

Lecture 6

Discrete Random Variables

Grace Tompkins

Last modified — 06 May 2026

Learning Outcomes

By the end of this lecture, students are anticipated to be able to:

- Define a discrete random variable
- Define and identify a probability mass function
- Define, identify, and apply common families of discrete distributions

Review of Random Variables

- A random variable is a *function* that maps from the sample space to \mathbb{R} .
- Typically denoted by X, Y, Z (letters near the end of the alphabet)
- The *distribution* of a random variable is the collection of probabilities associated with every subset of \mathbb{R} (we ignore details).

Why is this useful?

- Random variables will allow us to quit worrying about the sample space, and simply deal with the abstraction.
- Any properties we learn about the distribution of the random variable apply to any sample space that could possibly be “modelled” with that Random Variable.
- We can leave the world of dice and coins and just talk about math.

We start with **Discrete random variables**.

1 Discrete random variables

Discrete random variables

DEFINITION

We say that a random variable X is **discrete** if there exists a countable set $K = \{x_1, x_2, \dots\}$ such that

$$\mathbb{P}(X \in K) = 1.$$

- In particular, this means that $\{x : P(X = x) > 0\}$ is countable:

I can count the x that have positive probability.

- For discrete RVs, we call the set $\{x : P(X = x) > 0\}$ the **support** of X .
- These are the types of RV's we've worked with so far. We'll visit the other extreme next class

Examples of Experiments Described by Discrete Random Variables

- Toss a fair coin 3 times, count the total number of Heads.
- Roll a 6-sided die until you see 6, count the number of rolls.
- Count the number of people that arrive at the bus stop in some amount of time.
- Will it be rainy (-6), cloudy (2), or sunny ($+10$) on campus today?

Probability Mass Function (PMF)

- For discrete random variables, we can be explicit about the probability associated with each value in its support

DEFINITION

For a discrete random variable, its probability (mass) function is the function $p_X : \mathbb{R} \rightarrow [0, 1]$ defined by

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

It is occasionally written $f_X(x) = \mathbb{P}(X = x)$.

- Each of the examples of discrete random variables has a PMF. Once we write it down, we know everything there is to know.

A Rainy Day

EXERCISE: RAINY DAYS

Suppose it rains 30% of the time, is cloudy but not raining 30% of the time, and sunny 40% of the time. Your happiness is given by the random variable Z with $Z = -6, 2,$ and 10 respectively.

- What is the support of Z ?
- What is the PMF of Z ?

a) The values of z with a positive probability are $-6, 2, 10$

b)

z	$p_Z(z)$
-6	0.3
2	0.3
10	0.4
otherwise	0

or,

$$p_Z(z) = 0.3I_{(-6)}(z) + 0.3I_{(2)}(z) + 0.4I_{(10)}(z)$$

A Rainy Day

Fair Coins

EXERCISE: TOSS A COIN 3 TIMES

Toss a fair coin 3 times. Let X be the number of heads.

- What is the support of X ?
- What is the $p_X(x)$?

a) Support of X : $\{0, 1, 2, 3\}$

b) $\Omega = \{HHH, HHT, HTH, THH, TTH, THT, HTT, TTT\}$

x	$p(x)$
0	$1/8$
1	$3/8$
2	$3/8$
3	$1/8$
otherwise	0

Finding Probabilities of Events

The PMF can be used to find the probabilities of events.

We have that for some event A and a RV X ,

$$\mathbb{P}(X \in A) = \sum_{x \in A} p_X(x).$$

~~$A \subseteq \Omega$~~

$$A = \{z > 0\}$$

💡 EXAMPLE

Recall: it rains 30% of the time, is cloudy but not raining 30% of the time, and sunny 40% of the time. Your happiness is given by the random variable Z with $Z = -6, 2, \text{ and } 10$ respectively.

- Let A be the event that you are happier than 0. Find $\mathbb{P}(Z \in A)$.

$$P(Z \in A) = \sum_{x \in A} p_Z(z) = P_Z(2) + P_Z(10) = 0.3 + 0.4 = 0.7$$

$$\text{Aside: } P(A) = P(Z > 0) = 1 - P(Z \leq 0)$$

Another PMF

For a fixed number $\theta \in [0, 1]$, define a discrete random variable X with PMF given by:

$$p_X(x; \theta) = \begin{cases} (1 - \theta)^2 & x = 0, \\ 2\theta(1 - \theta) & x = 1, \\ \theta^2 & x = 2, \\ 0 & \text{else.} \end{cases}$$

- Different values of θ will give different PMFs.

$$p_X(x; \theta = 0.1) = \begin{cases} 0.81 & x = 0, \\ 0.18 & x = 1, \\ 0.01 & x = 2 \\ 0 & \text{else.} \end{cases}$$

By including a **parameter** θ , we are able to express many distributions (a **family of distributions**) with a single functional form.

2 Discrete families

Bernoulli

An experiment with 2 outcomes (success/failure).

I will provide you with
PMFs for these families
on an exam

DEFINITION

Suppose the probability of success is $\theta \in (0, 1)$. A RV X with PMF given by

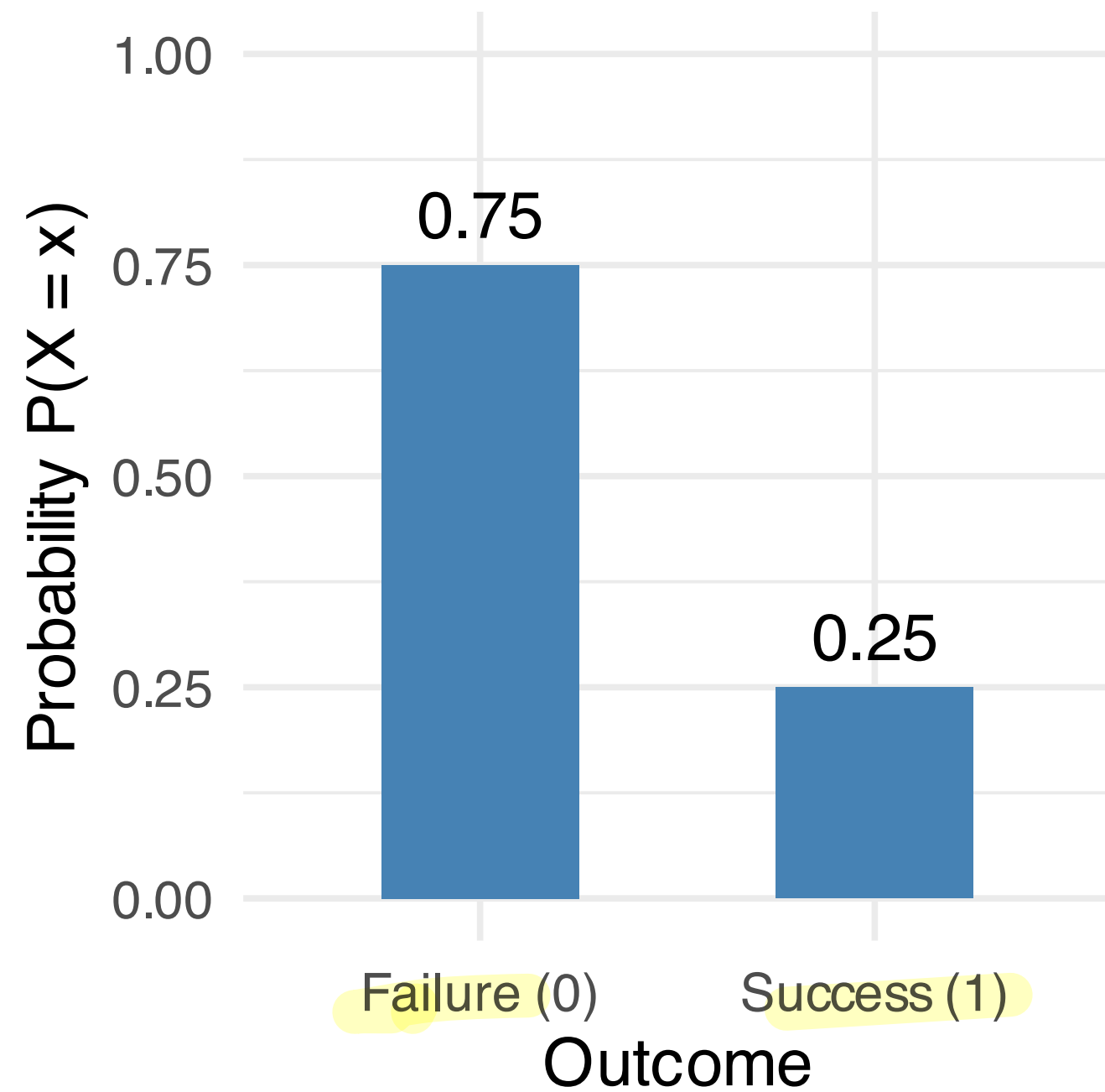
$$p_X(x; \theta) = \theta^x (1 - \theta)^{1-x} I_{\{0,1\}}(x),$$

