Continuous Probability Distributions

Lecture 24, CS70 Summer 2025

Framing: A Random Variable with Range [0, 2]

Imagine that we choose a real number X, uniformly at random from $[0, 2]$

What is the probability that $X = 0.3$?

- Zero! There's uncountably infinite numbers we could have chosen.

What is $P(X \leq 1/2)$?

- 25%. One quarter of the possibilities are in this range.

What is $P(X \in [0.7, 0.9])$?

- 10%, because one tenth of the possibilities are in this range.

Framing: A Random Variable with Range [0, 2]

Imagine that we choose a real number X, uniformly at random from $[0, 2]$

For other $0 \leq a \leq b \leq 2$: What is $P(X \in [a, b])$?

- Following the pattern from $P(X \in [0.7, 0.9])$:

$$P(X \in [a, b]) = \frac{b - a}{2}$$

$P(X = 0.3)$	0
$P(X \leq 1/2)$	25%
$P(X \in [0.7, 0.9])$	10%

Framing: A Random Variable with Range [0, 2]

Imagine that we choose a real number X, uniformly at random from $[0, 2]$

For other $0 \leq a \leq b \leq 2$: What is $P(X \in [a, b])$?

- Following the pattern from $P(X \in [0.7, 0.9])$:

$$P(X \in [a, b]) = \frac{b - a}{2}$$

$P(X = 0.3)$	0
$P(X \leq 1/2)$	25%
$P(X \in [0.7, 0.9])$	10%
$P(X \in [a, b])$	$\frac{b - a}{2}$

Review: Distributions of Discrete Random Variables

For a discrete random variable, we have the notion of the distribution of a random variable.

- Enumeration of every value and its probability.

The function that maps each value to its probability is called the Probability Mass Function.

Examples:

- The PMF of rolling two six-sided dice is to the right.
- The PMF of $X \sim \text{Binomial}(n, p)$ is $\binom{n}{i} p^i (1-p)^{n-i}$
- The PMF of $X \sim \text{Geometric}(p)$ is $(1-p)^{i-1} p$

PMF is just a new term for something we've seen before.

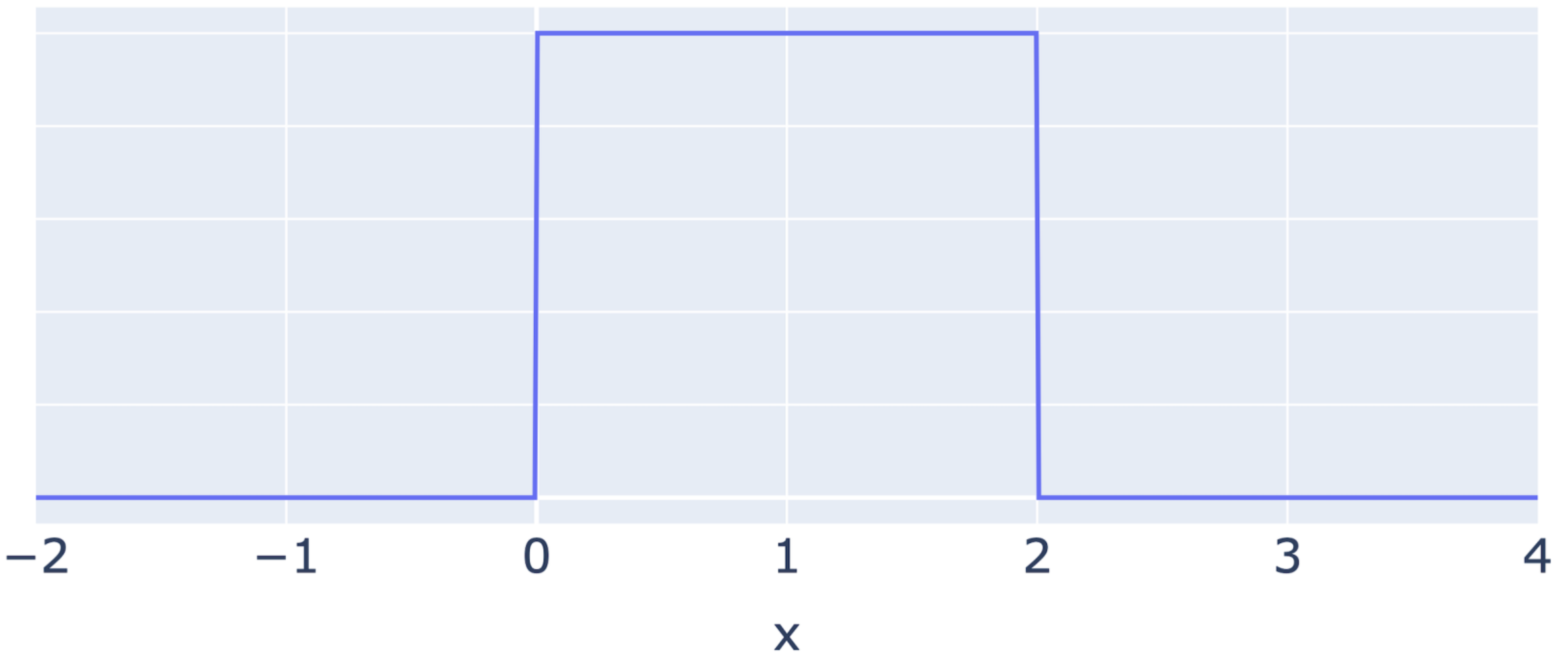
a	$P(X = a)$
2	1/36
3	2/36
4	3/36
5	4/36
6	5/36
7	6/36
8	5/36
9	4/36
10	3/36
11	2/36
12	1/36

Framing: A Random Variable with Range [0, 2]

Imagine that we choose a real number X, uniformly at random from $[0, 2]$

- For this random variable, the probability mass function is not useful. $P(X = 0.3) = 0$, and in general $P(X = x) = 0$ for all x.

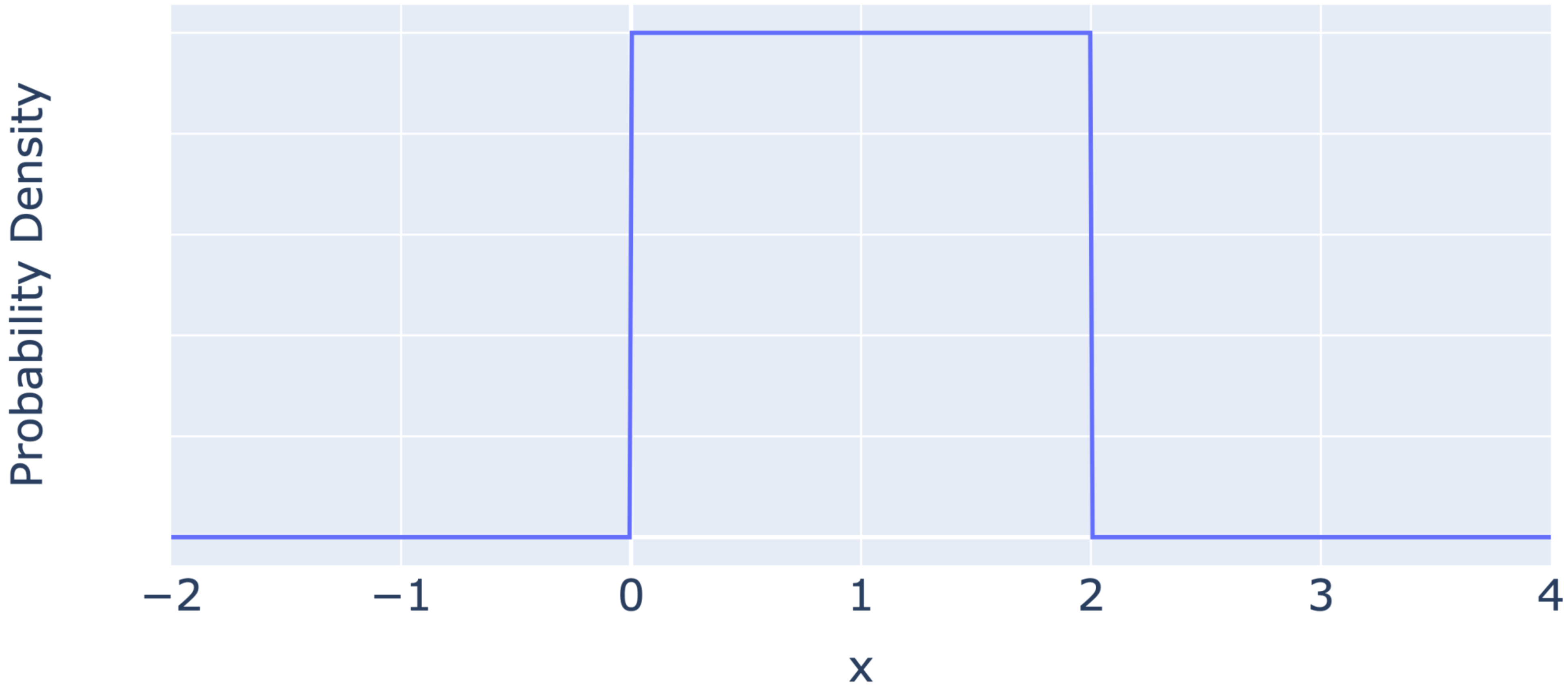
There is, however, a useful analogue called the "probability density function", that looks like this:



Framing: A Random Variable with Range [0, 2]

Imagine that we choose a real number X, uniformly at random from $[0, 2]$

- The "probability mass function" is not useful, always zero.
- The "probability density function" looks like this:



