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

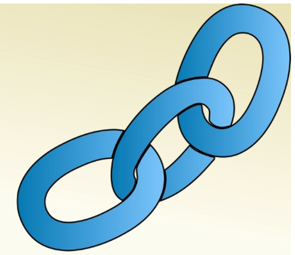
Bayes Theorem with LTP: Let $E_1, E_2, \dots, E_n$ be a partition of the sample space, and F, G events. Then,

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

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



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

Theorem. (Chain Rule) For events $\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_n$,

$$\mathbb{P}(\mathcal{A}_1 \cap \cdots \cap \mathcal{A}_n) = \mathbb{P}(\mathcal{A}_1) \cdot \mathbb{P}(\mathcal{A}_2|\mathcal{A}_1) \cdot \mathbb{P}(\mathcal{A}_3|\mathcal{A}_1 \cap \mathcal{A}_2)$$

$$\cdots \mathbb{P}(\mathcal{A}_n|\mathcal{A}_1 \cap \mathcal{A}_2 \cap \cdots \cap \mathcal{A}_{n-1})$$

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

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

Alternatively,

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

"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 : event that the main chute doesn't open $\qquad \mathbb{P}(A) = 0.02$

  B : event that the backup doesn't open $\qquad \mathbb{P}(B) = 0.1$

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

$$
\begin{aligned}
P(\text{at least one opens}) &= 1 - \Pr(\text{neither open}) \\
&= 1 - P(A \cap B) \\
&= 1 - P(A)P(B) \\
&= 1 - 0.02 \cdot 0.1 \\
&= 0.998
\end{aligned}
$$

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 $\qquad \mathbb{P}(A) = 0.02$
B : event that the backup doesn't open $\qquad \mathbb{P}(B) = 0.1$

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

$P(A \mid B) = 1$

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

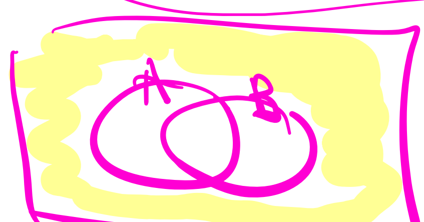
  A : event that the main chute doesn't open $\quad\quad \mathbb{P}(A) = 0.02$

  B : event that the backup doesn't open $\quad\quad \mathbb{P}(B) = 0.1$

- What is the chance that **both open** assuming independence?

$$P(\bar{A} \cap \bar{B})$$
$$= 1 - P(A \cup B)$$
$$= 1 - [P(A) + P(B) - P(A \cap B)]$$
$$= 1 - P(A) - P(B) + P(A)P(B)$$
$$= (1 - P(A))(1 - P(B))$$
$$= P(\bar{A})\, P(\bar{B})$$

A, B indp
$\Rightarrow$ $A, \bar{B}$ indp
$\bar{A}, B$ indp
$\boxed{\bar{A}, \bar{B} \text{ indp.}}$



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 C if $\mathbb{P}(C) \neq 0$ and $\mathbb{P}(\mathcal{A} \cap \mathcal{B} \mid C) = \mathbb{P}(\mathcal{A} \mid C) \cdot \mathbb{P}(\mathcal{B} \mid C)$.

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 $\mathbb{P}(\mathcal{A}|\mathcal{B}) = \mathbb{P}(\mathcal{A})$

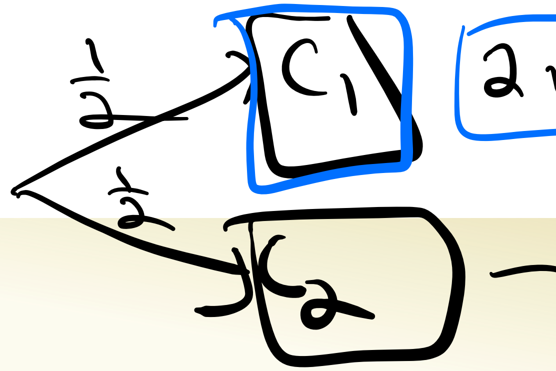
$$\frac{P(A\cap B\cap C)}{P(C)}$$

Conditional Independence

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

Equivalence:

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



2 indp tosses

Example – More coin tossing

Suppose there is a coin C1 with Pr(Head) = 0.3 and a coin C2 with Pr(Head) = 0.9. We pick one randomly with equal probability and flip that coin twice independently. What is the probability both tosses heads?

$$\Pr(H_1H_2) \;=\; \Pr(HH \mid C1)\Pr(C1) \;+\; \Pr(HH \mid C2)\Pr(C2) \qquad \text{LTP}$$

$$= P(H|C_1)P(H|C_1)\,P(C_1) + P(H|C_2)P(H|C_2)\,P(C_2)$$

Example – More coin tossing

Suppose there is a coin C1 with Pr(Head) = 0.3 and a coin C2 with Pr(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?

$$\Pr(HH) = \Pr(HH \mid C1)\Pr(C1) + \Pr(HH \mid C2)\Pr(C2)$$
$$= \Pr(H \mid C1)^2 \Pr(C1) + \Pr(H \mid C2)^2 \Pr(C2)$$
$$= 0.3^2 \cdot 0.5 + 0.9^2 \cdot 0.5 = 0.45$$

LTP

Conditional Independence

$$\text{Is } \Pr(H_1 H_2) = 0.45 \neq \Pr(H_1)\Pr(H_2)\,?$$

$$\Pr(H) = \Pr(H \mid C1)\Pr(C1) + \Pr(H \mid C2)\Pr(C2) = 0.6$$

$\Rightarrow P(H_1) = 0.6$

$P(H_2) = 0.6$

$$\Pr(H_1)\Pr(H_2) = 0.36 \neq \Pr(H_1 H_2)$$

H_1 and H_2 are ~~those~~ indep 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?*

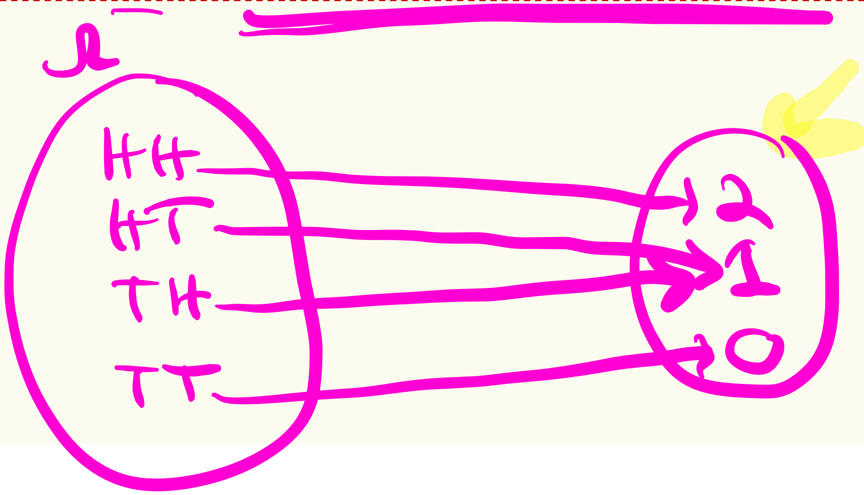
Random Variables

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

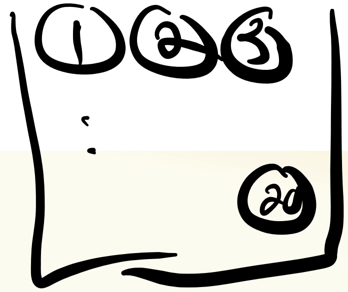
The set of values that X can take on is called its range/support Ω_X

X

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



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



$|\Omega| = \binom{20}{3}$

RV Example

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

- Draw a subset of 3 uniformly at random
- Let $X =$ 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 = maximum of the 3 numbers on the balls
  - Example: $X(2,7,5) = 7$
  - Example: $X(15,3,8) = 15$
- What is $|\Omega_X|$?

Poll: slido.com/4171468

A. 20^3

B. 20

C. $\boxed{18}$

D. $\binom{20}{3}$

$|\Omega_X| = 18$

1 2 3 4 5 6 - - - 19 20

Agenda

- Random Variables
- Probability Mass Function (pmf)
- Cumulative Distribution Function (CDF)

Ω, $P(\omega)\ \forall \omega \in \Omega$

r.v.s

$\mathcal{L}_X$

what is prob r.v. takes each value?

