Introduction to uncertainty

Announcements

- Next class (Wednesday, September 21) will be online only (same zoom link).
  - We will do a coding exercise (about minmax)
- Assignment 1 (almost) released

Probabilistic techniques

- AI is full of uncertainty
  - Can’t observe full state of system
  - Observations we can make are noisy
  - Our models of the world are imperfect

Martin-Shepard 2010

- Probabilistic frameworks give us a principled way of dealing with and reasoning about this uncertainty
  - Largely championed by Judea Pearl (2011 Turing Award)

But they're not a silver bullet!

- We'll still face challenges like…
  - Probability distributions that are impossibly complex, with intractably many dimensions
  - Parameter estimation problems that would require exponential amounts of data

- Much work is thus devoted to balancing between what we'd like to model and what we are able to model
  - *Probabilistic graphical models* are a popular framework

Probability 101

Probability definitions

- A *finite probability space* consists of:
  - A finite set *S* of mutually-exclusive *outcomes*
  - A function $P : S \to \mathbb{R}$ such that:

$$P(s) \geq 0, \forall s \in S \qquad \sum_{s \in S} p(s) = 1 \qquad P(\emptyset) = 0$$

- An *event A* is a subset of *S*, $A \subseteq S$.
  - The probability of an event is defined as

$$P(A) = \sum_{s \in A} P(s)$$

Basic identities

- For two events *A* and *B*…
  – What's the probability that either A or B (or both) occur?

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

If A and B are *disjoint*, their intersection is the empty set, and the last term is 0.

Basic identities

- For two events *A* and *B*…
  - What's the probability that either A or B (or both) occur?

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

If A and B are *disjoint,* their intersection is the empty set, and the last term is 0.

$$A = \texttt{even roll} = \{2, 4, 6\}$$
$$B = \texttt{roll} \geq 5 = 5, 6$$
$$P(A \cup B) = ?$$

Basic identities

- For two events *A* and *B*…
  - What's the probability that either A or B (or both) occur?

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

If A and B are *disjoint*, their intersection is the empty set, and the last term is 0.

$$A = \texttt{even roll} = \{2, 4, 6\}$$

$$B = \texttt{roll} \geq 5 = 5, 6$$

$$P(A \cup B) = \frac{3}{6} + \frac{2}{6} - \frac{1}{6} = \frac{2}{3}$$

Super simple example #1

- Suppose you roll a six-sided die 5 times. What’s the probability of rolling a “three” all 5 times?

Super simple example #1

- Suppose you roll a six-sided die 5 times. What’s the probability of rolling a “three” all 5 times?

$$S = \{11111, 11112, 11113...\}$$
$$A = \{33333\}$$
$$P(A) = \frac{|A|}{|S|} = \frac{1}{6^5}$$

Super simple example #2

- Suppose you roll a six-sided die 5 times. What's the probability of rolling a "three" during the first roll?

Super simple example #2

- Suppose you roll a six-sided die 5 times. What’s the probability of rolling a “three” during the first roll?

$$S = \{11111, 11112, 11113...\}$$

$$A = \{31111, 31112...\}$$

$$P(A) = \frac{|A|}{|S|} = \frac{6^4}{6^5} = \frac{1}{6}$$

Super simple example #3

- Suppose you roll a die 5 times. What's the probability of getting at least 1 six?

Example #3 (2^{nd} try)

- Suppose you roll a die 5 times. What's the probability of getting at least 1 six?
  - Answer 2: Sum probabilities of disjoint events

    P(at least 1 six) = P(1 six and 4 non-sixes) +
    P(2 sixes and 3 non-sixes) +
    P(3 sixes and 2 non-sixes) +
    P(4 sixes and 1 non-six) +
    P(5 sixes)
    = …

  - Right, but a lot of work.

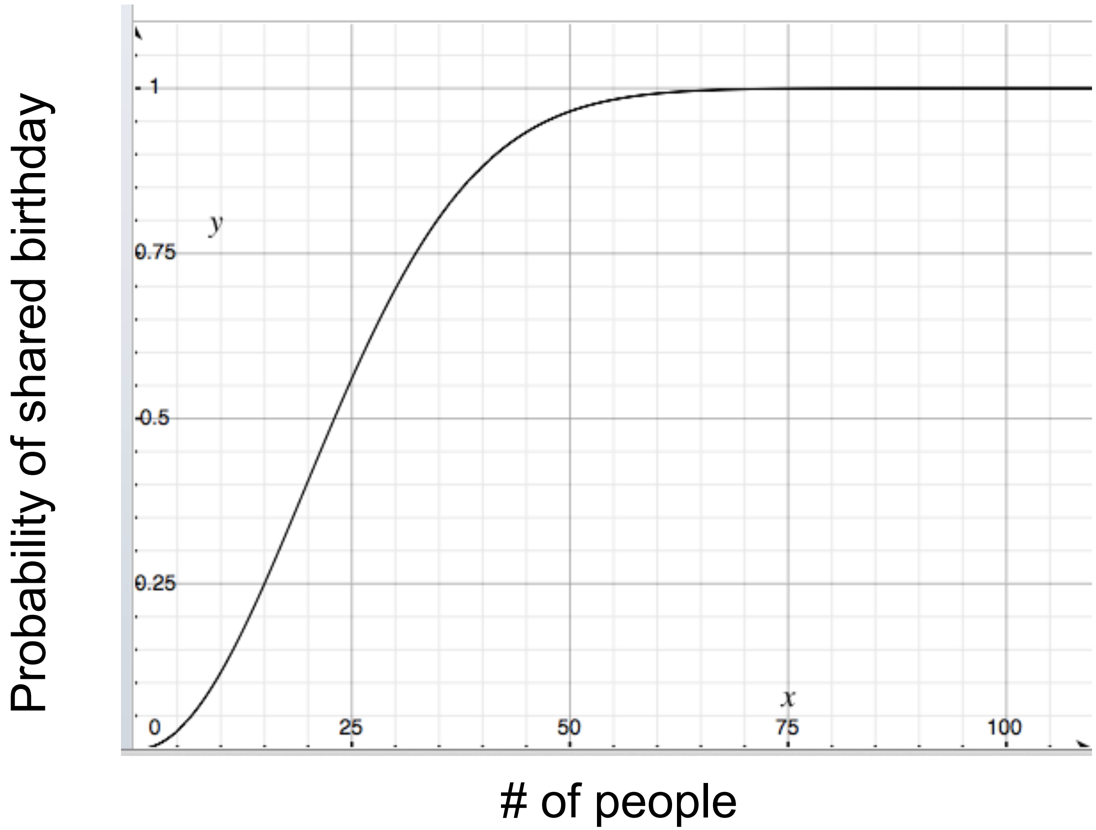
An example (3rd try)

