Probability

2.3 Independence

Anna Karlin
Most slides by Alex Tsun
Agenda

- Chain Rule
- Independence
- Conditional Independence not in class.
Chain Rule (Idea)

Have a Standard 52-Card Deck.
- 4 Suits (Clubs, Diamonds, Hearts, Spades)
- 13 ranks (A, 2, 3, ..., 9, 10, J, Q, K)
Chain Rule (Idea)

Have a Standard 52-Card Deck. Shuffle It, and draw the top 3 cards.

What is \( P(A, B, C) \)?

A: Ace of Spades First
B: 10 of Clubs Second
C: 4 of Diamonds Third
**Chain Rule (Idea)**

Have a Standard 52-Card Deck. Shuffle It, and draw the top 3 cards.

What is $P(\text{A, B, C}) = P(A, B, C)$?

- **A**: Ace of Spades First
- **B**: 10 of Clubs Second
- **C**: 4 of Diamonds Third

$P(A) \cdot P(B|A) \cdot P(C|A \land B)$
Chain Rule (Idea)

Have a Standard 52-Card Deck. Shuffle It, and draw the top 3 cards. (Uniform probability space).

What is \( P(\text{A}, \text{B}, \text{C}) \) = \( P(\text{A}, \text{B}, \text{C}) \)?

\[
\begin{align*}
P(A) & \cdot P(B|A) \cdot P(C|A, B) \\
\frac{1}{52} & \cdot \frac{1}{51} \cdot \frac{1}{50}
\end{align*}
\]

A: Ace of Spades First
B: 10 of Clubs Second
C: 4 of Diamonds Third

A, B, C \overset{\text{AND}}{\Rightarrow} A \cap B \cap C
Chain Rule

**Chain Rule:** Let $A_1, ..., A_n$ be events with nonzero probability. Then,

$$P(A_1, ..., A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1,A_2) ... P(A_n|A_1, ..., A_{n-1})$$

In the case of two events $A, B$,

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

An easy way to remember this formula: we need to do $n$ tasks, so we can perform them one at a time, conditioning on what we’ve done so far.
**Chain Rule (Idea)**

Have a standard 52-card deck. Shuffle it, and draw the top 3 cards. (Uniform probability space).

What is $P(\text{A, B, C}) = P(A, B, C)$?

$P(A) = \frac{1}{52}$

$P(B|A) = \frac{1}{51}$

$P(C|A,B) = \frac{1}{50}$

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

$P(B) = P(2^{nd} \text{ card is 10 of clubs})$

$P(C) = \frac{10C}{51 \cdot 50}$

A: Ace of Spades first

B: 10 of Clubs second

C: 4 of Diamonds third
Two people, A and B, are playing the following game.
A 6-sided die is thrown and each time it’s thrown, regardless of the history, it is equally likely to show any of the six numbers.
If it shows 5, A wins.
If it shows 1, 2 or 6, B wins.
Otherwise, they play a second round and so on.

What is \( \text{Pr}(A \text{ wins on } 4^{\text{th}} \text{ round})? \)

**Fun with Cards**

\[
\begin{align*}
\text{P}(B) &= \frac{51 \cdot 50}{52 \cdot 51.50} = \frac{1}{52} \\
\text{P}(B|A) &= \frac{\text{P}(A\cap B)}{\text{P}(A)} = \frac{1 \cdot 51}{51.50} = \frac{50}{51.50} = \frac{1}{51}
\end{align*}
\]
Quick question: In general, is

$$P(A \land B) = P(A, B) = P(A)P(B)?$$
The need for independence

Quick question: In general, is

$$P(A, B) = P(A)P(B)?$$

The chain rule says

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

So no, unless the special case when $$P(B|A) = P(B)$$. This case is so important it has a name.
Independence: Events $A, B$ are independent if any of the three equivalent conditions hold:

1. $P(A|B) = P(A)$
2. $P(B|A) = P(B)$
3. $P(A, B) = P(A)P(B)$
Independence

- Toss a coin 3 times. Each of 8 outcomes equally likely.

Define
- \( A = \{ \text{at most one T} \} = \{ \text{HHH, HHT, HTH, THH} \} \)
- \( B = \{ \text{at most 2 Heads} \} = \{ \text{HHH} \}^c \)

- Are \( A \) and \( B \) independent?

\[
\begin{align*}
A \cap B &= \{ \text{HHT, HTH, THH} \} \\
\Pr(A) \cdot \Pr(B) &\neq \Pr(A \cap B)
\end{align*}
\]

\[
\begin{array}{ccc}
P(A) & P(B) & P(A \cap B) \\
\hline
\frac{3}{4} & \frac{7}{8} & \frac{21}{32}
\end{array}
\]
Network Communication

Each link works with the probability given, independently. What's the probability A and D can communicate?

