabilità *a priori* per alcuni nodi del grafo, sulla base delle sue conoscenze come esperto del campo per il quale si costruisce il modello. I 'nodi' delle Reti bayesiane non sono 'neuroni' artificiali, sono variabili logiche e l'apprendimento probabilistico di cui le Reti bayesiane sono capaci consiste nell'aggiornamento automatico delle probabilità di tutti i nodi del grafo quando cambiano le probabilità di alcuni di tali nodi[4]. Ad esempio, nelle Reti usate nelle scienze forensi, tipicamente gli *input* sono cambiamenti nelle probabilità dei nodi che rappresentano le evidenze e gli *output* sono i cambiamenti conseguenti nelle probabilità dei nodi che rappresentano le ipotesi di interesse. Ma ciò che cambia sono solo le probabilità dei nodi e non la struttura del grafo, cioè i nodi stessi e le relazioni di dipendenza fra di loro: le Reti bayesiane 'imparano' solo all'interno del modello disegnato dall'agente umano[5].

La buona notizia è che noi comprendiamo perfettamente quello che fanno le Reti bayesiane, a differenza dei modelli della IA generativa, che eseguono inferenze probabilistiche *non bayesiane*, cioè che non fanno uso del Teorema di Bayes, basate fondamentalmente sulla scoperta di correlazioni statistiche tra i dati e in base alle quali il sistema effettua una previsione. Perché la macchina ha fatto quella previsione probabilistica? Non siamo in grado di spiegarlo, possiamo solo fidarci, perché il modello ha troppi parametri, che sono le connessioni fra i 'neuroni' artificiali (tipicamente i parametri sono dell'ordine di centinaia di miliardi), e macina una quantità astronomica di dati. È davvero impossibile per le limitate risorse cognitive degli esseri umani controllare quali correlazioni statistiche abbia 'visto' il modello e se non siano farlocche (quelle che i buoni vecchi statistici in carne ed ossa chiamano 'correlazioni spurie'). L'ipotesi di fondo dei sostenitori della nuova IA è che, proprio perché questi nuovi modelli della IA generativa sono in grado di passare in rassegna quantità enormi di dati, inevitabilmente finiranno per identificare le correlazioni statistiche genuine, scartando quelle 'spurie'. Uno dei primi saggi

---

[4] I 'neuroni' delle reti neuronali artificiali sono in realtà delle funzioni matematiche, una per ogni 'neurone', che simulano il comportamento dei veri neuroni. Mentre il cervello umano è l'organismo più complesso più sostenibile dell'Universo conosciuto (i consumi energetici del nostro cervello sono ridicolmente bassi), l'IA generativa richiede enormi potenze di calcolo, che attualmente richiedono enormi consumi di energia e di acqua per far funzionare i centri di calcolo (*server*).

[5] Un testo completo e aggiornato sulle Reti bayesiane è quello di CHERUBINI, *Guida al ragionamento probatorio con le reti bayesiane*, Milano, 2024.

che formulava questa ipotesi si intitolava significativamente, "L'irragionevole efficacia dei dati"[6]. I sostenitori del ragionamento bayesiano sono critici rispetto a questa ipotesi[7].
Tornando al ragionamento bayesiano come oggi è impiegato nelle scienze forensi, la distinzione fra la probabilità *a priori* e ciò che Kahneman chiama "la diagnosticità delle prove", ovvero il 'valore probatorio', è efficacemente rappresentato nella formula del Teorema di Bayes, che nella sua formulazione *standard* è come segue, dove *H* è l'ipotesi ed *E* l'evidenza:

P(H|E)=[P(E|H) x P(H)]/P(E)=[P(E|H) x P(H)]/[P(E| H) x P(H)] + [P(E|non H) x P(non H)]
La formula recita che la probabilità che *H* sia vera, data *E*, chiamata anche la probabilità *a posteriori* di *H*, è eguale alla probabilità che *E* sia vera, data *H*, chiamata anche la verosimiglianza di *E*, moltiplicata per la probabilità a priori che *H* sia vera, cioè la probabilità che si ha per *H* prima di osservare *E*, il tutto diviso per la probabilità totale che *E* sia vera, cioè la probabilità che aveva *E* di essere vera prima di venire osservata, sia essendo *H* vera oppure no.
Una formulazione equivalente, forse meno conosciuta ma usata correntemente dagli scienziati, inclusi gli scienziati forensi, è data nei termini del rapporto fra le probabilità *a posteriori* dell'ipotesi *H* e della sua negazione:

