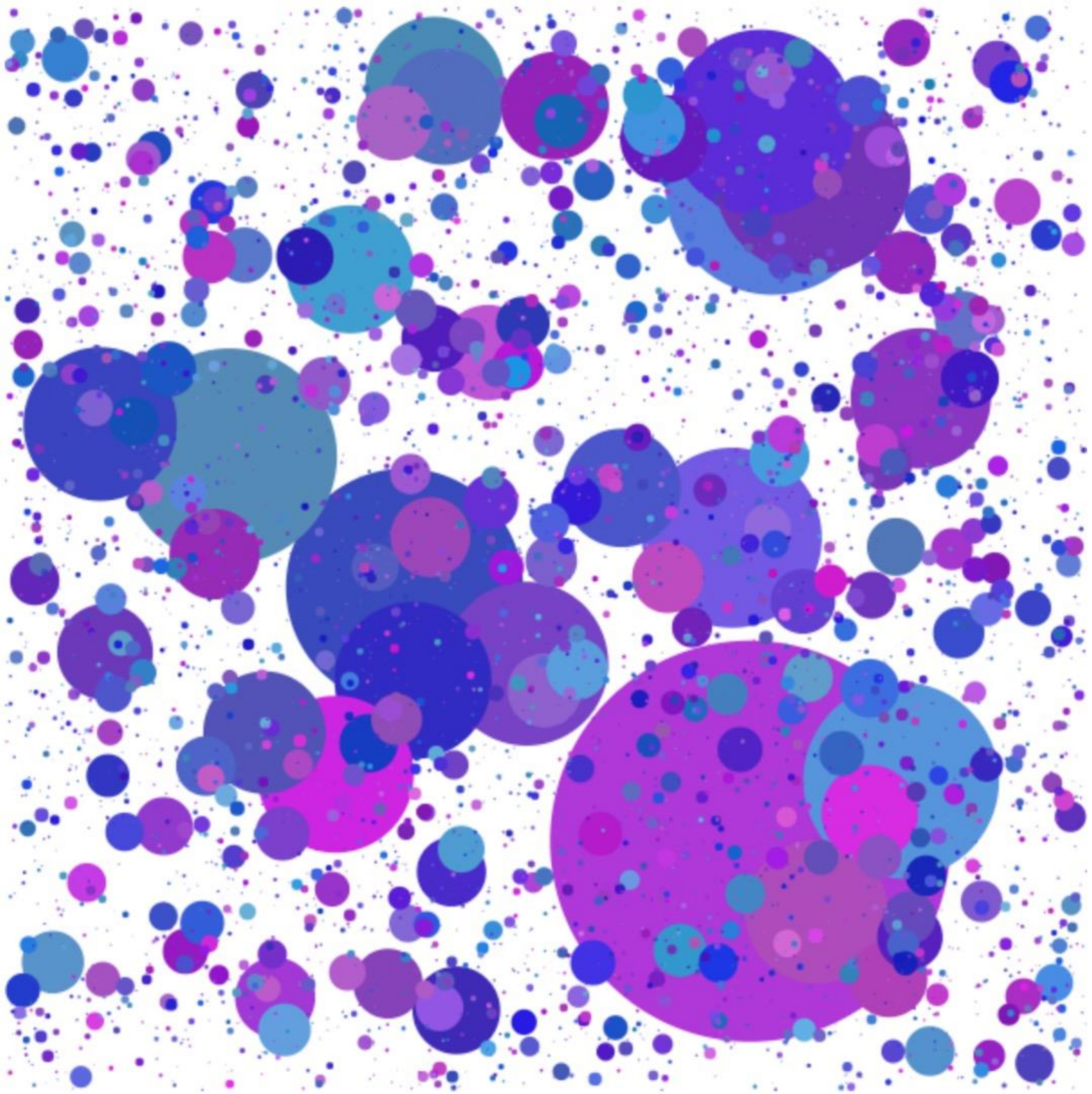
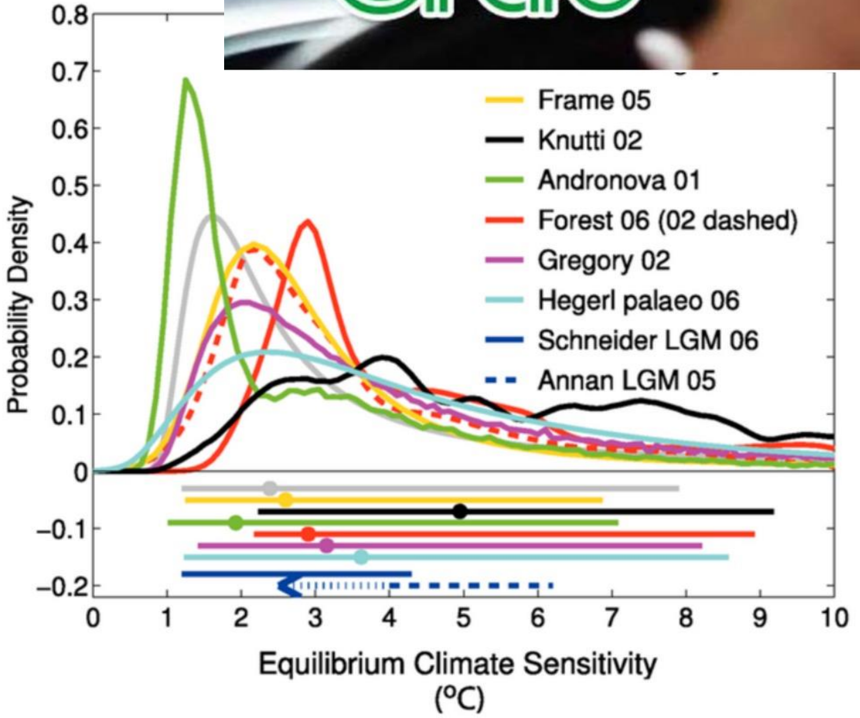
Continuous Variables

CS109, Stanford University

Pset #3 is Out




GrabCar
Grab

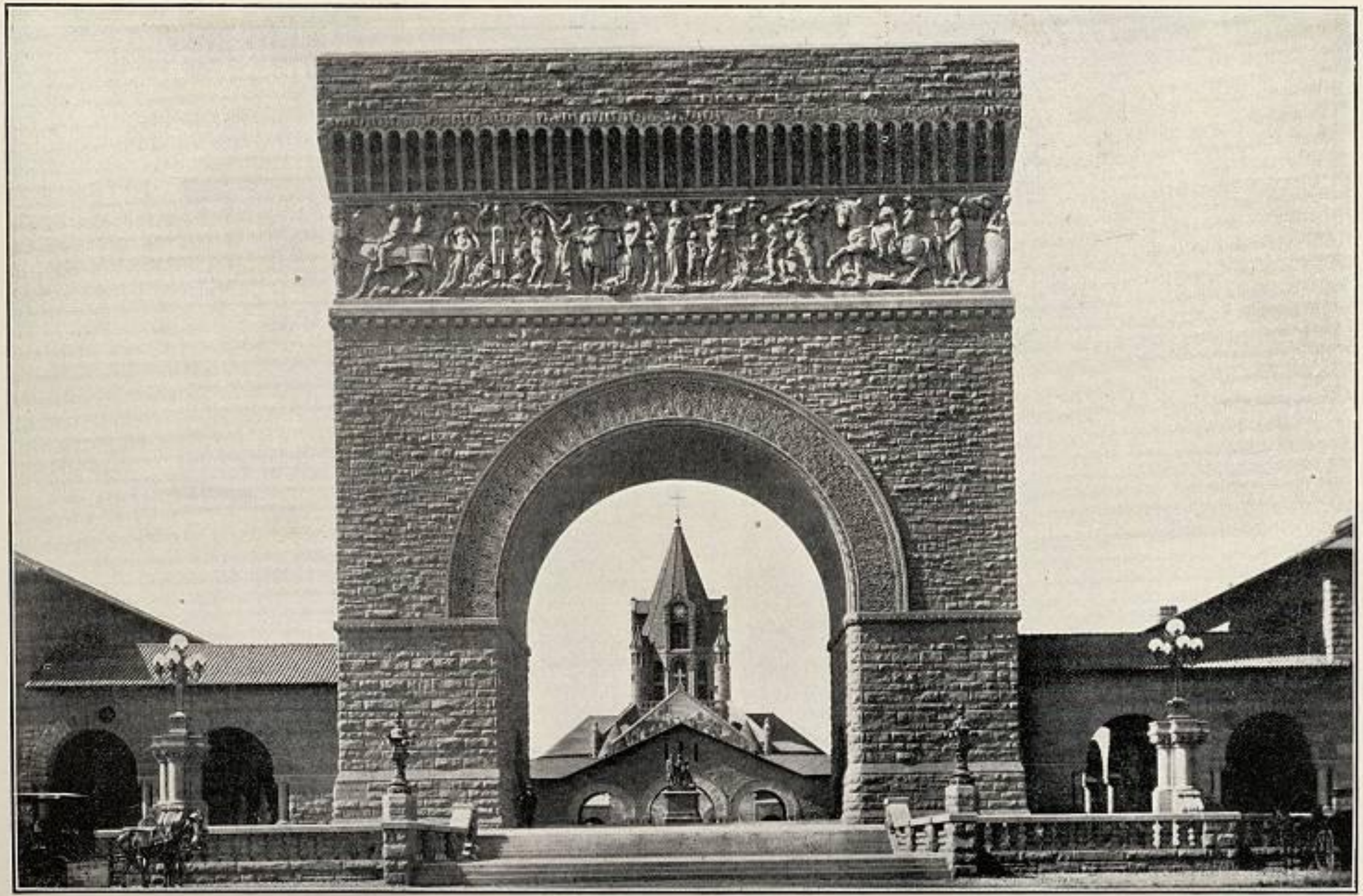

Frame 05
Knutti 02
Andronova 01
Forest 06 (02 dashed)
Gregory 02
Hegerl palaeo 06
Schneider LGM 06
Annan LGM 05
Probability Density
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
−0.1
−0.2
0 1 2 3 4 5 6 7 8 9 10
Equilibrium Climate Sensitivity
(°C)

Vote for Songs!

In a few weeks we are going to use those votes to motivate an important concept in probability theory.

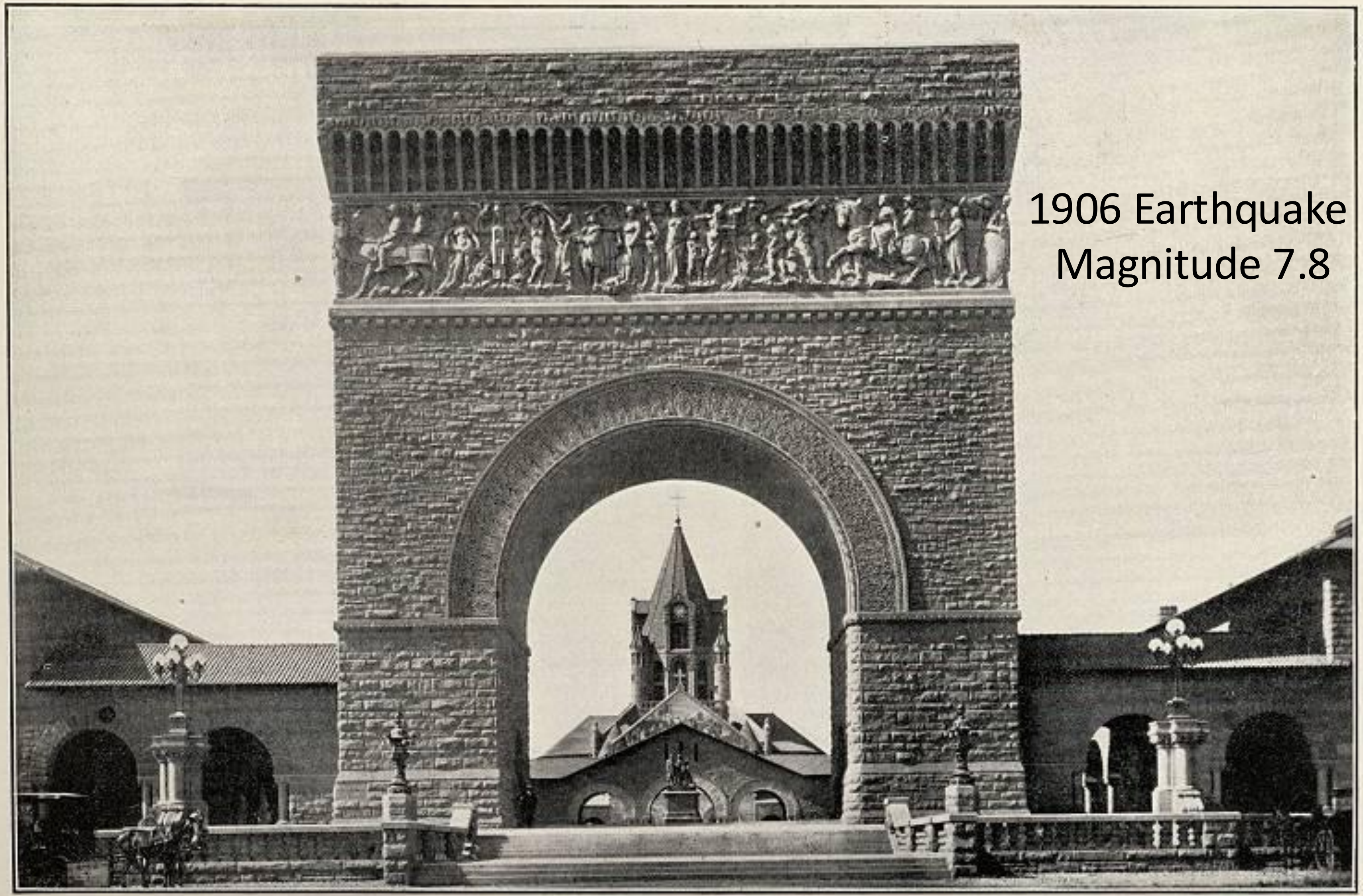
You can vote once a day !!

https://psetapp.stanford.edu/win26/music/

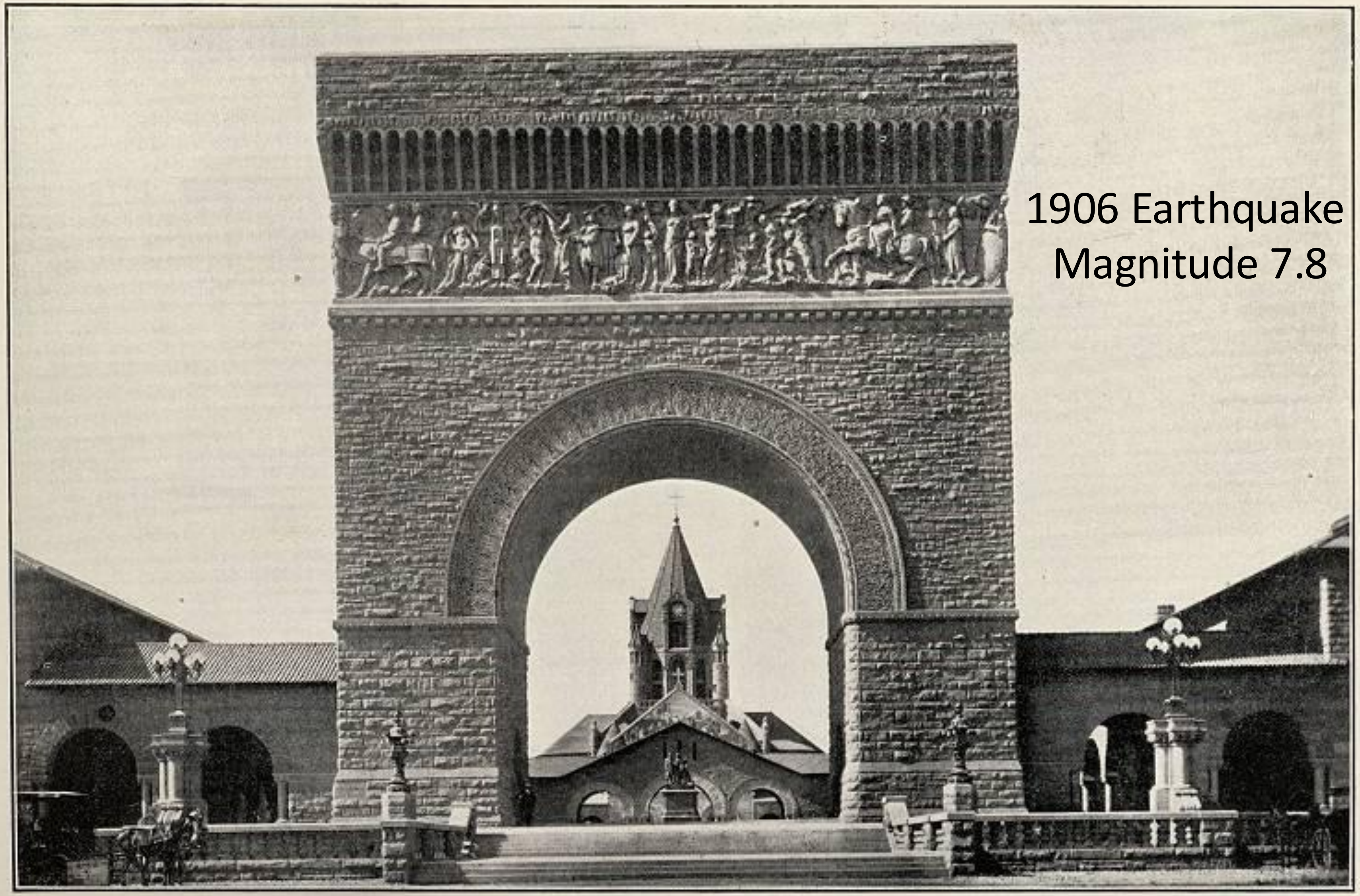


ILL. No. 65. MEMORIAL ARCH, WITH CHURCH IN BACKGROUND, STANFORD UNIVERSITY, SHOWING TYPES OF CARVED WORK WITH THE SANDSTONE.

1906 Earthquake
Magnitude 7.8



ILL. No. 65. MEMORIAL ARCH, WITH CHURCH IN BACKGROUND, STANFORD UNIVERSITY, SHOWING TYPES OF CARVED WORK WITH THE SANDSTONE.



ILL. No. 65. MEMORIAL ARCH, WITH CHURCH IN BACKGROUND, STANFORD UNIVERSITY, SHOWING TYPES OF CARVED WORK WITH THE SANDSTONE.

1906 Earthquake
Magnitude 7.8

How long until the next “big one”?

Review

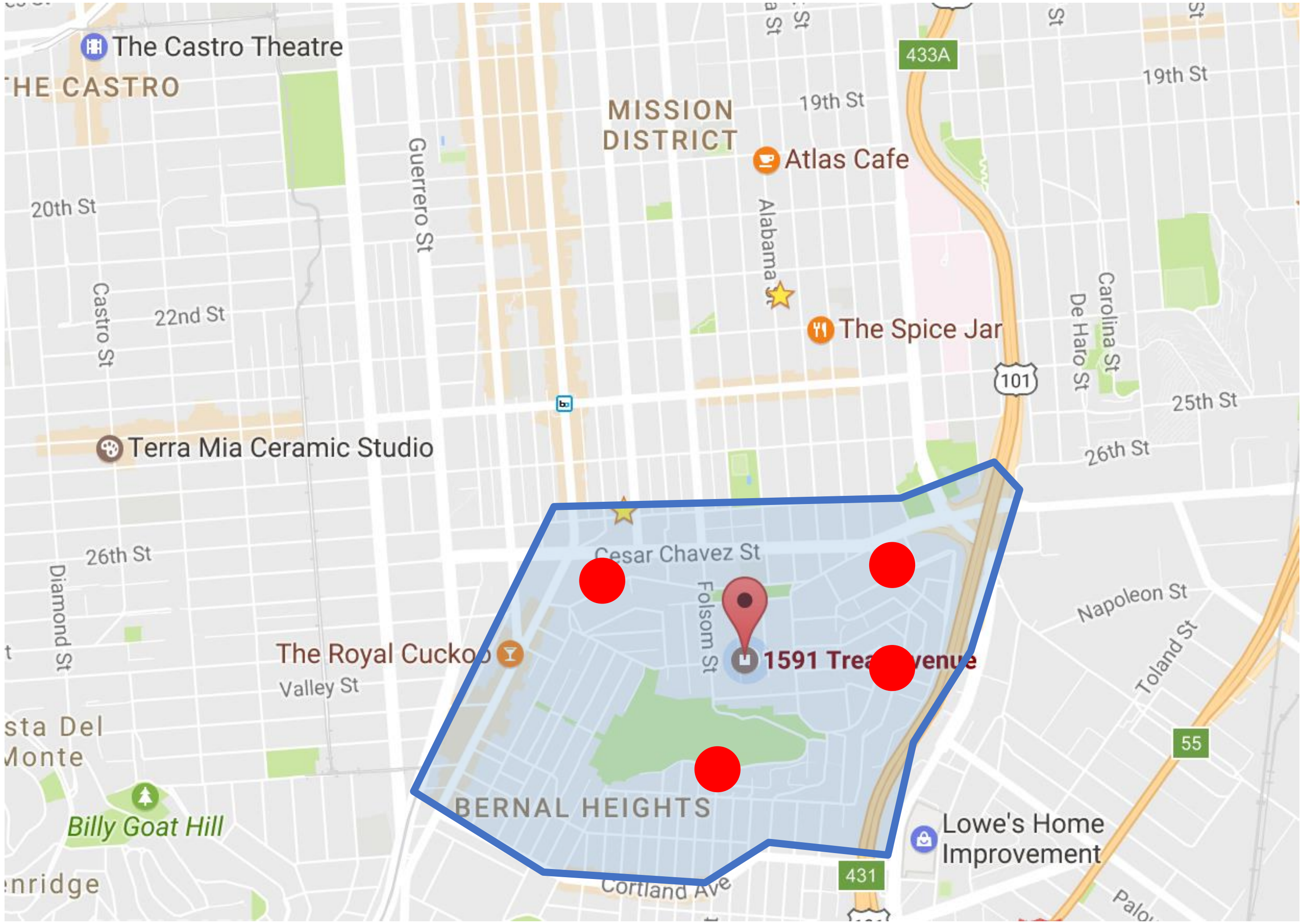
Binomial Random Variable

```
(H, H, H, H, T, T, T, T, T, T)
(H, H, H, T, H, T, T, T, T, T)
(H, H, H, T, T, H, T, T, T, T)
(H, H, H, T, T, T, H, T, T, T)
(H, H, H, T, T, T, T, H, T, T)
(H, H, H, T, T, T, T, T, H, T)
(H, H, H, T, T, T, T, T, T, H)
(H, H, T, H, H, T, T, T, T, T)
(H, H, T, H, T, H, T, T, T, T)
(H, H, T, H, T, T, H, T, T, T)
(H, H, T, H, T, T, T, H, T, T)
(H, H, T, H, T, T, T, T, H, T)
(H, H, T, H, T, T, T, T, T, H)
(H, H, T, T, H, H, T, T, T, T)
(H, H, T, T, H, T, H, T, T, T)
(H, H, T, T, H, T, T, H, T, T)
(H, H, T, T, H, T, T, T, H, T)
(H, H, T, T, H, T, T, T, T, H)
```

The number of **successes**, in n independent **trials**, where each **trial** is a **success** with probability p:

Poisson Random Variable

Probability of ***k* requests** from this area in the next 1 min



Poisson Random Variable

Poisson Random Variable

Notation: $X \sim \text{Poi}(\lambda)$

Description: Number of events in a fixed time frame if (a) the events occur with a constant mean rate and (b) they occur independently of time since last event.

Parameters: $\lambda \in \{0, 1, \ldots\}$, the constant average rate.