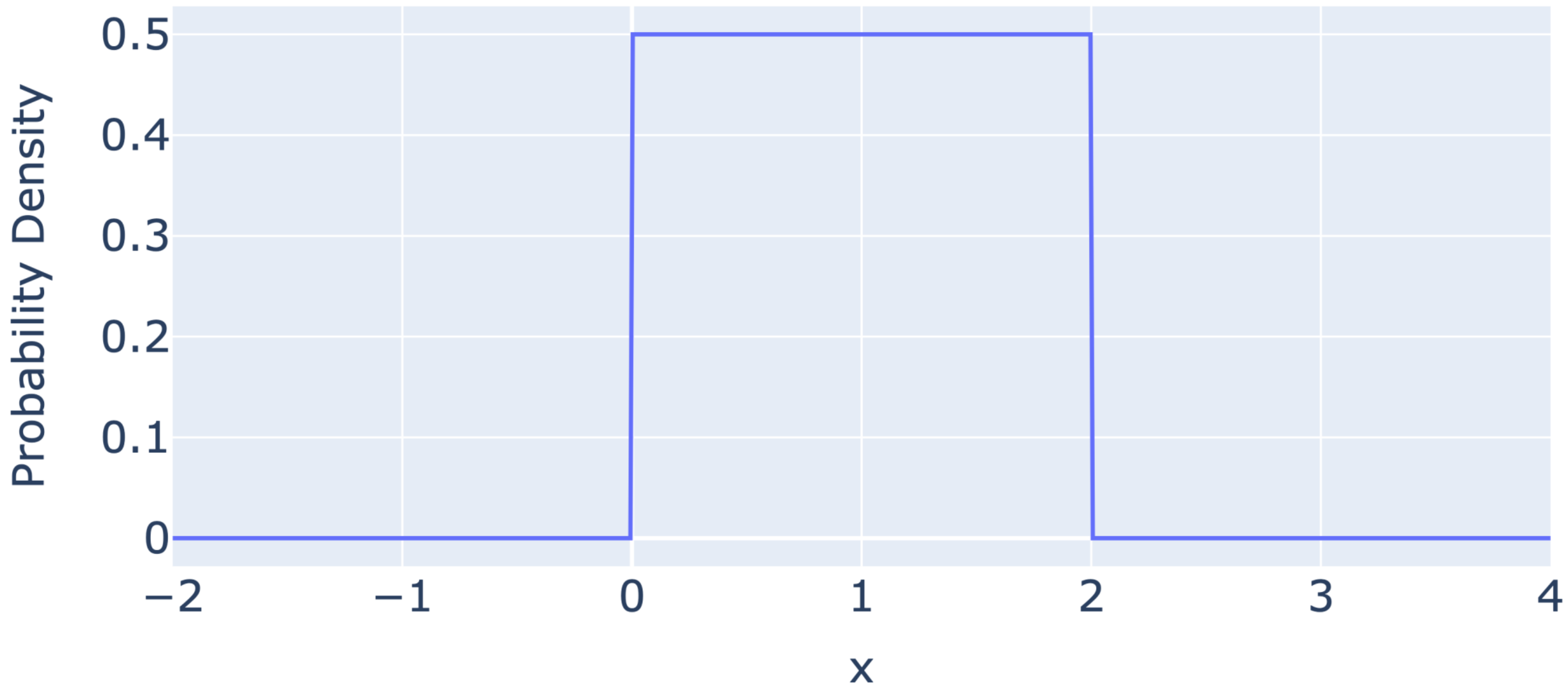
What should be the y-value of the PDF in the non-zero region?

Framing: A Random Variable with Range [0, 2]

Imagine that we choose a real number X, uniformly at random from $[0, 2]$

- To compute $P(X \in [a, b])$, we compute the area under the probability density function, i.e., $(b - a)/2$.
- For this PDF, we can find the area by inspection. It's a rectangle with width 2 and height 0.5, so area is $2 \times 0.5 = 1$. More generally, we can use integration.

I picked 0.5!



Probability Density Function (as described in the notes)

Definition 21.1 (Probability Density Function). A probability density function (PDF) for a real-valued random variable X is a function $f : \mathbb{R} \to \mathbb{R}$ satisfying:

1. f is non-negative: $f(x) \geq 0, \forall x \in \mathbb{R}$
2. The total integral of f is equal to 1: $\int_{-\infty}^{\infty} f(x)dx = 1$

PDF is usually lowercase f

Then the distribution of X is given by:

$$P(a \leq X \leq b) = \int_a^b f(x)dx, \qquad \text{for all } a < b$$

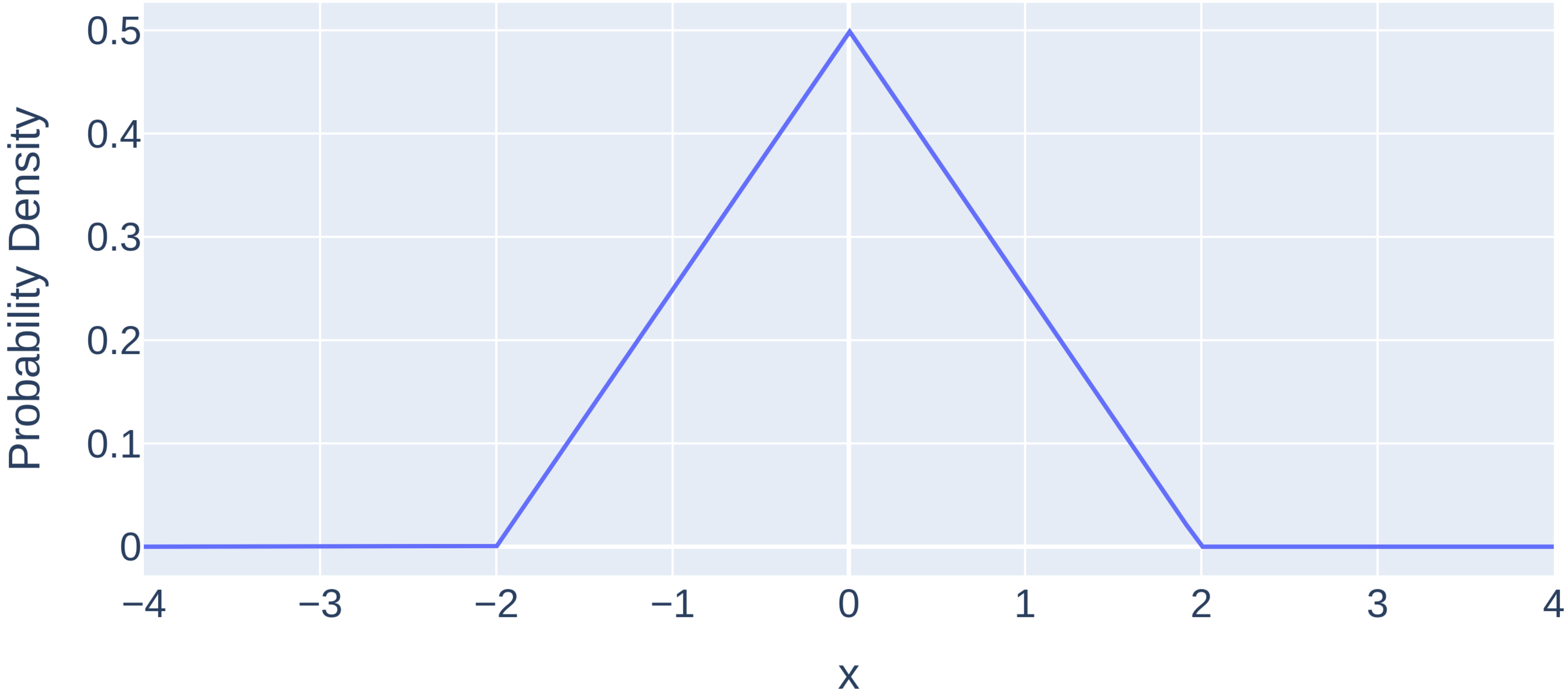
Can think of $f(x)$ as the "probability per unit length" in the vicinity of x.

- $f(x)$ values are not a probability, may be > 1.

Another Example Distribution: The Triangle Distribution

Consider the distribution below:

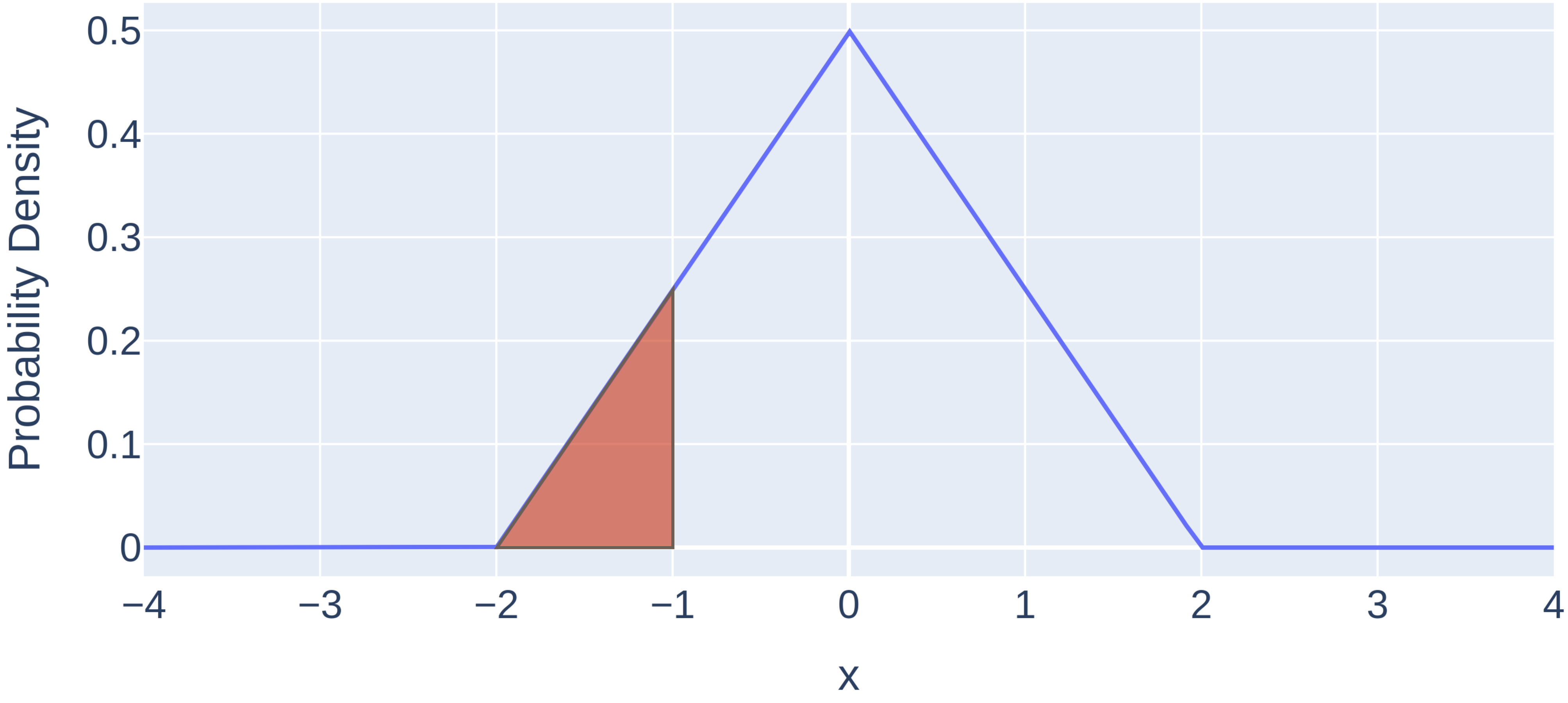
- What is $P(X \leq -1)$?



Another Example Distribution: The Triangle Distribution

Consider the distribution below:

- What is $P(X \leq -1)$?
- This is a right triangle with base 1, height 1/4, so area is $bh/2 = 1/8 = 0.125$



Another Example Distribution: The Triangle Distribution

Consider the distribution below:

- What is $P(X \leq -1)$? Can also compute using integration.

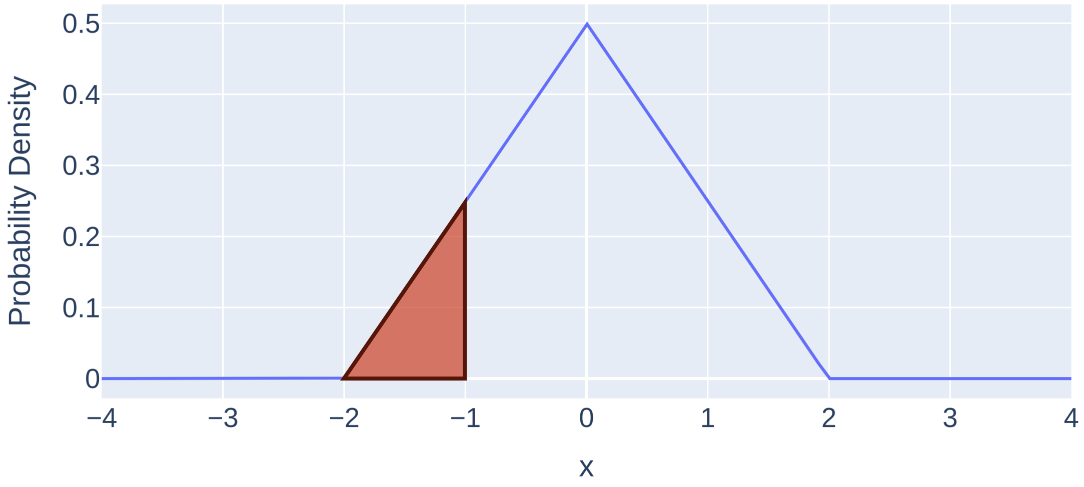
$$P(-\infty \leq X \leq -1) = \int_{-\infty}^{-1} f(x)dx = \int_{-2}^{-1} (0.25x + 0.5)dx$$

$$= (0.125x^2 + 0.5x)\Big|_{x=-2}^{x=-1}$$

$$= (0.125 - 0.5) - (0.5 - 1)$$

$$= (0.125 - 0.5 - 0.5 + 1)$$

$$= 0.125$$



Hypotenuse of triangle is a line with slope 0.25 and y-intercept 0.5

The Cumulative Density Function

In the previous exercise, we computed:

$$P(-\infty \leq X \leq a) = P(X \leq a) = \int_{-\infty}^{a} f(x)dx$$

This has a special name: The Cumulative Distribution Function (CDF).

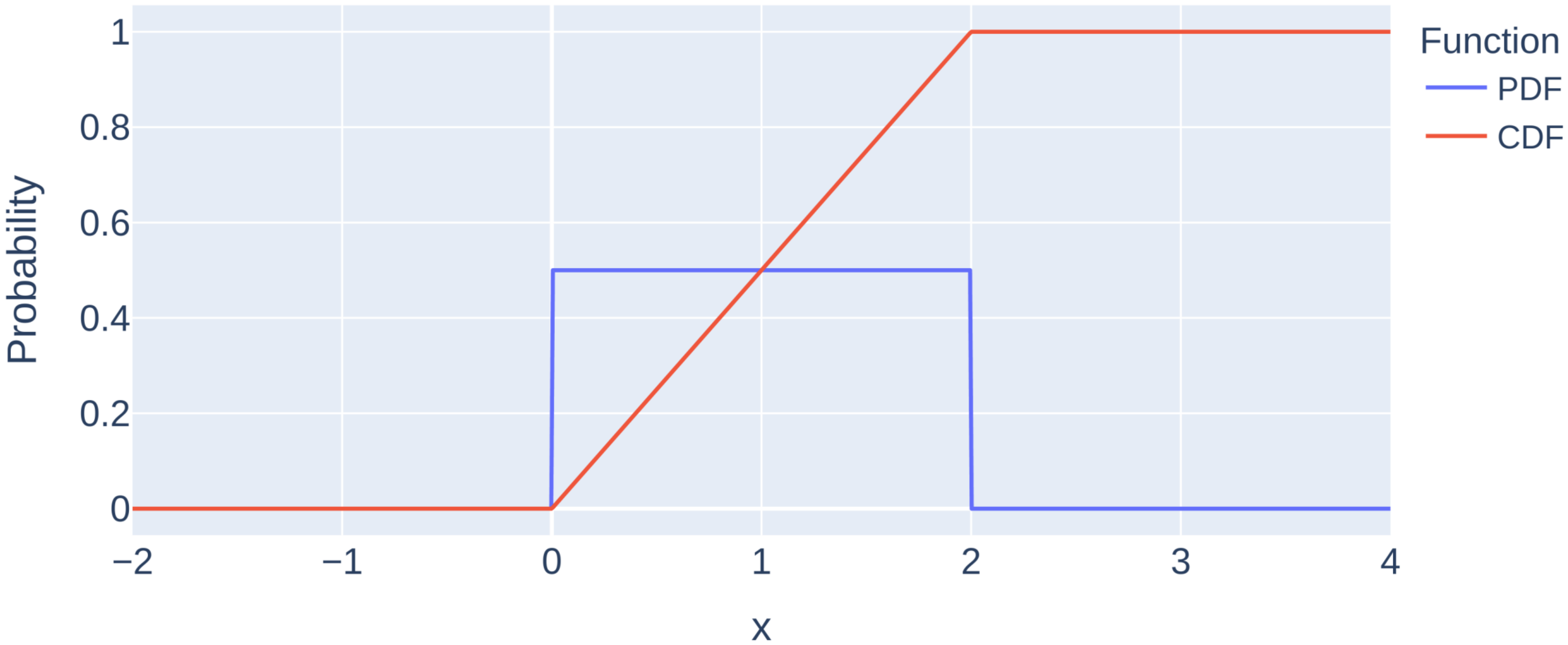
- The CDF tells us the probability that the random variable is $\leq a$.
- The CDF is the integral of the PDF from $-\infty$ to a.
- The CDF is often given as capital F, e.g., $F(a) = P(X \leq a)$.
- Likewise, the PDF is the derivative with respect to a of the CDF:

$$f(a) = \frac{dF(a)}{da}$$

Cumulative Density Function Example 1: Uniform Distribution

Below, we plot the PDF and CDF of the Uniform Distribution together.

- Note: The CDF is the integral of the PDF from $-\infty$ to x.
- Note: The CDF always grows towards 1 (and never exceeds it).



