CAPITOLO

15

Probabilità condizionata e formula di Bayes

Un dato nuovo non cancella quello che sapevamo: lo pesa insieme a ciò che sapevamo già. (testo della redazione)

OBIETTIVI

1. Interpretare la probabilità condizionata come una nuova probabilità sullo spazio ridotto e distinguere $p(A|B)$ da $p(B|A)$.
2. Applicare il teorema della probabilità composta e riconoscere gli eventi indipendenti, distinguendoli dagli eventi incompatibili.
3. Dimostrare e applicare il teorema delle probabilità totali, anche con i diagrammi ad albero.
4. Dimostrare la formula di Bayes e usarla per aggiornare una probabilità alla luce di nuove informazioni.
5. Analizzare test diagnostici, controlli di qualità e altri problemi reali con le tabelle di frequenze naturali.
6. Riconoscere le ipotesi di un modello probabilistico e valutare quanto le conclusioni dipendono dalla probabilità a priori.

INDICE DEL CAPITOLO

15.1 La probabilità condizionata 2
15.2 Probabilità composta e indipendenza 6
15.3 Il teorema delle probabilità totali 8
15.4 La formula di Bayes 10
15.5 Le applicazioni della formula di Bayes 15
15.6 Probabilità e modello matematico 21

Matematica 4 · Parte V · Calcolo combinatorio e probabilità · versione 1.0
Capitolo 15: Probabilità condizionata e formula di Bayes · pagine 417–444 del volume completo · estratto per lo studio

Un test medico risulta positivo. Il foglio dice che il test è «affidabile al 99%». Quanto è probabile, allora, che la persona sia davvero malata? Quasi tutti rispondono «al 99%», e quasi tutti sbagliano: con dati del tutto plausibili, che useremo in questo capitolo, la risposta corretta è circa 17%. L'errore non è di calcolo ma di ragionamento: si confonde la probabilità di un risultato positivo *sapendo* che la persona è malata con la probabilità che la persona sia malata *sapendo* che il risultato è positivo. Sono due domande diverse, e possono avere risposte lontanissime.

In *Matematica 2* hai incontrato la probabilità condizionata, la probabilità del prodotto di due eventi e l'indipendenza. In questo capitolo riprendiamo questi concetti con più rigore e li portiamo fino al risultato che le *Indicazioni nazionali* nominano espressamente: la **formula di Bayes** e le sue applicazioni. È la regola che dice come una probabilità deve cambiare quando arriva un'informazione nuova, ed è usata ogni giorno in medicina, nell'industria, nei tribunali e nei programmi che filtrano la posta elettronica. Alla fine torneremo sul concetto di modello matematico: una probabilità calcolata è sempre la conseguenza di ipotesi scelte da noi, e cambiare le ipotesi cambia la conclusione.

15.1 La probabilità condizionata

15.1.1 Quando un'informazione cambia la probabilità

Lanciamo due dadi equilibrati, uno rosso e uno blu, e scriviamo l'esito come la coppia $(i; j)$ dei numeri usciti. Lo spazio campionario Ω contiene 36 coppie equiprobabili. Consideriamo gli eventi

- A = «esce almeno un 6», formato da 11 coppie;
- B = «la somma dei due numeri è almeno 10», formato dalle 6 coppie (4; 6), (5; 5), (6; 4), (5; 6), (6; 5), (6; 6).

Prima del lancio $p(A) = \dfrac{11}{36}$, cioè circa il 31%. Supponiamo ora che un amico guardi i dadi e ci dica soltanto: «la somma è almeno 10». Che probabilità dobbiamo assegnare ad A? Gli esiti possibili non sono più 36: sono le 6 coppie di B, ancora equiprobabili fra loro, e di queste 5 contengono un 6. La nuova probabilità è $\dfrac{5}{6}$, circa l'83%: l'informazione ha quasi triplicato la probabilità di A. La situazione è rappresentata nella Figura 15.1.

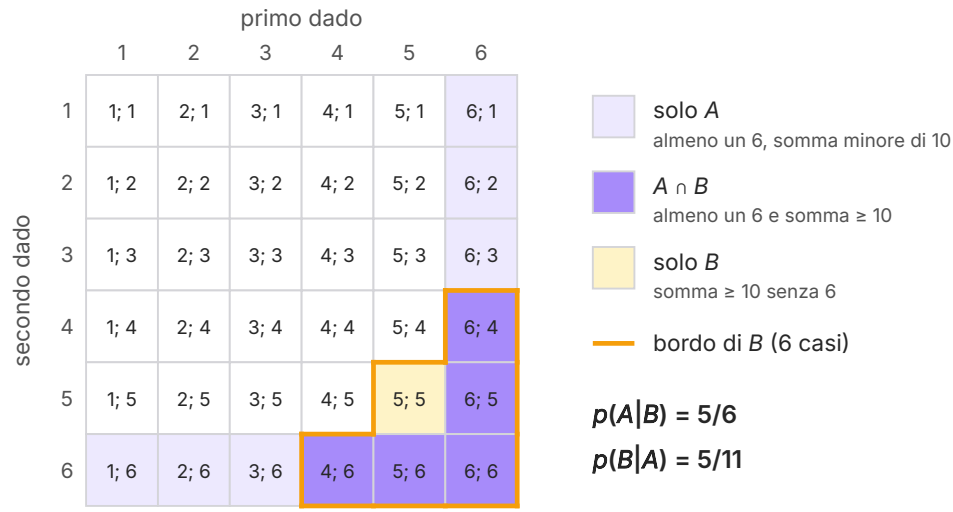


Figura 15.1 Lo spazio campionario del lancio di due dadi. Le caselle viola formano l'evento A («almeno un 6», 11 casi); il bordo arancione racchiude l'evento B («somma almeno 10», 6 casi). Sapendo che B si è verificato, i casi possibili sono le 6 caselle di B, e 5 di esse stanno anche in A: $p(A|B) = 5/6$. Sapendo invece che si è verificato A, i casi possibili sono le 11 caselle viola, e 5 di esse stanno in B: $p(B|A) = 5/11$.

Abbiamo calcolato il rapporto fra i casi di $A \cap B$ e i casi di B. Dividendo numeratore e denominatore per 36 lo stesso rapporto si scrive $\frac{5}{6} = \frac{5/36}{6/36}$, cioè come rapporto fra due probabilità calcolate in Ω. Questa forma non dipende dall'equiprobabilità degli esiti, e diventa la definizione generale.

IN SINTESI

Definizione. Siano A e B due eventi di uno stesso spazio campionario, con $p(B) > 0$. La **probabilità condizionata** di A dato B (si legge «p di A dato B») è

$$p(A|B) = \frac{p(A \cap B)}{p(B)}.$$

Se $p(B) = 0$ la probabilità condizionata a B non è definita.

15.1.2 Una nuova probabilità sullo spazio ridotto

Fissato un evento B con $p(B) > 0$, la formula precedente associa a ogni evento A un numero $p(A|B)$. Questa nuova funzione degli eventi è a sua volta una probabilità: soddisfa infatti i tre assiomi di Kolmogorov (vedi il capitolo precedente).

- **Non negatività.** $p(A|B) \geq 0$, perché è il rapporto fra un numero non negativo e un numero positivo.
- **Normalizzazione.** $p(\Omega|B) = \dfrac{p(\Omega \cap B)}{p(B)} = \dfrac{p(B)}{p(B)} = 1$.
- **Additività.** Se A_1 e A_2 sono incompatibili, lo sono anche $A_1 \cap B$ e $A_2 \cap B$; quindi $p((A_1 \cup A_2) \cap B) = p(A_1 \cap B) + p(A_2 \cap B)$ e, dividendo per $p(B)$, $p(A_1 \cup A_2|B) = p(A_1|B) + p(A_2|B)$.

Di conseguenza valgono per $p(\cdot|B)$ tutti i teoremi che valgono per p: per esempio $p(\overline{A}|B) = 1 - p(A|B)$, e $p(A \cup C|B) = p(A|B) + p(C|B) - p(A \cap C|B)$. Possiamo pensare che l'informazione «B si è verificato» trasformi B nel nuovo spazio campionario, lo **spazio ridotto**: gli esiti fuori da B diventano impossibili, quelli dentro B conservano le proporzioni che avevano fra loro, ma vengono «ingranditi» in modo che la loro probabilità totale torni a essere 1. È proprio ciò che fa la divisione per $p(B)$.

Osserva anche il caso particolare in cui $A \subset B$: allora $A \cap B = A$ e $p(A|B) = p(A) / p(B) \geq p(A)$. Sapere che si è verificato un evento più ampio non può far diminuire la probabilità di un evento che vi è contenuto.

15.1.3 Condizionare non vuol dire causare

Le probabilità condizionate compaiono in modo naturale nelle tabelle a doppia entrata della statistica. In *Matematica 3* hai calcolato le frequenze relative condizionate di un carattere rispetto all'altro: se si sceglie a caso un'unità della popolazione, quelle frequenze diventano esattamente probabilità condizionate.

Tabella 15.1 Gli 800 studenti di un liceo, classificati in base a due caratteri (dati inventati): S = «suona uno strumento», V = «ha almeno 8 in matematica».

	V	$\overline{V}$	Totale
S	72	128	200
$\overline{S}$	108	492	600
Totale	180	620	800

ESEMPIO SVOLTO

Leggere una tabella a doppia entrata. Si sceglie a caso uno studente fra quelli della Tabella 15.1. Calcola $p(V)$, $p(V|S)$, $p(V|\overline{S})$ e $p(S|V)$, e commenta i risultati.

Probabilità semplice. Gli studenti con almeno 8 sono 180 su 800: $p(V) = \dfrac{180}{800} = 0{,}225$.

Condizionate a S. Fra i 200 studenti che suonano, 72 hanno almeno 8: $p(V|S) = \dfrac{72}{200} = 0{,}36$. Con la definizione: $p(V|S) = \dfrac{72/800}{200/800} = 0{,}36$, lo stesso numero.

Condizionata a $\overline{S}$. Fra i 600 che non suonano, 108 hanno almeno 8: $p(V|\overline{S}) = \dfrac{108}{600} = 0{,}18$.

Condizionata inversa. Fra i 180 studenti con almeno 8, quelli che suonano sono 72: $p(S|V) = \dfrac{72}{180} = 0{,}4$.

Commento. In questa scuola chi suona uno strumento ha una probabilità doppia di avere almeno 8 rispetto a chi non suona (0,36 contro 0,18). I dati però non dicono *perché*: può darsi che lo studio della musica aiuti la matematica, ma può darsi anche che le stesse abitudini di studio, o lo stesso ambiente familiare, favoriscano entrambe le cose. Una probabilità condizionata misura un'associazione, non una causa. Nota infine che $p(V|S) = 0{,}36$ e $p(S|V) = 0{,}4$ sono numeri diversi.

L'ultima osservazione dell'esempio è il messaggio più importante del capitolo. Nella Figura 15.1 si vede bene perché $p(A|B)$ e $p(B|A)$ sono diverse: hanno lo stesso numeratore, i 5 casi di $A \cap B$, ma denominatori diversi, perché nel primo caso lo spazio ridotto è B (6 casi) e nel secondo è A (11 casi). In formule, poiché $p(A \cap B) = p(B)\cdot p(A|B) = p(A)\cdot p(B|A)$, si ha

$$p(A|B) = p(B|A) \cdot \frac{p(A)}{p(B)},$$

quindi le due probabilità condizionate coincidono solo quando $p(A) = p(B)$ (oppure quando entrambe sono nulle). Il rapporto $p(A) / p(B)$ che le collega è, come vedremo, il cuore della formula di Bayes.

