etherpad.wikimedia.org/p/312 for (anonymous) questions/comments!

Zoo of Continuous Random Variables

CSE 312 24Su

Lecture 14

Outline for Today

- Review concepts for continuous random variables
- Zoo of continuous random variables 🐻
  - *Continuous* uniform distribution
  - Exponential distribution
  - Normal distribution

Discrete Random Variables

The support has *finite* or *countably infinite values*

e.g., number of successes, number of trials till success, attendance at a class are all discrete because they take on a set of finite or countably infinite values

Some random experiments have <u>uncountably-infinite sample spaces</u>

> *How long until the next bus shows up?*

> *Throwing a dart on a board (what location does the dart land?)*

Continuous Random Variables

Random variables with a support of *uncountably-infinite values*

> *e.g., RVs that take on any **real** number in some interval(s) like distance, height, time, etc.*

Discrete RVs

Continuous RVs

Discrete RVs

Support is finite/countably infinite (e.g. *integers*)

Continuous RVs

Support is uncountably infinite (e.g., *real numbers*)

Discrete RVs

Support is finite/countably infinite (e.g. *integers*)

Continuous RVs

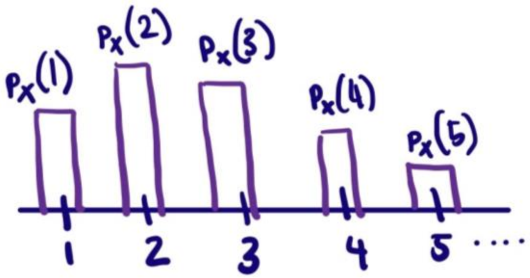
Support is uncountably infinite (e.g., *real numbers*)

$\mathbb{P}(\boldsymbol{X} = \boldsymbol{k}) = \frac{\mathbf{1}}{\infty} = \mathbf{0}$ so we don't use PMF. instead...

Discrete RVs

Support is finite/countably infinite (e.g. *integers*)

Probability mass function $p_X(k)$ gives probability of each value in support



$$\sum_{k \in \Omega_X} p_X(k) = 1$$

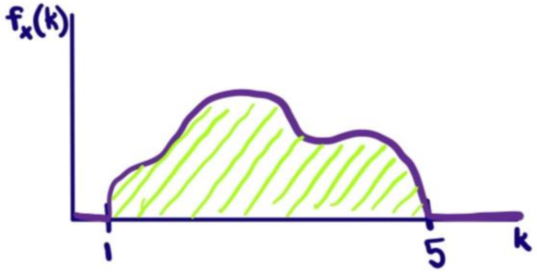
$$0 \le p_X(k) \le 1$$

Continuous RVs

Support is uncountably infinite (e.g., *real numbers*)

$\mathbb{P}(\boldsymbol{X} = \boldsymbol{k}) = \frac{\mathbf{1}}{\infty} = \mathbf{0}$ so we don't use PMF. instead...

Probability density function $\boldsymbol{f_X(k)}$ describes relative chances of taking values around k



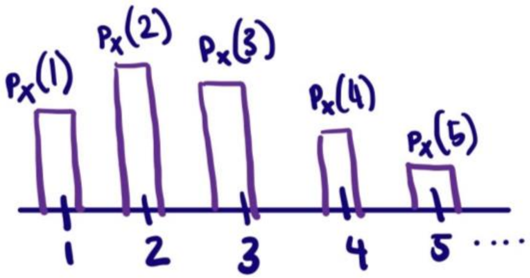
$$\int_{-\infty}^{\infty} f_X(k)\ \mathrm{d}k = 1$$

$$f_X(k) \ge 0$$

Discrete RVs

Support is finite/countably infinite (e.g. *integers*)

Probability mass function $p_X(k)$ gives probability of each value in support



$$\sum_{k \in \Omega_X} p_X(k) = 1$$

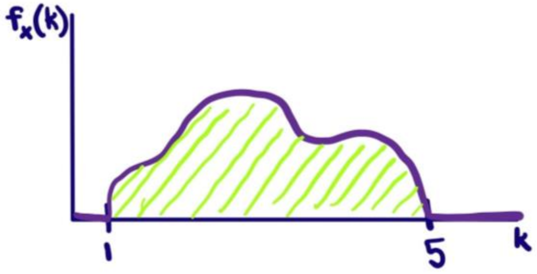
$$0 \le p_X(k) \le 1$$

Continuous RVs

Support is uncountably infinite (e.g., *real numbers*)

$\mathbb{P}(\boldsymbol{X} = \boldsymbol{k}) = \frac{\mathbf{1}}{\infty} = \mathbf{0}$ so we don't use PMF. instead…

Probability density function $\boldsymbol{f_X(k)}$ describes relative chances of taking values around k



$$\int_{-\infty}^{\infty} f_X(k)\ \mathrm{d}k = 1$$

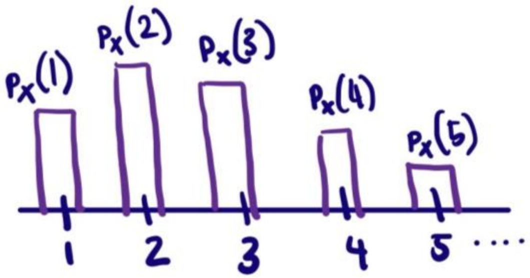
$$f_X(k) \ge 0$$

Cumulative Distribution Function (CDF) is the function $F_X(k) = \mathbb{P}(\boldsymbol{X} \le \boldsymbol{k})$

Discrete RVs

Support is finite/countably infinite (e.g. *integers*)

Probability mass function $p_X(k)$ gives probability of each value in support



$$\sum_{k \in \Omega_X} p_X(k) = 1$$

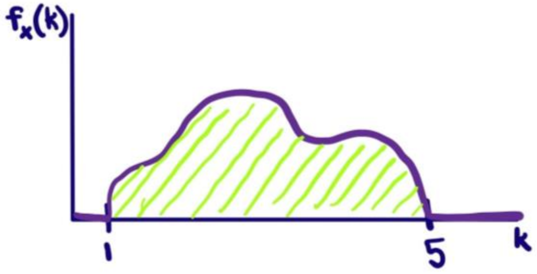
$$0 \leq p_X(k) \leq 1$$

Continuous RVs

Support is uncountably infinite (e.g., *real numbers*)

$\mathbb{P}(\boldsymbol{X} = \boldsymbol{k}) = \frac{1}{\infty} = \mathbf{0}$ so we don't use PMF. instead...

Probability density function $\boldsymbol{f_X(k)}$ describes relative chances of taking values around k



$$\int_{-\infty}^{\infty} f_X(k) \, \mathrm{d}k = 1$$

$$f_X(k) \geq 0$$

Cumulative Distribution Function (CDF) is the function $F_X(k) = \mathbb{P}(\boldsymbol{X} \leq \boldsymbol{k})$

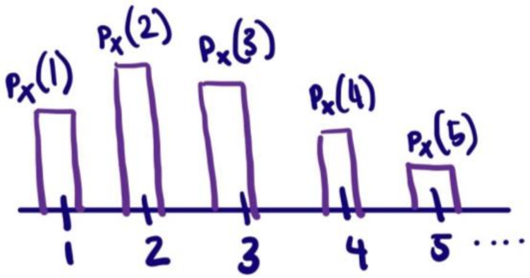
Sum up the probabilities of values $\leq k$

Integrate over values $\leq k$: $F_X(k) = \int_{-\infty}^{k} f_X(z) dz$

Discrete RVs

Support is <u>finite/countably infinite</u> (e.g. *integers*)

Probability mass function $p_X(k)$ gives probability of each value in support



$$\sum_{k \in \Omega_X} p_X(k) = 1$$

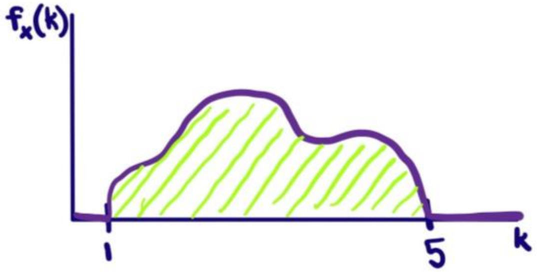
$$0 \leq p_X(k) \leq 1$$

Continuous RVs

Support is <u>uncountably infinite</u> (e.g., *real numbers*)

$\mathbb{P}(\boldsymbol{X} = \boldsymbol{k}) = \frac{\mathbf{1}}{\infty} = \mathbf{0}$ so we don't use PMF. instead...

Probability <u>density</u> function $\boldsymbol{f_X(k)}$ describes relative chances of taking values around k



$$\int_{-\infty}^{\infty} f_X(k)\ \mathrm{d}k = 1$$

$$f_X(k) \geq 0$$

Cumulative Distribution Function (CDF) is the function $F_X(k) = \mathbb{P}(\boldsymbol{X} \leq \boldsymbol{k})$

<u>Sum</u> up the probabilities of values $\leq k$

<u>Integrate</u> over values $\leq k$: $F_X(k) = \int_{-\infty}^{k} f_X(z) dz$

Expectation: $\mathbb{E}[X] = \sum_{k \in \Omega_X} (k \cdot p_X(k))$

$\mathbb{E}[g(X)] = \sum_{k \in \Omega_X} (g(k) \cdot p_X(k))$

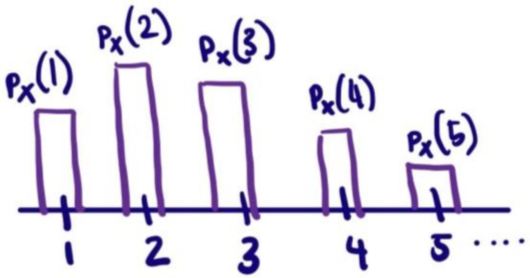
Expectation: $\mathbb{E}[X] = \int_{-\infty}^{\infty} z \cdot f_X(z)\ \mathrm{dz}$

$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(z) \cdot f_X(z)\ \mathrm{dz}$

Discrete RVs

Support is finite/countably infinite (e.g. *integers*)

Probability mass function $p_X(k)$ gives probability of each value in support



$$\sum_{k \in \Omega_X} p_X(k) = 1$$

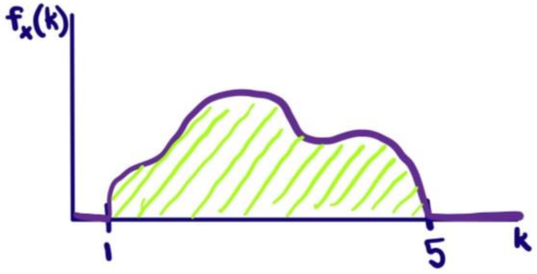
$$0 \le p_X(k) \le 1$$

Continuous RVs

Support is uncountably infinite (e.g., *real numbers*)

$\mathbb{P}(\boldsymbol{X} = \boldsymbol{k}) = \frac{1}{\infty} = \boldsymbol{0}$ so we don't use PMF. instead...

Probability density function $\boldsymbol{f_X(k)}$ describes relative chances of taking values around k



$$\int_{-\infty}^{\infty} f_X(k)\ \mathrm{d}k = 1$$

$$f_X(k) \ge 0$$

Cumulative Distribution Function (CDF) is the function $F_X(k) = \mathbb{P}(\boldsymbol{X} \le \boldsymbol{k})$

Sum up the probabilities of values $\le k$

Integrate over values $\le k$: $F_X(k) = \int_{-\infty}^{k} f_X(z)dz$

Expectation: $\mathbb{E}[X] = \sum_{k \in \Omega_X}(k \cdot p_X(k))$

$\mathbb{E}[g(X)] = \sum_{k \in \Omega_X}(g(k) \cdot p_X(k))$

Expectation: $\mathbb{E}[X] = \int_{-\infty}^{\infty} z \cdot f_X(z)\ \mathrm{d}z$

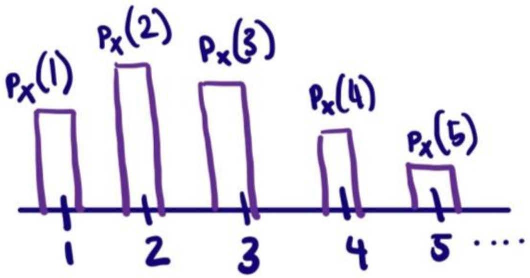
$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(z) \cdot f_X(z)\ \mathrm{d}z$

Variance is $Var(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$

Discrete RVs

Support is <u>finite/countably infinite</u> (e.g. *integers*)

Probability mass function $p_X(k)$ gives probability of each value in support



$$\sum_{k \in \Omega_X} p_X(k) = 1$$

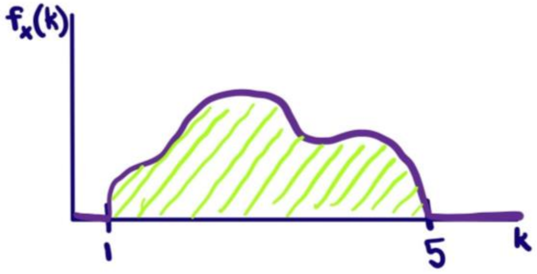
$$0 \leq p_X(k) \leq 1$$

Continuous RVs

Support is <u>uncountably infinite</u> (e.g., *real numbers*)

$\mathbb{P}(\boldsymbol{X} = \boldsymbol{k}) = \frac{\mathbf{1}}{\infty} = \mathbf{0}$ so we don't use PMF. instead...

