CSE 312

Foundations of Computing II

Lecture 7: Conditional Independence

PAUL G. ALLEN SCHOOL
OF COMPUTER SCIENCE & ENGINEERING

Aleks Jovcic

Slide Credit: Based on Stefano Tessaro's slides for 312 19au
incorporating ideas from Anna Karlin, Alex Tsun, Rachel Lin, Hunter Schafer & Anna ☺

Agenda

- Independence Example
- Conditional Independence
- Assumptions and Correlation
- Monty Hall Problem
- If time: Random Variables Introduction
  - Probability Mass Function (PMF)
  - Cumulative Distribution Function (CDF)

Example – Biased coin

We have a biased coin comes up Heads with probability 2/3; Each flip is independent of all other flips. Suppose it is tossed 3 times.

$\mathbb{P}(H_1 \cap H_2 \cap H_3) = \mathbb{P}(H_1)\mathbb{P}(H_2)\mathbb{P}(H_3)$

$\mathbb{P}(HHH) = \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{2}{3} = \frac{8}{27}$

$\mathbb{P}(TTT) = \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{3} = \left(\frac{1}{3}\right)^3$

$\mathbb{P}(HTT) = \frac{2}{3} \cdot \frac{1}{3} \cdot \frac{1}{3}$

Example – Biased coin

~~$|E|/|\Omega|$~~

We have a biased coin comes up Heads with probability 2/3, independently of other flips. Suppose it is tossed 3 times.

$\mathbb{P}(2\ heads\ in\ 3\ tosses) =$

$\mathbb{P}(HHT \cup HTH \cup THH) = \mathbb{P}(HHT) + \mathbb{P}(HTH) + \mathbb{P}(THH)$

$\left(\frac{2}{3}\right)^2 \frac{1}{3} \quad \left(\frac{2}{3}\right)^2 \frac{1}{3} \quad \text{'' ''}$

$$3 \cdot \left(\frac{2}{3}\right)^2 \frac{1}{3}$$

A) $(2/3)^2\ 1/3$

B) $2/3$

C) $3\ (2/3)^2\ 1/3$

D) $(1/3)^2$

Agenda

- Independence Example
- Conditional Independence
- Assumptions and Correlation
- Monty Hall Problem
- If time: Random Variables Introduction
  - Probability Mass Function (PMF)
  - Cumulative Distribution Function (CDF)

Conditional Independence

Definition. Two events $\mathcal{A}$ and $\mathcal{B}$ are **independent** conditioned on C if

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

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

Conditional Independence

Definition. Two events $\mathcal{A}$ and $\mathcal{B}$ are **independent** conditioned on $\mathcal{C}$ if

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

$\mathbb{P}(A_3 \mid A_2 \cap A_1)$

Equivalence:

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

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

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 we get all heads?

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

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

$$= \Pr(H \mid C_1)^2 \Pr(C_1) + \Pr(H \mid C_2)^2 \Pr(C_2)$$

$$(0.3)^2\, 0.5 + (0.9)^2\, 0.5$$

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 we get all heads?

$$\Pr(HH) = \Pr(HH \mid C1)\Pr(C1) + \Pr(HH \mid C2)\Pr(C2)$$ LTP

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?

$$\Pr(HH) = \Pr(HH \mid C1)\Pr(C1) + \Pr(HH \mid C2)\Pr(C2)$$ LTP

$$= \Pr(H \mid C2)^2 \Pr(C1) + \Pr(H \mid C2)^2 \Pr(C2)$$ Conditional Independence

$$= 0.3^2 \cdot 0.5 + 0.9^2 \cdot 0.5 = 0.45$$

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

Agenda

- Independence Example
- Conditional Independence
- Assumptions and Correlation
- Monty Hall Problem
- If time: Random Variables Introduction
  - Probability Mass Function (PMF)
  - Cumulative Distribution Function (CDF)

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?

~~$\mathbb{P}(A \cap B)$ 0.08~~ $1 - \mathbb{P}(A \cap B)$

$1 - (\mathbb{P}(A) \cdot \mathbb{P}(B))$

$0.02 \cdot 0.1 = 0.998$

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

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

- Assuming independence doesn't justify the assumption! Both chutes could fail because of the same rare event e.g., freezing rain.

Correlation

- Pick a person at random
- A : event that the person has lung cancer
- B : event that the person is a heavy smoker

- Fact: $\mathbb{P}(A|B) = 1.17 \cdot \mathbb{P}(A)$

- Conclusions?

