Week 5

Mathematics and Computational Methods for Complex Systems, 2023

Dhruva V. Raman

Probability

Experiments

A process with
measured, uncertain outcomes

Data scientist
Measure population's credit card scores

Climate scientist
Measure tomorrow's weather

and they all need…

Experiments

Mathematical framework to deal with
characterising and quantifying
uncertain events

= Lecture part 1

Terminology we will learn about:

1. Probability spaces

Mathematical setting for experiments

2. Random variables

Quantities within an experiment that
depend on uncertain events

Terminology we will learn about:

Necessary to properly understand random variables and their properties

1. Probability spaces

Mathematical setting for experiments

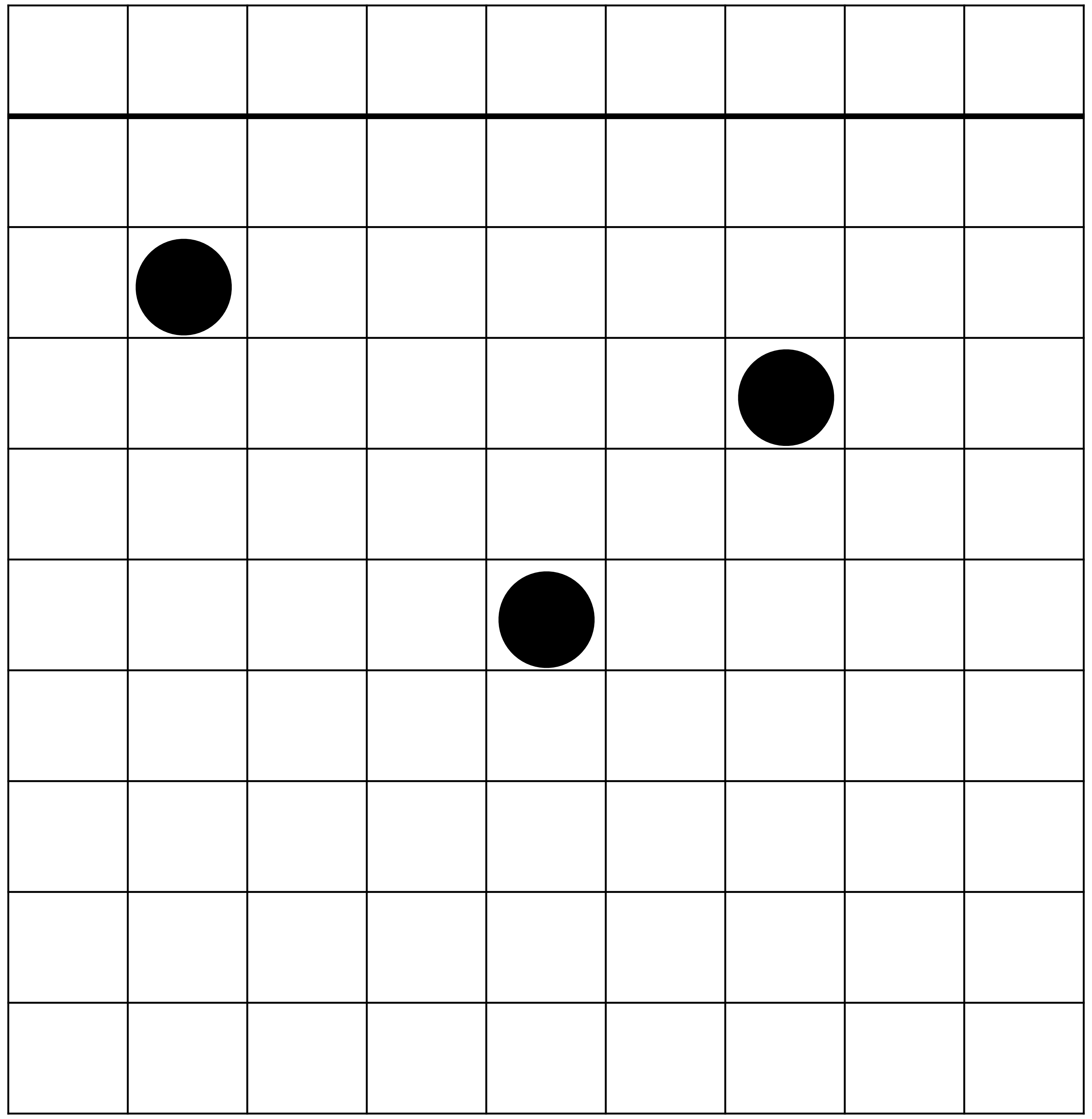
Necessary in basically any quantitative subject

2. Random variables

Quantities within an experiment that depend on uncertain events

Lecture theatre next week

● = Empty seat



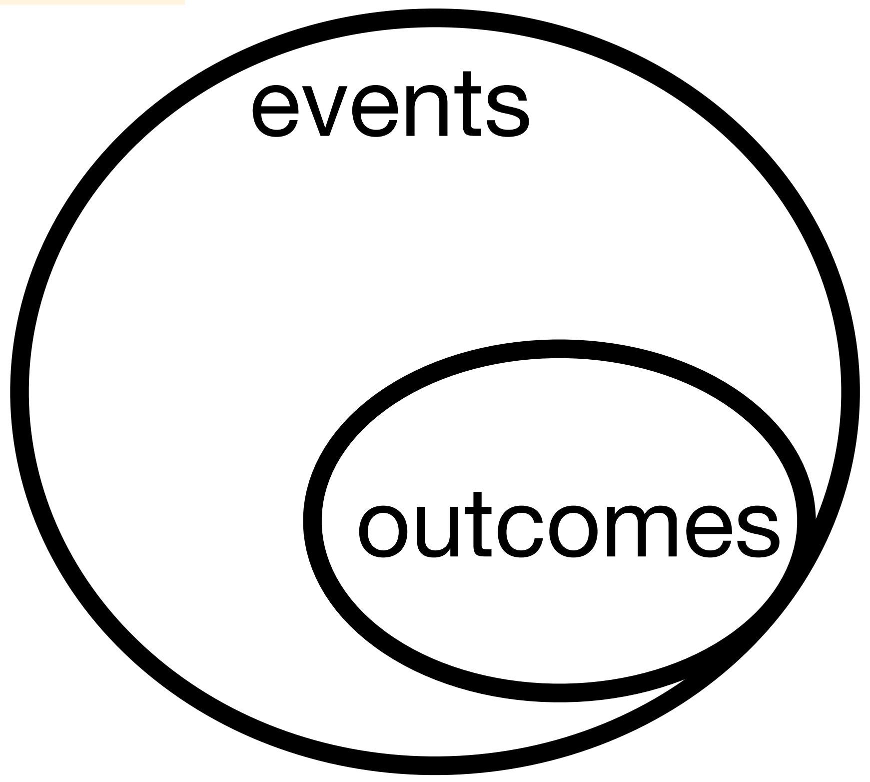
?

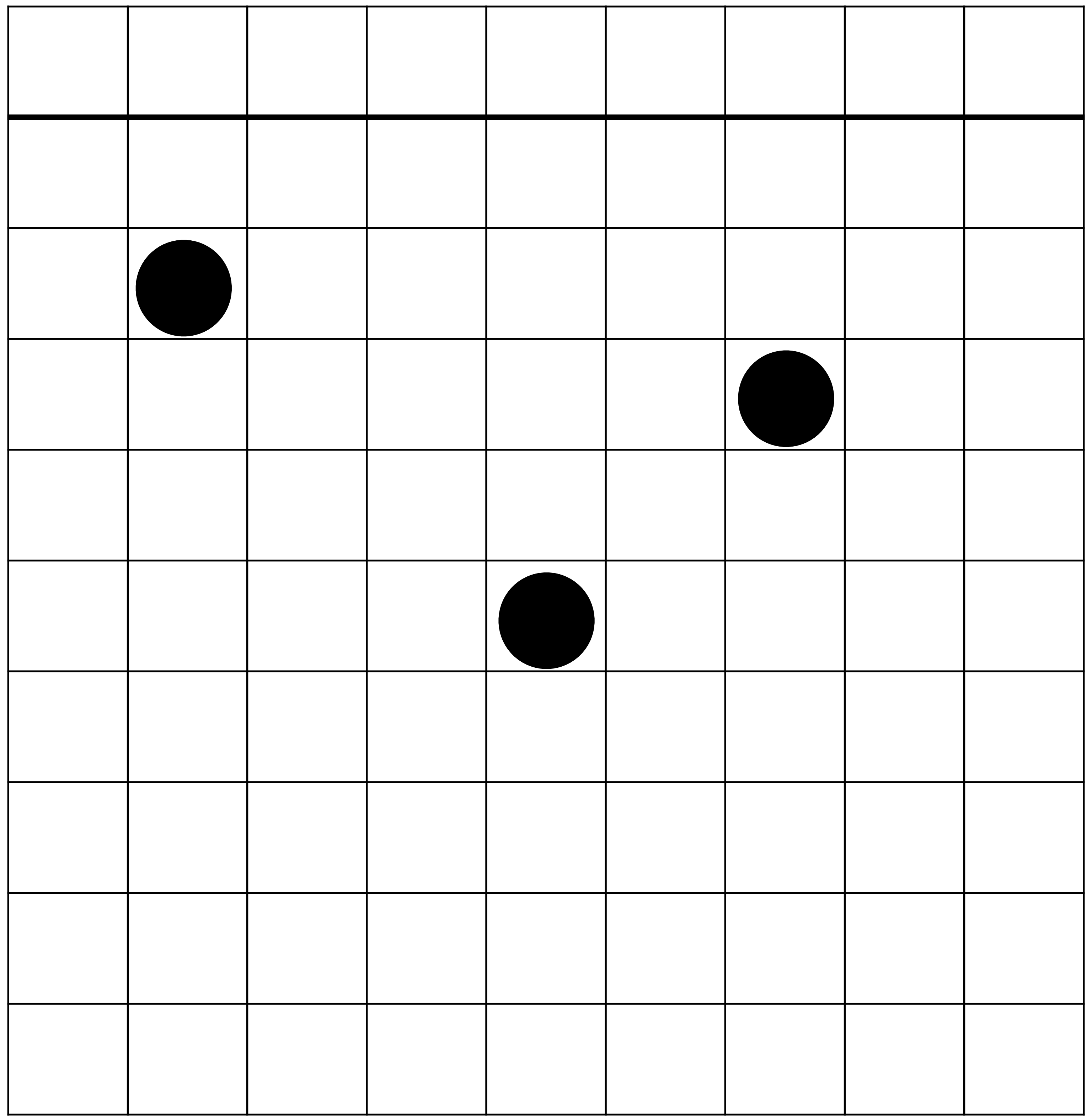
Lecture theatre next week

Formal terminology

Each seating combination is an outcome

Combinations of outcomes are events

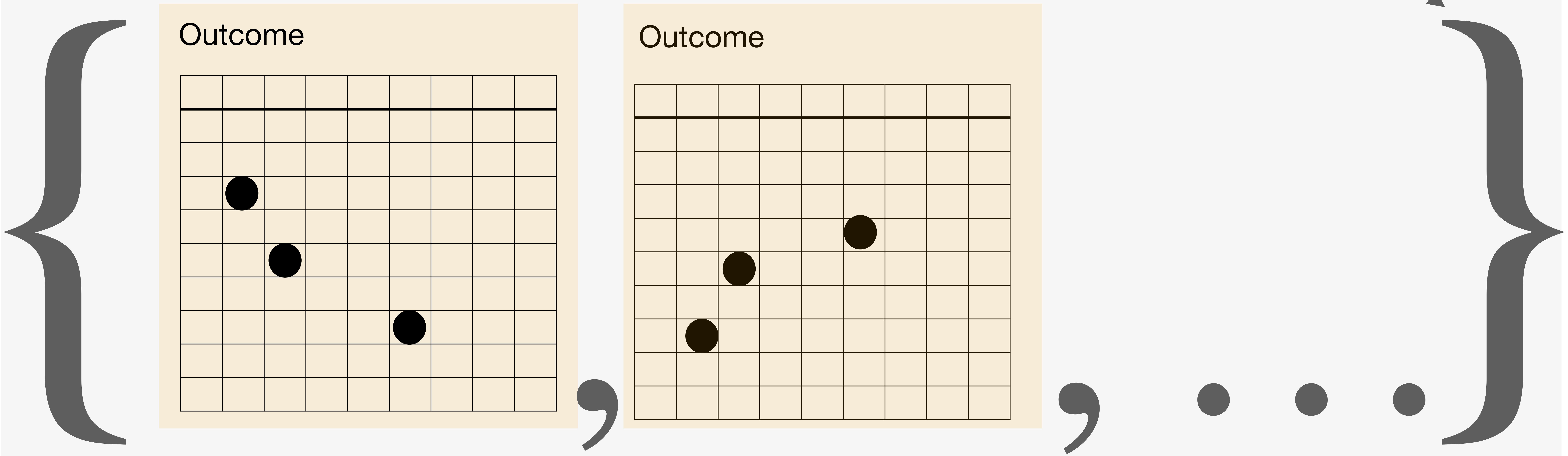




Events are sets of outcomes

Curly braces denote sets in maths notation

Event: the back row is completely filled



The set of events is the power set of outcomes

Power set of S

The set of subsets of S

$$S = \{1,2,4\}$$

$$\mathscr{P}(S) = \{\varnothing, \{1,2\}, \{2,4\}, \{4\}, \{1\}, \ldots\{1,2,4\}\}$$

…since any combination of outcomes is an event

The set of events is the power set of outcomes

Power set of S

The set of subsets of S

$$S = \{1,2,4\}$$

$$\mathscr{P}(S) = \{\varnothing, \{1,2\}, \{2,3\}, \{3,4\}, \{1\}, \ldots \{1,2,3,4\}\}$$

Homework

Cardinality of the power set of a set with n elements?

Induction is your friend!

Events have an algebra

Algebra

Rules for elements to interact with each other

…and make babies!

