1 Bayes’ Nets Representation and Probability

Suppose that a patient can have a symptom ($S$) that can be caused by two different diseases ($A$ and $B$). It is known that the variation of gene $G$ plays a big role in the manifestation of disease $A$. The Bayes’ Net and corresponding conditional probability tables for this situation are shown below. For each part, you may leave your answer as an arithmetic probability expression.

(a) Compute the following entry from the joint distribution:

\[ P(\text{+}g, \text{+}a, \text{+}b, \text{+}s) = P(\text{+}g)P(\text{+}a \mid \text{+}g)P(\text{+}b)P(\text{+}s \mid \text{+}a, \text{+}b) = (0.1)(1.0)(0.4)(1.0) = 0.04 \]

(b) What is the probability that a patient has disease $A$?

\[ P(\text{+}a) = P(\text{+}a \mid \text{+}g)P(\text{+}g) + P(\text{+}a \mid -\text{g})P(-\text{g}) = (1.0)(0.1) + (0.1)(0.9) = 0.19 \]

(c) What is the probability that a patient has disease $A$ given that they have disease $B$?

\[ P(\text{+}a \mid +b) = P(\text{+}a) = 0.19 \] The first equality holds true since $A \perp \perp B$, which can be inferred from the graph of the Bayes’ net.

(d) What is the probability that a patient has disease $A$ given that they have symptom $S$ and disease $B$?

\[ P(\text{+}a \mid +s, +b) = \frac{P(\text{+}a, \text{+}s, +b)}{P(\text{+}a, \text{+}s, +b)} = \frac{P(\text{+}a)p(\text{+}b)p(\text{+}s \mid \text{+}a, \text{+}b)}{P(\text{+}a)p(\text{+}b)p(\text{+}s \mid +a, +b)} = \frac{(0.1)(0.4)(1.0)}{0.076 + 0.2592} \approx 0.2267 \]

(e) What is the probability that a patient has the disease carrying gene variation $G$ given that they have disease $A$?

\[ P(\text{+}g \mid +a) = \frac{P(\text{+}g)p(\text{+}a \mid +g)}{P(\text{+}g)p(\text{+}a \mid +g) + P(-g)p(\text{+}a \mid -g)} = \frac{(0.1)(1.0)}{0.1 + 0.09} = 0.5263 \]
2 Bayes Nets: Representation

Parts (a), (b), and (c) pertain to the following Bayes' Net.

(a) Express the joint probability distribution as a product of terms from the Bayes Nets CPTs.

\[ P(A)P(C|A)P(B|A)P(D|B)P(E)P(F|D,E)P(G|D) \]

(b) Assume each node can take on 4 values. How many entries do the factors at A, D, and F have?

\begin{align*}
A & : 4 \\
D & : 4^2 \\
F & : 4^3
\end{align*}

(c) Mark all that are guaranteed to be true:

\[ \begin{array}{ccc}
\square & B & \perp \perp C \\
\square & A & \perp \perp F \\
\square & D & \perp \perp E|F \\
\square & E & \perp \perp D|A \\
\blacksquare & F & \perp \perp G|D \\
\blacksquare & B & \perp \perp F|D \\
\square & C & \perp \perp G \\
\blacksquare & D & \perp \perp E \\
\end{array} \]

Parts (d) and (e) pertain to the following CPTs.

(d) State all non-conditional independence assumptions that are implied by the probability distribution tables.

From the tables, we have \( A \perp \perp B \) and \( C \perp \perp D \), and \( B \perp \perp C \).

(e) Circle all the Bayes net(s) that can represent a distribution that is consistent with the tables given.

The question asks for Bayes Nets that can represent the distribution in the tables. So, in the nets we circle, the only requirement must be that A and B must not be independent, and C and D must not be independent.
The top left, bottom left, and bottom right nets have arrows between the A-B nodes and the C-D nodes, so we can circle those.

The top middle net has C and D as independent (D is not connected to anything), so we cannot circle it. The bottom middle net has A and B as independent (common effect), so we cannot circle it.

The top right net seems like it could represent the distribution, because D-separation finds that: A and B are not guaranteed to be independent (common cause), and C and D are not guaranteed to be independent (causal chain). However, since $B \perp \perp C$, the arrow between B and C is vacuous, thus A and B cannot actually be dependent.