

axioms

1. $0 \leq P(A) \leq 1$
2. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
3. $P(\Omega) = 1$, $P(\emptyset) = 0$

(joint probability) $(x, y) \in D_x \times D_y$, $P(x, y) = P(X=x \wedge Y=y)$

conditional probability

1. $P(A|B) = \frac{P(A \cap B)}{P(B)}$ assuming $P(B) > 0$
2. $P(A) = P(A|B)P(B) + P(A|B')P(B')$

chain rule

$$P(E_1 \wedge E_2 \wedge E_3 \dots) = \prod_{i=1}^n P(E_i | E_1 \wedge E_2 \dots E_{i-1})$$

→ derived from inductive application of Bayes' rule

independence

A & B are independent $\Leftrightarrow P(A \cap B) = P(A) \times P(B)$

$\Leftrightarrow P(A|B) = P(A)$ knowing B adds no information about A and vice versa

conditional independence

Given event B , event A is conditionally independent of event C if

$$P(A|B, C) = P(A|B)$$

$\Rightarrow A$ & C are related by a shared cause B

\Rightarrow if we take it into account, they become unrelated

Bayesian networks

→ suppose all n variables share a cause s . Then the full joint distribution

$$P(X_1 \wedge X_2 \dots \wedge X_n) = P(X_1|s) \cdot P(X_2|s) \dots \cdot P(X_n|s) \text{ by chain rule}$$

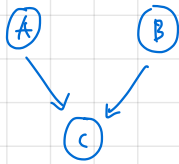
→ joint distributions represented by a graph

→ vertices are random variables, an edge from $X \rightarrow Y$ implies X directly influences Y (some corr. assume X causes Y)

→ then conditional distribution for each node is $P(X | \text{Parents}(X))$

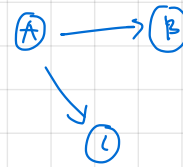
Structures in bayesian networks

independent causes



$$P(A \wedge B \wedge C) = P(A) \cdot P(B) \cdot P(C|A \wedge B)$$

conditionally independent effects



$$P(A \wedge B \wedge C) = P(A) \cdot P(C|A) \cdot P(B|A)$$

causal chain



$$P(A \wedge B \wedge C) = P(A) \cdot P(B|A) \cdot P(C|B)$$

Naive Bayes

↳ suppose we want to infer $P(\text{cause} | \underbrace{E_1 \wedge E_2 \dots}_{\text{percepts}})$ i.e. given these observations, we want to find out likelihood that x caused it

↳ additionally suppose all E_1, \dots, E_n are conditionally independent on cause.

$$1. P(\text{cause} | E_1 \wedge E_2 \dots) = \frac{P(\text{cause}) \cdot P(E_1 \wedge E_2 \dots | \text{cause})}{P(E_1 \wedge E_2 \dots)} \quad \text{by conditional prob.}$$

$$= \frac{P(\text{cause})}{P(E_1 \wedge E_2 \dots)} \prod_{i=1}^n P(E_i | \text{cause})$$

$$= \alpha P(\text{cause}) \prod_{i=1}^n P(E_i | \text{cause}), \quad \alpha = \frac{1}{P(E_1 \wedge E_2 \dots)}$$

2. Now, to compare, the probability of causes given observations, we can do it in relative terms of α

That is, relative likelihood $\frac{P(X | E_1 \wedge E_2 \dots)}{P(Y | E_1 \wedge E_2 \dots)} = \frac{P(X) \prod_{i=1}^n P(E_i | X)}{P(Y) \prod_{i=1}^n P(E_i | Y)}$