Conditional Probability and Bayes’ Theorem (2.4)
Independence (2.5)

Prof. Tesler

Math 186
Winter 2016
Scenario: Flip a fair coin three times

- Flip a coin 3 times. The sample space is
  \[ S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\} \]

- Define events
  \[ A = \text{“First flip is heads”} = \{HHH, HHT, HTH, HTT\} \]
  \[ B = \text{“Two flips are heads”} = \{HHT, HTH, THH\} \]

- Venn diagram:
A = “First flip is heads”  \[ P(A) = \frac{4}{8} \]
\[ = \{HHH, HHT, HTH, HTT\} \]

B = “Two flips are heads”  \[ P(B) = \frac{3}{8} \]
\[ = \{HHT, HTH, THH\} \]

Conditional probability:

- Flip a coin 3 times. If there are 2 heads, what’s the probability that the first flip is heads?
- **Rephrase:** Assuming B is true, what’s the probability of A?
- Since B is true, the coin flips are one of HHT, HTH, or THH.
- Out of those, the outcomes where A is true are HHT and HTH (which is \( A \cap B \)). So 2 out of the 3 possible outcomes in B give A.
- The probability of A, given that B is true, is
  \[ P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{2/8}{3/8} = \frac{2}{3} \]

\[ P(A \mid B) = \frac{P(A \cap B)}{P(B)} \]
Conditional probability

- $P(A)$ = probability of $A$
  measures $A$ as a fraction of the sample space.
- $P(A \mid B)$ = probability of $A$, given $B$
  measures $A \cap B$ as a fraction of $B$:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$
Scenario: rain and cancelling classes

Example

- There’s a 30% chance of rain tomorrow.
- **If it rains**, there’s a 10% chance classes will be cancelled.
- **If it doesn’t rain**, there’s only a 2% chance classes will be cancelled.

Questions

1. What’s the probability it will rain and classes will be cancelled?
2. What’s the total probability (whether or not it rains) that classes will be cancelled?
3. If classes are indeed cancelled tomorrow, what is the probability that it is due to rain?
1. Probability it will rain and classes will be cancelled?

- There’s a 30% chance of rain tomorrow.
- **If it rains**, there’s a 10% chance classes will be cancelled.
- **If it doesn’t rain**, there’s only a 2% chance classes will be cancelled.

Express the data using event notation

- Event $A =$ rains tomorrow
  
  $P(A) = .30 \quad P(A^c) = .70$

- Event $B =$ classes are cancelled tomorrow
  
  $P(B|A) = .10 \quad P(B|A^c) = .02$
1. Probability it will rain and classes will be cancelled?

- Events: $A =$ rains tomorrow, $B =$ classes cancelled tomorrow.
- There’s a 30% chance of rain tomorrow. $P(A) = .30$, $P(A^c) = .70$
- **If it rains**, there’s a 10% chance classes will be cancelled.
- **If it doesn’t rain**, there’s a 2% chance classes will be cancelled. $P(B|A) = .10$, $P(B|A^c) = .02$

Express the question using event notation: $P(A \cap B) =$ ?

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

Similarly, $P(A \cap B) = P(B|A)P(A)$.

So $P(A \cap B) = P(B|A)P(A) = ( .10) (.30) = \boxed{.03} = 3\%$. 

[Diagram showing Venn diagrams for events A and B]
2. Total probability classes will be cancelled?

- Events: $A = \text{rains tomorrow}$, $B = \text{classes cancelled tomorrow}$.
- There’s a 30% chance of rain tomorrow. $P(A) = .30$, $P(A^c) = .70$
- **If it rains**, there’s a 10% chance classes will be cancelled. $P(B|A) = .10$, $P(B|A^c) = .02$
- **If it doesn’t rain**, there’s a 2% chance classes will be cancelled.

Express the question using event notation: $P(B) = ?$

\[
P(B) = P(B \cap A) + P(B \cap A^c)
\]
\[
= P(B|A)P(A) + P(B|A^c)P(A^c)
\]
\[
= (.10)(.30) + (.02)(.70) = .03 + .014 = \boxed{.044} = 4.4\%
\]
3. If classes are cancelled, what’s the probability it’s due to rain?

