Probability Review

Part 1: Basic Theory

Rebecca Lin | Sunday, September 8th, 2024

Warmup

Example: Flip two "fair" 4-sided dice. How likely are they to sum to 5?

Solution: Let's write down all the possible outcomes.

	2nd dice is 1	2nd dice is 2	2nd dice is 3	2nd dice is 4
1st dice is 1	(1, 1)	(1, 2)	(1, 3)	(1, 4)
1st dice is 2	(2, 1)	(2, 2)	(2, 3)	(2, 4)
1st dice is 4	(3, 1)	(3, 2)	(3, 3)	(3, 4)
1st dice is 4	(4, 1)	(4, 2)	(4, 3)	(4, 4)

16 possible outcomes, **4** of them sum to 5

Definitions

Definition: The set of all possible outcomes is called the **sample space** (Ω).

Ex. A fair dice: $\Omega = \{\text{heads, tails}\}$

Ex. Warmup example: $\Omega = \{1,2,3,4\}^2 = \{(a, b) : a, b \in \{1,2,3,4\}\}$, $|\Omega| = 16$

The phrase "their sum is 5" servers to single out some outcomes of Ω

Definition: A subset E of Ω is called an **event**.

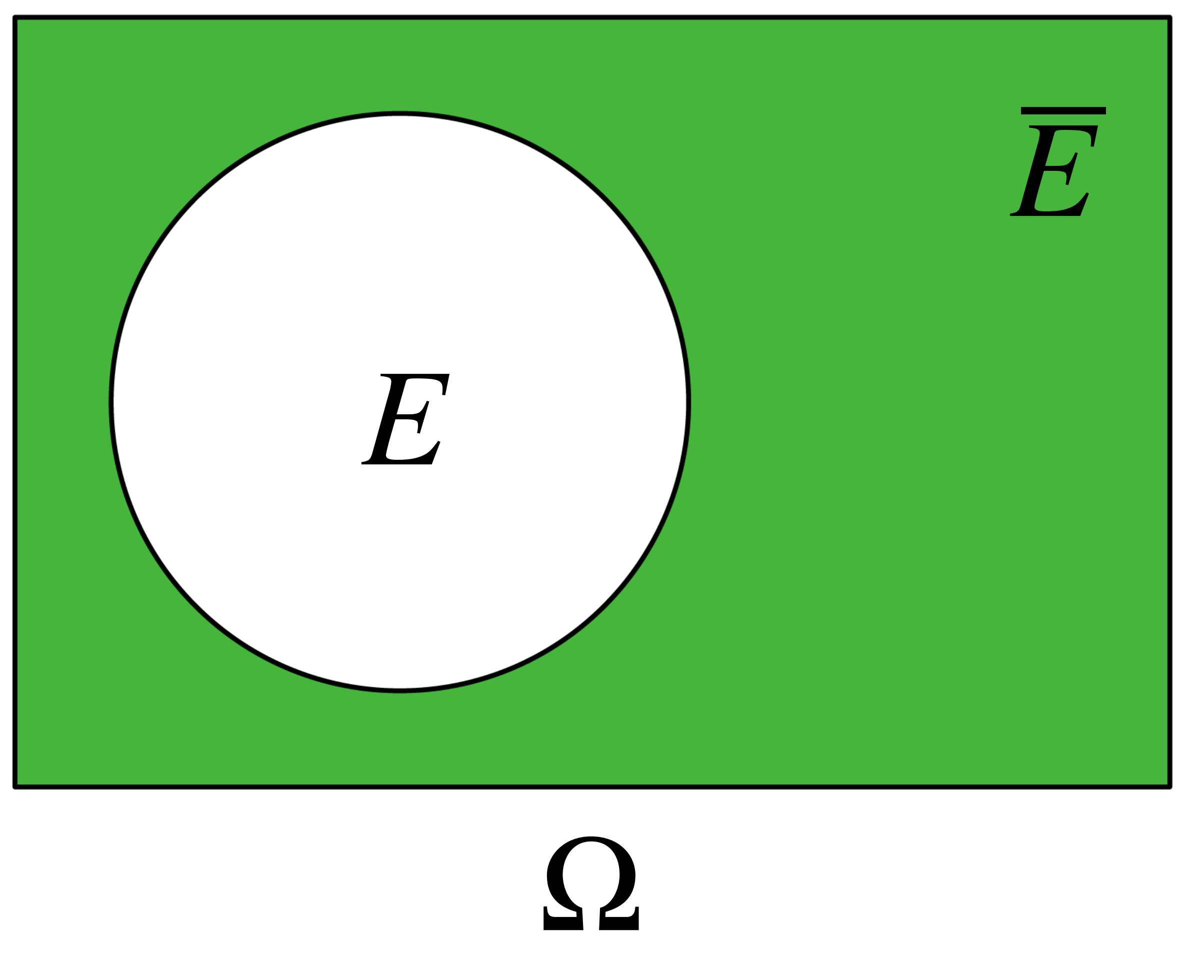
Ex. Warmup example: $E = \{(1,4), (2,3), (3,2), (4,1)\}$

Examples

- Toss a fair coin 3 times:
  $\Omega = \{\mathrm{H}, \mathrm{T}\}^3 = \{\mathrm{HHH}, \mathrm{HHT}, \mathrm{HTH}, \mathrm{THH}, \mathrm{HTT}, \mathrm{THT}, \mathrm{TTH}, \mathrm{TTT}\}$
- $E =$ "3rd toss is tails" $= \{\mathrm{HHT}, \mathrm{HTT}, \mathrm{THT}, \mathrm{TTT}\}$
- $\overline{E} =$ "3rd toss is NOT tails" $= \{\mathrm{HHH}, \mathrm{HTH}, \mathrm{THH}, \mathrm{TTH}\}$

Examples

- Toss a fair coin 3 times:
  $\Omega = \{\mathrm{H}, \mathrm{T}\}^3 = \{\mathrm{HHH}, \mathrm{HHT}, \mathrm{HTH}, \mathrm{THH}, \mathrm{HTT}, \mathrm{THT}, \mathrm{TTH}, \mathrm{TTT}\}$
- $E =$ "3rd toss is tails" $= \{\mathrm{HHT}, \mathrm{HTT}, \mathrm{THT}, \mathrm{TTT}\}$
- $\overline{E} =$ "3rd toss is not tails" $= \{\mathrm{HHH}, \mathrm{HTH}, \mathrm{THH}, \mathrm{TTH}\}$

