2.6 Legge delle alternative

Oggetto di questo paragrafo sono alcuni importanti risultati coinvolgenti le probabilità condizionate, molto utili dal punto di vista applicativo. In primo luogo osserviamo che se A e B sono elementi di $\mathscr{F}$, la relazione $A = (A \cap B) \cup (A \cap \overline{B})$ conduce ad affermare che $P(A) = P(A \cap B) + P(A \cap \overline{B})$ e, se si suppone $0 < P(B) < 1$, la regola moltiplicativa comporta $P(A) = P(B)\, P(A|B) + P(\overline{B})\, P(A|\overline{B})$. Gli eventi B e $\overline{B}$, che costituiscono una partizione di Ω, possono essere riguardati come due alternative poiché, per ogni fissato $\omega \in \Omega$ si ha che $\omega \in B$ oppure che $\omega \in \overline{B}$. Quanto detto può essere generalizzato introducendo la nozione di insieme di alternative.

Definizione 2.12 *Sia $(\Omega, \mathscr{F}, P)$ uno spazio di probabilità e sia B un evento di $\mathscr{F}$. Una famiglia $\{B_n;\ n = 1, 2, \ldots, K\}$ di eventi di $\mathscr{F}$ costituisce un insieme di alternative per l'evento B se sono soddisfatte le seguenti proprietà:*

(i) $B_i \cap B_j = \emptyset \quad (i, j = 1, 2, \ldots, K;\ i \neq j)$;

(ii) $\bigcup_{n=1}^{K} B_n = B$;

(iii) $P(B_n) > 0$ *per* $n = 1, 2, \ldots, K$.

Inoltre, se $B = \Omega$, le condizioni (i), (ii) e (iii) si dicono individuare un "insieme completo" di alternative.

In particolare la *(i)* afferma che gli eventi $B_1, B_2, \ldots, B_K$ sono incompatibili, e quindi alternativi, mentre la *(ii)* esprime la circostanza che tali eventi sono anche necessari per B. Sussiste il seguente

Teorema 2.7 *(Legge delle alternative) Sia $\{B_n;\ n = 1, 2, \ldots, K\}$ un insieme completo di alternative e sia A un evento di $\mathscr{F}$. Risulta:*

$$P(A) = \sum_{n=1}^{K} P(A|B_n)\ P(B_n). \tag{2.25}$$

Dimostrazione Poiché per ipotesi gli eventi dell'insieme $\{B_n;\ n = 1, 2, \ldots, K\}$ sono a due a due incompatibili, tali sono anche gli eventi dell'insieme $\{A \cap B_n;\ n = 1, 2, \ldots, K\}$. Inoltre, poiché $\bigcup_{n=1}^{K} B_n = \Omega$, si ha:

$$P(A) = P(A \cap \Omega) = P\Big(A \cap \bigcup_{n=1}^{K} B_n\Big) = P\Big(\bigcup_{n=1}^{K} (A \cap B_n)\Big) = \sum_{n=1}^{K} P(A \cap B_n),$$

da cui, facendo uso della regola moltiplicativa di cui al Teorema 2.6, segue la (2.25). $\square$

Esempio 2.29 Data un'urna contenente biglie numerate con gli interi da 1 a 9 si estragga con equiprobabilità una biglia e successivamente, senza reinserire la prima biglia nell'urna, se ne estragga una seconda. Si calcoli la probabilità che la seconda biglia estratta abbia numero pari.

Si considerino gli eventi $A_i = \{$*la i-esima biglia estratta ha numero pari*$\}$ $(i = 1, 2)$. Dalla legge delle alternative segue:

$$P(A_2) = P(A_1)\, P(A_2|A_1) + P(\overline{A_1})\, P(A_2|\overline{A_1}) = \frac{4}{9} \cdot \frac{3}{8} + \frac{5}{9} \cdot \frac{4}{8} = \frac{4}{9}.$$

$\diamondsuit$

Esempio 2.30 Supponiamo che quotidianamente vengano esaminate le condizioni metereologiche in una prefissata regione. Si consideri l'evento $A_j = \{$*il j-esimo giorno in esame non è piovoso*$\}$ $(j = 1, 2, \ldots)$ e si supponga che $P(A_{j+1}|A_j) = \beta$ e $P(A_{j+1}|\overline{A_j}) = \alpha$ con α e β numeri arbitrari interni all'intervallo $(0, 1)$. Si intende calcolare $P(A_j)\ (j = 1, 2, \ldots)$.

