Lecture 5: Bayes' rule and independence

Statistics 251

Yi Sun and Zhongjian Wang

Department of Statistics
The University of Chicago

Lecture Outline

Bayes' rule

Independence

Where are we?

Bayes' rule

Independence

Recollection on conditional probability

Remember that the conditional probability of E given F is

$$\mathbb{P}(E \mid F) = \frac{\mathbb{P}(E \cap F)}{\mathbb{P}(F)}.$$

Equivalently, we have

$$\mathbb{P}(E \cap F) = \mathbb{P}(F) \cdot \mathbb{P}(E \mid F).$$

Law of total probability

Suppose we want to compute $\mathbb{P}(E)$. For another event F, we have

$$\begin{aligned}\mathbb{P}(E) &= \mathbb{P}(E \cap F) + \mathbb{P}(E \cap F^c) \\ &= \mathbb{P}(E \mid F) \cdot \mathbb{P}(F) + \mathbb{P}(E \mid F^c) \cdot \mathbb{P}(F^c).\end{aligned}$$

Law of total probability

Suppose we want to compute $\mathbb{P}(E)$. For another event F, we have

$$\begin{aligned}\mathbb{P}(E) &= \mathbb{P}(E \cap F) + \mathbb{P}(E \cap F^c) \\ &= \mathbb{P}(E \mid F) \cdot \mathbb{P}(F) + \mathbb{P}(E \mid F^c) \cdot \mathbb{P}(F^c).\end{aligned}$$

Example: $D =$ "have disease", $+ =$ "test positive". Suppose $\mathbb{P}(D) = p$, $\mathbb{P}(+ \mid D) = 0.9$, and $\mathbb{P}(+ \mid D^c) = 0.1$. We have

$$\begin{aligned}\mathbb{P}(+) &= \mathbb{P}(+ \mid D) \cdot \mathbb{P}(D) + \mathbb{P}(+ \mid D^c) \cdot \mathbb{P}(D^c) \\ &= 0.9p + 0.1(1-p) = 0.1 + 0.8p.\end{aligned}$$

What we really care about is $\mathbb{P}(D \mid +)$, which is

$$\frac{\mathbb{P}(D \cap +)}{\mathbb{P}(+)} = \frac{\mathbb{P}(+ \mid D) \cdot \mathbb{P}(D)}{\mathbb{P}(+)} = \frac{0.9p}{0.9p + 0.1(1-p)}.$$

Bayes' rule

Bayes' rule: By $\mathbb{P}(B \mid A) \cdot \mathbb{P}(A) = \mathbb{P}(A \cap B) = \mathbb{P}(A \mid B) \cdot \mathbb{P}(B)$, we have

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

Bayes' rule

Bayes' rule: By $\mathbb{P}(B \mid A) \cdot \mathbb{P}(A) = \mathbb{P}(A \cap B) = \mathbb{P}(A \mid B) \cdot \mathbb{P}(B)$, we have

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

Interpretation: Start with estimate $\mathbb{P}(A)$ for A. After receiving new information, perform a **Bayesian update** to restrict the sample space to B.

- $\frac{\mathbb{P}(B|A)}{\mathbb{P}(B)}$ measures how strong the evidence is
- If $\mathbb{P}(B \mid A) = 0$, A and B are mutually exclusive.
- We have $\frac{\mathbb{P}(B|A)}{\mathbb{P}(B)} \leq \frac{1}{\mathbb{P}(A)}$, with equality if and only if $\mathbb{P}(A \cap B) = \mathbb{P}(B)$.

Bayesian updating

Draw a card at random from a deck. Define the events

$$A = \{\text{card is ace of spades}\}$$
$$B = \{\text{suit of cards is spades}\}.$$

To start, we know that $\mathbb{P}(A) = \frac{1}{52}$. If we now know that B occurred, we may update by

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(B \mid A)}{\mathbb{P}(B)}\mathbb{P}(A) = \frac{1}{1/4} \cdot \frac{1}{52} = \frac{1}{13}.$$

We may update as further information emerges about the card.