Probability <u>density</u> function $\boldsymbol{f_X(k)}$ describes relative chances of taking values around k



$$\int_{-\infty}^{\infty} f_X(k)\ \mathrm{d}k = 1$$

$$f_X(k) \geq 0$$

Cumulative Distribution Function (CDF) is the function $F_X(k) = \mathbb{P}(\boldsymbol{X} \leq \boldsymbol{k})$

<u>*Sum*</u> up the probabilities of values $\leq k$

<u>*Integrate*</u> over values $\leq k$: $F_X(k) = \int_{-\infty}^{k} f_X(z)dz$

Expectation: $\mathbb{E}[X] = \sum_{k \in \Omega_X} (k \cdot p_X(k))$

$\mathbb{E}[g(X)] = \sum_{k \in \Omega_X} (g(k) \cdot p_X(k))$

Expectation: $\mathbb{E}[X] = \int_{-\infty}^{\infty} z \cdot f_X(z)\ \mathrm{d}z$

$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(z) \cdot f_X(z)\ \mathrm{d}z$

Variance is $Var(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$

Linearity of expectation and properties of expectation and variance applies in both!

Zoo of **Continuous** Random variables

and get practice with working with continuous random variables!

Continuous Zoo

This zoo defines common patterns for continuous random variables and gives us the PDF, CDF, expectation, and variance, so we don't have to compute it every time!

$X \sim \mathrm{Unif}(a, b)$	$X \sim \mathrm{Exp}(\lambda)$	$X \sim \mathcal{N}(\mu, \sigma^2)$
$f_X(k) = \frac{1}{b-a}$ for $a \le k \le b$ $F_X(k) = \frac{x-a}{b-a}$ if $a \le k < b$ $\mathbb{E}[X] = \frac{a+b}{2}$ $\mathrm{Var}(X) = \frac{(b-a)^2}{12}$	$f_X(k) = \lambda e^{-\lambda k}$ for $k \ge 0$ $F_X(k) = 1 - e^{-\lambda k}$ if $k \ge 0$ $\mathbb{E}[X] = \frac{1}{\lambda}$ $\mathrm{Var}(X) = \frac{1}{\lambda^2}$	$f_X(k) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$ $F_X(k) = \Phi\left(\frac{k-\mu}{\sigma}\right)$ $\mathbb{E}[X] = \mu$ $\mathrm{Var}(X) = \sigma^2$

It's a smaller zoo, but it's just as much fun! :D

Continuous Uniform Distribution

Scenario: Pick a random *real* number between a and b. X is our choice.
> X is a **uniform random real number** between a and b -> $X \sim \mathrm{Unif}(a, b)$

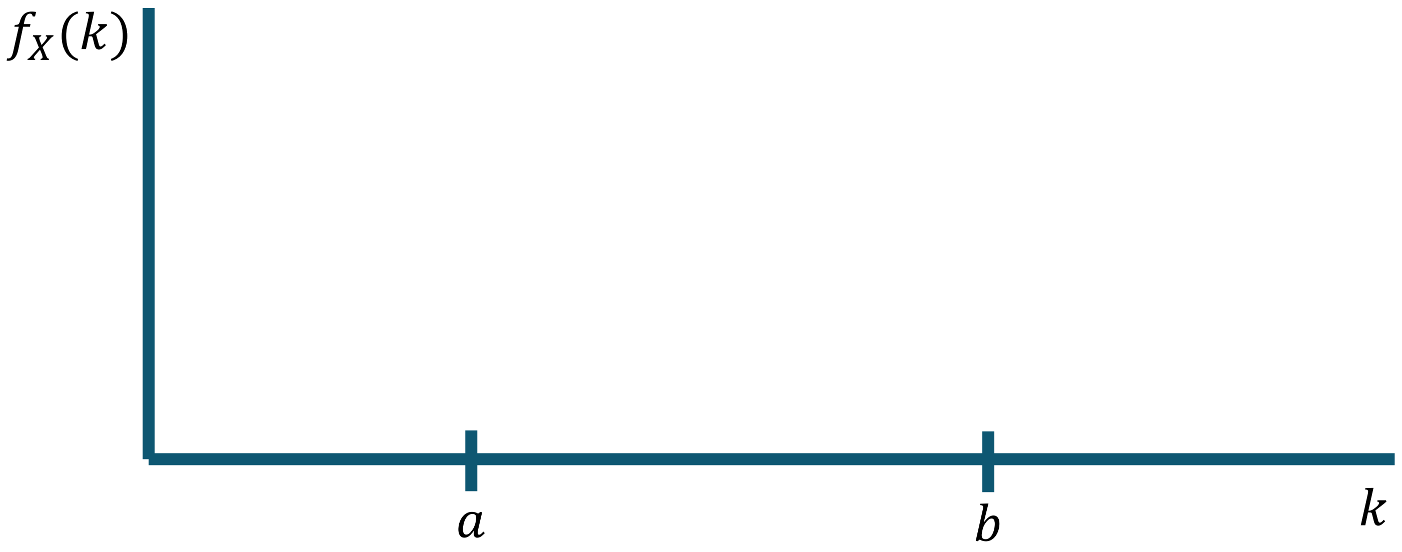
We saw the discrete uniform distribution before – it took on an integer between a and b
This is the **continuous** uniform distribution – it takes on a **real number** between a and b

Continuous Uniform Distribution - **PDF**

> X is a **uniform random real number** between a and b -> $X \sim \mathrm{Unif}(a, b)$

What is the PDF, $f_X(k)$?

Draw a picture and think about the properties the PDF must have!



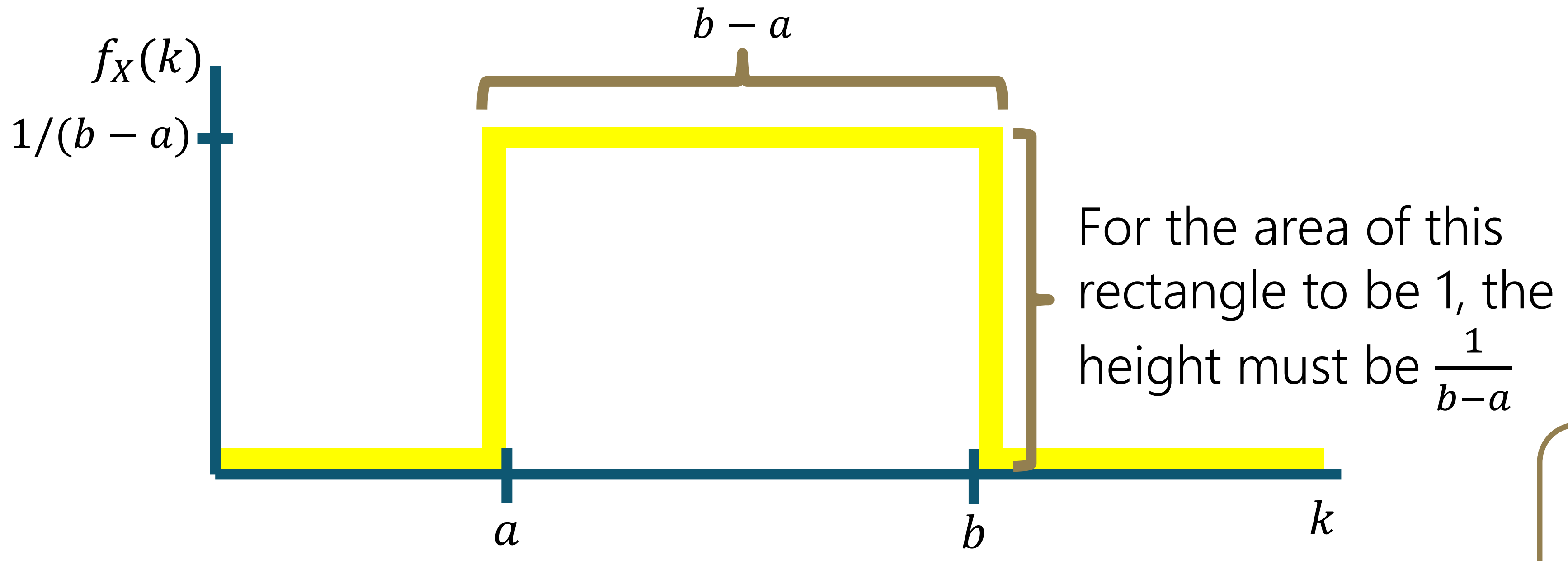
Fill out the poll everywhere:
pollev.com/cse312

Continuous Uniform Distribution - **PDF**

> X is a **uniform random real number** between a and b -> $X \sim \text{Unif}(a, b)$

What is the PDF, $f_X(k)$?

Draw a picture and think about the properties the PDF must have!



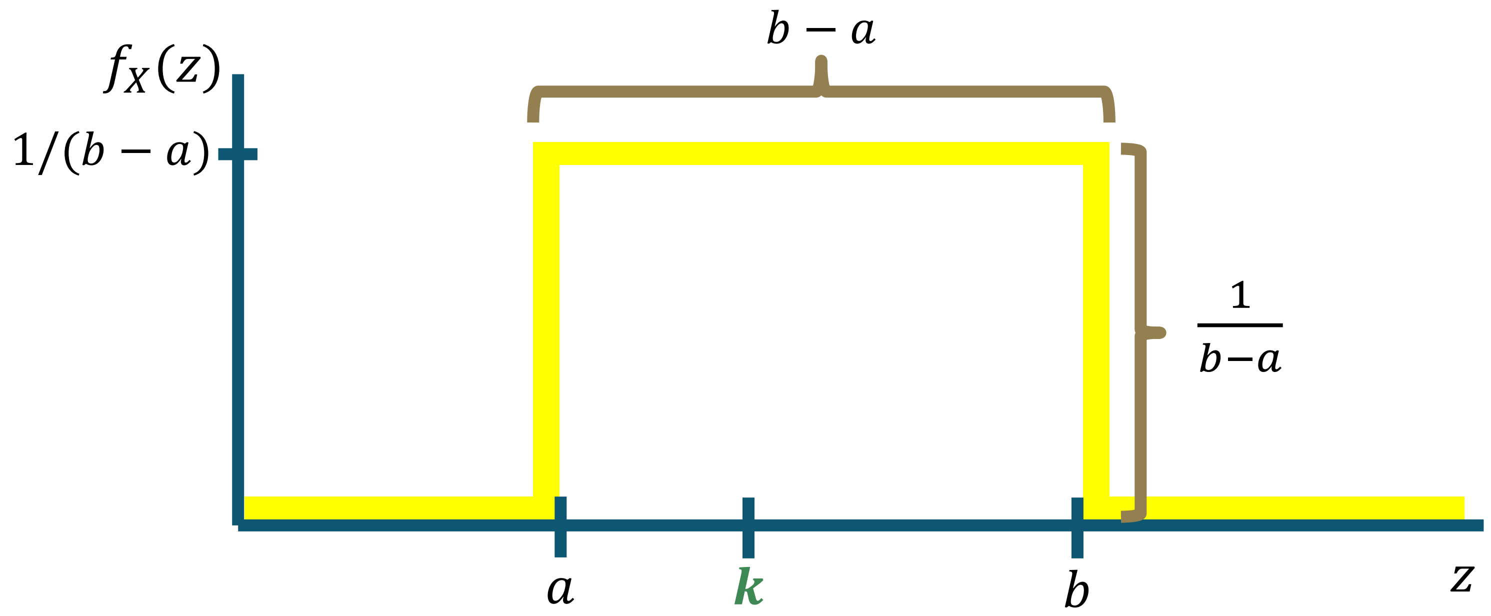
$$f_X(k) = \begin{cases} \dfrac{1}{b-a} & \text{if } a \le k \le b \\ 0 & \text{otherwise} \end{cases}$$

The area under a PDF must be 1. $\int_{-\infty}^{\infty} f_X(z)\, dz = 1$

Continuous Uniform Distribution – **CDF**(Visual)

> X is a **uniform random real number** between a and b -> $X \sim \text{Unif}(a, b)$

