Probability in Computing

LECTURE 8

Last time
• Probability Mass Functions (PMFs)
• Probability Density Functions (PDFs)
• Cumulative Distribution Functions (CDFs)

Today
• Conditional Probability

Reminders
• HW4 due Thursday
The annual death rate among people who know this statistic is one in six.

https://xkcd.com/795/
Incorporating New Information

• A patient has some unknown disease
• Based on the symptoms, the doctor estimates that the patient has:

  - Disease a: 55%
  - Disease b: 10%
  - Disease c: 30%

• Then a test reveals that the disease is neither b nor d
• Based on this information, what is the chance of having each disease?
Incorporating New Information

- A test eliminated diseases b and d

<table>
<thead>
<tr>
<th>Disease</th>
<th>Before the test</th>
<th>After the test</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>0.05</td>
<td>?</td>
</tr>
<tr>
<td>b</td>
<td>0.1</td>
<td>?</td>
</tr>
<tr>
<td>c</td>
<td>0.3</td>
<td>?</td>
</tr>
<tr>
<td>d</td>
<td>0.55</td>
<td>?</td>
</tr>
</tbody>
</table>
Incorporating New Information

- A test eliminated diseases b and d

<table>
<thead>
<tr>
<th>Disease</th>
<th>Before the test</th>
<th>After the test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disease a</td>
<td>0.05</td>
<td>?</td>
</tr>
<tr>
<td>Disease b</td>
<td>0.1</td>
<td>0</td>
</tr>
<tr>
<td>Disease c</td>
<td>0.3</td>
<td>?</td>
</tr>
<tr>
<td>Disease d</td>
<td>0.55</td>
<td>0</td>
</tr>
</tbody>
</table>
Incorporating New Information

- A test eliminated diseases b and d

<table>
<thead>
<tr>
<th>Disease</th>
<th>Before the test</th>
<th>After the test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disease a</td>
<td>0.05</td>
<td>0.05x</td>
</tr>
<tr>
<td>Disease b</td>
<td>0.1</td>
<td>0</td>
</tr>
<tr>
<td>Disease c</td>
<td>0.3</td>
<td>0.3x</td>
</tr>
<tr>
<td>Disease d</td>
<td>0.55</td>
<td>0</td>
</tr>
</tbody>
</table>

In the absence of any other information, the new probabilities of a and b should be proportional to their original probabilities.
Incorporating New Information

- A test eliminated diseases b and d

<table>
<thead>
<tr>
<th>Disease</th>
<th>Before the test</th>
<th>After the test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disease a</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>Disease b</td>
<td>0.1</td>
<td>0</td>
</tr>
<tr>
<td>Disease c</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td>Disease d</td>
<td>0.55</td>
<td>0</td>
</tr>
</tbody>
</table>

- Probabilities should add up to 1: $0.05x + 0.3x = 1$
  
  $x = \frac{1}{0.35}$
Incorporating New Information

- A test eliminated diseases $b$ and $d$

<table>
<thead>
<tr>
<th>Disease</th>
<th>Probability Before the Test</th>
<th>Probability After the Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a$</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>$b$</td>
<td>0.1</td>
<td>0</td>
</tr>
<tr>
<td>$c$</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td>$d$</td>
<td>0.55</td>
<td>0</td>
</tr>
</tbody>
</table>

Note: $\text{Pr("the disease is neither } b \text{ nor } d") = \text{Pr}\{a, c\} = 0.35$
Conditional Probability

- The information provided by the test is that the following event has happened:

  \[ B = "the disease is neither b nor d" = \{a, c\} \]

- We define a new probability function that assigns a probability \( \Pr(A \mid B) \) to every event \( A \subseteq \Omega \)

  \[ \Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)} \]

  **Read:** “Probability of \( A \) given \( B \)”

  **If** \( A = \{\omega\} \) **for an outcome** \( \omega \), **we also write:** \( \Pr(\omega \mid B) \)

  **If we specify** \( \Pr(\omega \mid B) \) **for every outcome** \( \omega \in \Omega \), **then** \( \Pr(A \mid B) \) **is determined for all events** \( A \) **(as the sum of probabilities of all outcomes in** \( A \)**)
The information provided by the test is that the following event has happened:

\[ B = "\text{the disease is neither } b \text{ nor } d" = \{a, c\} \]

We define a new probability function that assigns a probability \( P_r(A \mid B) \) to every event \( A \subseteq \Omega \):

\[
Pr(A \mid B) = \frac{Pr(A \cap B)}{Pr(B)}
\]

\[
Pr(a \mid B) = \frac{Pr(a)}{Pr(B)} = \frac{0.05}{0.35}
\]

\[ \Omega = \{a, b, c, d\} \]

\[ \Omega \]

\[ A \]

\[ B \]

\[ \Omega \]
Conditional Probability

• The information provided by the test is that the following event has happened:

\[ B = "the\ disease\ is\ neither\ b\ nor\ d" = \{a, c\} \]

• We define a new probability function that assigns a probability \( \Pr(A \mid B) \) to every event \( A \subseteq \Omega \)

\[
\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)}
\]

\[
\Pr(c \mid B) = \frac{\Pr(c)}{\Pr(B)} = \frac{0.3}{0.35}
\]

\[ \Omega = \{a, b, c, d\} \]
Conditional Probability

- The information provided by the test is that the following event has happened:
  
  \[ B = "the disease is neither b nor d" = \{a, c\} \]

- We define a new probability function that assigns a probability \( \Pr(A \mid B) \) to every event \( A \subseteq \Omega \)

\[
\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)}
\]

\[
\Pr(\{a, b\} \mid B) = \frac{\Pr(\{a\})}{\Pr(B)} = \frac{0.05}{0.35}
\]
Conditional Probability

