Probabilità e Statistica

Isadora Antoniano-Villalobos
isadora.antoniano@unive.it

Laurea in Informatica
(Data science/ Tecnologie e scienze dell'informazione)
Università Ca' Foscari di Venezia

Anno academico 2024/2025

Probabilità elementare: Testo Capitolo 2

Fenomeni aleatori

- La logica del **certo** è la logica della teoria degli insiemi e del calcolo su proposizioni (o eventi) che possono assumere il valore di vero o falso in modo **deterministico**.
- Il **calcolo delle probabilità** è invece la logica dell'**incerto**. Si usa per ragionare sui possibili risultati di un **fenomeno aleatorio** o **casuale**, del quale cioè non si può prevedere con certezza l'esito.
- Si può pensare al termine **aleatorio** come l'opposto di **deterministico**.
- **Esempi di fenomeni aleatori:**
  1. Il lancio di un dado.
  2. Il lancio di una stessa moneta 4 volte.
  3. La classificazione di 10 pezzi prodotti da una macchina in conformi o non conformi alle specifiche di progetto.
  4. L'estrazione di una mano di poker, cioè un insieme di cinque carte, da un mazzo di 52.

Fenomeni aleatori

5. L'osservazione del tempo di guasto [min] di un circuito elettrico formato da tre resistenze in serie.
6. La registrazione giornaliera dei livelli massimi di polveri [mcg/mc] nell'aria alla centralina del Parco di San Giuliano nel Gennaio 2015.

   ⋮

- **Esempi dell'incertezza in informatica:**
  1. Il tempo o lo spazio su un disco richiesti per l'installazione di un software
  2. Il numero di difetti di un nuovo software
  3. La memoria richiesta per processare un programma
  4. Il tempo richiesto per una stampa o il numero di lavori in coda di stampa prima di questo
  5. Il momento nel quale un virus attacca un sistema o il numero di file e directory infetti

     ⋮

Spazio campionario, risultati, eventi

Lo pazio campionario e gli eventi

Lo **spazio campionario** è l'insieme di tutti i possibili risultati di un esperimento o fenomeno aleatorio. In questo corso sarà rappresentato dalla lettera greca Ω. Un **evento** è un sottoinsieme $E \subset \Omega$.

- Un generico (singolo) risultato è un elemento dello spazio campionario e si può indicare con $\omega \in \Omega$.
- Una volta che il fenomeno aleatorio di interesse è stato osservato si può dire se un qualsiasi evento sia E vero o falso.
- Quando un evento è vero, si dice che si è **realizzato** o **verificato**
- I possibili risultati $\{\omega\}$, visti come singoletti, cioè insiemi contenenti un solo elemento, sono anch'essi eventi, detti **eventi elementari**.
- Ω viene anche chiamato l'**evento certo**, perché sicuramente si verificherà.

Esempi

1. Lancio di un dado: $\Omega = \{1, 2, 3, 4, 5, 6\}$.
   Un possibile evento: Il dado dà un punteggio superiore a quattro, ovvero E $= \{5, 6\}$
2. Lancio di una stessa moneta quattro volte: $\Omega =$ le sedici possibili sequenze di quattro dei simboli T e C, dove T indica "testa" e C indica "croce".
   Un possibile evento: Otteniamo almeno tre teste sui quattro lanci, ovvero
   $$\mathsf{E} = \{TTTC, TTCT, TCTT, CTTT, TTTT\}$$
3. La classificazione di 10 pezzi con 2 possibili risultati C e N, dove C indica 'conforme' e N indica 'non conforme': $\Omega =$ le 2^{10} possibili sequenze di dieci dei simboli C e N.
   Un possibile evento: Tutti i pezzi sono conformi, ovvero
   $$\mathsf{A} = \{CCCCCCCCCC\} \text{ (questo è anche un singoletto)}$$

Esempi

4. Una mano di poker: $\Omega =$ i $\binom{52}{5}$ possibili sottoinsiemi delle 52 carte. Un possibile evento: Si ottiene un poker, ovvero E dato da tutte le possibili mani contenenti un poker, che sono in numero di 13x48 perché 13 sono i possibili poker e 48 sono, per ogni dato poker, i modi di scegliere la quinta carta.
5. Il tempo di guasto di un circuito elettrico: $\Omega = \mathcal{R}^+ := [0, \infty)$, cioè tutti i numeri non negativi, visto che il tempo di guasto è un numero non negativo.
   Un possibile evento: Il circuito ha una durata di meno di 50 ore, ovvero E $= [0, 50)$

Esempi

6. I livelli massimi giornalieri di polveri nel Gennaio 2015: $\Omega =$ tutte le possibili sequenze di 31 numeri non negativi (la maggior parte contenuti tra 10 e 350).
   Un possibile evento: In nessun giorno si è superato il limite di 300 [mcg/mc], ovvero

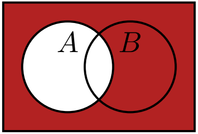
$$\mathsf{A} = \{(x_1, \ldots, x_{31}) : 0 \leq x_i \leq 300, i = 1, \ldots, 31\}$$

☞ **Nota:** Lo spazio campionario può essere **discreto**, se i suoi elementi sono finiti o al meno numerabili, o **continuo**, cuando contiene una quantità non numerabile di elementi.