Support: $x \in \{0, 1, \ldots\}$

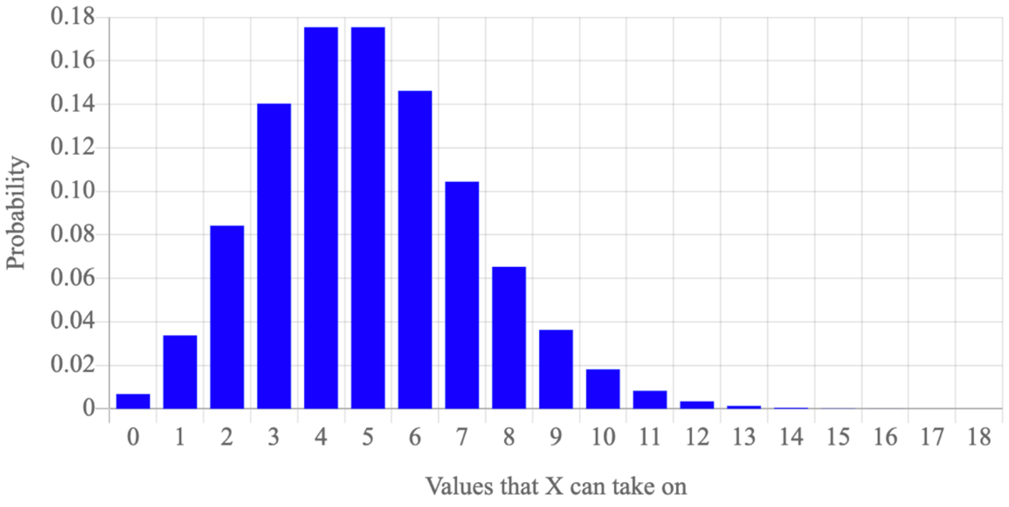
PMF equation: $\Pr(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}$

Expectation: $\mathrm{E}[X] = \lambda$

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

PMF graph:

Parameter λ: 5



Discrete Random Variables

$X \sim \text{Bern}(p)$

Successes in one trial

$X \sim \text{Geo}(p)$

Trials until one success

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

Events in one time interval

$Y \sim \text{Bin}(n, p)$

Successes in *n* trials

$Y \sim \text{NegBin}(r, p)$

Trials until *r* success

Expected Value

The probability of that value

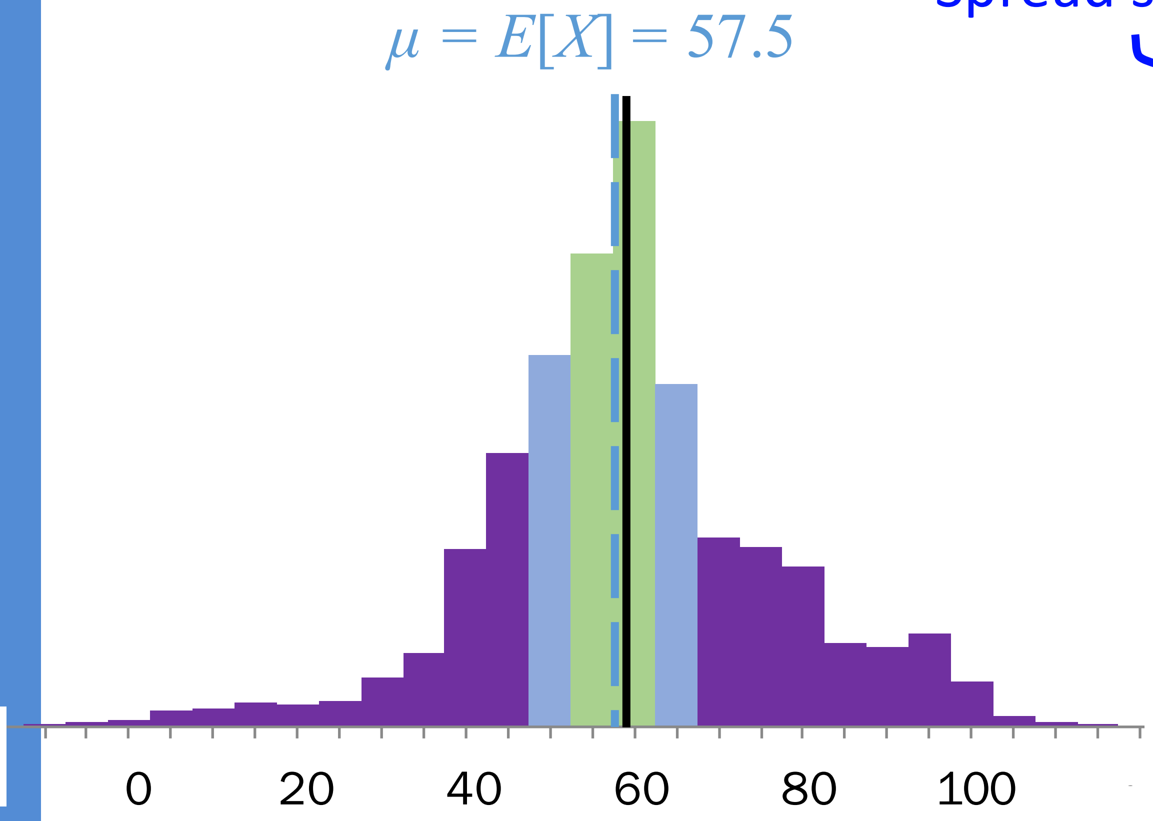
The value

$$E[X] = \sum_x x \cdot \mathrm{P}(X = x)$$

Loop over all values x that X can take on

How Should We Measure Spread? - Variance

Let X be a random variable



Spread stat.

On average..

distance

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

The random variable X

The mean of X

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

Properties of Variance

Variance of a Sum is a sum of Variances

Only if X_is are independent!

$$\mathrm{Var}(X) = \mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}(X_i)$$

Linearity of Variance

$$\mathrm{Var}(aX + b) = a^2 \mathrm{Var}(X)$$

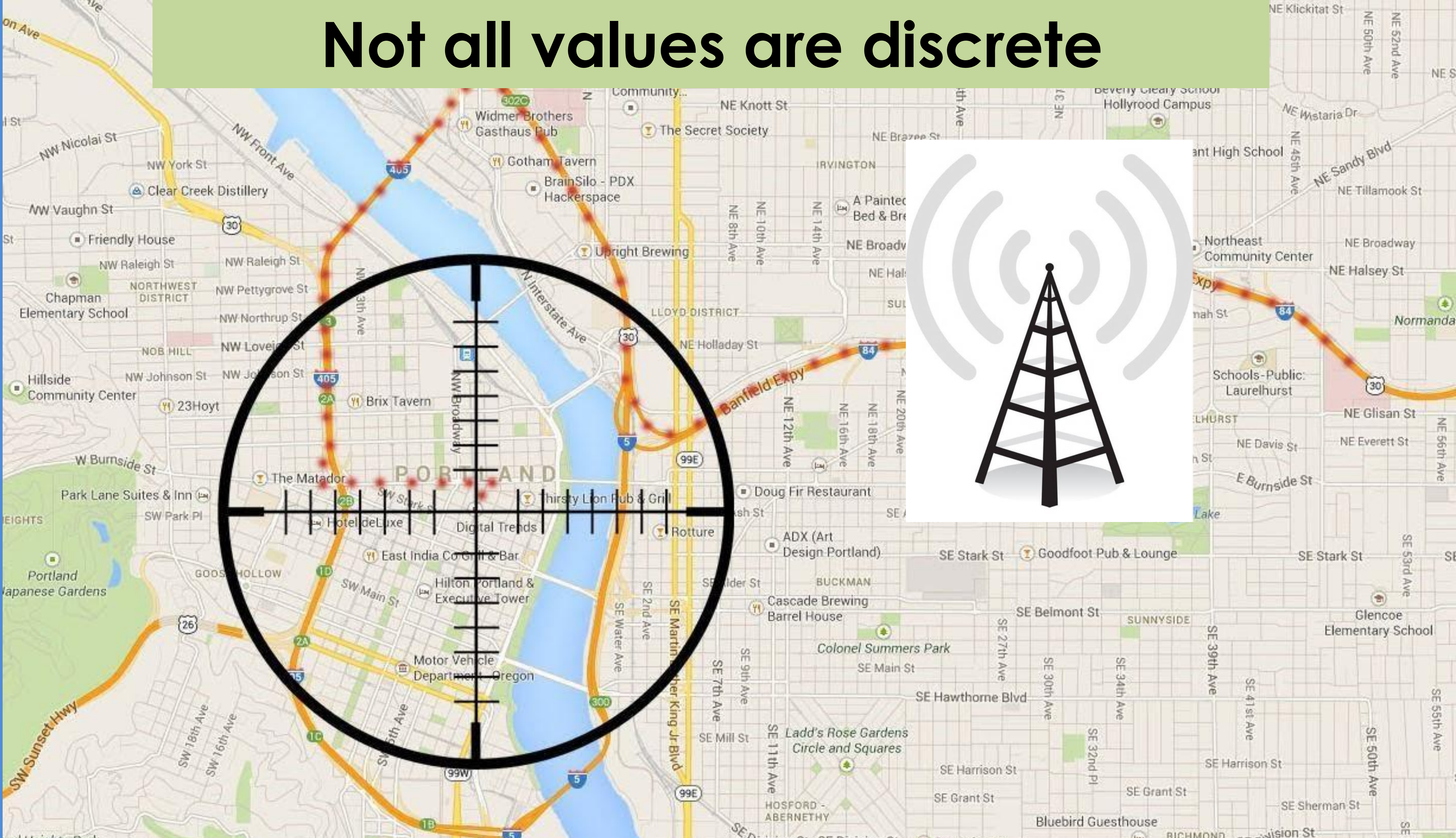
End Review

Learning Goals

1. Integrate a density function (PDF) to get a probability
2. Use a cumulative distribution function (CDF) to get a probability

Big hole in our knowledge

Not all values are discrete



Can’t Talk About Continuous Values

Say the average rate of earthquakes is 1 every 100 years.

We **can** talk about the probability distribution of different numbers of earthquakes next year.

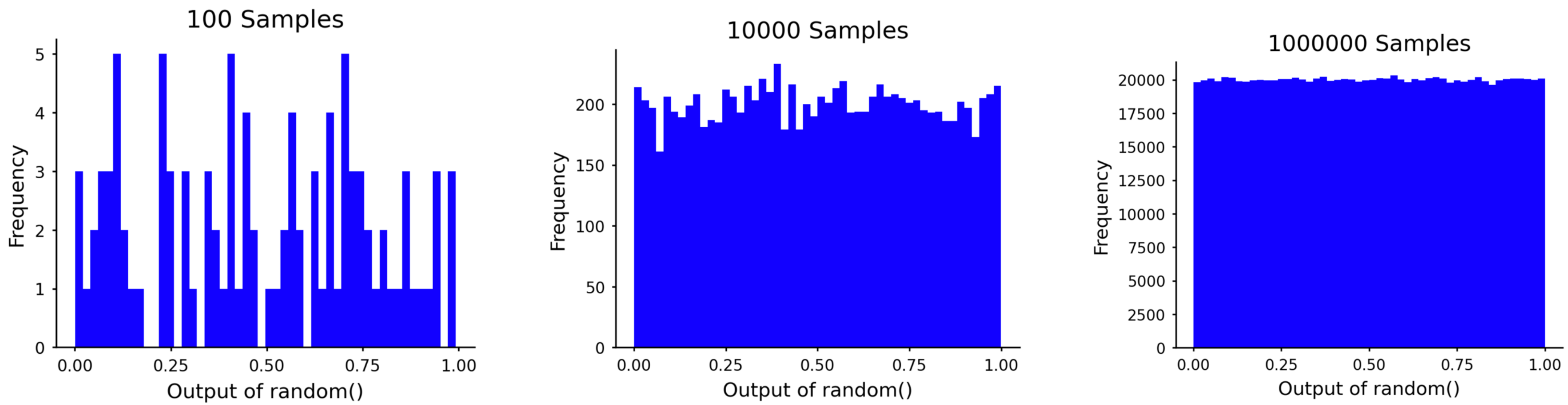
We **can’t** talk about the probability distribution of the amount of time until the next earthquake.

`random()?`

The `random()` Function

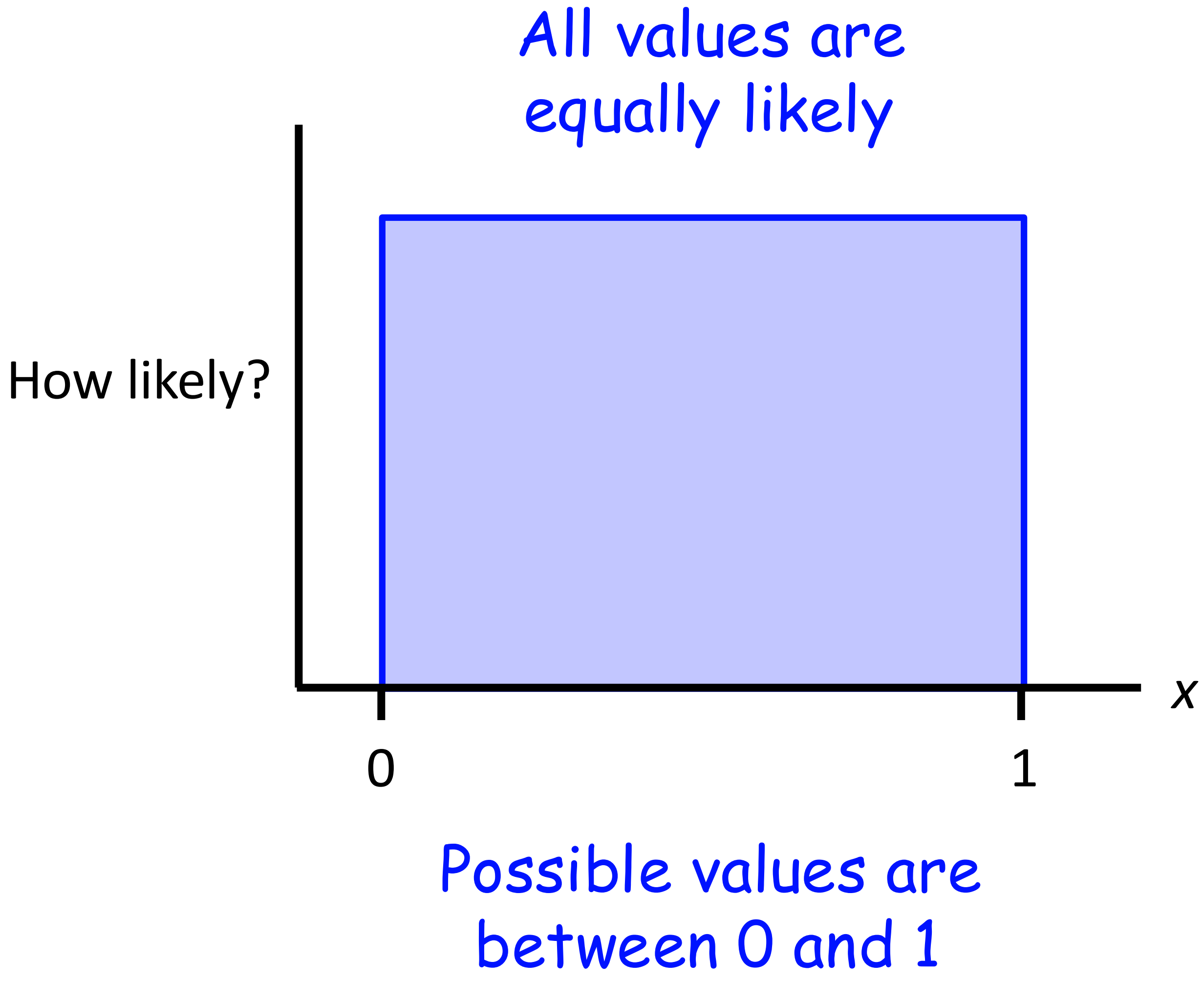
- Outputs values between 0 and 1
- All possible values are equally likely
- This is a continuous random variable!

```python
import random

samples_small = []
for i in range(100):
    samples_small.append(random.random())

samples_medium = []
for i in range(10000):
    samples_medium.append(random.random())

samples_large = []
for i in range(1000000):
    samples_large.append(random.random())
```

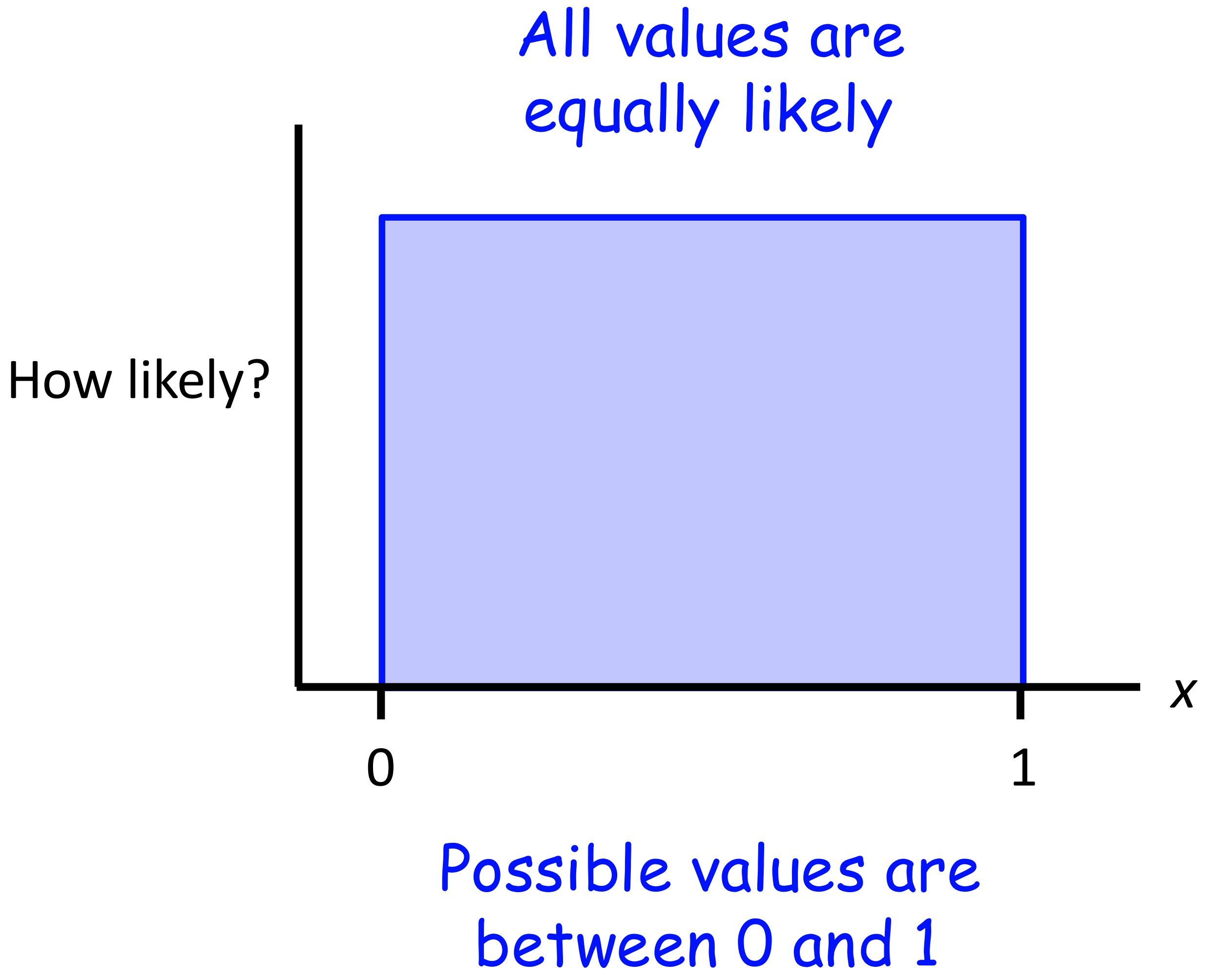


$X \sim$ Uniform(0,1): A Continuous Random Variable

$X \sim \text{Uniform}(0,1)$: A Continuous Random Variable

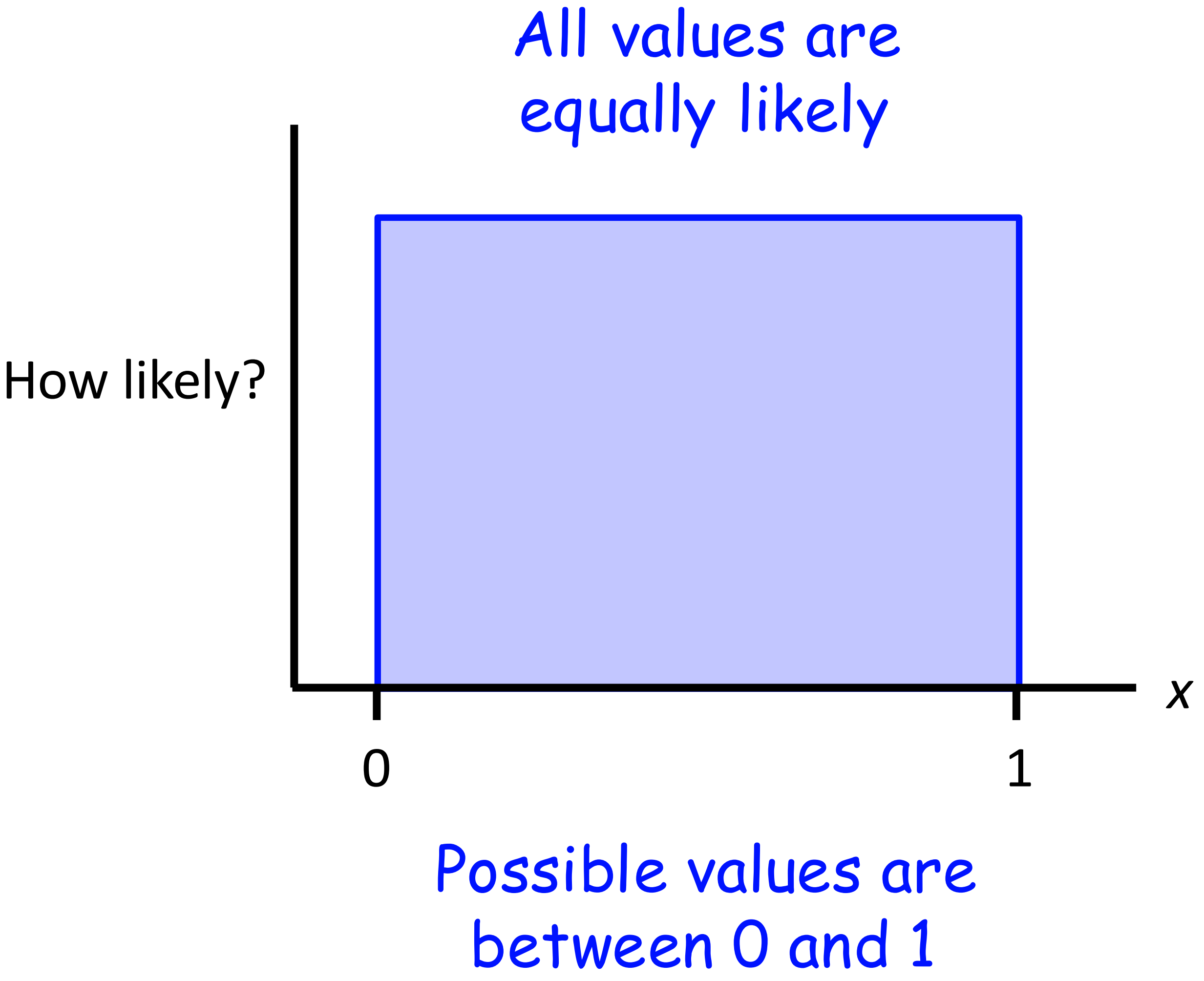


$X \sim \text{Uniform}(0,1)$: A Continuous Random Variable



$$P(0 \leq X \leq 1) = \ ?$$

$X \sim \text{Uniform}(0,1)$: A Continuous Random Variable

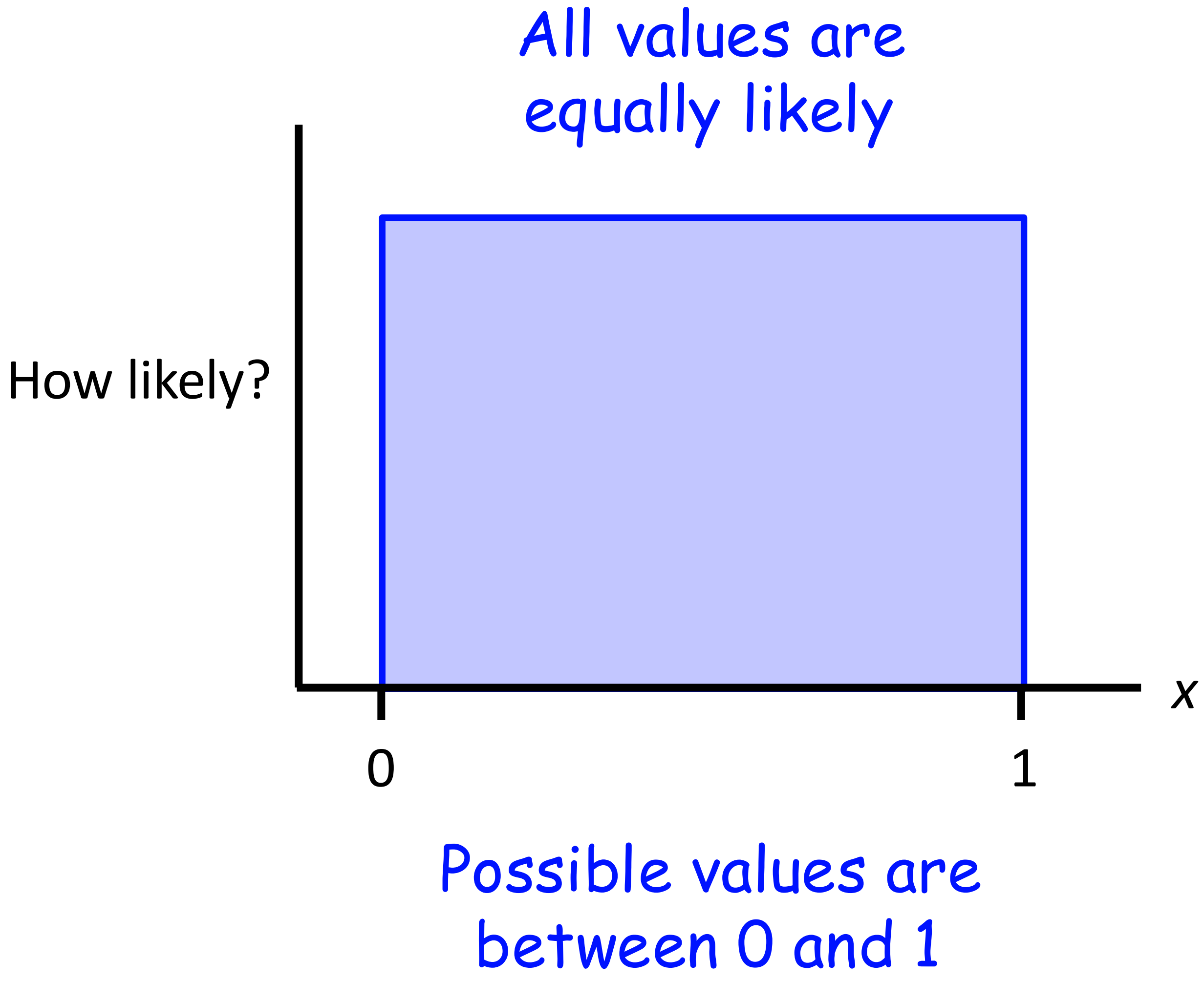


$$P(0 \leq X \leq 1) = 1$$

Probability of the whole sample space must equal 1

(Axiom 2)