What is the CDF, $F_X(k)$? *Let's compute this "visually"….*



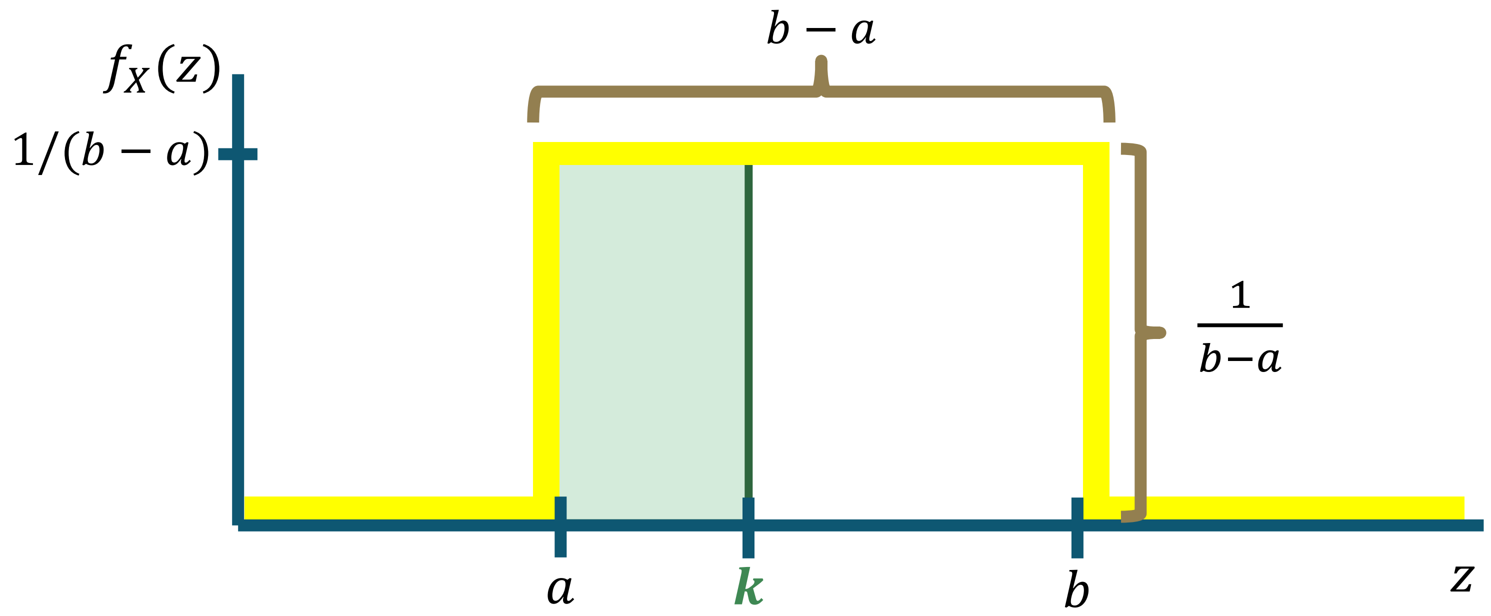
$$F_X(k) = \begin{cases} \\ \\ \end{cases}$$

Continuous Uniform Distribution – **CDF**(Visual)

> X is a **uniform random real number** between a and b -> $X \sim \text{Unif}(a, b)$

What is the CDF, $F_X(k)$? *Let's compute this "visually"….*

$F_X(k) = P(X \leq k)$ is the area of the green region below: $\frac{k-a}{b-a}$ if $a \leq k < b$



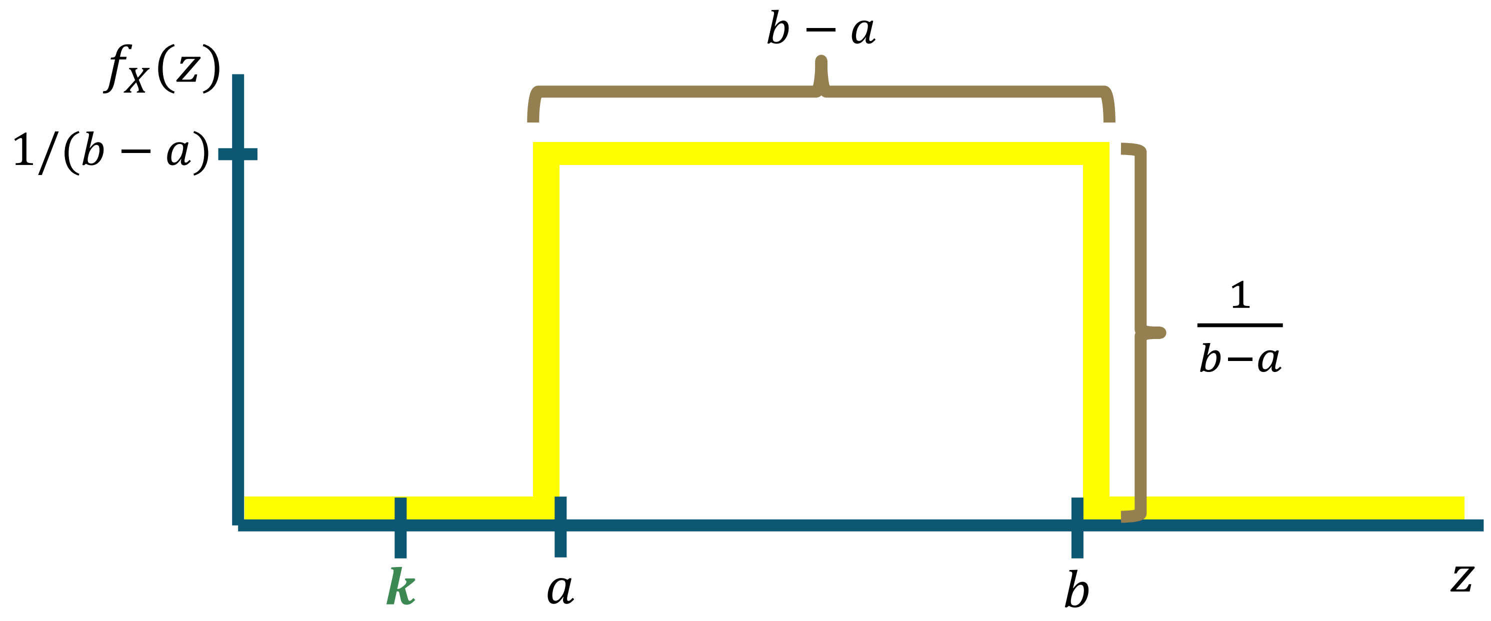
$$F_X(k) = \begin{cases} \frac{\boldsymbol{k-a}}{\boldsymbol{b-a}} & \textbf{if } \boldsymbol{a \leq k \leq b} \end{cases}$$

Continuous Uniform Distribution – **CDF**(Visual)

> X is a **uniform random real number** between a and b -> $X \sim \text{Unif}(a, b)$

What is the CDF, $F_X(k)$? *Let's compute this "visually"....*

$F_X(k) = P(X \leq k)$ is the area of the green region below: $\frac{k-a}{b-a}$ if $a \leq k < b$



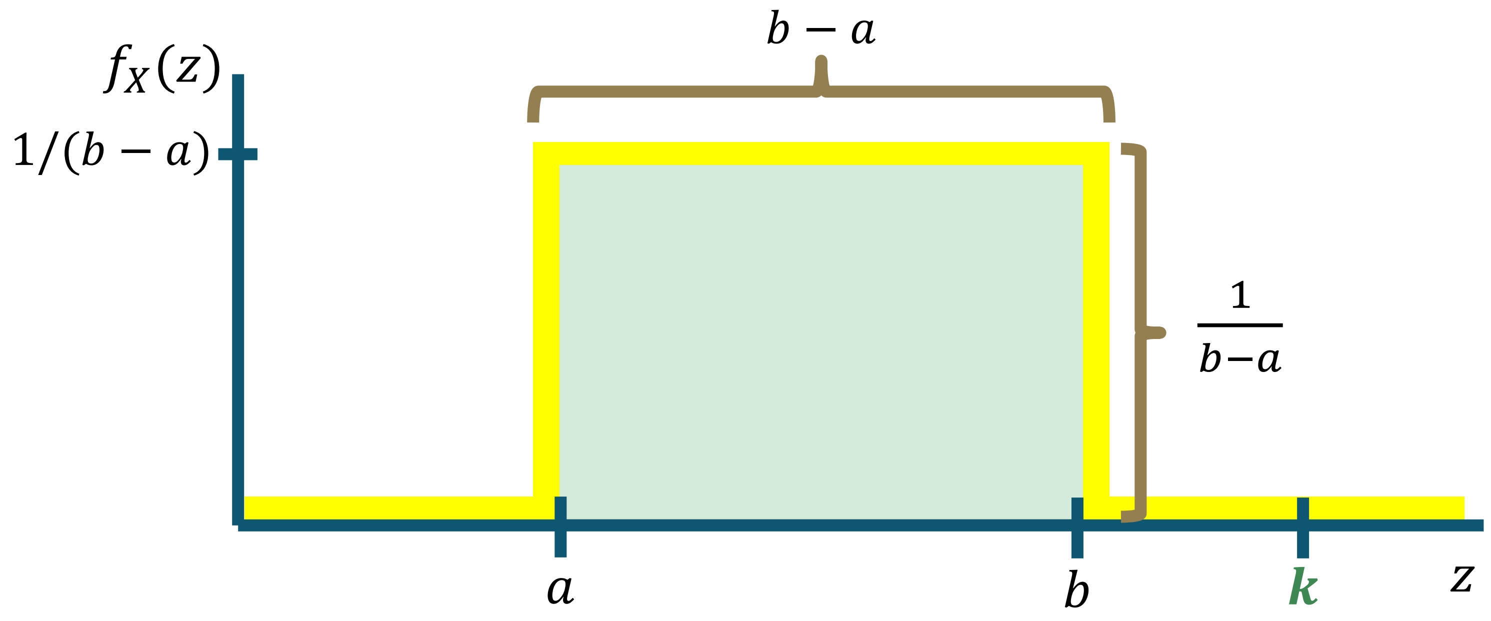
$$F_X(k) = \begin{cases} \mathbf{0} & \textbf{if } \boldsymbol{k < a} \\ \dfrac{k-a}{b-a} & \text{if } a \leq k \leq b \\ & \end{cases}$$

Continuous Uniform Distribution – **CDF**(**Visual**)

> X is a **uniform random real number** between a and b -> $X \sim \text{Unif}(a,b)$

What is the CDF, $F_X(k)$? *Let's compute this "visually"....*

$F_X(k) = P(X \leq k)$ is the area of the green region below: $\frac{k-a}{b-a}$ if $a \leq k < b$



$$F_X(k) = \begin{cases} 0 & \text{if } k < a \\ \dfrac{k-a}{b-a} & \text{if } a \leq k \leq b \\ \mathbf{1} & \textbf{if } \boldsymbol{k \geq b} \end{cases}$$

Continuous Uniform Distribution – **CDF(Integral)**

> X is a **uniform random real number** between a and b -> $X \sim \text{Unif}(a, b)$

What is the CDF, $F_X(k)$? *Let's compute this "algebraically"....*

The CDF is: $F_X(k) = \int_{-\infty}^{k} f_X(z)dz$

We can compute this for the ranges based on what $\boldsymbol{k}$ is and what $\boldsymbol{f_X(z)}$ is for the values of $\boldsymbol{z}$ less than $\boldsymbol{k}$

Case when $\boldsymbol{k \le a}$: $F_X(k) = \int_{-\infty}^{k} f_X(z)dz = \int_{-\infty}^{k} 0\, dz = 0$

Case when $\boldsymbol{a \le k \le b}$: $F_X(k) = \int_{-\infty}^{k} f_X(z)dz = \int_{-\infty}^{a} 0\, dz + \int_{\mathrm{a}}^{k} \frac{1}{b-a} dz = \frac{k-a}{b-a}$

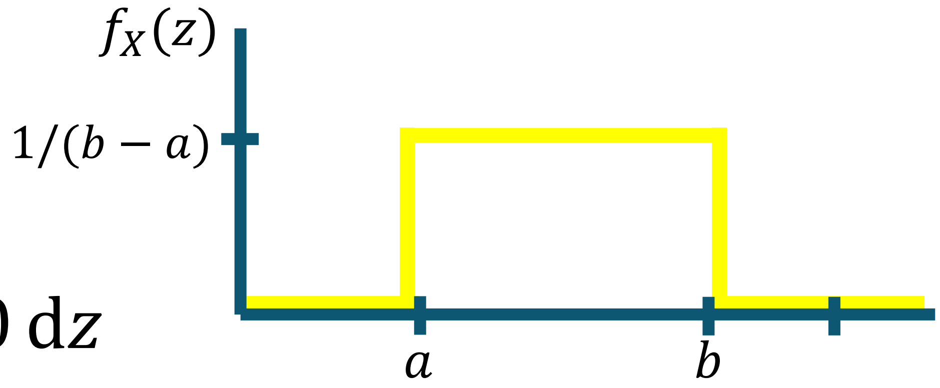
Case when $\mathbf{k} > \boldsymbol{b}$: $F_X(k) = \int_{-\infty}^{k} f_X(z)dz = \int_{\mathrm{a}}^{b} \frac{1}{b-a} dz + \int_{\mathrm{b}}^{k} 0 dz = 1$

$$F_X(k) = \begin{cases} 0 & \text{if } k < a \\ \dfrac{k-a}{b-a} & \text{if } a \le k \le b \\ \mathbf{1} & \textbf{if } \boldsymbol{k \ge b} \end{cases}$$