- Events: $A = \text{rains tomorrow}, \ B = \text{classes cancelled tomorrow}.$
- There’s a 30% chance of rain tomorrow. $P(A) = .30, \ P(A^c) = .70$
- **If it rains**, there’s a 10% chance classes will be cancelled.
- **If it doesn’t rain**, there’s a 2% chance classes will be cancelled. $P(B|A) = .10, \ P(B|A^c) = .02$

Express the question using event notation: $P(A|B) = ?$

\[
P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B|A)P(A)}{P(B)} = \frac{(.10)(.30)}{.044} = .6818 \approx 68.2%\]
Bayes’ Theorem (simple version)

**Theorem (Bayes’ Theorem)**

\[
P(A|B) = \frac{P(B|A)P(A)}{P(B)}
\]

*This lets us express the probability of A given B, in terms of the probability of B given A.*

**Alternate formulation of Bayes’ Theorem**

\[
P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)}
\]

where we used

\[
P(B) = P(B \cap A) + P(B \cap A^c) = P(B|A)P(A) + P(B|A^c)P(A^c)
\]
Scenario: location of genes on chromosomes

Length of chromosomes & fraction in genes

<table>
<thead>
<tr>
<th>Chromosome #</th>
<th>Length (nt)</th>
<th>Prob. gene @ each nt</th>
</tr>
</thead>
<tbody>
<tr>
<td>__</td>
<td>1,000,000</td>
<td>0.05</td>
</tr>
<tr>
<td>____</td>
<td>2,000,000</td>
<td>0.06</td>
</tr>
<tr>
<td>______</td>
<td>3,000,000</td>
<td>0.07</td>
</tr>
<tr>
<td>________</td>
<td>4,000,000</td>
<td>0.08</td>
</tr>
<tr>
<td>Total</td>
<td>10,000,000</td>
<td></td>
</tr>
</tbody>
</table>

Sample space, events, and probabilities

- \( S = \text{sample space} = \text{all positions} \)
  \( N(S) = 10,000,000 \)
- \( A_i = \text{positions on chromosome } i \)
  \( P(A_1) = \frac{1,000,000}{10,000,000} = 0.1 \)
  \( P(A_2) = \frac{2}{10} = 0.2 \)
  \( P(A_3) = \frac{3}{10} = 0.3 \)
  \( P(A_4) = \frac{4}{10} = 0.4 \)
- \( B = \text{positions in genes} \)
  \( P(B|A_1) = 0.05 \)
  \( P(B|A_2) = 0.06 \)
  \( P(B|A_3) = 0.07 \)
  \( P(B|A_4) = 0.08 \)
Sample space, events, and probabilities

- $A_i = \text{positions on chromosome } i$
  - $P(A_1) = .1$  $P(A_2) = .2$  $P(A_3) = .3$  $P(A_4) = .4$
- $B = \text{positions in genes}$
  - $P(B|A_1) = .05$  $P(B|A_2) = .06$  $P(B|A_3) = .07$  $P(B|A_4) = .08$

Venn Diagram
## Breaking down the probabilities of events

### Sample space, events, and probabilities

- $A_i = \text{positions on chromosome } i$
  
  $P(A_1) = .1 \quad P(A_2) = .2 \quad P(A_3) = .3 \quad P(A_4) = .4$

- $B = \text{positions in genes}$
  
  $P(B|A_1) = .05 \quad P(B|A_2) = .06 \quad P(B|A_3) = .07 \quad P(B|A_4) = .08$

### Venn diagram with probabilities

<table>
<thead>
<tr>
<th></th>
<th>$A_1$</th>
<th>$A_2$</th>
<th>$A_3$</th>
<th>$A_4$</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B$</td>
<td>$P(B \cap A_1)$</td>
<td>$P(B \cap A_2)$</td>
<td>$P(B \cap A_3)$</td>
<td>$P(B \cap A_4)$</td>
<td>$P(B)$</td>
</tr>
<tr>
<td>$B^c$</td>
<td>$P(B^c \cap A_1)$</td>
<td>$P(B^c \cap A_2)$</td>
<td>$P(B^c \cap A_3)$</td>
<td>$P(B^c \cap A_4)$</td>
<td>$P(B^c)$</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>$P(A_1)$</td>
<td>$P(A_2)$</td>
<td>$P(A_3)$</td>
<td>$P(A_4)$</td>
<td><strong>1</strong></td>
</tr>
</tbody>
</table>
Breaking down the probabilities of events

Sample space, events, and probabilities

- $A_i = \text{positions on chromosome } i$
  - $P(A_1) = .1 \quad P(A_2) = .2 \quad P(A_3) = .3 \quad P(A_4) = .4$

- $B = \text{positions in genes}$
  - $P(B|A_1) = .05 \quad P(B|A_2) = .06 \quad P(B|A_3) = .07 \quad P(B|A_4) = .08$

Venn diagram with probabilities

Fill in top row with $P(B \cap A_i) = P(B|A_i)P(A_i)$, and fill in column totals $P(A_i)$.

