Zoo of Continuous Random Variables

CSE 312 26Su

Lecture 14

Where Are We?

Last time:

- Continuous Random Variables
  - Probability Density Function
  - Cumulative Distribution Function
  - Expectation and Variance
- Comparing Discrete vs. Continuous
- Continuous Uniform Distribution

Today:

- Zoo of continuous random variables 🐻
  - 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

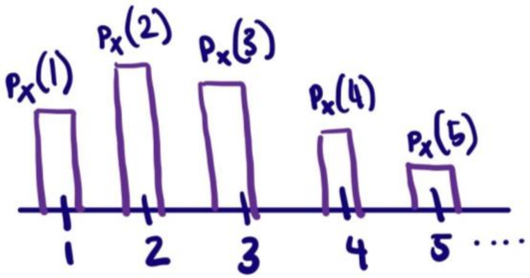
Support is uncountably infinite (e.g., *real numbers*)

$\mathbb{P}(\boldsymbol{X} = \boldsymbol{k}) = \frac{\mathbf{1}}{\infty} = \mathbf{0}$ 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



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

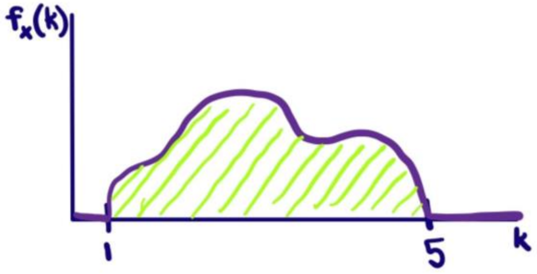
$$0 \le p_X(k) \le 1$$

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



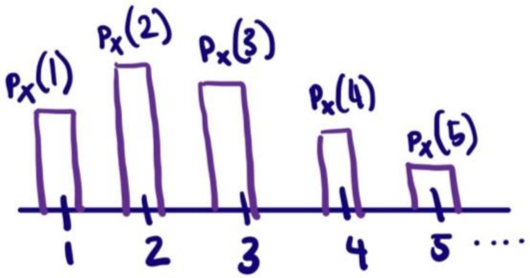
$$\int_{-\infty}^{\infty} f_X(k)\ \mathrm{d}k = 1$$

$$f_X(k) \ge 0$$

Discrete RVs

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

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



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

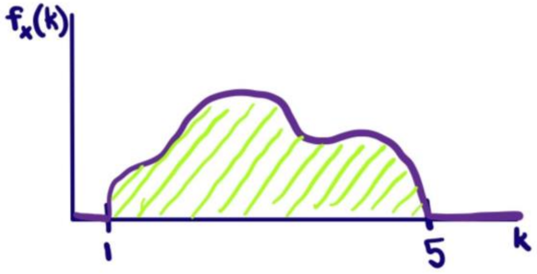
$$0 \leq p_X(k) \leq 1$$

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



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

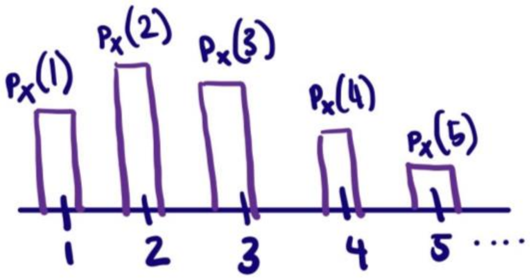
$$f_X(k) \geq 0$$

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

Discrete RVs

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

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



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

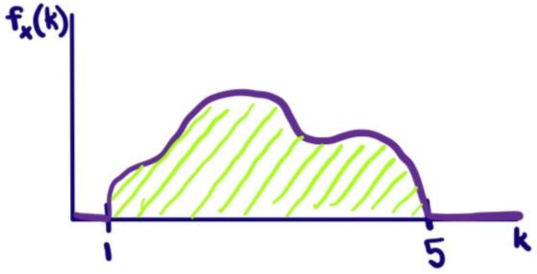
$$0 \leq p_X(k) \leq 1$$

Continuous RVs

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

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

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



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

$$f_X(k) \geq 0$$

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

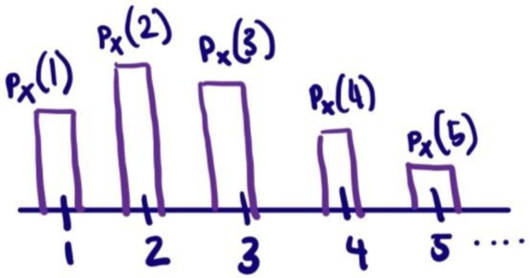
Sum up the probabilities of values $\leq k$

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



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

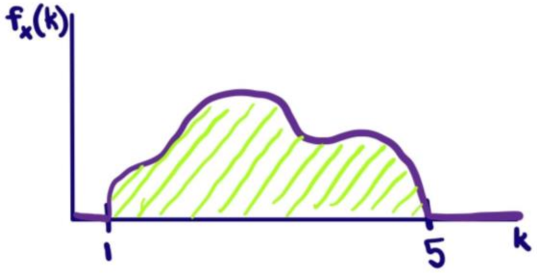
$$0 \leq p_X(k) \leq 1$$

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

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

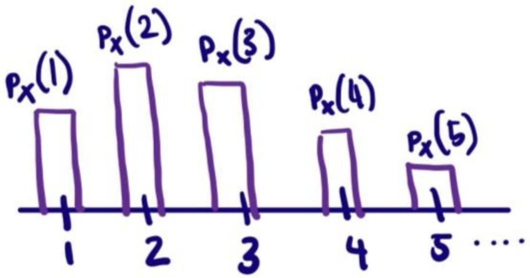
Expectation: $\mathbb{E}[X] = \int_{-\infty}^{\infty} z \cdot f_X(z)\ \mathrm{dz}$

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

Discrete RVs

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

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



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

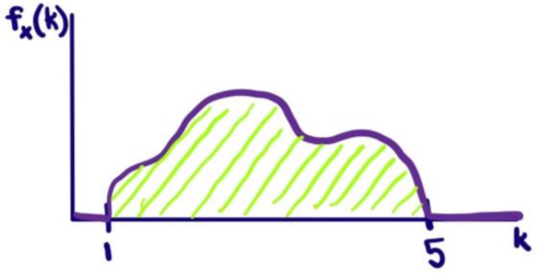
$$0 \le p_X(k) \le 1$$

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



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

$$f_X(k) \ge 0$$

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

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

<u>Integrate</u> 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: $\mathbb{E}[X] = \int_{-\infty}^{\infty} z \cdot f_X(z) \, \mathrm{d}z$

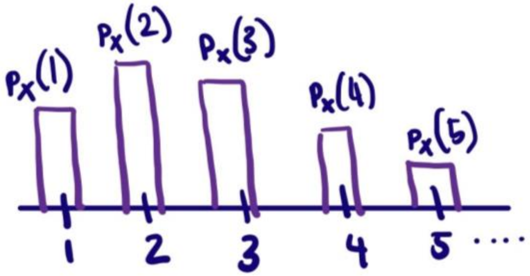
$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(z) \cdot f_X(z) \, \mathrm{d}z$

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

Discrete RVs

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

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



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

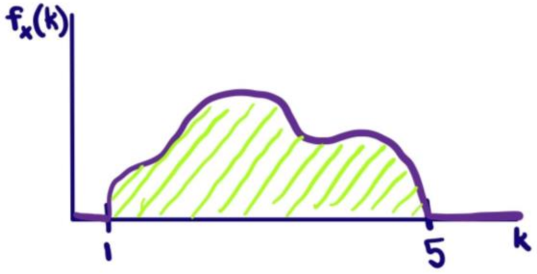
$$0 \leq p_X(k) \leq 1$$

Continuous RVs

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

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

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



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

$$f_X(k) \geq 0$$

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

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!

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

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

Continuous Uniform Distribution

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

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

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

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

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

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

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**

$X \sim \text{Exp}(\lambda)$ 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 $X > t$?

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

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 $X > t$?

there must be 0 events in the first t time units

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

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

Exponential Distribution- **CDF**

$X \sim \text{Exp}(\lambda)$ 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 $Y \sim \text{Poi}(\lambda t)$ (average λt events in t time units).
Then, $\mathbb{P}(\boldsymbol{X} > \boldsymbol{t}) = \mathbb{P}(\boldsymbol{Y} = \mathbf{0})$
"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: $F_X(k) = \begin{cases} 1 - e^{-\lambda k} & \text{if } k \geq 0 \\ 0 & \text{otherwise} \end{cases}$

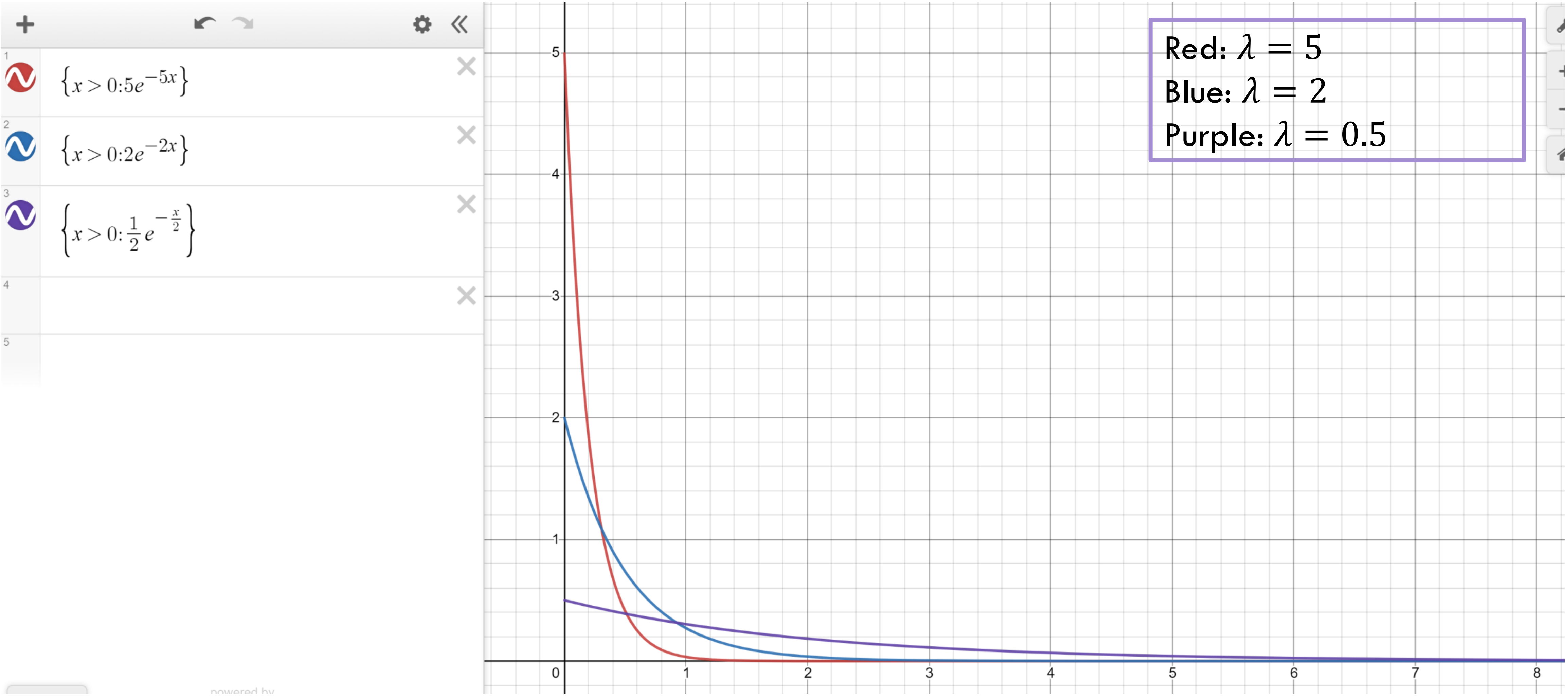
What's the PDF (probability density function)?

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

For t ≥ 0 it's that expression

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

Exponential Distribution- **PDF**



Memorylessness

$$\mathbb{P}(X \geq k+1 | X \geq 1) = \frac{\mathbb{P}(X \geq k+1 \cap X \geq 1)}{\mathbb{P}(X \geq 1)} = \frac{\mathbb{P}(X \geq k+1)}{1-(1-e^{-\lambda \cdot 1})}$$

$$= \frac{e^{-\lambda(k+1)}}{e^{-\lambda}} = e^{-\lambda k}$$

What about $\mathbb{P}(X \geq k)$ (without conditioning on the first step)?

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

It's the same!!!

More generally, for an exponential rv X, $\mathbb{P}(X \geq s+t | X \geq s) = \mathbb{P}(X \geq t)$

Side note

I hid a trick in that algebra,

$\mathbb{P}(X \geq 1) = 1 - \mathbb{P}(X < 1) = 1 - \mathbb{P}(X \leq 1)$

The first step is the complementary law.

