LECTURE 1: Probability models and axioms

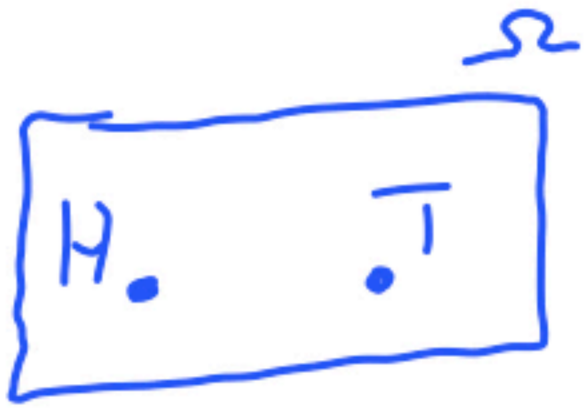
- Sample space
- Probability laws
  - Axioms
  - Properties that follow from the axioms
- Examples
  - Discrete
  - Continuous
- Discussion
  - Countable additivity
  - Mathematical subtleties
- Interpretations of probabilities

Sample space

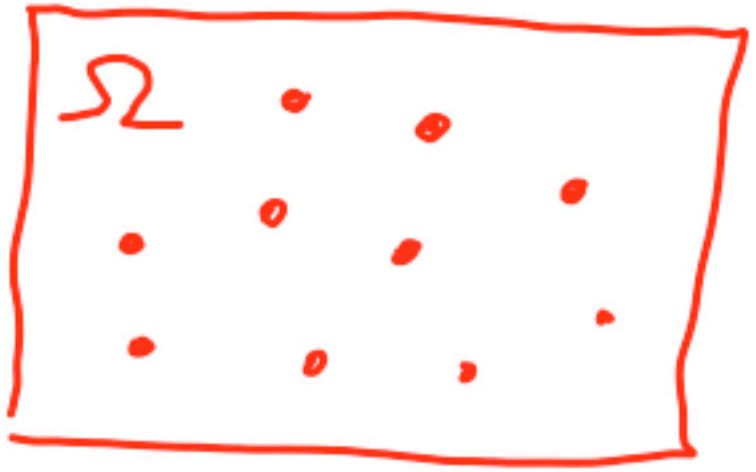
- Two steps:
  - Describe possible outcomes
  - Describe beliefs about likelihood of outcomes

Sample space

- List (set) of possible outcomes, Ω



- List must be:
  - Mutually exclusive
  - Collectively exhaustive
  - At the "right" granularity

