Formal Modeling in Cognitive Science

Lecture 17: Sample Spaces, Events, Probabilities

Steve Renals (notes by Frank Keller)

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

19 February 2007

Terminology

Terminology for probability theory:

- *experiment:* process of observation or measurement; e.g., coin flip;
- *outcome:* result obtained through an experiments; e.g., coin shows tail;
- *sample space:* set of all possible outcomes of an experiment; e.g., sample space for coin flip: $S = \{H, T\}$.

For now, we will only deal with *discrete sample spaces* (i.e., sample spaces whose elements can be mapped the set of integers).

Terminology

Example: Finite Sample Space

Roll two dice, each with numbers 1–6. Sample space:

$$S_1 = \{(x, y)|x = 1, 2, \dots, 6; y = 1, 2, \dots, 6\}$$

Alternative sample space for this experiment: sum of the dice:

$$S_2 = \{x|x = 2, 3, \dots, 12\}$$

Example: Infinite Sample Space

Flip a coin until head appears for the first time:

$$S_3 = \{H, TH, TTH, TTTH, TTTTH, \dots\}$$

Events

Often we are not interested in individual outcomes, but in events. An *event* is a subset of a sample space.

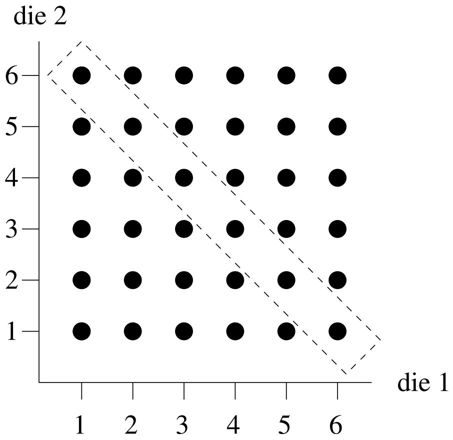
Example

With respect to S_1, describe the event B of rolling a total of 7 with the two dice.

$$B = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$$

Events

The event B can be represented graphically:



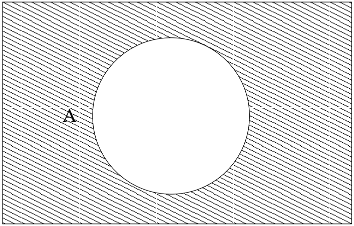
Events

Often we are interested in combinations of two or more events. This can be represented using set theoretic operations. Assume a sample space S and two events A and B:

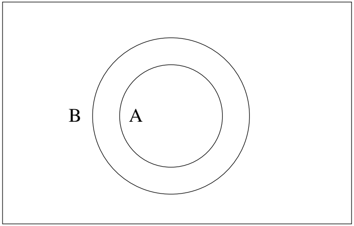
- *complement $\bar{A}$ (also A'):* all elements of S that are not in A;
- *subset $A \subset B$:* all elements of A are also elements of B;
- *union $A \cup B$:* all elements of S that are in A or B;
- *intersection $A \cap B$:* all elements of S that are in A and B.

These operations can be represented graphically using *Venn diagrams.*

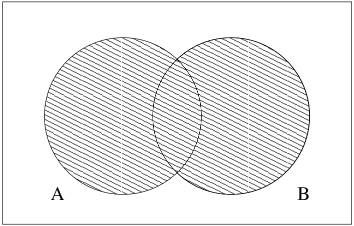
Venn Diagrams



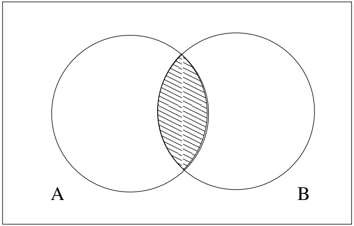
$\bar{A}$



$A \subset B$



$A \cup B$



$A \cap B$

Axioms of Probability

Events are denoted by capital letters A, B, C, etc. The *probability* of and event A is denoted by $P(A)$.

Axioms of Probability

1. The probability of an event is a nonnegative real number: $P(A) \geq 0$ for any $A \subset S$.
2. $P(S) = 1$.
3. If $A_1, A_2, A_3, \ldots,$ is a sequence of mutually exclusive events of S, then:

$$P(A_1 \cup A_2 \cup A_3 \cup \ldots) = P(A_1) + P(A_2) + P(A_3) + \ldots$$

Probability of an Event

Theorem: Probability of an Event

If A is an event in a discrete sample space S and $O_1, O_2, O_3, \ldots,$ are the individual outcomes comprising A, then $P(A) = P(O_1) + P(O_2) + P(O_3) + \ldots$

Example

We flip a fair coin twice. What's the probability of obtaining at least one head?