Dalla legge delle alternative segue:

$$P(A_{j+1}) = P(A_j)\,P(A_{j+1}|A_j) + P(\overline{A_j})\,P(A_{j+1}|\overline{A_j}) = P(A_j)\,\beta + [1 - P(A_j)]\,\alpha,$$

da cui si trae:

$$P(A_{j+1}) = (\beta - \alpha)\,P(A_j) + \alpha \qquad (j = 1, 2, \ldots).$$

Abbiamo ottenuto un'equazione ricorsiva, la cui soluzione, ricavata in modo iterativo, è la seguente:

$$\begin{aligned} P(A_j) &= (\beta - \alpha)^{j-1}\,P(A_1) + \alpha \sum_{i=0}^{j-2} (\beta - \alpha)^i \\ &= (\beta - \alpha)^{j-1}\,P(A_1) + \alpha\,\frac{1 - (\beta - \alpha)^{j-1}}{1 - (\beta - \alpha)} \qquad (j = 1, 2, \ldots). \end{aligned}$$

Si noti che il limite

$$\lim_{j \to +\infty} P(A_j) = \frac{\alpha}{1 - \beta + \alpha}$$

non dipende da $P(A_1)$; ciò indica che con il trascorrere dei giorni la probabilità di non avere pioggia tende a non dipendere dalla probabilità con cui ha piovuto il primo giorno. $\diamondsuit$

Teorema 2.8 *Sia $\{B_n;\ n = 1, 2, \ldots, K\}$ un insieme di alternative per l'evento $B \in \mathscr{F}$ e sia A un evento di $\mathscr{F}$. La probabilità di A condizionata da B è esprimibile al seguente modo:*

$$P(A|B) = \sum_{n=1}^{K} P(B_n|B)\,P(A|B_n). \tag{2.26}$$

Dimostrazione Poiché per ipotesi gli eventi dell'insieme $\{B_n;\ n = 1, 2, \ldots, K\}$ sono a due a due incompatibili e tali che $P(B_n) > 0$ per $n = 1, 2, \ldots, K$, usando la proprietà di additività delle probabilità si ha $P(B) = P\Big(\bigcup_{n=1}^{K} B_n\Big) = \sum_{n=1}^{K} P(B_n) > 0$. Si noti che gli eventi dell'insieme $\{A \cap B_n;\ n = 1, 2, \ldots, K\}$ sono incompatibili; poiché per ipotesi $\bigcup_{n=1}^{K} B_n = B$, si ricava:

$$\begin{aligned} P(A|B) &= \frac{P(A \cap B)}{P(B)} = \frac{P\left[A \cap \left(\bigcup_{n=1}^{K} B_n\right)\right]}{P(B)} = \frac{P\Big[\bigcup_{n=1}^{K} (A \cap B_n)\Big]}{P(B)} \\ &= \sum_{n=1}^{K} \frac{P(A \cap B_n)}{P(B)} = \sum_{n=1}^{K} \frac{P(B_n)}{P(B)}\,P(A|B_n). \end{aligned}$$

Osservando che per ogni n si ha $B_n = B_n \cap B$, risulta infine:

$$P(A|B) = \sum_{n=1}^{K} \frac{P(B_n \cap B)}{P(B)}\,P(A|B_n) = \sum_{n=1}^{K} P(B_n|B)\,P(A|B_n).$$

$\square$

Si noti che se nel Teorema 2.8 si pone $B = \Omega$, la (2.26) si identifica con la (2.25).

2.7 Teorema di Bayes

Introdurremo ora un teorema di particolare rilievo concernente le probabilità condizionate; esso riveste grande importanza sia concettuale che applicativa, ma la sua utilizzazione, se non correttamente effettuata, può dar luogo — ed ha invero spesso generato — equivoci e risultati paradossali. Si tratta del celebre *teorema* o *legge di Bayes*, che trovò la sua prima chiara formulazione ad opera appunto di Thomas Bayes nel 1763.

Teorema 2.9 *(Legge di Bayes) Sia $\{B_n;\ n = 1, 2, \ldots, K\}$ un insieme di eventi incompatibili di $\mathscr{F}$ tali che $P(B_n) > 0$ per $n = 1, 2, \ldots, K$ e sia $A \in \mathscr{F}$ un evento con $P(A) > 0$. Se $A \subset \bigcup_{n=1}^{K} B_n$, per $n = 1, 2, \ldots, K$ si ha:*

$$P(B_n|A) = \frac{P(B_n)\ P(A|B_n)}{\displaystyle\sum_{i=1}^{K} P(B_i)\ P(A|B_i)}\,. \tag{2.27}$$

Dimostrazione Essendo per ipotesi $P(A) > 0$ e $P(B_n) > 0$ per $n = 1, 2, \ldots, K$, dalla Definizione 2.11 segue:

$$P(B_n|A) = \frac{P(A \cap B_n)}{P(A)} = \frac{P(B_n)\ P(A|B_n)}{P(A)}. \tag{2.28}$$

Inoltre, poiché per ipotesi $A \subset \bigcup_{i=1}^{K} B_i$, si ha $A = A \cap \Big(\bigcup_{i=1}^{K} B_i\Big) = \bigcup_{i=1}^{K} (A \cap B_i)$. Per ipotesi gli eventi B_i $(i = 1, 2, \ldots, K)$ sono a due a due incompatibili e tali che $P(B_i) > 0$ per $i = 1, 2, \ldots, K$; pertanto risulta:

$$P(A) = P\Big(\bigcup_{i=1}^{K} (A \cap B_i)\Big) = \sum_{i=1}^{K} P(A \cap B_i) = \sum_{i=1}^{K} P(B_i)\ P(A|B_i). \tag{2.29}$$

Sostituendo la (2.29) nella (2.28) segue la (2.27). □

Si noti che la (2.27) è valida in particolare se gli eventi B_n $(n = 1, 2, \ldots, K)$ sono a due a due incompatibili e tali che $\bigcup_{n=1}^{K} B_n = \Omega$. Infatti, in tal caso la condizione $A \subset \bigcup_{n=1}^{K} B_n$ è soddisfatta per ogni evento $A \in \mathscr{F}$.

