

LECTURE 1: Probability models and axioms

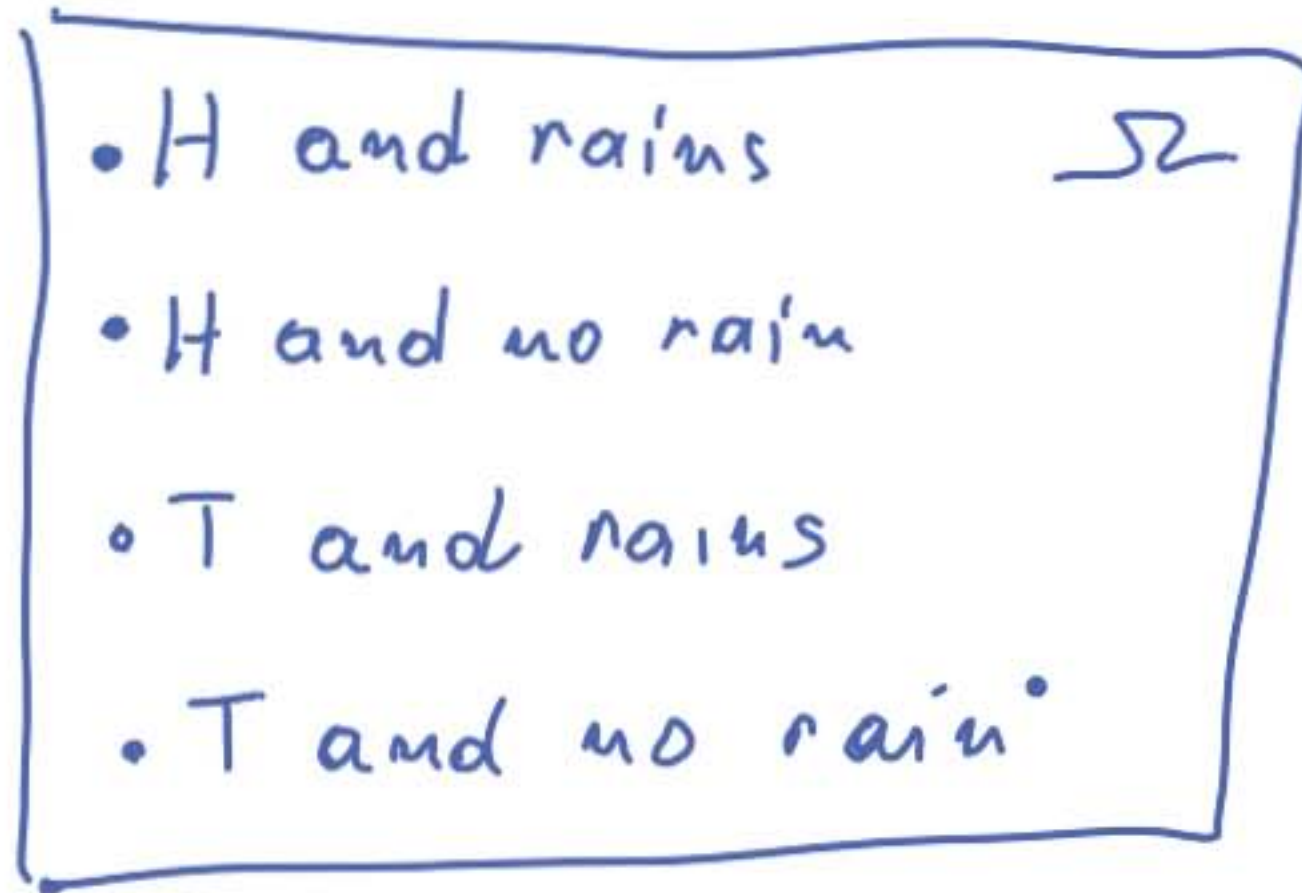
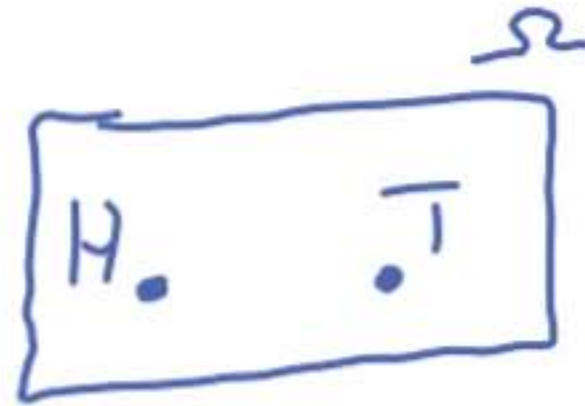
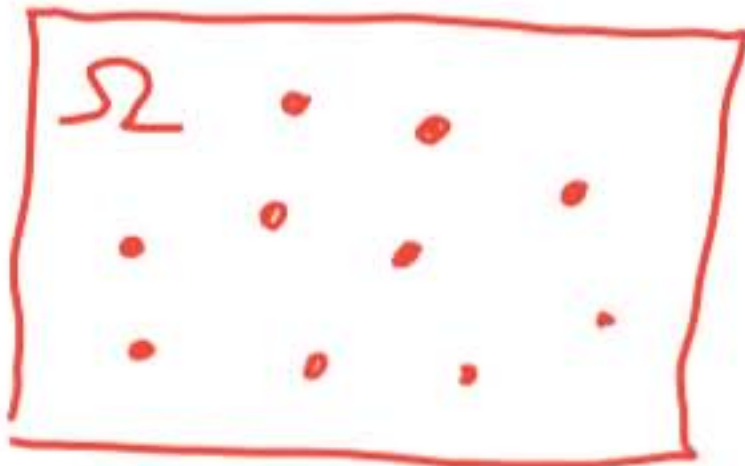
- Sample space
- Probability laws
 - Axioms
 - Properties that follow from the axioms
- Examples
 - Discrete
 - Continuous
- Discussion
 - Countable additivity
 - Mathematical subtleties
- Interpretations of probabilities

