Discrete Random Variables and Their Distributions

Administrivia

- **Reminder**: Homework 2 is posted and is due two Friday's from now
- If you didn't start early last time, please do so this time. Good **Milestones**:
    - Finish Problems 1-3 this week. More math, some programming.
    - Finish Problems 4-5 next week. Less math, more programming.

Previously on CSCI 3022

Def: a discrete random variable X is a function that maps the elements of the sample space Ω to a finite number of values $a_1, a_2, \ldots, a_n$ or an infinite number of values $a_1, a_2, \ldots$

Def: a probability mass function is the map between the random variable's values and the probabilities of those values

$$f(a) = P(X = a)$$

Def: a **cumulative distribution function** (CDF) is a function whose value at a point **a** is the cumulative sum of probability masses up until **a**.

$$F(a) = P(X \leq a)$$

PMF and CDF Warm-Up

Example: Suppose you roll two fair, six-sided dice. Let X be a random variable representing the maximum of the two dice.

Question: What are the possible values that X can take on?

$$X \in \{1, 2, 3, 4, 5, 6\}$$

PMF and CDF Warm-Up

Example: Suppose you roll two fair, six-sided dice. Let X be a random variable representing the maximum of the two dice.

Question: Which elements of the sample space map to which values of X ?

ω_2

ω_1

	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	2	$\boxed{3}$	4	5	6
3	3	3	3	4	5	6
4	4	4	4	4	5	6
5	5	5	5	5	5	6
6	6	6	6	6	6	6

PMF and CDF Warm-Up

Example: Suppose you roll two fair, six-sided dice. Let X be a random variable representing the maximum of the two dice.

Question: What is the PMF of random variable X?

	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	2	3	4	5	6
3	3	3	3	4	5	6
4	4	4	4	4	5	6
5	5	5	5	5	5	6
6	6	6	6	6	6	6

a	1	2	3	4	5	6
$f(a)$	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$

$\Sigma = 1$

PMF and CDF Warm-Up

Example: Suppose you roll two fair, six-sided dice. Let X be a random variable representing the maximum of the two dice.

Question: What is the probability that X is an even number?

a	1	2	3	4	5	6
$f(a)$	$1/36$	$3/36$	$5/36$	$7/36$	$9/36$	$11/36$

$P(\text{EVEN}) =$

$P(2) + P(4) + P(6)$

$= \frac{21}{36} = \frac{7}{12}$

PMF and CDF Warm-Up

Example: Suppose you roll two fair, six-sided dice. Let X be a random variable representing the maximum of the two dice.

Question: What is the probability that X is 3 or smaller?

a	1	2	3	4	5	6
$f(a)$	$1/36$	$3/36$	$5/36$	$7/36$	$9/36$	$11/36$

$$P(X \le 3) = \frac{1}{36} + \frac{3}{36} + \frac{5}{36} = \frac{9}{36} = \frac{1}{4}$$

PMF and CDF Warm-Up

Example: Suppose you roll two fair, six-sided dice. Let X be a random variable representing the maximum of the two dice.

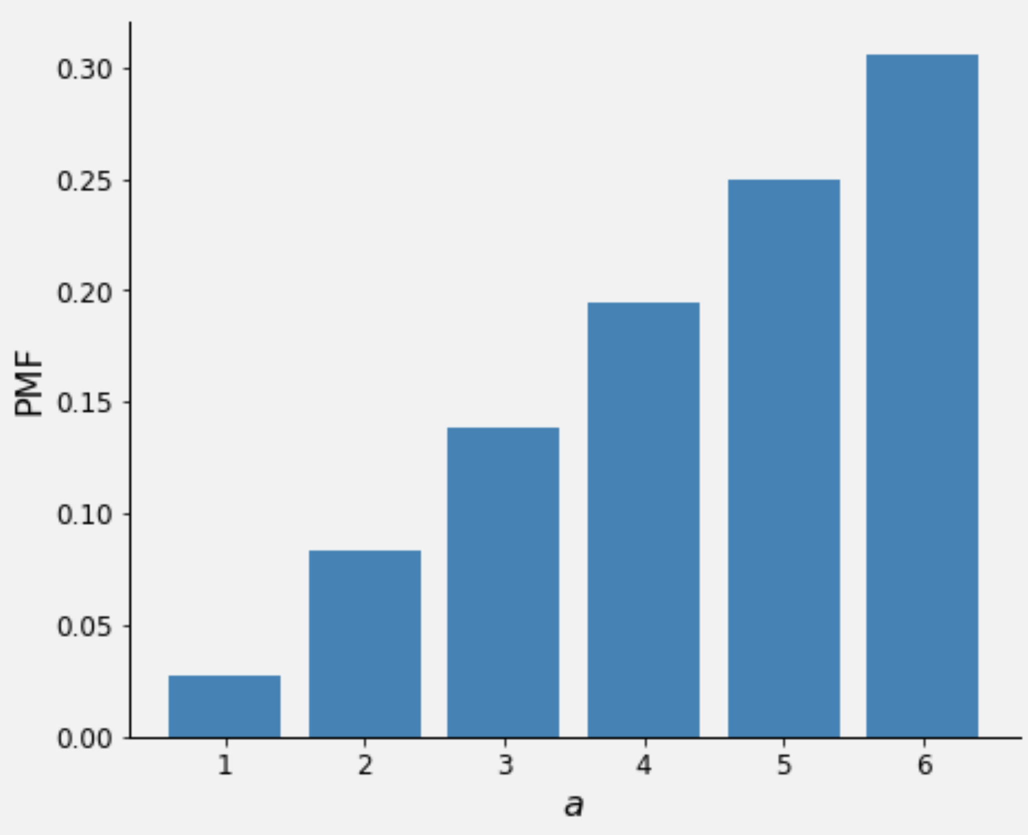
Question: What is the complete CDF of X ?

