Reasoning Under Uncertainty: Conditional Probability and Probabilistic Independence

CPSC 322 Lecture 24

March 10, 2006
Textbook §9.1 – §9.3
Recap

Conditional Probability

Bayes’ Theorem

Strict (or Marginal) Independence
Probability

- Probability is formal measure of uncertainty. There are two camps:
  - **Frequentists**: believe that probability represents something *objective*, and compute probabilities by counting the frequencies of different events
  - **Bayesians**: believe that probability represents something *subjective*, and understand probabilities as degrees of belief.
    - They compute probabilities by starting with prior beliefs, and then updating beliefs when they get new data.
    - **Example**: Your degree of belief that a bird can fly is your measure of belief in the flying ability of an individual based only on the knowledge that the individual is a bird.
    - Other agents may have different probabilities, as they may have had different experiences with birds or different knowledge about this particular bird.
    - An agent’s belief in a bird’s flying ability is affected by what the agent knows about that bird.
A random variable is a term in a language that can take one of a number of different values.

The domain of a variable $X$, written $\text{dom}(X)$, is the set of values $X$ can take.

A possible world specifies an assignment of one value to each random variable.

$w \models X = x$ means variable $X$ is assigned value $x$ in world $w$.

Let $\Omega$ be the set of all possible worlds.

Define a nonnegative measure $\mu(w)$ to each world $w$ so that the measures of the possible worlds sum to 1.

The probability of proposition $f$ is defined by:

$$P(f) = \sum_{w \models f} \mu(w).$$
Axioms of Probability: finite case

- Four axioms define what follows from a set of probabilities:
  - **Axiom 1** $P(f) = P(g)$ if $f \leftrightarrow g$ is a tautology. That is, logically equivalent formulae have the same probability.
  - **Axiom 2** $0 \leq P(f)$ for any formula $f$.
  - **Axiom 3** $P(\tau) = 1$ if $\tau$ is a tautology.
  - **Axiom 4** $P(f \lor g) = P(f) + P(g)$ if $\neg(f \land g)$ is a tautology.

- You can think of these axioms as constraints on which functions $P$ we can treat as probabilities.
- These axioms are sound and complete with respect to the semantics.
  - if you obey these axioms, there will exist some $\mu$ which is consistent with your $P$
  - there exists some $P$ which obeys these axioms for any given $\mu
Recap

Conditional Probability

Bayes’ Theorem

Strict (or Marginal) Independence
A probability distribution on a random variable $X$ is a function $\text{dom}(X) \rightarrow [0, 1]$ such that

$$x \mapsto P(X = x).$$

This is written as $P(X)$.

This also includes the case where we have tuples of variables. E.g., $P(X, Y, Z)$ means $P(\langle X, Y, Z \rangle)$.

When $\text{dom}(X)$ is infinite sometimes we need a probability density function...
Conditioning

- Probabilistic conditioning specifies how to revise beliefs based on new information.
- You build a probabilistic model taking all background information into account. This gives the prior probability.
- All other information must be conditioned on.
- If evidence $e$ is all of the information obtained subsequently, the conditional probability $P(h|e)$ of $h$ given $e$ is the posterior probability of $h$. 
Evidence \( e \) rules out possible worlds incompatible with \( e \).

We can represent this using a new measure, \( \mu_e \), over possible worlds:

\[
\mu_e(\omega) = \begin{cases} 
\frac{1}{P(e)} \times \mu(\omega) & \text{if } \omega \models e \\
0 & \text{if } \omega \not\models e
\end{cases}
\]

The conditional probability of formula \( h \) given evidence \( e \) is

\[
P(h|e) = \sum_{\omega \models h} \mu_e(\omega)
\]

\[
= \frac{P(h \land e)}{P(e)}
\]
Chain Rule

\[ P(f_1 \land f_2 \land \ldots \land f_n) \]
\[ = \quad P(f_n | f_1 \land \cdots \land f_{n-1}) \times \]
\[ P(f_1 \land \cdots \land f_{n-1}) \]
\[ = \quad P(f_n | f_1 \land \cdots \land f_{n-1}) \times \]
\[ P(f_{n-1} | f_1 \land \cdots \land f_{n-2}) \times \]
\[ P(f_1 \land \cdots \land f_{n-2}) \]
\[ = \quad P(f_n | f_1 \land \cdots \land f_{n-1}) \times \]
\[ P(f_{n-1} | f_1 \land \cdots \land f_{n-2}) \times \]
\[ \cdots \times P(f_3 | f_1 \land f_2) \times P(f_2 | f_1) \times P(f_1) \]
\[ = \quad \prod_{i=1}^{n} P(f_i | f_1 \land \cdots \land f_{i-1}) \]
Recap

Conditional Probability

Bayes' Theorem

Strict (or Marginal) Independence
Bayes’ theorem

The chain rule and commutativity of conjunction ($h \land e$ is equivalent to $e \land h$) gives us:

\[
P(h \land e) = P(h|e) \times P(e) \\
= P(e|h) \times P(h).
\]

If $P(e) \neq 0$, you can divide the right hand sides by $P(e)$:

\[
P(h|e) = \frac{P(e|h) \times P(h)}{P(e)}.
\]

This is Bayes’ theorem.
Why is Bayes’ theorem interesting?

Often you have causal knowledge:

\[ P(\text{symptom} \mid \text{disease}) \]
\[ P(\text{light is off} \mid \text{status of switches and switch positions}) \]
\[ P(\text{alarm} \mid \text{fire}) \]
\[ P(\text{image looks like } 🌳 \mid \text{a tree is in front of a car}) \]

and want to do evidential reasoning:

\[ P(\text{disease} \mid \text{symptom}) \]
\[ P(\text{status of switches} \mid \text{light is off and switch positions}) \]
\[ P(\text{fire} \mid \text{alarm}). \]
\[ P(\text{a tree is in front of a car} \mid \text{image looks like } 🌳) \]
Lecture Overview

Recap

Conditional Probability

Bayes’ Theorem

Strict (or Marginal) Independence
Random variable $X$ is independent of random variable $Y$ if, for all $x_i \in \text{dom}(X)$, $y_j \in \text{dom}(Y)$ and $y_k \in \text{dom}(Y)$,

$$
P(X = x_i | Y = y_j) = P(X = x_i | Y = y_k) = P(X = x_i).
$$

That is, knowledge of $Y$’s value doesn’t affect your belief in the value of $X$.
This is also called marginal independence.
Examples of probabilistic independence

- The probability that the Canucks will win the Stanley Cup is independent of whether light $l_1$ is lit.
  - remember the diagnostic assistant domain: the picture will recur in a minute!
- Whether there is someone in a room is independent of whether a light $l_2$ is lit.
- Whether light $l_1$ is lit is not independent of the position of switch $s_2$. 