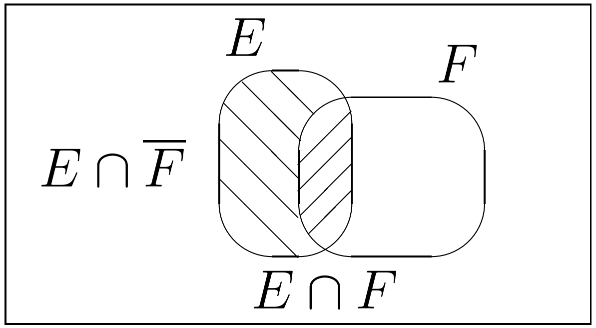
3.3 La formula delle alternative e la formula di Bayes

Proposizione. (Formula delle alternative) Sia F tale che $0 < P(F) < 1$. Se E è un evento qualsiasi risulta

$$P(E) = P(E|F)\, P(F) + P(E|\overline{F})\, P(\overline{F}).$$

Dimostrazione. L'evento E si può esprimere come $E = (E \cap F) \cup (E \cap \overline{F})$, con $E \cap F$ e $E \cap \overline{F}$ eventi incompatibili. Infatti, se un evento elementare appartiene all'evento E, esso inoltre appartiene o all'evento F o al suo complementare $\overline{F}$, e quindi appartiene o all'evento $E \cap F$ oppure a $E \cap \overline{F}$.



Usando la proprietà di additività finita e la regola del prodotto segue:

$$\begin{aligned} P(E) &= P(E \cap F) + P(E \cap \overline{F}) && (E \cap F \text{ e } E \cap \overline{F} \text{ sono incompatibili}) \\ &= P(E|F)\, P(F) + P(E|\overline{F})\, P(\overline{F}), && \text{da cui segue la tesi.} \end{aligned}$$

La formula delle alternative permette di determinare la probabilità di un evento condizionandolo prima alla realizzazione o meno di un altro evento.

Esempio. Da un'urna contenente 5 biglie bianche e 1 biglia rossa, 6 giocatori estraggono a turno 1 biglia a caso, senza reinserimento. Qual è la probabilità che il giocatore k-esimo estragga la biglia rossa?

Soluzione. Posto $A_k = \{$il giocatore k-esimo estrae la biglia rossa$\}$, risulta

$$P(A_1) = \frac{1}{6}, \qquad P(A_2) = P(A_2|A_1)P(A_1) + P(A_2|\overline{A_1})P(\overline{A_1}) = 0 \cdot \frac{1}{6} + \frac{1}{5} \cdot \frac{5}{6} = \frac{1}{6}.$$

Analogamente, si ottiene $P(A_k) = \dfrac{1}{6}$ per $k = 1, 2, \ldots, 6$.

Notiamo che risulta $P(A_k|\overline{A_1} \cap \overline{A_2} \cap \ldots \cap \overline{A_{k-1}}) = \dfrac{1}{6-(k-1)}$, per $k = 1, 2, \ldots, 6$.

Osserviamo inoltre che gli eventi $A_1, A_2, \ldots, A_6$ sono necessari e a 2 a 2 incompatibili.

Esempio. Una compagnia assicuratrice suddivide le persone in due classi: quelle propense a incidenti (il 30%) e quelle che non lo sono (il 70%). Le statistiche mostrano che le persone propense a incidenti hanno probabilità 0,4 di avere un incidente in un anno, mentre per le altre vale 0,2.

(a) Qual è la probabilità che un nuovo assicurato abbia un incidente entro un anno?

(b) Se un nuovo assicurato ha un incidente entro un anno, qual è la probabilità che si tratti di una persona propensa agli incidenti?

Soluzione. Definiamo gli eventi $F = \{$una persona è propensa a incidenti$\}$ ed $E = \{$un nuovo assicurato ha un incidente entro un anno$\}$. Per le ipotesi fatte risulta:

$$P(F) = 0{,}3 \qquad P(\overline{F}) = 0{,}7 \qquad P(E|F) = 0{,}4 \qquad P(E|\overline{F}) = 0{,}2.$$

(a) Si ha: $P(E) = P(E|F)\,P(F) + P(E|\overline{F})\,P(\overline{F}) = 0{,}4 \cdot 0{,}3 + 0{,}2 \cdot 0{,}7 = 0{,}26.$

(b) Risulta:

$$P(F|E) = \frac{P(E \cap F)}{P(E)} = \frac{P(E|F)\,P(F)}{P(E)} = \frac{0{,}4 \cdot 0{,}3}{0{,}26} = \frac{6}{13} \approx 0{,}461.$$

