

RAGIONARE SU PROBABILITA', POSSIBILITA' E FREQUENZE

1. Introduzione

Lo scopo di questa dispensa è quella di presentare alcuni recenti sviluppi della ricerca sul ragionamento probabilistico. Come abbiamo visto nella lezione precedente (Lezione 6 del Blocco “Ragionamento”), negli ultimi decenni numerose ricerche sperimentali hanno dimostrato i limiti delle nostre capacità di valutare eventi incerti. In particolare, i lavori di Kahneman e Tversky hanno dimostrato che i nostri giudizi probabilistici sono spesso guidati da procedure euristiche (v. ad es. Tversky & Kahneman, 1974) che, in qualche caso, ci portano a formulare giudizi probabilistici sistematicamente scorretti (i cosiddetti “biases” di giudizio). Basti ricordare a questo proposito, la fallacia della congiunzione prodotta dall’attivazione dell’euristica della disponibilità (v. il problema di Linda, Lezione 6 e Capitolo 3 di Girotto, 1994). La dimostrazione sperimentale dei limiti del nostro ragionamento ha contribuito a mettere in crisi la visione tradizionale del pensiero umano (quella difesa dalla teoria della scelta razionale), che presupponeva una nostra impeccabile capacità di fornire giudizi probabilistici corretti (v. per una rassegna, Piattelli-Palmarini, 1994). Negli ultimi anni, però, alcuni ricercatori (es. Cosmides & Tooby, 1996; Gigerenzer, 2003; Gigerenzer & Hoffrage, 1995) hanno proposto una teoria ancora più pessimistica del ragionamento probabilistico.

2. La teoria evoluzionista-frequentista

Secondo la teoria evoluzionista-frequentista, la mente umana sarebbe intrinsecamente incapace di ragionare correttamente sulle probabilità. L'argomento evoluzionista che sta alla base di questa teoria può essere riassunto come segue:

La probabilità di un evento singolo non è osservabile.

L'evoluzione naturale non può aver prodotto dei meccanismi cognitivi per elaborare delle informazioni non osservabili.

Quindi:

L'evoluzione naturale non può aver prodotto dei meccanismi cognitivi per elaborare l'informazione probabilistica.

Secondo i difensori della teoria evoluzionista, questa conclusione ha una solida base sperimentale. Si può infatti dimostrare che la grandissima maggioranza delle persone interrogate non è in grado risolvere problemi che presentano informazioni probabilistiche e che richiedono la valutazione della probabilità di un evento singolo. Considerate il problema 1:

Ecco i risultati di una ricerca condotta su una malattia e sul test che serve a diagnosticarla:

-Una persona sottoposta al test ha il 4% di probabilità di aver contratto la malattia.

-Se una persona ha contratto la malattia, ha il 75% di probabilità di avere una reazione positiva al test.

-Se una persona non ha contratto la malattia, ha comunque il 12.5% di probabilità di avere una reazione positiva al test.

Immagina che Paolo venga sottoposto al test. Se ha una reazione positiva, qual è la probabilità che abbia contratto la malattia?

La grande maggioranza dei partecipanti cui viene presentato questo problema, compresi studenti e docenti di prestigiose facoltà americane di medicina (v. Casscells, Schoenberger & Grayboys, 1978), non è in grado di fornire la risposta corretta, cioè "20%"¹. La risposta erronea più comune è "75%". Si tratta, come potete constatare, di

¹ Una soluzione formale del problema della malattia consiste nell'applicare la regola di Bayes (v. Girotto, 1994, Capitolo 3). Definendo H l'ipotesi "Paolo malato", $\neg H$ l'ipotesi "Paolo sano", e D il dato "Paolo positivo", la probabilità condizionale che Paolo sia malato se risulta positivo, cioè $p(H|D)$, sarà uguale a:

$$\begin{aligned} p(H|D) &= \\ &= \frac{p(D|H) \times p(H)}{[p(D|H) \times p(H) + p(D|\neg H) \times p(\neg H)]} \\ &= \frac{(75\% \times 4\%)}{[75\% \times 4\% + 12,5\% \times (100\% - 4\%)]} \end{aligned}$$

una valutazione molto lontana da quella corretta. Questo tipo di errori indica che anche persone potenzialmente esperte in valutazioni diagnostiche, come per esempio i medici, non sono sempre in grado di produrle (v. Eddy, 1982). Da un punto di vista teorico, questo tipo di errori sembra corroborare l'argomento evoluzionista secondo cui la mente umana è intrinsecamente incapace di valutare correttamente la probabilità di casi singoli.

