Lecture 21: Probability and Independence

Lecture 21: Probability and Independence

Why should you care?

- Many algorithms work better if they make random choices. For example, “quicksort” can avoid its worst case by making random choices. The Pollard-ρ factorisation algorithm makes random choices as it searches for integer factors. To understand such algorithms you must understand probability.
- Many mathematical objects can only be constructed randomly. For example, graphs with large chromatic number but no short cycles.
- Random models are useful for studying the structure of real world networks. And of traffic modelling on them.
- Probability also helps in machine learning, reliability modeling, simulation algorithms, data mining, speech recognition etc etc.
- All your life you will be making probabilistic judgements!

Probability gives us a way to model random processes mathematically. These processes could be anything from the rolling of dice, to radioactive decay of atoms, to the performance of a stock market index. The mathematical environment we work in when dealing with probabilities is called a probability space.

Probability spaces

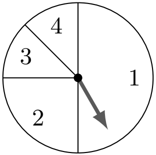
We'll start with a formal definition and then look at some examples of how the definition is used.

A *probability space* consists of

- a set S called a *sample space* which contains all the possible *outcomes* of the random process; and
- a *probability* function $\mathrm{Pr} : S \to [0, 1]$ such that the sum of the probabilities of the outcomes in S is 1.

Each time the process occurs it should produce exactly one outcome (never zero or more than one). The probability of an outcome is a measure of the likeliness that it will occur. It is given as a real number between 0 and 1 inclusive, where 0 indicates that the outcome cannot occur and 1 indicates that the outcome must occur.

Example.



The spinner above might be modeled by a probability space with sample space $S = \{1, 2, 3, 4\}$ and probability function given as follows.

$$\mathrm{Pr}(s) = \begin{cases} \frac{1}{2} & \text{for } s = 1 \\ \frac{1}{4} & \text{for } s = 2 \\ \frac{1}{8} & \text{for } s = 3 \\ \frac{1}{8} & \text{for } s = 4. \end{cases}$$

It can be convenient to give this as a table:

s	1	2	3	4
$\mathrm{Pr}(s)$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$

Example. Rolling a fair six-sided die could be modeled by a probability space with sample space $S = \{1, 2, 3, 4, 5, 6\}$ and probability function $\Pr$ given as follows.

s	1	2	3	4	5	6
$\Pr(s)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

A sample space like this one where every outcome has an equal probability is sometimes called a *uniform sample space*. Outcomes from a uniform sample space are said to have been taken *uniformly at random*.

Events

> An *event* is a subset of the sample space.

An event is just a collection of outcomes we are interested in for some reason.

Example. In the die rolling example with $S = \{1, 2, 3, 4, 5, 6\}$, we could define the event of rolling at least a 3. Formally, this would be the set $\{3, 4, 5, 6\}$. We could also define the event of rolling an odd number as the set $\{1, 3, 5\}$.

> The probability of an event A is the sum of the probabilities of the outcomes in A.

Example. In the spinner example, for the event $A = \{1, 2, 4\}$, we have

$$\begin{aligned}\Pr(A) &= \Pr(1) + \Pr(2) + \Pr(4) \\ &= \tfrac{1}{2} + \tfrac{1}{4} + \tfrac{1}{8} \\ &= \tfrac{7}{8}.\end{aligned}$$

In a uniform sample space (where all outcomes are equally likely) the probability of an event A can be calculated as:

$$\Pr(A) = \frac{\text{number of outcomes in } A}{\text{total number of outcomes}} = \frac{|A|}{|S|}.$$

Operations on events

Because events are defined as sets we can perform set operations on them. If A and B are events for a sample space S, then

- $A \cup B$ is the event "A or B"
- $A \cap B$ is the event "A and B"
- $\overline{A}$ is the event "not A"

We always take the sample space as our universal set, so $\overline{A}$ means $S - A$.

Probabilities of unions

We saw in the section on the inclusion-exclusion principal that $|A \cup B| = |A| + |B| - |A \cap B|$ for finite sets A and B. We have a similar law in probability.

For any two events A and B,

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B).$$

Example. In our die rolling example, let $A = \{1, 2\}$ and $B = \{2, 3, 4\}$ be events. Then

$$\Pr(A \cup B) = \frac{|A|}{|S|} + \frac{|B|}{|S|} - \frac{|A \cap B|}{|S|} = \frac{2}{6} + \frac{3}{6} - \frac{1}{6} = \frac{2}{3}.$$

Two events A and B are *mutually exclusive* if $\Pr(A \cap B) = 0$, that is, if A and B cannot occur together. For mutually exclusive events, we have

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

Independent events

We say that two events are *independent* when the occurrence or non-occurrence of one event does not affect the likelihood of the other occurring.

Two events A and B are *independent* if

$$\Pr(A \cap B) = \Pr(A)\Pr(B).$$

Example. A binary string of length 3 is generated uniformly at random. The event A that the first bit is a 1 is independent of the event B that the second bit is a 1. But A is *not* independent of the event C that the string contains exactly two 1s.

Formally, the sample space is
$S = \{111, 110, 101, 100, 011, 010, 001, 000\}$ and $\Pr(s) = \frac{1}{8}$ for any $s \in S$. So,

$$\begin{aligned}
A &= \{111, 110, 101, 100\} & \Pr(A) &= \tfrac{1}{2} \\
B &= \{111, 110, 011, 010\} & \Pr(B) &= \tfrac{1}{2} \\
C &= \{110, 101, 011\} & \Pr(C) &= \tfrac{3}{8} \\
A \cap B &= \{111, 110\} & \Pr(A \cap B) &= \tfrac{1}{4} \\
A \cap C &= \{110, 101\} & \Pr(A \cap C) &= \tfrac{1}{4}.
\end{aligned}$$

So $\Pr(A \cap B) = \Pr(A)\Pr(B)$ but $\Pr(A \cap C) \neq \Pr(A)\Pr(C)$.

Warning

Random processes can occur in both discrete and continuous settings, and probability theory can be applied in either setting. In this lecture, and in the next four lectures, we are discussing only the discrete case. Many of the definitions and results we state apply only in this case. Our definition of a probability space, for example, is actually the definition of a discrete probability space, and so on.

The discrete setting provides a good environment to learn most of the vital concepts and intuitions of probability theory. What you learn here is very useful in itself, and will act as a good base if you go on to study continuous probability.

Questions

An integer is chosen uniformly at random from the set $\{1, 2, \ldots, 30\}$. Let A be the event that the integer is at most 20. Let B be the event that the integer is divisible by 6. Let C be the event that the integer's last digit is a 5.

21.1 Write A, B and C as sets, and find their probabilities.

$A = \{1, 2, \ldots, 19, 20\}$. $\Pr(A) = 2/3$.
$B = \{6, 12, 18, 24, 30\}$. $\Pr(B) = 1/6$.
$C = \{5, 15, 25\}$. $\Pr(C) = 1/10$.

21.2 Find the probabilities of $A \cup B$, $A \cup C$, and $B \cup C$. Which pairs of A, B, C are mutually exclusive?

$\Pr(A \cup B) = 11/15$, $\Pr(A \cup C) = 7/10$, and $\Pr(B \cup C) = 4/15$.
The only pair which are mutually exclusive is B and C.

21.3 Find the probabilities of $A \cap B$, $A \cap C$, and $B \cap C$. Which pairs of A, B, C are independent?

$\Pr(A \cap B) = 1/10$, $\Pr(A \cap C) = 1/15$, and $\Pr(B \cap C) = 0$.
The only pair which are independent is A and C.