$X \sim \text{Uniform}(0,1)$: A Continuous Random Variable

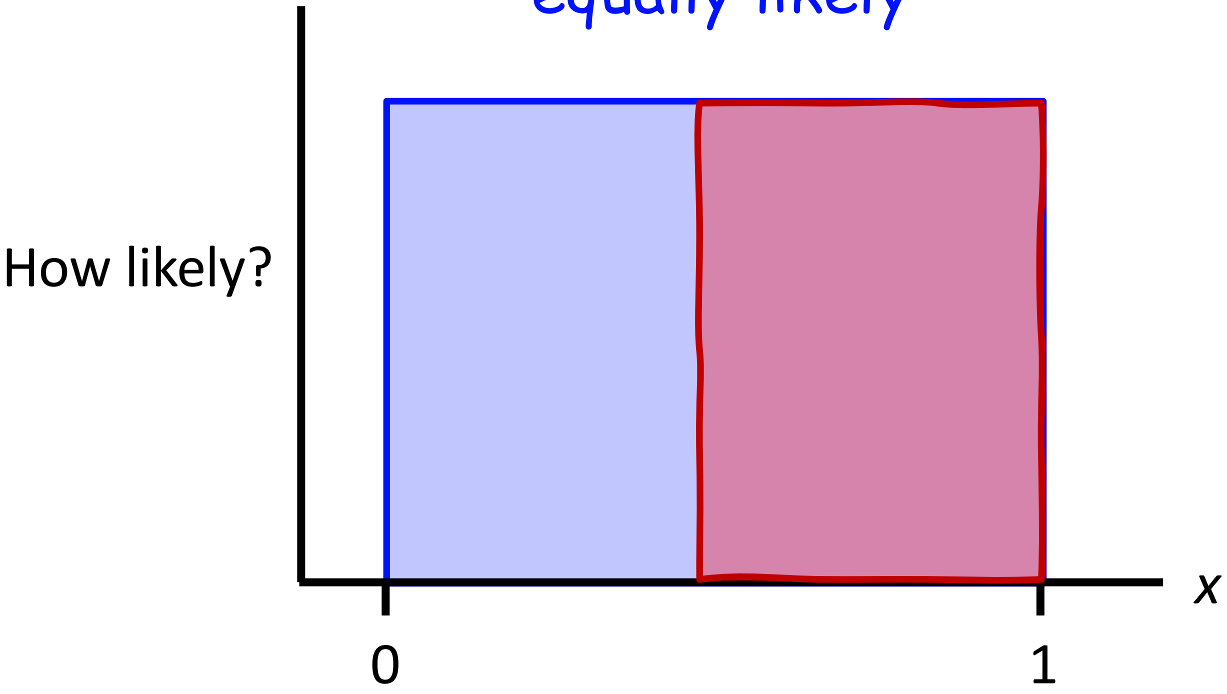


$$P(0 \leq X \leq 1) = 1$$

$$P(0.5 \leq X \leq 1) = ?$$

$X \sim \text{Uniform}(0,1)$: A Continuous Random Variable

All values are
equally likely



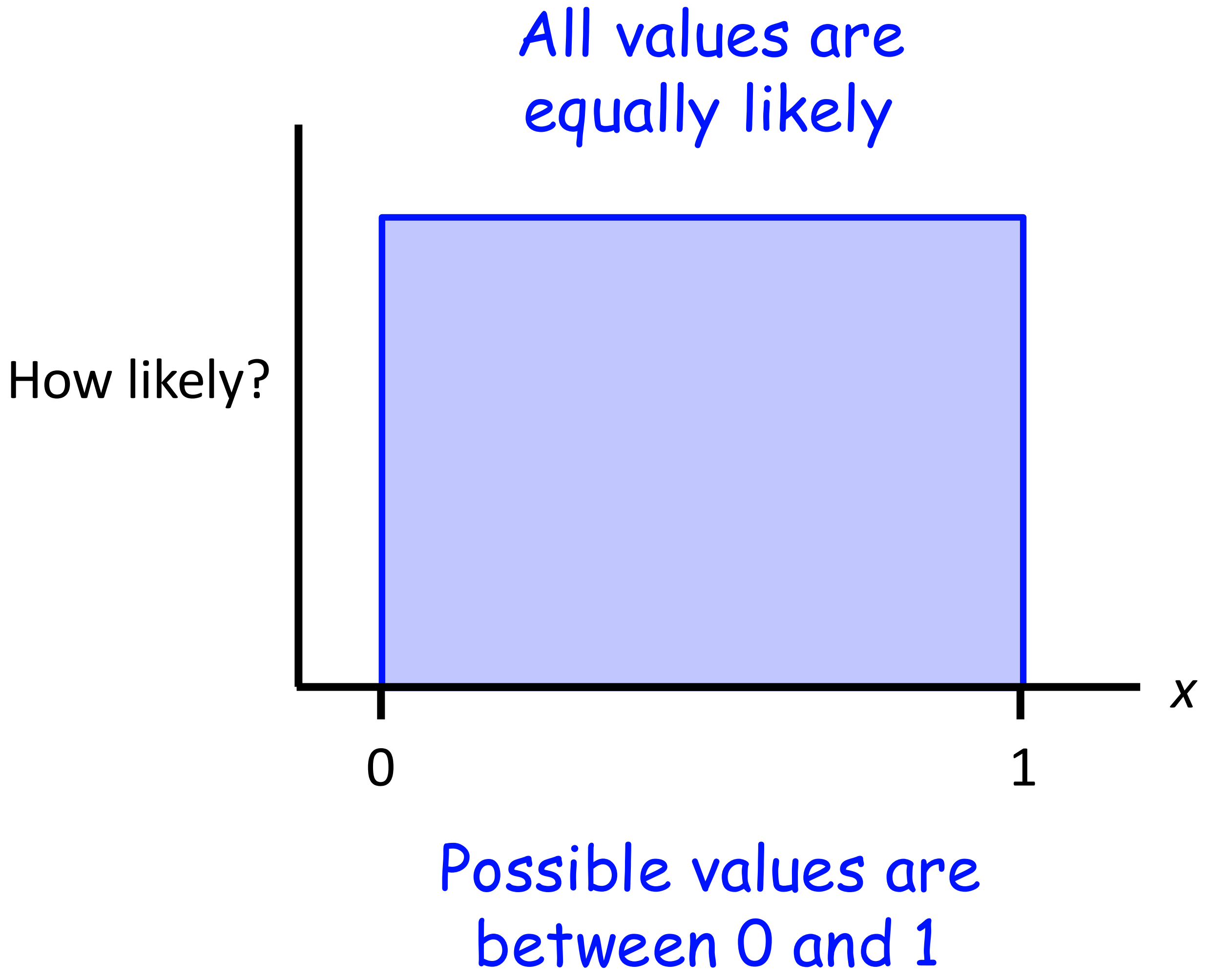
Possible values are
between 0 and 1

$$P(0 \leq X \leq 1) = 1$$

$$P(0.5 \leq X \leq 1) = 0.5$$

Half of all possible outcomes
are between 0.5 and 1

$X \sim$ Uniform(0,1): A Continuous Random Variable

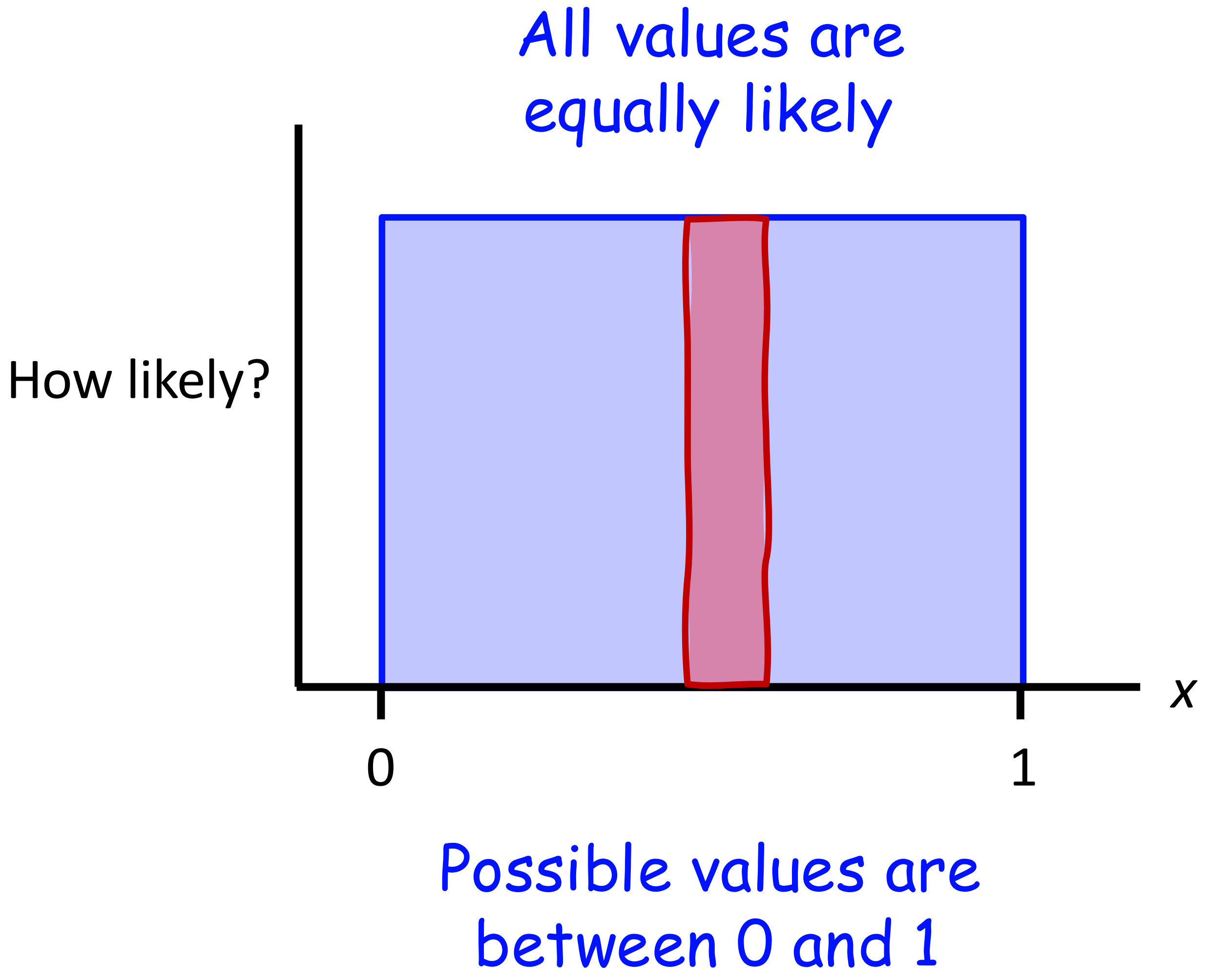


$$P(0 \leq X \leq 1) = 1$$

$$P(0.5 \leq X \leq 1) = 0.5$$

$$P(0.5 \leq X \leq 0.6) = ?$$

$X \sim \text{Uniform}(0,1)$: A Continuous Random Variable



$$P(0 \leq X \leq 1) = 1$$

$$P(0.5 \leq X \leq 1) = 0.5$$

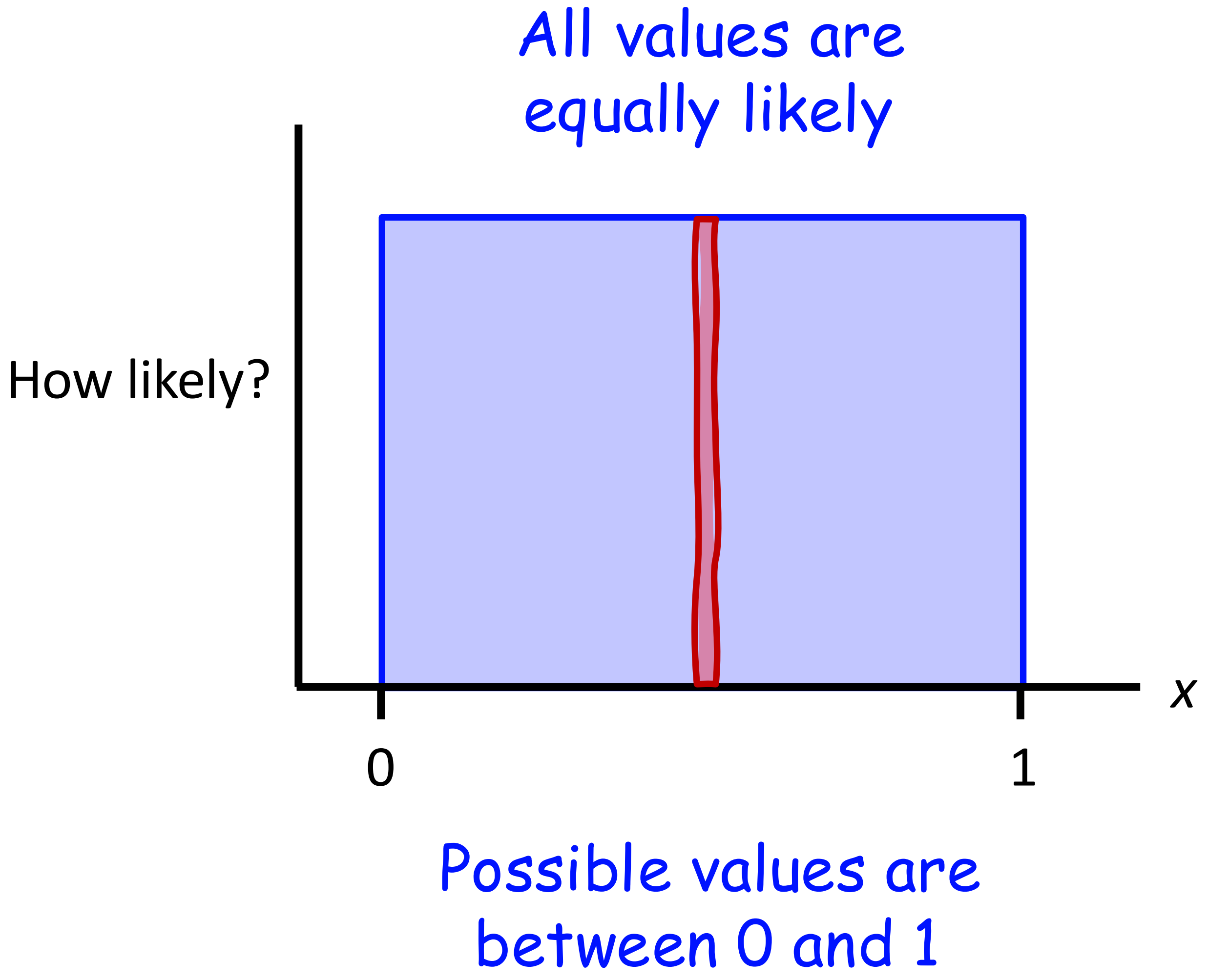
$$P(0.5 \leq X \leq 0.6) = 0.1$$

1/10 of all possible outcomes are between 0.5 and 0.6

So far, the pattern looks like:

P(start ≤ X ≤ end) = end - start

$X \sim \text{Uniform}(0,1)$: A Continuous Random Variable



$$P(0 \leq X \leq 1) = 1$$

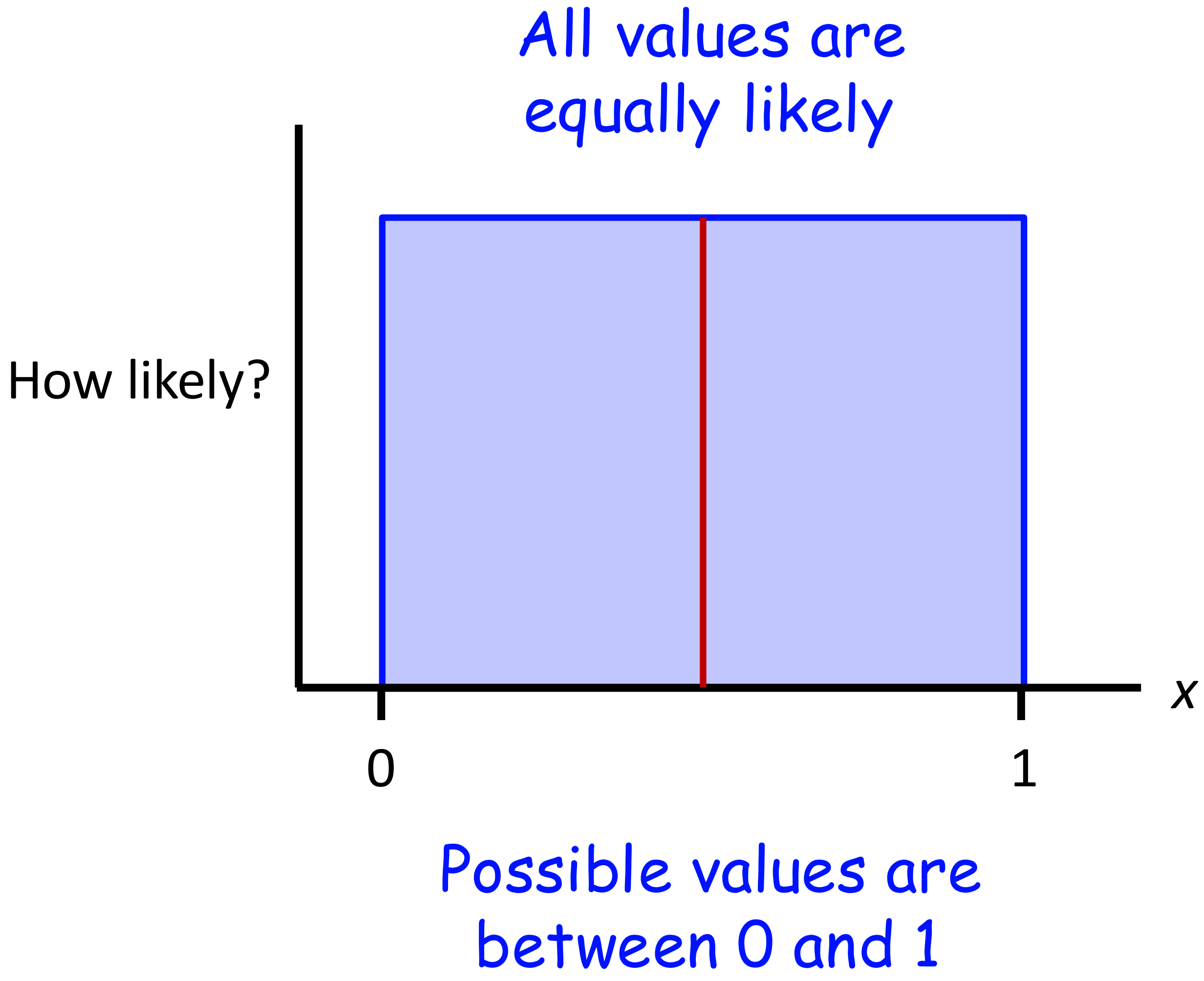
$$P(0.5 \leq X \leq 1) = 0.5$$

$$P(0.5 \leq X \leq 0.6) = 0.1$$

$$P(0.5 \leq X \leq 0.5001) = 0.0001$$

As we get more precise, probabilities keep shrinking…

$X \sim \text{Uniform}(0,1)$: A Continuous Random Variable



$$P(0 \le X \le 1) = 1$$

$$P(0.5 \le X \le 1) = 0.5$$

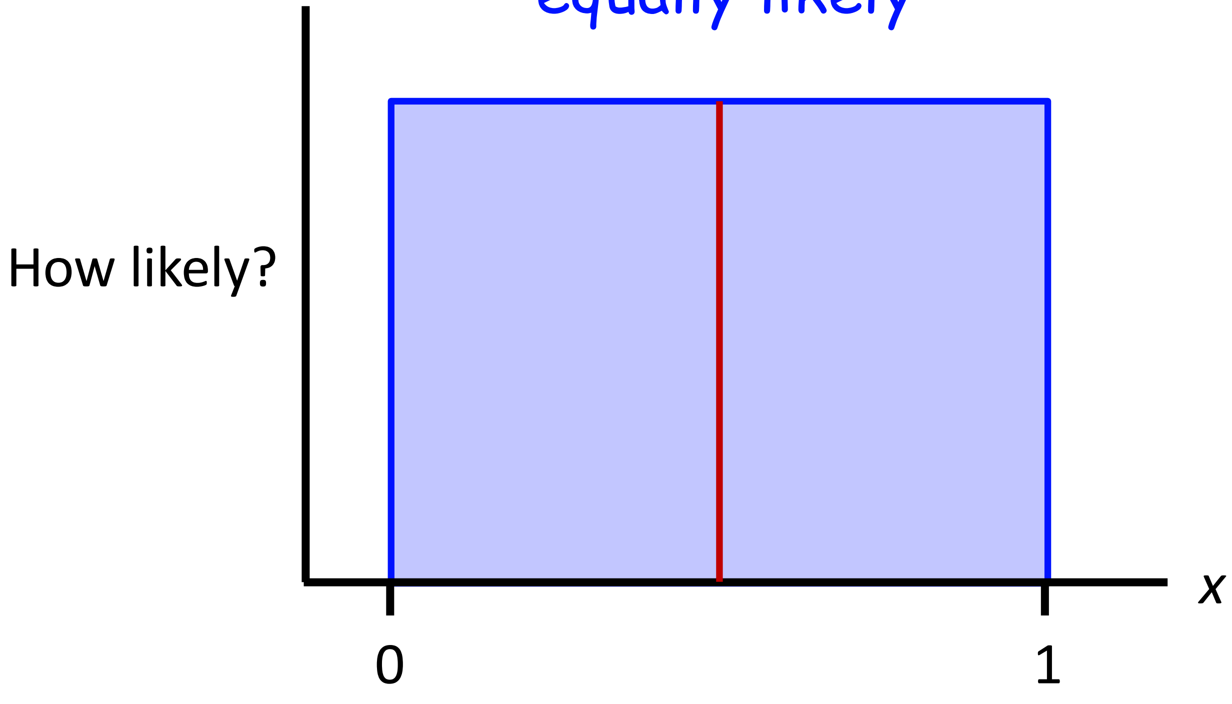
$$P(0.5 \le X \le 0.6) = 0.1$$

$$P(0.5 \le X \le 0.5001) = 0.0001$$

$$P(X = 0.5) = ?$$

$X \sim$ Uniform(0,1): A Continuous Random Variable

All values are equally likely



Possible values are between 0 and 1

$$P(0 \leq X \leq 1) = 1$$

$$P(0.5 \leq X \leq 1) = 0.5$$

$$P(0.5 \leq X \leq 0.6) = 0.1$$

$$P(0.5 \leq X \leq 0.5001) = 0.0001$$

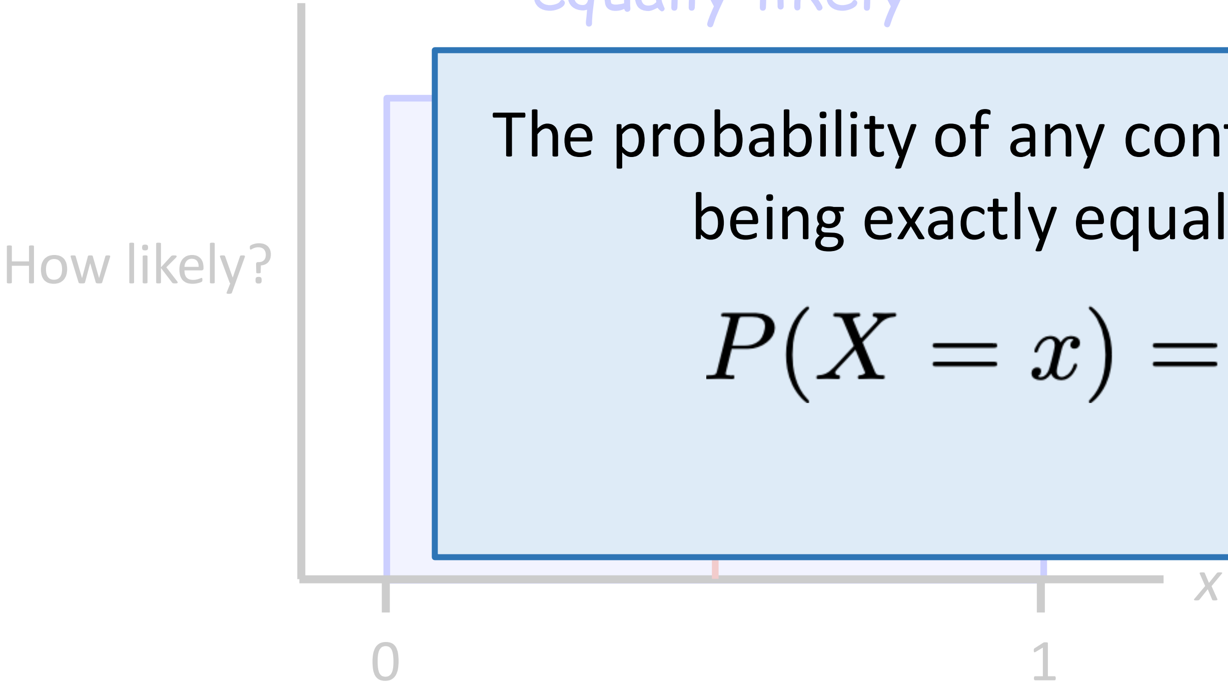
$$P(X = 0.5) = 0$$

The probability of any exact outcome, with infinite precision…is zero

$X \sim$ Uniform(0,1): A Continuous Random Variable

All values are equally likely

$P(0 \le X \le 1) = 1$



The probability of any continuous random variable being exactly equal to any value is 0.