Sulla base dell'ipotesi evoluzionista, però, si deve supporre che la selezione naturale abbia dotato la mente umana di meccanismi cognitivi atti ad elaborare le informazioni osservabili. Abbiamo visto sopra che non possiamo attribuire ai nostri antenati ominidi la capacità di ragionare sulle probabilità, visto che queste non sono osservabili. Per esempio, i nostri antenati non potevano certo valutare la probabilità di catturare una preda in una battuta di caccia, formulando una frase come la seguente:

“La probabilità di tornare con la preda andando a caccia nella foresta è del 20%”.

I nostri antenati, tuttavia, potevano osservare la ripetizione di eventi reali e registrarne le frequenze. Per esempio, è possibile che qualcuno di loro abbia prodotto una frase come la seguente

“Siamo andati a caccia 20 volte nella foresta e siamo tornati 4 volte con la preda”.

In altri termini, possiamo attribuire anche ai nostri antenati ominidi la capacità di ragionare su frequenze. Di conseguenza, facendo l'ipotesi che la nostra mente sia dotata di un meccanismo innato per trattare le frequenze osservate, possiamo predire che le persone non esperte saranno in grado di risolvere problemi che presentano informazioni su frequenze e che chiedono di predire frequenze. Considerate, ad esempio, il problema 2:

Ecco i risultati di una ricerca condotta su una malattia e sul test che serve a diagnosticarla:

-4 persone su 100 avevano contratto la malattia.

-3 delle 4 persone che avevano contratto la malattia hanno avuto una reazione positiva al test.

-12 delle 96 persone che non avevano contratto la malattia hanno comunque avuto una reazione positiva al test.

$$= 3\% / (3\% + 12\%)$$
$$= 20\%$$

100 persone sono ora sottoposte al test. In questo gruppo di 100 persone, ci attendiamo che ___ persone avranno una reazione positiva, tra le quali ___ avranno contratto la malattia.

Rispetto al problema 1, nel quale le informazioni erano relative a probabilità di casi singoli (per es. “Una persona ha il 4% di probabilità di aver contratto la malattia”), nel problema 2 tutte le informazioni riguardano frequenze di osservazioni (per es. “4 persone su 100 avevano contratto la malattia”). Inoltre, invece di porre una domanda sulla probabilità di un caso singolo (“Se Paolo ha una reazione positiva, qual è la probabilità che abbia contratto la malattia?”), questo problema chiede di predire la frequenza delle persone ammalate e con reazione positiva sul totale delle persone con reazione positiva. Ora, come predetto dai sostenitori della teoria evoluzionista-frequentista, una buona parte dei partecipanti risolve correttamente questa versione del problema della malattia. Dato che, sul totale di 100 persone osservate, c'erano 3 persone malate e positive e 12 sane e positive, sembra abbastanza semplice predire che, nel nuovo campione, ci saranno in tutto 15 (cioè 3 + 12) persone con reazione positiva e che 3 di queste saranno realmente ammalate. Risultati come questo dimostrano, quindi, che si possono rendere semplici dei difficili problemi di ragionamento probabilistico, trasformandoli in problemi di frequenze. Più in generale, questi risultati sembrano dimostrare che la mente umana, anche se intrinsecamente incapace di ragionare sulle probabilità di casi singoli, possiede un meccanismo innato per ragionare sulle frequenze.

3. Il ragionamento estensionale sulle possibilità

La tesi evoluzionista è molto affascinante e, in effetti, ha convinto molti non specialisti di psicologia del pensiero (ad es. Pinker, 1997/2000). E' però veramente fondata la conclusione che i sostenitori di tale tesi hanno tratto dai risultati appena presentati? In altre parole, la mente umana è veramente cieca alle probabilità? Non c'è alcun modo per i non esperti per valutare la probabilità di un evento singolo?

Come abbiamo visto nella Lezione 4, nel campo del ragionamento deduttivo la teoria che sembra meglio spiegare le capacità inferenziali umane è quella dei modelli mentali. Secondo un'ipotesi derivata da questa teoria, le inferenze probabilistiche spontanee sono tratte, al pari di quelle deduttive, in modo estensionale (Johnson-Laird, Legrenzi, Girotto, Sonino-Legrenzi & Caverni, 1999). In altri termini, le persone non esperte ragionano non applicando regole (equivalenti a quelle della logica classica o a quelle del calcolo delle probabilità), ma sulla base di rappresentazioni (modelli) mentali di possibilità. Considerate il seguente un problema:

In una mano di carte c'è un asso o un re o entrambe le carte. Qual è la probabilità dell'evento seguente “Nella mano c'è un asso e un re”?