Events have an algebra

Algebra

Rules for elements to interact with each other

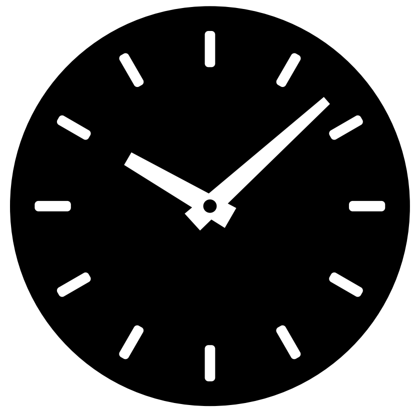
…and make babies!

Recap:
algebra of vector spaces

Closure under addition

$$x + y \in V \quad \forall x, y \in V$$

$$\vdots$$



"plus, minus, times"

Events have an algebra

Algebra

Rules for elements to interact with each other

…and make babies!

Events

$$X \cup Y = Z$$ **Union (or)**

$$X \cap Y = Z$$ **Intersection (and)**

$$X \backslash Y = Z$$ **Complement (not/without)**

Questions for the audience

Event X: the back row is completely filled

Event Y1: the front row is completely filled

Event Y2: the back row is partially filled

Express as grammatically correct sentences:

$X \cap Y_1$ $\quad$ $X \cap Y_2$

$X \cup Y_1$ $\quad$ $X \cup Y_2$

$X \backslash Y_1$ $\quad$ $X \backslash Y_2$

- can any of these be simplified?

Questions for the audience

Event X: the back row is completely filled

Event Y1: the front row is completely filled

Event Y2: the back row is partially filled

Express as grammatically correct sentences:

$X \cap Y_1$ $\qquad$ $X \cap Y_2 = X$

$X \cup Y_1$ $\qquad$ $X \cup Y_2 = X$

$X \backslash Y_1$ $\qquad$ $X \backslash Y_2 = \varnothing$

The empty set. Learn this notation!

Each event has a probability

Event: the back row is completely filled

$\downarrow$

Probability 0.3

Probability function

$$\mathbb{P} : \text{events} \to [0,1]$$

$$0 \leq \mathbb{P}(x) \leq 1$$

What does it mean to have a probability?

Probability is a way of expressing partial knowledge of an event

With enough knowledge and insight, could I predict tomorrow's weather without uncertainty?

✗ Correct probability

✓ Best probability given my knowledge of the world

Probability Space

is three things:

$$(\Omega, \mathscr{F}, \mathbb{P})$$

Sample space:
the set containing all possible outcomes

Event space:
the set containing all possible events (combinations of outcomes)

Probability function:
Assigns a probability to every single event

Probability Space

for next week's lecture seating:

$$(\Omega, \mathscr{F}, \mathbb{P})$$

Sample space:
Every possible seating combination

Event space:
Sets of seating combinations

Example set:
All seating combinations where the back row is filled

Probability function:
Assigns a probability to every single event

Probability Space

for next week's lecture seating:

…depends upon what I'm modelling

$$(\Omega, \mathscr{F}, \mathbb{P})$$

Homework

Sample/event space if I was also interested in who is sitting where?

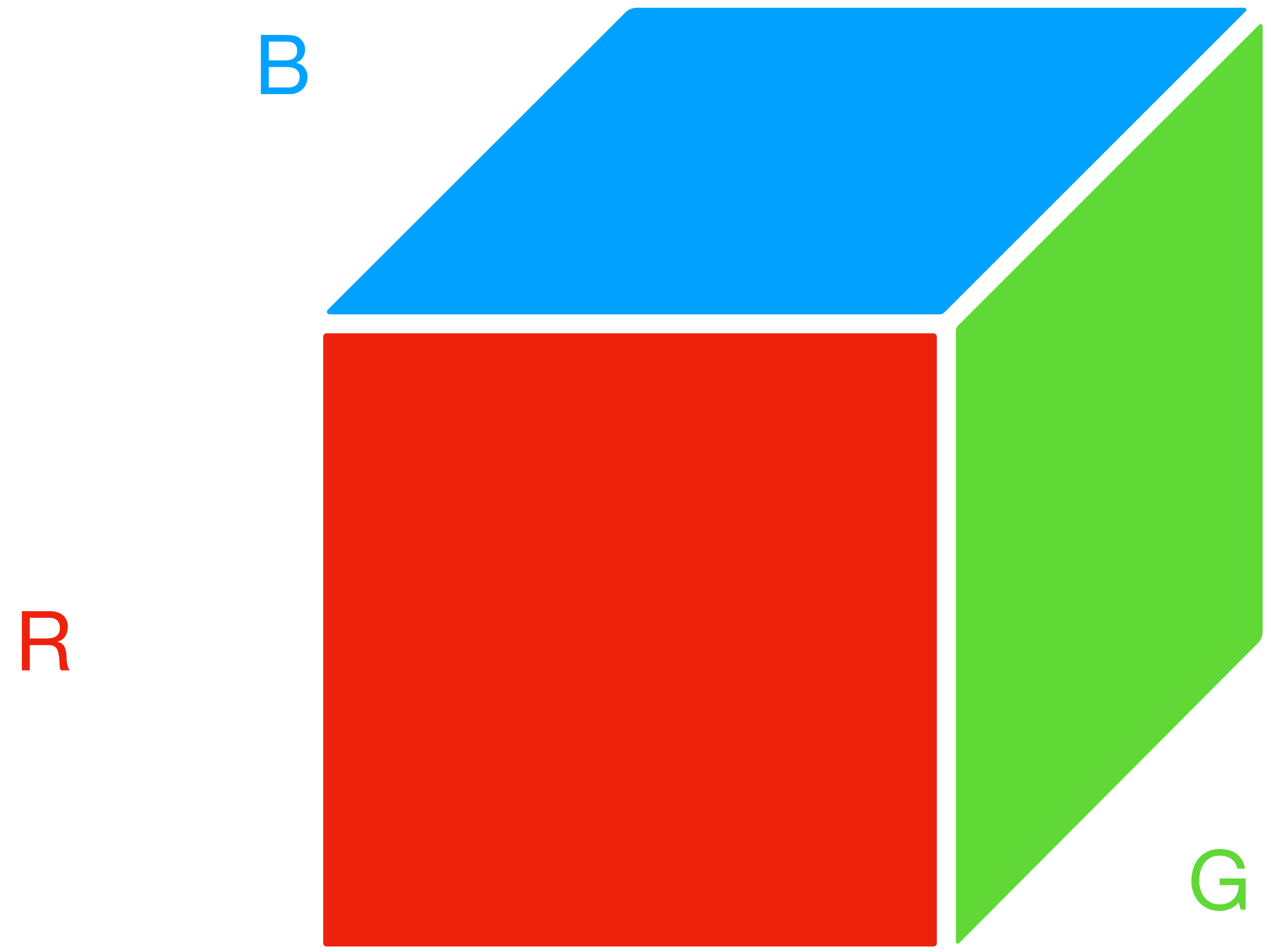
Generating a random colour

$$\Omega = \{\omega : \underset{R}{\omega_1}, \underset{G}{\omega_2}, \underset{B}{\omega_3} \in [0,1]\}$$

Example events
can be drawn

“Red is saturated”

“Blue > 0.5”



Generating a random colour

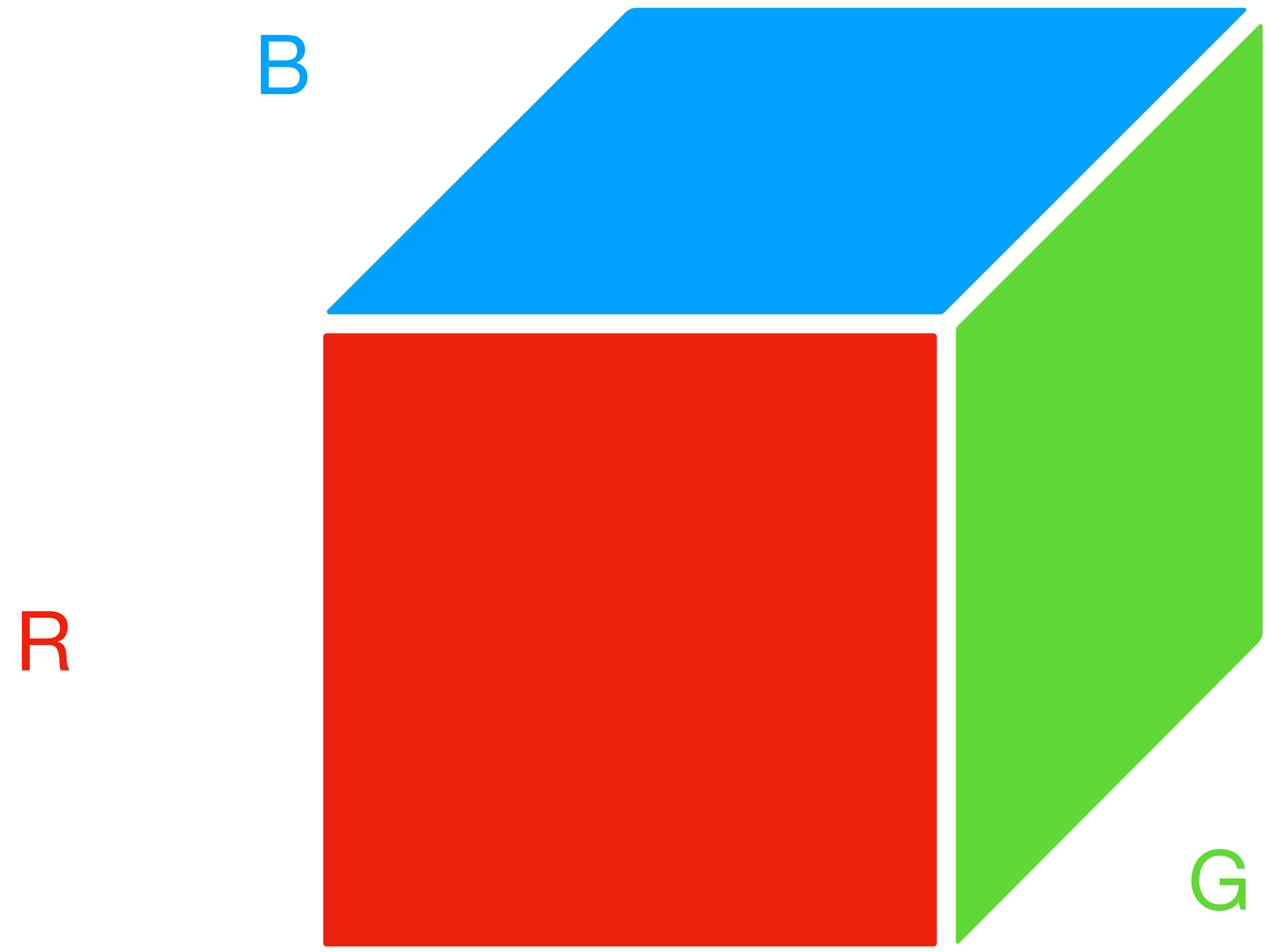
$$\Omega = \{\omega : \omega_1, \omega_2, \omega_3 \in [0,1]\}$$

R G B

Example events
can be drawn

"Red is saturated"

"Blue > 0.5"



Events have a 'volume'!!!

Volumes and probabilities are both measures

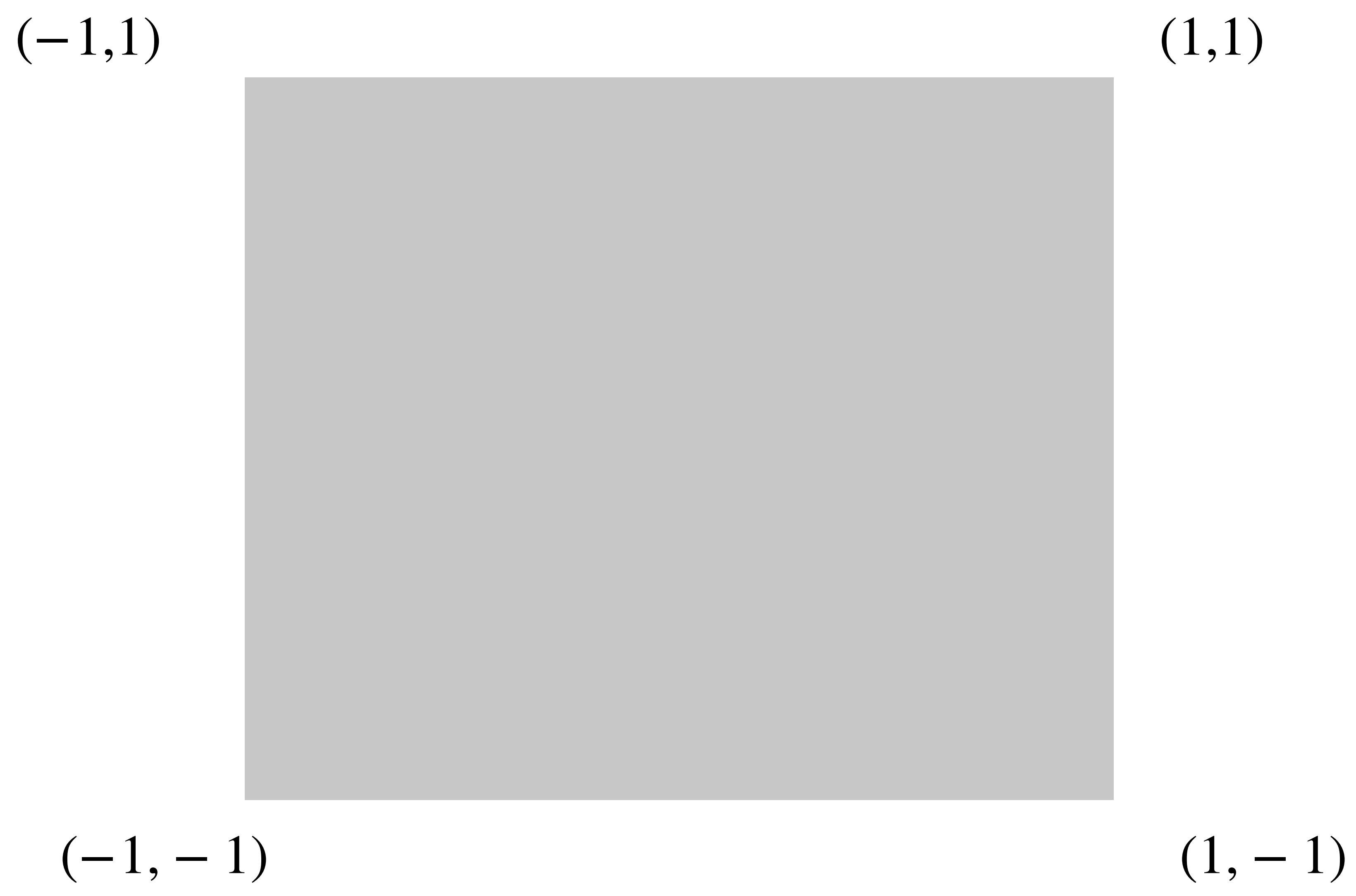
also lengths, areas, etc …

How "big" is an event? $\Leftrightarrow$ How probable is an event?

