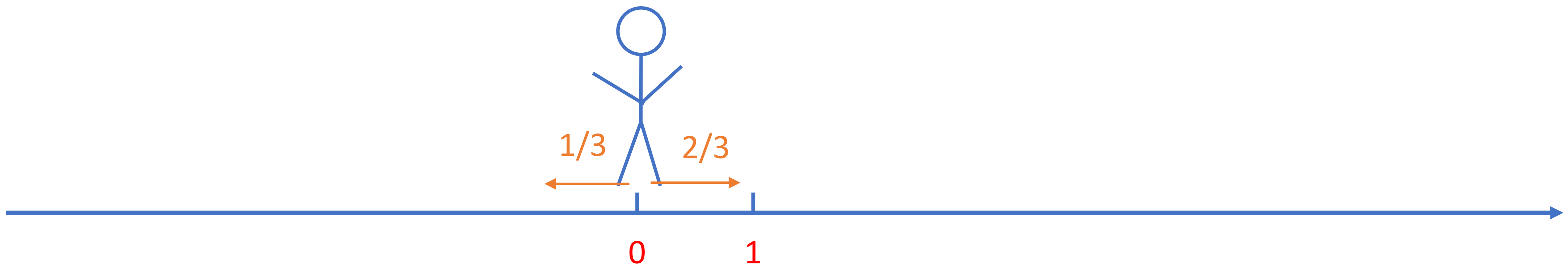
Lecture 16: Expectation



Expectation

Intuition (Lottery Example).

Say there are two lotteries:

1. 10% prob. of winning \$1000
2. 0.01\$ prob. of winning \$2000

Which one is more preferable?

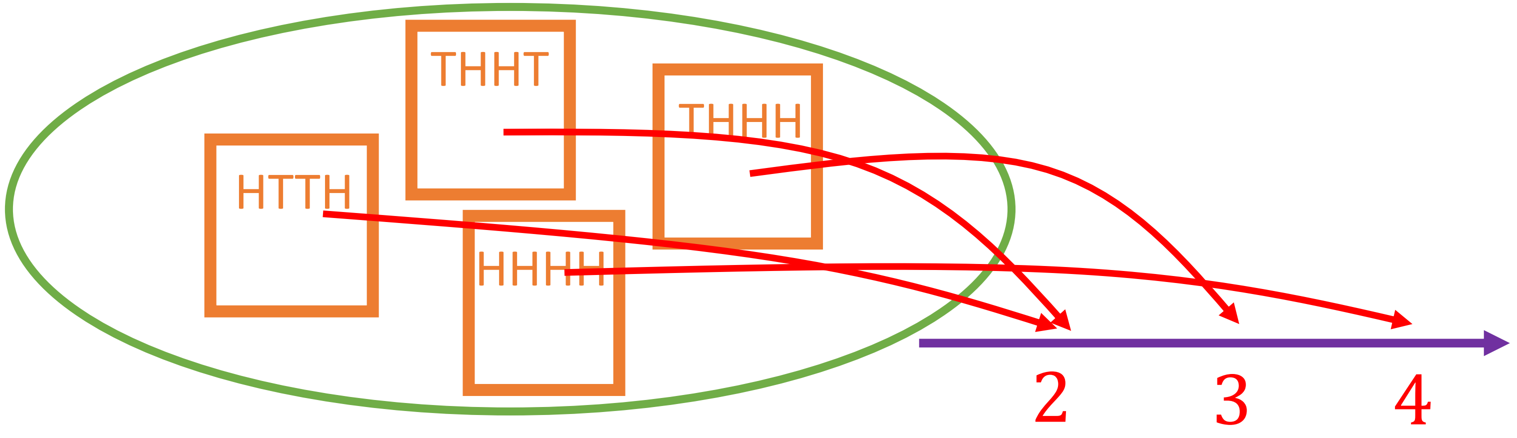
$10\% \cdot 1000 = 100$ >>. $0.01\% \cdot 2000 = 0.2$

Expectation

Definition-1.

The expectation of a random variable X is defined as,

$$\mathbb{E}[X] = \sum_{\omega} X(\omega) \cdot \mathbb{P}(\omega).$$



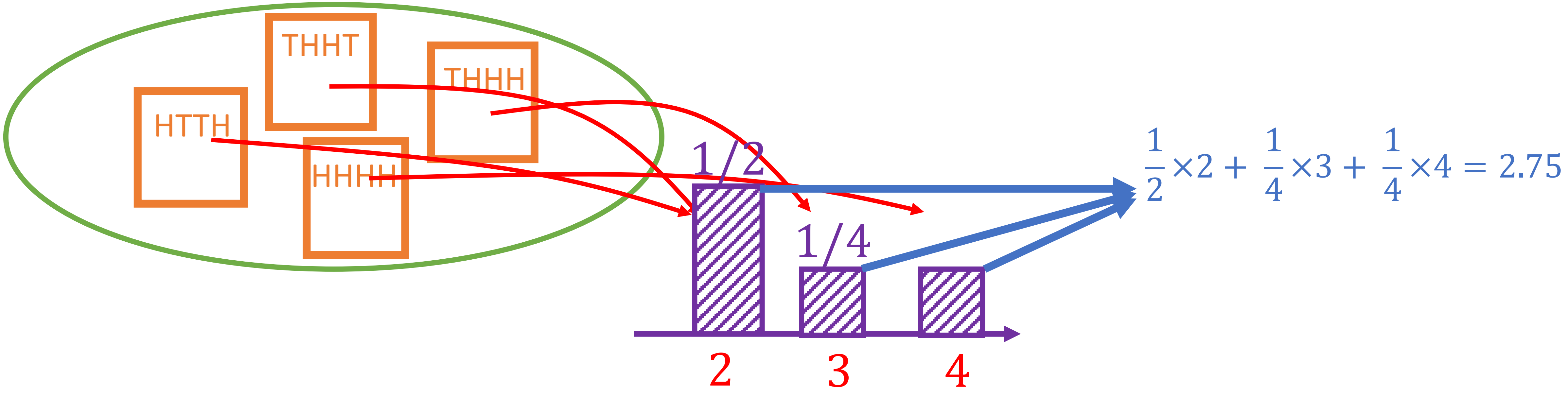
$$\frac{1}{4}\times 2 + \frac{1}{4}\times 2 + \frac{1}{4}\times 3 + \frac{1}{4}\times 4 = 2.75$$

Expectation

Definition-2.

The expectation of a random variable X is defined as,

$$\mathbb{E}[X] = \sum_a \mathbb{P}[X = a] \cdot a\,.$$



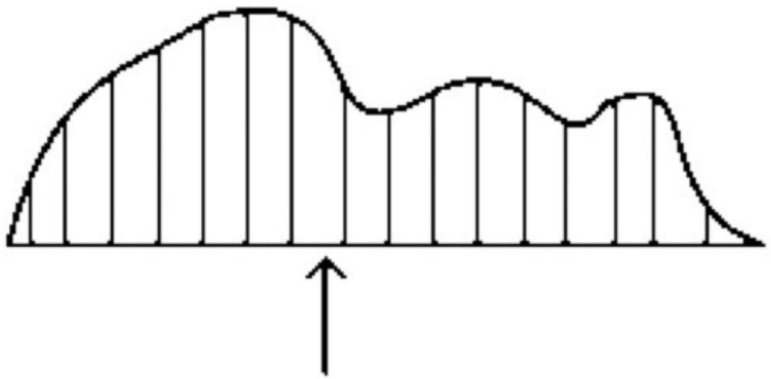
Expectation

Definition.

