VANDERBILT
School *of* Engineering

CE 3300-01 – RISK, RELIABILITY, AND RESILIENCE ENGINEERING

Discrete Random Variables

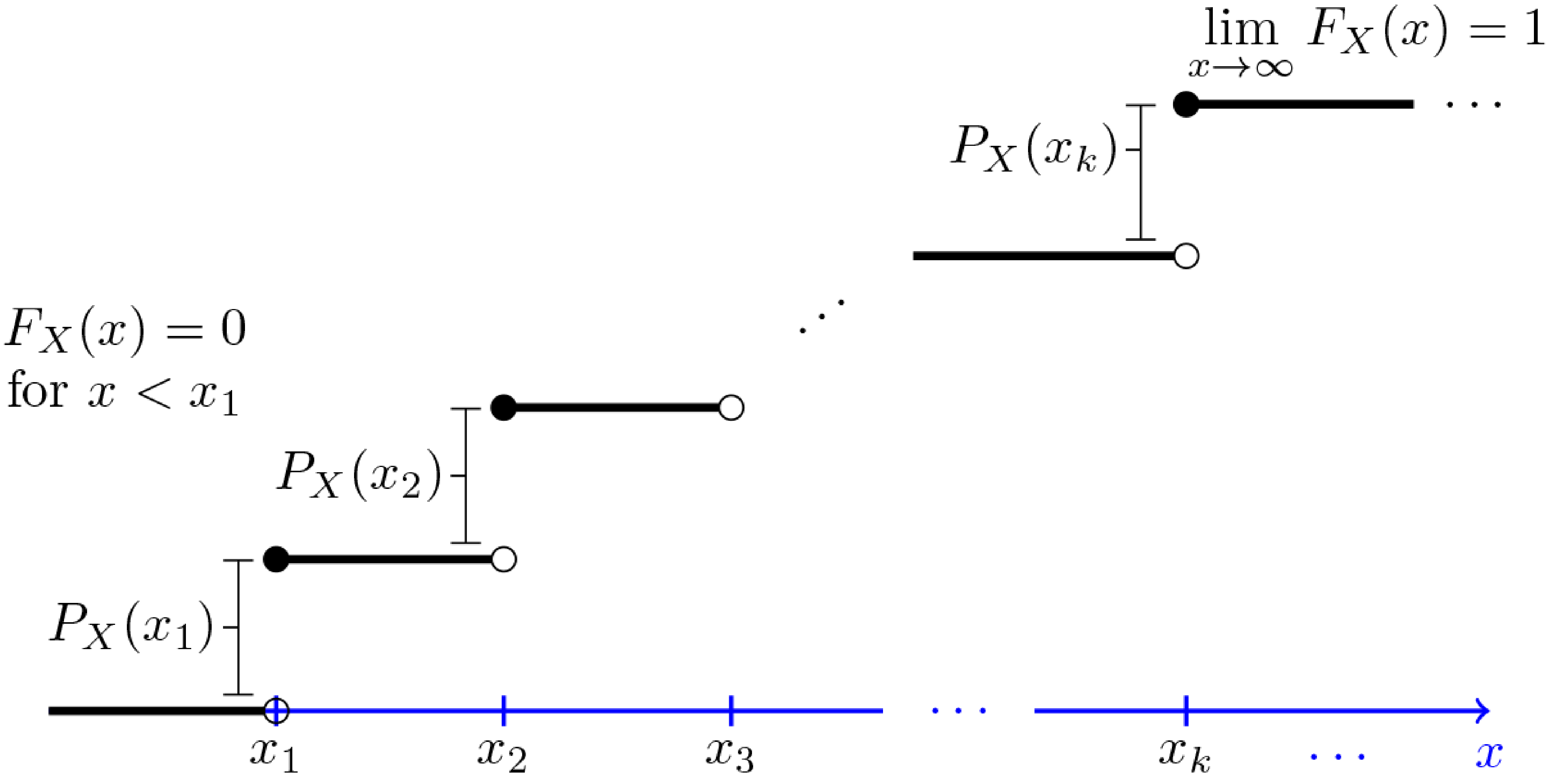
Book Reference: Chapter 3

TA: Mohamad Kazma

February 4th-5th, 2025

Instructor: Dr. Hiba Baroud

Today: Discrete Random Variables



Discrete RVs – Learning Objectives

- Random Variables and Probability Distribution
  - **Define** Random Events and Random Variables
  - **Identify** Probability Distribution of a Random Variable

- Main Descriptors of a Random Variable
  - **Calculate** Mean and Standard deviation of Probability Distributions

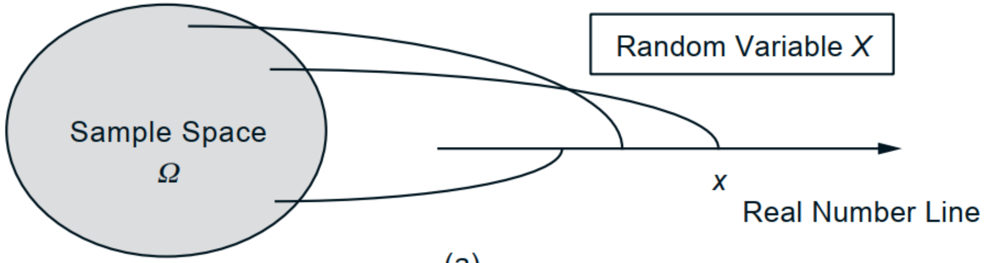
- **Count!**

- **Describe** discrete probability distributions

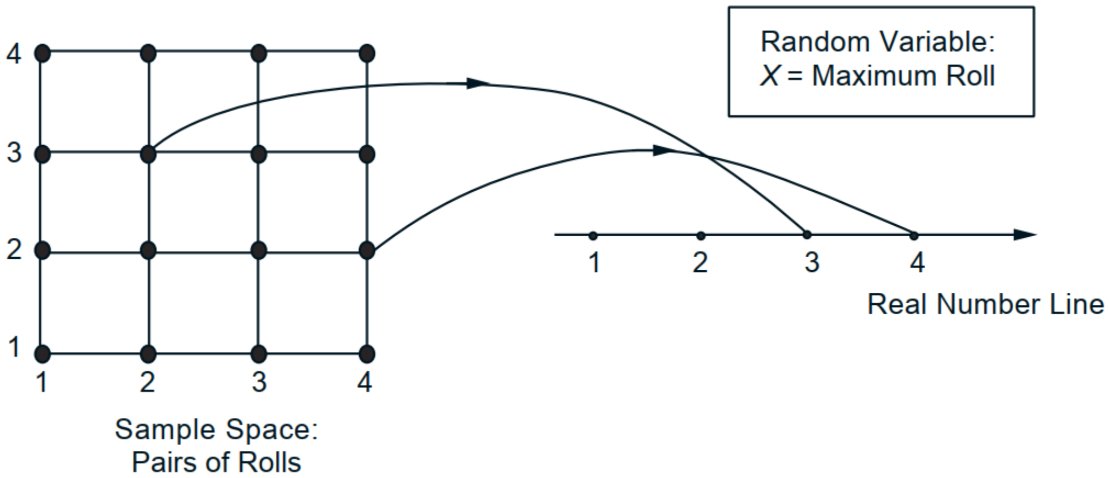
Probability Distributions and Random Variables

Extension of probability concepts to probability distributions

- The value of a **random variable (RV)** may be defined by a range of possible outcomes
- A RV is considered a mapping function that transforms (or maps) the events in a possibility space into the number system (the real line).