Proposizione. (Formula delle alternative, con n alternative) Se gli eventi $F_1, F_2, \ldots, F_n$ sono a due a due incompatibili, necessari, e ciascuno con probabilità positiva, e se E è un evento qualsiasi, allora risulta

$$P(E) = \sum_{i=1}^{n} P(E|F_i)\, P(F_i).$$

Dimostrazione. Scrivendo

$$E = E \cap S = E \cap \left(\bigcup_{i=1}^{n} F_i \right) = \bigcup_{i=1}^{n} (E \cap F_i) \qquad \text{(con } E \text{ evento qualsiasi)}$$

e osservando che gli eventi $E \cap F_i$, $i = 1, 2, \ldots, n$ sono a due a due incompatibili, per la proprietà di additività finita e per la regola del prodotto si ha infine

$$P(E) = P\left(\bigcup_{i=1}^{n} (E \cap F_i) \right) = \sum_{i=1}^{n} P\left(E \cap F_i\right) = \sum_{i=1}^{n} P(E|F_i)\, P(F_i).$$

Nella formula delle alternative

$$P(E) = \sum_{i=1}^{n} P(E|F_i)\, P(F_i)$$

la probabilità di E viene espressa come media ponderata delle $P(E|F_i)$, dove il peso di ciascun termine è uguale alla probabilità dell'evento F_i, rispetto al quale si condiziona.

Dalle ipotesi che gli eventi $F_1, F_2, \ldots, F_n$ sono a due a due incompatibili e necessari segue che in un esperimento si realizza uno e uno solo degli eventi $F_1, F_2, \ldots, F_n$, che evidentemente costituiscono una partizione dello spazio campionario, e quindi

$$\sum_{i=1}^{n} P(F_i) = P\left(\bigcup_{i=1}^{n} F_i\right) = P(S) = 1,$$

per la proprietà di additività finita.

Esempio. Un'urna contiene 3 monete; la 1^a è non truccata, la 2^a mostra testa con probabilità p, mentre la 3^a dà testa con probabilità $1-p$, con $0 < p < 1$. Se si sceglie una moneta a caso qual è la probabilità che lanciata mostri testa? Se la moneta lanciata mostra testa, qual è la probabilità che si tratti della 2^a?

Soluzione. Definiamo gli eventi $T = \{\text{esce testa}\}$ e $F_j = \{\text{si sceglie la moneta } j\text{-esima}\}$, $j = 1, 2, 3$. Dalle ipotesi fatte segue

$$P(F_j) = \frac{1}{3} \qquad (j = 1, 2, 3),$$

e inoltre

$$P(T|F_1) = 0{,}5 \qquad P(T|F_2) = p \qquad P(T|F_3) = 1 - p.$$

La probabilità di avere testa è quindi

$$P(T) = \sum_{j=1}^{3} P(T|F_j)\, P(F_j) = 0{,}5 \cdot \frac{1}{3} + p \cdot \frac{1}{3} + (1-p) \cdot \frac{1}{3} = \frac{1}{2}.$$

Pertanto,

$$P(F_2|T) = \frac{P(T \cap F_2)}{P(T)} = \frac{P(T|F_2)P(F_2)}{P(T)} = p\,\frac{2}{3}.$$

Esempio. Un vettore booleano di lunghezza 5 contiene 2 bit pari a **1** e 3 bit pari a **0**. Un algoritmo esamina i bit del vettore uno per volta, e si ferma in corrispondenza del primo bit pari a **1**. Qual è la probabilità che l'algoritmo si fermi al passo k-esimo $(1 \leq k \leq 4)$? Qual è la probabilità che il bit successivo al primo **1** sia pari a **0**?

Soluzione. Lo spazio campionario è costituito da $|S| = \binom{5}{2} = 10$ vettori booleani. Ponendo $F_k = \{$l'algoritmo si ferma al passo k-esimo$\}$ $(1 \leq k \leq 4)$, tale evento è costituito da tutte le sequenze di S aventi **0** nei primi $k-1$ bit, **1** nel bit k-esimo, e che negli ultimi $5-k$ bit contengono un solo bit pari a **1**, pertanto:

$$P(F_1) = \frac{4}{10}, \qquad P(F_2) = \frac{3}{10}, \qquad P(F_3) = \frac{2}{10}, \qquad P(F_4) = \frac{1}{10}.$$

Posto $E = \{$il bit successivo al primo **1** è pari a **0**$\}$, si ha

$$P(E|F_1) = \frac{3}{4}, \qquad P(E|F_2) = \frac{2}{3}, \qquad P(E|F_3) = \frac{1}{2}, \qquad P(E|F_4) = 0,$$

e quindi $P(E) = \sum_{k=1}^{4} P(E|F_k)\, P(F_k) = \frac{4}{10} \cdot \frac{3}{4} + \frac{3}{10} \cdot \frac{2}{3} + \frac{2}{10} \cdot \frac{1}{2} + \frac{1}{10} \cdot 0 = \frac{6}{10}.$