P(H|E)/P(non H|E)=[P(E|H)/P(E|non H)] x [P(H)/P(non H)]

Il rapporto *[P(E |H)/P(E|non H)]* è chiamato il rapporto delle verosimiglianze (*likelihood ratio*) ed è quello che l'*European Network of Forensic*

[6] ALON – NORVIG – PEREIRA, *The unreasonable effectiveness of data*, in *IEEE Intelligent Systems*, 2009, 24, 8 ss.

[7] Si vedano, ad esempio, le critiche formulate da uno dei padri delle Reti bayesiane, Judea Pearl: PEARL, *The limitations of opaque learning machines*, in *Possible Minds. Twenty-Five ways of Looking at AI*, a cura di Brockman, New York, 2019, 31 ss.; PEARL, *Radical empiricism and machine learning research*, in *Journal of Causal Inference*, 2021, 9, 78 ss. Naturalmente vengono eseguiti dei controlli di affidabilità su questi modelli, prima che siano resi pubblici, ma è già esperienza comune di come essi possano imparare ad aggirare i controlli. Inoltre, il grande elefante nella stanza è il fatto che la ricerca di punta in questo settore è completamente in mano a imprese private che selezionano anche quali risultati delle loro ricerche mettere a disposizione della comunità scientifica.

Science Institute[8] raccomanda di prendere come misura del valore probatorio dell'evidenza *E*.
Se il rapporto è maggiore di 1, *[P(E|H)/P(E|non H)]* > 1, diremo che la prova *E* è a favore dell'ipotesi *H*; se è minore di 1, *[P(E|H)/P(E|non H)=]* < 1, diremo che la prova è a favore dell'ipotesi *non H*; se è, anche approssimativamente, pari a 1, diremo che la prova *E* è irrilevante (epistemicamente) rispetto alle due ipotesi.
Un semplice esempio è quello della prova del *DNA*, dove *E* sia l'osservazione di una coincidenza tra la traccia biologica e il DNA di un sospettato e l'ipotesi *H* è che il sospettato sia l'origine della traccia. Se non consideriamo la possibilità di contaminazioni ed escludiamo errori di laboratorio, allora *P(E|H)=1,* mentre *P(E|non H)=p* è la cosiddetta 'probabilità di una coincidenza casuale' (*random match probability*). Tipicamente, *p* è molto piccola e quindi il valore probatorio in questo caso è molto alto. Ma attenzione, la probabilità di un falso positivo (errore di laboratorio) deve sempre essere considerata (nessun laboratorio è perfetto), la possibilità di contaminazione nella manipolazione del reperto deve essere valutata ed infine, *last but not least*, occorre sempre valutare la *rilevanza* della prova, ovvero la possibilità che la traccia biologica sia sì stata lasciata dal sospettato, che non ha però commesso il fatto, cioè sia stata lasciata in circostanze estranee al fatto da provare. Una Rete bayesiana è estremamente utile per poter prendere in considerazione tutte queste variabili.
Il Teorema di Bayes è molto utile anche senza bisogno di effettuare calcoli numerici, perché serve a salvarci da alcune fallacie ben note nella letteratura, come la cosiddetta euristica della rappresentatività e la connessa tendenza a non considerare le probabilità *a priori*: quando un individuo è giudicato rappresentativo, cioè avente caratteristiche tipiche della classe a cui appartiene, il giudizio viene influenzato dal grado stimato di rappresentatività (statisticamente irrilevante) mentre sono trascurate variabili statistiche rilevanti, come la frequenza degli individui appartenenti a quella classe presenti nel campione considerato.
Giudicare la probabilità in base alla rappresentatività comporta alcuni vantaggi: le impressioni intuitive prodotte da questo tipo di giudizio sono spesso più

---

[8] ENFSI, *Guideline for Evaluating Reporting in Forensic Science*, Bruxelles, 2015. Dell'istituto fanno parte, in Italia, il RIS dei Carabinieri e il Laboratorio Centrale della Polizia scientifica.

precise di quanto non lo siano intuizioni casuali ed è per questo che le generalizzazioni di senso comune o massime di comune esperienza possono avere un certo valore 'diagnostico', per dirla con Kahneman.

