

CSE 312

Foundations of Computing II

**Lecture 15: Expectation & Variance of Continuous RVs
Exponential and Normal Distributions**

Announcements

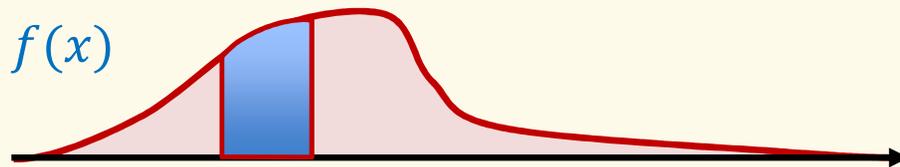
- See EdStem posts related to next week's midterm on May 4 in class:
 - Midterm General Information
 - Midterm Review (including Practice Midterm)
 - Practice Midterm and other Solutions
- Midterm Q&A session next Tuesday 4pm on Zoom

Review – Continuous RVs

Probability Density Function (PDF).

$f: \mathbb{R} \rightarrow \mathbb{R}$ s.t.

- $f(x) \geq 0$ for all $x \in \mathbb{R}$
- $\int_{-\infty}^{+\infty} f(x) dx = 1$



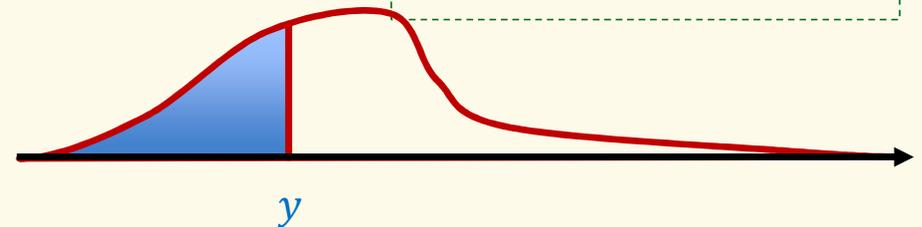
Density \neq Probability !

$$\begin{aligned} P(X \in [a, b]) &= \int_a^b f_X(x) dx \\ &= F_X(b) - F_X(a) \end{aligned}$$

Cumulative Distribution Function (CDF).