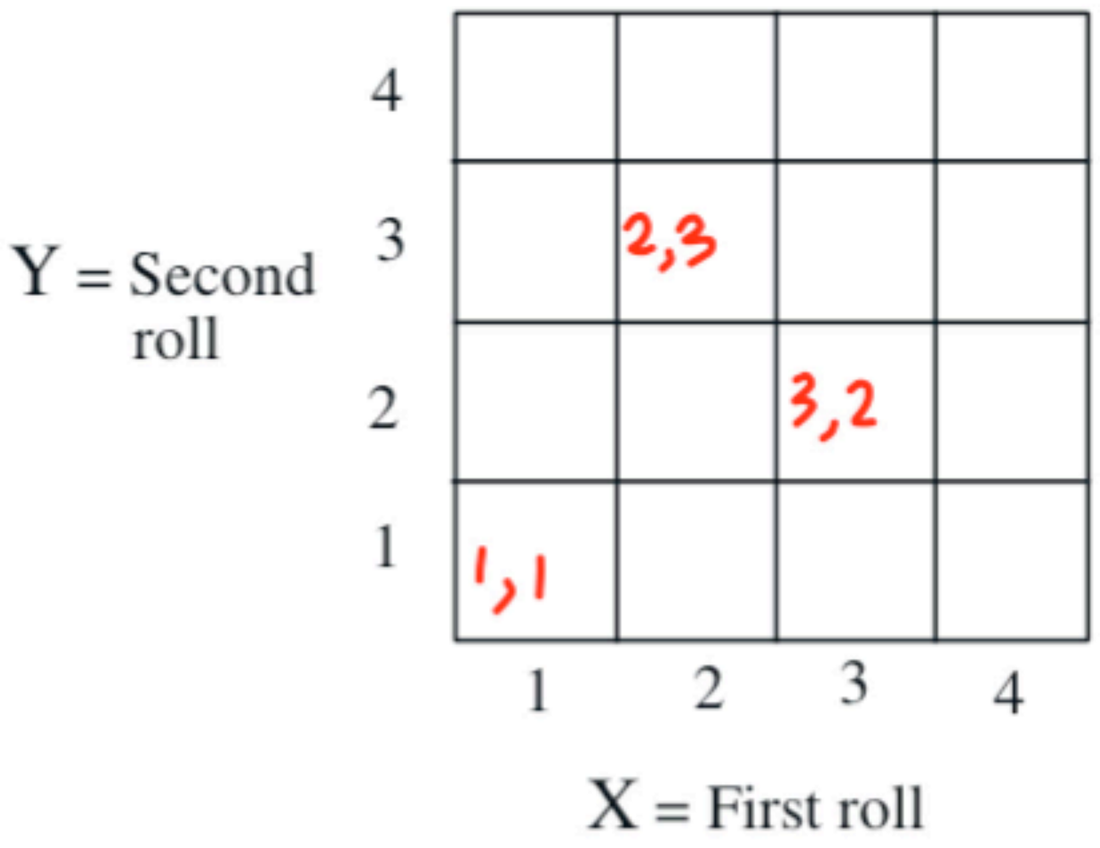


Ω

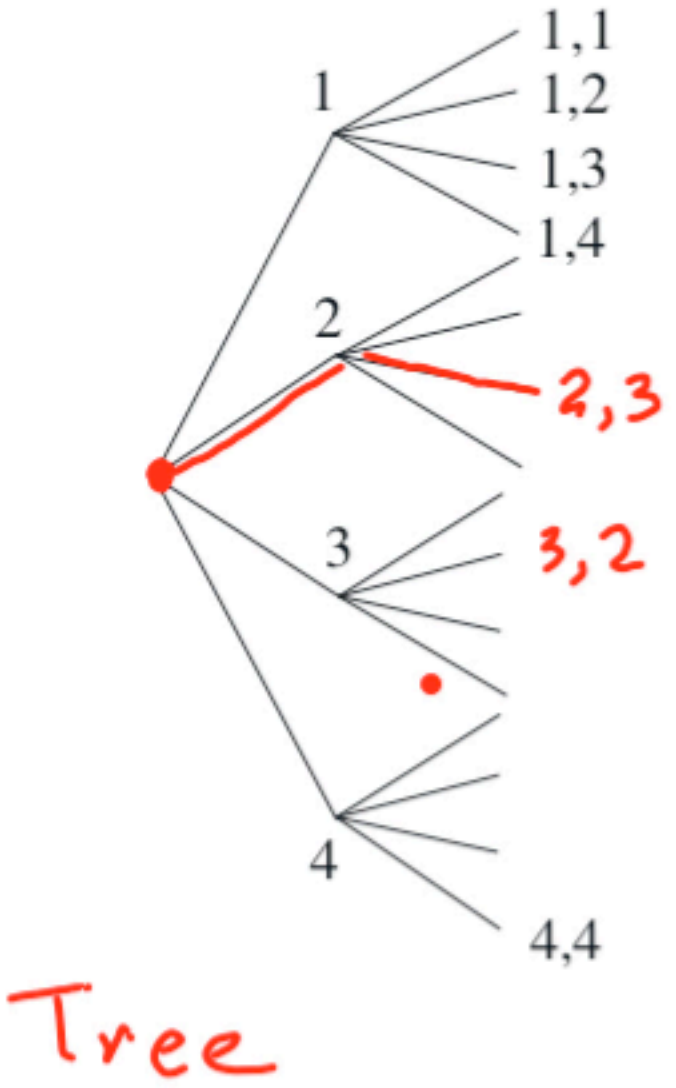
- H and rains
- H and no rain
- T and rains
- T and no rain

