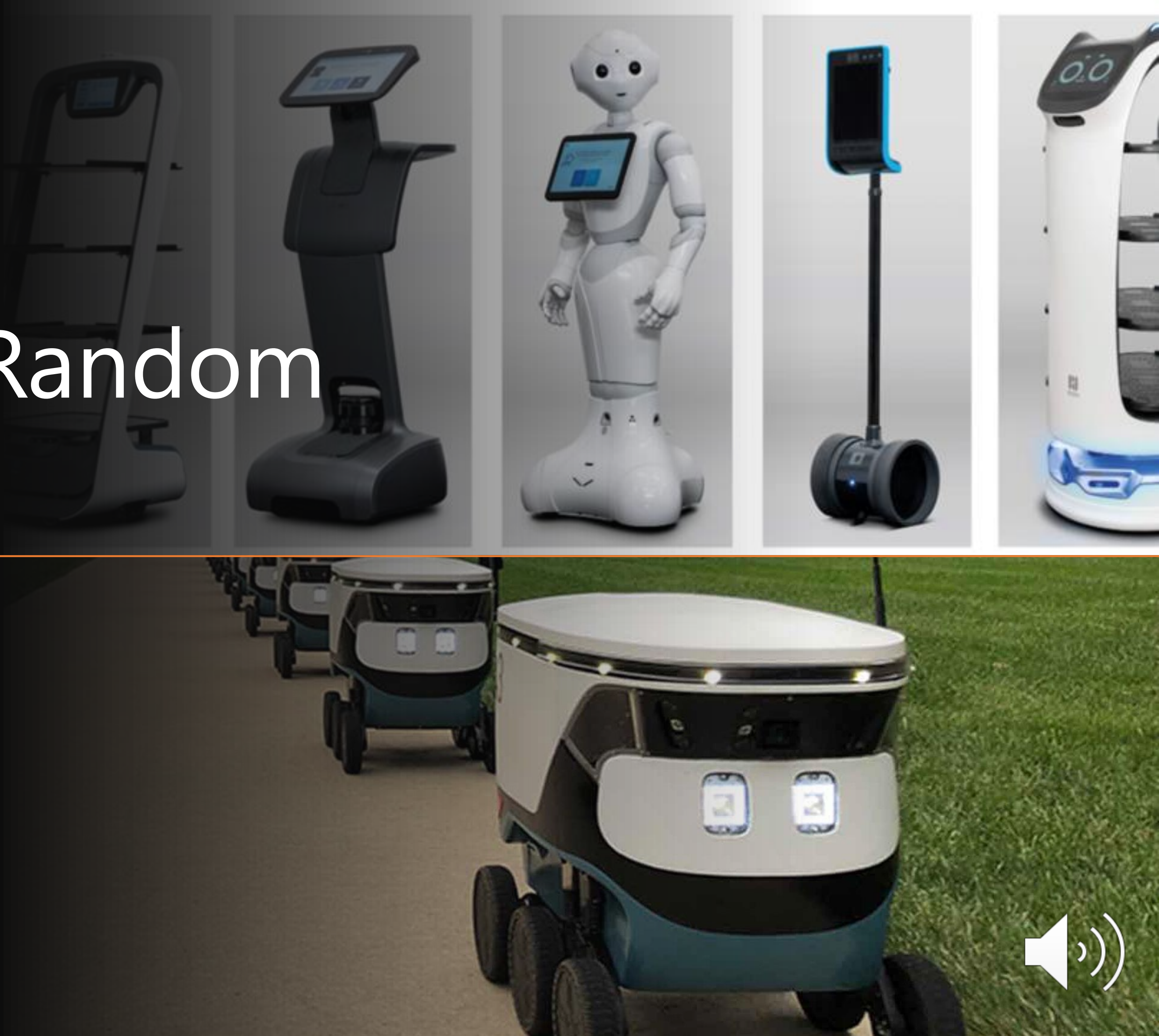
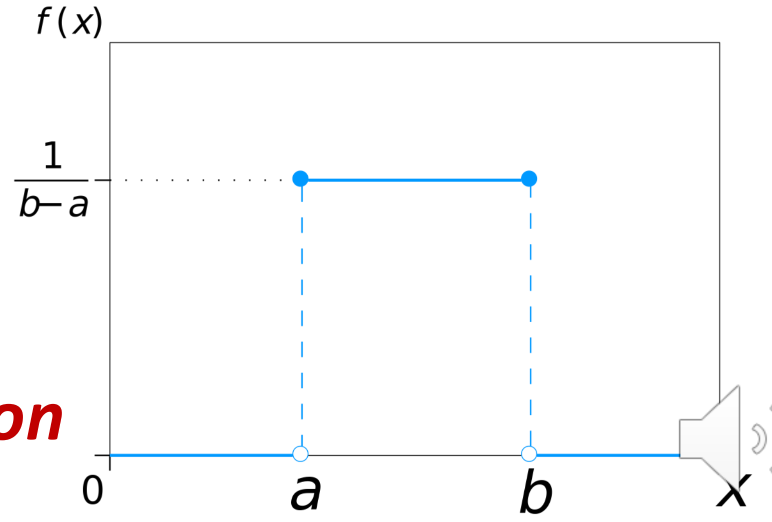
Continuous Random Variables

