Lecture 6: Independence and Probability Calculation

Ye Tian

Department of Statistics, Columbia University
Calculus-based Introduction to Statistics (S1201)

July 13, 2022
Recap: conditional probability

**Definition:** Conditional probability of $A$ given $B$ is defined as

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

**Multiplication rule:**

- $P(A \cap B) = P(A)P(B|A)$
- $P(A \cap B \cap C) = P(A)P(B|A)P(C|A \cap B)$

**Rule of total probability:** If $\Omega = A_1 \cup A_2 \cup \cdots \cup A_k$ and $A_1, \ldots, A_k$ are mutually exclusive (such $A_1, \ldots, A_k$ are called a partition of the sample space), then for any event $B$,

$$
P(B) = P(A_1)P(B|A_1) + P(A_2)P(B|A_2) + \ldots + P(A_k)P(B|A_k)$$
Recap: Bayes' Theorem

**Bayes' Theorem:** If $A_1, \ldots, A_k$ are a partition of sample space $\Omega$, then for any event $B$,

$$
P(A_j|B) = \frac{P(A_j \cap B)}{P(B)} = \frac{P(B|A_j)P(A_j)}{\sum_{i=1}^{k} P(B|A_i)P(A_i)}, \quad j = 1, \ldots, k.
$$

Tree diagram:
Last lecture: Monty Hall problem

You're given the choice of three doors: Behind one door is an iPhone 13; behind the others, goats. You **randomly** pick a door, say No.1, and the host, **who knows what's behind the doors**, opens another door, say No.3, which has a goat. He then says to you, "Do you want to pick door No.2?"

**Question:** Would you switch your choice?

**Underlying assumptions:**
- If the iPhone is behind the door we pick, then the host **randomly** picks one door from the remaining two to open.
- Otherwise, the host opens the door of goat.
Last lecture: Monty Hall problem

Solution (by Bayes Theorem):

- \( P(\text{iPhone is behind door } i) = \frac{1}{3}, \ i = 1, 2, 3 \)

Then

\[
P(\text{win iPhone by staying with } \#1) = P(\text{iPhone behind } \#1|\text{host opens } \#3 \text{ and see a goat})
\]

\[
= \frac{P(\text{host opens } \#3 \text{ and see a goat}|\text{iPhone behind } \#1)P(\text{iPhone behind } \#1)}{\sum_{i=1}^{3} P(\text{host opens } \#3 \text{ and see a goat}|\text{iPhone behind } \#i)P(\text{iPhone behind } \#i)}
\]

\[
= \frac{\frac{1}{2} \times \frac{1}{3}}{\frac{1}{2} \times \frac{1}{3} + 1 \times \frac{1}{3} + 0 \times \frac{1}{3}}
\]

\[
= \frac{\frac{1}{3}}{\frac{1}{3}} = \frac{1}{3}
\]
Last lecture: Monty Hall problem

Solution (by Bayes Theorem):

\[ \text{P(iPhone is behind door } i \text{)} = \frac{1}{3}, \quad i = 1, 2, 3 \]

Then

\[ \text{P(win iPhone after switching to } \#2) = \text{P(iPhone behind } \#2 \text{ | host opens } \#3 \text{ and see a goat)} \]

\[ = \frac{\text{P(host opens } \#3 \text{ and see a goat | iPhone behind } \#2) \text{P(iPhone behind } \#2)}{\sum_{i=1}^{3} \text{P(host opens } \#3 \text{ and see a goat | iPhone behind } \#i) \text{P(iPhone behind } \#i)} \]

\[ = \frac{1 \times \frac{1}{3}}{\frac{1}{2} \times \frac{1}{3} + 1 \times \frac{1}{3} + 0 \times \frac{1}{3}} \]

\[ = \frac{2}{3} \times \frac{1}{3} \quad \Rightarrow \quad \text{We should switch!} \]
Last lecture: Monty Hall problem (modified)

You **randomly** pick a door, say No.1, and the host, **who doesn't know what's behind the doors, randomly** opens another door among the remaining two, say No.3, which has a goat. He then says to you, "Do you want to pick door No.2?"

