Gov 51: Bayes Rule

Matthew Blackwell

Harvard University

QAnon

You meet a man named Steve and he tells you that he is a Republican. You have been interested in meeting someone who believes in the QAnon conspiracy theory. Given what you know about Steve, would you guess that he believes in QAnon or not?

QAnon

You meet a man named Steve and he tells you that he is a Republican. You have been interested in meeting someone who believes in the QAnon conspiracy theory. Given what you know about Steve, would you guess that he believes in QAnon or not?

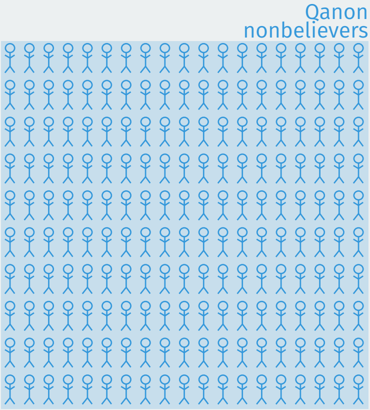
- Common response: probably believes in QAnon since believers tend to be Republicans.

QAnon

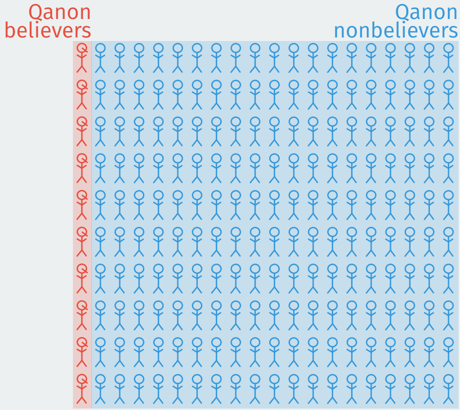
You meet a man named Steve and he tells you that he is a Republican. You have been interested in meeting someone who believes in the QAnon conspiracy theory. Given what you know about Steve, would you guess that he believes in QAnon or not?

- Common response: probably believes in QAnon since believers tend to be Republicans.
- **Base rate fallacy**: ignores how uncommon QAnon believers are!

