2.4. Probability spaces

Axiom 2.1 (Probability)

Probabilities are represented by real numbers between 0 and 1, inclusive.

Axiom 2.2 (Probability)

The probability that *some* outcome occurs is 1.

Axiom 2.3 (Probability)

The probability of one or the other of two *disjoint* events occurring is the sum of the individual probabilities.

Definition 2.3

A *probability space* consists of three things:

1. A set S called the *sample space*.
2. A collection of subsets of S, called *events*.
3. A *probability measure* P.

Definition 2.4

Let S be a sample space.

- The elements of S are called *sample points*.
- Subsets of S are called *events*.
- An event containing just one sample point is called a *simple event*. All other events are called *compound events*.

Theorem 2.1 (Properties of Events)

Let S be a sample space.

1. If A is an event, then so too is its complement $S \setminus A$.
2. The entire sample space itself is always an event, and so is the empty set $\emptyset$.
3. If A and B are events, then so too is the union $A \cup B$ and intersection $A \cap B$.
4. In fact, if $A_1, A_2, A_3, \ldots$ is an infinite sequence of events, then the infinite union

   $$A_1 \cup A_2 \cup A_3 \cup \cdots$$

   and the infinite intersection

   $$A_1 \cap A_2 \cap A_3 \cap \cdots$$

   are also events.

Definition 2.5

Let S be a sample space. A *probability measure* P (also called a *probability distribution*) is a function that to each event A in S assigns a number $P(A)$, called the *probability of* A, subject to the following axioms:

1. $P(A) \geq 0$ for all events A.
2. $P(S) = 1$.
3. If $A_1, A_2, A_3, \ldots$ is a sequence of pairwise disjoint events in S, then

$$P(A_1 \cup A_2 \cup A_3 \cup \cdots) = \sum_{i=1}^{\infty} P(A_i).$$

$$P(\text{event}) = \text{probability}$$

input

output

Theorem 2.2 (Properties of Probability Measures)

Let P be a probability measure on a sample space S.

1. If A is an event, then $P(A) \leq 1$.
2. If A and B are events and $A \subset B$, then $P(A) \leq P(B)$.
3. If A is an event, then $P(A^c) = 1 - P(A)$.
4. If A and B are disjoint events, then $P(A \cup B) = P(A) + P(B)$.
5. We have $P(\emptyset) = 0$.

Definition 2.6

Let S be any set and $p : S \to \mathbb{R}$ a function.

1. The *support* of p is the set of all points $s \in S$ where $p(s)$ is nonzero, i.e., it is the set

$$\{s \in S : p(s) \neq 0\}. \tag{2.1}$$

2. The function p is said to have *discrete* support if its support (2.1) is either finite or countably infinite.

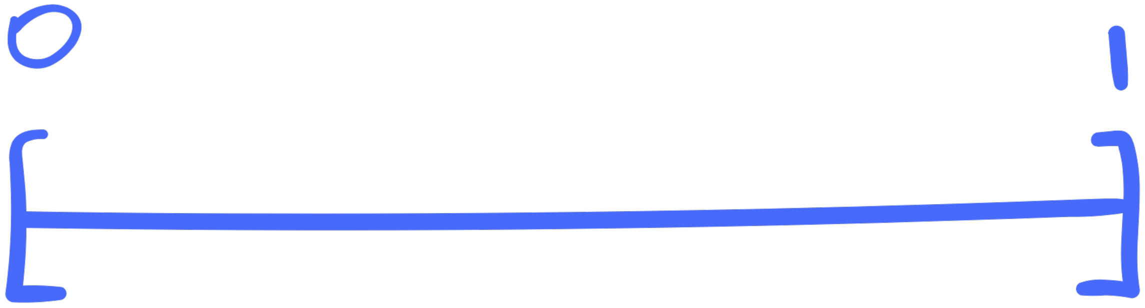
Definition 2.7 (sort of)

A set A is *countably infinite* if, given an infinite amount of time, I could count the elements of A one at a time, counting one element per second.

countably infinite

uncountably infinite

all numbers between 0 and 1



one!

but where is "two"?!

Definition 2.8

Let P be a probability measure on a sample space S. We shall say P is *discrete* if every subset $A \subset S$ is an event and there is a function $p : S \to \mathbb{R}$ with discrete support such that