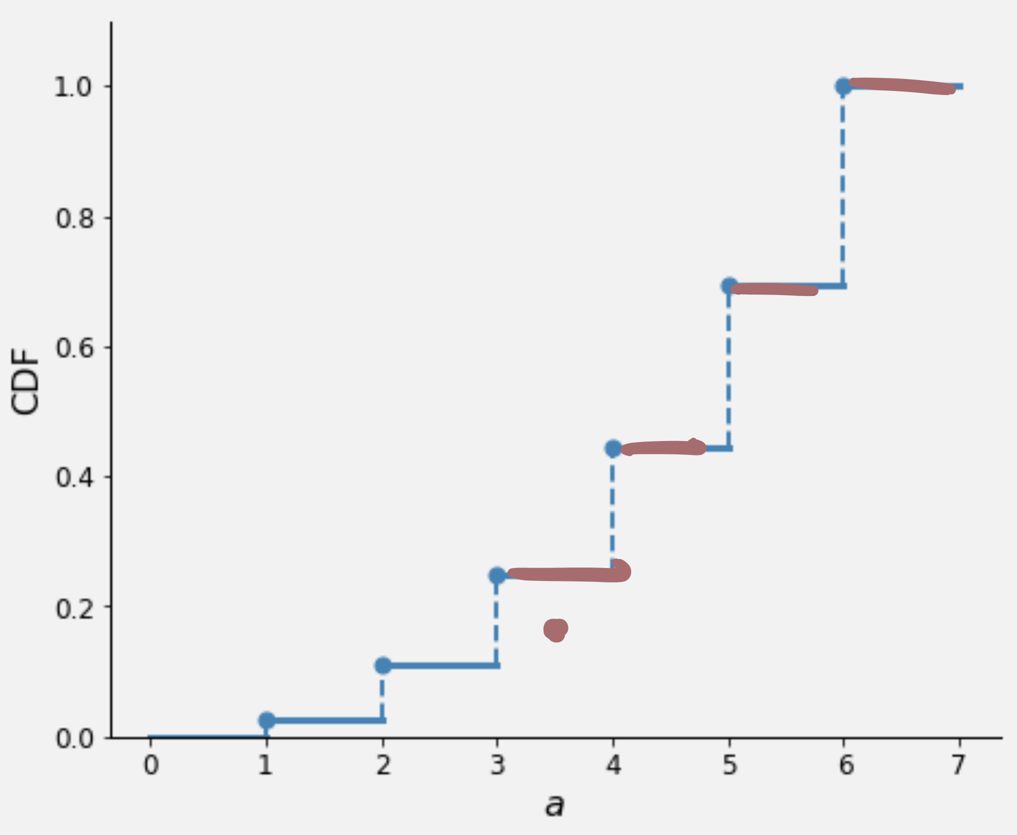
a	1	2	3	4	5	6
$f(a)$	$1/36$	$3/36$	$5/36$	$7/36$	$9/36$	$11/36$

a	1	2	3	4	5	6
$F(a)$	$\frac{1}{36}$	$\frac{4}{36}$	$\frac{9}{36}$	$\frac{16}{36}$	$\frac{25}{36}$	$\frac{36}{36}$