In parole: $p(A|B)$ e $p(B|A)$ rispondono a domande diverse: «fra i casi in cui accade B, quanto spesso accade A?» e «fra i casi in cui accade A, quanto spesso accade B?». La probabilità che un giocatore di basket sia alto più di 1,90 m è grande; la probabilità che una persona alta più di 1,90 m giochi a basket è piccola. Prima di calcolare, chiediti sempre quale sia lo spazio ridotto.

15.2 Probabilità composta e indipendenza

15.2.1 Il teorema della probabilità composta

Moltiplicando per $p(B)$ entrambi i membri della definizione di probabilità condizionata si ottiene una formula per la probabilità che due eventi si verifichino *entrambi*. Scambiando i ruoli di A e B si ottiene una seconda forma, utile quando è nota la condizionata nell'altro verso.

IN SINTESI

Teorema della probabilità composta. Ipotesi: A e B eventi di uno stesso spazio campionario, con $p(A) > 0$ e $p(B) > 0$. **Tesi:**

$$p(A \cap B) = p(B) \cdot p(A|B) = p(A) \cdot p(B|A).$$

Per tre eventi, se $p(A \cap B) > 0$: $p(A \cap B \cap C) = p(A) \cdot p(B|A) \cdot p(C|A \cap B)$. In generale la probabilità che si verifichino più eventi uno dopo l'altro è il prodotto delle probabilità di ciascuno, condizionate a tutti quelli precedenti.

La prima uguaglianza è la definizione di $p(A|B)$ moltiplicata per $p(B)$; la seconda è la definizione di $p(B|A)$ moltiplicata per $p(A)$, ricordando che $B \cap A = A \cap B$. Per tre eventi applichiamo due volte il teorema: $p(A \cap B \cap C) = p(A \cap B) \cdot p(C|A \cap B) = p(A) \cdot p(B|A) \cdot p(C|A \cap B)$. c.v.d.

È la regola che usi quando percorri un diagramma ad albero: su ogni ramo si scrive la probabilità condizionata a tutto ciò che è accaduto prima, e la probabilità di un percorso è il prodotto delle probabilità dei suoi rami. Il teorema della probabilità composta è dunque la giustificazione della regola dei percorsi.

ESEMPIO SVOLTO

Tre carte dello stesso seme. Da un mazzo di 40 carte (4 semi da 10 carte) si estraggono tre carte una dopo l'altra, senza reimmissione. Calcola la probabilità che siano tutte e tre di denari.

Eventi. Sia D_k = «la k-esima carta è di denari», per $k = 1, 2, 3$. Cerchiamo $p(D_1 \cap D_2 \cap D_3)$.

Rami dell'albero. $p(D_1) = \frac{10}{40}$. Se la prima è di denari, restano 39 carte di cui 9 di denari: $p(D_2|D_1) = \frac{9}{39}$. Se le prime due sono di denari, restano 38 carte di cui 8 di denari: $p(D_3|D_1 \cap D_2) = \frac{8}{38}$.

$$p(D_1 \cap D_2 \cap D_3) = \frac{10}{40} \cdot \frac{9}{39} \cdot \frac{8}{38} = \frac{720}{59\,280} = \frac{3}{247} \approx 0{,}0121.$$

Controllo con il calcolo combinatorio. Contando le terne non ordinate: $C_{10,3} / C_{40,3} = \frac{120}{9880} = \frac{3}{247}$. I due metodi danno lo stesso risultato, come abbiamo visto nel capitolo precedente per le estrazioni successive e contemporanee.

15.2.2 Eventi indipendenti

Può accadere che l'informazione «B si è verificato» non cambi affatto la probabilità di A. Lanciamo un dado e consideriamo A = «esce un numero pari» = {2, 4, 6} e B = «esce un numero minore o uguale a 4» = {1, 2, 3, 4}. Allora $p(A) = \frac{1}{2}$, mentre nello spazio ridotto B i casi favorevoli sono 2 e 4, e $p(A|B) = \frac{2}{4} = \frac{1}{2}$: sapere che B si è verificato non dice nulla su A. Se invece C = «esce un numero minore o uguale a 3», si ha $p(A|C) = \frac{1}{3} < p(A)$: C rende A meno probabile.

Diciamo allora che due eventi A e B sono **eventi indipendenti** se

$$p(A \cap B) = p(A) \cdot p(B).$$

Se $p(B) > 0$, la condizione equivale a $p(A|B) = p(A)$; se $p(A) > 0$, equivale a $p(B|A) = p(B)$. Due eventi che non sono indipendenti si dicono **eventi dipendenti**.

La definizione con il prodotto ha due vantaggi. È **simmetrica**: se A è indipendente da B, anche B è indipendente da A, cosa che con le condizionate andrebbe dimostrata. E ha senso anche quando uno dei due eventi ha probabilità nulla (in quel caso i due eventi risultano sempre indipendenti). Nel caso del dado, $p(A \cap B) = \frac{2}{6} = \frac{1}{3}$ e $p(A) \cdot p(B) = \frac{1}{2} \cdot \frac{4}{6} = \frac{1}{3}$: A e B sono indipendenti.

L'esempio del dado mostra che l'indipendenza è una proprietà *numerica* delle probabilità, non una proprietà fisica degli eventi: A e B riguardano lo stesso lancio dello stesso dado, eppure sono indipendenti. Nella pratica, però, l'indipendenza viene spesso *assunta* come ipotesi del modello, quando gli eventi riguardano esperimenti che non si influenzano: due lanci successivi di una moneta, le estrazioni con reimmissione, i pezzi prodotti da macchine diverse. In questi casi si usa direttamente la regola del prodotto. Quando gli eventi sono più di due, l'indipendenza richiede che la regola del prodotto valga per ogni gruppo di eventi, non solo per le coppie (ne vedrai un esempio negli esercizi).

ERRORE FREQUENTE

Incompatibili non vuol dire indipendenti. Molti pensano che due eventi «che non hanno niente in comune» siano indipendenti. È vero il contrario. Se A e B sono incompatibili e hanno entrambi probabilità positiva, allora $p(A \cap B) = 0$, mentre $p(A) \cdot p(B) > 0$: sono **dipendenti**. Infatti, se si verifica B, l'evento A diventa impossibile: $p(A|B) = 0 \neq p(A)$. L'informazione su B dice tutto su A, altro che indipendenza!

Un secondo errore è opposto: trattare come indipendenti le estrazioni *senza* reimmissione. Nell'esempio delle tre carte, moltiplicare $\frac{10}{40}$ per sé stesso tre volte darebbe $\frac{1}{64} \approx 0{,}0156$, un valore sbagliato: il risultato corretto è $\frac{3}{247} \approx 0{,}0121$.

15.3 Il teorema delle probabilità totali

15.3.1 Partizioni dello spazio campionario

In molti problemi lo spazio campionario si divide in modo naturale in casi che si escludono a vicenda: un pezzo è prodotto dalla macchina 1, dalla 2 o dalla 3; una persona è malata oppure sana; una pallina viene estratta dalla prima urna oppure dalla seconda. Questa situazione ha un nome.

Gli eventi $H_1, H_2, \ldots, H_n$ formano una **partizione** dello spazio campionario Ω se

- sono incompatibili a due a due: $H_i \cap H_j = \varnothing$ per $i \neq j$;
- la loro unione è Ω: $H_1 \cup H_2 \cup \ldots \cup H_n = \Omega$;
- ciascuno ha probabilità positiva: $p(H_i) > 0$.

In altre parole, si verifica sempre *uno e uno solo* degli eventi H_i. Ne segue $p(H_1) + p(H_2) + \ldots + p(H_n) = 1$.

L'esempio più semplice di partizione è formato da un evento H e dal suo contrario $\overline{H}$, purché entrambi abbiano probabilità positiva. Gli eventi H_i si chiamano spesso *ipotesi* (da qui la lettera H): sono le possibili «cause» o «situazioni» fra cui non sappiamo quale sia quella vera.

15.3.2 Il teorema

Supponiamo di conoscere le probabilità delle ipotesi e, per ciascuna ipotesi, la probabilità condizionata di un evento E. Possiamo allora calcolare la probabilità di E: è la media delle probabilità condizionate, «pesata» con le probabilità delle ipotesi.

IN SINTESI

Teorema delle probabilità totali. Ipotesi: $H_1, \dots, H_n$ partizione di Ω; E evento di Ω. **Tesi:**

$$p(E) = p(H_1) \cdot p(E|H_1) + p(H_2) \cdot p(E|H_2) + \dots + p(H_n) \cdot p(E|H_n) = \sum_{i=1}^{n} p(H_i) \cdot p(E|H_i).$$

Dimostrazione. Poiché le ipotesi ricoprono Ω, ogni esito di E sta in uno degli eventi H_i; quindi $E = (E \cap H_1) \cup (E \cap H_2) \cup \dots \cup (E \cap H_n)$. Gli eventi di questa unione sono incompatibili a due a due, perché lo sono le ipotesi. Per l'additività (estesa a n eventi) si ha

$$p(E) = p(E \cap H_1) + p(E \cap H_2) + \dots + p(E \cap H_n),$$

e per il teorema della probabilità composta ogni addendo è $p(E \cap H_i) = p(H_i) \cdot p(E|H_i)$, che è definito perché $p(H_i) > 0$. Sostituendo si ottiene la tesi. c.v.d.

Vediamo il teorema al lavoro. Gli studenti di quarta di un liceo (dati inventati) frequentano tre indirizzi: il corso di ordinamento (evento H_1, il 50% degli studenti), le scienze applicate (H_2, 30%) e l'indirizzo sportivo (H_3, 20%). Usano i mezzi pubblici per venire a scuola il 40% degli studenti del primo indirizzo, il 70% del secondo e il 50% del terzo. Qual è la probabilità che uno studente di quarta scelto a caso usi i mezzi pubblici (evento E)?

$$p(E) = 0{,}5 \cdot 0{,}4 + 0{,}3 \cdot 0{,}7 + 0{,}2 \cdot 0{,}5 = 0{,}20 + 0{,}21 + 0{,}10 = 0{,}51.$$

La Figura 15.2 dà una lettura geometrica del teorema. Rappresentiamo Ω con un rettangolo di area 1 e lo dividiamo in strisce verticali di larghezze $p(H_1)$, $p(H_2)$, $p(H_3)$. In ogni striscia coloriamo una parte alta quanto $p(E|H_i)$: l'area colorata è $p(H_i) \cdot p(E|H_i) = p(E \cap H_i)$, e la somma delle aree colorate è $p(E)$. Osserva che $p(E) = 0{,}51$ è compresa fra la più piccola e la più grande delle probabilità condizionate (0,4 e 0,7), come deve essere per una media pesata.