**Question:** Would you switch your choice?

**Underlying assumptions:**
- No matter which one you originally pick, the host always **randomly** picks one door from the remaining two to open.
Last lecture: Monty Hall problem (modified)

Solution (by Bayes Theorem):

\[ \Pr(\text{iPhone is behind door } i) = \frac{1}{3}, \quad i = 1, 2, 3 \]

Then

\[ \Pr(\text{win iPhone by staying with } \#1) = \Pr(\text{iPhone behind } \#1 | \text{host opens } \#3 \text{ and see a goat}) \]

\[ = \frac{\Pr(\text{host opens } \#3 \text{ and see a goat} | \text{iPhone behind } \#1) \Pr(\text{iPhone behind } \#1)}{\sum_{i=1}^{3} \Pr(\text{host opens } \#3 \text{ and see a goat} | \text{iPhone behind } \#i) \Pr(\text{iPhone behind } \#i)} \]

\[ = \frac{\frac{1}{2} \times \frac{1}{3}}{\frac{1}{2} \times \frac{1}{3} + \frac{1}{2} \times \frac{1}{3} + 0 \times \frac{1}{3}} \]

\[ = \frac{1}{2}. \]
Last lecture: Monty Hall problem (modified)

Solution (by Bayes Theorem):

- $\mathbb{P}(\text{iPhone is behind door } i) = 1/3$, $i = 1, 2, 3$

Then

$$
\mathbb{P}(\text{win iPhone after switching to } #2) = \mathbb{P}(\text{iPhone behind } #2 \text{ and see a goat}|\text{host opens } #3) \\
= \frac{\mathbb{P}(\text{host opens } #3 \text{ and see a goat}|\text{iPhone behind } #2) \mathbb{P}(\text{iPhone behind } #2)}{\sum_{i=1}^{3} \mathbb{P}(\text{host opens } #3 \text{ and see a goat}|\text{iPhone behind } #i) \mathbb{P}(\text{iPhone behind } #i)}
$$

$$
= \frac{\frac{1}{2} \times \frac{1}{3}}{\frac{1}{2} \times \frac{1}{3} + \frac{1}{2} \times \frac{1}{3} + 0 \times \frac{1}{3}}
$$

$$
= \frac{1}{2} \quad \Rightarrow \text{It doesn't matter!!}
$$
Reflection on Monty Hall problem

○ Why is the answer so different?
○ A more intuitive explanation of the original Monty Hall problem:

<table>
<thead>
<tr>
<th>Behind door 1</th>
<th>Behind door 2</th>
<th>Behind door 3</th>
<th>Result if staying at door #1</th>
<th>Result if switching to the door offered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goat</td>
<td>Goat</td>
<td>iPhone</td>
<td>Wins goat</td>
<td>Wins iPhone</td>
</tr>
<tr>
<td>Goat</td>
<td>iPhone</td>
<td>Goat</td>
<td>Wins goat</td>
<td>Wins iPhone</td>
</tr>
<tr>
<td>iPhone</td>
<td>Goat</td>
<td>Goat</td>
<td>Wins iPhone</td>
<td>Wins goat</td>
</tr>
</tbody>
</table>

The host always helps us kick out a wrong choice in the remaining two!

○ But in the modified version, the host selects the door randomly, which doesn't benefit switching.
Today's goal

- Understand the definition of independence of events
- Know how to decide whether events are independent or not
- Comprehensively use all properties of probability we mentioned before to calculate relative problems
Independence
**Intuition**

In previous examples, often $P(A|B) \neq P(A)$, which means the information "B has occurred" resulted in a change in the likelihood of A occurring.

But sometimes, $P(A|B) = P(A)$, which means the chance that A will occur or has occurred is not affected by knowledge that B has occurred.

