



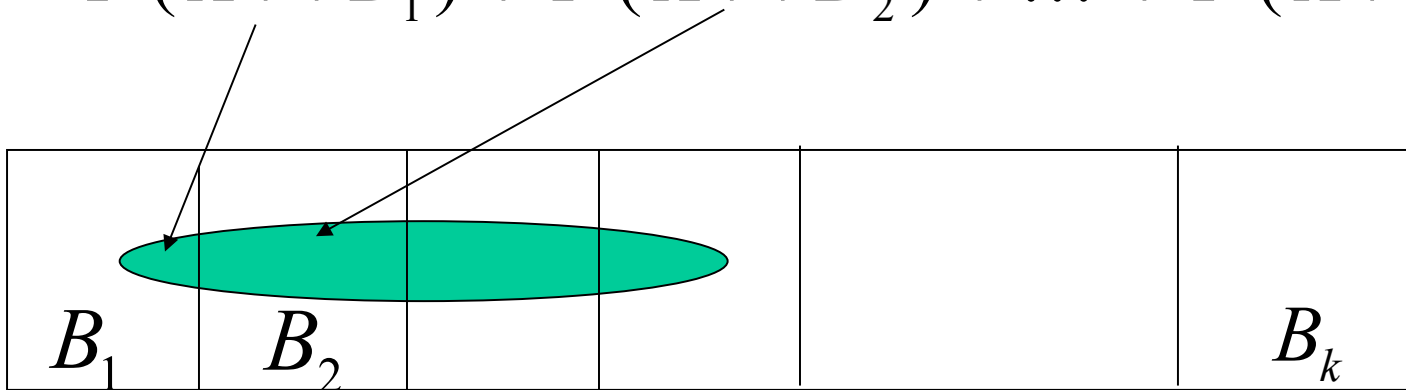
Formula delle probabilità totali (ovvero, come calcolare la probabilità marginale)



Probabilità marginale

Data una partizione $B_1, B_2, \dots, B_k$, e dato un evento A di cui si conoscono le probabilità delle intersezioni con gli elementi della partizione, la probabilità di A è

$$\begin{aligned} P(A) &= P(A \cap S) = P(A \cap (B_1 \cup B_2 \cup \dots \cup B_k)) \\ &= P((A \cap B_1) \cup (A \cap B_2) \cup \dots \cup (A \cap B_k)) \\ &= P(A \cap B_1) + P(A \cap B_2) + \dots + P(A \cap B_k) \end{aligned}$$



Questo metodo si chiama marginalizzazione perché è proprio quello che si usa nelle tabelle doppie per calcolare le frequenze marginali



Formula delle probabilità totali

Riprendendo la formula della probabilità marginale

$$P(A) = P(A \cap B_1) + P(A \cap B_2) + \dots + P(A \cap B_k)$$

e applicando la regola moltiplicativa si ottiene la **formula delle probabilità totali**

$$\begin{aligned} P(A) &= P(A | B_1)P(B_1) + P(A | B_2)P(B_2) + \dots + P(A | B_k)P(B_k) \\ &= \sum_{i=1}^k P(A | B_i)P(B_i) \end{aligned}$$

La prob di A si ottiene come *media pesata* delle prob di A dato B_i , con pesi pari alle prob di B_i



Probabilità totali – esempio

In una città il partito Blu ha il 75% dei consensi e il partito Giallo il 25%. Il sindaco propone la pedonalizzazione del centro storico: da un sondaggio risultano favorevoli il 80% degli elettori del partito Blu e il 40% degli elettori del partito Giallo. Estraendo un elettore a caso, qual è la probabilità che sia favorevole alla pedonalizzazione?