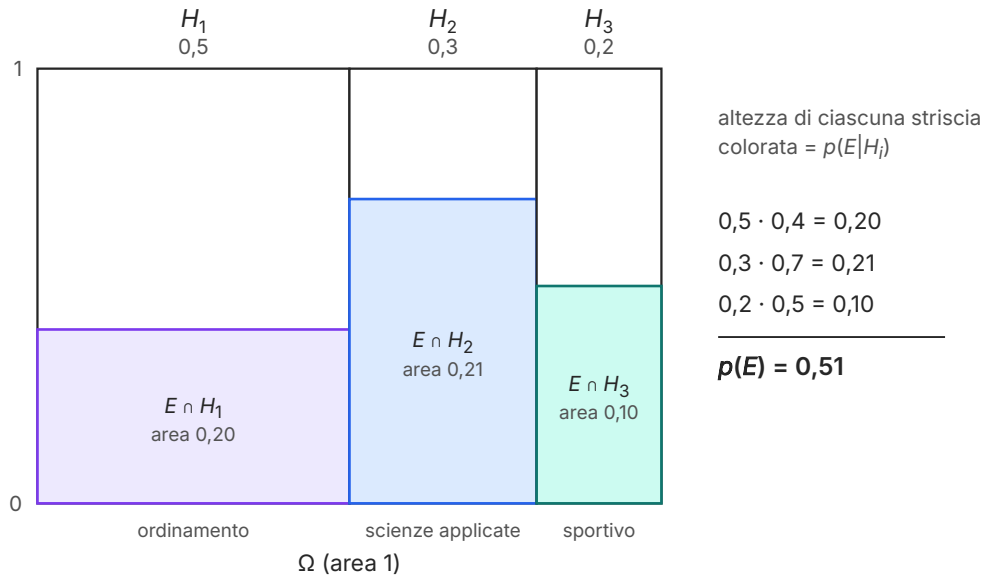


Figura 15.2 Il teorema delle probabilità totali come somma di aree (dati dell'esempio sugli studenti di quarta). Il rettangolo Ω ha area 1 ed è diviso nelle tre ipotesi, con larghezze 0,5, 0,3 e 0,2; in ciascuna striscia la parte colorata ha altezza $p(E|H_i)$ e area $p(E \cap H_i)$. Le aree colorate sommano a $p(E) = 0,51$.

Lo stesso calcolo si legge su un diagramma ad albero: dalla radice partono i rami delle ipotesi, da ciascuna ipotesi partono i rami E ed $\overline{E}$; la probabilità di E è la somma delle probabilità di tutti i percorsi che terminano in E. La Figura 15.4 mostra un albero di questo tipo.

15.4 La formula di Bayes

15.4.1 Dall'effetto alla causa

Il teorema delle probabilità totali va «dalle cause agli effetti»: conoscendo le probabilità delle ipotesi e la probabilità di E sotto ciascuna ipotesi, calcola la probabilità di E. Spesso la domanda è quella inversa. Abbiamo osservato E: quale ipotesi l'ha prodotto, e con quale probabilità?

Un esempio semplice. Abbiamo due urne: la prima contiene 3 palline bianche e 2 nere, la seconda 1 bianca e 4 nere. Lanciamo una moneta per scegliere l'urna e ne estraiamo una pallina, senza guardare quale urna sia stata scelta. La pallina è bianca. Qual è la probabilità che venga dalla prima urna? Prima dell'estrazione le due urne erano equiprobabili; ma la prima urna produce palline bianche più spesso della seconda, quindi l'osservazione «bianca» deve rendere più credibile la prima urna. Di quanto?

Siano U_1 e U_2 le due urne e W l'evento «esce una pallina bianca». Per il teorema delle probabilità totali, $p(W) = \frac{1}{2} \cdot \frac{3}{5} + \frac{1}{2} \cdot \frac{1}{5} = \frac{3}{10} + \frac{1}{10} = \frac{2}{5}$. Per la definizione di probabilità condizionata e per il teorema della probabilità composta,

$$p(U_1|W) = \frac{p(U_1 \cap W)}{p(W)} = \frac{p(U_1) \cdot p(W|U_1)}{p(W)} = \frac{3/10}{2/5} = \frac{3}{4}.$$

La probabilità della prima urna è salita da $\frac{1}{2}$ a $\frac{3}{4}$. La Figura 15.3 mostra il ragionamento con le aree: l'informazione «è uscita una pallina bianca» elimina le parti bianche del quadrato, e il nuovo spazio è la sola parte colorata, di area 4/10. Di questa, 3/10 stanno nella prima urna: i tre quarti.

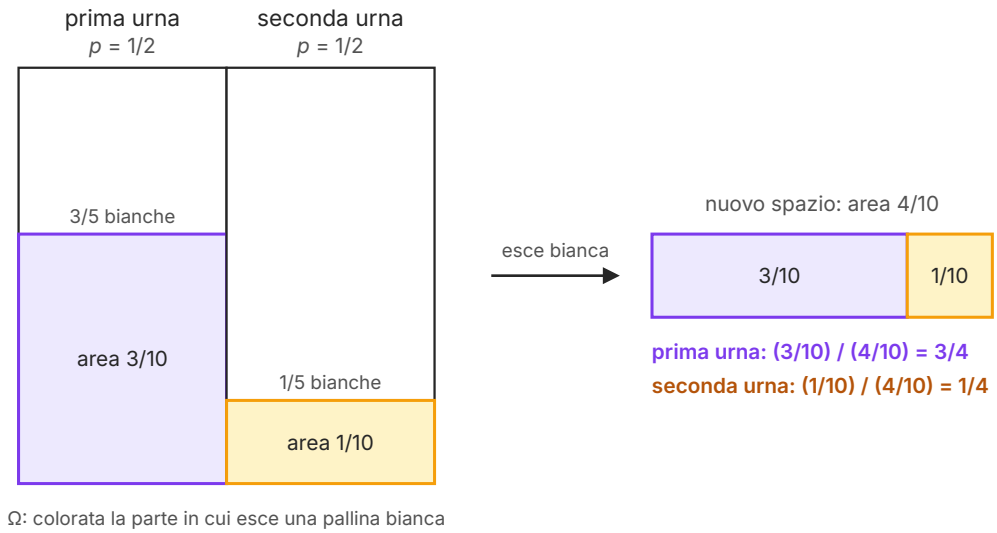


Figura 15.3 La formula di Bayes vista con le aree. A sinistra, il quadrato di area 1 è diviso fra le due urne (probabilità 1/2 ciascuna); in ogni metà è colorata la parte in cui esce una pallina bianca (3/5 e 1/5 della striscia). A destra, l'osservazione «bianca» lascia solo le parti colorate, di area totale 4/10: la frazione che spetta alla prima urna è (3/10) / (4/10) = 3/4.

Il ragionamento è del tutto generale, e la formula che ne risulta porta il nome di Thomas Bayes.

IN SINTESI

Teorema (formula di Bayes). Ipotesi: $H_1, \dots, H_n$ partizione di Ω; E evento con $p(E) > 0$. **Tesi:** per ogni $i = 1, \dots, n$

$$p(H_i|E) = \frac{p(H_i) \cdot p(E|H_i)}{p(E)} = \frac{p(H_i) \cdot p(E|H_i)}{\sum_{j=1}^{n} p(H_j) \cdot p(E|H_j)}.$$

Con due sole ipotesi, H ed $\overline{H}$: $p(H|E) = \frac{p(H) \cdot p(E|H)}{p(H) \cdot p(E|H) + p(\overline{H}) \cdot p(E|\overline{H})}$.

Dimostrazione. Per la definizione di probabilità condizionata, $p(H_i|E) = p(H_i \cap E) / p(E)$. Per il teorema della probabilità composta il numeratore è $p(H_i \cap E) = p(H_i) \cdot p(E|H_i)$; per il teorema delle probabilità totali il denominatore è $p(E) = \sum p(H_j) \cdot p(E|H_j)$. Sostituendo si ottiene la tesi. c.v.d.

La dimostrazione è breve, perché la formula non contiene idee nuove: mette insieme la definizione di probabilità condizionata e i due teoremi delle sezioni precedenti. La sua forza sta nell'interpretazione. Ogni termine ha un nome:

- $p(H_i)$ è la **probabilità a priori** dell'ipotesi H_i, quella che le assegnavamo *prima* di osservare E;
- $p(H_i|E)$ è la **probabilità a posteriori**, quella che le assegniamo *dopo* aver osservato E;
- $p(E|H_i)$ è la **verosimiglianza** dell'ipotesi H_i: misura quanto l'ipotesi rende probabile ciò che abbiamo osservato.

In breve: *probabilità a posteriori = probabilità a priori × verosimiglianza, diviso un numero che è lo stesso per tutte le ipotesi*. Il denominatore serve solo a fare in modo che le probabilità a posteriori sommino a 1. Un'ipotesi diventa più credibile se rende l'osservazione più probabile di quanto la rendano, in media, le altre; e il punto di partenza, la probabilità a priori, conta quanto la verosimiglianza.

15.4.2 Un esempio industriale

ESEMPIO SVOLTO

Quale macchina ha prodotto il pezzo difettoso? Un'azienda produce bulloni con tre macchine (dati inventati): la macchina M_1 produce il 50% dei bulloni, M_2 il 30%, M_3 il 20%. I bulloni difettosi sono il 2% di quelli prodotti da M_1, il 3% di quelli di M_2 e il 5% di quelli di M_3. Un bullone scelto a caso risulta difettoso. Qual è la probabilità che sia stato prodotto da ciascuna macchina?

Ipotesi e dati. Le tre macchine formano una partizione: $p(M_1) = 0{,}5$, $p(M_2) = 0{,}3$, $p(M_3) = 0{,}2$. Detto D l'evento «il bullone è difettoso», le verosimiglianze sono $p(D|M_1) = 0{,}02$, $p(D|M_2) = 0{,}03$, $p(D|M_3) = 0{,}05$.

Probabilità totale. Sommiamo le probabilità dei percorsi che terminano in D (Figura 15.4):

$$p(D) = 0{,}5 \cdot 0{,}02 + 0{,}3 \cdot 0{,}03 + 0{,}2 \cdot 0{,}05 = 0{,}010 + 0{,}009 + 0{,}010 = 0{,}029.$$

Formula di Bayes.

$$p(M_1|D) = \frac{0{,}010}{0{,}029} = \frac{10}{29} \approx 0{,}345; \quad p(M_2|D) = \frac{0{,}009}{0{,}029} = \frac{9}{29} \approx 0{,}310; \quad p(M_3|D) = \frac{0{,}010}{0{,}029} = \frac{10}{29} \approx 0{,}345.$$

Controllo e commento. La somma è $\frac{10+9+10}{29} = 1$, come deve essere. La macchina M_3 produce solo un bullone su cinque, ma è responsabile di più di un terzo dei difetti: la sua probabilità è passata da 0,2 (a priori) a circa 0,345 (a posteriori), perché la sua verosimiglianza è la più alta. Al contrario M_1 scende da 0,5 a circa 0,345. Se il responsabile della qualità deve scegliere quale macchina controllare per prima, questi numeri sono più utili delle sole percentuali di difettosi.