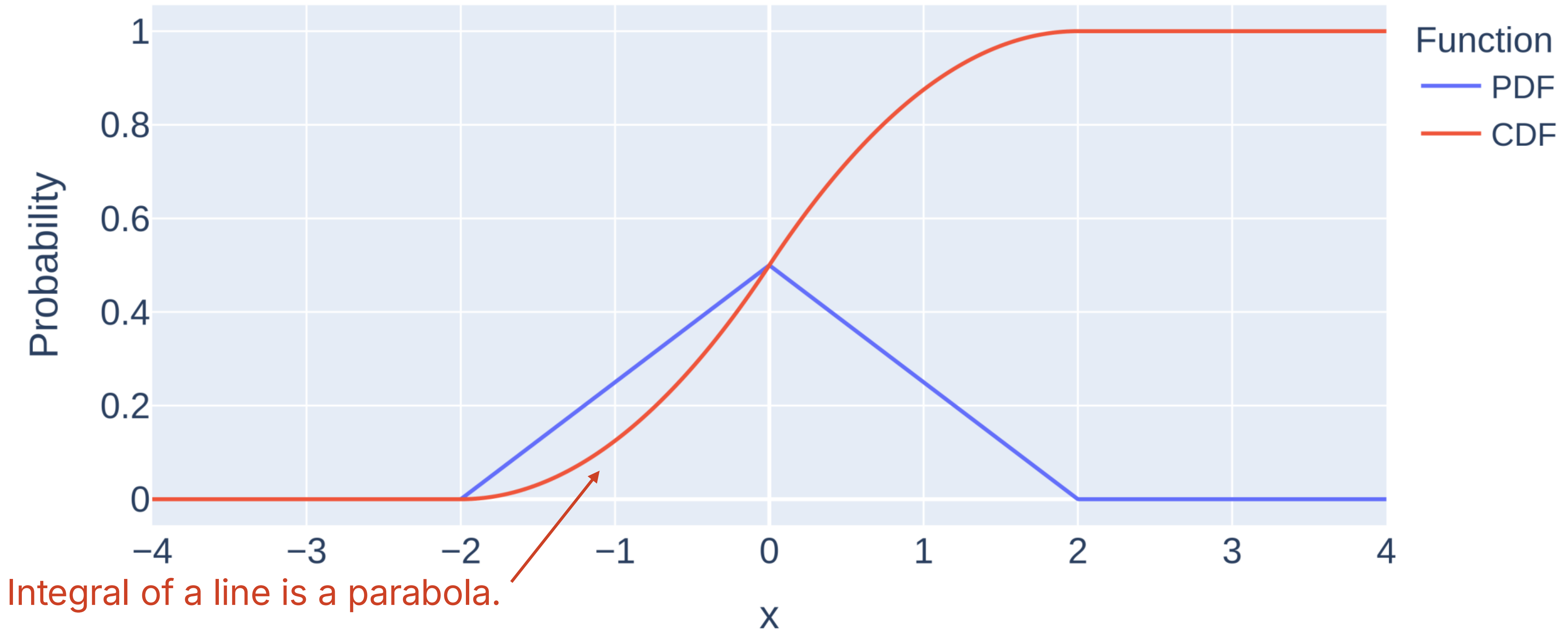
PDF: Probability per unit length

CDF: $P(X \leq x)$

Cumulative Density Function Example 2: Uniform Distribution

Below, we plot the PDF and CDF of the Uniform Distribution together.

- Note: The CDF is the integral of the PDF from $-\infty$ to x.
- Note: The CDF always grows towards 1 (and never exceeds it).



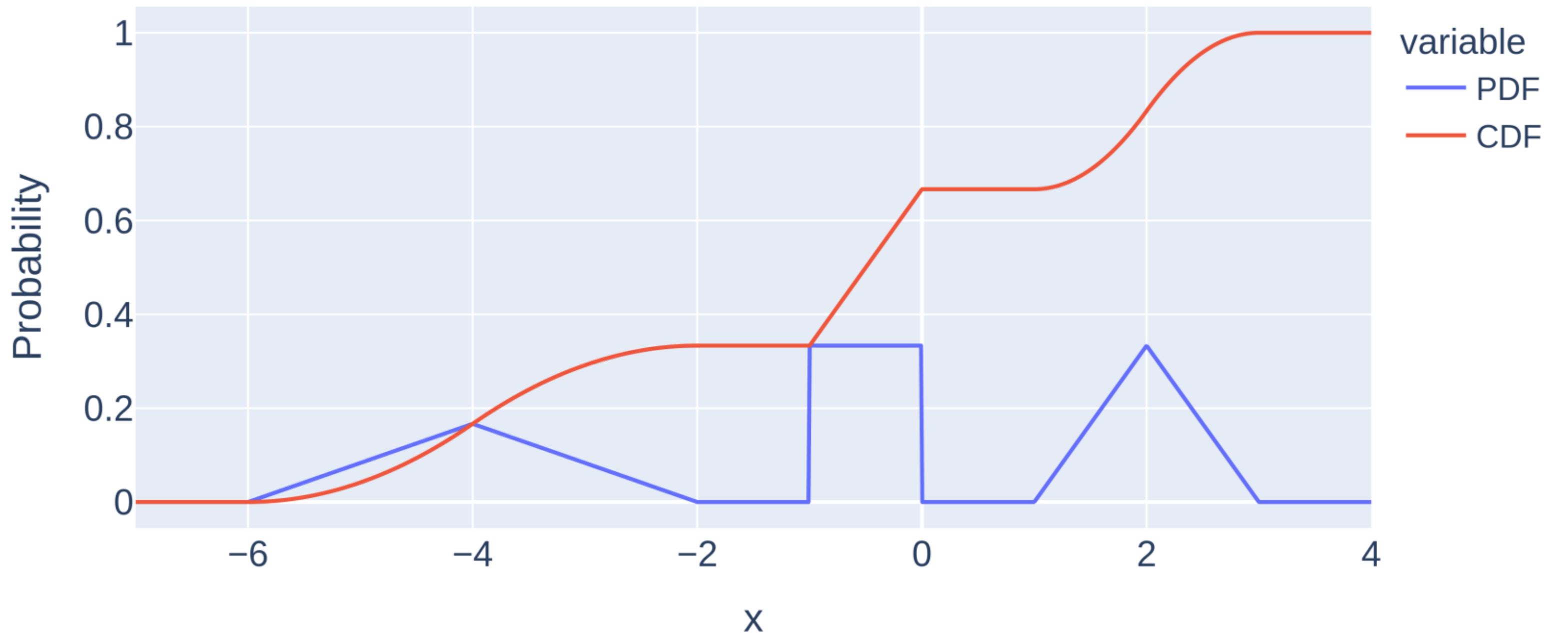
PDF: Probability per unit length

CDF: $P(X \leq x)$

Applying a CDF

Given that CDFs are already the integral of the PDF, we can use the CDF directly to compute the probability of events without integrating.

Example, consider the distribution X with PDF $f(x)$ and CDF $F(x)$ given below:

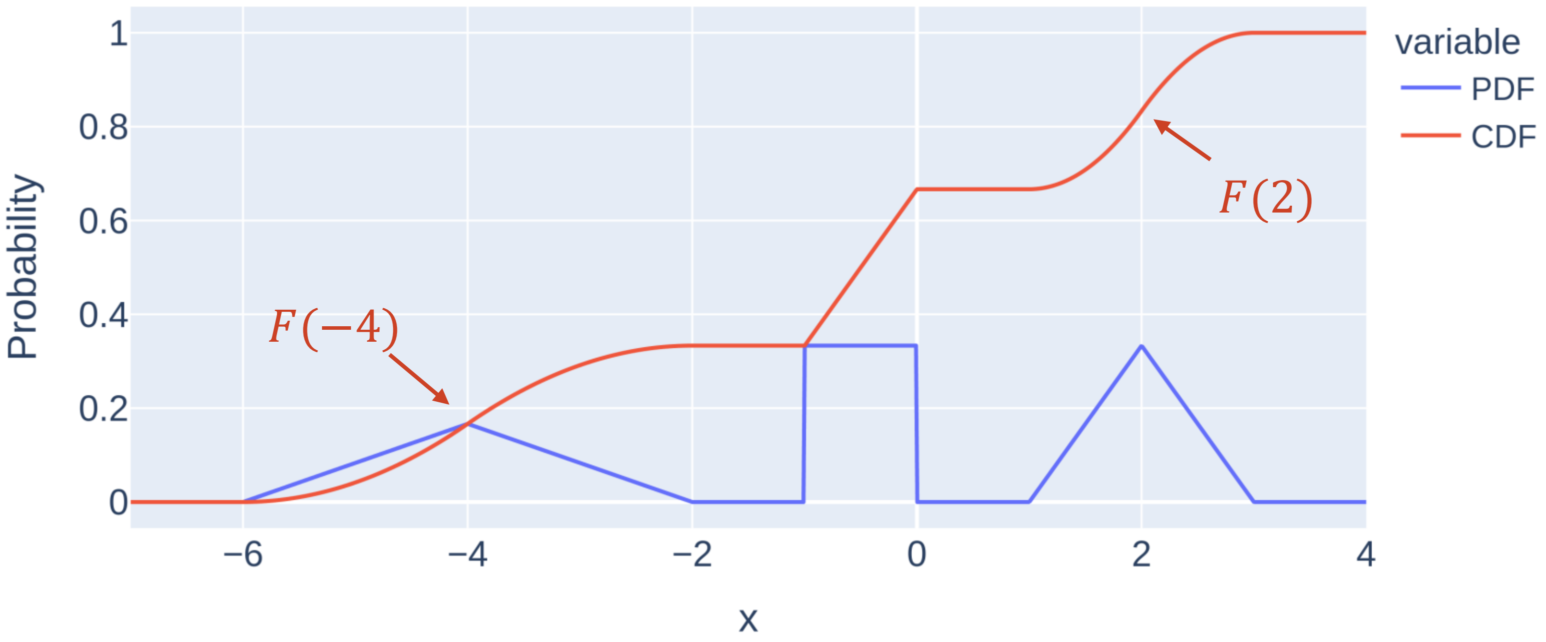


How would we compute $P(x \in [-4, 2])$?

Applying a CDF

Given that CDFs are already the integral of the PDF, we can use the CDF directly to compute the probability of events without integrating.

Example, consider the distribution X with PDF $f(x)$ and CDF $F(x)$ given below:



How would we compute $P(x \in [-4, 2])$? Just compute $F(2) - F(-4)$.

Continuous vs. Discrete Probability Distributions, Exponential Distribution

Lecture 24, CS70 Summer 2025

Discrete vs. Continuous Probability Distributions

We characterize discrete probability distributions by probabilities of outcomes.

