0.1 Learning Objectives

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

- Distinguish between discrete and continuous random variables
- Understand why P(X = exact value) = 0 for continuous RVs
- Interpret probability density functions (PDF)
- Apply the continuous uniform distribution
- Compute probabilities using PDF and CDF

0.2

So far, we've studied **discrete** random variables:

- Number of heads in n coin flips (Binomial)
- Number of trials until first success (Geometric)
- Number of events in a time period (Poisson)

0.2

So far, we've studied **discrete** random variables:

- Number of heads in n coin flips (Binomial)
- Number of trials until first success (Geometric)
- Number of events in a time period (Poisson)

These are great for modeling **countable** outcomes.

0.2

So far, we've studied **discrete** random variables:

- Number of heads in n coin flips (Binomial)
- Number of trials until first success (Geometric)
- Number of events in a time period (Poisson)

These are great for modeling **countable** outcomes.

But how would we model these?

- Height of a randomly selected person
- Time until the next customer arrives
- Temperature at noon tomorrow
- Position of a particle after random motion

0.2

So far, we've studied **discrete** random variables:

- Number of heads in n coin flips (Binomial)
- Number of trials until first success (Geometric)
- Number of events in a time period (Poisson)

These are great for modeling **countable** outcomes.

But how would we model these?

- Height of a randomly selected person
- Time until the next customer arrives
- Temperature at noon tomorrow
- Position of a particle after random motion

These quantities can take **any value** in a range – we need **continuous** random variables!

0.3 Quick Recall: The Sand Analogy

Recall: The **sand analogy** for distributions:

- We have **1 unit of probability mass** (like 1 kg of sand)
- A distribution describes **how that mass is spread out**

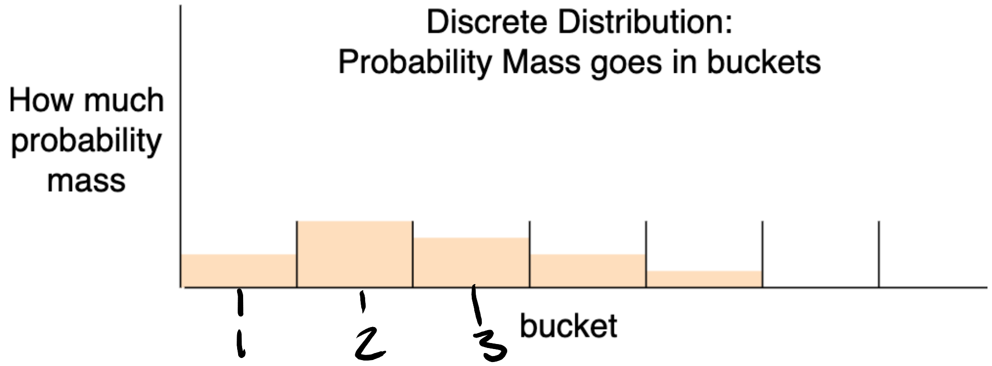
0.3 Quick Recall: The Sand Analogy

Recall: The **sand analogy** for distributions:
- We have **1 unit of probability mass** (like 1 kg of sand)
- A distribution describes **how that mass is spread out**

Discrete distributions:

Put sand in buckets

$P(X = 3) = 0.25 \longrightarrow 25\%$ of the sand is in the "$k = 3$" bucket



0.4 From Buckets to Sand Piles

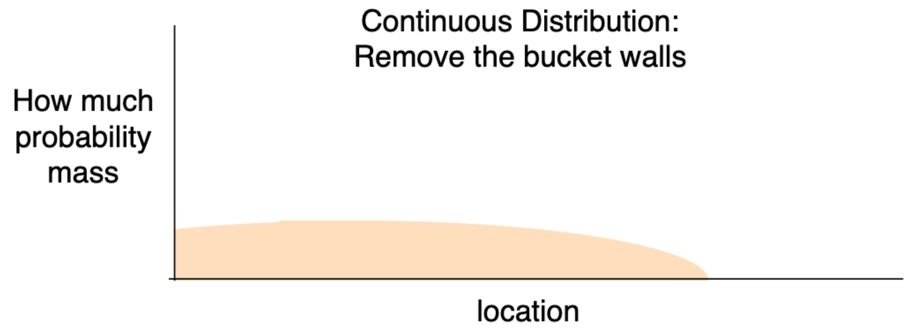
Continuous: Probability mass is spread continuously

- No discrete "buckets" – mass is everywhere

0.4 From Buckets to Sand Piles

Continuous: Probability mass is spread continuously

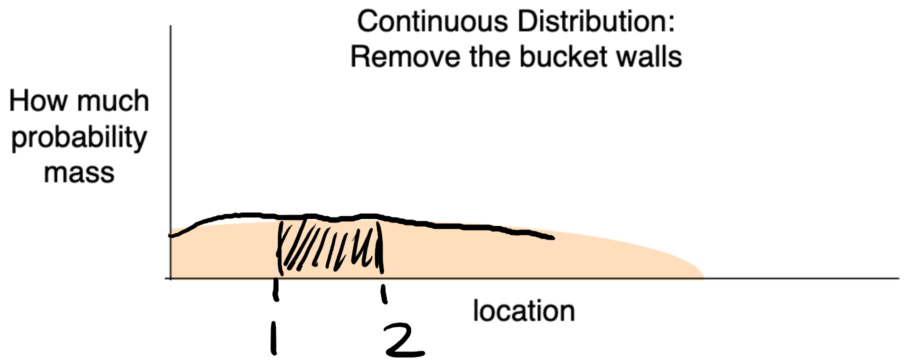
- No discrete "buckets" – mass is everywhere



0.4 From Buckets to Sand Piles

Continuous: Probability mass is spread continuously

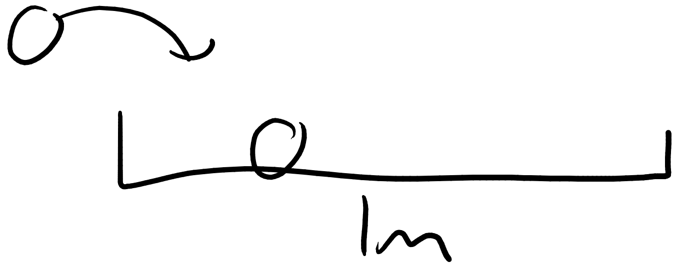
- No discrete "buckets" – mass is everywhere



How do we measure probability when there are no buckets?

1. The Density Paradox

1.1 A Marble in a Drawer



You have an empty drawer with a width of 1 meter.

You place a marble inside, close your eyes, and roll it.

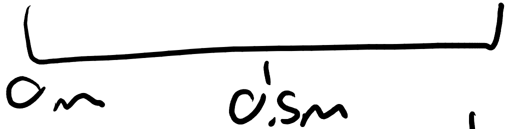
It bounces around until it comes to rest.

1.1 A Marble in a Drawer

You have an empty drawer with a width of 1 meter.

You place a marble inside, close your eyes, and roll it.

It bounces around until it comes to rest.