- Suppose you roll a die 5 times. What's the probability of getting at least 1 six?

$$
\begin{aligned}
S &= \{11111, 11112, 11113 \ldots\} \\
A &= \texttt{at least one six} \\
B &= \texttt{no six} \\
A \cap B = \emptyset, A \cup B &= S \\
P(A) = 1 - P(B) = 1 - \left(\frac{5}{6}\right)^5 &\approx 0.6
\end{aligned}
$$

The Birthday Problem

- Given a class of ~25 people, what's the probability that at least two of us share the same birthday?



Conditional probabilities and Bayes’ Law

Conditional probabilities

- Probability that one event occurs, given that another event is known to have occurred
  - Denoted $P(A|B)$. "Probability of A given B"
  - Defined as:

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

Conditional probabilities

- Probability that one event occurs, given that another event is known to have occurred
  - Denoted $P(A|B)$. "Probability of A given B"
  - Defined as:

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

```
A = role an even die
        B = rolling 6
```

$$P(B|A) = \frac{P(\{6\})}{P(\{2,4,6\})} = \frac{1/6}{3/6} = \frac{1}{3}$$

Conditional probabilities

- Probability that one event occurs, given that another event is known to have occurred
  - Denoted $P(B|A)$. "Probability of B given A"
  - Defined as:

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

- Leads directly to the *chain rule*:

$$P(A \cap B) = P(B|A)P(A)$$

  - More generally:

$$P(A_1 \cap A_2 \cap ... \cap A_n) = P(A_1)P(A_2|A_1)...P(A_n|A_1 \cap ... \cap A_{n-1})$$

Independence of events

- Two events are *independent* if $P(A|B) = P(A)$
  - Or, equivalently, if $P(B|A) = P(B)$
  - Independence denoted $A \perp B$

- The joint probability of independent events A and B both occurring is then simply:

$$P(A \cap B) = P(A)P(B)$$

  - This idea of *factoring* a distribution into a product of two simpler distributions will be a recurring theme!

Conditional independence

- Sometimes events are both conditioned on the same event, but otherwise are independent
  - Mary and Bob live in same city but independently
  - A denotes event that it's raining, B denotes event that Mary has an umbrella, C denotes event that Bob has an umbrella
  - Events B and C are **not** independent
  - But B and C are **conditionally independent** given A,

$$P(B|A,C) = P(B|A)$$

$$P(C|A,B) = P(C|A)$$

  - Denoted

$$B \perp C|A$$

Bayes' Law

- For two events *A* and *B*,

Likelihood

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Priors

Posterior

- Useful when you want to know something about A, but all you can directly observe is B
  - This process is called *Bayesian inference*

Bayes’ Law Example #1

- I have two coins, one fair and one with heads on both sides. I choose a coin at random, flip it, and get heads. What is the probability that it is the fair coin?

Bayes’ Law Example #1

- I have two coins, one fair and one with heads on both sides. I choose a coin at random, flip it, and get heads. What is the probability that it is the fair coin?

Bayes' Law Example #1

- I have two coins, one fair and one with heads on both sides. I choose a coin at random, flip it, and get heads. What is the probability that it is the fair coin?

```
F = Fair
H = Heads
```

$$P(F|H) = \frac{P(H|F)P(F)}{P(H)} = \frac{\frac{1}{2}\frac{1}{2}}{\frac{3}{4}} = \frac{1}{3}$$

$$P(H|F) = \frac{1}{2}$$

$$P(F) = \frac{1}{2}$$

$$P(H) = P(H \cap F) + P(H \cap \bar{F})$$

$$= P(H|F)P(F) + P(H|\bar{F})P(\bar{F})$$

$$= \frac{1}{2}\frac{1}{2} + 1\frac{1}{2} = \frac{3}{4}$$

Bayes’ Law Example #2

- A doctor says you have an illness that afflicts 0.01% of the population. Her diagnoses are right 99% of the time. What’s the probability that you have the illness?

Bayes' Law Example #2

- A doctor says you have an illness that afflicts 0.01% of the population. Her diagnoses are right 99% of the time. What's the probability that you have the illness?

```
i = Ill
p = Positive
```

$$P(i|p) = \frac{P(p|i)P(i)}{P(p)} = \frac{(0.99)(0.0001)}{0.010098} \approx 0.0098$$

$$P(p|i) = 0.99$$

$$P(i) = 0.0001$$

$$\begin{aligned} P(p) &= P(p \cap i) + P(p \cap \bar{i}) \\ &= P(p|i)P(i) + P(p|\bar{i})P(\bar{i}) \\ &= (0.99)(0.0001) + (0.01)(0.9999) \approx 0.010098 \end{aligned}$$

Next class

- **Inference on Bayes Nets**