

Probabilità e Processi Stocastici (455AA)

Lezione 2

Dario Trevisan

28/09/2023

Section 1

Richiami

Regole principali

- $0 \leq P(A|I) \leq 1$

Regole principali

- $0 \leq P(A|I) \leq 1$
- **Somma** $P(A \text{ oppure } B) = P(A) + P(B)$ purché A, B incompatibili,

$$P(A \text{ e } B) = 0.$$

Regole principali

- $0 \leq P(A|I) \leq 1$
- **Somma** $P(A \text{ oppure } B) = P(A) + P(B)$ purché A, B incompatibili,

$$P(A \text{ e } B) = 0.$$

- Negazione $P(\text{ non } A) = 1 - P(A)$

Regole principali

- $0 \leq P(A|I) \leq 1$
- **Somma** $P(A \text{ oppure } B) = P(A) + P(B)$ purché A, B incompatibili,

$$P(A \text{ e } B) = 0.$$

- Negazione $P(\text{ non } A) = 1 - P(A)$
- **Prodotto** $P(A \text{ e } B) = P(A)P(B|A)$.

Regole principali

- $0 \leq P(A|I) \leq 1$
- **Somma** $P(A \text{ oppure } B) = P(A) + P(B)$ purché A, B incompatibili,

$$P(A \text{ e } B) = 0.$$

- Negazione $P(\text{ non } A) = 1 - P(A)$
- **Prodotto** $P(A \text{ e } B) = P(A)P(B|A)$.
- Sistema di alternative $(A_i)_{i=1}^n$ se *esattamente una* delle A_i si realizza (ma non si sa quale)

Regole principali

- $0 \leq P(A|I) \leq 1$
- **Somma** $P(A \text{ oppure } B) = P(A) + P(B)$ purché A, B incompatibili,

$$P(A \text{ e } B) = 0.$$

- Negazione $P(\text{ non } A) = 1 - P(A)$
- **Prodotto** $P(A \text{ e } B) = P(A)P(B|A)$.
- Sistema di alternative $(A_i)_{i=1}^n$ se *esattamente una* delle A_i si realizza (ma non si sa quale)
- Densità discreta associata a un sistema $(A_i)_{i=1}^n$ e all'informazione I :

$$i \mapsto P(A_i|I)$$

Regole principali

- $0 \leq P(A|I) \leq 1$
- **Somma** $P(A \text{ oppure } B) = P(A) + P(B)$ purché A, B incompatibili,

$$P(A \text{ e } B) = 0.$$

- Negazione $P(\text{ non } A) = 1 - P(A)$
- **Prodotto** $P(A \text{ e } B) = P(A)P(B|A)$.
- Sistema di alternative $(A_i)_{i=1}^n$ se *esattamente una* delle A_i si realizza (ma non si sa quale)
- Densità discreta associata a un sistema $(A_i)_{i=1}^n$ e all'informazione I :

$$i \mapsto P(A_i|I)$$

- Densità Bernoulli, densità uniforme

Regole principali

- $0 \leq P(A|I) \leq 1$
- **Somma** $P(A \text{ oppure } B) = P(A) + P(B)$ purché A, B incompatibili,

$$P(A \text{ e } B) = 0.$$

- Negazione $P(\text{ non } A) = 1 - P(A)$
- **Prodotto** $P(A \text{ e } B) = P(A)P(B|A)$.
- Sistema di alternative $(A_i)_{i=1}^n$ se *esattamente una* delle A_i si realizza (ma non si sa quale)
- Densità discreta associata a un sistema $(A_i)_{i=1}^n$ e all'informazione I :

$$i \mapsto P(A_i|I)$$

- Densità Bernoulli, densità uniforme
- Diagrammi ad albero

Estrazioni da un'urna (senza rimpiazzo)

- **Probabilità di estrarre una precisa sequenza ordinata** di $n \leq N$ palline colorate, di cui $r \leq R$ sono rosse e le rimanenti $b \leq B$ sono blu:

$$\frac{R(R-1) \cdot \dots \cdot (R-r+1) \cdot B(B-1) \cdot \dots \cdot (B-b+1)}{N(N-1) \cdot \dots \cdot (N-n+1)}.$$

- **Probabilità di estrarre una qualsiasi sequenza ordinata** di $n \leq N$ palline colorate, di cui $r \leq R$ sono rosse e le rimanenti $b \leq B$ sono blu:

$$\frac{\binom{R}{r} \binom{B}{b}}{\binom{N}{n}}.$$

Section 2

Formula di Bayes e applicazioni

Formula di Bayes

Date affermazioni A , B e l'informazione nota I , vale

$$P(A|B, I) = P(A|I) \frac{P(B|A, I)}{P(B|I)}$$

(purché $P(B|I) > 0$).