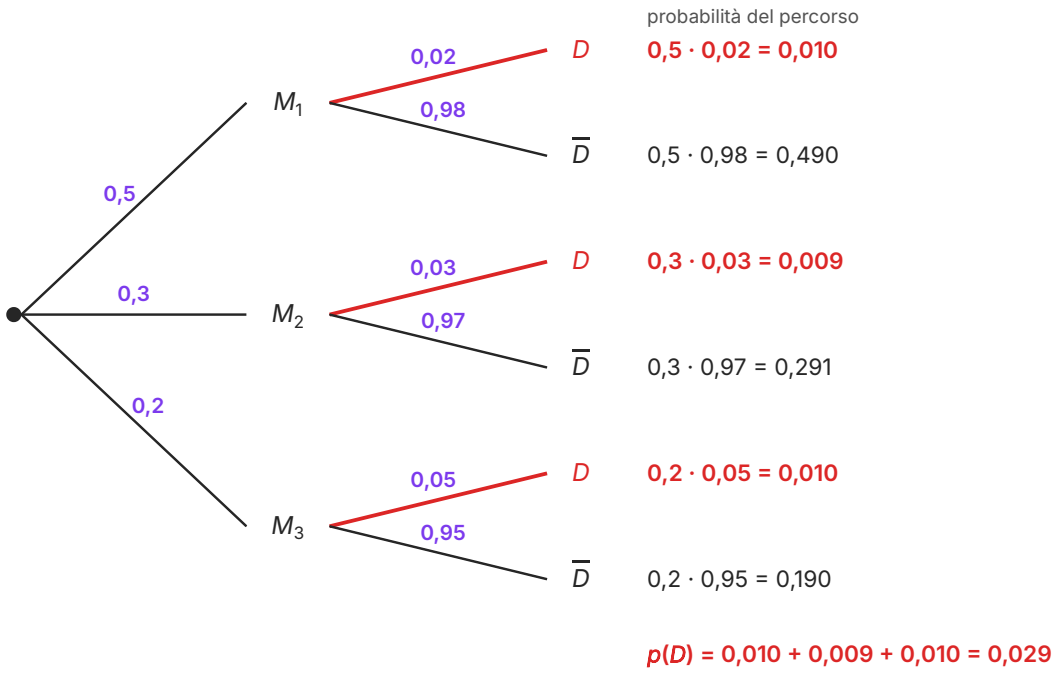


Figura 15.4 Il diagramma ad albero dell'esempio delle tre macchine. Sui rami del primo livello ci sono le probabilità a priori $p(M_i)$, su quelli del secondo livello le probabilità condizionate $p(D|M_i)$ e $p(\overline{D}|M_i)$; a destra la probabilità di ciascun percorso. I percorsi in rosso terminano in D: la loro somma è $p(D) = 0{,}029$, e ciascuno di essi, diviso per 0,029, dà la probabilità a posteriori della macchina corrispondente.

15.4.3 Aggiornare una probabilità

La formula di Bayes si può applicare più volte di seguito. Dopo aver osservato E_1, le probabilità a posteriori $p(H_i|E_1)$ descrivono ciò che sappiamo; se arriva una seconda osservazione E_2, le usiamo come *nuove probabilità a priori* e applichiamo di nuovo la formula. Questo procedimento si chiama **aggiornamento bayesiano**: ogni dato sposta le probabilità delle ipotesi, e la probabilità a posteriori di oggi è la probabilità a priori di domani.

Torniamo alle urne. Dopo la prima pallina bianca, $p(U_1) = \dfrac{3}{4}$. Rimettiamo la pallina nella stessa urna ed estraiamo di nuovo dalla *stessa* urna: esce un'altra pallina bianca. Con le nuove probabilità a priori $\dfrac{3}{4}$ e $\dfrac{1}{4}$:

$$p(U_1|\text{seconda bianca}) = \frac{\frac{3}{4}\cdot\frac{3}{5}}{\frac{3}{4}\cdot\frac{3}{5}+\frac{1}{4}\cdot\frac{1}{5}} = \frac{9/20}{9/20+1/20} = \frac{9}{10}.$$

Due palline bianche portano la prima urna al 90%. Avremmo ottenuto lo stesso numero con un solo passo, considerando l'evento «due bianche» con probabilità a priori $\frac{1}{2}$ e verosimiglianze $(\frac{3}{5})^2$ e $(\frac{1}{5})^2$: $\frac{(1/2)(9/25)}{(1/2)(9/25)+(1/2)(1/25)} = \frac{9}{10}$. Aggiornare un passo alla volta o tutto insieme dà lo stesso risultato, a patto che le due estrazioni, *per ciascuna urna*, siano indipendenti: qui lo sono, perché la pallina viene rimessa nell'urna.

APPROFONDIMENTO · THOMAS BAYES E PIERRE-SIMON LAPLACE

Thomas Bayes era un pastore presbiteriano inglese con interessi matematici. Non pubblicò mai il lavoro per cui è ricordato: dopo la sua morte, avvenuta nel 1761, l'amico Richard Price ne trovò il manoscritto, lo curò e lo presentò alla Royal Society di Londra, che lo pubblicò nel 1763 con il titolo *An Essay towards solving a Problem in the Doctrine of Chances*. Il problema era proprio quello di risalire dai dati osservati alla probabilità di una causa sconosciuta.

Indipendentemente da Bayes, il matematico francese Pierre-Simon Laplace formulò la stessa regola in una memoria del 1774 sulla «probabilità delle cause a partire dagli eventi», e la sviluppò nella *Théorie analytique des probabilités* del 1812, applicandola a problemi di astronomia e di demografia. La forma generale della formula, con molte ipotesi e il denominatore dato dalle probabilità totali, è sostanzialmente quella di Laplace; il nome di Bayes le è rimasto per la priorità dell'idea.

15.5 Le applicazioni della formula di Bayes

15.5.1 I test diagnostici

Un test diagnostico serve a stabilire se una persona ha una certa malattia. Nessun test è perfetto: può dare un risultato positivo a una persona sana (un **falso positivo**) o negativo a una persona malata (un **falso negativo**). Indichiamo con M l'evento «la persona è malata», con $\overline{M}$ il suo contrario e con + e − i due risultati del test. La qualità di un test si descrive con due probabilità condizionate.

- La **sensibilità** è $p(+|M)$: la probabilità che il test riconosca un malato.
- La **specificità** è $p(-|\overline{M})$: la probabilità che il test riconosca un sano. Le probabilità di errore sono dunque $p(-|M) = 1 - \text{sensibilità}$ e $p(+|\overline{M}) = 1 - \text{specificità}$.

A queste si aggiunge una caratteristica della popolazione in cui il test si usa, la **prevalenza** della malattia, cioè la frazione di persone malate: $p(M)$ per una persona scelta a caso. Chi riceve un risultato, però, vuole conoscere altre due probabilità: il **valore predittivo positivo** $p(M|+)$, la probabilità di essere malato quando il test è positivo, e il **valore predittivo negativo** $p(\overline{M}|-)$, la probabilità di essere sano quando il test è negativo. Il passaggio dalle prime alle seconde è esattamente il lavoro della formula di Bayes.

ESEMPIO SVOLTO

Un test positivo. Una malattia ha prevalenza 1% in una certa popolazione; un test ha sensibilità 99% e specificità 95% (dati inventati). Una persona scelta a caso fa il test e risulta positiva. Qual è la probabilità che sia malata? E se il test fosse risultato negativo, quale sarebbe la probabilità che sia sana?

Dati. $p(M) = 0{,}01$, $p(\overline{M}) = 0{,}99$, $p(+|M) = 0{,}99$, $p(+|\overline{M}) = 1 - 0{,}95 = 0{,}05$.

Probabilità di un positivo. Dall'albero della Figura 15.5:

$$p(+) = 0{,}01 \cdot 0{,}99 + 0{,}99 \cdot 0{,}05 = 0{,}0099 + 0{,}0495 = 0{,}0594.$$

Valore predittivo positivo.

$$p(M|+) = \frac{0{,}0099}{0{,}0594} = \frac{1}{6} \approx 16{,}7\%.$$

Valore predittivo negativo. $p(-) = 1 - 0{,}0594 = 0{,}9406$ e $p(\overline{M} \cap -) = 0{,}99 \cdot 0{,}95 = 0{,}9405$, quindi

$$p(\overline{M}|-) = \frac{0{,}9405}{0{,}9406} \approx 0{,}99989.$$

Commento. Un test negativo è quasi una garanzia di salute; un test positivo, invece, lascia la persona molto più probabilmente sana che malata: 5 positivi su 6 sono falsi positivi. Il motivo è che i sani sono 99 volte più numerosi dei malati, e il 5% di un gruppo molto grande supera il 99% di un gruppo molto piccolo. Il test è comunque utile: ha fatto salire la probabilità di malattia dall'1% a quasi il 17%, cioè di circa 17 volte, e giustifica un secondo esame più accurato.

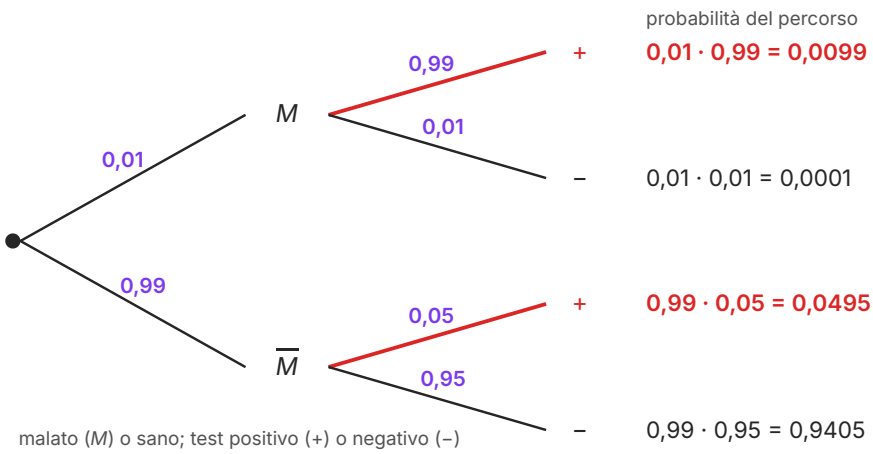


Figura 15.5 Il diagramma ad albero del test diagnostico dell'esempio svolto: prevalenza 1%, sensibilità 99%, specificità 95%. I due percorsi in rosso terminano con un test positivo; la loro somma è $p(+) = 0{,}0594$, e il percorso dei malati ne costituisce la sesta parte.

Lo stesso calcolo diventa quasi evidente se, invece delle probabilità, si ragiona con le **frequenze naturali**: si immagina un gruppo di persone, per esempio 10 000, e si contano quante cadono in ciascun caso (Figura 15.6). Dei 10 000, l'1% è malato: 100 persone, delle quali il test riconosce il 99%, cioè 99. I sani sono 9 900, e il 5% di loro, cioè 495, risulta positivo. In tutto i positivi sono $99 + 495 = 594$, e solo 99 di essi sono malati: $\frac{99}{594} = \frac{1}{6}$. Gli studi di psicologia del ragionamento hanno mostrato che con le frequenze naturali anche chi non conosce la formula di Bayes sbaglia molto meno: è il metodo da preferire quando si deve spiegare un risultato a chi non è un matematico.

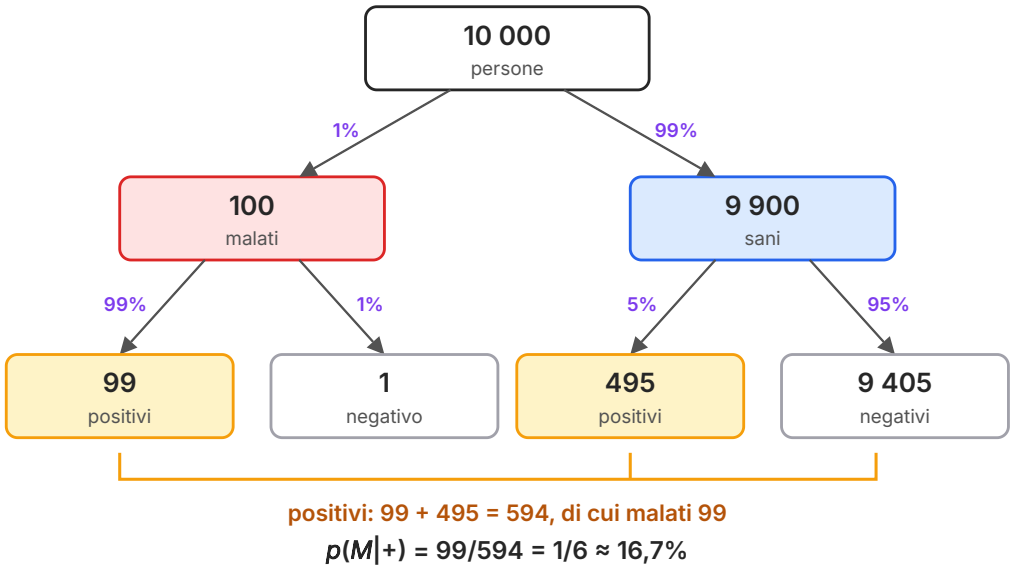


Figura 15.6 Le frequenze naturali per lo stesso test, su un gruppo immaginario di 10 000 persone. I 594 positivi sono formati da 99 malati e 495 sani: la frazione di malati fra i positivi è 99/594 = 1/6.