The second step is using that $\int_1^1 f_X(z)\mathrm{d}z = 0$

In general, for continuous random variables we can switch out $\leq$ and $<$ without anything changing.

We can't make those switches for discrete random variables.

Exponential Distribution- **Expectation**

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

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

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

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.

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

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

$=$ …after a bunch of calculus

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

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

Exponential Distribution

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

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

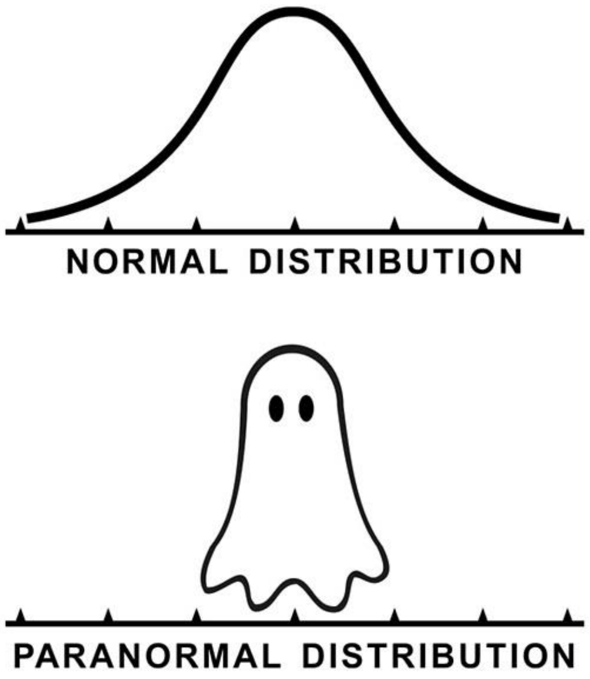
PDF: $f_X(k) = \begin{cases} \lambda e^{-\lambda k} & \text{if } k \geq 0 \\ 0 & \text{otherwise} \end{cases}$

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

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:

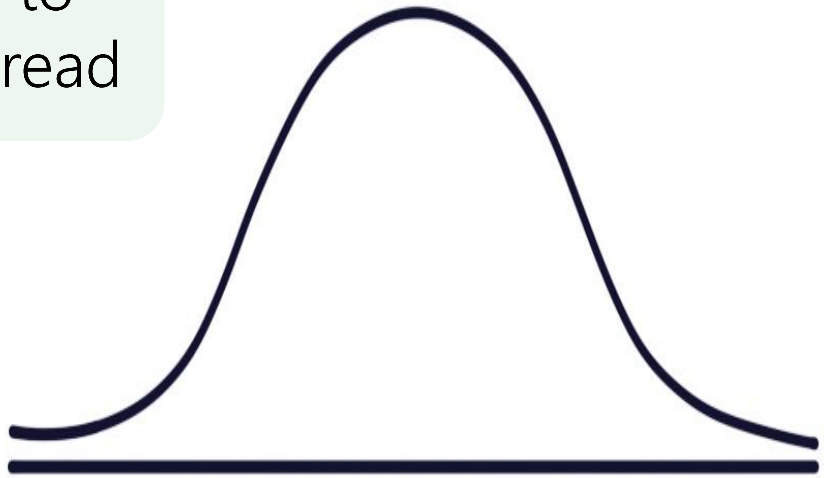
symmetric
around the mean

constant for
normalization

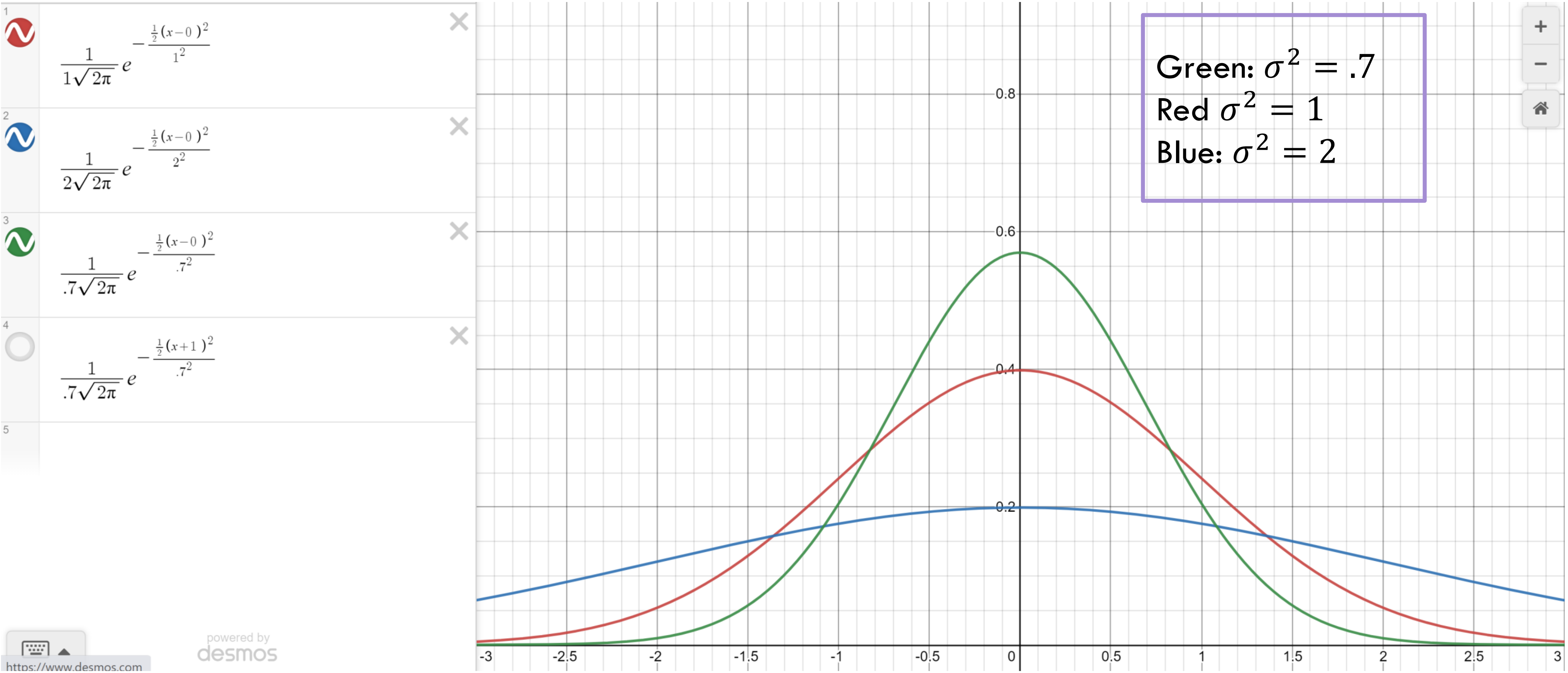
$$f_X(k) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(k-\mu)^2}{2\sigma^2}}$$

variance to
control spread

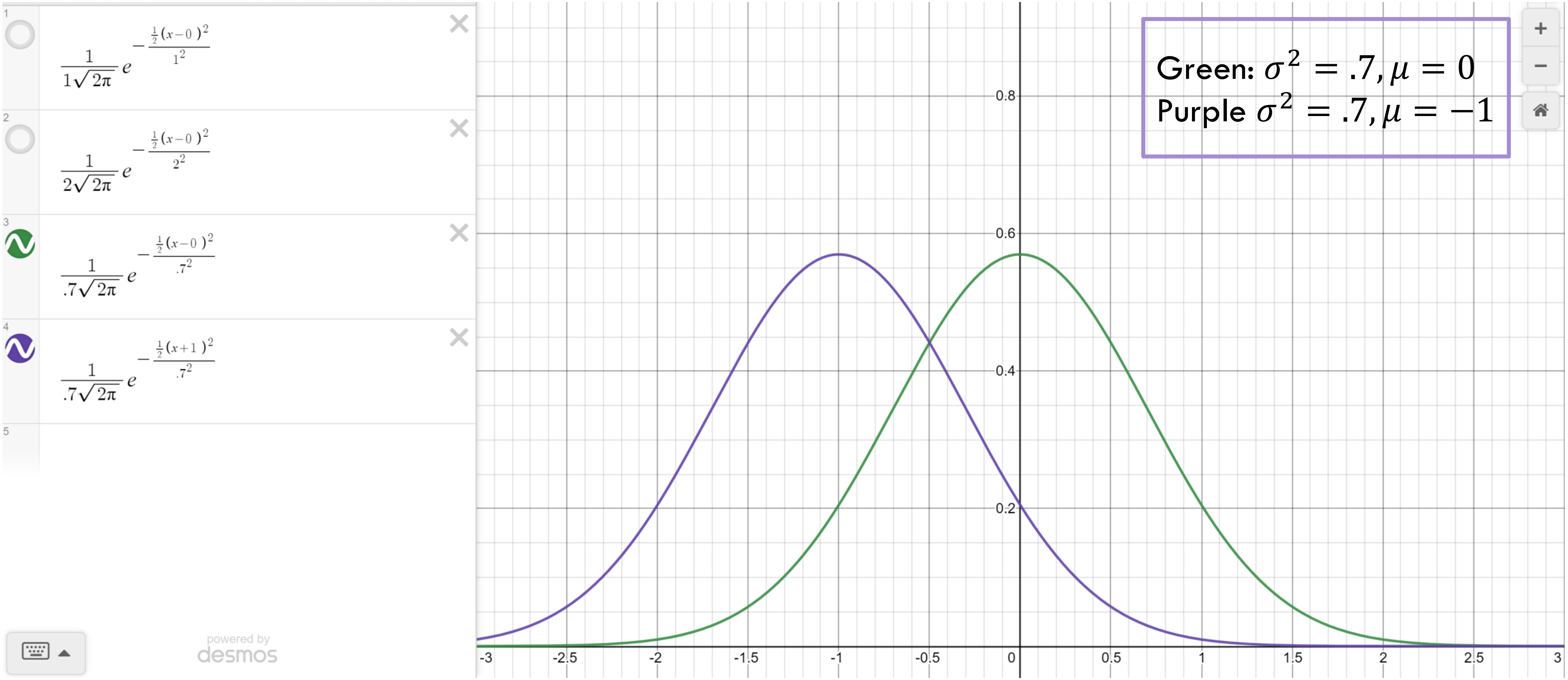
exponential
term for tails



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 independent 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 not 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 $Z \sim \mathcal{N}(0,1)$**

Yes, we're going to use a table in 2026 🤩
Mainly for consistency in this class.

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

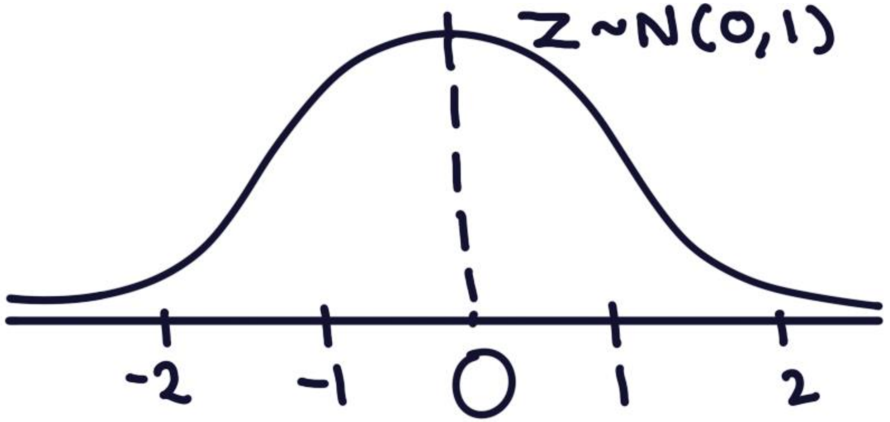
We have a table with precomputed values!

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!