The expectation of a random variable X is defined as,

$$\mathbb{E}[X] = \sum_a \mathbb{P}[X = a] \cdot a\,.$$

$E[X]$ measures the "center of mass" of the distribution



Linearity of Expectation

Theorem (Linearity).

For two jointly distributed random variables X, Y,

$$\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y].$$

Note X, Y do not need to be independent.

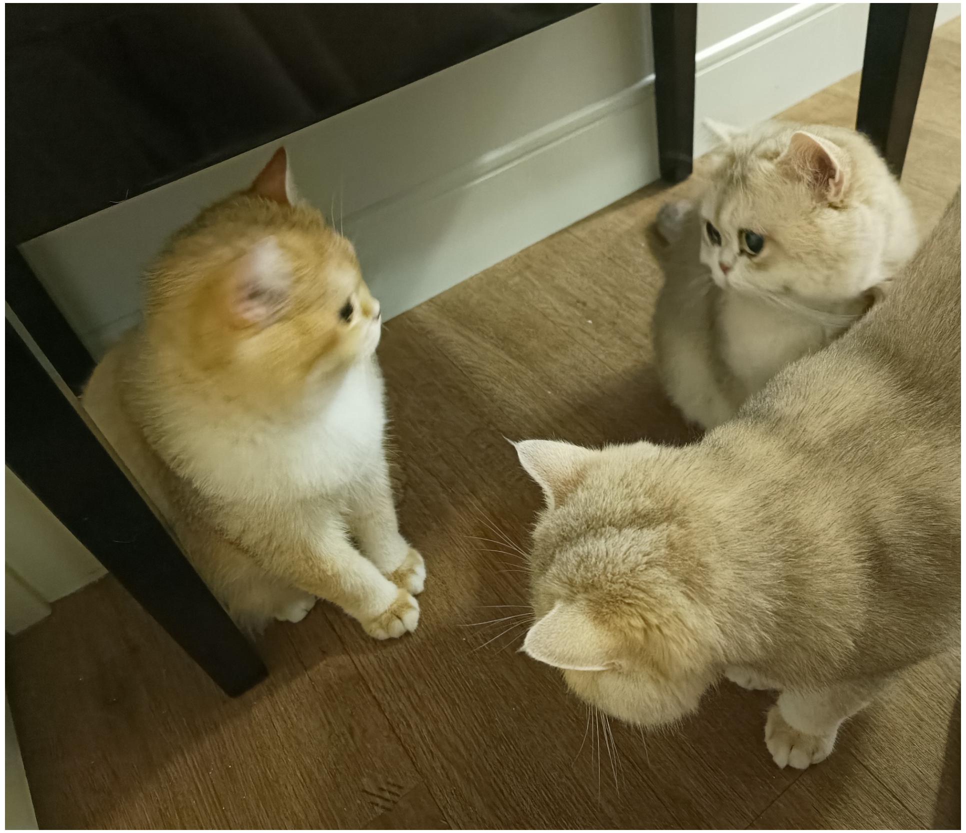
Proof.

$$\begin{aligned}
\mathbb{E}[X + Y] &= \sum_{a,b} \mathbb{P}[X = a, Y = b] \cdot (a + b) \\
&= \sum_{a,b} \mathbb{P}[X = a, Y = b] \cdot a + \sum_{a,b} \mathbb{P}[X = a, Y = b] \cdot b \\
&= \sum_{a} \mathbb{P}[X = a] \cdot a + \sum_{b} \mathbb{P}[Y = b] \cdot b \\
&= \mathbb{E}[X] + \mathbb{E}[Y]
\end{aligned}$$

Linearity of Expectation (Example)

Example (The matching problem).

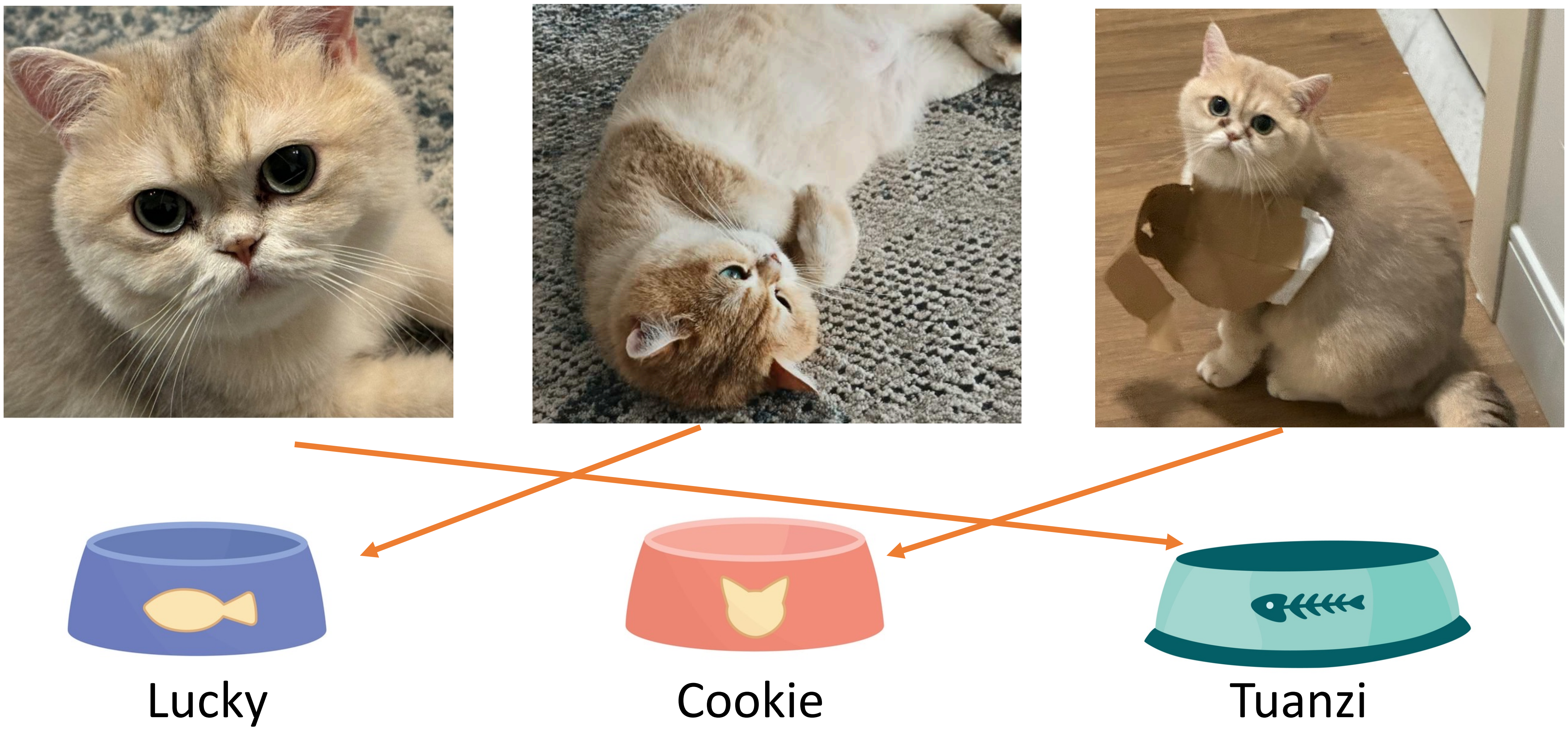
n cats each has a bowl with their name on it. However, when it comes time for dinner, each cat i goes to a random bowl p_i such that no two cats select the same bowl.



Suppose p is uniformly random over all permutations. Let X be the number of cats who got their own bowl.

What is $\mathbb{E}[X]$?

Linearity of Expectation (Example)



Linearity of Expectation (Example)

Example (The matching problem).