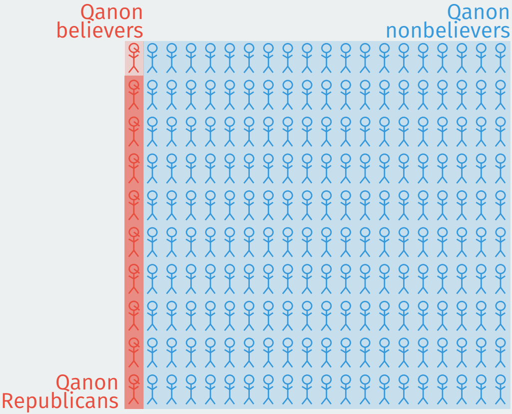
Visualizing QAnon support



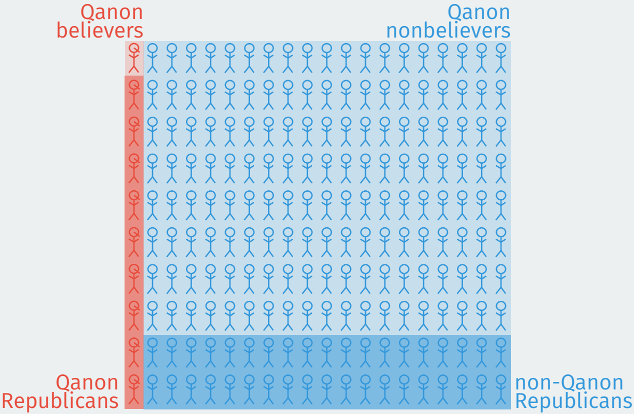
Visualizing QAnon support



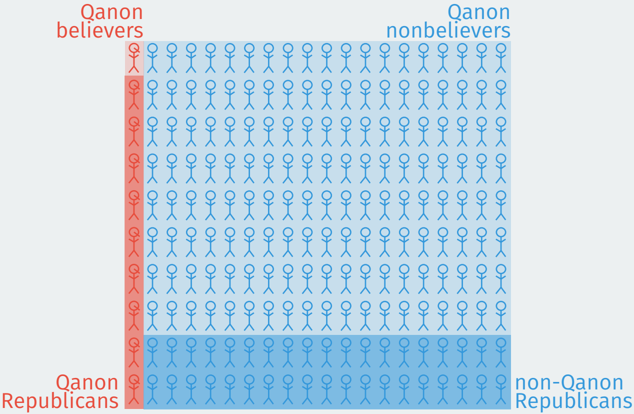
Visualizing QAnon support



Visualizing QAnon support

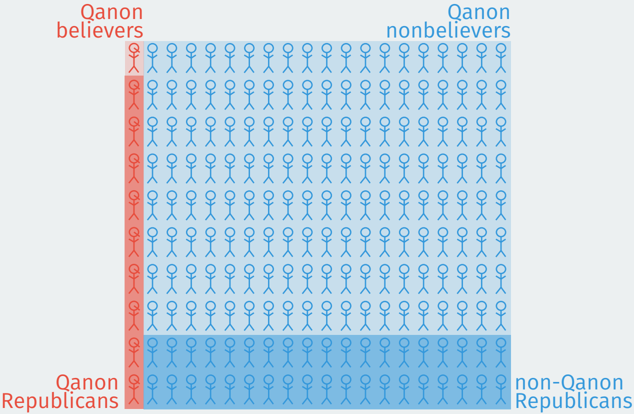


Visualizing QAnon support



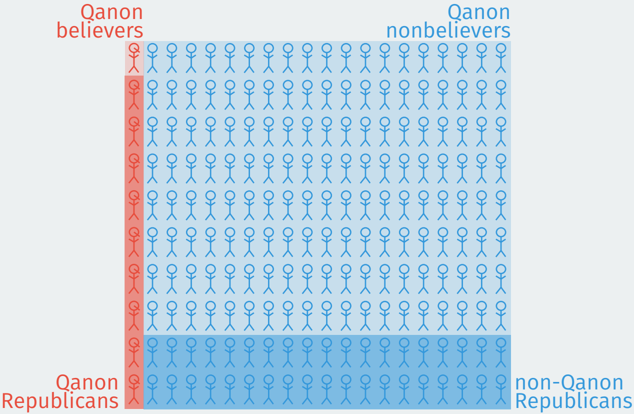
Chance a random Republican believes QAnon =