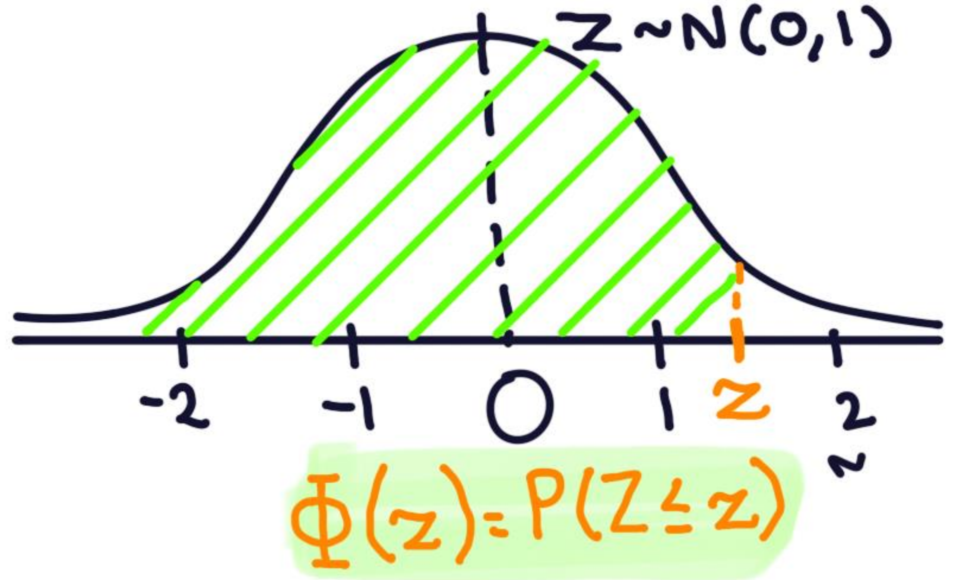
AKA the "z-table", "phi-table"

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 $Z \sim \mathcal{N}(0,1)$**

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

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



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 standard normal random variable: $Z \sim \mathcal{N}(0,1)$

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 $Z \sim \mathcal{N}(0,1)$.

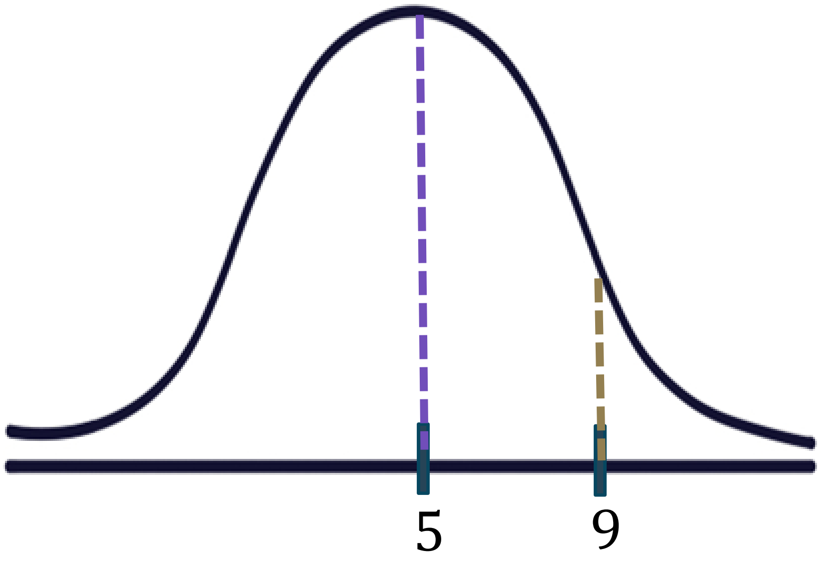
Let $X \sim \mathcal{N}(5,4)$. What is $\mathbb{P}(X \leq 9)$?

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

$= \mathbb{P}\left(\frac{Y-5}{2} \leq \frac{9-5}{2}\right)$ **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 $\mathbb{P}(X > 9)$?

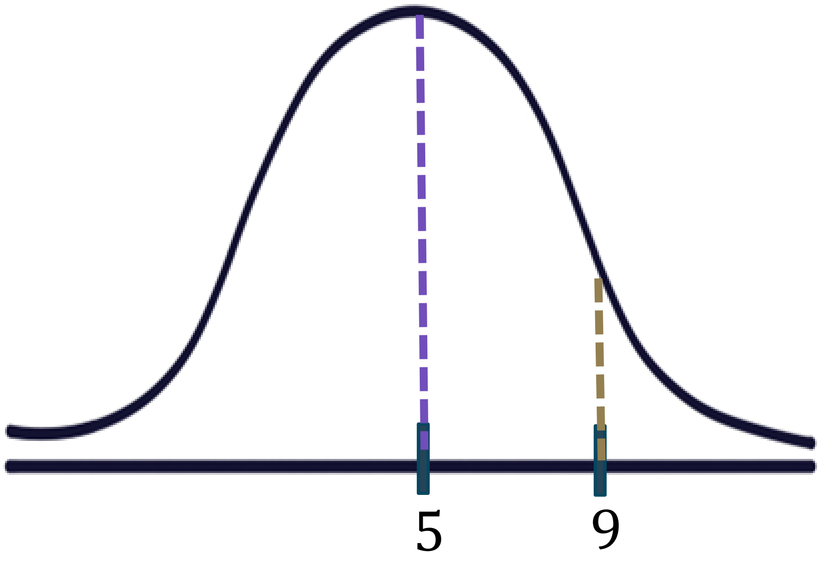
$\mathbb{P}(X > 9) = 1 - \mathbb{P}(X \leq 9)$

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

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

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

$= 1 - 0.97725 = 0.02275$



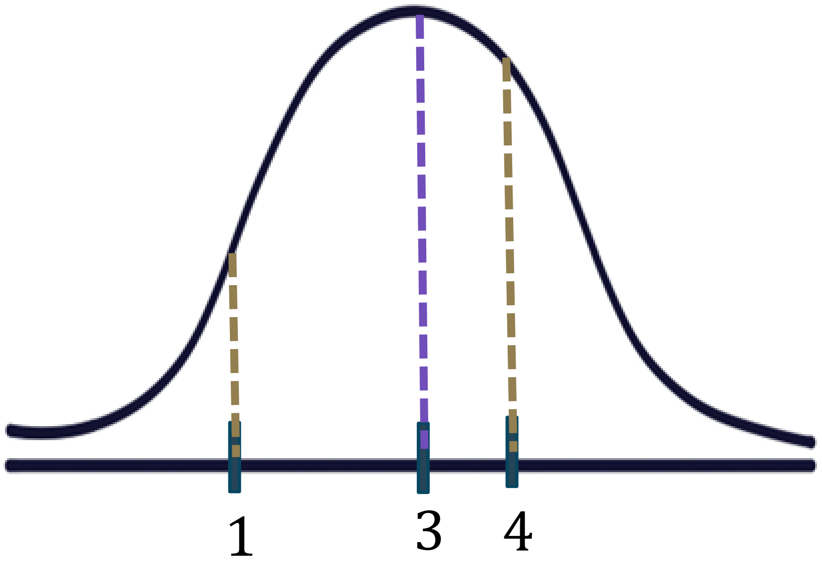
z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.5279	0.53188	0.53586
0.1	0.53983	0.5438	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	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 $1 \leq X \leq 4$?

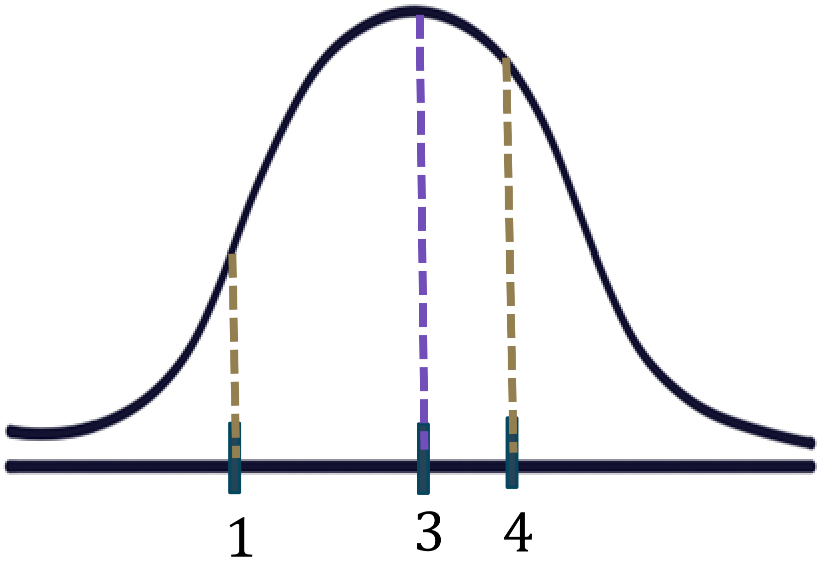
$$\mathbb{P}(1 \leq X \leq 4)$$

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

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

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

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

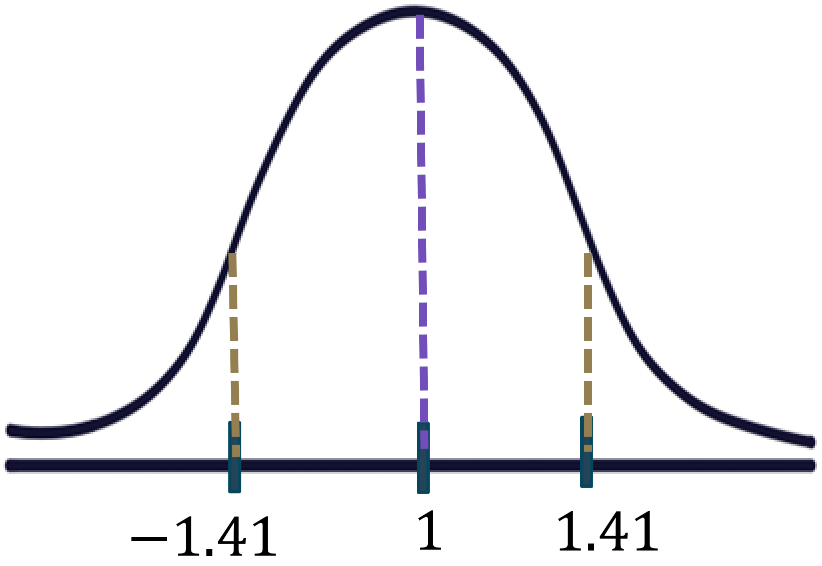


More practice

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

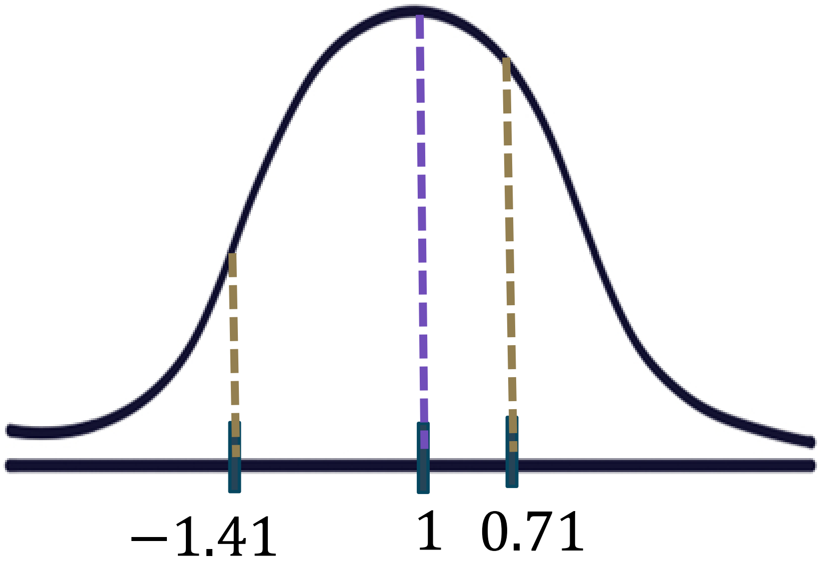
How do we find $\Phi(-1.41) = \mathbb{P}(Z \leq -1.41)$? Our table only has CDF values for positive numbers! 😱

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 $1 \leq X \leq 4$?

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

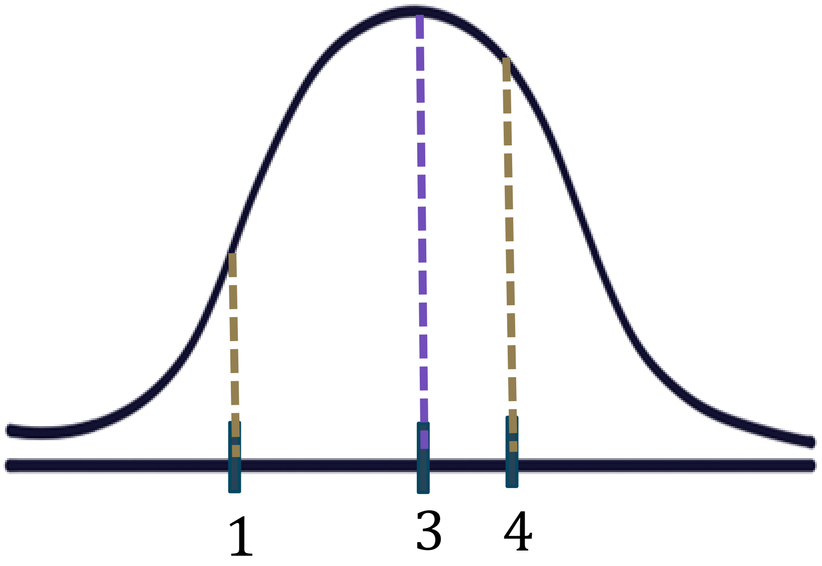
$= \mathbb{P}(X \leq 4) - \mathbb{P}(X \leq 1)$

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

$= \mathbb{P}\left(Z \leq \frac{4-3}{\sqrt{2}}\right) - \mathbb{P}\left(Z \leq \frac{1-3}{\sqrt{2}}\right)$ 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?