Question: What is the probability the marble is on the left half of the drawer? $\frac{1}{2}$

1.1 A Marble in a Drawer

You have an empty drawer with a width of 1 meter.

You place a marble inside, close your eyes, and roll it.

It bounces around until it comes to rest.

Question: What is the probability the marble is on the left half of the drawer?

Answer: $P(X \leq 0.5) = 0.5$ (equally likely anywhere)

1.1 A Marble in a Drawer

You have an empty drawer with a width of 1 meter.

You place a marble inside, close your eyes, and roll it.

It bounces around until it comes to rest.

Question: What is the probability the marble is on the left half of the drawer?

Answer: $P(X \leq 0.5) = 0.5$ (equally likely anywhere)

Question: What is the probability the marble is on the left quarter of the drawer?

1.1 A Marble in a Drawer

You have an empty drawer with a width of 1 meter.

You place a marble inside, close your eyes, and roll it.

It bounces around until it comes to rest.

Question: What is the probability the marble is on the left half of the drawer?

Answer: $P(X \leq 0.5) = 0.5$ (equally likely anywhere)

Question: What is the probability the marble is on the left quarter of the drawer?

Answer: $P(X \leq 0.25) = 0.25$ (equally likely anywhere)

1.2 Getting More Precise

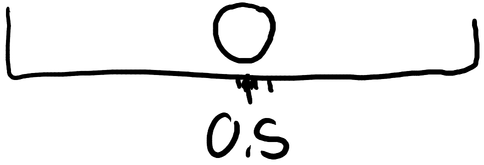
> ***Try it yourself***
>
> Talk to your neighbor and try to solve this problem.

$\frac{1}{10}$? 0?

What is the probability the marble is at **exactly** 0.5 meters?

0.5000000000 →

0.49 0.51



1.2 Getting More Precise

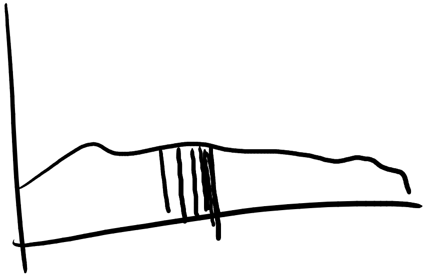
1.2 Getting More Precise

> ***Try it yourself***
>
> Talk to your neighbor and try to solve this problem.

What is the probability the marble is at **exactly** 0.5 meters?

Let's think about this by shrinking the interval:

Region	Width	Probability
Left 50%	0.5 m	0.5
Left 25%	0.25 m	0.25
Left 10%	0.1 m	0.1
Left 1%	0.01 m	0.01
…	…	…
Exactly 0.5	**0 m**	**0**



1.2 Getting More Precise

1.2 Getting More Precise

Try it yourself

Talk to your neighbor and try to solve this problem.

What is the probability the marble is at **exactly** 0.5 meters?

Let's think about this by shrinking the interval:

Region	Width	Probability
Left 50%	0.5 m	0.5
Left 25%	0.25 m	0.25
Left 10%	0.1 m	0.1
Left 1%	0.01 m	0.01
…	…	…
Exactly 0.5	**0 m**	**0**

1.2 Getting More Precise

As the region shrinks to a single point, the probability goes to **zero**!

1.3 The Paradox

Definition: The Density "Paradox"

For any continuous random variable X:

$$P(X = x) = 0 \quad \text{for any specific value } x$$

1.3 The Paradox

Definition: The Density Paradox

For any continuous random variable X:

$$P(X = x) = 0 \quad \text{for any specific value } x$$

This seems strange!

- The marble **is** somewhere
- Every location has probability 0
- Yet the marble must be at one of these locations

1.3 The Paradox

Definition: The Density Paradox

For any continuous random variable X:

$$P(X = x) = 0 \quad \text{for any specific value } x$$

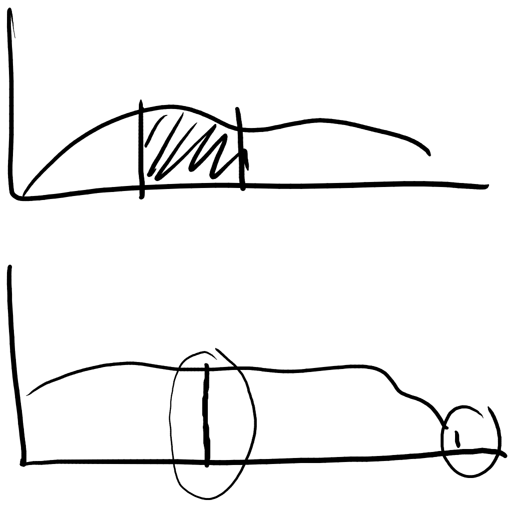
This seems strange!

- The marble **is** somewhere
- Every location has probability 0
- Yet the marble must be at one of these locations

How do we resolve this?

1.4 Resolving the Paradox: Density

The answer: We use **probability density** instead of probability mass.



1.4 Resolving the Paradox: Density

The answer: We use **probability density** instead of probability mass.

Definition: Probability Density Function (PDF)

For a continuous random variable X, the **PDF** $f_X(x)$ describes how probability is **spread out** (distributed) across values.

1.4 Resolving the Paradox: Density

The answer: We use **probability density** instead of probability mass.

Definition: Probability Density Function (PDF)

For a continuous random variable X, the **PDF** $f_X(x)$ describes how probability is **spread out** (distributed) across values.

The PDF does **not** return a probability.

1.4 Resolving the Paradox: Density

The answer: We use **probability density** instead of probability mass.

Definition: Probability Density Function (PDF)

For a continuous random variable X, the **PDF** $f_X(x)$ describes how probability is **spread out** (distributed) across values.

The PDF does **not** return a probability.

Evaluating the PDF at a point gives a **density** – probability per unit length.

1.4 Resolving the Paradox: Density

The answer: We use **probability density** instead of probability mass.

Definition: Probability Density Function (PDF)

For a continuous random variable X, the **PDF** $f_X(x)$ describes how probability is **spread out** (distributed) across values.

The PDF does **not** return a probability.

Evaluating the PDF at a point gives a **density** – probability per unit length.

Think of it like **population density**:

- "10,000 people per square mile" isn't a count – it's a density
- To get an actual count, you must multiply by an area

1.5 PDF vs PMF: Key Differences

PMF (Discrete)

$p_X(x) = P(X = x)$

Direct probability at each point

Sum to 1:

$$\sum_x p_X(x) = 1$$

PDF (Continuous)

$f_X(x) =$ "density" at x

Not a probability!

Integrate to 1:

$$\int_{-\infty}^{\infty} f_X(x)dx = 1$$

1.5 PDF vs PMF: Key Differences

PMF (Discrete)

$p_X(x) = P(X = x)$

Direct probability at each point

Sum to 1:

$$\sum_x p_X(x) = 1$$

PDF (Continuous)

$f_X(x) =$ "density" at x

Not a probability!

Integrate to 1:

$$\int_{-\infty}^{\infty} f_X(x)dx = 1$$

Note: A PDF can have values > 1! It's not a probability.

