Probability

Precise model for Random Experiments

Randomness in computation
- randomness in data
- probabilistic algorithms
  e.g. Quicksort.

Probability: Precise, unambiguous way to argue about uncertainty

intuitive?

Summary

We model random experiment as

Probability Space
- **Sample Space** Ω – set of possible outcomes
- **Probabilities** assigned to each outcome $\omega \in \Omega$.
  - $\forall \omega \in \Omega \qquad 0 \le \Pr[\omega] \le 1$
  - add up to 1: $\quad \sum_{\omega \in \Omega} \Pr[\omega] = 1$

An **Event** E is a subset of Ω.

$$\Pr[E] = \sum_{\omega \in \Omega} \Pr[\omega].$$

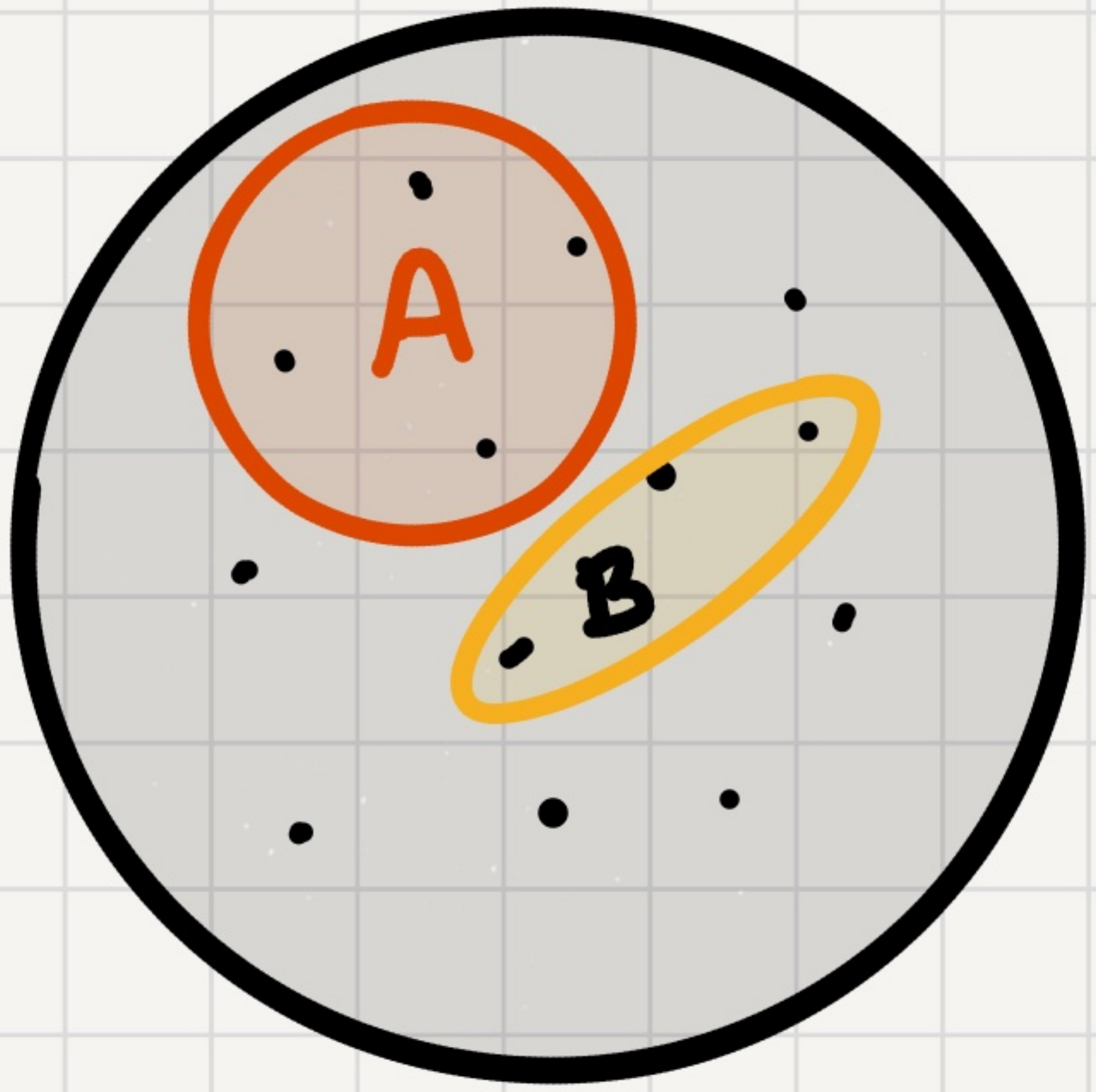
Uniform probability space $\quad \forall \omega \in \Omega: \quad \Pr[\omega] = \frac{1}{|\Omega|}$

$$\Pr[E] = \frac{|E|}{|\Omega|}$$

Probability is Additive

Theorem: If A and B are disjoint events (i.e. $A \cap B = \emptyset$), then

$$\Pr[A \cup B] = \Pr[A] + \Pr[B]$$



Proof:

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

$$= \sum_{\omega \in A} \Pr[\omega] + \sum_{\omega \in B} \Pr[\omega] \quad \left(\text{since } A \cap B = \emptyset\right)$$

$$= \Pr[A] + \Pr[B].$$

Probability is Additive

Theorem: If $A_1, A_2, \dots, A_n$ are pairwise disjoint, i.e. $\forall i \neq j : A_i \cap A_j = \emptyset$, then

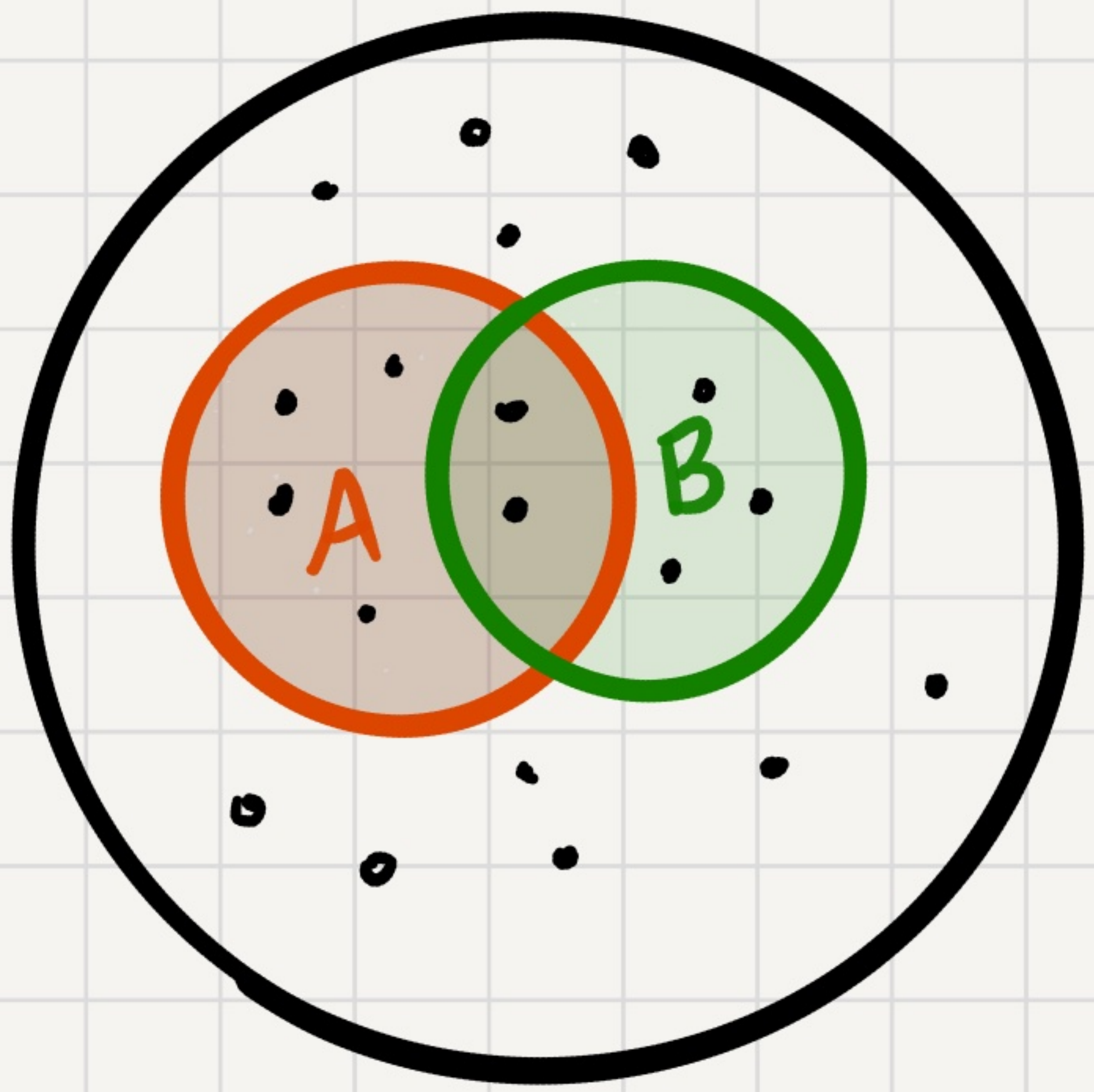
$$\Pr[A_1 \cup A_2 \cup \dots \cup A_n] = \Pr[A_1] + \dots + \Pr[A_n]$$

Proof: By induction on n ...

Inclusion-Exclusion

Theorem: If A, B are events then

$$\Pr[A \cup B] = \Pr[A] + \Pr[B] - \Pr[A \cap B]$$



Proof:

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

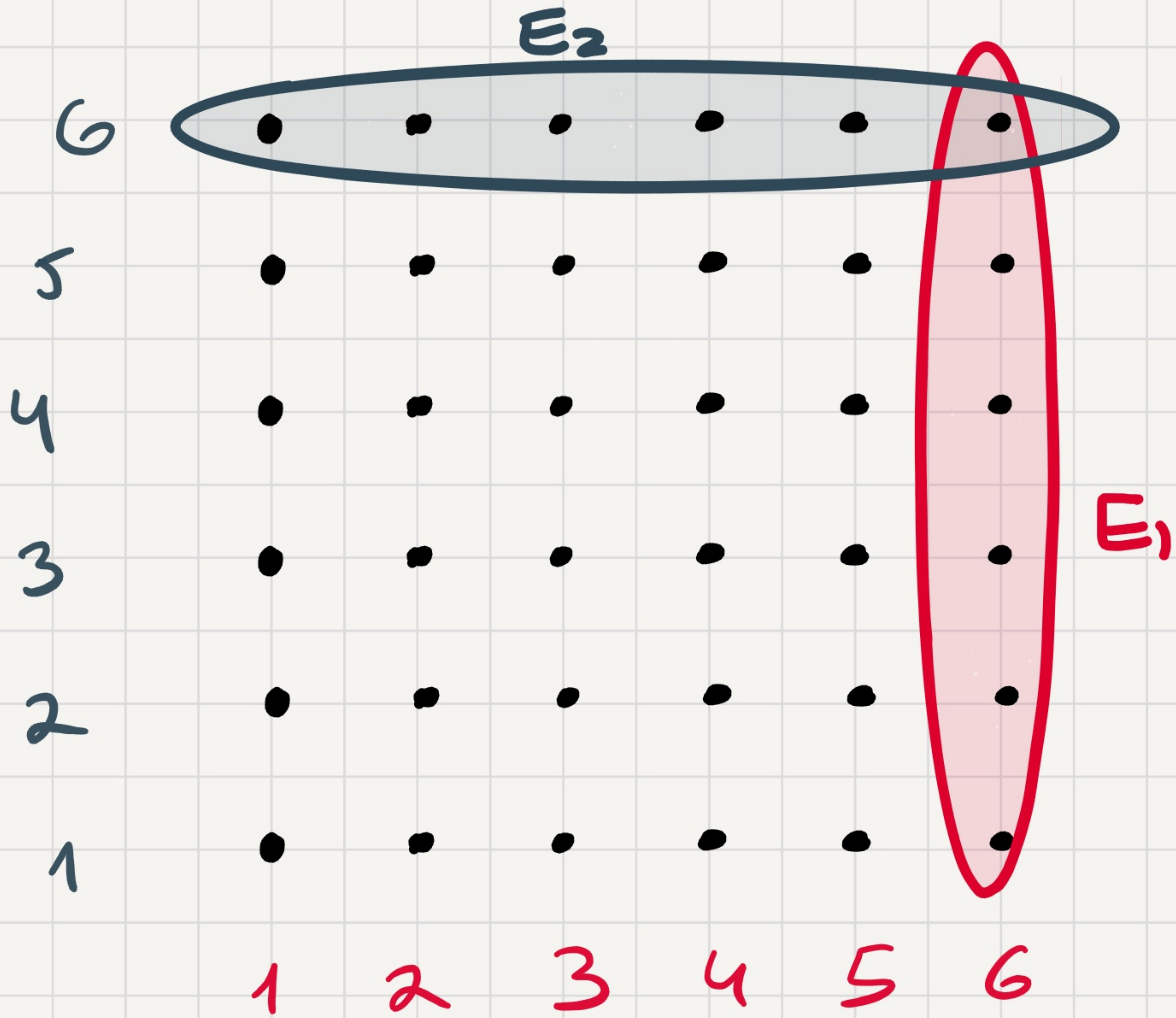
$$= \sum_{\omega \in A} \Pr[\omega] + \sum_{\omega \in B} \Pr[\omega] - \sum_{\omega \in A \cap B} \Pr[\omega]$$

