Z. Abel, B. Chapman, E. Demaine
revised Wednesday 24th January, 2024

Lecture 18: Introduction to Probability

1 Probability

One of the most important disciplines in all of the sciences

- Randomized Algorithms (Miller-Rabin, 6.1210, 6.1220, 6.5220)
- Game Theory
- Information Theory
- Signal Processing
- Cryptography
- Machine Learning
- Medicine
- Statistics
- Forensics

Also one of the least understood

- Mark Twain: "There are three kinds of lies: lies, damned lies, and statistics."
- "Common sense" demonstrably unreliable
- Many graduate students don't know where to start thinking about a probability question
- Many garbage papers based entirely on fundamental misunderstandings of probability
- Even colleagues in CS who are pretty uncomfortable with probability and statistics
- Monty Hall Problem

Solution? Throw away intuition, and simply fall back to rigorous, step-by-step analysis.

2 Monty Hall Problem

Let's examine an illustrative example. In 1990, *Parade Magazine* columnist Marilyn vos Savant received the following letter from one Craig Whitaker:

> *Suppose you're on a game show, and you're given the choice of three doors. Behind one door is a car, behind the others, goats. You pick a door, say number 1, and the host, who knows what's behind the doors, opens another door, say number 3, which has a goat. He says to you, "Do you want to pick door number 2?" Is it to your advantage to switch your choice of doors?*

Marilyn replied that it is indeed advantageous to switch; if the car is behind the originally picked door, then you lose by switching, but if it's not (which is twice as likely), then you win by switching. She soon received a torrent of hate mail, including from professional mathematicians, adamantly asserting that she was wrong. The hypo described became known as the Monty Hall Problem, after the host of the game show Let's Make a Deal, in which contestants faced very similar situations.

3 The Tree Method

We will solve the Monty Hall Problem using the Tree Method, a simple, elementary, and rigorous approach that doesn't rely on intuition!

Step 0: The Question

Before we can even think about solving a mathematical problem, we need to make sure we really understand the setup and what exactly we're trying to ask. Craig's letter is not entirely precise, so we need to make some clarifying assumptions. We will assume:

- The prize is equally likely to be behind any one of three doors: A, B, or C.
- Marilyn is equally likely to pick any of the three doors, regardless of the prize's location.
- After Marilyn picks a door, Monty must open an unpicked door with a zonk (non-prize) behind it, and offer Marilyn the option to stay with the originally picked door or switch to the other unopened door.
- If Monty has a choice of two unpicked doors with zonks behind them, he is equally likely to open either door.

Note that these are *axioms*; if you accept them, then you must also accept whatever conclusions we reach, but there are perfectly reasonable axioms that lead to different conclusions! At one extreme, a possible axiom is that Monty is tricksy and only offers the option to switch if Marilyn originally picked the prize, in which case switching always loses. At the other extreme, another possible axiom is that Monty is benevolent and only offers the option to switch if Marilyn originally picked a zonk, in which case switching always wins.

With our axioms, we can now pose the precise mathematical question: "What is the probability that Marilyn (who accepts the offer to switch) wins the prize?"

Step 1: The Sample Space

A probability problem models some kind of random process, experiment, or game. Underlying every probability problem is a *probability space*, which consists of a *sample space* and a *probability function.*

Definition 1. *A* discrete probability space *is a pair* $(\mathcal{S}, \Pr)$*, where:*

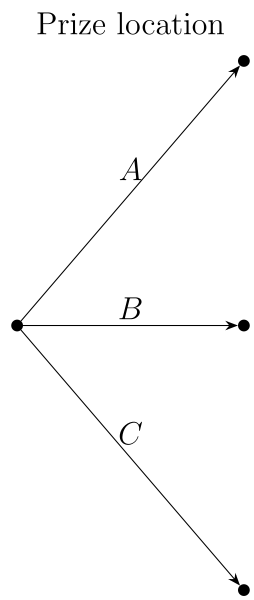
- $\mathcal{S}$ *is a non-empty countable set, called the* (discrete) sample space*, and*
- $\Pr : \mathcal{S} \to [0, 1]$ *is a total function with* $\sum_{\omega \in \mathcal{S}} \Pr[\omega] = 1$*, called the* probability function.

Definition 2. *An element* $\omega \in \mathcal{S}$ *is called an* outcome.