Verosimiglianza (*Likelihood*) di A rispetto a B ,

$$L(A; B) = P(B|A).$$

Statistica bayesiana

Dato un sistema di alternative $(A_i)_{i=1}^n$ e una qualsiasi affermazione B , otteniamo

$$P(A_i|B) = P(A_i)P(B|A_i) \cdot \frac{1}{P(B)}. \quad (1)$$

Notazione compatta (più facile da ricordare)

$$P(A_i|B) \propto P(A_i)P(B|A_i) = P(A_i)L(A_i; B). \quad (2)$$

Massimo a posteriori stima

$$i_{\text{MAP}} \in \arg \max \{P(A_i)L(A_i; B) : i \in \{1, \dots, n\}\},$$

Stima di **massima verosimiglianza**

$$i_{\text{MLE}} \in \arg \max_{i=1, \dots, n} L(A_i; B).$$

Esempio: estrazioni da un'urna non nota

Il robot *non* è informato sul numero di palline rosse R , ma solamente sul totale $N = 3$. Dopo un'estrazione, si osserva una pallina rossa è stata estratta. Cosa può dedurre circa il contenuto dell'urna?

Introduciamo un sistema di alternative relativo al contenuto dell'urna:

$A_i =$ l'urna contiene $R = i$ palline rosse,

per $i \in \{0, 1, 2, 3\}$.

Probabilità a priori uniforme

$$P(A_i|\Omega) = \frac{1}{4}.$$

Verosimiglianza

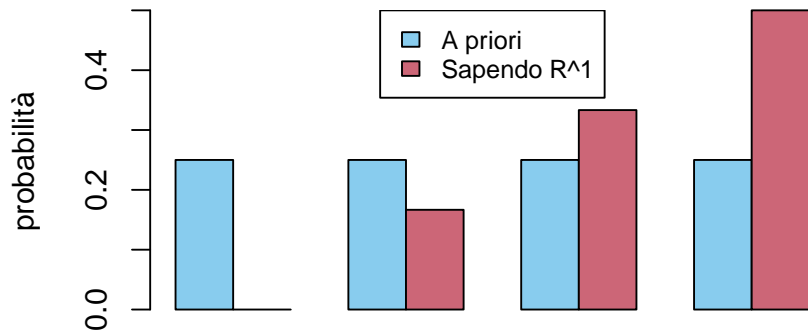
$$L(A_i; R^1) = P(R^1|A_i) = \frac{i}{3},$$

e formula di Bayes:

Possiamo confrontare le densità discrete mediante un grafico a barre:

```
## [1] 0.25 0.25 0.25 0.25
```

```
## [1] 0.0000000 0.1666667 0.3333333 0.5000000
```



Decisioni e test statistici

Obiettivo: scartare (rifiutare) alcune alternative A_i sulla base dell'osservazione di B . **Intuizione:** eliminare quelle con probabilità a posteriori nulla (ovvio) o bassa.

Nella pratica ci si appoggia alla verosimiglianza invece delle probabilità a posteriori (simile se a priori sono uniformi).

Ipotesi nulla \mathcal{H}_0 (da rifiutare)) e alternativa \mathcal{H}_1 . **Valore p** (p -value) di Fisher:

$$p = \max_{A_i \in \mathcal{H}_0} P(B|A_i) = \max_{A_i \in \mathcal{H}_0} L(A_i; B).$$

Più piccolo è il valore p , minore è la probabilità che B sia vero rispetto a l'ipotesi \mathcal{H}_0 , e quindi, invocando Bayes, anche che A_i sia vero sapendo B . Quindi possiamo rifiutare \mathcal{H}_0 con maggiore sicurezza.

Torniamo all'esempio dell'urna: seconda estrazione

Cosa deduce il robot se viene informato che alla seconda estrazione (senza rimpiazzo) la pallina estratta è blu (B^2)?

Non serve tornare alla densità iniziale (uniforme), ma basta usare come nuova densità *a priori* la densità discreta rispetto all'informazione R^1 . Bayes:

$$P(A_i|R^1, B^2) = P(A_i|R^1)P(B^2|R^1, A_i) \cdot \frac{1}{P(B^2|R^1)}.$$

Otteniamo

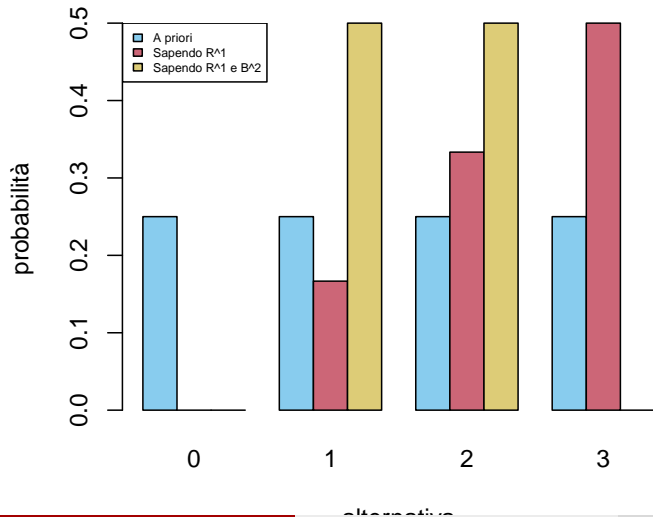
$$P(B^2|R^1, A_i) = \frac{3-i}{2}.$$

e

$$P(A_i|R^1, B^2) \propto i(3-i)$$

.