is said to have the $\text{Bern}(\theta)$ distribution.

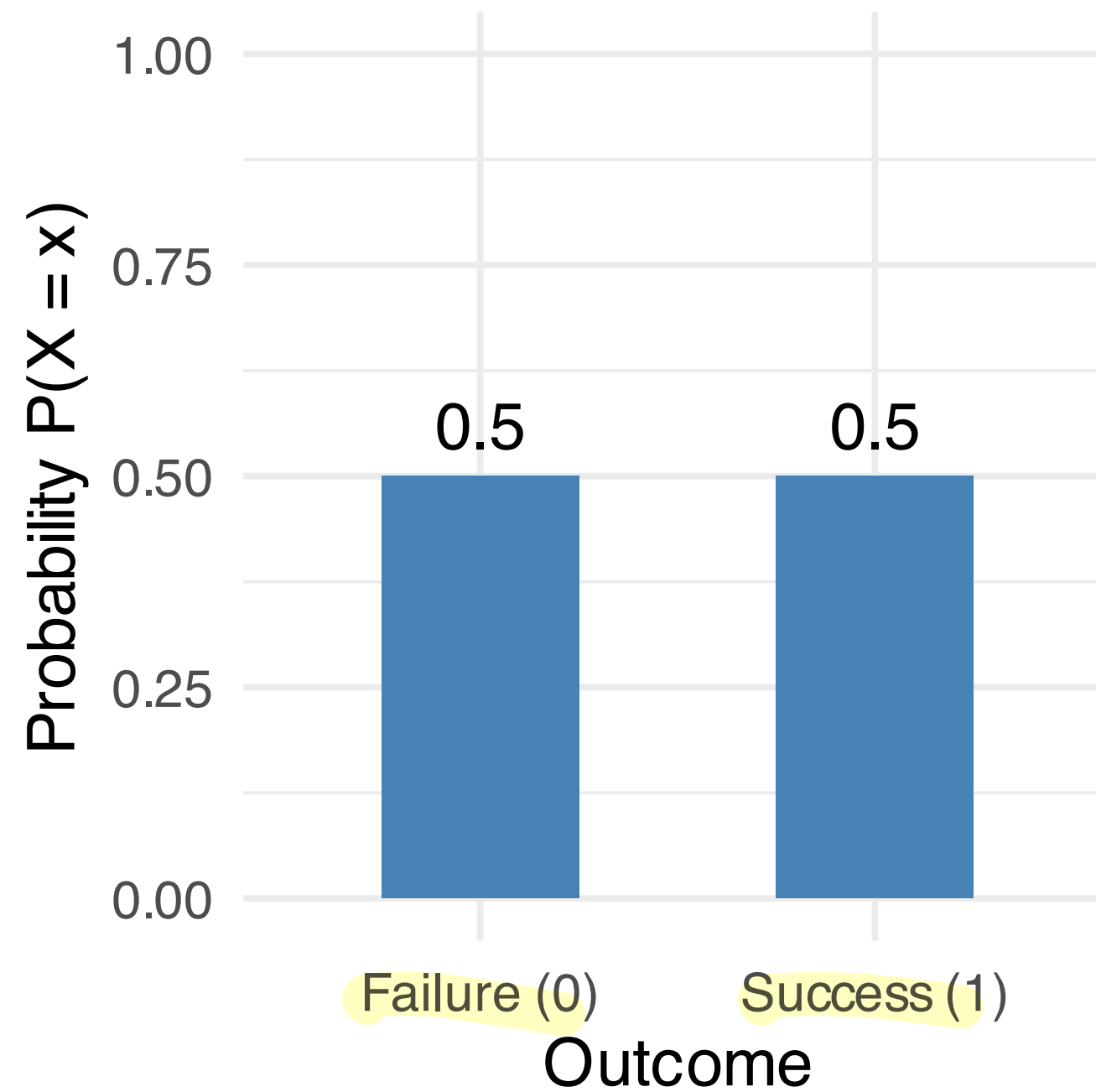
- If $\theta = 0.5$, then we have our favourite fair-coin friend.
- But we can let θ be numbers other than 0.5.

Bernoulli

Bernoulli PMF ($\theta = 0.25$)



Bernoulli PMF ($\theta = 0.5$)



Binomial

Used to determine the number of “successes” (x) in a sequence of n independent Bernoulli trials, each with the same probability of success (θ).

