Introduction to Probability

David Gerard
Some slides are borrowed from Linda Collins

2017-10-10

Learning Objective

- Sample Space/Events
- Axioms of Probability
- Some Probability Rules
- Sections 2.1.1 to 2.1.5 in DBC

Recall: Sample vs Population

population

A population is a group of cases for which you want information.

sample

A sample is a subgroup of cases of the population.

A simple random sample is when we choose a subset of the population where each subset is equally likely to be chosen.

Probability vs. Statistics

- **Statistics (Inference):**
  - Just observe a sample. What can we conclude (probabilistically) about the population?
  - Sample $\longrightarrow$ Population?
  - Messy and more of an art.
  - No correct answers. Lots of wrong answers. Some "good enough" answers.
- **Probability (from the viewpoint of Statisticians):**
  - Logically self-contained, a subset of Mathematics.
  - One correct answer.
  - We know the population. What is the probability of the sample?
  - Population $\longrightarrow$ Sample?

What is Probability?

Random

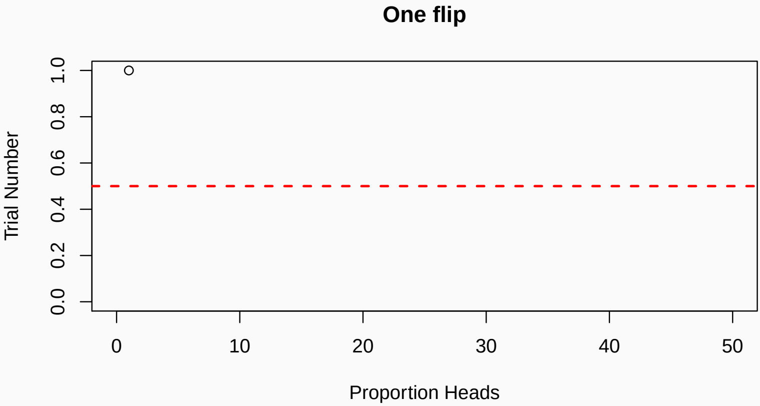
We call a phenomenon random if individual outcomes are uncertain but there is, nonetheless, a regular distribution of outcomes in a large number of repetitions.

Probability

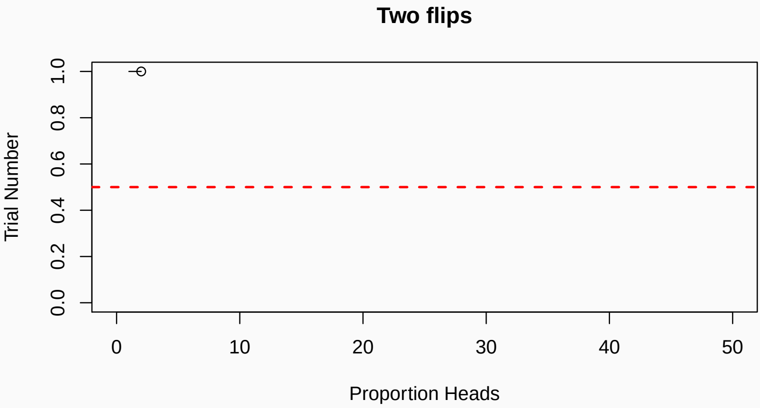
The probability of any outcome of a random phenomenon is the proportion of times the outcome would occur in a very long series of repetitions.

Flipping a coin

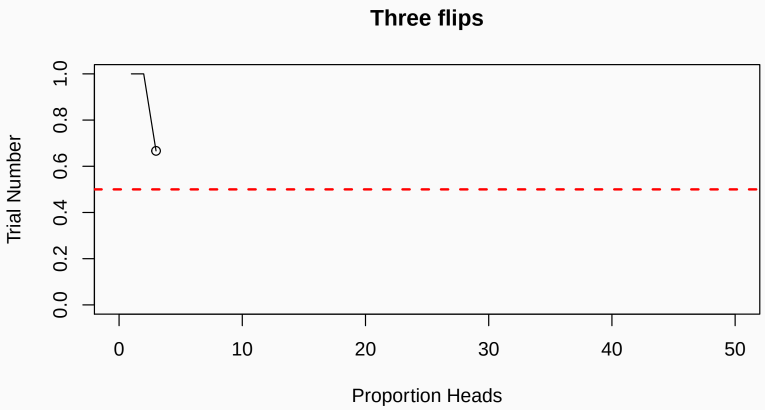
The probability of a fair coin landing heads is 0.5 because, in the long run, half of the time the coin will land heads.