Che cosa succede se una persona positiva ripete il test, e anche il secondo risultato è positivo? Supponiamo che i due risultati, per una persona malata e per una persona sana, siano indipendenti (è un'ipotesi da discutere, come vedremo). La probabilità a priori ora è $\frac{1}{6}$, e la formula di Bayes dà

$$p(M|\text{due positivi}) = \frac{\frac{1}{6}\cdot 0{,}99}{\frac{1}{6}\cdot 0{,}99 + \frac{5}{6}\cdot 0{,}05} = \frac{0{,}99}{0{,}99 + 0{,}25} = \frac{99}{124} \approx 79{,}8\%.$$

Il secondo positivo ha portato la probabilità dal 16,7% a quasi l'80%; un terzo positivo la porterebbe oltre il 98%. È per questo che, nella pratica medica, un primo test di *screening* positivo viene di norma seguito da esami di conferma.

ERRORE FREQUENTE

Scambiare $p(A|B)$ con $p(B|A)$. «Il test è affidabile al 99%, quindi se sono positivo ho il 99% di probabilità di essere malato» confonde la sensibilità $p(+|M)$ con il valore predittivo $p(M|+)$. La prima descrive il test; la seconda dipende anche dalla popolazione.

Ignorare la frequenza di base. Chi stima $p(M|+)$ guardando solo sensibilità e specificità dimentica la prevalenza $p(M)$, cioè la probabilità a priori. Questo errore, molto studiato dagli psicologi, si chiama **fallacia della frequenza di base**. Una probabilità a posteriori non si può calcolare senza una probabilità a priori.

15.5.2 I filtri antispam

Molti programmi di posta elettronica separano i messaggi indesiderati (lo *spam*) da quelli normali con un metodo che si basa sulla formula di Bayes, e che per questo si chiama **filtro bayesiano**. Il programma analizza un grande archivio di messaggi già classificati e stima, per ogni parola, la frequenza con cui compare nello spam e nei messaggi normali. Quando arriva un messaggio nuovo, le parole che contiene sono le osservazioni, e le due ipotesi sono «spam» (S) e «non spam» ($\overline{S}$).

Un esempio con un'unica parola (dati inventati). In un archivio, il 40% dei messaggi è spam; la parola «gratis» compare nel 25% dei messaggi di spam e nel 2% degli altri. Detto G l'evento «il messaggio contiene “gratis”»,

$$p(S|G) = \frac{0{,}4 \cdot 0{,}25}{0{,}4 \cdot 0{,}25 + 0{,}6 \cdot 0{,}02} = \frac{0{,}1}{0{,}112} = \frac{25}{28} \approx 89{,}3\%.$$

Un filtro reale combina le verosimiglianze di molte parole, di solito supponendo, per semplicità, che le parole compaiano in modo indipendente le une dalle altre (un'ipotesi certamente falsa, ma che funziona sorprendentemente bene), e sposta il messaggio nella cartella dello spam se la probabilità a posteriori supera una soglia fissata. Il filtro, inoltre, impara: ogni volta che l'utente corregge una classificazione, le frequenze vengono aggiornate.

15.5.3 La fallacia del procuratore

In un processo (immaginario) l'accusa presenta questa prova: sulla scena del reato è stata trovata una traccia biologica con caratteristiche che si riscontrano in una persona su 10 000, e l'imputato le possiede. «La probabilità che un innocente abbia queste caratteristiche è 1 su 10 000», dice l'accusa, «quindi la probabilità che l'imputato sia innocente è 1 su 10 000». Il ragionamento scambia $p(T|I)$, la probabilità della traccia compatibile (T) dato che la persona è innocente (I), con $p(I|T)$, la probabilità di innocenza data la traccia compatibile. È la **fallacia del procuratore**.

Per vedere quanto possano essere diverse, supponiamo che l'autore del reato possa essere, con uguale probabilità, uno qualsiasi di 200 001 abitanti di una zona, e che non ci siano altre

prove. Fra i 200 000 innocenti, in media 20 hanno le caratteristiche della traccia; con il colpevole, le persone compatibili sono circa 21, e la formula di Bayes conferma il conto:

$$p(\text{colpevole}|T) = \frac{1/200\,001 \cdot 1}{1/200\,001 \cdot 1 + 200\,000/200\,001 \cdot 1/10\,000} = \frac{1}{1+20} = \frac{1}{21} \approx 4{,}8\%.$$

Senza altre prove, la probabilità di colpevolezza è meno del 5%, non il 99,99%. Naturalmente la prova non è inutile: ha fatto salire la probabilità di colpevolezza da 1 su 200 001 a 1 su 21. Ma un tribunale deve valutarla insieme alle altre prove, che modificano la probabilità a priori. L'errore opposto, la *fallacia della difesa*, consiste nel dire «la prova vale poco, perché le persone compatibili sono 21», dimenticando che in un processo reale l'imputato non è stato scelto a caso fra gli abitanti.

15.5.4 Il problema di Monty Hall

In *Matematica 2* hai incontrato il problema di Monty Hall e l'hai risolto con un albero. Ricordiamolo: un concorrente sceglie una di tre porte chiuse, dietro una sola delle quali c'è un'automobile. Il conduttore, che sa dov'è l'auto, apre sempre un'altra porta dietro cui c'è una capra (scegliendo a caso se può scegliere) e offre al concorrente di cambiare porta. Conviene cambiare? Ora possiamo rispondere con la formula di Bayes.

Il concorrente sceglie la porta 1. Le ipotesi sono H_1, H_2, H_3 = «l'auto è dietro la porta 1, 2, 3», ciascuna con probabilità a priori $\frac{1}{3}$. Supponiamo che il conduttore apra la porta 3 (evento E). Le verosimiglianze sono:

- $p(E|H_1) = \frac{1}{2}$: se l'auto è dietro la porta 1, il conduttore può aprire la 2 o la 3 e sceglie a caso;
- $p(E|H_2) = 1$: se l'auto è dietro la porta 2, il conduttore è costretto ad aprire la 3;
- $p(E|H_3) = 0$: il conduttore non apre mai la porta dell'auto.

$$p(E) = \frac{1}{3}\cdot\frac{1}{2} + \frac{1}{3}\cdot 1 + \frac{1}{3}\cdot 0 = \frac{1}{2}, \quad p(H_1|E) = \frac{1/6}{1/2} = \frac{1}{3}, \quad p(H_2|E) = \frac{1/3}{1/2} = \frac{2}{3}.$$

Cambiare porta raddoppia la probabilità di vincere: $\frac{2}{3}$ contro $\frac{1}{3}$. La formula mostra anche *perché* l'intuizione «ormai è 50 e 50» sbaglia: l'apertura della porta 3 è un'osservazione che l'ipotesi H_2 rende certa e l'ipotesi H_1 rende probabile solo al 50%. La verosimiglianza di H_2 è doppia, e la sua probabilità a posteriori è doppia.

COLLEGAMENTI

Medicina. I programmi di *screening*, che sottopongono a un test grandi gruppi di persone senza sintomi, si progettano con i ragionamenti di questa sezione: in una popolazione in cui la malattia è rara anche un buon test produce molti falsi positivi, e il percorso deve prevedere esami di conferma.

Diritto. Nei processi le prove scientifiche, come le analisi del DNA, si esprimono sempre più spesso come rapporti di verosimiglianza (la probabilità della prova se l'imputato è colpevole divisa per la probabilità della stessa prova se è innocente), proprio per evitare la fallacia del procuratore.

Informatica. Molti sistemi di apprendimento automatico (i *classificatori*) assegnano un oggetto alla classe che ha la probabilità a posteriori più alta, stimando probabilità a priori e verosimiglianze dai dati di addestramento.

15.6 Probabilità e modello matematico

15.6.1 Il peso della probabilità a priori

Il valore predittivo di un test non è una proprietà del test: dipende dalla popolazione a cui si applica. La Tabella 15.2 riporta, per lo stesso test di sensibilità 99% e specificità 95%, il valore predittivo positivo per diverse prevalenze.

Tabella 15.2 Valore predittivo positivo $p(M|+)$ di un test con sensibilità 99% e specificità 95%, al variare della prevalenza $p(M)$ (valori arrotondati).

Prevalenza $p(M)$	$p(M \cap +)$	$p(\overline{M} \cap +)$	$p(M\|+)$
0,1%	0,00099	0,04995	1,9%
1%	0,0099	0,0495	16,7%
10%	0,099	0,045	68,75%
50%	0,495	0,025	95,2%

Il valore predittivo, come funzione della prevalenza x, è

$$f(x) = \frac{0{,}99x}{0{,}99x + 0{,}05(1-x)} = \frac{0{,}99x}{0{,}94x + 0{,}05},$$

una funzione omografica, il cui grafico è un ramo di iperbole. La Figura 15.7 lo confronta con quello di un test con la stessa sensibilità ma specificità 99%. Due fatti saltano all'occhio. Quando la malattia è rara il valore predittivo crolla, qualunque sia la qualità del test; e per le malattie

rare conta soprattutto la specificità: portandola dal 95% al 99%, con prevalenza 1% il valore predittivo passa dal 16,7% al 50%. Per questo un test di screening si usa di solito nei gruppi a rischio, dove la prevalenza è più alta.

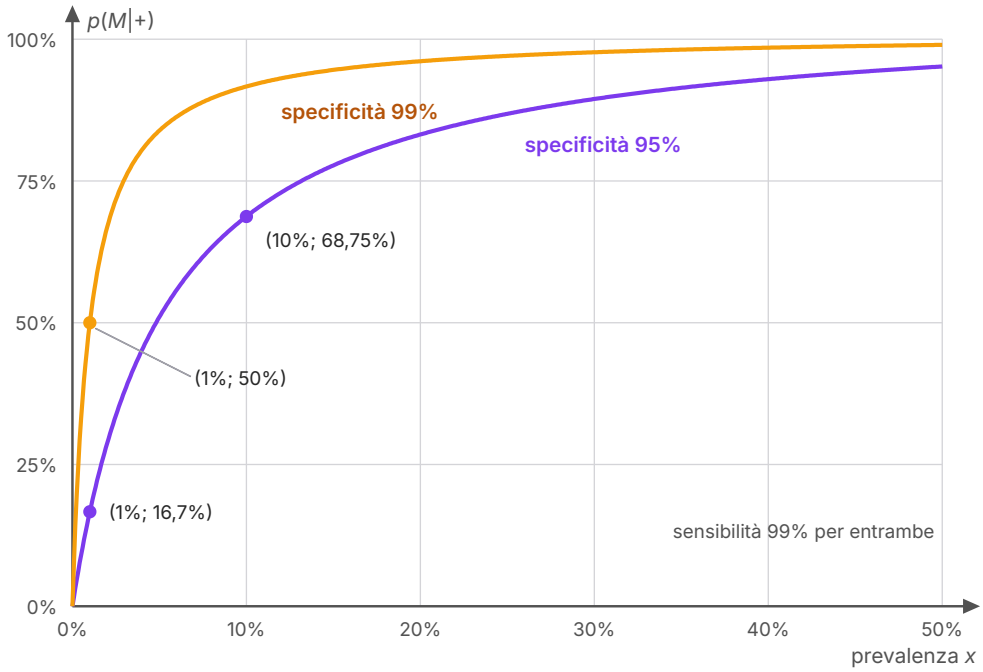


Figura 15.7 Il valore predittivo positivo $p(M|+)$ in funzione della prevalenza, per due test con sensibilità 99% e specificità 95% (viola) o 99% (arancione). Le curve sono rami delle iperboli $y = 0{,}99x / (0{,}94x + 0{,}05)$ e $y = 0{,}99x / (0{,}98x + 0{,}01)$; i punti segnati corrispondono ai valori calcolati nel testo.