- The information provided by the test is that the following event has happened:

  \[ B = \"the disease is neither b nor d\" = \{a, c\} \]

- We define a new probability function that assigns a probability \( \Pr(A \mid B) \) to every event \( A \subseteq \Omega \)

  \[
  \Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)}
  \]

  \[
  \Pr(\{a, c\} \mid B) = \frac{\Pr(\{a, c\})}{\Pr(B)} = 1
  \]
**Definition: Conditional Probability**

Let $B$ be an event such that $\Pr(B) \neq 0$. For every event $A$, we define the conditional probability of event $A$ given event $B$:

$$
\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)}.
$$

$\Pr(A \mid B)$ is undefined if $\Pr(B) = 0$. 
Experiment: roll two fair 6-sided dice

A = “first die is 6”
B = “sum of the two rolls is 9”

What is $\Pr(A \mid B)$?

A. $\frac{1}{36}$
B. $\frac{1}{6}$
C. $\frac{1}{4}$
D. None of the above
• Experiment: toss a fair coin three times
• A = “more H than T”
• B = “first toss is H”

What is \( \Pr(A \mid B) \)?

A. \( \frac{3}{8} \)
B. \( \frac{1}{2} \)
C. \( \frac{3}{4} \)
D. None of the above
• Experiment: spin the dial of the spinner
• A = “dial stopped in the upper half of the circle”
• B = “outcome is in [1/3, 2/3]”

What is $\Pr(A \mid B)$?

A. $\frac{1}{3}$
B. $\frac{1}{2}$
C. $\frac{3}{4}$
D. None of the above
• Experiment: spin the dial of the spinner
• A = “dial stopped in the upper half of the circle”
• C = “outcome is in [1/6, 1/3]”

What is $\Pr(A | C)$?

A. $\frac{1}{12}$
B. $\frac{1}{6}$
C. 1
D. None of the above
Mr. Smith has two children (boys and girls are equally likely)

- A = “both children are boys”
- B = “at least one child is a boy”
- C = “older child is a boy”

True or false: \( \Pr(A \mid B) = \Pr(A \mid C) \)?

A. True
B. False
Conditional probabilities satisfy the axioms of probability

- **Non-negativity**: $\Pr(A \mid E) \geq 0$ for all events $A \subseteq \Omega$

- **Additivity**: if $A$ and $B$ are *disjoint* events then
  \[
  \Pr(A \cup B \mid E) = \Pr(A \mid E) + \Pr(B \mid E)
  \]

- **Normalization**: $\Pr(\Omega \mid E) = 1$

- This applies only to conditional probabilities with the same conditioning event: $\Pr(\cdot \mid E)$
Conditional probabilities satisfy the axioms of probability

- **Non-negativity:** $\Pr(A \mid E) \geq 0$ for all events $A \subseteq \Omega$

Proof:

$$\Pr(A \mid E) = \frac{\Pr(A \cap E)}{\Pr(E)}$$

- **Normalization:** $\Pr(\Omega \mid E) = 1$

Proof:

$$\Pr(\Omega \mid E) = \frac{\Pr(\Omega \cap E)}{\Pr(E)}$$

$$= \frac{\Pr(E)}{\Pr(E)} = 1$$

by definition of conditional probability

by definition of conditional probability

since $E \subseteq \Omega$
Conditional probabilities satisfy the axioms of probability

- **Additivity:** if A and B are *disjoint* events then
  \[
  \Pr(A \cup B \mid E) = \Pr(A \mid E) + \Pr(B \mid E)
  \]

**Proof:**

\[
\Pr(A \cup B \mid E) = \frac{\Pr((A \cup B) \cap E)}{\Pr(E)}
\]

\[
= \frac{\Pr((A \cap E) \cup (B \cap E))}{\Pr(E)}
\]

\[
= \frac{\Pr(A \cap E) + \Pr(B \cap E))}{\Pr(E)}
\]

\[
= \Pr(A \mid E) + \Pr(B \mid E)
\]
Conditional Probability Rules

Since the conditional probability function is a valid probability function, all the probability rules remain valid.

- Example: Inclusion-Exclusion Principle

  **Standard:** \( \Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B) \)

  **Conditional:** \( \Pr(A \cup B \mid E) = \Pr(A \mid E) + \Pr(B \mid E) - \Pr(A \cap B \mid E) \)
Tree Diagrams: Example

- Best-of-three tournament between BU Terriers and Harvard Huskies
- **First game:** Terriers win with probability $\frac{1}{2}$
- **Subsequent games:**

  Terriers win a game with probability
  \[
  \begin{cases} 
  \frac{2}{3} & \text{if they won the previous game} \\
  \frac{1}{3} & \text{if they lost the previous game}
  \end{cases}
  \]

What is the probability that the Terriers win the tournament **given** that they win the first game?
Event $T =$ “Terriers win the tournament”
Event $E_1 =$ “Terriers win the first game”
Tree Diagram: Justification

• Why do we multiply probabilities along each branch in a tree diagram?

• Example:
  – Event $E_1 =$ "Terriers win the first game"
  – Event $E_2 =$ "Terriers win the second game"

• We calculated $\Pr(WW) = \Pr(E_1 \cap E_2)$
  $$= \Pr(E_1) \cdot \Pr(E_2 \mid E_1)$$

• Product rule follows from the definition of conditional probability:
  $$\Pr(E_2 \mid E_1) = \frac{\Pr(E_1 \cap E_2)}{\Pr(E_1)}$$