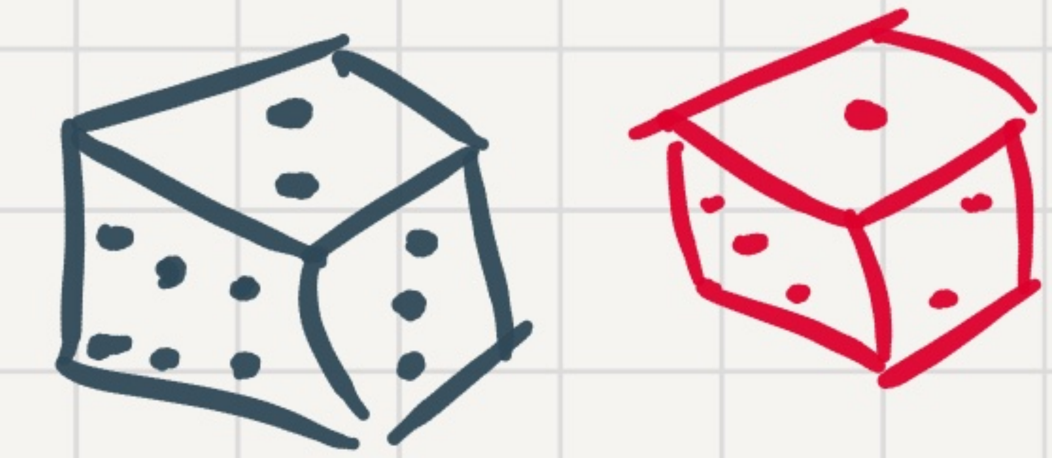
$$= \Pr[A] + \Pr[B] - \Pr[A \cap B]$$

Rolling two dice (Inclusion - Exclusion Example)

E_1 = 'red die shows 6'

E_2 = 'blue die shows 6'





Pr. of each outcome $= 1/36$.

$$\Pr[E_1 \cup E_2] = \frac{|E_1| + |E_2| - |E_1 \cap E_2|}{36} = \frac{11}{36}$$

Union Bound

If $A_1, \ldots, A_n$ are events, then

$$\Pr[A_1 \cup \ldots \cup A_n] \leq \Pr[A_1] + \ldots + \Pr[A_n]$$

Proof: $$\sum_{\omega \in A_1 \cup \ldots \cup A_n} \Pr[\omega] \leq \sum_{\omega \in A_1} \Pr[\omega] + \ldots + \sum_{\omega \in A_n} \Pr[\omega]$$

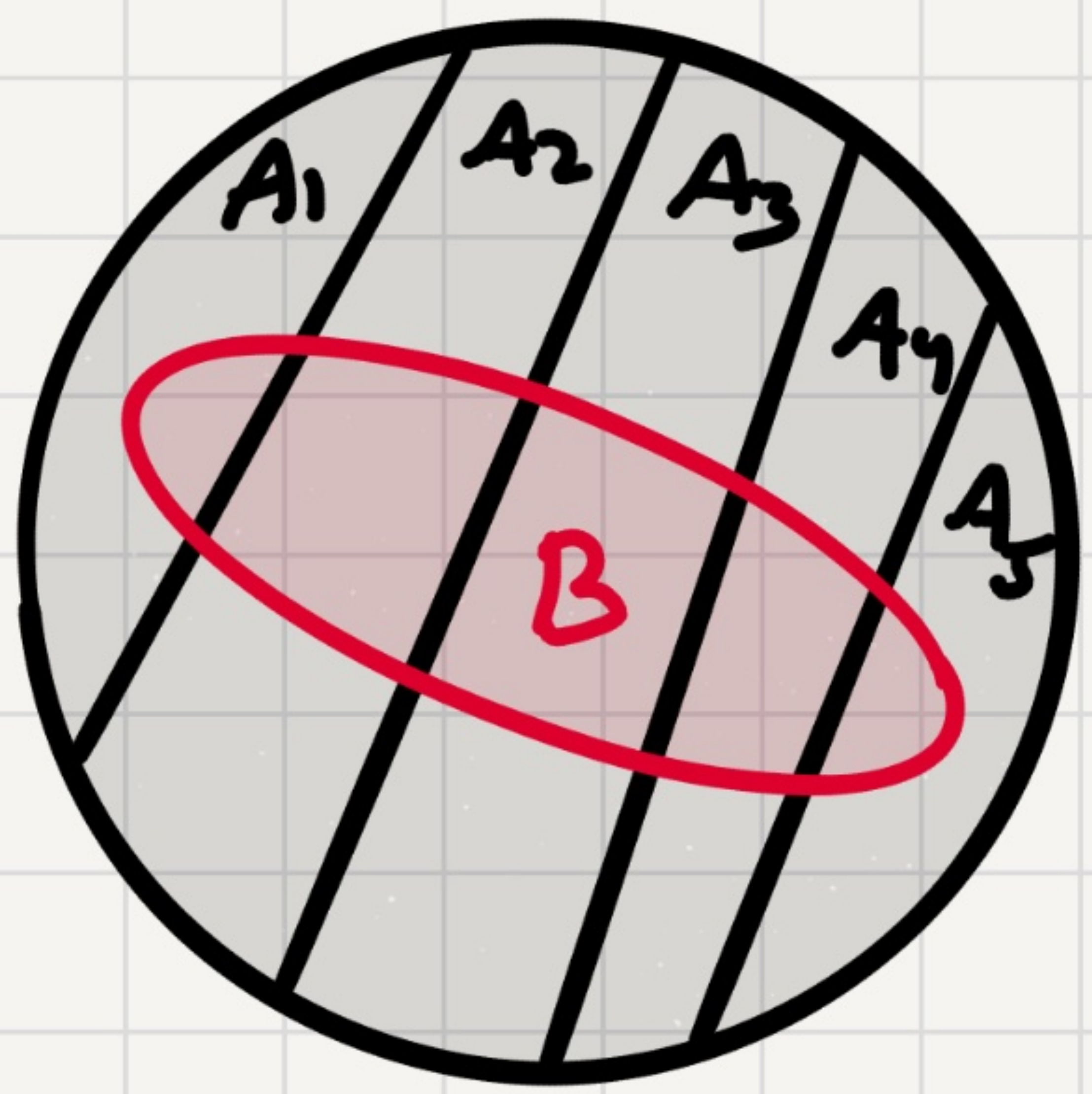
since every $\omega \in A_1 \cup \ldots \cup A_n$ is counted at least once in the right hand side.

Law of Total Probability

If $A_1, \dots, A_n$ are pairwise disjoint and

$$A_1 \cup A_2 \cup \dots \cup A_n = \Omega$$

then $$\Pr[B] = \Pr[B \cap A_1] + \dots + \Pr[B \cap A_n]$$

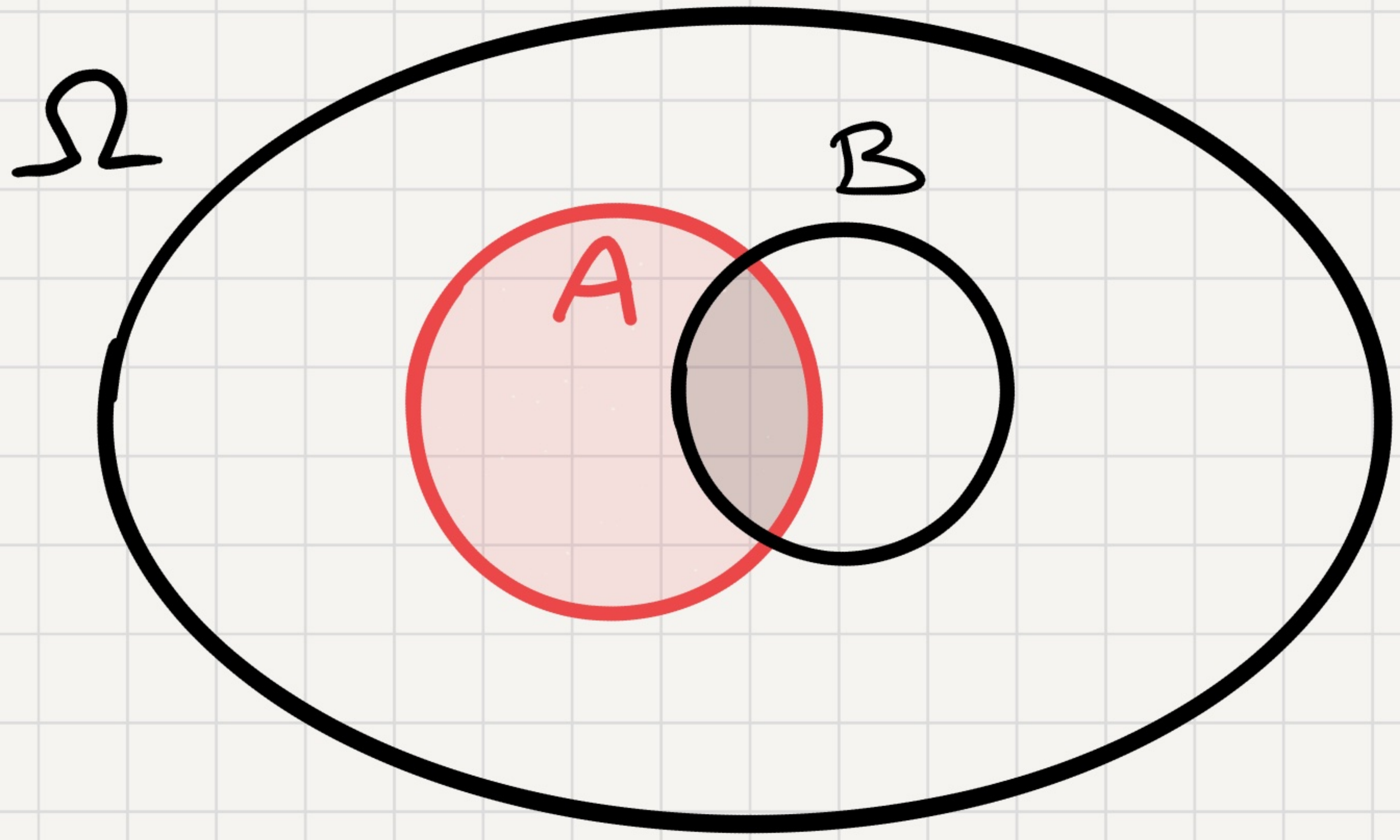


<u>Proof:</u> follows from additivity of probability since $B \cap A_1, \dots, B \cap A_n$ are pairwise disjoint.

Conditional Probability

<u>Def'n:</u> The conditional probability of event B given event A is

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



For Example: You draw a random card from a deck (with 52 cards)

I tell you that your card is a heart.

What's the prob. it is a 3?

$$\Pr[\text{"3"} \mid \text{"}\heartsuit\text{"}] = \frac{\Pr[\text{"3"} \cap \text{"}\heartsuit\text{"}]}{\Pr[\text{"}\heartsuit\text{"}]} =$$

