LESSON 8

Total Probability and Bayes' Rule

CSCI 3022

Roadmap

Lesson 8, CSCI 3022

- Total Probability
- Bayes' Rule

See Lecture Slides From Last Class for Filled In Solution:

Suppose you have a biased coin, with probability of heads given by **q**

a). You flip the coin 5 times. What is the probability that you get the sequence HTTHT?

$$
\begin{aligned}
P(\text{you get the sequence } HTTHT) =& P(H_1, T_2, T_3, H_4, T_5) \text{ (because this is a joint probability)} \\
=& P(H_1)P(T_2|H_1)P(T_3|H_1, T_2)P(H_4|H_1, T_2, T_3)P(T_5|H_1, T_2, T_3, H_4) \text{ (by the multiplication rule)} \\
=& q(1-q)(1-q)q(1-q) \\
=& q^2(1-q)^3
\end{aligned}
$$

b). You flip the coin 5 times. What is the probability that you get 2 heads?

Ways you can get 2 heads:

$$E1 = HTTHT$$
$$E2 = HHTTT$$
$$E3 = HTTTH$$
$$E4 = HTHTT$$
$$E5 = THHTT$$
$$E6 = THTHT$$
$$E7 = THTTH$$
$$E8 = TTHHT$$
$$E9 = TTHTH$$
$$E10 = TTTHH$$

For this problem we're interested in different possible ways we can get exactly 2 heads when flipping a coin 5 times.

We've provided the full list, but we don't actually need the list, just the number of ways.
We can count this using combinations. The number of ways is equivalent to $\binom{5}{2} = \frac{5!}{2!3!} = 10$

$$
\begin{aligned}
P(\text{exactly two heads}) &= P(E1 \cup E2 \cup E3 \cup E4 \ldots \cup E10) \\
&= P(E1) + P(E2) + P(E3) + \ldots P(E10) \text{ (by the addition rule for disjoint events)} \\
&= q^2(1-q)^3 + q^2(1-q)^3 + q^2(1-q)^3 \ldots q^2(1-q)^3 \text{ (because the probability for EACH of these events is } q^2(1-q)^3\text{)} \\
&= \binom{5}{2} q^2(1-q)^3 \\
&= 10q^2(1-q)^3
\end{aligned}
$$

Warm-Up: Serendipity

- There population of CU undergraduates is n=31,000 students
- Suppose you are friends with r=100 people.
- You walk into a classroom and you see k=160 random people.
- Assume each group of k CU undergrads is equally likely to be in the room.

What is the probability that you see at least one friend in the room?

https://web.stanford.edu/class/cs109/demos/serendipity.html

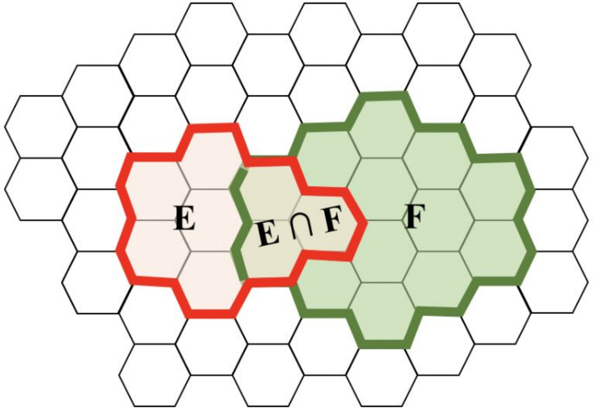
Example

Suppose there's a rare disease with prevalence of 1/1000 in the population.
A test for the disease has a false positive rate of 5% and a false negative rate of 1%.

a). What's the probability you test positive?

b). You test positive. What's the probability you have the disease?

Law of Total Probability



$$\mathrm{P}(E) = \mathrm{P}(E \text{ and } F) + \mathrm{P}(E \text{ and } F^{C})$$

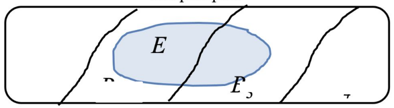
Law of Total Probability

<u>Thm</u> Let F be an event where $P(F) > 0$. For any event E,

$$P(E) = P(E|F)P(F) + P(E|F^C)P(F^C)$$

General Law of Total Probability

<u>Thm</u> For mutually exclusive events $F_1, F_2, \ldots, F_n$ such that $F_1 \cup F_2 \cup \cdots \cup F_n = S$,



$$P(E) = \sum_{i=1}^{n} P(E|F_i)P(F_i)$$

Practice

Example:
Suppose there's a rare disease with prevalence of 1/1000 in the population.
A test for the disease has a false positive rate of 5% and a false negative rate of 1%.

a). What's the probability you test positive?

$$P(\text{Pos}) = P(\text{Pos, D}) + P(\text{Pos, No D})$$

Positive Test

Disease

Practice

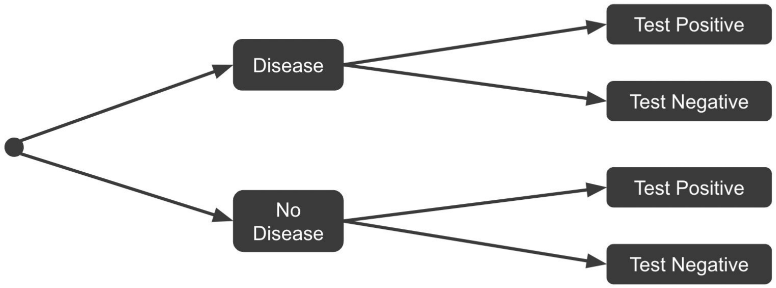
$P(E) = P(E|F)P(F) + P(E|F^C)P(F^C)$ Law of Total Probability

Example:
Suppose there's a rare disease with prevalence of 1/1000 in the population.
A test for the disease has a false positive rate of 5% and a false negative rate of 1%.

a). What's the probability you test positive?