Sample space: discrete/finite example

- Two rolls of a tetrahedral die

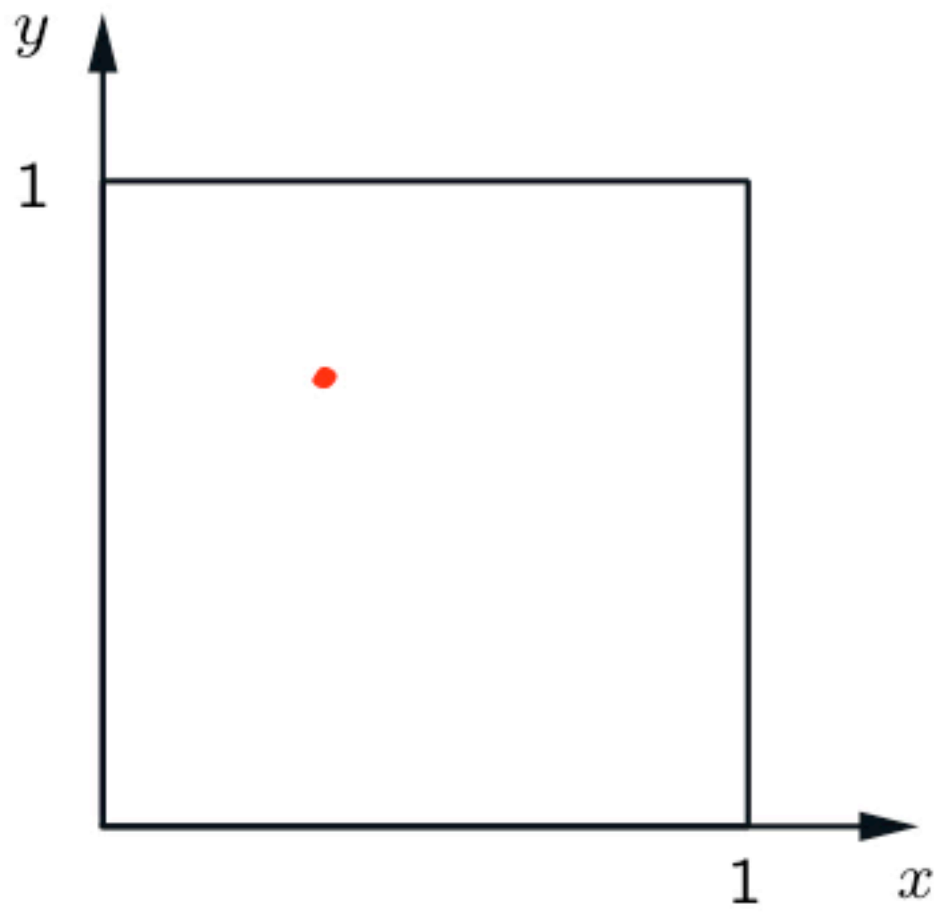


sequential description



Sample space: continuous example

- (x, y) such that $0 \leq x, y \leq 1$

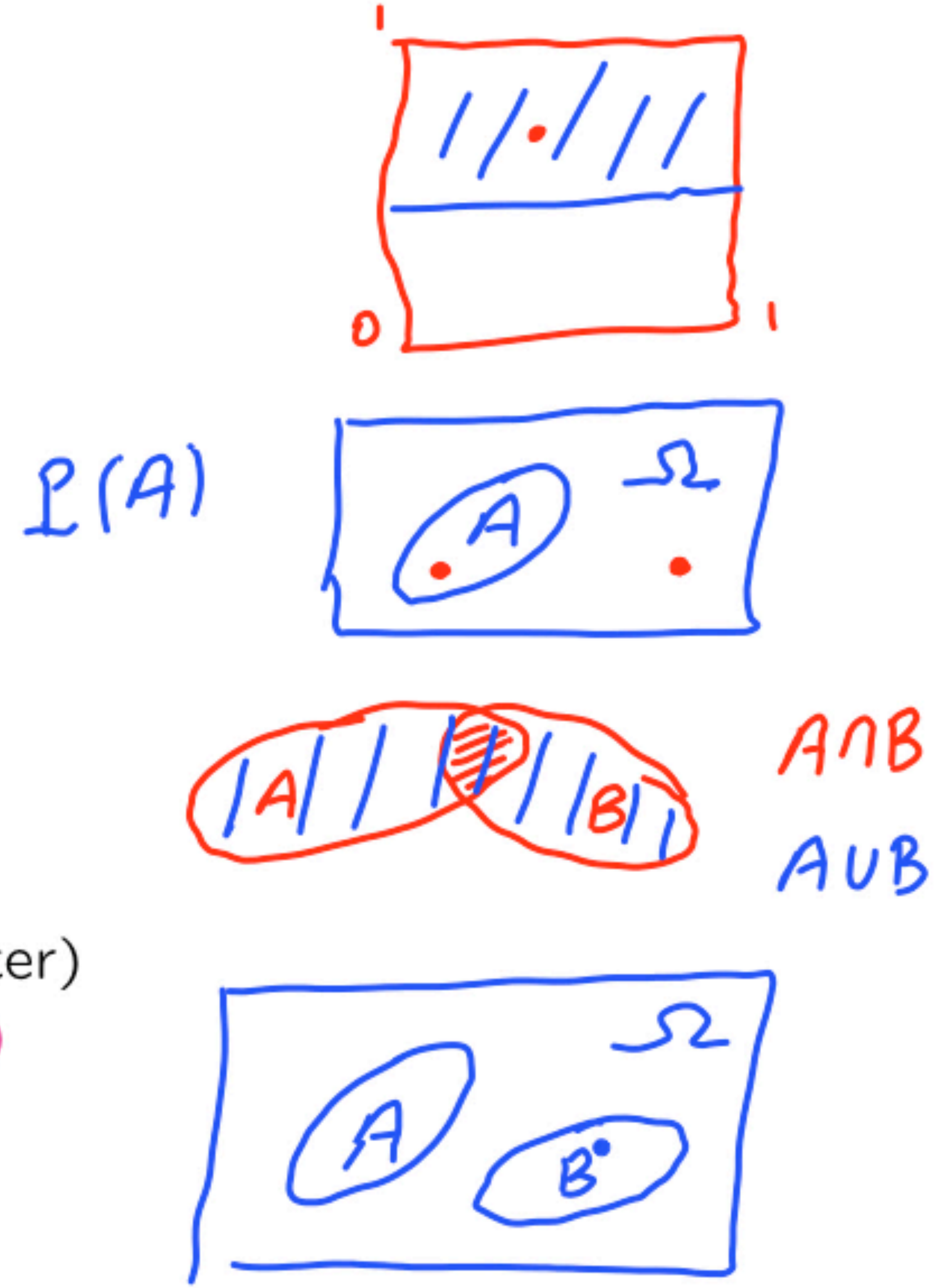


Probability axioms

- **Event:** a subset of the sample space
  - Probability is assigned to events

- **Axioms:**
  - Nonnegativity: $\mathbf{P}(A) \geq 0$
  - Normalization: $\mathbf{P}(\Omega) = 1$
  - (Finite) additivity: (to be strengthened later)
    If $A \cap B = \emptyset$, then $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B)$

    empty set



Some simple consequences of the axioms

Axioms

$\mathbf{P}(A) \geq 0$

$\mathbf{P}(\Omega) = 1$

For disjoint events:

$\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B)$

Consequences

$\mathbf{P}(A) \leq 1$

$\mathbf{P}(\varnothing) = 0$

$\mathbf{P}(A) + \mathbf{P}(A^c) = 1$

$\mathbf{P}(A \cup B \cup C) = \mathbf{P}(A) + \mathbf{P}(B) + \mathbf{P}(C)$

and similarly for k disjoint events

$$\mathbf{P}(\{s_1, s_2, \ldots, s_k\}) = \mathbf{P}(\{s_1\}) + \cdots + \mathbf{P}(\{s_k\})$$

$$= \mathbf{P}(s_1) + \cdots + \mathbf{P}(s_k)$$

Some simple consequences of the axioms

Axioms

(a) $\mathbf{P}(A) \geq 0$

(b) $\mathbf{P}(\Omega) = 1$

For disjoint events:

(c) $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B)$

$A \cup A^c = \Omega$

$A \cap A^c = \emptyset$

$1 \overset{(b)}{=} \mathbf{P}(\Omega) = \mathbf{P}(A \cup A^c)$

$\overset{(c)}{=} \mathbf{P}(A) + \mathbf{P}(A^c)$

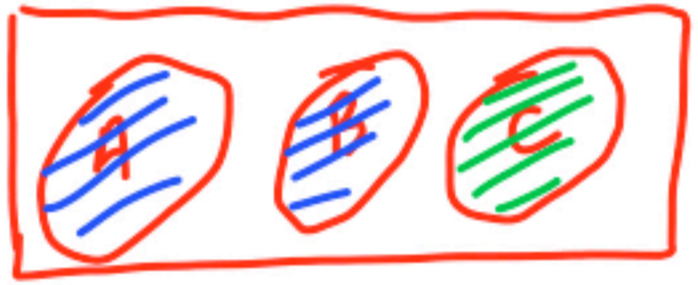
$\mathbf{P}(A) = 1 - \underbrace{\mathbf{P}(A^c)}_{} \overset{(a)}{\leq} 1$

$1 = \mathbf{P}(\Omega) + \mathbf{P}(\Omega^c)$

$1 = 1 + \mathbf{P}(\emptyset) \Rightarrow \mathbf{P}(\emptyset) = 0.$

Some simple consequences of the axioms

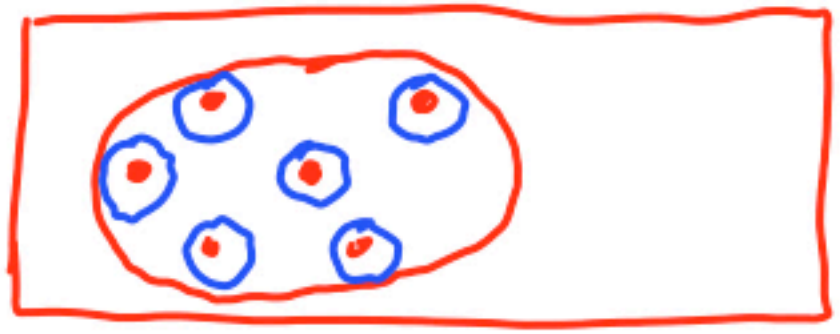
- A, B, C disjoint: $\mathbf{P}(A \cup B \cup C) = \mathbf{P}(A) + \mathbf{P}(B) + \mathbf{P}(C)$



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

If $A_1, \ldots, A_k$ disjoint $\Rightarrow P(A_1 \cup \cdots \cup A_k) = \sum_{i=1}^{k} P(A_i)$

- $\mathbf{P}(\{s_1, s_2, \ldots, s_k\}) =$ $P(\{s_1\} \cup \{s_2\} \cup \cdots \cup \{s_k\})$

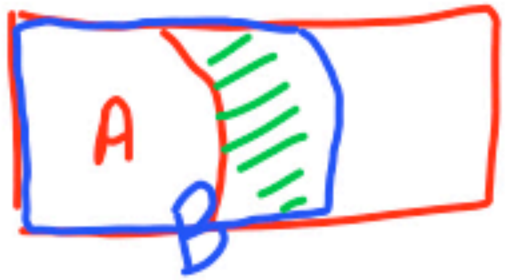


$$= P(\{s_1\}) + \cdots + P(\{s_k\})$$
$$= P(s_1) + \cdots + P(s_k)$$

More consequences of the axioms

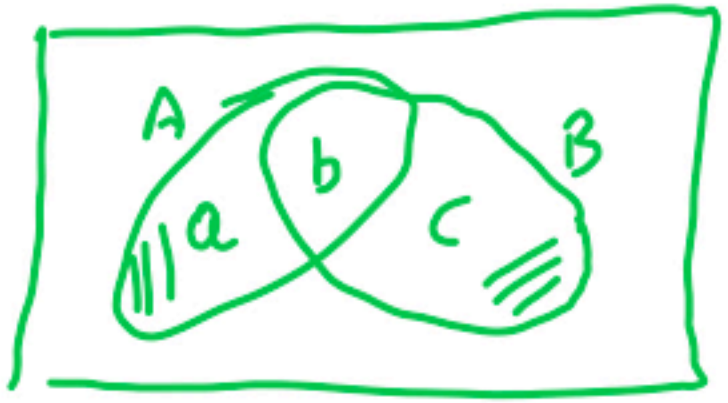
- If $A \subset B$, then $\mathbf{P}(A) \leq \mathbf{P}(B)$

