Chapter 2
Probability on Events

Sample Space and Events

Probability is defined in terms of some experiment.

Ω = Sample space of the experiment = Set of all possible outcomes

<u>Defn</u>: An **event**, E, is any subset of the sample space, Ω.

<u>Example</u>: Roll die twice

Q: What does event E_1 represent?
Q: What is $E_1 \cup E_2$?
Q: What is $\overline{E_1}$?
Q: Are E_1 and E_2 independent? (we'll see)

$\Omega =$

	E_1			E_2	
(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

Sample Space and Events

<u>Defn</u>: If $E_1 \cap E_2 = \emptyset$, then E_1 and E_2 are **mutually exclusive**.

<u>Defn</u>: If $E_1, E_2, \ldots, E_n$ are events such that $E_i \cap E_j = \emptyset, \quad \forall i \neq j$, and such that $\bigcup_{i=1}^{n} E_i = F$ then we say that events $E_1, E_2, \ldots, E_n$ **partition** set F.

Q: What is an example of events that partition Ω for 2 rolls of a die?

Sample Space and Events

Defn: A sample space is **discrete** if the number of outcomes is:
_____countable_______.
A sample space is **continuous** if the number of outcomes is:
_____uncountable_____.

Q: Which of these experiments have a discrete/continuous sample space?

- ❑ Roll a die 2 times discrete
- ❑ Throw a dart at a unit interval. continuous
- ❑ Flip a coin until we see the first head. discrete
- ❑ Mark the time when the 100th email arrives. continuous

Probability Defined on Events

$\boldsymbol{P}\{E\}$ = probability of event E
= probability that the outcome of the experiment lies in set E

The 3 Probability Axioms:

Non-negativity: $\boldsymbol{P}\{E\} \geq 0$ for any event E.

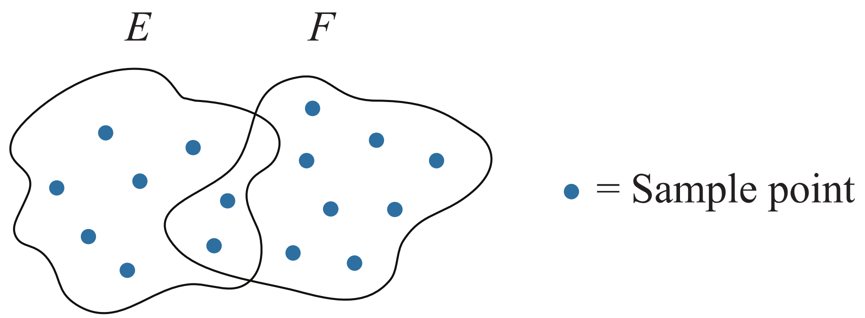
Additivity: If $E_1, E_2, E_3, \ldots$ is a countable sequence of disjoint events, then

$$\boldsymbol{P}\{E_1 \cup E_2 \cup E_3 \cup \cdots\} = \boldsymbol{P}\{E_1\} + \boldsymbol{P}\{E_2\} + \boldsymbol{P}\{E_3\} + \cdots$$

Normalization: $\boldsymbol{P}\{\Omega\} = 1$

Consequences of the 3 Probability Axioms

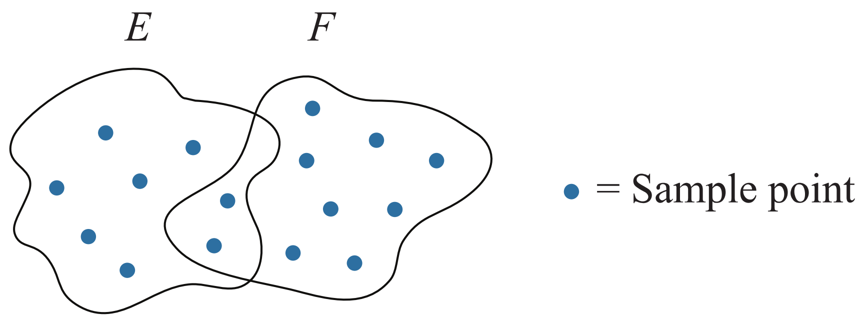
<u>Lemma 2.5</u>: $\boldsymbol{P}\{E \cup F\} = \boldsymbol{P}\{E\} + \boldsymbol{P}\{F\} - \boldsymbol{P}\{E \cap F\}$