(According to some notion of volume)

Standard measure: $x \times y$

Another measure: $2x \times \frac{1}{3}y$



What makes a measure?

Optional material

(Practise reading maths)

$\mathbb{P} : \mathscr{F} \to \mathbb{R}$ is a measure on $\mathscr{F}$ if:

$$\mathbb{P}[\varnothing] = 0$$

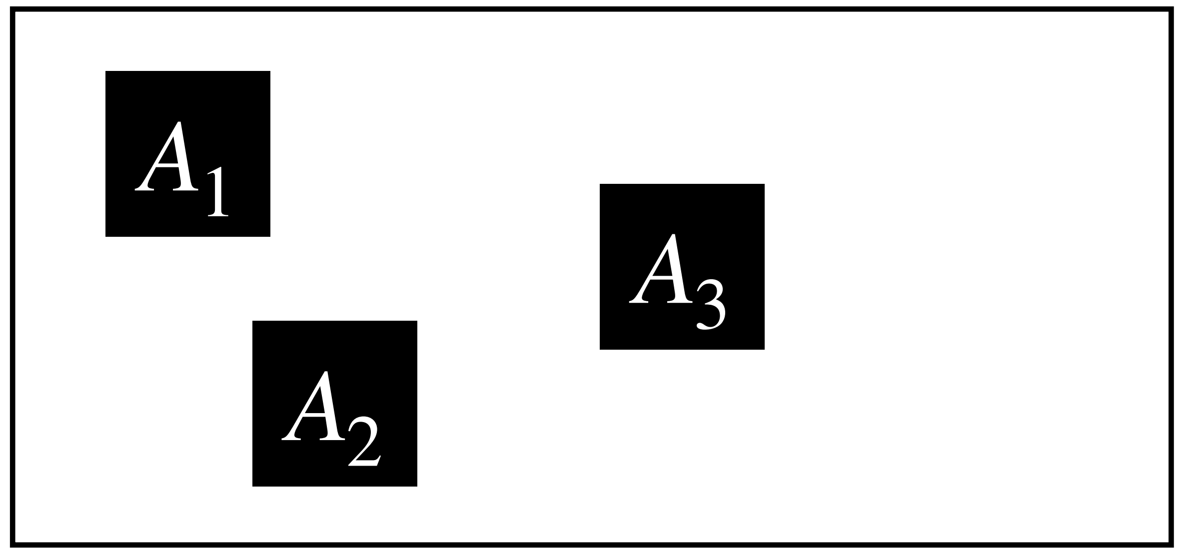
Non-negativity

$$\mathbb{P}[f] \geq 0 \ \forall f \in \mathscr{F}$$

Countable additivity for disjoint sets

$$\mathbb{P}\Big[\bigcup_{i=1}^{n} A_i\Big] = \sum_{i=1}^{N} \mathbb{P}[A_i]$$

$$A_i \cap A_j \quad \forall i, j \in \{1, \ldots, n\}$$



Total measure = sum of individual measures

$\Leftrightarrow$

What makes a measure?

Optional material

(Practise reading maths)

$\mathbb{P} : \mathscr{F} \to \mathbb{R}$ is a measure on $\mathscr{F}$ if:

$$\mathbb{P}[\varnothing] = 0$$

Non-negativity

$$\mathbb{P}[f] \geq 0 \ \forall f \in \mathscr{F}$$

Countable additivity for disjoint sets

$$\mathbb{P}\Big[\bigcup_{i=1}^{n} A_i\Big] = \sum_{i=1}^{N} \mathbb{P}[A_i]$$

$$A_i \cap A_j \quad \forall i, j \in \{1, \ldots, n\}$$

Probability measures also require:

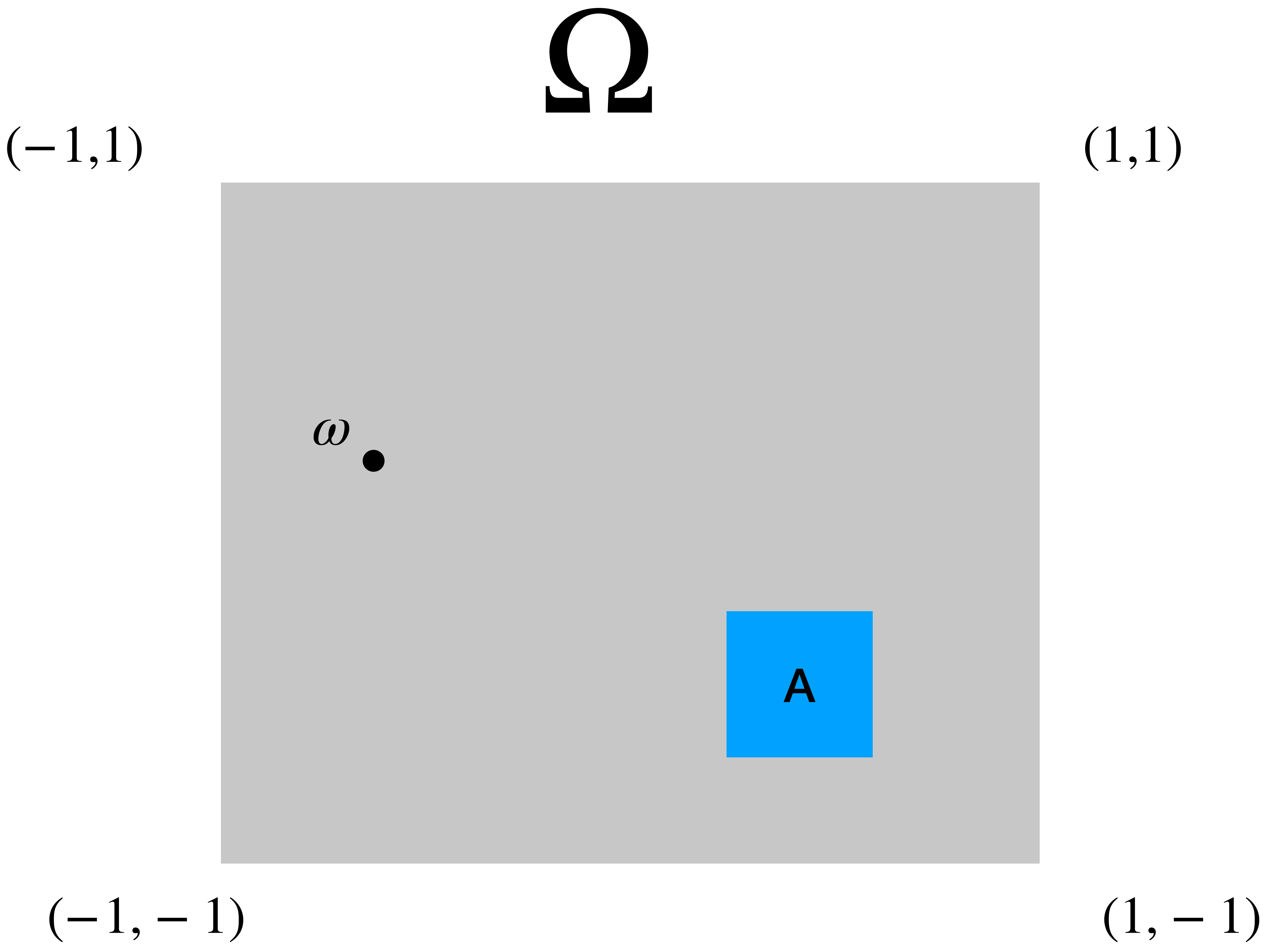
Normalised:

$$\mathbb{P}[\Omega] = 1$$

Measuring sets

$$\Omega = \{\omega : \omega_1, \omega_2 \in [-1,1]\}$$

Area of A? Easy!



Measuring sets

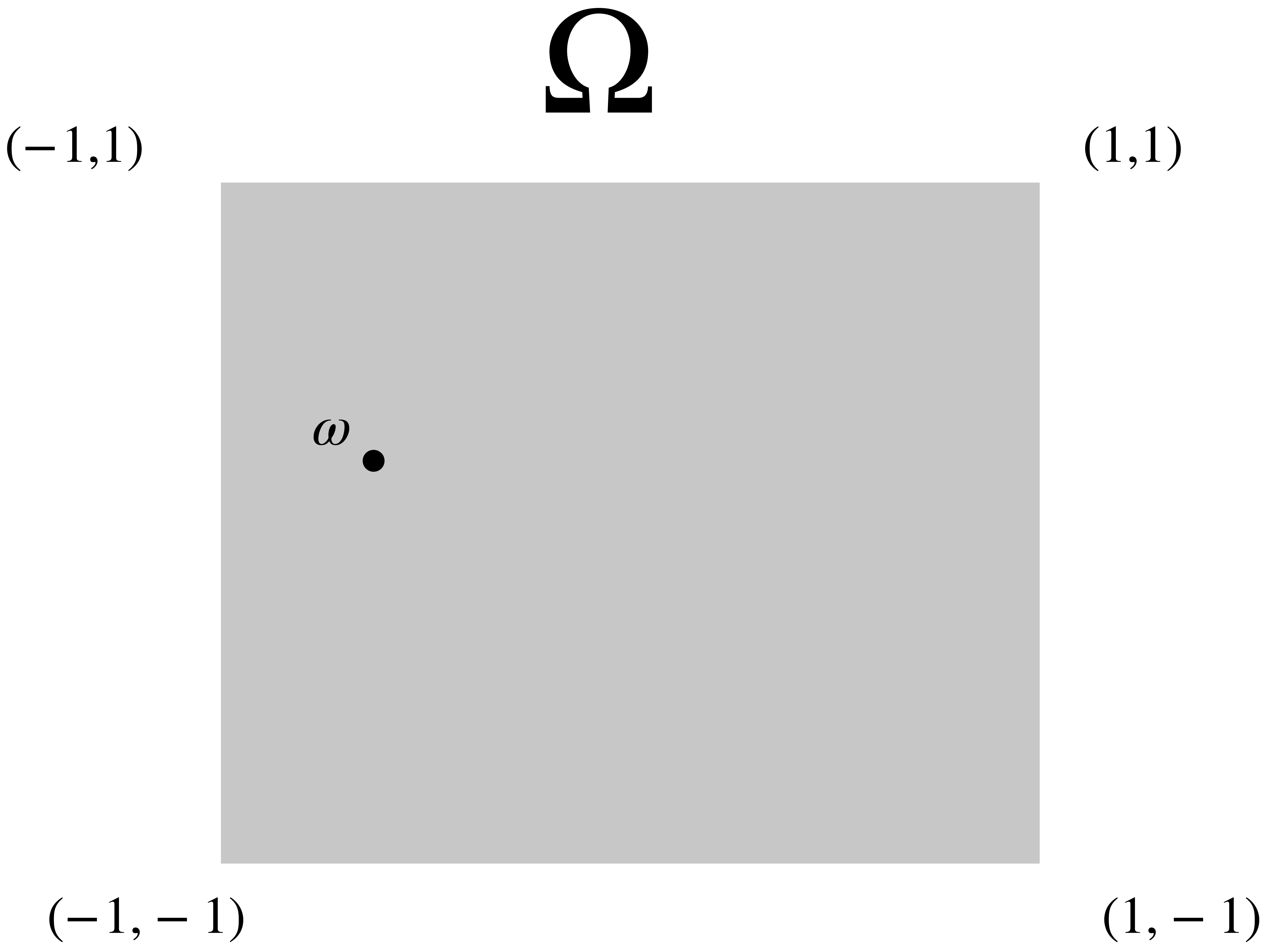
$$\Omega = \{\omega : \omega_1, \omega_2 \in [-1,1]\}$$

$$A = \mathbb{Q}^2 \cap \Omega$$

(Set of rationals in the square)

$$\mathbb{P}[A]?$$

….i.e. area of A?