$$B = A \cup (B \cap A^c)$$



$$P(B) = P(A) + \underbrace{P(B \cap A^c)} \geq P(A)$$

- $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \overbrace{\mathbf{P}(A \cap B)}^{\geq 0}$



$a = P(A \cap B^c)$ $\quad b = P(A \cap B)$ $\quad c = P(B \cap A^c)$

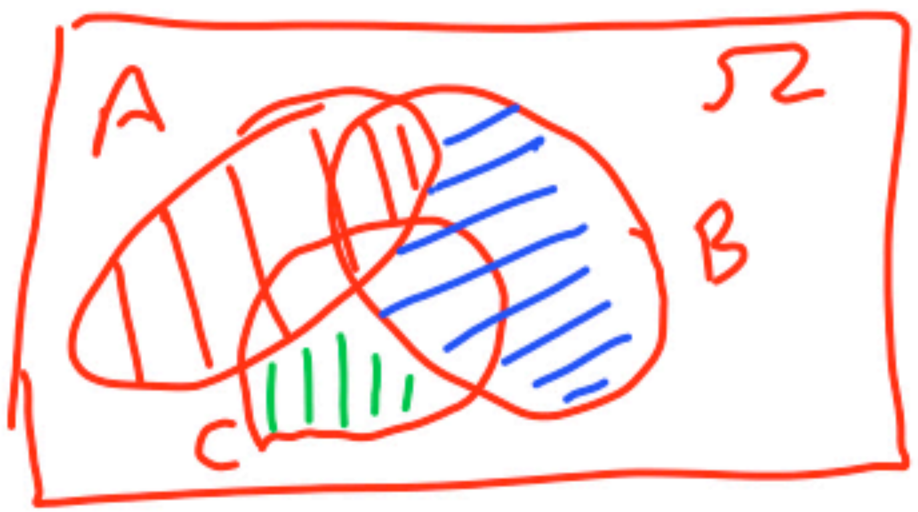
$$P(A \cup B) = a + b + c$$

$$P(A) + P(B) - P(A \cap B) = (a+b) + (\cancel{b} + c) - \cancel{b}$$

- $\mathbf{P}(A \cup B) \leq \mathbf{P}(A) + \mathbf{P}(B)$ union bound $= a + b + c$

More consequences of the axioms

- $\mathbf{P}(A \cup B \cup C) = \mathbf{P}(A) + \mathbf{P}(\underline{A^c \cap B}) + \mathbf{P}(\underline{A^c \cap B^c \cap C})$.