15.6.2 Le ipotesi del modello

Ogni calcolo di questo capitolo è la conclusione di un **modello probabilistico**: uno spazio campionario, una partizione in ipotesi, delle probabilità a priori e delle verosimiglianze, e alcune ipotesi di indipendenza. I teoremi garantiscono che, *se* le ipotesi sono vere, le conclusioni sono corrette; non dicono nulla sulla verità delle ipotesi. Vediamo dove si nascondono le scelte.

- **La probabilità a priori.** Nel test diagnostico abbiamo usato la prevalenza nella popolazione. Ma se la persona fa il test perché ha dei sintomi, non è «scelta a caso»: la sua probabilità a priori è più alta, e la probabilità a posteriori di un positivo cambia di conseguenza.
- **L'indipendenza.** Per due test positivi abbiamo moltiplicato le verosimiglianze. Se però un sano risulta positivo per una caratteristica stabile del suo organismo, il secondo test sbaglierà probabilmente come il primo: i due errori non sono indipendenti, e il 79,8% è una sovrastima.

- **L'equiprobabilità.** Nel problema di Monty Hall abbiamo supposto che il conduttore, quando può, scelga a caso. Se preferisse sempre la porta con il numero più alto, la verosimiglianza $p(E|H_1)$ sarebbe 1 e il risultato diventerebbe $\frac{1}{2}$ e $\frac{1}{2}$.

Cambiare il modello cambia la conclusione, ed è giusto così: è la differenza fra la matematica, che deduce, e la modellistica, che sceglie. Lo stesso fenomeno si può rappresentare con modelli diversi, e un buon modello è quello le cui ipotesi sono dichiarate e controllabili. La Tabella 15.2 serve proprio a questo: mostra quanto la conclusione sia sensibile alle ipotesi, e quindi quanto valga la pena di conoscerle bene.

C'è infine il caso in cui la probabilità a priori non viene da una frequenza ma da un giudizio. In *Matematica 2* hai incontrato la definizione soggettiva di Bruno de Finetti: la probabilità è il grado di fiducia di una persona in un evento, purché le sue valutazioni siano coerenti. La formula di Bayes completa quella definizione: dice come una persona coerente deve modificare le proprie probabilità quando osserva un dato. Due persone che partono da probabilità a priori diverse arrivano a probabilità a posteriori diverse; ma se osservano entrambe molti dati, e nessuna delle due esclude in partenza un'ipotesi (assegnandole probabilità zero), le loro valutazioni tendono ad avvicinarsi. È in questo senso che l'esperienza «corregge» le opinioni.

SINTESI

- Se $p(B) > 0$, la probabilità condizionata è $p(A|B) = \dfrac{p(A \cap B)}{p(B)}$: è una nuova probabilità, definita sullo spazio ridotto B, e soddisfa gli assiomi.
- In generale $p(A|B) \neq p(B|A)$: le due condizionate sono legate da $p(A|B) = p(B|A) \cdot p(A) \,/\, p(B)$. Una probabilità condizionata misura un'associazione, non una relazione di causa.
- Teorema della probabilità composta: $p(A \cap B) = p(B) \cdot p(A|B) = p(A) \cdot p(B|A)$; per più eventi si moltiplicano le probabilità condizionate a tutti gli eventi precedenti (regola dei percorsi di un albero).
- A e B sono indipendenti se $p(A \cap B) = p(A) \cdot p(B)$, cioè (se $p(B) > 0$) se $p(A|B) = p(A)$. Due eventi incompatibili di probabilità positiva sono sempre dipendenti.
- Una partizione di Ω è formata da eventi incompatibili a due a due, di probabilità positiva, la cui unione è Ω.
- Teorema delle probabilità totali: $p(E) = \sum p(H_i) \cdot p(E|H_i)$, una media delle condizionate pesata con le probabilità delle ipotesi.
- Formula di Bayes: $p(H_i|E) = \dfrac{p(H_i) \cdot p(E|H_i)}{\sum p(H_j) \cdot p(E|H_j)}$. La probabilità a posteriori è proporzionale al prodotto della probabilità a priori per la verosimiglianza.
- L'aggiornamento bayesiano usa le probabilità a posteriori come nuove probabilità a priori; con osservazioni indipendenti per ciascuna ipotesi, aggiornare un passo alla volta o tutto insieme dà lo stesso risultato.
- In un test diagnostico sensibilità $p(+|M)$ e specificità $p(-|\overline{M})$ descrivono il test; il valore predittivo $p(M|+)$ dipende anche dalla prevalenza. Le frequenze naturali rendono il calcolo trasparente.
- Scambiare $p(A|B)$ con $p(B|A)$ (fallacia del procuratore) e ignorare la probabilità a priori (fallacia della frequenza di base) sono gli errori più comuni.
- Ogni risultato dipende dalle ipotesi del modello: probabilità a priori, indipendenza, equiprobabilità. Cambiando le ipotesi cambiano le conclusioni.

ESERCIZI

Comprendere

1. ● Se $p(B) > 0$, la probabilità condizionata $p(A|B)$ è uguale a:
   - **A.** $\dfrac{p(A \cap B)}{p(A)}$
   - **B.** $\dfrac{p(A \cap B)}{p(B)}$
   - **C.** $p(A) \cdot p(B)$
   - **D.** $p(B|A)$

2. ● Due eventi indipendenti A e B hanno $p(A) = 0{,}4$ e $p(B) = 0{,}5$. Allora $p(A \cap B)$ vale:
   - **A.** 0,9
   - **B.** 0,45
   - **C.** 0,1
   - **D.** 0,2

3. ●● Due eventi A e B sono incompatibili e hanno entrambi probabilità positiva. Allora:
   - **A.** sono certamente dipendenti
   - **B.** sono certamente indipendenti
   - **C.** possono essere dipendenti o indipendenti
   - **D.** si ha $p(A|B) = p(A)$

4. ●● Nella formula di Bayes $p(H_i|E) = p(H_i) \cdot p(E|H_i) \,/\, p(E)$, il denominatore $p(E)$ si calcola:
   - **A.** come probabilità a priori dell'ipotesi H_i
   - **B.** come verosimiglianza dell'ipotesi H_i
   - **C.** con il teorema delle probabilità totali
   - **D.** come somma delle probabilità a posteriori

5. ●● «Il test ha sensibilità 90%» significa che, detto M l'evento «la persona è malata»:
   - **A.** $p(M|+) = 0{,}9$
   - **B.** $p(+|M) = 0{,}9$
   - **C.** $p(-|\overline{M}) = 0{,}9$
   - **D.** il 90% dei positivi è malato

6. ●● Si lancia due volte una moneta equilibrata. Gli eventi A = «testa al primo lancio» e B = «esattamente una testa nei due lanci» sono:
   - **A.** incompatibili
   - **B.** dipendenti, con $p(A|B) > p(A)$
   - **C.** dipendenti, con $p(A|B) < p(A)$
   - **D.** indipendenti

7. ● Vero o falso? Correggi le affermazioni false. a) Per ogni coppia di eventi con probabilità positiva si ha $p(A|B) = p(B|A)$. b) Se $p(B) > 0$, allora $p(\overline{A}|B) = 1 - p(A|B)$. c) Se $A \subset B$ e $p(B) > 0$, allora $p(A|B) = p(A) / p(B)$. d) Per ogni coppia di eventi, $p(A|B) + p(A|\overline{B}) = 1$.

8. ● Completa. Gli eventi H_1, ..., H_n formano una partizione di Ω se sono _____ a due a due, se la loro unione è _____ e se ciascuno di essi ha probabilità _____ .

9. ●● Completa. Nella formula di Bayes la probabilità $p(H)$ si chiama probabilità _____ , la probabilità $p(H|E)$ si chiama probabilità _____ e la probabilità $p(E|H)$ si chiama _____ dell'ipotesi H.

Calcolare

10. ● Sono dati $p(A) = 0{,}6$, $p(B) = 0{,}5$ e $p(A \cap B) = 0{,}2$. Calcola $p(A|B)$, $p(B|A)$, $p(A \cup B)$ e $p(\overline{A}|B)$. Gli eventi A e B sono indipendenti?

11. ● Si lancia un dado equilibrato. a) Sapendo che è uscito un numero pari, qual è la probabilità che sia maggiore di 3? b) Sapendo che è uscito un numero primo, qual è la probabilità che sia pari?

12. ●● Si lanciano due dadi equilibrati. a) Sapendo che almeno uno dei due mostra 2, qual è la probabilità che la somma sia 7? b) Sapendo che la somma è 7, qual è la probabilità che almeno uno dei due mostri 2?

13. ●● In un gruppo di 300 persone (dati inventati), 180 si sono vaccinate contro l'influenza e 120 no. In inverno si sono ammalate 18 persone vaccinate e 30 non vaccinate. Si sceglie una persona a caso; siano V = «è vaccinata» e I = «si è ammalata». a) Costruisci la tabella a doppia entrata. b) Calcola $p(I)$, $p(I|V)$, $p(I|\overline{V})$ e $p(V|I)$. c) Gli eventi I e V sono indipendenti?

14. ●● Da un mazzo di 40 carte si estraggono due carte, una dopo l'altra, senza reimmissione. Calcola la probabilità che: a) siano entrambe assi; b) la seconda carta sia un asso.

15. ●● L'urna A contiene 4 palline rosse e 6 blu, l'urna B 7 rosse e 3 blu. Si lancia un dado: se esce 1 o 2 si estrae una pallina da A, altrimenti da B. a) Calcola la probabilità che la pallina sia rossa. b) La pallina è rossa: qual è la probabilità che venga da A?

16. ●● Gli eventi A e B hanno $p(A) = 0{,}3$, $p(B) = 0{,}5$ e $p(A \cup B) = 0{,}65$. Stabilisci se sono indipendenti e calcola $p(A|\overline{B})$.

17. ●● Un'urna contiene 5 palline bianche e 3 nere. Se ne estraggono tre, una dopo l'altra, senza reimmissione. Calcola la probabilità di estrarre, nell'ordine, bianca, bianca, nera, e la probabilità di estrarre esattamente una pallina nera.