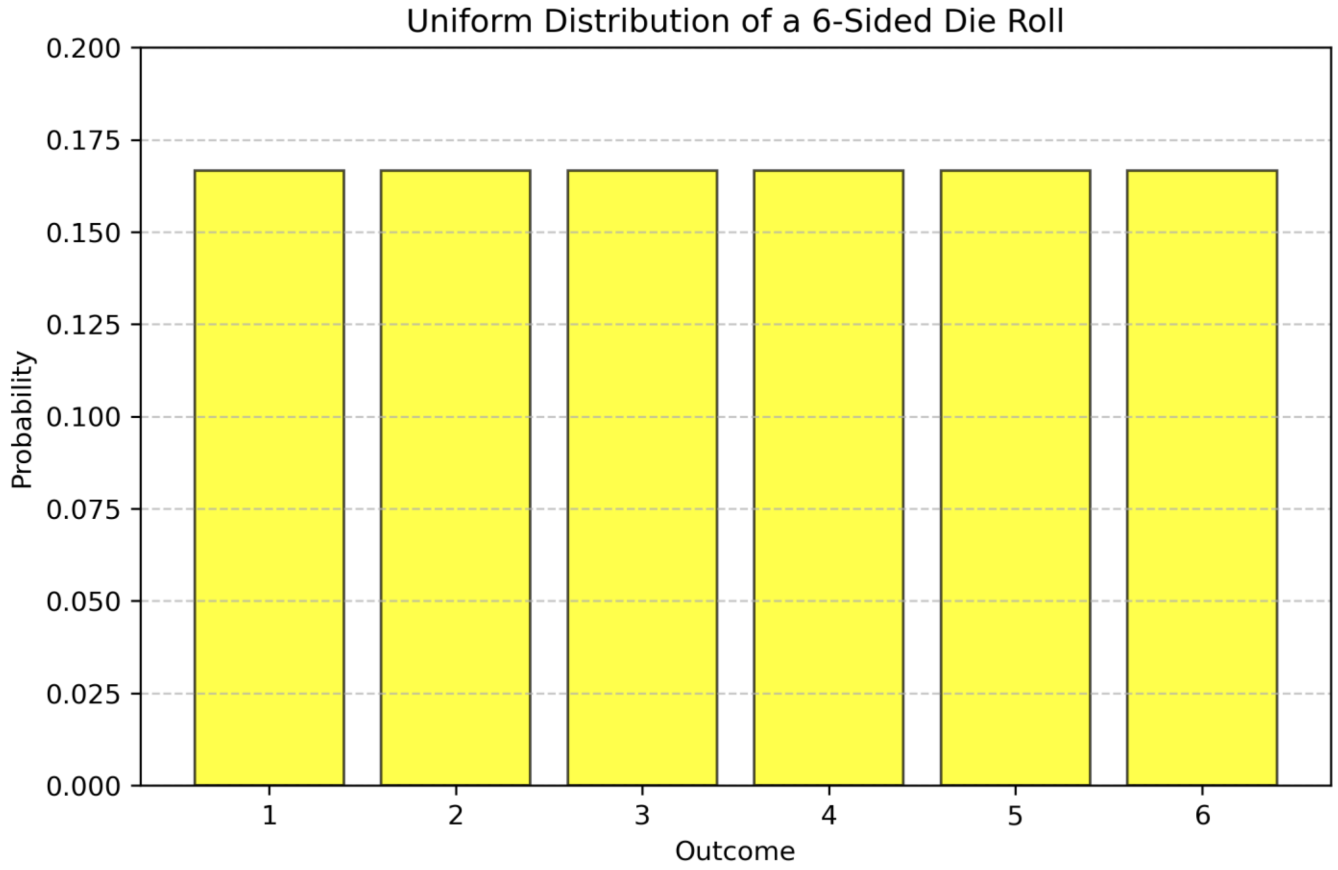


(a)



The Uniform Distribution of a RV

x_i Rolling a die	$P(X = x_i)$
1	1/6
2	1/6
3	1/6
4	1/6
5	1/6
6	1/6



Example 1 – Random Experiment

Let X random variable (RV) that represents the number of vehicles owned by a randomly selected family when an experiment of choosing a family at random is conducted.

Number of Vehicles Owned	Frequency	Relative Frequency
0	30	30/2000 = 0.015
1	470	470/2000 = 0.235
2	850	850/2000 = 0.425
3	490	490/2000 = 0.245
4	160	160/2000 = 0.08
	n=2000	Sum = 1

Discrete Vs. Continuous RVs

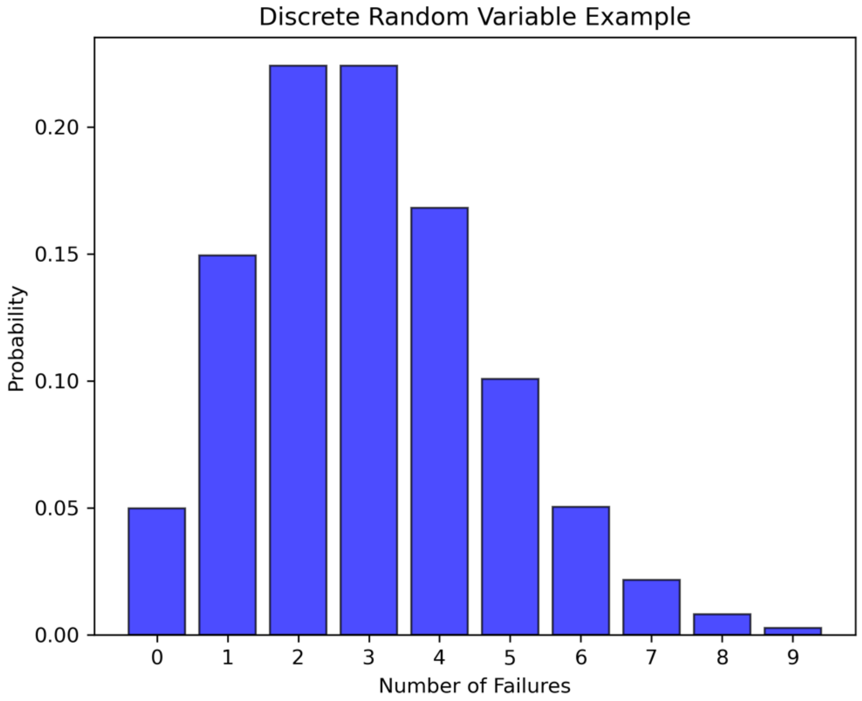
- A RV is a variable whose value can take on one of several realizations such that each realization is associated with a likelihood of occurrence

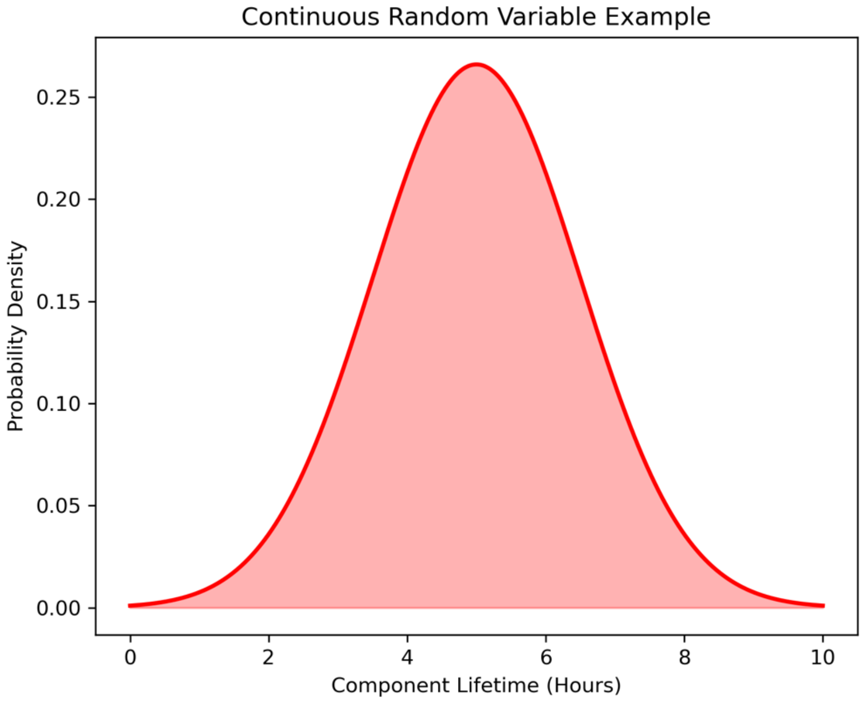
A **discrete** RV takes on a finite or countably infinite set of values. This means that the possible outcomes are distinct and can be listed individually.

A **continuous** RV can take on **any value** within an interval. Since there are infinitely many possible values, the probability of a single exact outcome is **zero.**

Discrete Vs. Continuous RVs

Feature	Discrete RV	Continuous RV
Possible Values	Countable (finite or infinite)	Uncountable (infinite real values)
Probability Assigned To	Specific values via PMF	Intervals via PDF
Example in Reliability	Number of system failures in a day	Time until the next system failure





Examples – Discrete Vs. Continuous RVs

RV Type	Example
Discrete	Number of system failures in a day
Discrete	Number of defective items in a batch
Discrete	Number of accidents at an intersection per day
Discrete	Number of calls received in a call center per hour
Discrete	Number of customers in a queue
Continuous	Lifetime of a machine component (in hours)
Continuous	Time between failures of a system
Continuous	Temperature of a sensor
Continuous	Voltage fluctuations in a circuit
Continuous	Stress level of a material under load

Example 2 – Type of RV

Determine if a RV is continuous or discrete

1. Time left on a parking meter. — Continuous
2. Total pounds of fish caught. — Continuous
3. Points scored in a football game. — Discrete
4. Weight of a randomly selected player. — Continuous
5. Number of defective items in a production facility. — Discrete

The Probability Distribution of a RV

A **Probability Distribution** describes how the values of a **random variable (RV)** are distributed.