$$P(X = x) = 0, \text{ for all } x$$

$= 0.5$

$= 0.1$

$= 0.0001$

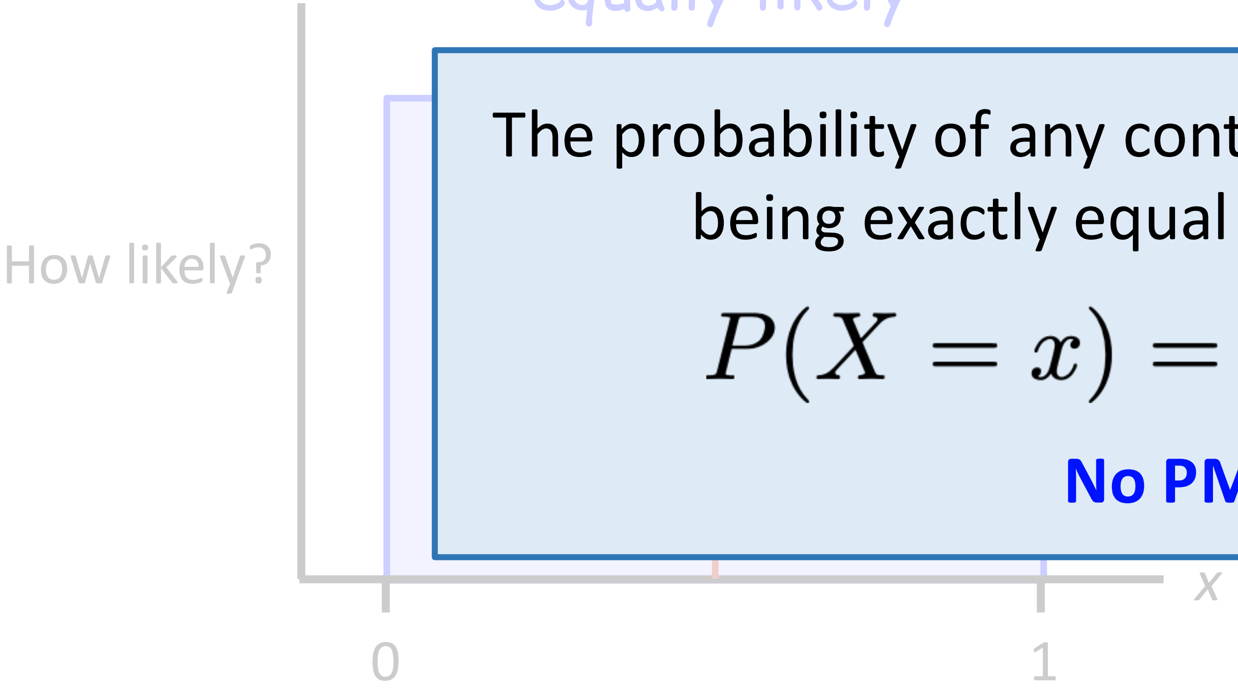
$P(X = 0.5) = 0$

Possible values are between 0 and 1

The probability of any exact outcome, with infinite precision…is zero

$X \sim$ Uniform(0,1): A Continuous Random Variable

All values are equally likely



Possible values are between 0 and 1

The probability of any continuous random variable being exactly equal to any value is 0.

$$P(X = x) = 0, \text{ for all } x$$

No PMFs!

$$P(0 \leq X \leq 1) = 1$$

$= 0.5$

$= 0.1$

$= 0.0001$

$$P(X = 0.5) = 0$$

The probability of any exact outcome, with infinite precision…is zero

$X \sim \text{Uniform}(0,1)$: A Continuous Random Variable

$$P(0 \le X \le 1) = 1$$

$$P(0.5 \le X \le 1) = 0.5$$

$$P(0.5 \le X \le 0.6) = 0.1$$

$$P(0.5 \le X \le 0.5001) = 0.0001$$

$$P(X = 0.5) = 0$$

ability of any exact outcome,
infinite precision…is zero

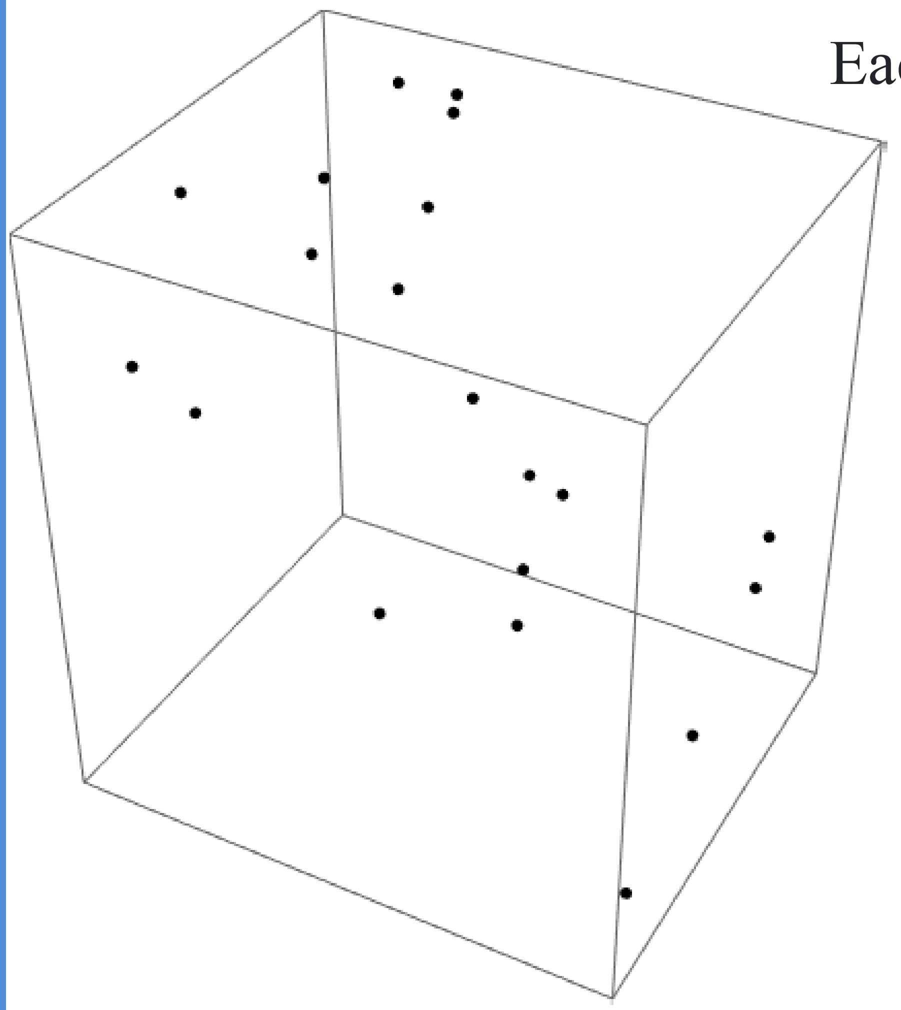
The only way to talk about probabilities of outcomes for *continuous* random variables is using ranges of possible values.

Curse of Dimensionality

A random *point* of dimension d is a list of d random values: $[X_1 \ldots X_d]$

$X_i \sim \text{Uni}(0, 1)$ for all i

Each value X_i is independent of other values



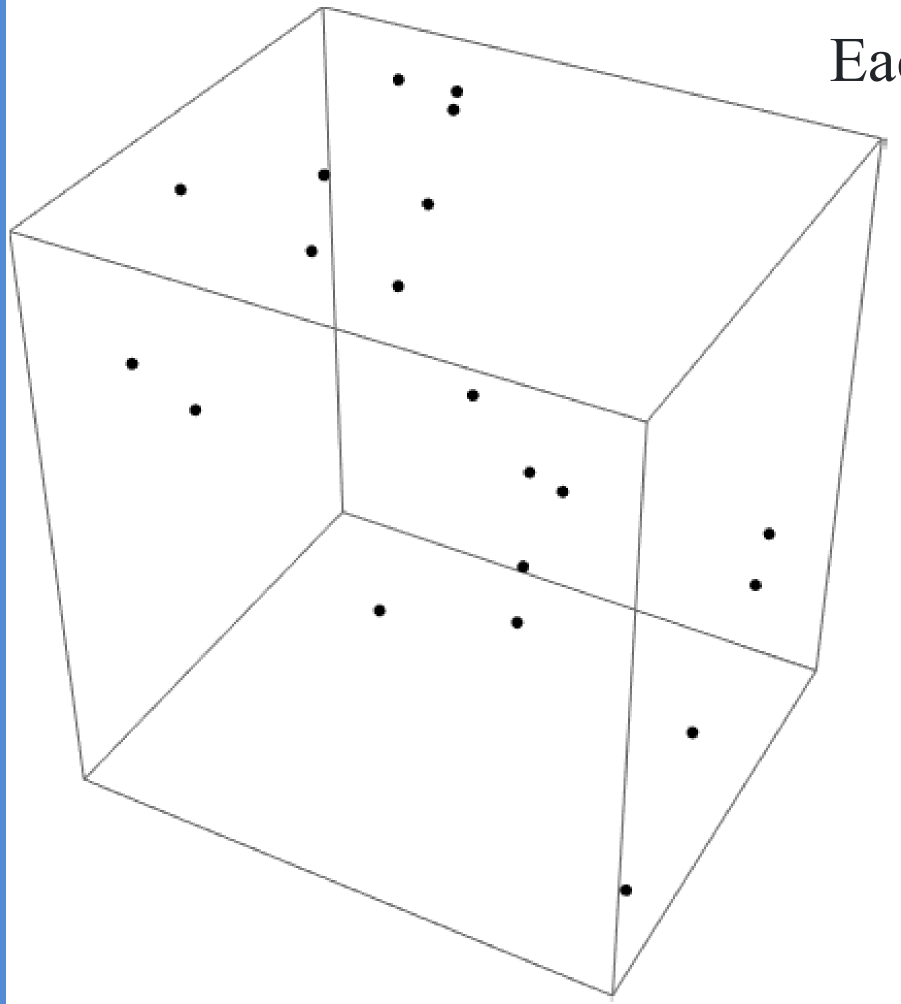
X_i is close to an edge if X_i is less than 0.01 **or** X_i is greater than 0.99. What is the probability that X_i is close to an edge?

Curse of Dimensionality

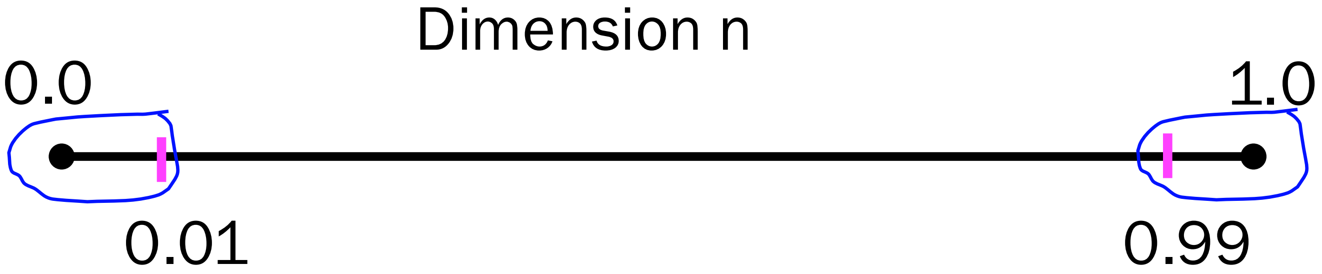
A random *point* of dimension d is a list of d random values: $[X_1 \ldots X_d]$

$X_i \sim \text{Uni}(0, 1)$ for all i

Each value X_i is independent of other values



X_i is close to an edge if X_i is less than 0.01 **or** X_i is greater than 0.99. What is the probability that X_i is close to an edge?

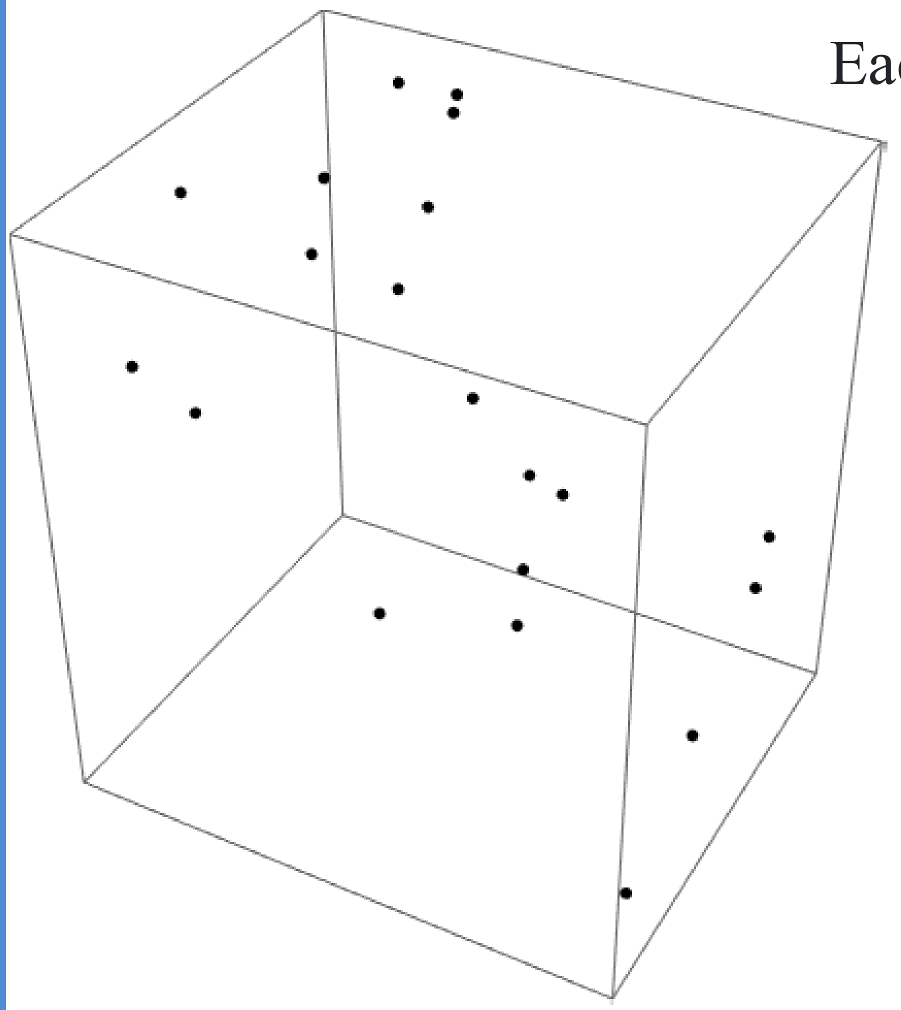


Curse of Dimensionality

A random *point* of dimension d is a list of *d* random values: $[X_1 \dots X_d]$

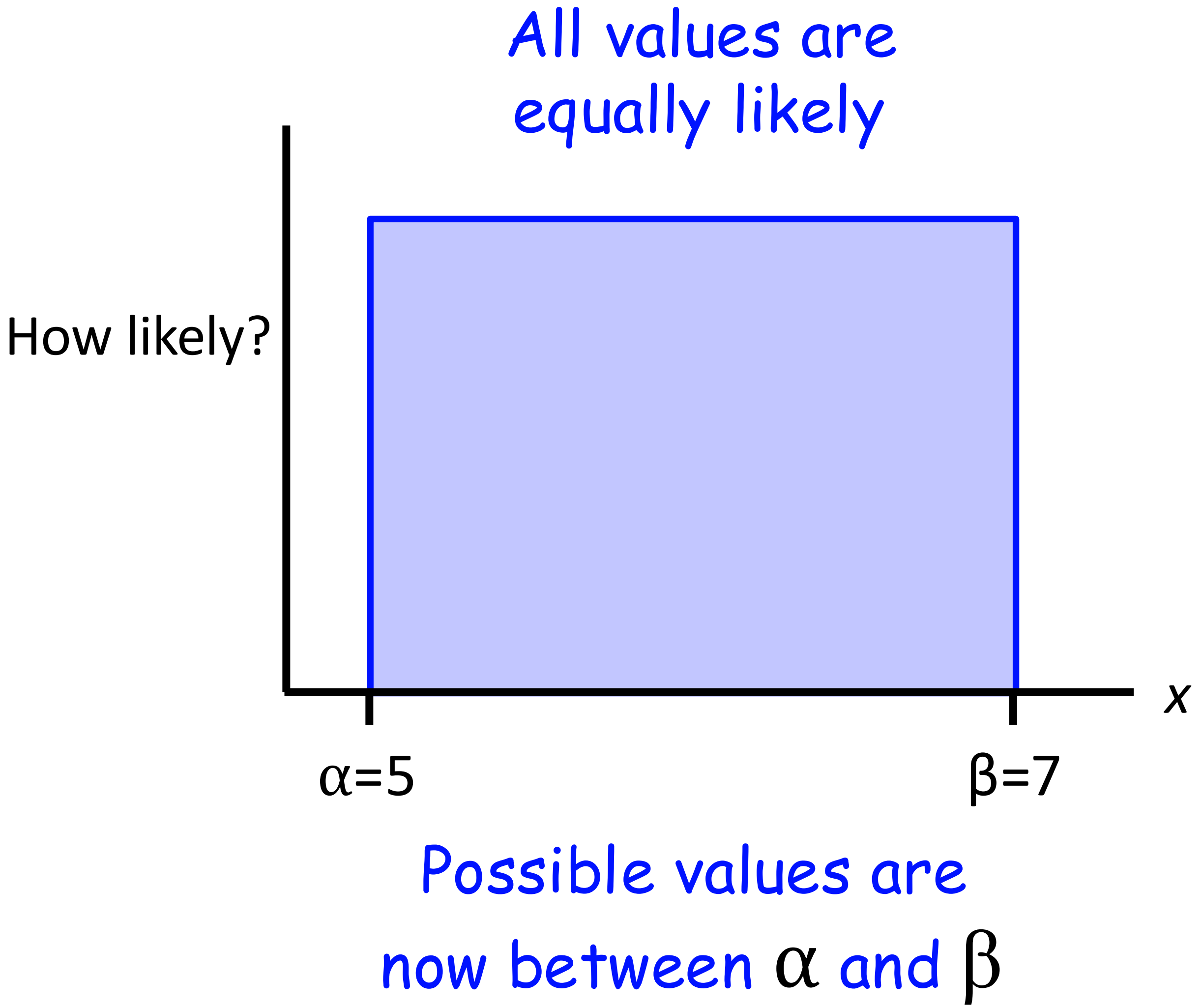
$X_i \sim \text{Uni}(0, 1)$ for all i

Each value X_i is independent of other values



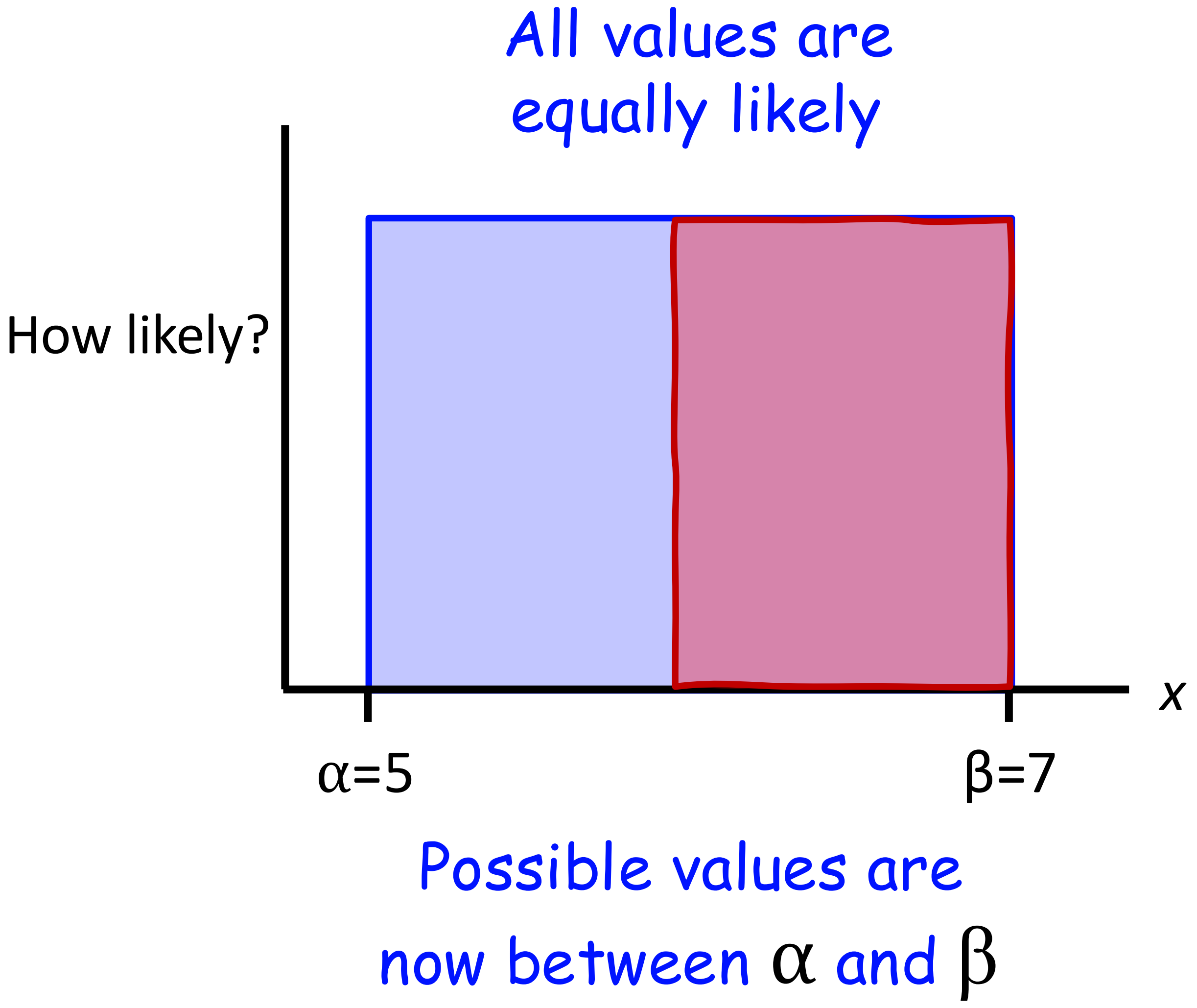
A random *point* $[X_1 \dots X_{100}]$ of dimension 100 is close to an edge if *any* of its values are close to an edge. What is the probability that a 100 dimensional point is close to an edge?