Let's try the definition first.

$$\mathbb{E}[X] = \sum_{a=0}^{n} \mathbb{P}[X = a] \cdot a$$

where $\mathbb{P}[X = a] = \mathbb{P}[\text{Exactly } a \text{ cats got their bowl}]$

$$= \frac{\#\{\text{permutation } p \text{ with exactly } a \text{ fixed points}\}}{n!}$$

Too hard!

Linearity of Expectation (Example)

Example (The hat-check problem).

Let's try the linearity approach. Let

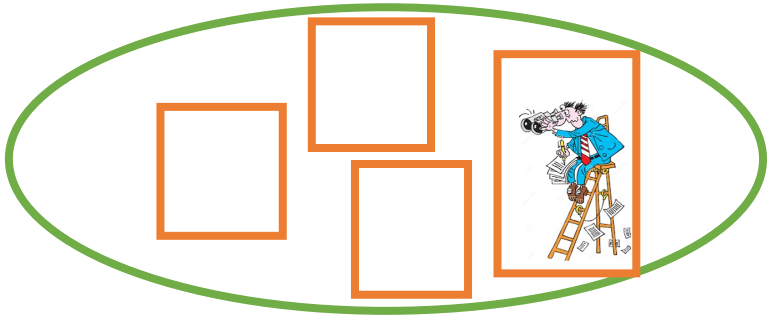
$$X_i = \mathbf{1}[\text{the } i\text{-th cat got its own bowl}]$$

be a 0/1 indicator random variable.

(This is also called the method of indicators.)

We know in any possible world (any outcome),

$$X = X_1 + X_2 + \cdots + X_n.$$



Linearity of Expectation (Example)

Example (The hat-check problem).

Thus, by linearity, we know in expectation,

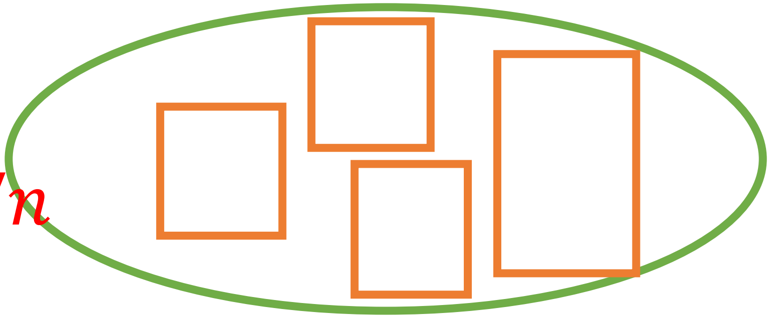
$$\mathrm{E}[X] = \mathrm{E}[X_1] + \mathrm{E}[X_2] + \cdots + \mathrm{E}[X_n].$$

By symmetry, we know

$$\mathrm{E}[X_1] = \mathrm{E}[X_2] = \cdots = \mathrm{E}[X_n]$$

What is $\mathrm{E}[X_1]$?

$$\mathrm{E}[X_1] = \mathbb{P}[1^{\text{st}} \text{ cat gets its bowl}] = 1/n$$



Linearity of Expectation (Example)

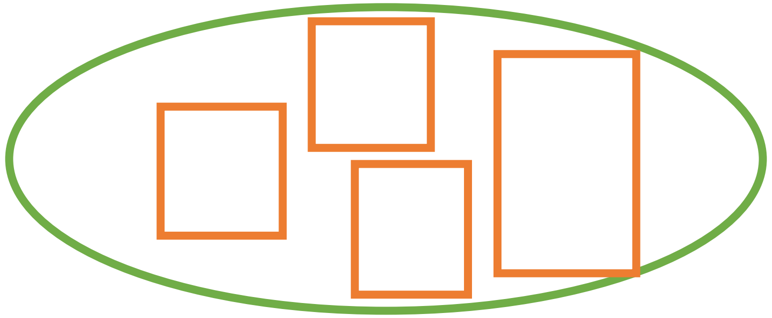
Example (The hat-check problem).

Thus, by linearity, we know in expectation,

$$\mathrm{E}[X] = \mathrm{E}[X_1] + \mathrm{E}[X_2] + \cdots + \mathrm{E}[X_n] = 1.$$

By symmetry, we know

$$\mathrm{E}[X_1] = \mathrm{E}[X_2] = \cdots = \mathrm{E}[X_n] = 1/n$$



Conditional Expectation

Definition.

Conditioning on an event E, the distribution of $X|E$ becomes

$$\mathbb{P}[\, X = a \mid E \,] = \frac{\mathbb{P}[\, X = a \wedge E \,]}{\mathbb{P}[E]}.$$

and the conditional expectation,

$$\mathbb{E}[\, X \mid E \,] = \sum_a \mathbb{P}[\, X = a \mid E \,] \cdot a$$

$$= \sum_{\omega \in E} \mathbb{P}[\, \omega \mid E \,] \cdot a$$

Conditional Expectation

Intuitively, $\mathbb{E}[X \cdot \mathbf{1}_E]$ is the event E part of X.
$\mathbb{E}[\, X \mid E \,]$ is just take that part out, and multiply it by a factor of $\frac{1}{\mathbb{P}[E]}$ due to conditioning.

A sometimes useful formula.

$$\mathbb{E}[\, X \mid E \,] = \frac{\mathbb{E}[X \cdot \mathbf{1}_E]}{\mathbb{P}[E]}$$

where $\mathbf{1}_E$ is the indicator variable where $\mathbf{1}_E(\omega) = \mathbf{1}[\omega \in E]$.

Proof.

$$\begin{aligned}
\mathbb{E}[\, X \mid E \,] &= \textstyle\sum_{\omega \in E} \mathbb{P}[\omega | E] \cdot X(\omega) \\
&= \textstyle\sum_{\omega \in \Omega} \mathbb{P}[\omega | E] \cdot X(\omega) \cdot \mathbf{1}_E(\omega) \\
&= \textstyle\sum_{\omega \in \Omega} \frac{\mathbb{P}[\omega]}{\mathbb{P}[E]} \cdot X(\omega) \cdot \mathbf{1}_E(\omega) \\
&= \frac{1}{\mathbb{P}[E]} \cdot \textstyle\sum_{\omega \in \Omega} \mathbb{P}[\omega] \cdot X(\omega) \cdot \mathbf{1}_E(\omega) = \frac{\mathbb{E}[X \cdot \mathbf{1}_E]}{\mathbb{P}[E]}
\end{aligned}$$

Law of total expectation

Previously, we have learned law of total probability:

Theorem (Law of total probability).

For any event E and F,

$$\mathbb{P}[F] = \mathbb{P}[F \mid E] \cdot \mathbb{P}[E] + \mathbb{P}[F \mid \neg E] \cdot \mathbb{P}[\neg E].$$

Theorem (Law of total expectation).

For any event E and variable X,

$$\mathbb{E}[X] = \mathbb{E}[X \mid E] \cdot \mathbb{P}[E] + \mathbb{E}[X \mid \neg E] \cdot \mathbb{P}[\neg E].$$

Law of total expectation

Theorem (Law of total expectation).