- It lists of all possible values that a R.V can assume with their corresponding likelihood of occurrence.

Discrete Probability Distributions (e.g., Binomial, Poisson)

- Associated with discrete random variables.
- Uses a Probability Mass Function (PMF).

Continuous Probability Distributions (e.g., Normal, Exponential)

- Associated with continuous random variables.
- Uses a Probability Density Function (PDF).

The PMF and CDF of a Discrete RV

The **Probability Mass Function (PMF)** defines the probability that a discrete random variable X takes on a specific value x.

$$f_x(x) = p_X(x_i) = \mathbf{P}(X = x_i)$$

Axioms of Probability

$$0 \leq p_X(x_i) \leq 1 \qquad \sum_{x_i} p_X(x_i) = 1$$

The **Cumulative Distribution Function (CDF)** of a discrete random variable X, gives the probability that X takes a value **less than or equal to** x.

$$F_X(x) = \sum_{\text{all } x_i \leq x} P(X = x_i) = \sum_{\text{all } x_i \leq x} p_X(x_i)$$

Computation of a Discrete PMF and CDF

The Probability Mass Function (PMF) defines the probability that a discrete random variable X takes on a specific value x.

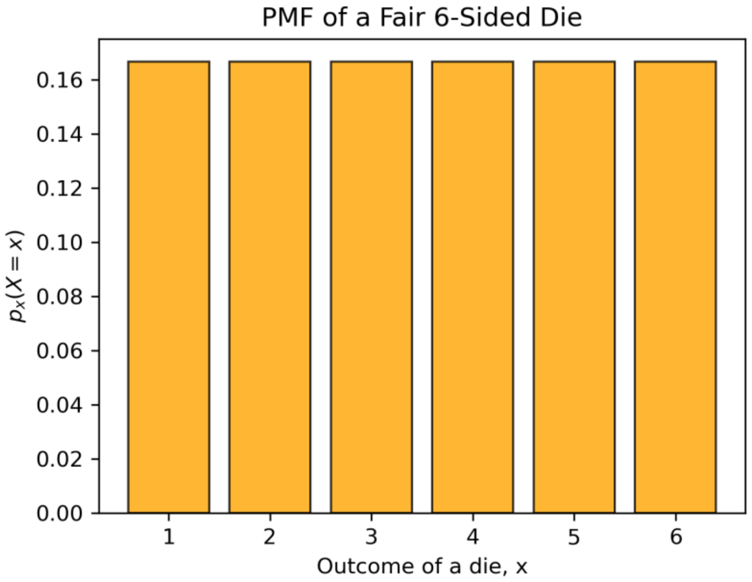
List/Table	List/Table	PMF	CDF $F_X(x)$ is non-decreasing

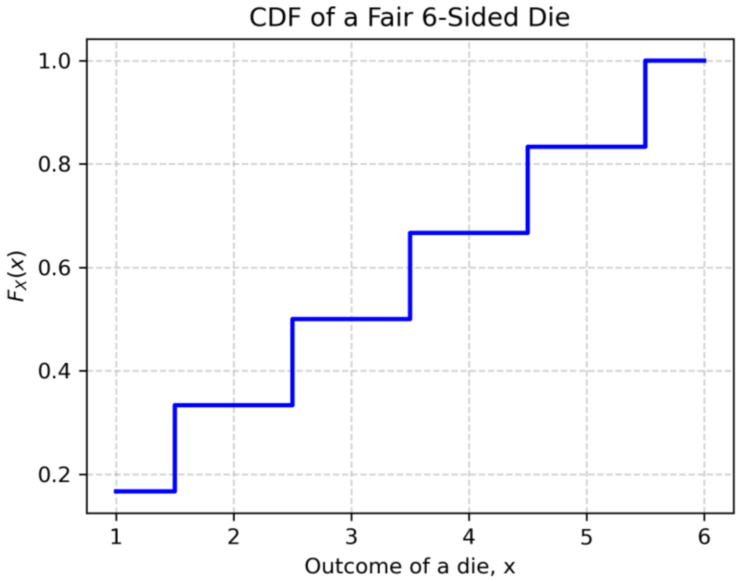
$$F_X(x) = \sum_{\text{all } x_i \leq x} P(X = x_i) = \sum_{\text{all } x_i \leq x} p_X(x_i)$$

$$p_X(x) = \mathbf{P}(\{X = x\})$$

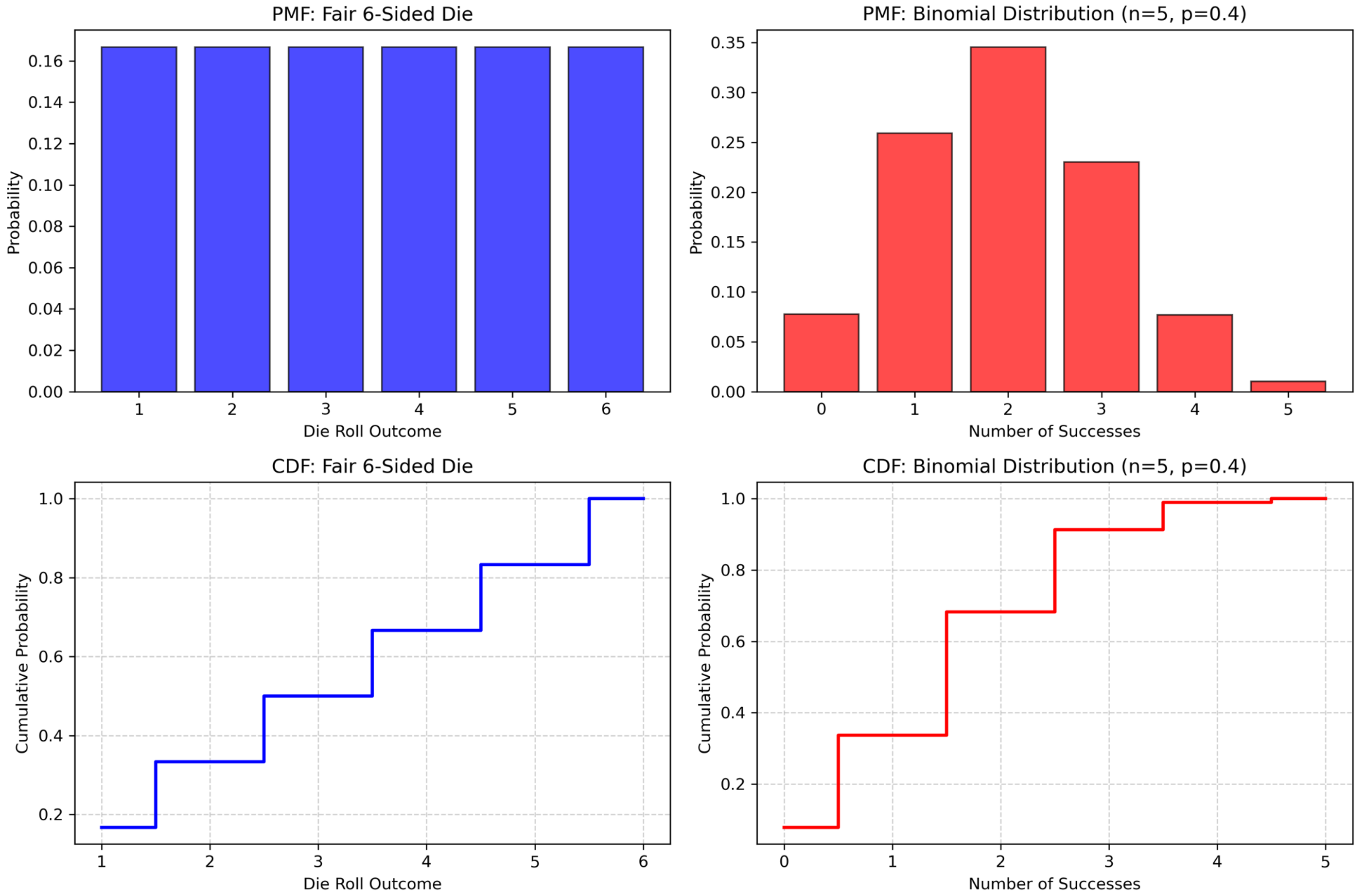
$$P(X = x) = \frac{1}{6}$$

x_i Rolling a die	$P(X = x_i)$
1	1/6
2	1/6
3	1/6
4	1/6
5	1/6
6	1/6





Examples – Discrete PMFs and CDFs



Descriptors of a RV – Expected Value

We define the Expected Value (also called the expectation or the mean) of a random variable X, with PMF $p\_X(x)$.

