0.1 Learning Objectives

By the end of this lecture, you will be able to:

- Model counts of successes using the Binomial distribution
- Model rare events using the Poisson distribution
- Model trials until r successes using the Negative Binomial distribution
- Choose the appropriate distribution for a given counting problem

0.2 Quick Recall

Recall: From last lecture, we introduced:

- **Bernoulli(p)**: Binary outcome (0 or 1), with $P(X = 1) = p$
- **Geometric(p)**: Trials until first success, $E(X) = 1/p$

0.2 Quick Recall

Recall: From last lecture, we introduced:

- **Bernoulli(p)**: Binary outcome (0 or 1), with $P(X = 1) = p$
- **Geometric(p)**: Trials until first success, $E(X) = 1/p$

Today: Three more distributions that build on these ideas:

- **Binomial**: Count successes in n trials
- **Poisson**: Count events in a fixed interval
- **Negative Binomial**: Count trials until r successes

1. The Binomial Distribution

1.1 Motivating Example

You’re a basketball player who makes 70% of your free throws.

1.1 Motivating Example

You’re a basketball player who makes 70% of your free throws.

In a game, you shoot 10 free throws.

1.1 Motivating Example

You're a basketball player who makes 70% of your free throws.

In a game, you shoot 10 free throws.

Questions:

- What's the probability you make exactly 7?
- What's your expected number of makes? 7 / 10
- What's a "typical" range of makes?

stdev

$\longleftrightarrow$ variance

1.1 Motivating Example

You’re a basketball player who makes 70% of your free throws.

In a game, you shoot 10 free throws.

Questions:

- What’s the probability you make exactly 7?
- What’s your expected number of makes?
- What’s a “typical” range of makes?

This is a **Binomial** scenario!

1.2 Building Intuition

Let’s start simpler: 3 free throws, 50% shooter.

1.2 Building Intuition

Let’s start simpler: 3 free throws, 50% shooter.

Let X = number of makes. What are the possible outcomes?

1.2 Building Intuition

Let's start simpler: 3 free throws, 50% shooter.

Let X = number of makes. What are the possible outcomes?

Outcome (✅ = make, ❌ = miss)	X	Probability
→ ✅✅✅	→3	$0.5^3 = 0.125$
→ ✅✅❌, ✅❌✅, ❌✅✅	→2	$3 \times 0.5^3 = 0.375$
→ ✅❌❌, ❌✅❌, ❌❌✅	→1	$3 \times 0.5^3 = 0.375$
→ ❌❌❌	→0	$0.5^3 = 0.125$

$P(\text{make } 1 \,\&\, 2 \,\&\, 3)$

$= \; ?$

$P(\text{make } 1) \cdot P(\text{make } 2) \cdot P(\text{make } 3)$

$P(A \cap B)$

$= P(A) \cdot P(B)$

?

"hot hand"

1.2 Building Intuition

Let’s start simpler: 3 free throws, 50% shooter.

Let X = number of makes. What are the possible outcomes?

Outcome (✅ = make, ❌ = miss)	X	Probability
✅✅✅	3	$0.5^3 = \boxed{0.125}$
✅✅❌, ✅❌✅, ❌✅✅	2	$\boxed{3 \times} 0.5^3 = 0.375$
✅❌❌, ❌✅❌, ❌❌✅	1	$\boxed{3 \times} 0.5^3 = 0.375$
❌❌❌	0	$0.5^3 = 0.125$

Notice: The coefficient counts the **number of ways** to arrange k successes in n trials.

1.3 The Counting Part

How many ways to get exactly 2 makes in 3 shots?

1.3 The Counting Part

How many ways to get exactly 2 makes in 3 shots?

We need to **choose** which 2 of the 3 shots are makes:

$$\binom{3}{2} = \frac{3!}{2! \cdot 1!} = \boxed{3}$$

1.3 The Counting Part

How many ways to get exactly 2 makes in 3 shots?

We need to **choose** which 2 of the 3 shots are makes:

$$\binom{3}{2} = \frac{3!}{2! \cdot 1!} = 3$$

In general, the number of ways to get k successes in n trials:

$$\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!}$$

1.3 The Counting Part

How many ways to get exactly 2 makes in 3 shots?

We need to **choose** which 2 of the 3 shots are makes:

$$\binom{3}{2} = \frac{3!}{2! \cdot 1!} = 3$$

In general, the number of ways to get k successes in n trials:

$$\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!}$$

Note: This is “n choose k” — the binomial coefficient we saw in counting!

