CAPITOLO 2

IL CALCOLO DELLE PROBABILITA'

1. INTRODUZIONE

Il calcolo delle probabilità è il ramo della matematica che si occupa di elaborare dei modelli per descrivere situazioni caratterizzate dall'incertezza su ciò che accadrà nel futuro. Nata dallo studio dei giochi d'azzardo, oggi questa disciplina trova applicazione in molti settori: fisica, ingegneria, informatica, statistica, medicina, controllo della qualità, gestione della sicurezza delle comunicazioni, ecc. Fenomeni il cui risultato non può essere previsto con certezza, come ad esempio il lancio di un dado o l'estrazione di un numero al lotto, sono detti **esperimenti aleatori** o **casuali**. Si dice **spazio campionario** o **spazio degli eventi** l'insieme di tutti i possibili esiti di un esperimento aleatorio ed **evento** ogni suo sottoinsieme. Eventi particolari:

- evento **elementare**: evento rappresentato da un sottoinsieme dello spazio campionario costituito solo da un elemento;
- evento **certo**: evento rappresentato dall'intero spazio campionario;
- evento **impossibile**: evento rappresentato dall'insieme vuoto.

ESEMPIO

Esperimento aleatorio: si lancia un dado regolare a 6 facce e si osserva che numero esce.

I possibili esiti dell'esperimento sono: 1, 2, 3, 4, 5, 6, quindi lo **spazio campionario** è l'insieme $\{1,2,3,4,5,6\}$.

Un **evento** possibile è E = "esce un multiplo di 3", per cui E = $\{3,6\}$.

Un **evento elementare** è A = "esce il numero 5", per cui A = $\{5\}$.

Un **evento certo** è B = "esce un numero minore di 10", per cui B = $\{1,2,3,4,5,6\}$.

Un **evento impossibile** è C = "esce un numero maggiore di 10", per cui C = $\varnothing$.

Avendo definito gli eventi come insiemi, si possono definire le operazioni tra eventi come le operazioni ordinarie tra insiemi.

OPERAZIONI TRA EVENTI

Dati due eventi A e B appartenenti ad uno spazio campionario S, si definisce:

- **evento unione** (o **somma logica**) di A e B, e si indica con $A \cup B$, l'evento che si realizza quando si realizzano A **o** B o entrambi;
- **evento intersezione** (o **prodotto logico**) di A e B, e si indica con $A \cap B$, l'evento che si realizza quando si realizzano entrambi gli eventi A **e** B;

- **evento contrario** di A, e si indica con $\overline{A}$, l'evento che si realizza quando **non** si realizza A.

ESEMPIO

Si lanci un dado regolare a sei facce e si considerino i seguenti eventi:

A = "esce un numero maggiore o uguale a 3"

B = "esce un numero pari"

Si esprima a parole ed in notazione insiemistica l'evento unione, l'evento intersezione e gli eventi contrari di A e B.

A∪B = "esce un numero pari **o** maggiore o uguale a 3" = $\{2, 3, 4, 5, 6\}$

A∩B = "esce un numero pari maggiore o uguale a 3" = $\{4, 6\}$

$\overline{A}$ = "esce un numero minore di 3" = $\{1, 2\}$

$\overline{B}$ = "non esce un numero pari" = $\{1, 3, 5\}$

Due eventi si dicono **incompatibili** se il verificarsi di uno esclude il verificarsi contemporaneo dell'altro, ossia se la loro intersezione è l'evento impossibile. In caso contrario gli eventi si dicono **compatibili**. Gli eventi A e B dell'esempio precedente sono compatibili.

2. IL CONCETTO DI PROBABILITA'

Intuitivamente la probabilità di un evento E è un numero che esprime il grado di fiducia attribuito al verificarsi di E. Ad esempio la probabilità che esca il numero 5 nel lancio di un dado è data da 1 caso su 6 cioè da $\frac{1}{6}$. I casi concreti però non sono sempre così semplici e non possono essere affrontati tutti nello stesso modo. Si danno perciò tre definizioni diverse di probabilità: classica, frequentista e soggettiva da adattare a tutti i casi da analizzare.

