Azzolini Riccardo 2020-02-18

Spazi di probabilità

1 Tipi di fenomeni

Si può distinguere tra:

fenomeni deterministici: si ha una legge che associa uno stato al sistema in ogni sua condizione (un esempio è il lancio di una pallina, la cui posizione può essere calcolata istante per istante);

fenomeni aleatori: non si ha una descrizione deterministica, o perché non è possibile, oppure perché vi si rinuncia (ad esempio, l'estrazione di una pallina numerata da un'urna).

Nell'ambito del calcolo delle probabilità, si considerano i fenomeni aleatori.

In assenza di una descrizione deterministica, si può comunque avere una capacità predittiva: la probabilità fornisce appunto un metodo predittivo per i fenomeni aleatori.

2 Problema d'esempio

Individuare le differenze nel modo di introdurre un modello predittivo nei due casi:

1. Un'urna contiene 6 palline, numerate da 1 a 6, ma per il resto identiche. Una pallina viene estratta a caso, e se ne guarda il numero. In che modo si possono scrivere i possibili esiti?
2. Considerare l'istante in cui un componente elettronico si guasta. In che modo si possono scrivere i possibili esiti?

3 Modellazione del caso 1

I possibili esiti sono del tipo

- "esce il numero 2",
- "esce il numero 6",

ma anche

- "esce un numero dispari",
- "esce un numero ≤ 3",
- "non esce il 4",
- "esce un numero pari ≤ 4".

Per modellare il problema, si introduce l'insieme Ω di tutti gli esiti singoli:

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

Gli eventi possibili sono allora i sottoinsiemi $A \subseteq \Omega$ di qualunque tipo (compresi $\varnothing$, cioè l'"evento impossibile" che non si verifica mai, e Ω stesso, ovvero l'"evento certo", che si verifica in ogni caso). In altre parole, $A \in \mathscr{P}(\Omega)$: gli eventi sono gli elementi dell'*insieme delle parti* $\mathscr{P}(\Omega)$ di Ω.

Sugli eventi, descritti in questo modo, è possibile applicare le operazioni di unione, intersezione e complemento. Intuitivamente:

- $A \cap B$ significa "si verificano entrambi gli eventi A e B";
- $A \cup B$ significa "almeno uno dei due eventi si verifica";
- A^{c} significa "l'evento A non si verifica".

4 Modellazione del caso 2

In questo caso, si può introdurre l'insieme dei possibili esiti, che sono tutti gli infiniti istanti in cui il componente si può guastare, ma questi non corrisponderanno a nessun evento, in quando ciascuno di essi ha probabilità di verificarsi nulla. Infatti, un esempio di esito è "il componente si guasta *esattamente* a $t = 1$ s" (inteso come $1.00000\dots$), e la probabilità che ciò avvenga è effettivamente 0. Per avere probabilità non nulla, ogni evento dovrà quindi essere un insieme (ad esempio, un intervallo) di istanti.

L'insieme degli esiti è $\Omega = \mathbb{R}_0^+$.[1] Da questo, si possono provare vari approcci per definire l'insieme degli eventi. Nel farlo, si vuole ottenere la proprietà che unioni, intersezioni, e complementi di eventi siano ancora eventi (per permettere, ad esempio, di ragionare sulla probabilità che due eventi si verifichino entrambi, oppure che un dato evento non si verifichi). Alcuni dei possibili approcci sono:

[1] Il fatto che lo 0 sia compreso non è rilevante per la soluzione del problema, perché, comunque, la probabilità di un guasto esattamente all'istante 0 è, appunto, nulla.

- Si potrebbero considerare gli intervalli di $\mathbb{R}$ del tipo $[a, b]$ (intervalli chiusi), ma si osserva che:
$$\bigcup_{n=1}^{\infty} \left[0, 1 - \frac{1}{n}\right] = [0, 1)$$
Allora, non sarebbe consentito fare unioni infinite (numerabili) di intervalli, poiché eventualmente si otterrebbe un intervallo non chiuso, cioè non del tipo $[a, b]$, il quale non apparterrebbe più all'insieme degli intervalli (eventi) considerati.
- Un'altra possibilità sarebbe considerare l'insieme delle parti, $\mathscr{P}(\mathbb{R}_0^+)$, ma esso è scomodo da trattare. In particolare, alcuni dei suoi elementi (cioè alcuni sottoinsiemi di $\mathbb{R}_0^+$) non sono misurabili, cioè non è possibile assegnarvi un valore di probabilità. Un esempio è l'insieme dei numeri irrazionali compresi nell'intervallo $[0, 1]$,
$$\tilde{I} = \mathbb{R}_0^+ \setminus \mathbb{Q}_0^+ \cap [0, 1]$$
che è formato da infiniti punti, ma, rispetto all'intero intervallo $[0, 1]$, ha anche infiniti punti "mancanti".
- Ci sono molte soluzioni valide. Tipicamente, si considerano gli intervalli aperti di $\mathbb{R}$ (o, in questo caso, $\mathbb{R}_0^+$), cioè quelli del tipo (a, b), e si costruisce, a partire da essi, il più piccolo insieme $\mathscr{A}$ i cui elementi (sottoinsiemi di $\mathbb{R}$) siano eventi, ossia siano tali che si possano applicare su di essi le operazioni di unione, intersezione e complemento, ottenendo sempre insiemi ancora appartenenti ad $\mathscr{A}$.

