

Lecture 8

Discrete Distributions, Part 2

Grace Tompkins

Last modified — 05 Oct 2026

Learning Outcomes

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

- Apply the Poisson, geometric, negative binomial, and hypergeometric distributions to solve word problems.

1 Discrete Random Variables, continued

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 $\text{Pois}(\lambda)$ distribution.

Sometimes the PMF is written as:

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

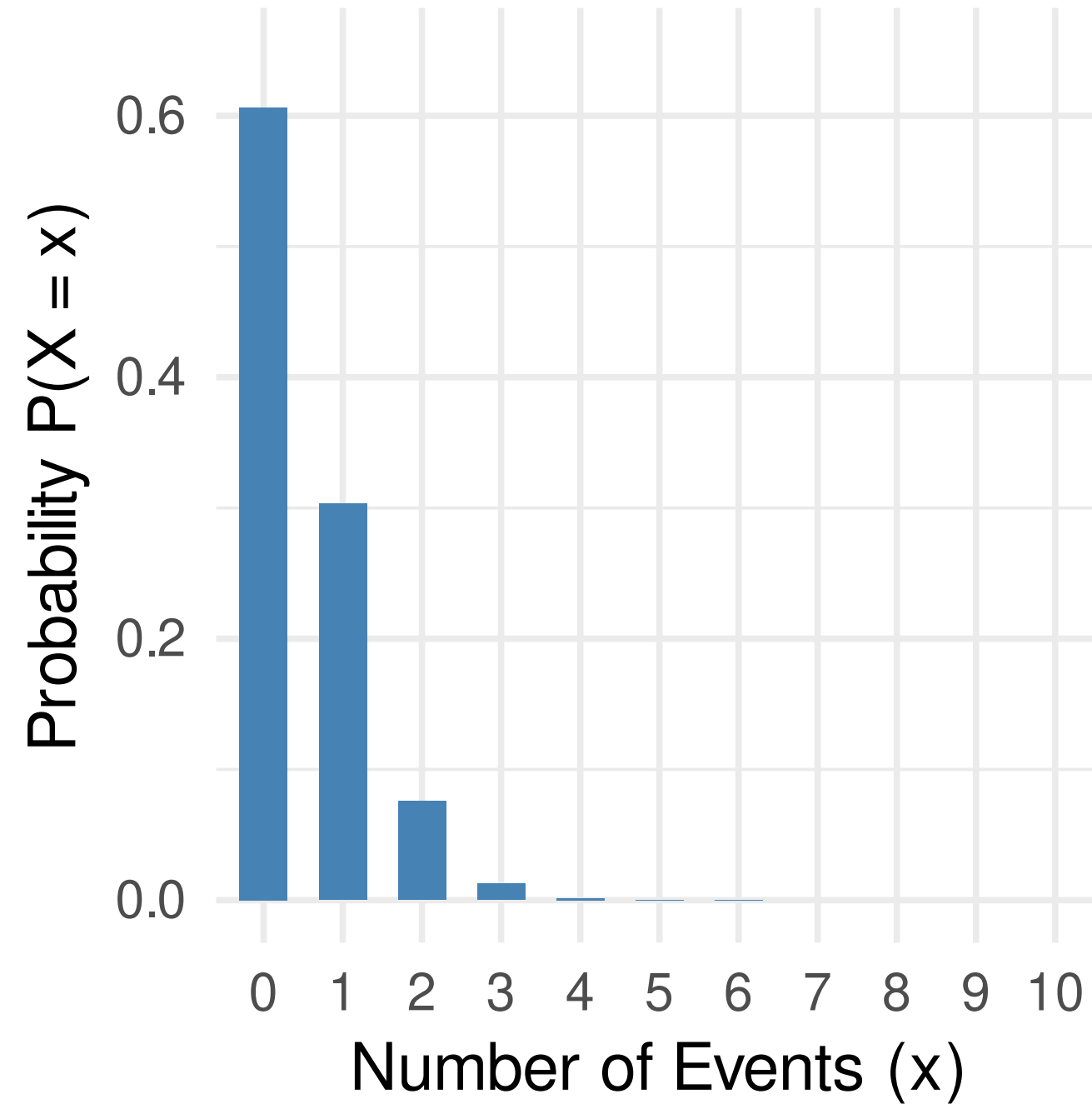
where λ is “per unit time” and t is the number of seconds/minutes/days/etc.