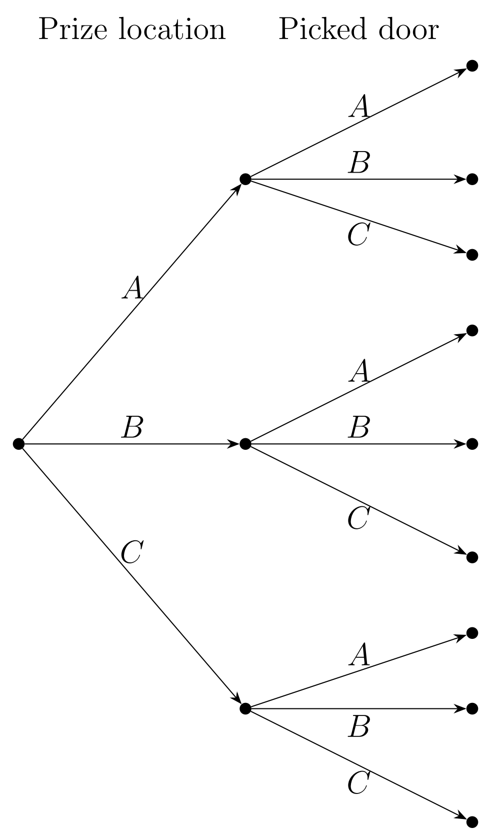
Our first objective is to identify the relevant sample space using a Tree Diagram. (We will ignore the probability function for the moment, but we will return to it later.) Each level of the Tree Diagram models a step of the random process, and it will branch to represent the different possible results of that step. (Recall Counting by Recipe!) In the case of Monty Hall, there are three steps:

- Monty hides a prize behind one door
- Marilyn picks a door
- Monty opens an unpicked door to reveal a zonk

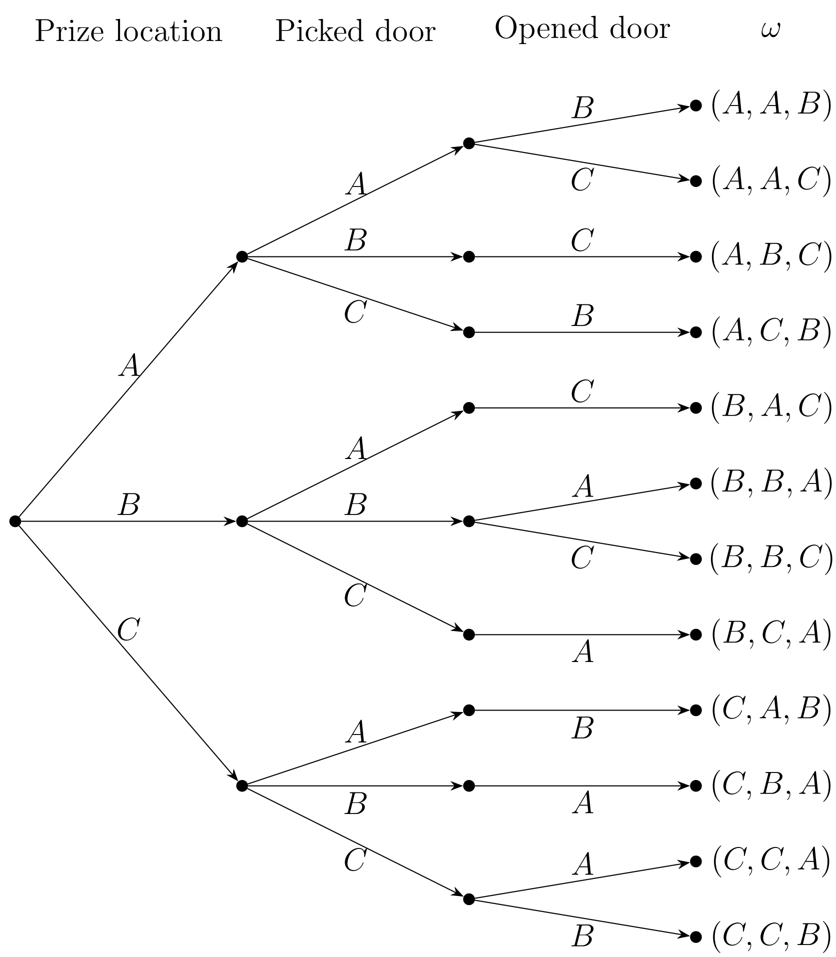
The first step could result in a prize behind door A, behind door B, or behind door C, so the first level of the Tree Diagram looks like the following:



At the second step, the player can pick any of the three doors, so each branch of the tree must branch three times.