DEFINITION

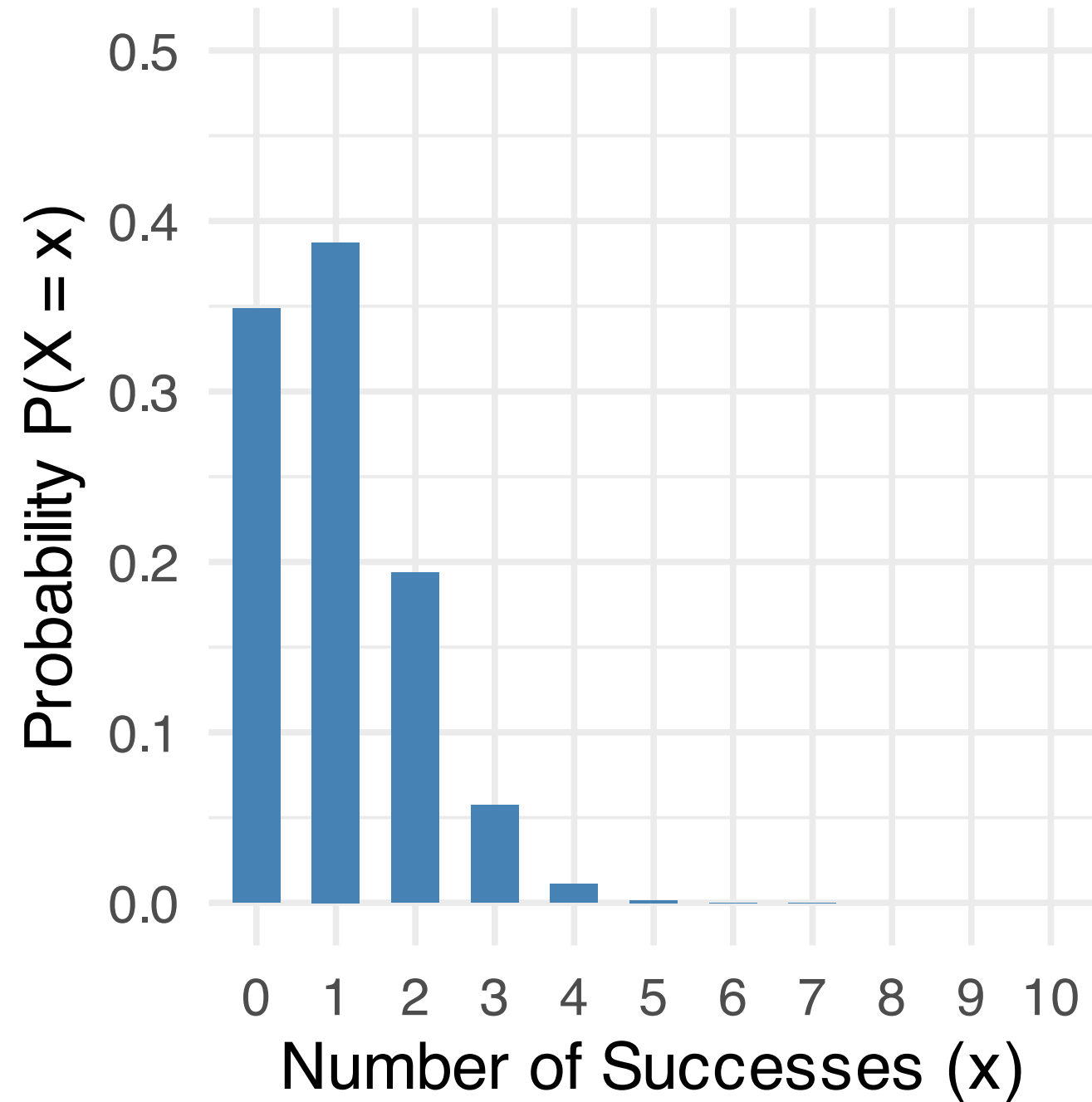
Suppose the probability of success is $\theta \in (0, 1)$. A RV X with PMF given by

$$p_X(x; n, \theta) = \binom{n}{x} \theta^x (1 - \theta)^{n-x} I_{\{0,1,\dots,n\}}(x),$$

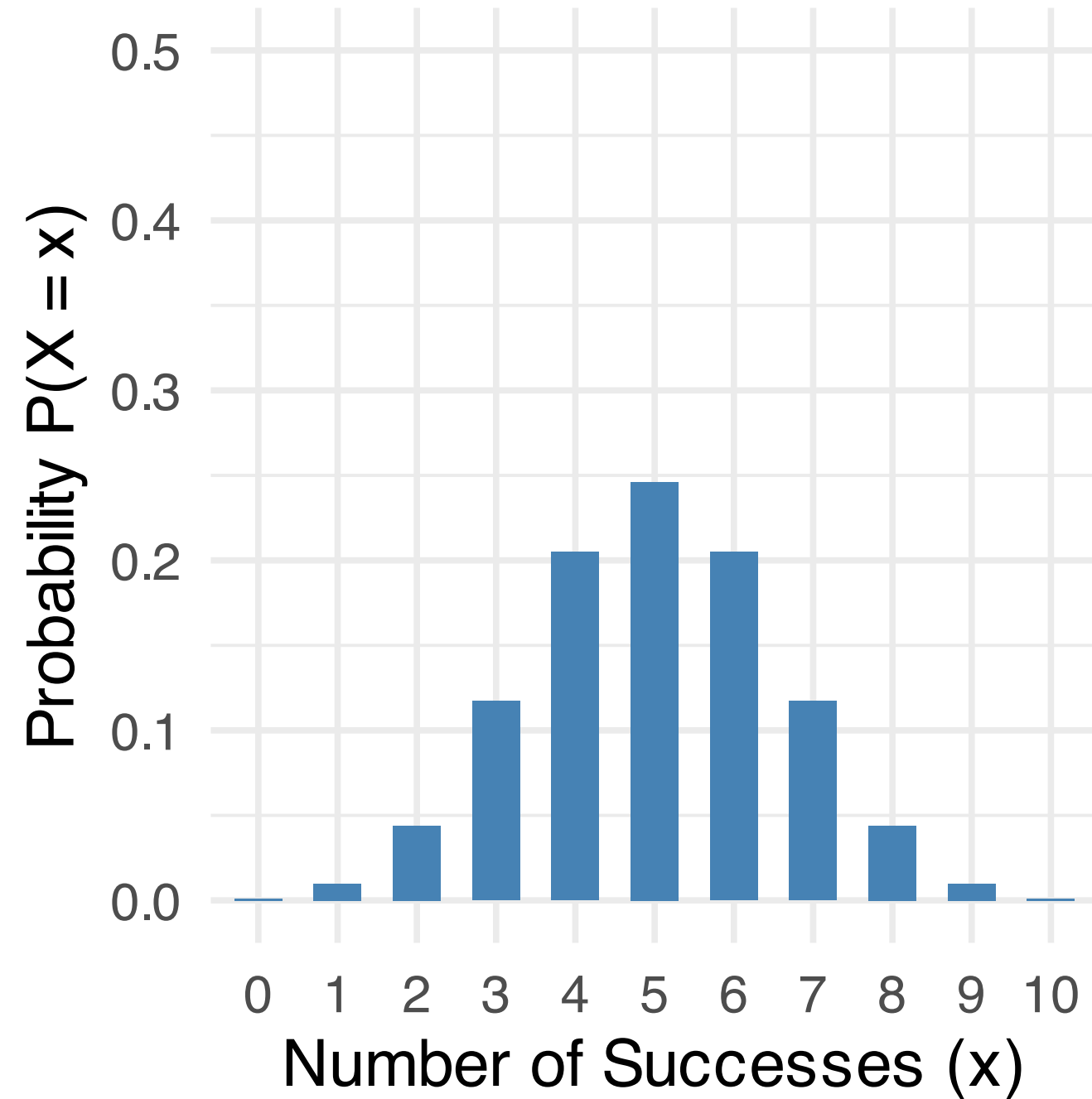
is said to have the **Binom**(n, θ) distribution.