Visualizing QAnon support

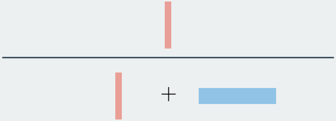


Chance a random Republican believes QAnon =

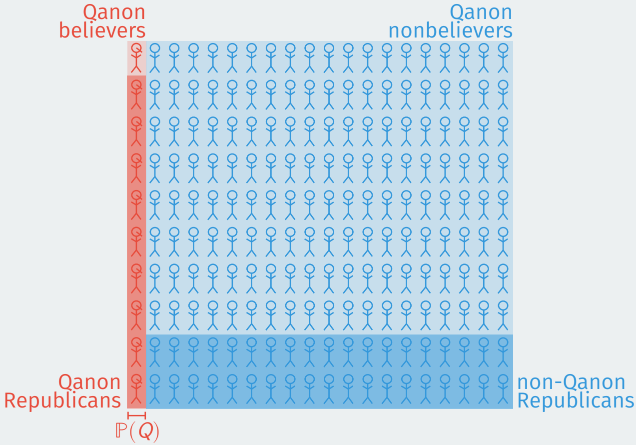
Visualizing QAnon support

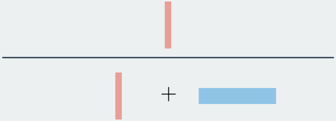


Chance a random Republican believes QAnon =

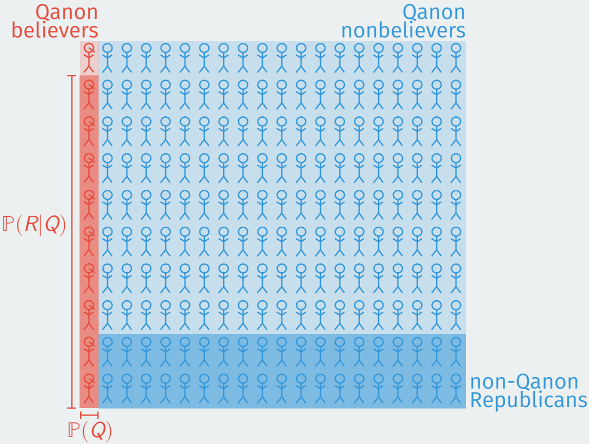


Visualizing QAnon support

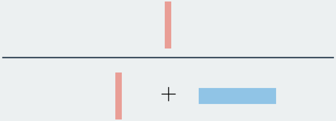


Chance a random Republican believes QAnon = 

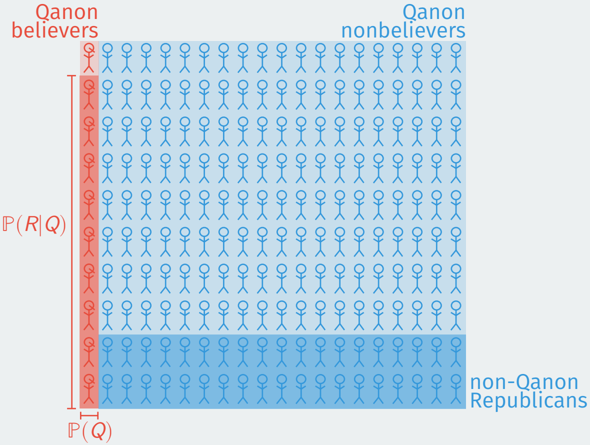
Visualizing QAnon support



Chance a random Republican believes QAnon =

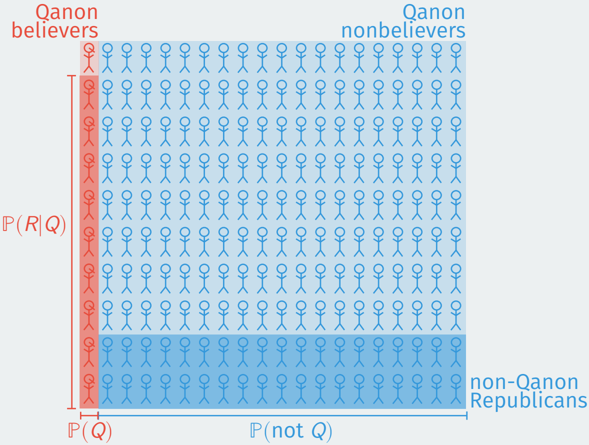


Visualizing QAnon support



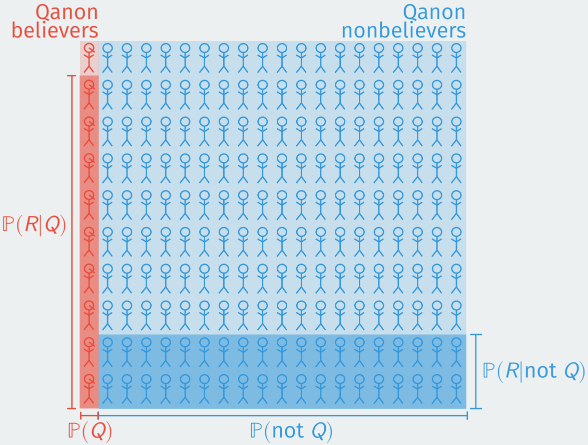
Chance a random Republican believes QAnon = $\dfrac{\mathbb{P}(R|Q)\mathbb{P}(Q)}{\mathbb{P}(R|Q)\mathbb{P}(Q) + \quad\quad}$