Operazioni sugli eventi e Diagrammi di Venn

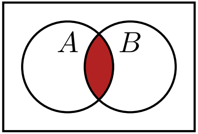
Nota: Unioni e intersezioni e complementi di eventi sono, a loro volta, eventi.

- La negazione o **complementare** di un evento A, indicato con $\bar{A}$, è l'evento che è vero quando A è falso ed è falso quando A è vero.



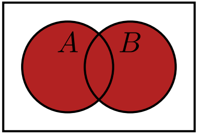
La negazione dell'evento certo è l'**evento impossibile**: $\bar{\Omega} = \emptyset$ (evento impossibile = insieme vuoto).

- L'**intersezione** di due eventi A e B, indicata con $A \cap B$, è l'evento che è vero quando sia A che B sono veri.

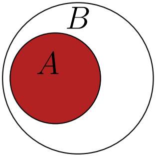


Operazioni sugli eventi e Diagrammi di Venn

- L'**unione** di due eventi A e B, indicata con $A \cup B$, è l'evento che è vero quando o A oppure B oppure entrambi sono veri.

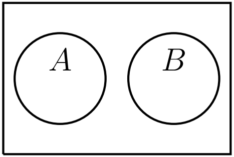


- L'evento A è **incluso** nell'evento B, in simboli $A \subset B$, se il verificarsi di A implica il verificarsi di B.



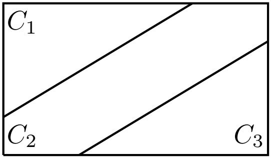
Partizioni

- Due eventi A e B si dicono **incompatibili**, o **disgiunti**, se non è possibile che siano entrambi veri, cioè se $A \cap B = \emptyset$.



- Una famiglia di eventi si dice una **partizione** dello spazio campionario se ogni coppia di insiemi della famiglia ha intersezione vuota e l'unione di tutti i componenti della famiglia è Ω stesso.

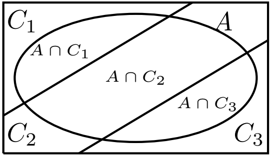
☞ Una partizione può essere **finita**, ad esempio, $\{C_1, C_2, C_3\}$ è una partizione di 3 elementi se $C_1 \cap C_2 = C_1 \cap C_3 = C_2 \cap C_3 = \emptyset$ e $C_1 \cup C_2 \cup C_3 = \Omega$.



Partizioni

- Un qualsiasi evento A si può scrivere come unione delle sue intersezioni con gli elementi di una partizione:

$$A = (A \cap C_1) \cup (A \cap C_2) \cup (A \cap C_3)$$



- In generale, si può pensare a una partizione **numerabile** $C_1, C_2, \ldots$:

$$C_i \cap C_j = \emptyset \quad \forall i, j; \qquad \bigcup_{i=1}^{\infty} C_i = \Omega$$

Scrivendo un qualsiasi evento A come unione numerabile delle sue intersezioni con gli elementi di una partizione:

$$A = \bigcup_{i=1}^{\infty} (A \cap C_i)$$

Esempio

Fenomeno aleatorio: lancio di un dado.
Eventi: Ne consideriamo i seguenti:

$A = \{5, 6\}$ ➔ il risultato del lancio è superiore a 4

$B = \{2, 4, 6\}$ ➔ il risultato del lancio è pari

Allora:

$A \cap B = \{6\}$ ➔ il risultato del lancio è pari e superiore a 4

$A \cup B = \{2, 4, 5, 6\}$ ➔ il risultato del lancio è pari oppure superiore a 4

Partizione: numeri divisibili per 3 e non:

$$C_1 = \{3, 6\}; \quad C_2 = \{1, 2, 4, 5\}$$

Allora:

$$A = (A \cap C_1) \cup (A \cap C_2) = \{6\} \cup \{5\}$$

$$B = (B \cap C_1) \cup (B \cap C_2) = \{6\} \cup \{2, 4\}$$

Definizione assiomatica di probabilità

Formalmente, la **probabilità** è una funzione che assegna ad ogni evento di uno spazio campionario un valore in $\mathbb{R}^+$, ossia un numero non negativo, e deve sodisfare i seguenti assiomi:

1. **Non negatività:** $0 \leq \mathbb{P}\left[E\right] \leq 1$
2. **Normalizzazione:** $\mathbb{P}\left[\Omega\right] = 1$
3. **Aditività:** Se $E_1, E_2, \ldots$ è una sequenza di eventi incompatibili, cioè se $E_i \cap E_j = \emptyset \quad \forall i \neq j$, allora

$$\mathbb{P}\left[\bigcup_{n=1}^{\infty} E_n\right] = \sum_{n=1}^{\infty} \mathbb{P}\left[E_n\right]$$

Definizione assiomatica di probabilità

Interpretazione della probabilità

La probabilità dell'evento A, $\mathbb{P}[A]$, è un numero tra 0 e 1 che indica il grado di fiducia del individuo nell'avverarsi dell'evento A. Più $\mathbb{P}[A]$ è vicina a 1, più ci aspettiamo che l'evento si avveri (minore la nostra incertezza sul avverarsi del evento).