Continuous Uniform Distribution – **Expectation**

> X is a **uniform random real number** between a and b -> $X \sim \text{Unif}(a, b)$

Using the formula for expectation….



$$\mathbb{E}[X] = \int_{-\infty}^{\infty} z \cdot f_X(z)\, \mathrm{dz}$$

$$= \int_{-\infty}^{a} z \cdot 0\, \mathrm{dz} + \int_{a}^{b} z \cdot \frac{1}{b-a}\, \mathrm{dz} + \int_{b}^{\infty} z \cdot 0\, \mathrm{dz}$$

$$= 0 + \int_{a}^{b} \frac{z}{b-a}\, \mathrm{dz} + 0$$

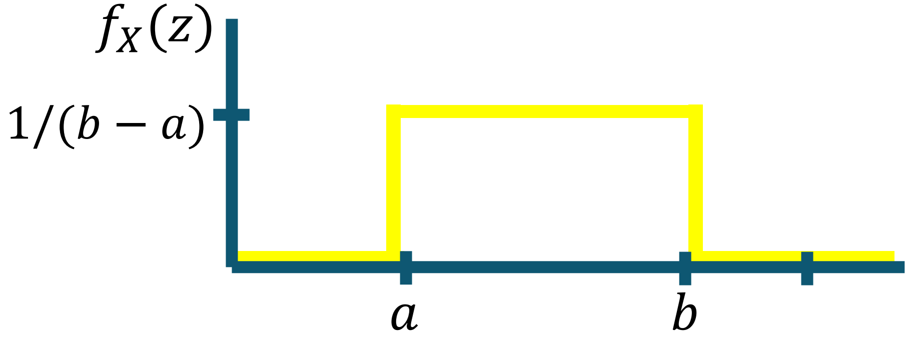
$$= \frac{z^2}{2(b-a)}\bigg|_{z=a}^{b} = \frac{b^2}{2(b-a)} - \frac{a^2}{2(b-a)} = \frac{b^2-a^2}{2(b-a)} = \frac{(b+a)(b-a)}{2(b-a)}$$

$$= \frac{a+b}{2}$$

Continuous Uniform Distribution – **Variance**

> X is a **uniform random real number** between a and b -> $X \sim \mathrm{Unif}(a, b)$

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



Computing $\mathbb{E}[X^2]$

$$\mathbb{E}[X^2] = \int_{-\infty}^{\infty} z^2 f_X(z)\mathrm{d}z = \int_{-\infty}^{a} z^2 \cdot 0 \,\mathrm{d}z + \int_a^b z^2 \cdot \frac{1}{b-a}\mathrm{d}z + \int_b^{\infty} z^2 \cdot 0\,\mathrm{d}z$$

$$= 0 + \int_a^b z^2 \cdot \frac{1}{b-a}\mathrm{d}z + 0$$

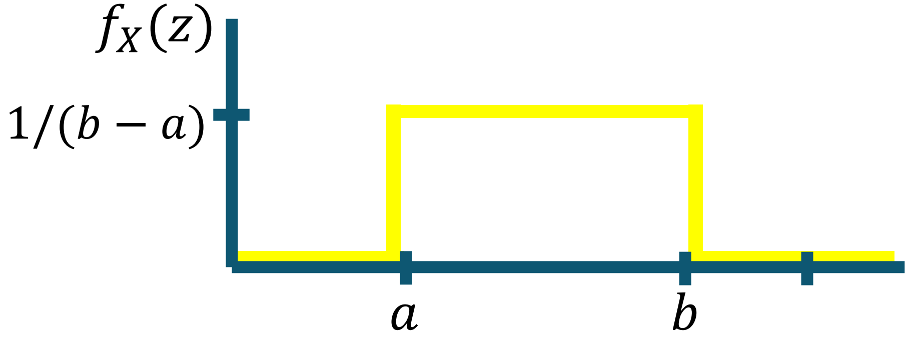
$$= \frac{1}{b-a} \cdot \frac{z^3}{3}\Big|_{z=a}^{b} = \frac{1}{b-a}\left(\frac{b^3}{3} - \frac{a^3}{3}\right) = \frac{1}{3(b-a)} \cdot (b-a)(a^2 + ab + b^2)$$

$$= \frac{a^2+ab+b^2}{3}$$

Continuous Uniform Distribution – **Variance**

> X is a **uniform random real number** between a and b -> $X \sim \mathrm{Unif}(a, b)$

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



$\mathbb{E}[X^2] = \frac{a^2+ab+b^2}{3}$

Plug into the formula:

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

$= \frac{a^2+ab+b^2}{3} - \left(\frac{a+b}{2}\right)^2 = \frac{(b-a)^2}{12}$

Continuous Uniform Distribution

$X \sim \text{Unif}(a, b)$ (uniform **real** number between a and b)

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

CDF: $F_X(k) = \begin{cases} 0 & \text{if } k < a \\ \frac{k-a}{b-a} & \text{if } a \leq k \leq b \\ 1 & \text{if } k \geq b \end{cases}$

Expectation: $\mathbb{E}[X] = \frac{a+b}{2}$

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

Exponential Distribution

How much **time** till an event occurs?

e.g., seconds till thunder, time till the first customer

This sounds very similar to a geometric distribution!
> Geometric random variable is the **number of trials** till success (discrete).
> Exponential random variable is ***time*** (a real number, continuous) till success

With the geometric distribution, we said trials must be independent ->
"If the flip 1 is tails, the coin doesn't remember it was tails, you've made no progress"

Here, *waiting must not make the event happen any sooner ->*
"If we don't get success in the first 3.87sec, chances of seeing success doesn't change"

This means **memorylessness**! $\mathbb{P}(X \geq k + 1 \mid X \geq 1) = \mathbb{P}(Y \geq k)$

Exponential Distribution- **CDF**

$X \sim \text{Exp}(\lambda)$ is time till the first event. Average of λ events per time unit.

It would be hard to come up with the PDF directly here, so start with CDF.

We want $F_X(t) = \mathbb{P}(X \leq t) = 1 - \mathbb{P}(X > t)$

Exponential Distribution- **CDF**

$X \sim \text{Exp}(\lambda)$ is time till the first event. Average of λ events per time unit.

It would be hard to come up with the PDF directly here, so start with CDF.

We want $F_X(t) = \mathbb{P}(X \le t) = 1 - \mathbb{P}(X > t)$

What distribution do we know about the also deals with time?

Exponential Distribution- **CDF**

$X \sim \mathrm{Exp}(\lambda)$ is time till the first event. Average of λ events per time unit.

It would be hard to come up with the PDF directly here, so start with CDF.

We want $F_X(t) = \mathbb{P}(X \leq t) = 1 - \mathbb{P}(X > t)$

*What distribution do we know about the also deals with time…***Poisson!**

Poisson random variable gives us the number of events in a unit of time

Exponential Distribution- **CDF**

$X \sim \mathrm{Exp}(\lambda)$ is time till the first event. Average of λ events per time unit.

It would be hard to come up with the PDF directly here, so start with CDF.

We want $F_X(t) = \mathbb{P}(X \leq t) = 1 - \mathbb{P}(X > t)$

*What distribution do we know about the also deals with time…*Poisson!

Poisson random variable gives us the number of events in a unit of time

What Poisson are we waiting on, and what event for it tells you that $X > t$?

what must be true about the number of successes in a certain time interval?

Exponential Distribution- **CDF**

$X \sim \text{Exp}(\lambda)$ is time till the first event. Average of λ events per time unit.

It would be hard to come up with the PDF directly here, so start with CDF.

We want $F_X(t) = \mathbb{P}(X \leq t) = 1 - \mathbb{P}(X > t)$

*What distribution do we know about the also deals with time…*Poisson!

Poisson random variable gives us the number of events in a unit of time

What Poisson are we waiting on, and what event for it tells you that $X > t$?

there must be 0 events in the first t time units

$Y \sim \text{Poi}(\lambda t)$ (average λt events in t time units).

Then, $\mathbb{P}(\boldsymbol{X} > \boldsymbol{t}) = \mathbb{P}(\boldsymbol{Y} = \mathbf{0})$

Exponential Distribution- **CDF**

$X \sim \mathrm{Exp}(\lambda)$ is time till the first event. Average of λ events per time unit.

We want $F_X(t) = \mathbb{P}(X \leq t) = 1 - \mathbb{P}(X > t)$

Let $Y \sim \mathrm{Poi}(\lambda t)$ (average λt events in t time units).
Then, $\mathbb{P}(\boldsymbol{X > t}) = \mathbb{P}(\boldsymbol{Y = 0})$
"its take more than t time units for first event" = "0 successes in the first t time units"

Putting it all together…

$$\begin{aligned} F_X(t) &= 1 - \mathbb{P}(X > t) \\ &= 1 - \mathbb{P}(Y = 0) \\ &= 1 - e^{-\lambda t} \frac{(\lambda t)^0}{0!} \\ &= 1 - e^{-\lambda t} \end{aligned}$$

$$\boxed{F_X(k) = \begin{cases} 1 - e^{-\lambda k} & \text{if } k \geq 0 \\ 0 & \text{otherwise} \end{cases}}$$

Exponential Distribution- **PDF**

Now we know the CDF: $F_X(k) = \begin{cases} 1 - e^{-\lambda k} & \text{if } k \geq 0 \\ 0 & \text{otherwise} \end{cases}$

What's the PDF (probability density function)?

$f_Y(t) =$

Exponential Distribution- **PDF**

Now we know the CDF: $F_X(k) = \begin{cases} 1 - e^{-\lambda k} & \text{if } k \geq 0 \\ 0 & \text{otherwise} \end{cases}$

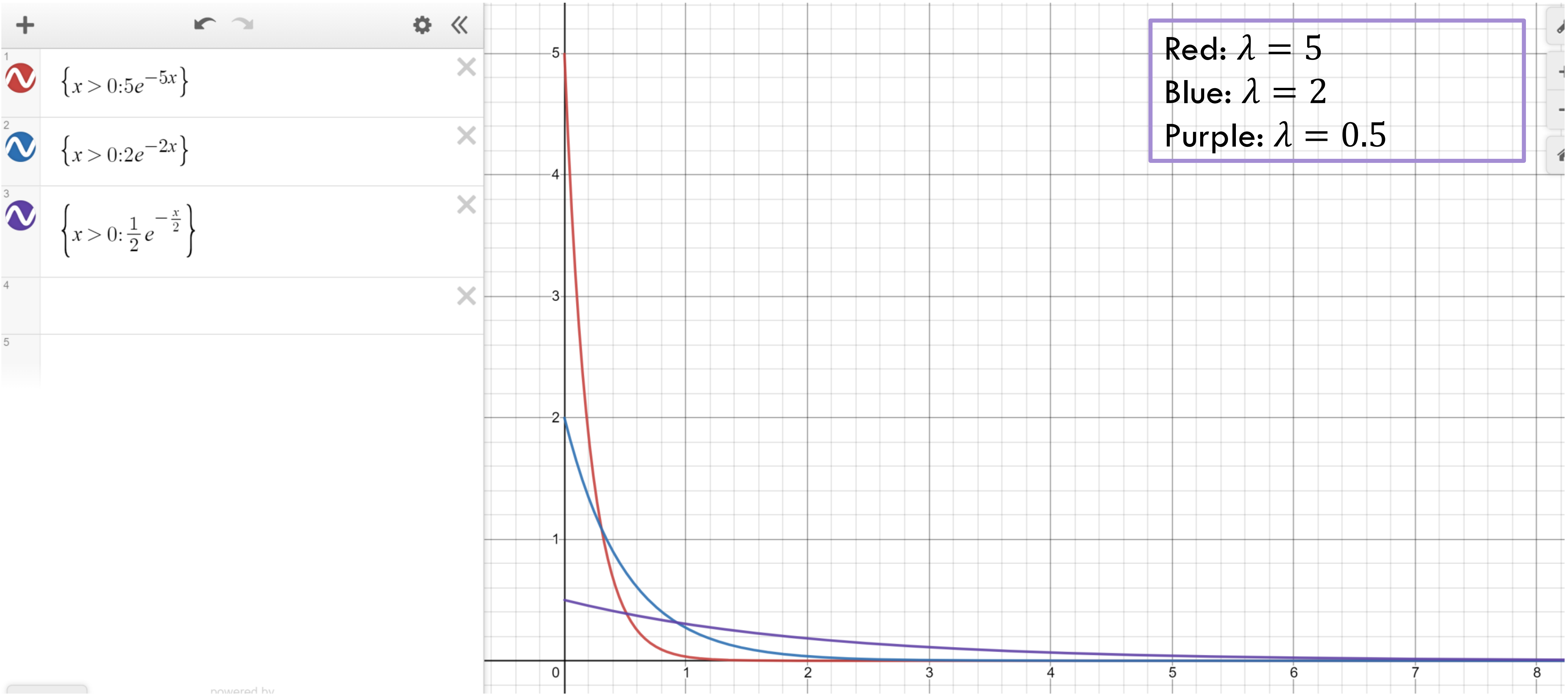
What's the PDF (probability density function)?

