From last lecture, we got this result: \( p(x_v) = \prod_{i=1}^{n} P(x_i|x_{\pi_i}). \)

Now we take a look at the joint probability of a six-node example:

\[
p(x_1, x_2, x_3, x_4, x_5, x_6) = p(x_1)p(x_2|x_1)p(x_3|x_1)p(x_4|x_2)p(x_5|x_3)p(x_6|x_5, x_2)
\]

Or by "chain rule" and the property of Markov-chain, we can get the same result:

\[
p(x_1, x_2, x_3, x_4, x_5, x_6) \\
= p(x_1)p(x_2|x_1)p(x_3|x_2, x_1)p(x_4|x_3, x_2, x_1)p(x_5|x_4, x_3, x_2, x_1)p(x_6|x_5, x_4, x_3, x_2, x_1) \\
= p(x_1)p(x_2|x_1)p(x_3|x_1)p(x_4|x_2)p(x_5|x_3)p(x_6|x_5, x_2)
\]
0.1 Independence

- Marginal independence

\[ x_A \perp x_B : p(x_A, x_B) = p(x_A)p(x_B) = p(x_A|x_B) = p(x_B) \]

- Conditional independence

\[ x_A \perp x_B|x_C : \]

\[ p(x_A, x_B|x_C) = p(x_A|x_C)p(x_B|x_C) = p(x_A|x_B, x_C) = p(x_A|x_C) \]

- Aside: before we move on further, we first define the following terms:

1. I is defined as a set of ordering nodes in a graph C;
2. For each \( i \in V \), \( V_i \) is defined as a set of all nodes that appear earlier than \( i \) excluding \( \pi_i \).

For example, for the case of the six-noded figure given above,

\[ I = \{1, 2, 3, 4, 5, 6\}, V_3 = \{2\}, V_6 = \{1, 3, 4\} \]

- Look at the conditional probabilities \( x_i \perp x_v|x_x \), in the six-noded example:

\[ x_1 \perp \phi | \phi, x_2 \perp \phi | x_1, x_3 \perp x_2 | x_1 \]

\[ x_4 \perp \{x_1, x_3\} | x_2, \]

\[ x_5 \perp \{x_1, x_2, x_4\} | x_3, \]

\[ x_6 \perp \{x_1, x_3, x_4\} | \{x_2, x_5\} \]
• For the purpose of illustration, here we want to show

\[ p(x_4|x_1, x_2, x_3) = p(x_4|x_2) \]

Proof: first, we know

\[ p(x_1, x_2, x_3, x_4, x_5, x_6) \]

\[ = p(x_1)p(x_2|x_1)p(x_3|x_1)p(x_4|x_2)p(x_5|x_3)p(x_6|x_5, x_2) \]

then

\[
p(x_4|x_1, x_2, x_3) = \frac{p(x_1, x_2, x_3, x_4)}{p(x_1, x_2, x_3)} \sum_{x_5} \sum_{x_6} p(x_1, x_2, x_3, x_4, x_5, x_6) \\
= \frac{p(x_1)p(x_2|x_1)p(x_3|x_1)p(x_4|x_2)}{p(x_1)p(x_2|x_1)p(x_3|x_1)} \\
= p(x_4|x_2)
\]

• Other conditional independence

0.2 Bayes Ball

There are three canonical graphs. Now let us first look at three three-noded graphs:

• Case 1: markov chain: \( x \perp Z \mid Y \)
The proof of the independence:

\[
p(z|x, y) = \frac{p(x, y, z)}{p(x, y)} = \frac{p(x)p(y|x)p(z|y)}{p(x)p(y|x)} = p(z|y)
\]

Where

\[
p(x, y) = \sum_z p(x, y, z) = \sum_z p(x)p(y|x)p(z|y) = p(x)p(y|x)\sum_z p(z|y) = p(x)p(y|x)
\]

- Case 2: hidden cause: \(X \perp z | y\)

The proof of the independence:
\begin{align*}
    p(z|x, y) &= \frac{p(x, y, z)}{p(x, y)} \\
               &= \frac{p(x)p(y|x)p(z|y)}{p(x)p(y|x)} \\
               &= p(z|y)
\end{align*}

- An example of the situation of hidden case:

The variables of "Shoe size" and "Grey hair" are dependent in some sense, if there is no "Age" in the picture. However, when "Age" is observed, there is no dependence between "Shoe size" and "Grey hair" given that "Age" is known.