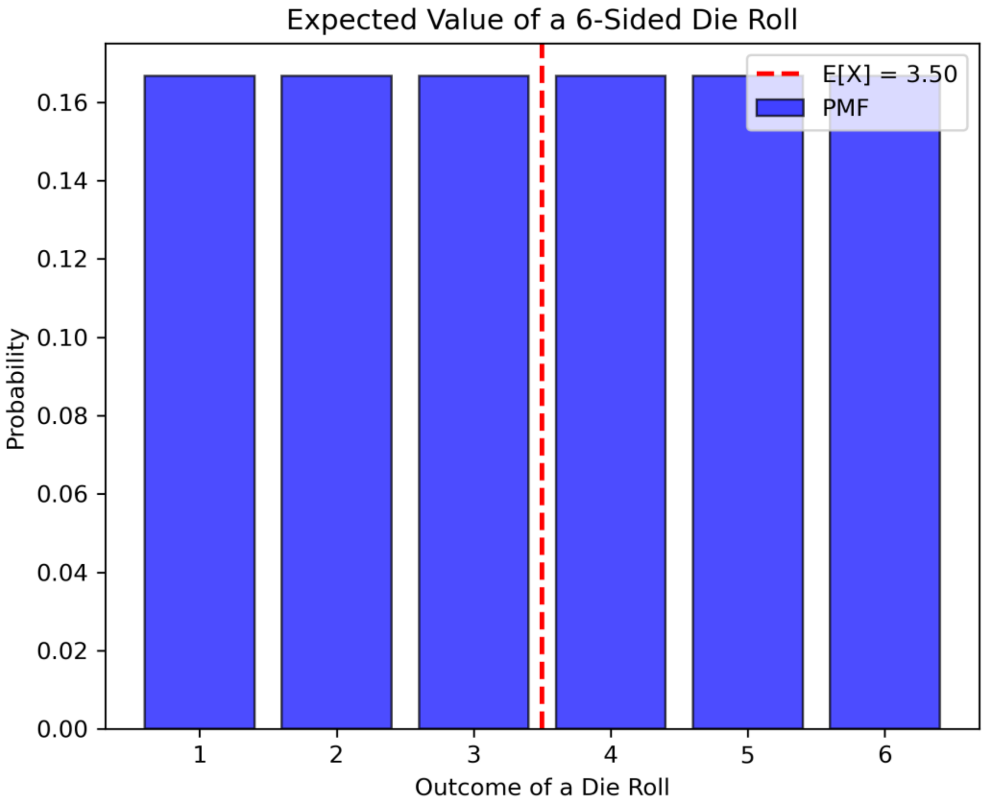
$$\mu_X = \mathbf{E}[X] = \sum_{x_i} x_i p_X(x_i)$$

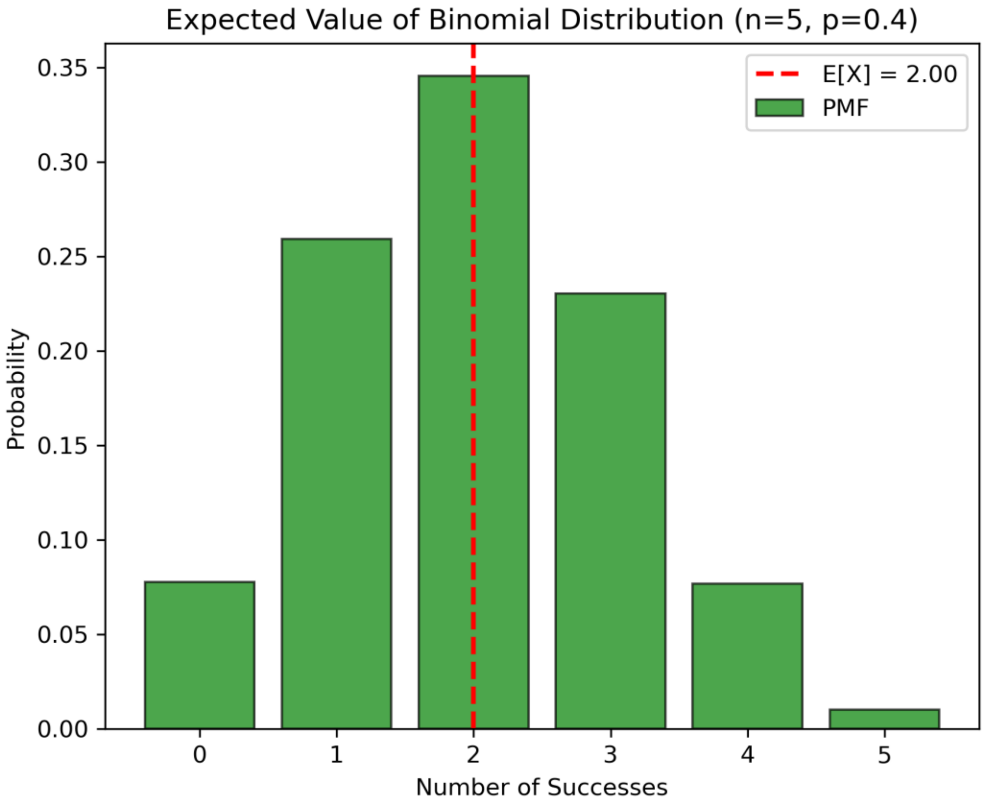
1st moment
Center of gravity of the PMF

Expected Value in Reliability Engineering

- Mean Time to Failure (MTTF): Average time until a system or component fails.
- Risk Analysis: Used to estimate expected losses or failure rates.

Examples– Expected Value of a Discrete RV





Descriptors of a RV – Variance

The variance provides a measure of dispersion of X around its mean.
How widely or narrowly the values of the random variable are dispersed.

$$\text{var}(X) = \mathbf{E}\left[(X - \mathbf{E}[X])^2\right]$$

2^{nd} moment
Moment of inertia of the area under the PDF, about the centroid (mean)

Higher variance means values are **spread out** from the mean.

- **Lower variance** means values are **closer** to the mean.
- **Variance is always non-negative**

Descriptors of a RV – Variance

$$\text{var}(X) = \mathbf{E}\left[(X - \mathbf{E}[X])^2\right]$$

Step 1: Expanding the squared term: $(X - \mu_X)^2 = X^2 - 2X\mu_X + \mu_X^2$
where $\mu_X = \mathbf{E}[X]$ is the expected value.

Step 2: Taking the expectation: $\text{Var}(X) = \mathbf{E}\left[X^2 - 2X\mu_X + \mu_X^2\right]$
Step 3: Linearity of Expectation: $\text{Var}(X) = \mathbf{E}[X^2] - 2\mu_X\mathbf{E}[X] + \mathbf{E}[\mu_X^2]$.

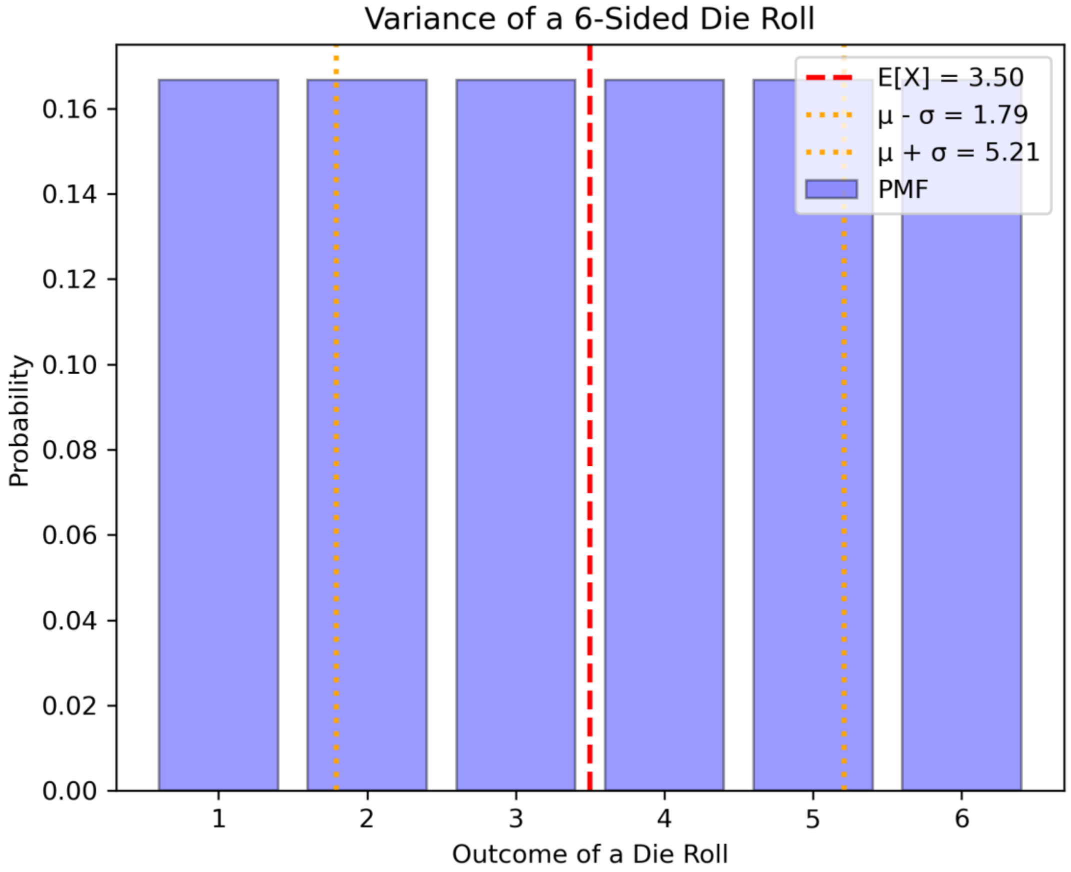
Since $\mathbf{E}[X] = \mu_X$, we substitute:

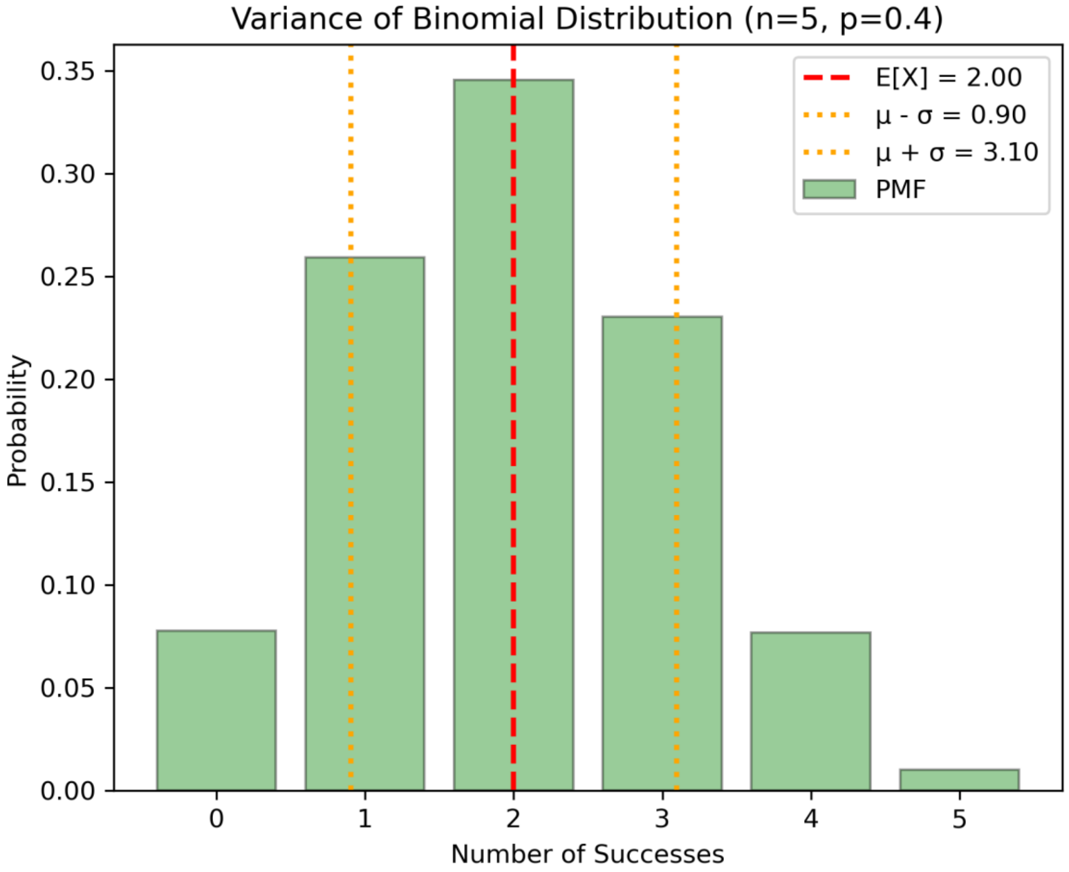
$$\text{Var}(X) = \mathbf{E}[X^2] - 2\mu_X^2 + \mu_X^2$$

Step 4: Simplification: $\text{Var}(X) = \mathbf{E}[X^2] - (\mathbf{E}[X])^2 = \mathbf{E}[X^2] - \mu_x^2$
Final Expression:

$$\text{Var}(X) = \sum x_i^2 p_X(x_i) - \left(\sum x_i p_X(x_i)\right)^2$$

Examples– Variance of a Discrete RV





$$\sigma_X = \sqrt{\mathrm{Var}(X)}$$

Example 3 – Vehicles Owned

Calculate the average number of vehicles owned/Family

$$E[X] = \sum x_i p_{x_i} \, (X = x_i)$$

$$E[X] = (0 \times 0.015) + (1 \times 0.235) + (2 \times 0.425) + (3 \times 0.245) + (4 \times 0.08)$$

$$E[X] = 0 + 0.235 + 0.85 + 0.735 + 0.32 = 2.14$$

Number of Vehicles Owned	Frequency	Relative Frequency
0	30	30/2000 = 0.015
1	470	470/2000 = 0.235
2	850	850/2000 = 0.425
3	490	490/2000 = 0.245
4	160	160/2000 = 0.08
	n=2000	Sum = 1

Example 3 – Continued

$$\text{var}(X) = \mathbf{E}\left[(X - \mathbf{E}[X])^2\right]$$

We first compute $x_i^2 P(X = x_i)$

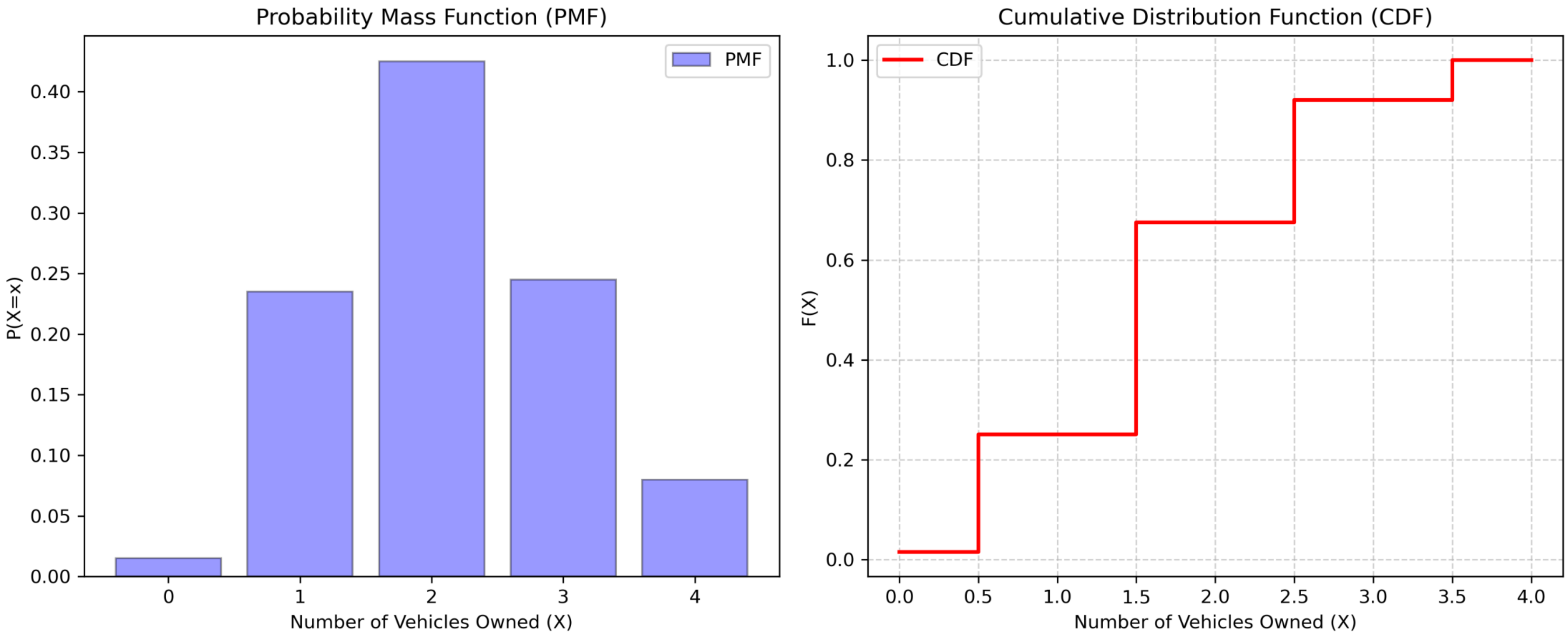
$$E\left[X^2\right] = \sum x_i^2 P(X = x_i)$$

$$E\left[X^2\right] = (0^2 \times 0.015) + (1^2 \times 0.235) + (2^2 \times 0.425) + (3^2 \times 0.245) + (4^2 \times 0.08)$$

$$E\left[X^2\right] = 0 + 0.235 + 1.7 + 2.205 + 1.28 = 5.42$$

$$\text{Var}(X) = 5.42 - (2.14)^2 = 0.8404$$

Example 3 – Continued



$$P(X \leq 2) = 0.425 + 0.235 + 0.015$$

Example 4 – Blood Type

The American Red Cross says that about
45% of the US population has type O, 40% has type A, 11% has type B,
and the rest has type A B.