1.4 The Binomial Distribution

Definition: Binomial Distribution

$X \sim \text{Binomial}(n, p)$ counts the number $\overset{k}{\underline{\text{of successes}}}$ in n independent trials, each with success probability p.

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

permutations

probability of successes

probability of failure

1.4 The Binomial Distribution

Definition: Binomial Distribution

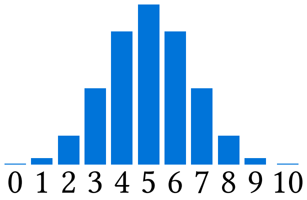
$X \sim \text{Binomial}(n, p)$ counts the number of successes in n independent trials, each with success probability p.

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

Components:

- $\binom{n}{k}$: ways to arrange k successes
- p^k: probability of k successes
- $(1 - p)^{n-k}$: probability of $(n - k)$ failures

$\text{Binomial}(10, 0.5)$



1.5 Binomial = Sum of Bernoullis

If $X_1, X_2, \ldots, X_n$ are independent Bernoulli(p) random variables, then:

$$\underline{X} = \underset{\text{1st shot}}{\underline{X_1}} + \underline{X_2} + \ldots + \underset{\text{nth shot}}{X_n} \sim \text{Binomial}(\underset{\text{\# trials}}{\boxed{n}}, \overset{\text{probability of success}}{\boxed{p}})$$

1.5 Binomial = Sum of Bernoullis

If $X_1, X_2, \ldots, X_n$ are independent Bernoulli(p) random variables, then:

$$X = X_1 + X_2 + \ldots + X_n \sim \text{Binomial}(n, p)$$

Each X_i is an **indicator** for success on trial i.

The sum counts the total number of successes!

1.6 Variance of Independent Sums

Recall: $E(X+Y) = E(X) + E(Y)$ always (linearity).

1.6 Variance of Independent Sums

Recall: $E(X+Y)=E(X)+E(Y)$ always (linearity).

What about variance? In general, $\mathrm{Var}(X+Y)\neq \mathrm{Var}(X)+\mathrm{Var}(Y)$.

1.6 Variance of Independent Sums

Recall: $E(X + Y) = E(X) + E(Y)$ always (linearity).

What about variance? In general, $\text{Var}(X + Y) \neq \text{Var}(X) + \text{Var}(Y)$.

But for **independent** random variables:

Definition: Variance of Independent Sums

If X and Y are **independent**, then:

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$$

1.6 Variance of Independent Sums

Recall: $E(X + Y) = E(X) + E(Y)$ always (linearity).

What about variance? In general, $\mathrm{Var}(X + Y) \neq \mathrm{Var}(X) + \mathrm{Var}(Y)$.

But for **independent** random variables:

Definition: Variance of Independent Sums

If X and Y are **independent**, then:

$$\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$$

This extends to n independent variables:

$$\mathrm{Var}(X_1 + \ldots + X_n) = \mathrm{Var}(X_1) + \ldots + \mathrm{Var}(X_n)$$

1.7 Binomial Mean and Variance

Using linearity of expectation and the Bernoulli connection:

1.7 Binomial Mean and Variance

Using linearity of expectation and the Bernoulli connection:

$$E(X) = E(X_1 + \ldots + X_n) = \underbrace{E(X_1) + \ldots + E(X_n)}_{n} = np$$

1.7 Binomial Mean and Variance

Using linearity of expectation and the Bernoulli connection:

$$E(X) = E(X_1 + \dots + X_n) = E(X_1) + \dots + E(X_n) = np$$

Since the X_i are **independent**:

$$\mathrm{Var}(X) = \underbrace{\mathrm{Var}(X_1) + \dots + \mathrm{Var}(X_n)}_{n} = \underbrace{n}_{}\underbrace{p(1-p)}_{}$$

1.7 Binomial Mean and Variance

Using linearity of expectation and the Bernoulli connection:

$$E(X) = E(X_1 + \ldots + X_n) = E(X_1) + \ldots + E(X_n) = np$$

Since the X_i are **independent**:

$$\mathrm{Var}(X) = \mathrm{Var}(X_1) + \ldots + \mathrm{Var}(X_n) = np(1 - p)$$

Mean:

$$E(X) = np$$

Variance:

$$\mathrm{Var}(X) = np(1 - p)$$

1.8 Back to Free Throws

You're a 70% free throw shooter taking 10 shots.

$$X \sim \text{Binomial}(\underline{10}, \underline{0.7})$$

1.8 Back to Free Throws

You're a 70% free throw shooter taking 10 shots.