Sample space

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

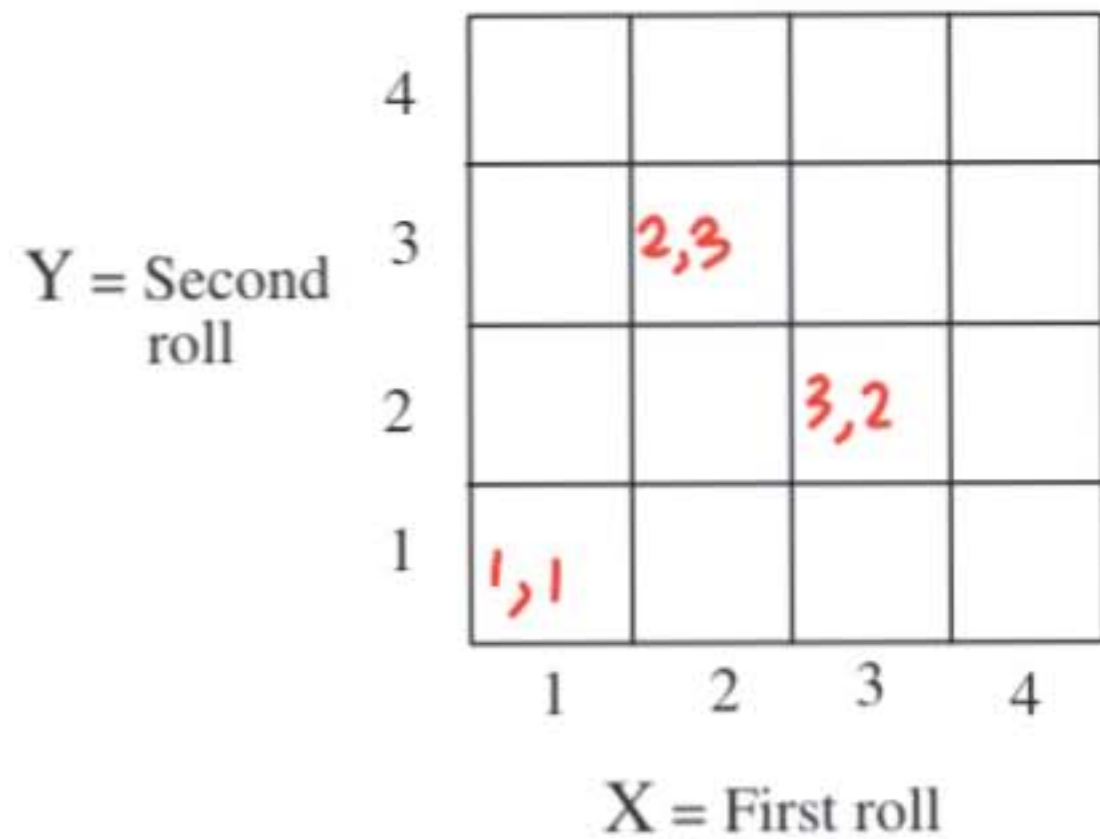
Sample space

- List (set) of possible outcomes, Ω
- List must be:
 - Mutually exclusive
 - Collectively exhaustive
 - At the “right” granularity

