Il Teorema di Bayes

Paolo Sciattella

Teorema di Bayes: introduzione

Teorema di Bayes: introduzione

Nelle lezioni precedenti sono stati introdotti i concetti di eventi incompatibili e di eventi indipendenti

- A e B sono eventi **incompatibili** se

$$A \cap B = \emptyset$$

- A e B sono eventi **indipendenti** se

$$P(A \cap B) = P(A) \cdot P(B)$$

Teorema di Bayes: introduzione

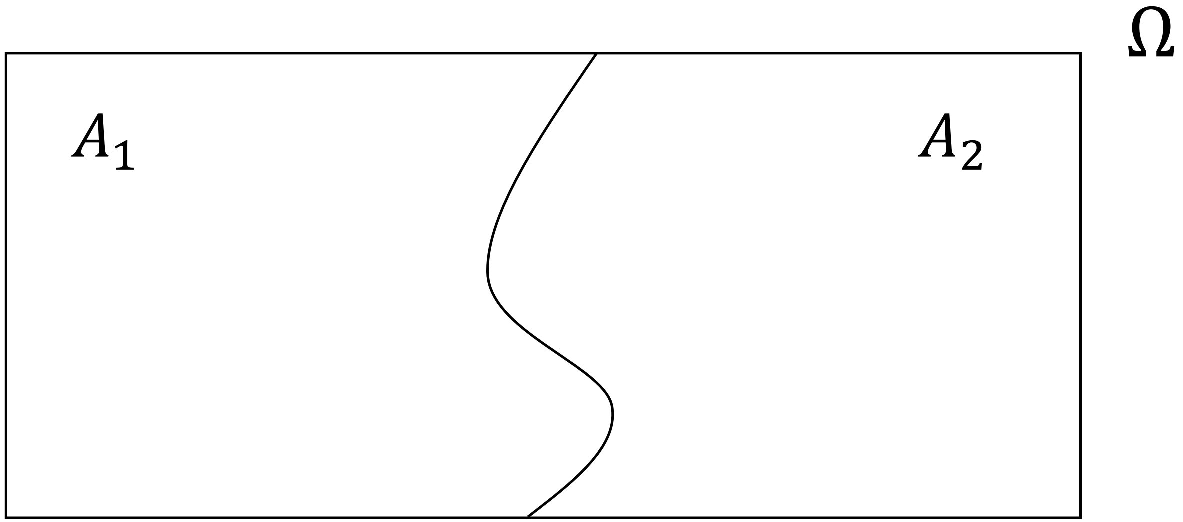
- Il teorema di Bayes rappresenta uno dei fondamenti della teoria della probabilità e della statistica. Lo presentiamo qui considerando prima un caso specifico per poi descriverlo nella sua forma più generale.

Teorema di Bayes: introduzione

Sia $\{A_1, A_2\}$ una partizione dello spazio campionario Ω, tale che:

$$A_1 \cap A_2 = \emptyset$$

$$A_1 \cup A_2 = \Omega$$

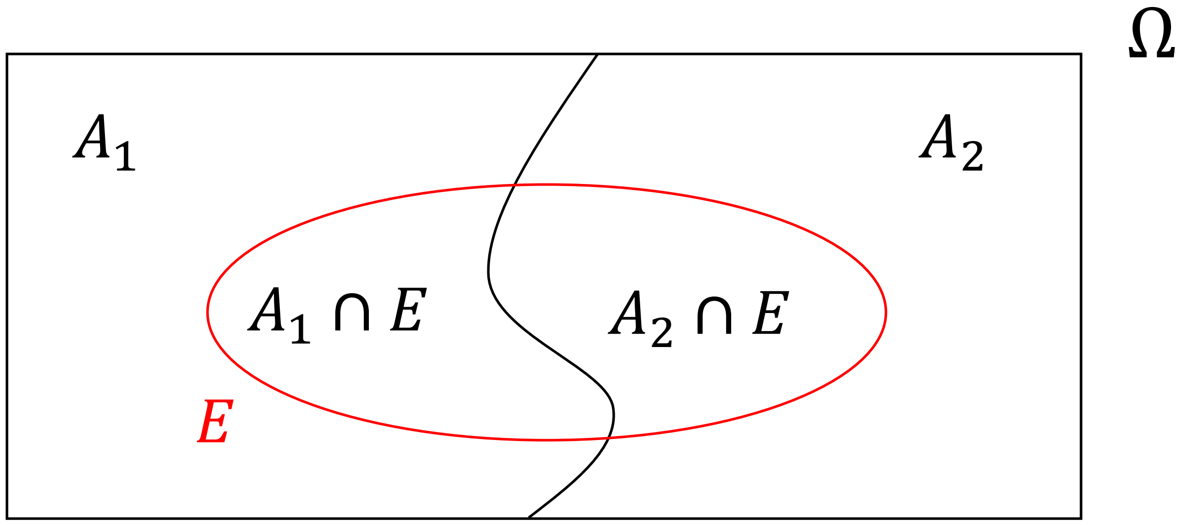


Teorema di Bayes: introduzione

Consideriamo un terzo evento $E \subset \Omega$, con $\mathrm{P}(E) > 0$, di cui si conoscono le probabilità condizionate ad A_1 e A_2, ossia

$P(E|A_1)$ e $P(E|A_2)$.

È chiaro, per le ipotesi fatte, che se si verifica E deve essersi verificato almeno uno degli eventi A_1 e A_2.



Teorema di Bayes: introduzione

Supponendo che si sia verificato l'evento E, ci si chiede: qual è la probabilità che si sia verificato A_1 piuttosto che A_2?

Per rispondere alla domanda si deve calcolare la probabilità:

$$P(A_1|E) = \frac{P(A_1 \cap E)}{P(E)}$$

Sappiamo che $P(A_1 \cap E) = P(E|A_1) \cdot P(A_1)$, quindi:

$$P(A_1|E) = \frac{P(A_1 \cap E)}{P(E)} = \frac{P(E|A_1) \cdot P(A_1)}{P(E)}$$

Teorema di Bayes: introduzione

Sapendo che $\mathrm{E} = (A_1 \cap E) \cup (A_2 \cap E)$ e che A_1 e A_2 sono eventi disgiunti, ossia $A_1 \cap A_2 = \emptyset$

Si può calcolare $\mathrm{P}(E)$ come:

$$P(E) = P(A_1 \cap E) + P(A_2 \cap E) =$$

$$= P(E|A_1) \cdot P(A_1) + P(E|A_2) \cdot P(A_2)$$

Teorema di Bayes: introduzione

Sostituendo in $\mathrm{P}(E)$ si ottiene:

$$P(A_1|E) = \frac{P(A_1 \cap E)}{P(E)} = \frac{P(E|A_1) \cdot P(A_1)}{P(E|A_1) \cdot P(A_1) + P(E|A_2) \cdot P(A_2)}$$

*

Il Teorema di Bayes: generalizzazione

Il Teorema di Bayes: generalizzazione

Consideriamo una partizione dello spazio campionario Ω, cioè un insieme di eventi:

$$A_1, A_2, \dots, A_K$$

Tali che:

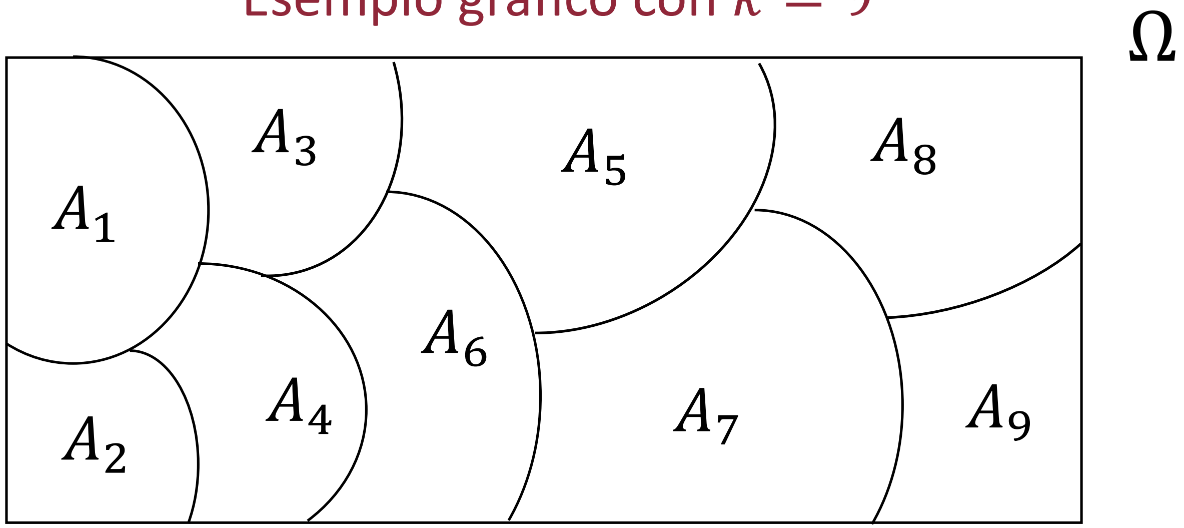
- $A_i \cap A_j = \emptyset \ \ \forall i \neq j$ (eventi **esclusivi)**
- $\bigcup_{i=1}^{K} A_i = \Omega$ ossia la loro unione ricostruisce lo spazio campionario Ω (eventi esaustivi)

Il Teorema di Bayes: generalizzazione

Consideriamo una partizione dello spazio campionario Ω, cioè un insieme di eventi:

$$A_1, A_2, \dots, A_K$$

Esempio grafico con $k = 9$



Il Teorema di Bayes: generalizzazione

Sia poi B un altro evento incluso in Ω.

Le probabilità a posteriori $P(A_i|\, B)$ sono calcolabili utilizzando l'espressione delle probabilità condizionate, ossia:

$$P(A_i|\, B) = \frac{P(A_i \cap B)}{P(B)} \qquad i = 1, 2, \dots, K$$

Il Teorema di Bayes: generalizzazione

$$P(A_i|\,B) = \frac{P(A_i \cap B)}{P(B)} \qquad i = 1, 2, \dots, K$$

sappiamo che

$$P(A_i \cap B) = P(A_i) \cdot P(B|A_i)$$

Inoltre ricordando l'incompatibilità degli eventi Ai si può scrivere:

$$P(B) = P(B \cap \Omega) = P(B \cap \bigcup_{i=1}^{K} A_i)$$

(*continua…*)

Il Teorema di Bayes: generalizzazione

$$P(B) = P(B \cap \Omega) = P(B \cap \bigcup_{i=1}^{K} A_i) =$$

$$= P\big((A_1 \cap B) \cup (A_2 \cap B) \cup \ldots \cup (A_K \cap B)\big) =$$

$$= P(A_1 \cap B) + P(A_2 \cap B) + \ldots + P(A_K \cap B) =$$

$$= P(A_1) \cdot P(B|A_1) + P(A_2) \cdot P(B|A_2) + \ldots + P(A_K) \cdot P(B|A_K)$$

Il Teorema di Bayes: generalizzazione

Pertanto sfruttando le uguaglianze
$P(A_i \cap B) = P(A_i) \cdot P(B|A_i)$

e

$P(B) = P(A_1) \cdot P(B|A_1) + P(A_2) \cdot P(B|A_2) + \ldots + P(A_K) \cdot P(B|A_K)$

si ottiene il risultato:

$$P(A_i|B) = \frac{P(A_i) \cdot P(B|A_i)}{P(A_1) \cdot P(B|A_1) + P(A_2) \cdot P(B|A_2) + \ldots + P(A_K) \cdot P(B|A_K)}$$