$X \sim$ Uniform(α, β): More General Case



$$P(5 \leq X \leq 7) = 1$$

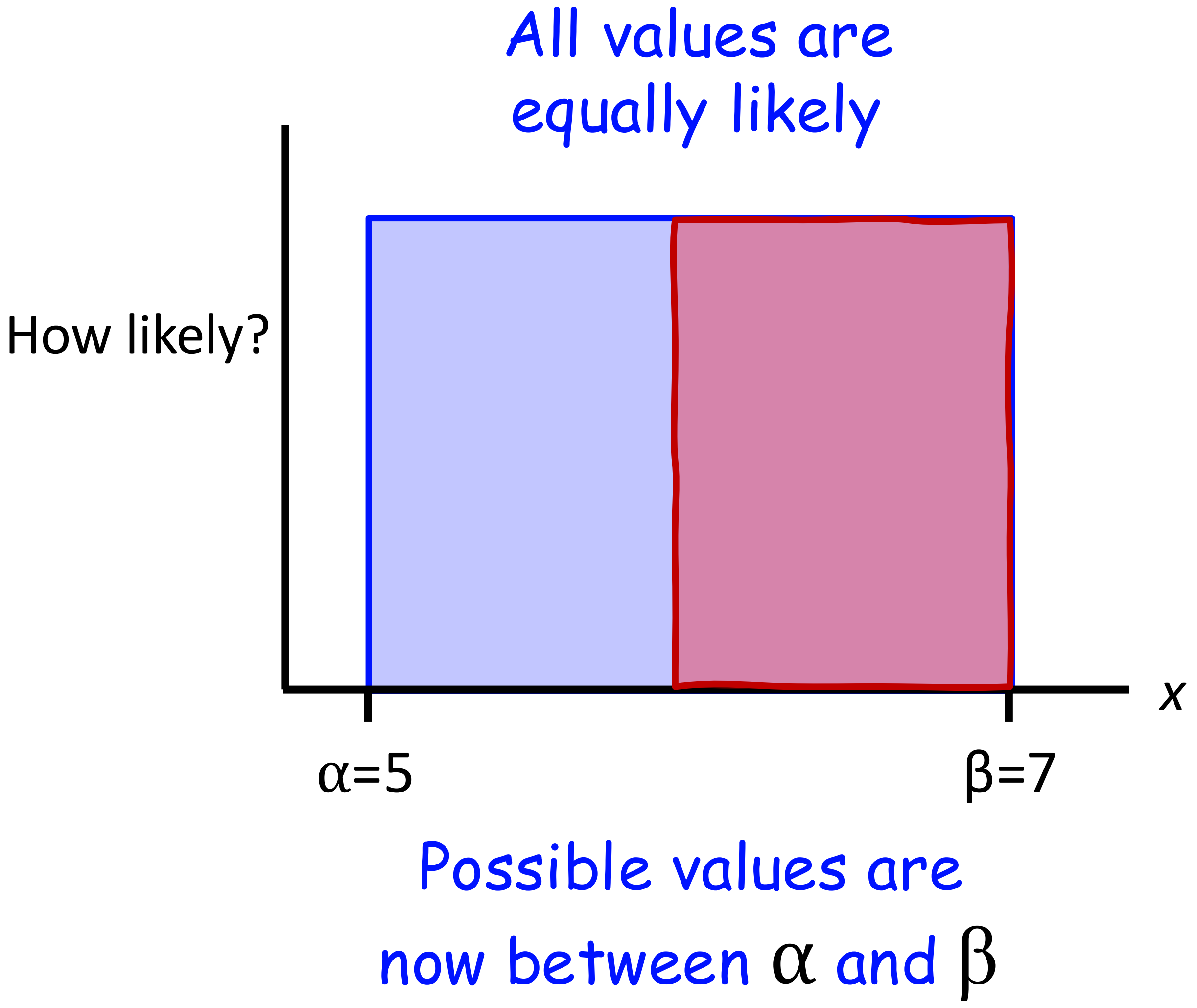
$X \sim \text{Uniform}(\alpha, \beta)$: More General Case



$$P(5 \leq X \leq 7) = 1$$

$$P(6 \leq X \leq 7) = \ ?$$

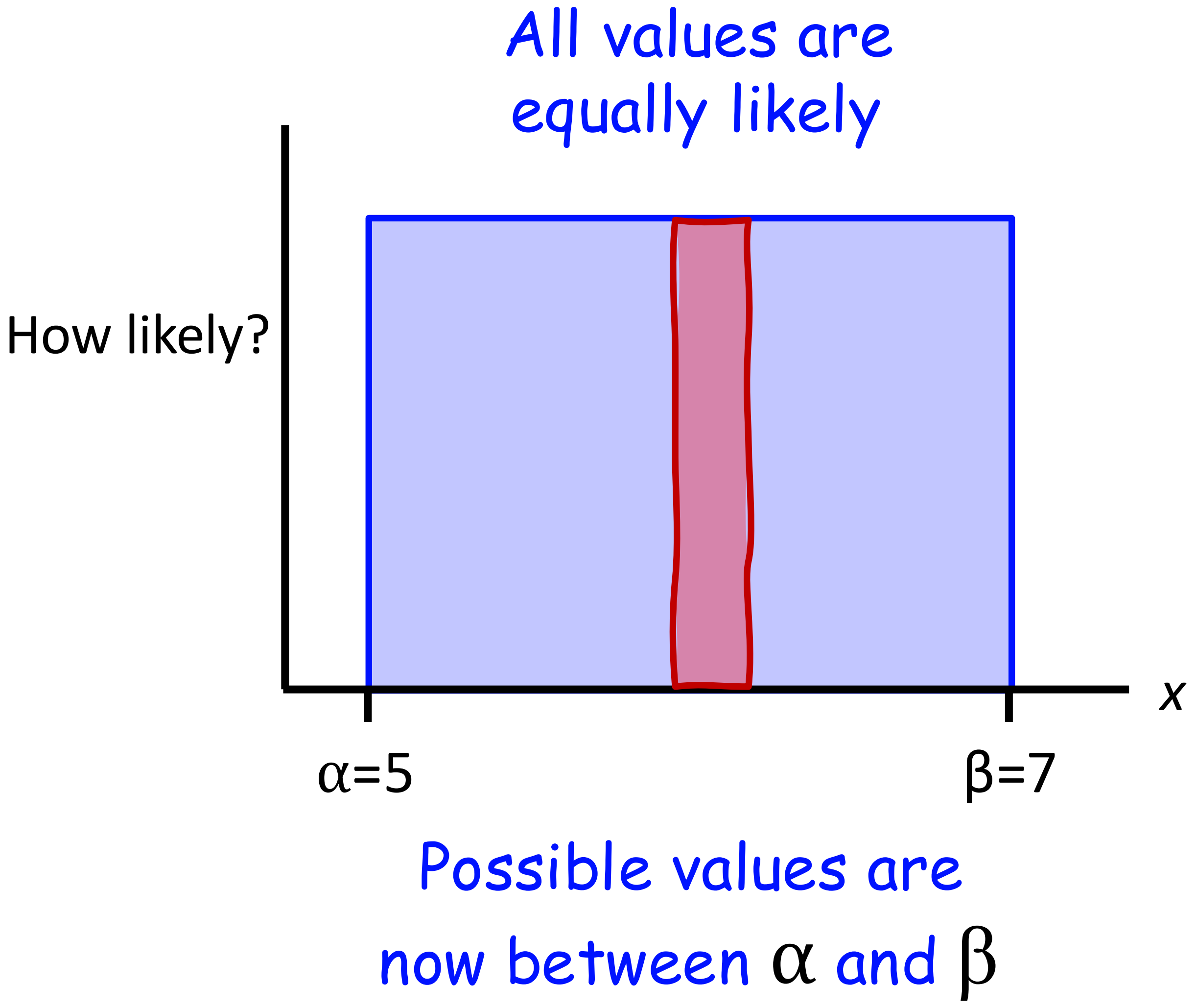
$X \sim$ Uniform(α, β): More General Case



$$P(5 \leq X \leq 7) = 1$$

$$P(6 \leq X \leq 7) = 0.5$$

$X \sim$ Uniform(α, β): More General Case

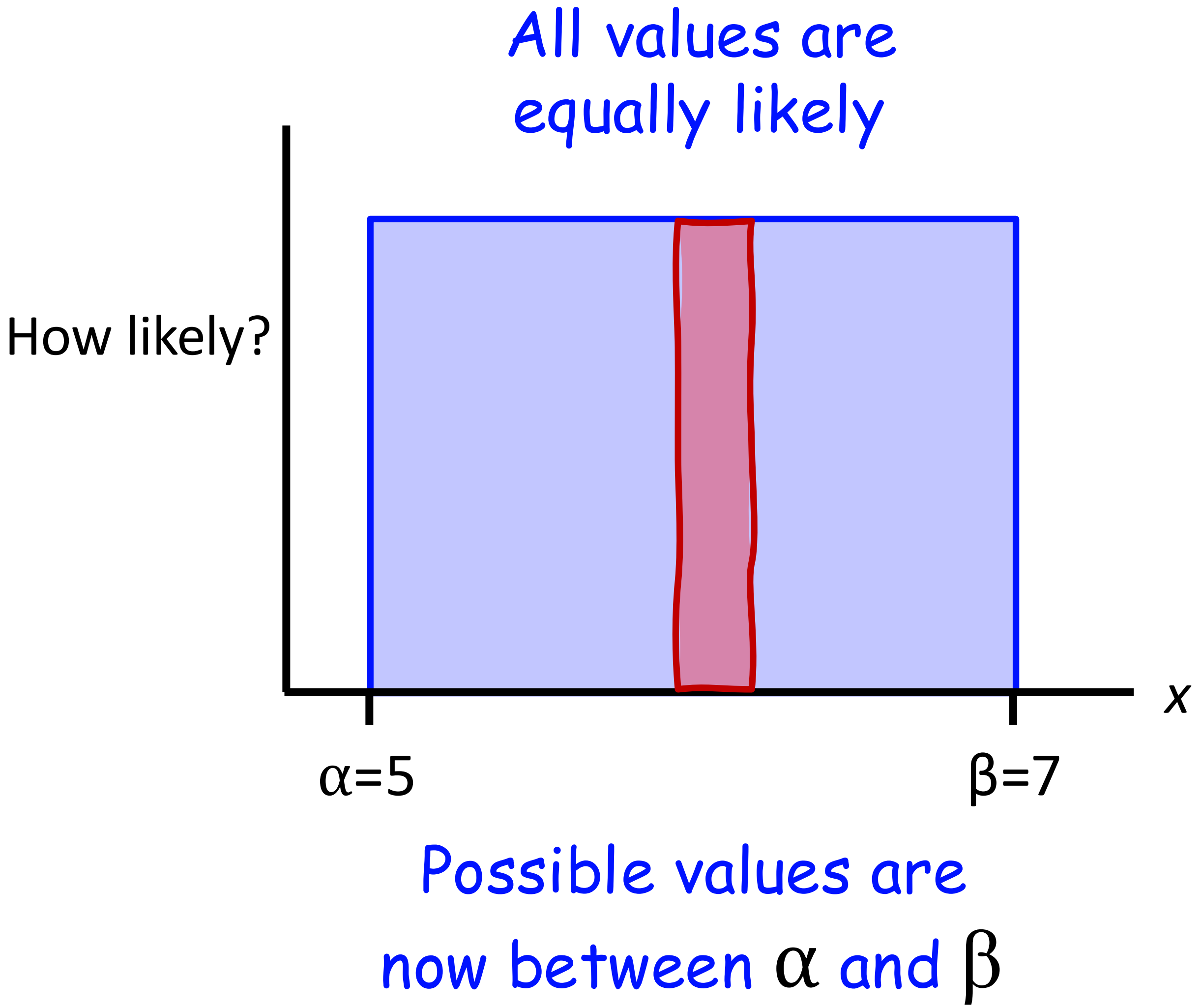


$$P(5 \leq X \leq 7) = 1$$

$$P(6 \leq X \leq 7) = 0.5$$

$$P(6 \leq X \leq 6.1) = \ ?$$

$X \sim$ Uniform(α, β): More General Case

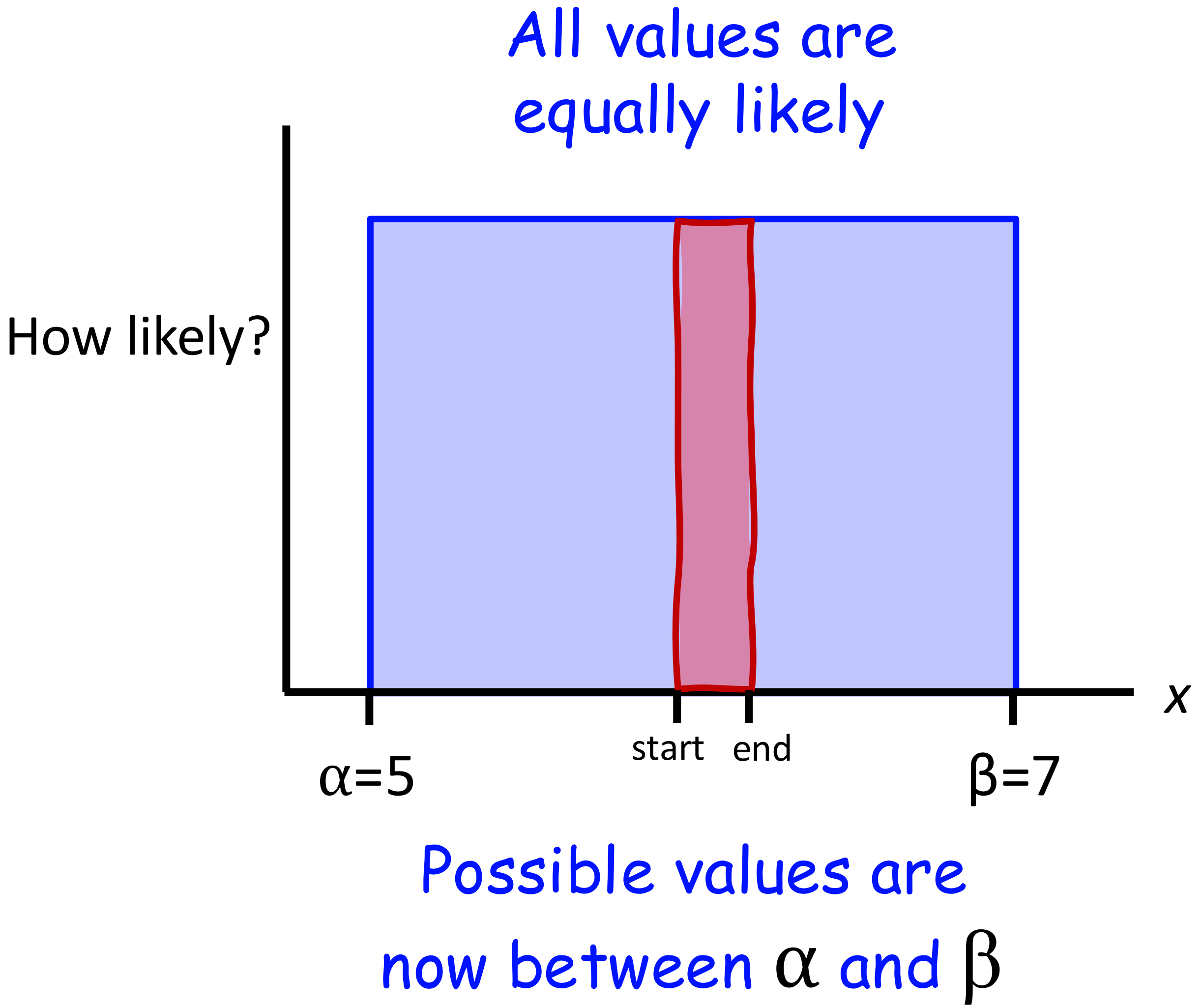


$$P(5 \leq X \leq 7) = 1$$

$$P(6 \leq X \leq 7) = 0.5$$

$$P(6 \leq X \leq 6.1) = 0.05$$

$X \sim$ Uniform(α, β): More General Case



$$P(5 \leq X \leq 7) = 1$$

$$P(6 \leq X \leq 7) = 0.5$$

$$P(6 \leq X \leq 6.1) = 0.05$$

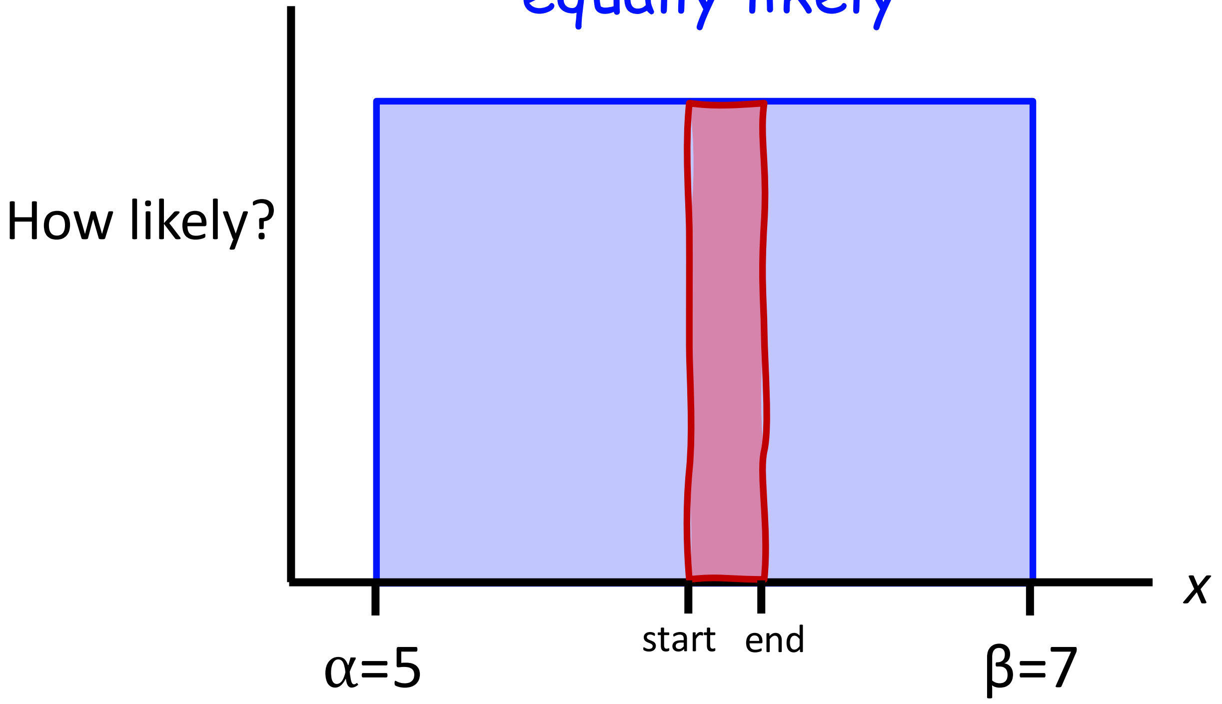
For Uniform(0,1):

P(start ≤ X ≤ end) = end - start

Does that still work?

$X \sim$ Uniform(α, β): More General Case

All values are equally likely



Possible values are now between α and β

$$P(5 \leq X \leq 7) = 1$$

$$P(6 \leq X \leq 7) = 0.5$$

$$P(6 \leq X \leq 6.1) = 0.05$$

For Uniform(0,1):

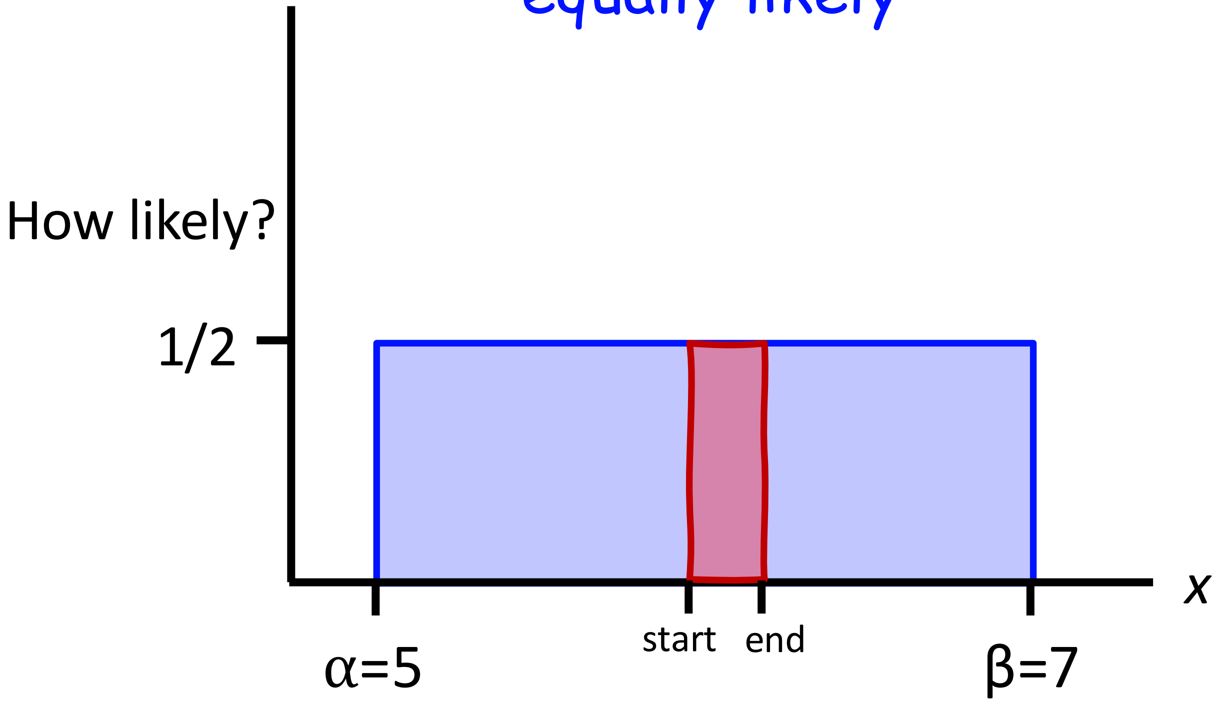
P(start ≤ X ≤ end) = end - start

Does that still work? No!

Need to divide by 2?

$X \sim$ Uniform(α, β): More General Case

All values are equally likely



Possible values are now between α and β

$$P(5 \leq X \leq 7) = 1$$

$$P(6 \leq X \leq 7) = 0.5$$

$$P(6 \leq X \leq 6.1) = 0.05$$

For Uniform(α, β):

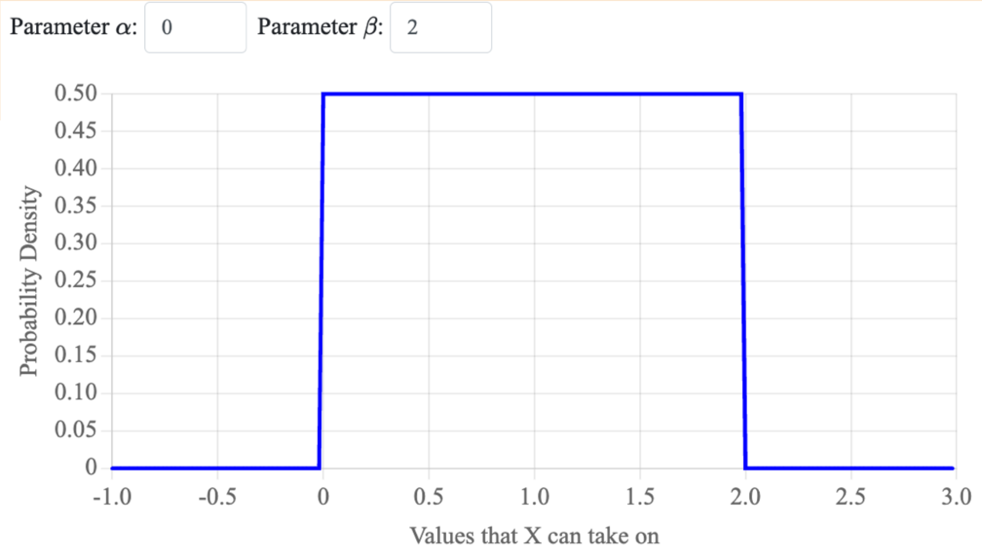
$$P(\text{start} \leq X \leq \text{end}) = \frac{\text{end - start}}{\beta - \alpha}$$

Uniform Random Variable

A **Uniform** random variable X takes on a value, with equal likelihood between α and β.

$$X \sim \text{Uni}(\alpha, \beta) \qquad P(x_1 < X < x_2) = \frac{x_2 - x_1}{\beta - \alpha}$$

Support: $[\alpha, \beta]$



Examples:
- Result of python random()
- Random points

Can we generalize to other continuous random variables?

Riding the Marguerite

Riding the Marguerite

You're running to the bus stop. You don't know exactly when the bus arrives.

You have a probability distribution for bus arrival times -- some times are more likely than others.

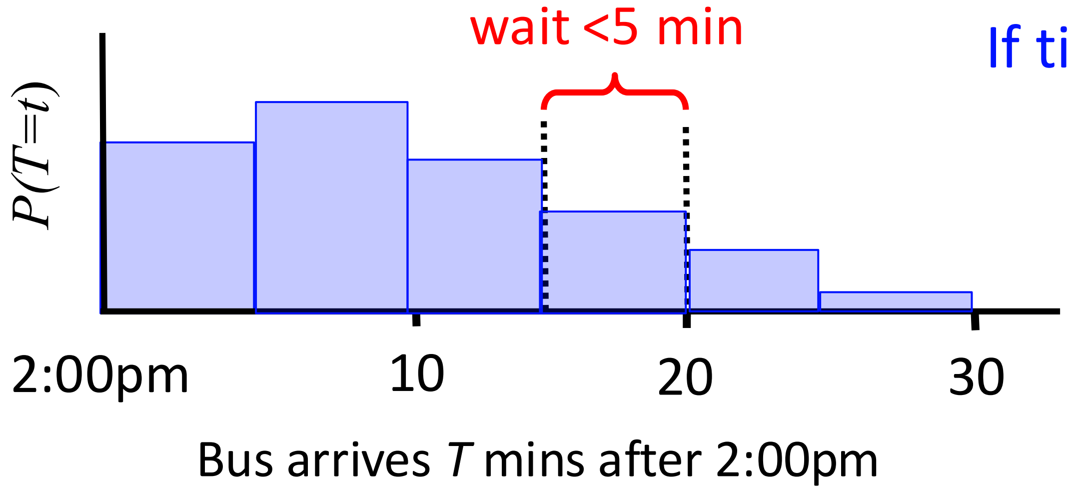
You show up at 2:15pm. What is P(wait < 5 min)?

Riding the Marguerite

You're running to the bus stop. You don't know exactly when the bus arrives.

You have a probability distribution for bus arrival times -- some times are more likely than others.

You show up at 2:15pm. What is P(wait < 5 min)?



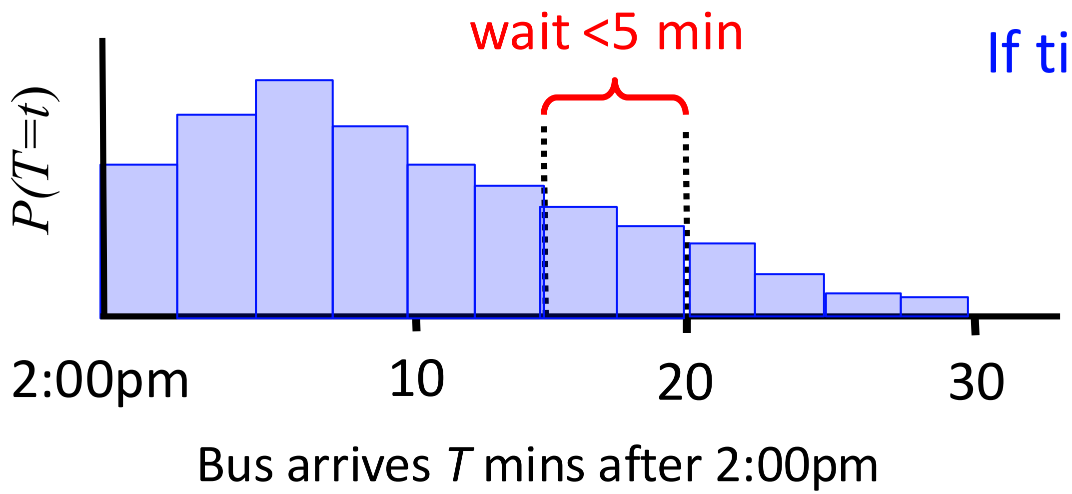
If time was discrete: a PMF could look like this.