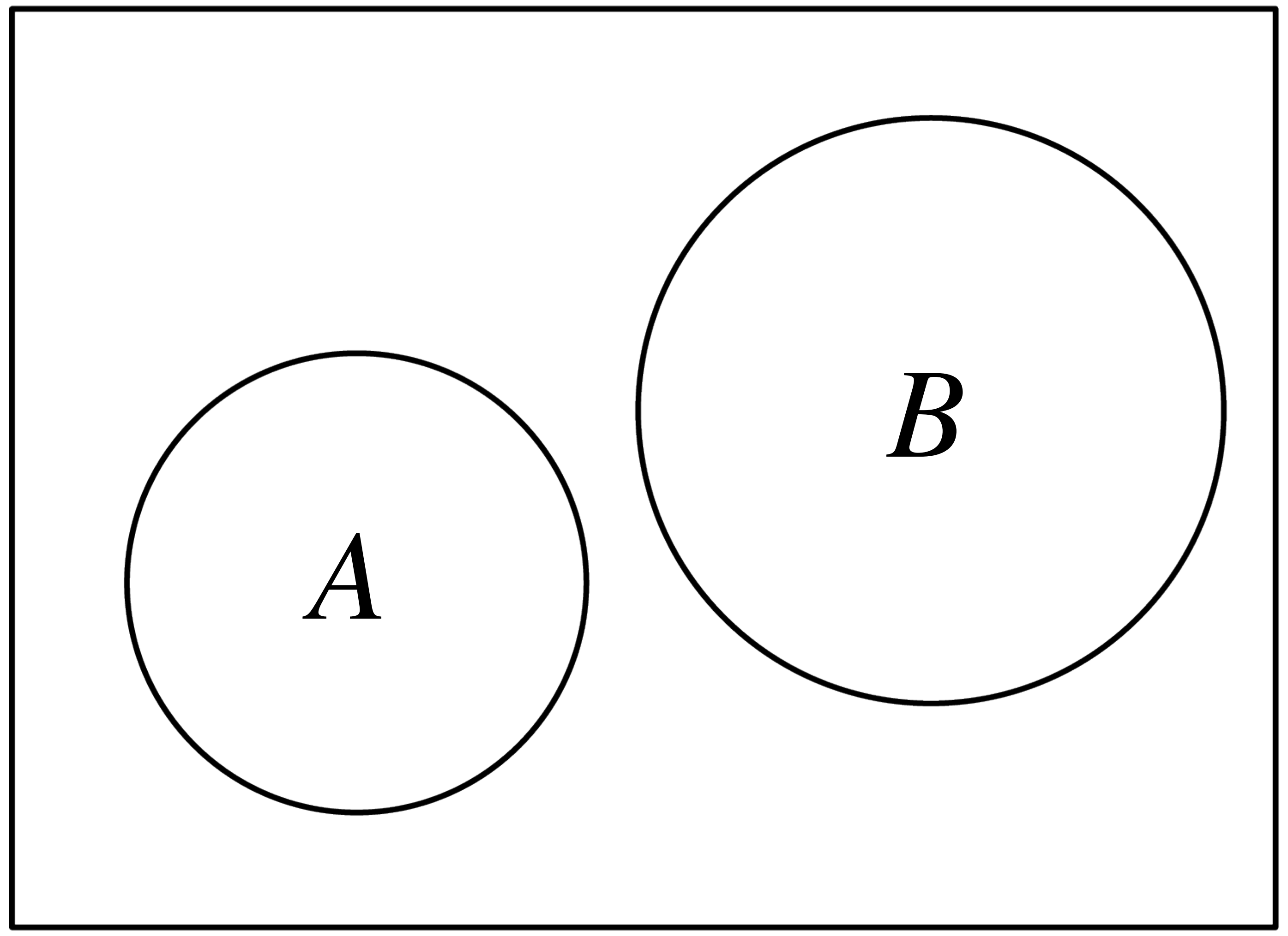


Examples

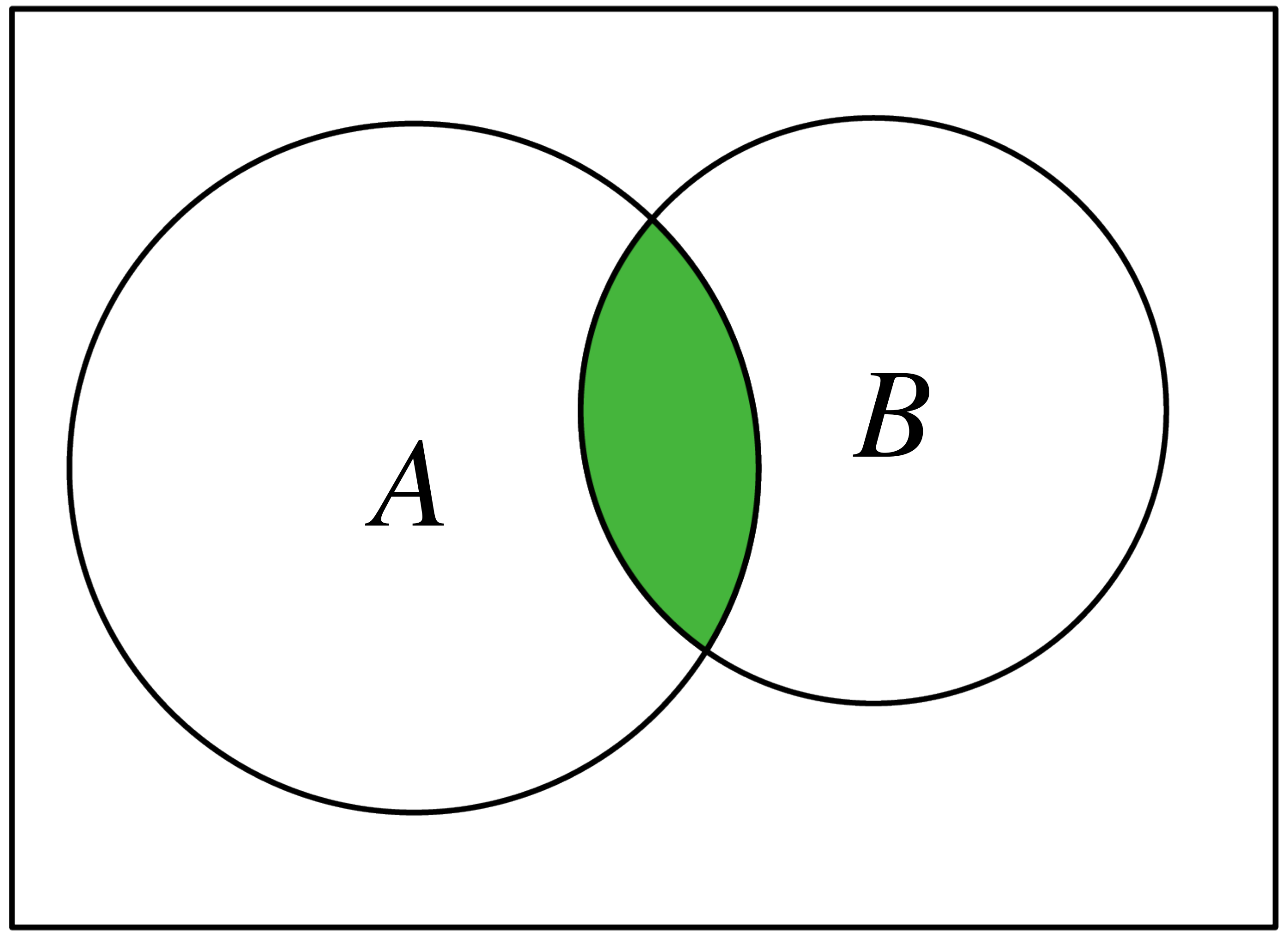
- Toss a fair coin 3 times:
  $\Omega = \{\mathrm{H}, \mathrm{T}\}^3 = \{\mathrm{HHH}, \mathrm{HHT}, \mathrm{HTH}, \mathrm{THH}, \mathrm{HTT}, \mathrm{THT}, \mathrm{TTH}, \mathrm{TTT}\}$
- $E =$ "3rd toss is tails" $= \{\mathrm{HHT}, \mathrm{HTT}, \mathrm{THT}, \mathrm{TTT}\}$
- $\overline{E} =$ "3rd toss is not tails" $= \{\mathrm{HHH}, \mathrm{HTH}, \mathrm{THH}, \mathrm{TTH}\}$
- $F =$ "exactly one toss is heads" $= \{\mathrm{HTT}, \mathrm{THT}, \mathrm{TTH}\}$