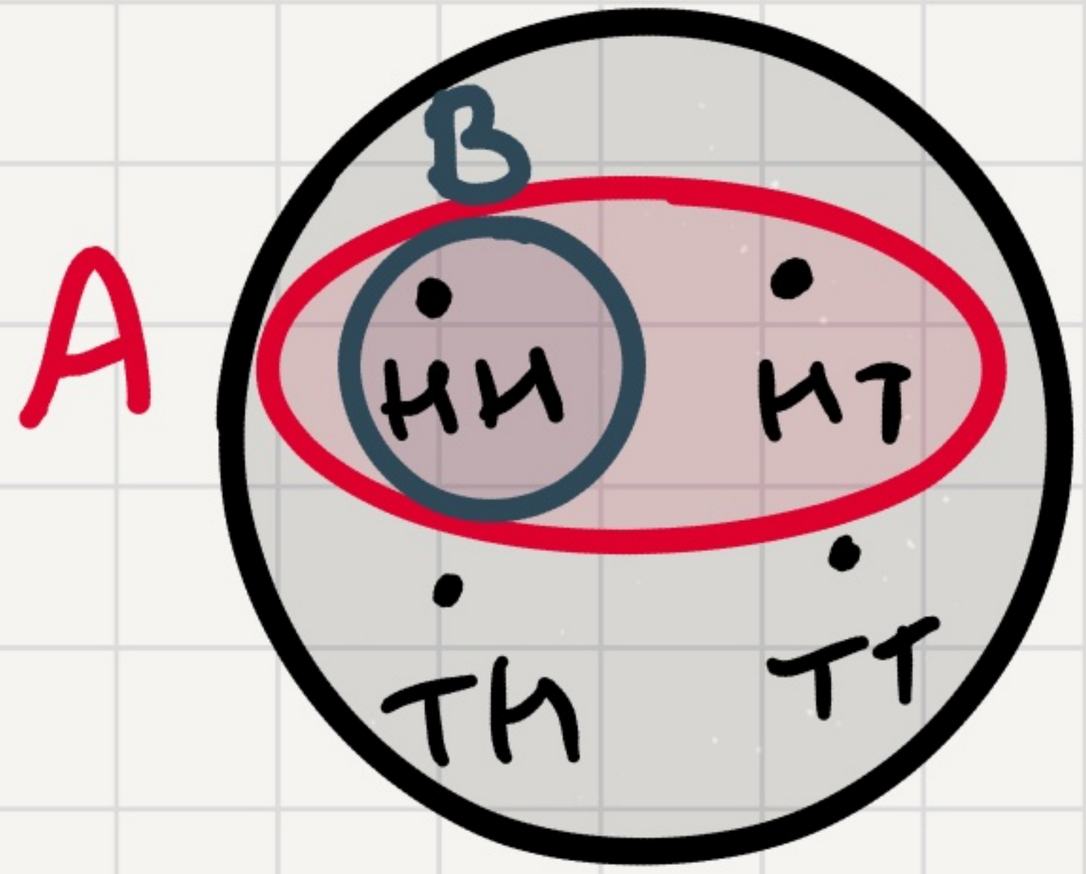
$$\Pr[\text{"3"}] =$$

Conditional Probability Example

Two fair coin flips. $\Omega = \{HH, HT, TH, TT\}$

A - first flip is heads

B - two heads



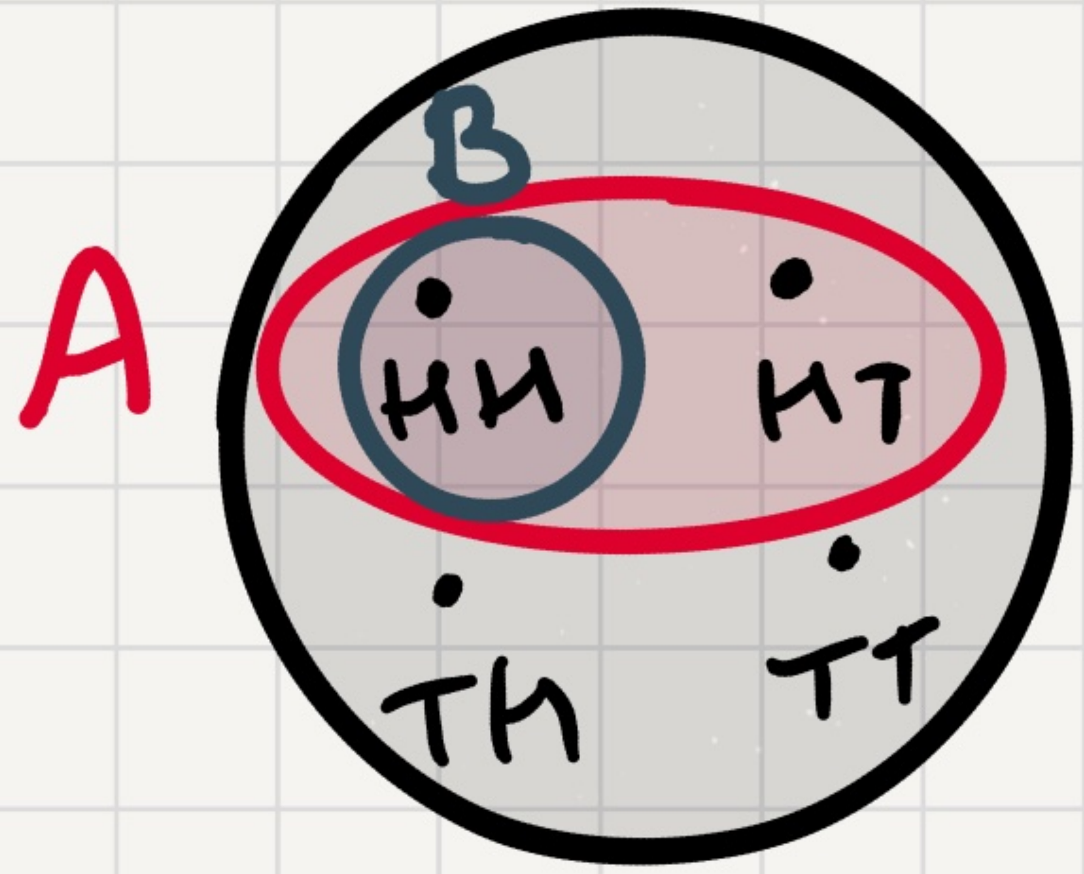
$$\Pr[B \mid A] = \frac{\Pr[A \cap B]}{\Pr[A]}$$

Alternative Way to think on Conditional Probabilities

Two fair coin flips. $\Omega = \{HH, HT, TH, TT\}$

A - first flip is heads

B - two heads



Given A, our probabilities of each outcome are updated

$$\forall \omega \in A \qquad Pr[\omega \mid A] = \frac{Pr[\omega]}{Pr[A]}$$

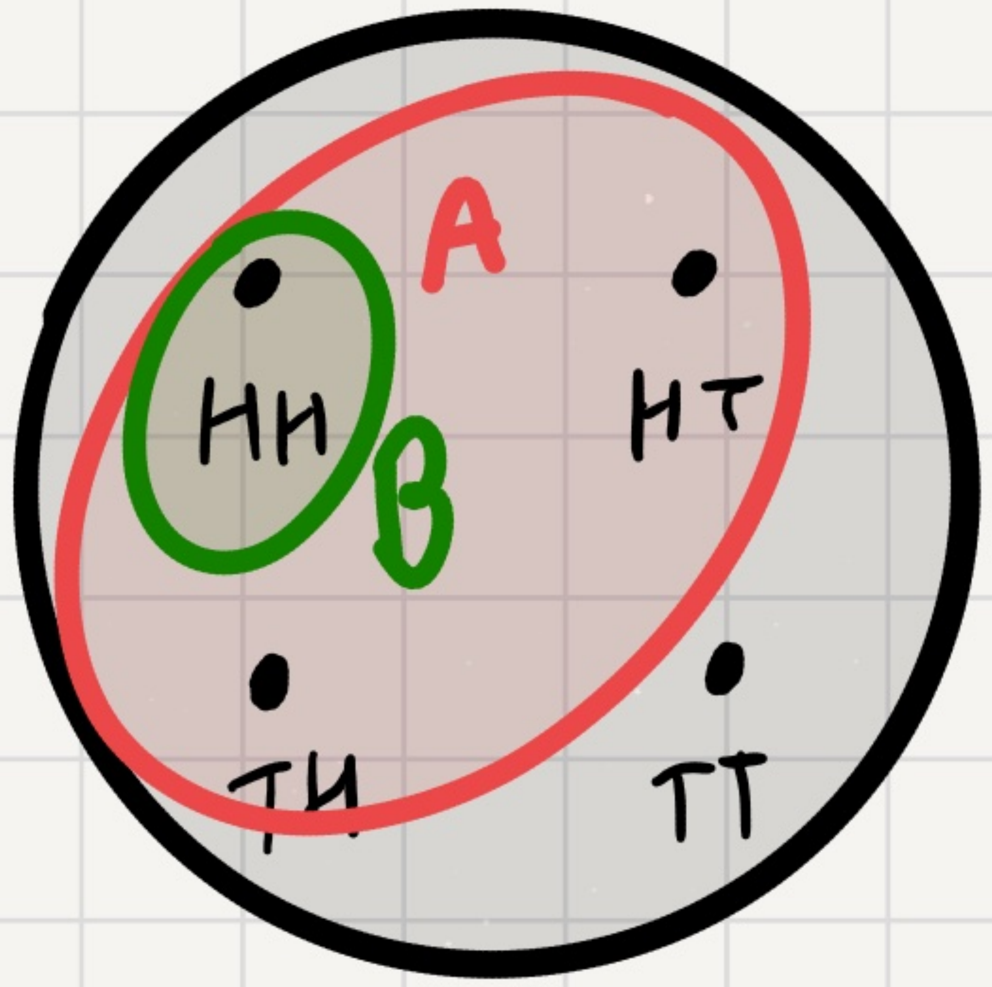
$$\forall \omega \notin A \qquad Pr[\omega \mid A] = 0.$$

Conditional Probability Example #2

Two fair coin flips. $\Omega = \{HH, HT, TH, TT\}$

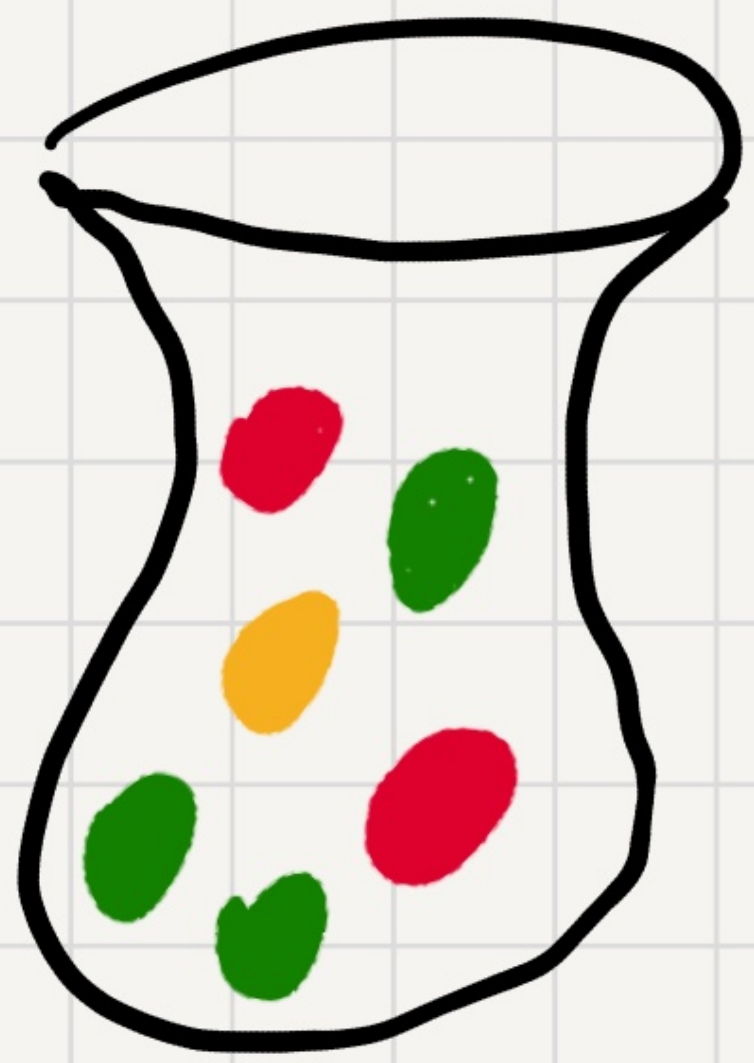
A - at least one head

B - two heads



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

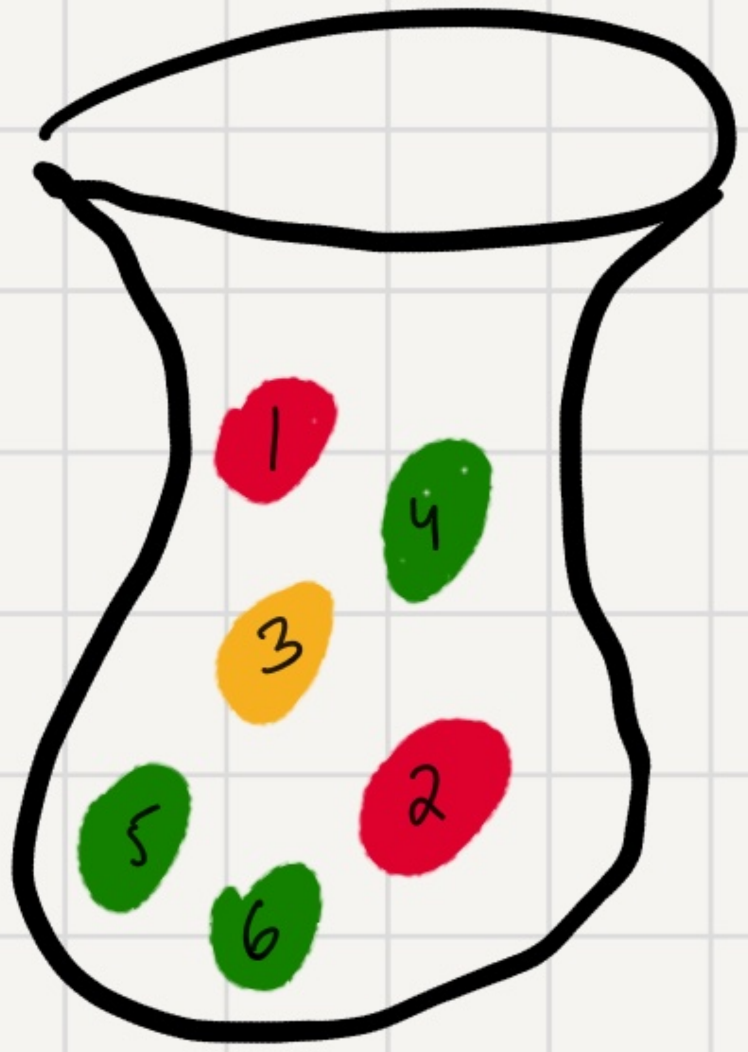
Bag of Marbles



$$\Pr\left[\text{"red marble"} \mid \text{"red or orange marble"}\right]$$

$$= \frac{2/6}{3/6}$$

Bag of Marbles



$$\Pr[\text{"red marble"} \mid \text{"red or orange marble"}]$$

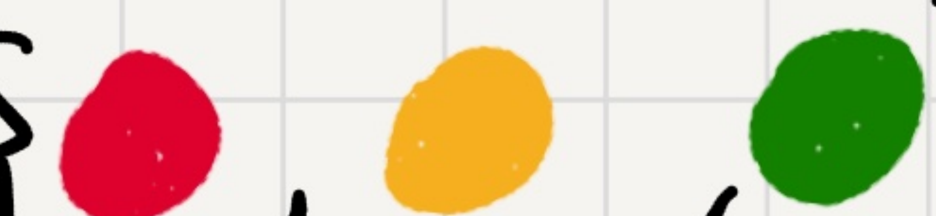
$$= \frac{2/6}{3/6}$$

Two ways to model this experiment:

1. Labeled marbles

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

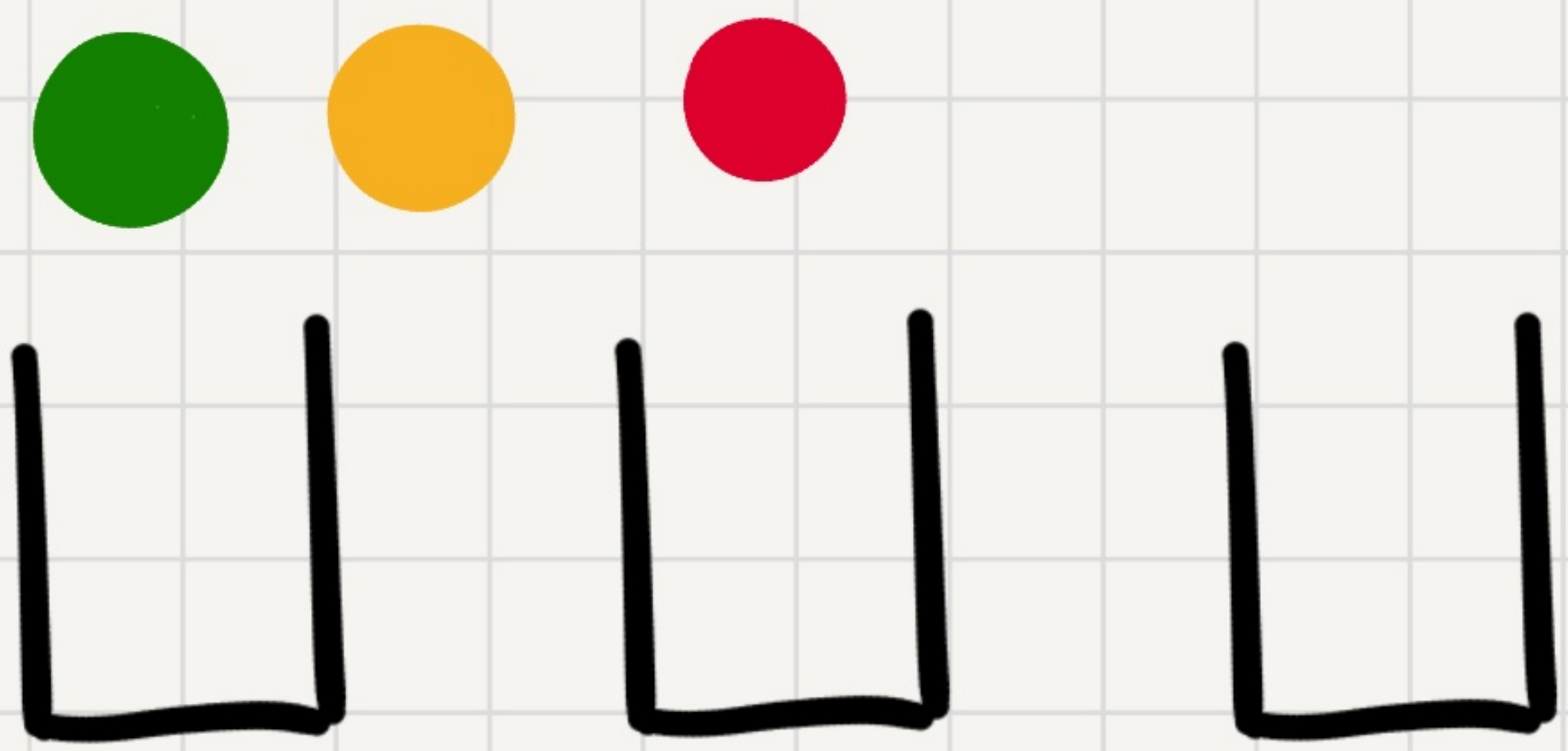
uniform prob. space.

2. Unlabeled marbles $\Omega = \{$  $\}$

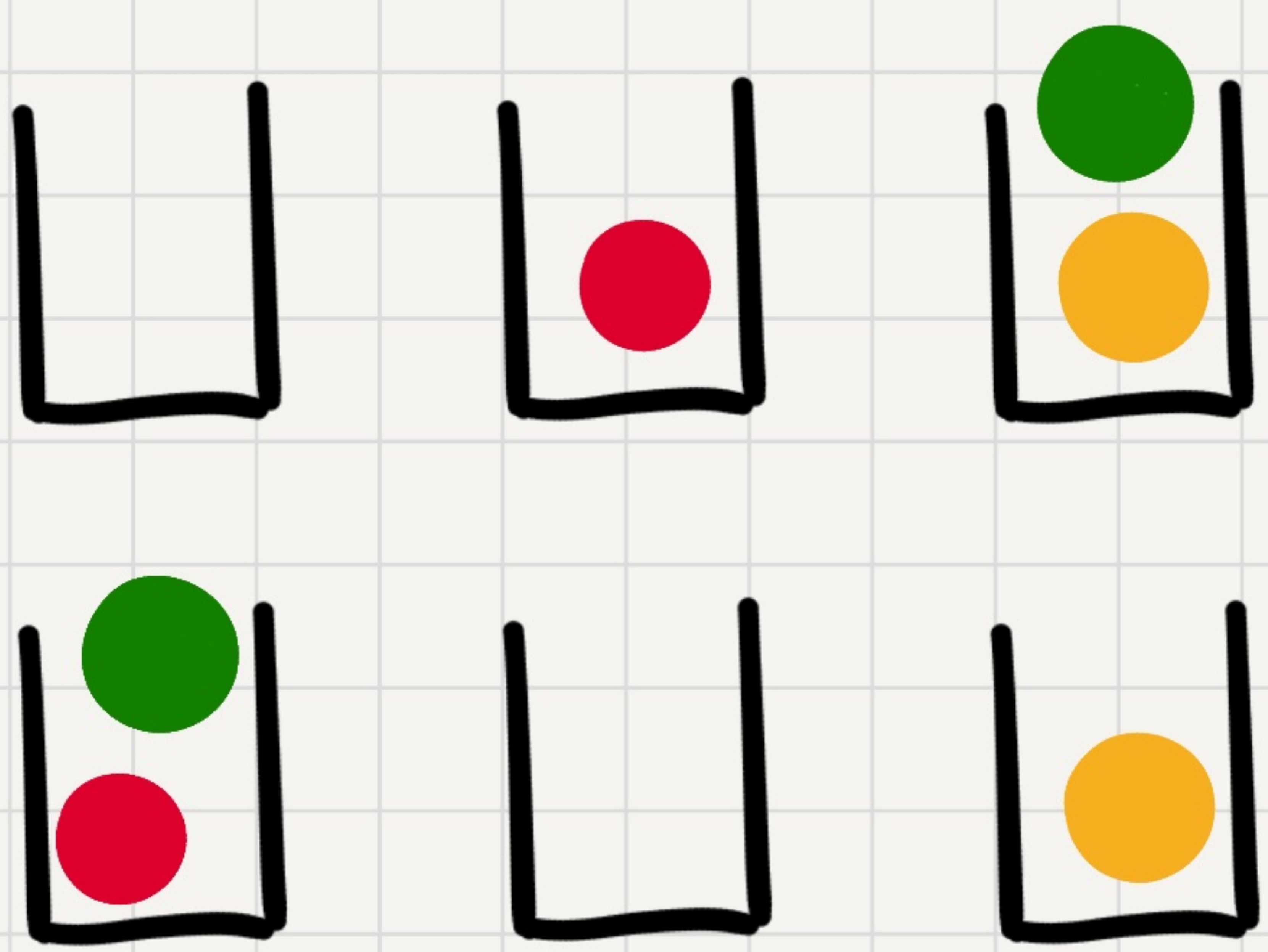
likelihoods $2/6$, $1/6$, $3/6$

Example #4: Balls in bins

Suppose I toss 3 balls in 3 bins uniformly at random



Possible outcomes ω = (bin of red ball, bin of orange ball, bin of green ball)



A = "bin 1 is empty"

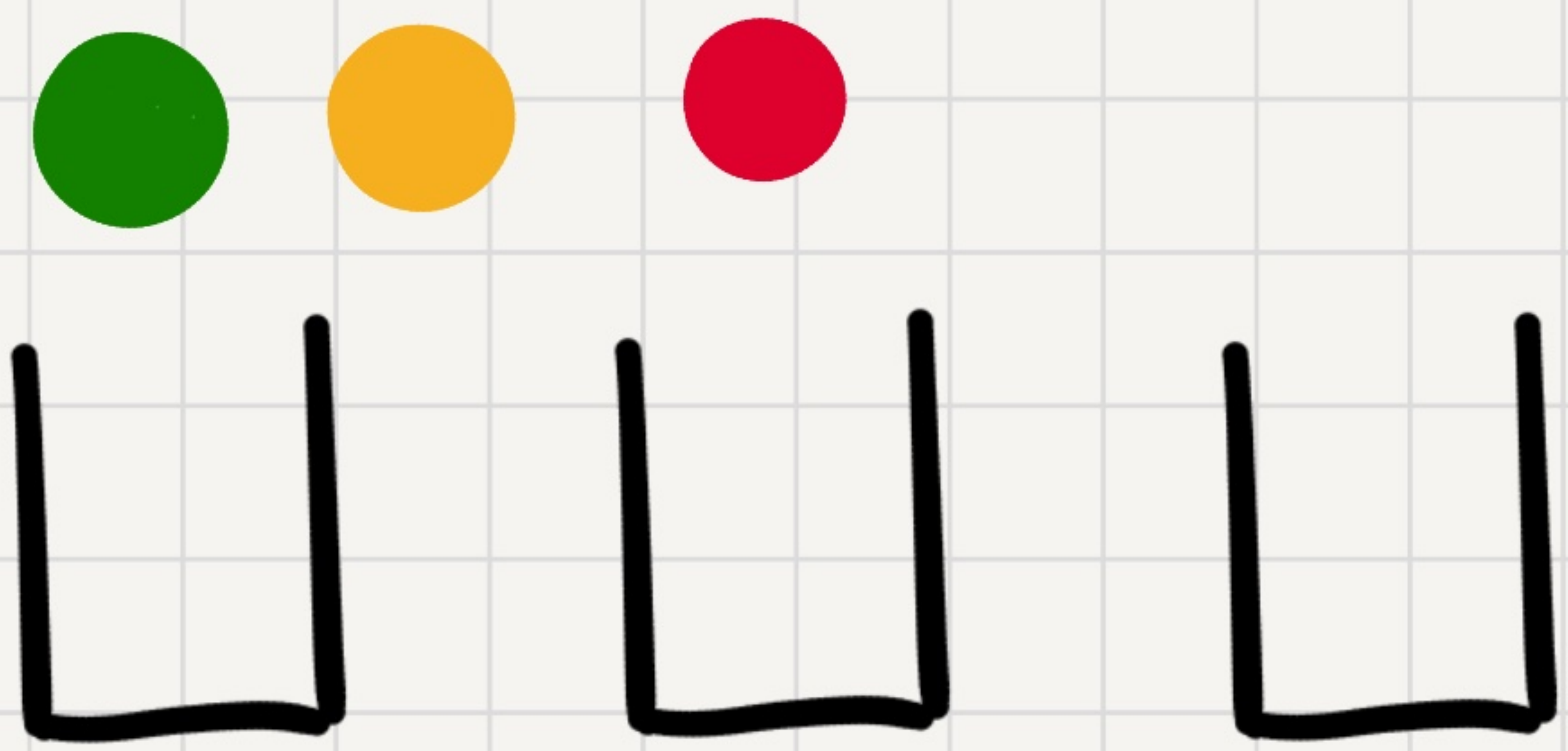
B = "bin 2 is empty"