La legge di Bayes assume un significato particolarmente importante se agli eventi coinvolti si assegna il ruolo di *cause* ed *effetto*. Precisamente, gli eventi B_n possono essere considerati come possibili cause del verificarsi dell'evento A (effetto delle cause considerate) o essere riguardati come ipotesi che ne rendono conto. La condizione $A \subset \bigcup_{n=1}^{K} B_n$ comporta che se si verifica A, si verifica necessariamente uno ed uno solo degli eventi B_n, essendo questi ultimi incompatibili. Quindi, se l'evento A si verifica, una ed una sola delle cause, o delle ipotesi, B_n deve essere intervenuta. Una volta osservato l'evento A, ci si può dunque chiedere quale sia stata la causa o l'ipotesi responsabile della sua occorrenza.

Le probabilità *$P(B_n)$* sono dette probabilità apriori, le probabilità condizionate *$P(B_n|A)$* si dicono a posteriori perché rappresentano le probabiltià che si verifichino gli eventi B_n dopo che si è verificato l'evento A, infine le probabilità *$P(A|B_i)$* sono dette verosimiglianze o probabilità probative.

Esempio 2.31 Si considerino due urne indistinguibili, l'una contenente una biglia bianca e una biglia rossa e l'altra tre biglie rosse e una biglia verde. L'esperimento consiste nello scegliere a caso un'urna e da essa estrarre una biglia. Si vuole calcolare la probabilità che, scegliendo a caso un'urna ed estraendo da essa una biglia che risulta essere rossa, sia stata scelta l'urna i-esima $(i = 1, 2)$.

Si considerino i seguenti eventi: $A = \{$*è stata estratta una biglia rossa*$\}$, $B_i = \{$*è stata scelta l'urna i-esima*$\}$ $(i = 1, 2)$. Poiché non vi è alcuna preferenza nella scelta dell'una o dell'altra urna, $P(B_i) = 1/2$ $(i = 1, 2)$; inoltre, dai dati forniti, risulta $P(A|B_1) = 1/2$ e $P(A|B_2) = 3/4$. Dalla (2.27) segue allora:

$$P(B_1|A) = \frac{P(A|B_1)}{P(A|B_1) + P(A|B_2)} = \frac{2}{5}\,,$$

e quindi $P(B_2|A) = 3/5$. ♢

Esempio 2.32 Due tiratori sparano indipendentemente un colpo ciascuno su di un bersaglio. Siano rispettivamente p e q le probabilità che essi colpiscano il centro del bersaglio. Si è interessati a calcolare la probabilità che l'autore del centro sia stato il primo (secondo) tiratore sapendo che il centro del bersaglio è stato colpito una sola volta.

Si definiscano i seguenti eventi: $A = \{$*il centro del bersaglio è stato colpito una sola volta*$\}$, $B_1 = \{$*i due tiratori sparano un colpo a vuoto*$\}$, $B_2 = \{$*i due tiratori fanno entrambi centro*$\}$, $B_3 = \{$*solo il primo tiratore colpisce il bersaglio*$\}$, $B_4 = \{$*solo il secondo tiratore colpisce il bersaglio*$\}$. Si nota che gli eventi B_1, B_2, B_3, B_4 sono necessari ed incompatibili e tali che $P(B_1) = (1-p)(1-q)$, $P(B_2) = pq$, $P(B_3) = p(1-q)$, $P(B_4) = (1-p)q$. Dal Teorema 2.9 si ricava:

$$P(B_n|A) = \frac{P(B_n)P(A|B_n)}{\sum_{i=1}^{4} P(B_i)\,P(A|B_i)} \qquad (n = 1, 2, 3, 4). \tag{2.30}$$

Dai dati forniti risulta evidentemente $P(A|B_1) = P(A|B_2) = 0$, $P(A|B_3) = P(A|B_4) = 1$, così che dalla (2.30) segue $P(B_1|A) = P(B_2|A) = 0$, mentre per le probabilità richieste si ottiene:

$$P(B_3|A) = \frac{p\,(1-q)}{p+q-2\,p\,q}\,, \qquad P(B_4|A) = \frac{(1-p)\,q}{p+q-2\,p\,q}\cdot$$

◊

È infine qui opportuno menzionare che la legge delle alternative e il teorema di Bayes si estendono facilmente al caso di insiemi numerabili di alternative.