1 Probabilità condizionale

Introduciamo il concetto tramite un esempio.

Esempio 1.1. *Calcolare la probabilità di vincere al gioco della roulette se si punta sui numeri* $\{1, 5, 32\}$.

Costruiamo un modello probabilistico. Uno spazio di probabilità possibile è $\Omega = \{0, 1, 2, ..., 36\}$ con la sigma-algebra delle parti $\wp(\Omega)$. Pensando che il gioco sia condotto in modo onesto, è naturale considerare su $(\Omega, \wp(\Omega))$ la distribuzione di probabilità uniforme.
Con questa scelta, indicando con $B = \{1, 5, 32\}$ risulta

$$P(B) = \frac{\#B}{\#\Omega} = \frac{3}{37}.$$

Supponiamo ora di venire a conoscenza che la roulette è truccata in modo tale che escano solo numeri dispari. In tal modo, posto $A = \{1, 3, 5, 7, ..., 35\}$, nasce il problema di

ASSEGNARE SULLO STESSO SPAZIO $(\Omega, \wp(\Omega))$ UNA NUOVA DISTRIBUZIONE DI PROBABILITÀ Q CHE ASSEGNI ALL'EVENTO A MISURA MASSIMA 1.

Domanda: questa nuova misura Q che probabilità dovrebbe assegnare all' evento B?

Ragionevomente il rapporto tra la cardinalità dei numeri dispari che stanno in B diviso la cardinalità di A, ovvero il rapporto tra la cardinalità dell'evento $A \cap B$ diviso la cardinalità di A:

$$Q(B) = \frac{\#A \cap B}{\#A} = \frac{2}{18}.$$

moltiplichiamo e dividiamo ambo i membri della predecente uguaglianza per $\#\Omega$;

$$Q(B) = \frac{\#B \cap A}{\#A} \times \frac{\#\Omega}{\#\Omega} = \frac{\frac{\#B \cap A}{\#\Omega}}{\frac{\#A}{\#\Omega}} = \frac{P(B \cap A)}{P(A)}.$$

Abbiamo quindi introdotto una nuova misura di probabilità basata sulla certezza che A si verifichi; tale misura assegna probabilità 1 all' evento A e assegna all' evento B una misura $Q(B)$ che si esprime tramite la distribuzione di probabilità iniziale P. Questa formula, introdotta in un caso molto particolare ha valenza generale:

Definizione 1.2. *Sia* $(\Omega, \mathcal{A}, P)$ *uno spazio di probabilità e siano* $A, B \in \mathcal{A}$ *tali che* $P(A) > 0$. *Si dice probabilità di* B *condizionata ad* A *la quantità* $P(B|A)$ *definita come*

$$P(B|A) = \frac{P(B \cap A)}{P(A)}. \tag{1}$$

Esercizio 1.1. *Dimostrare che l' applicazione* $P(\cdot|A) : \mathcal{A} \to [0, 1]$ *tale che* $B \to P(B|A)$ *è una distribuzione di probabilità su* $(\Omega, \mathcal{A})$.

Definizione 1.3. *Sia* $(\Omega, \mathcal{A}, P)$ *uno spazio di probabilità. Si definisce partizione di* Ω *una sequenza finita di eventi* $\{A_1, A_2, ..., A_n\}$ *a due a due disgiunti ed esaustivi, ovvero tali che*

1 $A_u \cap A_j = \phi$ *se* $i \neq j$;

2 $\cup_{i=1}^n A_i = \Omega$.

Come immediata conseguenza degli assiomi si ha che se $\{A_1, A_2, ..., A_n\}$ è una partizione di Ω allora

$$\sum_{i=1}^n P(A_i) = 1.$$

Vale inoltre il

Teorema delle Probabilità Totali

Teorema 1.4. *Su uno spazio di probabilità* $(\Omega, \mathcal{A}, P)$ *si consideri una qualunque partizione* $\{A_1, A_2, ..., A_n\}$ *tale che* $P(A_i) > 0, \quad i = 1, ..., n$. *Allora, per ogni* $B \in \mathcal{A}$ *si ha*

$$P(B) = \sum_{i=1}^n P(B|A_i)P(A_i).$$

Proof. Basta decomporre B come unione finita di eventi disgiunti secondo la rappresentazione