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

$\Pr(\omega)$	ω	$X(\omega)$
$1/6$	1, 2, 3	3
$1/6$	1, 3, 2	1
$1/6$	2, 1, 3	1
$1/6$	2, 3, 1	0
$1/6$	3, 1, 2	0
$1/6$	3, 2, 1	1

$\Omega_X = \{0, 1, 3\}$

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

$\Pr(\omega)$	ω	$X(\omega)$
1/6	1, 2, 3	3
1/6	1, 3, 2	1
1/6	2, 1, 3	1
1/6	2, 3, 1	0
1/6	3, 1, 2	0
1/6	3, 2, 1	1

Ω

$P(X=0) = \frac{1}{3} \longrightarrow$ 2, 3, 1; 3, 1, 2 — $X(\omega) = 0$

$P(X=1) = 3 \cdot \frac{1}{6} = \frac{1}{2}$ — 1, 3, 2; 2, 1, 3; 3, 2, 1 — $X(\omega) = 1$

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

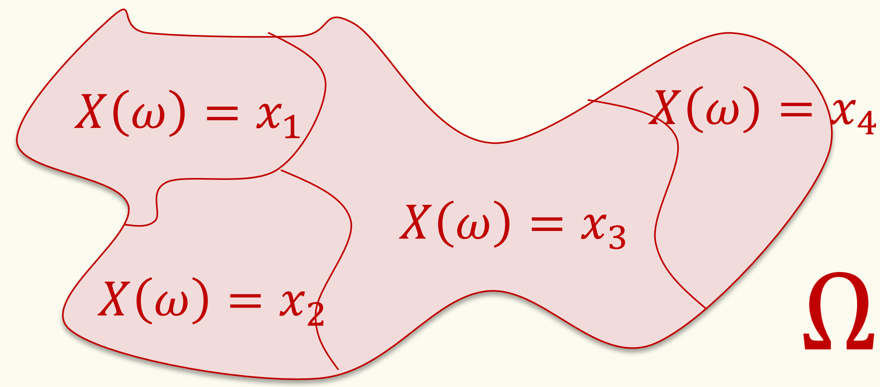
$\{X=1\} = \{(132)(213)(321)\}$

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 Ω_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 Ω_X

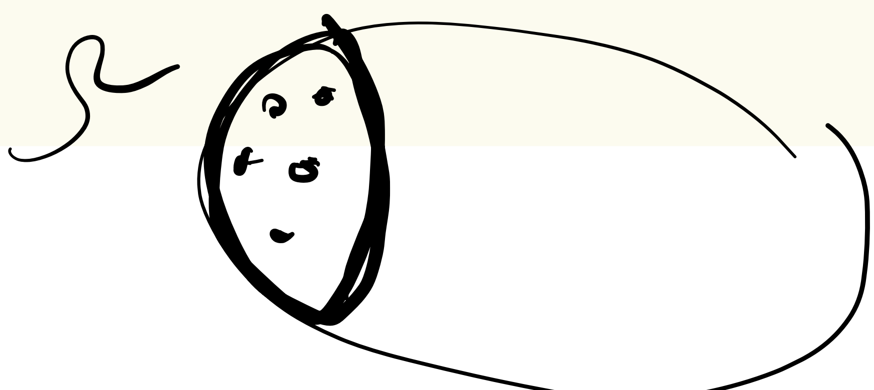
We use the notation

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

For the probability mass function

$p_X(x)$

$$\boxed{\sum_{x \in \Omega_X} \mathbb{P}(X = x) = 1}$$



Probability Mass Function

prob $\frac{1}{2}$ of coming up Hs

Flipping two independent coins $\Omega = \{HH, HT, TH, TT\}$

X = number of heads in the two flips

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

$\boxed{\Omega_X = \{0, 1, 2\}}$

What is the pmf of X?