Finally, Monty opens a door that doesn't match either of the previous two results. This could be a choice between either one or two doors, depending on whether the previous results were the same.



If we start at the root (source), and we follow edges based on the result of each step, then the leaf (sink) we reach (or equivalently, the path we follow) will tell us everything we need to know about how the random process turned out. Each leaf (sink) in our Tree Diagram is therefore an outcome, and the sample space is the set of leaves! In the case of Monty Hall, the top leaf represents the outcome where the prize is behind door A, Marilyn picks A, and Monty reveals a zonk behind B.

We can represent our sample space concisely as $\{(x, y, z) \in \{A, B, C\}^3 : z \neq x \wedge z \neq y\}$.

Step 2: The Probability Function

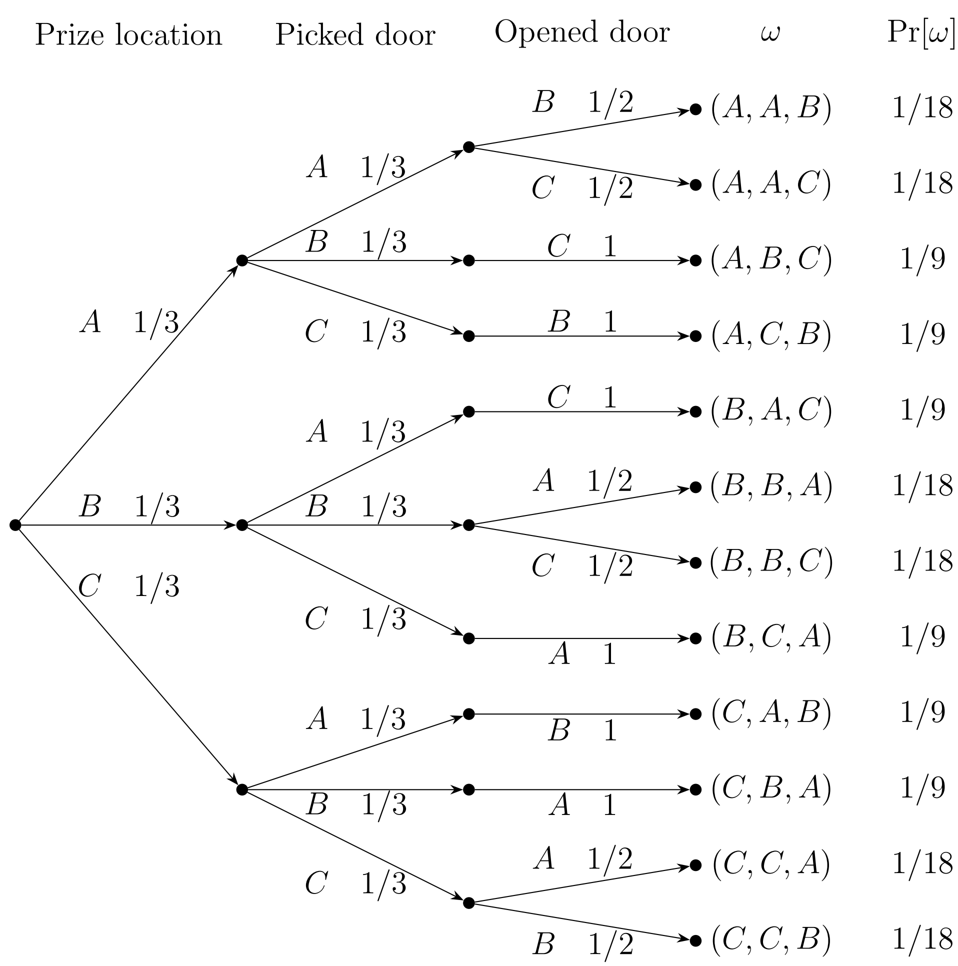
Now that we have our sample space, we must compute our probability function. To do this, we will assign a probability to each edge in our Tree Diagram. If we have a probability p

on edge (u, v), that means that assuming we have already reached u, we proceed to v with probability p. The probabilities of the edges leaving u should always sum to 1, and they should all be immediate from the process being modeled. (If they are not immediate, you may want to restructure your Tree Diagram! E.g., if the first level of the tree were the door opened instead of the prize location, the probabilities wouldn't be immediately obvious.)