CS 3630



The uniform distribution

- The uniform distribution is the simplest example of a continuous random variable.
- We saw this distribution in our sampling algorithm.
- We use the notation $X \sim U(a, b)$ to denote that X is a continuous random variable with uniform distribution on the interval $[a, b]$.
- The pdf for such a random variable is given by:



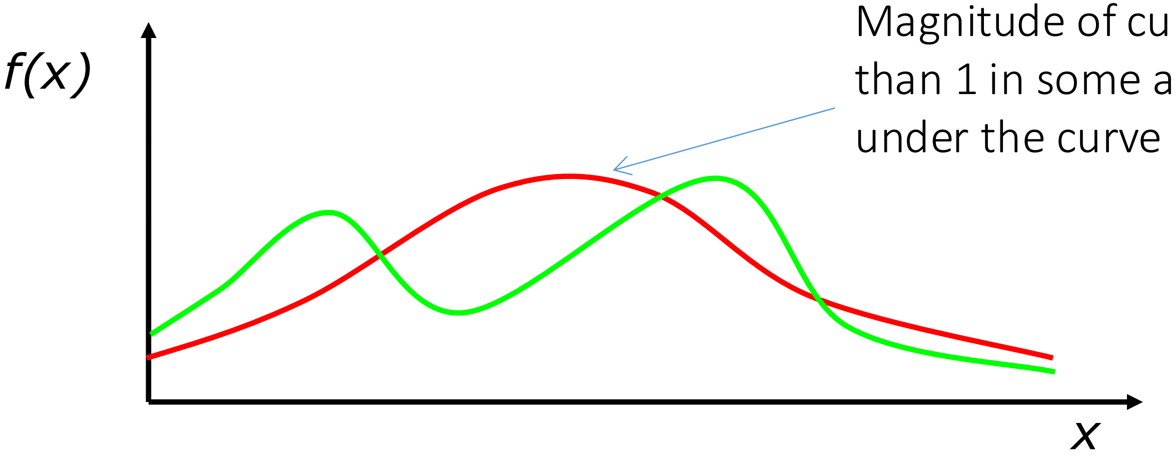
pdf for the uniform distribution

More about pdf's

The total area under a pdf equals 1, always, for every pdf.

$$\int_{-\infty}^{\infty} f_X(u)du = F_X(\infty) - F_X(-\infty) = 1 - 0 = 1$$

But the magnitude of $f_X(u)$ can take any non-negative value – so long as the total area under the curve integrates to one!