Measuring sets

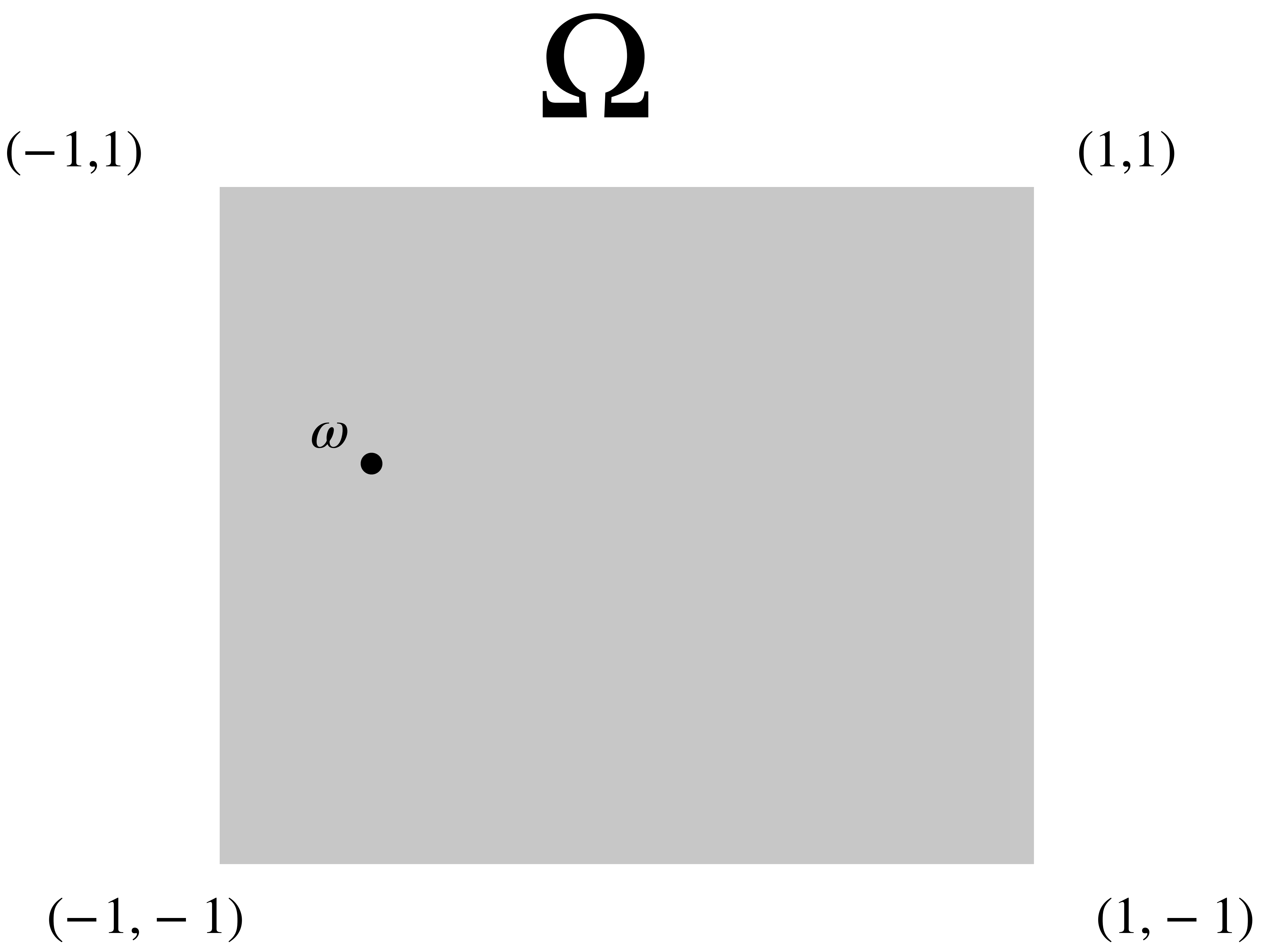
$$\Omega = \{\omega : \omega_1, \omega_2 \in [-1,1]\}$$

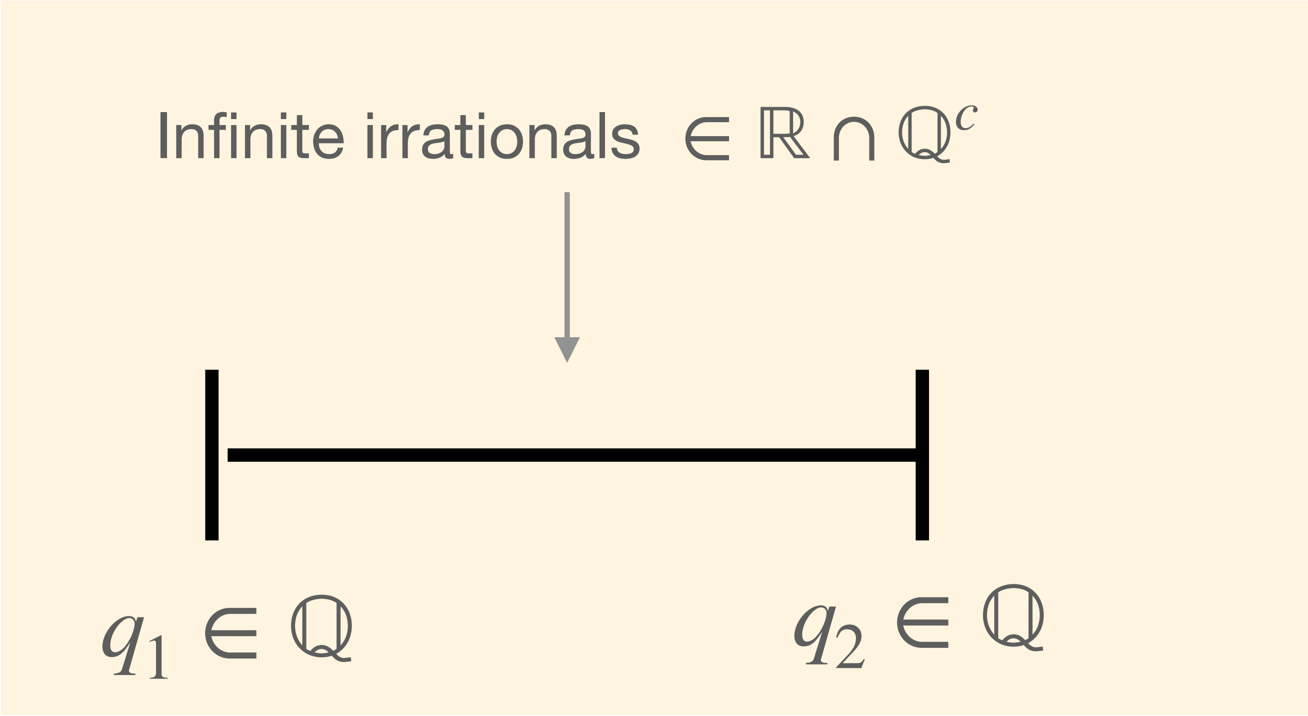
$$A = \{\omega : \omega_1, \omega_2 \in \mathbb{Q}\} \cap \Omega$$

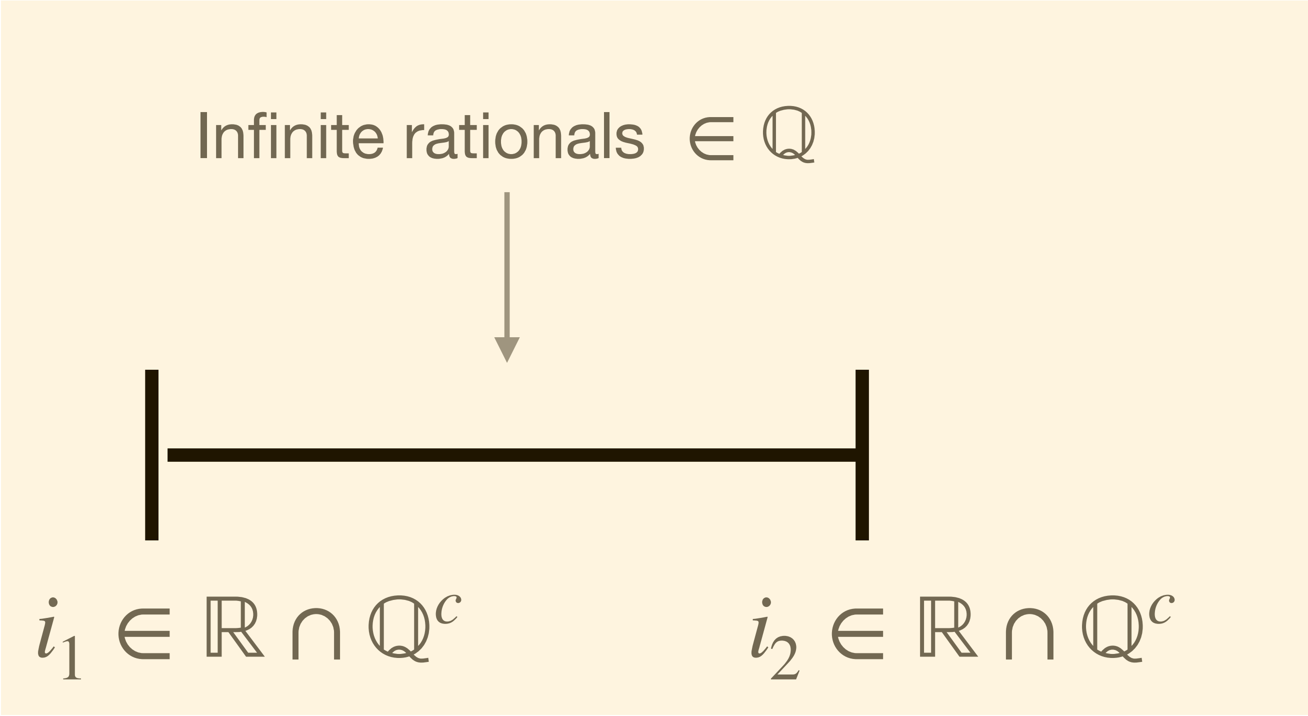
(Set of rationals in the square)

$$\mathbb{P}[A] = 0$$

(but not obvious or logical)







Banach Tarski paradox

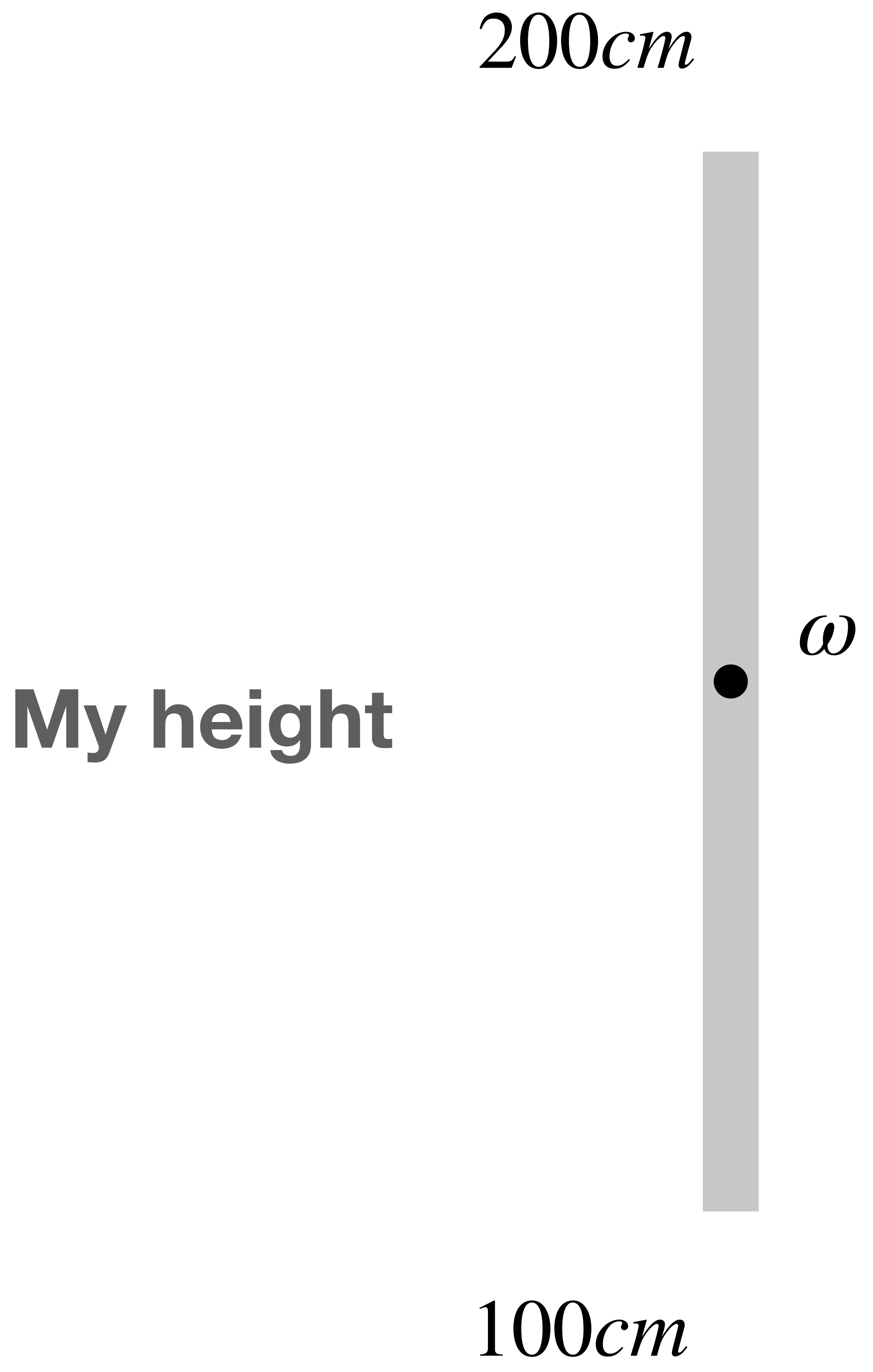
…you can't measure everything!