Riding the Marguerite

You're running to the bus stop. You don't know exactly when the bus arrives.

You have a probability distribution for bus arrival times -- some times are more likely than others.

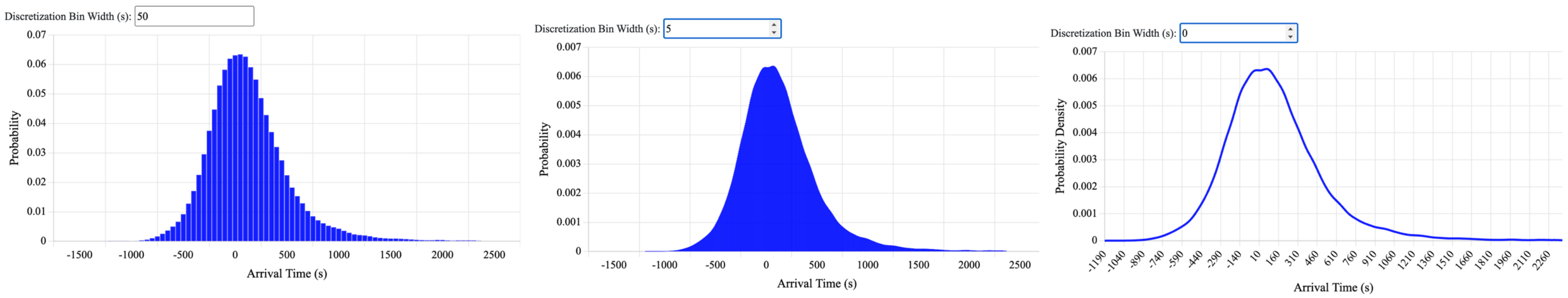
You show up at 2:15pm. What is P(wait < 5 min)?



If time was discrete: a PMF could look like this.

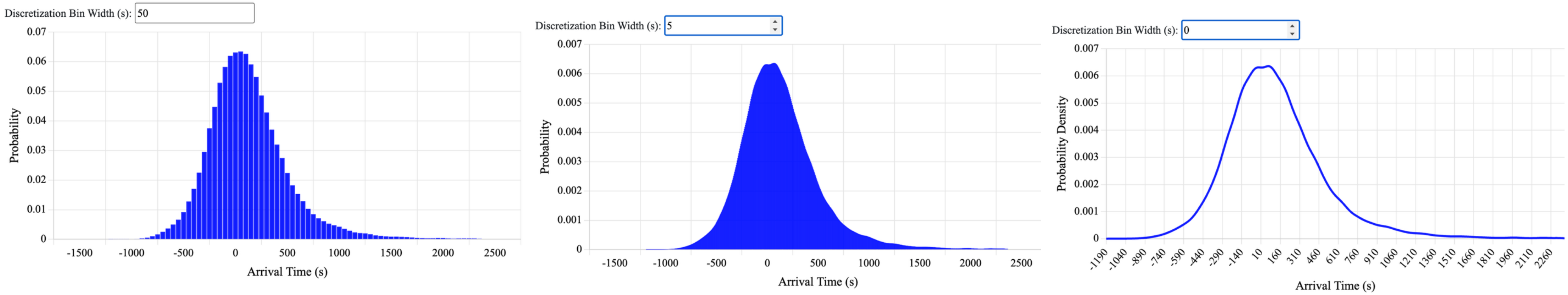
To the course reader!

What Happens as Bin Width Goes to 0 ?



As the bin width goes to 0.....

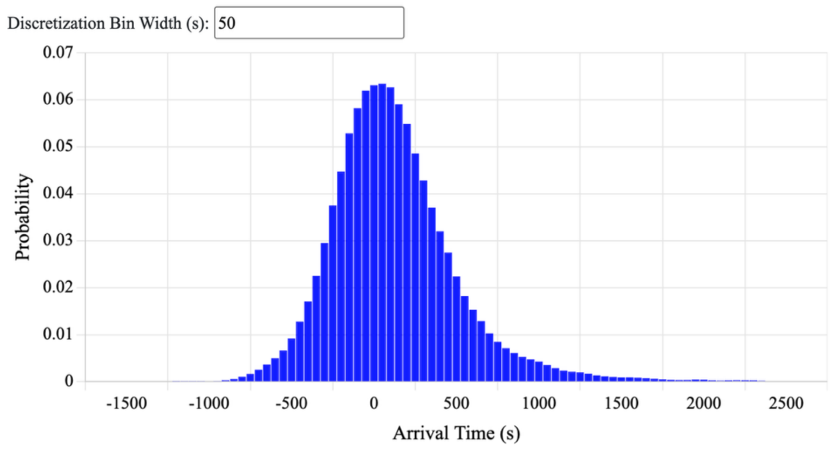
What Happens as Bin Width Goes to 0 ?

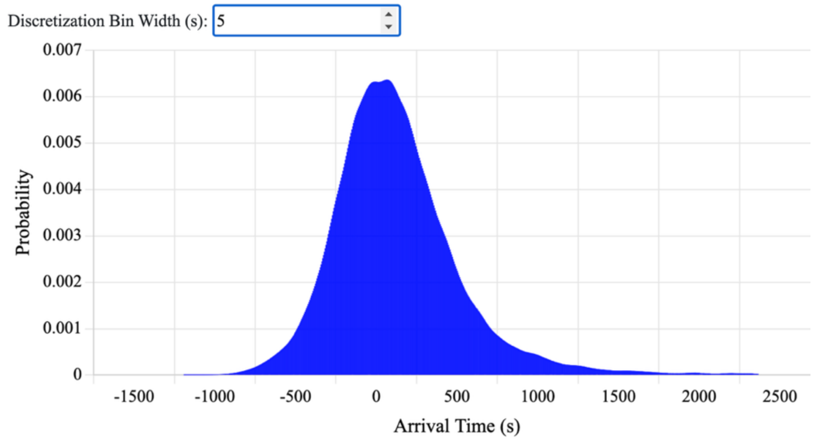


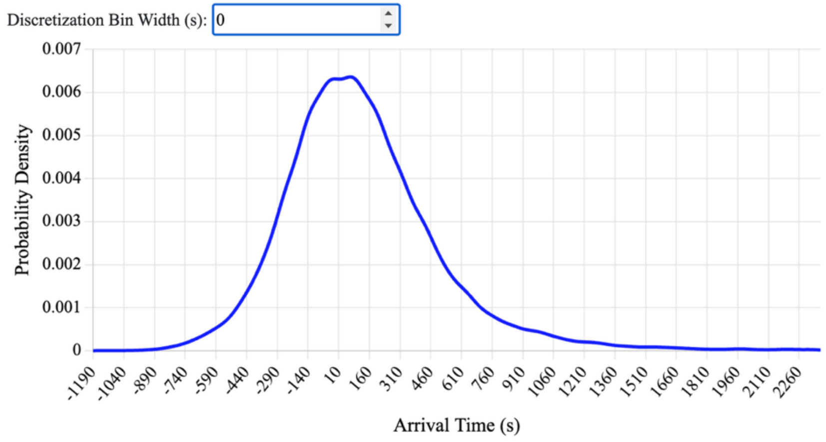
As the bin width goes to 0.....

This limit is where calculus is born!!

What Happens as Bin Width Goes to 0 ?







As the bin width goes to 0.....

This limit is where calculus is born!!

Summing probabilities over many tiny intervals becomes an integral.

Reimann sums

A PMF turns into a PDF, which is a derivative of a probability.

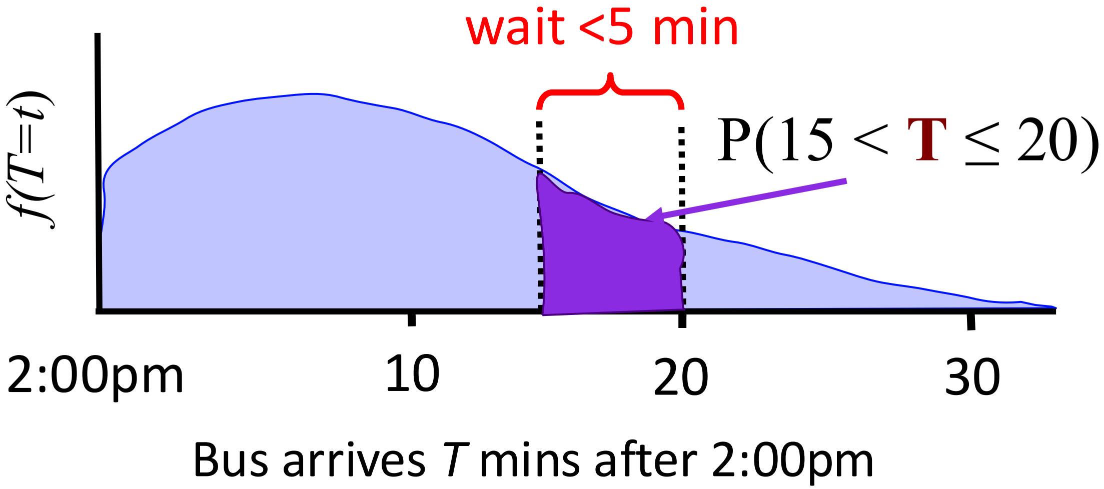
Finite differences become a derivative

Riding the Marguerite

You're running to the bus stop. You don't know exactly when the bus arrives.

You have a probability distribution for bus arrival times -- some times are more likely than others.

You show up at 2:15pm. What is P(wait < 5 min)?



When interval sizes tend towards 0:

- Time is now a **continuous** variable
- The **probability mass function (PMF)** becomes a **derivative** called a **probability density function (PDF)**
- Probability are now calculated as **area under the curve**

Time For Integrals!!!!

Probability Density Function

The **probability density function** (PDF) of a continuous random variable represents the relative likelihood of various values.

Units of probability *divided by units of X*.
Integrate it to get probabilities!

$$P(a < X < b) = \int_{x=a}^{b} f(X = x) dx$$

Probability Density Function

The **probability density function** (PDF) of a continuous random variable represents the relative likelihood of various values.

Units of probability *divided by units of X*.
Integrate it to get probabilities!

$$P(a < X < b) = \int_{x=a}^{b} \boxed{f_X(x)}\, dx$$

This is another way to write the PDF

Probability Density Function

The **probability density function** (PDF) of a continuous random variable represents the relative likelihood of various values.

Units of probability *divided by units of X*.
Integrate it to get probabilities!

$$P(a < X < b) = \int_{x=a}^{b} \boxed{f(x)}\, dx$$

This is another way to write the PDF

Probability Density Function

The **probability density function** (PDF) of a continuous random variable represents the relative likelihood of various values.

Units of probability *divided by units of X*.
Integrate it to get probabilities!

$$P(a < X < b) = \int_{x=a}^{b} f(X = x)dx$$

PDFs like $f(X = x)$ vs. PMFs like $P(X = x)$

$$P(X = x)$$

"The probability that a **discrete** random variable X takes on the value *x*."

$$f(X = x)$$

"The ***derivative*** of the probability that a **continuous** random variable X takes at the value *x*."

*They are both measures of how **likely** X is to take on the value x.*
*Sometimes called the **distribution** function.*

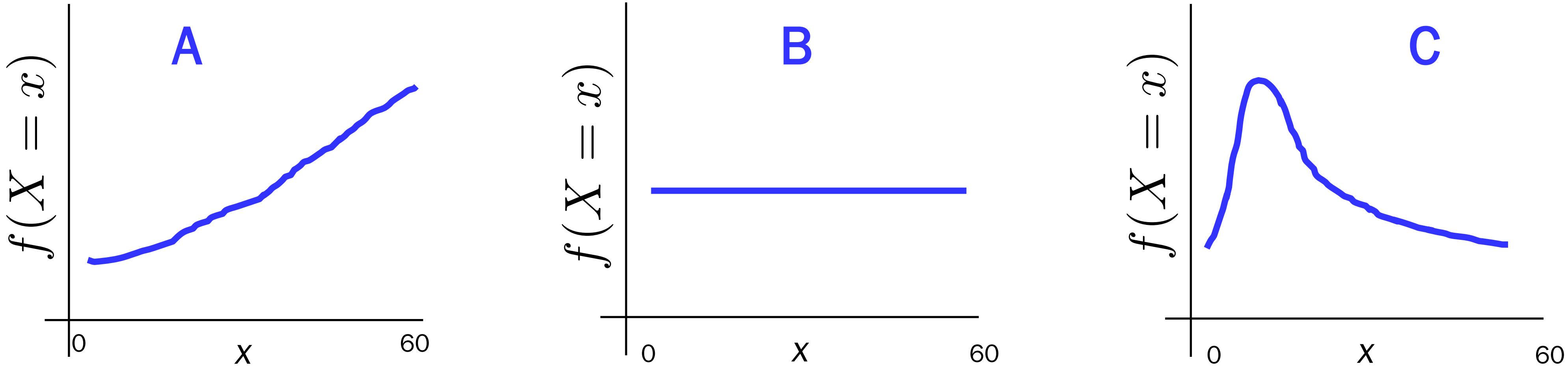
What do you get if you integrate over a probability *density* function?

A probability!

The Relative Values of PDFs Are Meaningful

Probability density functions are derivatives that articulate ***relative*** belief.

Let X be the # of minutes after 2pm that the bus arrives at a stop.

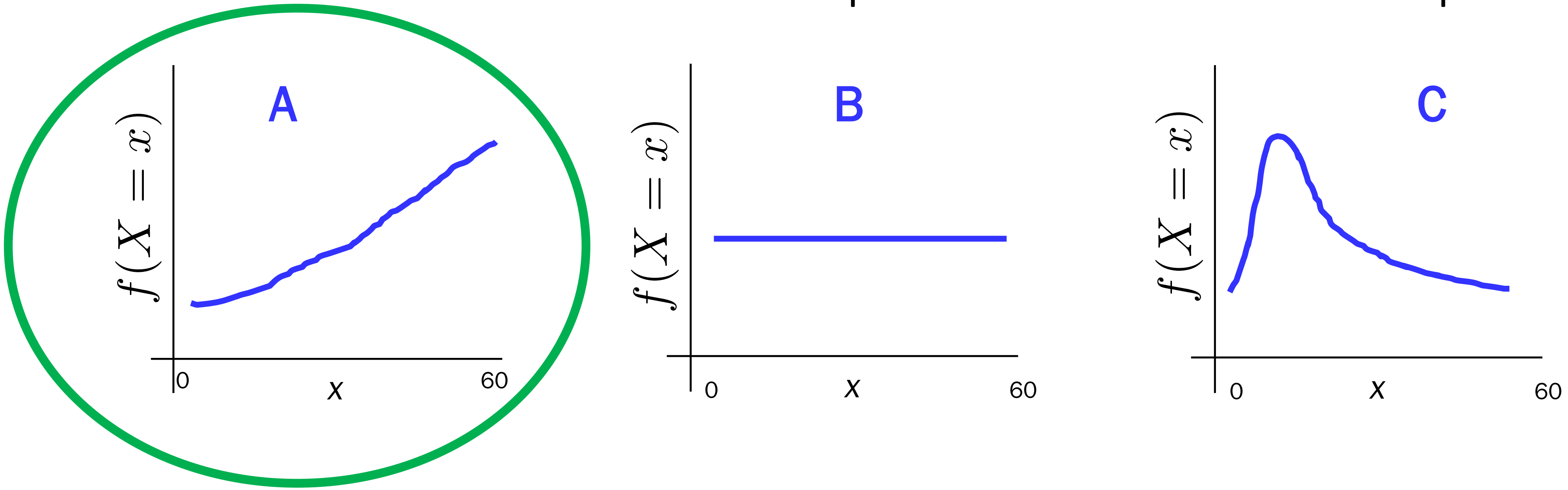


Which of these represent that the bus's arrival
is more likely to be close to 3:00pm?

The Relative Values of PDFs Are Meaningful

Probability density functions are derivatives that articulate ***relative*** belief.

Let X be the # of minutes after 2pm that the bus arrives at a stop.



Which of these represent that the bus's arrival
is more likely to be close to 3:00pm?

The relative value of probability densities is meaningful

Truths of Probability For Continuous Random Variables

Truth 1: $$P(a < X < b) = \int_{x=a}^{b} f(X = x)\ dx$$ Area under the curve!

Truths of Probability For Continuous Random Variables

Truth 1:

$$P(a < X < b) = \int_{x=a}^{b} f(X = x)\ dx$$

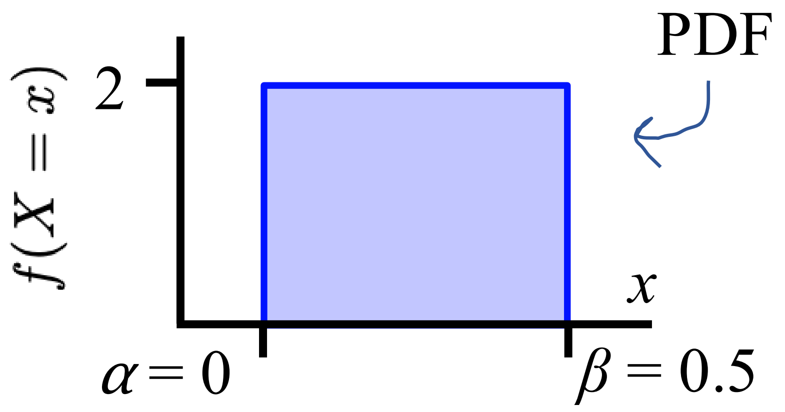
Area under the curve!

Truth 2:

$$0 \leq \int_{x=a}^{b} f(X = x)\ dx \leq 1$$

Since the integral is a probability (Axiom 1)

Can a PDF ever have a value > 1?
Yes!



Truths of Probability For Continuous Random Variables

Truth 1:

$$P(a < X < b) = \int_{x=a}^{b} f(X = x)\ dx$$

Area under the curve!

Truth 2:

$$0 \leq \int_{x=a}^{b} f(X = x)\ dx \leq 1$$

Since the integral is a probability (Axiom 1)

Truth 3:

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

That's all possible values (Axiom 2)

Truths of Probability For Continuous Random Variables

Truth 1: $$P(a < X < b) = \int_{x=a}^{b} f(X = x)\ dx$$ Area under the curve!

Truth 2: $$0 \leq \int_{x=a}^{b} f(X = x)\ dx \leq 1$$ Since the integral is a probability (Axiom 1)

Truth 3: $$\int_{x=-\infty}^{\infty} f(X = x)\ dx = 1$$ That's all possible values (Axiom 2)

Truth 4: $$P(X = x) = 0$$ What a time to be alive…

Solve for K $f(X = x) = K \cdot x^2$ $0 \leq x \leq 1$

Truth 1:

$$P(a < X < b) = \int_{x=a}^{b} f(X = x)\ dx$$

Area under the curve!

Truth 2:

$$0 \leq \int_{x=a}^{b} f(X = x)\ dx \leq 1$$

Since the integral is a probability (Axiom 1)

Truth 3:

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

That's all possible values (Axiom 2)

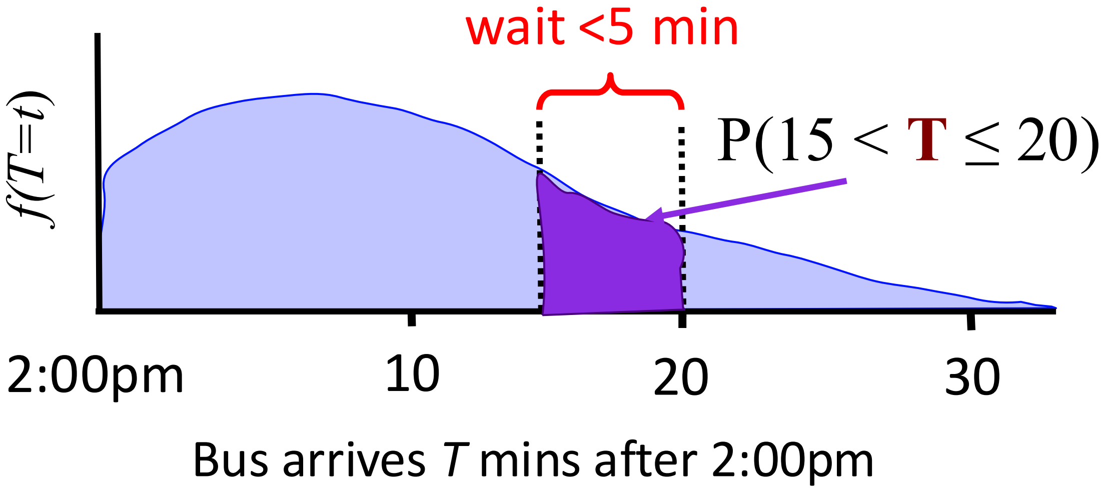
Truth 4:

$$P(X = x) = 0$$

What a time to be alive…

PDFs Need an Integral

What is the probability that the bus arrives at: 12.12332343234… mins after 2pm?



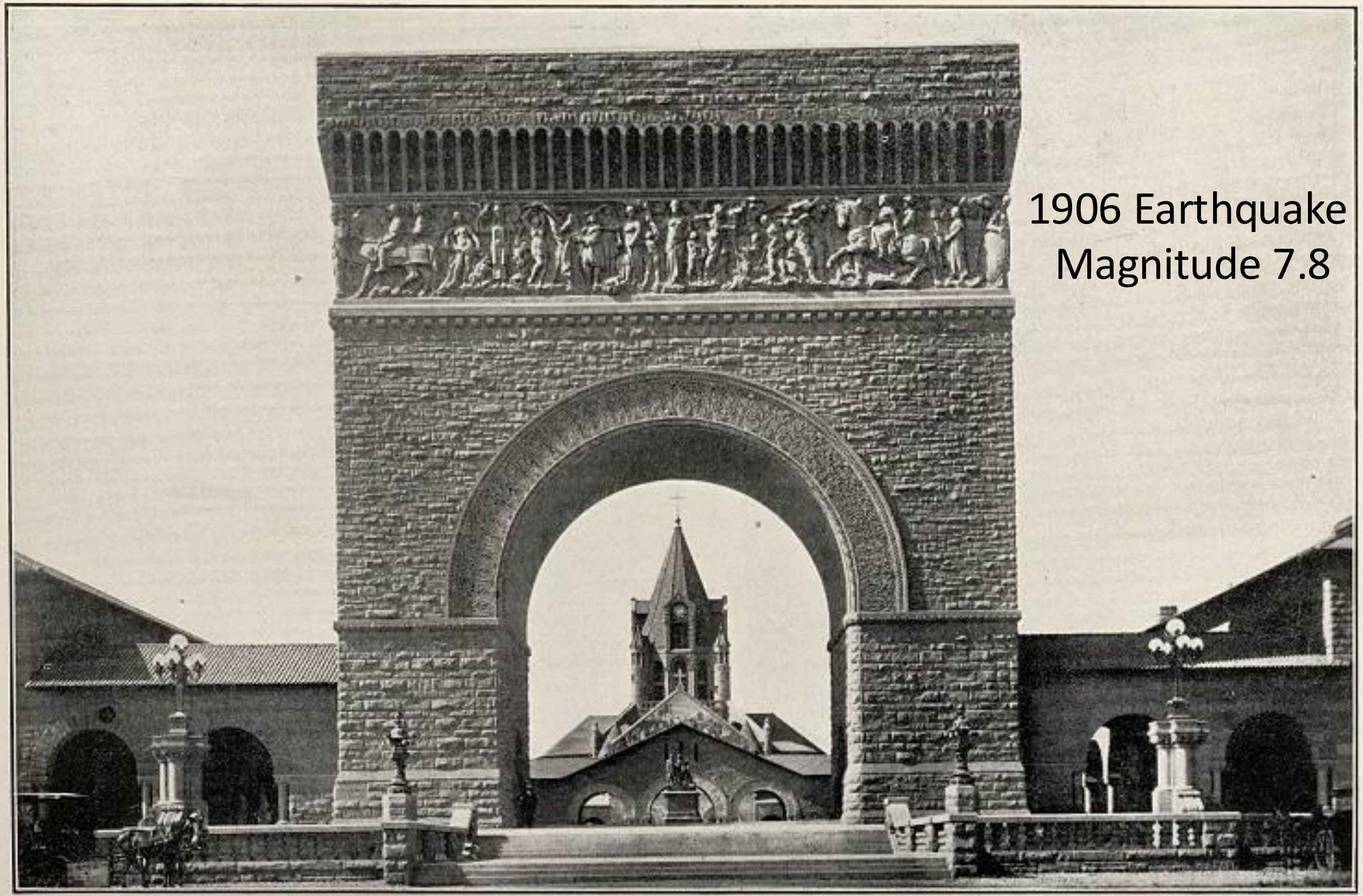
What do you get if you integrate over a probability *density* function?

A probability!

Pedagogic Pause

You are ready for the classic continuous random variables

It’s Time
To Talk About Time, Again



ILL. No. 65. MEMORIAL ARCH, WITH CHURCH IN BACKGROUND, STANFORD UNIVERSITY, SHOWING TYPES OF CARVED WORK WITH THE SANDSTONE.