Visualizing QAnon support



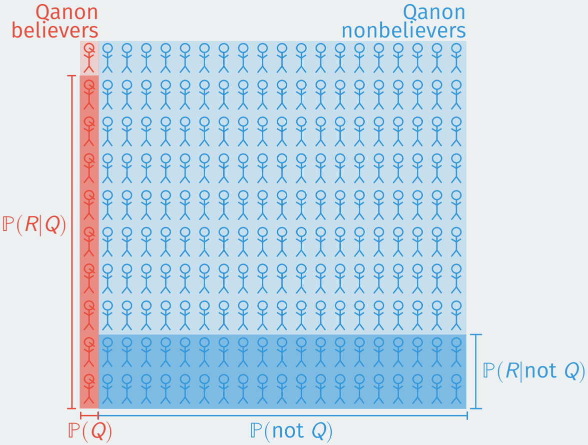
Chance a random Republican believes QAnon = $\dfrac{\mathbb{P}(R|Q)\mathbb{P}(Q)}{\mathbb{P}(R|Q)\mathbb{P}(Q) + \quad\quad}$

Visualizing QAnon support



Chance a random Republican believes QAnon = $\dfrac{\mathbb{P}(R|Q)\mathbb{P}(Q)}{\mathbb{P}(R|Q)\mathbb{P}(Q) + \qquad}$

Visualizing QAnon support



$$\text{Chance a random Republican believes QAnon} = \frac{\mathbb{P}(R|Q)\mathbb{P}(Q)}{\mathbb{P}(R|Q)\mathbb{P}(Q) + \mathbb{P}(R|\text{not } Q)\mathbb{P}(\text{not } Q)}$$

Bayes' rule

- Reverend Thomas Bayes (1701–61): English minister and statistician

Bayes' rule

- Reverend Thomas Bayes (1701–61): English minister and statistician
- **Bayes' rule**: if $\mathbb{P}(B) > 0$, then:

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(B \mid A)\mathbb{P}(A)}{\mathbb{P}(B)}$$

Bayes' rule

- Reverend Thomas Bayes (1701–61): English minister and statistician
- **Bayes' rule**: if $\mathbb{P}(B) > 0$, then:

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(B \mid A)\mathbb{P}(A)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B \mid A)\mathbb{P}(A)}{\mathbb{P}(B \mid A)\mathbb{P}(A) + \mathbb{P}(B \mid \text{not } A)\mathbb{P}(\text{not } A)}$$

Why is Bayes' rule useful?

- What is the probability of some hypothesis given some evidence?

Why is Bayes' rule useful?

- What is the probability of some hypothesis given some evidence?
    - $\mathbb{P}(\text{QAnon} \mid \text{Republican})$?

Why is Bayes' rule useful?

- What is the probability of some hypothesis given some evidence?
  - $\mathbb{P}(\text{QAnon} \mid \text{Republican})$?
- Often easier to know probability of evidence given hypothesis.

Why is Bayes' rule useful?

- What is the probability of some hypothesis given some evidence?
    - $\mathbb{P}(\text{QAnon} \mid \text{Republican})$?
- Often easier to know probability of evidence given hypothesis.
    - $\mathbb{P}(\text{Republican} \mid \text{QAnon})$

Why is Bayes' rule useful?

- What is the probability of some hypothesis given some evidence?
    - $\mathbb{P}(\text{QAnon} \mid \text{Republican})$?
- Often easier to know probability of evidence given hypothesis.
    - $\mathbb{P}(\text{Republican} \mid \text{QAnon})$
- Combine this with the **prior probability** of the hypothesis.

Why is Bayes' rule useful?

- What is the probability of some hypothesis given some evidence?
    - $\mathbb{P}(\text{QAnon} \mid \text{Republican})$?
- Often easier to know probability of evidence given hypothesis.
    - $\mathbb{P}(\text{Republican} \mid \text{QAnon})$
- Combine this with the **prior probability** of the hypothesis.
    - Prior: $\mathbb{P}(\text{QAnon})$

Why is Bayes' rule useful?

- What is the probability of some hypothesis given some evidence?
    - $\mathbb{P}(\text{QAnon} \mid \text{Republican})$?
- Often easier to know probability of evidence given hypothesis.
    - $\mathbb{P}(\text{Republican} \mid \text{QAnon})$
- Combine this with the **prior probability** of the hypothesis.
    - Prior: $\mathbb{P}(\text{QAnon})$
    - **Posterior**: $\mathbb{P}(\text{QAnon} \mid \text{Republican})$

Why is Bayes' rule useful?

