Week 3

Continuous Random Variables

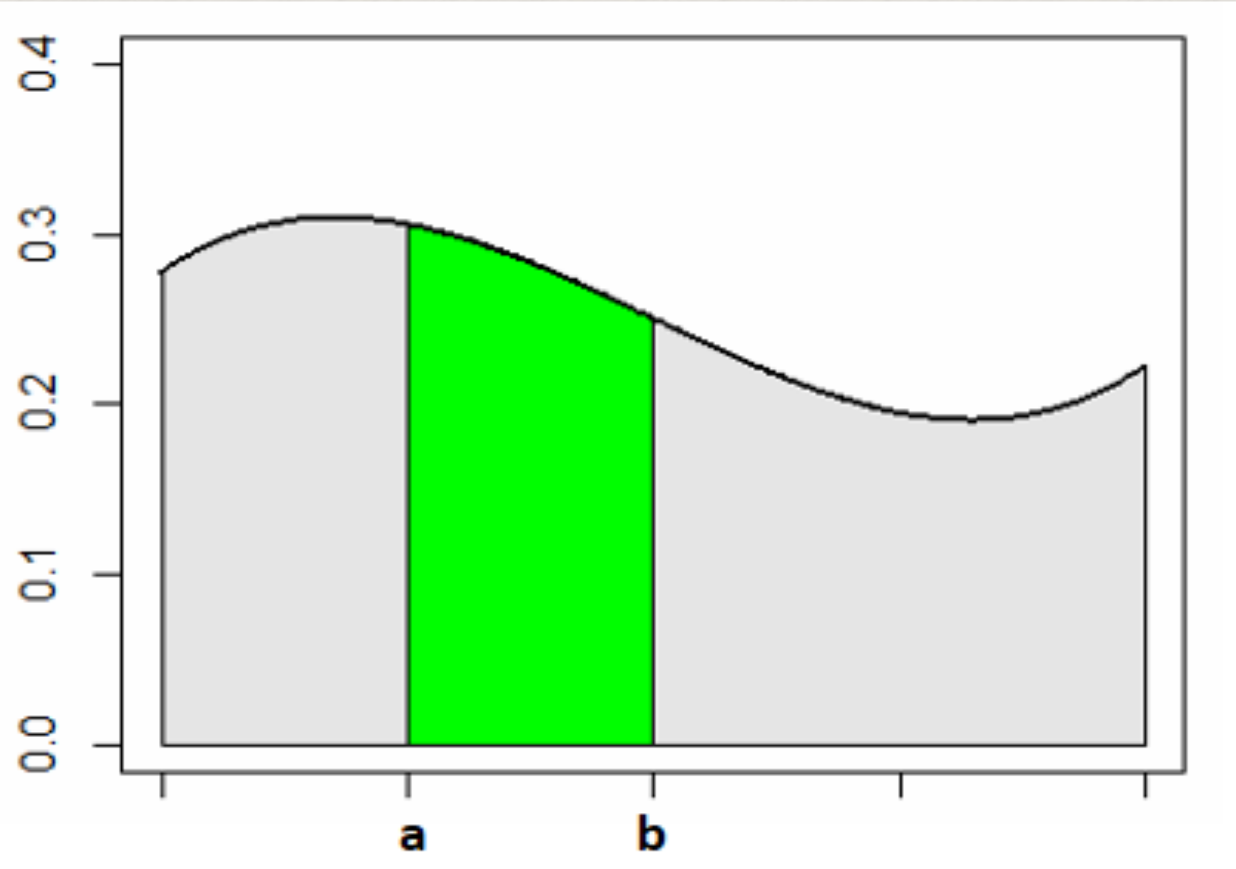
ST 314
Introduction to Statistics for Engineers

Oregon State University

Continuous Probability Distributions

For a continuous random variable X, the **Probability Density Function (PDF)**, denoted by $f(X)$, is a mathematical expression for the shape of the distribution of X.

We use the pdf to calculate probabilities between two values of X and expected values of X.



Probability Density Function (PDF)

- The probability density function is a continuous function of the random variable X such that