DEFINIZIONE CLASSICA DI PROBABILITÀ

Sia S uno spazio campionario formato da n eventi elementari, detti **casi possibili**, aventi tutti la stessa probabilità di verificarsi e sia E un evento relativo ad S formato da k eventi elementari, detti **casi favorevoli**. Si definisce probabilità dell'evento E, e si indica con $p(E)$, il rapporto tra il numero dei casi favorevoli ed il numero dei casi possibili:

$$p(E) = \frac{k}{n}$$

<u>Casi particolari</u>: se un evento è certo la sua probabilità è uguale ad 1; se un evento è impossibile, la sua probabilità è uguale a 0. Inoltre, poiché il numero di casi favorevoli non può mai superare quello dei casi possibili, la probabilità di un evento E è un numero sempre compreso tra 0 e 1. Pertanto:

$$0 \leq p(E) \leq 1$$

ESEMPI

1. Qual è la probabilità che lanciando due monete esca due volte testa?

Lo spazio campionario è dato da:

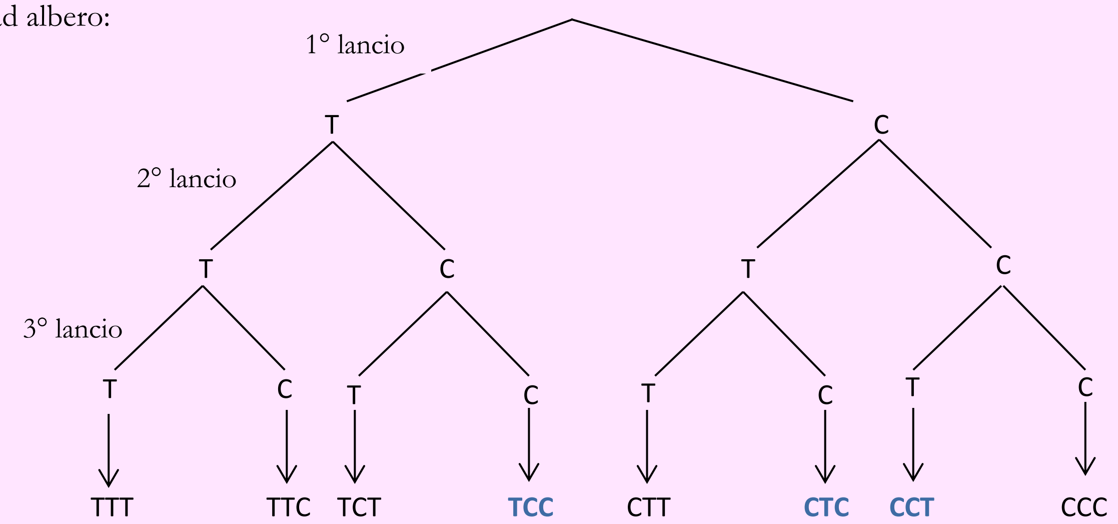
$$S = \{TT; TC; CT; CC\}$$

dove C sta per croce e T per testa. Quindi

$$p(TT) = \frac{1}{4} = 0.25 = 25\%$$

2. Si lancia una moneta non truccata per tre volte consecutive. Calcolare la probabilità che esca esattamente due volte "croce".

Per individuare tutti i casi possibili, ossia lo spazio campionario, si può utilizzare un diagramma ad albero:



Quindi vi sono 3 casi favorevoli (in blu) su 8 casi possibili. Perciò:

$$p(esattamente\ 2\ croci) = \frac{3}{8} = 0.375 = 37.5\%$$

3. Si lancino due dadi non truccati a sei facce. Calcolare la probabilità che la somma dei numeri ottenuti sia maggiore di 9.