  Allora, è nota in letteratura la costruzione $(\Omega, \mathscr{A})$, dove $\mathscr{A}$ è, appunto, un'opportuna classe di insiemi con operazioni (in questo caso, gli intervalli aperti di $\Omega = \mathbb{R}_0^+$, con gli insiemi che si ricavano applicando le operazioni di unione, intersezione e complemento).

5 Modello comune

È possibile definire un modello comune ai due casi, che permette di affrontare entrambi allo stesso modo.

5.1 σ-algebra

Definizione: Una famiglia $\mathscr{A} \subseteq \mathscr{P}(\Omega)$ di parti di Ω, cioè di sottoinsiemi dell'insieme Ω, si dice **$\boldsymbol{\sigma}$-algebra** ("sigma-algebra") se e solo se:

1. $\varnothing, \Omega \in \mathscr{A}$
2. se $A \in \mathscr{A}$, allora anche $A^\mathsf{c} \in \mathscr{A}$
3. se $A_1, \dots, A_n, \dots \in \mathscr{A}$, allora anche:

a) $\bigcup_{n=1}^{\infty} A_n \in \mathscr{A}$

b) $\bigcap_{n=1}^{\infty} A_n \in \mathscr{A}$

Osservazioni:

- La proprietà 3a è il motivo per cui bisogna scartare la soluzione basata sugli intervalli chiusi $[a, b]$.
- La proprietà 3b è in realtà ridondante, in quanto può essere dimostrata dalle altre, applicando la legge di De Morgan. Tale legge afferma che[2]

$$A^{\mathsf{c}} \cap B^{\mathsf{c}} = (A \cup B)^{\mathsf{c}}$$

e vale anche per intersezioni/unioni infinite:

$$\bigcap_{n=1}^{\infty} A_n^{\mathsf{c}} = \left(\bigcup_{n=1}^{\infty} A_n\right)^{\mathsf{c}}$$

Dimostrazione: $\forall A_n \in \mathscr{A}$, per la proprietà 2, $\exists B_n \in \mathscr{A}$ tale che $A_n = B_n^{\mathsf{c}}$. Allora, sostituendo A_n con l'equivalente B_n^{c} nell'intersezione infinita della proprietà 3b, si ottiene

$$\bigcap_{n=1}^{\infty} A_n = \bigcap_{n=1}^{\infty} B_n^{\mathsf{c}}$$

e, applicando la legge di De Morgan, si ricava

$$\bigcap_{n=1}^{\infty} A_n = \bigcap_{n=1}^{\infty} B_n^{\mathsf{c}} = \left(\bigcup_{n=1}^{\infty} B_n\right)^{\mathsf{c}} \in \mathscr{A}$$

che appartiene ad $\mathscr{A}$ perché:

  - $B_n \in \mathscr{A}$, quindi $\bigcup_{n=1}^{\infty} B_n \in \mathscr{A}$ (proprietà 3a);
  - $\bigcup_{n=1}^{\infty} B_n \in \mathscr{A}$, quindi $\left(\bigcup_{n=1}^{\infty} B_n\right)^{\mathsf{c}} \in \mathscr{A}$ (proprietà 2). $\square$

- In alternativa, si può considerare ridondante la proprietà 3a, ricavandola dalle proprietà 2 e 3b, sempre sfruttando la legge di De Morgan.

[2] La legge di De Morgan che si applica agli insiemi è equivalente a quella che vale nell'algebra booleana: $(\textsc{not}\ A)\ \textsc{and}\ (\textsc{not}\ B) = \textsc{not}(A\ \textsc{or}\ B)$.