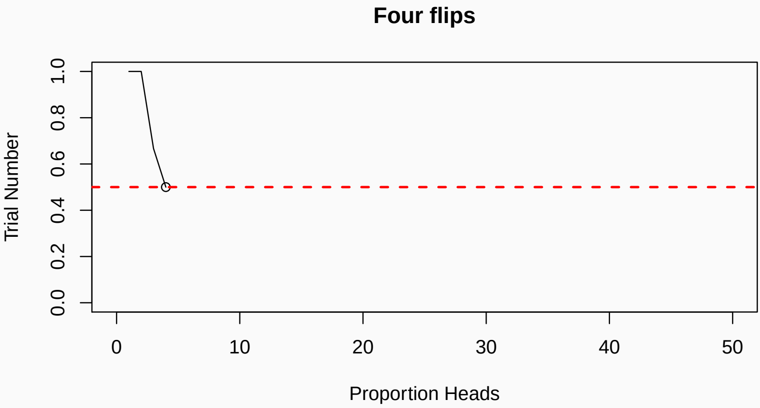
Flipping a coin



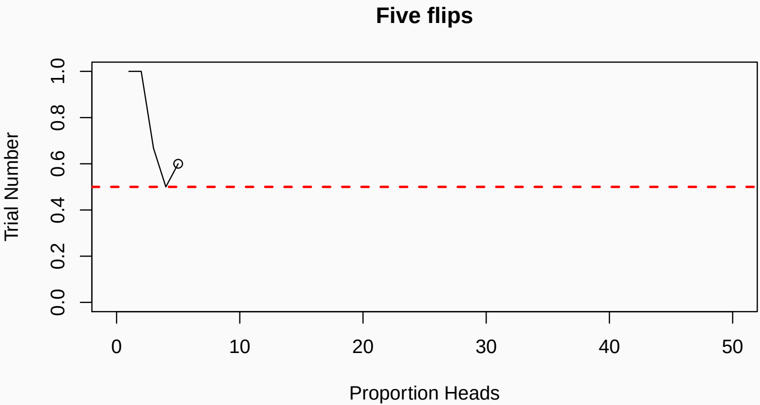
Flipping a coin



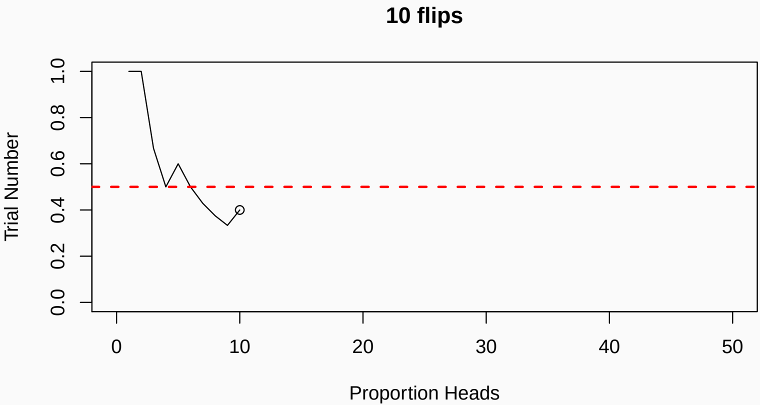
Flipping a coin



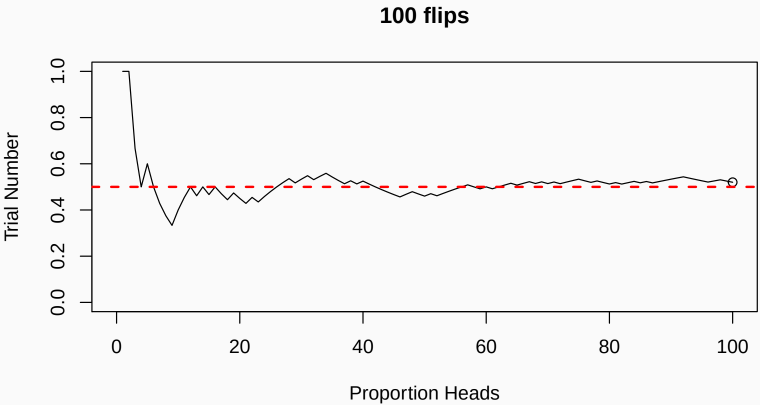
Flipping a coin



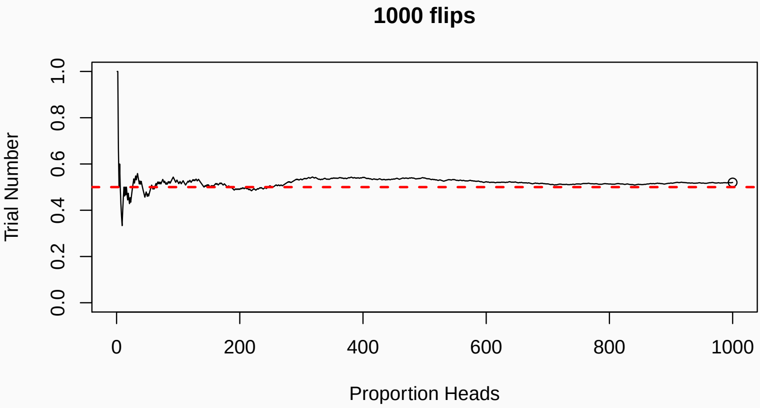
Flipping a coin



Flipping a coin



Flipping a coin



This just illustrated a fact known as the

Law of Large Numbers

If a situation, trial, or experiment is repeated, the proportion of outcomes of interest is more and more likely to be close to the probability of the outcome of interest as we repeat the experiment more and more times.

Sample Spaces and Events

Sample Space

The sample space S of a random phenomenon is the set of all possible outcomes.

Event

An event is an outcome or a set of outcomes of a random phenomenon. That is, an event is a subset of the sample space. We will denote the set of all events by $\mathcal{A}$.

E.g.: Suppose you roll a fair six-sided die once, what is the sample space?
Answer:

E.g.: What are all possible events?
Answer:

Sample Spaces and Events: Another Example

Suppose you flip a fair coin twice and observe the sequence of heads and tails, what is the sample space?
Answer: $S = \{HH, HT, TH, TT\}$,

What are all possible events?
Answer: All subsets of S. E.g. $\{HH, HT\}$, $\{HH, TH, TT\}$, etc...

Suppose instead of the sequence, you measure the number of heads. What is the sample space?
Answer:

What are the possible events?