L'analisi dei casi possibili si può semplificare utilizzando una tabella a doppia entrata in cui i dati in entrata sono i punteggi dei due dadi rispettivamente ed in uscita si ha la somma dei dadi:

Primo dado / Secondo dado	**1**	**2**	**3**	**4**	**5**	**6**
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	**10**
5	6	7	8	9	**10**	**11**
6	7	8	9	**10**	**11**	**12**

$$p(somma > 9) = \frac{6}{36} = \frac{1}{6} = 0.1\overline{6} \approx 16.7\%$$

4. Un'urna contiene 12 palline numerate da 1 a 12. Si estraggono a caso successivamente tre palline, reimmettendo nell'urna la pallina estratta prima della successiva estrazione. Calcolare la probabilità che la prima e la seconda pallina estratta abbiano numeri dispari.

Per calcolare il numero di casi favorevoli e di casi possibili è necessario ricorrere al teorema fondamentale del calcolo combinatorio, visto nel capitolo 6.

Casi possibili: $12 \cdot 12 \cdot 12 = 1.728$

Casi favorevoli: $6 \cdot 6 \cdot 12 = 432$

$$p = \frac{432}{1.728} = \frac{1}{4} = 0.25 = 25\%$$

DEFINIZIONE FREQUENTISTA (O STATISTICA) DI PROBABILITÀ

Si consideri un evento E relativo ad un esperimento ripetibile nelle medesime condizioni quante volte si vuole. Si definisce probabilità dell'evento E la sua frequenza relativa osservata in un numero sufficientemente grande di prove, ossia il rapporto fra il numero di prove in cui l'evento si è verificato ed il numero totale delle prove effettuate.

Un importante punto di contatto tra la definizione classica e quella frequentista è dato dal teorema noto come:

LEGGE DEI GRANDI NUMERI

Ad un dato evento E sia attribuibile una probabilità sia secondo la definizione classica sia secondo quella frequentista; allora, in una serie di prove ripetute, l'evento E si manifesta con una frequenza relativa che tende a coincidere, al crescere del numero delle prove, con il valore teorico della sua probabilità.

Questo teorema è alla base di un importante metodo per valutare la probabilità di un evento in tutti i casi in cui l'approccio matematico non è noto o comunque troppo complesso. Esso consiste nell'effettuare al computer un gran numero di simulazioni del fenomeno in esame e calcolare la frequenza relativa dell'evento di cui si vuole calcolare la probabilità. Tale metodo è detto **metodo Monte Carlo**.

DEFINIZIONE SOGGETTIVA DI PROBABILITÀ

Si definisce probabilità di un evento E la misura del grado di fiducia che un individuo coerente assegna, in base alle informazioni ed alle proprie opinioni, al verificarsi dell'evento E. Tale numero si può pensare come il prezzo p che l'individuo è disposto a pagare per ricevere l'importo 1 se l'evento E si verifica e l'importo 0 se l'evento E non si verifica.

Ciascuna di queste tre definizioni presenta degli inconvenienti.

- La definizione classica prevede che la spazio degli eventi sia finito e che gli eventi abbiano tutti la stessa probabilità di verificarsi e questo restringe non poco il campo di applicabilità.
- La definizione frequentista richiede la ripetibilità dell'evento nelle stesse condizioni un numero sufficientemente grande di volte. Ma quanto deve essere questo numero? Ed è possibile ripetere le prove sempre nelle stesse identiche condizioni? Inoltre, se cambiasse il numero di prove, cambierebbe il risultato.
- La definizione soggettiva, proprio per definizione, dipende dalle informazioni che ogni individuo possiede circa l'evento da analizzare.

