Overview of Second Half of the Course

Lecture 16, CS70 Summer 2025

Slides credit: Originals from Josh Hug (CS70 Fall 2024)

A 9 Percent Chance?

"A team of researchers from the University of Hawaiʻi at Mānoa published a study that estimated the probability of a magnitude 9+ earthquake in the Aleutian Islands—an event with sufficient power to create a mega-tsunami especially threatening to Hawaiʻi. In the next 50 years, they report, there is a 9 percent chance of such an event." - The School of Ocean and Earth Science and Technology at the University of Hawaiʻi at Mānoa

Probability and Intuition

Suppose we flip a coin 20 times.

Consider two outcomes:

a) HHHHHHHHHHHHHHHHHHHHH

b) HHTTHTHHHTTHTHTHHTHTH

Which outcome is *more likely*?

Neither! Both equally likely...

Which outcome is *more random*?

This is a deeper question...

Kolmogorov complexity says (b) is – we won't deal with this notion of "random"

Probability and Intuition

Suppose we have two shopping centers.

Shopping center 1:

3 Indian restaurants
2 Moroccan restaurants

Shopping center 2:

1 Indian restaurant
1 Moroccan restaurant

Suppose we pick a center randomly with probability 50%, and then pick a restaurant at that center with equal probability.

- If we eat at a Moroccan restaurant, what is the chance we picked center 1?

Probability and CS70

In CS70, we'll give you a solid foundation for probability.

- What did the "9%" mean in that chance for a mega-tsunami?
- Which outcome of coin flips is more likely?
- What is the chance we picked shopping center 1?

Probabilistic reasoning is famously counterintuitive!

- A firm foundation will keep you safe from easy mistakes.

Probability

Second part of the class – time for a fresh start!

Most important topics/skills from the first part:

- Rigorous reasoning and writing formal proofs
- Counting

Others do come up.... Some we'll talk about in CS70, others not...

- Graphs: Random graph generation or random traversals
- Random transmission errors (for error detection/correction)
- ...

Counting Summary

Parts of the counting lecture we'll need today:

- Given n items, there are $n!$ permutations.
- If we draw k samples from a set of n objects, the number of possible draws is:
  - With replacement, where order matters: n^k
  - Without replacement, where order matters: $n \times (n-1) \times (n-2) \times \cdots \times (n-k+1)$
  - Without replacement, where order doesn't matter: $\binom{n}{k}$

Probability Spaces and Events

Lecture 16, CS70 Summer 2025

Probability Basics

Much like part one of this class, only a few terms/concepts

It's the variations and application that you need to practice with!

Concepts for today:

- Random experiments
- Outcomes
- Sample spaces (just the set of all outcomes)
- Probability spaces (just a function/measure defined on a sample space)
- Events (just a subset of a sample space)

That's it!

The rest is just counting...

Random Experiment, Sample Point, Sample Space

Random Experiment: Drawing a sample of k items from a set S with cardinality n.

Example: Rolling two 4-sided dice. In this case $S = \{1, 2, 3, 4\}$, $k = 2$, $n = 4$, and we are drawing with replacement.

Sample Point: The outcome of a single experiment, e.g. 22.

Sample Space: Set of all possible outcomes, given by Ω.

$\Omega =$

11	12	13	14
21	22	23	24
31	32	33	34
41	42	43	44

Probability Space

A Probability Space consists of:

- A sample space Ω
- A probability function $\boldsymbol{P}(\omega)$ that maps $\omega \in \Omega$ to a real number such that:
  - $0 \leq \boldsymbol{P}(\omega) \leq 1$ for all $\omega \in \Omega$
  - $\sum_{w \in \Omega} \boldsymbol{P}(\omega) = 1$

Example: Rolling two fair 4-sided dice. In this case, Ω and $\boldsymbol{P}$ are:

$\Omega =$

11	12	13	14
21	22	23	24
31	32	33	34
41	42	43	44

$$\boldsymbol{P}(\omega) = \frac{1}{16}$$

This is a "uniform probability space".

Event

Example: Rolling two 4-sided dice. In this case S = {1, 2, 3, 4}, k = 2, and we are drawing without replacement. If dice are fair then:

$\Omega =$

11	12	13	14
21	22	23	24
31	32	33	34
41	42	43	44

$$\boldsymbol{P}(\omega) = \frac{1}{16}$$

An event A is a subset of our sample space Ω.

- Example: Let $A_{2ndgreater}$ be the event where the second roll is greater than first.
  - $A_{2ndgreater} = \{12, 13, 14, 23, 24, 34\}$

Naturally, $\boldsymbol{P}(A) = \sum_{\omega \in A} \boldsymbol{P}(\omega)$, e.g. $\boldsymbol{P}(A_{2ndgreater}) = 6/16 = 3/8$

Quick Note on Notation

Slides use parentheses for probability functions, e.g.

$$\boldsymbol{P}(\omega)$$

The notes use brackets and a blackboard P.

$$\mathbb{P}[\boldsymbol{\omega}]$$

Reminder: Mathematical symbols are just placeholders for ideas.

- These two notations are exactly equivalent.

Sample Space Ignoring Order

Note: It is possible to define your sample space such that it ignores the order of the samples, e.g. treats 21 and 12 as the same object.

For two four-sided dice: $\Omega = \{11, 12, 13, 14, 22, 23, 24, 33, 34, 44\}$

- 10 outcomes rather than 16.
- Probability space is non-uniform and thus harder to work with.
  - $P(11) = 1/16$
  - $P(12) = 1/8$

We'll almost always take into account order in defining our sample spaces.

$\Omega =$

11	12	13	14
21	22	23	24
31	32	33	34
41	42	43	44

vs.

$\Omega =$

11	12	13	14
	22	23	24
		33	34
			44

Non-uniform Probability Spaces

Lecture 16, CS70 Summer 2025

Another Example Probability Space

Example: Rolling two unfair 4-sided dice. In this case S = {1, 2, 3, 4}, k = 2, and we are drawing without replacement. Suppose that dice are loaded so that 1 appears 5/8 times, and other numbers occurs 1/8 times.

In this example $\boldsymbol{P}(\omega)$ is slightly more complicated:

- $\boldsymbol{P}(\mathbf{11}) = ??$

$\Omega =$

11	12	13	14
21	22	23	24
31	32	33	34
41	42	43	44

Another Example Probability Space

Example: Rolling two unfair 4-sided dice. In this case S = {1, 2, 3, 4}, k = 2, and we are drawing without replacement. Suppose that dice are loaded so that 1 appears 5/8 times, and other numbers occurs 1/8 times.

In this example $\boldsymbol{P}(\omega)$ is slightly more complicated:

- $\boldsymbol{P}(\mathbf{11}) = 25/64$

$\Omega =$

11	12	13	14
21	22	23	24
31	32	33	34
41	42	43	44

Another Example Probability Space

Example: Rolling two unfair 4-sided dice. In this case S = {1, 2, 3, 4}, k = 2, and we are drawing without replacement. Suppose that dice are loaded so that 1 appears 5/8 times, and other numbers occurs 1/8 times.

In this example $\boldsymbol{P}(\omega)$ is slightly more complicated:

- $\boldsymbol{P}(11) = 25/64$
- $\boldsymbol{P}(21) = \boldsymbol{P}(31) = \boldsymbol{P}(41) = \boldsymbol{P}(12) = \boldsymbol{P}(13) = \boldsymbol{P}(14) = ??$

$$\Omega = \begin{matrix} 11 & 12 & 13 & 14 \\ 21 & 22 & 23 & 24 \\ 31 & 32 & 33 & 34 \\ 41 & 42 & 43 & 44 \end{matrix}$$

Another Example Probability Space

Example: Rolling two unfair 4-sided dice. In this case S = {1, 2, 3, 4}, k = 2, and we are drawing without replacement. Suppose that dice are loaded so that 1 appears 5/8 times, and other numbers occurs 1/8 times.

In this example $\boldsymbol{P}(\omega)$ is slightly more complicated:

- $\boldsymbol{P}(11) = 25/64$
- $\boldsymbol{P}(21) = \boldsymbol{P}(31) = \boldsymbol{P}(41) = \boldsymbol{P}(12) = \boldsymbol{P}(13) = \boldsymbol{P}(14) = 5/64$

$\Omega =$

11	12	13	14
21	22	23	24
31	32	33	34
41	42	43	44

This is a non-uniform probability space.

Another Example Probability Space

Example: Rolling two unfair 4-sided dice. In this case S = {1, 2, 3, 4}, k = 2, and we are drawing without replacement. Suppose that dice are loaded so that 1 appears 5/8 times, and other numbers occurs 1/8 times.

In this example $\boldsymbol{P}(\omega)$ is slightly more complicated:

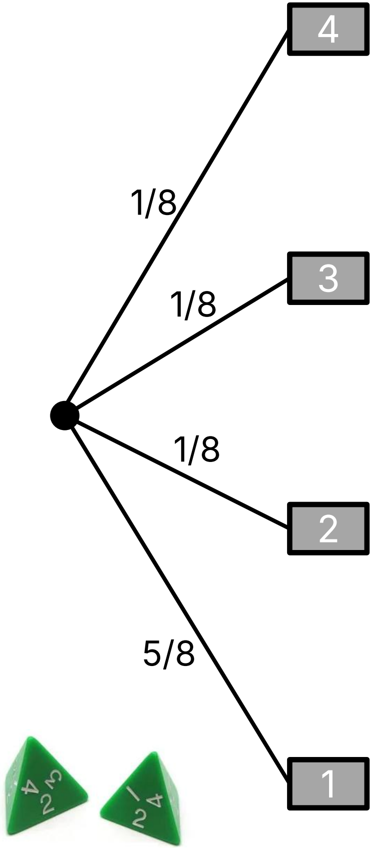
- $\boldsymbol{P}(11) = 25/64$
- $\boldsymbol{P}(21) = \boldsymbol{P}(31) = \boldsymbol{P}(41) = \boldsymbol{P}(12) = \boldsymbol{P}(13) = \boldsymbol{P}(14) = 5/64$
- $\boldsymbol{P}(22) = \boldsymbol{P}(23) = \cdots = \boldsymbol{P}(44) = 1/64$

$\Omega =$

11	12	13	14
21	22	23	24
31	32	33	34
41	42	43	44

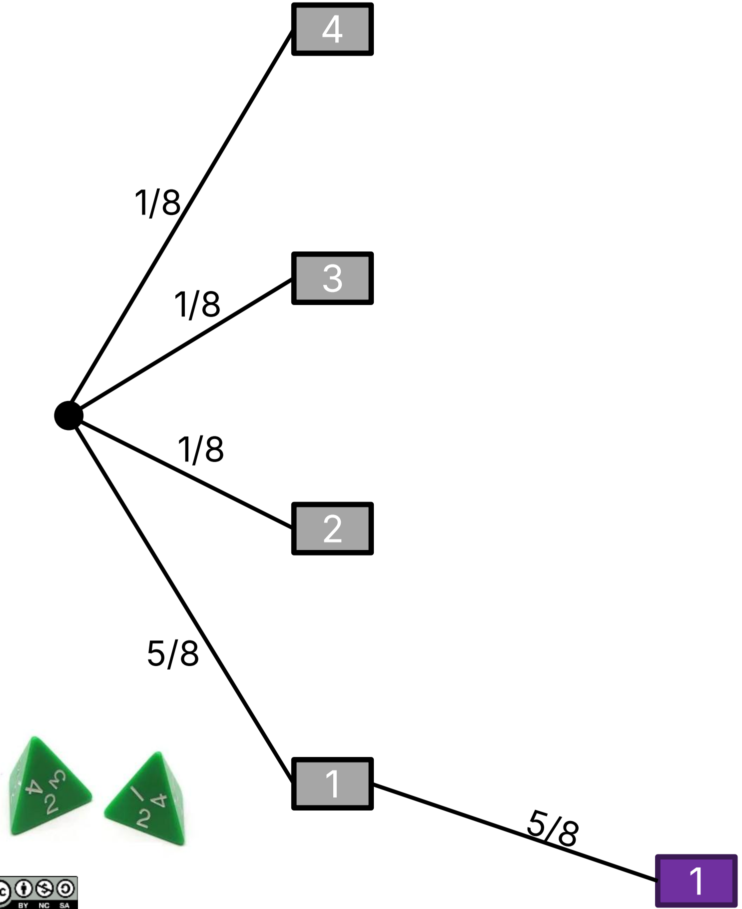
This is a non-uniform probability space.