5.2 Probabilità

Definizione: Sia Ω un insieme, e sia $\mathscr{A} \subseteq \mathscr{P}(\Omega)$ una σ-algebra di parti di Ω. Una (**mappa di**) **probabilità** su $\mathscr{A}$ è un'applicazione $P : \mathscr{A} \to [0, 1]$ tale che:

1. $P(\Omega) = 1$
2. se la successione $\{A_n\}$ di eventi $A_n \in \mathscr{A}$ è tale che $A_i \cap A_j = \varnothing$ per $i \neq j$ (cioè se tali eventi sono a due a due disgiunti), allora

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$$

(ovvero la probabilità che si verifichi uno qualsiasi tra gli eventi disgiunti A_n è uguale alla somma delle probabilità dei singoli eventi)

Nota: Dalle altre proprietà di $\mathscr{A}$ e P, si deduce che $P(\varnothing) = 0$ (quindi, non è necessario imporre separatamente che ciò valga).

5.3 Spazio di probabilità

Definizione: Uno **spazio di probabilità** è una terna $(\Omega, \mathscr{A}, P)$, con le caratteristiche date in precedenza, cioè dove:

- Ω è l'insieme di tutti gli esiti, che prende il nome di **spazio totale** o **spazio campionario**;
- $\mathscr{A}$, una σ-algebra di parti di Ω, è l'insieme di tutti gli eventi;
- P è una probabilità su $\mathscr{A}$.

6 Applicazione del modello al caso 1

Nel caso 1, si hanno:

$$\begin{aligned} \Omega &= \{1, 2, 3, 4, 5, 6\} \\ \mathscr{A} &= \mathscr{P}(\Omega) \end{aligned}$$

Infatti, $\mathscr{P}(\Omega)$ soddisfa le condizioni di σ-algebra (ciò è facilmente verificabile, perché l'insieme è finito).

Rimane da definire P. Assumendo che ogni singolo esito sia equiprobabile, si ottiene dalle regole di P (oppure si postula, verificandone poi la consistenza con le regole) che la probabilità di un evento è data dal rapporto tra il numero di casi favorevoli e il numero di casi possibili:

$$P(A) = \frac{\# A}{\# \Omega}$$

(dove $\#$ indica la cardinalità, cioè il numero di elementi, di un insieme).

Quindi, ad esempio:

- L'evento A_1 = "esce il numero 2" = $\{2\} \in \mathscr{A}$ ha probabilità
$$P(A_1) = \frac{\# A_1}{\# \Omega} = \frac{1}{6}$$
- Per A_2 = "esce un numero dispari" = $\{1, 3, 5\} \in \mathscr{A}$ si ha
$$P(A_2) = \frac{\# A_2}{\# \Omega} = \frac{3}{6} = \frac{1}{2}$$
- La probabilità di A_3 = "non esce il 4" = $\{1, 2, 3, 5, 6\} \in \mathscr{A}$ è
$$P(A_3) = \frac{\# A_3}{\# \Omega} = \frac{5}{6}$$
- Per A_4 = "esce un numero pari ≤ 4" = $\{2, 4\} \in \mathscr{A}$ vale
$$P(A_4) = \frac{\# A_4}{\# \Omega} = \frac{2}{6} = \frac{1}{3}$$

7 Applicazione del modello al caso 2

Anche il caso 2 può essere modellato come uno spazio di probabilità, ponendo $\Omega = \mathbb{R}_0^+$, e considerando la σ-algebra $\mathscr{A}$ generata dagli intervalli aperti di $\mathbb{R}_0^+$.

La probabilità di un evento (a, b) è di fatto uguale a quella di $[a, b]$, dato che i singoli punti a e b hanno probabilità nulla, quindi considerarli o meno non influisce sul risultato. Nel caso in cui $a = 0$, può essere ragionevole adottare il modello

$$P\big([0, b]\big) = 1 - e^{-\lambda b}$$

(dove λ è un parametro positivo da misurare sperimentalmente, che dipende dalle caratteristiche del componente elettronico considerato). Allora, osservando che $[a, b] = [0, b] \setminus [0, a)$, nel caso generale vale

$$\begin{aligned} P\big([a, b]\big) &= P\big([0, b]\big) - P\big([0, a)\big) \\ &= 1 - e^{-\lambda b} - \Big(1 - e^{-\lambda a}\Big) \\ &= e^{-\lambda a} - e^{-\lambda b} \end{aligned}$$

Osservazione: In generale, non esiste un modello (cioè uno spazio di probabilità) "corretto" per un determinato fenomeno aleatorio: la scelta è soggettiva, empirica, e lo stesso problema può essere studiato con modelli diversi. Per questo, si può parlare al massimo di spazi di probabilità "ragionevoli" o "naturali" per un certo fenomeno.