Formal Modeling in Cognitive Science

Lecture 18: Conditional Probability; Bayes' Theorem

Steve Renals (notes by Frank Keller)

School of Informatics
University of Edinburgh
s.renals@ed.ac.uk

20 February 2007

1. Conditional Probability and Independence
    - Conditional Probability
    - Independence
2. Bayes' Theorem
    - Total Probability
    - Bayes' Theorem

Conditional Probability

Definition: Conditional Probability

If A and B are two events in a sample space S, and $P(A) \neq 0$ then the conditional probability of B given A is:

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

Intuitively, the conditional probability $P(B|A)$ is the probability that the event B will occur given that the event A has occurred.

Examples

The probability of having a traffic accident given that it snows: $P(\text{accident}|\text{snow})$.

The probability of reading the word *amok* given that the previous word was *run*: $P(\text{amok}|\text{run})$.

Conditional Probability

Example

A manufacturer knows that the probability of an order being ready on time is 0.80, and the probability of an order being ready on time and being delivered on time is 0.72. What is the probability of an order being delivered on time, given that it is ready on time?

R: order is ready on time; D: order is delivered on time.
$P(R) = 0.80$, $P(R \cap D) = 0.72$. Therefore:

$$P(D|R) = \frac{P(R \cap D)}{P(R)} = \frac{0.72}{0.80} = 0.90$$

Conditional Probability

From the definition of conditional probability, we obtain:

Theorem: Multiplication Rule

If A and B are two events in a sample space S, and $P(A) \neq 0$ then:

$$P(A \cap B) = P(A)P(B|A)$$

As $A \cap B = B \cap A$, it follows also that:

$$P(A \cap B) = P(B)P(A|B)$$

Example

Back to lateralization of language (see last lecture). Let $P(A) = 0.15$ be the probability of being left-handed, $P(B) = 0.05$ be the probability of language being right-lateralized, and $P(A \cap B) = 0.04$.

The probability of language being right-lateralized given that a person is left-handed:

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{0.04}{0.15} = 0.267$$

The probability being left-handed given that language is right-lateralized:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.04}{0.05} = 0.80$$

Independence

Definition: Independent Events

Two events A and B are independent if and only if:

$$P(B \cap A) = P(A)P(B)$$

Intuitively, two events are independent if the occurrence of non-occurrence of either one does not affect the probability of the occurrence of the other.

Theorem: Complement of Independent Events

If A and B are independent, then A and $\bar{B}$ are also independent.

This follows straightforwardly from set theory.

Independence

Example

A coin is flipped three times. Each of the eight outcomes is equally likely. A: head occurs on each of the first two flips, B: tail occurs on the third flip, C: exactly two tails occur in the three flips. Show that A and B are independent, B and C dependent.

$$\begin{array}{ll}
A = \{HHH, HHT\} & P(A) = \frac{1}{4} \\
B = \{HHT, HTT, THT, TTT\} & P(A) = \frac{1}{2} \\
C = \{HTT, THT, TTH\} & P(C) = \frac{3}{8} \\
A \cap B = \{HHT\} & P(A \cap B) = \frac{1}{8} \\
B \cap C = \{HTT, THT\} & P(B \cap C) = \frac{1}{4}
\end{array}$$

$P(A)P(B) = \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{8} = P(A \cap B)$, hence A and B are independent.
$P(B)P(C) = \frac{1}{2} \cdot \frac{3}{8} = \frac{3}{16} \neq P(B \cap C)$, hence B and C are dependent.

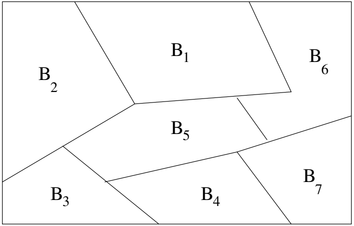
Total Probability

Theorem: Rule of Total Probability

If events $B_1, B_2, \ldots, B_k$ constitute a partition of the sample space S and $P(B_i) \neq 0$ for $i = 1, 2, \ldots, k$, then for any event A in S:

$$P(A) = \sum_{i=1}^{k} P(B_i)P(A|B_i)$$

$B_1, B_2, \ldots, B_k$ form a *partition* of S if they are pairwise mutually exclusive and if $B_1 \cup B_2 \cup \ldots \cup B_k = S$.



Total Probability

Example

In an experiment on human memory, participants have to memorize a set of words (B_1), numbers (B_2), and pictures (B_3). These occur in the experiment with the probabilities $P(B_1) = 0.5$, $P(B_2) = 0.4$, $P(B_3) = 0.1$.

Then participants have to recall the items (where A is the recall event). The results show that $P(A|B_1) = 0.4$, $P(A|B_2) = 0.2$, $P(A|B_3) = 0.1$. Compute $P(A)$, the probability of recalling an item.

By the theorem of total probability:

$$\begin{aligned} P(A) &= \textstyle\sum_{i=1}^{k} P(B_i)P(A|B_i) \\ &= P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + P(B_3)P(A|B_3) \\ &= 0.5 \cdot 0.4 + 0.4 \cdot 0.2 + 0.1 \cdot 0.1 = 0.29 \end{aligned}$$

Bayes' Theorem

Bayes' Theorem

If $B_1, B_2, \dots, B_k$ are a partition of S and $P(B_i) \neq 0$ for $i = 1, 2, \dots, k$, then for any A in S such that $P(A) \neq 0$:

$$P(B_r|A) = \frac{P(B_r)P(A|B_r)}{\sum_{i=1}^{k} P(B_i)P(A|B_i)}$$

This can be simplified by renaming $B_r = B$ and by substituting $P(A) = \sum_{i=1}^{k} P(B_i)P(A|B_i)$ (theorem of total probability):

Bayes' Theorem (simplified)

$$P(B|A) = \frac{P(B)P(A|B)}{P(A)}$$

Bayes' Theorem

Example

Reconsider the memory example. What is the probability that an item that is correctly recalled (A) is a picture (B_3)?

By Bayes' theorem:

$$\begin{aligned} P(B_3|A) &= \frac{P(B_3)P(A|B_3)}{\sum_{i=1}^{k} P(B_i)P(A|B_i)} \\ &= \frac{0.1 \cdot 0.1}{0.29} = 0.0345 \end{aligned}$$

The process of computing $P(B|A)$ from $P(A|B)$ is sometimes called *Bayesian inversion.*

Summary

- Conditional probability: $P(B|A) = \frac{P(A \cap B)}{P(A)}$;
- independence: $P(B \cap A) = P(A)P(B)$.
- rule of total probability: $P(A) = \sum_{i=1}^{k} P(B_i)P(A|B_i)$;
- Bayes' theorem: $P(B|A) = \frac{P(B)P(A|B)}{P(A)}$.