P(Pos) = P(Pos, D) + P(Pos, No D)

= $\boxed{\text{P(D.) P(Pos. | D.) + P(No D.) P(Pos. | No D.)}}$



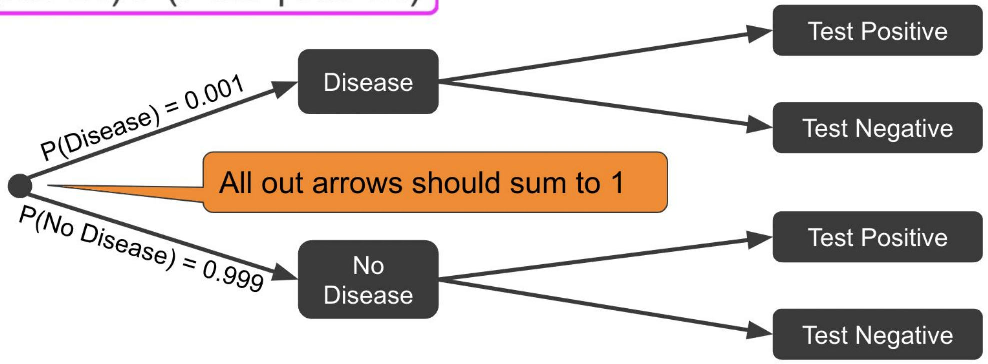
$$P(E) = P(E|F)P(F) + P(E|F^C)P(F^C)$$ Law of Total Probability

Example:
Suppose there's a rare disease with prevalence of 1/1000 in the population.
A test for the disease has a false positive rate of 5% and a false negative rate of 1%.

a). What's the probability you test positive?

P(Pos) = P(Pos, D) + P(Pos, No D)

= $\boxed{\text{P(D.) P(Pos. | D.) + P(No D.) P(Pos. | No D.)}}$



- Rare disease with **prevalence of 1/1000** in population

Practice

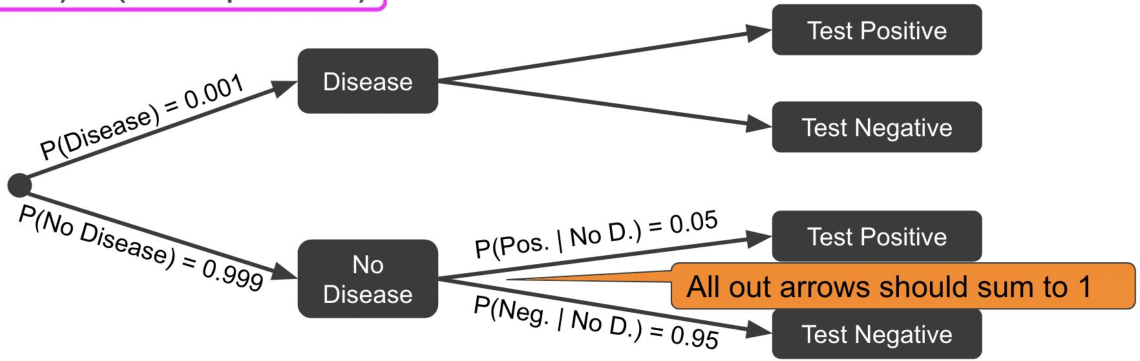
$$P(E) = P(E|F)P(F) + P(E|F^C)P(F^C)$$ Law of Total Probability

Example:
Suppose there's a rare disease with prevalence of 1/1000 in the population.
A test for the disease has a false positive rate of 5% and a false negative rate of 1%.

a). What's the probability you test positive?

P(Pos) = P(Pos, D) + P(Pos, No D)

= $\boxed{\text{P(D.) P(Pos. | D.) + P(No D.) P(Pos. | No D.)}}$



- **False Positive Rate of 5%**
  - ***If you do NOT have the disease then*** **5%** of the time the test says you do.

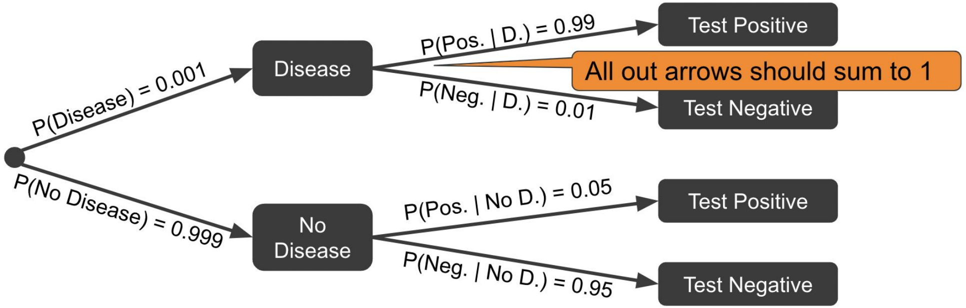
$$P(E) = P(E|F)P(F) + P(E|F^C)P(F^C)$$ Law of Total Probability

Example:
Suppose there's a rare disease with prevalence of 1/1000 in the population.
A test for the disease has a false positive rate of 5% and a false negative rate of 1%.

a). What's the probability you test positive?

P(Pos) = P(Pos, D) + P(Pos, No D)

= $\boxed{\text{P(D.) P(Pos. | D.) + P(No D.) P(Pos. | No D.)}}$



- **False Negative Rate of 1%**
  - ***If* <u>you DO have the disease</u>** then 1% of the time the test says you do not have the disease.