1.6 Intuitive Meaning of Density

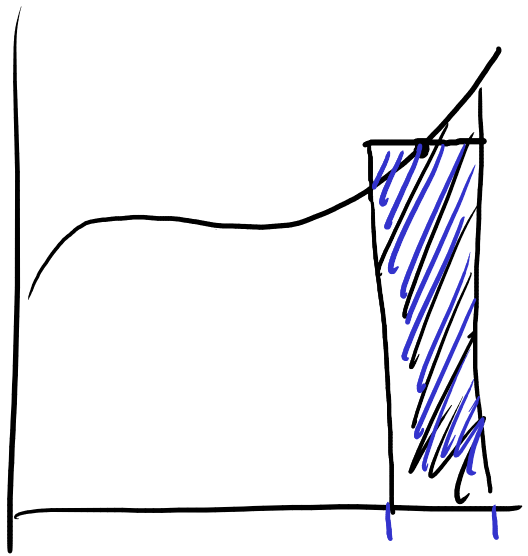
Recall: Think of $f_X(x)$ as **"probability mass per unit length"** around x.

1.6 Intuitive Meaning of Density

Recall: Think of $f_X(x)$ as **"probability mass per unit length"** around x.

If δ is very small:

$$P(x \leq X \leq x + \delta) \approx f_X(x) \cdot \delta$$



1.6 Intuitive Meaning of Density

Recall: Think of $f_X(x)$ as **“probability mass per unit length”** around x.

If δ is very small:

$$P(x \leq X \leq x + \delta) \approx f_X(x) \cdot \delta$$

Key insight:

- Regions with **high density** have **high probabilities**
- Regions with **low density** have **low probabilities**

1.6 Intuitive Meaning of Density

Recall: Think of $f_X(x)$ as **"probability mass per unit length"** around x.

If δ is very small:

$$P(x \leq X \leq x + \delta) \approx f_X(x) \cdot \delta$$

Key insight:

- Regions with **high density** have **high probabilities**
- Regions with **low density** have **low probabilities**

Even though $f_X(x)$ isn't a probability, it tells us which values are common vs rare.

1.7 Getting Probability from Density

To get a probability from a PDF, we **integrate** over a region:

Definition: Probability from PDF

$$P(a \leq X \leq b) = \int_a^b f_X(x)dx$$

1.7 Getting Probability from Density

To get a probability from a PDF, we **integrate** over a region:

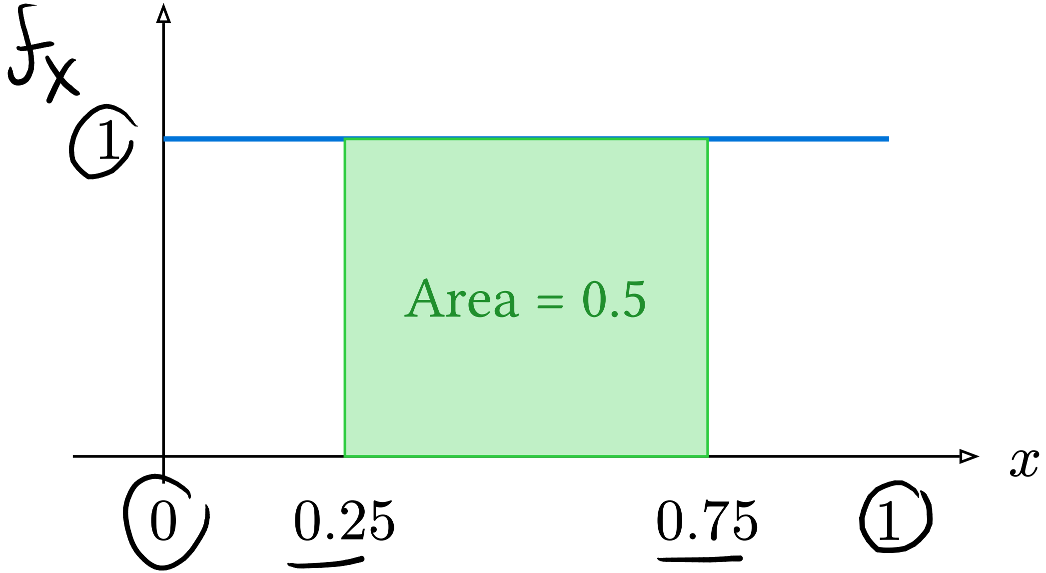
Definition: Probability from PDF

$$P(a \leq X \leq b) = \int_a^b f_X(x)dx$$

1.8 Example: Marble in a Drawer

Example: For our marble in a 1-meter drawer:

- PDF: $f_X(x) = 1$ for $x \in [0, 1]$
- $P(0.25 \leq X \leq 0.75) = \int_{0.25}^{0.75} 1 dx = 0.5$



2. The Continuous Uniform Distribution

2.1 Another Example: Wheel of Fortune

A gambler spins a wheel of fortune, continuously calibrated between 0 and 1.

2.1 Another Example: Wheel of Fortune

A gambler spins a wheel of fortune, continuously calibrated between 0 and 1.

Question: What's the probability the result is between 0.3 and 0.7?

2.1 Another Example: Wheel of Fortune

A gambler spins a wheel of fortune, continuously calibrated between 0 and 1.

Question: What's the probability the result is between 0.3 and 0.7?

Assuming all subintervals of $[0, 1]$ of the same length are equally likely:

$$P(0.3 \leq X \leq 0.7) = 0.7 - 0.3 = 0.4$$

2.1 Another Example: Wheel of Fortune

A gambler spins a wheel of fortune, continuously calibrated between 0 and 1.

Question: What's the probability the result is between 0.3 and 0.7?

Assuming all subintervals of $[0, 1]$ of the same length are equally likely:

$$P(0.3 \leq X \leq 0.7) = 0.7 - 0.3 = 0.4$$

This is the **continuous uniform distribution** in action.

2.2 Motivating the Uniform

Our marble example is the **continuous uniform distribution**:

2.2 Motivating the Uniform

Our marble example is the **continuous uniform distribution**:

Definition: Continuous Uniform Distribution

$X \sim \text{Uniform}(a, b)$ means X is equally likely to take any value in $[a, b]$.

PDF:

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{if } a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}$$

2.2 Motivating the Uniform

Our marble example is the **continuous uniform distribution**:

Definition: Continuous Uniform Distribution

$X \sim \text{Uniform}(a, b)$ means X is equally likely to take any value in $[a, b]$.