- First, we thought about $P(\omega)$.
- We later defined events as a collection of samples in Ω. For an event A, we said that $P(A) = \sum_{\omega \in A} P(\omega)$

Example:

- For $X \sim \text{Binomial}(n, p)$, PMF is $\binom{n}{i} p^i (1-p)^{n-i}$
- $P(X \leq 2) = \binom{n}{0} p^0 (1-p)^{n-0} + \binom{n}{1} p^1 (1-p)^{n-1} + \binom{n}{2} p^2 (1-p)^{n-2}$

$$= \sum_{i=0}^{2} \binom{n}{i} p^i (1-p)^{n-i}$$

Sum for cumulative probability in discrete case – integral for continuous.

Discrete vs. Continuous Probability Distributions

By contrast, for continuous probability distributions, we look at events first.

- CDF gives $P(X \leq a)$, which is the probability of the event that X is less than or equal to a.
- More generally, we only ever care about the probability of intervals (or unions of intervals).
- PDF is the continuous analog of the PMF. PDF outputs are not probabilities.

Continuous Uniform as the Limit of Discrete Uniform

We can think of continuous distributions as the large n limit of discrete distributions. For example:

Let C be uniformly random (continuously) over $(0, 1]$. Let n be some integer.

Define a new random variable $D = \frac{\lceil C \cdot n \rceil}{n}$

- Example: If $C = 0.53557162$, and $n = 1000$, then $D = \frac{\lceil 535.57162 \rceil}{1000} = \frac{536}{1000}$

Then D is a discrete random variable that is uniform in $\left\{\frac{1}{n}, \frac{2}{n}, \frac{3}{n}, \ldots, \frac{n}{n}\right\}$

Naturally $|C - D| \leq \frac{1}{n}$, so as n goes to infinity, D becomes like C.

Goal: Exponential Random Variable. Recall: Geometric

Let's use this large n limit idea to derive an important continuous distribution: The exponential distribution is the continuous analog to the geometric distribution.

Recall, if $X \sim \text{Geometric}(p)$, we are effectively modeling the number of coin tosses we need to make until we get H, where $P(H) = p$.

- $P(X = i) = (1 - p)^{i-1} p$
- $P(X > i) = (1 - p)^{i}$ (this is the probability the first i flips are all tails)

Time Interpretation of Geometric Random Variables

Imagine that we are flipping coins until we get heads at some fixed rate of time, e.g., one flip per minute. Suppose we model this process with a random variable $X \sim \text{Geometric}(\lambda_1)$.

- Here I'm just using λ_1 instead of p. There's nothing important about this.

Now suppose we model the process of flipping coins until we get heads at a rate of one flip per second. Suppose we model this process as the random variable $X_{60} \sim \text{Geometric}(\lambda_2)$.

- What does X_{60} tell you?

Time Interpretation of Geometric Random Variables

Imagine that we are flipping coins until we get heads at some fixed rate of time, e.g., one flip per minute. Suppose we model this process with a random variable $X \sim \text{Geometric}(\lambda_1)$.

- Here I'm just using λ_1 instead of p. There's nothing important about this.

Now suppose we model the process of flipping coins until we get heads at a rate of one flip per second. Suppose we model this process as the random variable $X_{60} \sim \text{Geometric}(\lambda_2)$.

- What does X_{60} tell you? The number of seconds you have to wait to get the first heads when flipping with heads probability λ_2.

Time Interpretation of Geometric Random Variables

Imagine that we are flipping coins until we get heads at some fixed rate of time, e.g., one flip per minute. Suppose we model this process with a random variable $X \sim \mathrm{Geometric}(\lambda)$.

- Here I'm just using λ instead of p. There's nothing important about this.

Now suppose we model the process of flipping coins until we get heads at a rate of one flip per second. Suppose we model this process as the random variable $X_{60} \sim \mathrm{Geometric}\left(\frac{\lambda}{60}\right)$.

- What does X_{60} tell you? The number of seconds you have to wait to get the first heads when flipping with heads probability $\lambda/60$. Note: We've here chosen $\lambda_2 = \lambda_1/60$

If we multiply X by 60, do we get X_{60}?

Time Interpretation of Geometric Random Variables

Imagine that we are flipping coins until we get heads at some fixed rate of time, e.g., one flip per minute. Suppose we model this process with a random variable $X \sim \text{Geometric}(\lambda)$.

- Here I'm just using λ instead of p. There's nothing important about this.

Now suppose we model the process of flipping coins until we get heads at a rate of one flip per second. Suppose we model this process as the random variable $X_{60} \sim \text{Geometric}\left(\frac{\lambda}{60}\right)$.

- What does X_{60} tell you? The number of seconds you have to wait to get the first heads when flipping with heads probability $\lambda/60$. Note: We've here chosen $\lambda_2 = \lambda_1/60$

If we multiply X by 60, do we get X_{60}? No. $\text{range}(60X) = \{60, 120, \ldots\}$, but $\text{range}(X_{60}) = \{1, 2, \ldots\}$

Interpretation of Increasing N

Define $X_n \sim \text{Geometric}(\lambda/n)$, this is the number of time intervals we need to wait if we flip n times per minute, with a probabililty of heads of λ/n.

- Effectively, we're increasing the temporal resolution of our experiment to an arbitrary degree.
- X_1: Models the number of minutes we have to wait for heads, can only yield number of minutes we have to wait.
- X_{60}: Models the number of seconds we have to wait for heads, can only yield number of seconds we have to wait.
- $X_{60{,}000{,}000}$: Models the number of microseconds we have to wait for heads, can only yield number of microseconds we have to wait.

Interesting question: What happen as n goes to infinity?

- We'll use our old friend $\lim_{n\to\infty}(1-1/x)^x = e^{-1}$

Fixed Time Geometric Random Variable as $n \to \infty$

Define $X_n \sim \text{Geometric}(\lambda/n)$, this is the number of time intervals we need to wait if we flip n times per minute, with a chance of heads of λ/n.

We know for $X \sim \text{Geometric}(p)$, that $P(X > i) = (1-p)^i$ for $i \in \{0, 1, \dots\}$

We have that $P\left(X_n > \frac{i}{n}\right) = \left(1 - \frac{\lambda}{n}\right)^i$ for $i/n \in \{0, 1, \dots\}$

We have that for any value $a \in \text{range}(X_n)$:

$$P(X_n > a) = P\left(X_n > \frac{an}{n}\right)$$

Why? We can just trivially multiply a by n/n. This will be useful in a moment.

Fixed Time Geometric Random Variable as $n \to \infty$

Define $X_n \sim \text{Geometric}(\lambda/n)$, this is the number of time intervals we need to wait if we flip n times per minute, with a chance of heads of λ/n.

We know for $X \sim \text{Geometric}(p)$, that $P(X > i) = (1-p)^i$ for $i \in \{0, 1, \ldots\}$

We have that $P\left(X_n > \frac{i}{n}\right) = \left(1 - \frac{\lambda}{n}\right)^i$ for $i/n \in \{0, 1, \ldots\}$

We have that for any value $a \in \text{range}(X_n)$:

$$P(X_n > a) = P\left(X_n > \frac{an}{n}\right) = \left(1 - \frac{\lambda}{n}\right)^{an} = \left(\left(1 - \frac{\lambda}{n}\right)^{n/\lambda}\right)^{a\lambda} \to e^{-a\lambda}$$

$$\lim_{n\to\infty}(1 - 1/x)^x = e^{-1}$$

Exponential Distribution

We've shown that if we crank the temporal resolution of our coin flipping experiment to infinity (i.e., informally $\lim_{n\to\infty} X_n$), we get:

$$P(X_n > a) \to e^{-a\lambda} \qquad \text{for } a \geq 0$$

The random variable X_∞ we end up with is the exponential distribution.

- This is a continuous distribution, i.e., $P(X_\infty = a) = 0$ for any specific a.
- The equation above gives the probability of the event $X_\infty > a$.

Note: This approach of thinking of X_n as a family of random variables is not in the notes. The terminology X_∞ as representing the "final" random variable in the family is non-standard.

Exponential Distribution

If $Y \sim \mathrm{Exp}(\lambda)$, then we know:

$$P(Y > a) = \begin{cases} e^{-a\lambda}, & \text{if } a \geq 0 \\ 1, & \text{otherwise} \end{cases}$$

What is the CDF of Y for $a \geq 0$? Reminder, a CDF gives the probability that a random variable is less than or equal to some quantity.

Exponential Distribution

If $Y \sim \text{Exp}(\lambda)$, then we know:

$$P(Y > a) = \begin{cases} e^{-a\lambda}, & \text{if } a \geq 0 \\ 1, & \text{otherwise} \end{cases}$$

Since $P(Y \leq a) = 1 - P(Y > a)$, we have that the CDF $F_Y(a)$ is given by:

$$P(Y \leq a) = \begin{cases} 1 - e^{-a\lambda}, & \text{if } a \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

How do we get the PDF $f_Y(y)$ given the CDF?

What is the PDF $f_Y(y)$?

Exponential Distribution

If $Y \sim \mathrm{Exp}(\lambda)$, then we know:

$$P(Y > a) = \begin{cases} e^{-a\lambda}, & \text{if } a \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

Since $P(Y \leq a) = 1 - P(Y > a)$, we have that the CDF $F_Y(a)$ is given by:

$$F_Y(a) = \begin{cases} 1 - e^{-a\lambda}, & \text{if } a \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

The PDF is given by the derivative of the CDF:

$$f_Y(a) = \begin{cases} \lambda e^{-a\lambda}, & \text{if } a \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

Application of an Exponential Random Variable

Suppose that the that the number of minutes until the next neutrino detection is given by $X \sim \mathrm{Exp}(0.02)$.