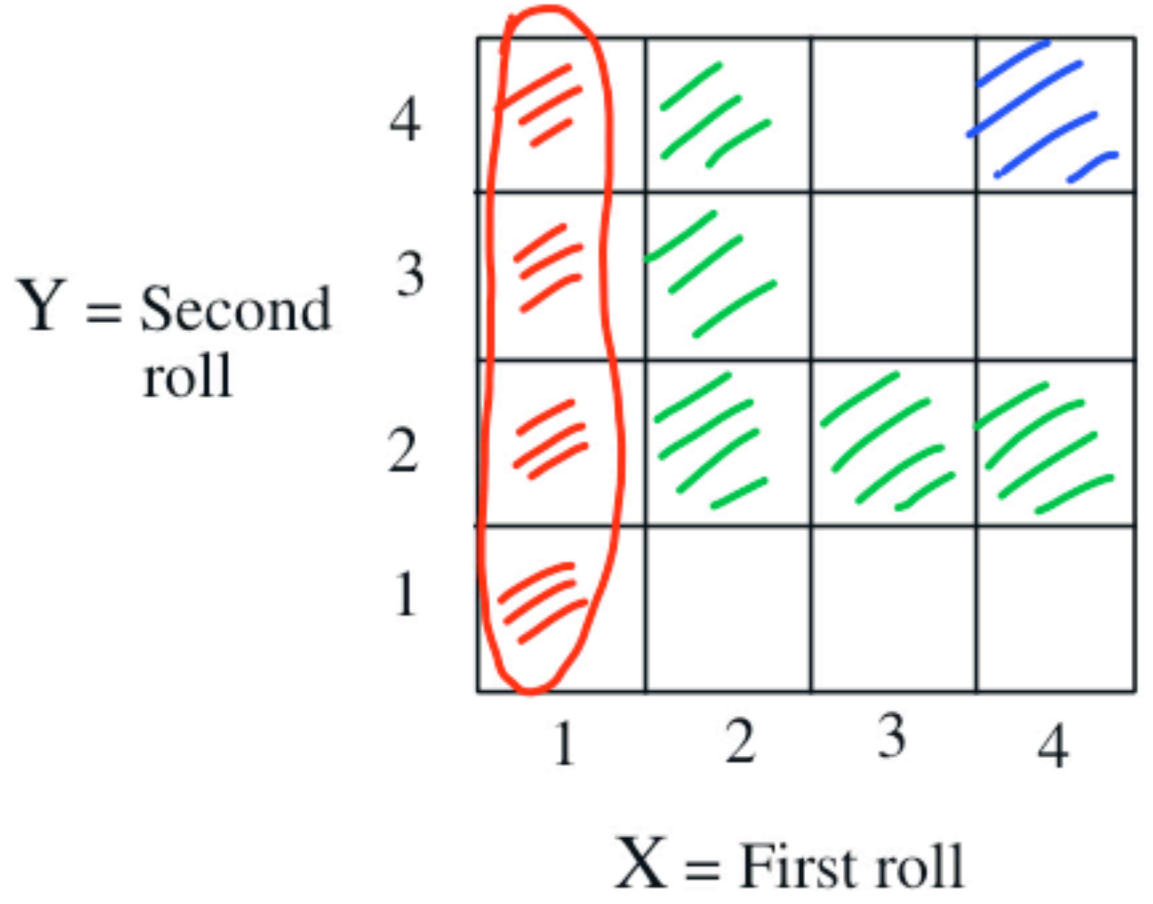


$$\mathbb{P}(A \cup B \cup C) =$$

$$= A \cup (\underline{B \cap A^c}) \cup (\underline{C \cap A^c \cap B^c})$$

Probability calculation: discrete/finite example

- Two rolls of a tetrahedral die
- Let every possible outcome have probability 1/16



- $\mathrm{P}(X = 1) =$ $4 \cdot \frac{1}{16} = \frac{1}{4}$

Let $Z = \min(X, Y)$

$x = 2,\ y = 3,\ z = 2$

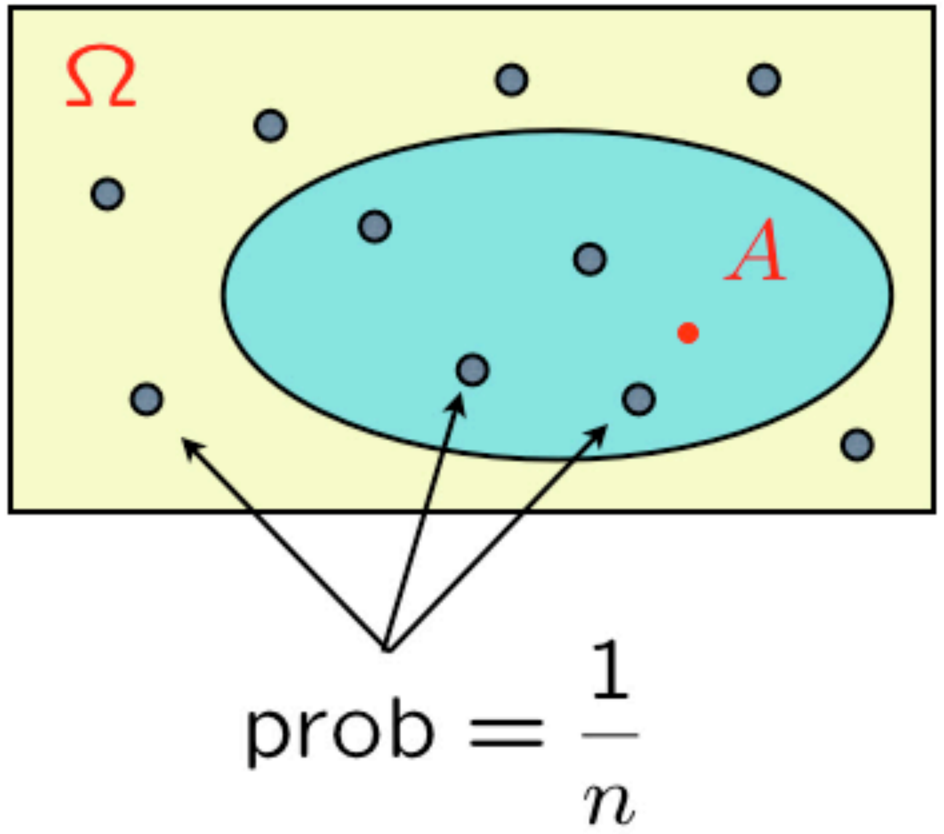
- $\mathrm{P}(Z = 4) =$ $1/16$
- $\mathrm{P}(Z = 2) =$ $5 \cdot \frac{1}{16}$.