Il Teorema di Bayes: generalizzazione

Teorema di Bayes

Dato un insieme **esclusivo** ed **esaustivo** di eventi: $A_1, A_2, \dots, A_K$ e un evento B, si ha: $P(A_i \cap B) = P(A_i) \cdot P(B|A_i)$

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

- Le probabilità $P(A_i)$ sono dette **probabilità a priori**
- Le probabilità $P(B|A_i)$ sono dette **verosimiglianze** degli eventi A_i
- Le probabilità $P(A_i|B)$ sono dette **probabilità a posteriori**

*

Teorema di Bayes: applicazione

Teorema di Bayes: applicazione

- Supponiamo di sapere che in una data popolazione il **10%** degli individui è **affetto** da una determinata **patologia**.
- Per diagnosticare la presenza della patologia si deve effettuare un **test ematico**.
- È noto tuttavia che il test risulta **negativo** anche per il **10% dei malati** (falsi negativi) mentre risulta **positivo** nel **20% dei sani** (falso positivo).
- A questo punto è lecito chiedersi: se un individuo risulta positivo al test, qual è la probabilità che esso sia effettivamente malato?

Teorema di Bayes: applicazione

Elenchiamo gli eventi che stiamo considerando:

- A_1: l'individuo è malato
- A_2: l'individuo è sano
- B_1: il test è negativo
- B_2: il test è positivo

La risposta alla nostra domanda richiede la determinazione della probabilità:

$\boldsymbol{P(A_1|B_2)}$ = ***probabilità che un individuo positivo al test sia effettivamente malato***.

Teorema di Bayes: applicazione

Dai dati disponibili emerge che sono note le seguenti probabilità:

- $P(A_1) = 0.1$ probabilità di estrarre un individuo **malato**
- $P(A_2) = 0.9$ probabilità di estrarre un individuo **sano**
- $P(B_2|A_2) = 0.2$ probabilità che il test dia un **falso-positivo**
- $P(B_1|A_1) = 0.1$ probabilità che il test dia un **falso-negativo**

$$P(A_1|B_2) = \frac{P(B_2|A_1) \cdot P(A_1)}{P(B_2|A_1) \cdot P(A_1) + P(B_2|A_2) \cdot P(A_2)}$$

Teorema di Bayes: applicazione

Considerando che:

$P(B_1|A_1) = 0.1$ probabilità che il test dia un **falso-negativo**

Ne consegue:

$P(B_2|A_1) = 1 - P(B_1|A_1) = 0.9$ probabilità che il test dia un **vero negativo**

$$P(A_1|B_2) = \frac{0.9 \cdot 0.1}{0.9 \cdot 0.1 + 0.2 \cdot 0.9} = 0.33$$

Quindi la probabilità che un individuo che risulti positivo al test sia effettivamente malato è pari a 0.33.

Teorema di Bayes: applicazione

È noto che in un determinato comune A, il 90% dei laureati è occupato, mentre per i non laureati tale quota scende al 70%. Sapendo che nel comune il 50% dei residenti è laureato, calcolare la probabilità di estrarre casualmente un laureato dalla sottopopolazione degli occupati

Teorema di Bayes: applicazione

Elenchiamo gli eventi che stiamo considerando:

- A_1: l'individuo è laureato
- A_2: l'individuo non è laureato
- B_1: l'individuo è occupato
- B_2: l'individuo è non occupato

La risposta alla nostra domanda richiede la determinazione della probabilità:

$\boldsymbol{P(A_1|B_1)}$ = ***probabilità di estrarre casualmente un laureato dalla sottopopolazione degli occupati***.