**Example:** A box contains 4 red tickets and 6 white tickets. Pick a ticket at random from the box.

- What is the chance of getting a winning ticket? (5/10)
- If you see that it is a white one, what is the chance of getting a winning ticket? (3/6)
Independence between two events

**Definition:** Two events $A$ and $B$ are independent if $P(A|B) = P(A)$, denoted as $A ⊥⊥ B$. They are dependent otherwise, denoted as $A ∥ B$.

**Question:** What can we say about $\emptyset$ and $\Omega$?

**Equivalent definitions for $A ⊥⊥ B$:**
- $P(B|A) = P(B)$
- $P(A \cap B) = P(A)P(B)$ (very useful!)

**Properties:** $A$ and $B$ are independent iff (if and only if) any of the following holds:
- $A$ and $B^c$ are independent
- $A^c$ and $B$ are independent
- $A^c$ and $B^c$ are independent

You will prove these in HW2!
Example

Consider a gas station with six pumps numbered 1, 2,..., 6. Suppose that
\[ P(\{1\}) = P(\{6\}) = 0.1, \; P(\{2\}) = P(\{5\}) = 0.15, \; P(\{3\}) = P(\{4\}) = 0.25. \]

Define events \( A = \{2, 4, 6\} \), \( B = \{1, 2, 3\} \), \( C = \{2, 3, 4, 5\} \).

Therefore:
- \( P(A) = 0.5, \; P(B) = 0.5, \; P(C) = 0.8. \)
- \( P(A \cap B) = P(\{2\}) = 0.15 \)
- \( P(B \cap C) = P(\{2, 3\}) = 0.4 \)
- \( P(A \cap C) = P(\{2, 4\}) = 0.4 \)

Hence:
- \( P(A \cap B) \neq P(A)P(B) \Rightarrow A \not\perp\!
\perp B \)
- \( P(B \cap C) = P(B)P(C) \Rightarrow B \perp\!
\perp C \)
- \( P(A \cap C) = P(A)P(C) \Rightarrow A \perp\!
\perp C \)
Independence between multiple events

**Definition:** Events $A_1, \ldots, A_n$ are (mutually) independent if for every $k = 2, 3, \ldots, n$ and every subset of indices $i_1, i_2, \ldots, i_k$,

$$
P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}) = P(A_{i_1})P(A_{i_2}) \cdots P(A_{i_k}).$$

- When $n = 2$: it reduces to $P(A_1 \cap A_2) = P(A_1)P(A_2)$
- When $n = 3$: it reduces to (all the following conditions have to hold at the same time):
  1. $P(A_1 \cap A_2) = P(A_1)P(A_2)$
  2. $P(A_2 \cap A_3) = P(A_2)P(A_3)$
  3. $P(A_1 \cap A_3) = P(A_1)P(A_3)$
  4. $P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3)$

Therefore, independence is often an idealistic assumption in practice!
Independence between multiple events

**Theorem:** If $A_1, \ldots, A_n$ are mutually independent, then for any index sets \{i_1, \ldots, i_k\} and \{j_1, \ldots, j_r\} where \{i_1, \ldots, i_k\} \cap \{j_1, \ldots, j_r\} = \emptyset,$

intersections/uniions between $A_{i_1}, \ldots, A_{i_k} \perp$

intersections/uniions between $A_{j_1}, \ldots, A_{j_r}.$

The events $A_{i_1}, \ldots, A_{i_k}$ and $A_{j_1}, \ldots, A_{j_r}$ can be replaced by their complements.

For example:

- $(A_1 \cup A_2) \perp (A_3 \cap A_4)$
- $(A_1 \cap A_2^c \cup A_5) \perp (A_3 \cup A_4)$
- $A_2 \perp (A_1^c \cap A_4 \cap A_3)$
- ...