How long until the next “big one”?

Exponential Random Variable

For any Poisson Process, the **Exponential** RV models *time until an event*:

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

PDF:

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

Exponential Random Variable

For any Poisson Process, the **Exponential** RV models *time until an event*:

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

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

Examples:

- Time until next earthquake
- Time until a ping reaches a web server
- Time until next Uber request

Exponential Random Variable

For any Poisson Process, the **Exponential** RV models *time until an event*:

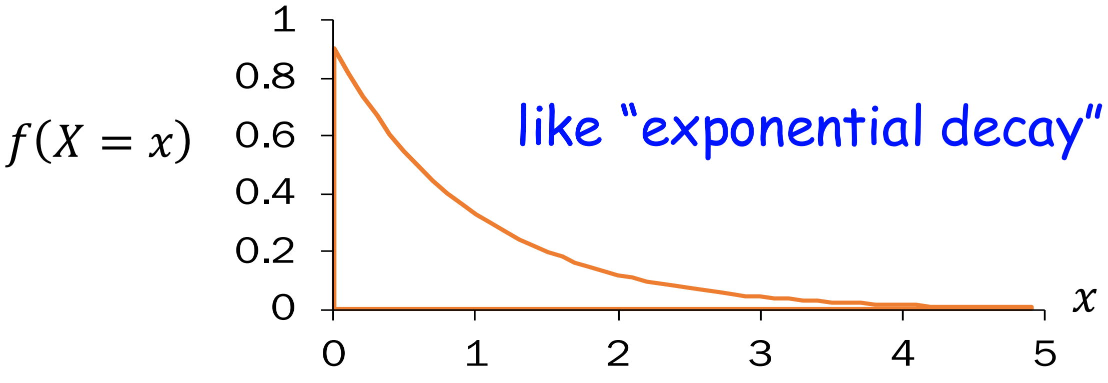
$$X \sim \mathrm{Exp}(\lambda)$$

PDF:

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

Examples:

- Time until next earthquake
- Time until a ping reaches a web server
- Time until a Uranium atom decays



The process for an Exponential and a Poisson are the **same**. So is the parameter λ. The question is different

ILL. No. 65. MEMORIAL ARCH, WITH CHURCH IN BACKGROUND, STANFORD UNIVERSITY, SHOWING TYPES OF CARVED WORK WITH THE SANDSTONE

How Long Until the Next Earthquake

Based on historical data, major earthquakes (magnitude 8.0+) happen at a **rate of 0.002** per year*. What is the probability of **a major earthquake in the next 30 years?**

Y = Years until the next earthquake of magnitude 8.0+

$$Y \sim \text{Exp}(\lambda = 0.002)$$

$$f_Y(y) = \lambda e^{-\lambda y} = 0.002 e^{-0.002y}$$

$$P(Y < 30) = \int_0^{30} 0.002 e^{-0.002y} \, dy$$

*In California, according to the USGS, 2015 ty

Integral Review

$$\int e^{cx}\ dx = \frac{1}{c}e^{cx}$$

How Long Until the Next Earthquake

Based on historical data, major earthquakes (magnitude 8.0+) happen at a **rate of 0.002** per year*. What is the probability of **a major earthquake in the next 30 years?**

Y = Years until the next earthquake of magnitude 8.0+

$$Y \sim \text{Exp}(\lambda = 0.002)$$

$$f_Y(y) = \lambda e^{-\lambda y}$$
$$= 0.002^{-0.002y}$$

$$P(Y < 30) = \int_0^{30} 0.002 e^{-0.002y}\, dy$$
$$= 0.002 \left[-\frac{1}{0.002} e^{-0.002y} \right]_0^{30}$$
$$= \left[-e^{-0.002y} \right]_0^{30}$$
$$= -e^{-0.002 \cdot 30} + e^0$$
$$\approx 0.058$$

*In California, according to the USGS, 2015

ty

How Long Until the Next Earthquake

Based on historical data, major earthquakes (magnitude 8.0+) happen at a **rate of 0.002** per year*. What is the **expected number of years until the next earthquake?**

Y = Years until the next earthquake of magnitude 8.0+

$$Y \sim \text{Exp}(\lambda = 0.002)$$

$$E[Y] = \frac{1}{\lambda} = \frac{1}{0.002} = 500$$

*In California, according to the USGS, 2015

ty

How Long Until the Next Earthquake

Based on historical data, major earthquakes (magnitude 8.0+) happen at a **rate of 0.002** per year*. What is the **standard deviation of years until the next earthquake?**

Y = Years until the next earthquake of magnitude 8.0+

$$Y \sim \text{Exp}(\lambda = 0.002)$$

$$\text{Var}(Y) = \frac{1}{\lambda^2} = \frac{1}{0.002^2} = 250,000 \text{ years}^2$$

$$\text{Std}(Y) = \sqrt{\text{Var}(X)} = 500 \text{ years}$$

*In California, according to the USGS, 2015

ty

How Many Earthquakes

Based on historical data, major earthquakes (magnitude 8.0+) happen at a **rate of 0.002** per year*. What is the probability of **zero major earthquakes magnitude next year?**

$X =$ Number of major earthquakes next year

$$X \sim \text{Poi}(\lambda = 0.002)$$

$$P(X = 0) = \frac{\lambda^0 e^{-\lambda}}{0!} = \frac{0.002^0 e^{-0.002}}{0!} \approx 0.998$$

*In California, according to the USGS, 2015 ty

Is there a way to avoid integrals?

Cumulative Distribution Function

A cumulative distribution function (CDF) is a "closed form" equation for the probability that a random variable is less than a given value

$$F(x) = P(X < x)$$

If you learn how to use a cumulative distribution function, you can avoid integrals!

$$F_X(x)$$

This is also shorthand notation for the CDF

Cumulative Distribution Function

$$F(x) = P(X < x)$$

$$x = 2$$

$$0.03125$$

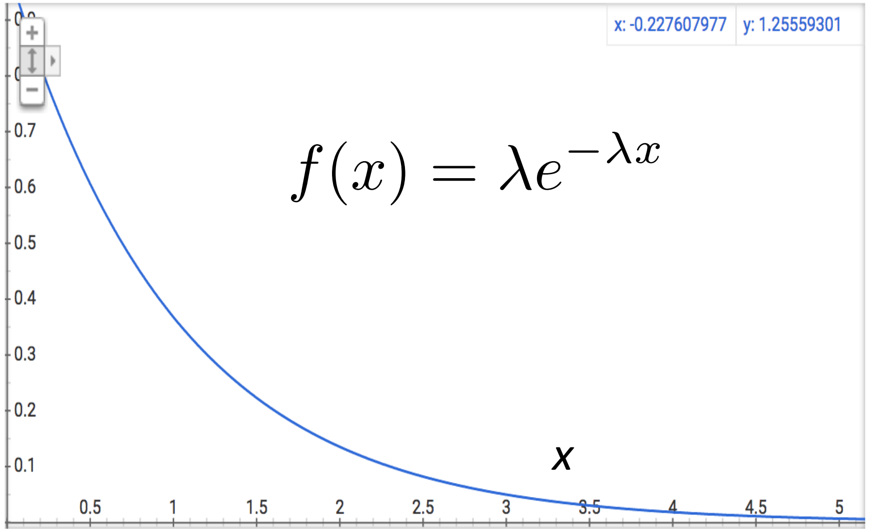
CDF of an Exponential

$$F_X(x) = 1 - e^{-\lambda x}$$

$$\begin{aligned}
P(X < x) &= \int_{y=-\infty}^{x} f(y)\, dy \\
&= \int_{y=0}^{x} \lambda e^{-\lambda y}\, dy \\
&= \frac{\lambda}{\lambda}\Big[-e^{-\lambda y}\Big]_0^x \\
&= [-e^{-\lambda x}] - [-e^{\lambda 0}] \\
&= 1 - e^{-\lambda x}
\end{aligned}$$

Using CDF Example. X is Exp(λ = 1)

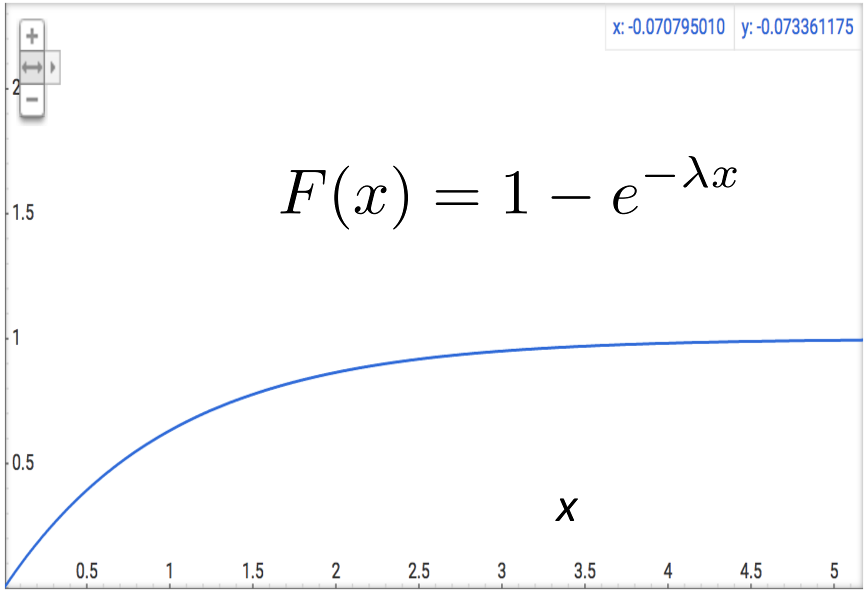
Probability density function



Cumulative distribution function

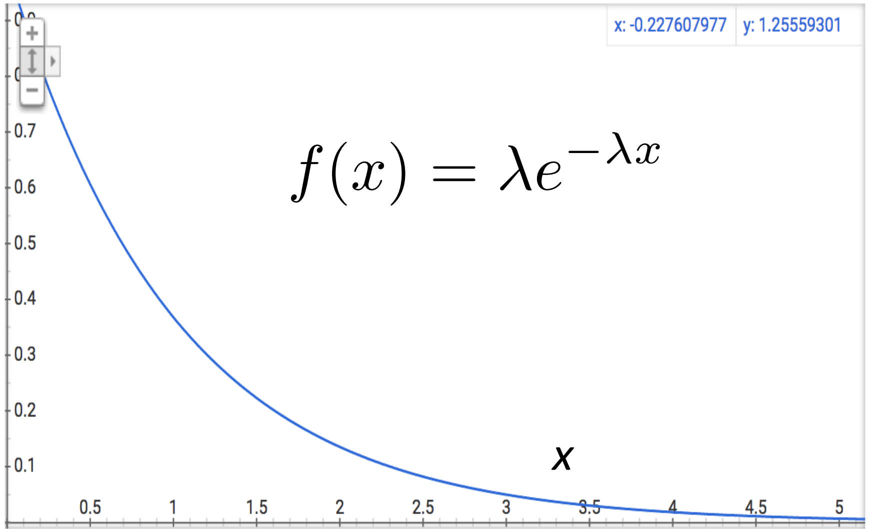
$$F_X(x) = P(X < x)$$

$$= \int_{y=-\infty}^{x} f(y)\, dy$$



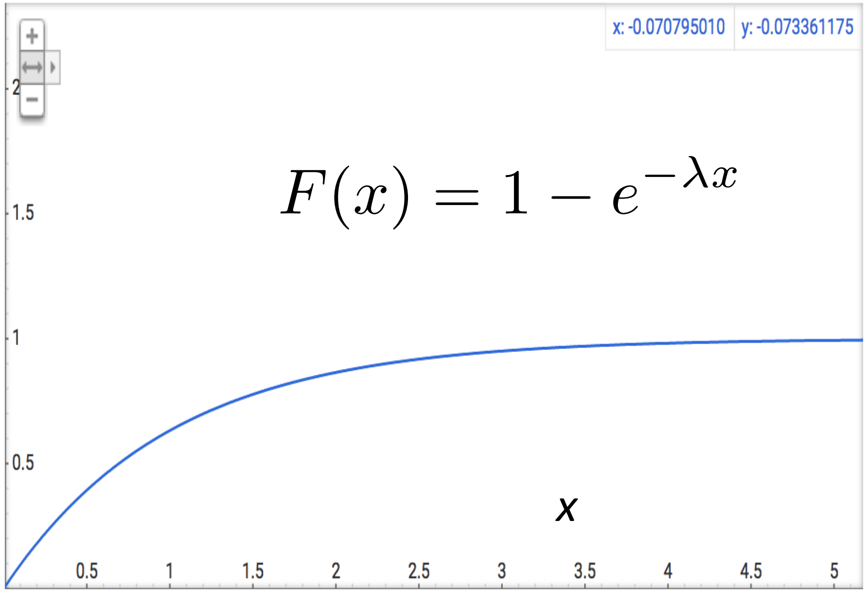
Using CDF Example. X is Exp(λ = 1)

Probability
density
function



P(X < 2)

Cumulative
distribution
function

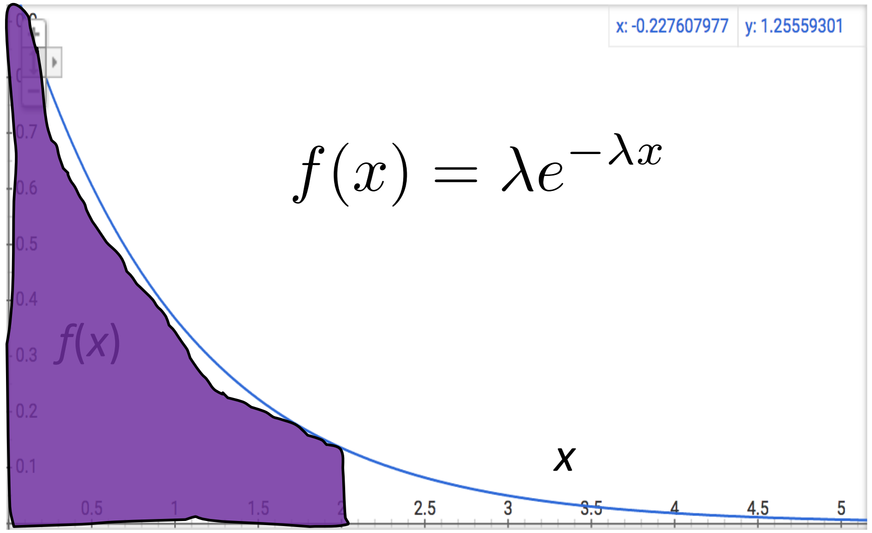


$$F_X(x) = P(X < x)$$

$$= \int_{y=-\infty}^{x} f(y)\, dy$$

Using CDF Example. X is Exp(λ = 1)

Probability density function



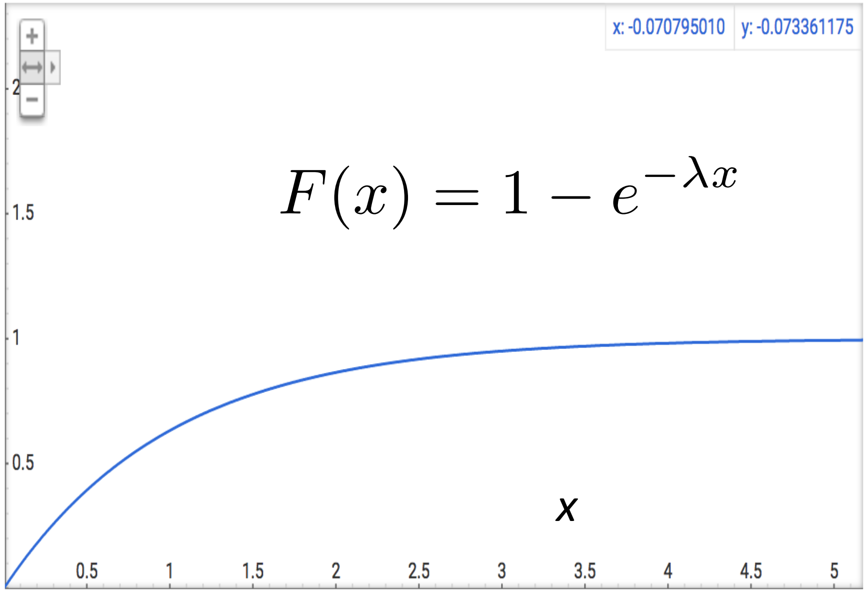
P(X < 2)

$$= \int_{x=-\infty}^{2} f(x)\, dx$$

Cumulative distribution function

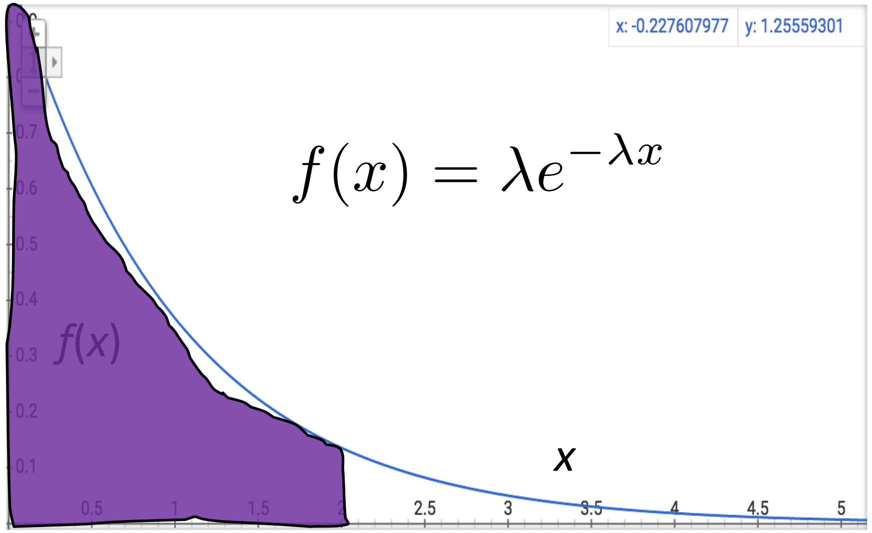
$$F_X(x) = P(X < x)$$

$$= \int_{y=-\infty}^{x} f(y)\, dy$$



Using CDF Example. X is Exp(λ = 1)

Probability density function



P(X < 2)

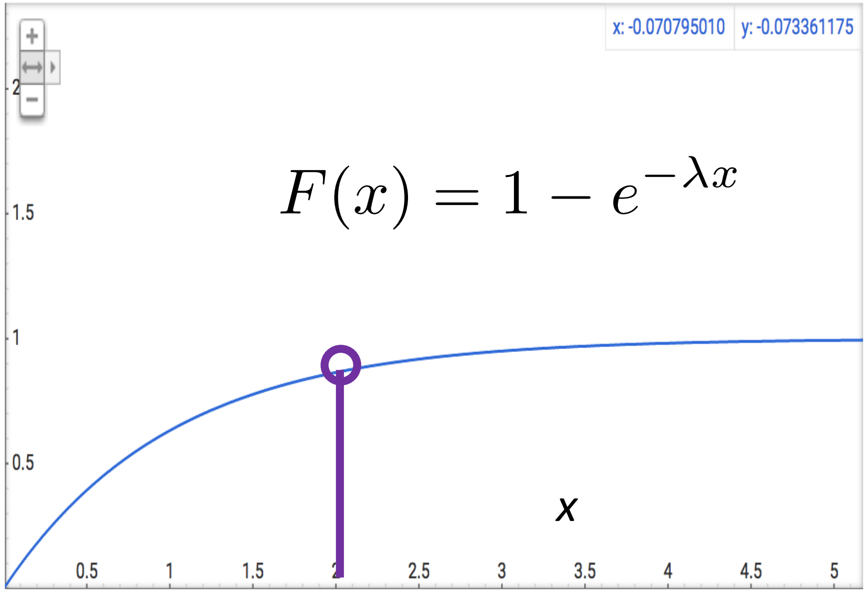
$$= \int_{x=-\infty}^{2} f(x)\, dx$$

or

Cumulative distribution function

$$F_X(x) = P(X < x)$$

$$= \int_{y=-\infty}^{x} f(y)\, dy$$



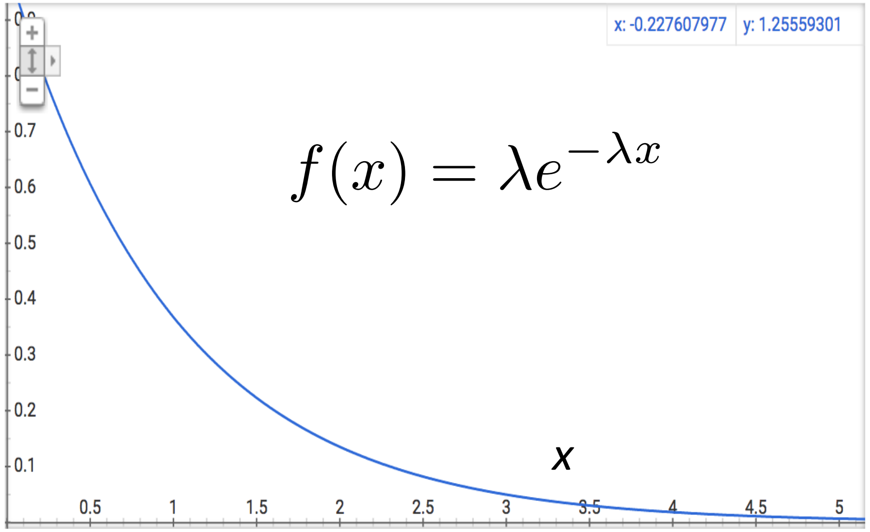
$$= F(2)$$

$$= 1 - e^{-2}$$

$$\approx 0.84$$

Using CDF Example. X is Exp(λ = 1)

Probability density function

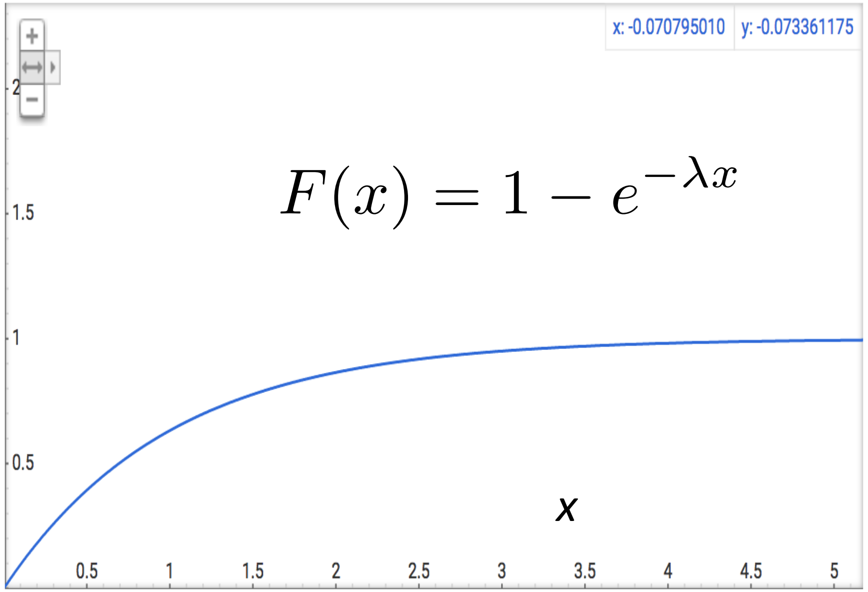


P(X > 1)

Cumulative distribution function

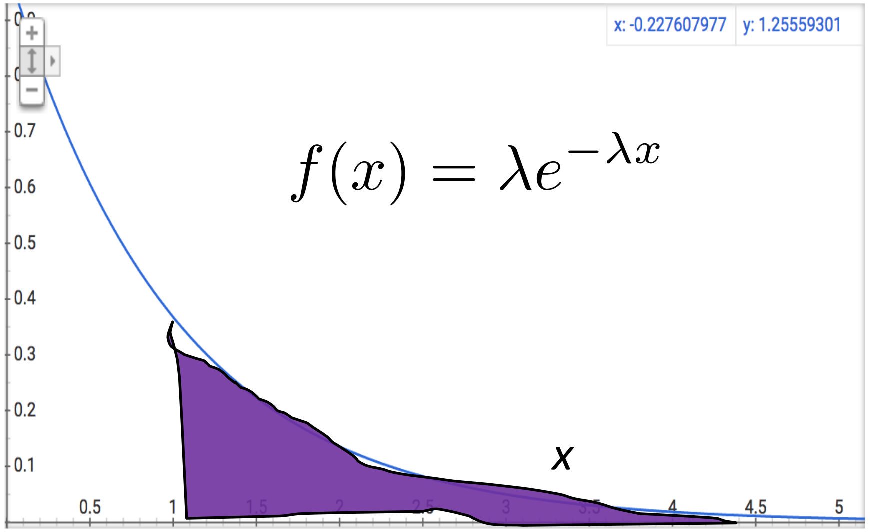
$$F_X(x) = P(X < x)$$

$$= \int_{y=-\infty}^{x} f(y)\, dy$$



Using CDF Example. X is Exp(λ = 1)