Binomial

Binomial PMF ($\theta = 0.1, n = 10$)



Binomial PMF ($\theta = 0.5, n = 10$)



Using the Binomial

$$p_X(x; n, \theta) = \binom{n}{x} \theta^x (1 - \theta)^{n-x} I_{\{0,1,\dots,n\}}(x)$$

EXERCISE: GUESSING ON THE MIDTERM

- Suppose Midterm λ has 5 True/False questions. What is the probability that you get at least 4 correct answers by random guessing?
- Suppose Midterm λ has 2 multiple choice questions with 5 options. What is the probability that you get both correct by random guessing?

$$\text{a) } P(\text{Success}) = \theta = 0.5 \\ n = 5$$

$$\text{b) } P(\text{Success}) = \theta = 1/5 = 0.2 \\ n = 2$$

Using the Binomial

$$\text{a) } P(\text{Success}) = \theta = 0.5$$

$$n = 5$$

X = number of successes

$$\begin{aligned} P(X \geq 4) &= P(X=4) + P(X=5) \\ &= p_X(4; n=5, \theta=0.5) + p_X(5; n=5, \theta=0.5) \\ &= \binom{5}{4} (0.5)^4 (1-0.5)^{5-4} + \binom{5}{5} (0.5)^5 (1-0.5)^{5-5} \\ &= 5/32 + 1/32 \\ &= 6/32 = 3/16 \approx 0.1875 \end{aligned}$$

$$\text{b) } P(\text{Success}) = \theta = 1/5 = 0.2$$

$$n = 2$$

X = number of successes

$$\begin{aligned} P(X=2) &= p_X(2; n=2, \theta=0.2) \\ &= \binom{2}{2} (0.2)^2 (1-0.2)^{2-2} \\ &= 0.2^2 (1) \\ &= 1/25 = 0.04 \end{aligned}$$

Poisson

DEFINITION

Let $\lambda > 0$. A RV Y with PMF given by

$$p_Y(y; \lambda) = \frac{\lambda^y}{y!} e^{-\lambda} I_{\{0,1,2,\dots\}}(y),$$

is said to have the Poiss(λ) distribution.

Sometimes the PMF is written as:

$$p_Y(y; \lambda t) = \frac{(\lambda t)^y}{y!} e^{-(\lambda t)} I_{\{0,1,2,\dots\}}(y),$$

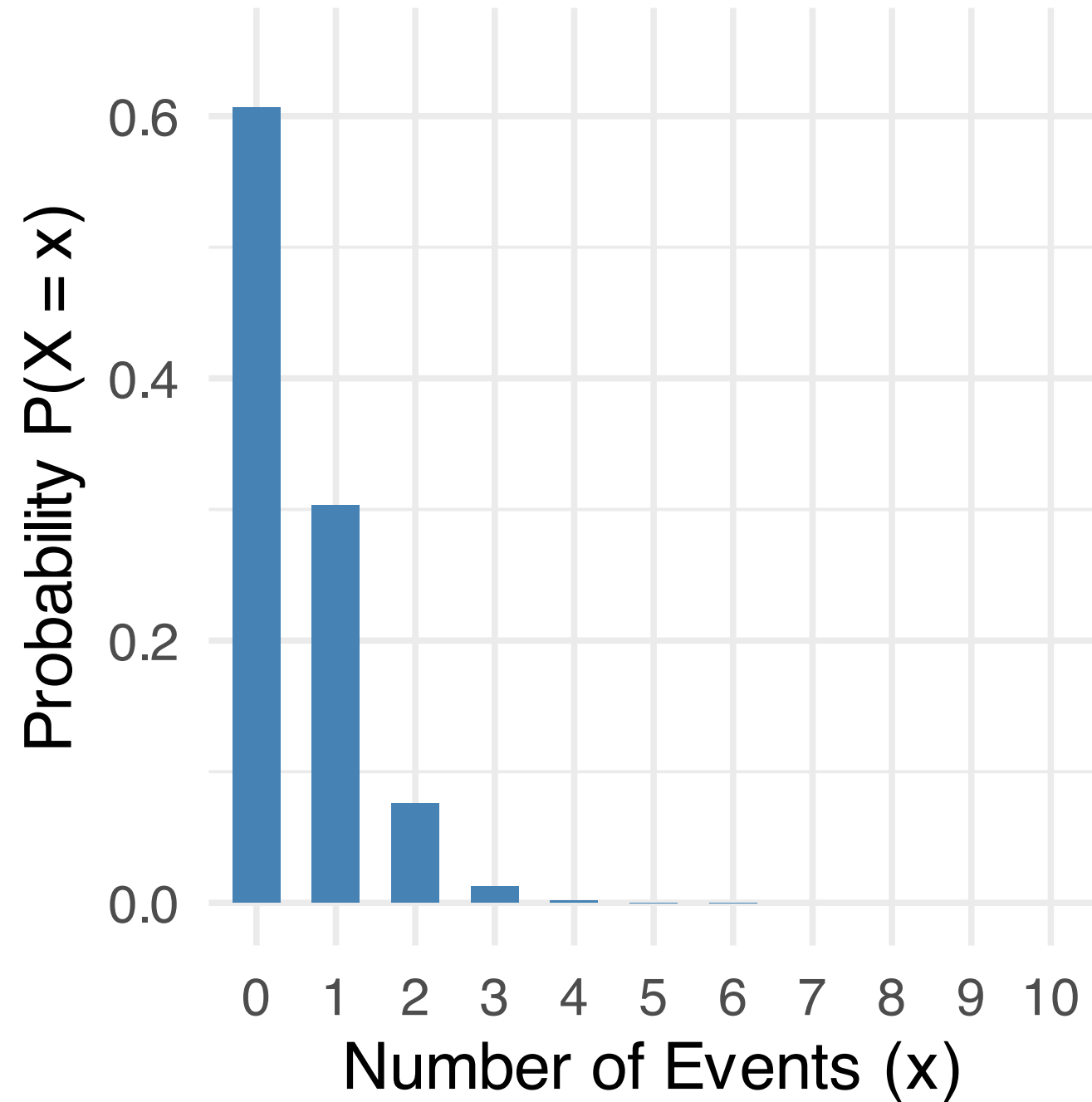
where λ is “per unit time” and t is the number of seconds/minutes/days/etc.