Example: parallel system

A system consists of 5 independent components $C_1, \ldots, C_5$. $\mathbb{P}(C_i \text{ works properly}) = 0.9$, $i = 1, \ldots, 5$. The chance that the system will work properly $= ?$

\[
\{\text{The whole system works}\} = (C_1 \cup C_2) \cap (C_3 \cup C_4 \cup C_5)
\]

\[
\mathbb{P}(\text{The whole system works}) = \mathbb{P}((C_1 \cup C_2) \cap (C_3 \cup C_4 \cup C_5))
\]

\[
= \mathbb{P}(C_1 \cup C_2) \times \mathbb{P}(C_3 \cup C_4 \cup C_5)
\]

- $\mathbb{P}(C_1 \cup C_2) = 1 - \mathbb{P}(C_1^c \cap C_2^c) = 1 - \mathbb{P}(C_1^c)\mathbb{P}(C_2^c) = 0.99$;
- $\mathbb{P}(C_3 \cup C_4 \cup C_5) = 1 - \mathbb{P}(C_3^c)\mathbb{P}(C_4^c)\mathbb{P}(C_5^c) = 1 - 0.1^3 = 0.999$

$\implies \mathbb{P}(\text{The whole system works}) = 0.99 \times 0.999 = 0.989$
Example: parallel system 2

A system consists of 3 **independent** components $C_1$, $C_2$, $C_3$.
$\mathbb{P}(C_i \text{ works properly}) = 0.7$, $i = 1, \ldots, 3$. The chance that the system will work properly =?

\[
\{\text{The whole system works}\} = (C_1 \cap C_2) \cup C_3
\]

\[
\mathbb{P}(\text{The whole system works}) = \mathbb{P}((C_1 \cap C_2) \cup C_3)
\]
\[
= \mathbb{P}(C_1 \cap C_2) + \mathbb{P}(C_3) - \mathbb{P}(C_1 \cap C_2 \cap C_3)
\]
\[
= \mathbb{P}(C_1)\mathbb{P}(C_2) + \mathbb{P}(C_3) - \mathbb{P}(C_1)\mathbb{P}(C_2)\mathbb{P}(C_3)
\]
\[
= 0.7^2 + 0.7 - 0.7^3
\]
\[
= 0.847
\]
**Example: coffee purchase**

The accompanying table gives information on the type of coffee selected by someone purchasing a single cup.

<table>
<thead>
<tr>
<th></th>
<th>Small</th>
<th>Medium</th>
<th>Large</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regular</td>
<td>14%</td>
<td>20%</td>
<td>26%</td>
</tr>
<tr>
<td>Decaf</td>
<td>20%</td>
<td>10%</td>
<td>10%</td>
</tr>
</tbody>
</table>

Consider randomly selecting such a coffee purchaser.

- Is "buying decaf coffee" independent with "buying medium-size coffee"?
- Are "buying regular coffee", "buying small-size coffee" and "buying medium-size coffee" mutually independent?

\[
P(\text{decaf} \cap \text{medium}) = 10\% \neq 40\% \times 30\% = P(\text{decaf}) \times P(\text{medium})
\]

\[\Rightarrow \text{decaf} \not\perp \text{medium}\]

Can we claim that \{regular\}, \{small\} and \{medium\} are NOT mutually independent?
Review: Probability Calculation
Review: Probability Calculation

\( \mathbb{P}(\cdot) : A \in \) a set of all events \( \rightarrow \) a number \( \mathbb{P}(A) \) (a mapping/function which maps an event to a number)

**Three axioms:**

- \( 0 \leq \mathbb{P}(A) \leq 1 \) for any event \( A \subseteq \Omega \)
- \( \mathbb{P}(\Omega) = 1, \mathbb{P}(\emptyset) = 0 \)
- If \( A_1, A_2, A_3, \ldots, A_n \) is a collection of disjoint (mutually exclusive) events, then

\[
\mathbb{P}(A_1 \cup A_2 \cup \ldots \cup A_n) = \sum_{i=1}^{n} \mathbb{P}(A_i).
\]

**Theorem:** In an experiment consisting of \( N \) outcomes with equal probability, for any event \( A \),

\[
\mathbb{P}(A) = \frac{N(A)}{N},
\]

where we use counting techniques to calculate \( N(A) \) and \( N \).
Review: Probability Calculation

**General idea to calculate probability:**

1. Translate the conditions and the event of interest by probability language
2. Calculate.
   - For equally likely outcomes, consider $\Pr(A) = \frac{N(A)}{N}$ combined with counting techniques
   - If conditioning probability is easier to calculate or it is given, consider conditional multiplication rule, rule of total probability, and Bayes' Theorem
   - See whether there are independence or not, which may help us to do the calculation
Example 1

Given $A$ and $B$, such that $\mathbb{P}(A) = \mathbb{P}(A^c \cap B) = 0.4$ and $\mathbb{P}(A \cap B) = 0.1$.