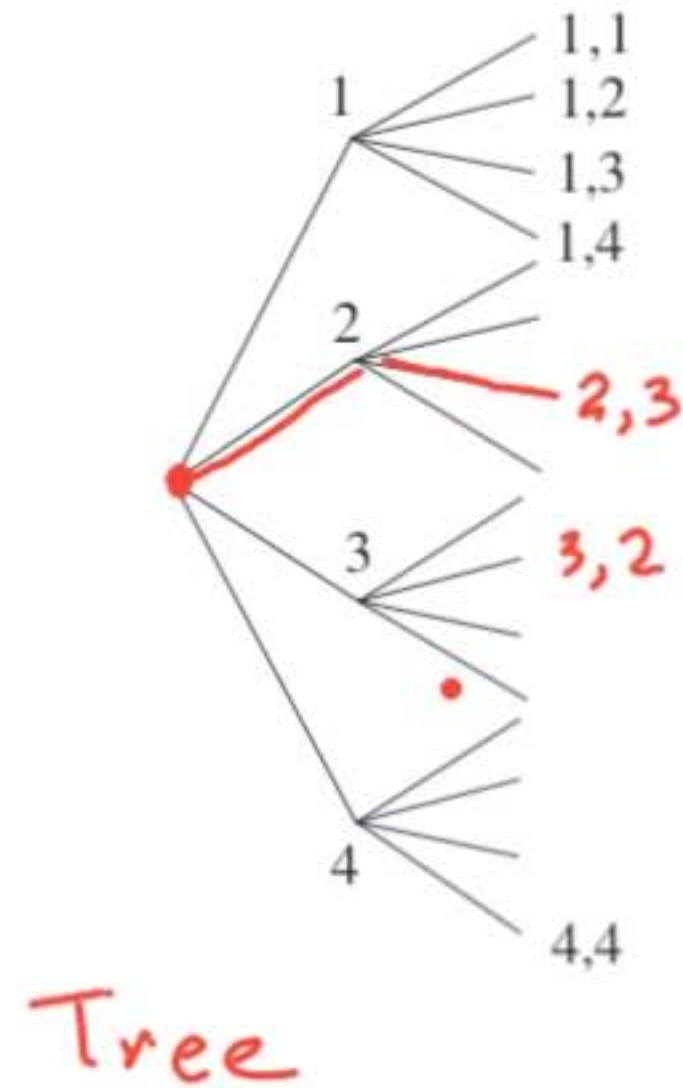


Sample space: discrete/finite example

- Two rolls of a tetrahedral die

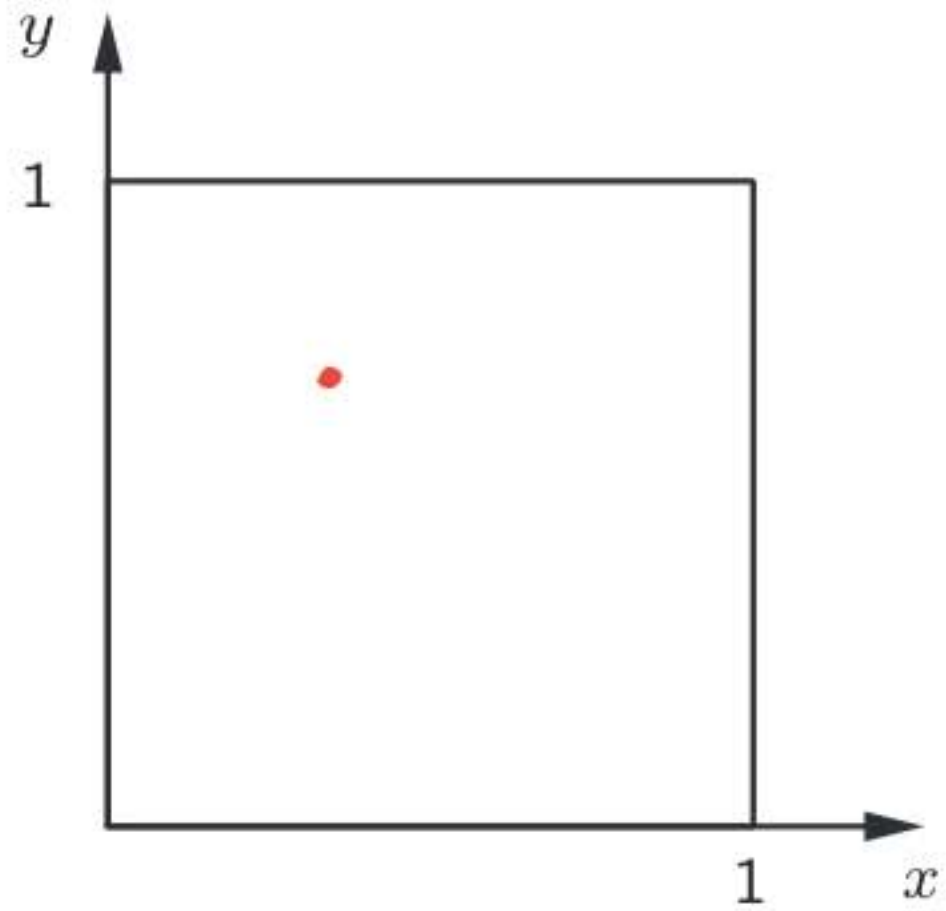


sequential description



Sample space: continuous example

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



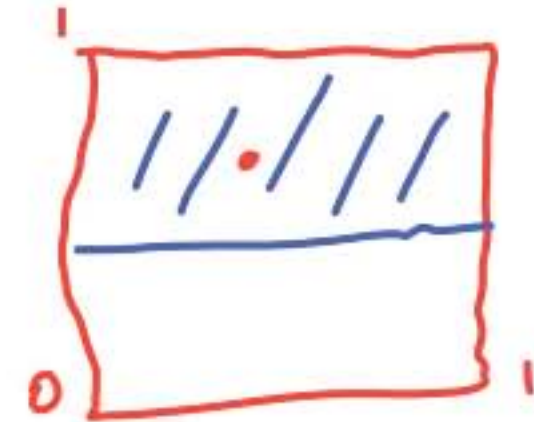
Probability axioms

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

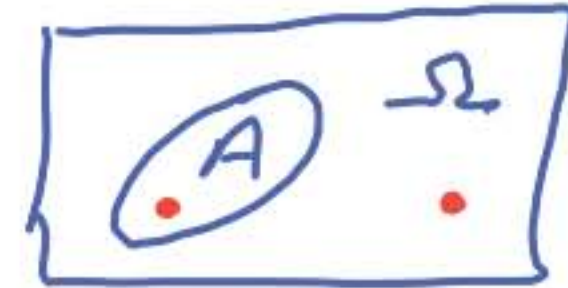
- **Axioms:**

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

\uparrow
empty set

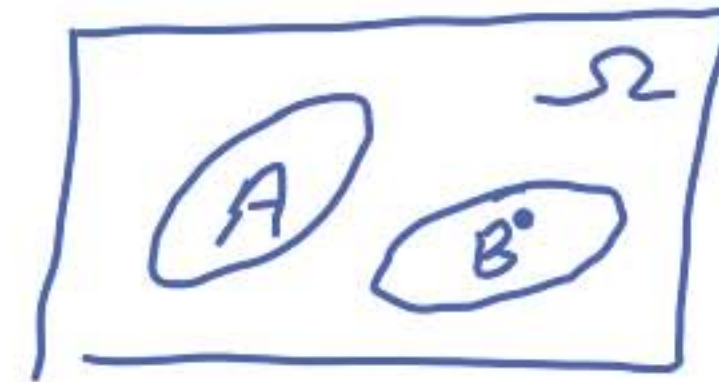


$P(A)$



$A \cap B$

$A \cup B$



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

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

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

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

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

and similarly for k disjoint events

$$\begin{aligned} \mathbf{P}(\{s_1, s_2, \dots, s_k\}) &= \mathbf{P}(\{s_1\}) + \dots + \mathbf{P}(\{s_k\}) \\ &= \mathbf{P}(s_1) + \dots + \mathbf{P}(s_k) \end{aligned}$$