E' stato quindi necessario formulare una definizione di probabilità che possa adattarsi a tutte le situazioni: **LA DEFINIZIONE ASSIOMATICA DELLA PROBABILITÀ**. In quest'ottica il concetto di probabilità diventa un concetto primitivo, che non viene definito esplicitamente, ma è descritto dagli assiomi assunti a fondamento della teoria stessa; cioè:

1. Dato uno spazio campionario S, la probabilità $p(E)$ di un evento E di S è un numero reale tale che:
   - $0 \leq p(E) \leq 1$;
   - se E = S, allora $p(E) = 1$.
2. Se A e B sono due eventi incompatibili, allora $p(A \cup B) = p(A) + p(B)$

Quindi si è liberi di seguire l'approccio (classico, frequentista o soggettivo) che si ritiene più consono al caso da studiare a patto che vengano sempre rispettati gli assiomi 1 e 2 sopra citati.

OSSERVAZIONE: qui di seguito si forniscono alcuni esempi di calcolo della probabilità di eventi non equiprobabili o in caso di spazi campionari infiniti.

ESEMPI

1. **DADO TRUCCATO**. Si lancia una volta un dado a sei facce truccato in modo tale che la probabilità di ottenere ciascun numero sia proporzionale al numero stesso. Qual è la probabilità che esca un numero dispari?

   Si sa che la probabilità che esca un numero è proporzionale al numero stesso, cioè:

n°	1	2	3	4	5	6
p(n°)	t	$2t$	$3t$	$4t$	$5t$	$6t$

   con t costante di proporzionalità.

   Per gli assiomi 1 e 2, la somma delle probabilità di tutti gli eventi elementari deve essere 1, quindi:

$$t + 2t + 3t + 4t + 5t + 6t = 1 \quad \Rightarrow \quad 21t = 1 \quad \Rightarrow \quad t = \frac{1}{21}$$

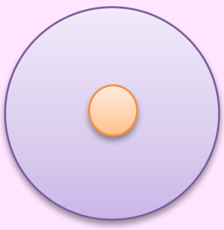
Aggiornando la tabella:

n°	1	2	3	4	5	6
p(n°)	$\frac{1}{21}$	$\frac{2}{21}$	$\frac{3}{21}$	$\frac{4}{21}$	$\frac{5}{21}$	$\frac{6}{21}$

Dal momento che i tre eventi "esce il n° 1", "esce il numero 3" ed "esce il numero 5" sono incompatibili, la probabilità che esca un numero dispari è data da:

$$p(dispari) = p(1) + p(3) + p(5) = \frac{1}{21} + \frac{3}{21} + \frac{5}{21} = \frac{9}{21} = \frac{3}{7} \approx 0.429 = 42.9\%$$

2. **SPAZI INFINITI**. Un tiratore lancia una freccetta su un bersaglio circolare del raggio di 40 cm. Sapendo che il tiratore colpisce il bersaglio e che tutti i punti di esso hanno la stessa probabilità di essere colpiti, qual è la probabilità di colpire un punto distante meno di 8 cm dal centro?



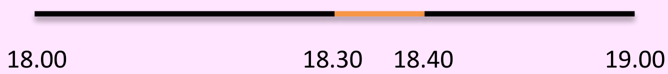
Casi possibili: tutti i punti del cerchio viola

Casi favorevoli: tutti i punti del cerchio arancione

Quindi: $p = \frac{area\ cerchio\ arancio}{area\ cerchio\ viola} = \frac{\pi \cdot 8^2}{\pi \cdot 40^2} = \frac{1}{25} = 0.04 = 4\%$

3. **SPAZI INFINITI**. Federica e Marco si danno appuntamento in piazza del Duomo alle ore 18.00, con l'accordo che se uno dei due fosse in ritardo l'altro lo aspetterà fino alle 19.00. Federica arriva puntuale all'appuntamento, ma Marco non c'è. Calcolare la probabilità che Marco arrivi tra le 18.30 e le 18.40.