<u>Proof</u>: (Hint: Think about Additivity Axiom)

Consequences of the 3 Probability Axioms

Lemma 2.5: $\boldsymbol{P}\{E \cup F\} = \boldsymbol{P}\{E\} + \boldsymbol{P}\{F\} - \boldsymbol{P}\{E \cap F\}$



Proof:

Express $E \cup F$ as a union of mutually exclusive sets

$$E \cup F = E \cup \big(F \setminus (E \cap F)\big)$$

Then, by the Additivity Axiom we have 2 observations:

$$\boldsymbol{P}\{E \cup F\} = \mathbf{P}\{E\} + \boldsymbol{P}\{F \setminus (E \cap F)\}$$

$$\boldsymbol{P}\{F\} = \boldsymbol{P}\{F \setminus (E \cap F)\} + \boldsymbol{P}\{E \cap F\}$$

Now substitute the 2nd equation into the 1st . ■

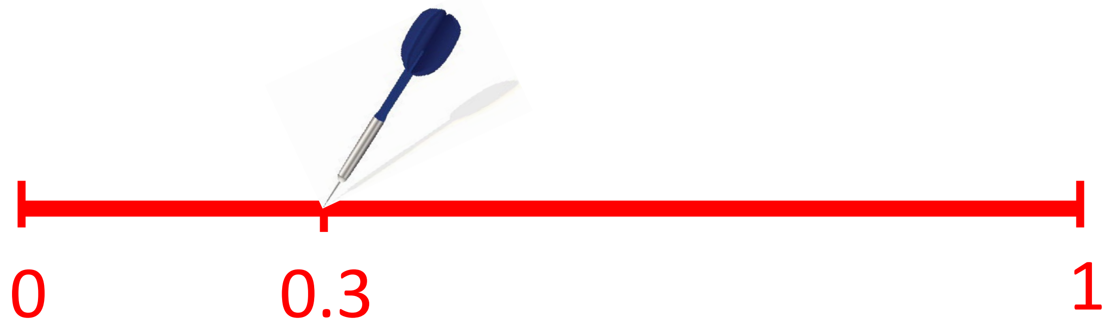
Lemma 2.6: $\boldsymbol{P}\{E \cup F\} \leq \boldsymbol{P}\{E\} + \boldsymbol{P}\{F\}$

Proof: WHY??

Consequences of the 3 Probability Axioms

Q: You throw a dart, equally likely to land anywhere in [0,1].
What is $\boldsymbol{P}\{\text{Dart lands at } 0.3\}$?

(Argue using the Probability Axioms.)