Correlation

- Pick a person at random
- A : event that the person has lung cancer
- B : event that the person is a heavy smoker

- Fact: $\mathbb{P}(A|B) = 1.17 \cdot \mathbb{P}(A)$

- Conclusions?
  - Smoking causes lung cancer.
    - Smoking increases the probability of smoking by 17%.

Causality vs. Correlation

- Events A and B are **positively correlated** if

$$\mathbb{P}(A \cap B) > \mathbb{P}(A) \cdot \mathbb{P}(B)$$

$$\mathbb{P}(A|B) > \mathbb{P}(A)$$

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

$$\mathbb{P}(A \cap B) = \mathbb{P}(A|B)\,\mathbb{P}(B) > \mathbb{P}(A)\,\mathbb{P}(B)$$

- E.g. smoking and lung cancer.

- But A and B being positively correlated does not mean that A causes B or B causes A.

Causality vs. Correlation

- Events A and B are **positively correlated** if

$$\mathbb{P}(A \cap B) > \mathbb{P}(A) \cdot \mathbb{P}(B)$$

- But A and B being positively correlated does not mean that A causes B or B causes A.

Other examples:

- Tesla owners are more likely to be rich. That does not mean poor people should buy a Tesla to get rich.
- People who go to the opera are more likely to have a good career. That does not mean that going to the opera will improve your career.
- Rabbits eat more carrots and do not wear glasses. Are carrots good for eyesight?

Agenda

- Independence Example
- Conditional Independence
- Assumptions and Correlation
- Monty Hall Problem
- If time: Random Variables Introduction
  - Probability Mass Function (PMF)
  - Cumulative Distribution Function (CDF)

Monty Hall Problem

Suppose you're on a game show, and you're given the choice of three doors. Behind one of the doors is a car, behind the other, goats. You pick a door, say number 1, and the host, who knows what's behind the doors, opens another door, say number 3, which has a goat. He says to you, "Do you want to switch to door number 2?" Is it to your advantage to switch your choice of doors?

Assumptions

- The player is equally likely to pick each of the three doors.
- After the player picks a door, the host must open a different door with a goat behind it and offer the player the choice of staying with the original door or switching.
- If the host has a choice of which door to open, then he is equally likely to select each of them.

Should you switch or stay?

W = winning by switching after host opens door #1

$\$_1, \$_2, \$_3$

$P(W)$

$$P(W \mid \$_1)P(\$_1) + P(W \mid \$_2)P(\$_2) + P(W \mid \$_3)P(\$_3)$$

$\rightarrow 0 \cdot \frac{1}{3}$

$\rightarrow 1 \cdot \frac{1}{3}$

$\rightarrow 1 \cdot \frac{1}{3}$

$$= \frac{2}{3}$$

Agenda

- Independence Example
- Conditional Independence
- Assumptions and Correlation
- Monty Hall Problem
- If time: Random Variables Introduction
  - 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.:

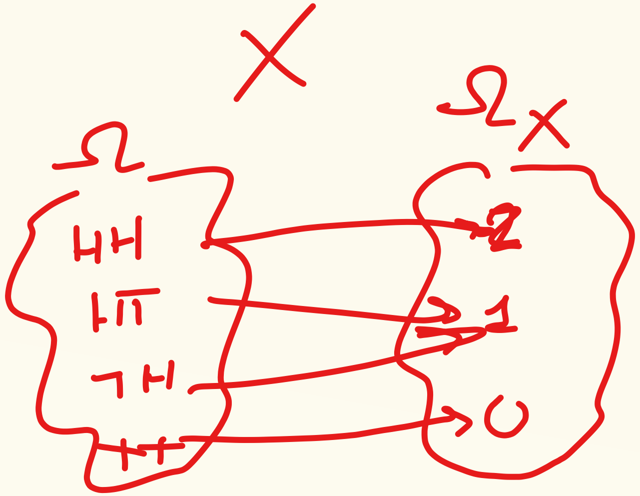
- *What is the total of two dice rolls?*
- *What is the number of coin tosses needed to see the first head?*
- *What is the number of heads among 2 coin tosses?*

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

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



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
  - Example: $X(2, 7, 5) = 7$
  - Example: $X(15, 3, 8) = 15$
- What is $|\Omega_X|$? = 18