Some events can't have a probability

What's the probability our heights satisfy the following property:

$$V = \{v : \forall r \in \mathbb{R}, \ \ \exists! v \in V : v - r \in \mathbb{Q}\}$$

(Vitali set, don't worry if uninterested)



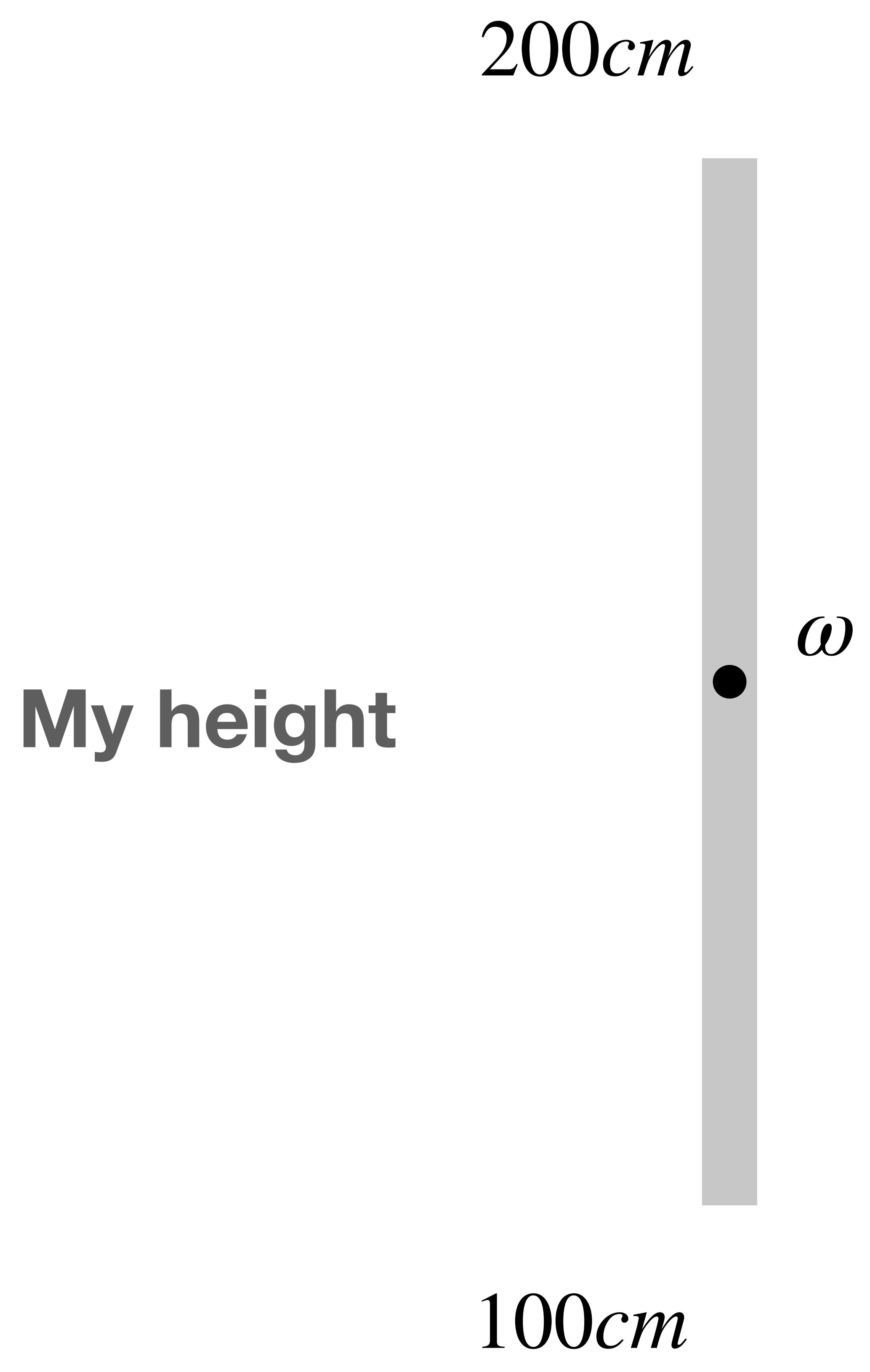
Some events can't have a probability

What's the probability our heights satisfy the following property:

$$V = \{v : \forall r \in \mathbb{R}, \quad \exists! v \in V : v - r \in \mathbb{Q}\}$$

(Vitali set, don't worry if uninterested)

No answer to this question!!!



How do we know what events we can measure the probability of?

Events form a $\sigma-$algebra

(i.e. a space of safe-to-measure things)

Ω

Any set (e.g. outcomes)

$\mathcal{F}$

Any set of sets (e.g. events)

$\mathcal{F}$ **is a $\sigma-$algebra on Ω if:**

$$\Omega \in \mathcal{F}$$

Closure under complements

$$S \in \mathcal{F} \Rightarrow S^c \in \mathcal{F}$$

Closure under countable unions

$$\{A\}_{i=1}^{n \in \mathbb{N}} \in \mathcal{F} \Rightarrow \bigcup_{i=1}^{n} A_i \in \mathcal{F}$$

Probability Space

is three things:

$\sigma-$algebras are optional topic

$$(\Omega, \mathscr{F}, \mathbb{P})$$

Sample space:
the set containing all possible outcomes

Probability function:
Assigns a probability to every single event

Event space:
the ~~set~~ $\sigma-$algebra containing all possible events (combinations of outcomes)

Random variables

are quantitative questions
about the experiment

Random variables

are quantitative questions
about the experiment

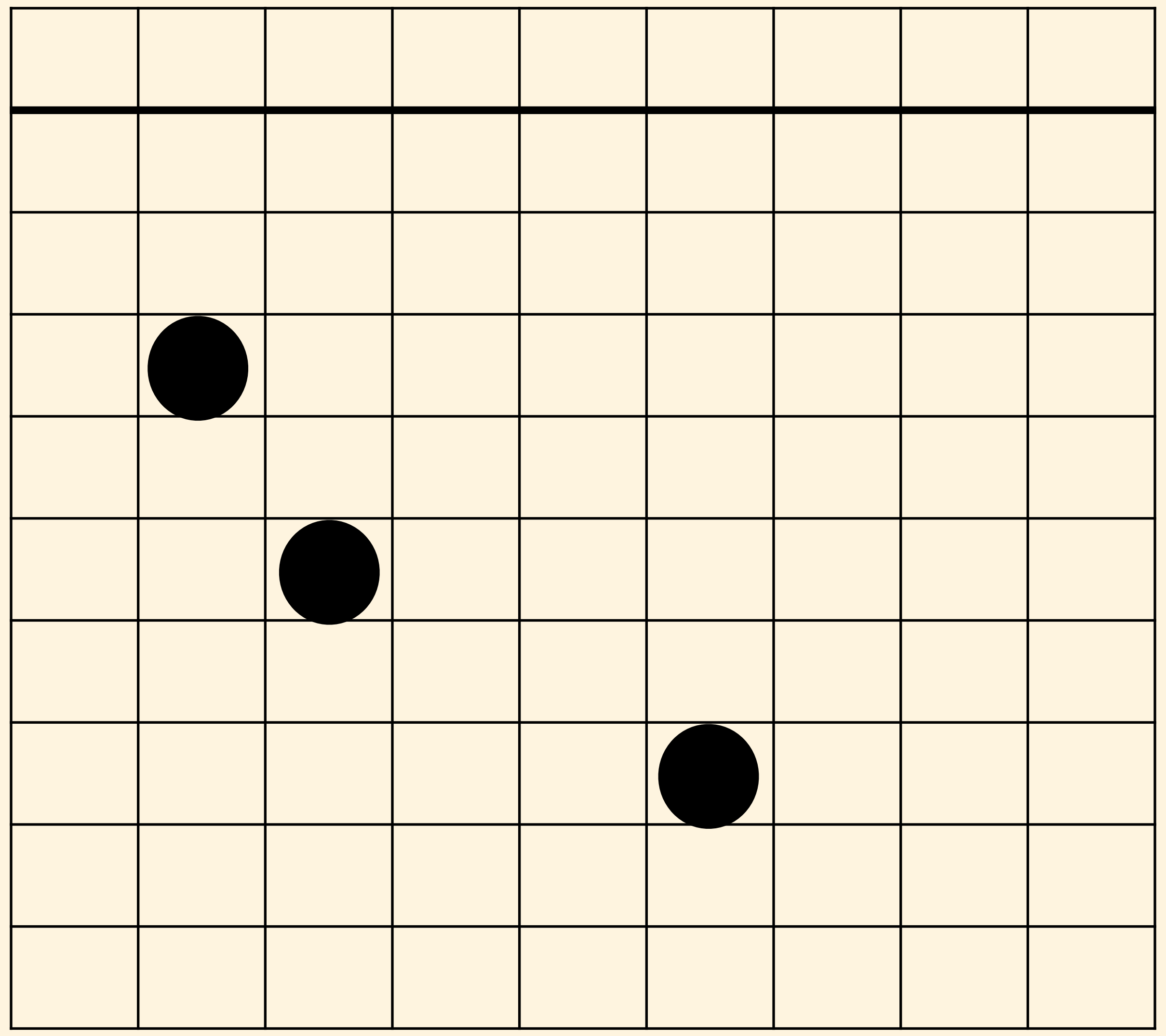
are functions that map from
outcomes to numbers

(or to any "measurable space")

Random variables example

What was the number of unfilled seats? $X(\omega)$

Outcome ω



$\longrightarrow$

Number

87

(Or any integer between
0 and *number of seats*)

Random variables example

Probability of four unfilled seats?

Correct expression

$$\mathbb{P}[\{\omega \in \Omega : X(\omega) = 4\}]$$

Shorthand

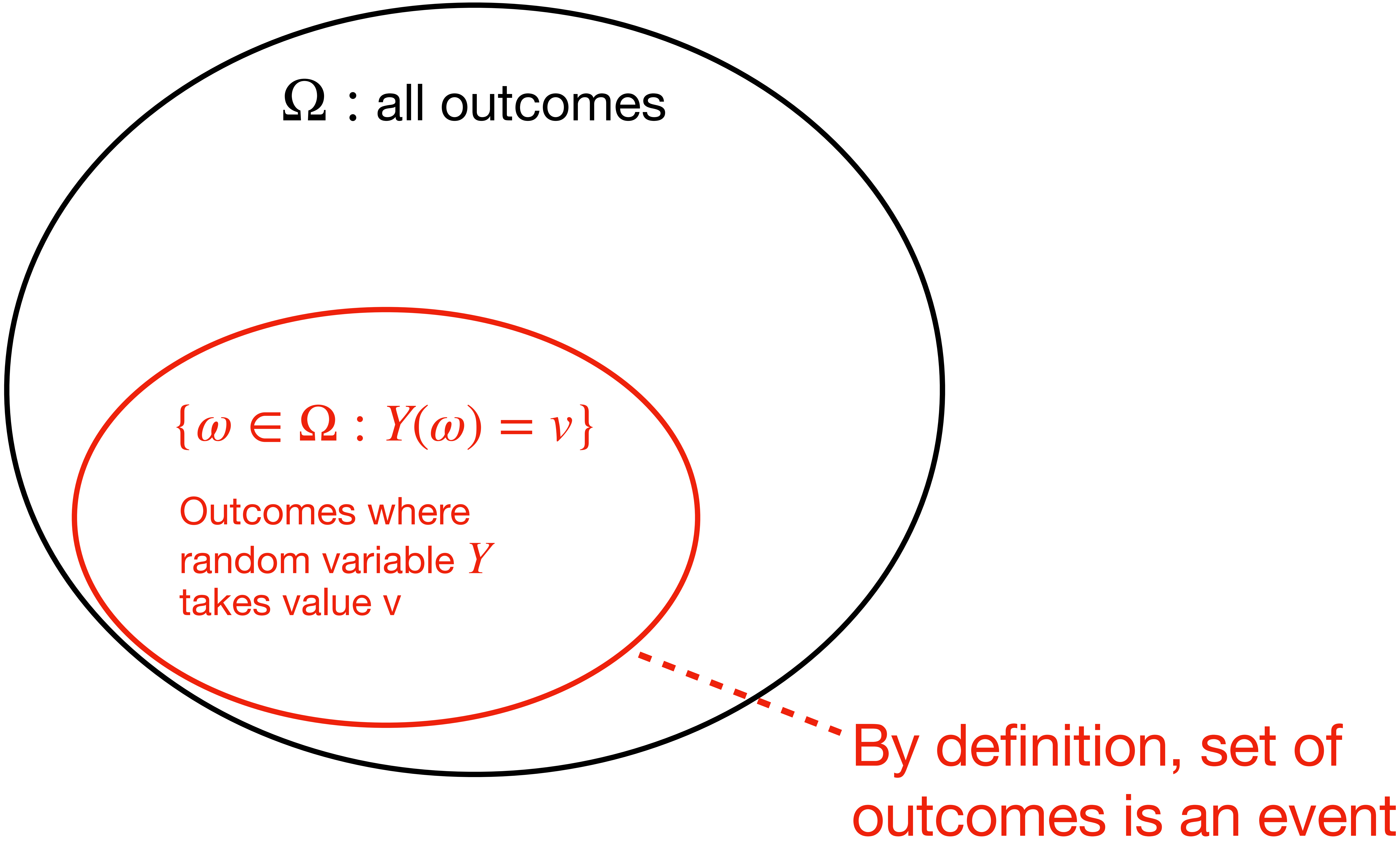
$$\mathbb{P}[X = 4]$$

Random variable taking value is an event

$$\begin{tikzcd} \text{Outcome} \arrow[r] \arrow[d] & \text{Probability} \\ \text{Random variable takes value} & \end{tikzcd}$$

