CSE 312
Foundations of Computing II

Lecture 7: More on independence; start random variables

slido.com/4171468
for polls and anonymous questions
Agenda

• Recap
• Independence As An Assumption
• Conditional Independence

• New Topic: Random Variables
**Bayes Theorem with Law of Total Probability**

**Bayes Theorem with LTP:** Let \( E_1, E_2, \ldots, E_n \) be a partition of the sample space, and \( F, G \) events. Then,

\[
P(G|F) = \frac{P(F|G)P(G)}{P(F)} = \frac{P(F|G)P(G)}{\sum_{i=1}^{n} P(F|E_i)P(E_i)}
\]

**Simple Partition:** In particular, if \( E \) is an event with non-zero probability, then

\[
P(G|F) = \frac{P(F|G)P(G)}{P(F|E)P(E) + P(F|E^C)P(E^C)}
\]
Chain Rule

\[ P(B|A) = \frac{P(A \cap B)}{P(A)} \quad \Rightarrow \quad P(A)P(B|A) = P(A \cap B) \]

**Theorem. (Chain Rule)** For events \( A_1, A_2, \ldots, A_n \),

\[ P(A_1 \cap \cdots \cap A_n) = P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1 \cap A_2) \]

\[ \cdots P(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1}) \]

An easy way to remember: We have \( n \) events and we can evaluate their probabilities sequentially, conditioning on the occurrence of previous events.
Independence

**Definition.** Two events $\mathcal{A}$ and $\mathcal{B}$ are (statistically) **independent** if

$$P(\mathcal{A} \cap \mathcal{B}) = P(\mathcal{A}) \cdot P(\mathcal{B}).$$

Alternatively,

- If $P(\mathcal{A}) \neq 0$, equivalent to $P(\mathcal{B} | \mathcal{A}) = P(\mathcal{B})$
- If $P(\mathcal{B}) \neq 0$, equivalent to $P(\mathcal{A} | \mathcal{B}) = P(\mathcal{A})$

"The probability that $\mathcal{B}$ occurs after observing $\mathcal{A}$" -- Posterior
= "The probability that $\mathcal{B}$ occurs" -- Prior
Agenda

• Recap
• Independence As An Assumption
• Conditional Independence

• New Topic: Random Variables
Independence as an assumption

- People often assume it **without justification.**
- Example: A sky diver has two chutes
  
  \[ A : \text{event that the main chute doesn’t open} \quad \mathbb{P}(A) = 0.02 \]
  
  \[ B : \text{event that the backup doesn’t open} \quad \mathbb{P}(B) = 0.1 \]
  
- What is the chance that at least one opens assuming independence?

\[
\mathbb{P}(\text{at least one opens}) = 1 - \mathbb{P}(\text{neither open}) \\
= 1 - \mathbb{P}(A \cap B) \\
= 1 - \mathbb{P}(A) \mathbb{P}(B) \\
= 1 - 0.02 \cdot 0.1 \\
= 0.998
\]
Independence as an assumption

- People often assume it **without justification**.
- Example: A sky diver has two chutes
  
  $A$ : event that the main chute doesn’t open $\quad \mathbb{P}(A) = 0.02$
  $B$ : event that the backup doesn’t open $\quad \mathbb{P}(B) = 0.1$

- What is the chance that at least one opens assuming independence? 

\[
\mathbb{P}(A \cup B) = 1
\]

Assuming independence doesn’t justify the assumption! Both chutes could fail because of the same rare event e.g., freezing rain.
Corollaries of independence of two events

- Example: A sky diver has two chutes

  \[ A : \text{event that the main chute doesn't open} \quad \mathbb{P}(A) = 0.02 \]
  \[ B : \text{event that the backup doesn't open} \quad \mathbb{P}(B) = 0.1 \]

- What is the chance that both open assuming independence?

\[
P(A \cap B) = 1 - P(A \cup B) = 1 - \left[ P(A) + P(B) - P(A \cap B) \right]
\]
\[
= 1 - P(A) - P(B) + P(A)P(B)
\]
\[
= (1 - P(A))(1 - P(B)) = \mathbb{P}(\overline{A}) \mathbb{P}(\overline{B})
\]

\[ A, B \text{ independent} \Rightarrow A \text{ and } B \text{ are independent} \]
Agenda

- Recap
- Sometimes Independence Occurs for Nonobvious Reasons
- Independence As An Assumption
- Conditional Independence

- New Topic: Random Variables
Conditional Independence

**Definition.** Two events $\mathcal{A}$ and $\mathcal{B}$ are independent conditioned on $\mathcal{C}$ if $\mathbb{P}(\mathcal{C}) \neq 0$ and $\mathbb{P}(\mathcal{A} \cap \mathcal{B} | \mathcal{C}) = \mathbb{P}(\mathcal{A} | \mathcal{C}) \cdot \mathbb{P}(\mathcal{B} | \mathcal{C})$.

