CMU 15-381
Lecture 7: Probability

Teachers:
Emma Brunskill
Ariel Procaccia (this time)
Gambling 101

- You choose a die first, I choose second
- We both throw; higher number wins
- Which die would you choose?

A
\[
\begin{array}{ccc}
5 & 5 & 5 \\
5 & & 5 \\
5 & & \\
\end{array}
\]

B
\[
\begin{array}{ccc}
7 & 1 & 7 \\
7 & & \\
7 & & \\
\end{array}
\]

C
\[
\begin{array}{ccc}
8 & 2 & 8 \\
2 & & 2 \\
2 & & \\
\end{array}
\]

D
\[
\begin{array}{ccc}
3 & 3 & 3 \\
3 & & 9 \\
3 & & 9 \\
\end{array}
\]
GAMBLING 101

• Antoine Gombaud (1607-1684) made history for being a loser

  I will roll a die four times; I win if I get a 1

• After a while no one would take the bet

• \[1 - \left(\frac{5}{6}\right)^4 = 0.518\]
GAMBLING 101

• Gombaud invented a new scam:

  I will roll two dice 24 times; I win if I get a double 1

• Why was he losing money?

  \[1 - \left(\frac{35}{36}\right)^{24} = 0.491\]

• Gombaud wrote to Pascal and Fermat, who subsequently created probability theory
GAMBLING 101

Morale of the story:
Analyzing gambling is not a side-benefit of probability theory; probability theory was created to analyze gambling!
PENNIES AND GOLD

- Three bags contain two gold coins, two pennies, and one of each
- Bag is chosen at random, and one coin from it is selected at random; the coin is gold
- **Poll 1:** What is the probability that the other coin is gold?
  1. 1/6
  2. 1/3
  3. 2/3
  4. 1
LANGUAGE OF PROBABILITY

Probability can be counterintuitive; we need a formal language!
The sample space is a finite set of elements $S$

A probability distribution $p$ assigns a non-negative real probability to each element, such that

$$\sum_{x \in S} p(x) = 1$$
**Language of Probability**

- An event is a subset $E \subseteq S$
- $\Pr[E] = \sum_{x \in E} p(x)$
- If each element $x \in S$ has equal probability, the distribution is uniform:
  \[
  \Pr[E] = \sum_{x \in E} p(x) = \frac{|E|}{|S|}
  \]
LANGUAGE OF PROBABILITY

• We roll a white die and black die

• \( S = \{(1,1), (1,2), (1,3), (1,4), (1,5), (1,6), (2,1), (2,2), (2,3), (2,4), (2,5), (2,6), (3,1), (3,2), (3,3), (3,4), (3,5), (3,6), (4,1), (4,2), (4,3), (4,4), (4,5), (4,6), (5,1), (5,2), (5,3), (5,4), (5,5), (5,6), (6,1), (6,2), (6,3), (6,4), (6,5), (6,6)\}\)

• **Poll 2:** Probability that the sum is 7 or 11?
  1. 1/9
  2. 2/9
  3. 3/9
  4. 4/9
**Conditional probability**

- The probability of event $A$ given event $B$ is defined as
  \[
  \Pr[A|B] = \frac{\Pr[A \cap B]}{\Pr[B]}
  \]

- Think of it as the proportion of $A \cap B$ to $B$
Pennies and Gold, Revisited

- Three bags contain two gold coins, two pennies, and one of each.
- Bag is chosen at random, and one coin from it is selected at random; the coin is gold.
- \( G_i \): coin \( i \in \{1, 2\} \) is gold.
- \( \Pr[G_1] = \frac{1}{2}, \Pr[G_1 \cap G_2] = \frac{1}{3} \)
- \( \Pr[G_2 | G_1] = \frac{1/3}{1/2} = \frac{2}{3} \)
The Chain Rule

- For $A$ and $B$ to occur, $B$ must occur, and $A$ must occur given that $B$ occurred; formally:

\[
\Pr[A \cap B] = \Pr[B] \times \Pr[A|B]
\]

- Applying iteratively:
  \[
  \Pr[A_1 \cap \cdots \cap A_n] = \Pr[A_1] \times \Pr[A_2|A_1] \times \cdots \Pr[A_n|A_1,\ldots,A_{n-1}]
  \]
BAYES’ RULE

- $\Pr[B] \times \Pr[A|B] = \Pr[A \cap B] = \Pr[A] \times \Pr[B|A]$
- Therefore,

Bayes’ Rule:

$$\Pr[A|B] = \frac{\Pr[A] \Pr[B|A]}{\Pr[B]}$$
Monty Hall problem

- Announcer hides prize behind one of three doors at random
- You choose a door
- Announcer opens a door with no prize
- Should you stay with your choice or switch?
Monty Hall Problem

• Choose door 1, door 2 opens

\[ \Pr[P_3 | O_2] = \frac{\Pr[P_3] \Pr[O_2 | P_3]}{\Pr[O_2]} \]

• \( \Pr[P_3] = \frac{1}{3} \), \( \Pr[O_2 | P_3] = 1 \),
  \( \Pr[O_2] = 1/2 \)

• Therefore, \( \Pr[P_3 | O_2] = 2/3 \)

• Poll 3: Assuming there are five doors, what is the probability of winning when switching?
  