a. In a sample of four donors, what is the probability that 4 are type O

$$P(O \cap O \cap O \cap O)$$

Let X be the number of type O blood donors in a sample of 4 . Since each donor has an independent probability of being type O , we define X as a binomial random variable:

$$X \sim \text{Binomial}(n = 4, p = 0.45)$$

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

$$P(X = 4) = \binom{4}{4} (0.45)^4 (1 - 0.45)^{4-4} = 0.45^4$$

Example 4 – Blood Type

b. In a sample of four donors, what is the probability that 3 are type O

$$P(O \cap O \cap O \cap \bar{O})$$

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

$$P(X = 3) = \binom{4}{3} (0.45)^3 (1 - 0.45)^{4-3} = 0.45^3 \times 0.55^1$$

Combinatorics

Combinatorics is the study of **counting and arranging objects**. It plays a fundamental role in probability theory and statistics.

- Used in the calculation of **discrete probabilities**.
- Helps determine **the total possible realizations** of a discrete random variable.
- Finds the total occurrences of a realization.

- **Computing the chances of winning** the lottery.
- Analyzing the **number of ways to form** a team.
- Determining **possible arrangements** of objects.

Multiplication Rule

The multiplication rule is used to **count the total number of ways multiple independent operations** can be performed.

- Operation 1 can be performed in n_1 ways.
- Operation 2 can be performed in n_2 ways.

The total number of ways to perform both operations together is:

$$\text{Total ways} = n_1 \times n_2$$

- This rule only applies to **independent events**.
- When events are dependent, **conditional probability** is required.
- **Not applicable** when order matters (this is where permutations are used instead).
- Decision trees can be used to visualize choices (covered in another class).

Example 5 – Multiplication Rule

A customer is selecting a car with the following options:

- Car brands: Honda, Toyota, Nissan (3 choices)
- Models: Two-door or four-door (2 choices)
- Exterior colors: (5 choices)
- Interior colors: (4 choices)

What are the different Configurations the customer can choose from?

$$\text{Total Car Options} = 3 \times 2 \times 5 \times 4 = 120 \text{ different car configurations}$$

Permutation (Arrangement)

A permutation is an ordered arrangement of r elements chosen from n elements, meaning that order matters.

$$_nP_r = \frac{n!}{(n-r)!}$$

- $n\,!$ represents the total number of arrangements of n elements,
- $(n-r)\,!$ adjusts for the elements not selected.

Assigning 5 different medals (gold, silver, bronze, etc.) to 5 athletes from a group of 12 :

$$_{12}P_5 = \frac{12!}{(12-5)!}$$

Combination (Selection)

A combination is a way of **selecting r elements from a set of n** elements **without considering order**.

$$ {}_nC_r = \binom{n}{r} = \frac{n!}{r!(n-r)!} $$

- n ! (factorial) represents the number of ways to arrange n elements,
- $r!$ accounts for the arrangements of the selected r elements,
- $(n-r)!$ adjusts for the remaining unselected elements.

Selecting a committee of **3** people from a group of 10 :

$$ \binom{10}{3} = \frac{10!}{3!(10-3)!} = \frac{10!}{3!7!} = \frac{10 \times 9 \times 8}{3 \times 2 \times 1} = 120 $$

Example 6 – Combinations and Permutations

A school is forming a **student council** from a group of **8 students**, and they need to:

1. **Select 3 students to be part of the council (order doesn't matter)**.
2. **Assign 3 specific leadership positions (President, Vice President, and Treasurer) to 3 students (order matters)**.

Part 1: Solving Using Combinations

Since order **does not matter**, we use the combination formula:

$$\binom{8}{3} = \frac{8!}{3!(8-3)!} = \frac{8!}{3!5!} = \frac{8 \times 7 \times 6}{3 \times 2 \times 1} = 56$$

Part 2: Solving Using Permutations

Since we are assigning leadership positions **(order matters),** we use the permutation formula:

$${}_8P_3 = \frac{8!}{(8-3)!} = \frac{8!}{5!} = 8 \times 7 \times 6 = 336$$

Binomial Distribution - Definition

Bernoulli trial is a problem involving the occurrence or recurrence of an event, which is unpredictable, in a sequence of discrete "trials."

A **Bernoulli sequence** is a series of independent Bernoulli trials, forming the basis of the Binomial distribution.

The **Bernoulli distribution** models n identical trial of a binary experiment, where there are only two possible outcomes:

Examples of Bernoulli trials:

- Flipping a coin ($p = 0.5$)
- Checking if a machine part is defective ($p = 0.02$)
- Passing or failing a test ($p = 0.8$)

Binomial Distribution – PMF and Parameters

The PMF of a Bernoulli-distributed random variable X is:

$$\begin{cases} p, & \text{if } x_1 = 1 \quad \text{(success)} \\ 1-p, & \text{if } x_0 = 0 \quad \text{(failure)} \end{cases}$$

PMF Binomial

$$X \sim \text{Binomial}(n, p)$$

$$p_x(X = x) = f_x(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

- n is the number of trials,
- p is the probability of "success"

- ***n* identical** Bernoulli trials
- Each trial has **only two possible outcomes**
- Probability of each **outcome is constant**
- Trials are independent
- Range is from 0 to n