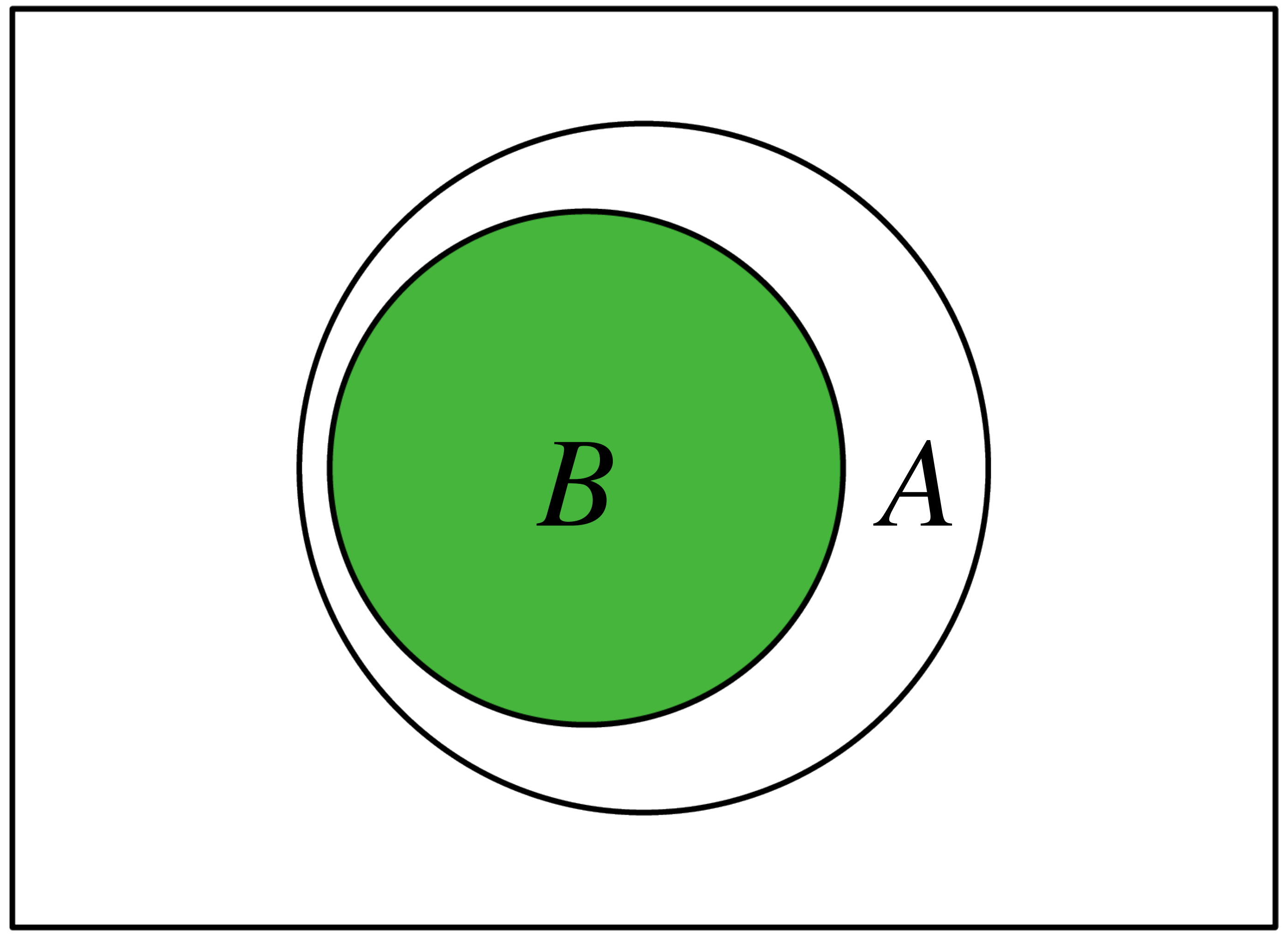
Examples

- Toss a fair coin 3 times:
  $\Omega = \{\text{H}, \text{T}\}^3 = \{\text{HHH}, \text{HHT}, \text{HTH}, \text{THH}, \text{HTT}, \text{THT}, \text{TTH}, \text{TTT}\}$
- $E =$ "3rd toss is tails" $= \{\text{HHT}, \text{HTT}, \text{THT}, \text{TTT}\}$
- $\overline{E} =$ "3rd toss is not tails" $= \{\text{HHH}, \text{HTH}, \text{THH}, \text{TTH}\}$
- $F =$ "exactly one toss is heads" $= \{\text{HTT}, \text{THT}, \text{TTH}\}$
- $G =$ "3rd toss is tails and exactly one toss is heads" $= E \cap F = \{\text{HTT}, \text{THT}\}$



e.g., G





$A \cap B = \varnothing$

A and B cannot occur simultaneously—they are called **disjoint** or **mutually exclusive**

$B \subset A$

if B occurs then A must occur

Examples

- Toss a fair coin 3 times:
  $\Omega = \{\text{H}, \text{T}\}^3 = \{\text{HHH}, \text{HHT}, \text{HTH}, \text{THH}, \text{HTT}, \text{THT}, \text{TTH}, \text{TTT}\}$
- $E =$ "3rd toss is tails" $= \{\text{HHT}, \text{HTT}, \text{THT}, \text{TTT}\}$
- $\overline{E} =$ "3rd toss is not tails" $= \{\text{HHH}, \text{HTH}, \text{THH}, \text{TTH}\}$
- $F =$ "exactly one toss is heads" $= \{\text{HTT}, \text{THT}, \text{TTH}\}$
- $G =$ "3rd toss is tails and exactly one toss is heads" $= E \cap F = \{\text{HTT}, \text{THT}\}$
- $H =$ "3rd toss is tails or exactly one toss is heads" $= E \cup F = \{\text{HHT}, \text{HTT}, \text{THT}, \text{TTH}, \text{TTT}\}$