PDF:

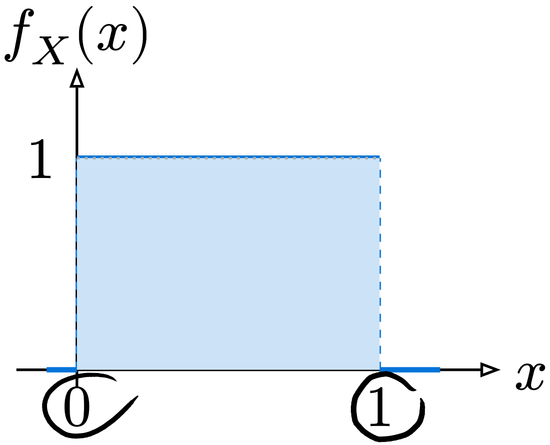
$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{if } a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}$$

Why $\frac{1}{b-a}$? So the area under the PDF equals 1:

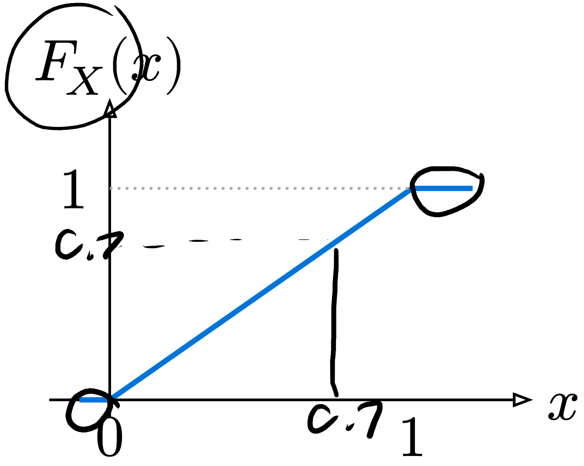
$$\int_a^b \frac{1}{b-a} dx = \frac{b-a}{b-a} = 1 \quad \checkmark$$

2.3 Visualizing Uniform PDF

For $X \sim \text{Uniform}(0, 1)$:



PDF: height = 1, area = 1

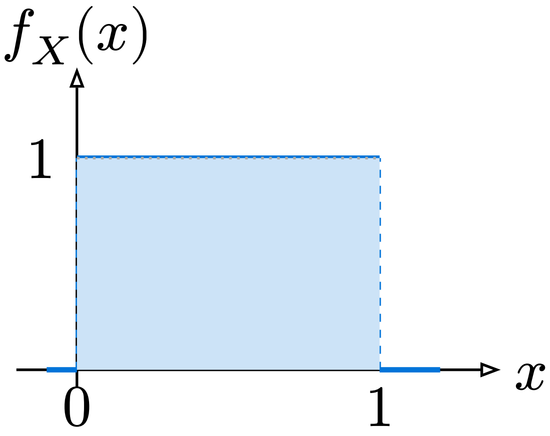


CDF: straight line, slope = 1

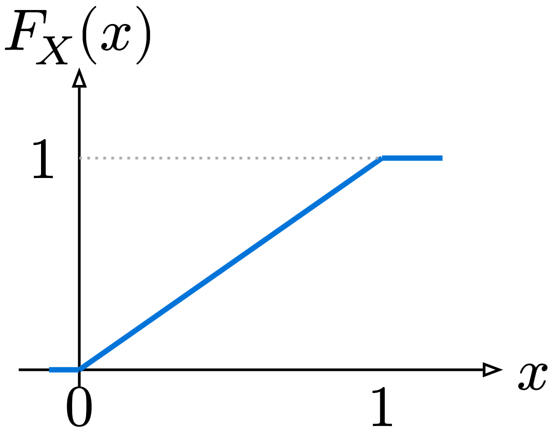
$F_X = CDF(k) = P(X \leq k)$

2.3 Visualizing Uniform PDF

For $X \sim \text{Uniform}(0, 1)$:



PDF: height = 1, area = 1



CDF: straight line, slope = 1

The uniform distribution is the **simplest** continuous distribution.

3. PDF to CDF and Back

3.1 The Fundamental Relationship

PDF and CDF are two views of the same distribution:

3.1 The Fundamental Relationship

PDF and CDF are two views of the same distribution:

CDF from PDF:

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt$$

Integration: area under PDF up to x

PDF from CDF:

$$f_X(x) = \frac{d}{dx}F_X(x)$$

Differentiation: slope of CDF at x

3.1 The Fundamental Relationship

PDF and CDF are two views of the same distribution:

CDF from PDF:

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt$$

Integration: area under PDF up to x

PDF from CDF:

$$f_X(x) = \frac{d}{dx}F_X(x)$$

Differentiation: slope of CDF at x

Note: CDF is the "running total" of probability density.

3.2 Computing Probabilities via CDF

Using the CDF is often easier than integrating the PDF:

Definition: Probability via CDF Assume $b > a$

$$P(a \leq X \leq b) = F_X(b) - F_X(a)$$

3.2 Computing Probabilities via CDF

Using the CDF is often easier than integrating the PDF:

Definition: Probability via CDF

$$P(a \leq X \leq b) = F_X(b) - F_X(a)$$

Why this works:

$$F_X(b) - F_X(a) = \int_{-\infty}^{b} f_X(x)dx - \int_{-\infty}^{a} f_X(x)dx = \int_{a}^{b} f_X(x)dx$$

3.2 Computing Probabilities via CDF

Using the CDF is often easier than integrating the PDF:

Definition: Probability via CDF

$$P(a \leq X \leq b) = F_X(b) - F_X(a)$$

Why this works:

$$F_X(b) - F_X(a) = \int_{-\infty}^{b} f_X(x)dx - \int_{-\infty}^{a} f_X(x)dx = \int_{a}^{b} f_X(x)dx$$

Example: Bus Wait

$X \sim \text{Uniform}(0, 15)$. What's $P(3 \leq X \leq 10)$?

$P(3 \leq X \leq 10) = F_X(10) - F_X(3) = \frac{10}{15} - \frac{3}{15} = \frac{7}{15}$

3.3 Uniform CDF

Definition: Uniform CDF

For $X \sim \text{Uniform}(a, b)$:

$$F_X(x) = P(X \leq x) = \begin{cases} 0 & \text{if } x < a \\ \frac{x-a}{b-a} & \text{if } a \leq x \leq b \\ 1 & \text{if } x > b \end{cases}$$

3.3 Uniform CDF

Definition: Uniform CDF

For $X \sim \text{Uniform}(a, b)$:

$$F_X(x) = P(X \le x) = \begin{cases} 0 & \text{if } x < a \\ \frac{x-a}{b-a} & \text{if } a \le x \le b \\ 1 & \text{if } x > b \end{cases}$$

Derivation:

$$F_X(x) = \int_a^x \frac{1}{b-a} dt = \frac{x-a}{b-a}$$

3.3 Uniform CDF

Definition: Uniform CDF

For $X \sim \text{Uniform}(a, b)$:

$$F_X(x) = P(X \leq x) = \begin{cases} 0 & \text{if } x < a \\ \frac{x-a}{b-a} & \text{if } a \leq x \leq b \\ 1 & \text{if } x > b \end{cases}$$

Derivation:

$$F_X(x) = \int_a^x \frac{1}{b-a} dt = \frac{x-a}{b-a}$$

Example: For $X \sim \text{Uniform}(0, 1)$, $F_X(0.7) = \frac{0.7-0}{1-0} = 0.7$

3.4 Computing Probabilities

For $X \sim \text{Uniform}(a, b)$, probability = fraction of the interval:

$$P(c \leq X \leq d) = \frac{d - c}{b - a}$$

3.4 Computing Probabilities

For $X \sim \text{Uniform}(a, b)$, probability = fraction of the interval:

$$P(c \le X \le d) = \frac{d - c}{b - a}$$

Example: Marble in Drawer

$X \sim \text{Uniform}(0, 1)$.

