Definizione classica di probabilità

Immaginiamo di avere un esperimento, cioè dal punto di vista della statistica, un qualche cosa che non ha un risultato certo (e.g. lancio di moneta, evento naturale, studio caratteristica popolazione…); presenta esiti diversi in maniera casuale, che non riusciamo a controllare completamente.

L'insieme di tutti gli esiti di un esperimento costituisce lo spazio campione.

L'evento è un sottoinsieme dello spazio campione; contiene un numero di esiti dell'esperimento sottoinsieme di tutti gli esiti possibili.

NOTAZIONE E DEF. PRELIMINARI

ESPERIMENTO (CASUALE) presenta esiti diversi in maniera casuale

ESITI $e_1, e_2, \ldots e_N$ (N = n° esiti dell'esperimento)

SPAZIO CAMPIONE: insieme di tutti gli esiti di un dato esperimento

$\{e_1, e_2, \ldots e_N\}$

con notazione S o Ω

EVENTO: sottoinsieme dello spazio campione

esempio

LANCIO DADO CUBICO ≡ esperimento

$e_1, e_2, e_3, e_4, e_5, e_6$ esiti se e_k = uscita di k puntini

$S = \{e_1, e_2, e_3, e_4, e_5, e_6\}$ SPAZIO CAMPIONE

evento A = 'uscita di un numero multiplo di 3'

$A = \{e_3, e_6\}$

Un evento può essere un qualunque sottoinsieme dello spazio campione. Questo illustrato è un caso particolare. Con questa terminologia si può dare la definizione classica di probabilità.

DEFINIZIONE CLASSICA DI PROBABILITÀ

Dato un esperimento ed un evento E associato all'esperimento, la probabilità di E (P(E)) è un numero tale che

$$P(E) = \frac{n^\circ \text{ esiti contenuti in } E}{\begin{array}{c} n^\circ \text{ esiti possibili} \\ (\text{contenuti nello spazio campione}) \end{array}} = \frac{n^\circ \text{ esiti favorevoli}}{n^\circ \text{ esiti totali}}$$

se il numero di esiti totali è FINITO e gli esiti sono EQUIPROBABILI

Favorevoli nel linguaggio comune è una cosa positiva, in realtà indica il numero di esiti contenuti nell'evento studiato.

Gli esiti possibili sono sempre conosciuti a priori? Quando parliamo di probabilità discreta bisogna immaginare di conoscere gli esiti possibili, e supporre il numero di esiti finito. In caso di numero di esiti totali infiniti, ci sono dei problemi matematici (tutto diviso infinito fa 0). Con le variabili casuali cadrà.

Senza presupporre che il numero di esiti totali sia finito, avremmo un assurdo. Inoltre, immaginando il dado, il dado non deve essere truccato. In tal caso, non si avrebbe il risultato corretto: c'è una faccia che ha molta più probabilità di uscire delle altre. Bisogna quindi presupporre che tutte le facce siano equiprobabili.

Questa definizione non funziona sempre (esiti infiniti, non equiprobabili, quindi situazioni "fisiche"). Nei casi semplici funziona. Dal punto di vista della logica ha un problema, si definisce la probabilità come un numero, usando il concetto di uguale probabilità, matematicamente orribile, ma funziona. Con la definizione assiomatica, questa si potrà recuperare come definizione di un caso particolare.

Essenzialmente la definizione classica di probabilità richiede di saper contare tutti gli esiti (favorevoli e totali).

OSS. È indispensabile saper 'contare' gli esiti per utilizzare questa definizione

$\Downarrow$

RICHIAMI DI CALCOLO COMBINATORIO

Analizziamo vantaggi e svantaggi di questa definizione. C'è un solo vantaggio: è semplicissima, chiunque la può capire.

PRO E CONTRO DELLA DEF. CLASSICA:

PRO

FACILITÀ DAL PUNTO DI VISTA DELLA COMPRENSIONE E DEGLI STRUMENTI MATEMATICI

CONTRO

- NON UTILIZZABILE PER ESITI CHE NON SIANO EQUIPROBABILI
- GLI ESITI DEVONO ESSERE FINITI

ESSERE FINITI

- DEVE ESSERE POSSIBILE "CONTARE" GLI ESITI

Non funziona se gli esiti non sono equiprobabili, infiniti e non è sempre banale contare gli esiti (e.g. nella medicina a nessun esperimento può essere applicata questa definizione, negli eventi fisici nemmeno, situazioni reali interessanti).