Practice

$$P(E) = P(E|F)P(F) + P(E|F^C)P(F^C)$$ Law of Total Probability

Example:

Suppose there's a rare disease with prevalence of 1/1000 in the population.
A test for the disease has a false positive rate of 5% and a false negative rate of 1%.

a). What's the probability you test positive?

$$P(\text{Pos}) = P(\text{Pos, D}) + P(\text{Pos, No D})$$

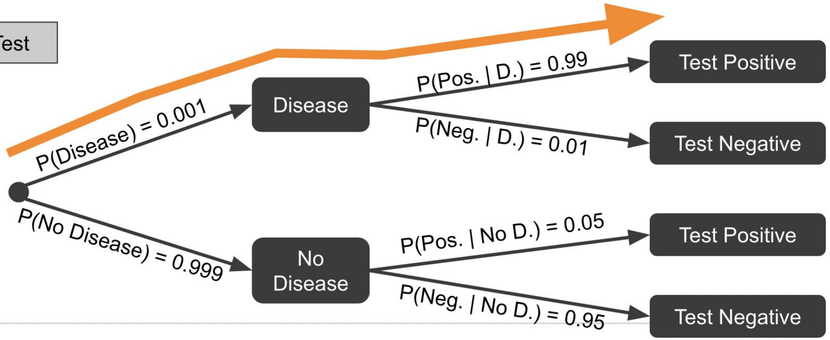
$$= \boxed{\underbrace{P(\text{D.})\, P(\text{Pos.} \mid \text{D.})}_{P(\text{D., Pos.})} + P(\text{No D.})\, P(\text{Pos.} \mid \text{No D.})}$$

$$P(\text{D., Pos.}) = P(\text{D.})\, P(\text{Pos.} \mid \text{D.})$$

$$= 0.001 * 0.99 = 0.00099$$

Disease

Positive Test



Practice

$P(E) = P(E|F)P(F) + P(E|F^C)P(F^C)$ Law of Total Probability

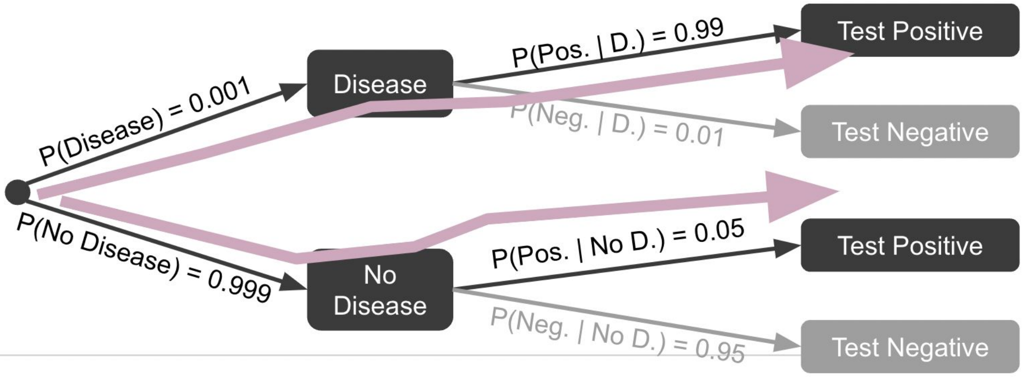
Example:
Suppose there's a rare disease with prevalence of 1/1000 in the population.
A test for the disease has a false positive rate of 5% and a false negative rate of 1%.

a). What's the probability you test positive?

P(Pos) = P(Pos, D) + P(Pos, No D)

= $\boxed{\text{P(D.) P(Pos. | D.) + P(No D.) P(Pos. | No D.)}}$

= (.001)*(0.99) + (0.999) * (0.05) = **0.05094**



Example

Suppose there's a rare disease with prevalence of 1/1000 in the population.
A test for the disease has a false positive rate of 5% and a false negative rate of 1%.

a). What's the probability you test positive?

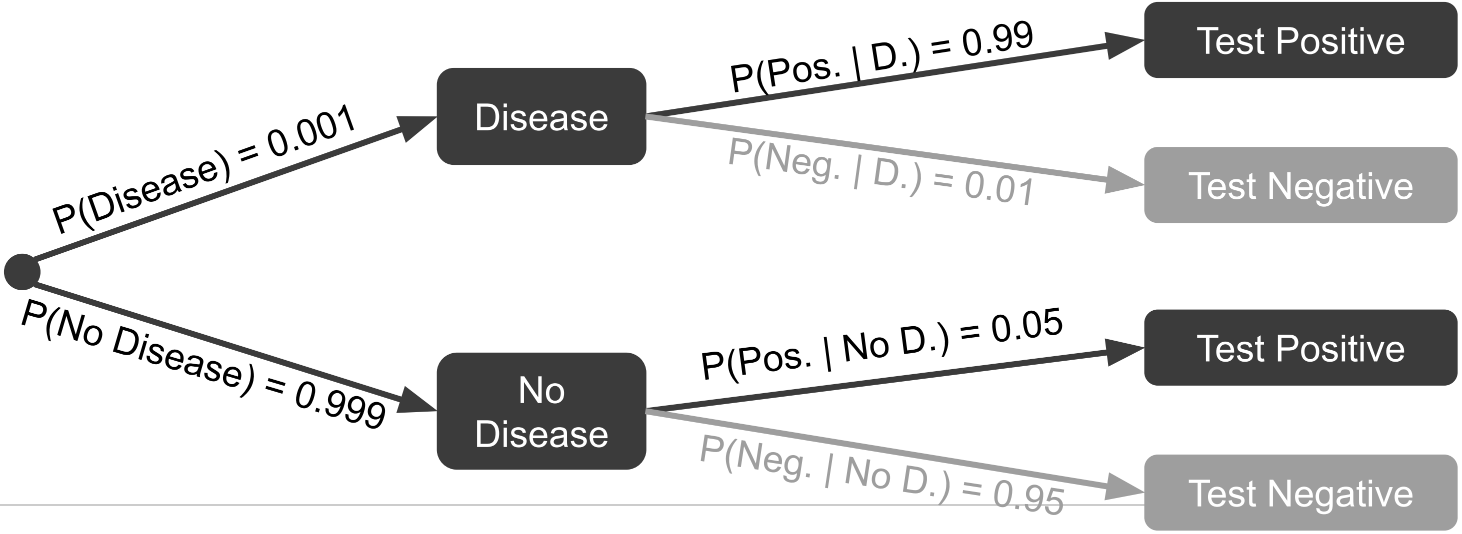
$$P(Pos) = P(Pos, D) + P(Pos, No\ D)$$
$$= \boxed{P(D.)\,P(Pos.\mid D.) + P(No\ D.)\,P(Pos.\mid No\ D.)}$$
$$= (.001)*(0.99) + (0.999)*(0.05) = \mathbf{0.05094}$$

b). You test positive. What's the probability you have the disease?