$P(0.2 \le X \le 0.8) =$

3.4 Computing Probabilities

For $X \sim \text{Uniform}(a, b)$, probability = fraction of the interval:

$$P(c \leq X \leq d) = \frac{d - c}{b - a}$$

Example: Marble in Drawer

$X \sim \text{Uniform}(0, 1)$.

$P(0.2 \leq X \leq 0.8) = \frac{0.8-0.2}{1-0}$

3.4 Computing Probabilities

For $X \sim \text{Uniform}(a, b)$, probability = fraction of the interval:

$$P(c \leq X \leq d) = \frac{d - c}{b - a}$$

Example: Marble in Drawer

$X \sim \text{Uniform}(0, 1)$.

$P(0.2 \leq X \leq 0.8) = \frac{0.8 - 0.2}{1 - 0} = 0.6$

3.5 Another Example: Waiting Time

3.5 Another Example: Waiting Time

Example: Waiting Time

A bus arrives uniformly at random between 0 and 15 minutes.

$X \sim \text{Uniform}(0, 15).$

$P(X \leq 5) = \frac{5-0}{15-0} = \frac{1}{3}$

3.6 Expected Value for Continuous RVs

For continuous RVs, replace sums with integrals:

$$E(X) = \sum_{x \in X} x P(X = x)$$

Definition: Expected Value (Continuous)

$$E(X) = \int_{-\infty}^{\infty} x f_X(x) dx$$

PDF

3.6 Expected Value for Continuous RVs

For continuous RVs, replace sums with integrals:

Definition: Expected Value (Continuous)

$$E(X) = \int_{-\infty}^{\infty} x \cdot f_X(x) dx$$

For Uniform(a, b):

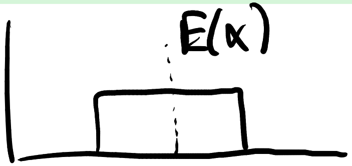
$$E(X) = \int_a^b x \cdot \frac{1}{b-a} dx = \frac{1}{b-a} \cdot \left[\frac{x^2}{2}\right]_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{a+b}{2}$$

3.6 Expected Value for Continuous RVs

For continuous RVs, replace sums with integrals:

Definition: Expected Value (Continuous)

$$E(X) = \int_{-\infty}^{\infty} x \cdot f_X(x) dx$$



For Uniform(a, b):

$$E(X) = \int_a^b x \cdot \frac{1}{b-a} dx = \frac{1}{b-a} \cdot \left[\frac{x^2}{2}\right]_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{a+b}{2}$$

Note: The expected value is the midpoint – exactly what we'd expect for a symmetric distribution!

3.7 Variance for Continuous RVs

Definition: Variance (Continuous)

$$\mathrm{Var}(X) = E(X^2) - (E(X))^2 = \int_{-\infty}^{\infty} x^2 \cdot f_X(x)dx - \mu^2$$

3.7 Variance for Continuous RVs

Definition: Variance (Continuous)

$$\text{Var}(X) = E(X^2) - (E(X))^2 = \int_{-\infty}^{\infty} x^2 \cdot f_X(x)dx - \mu^2$$

For Uniform(a, b):

$$E(X^2) = \int_a^b x^2 \cdot \frac{1}{b-a}dx = \frac{b^3 - a^3}{3(b-a)}$$

3.7 Variance for Continuous RVs

Definition: Variance (Continuous)

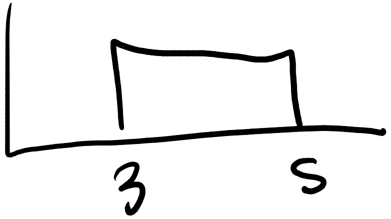
$$\mathrm{Var}(X) = E(X^2) - (E(X))^2 = \int_{-\infty}^{\infty} x^2 \cdot f_X(x)dx - \mu^2$$

For Uniform(a, b):

$$E(X^2) = \int_a^b x^2 \cdot \frac{1}{b-a}dx = \frac{b^3 - a^3}{3(b-a)}$$

After algebra:

$$\mathrm{Var}(X) = \frac{(b-a)^2}{12}$$



$$\frac{(5-3)^2}{12}$$

3.8 Uniform Distribution Summary

Definition: Uniform Properties

For $X \sim \text{Uniform}(a, b)$:

Expected Value:

$$E(X) = \frac{a+b}{2}$$

Variance:

$$\text{Var}(X) = \frac{(b-a)^2}{12}$$

3.8 Uniform Distribution Summary

Definition: Uniform Properties

For $X \sim \text{Uniform}(a, b)$:

Expected Value:

$$E(X) = \frac{a+b}{2}$$

Variance:

$$\text{Var}(X) = \frac{(b-a)^2}{12}$$

Example: $X \sim \text{Uniform}(0, 1)$

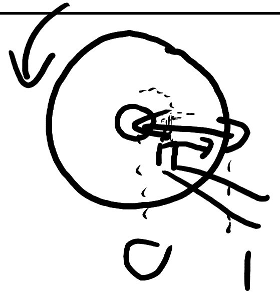
- $E(X) = \frac{0+1}{2} = 0.5$
- $\text{Var}(X) = \frac{(1-0)^2}{12} = \frac{1}{12} \approx 0.083$
- $\sigma = \frac{1}{\sqrt{12}} \approx 0.289$

3.9 Example: Disk Seek Time

Example: Disk Performance

A disk has read/write head position X modeled as $\text{Uniform}(0, 1)$.

The head moves between positions for consecutive reads X_1 and X_2.



Var(seek time)

3.9 Example: Disk Seek Time

Example: Disk Performance

A disk has read/write head position X modeled as Uniform$(0, 1)$.

The head moves between positions for consecutive reads X_1 and X_2.

If X_1, X_2 are independent:

$$\mathrm{Var}(X_1 + X_2) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2) = \frac{1}{12} + \frac{1}{12} = \frac{1}{6}$$

3.9 Example: Disk Seek Time

Example: Disk Performance

A disk has read/write head position X modeled as $\text{Uniform}(0, 1)$.

The head moves between positions for consecutive reads X_1 and X_2.

If X_1, X_2 are independent:

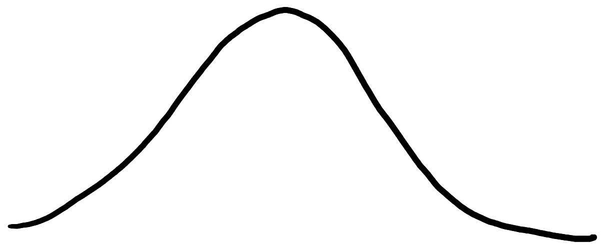
$$\text{Var}(X_1 + X_2) = \text{Var}(X_1) + \text{Var}(X_2) = \frac{1}{12} + \frac{1}{12} = \frac{1}{6}$$

Note: Uniform distributions also help detect **denial-of-service attacks**: attackers often use uniformly distributed fake IP addresses–a pattern not seen in normal traffic.