- Una volta osservato il fenomeno aleatorio, sappiamo se A si è verificato o meno, e la sua probabilità non serve più (diventa 1 se l'evento si è verificato e 0 in caso contrario)
- Si può pensare alla probabilità come a una massa unitaria (in virtù del assioma (ii)) da spargere sullo spazio campionario.Se un evento si può scomporre in più pezzi (eventi disgiunti) la sua massa sarà uguale alla somma delle singole masse sui pezzi (singoli eventi).

Alcune proprietà della probabilità

- Probabilità del **complementare**: dato un evento A,

$$\mathbb{P}\left[\bar{A}\right] = 1 - \mathbb{P}\left[A\right].$$

- Probabilità dell'**evento impossibile**:

$$\mathbb{P}\left[\emptyset\right] = \mathbb{P}\left[\bar{\Omega}\right] = 1 - \mathbb{P}\left[\Omega\right] = 0.$$

- Probabilità dell'**unione**: dati due eventi A e B,

$$\mathbb{P}\left[A \cup B\right] = \mathbb{P}\left[A\right] + \mathbb{P}\left[B\right] - \mathbb{P}\left[A \cap B\right].$$

  ☞ Una conseguenza di questa proprietà:

$$\mathbb{P}\left[A \cup B\right] \leq \mathbb{P}\left[A\right] + \mathbb{P}\left[B\right]$$

Alcune proprietà della probabilità

- Probabilità di una **partizione**: se $C_1, C_2, \ldots$ sono una partizione, allora
$$\mathbb{P}\left[\bigcup_{i=1}^{\infty} C_i\right] = \mathbb{P}[\Omega] = 1.$$

Legge della probabilità totale (versione facile)

Se $C_1, C_2, \ldots$ sono una partizione dell'evento certo, la probabilità di un qualsiasi evento A può essere scritta come

$$\mathbb{P}[A] = \sum_i \mathbb{P}[A \cap C_i]$$

Spazi campionari finiti

Se lo spazio campionario costituisce un insieme finito, $\Omega = \{\omega_1, \ldots, \omega_n\}$, allora un'assegnazione di probabilità è data da n valori $p_1, \ldots, p_n$ tali che:

1. $p_i \in [0, 1]$, $\forall i = 1, \ldots, n$;
2. $\sum_{i=1}^{n} p_i = 1$;
3. $p_i = \mathbb{P}\left[\{\omega_i\}\right]$, $\forall i = 1, \ldots, n$.

Dato che ogni evento $A \subset \Omega$ si può scrivere come unione (finita) degli eventi elementari (disgiunti) che lo costituiscono,

$$A = \{\omega_{i_1}, \ldots, \omega_{i_r}\} = \bigcup_{k=1}^{r} \{\omega_{i_k}\},$$

si ha che

$$\mathbb{P}\left[A\right] = \sum_{k=1}^{r} \mathbb{P}\left[\{\omega_{i_k}\}\right] = \sum_{k=1}^{r} p_{i_k}.$$

Eventi elementari equiprobabili

In particolare, se possiamo supporre (per ragioni di simmetria) che tutti gli eventi elementari abbiano la stessa probabilità, allora

$$p_i = \mathbb{P}\left[\{\omega_i\}\right] = \frac{1}{n}, \quad \forall i = 1, \ldots, n.$$

Per ogni evento $A = \{\omega_{i_1}, \ldots, \omega_{i_r}\}$ si può dunque scrivere

$$\mathbb{P}\left[A\right] = \frac{r}{n} = \frac{\#A}{\#\Omega} = \frac{\text{numero casi favorevoli}}{\text{numero casi possibili}}.$$

Attenzione! Questa formula vale solo se gli eventi elementari sono **equiprobabili**.

Eventi elementari equiprobabili: Esempi

1. **Dado:** Qual è la probabilità che il risultato del lancio di un dado equilibrato sia un numero divisibile per 3?

   - Dato che il dado non è truccato, si può assumere che ognuno dei 6 possibili risultati abbia la stessa probabilità pari a $1/6$.
   - I casi favorevoli al nostro evento sono 2, $\{3\}$ e $\{6\}$, mentre quelli possibili sono 6. Il risultato è dunque $2/6 = 1/3$.

☞ Se il dado fosse truccato questo procedimento di calcolo non sarebbe corretto.