Answer:

The Axioms of Probability

A **probability** on a sample space S (and a set $\mathcal{A}$ of events) is a function which assigns each event A (in $\mathcal{A}$) a value in $[0, 1]$ and satisfies the following rules:

- **Axiom 1:** All probabilities are nonnegative:

$$P(A) \geq 0 \qquad \text{for all events } A.$$

- **Axiom 2:** The probability of the whole sample space is 1:

$$P(S) = 1.$$

- **Axiom 3 (Addition Rule):** If two events A and B are disjoint (have no outcomes in common) then

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

The Axioms of Probablity: Example of Application

Suppose a random experiment has N different outcomes, such that ith outcomes occurs with probability p_i, $i = 1, \ldots, N$. It is natural to define the probability of an event as the sum of the probabilities of the distinct outcomes making up the event.

Why do we need to learn techniques for counting? If each outcome equally likely, $p_1 = \ldots = p_N = 1/N$, then for any event A,

$$P(A) = \frac{\#(A)}{\#(S)} = \frac{\text{number of outcomes in } A}{\text{number of outcomes in } S}$$

This setup satisfies the 3 axioms

- $P(A) \geq 0$
- $P(S) = \frac{\#(S)}{\#(S)} = 1$
- If A and B are disjoint then

$$P(A \cup B) = \frac{\#(A \cup B)}{\#(S)} = \frac{\#(A)}{\#(S)} + \frac{\#(B)}{\#(S)} = P(A) + P(B).$$

In some cases

Symmetry of outcomes: all outcomes of an experiment are assumed equally likely

Assume N outcomes: Probability of an outcome: $1/N$.