- What is the probability of some hypothesis given some evidence?
  - $\mathbb{P}(\text{QAnon} \mid \text{Republican})$?
- Often easier to know probability of evidence given hypothesis.
  - $\mathbb{P}(\text{Republican} \mid \text{QAnon})$
- Combine this with the **prior probability** of the hypothesis.
  - Prior: $\mathbb{P}(\text{QAnon})$
  - **Posterior**: $\mathbb{P}(\text{QAnon} \mid \text{Republican})$
- Applying Bayes' rule is often called **updating the prior**.

Why is Bayes' rule useful?

- What is the probability of some hypothesis given some evidence?
  - $\mathbb{P}(\text{QAnon} \mid \text{Republican})$?
- Often easier to know probability of evidence given hypothesis.
  - $\mathbb{P}(\text{Republican} \mid \text{QAnon})$
- Combine this with the **prior probability** of the hypothesis.
  - Prior: $\mathbb{P}(\text{QAnon})$
  - **Posterior**: $\mathbb{P}(\text{QAnon} \mid \text{Republican})$
- Applying Bayes' rule is often called **updating the prior**.
  - $\mathbb{P}(\text{QAnon}) \rightsquigarrow \mathbb{P}(\text{QAnon} \mid \text{Republican})$

Why is Bayes' rule useful?

- What is the probability of some hypothesis given some evidence?
  - $\mathbb{P}(\text{QAnon} \mid \text{Republican})$?
- Often easier to know probability of evidence given hypothesis.
  - $\mathbb{P}(\text{Republican} \mid \text{QAnon})$
- Combine this with the **prior probability** of the hypothesis.
  - Prior: $\mathbb{P}(\text{QAnon})$
  - **Posterior**: $\mathbb{P}(\text{QAnon} \mid \text{Republican})$
- Applying Bayes' rule is often called **updating the prior**.
  - $\mathbb{P}(\text{QAnon}) \rightsquigarrow \mathbb{P}(\text{QAnon} \mid \text{Republican})$
  - How does the evidence change the chance of the hypothesis being true?

Uses of Bayes’ rule

- Medical testing:

Uses of Bayes' rule

- Medical testing:
    - Want to know: $\mathbb{P}(\text{Disease} \mid \text{Test Positive})$

Uses of Bayes' rule

- Medical testing:
  - Want to know: $\mathbb{P}(\text{Disease} \mid \text{Test Positive})$
  - Have: $\mathbb{P}(\text{Test Positive} \mid \text{Disease})$ and $\mathbb{P}(\text{Disease})$

Uses of Bayes' rule

- Medical testing:
  - Want to know: $\mathbb{P}(\text{Disease} \mid \text{Test Positive})$
  - Have: $\mathbb{P}(\text{Test Positive} \mid \text{Disease})$ and $\mathbb{P}(\text{Disease})$
- Predicting traits from names:

Uses of Bayes' rule

- Medical testing:
  - Want to know: $\mathbb{P}(\text{Disease} \mid \text{Test Positive})$
  - Have: $\mathbb{P}(\text{Test Positive} \mid \text{Disease})$ and $\mathbb{P}(\text{Disease})$
- Predicting traits from names:
  - Want to know: $\mathbb{P}(\text{African American} \mid \text{Last Name})$

Uses of Bayes' rule

- Medical testing:
  - Want to know: $\mathbb{P}(\text{Disease} \mid \text{Test Positive})$
  - Have: $\mathbb{P}(\text{Test Positive} \mid \text{Disease})$ and $\mathbb{P}(\text{Disease})$
- Predicting traits from names:
  - Want to know: $\mathbb{P}(\text{African American} \mid \text{Last Name})$
  - Have: $\mathbb{P}(\text{Last Name} \mid \text{African American})$ and $\mathbb{P}(\text{African American})$