$\{2,3,7\} \to 7$

$\{2,7,5\} \to 7$

$\{18, 19, 20\}$

$\{3, 4, 17\}$

Ω_X

A. 20^3

B. 20

C. 18

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

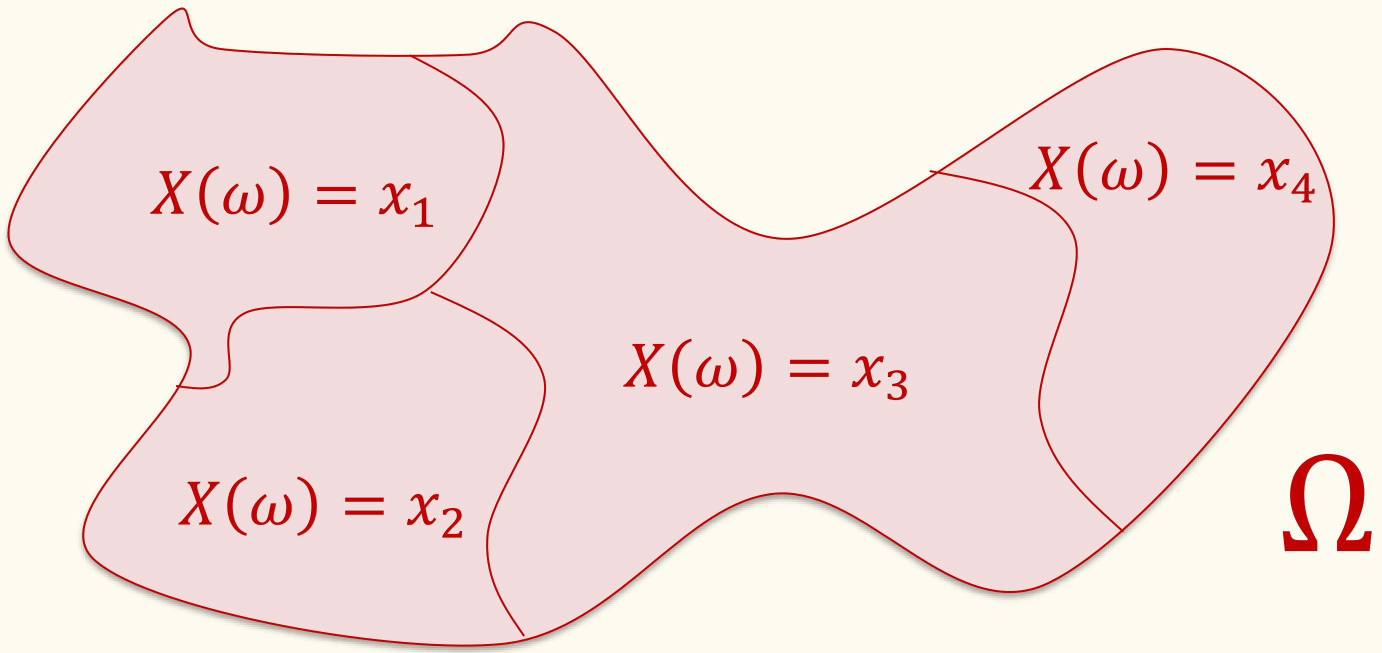
Agenda

- Independence Example
- Conditional Independence
- Assumptions and Correlation
- Monty Hall Problem
- If time: Random Variables Introduction
  - Probability Mass Function (PMF)
  - Cumulative Distribution Function (CDF)

Probability Mass Function (PMF)

Random variables partition the sample space.

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



Probability Mass Function (PMF)

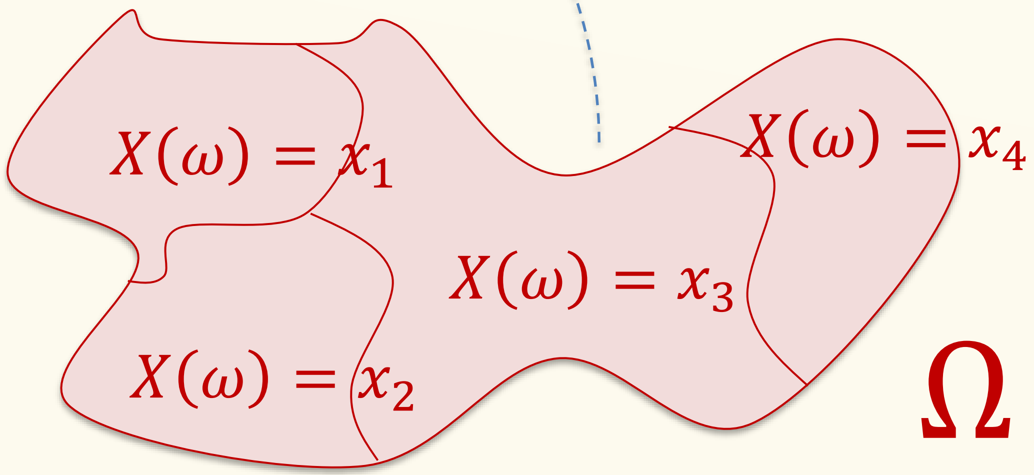
Definition. For a RV $X: \Omega \to \mathbb{R}$, we define the event

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

We write $\mathbb{P}(X = x) = \mathbb{P}(\{X = x\}) = \mathbb{P}(\{\omega \in \Omega \mid X(\omega) = x\})$ where $\mathbb{P}(X = x)$ is the probability mass function (PMF) of X

Random variables partition the sample space.