[1] 0.0 0.5 0.5 0.0



Section 3

Cenni agli assiomi di Kolmogorov

Problemi dell'approccio elementare

- 1 Come attribuire le **probabilità iniziali** (quelle che abbiamo chiamato *a priori*)?

Problemi dell'approccio elementare

- 1 Come attribuire le **probabilità iniziali** (quelle che abbiamo chiamato *a priori*)?
- 2 Come garantire la **consistenza** del calcolo, ossia che $P(A|I)$ sia ben definita?

Problemi dell'approccio elementare

- 1 Come attribuire le **probabilità iniziali** (quelle che abbiamo chiamato *a priori*)?
- 2 Come garantire la **consistenza** del calcolo, ossia che $P(A|I)$ sia ben definita?
- 3 Come trattare i **passaggi al limite**, in particolare, nel caso di infinite affermazioni?

Alcune risposte sono fornite dalla **descrizione assiomatica** della probabilità proposta da Kolmogorov.

L'idea: formalizzare i diagrammi di Eulero-Venn, identificando

- le affermazioni (eventi) A, I ecc. con *sottoinsiemi* di un insieme “universo” Ω (l'informazione iniziale)

Alcune risposte sono fornite dalla **descrizione assiomatica** della probabilità proposta da Kolmogorov.

L'idea: formalizzare i diagrammi di Eulero-Venn, identificando

- le affermazioni (eventi) A, I ecc. con *sottoinsiemi* di un insieme “universo” Ω (l'informazione iniziale)
- la probabilità $P(A|I)$ con una nozione astratta di *area* (relativa), definendo prima la probabilità rispetto all'informazione iniziale Ω .

a) insieme Ω

Si fissa un insieme “universo” Ω che rappresenta tutte le possibili situazioni (scenari) che si potrebbero presentare nel problema che si sta considerando.

Ad esempio, nel caso di un lancio di dado,

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

che corrisponde ai possibili esiti (ovviamente altre scelte sono ragionevoli).

b) Eventi

Si identificano quali affermazioni $A \subseteq \Omega$ sono d'interesse per il problema che si vuole affrontare.

Insieme \mathcal{A} i cui elementi $A \in \mathcal{A}$ sono sottoinsiemi di Ω , detto la *σ -algebra degli eventi*:

- $\Omega \in \mathcal{A}$,

b) Eventi

Si identificano quali affermazioni $A \subseteq \Omega$ sono d'interesse per il problema che si vuole affrontare.

Insieme \mathcal{A} i cui elementi $A \in \mathcal{A}$ sono sottoinsiemi di Ω , detto la σ -algebra degli eventi:

- $\Omega \in \mathcal{A}$,
- se $A, B \in \mathcal{A}$, anche A^c ("non A "), $A \cap B$ (" A e B ") e $A \cup B$ (" A oppure B ") sono eventi in \mathcal{A} .

b) Eventi

Si identificano quali affermazioni $A \subseteq \Omega$ sono d'interesse per il problema che si vuole affrontare.

Insieme \mathcal{A} i cui elementi $A \in \mathcal{A}$ sono sottoinsiemi di Ω , detto la *σ -algebra degli eventi*:

- $\Omega \in \mathcal{A}$,
- se $A, B \in \mathcal{A}$, anche A^c ("non A "), $A \cap B$ (" A e B ") e $A \cup B$ (" A oppure B ") sono eventi in \mathcal{A} .
- Inoltre, per permettere di passare al limite, si richiede che valga lo stesso per l'unione infinita di eventi: dati $A_n \in \mathcal{A}$, pure $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.

c) funzione di probabilità (iniziale)

Si introduce una $P : \mathcal{A} \rightarrow [0, 1]$ tale che $P(\Omega) = 1$ e, per ogni $A, B \in \mathcal{A}$ con $A \cap B = \emptyset$ valga

$$P(A \cup B) = P(A) + P(B).$$

- $P(A)$ corrisponde alla probabilità $P(A|\Omega)$ rispetto all'informazione iniziale, di cui si richiede valga la regola della somma (per eventi incompatibili).

c) funzione di probabilità (iniziale)

Si introduce una $P : \mathcal{A} \rightarrow [0, 1]$ tale che $P(\Omega) = 1$ e, per ogni $A, B \in \mathcal{A}$ con $A \cap B = \emptyset$ valga

$$P(A \cup B) = P(A) + P(B).$$

- $P(A)$ corrisponde alla probabilità $P(A|\Omega)$ rispetto all'informazione iniziale, di cui si richiede valga la regola della somma (per eventi incompatibili).
- Per passare al limite, si richiede in più che la regola della somma si estenda ad infiniti $A_n \in \mathcal{A}$ tali che $A_n \cap A_m = \emptyset$ per $n \neq m$:

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

d) Probabilità condizionata ad I

Per ogni $A, I \in \mathcal{A}$ tale che $P(I) > 0$, si pone

$$P(A|I) = \frac{P(A \cap I)}{P(I)}.$$

(Ω, \mathcal{A}, P) è uno **spazio di probabilità** secondo Kolmogorov.