Poisson

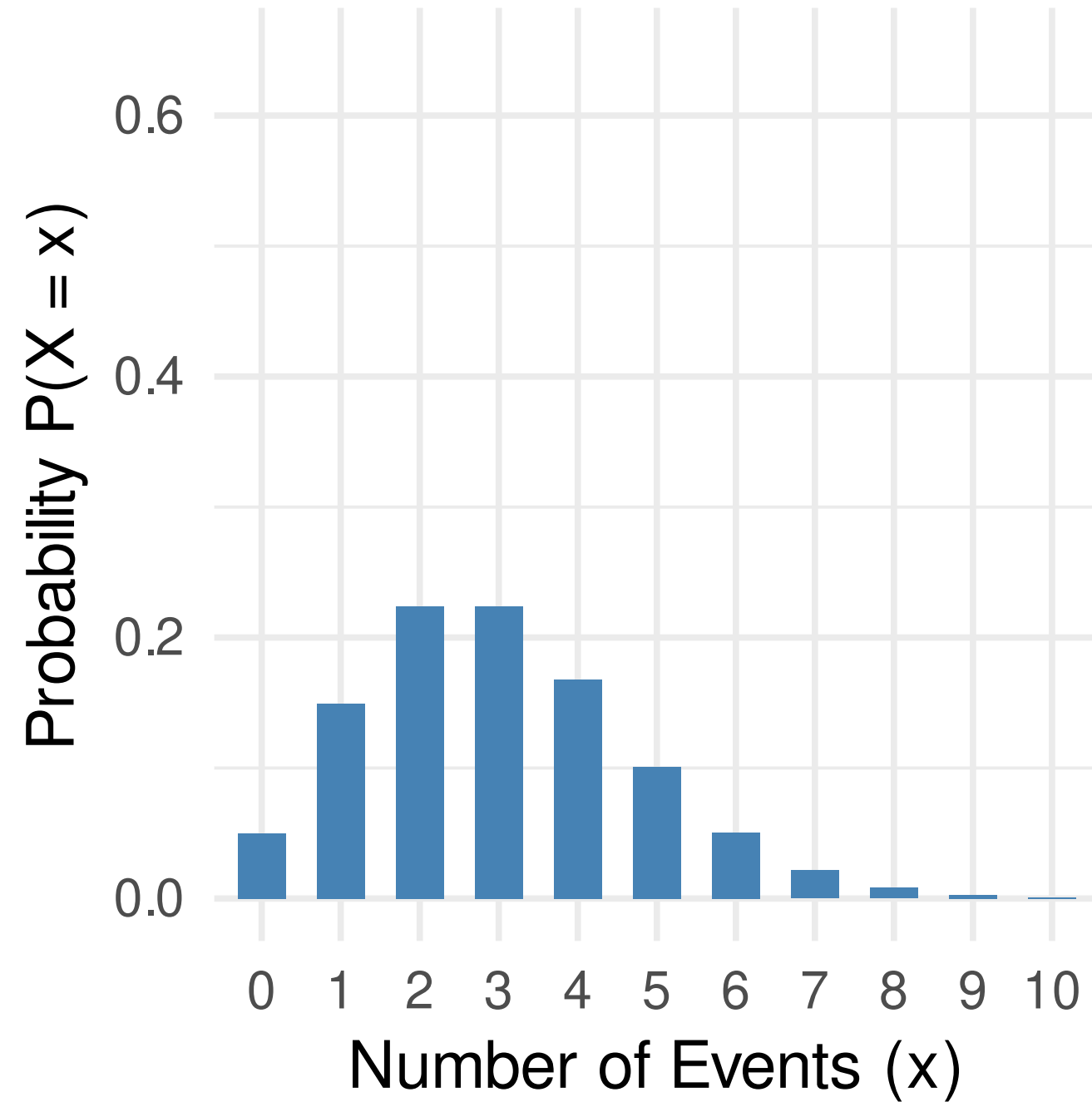
- The Poisson distribution is good for modelling counts of things, often per unit of time. This is not the only distribution that is good for modelling counts, and we'll see reasons to potentially choose a different one later.
- λ is often the average number of events per unit of time. X can be the count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

Poisson Maximization

EXERCISE: MAXIMIZING THE POISSON

Suppose that $Y \sim \text{Pois}(\lambda)$, 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!