$X \sim \text{Binomial}(10, 0.7)$

Expected makes: $E(X) = \underline{10} \times \underline{0.7} = \textcircled{7}$ ✓

1.8 Back to Free Throws

You're a 70% free throw shooter taking 10 shots.

$$X \sim \text{Binomial}(10, 0.7)$$

Expected makes: $E(X) = 10 \times 0.7 = 7$

Standard deviation: $\sigma = \sqrt{\underline{10} \times \underline{0.7} \times \underline{0.3}} = \sqrt{2.1} \approx 1.45$

1.8 Back to Free Throws

You're a 70% free throw shooter taking 10 shots.

$X \sim \text{Binomial}(10, 0.7)$

Expected makes: $E(X) = 10 \times 0.7 = 7$

Standard deviation: $\sigma = \sqrt{10 \times 0.7 \times 0.3} = \sqrt{2.1} \approx 1.45$

Probability of exactly 7 makes:

PMF

$$P(X = \underset{\uparrow}{7}) = \binom{10}{7}\underline{(0.7)}^7(0.3)^3 = 120 \times 0.082 \times 0.027 \approx 0.267$$

k # successes

1.9 Your Turn: Binomial

Try it yourself

Talk to your neighbor and try to solve this problem.

A multiple-choice test has 5 questions, each with 4 choices. You guess randomly on all questions.

X : Score on exam

$X \sim ?$

$\mathbb{E}(X) ?$

$P(X=k) ?$

P

$X \sim \text{Binom}(5, \frac{1}{4})$

$\mathbb{E}(X) = 5 \cdot \frac{1}{4} = \frac{5}{4}$

1.9 Your Turn: Binomial

Try it yourself

Talk to your neighbor and try to solve this problem.

A multiple-choice test has 5 questions, each with 4 choices. You guess randomly on all questions.

What is the probability of getting **at least 2 correct**?

$$P(X \geq 2) = P(x=2) + P(x=3) + P(x=4) + P(x=5)$$

$$= 1 - P(X<2)$$

$$= 1 - P(X=0) - P(X=1) \longrightarrow \binom{5}{1} .25^1 \, .75^4$$

$$\hookrightarrow \binom{5}{0} .25^0 (0.75)^5$$

1.9 Your Turn: Binomial

Try it yourself

Talk to your neighbor and try to solve this problem.

A multiple-choice test has 5 questions, each with 4 choices. You guess randomly on all questions.

What is the probability of getting **at least 2 correct**?

Let $X \sim \mathrm{Binomial}(5, 1/4)$

1.9 Your Turn: Binomial

Try it yourself

Talk to your neighbor and try to solve this problem.

A multiple-choice test has 5 questions, each with 4 choices. You guess randomly on all questions.

What is the probability of getting **at least 2 correct**?

Let $X \sim \text{Binomial}(5, 1/4)$

$$P(X \geq 2) = 1 - P(X = 0) - P(X = 1)$$

1.9 Your Turn: Binomial

Try it yourself

Talk to your neighbor and try to solve this problem.

A multiple-choice test has 5 questions, each with 4 choices. You guess randomly on all questions.

What is the probability of getting **at least 2 correct**?

Let $X \sim \text{Binomial}(5, 1/4)$

$$P(X \geq 2) = 1 - P(X = 0) - P(X = 1)$$

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

1.9 Your Turn: Binomial

Try it yourself

Talk to your neighbor and try to solve this problem.

A multiple-choice test has 5 questions, each with 4 choices. You guess randomly on all questions.

What is the probability of getting **at least 2 correct**?

Let $X \sim \text{Binomial}(5, 1/4)$

$$P(X \geq 2) = 1 - P(X = 0) - P(X = 1)$$

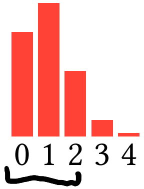
$$= 1 - \binom{5}{0}\left(\frac{1}{4}\right)^0\left(\frac{3}{4}\right)^5 - \binom{5}{1}\left(\frac{1}{4}\right)^1\left(\frac{3}{4}\right)^4$$

$$= 1 - 0.237 - 0.396 = \boxed{0.367}$$

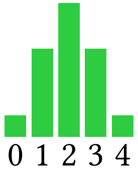
1.10 Binomial Shape Changes

How does the shape depend on p? — success likelihood

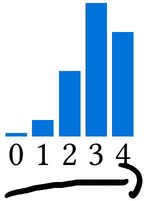
$p = 0.2$ (right -skewed)



$p = 0.5$ (symmetric)



$p = 0.8$ (left -skewed)

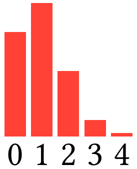


success

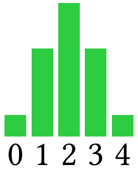
1.10 Binomial Shape Changes

How does the shape depend on p?

