UNITELMA SAPIENZA

SAPIENZA
UNIVERSITÀ DI ROMA
DIPARTIMENTO DI INFORMATICA

Probabilità

Marco Isopi

11. Probabilità totale e teorema di Bayes

Scopo della lezione

Obiettivi della lezione:

1. presentare la formula della probabilità totale che permette di scrivere la probabilità di un evento tenendo conto esplicitamente di “tutti i casi possibili”.

Scopo della lezione

Obiettivi della lezione:

1. presentare la formula della probabilità totale che permette di scrivere la probabilità di un evento tenendo conto esplicitamente di "tutti i casi possibili".
2. affrontare il problema di calcolare la probabilità delle cause di eventi osservati, tramite il teorema di Bayes.

Primi esempi

Consideriamo il seguente esperimento.
Un'urna contiene due palle bianche e tre nere. Una palla viene estratta dall'urna.

Se è bianca si pesca un carta da un mazzo che è stato privato delle carte di cuori.
Se è nera si pesca un carta da un mazzo che è stato privato delle carte di picche.

Vogliamo calcolare la probablità di ottenere una carta rossa.
Vediamo subito che se la pallina pescata è bianca, tale probablità è $\frac{1}{3}$, mentre se la pallina pescata è nera, è $\frac{2}{3}$.

Questa però non è una risposta completa alla nostra domanda.

Primi esempi

Definiamo i seguenti eventi:

$$B = \{\text{viene estratta una palla bianca}\}$$
$$N = \{\text{viene estratta una palla nera}\}$$
$$R = \{\text{viene estratta una carta rossa}\}$$

Allo scopo di calcolare $\mathbf{P}(R)$, scriviamo $R = (R \cap B) \cup (R \cap N)$.

Dato che B e N sono eventi incompatibili, lo sono anche $R \cap B$ e $R \cap N$.

Primi esempi

Quindi

$$\begin{aligned}\mathbf{P}(R) &= \mathbf{P}(R \cap B) + \mathbf{P}(R \cap N)\\ &= \mathbf{P}(R \mid B)\mathbf{P}(B) + \mathbf{P}(R \mid N)\mathbf{P}(N)\\ &= \frac{1}{3}\cdot\frac{2}{5} + \frac{2}{3}\cdot\frac{3}{5} = \frac{8}{15}.\end{aligned}$$

Possiamo ricavare una regola generale, adattando l'esempio appena visto.

Probabilità totale

Facciamo alcune osservazioni:

- è indispensabile che gli eventi presi in considerazione per il primo stadio dell'esperimento coprano tutti gli esiti possibili;

Probabilità totale

Facciamo alcune osservazioni:

- è indispensabile che gli eventi presi in considerazione per il primo stadio dell'esperimento coprano tutti gli esiti possibili;
- serve anche che tali eventi siano mutualmente esclusivi, altrimenti non potremmo sommare le probabilità delle loro intersezioni con R;

Probabilità totale

Facciamo alcune osservazioni:

- è indispensabile che gli eventi presi in considerazione per il primo stadio dell'esperimento coprano tutti gli esiti possibili;
- serve anche che tali eventi siano mutualmente esclusivi, altrimenti non potremmo sommare le probabilità delle loro intersezioni con R;
- non è invece importante che gli esiti del primo stadio siano stati raggruppati solo in due eventi.

Probabilità totale

Prima di dare una formulazione generale dell'idea illustrata nell'esempio precedente ci serve una

Definizione:
Una famiglia di insiemi $A_1, A_2, \ldots A_n$ si dice partizione dell'insieme S se

$$A_i \cap A_j = \emptyset \text{ quando } i \neq j; \qquad \bigcup_{i=1}^{n} A_i = S.$$

Probabilità totale

Siamo ora in grado di enunciare la **formula della probabilità totale**:

Se $A_1, A_2, \dots A_n$ è una partizione di S e B un evento qualunque,

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

Inoltre, se $\mathbf{P}(A_i) > 0$ per ogni i,

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

Formula di Bayes

Consideriamo di nuovo l'esperimento visto in precedenza.