The sample space is $S = \{HH, HT, TH, TT\}$. As the coin is fair, all outcomes are equally likely: $P(HH) = P(HT) = P(TH) = P(TT) = \frac{1}{4}$. The event of obtaining at least one head is $A = \{HH, HT, TH\}$ and $P(A) = P(HH) + P(HT) + P(TH) = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4}$.

Probability of an Event

This leads us to the following special case of the previous theorem:

Theorem: Equally Likely Outcomes

If an experiment can result in N equally likely outcomes, and if n of these outcomes constitute an event A, then $P(A) = \frac{n}{N}$.

This theorem is consistent with the *frequency interpretation* of probability theory: the probability of an event is the proportion of the time that events of the same kind will occur in the long run.

This will become important later in this course.

Probability of an Event

Example

Assume all letters occur equally often in English. Then what's the probability of a three-letter word only consisting of vowels?

There are $N = 26^3$ three letter words. The set of vowels is $\{a, e, i, o, u\}$. Then the event A of having a three letter word consisting only of vowels is $A = \{aaa, aae, aai, aao, \ldots\}$. The size of this set is $n = 5^3$. By the theorem of equally likely outcomes, $P(A) = \frac{n}{N} = \frac{5^3}{26^3} = 0.00711$.

Rules of Probability

Theorems: Rules of Probability

1. If A and $\bar{A}$ are complementary events in the sample space S, then $P(\bar{A}) = 1 - P(A)$.
2. $P(\emptyset) = 0$ for any sample space S.
3. If A and B are events in a sample space S and $A \subset B$, then $P(A) \leq P(B)$.
4. $0 \leq P(A) \leq 1$ for any event A.

Examples and Explanations

1. What's the probability of a three letter word not consisting of three vowels? $P(\bar{A}) = 1 - P(A) = 1 - 0.00711 = 0.99289$, where A is the set of all three letter words containing only vowels (see example above).
2. This follows from set theory: $S \cup \emptyset = S$, hence $P(S) + P(\emptyset) = P(S)$, hence $P(\emptyset) = 0$.
3. Let $A = \{HT, TH\}$, the event of getting exactly one head when flipping a coin twice, and $B = \{HH, HT, TH\}$, the event of getting at least one head. Then $P(A) = \frac{1}{2}$ and $P(B) = \frac{3}{4}$, i.e., $P(A) \leq P(B)$.
4. Again, this follows from set theory: $\emptyset \subset A \subset S$ for any event A. Hence $P(\emptyset) \leq A \leq P(S)$, and therefore $0 \leq A \leq 1$.

Addition Rule

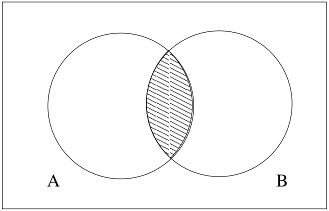
Axiom 3 allows us to add the probabilities of mutually exclusive events. This is called the *special addition rule.* But what about events that are not mutually exclusive?

Theorem: General Addition Rule

If A and B are two events in a sample space S, then:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Illustrated using a Venn diagram: shaded area occurs twice and has to be subtracted.



Addition Rule

Example

Language is lateralized in the brain: in most people, language processing mainly takes place in the left hemisphere. In some people, however, language is right-lateralized, i.e., it is mainly processed in the right hemisphere.

Assume the probability of being left-handed is $P(A) = 0.15$, and the probability of language being right-lateralized is $P(B) = 0.05$.

If A and B are mutually exclusive then the probability of being either left-handed or right-lateralized is $P(A \cup B) = P(A) + P(B) = 0.2$.

However, the two events are not mutually exclusive: there are left-handers that are right-lateralized (in fact, this is more likely in left-handers than in right-handers). We know that $P(A \cap B) = 0.04$.

Now the probability of being left-handed or right-lateralized is $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.16$.

Summary

- Sample space S contains all possible outcomes of an experiment; events A and B are subsets of S;
- for equally likely outcomes: $P(A) = \frac{n}{N}$;
- rules of probability:
  - $P(\bar{A}) = 1 - P(A)$;
  - $P(\emptyset) = 0$;
  - if $A \subset B$, then $P(A) \leq P(B)$;
  - $0 \leq P(B) \leq 1$;
- addition rule: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.