Poisson Maximization

Geometric Distribution

DEFINITION

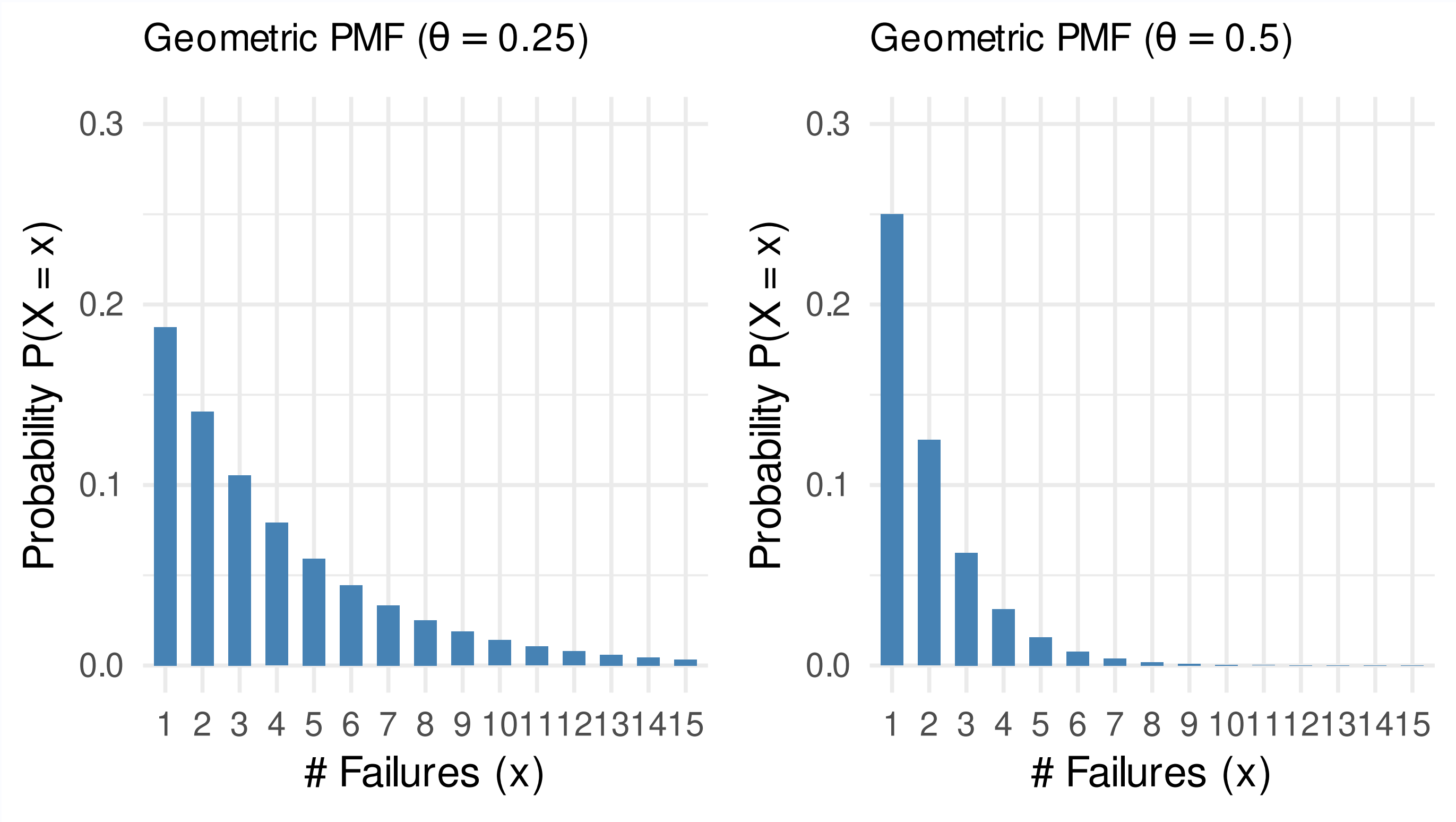
Suppose the probability of success is $\theta \in (0, 1)$. A RV X with PMF given by

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

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 (θ).