$$P(c \leq X \leq d) = \int_c^d f(x)\,dx$$

$P(X = c) = 0$

$$P(X = c) = P(c \leq X \leq c) = \int_c^c f(x)\,dx$$

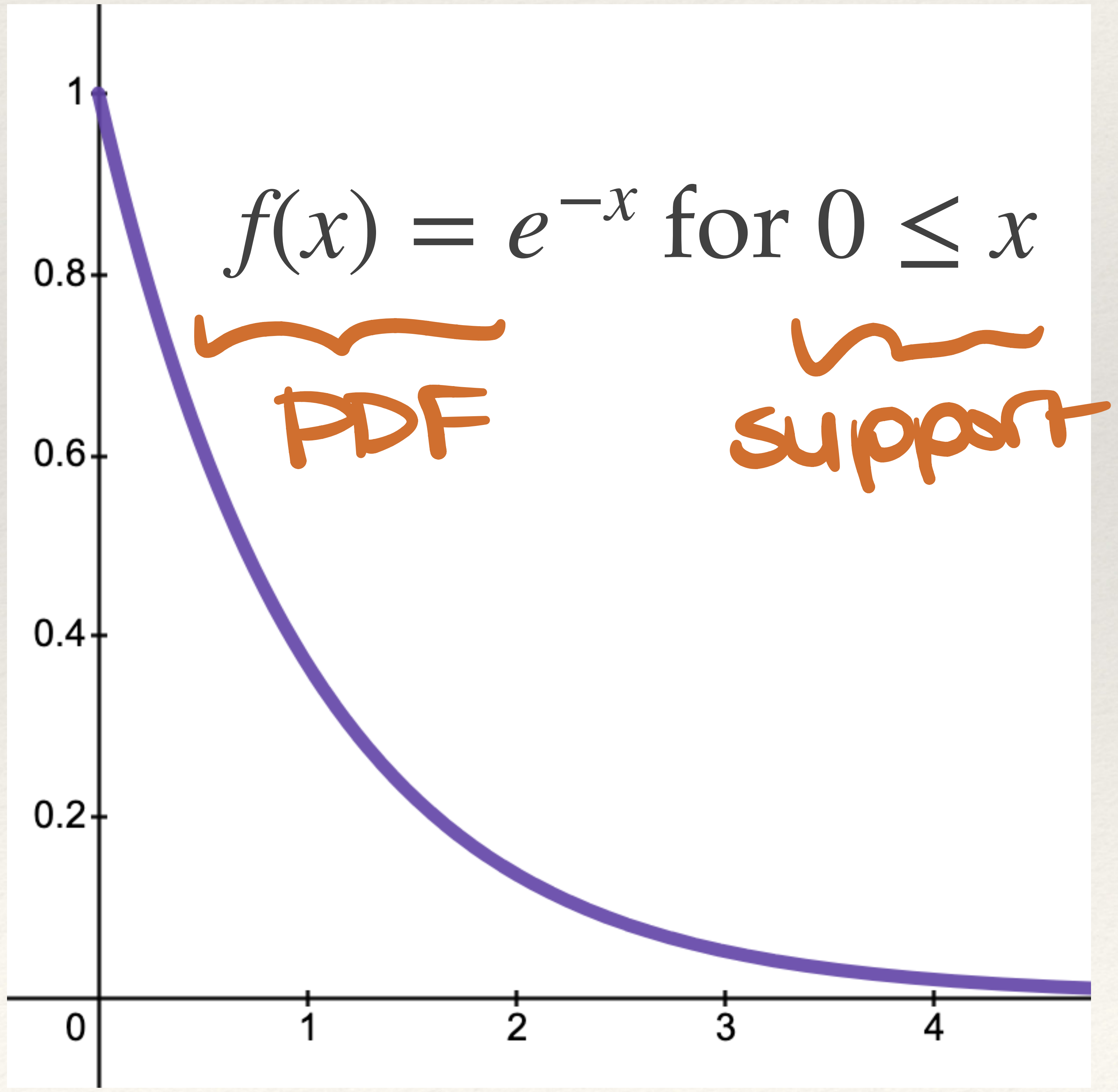
- $f(x) \geq 0$

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

change to the bounds (support)
that the random can take on

Probability Density Function (PDF)

Let X be the random variable that represents the number of minutes between customers who enter the drive thru of a fast food restaurant. Let's model the time between customers with the following probability density function.



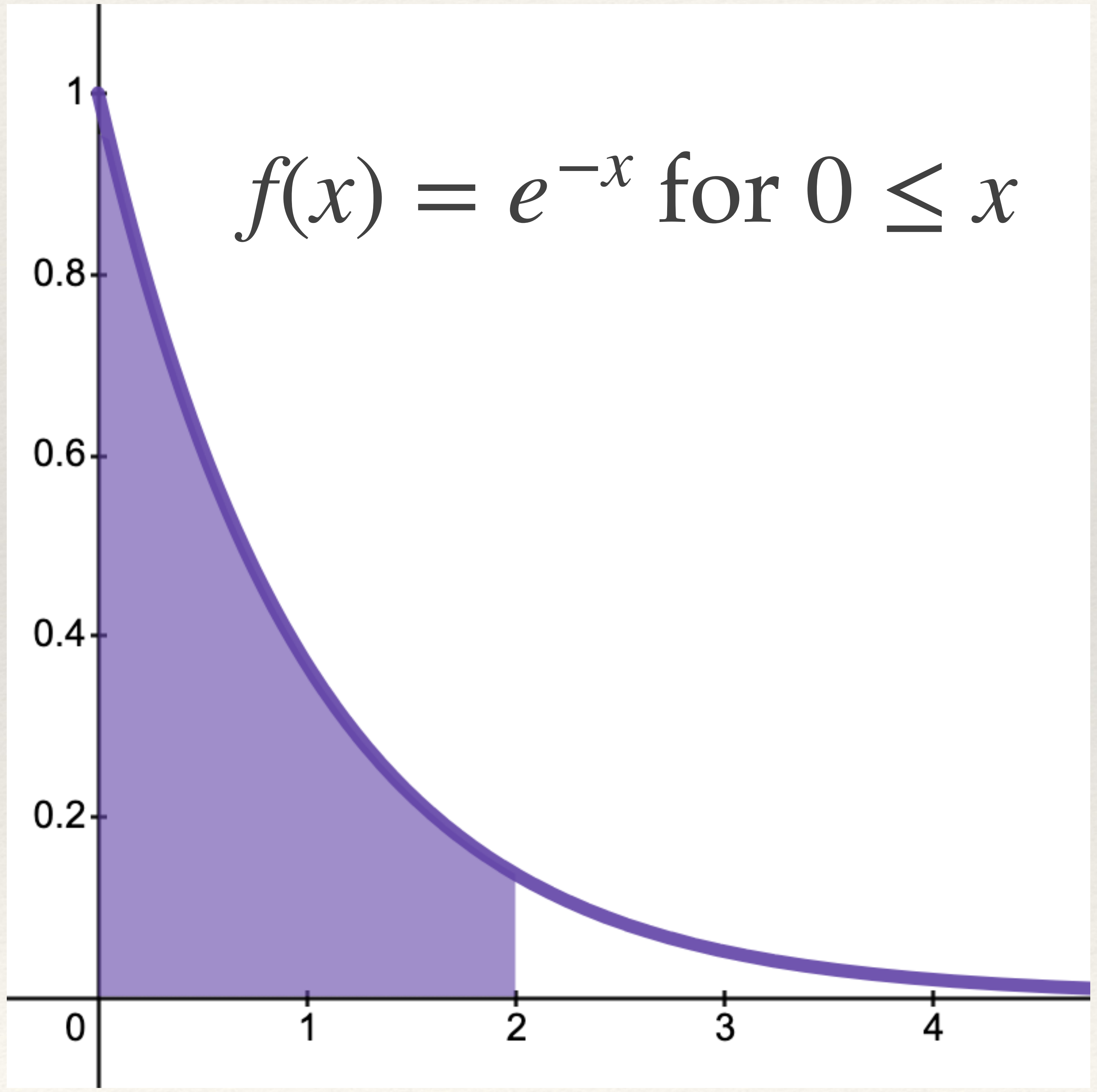
$$\underbrace{f(x) = e^{-x}}_{\text{PDF}} \text{ for } \underbrace{0 \leq x}_{\text{support}}$$