Eventi elementari equiprobabili: Esempi

2. **Urna:** Si consideri un'urna composta da quattro palle bianche numerate da 1 a 4 e tre palle nere numerate da 1 a 3. Si campioni casualmente una palla dall'urna.
   - È ragionevole assumere che ciascuna palla abbia probabilità 1/7 di essere estratta.
   - ➔ Consideriamo i seguenti eventi:

$$\begin{aligned} B &= \text{"viene estratta una palla bianca"};\\ N &= \text{"viene estratta una palla nera"}\\ C_i &= \text{"viene estratto il numero } i\text{"}, i = 1, 2, 3, 4;\\ D &= \text{"viene estratto un numero dispari"}. \end{aligned}$$

☞ **Nota:** $N = \bar{B}$

Eventi elementari equiprobabili: Esempi

- Appplicando gli assiomi e altre proprietà del calcolo della probabilità possiamo calcolare le probabilità di tutti questi eventi e di altri ancora:

$$
\begin{aligned}
\mathbb{P}[B] &= 4/7\\
\mathbb{P}[N] &= 1-\mathbb{P}[B] = 3/7\\
\mathbb{P}[C_i] &= 2/7, \quad i=1,2,3\\
\mathbb{P}[C_4] &= 1/7,\\
\mathbb{P}[D] &= 4/7\\
\mathbb{P}[B\cap D] &= 2/7\\
\mathbb{P}[B\cup D] &= \mathbb{P}[B]+\mathbb{P}[D]-\mathbb{P}[B\cap D] = 6/7\\
\mathbb{P}[B\cap C_4] &= \mathbb{P}[C_4] = 1/7 \;\; (C_4\subset B)\\
\mathbb{P}[B\cap C_2] &= 1/7 \;\; (C_2\not\subset B)
\end{aligned}
$$

Probabilità condizionata e indipendenza

Probabilità condizionata

Probabilità condizionata

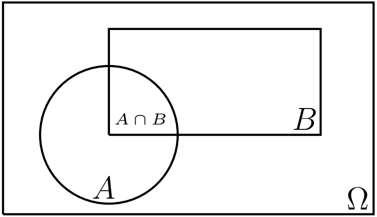
Sia B un evento di probabilità positiva. La **probabilità condizionata** dell'evento A dato l'evento B è

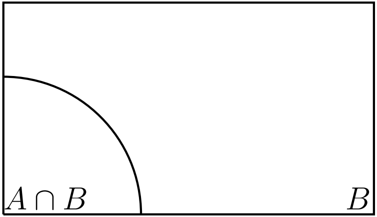
$$\mathbb{P}[A|B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}, \quad \mathbb{P}[B] > 0.$$

- $\mathbb{P}[A|B]$ è anche chiamata la **probabilità subordinata** o **condizionale** di A subordinatamente a B.
  ☞ Da notare l'uso della sbarra verticale $|$ che collega l'evento condizionato A e l'evento condizionante B.
- $\mathbb{P}[A|B]$ rappresenta la probabilità di A valutata in presenza dell'informazione aggiuntiva che B si verifichi.

Probabilità condizionata

- Intuitivamente, si restringe il campo delle possibilità non alla totalità dei possibili risultati Ω ma ad un suo sottoinsieme proprio $B \subset \Omega$.



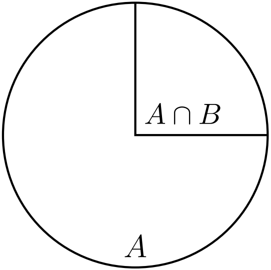


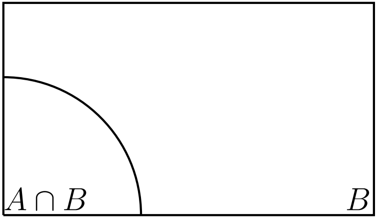
Prima del esperimento, si sa soltanto che B potrebbe avverarsi: $0 \leq \mathbb{P}[B] \leq 1$.
Per qualsiasi evento $A \subset \Omega$,
$0 \leq \mathbb{P}[A \cap B] \leq \mathbb{P}[B]$.

Dopo l'esperimento, si sa con certezza che B è vero:
$\mathbb{P}[B|B] = \frac{\mathbb{P}[B]}{\mathbb{P}[B]} = 1$.
Per qualsiasi evento $A \subset \Omega$,
$0 \leq \mathbb{P}[A|B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]} \leq 1$.

Probabilità condizionata

Attenzione! $\mathbb{P}[A|B]$ e $\mathbb{P}[B|A]$ significano cose molto diverse:





$\mathbb{P}[B|A] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[A]}$ si riferisce all'incertezza residua su B dopo il resultato del esperimento che ha reso A vero.

$\mathbb{P}[A|B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}$ si riferisce all'incertezza residua su A dopo il resultato del esperimento che ha reso B vero.

Probabilità condizionata

- $\mathbb{P}\left[A|B\right]$ e $\mathbb{P}\left[\bar{A}|B\right]$ sono in relazione diretta:

$$\mathbb{P}\left[\bar{A}|B\right] = \frac{\mathbb{P}\left[\bar{A}\cap B\right]}{\mathbb{P}\left[B\right]} = \frac{\mathbb{P}\left[B\right] - \mathbb{P}\left[A\cap B\right]}{\mathbb{P}\left[B\right]} = 1 - \mathbb{P}\left[A|B\right],$$

Le probabilità **condizionate allo stesso evento** obbediscono alle leggi (assiomi) della probabilità!

- $\mathbb{P}\left[A|B\right]$ e $\mathbb{P}\left[A|\bar{B}\right]$ non sono in relazione diretta: $\mathbb{P}\left[A|B\right]$ potrebbe essere uguale a $1 - \mathbb{P}\left[A|\bar{B}\right]$ in qualche esempio, ma in generale non è così.