Teorema di Bayes: applicazione

Dai dati disponibili emerge che sono note le seguenti probabilità:

- $P(A_1) = 0.5$ probabilità di estrarre un individuo **laureato**
- $P(A_2) = 0.5$ probabilità di estrarre un individuo **non laureato**
- $P(B_1|A_1) = 0.9$ probabilità di essere occupato essendo laureato
- $P(B_1|A_2) = 0.7$ probabilità di essere occupato non essendo laureato

$$P(A_1|B_1) = \frac{P(B_1|A_1) \cdot P(A_1)}{P(B_1|A_1) \cdot P(A_1) + P(B_1|A_2) \cdot P(A_2)}$$

Teorema di Bayes: applicazione

Considerando che:

$P(B_1|A_1) = 0.9$ probabilità di essere occupato essendo laureato

e

$P(B_1|A_2) = 0.7$ probabilità di essere occupato non essendo laureato

$$P(A_1|B_1) = \frac{P(B_1|A_1) \cdot P(A_1)}{P(B_1|A_1) \cdot P(A_1) + P(B_1|A_2) \cdot P(A_2)}$$

$$= \frac{0.9 \cdot 0.5}{0.9 \cdot 0.5 + 0.7 \cdot 0.5} = 0.5625$$

Quindi la probabilità di estrarre casualmente un laureato dalla sottopopolazione degli occupati è pari al 56.25%.

Teorema di Bayes: applicazione

In una partita di calcio, le probabilità dei tre eventi: la squadra di casa “vince”, “pareggia”, o “perde” sono stimate rispettivamente con 0.5, 0.3 e 0.2. Avvicinandosi al ritrovo dei tifosi locali si vede esposta la bandiera della squadra di casa, il che accade quando essa non ha perso. Come si valuta adesso le probabilità dei tre eventi?

Teorema di Bayes: applicazione

Elenchiamo gli eventi che stiamo considerando:

- A_1: la squadra di casa vince
- A_2: la squadra di casa pareggia
- A_3: la squadra di casa perde
- B: bandiera esposta

La risposta alla nostra domanda richiede la determinazione delle probabilità:

$\boldsymbol{P(A_1|B)}$ = ***probabilità di aver vinto data la bandiera esposta***

$\boldsymbol{P(A_2|B)}$ = ***probabilità di aver pareggiato data la bandiera esposta***

$\boldsymbol{P(A_3|B)}$ = ***probabilità di aver perso data la bandiera esposta***

Teorema di Bayes: applicazione

Dai dati disponibili emerge che sono note le seguenti probabilità:

- $P(A_1) = 0.5$ probabilità di vittoria
- $P(A_2) = 0.3$ probabilità di pareggio
- $P(A_3) = 0.2$ probabilità di sconfitta
- $P(B|A_1) = 1$ probabilità di esporre la bandiera avendo vinto
- $P(B|A_2) = 1$ probabilità di esporre la bandiera avendo pareggiato
- $P(B|A_3) = 0$ probabilità di esporre la bandiera avendo perso

$$P(A_i|B) = \frac{P(B|A_i) \cdot P(A_i)}{P(B|A_1) \cdot P(A_1) + P(B|A_2) \cdot P(A_2) + P(B|A_3) \cdot P(A_3)}$$

Teorema di Bayes: applicazione

Calcoliamo $P(B)$:

$$P(B) = P(B|A_1) \cdot P(A_1) + P(B|A_2) \cdot P(A_2) + P(B|A_3) \cdot P(A_3)$$
$$= 1 \cdot 0.5 + 1 \cdot 0.3 + 0 \cdot 0.2 = 0.8$$

Possiamo calcolare quindi:

$$P(A_1|B) = \frac{P(B|A_1) \cdot P(A_1)}{P(B)} = \frac{0.5}{0.8} = 0.625$$

$$P(A_2|B) = \frac{P(B|A_2) \cdot P(A_2)}{P(B)} = \frac{0.3}{0.8} = 0.375$$

$$P(A_3|B) = \frac{P(B|A_3) \cdot P(A_3)}{P(B)} = \frac{0}{0.8} = 0$$

*