Uses of Bayes' rule

- Medical testing:
  - Want to know: $\mathbb{P}(\text{Disease} \mid \text{Test Positive})$
  - Have: $\mathbb{P}(\text{Test Positive} \mid \text{Disease})$ and $\mathbb{P}(\text{Disease})$
- Predicting traits from names:
  - Want to know: $\mathbb{P}(\text{African American} \mid \text{Last Name})$
  - Have: $\mathbb{P}(\text{Last Name} \mid \text{African American})$ and $\mathbb{P}(\text{African American})$
- Spam filtering:

Uses of Bayes' rule

- Medical testing:
  - Want to know: $\mathbb{P}(\text{Disease} \mid \text{Test Positive})$
  - Have: $\mathbb{P}(\text{Test Positive} \mid \text{Disease})$ and $\mathbb{P}(\text{Disease})$
- Predicting traits from names:
  - Want to know: $\mathbb{P}(\text{African American} \mid \text{Last Name})$
  - Have: $\mathbb{P}(\text{Last Name} \mid \text{African American})$ and $\mathbb{P}(\text{African American})$
- Spam filtering:
  - Want to know: $\mathbb{P}(\text{Spam} \mid \text{Email text})$

Uses of Bayes' rule

- Medical testing:
  - Want to know: $\mathbb{P}(\text{Disease} \mid \text{Test Positive})$
  - Have: $\mathbb{P}(\text{Test Positive} \mid \text{Disease})$ and $\mathbb{P}(\text{Disease})$
- Predicting traits from names:
  - Want to know: $\mathbb{P}(\text{African American} \mid \text{Last Name})$
  - Have: $\mathbb{P}(\text{Last Name} \mid \text{African American})$ and $\mathbb{P}(\text{African American})$
- Spam filtering:
  - Want to know: $\mathbb{P}(\text{Spam} \mid \text{Email text})$
  - Have: $\mathbb{P}(\text{Email text} \mid \text{Spam})$ and $\mathbb{P}(\text{Spam})$

Medical tests

- Suppose you go and get a COVID-19 test and it comes back positive!

Medical tests

- Suppose you go and get a COVID-19 test and it comes back positive!
    - Let a positive test be PT.

Medical tests

- Suppose you go and get a COVID-19 test and it comes back positive!
  - Let a positive test be PT.
- What's the probability you actually have COVID-19?

Medical tests

- Suppose you go and get a COVID-19 test and it comes back positive!
    - Let a positive test be PT.
- What's the probability you actually have COVID-19?
    - Let having COVID be labeled C.

Medical tests

- Suppose you go and get a COVID-19 test and it comes back positive!
    - Let a positive test be PT.
- What's the probability you actually have COVID-19?
    - Let having COVID be labeled C.
    - Question: What is $\mathbb{P}(C \mid PT)$?

Medical tests

- Suppose you go and get a COVID-19 test and it comes back positive!
  - Let a positive test be PT.
- What's the probability you actually have COVID-19?
  - Let having COVID be labeled C.
  - Question: What is $\mathbb{P}(C \mid PT)$?
- Components for calculating Bayes' rule:

Medical tests

- Suppose you go and get a COVID-19 test and it comes back positive!
    - Let a positive test be PT.
- What's the probability you actually have COVID-19?
    - Let having COVID be labeled C.
    - Question: What is $\mathbb{P}(C \mid PT)$?
- Components for calculating Bayes' rule:
    - $\mathbb{P}(PT|C) = 0.7$: true positive rate

Medical tests

- Suppose you go and get a COVID-19 test and it comes back positive!
  - Let a positive test be PT.
- What's the probability you actually have COVID-19?
  - Let having COVID be labeled C.
  - Question: What is $\mathbb{P}(C \mid PT)$?
- Components for calculating Bayes' rule:
  - $\mathbb{P}(PT|C) = 0.7$: true positive rate
  - $\mathbb{P}(PT \mid \text{not } C) = 0.05$: false positive rate

Medical tests

- Suppose you go and get a COVID-19 test and it comes back positive!
    - Let a positive test be PT.
- What's the probability you actually have COVID-19?
    - Let having COVID be labeled C.
    - Question: What is $\mathbb{P}(C \mid PT)$?
- Components for calculating Bayes' rule:
    - $\mathbb{P}(PT|C) = 0.7$: true positive rate
    - $\mathbb{P}(PT \mid \text{not } C) = 0.05$: false positive rate
    - $\mathbb{P}(C) = 0.001$ rough prevalance of active COVID cases.