Poll: which one is true

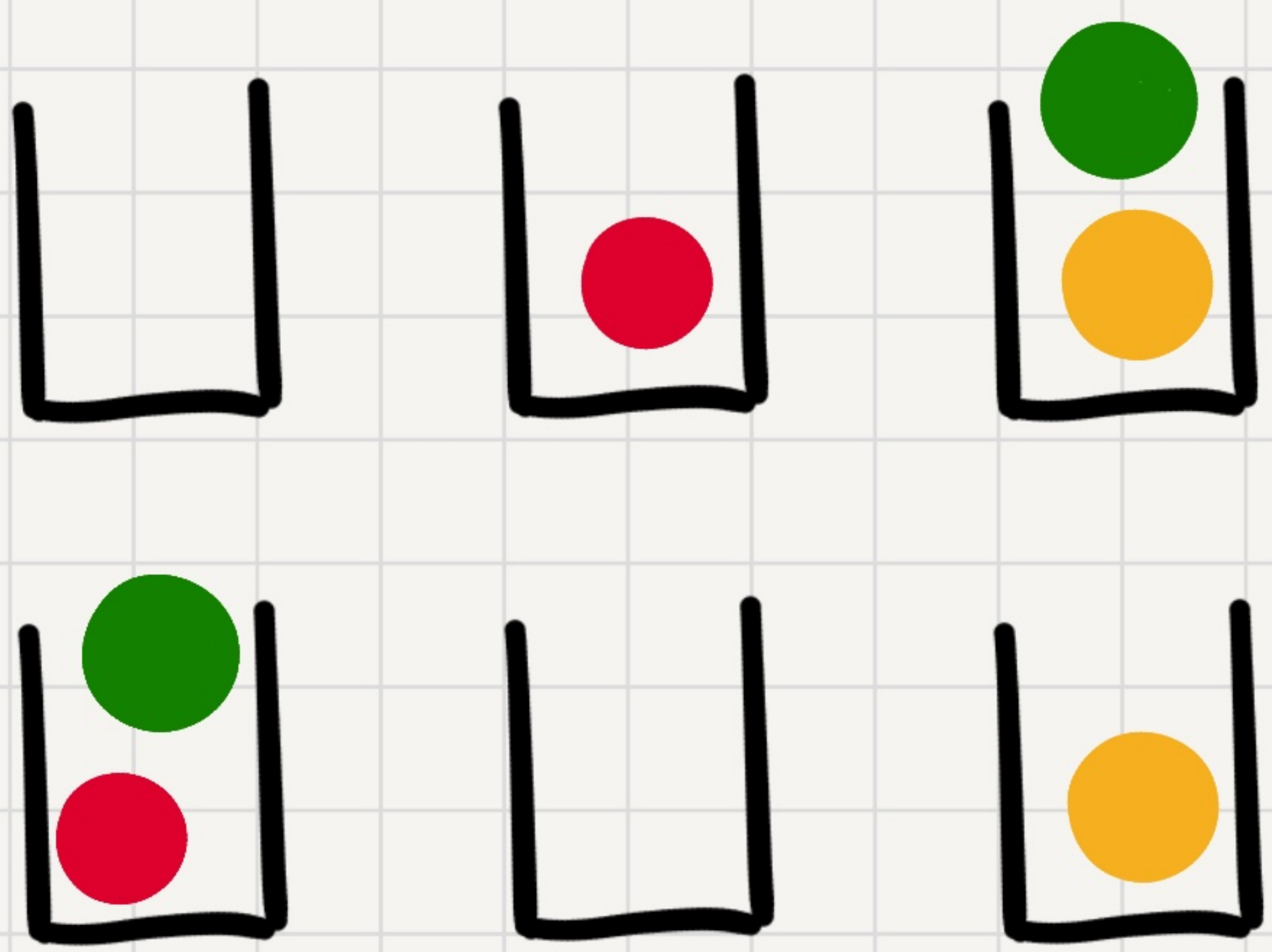
1. $\Pr[A \mid B] = \Pr[A]$
2. $\Pr[A \mid B] < \Pr[A]$
3. $\Pr[A \mid B] > \Pr[A]$

Example #4: Balls in bins

Suppose I toss 3 balls in 3 bins uniformly at random



Possible outcomes ω = (bin of red ball, bin of orange ball, bin of green ball)



A = "bin 1 is empty"

B = "bin 2 is empty"

$\Pr[A] =$

$\Pr[A|B] = \dfrac{\Pr[A \cap B]}{\Pr[B]} =$

Example #5: Coin Flips

Suppose I toss a fair coin 50 times.

$A =$ "the first 49 flips are H"

$B =$ "the last flip is H"

Poll: Is $\Pr[B \mid A]$

1. smaller than $1/2$
2. equals $1/2$
3. larger than $1/2$

Example #5: Coin Flips

Suppose I toss a fair coin 50 times.

$A =$ "the first 49 flips are H"

$B =$ "the last flip is H"

Poll: Is $\Pr[B \mid A]$

1. smaller than $1/2$
2. equals $1/2$
3. larger than $1/2$

$$\Pr[B \mid A] = \frac{\Pr[A \cap B]}{\Pr[A]} = \frac{1/2^{50}}{2/2^{50}} = 1/2$$

The Product Rule

Recall that by definition

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

$$\Rightarrow \quad \Pr[A \cap B] = \Pr[A] \cdot \Pr[B \mid A]$$

In words, the prob. of both A & B to happen is the prob. that A happens times the prob. that B happens given that A happened.

The Product Rule

Recall that by definition

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

$$\Rightarrow \quad \Pr[A \cap B] = \Pr[A] \cdot \Pr[B \mid A]$$

Generalize to more than two events

$$\Pr[A \cap B \cap C] = \Pr[A \cap B] \cdot \Pr[C \mid A \cap B]$$

$$= \Pr[A] \cdot \Pr[B \mid A] \cdot \Pr[C \mid A \cap B]$$

The Product Rule

Theorem: Let $A_1, \ldots, A_n$ be events. Then

$$\Pr[A_1 \cap \ldots \cap A_n] = \Pr[A_1] \cdot \Pr[A_2 | A_1] \cdots \Pr[A_n | A_1 \cap \ldots A_{n-1}]$$

Proof: By Induction on n.

Base case: $n=2$. We showed in previous slide.

General case. $n \geq 3$. Assume true for $n-1$ events.

$$\Pr[(A_1 \cap \ldots \cap A_{n-1}) \cap A_n] = \Pr[A_1 \cap \ldots \cap A_{n-1}] \cdot \Pr[A_n | A_1 \cap \ldots \cap A_{n-1}]$$

$$\underset{\text{ind. hypothesis}}{=} \Pr[A_1] \cdot \Pr[A_2 | A_1] \cdots \Pr[A_{n-1} | A_1 \cap \ldots \cap A_{n-2}] \cdot \Pr[A_n | A_1 \cap \ldots \cap A_{n-1}]$$

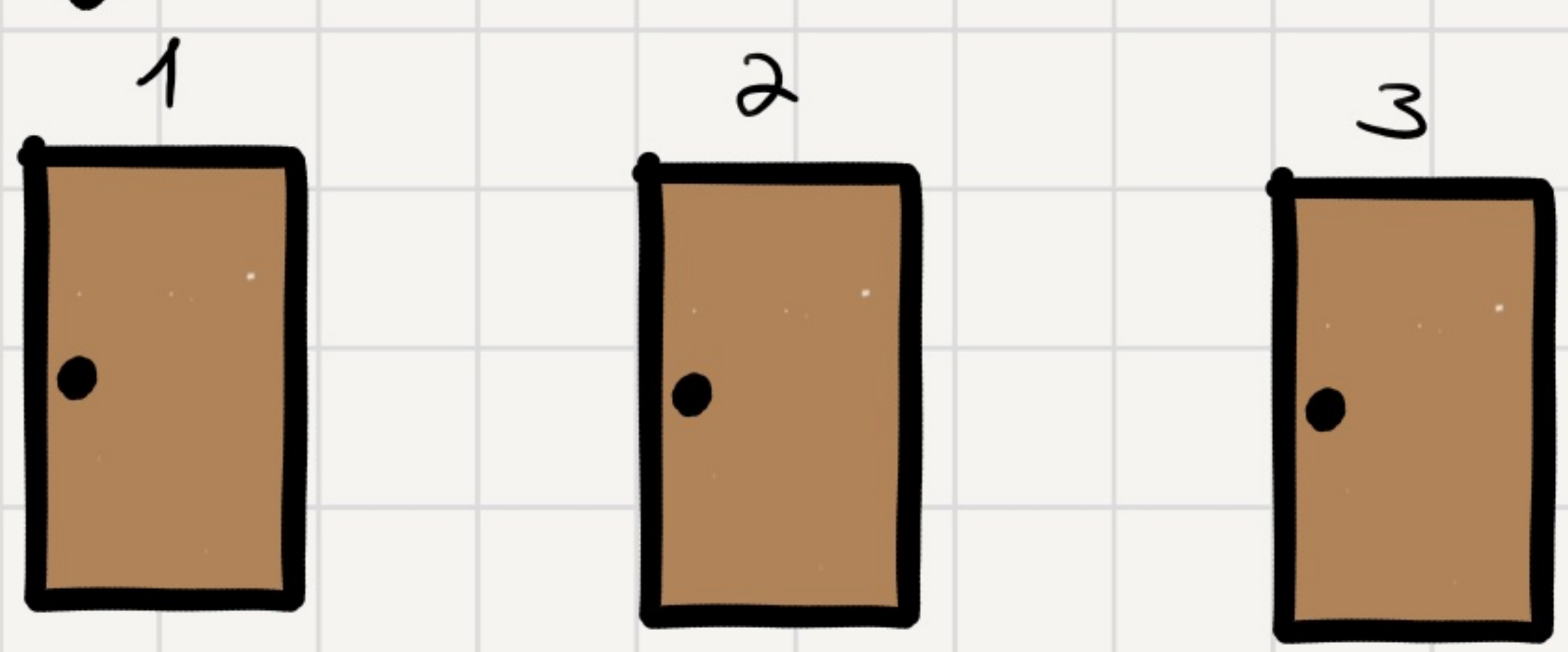
Random Experiment (Somewhat Confusing)

Monty Hall: Gameshow

3 doors

1 prize = car

2 goats



1. The host places the prize behind a randomly selected door.
2. You initially pick door # 1.
3. The host, who knows where the prize is, opens one of the doors $\{2, 3\}$ that has a goat.
4. The host offers you to switch doors. Should you?

$$\Omega = \{ (1,2), (1,3), (2,3), (3,2) \}$$

prize location

door opened by host

$$\Pr[(1,2)] = \Pr[\text{Prize at } 1] \cdot \Pr\left[\text{Host opened door \#2} \mid \text{Prize at } 1\right] = \tfrac{1}{3} \cdot \tfrac{1}{2}$$

$$\Pr[(1,3)] = \Pr[\text{Prize at } 1] \cdot \Pr\left[\text{Host opened door \#3} \mid \text{Prize at } 1\right] = \tfrac{1}{3} \cdot \tfrac{1}{2}$$

$$\Pr[(2,3)] = \Pr[\text{Prize at } 2] \cdot \Pr\left[\text{Host opened door \#3} \mid \text{Prize at } 2\right] = \tfrac{1}{3}$$

$$\Pr[(3,2)] = \Pr[\text{Prize at } 3] \cdot \Pr\left[\text{Host opened door \#2} \mid \text{Prize at } 3\right] = \tfrac{1}{3}$$

$$\Omega = \{ (1,2), (1,3), (2,3), (3,2) \}$$

prize location

door opened by host

$\Pr[(1,2)] = \frac{1}{6}$

$\Pr[(1,3)] = \frac{1}{6}$

$\Pr[(2,3)] = \frac{1}{3}$

$\Pr[(3,2)] = \frac{1}{3}$

Last time we showed that "always switching" strategy wins w.p. $\frac{2}{3}$

but maybe we'd like to sometimes switch and sometime stay conditioned on which door was opened.

$$\Omega = \{ \underset{\substack{\uparrow \\ \text{prize} \\ \text{location}}}{(1,} \underset{\substack{\nwarrow \\ \text{door opened} \\ \text{by host}}}{2)},\ (1,3),\ (2,3),\ (3,2) \}$$

$$\Pr[(1,2)] = \frac{1}{6}$$

$$\Pr[(1,3)] = \frac{1}{6}$$

$$\Pr[(2,3)] = \frac{1}{3}$$

$$\Pr[(3,2)] = \frac{1}{3}$$

$$\Pr\left[\text{Prize at } 2 \,\middle|\, \begin{array}{r}\text{Host opened} \\ \text{door \#3}\end{array}\right]$$

$$= \frac{\Pr[(2,3)]}{\Pr[(2,3)] + \Pr[(1,3)]} = \frac{1/3}{1/3 + 1/6}$$

$$= \frac{2}{3}$$

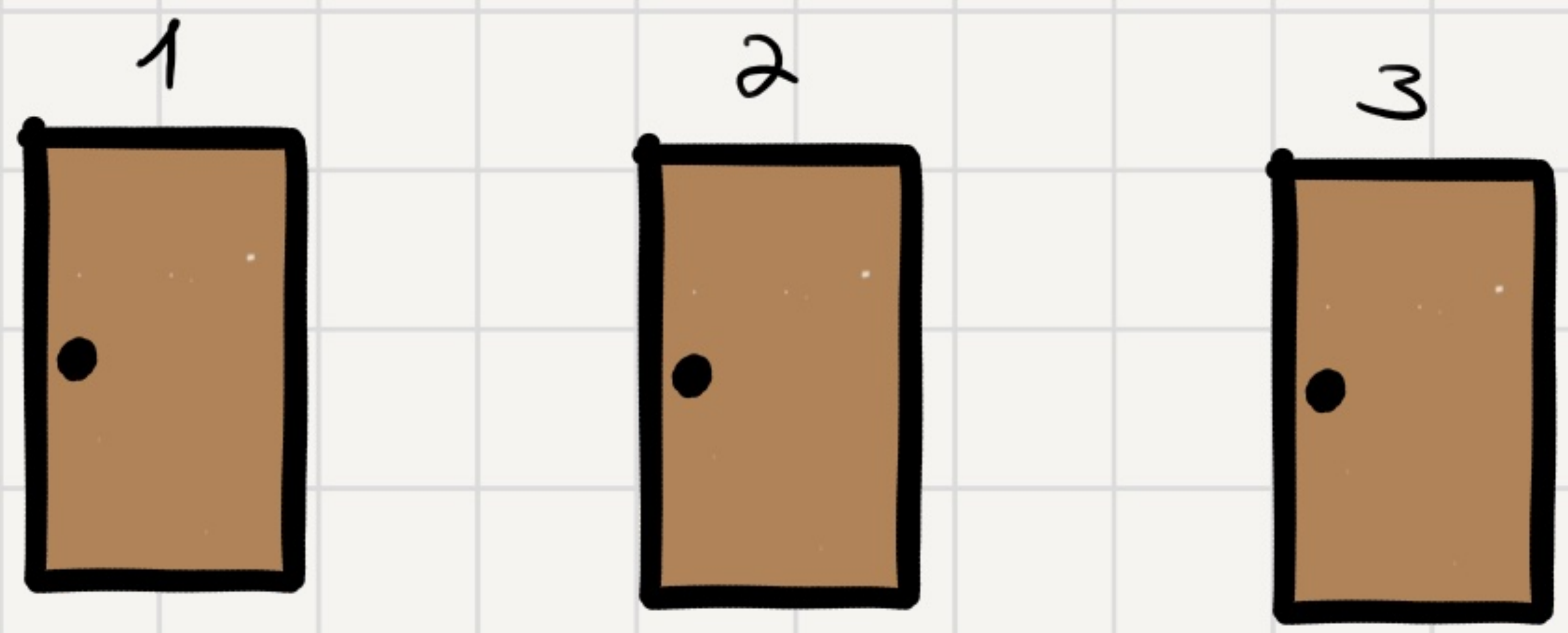
Random Experiment (Somewhat Confusing)

Monty Hall: Gameshow

3 doors

1 prize = car

2 goats



1. The host places the prize behind a randomly selected door.
2. You initially pick door # 1.
3. Carol, who knows where the prize is, opens one of the doors {2, 3} that has a goat.
4. The host offers you to switch doors. Should you?