What is the probability that the next neutrino detection occurs in the next 100 minutes?

$$F_Y(y) = \begin{cases} 1 - e^{-\lambda y}, & \text{if } y \geq 0 \\ 0, & \text{otherwise} \end{cases} \qquad f_Y(y) = \begin{cases} \lambda e^{-\lambda y}, & \text{if } a \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

Application of an Exponential Random Variable

Suppose that the that the number of minutes until the next neutrino detection is given by $X \sim \mathrm{Exp}(0.02)$.

What is the chance that the next neutrino detection occurs in the next 100 minutes?

- Easiest to use the CDF! We want $F_Y(100) = P(Y < 100)$

$$F_Y(y) = \begin{cases} 1 - e^{-\lambda y}, & \text{if } y \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

$$= 1 - e^{-100 * 0.02}$$

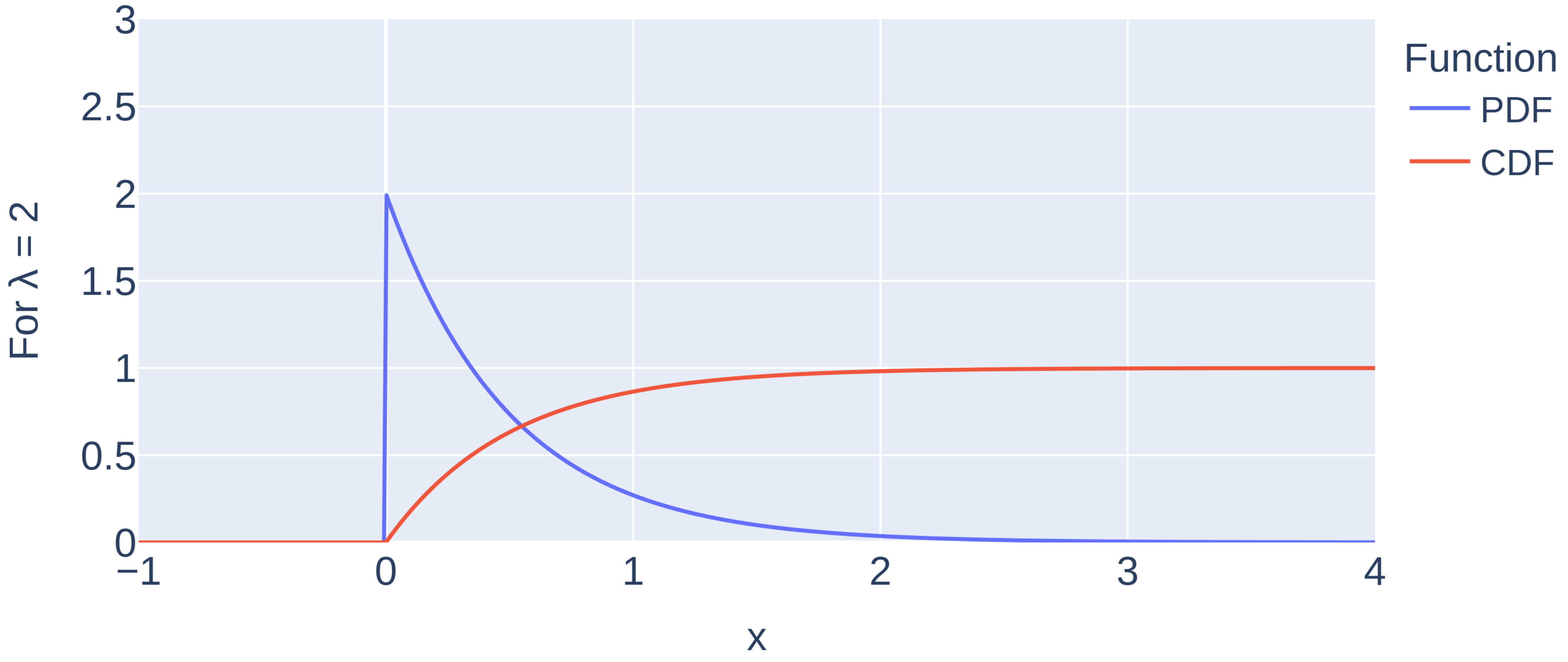
$$= 86.47\%$$

The PDF and CDF of an Exponential Random Variable (Visually)

Below are the PDF and CDF of $X \sim \mathrm{Exp}(2)$.

- From the plot, how do we verify the area under the PDF is 1?

Integrate!
(*exercise left for the reader*)

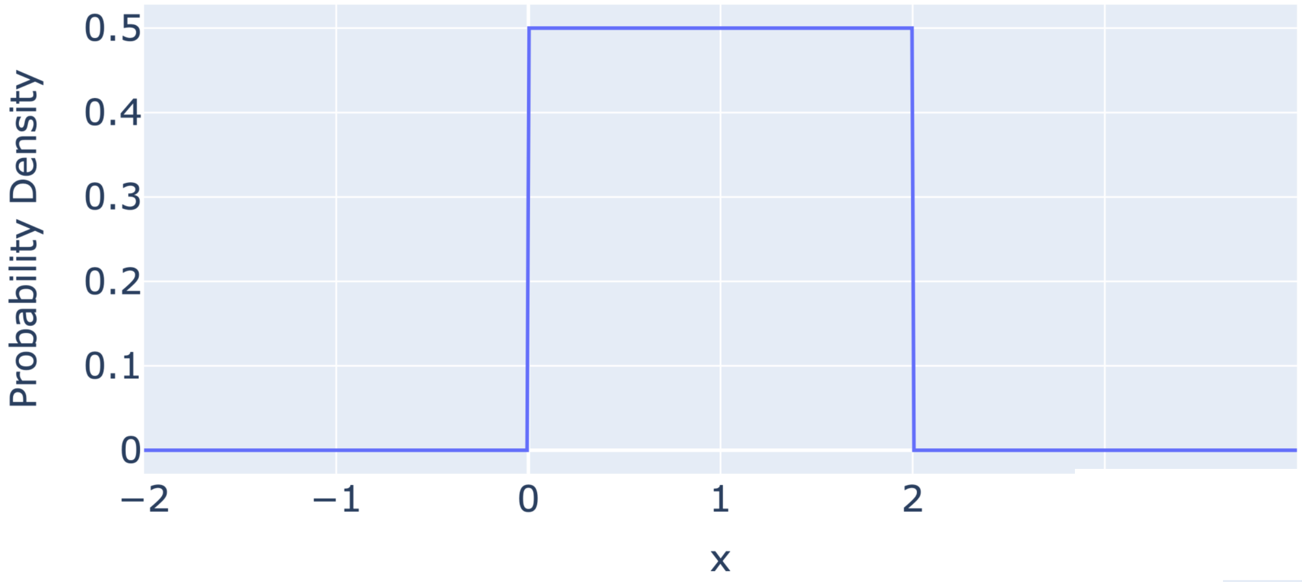


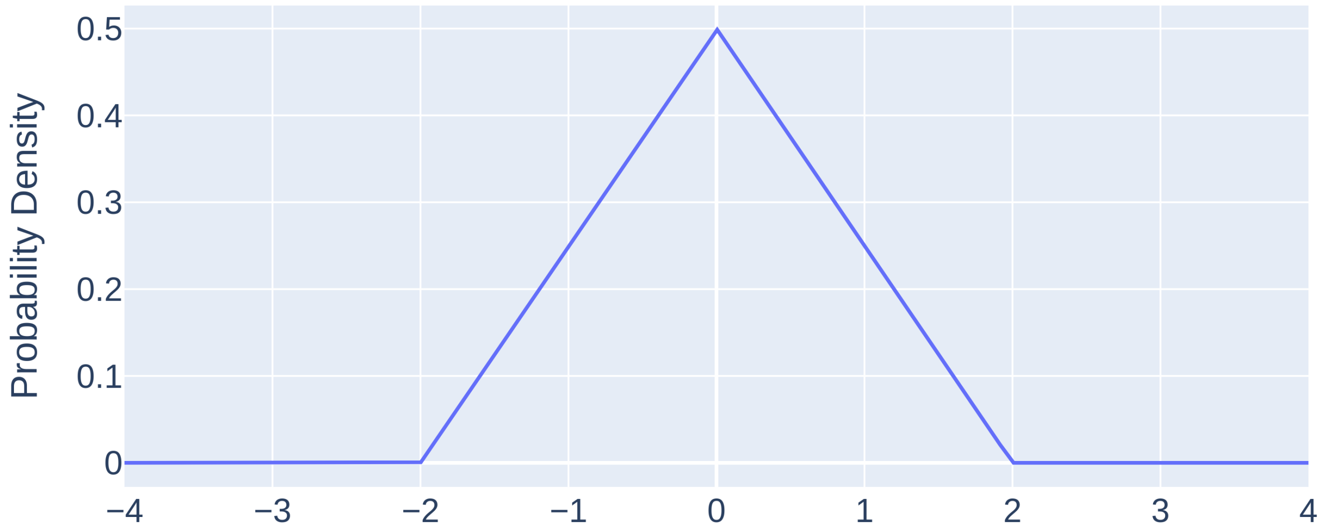
Expectation and Variance

Lecture 24, CS70 Summer 2025

Expectation of a Continuous Random Variable

What do you think are the expectations of the random variables below?