Un'urna contiene due palle bianche e tre nere. Una palla viene estratta dall'urna.

Se è bianca si pesca un carta da un mazzo che è stato privato delle carte di cuori.
Se è nera si pesca un carta da un mazzo che è stato privato delle carte di picche.

Ora ci poniamo una domanda diversa, supponendo di conoscere l'esito finale dell'esperimento.

Formula di Bayes

Supponiamo sia stata estratta una carta rossa.
Con quale probabilità la palla estratta al primo stadio era bianca?

$$\mathbf{P}(B \mid R) = \frac{\mathbf{P}(B \cap R)}{\mathbf{P}(R)} = \frac{\mathbf{P}(R \mid B)\mathbf{P}(B)}{\mathbf{P}(R)}.$$

Le quantità che compaiono nel membro di destra sono tutte note, quindi il problema è risolto:

$$\mathbf{P}(B \mid R) = \mathbf{P}(R \mid B)\frac{\mathbf{P}(B)}{\mathbf{P}(R)} = \frac{1}{3} \cdot \frac{\frac{2}{5}}{\frac{8}{15}} = \frac{1}{4}.$$

Formula di Bayes

Nell'esempio appena visto abbiamo scritto la probabilità di una causa dato l'effetto.

Questa è un'istanza della regola di Bayes.

Per due eventi A e B tali che $\mathbf{P}(A) > 0$ e $\mathbf{P}(B) > 0$, abbiamo:

$$\mathbf{P}(B \mid A) = \mathbf{P}(A \mid B)\frac{\mathbf{P}(B)}{\mathbf{P}(A)}.$$

Esempio 1

In un test a scelta multipla ci sono 5 risposte possibili per ogni domanda.

Per ogni domanda la probabilità di essere convinti di conoscere la risposta è $\frac{3}{4}$, ma con probabilità $\frac{1}{5}$ la nostra convinzione è errata.

Per le domande di cui non conosciamo la risposta ne scegliamo una caso.

Calcoliamo la probabilità di dare la risposta esatta a una fissata domanda.

Esempi

Chiamiamo N l'evento "la so" (o almeno credo). E l'evento "la risposta è corretta".

Sappiamo: $\mathbf{P}(N) = \frac{3}{4}$, $\mathbf{P}(E \mid N) = \frac{4}{5}$, $\mathbf{P}(E \mid N^C) = \frac{1}{5}$.

Dalla formula della probabilità totale:

$$\mathbf{P}(E) = \mathbf{P}(E \mid N)\mathbf{P}(N) + \mathbf{P}(E \mid N^C)\mathbf{P}(N^C) = \frac{4}{5} \cdot \frac{3}{4} + \frac{1}{5} \cdot \frac{1}{4} = \frac{13}{20}.$$

Esempi

Se diamo la risposta esatta, con quale probabilità si tratta di un colpo di fortuna?

La quantità che ci interessa è ora $\mathbf{P}(N^C \mid E)$.

Usando la regola di Bayes

$$\mathbf{P}(N^C \mid E) = \mathbf{P}(E \mid N^C)\frac{\mathbf{P}(N^C)}{\mathbf{P}(E)} = \frac{1}{5}\frac{\frac{1}{4}}{\frac{13}{20}} = \frac{1}{13}.$$

Esempio 2

Di dieci confezioni di lampadine, nove contengono 10 lampadine funzionanti e 2 guaste, mentre la decima ne contiene 2 funzionanti e 10 guaste.

Scegliamo una confezione a caso ne prendiamo due lampadine. Se entrambe sono guaste, con quale probabilità provengono dalla decima confezione?

Chiamiamo *G* l'evento "entrambe le lampadine scelte sono guaste, e *B* l'evento "la confezione è buona".

