Prelude to probability: operations on sets

In the next math circle session we'll see an introduction to some basic concepts of probability theory. This provides a mathematical framework for understanding a huge range of real-world processes, from statistics about human populations to the effects of 'random' external influences on scientific experiments. However, it's essentially just an abstract notion of counting the sizes of different sets, and so an important first step is to have an understanding of the basic operations we can perform on sets: these are recalled below.

First off, suppose that Ω is a set. It won't really matter what it's a set *of*, but for intuition we often think of it as the set of all possible outcomes from some experiment: for example,

1. the set of values on the faces of a die, conventionally $\{1, 2, \ldots, 6\}$, which are the possible outcomes upon rolling the die;
2. a set of coloured balls in a bag, where our experiment might be to remove a ball from the bag without looking;
3. the set of possible locations on a dartboard, which are the possible outcomes of an 'experimental' dart-throw.

In probability we usually refer to the members of Ω as **samples**, corresponding to the idea that our experiment is drawing a 'sample' from some collection of theoretically possible outcomes. Correspondingly, the set Ω is the **sample space**.

Subsets of Ω are called **events**: think of them as the sets of possible samples that have some additional feature in common. For instance, in example 2. above, the set $\{\text{red balls in the bag}\}$ is such a subset, and corresponds to the event that the sample ball we remove from the bag turns out to be a red one. We sometimes denote that A is an event by $A \subseteq \Omega$.

Importantly, there are various ways in which we can combine events to make new ones. If A and B are events, then we can also define

- the **complement** of A:

$$A^{\mathrm{c}} = \text{the set of samples in } \Omega \text{ that are } \textit{not} \text{ in } A$$

  (this is sometimes written $\mathrm{not}A$);

- the **intersection** of A and B:

$$A \cap B = \text{the set of samples in } \Omega \text{ that are in both } A \text{ and } B$$

  (sometimes written A AND B);

- the **union** of A and B:

$$A \cup B = \text{the set of samples in } \Omega \text{ that are in either } A \text{ or } B \text{ or possibly both}$$

  (sometimes written A OR B);

- the **difference** of A and B:

$$A \setminus B = \text{the set of samples in } \Omega \text{ that are in } A \text{ but not in } B.$$

We write $\emptyset$ for the empty set: the set with no elements, or equivalently the event that never occurs. If $A \cap B = \emptyset$ is the empty set — that is, no sample is simultaneously in both A and B — then we say A and B are **disjoint**. A larger collection of events A_1, A_2, ..., A_k is **disjoint** if any two of them are disjoint.

Finally, we write $A \subseteq B$ and say A is a **subset** of B if every element of A is also an element of B, and if ω is an element of A then we write $\omega \in A$.

After reviewing these operations a little, in the math circle session we'll introduce the idea of the 'probability' of an event A, say $P(A)$. This is a number between 0 and 1 that indicates how likely it is that a random sample ω from the sample space will be in A. Equivalently, it could be the proportion of times that we see samples that come from A if we run our experiment a huge number of times to get a different, independent sample each time. Mathematically, we think of P as a function from the set of subsets of Ω to the set of all real numbers $0 \leq p \leq 1$, and assume that it has the following basic properties:

[a] since $\emptyset$ is "the event that never occurs", we must have $P(\emptyset) = 0$;

[b] similarly, every possible sample is in Ω, so $P(\Omega) = 1$;

[c] lastly, if A and B are disjoint, then

$$P(A \cup B) = P(A) + P(B).$$

To see the sense in this, observe that if we run our experiment 1000 times and let $P(A)$ be the fraction of those times that the resulting sample was in A, and similarly for $P(B)$, then since A and B never occur at the same time, the faction of times we obtain $A \cup B$ is the sum of their individual fractions.

If, for example, $\Omega = \{1, 2, \ldots, N\}$, can you think of any examples of functions P that satisfy these rules [Hint: what if 'all outcomes are equally likely'?]? We'll see some on Sunday, and then begin to explore the many consequences of the simple rules [a], [b] and [c].

Further reading

Many good introductions to mathematical probability are available; for example, I like Kai Lai Chung's *Elementary Probability Theory with Stochastic Processes*.

Tim Austin
DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, LOS ANGELES, CA 90095-1555, USA

Email: `timaustin@math.ucla.edu`
URL: `http://www.math.ucla.edu/~timaustin`