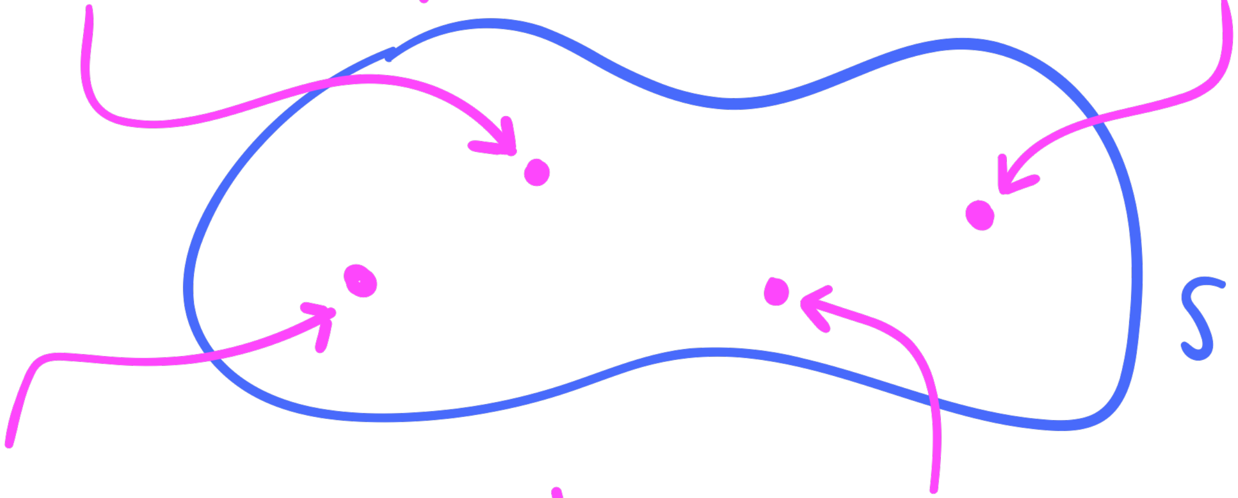
$$P(A) = \sum_{s \in A} p(s), \tag{2.2}$$

for all events A. In this case, the function p is called the *probability mass function* (PMF) of the probability measure P (or sometimes just the *probability function*), and S is called a *discrete probability space* (when equipped with P).

Discrete probability space

you get probability $\frac{1}{8}$

you get probability $\frac{1}{4}$



(only 4 sample points)

you get probability $\frac{1}{4}$

you get probability $\frac{3}{8}$

$$\frac{1}{8} + \frac{1}{4} + \frac{3}{8} + \frac{1}{4} = 1 \checkmark$$

Definition 2.8

Let P be a probability measure on a sample space S. We shall say P is *discrete* if every subset $A \subset S$ is an event and there is a function $p : S \to \mathbb{R}$ with discrete support such that

$$P(A) = \sum_{s \in A} p(s), \tag{2.2}$$

#

for all events A. In this case, the function p is called the *probability mass function* (PMF) of the probability measure P (or sometimes just the *probability function*), and S is called a *discrete probability space* (when equipped with P).

- If you want to compute the probability $P(A)$ of an event A, simply find all sample points $s \in A$, compute $p(s)$ for each of these $s \in A$, and then add them all up.

Theorem 2.3 (Properties of Probability Mass Functions)

Let $p(s)$ be the probability mass function of a discrete probability measure P. Then:

1. $p(s) \geq 0$ for all $s \in S$, and
2. $\sum_{s \in S} p(s) = 1$.

Theorem 2.4 (Discrete Probability Construction Lemma)

Let S be a set and $p : S \to \mathbb{R}$ a function with discrete support. If

1. $p(s) \geq 0$ for all $s \in S$, and
2. $\sum_{s \in S} p(s) = 1$,

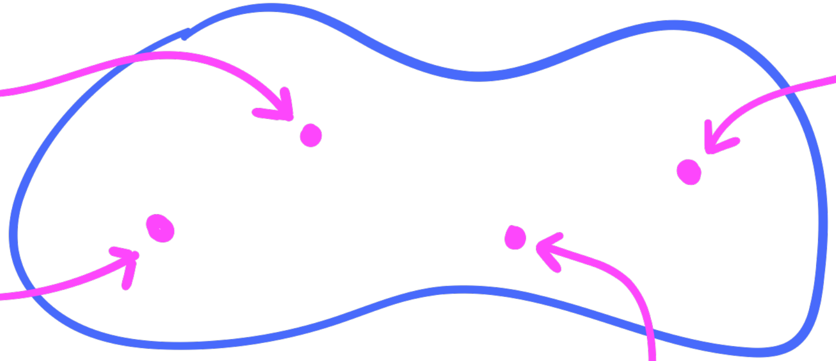
then there is a unique discrete probability measure P on S such that

$$P(A) = \sum_{s \in A} p(s)$$

for all $A \subset S$.

uniform probability space

you get probability $\frac{1}{4}$ you get probability $\frac{1}{4}$



S (only 4 sample points)

you get probability $\frac{1}{4}$ you get probability $\frac{1}{4}$

$$\frac{1}{4}+\frac{1}{4}+\frac{1}{4}+\frac{1}{4}=1 \checkmark$$

all probabilities equal

Definition 2.9

Let P be a discrete probability measure on a sample space S with probability mass function $p(s)$. Then P is called a *uniform probability measure* if the support of $p(s)$ has finite cardinality $n > 0$, and if

$$p(s) = \frac{1}{n}$$

for each s in the support of p.