18. ●● Una ditta acquista componenti da due fornitori (dati inventati): il 60% dal fornitore A, con il 3% di pezzi difettosi, e il 40% dal fornitore B, con il 5% di pezzi difettosi. a) Qual è la probabilità che un componente scelto a caso sia difettoso? b) Un componente è difettoso: qual è la probabilità che venga da B?

19. ●● Una malattia ha prevalenza 2%; un test ha sensibilità 95% e specificità 90% (dati inventati). a) Costruisci la tabella delle frequenze naturali per 10 000 persone. b) Calcola il valore predittivo positivo $p(M|+)$ e il valore predittivo negativo $p(\overline{M}|-)$.

20. ●●● Con i dati dell'esercizio precedente, una persona risulta positiva a due test successivi, i cui risultati si suppongono indipendenti sia per i malati sia per i sani. Calcola la probabilità che sia malata, prima aggiornando un test alla volta, poi con un solo passo.

Dimostrare e argomentare

21. ●● Dimostra che, se A e B sono indipendenti, anche A e $\overline{B}$ sono indipendenti. (Suggerimento: $A = (A \cap B) \cup (A \cap \overline{B})$.) Deduci che sono indipendenti anche $\overline{A}$ e $\overline{B}$.

22. ●● Dimostra, partendo dalla definizione, che se $p(B) > 0$ allora $p(\overline{A}|B) = 1 - p(A|B)$.

23. ●● Siano A e B eventi con $p(A \cap B) > 0$. Dimostra che $p(A|B) = p(B|A)$ se e solo se $p(A) = p(B)$.

24. ●● Trova l'errore. Un articolo afferma: «Il 90% dei ciclisti feriti in incidenti stradali indossava il casco: dunque il casco non protegge». Spiega quale probabilità condizionata riporta l'articolo, quale servirebbe per trarre una conclusione, e quale informazione manca.

25. ●●● Si lanciano due monete equilibrate. Siano A = «testa con la prima moneta», B = «testa con la seconda moneta», C = «le due monete mostrano la stessa faccia». Dimostra che i tre eventi sono indipendenti a due a due, ma che $p(A \cap B \cap C) \neq p(A) \cdot p(B) \cdot p(C)$.

Problemi

26. ●● In un archivio di posta (dati inventati) il 30% dei messaggi è spam. La parola «offerta» compare nel 40% dei messaggi di spam e nel 5% degli altri. a) Un messaggio contiene «offerta»: qual è la probabilità che sia spam? b) E se non la contiene?

27. ●● In una località immaginaria piove nel 25% dei giorni. Nei giorni di pioggia il servizio meteorologico aveva previsto pioggia l'80% delle volte; nei giorni senza pioggia l'aveva prevista il 10% delle volte (dati inventati). a) Domani è prevista pioggia: qual è la probabilità che piova? b) E se la pioggia non è prevista?

28. ●● In una specie animale il colore del pelo dipende da un solo gene con due alleli: N (pelo nero) è dominante, n (pelo bianco) è recessivo. Due genitori Nn hanno un figlio con il pelo nero. a) Qual è la probabilità che il figlio sia Nn? b) Il figlio viene incrociato con un individuo nn e nascono 3 piccoli, tutti neri. Aggiorna la probabilità che il figlio sia Nn.

29. ●● Sulla scena di un reato (caso immaginario) si trova una traccia di un gruppo sanguigno presente nel 2% della popolazione. L'autore può essere, con uguale probabilità, una qualsiasi di 5001 persone, e l'imputato ha quel gruppo sanguigno. a) Senza altre prove, qual è la probabilità che sia colpevole? b) Altre prove riducono i possibili autori a 51 persone, di cui l'imputato è una, ugualmente probabili: rifai il calcolo. c) Spiega l'errore di chi afferma: «la probabilità che l'imputato sia innocente è il 2%».

30. ●●● In una linea di produzione il 4% dei pezzi è difettoso (dati inventati). Un sistema di controllo automatico segnala il 95% dei pezzi difettosi, ma segnala per errore anche il 2% dei pezzi buoni. a) Un pezzo viene segnalato: qual è la probabilità che sia davvero difettoso? b) I pezzi segnalati passano per un secondo controllo, identico e indipendente dal primo: qual è la probabilità che un pezzo segnalato due volte sia difettoso? c) Commenta l'utilità del secondo controllo.

Con il computer

31. ●● In un foglio di calcolo scrivi in tre celle la sensibilità, la specificità e la prevalenza di un test, e in una quarta cella la formula del valore predittivo positivo. Poi costruisci una colonna di prevalenze da 0,001 a 0,5 e la colonna dei valori predittivi corrispondenti, e rappresentale con un grafico a dispersione. a) Con sensibilità 99% e specificità 95%, trova per tentativi la prevalenza per cui $p(M|+) = 50\%$, e verifica il risultato risolvendo un'equazione. b) Ripeti con specificità 99%. c) Che cosa cambia se si migliora la sensibilità invece della specificità?

Verso la verifica

32. ●● Sono dati $p(A) = \frac{1}{3}$, $p(B|A) = \frac{1}{4}$ e $p(B|\overline{A}) = \frac{1}{2}$. Calcola $p(B)$, $p(A|B)$ e $p(A \cup B)$, e stabilisci se A e B sono indipendenti.

33. ●● In una classe il 60% degli studenti sono ragazze. Porta gli occhiali il 30% delle ragazze e il 20% dei ragazzi. Si sceglie a caso uno studente con gli occhiali: qual è la probabilità che sia una ragazza?

34. ●●● **Sfida.** Il gioco di Monty Hall si fa con quattro porte: dietro una c'è un'automobile, dietro le altre tre una capra. Il concorrente sceglie la porta 1; il conduttore, che sa dov'è l'auto, apre una delle altre porte dietro cui c'è una capra, scegliendo a caso fra quelle possibili, e offre di cambiare con una delle due porte rimaste chiuse. Supponi che apra la porta 4. Calcola con la formula di Bayes la probabilità che l'auto sia dietro ciascuna porta e stabilisci se conviene cambiare.

PER APPROFONDIRE

- [FEL] Feller, *An Introduction to Probability Theory and Its Applications*, vol. I: il capitolo sulla probabilità condizionata e l'indipendenza, con molti esempi classici di applicazione della formula di Bayes.
- [BOY] Boyer, *Storia della matematica*: la nascita del calcolo delle probabilità nel Settecento e l'opera di Laplace.

Glossario del capitolo

Aggiornamento bayesiano
L'uso ripetuto della formula di Bayes, in cui le probabilità a posteriori diventano le nuove probabilità a priori.

Eventi dipendenti
Eventi che non sono indipendenti.

Eventi indipendenti
Eventi per cui $p(A \cap B) = p(A) \cdot p(B)$; se $p(B) > 0$, equivale a $p(A|B) = p(A)$.

Fallacia del procuratore
L'errore di scambiare $p(\text{prova}|\text{innocente})$ con $p(\text{innocente}|\text{prova})$.

Falso positivo
Un risultato positivo del test per una persona sana (falso negativo: negativo per una persona malata).

Formula di Bayes
$p(H_i|E) = p(H_i) \cdot p(E|H_i) \,/\, \sum p(H_j) \cdot p(E|H_j)$: calcola la probabilità di un'ipotesi dopo aver osservato E.

Frequenze naturali
Conteggi su un gruppo immaginario di persone (per esempio 10 000) usati al posto delle probabilità per rendere evidente un calcolo.

Modello probabilistico
Uno spazio campionario con le probabilità assegnate e le ipotesi (indipendenza, equiprobabilità, probabilità a priori) da cui si deducono le conclusioni.

Partizione
Una famiglia di eventi di probabilità positiva, incompatibili a due a due, la cui unione è lo spazio campionario.

Prevalenza
La frazione di persone malate in una popolazione.

Probabilità a posteriori
La probabilità di un'ipotesi dopo un'osservazione, calcolata con la formula di Bayes.

Probabilità a priori
La probabilità di un'ipotesi prima di un'osservazione.

Probabilità condizionata
Se $p(B) > 0$, la probabilità di A sapendo che si è verificato B: $p(A|B) = p(A \cap B) \,/\, p(B)$.

Sensibilità
La probabilità $p(+|M)$ che un test sia positivo per una persona malata.

Spazio ridotto
L'evento B considerato come nuovo spazio campionario quando si sa che B si è verificato.

Specificità
La probabilità $p(-|\overline{M})$ che un test sia negativo per una persona sana.

Teorema della probabilità composta
$p(A \cap B) = p(B) \cdot p(A|B) = p(A) \cdot p(B|A)$.

Teorema delle probabilità totali
Se $H_1, \ldots, H_n$ è una partizione di Ω, $p(E) = \sum p(H_i) \cdot p(E|H_i)$.

Valore predittivo positivo
La probabilità $p(M|+)$ di essere malati quando il test è positivo.

Verosimiglianza
La probabilità $p(E|H)$ dell'osservazione E se l'ipotesi H è vera.

Soluzioni

1. B. È la definizione, valida per $p(B) > 0$.

2. D. Per eventi indipendenti $p(A \cap B) = p(A) \cdot p(B) = 0{,}4 \cdot 0{,}5 = 0{,}2$.

3. A. $p(A \cap B) = 0$ mentre $p(A) \cdot p(B) > 0$; infatti $p(A|B) = 0 \neq p(A)$.

4. C. $p(E) = \sum p(H_j) \cdot p(E|H_j)$.

5. B. La sensibilità è la probabilità che il test sia positivo per un malato.

6. D. $p(A) = \frac{1}{2}$, $p(B) = \frac{2}{4} = \frac{1}{2}$ (esiti TC e CT), $A \cap B$ contiene il solo esito TC: $p(A \cap B) = \frac{1}{4} = p(A) \cdot p(B)$.

7. a) Falso: $p(A|B) = p(B|A) \cdot p(A) / p(B)$, e le due condizionate coincidono solo se $p(A) = p(B)$ oppure se $p(A \cap B) = 0$ (e allora sono entrambe nulle). b) Vero: $p(\cdot|B)$ è una probabilità. c) Vero: $A \cap B = A$. d) Falso: è vero invece che $p(A|B) + p(\overline{A}|B) = 1$. Controesempio con un dado: A = «6», B = «pari»: $p(A|B) = \frac{1}{3}$, $p(A|\overline{B}) = 0$, somma $\frac{1}{3}$.

8. Incompatibili; Ω; positiva.

9. A priori; a posteriori; verosimiglianza.

10. $p(A|B) = \frac{0{,}2}{0{,}5} = 0{,}4$; $p(B|A) = \frac{0{,}2}{0{,}6} = \frac{1}{3}$; $p(A \cup B) = 0{,}6 + 0{,}5 - 0{,}2 = 0{,}9$; $p(\overline{A}|B) = 1 - 0{,}4 = 0{,}6$. Non sono indipendenti: $p(A) \cdot p(B) = 0{,}3 \neq 0{,}2$.

11. a) Spazio ridotto {2, 4, 6}, casi favorevoli 4 e 6: $\frac{2}{3}$. b) Spazio ridotto {2, 3, 5}, caso favorevole 2: $\frac{1}{3}$.

12. a) Le coppie con almeno un 2 sono 11; di queste hanno somma 7 (2; 5) e (5; 2): $\frac{2}{11}$. b) Le coppie con somma 7 sono 6, e 2 contengono un 2: $\frac{2}{6} = \frac{1}{3}$.