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

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

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

$$= \Phi(2) - \left(1 - \Phi(2)\right)$$

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

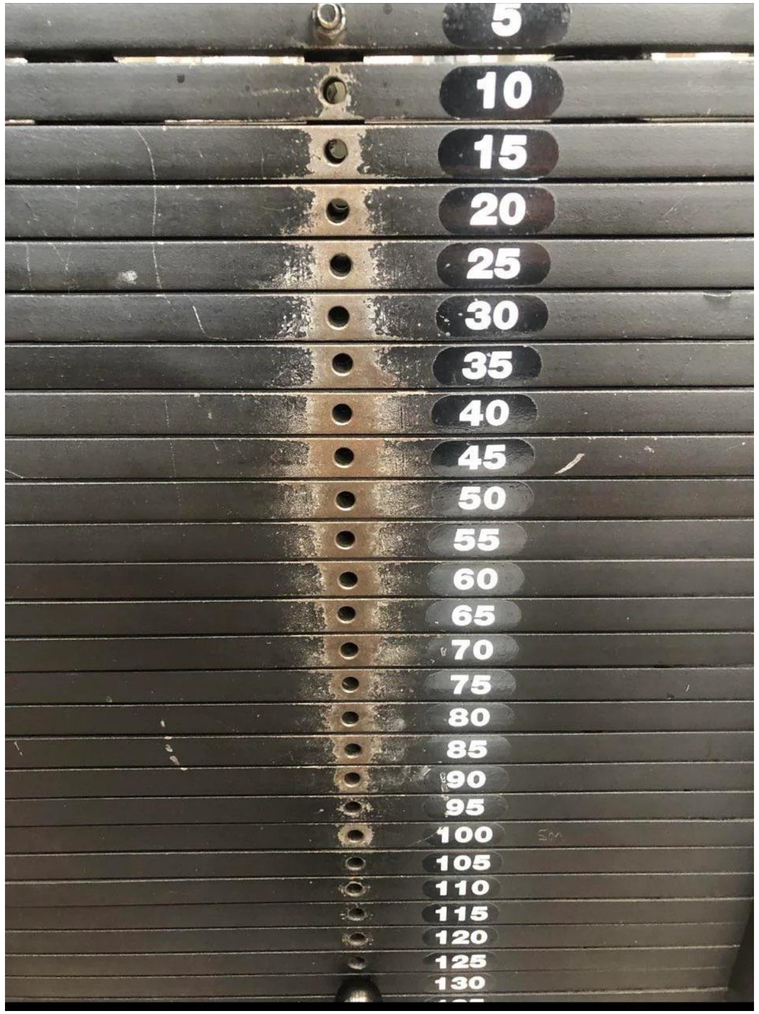
You’ll sometimes hear statisticians refer to the “68-95-99.7 rule” which is the probability of being within 1,2, or 3 standard deviations of the mean.

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

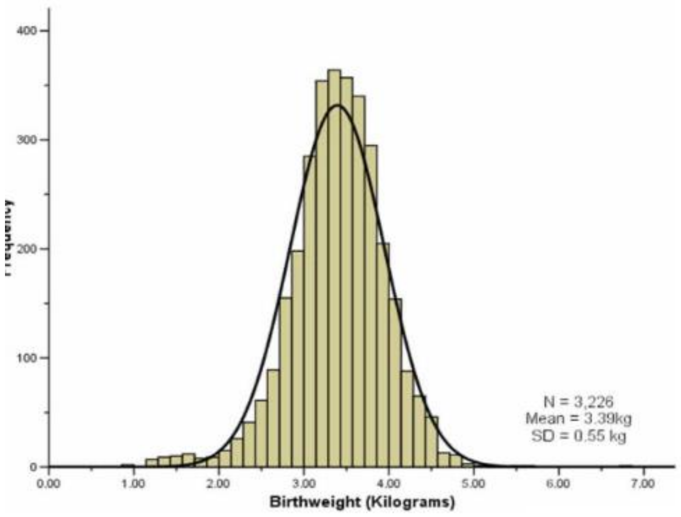
Can you spot the normal distribution?

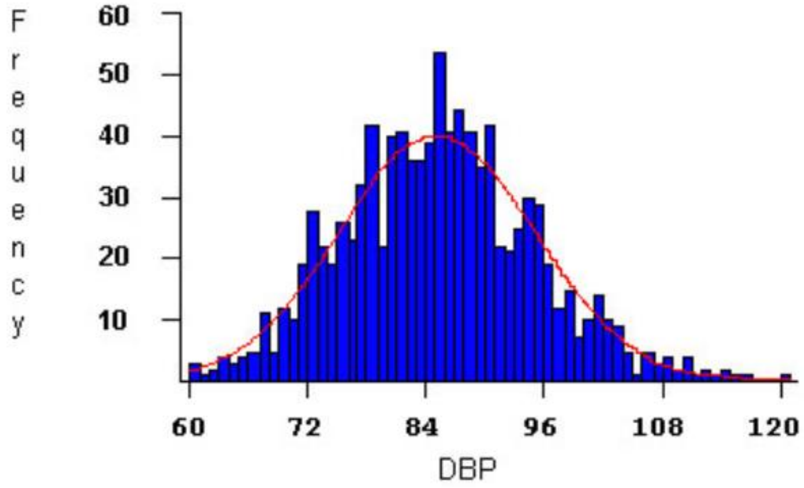


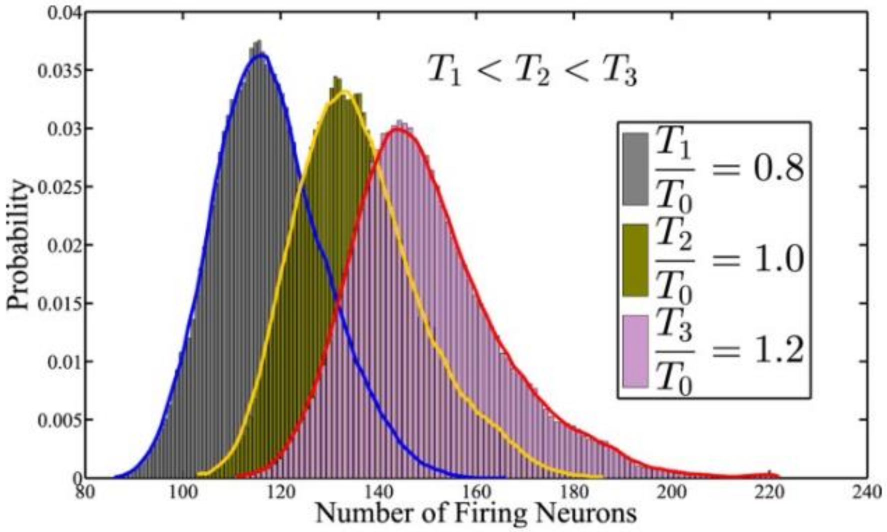
Picture from reddit

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

Normal distributions show up everywhere!







But…why?

This is because of what we call, the *central limit theorem*!

“The sum of **any** independent random variables **approaches** a normal distribution. It becomes closer to normal as we sum more RVs together.”

More formally, the Central Limit Theorem!

This is because of what we call, the *central limit theorem*!

"The sum of **any** independent random variables **approaches** a normal distribution. It becomes closer to normal as we sum more RVs together."

Central Limit Theorem

If $X_1, X_2, \ldots, X_n$ are i.i.d. random variables, each with mean μ and variance σ^2

Let $Y_n = X_1 + X_2 + \cdots + X_n$

As $n \to \infty$, Y_n approaches a normal distribution $\mathcal{N}(n \cdot \mu, n \cdot \sigma^2)$

(i.e., CDF of Y_n converges to the CDF of $\mathcal{N}(n \cdot \mu, n \cdot \sigma^2)$)

What does i.i.d mean?

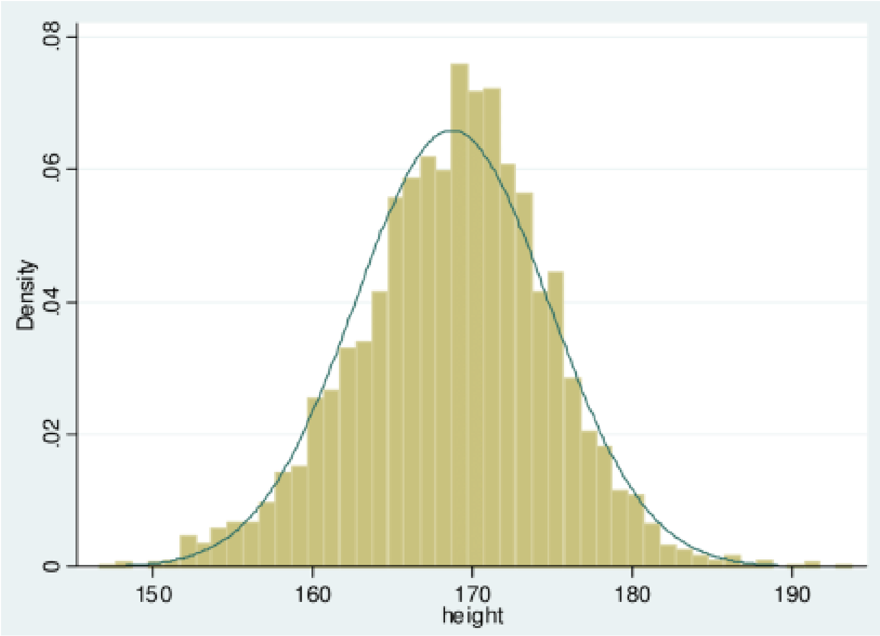
Independent and Identically Distributed (i.i.d)

For random variables $X_1, X_2, \ldots, X_n$ to be i.i.d., they must

- Be <u>mutually independent</u>
  "knowing value of one random variable doesn't give us info about the value of others"
- All have the <u>same PMF</u> (if discrete) or <u>PDF</u> (if continuous)
  "they follow the same probability distribution"

CLT with ***RVs that are <u>NOT</u> i.i.d***

(R.A. Fisher (1918))



Height may be expressed as the *sum of many independent, random factors*

How much milk you drank per day as a child

Which variant of GH1 (a growth hormone) you have

How much protein/calcium/vitamins/minerals you had as a child

How many hours of sleep you averaged

How many hours of physical activity you averaged

$H = H_1 + H_2 +,,, + H_n$ -> $H \sim N(\ldots)$

A version of CLT does work here! But it's outside the scope of this class.

CLT with ***RVs that are i.i.d***

"number of firing neurons" – sum of indicator random variables for whether each neuron fired

Assume: each neuron is **independent**, and has the **same probability**

"number of people who voted for someone" – sum of indicator random variables for (assume people are independent)

Assume: each person makes **independent**, and has the **same probability**

"total amount invested in a year" – sum of random variables four amount invested each day (assume each investment is independent)

Assume: each day's investment is **independent** and follows the **same distribution**

We will use CLT in this class on problems like this.

A Sum of i.i.d Random Variables

If we have $X_1, X_2, \dots, X_n$ as i.i.d RVs each with mean μ and variance σ^2

$S_n = X_1 + X_2 + \cdots + X_n$ is the sum of those RVs. Then…

> **Expectation.** *by linearity of expectation…*

$$\mathbb{E}[S_n] = \mathbb{E}[X_1 + X_2 + \cdots + X_n] = \mathbb{E}[X_1] + \mathbb{E}[X_2] + \cdots + \mathbb{E}[X_n] = n\mu$$

> **Variance**. *by linearity of variance <u>because of independence</u>*

$$Var(S_n) = Var(X_1 + X_2 + \cdots + X_n) = Var(X_1) + Var(X_2) + \cdots + Var(n\sigma^2$$

Proof of the CLT?

We're not going to cover the proof here.

How is the proof done?