Bayesian updating

Draw a card at random from a deck. Define the events

$$A = \{\text{card is ace of spades}\}$$
$$B = \{\text{suit of cards is spades}\}.$$

To start, we know that $\mathbb{P}(A) = \frac{1}{52}$. If we now know that B occurred, we may update by

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(B \mid A)}{\mathbb{P}(B)}\mathbb{P}(A) = \frac{1}{1/4} \cdot \frac{1}{52} = \frac{1}{13}.$$

We may update as further information emerges about the card.

Bayesian updating in the real world...

We can assign probabilites to events which have not yet happened:

$$\mathbb{P}(\text{Biden wins election})$$
$$\mathbb{P}(\text{Cubs win the World Series})$$
$$\mathbb{P}(\text{stock prices will go up this year}).$$

According to Thomas Bayes:

$$\mathbb{P}(A) := \{\text{value of right to get \$1 if event occurs}\}.$$

This creates philosophical questions:

- Does this “value” have a well-defined price?
- How is $\mathbb{P}(A)$ defined when there are no enforceable financial contracts?
- Can we use this interpretation (and Bayes' rule) in everyday reasoning?

Where are we?

Bayes' rule

Independence

Independent events

Events E and F are **independent** if $\mathbb{P}(E \cap F) = \mathbb{P}(E)\mathbb{P}(F)$.

- Equivalent formulation: $\mathbb{P}(E \mid F) = \mathbb{P}(E)$
- Equivalent formulation: $\mathbb{P}(F \mid E) = \mathbb{P}(F)$

Toss two coins with sample space $\{(H,H),(H,T),(T,H),(T,T)\}$.

- {first coin H} and {second coin H} are independent
- {first coin H} and {odd number of H} are independent

Probability of each event is $\frac{1}{2}$, and probability of both is $\frac{1}{4}$.

Indpendence of multiple events

Events $E_1, \ldots, E_n$ are **independent** if for each $\{i_1, \ldots, i_k\} \subset \{1, \ldots, n\}$ we have

$$\mathbb{P}\Big(E_{i_1} \cap \cdots \cap E_{i_k}\Big) = \mathbb{P}(E_{i_1})\mathbb{P}(E_{i_2}) \cdots \mathbb{P}(E_{i_k}).$$

Implies statements like $\mathbb{P}(E_1 \cap E_2 \mid E_3 \cap E_4 \cap E_5) = \mathbb{P}(E_1 \cap E_2)$.

Indpendence of multiple events

Events $E_1, \ldots, E_n$ are **independent** if for each $\{i_1, \ldots, i_k\} \subset \{1, \ldots, n\}$ we have

$$\mathbb{P}\Big(E_{i_1} \cap \cdots \cap E_{i_k}\Big) = \mathbb{P}(E_{i_1})\mathbb{P}(E_{i_2}) \cdots \mathbb{P}(E_{i_k}).$$

Implies statements like $\mathbb{P}(E_1 \cap E_2 \mid E_3 \cap E_4 \cap E_5) = \mathbb{P}(E_1 \cap E_2)$.

Does pairwise independence imply independence?

Indpendence of multiple events

Events $E_1, \ldots, E_n$ are **independent** if for each $\{i_1, \ldots, i_k\} \subset \{1, \ldots, n\}$ we have

$$\mathbb{P}\Big(E_{i_1} \cap \cdots \cap E_{i_k}\Big) = \mathbb{P}(E_{i_1})\mathbb{P}(E_{i_2}) \cdots \mathbb{P}(E_{i_k}).$$

Implies statements like $\mathbb{P}(E_1 \cap E_2 \mid E_3 \cap E_4 \cap E_5) = \mathbb{P}(E_1 \cap E_2)$.

Does pairwise independence imply independence?

- No: Consider {first coin H}, {second coin H}, {odd number of H}

Indpendence: examples

Shuffle 4 cards labeled $1, 2, 3, 4$. Let

$$E_{i,j} = \{\text{card } i \text{ comes before card } j\}.$$

Is $E_{1,2}$ independent of $E_{3,4}$?

Is $E_{1,2}$ independent of $E_{1,3}$?

Indpendence: examples

What is $\mathbb{P}(E_{1,2} \mid E_{1,3})$?

What is $\mathbb{P}(E_{1,7} \mid E_{1,2} \cap E_{1,3} \cap \cdots \cap E_{1,6})$?