4. Other Continuous Distributions

4.1 The Normal (Gaussian) Distribution

The famous “bell curve” – central to statistics.



4.1 The Normal (Gaussian) Distribution

The famous “bell curve” – central to statistics.

Definition: Normal Distribution

$X \sim N(\mu, \sigma^2)$ has PDF:

mean / expected val — variance

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad -\infty < x < \infty$$

4.1 The Normal (Gaussian) Distribution

The famous "bell curve" – central to statistics.

Definition: Normal Distribution

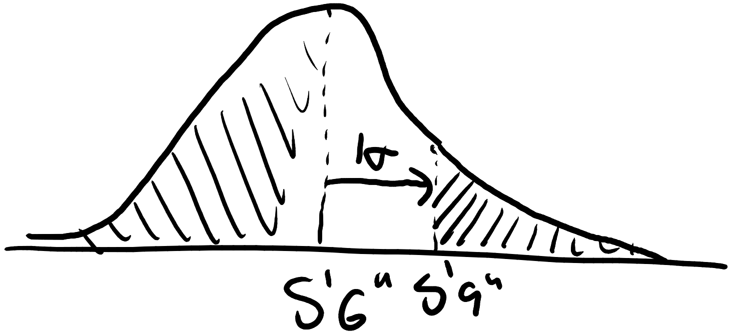
$X \sim N(\mu, \sigma^2)$ has PDF:

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad -\infty < x < \infty$$

Parameters:

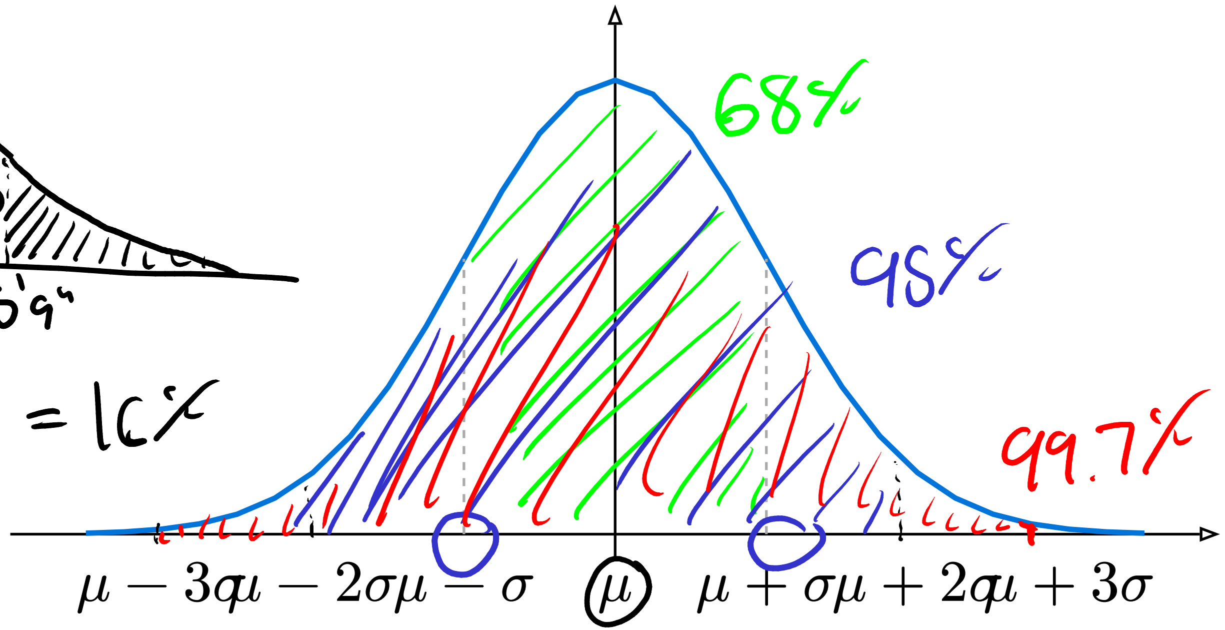
- μ: mean (center of the bell)
- σ^2: variance (spread of the bell)
- σ: standard deviation

4.2 Visualizing the Normal Distribution



$P(X > 5'9") = 16\%$

$1 - (.5 + .34)$
$= .16$



68-95-99.7 Rule: About 68% within $\mu \pm \sigma$, 95% within $\mu \pm 2\sigma$, 99.7% within $\mu \pm 3\sigma$

$\sigma = 3"$

5'6"

$\mu = 5'6"$
$\sigma = 3"$

$P(X > 5'9")$?

4.3 Normal Distribution Properties

$N(\mu, \sigma)$

Definition: Normal Properties

For $X \sim N(\mu, \sigma^2)$:

- **Mean:** $E(X) = \mu$
- **Variance:** $\mathrm{Var}(X) = \sigma^2$
- **Standard Deviation:** σ

4.3 Normal Distribution Properties

Definition: Normal Properties

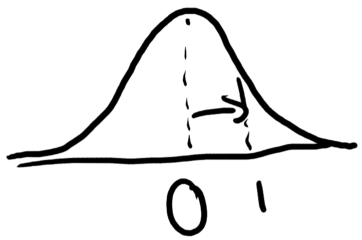
For $X \sim N(\mu, \sigma^2)$:

- **Mean:** $E(X) = \mu$
- **Variance:** $\mathrm{Var}(X) = \sigma^2$
- **Standard Deviation:** σ

Note: Interestingly, sums of many independent RVs tend to be normal. We'll come back to this later.

4.4 Standardization

Any normal random variable can be transformed to the **standard normal** $N(0, 1)$:



4.4 Standardization

4.4 Standardization

Any normal random variable can be transformed to the **standard normal** $N(0,1)$:

Definition: Standardization

If $X \sim N(\mu, \sigma^2)$, then:

$$Z = \frac{X - \mu}{\sigma} \sim N(0,1)$$

4.4 Standardization

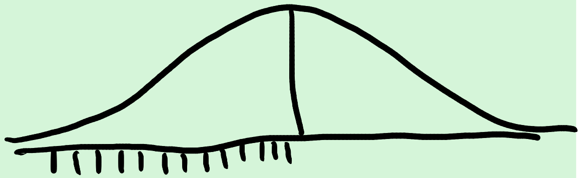
4.4 Standardization

Any normal random variable can be transformed to the **standard normal** $N(0, 1)$:

Definition: Standardization

If $X \sim N(\mu, \sigma^2)$, then:

$$Z = \frac{X - \mu}{\sigma} \sim N(0, 1)$$



Why is this useful?

Z-table

- Tables and software often give probabilities for $N(0, 1)$
- To find $P(X \leq x)$, convert to $P\left(Z \leq \frac{x-\mu}{\sigma}\right)$

4.4 Standardization

4.4 Standardization

Any normal random variable can be transformed to the **standard normal** $N(0, 1)$:

Definition: Standardization

If $X \sim N(\mu, \sigma^2)$, then:

$$Z = \frac{X - \mu}{\sigma} \sim N(0, 1)$$

Why is this useful?