Some simple consequences of the axioms

Axioms

(a) $P(A) \geq 0$

(b) $P(\Omega) = 1$

For disjoint events:
(c) $P(A \cup B) = P(A) + P(B)$



$$A \cup A^c = \Omega$$

$$A \cap A^c = \emptyset$$

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

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

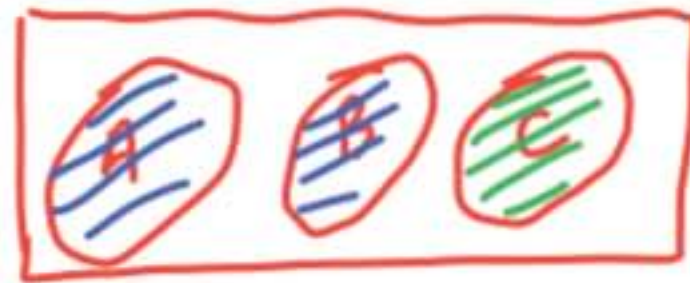
$$P(A) = 1 - \underbrace{P(A^c)}_{(a)} \leq 1$$

$$1 = P(\Omega) + P(\Omega^c)$$

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

Some simple consequences of the axioms

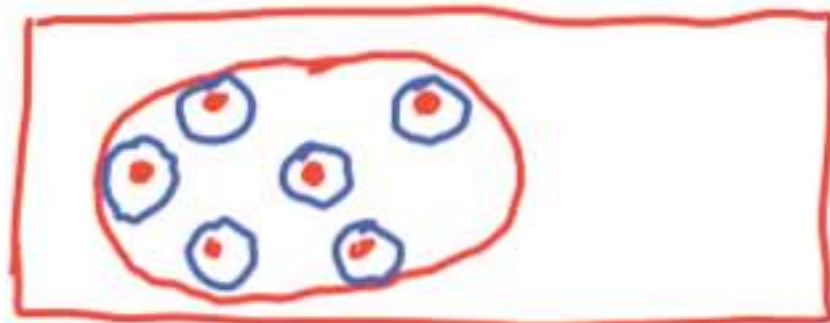
- A, B, C disjoint: $P(A \cup B \cup C) = P(A) + P(B) + P(C)$



$$\begin{aligned} P(A \cup B \cup C) &= P((A \cup B) \cup C) = P(A \cup B) + P(C) \\ &= P(A) + P(B) + P(C) \end{aligned}$$

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

- $P(\{s_1, s_2, \dots, s_k\}) = P(\{s_1\} \cup \{s_2\} \cup \dots \cup \{s_k\})$



$$\begin{aligned} &= P(\{s_1\}) + \dots + P(\{s_k\}) \\ &= P(s_1) + \dots + P(s_k) \end{aligned}$$