For any event E and variable X,

$$\mathbb{E}[X] = \mathbb{E}[X \mid E] \cdot \mathbb{P}[E] + \mathbb{E}[X \mid \neg E] \cdot \mathbb{P}[\neg E]$$

Proof

$$\begin{aligned}
\mathbb{E}[X] &= \textstyle\sum_a \mathbb{P}[X = a] \cdot a \\
&= \textstyle\sum_a (\mathbb{P}[X = a \mid E] \cdot \mathbb{P}[E] + \mathbb{P}[X = a \mid \neg E] \cdot \mathbb{P}[\neg E]) \cdot a \\
&= \mathbb{P}[E] \cdot \textstyle\sum_a \mathbb{P}[X = a \mid E] \cdot a + \mathbb{P}[\neg E] \cdot \textstyle\sum_a \mathbb{P}[X = a \mid \neg E] \cdot a \\
&= \mathbb{P}[E] \cdot \mathbb{E}[X \mid E] + \mathbb{P}[\neg E] \cdot \mathbb{E}[X \mid \neg E]
\end{aligned}$$

Multiplication Law of Independent Random Variables

Theorem.

For any two independent random variables X and Y,

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

Analogue w/ Counting

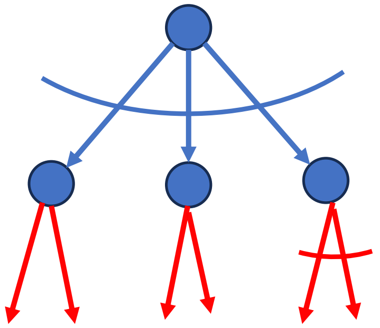
Multiplication rule in expectation

For any two independent random variables X and Y,

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

Multiplication rule in counting

If every one of A cases has B subcases, in total there are AB subcases.



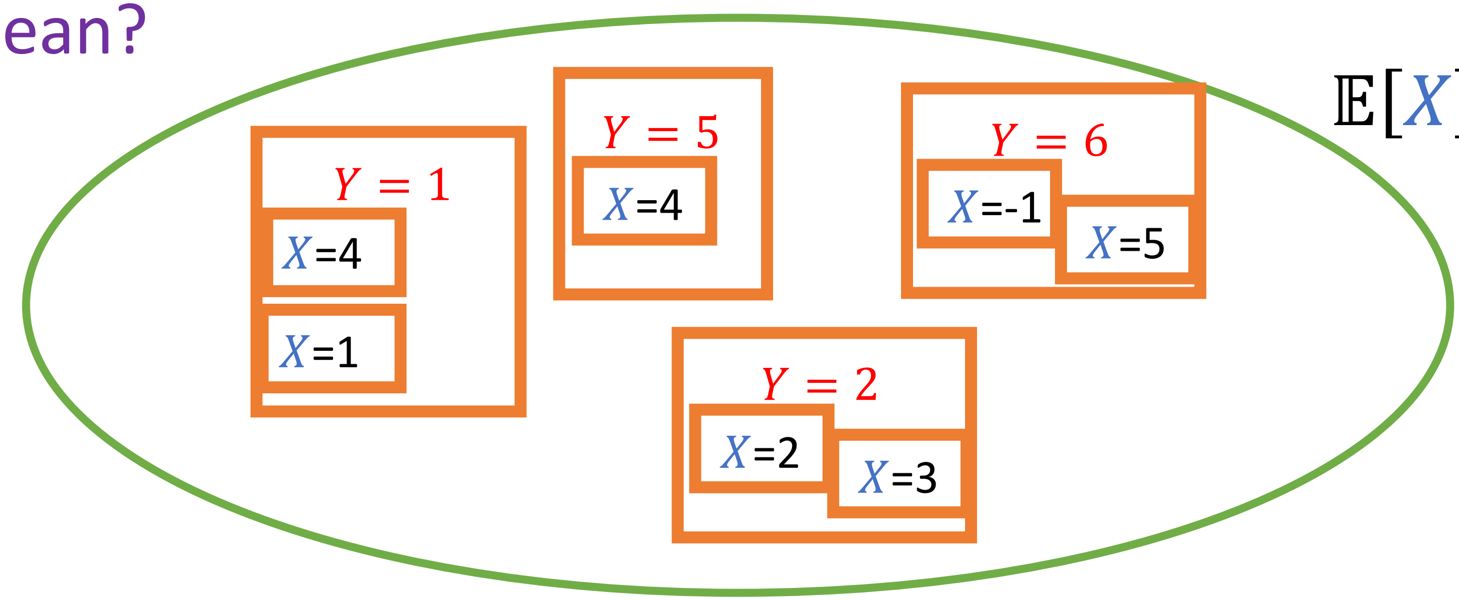
Law of iterated expectation

Lemma.

For any two random variables X and Y,

$$\mathbb{E}_Y[\mathbb{E}_X[X|Y]] = \mathbb{E}[X]$$

What does this mean?



$\mathbb{E}[X]$

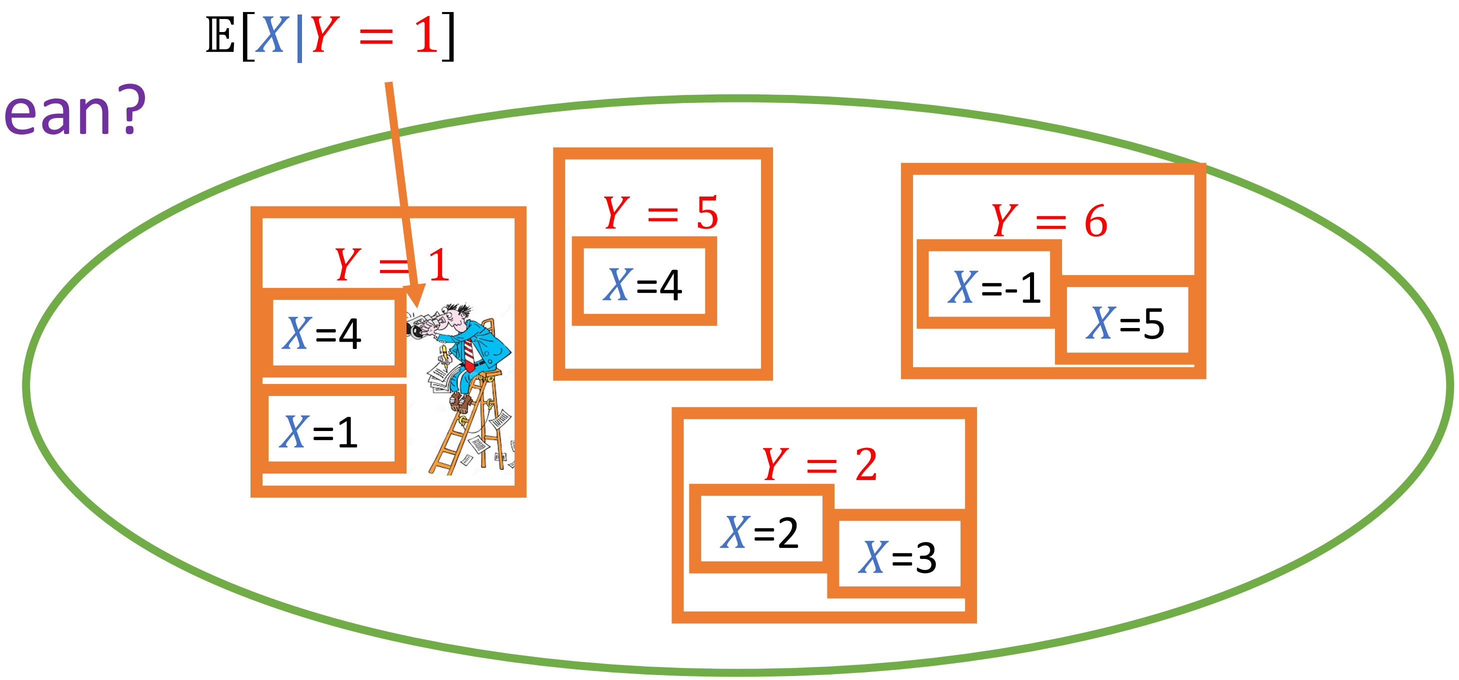
Law of iterated expectation

Lemma.