Geometric Distribution



Geometric Distribution

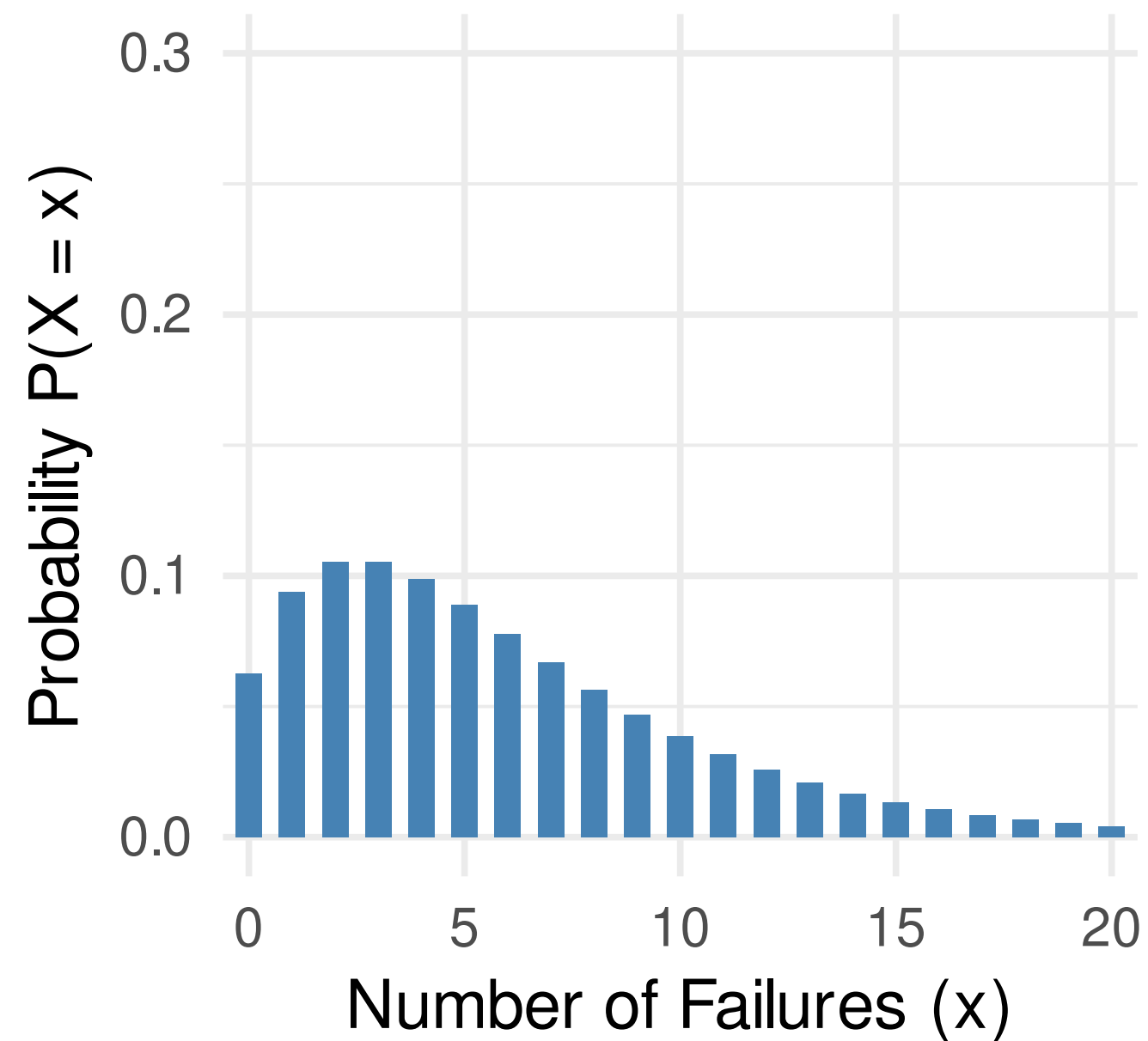
EXAMPLE

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

Negative Binomial

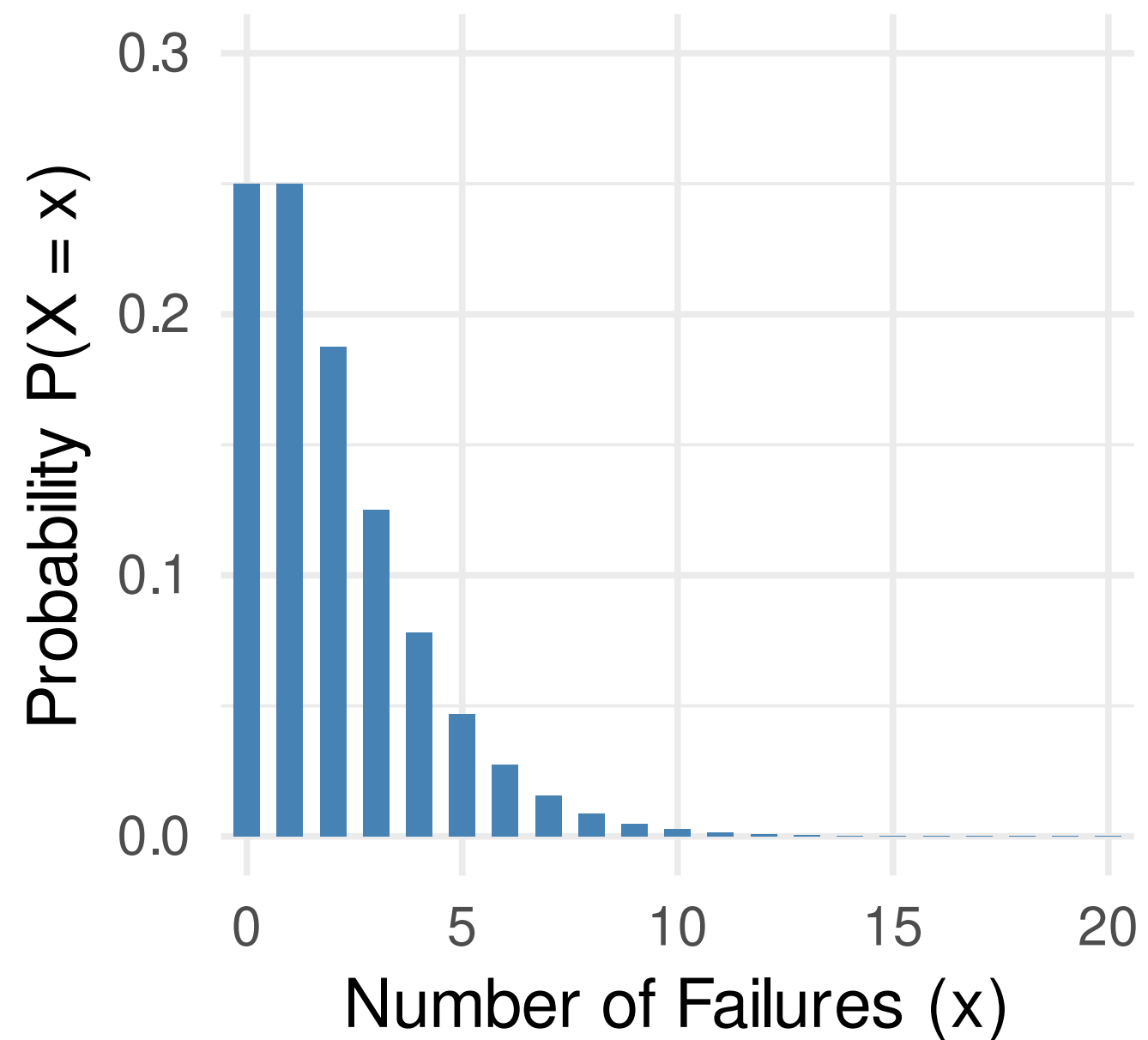
Negative Binomial PMF

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



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 $\text{NegBinom}(r, \theta)$ 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?