We collude with Carol. If prize is behind door #1 Carol will open door #2

$$\Omega = \{ (1,2), (2,3), (3,2) \}$$

prize location

door opened by host

$$\Pr[(1,2)] = \frac{1}{3}$$

different

$$\Pr[(2,3)] = \frac{1}{3}$$

$$\Pr[(3,2)] = \frac{1}{3}$$

$$\Pr\left[\text{Prize at } 2 \,\middle|\, \begin{array}{c}\text{Host opened} \\ \text{door \#3}\end{array} \right]$$

$$=$$

$$\Pr\left[\text{Prize at } 3 \,\middle|\, \begin{array}{c}\text{Host opened} \\ \text{door \#2}\end{array} \right]$$

$$=$$

Correlation

An Example: Pick a random person in the US

Event A: the person has lung cancer

Event B: the person is a heavy smoker

$$\Pr[A \mid B] = 1.17 \cdot \Pr[A]$$

What can you conclude?

- Heavy smoking increases prob. of lung cancer by 17%
- Smoking causes lung cancer.

Correlation

An Example: Pick a random person in the US

Event A: the person has lung cancer

Event B: the person is a heavy smoker

$$\Pr[A|B] = 1.17 \cdot \Pr[A]$$

$$\Pr[B|A] = \frac{\Pr[A \cap B]}{\Pr[A]} = \frac{\Pr[B] \cdot \Pr[A|B]}{\Pr[A]} = 1.17 \cdot \Pr[B]$$

Conclusions:
- Lung cancer increases prob. of smoking by 17%.
- Lung cancer causes smoking.

Correlation vs. Causation

Events A and B are positively correlated if

$$\Pr[A \cap B] > \Pr[A] \cdot \Pr[B]$$

A and B being positively correlated doesn't mean that A causes B or that B causes A.

Independence

Definition: We say that events A and B are **independent** if

$$\Pr[A \cap B] = \Pr[A] \cdot \Pr[B].$$

or equivalently $\Pr[B \mid A] = \Pr[B]$

Independence

Definition: We say that events A and B are **independent** if

$$\Pr[A \cap B] = \Pr[A] \cdot \Pr[B].$$

or equivalently $\Pr[B|A] = \Pr[B]$

Poll: Which of the following are independent?

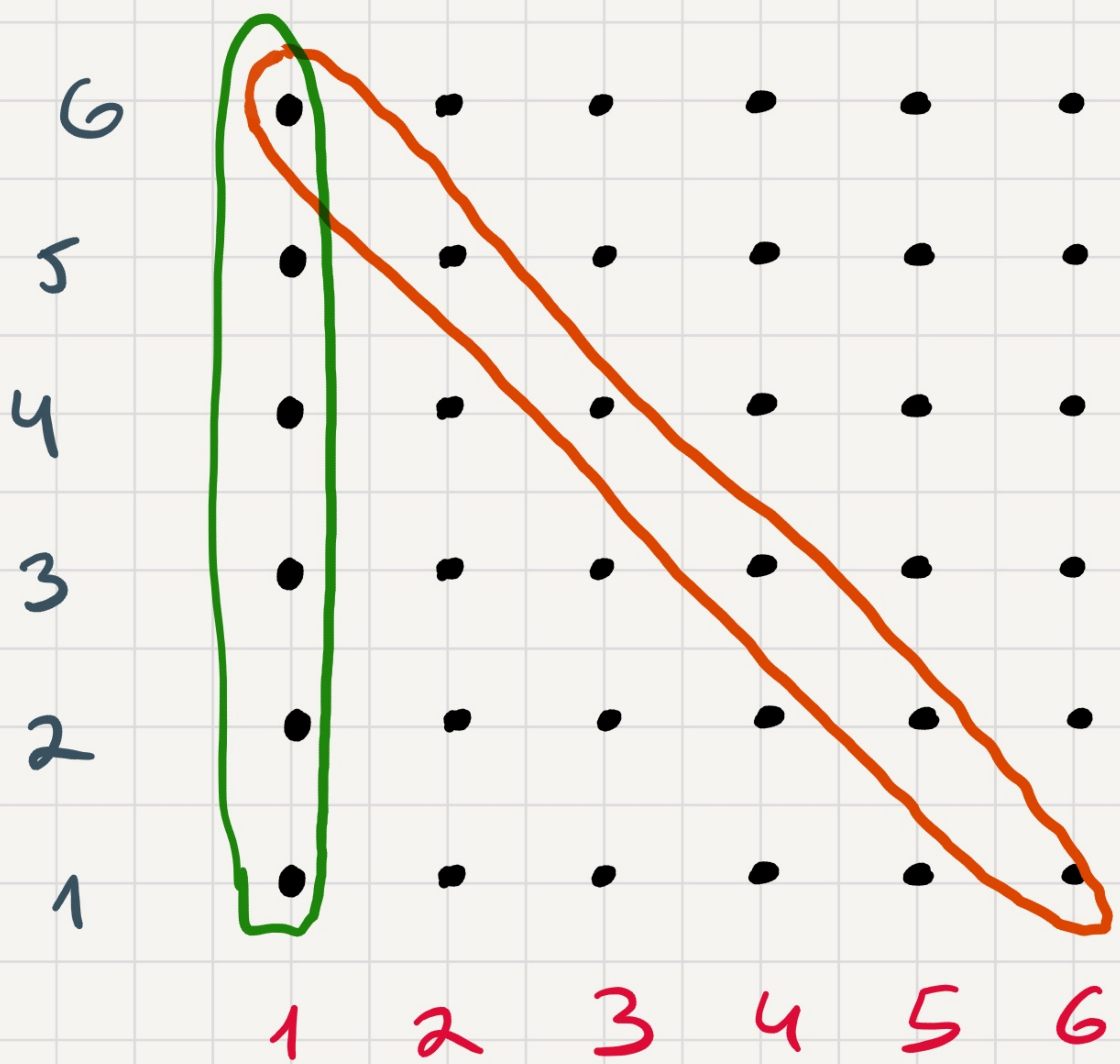
1. Flip two fair coins. A = "first coin is H" B= "second coin is H"

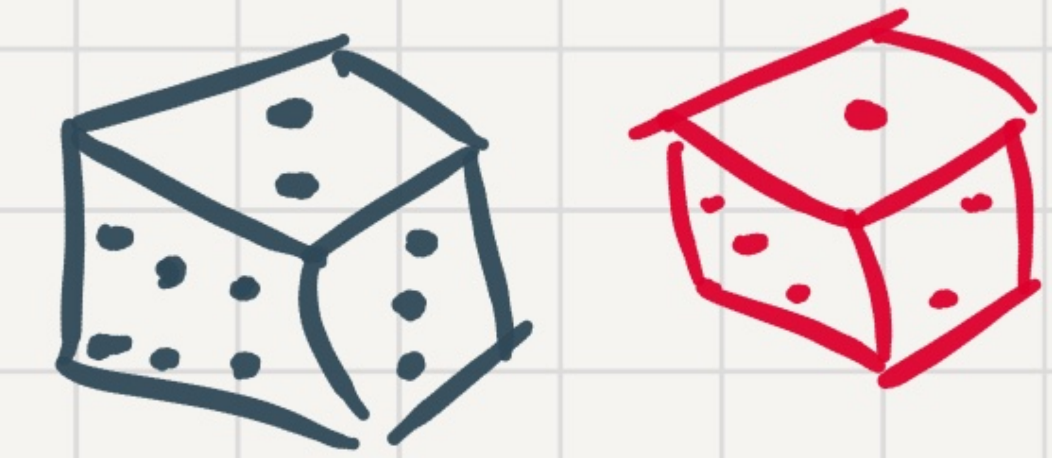
2. Flip two dice. A="sum is 7". B= "first die is 1"

3. Flip two dice A = "sum is 2". B= "first die is 1"

Rolling two dice

A = "sum is 7" B = " red die is 1"





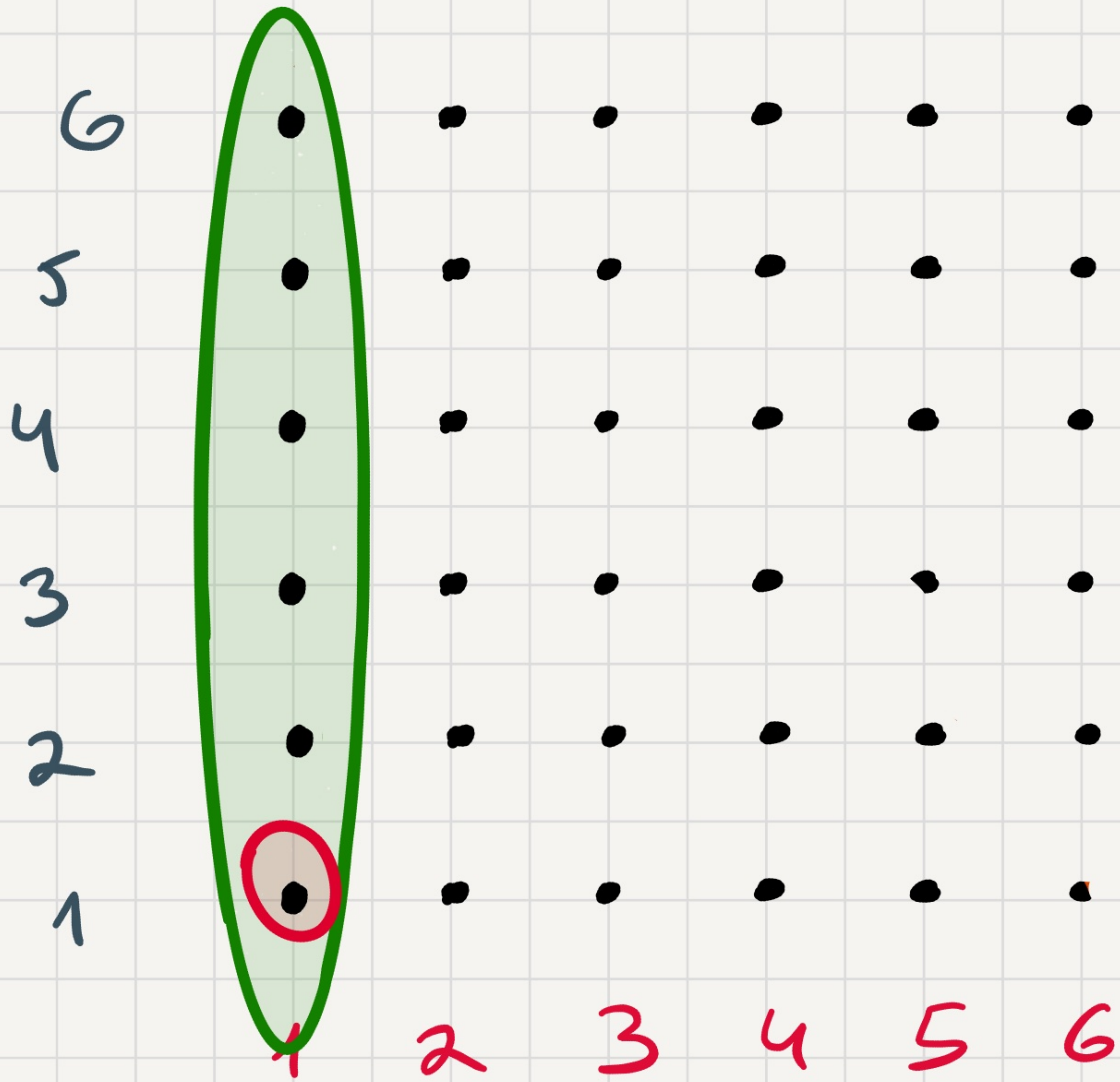
Pr. of each outcome $= 1/36$.