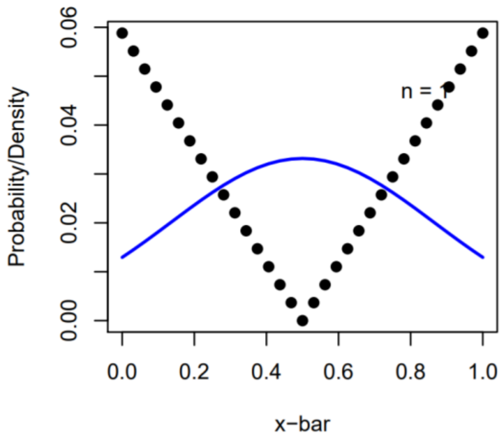
Step 1: Prove that for all positive integers $k, [(Y_n)^k] \to \mathbb{E}[Z^k]$

Step 2: Prove that if $\mathbb{E}[(Y_n)^k] = \mathbb{E}[Z^k]$ for all k then $F_{Y_n}(z) = F_Z(z)$

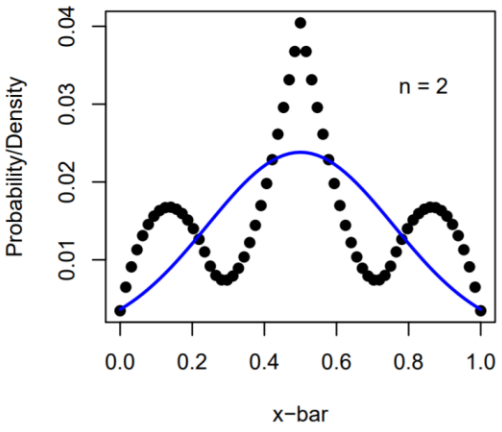
“Proof by example”

CLT says a sum of $\boldsymbol{n}$ i.i.d RVs approaches a normal distribution as $\boldsymbol{n}$ gets larger

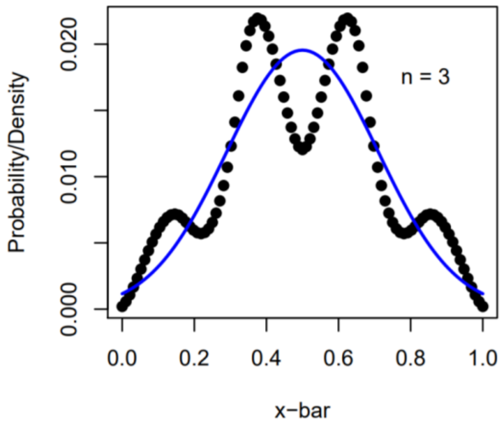
$n = 1$



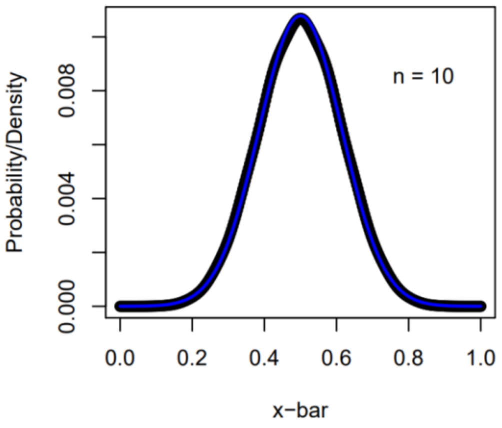
$n = 2$



$n = 3$



$n = 10$



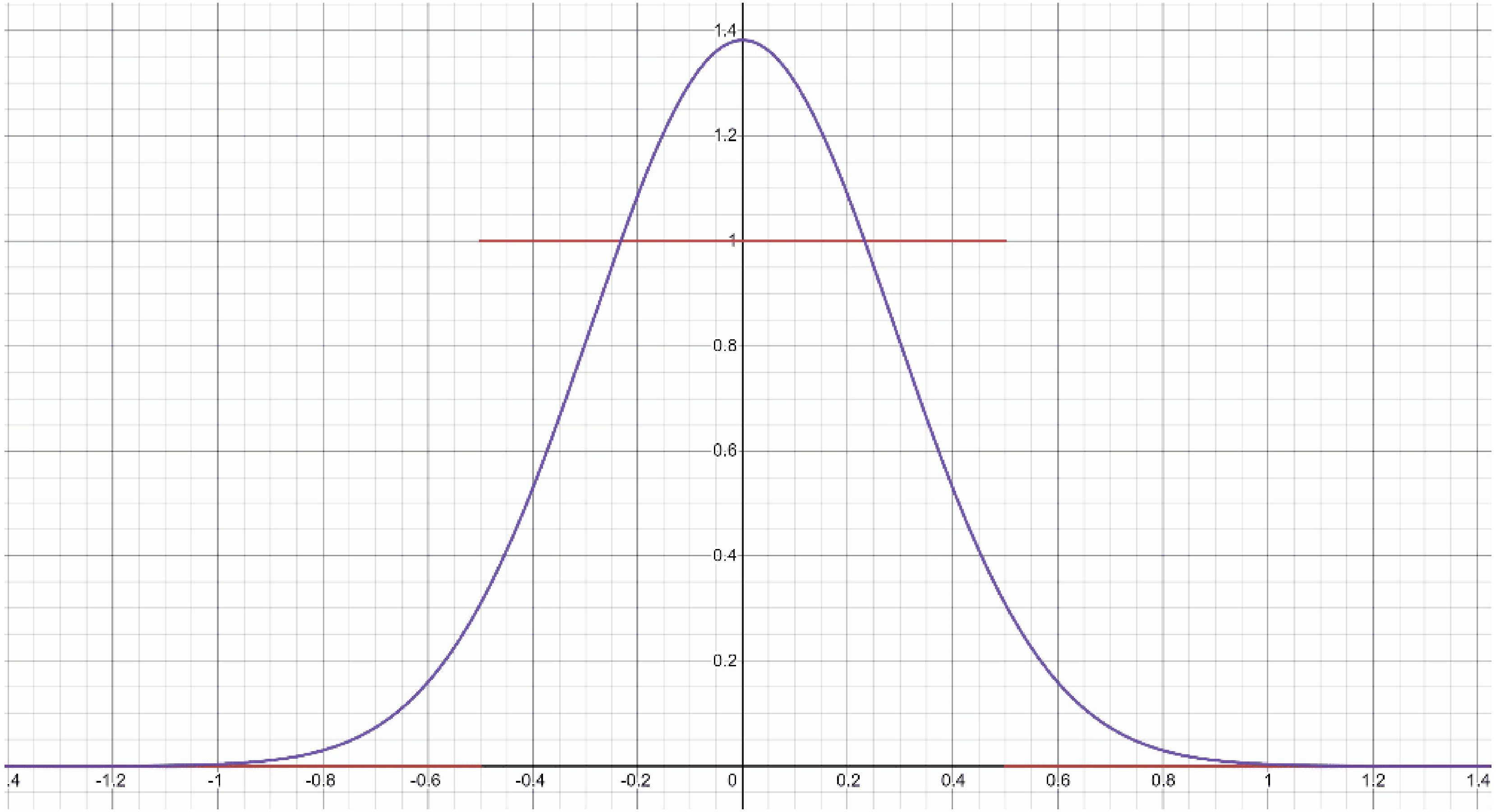
n is the number of i.i.d RVs summed

The **dotted lines** show an “empirical PMF” – a PMF estimated by running the experiment a large number of times.

The blue line is the normal RV that the CLT predicts.

Shown are $n = 1,2,3,10$

“Proof by example” -- uniform

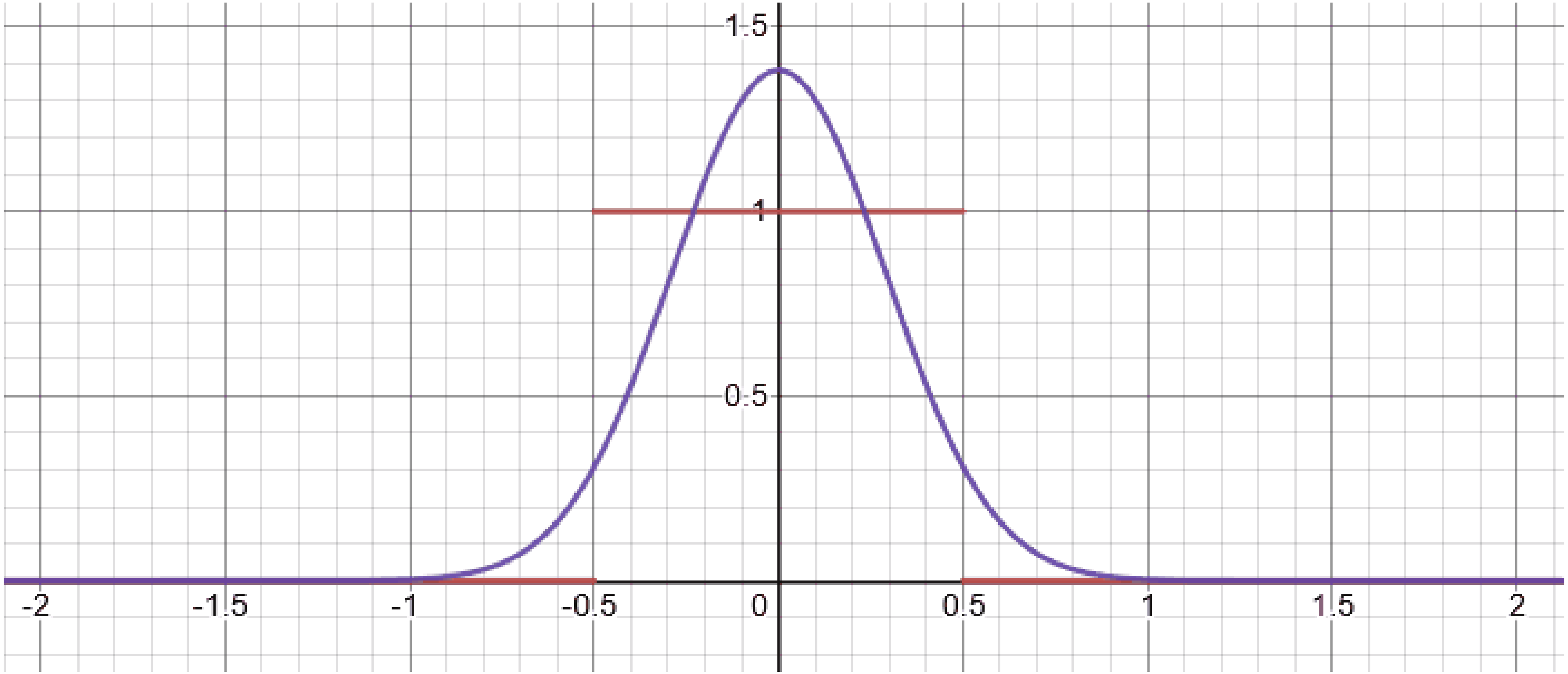


https://www.desmos.com/calculator/2n2m05a9km

"Proof by example" -- uniform

$$X = X_1$$

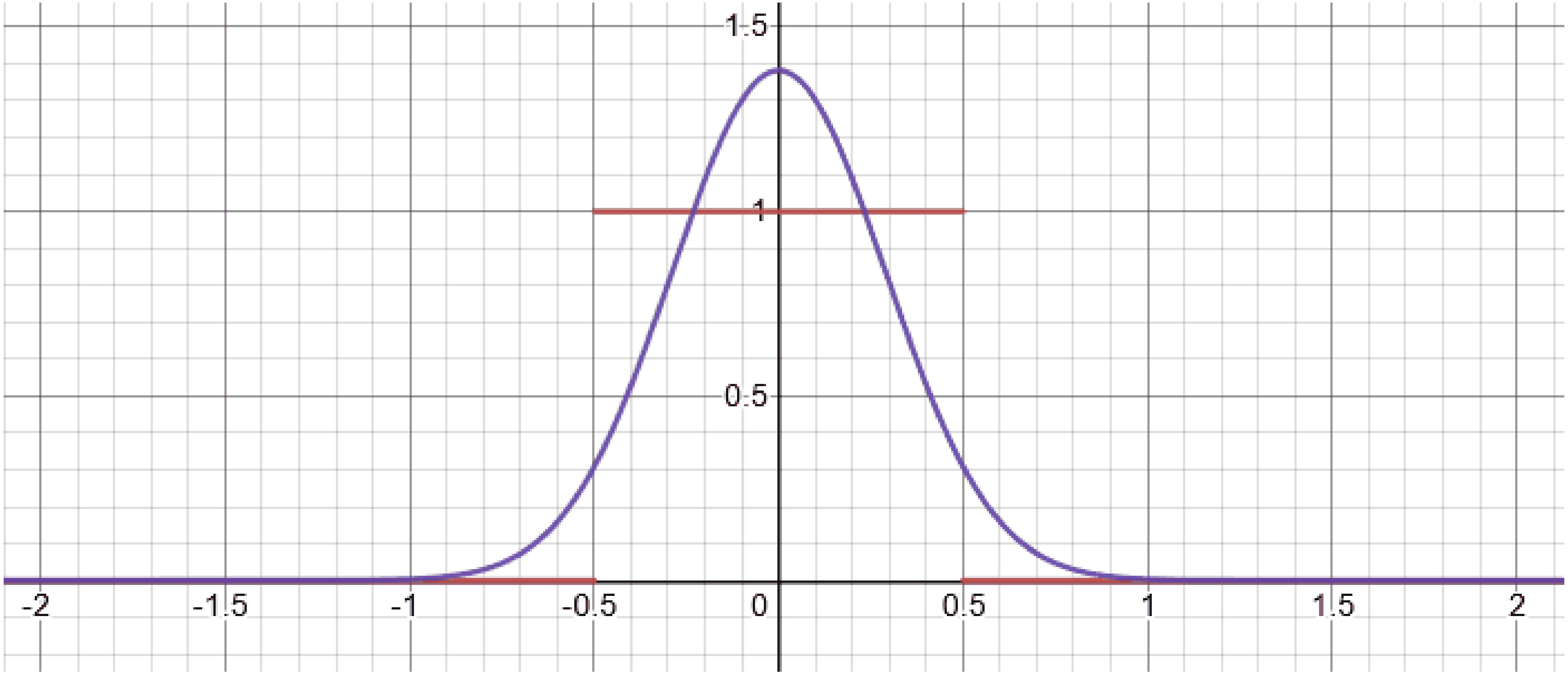
where each $X_i \sim \mathrm{Unif}(0,1)$ and is independent