Definizione frequentista di probabilità

Un'altra definizione proposta spesso e che è sicuramente conosciuta viene introdotta per considerare i casi più realistici: la definizione frequentista di probabilità. Volendo è una definizione sperimentale.

DEFINIZIONE FREQUENTISTA DI PROBABILITÀ

Dato un esperimento ed un evento E associato all'esperimento, si ripeta l'esperimento in maniera identica e indipendente un numero "elevato" di volte (N), allora

$$P(E) = \frac{n° \text{ di esperimenti in cui si osserva } E}{n° \text{ totale di esperimenti eseguiti}}$$

sempre supponendo che il numero di esiti dello spazio campione sia finito

Con un dado truccato, si lancia tante volte e si fa il rapporto tra numero di uscite con la faccia desiderata ed il numero totale di lanci.

Il vantaggio è evidente: non si deve contare niente, si può fare un esperimento di osservazione dei fenomeni naturali… Alla fine per decidere la probabilità basta un rapporto, semplice.

Gli svantaggi sono che non sempre si hanno a disposizione tanti esperimenti.

Affinché questa definizione abbia senso, deve essere garantito che il numero di esperimenti sia finito.

PRO E CONTRO DELLA DEFINIZIONE FREQUENTISTA

PRO
- "FACILE" DAL PUNTO DI VISTA DEL CALCOLO
- SI PUÒ USARE ANCHE PER ESITI CHE NON SIANO EQUIPROBABILI

CONTRO
- NON SEMPRE È POSSIBILE RIPETERE MOLTE VOLTE L'ESPERIMENTO
- GLI ESITI DEVONO ESSERE IN NUMERO

FINITO

- IL RISULTATO DELLA RIPETIZIONE DEGLI ESPERIMENTI È SPESSO NON MOLTO PRECISO ($N \gg 1$)

In alcuni contesti è difficile essere indipendenti da valori esterni? Chiaramente l'esperimento va ripetuto sempre in maniera identica, quindi è vero (e.g. nei test di un farmaco bisogna ripetere sempre lo stesso iter per i pazienti). Sicuramente questo rende complessa l'applicazione anche di questa definizione, seppur molto utilizzata nel metodo scientifico.

È una definizione imprecisa? Un contro di questa definizione è che, la precedente era un calcolo rigoroso, qui a seconda del numero di prove ho delle situazioni, quindi una probabilità approssimata, non completamente corretta. La legge dei grandi numeri permetterà di capire quando usare la definizione frequentista e l'errore presente quando si usa. Non facendo infinite prove, rimane una definizione imprecisa.

Richiami di calcolo combinatorio

PRINCIPIO FONDAMENTALE DEL CALCOLO COMBINATORIO O PRINCIPIO DI ENUMERAZIONE

Dati un esperimento 1 con N esiti possibili ed un esperimento 2 con M esiti possibili indipendenti dagli esiti di 1, il numero di coppie ordinate (esito esperimento 1, esito esperimento 2) è

$$N \cdot M$$

Immaginiamo di avere 2 esperimenti: nel primo ho N esiti, nel secondo M esiti. Immaginiamo di costruire le coppie ordinate di tutti gli esiti. Quante sono?

GENERALIZZAZIONE A K ($K \geq 2$) ESPERIMENTI

n_i = n° esiti dell'i-esimo esperimento $(i = 1, 2, \dots k)$

$\Downarrow$

n° coppie ordinate $(\text{esito}_{\text{esp. 1}}, \text{esito}_{\text{esp. 2}}, \dots \text{esito}_{\text{esp. } k}) =$

Facciamo che siano indipendenti per semplificare il calcolo (diventa più complicato se 1 influenza 2 o viceversa).

$$= n_1 \cdot n_2 \cdot n_3 \cdots n_k = \prod_{i=1}^{k} n_i$$

<u>esempio</u> n° targhe automobilistiche formate da 2 lettere, 3 cifre, 2 lettere

$$26 \times 26 \times 10 \times 10 \times 10 \times 26 \times 26 = 26^4 \times 10^3$$

Il numero di targhe automobilistiche costruibili come quelle che abbiamo adesso (26 lettere alfabeto, 10 numeri). Con lettere su le prime e le ultime 2 caselle e numeri sulle centrali. Anziché contare tutti gli esiti, si ha un calcolo veloce.