Does the area under $f(x)$ equal 1?

$$\int_0^\infty e^{-x}\,dx = -e^{-x}\Big|_0^\infty$$

$$= 0-(-1) = 1$$

Probability Density Function (PDF)



A customer has just entered the drive thru. Use the PDF to determine the probability that the next customer will arrive within 2 minutes.

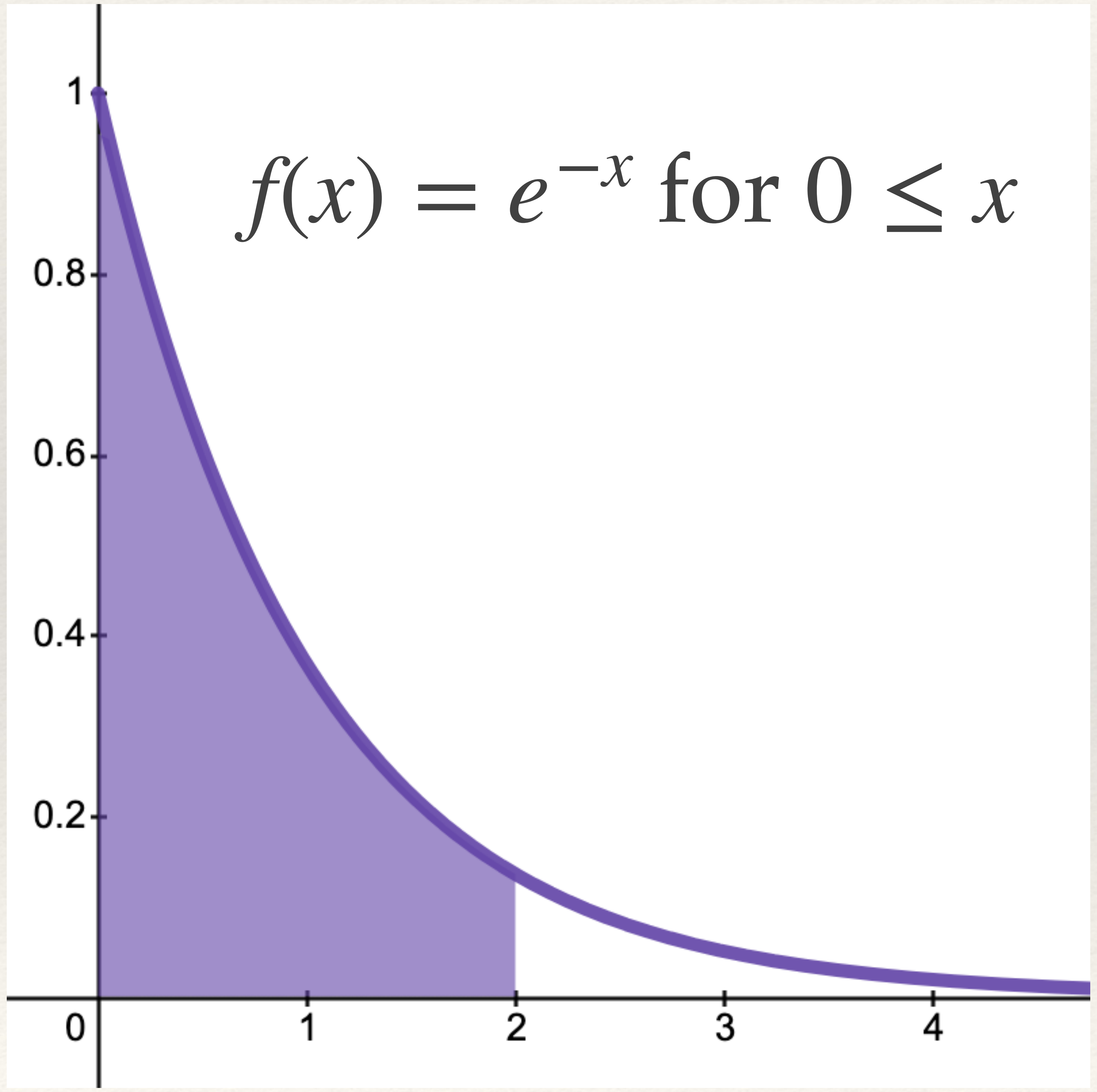
$$P(0 \leq X \leq 2) = \int_0^2 e^{-x}\,dx$$

$$= -e^{-x} \Big|_0^2$$

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

$$= 0.8647$$

Probability Density Function (PDF)



A customer has just entered the drive thru. Use the PDF to determine the probability that the next customer will arrive ~~within~~ 2 minutes.