Plain Independence. Two events $\mathcal{A}$ and $\mathcal{B}$ are independent if

$$
\mathbb{P}(\mathcal{A} \cap \mathcal{B}) = \mathbb{P}(\mathcal{A}) \cdot \mathbb{P}(\mathcal{B}).
$$

Equivalence:

- If $\mathbb{P}(\mathcal{A}) \neq 0$, equivalent to $\mathbb{P}(\mathcal{B} | \mathcal{A}) = \mathbb{P}(\mathcal{B})$
- If $\mathbb{P}(\mathcal{B}) \neq 0$, equivalent to $\mathbb{P}(\mathcal{A} | \mathcal{B}) = \mathbb{P}(\mathcal{A})$
Conditional Independence

**Definition.** Two events $A$ and $B$ are independent conditioned on $C$ if $P(C) \neq 0$ and

$$P(A \cap B \mid C) = P(A \mid C) \cdot P(B \mid C).$$

**Equivalence:**

- If $P(A \cap C) \neq 0$, equivalent to $P(B \mid A \cap C) \equiv P(B \mid C)$
- If $P(B \cap C) \neq 0$, equivalent to $P(A \mid B \cap C) \equiv P(A \mid C)$
Example – More coin tossing

Suppose there is a coin $C_1$ with $\Pr(\text{Head}) = 0.3$ and a coin $C_2$ with $\Pr(\text{Head}) = 0.9$. We pick one randomly with equal probability and flip that coin twice independently. What is the probability both tosses heads?

$$\Pr(HH) = \Pr(HH \mid C_1) \Pr(C_1) + \Pr(HH \mid C_2) \Pr(C_2)$$

$$= \Pr(H \mid C_1) \Pr(H \mid C_1) \Pr(C_1) + \Pr(H \mid C_2) \Pr(H \mid C_2) \Pr(C_2)$$
Example – More coin tossing

Suppose there is a coin C1 with \( \Pr(\text{Head}) = 0.3 \) and a coin C2 with \( \Pr(\text{Head}) = 0.9 \). We pick one randomly with equal probability and flip that coin 2 times independently. What is the probability we get all heads?

\[
\Pr(HH) = \Pr(HH | C1) \Pr(C1) + \Pr(HH | C2) \Pr(C2) \\
= \Pr(H | C1)^2 \Pr(C1) + \Pr(H | C2)^2 \Pr(C2) \\
= 0.3^2 \cdot 0.5 + 0.9^2 \cdot 0.5 = 0.45
\]

Is \( \Pr(H_1H_2) = 0.45 = \Pr(H_1) \Pr(H_2) \)?

\[
\Pr(H) = \Pr(H | C1) \Pr(C1) + \Pr(H | C2) \Pr(C2) = 0.6 \\
\Pr(H_1) \Pr(H_2) = 0.36 = \Pr(H_1H_2)
\]

\( H_1 \) and \( H_2 \) are *not* independent events?
New topic: random variables

- Random Variables
- Probability Mass Function (PMF)
- Cumulative Distribution Function (CDF)
Random Variables (Idea)

Often: We want to capture quantitative properties of the outcome of a random experiment, e.g.:

– What is the total of two dice rolls?
– What is the number of coin tosses needed to see the first head?
– What is the number of heads among 5 coin tosses?
**Random Variables**

**Definition.** A random variable (RV) for a probability space \((\Omega, \mathbb{P})\) is a function \(X: \Omega \rightarrow \mathbb{R}\).

The set of values that \(X\) can take on is called its range/support \(\Omega_X\).

**Example.** Number of heads in 2 independent coin flips \(\Omega = \{HH, HT, TH, TT\}\)

\(\Omega_X = \{0, 1, 2\}\)
RV Example

20 balls labeled 1, 2, ..., 20 in an urn

- Draw a subset of 3 uniformly at random
- Let $X = \text{maximum of the 3 numbers on the balls}$
  - Example: $X(2, 7, 5) = 7$
  - Example: $X(15, 3, 8) = 15$
RV Example

20 balls labeled 1, 2, ..., 20 in an urn

- Draw a subset of 3 uniformly at random
- Let $X = \text{maximum of the 3 numbers on the balls}$
  - Example: $X(2, 7, 5) = 7$
  - Example: $X(15, 3, 8) = 15$

- What is $|\Omega_X|$?

$|\mathcal{X}| = 18$

Poll: [Slido link removed]

A. $20^3$
B. 20
C. 18
D. $\binom{20}{3}$
Agenda

• Random Variables
• Probability Mass Function (pmf)
• Cumulative Distribution Function (CDF)
Example: Returning Homeworks

- Class with 3 students, randomly hand back homeworks. All permutations equally likely.
- Let $X$ be the number of students who get their own HW

<table>
<thead>
<tr>
<th>$\Pr(\omega)$</th>
<th>$\omega$</th>
<th>$X(\omega)$</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>

$\mathcal{X} = \{0, 1, 3\}$
Example: Returning Homeworks

- Class with 3 students, randomly hand back homeworks. All permutations equally likely.
- Let $X$ be the number of students who get their own HW.