Visualizing PMFs and CDFs

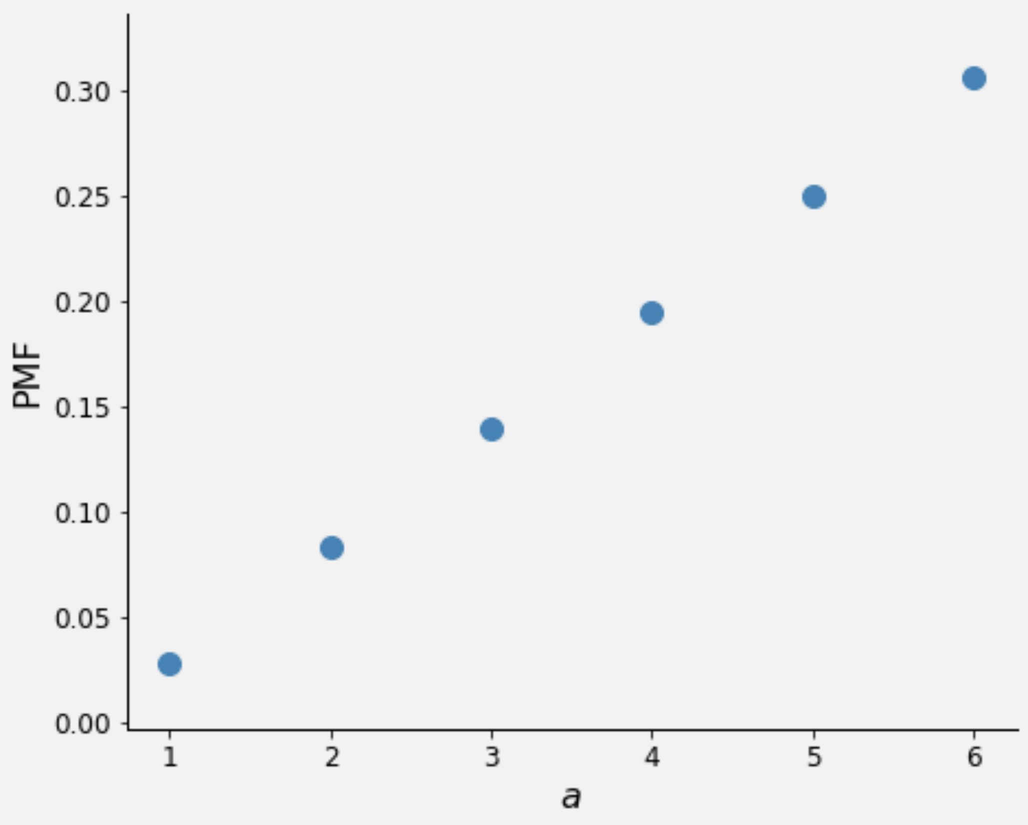
Example: Suppose you roll two fair, six-sided dice. Let X be a random variable representing the maximum of the two dice.