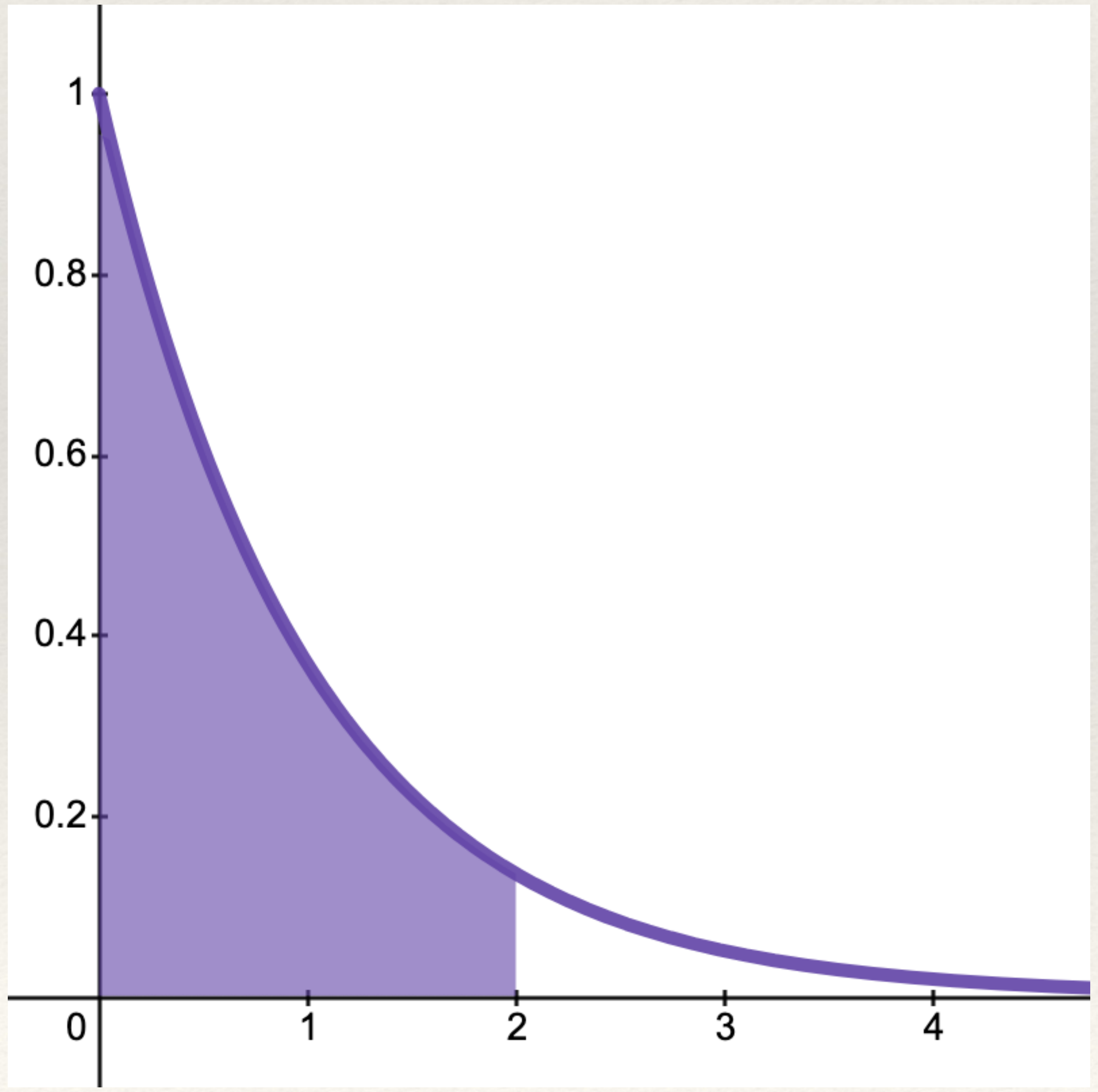
in more than

$$P(X > 2) = P(X \geq 2)$$

$$= \int_2^\infty e^{-x}\,dx$$

Cumulative Distribution Function

For a continuous random variable X, the **Cumulative Density Function (CDF)** is used to calculate the probability the random variable X will be less than or equal to some real number x.



$$F(x) = P(X \leq x) = \int_{\underbrace{-\infty}_{\text{change to lower bound of the support}}}^{x} f(t)\,dt$$