Probability

Definition

Definition: A **probability** is a way of assigning values $\Pr[E]$ to each event $E \subseteq \Omega$ such that the following axioms hold:

1. $0 \leq \Pr[E] \leq 1$ for all events $E \subseteq \Omega$

2. $\Pr[\Omega] = 1$

3. For any pairwise disjoint sets $E_1, E_2, \ldots$, $\Pr\left[\bigcup_i E_i\right] = \sum_i \Pr[E_i]$

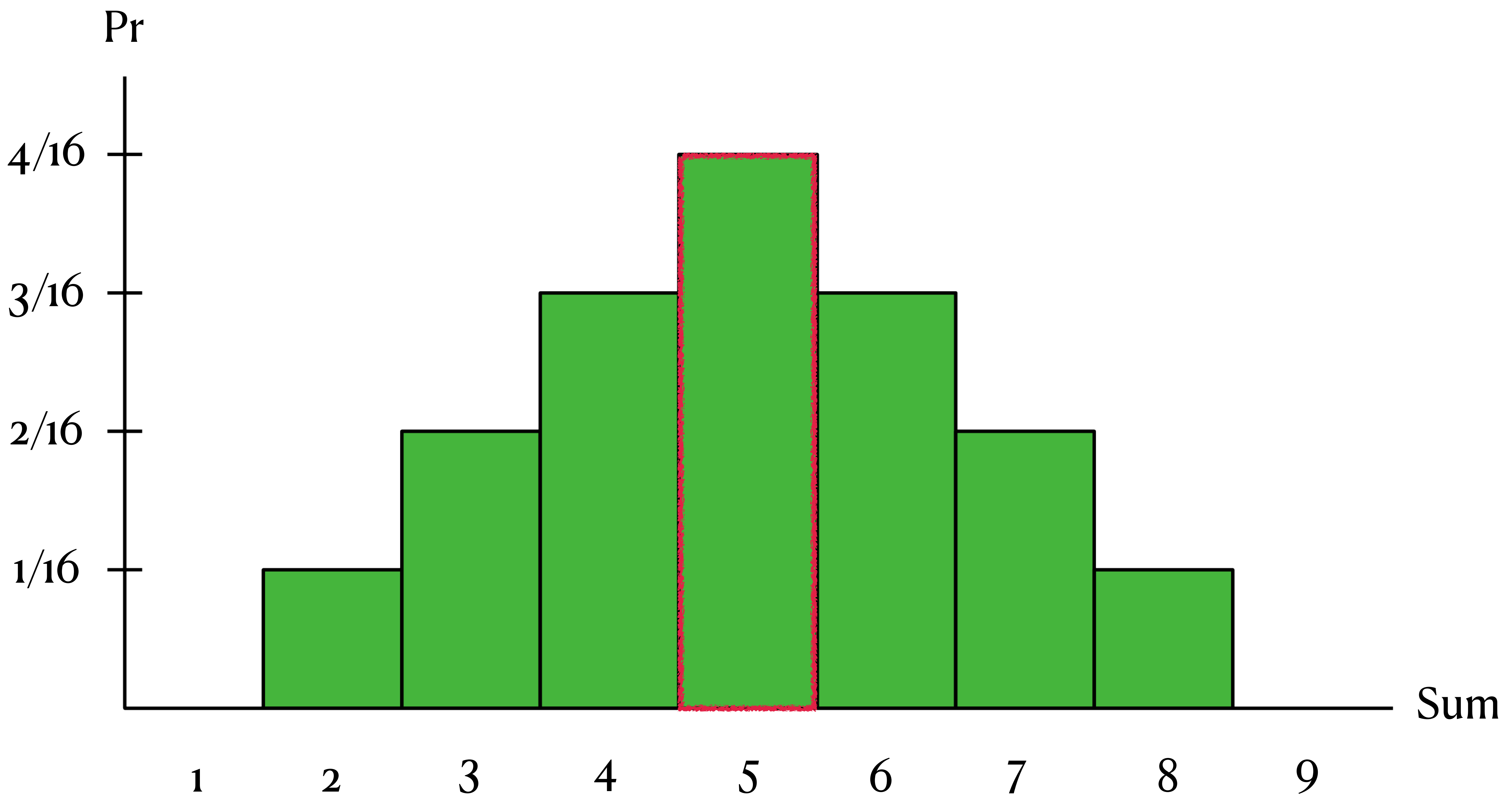
Probability

Recall: Warmup Example

	2nd dice is 1	2nd dice is 2	2nd dice is 3	2nd dice is 4
1st dice is 1	(1, 1)	(1, 2)	(1, 3)	(1, 4)
1st dice is 2	(2, 1)	(2, 2)	(2, 3)	(2, 4)
1st dice is 4	(3, 1)	(3, 2)	(3, 3)	(3, 4)
1st dice is 4	(4, 1)	(4, 2)	(4, 3)	(4, 4)

Probability

Recall: Warmup Example



Probability

Example: Uniform Probability

Suppose Ω is finite. Set $\Pr[E] = \frac{|E|}{|\Omega|}$ for all events E.

Let's check the axioms:

1. $E \subseteq \Omega$, so $0 \leq \frac{|E|}{|\Omega|} \leq 1$
2. $\Pr[\Omega] = \frac{|\Omega|}{|\Omega|} = 1$
3. $\Pr\left[\bigcup_i E_i\right] = \frac{|\sum_i E_i|}{|\Omega|} = \frac{\sum_i |E_i|}{|\Omega|} \underbrace{=}_{def.} \sum_i \Pr[E_i]$

Probability

Example: Uniform Probability

Suppose Ω is finite. Set $\Pr[E] = \dfrac{|E|}{|\Omega|}$ for all events E.

Let's check the axioms:

✓ 1. $E \subseteq \Omega$, so $0 \leq \dfrac{|E|}{|\Omega|} \leq 1$

✓ 2. $\Pr[\Omega] = \dfrac{|\Omega|}{|\Omega|} = 1$

✓ 3. $\Pr\left[\bigcup_i E_i\right] = \dfrac{|\sum_i E_i|}{|\Omega|} = \dfrac{\sum_i |E_i|}{|\Omega|} \underbrace{\equiv}_{def.} \sum_i \Pr[E_i]$

Law of Total Probability

On the Board

Conditional Probability

Intuition

Idea: We're interested in event A. What if someone tells us that event B occurred. What does this this tell us about A?