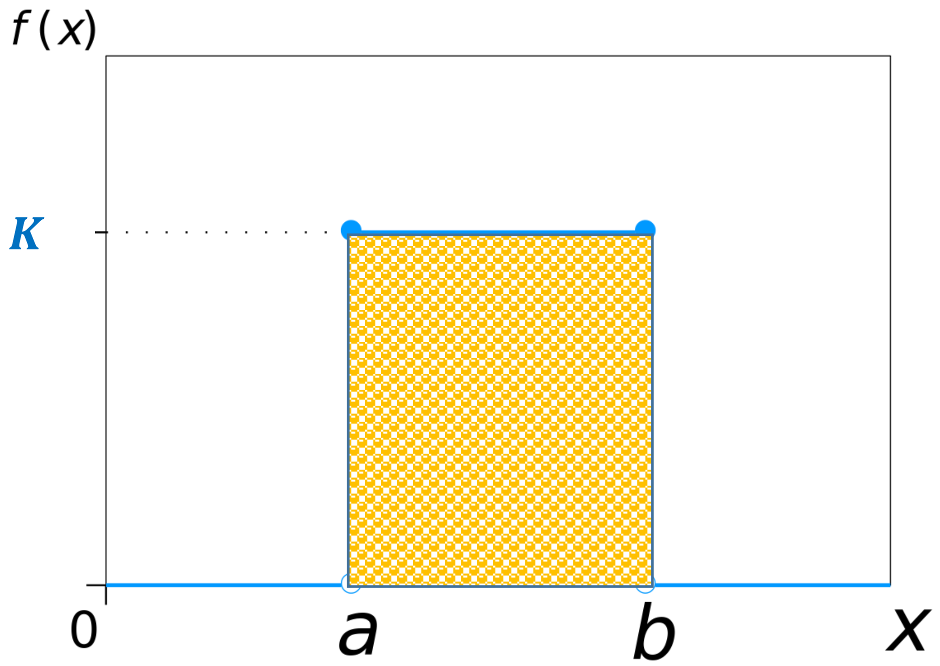


Magnitude of curve could be greater than 1 in some areas. The total area under the curve must add up to 1.

The uniform distribution (again)

- It is now easy to understand why the "height" of the uniform distribution pdf is $\frac{1}{b-a}$:

$$\mathbf{1 = P(a \leq X \leq b) = \int_a^b K\,du = Kb - Ka \rightarrow K = \frac{1}{b-a}}$$



In this case, the geometry of rectangles is enough to tell us the answer:

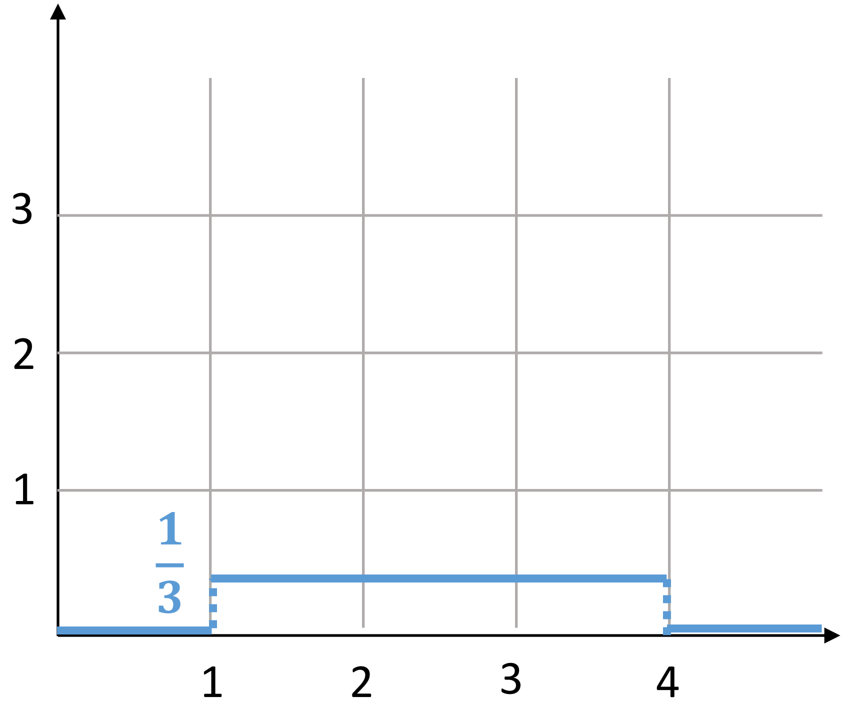
$$Area = K\,(b-a)$$

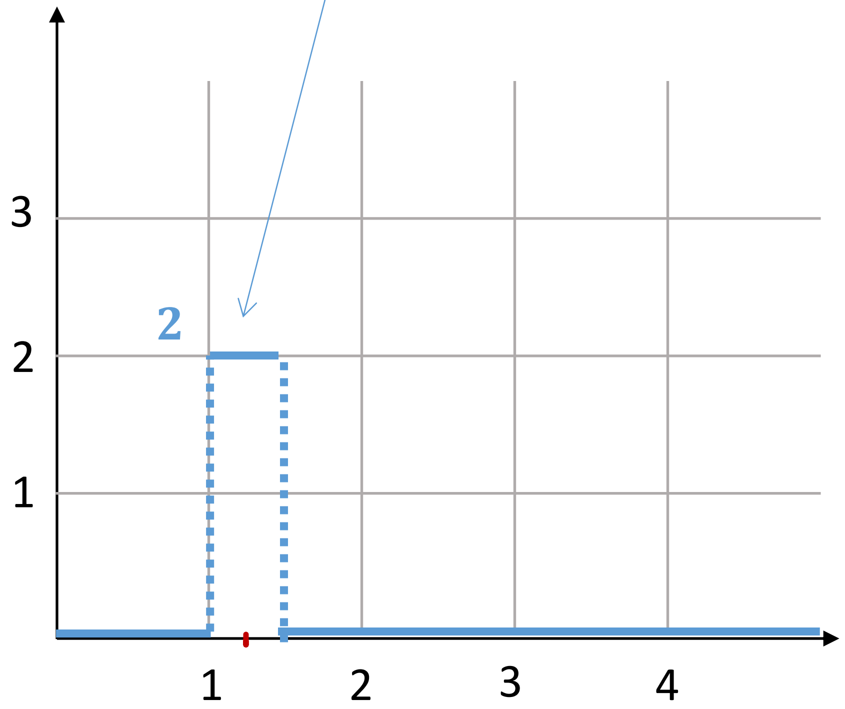
So, if $Area = 1$, then we must have

$$\mathbf{K = \frac{1}{b-a}}$$

The uniform distribution

Magnitude of curve could be greater than 1 in some areas. The total area under the curve must add up to 1.





Continuous random variables

- Recall the cumulative distribution function (CDF):

$$F_X(\alpha) = P(X \leq \alpha) = \sum_{i=0}^{k-1} p_X(x_i)$$

r.v. $\boldsymbol{x}$	$\boldsymbol{p_X(x)}$
0	0.20
1	0.30
2	0.25
3	0.20
4	0.05

Category ($\boldsymbol{\omega}$)	**r.v. $\boldsymbol{x}$**	$\boldsymbol{F_X(\alpha)}$
Cardboard	0	$P(X \leq 0) = 0.20$, $\boldsymbol{\alpha} = 0$
Paper	1	$P(X \leq 1) = 0.50$, $\boldsymbol{\alpha} = 1$
Cans	2	$P(X \leq 2) = 0.75$, $\boldsymbol{\alpha} = 2$
Scrap Metal	3	$P(X \leq 3) = 0.90$, $\boldsymbol{\alpha} = 3$
Bottle	4	$P(X \leq 4) = 1.00$, $\boldsymbol{\alpha} = 4$

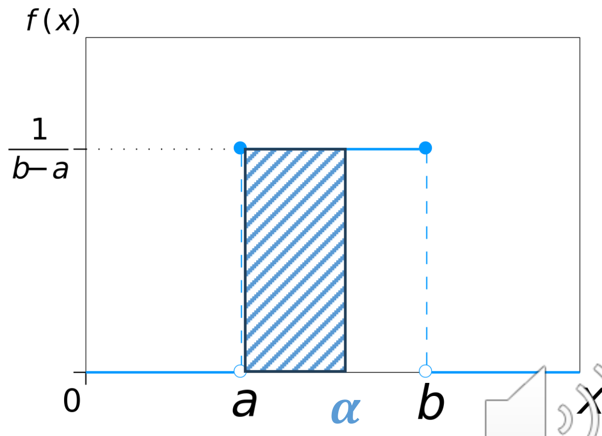
Continuous random variables

- Recall the cumulative distribution function (CDF):

$$F_X(\alpha) = P(X \leq \alpha)$$

- If F_X is continuous everywhere, then X is a ***continuous random variable***.