Probabilità condizionata

- **Esempio:** Si ricordi l'esempio dell'urna composta da quattro palle bianche numerate da 1 a 4 e tre palle nere numerate da 1 a 3, dalla quale si campiona casualmente una palla, con eventi:

$$
\begin{aligned}
B &= \text{"viene estratta una palla bianca"};\\
N &= \text{"viene estratta una palla nera"}\\
C_i &= \text{"viene estratto il numero } i\text{"}, i = 1, 2, 3, 4.
\end{aligned}
$$

Si valutino le probabilità condizionate che la palla estratta abbia il numero 1 dato che è bianca, che abbia il numero 1 dato che è nera e che la palla estratta sia nera dato che ha il numero 1.

Probabilità condizionata

Formalmente:

$$\mathbb{P}[C_1|B] = \frac{\mathbb{P}[C_1 \cap B]}{\mathbb{P}[B]} = \frac{1/7}{4/7} = 1/4$$

$$\mathbb{P}[C_1|N] = \frac{\mathbb{P}[C_1 \cap N]}{\mathbb{P}[N]} = \frac{1/7}{3/7} = 1/3$$

$$\mathbb{P}[N|C_1] = \frac{\mathbb{P}[N \cap C_1]}{\mathbb{P}[C_1]} = \frac{1/7}{2/7} = 1/2.$$

- $\mathbb{P}[N|C_1] \neq \mathbb{P}[C_1|N]$.
- Invece, per esempio, $\mathbb{P}[N|C_1] = 1 - \mathbb{P}[B|C_1] = 1/2$, $(N = \bar{B})$.

La formula delle probabilità composte

- La definizione di probabilità condizionata si può anche usare come formula pratica per la fattorizzazione della probabilità di un'intersezione, sempre che $\mathbb{P}[A|B]$ sia ben definita:

$$\mathbb{P}[A \cap B] = \mathbb{P}[A|B]\,\mathbb{P}[B],$$

- Questa formula si generalizza ad un qualsiasi numero di eventi $A_1, \ldots, A_n$ e viene anche chiamata la **formula delle probabilità composte**.

Formula delle probabilità composte

$$\begin{aligned}\mathbb{P}[A_1 \cap A_2 \cap \cdots \cap A_n] =& \mathbb{P}[A_1]\,\mathbb{P}[A_2|A_1]\,\mathbb{P}[A_3|A_1 \cap A_2] \ldots \\ & \ldots \mathbb{P}[A_n|A_1 \cap \cdots \cap A_{n-1}].\end{aligned}$$

La formula delle probabilità composte

➔ **Esempio:** Si consideri ora l'esperimento che consiste nell'estrazione di 3 palline senza reinserimento dalla solita urna. Qual è la probabilità che le prime due siano bianche e la terza nera?

Risulta utile considerare i seguenti eventi:

B_i = "pallina bianca all'i-esima estrazione"
N_i = "pallina nera all'i-esima estrazione"

Applicando la forula delle probabilità composte:

$$\begin{aligned}\mathbb{P}\left[B_1 \cap B_2 \cap N_3\right] &= \mathbb{P}\left[B_1\right]\mathbb{P}\left[B_2 | B_1\right]\mathbb{P}\left[N_3 | B_1 \cap B_2\right] \\ &= \frac{4}{7}\frac{3}{6}\frac{3}{5} = \frac{6}{35}.\end{aligned}$$

☞ Qual è la probabilità che siano tutte tre nere?

Eventi indipendenti

2 eventi indipendenti

Si dice che A e B sono **indipendenti** nel caso particolare in cui

$$\mathbb{P}[A|B] = \mathbb{P}[A].$$

In questo caso si scrive $A \perp B$, e vale anche $\mathbb{P}\left[A|\bar{B}\right] = \mathbb{P}[A]$

- Se A e B sono indipendenti, si ha allora

$$\mathbb{P}[A \cap B] = \mathbb{P}[A]\,\mathbb{P}[B]$$

che può anche essere presa come **definizione di eventi indipendenti**.

Eventi indipendenti

La definizione si può estendere a più di 2 eventi

n eventi reciprocamente indipendenti

Gli eventi $A_1, \ldots, A_n$, si dicono reciprocamente **indipendenti** se comunque si prendono $k > 1$ di essi, si ha

$$\mathbb{P}\left[A_{i_1} \cap \cdots \cap A_{i_k}\right] = \mathbb{P}\left[A_{i_1}\right] \ldots \mathbb{P}\left[A_{i_k}\right].$$

➔ **Esempio:** Tornando alla urna, se le estrazioni dell'esempio al lucido precedente si effettuano con reinserimento, allora i tre eventi B_1, B_2 e N_3 sono indipendenti e si ha:

$$\mathbb{P}\left[B_1 \cap B_2 \cap N_3\right] = \mathbb{P}\left[B_1\right] \mathbb{P}\left[B_2\right] \mathbb{P}\left[N_3\right] = \frac{4}{7}\frac{4}{7}\frac{3}{7} = \frac{48}{147}.$$

Eventi indipendenti

Esiste anche una nozione più debole (meno restrittiva) di indipendenda per più di due eventi:

n eventi indipendenti due a due

Gli eventi $A_1, \ldots, A_n$, si dicono **indipendenti due a due** se per ogni coppia è indipendente, ovvero

$$\mathbb{P}[A_i \cap A_j] = \mathbb{P}[A_i] \ldots \mathbb{P}[A_j] \quad \forall i \neq j.$$

- Se $A_1, \ldots, A_n$, sono reciprocamente indipendenti, allora sono indipendenti due a due ma il contrario non è necessariamente vero.
- Le definizioni si possono estendere ulteriormente a una collezione infinita di eventi $A_1, \ldots, A_n, \ldots$, che saranno indipendenti se ogni sottoinsieme finito di essi è formato da eventi indipendenti.