$f_Y(t) = \frac{d}{dt}\left(1 - e^{-\lambda t}\right) = 0 - \frac{d}{dt}\left(e^{-\lambda t}\right) = \lambda e^{-\lambda t}.$

For t ≥ 0 it's that expression

For $t < 0$ it's just 0.

Exponential Distribution- **PDF**



Exponential Distribution- **Expectation**

$X \sim \text{Exp}(\lambda)$ is time till the first event. Average of λ events per time unit.

$\mathbb{E}[X] = \int_{-\infty}^{\infty} z \cdot f_X(z)\mathrm{dz}$

$= \int_0^{\infty} z \cdot \lambda e^{-\lambda z}\, dz$

Let $u = z;\ dv = \lambda e^{-\lambda z} dz\ (v = -e^{-\lambda z})$

Don't worry about the derivation (it's here if you're interested; you're not responsible for the derivation. Just the value.

Integrate by parts: $-ze^{-\lambda z} - \int -e^{-\lambda z}\, dz = -ze^{-\lambda z} - \frac{1}{\lambda} e^{-\lambda z}$

Definite Integral: $-ze^{-\lambda z} - \frac{1}{\lambda} e^{-\lambda z} \Big|_{z=0}^{\infty} = (\lim_{z\to\infty} -ze^{-\lambda z} - \frac{1}{\lambda} e^{-\lambda z}) - (0 - \frac{1}{\lambda})$

By L'Hopital's Rule $(\lim_{z\to\infty} -\frac{z}{e^{\lambda z}} - \frac{1}{\lambda e^{\lambda z}}) - (0 - \frac{1}{\lambda}) = \left(\lim_{z\to\infty} -\frac{1}{\lambda e^{\lambda z}}\right) + \frac{1}{\lambda} = \frac{1}{\lambda}$

Exponential Distribution- **Variance**

$X \sim \text{Exp}(\lambda)$ is time till the first event. Average of λ events per time unit.

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

$$= \int_{-\infty}^{\infty} z^2 \cdot f_X(z)\mathrm{d}z - \left(\frac{1}{\lambda}\right)^2$$

$=$ ...after a bunch of calculus

$$= \frac{1}{\lambda^2}$$

Don't worry about the actual calculus here as well ☺

Exponential Distribution

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

Parameter $\lambda \geq 0$ is the average number of events in a unit of time.

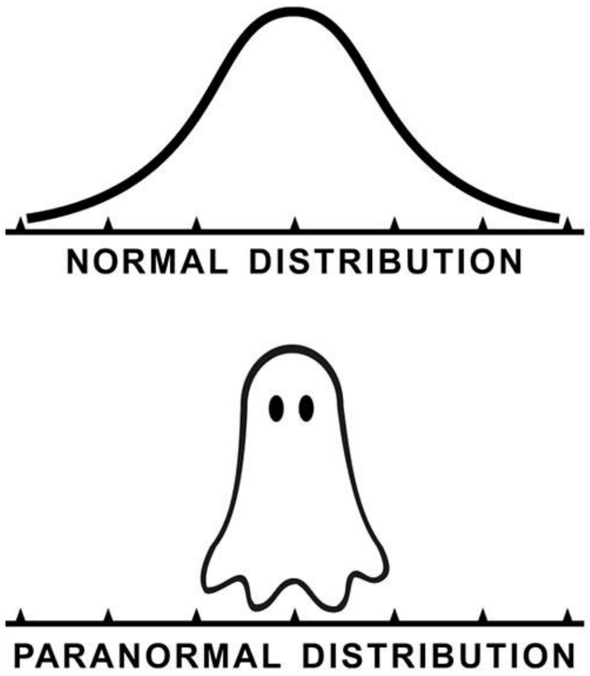
PDF: $f_X(k) = \begin{cases} \lambda e^{-\lambda k} & \text{if } k \geq 0 \\ 0 & \text{otherwise} \end{cases}$

CDF: $F_X(k) = \begin{cases} 1 - e^{-\lambda k} & \text{if } k \geq 0 \\ 0 & \text{otherwise} \end{cases}$

Expectation: $\mathbb{E}[X] = \frac{1}{\lambda}$

Variance: $\mathrm{Var}(X) = \frac{1}{\lambda^2}$

Normal Distribution



Normal Random Variable *(AKA Gaussian)*

There's not a single scenario that follows a normal distribution…
But we're going to see that it shows up in a lot of real world situations!

A **normal random variable** $X \sim \mathcal{N}(\mu, \sigma^2)$ has two parameters:

- $\mu = \mathbb{E}[X]$ is the mean
- $\sigma^2 = \mathrm{Var}(X)$ is the variance ($\sigma = \sqrt{Var(X)}$ is *standard deviation*)

and follows this ***probability density function*** (a bell curve!):

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

Let's take a closer look at that PDF…

$X \sim \mathcal{N}(\mu, \sigma^2)$ iff X follows the following PDF:

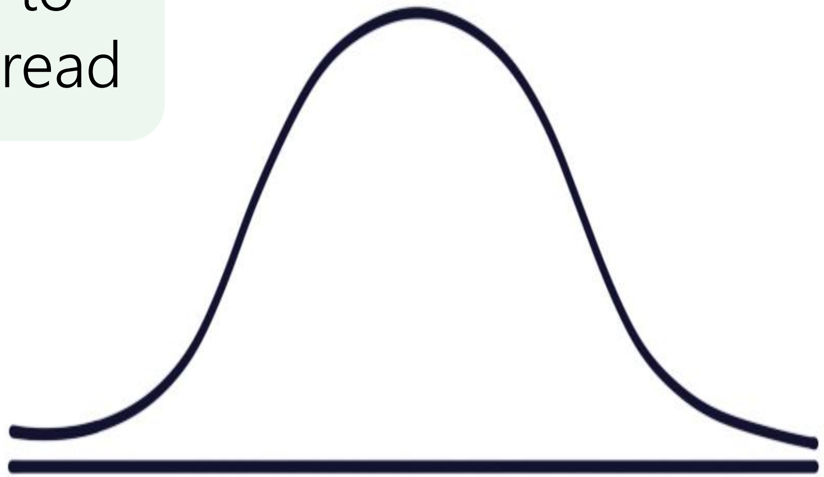
constant for normalization

symmetric around the mean

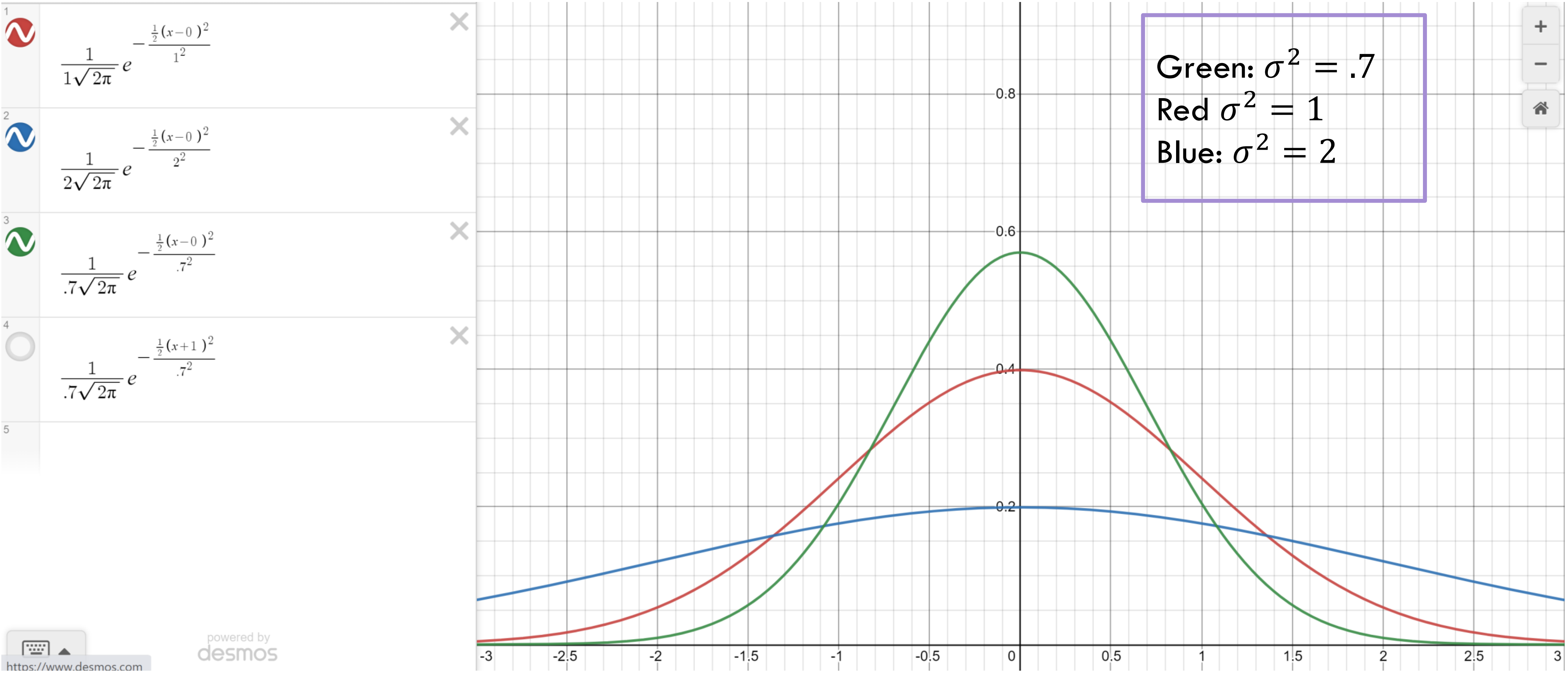
variance to control spread

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

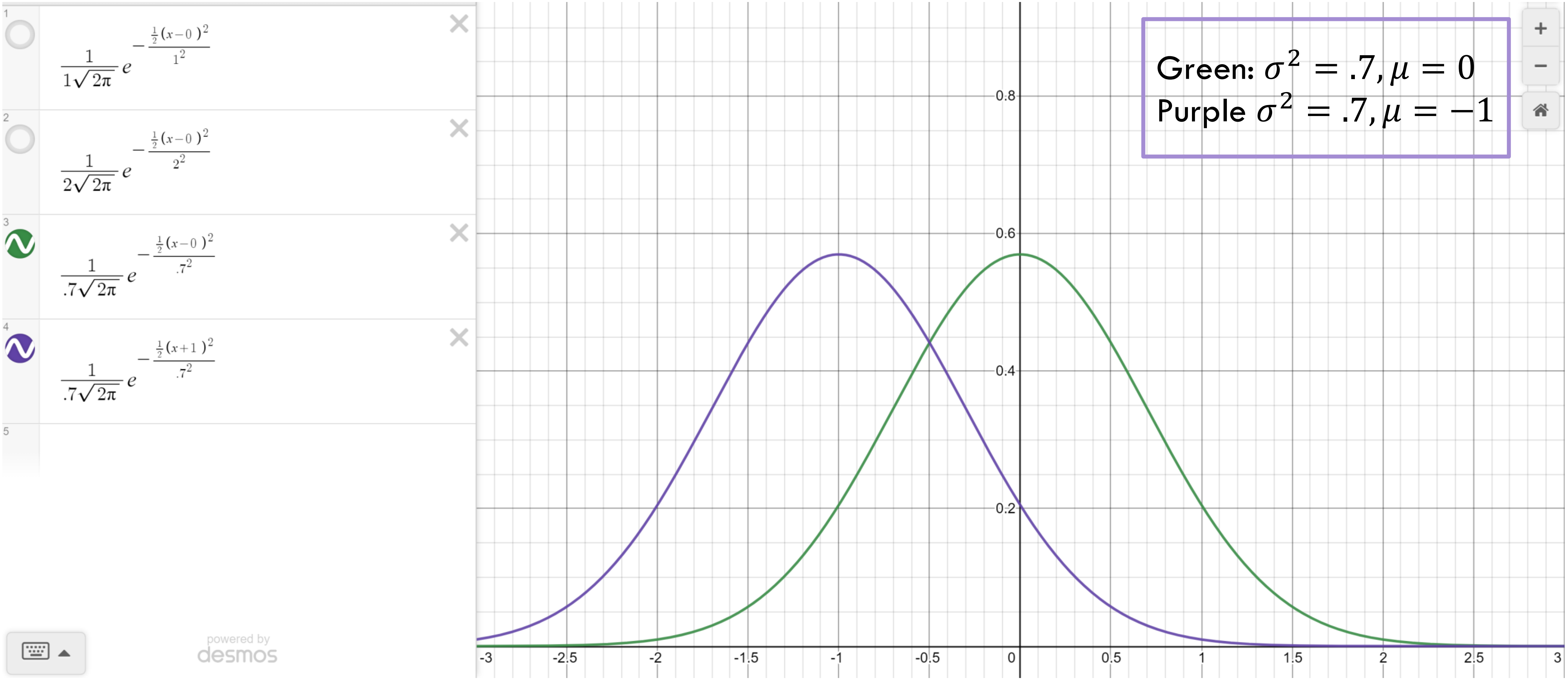
exponential term for tails



Changing the variance



Changing the mean



Closure of Normals *Under Scale and Shift*

When we scale a normal (multiplying by a constant) or shift it (adding a constant) *we get a normal random variable back*!