For any two random variables X and Y,

$$\mathbb{E}_Y[\mathbb{E}_X[X|Y]] = \mathbb{E}[X]$$

What does this mean?



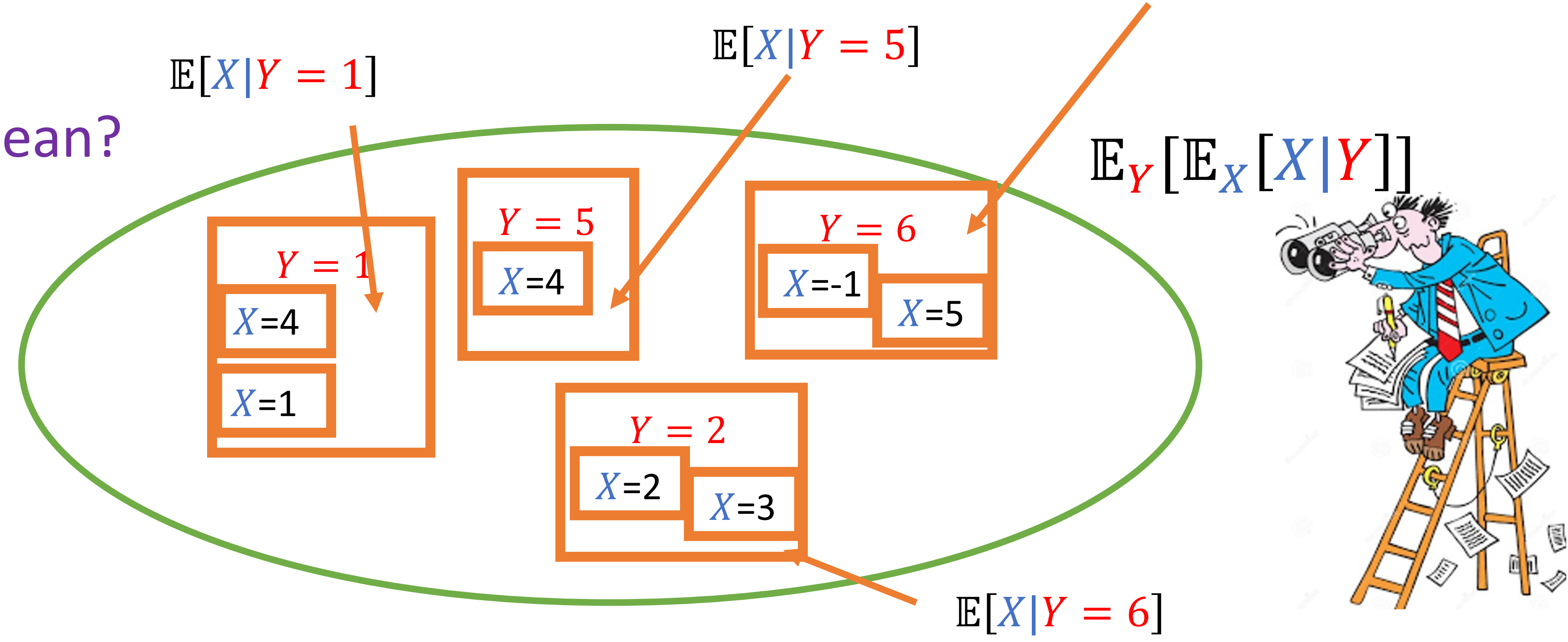
Law of iterated expectation

Lemma.

For any two random variables X and Y,

$$\mathbb{E}_Y[\mathbb{E}_X[X|Y]] = \mathbb{E}[X]$$

What does this mean?



Self-reference trick

Problem.

On an axis that is infinitely long on both ends, you start from 0.

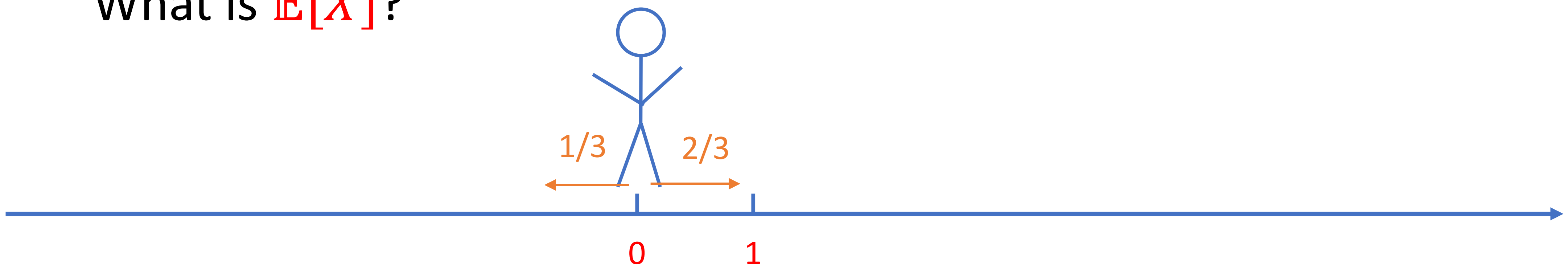
Each step:

With probability 2/3, you walk length one right.

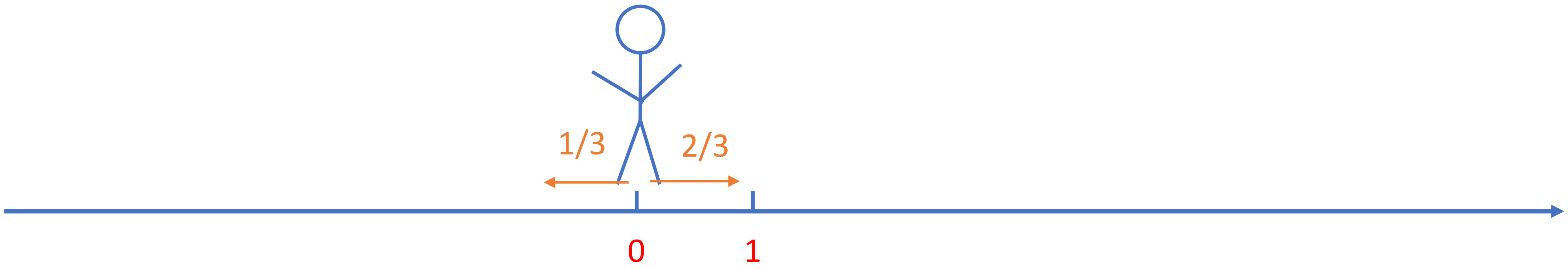
With probability 1/3, you walk length one left.

Let X be the number of steps you take to reach 1 for the first time.

What is $\mathbb{E}[X]$?



Self-reference trick



Let X be the number of steps you take to reach 1 from 0 for the first time.