For Monty Hall:

- Edges in leftmost layer have probability 1/3.
- Edges in middle layer have probability 1/3.
- Edges in rightmost layer have probability 1 or 1/2, depending on whether Monty had one or two choices.

For an outcome ω, the probability $\Pr[\omega]$ is the product of the probabilities along the path from the root to ω. This will be justified more formally in Lecture 19.



Step 3: The Events

Definition 3. *A subset $A \subseteq \mathcal{S}$ is called an* event.

Examples:

- $\{(A, B, C), (A, C, B), (B, A, C), (B, C, A), (C, A, B), (C, B, A)\}$ is the event [Marilyn wins]
- $\{(A, A, C), (A, B, C), (B, A, C), (B, B, C)\}$ is the event [Monty opens door C]
- $\{(A, A, B), (A, A, C), (A, B, C), (A, C, B)\}$ is the event [the prize is behind door A]

Note the use of the notation $[X]$ to mean "the event (set of outcomes) in which X occurs".

Step 4: The Answer

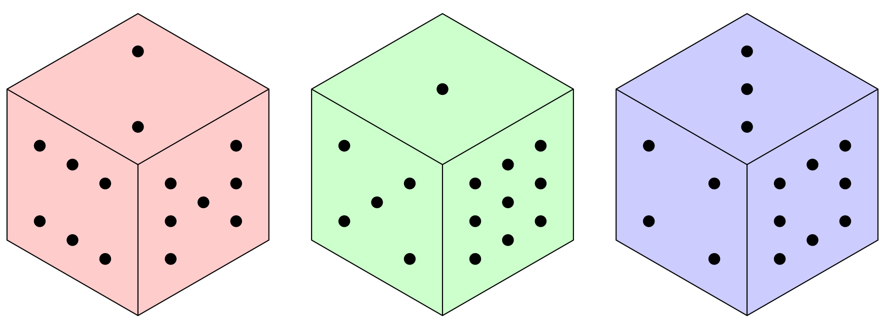
Definition 4. *We extend the probability function* Pr *to events: for an event A, we define the* probability of A *as*

$$\Pr[A] := \sum_{\omega \in A} \Pr[\omega]$$

We know the probabilities of each outcome, so we can compute the probabilities of events! In this case, each of the six outcomes in the event [Marilyn wins] occurs with probability 1/9, so Pr[Marilyn wins] $= 6/9 = 2/3$. Marilyn was correct! No intuition, no ingenious analogy, no fuss, no mess, just arithmetic over $\mathbb{Q}$. Hardest part: resisting temptation to jump to an "obvious" conclusion.

4 Strange Dice

Suppose we have three strange dice:



Consider the following game:

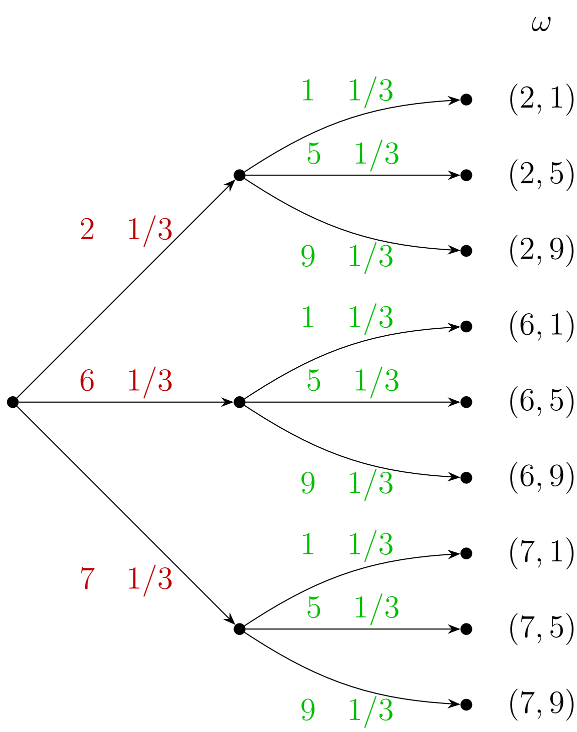
- Player 1 picks a die.

