Probability Rules, Factors, and Factor Operations

Daniel Lowd

Conditional Probability Rules

• Conditional probability: 
  \[ P(X|Y,Z) = \frac{P(X,Y|Z)}{P(Y|Z)} \]

• Product rule: 
  \[ P(X,Y|Z) = P(X|Y,Z)P(Y|Z) \]

• Bayes’ rule: 
  \[ P(X|Y,Z) = \frac{P(Y|X,Z)P(X|Z)}{P(Y|Z)} \]

• Chain rule: 
  \[ P(X_1,X_2,\ldots,X_n) = \prod_{i=1}^{n} P(X_i|X_{i-1},\ldots,X_1) \]
  \[ = P(X_1)P(X_2|X_1)P(X_3|X_1,X_2)\ldots \]

Probability Summation Rules

• Summing out: 
  \[ P(X|Z) = \sum_{y} P(y,X|Z) \]

• Law of total probability: 
  \[ P(X|Z) = \sum_{y} P(X|y,Z)P(y|Z) \]

• Probabilities sum to one: 
  \[ P(\Omega|Z) = \sum_{x} P(x|Z) = 1 \]

Probability Independence Rules

• When \( X \perp Y|Z \)
  (X and Y are conditionally independent given Z):
  \[ P(X|Y,Z) = P(X|Z) \]
  \[ P(Y|X,Z) = P(Y|Z) \]
  \[ P(X,Y|Z) = P(X|Z)P(Y|Z) \]

• When groups of variables are independent from each other, so are subsets of those groups:
  \[ X \perp Y, W|Z \Rightarrow X \perp Y|Z \]

Joint Distribution

• Intelligence (I)
  – \( I \) (low), \( I' \) (high)

• Difficulty (D)
  – \( D' \) (easy), \( D' \) (hard)

• Grade (G)
  – \( G' \) (A), \( G' \) (B), \( G' \) (C)

Conditioning: Reduction

Condition on \( G' \)
Conditioning: Renormalization

<table>
<thead>
<tr>
<th>I</th>
<th>D</th>
<th>G</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>i0</td>
<td>d0</td>
<td>g1</td>
<td>0.126</td>
</tr>
<tr>
<td>i0</td>
<td>d1</td>
<td>g1</td>
<td>0.009</td>
</tr>
<tr>
<td>i1</td>
<td>d0</td>
<td>g1</td>
<td>0.252</td>
</tr>
<tr>
<td>i1</td>
<td>d1</td>
<td>g1</td>
<td>0.06</td>
</tr>
</tbody>
</table>

Sum: 0.447

Marginalization

Marginalize I to compute \( P(D|g^1) \):

<table>
<thead>
<tr>
<th>D</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>d0</td>
<td>0.846</td>
</tr>
<tr>
<td>d1</td>
<td>0.154</td>
</tr>
</tbody>
</table>

Factors

- A factor \( \phi(X_1,\ldots,X_k) \) is a function \( \phi : \text{Val}(X_1,\ldots,X_k) \to R \)
- Scope = \{X_1, \ldots, X_k\}

Example Factors

<table>
<thead>
<tr>
<th>I</th>
<th>D</th>
<th>G</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>i0</td>
<td>d0</td>
<td>g1</td>
<td>0.126</td>
</tr>
<tr>
<td>i0</td>
<td>d0</td>
<td>g2</td>
<td>0.168</td>
</tr>
<tr>
<td>i1</td>
<td>d1</td>
<td>g1</td>
<td>0.252</td>
</tr>
<tr>
<td>i1</td>
<td>d1</td>
<td>g2</td>
<td>0.06</td>
</tr>
</tbody>
</table>

Factor Product

\[ \begin{array}{c|c|c}
A & B & \text{Prob.} \\
\hline
a^0 & b^0 & 0.5 \\
&a^0 & b^1 & 0.8 \\
&a^1 & b^0 & 0.1 \\
&a^1 & b^1 & 0.3 \\
&a^1 & b^2 & 0.9 \\
\end{array} \]

\[ \begin{array}{c|c|c}
B & C & \text{Prob.} \\
\hline
b^0 & c^0 & 0.5 \\
&b^1 & c^0 & 0.7 \\
&b^0 & c^1 & 0.1 \\
&b^0 & c^2 & 0.2 \\
\end{array} \]