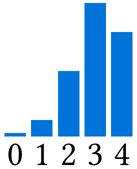
$p = 0.2$ (left-skewed)



$p = 0.5$ (symmetric)



$p = 0.8$ (right-skewed)

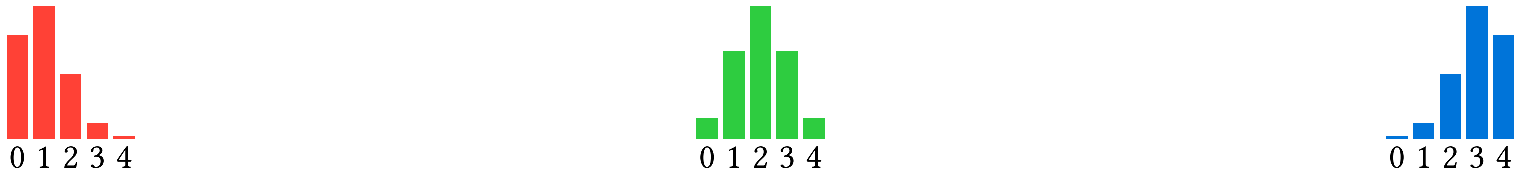


All three are $\text{Binomial}(4, p)$ — same n, different p.

1.10 Binomial Shape Changes

How does the shape depend on p?

$p = 0.2$ (left-skewed) $p = 0.5$ (symmetric) $p = 0.8$ (right-skewed)



All three are $\text{Binomial}(4, p)$ — same n, different p.

1.11 The Quantile Function

We’ve seen two key functions for working with distributions:

$$\begin{tikzcd} \text{PMF} \arrow[r, bend left] & \text{CDF} \arrow[l, bend left] \end{tikzcd}$$

1.11 The Quantile Function

We've seen two key functions for working with distributions:

- **PMF**: Given k, find $P(X = k)$
- **CDF**: Given k, find $P(X \leq k)$

1.11 The Quantile Function

We've seen two key functions for working with distributions:

- **PMF**: Given k, find $P(X = k)$
- **CDF**: Given k, find $P(X \leq k)$

The **quantile function** goes the other direction:

Definition: Quantile Function

Given a probability q, find the smallest value k such that $P(X \leq k) \geq q$.

This is the "inverse CDF."

1.11 The Quantile Function

We've seen two key functions for working with distributions:

- **PMF**: Given k, find $P(X = k)$
- **CDF**: Given k, find $P(X \leq k)$

The **quantile function** goes the other direction:

Definition: Quantile Function

Given a probability q, find the smallest value k such that $P(X \leq k) \geq q$.

This is the "inverse CDF."

Example: For $X \sim \text{Binomial}(\underline{10}, \underline{0.7})$, the median is the value k where $P(X \leq k) \geq 0.5$.

1.12 Computing Binomial in Code

Python:

```python
from scipy.stats import binom

# P(X = 7) for Binomial(10, 0.7)
binom.pmf(7, n=10, p=0.7)

# P(X <= 7)
binom.cdf(7, n=10, p=0.7)

# Smallest k where P(X <= k) >= 0.5
binom.ppf(0.5, n=10, p=0.7)
```

R:

PMF

density

```r
# P(X = 7) for Binomial(10, 0.7)
dbinom(7, size=10, prob=0.7)
```

CDF

```r
# P(X <= 7)
pbinom(7, size=10, prob=0.7)

# Smallest k where P(X <= k) >= 0.5
qbinom(0.5, size=10, prob=0.7)
```

2. The Poisson Distribution

2.1 A Different Kind of Counting

Binomial counts successes in a **fixed number of trials**.

2.1 A Different Kind of Counting

Binomial counts successes in a **fixed number of trials**.

But what if we're counting events over a **fixed time interval**?

2.1 A Different Kind of Counting

Binomial counts successes in a **fixed number of trials**.

But what if we're counting events over a **fixed time interval**?

Example: Counting Rare Events

- Number of customers arriving at a store per hour
- Number of typos on a page
- Number of car accidents at an intersection per day
- Number of emails received per minute

2.1 A Different Kind of Counting

Binomial counts successes in a **fixed number of trials**.

But what if we're counting events over a **fixed time interval**?