  1. 3/15
  2. 4/15
  3. 5/15
  4. 6/15
INDEPENDENCE

- Events $A$ and $B$ are independent if and only if $\Pr[A|B] = \Pr[A]$.
- Poll 4: Which of the following events are independent when rolling black die and white die?
  1. Black die is 1, white die is 1
  2. Sum is 2, sum is 3
  3. Black die is 1, product is 2
  4. Black die is 1, sum is 2
THE BIRTHDAY PARADOX

- $m$ people in a room; suppose all birthdays are equally likely (excluding Feb 29); what is the probability that two people have the same birthday?
- $S = \{1, \ldots, 365\}^m$, sample $\tilde{x} = (x_1, \ldots x_m)$
- $E = \{\tilde{x} \in S \mid \exists i, j, \text{s.t. } x_i = x_j\}$
- Let $A_i$ be the event that person $i$’s birthday differs from the birthdays of $1, \ldots, i - 1$
- $\bar{E} = A_1 \cap \cdots \cap A_m$
- Using the chain rule:
  \[
  \Pr[\bar{E}] = \Pr[A_1] \times \Pr[A_2|A_1] \times \cdots \Pr[A_m|A_1, \ldots, A_{m-1}]
  \]
THE BIRTHDAY PARADOX

• $A_1 \cap \cdots \cap A_{i-1}$ means first $i - 1$ students had different birthdays

• $i - 1$ out of 365 occupied when $i$th birthday is chosen

• $\Pr[A_i|A_1, \ldots, A_{i-1}] = \frac{365-(i-1)}{365} = 1 - \frac{i-1}{365}$

• $\Pr[\bar{E}] = 1 \times \left(1 - \frac{1}{365}\right) \times \cdots \times \left(1 - \frac{m-1}{365}\right)$

• $\Pr[E] = 1 - \Pr[\bar{E}]$
THE BIRTHDAY PARADOX

Pr[E]: prob. of matching bdays
Pr[\bar{E}]: prob. of no matching bdays

Number of people
Probability
0.0
0.2
0.4
0.6
0.8
1.0
0 10 20 30 40 50 60 70 80 90

~23
The birthday paradox

- **Poll 5:** What is the probability that two people have the same birthday if there are 730 people?
  1. 1/2
  2. 0.75
  3. $0.99999999999997$
  4. 1
**Joint Probability Distribution**

<table>
<thead>
<tr>
<th></th>
<th>Car 1</th>
<th>Car 2</th>
<th>Car 3</th>
<th>Truck</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red</td>
<td>0.05</td>
<td>0.2</td>
<td>0</td>
<td>0.1</td>
</tr>
<tr>
<td>Blue</td>
<td>0.1</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
</tr>
<tr>
<td>Yellow</td>
<td>0</td>
<td>0.1</td>
<td>0.05</td>
<td>0.1</td>
</tr>
<tr>
<td>Black</td>
<td>0.1</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
</tr>
</tbody>
</table>
The Sum Rule

• How do we answer the question “What is the probability of a red car of any type?”

The Sum Rule:

\[ \Pr[X] = \sum_Y \Pr[X \cap Y] \]

• \( \Pr[X] \) is known as the marginal probability of \( X \)
### Joint Probabilities:

<table>
<thead>
<tr>
<th></th>
<th>Red</th>
<th>Blue</th>
<th>Yellow</th>
<th>Black</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red</td>
<td>0.05</td>
<td>0.2</td>
<td>0</td>
<td>0.1</td>
</tr>
<tr>
<td>Blue</td>
<td>0.1</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
</tr>
<tr>
<td>Yellow</td>
<td>0</td>
<td>0.1</td>
<td>0.05</td>
<td>0.1</td>
</tr>
<tr>
<td>Black</td>
<td>0.1</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
</tr>
</tbody>
</table>

### Probability of color given type:

<table>
<thead>
<tr>
<th></th>
<th>Red</th>
<th>Blue</th>
<th>Yellow</th>
<th>Black</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red</td>
<td>0.2</td>
<td>0.667</td>
<td>0</td>
<td>0.5</td>
</tr>
<tr>
<td>Blue</td>
<td>0.4</td>
<td>0</td>
<td>0.4</td>
<td>0</td>
</tr>
<tr>
<td>Yellow</td>
<td>0</td>
<td>0.333</td>
<td>0.2</td>
<td>0.5</td>
</tr>
<tr>
<td>Black</td>
<td>0.4</td>
<td>0</td>
<td>0.4</td>
<td>0</td>
</tr>
</tbody>
</table>
SUMMARY

• Definitions / facts
  o Language of probability
  o Conditional probability
  o Independence

• Three useful rules:
  o Chain Rule
  o Bayes’ Rule
  o Sum Rule