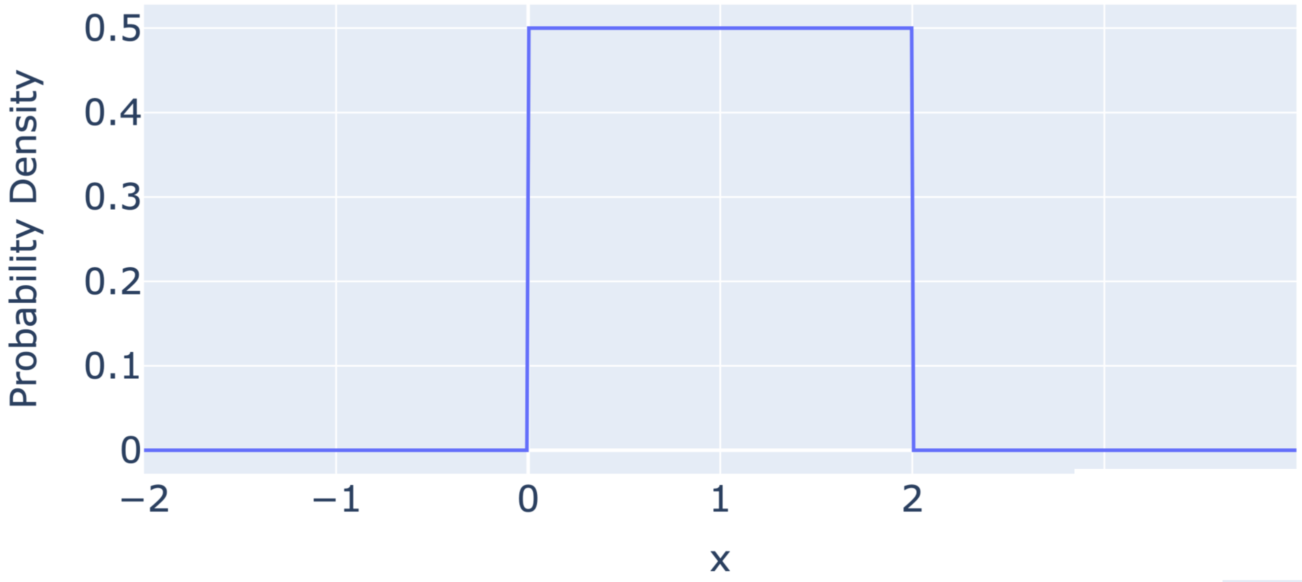




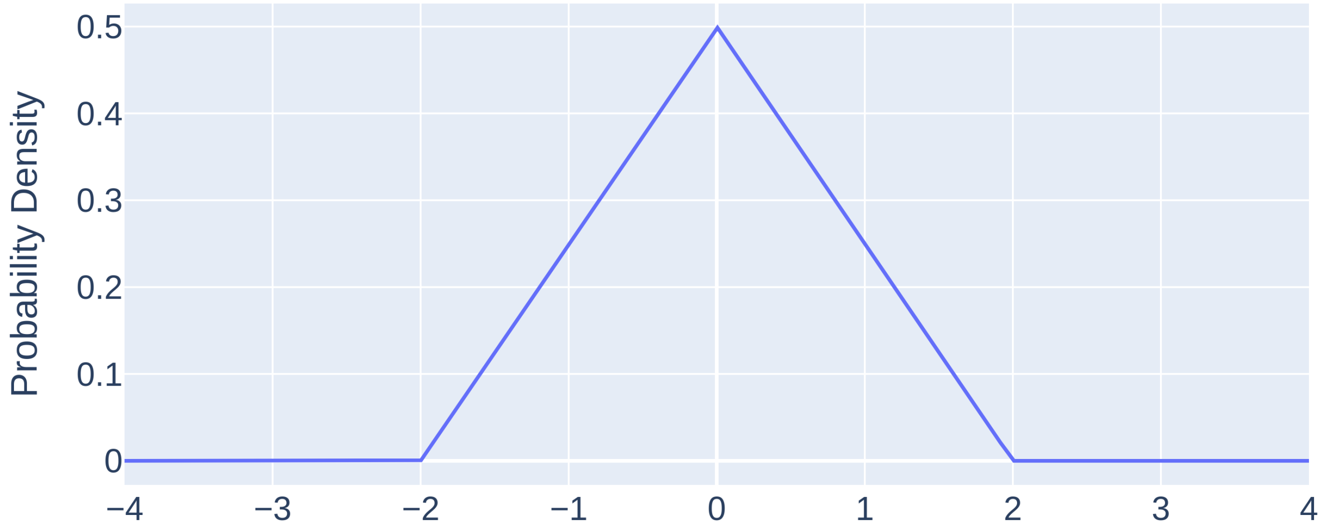
Expectation of a Continuous Random Variable

What do you think are the expectations of the random variables below?

x=1

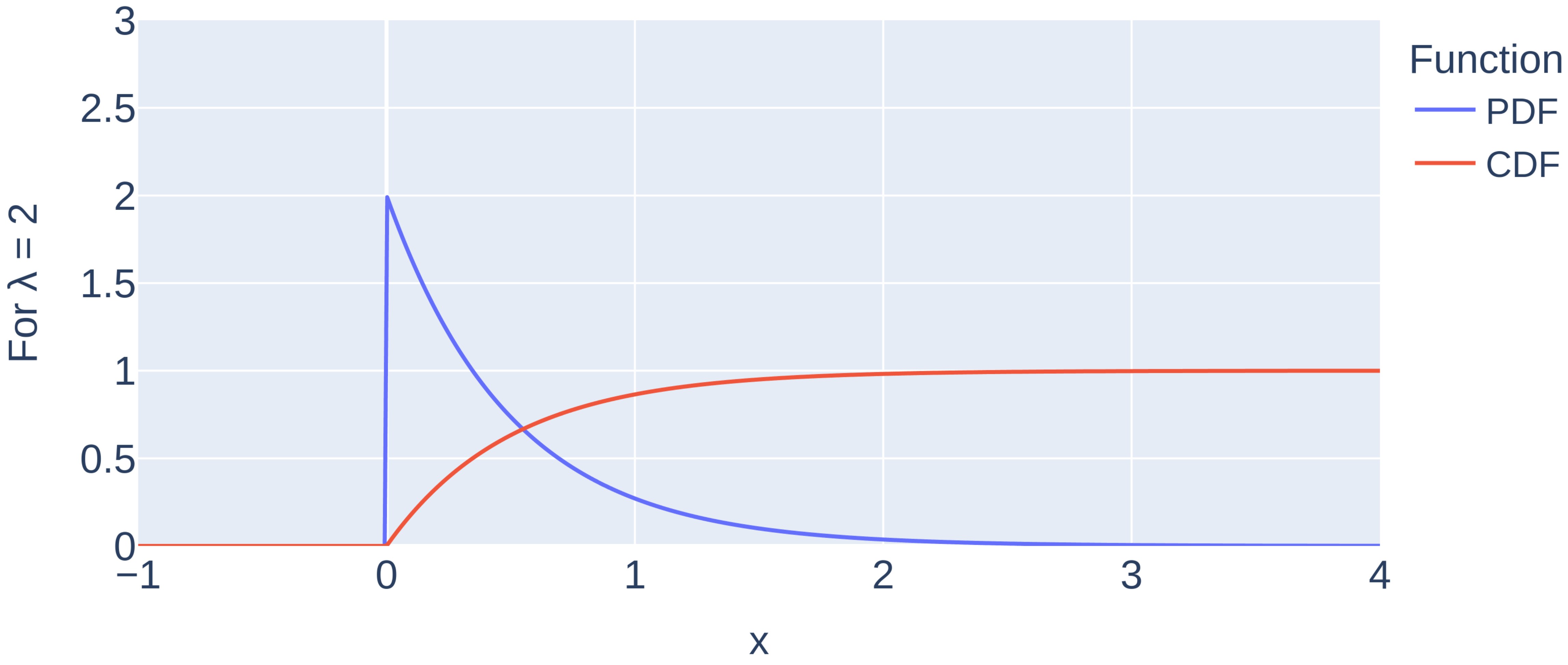


x=0



Expectation of an Exponential Random Variable

What do you think is the expectation of $X \sim \mathrm{Exp}(2)$?



Definition of Expectation for Continuous Random Variable

If X is a random variable with PDF $f_X(x)$, then $E[X]$ is:

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

Contrast with our discrete expectation formula:

$$E[X] = \sum_{x \in \text{range}(X)} x \cdot P(X = x)$$

Expectation for $X \sim \text{Exp}(\lambda)$

To compute the $E[X]$ where $X \sim \text{Exp}(\lambda)$, we integrate over the PDF of our exponential random variable.

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{if } x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

$$E[X] = \int_0^\infty x \cdot \lambda e^{-\lambda x} dx$$

Recall Integration by Parts:

$$\int_a^b u(x) \cdot v'(x) dx = u(x) \cdot v(x) \Big|_a^b - \int_a^b u'(x) \cdot v(x) dx$$

Expectation for $X \sim \text{Exp}(\lambda)$

To compute the $E[X]$ where $X \sim \text{Exp}(\lambda)$, we integrate over the PDF of our exponential random variable.