$A = \{\text{favorevole alla pedonalizzazione}\}$

$B = B_1 = \{\text{elettore del partito Blu}\}$

$B' = B_2 = \{\text{elettore del partito Giallo}\}$

$$\begin{aligned} P(A) &= P(A | B_1)P(B_1) + P(A | B_2)P(B_2) \\ &= 0.8 \times 0.75 + 0.4 \times 0.25 = 0.7 \end{aligned}$$



Ragionamento per scenari

- La formula delle probabilità totali corrisponde ad un *ragionamento per scenari* usato molto spesso
- Es. $A = \{\text{l'anno prossimo il fatturato della mia azienda aumenta}\}$ e B_1 , B_2 , B_3 sono 3 scenari relativi all'andamento del settore economico in cui opera l'azienda (crescita, stazionario, recessione). Ragionare per scenari significa attribuire la probabilità non direttamente ad A , ma ad A dato B_1 , A dato B_2 ecc. (perché l'attribuzione è più facile) e poi derivare la probabilità di A tramite la formula delle probabilità totali
 - Supponiamo di attribuire le seguenti probabilità: $P(A|B_1)=0.8$, $P(A|B_2)=0.5$, $P(A|B_3)=0.1$
 - E' necessario attribuire una probabilità anche agli scenari, supponiamo $P(B_1)=0.1$, $P(B_2)=0.7$, $P(B_3)=0.2$ (ovviamente la somma è 1)
 - Pertanto $P(A) = 0.8 \times 0.1 + 0.5 \times 0.7 + 0.1 \times 0.2 = 0.45$



Formula di Bayes

Thomas Bayes (Londra, 1702 – Tunbridge Wells, 17 aprile 1761) è stato un matematico e ministro presbiteriano britannico. Deve la sua fama ai suoi studi nel campo della matematica e della filosofia; è noto soprattutto nella statistica per il suo teorema sulla probabilità condizionata, pubblicato postumo nel 1763.



Formula di Bayes /1

$$P(B_i | A) = \frac{P(A | B_i)P(B_i)}{P(A)}$$
$$= \frac{P(A | B_i)P(B_i)}{P(A | B_1)P(B_1) + P(A | B_2)P(B_2) + \dots + P(A | B_k)P(B_k)}$$

dove B_i è l' i -mo di k eventi costituenti una partizione

Interpretazione: gli eventi della partizione $B_1, \dots, B_k$ sono i possibili *stati di natura*, mentre A è un nuovo fatto (*evidenza empirica*)

Prima di osservare il fatto A , gli stati di natura hanno certe probabilità $P(B_i)$, dette **a priori**. La formula di Bayes consente di *aggiornare* tali probabilità alla luce del fatto A , ottenendo le probabilità $P(B_i | A)$, dette **a posteriori**



Formula di Bayes /2

$$P(B_i | A) = \frac{P(A | B_i)P(B_i)}{P(A | B_1)P(B_1) + P(A | B_2)P(B_2) + \dots + P(A | B_k)P(B_k)}$$

La formula è utile quando si dispone delle

- probabilità marginali $P(B_i)$, dette **probabilità a priori**, e
- probabilità condizionate $P(A | B_i)$, dette **verosimiglianze**.

La formula è detta anche “della probabilità inversa” perché consente di invertire il verso del condizionamento nei casi in cui si dispone di $P(A | B_i)$ ma siamo invece interessati a $P(B_i | A)$



Esempio: chi è il colpevole?

- In un problema di investigazione per omicidio
 - $B1 = \{\text{il colpevole è Tizio}\}$
 - $B2 = \{\text{il colpevole è Caio}\}$
 - $B3 = \{\text{il colpevole non è né Tizio, né Caio}\}$
 - $A = \{\text{uno specifico indizio}\}$
- $B1, B2, B3$ costituiscono una partizione
- L'investigatore attribuisce a $B1, B2, B3$ delle probabilità a priori (cioè prima di osservare l'indizio A)
- Se l'indizio è rilevante le probabilità si modificano, cioè $P(Bi|A) \neq P(Bi)$, e la formula di Bayes consente di calcolarle



Conoscenza probabilistica

- Nella formula di Bayes la conoscenza è rappresentata dalle probabilità assegnate agli stati del mondo $B_1, \dots, B_k$, prima e dopo l'osservazione di un nuovo evento A

- A priori
$$P(B_1) + P(B_2) + \dots + P(B_k) = 1$$

- A posteriori
$$P(B_1 | A) + P(B_2 | A) + \dots + P(B_k | A) = 1$$

- La somma delle probabilità degli stati del mondo è 1 sia a priori che a posteriori
- Pertanto, l'osservazione del nuovo evento A fa aumentare la probabilità di alcuni stati del mondo e diminuire quella di altri (es. un indizio contro Caio aumenta la probabilità che il colpevole sia Caio e riduce la probabilità che sia Tizio)