Nella maggior parte di casi, le persone che si comportano in maniera amichevole (evidenza osservabile *E*) sono effettivamente amichevoli (ipotesi *H*); è più probabile che un atleta professionista molto alto e magro (evidenza osservabile *E*) giochi a basket (ipotesi *H*) piuttosto che a calcio; è più probabile che leggano un giornale cartaceo (evidenza osservabile *E*) i laureati (ipotesi *H*) piuttosto che i diplomati. Questi giudizi trovano espressione nel rapporto delle verosimiglianze:

$P(E/H) > P(E/non\ H)$ e quindi l'evidenza ha un peso positivo: $[P(E/H)/P(E/non\ H)] > 1$

Ma se dobbiamo valutare un'ipotesi riguardante un singolo individuo, allora dobbiamo prendere in considerazione anche le probabilità *a priori*, e il Teorema di Bayes ci obbliga a farlo. Vedo per strada qualcuno leggere il giornale (*E*): quale delle seguenti ipotesi, questa persona è laureata (*H*) oppure non è laureata (non *H*), è più probabile *a posteriori*, cioè sulla base di questa osservazione *ma anche* sulla base di tutte le altre informazioni che posso avere? La rappresentatività ci spingerebbe a optare per l'ipotesi laurea, ma dobbiamo tenere in conto la probabilità *a priori* di incontrare un laureato per strada, e, per strada, è più frequente incontrare persone non laureate.

Se è vero che $[P(E/H)/P(E/non\ H)] > 1$, è anche vero che $[P(H)/P(non\ H)] < 1$ e questo secondo rapporto è, ragionevolmente, molto più basso del primo e quest'ultimo può non essere sufficiente ad alzare la bilancia del giudizio finale fino a farla pendere a favore di *H*, cioè a far sì che $P(H/E) > P(non\ H/E)$.

Naturalmente, se questa persona si trovasse, poniamo, nell'atrio della sede di una grande azienda, le probabilità *a priori* cambierebbero. Le probabilità *a priori* si basano su tutta la conoscenza disponibile, quindi sulle frequenze statistiche note, ma anche sull'informazione non statistica e quindi anche sulle circostanze ambientali. Questo è un punto fondamentale della logica bayesiana: i giudizi di probabilità non si basano solamente sulle informazioni statistiche e, spesso, è sufficiente formulare dei giudizi *qualitativi*, non necessariamente numerici, in termini di 'maggiore di' e 'minore di'.

Un celebre caso di scuola è quello del taxi blu. Un taxi è stato coinvolto in un incidente notturno e in città ci sono solo due compagnie di taxi, i taxi verdi e i taxi blu. Le informazioni che abbiamo sono: (a) l'85% dei taxi sono verdi e il 15% sono blu; (b) un testimone ha identificato come blu il taxi coinvolto nell'incidente e questo testimone ha identificato correttamente i taxi nell'80% dei casi e ha sbagliato nel 20% dei casi, dunque la probabilità di errore è del 20%. Qual è la probabilità che il taxi coinvolto nell'incidente fosse blu? Siano *B* e *V*, rispettivamente, le ipotesi che il taxi coinvolto fosse blu o verde e 'blu' l'evidenza che il testimone ha riferito che il taxi era blu; il Teorema di Bayes dice:

$$
\begin{aligned}
P(B \mid \text{blu}) &= [P(\text{blu} \mid B) \times P(B)] / [P(\text{blu} \mid B) \times P(B)] + [P(\text{blu} \mid V) \times P(V)] \\
&= [(0.80) \times (0.15)] / [(0.80) \times (0.15)] + [(0.20) \times (0.85)] \\
&= 0.12 / (0.12 + 0.17) = 0.41
\end{aligned}
$$

Negli esperimenti fatti, la maggioranza dei soggetti coinvolti risponde 80%, ma quando la domanda viene posta eliminando l'informazione sull'attendibilità del testimone, quasi tutti i soggetti rispondono che la probabilità è 15%[9].