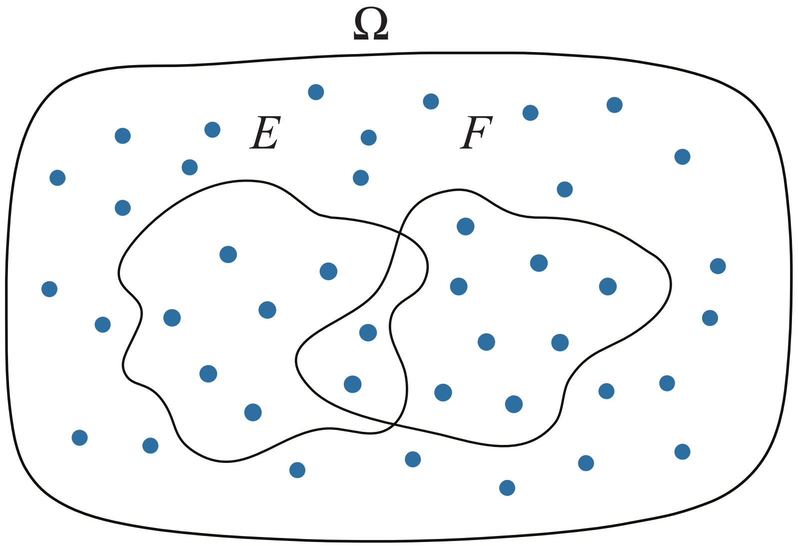


Conditional Probability on Events

<u>Defn</u>: The conditional probability of event E given event F is

$$\boldsymbol{P}\{E|F\} = \frac{\boldsymbol{P}\{E \cap F\}}{\boldsymbol{P}\{F\}}$$

assuming $\boldsymbol{P}\{F\} > 0$.



<u>Two equivalent views</u>:

$$\boldsymbol{P}\{E \mid F\} = \frac{2}{10}$$

(of the 10 outcomes in set F, only 2 of these are in set E)

$$\boldsymbol{P}\{E \mid F\} = \frac{\boldsymbol{P}\{E \cap F\}}{\boldsymbol{P}\{F\}} = \frac{\frac{2}{42}}{\frac{10}{42}} = \frac{2}{10}$$

Conditional Probability on Events

<u>Defn</u>: The conditional probability of event E given event F is

$$\boldsymbol{P}\{E|F\} = \frac{\boldsymbol{P}\{E \cap F\}}{\boldsymbol{P}\{F\}}$$

assuming $P\{F\} > 0$.

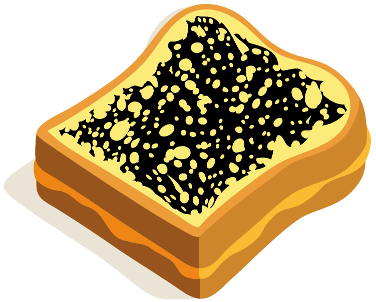
<u>Sandwich choices</u>:

Mon – Jelly
Tues – Cheese
Wed – Turkey
Thur – Cheese
Fri – Turkey
Sat – Cheese
Sun – None

1st half of week (Mon – Wed)

2nd half of week (Thur – Sun)

Q: What is $\boldsymbol{P}$\{Cheese | 2nd half of week\} ?
Argue this from 2 views.



Conditional Probability on Events

Defn: The conditional probability of event E given event F is

$$\boldsymbol{P}\{E|F\} = \frac{\boldsymbol{P}\{E \cap F\}}{\boldsymbol{P}\{F\}}$$

assuming $P\{F\} > 0$.

Sandwich choices:

Mon – Jelly
Tues – Cheese
Wed – Turkey
Thur – Cheese
Fri – Turkey
Sat – Cheese
Sun – None

1st half of week (Mon–Wed)

2nd half of week (Thur–Sun)

Q: What is $\boldsymbol{P}$\{Cheese | 2nd half of week\} ?
Argue this from 2 views.

$\boldsymbol{P}\{\text{Cheese} \mid 2^{\text{nd}} \text{ half}\} = \frac{2}{4}$ (of the 4 days in 2nd half, 2 are cheese sandwiches)

$$\boldsymbol{P}\{\text{Cheese} \mid 2^{\text{nd}} \text{ half}\} = \frac{\boldsymbol{P}\{\text{Cheese} \cap 2^{\text{nd}} \text{ half}\}}{\boldsymbol{P}\{2^{\text{nd}} \text{ half}\}} = \frac{\frac{2}{7}}{\frac{4}{7}} = \frac{2}{4}$$

Conditional Probability on Events

Q: What is $\boldsymbol{P}\{\text{both are colts} \mid\ \geq 1 \text{ colt}\}$?

