Probability

2.3 Independence

Anna Karlin
Most slides by Alex Tsun
Agenda

- Chain Rule
- Independence
- Conditional Independence
**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 \( P(\text{A, B, C}) \) = \( P(A, B, C) \)?

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

\[
\frac{1}{52} \cdot \frac{1}{51} \cdot \frac{1}{50}
\]
Chain Rule (Idea)

Have a standard 52-card deck. Shuffle it, and draw the top 3 cards.

(uniform probability space).

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

\[
\frac{1}{52} \cdot \frac{1}{51} \cdot \frac{1}{50}
\]

\( P(A) \cdot P(B|A) \cdot P(C|A, B) \)

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

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

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

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**.

(U**niform probability space**).

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

\[
P(A) \cdot P(B|A) \cdot P(C|A, B)
\]

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 Pr(A wins on 4th round)?
The need for independence

Quick question: In general, is

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

**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 = \{at most one T\} = \{HHH, HHT, HTH, THH\}
  - B = \{at most 2 Heads\} = \{HHH\}^c

- Are A and B independent?
Network Communication

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

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

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

\[ P(\text{top} \cup \text{bottom}) = P(\text{top}) + P(\text{bottom}) - P(\text{top} \cap \text{bottom}) \]
\[ = P(\text{top}) + P(\text{bottom}) - P(\text{top})P(\text{bottom}) \]
\[ = pq + rs - pqrs \]
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.

- \( \Pr (HHH) = ? \)
- \( \Pr (TTT) = ? \)
- \( \Pr (HHT) = ? \)
- \( \Pr (HTH) = ? \)
- \( \Pr (2 \text{ heads}) = ? \)
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.
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, $\omega \mapsto X(\omega)$. 
Random Variable

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, $\omega \mapsto X(\omega)$.

The set of possible values $X$ can take on is its **range/support**, denoted $\Omega_X$. 
20 balls numbered 1..20

- Draw a subset of 3 uniformly at random.
- Let $X =$ maximum of the numbers on the 3 balls.
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, $\omega \mapsto X(\omega)$.

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

If $\Omega_X$ is finite or countably infinite (typically integers or a subset), $X$ is a discrete random variable (drv). Else if $\Omega_X$ is uncountably large (the size of real numbers), $X$ is continuous random variable.
**Identify those RVs**

For each of the following random variables, identify its range $\Omega_X$ and whether it is discrete or continuous.

<table>
<thead>
<tr>
<th>RV Description</th>
<th>Range</th>
<th>drv or crv?</th>
</tr>
</thead>
<tbody>
<tr>
<td>The number of heads in $n$ flips of a fair coin.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The number of people born this year.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The number of flips of a fair coin up to and including my first head.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The amount of time I wait for the next bus in seconds.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Random Picture
Flipping two coins

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

$X(HH) = 2 \quad X(HT) = 1 \quad X(TH) = 1 \quad X(TT) = 0$

What is the support/range $\Omega_x$? $\Omega_x = \{0, 1, 2\}$
Flipping two coins

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

\[ X(HH) = 2 \quad X(HT) = 1 \quad X(TH) = 1 \quad X(TT) = 0 \]

What is the support/range \( \Omega_X \)? \( \Omega_X = \{0, 1, 2\} \)

But what are the probabilities \( X \) takes on these values? For this, we define the **probability mass function (pmf)** of \( X \), as \( p_X: \Omega_X \rightarrow [0, 1] \)

\[ p_X(k) = P(X = k) \]
Flipping two coins

\[ \Omega = \{HH, HT, TH, TT\} \]
\[ X(HH) = 2 \quad X(HT) = 1 \quad X(TH) = 1 \quad X(TT) = 0 \]

What is the support/range \( \Omega_X \)? \( \Omega_X = \{0, 1, 2\} \)

But what are the probabilities \( X \) takes on these values? For this, we define the **probability mass function (pmf)** of \( X \), as \( p_X : \Omega_X \to [0, 1] \)

\[ p_X(k) = P(X = k) \]

\[ p_X(k) = \begin{cases} 
 0 & k = 0 \\
 1 & k = 1 \\
 2 & k = 2 
\end{cases} \]
Flipping two coins

\[ \Omega = \{HH, HT, TH, TT\} \]
\[ X(HH) = 2 \quad X(HT) = 1 \quad X(TH) = 1 \quad X(TT) = 0 \]