Poisson

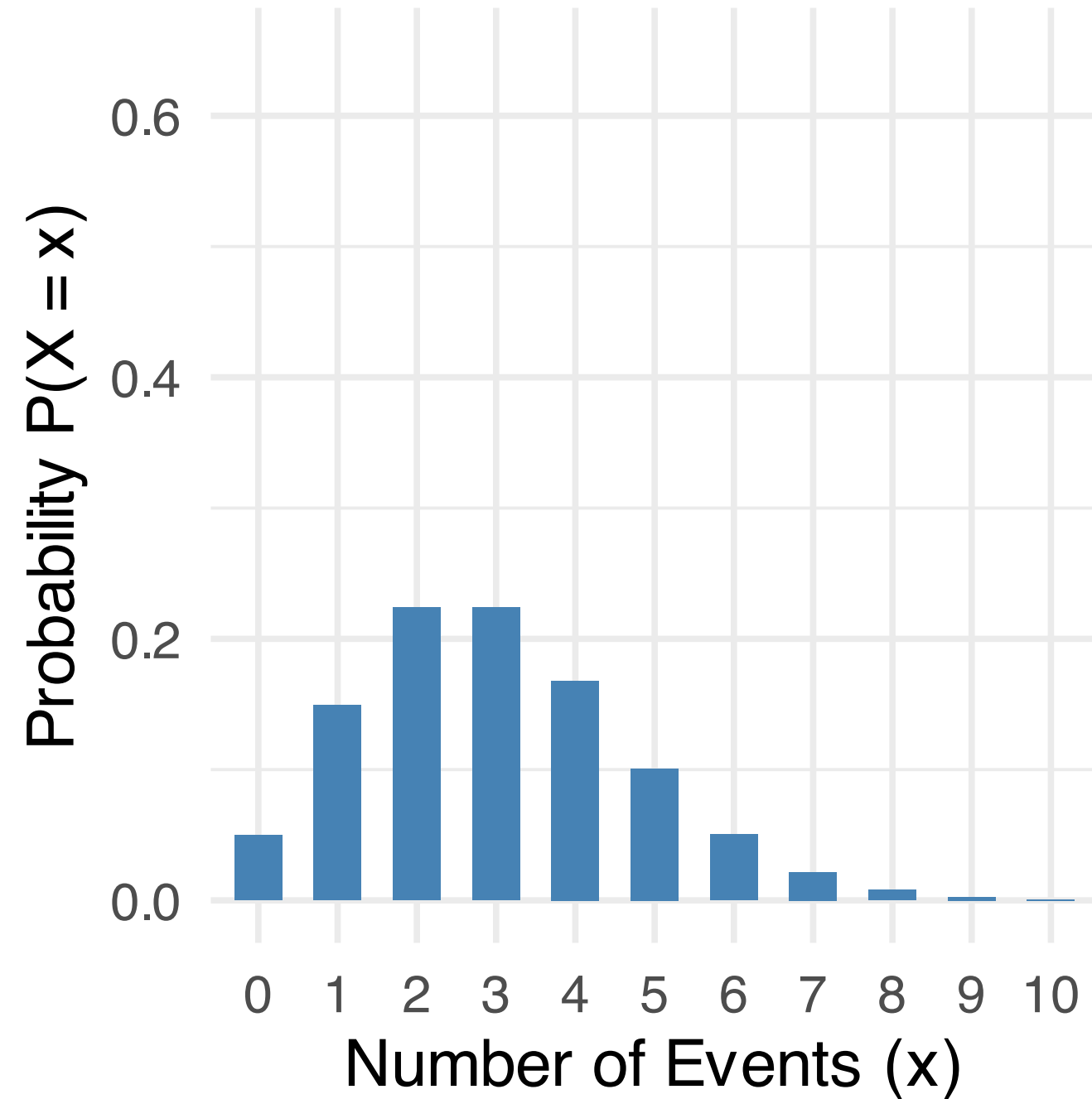
- The Poisson distribution is good for modelling counts of things, often per unit of time. This is not the only distribution that is good for modelling counts, and we'll see reasons to potentially choose a different one later.
- λ is often the average number of events per unit of time. X can be the count of events in a particular unit of time

Poisson

Poisson PMF ($\lambda = 0.5$)



Poisson PMF ($\lambda = 3$)



Poisson

💡 EXAMPLE

A coffee shop receives an average of 3 customers per minute during the morning rush. What is the probability that exactly 5 customers arrive in a given minute?

Here, λ is the average number of customers per minute. If we let X be the number of customers in a given minute, then $X \sim \text{Poiss}(3)$. Using the formula:

$$p_X(x; \lambda) = \frac{\lambda^x}{x!} e^{-\lambda} I_{\{0,1,2,\dots\}}(x)$$

$$p_X(x=5) = \frac{3^5}{5!} e^{-3} = 0.1008$$

Poisson Maximization

EXERCISE: MAXIMIZING THE POISSON

Suppose that $Y \sim \text{Poiss}(\lambda)$, and you observe $Y = 5$. What value of λ makes $\mathbb{P}(Y = 5)$ as large as possible?

$$p_Y(y; \lambda) = \frac{\lambda^y}{y!} e^{-\lambda} I_{\{0,1,2,\dots\}}(y)$$

Hint: Take the log! ... and a derivative

Poisson Maximization

$$P(Y=5) = \frac{\lambda^5}{5!} e^{-\lambda}$$

$$\hat{\lambda} = \arg \max_{\lambda} \left\{ \frac{\lambda^5}{5!} e^{-\lambda} \right\}$$

$$= \arg \max_{\lambda} \left\{ \log \left(\frac{\lambda^5}{5!} e^{-\lambda} \right) \right\}$$

$$= \arg \max_{\lambda} \left\{ 5 \log(\lambda) - \log(5!) - \lambda \right\}$$

$$\Rightarrow 0 = \frac{d}{d\lambda} \left\{ 5 \log(\lambda) - \log(5!) - \lambda \right\}$$

$$0 = \frac{5}{\lambda} - 1$$



$$\hat{\lambda} = 5$$

Poisson Maximization

Try at home: ensure that this is a maximum with the second derivative test!

Geometric Distribution

DEFINITION

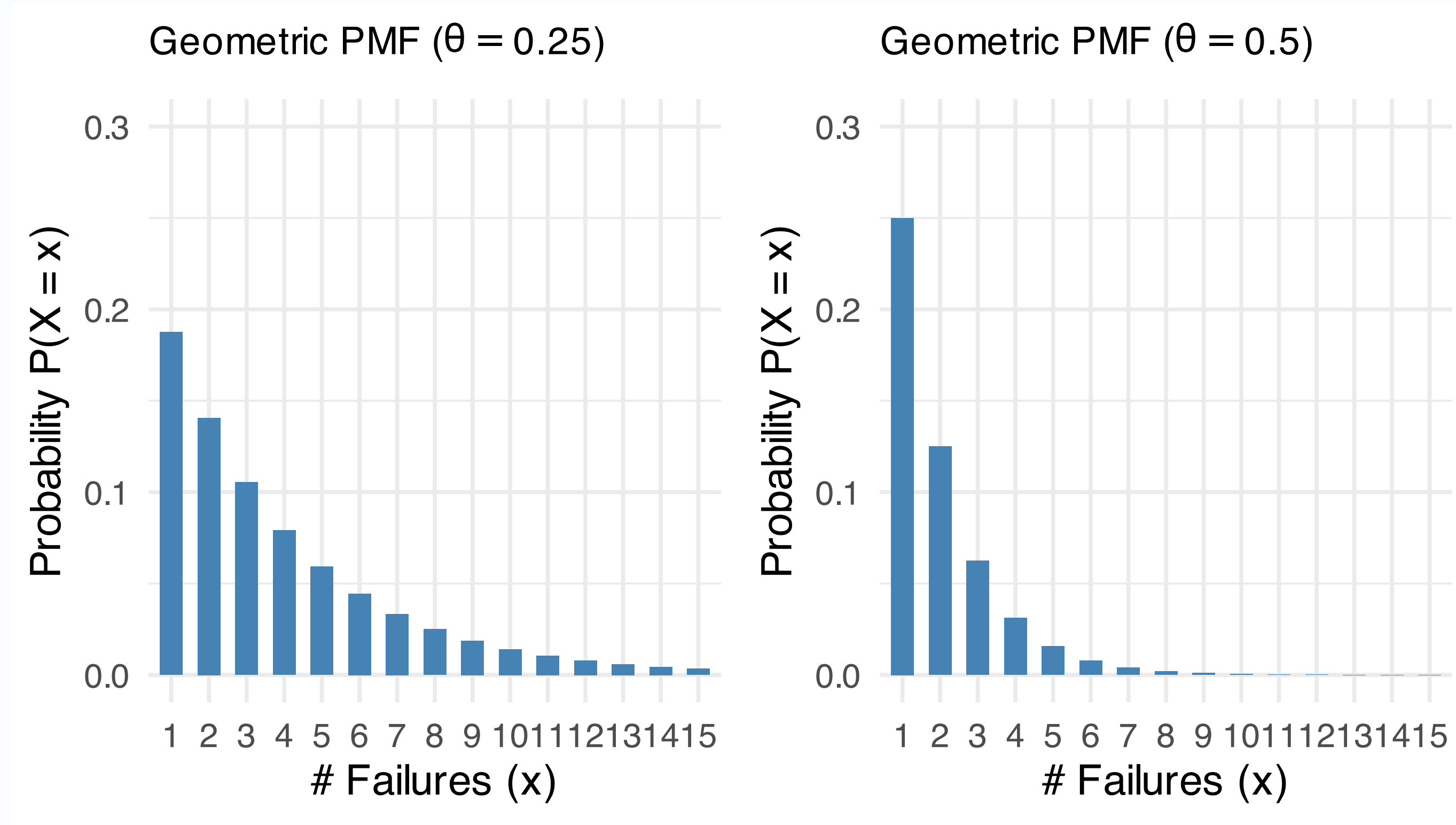
Suppose the probability of success is $\theta \in (0, 1)$. A RV X with PMF given by