Ex. Rolling two dice:

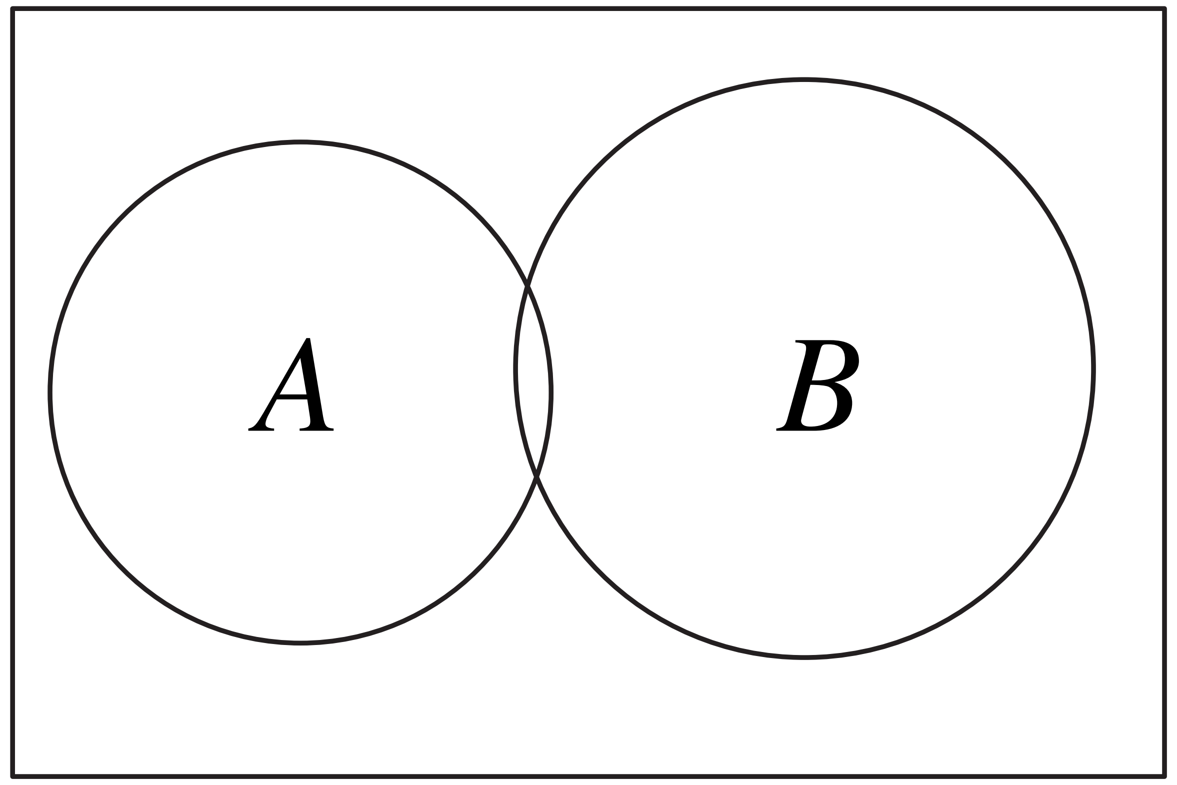
$A =$ "their sum is 3"

$B =$ "1st dice is 4"

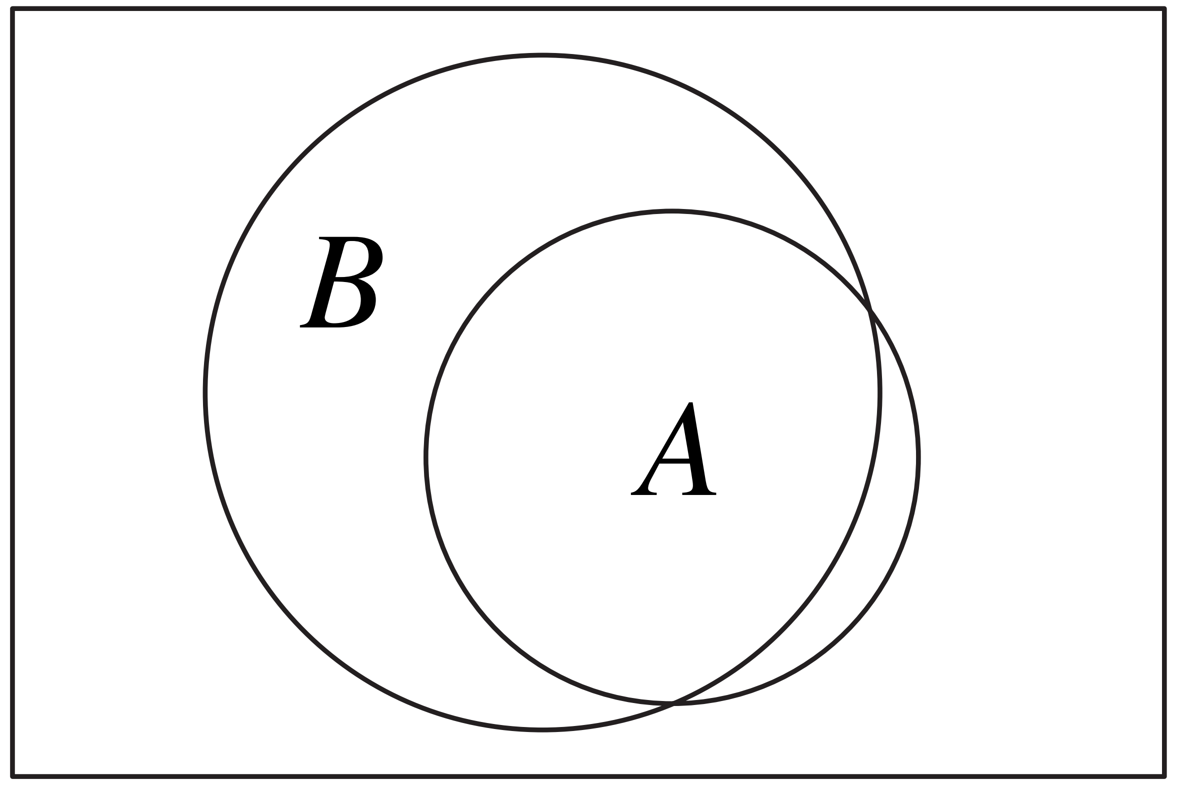
$\Pr[A \text{ given } B] = \Pr[A \mid B] = 0$