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



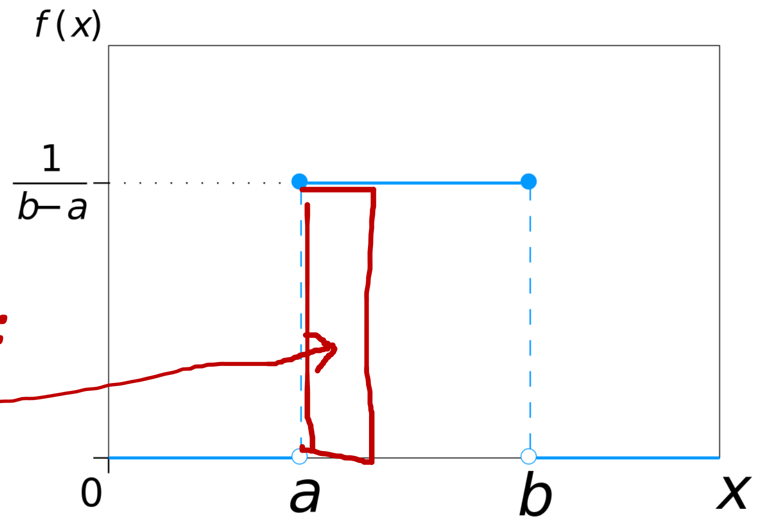
The uniform distribution's CDF

- It's easy to compute the CDF for the uniform distribution given its pdf.
- For $(a \leq \beta \leq b)$, simply evaluate the integral

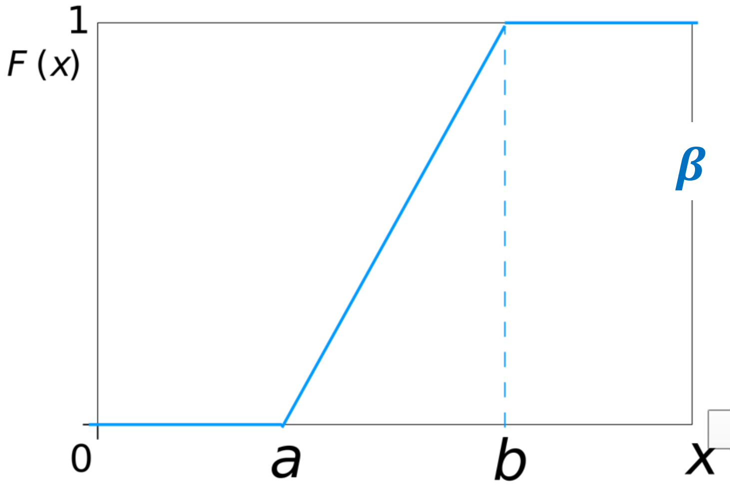
$$F_X(\beta) = P(X \leq \beta) = \int_a^\beta \frac{1}{b-a}\,du$$

1. For $(a \leq \beta \leq b)$ the integral evaluates to $\boldsymbol{\frac{\beta - a}{b - a}}$.
2. For $(\beta \leq a)$, we have $F_X(\beta) = 0$.
3. For $(b \leq \beta)$, we have $F_X(\beta) = 1$.

PDF:



CDF:



Notice that the CDF is continuous everywhere, even though the pdf has discontinuities at a and b.

Relationship between PDF and CDF

Let X be a continuous random variable with pdf f and cdf F.

- The cdf is found by integrating the pdf: .