Tree Diagram for a Non-Uniform Probability Space



A tree diagram gives a somewhat more formal explanation of the probability of each outcome.

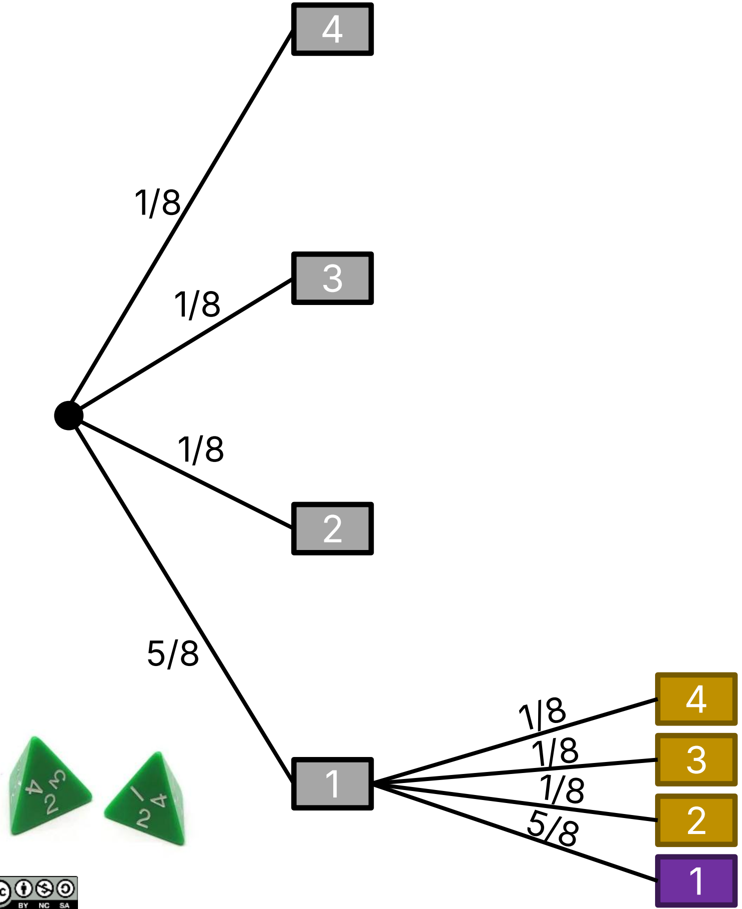
Tree Diagram for a Non-Uniform Probability Space



- $\boldsymbol{P}(11) = \frac{5}{8} \times \frac{5}{8} = 25/64$

The tree diagram gives a somewhat more formal explanation of the probability of each outcome.

Tree Diagram for a Non-Uniform Probability Space

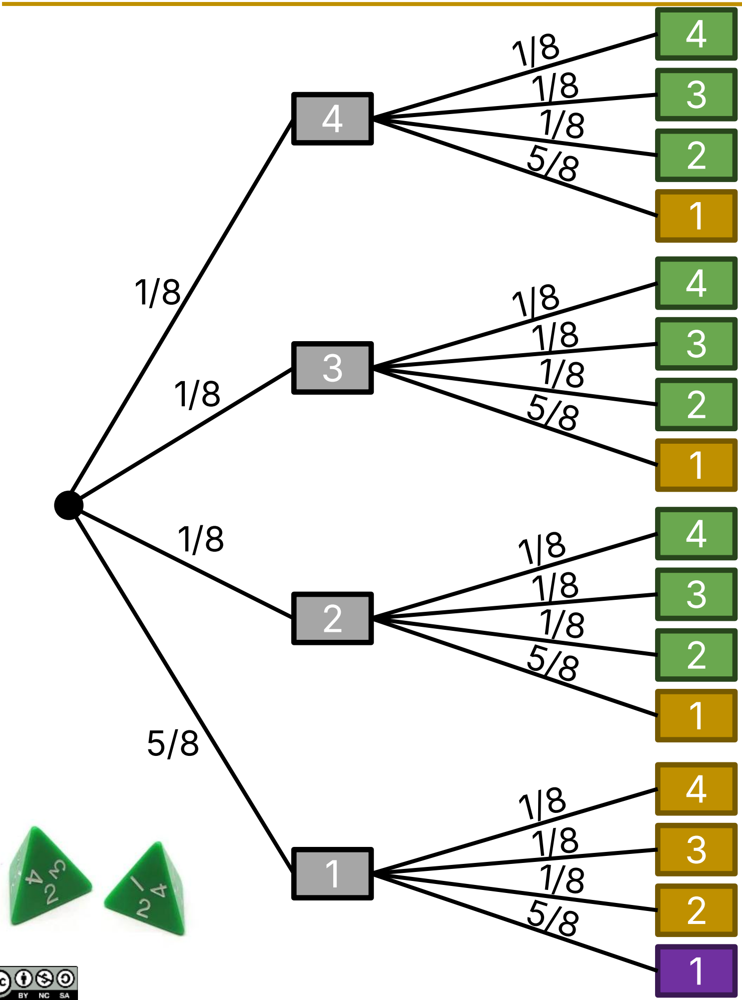


- $\boldsymbol{P}(11) = \frac{5}{8} \times \frac{5}{8} = 25/64$
- $\boldsymbol{P}(12) = \boldsymbol{P}(13) = \boldsymbol{P}(14) = \frac{5}{8} \times \frac{1}{8} = 5/64$

The tree diagram gives a somewhat more formal explanation of the probability of each outcome.

Here, we can multiply the probabilities together because the dice rolls are independent. More on this in a later lecture.

Tree Diagram for a Non-Uniform Probability Space



- $\boldsymbol{P}(11) = 25/64$
- $\boldsymbol{P}(12) = \boldsymbol{P}(13) = \boldsymbol{P}(14) = \boldsymbol{P}(21) = \boldsymbol{P}(31) = \boldsymbol{P}(41) = 5/64$
- $\boldsymbol{P}(22) = \boldsymbol{P}(23) = \cdots = \boldsymbol{P}(44) = 1/64$

The tree diagram gives a somewhat more formal explanation of the probability of each outcome.

Here, we can multiply the probabilities together because the dice rolls are independent. More on this in a later lecture.

Events in a Non Uniform Probability Space

Example: Rolling two unfair 4-sided dice. In this case S = {1, 2, 3, 4}, k = 2, and we are drawing without replacement. Suppose that dice are loaded so that 1 appears 5/8 times, and other numbers occurs 1/8 times.

In this example $\boldsymbol{P}(\omega)$ yields a non-uniform probability space:

- $\boldsymbol{P}(11) = 25/64$
- $\boldsymbol{P}(21) = \boldsymbol{P}(31) = \boldsymbol{P}(41) = \boldsymbol{P}(12) = \boldsymbol{P}(13) = \boldsymbol{P}(14) = 5/64$
- $\boldsymbol{P}(22) = \boldsymbol{P}(23) = \cdots = \boldsymbol{P}(44) = 1/64$

$\Omega =$

11	12	13	14
21	22	23	24
31	32	33	34
41	42	43	44

What is $\boldsymbol{P}(A_{2ndgreater})$?

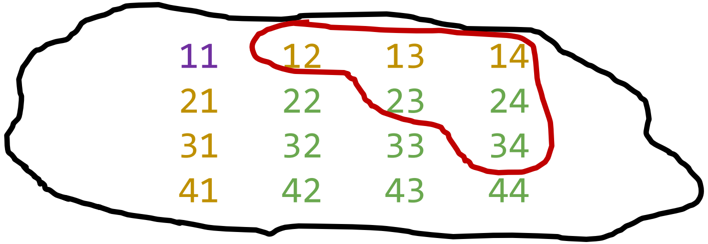
Events in a Non Uniform Probability Space

Example: Rolling two unfair 4-sided dice. In this case S = {1, 2, 3, 4}, k = 2, and we are drawing without replacement. Suppose that dice are loaded so that 1 appears 5/8 times, and other numbers occurs 1/8 times.

In this example $\boldsymbol{P}(\omega)$ yields a non-uniform probability space:

- $\boldsymbol{P}(11) = 25/64$
- $\boldsymbol{P}(21) = \boldsymbol{P}(31) = \boldsymbol{P}(41) = \boldsymbol{P}(12) = \boldsymbol{P}(13) = \boldsymbol{P}(14) = 5/64$
- $\boldsymbol{P}(22) = \boldsymbol{P}(23) = \cdots = \boldsymbol{P}(44) = 1/64$

$\Omega =$



What is $\boldsymbol{P}(A_{2nd greater})$?

$$\boldsymbol{P}(A_{2nd greater}) = 3 \times 5/64 + 3 \times 1/64$$
$$= 18/64$$
$$= 9/32$$

Example: Four Biased Coins

Lecture 16, CS70 Summer 2025

Four Biased Coins

One common model in engineering applications is the biased coin. Can be used to model, e.g. chance that a system will fail.

Suppose we have a coin where $\boldsymbol{P}(H) = 1/3$ and $\boldsymbol{P}(T) = 2/3$.

Suppose that our experiment consists of four consecutive flips.

What is the chance that we get HHHH? The chance that we get HTHT?

$\Omega =$

HHHH	HHHT	HHTH	HHTT
HTHH	HTHT	HTTH	HTTT
THHH	THHT	THTH	THTT
TTHH	TTHT	TTTH	TTTT

Four Biased Coins

One common model in engineering applications is the biased coin. Can be used to model, e.g. chance that a system will fail.

Suppose we have a coin where $\boldsymbol{P}(H) = 1/3$ and $\boldsymbol{P}(T) = 2/3$.

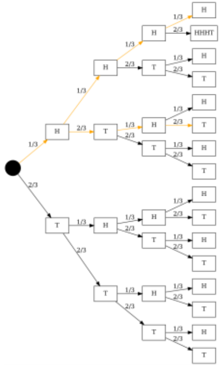
Suppose that our experiment consists of four consecutive flips.

What is the chance that we get HHHH? The chance that we get HTHT?

- $P(HHHH) = (1/3)^4$, $P(HTHT) = (1/3)^2(2/3)^2$

$\Omega =$

HHHH	HHHT	HHTH	HHTT
HTHH	HTHT	HTTH	HTTT
THHH	THHT	THTH	THTT
TTHH	TTHT	TTTH	TTTT



Four Biased Coins

One common model in engineering applications is the biased coin. Can be used to model, e.g. chance that a system will fail.

Suppose we have a coin where $\boldsymbol{P}(H) = 1/3$ and $\boldsymbol{P}(T) = 2/3$.

Suppose that our experiment consists of four consecutive flips.

What is the chance that all the flips are the same?

- $P(A_{same}) = P(HHHH) + P(TTTT) = (1/3)^4 + (2/3)^4 = 17/81$

$\Omega =$

HHHH	HHHT	HHTH	HHTT
HTHH	HTHT	HTTH	HTTT
THHH	THHT	THTH	THTT
TTHH	TTHT	TTTH	TTTT

Four Biased Coins

One common model in engineering applications is the biased coin. Can be used to model, e.g. chance that a system will fail.

Suppose we have a coin where $\boldsymbol{P}(H) = 1/3$ and $\boldsymbol{P}(T) = 2/3$.

Suppose that our experiment consists of four consecutive flips.

What is the chance that we have two heads?

$\Omega =$

HHHH	HHHT	HHTH	HHTT
HTHH	HTHT	HTTH	HTTT
THHH	THHT	THTH	THTT
TTHH	TTHT	TTTH	TTTT

Four Biased Coins

One common model in engineering applications is the biased coin. Can be used to model, e.g. chance that a system will fail.

Suppose we have a coin where $\boldsymbol{P}(H) = 1/3$ and $\boldsymbol{P}(T) = 2/3$.

Suppose that our experiment consists of four consecutive flips.

What is the chance that we have two heads?