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

$$p_X(x; N, K, n) = \frac{\binom{K}{x} \binom{N-K}{n-x}}{\binom{N}{n}}, \text{ for } \max(0, n + M - N) \leq x \leq \min(n, M)$$

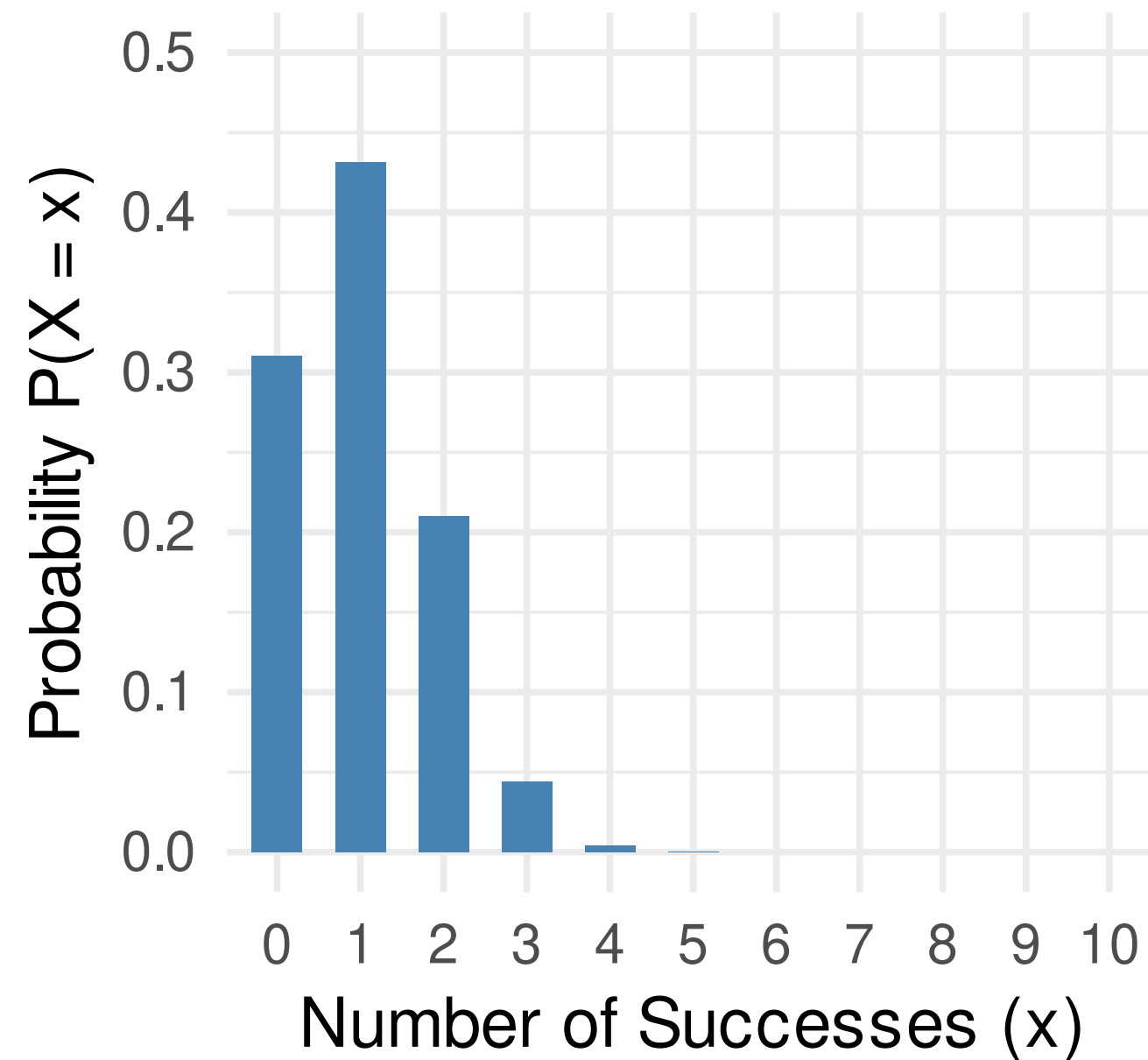
is said to have a hypergeometric distributions