- Player 2 picks a different die (after seeing Player 1 pick).
- Both players roll their chosen die.
- Higher number wins.

Who should win?

4.1 Red vs. Green

Let's draw the Tree Diagram!

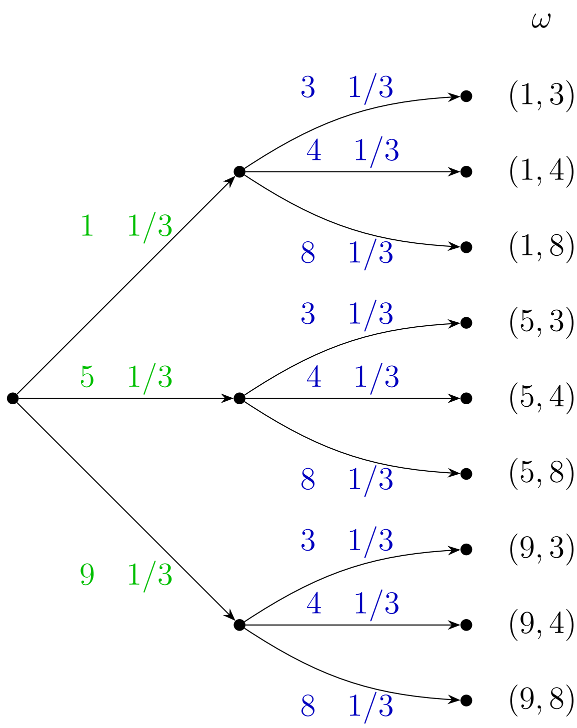


$\mathcal{S} = \{2, 6, 7\} \times \{1, 5, 9\}$. Each outcome is a pair of die rolls (r, g) where r is the red result and g is the green result. Each die has 1/3 probability of showing each possible number, regardless of the other die roll. Therefore, every edge has probability 1/3, and every outcome has probability 1/9.

Definition 5. *A probability space* $(\mathcal{S}, \Pr)$ *is* uniform *if* $\Pr$ *is the constant function.*

Uniform probability spaces are especially nice, because computing probabilities simply means counting outcomes! The event [Red wins] is $\{(2, 1), (6, 1), (6, 5), (7, 1), (7, 5)\}$, with 5 outcomes. There are 9 outcomes in total, so $\Pr[\text{Red wins}] = 5/9$.

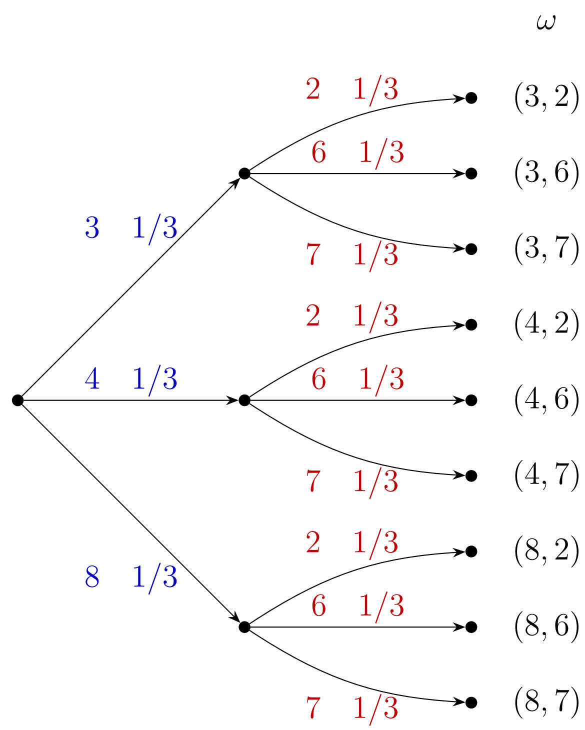
4.2 Green vs. Blue



The event [Green wins] is $\{(5,3),(5,4),(9,3),(9,4),(9,8)\}$, with 5 outcomes. We again have a uniform space with 9 outcomes in total, so Pr[Green wins] $= 5/9$.