Discrete uniform law

finite

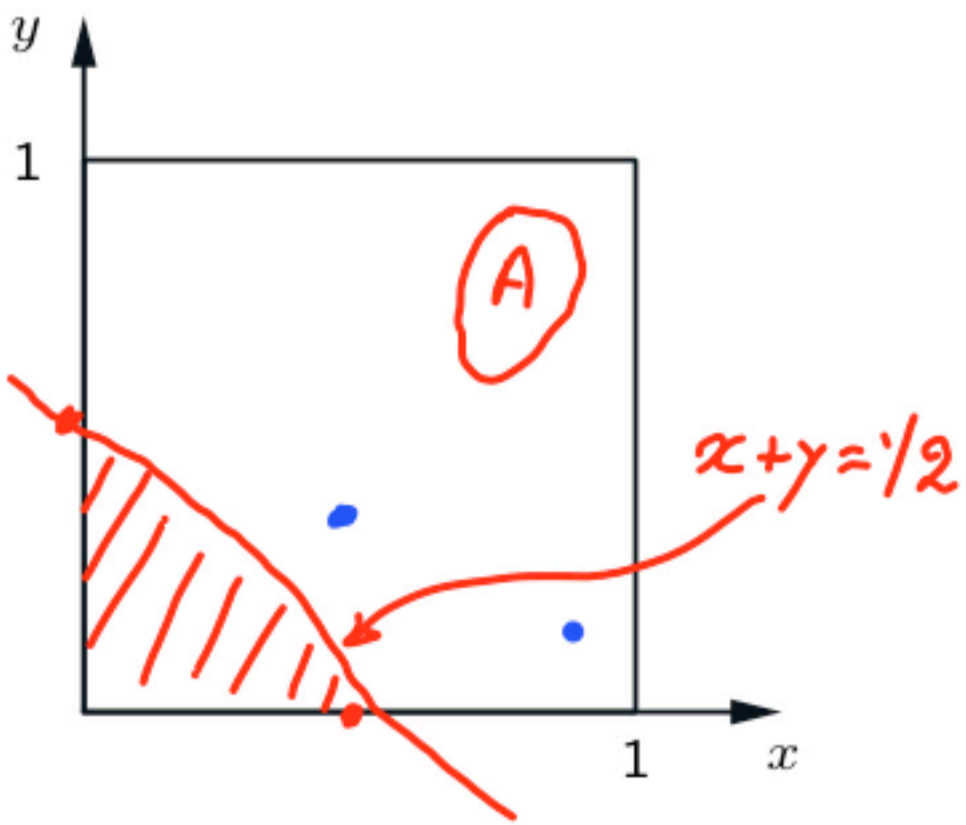
- Assume Ω consists of n equally likely elements
- Assume A consists of k elements

$$P(A) = k \cdot \frac{1}{n}$$



Probability calculation: continuous example

- (x, y) such that $0 \le x, y \le 1$



- **Uniform** probability law: Probability = Area

$$\mathbf{P}\Big(\{(x,y) \mid x + y \le 1/2\}\Big) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$$

$$\mathbf{P}\Big(\{(0.5, 0.3)\}\Big) = 0$$

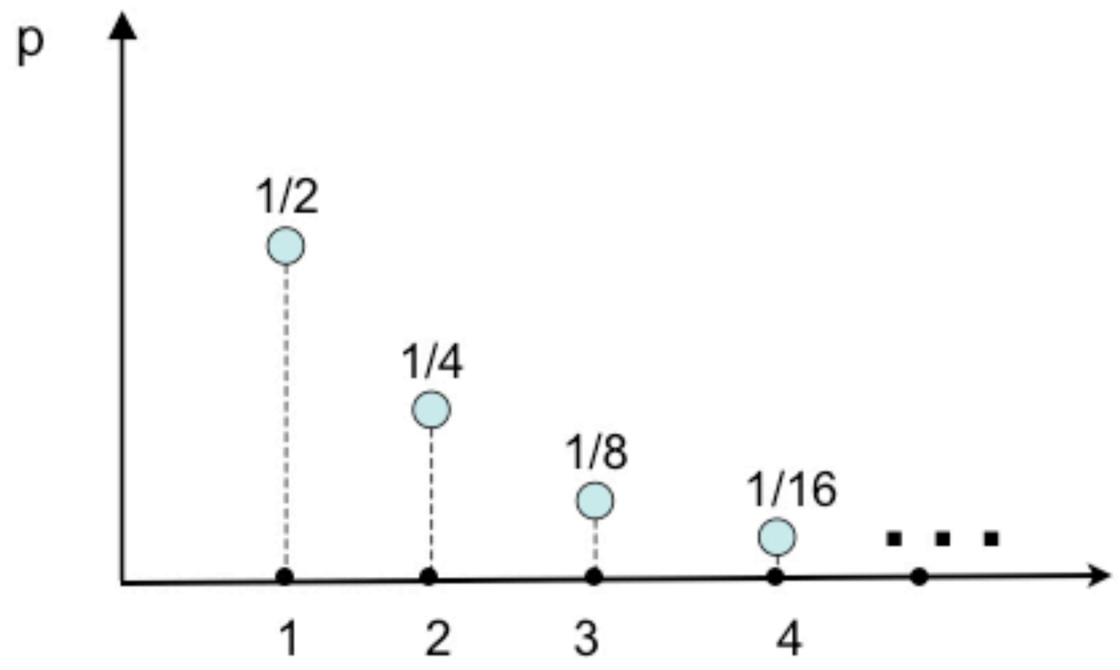
Probability calculation steps

- Specify the sample space
- Specify a probability law
- Identify an event of interest
- Calculate...