The offspring of a horse is called a foal.
Horse couples have one foal at a time.
Each foal is equally likely to be a "colt" or a "filly."

We're told that a horse couple had 2 foals,
and at least one of these is a colt.

Conditional Probability on Events

The offspring of a horse is called a foal.
Horse couples have one foal at a time.
Each foal is equally likely to be a "colt" or a "filly."

We're told that a horse couple had 2 foals,
and at least one of these is a colt.

$$\mathbf{P}\{\text{both are colts} \mid \ \geq 1 \text{ colt}\}$$

$$= \frac{\mathbf{P}\{\text{both are colts} \ \& \geq 1 \text{ colt}\}}{\mathbf{P}\{\geq 1 \text{ colt}\}}$$

$$= \frac{\mathbf{P}\{\text{both are colts}\}}{\mathbf{P}\{\geq 1 \text{ colt}\}}$$

$$= \frac{\frac{1}{4}}{\frac{3}{4}}$$

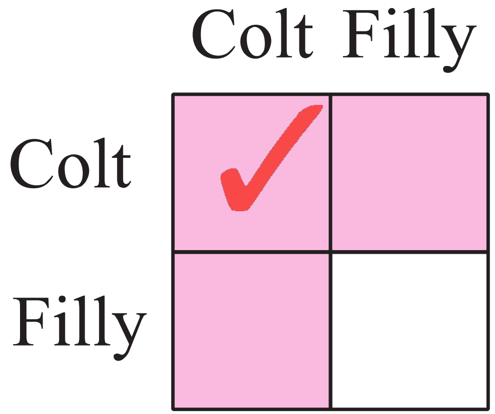
$$= \frac{1}{3}$$

Conditional Probability on Events

$$\boldsymbol{P}\{\text{both are colts} \mid\ \geq 1 \text{ colt}\} = \frac{1}{3}$$

The offspring of a horse is called a foal.
Horse couples have one foal at a time.
Each foal is equally likely to be a “colt” or a “filly.”

We’re told that a horse couple had 2 foals,
and at least one of these is a colt.



Conditional Probability on Events

If $\boldsymbol{P}\{E_1 \cap E_2\} > 0$, then:

$$\boldsymbol{P}\{E_2 | E_1\} = \frac{\boldsymbol{P}\{E_1 \cap E_2\}}{\boldsymbol{P}\{E_1\}}$$

Equivalently, we can write:

If $\boldsymbol{P}\{E_1 \cap E_2\} > 0$, then:

$$\boldsymbol{P}\{E_1 \cap E_2\} = \boldsymbol{P}\{E_1\} \cdot \boldsymbol{P}\{E_2 | E_1\}$$

Likewise: $\boldsymbol{P}\{E_1 \cap E_2\} = \boldsymbol{P}\{E_2\} \cdot \boldsymbol{P}\{E_1 | E_2\}$

- Probability outcome is in both E_1 and E_2
- Probability outcome is in E_1
- Probability outcome is in E_2 given that it's in E_1

Chain Rule for Conditioning

If $\boldsymbol{P}\{E_1 \cap E_2\} > 0$, then:

$$\boldsymbol{P}\{E_1 \cap E_2\} = \boldsymbol{P}\{E_1\} \cdot \boldsymbol{P}\{E_2 | E_1\}$$

This can be generalized!

Theorem 2.10: **[Chain Rule for Conditioning]**

If $\boldsymbol{P}\{E_1 \cap E_2 \cap E_3 \cap \cdots \cap E_n\} > 0$, then

$$\boldsymbol{P}\{E_1 \cap E_2 \cap E_3 \cap \cdots \cap E_n\}$$

$$= \boldsymbol{P}\{E_1\} \cdot \boldsymbol{P}\{E_2 \mid E_1\} \cdot \boldsymbol{P}\{E_3 \mid E_1 \cap E_2\} \cdots \boldsymbol{P}\{E_n \mid E_1 \cap E_2 \cap E_3 \cap \cdots \cap E_{n-1}\}$$