4.3 Blue vs. Red

If Red beats Green and Green beats Blue, then Red must beat Blue, right? NO!



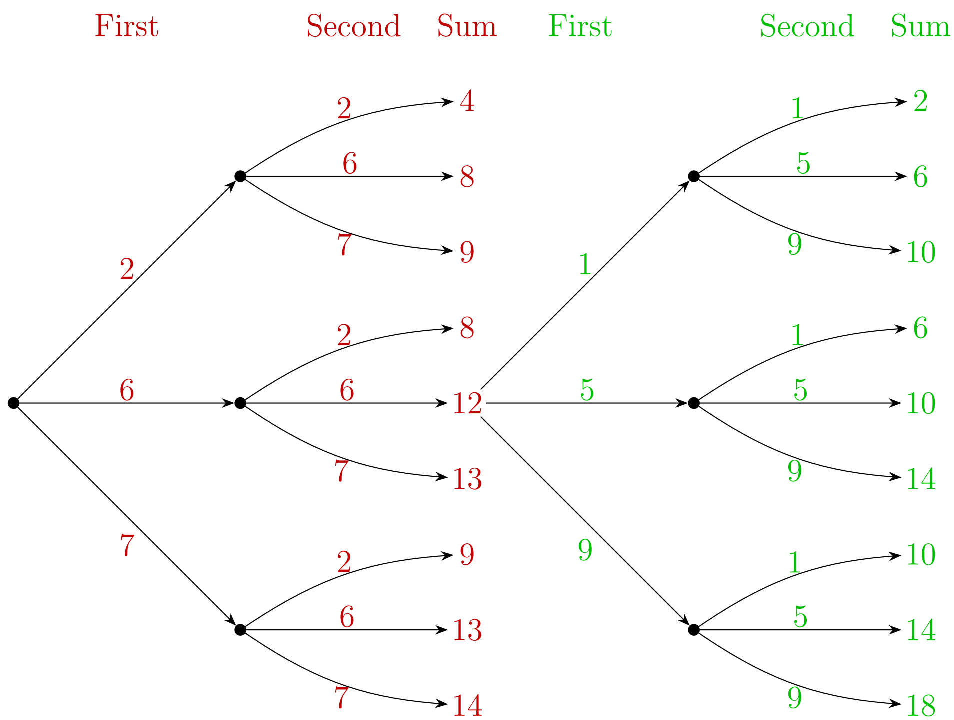
This time, Blue wins with probability 5/9. It may be counterintuitive that Red beats Green with probability 5/9, Green beats Blue with probability 5/9, and Blue beats Red with probability 5/9. However, there's no reason that this "beats" relation should be transitive; it's just faulty intuition!

So Player 2 should win with probability 5/9; regardless of which die Player 1 chooses, Player 2 can counter it.

But wait; it gets even weirder...

4.4 Red 2.0 vs. Green 2.0

Suppose each player rolls their chosen die *twice*, and whoever has the higher sum wins. This time, we have a uniform probability space with 81 outcomes, so the Tree Diagram will be rather cumbersome to draw in its entirety. Instead, we'll look for a pattern and abbreviate. We'll also omit the probabilities, since we've already established that we have a uniform probability space.



After rolling the red die twice, we reach one of the nine red vertices labeled with the sum of the two rolls. Each such vertex is the root of a subtree identical to the green one drawn. An outcome therefore combines a leaf of the red subtree with a leaf of the green one. How many outcomes are in the event [Red wins]? There is 1 in the topmost subtree ($4 > 2$). There are 3 in the second subtree ($8 > 6$ for two outcomes, and $8 > 2$ for one). We similarly count 3 in the third, fourth, and seventh subtrees and 6 in the fifth, sixth, eighth, and ninth subtrees. This gives a total of $1+3+3+3+6+6+3+6+6 = 37$ outcomes. The event [Red loses] contains $8+6+6+6+3+3+6+3+1 = 42$ outcomes. The event [Draw] contains the remaining 2 outcomes. We conclude that Pr[Red wins] $= 37/81$, and Pr[Red loses] $= 42/81$. With one roll, Red beats Green more often than not, but with two rolls, Green beats Red more often than not!

In fact, if we roll twice, the dice continue to be intransitive, but the order reverses! For more fun with Intransitive Dice, see https://arxiv.org/abs/1311.6511.