$$p_X(x; \theta) = (1 - \theta)^x \theta I_{\{0,1,2,\dots\}}(x)$$

is said to have a $\text{Geom}(\theta)$ distribution.

- X could be the number of failures before a success, where we have a sequence of n independent Bernoulli trials, each with the same probability of success.

Geometric Distribution



Geometric Distribution

EXAMPLE

What is the probability of flipping 3 coins before seeing the first tails?

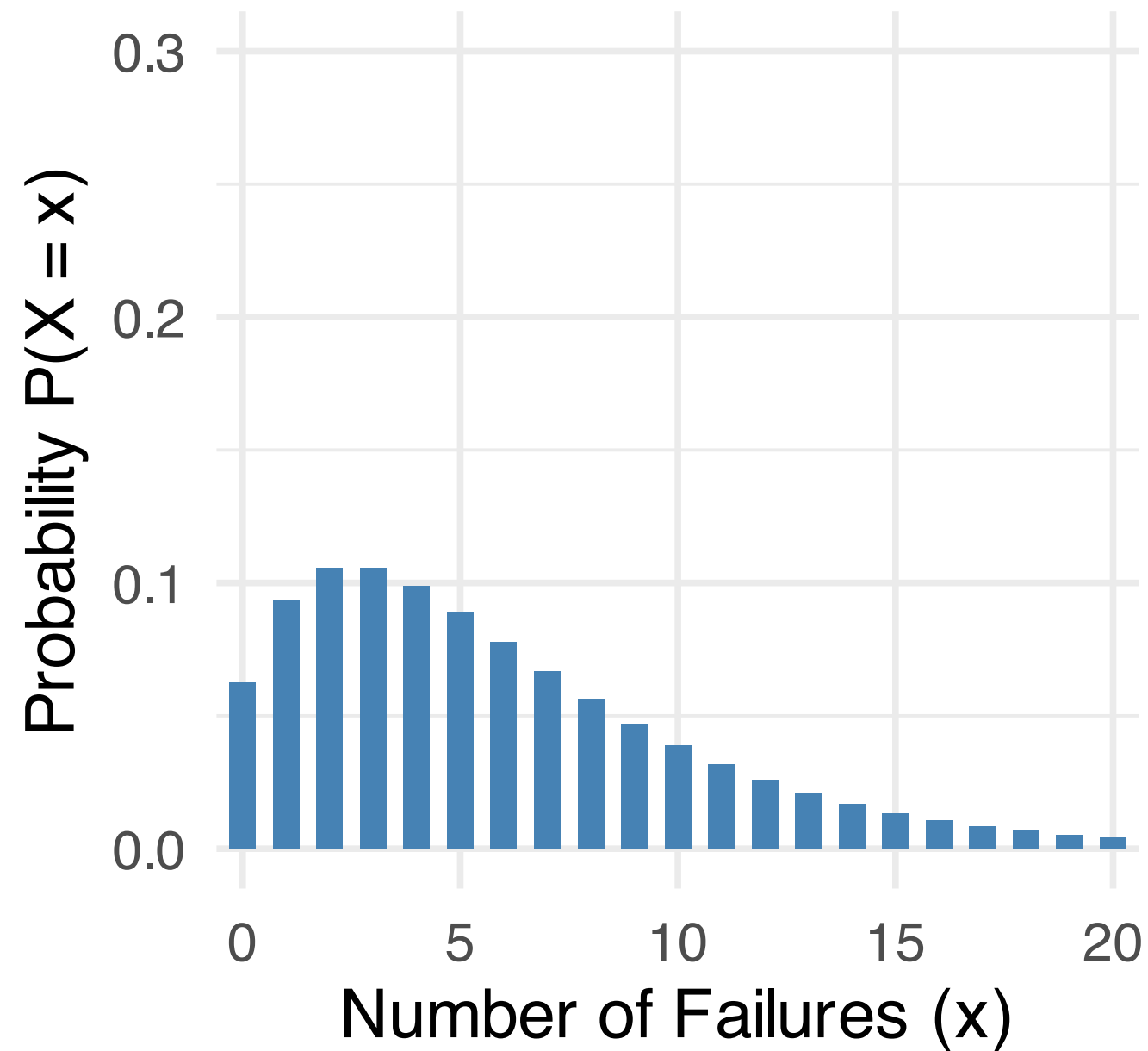
We can let X be the number of heads (failures) before the first tail (success). Then:

$$p_X(x = 3; 0.5) = (1 - 0.5)^3(0.5) = 0.0625$$

Negative Binomial

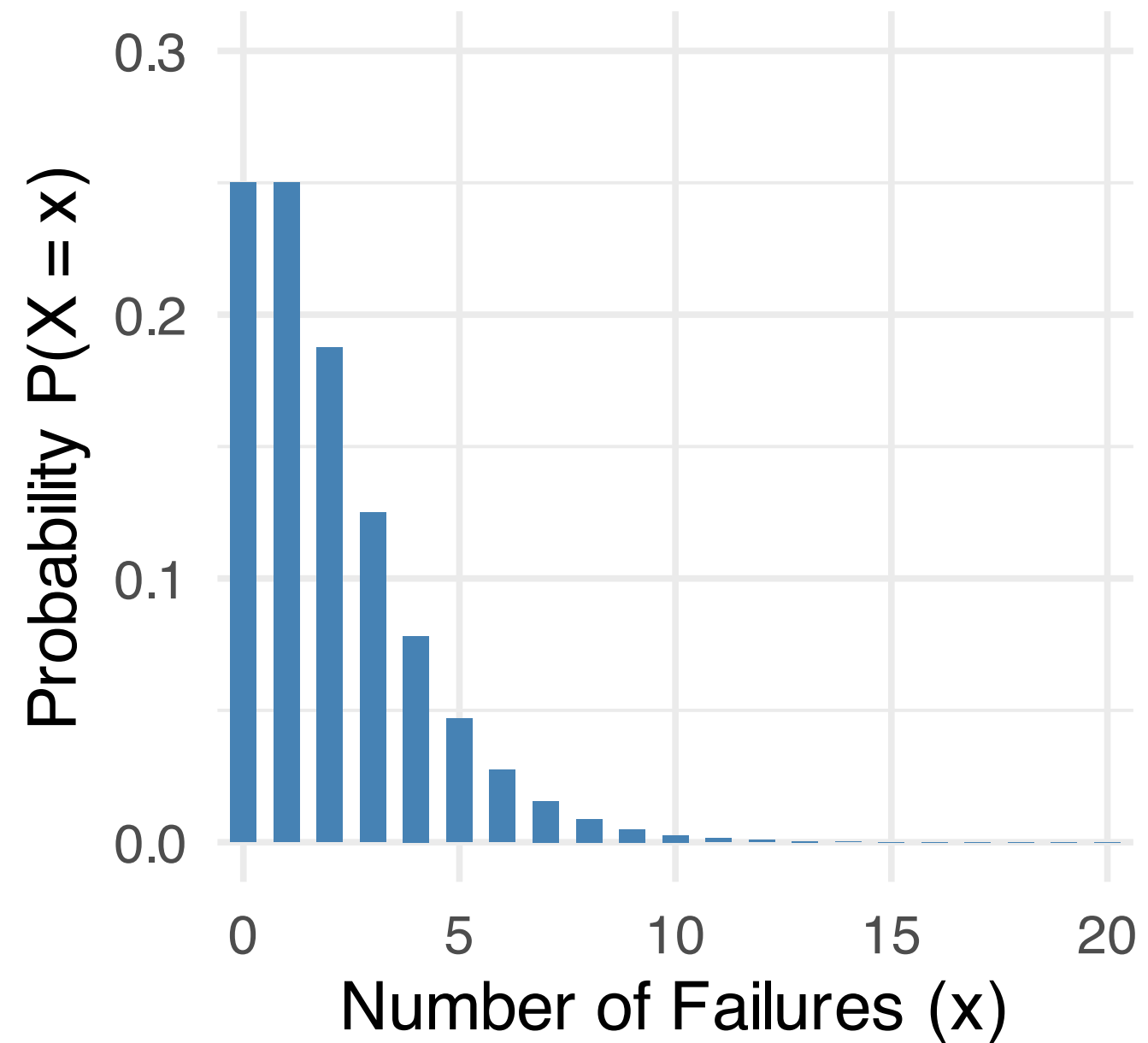
Negative Binomial PMF

$$\theta = 0.25, r = 2$$



Negative Binomial PMF

$$\theta = 0.5, r = 2$$



Negative Binomial

DEFINITION

Suppose again the probability of success is $\theta \in (0, 1)$. The random variable Y with PMF given by

$$p_Y(y; \theta, r) = \binom{r - 1 + y}{y} \theta^r (1 - \theta)^y I_{\{0,1,2,\dots\}}(y)$$

is said to have a NegBinom(r, θ) distribution.

- This distribution is useful in determining the probability that y failures appear before the r th success in a repeated experiment, like a coin flip.
- This is a generalization of the geometric distribution.

Negative Binomial

EXERCISE: FLIPS

What is the probability of flipping 3 heads before seeing the 2nd tails?

$$\theta = 0.5$$

$$r = 2$$

Y = number of heads (fails) before 2nd tail (success)

$$P(Y = 3) = P_Y(3; r = 2, \theta = 0.5) = \binom{2 - 1 + 3}{3} 0.5^2 (1 - 0.5)^3$$
$$\approx 0.125$$

Hypergeometric

📖 DEFINITION

Suppose we have a finite population of size N containing K successes, and a sample size of n . A random variable X with PMF

$$p_X(x; N, K, n) = \frac{\binom{K}{x} \binom{N-K}{n-x}}{\binom{N}{n}} \mathbb{1}_{\{x \in \mathcal{Z}\}}, \quad \max(0, n+M-N) \leq y \leq \min(n, M)$$

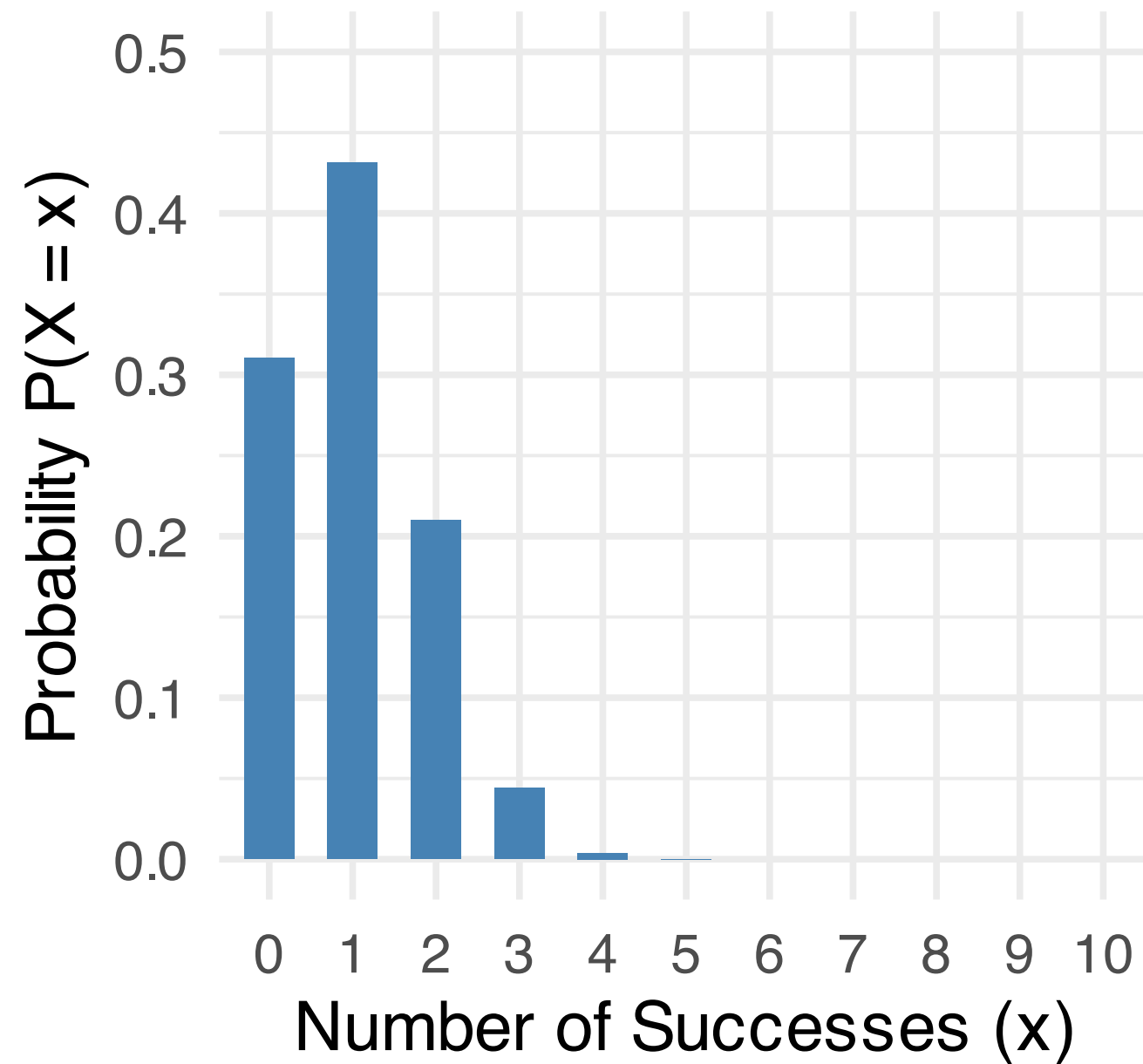
is said to have a hypergeom distribution.