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

Theorem. $f(x) = \frac{dF(x)}{dx}$



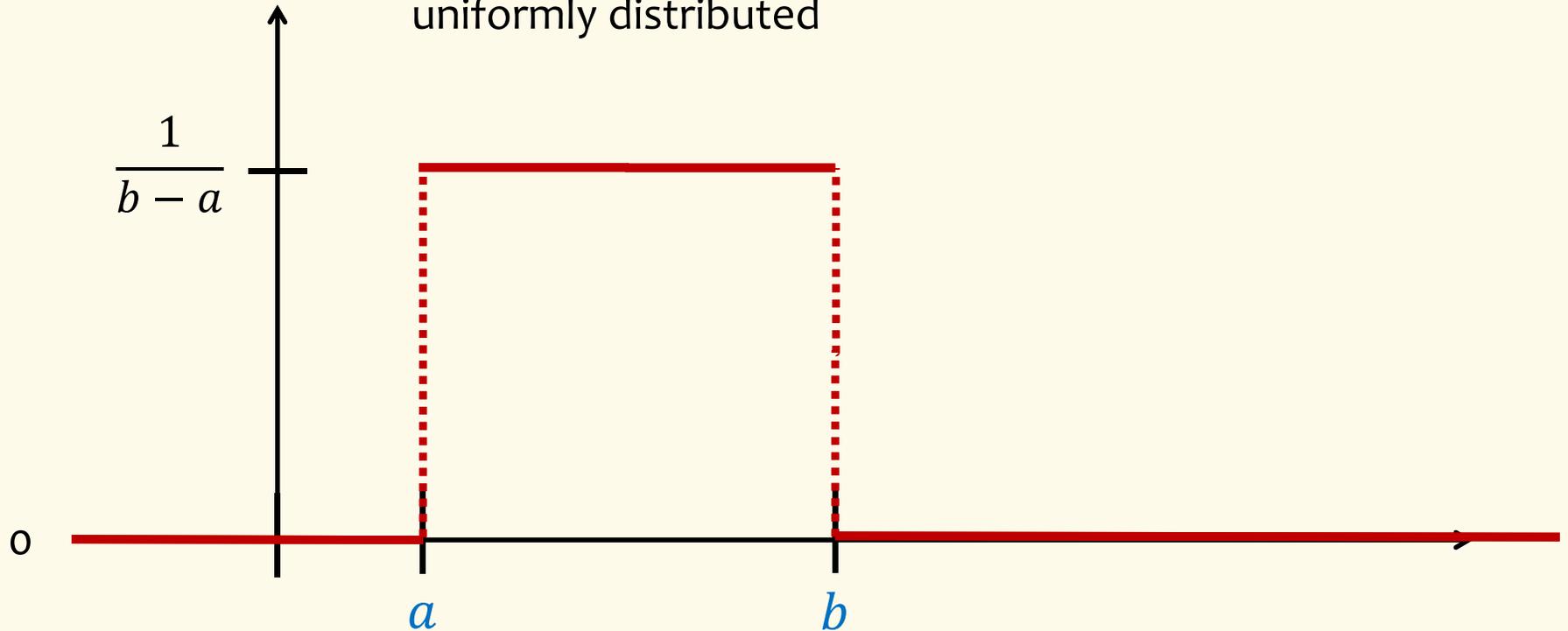
$$F_X(y) = P(X \leq y)$$

Review: Uniform Distribution

$$X \sim \text{Unif}(a, b)$$

We also say that X follows the uniform distribution / is uniformly distributed

$$f_X(x) = \begin{cases} \frac{1}{b-a} & x \in [a, b] \\ 0 & \text{else} \end{cases}$$



Review: From Discrete to Continuous

	Discrete	Continuous
PMF/PDF	$p_X(x) = P(X = x)$	$f_X(x) \neq P(X = x) = 0$
CDF	$F_X(x) = \sum_{t \leq x} p_X(t)$	$F_X(x) = \int_{-\infty}^x f_X(t) dt$
Normalization	$\sum_x p_X(x) = 1$	$\int_{-\infty}^{\infty} f_X(x) dx = 1$
Expectation	$\mathbb{E}[g(X)] = \sum_x g(x) p_X(x)$	$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx$

Expectation of a Continuous RV

Definition. The **expected value** of a continuous RV X is defined as

$$\mathbb{E}[X] = \int_{-\infty}^{+\infty} f_X(x) \cdot x \, dx$$

Fact. $\mathbb{E}[aX + bY + c] = a\mathbb{E}[X] + b\mathbb{E}[Y] + c$

Proofs follow same ideas as discrete case

Definition. The **variance** of a continuous RV X is defined as

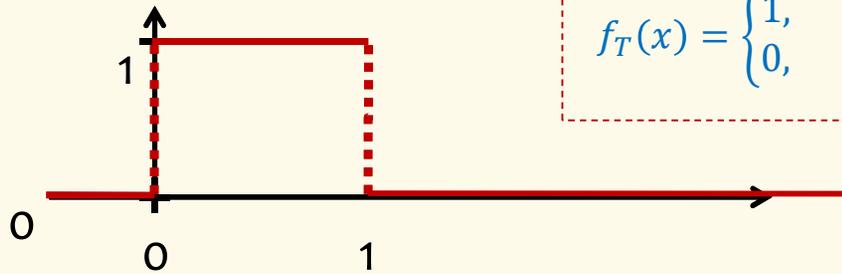
$$\text{Var}(X) = \int_{-\infty}^{+\infty} f_X(x) \cdot (x - \mathbb{E}[X])^2 \, dx = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

Agenda

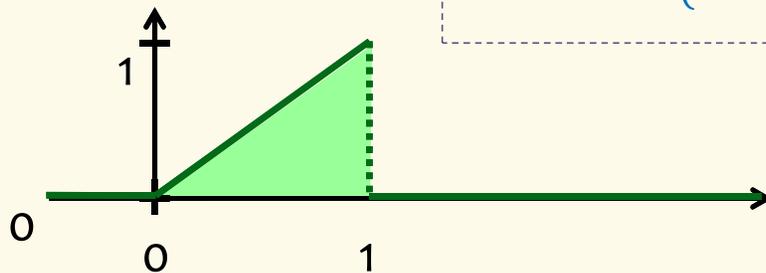
- Uniform Distribution ◀
- Exponential Distribution
- Normal Distribution

Expectation of a Continuous RV

Example. $T \sim \text{Unif}(0,1)$



$$f_T(x) = \begin{cases} 1, & x \in [0,1] \\ 0, & x \notin [0,1] \end{cases}$$



$$f_T(x) \cdot x = \begin{cases} x, & x \in [0,1] \\ 0, & x \notin [0,1] \end{cases}$$

Definition.

$$\mathbb{E}[X] = \int_{-\infty}^{+\infty} f_X(x) \cdot x \, dx$$

$$\mathbb{E}[T] = \underbrace{\frac{1}{2} 1^2}_{\text{Area of triangle}} = \frac{1}{2}$$

Area of triangle

Uniform Density – Expectation

$X \sim \text{Unif}(a, b)$

$$f_X(x) = \begin{cases} \frac{1}{b-a} & x \in [a, b] \\ 0 & \text{else} \end{cases}$$

$$\begin{aligned} \mathbb{E}[X] &= \int_{-\infty}^{+\infty} f_X(x) \cdot x \, dx \\ &= \frac{1}{b-a} \int_a^b x \, dx = \frac{1}{b-a} \left(\frac{x^2}{2} \right) \Big|_a^b = \frac{1}{b-a} \left(\frac{b^2 - a^2}{2} \right) \\ &= \frac{(b-a)(a+b)}{2(b-a)} = \frac{a+b}{2} \end{aligned}$$