"Proof by example" -- uniform

$$\boldsymbol{X} = \boldsymbol{X_1}$$

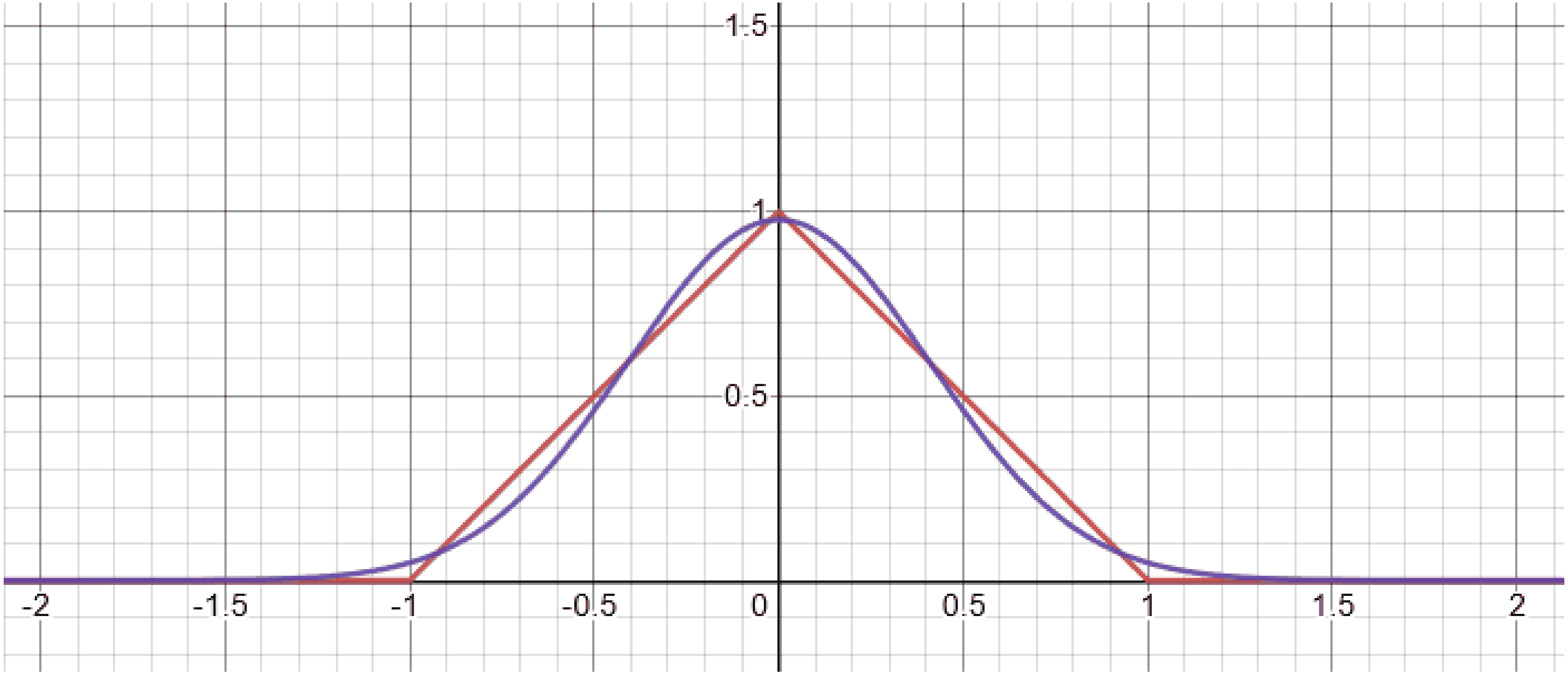
where each $X_i \sim \mathrm{Unif}(0,1)$ and is independent



“Proof by example” -- uniform

$$X = X_1 + X_2$$

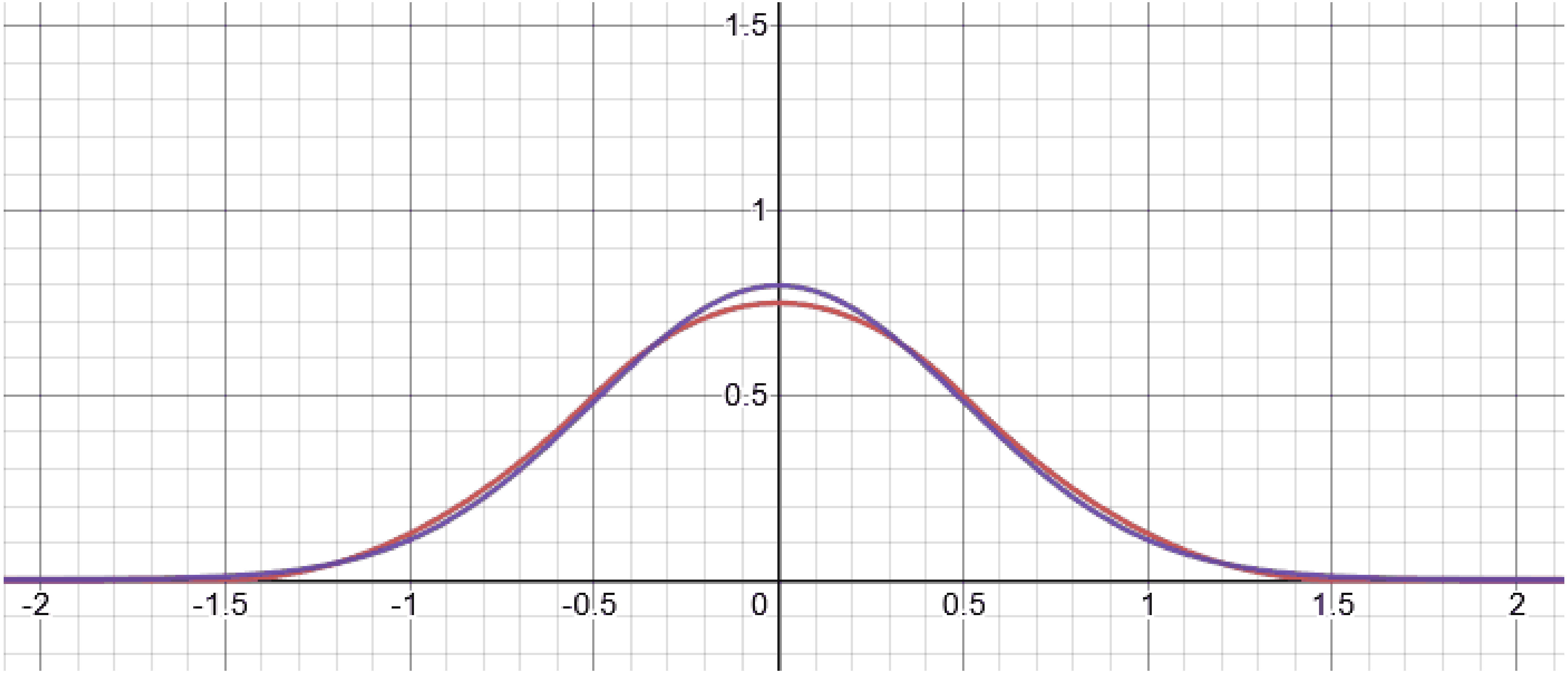
where each $X_i \sim \mathrm{Unif}(0,1)$ and is independent



"Proof by example" -- uniform

$$X = X_1 + X_2 + X_3$$

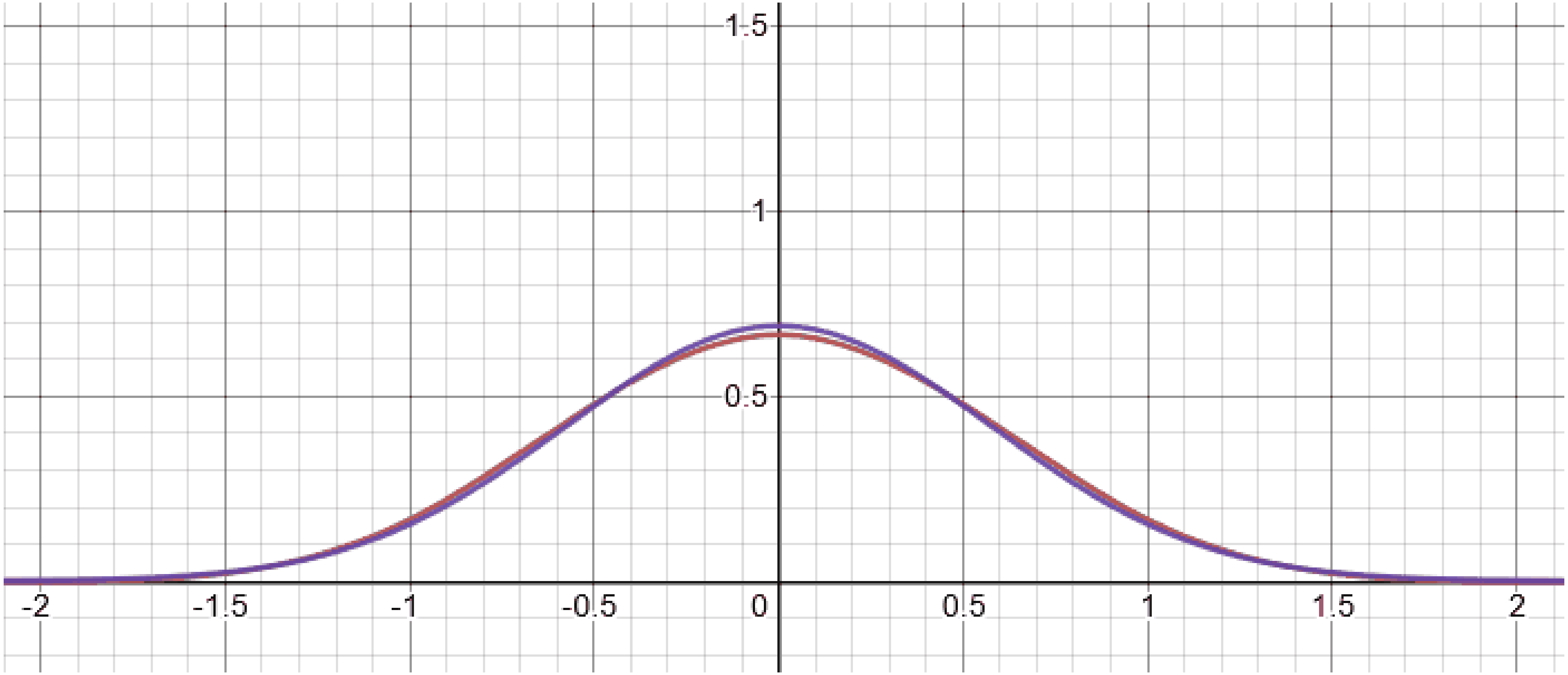
where each $X_i \sim \text{Unif}(0,1)$ and is independent