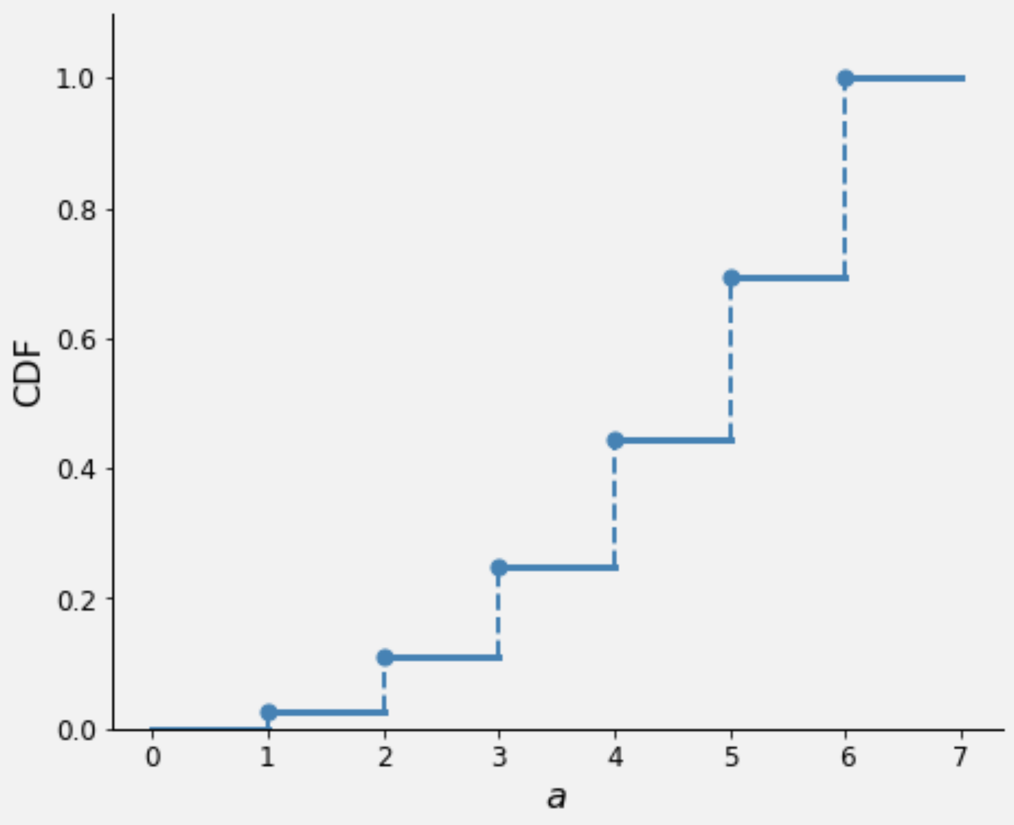




Visualizing PMFs and CDFs

Example: Suppose you roll two fair, six-sided dice. Let X be a random variable representing the maximum of the two dice.





Common Discrete RV Distributions

Discrete random variables can be categorized into different types or classes that each **model** different real-world situations.

The Bernouilli Distribution

The Bernouilli distribution is used to model experiments with only two possible outcomes, often referred to as "success" and "failure" and encoded as 1 and 0, respectively

Def: A discrete random variable X has a Bernouilli distribution with parameter p, where $0 \leq p \leq 1$, if its probability mass function is given by

$$f(1) = p_X(1) = P(X = 1) = p \quad \text{and} \quad p_X(0) = P(X = 0) = 1 - p$$

We denote this distribution by $Ber(p)$

The Bernouilli Distribution

The Bernouilli distribution is used to model experiments with only two possible outcomes, often referred to as "success" and "failure" and encoded as 1 and 0, respectively

Def: A discrete random variable X has a Bernouilli distribution with parameter p, where $0 \leq p \leq 1$, if its probability mass function is given by

$$p_X(1) = P(X = 1) = p \quad \textbf{and} \quad p_X(0) = P(X = 0) = 1 - p$$

We denote this distribution by $Ber(p)$

Question: Wouldn't it be nice if we could describe the PMF with a single equation?

$$p_x(x) = p^x (1-p)^{(1-x)}$$

$$p_x(1) = p$$

$$p_x(0) = 1-p$$

Counting Interludes

- We'll come back to the Bernouilli distribution in a minute.
- Believe it or not, counting comes up all over the place in probability, and therefore in data science, computer science, math, physics, etc.
- Some counting is easy: how many integers are in the interval [0,9] ?
- We're interested in counting problems that require more thought: Dan, Chris, Dave, Rhonda, and Tony line up at the coffee stand. How many different orders could they stand in?
- If there are 10 problems on an exam, and you need to get 7 correct to pass, how many different ways are there to pass?

Counting Interludes

- We'll talk about two important kinds of counting problems today:

- Counting **permutations** means counting the number of ways that a set of objects can be ordered or permuted.

Example: Dan, Chris, Dave, Rhonda, and Tony line up at the coffee stand. How many different orders could they stand in?

- Counting **combinations** means counting the number of ways that a set of objects can be combined into subsets

Example: If there are 10 problems on an exam, and you need to get 7 correct to pass, how many different ways are there to pass?

Permutations

abc

Questions:

- How many ways are there of ordering 1 object?

a ⇒ 1

- How many ways are there of ordering 2 objects?

ab, ba ⇒ 2

- How many ways are there of ordering 3 objects?

abc, bac, acb, bca
cba, cab

The Big Question: What is the formula for the number of ways you can order n objects?

$n!$

Permutations

Question: What if we have n objects, but want to count permutations of only r of them?

Example: How many three-character strings can we make if each character is a distinct letter from the English alphabet?

$$\underline{26} \times \underline{25} \times \underline{24}$$

Question: What is the general formula for r-permutations of n objects?

$$P(n,r) = \frac{n!}{(n-r)!} = P_{n,r}$$

Combinations

Counting **combinations** means counting the number of ways that a set of objects can be combined into subsets

Key Difference: When counting combinations, order **does not** matter

Example: How many three-character **combinations** can we make if each character is a distinct letter from the English alphabet?