- Tables and software often give probabilities for $N(0, 1)$
- To find $P(X \leq x)$, convert to $P\left(Z \leq \frac{x-\mu}{\sigma}\right)$

4.4 Standardization

Example: Standardization

Adult male heights: $X \sim N(70, 3^2)$ inches. What is $P(X \leq 73)$?

$P(X \leq 73) = P\left(Z \leq \frac{73-70}{3}\right) = P(Z \leq 1) \approx 0.84$

4.5 The Exponential Distribution

Models **time until an event** (e.g., time until next customer arrives).

4.5 The Exponential Distribution

Models **time until an event** (e.g., time until next customer arrives).

Definition: Exponential Distribution

$X \sim \text{Exp}(\lambda)$ has PDF:

$$f_X(x) = \lambda e^{-\lambda x}, \quad x > 0$$

where $\lambda > 0$ is the **rate** parameter.

4.5 The Exponential Distribution

Models **time until an event** (e.g., time until next customer arrives).

Definition: Exponential Distribution

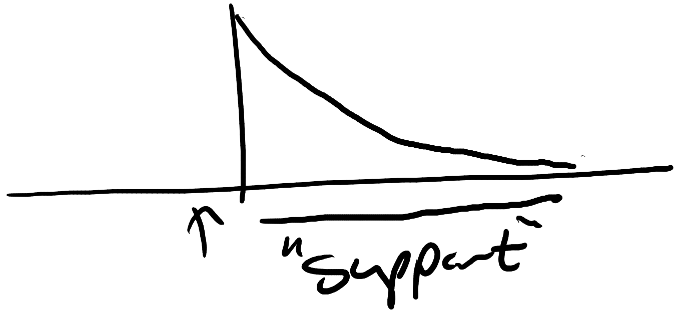
$X \sim \text{Exp}(\lambda)$ has PDF:

$$f_X(x) = \lambda e^{-\lambda x}, \quad x > 0$$

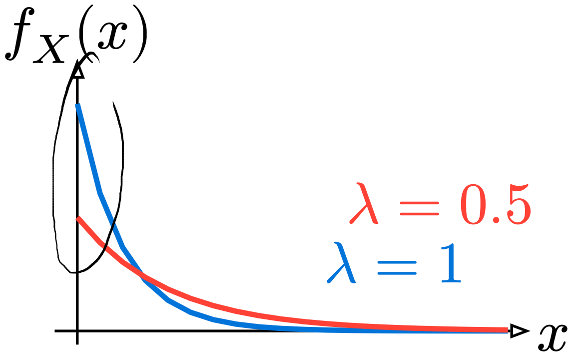
where $\lambda > 0$ is the **rate** parameter.

Properties:

- $E(X) = \frac{1}{\lambda}$ (mean time until event)
- $\text{Var}(X) = \frac{1}{\lambda^2}$
- Support: $[0, \infty)$



4.6 Visualizing Exponential Distributions



Higher $\lambda \longrightarrow$ faster decay $\longrightarrow$ shorter expected time

4.7 Exponential: The Memoryless Property

Definition: Memoryless Property

For $X \sim \mathrm{Exp}(\lambda)$:

$$P(X > s + t \mid X > s) = P(X > t)$$

4.7 Exponential: The Memoryless Property

Definition: Memoryless Property

For $X \sim \text{Exp}(\lambda)$:

$$P(X > s + t \mid X > s) = P(X > t)$$

Sound familiar? This is just like the geometric distribution!

4.7 Exponential: The Memoryless Property

Definition: Memoryless Property

For $X \sim \text{Exp}(\lambda)$:

$$P(X > s + t \mid X > s) = P(X > t)$$

Sound familiar? This is just like the geometric distribution!

Geometric (discrete)	**Exponential** (continuous)
Trials until first success	Time until first event
Memoryless	Memoryless

4.7 Exponential: The Memoryless Property

Definition: Memoryless Property

For $X \sim \text{Exp}(\lambda)$:

$$P(X > s + t \mid X > s) = P(X > t)$$

Sound familiar? This is just like the geometric distribution!

Geometric (discrete)	**Exponential** (continuous)
Trials until first success	Time until first event
Memoryless	Memoryless

Note: Exponential is the **continuous analog** of the geometric distribution!

4.8 Exponential: Applications

Common applications:

- Time until next customer arrives at a bank
- Lifetime of electronic components
- Time between network packet arrivals
- Interarrival times in queuing systems

4.8 Exponential: Applications

Common applications:

- Time until next customer arrives at a bank
- Lifetime of electronic components
- Time between network packet arrivals
- Interarrival times in queuing systems

Example: Component Lifetime

A component has lifetime $L \sim \text{Exp}(\lambda)$ with mean 10,000 hours.

So $\lambda = \frac{1}{10000}$.

$P(L < 500) = 1 - e^{-\frac{500}{10000}} = 1 - e^{-0.05} \approx 0.049$

4.9 The Gamma Distribution

What if we wait for multiple events?

4.9 The Gamma Distribution

What if we wait for multiple events?

Definition: Gamma Distribution

If $X_1, ..., X_r$ are independent $\text{Exp}(\lambda)$ random variables, then:

$$T_r = X_1 + ... + X_r \sim \text{Gamma}(r, \lambda)$$

PDF: $f_X(x) = \frac{\lambda^r x^{r-1} e^{-\lambda x}}{\Gamma(r)}, \quad x > 0$

4.9 The Gamma Distribution

What if we wait for multiple events?

Definition: Gamma Distribution

If $X_1, \ldots, X_r$ are independent $\text{Exp}(\lambda)$ random variables, then:

$$T_r = X_1 + \ldots + X_r \sim \text{Gamma}(r, \lambda)$$

PDF: $f_X(x) = \frac{\lambda^r x^{r-1} e^{-\lambda x}}{\Gamma(r)}, \quad x > 0$

Parameters:

- r: shape parameter (number of events)
- λ: rate parameter

4.10 Gamma Distribution Properties

Definition: Gamma Properties

For $X \sim \text{Gamma}(r, \lambda)$:

- $E(X) = \frac{r}{\lambda}$
- $\text{Var}(X) = \frac{r}{\lambda^2}$

4.10 Gamma Distribution Properties

Definition: Gamma Properties

For $X \sim \text{Gamma}(r, \lambda)$:

- $E(X) = \frac{r}{\lambda}$
- $\text{Var}(X) = \frac{r}{\lambda^2}$

Special case: $\text{Gamma}(1, \lambda) = \text{Exp}(\lambda)$

4.10 Gamma Distribution Properties

Definition: Gamma Properties

For $X \sim \text{Gamma}(r, \lambda)$:

- $E(X) = \frac{r}{\lambda}$
- $\text{Var}(X) = \frac{r}{\lambda^2}$

Special case: $\text{Gamma}(1, \lambda) = \text{Exp}(\lambda)$

Example: Network Buffer

A node waits until 5 messages arrive before transmitting.