There's a rare disease with prevalence of 1/1000 in the population.
A test for the disease has a false positive rate of 5% and a false negative rate of 1%.
b). You test positive. What's the prob you have the disease?

$$P(D. \mid Pos.) = \frac{P(D., Pos.)}{P(Pos.)} = \frac{P(D.)\, P(Pos. \mid D.)}{P(Pos.)}$$

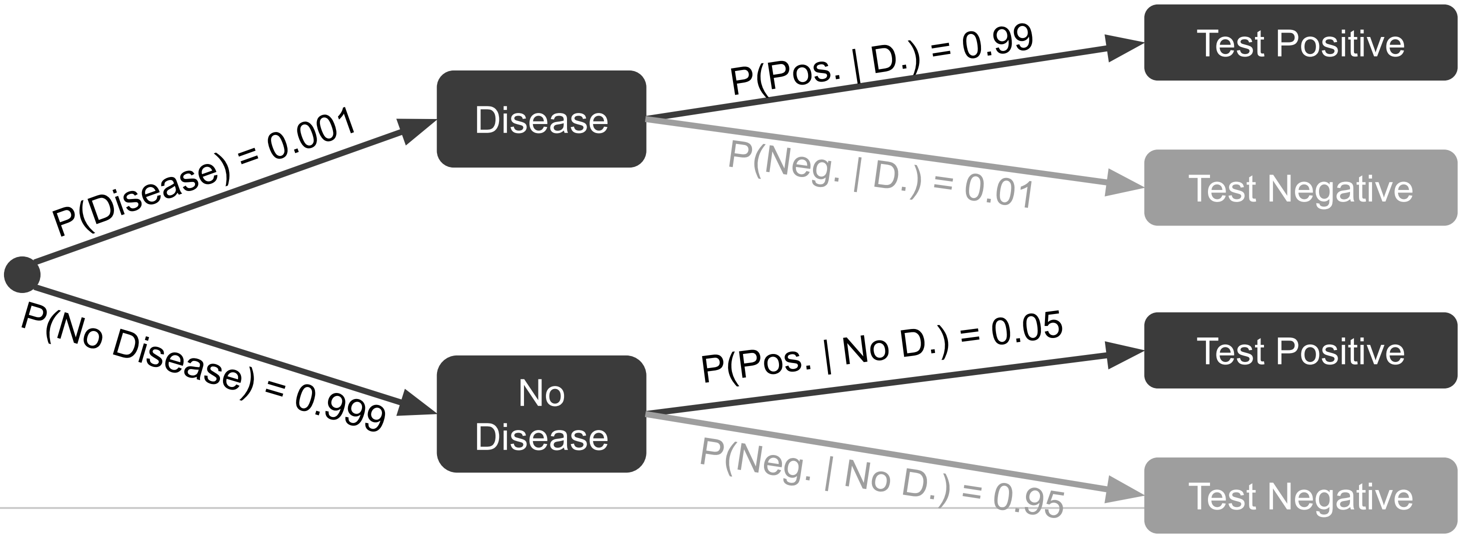
Famous
Equation Alert!
Bayes' Thm



There's a rare disease with prevalence of 1/1000 in the population.
A test for the disease has a false positive rate of 5% and a false negative rate of 1%.
b). You test positive. What's the prob you have the disease?

$$P(D., Pos.) = P(D.)\,P(Pos. \mid D.) = 0.00099$$

$$P(D. \mid Pos.) = \frac{0.00099}{0.00099 + 0.04995} = 0.0194 \approx 2\%$$



Bayes' Theorem

$P(E|F) \Rightarrow P(F|E)$

<u>Thm</u> For any events E and F where $P(E) > 0$ and $P(F) > 0$,

$$P(F|E) = \frac{P(E|F)P(F)}{P(E)}$$

<u>Proof</u>

2 steps!

<u>Expanded form</u>:

$$P(F|E) = \frac{P(E|F)P(F)}{P(E|F)P(F) + P(E|F^C)P(F^C)}$$

<u>Proof</u>

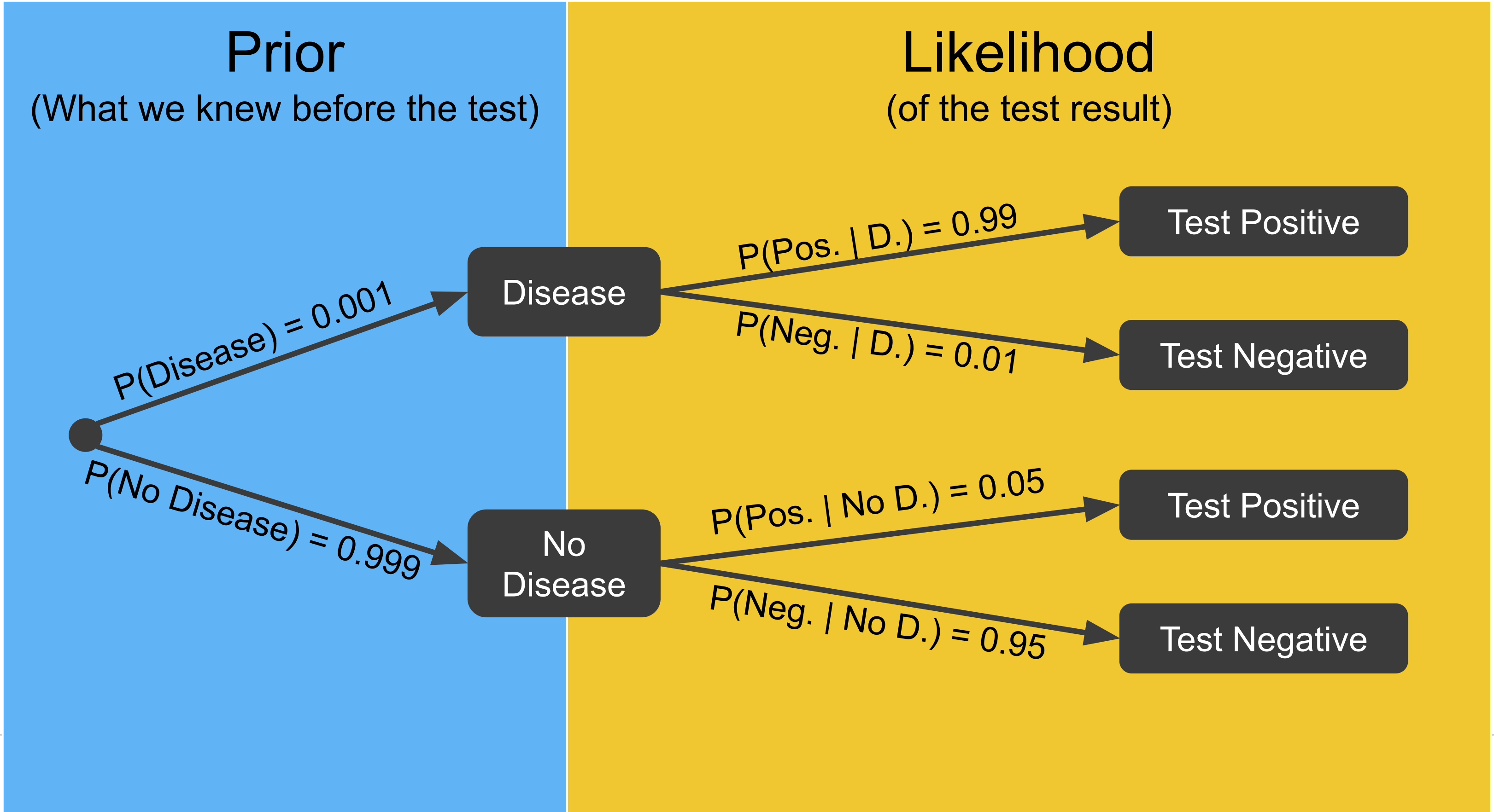
1 more step!