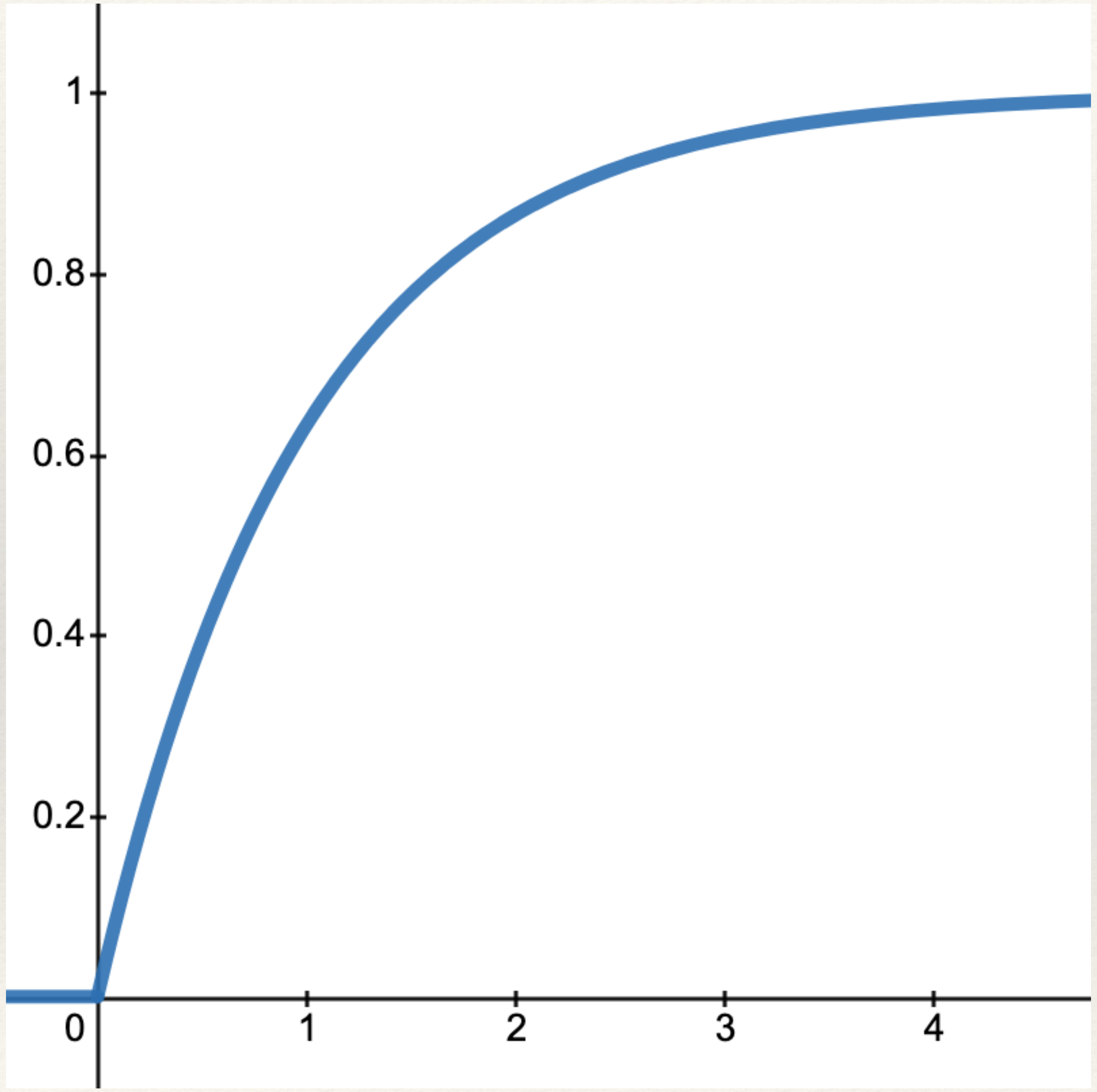
Cumulative Distribution Function

Properties

- $F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt$ PDF (def).
- $0 \leq F(x) \leq 1$
- $P(c \leq X \leq d) = F(d) - F(c) = \int_{c}^{d} f(x)\,dx$
- $F'(x) = f(x)$

  1st derivative of CDF equals the PDF

Cumulative Distribution Function (CDF)



The random variable X has the following probability density function. Determine its CDF.

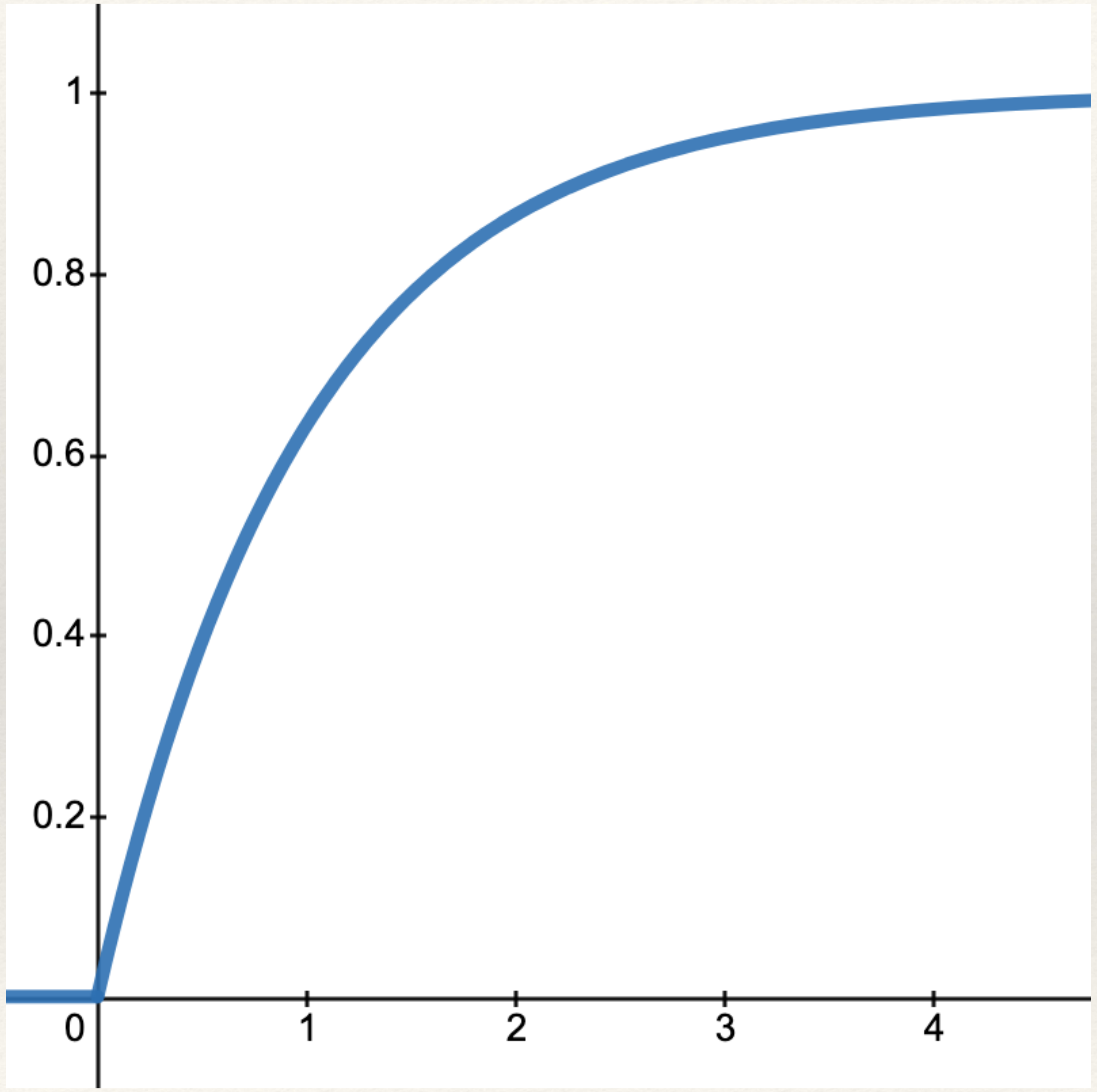
$f(x) = e^{-x}$ for $0 \leq x$

$$F(x) = \int_0^x e^{-t}\,dt$$

$$= -e^{-t}\Big|_0^x = -e^{-x} - (-1)$$

$$F(x) = P(X \leq x) = 1 - e^{-x}$$

Cumulative Distribution Function (CDF)



A customer has just entered the drive thru. Use the CDF to determine the probability that the next customer will arrive within 2 minutes.

$$F(x) = 1 - e^{-x}$$

$$P(X \leq 2) = F(2) = 1 - e^{-2}$$

$$= 0.8647$$

Expectation

Let X be a continuous random variable with PDF $f(x)$ and possible outcomes from $-\infty$ to ∞.