More consequences of the axioms

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



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

$$P(B) = P(A) + \underbrace{P(B \cap A^c)}_{\geq 0} \geq P(A)$$

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



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

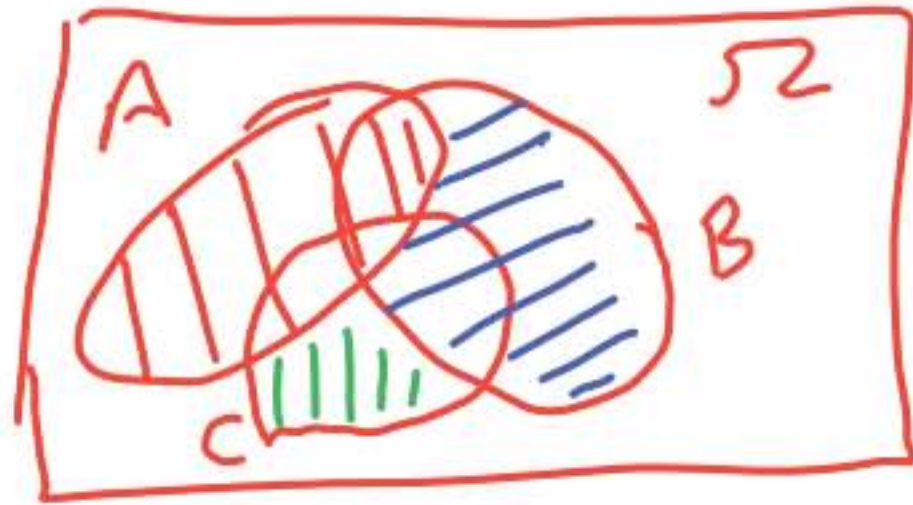
$$P(A \cup B) = a + b + c$$

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

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

More consequences of the axioms

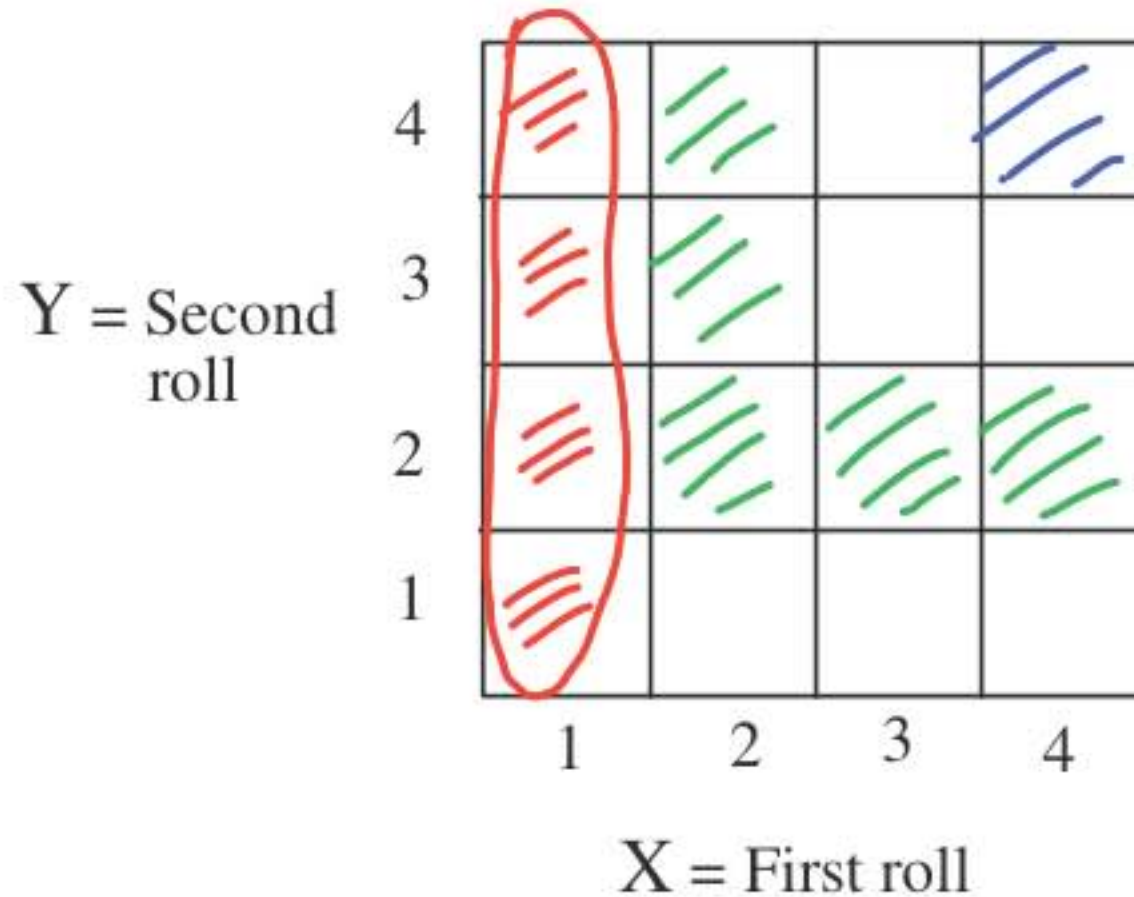
- $P(A \cup B \cup C) = P(A) + P(\underline{A^c \cap B}) + P(\underline{A^c \cap B^c \cap C})$ • ←



$$\begin{aligned} P(A \cup B \cup C) &= \\ &= A \cup (\underline{B \cap A^c}) \cup (\underline{C \cap A^c \cap B^c}) \end{aligned}$$