Example: Counting Rare Events

- Number of customers arriving at a store per hour
- Number of typos on a page
- Number of car accidents at an intersection per day
- Number of emails received per minute

These don't have a fixed "number of trials"!

2.2 The Poisson Distribution

Definition: Poisson Distribution

$X \sim \text{Poisson}(\lambda)$ counts the number of events in a fixed interval, where $\lambda > 0$ is the average rate of events.

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}, \quad k = 0, 1, 2, \ldots$$

2.2 The Poisson Distribution

Definition: Poisson Distribution

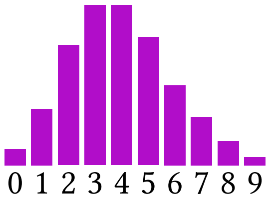
$X \sim \text{Poisson}(\lambda)$ counts the number of events in a fixed interval, where $\lambda > 0$ is the average rate of events.

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}, \quad k = 0, 1, 2, \ldots$$

Key features:

- λ (lambda) = expected number of events
- k can be any non-negative integer
- No upper bound (unlike Binomial)!

$\text{Poisson}(4)$



2.3 Poisson Mean and Variance

A remarkable property of the Poisson:

2.3 Poisson Mean and Variance

A remarkable property of the Poisson:

$$E(X) = \lambda$$

2.3 Poisson Mean and Variance

A remarkable property of the Poisson:

$$E(X) = \boxed{\lambda}$$

$$\mathrm{Var}(X) = \boxed{\lambda}$$

2.3 Poisson Mean and Variance

A remarkable property of the Poisson:

$$E(X) = \lambda$$

$$\mathrm{Var}(X) = \lambda$$

Note: The mean equals the variance! This is a signature property of Poisson.

2.3 Poisson Mean and Variance

A remarkable property of the Poisson:

$$E(X) = \lambda$$

$$\mathrm{Var}(X) = \lambda$$

Note: The mean equals the variance! This is a signature property of Poisson.

If you observe data where mean $\approx$ variance, Poisson might be a good model.

2.4 Example: Server Requests

A web server receives an average of 3 requests per second.

λ

$X \sim \text{Poisson}(3)$

2.4 Example: Server Requests

A web server receives an average of 3 requests per second.

$X \sim \text{Poisson}(3)$

What's the probability of receiving exactly 5 requests in a given second?

2.4 Example: Server Requests

A web server receives an average of 3 requests per second.

$X \sim \text{Poisson}(3)$

What's the probability of receiving exactly 5 requests in a given second?

$$P(X = \textcircled{5}) = \frac{3^5 e^{-3}}{5!} = \frac{243 \times 0.0498}{120} \approx \underline{0.101}$$

2.4 Example: Server Requests

A web server receives an average of 3 requests per second.

$X \sim \text{Poisson}(3)$

What's the probability of receiving exactly 5 requests in a given second?

$$P(X = 5) = \frac{3^5 e^{-3}}{5!} = \frac{243 \times 0.0498}{120} \approx 0.101$$

What's the probability of receiving no requests?

2.4 Example: Server Requests

A web server receives an average of 3 requests per second.

$X \sim \text{Poisson}(3)$

What's the probability of receiving exactly 5 requests in a given second?

$$P(X = 5) = \frac{3^5 e^{-3}}{5!} = \frac{243 \times 0.0498}{120} \approx 0.101$$

What's the probability of receiving no requests?

$$P(X = 0) = \frac{3^0 e^{-3}}{0!} = e^{-3} \approx 0.050$$

2.5 Your Turn: Poisson

Try it yourself

Talk to your neighbor and try to solve this problem.

A call center receives an average of 2 calls per minute.

1. What’s the probability of exactly 4 calls in the next minute?

2.5 Your Turn: Poisson

Try it yourself

Talk to your neighbor and try to solve this problem.

$\lambda = 2$

A call center receives an average of 2 calls per hour

1. What's the probability of exactly 4 calls in the next hour ?

$$P(X = 4) = \frac{2^4 e^{-2}}{4!} = \frac{16 \times 0.135}{24} \approx 0.\underline{090}$$

2.5 Your Turn: Poisson

Try it yourself

Talk to your neighbor and try to solve this problem.

A call center receives an average of 2 calls per minute.

1. What's the probability of exactly 4 calls in the next minute?

$$P(X = 4) = \frac{2^4 e^{-2}}{4!} = \frac{16 \times 0.135}{24} \approx 0.090$$

2. What's the probability of **no** calls in the next minute?

2.5 Your Turn: Poisson

Try it yourself

Talk to your neighbor and try to solve this problem.

A call center receives an average of 2 calls per minute.