$$B = B \cap \Omega = B \cap \cup_{i=1}^n A_i = \cup_{i=1}^n (B \cap A_i).$$

Quindi dagli assiomi della misura di probabilità si ottiene

$$P(B) = \sum_{i=1}^n P(B \cap A_i).$$

Si conclude la dimostrazione utilizzando la formula inversa della (2). $\square$

La probabilità condizionata ha una duplice applicazione; da una parte permette di aggiornare la probabilità di alcuni fenomeni alla luce dell' osservazione di alcuni eventi collegati. Pensate ad esempio di lanciare una moneta 20 volte e vedere sempre come risultato Testa. Se vi chiedessero di puntare su Testa o Croce al 21-simo lancio che fareste? Avreste il dubbio che forse i due soli esiti possibili non avvengano con la stessa probabilità? La formula (2) permette di ricalcolare le probabilità dell' evento "esce Testa al 21-simo lancio" avendo osservato l' evento $A =$"esce Testa nei primi venti lanci".
Inoltre in alcuni modelli le probabilità di eventi di interesse si riescono a calcolare, grazie al teorema delle probabilità totali, per mezzo di probabilità condizionate talvolta più semplici da calcolare.

Esempio 1.5. *Si vuole studiare l' incidenza di una certa malattia respiratoria su una popolazione specifica. È noto che il 40% degli individui di questa popolazione fuma. Si sa inoltre che*

- *Il 25% dei fumatori ha contratto la malattia respiratoria;*
- *il 7% dei non fumatori ha contratto la malattia respiratoria.*

Si vuole calcolare la probabilità che un individuo scelto a caso nella popolazione sia malato.

Gli eventi di interesse sono:

$$\begin{aligned} M &= \{\textit{L' individuo selezionato è malato }\} \\ M^c &= \{\textit{L' individuo selezionato non è malato }\} \\ F &= \{\textit{L' individuo selezionato fuma }\} \\ F^c &= \{\textit{L' individuo selezionato non fuma }\} \end{aligned}$$

Dobbiamo calcolare $P(M)$. I dati del problema permettono di assegnare le seguenti misure:

$$\begin{aligned} P(F) =& \frac{2}{5};\; P(F^c) = \frac{3}{5} \\ P(M|F) =& \frac{25}{100};\; P(M|F^c) = \frac{7}{100} \end{aligned}$$

*Osserviamo che, qualunque sia lo spazio di probabilità prescelto, un **evento ed il suo complementare formano sempre una partizione**. Per applicare il teorema delle Probabilità totali **abbiamo bisogno di una***

***partizione di cui siano noti le probabilità degli eventi che la compongono**, nel nostro caso $\{F, F^c\}$. Allora applicando il teorema delle Probabilità totali si ottiene*

$$P(M) = P(M|F)P(F) + P(M|F^c)P(F^c) = \frac{25}{100} \times \frac{2}{5} + \frac{7}{100} \times \frac{3}{5} = \frac{71}{500}.$$

Esiste una formula che fornisce una sorta di inversione del Teorema delle Probabilità totali nota come

Formula di Bayes

Su uno spazio di probabilità $(\Omega, \mathcal{A}, P)$ si consideri una partizione $\{A_1, A_2, ..., A_n\}$ tale che $P(A_i) > 0, \quad i = 1, ..., n$ ed un evento $B \in \mathcal{A}$ tale che $P(B) > 0$. Allora, per ogni $j = 1, ..., n$

$$P(A_j|B) = \frac{P(B|A_j)P(A_j)}{P(B)} = \frac{P(B|A_j)P(A_j)}{\sum_{i=1}^n P(B|A_i)P(A_i)}.$$

La dimostrazione segue dalla definizione di probabilità condizionata e dal Teorema delle Probabilità totali:

$$P(A_j|B) = \frac{P(A_j \cap B)}{P(B)} = \frac{P(B|A_j)P(A_j)}{\sum_{i=1}^n P(B|A_i)P(A_i)}.$$

Torniamo all'esempio precedente. Qual è la probabilità che una persona malata sia un fumatore? Formalmente, dobbiamo calcolare $P(F|M)$. Utilizzando la formula di Bayes otteniamo:

$$P(F|M) = \frac{P(M|F)P(F)}{P(M)} = \frac{\frac{25}{100} \times \frac{2}{5}}{\frac{71}{500}} = \frac{50}{71}.$$

Esercizi per casa: 1.33 e 1.42 del libro di Abundo.

Esempio 1.6. *(Quando l' intuito non aiuta...)*
Tre mobili indistinguibili contengono ciascuno due cassetti. Nel primo mobile c'è una moneta d' oro in ciascun cassetto, nel secondo una d' oro in un cassetto ed una d' argento nell'altro e nel terzo mobile c'è una moneta d' argento in ciascun cassetto. Apro un cassetto e trovo una moneta d'oro. Calcolare la probabilità che anche l' altro cassetto del mobile scelto contenga una moneta d' oro.