Eventi indipendenti

Nota: Eventi indipendenti e eventi disgiunti sono cose molto diverse. Due eventi sono disgiunti o meno a prescindere dalle loro probabilità. Due eventi disgiunti possono essere indipendenti soltanto se uno di essi ha probabilità zero.

➔ **Esempio:** Si consideri l'esperimento di lanciare un dado equo due volte. Si definiscano i seguenti eventi:

$A =$ "la somma dei dadi è 6"
$B =$ "la somma dei dadi è 7"
$C =$ "il primo dado dà 4"

Si ha

$$\mathbb{P}[A] = \frac{5}{36}, \quad \mathbb{P}[B] = \frac{1}{6}, \quad \mathbb{P}[C] = \frac{1}{6}.$$

Eventi indipendenti

- Poiché
$$\mathbb{P}\left[A \cap C\right] = \mathbb{P}\left[(4,2)\right] = \frac{1}{36}\ ,$$
allora A e C non sono indipendenti.
- B e C sono invece indipendenti, ma non disgiunti. Infatti,
$$\mathbb{P}\left[B \cap C\right] = \frac{1}{36} = \mathbb{P}\left[B\right]\mathbb{P}\left[C\right].$$
- Infine, A e B sono disgiunti ma non indipendenti:
$$\mathbb{P}\left[A \cap B\right] = \mathbb{P}\left[\emptyset\right] = 0.$$

Esempio: Sistemi in serie

Un sistema formato da n componenti separati si dice **in serie** se funziona quando tutti gli n componenti funzionano.

$$\boxed{in} \longrightarrow \boxed{C_1} \longrightarrow \boxed{C_2} \longrightarrow \boxed{\cdots} \longrightarrow \boxed{C_n} \longrightarrow \boxed{out}$$

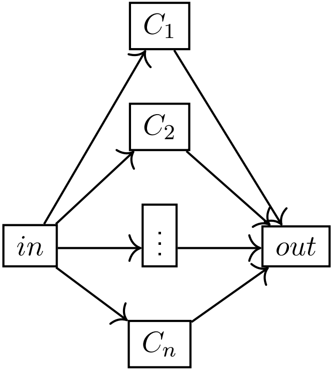
- Si supponga che i componenti del sistema si guastino in modo **indipendente** e che la probabilità di guasto del componente i-esimo sia p_i.
- Sia inoltre A_i l'evento in cui il componente i-esimo funziona e A l'evento in cui l'intero sistema funziona (quindi $A = \cap_{i=1}^n A_i$)

☞ Allora, usando l'ipotesi di indipendenza dei componenti:

$$\mathbb{P}[A] = \mathbb{P}\left[\bigcap_{i=1}^n A_i\right] = \prod_{i=1}^n \mathbb{P}[A_i] = \prod_{i=1}^n (1 - p_i),$$

Esempio: Sistemi in parallelo

Un sistema formato da n componenti separati si dice **in parallelo** se funziona quando almeno uno degli n componenti funziona.



Esempio: Sistemi in parallelo

- Si supponga che i componenti del sistema si guastino in modo **indipendente** e che la probabilità di guasto del componente i-esimo sia p_i.
- Sia inoltre A_i l'evento in cui il componente i-esimo funziona e B l'evento in cui l'intero sistema funziona (quindi $B = \cup_{i=1}^{n} A_i$)

☞ Allora, usando l'ipotesi di indipendenza dei componenti:

$$\mathbb{P}[B] = 1 - \mathbb{P}\left[\bar{B}\right] = 1 - \mathbb{P}\left[\bigcap_{i=1}^{n} \bar{A}_i\right] = 1 - \prod_{i=1}^{n} \mathbb{P}\left[\bar{A}_i\right] = 1 - \prod_{i=1}^{n} p_i,$$

Indipendenza condizionata

Finalmente, due (o più) eventi possono essere indipendenti condizionatamente a un terzo:

2 eventi condizionatamente indipendenti

Si dice che A e B sono **condizionalmente indipendenti** dato l'evento C se

$$\mathbb{P}\left[A \cap B | C\right] = \mathbb{P}\left[A | C\right] \mathbb{P}\left[A | C\right].$$

In questo caso si scrive $A \perp B | C$

- Du eventi indipendenti potrebbero non essere condizionatamente indipendenti, e viceversa.

La legge della probabilità totale

Legge della probabilità totale

Se $C_1, C_2, \ldots$ sono una partizione dell'evento certo, la probabilità di un qualsiasi evento A può essere scritta come

$$\mathbb{P}[A] = \sum_i \mathbb{P}[A \cap C_i] = \sum_i \mathbb{P}[C_i]\,\mathbb{P}[A|C_i]$$

che si dice **legge della probabilità totale**.