PERMS: $\dfrac{26!}{(26-3)!}$

abc
cab
bac --- acb $\Big\}$ 6

COMBOS: $\dfrac{n!}{(n-r)!\,r!} = C_{n,r} = \binom{n}{r}$

Combinations

Example: How many three-character **combinations** can we make if each character is a distinct letter from the English alphabet?

Start with the number of 3-permutations of 26 letters:

If order doesn't matter, we're counting combinations multiple times.

Combinations

There are many different notations for combinations

The number of ways to choose r objects from a set of n can be written as

$$C(n, r) \quad \textbf{and} \quad C_{n,r} \quad \textbf{and} \quad \binom{n}{r}$$

Example: If there are 10 problems on an exam and you need to answer at least 7 problems correctly to pass, how many different ways are there to pass the test?

$$\binom{10}{7} + \binom{10}{8} + \binom{10}{9} + \binom{10}{10} =$$

Combinations

Example: A coin is flipped 10 times. How many possible outcomes have exactly 2 Heads?

$$\binom{10}{2}$$

Example: A coin is flipped 10 times. How many possible outcomes have Heads or fewer?

$$2^{10}$$

The Sum of Bernoulli Random Variables

Example: Suppose you show up to a quiz completely unprepared. The quiz has 5 questions, each with 4 multiple choice options. You decide to guess at the answer to each question in a completely random way. What is the probability that you get 3 questions correct?

$R_i \sim \text{Ber}(p = \frac{1}{4})$

For $i = 1, 2, \ldots, 5$ let

$$R_i = \begin{cases} 1 & \text{if the } i\text{th answer is correct} \\ 0 & \text{if the } i\text{th answer is incorrect} \end{cases}$$

Question: What can you say about R_i ?

The Sum of Bernoulli Random Variables

Example: Suppose you show up to a quiz completely unprepared. The quiz has 5 questions, each with 4 multiple choice options. You decide to guess at the answer to each question in a completely random way. What is the probability that you get 3 questions correct?

The number of correct answers on the quiz is given by the random variable X

$$X = R_1 + R_2 + R_3 + R_4 + R_5$$

Question: What values can the random variable X take on? $0, 1, 2, 3, 4, 5$

Question: What is the probability that you get 0 problems correct?

$$P(X=0) = \left(1-\frac{1}{4}\right)^5 = \left(\frac{3}{4}\right)^5$$

The Sum of Bernoulli Random Variables

Question: What is the probability that you get 0 problems correct?

$$\begin{aligned} P(X=0) &= P(R_1=0, R_2=0, R_3=0, R_4=0, R_5=0) \\ &= P(R_1=0)P(R_2=0)P(R_3=0)P(R_4=0)P(R_5=0) \\ &= \left(\frac{3}{4}\right)^5 \end{aligned}$$

Question: What is the probability that you get 1 problem correct?

$$P(X=1) = 5 \cdot \left(\frac{1}{4}\right)\left(\frac{3}{4}\right)^4$$

The Sum of Bernoulli Random Variables

Question: What is the probability that you get k problems correct?

$$P_X(k) = \binom{n}{k} p^k (1-p)^{n-k} \qquad k = 0, 1, \ldots, 5$$

$$\binom{n}{n} = \frac{n!}{(n-n)!\,n!} = \frac{n!}{0!\,n!}$$

The Binomial Distribution

Question: What is the probability that you get k problems correct?

$$P(X = k) = \binom{n}{k} \left(\frac{1}{4}\right)^k \left(\frac{3}{4}\right)^{n-k}$$

