

Indice

1	Concetti preliminari	1
1.1	Metodologia bayesiana	1
1.2	Funzioni belief e probabilità	2
1.2.1	Funzione belief (belief function)	2
1.2.2	Assiomi di credibilità	2
1.3	Partizioni e regola di Bayes	3

Capitolo 1

Concetti preliminari

Sotto il teorema di Bayes, nessuna teoria è perfetta. Piuttosto, è un lavoro in corso, sempre soggetto a ulteriori perfezionamenti e prove.

Nate Silver

La statistica bayesiana è un sottocampo della statistica in cui la verità di un enunciato è espressa in termini di credibilità (*believe*) o più specificatamente di **probabilità bayesiana**.

Definizione 1.0.1 (Probabilità bayesiana). La probabilità bayesiana è un'interpretazione del concetto di probabilità, in cui, anziché la frequenza di un qualche evento, la probabilità viene vista come aspettativa razionale che rappresenta un stato di conoscenza o come quantificazione di una convinzione personale.

Nella visione bayesiana assegnamo una probabilità ad un'ipotesi, che è la probabilità che l'ipotesi sia vera. Nell'approccio frequentista invece, l'ipotesi viene solo verificata senza che le venga assegnata una probabilità. Per valutare la probabilità di un'ipotesi si deve fornire una **probabilità iniziale**(*probabilità a priori*). Questa a sua volta viene aggiornata con l'arrivo di nuovi dati, che vengono utilizzati per calcolare la **probabilità aggiornata**(*probabilità a posteriori*).

1.1 Metodologia bayesiana

L'interpretazione bayesiana fornisce un insieme standard di procedure e formule per aggiornare la probabilità iniziale con l'arrivo di nuovi dati. Queste procedure sono chiamate **metodologia bayesiana** e sono le seguenti:

- Uso di variabili casuali, per modellare le fonti di incertezza nei modelli statistici, compresa l'incertezza risultante dalla mancanza di informazioni.

- La necessità di determinare la distribuzione di probabilità iniziale, tenendo conto delle informazioni disponibili.
- Uso ricorsivo del **teorema di Bayes** per aggiornare la probabilità iniziale con l'arrivo di nuovi dati.
- La probabilità che può essere assegnata ad un evento è un numero reale compreso tra 0 e 1.

La distribuzione finale di probabilità può a sua volta diventare una probabilità iniziale che può essere aggiornata con nuovi dati.

1.2 Funzioni belief e probabilità

La probabilità è un modo per esprimere numericamente un'aspettativa razionale. Mostriamo adesso che molte proprietà che hanno le nostre aspettative numeriche sono anche valide per le probabilità.

1.2.1 Funzione belief (belief function)

Definizione 1.2.1 (Funzione belief). La funzione belief è una funzione che associa ad ogni possibile evento o enunciato un valore reale compreso tra 0 e 1.

Siano F , G e H tre possibili enunciati che possono sovrapporsi. Sia $Be()$ una funzione belief che assegna un numero agli enunciati tale che maggiore è il numero assegnato, maggiore è la credibilità dell'enunciato (*degree of belief*). Ad esempio, se $Be(F) > Be(G)$ allora F è più credibile di G .

Vogliamo che la funzione descriva le nostre credenze sotto certe condizioni:

- $Be(F|H) > Be(G|H)$ significa che dato H vero, allora è più credibile F che G .
- $Be(F|G) > Be(F|H)$ significa che se siamo costretti a scommettere su F , lo facciamo più volentieri sotto la condizione che G sia vero rispetto alla condizione che H sia vero.

1.2.2 Assiomi di credibilità

Ogni funzione che rappresenta le nostre credenze deve avere le seguenti proprietà:

$$Be(\neg H|H) \leq Be(F|H) \leq Be(H|H) \quad (1.1)$$

$$Be(F \cup G|H) \geq \max\{Be(F|H), Be(G|H)\} \quad (1.2)$$

$$Be(F \cap G|H) \text{ può essere derivata da } Be(G|H) \text{ e } Be(F|G \cap H) \quad (1.3)$$

Adesso vediamo cosa significano le proprietà sopra elencate.

