Esercizi
Probabilità Condizionata

Esercizio. Da un rilevamento statistico è noto che una certa popolazione è composta per il 40% da fumatori abituali. È noto inoltre che il 5% dei decessi avviene a causa di un certo tipo di tumore. Infine, si è constatato che tra quanti sono deceduti a causa di quel tipo di tumore il 60% erano fumatori abituali. Calcolare la probabilità per i fumatori abituali di morire per un tumore di quel tipo.

Soluzione: consideriamo gli eventi:

$$F = \{\textit{essere fumatore abituale}\}, \quad N = \{\textit{morire di tumore}\}$$

A priori sappiamo che

$$p(F) = \frac{40}{100}, \quad p(N) = \frac{5}{100}, \quad p(F \mid N) = \frac{60}{100}$$

Applicando la formula di Bayes:

$$p(N \mid F) = \frac{p(F \mid N) \cdot p(N)}{p(F)} = \frac{3}{40} \approx 0.075$$

La probabilità per i fumatori di morire per un tumore di quel tipo è del 7.5%.

Un costruttore viene fornito per gli stessi tipi di pezzi per l'80% dalla ditta A e per il restante 20% dalla ditta B. Tali pezzi vengono poi depositati assieme nello stesso magazzino. Per il passato è stato notato che i prodotti di A erano per il 5% difettosi, mentre quelli della ditta B lo erano nella misura del 9%.
Avendo scelto un pezzo a caso dal magazzino ed avendo riscontrato che è difettoso, qual è la probabilità che sia stato fornito da B?

Indicando con D l'evento "pezzo difettoso", si ha:

P(D | A) = 0,05 P(A) = 0,8
P(D | B) = 0,09 P(B) =0,2

Si può quindi determinare la probabilità che avendo estratto a caso un pezzo difettoso, questo sia stato fornito dalla ditta B applicando il teorema di Bayes:

$$P(B/D)=\frac{P(D/B)\cdot P(B)}{P(D/A)\cdot P(A)+P(D/B)\cdot P(B)}=$$

$$P(B/D)=\frac{0{,}09\cdot 0{,}20}{0{,}09\cdot 0{,}20+0{,}05\cdot 0{,}80}=$$

$$=\frac{0{,}018}{0{,}0518}=0{,}3475 \checkmark$$

Esempio 1.49 Un nostro amico lancia due dadi regolari a sei facce, noi scommettiamo che esca almeno un 6. L'amico lancia i dadi, senza mostrarceli, e ci annuncia che la somma dei due dadi vale 9. Qual è la probabilità che vinciamo la scommessa, tenendo conto dell'informazione ricevuta? Qual è la probabilità in assenza dell'informazione?

Lo spazio campionario naturale per questo esperimento aleatorio è $\Omega = \{1, \ldots, 6\}^2$, l'insieme delle coppie (i, j) di numeri in $\{1, \ldots, 6\}$, che muniamo della probabilità P uniforme; osserviamo che $|\Omega| = 36$. I due eventi che appaiono nell'enunciato del problema sono

$$\begin{aligned} A &= \text{“esce almeno un 6”} = \{(i, j) \in \Omega,\ i = 6 \text{ o } j = 6\} \\ B &= \text{“la somma dei due dadi vale 9”} = \{(i, j) \in \Omega,\ i + j = 9\}\,. \end{aligned}$$

In assenza di informazioni sul verificarsi di B, la probabilità di A vale

$$\mathrm{P}(A) = 1 - \mathrm{P}(A^c) = 1 - \frac{|A^c|}{|\Omega|} = 1 - \frac{25}{36} = \frac{11}{36} \simeq 30,6\%\,,$$

avendo utilizzato il fatto che $A^c =$ "non esce nessun 6"$= \{(i, j) \in \Omega,\ 1 \leq i, j \leq 5\}$ e dunque $|A^c| = 25$. In alternativa possiamo scrivere scrivere $A = A_1 \cup A_2$ avendo posto $A_1 :=$ "il primo dado vale 6" e $A_2 :=$ "il secondo dado vale 6", così che $\mathrm{P}(A) = \mathrm{P}(A_1) + \mathrm{P}(A_2) - \mathrm{P}(A_1 \cap A_2) = \frac{1}{6} + \frac{1}{6} - \frac{1}{36} = \frac{11}{36}$.

Per tenere conto dell'informazione che B si è verificato, calcoliamo la probabilità di A *condizionale* a B che è data da

$$\mathrm{P}(A \mid B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} = \frac{|A \cap B|/|\Omega|}{|B|/|\Omega|} = \frac{|A \cap B|}{|B|}.$$

Per determinare la cardinalità di B, possiamo semplicemente elencare i suoi elementi: si ha $B = \{(3,6), (4,5), (5,4), (6,3)\}$ e dunque $|B| = 4$. Analogamente si ha $A \cap B = \{(3,6), (6,3)\}$ (gli elementi di B che contengono almeno un 6), quindi

$$\mathrm{P}(A \mid B) = \frac{|A \cap B|}{|B|} = \frac{1}{2} = 50\%\,,$$

che è maggiore della probabilità in assenza di informazioni.

In altri termini, il fatto di avere un'informazione sull'esperimento (B si è verificato) influenza il "grado di fiducia" che attribuiamo ad A. Ad esempio, se il nostro amico ci comunicasse che la somma dei dadi vale 12, saremmo sicuri che sono usciti due 6; se invece ci comunicasse che la somma dei dadi è inferiore a 6, saremmo sicuri che non è uscito nessun 6. □

Esempio 1.58 (Test clinico) Per determinare la presenza di un virus, viene elaborato un test clinico avente le seguenti caratteristiche:

- *sensibilità*: se il virus è presente, il test risulta positivo nel 99% dei casi;
- *specificità*: se il virus è assente, il test risulta negativo nel 98% dei casi.

È noto che 4 persone su 10 000 hanno il virus, ossia lo 0.04% (*prevalenza* del virus). Supponiamo che un individuo scelto a caso nella popolazione risulti positivo al test: con quale grado di fiducia possiamo affermare che abbia il virus?

Come accade sovente, non è rilevante descrivere nel dettaglio lo spazio campionario. Si considerino gli eventi descritti informalmente da $A =$ "l'individuo ha il virus" e $B =$ "il test dà esito positivo". I dati del problema sono:

$$\mathrm{P}(A) = 4 \cdot 10^{-4}\,, \qquad \mathrm{P}(B|A) = 0.99\,, \qquad \mathrm{P}(B|A^c) = 0.02\,. \tag{1.50}$$

Calcoliamo $\mathrm{P}(A|B)$. Utilizzando la formula di Bayes e la formula delle probabilità totali, si ha

$$\begin{aligned}\mathrm{P}(A|B) &= \frac{\mathrm{P}(B|A)\,\mathrm{P}(A)}{\mathrm{P}(B)} = \frac{\mathrm{P}(B|A)\,\mathrm{P}(A)}{\mathrm{P}(B|A)\,\mathrm{P}(A) + \mathrm{P}(B|A^c)\,\mathrm{P}(A^c)} \\ &= \frac{3.96 \cdot 10^{-4}}{3.96 \cdot 10^{-4} + 1.9992 \cdot 10^{-2}} \simeq \frac{4}{4+200} \simeq 0.02 = 2\%\,,\end{aligned}$$

che è estremamente bassa. Quindi, anche se un individuo risulta positivo al test, è estremamente improbabile che abbia il virus! Questo test dunque darà un grande numero di falsi positivi. □