The **expected value** is the value we would expect X to take on. This is the average or mean value, denoted by $E(X)$ or μ_x.

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

Discrete : $E(X) = \sum x P(X=x)$

Variance

The **variance**, σ_x^2, is a measurement of how much spread there is in the distribution of X.

$$\mathrm{Var}(X) = \underbrace{E(X^2)} - (E(X))^2 = E((X - E(X))^2)$$

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

The **standard deviation**, σ_x, is the size of a typical deviation away from μ_x and has the same units as X and μ_x.

$$SD(X) = \sigma_X = \sqrt{\mathrm{var}(X)}$$

Expectation & Variance Example

support

Suppose X is a continuous random variable defined on the interval [0,1] with the following probability density function: $f(x) = 3x^2$.

Find the expectation, $E(X)$, and the variance, $\text{Var}(X)$, of X.

$$E(X) = \int_0^1 x \cdot 3x^2\,dx = \int_0^1 3x^3\,dx = \frac{3}{4}x^4 \Big|_0^1 = \frac{3}{4}$$

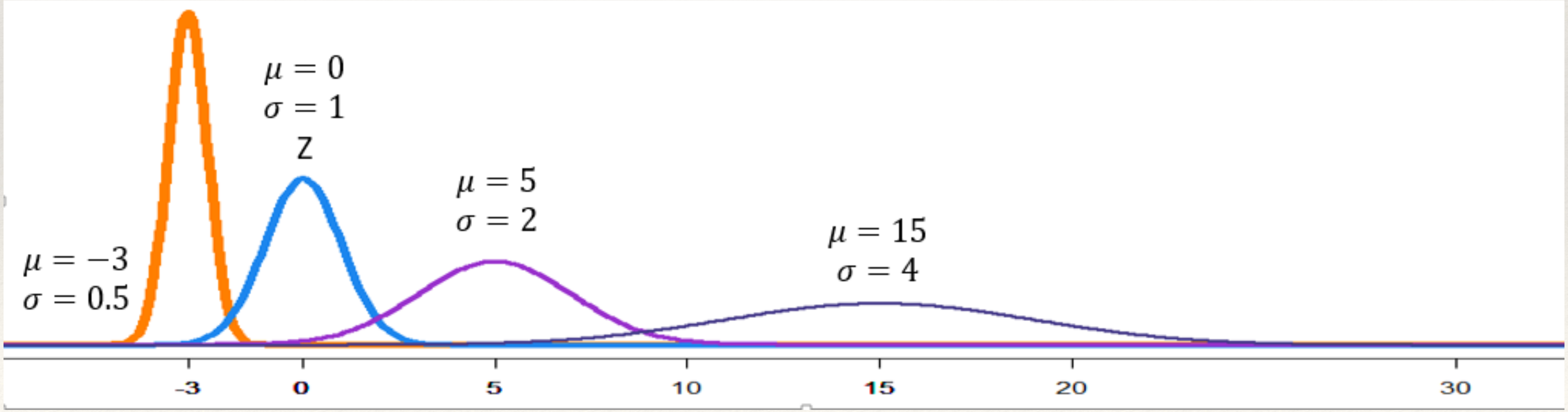
$$\text{var}(X) = E(X^2) - (E(X))^2$$

$$E(X^2) = \int_0^1 x^2 \cdot 3x^2\,dx = \int_0^1 3x^4\,dx = \frac{3}{5}x^5 \Big|_0^1 = \frac{3}{5}$$

$$\text{Var}(X) = \frac{3}{5} - \left(\frac{3}{4}\right)^2 = 0.0375 \qquad SD(X) = \sqrt{\text{Var}(X)} = \sqrt{0.0375}$$

The Normal Distribution

- Defined by two parameters: μ, σ $\quad X \sim N(\mu, \sigma)$
- Symmetric, single-peaked (uni-modal), bell-shaped
- Models many naturally occurring random variables



Normal PDF

A random variable X that follows a Normal distribution has a probability density function:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