<table>
<thead>
<tr>
<th>$B$</th>
<th>$A_1$</th>
<th>$A_2$</th>
<th>$A_3$</th>
<th>$A_4$</th>
<th>Total $P(B)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B$</td>
<td>(.05)(.1)</td>
<td>(.06)(.2)</td>
<td>(.07)(.3)</td>
<td>(.08)(.4)</td>
<td>$P(B)$</td>
</tr>
<tr>
<td></td>
<td>= .005</td>
<td>= .012</td>
<td>= .021</td>
<td>= .032</td>
<td></td>
</tr>
<tr>
<td>$B^c$</td>
<td>$P(B^c \cap A_1)$</td>
<td>$P(B^c \cap A_2)$</td>
<td>$P(B^c \cap A_3)$</td>
<td>$P(B^c \cap A_4)$</td>
<td>$P(B^c)$</td>
</tr>
<tr>
<td>Total</td>
<td>.1</td>
<td>.2</td>
<td>.3</td>
<td>.4</td>
<td>1</td>
</tr>
</tbody>
</table>
Breaking down the probabilities of events

Sample space, events, and probabilities

- $A_i =$ positions on chromosome $i$
  
  
  
  
  
  

  \[
P(A_1) = .1 \quad P(A_2) = .2 \quad P(A_3) = .3 \quad P(A_4) = .4
  \]

- $B =$ positions in genes
  
  \[
P(B|A_1) = .05 \quad P(B|A_2) = .06 \quad P(B|A_3) = .07 \quad P(B|A_4) = .08
  \]

Venn diagram with probabilities

Fill in rest of table to complete the row and column totals.

<table>
<thead>
<tr>
<th></th>
<th>$A_1$</th>
<th>$A_2$</th>
<th>$A_3$</th>
<th>$A_4$</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B$</td>
<td>.005</td>
<td>.012</td>
<td>.021</td>
<td>.032</td>
<td>.070</td>
</tr>
<tr>
<td>$B^c$</td>
<td>.095</td>
<td>.188</td>
<td>.279</td>
<td>.368</td>
<td>.930</td>
</tr>
<tr>
<td>Total</td>
<td>.1</td>
<td>.2</td>
<td>.3</td>
<td>.4</td>
<td>1</td>
</tr>
</tbody>
</table>

\[
P(B) = P(B \cap A_1) + \cdots + P(B \cap A_4)
\]

\[
= P(B|A_1)P(A_1) + \cdots + P(B|A_4)P(A_4)
\]
### Events and Probabilities

| Event | Probability | Conditional Probability
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(A_1)$</td>
<td>.1</td>
<td>$P(B</td>
</tr>
<tr>
<td>$P(A_2)$</td>
<td>.2</td>
<td>$P(B</td>
</tr>
<tr>
<td>$P(A_3)$</td>
<td>.3</td>
<td>$P(B</td>
</tr>
<tr>
<td>$P(A_4)$</td>
<td>.4</td>
<td>$P(B</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Event</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B$</td>
<td>.005</td>
</tr>
<tr>
<td>$B^c$</td>
<td>.095</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$A_1$</th>
<th>$A_2$</th>
<th>$A_3$</th>
<th>$A_4$</th>
<th><strong>Total</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(B)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>.070</strong></td>
</tr>
<tr>
<td></td>
<td>.012</td>
<td>.021</td>
<td>.032</td>
<td>.368</td>
<td><strong>.930</strong></td>
</tr>
<tr>
<td>$P(A_1)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>.1</strong></td>
</tr>
<tr>
<td>$P(A_2)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>.2</strong></td>
</tr>
<tr>
<td>$P(A_3)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>.3</strong></td>
</tr>
<tr>
<td>$P(A_4)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>.4</strong></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>.458</strong></td>
<td><strong>.523</strong></td>
<td><strong>.588</strong></td>
<td><strong>.653</strong></td>
<td><strong>1</strong></td>
</tr>
</tbody>
</table>

- Over the whole genome, what fraction of positions are in genes? 
  \[ P(B) = .070 \]
- How many positions are in genes? 
  \[ (10,000,000) \cdot (.070) = 700,000 \]
- If site $x$ is in a gene, what’s the probability $x$ is on chromosome $i$?
  - $P(A_1|B) = \frac{P(B|A_1)P(A_1)}{P(B)} = \frac{(.05)(.10)}{.070} \approx .0714$
  - $P(A_2|B) = \frac{(.06)(.20)}{.070} \approx .1714$
  - $P(A_3|B) = \frac{(.07)(.30)}{.070} \approx .3000$
  - $P(A_4|B) = \frac{(.08)(.40)}{.070} \approx .4571$
### Full version of Bayes’ Theorem

#### Definition (Partition of $S$)

Events $A_1, \ldots, A_n$ partition the sample space $S$ when

- $S = A_1 \cup \cdots \cup A_n$.
- $A_i \cap A_j = \emptyset$ for $i \neq j$. \textit{(pairwise mutually exclusive)}
- $P(A_i) > 0$ for all $i$.