Probability calculation: discrete/finite example

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



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

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

- $X = 2, Y = 3, Z = 2$
 $P(\underline{Z} = 4) = 1/16$

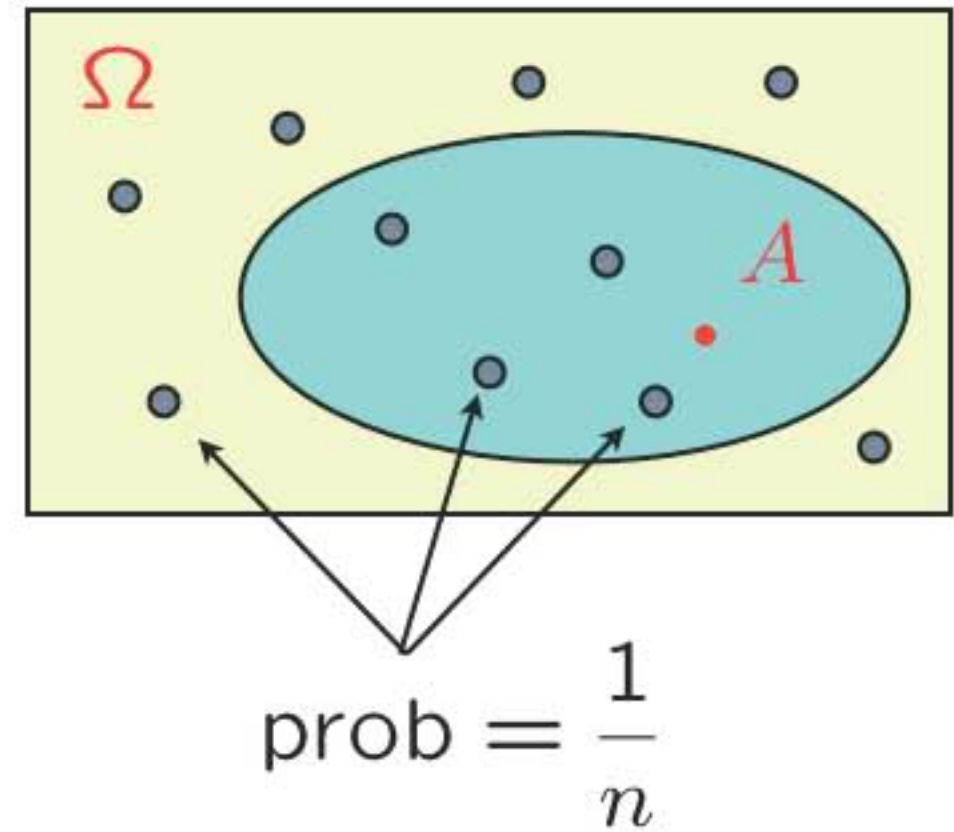
- $P(Z = 2) = 5 \cdot \frac{1}{16}$

Discrete uniform law

✓ finite

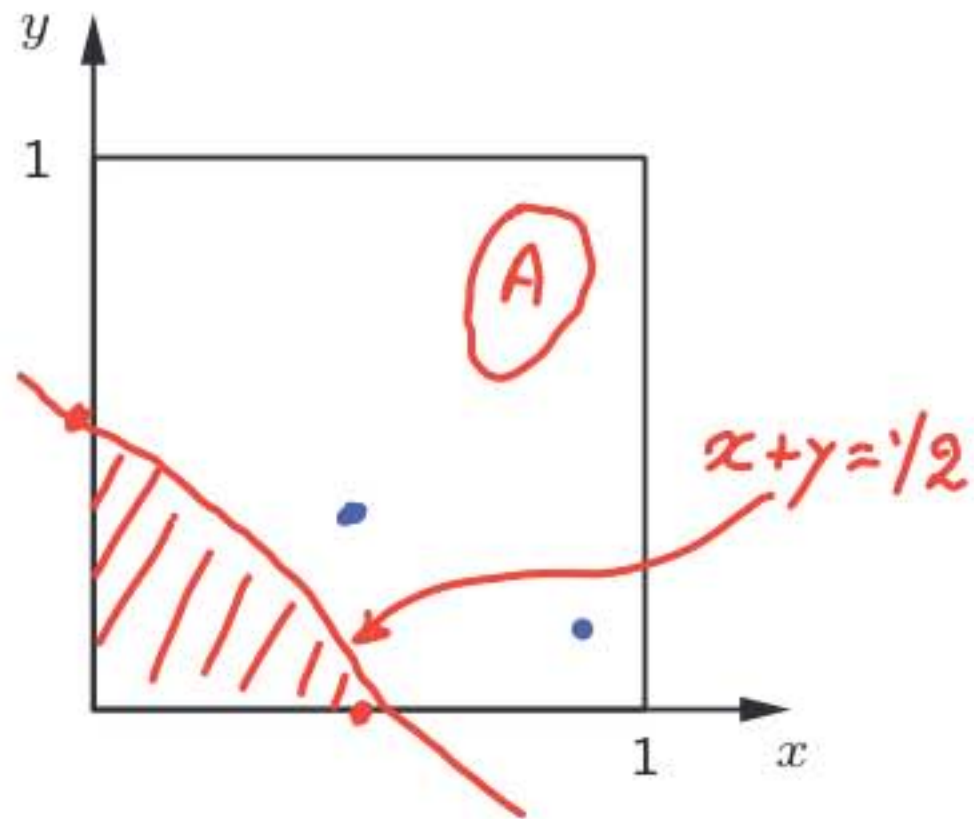
- Assume Ω consists of n equally likely elements
- Assume A consists of k elements

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



Probability calculation: continuous example

- (x, y) such that $0 \leq x, y \leq 1$
- **Uniform** probability law: Probability = Area