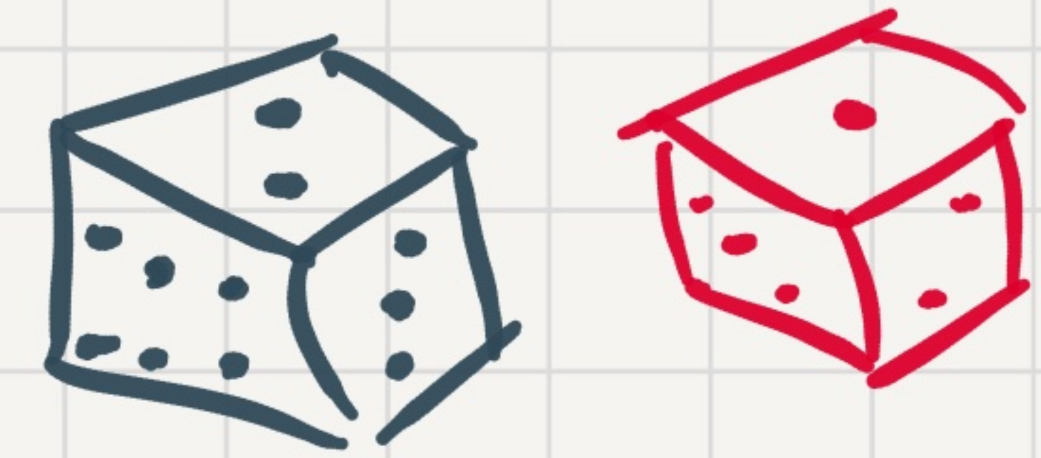
$$\Pr[A \cap B] = 1/36 = \Pr[A] \cdot \Pr[B]$$

independent!

Rolling two dice

A = "sum is 2" B = "red die is 1"





Pr. of each outcome = 1/36.

$$\Pr[A\cap B] = 1/36 > \frac{1}{36}\cdot\frac{1}{6} = \Pr[A]\cdot\Pr[B]$$

not independent

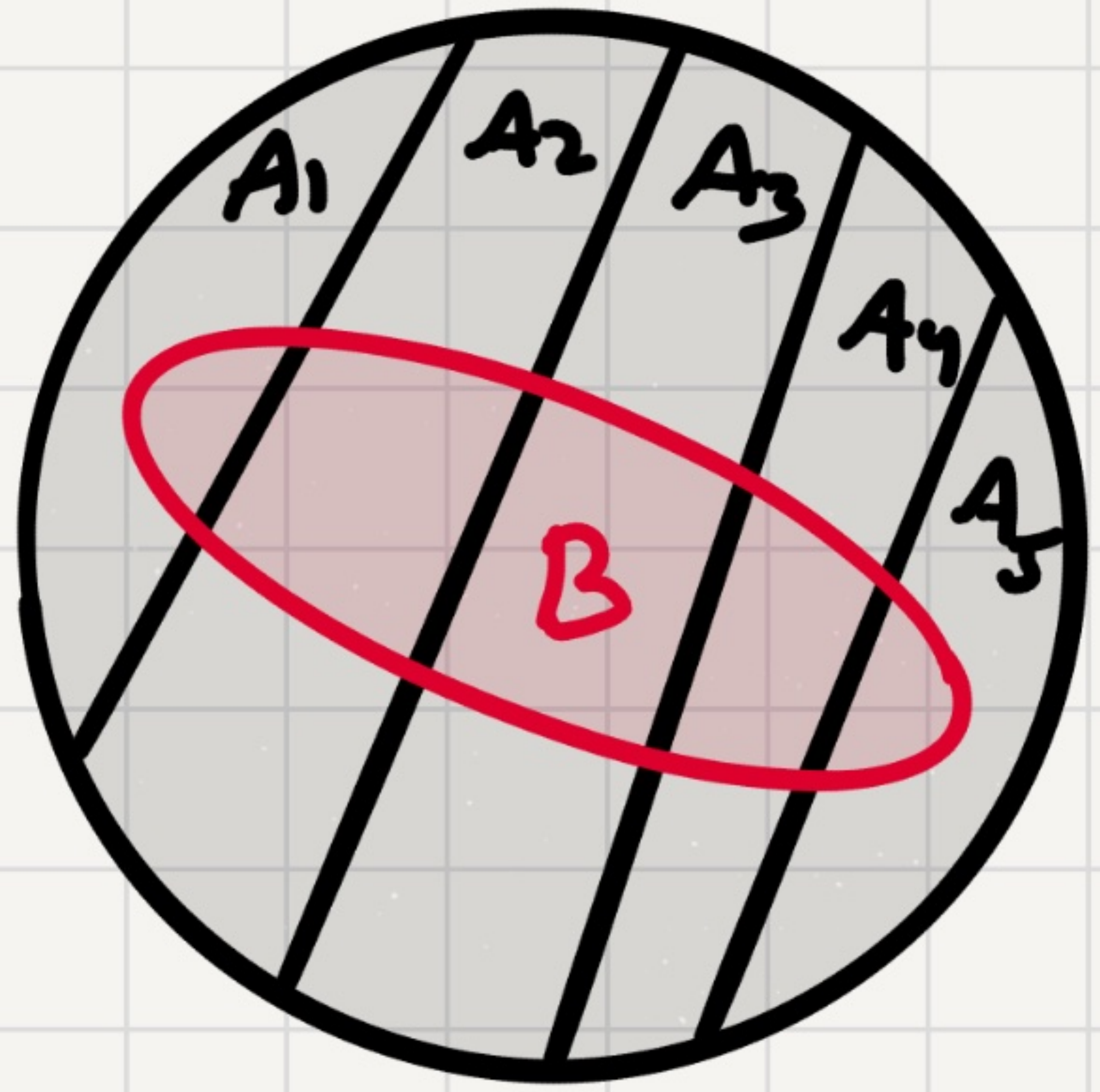
Recall

Law of Total Probability

If $A_1, \ldots, A_n$ are pairwise disjoint and

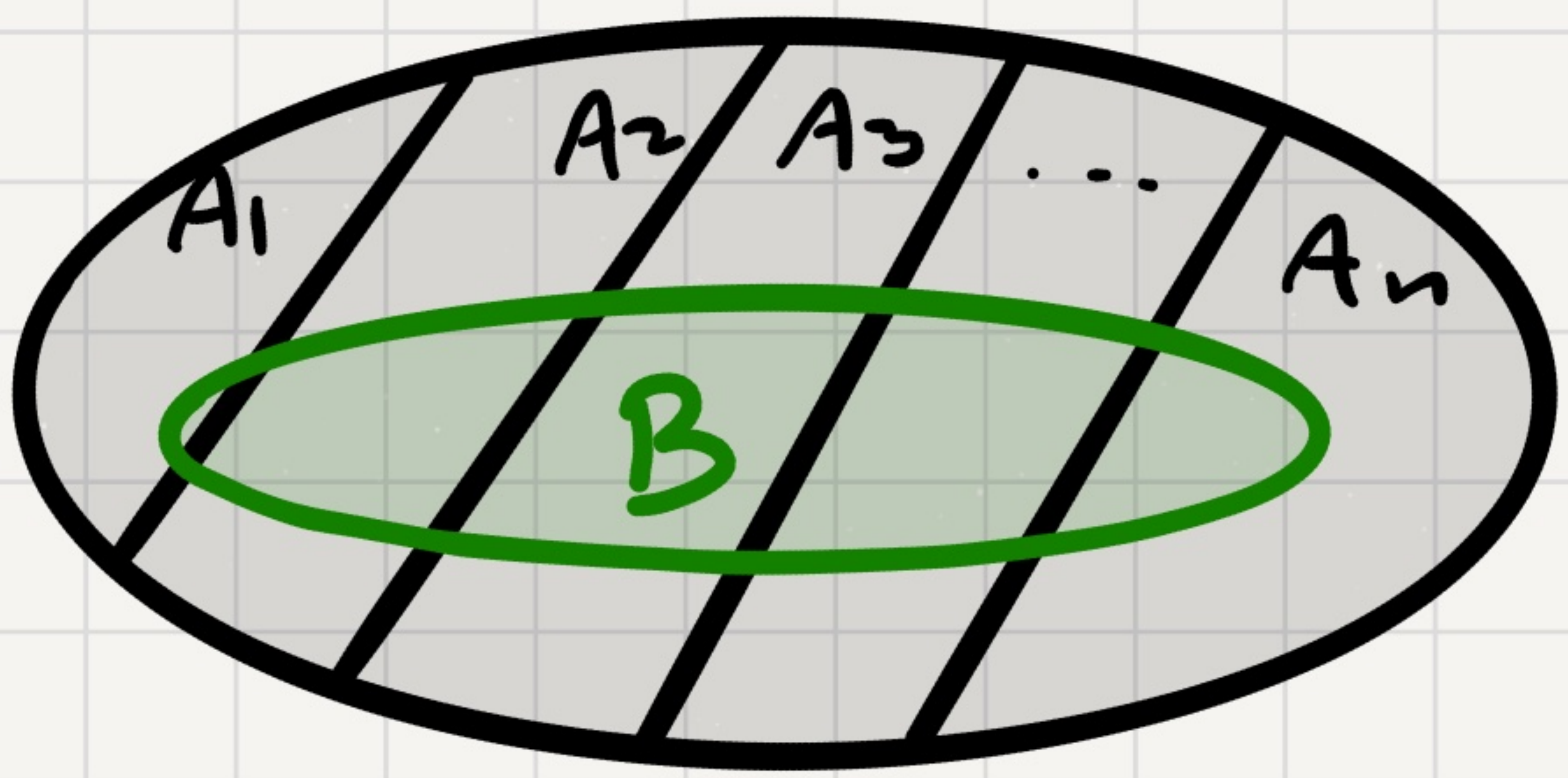
$$A_1 \cup A_2 \cup \ldots \cup A_n = \Omega$$

then $$\Pr[B] = \Pr[B \cap A_1] + \ldots + \Pr[B \cap A_n]$$



Total Probability with Conditional Probability

Assume Ω is a union of disjoint events $A_1, \dots, A_n$



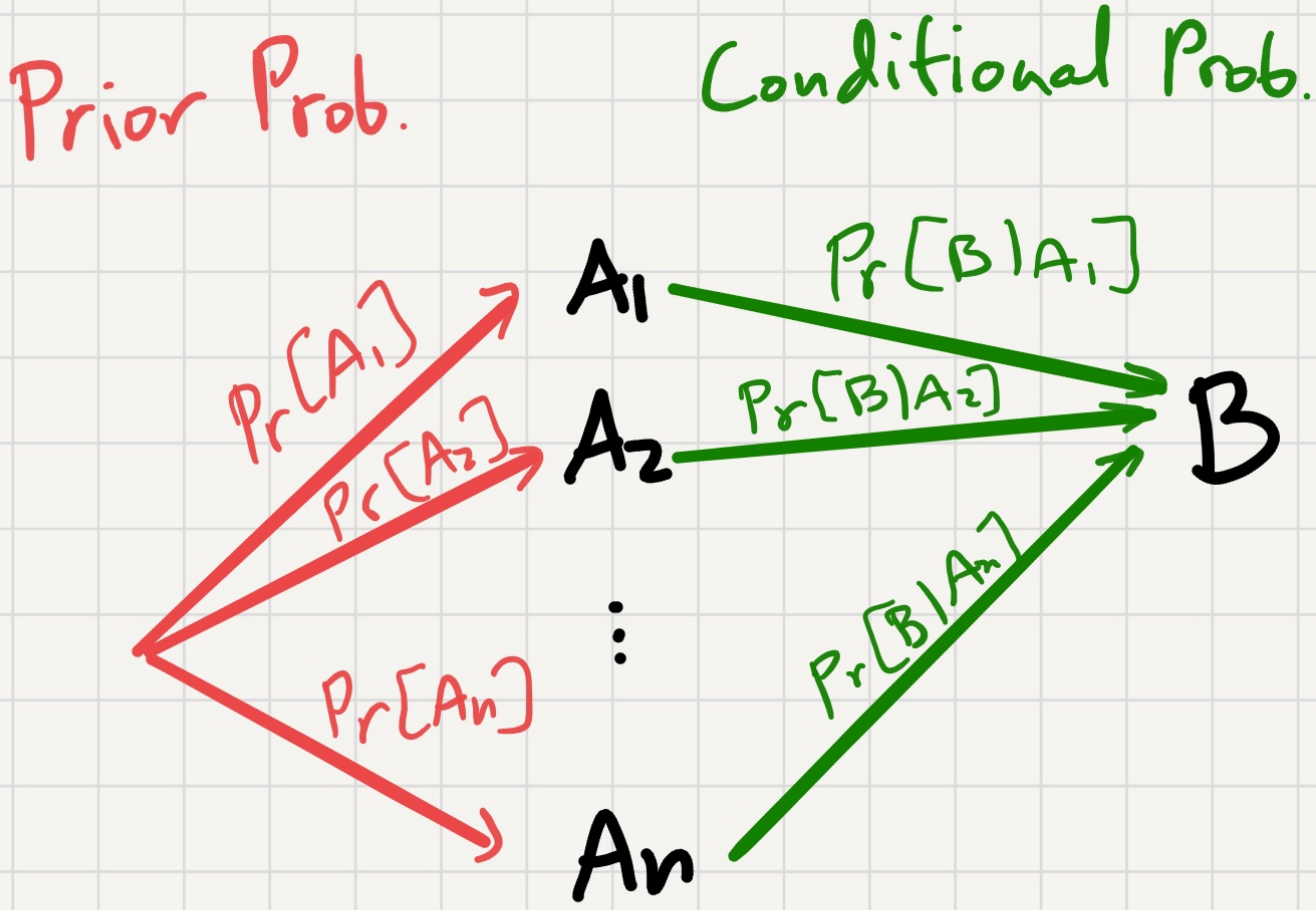
Since B is the disjoint union of $B \cap A_1, \dots, B \cap A_n$

$$\Pr[B] = \Pr[B \cap A_1] + \dots + \Pr[B \cap A_n]$$

Thus, by the product rule

$$\boxed{\Pr[B] = \Pr[A_1] \cdot \Pr[B \mid A_1] + \dots + \Pr[A_n] \cdot \Pr[B \mid A_n]}$$

Total Probability Rule with Conditional Probability



$$\boxed{Pr[B] = Pr[A_1] \cdot Pr[B|A_1] + \ldots + Pr[A_n] \cdot Pr[B|A_n]}$$

Bayes Rule

Suppose you know $\Pr[B|A]$, $\Pr[A]$, $\Pr[B]$

What's $\Pr[A|B]$?