- $\boldsymbol{P}(HHTT) = (2/3)^2(1/3)^2 = 4/81$, $\boldsymbol{P}(A_{twoheads}) = 6 \times 4/81 = 24/81$

$\Omega =$

HHHH	HHHT	HHTH	HHTT
HTHH	HTHT	HTTH	HTTT
THHH	THHT	THTH	THTT
TTHH	TTHT	TTTH	TTTT

Note: This "manual counting" approach doesn't scale to larger k!

Example: Sixteen Biased Coins

Lecture 16, CS70 Summer 2025

Sixteen Biased Coins

Suppose we have a coin where $\boldsymbol{P}(H) = 1/3$ and $\boldsymbol{P}(T) = 2/3$.

Suppose that our experiment consists of 16 consecutive flips.

- What is the chance that we have 6 heads?

We can't just enumerate events individually! Way too many to list.

Back to Four Biased Coins

Suppose we have a coin where $\boldsymbol{P}(H) = 1/3$ and $\boldsymbol{P}(T) = 2/3$.

Suppose that our experiment consists of four consecutive flips.

What is the chance that we have two heads?

- Probability of any two heads event: $\boldsymbol{P}(HHTT) = (2/3)^2(1/3)^2 = 4/81$
- Given 2 heads out of 4, there are $\binom{4}{2} = 6$ ways to choose the position of the heads. Thus overall probability is $6 \times 4/81 = 24/81$

$\Omega =$

HHHH	HHHT	HHTH	HHTT
HTHH	HTHT	HTTH	HTTT
THHH	THHT	THTH	THTT
TTHH	TTHT	TTTH	TTTT

We just did this process of counting the 6 sample points with 2 heads, but without drawing it out visually.

Sixteen Biased Coins

One common model in engineering applications is the biased coin. Can be used to model, e.g. chance that a system will fail.

Suppose we have a coin where $\boldsymbol{P}(H) = 1/3$ and $\boldsymbol{P}(T) = 2/3$.

Suppose that our experiment consists of 16 consecutive flips.

- What is the chance that we have exactly 6 heads?

Sixteen Biased Coins

One common model in engineering applications is the biased coin. Can be used to model, e.g. chance that a system will fail.

Suppose we have a coin where $\boldsymbol{P}(H) = 1/3$ and $\boldsymbol{P}(T) = 2/3$.

Suppose that our experiment consists of 16 consecutive flips.

- What is the chance that we have exactly 6 heads?
  - Chance of any given outcome with 6 heads: $(1/3)^6(2/3)^{10}$
  - Number of ways to choose position of 6 out of 16 heads? $\binom{16}{6}$

Sixteen Biased Coins

One common model in engineering applications is the biased coin. Can be used to model, e.g. chance that a system will fail.

Suppose we have a coin where $\boldsymbol{P}(H) = 1/3$ and $\boldsymbol{P}(T) = 2/3$.

Suppose that our experiment consists of 16 consecutive flips.

- What is the chance that we have exactly 6 heads?
  - Chance of any given outcome with 6 heads: $(1/3)^6(2/3)^{10}$
  - Number of ways to choose position of 6 out of 16 heads: $\binom{16}{6}$
  - Overall chance is therefore $(1/3)^6(2/3)^{10}\binom{16}{6}$

Sixteen Biased Coins

One common model in engineering applications is the biased coin. Can be used to model, e.g. chance that a system will fail.

Suppose we have a coin where $\boldsymbol{P}(H) = 1/3$ and $\boldsymbol{P}(T) = 2/3$.

Suppose that our experiment consists of 16 consecutive flips.

- What is the chance that we have exactly 6 heads?
  - Chance of any given outcome with 6 heads: $(1/3)^6(2/3)^{10}$
  - Number of ways to choose position of 6 out of 16 heads: $\binom{16}{6}$
  - Overall chance is therefore $(1/3)^6(2/3)^{10}\binom{16}{6}$

$(1/3)^6(2/3)^{10}\binom{16}{10}$ is also correct!

$$\binom{16}{10} = \binom{16}{6} = 8008$$

Poker Hands

Lecture 16, CS70 Summer 2025

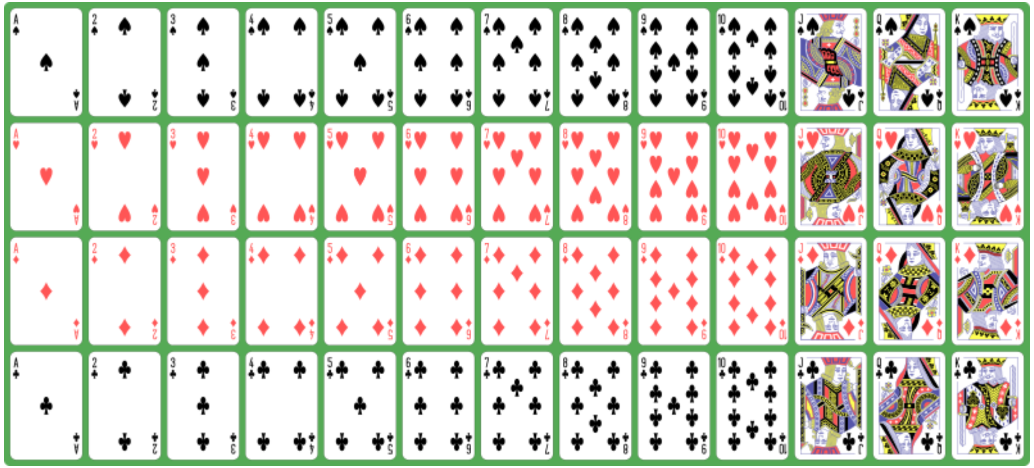
Playing Cards

In a standard deck of American playing cards, there are 52 cards.

- Each card has a "rank" and a "suit".
- Ranks: {A, 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K}
- Suits: {♠, ♥, ♦, ♣}

When shuffled, there are 52! (roughly 8×10^{67}) permutations.

- No two shuffled decks have ever been the same!



Poker Hands as a Random Experiment

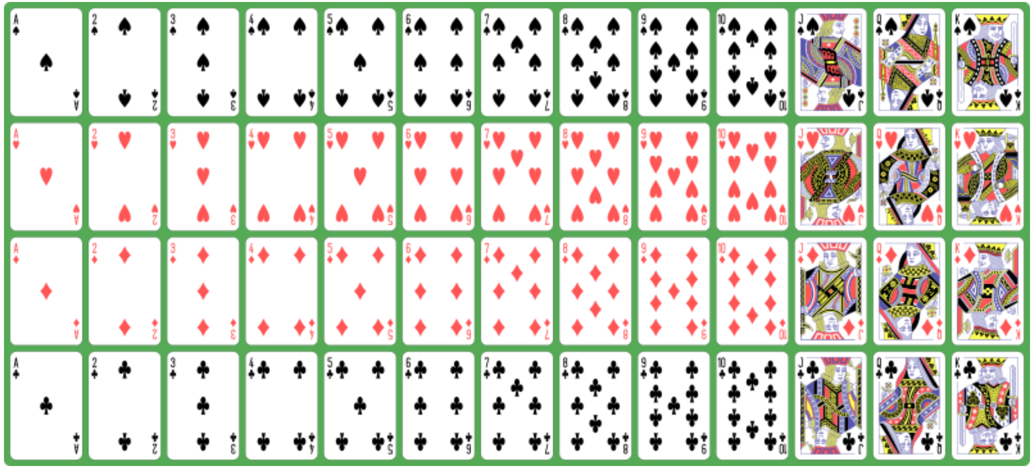
In the game of poker, each player is given a "hand" consisting of 5 cards from the set of 52 cards shown below.

- Each card has a "rank" and a "suit".
- Ranks: {A, 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K}
- Suits: {♠, ♥, ♦, ♣}

Example Hand: {A♥, 2♣, 5♣, K♣, Q♦}

Questions:

- What is n (the cardinality of the set we're drawing from)? What is k? What is |Ω|?
- Is this with/without replacement?
- Is the probability space uniform?



Poker Hands as a Random Experiment

In the game of poker, each player is given a "hand" consisting of 5 cards from the set of 52 cards shown below.

- Each card has a "rank" and a "suit".
- Ranks: {A, 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K}
- Suits: {♠, ♥, ♦, ♣}

Example Hand: {A♥, 2♣, 5♣, K♣, Q♦}

Questions:

- What is n (the cardinality of the set we're drawing from)? What is k? What is $|\Omega|$?
- Is this with/without replacement?
- Is the probability space uniform?

Answers:

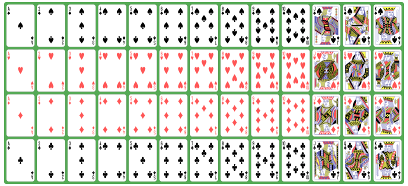
- n is 52. k is 5. Drawing without replacement.
- Drawing 5 items without replacement from a set of cardinality 52: $52 \times 51 \times 50 \times 49 \times 48$ or 311,875,200 possible hands.
- The space is uniform.

Poker Hands as a Random Experiment

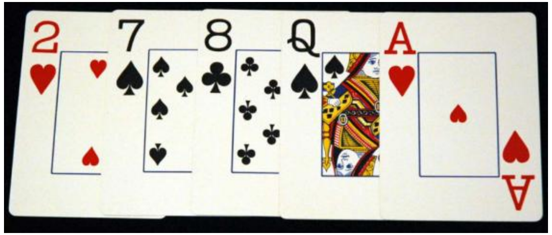
In the game of poker, each player is given a “hand” consisting of 5 cards from the set of 52 cards shown below.

- Each card has a “rank” and a “suit”.
- Ranks: {A, 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K}
- Suits: {♠, ♥, ♦, ♣}

Example Hand: {A♥, 2♣, 5♣, K♣, Q♦}



Question: Which hand is more likely?

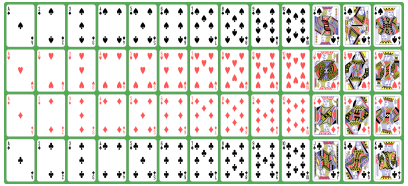


Poker Hands as a Random Experiment

In the game of poker, each player is given a “hand” consisting of 5 cards from the set of 52 cards shown below.

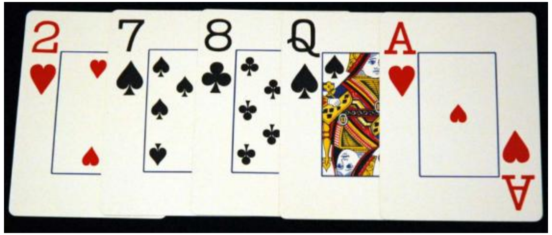
- Each card has a “rank” and a “suit”.
- Ranks: {A, 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K}
- Suits: {♠, ♥, ♦, ♣}

Example Hand: {A♥, 2♣, 5♣, K♣, Q♦}



Question: Which hand is more likely?

- They are equally likely. It’s a uniform probability space.



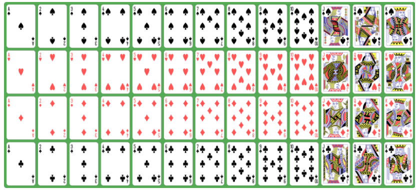
Poker Hands: Probability of a "Flush"

In the game of poker, each player is given a "hand" consisting of 5 cards from the set of 52 cards shown below.

Example Flushes: {A♥, 2♥, 5♥, K♥, Q♥}   {5♥, A♥, 2♥, Q♥, K♥}

What is the chance that your hand is a "flush", i.e. has all the same suit?

- Since all hands are equally likely, we just need to count the number of different flushes.
- Number of possible hands (taking into account order): $52 \times 51 \times 50 \times 49 \times 48$ or 311,875,200.
- Number of different ♥ flushes:



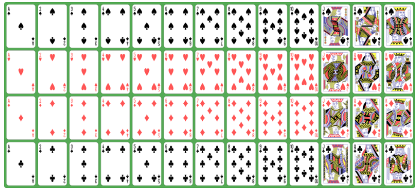
Poker Hands: Probability of a “Flush”

In the game of poker, each player is given a “hand” consisting of 5 cards from the set of 52 cards shown below.