L'utilizzo del Teorema di Bayes permette anche di evitare la celebre 'fallacia dell'accusatore', il cui primo esempio conosciuto occorse nel caso Dreyfus[10]. Una delle prove a carico di Albert Dreyfus portate al processo era un *memorandum*, noto come il *bordereau*, che i servizi segreti francesi avevano recuperato dalla spazzatura dell'ambasciata tedesca. Secondo l'accusa il *bordereau* era stato scritto da Dreyfus e conteneva un messaggio in codice, e uno degli argomenti portati dall'accusa si basava sull'esame di due caratteristiche del documento, la posizione di alcune parole nel testo e la frequenza con cui comparivano. Nel corso del processo del 1894 Alphonse Bertillon, colui che è considerato uno dei fondatori della scienza forense, depose come teste dell'accusa, testimoniando come la piccolissima probabilità che tali caratteristiche si presentassero per caso in un testo qualunque andasse a favore dell'ipotesi che fossero intenzionali e dunque, prova che il testo fosse scritto in codice.

[9] KAHNEMAN, op. cit., 181 ss.
[10] GARBOLINO, op. cit., 103 ss.

Nel secondo processo d'appello, nel 1904, venne richiesto un parere a un comitato di saggi composto da tre matematici membri dell'Accademia delle scienze, fra i quali Henri Poincaré. Il parere del comitato sulla perizia probabilistica di Bertillon venne recepito poi dalla sentenza della Corte di cassazione nel 1906. I tre saggi rilevarono come la perizia commettesse la *fallacia dell'accusatore* (anche se Poincaré e i suoi colleghi non usarono questo termine che sarà coniato molto più tardi), cioè l'errore di scambiare la *verosimiglianza* dell'ipotesi della difesa, ovvero che le caratteristiche sono presenti per caso dato che Dreyfus è innocente, cioè *P(E/non H)*, con la probabilità *a posteriori* della stessa ipotesi, cioè *P(non H/E)*. L'intervento dei tre saggi segnava un passaggio molto importante per la neonata scienza forense perché formulava per la prima volta la regola per cui il valore della prova è espresso dal rapporto delle verosimiglianze. Poincaré era un fautore del Teorema di Bayes che, per un francese dell'epoca, era conosciuto come il Teorema di Laplace, perché anche il grande scienziato francese aveva dato, nel suo saggio *Sulla probabilità delle cause* del 1774, una formula analoga a quella del matematico inglese[11].

Per tornare alle aule dei tribunali dei nostri giorni, supponiamo che un *match* sia stato dichiarato fra un campione di sangue del sospettato e il campione di sangue prelevato dall'abito della vittima, e che la probabilità di una coincidenza casuale, supposto che l'origine del campione non sia il sospettato, sia pari a 1 su 7 milioni. Immaginiamo che la deposizione dell'esperto in aula sia del seguente tenore: "Si può dire che il sangue sul vestito della vittima potrebbe essere di qualcun altro che non sia il sospettato con probabilità 1 su 7 milioni".

Cosa potrebbero pensare, di fronte a una testimonianza resa in tal modo, un giudice o un giurato? Potrebbero facilmente essere indotti a pensare che la probabilità che il sangue *non* sia quello del sospettato è 1 su 7 milioni, incorrendo così nella "fallacia dell'accusatore". Ma il perito ha riferito la probabilità di osservare una coincidenza (*E*) se l'origine della traccia non è il sospettato (*non H*), *P(E/non H)*, non ha riferito che la probabilità che l'origine della traccia *non* sia il sospettato data l'osservazione della coincidenza, cioè *P(non H/E)*.

[11] Thomas Bayes molto probabilmente scrisse il suo celebre saggio attorno al 1749-50, ma venne pubblicato solamente postumo nel 1763.

Una comunicazione corretta ed efficace del valore probatorio è un problema di grande importanza e l'utilizzo del Teorema di Bayes per comunicare il valore probatorio permette di distinguere chiaramente le verosimiglianze dalle probabilità *a posteriori*.