Si consideri l'intervallo temporale schematizzato:



Casi possibili: 60 minuti

Casi favorevoli: 10 minuti

Quindi: $p = \frac{10}{60} = \frac{1}{6} = 0.1\overline{6} \approx 16.7\%$

3. TEOREMI SUL CALCOLO DELLE PROBABILITA'

I teoremi che verranno enunciati in seguito si fondano solo sugli assiomi 1 e 2 del paragrafo precedente, quindi sono applicabili indipendentemente dall'approccio scelto per il calcolo della probabilità.

PROBABILITÀ DELL'EVENTO CONTRARIO.

Sia A un evento ed $\overline{A}$ il suo evento contrario, allora:

$$p(\bar{A}) = 1 - p(A)$$

Come conseguenza:

$$p(\emptyset) = 0$$

PROBABILITÀ DELLA DIFFERENZA DI DUE EVENTI.

Siano A e B due eventi tali che A $\subseteq$ B, allora:

$$p(B - A) = p(B) - p(A)$$

PROBABILITÀ DELL'UNIONE DI DUE EVENTI.

Siano A e B due eventi, allora:

$$p(A \cup B) = p(A) + p(B) - p(A \cap B)$$ **relazione di Boole**

PROBABILITÀ CONDIZIONATA

Siano A e B due eventi, con B tale che $p(B) \neq 0$. Si definisce probabilità condizionata dell'evento A, sapendo che si è già verificato un evento B, e si indica $p(A|B)$, il rapporto seguente:

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

Come formula inversa si ha la **formula delle probabilità composte**:

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

oppure:

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

Due eventi si dicono **indipendenti** se il verificarsi dell'uno non altera la probabilità che si verifichi l'altro, cioè se $p(A|B) = p(A)$ o, analogamente, se $p(B|A) = p(B)$. In caso contrario gli eventi si dicono **dipendenti**. Nel caso di eventi indipendenti la formula delle probabilità composte diventa:

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

ESEMPI

1. Si lancia per 5 volte una moneta non truccata. Calcolare la probabilità che esca croce almeno una volta.

L'evento E="esce croce almeno una volta" è il contrario dell'evento A="non esce mai croce" cioè "esce 5 volte testa". Quindi

$$p(E) = p(\bar{A}) = 1 - p(A) = 1 - \frac{1}{2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2} = 1 - \frac{1}{64} = \frac{63}{64}$$

2. In un'urna vi sono 10 palline numerate da 1 a 10. Calcolare la probabilità che, estraendo a caso una pallina, su di essa vi sia un numero divisibile per 2 o per 3.

Gli eventi in gioco sono:

A = "n° divisibile per 2"; B = "n° divisibile per 3"; A$\cap$B = "n° divisibile per 6", quindi:

$$p(A \cup B) = p(A) + p(B) - p(A \cap B) = \frac{5}{10} + \frac{3}{10} - \frac{1}{10} = \frac{7}{10}$$

3. Da un'urna contenente 30 palline numerate da 1 a 30 si estraggono successivamente due palline, senza rimettere la prima pallina nell'urna. Determinare la probabilità che il primo numero estratto sia divisibile per 6 ed il secondo sia pari?

Gli eventi in gioco sono:

A = "primo n° divisibile per 6"; B = "secondo n° pari".

Si deve calcolare la probabilità della loro intersezione. Quindi:

$p(A) = \frac{5}{30} = \frac{1}{6}$

$p(B|A)$ è la probabilità di estrarre un numero pari dopo che dall'urna è stata estratta una pallina con multiplo di 6 (e quindi pari), quindi:

$p(B|A) = \frac{14}{29}$

Da cui:

$$p(A \cap B) = P(A) \cdot P(B|A) = \frac{1}{6} \cdot \frac{14}{29} = \frac{7}{87}$$

4. Si lancia un dado per due volte consecutive. Gli eventi A = "primo n° uguale a 2" e B = "secondo n° uguale a 5" sono indipendenti?

$$p(A) = \frac{1}{6} \quad \text{e} \quad p(B) = \frac{1}{6}$$

$$p(A \cap B) = \frac{1}{36} = p(A) \cdot p(B)$$

quindi i due eventi sono indipendenti, infatti il numero che è uscito la prima volta non influenza il numero che uscirà col secondo lancio.