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{if } x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

$$E[X] = \int_0^\infty \underbrace{x}_{u(x)} \cdot \underbrace{\lambda e^{-\lambda x}}_{v'(x)} dx = \underbrace{x \cdot \left(-e^{-\lambda x}\right)\Big|_0^\infty}_{=0} - \underbrace{\int_0^\infty 1 \cdot \left(-e^{-\lambda x}\right) dx}_{-\frac{e^{-\lambda x}}{\lambda}\Big|_0^\infty = \frac{1}{\lambda}}$$

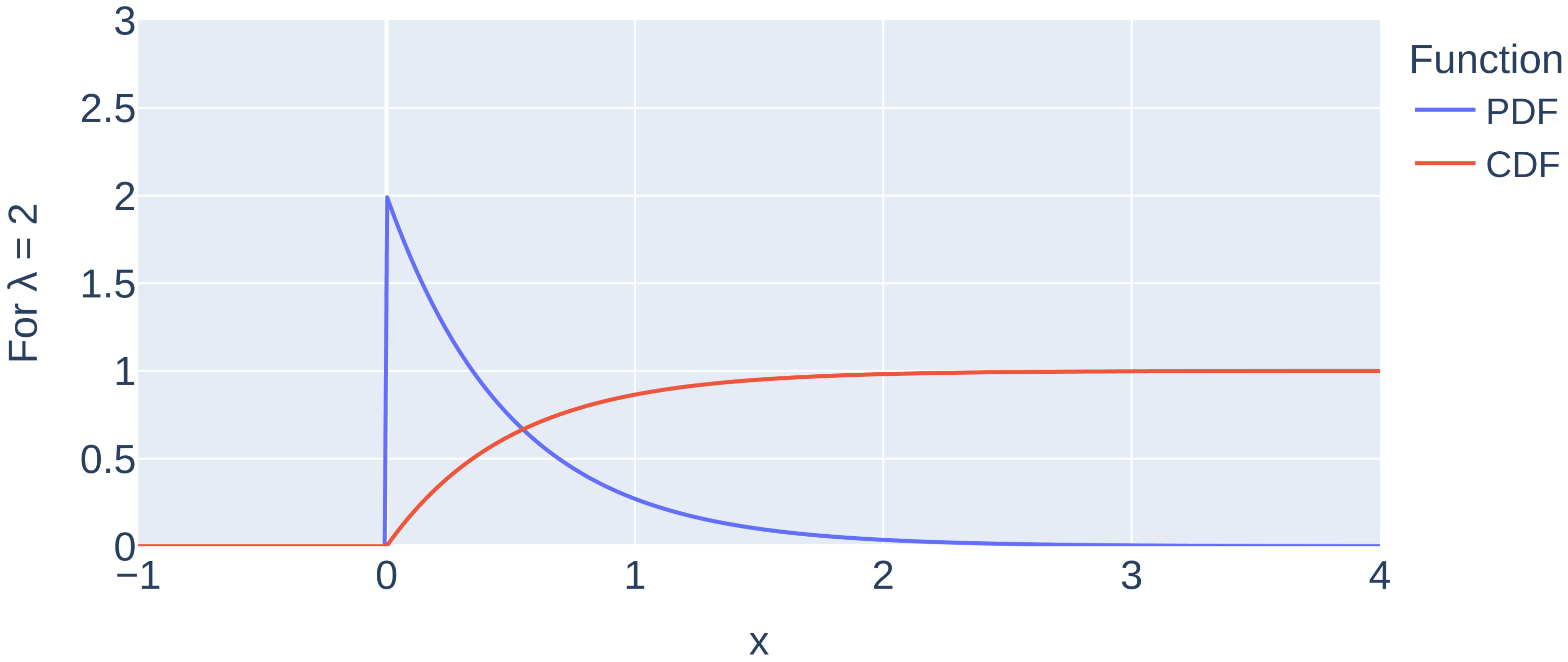
Recall Integration by Parts:

$$\int_a^b u(x) \cdot v'(x) dx = u(x) \cdot v(x)\Big|_a^b - \int_a^b u'(x) \cdot v(x) dx$$

Expectation of an Exponential Random Variable

What do you think is the expectation of $X \sim \text{Exp}(2)$?

- Answer: $1/\lambda = 1/2$



Definition of Variance for Continuous Random Variable

If X is a random variable with PDF $f_X(x)$, then $\mathrm{var}(\mathrm{X})$ is:

$$\mathrm{var}(X) = E[X^2] - E[X]^2$$

$$= \int_{-\infty}^{\infty} x^2 \cdot f_X(x)dx - \left(\int_{-\infty}^{\infty} x \cdot f_X(x)dx\right)^2$$

Expectation for $X \sim \text{Exp}(\lambda)$

To compute the variance, we'll need $E[X^2]$. Luckily, this isn't so bad.

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{if } x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

$$E[X] = \int_0^\infty \lambda x e^{-\lambda x} dx = \frac{1}{\lambda}$$

$$E[X^2] = \int_0^\infty \underbrace{x^2}_{u(x)} \cdot \underbrace{\lambda e^{-\lambda x}}_{v'(x)} dx = \underbrace{x^2 \cdot \left(-e^{-\lambda x}\right)\Big|_0^\infty}_{=0} - \int_0^\infty 2x \cdot \left(-e^{\lambda x}\right) dx$$

$$+ \frac{2}{\lambda} \int_0^\infty \lambda x e^{-\lambda x} dx = \frac{2}{\lambda} \cdot \frac{1}{\lambda} = \frac{2}{\lambda^2}$$

Recall Integration by Parts from HW1:

$$\int_a^b u(x) \cdot v'(x) dx = u(x) \cdot v(x)\Big|_a^b - \int_a^b u'(x) \cdot v(x) dx$$

Expectation for $X \sim \text{Exp}(\lambda)$

To compute the variance, we'll need $E[X^2]$. Luckily, this isn't so bad.

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{if } x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

$$E[X] = \int_0^\infty \lambda x e^{-\lambda x} dx = \frac{1}{\lambda}$$

$$E[X^2] = \frac{2}{\lambda^2}$$

$$\text{var}(X) = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

Geometric vs. Exponential

Compare $\mathrm{C}\sim\mathrm{Exp}(\lambda)$ and $D\sim\mathrm{Geometric}(\lambda)$

$$E[C] = \frac{1}{\lambda} \qquad E[D] = \frac{1}{\lambda}$$

$$\mathrm{var}(C) = \frac{1}{\lambda^2} \qquad \mathrm{var}(D) = \frac{1-\lambda}{\lambda^2}$$

Expectation and Variance of a Uniform Random Variable

In the notes, it is shown that if X is a uniform random variable over $[0, \ell]$:

$$E[X] = \frac{\ell}{2} \qquad \mathrm{var}(X) = \frac{\ell^2}{12}$$

See notes (or just show yourself) for a proof. Contrast with discrete D uniform over $\{1, 2, \ldots, n\}$, which was:

$$E[D] = \frac{n+1}{2} \qquad \mathrm{var}(D) = \frac{n^2 - 1}{12}$$

Joint Distributions and Independence

Lecture 24, CS70 Summer 2025

Joint Distributions and Independence

We'll take a brief look at joint distributions and independence.

- We'll cover these in more detail in the next lecture.
- Today sets you up with the basic definitions you'll need for the discussion.

Joint Density

Suppose we have random variables X and Y which are continuous. Then their joint PDF is a non-negative function $f_{X,Y}(x, y)$ such that:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dx\, dy = 1$$

And for all $a < b$ and $c < d$, we have:

$$P(a \leq X \leq b, c \leq Y \leq d) = \int_{c}^{d}\int_{a}^{b} f_{X,Y}(x, y)\, dx\, dy$$

Can interpret PDF as probability that we are in a small rectangle around x, y.

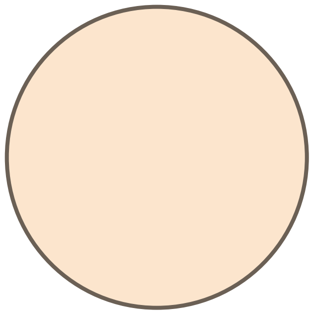
- Actual probability is zero since the rectangle infinitesimal.

Joint Density Example

Suppose we have random variables X and Y which are continuous. Then their joint PDF is a non-negative function $f_{X,Y}(x, y)$ such that:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dx\, dy = 1$$

Suppose we throw a dart at the origin. Suppose we land uniformly within 2 feet of the origin.



2 foot radius

$$f_{X,Y}(x, y) = \begin{cases} \dfrac{1}{4\pi}, & \text{if } x^2 + y^2 \le 4 \\ 0, & \text{otherwise} \end{cases}$$

Marginal Probabilities

Suppose we have random variables X and Y which are continuous. Then their joint PDF is a non-negative function $f_{X,Y}(x, y)$ such that:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dx\, dy = 1$$

Similar to marginal probabilities for discrete random variables, we can compute $f_X(x)$ and $f_Y(y)$. How?

Marginal Probabilities

Suppose we have random variables X and Y which are continuous. Then their joint PDF is a non-negative function $f_{X,Y}(x, y)$ such that:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dx\, dy = 1$$

Similar to marginal probabilities for discrete random variables, we can compute $f_X(x)$ and $f_Y(y)$. How?

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dy \qquad f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dx$$

Independent Random Variables

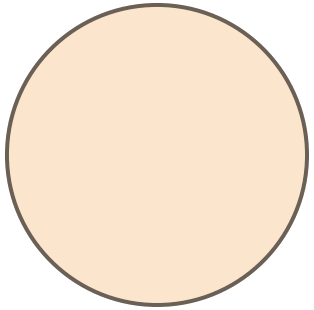
We saw X and Y are independent continuous random variables if for all $a < b$ and $c < d$.

$$P(a \le X \le b, c \le Y \le d) = P(a \le X \le b) \cdot P(c \le Y \le d)$$

Or equivalently: If $f_{X,Y}(x, y) = f_X(x) \cdot f_Y(y)$ for all x, y

For the dart example earlier, are X and Y independent?

- Why or why not?



2 foot radius

Independent Random Variables

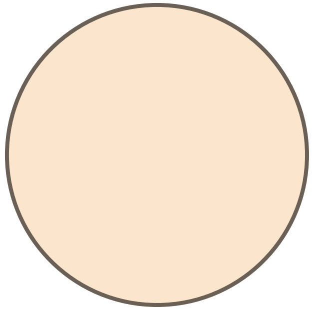
We saw X and Y are independent continuous random variables if for all $a < b$ and $c < d$.

$$P(a \leq X \leq b, c \leq Y \leq d) = P(a \leq X \leq b) \cdot P(c \leq Y \leq d)$$

Or equivalently: If $f_{X,Y}(x, y) = f_X(x) \cdot f_Y(y)$ for all x, y

For the dart example earlier, are X and Y independent?

- Not independent! Example: if $X = 2$, Y must be zero.



2 foot radius

Independent Random Variables

We've only touched on joint PDFs, marginal PDFs, and independence.

We'll return to these ideas in the next lecture.