In other words, $A_1, \ldots, A_n$ are all nonempty with positive probability, and every element of the sample space is in exactly one of them.

#### Theorem (Bayes’ Theorem)

\textit{Let} $A_1, \ldots, A_n$ \textit{be mutually exclusive events that partition sample space $S$, and $B$ be any event on $S$. Then}

- $P(B) = \sum_{i=1}^{n} P(B|A_i)P(A_i)$
- \textit{If} $P(B) > 0$ \textit{then for each} $j = 1, \ldots, n$,

$$P(A_j|B) = \frac{P(B|A_j)P(A_j)}{P(B)} = \frac{P(B|A_j)P(A_j)}{\sum_{i=1}^{n} P(B|A_i)P(A_i)}$$
Independence (2.5)
**Independence**

Events $A$ and $B$ are independent when

$$P(A \cap B) = P(A)P(B)$$

**Derivation from conditional probability**

$A$ and $B$ are independent when knowledge of one event doesn’t affect the probability of the other event:

$$P(A|B) = P(A) \iff \frac{P(A \cap B)}{P(B)} = P(A) \iff P(A \cap B) = P(A)P(B)$$
Independence examples

Rolling two dice (red and green)

- $P(\text{red} = 1) = 1/6$
- $P(\text{green} = 2) = 1/6$
- $P(\text{red} = 1 \text{ and green} = 2) = (1/6)(1/6) = 1/36$
- The two rolls are independent.

Dealing cards

- Draw two cards $X, Y$ from a standard 52 card deck.
- $P(\text{$X$ is red}) = 1/2$
- $P(\text{$Y$ is red}) = 1/2$
- $P(\text{$X$ is red and $Y$ is red}) =$
  
  $P(\text{$X$ is red | $Y$ is red})P(\text{$Y$ is red}) = (25/51)(1/2) = \frac{25}{102}$
- This doesn’t equal $(1/2)(1/2) = 1/4$, so the cards are dependent.
Independence for multiple events

Rolling two dice (red and green)

\[ A = \text{“red is even”} \quad P(A) = 1/2 \]
\[ B = \text{“green is even”} \quad P(B) = 1/2 \]
\[ C = \text{“red+green is even”} \quad P(C) = 1/2 \]

- Any two of the above imply the third, so they are not independent.
- We need a way to check this.
Independence for multiple events

Rolling two dice (red and green)

- $A = \text{“red is even”}$, $B = \text{“green is even”}$, $C = \text{“red+green is even”}$

- $S = \{(r, g) : r = 1, \ldots, 6 \text{ and } g = 1, \ldots, 6\}$

- $A \cap B = \{(r, g) : r = 2, 4, 6 \text{ and } g = 2, 4, 6\}$

- $P(A \cap B) = \frac{3^2}{6^2} = \frac{9}{36} = \frac{1}{4}$

- $P(A)P(B) = \left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \frac{1}{4}$ so $A$ and $B$ are independent.

- $A \cap B = A \cap C = B \cap C = \{(r, g) : r = 2, 4, 6 \text{ and } g = 2, 4, 6\}$

Likewise, $A$ and $C$ are independent, and $B$ and $C$ are independent.

- **Three-way intersection:**

  $A \cap B \cap C = \{(r, g) : r = 2, 4, 6 \text{ and } g = 2, 4, 6\}$

  $P(A \cap B \cap C) = \frac{1}{4} \neq P(A)P(B)P(C) = \left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \frac{1}{8}$

$A, B, C$ are dependent.
Independence for multiple events

Events $A_1, A_2, \ldots, A_n$ are independent if all combinations of them have multiplicative probabilities:

- **All pairs:** $P(A_i \cap A_j) = P(A_i)P(A_j)$ \hspace{1cm} $i, j$ distinct

- **All triples:** $P(A_i \cap A_j \cap A_k) = P(A_i)P(A_j)P(A_k)$ \hspace{1cm} $i, j, k$ distinct

**All 4-way, All 5-way, \ldots, All $n$-way**

If any of the above equations fail to hold, then $A_1, A_2, \ldots, A_n$ are dependent.
Venn diagram of independence