- Can be used to determine the probability that there are x successes out of n draws without replacement.

Hypergeometric

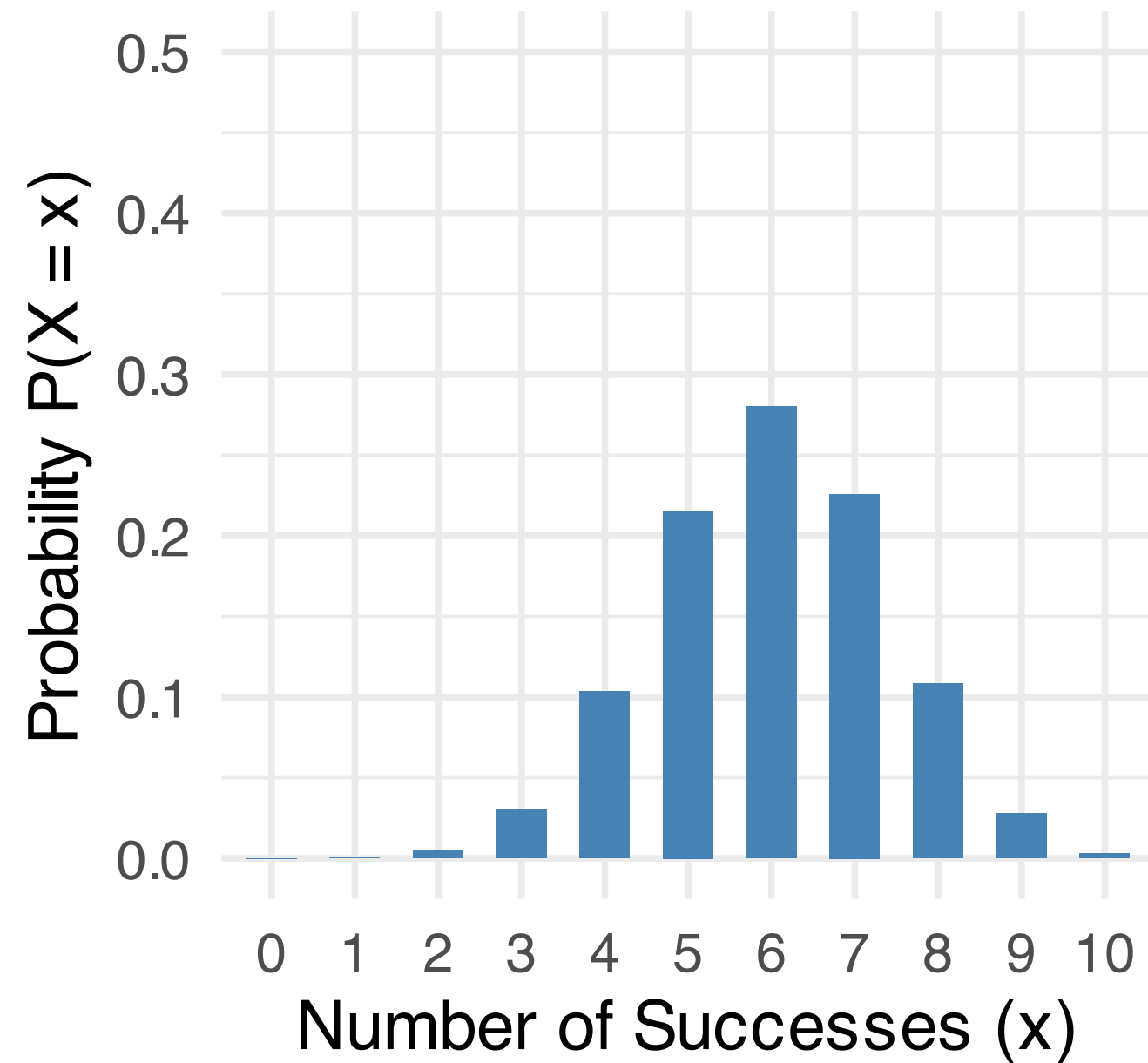
Hypergeometric PMF

$N = 50$, $K = 5$, $n = 10$



Hypergeometric PMF

$N = 50$, $K = 30$, $n = 10$



Hypergeometric

EXERCISE: LAPTOPS

A shipment of 50 laptops arrives at a warehouse, 15 of which are defective. A quality control inspector randomly selects 10 laptops to inspect. What is the probability that exactly 3 of the selected laptops are defective?

$$N = 50$$

$$K = 15$$

$$n = 10$$

$X = \#$ defective in sample

$$P(X=3) = p_X(3; N=50, K=15, n=10) = \frac{\binom{15}{3} \binom{50-15}{10-3}}{\binom{50}{10}} = 0.299$$

Other important (named) discrete distributions

- Categorical($\{p_1, \dots, p_K\}$) — Generalization of Bern(θ) to K outcomes instead of 2.
- Multinom($n, \{p_1, \dots, p_K\}$) — Generalization of Binom(n, θ) to K outcomes instead of 2.

For each of these (and many others), we don't need to “count” stuff. We just ask the math formula (or software) to calculate the probabilities of outcomes.

The trick is to translate the word problem to the correct named distribution.

Translating Word Problems

EXERCISE: TRANSLATIONS

For each of the following experiments, name the distribution that would make an appropriate model, and if obvious, identify the value(s) of any parameters.

- a. How many fries can you throw at the seagulls on Granville Island before they eat six? $\text{NegBin}(6, \theta)$
- b. Roll a die until you see 6. How many rolls? $\text{Geom}(1/6)$
- c. Given 1000 bank transactions, an auditor samples 50 to check for fraud. If all 50 are legitimate, what is the probability that the remainder contains fraud? **tricky!* $\text{Hyp Geom}(N=1000, n=50, k=?)$
- d. How many patients arrive at the ER in 3 hours? $\text{Poiss}(3\lambda)$ where $\lambda = \# \text{ patient per hour}$

Translating Word Problems

To do:

- Review calculus: ensure you can solve **these problems** (linked on website)
- Read **Chapter 2.4** [↗](#) before Tuesday's class
- Assignment 2 due May 27th, 11:59pm.