13. a) V: 18 malati, 162 sani (totale 180); $\overline{V}$: 30 malati, 90 sani (totale 120); totali 48 e 252, in tutto 300. b) $p(I) = \frac{48}{300} = 0{,}16$; $p(I|V) = \frac{18}{180} = 0{,}1$; $p(I|\overline{V}) = \frac{30}{120} = 0{,}25$; $p(V|I) = \frac{18}{48} = 0{,}375$. c) No: $p(I|V) = 0{,}1 \neq p(I) = 0{,}16$.

14. a) $\frac{4}{40} \cdot \frac{3}{39} = \frac{12}{1560} = \frac{1}{130} \approx 0{,}0077$. b) Probabilità totali: $\frac{4}{40} \cdot \frac{3}{39} + \frac{36}{40} \cdot \frac{4}{39} = \frac{12 + 144}{1560} = \frac{1}{10}$: la stessa probabilità della prima carta.

15. a) $p(A) = \frac{1}{3}$, $p(B) = \frac{2}{3}$: $p(R) = \frac{1}{3} \cdot \frac{4}{10} + \frac{2}{3} \cdot \frac{7}{10} = \frac{4 + 14}{30} = \frac{3}{5}$. b) $p(A|R) = \frac{4/30}{3/5} = \frac{2}{9} \approx 0{,}222$.

16. $p(A \cap B) = 0{,}3 + 0{,}5 - 0{,}65 = 0{,}15 = 0{,}3 \cdot 0{,}5$: indipendenti. $p(A|\overline{B}) = p(A \cap \overline{B}) / p(\overline{B}) = (0{,}3 - 0{,}15) / 0{,}5 = 0{,}3 = p(A)$, come deve essere per eventi indipendenti.

17. $\frac{5}{8} \cdot \frac{4}{7} \cdot \frac{3}{6} = \frac{60}{336} = \frac{5}{28} \approx 0{,}179$. La nera può essere la prima, la seconda o la terza, e i tre ordini hanno la stessa probabilità: $3 \cdot \frac{5}{28} = \frac{15}{28} \approx 0{,}536$ (controllo: $C_{3,1} \cdot C_{5,2} / C_{8,3} = \frac{30}{56} = \frac{15}{28}$).

18. a) $p(D) = 0{,}6 \cdot 0{,}03 + 0{,}4 \cdot 0{,}05 = 0{,}018 + 0{,}020 = 0{,}038$. b) $p(B|D) = \frac{0{,}020}{0{,}038} = \frac{10}{19} \approx 0{,}526$.

19. a) 200 malati: 190 positivi, 10 negativi; 9 800 sani: 980 positivi, 8 820 negativi. b) $p(M|+) = \frac{190}{190 + 980} = \frac{19}{117} \approx 16{,}2\%$; $p(\overline{M}|-) = \frac{8820}{8820 + 10} = \frac{882}{883} \approx 99{,}89\%$.

20. Un passo alla volta: dopo il primo positivo p = $\frac{19}{117}$; dopo il secondo $\frac{(19/117) \cdot 0{,}95}{(19/117) \cdot 0{,}95 + (98/117) \cdot 0{,}1} = \frac{18{,}05}{18{,}05 + 9{,}8} =$

$\frac{361}{557} \approx 64{,}8\%$. In un solo passo: $\frac{0{,}02 \cdot 0{,}95^2}{0{,}02 \cdot 0{,}95^2 + 0{,}98 \cdot 0{,}1^2} = \frac{0{,}01805}{0{,}02785} = \frac{361}{557}$, lo stesso risultato.

21. Poiché $A = (A \cap B) \cup (A \cap \overline{B})$ con i due eventi incompatibili, $p(A \cap \overline{B}) = p(A) - p(A \cap B) = p(A) - p(A) \cdot p(B) = p(A)(1 - p(B)) = p(A) \cdot p(\overline{B})$: A e $\overline{B}$ sono indipendenti. Applicando lo stesso risultato alla coppia $\overline{B}$, A (l'indipendenza è simmetrica) si ottiene che $\overline{B}$ e $\overline{A}$ sono indipendenti. c.v.d.

22. $A \cap B$ e $\overline{A} \cap B$ sono incompatibili e la loro unione è B; quindi $p(A \cap B) + p(\overline{A} \cap B) = p(B)$. Dividendo per $p(B) > 0$: $p(A|B) + p(\overline{A}|B) = 1$, cioè $p(A|B) = 1 - p(\overline{A}|B)$. c.v.d.

23. Per il teorema della probabilità composta $p(A|B) \cdot p(B) = p(A \cap B) = p(B|A) \cdot p(A)$, e il numero $p(A \cap B)$ è positivo, quindi lo sono anche le due condizionate. Se $p(A|B) = p(B|A)$, dividendo si ottiene $p(B) = p(A)$; viceversa, se $p(A) = p(B)$, dividendo per $p(A) = p(B)$ si ottiene $p(A|B) = p(B|A)$. c.v.d.

24. Traccia. L'articolo riporta $p(\text{casco}|\text{ferito}) = 0{,}9$, cioè la frazione di ciclisti con il casco *fra i feriti*. Per giudicare se il casco protegge bisogna confrontare $p(\text{ferito}|\text{casco})$ e $p(\text{ferito}|\overline{\text{casco}})$. Manca la frazione di ciclisti che usano il casco: se, per esempio, lo usasse il 95% di tutti i ciclisti, fra i feriti i ciclisti con il casco sarebbero meno rappresentati che fra tutti i ciclisti. Inoltre l'articolo non distingue la gravità delle ferite.

25. $\Omega = \{TT, TC, CT, CC\}$, esiti equiprobabili; $A = \{TT, TC\}$, $B = \{TT, CT\}$, $C = \{TT, CC\}$, ciascuno di probabilità $\frac{1}{2}$. Ogni intersezione di due di essi contiene il solo esito TT, di probabilità $\frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2}$: sono indipendenti a due a due. Ma anche $A \cap B \cap C$ contiene il solo esito TT e ha probabilità $\frac{1}{4}$, mentre $p(A) \cdot p(B) \cdot p(C) = \frac{1}{8}$. Infatti, se si verificano A e B, C è certo: C non è indipendente dalla coppia.

26. a) $p(S|O) = \frac{0{,}3 \cdot 0{,}4}{0{,}3 \cdot 0{,}4 + 0{,}7 \cdot 0{,}05} = \frac{0{,}12}{0{,}155} = \frac{24}{31} \approx 77{,}4\%$. b) $p(S|\overline{O}) = \frac{0{,}3 \cdot 0{,}6}{0{,}3 \cdot 0{,}6 + 0{,}7 \cdot 0{,}95} = \frac{0{,}18}{0{,}845} = \frac{36}{169} \approx 21{,}3\%$.

27. Siano R = «piove», F = «è prevista pioggia». a) $p(R|F) = \frac{0{,}25 \cdot 0{,}8}{0{,}25 \cdot 0{,}8 + 0{,}75 \cdot 0{,}1} = \frac{0{,}2}{0{,}275} = \frac{8}{11} \approx 72{,}7\%$. b) $p(R|\overline{F}) = \frac{0{,}25 \cdot 0{,}2}{0{,}25 \cdot 0{,}2 + 0{,}75 \cdot 0{,}9} = \frac{0{,}05}{0{,}725} = \frac{2}{29} \approx 6{,}9\%$.

28. a) I figli di Nn × Nn sono NN, Nn, nN, nn con probabilità $\frac{1}{4}$ ciascuno; il pelo nero esclude nn: $p(\text{Nn} \mid \text{nero}) = \frac{1/2}{3/4} = \frac{2}{3}$. b) Se il figlio è NN, i piccoli sono tutti Nn e neri (verosimiglianza 1); se è Nn, ogni piccolo è nero con probabilità $\frac{1}{2}$, e tre neri hanno probabilità $\frac{1}{8}$. $p(\text{Nn} \mid 3 \text{ neri}) = \frac{(2/3) \cdot (1/8)}{(2/3) \cdot (1/8) + (1/3) \cdot 1} = \frac{1/12}{5/12} = \frac{1}{5}$.

29. a) Gli innocenti sono 5000; quelli compatibili sono in media $5000 \cdot 0{,}02 = 100$: $p = \frac{1}{1 + 100} = \frac{1}{101} \approx 1\%$. b) Con 50 innocenti, i compatibili sono in media 1: $p = \frac{1}{1 + 1} = \frac{1}{2}$. c) Il 2% è la probabilità che un innocente abbia quel gruppo sanguigno, $p(\text{compatibile}|\text{innocente})$, non la probabilità $p(\text{innocente}|\text{compatibile})$: è la fallacia del procuratore.

30. a) $p(D|S) = \frac{0{,}04 \cdot 0{,}95}{0{,}04 \cdot 0{,}95 + 0{,}96 \cdot 0{,}02} = \frac{0{,}038}{0{,}0572} = \frac{95}{143} \approx 66{,}4\%$. b) In un solo passo: $\frac{0{,}04 \cdot 0{,}95^2}{0{,}04 \cdot 0{,}95^2 + 0{,}96 \cdot 0{,}02^2} = \frac{0{,}0361}{0{,}036484} = \frac{9025}{9121} \approx 98{,}9\%$. c) Il primo controllo da solo scarterebbe un pezzo buono ogni due difettosi circa; il secondo controllo riduce i pezzi buoni scartati a circa l'1% degli scarti, purché i due controlli sbaglino davvero in modo indipendente.

31. a) Dall'uguaglianza $0{,}99x = 0{,}05(1 - x)$ si ottiene $x = \frac{0{,}05}{1{,}04} \approx 0{,}048$, cioè una prevalenza del 4,8% circa. b) $0{,}99x = 0{,}01(1 - x)$ dà $x = 0{,}01$: l'1%. c) Per le malattie rare il denominatore è dominato dai falsi positivi, $(1 - \text{specificità})(1 - x)$: migliorare la sensibili-

tà dal 99% al 100% cambia poco, migliorare la specificità cambia molto.

32. $p(B) = \frac{1}{3} \cdot \frac{1}{4} + \frac{2}{3} \cdot \frac{1}{2} = \frac{1}{12} + \frac{4}{12} = \frac{5}{12}$; $p(A|B) = \frac{1/12}{5/12} = \frac{1}{5}$; $p(A \cup B) = \frac{1}{3} + \frac{5}{12} - \frac{1}{12} = \frac{2}{3}$. Non sono indipendenti: $p(A|B) = \frac{1}{5} \neq p(A) = \frac{1}{3}$.

33. $p(\text{ragazza}|\text{occhiali}) = \frac{0{,}6 \cdot 0{,}3}{0{,}6 \cdot 0{,}3 + 0{,}4 \cdot 0{,}2} = \frac{0{,}18}{0{,}26} = \frac{9}{13} \approx 69{,}2\%$.

34. Probabilità a priori $\frac{1}{4}$ per ciascuna porta. Verosimiglianze dell'evento «apre la 4»: con l'auto dietro la 1 il conduttore sceglie fra 2, 3, 4: $\frac{1}{3}$; con l'auto dietro la 2 sceglie fra 3 e 4: $\frac{1}{2}$; dietro la 3: $\frac{1}{2}$; dietro la 4: 0. $p(E) = \frac{1}{4}\left(\frac{1}{3} + \frac{1}{2} + \frac{1}{2}\right) = \frac{1}{3}$. A posteriori: porta 1 $\frac{1/12}{1/3} = \frac{1}{4}$; porte 2 e 3 $\frac{1/8}{1/3} = \frac{3}{8}$ ciascuna; porta 4: 0. Conviene cambiare: scegliendo a caso una delle due porte chiuse si vince con probabilità $\frac{3}{8} > \frac{1}{4}$.