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

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

Probability calculation steps

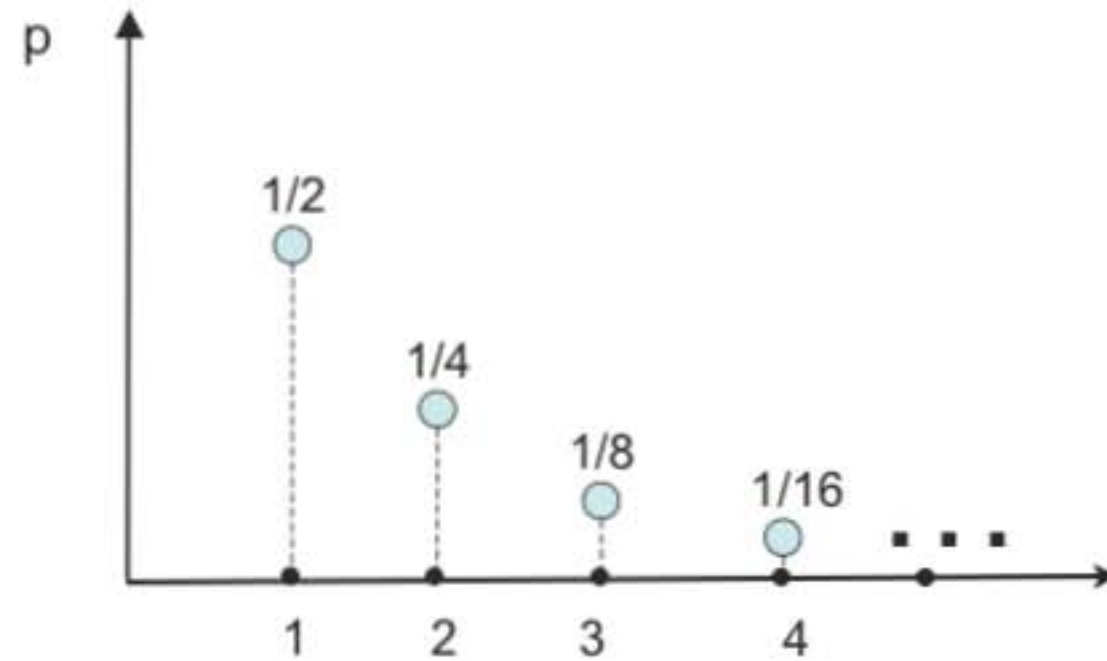
- Specify the sample space
- Specify a probability law
- Identify an event of interest
- Calculate... ●

Probability calculation: discrete but infinite sample space

- Sample space: $\{1, 2, \dots\}$

– We are given $P(n) = \frac{1}{2^n}$, $n = 1, 2, \dots$

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



- $P(\text{outcome is even}) = P(\{2, 4, 6, \dots\})$
 $= P(\{2\} \cup \{4\} \cup \{6\} \cup \dots) \stackrel{=}{=} P(2) + P(4) + P(6) + \dots$

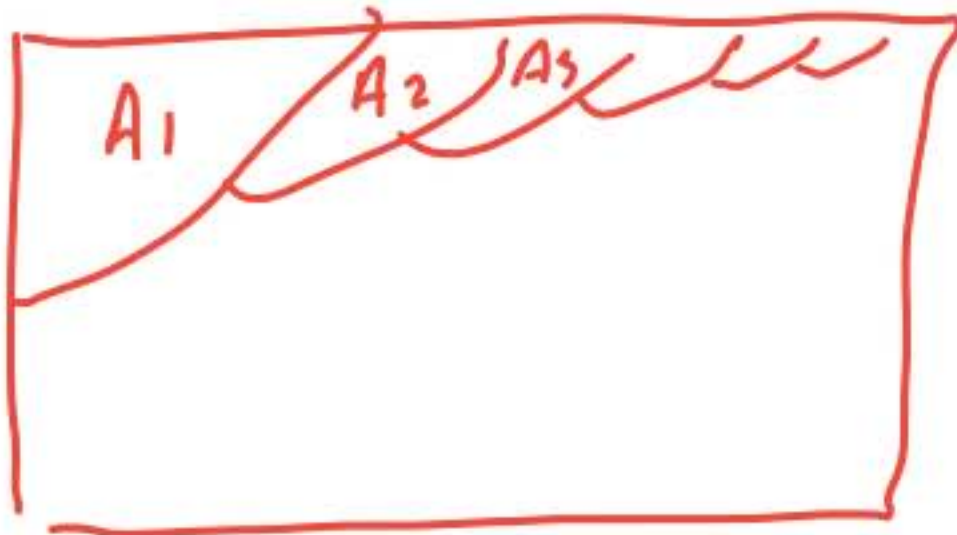
$$= \frac{1}{2^2} + \frac{1}{2^4} + \frac{1}{2^6} + \dots = \frac{1}{4} \left(1 + \frac{1}{4} + \frac{1}{4^2} + \dots \right) = \frac{1}{4} \cdot \frac{1}{1 - \frac{1}{4}} = \frac{1}{3}$$

Countable additivity axiom

- Strengthens the finite additivity axiom

Countable Additivity Axiom:

If A_1, A_2, A_3, \dots is an infinite **sequence** of **disjoint** events, then $P(A_1 \cup A_2 \cup A_3 \cup \dots) = P(A_1) + P(A_2) + P(A_3) + \dots$



Mathematical subtleties

Countable Additivity Axiom:

If A_1, A_2, A_3, \dots is an infinite sequence of disjoint events, then $P(A_1 \cup A_2 \cup A_3 \cup \dots) = P(A_1) + P(A_2) + P(A_3) + \dots$.