Un esperto potrebbe rispondere che la probabilità dell'evento in questione può variare da 0 a 1. Quasi tutti i partecipanti (studenti universitari non esperti di calcolo probabilistico) dell'esperimento di Johnson-Laird e colleghi risposero, invece, che la probabilità dell'evento in questione era pari a $1/3$. Secondo la teoria dei modelli mentali, le persone non esperte risolvono questo problema rappresentandosi il contenuto del primo enunciato nel modo seguente:

asso	
	re
asso	re

In questa rappresentazione, le tre righe indicano tre diversi modelli mentali, ognuno dei quali rappresenta una possibilità vera: la prima riga rappresenta la possibilità che nella mano ci sia un asso, la seconda che ci sia un re, la terza che ci sia un asso e un re. Dato che solo uno dei tre modelli rappresenta la possibilità che ci sia un asso e un re, la conclusione che le persone ne traggono è che c'è $1/3$ di probabilità che nella mano ci sia un asso e un re. Insomma, i partecipanti traggono delle inferenze probabilistiche calcolando la proporzione di modelli, tra quelli costruiti, che rappresentano l'evento che devono valutare. Queste inferenze sono "estensionali", perché tratte sulla base dell'enumerazione delle diverse possibilità in cui gli eventi possono verificarsi, e "non esperte" dato che non dipendono dall'applicazione delle regole del calcolo probabilistico.

La capacità dei non esperti di ragionare in modo estensionale sulle probabilità è stata dimostrata anche con problemi nei quali le informazioni probabilistiche sono presentate in termini numerici. In particolare, è stato dimostrato che le persone non esperte sono in grado di risolvere versioni del problema della malattia come il seguente problema 3:

Ecco i risultati di una ricerca condotta su una malattia e sul test che serve a diagnosticarla:

- Una persona sottoposta al test ha 4 possibilità di aver contratto la malattia.
- 3 di queste 4 possibilità di aver contratto la malattia sono associate ad una reazione positiva al test.
- 12 delle rimanenti 96 possibilità di non aver contratto la malattia sono ugualmente associate ad una reazione positiva al test.

Immagina che Paolo venga sottoposto al test. Sul totale di 100 possibilità, Paolo avrà ___ possibilità di avere una reazione positiva al test, ___ delle quali saranno associate alla presenza della malattia.

Questa versione del problema della malattia mantiene le caratteristiche fondamentali del problema originale (il problema 1). Si tratta, cioè, di un problema che presenta informazioni relative alla probabilità di un caso singolo (la possibilità che una persona contragga la malattia) e richiede la valutazione della probabilità di un caso singolo (Paolo ha contratto la malattia). Questo problema, però, si differenzia dal

problema 1 perché esprime le probabilità in termini di possibilità, cioè con dei numeri interi e non con delle percentuali, proprio come il problema 2, che esprimeva le frequenze con dei numeri interi. Inoltre, il problema 3 pone una domanda che ha la stessa forma della domanda posta nel problema 2, cioè chiede esplicitamente ai partecipanti il calcolo di un rapporto. In particolare, il problema 3 chiede prima di calcolare il denominatore (cioè l'insieme delle possibilità di avere una reazione positiva) e poi di calcolare il numeratore (cioè il sottoinsieme di possibilità che la reazione positiva sia accompagnata dalla presenza della malattia). Come la domanda del problema 2, anche quella del problema 3 chiede quindi di calcolare separatamente i due termini del rapporto di probabilità che conduce alla risposta corretta. Sulla base dei dati presentati, che riguardano un insieme finito di 100 possibilità, dal quale è facile ricavare le possibilità rilevanti per la soluzione del problema (3 e 12), e sulla base della domanda posta, che chiede di calcolare un rapporto (non una percentuale), il problema dovrebbe essere facilmente risolvibile in modo estensionale. In effetti, la maggioranza delle persone interrogate risponde correttamente che Paolo ha 15 possibilità di avere una reazione positiva e che 3 di queste possibilità saranno associate alla malattia (v. Girotto & Gonzalez, 2001/Esperimento 1). Insomma, anche un problema di probabilità può essere risolto dalle persone non esperte in modo estensionale, cioè considerando ed enumerando le possibilità di accadimento dei vari eventi.