- Find $\mathbb{P}(A \cup B)$
- Are $A$ and $B$ independent?

- Venn diagram:

Or:

- $\mathbb{P}(B) = \mathbb{P}(A^c \cap B) + \mathbb{P}(A \cap B) = 0.4 + 0.1 = 0.5$
- Then $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$
  \[= 0.4 + 0.5 - 0.1 = 0.8\]
- No, because $\mathbb{P}(A \cap B) = 0.1 \neq 0.4 \times 0.5 = \mathbb{P}(A) \times \mathbb{P}(B)$. 
Example 2

5 faculty members (3 pro-A and 2 pro-B) vote for the department chair from candidate A and B one by one in a random order.

The probability that "A remains ahead of B throughout the vote count" = ? (e.g., "AABAB" is OK, but not "ABBAA")

Some observations:

- Permutation problem
- The last one has to vote for B. (Why?)
- The first one must vote for A. (Why?)
- So it has to be "ABABA", "AABAB" or "AAABB"

\[
N(\Omega) = P_{5,5} = 5! = 120
\]
\[
N(\text{A remains ahead of B}) = 3 \times P_{3,3} \times P_{2,2} = 36
\]
\[
P(\text{A remains ahead of B}) = \frac{N(\text{A remains ahead of B})}{N(\Omega)} = 0.3
\]
Example 3
Randomly select a student at Columbia:
- $A = \{\text{the selected student has a American Express card}\}$
- $B = \{\text{the selected student has a Barclays card}\}$
- $C = \{\text{the selected student has a Chase card}\}$
- $\mathbb{P}(A) = 0.6, \mathbb{P}(B) = 0.4, \mathbb{P}(C) = 0.2$
- $\mathbb{P}(A \cap B) = 0.3, \mathbb{P}(A \cap C) = 0.15, \mathbb{P}(B \cap C) = 0.1$
- $\mathbb{P}(A \cap B \cap C) = 0.08$

(1) The probability that the selected student has $\geq 1$ of 3 types of cards = ?
(2) $\mathbb{P}(B|A) =$ ?, $\mathbb{P}(A|B) =$ ?

\[ P(A \cup B \cup C) = 0.73 \]
\[ P(B|A) = \frac{P(A \cap B)}{P(A)} = 0.5, \]
\[ P(A|B) = \frac{P(A \cap B)}{P(B)} = 0.75 \]
Example 3

Randomly select a student at Columbia:

- $A = \{\text{the selected student has a American Express card}\}$
- $B = \{\text{the selected student has a Barclays card}\}$
- $C = \{\text{the selected student has a Chase card}\}$

- $\mathbb{P}(A) = 0.6$, $\mathbb{P}(B) = 0.4$, $\mathbb{P}(C) = 0.2$.
  $\mathbb{P}(A \cap B) = 0.3$, $\mathbb{P}(A \cap C') = 0.15$, $\mathbb{P}(B \cap C') = 0.1$
  $\mathbb{P}(A \cap B \cap C') = 0.08$

(3) Given that the student has an American Express card, what is the probability that she/he has $\geq 1$ of the other two types of cards?

\[
\mathbb{P}(B \cup C | A) = \frac{\mathbb{P}((B \cup C) \cap A)}{\mathbb{P}(A)} = \frac{0.22 + 0.08 + 0.07}{0.6} = 0.617
\]
Example 4

A system consists of 7 independent components $A_1, ..., D$.