With probability 2/3, first step you get to 1,

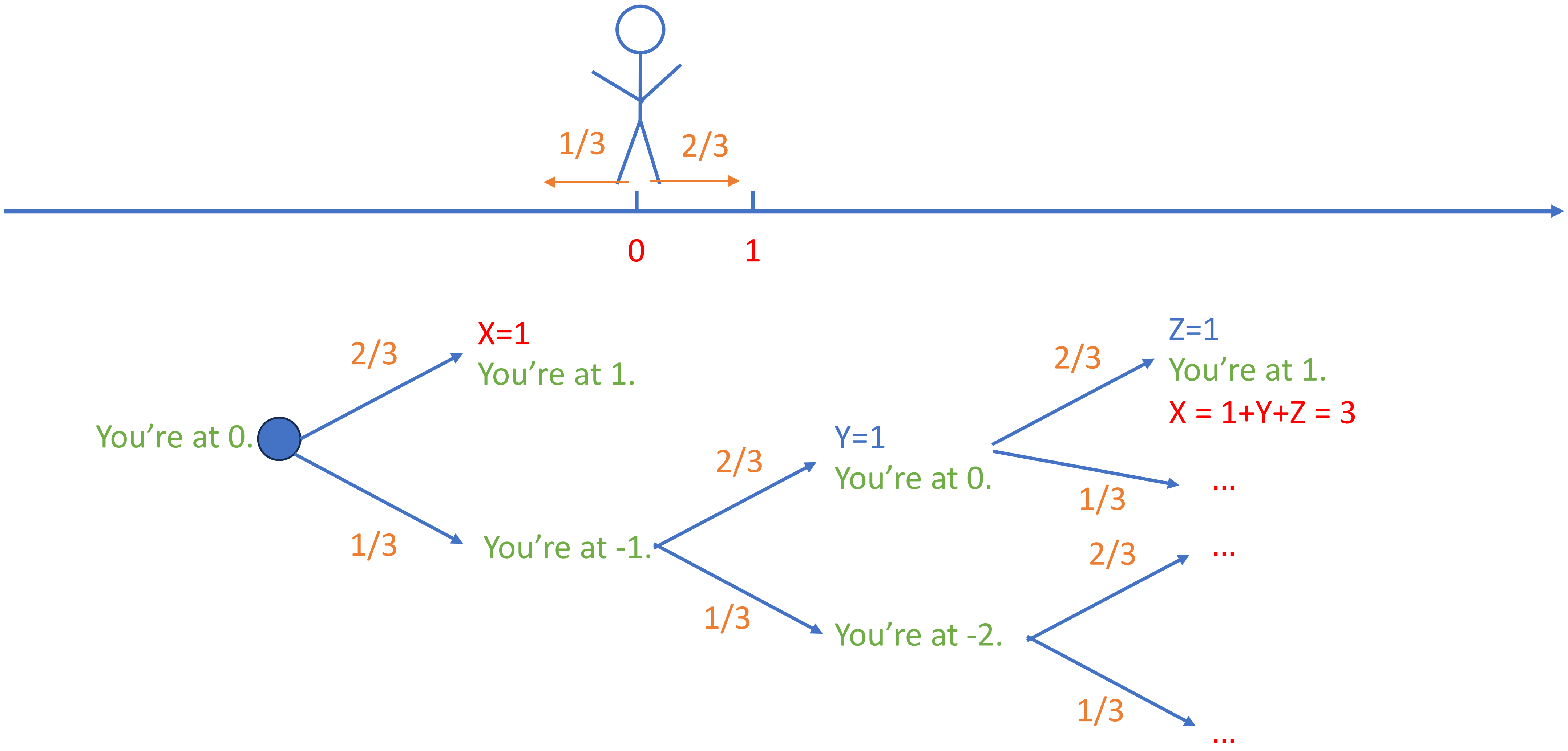
in that world, $X = 1$.

With probability 1/3, first step you get to -1.

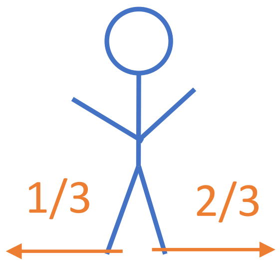
in that world, next, you need to first get from -1 ->0. (Y steps.)

then, you need to get from 0 -> 1. (Z steps.)

Self-reference trick



Self-reference trick



Let X be the number of steps 0 -> 1 for the first time.

Y be the number of steps -1 -> 0 for the first time (after you reach -1)

Z be the number of steps 0 -> 1 for the first time (after you reach 0 again.)

We know $\mathbb{E}[X] = \frac{2}{3} \cdot \mathbb{E}[X \mid \text{first step } 0 \to 1] + \frac{1}{3} \cdot \mathbb{E}[X \mid \text{first step } 0 \to -1]$

(law of total expectation)

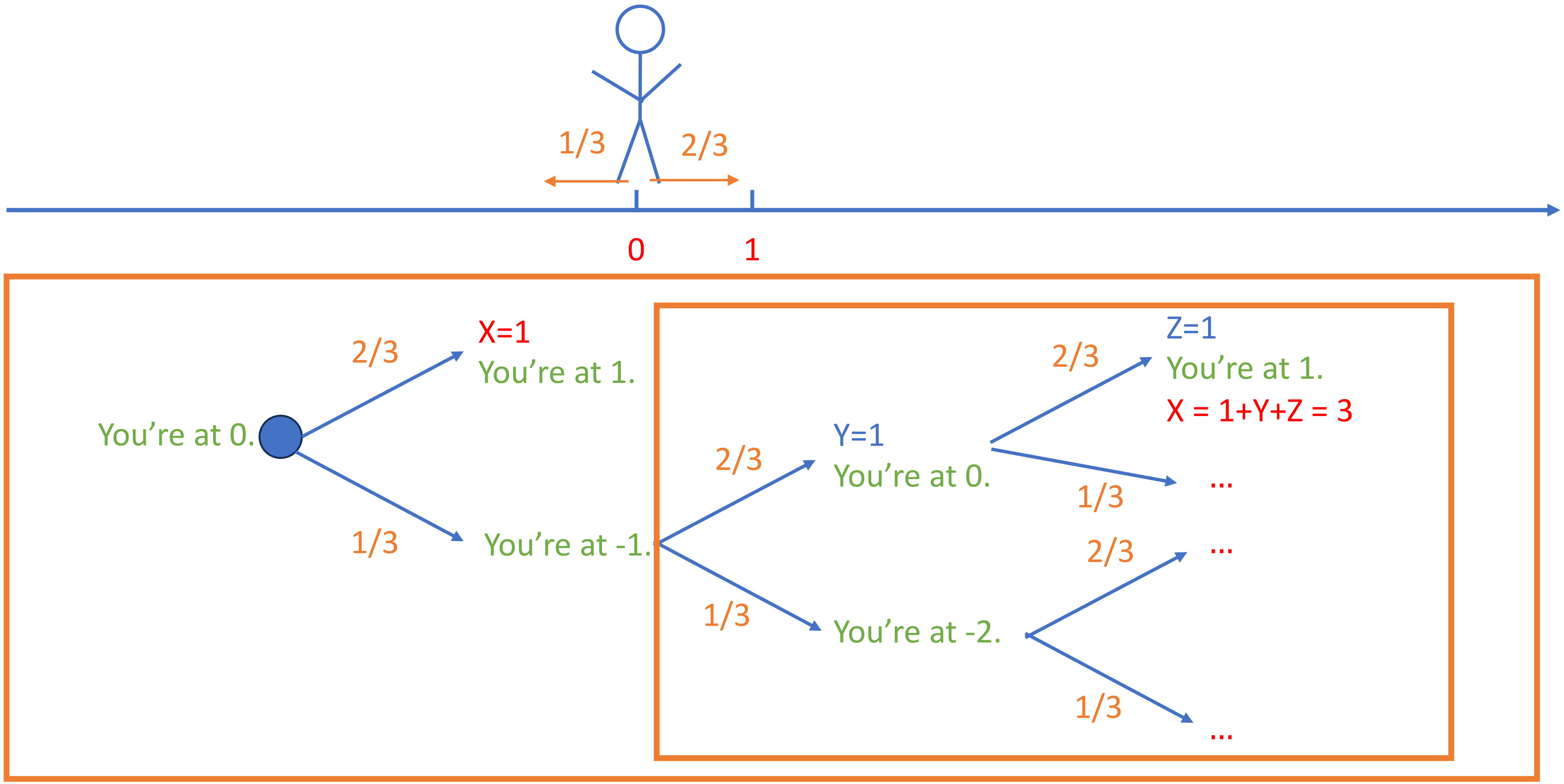
$$= \frac{2}{3} \cdot 1 + \frac{1}{3} \cdot (1 + \mathbb{E}[Y + Z \mid \text{first step } 0 \to -1]).$$

$$= \frac{2}{3} \cdot 1 + \frac{1}{3} \cdot (1 + \mathbb{E}[Y \mid \text{first step } 0 \to -1] + \mathbb{E}[Z \mid \text{first step } 0 \to -1]).$$