Event is defined as a subset of possible outcomes.

Probability of event containing n outcomes: n/N

- requires finitely many and equally likely outcomes
- can be determined by counting outcomes

E.g.: Roll a six-sided die, then
$P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = 1/6$
E.g.: Roll a six-sided die twice, then
$P(2 \text{ and } 3) = P(1 \text{ and } 1) = P(6 \text{ and } 3) = P(4 \text{ and } 5) = \cdots = 1/36$

Set Theory Primer

A Set Theory Primer

A set is *"a collection of definite, well distinguished objects of our perception or of our thought"*. (GEORG CANTOR, 1845-1918)

Some important sets:

- $\mathbb{N} = \{1, 2, 3, \ldots\}$, the set of *natural numbers*
- $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$, the set of *integers*
- $\mathbb{R} = (-\infty, \infty)$, the set of *real numbers*

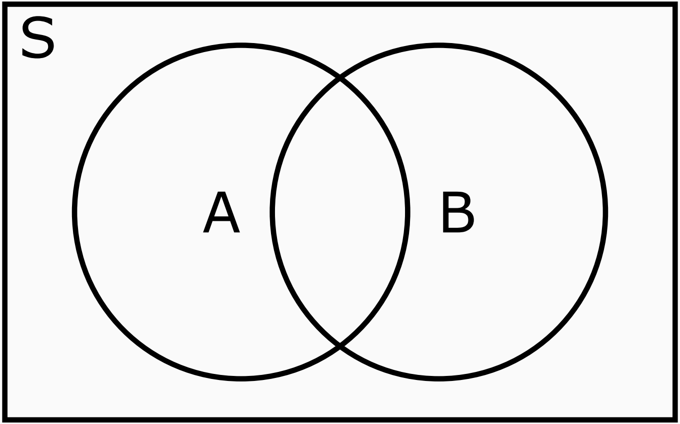
Intervals are denoted as follows:

$[0, 1]$	the interval from 0 to 1 including 0 and 1
$[0, 1)$	the interval from 0 to 1 including 0 but not 1
$(0, 1)$	the interval from 0 to 1 not including 0 and 1

If a is an element of the set A then we write $a \in A$.

Set's are intuitively thought of in terms of Venn diagrams

- S = a set, A = a subset of S, B = another subset of S.
- Denoted $A \subseteq S$ and $B \subseteq S$.

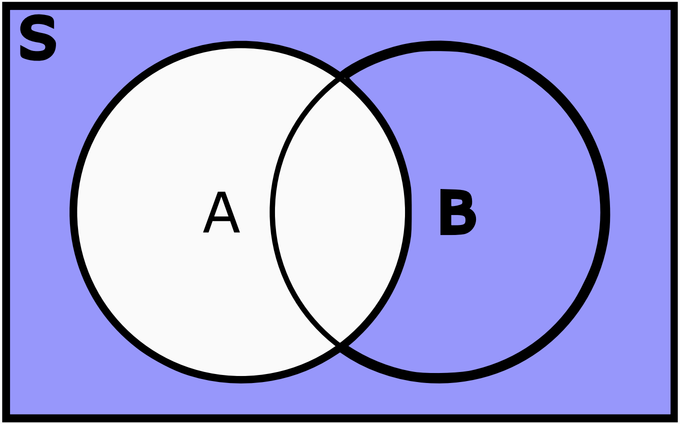


Empty Set

- The set with no elements is called the empty set.
- Denoted $\emptyset$.
- For any set A, we have $\emptyset \subseteq A$.