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

Bayes Rule Example #1

Experiment:

1. Pick at random either a fair coin or a biased coin with 60% chance heads.
2. Toss the coin you picked

$\Omega = \{(\text{fair}, H), (\text{fair}, T), (\text{biased}, H), (\text{biased}, T)\}$

A = "coin is fair" B = "got H"

WTK: $\Pr[A|B]$

Bayes Rule Example #1

Experiment:

1. Pick at random either a fair coin or a biased coin with 60% chance heads.
2. Toss the coin you picked

$\Omega = \{ (\text{fair}, H), (\text{fair}, T), (\text{biased}, H), (\text{biased}, T) \}$

$A =$ "coin is fair" $\qquad B =$ "got H"

WTK: $\Pr[A \mid B] = \dfrac{\Pr[A \cap B]}{\Pr[B]} = \dfrac{\Pr[A] \cdot \Pr[B \mid A]}{\Pr[B]} = \dfrac{0.5 \times 0.5}{\Pr[B]}$

What's $\Pr[B]$?

Bayes Rule Example #1

Experiment:

1. Pick at random either a fair coin or a biased coin with 60% chance heads.
2. Toss the coin you picked

$$\Omega = \{ (\text{fair}, H), (\text{fair}, T), (\text{biased}, H), (\text{biased}, T) \}$$

$A =$ "coin is fair" $\qquad B =$ "got H"

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

$$\begin{aligned} \Pr[B] &= \Pr[A \cap B] + \Pr[\bar{A} \cap B] \\ &= \Pr[A]\Pr[B \mid A] + \Pr[\bar{A}] \cdot \Pr[B \mid \bar{A}] \end{aligned}$$

Bayes Rule Example #1

Experiment:

1. Pick at random either a fair coin or a biased coin with 60% chance heads.
2. Toss the coin you picked

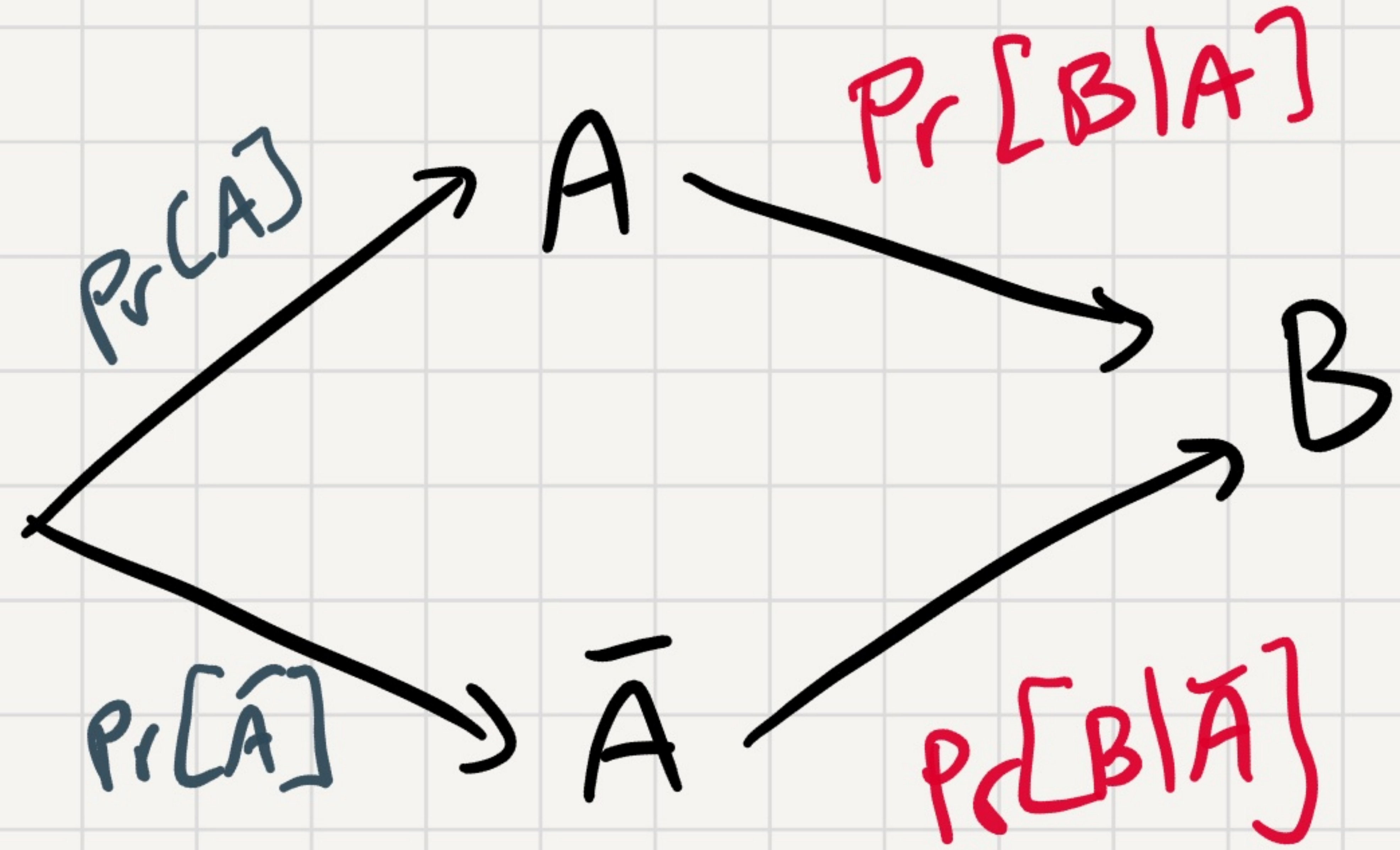
$\Omega = \{ (\text{fair}, H), (\text{fair}, T), (\text{biased}, H), (\text{biased}, T) \}$

$A =$ "coin is fair" $\qquad B =$ "got H"

WTK: $\Pr[A \mid B] = \dfrac{\Pr[A \cap B]}{\Pr[B]} = \dfrac{\Pr[A] \cdot \Pr[B \mid A]}{\Pr[B]} = \dfrac{0.25}{\Pr[B]}$

$\Pr[B] = \Pr[A] \Pr[B \mid A] + \Pr[\bar{A}] \cdot \Pr[B \mid \bar{A}]$

$= 0.5 \times 0.5 \quad + \quad 0.5 \times 0.6 \qquad = 0.55$



$$\Pr[B] = \Pr[A]\cdot\Pr[B|A] + \Pr[\bar{A}]\cdot\Pr[B|\bar{A}]$$

Bayes Rule (updated)

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

$$= \frac{\Pr[A]\cdot\Pr[B|A]}{\Pr[A]\cdot\Pr[B|A] + \Pr[\bar{A}]\cdot\Pr[B|\bar{A}]}$$

Bayes Rule Example #2

Suppose there's a disease that occur in 0.001 of the population.

There's a test for the disease.

For a random person: $\Pr[\text{test positive} \mid \text{sick}] = 0.99$

$\Pr[\text{test positive} \mid \text{not sick}] = 0.01$

A random person arrives and tests positive

Q: what's the likelihood that he has the disease?

$\Pr[\text{sick}] = 0.001$

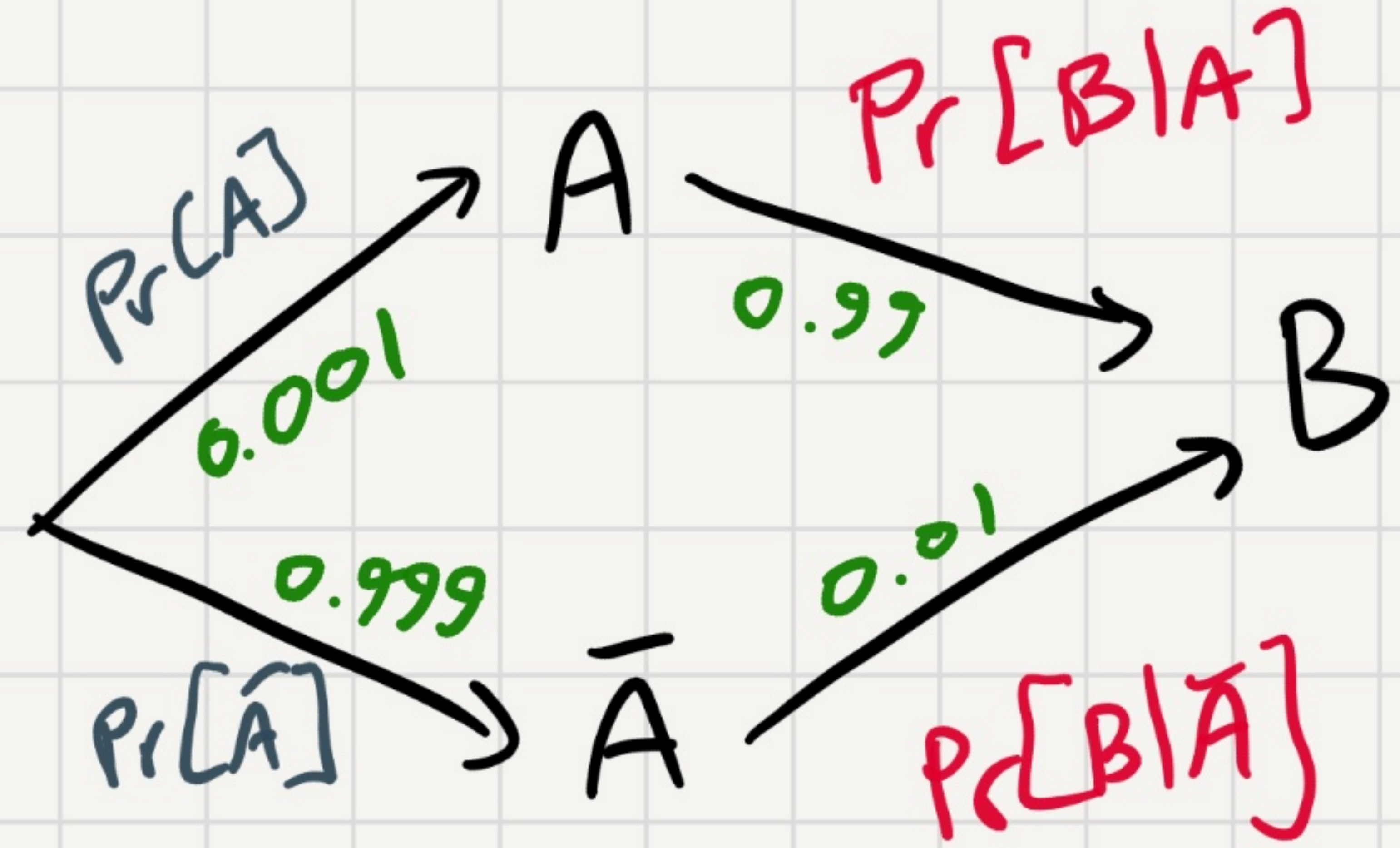
$\Pr[\text{positive} \mid \text{sick}] = 0.99$

$\Pr[\text{positive} \mid \text{not sick}] = 0.01$

$A =$ "sick"

$B =$ "tested positive"

WTK: $\Pr[\text{sick} \mid \text{positive}]$



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

$$= \frac{\Pr[A] \cdot \Pr[B \mid A]}{\Pr[A] \cdot \Pr[B \mid A] + \Pr[\bar{A}] \Pr[B \mid \bar{A}]} = \frac{0.001 \times 0.99}{0.001 \times 0.99 + 0.999 \times 0.01}$$

$$\approx 0.09$$

Summary

- Probability is additive
- Union Bound $\Pr[A_1 \cup \ldots \cup A_n] \le \Pr[A_1] + \ldots + \Pr[A_n]$
- Inclusion-Exclusion $\Pr[A \cup B] = \Pr[A] + \Pr[B] - \Pr[A \cap B]$
- Total probability: if $A_1, \ldots, A_n$ partition Ω

  then $\Pr[B] = \Pr[A_1 \cap B] + \ldots + \Pr[A_n \cap B]$
- Conditional probability: $\Pr[B \mid A] = \frac{\Pr[A \cap B]}{\Pr[A]}$
- Independence: $\Pr[A \cap B] = \Pr[A] \cdot \Pr[B]$
- Correlation $\Pr[A \cap B] > \Pr[A] \cdot \Pr[B]$
- Bayes Rule.