Independent Events

<u>Defn</u>: Events E and F are **independent**, written $E \perp F$, if:

$$\boldsymbol{P}\{E \cap F\} = \boldsymbol{P}\{E\} \cdot \boldsymbol{P}\{F\}$$

Here's an equivalent and more intuitive definition:

<u>Defn</u>: Assuming $\boldsymbol{P}\{F\} > 0$, Events E and F are **independent**, if:

$$\boldsymbol{P}\{E \mid F\} = \boldsymbol{P}\{E\}$$

See the book for a proof of the equivalence.

Practice with Independent Events

Defn: Events E and F are **independent**, written $E \perp F$, if:

$$\boldsymbol{P}\{E \cap F\} = \boldsymbol{P}\{E\} \cdot \boldsymbol{P}\{F\}$$

Defn: Assuming $\boldsymbol{P}\{F\} > 0$, Events E and F are **independent**, if:

$$\boldsymbol{P}\{E \mid F\} = \boldsymbol{P}\{E\}$$

Q: Can two mutually exclusive, non-null events be independent? No!

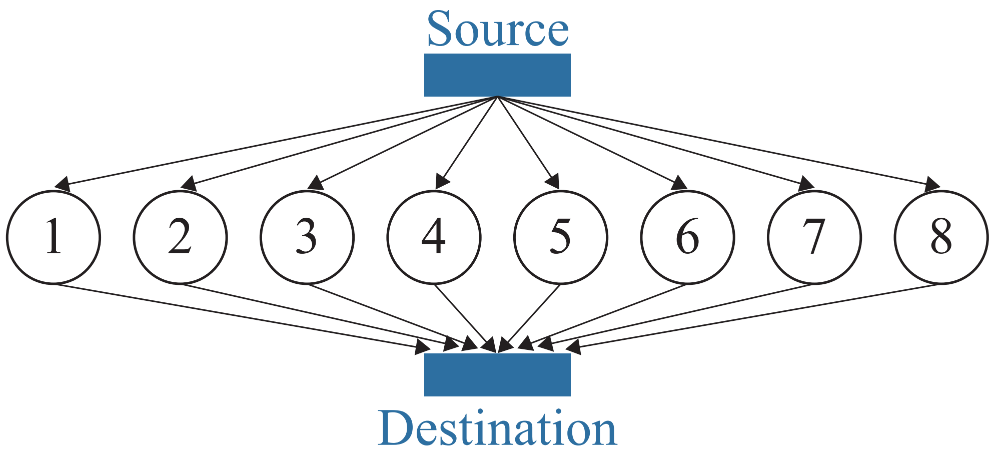
Q: Suppose we roll a die twice. Which of these pairs of events are independent:

a. Let E="1st roll is 6." Let F ="2nd roll is 6"
b. Let E="Sum of rolls is 7." Let F ="2nd roll is 4"

Both!

Practice with Independent Events

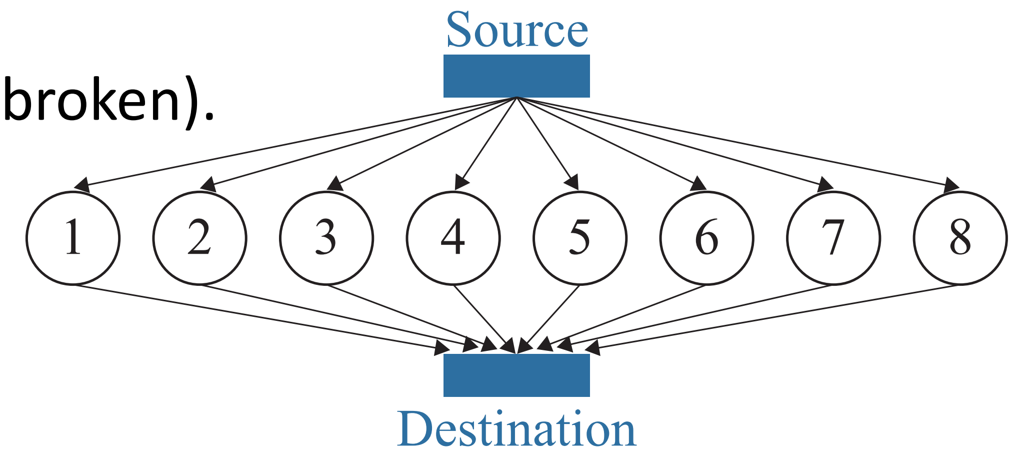
You are routing a packet from the source to the destination.
But each of the 16 edges in the network only works with probability p.