1. What's the probability of exactly 4 calls in the next minute?

$$P(X=4)=\frac{2^4e^{-2}}{4!}=\frac{16\times 0.135}{24}\approx 0.090$$

2. What's the probability of **no** calls in the next minute?

$$P(X=0)=e^{-2}\approx 0.\underline{13}5$$

2.6 Computing Poisson in Code

Python:

```
from scipy.stats import poisson

# P(X = 5) for Poisson(3)
poisson.pmf(5, mu=3)

# P(X <= 5)
poisson.cdf(5, mu=3)

# Smallest k where P(X <= k) >=
0.9
poisson.ppf(0.9, mu=3)
```

R:

```
# P(X = 5) for Poisson(3)
dpois(5, lambda=3)

# P(X <= 5)
ppois(5, lambda=3)

# Smallest k where P(X <= k) >=
0.9
qpois(0.9, lambda=3)
```

3. The Negative Binomial Distribution

3.1 From Geometric to Negative Binomial

Recall: Geometric(p) counts trials until the **first** success.

3.1 From Geometric to Negative Binomial

Recall: Geometric(p) counts trials until the **first** success.

What if we want trials until the **r-th** success?

3.1 From Geometric to Negative Binomial

Recall: Geometric(p) counts trials until the **first** success.

What if we want trials until the **r-th** success?

Example: Backup Batteries

A device needs 3 working batteries. Each battery works with probability 0.9.

How many batteries do you need to test until you find 3 working ones?

3.2 The Negative Binomial Distribution

Definition: Negative Binomial Distribution

$X \sim \mathrm{NegBin}(r, p)$ counts the number of trials until the r-th success, where each trial succeeds independently with probability p.

$$P(X = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}, \quad k = r, r+1, r+2, \ldots$$

3.2 The Negative Binomial Distribution

Definition: Negative Binomial Distribution

$X \sim \text{NegBin}(r, p)$ counts the number of trials until the r-th success, where each trial succeeds independently with probability p.

$$P(X = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}, \quad k = r, r+1, r+2, \ldots$$

Intuition: To get the r-th success on trial k:

- Need exactly $r-1$ successes in the first $k-1$ trials: $\binom{k-1}{r-1}$
- Then a success on trial k: p
- With $k-r$ failures total: $(1-p)^{k-r}$

3.3 Negative Binomial = Sum of Geometrics

Just like Binomial is a sum of Bernoullis…

3.3 Negative Binomial = Sum of Geometrics

Just like Binomial is a sum of Bernoullis…

$$\boxed{X} = G_1 + G_2 + \dots + G_{\boxed{r}}$$

where $G_i \sim \text{Geometric}(p)$ is the number of trials between success $i - 1$ and success i.

3.3 Negative Binomial = Sum of Geometrics

Just like Binomial is a sum of Bernoullis…

$$X = G_1 + G_2 + \ldots + G_r$$

where $G_i \sim \text{Geometric}(p)$ is the number of trials between success $i - 1$ and success i.

Using linearity and independence:

Mean:

$$E(X) = \frac{r}{p}$$

Variance:

$$\text{Var}(X) = \frac{r(1-p)}{p^2}$$

3.3 Negative Binomial = Sum of Geometrics

Just like Binomial is a sum of Bernoullis…

$$X = G_1 + G_2 + \ldots + G_r$$

where $G_i \sim \text{Geometric}(p)$ is the number of trials between success $i - 1$ and success i.

Using linearity and independence:

Mean:

$$E(X) = \frac{r}{p}$$

Variance:

$$\text{Var}(X) = \frac{r(1 - p)}{p^2}$$

Note: When $r = 1$, this reduces to the Geometric distribution!

3.4 Example: Backup Batteries

You need 3 working batteries. Each works with probability $p = 0.9$.

$$X \sim \text{NegBin}(3, 0.9)$$

3.4 Example: Backup Batteries

You need 3 working batteries. Each works with probability $p = 0.9$.

$X \sim \text{NegBin}(3, 0.9)$

Expected number of batteries to test:

$$E(X) = \frac{3}{0.9} = 3.33$$

3.4 Example: Backup Batteries

You need 3 working batteries. Each works with probability $p = 0.9$.