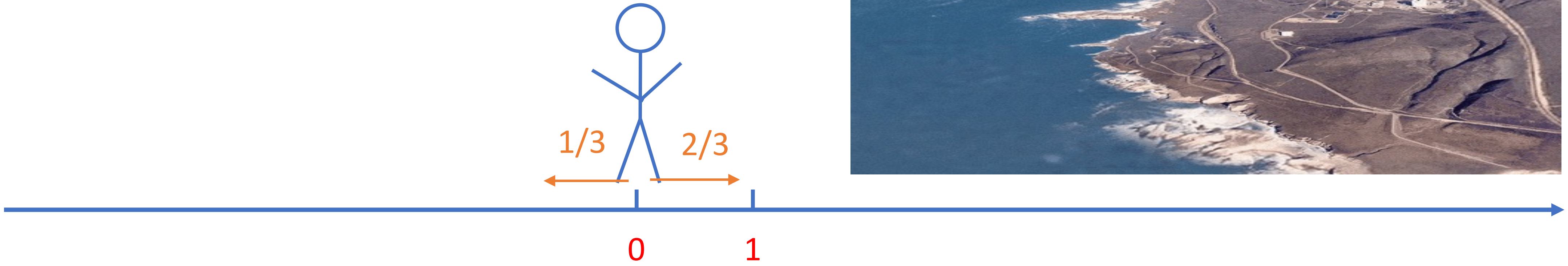
(linearity of expectation)

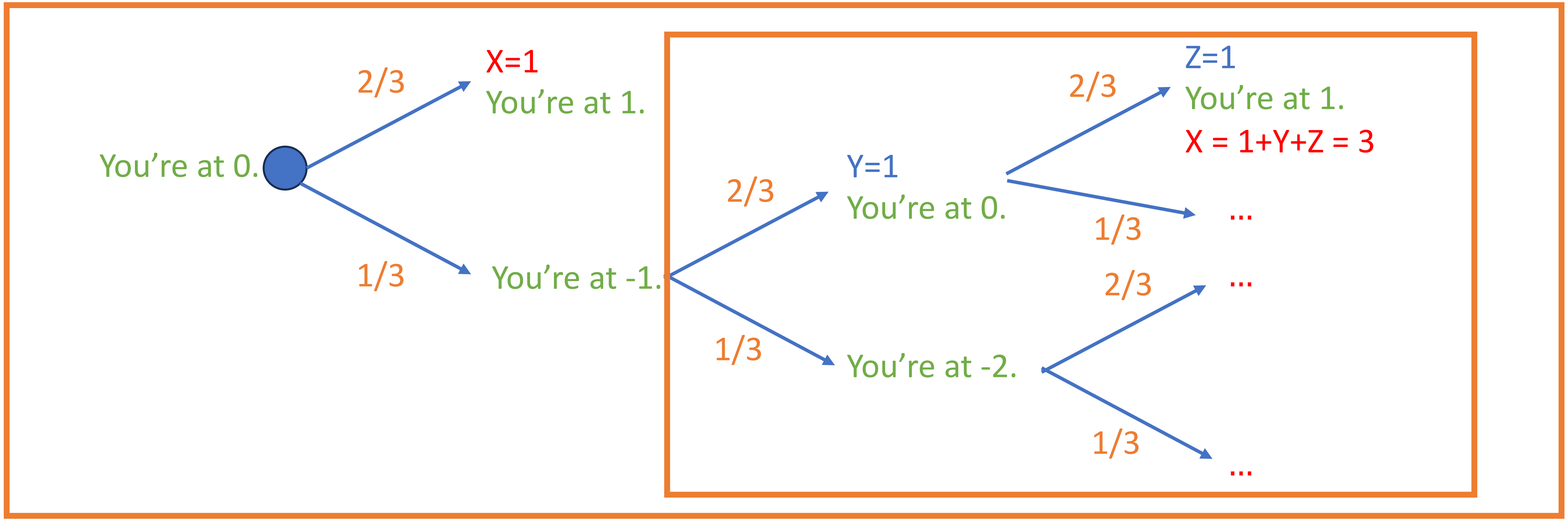
$$= \frac{2}{3} \cdot 1 + \frac{1}{3} \cdot (1 + \mathbb{E}[X] + \mathbb{E}[X])..$$

Self-reference trick

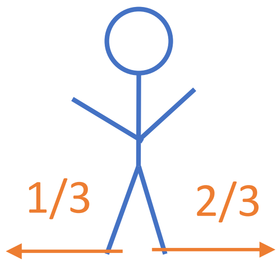


Self-reference trick





Self-reference trick



Let X be the number of steps 0 -> 1 for the first time.

We know $\mathbb{E}[X] = \frac{2}{3} \cdot 1 + \frac{1}{3} \cdot (1 + 2\mathbb{E}[X])$.

=> $\frac{1}{3}\mathbb{E}[X] = 1$.

Thus $\mathbb{E}[X] = 3$.

Crazy, right?

It's ok if you don't feel like fully understand it. Try to revisit this example after we learn Markov Chains.

Envelope Paradox

Say I have two envelopes that contain $$$

One contains twice the money of the other one and I randomly swapped two.

Strategy 1: Pick one envelope, get x dollars.

Strategy 2: Switch to other one,

with ½ probability, get $x/2$ dollars.

