Review

- Probabilities
  - Unconditional: \( P(X) \)
  - Conditional: \( P(X|Y) \)
  - Joint: \( P(X,Y) \)
- Conditional independence
  \[
  \begin{align*}
  P(X|Y) &= P(X) \\
  P(Y|X) &= P(Y) \\
  P(X,Y) &= P(X)P(Y)
  \end{align*}
  \]
  equivalent

- Rules
  \[
  P(A,B,C,...) = P(A)P(B|A)P(C|A,B) \text{ (product rule)}
  \]

- Bayes' rule
  \[
  P(X|Y) = \frac{P(Y|X)P(X)}{P(Y)}
  \]

- Marginalization
  \[
  P(X) = \sum_{Y}P(X,Y=y)
  \]

Probabilistic inference

Do probabilities capture patterns of commonsense reasoning?

Today — reasoning about:

1) multiple explanations of a single event
2) multiple events with a single explanation
3) intervening events
Binary random variables

B = burglary
E = earthquake
A = alarm

Joint distribution

\[ P(B, E, A) = P(B) \ P(E|B) \ P(A|B, E) \]

Domain knowledge

\[ P(B=1) = 0.001 \]
\[ P(E=1|B=0) = 0.002 \]
\[ P(E=1|B=1) = 0.002 \]
\[ P(E=1|B)= P(E=1) = 0.002 \]

<table>
<thead>
<tr>
<th></th>
<th>E = 0</th>
<th>E = 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>B = 0</td>
<td>0</td>
<td>0.001</td>
</tr>
<tr>
<td>B = 1</td>
<td>0.29</td>
<td>0.94</td>
</tr>
<tr>
<td>B = 1</td>
<td>0.95</td>
<td></td>
</tr>
</tbody>
</table>

1. Reasoning about multiple explanations

\[ P(B=1|A=1) = ? \]

\[ P(B=1|A=1, E=1) = ? \]

\[ P(B=1|A=1, E=1) \rightarrow 0.387 \]

\[ P(B=1|A=1) \rightarrow 0.0033 \]

Bayes' rule:

\[ P(B=1|A=1) = \frac{P(A=1|B=1) \ P(B=1)}{P(A=1)} \]

\[ P(A=1) = \sum \ P(A=1, E=e, B=b) = \sum \ P(B=b) \ P(E=e|B=b) \ P(A=1|B=b, E=e) \]

Term in numerator

\[ = P(B=0) \ P(E=0) \ P(A=1|B=0, E=0) + \]
\[ P(B=0) \ P(E=1) \ P(A=1|B=0, E=1) + \]
\[ P(B=1) \ P(E=0) \ P(A=1|B=1, E=0) + \]
\[ P(B=1) \ P(E=1) \ P(A=1|B=1, E=1) \]

\[ = (1-0.001) \ (1-0.002) \ (0.001) + \]
\[ (1-0.001) \ (0.002) \ (0.29) + \]
\[ (0.001) \ (1-0.002) \ (0.94) + \]
\[ (0.001) \ (0.002) \ (0.95) = 0.00252 \]
Term in numerator:

\[ P(A=1 \mid B=1) = \sum_{e \in \{0,1\}} P(A=1, E=e \mid B=1) \]

\[ = \sum_{e} P(A=1 \mid E=e, B=1) P(E=e \mid B=1) \]

\[ = P(A=1 \mid E=0, B=1) P(E=0) \]

\[ + P(A=1 \mid E=1, B=1) P(E=1) \]

\[ = (0.94) (1-0.002) + (0.95) (0.002) \]

\[ = 0.94002 \]

So...

\[ P(B=1 \mid A=1) = \frac{P(A=1 \mid B=1) P(B=1)}{P(A=1)} = \frac{(0.94002)(0.001)}{(0.00252)} = 0.37 \]

Conditionalized form of Bayes rule:

\[ P(B=1 \mid A=1, E=1) = \frac{P(A=1 \mid B=1, E=1) P(B=1 \mid E=1)}{P(A=1 \mid E=1)} \]

\[ = \frac{(0.94002)(0.001)}{(0.00252)} = 0.0033 \]