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

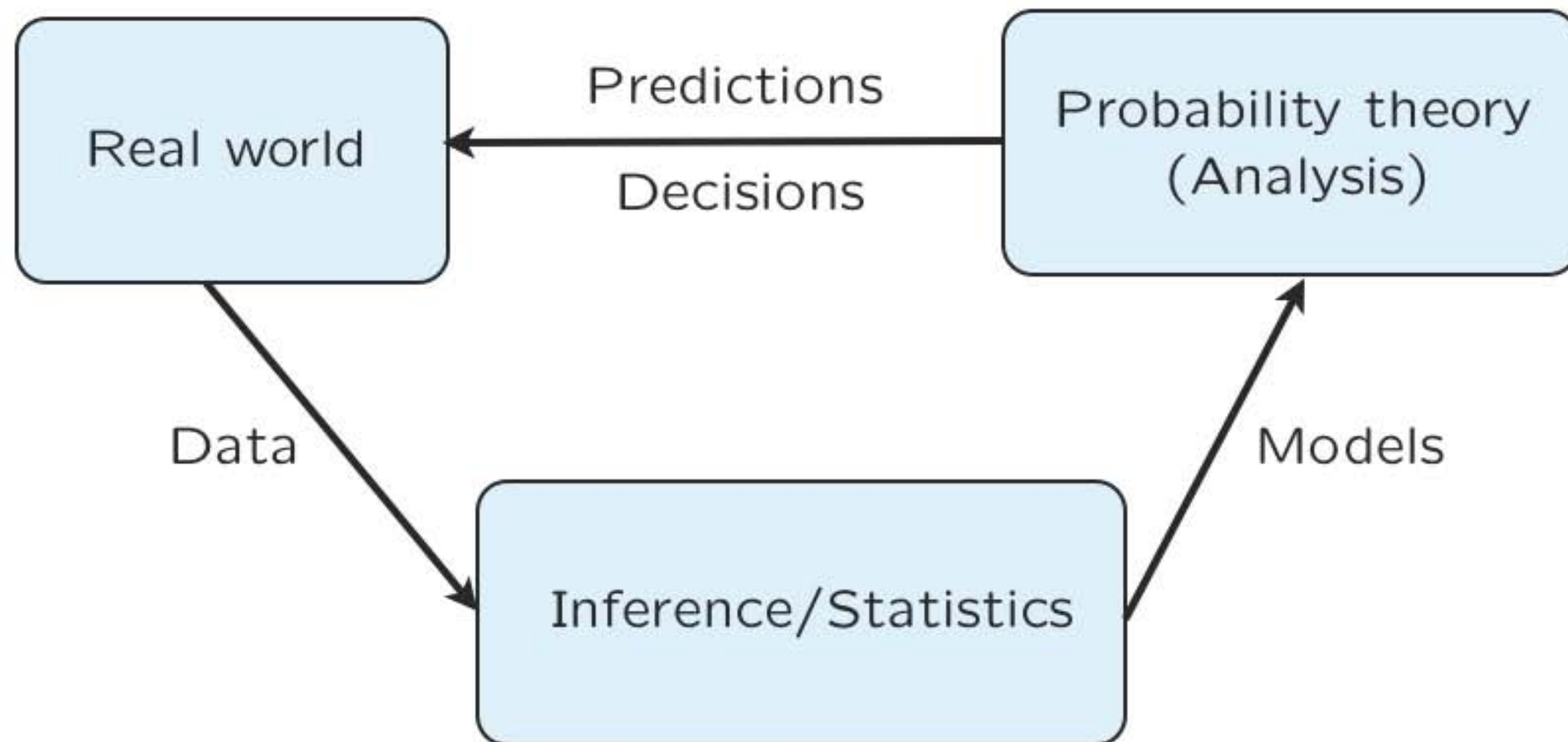
- Additivity holds only for “countable” sequences of events
- The unit square (similarly, the real line, etc.) is **not countable** (its elements cannot be arranged in a sequence)
- “Area” is a legitimate probability law on the unit square, as long as we do not try to assign probabilities/areas to “very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

Interpretations of probability theory

- A narrow view: a branch of math
 - Axioms \Rightarrow theorems
- “Thm:” “Frequency” of event A “is” $P(A)$
- Are probabilities frequencies?
 - $P(\text{coin toss yields heads}) = 1/2$
 - $P(\text{the president of ... will be reelected}) = 0.7$
- Probabilities are often interpreted 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



MIT OpenCourseWare
<https://ocw.mit.edu>

Resource: Introduction to Probability
John Tsitsiklis and Patrick Jaillet

The following may not correspond to a particular course on MIT OpenCourseWare, but has been provided by the author as an individual learning resource.

For information about citing these materials or our Terms of Use, visit: <https://ocw.mit.edu/terms>.