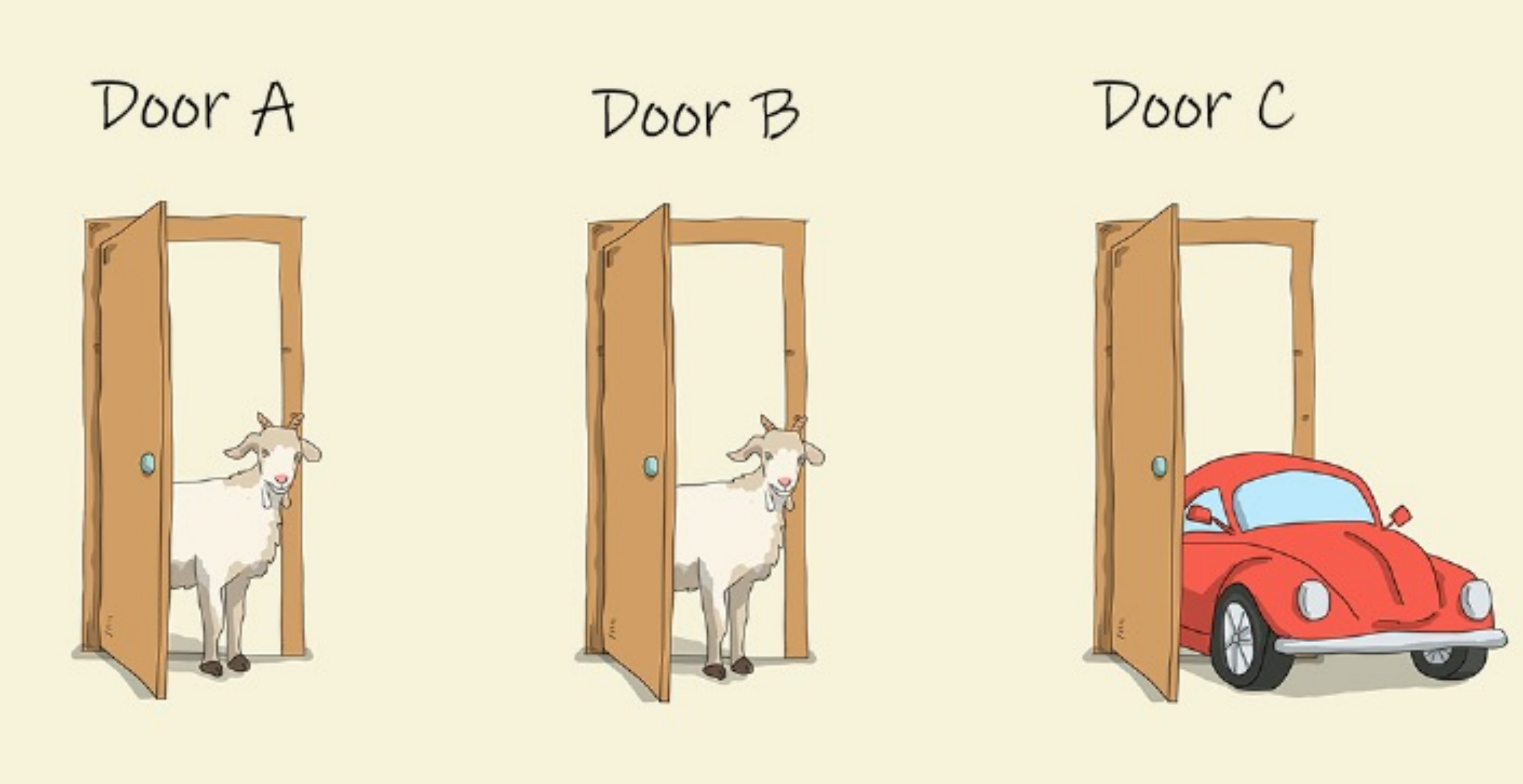
with ½ probability, get $2x$ dollars.

In expectation, $\frac{1}{2} \cdot \frac{x}{2} + \frac{1}{2} \cdot 2x = 1.5\,x$

What's wrong?

Monty Hall Problem

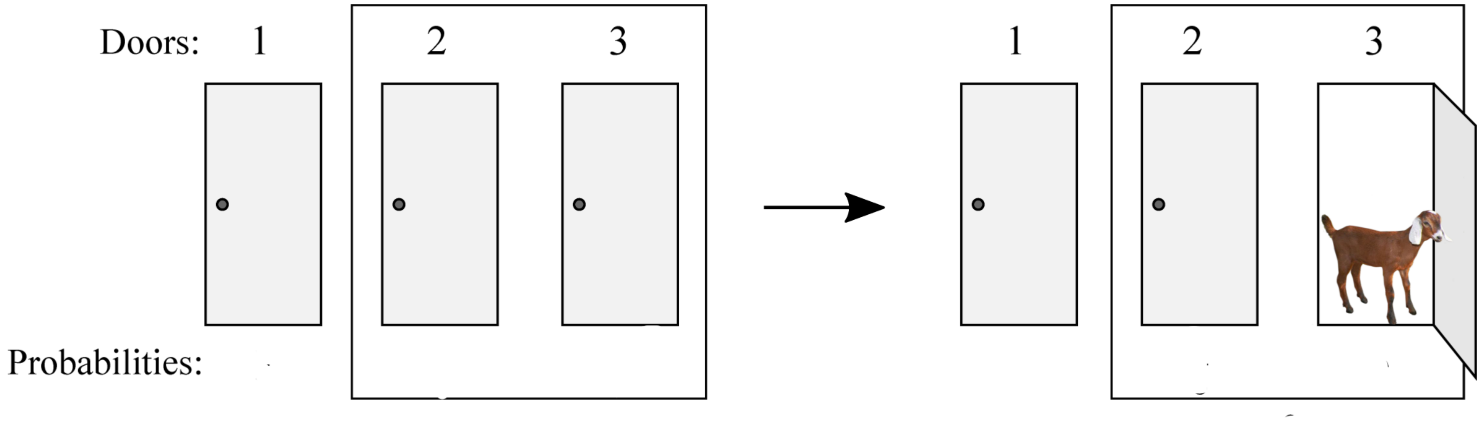
There are three doors. Behind one there is a car. Behind the other two are just goats.



Monty Hall Problem

There are three doors. Behind one there is a car. Behind the other two are just goats.

You choose one door (not opened). Then the host opens a door in the remaining two doors that has a goat behind it.

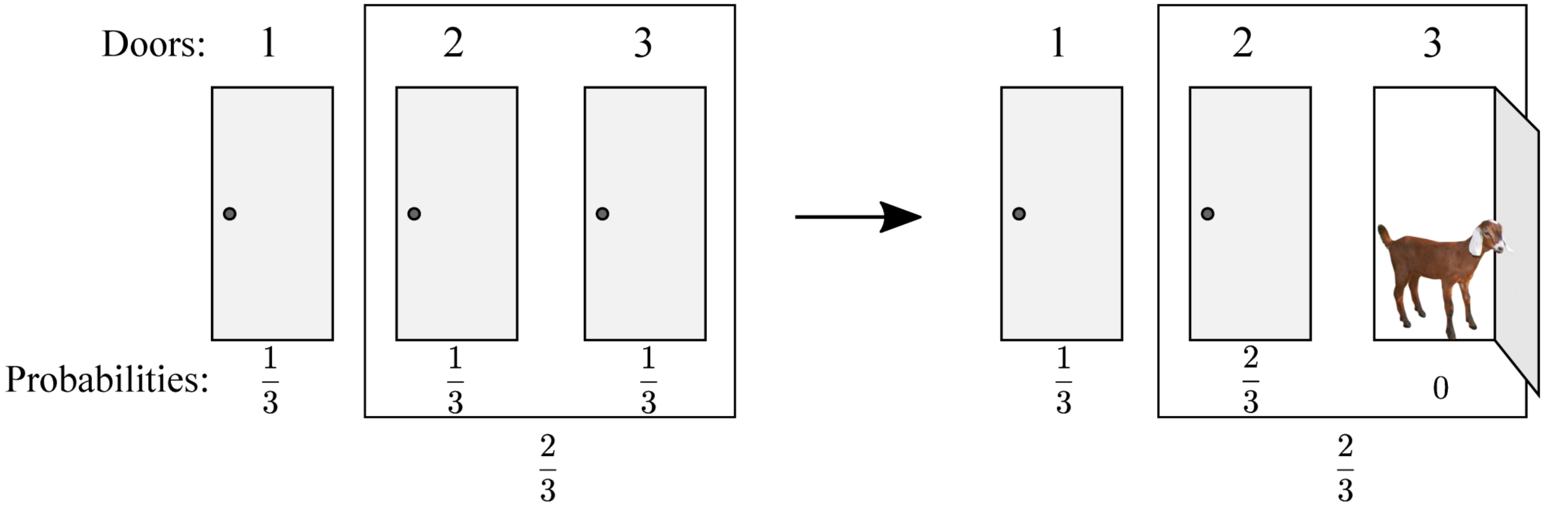


Then the host asks you whether you'd like to switch.

Monty Hall Problem

There are three doors. Behind one there is a car. Behind the other two are just goats.

You choose one door (not opened). Then the host opens a door in the remaining two doors that has a goat behind it.



The smart thing to do: Always switch!