Aggiornare, aggiornare, aggiornare

- La formula di Bayes esprime in termini matematici la procedura di aggiornamento della conoscenza (probabilità a priori) alla luce della nuova evidenza empirica (evento osservato)
- Un individuo che non segue questa procedura non è razionale (nel senso che viola gli assiomi della probabilità)
- Quando si osserva un nuovo evento le probabilità a posteriori diventano a priori e vengono nuovamente aggiornate tramite la formula di Bayes, e così via in un ciclo senza fine (non si finisce mai di imparare ...)
- Due persone possono partire da probabilità a priori diverse e per questo avere probabilità a posteriori diverse; tuttavia, se due persone osservano la stessa sequenza di eventi e attribuiscono lo stesso valore alle verosimiglianze, mano a mano che gli eventi si verificano le probabilità aggiornate che attribuiscono agli stati del mondo necessariamente convergono



Esempio: test diagnostico /1

- Tipica applicazione in medicina
 - D = soggetto affetto da una certa malattia (Disease)
 - D' = soggetto non affetto da quella malattia
 - T = Test positivo (cioè segnala la malattia)
 - T' = Test negativo
- Si vuole determinare $P(D|T)$ = probabilità che un soggetto per il quale il test dà esito positivo sia effettivamente affetto dalla malattia in questione



Esempio: test diagnostico /2

- Fino a che non si conosce l'esito del test, il soggetto ha una probabilità $P(D)$ di avere la malattia in questione. Tale probabilità viene stimata tramite la ***prevalenza*** nella popolazione
 - Supponiamo che la prevalenza sia di 3 persone ogni 100: pertanto la probabilità che il soggetto abbia la malattia è $P(D) = 0.03$, mentre la probabilità che sia sano (nel senso che non ha quella malattia) è $P(D') = 0.97$



Esempio: test diagnostico /3

- Il secondo elemento necessario per il calcolo è costituito dalla capacità del test di segnalare correttamente chi è sano e chi è malato:
 - Sensitività: probabilità di segnalare correttamente che un soggetto ha quella malattia: $P(T | D) = 0.90$
 - da ciò segue $P(T' | D) = 0.10$ (nel 10% dei soggetti malati il test sbaglia perché non rivela la malattia – *falso negativo*)
 - Specificità: probabilità di segnalare correttamente che un soggetto non ha quella malattia: $P(T' | D') = 0.98$
 - da ciò segue $P(T | D') = 0.02$ (nel 2% dei soggetti sani il test sbaglia perché segnala la malattia – *falso positivo*)



Gli errori dei test diagnostici

Realtà	Esito test	
	<i>Positivo (T)</i> <i>(→ Malato)</i>	<i>Negativo (T')</i> <i>(→ Sano)</i>
<i>Malato (D)</i>	OK P(T D)=0.90 sensitività	Falso negativo P(T' D)=0.10 1-sensitività
<i>Sano (D')</i>	Falso positivo P(T D')=0.02 1-specificità	OK P(T' D')=0.98 specificità

I due tipi di errore sono ben diversi, sia concettualmente che per le conseguenze!



Esempio: test diagnostico /cont.

$$\begin{aligned} P(D | T) &= \frac{P(T | D) P(D)}{P(T | D) P(D) + P(T | D') P(D')} \\ &= \frac{(0.90)(0.03)}{(0.90)(0.03) + (0.02)(0.97)} \\ &= \frac{0.0270}{0.0270 + 0.0194} = \frac{0.0270}{0.0464} = 0.582 \end{aligned}$$