$X \sim \mathrm{NegBin}(3, 0.9)$

Expected number of batteries to test:

$$E(X) = \frac{3}{0.9} = 3.33$$

Probability you need exactly 4 batteries (i.e., one dud):

$$P(X = 4) = \binom{3}{2}(0.9)^3(0.1)^1 = 3 \times 0.729 \times 0.1 = \underline{0.219}$$

3.5 Distribution Functions: Quick Reference

Function	Description	Python (scipy.stats)	R
PMF	$P(X = k)$	`binom.pmf(k, n, p)`	`dbinom(k, n, p)`
		`poisson.pmf(k, mu)`	`dpois(k, lambda)`
		`nbinom.pmf(k, r, p)`	`dnbinom(k, r, p)`
CDF	$P(X \leq k)$	`binom.cdf(k, n, p)`	`pbinom(k, n, p)`
		`poisson.cdf(k, mu)`	`ppois(k, lambda)`
		`nbinom.cdf(k, r, p)`	`pnbinom(k, r, p)`
Quantile	inverse CDF	`binom.ppf(q, n, p)`	`qbinom(q, n, p)`
		`poisson.ppf(q, mu)`	`qpois(q, lambda)`
		`nbinom.ppf(q, r, p)`	`qnbinom(q, r, p)`
Random	samples	`binom.rvs(n, p, size=m)`	`rbinom(m, n, p)`
		`poisson.rvs(mu, size=m)`	`rpois(m, lambda)`
		`nbinom.rvs(r, p, size=m)`	`rnbinom(m, r, p)`

"DPQR"

$X \sim \text{NegBin}(3, 0.9)$

$[3, 5, 4, 3, 5, 7, 3, 4, 3]$

3.5 Distribution Functions: Quick Reference

Function	Description	Python (scipy.stats)	R
PMF	$P(X = k)$	`binom.pmf(k, n, p)`	`dbinom(k, n, p)`
		`poisson.pmf(k, mu)`	`dpois(k, lambda)`
		`nbinom.pmf(k, r, p)`	`dnbinom(k, r, p)`
CDF	$P(X \leq k)$	`binom.cdf(k, n, p)`	`pbinom(k, n, p)`
		`poisson.cdf(k, mu)`	`ppois(k, lambda)`
		`nbinom.cdf(k, r, p)`	`pnbinom(k, r, p)`
Quantile	inverse CDF	`binom.ppf(q, n, p)`	`qbinom(q, n, p)`
		`poisson.ppf(q, mu)`	`qpois(q, lambda)`
		`nbinom.ppf(q, r, p)`	`qnbinom(q, r, p)`
Random	samples	`binom.rvs(n, p, size=m)`	`rbinom(m, n, p)`
		`poisson.rvs(mu, size=m)`	`rpois(m, lambda)`
		`nbinom.rvs(r, p, size=m)`	`rnbinom(m, r, p)`

Note: R pattern: **d** (density/PMF), **p** (CDF), **q** (quantile), **r** (random)

4. Comparing Distributions

4.1 When to Use Which Distribution?

Binomial(n, p)

- Fixed number of trials n
- Each trial: success/failure
- Counting: total successes
- Examples:
  - Coin flips
  - Defective items in batch
  - Multiple choice guessing

Poisson(λ)

- Events over a fixed time/space interval
- No fixed "number of trials"
- Counting: events per unit time/space
- Examples:
  - Arrivals per hour
  - Typos per page
  - Accidents per month

4.2 When to Use Which Distribution? (cont'd)

Geometric(p)

- Trials until **first** success
- Counting: number of trials
- Examples:
  - Rolls until first 6
  - Attempts until first win
  - Calls until first answer

NegBin(r, p)

- Trials until **r-th** success
- Counting: number of trials
- Examples:
  - Interviews until 3 hires
  - Games until 4 wins
  - Tests until r working parts

4.3 Distribution Summary

Distribution	**PMF**	**Mean**	**Variance**	**Use When**
Bernoulli(p)	$P(X = 1) = p$	p	$p(1 - p)$	Single yes/no trial
Binomial(n, p)	$\binom{n}{k}p^{k(1-p)^{n-k}}$	np	$np(1 - p)$	Successes in n trials
Geometric(p)	$(1 - p)^{k-1}p$	$1/p$	$(1 - p)/p^2$	Trials until 1st success
NegBin(r, p)	$\binom{k-1}{r-1}p^{r(1-p)^{k-r}}$	r/p	$r(1 - p)/p^2$	Trials until rth success
Poisson(λ)	$\left(\lambda^k e^{-\lambda}\right)/k!$	λ	λ	Rare events, arrivals

CDF PPF

4.4 Practice Problem: Which Distribution?

Scenario 1: A website has 1000 visitors/day. Each has 0.2% chance of making a purchase. Model the number of purchases.

X : #purchases

$\mathbb{E}(X)$?