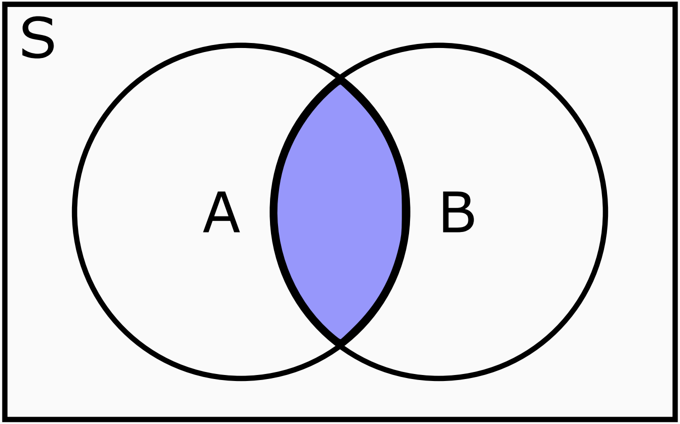
Set complement

- Set Complement: Set of all elements in S that are not in A.
- Denoted A (A^c or A')



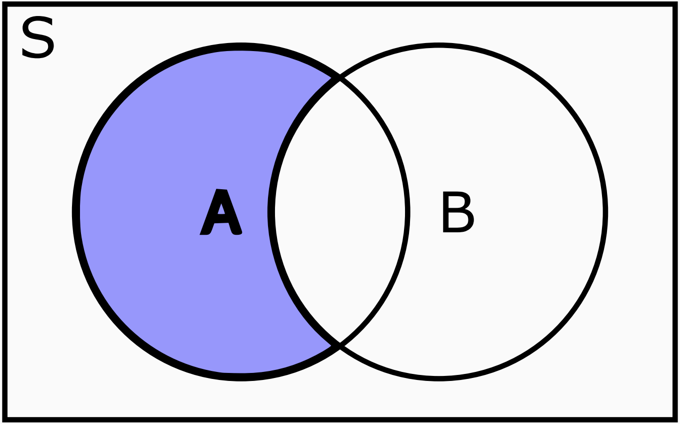
Intersection

- The intersection of A and B: Set of all elements in S which are both in A and in B.
- Denoted $A \cap B$.



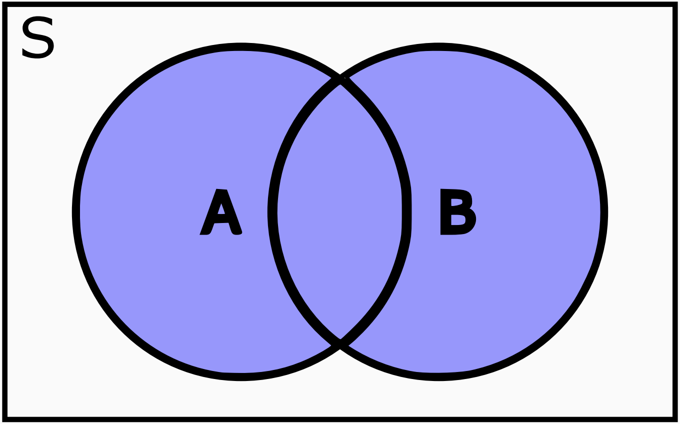
Set difference

- The set difference of A and B: Set of all elements in A which are not in B.
- Denoted $A \setminus B = A \cap B^c$.



Set union

- The union of A and B: Set of all elements in S that are in A or in B or in both.
- Denoted $A \cup B$.

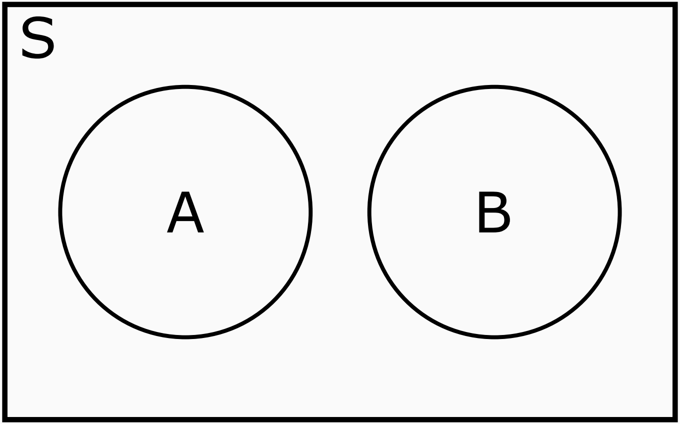


Two questions

- What is $A \cup A^c$?
- What is $A \cap A^c$?

