Probability Refresher

CMPUT 365
Fall 2021

Rolling a 6-sided dice

- The outcomes of a **single** dice roll is {1, 2, 3, 4, 5, 6}
- Some events can be:
  - seeing number 1, seeing number 2, seeing 1 or 2, seeing 1 or 3, etc
- The probability of seeing number 1 is 1/6, the probability of seeing number 1 or 2 is 1/6 + 1/6 = 1/3

Modelling random phenomena

- Define a set of possible outcomes, a collection of events and probability for each event
- 6-sided dices experiment**:**
    - the **outcomes** of a single dice roll are 1, 2, 3, 4, 5, 6
    - A collection of all **events** for 6-sided dice experiment: {{},{1},{2},{3},{4},{5},{6},{1,2}, {1,3}, …, {1,2,3,4,5,6}}
        - {1} translates to “seeing number 1”
        - {1,2} translates to “seeing number 1 or 2” and the **probability**(“seeing number 1 or 2”) = 1/3
        - {} translates to “seeing nothing” and the probability{“seeing nothing”} = 0
        - {1, 2, 3, 4, 5, 6} translates to “seeing 1 or 2 or 3 or 4 or 5 or 6” and the probability of this event is 1

Flipping a 2-sided coin twice

- The possible outcomes are {HH, HT, TH, TT}
- Some examples of events:
    - Seeing head on the first flip = {HH, HT} , seeing only tails = {TT}
- Probability on the events:
    - Prob("seeing head on the first flip") = 1/4 + 1/4 = 1/2
    - Prob("seeing only tails") = 1/4

Random variable

- A random variable X assigns a numerical value to each of the outcomes in the outcome space. In other words, it is a function that maps an outcome (not event) to the real
- Examples:
  - 6-sided dice roll:
    - X(outcome 1) = 1, X(outcome 2) = 2, X(outcome 3) = 3, etc

Random variable continued

- Let's examine the 2-sided coin flip experiment
- Outcomes are {HH, HT, TH, TT}
- A table of events and their corresponding probabilities on the right
- Define X(HH) = 1, X(HT) = 2, X(TH) = 3, X(TT) = 4
- After flipping a coin twice, now we see 1 instead of the outcome HH, HT, TH, or TT, what is the probability of seeing 1?
    - $P(X = 1)$?

Event	Probability
nothing	0
HH	1/4
HT	1/4
TH	1/4
TT	1/4
HH, HT	1/2
HH, TH	1/2
HH, TT	1/2
HT, TH	1/2
HT, TT	1/2
TH, TT	1/2
HH, HT, TH	3/4
HH, HT, TT	3/4
HT, TH, TT	3/4
HH, HT, TH, TT	1

What does P(X=1) mean?

- Let $\omega = \{\text{HH, HT, TH, TT}\}$, then $P(X = 1)$ really mean:
    - $P(\{\omega \,|\, X(\omega) = 1\}) = P(\{HH\}) = 1/4$
    - Recall X(HH) = 1, X(HT) = 2, X(TH) = 3, X(TT) = 4

Event	Probability
nothing	0
HH	1/4
HT	1/4
TH	1/4
TT	1/4
HH, HT	1/2
HH, TH	1/2
HH, TT	1/2
HT, TH	1/2
HT, TT	1/2
TH, TT	1/2
HH, HT, TH	3/4
HH, HT, TT	3/4
HT, TH, TT	3/4
HH, HT, TH, TT	1

Random variable continued

- Let's examine the 2-sided coin flip experiment again
- Outcomes are {HH, HT, TH, TT}
- A table of events and their corresponding probabilities on the right
- Define X(HT) = 1 = X(TH) and X(HH) = 2 and X(TT) = 3
- After flipping a coin twice, now we see 1 instead of the outcome HH or TH, what is the probability of seeing 1?
    - $P(X = 1)$?