Se nell'esempio precedente il vettore booleano ha lunghezza n, allora $|S| = \binom{n}{2}$ e l'evento F_k è costituito dalle sequenze di S aventi **0** nei primi $k-1$ bit, **1** nel bit k-esimo, e che negli ultimi $n-k$ bit contengono un solo bit pari a **1**. Pertanto:

$$P(F_k) = \frac{n-k}{\binom{n}{2}} = \frac{2(n-k)}{n(n-1)}, \qquad P(E|F_k) = \frac{n-k-1}{n-k}, \qquad 1 \leq k \leq n-1,$$

e quindi

$$P(E) = \sum_{k=1}^{n-1} P(E|F_k)\, P(F_k) = \sum_{k=1}^{n-1} \frac{n-k-1}{n-k} \cdot \frac{2(n-k)}{n(n-1)} = \frac{2}{n(n-1)} \sum_{j=0}^{n-2} j,$$

avendo posto $j = n-k-1$. Ricordando che $\sum_{j=1}^{N} j = \frac{N(N+1)}{2}$, si ha

$$P(E) = \frac{2}{n(n-1)} \cdot \frac{(n-1)(n-2)}{2} = \frac{n-2}{n}.$$

Notiamo che: $\displaystyle\sum_{k=1}^{n-1} P(F_k) = \sum_{k=1}^{n-1} \frac{2(n-k)}{n(n-1)} = \frac{2}{n(n-1)} \sum_{j=1}^{n-1} j = \frac{2}{n(n-1)} \frac{n(n-1)}{2} = 1.$

La formula delle alternative $P(E) = \sum_{i=1}^{n} P(E|F_i)\, P(F_i)$ permette di determinare la probabilità di un evento condizionandolo prima alla realizzazione di uno, e uno solo, degli n eventi $F_1, F_2, \ldots, F_n$. Supponiamo ora che E si sia verificato e di voler determinare quali degli eventi alternativi $F_1, F_2, \ldots, F_n$ si sia anch'esso verificato.

Proposizione. (Formula di Bayes) Se E è un evento avente probabilità positiva, e $F_1, F_2, \ldots, F_n$ sono eventi a due a due incompatibili, ciascuno avente probabilità positiva, e necessari, allora

$$P(F_j|E) = \frac{P(E|F_j)\, P(F_j)}{\sum_{i=1}^{n} P(E|F_i)\, P(F_i)} \qquad (j = 1, 2, \ldots, n).$$

Dimostrazione. Dalla definizione di probabilità condizionata, dalla regola del prodotto e dalla formula delle alternative segue immediatamente

$$P(F_j|E) = \frac{P(E \cap F_j)}{P(E)} = \frac{P(E|F_j)\, P(F_j)}{\sum_{i=1}^{n} P(E|F_i)\, P(F_i)} \qquad (j = 1, 2, \ldots, n).$$

Verifichiamo che le probabilità della formula di Bayes sommano all'unità; infatti risulta

$$\sum_{j=1}^{n} P(F_j|E) = \sum_{j=1}^{n} \frac{P(E|F_j)\, P(F_j)}{\sum_{i=1}^{n} P(E|F_i)\, P(F_i)} = 1.$$

Esempio. In un gioco vi sono 3 carte identiche per la forma, la prima con entrambe le facce di colore rosso, la seconda con entrambe le facce di colore nero, la terza con una faccia rossa e una nera. Si sceglie a caso una carta e la si appoggia sul tavolo; se la faccia superiore della carta è rossa, qual è la probabilità che l'altra faccia sia nera?

Soluzione. Indichiamo con F_1, F_2 e F_3 gli eventi riferiti alle 3 carte, e poniamo $R = \{$la faccia superiore della carta scelta è rossa$\}$. Dalla formula di Bayes segue

$$P(F_3|R) = \frac{P(R|F_3)\, P(F_3)}{\sum_{i=1}^{3} P(R|F_i)\, P(F_i)} = \frac{\dfrac{1}{2} \cdot \dfrac{1}{3}}{1 \cdot \dfrac{1}{3} + 0 \cdot \dfrac{1}{3} + \dfrac{1}{2} \cdot \dfrac{1}{3}} = \frac{1}{3}.$$

Notiamo che tale risultato si può ottenere anche come rapporto di casi favorevoli su casi possibili, in quanto una sola delle tre facce rosse ha una faccia nera sul retro.

Esempio. Un sistema di gestione della posta elettronica riceve un messaggio, che si suppone sia *spam* con probabilità 0,7 e *non spam* con probabilità 0,3. Il sistema effettua un controllo su ogni messaggio ricevuto; se riceve un messaggio *spam* lo valuta come tale con probabilità 0,9 (e lo valuta come *non spam* con probabilità 0,1) mentre se riceve un messaggio *non spam* lo valuta come tale con probabilità 0,8 (e lo valuta come *spam* con probabilità 0,2).