What is the support/range \( \Omega_X \)? \( \Omega_X = \{0, 1, 2\} \)

But what are the probabilities \( X \) takes on these values? For this, we define the **probability mass function (pmf)** of \( X \), as \( p_X : \Omega_X \to [0,1] \)

\[ p_X(k) = P(X = k) \]

\[ p_X(k) = \begin{cases} 1/4, & k = 0 \\ 1/2, & k = 1 \\ 1/4, & k = 2 \end{cases} \]
The **probability mass function (pmf)** of a discrete random variable $X$ assigns probabilities to the possible values of the random variable. That is, $p_X: \Omega_x \rightarrow [0,1]$ where

\[ p_X(k) = P(X = k) \]
The probability mass function (pmf) of a discrete random variable $X$ assigns probabilities to the possible values of the random variable. That is, $p_x: \Omega_x \rightarrow [0,1]$ where

$$p_x(k) = P(X = k)$$

Note that $\{X = a\}$ for $a \in \Omega_x$ form a partition of $\Omega$, since each outcome $\omega \in \Omega$ is mapped to exactly one number.
The **probability mass function (pmf)** of a discrete random variable $X$ assigns probabilities to the possible values of the random variable. That is, $p_X : \Omega_X \rightarrow [0,1]$ where

$$p_X(k) = P(X = k)$$

Note that $\{X = a\}$ for $a \in \Omega_X$ form a partition of $\Omega$, since each outcome $\omega \in \Omega$ is mapped to exactly one number. Hence,

$$\sum_{z \in \Omega_X} p_X(z) = 1$$
20 balls numbered 1..20

- Draw a subset of 3 uniformly at random.
- Let $X =$ maximum of the numbers on the 3 balls.

- $\Pr(X = 20)$
- $\Pr(X = 18)$
- $\Pr(X < 17)$
Flip a biased coin until get heads (flips independent)

With probability $p$ of coming up heads
Keep flipping until the first Heads observed.

Let $X$ be the number of flips until done.

- $\Pr(X = 1)$
- $\Pr(X = 2)$
- $\Pr(X = k)$
Flip a biased coin independently

Probability $p$ of coming up heads, $n$ coin flips
$X$: number of Heads observed.

- $\Pr(X = k)$
Flip a biased coin independently

- Probability $p$ of coming up heads, $n$ coin flips
- $X$: number of Heads observed.
The cumulative distribution function (CDF) of a random variable $F_X(x)$ specifies for each possible real number $x$, the probability that $X \leq k$, that is

$$F_X(x) = P(X \leq k)$$

Ex: if $X$ has probability mass function given by:

- $p(1) = \frac{1}{4}$
- $p(2) = \frac{1}{2}$
- $p(3) = \frac{1}{8}$
- $p(4) = \frac{1}{8}$

$$F(a) = \begin{cases} 0 & a < 1 \\ \frac{1}{4} & 1 \leq a < 2 \\ \frac{3}{4} & 2 \leq a < 3 \\ \frac{7}{8} & 3 \leq a < 4 \\ 1 & 4 \leq a \end{cases}$$
Homeworks of 3 students returned randomly

- Each permutation equally likely
- X: # people who get their own homework

<table>
<thead>
<tr>
<th>Prob</th>
<th>Outcome w</th>
<th>X(w)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/6</td>
<td>1 2 3</td>
<td>3</td>
</tr>
<tr>
<td>1/6</td>
<td>1 3 2</td>
<td>1</td>
</tr>
<tr>
<td>1/6</td>
<td>2 1 3</td>
<td>1</td>
</tr>
<tr>
<td>1/6</td>
<td>2 3 1</td>
<td>0</td>
</tr>
<tr>
<td>1/6</td>
<td>3 1 2</td>
<td>0</td>
</tr>
<tr>
<td>1/6</td>
<td>3 2 1</td>
<td>1</td>
</tr>
</tbody>
</table>