Vi è un'ulteriore possibilità per facilitare il ragionamento in questo tipo di problemi. Invece di chiedere ai partecipanti di valutare la probabilità che Paolo sia ammalato se ha una reazione positiva, possiamo chiedere loro di valutare la probabilità che Paolo sia ammalato contro la probabilità che sia sano. In altre parole, invece di chiedere il calcolo di un rapporto, si può chiedere la semplice assegnazione di possibilità numeriche a due ipotesi alternative. Per esempio, possiamo chiedere:

Immagina che Paolo venga sottoposto al test. Se ha una reazione positiva, ci saranno ___ possibilità che la malattia sia associata a questa reazione, contro ___ possibilità che nessuna malattia sia associata a questa reazione.

Questo tipo di domanda, che possiamo definire comparativa o distributiva, riflette un modo usuale di esprimere l'incertezza nella vita quotidiana. Per esempio, quando facciamo una scommessa, "distribuiamo" il nostro giudizio su due ipotesi diverse: "Scommetto 20 a 1 che la squadra A batterà la squadra B". Tali espressioni appaiono del tutto naturali, e ne troviamo delle tracce anche in opere precedenti lo sviluppo della moderna teoria della probabilità, come testimonia, per esempio, il seguente verso di Shakespeare:

"Twenty to one then he is shipped already".
("Venti a uno, allora, è già partito con la nave").
(Shakespeare, The Two Gentlemen of Verona, I, 1)

Dato che il problema 3 indica le possibilità di avere la reazione positiva se si è malati (3) e se non si è malati (12), non dovrebbe essere difficile rispondere alla domanda distributiva, assegnando 3 possibilità all'ipotesi di avere la malattia e 12 possibilità all'ipotesi di non averla. In effetti la quasi totalità delle persone interrogate (85%) risolve correttamente il problema della malattia quando la domanda è di tipo distributivo (Girotto & Gonzalez, 2001/Esperimento 4).

Questi e altri risultati (v. Girotto & Gonzalez, 2002; Johnson-Laird et al., 1999) dimostrano che gli errori di ragionamento probabilistico non dipendono dal tipo d'informazione da elaborare (probabilità di casi singoli vs. frequenze di osservazione), ma dal modo in cui viene presentata l'informazione nei problemi e dal tipo di domanda posta ai partecipanti. Se le persone possono applicare le loro intuizioni estensionali alla rappresentazione di un insieme finito di unità discrete (come le possibilità o le frequenze), sono in grado di trarre delle inferenze corrette, indipendentemente dalla natura dell'informazione trattata. Più in generale, questi risultati dimostrano l'esistenza di una competenza estensionale nel ragionamento sulle possibilità.

4. Bibliografia

- Casscells, W., Schoenberger, A., & Grayboys, T. (1978). Interpretation by physicians of clinical laboratory results. New England Journal of Medicine, 299, 999-1000
- Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. Cognition, 58, 1-73.
- Eddy, D. M. (1982). Probabilistic reasoning in clinical medicine: Problems and opportunities. In D. Kahneman, P. Slovic, & A. Tversky (Eds) Judgment under uncertainty: Heuristics and biases. Cambridge: Cambridge University Press.
- Gigerenzer, G. (2003) Quando i numeri ci fanno sbagliare, Milano: Cortina.
- Gigerenzer, G., & Hoffrage, U. (1995). How to improve bayesian reasoning without instruction: Frequency format. Psychological Review, 102, 684-704.
- Girotto, V. (1994). Il ragionamento. Bologna: Il Mulino.
- Girotto, V., & Gonzalez, M. (2001). Solving probabilistic and statistical problems: A matter of information structure and question form. Cognition, 78, 247-276.

Girotto, V., & Gonzalez, M. (2002). Chances and frequencies in probabilistic reasoning. Cognition, 84, 353-359.

Johnson-Laird, P. N., Legrenzi, P., Girotto, V., Sonino-Legrenzi, M., & Caverni, J. P. (1999). Naive probability: A model theory of extensional reasoning.

Psychological Review, 106, 62-88.

Piattelli-Palmarini, M. (1994). Inevitable illusions. New York, Wiley.

Pinker, S. (1997). How the mind works. New York: Norton; trad. it. Come funziona la mente. Milano: Mondadori, 2000.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. Science, 185, 1124-1131.