Conditional Probability

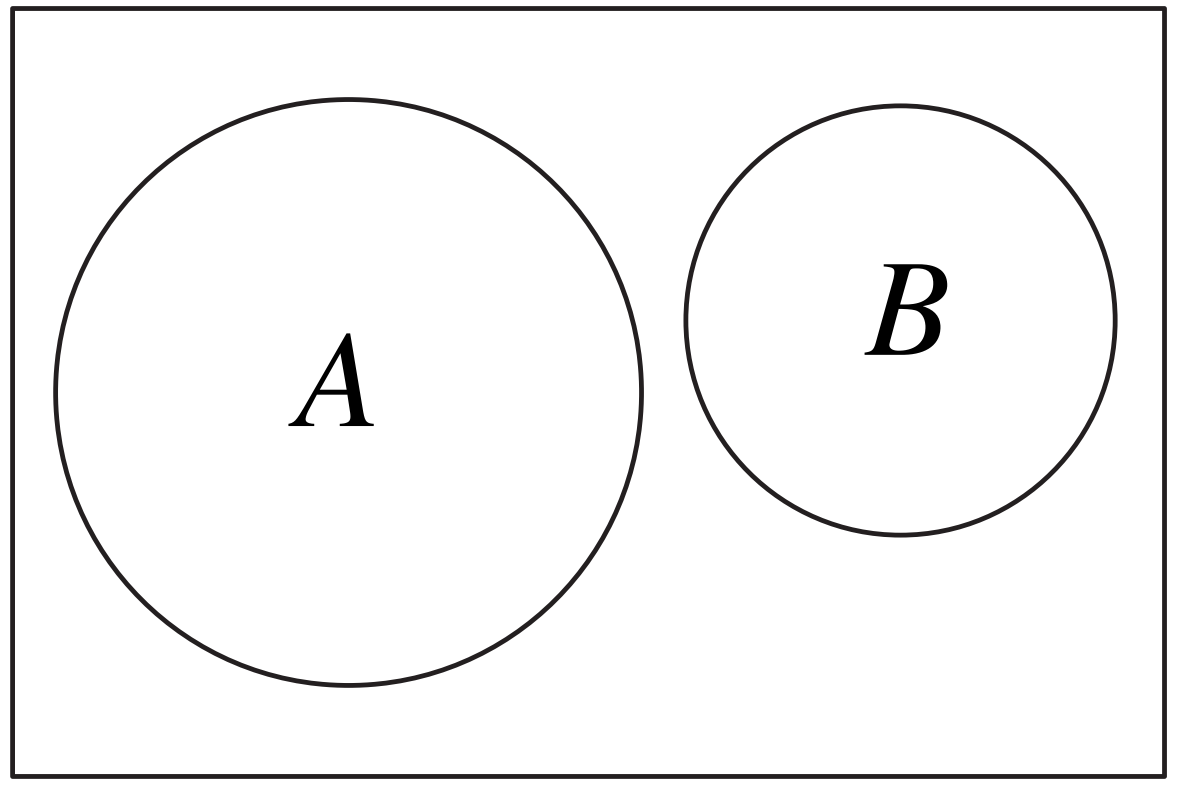
Intuition



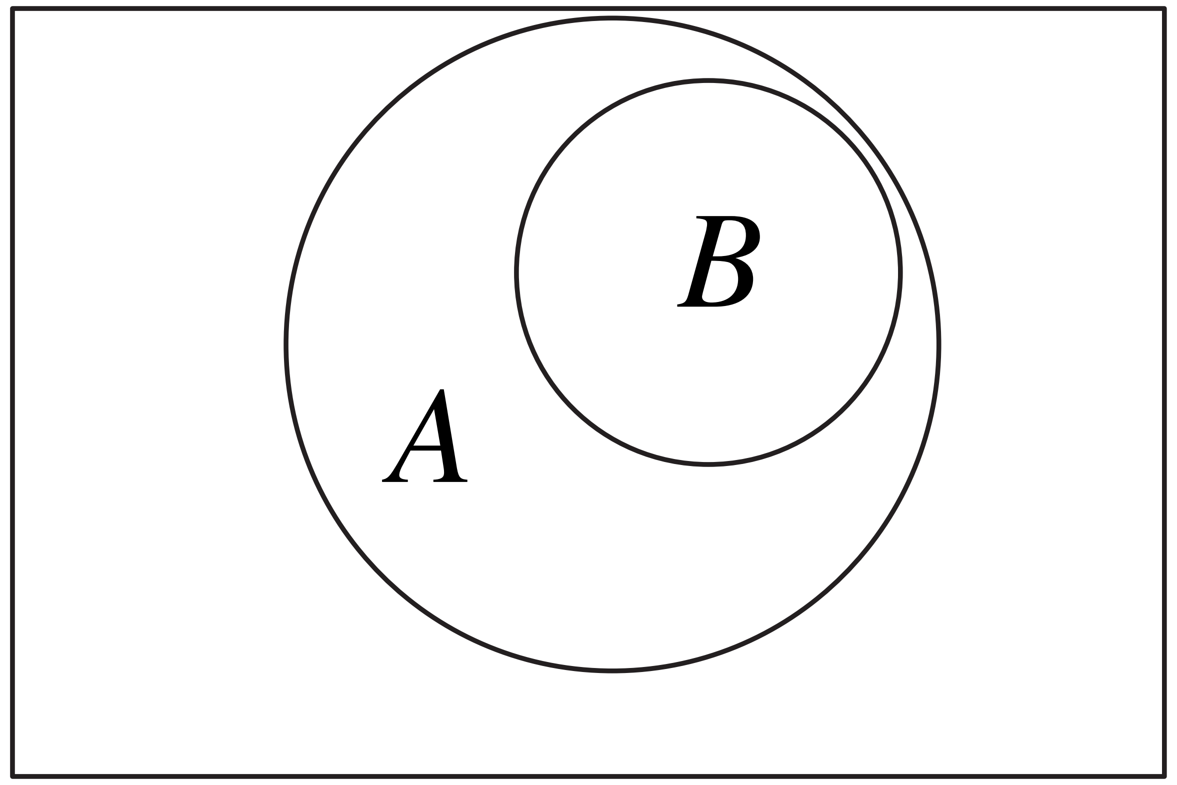
knowing that *B* occurs reduces the likelihood of *A* occurring



given that *B* occurs, it's more likely for *A* to occur



if *B* occurs, then *A* cannot occur



if *B* occurs, then *A* must occur

Conditional Probability

Definition

Definition: Let A and B be two events. Suppose $\Pr[B] > 0$. Then

$$\Pr[A \mid B] = \frac{\Pr[A \cap B]}{\Pr[B]}$$

"the conditional probability of A given that B occurs"
or
"the probability of A given B"

Conditional Probability

Definition

Definition: Let A and B be two events. Suppose $\Pr[B] > 0$. Then

$$\Pr[A \mid B] = \frac{\Pr[A \cap B]}{\Pr[B]}$$

"the conditional probability of A given that B occurs"
or
"the probability of A given B"

$\Pr[B]$ becomes the new sample space, so we normalize by it

Conditional Probability

Definition

Definition: Let A and B be two events. Suppose $\Pr[B] > 0$. Then

$$\Pr[A \mid B] = \frac{\Pr[A \cap B]}{\Pr[B]}$$

"the conditional probability of A given that B occurs"
or
"the probability of A given B"

$\Pr[B]$ becomes the new sample space, so we normalize by it

Bayes' Rule:

$$\Pr[A \mid B] = \frac{\Pr[B \mid A]\Pr[A]}{\Pr[B]}$$

Independence

Idea: B has nothing to do with A, i.e., $\Pr[A] = \Pr[A \mid B]$

Definition: A and B are **independent** if $\Pr[A \cap B] = \Pr[A]\Pr[B]$

Because: $\Pr[A] = \Pr[A \mid B] = \dfrac{\Pr[A \cap B]}{\Pr[B]}$

Question: Suppose we roll two dice. Which of the following events are independent?