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



Probability Mass Function (PMF)

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

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

We write $\mathbb{P}(X = x) = \mathbb{P}(\{X = x\}) = \mathbb{P}(\{\omega \in \Omega \mid X(\omega) = x\})$ where $\mathbb{P}(X = x)$ is the probability mass function (PMF) of X

Random variables partition the sample space.

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

You also see this notation (e.g. in book):

$$\mathbb{P}(\boldsymbol{X} = \boldsymbol{x}) = \boldsymbol{p_X}(\boldsymbol{x})$$

Probability Mass Function

Flipping two independent coins

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

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

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

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

What is $Pr(X = k)$?

RV Example

20 balls labeled 1, 2, …, 20 in a bin

- Draw a subset of 3 uniformly at random
- Let $X = \text{maximum of the 3 numbers on the balls}$

What is $Pr(X = 20)$?

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

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

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

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

Agenda

- Independence Example
- Conditional Independence
- Assumptions and Correlation
- Monty Hall Problem
- If time: Random Variables Introduction
  - Probability Mass Function (PMF)
  - Cumulative Distribution Function (CDF)

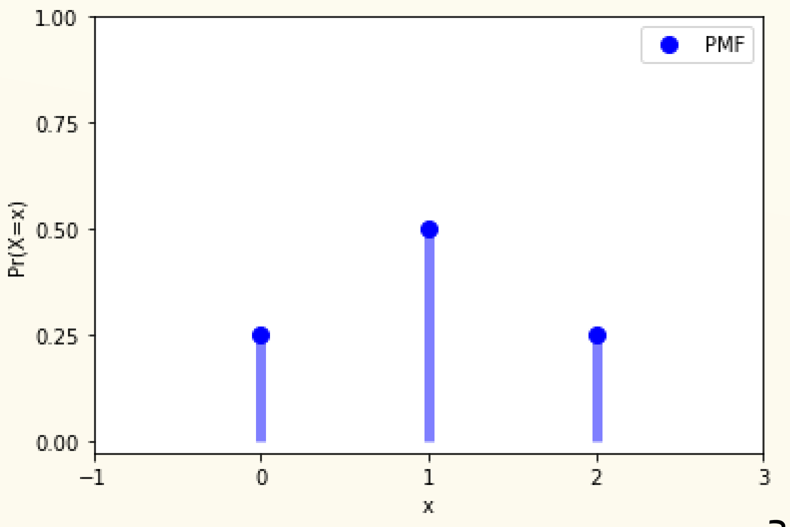
Cumulative Distribution Function (CDF)

Definition. For a RV $X: \Omega \to \mathbb{R}$, the cumulative distribution function of where 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 clips, where X is the number of heads

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



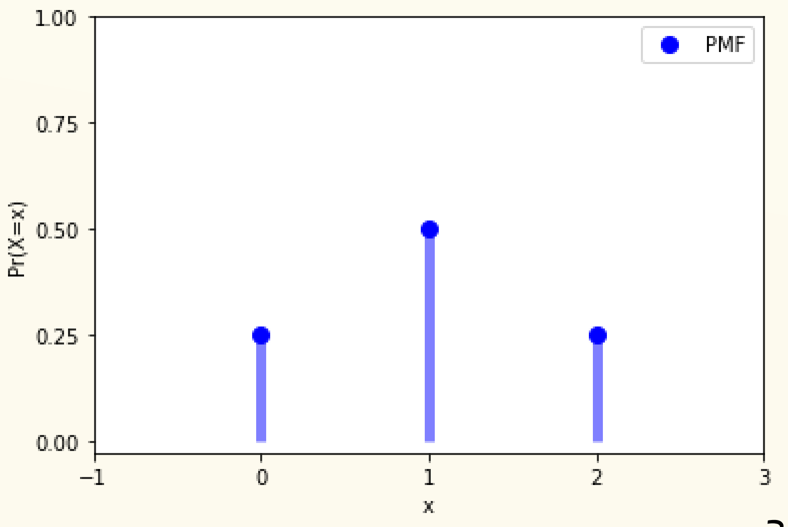
Cumulative Distribution Function (CDF)

Definition. For a RV $X: \Omega \to \mathbb{R}$, the cumulative distribution function of where 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 clips, where X is the number of heads

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



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