Numerator:

\[ P(B=1 \mid E=1) = P(B=1) = 0.001 \]

Denominator:

\[ P(A=1 \mid E=1) = \sum_{b} P(A=1, B=b \mid E=1) \]

\[ = \sum_{b \in \{0,1\}} P(A=1 \mid B=b, E=1) P(B=b \mid E=1) \]

\[ = P(A=1 \mid B=0, E=1) P(B=0) + P(A=1 \mid B=1, E=1) P(B=1) \]

\[ = (0.29)(1-0.001) + (0.95)(0.001) \]

\[ = 0.29 \]

Summary:

\[ P(B=1) = 0.001 \]

\[ P(B=1 \mid A=1) = 0.37 \]

\[ P(B=1 \mid A=1, E=1) = 0.0033 \]

\[ \Rightarrow \text{Earthquake "explains away" the alarm, weakening our belief in the burglary.} \]
"Explaining away" is an example of non-monotonic reasoning.

\[ P(B=1) < P(B=1 \mid A=1) \]
\[ P(B=1 \mid A=1) > P(B=1 \mid A=1, E=1) \]

Arises from multiple (causal) explanations of an observed event.

2) Multiple events with a common explanation
   
   More random variables
   - \( J = \text{John calls} \)
   - \( M = \text{Mary calls} \)

   Conditional independence assumptions
   
   Already: \( P(B \mid E) = P(B) \)
   
   Also: \( P(J \mid A) = P(J \mid A, B, E) \)
   \( P(M \mid A) = P(M \mid A, J, B, E) \)

   Joint distribution
   
   \[ P(B, E, A, J, M) = P(B) P(E \mid B) P(A \mid B, E) P(J \mid B, E, A) P(M \mid B, E, A, J) \]

   \[ = P(B) \underbrace{P(E) P(A \mid B) P(J \mid A)}_{\text{conditional independence assumptions}} P(M \mid A) \]

   Conditional probabilities
   
   \( P(J=1 \mid A=0) = 0.05 \)
   \( P(J=1 \mid A=1) = 0.9 \)
   \( P(M=1 \mid A=0) = 0.01 \)
   \( P(M=1 \mid A=1) = 0.7 \)

   Compare \( P(A=1) = 0.00252 \)
   \( P(A=1 \mid J=1) = ? \rightarrow 0.0435 \) \( \uparrow \) non-monotonic
   \( P(A=1 \mid J=1, M=0) = ? \rightarrow 0.0136 \) \( \downarrow \)
Bayes rule:

\[ P(A=1 | J=1) = \frac{P(J=1 | A=1) P(A=1)}{P(J=1)} = 0.435 \]

Denominator:

\[ P(J=1) = \sum_a P(A=a, J=1) \text{ marginalization} \]

\[ = \sum_a P(J=1 | A=a) P(A=a) \text{ product rule} \]

\[ = P(J=1 | A=0) P(A=0) + P(J=1 | A=1) P(A=1) \]

\[ = (0.05) (1 - 0.00252) + (0.9) (0.00252) \]

\[ = 0.0521 \]

Bayes rules (with multiple pieces of evidence)

\[ P(A=1 | J=1, M=0) = \frac{P(J=1, M=0 | A=1) P(A=1)}{P(J=1, M=0)} \]

\[ = \frac{P(J=1 | A=1) P(M=0 | A=1) P(A=1)}{P(J=1, M=0)} \]

\[ = 0.0136 \]

Denominator:

\[ P(J=1, M=0) = \sum_a P(...) \]

\[ = 0.05 \]

3) Reasoning about intervening events

Compare:

\[ P(A=1) = 0.00252 \]

\[ P(A=1 | J=1) = 0.435 \]

\[ P(A=1 | J=1, B=1) = P(J=1 | A=1, B=1) P(A=1 | B=1) \]

\[ = 0.9965 \]

Conditionalized Bayes rule:

\[ P(J=1 | B=1) \]

\[ = 0.849 \]

Conditional marginalization