<table>
<thead>
<tr>
<th>$\Pr(\omega)$</th>
<th>$\omega$</th>
<th>$X(\omega)$</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>

$P(X=0) = \frac{1}{6}$

$P(X=1) = \frac{3 \cdot \frac{1}{6}}{2} = \frac{1}{2}$

$P(X=3) = \frac{1 \cdot \frac{1}{6}}{6} = \frac{1}{6}$

$\{X=1\} = \{ (132) (213) (321) \}$

$X(\omega) = 0$  \hspace{1cm}  $X(\omega) = 1$  \hspace{1cm}  $X(\omega) = 3$
**Probability Mass Function (PMF)**

**Definition.** A random variable (RV) for a probability space \((\Omega, \mathbb{P})\) is a function \(X: \Omega \to \mathbb{R}\).

The set of values that \(X\) can take on is called its range/support \(\Omega_X\).

Random variables partition the sample space.

**Definition.** For a RV \(X: \Omega \to \mathbb{R}\), we define the event
\[
\{X = x\} \overset{\text{def}}{=} \{\omega \in \Omega \mid X(\omega) = x\}.
\]
Probability Mass Function (PMF)

**Definition.** For a RV $X: \Omega \to \mathbb{R}$, we define the event

$$\{X = x\} \overset{\text{def}}{=} \{\omega \in \Omega \mid X(\omega) = x\}$$

The **probability mass function** (PMF) of $X$ tells us the probabilities of these events, i.e., the probability that $X$ takes each value in $\Omega_X$.

We use the notation

$$p_X(x) = \mathbb{P}(X = x) = \mathbb{P}(\{\omega \in \Omega \mid X(\omega) = x\})$$

For the probability mass function

$$\sum_{x \in \Omega_X} \mathbb{P}(X = x) = 1$$
Probability Mass Function

Flipping two independent coins

$X = \text{number of heads in the two flips}$

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

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

$\Omega_x = \{0, 1, 2\}$

What is the pmf of $X$?

$p_X(x) = P(X = x) = \begin{cases} 
\frac{1}{4} & x = 0 \\
\frac{1}{4} & x = 1 \\
\frac{1}{4} & x = 2 \\
0 & \text{otherwise}
\end{cases}$
Probability Mass Function

Flipping two independent coins

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

\[ X = \text{number of heads in the two flips} \]

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

\[ \Omega_X = \{0, 1, 2\} \]

\[
\Pr(X = x) = \begin{cases} 
\frac{1}{4}, & x = 0 \\
\frac{1}{2}, & x = 1 \\
\frac{1}{4}, & x = 2 \\
0, & \text{o.w.}
\end{cases}
\]
RV Example

20 balls labeled 1, 2, …, 20 in a bin
– Draw a subset of 3 uniformly at random
– Let $X =$ maximum of the 3 numbers on the balls

What is $p(x(20)) = P(X = 20)$?

$= \Pr(\text{largest of 3 numbers is 20}) = \frac{\binom{1}{1}}{\binom{20}{3}}$

$= \binom{20}{2} / \binom{20}{3}$

Poll: slido.com/4171468

A. $\binom{20}{2} / \binom{20}{3}$
B. $\binom{19}{2} / \binom{20}{3}$
C. $19^2 / \binom{20}{3}$
D. $19 \cdot 18 / \binom{20}{3}$
Agenda

• Random Variables
• Probability Mass Function (PMF)
• Cumulative Distribution Function (CDF)
Cumulative Distribution Function (CDF)

**Definition.** For a RV $X: \Omega \rightarrow \mathbb{R}$, the cumulative distribution function of $X$ specifies for any real number $x$, the probability that $X \leq x$.

$$F_X(x) = \Pr(X \leq x)$$

Go back to 2 coin flips, where $X$ is the number of heads

$$\Pr(X = x) = \begin{cases} 
\frac{1}{4}, & x = 0 \\
\frac{1}{2}, & x = 1 \\
\frac{1}{4}, & x = 2 \\
0, & o.w.
\end{cases}$$
Cumulative Distribution Function (CDF)

**Definition.** For a RV $X: \Omega \rightarrow \mathbb{R}$, the cumulative distribution function of $X$ specifies for any real number $x$, the probability that $X \leq x$.

$$F_X(x) = \text{Pr}(X \leq x)$$

Go back to 2 coin clips, where $X$ is the number of heads

$$\Pr(X = x) = \begin{cases} \frac{1}{4}, & x = 0 \\ \frac{1}{2}, & x = 1 \\ \frac{1}{4}, & x = 2 \end{cases}$$

$$F_X(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{4}, & 0 \leq x < 1 \\ \frac{3}{4}, & 1 \leq x < 2 \\ 1, & 2 \leq x \end{cases}$$