- La prima uguaglianza viene dal fatto che

$$A = \bigcup_i (A \cap C_i),$$

  che sono eventi a due a due disgiunti.
- La seconda uguaglianza segue dalla definizione di probabilità condizionata, sempre che $\mathbb{P}[C_i] > 0$, $\forall i = 1, 2, \ldots$.

La legge della probabilità totale

- La legge delle probabilità totali si può applicare anche alla probabilità condizionata:

$$\mathbb{P}\left[A|B\right] = \sum_i \mathbb{P}\left[A \cap B \cap C_i\right] = \sum_i \mathbb{P}\left[C_i|B\right] \mathbb{P}\left[A|B \cap C_i\right]$$

Il teorema di Bayes

Formula o teorema di Bayes

Sia data la partizione $C_1, C_2, \ldots$ e tutti i suoi elementi abbiano probabilità positiva. Sia A un ulteriore evento, anch'esso con probabilità positiva. Fissiamo l'attenzione su uno specifico elemento C_m della partizione. Abbiamo allora

$$\mathbb{P}\left[C_m | A\right] == \frac{\mathbb{P}\left[A | C_m\right] \mathbb{P}\left[C_m\right]}{\sum_i \mathbb{P}\left[A | C_i\right] \mathbb{P}\left[C_i\right]}$$

- **Thomas Bayes (1702–1761):** matematico e ministro presbiteriano britannico. Deve la sua fama ai suoi studi nel campo della matematica e della filosofia; è noto soprattutto nella statistica per il teorema di Bayes, vertente sulla probabilità condizionata, pubblicato postumo nel 1763 (Wikipedia)

Thomas Bayes?

Il teorema di Bayes

La formula si può derivare da:

1. La definizione di probabilità condizionata:

$$\mathbb{P}\left[C_m|A\right] = \frac{\mathbb{P}\left[C_m \cap A\right]}{\mathbb{P}\left[A\right]}$$

2. La formula delle probabilità composte (per il numeratore):

$$\mathbb{P}\left[C_m \cap A\right] = \mathbb{P}\left[C_m\right]\mathbb{P}\left[A|C_m\right]$$

3. La legge della probabilità totale (per il denominatore):

$$\mathbb{P}\left[A\right] = \sum_i \mathbb{P}\left[A|C_i\right]\mathbb{P}\left[C_i\right]$$

Interpretazione del teorema di Bayes

- Permette di aggiornare l'assegnazione di probabilità data **a priori** a certi eventi C_m, alla luce di nuova informazione (A si è verificato).
- Il risultato di questa operazione di aggiornamento sono le nuove probabilità $\mathbb{P}[C_m|A]$, dette anche **a posteriori**.
- Le probabilità a posteriori costituiscono una sintesi dell'informazione disponibile a priori su un certo fenomeno e dell'informazione empirica.

➔ **Esempio:** Consideriamo un filtro automatico che blocchi i messaggi *spam* in arrivo alla casella di posta elettronica.

- Il metodo per selezionare i messaggi spam si basa sulla presenza di alcune parole chiave all'interno del messaggio stesso (ad esempio `viagra`, `vi@gra`, `v1@gr@`, `V.iagra`, ecc.).
- Consideriamo gli eventi:
  $S =$ "il messaggio è spam"
  $W =$ "il messaggio contiene determinate parole chiave"

Interpretazione del teorema di Bayes

Per il teorema di Bayes possiamo fissare a priori la probabilità che un messaggio sia spam e poi aggiornarla nel caso in cui il messaggio contenga le parole chiave considerate:

$$\mathbb{P}[S|W] = \frac{\mathbb{P}[W|S]\,\mathbb{P}[S]}{\mathbb{P}[W|S]\,\mathbb{P}[S] + \mathbb{P}\left[W|\bar{S}\right]\mathbb{P}\left[\bar{S}\right]},$$

- In mancanza di ulteriore informazione, possiamo fissare $\mathbb{P}[S] = 1/2$, cioè a priori possiamo immaginare di avere la stessa probabilità di ricevere un messaggio spam o uno buono;
- $\mathbb{P}[W|S]$ e $\mathbb{P}\left[W|\bar{S}\right]$ possiamo **stimarle** utilizzando l'esperienza, cioè i dati. Ad esempio, immaginiamo di aver analizzato un campione di 2000 messaggi spam e di aver trovato le parole chiave in 250 di questi. Mentre su 1000 messaggi non spam solo 5 contenevano le parole chiave. Allora, possiamo fissare $\mathbb{P}[W|S] \simeq 250/2000$ e $\mathbb{P}\left[W|\bar{S}\right] \simeq 5/1000$.