Q: What is the probability that you can get the packet from the source to the destination?

Practice with Independent Events

Each edge works with probability p. There are 8 paths.
Let E_i denote the event that the i^{th} path is usable (not broken).



Q: What is $\boldsymbol{P}\{E_i\}$?   **Q**: What is $\boldsymbol{P}\{\overline{E_1}\}$?

Q: What is $\boldsymbol{P}\{\text{Can get from source to destination}\}$?

$$\boldsymbol{P}\{\text{Can get from source to destination}\} = \boldsymbol{P}\{\text{At least one path works}\}$$

$$= \boldsymbol{P}\{E_1 \cup E_2 \cup \cdots \cup E_8\}$$

$$= 1 - \boldsymbol{P}\{\text{All paths are broken}\}$$

$$= 1 - \boldsymbol{P}\{\overline{E_1}\} \cdot \boldsymbol{P}\{\overline{E_2}\} \cdots \boldsymbol{P}\{\overline{E_8}\} = 1 - (1 - p^2)^8$$

More Independence Definitions

Defn 2.15: Events $A_1, A_2, \ldots, A_n$ are **independent** if, for every subset S of $\{1, 2, \ldots, n\}$:

$$\mathbf{P}\left\{\bigcap_{i \in S} A_i\right\} = \prod_{i \in S} \boldsymbol{P}\{A_i\}$$

Defn 2.16: Events $A_1, A_2, \ldots, A_n$ are **pairwise independent** if every pair of events is independent.

Defn 2.17: Two events E and F are said to be **conditionally independent given $\boldsymbol{G}$**, where $\boldsymbol{P}\{G\} > 0$, if

$$\boldsymbol{P}\{E \cap F \mid G\} = \boldsymbol{P}\{E \mid G\} \cdot \boldsymbol{P}\{F \mid G\}$$

Law of Total Probability

For any sets E and F: $E = (E \cap F) \cup (E \cap \bar{F})$

$$\boldsymbol{P}\{E\} = \boldsymbol{P}\{E \cap F\} + \boldsymbol{P}\{E \cap \bar{F}\}$$

$$= \boldsymbol{P}\{E \mid F\} \cdot \boldsymbol{P}\{F\} + \boldsymbol{P}\{E \mid \bar{F}\} \cdot \boldsymbol{P}\{\bar{F}\}$$

Generalizing, we have:

Theorem 2.18: **[Law of Total Probability]**
Let $F_1, F_2, \dots, F_n$ partition the state space Ω. Then:

$$\boldsymbol{P}\{E\} = \sum_{i=1}^{n} \boldsymbol{P}\{E \cap F_i\} = \sum_{i=1}^{n} \boldsymbol{P}\{E | F_i\} \cdot \boldsymbol{P}\{F_i\}$$

Law of Total Probability

The Law of Total Probability applies to conditional probability as well:

Theorem 2.19: **[Law of Total Probability for Conditional Probability]**
Let $F_1, F_2, \ldots, F_n$ partition the state space Ω. Then:

$$\boldsymbol{P}\{A \mid B\} = \sum_{i=1}^{n} \boldsymbol{P}\{A \mid B \cap F_i\} \cdot \boldsymbol{P}\{F_i \mid B\}$$

Bayes' Law

Sometimes we want to know $\boldsymbol{P}\{F \mid E\}$ but all we know is the reverse direction, $\boldsymbol{P}\{E \mid F\}$.