Example Flushes: {A❤, 2❤, 5❤, K❤, Q❤} {5❤, A❤, 2❤, Q❤, K❤}

What is the chance that your hand is a “flush”, i.e. has all the same suit?

- Since all hands are equally likely, we just need to count the number of different flushes.
- Number of possible hands (taking into account order): $52 \times 51 \times 50 \times 49 \times 48$ or 311,875,200.
- Number of different ❤ flushes: There are 13 hearts, and we are drawing 5 without replacement: there are $13 \times 12 \times 11 \times 10 \times 9$ or 154,440 heart flushes.
- Four different suits, so total number of flushes = 617,760
- Thus, chance of a flush is 617,760/311,875,200 ≈ 0.2%



Alternate Framing: Ignoring Order

With hands playing cards it's also perfectly reasonable to define the sample space so that we ignore order.

$$\{1♥, 2♥, 3♥, 4♥, 5♥\} \equiv \{5♥, 3♥, 1♥, 2♥, 4♥\}$$

Questions:

- What is n (the cardinality of the set we're drawing from)? What is k? What is $|\Omega|$?
- Is this with/without replacement?
- Is the probability space uniform?

Answers:

- n is 52. k is 5. Drawing without replacement.
- Drawing 5 items without replacement from a set of cardinality 52, ignoring order: ??
- The space is ??

Alternate Framing: Ignoring Order

With hands playing cards it's also perfectly reasonable to define the sample space so that we ignore order.

$$\{1♥, 2♥, 3♥, 4♥, 5♥\} \equiv \{5♥, 3♥, 1♥, 2♥, 4♥\}$$

Questions:

- What is n (the cardinality of the set we're drawing from)? What is k? What is $|\Omega|$?
- Is this with/without replacement?
- Is the probability space uniform?

Answers:

- n is 52. k is 5. Drawing without replacement.
- Drawing 5 items without replacement from a set of cardinality 52, ignoring order: $\binom{52}{5}$ or 2,598,960 possible hands.
- The space is .... ?

Alternate Framing: Ignoring Order

With hands playing cards it's also perfectly reasonable to define the sample space so that we ignore order.

$$\{1♥, 2♥, 3♥, 4♥, 5♥\} \equiv \{5♥, 3♥, 1♥, 2♥, 4♥\}$$

Why is this probability space uniform, but the probability space of two four-sided dice ignoring order was not uniform?

Questions:

- What is n (the cardinality of the set we're drawing from)? What is k? What is $|\Omega|$?
- Is this with/without replacement?
- Is the probability space uniform?

Answers:

- n is 52. k is 5. Drawing without replacement.
- Drawing 5 items without replacement from a set of cardinality 52, ignoring order: $\binom{52}{5}$ or 2,598,960 possible hands.
- The space is uniform.

Alternate Framing: Ignoring Order

With hands playing cards it's also perfectly reasonable to define the sample space so that we ignore order.

$$\{1♥, 2♥, 3♥, 4♥, 5♥\} \equiv \{5♥, 3♥, 1♥, 2♥, 4♥\}$$

Why is this probability space uniform, but the probability space of two four-sided dice ignoring order was not uniform? We're drawing without replacement, so duplicates are impossible! No 12 vs. 21 situation.

Questions:

- What is n (the cardinality of the set we're drawing from)? What is k? What is $|\Omega|$?
- Is this with/without replacement?
- Is the probability space uniform?

Answers:

- n is 52. k is 5. Drawing without replacement.
- Drawing 5 items without replacement from a set of cardinality 52, ignoring order: $\binom{52}{5}$ or 2,598,960 possible hands.
- The space is uniform.

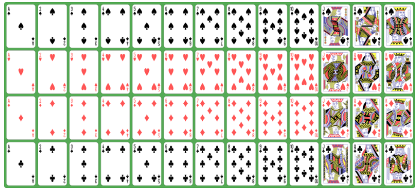
Poker Hands: Probability of a "Flush"

We can analyze the probability of a flush in this sample space as well.

Example Flush: {A❤, 2❤, 5❤, K❤, Q❤}

What is the chance that your hand is a "flush", i.e. has all the same suit?

- Since probability space is uniform, we just need to count the number of different flushes.
- Number of different ❤ flushes:



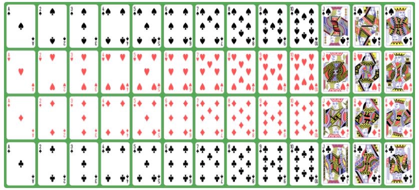
Poker Hands: Probability of a "Flush"

We can analyze the probability of a flush in this sample space as well.

Example Flush: {A♥, 2♥, 5♥, K♥, Q♥}

What is the chance that your hand is a "flush", i.e. has all the same suit?

- Since probability space is uniform, we just need to count the number of different flushes.
- Number of different ♥ flushes: There are 13 hearts, so there are $\binom{13}{5} = 1287$ heart flushes.
- Four different suits, so total number of flushes = 5,148.
- Total number of hands is 2,598,960.
- Thus, chance of a flush is 5,148/2,598,960 $\approx$ 0.2%



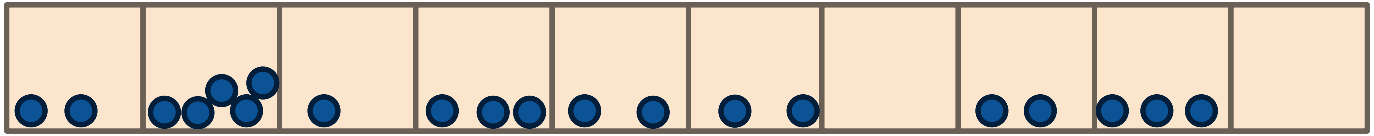
Balls and Bins

Lecture 16, CS70 Summer 2025

Example: Balls and Bins

Suppose we throw 20 balls independently into 10 bins. Assume each ball is equally likely to land in any of the ten bins.

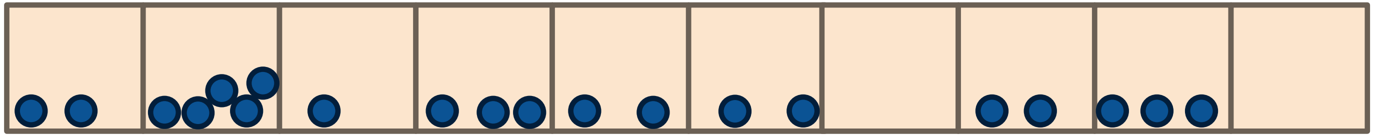
Let's do a quick demo.



20 balls thrown!

Example: Balls and Bins

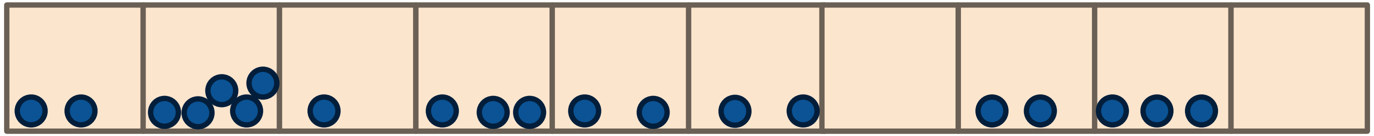
Suppose we throw 20 balls independently into 10 bins. Assume each ball is equally likely to land in any of the ten bins.



What is k, n, is this with/without replacement, and what is $|\Omega|$? Uniform p.s.?

Example: Balls and Bins

Suppose we throw 20 balls independently into 10 bins. Assume each ball is equally likely to land in any of the ten bins.

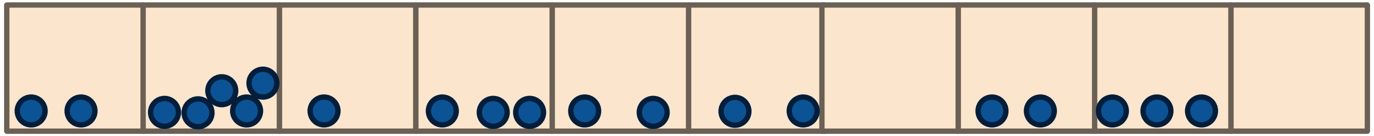


What is k, n, is this with/without replacement, and what is $|\Omega|$? Uniform p.s.?

- k is 20 (samples drawn).
- n is 10 (set of possible outcomes are {1, 2, 3, 4, 5, 6, 7, 8, 9, 10}.
- Drawings are with replacement (can get same number multiple times).
- Sample space has 10^{20} elements
- Probability space is uniform, i.e., specific distribution of balls above (ignoring who is on top of whom, etc.) occurs 1 out of 10^{20} times.

Example: Balls and Bins

Suppose we throw 20 balls independently into 10 bins. Assume each ball is equally likely to land in any of the ten bins.

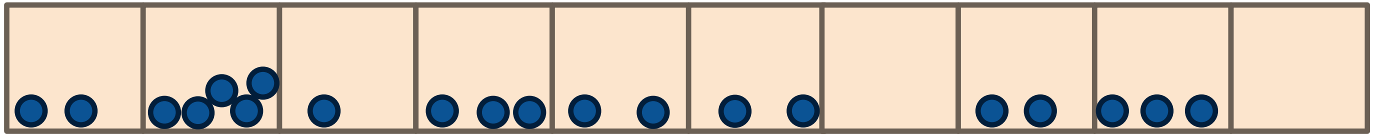


What is k, n, is this with/without replacement, and what is $|\Omega|$? Uniform p.s.?

- k is 20 (samples drawn).
- n is 10 (set of possible outcomes are {1, 2, 3, 4, 5, 6, 7, 8, 9, 10}.
- Drawings are with replacement (can get same number multiple times).
- Sample space has 10^{20} elements – does this consider order?
- Probability space is uniform, i.e., specific distribution of balls above (ignoring who is on top of whom, etc.) occurs 1 out of 10^{20} times.

Example: Balls and Bins

Suppose we throw 20 balls independently into 10 bins. Assume each ball is equally likely to land in any of the ten bins.

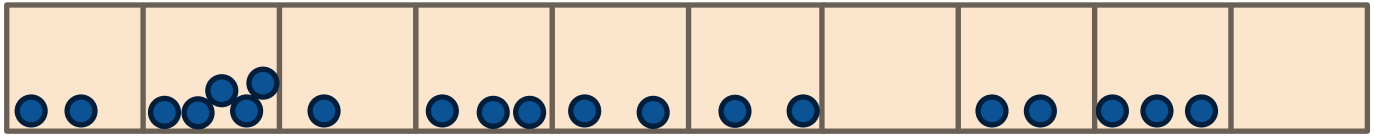


What is k, n, is this with/without replacement, and what is $|\Omega|$? Uniform p.s.?

- k is 20 (samples drawn).
- n is 10 (set of possible outcomes are {1, 2, 3, 4, 5, 6, 7, 8, 9, 10}.
- Drawings are with replacement (can get same number multiple times).
- Sample space has 10^{20} elements. This considers order, despite balls being indistinguishable.
- Probability space is uniform, i.e., specific distribution of balls above (ignoring who is on top of whom, etc.) occurs 1 out of 10^{20} times.

Example: Balls and Bins

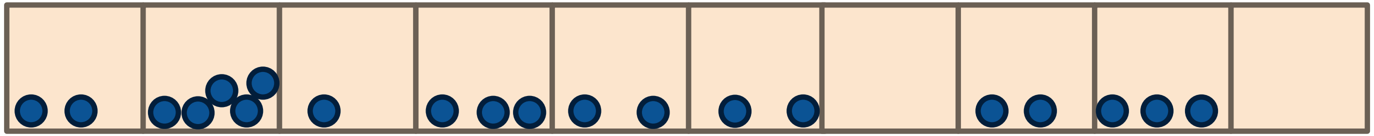
Suppose we throw 20 balls independently into 10 bins. Assume each ball is equally likely to land in any of the ten bins.



Why do you think I chose to consider order in defining my sample space?

Example: Balls and Bins

Suppose we throw 20 balls independently into 10 bins. Assume each ball is equally likely to land in any of the ten bins.



Why do you think I chose to consider order in defining my sample space?

- So that the probability space is uniform!
- Otherwise, some sample points would be more likely than others.
  - Example: Only one configuration where all balls go in bin 1 (unlikely). Many configurations where 2 balls go in each bin (thus more likely).