T: top path works
   \[ P(T) = P(AB \cap BD) = P(AB) P(BD) = pq \]

B: bottom path works
   \[ P(B) = r \cdot s \]

\[ P(3 \text{ working path}) = P(T \cup B) = P(T) + P(B) - P(T \cap B) = pq + rs - P(T) P(B) = pq + rs - \frac{pq \cdot rs}{pq \cdot rs} \]

\[ 1 - P(\text{no path works}) = P(\overline{T} \cup \overline{B}) = P(\overline{T}) P(\overline{B}) = (1-pq)(1-rs) \]

\[ P(C \cup D) = P(C) + P(D) - P(C \cap D) \]

when C \& D are independent

\[ P(C \cap D) = P(C) P(D) \]

C, D are mutually exclusive event

a) indep
b) not indep
Network Communication

Each link works with the probability given, independently. What's the probability A and D can communicate?

\[ P(\text{top}) = P(AB \cap BD) = P(AB)P(BD) = pq \]
\[ P(\text{bottom}) = P(AC \cap CD) = P(AC)P(CD) = rs \]

\[ P(\text{top} \cup \text{bottom}) = P(\text{top}) + P(\text{bottom}) - P(\text{top} \cap \text{bottom}) \]
\[ = P(\text{top}) + P(\text{bottom}) - P(\text{top})P(\text{bottom}) \]
\[ = pq + rs - pqrst = 0 \]
Using Independence to Define a Probabilistic Model

- We can define our probability model via independence.

- Example: suppose a biased coin comes up heads with probability 2/3, independent of other flips.

- Sample space: sequences of 3 coin tosses.

- \( \Pr(\text{HHH}) = ? \)
- \( \Pr(\text{TTT}) = ? \)
- \( \Pr(\text{HHT}) = ? \)
- \( \Pr(\text{HTH}) = ? \)
- \( \Pr(\text{2 heads}) = ? \)

\[
\mathcal{S} = \{\text{HHH}, \text{HHT}, \text{HTH}, \text{TTT}\}
\]

\[
\begin{align*}
\Pr(\text{HHH}) &= (\frac{2}{3})^3 \\
\Pr(\text{TTT}) &= (\frac{1}{3})^3 \\
\Pr(\text{HHT}) &= \frac{2}{3} \cdot \frac{1}{3} \cdot \frac{2}{3} = \frac{4}{9} \\
\Pr(\text{HTH}) &= \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{1}{3} = \frac{4}{9} \\
\Pr(\text{2 heads}) &= \frac{2}{3} \cdot \frac{1}{3} \cdot \frac{2}{3} + \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{1}{3} = \frac{4}{9} + \frac{4}{9} = \frac{8}{9}
\end{align*}
\]
Probability

3.1 Discrete Random Variables Basics

Anna Karlin

Most slides by Alex Tsun
Agenda

- Intro to Discrete Random Variables
- Probability Mass Functions
- Cumulative Distribution function
Flipping two coins

\[ \Omega = \{HH, HT, TH, TT\} \]

Let \( X \) be the number of heads in two independent flips of a fair coin.
Flipping two coins

$\Omega = \{HH, HT, TH, TT\}$

Let $X$ be the number of heads in two independent flips of a fair coin.

$X$ is a function, $X: \Omega \to \mathbb{R}$ which takes outcomes $\omega \in \Omega$ and maps them to a number.

$\mathbb{S}_X = \{0, 1, 2\}$
Suppose we conduct an experiment with sample space $\Omega$. A **random variable (rv)** is a numeric function of the outcome, $X: \Omega \to \mathbb{R}$. That is, it maps outcomes $\omega \in \Omega$ to numbers, $\omega \mapsto X(\omega)$. 
Random Variable

Suppose we conduct an experiment with sample space \( \Omega \). A **random variable (rv)** is a numeric function of the outcome, \( X : \Omega \rightarrow \mathbb{R} \). That is, it maps outcomes \( \omega \in \Omega \) to numbers, \( \omega \mapsto X(\omega) \).

The set of possible values \( X \) can take on is its **range/support**, denoted \( \Omega_X \).