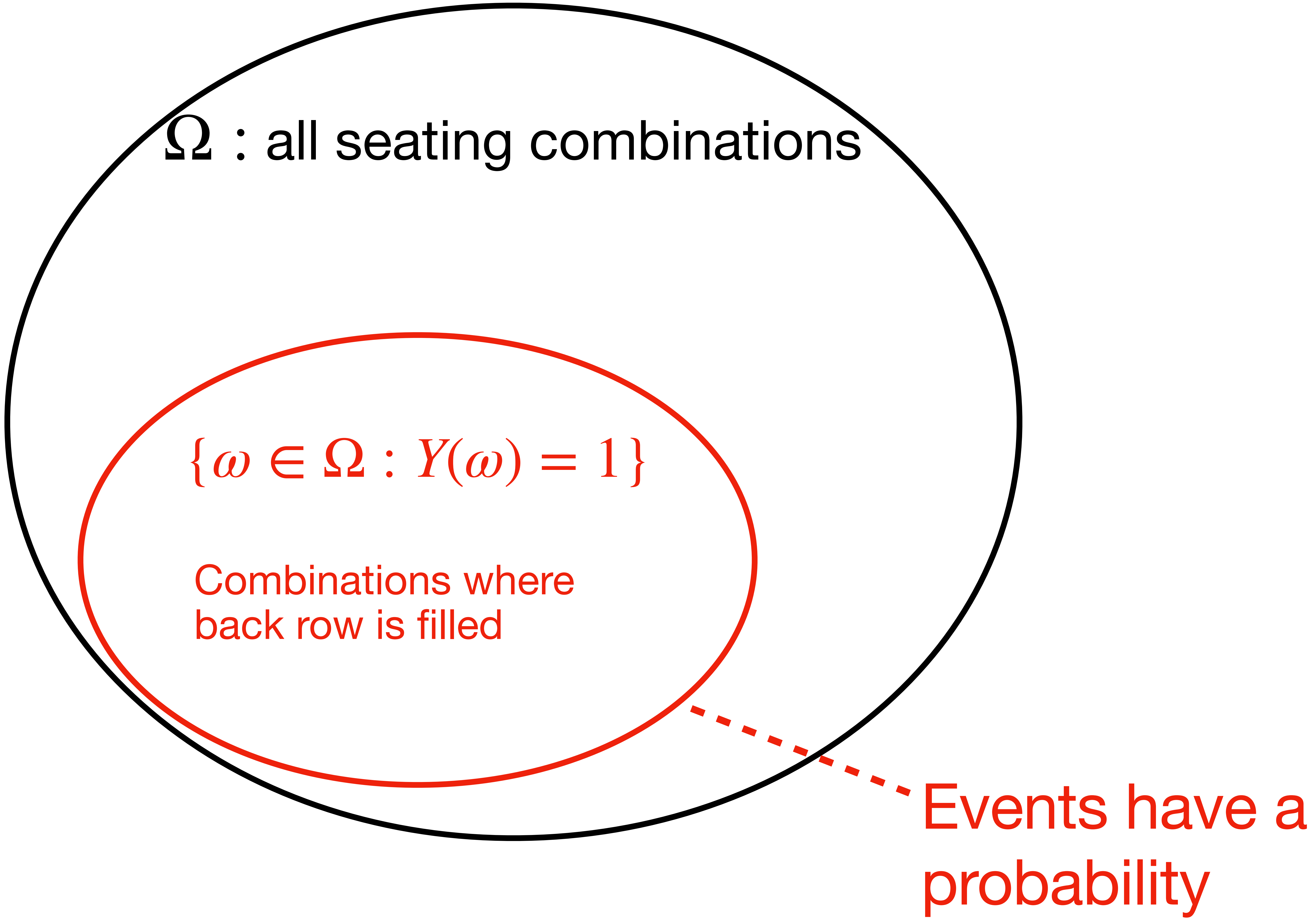
Eg $\{$ outcomes where number of unfilled seats $= 5\}$

Random variable taking value is an event



Random variable taking value is an event

Y is whether back row is filled



Support of a random variable

Set of plausible values a random variable can take

Support of a random variable

Set of plausible values a random variable can take

$\Updownarrow$

Smallest set S such that $\mathbb{P}[X \in S] = 1$

Support of a random variable

Set of plausible values a random variable can take

$\Updownarrow$

Smallest set S such that

$\mathbb{P}[X \in S] = 1$

Long form:

$\mathbb{P}[\{\omega \in \Omega : X(\omega) \in S\}] = 1$

"Set of outcomes where X takes a value inside S"

Support of a random variable

X is the number of unfilled seats

Smallest set S such that
$\mathbb{P}[X \in S] = 1$

With probability one:

$0 \leq X \leq$ number of seats
$X \in \mathbb{Z}$

$\text{supp}(X) = \{0,1,\ldots \text{number of seats}\}$

Support of a random variable

Y is whether the back row is filled

What’s the support of Y?

Support of a random variable

Y is whether the back row is filled

What's the support of Y?

$$\text{supp}(Y) = \{0,1\}$$

Support of a random variable

Z is height of 2nd person in 3rd row

What’s the support of Z?

Support of a random variable

Z is height of 2nd person in 3rd row

What's the support of Z?

$$\text{supp}(Z) = (l, h)$$

l : height of shortest person on course

h : height of tallest person on course

Two flavours of random variables

Support is finite set

Support is infinite set

Discrete random variables

X is the number of unfilled seats

$\text{supp}(X) = \{0,1,2,\ldots 200\}$

Y is whether the back row is filled

$\text{supp}(Y) = \{0,1\}$

Continuous random variables

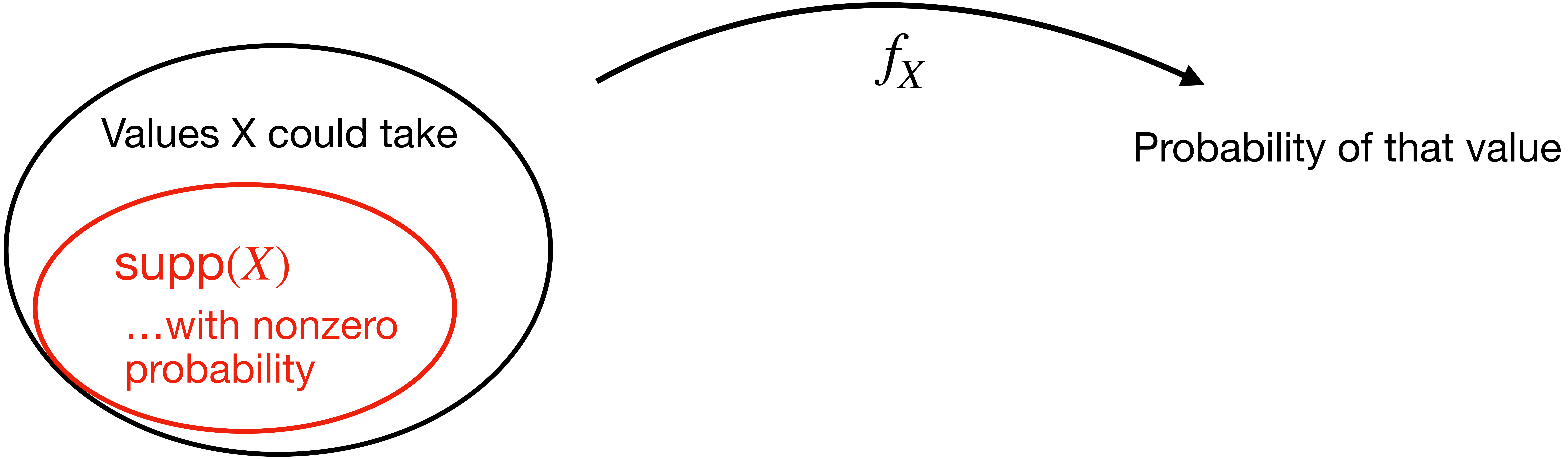
Z is height of 2nd person in 3rd row

$\text{supp}(Z) = (l, u)$

Probability mass function

…discrete random variables only!

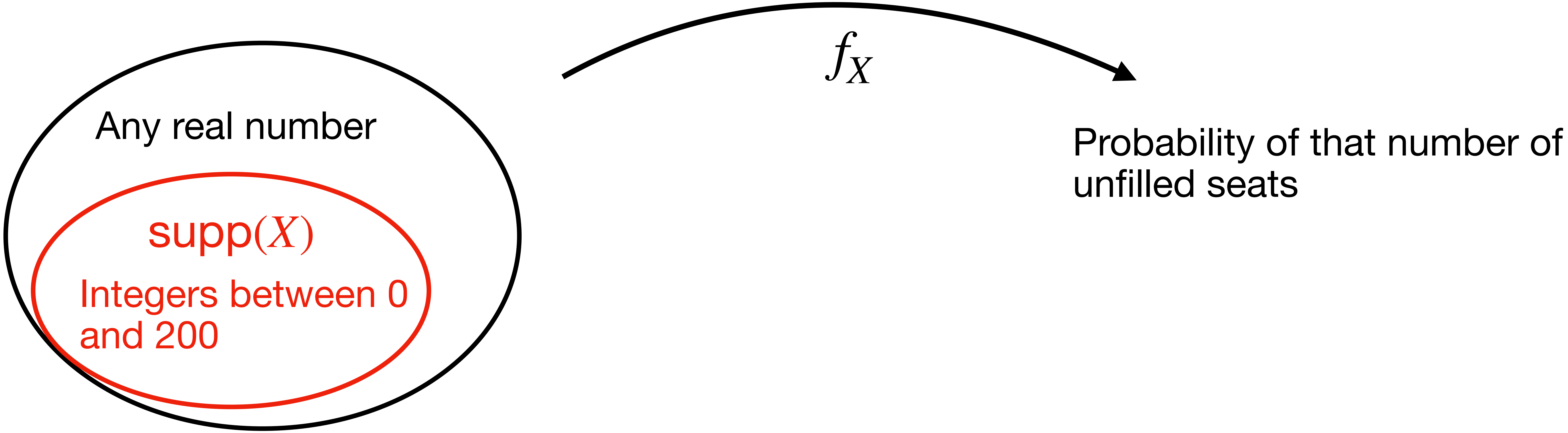
$$f_X : \mathbb{R} \to [0,1]$$



Probability mass function

…discrete random variables only!

X is the number of unfilled seats



Probability mass function

…discrete random variables only!

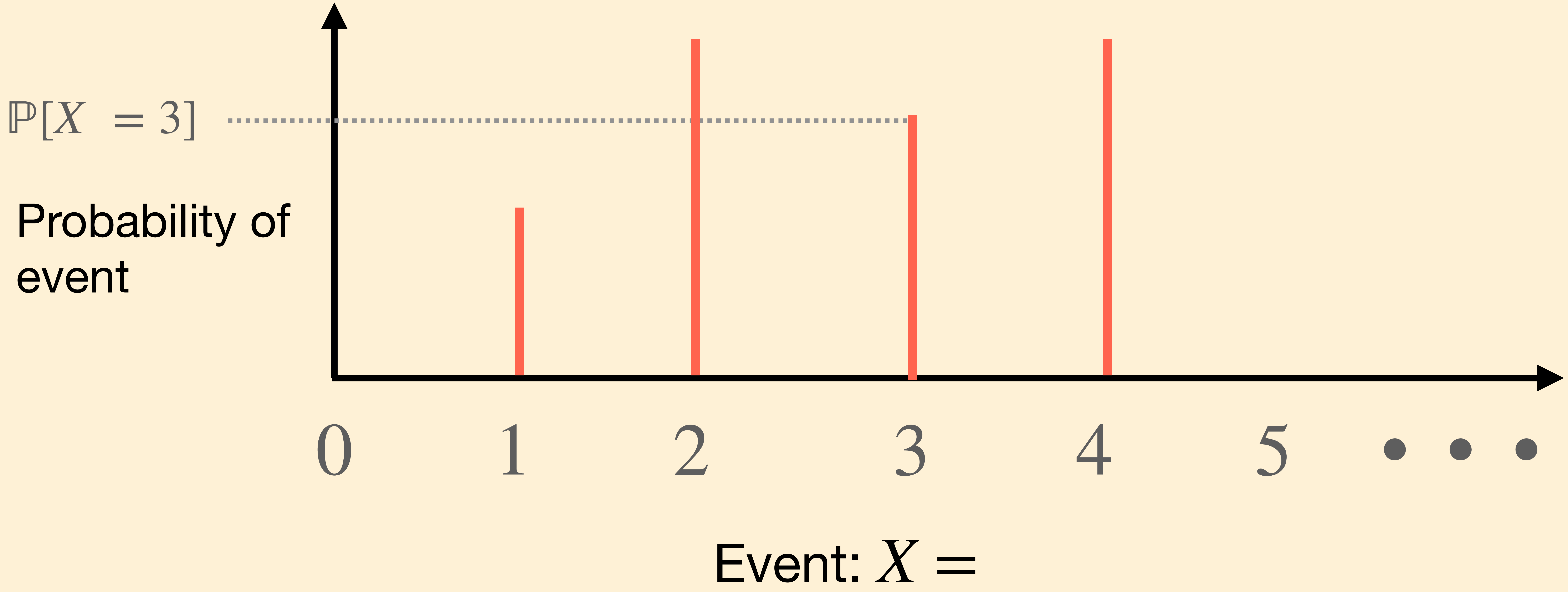
Not explicitly tied to probability space → Can talk about it without defining probability space

…but keep latter in the back of your mind!