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

- By the Fundamental Theorem of Calculus, the pdf can be found by differentiating the cdf:

$$f(x) = \frac{d}{dx} F_X(x)$$

Computing probabilities

More generally, applying the fundamental theorem of calculus, we obtain:

$$\int_{\alpha}^{\beta} f_X(u)du = F_X(\beta) - F_X(\alpha)$$

$$= P(X \leq \beta) - P(X \leq \alpha)$$

$$= P(\alpha \leq X \leq \beta)$$

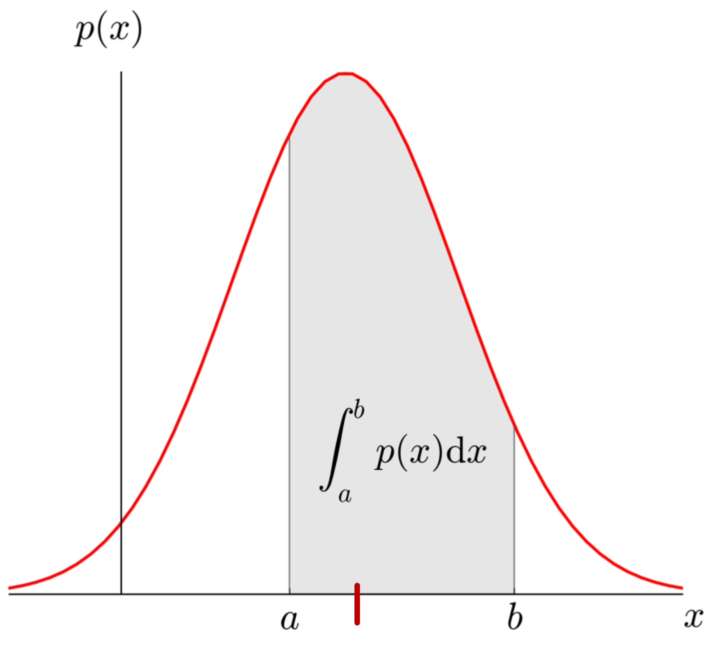
which gives

$$\boldsymbol{P(\alpha \leq X \leq \beta) = \int_{\alpha}^{\beta} f_X(u)du}$$

The probability that $\alpha \leq X \leq \beta$ is equal to the area under the pdf f_X between α and β.

In pictures:

- X takes on values in the continuum.
- $p(x)$, is a probability density function.



What happens when $a = b$?

Since f is continuous

$$\int_a^a f(u)du = 0$$

This leads to the possibly surprising result:

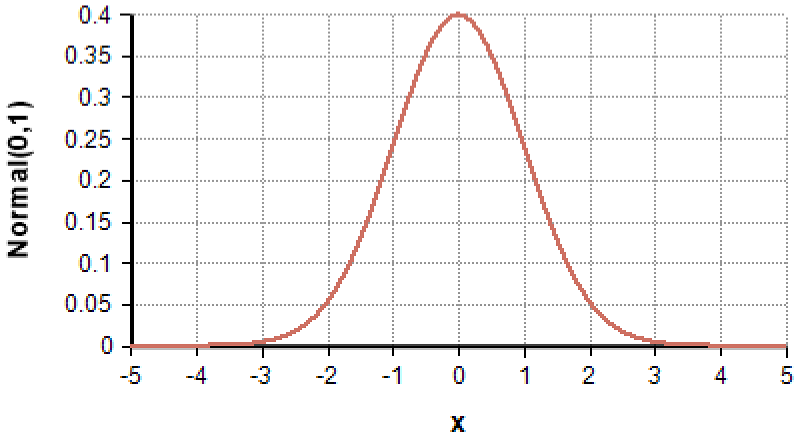
$$P(X = a) = 0$$

for any scalar a.

The Gaussian (aka normal) distribution

The Gaussian distribution is the most famous of all probability distributions, so famous that the Germans put Gauss and his pdf on their money!

You have likely seen or worked with the famous Bell Curve.



The Gaussian distribution

- The Gaussian has two defining parameters.
- The mean, μ
  - Defines the "location" of the pdf.
  - The pdf is symmetric about the mean.
- The variance, σ^2
  - Defines the "spread" of the pdf.
  - Can be specified also in terms of standard deviation, σ.
- The defining equation is given by:

$$f_X(x) = \boldsymbol{\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}} = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

The Gaussian distribution

Let's take a closer look:

- The leading term, $\frac{1}{\sigma\sqrt{2\pi}}$, is a normalizing term, so that $\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = 1$.

- So, let's simplify notation by writing

$$f_X(x) = Ke^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

The Gaussian distribution

Let's take a closer look:

- The leading term, $\frac{1}{\sigma\sqrt{2\pi}}$, is a normalizing term, so that $\int_{-\infty}^{\infty}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx = 1.$

- So, let's simplify notation by writing

$$f_X(x) = Ke^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

f_X is a decreasing exponential function.