Cominciamo con l' osservare che, qualunque Spazio di Probabilità si consideri, la terna di eventi $\{A_1, A_2, A_3\}$, con $A_i =$ " scelgo l' i-simo cassetto ", $i =$

$1, 2, 3$ ne è una partizione.
Poniamo $B =$ " il cassetto contiene una moneta d'oro ". Il problema è calcolare $P(A_1|B)$.
Utilizzando la formula di Bayes:

$$P(A_1|B) = \frac{P(B|A_1)P(A_1)}{\sum_{i=1}^{3} P(B|A_i)P(A_i)}$$

dove $P(A_i) = \frac{1}{3}, \quad i = 1, 2, 3,\ P(B|A_1) = 1,\ P(B|A_2) = \frac{1}{2},\ P(B|A_3) = 0.$
Sostituendo si ottiene quindi $P(A_1|B) = \frac{1\times\frac{1}{3}}{1\times\frac{1}{3}+\frac{1}{2}\times\frac{1}{3}+0\times\frac{1}{3}} = \frac{2}{3}$.

Esempio 1.7. *Consideriamo un'urna contenente b palline bianche ed r palline rosse. Estraiamo senza reimmissione due palline dall' urna. Calcolare la probabilità dell' evento $B =$ "la seconda pallina estratta è bianca".*

Costruiamo lo spazio campionario Ω come l'insieme dei possibili esiti di questo esperimento. Vedremo che, con l' aiuto della probabilità condizionale, non avremo bisogno di definire la misura di Probabilità sull' insieme delle parti di Ω.

$$\Omega = \{\omega = (\omega_1, \omega_2) : \omega_i \in \{bianca, rossa\}, i = 1, 2\}.$$

$$B = \{\omega = (\omega_1, \omega_2) : \omega_2 = bianca\}$$

Osserviamo che non possiamo considerare una misura di probabilità uniforme. Infatti, poiché in generale $b \neq r$ non tutti gli eventi elementari sono equiprobabili. Però, se consideriamo gli eventi

$$R_1 = \{\omega = (\omega_1, \omega_2) : \omega_1 = rossa\}, R_1^c = \{\omega = (\omega_1, \omega_2) : \omega_1 = bianca\}$$

allora $\{R_1, R_1^c\}$ sono una partizione di Ω tale che

$$P(R_1) = \frac{r}{b+r} \text{ e } P(R_1^c) = \frac{b}{b+r}.$$

Inoltre

$$P(B|R_1) = \frac{b}{b+r-1} \text{ e } P(B|R_1^c) = \frac{b-1}{b+r-1}.$$

Quindi, applicando il teorema delle Probabilità totali

$$\begin{aligned} P(B) =& P(B|R_1)P(R_1) + P(B|R_1^c)P(R_1^c) = \\ =& \frac{b}{b+r-1} \times \frac{r}{b+r} + \frac{b-1}{b+r-1}\frac{b}{b+r} = \frac{b}{b+r}. \end{aligned}$$

È interessante risolvere lo stesso esercizio con reimmissione della prima pallina estratta. Il risultato è uguale. Cosa potete dedurne?

2 Indipendenza

Definizione 2.1. *Sia dato uno spazio di probabilità* $(\Omega, \mathcal{A}, P)$*. Due eventi* A, B *si dicono indipendenti se*

$$P(A \cap B) = P(A)P(B). \tag{2}$$

Cosa ci si aspetta da due eventi indipendenti? Che il verificarsi dell' uno non fornisce informazioni aggiuntive sul verificarsi dell' altro, e questo è appunto quello che accade. Infatti, supponiamo che $P(A) > 0$. Allora se A e B sono indipendenti accade che

$$P(B|A) = P(B).$$

La verifica segue immediatamente dall' assunzione di indipendenza e dalla definizione di probabilità condizionata. Infatti

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

La proprietà è simmetrica; infatti, se $P(B) > 0$ risulta anche $P(A|B) = P(A)$.