Applying Bayes' rule to COVID tests

- Use the law of total probability to get the denominator:

Applying Bayes' rule to COVID tests

- Use the law of total probability to get the denominator:

$$\mathbb{P}(PT)$$

Applying Bayes' rule to COVID tests

- Use the law of total probability to get the denominator:

$$\mathbb{P}(PT) = \mathbb{P}(PT \mid C)\mathbb{P}(C)$$

Applying Bayes' rule to COVID tests

- Use the law of total probability to get the denominator:

$$\mathbb{P}(PT) = \mathbb{P}(PT \mid C)\mathbb{P}(C) + \mathbb{P}(PT|\text{not } C)\mathbb{P}(\text{not } C)$$

Applying Bayes' rule to COVID tests

- Use the law of total probability to get the denominator:

$$\mathbb{P}(PT) = \mathbb{P}(PT \mid C)\mathbb{P}(C) + \mathbb{P}(PT|\text{not } C)\mathbb{P}(\text{not } C)$$
$$= (0.7 \times 0.001) + (0.05 \times 0.999)$$

Applying Bayes' rule to COVID tests

- Use the law of total probability to get the denominator:

$$
\begin{aligned}
\mathbb{P}(PT) &= \mathbb{P}(PT \mid C)\mathbb{P}(C) + \mathbb{P}(PT|\text{not } C)\mathbb{P}(\text{not } C) \\
&= (0.7 \times 0.001) + (0.05 \times 0.999) \\
&= 0.051
\end{aligned}
$$

Applying Bayes' rule to COVID tests

- Use the law of total probability to get the denominator:

$$
\begin{aligned}
\mathbb{P}(PT) &= \mathbb{P}(PT \mid C)\mathbb{P}(C) + \mathbb{P}(PT|\text{not } C)\mathbb{P}(\text{not } C) \\
&= (0.7 \times 0.001) + (0.05 \times 0.999) \\
&= 0.051
\end{aligned}
$$

- Now plug in all values to Bayes' rule:

Applying Bayes' rule to COVID tests

- Use the law of total probability to get the denominator:

$$\begin{aligned}\mathbb{P}(PT) &= \mathbb{P}(PT \mid C)\mathbb{P}(C) + \mathbb{P}(PT|\text{not } C)\mathbb{P}(\text{not } C)\\ &= (0.7 \times 0.001) + (0.05 \times 0.999)\\ &= 0.051\end{aligned}$$

- Now plug in all values to Bayes' rule:

$$\mathbb{P}(C \mid PT)$$

Applying Bayes' rule to COVID tests

- Use the law of total probability to get the denominator:

$$\begin{aligned}\mathbb{P}(PT) &= \mathbb{P}(PT \mid C)\mathbb{P}(C) + \mathbb{P}(PT|\text{not } C)\mathbb{P}(\text{not } C) \\ &= (0.7 \times 0.001) + (0.05 \times 0.999) \\ &= 0.051\end{aligned}$$

- Now plug in all values to Bayes' rule:

$$\mathbb{P}(C \mid PT) = \frac{\mathbb{P}(PT \mid C)\mathbb{P}(C)}{\mathbb{P}(PT)}$$

Applying Bayes' rule to COVID tests

- Use the law of total probability to get the denominator:

$$\begin{aligned}\mathbb{P}(PT) &= \mathbb{P}(PT \mid C)\mathbb{P}(C) + \mathbb{P}(PT|\text{not } C)\mathbb{P}(\text{not } C) \\ &= (0.7 \times 0.001) + (0.05 \times 0.999) \\ &= 0.051\end{aligned}$$

- Now plug in all values to Bayes' rule:

$$\mathbb{P}(C \mid PT) = \frac{\mathbb{P}(PT \mid C)\mathbb{P}(C)}{\mathbb{P}(PT)} = \frac{0.7 \times 0.001}{0.051} \approx 0.014$$