$f_X:$ Value v → Event: $X = v$ → Probability of event

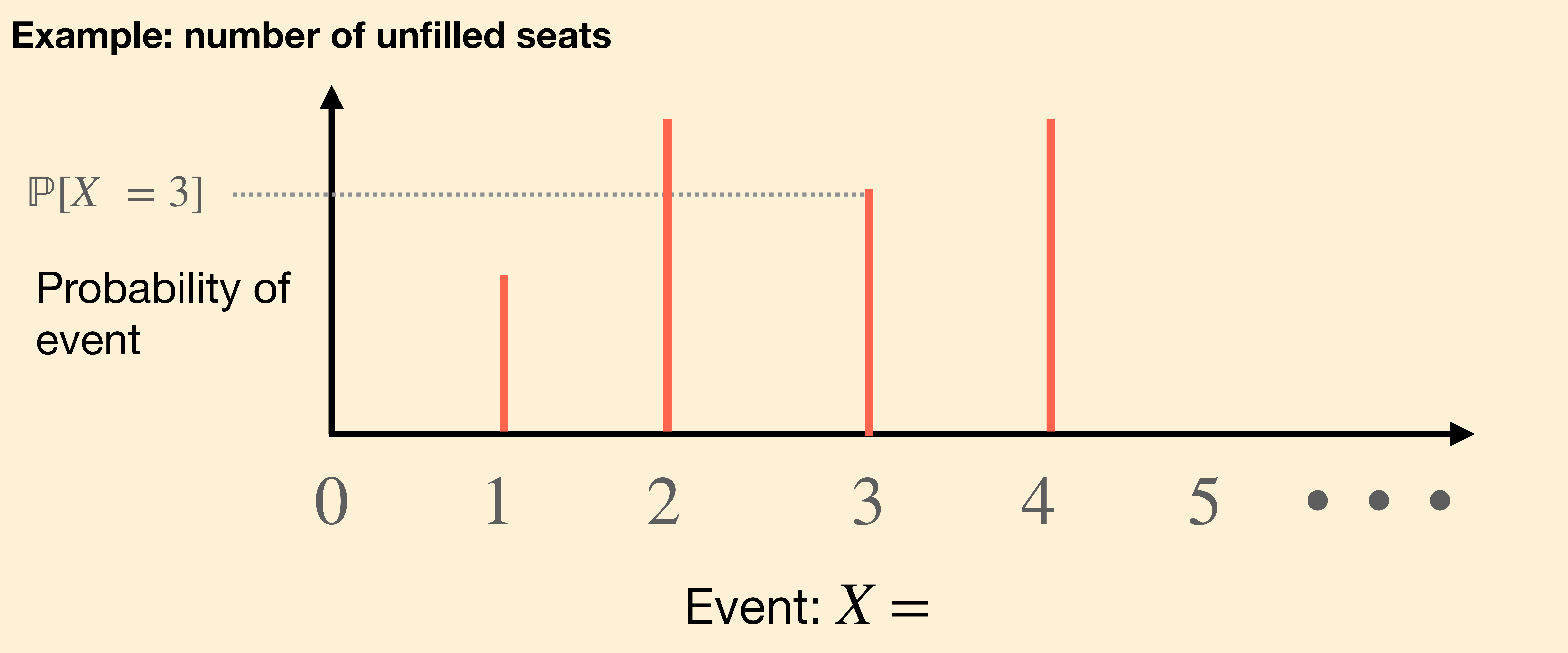
Graph of probability mass function f_X

Example: number of unfilled seats



NB: could have mathematically identical PMF for completely different experiment

Properties of Probability mass functions



Sum of red line lengths adds to one

Probability of X taking some value = 1

Random variables with particular
PMFs pop up quite often…

We give then names

Bernoulli random variable

Random variables are quantitative questions about the experiment

Bernoulli random variables are binary questions (yes/no)

Y is whether the back row is filled

f_Y?

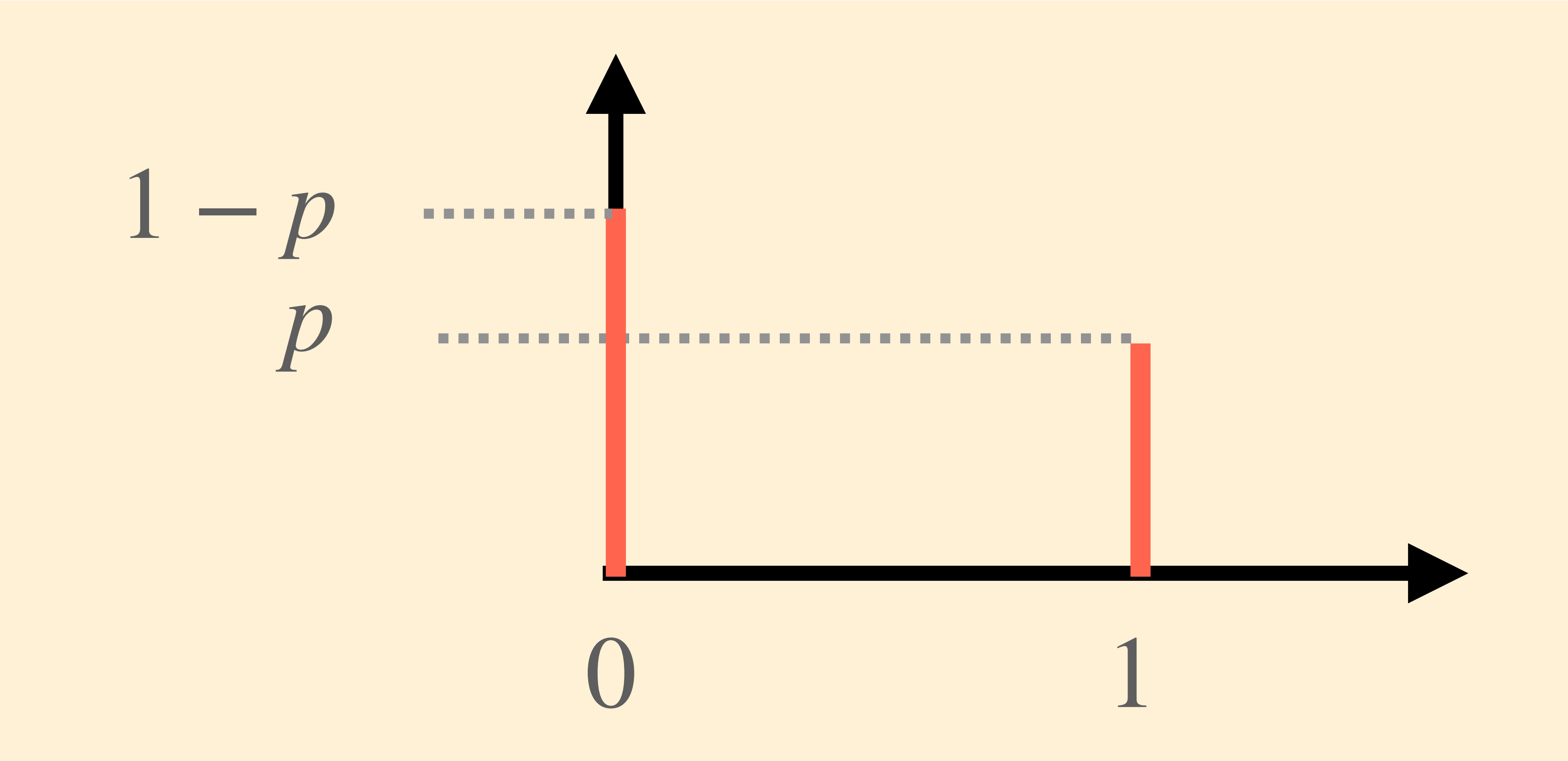
Bernoulli random variable

```
numpy.random.binomial(1,p)
```

$$Y \sim \text{Bern}(p)$$

Y **is distributed as a Bernoulli random variable,**
with probability p

$$f_Y = \begin{cases} p & \text{if } x = 1 \\ 1 - p & \text{if } x = 0 \end{cases}$$



Uniform random variables

`numpy.random.rand`

Probability of every outcome in support is equal

(Probably means we know very little about experiment)

Uniform random variables

Probability of every outcome in support is equal

(Probably means we know very little about experiment)

NB: S is support

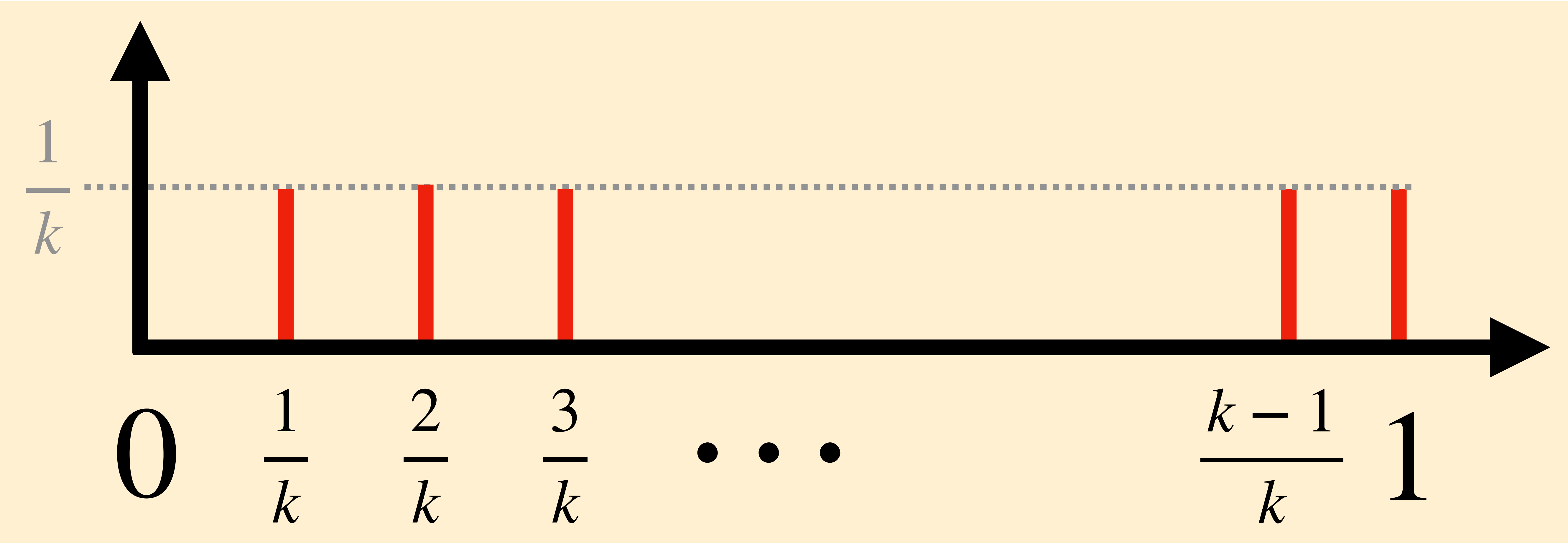
X is uniformly distributed on the set S	$\Longleftrightarrow$	$X \sim U(S)$
X is a discrete random variable	$\Longleftrightarrow$	$\lvert S \rvert < \infty$

Example of a Uniform random variable

$$X \sim U\left(\left\{\frac{i}{k}\right\}_{i=1}^{k}\right)$$

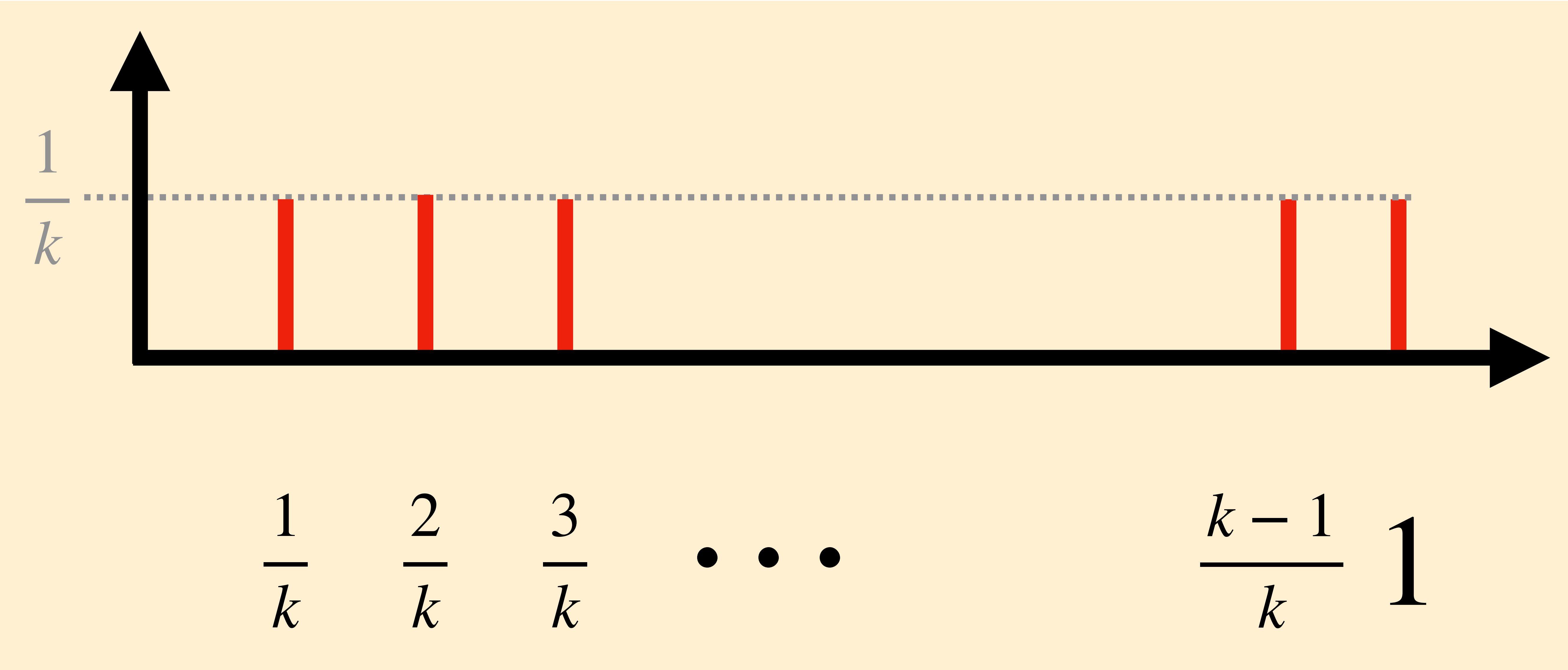
NB: { • } is set

(For some fixed $k \in \mathbb{N}$)



More outcomes means less probability per outcome

What happens as $k \to \infty$?