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

$P(X=1)$

Probability Mass Function

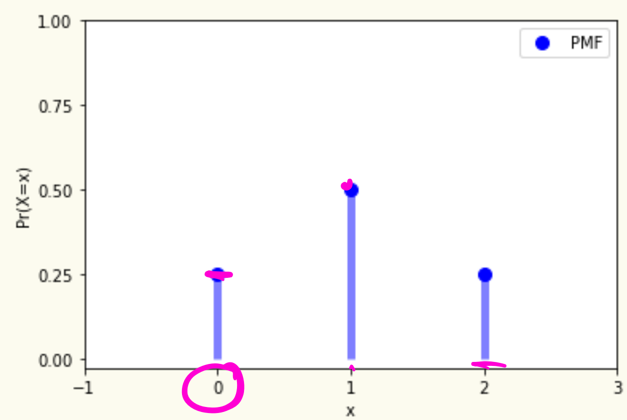
Flipping two independent coins $\quad \Omega = \{\mathrm{HH}, \mathrm{HT}, \mathrm{TH}, \mathrm{TT}\}$

$X =$ number of heads in the two flips

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

$$\Omega_{\mathrm{X}} = \{0, 1, 2\}$$

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



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

Poll: **slido.com/4171468**

What is $p_X(20) = P(X = 20)$?

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

E

A. $\binom{20}{2} / \binom{20}{3}$

B. $\boxed{\binom{19}{2} / \binom{20}{3}}$

C. $19^2 / \binom{20}{3}$

D. $19 \cdot 18 / \binom{20}{3}$

$$p_X(k) = P(X = k) = \frac{\binom{k-1}{2}}{\binom{20}{3}}$$

$k \in \{3, \ldots, 20\}$

(_ _ k)

1 2 3 . . k-1 $\boxed{k}$ 20

Agenda

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

$P_X(x) = P(X=x) \quad \forall x \in \Omega_X$

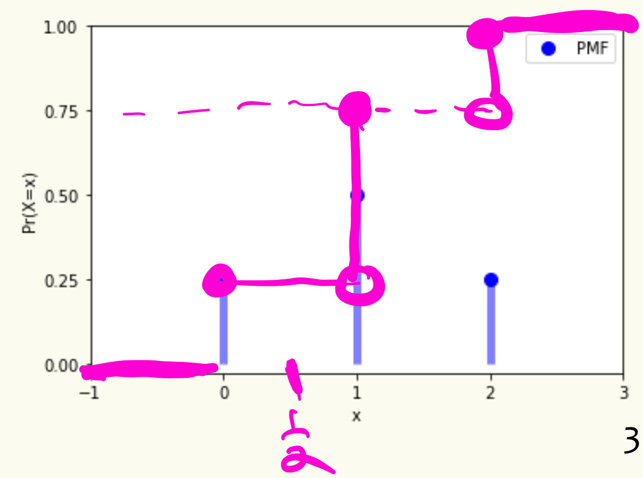
Cumulative Distribution Function (CDF)

Definition. For a RV $X: \Omega \to \mathbb{R}$, 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)

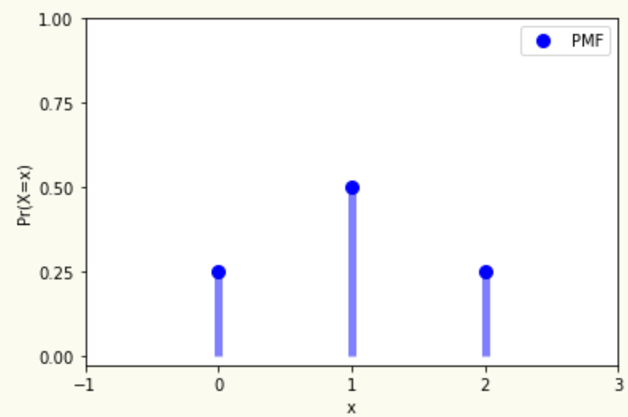
Definition. For a RV $X: \Omega \to \mathbb{R}$, the cumulative distribution function of X specifies for any real number x, the probability that $X \leq x$.

$$\mathrm{F}_X(x) = \Pr(X \leq x)$$

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

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

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




gifak-net
h2
h1
b
a
S=a*h1
S=b*h2