Binomial Distribution – Moments

The expectation of a binomially distributed random variable is

$$E[X] = E\left[X_1 + X_2 + \ldots + X_n\right]$$

$$= E\left[X_1\right] + E\left[X_2\right] + \ldots + E\left[X_n\right]$$

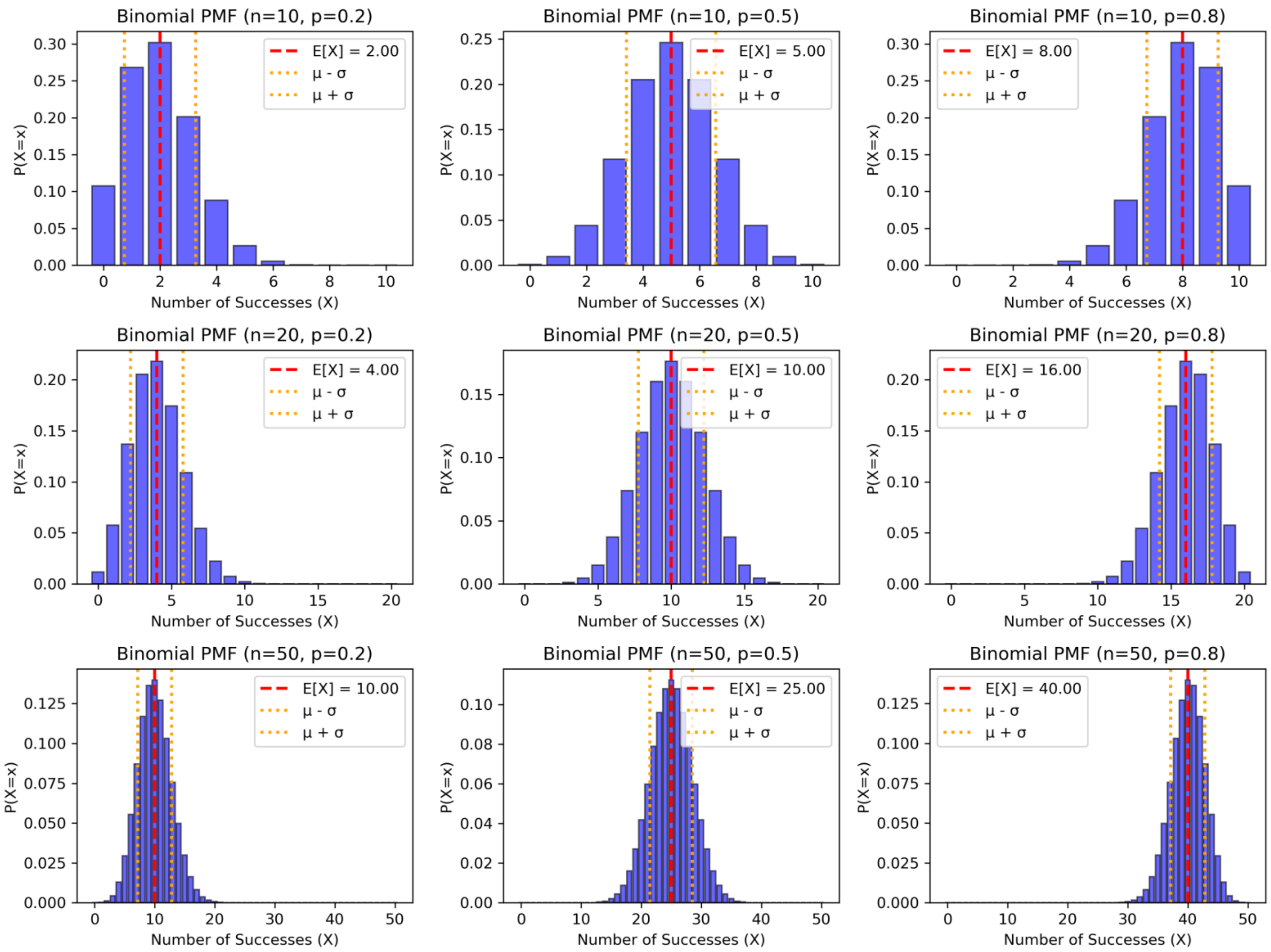
$$E[X] = np$$

The Variance of a Binomial-distributed variable is

$$\mathrm{Var}(X) = E\left[X^2\right] - \mu_x^2 = np(1-p)$$

$$\sigma = \sqrt{np(1-p)}$$

Binomial Distributions



Example 7.1 – Patient Recovery

The probability that a patient recovers from a rare disease is known to be 0.5.

- If 15 people are known to have contracted the disease.
- **What is the probability that two survive**?

$$x : \# \text{ of survivors} \qquad X \sim \text{Binomial}(n = 15,\ \text{p=0.5})$$

$$\begin{aligned} P(x = 2) &= C_2^{15} 0.5^2 (1 - 0.5)^{15-2} \\ &= \frac{15!}{2!13!} 0.5^2 0.5^{13} \end{aligned}$$

Example 7.2 – Patient Recovery

The probability that a patient recovers from a rare disease is known to be 0.5.

- If 15 people are known to have contracted the disease.
- **What is the probability that at least 10 survive**?

$$P(x \geq 10) = P(x = 10) + P(x = 11) + \ldots + P(x = 15)$$

$$= C_{10}^{15} 0.5^{10} 0.5^{5} + \cdots + C_{15}^{15} 0.5^{15} 0.5^{0}$$

Poisson Distribution – Definition

The Poisson distribution models the **number of times an event occurs** in a **fixed interval** of time or space

- Events occur independently.
- Occurrences are random
- The rate (# occurrences per unit time or space) **is constant and is not necessarily discrete.**
- The probability of multiple occurrences in a very small interval is negligible.

Examples
- Occurrence of hurricanes in a particular region (space process)
- Occurrence of flood over the lifetime of a structure (time process)
- Cracks in a weld (space process)

Siméon Denis Poisson (1781-1840)

Poisson Distribution – PMF and Parameters

A random variable X follows a Poisson distribution if it counts the number of occurrences in an interval with a **constant average rate λ**

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

The PMF of a Poisson-distributed variable X is

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

where λ is the expected number of occurrences per unit.

Time-Based Process:
- Flood occurrence over the lifetime of a structure.
- Customer arrivals at a store per hour.

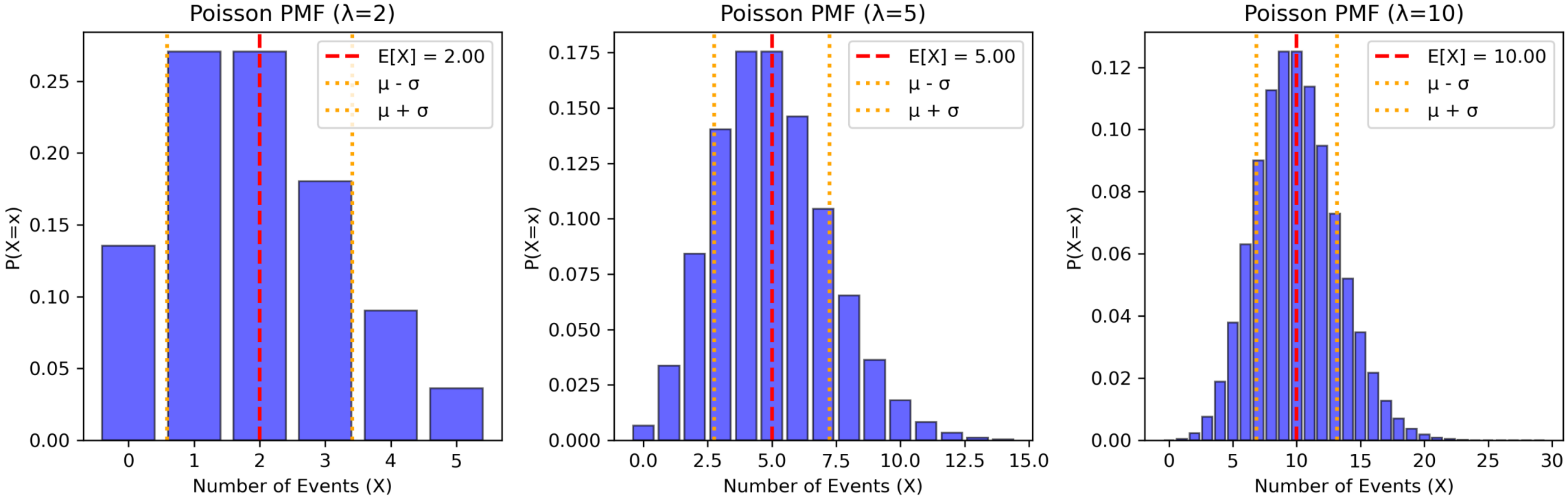
Space-Based Process:
- Hurricanes in a particular region.
- Cracks in a weld per meter.

$x! =$ factorial of x, which accounts for different ways events can occur

Poisson Distribution – Moments

$$E[X] = V[X] = \lambda$$

Poisson Distribution



Example 8.1 – Number of Emergency Calls

The number of emergency calls arriving at the Nashville fire department for various incidents (fires, crashes, etc.) is modeled as a **Poisson RV**. Assume there are on average 10 calls/hour.

- **Calculate P(exactly 5 calls in one hour)**

$$X = \#\text{ of calls}$$
$$x \sim \text{POI}(\lambda = 10 \text{ calls / hour })$$

$$P(X = 5) = \frac{10^5 e^{-10}}{5!}$$

Example 8.2 – Number of Emergency Calls

The number of emergency calls arriving at the Nashville fire department for various incidents (fires, crashes, etc.) is modeled as a **Poisson RV**.
Assume there are on average 10 calls/hour.

- **Calculate P(3 calls or less in one hour)**

$$X = \# \text{ of calls}$$
$$x \sim \text{POI}(\lambda = 10 \text{ calls } / \text{ hour })$$

$$P(X \leq 3) = \frac{10^3 e^{-10}}{3!} + \frac{10^2 e^{-10}}{2!} + \frac{10e^{-10}}{1!} + \frac{1e^{-10}}{0!}$$

Example 8.3 – Number of Emergency Calls

The number of emergency calls arriving at the Nashville fire department for various incidents (fires, crashes, etc.) is modeled as a **Poisson RV**. Assume there are on average 10 calls/hour.

- **Calculate P(exactly 15 calls in 2 hours)**

x : # of calls

$x \sim \text{POI}/(\lambda = 20/hours)$

$$P(X = 15) = \frac{20^{15} e^{-20}}{15!}$$

Example 8.4 – Number of Emergency Calls

The number of emergency calls arriving at the Nashville fire department for various incidents (fires, crashes, etc.) is modeled as a **Poisson RV**. Assume there are on average 10 calls/hour.