Event	Probability
nothing	0
HH	1/4
HT	1/4
TH	1/4
TT	1/4
HH, HT	1/2
HH, TH	1/2
HH, TT	1/2
HT, TH	1/2
HT, TT	1/2
TH, TT	1/2
HH, HT, TH	3/4
HH, HT, TT	3/4
HT, TH, TT	3/4
HH, HT, TH, TT	1

Discrete and continuous random variables

- A **discrete** random variable takes a distinct number of values, which can be finite or countably infinite:
    - Example: 6-sided dice ($\mathscr{X} = \{1,2,3,4,5,6\}$)
- A **continuous** random variable takes an uncountably infinite number of values (like an interval)
    - Example: temperature of a thermometer

Notations

- Discrete random variable:
  - $P(X = x)$
- Continuous random variable:
  - $P(X \leq x)$ or $P(a \leq X \leq b)$
  - $P(X = x) = 0$

Probability mass function

- Let X be a discrete random variable that takes values in $\mathscr{X} = \{x_1, x_2, \ldots\}$ (finite or countably infinite). The probability

    - $P(X = x_k)$ for $k = 1,2,3,...,$

  is called the probability mass function (PMF) of X

- Equivalently we write: p(x) = P(X = x)

Properties of the PMF

1. $p : \mathcal{X} \to [0, 1]$

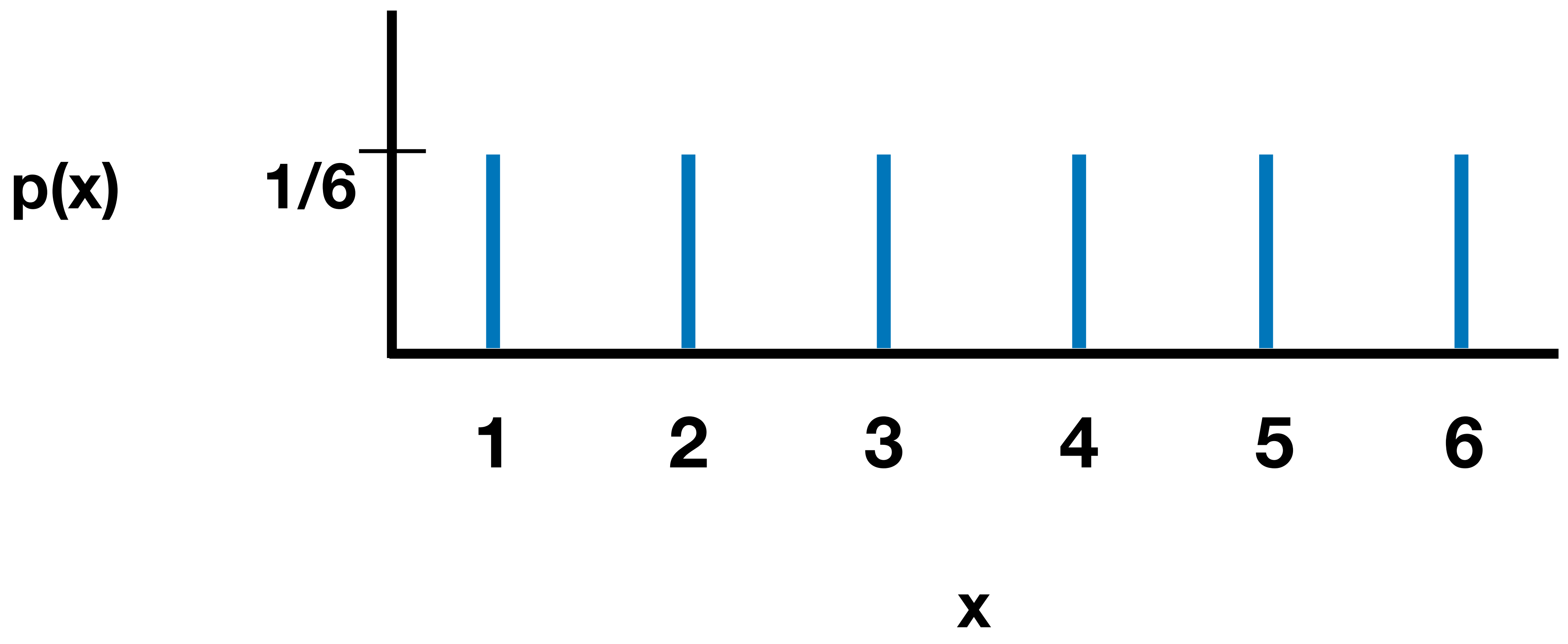
   i.e., $0 \leq p(x) \leq 1$

2. $\displaystyle\sum_{x \in \mathcal{X}} p(x) = 1$

e.g., $\mathcal{X} = \{1, 2, 3, 4, 5, 6\}$

An example of PMF

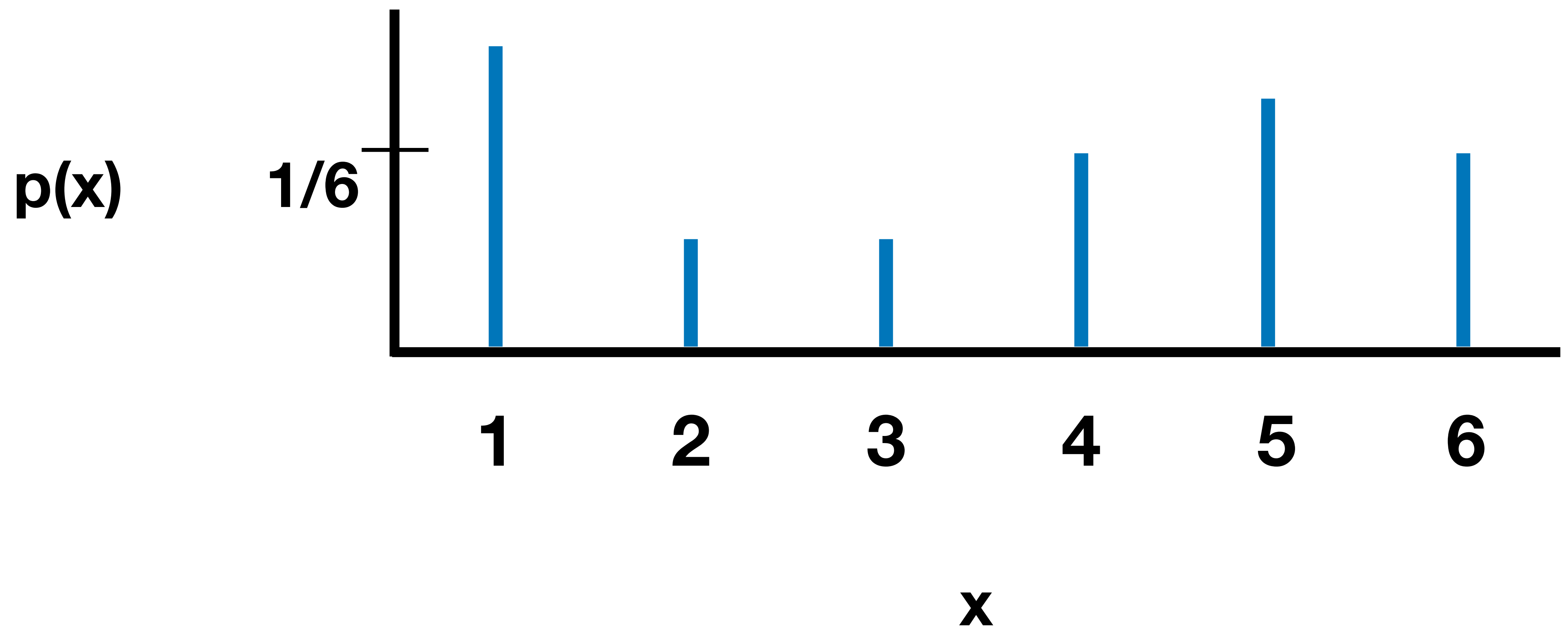
- Recall the 6-sided dice roll
- p(x) = 1/6 for all x in {1, 2, 3, 4, 5, 6}



- What is $\sum_{x \in \mathcal{X}} p(x)$?