$P(X=50)$?

$X \sim$

4.4 Practice Problem: Which Distribution?

4.4 Practice Problem: Which Distribution?

Scenario 1: A website has 1000 visitors/day. Each has 0.2% chance of making a purchase. Model the number of purchases.

$$\boxed{\mathrm{Binomial}(1000, 0.002) \text{ or } \mathrm{Poisson}(2) \text{ (}\underline{\text{large } n}\text{, } \underline{\text{small } p}\text{)}}$$

4.4 Practice Problem: Which Distribution?

4.4 Practice Problem: Which Distribution?

Scenario 1: A website has 1000 visitors/day. Each has 0.2% chance of making a purchase. Model the number of purchases.

$\boxed{\text{Binomial}(1000, 0.002) \text{ or Poisson}(2) \text{ (large } n\text{, small } p\text{)}}$

Scenario 2: You roll a die until you get a 6.

4.4 Practice Problem: Which Distribution?

4.4 Practice Problem: Which Distribution?

Scenario 1: A website has 1000 visitors/day. Each has 0.2% chance of making a purchase. Model the number of purchases.

$\boxed{\text{Binomial}(1000, 0.002) \text{ or Poisson}(2) \text{ (large } n\text{, small } p\text{)}}$

Scenario 2: You roll a die until you get a 6.

$\boxed{\text{Geometric}(1/6)}$

4.4 Practice Problem: Which Distribution?

4.4 Practice Problem: Which Distribution?

Scenario 1: A website has 1000 visitors/day. Each has 0.2% chance of making a purchase. Model the number of purchases.

$\boxed{\text{Binomial}(1000, 0.002) \text{ or Poisson}(2) \text{ (large } n\text{, small } p\text{)}}$

Scenario 2: You roll a die until you get a 6.

$\boxed{\text{Geometric}(1/6)}$

Scenario 3: A radioactive source emits particles at an average rate of 5 per minute.

4.4 Practice Problem: Which Distribution?

4.4 Practice Problem: Which Distribution?

Scenario 1: A website has 1000 visitors/day. Each has 0.2% chance of making a purchase. Model the number of purchases.

$$\boxed{\text{Binomial}(1000, 0.002) \text{ or Poisson}(2) \text{ (large } n\text{, small } p\text{)}}$$

Scenario 2: You roll a die until you get a 6.

$$\boxed{\text{Geometric}(1/6)}$$

Scenario 3: A radioactive source emits particles at an average rate of 5 per minute.

$$\boxed{\text{Poisson}(5)}$$

4.4 Practice Problem: Which Distribution?

4.4 Practice Problem: Which Distribution?

Scenario 1: A website has 1000 visitors/day. Each has 0.2% chance of making a purchase. Model the number of purchases.

$$\boxed{\text{Binomial}(1000, 0.002) \text{ or Poisson}(2) \text{ (large } n\text{, small } p\text{)}}$$

Scenario 2: You roll a die until you get a 6.

$$\boxed{\text{Geometric}(1/6)}$$

Scenario 3: A radioactive source emits particles at an average rate of 5 per minute.

$$\boxed{\text{Poisson}(5)}$$

Scenario 4: You need to interview candidates until you find 3 qualified ones. Each candidate is qualified with probability 0.2.

4.4 Practice Problem: Which Distribution?

4.4 Practice Problem: Which Distribution?

Scenario 1: A website has 1000 visitors/day. Each has 0.2% chance of making a purchase. Model the number of purchases.

$$\boxed{\text{Binomial}(1000, 0.002) \text{ or Poisson}(2) \text{ (large } n\text{, small } p\text{)}}$$

Scenario 2: You roll a die until you get a 6.

$$\boxed{\text{Geometric}(1/6)}$$

Scenario 3: A radioactive source emits particles at an average rate of 5 per minute.

$$\boxed{\text{Poisson}(5)}$$

Scenario 4: You need to interview candidates until you find 3 qualified ones. Each candidate is qualified with probability 0.2.

4.4 Practice Problem: Which Distribution?

$$\boxed{\text{NegBin}(\underline{3}, \underline{0.2}), \text{ expected interviews: } 3/0.2 = \underline{15}}$$

4.5 Recap

Today we covered:

- Binomial(n, p): successes in n trials; NegBin(r, p): trials until r successes
- Binomial = sum of Bernoullis; NegBin = sum of Geometrics
- Poisson(λ): rare events; $E(X) = \mathrm{Var}(X) = \lambda$
- Poisson approximates Binomial when n large, p small
- Choose distribution based on: what are you counting?