Famous Equation: Bayes' Rule

Bayes' Rule allows us to **update probabilities** by incorporating observations:

$$\underbrace{P(\text{Disease} \mid \text{Pos. Test})}_{\text{Posterior}} = \frac{\underbrace{P(\text{Disease})}_{\text{Prior}} \; \underbrace{P(\text{Pos. Test} \mid \text{Disease})}_{\text{Likelihood (Observation)}}}{P(\text{Pos. Test})}$$

Tree Diagrams and Terminology



Bayes' Theorem

likelihood prior

posterior

$$P(F|E) = \frac{P(E|F)P(F)}{P(E)}$$

normalization constant

Mathematically:

$$P(E|F) \rightarrow P(F|E)$$

Real-life application:

Given new evidence E, update belief of fact F
Prior belief → Posterior belief
$P(F) \rightarrow P(F|E)$

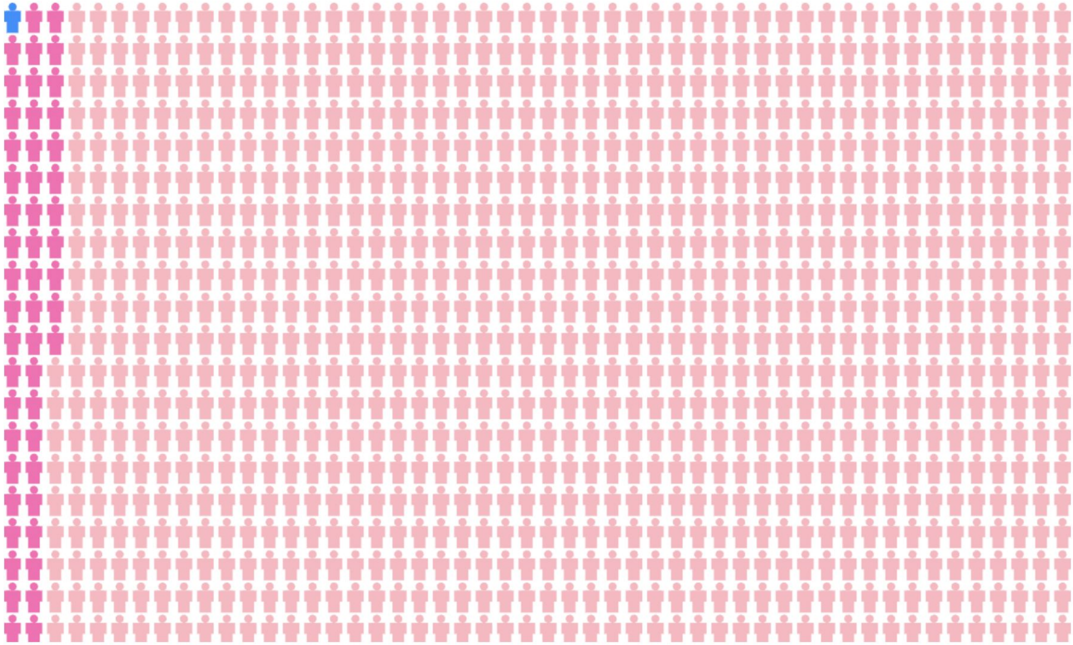
A Closer Look at the Answer

Assume a patient is picked at random.

- Prior probability of disease
  - P(Disease) = 0.001 = one-tenth of 1%
- Posterior probability of disease given positive test
  - P(Disease | Test positive) = 0.0194... ≅ 2%
- Bigger than the prior, but still pretty small
- Should we approve such a test?
  - The test has **low error rates** compared to most tests
- How can this be?

There's a rare disease with prevalence of 1/1000 in the population.
A test for the disease has a false positive rate of 5% and a false negative rate of 1%.
b). You test positive. What's the prob you have the disease?