- $E(X) = \mu$
- $Var(X) = \sigma^2$ $\quad SD(X) = \sigma$

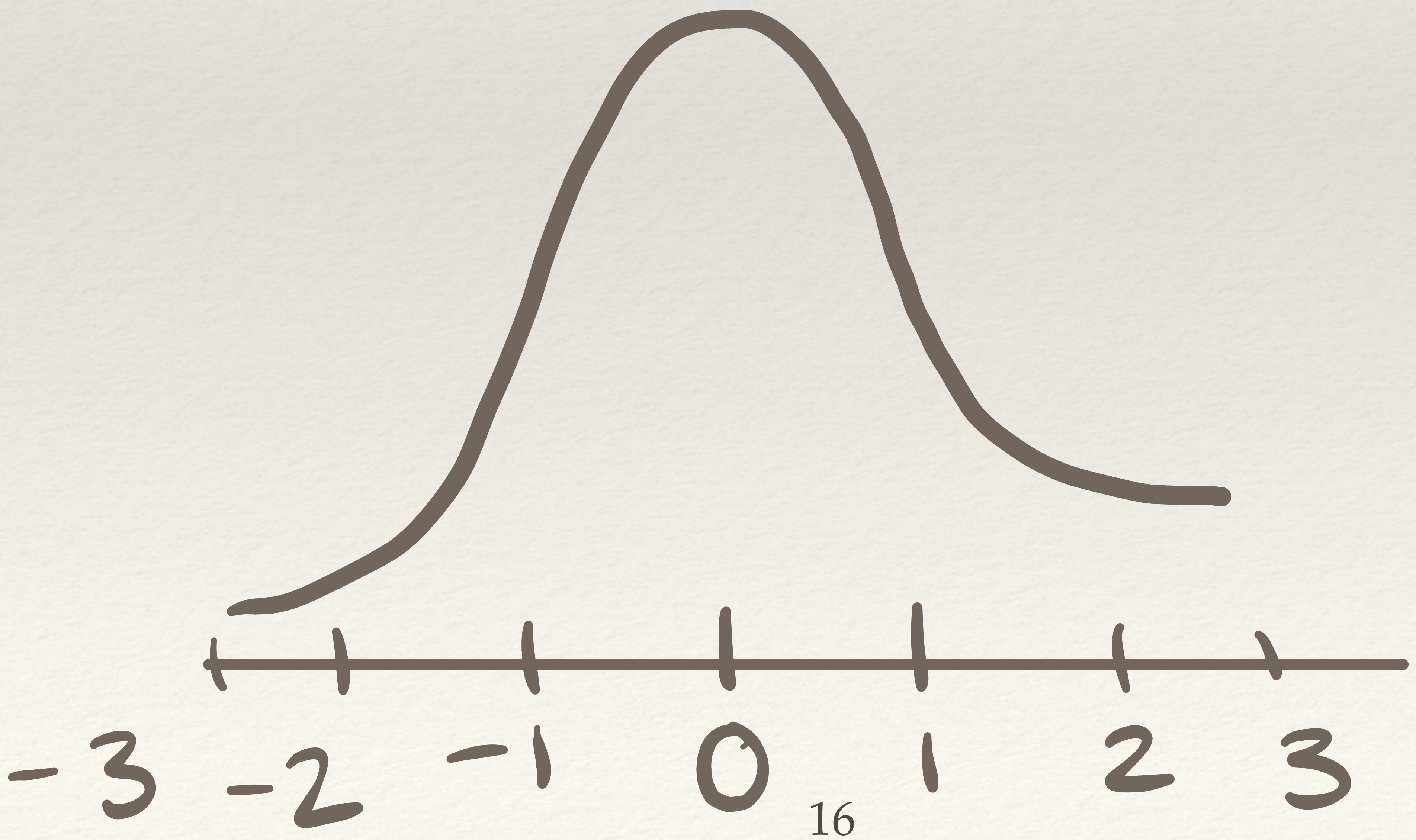
Standardizing Normal Distributions

For a Normal random variable *X*, a z-score represents the number of standard deviations any observation *x* is from the mean:

$$z = \frac{x - \mu}{\sigma}$$

The Standard Normal Distribution

- Any normally distributed random variable can be "standardized" or transformed into a **standard normal distribution**.
- The **standard normal** random variable, denoted by Z, has $\mu =$ ____0____ and $\sigma =$ ____1____.



Normal Random Variable Example

For a particular bridge, recorded vehicle speeds are normally distributed with a mean of 58 mph and a standard deviation of 10 mph.

Suppose a randomly chosen vehicle is going 40 miles per hour.

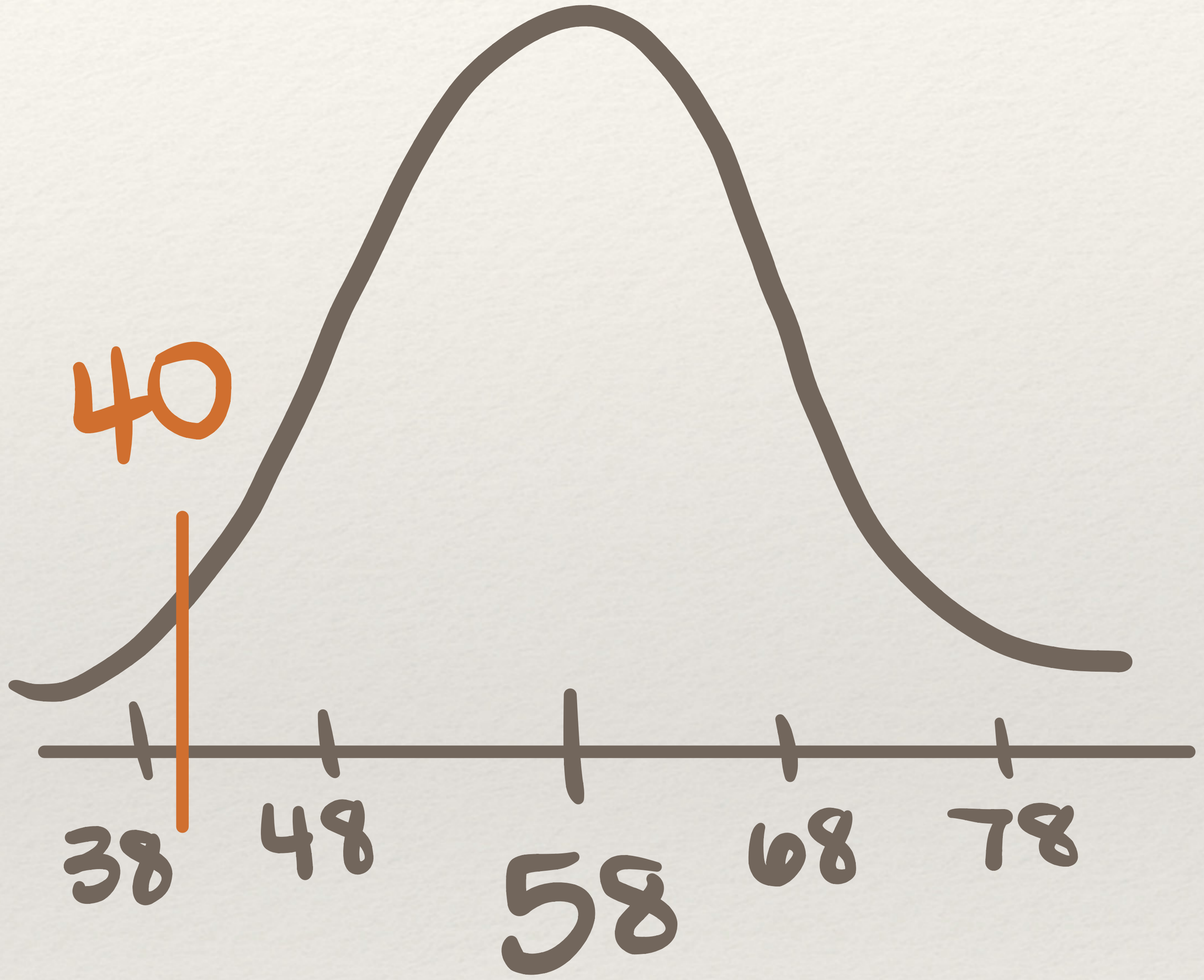
How many standard deviations away from the mean is 40 mph?