“Proof by example” -- uniform

$$X = X_1 + X_2 + X_3 + X_4$$

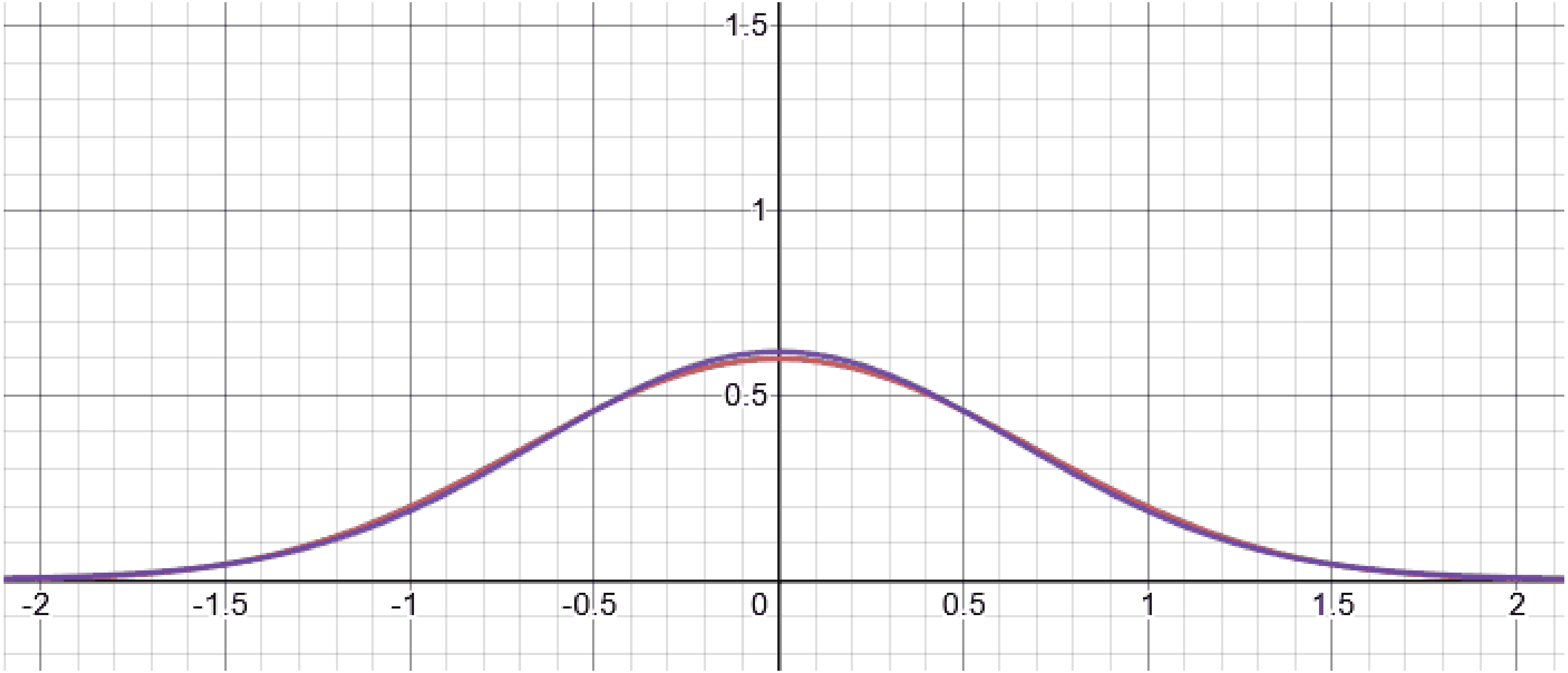
where each $X_i \sim \mathrm{Unif}(0,1)$ and is independent



“Proof by example” -- uniform

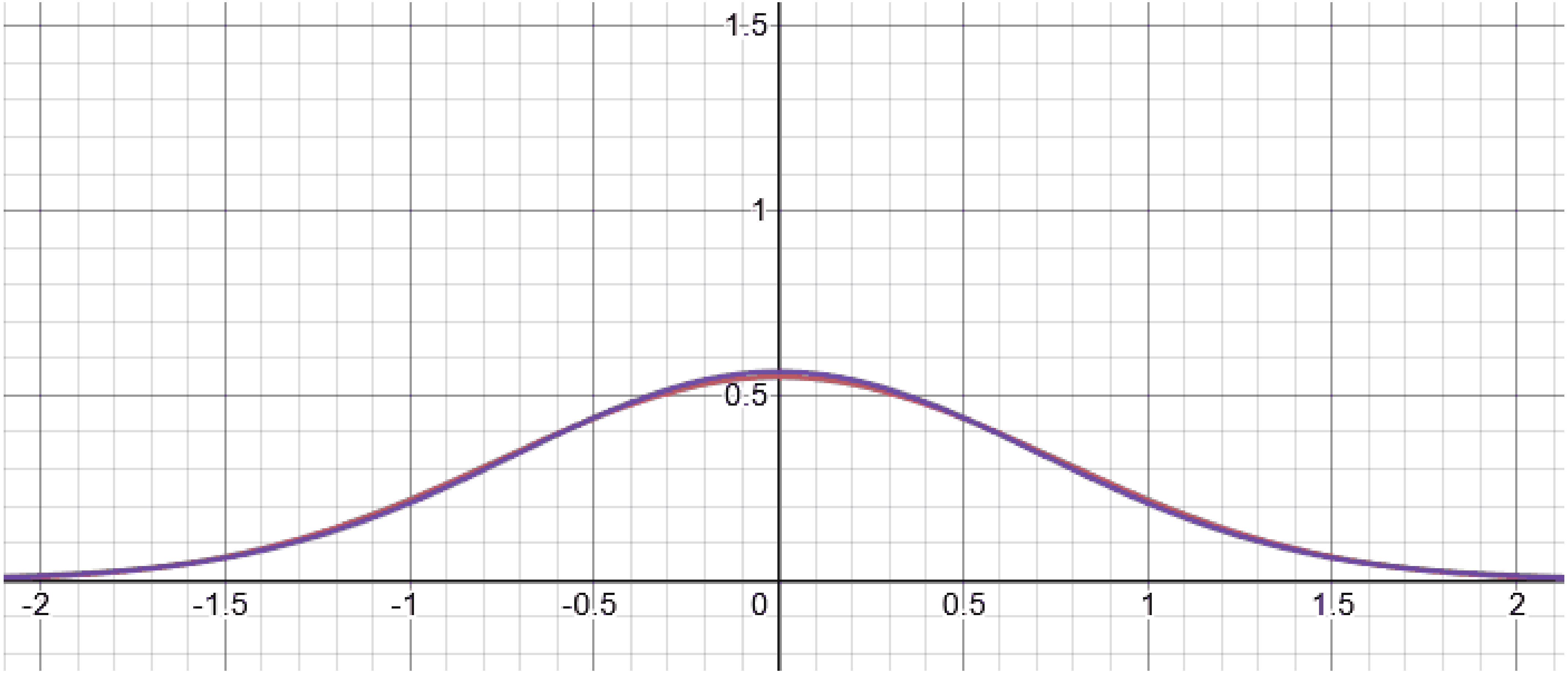
$$\boldsymbol{X} = \boldsymbol{X_1} + \boldsymbol{X_2} + \boldsymbol{X_3} + \boldsymbol{X_4} + \boldsymbol{X_5}$$

where each $X_i \sim \text{Unif}(0,1)$ and is independent



"Proof by example" -- uniform

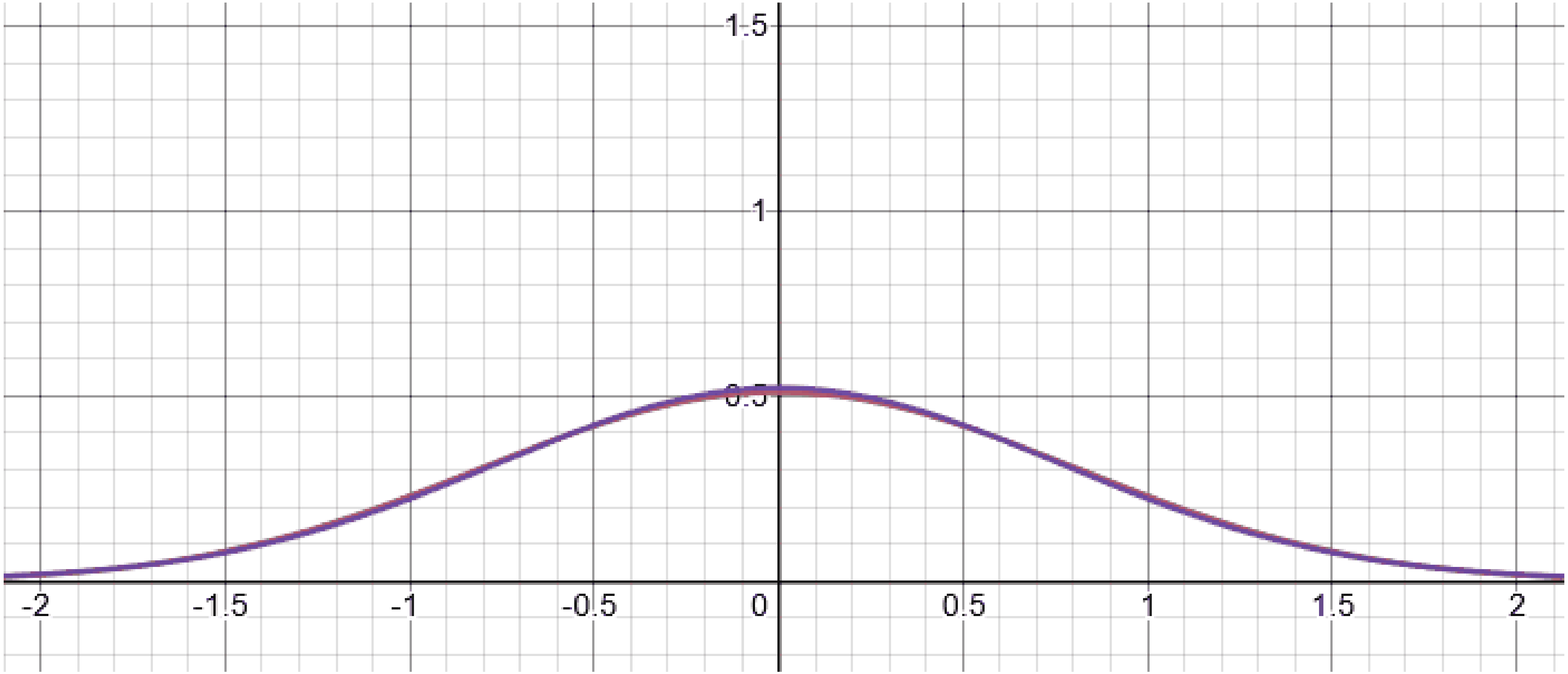
$$X = X_1 + X_2 + X_3 + X_4 + X_5 + X_6$$
where each $X_i \sim \mathrm{Unif}(0,1)$ and is independent



"Proof by example" -- uniform

$$X = X_1 + X_2 + X_3 + X_4 + X_5 + X_6 + X_7$$

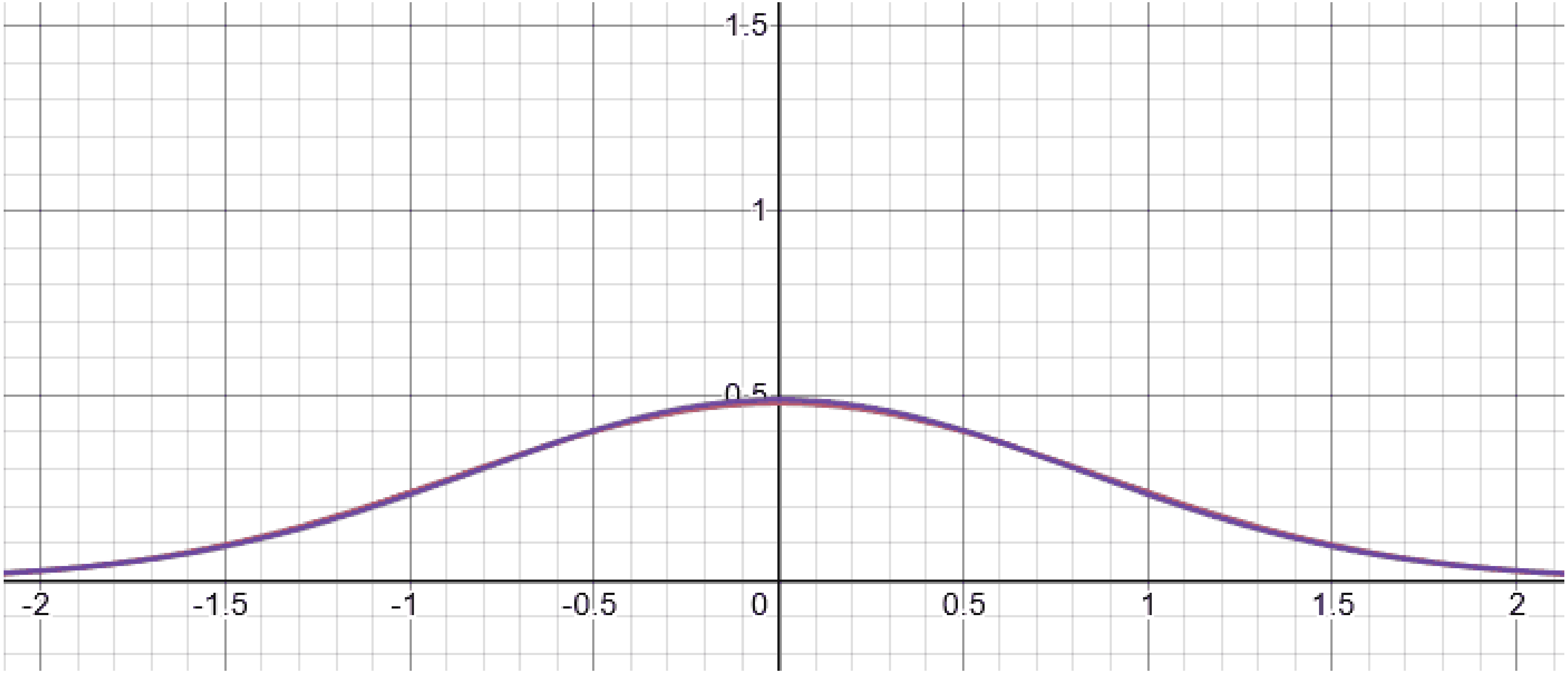
where each $X_i \sim \mathrm{Unif}(0,1)$ and is independent