TEOREMA DELLA PROBABILITÀ TOTALE.

Si consideri lo spazio campionario S e siano $H_1, H_2, \ldots, H_n$ una partizione di S, cioè siano eventi di probabilità non nulla di S a due a due disgiunti, la cui unione è S stesso:

$$S = H_1 \cup H_2 \cup \ldots \cup H_n$$

Considerato un qualunque evento A di S gli insiemi $A \cap H_1$, $A \cap H_2$, …, $A \cap H_n$ risultano a loro volta a due a due disgiunti e la loro unione è A. Proprio per il fatto che sono disgiunti:

$$p(A) = p(A \cap H_1) + p(A \cap H_2) + \cdots + p(A \cap H_n)$$

Per la formula delle probabilità composte è:

$$p(A) = p(A|H_1) \cdot p(H_1) + p(A|H_2) \cdot p(H_2) + \cdots + p(A|H_n) \cdot p(H_n)$$

che è detta formula della **probabilità totale** o **completa** o **formula di disintegrazione**.

ESEMPIO

Un esperto di cavalli ritiene che il purosangue Star sia più forte se corre con la pioggia. In particolare, l'esperto stima che Star possa vincere con probabilità del 15% in caso di tempo asciutto e con probabilità del 35% in caso di pioggia. Il servizio meteorologico prevede , per l'ora della gara, tempo asciutto con probabilità del 40%. Determinare la probabilità che Star vinca la gara.

A = "tempo asciutto" $\bar{A}$ = "pioggia"

B = "Star vince"

Si sa che:

$p(B|A) = 0.15$: Star vince con probabilità del 15% in caso di tempo asciutto;

$p(B|\bar{A}) = 0.35$: Star vince con probabilità del 35% in caso di pioggia;

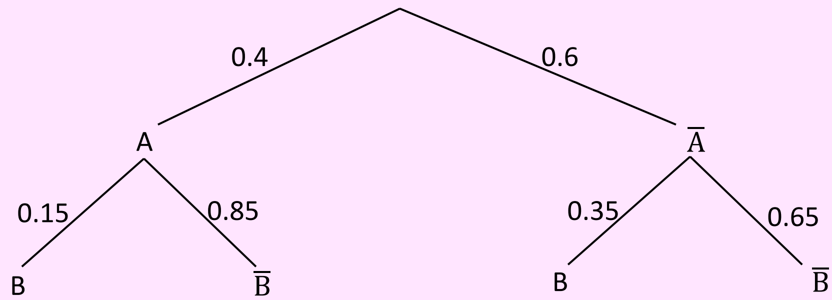
$p(A) = 0.40$: tempo asciutto con probabilità del 40%;

$p(B) = ?$

Gli eventi A e $\bar{A}$ costituiscono una partizione dello spazio campionario "tempo durante la gara", quindi per il teorema della probabilità totale:

$$p(B) = p(B|A) \cdot p(A) + p(B|\bar{A}) \cdot p(\bar{A}) = p(B|A) \cdot p(A) + p(B|\bar{A}) \cdot \big(1 - p(A)\big) =$$
$$= 0.15 \cdot 0.4 + 0.35 \cdot (1 - 0.4) = 0.27 = 27\%$$

Si può rappresentare il problema anche mediante un diagramma ad albero:



OSSERVAZIONI: nel caso si rappresenti un problema, come nel caso precedente, con un diagramma ad albero, è importante tenere presente che:

1. La somma delle probabilità corrispondenti ai rami che escono da uno stesso nodo è sempre uguale ad 1.
2. La probabilità dell'evento rappresentato da un cammino (cioè da una sequenza di rami) è il prodotto delle probabilità segnate sui rami da cui è costituito.
3. La probabilità di un evento è la somma delle probabilità di tutti i cammini che conducono ad esso.