- **Proprietà 1.1:** Il numero che assegno a $Be(F|H)$ è compreso tra i numeri che assegniamo all'assenza di credenza e alla piena credenza.
- **Proprietà 1.2:** La credibilità di F o G dato H è almeno uguale alla credibilità di F o G .
- **Proprietà 1.3:** Possiamo decidere se F e G sono entrambi credibili dato H , prima decidendo se G è credibile dato H e successivamente decidendo se F è credibile dato G e H .

Possiamo confrontare gli assiomi di credibilità con gli assiomi di probabilità:

$$0 = P(\neg H|H) \leq P(F|H) \leq P(H|H) = 1 \quad (1.4)$$

$$P(F \cup G|H) = P(F|H) + P(G|H) \text{ se } F \cap G = \emptyset \quad (1.5)$$

$$P(F \cap G|H) = P(G|H)P(F|G \cap H) \quad (1.6)$$

Notiamo che una funzione che soddisfa gli assiomi di credibilità soddisfa anche gli assiomi di probabilità. Quindi se usiamo una funzione di probabilità per descrivere le nostre credenze, allora la funzione soddisfa gli assiomi di credibilità.

1.3 Partizioni e regola di Bayes

Definizione 1.3.1 (Partizione di un insieme). Una partizione dell'insieme H è una famiglia di sottoinsiemi $\{H_1, H_2, \dots, H_k\}$ che soddisfa le seguenti proprietà:

1.

$$H_i \cap H_j = \emptyset \quad \forall i \neq j \quad (1.7)$$

2.

$$\bigcup_{i=1}^k H_i = H \quad (1.8)$$

La famiglia $\{H_1, H_2, \dots, H_k\}$ è detta **disgiunta** se vale la proprietà 1.7, mentre è detta **completa** se vale la proprietà 1.8. Possiamo dire che $\{H_1, H_2, \dots, H_k\}$ è una **partizione di H** se è disgiunta e completa.

Prima di introdurre la regola di Bayes, vediamo due proprietà che derivano direttamente dagli assiomi di credibilità visti nella sezione 1.2.2.

Lemma 1.3.1. Sia $\{H_1, H_2, \dots, H_k\}$ un partizione di H , $P(H) = 1$ e sia E un evento specifico. Allora gli assiomi di probabilità ci dicono che:

$$\sum_{i=1}^k P(H_i) = 1. \quad (1.9)$$

$$P(E) = \sum_{i=1}^k P(E \cap H_i) = \sum_{i=1}^k P(E|H_i)P(H_i). \quad (1.10)$$

Dove 1.9 è detta **Regola di probabilità totale** e 1.10 è detta **Regola di probabilità marginale**.

Teorema 1.3.2 (Teorema di Bayes). Sia $\{H_1, H_2, \dots, H_k\}$ un partizione di H e sia $E \subset H \mid P(E) \neq 0$. Il teorema afferma che:

$$P(H_j|E) = \frac{P(E|H_j)P(H_j)}{P(E)} \stackrel{1.10}{=} \frac{P(E|H_j)P(H_j)}{\sum_{i=1}^k P(E|H_i)P(H_i)}. \quad (1.11)$$

L'evento E implica che si è avverata almeno un'alternativa e quindi $P(E) \neq 0$. Cosa importante, altrimenti il teorema non avrebbe senso, perchè non possiamo dividere per zero.

Dimostrazione. Il teorema deriva dalla definizione di probabilità condizionata e dalla regola di probabilità totale. La probabilità di un evento A dato un evento B è definita come:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}. \quad (1.12)$$

Analogamente, la probabilità di un evento B dato un evento A è definita come:

$$P(B|A) = \frac{P(A \cap B)}{P(A)}. \quad (1.13)$$

Usando la regola di probabilità marginale 1.10 possiamo scrivere la probabilità congiunta come:

$$P(A \cap B) = P(B|A)P(A). \quad (1.14)$$

Sostituendo 1.14 in 1.12 troviamo il teorema di bayes:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B|A)P(A)}{P(B)}. \quad (1.15)$$

□