“Proof by example” -- uniform

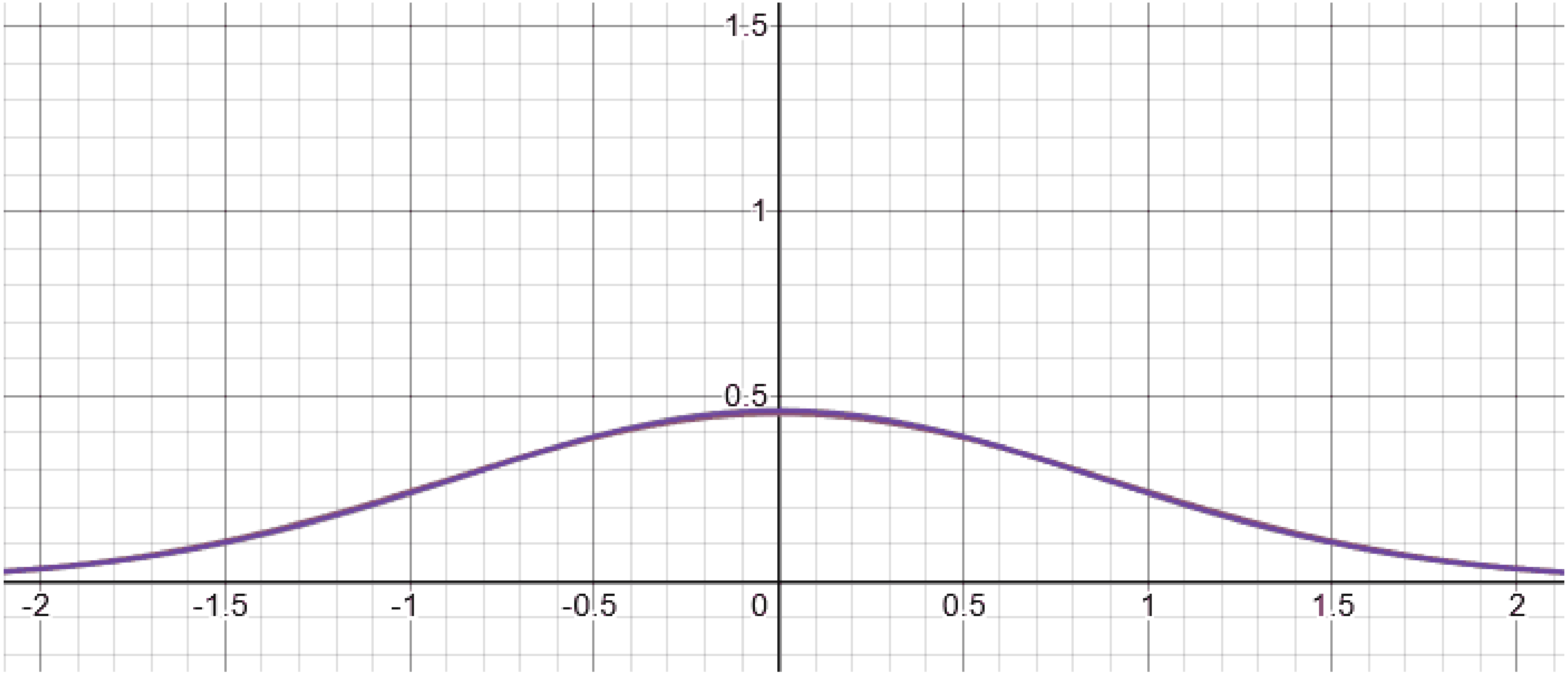
$$\boldsymbol{X = X_1 + X_2 + X_3 + X_4 + X_5 + X_6 + X_7 + X_8}$$

where each $X_i \sim \mathrm{Unif}(0,1)$ and is independent



“Proof by example” -- uniform

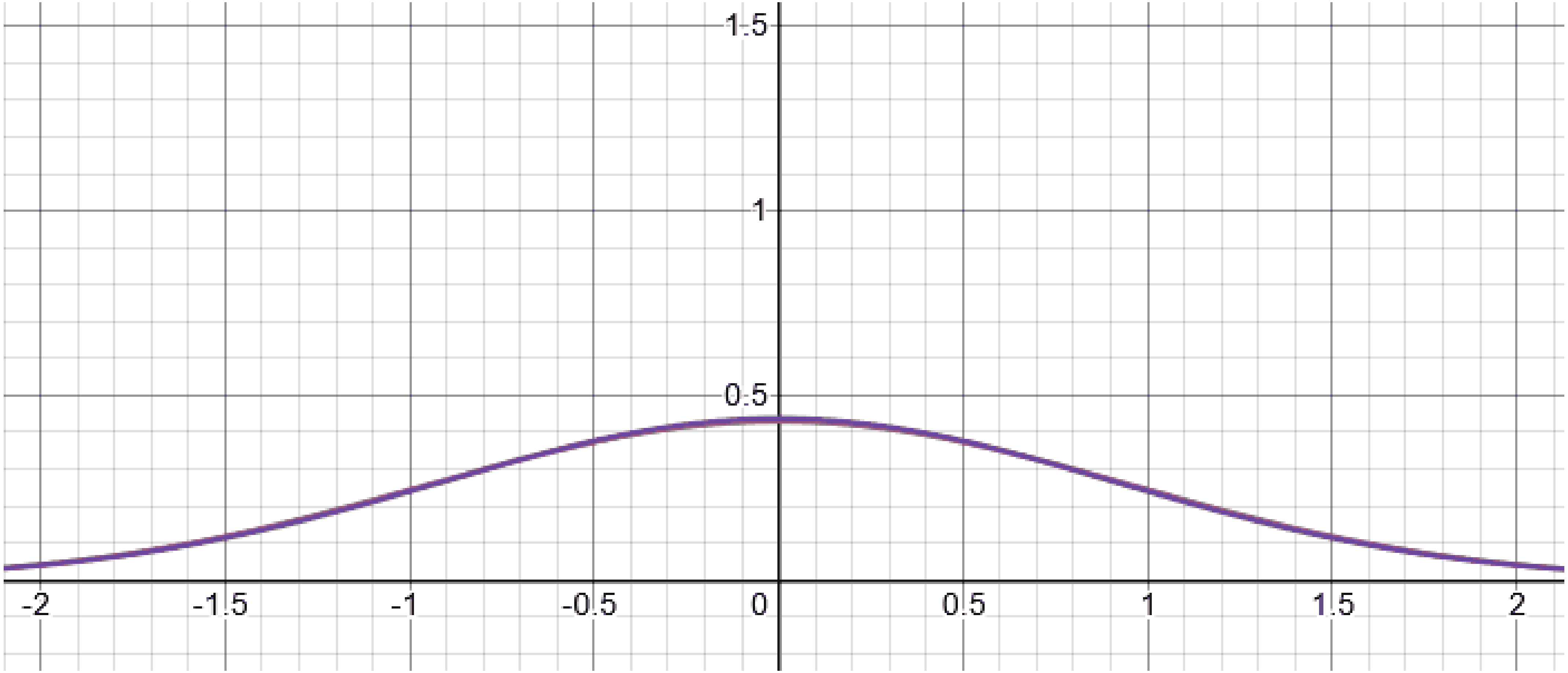
$$\boldsymbol{X = X_1 + X_2 + X_3 + X_4 + X_5 + X_6 + X_7 + X_8 + X_9}$$
where each $X_i \sim \text{Unif}(0,1)$ and is independent



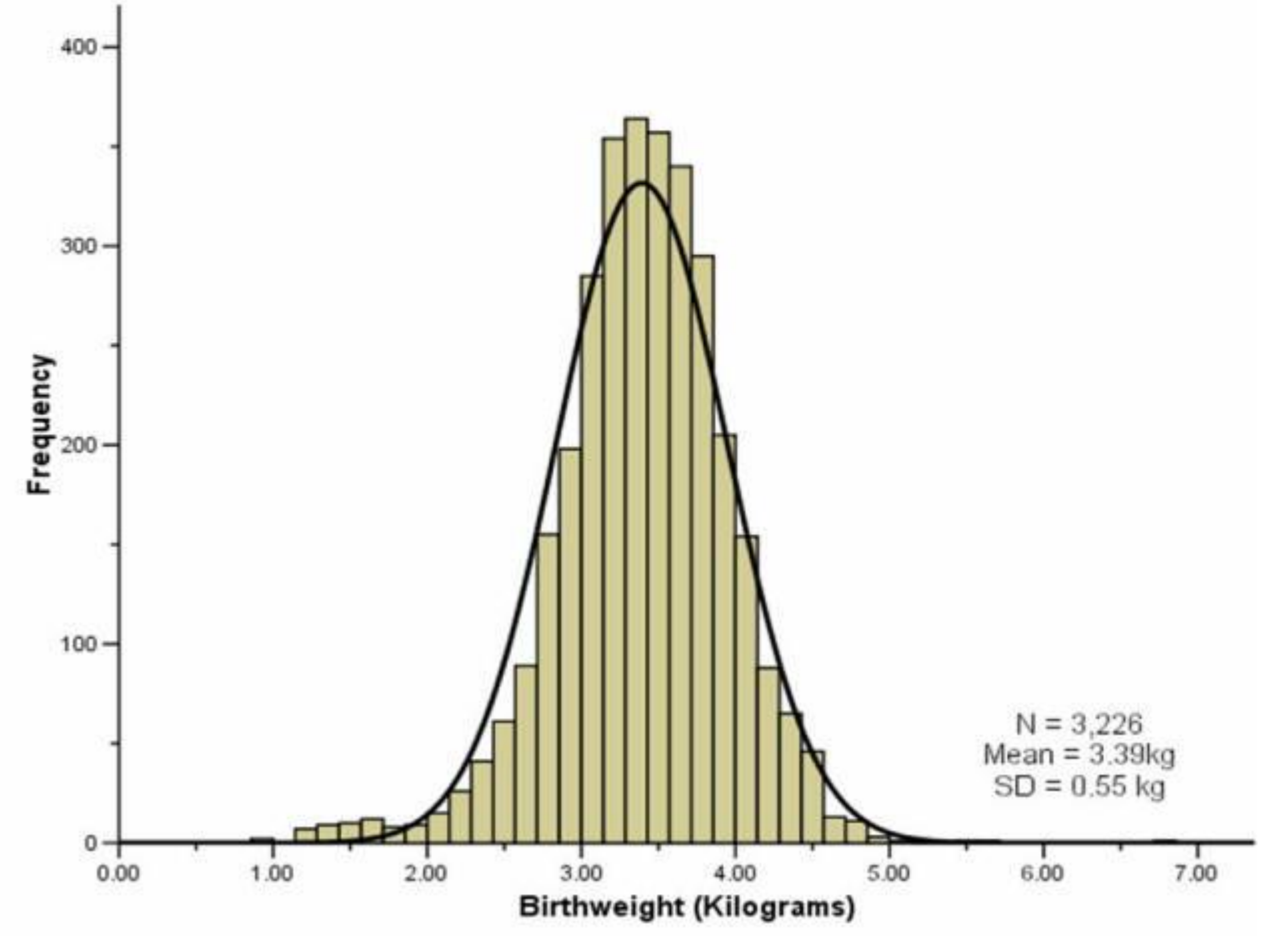
"Proof by example" -- uniform

$$\boldsymbol{X = X_1 + X_2 + X_3 + X_4 + X_5 + X_6 + X_7 + X_8 + X_9 + X_{10}}$$

where each $X_i \sim \mathrm{Unif}(0,1)$ and is independent



“Proof by real-world”



birthweight

A lot of real-world bell-curves can be explained as:

1. The random variable comes from a combination of independent factors.

2. The CLT says the distribution will become like a bell curve.

Theory vs. Practice

> The formal theorem statement is "in the limit"

You might not get exactly a normal distribution for any finite n (e.g. if you sum discrete, the sum is always discrete and will be discontinuous for every finite n.

> In practice, the approximations get very accurate very quickly (at least with a few tricks we'll see soon).

They won't be exact (unless the X_i are normals) but it's close enough to use even with relatively small n.