Another PMF

- Recall the 6-sided dice roll
- What does this PMF say?



Another example

- Imagine that you are a medical doctor, prescribing treatments to patients
- You give the treatment for 100 patients and note the treatment outcome as {Bad, Neutral, Good}.
- 20 patients have a Bad outcome
- 50 a Neutral outcome
- 30 a Good outcome
- Obtain p(x) using 20/100, 50/100 and 30/100

$$p(x) = \begin{cases} 0.2 & \text{if } x \text{ is Bad} \\ 0.5 & \text{if } x \text{ is Neutral} \\ 0.3 & \text{if } x \text{ is Good.} \end{cases}$$

Another example

- Imagine that you are a medical doctor, prescribing treatments to patients
- You give the treatment for many patients and note the outcome X = treatment outcome in {Bad, Neutral, Good}. You find that you have the following probabilities:

$$p(x) = \begin{cases} 0.2 & \text{if } x \text{ is Bad} \\ 0.5 & \text{if } x \text{ is Neutral} \\ 0.3 & \text{if } x \text{ is Good.} \end{cases}$$

- But that's pretty stochastic…

How can we make this less stochastic?

- What if we knew something about the patient?
- What information could help?

$$p(x) = \begin{cases} 0.2 & \text{if } x \text{ is Bad} \\ 0.5 & \text{if } x \text{ is Neutral} \\ 0.3 & \text{if } x \text{ is Good.} \end{cases}$$

Conditional Probabilities

- X = outcome of treatment
- Y = age of the patient
- P(X = Good | Y = 22) can a different value than P(X = Good | Y = 78)
- $P(X = x \mid Y = y)$ can be read as the probability that the random variable ‘X’ takes the value ‘x’, given that the random variable ‘Y” has taken the value ‘y’.

Another example of conditional probabilities

- Often the value of a random variable is dependent on or correlated with another random variable

- Example: a store restocks on Wednesday, so the probability that they have pencils in stock depends on the day of the week

1	Wed	**Pencil Delivery!**
2	Thu	
3	Fri	
4	Sat	
5	Sun	
6	Mon	
7	Tues	

$$P(\text{Pencils in stock} \mid \text{Tuesday}) = 0.2 \qquad P(\text{Pencils in stock} \mid \text{Wednesday}) = 1.0$$

"if" statement of probability

Same properties for conditional probabilities as for unconditioned probabilities

$$1.\ \ p(\cdot|Y=y) : \mathcal{X} \to [0,1]$$

Could write it $p_{Y=y}(x)$

$$2.\ \ \sum_{x \in \mathcal{X}} p(x|Y=y) = 1$$

e.g. If $P(\text{Pencils in stock} \mid \text{Tuesday}) = 0.2$, then
$P(\text{Pencils NOT in stock} \mid \text{Tuesday}) = 0.8$

Exercise 1

- Find $P(X = 3 \mid X = 3)$ and $P(X = 9 \mid X = 4)$.

Known

Known

Must be true

Impossible

Exercise 1: Answer

- Find $P(X = 3 \mid X = 3)$ and $P(X = 9 \mid X = 4)$.

Known

Must be true

$$P(X = 3 \mid X = 3) = 1$$

Multiple random variables

- We now have seen two random variables, X and Y, we can reason about more than just $p(x \mid y)$
- We can ask about their joint probability: $p(x, y)$
- We might want to ask about the other conditional: $p(y \mid x)$

Back to the Pencils example

- Let X in {0, 1} (Pencils not in stock, Pencils in stock)
- Let Y in {Monday, Tuesday, …, Sunday} (Notice p(y) = 1/7 for all y)
- We can ask P(X = 1, Y = Tuesday) or P(Y = Tuesday | X = 1)
- Do you think P(Y = Wednesday | X = 1) > P(Y = Tuesday | X = 1)