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

Hypergeometric

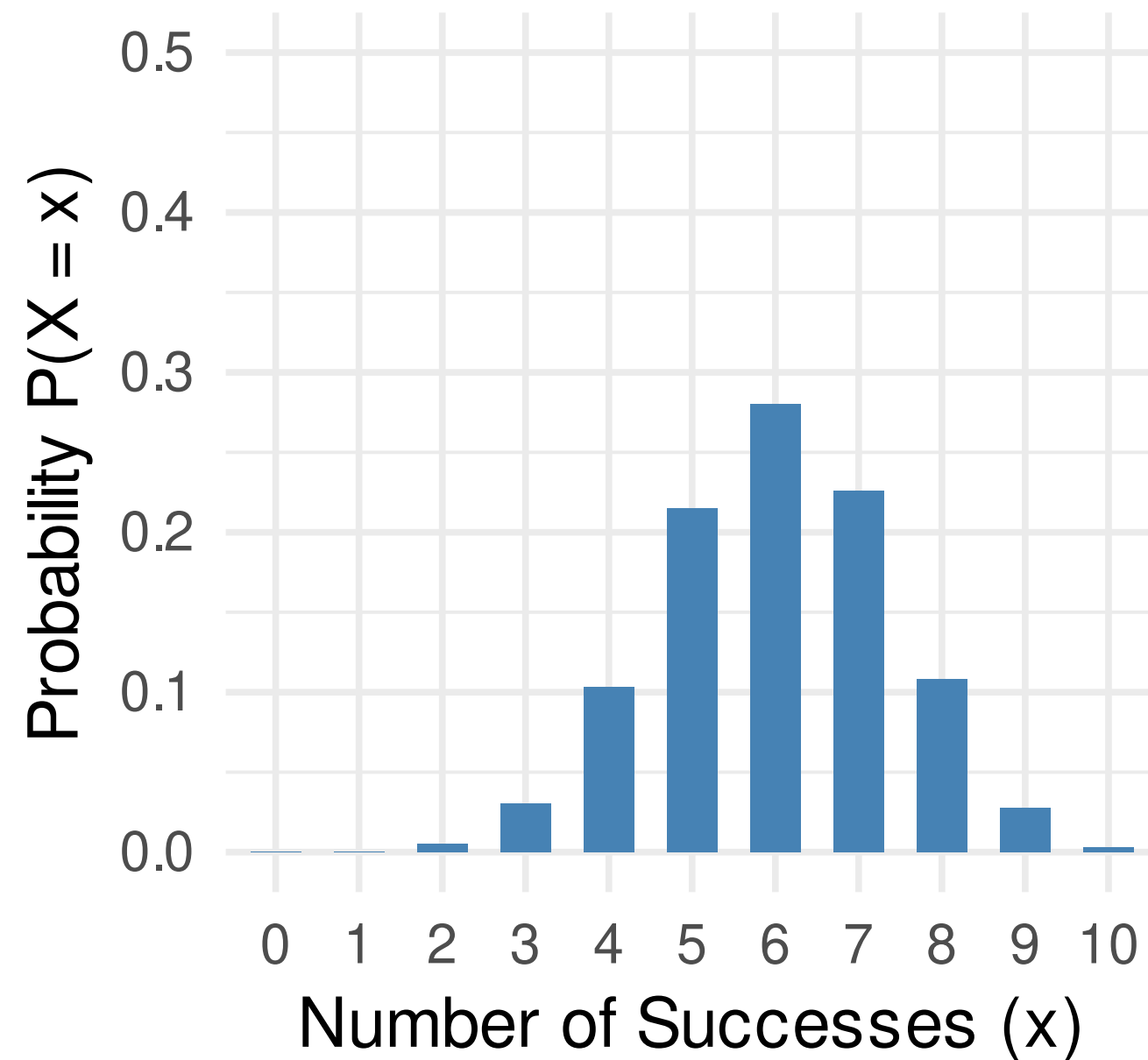
Hypergeometric PMF

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



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?

Other important (named) discrete distributions

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

Name the distribution that could be used to model the appropriate experiments and (if given) identify any values of the parameters

- a. How many fries can you throw at a seagull on Granville island before it eats six?
- b. Roll a die until you see a six. How many rolls did it take?
- c. Given 1000 bank transactions, an auditor samples 50 to check for fraud. If all 50 are legitimate, what is the probability that the remaining 950 contain fraud?
- d. How many patients will arrive in the ER in 3 hours?
- e. What is the probability of correctly guessing on 10 true/false questions?

Upcoming Assessments:

- Checkpoint 3 (Sections 2.1 - 2.3: Random Variables, Distributions, and Discrete Distributions) is on Friday Oct 09
- Midterm 1 (Sections 1.1 - 2.4) is on Friday Oct 16