Example: Balls and Bins

Suppose we throw 20 balls independently into 10 bins. Assume each ball is equally likely to land in any of the ten bins.

- k=20, n=10, with replacement, $|\Omega| = 10^{20}$

What is $\boldsymbol{P}(A_{bin1empty})$, i.e. the probability bin 1 has no ball?

- Since all outcomes are equally likely, just need to count number of outcomes where no ball lands in bin 1.
- Number of ways 20 balls can land in 9 bins: 9^{20}
  - Can think of this as "Assume that the first bin is empty, how many different configurations are there?"
- So $\boldsymbol{P}\left(A_{bin1empty}\right) = 9^{20}/10^{20} \approx 12\%$

Example: Balls and Bins

Suppose we throw 20 balls independently into 10 bins. Assume each ball is equally likely to land in any of the ten bins.

- k=20, n=10, with replacement, $|\Omega| = 10^{20}$

What is $\boldsymbol{P}(A_{bin1empty})$, i.e. the probability bin 1 has no ball?

- $9^{20}/10^{20} \approx 12\%$

What is $\boldsymbol{P}(A_{bin1nonempty})$, i.e. the probability at least one ball lands in bin 1?

- This is just $1 - \boldsymbol{P}\left(A_{bin1empty}\right) = 1 - 9^{20}/10^{20} \approx 88\%$

In general, for an event A, $P(A) = 1 - P(\bar{A})$

Example: Balls and Bins

Suppose we throw m balls independently into n bins. Assume each ball is equally likely to land in any of the n bins.

- k=m, n=n, with replacement, $|\Omega| = n^m$

What is $\boldsymbol{P}(A_{bin1empty})$, i.e. the probability bin 1 has no ball?

- For k = 20, n = 10, we had $\boldsymbol{P}\left(A_{bin1empty}\right) = 9^{20}/10^{20}$
- For k = m, n = n, we have $\boldsymbol{P}\left(A_{bin1empty}\right) = (n-1)^m/n^m$

So $\boldsymbol{P}\left(A_{bin1empty}\right) = \frac{(n-1)^m}{n^m} = \left(\frac{n-1}{n}\right)^m = \left(1 - \frac{1}{n}\right)^m$

Number of ways m balls can land in n-1 bins: $(n-1)^m$

Buckets and Balls and the World

Suppose we throw m balls independently into n bins. Assume each ball is equally likely to land in any of the n bins.

This problem is a common model in computer science, e.g., modeling jobs being sent to one of many servers. Suppose m jobs are sent to one of n servers randomly.

- $P(A_{server1notused}) = \left(\frac{n-1}{n}\right)^m = \left(1 - \frac{1}{n}\right)^m$

Or m items are stored in one of n hash table buckets:

- $P(A_{bucket1notused}) = \left(\frac{n-1}{n}\right)^m = \left(1 - \frac{1}{n}\right)^m$

Example: Buckets and Balls and Dice and Coins

Suppose we throw m balls independently into n bins. Assume each ball is equally likely to land in any of the n bins.

This is also a more general framing of some previous problems, e.g.

- Rolling a 4-sided die twice: $m = ?$, $n = ?$

Example: Buckets and Balls and Dice and Coins

Suppose we throw m balls independently into n bins. Assume each ball is equally likely to land in any of the n bins.

This is also a more general framing of some previous problems, e.g.

- Rolling a 4-sided die twice: $m = 2$, $n = 4$.
- Flipping a coin 16 times, $m = ?$, $n = ?$.

Example: Buckets and Balls and Dice and Coins

Suppose we throw m balls independently into n bins. Assume each ball is equally likely to land in any of the n bins.

This is also a more general framing of some previous problems, e.g.

- Rolling a 4-sided die twice: $m = 2$, $n = 4$.
- Flipping a coin 16 times, $m = 16$, $n = 2$.
- Drawing a hand of 5 cards from a deck of 52 cards? $m = ?$, $n = ?$.

Example: Buckets and Balls and Dice and Coins

Suppose we throw m balls independently into n bins. Assume each ball is equally likely to land in any of the n bins.

This is also a more general framing of some previous problems, e.g.

- Rolling a 4-sided die twice: $m = 2$, $n = 4$.
- Flipping a coin 16 times, $m = 16$, $n = 2$.
- Drawing a hand of 5 cards from a deck of 52 cards? Doesn't apply!
    - Drawing a hand is without replacement! Doesn't match independent ball throwing.

The Birthday Paradox (n=50 case)

Lecture 16, CS70 Summer 2025

Birthdays

Suppose we have 50 people in a room. What is the chance that two share a birthday? (For simplicity, ignore leap year birthdays, i.e., assume each year has 365 days).

What's your gut feeling?

- Less than 20%?
- Decent chance, say 50%ish?
- Very likely, say >90%?

Birthdays

Suppose we have 50 people in a room who all have a birthday between day 1 and day 365.

In terms of probability spaces:

- What is the cardinality of the set S that we're drawing from?
- How many samples are we drawing?
- Are we drawing with or without replacement?
- What is the cardinality of the sample space Ω?

Birthdays

Suppose we have 50 people in a room who all have a birthday between day 1 and day 365.

What is the probability space for the possible birthdays?

- What is the cardinality of the set S that we're drawing from? 365
- How many samples are we drawing? 50
- Are we drawing with or without replacement? With
- What is the cardinality of the sample space Ω? $|\Omega| = 365^{50}$

The probability space is again uniform.

Birthdays

Suppose we have 50 people in a room who all have a birthday between day 1 and day 365.

- Cardinality of the sample space Ω? $|\Omega| = 365^{50}$

Let A be the event where two (or more) of the birthdays are the same. Then:

$$P(A) = \frac{|A|}{|\Omega|}$$

Examples from A: [1, 1, ...], [5, 5, ...5, ...], [66, 12, 51, 66, ...]

Problem: How the heck do we figure out $|A|$?

Birthdays

Suppose we have 50 people in a room who all have a birthday between day 1 and day 365.

- Cardinality of the sample space Ω? $|\Omega| = 365^{50}$

Let A be the event where two (or more) of the birthdays are the same. Hard to compute $|A|$.

As often the case, $|\bar{A}|$ is easier to compute.

- Example from $|\bar{A}|$: [55, 1, 67, 2, 99, ..., 33, 41], with no repeats.

How many samples are there like this with no repeats?

- Let's go back to the random experiments framework.

Birthdays with no Repeats

Suppose we have 50 people in a room who all have a birthday between day 1 and day 365 AND THERE ARE NO REPEAT BIRTHDAYS.

Goal: Figure out the size of this sample space.

- What is the cardinality of the set S that we're drawing from?
- How many samples are we drawing?
- Are we drawing with or without replacement?
- What is the cardinality of this sample space?

Birthdays with no Repeats

Suppose we have 50 people in a room who all have a birthday between day 1 and day 365 AND THERE ARE NO REPEAT BIRTHDAYS.

Goal: Figure out the size of this sample space.

- What is the cardinality of the set S that we're drawing from? 365
- How many samples are we drawing? 50
- Are we drawing with or without replacement? Without
- What is the cardinality of this sample space?

Birthdays with no Repeats

Suppose we have 50 people in a room who all have a birthday between day 1 and day 365 AND THERE ARE NO REPEAT BIRTHDAYS.

In terms of probability spaces:

- What is the cardinality of the set S that we're drawing from? 365
- How many samples are we drawing? 50
- Are we drawing with or without replacement? Without
- What is the cardinality of this sample space? $365 \times 364 \times \cdots \times 316$

Note that this sample space is just $\bar{A}$, i.e. $|\bar{A}| = 365 \times 364 \times \cdots \times 316$

Birthdays

Suppose we have 50 people in a room who all have a birthday between day 1 and day 365. What is the chance that two of those 50 people share the same birthday?

Let A be the event where two (or more) of the birthdays are the same. Then:

$$P(A) = \frac{|A|}{|\Omega|} = 1 - P(\bar{A}) = 1 - \frac{|\bar{A}|}{|\Omega|}$$

For $k = 50$, $|\bar{A}| = 365 \times 364 \times \cdots \times 316$, and $|\Omega| = 365^{50}$, so

$$P(A) = 1 - \frac{365 \times 364 \times \cdots \times 316}{365^{50}}$$

Can use e.g. Wolfram alpha to compute.

Birthdays

Suppose we have 50 people in a room who all have a birthday between day 1 and day 365. What is the chance that two of those 50 people share the same birthday?

Let A be the event where two (or more) of the birthdays are the same. Then:

$$P(A) = \frac{|A|}{|\Omega|} = 1 - P(\bar{A}) = 1 - \frac{|\bar{A}|}{|\Omega|}$$

For $k = 50$, $|\bar{A}| = 365 \times 364 \times \cdots \times 316$, and $|\Omega| = 365^{50}$, so

$$P(A) = 1 - \frac{365 \times 364 \times \cdots \times 316}{365^{50}} \approx 1 - 0.03 \approx 0.97$$

The Birthday Paradox (general case)

Lecture 16, CS70 Summer 2025

Birthdays

Suppose we have 50 people in a room who all have a birthday between day 1 and day 365. What is the chance that two of those 50 people share the same birthday?

- Compute by counting number of birthday combinations with no repeats.

$$|\bar{A}| = 365 \times 364 \times \cdots \times 316$$

Suppose we want to generalize to the case where we have n people. What is $|\bar{A}|$ in that case?

$$P(A) = 1 - \frac{365 \times 364 \times \cdots \times 316}{365^{50}} \approx 1 - 0.03 \approx 0.97$$

Birthdays

Suppose we have 50 people in a room who all have a birthday between day 1 and day 365. What is the chance that two of those 50 people share the same birthday?

- Compute by counting number of birthday combinations with no repeats.

$$|\bar{A}| = 365 \times 364 \times \cdots \times 316$$

Suppose we want to generalize to the case where we have n people. What is $|\bar{A}|$ in that case?

$$|\bar{A}| = 365 \times 364 \times \cdots \times (365 - n + 1)$$

Birthdays

Suppose we have n people in a room who all have a birthday between day 1 and day 365. What is the chance that two of those n people share the same birthday?

$$P(A) = 1 - \frac{365 \times 364 \times \cdots \times (365 - n + 1)}{365^n}$$

Quick Aside: Computing This Function

Suppose we want to compute the results of the function below.

- Why might it be a bad idea to compute the numerator and denominator separately?
- How can we compute the result more efficiently?

$$P(A) = 1 - \frac{365 \times 364 \times \cdots \times (365 - n + 1)}{365^n}$$

Quick Aside: Computing This Function

Suppose we want to compute the results of the function below.