Interarrival times are $\text{Exp}(\lambda = 0.01)$ (mean 100 ms).

Time until transmission: $T \sim \text{Gamma}(5, 0.01)$

$E(T) = \frac{5}{0.01} = 500$ ms

4.11 The Beta Distribution

For modeling **proportions** and **probabilities** (values in $(0, 1)$).

4.11 The Beta Distribution

For modeling **proportions** and **probabilities** (values in $(0, 1)$).

Definition: Beta Distribution

$X \sim \text{Beta}(\alpha, \beta)$ has PDF:

$$f_X(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \quad 0 < x < 1$$

4.11 The Beta Distribution

For modeling **proportions** and **probabilities** (values in $(0, 1)$).

Definition: Beta Distribution

$X \sim \text{Beta}(\alpha, \beta)$ has PDF:

$$f_X(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \quad 0 < x < 1$$

Properties:

- $E(X) = \frac{\alpha}{\alpha+\beta}$
- Support: $(0, 1)$
- Very flexible shape (U-shaped, uniform, bell-shaped, skewed)

4.11 The Beta Distribution

For modeling **proportions** and **probabilities** (values in $(0,1)$).

Definition: Beta Distribution

$X \sim \text{Beta}(\alpha, \beta)$ has PDF:

$$f_X(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \quad 0 < x < 1$$

Properties:

- $E(X) = \frac{\alpha}{\alpha+\beta}$
- Support: $(0,1)$
- Very flexible shape (U-shaped, uniform, bell-shaped, skewed)

Note: Special case: $\text{Beta}(1,1) = \text{Uniform}(0,1)$

4.12 Summary: Continuous Distribution Families

Distribution	Support	Mean	Use Case
$\text{Uniform}(a, b)$	$[a, b]$	$\frac{a+b}{2}$	Equally likely outcomes
$N(\mu, \sigma^2)$	$(-\infty, \infty)$	μ	Bell curves, CLT
$\text{Exp}(\lambda)$	$[0, \infty)$	$1/\lambda$	Time until event
$\text{Gamma}(r, \lambda)$	$[0, \infty)$	r/λ	Sum of exponentials
$\text{Beta}(\alpha, \beta)$	$(0, 1)$	$\alpha/(\alpha + \beta)$	Proportions

4.13 Discrete vs Continuous: Boundaries

A subtle but important point:

4.13 Discrete vs Continuous: Boundaries

A subtle but important point:

Discrete: Boundary points matter!

$P(X \leq 3) \neq P(X < 3)$ in general

4.13 Discrete vs Continuous: Boundaries

A subtle but important point:

Discrete: Boundary points matter!

$P(X \leq 3) \neq P(X < 3)$ in general

Continuous: Boundary points don't matter!

$P(X \leq 3) = P(X < 3)$ because $P(X = 3) = 0$

4.13 Discrete vs Continuous: Boundaries

A subtle but important point:

Discrete: Boundary points matter!

$P(X \leq 3) \neq P(X < 3)$ in general

Continuous: Boundary points don't matter!

$P(X \leq 3) = P(X < 3)$ because $P(X = 3) = 0$

So for continuous RVs:

- $P(a \leq X \leq b) = P(a < X < b) = P(a \leq X < b) = P(a < X \leq b)$

All are equal because single points have probability zero.

5. Computing with Code

5.1 Python: scipy.stats.uniform

```python
from scipy.stats import uniform

# Uniform(0, 1) - note: scipy uses loc, scale parametrization
# Uniform(a, b) = uniform(loc=a, scale=b-a)
X = uniform(loc=0, scale=1)  # Uniform(0, 1)

# PDF at x = 0.5
X.pdf(0.5)  # 1.0

# CDF: P(X <= 0.7)
X.cdf(0.7)  # 0.7

# Inverse CDF (quantile): value x where P(X <= x) = 0.9
X.ppf(0.9)  # 0.9

# Generate 5 random samples
X.rvs(size=5)  # e.g., [0.42, 0.78, 0.15, 0.91, 0.33]
```

5.2 Python: Uniform(5, 15) Example

```
from scipy.stats import uniform

# Uniform(5, 15): loc=5, scale=15-5=10
X = uniform(loc=5, scale=10)

# P(X > 12) = 1 - P(X <= 12)
1 - X.cdf(12)  # 0.3

# P(3 <= X <= 10) = P(X <= 10) - P(X <= 3)
# Note: 3 < 5, so P(X <= 3) = 0
X.cdf(10) - X.cdf(3)  # 0.5

# Expected value
X.mean()  # 10.0

# Standard deviation
X.std()   # 2.887 (= sqrt(100/12))
```

5.3 R: dunif, punif, qunif, runif

```r
# Uniform(0, 1) - R uses min, max parametrization
# dunif = density (PDF), punif = CDF, qunif = quantile, runif = random

# PDF at x = 0.5
dunif(0.5, min = 0, max = 1)  # 1

# CDF: P(X <= 0.7)
punif(0.7, min = 0, max = 1)  # 0.7

# Quantile: value x where P(X <= x) = 0.9
qunif(0.9, min = 0, max = 1)  # 0.9

# Generate 5 random samples
runif(5, min = 0, max = 1)  # e.g., 0.42, 0.78, ...
```

5.4 R: Uniform(5, 15) Example

```r
# Uniform(5, 15)
a <- 5
b <- 15

# P(X > 12) = 1 - P(X <= 12)
1 - punif(12, min = a, max = b)  # 0.3

# P(7 <= X <= 10)
punif(10, min = a, max = b) - punif(7, min = a, max = b)  # 0.3

# Expected value (manually)
(a + b) / 2  # 10

# Variance (manually)
(b - a)^2 / 12  # 8.333
```

5.5 Connecting Discrete and Continuous

Discrete

- PMF: $P(X = x)$
- Sum to 1
- Probability at points
- CDF has jumps
- Examples: Bernoulli, Binomial, Geometric, Poisson

Continuous

- PDF: density at x
- Integrate to 1
- Probability over intervals
- CDF is smooth
- Examples: Uniform, (next: Exponential, Normal)

5.5 Connecting Discrete and Continuous

Discrete

- PMF: $P(X = x)$
- Sum to 1
- Probability at points
- CDF has jumps
- Examples: Bernoulli, Binomial, Geometric, Poisson

Continuous

- PDF: density at x
- Integrate to 1
- Probability over intervals
- CDF is smooth
- Examples: Uniform, (next: Exponential, Normal)

Note: Same concepts (expectation, variance, CDF) apply to both – just replace sums with integrals!

5.6 Recap

Today we covered:

- Continuous RVs: $P(X = x) = 0$ for exact values; use density instead
- PDF gives density, not probability; integrate to get probability
- Uniform(a, b): flat density $\frac{1}{b-a}$; $E(X) = \frac{a+b}{2}$
- CDF: $F(x) = P(X \leq x)$; $P(a \leq X \leq b) = F(b) - F(a)$
- Sand analogy: discrete = buckets, continuous = smooth piles