Iniziamo trovando $\mathbf{P}(G \mid B)$ e $\mathbf{P}(G \mid B^C)$:

$$\mathbf{P}(G \mid B) = \frac{\binom{2}{2}}{\binom{12}{2}} = \frac{1}{66},$$

$$\mathbf{P}(G \mid B^C) = \frac{\binom{10}{2}}{\binom{12}{2}} = \frac{15}{22}.$$

Dalla formula della probabilità totale:

$$\mathbf{P}(G) = \mathbf{P}(G \mid B)\mathbf{P}(B) + \mathbf{P}(G \mid B^C)\mathbf{P}(B^C) = \frac{1}{66} \cdot \frac{9}{10} + \frac{15}{22} \cdot \frac{1}{10} = \frac{9}{110}.$$

Adesso usiamo la formula di Bayes:

$$\mathbf{P}(B^C \mid G) = \mathbf{P}(G \mid B^C)\frac{\mathbf{P}(B^C)}{\mathbf{P}(G)} = \frac{15}{22} \cdot \frac{\frac{1}{10}}{\frac{9}{110}} = \frac{5}{6}.$$

Esempi

Esempio 3

Un’analisi del sangue è efficace al 95% nell’individuare una certa malattia, quando questa è presente. Inoltre dà luogo all’1% di falsi positivi.

Si stima che la patologia sia presente nello 0.5% della popolazione.

Se una persona risulta positiva al test, con quale probabilità è malata?

Chiamiamo M l’evento “il paziente ha la patologia”, P l’evento “il paziente risulta positivo all’analisi”.

Sappiamo che $\mathbf{P}(P \mid M) = \frac{19}{20}$, $\mathbf{P}(P \mid M^C) = \frac{1}{100}$ e $\mathbf{P}(M) = \frac{1}{200}$.

Troviamo

$$\mathbf{P}(P) = \mathbf{P}(P \mid M)\mathbf{P}(M) + \mathbf{P}(P \mid M^C)\mathbf{P}(M^C) = \frac{19}{20} \cdot \frac{1}{200} + \frac{1}{100} \cdot \frac{199}{200} = \frac{294}{20000}.$$

Per trovare la probabilità che un individuo positivo al test sia effettivamente malato usiamo la formula di Bayes:

$$\mathbf{P}(M \mid P) = \mathbf{P}(P \mid M) \cdot \frac{\mathbf{P}(M)}{\mathbf{P}(P)} = \frac{19}{20} \cdot \frac{\frac{1}{200}}{\frac{294}{20000}} = \frac{95}{294} \simeq 0.323.$$

Esempio 4

In un concorso a premi ci sono tre porte chiuse. Dietro una di queste porte c'è un premio. Dopo che ne abbiamo scelta una il presentatore, che sa dove si trova il premio, ne apre una diversa da quella scelta da noi e dietro la quale non si trova il premio.

A questo punto ci viene data la possibilità di cambiare la scelta fatta in partenza. Cosa conviene fare?

Chiamiamo A l'evento "il premio è dietro la prima porta" e definiamo analogamente B e C. Naturalmente $\mathbf{P}(A) = \frac{1}{3}$.

Chiamiamo poi R_C l'evento "“il presentatore apre la terza porta".

Esempi

Se $\mathbf{P}(A \mid R_C) < \frac{1}{2}$, conviene cambiare.

Quindi il problema si riduce al calcolo di $\mathbf{P}(A \mid R_C)$.

Per prima cosa calcoliamo $\mathbf{P}(R_C)$:

$$\begin{aligned}\mathbf{P}(R_C) &= \mathbf{P}(R_C \mid A)\mathbf{P}(A) + \mathbf{P}(R_C \mid B)\mathbf{P}(B) + \mathbf{P}(R_C \mid C)\mathbf{P}(C) \\ &= \frac{1}{2} \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} + 0 \cdot \frac{1}{3} = \frac{1}{2}.\end{aligned}$$

Esempi

Usando anche stavolta la regola di Bayes,

$$\mathbf{P}(A \mid R_C) = \mathbf{P}(R_C \mid A) \cdot \frac{\mathbf{P}(A)}{\mathbf{P}(R_C)} = \frac{1}{2} \cdot \frac{\frac{1}{3}}{\frac{1}{2}} = \frac{1}{3}.$$

Conviene di certo cambiare! La probabilità che il premio si trovi dietro la porta chiusa che non abbiamo scelto è infatti $\frac{2}{3}$.