Probability calculation: discrete but infinite sample space

- Sample space: $\{1, 2, \ldots\}$
  - We are given $\mathrm{P}(n) = \dfrac{1}{2^n}, \quad n = 1, 2, \ldots$

$$\sum_{n=1}^{\infty} \frac{1}{2^n} = \frac{1}{2} \sum_{n=0}^{\infty} \frac{1}{2^n} = \frac{1}{2} \cdot \frac{1}{1-(1/2)} = 1$$



- $\mathrm{P}(\text{outcome is even}) = P(\{2, 4, 6, \ldots\})$

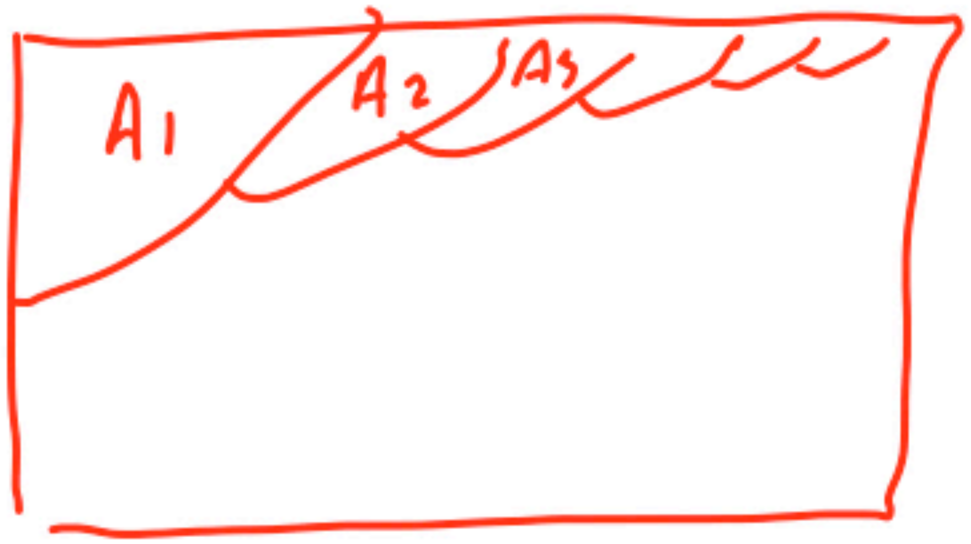
$$= P(\{2\} \cup \{4\} \cup \{6\} \cup \ldots) = P(2) + P(4) + P(6) + \ldots$$

$$= \frac{1}{2^2} + \frac{1}{2^4} + \frac{1}{2^6} + \ldots = \frac{1}{4}\left(1 + \frac{1}{4} + \frac{1}{4^2} + \ldots\right) = \frac{1}{4} \cdot \frac{1}{1 - \frac{1}{4}} = \frac{1}{3}$$

Countable additivity axiom

- Strengthens the finite additivity axiom

Countable Additivity Axiom:
If A_1, A_2, A_3,... is an infinite **sequence** of disjoint events,
then $\mathbf{P}(A_1 \cup A_2 \cup A_3 \cup \cdots) = \mathbf{P}(A_1) + \mathbf{P}(A_2) + \mathbf{P}(A_3) + \cdots$



Mathematical subtleties

Countable Additivity Axiom:
If A_1, A_2, A_3,... is an infinite sequence of disjoint events,
then $\mathbf{P}(A_1 \cup A_2 \cup A_3 \cup \cdots) = \mathbf{P}(A_1) + \mathbf{P}(A_2) + \mathbf{P}(A_3) + \cdots$

$$1 = P(\Omega) = P\left(\bigcup \{(x,y)\}\right) \overset{?}{=} \sum P(\{(x,y)\}) = \sum 0 = 0$$

- Additivity holds only for "countable" sequences of events
- The unit square (simlarly, the real line, etc.) is **not countable** (its elements cannot be arranged in a sequence)
- "Area" is a legitimate probability law on the unit square, as long as we do not try to assign probabilities/areas to "very strange" sets

Interpretations of probability theory

- A narrow view: a branch of math
  - Axioms $\Rightarrow$ theorems

"Thm:" "Frequency" of event A "is" $\mathbf{P}(A)$

- Are probabilities frequencies?
  - $\mathbf{P}(\text{coin toss yields heads}) = 1/2$
  - $\mathbf{P}(\text{the president of \ldots will be reelected}) = 0.7$

- Probabilities are often intepreted as:
  - Description of beliefs
  - Betting preferences

The role of probability theory

- A framework for analyzing phenomena with uncertain outcomes
  - Rules for consistent reasoning
  - Used for predictions and decisions