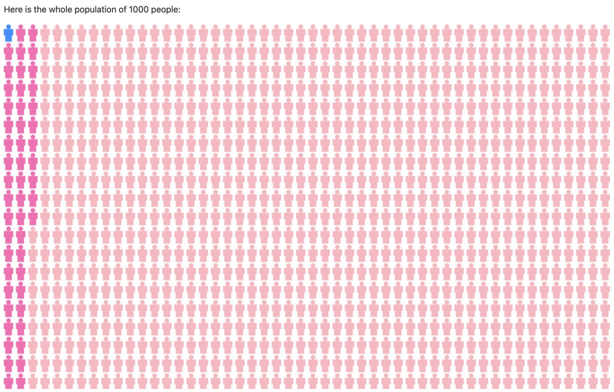
Here is the whole population of 1000 people:

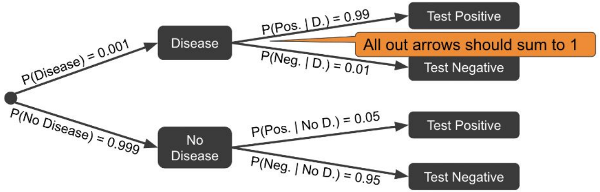


- Out of 1000 people:
  - # of people with disease: $1000*(1/1000) = 1$
  - # of people who test positive and have the disease: $1000*(.001*.99) \cong 1$
  - # of people who test positive and don't have the disease: $1000*(.999*.05) \cong 50$
  - # of people who test positive: $1+50 = 51$

$P(D \mid Pos) \cong 1/51 \cong 2\%$

There’s a rare disease with prevalence of 1/1000 in the population.
A test for the disease has a false positive rate of 5% and a false negative rate of 1%.
b). You test positive. What’s the prob you have the disease?





- Out of 1000 people:

- # of people with disease:
  1000*(1/1000) = 1
- # of people who test positive and have the disease:
  1000*(.001*.99) ≅ 1
- # of people who test positive and don’t have the disease:
  1000*(.999*.05) ≅ 50
- # of people who test positive:
  1+50 = 51

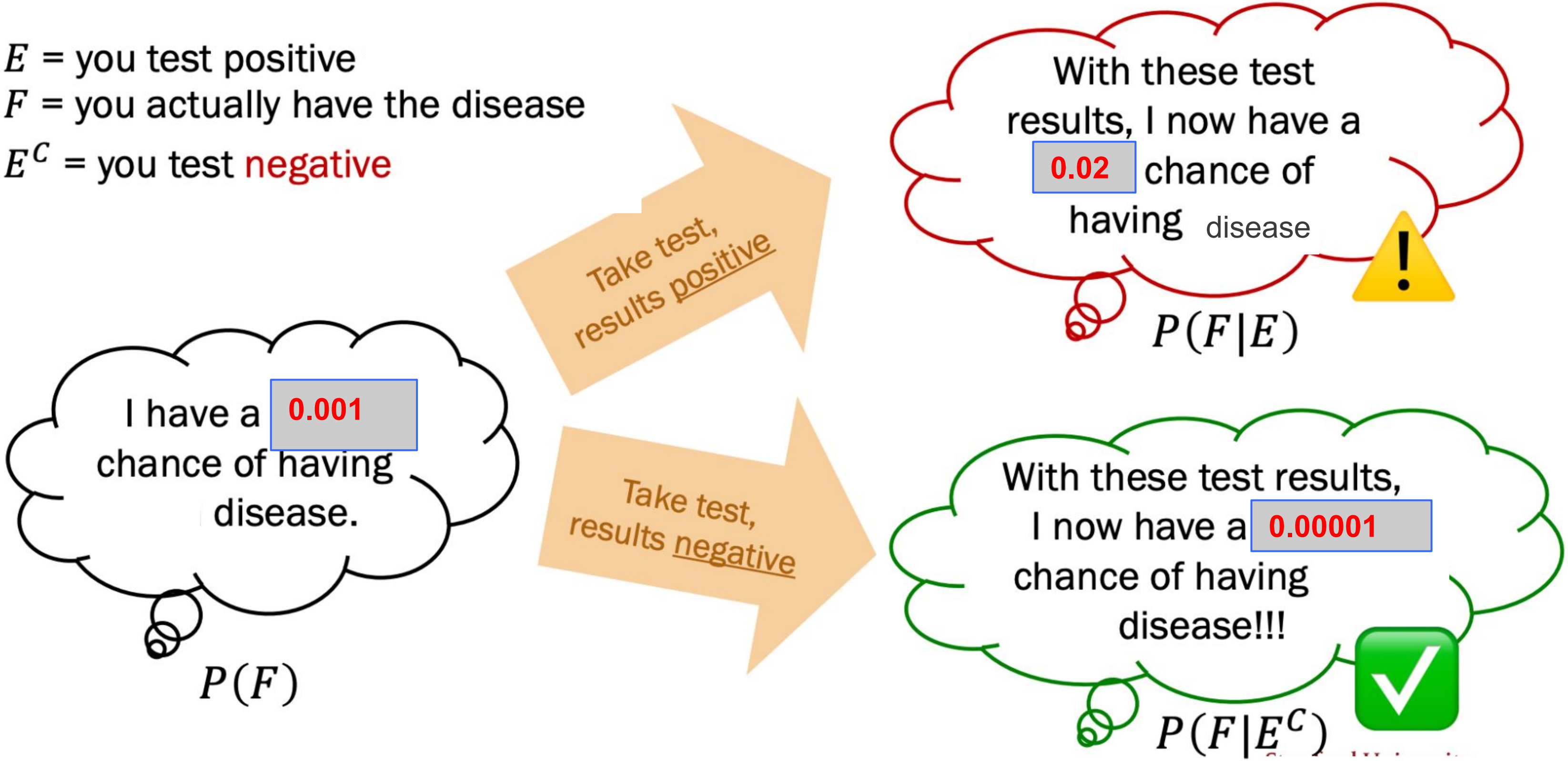
P(D | Pos) ≅ 1/51≅2%

- # of people who test negative and DO have the disease:
  1000*(.001*.01) =0.01 ≅ 0
- # of people who test negative and don’t have the disease:
  1000*(.999*.95) ≅ 949
- # of people who test negative:
  0+949= 949