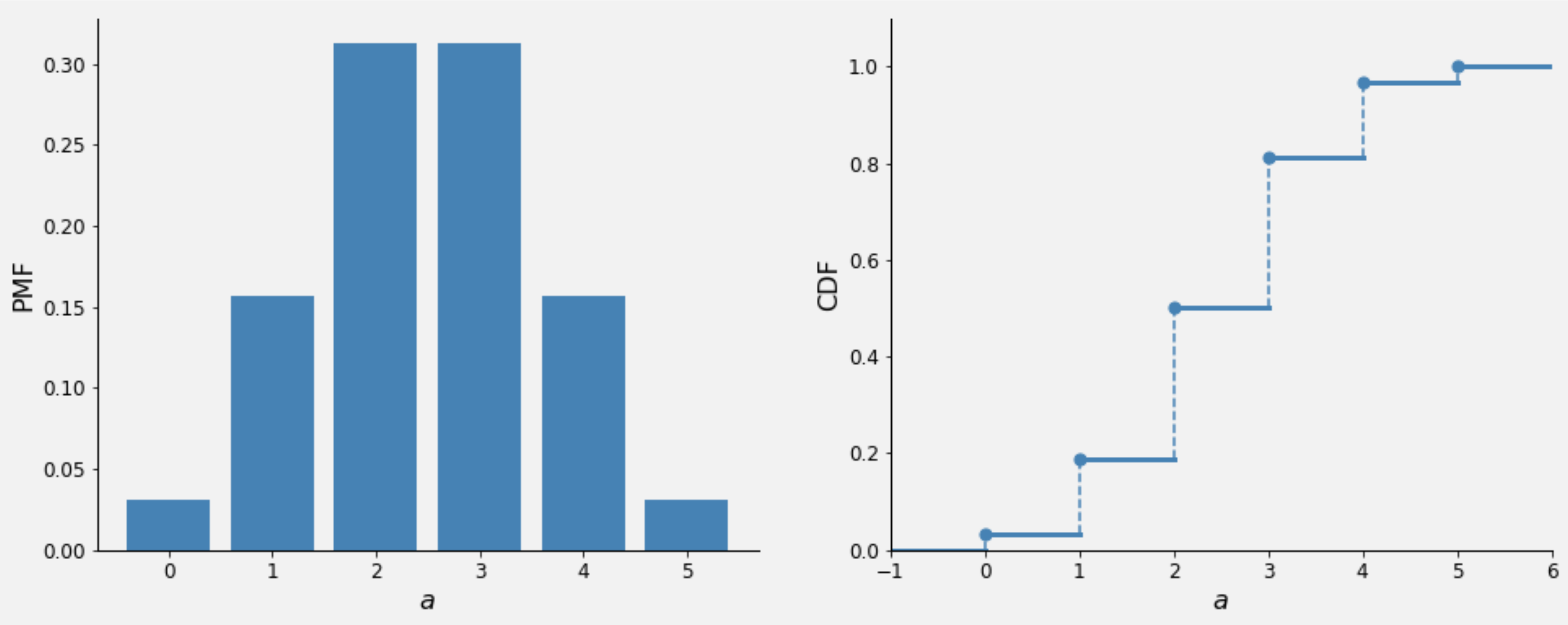
Def: A discrete random variable X has a binomial distribution with parameters n and p, where $n = 1, 2, \ldots$ and $0 \leq p \leq 1$, if it's probability mass function is given by

$$p_X(k) = P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \quad \text{for } k = 0, 1, \ldots, n$$