A. "1st dice is 1"

B. "2nd dice is 1"

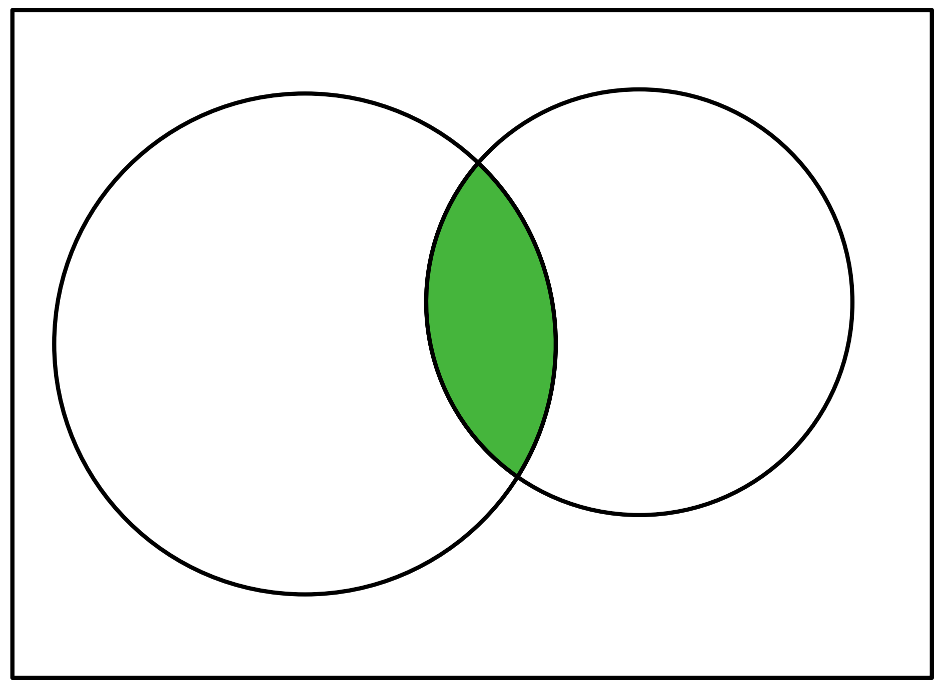
C. "sum is 3"

D. "sum is 7"

Union Bound

Warmup: $A_1, A_2 \subseteq \Omega$

$$\Pr[A_1 \cup A_2] = \Pr[A_1] + \Pr[A_2] - \Pr[A_1 \cap A_2] \leq \Pr[A_1] + \Pr[A_2]$$

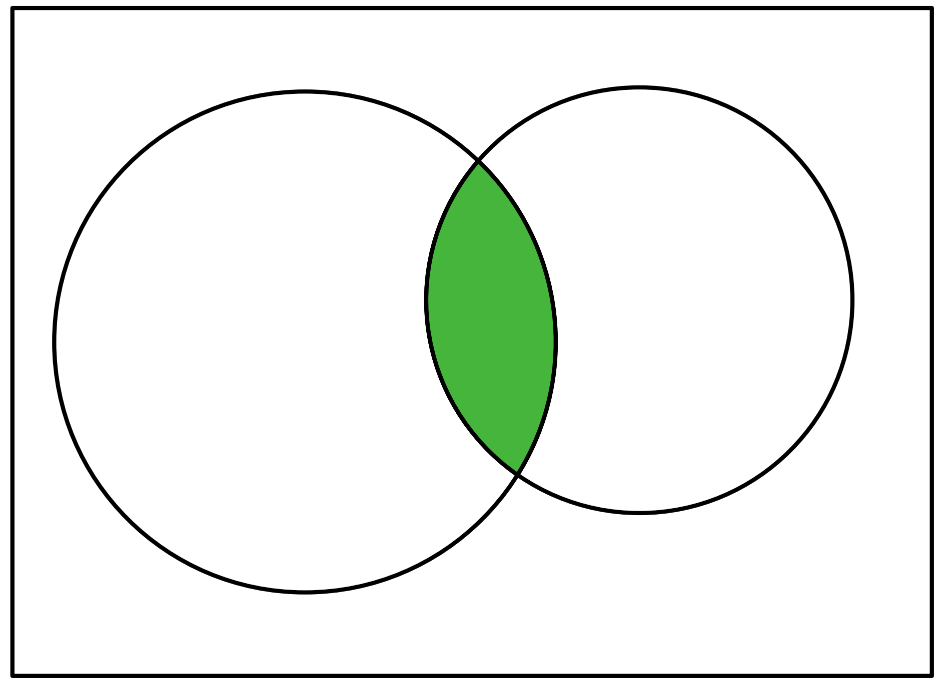


Union Bound

Warmup: $A_1, A_2 \subseteq \Omega$

$$\Pr[A_1 \cup A_2] = \Pr[A_1] + \Pr[A_2] + \Pr[A_1 \cap A_2] \leq \Pr[A_1] + \Pr[A_2]$$

tight when A_1 and A_2 are disjoint

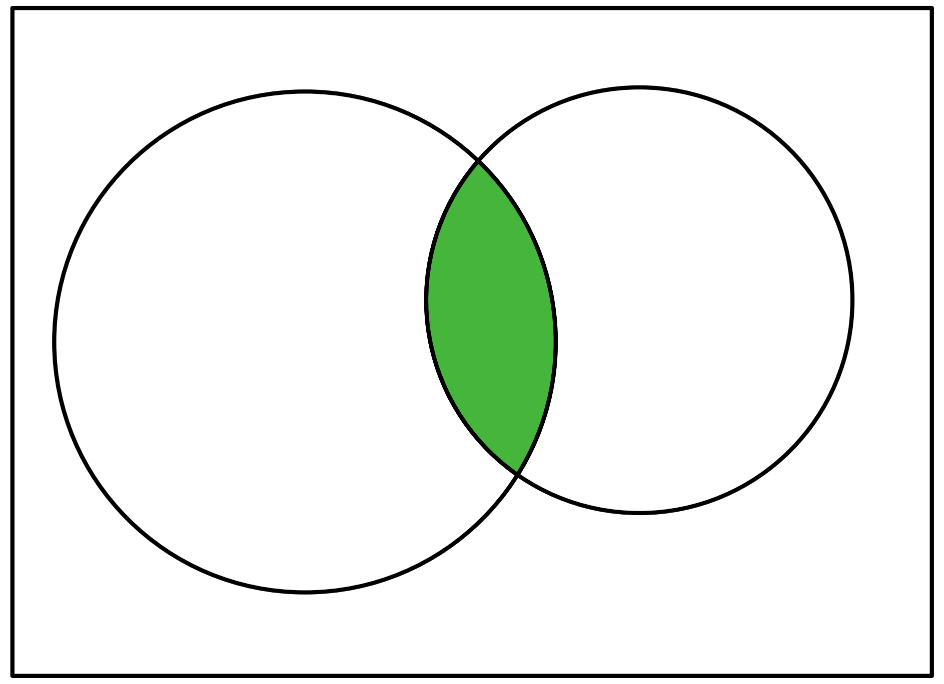


Union Bound

Warmup: $A_1, A_2 \subseteq \Omega$

$$\Pr[A_1 \cup A_2] = \Pr[A_1] + \Pr[A_2] + \boxed{\Pr[A_1 \cap A_2]} \leq \Pr[A_1] + \Pr[A_2]$$

tight when A_1 and A_2 are disjoint



Now let's generalize:

Definition (Union Bound): Given $A_1, A_2, \ldots, A_n \subseteq \Omega$,

$$\Pr[A_1 \cup A_2 \cup \ldots A_n] \leq \sum_{i=1}^{n} \Pr[A_i]$$

Union Bound

Application

Problem: The probability that it rains on any given day it at most 0.01. Bound the probability that at it rains at least once over the next two weeks.

Solution: Let A_i be the probability that it rains on day i for $i = 1,\ldots,14$.

$$\Pr[\text{rains at least once}] = \Pr[A_1 \cap A_2 \cap \ldots \cap A_{14}] \underbrace{\leq}_{u.b.} \sum_{i=1}^{14} A_i = \sum_{i=1}^{14} 0.01 = 0.14$$

Random Variables

On the Board

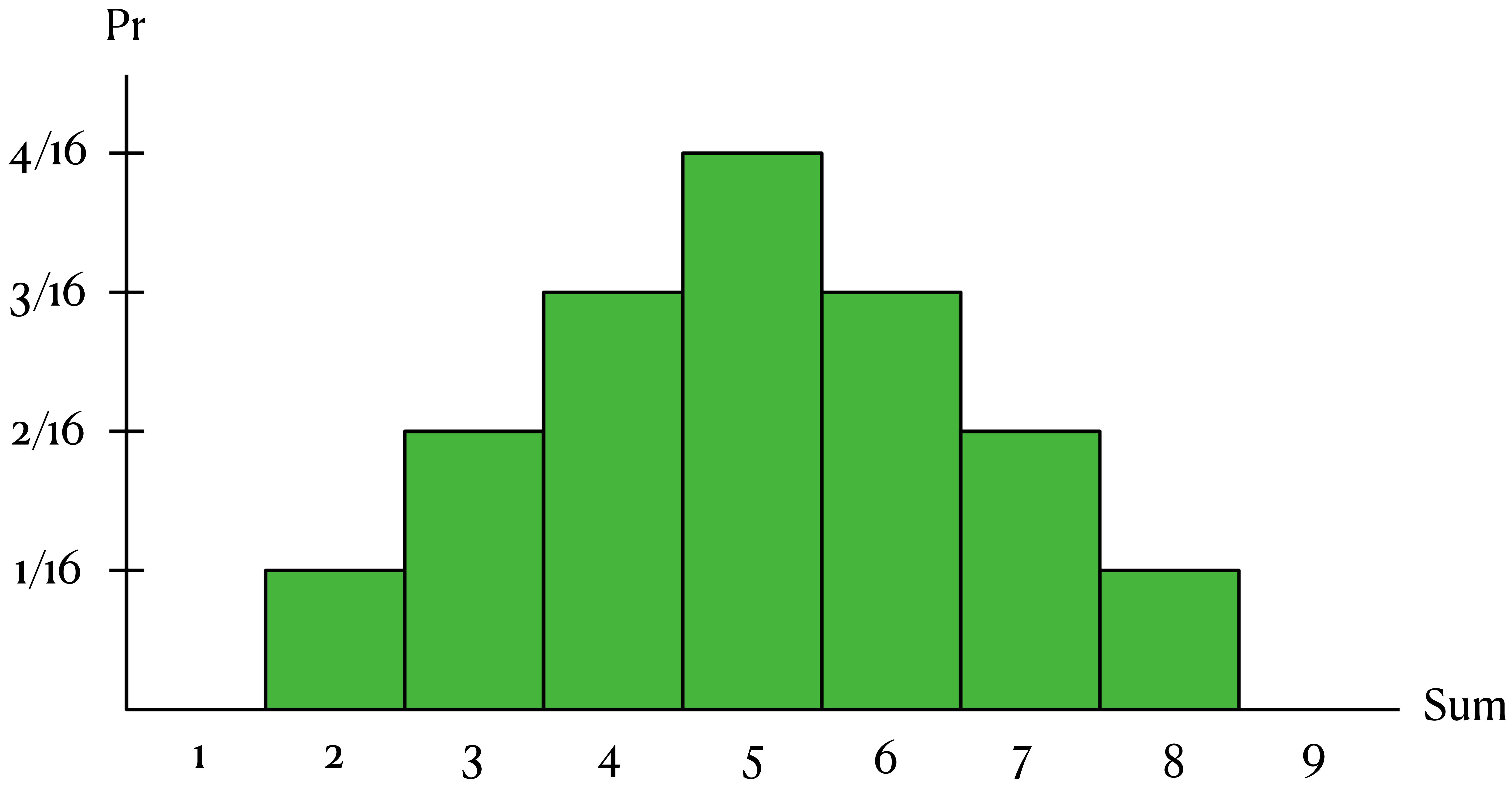
Expectation

Definition: Given a discrete-valued random variables X taking on possible values $x_1, \ldots, x_n$, its **expectation** is defined as

$$\mathbb{E}[X] = \sum_{i=1}^{n} x_i \cdot \Pr[X = x_i]$$

Expectation

Recall: Warmup Example



$$\mathbb{E}[X] = 2 \cdot \frac{1}{16} + 3 \cdot \frac{2}{16} + 4 \cdot \frac{3}{16} + 5 \cdot \frac{4}{16} + 6 \cdot \frac{3}{16} + 7 \cdot \frac{2}{16} + 8 \cdot \frac{1}{16} = 5$$

Expectation

Linearity of Expectation: Given random variables $X_1, \ldots, X_n$ and $X = \sum_{i=1}^{n} X_i$, we have

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathbb{E}[X_i]$$

These variables need not be independent!

Multiplicity of Expectation Under Independence: If X and Y are independent, then

$$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$$

Expectation

Example: On the Board

Variance

Idea: Measures how far the random variable is from its expected value.

Definition:

$$\mathrm{Var}[X] = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right]$$

Linearity of Variance Under Pairwise Independence: Given pairwise independent random variables $X_1, \ldots, X_n$,

$$\mathrm{Var}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathrm{Var}\left[X_i\right]$$

Variance

Examples: On the Board

More Examples

On the Board