che è molto più grande di $P(D)=0.03$

ma molto più piccolo di $P(T | D) = 0.90$

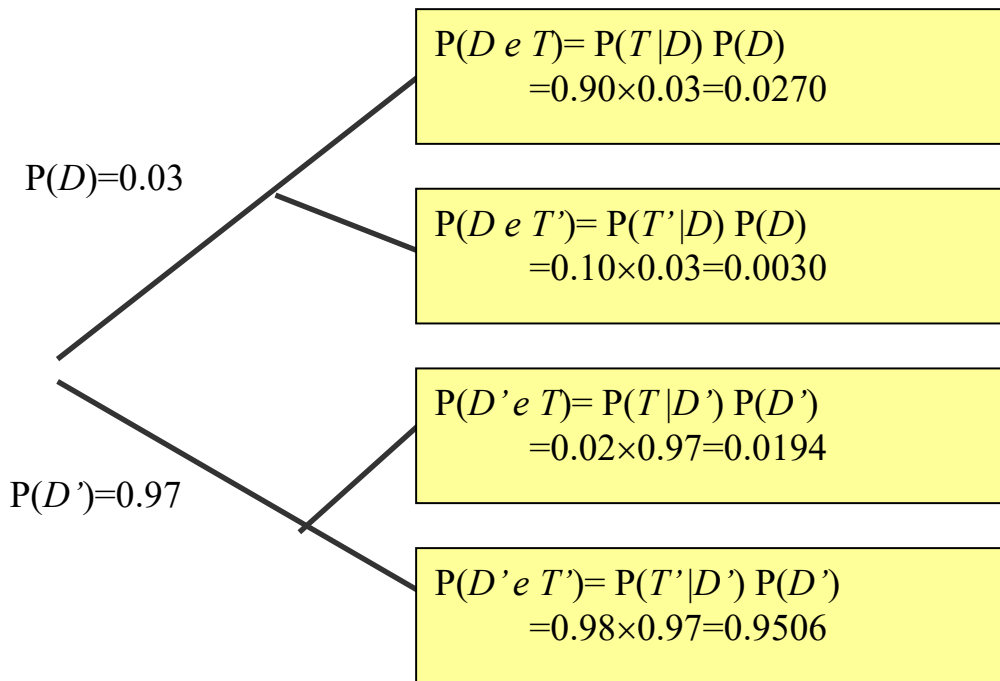
Talvolta i medici confondono $P(D | T)$ con $P(T | D)$, affermando che il paziente positivo al test ha una prob. 0.90 di avere la malattia in questione

Applicando la formula di Bayes ...



Esempio: test diagnostico /cont.

Evento D_i	Probabilità a priori $P(D_i)$	Probabilità condizionata $P(T D_i)$	Probabilità congiunta $P(T D_i) P(D_i)$	Probabilità aggiornata $P(D_i T)$
D = soggetto malato	0.03	0.90	0.0270	0.582
D' = soggetto sano	0.97	0.02	0.0194	0.418



	T	T'	
D	0.0270	0.0030	0.0300
D'	0.0194	0.9506	0.9700
	0.0464	0.9536	



Esempio: test diagnostico /cont.

Cosa succederebbe se lo stesso test fosse usato per una malattia piuttosto rara, con prevalenza di 3 su 1000?

Applicando la formula di Bayes con $P(D)=0.003$:

$$P(D | T) = \frac{(0.90)(0.003)}{(0.90)(0.003) + (0.02)(0.997)} = 0.119$$

Il test produrrebbe una quantità enorme di falsi positivi, cioè soggetti sani per i quali il test segnala la malattia (questo fa capire perché gli screening di massa siano problematici)



La probabilità stimata dai medici

G. Gigerenzer nel libro *Quando i numeri ingannano* riporta i risultati di un esperimento, condotto negli anni Novanta, in cui ha sottoposto ad alcuni medici tedeschi ($\frac{2}{3}$ ospedalieri e $\frac{1}{3}$ liberi professionisti) il seguente quesito:



Per facilitare la diagnosi precoce del cancro al seno le donne, da una certa età in poi, vengono incoraggiate a sottoporsi a intervalli regolari a controlli sistematici, anche se non hanno alcun sintomo conclamato. Lei immagini di condurre, in una certa regione del paese, uno screening mammografico del cancro al seno, e supponga che riguardo alle donne fra i 40 e i 50 anni di questa regione, asintomatiche, che si sottopongono a una mammografia regolare si sappiano le seguenti cose:

*La probabilità che una di loro abbia il cancro al seno è dello **0.8%**. Se una donna ha il cancro al seno, la probabilità che il suo mammogramma risulti positivo è del **90%**; se non ha il cancro al seno, c'è comunque una probabilità del **7%** che il suo mammogramma sia positivo. Immaginiamo, dunque, una donna con un mammogramma positivo: quanto è probabile che abbia effettivamente il cancro?*



La probabilità stimata dai medici /cont.