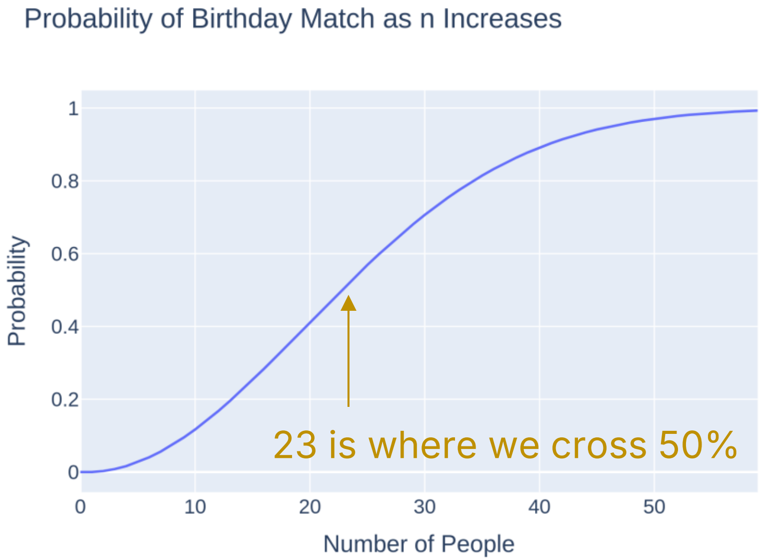
- Why might it be a bad idea to compute the numerator and denominator separately?
  - In many programming languages like Java, we'd get overflow issues.
  - In languages that support huge integers, performance will be slow. 365^{100} is an 852 bit number, much bigger than 64 bits for typical operations.
- How can we compute the result more efficiently?
  - One idea: Compute product of $\frac{365}{365} \times \frac{364}{365} \times \frac{363}{365} \times \cdots$
  - You can compute this approximately in constant time (maybe try to come up with a technique, happy to discuss at office hours!)

$$P(A) = 1 - \frac{365 \times 364 \times \cdots \times (365 - n + 1)}{365^n}$$

Birthday Paradox Table

For various n, the probability that at least two people share a birthday is given below.

n	$P(A)$	%
1	0	0%
2	0.0027	0.27%
3	0.008	0.8%
4	0.016	1.6%
10	0.117	11.7%
15	0.252	25.2%
20	0.411	41.1%
23	0.507	50.7%
50	0.97	97.0%
60	0.994	99.4%



The Monty Hall Problem: Simple Analysis

Lecture 16, CS70 Summer 2025

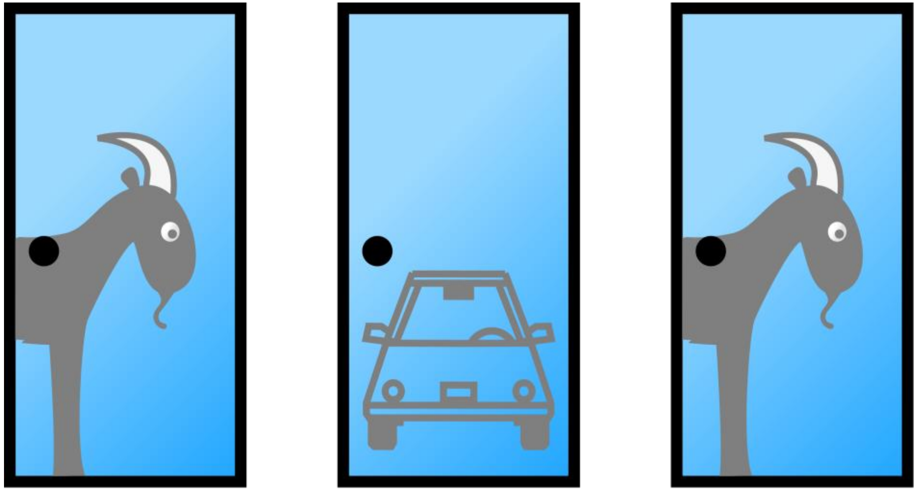
Probability Spaces and Event Probabilities

Let's consider a much trickier problem using the same framework as before.

The Monty Hall Problem

Let's start by imagining a game. In this game there are 3 doors with prizes behind them, and you want to win the car.

- One door has a car behind it.
- Two doors have a goat behind them.



The Monty Hall Problem

Let’s start by imagining a game. In this game there are 3 doors with prizes behind them, and you want to win the car.

- One door has a car behind it.
- Two doors have a goat behind them.

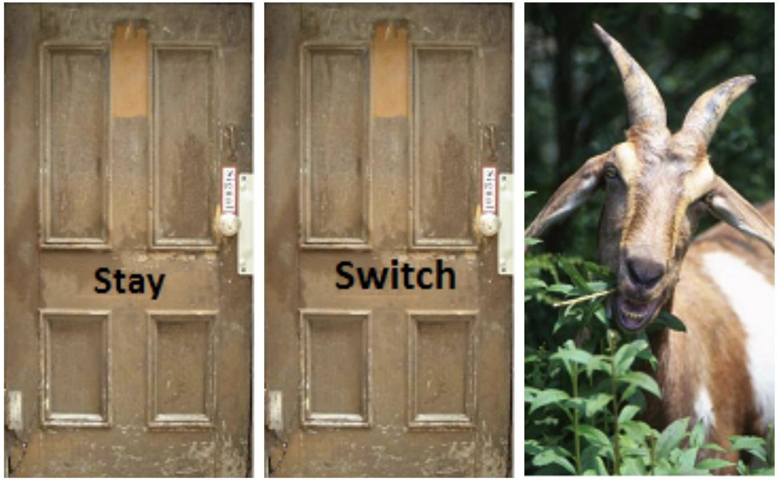
The rules are as follows:

- You pick one of 3 doors.
- The host reveals what is behind one of the other doors. The answer is ALWAYS a goat.
- You get a chance to switch doors after the host shows you the goat.

Let’s try to play a few rounds:
https://www.rossmanchance.com/applets/2021/montyhall/Monty.html

The Monty Hall Problem

Suppose we're playing, pick the left door, and we get to this state:



Should you switch doors? (Note: You are ALWAYS shown a goat)

Monty Hall Problem

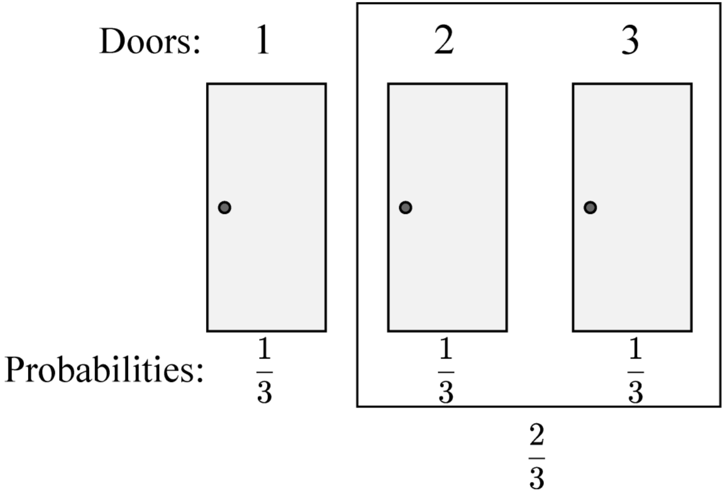
Strangely the answer is yes - your chance of winning is better if you switch!

Explanation (1/4)

The chance that you were correct with your first guess is 1/3.

Only two possibilities exist:

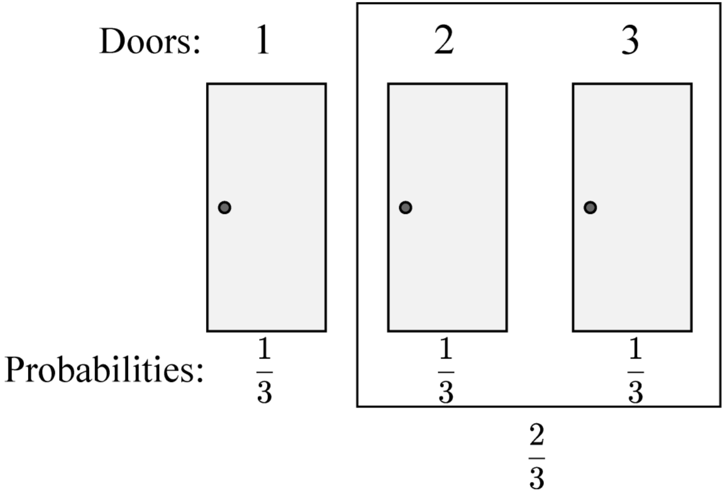
- You were right (with probability 1/3).
- You were wrong (with probability 2/3).

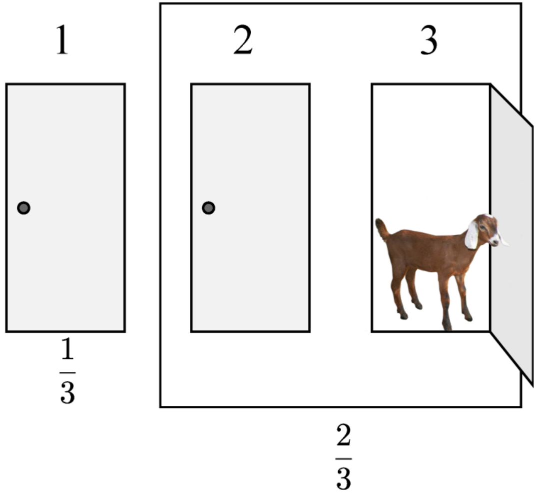


Explanation (2/4)

Suppose the host opens door 3, how do the probabilities change?

- Chance that your original choice is correct is unchanged. It's still 1/3.

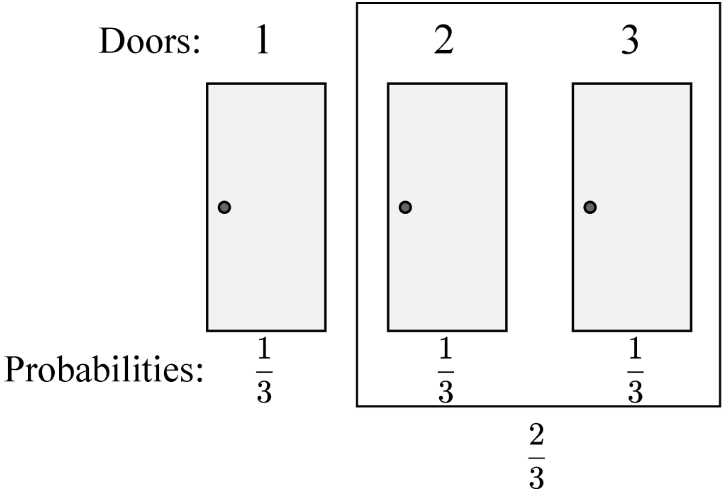


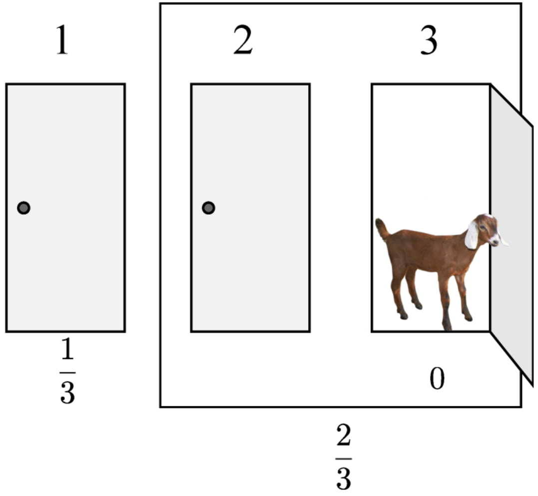


Explanation (3/4)

Suppose the host opens door 3, how do the probabilities change?

- Chance that your original choice is correct is unchanged. It's still 1/3.
- Chance that door 3 has the car is now zero.

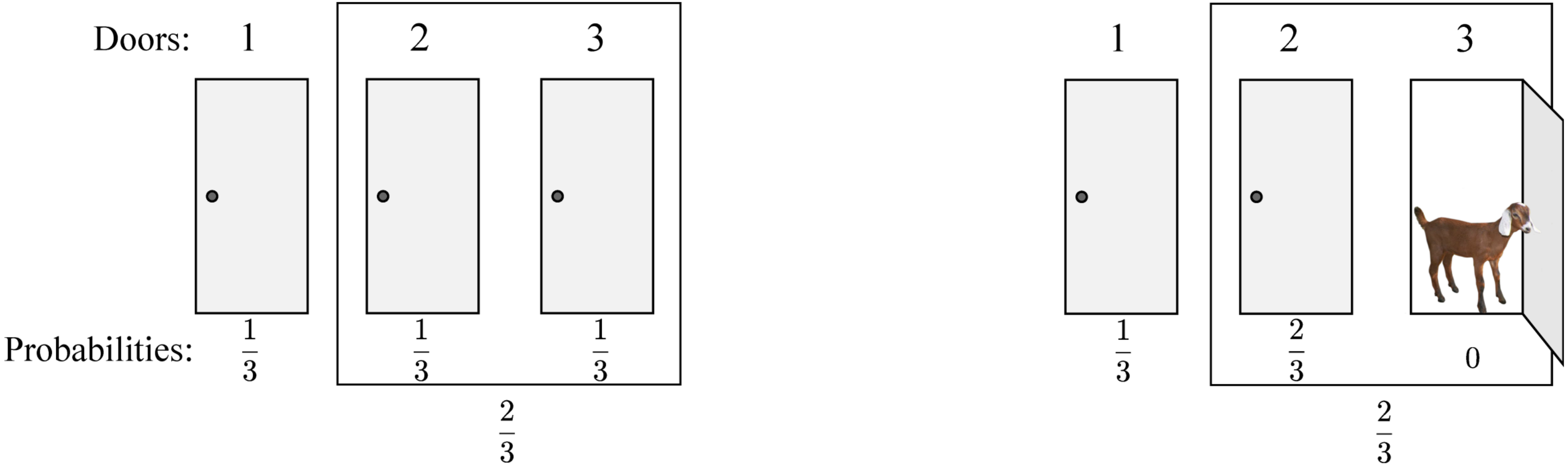




Explanation (4/4)

Suppose the host opens door 3, how do the probabilities change?