Interpretazione del teorema di Bayes

- Allora possiamo calcolare

$$\mathbb{P}[S|W] = \frac{0.125 \times 0.5}{0.125 \times 0.5 + 0.005 \times 0.5} = 0.962.$$

- Il filtro automatico confronta questa probabilità con una soglia prefissata. Se $\mathbb{P}[S|W]$ supera la soglia allora il messaggio viene classificato come spam e dunque bloccato. Se invece $\mathbb{P}[S|W]$ sta sotto la soglia, allora il messaggio viene consegnato nella casella di posta.
- Il valore della soglia dipende da quanto sono infastidito dai messaggi spam (soglia bassa) e quanto invece mi preoccupa perdere un messaggio importante (soglia alta).
- Se nell'esempio fissiamo una soglia pari a 0.9, allora il messaggio considerato viene bloccato come spam: $\mathbb{P}[S|W] = 0.962 > 0.9$.

Esempio: Test diagnostici

- La frazione dei soggetti affetti da una certa malattia (per esempio, la sieropositività HIV oppure la tubercolosi) in una popolazione si chiama **prevalenza**.
- Si consideri un test diagnostico per la malattia. La **sensitività** di un test è la probabilità che il test, somministrato a un malato, sia positivo.
- La **specificità** di un test è la probabilità che il test, somministrato a un non malato, sia negativo.

☞ **Situazione ideale:** sensitività $=$ specificità $= 1$. Purtroppo, la situazione ideale non è raggiungibile, e i test reali sono imperfetti, cioè con sensitività < 1 e specificità < 1. Questo da origine a **Falsi positivi e falsi negativi**

Esempio: Test diagnostici

- Si immagini di somministrare un test diagnostico non perfetto a una persona estratta a caso dalla popolazione e si considerino gli eventi:

$$
\begin{aligned}
M &= \text{la persona estratta è malata} \\
+ &= \text{il test dà risultato positivo} \\
- &= \text{il test dà risultato negativo} \\
\bar{M} \cap + &= \text{il test dà un falso positivo} \\
M \cap - &= \text{il test dà un falso negativo}
\end{aligned}
$$

- Si ha allora

$$
\mathbb{P}\left[M\right] = \text{prevalenza} \quad \mathbb{P}\left[+|M\right] = \text{sensitività} \quad \mathbb{P}\left[-|\bar{M}\right] = \text{specificità}
$$

Esempio: Test diagnostici

- Probabilità di un falso positivo:

$$\begin{aligned}\mathbb{P}\left[\bar{M} \cap +\right] &= \mathbb{P}\left[\bar{M}\right]\mathbb{P}\left[+|\bar{M}\right] \\ &= (1 - \text{prevalenza}) \times (1 - \text{specificità}).\end{aligned}$$

- Probabilità di un falso negativo:

$$\begin{aligned}\mathbb{P}\left[M \cap -\right] &= \mathbb{P}\left[M\right]\mathbb{P}\left[-|M\right] \\ &= \text{prevalenza} \times (1 - \text{sensitività}).\end{aligned}$$

- Si studi un test per l'HIV:

$$\begin{aligned}\text{prevalenza} &= \mathbb{P}\left[HIV\right] = 0.001 \\ \text{sensitività} &= \mathbb{P}\left[+|HIV\right] = .95 \\ \text{specificità} &= \mathbb{P}\left[-|\overline{HIV}\right] = .98\end{aligned}$$

Esempio: Test diagnostici

- Probabilità di falso positivo:

$$\begin{aligned}\mathbb{P}\left[\overline{HIV}\cap +\right] &= \mathbb{P}\left[\overline{HIV}\right]\mathbb{P}\left[+|\overline{HIV}\right]\\ &= (1-0.001)(1-0.98) = 0.01998.\end{aligned}$$

- Probabilità di falso negativo:

$$\begin{aligned}\mathbb{P}\left[HIV\cap -\right] &= \mathbb{P}\left[HIV\right]\mathbb{P}\left[-|HIV\right]\\ &= 0.001(1-0.95) = 0.00005.\end{aligned}$$

- Probabilità che il test somministrato ad una persona campionata a caso dalla popolazione sia positivo:

$$\begin{aligned}\mathbb{P}\left[+\right] &= \mathbb{P}\left[+|HIV\right]\mathbb{P}\left[HIV\right] + \mathbb{P}\left[+|\overline{HIV}\right]\mathbb{P}\left[\overline{HIV}\right]\\ &= 0.95\times 0.001 + (1-0.98)\times(1-0.001)\\ &= 0.02093\end{aligned}$$

cioè in pratica avremo, a lungo andare, il 2 per cento di positivi, siano essi veri positivi o falsi positivi.

Esempio: Test diagnostici

- Probabilità che una persona risultata positiva sia effettivamente malata:

$$\mathbb{P}[HIV|+] = \frac{\mathbb{P}[HIV]\,\mathbb{P}[+|HIV]}{\mathbb{P}[HIV]\,\mathbb{P}[+|HIV] + \mathbb{P}\left[\overline{HIV}\right]\mathbb{P}\left[+|\overline{HIV}\right]}$$
$$= \frac{0.001 \cdot 0.95}{0.001 \cdot 0.95 + (1-0.001)(1-0.98)} \approx 0.045$$

- Nel campo della diagnostica, $\mathbb{P}[M|+]$ viene chiamata **valore predittivo positivo**, mentre $\mathbb{P}\left[\overline{M}|-\right]$ è il **valore predittivo negativo**.