- $P(A_1 \text{ works}) = P(A_2 \text{ works}) = 0.9$
- $P(B_1 \text{ works}) = P(B_2 \text{ works}) = 0.8$
- $P(C_1 \text{ works}) = P(C_2 \text{ works}) = 0.7$, $P(D \text{ works}) = 0.95$

The chance that the system will work properly = ?

\[
P(\{\text{the system works}\}) = P(((A_1 \cap B_1 \cap C_1) \cup (A_2 \cap B_2 \cap C_2)) \cap D) \\
= P((A_1 \cap B_1 \cap C_1) \cup (A_2 \cap B_2 \cap C_2))P(D) \\
= [2P(A_1 \cap B_1 \cap C_1) - P((A_1 \cap B_1 \cap C_1) \cap (A_2 \cap B_2 \cap C_2))]P(D) \\
= [2P(A_1)P(B_1)P(C_1) - P(A_1 \cap B_1 \cap C_1)P(A_2 \cap B_2 \cap C_2)]P(D) \\
= [2 \times 0.9 \times 0.8 \times 0.7 - (0.9 \times 0.8 \times 0.7)^2] \times 0.95 \\
= 0.717
\]
**Conditional independence**

**Independence:** Two events $A$ and $B$ are **independent** if
\[ P(A \cap B) = P(A)P(B), \]
denoted as $A \perp B$. They are **dependent** otherwise, denoted as $A \not\perp B$.

**Conditional independence:** Two events $A$ and $B$ are **independent** conditioning on event $C$ if
\[ P(A \cap B|C) = P(A|C)P(B|C), \]
denoted as $(A \perp B)|C$. They are **dependent** otherwise, denoted as $(A \not\perp B)|C$.

Can be extended to the case of multiple ($\geq 2$) events too.
Example 5

- 5% of the population has a certain disease
- A diagnostic test:
  - correctly detects the presence of the disease 98% of the time
  - correctly detects the absence of the disease 99% of the time
- Randomly select an individual → run tests twice → Both positive
- The two test results are independent given the presence/absence of the disease on the selected individual

Question: The probability that this individual has the disease = ?

- $P_i = \{i\text{-th test positive}\}, \ i = 1, 2$
- $D = \{\text{this individual has the disease}\}$
- Goal: $P(D|P_1 \cap P_2) =$?
- What we know:
  - $P(D) = 0.05, P(P_i|D) = 0.98, P(P_i^c|D^c) = 0.99, i = 1, 2$
  - $P_1 \perp \perp P_2$ given $D, P_1 \perp \perp P_2$ given $D^c$
- Bayes' Theorem!
Example 5

- $P_i = \{i\text{-th test positive}\}, \ i = 1, 2$
- $D = \{\text{this individual has the disease}\}$
- Goal: $P(D|P_1 \cap P_2) = ?$
- What we know:
  - $P(D) = 0.05, \ P(P_i|D) = 0.98, \ P(P_i^c|D^c) = 0.99, \ i = 1, 2$
  - $P_1 \perp P_2 \text{ given } D, \ P_1 \perp P_2 \text{ given } D^c$

By Bayes' Theorem,

$$P(D|P_1 \cap P_2) = \frac{P(P_1 \cap P_2|D)P(D)}{P(P_1 \cap P_2|D)P(D) + P(P_1 \cap P_2|D^c)P(D^c)}$$

$$= \frac{P(P_1|D)P(P_2|D)P(D)}{P(P_1|D)P(P_2|D)P(D) + P(P_1|D^c)P(P_2|D^c)P(D^c)}$$

$$= \frac{0.98 \times 0.98 \times 0.05}{0.98 \times 0.98 \times 0.05 + 0.01 \times 0.01 \times 0.95} = 0.998.$$
Reading list (optional)

- "Probability and Statistics for Engineering and the Sciences" (9th edition):
  - Chapter 2.5
- "OpenIntro statistics" (4th edition, free online, download [here]):
  - Chapter 3.1.7 and 3.2.6
Many thanks to
- Yang Feng
- Joyce Robbins
- Chengliang Tang
- Owen Ward
- Wenda Zhou
- And all my teachers in the past 25 years