- Chance that your original choice is correct is unchanged. It's still 1/3.
- Chance that door 3 has the car is now zero.
- Chance that door 2 has the car is now 2/3 (since door 3 can't have car).



Explanation Summary

The chance that you were correct with your first guess is 1/3.

That is, only two possibilities exist before the door is open:

- You are right (with probability 1/3).
- You are wrong (with probability 2/3).

After opening a door, the chance you were right is STILL 1/3.

- So switching (to the other binary choice) must have a probability of 2/3.

Monty Hall Problem

When this problem was answered in Parade magazine by columnist Marilyn vos Savant in 1990, the answer drew huge amounts of angry mail including from folks holding PhDs.

Let's take a walk down memory lane:
https://web.archive.org/web/20130121183432/http://marilynvossavant.com/game-show-problem/

The Monty Hall Problem: Sample Space Analysis

Lecture 16, CS70 Spring 2025

Explanation in Terms of Sample Spaces

We can also think about this problem in terms of a sample space where we draw 3 samples from the set S = {1, 2, 3}:

- i: the door with the car prize (randomly chosen by TV show crew)
- j: the door that you select (randomly chosen by you)
- k: the door that gets opened (randomly chosen by host, with caveat that they will never select the door with the car or the door chosen by you)

Explanation in Terms of Sample Spaces

We can also think about this problem in terms of a sample space where we draw 3 samples from the set S = {1, 2, 3}:

- i: the door with the car prize (randomly chosen by TV show crew)
- j: the door that you select (randomly chosen by you)
- k: the door that gets opened (randomly chosen by host, with caveat that they will never select the door with the car or the door chosen by you)

Our usual questions:

- What is the cardinality of the set S that we're drawing from?
- How many samples are we drawing?
- Are we drawing with or without replacement?
- What is the cardinality of this sample space?

Explanation in Terms of Sample Spaces

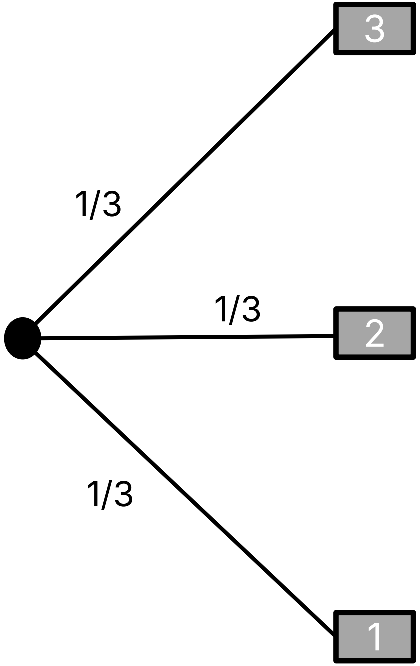
We can also think about this problem in terms of a sample space where we draw 3 samples from the set S = {1, 2, 3}:

- i: the door with the car prize (randomly chosen by TV show crew)
- j: the door that you select (randomly chosen by you)
- k: the door that gets opened (randomly chosen by host, with caveat that they will never select the door with the car or the door chosen by you)

Our usual questions:

- What is the cardinality of the set S that we're drawing from? 3
- How many samples are we drawing? 3
- Are we drawing with or without replacement? Neither. j is drawn with replacement, and k is drawn without replacement (after i and j).
- What is the cardinality of this sample space? See next slide.