- La risposta si ottiene calcolando la probabilità a posteriori con i seguenti dati:
 - I soggetti malati sono lo 0.8% (*probabilità a priori*)
 - Tra i malati, il test viene positivo nel 90% dei casi (*sensibilità*)
 - Tra i sani, il test viene positivo nel 7% dei casi (ovvero la *specificità* è del 93%)
- La risposta corretta è che una donna con mammogramma positivo ha una probabilità di avere effettivamente il cancro pari a **9.39%**
- Le risposte fornite dai **24 medici** sono state le seguenti:
 - 8 hanno risposto 90% (cioè la sensibilità)
 - 4 hanno risposto 1% (cioè la probabilità a priori)
 - 8 hanno risposto tra 50% e 80%
 - **solo 4 hanno dato una risposta vicina al vero (di cui 2 in modo fortuito)**





Test diagnostico per la Stupidite

- Consideriamo una ipotetica malattia chiamata Stupidite che colpisce il 10% della popolazione (*prevalenza*)



Prendiamo una persona a caso dalla popolazione. Se non conosciamo l'esito del test diagnostico, la probabilità che abbia la malattia è detta a **probabilità a priori** (coincide con la *prevalenza*, in questo caso 10%)

- Supponiamo di fare la diagnosi usando il test di laboratorio *StupiTest*. Come tutti i test diagnostici, lo *StupiTest* è utile ma non è perfetto:

Il test viene **positivo** (cioè, indica presenza di malattia)

- nel 70% dei soggetti malati (*sensibilità del test*)
- nel 20% dei soggetti sani (*percentuale di falsi positivi*)



Test diagnostico per la Stupidite /cont.

Realtà	Esito test	
	<i>Positivo</i> (→ <i>Malato</i>)	<i>Negativo</i> (→ <i>Sano</i>)
<i>Malato</i>	OK sensitività = 70%	Falso negativo 1-sensitività = 30%
<i>Sano</i>	Falso positivo 1-specificità = 20%	OK specificità = 80%

Dunque, ogni 100 malati sottoposti a test, 70 risultano positivi (giusto!) e 30 negativi (sbagliato!)

E ogni 100 sani sottoposti a test, 20 risultano positivi (sbagliato!) e 80 negativi (giusto!)



Test diagnostico per la Stupidite /cont.

- Supponiamo che Tizio abbia effettuato il test, ottenendo un esito positivo: qual è la probabilità che Tizio sia davvero affetto da Stupidite? Dobbiamo calcolare la probabilità a posteriori

Prendiamo una persona a caso dalla popolazione. Se sappiamo che ha effettuato il test ed è risultato positivo, la probabilità che abbia la malattia è detta a ***probabilità a posteriori***

- Quanto vale la probabilità a posteriori? In questa situazione ipotetica, siccome il test è risultato positivo, la probabilità a posteriori deve essere maggiore di quella a priori (10%). E' forse uguale alla sensibilità del test (70%)?
- La risposta si ottiene con la famosa ***formula di Bayes*** oppure con il ***metodo delle frequenze***, cioè convertendo le probabilità in frequenze di una ipotetica popolazione – vediamo come si applica questo metodo





Test diagnostico per la Stupidite /cont.

- Ricapitoliamo:
 - I soggetti malati sono il 10%
 - Tra i malati, il test viene positivo nel 70% dei casi
 - Tra i sani, il test viene positivo nel 20% dei casi
- Cosa ci aspettiamo sottoponendo al test 100 soggetti?
 - 10 sono malati e 90 sono sani
 - Tra i 10 malati, il test viene positivo a 7 soggetti
 - Tra i 90 sani, il test viene positivo a 18 soggetti



Nel complesso il test viene positivo a 25 soggetti, di cui 7 malati e 18 sani → se prendiamo un soggetto a caso tra i 25 che hanno effettuato il test con risultato positivo, la probabilità che sia malato è $7/25 = 28\%$: questa è la probabilità a posteriori!

 = test positivo  = test negativo

m	m	m	m	m	m	m	m	m	m