Disjoint/mutually exclusive

- A and B are disjoint if A and B have no common elements, that is $A \cap B = \emptyset$. Two events A and B with this property are said to be mutually exclusive.



Some rules of probability

Set of all events satisifies

We assume the collection of events $\mathcal{A}$ satisfies:

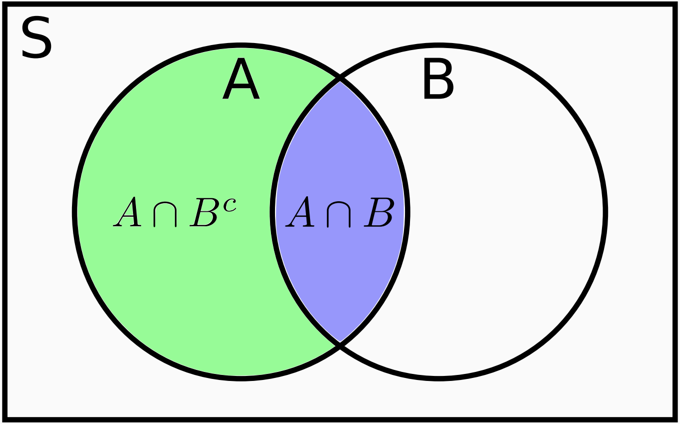
- $\phi, S \in \mathcal{A}$
- if $A \in \mathcal{A}$ then $A^c \in \mathcal{A}$
- If $A, B \in \mathcal{A}$ then $A \cap B \in \mathcal{A}$
- If $A, B \in \mathcal{A}$ then $A \cup B \in \mathcal{A}$

Partition Rule

Let A and B be events in a outcome set S.

Partition rule

$P(A) = P(A \cap B) + P(A \cap B^c)$



Partition Rule

Example: Roll a pair of fair dice

$$P(\text{Total of 10})$$
$$= P(\text{Total of 10 and double}) + P(\text{Total of 10 and no double})$$
$$= \frac{1}{36} + \frac{2}{36} = \frac{3}{36} = \frac{1}{12}$$

Complement Rule

Complement rule:

$P(A^c) = 1 - P(A)$

Example: Often useful for events of the type "at least one" or "at least as large as some small number"

$$P(\text{Total is at least 4}) = 1 - P(\text{Total is less than 4}) = 1 - \frac{3}{36} = \frac{33}{36}$$

Containment Rule

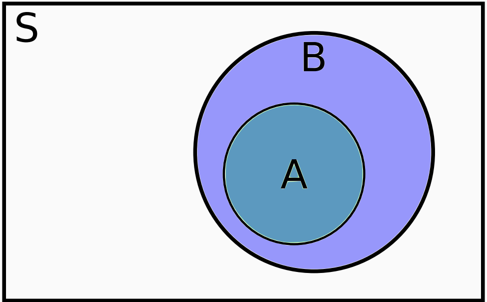
Let A and B be events in an outcome set S.

Containment rule

$P(A) \leq P(B)$ for all $A \subseteq B$

Example: Compare "two ones" with "any double",

$$\frac{1}{36} = P(\text{two ones}) \leq P(\text{any double}) = \frac{6}{36} = \frac{1}{6}$$



Inclusion Exclusion Formula

Inclusion exclusion formula

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

Example: Roll a pair of fair dice

$$
\begin{aligned}
&P(\text{Total of 10 or double}) \\
&\quad = P(\text{Total of 10}) + P(\text{Double}) - P(\text{Total of 10 and double}) \\
&\quad = \frac{3}{36} + \frac{6}{36} - \frac{1}{36} = \frac{8}{36} = \frac{2}{9}
\end{aligned}
$$

The two events are Total of $10 = \{(4,6),(6,4),(5,5)\}$ and

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

The intersection is Total of 10 and double $= \{(5,5)\}$.

Adding the probabilities for the two events, the probability for the event $\{(5,5)\}$ is added twice, so we need to subtract this probability back out.

Inclusion Exclusion