If $X \sim \mathcal{N}(\mu, \sigma^2)$

Then for $Y = aX + b$, $Y \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$

intuitively: we are just stretching and squishing the distribution – it's symmetric without major disruptions it still follows the same general shape

Normals are unique in that you get a NORMAL back.

If you multiply a binomial by $3/2$ *you don't get a binomial (it's support isn't even integers!)*

Closure of Normals *Under Addition*

When we add two <u>independent</u> normal random variables, you get another normal random variable.

If $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$ and $Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$ and X and Y are independent,

Then, for $Z = aX + bY + c$, $\mathrm{Z} \sim \mathcal{N}(a\mu_X + b\mu_Y + c, a^2\sigma_X^2 + b^2\sigma_Y^2)$

Normals are unique in that you get a NORMAL back.

The sum of two dice rolls (sum of two uniform distributions) does <u>not</u> follow a uniform distribution

Ok…what about the **CDF**?

There is no closed form for the CDF ☹

So how *can* we find the values $F_X(k) = \mathbb{P}(\mathrm{X} \leq \mathrm{k})$?

And for finding the probability of X being in other ranges, we certainly don't want to bother integrating over that PDF…

We have a table with precomputed values!

AKA the "z-table", "phi-table"

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.5279	0.53188	0.53586
0.1	0.53983	0.5438	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61791	0.62172	0.62552	0.6293	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
0.4	0.65542	0.6591	0.66276	0.6664	0.67003	0.67364	0.67724	0.68082	0.68439	0.68793
0.5	0.69146	0.69497	0.69847	0.70194	0.7054	0.70884	0.71226	0.71566	0.71904	0.7224
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.7549
0.7	0.75804	0.76115	0.76424	0.7673	0.77035	0.77337	0.77637	0.77935	0.7823	0.78524
0.8	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83398	0.83646	0.83891
1.0	0.84134	0.84375	0.84614	0.84849	0.85083	0.85314	0.85543	0.85769	0.85993	0.86214
1.1	0.86433	0.8665	0.86864	0.87076	0.87286	0.87493	0.87698	0.879	0.881	0.88298
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	0.89435	0.89617	0.89796	0.89973	0.90147
1.3	0.9032	0.9049	0.90658	0.90824	0.90988	0.91149	0.91309	0.91466	0.91621	0.91774
1.4	0.91924	0.92073	0.9222	0.92364	0.92507	0.92647	0.92785	0.92922	0.93056	0.93189
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943	0.94062	0.94179	0.94295	0.94408
1.6	0.9452	0.9463	0.94738	0.94845	0.9495	0.95053	0.95154	0.95254	0.95352	0.95449
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	0.95994	0.9608	0.96164	0.96246	0.96327
1.8	0.96407	0.96485	0.96562	0.96638	0.96712	0.96784	0.96856	0.96926	0.96995	0.97062
1.9	0.97128	0.97193	0.97257	0.9732	0.97381	0.97441	0.975	0.97558	0.97615	0.9767
2.0	0.97725	0.97778	0.97831	0.97882	0.97932	0.97982	0.9803	0.98077	0.98124	0.98169
2.1	0.98214	0.98257	0.983	0.98341	0.98382	0.98422	0.98461	0.985	0.98537	0.98574
2.2	0.9861	0.98645	0.98679	0.98713	0.98745	0.98778	0.98809	0.9884	0.9887	0.98899
2.3	0.98928	0.98956	0.98983	0.9901	0.99036	0.99061	0.99086	0.99111	0.99134	0.99158
2.4	0.9918	0.99202	0.99224	0.99245	0.99266	0.99286	0.99305	0.99324	0.99343	0.99361
2.5	0.99379	0.99396	0.99413	0.9943	0.99446	0.99461	0.99477	0.99492	0.99506	0.9952
2.6	0.99534	0.99547	0.9956	0.99573	0.99585	0.99598	0.99609	0.99621	0.99632	0.99643
2.7	0.99653	0.99664	0.99674	0.99683	0.99693	0.99702	0.99711	0.9972	0.99728	0.99736
2.8	0.99744	0.99752	0.9976	0.99767	0.99774	0.99781	0.99788	0.99795	0.99801	0.99807
2.9	0.99813	0.99819	0.99825	0.99831	0.99836	0.99841	0.99846	0.99851	0.99856	0.99861
3.0	0.99865	0.99869	0.99874	0.99878	0.99882	0.99886	0.99889	0.99893	0.99896	0.999

We have a table containing values for the CDF of the **standard normal random variable** $\boldsymbol{Z} \sim \mathcal{N}(0,1)$

Yes, we're going to use a table in 2024 🤩
Mainly for consistency in this class.

(In the real world, we have programming libraries like Python's scipy: stats.norm.cdf)

We have a table with precomputed values!

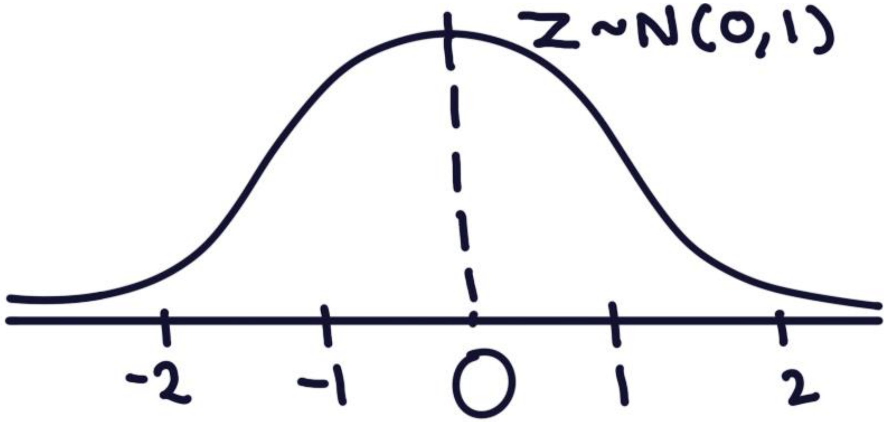
AKA the "z-table", "phi-table"

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.5279	0.53188	0.53586
0.1	0.53983	0.5438	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61791	0.62172	0.62552	0.6293	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
0.4	0.65542	0.6591	0.66276	0.6664	0.67003	0.67364	0.67724	0.68082	0.68439	0.68793
0.5	0.69146	0.69497	0.69847	0.70194	0.7054	0.70884	0.71226	0.71566	0.71904	0.7224
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.7549
0.7	0.75804	0.76115	0.76424	0.7673	0.77035	0.77337	0.77637	0.77935	0.7823	0.78524
0.8	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83398	0.83646	0.83891
1.0	0.84134	0.84375	0.84614	0.84849	0.85083	0.85314	0.85543	0.85769	0.85993	0.86214
1.1	0.86433	0.8665	0.86864	0.87076	0.87286	0.87493	0.87698	0.879	0.881	0.88298
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	0.89435	0.89617	0.89796	0.89973	0.90147
1.3	0.9032	0.9049	0.90658	0.90824	0.90988	0.91149	0.91309	0.91466	0.91621	0.91774
1.4	0.91924	0.92073	0.9222	0.92364	0.92507	0.92647	0.92785	0.92922	0.93056	0.93189
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943	0.94062	0.94179	0.94295	0.94408
1.6	0.9452	0.9463	0.94738	0.94845	0.9495	0.95053	0.95154	0.95254	0.95352	0.95449
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	0.95994	0.9608	0.96164	0.96246	0.96327
1.8	0.96407	0.96485	0.96562	0.96638	0.96712	0.96784	0.96856	0.96926	0.96995	0.97062
1.9	0.97128	0.97193	0.97257	0.9732	0.97381	0.97441	0.975	0.97558	0.97615	0.9767
2.0	0.97725	0.97778	0.97831	0.97882	0.97932	0.97982	0.9803	0.98077	0.98124	0.98169
2.1	0.98214	0.98257	0.983	0.98341	0.98382	0.98422	0.98461	0.985	0.98537	0.98574
2.2	0.9861	0.98645	0.98679	0.98713	0.98745	0.98778	0.98809	0.9884	0.9887	0.98899
2.3	0.98928	0.98956	0.98983	0.9901	0.99036	0.99061	0.99086	0.99111	0.99134	0.99158
2.4	0.9918	0.99202	0.99224	0.99245	0.99266	0.99286	0.99305	0.99324	0.99343	0.99361
2.5	0.99379	0.99396	0.99413	0.9943	0.99446	0.99461	0.99477	0.99492	0.99506	0.9952
2.6	0.99534	0.99547	0.9956	0.99573	0.99585	0.99598	0.99609	0.99621	0.99632	0.99643
2.7	0.99653	0.99664	0.99674	0.99683	0.99693	0.99702	0.99711	0.9972	0.99728	0.99736
2.8	0.99744	0.99752	0.9976	0.99767	0.99774	0.99781	0.99788	0.99795	0.99801	0.99807
2.9	0.99813	0.99819	0.99825	0.99831	0.99836	0.99841	0.99846	0.99851	0.99856	0.99861
3.0	0.99865	0.99869	0.99874	0.99878	0.99882	0.99886	0.99889	0.99893	0.99896	0.999

We have a table containing values for the CDF of the **standard normal random variable $\boldsymbol{Z} \sim \mathcal{N}(0,1)$**

> Φ is a function for CDF of $\mathcal{N}(0,1)$

> $\Phi(\mathrm{z}) = \mathrm{F_Z}(\mathrm{z}) = \mathbb{P}(\mathrm{Z} \leq z)$



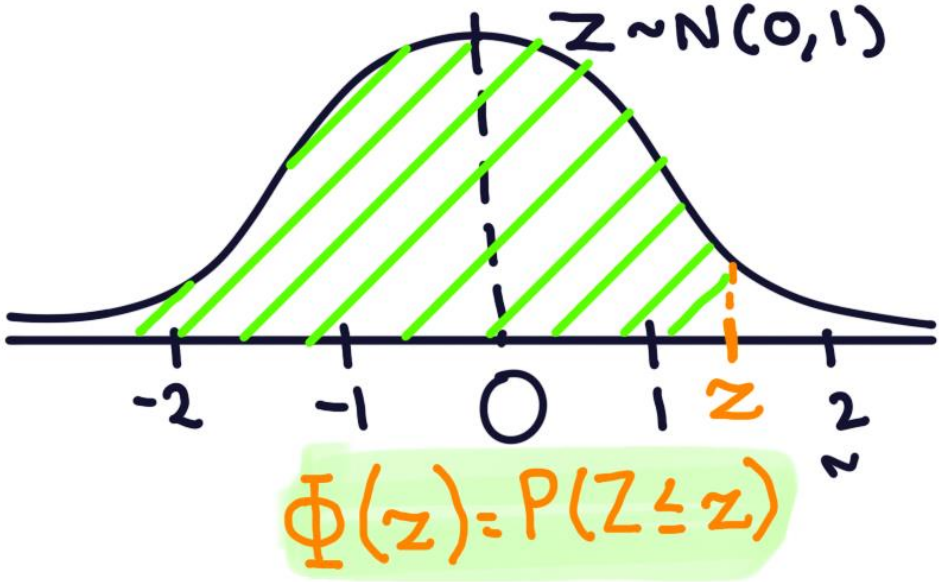
We have a table with precomputed values!