We denote this distribution by $Bin(n, p)$

The Binomial Distribution

Example: n = 5, p = 0.5

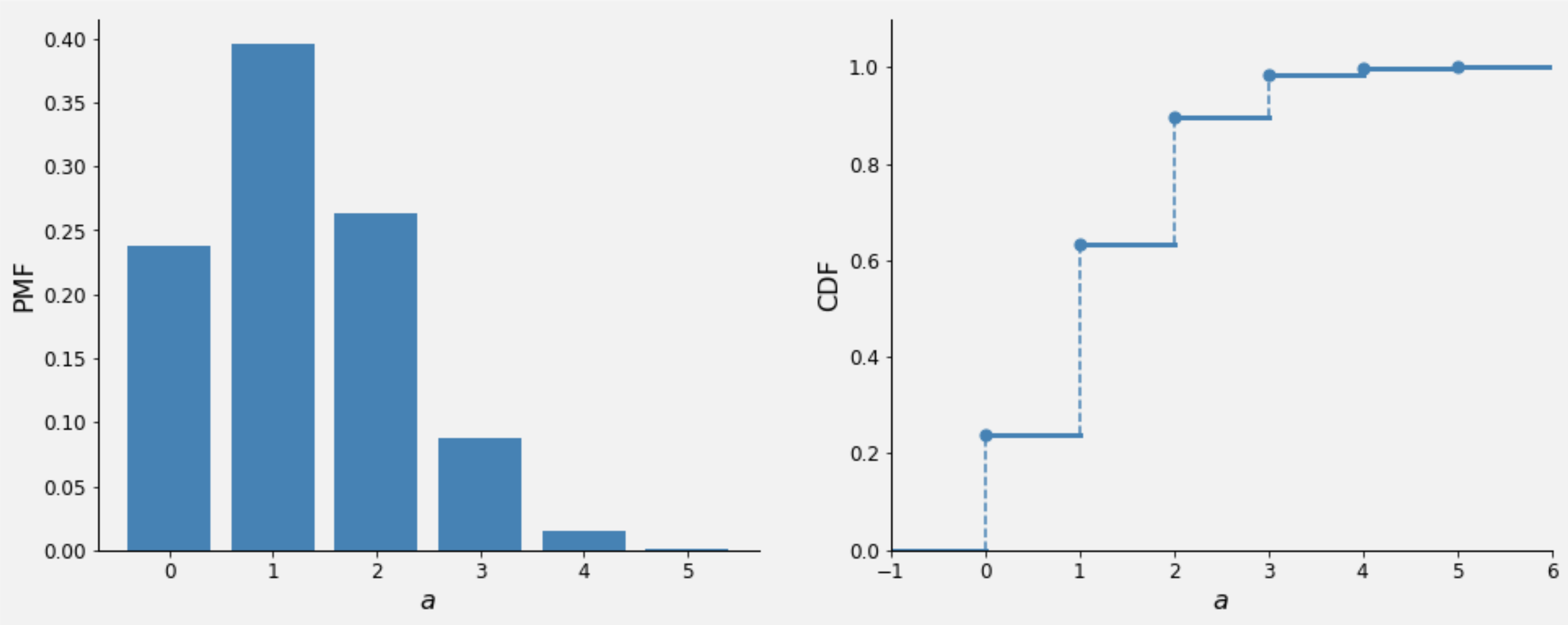


The Binomial Distribution

Example: $n = 5$, $p = 0.25$

The Binomial Distribution

Example: n = 5, p = 0.25

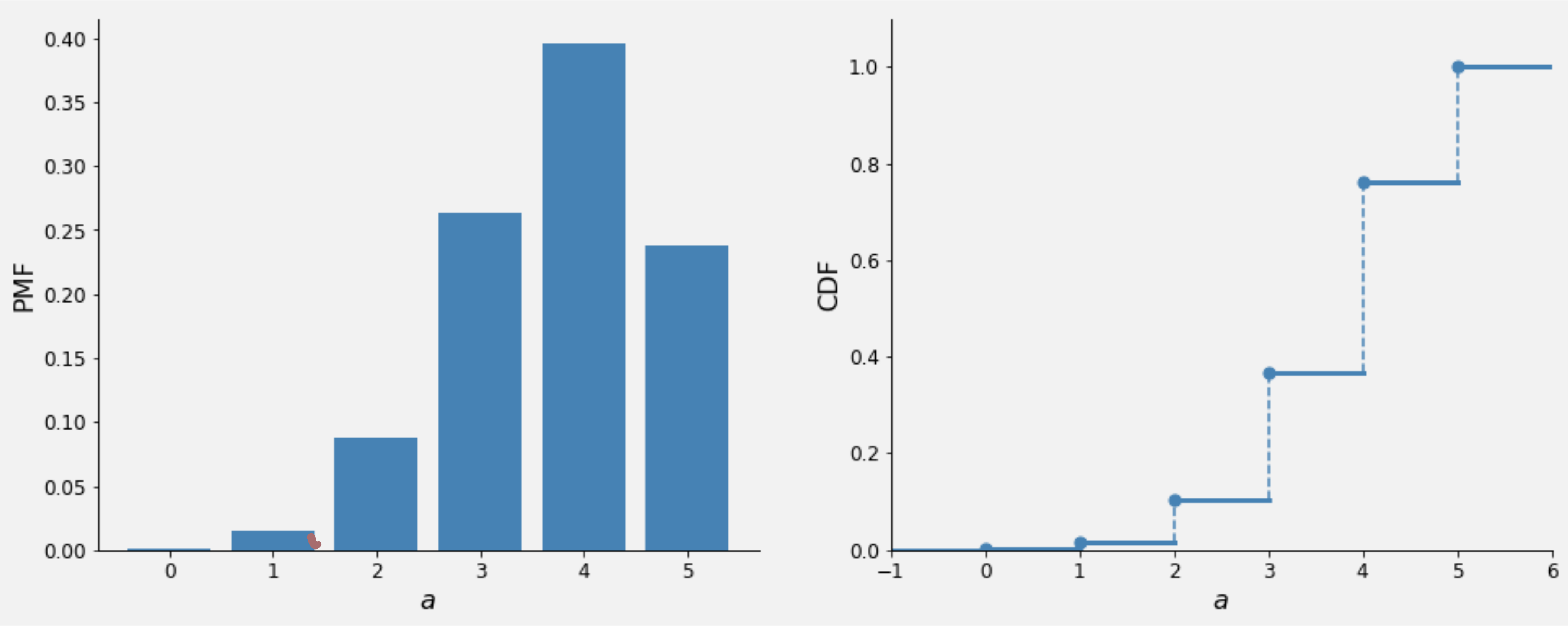


The Binomial Distribution

Example: $n = 5$, $p = 0.75$

The Binomial Distribution

Example: n = 5, p = 0.75



The Binomial Distribution

What **Assumptions** did we make in going from $Ber(p)$ to $Bin(n, p)$?