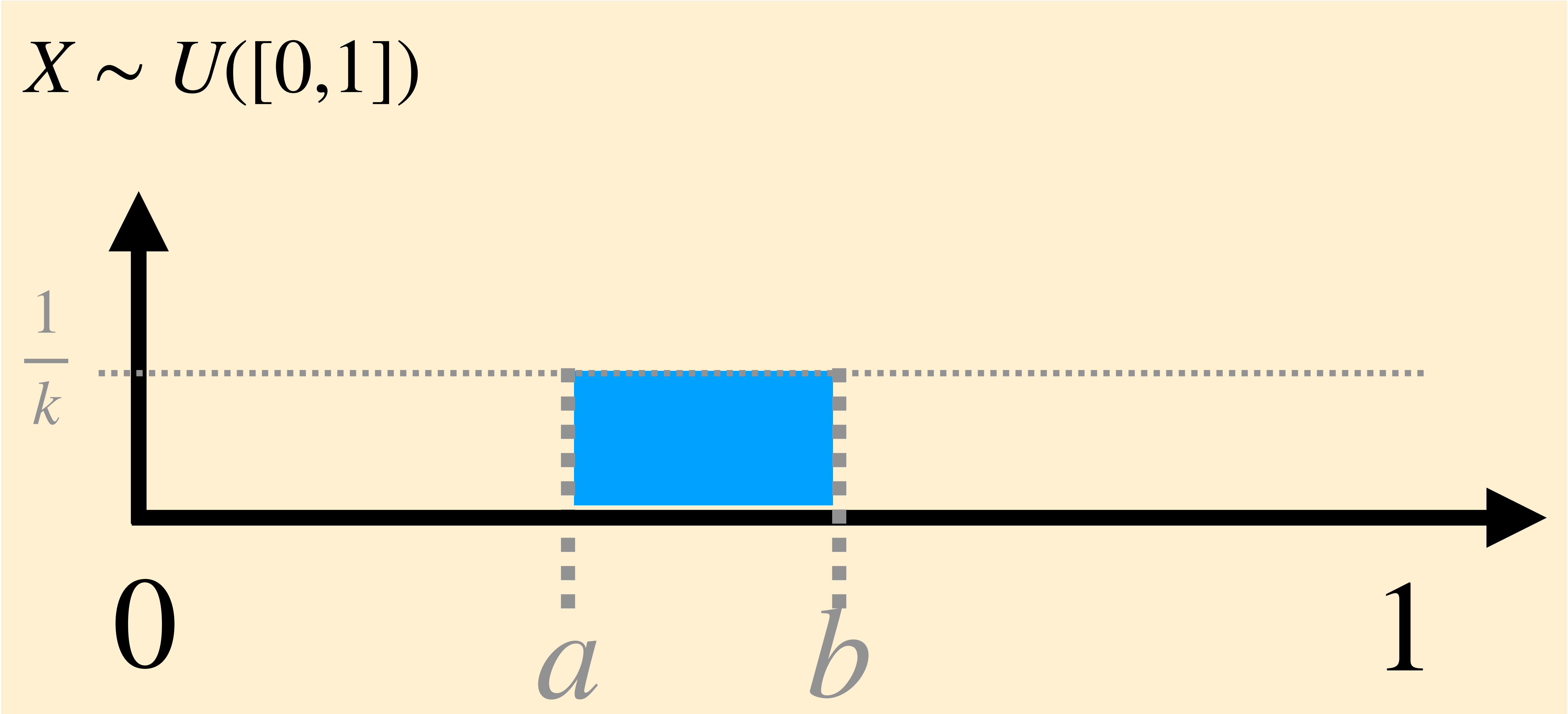


Random variable with continuous support

Every possible (single) outcome is *impossible*

Only reasonable quantity: ranges of outcomes

Uniform random variable with continuous support



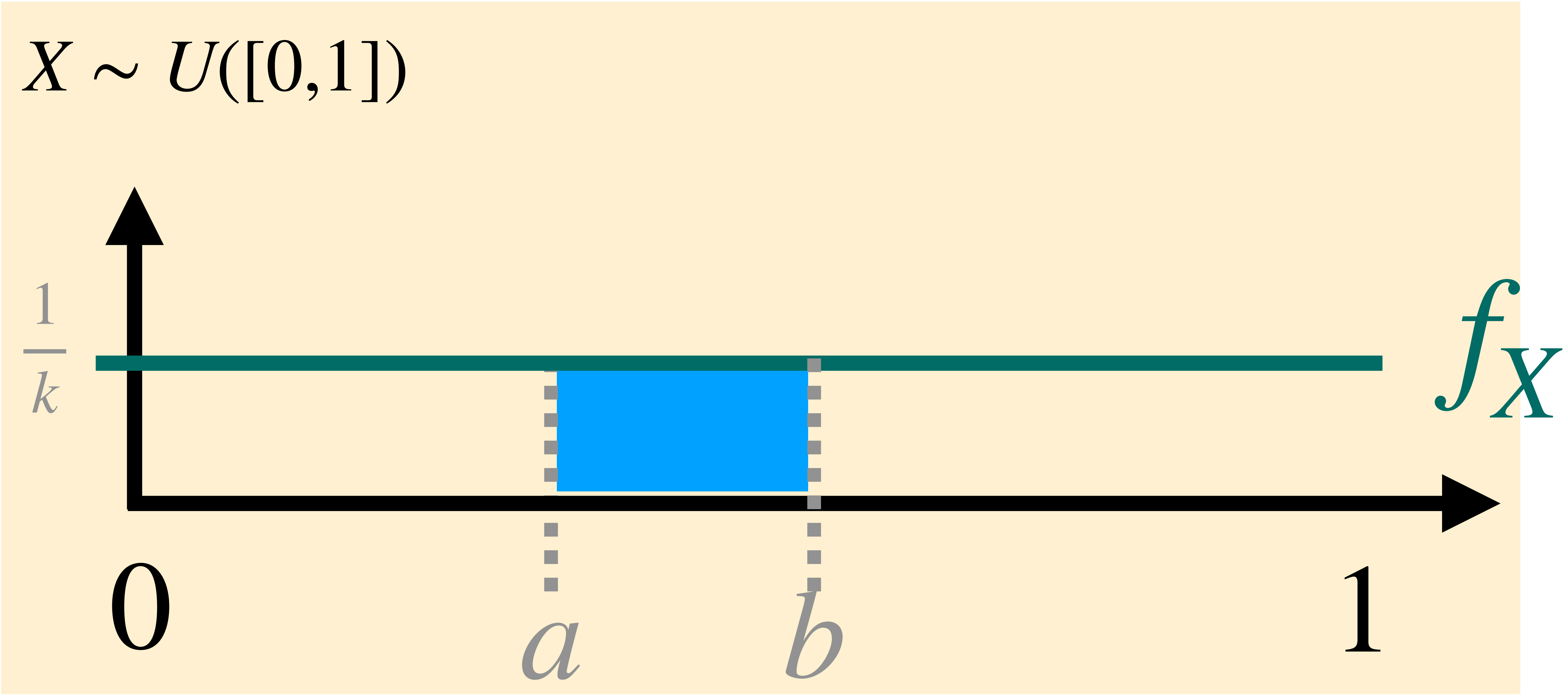
$$\mathbb{P}[X \in \{a \cup b\}] = 0$$

Probability X equals a or b

$$\mathbb{P}[X \in [a, b]] = b - a$$

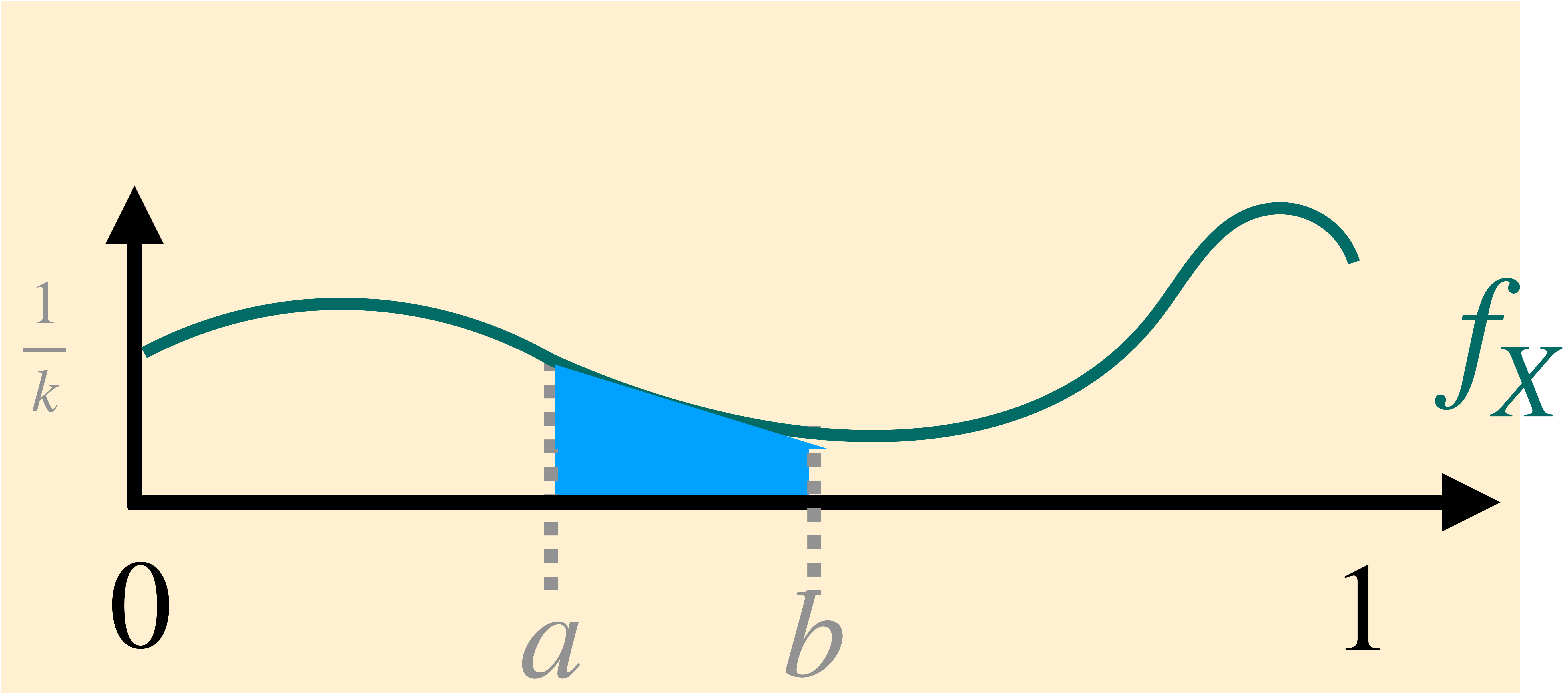
Probability X between a and b

Uniform random variable with continuous support



$$\boxed{\mathbb{P}[x \in (a,b)] = \int_a^b f_X(x)\ dx}$$

Probability density function of continuous random variable



$$\mathbb{P}[x \in (a, b)] = \int_a^b f_X(x)\, dx$$

Summary

Probability space

↓

Random variable

↙ ↘

Discrete
finite support

↓

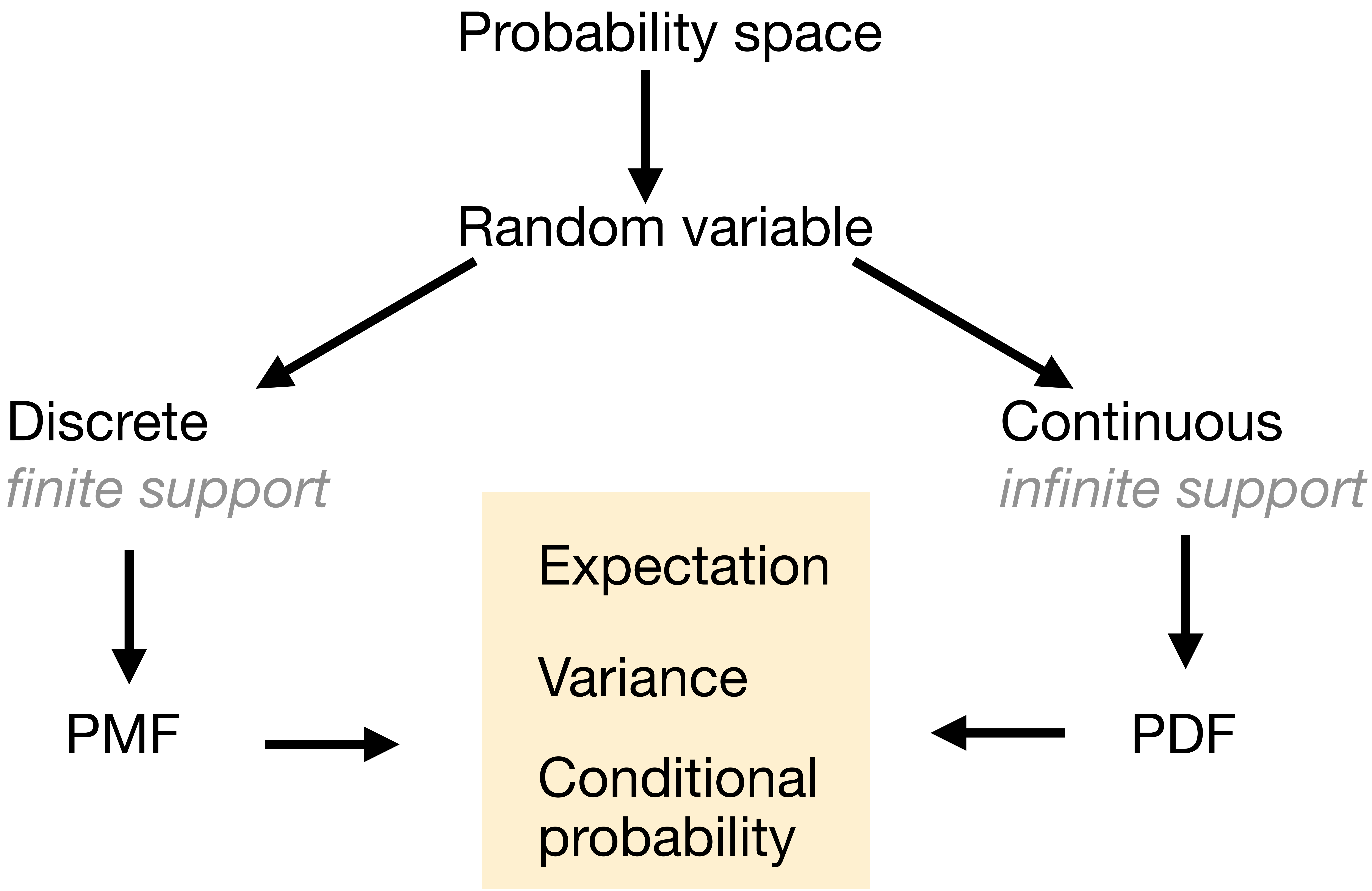
PMF

Continuous
infinite support

↓

PDF

What’s left


Probability space
Random variable
Discrete
finite support
Continuous
infinite support
PMF
PDF
Expectation
Variance
Conditional probability