P(D | Neg)=0.01/949
=0.00001≅0%

Why it's still good to get tested

E = you test positive
F = you actually have the disease
E^C = you test negative

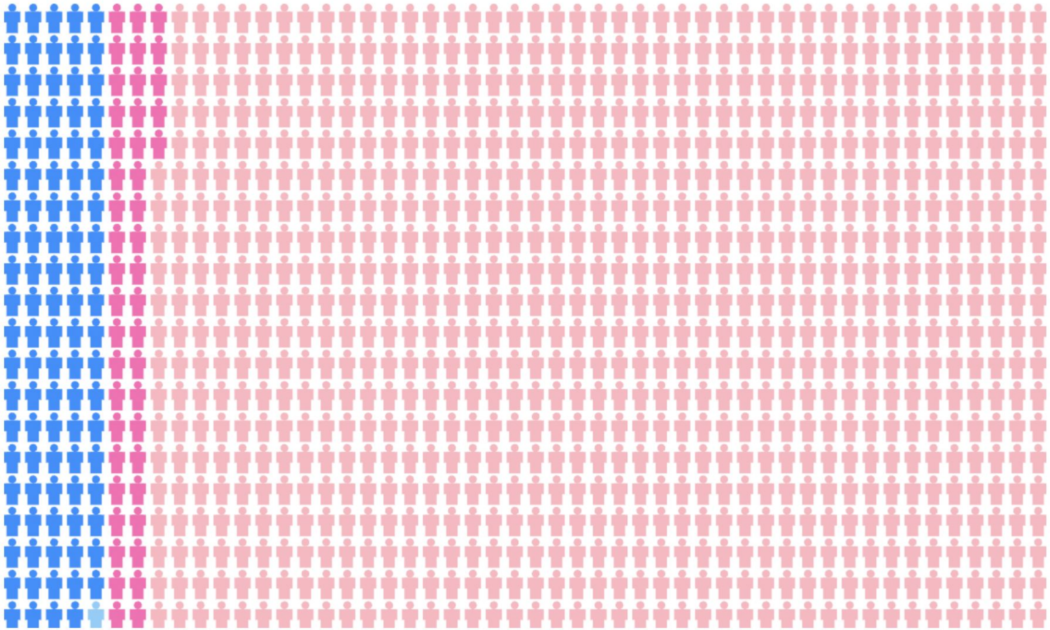


Assumptions Matter

- **“Assume a patient is picked at random.”**
  - But usually, people aren’t picked at random for medical tests
  - So our intuition about randomly picked patients may not be great
- For a ***randomly picked patient, the result does make sense***, because the disease is very rare.
- What if the doctor believes there is a 10% chance the patient has the disease?

There's a rare disease with prevalence of 1/1000 in the population.
A test for the disease has a false positive rate of 5%
and a false negative rate of 1%.
b). You test positive. What's the prob you have the disease?

- What if, based on additional information your doctor believes you are in a subpopulation where there is a 10% chance you have the disease?



- Out of 1000 people:

- # of people with disease:
  1000*(1/10) = 100

- # of people who test positive and have the disease:

  1000*(0.1*0.99) ≅ 99

- # of people who test positive and don't have the disease:

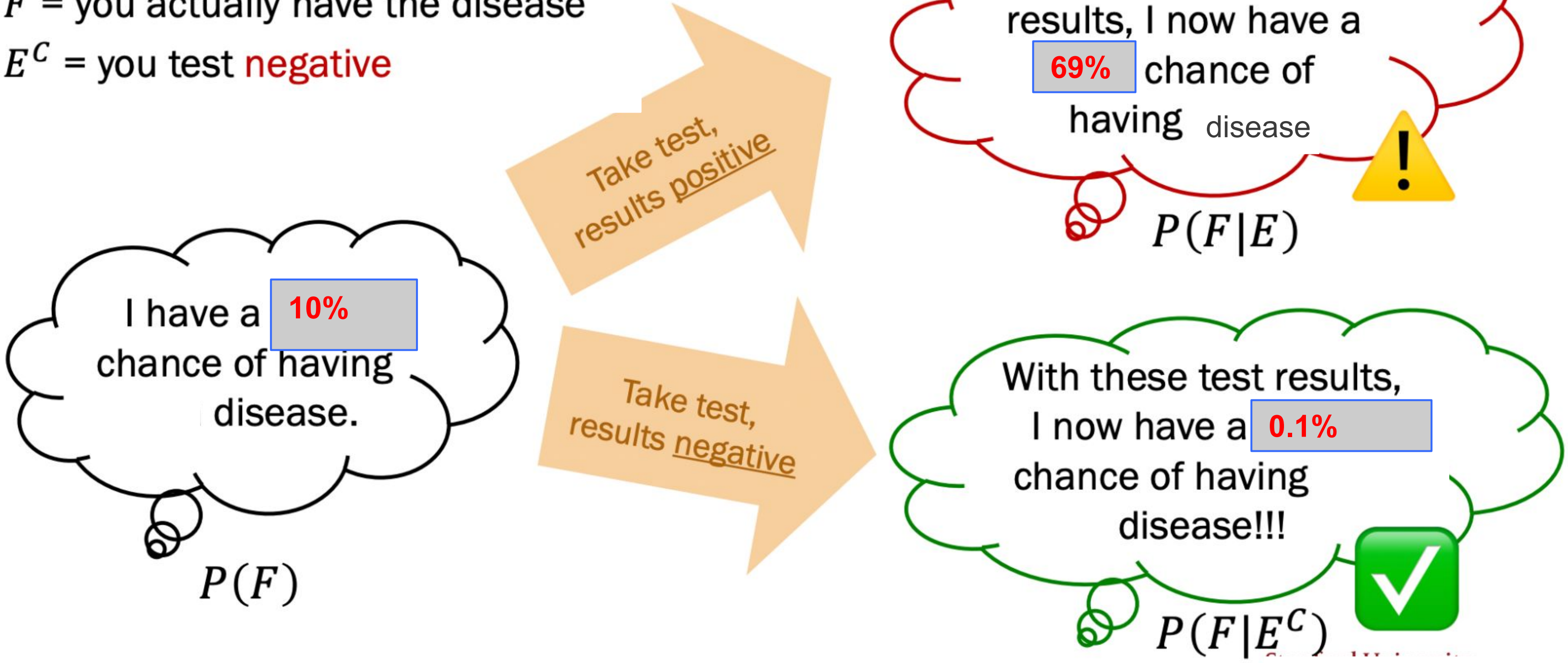
  1000*(.90*.05) ≅ 45

- # of people who test positive:
  99+45 = 144

P(D | Pos) ≅ 99/144 ≅ 69%

Assumptions Matter

E = you test positive
F = you actually have the disease
E^C = you test negative



Subjective Probabilities

posterior likelihood prior

$$P(F|E) = \frac{P(E|F)P(F)}{P(E)}$$

normalization constant

Recall: the probability of an outcome can be defined as:

[Frequentist] The **frequency** with which it will occur in repeated trials:

$$\mathrm{P(Event)} = \lim_{n\to\infty} \frac{\mathrm{count(Event)}}{n}$$

[Bayesian] Someone's **subjective degree** of belief that it will occur

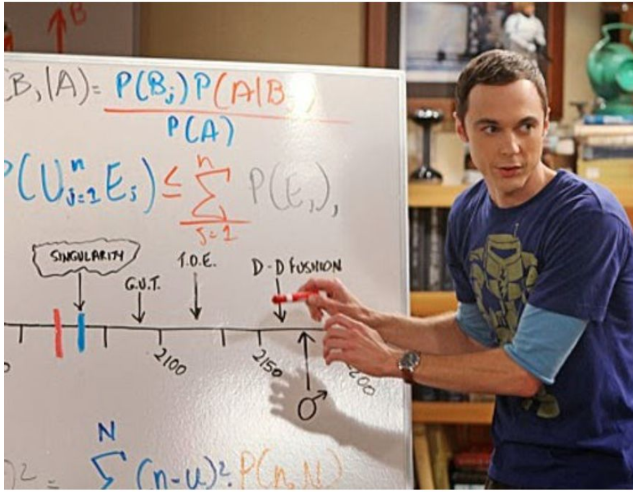
Why use subjective priors?

- To quantify your degree of uncertainty about an outcome, even when there is no physical randomization
  - i.e. chance of CU football team getting into a Bowl game next year
  - i.e. chance of the "Big One" in the next 30 years

Purpose of Bayes' Rule

- Update your prediction based on new information

- In a multi-stage experiment,
  find the chance of an event at an earlier stage,
  given the result of a later stage



More Practice

A car heading from Berkeley to San Francisco is pulled over on the freeway **for speeding**. Which type of car is it more likely to be:

- a **Tesla** which is **relatively common** in California
- or a **Lamborghini** which is **a rare** car that is **known for speeding**

you don't have enough information to **calculate the answer directly.**

What would you guess, and why? Make some reasonable assumptions (data scientists often have to do this) and explain your thought process.

More Practice

A car heading from Berkeley to San Francisco is pulled over on the freeway **for speeding**. Which type of car is it more likely to be:

- a **Tesla** which is **relatively common** in California
- or a **Lamborghini** which is **a rare** car that is **known for speeding**

you don't have enough information to **calculate the answer directly.**

What would you guess, and why? Make some reasonable assumptions (data scientists often have to do this) and explain your thought process.

Let T: Tesla, L: Lamborghini and S: speeding

We're interested in which one is bigger: $P(T|S)$ vs $P(L|S)$

$P(T|S) = P(S|T)P(T)/P(S)$

$P(L|S) = P(S|L)P(L)/P(S)$

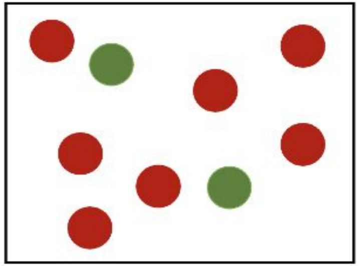
Notice - they both have the same denom, so we're just comparing numerators.

$P(L)$ is much much smaller than $P(T)$

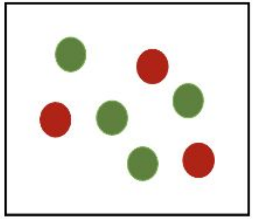
Even if we assume that $P(S|L)$ is greater than $P(T|L)$ it won't be SO much greater to make up for the fact that $P(L)$ is much much smaller than $P(T)$. So we would expect it's more likely to be a TESLA.

More Practice

Suppose we have two boxes filled with green and red balls. Paul selects a ball by first choosing one of the two boxes. Since box 1 is larger he is twice as likely to choose box 1 than he is box 2. He then selects one of the balls in the box at random. What is the probability Paul has selected a red ball?



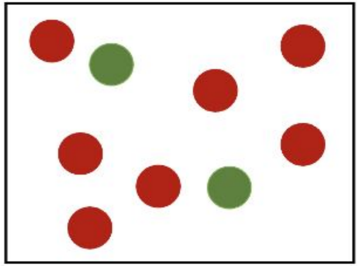
Box 1: 2 greens, 7 reds



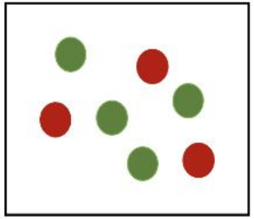
Box 2: 4 greens, 3 reds

More Practice

Suppose we have two boxes filled with green and red balls. Paul selects a ball by first choosing one of the two boxes. Since box 1 is larger he is twice as likely to choose box 1 than he is box 2. He then selects one of the balls in the box at random. What is the probability Paul has selected a red ball?



Box 1: 2 greens, 7 reds



Box 2: 4 greens, 3 reds

Let R: Red, B1: Box 1, B2 :Box 2

$$P(R) = P(R \cap B_1) + P(R \cap B_2)$$

Law of Total Prob

$$P(R) = P(R|B_1)P(B_1) + P(R|B_2)P(B_2)$$

$$\underbrace{(7/9)\cdot(2/3)} + \underbrace{(3/7)\left(\tfrac{1}{3}\right)} = \frac{14}{27} + \frac{1}{7}$$

$\approx .66$