1	Wed	**Pencil Delivery!**
2	Thu	
3	Fri	
4	Sat	
5	Sun	
6	Mon	
7	Tues	

$P(\text{Pencils in stock} \mid \text{Tuesday}) = 0.2$ $P(\text{Pencils in stock} \mid \text{Wednesday}) = 1.0$

"if" statement of probability

Chain Rule

- Can express joint probability using conditional probabilities
- $p(x, y) = p(x \mid y)\, p(y)$
- $p(x, y) = p(y \mid x)\, p(x)$

Back to the Pencils example

- Do you think P(Y = Wednesday, X = 1) > P(Y = Tuesday, X = 1) ?
- Why or why not? Use P(Y = Wednesday, X = 1) **=** P(Y = Wednesday | X = 1) P(X = 1)

1	Wed	**Pencil Delivery!**
2	Thu	
3	Fri	
4	Sat	
5	Sun	
6	Mon	
7	Tues	

$P(\text{Pencils in stock} \mid \text{Tuesday}) = 0.2$ $P(\text{Pencils in stock} \mid \text{Wednesday}) = 1.0$

"if" statement of probability

Independence and Conditional Independence

- X and Y are **independent** if and only if $p(x, y) = p(x)\,p(y)$
  - equivalently, if $p(x \mid y) = p(x)$
  - recall: $p(x, y) = p(x \mid y)\,p(y)$
- X and Y are **conditionally independent**, given Z, if and only if $p(x, y \mid z) = p(x \mid z)\,p(y \mid z)$

Example of Independence

- 2-sided fair coin
- Let X denote the first flip, which results in any of the outcome {H, T}
- Let Y denote the second flip, which results in any of the outcome {H, T}
- Both flips are independent of each other, and if the probability of flipping H is 1/2, then
    - $p(x, y) = p(x)\,p(y) = 1/2 \times 1/2 = 1/4$
    - i.e., $P(X = H, Y = H) = P(X = H)\,P(Y = H)$, and $P(X = T, Y = H) = P(X = T)\,P(Y = H)$, …

Example of Conditional Independence

- For the pencils example, imagine you had an random variable Z where
  - Z = 0 if the day of the week is in the range Saturday to Tuesday
  - Z = 1 if the day of the week is in the range Wednesday to Friday
- Z is not independent of X (recall, X is whether Pencils are in stock or not)
  - P(Z = 0 | X = 1) not equal to P(Z = 0) (P(Z = 0) = 4/7 whereas P(Z = 0 | X = 1) < 4/7)
- Z is conditionally independent of X, given Y (Y is the day of the week)
  - $P(Z=0|X=1,Y=D) = P(Z=0|Y=D) = \begin{cases} 0 & \text{if D in Sat to Tues} \\ 1 & \text{if D in Wed to Fri} \end{cases}$

Marginalization

- If we have the joint distribution p(x, y), we can find the marginals p(x) and p(y)

$$p(x) = \sum_{y \in \mathcal{Y}} p(x, y) \qquad p(y) = \sum_{x \in \mathcal{X}} p(x, y)$$

Exercise 2 (Marginalization)

- Imagine someone gives you $P(X = 1 \mid Y = y)$ for y in {Monday,…, Sunday}
- You already know $p(y) = 1/7$
- How would you get $p(X = 1)$? (i.e, the probability that Pencils are in stock)

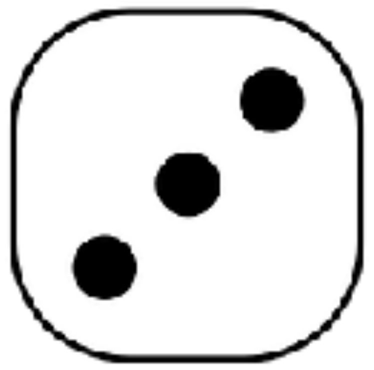
Exercise 3 (Conditional Probs)

- You flip a coin and get two heads in a row. What is the probability that your third coin flip also results in heads?

Exercise 4
(Formalizing probability questions)

- You roll two standard dice. One of the die shows a 3. What is the probability that the sum of the dice is greater than 7?

Roll 1



Roll 2