Last time
- Finish Entropy
- Start Cond. Prob.
  - Bayes Law

Today
- Finish Cond. Prob.
  & Bayes Law
- Start Markov Chains
  & Page Rank

Next time
- Finish MC & PR
  - Start Module 4: Algorithms & Analysis
Conditional Probability

\[ \Pr[H \land E] = \Pr[H] \cdot \Pr[E | H] \]

\[ \Rightarrow \Pr[E | H] = \frac{\Pr[H \land E]}{\Pr[H]} \]

\[ = \Pr[E] \cdot \Pr[H | E] \]

\[ \Rightarrow \Pr[H | E] = \frac{\Pr[H \land E]}{\Pr[E]} \]

Bayes Law:

\[ \Pr[E] \cdot \Pr[H | E] = \Pr[H \land E] = \Pr[H] \cdot \Pr[E | H] = \Pr(H) \cdot \Pr[E] \]

\[ \Pr[H | E] = \frac{\Pr[E | H] \cdot \Pr[H]}{\Pr[E]} \]

**H** = hypothesis: do have Zika?

**E** = evidence: did blood test come back pos?
Bayes Law

\[ \Pr(B \mid A) \cdot \Pr(A) = \Pr(A \cap B) = \Pr(A \mid B) \cdot \Pr(B) \]

\[ \Pr(A \mid B) = \frac{\Pr(B \mid A) \cdot \Pr(A)}{\Pr(B)} \]

\[ = \frac{\Pr(B \mid A) \cdot \Pr(A)}{\Pr(B \mid A) \cdot \Pr(A) + \Pr(B \mid \overline{A}) \cdot \Pr(\overline{A})} \]

What is \( \Pr(B) \)?

\[ \Pr(B) = \begin{array}{c} \boxed{\text{\#1}} \end{array} + \boxed{\text{\#2}} \]

\[ = \Pr(B \mid A) \cdot \Pr(A) + \Pr(B \mid \overline{A}) \cdot \Pr(\overline{A}) \]

\[ \Pr(H \mid E) = \frac{\Pr(E \mid H) \cdot \Pr(H)}{\Pr(E \mid H) \cdot \Pr(H) + \Pr(E \mid \overline{H}) \cdot \Pr(\overline{H})} \]
Example: Zika in FL 2016

- Prevalence of Zika in S. FL. \( p(\text{Zika}) = 10^{-5} \) (1 in 100,000)

- Accuracy of blood test is 99%

  i.e. \( p(\text{pos. test} | \text{Zika}) = 0.99 \) ← test subjects who had Zika

  \( p(\text{pos. test} | \text{no Zika}) = 0.01 \) ← control group who don't have Zika

- You test positive: what is the chance that you have Zika?

  \[
  \Rightarrow \frac{\text{Not}}{\text{Not}} \quad p(\text{pos. test} | \text{Zika}) = 0.99
  \]

  \[
  \Rightarrow \text{you want} \quad p(\text{Zika} | \text{pos. test})
  \]
\[
p(zika | \text{pos test}) = \frac{p(\text{pos test} | zika) \cdot p(zika)}{p(\text{pos test})} = \frac{p(\text{pos test} | zika) \cdot p(zika)}{p(\text{pos test} | zika) \cdot p(zika) + p(\text{pos test} | \text{not zika}) \cdot p(\text{not zika})}
\]

\[
= \frac{0.99 \times 10^{-5}}{0.99 \times 10^{-5} + 0.01 \times (1-10^{-5})}
\]

\[
= \frac{0.0000099}{0.0000099 + 0.0099999}
\]

\[
\approx 0.00099
\]

\[
\approx 0.1\% \quad \text{i.e. only 1 in 1,000!}
\]
Seems wildly counterintuitive, but…

10,000,000 people in FL. \(10^7\)

w/ Zika: \(10^7 \cdot 10^{-5} = 10^2 = 100\)

w/o Zika: 9,999,900

\[
\text{test pos w/ Zika: } 100 \cdot 0.99 = 99
\]

\[
\text{test pos. w/o Zika: } 9,999,900 \times 0.01 = 99,999
\]

\[\therefore\] If you test pos., you're about 1,000 times more likely to be among 99,999 who don't have Zika but test pos. than among the 99 who do have Zika and test pos.
Markov Chain

B: Bertucci's
M: Margaritas
S: Sato

Pr[B|B] = 0.7
Pr[M|B] = 0.2
Pr[S|B] = 0.1

Pr[B|M] = 0.2
Pr[M|S] = 0.1
Pr[S|B] = 0.5

3 reds ⇒ must add +0.1