AKA the "z-table", "phi-table"

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.5279	0.53188	0.53586
0.1	0.53983	0.5438	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61791	0.62172	0.62552	0.6293	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
0.4	0.65542	0.6591	0.66276	0.6664	0.67003	0.67364	0.67724	0.68082	0.68439	0.68793
0.5	0.69146	0.69497	0.69847	0.70194	0.7054	0.70884	0.71226	0.71566	0.71904	0.7224
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.7549
0.7	0.75804	0.76115	0.76424	0.7673	0.77035	0.77337	0.77637	0.77935	0.7823	0.78524
0.8	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83398	0.83646	0.83891
1.0	0.84134	0.84375	0.84614	0.84849	0.85083	0.85314	0.85543	0.85769	0.85993	0.86214
1.1	0.86433	0.8665	0.86864	0.87076	0.87286	0.87493	0.87698	0.879	0.881	0.88298
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	0.89435	0.89617	0.89796	0.89973	0.90147
1.3	0.9032	0.9049	0.90658	0.90824	0.90988	0.91149	0.91309	0.91466	0.91621	0.91774
1.4	0.91924	0.92073	0.9222	0.92364	0.92507	0.92647	0.92785	0.92922	0.93056	0.93189
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943	0.94062	0.94179	0.94295	0.94408
1.6	0.9452	0.9463	0.94738	0.94845	0.9495	0.95053	0.95154	0.95254	0.95352	0.95449
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	0.95994	0.9608	0.96164	0.96246	0.96327
1.8	0.96407	0.96485	0.96562	0.96638	0.96712	0.96784	0.96856	0.96926	0.96995	0.97062
1.9	0.97128	0.97193	0.97257	0.9732	0.97381	0.97441	0.975	0.97558	0.97615	0.9767
2.0	0.97725	0.97778	0.97831	0.97882	0.97932	0.97982	0.9803	0.98077	0.98124	0.98169
2.1	0.98214	0.98257	0.983	0.98341	0.98382	0.98422	0.98461	0.985	0.98537	0.98574
2.2	0.9861	0.98645	0.98679	0.98713	0.98745	0.98778	0.98809	0.9884	0.9887	0.98899
2.3	0.98928	0.98956	0.98983	0.9901	0.99036	0.99061	0.99086	0.99111	0.99134	0.99158
2.4	0.9918	0.99202	0.99224	0.99245	0.99266	0.99286	0.99305	0.99324	0.99343	0.99361
2.5	0.99379	0.99396	0.99413	0.9943	0.99446	0.99461	0.99477	0.99492	0.99506	0.9952
2.6	0.99534	0.99547	0.9956	0.99573	0.99585	0.99598	0.99609	0.99621	0.99632	0.99643
2.7	0.99653	0.99664	0.99674	0.99683	0.99693	0.99702	0.99711	0.9972	0.99728	0.99736
2.8	0.99744	0.99752	0.9976	0.99767	0.99774	0.99781	0.99788	0.99795	0.99801	0.99807
2.9	0.99813	0.99819	0.99825	0.99831	0.99836	0.99841	0.99846	0.99851	0.99856	0.99861
3.0	0.99865	0.99869	0.99874	0.99878	0.99882	0.99886	0.99889	0.99893	0.99896	0.999

We have a table containing values for the CDF of the **standard normal random variable $\boldsymbol{Z}\sim\mathcal{N}(0,1)$**

> Φ is a function for CDF of $\mathcal{N}(0,1)$
> $\Phi(\mathrm{z}) = \mathrm{F_Z}(\mathrm{z}) = \mathbb{P}(\mathrm{Z} \le z)$



But how to go from $\mathcal{N}(\mu, \sigma^2)$ to $\mathcal{N}(0,1)$?

We have a table for the values of $\mathcal{N}(0,1)$. How to use this for $\mathcal{N}(\mu, \sigma^2)$?

We will **standardize** X! If we have $\mathrm{X} \sim \mathcal{N}(\mu, \sigma^2)$

1. **Subtract $\boldsymbol{\mu}$ to shift the distribution to have mean of 0**

   $\mathbb{E}[X - \mu] = \mathbb{E}[X] - \mu = \mu - \mu = 0$

2. **Divide by $\boldsymbol{\sigma}$ to squish/stretch the distribution to have variance of 1**

   $Var\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma^2} Var(X - \mu) = \frac{1}{\sigma^2} Var(X) = \frac{1}{\sigma^2} \sigma^2 = 1$

$Z = \frac{X-\mu}{\sigma}$ is a <u>standard normal random variable</u>: $Z \sim \mathcal{N}(0,1)$

Computing Probabilities of Normal RVs

1. **Write the probability we're interested in in terms of the CDF**
2. **Standardize** the normal random variable: $Z = \frac{X-\mu}{\sigma}$
2. **Round** the "z-score"(s) to the hundredths place.
3. **Look up the value(s)** in the table

Practice!

We use $\Phi(z)$ to mean $F_Z(z)$ where $Z \sim \mathcal{N}(0,1)$.

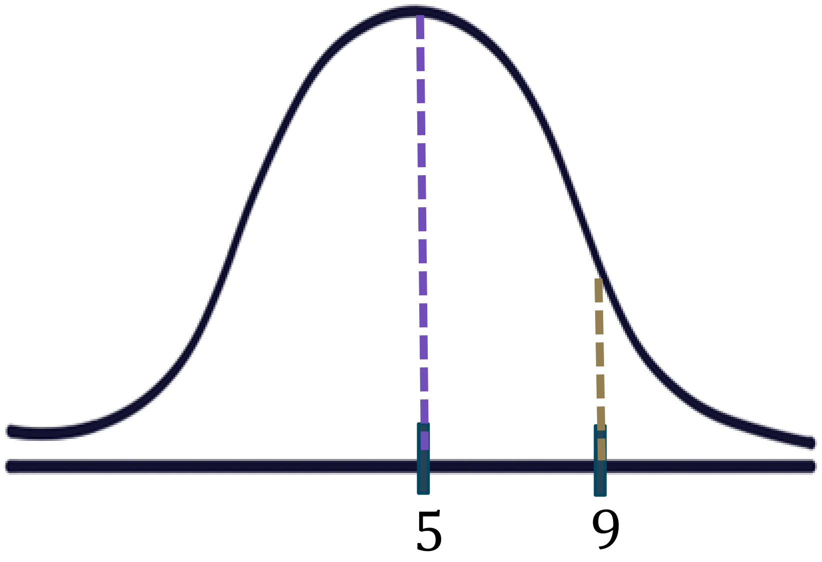
Let $X \sim \mathcal{N}(5,4)$. What is $\mathbb{P}(X \leq 9)$?

$\mathbb{P}(X \leq 9)$

$= \mathbb{P}\left(\frac{Y-5}{2} \leq \frac{9-5}{2}\right)$ **standardize** (algebra on both sides)

$= \mathbb{P}(Z \leq \frac{9-5}{2})$ where $Z \sim N(0,1)$.

$= \mathbb{P}\left(Z \leq \frac{9-5}{2}\right) = \Phi(2.00) = 0.97725$



Practice!

We use $\Phi(z)$ to mean $F_Z(z)$ where $Z \sim \mathcal{N}(0,1)$.

Let $X \sim \mathcal{N}(5,4)$. What is $\mathbb{P}(X > 9)$?

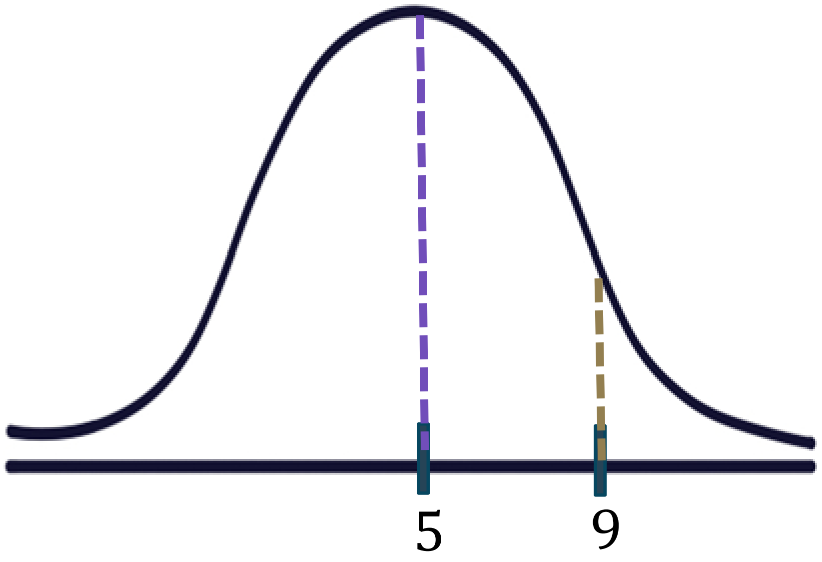
$\mathbb{P}(X > 9) = 1 - \mathbb{P}(X \leq 9)$

$= 1 - \mathbb{P}\left(\frac{Y-5}{2} \leq \frac{9-5}{2}\right)$ **standardize** (algebra on both sides)

$= 1 - \mathbb{P}(Z \leq \frac{9-5}{2})$ where $Z \sim N(0,1)$.

$= 1 - \mathbb{P}\left(Z \leq \frac{9-5}{2}\right) = 1 - \Phi(2.00)$

$= 1 - 0.97725 = 0.02275$



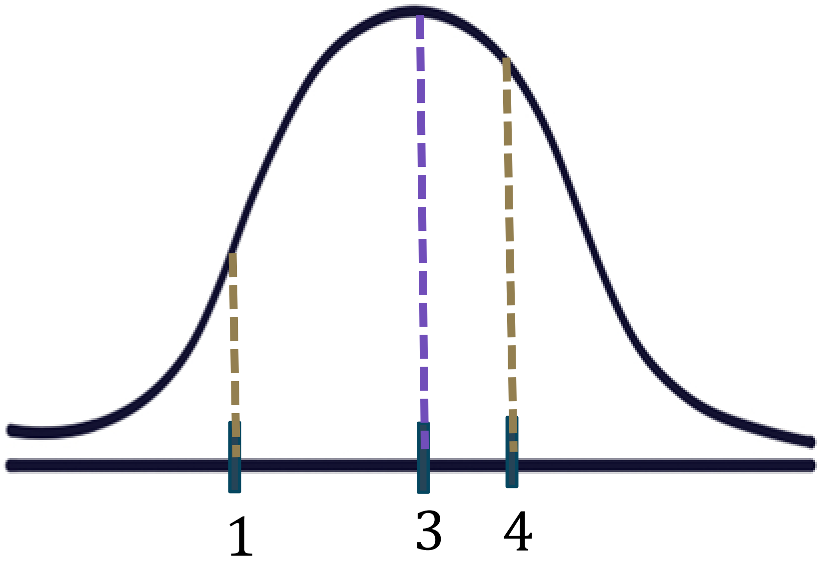
z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.5279	0.53188	0.53586
0.1	0.53983	0.5438	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61791	0.62172	0.62552	0.6293	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
0.4	0.65542	0.6591	0.66276	0.6664	0.67003	0.67364	0.67724	0.68082	0.68439	0.68793
0.5	0.69146	0.69497	0.69847	0.70194	0.7054	0.70884	0.71226	0.71566	0.71904	0.7224
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.7549
0.7	0.75804	0.76115	0.76424	0.7673	0.77035	0.77337	0.77637	0.77935	0.7823	0.78524
0.8	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83398	0.83646	0.83891
1.0	0.84134	0.84375	0.84614	0.84849	0.85083	0.85314	0.85543	0.85769	0.85993	0.86214
1.1	0.86433	0.8665	0.86864	0.87076	0.87286	0.87493	0.87698	0.879	0.881	0.88298
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	0.89435	0.89617	0.89796	0.89973	0.90147
1.3	0.9032	0.9049	0.90658	0.90824	0.90988	0.91149	0.91309	0.91466	0.91621	0.91774
1.4	0.91924	0.92073	0.9222	0.92364	0.92507	0.92647	0.92785	0.92922	0.93056	0.93189
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943	0.94062	0.94179	0.94295	0.94408
1.6	0.9452	0.9463	0.94738	0.94845	0.9495	0.95053	0.95154	0.95254	0.95352	0.95449
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	0.95994	0.9608	0.96164	0.96246	0.96327
1.8	0.96407	0.96485	0.96562	0.96638	0.96712	0.96784	0.96856	0.96926	0.96995	0.97062
1.9	0.97128	0.97193	0.97257	0.9732	0.97381	0.97441	0.975	0.97558	0.97615	0.9767
2.0	0.97725	0.97778	0.97831	0.97882	0.97932	0.97982	0.9803	0.98077	0.98124	0.98169
2.1	0.98214	0.98257	0.983	0.98341	0.98382	0.98422	0.98461	0.985	0.98537	0.98574
2.2	0.9861	0.98645	0.98679	0.98713	0.98745	0.98778	0.98809	0.9884	0.9887	0.98899
2.3	0.98928	0.98956	0.98983	0.9901	0.99036	0.99061	0.99086	0.99111	0.99134	0.99158
2.4	0.9918	0.99202	0.99224	0.99245	0.99266	0.99286	0.99305	0.99324	0.99343	0.99361
2.5	0.99379	0.99396	0.99413	0.9943	0.99446	0.99461	0.99477	0.99492	0.99506	0.9952
2.6	0.99534	0.99547	0.9956	0.99573	0.99585	0.99598	0.99609	0.99621	0.99632	0.99643
2.7	0.99653	0.99664	0.99674	0.99683	0.99693	0.99702	0.99711	0.9972	0.99728	0.99736
2.8	0.99744	0.99752	0.9976	0.99767	0.99774	0.99781	0.99788	0.99795	0.99801	0.99807
2.9	0.99813	0.99819	0.99825	0.99831	0.99836	0.99841	0.99846	0.99851	0.99856	0.99861
3.0	0.99865	0.99869	0.99874	0.99878	0.99882	0.99886	0.99889	0.99893	0.99896	0.999

More practice

Let $X \sim \mathcal{N}(3, 2)$.