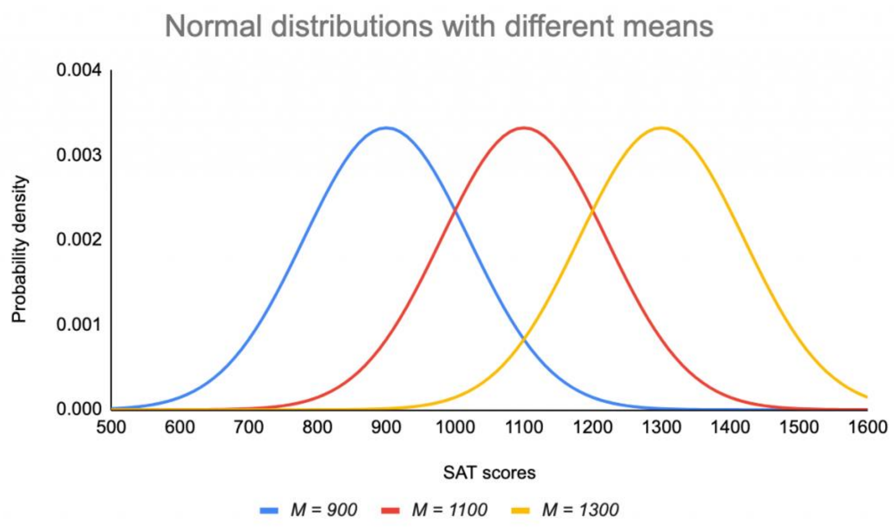
The Gaussian distribution

Let's take a closer look:

- The leading term, $\frac{1}{\sigma\sqrt{2\pi}}$, is a normalizing term, so that $\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = 1$.

- So, let's simplify notation by writing

$$f_X(x) = Ke^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$



f_X decreases exponentially with the square of the distance to the mean.

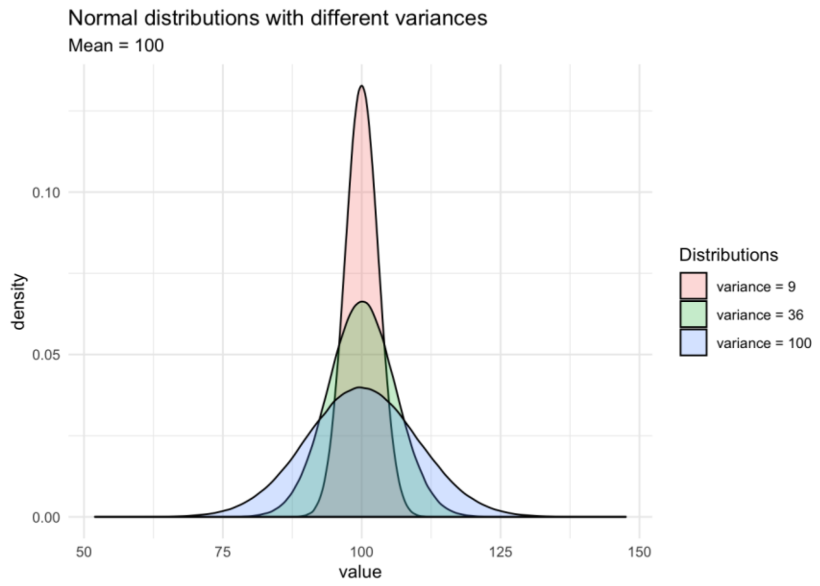
The Gaussian distribution

Let's take a closer look:

- The leading term, $\frac{1}{\sigma\sqrt{2\pi}}$, is a normalizing term, so that $\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = 1$.