- **Calculate P(exactly 5 calls in 30 mins)**

$$X : \# \text{ of calls}$$

$$X \sim \text{POI}(\lambda = 5 \text{ calls } /30\text{mns})$$

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

Special Cases

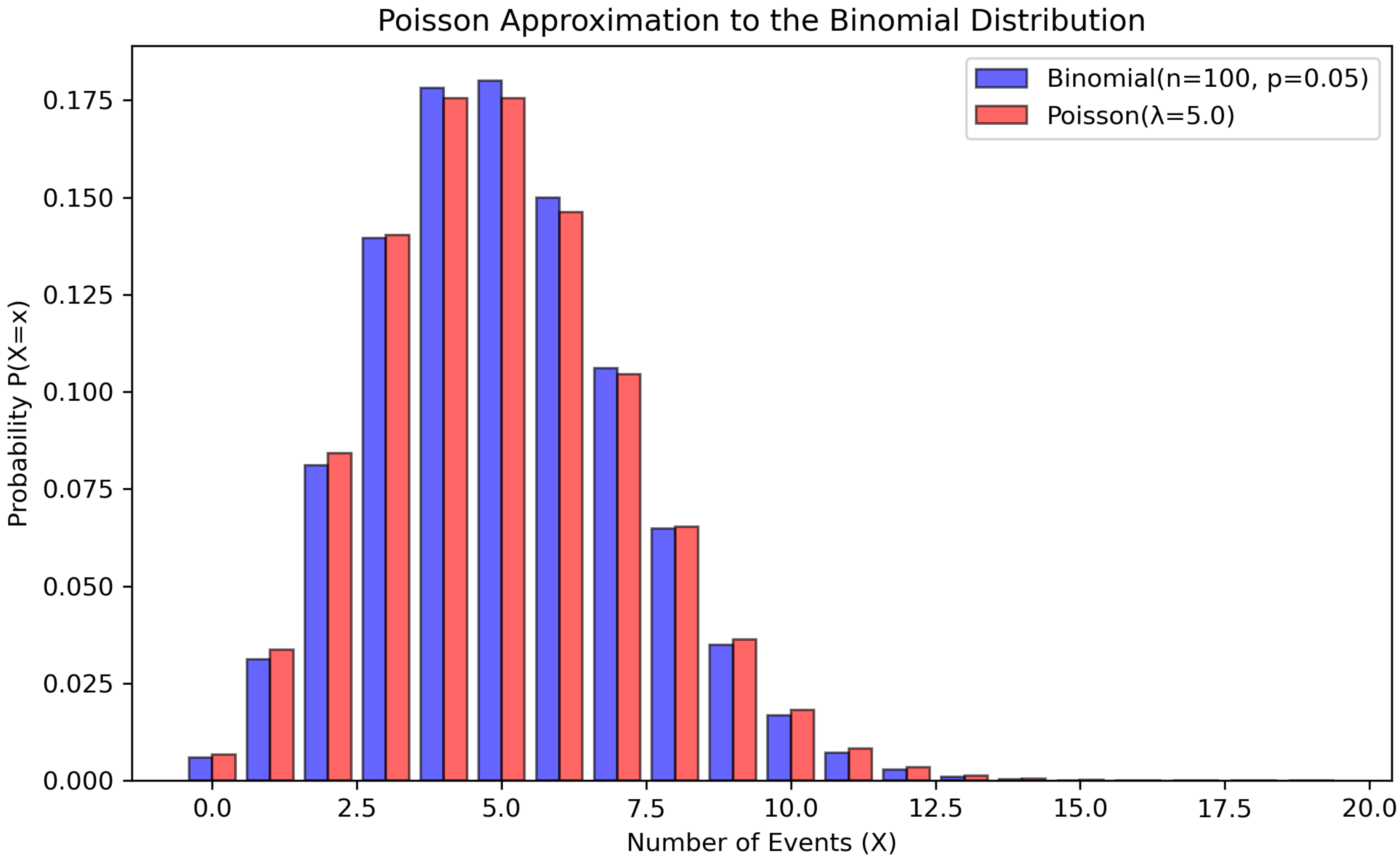
- For large *n*, the binomial distribution can be approximated by the normal distribution.
- For large *n* and small *p*, the binomial distribution can be approximated by the Poisson distribution

$$X \sim \text{binomial}(n, p)$$

$$X \sim normal(np, \sqrt{npq}) \quad \text{if } npq \text{ not small}$$

$$X \sim poisson(\lambda) \quad \lambda = np \quad \text{if } np \text{ small}$$

Special Cases



Example 9 – Page Error

A publisher ensures that the probability of any **given page containing one or more errors is 0.005.**

- Assuming that errors are independent from page to page, what is the **probability that one of its 400-page novels will contain exactly one page with errors**?

Option 1

$$P(1 \text{ out } 400 \text{ pages with errors })$$

$$= 400 \times 0.005 \times (1 - 0.005)^{399}$$

Example 9 – Continued

A publisher ensures that the probability of any **given page containing one or more errors is 0.005.**

- Assuming that errors are independent from page to page, what is the **probability that one of its 400-page novels will contain exactly one page with errors**?

Option 2

$$X : \# \text{ of pages with } \geqslant 1 \text{ errors}$$

$$X \sim \text{Bin}(n = 400, p = 0.005)$$

$$P(X = 1) = C_1^{400} 0.005^1 (1 - 0.005)^{399}$$
$$= 0.27$$

Example 9 – Continued

A publisher ensures that the probability of any **given page containing one or more errors is 0.005.**

- Assuming that errors are independent from page to page, what is the **probability that one of its 400-page novels will contain exactly one page with errors**?

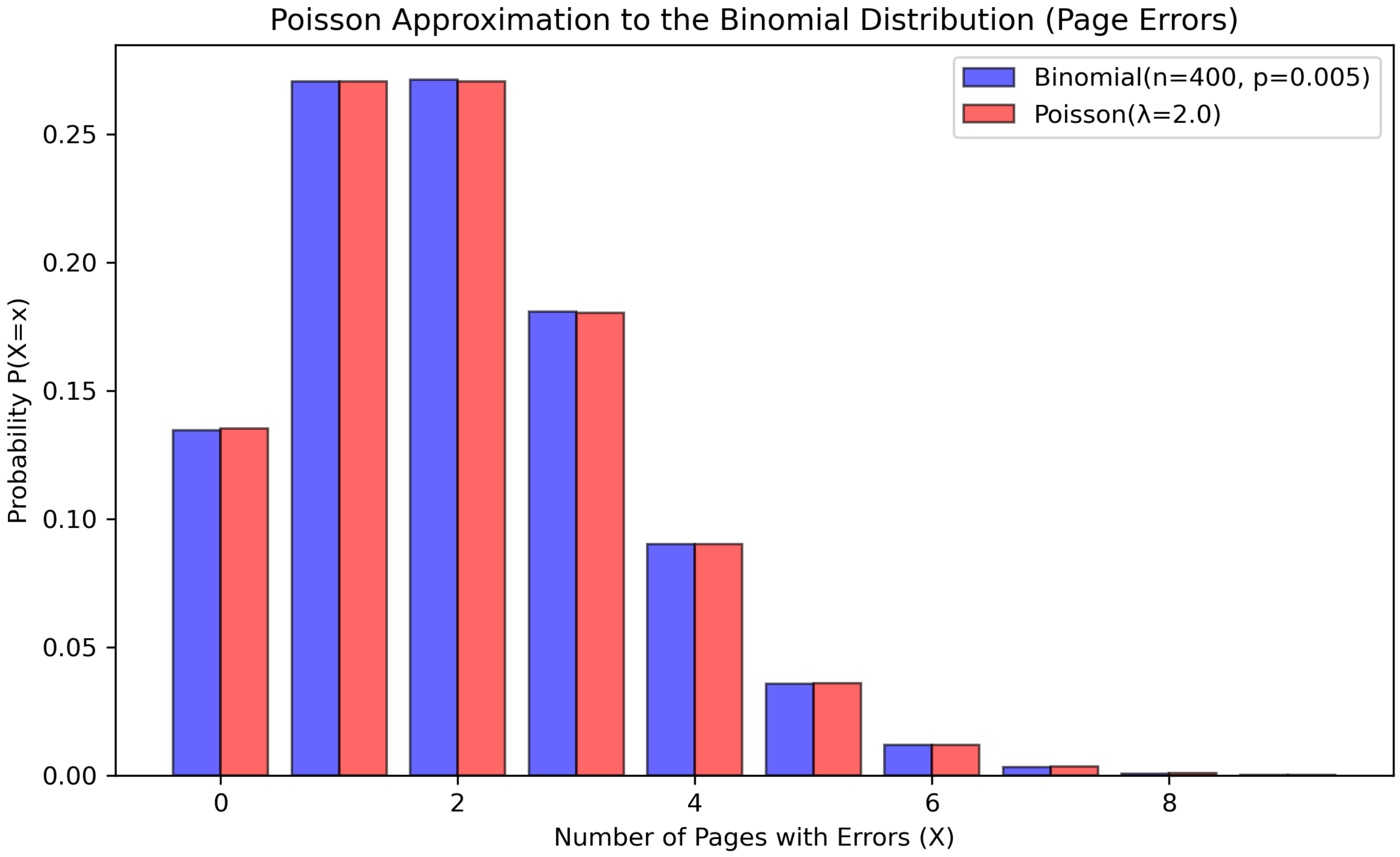
Option 3

X: # of pages with $\geqslant$ 1errors

$$X : \mathrm{POI}(\lambda = 400 \times 0.005 = 2)$$

$$P(X = 1) = \frac{2^1 e^{-2}}{1!} = 0.27$$

Example 9 – Continued



Next: Quiz + Continuous Random Variables