Basics of Probability

Gilles Cazelais

Lots of problems that come up in applications involve phenomena for which exact predictions are either impossible or very difficult. The best we can do in such cases is to determine the *probability* of the possible outcomes. We can think of probability as a number between 0 and 1 that measures how likely an event is to happen. An event with probability zero would be an impossible event and an event with probability one would be certain to happen.

An **experiment** is an activity with an observable result. A **sample space** of an experiment is a set of outcomes such that in each trial of the experiment, *one and only one* of these outcomes occurs. Each element of a sample space is called a **simple event**. An **event** is a subset of a sample space.

Given a sample space $S = \{E_1, E_2, \dots, E_n\}$ with n simple events, a **probability distribution** for S is an assignment of a number $P(E_i)$ to each simple event E_i so that

1. $0 \leq P(E_i) \leq 1$, for all $i = 1, \dots, n$
2. $P(E_1) + P(E_2) + \cdots + P(E_n) = 1$.

The **probability of an event** E, denoted by $P(E)$, is the sum of the probabilities of all simple events contained in E.

In practice, probability distributions are often estimated by experiment. For example, if you have a coin and want to know what $P(heads)$ is, you could flip it one million times and collect statistics. The number of *heads* obtained divided by one million could then be used as a very good estimate of $P(heads)$. If an experiment is performed N times and the event E occurs $n(E)$ times, then the ratio $n(E)/N$ is called the **relative frequency** of E. In general, if the number of trials N is high enough, the relative frequency is a good approximation of $P(E)$. In fact, we could have defined $P(E)$ by

$$P(E) = \lim_{N \to \infty} \frac{n(E)}{N}.$$

Another method commonly used to assign a probability distribution to a sample space is by simply assuming it. For instance, in many situations it is reasonable to assume that all outcomes in the sample space are *equally likely* to occur. If you have a coin and have no reason to suspect that it is not fair, it is then reasonable to assume that both outcomes *heads* and *tails* are equally likely to occur. In this case we would assign

$$P(heads) = P(tails) = 1/2.$$

Statistical methods could then be used to test if the assumption that the coin is fair is acceptable or not.

If $S = \{E_1, E_2, \dots, E_n\}$ and if we assume that all simple events are **equally likely** to occur, then

$$P(E_1) = P(E_2) = \cdots = P(E_n) = \frac{1}{n}.$$

In this case we can calculate the probability of any event E as follows

$$P(E) = \frac{\text{number of simple events in } E}{\text{number of simple events in } S}.$$

► The above formula works *only* when the simple events in the sample space are equally likely. For example, if a die is weighted and the outcomes are not equally likely, then the above formula would not apply.

Typeset with LaTeX on April 20, 2006.