- So, let's simplify notation by writing

$$f_X(x) = Ke^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$



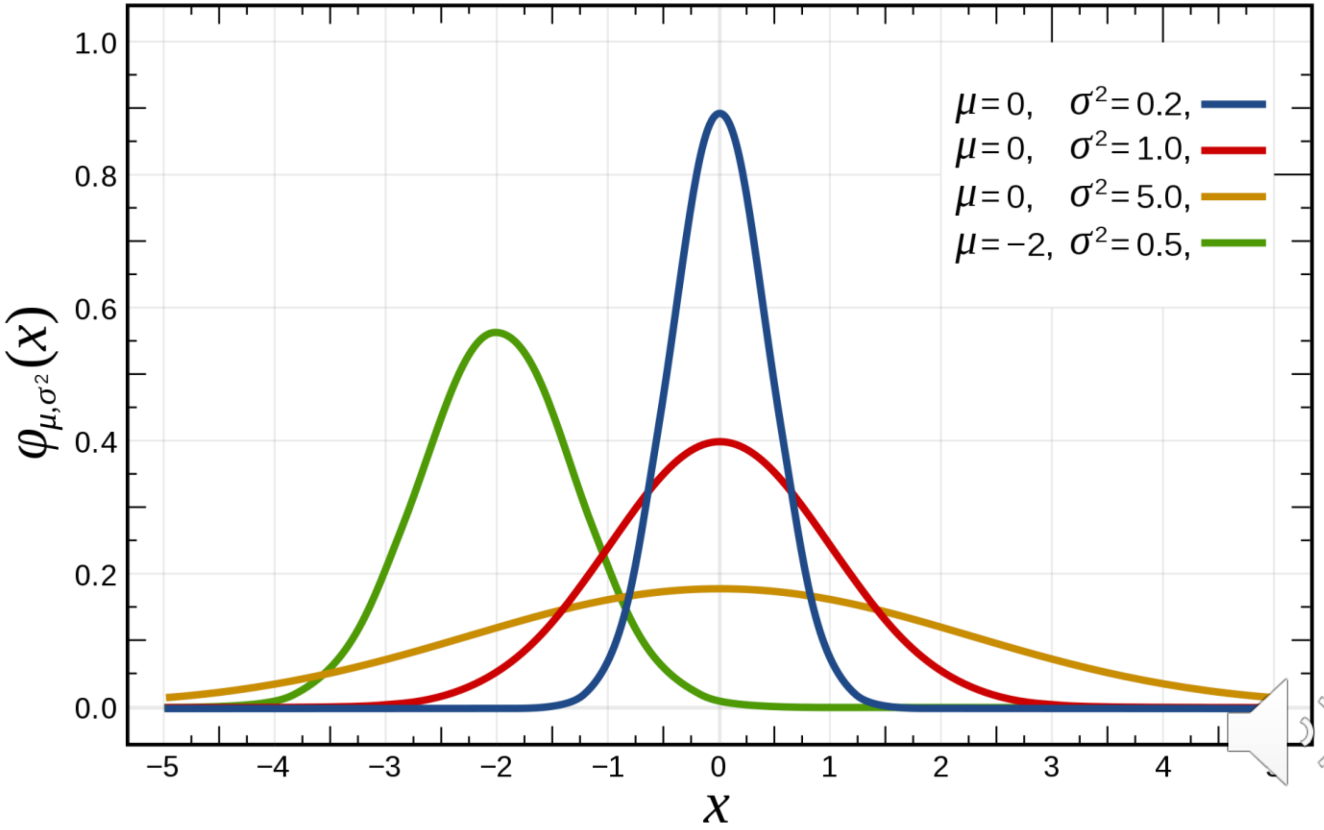
The "rate" of decrease depends on σ^2:

- ***If σ^2 is very large, f_X decreases slowly, thus, a wide spread.***
- ***If σ^2 is very small, f_X decreases quickly, thus, a narrow peak.***

The Gaussian Distribution

- Since the Gaussian is parameterized by its mean and variance, we often write $N(\mu, \sigma^2)$ to denote the Gaussian distribution.
- The special case when $\boldsymbol{\mu = 0, \sigma^2 = 1}$ is called the ***standard normal distribution*** (the red curve in the figure).

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



The Gaussian distribution

The standard deviation is a handy way to characterize probabilities.

- Approximately 68% of the probability mass lies within one standard deviation of the mean.
- Approximately 99.99966% of the probability mass lies within six standard deviations of the mean (for business majors, six sigma is a big thing).