$\mu = 58$

$\sigma = 10$

$x = 40$

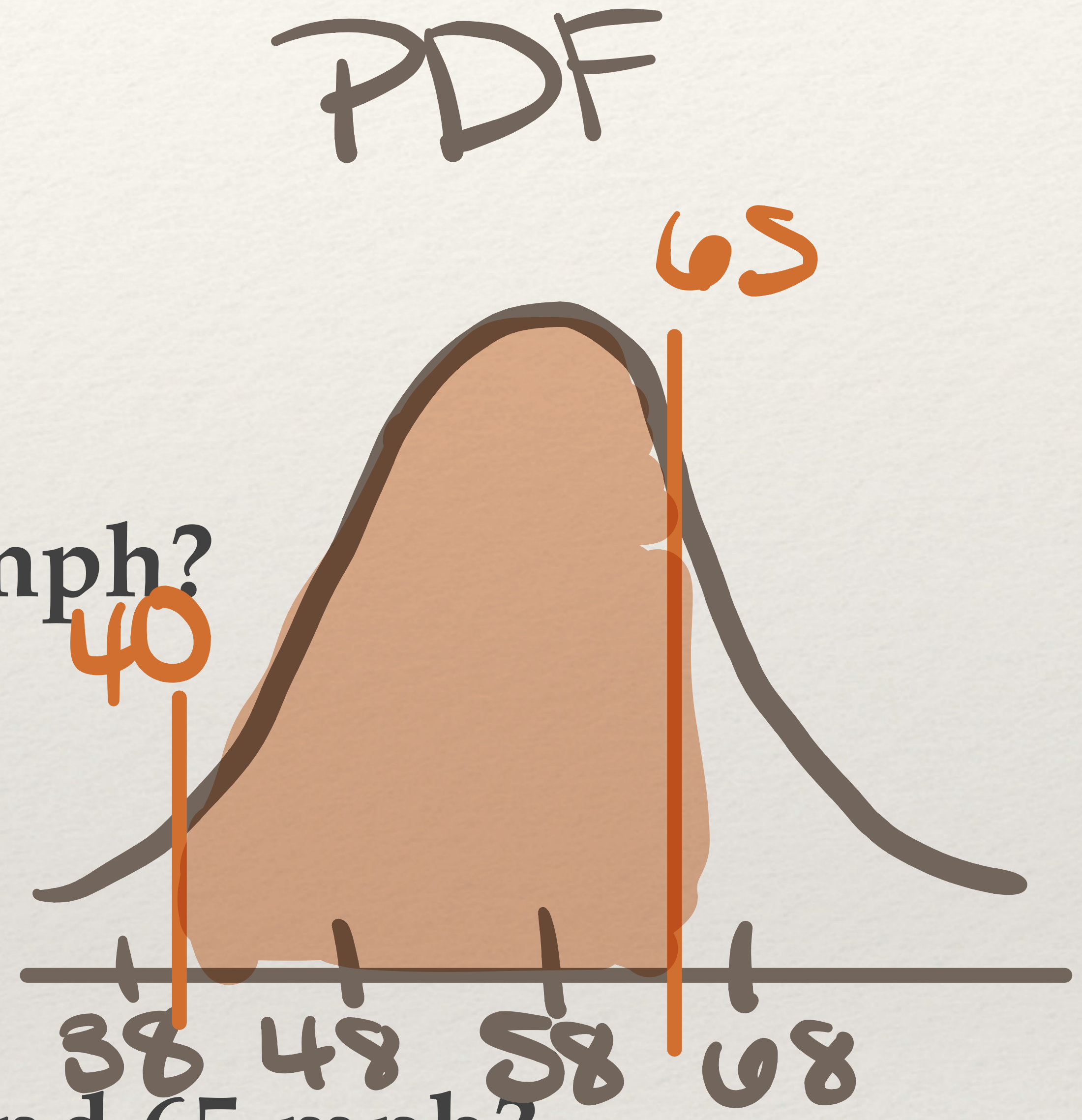
$$z = \frac{x-\mu}{\sigma} = \frac{40-58}{10} = -1.8$$



Normal Random Variable Example

For a particular bridge, recorded vehicle speeds are normally distributed with a mean of 58 mph and a standard deviation of 10 mph.

Suppose a randomly chosen vehicle is going 40 miles per hour.



What is the probability of a randomly selecting a vehicle going less than 40 mph?

$$P(X<40) = \int_{-\infty}^{40} \frac{1}{\sqrt{2\pi 10^2}} e^{-\frac{(x-58)^2}{2(10)^2}} \, dx$$

What is the probability of a randomly selecting a vehicle going between 40 and 65 mph?

$$P(40<X<65) = \text{pnorm}(65, 58, 10) - \text{pnorm}(40, 58, 10)$$