IL TEOREMA DI BAYES.

Dalla formula delle probabilità composte:

$$p(A \cap B) = p(B) \cdot p(A|B) \quad \text{e} \quad p(B \cap A) = p(A) \cdot p(B|A)$$

ma $p(A \cap B) = p(B \cap A)$, quindi:

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

da cui

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

che rappresenta il **teorema di Bayes**. In altre parole, se è alta la probabilità che l'evento A sia effetto della causa B, il fatto che l'evento A si sia verificato aumenta la probabilità, anche se non dà la certezza, che a produrlo sia proprio la causa B. Viceversa, se è bassa la probabilità che l'evento A sia effetto della causa B, il fatto che l'evento A si sia verificato diminuisce la probabilità che a produrre A sia stata proprio la causa B.

ESEMPI

1. Una fabbrica di lampadine ha due linee di produzione: la prima linea produce 400 pezzi al giorno, di cui il 3% difettosi; la seconda linea produce 300 pezzi al giorno di cui l'1% difettosi. Facendo un controllo a caso su tutta la produzione giornaliera si trova una lampadina difettosa. Qual è la probabilità che provenga dalla prima linea?

Eventi in gioco:

L_1: "lampadina scelta proviene dalla linea 1";

L_2: "lampadina scelta proviene dalla linea 2";

D: "lampadina difettosa".

Si sa che:

$$p(L_1) = \frac{400}{700} = \frac{4}{7} \qquad p(L_2) = \frac{300}{700} = \frac{3}{7}$$

$$p(D|L_1) = 0.03 \qquad p(D|L_1) = 0.01$$

Si deve calcolare:

$$p(L_1|D) = \frac{p(L_1) \cdot p(D|L_1)}{p(D)}$$

Per il teorema delle probabilità totali:

$$p(D) = p(D|L_1) \cdot p(L_1) + p(D|L_2) \cdot p(L_2) = 0.03 \cdot \frac{4}{7} + 0.01 \cdot \frac{3}{7} = \frac{3}{140}$$

Quindi:

$$p(L_1|D) = \frac{p(L_1) \cdot p(D|L_1)}{p(D)} = \frac{\frac{4}{7} \cdot 0.03}{\frac{3}{140}} = \frac{4}{5}$$

2. Un test clinico è efficace nel 96% dei casi nell'individuare una data malattia nelle persone effettivamente malate. Il test può però generare dei falsi positivi, cioè può dare esito positivo nell'1% delle persone sane che si sottopongono al test. La malattia colpisce lo 0.8% della popolazione. Se una persona risulta positiva al test, qual è la probabilità che sia effettivamente malata?

Eventi in gioco:

M = "la persona è malata" T^+ = "test risulta positivo"

Si sa che:

$$p(T^+|M) = 0.96 \qquad p(T^+|\bar{M}) = 0.01 \qquad p(M) = 0.008 \qquad p(M|T^+) = ?$$

Per il teorema di Bayes:

$$p(M|T^+) = \frac{p(T^+|M) \cdot p(M)}{p(T^+)}$$

Per la formula della probabilità totale:

$$p(T^+) = p(T^+|M) \cdot p(M) + p(T^+|\bar{M}) \cdot p(\bar{M}) =$$
$$= p(T^+|M) \cdot p(M) + p(T^+|\bar{M}) \cdot \big(1 - p(M)\big) =$$
$$= 0.96 \cdot 0.008 + 0.01 \cdot (1 - 0.008) = 0.0176$$

Quindi:

$$p(M|T^+) = \frac{p(T^+|M) \cdot p(M)}{p(T^+)} = \frac{0.96 \cdot 0.008}{0.0176} = 0.43\overline{63} \approx 43.6\%$$

Quindi una persona che risulta positiva al test ha soltanto circa il 43% di probabilità di risultare effettivamente malata.