- Each of the n Bernouilli trials are independent
- Each Bernouilli trial has the same probability of success, p.

The Most Boring (but Common) of Them All

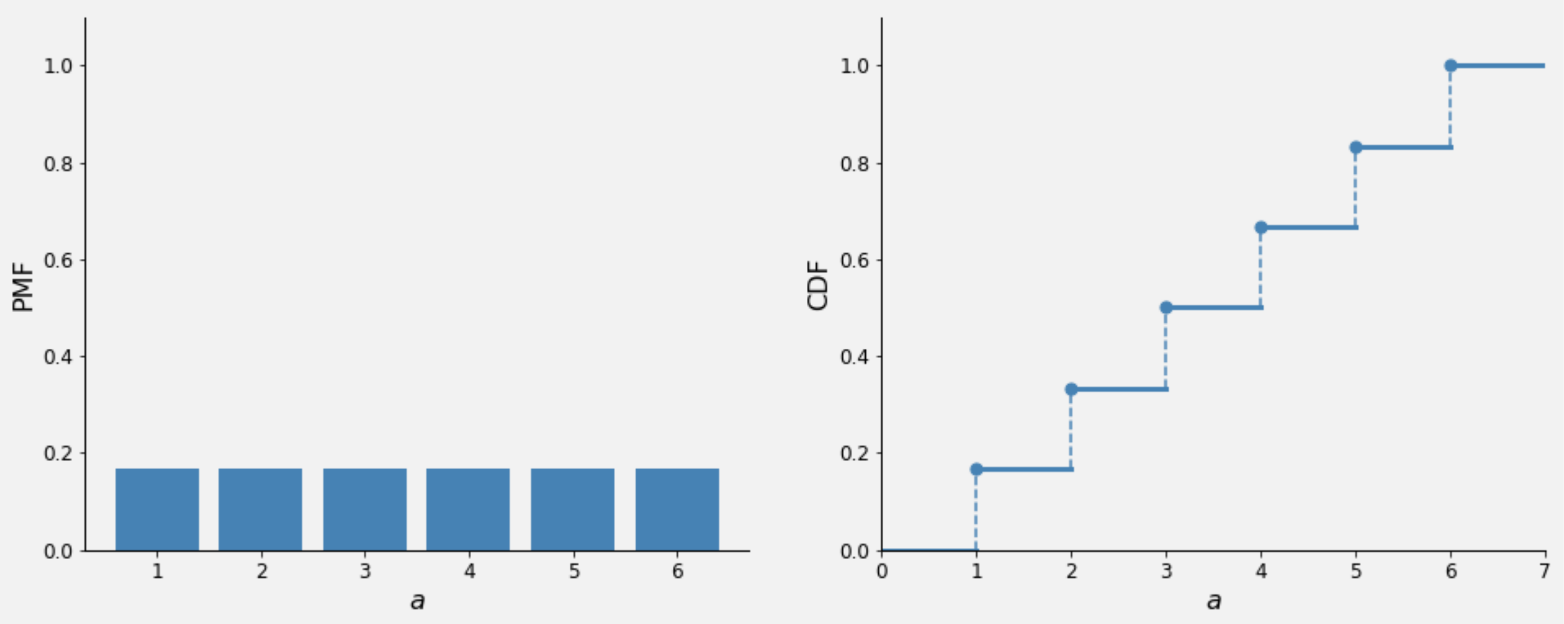
What is the distribution of a fair die?

The Most Boring (but Common) of Them All

What is the distribution of a fair die?

Def: A discrete random variable X has a discrete uniform distribution with parameters a, b, and n = b − a + 1 if

$$p_X(k) = \frac{1}{n} \quad \text{for } k = a, a+1, \ldots, b$$



OK! Let's Go to Work!

Get in groups, get out laptop, and open the Lecture 7 In-Class Notebook

Let's:

- See some more examples of computing PMFs and CDFs
- Look at some more examples of the Binomial distribution
- Learn how to sample from the Bernoulli and Binomial distributions in Numpy