What is the probability that $1 \leq X \leq 4$?

$\mathbb{P}(1 \leq X \leq 4)$



More practice

Let $X \sim \mathcal{N}(3, 2)$.

What is the probability that $1 \le X \le 4$?

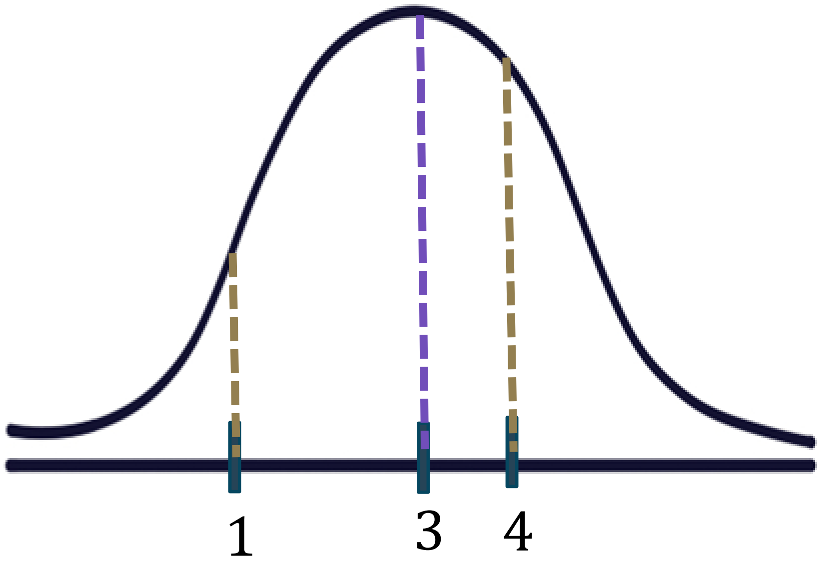
$\mathbb{P}(1 \le X \le 4)$

$= \mathbb{P}(X \le 4) - \mathbb{P}(X \le 1)$

$= \mathbb{P}\left(\frac{X-3}{\sqrt{2}} \le \frac{4-3}{\sqrt{2}}\right) - \mathbb{P}\left(\frac{X-3}{\sqrt{2}} \le \frac{1-3}{\sqrt{2}}\right) =$

$= \mathbb{P}\left(Z \le \frac{4-3}{\sqrt{2}}\right) - \mathbb{P}\left(Z \le \frac{1-3}{\sqrt{2}}\right)$ where $Z \sim \mathcal{N}(0,1)$

$= \Phi(.71) - \Phi(-1.41)$

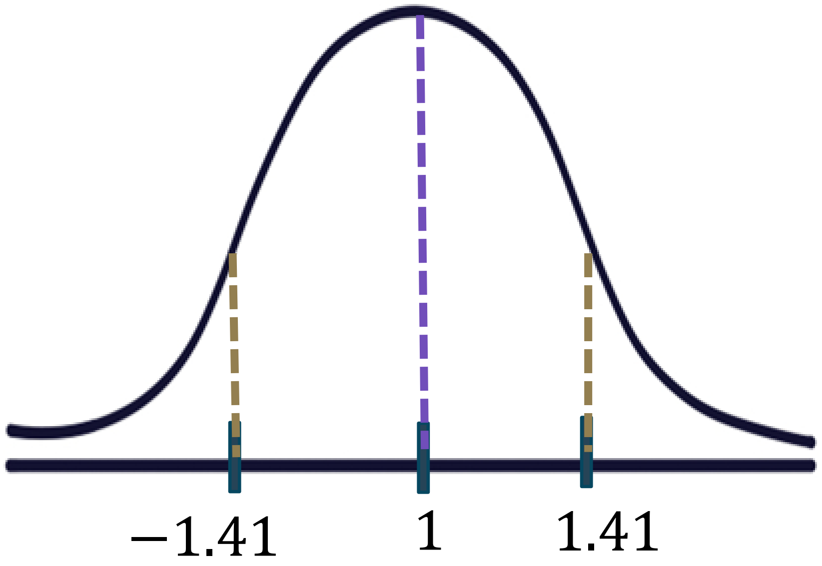


More practice

How do we find $\Phi(-1.41) = \mathbb{P}(Z \leq -1.41)$? Our table only has CDF values for positive numbers! 😱

Recall: the normal distribution is <u>symmetric</u>

$\mathbb{P}(Z \leq -1.41) =$



More practice

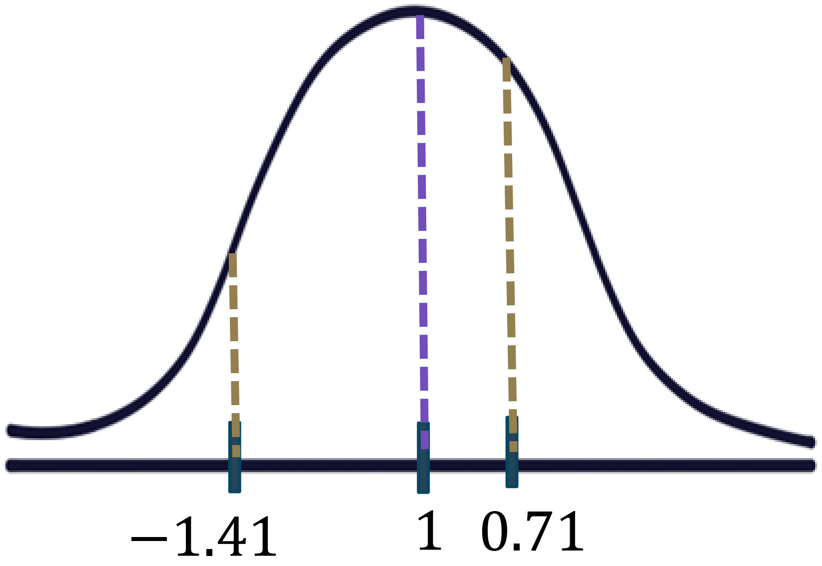
How do we find $\Phi(-1.41) = \mathbb{P}(Z \leq -1.41)$? Our table only has CDF values for positive numbers! 😱

Recall: the normal distribution is <u>symmetric</u>

$\mathbb{P}(Z \leq -1.41) = \mathbb{P}(Z \geq 1.41)$

$= 1 - \mathbb{P}(Z \leq 1.41)$

$= 1 - \Phi(1.41)$ *this is something we can do…*



More practice

Let $X \sim \mathcal{N}(3, 2)$.

What is the probability that $1 \le X \le 4$?

$\mathbb{P}(1 \le X \le 4)$

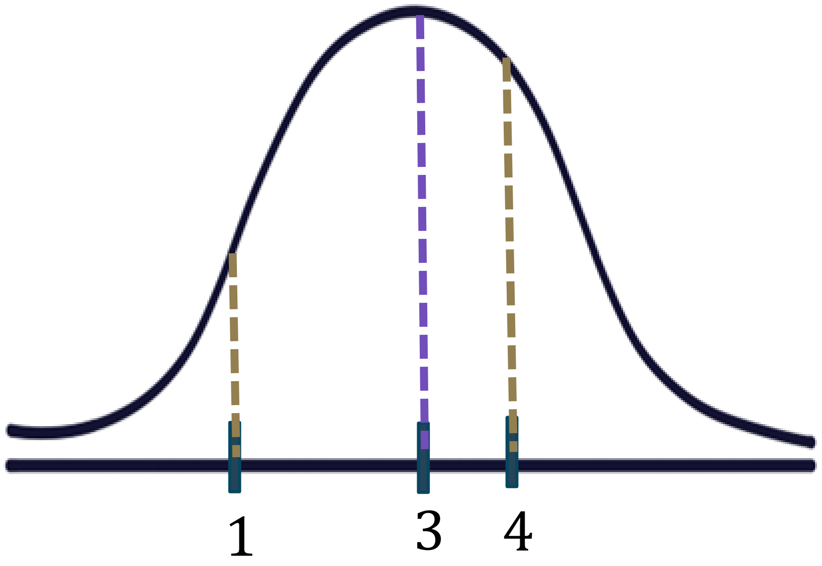
$= \mathbb{P}(X \le 4) - \mathbb{P}(X \le 1)$

$= \mathbb{P}\left(\frac{X-3}{\sqrt{2}} \le \frac{4-3}{\sqrt{2}}\right) - \mathbb{P}\left(\frac{X-3}{\sqrt{2}} \le \frac{1-3}{\sqrt{2}}\right) =$

$= \mathbb{P}\left(Z \le \frac{4-3}{\sqrt{2}}\right) - \mathbb{P}\left(Z \le \frac{1-3}{\sqrt{2}}\right)$ where $Z \sim \mathcal{N}(0,1)$

$= \Phi(.71) - \Phi(-1.41) = \Phi(.71) - (1 - \Phi(1.41))$

$= 0.76115 - (1 - 0.92073) = \mathbf{0.68188}$



“within __ standard deviations from the mean”

What’s the probability of being within two standard deviations from the mean?

$$\mathbb{P}(\mu - 2\sigma \leq X \leq \mu + 2\sigma)$$

$$= \mathbb{P}(X \leq \mu + 2\sigma) - \mathbb{P}(X \leq \mu - 2\sigma) = \mathbb{P}\left(Z \leq \frac{\mu+2\sigma-\mu}{\sigma}\right) - \mathbb{P}\left(Z \leq \frac{\mu-2\sigma-\mu}{\sigma}\right)$$

$$= \Phi(2) - \Phi(-2)$$

$$= \Phi(2) - \big(1 - \Phi(2)\big)$$

$$= .97725 - (1 - .97725) = .9545$$

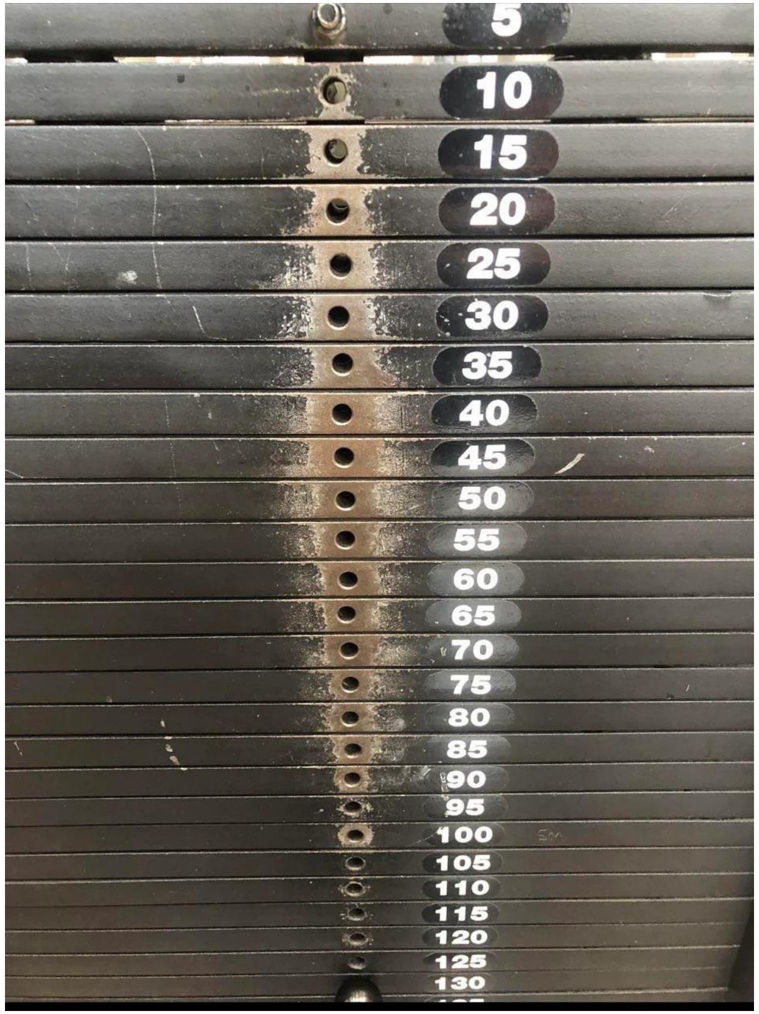
You’ll sometimes hear statisticians refer to the “68-95-99.7 rule” which is the probability of being within 1,2, or 3 standard deviations of the mean.

Normal (aka Guassian)

When finding the probability of a normal random variable, draw a picture!! It can help reasoning about how we can use the CDF and the z-table to compute the desired region.

1. **Write the probability we're interested in in terms of the CDF**
2. **Standardize** the normal random variable: $Z = \frac{X-\mu}{\sigma}$
2. **Round** the "z-score"(s) to the hundredths place.
3. **Look up the value(s)** in the table

Can you spot the normal distribution?



Picture from reddit

It turns out the normal distribution appear a LOT in the real world. Like…in the gym!

On Monday, we will talk about how and why!