- Event $A$ is split into $A \cap B$ and $A \cap B^c$.
- If $A$ and $B$ are independent, then
  \[
  P(A \cap B^c) = P(A) - P(A \cap B)
  \]
  \[
  = P(A) - P(A)P(B) = P(A)(1 - P(B)) = P(A)P(B^c)
  \]

$A$ and $B$ are independent iff all regions of the Venn diagram have multiplicative probabilities

\[
\begin{align*}
P(A \cap B^c) &= P(A)P(B^c) \\
P(A^c \cap B) &= P(A^c)P(B) \\
P(A \cap B) &= P(A)P(B) \\
P(A^c \cap B^c) &= P(A^c)P(B^c)
\end{align*}
\]
A, B, C are independent iff all 8 regions follow the multiplication rule; e.g., for the region indicated,

\[ P(A^c \cap B \cap C) = P(A^c)P(B)P(C) \]

For a Venn diagram on \( n \) sets, the sets are independent iff all \( 2^n \) regions obey the multiplication rule.
Independent vs. mutually exclusive

- **A, B, C independent:**
  - Full Venn diagram with intersecting sets.
  - Intersections and Venn diagram regions have probabilities given by the multiplication formulas.

- **A, B, C mutually exclusive:** No overlaps between sets.
Repeated independent trials

Repeat an experiment over and over, with all trials independent.

**Roll a die over and over**

The probabilities of the values of the rolls are not influenced by previous rolls, so they are independent.

**Draw cards from a deck without replacement**

The card values are influenced by previous draws, so they are not independent.
Roll a die 10 times

**Probability of at least one 3**

The rolls are $R_1, R_2, \ldots, R_{10}$.

\[
P(\text{rolling at least one 3}) = 1 - P(\text{no 3's})
\]

\[
P(\text{no 3's}) = P(R_1 \neq 3)P(R_2 \neq 3) \cdots P(R_{10} \neq 3) = (5/6)^{10}
\]

\[
P(\text{rolling at least one 3}) = 1 - (5/6)^{10}
\]

**Probability of rolling exactly one 3**

\[
P(\text{roll exactly one 3}) = \sum_{i=1}^{10} P(R_i = 3, \text{others } \neq 3)
\]

\[
= \sum_{i=1}^{10} (1/6)(5/6)^9 = 10(1/6)(5/6)^9
\]
Review of geometric series

Geometric series

\[ a + ar + ar^2 + ar^3 + \cdots = \sum_{i=0}^{\infty} ar^i = \frac{a}{1 - r} \]

where \( a \) is the initial term and \( r \) is the ratio, with \(|r| < 1\).
A solitaire game

The Rules

Roll a die repeatedly.
- Win if it shows 3.
- Lose if it shows 4.
- Try again otherwise.

Events

- What are the probabilities of winning; losing; and playing forever without winning or losing?
- Events: $A =$ “win”, $B =$ “lose”, $C =$ “play forever”.
- $A, B, C$ are mutually exclusive and $C = (A \cup B)^c$. 
Probability of winning

- \( A = \text{“win”} = A_1 \cup A_2 \cup A_3 \cup \cdots = \bigcup_{k=1}^{\infty} A_k \)
  - where \( A_k \) is the event that you win on the \( k \)th roll.

- To win on the \( k \)th roll,
  - each of the first \( k - 1 \) rolls must be one of 1, 2, 5, or 6,
  - and the \( k \)th roll must be 3.

- \( P(A_k) = \left(\frac{4}{6}\right)^{k-1}\left(\frac{1}{6}\right) \)

- \( P(A) = \sum_{k=1}^{\infty} P(A_k) = \sum_{k=1}^{\infty} \left(\frac{4}{6}\right)^{k-1}\left(\frac{1}{6}\right) \)

  - Geometric series:
    - First term (plug in \( k = 1 \)): \( \left(\frac{4}{6}\right)^0\left(\frac{1}{6}\right) = \frac{1}{6} \)
    - Ratio: \( \frac{4}{6} \)
    - Sum: \( P(A) = \frac{\frac{1}{6}}{1-\left(\frac{4}{6}\right)} = \frac{\frac{1}{6}}{\frac{2}{6}} = \frac{1}{2} \)

- Probability of losing is similarly computed as \( P(B) = \frac{1}{2} \).

- Probability of never winning or losing:
  - \( C = (A \cup B)^c \)
  - \( P(C) = 1 - P(A) - P(B) = 1 - \left(\frac{1}{2}\right) - \left(\frac{1}{2}\right) = 0 \).