(i) Calcolare la probabilità che il sistema valuti come *spam* il messaggio ricevuto.

(ii) Se il sistema ha valutato come *spam* il messaggio ricevuto, qual è la probabilità che invece sia *non spam*?

(iii) Se il sistema ha valutato come *non spam* il messaggio ricevuto, qual è la probabilità che invece sia *spam*?

Soluzione. Per le ipotesi sull'evento $F = \{\text{il messaggio è } spam\}$ si ha

$$P(F) = 0{,}7 \qquad P(\overline{F}) = 0{,}3.$$

Inoltre, posto $E = \{$il sistema valuta come *spam* il messaggio ricevuto$\}$, risulta

$$P(E \,|\, F) = 0{,}9 \qquad P(\overline{E} \,|\, F) = 0{,}1 \qquad P(\overline{E} \,|\, \overline{F}) = 0{,}8 \qquad P(E \,|\, \overline{F}) = 0{,}2.$$

Pertanto, dalla formula delle alternative segue

$$P(E) = P(E \,|\, F)\, P(F) + P(E \,|\, \overline{F})\, P(\overline{F}) = 0{,}9 \cdot 0{,}7 + 0{,}2 \cdot 0{,}3 = 0{,}69.$$

Dalla formula di Bayes si ricavano le probabilità condizionate richieste:

$$P(\overline{F} \,|\, E) = \frac{P(E \,|\, \overline{F})\, P(\overline{F})}{P(E)} = \frac{0{,}2 \cdot 0{,}3}{0{,}69} = \frac{0{,}06}{0{,}69} = 0{,}0870$$

e

$$P(F \,|\, \overline{E}) = \frac{P(\overline{E} \,|\, F)\, P(F)}{P(\overline{E})} = \frac{0{,}1 \cdot 0{,}7}{0{,}31} = \frac{0{,}07}{0{,}31} = 0{,}2258.$$

Esempio. Si lanciano a caso n monete non truccate; per ogni moneta che mostra testa si inserisce una biglia nera in un'urna, mentre per ogni croce si inserisce una biglia bianca. Se poi si estrae a caso una biglia dall'urna, qual è la probabilità che sia nera? Se la biglia estratta è nera, qual è la probabilità che nell'urna vi erano k biglie nere?

Soluzione. Sia $E_k = \{\text{esce testa } k \text{ volte}\} = \{\text{nell'urna vi sono } k \text{ biglie nere e } n-k \text{ bianche}\}$, $k = 0, 1, \ldots, n$. Tali eventi sono a due a due incompatibili, sono necessari, e $P(E_k) = \dfrac{\binom{n}{k}}{2^n} > 0$. Sia $A = \{\text{la biglia è nera}\}$; per la formula delle alternative si ha

$$P(A) = \sum_{k=0}^{n} P(A|E_k)\, P(E_k) = \sum_{k=1}^{n} \frac{k}{n} \cdot \frac{\binom{n}{k}}{2^n} = \frac{1}{2^n} \sum_{k=1}^{n} \binom{n-1}{k-1},$$

essendo $\dfrac{k}{n} \cdot \dbinom{n}{k} = \dfrac{k}{n} \cdot \dfrac{n!}{k!\,(n-k)!} = \dfrac{(n-1)!}{(k-1)!\,(n-k)!} = \dbinom{n-1}{k-1}$. Ponendo $r = k-1$, dal teorema del binomio segue

$$P(A) = \frac{1}{2^n} \sum_{r=0}^{n-1} \binom{n-1}{r} = \frac{1}{2^n} \cdot 2^{n-1} = \frac{1}{2}.$$

Per ricavare $P(E_k|A)$ facciamo uso della formula di Bayes:

$$P(E_k|A) = \frac{P(A|E_k)\,P(E_k)}{\displaystyle\sum_{i=0}^{n} P(A|E_i)\,P(E_i)} = \frac{\dfrac{k}{n}\cdot\dfrac{\binom{n}{k}}{2^n}}{\dfrac{1}{2}} = \begin{cases} 0 & \text{se } k=0 \\ \dfrac{\binom{n-1}{k-1}}{2^{n-1}} & \text{se } k=1,2,\ldots,n. \end{cases}$$

È facile verificare che la somma delle probabilità $P(E_k|A)$ è unitaria:

$$\sum_{k=0}^{n} P(E_k|A) = \sum_{k=1}^{n} \frac{\binom{n-1}{k-1}}{2^{n-1}} = \frac{1}{2^{n-1}}\sum_{r=0}^{n-1}\binom{n-1}{r} = \frac{1}{2^{n-1}}\cdot 2^{n-1} = 1.$$