Esempio 2.2. *Consideriamo il lancio di due dadi. Dimostriamo che gli eventi* A_1 = *"esce 6 sul primo dado" e* A_2 = *"esce 6 sul secondo dado" sono indipendenti.*

Già sappiamo che questo gioco può essere modellizzato con una distribuzione di probabilità uniforme sullo spazio

$$\Omega = \{\omega = (\omega_1, \omega_2) : \omega_i \in \{1, 2, 3, 4, 5, 6\}, i = 1, 2\}, \#\Omega = 36$$

Inoltre

$$\begin{aligned} A_1 =& \{\omega = (\omega_1, \omega_2) : \omega_1 = 6, \},\ \#A_1 = 6, \\ A_2 =& \{\omega = (\omega_1, \omega_2) : \omega_2 = 6\},\ \#A_2 = 6, \\ \#A_1 \cap& A_2 = 1. \end{aligned}$$

Quindi

$$P(A_1) = \frac{\#A_1}{\#\Omega} = \frac{1}{6}, \quad P(A_2) = \frac{\#A_2}{\#\Omega} = \frac{1}{6}, \quad P(A_1 \cap A_2) = \frac{\#A_1 \cap A_2}{\#\Omega} = \frac{1}{36}$$

e di conseguenza

$$P(A_1|A_2) = \frac{P(A_1 \cap A_2)}{P(A_2)} = \frac{\frac{1}{36}}{\frac{1}{6}} = \frac{1}{6} = P(A_1),$$

ovvero $P(A_1 \cap A_2) = P(A_1)P(A_2)$*.*

Esercizio 2.1. *Si consideri il lancio di un dado. Dimostrare che gli eventi $A_1 = \{2, 3, 5\}$ e $A_2 = \{4, 5\}$ sono indipendenti.*

Osserviamo che **due eventi disgiunti A e B non sono mai indipendenti, a meno che almeno uno dei due abbia probabilità 0 (evento quasi nullo) o 1 (evento quasi certo)**.

Proprietà 2.1. ***1*** *Gli eventi quasi nulli e quasi certi sono indipendenti da ciascun altro evento;*
Infatti, se $A \in \mathcal{A}$ è tale che $P(A) = 0$, allora scelto un qualunque evento $B \in \mathcal{A}$ si ha

$$0 \leq P(A \cap B) \leq P(A) = 0, \text{ ovvero } 0 = P(A \cap B) = P(A)P(B) = 0.$$

Se invece $P(A) = 1$, allora dalla rappresentazione $B = (A \cap B) \cup (A^c \cap B)$ si ottiene

$$P(A \cap B) = P(B) - P(A^c \cap B) = P(B) \text{ dal momento che } P(A^c) = 1 - P(A) = 0.$$

2 *Se (A, B) è una coppia di eventi indipendenti, allora anche (A^c, B), (A, B^c), (A^c, B^c) sono coppie di eventi indipendenti.*
Verifichiamolo per la coppia (A^c, B).

$$\begin{aligned} P(A^c \cap B) =& P(B) - P(A \cap B) = P(B) - P(A)P(B) = \\ =& P(B)(1 - P(A)) = P(B)P(A^c). \end{aligned}$$

Definizione 2.3. *Sia dato uno spazio di probabilità $(\Omega, \mathcal{A}, P)$. Gli eventi $A_1, A_2, ..., A_n$ si dicono indipendenti se e solo se per ogni $k \leq n$ e per ogni $\{i_1, ..., i_k\} \in \{1, ..., n\}$ accade che*

$$P(A_{i_1} \cap A_{i_2} \cap .... \cap A_{i_k}) = P(A_{i_1})P(A_{i_1}) \cdots P(A_{i_k}). \tag{3}$$

Esempio 2.4. *Consideriamo il lancio di due dadi. Dimostriamo che gli eventi E = "la somma delle facce è 7", F = "esce 4 sul primo dado" e G = "esce 3 sul secondo dado" non sono indipendenti.*

Notiamo che $E = \{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}$ e quindi $P(E) = \frac{1}{6}$, Inoltre $P(F) = P(G) = \frac{1}{6}$. Verifichiamo che la 2.3, non vale per tutti i possibili sottogruppi di $\{E, F, G\}$.

$$P(E \cap F) = P((4, 3)) = \frac{1}{36} = P(E)P(F)$$

$$P(E \cap G) = P((4,3)) = \frac{1}{36} = P(E)P(G)$$
$$P(F \cap G) = P((4,3)) = \frac{1}{36} = P(F)P(G)$$

Ma

$$P(E \cap F \cap G) = P((4,3)) = \frac{1}{36} \neq P(E)P(F)P(G) = \frac{1}{216}$$

Esercizio 2.2. *Se* (A, B, C) *sono indipendenti, allora le seguenti coppie di eventi sono indipendenti:* $(A, B \cap C)$, $(A, B \cup C)$, $(A, B^c \cap C^c)$, $(A, B^c \cup C^c)$.