Probability density function



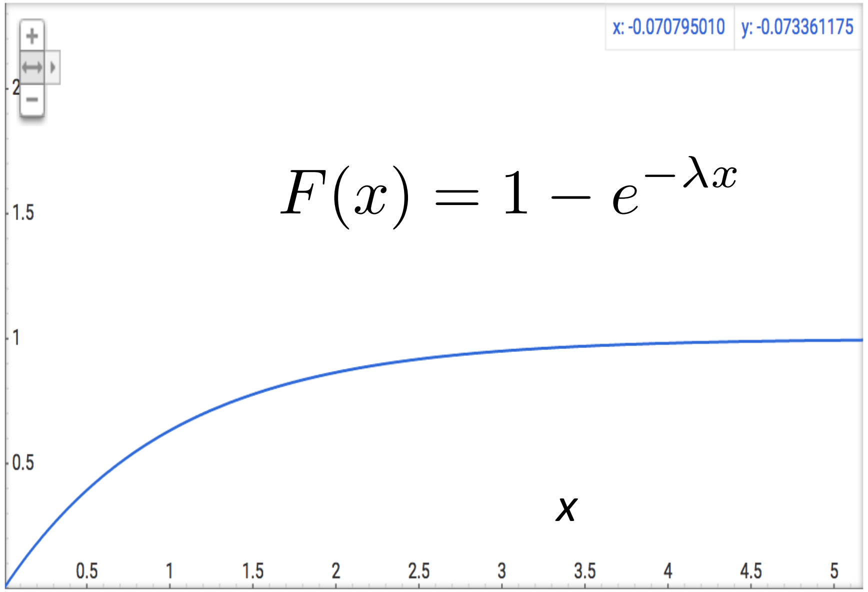
P(X > 1)

$$= \int_{x=1}^{\infty} f(x)\, dx$$

Cumulative distribution function

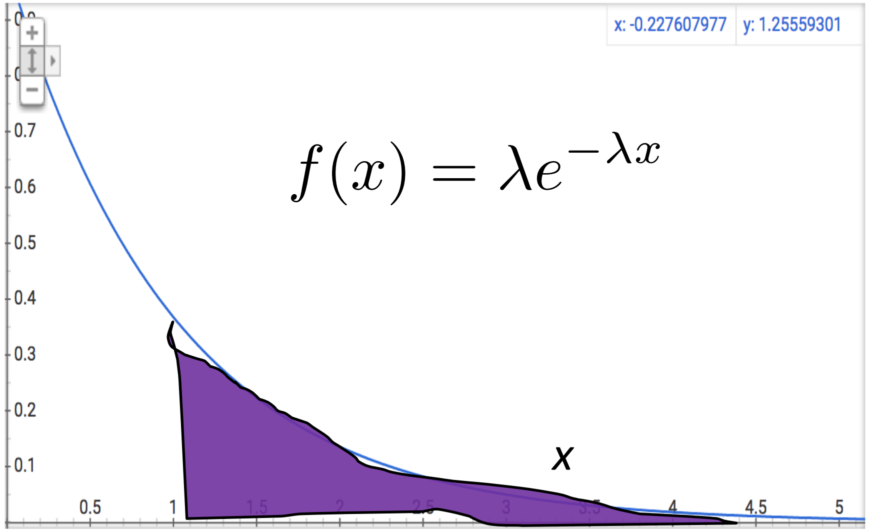
$$F_X(x) = P(X < x)$$

$$= \int_{y=-\infty}^{x} f(y)\, dy$$



Using CDF Example. X is Exp($\lambda = 1$)

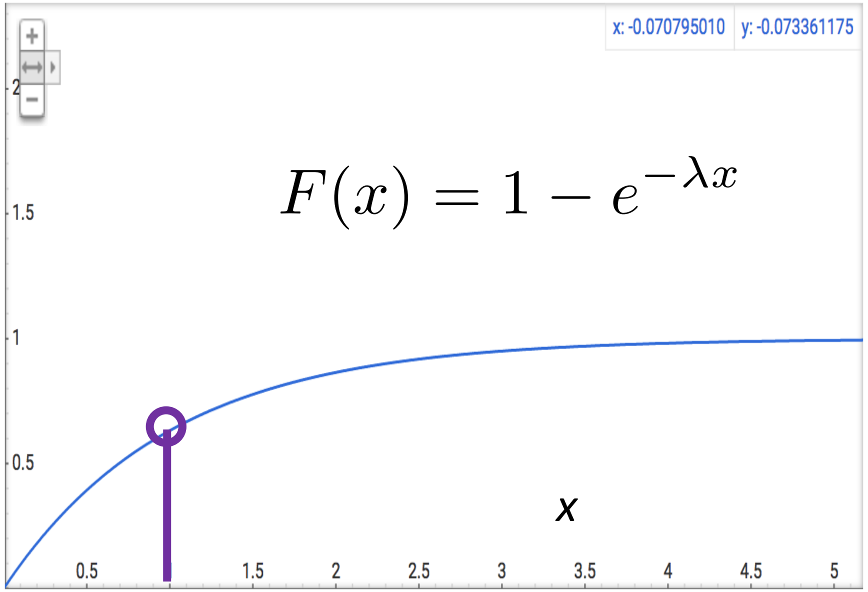
Probability density function



Cumulative distribution function

$$F_X(x) = P(X < x)$$

$$= \int_{y=-\infty}^{x} f(y)\, dy$$



P(X > 1)

$$= \int_{x=1}^{\infty} f(x)\, dx$$

or

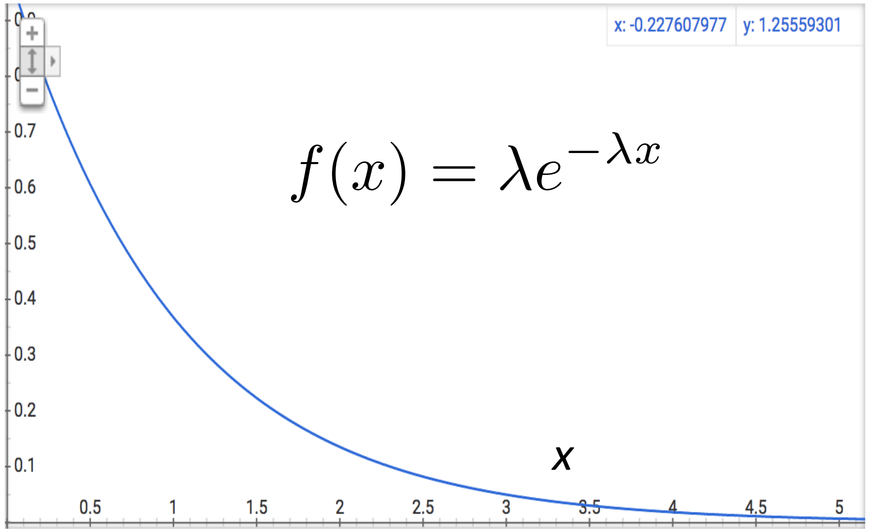
$$= 1 - F(1)$$

$$= e^{-1}$$

$$\approx 0.37$$

Using CDF Example. X is Exp(λ = 1)

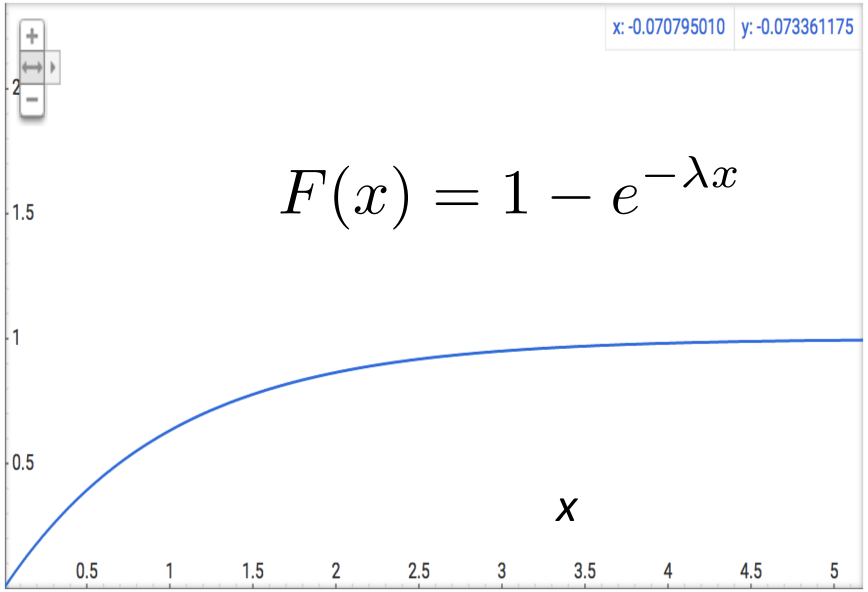
Probability density function



Cumulative distribution function

$$F_X(x) = P(X < x)$$

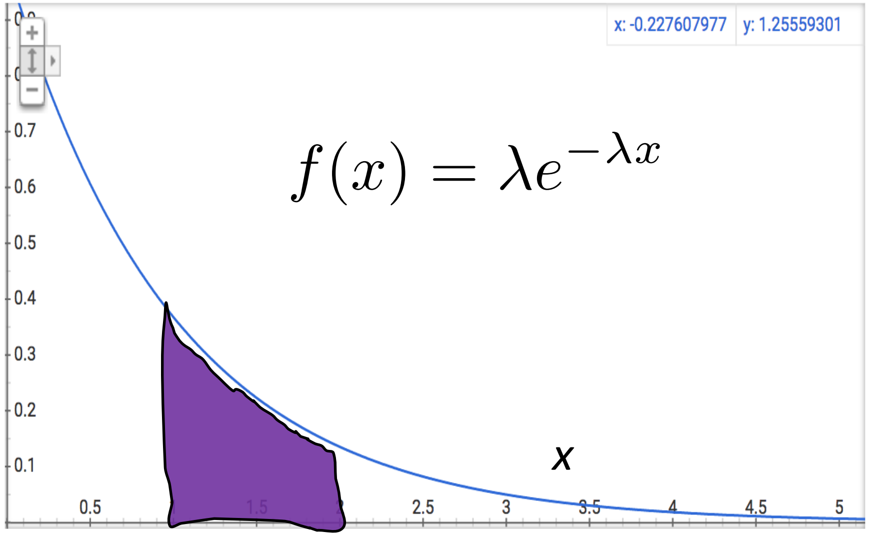
$$= \int_{y=-\infty}^{x} f(y)\, dy$$



P(1 < X < 2)

Using CDF Example. X is Exp($\lambda = 1$)

Probability density function



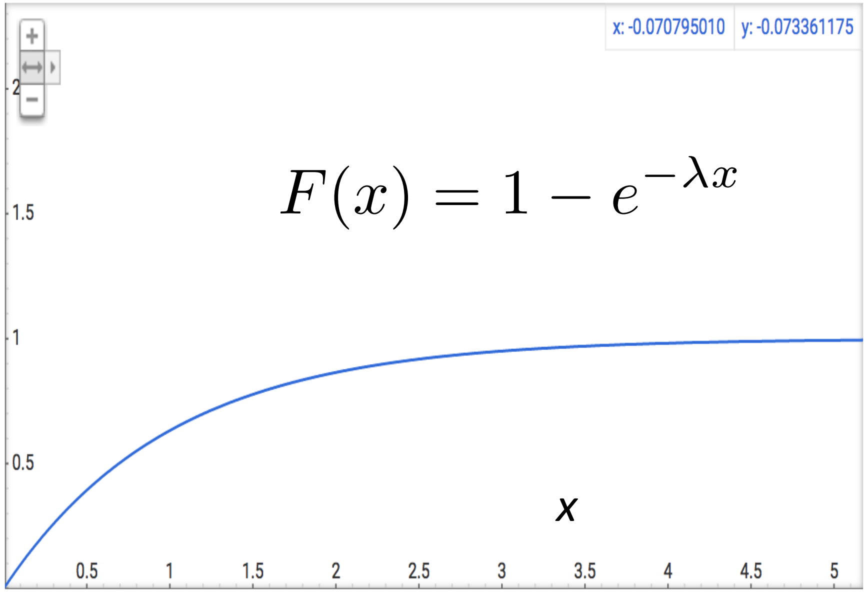
P(1 < X < 2)

$$= \int_{x=1}^{2} f(x)\, dx$$

Cumulative distribution function

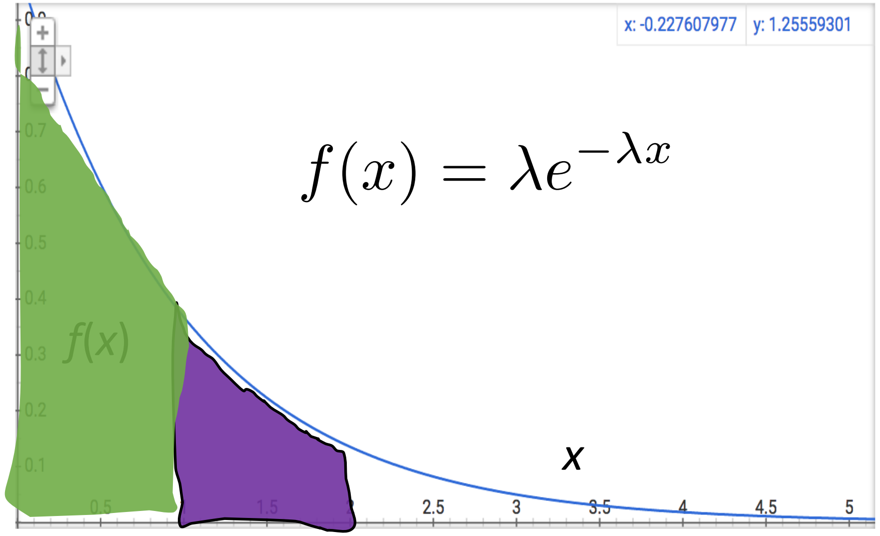
$$F_X(x) = P(X < x)$$

$$= \int_{y=-\infty}^{x} f(y)\, dy$$



Using CDF Example. X is Exp($\lambda = 1$)

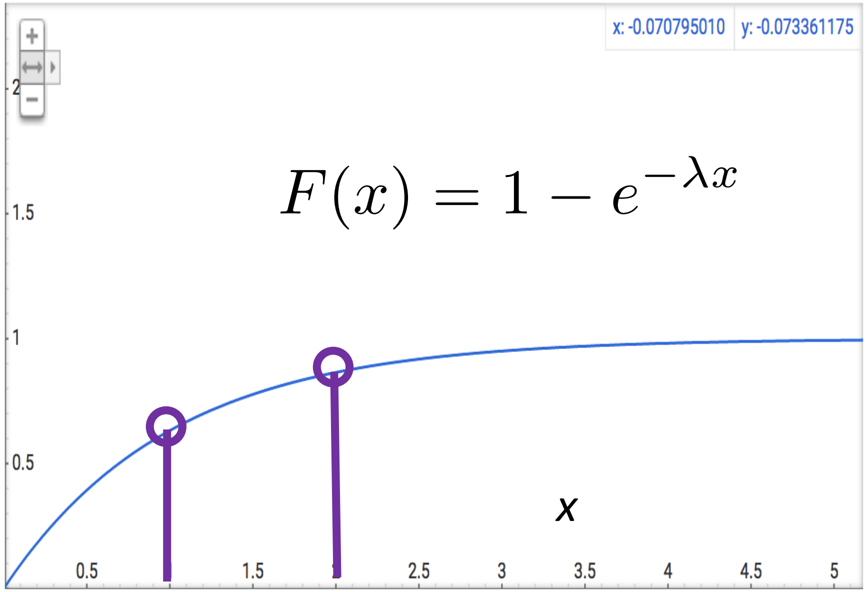
Probability density function



Cumulative distribution function

$$F_X(x) = P(X < x)$$

$$= \int_{y=-\infty}^{x} f(y)\, dy$$



P(1 < X < 2)

$$= \int_{x=1}^{2} f(x)\, dx$$

or

$$= F(2) - F(1)$$

$$= (1 - e^{-2}) - (1 - e^{-1})$$

$$\approx 0.23$$

Probability of Earthquake in Next 4 Years?

Based on historical data, earthquakes of magnitude 8.0+ happen at a **rate of 0.002** per year*. What is the probability of **a major earthquake in the next 4 years?**

Y = Years until the next earthquake of magnitude 8.0+

$$Y \sim \text{Exp}(\lambda = 0.002) \qquad F(y) = 1 - e^{-0.002y}$$

$$\begin{aligned} P(Y < 4) &= F(4) \\ &= 1 - e^{-0.002 \cdot 4} \\ &\approx 0.008 \end{aligned}$$

Feeling lucky?

*According to USGS, 2015

Two Classic Random Variables

Uniform Random Variable

Notation: $X \sim \text{Uni}(\alpha, \beta)$

Description: A continuous random variable that takes on values, with equal likelihood, between α and β

Parameters: $\alpha \in \mathbb{R}$, the minimum value of the variable.
$\beta \in \mathbb{R}$, $\beta > \alpha$, the maximum value of the variable.

Support: $x \in [\alpha, \beta]$

PDF equation: $f(x) = \begin{cases} \frac{1}{\beta-\alpha} & \text{for } x \in [\alpha, \beta] \\ 0 & \text{else} \end{cases}$

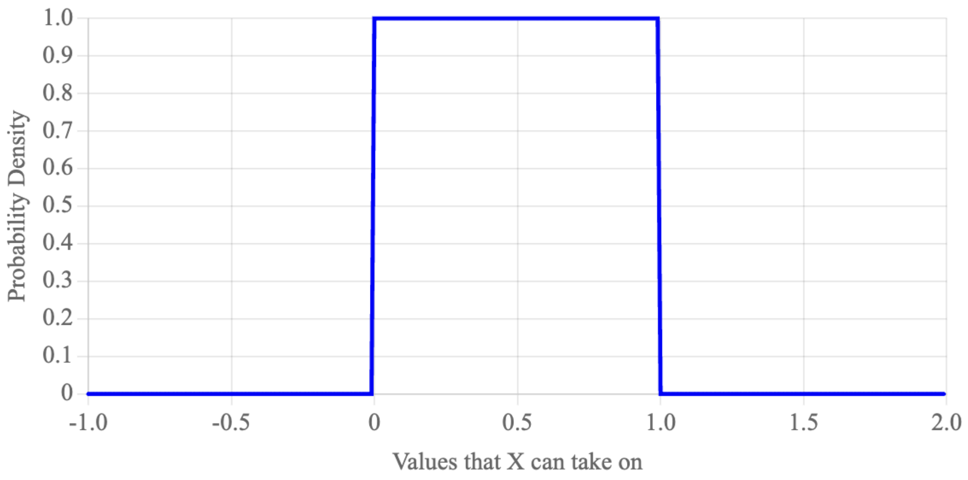
CDF equation: $F(x) = \begin{cases} \frac{x-\alpha}{\beta-\alpha} & \text{for } x \in [\alpha, \beta] \\ 0 & \text{for } x < \alpha \\ 1 & \text{for } x > \beta \end{cases}$

Expectation: $\text{E}[X] = \frac{1}{2}(\alpha + \beta)$

Variance: $\text{Var}(X) = \frac{1}{12}(\beta - \alpha)^2$

PDF graph:

Parameter α: 0 Parameter β: 1



Exponential Random Variable

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

Description: Time until next events if (a) the events occur with a constant mean rate and (b) they occur independently of time since last event.

Parameters: $\lambda \in \{0, 1, \ldots\}$, the constant average rate.

Support: $x \in \mathbb{R}^+$

PDF equation: $f(x) = \lambda e^{-\lambda x}$

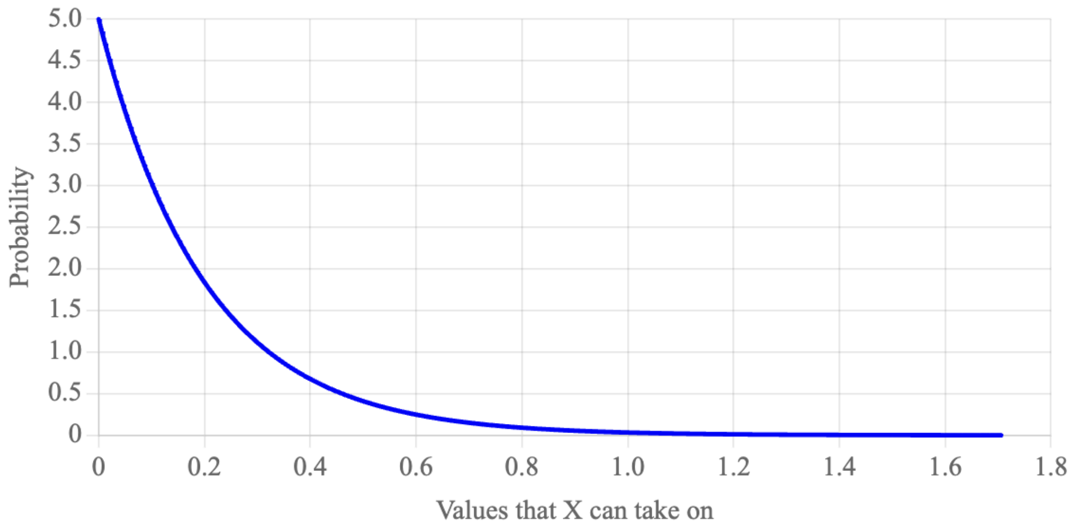
CDF equation: $F(x) = 1 - e^{-\lambda x}$

Expectation: $\text{E}[X] = 1/\lambda$

Variance: $\text{Var}(X) = 1/\lambda^2$

PDF graph:

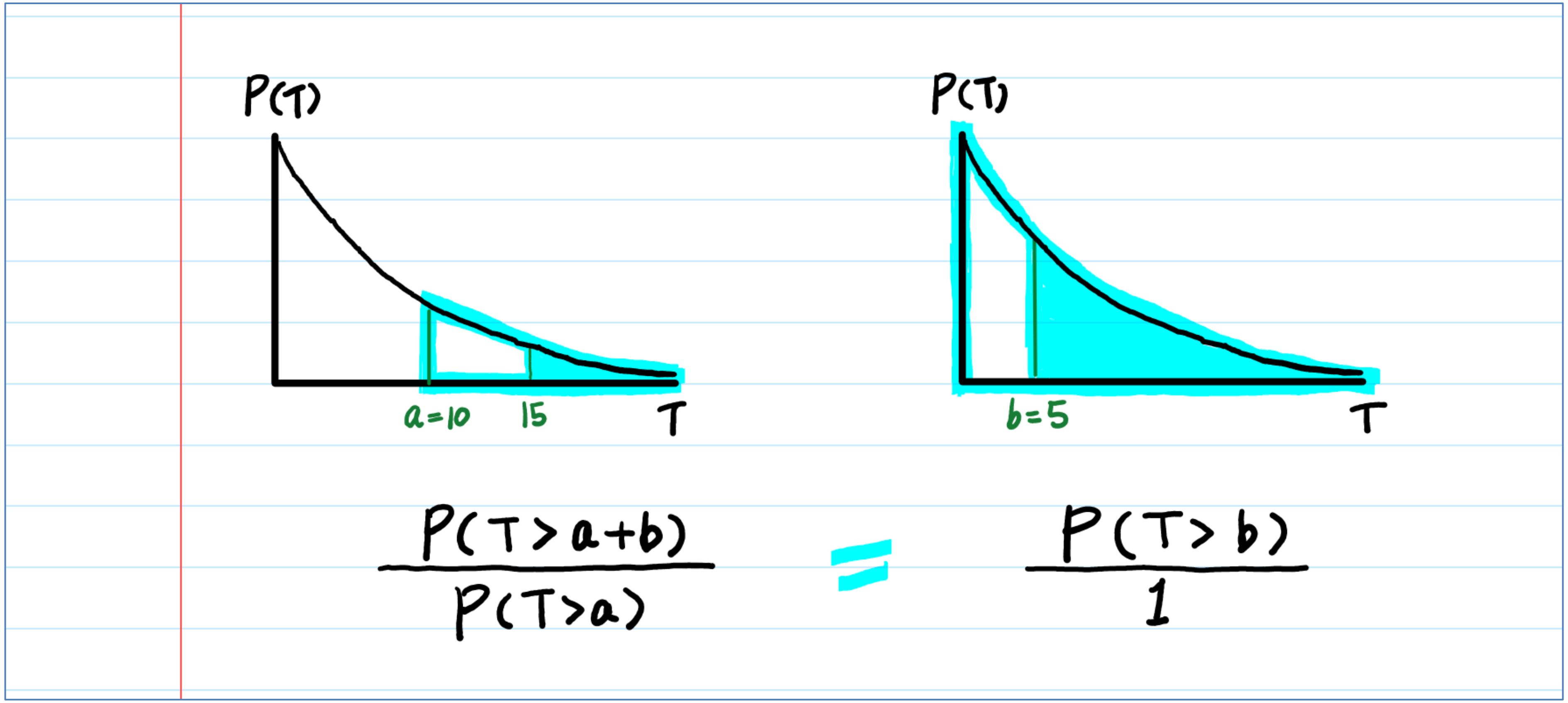
Parameter λ: 5



Funnest Fact: Exponential is Memoryless!

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

$$\text{P}(X > s + t | X > s) = \text{P}(X > t)$$ What if *s* time has passed?



Funnest Fact: Exponential is Memoryless!

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

$$P(X > s + t | X > s) = P(X > t)$$ What if *s* time has passed?

Which is something we can prove:

$P(X > s+t \mid X > s) = \dfrac{P(X > s+t \text{ and } X > s)}{P(X > s)}$	Def of conditional prob.
$= \dfrac{P(X > s+t)}{P(X > s)}$	Because $X > s + t$ implies $X > s$
$= \dfrac{1 - F_X(s+t)}{1 - F_X(s)}$	Def of CDF
$= \dfrac{e^{-\lambda(s+t)}}{e^{-\lambda s}}$	By CDF of Exp
$= e^{-\lambda t}$	Simplify
$= 1 - F_X(t)$	By CDF of Exp
$= P(X > t)$	Def of CDF