Theorem 2.20 : **[Bayes' Law]** Assuming $\boldsymbol{P}\{E\} > 0$,

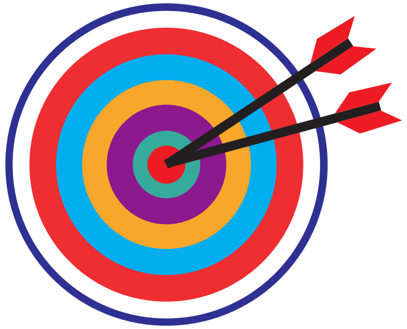
$$\boldsymbol{P}\{F \mid E\} = \frac{\boldsymbol{P}\{E \cap F\}}{\boldsymbol{P}\{E\}} = \frac{\boldsymbol{P}\{E \mid \mathrm{F}\} \cdot \boldsymbol{P}\{F\}}{\boldsymbol{P}\{E\}}$$

Theorem 2.21: **[Extended Bayes' Law]** Let $F_1, F_2, \dots, F_n$ partition Ω. Assuming $\boldsymbol{P}\{E\} > 0$,

$$\boldsymbol{P}\{F \mid E\} = \frac{\boldsymbol{P}\{E \mid F\} \cdot P\{F\}}{\boldsymbol{P}\{E\}} = \frac{\boldsymbol{P}\{E \mid \mathrm{F}\} \cdot \boldsymbol{P}\{F\}}{\sum_{j=1}^{n} \boldsymbol{P}\{E \mid F_j\} \cdot P\{F_j\}}$$

Bayes' Law Example

There's a rare child cancer that occurs in one out of a million kids.
There's a test for this cancer that is 99.9% effective:



<u>High Accuracy Cancer Screening</u>

If CANCER ⟶ Test Postive w.p. 99.9%

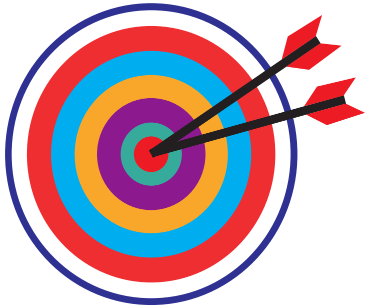
If $\overline{\text{CANCER}}$ ⟶ Test Negative w.p. 99.9%

Q: Suppose my child's test result is positive. How worried should I be?

Bayes' Law Example

Rare cancer occurs in 1 out of 10^6 kids.
Test for this cancer is 99.9% effective:

Q: My child's test result is positive.
How worried should I be?



High Accuracy Cancer Screening

If CANCER ⟶ Test Postive w.p. 99.9%

If $\overline{\text{CANCER}}$ ⟶ Test Negative w.p. 99.9%

$\boldsymbol{P}\{\text{Cancer} \mid \text{Test Pos}\}$

$$= \frac{\boldsymbol{P}\{\text{Test pos} \mid \text{Cancer}\} \cdot \boldsymbol{P}\{\text{Cancer}\}}{\boldsymbol{P}\{\text{Test pos} \mid \text{Cancer}\} \cdot \boldsymbol{P}\{\text{Cancer}\} + \boldsymbol{P}\{\text{Test pos} \mid \text{No cancer}\} \cdot \boldsymbol{P}\{\text{No cancer}\}}$$

$$= \frac{0.999 \cdot 10^{-6}}{0.999 \cdot 10^{-6} + 10^{-3} \cdot (1 - 10^{-6})} \approx \frac{10^{-6}}{10^{-6} + 10^{-3}} = \frac{1}{1001}$$

Why so low?