Math-UA.233: Theory of Probability
Lecture 5

Tim Austin

tim@cims.nyu.edu
cims.nyu.edu/~tim
To model an experiment, we choose:

- A sample space $S$, containing all possible outcomes.
- A probability value $P(E)$ for each event $E \subset S$, quantifying how ‘likely’ that event is. We call $P$ a ‘probability function’ or ‘probability distribution’. 
In order to apply probability theory, $P$ must satisfy the axioms of probability:

**Axiom 1**: Any event $E$ satisfies $0 \leq P(E) \leq 1$.

**Axiom 2**: $P(S) = 1$.

**Axiom 3**: If the sequence of events $E_1, E_2, \ldots$ are mutually exclusive, then

$$P\left( E_1 \cup E_2 \cup \cdots \right) = P(E_1) + P(E_2) + \cdots.$$
In many simple cases, we assume ‘equally likely outcomes’: $S$ is finite, and

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

We also call this choice of $P$ the ‘uniform distribution on $S$’.
Example (Ross 2.5b)

Three balls are \textbf{randomly drawn} from a bowl containing 6 white and 5 black balls. What is the probability that one of the balls is white and the other two are black? [More than one approach works here: depends whether we order the balls]

NOTE: The phrase ‘randomly drawn’ in this problem really means ‘\textit{assume equally likely outcomes}’. 
Example (Ross 2.5c)

A committee of 5 is to be selected from a group of 6 men and 9 women. If the selection is made randomly, what is the probability that the committee consists of 3 men and 2 women? [Again, two possible approaches, but one is clearly better!]

Example (Ross 2.5d)

An urn contains \( n \) balls, one of which is special. If \( k \) of these balls are withdrawn one at a time, with each selection being equally likely to be any of the balls that remains at that time, what is the probability that the special ball is chosen? [Again, two possible approaches.]
Example (Ross 2.5e)

Suppose that $n + m$ distinguishable balls, of which $n$ are red and $m$ are blue, are arranged in a linear order in such a way that all $(n + m)!$ possible orderings are equally likely [think: a thorough shuffling of $n$ red cards and $m$ blue cards].

If we record the result of this experiment by listing only the colours of the successive balls, show that all the possible results remain equally likely.
Example (Ross 2.5h)

In the game of bridge, the entire deck of 52 cards is dealt out to 4 players. What is the probability that

(a) one of the players receives all 13 spades?
(b) each player receives 1 ace?

(For more gambling examples, see 2.5f, 2.5g and 2.5j.)
Example

What is the probability that two members of our class have the same birthday?

NOTE: For calculations, we assume equally likely outcomes among 365 possible days. But are they really equally likely? What about leap years?
Beyond equally likely outcomes

Some experiments clearly require a different choice of probabilities than ‘equally likely outcomes’.

Suppose we flip a coin, so \( S = \{ \text{H}, \text{T} \} \). Then we must have

\[
P(\{\text{H}\}) = p \quad \text{and} \quad P(\{\text{T}\}) = 1 - p
\]

for some \( 0 \leq p \leq 1 \). (Because these two single-outcome probabilities must be non-negative and their sum must equal \( P(S) = 1 \)).

If the coin is fair, then we assume \( p = 1/2 \), and in this case the outcomes are equally likely.

But if the coin is biased, then \( p \neq 1/2 \). This basic example will come up repeatedly later. It is called the ‘\( p \)-biased coin’.
Outcomes with different probabilities can also appear when we start with equally likely outcomes, but then we change our description of the experiment.
Example

Alice and Bob play a game. They roll a die, and if it gives 5 or 6 then Alice wins, otherwise Bob wins.

- ‘Complete’ description of the game: \( S = \{1, 2, 3, 4, 5, 6\} \), \( P = \text{uniform} \) (i.e. outcomes are equally likely).

But if we only care who wins, then we could switch to:

- ‘Reduced’ description: \( S' = \{A, B\} \), where \( A \) means ‘Alice wins’. Now we treat this as a single outcome, and ignore the exact value shown by the die. Now the correct probability function \( P' \) is specified by

\[
P'(\{A\}) = 1/3, \quad P'(\{B\}) = 2/3.
\]

This reduced description is mathematically the same as the 1/3-biased coin.
Similar but more complicated:

**Example**

We roll two fair dice, and record the sum of the values shown.

- ‘Complete’ description:
  
  \[ S = \{(i, j) : 1 \leq i, j \leq 6\}, \quad P = \text{uniform}. \]

- ‘Reduced’ description: \( S' = \{2, 3, \ldots, 12\} \) and
  
  \[ P'(\{2\}) = \frac{1}{36}, \quad P'(\{3\}) = \frac{1}{18}, \quad \ldots, \quad P'(\{12\}) = \frac{1}{36}. \]
A variant of this idea appears with experiments in which we are waiting for something to happen.

Example

An urn contains 2 indistinguishable red and 2 indistinguishable blue balls. They are withdrawn one-by-one at random (and not replaced) until a red ball is obtained.

Possible sample space: \( S = \{1, 2, 3\} \), indicating how many balls are withdrawn up to and including the first red.

Single-outcome probabilities:

\[
P(\{1\}) = \frac{2}{4} = \frac{1}{2}, \quad P(\{2\}) = \frac{2 \cdot 2}{4 \cdot 3} = \frac{1}{3}, \quad P(\{3\}) = \frac{2 \cdot 1 \cdot 2}{4 \cdot 3 \cdot 2} = \frac{1}{6}.
\]

IDEA: Each of these probabilities is obtained by considering the withdrawal of a fixed number of balls (1, 2 or 3, respectively).