The Sample Space for Monty Hall

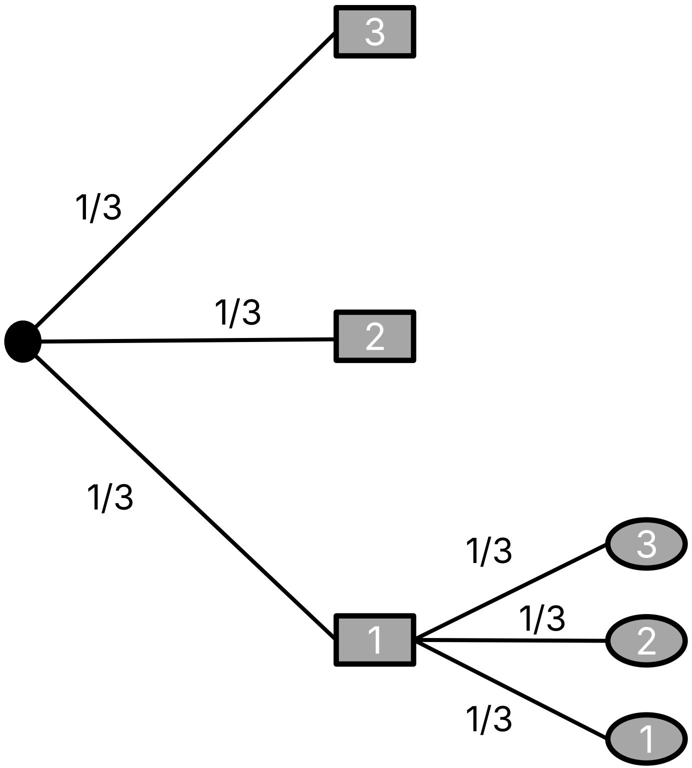


i: the door with the car prize

j: the door that you select first

k: the door that gets opened

The Sample Space for Monty Hall

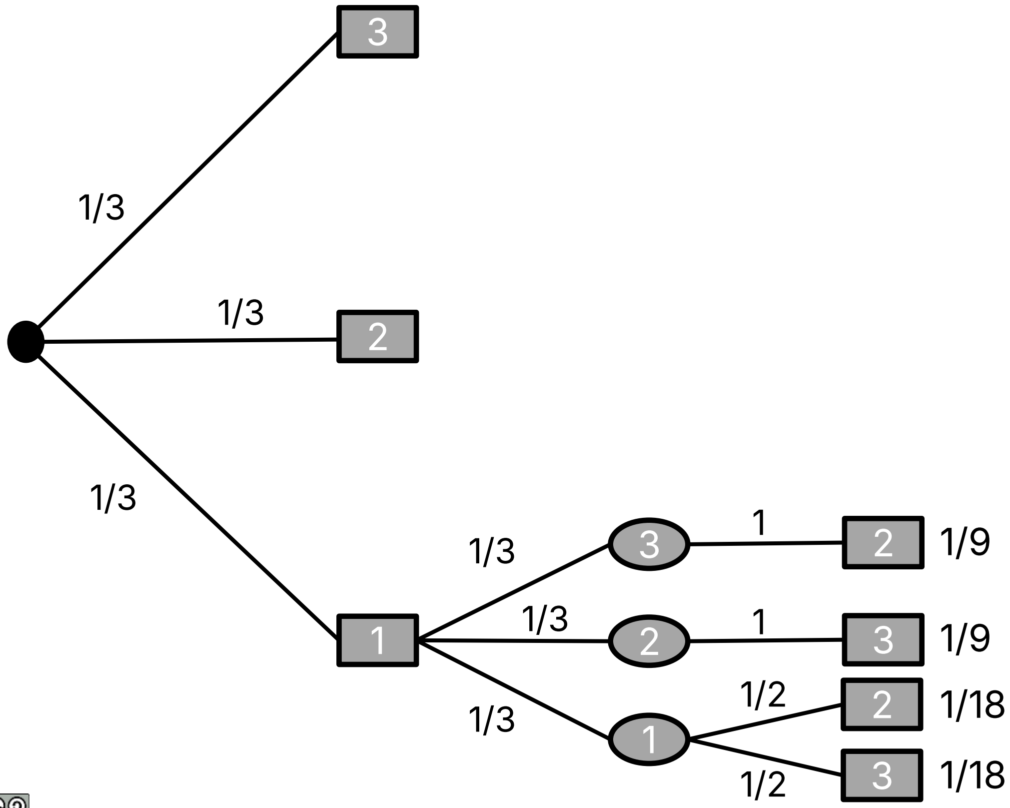


i: the door with the car prize

j: the door that you select first

k: the door that gets opened

The Sample Space for Monty Hall

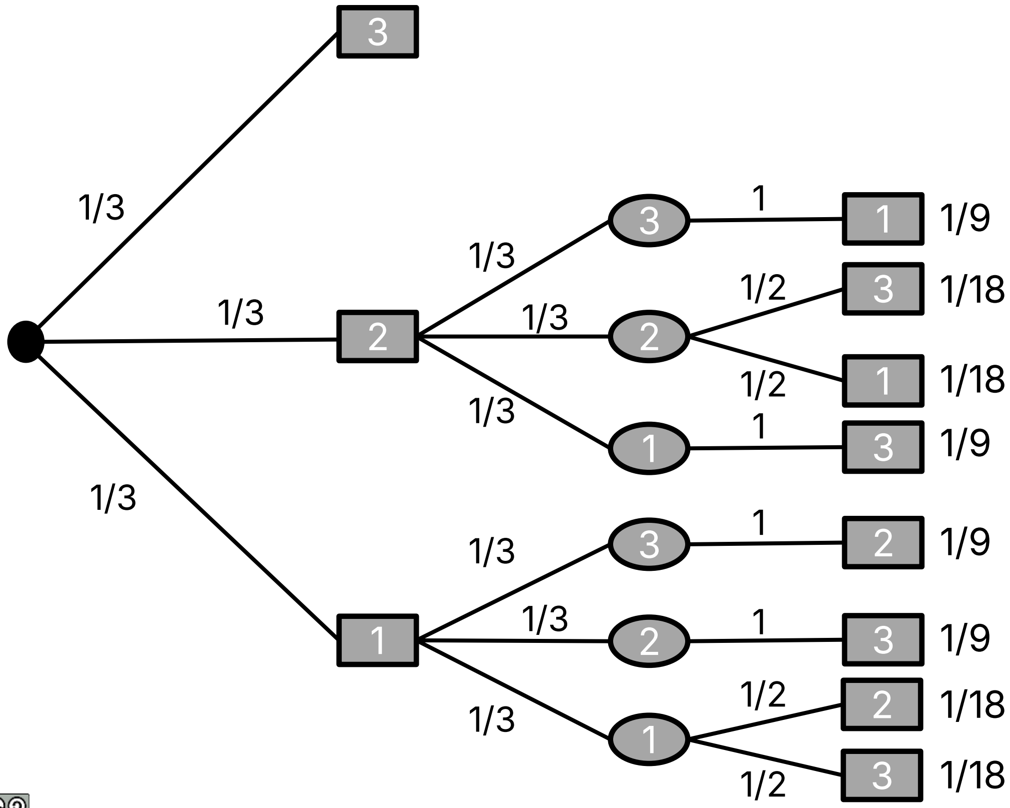


i: the door with the car prize

j: the door that you select first

k: the door that gets opened

The Sample Space for Monty Hall

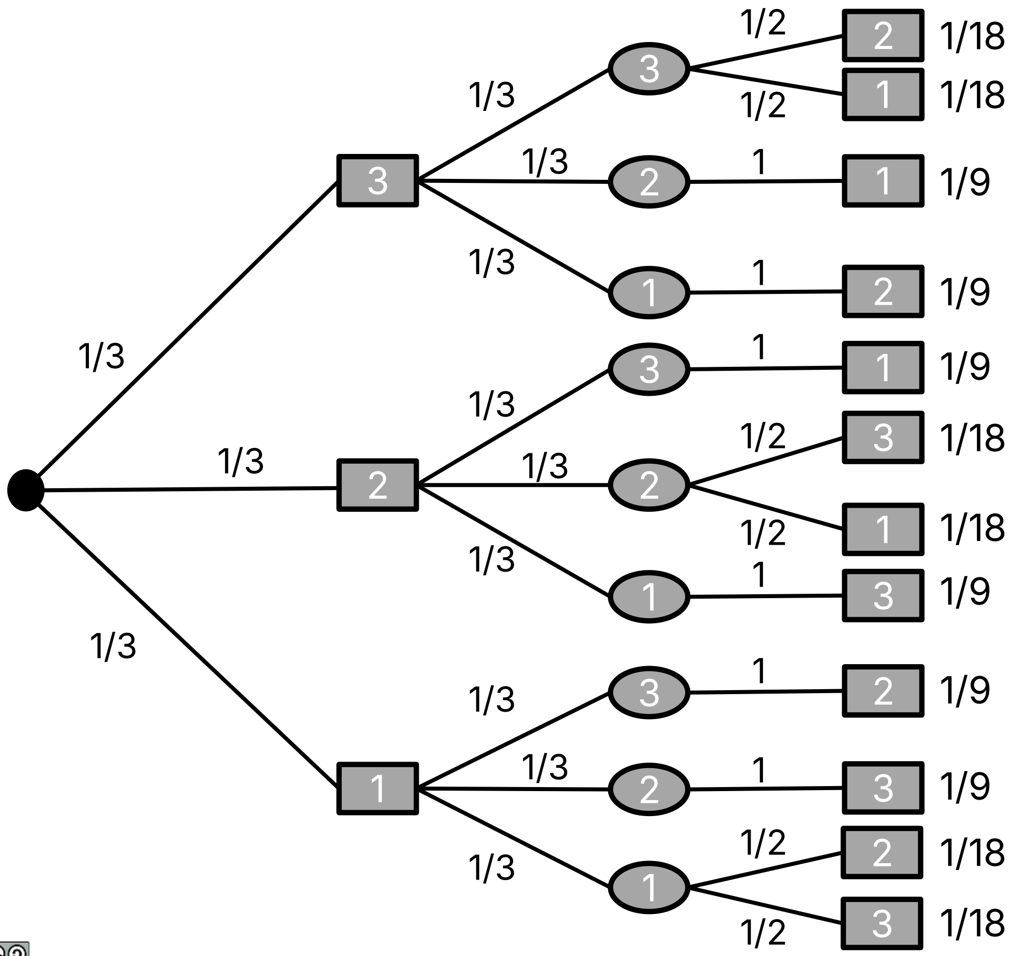


i: the door with the car prize

j: the door that you select first

k: the door that gets opened

The Sample Space for Monty Hall



i: the door with the car prize

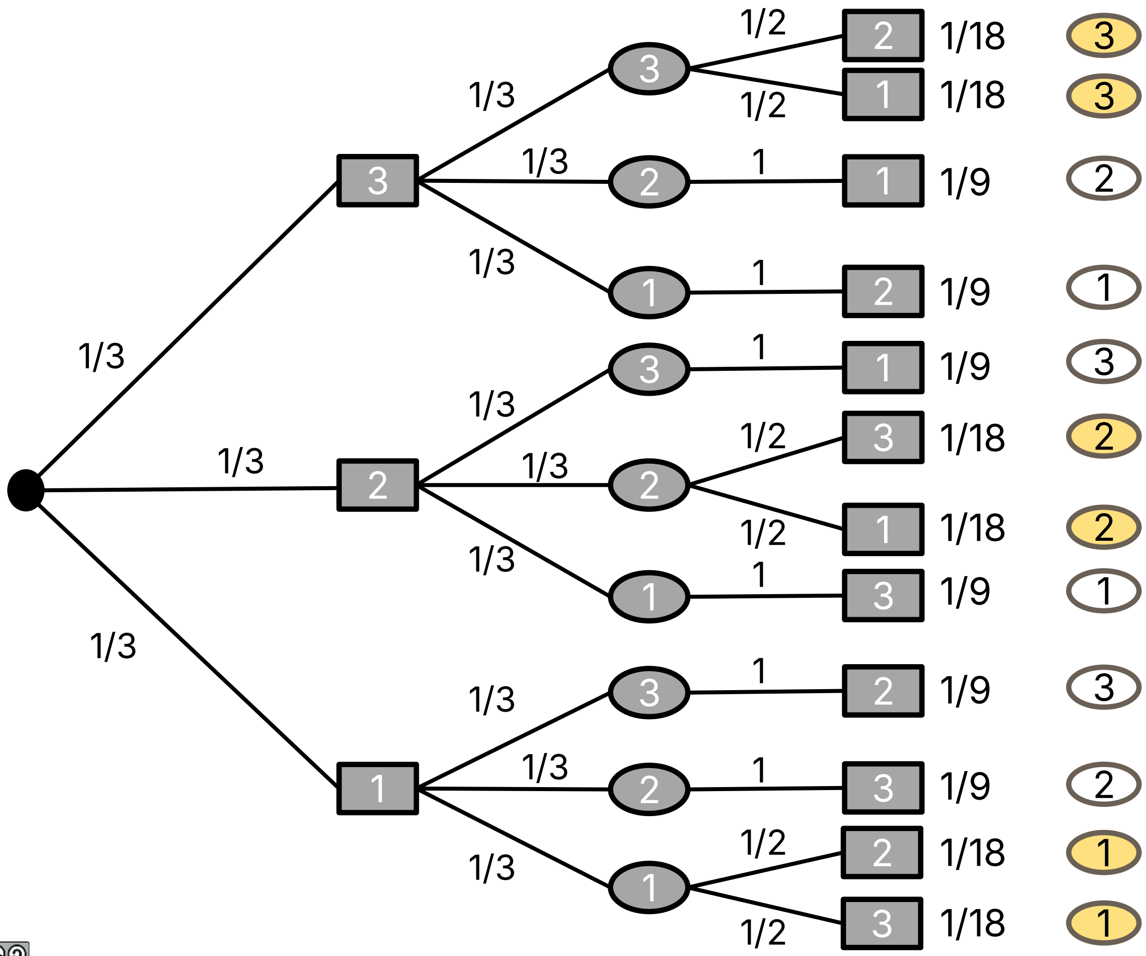
j: the door that you select first

k: the door that gets opened

So what? How is this useful for understanding what to do?

Have to consider the final and deterministic step – our strategy.

The Sample Space for Monty Hall + Not Switching Strategy



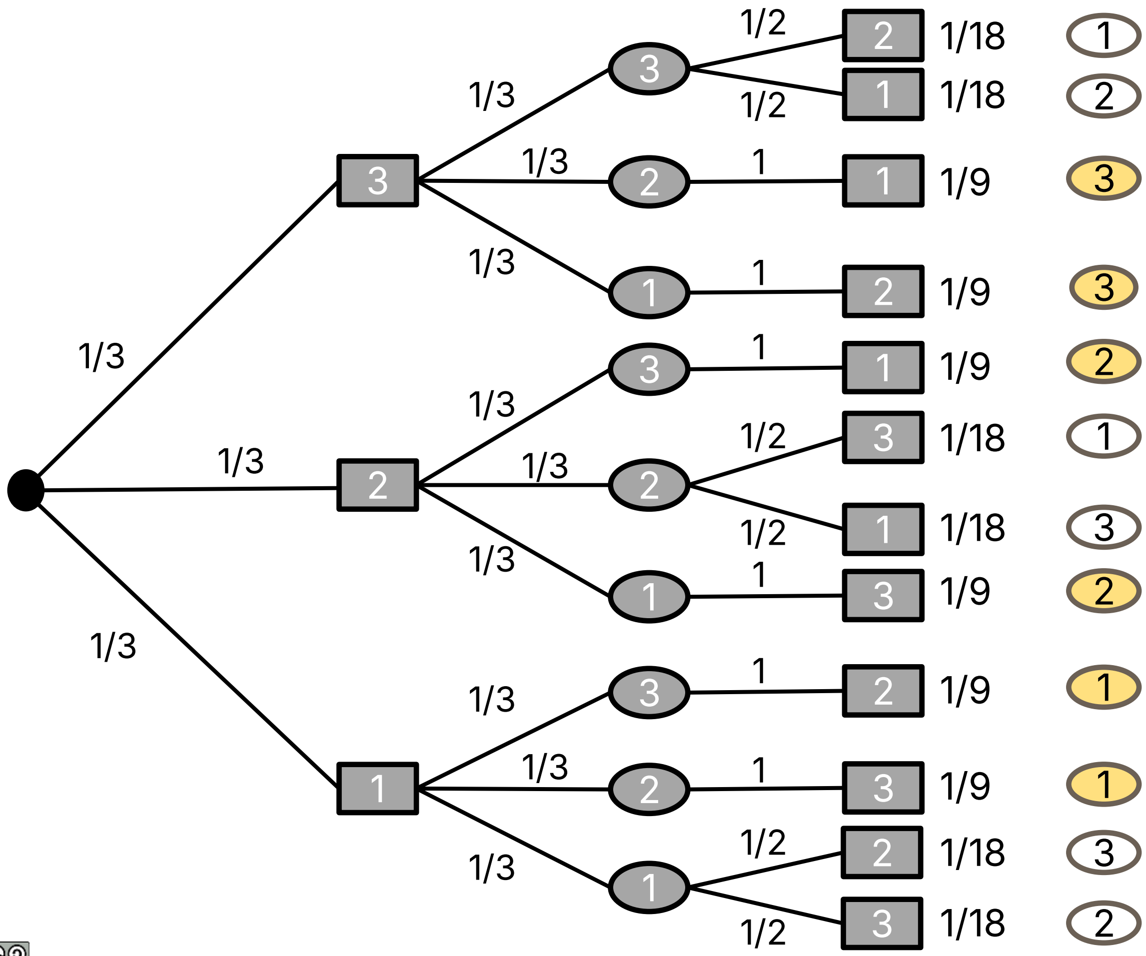
i: the door with the car prize

j: the door that you select first

k: the door that gets opened

Chance of winning if not switching: $6 \times \frac{1}{18} = 1/3$

The Sample Space for Monty Hall + Switching Strategy



i: the door with the car prize

j: the door that you select first

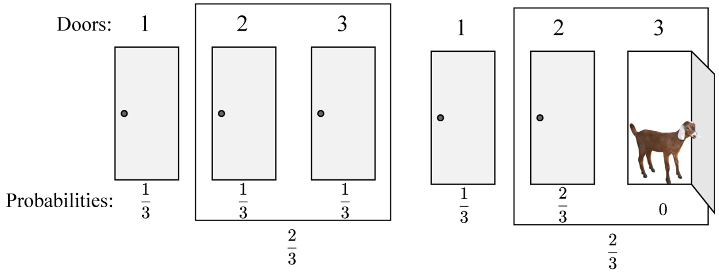
k: the door that gets opened

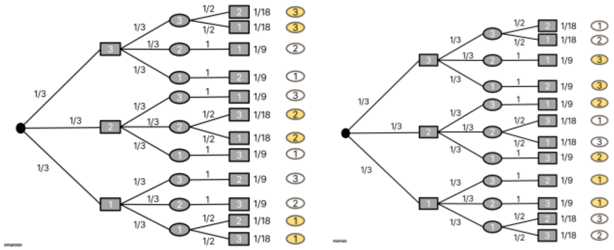
Chance of winning not switching: $6 \times \frac{1}{9} = 2/3$

Quick Note on Monty Hall

The first argument from earlier (on the left below) is much easier to follow than the sample space argument (on the right below).

- Both are rigorous. First is clearer.
- Why bother with the second argument? To show the universal applicability of the sample space framework.





Conclusion

Lecture 16, CS70 Summer 2025

Conclusion

Probability is hard!

- Intuitions can be wildly wrong.
- Rigor is critical for correctness.

Recall the key steps in all our calculations (framed slightly differently):

- Determine Ω, the set of all possible outcomes.
- Determine $\mathrm{P}(\omega)$, the probability of each outcome $\omega \in \Omega$?
  - For uniform probability spaces, $\mathrm{P}(\omega) = 1/|\Omega|$
- Determine the event A that we're interested in.
- Compute the probability of A by adding up the probabilities of the sample points contained in it.