Uniform Density – Variance

$$X \sim \text{Unif}(a, b)$$

$$f_X(x) = \begin{cases} \frac{1}{b-a} & x \in [a, b] \\ 0 & \text{else} \end{cases}$$

$$\mathbb{E}[X^2] = \int_{-\infty}^{+\infty} f_X(x) \cdot x^2 \, dx$$

$$= \frac{1}{b-a} \int_a^b x^2 \, dx = \frac{1}{b-a} \left(\frac{x^3}{3} \right) \Big|_a^b = \frac{b^3 - a^3}{3(b-a)}$$

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

Uniform Density – Variance

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

$$X \sim \text{Unif}(a, b)$$

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

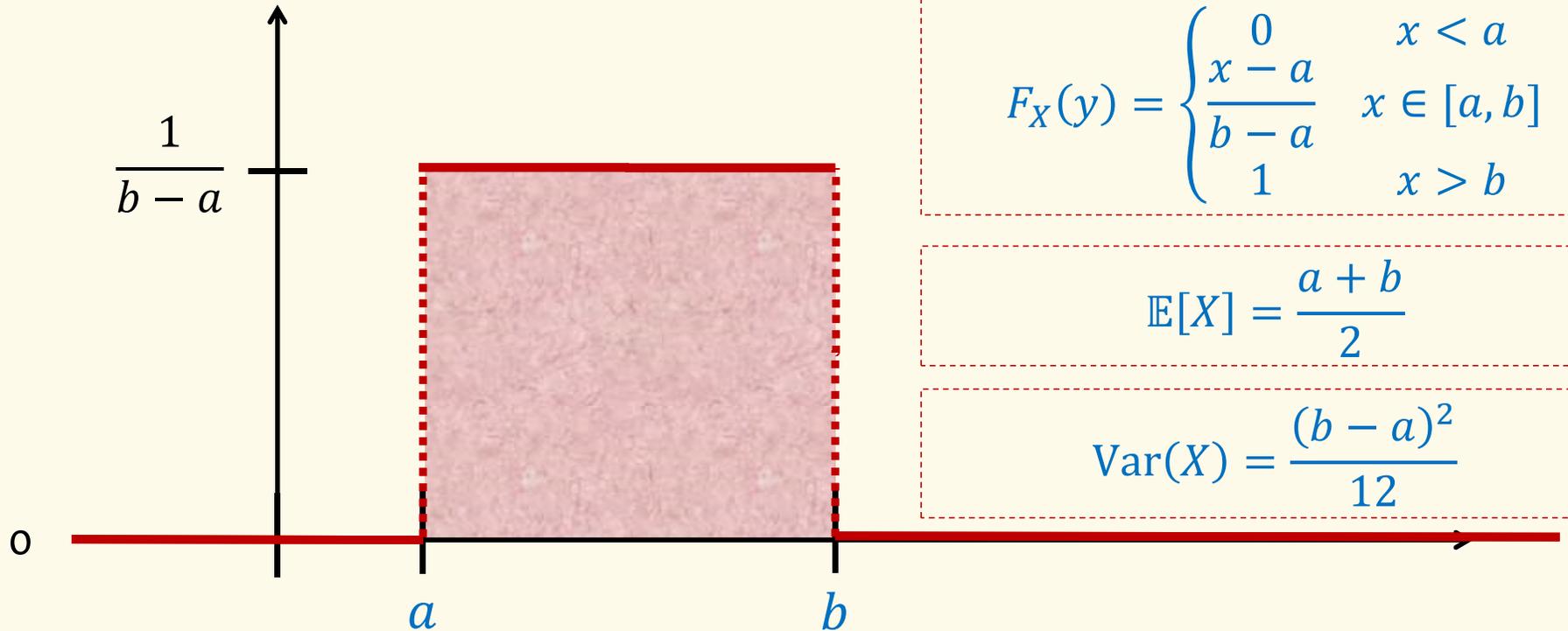
$$= \frac{b^2 + ab + a^2}{3} - \frac{a^2 + 2ab + b^2}{4}$$

$$= \frac{4b^2 + 4ab + 4a^2}{12} - \frac{3a^2 + 6ab + 3b^2}{12}$$

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

Uniform Distribution Summary

$X \sim \text{Unif}(a, b)$



$$f_X(x) = \begin{cases} \frac{1}{b-a} & x \in [a, b] \\ 0 & \text{else} \end{cases}$$

$$F_X(y) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & x \in [a, b] \\ 1 & x > b \end{cases}$$

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

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

Agenda

- Uniform Distribution
- Exponential Distribution ◀
- Normal Distribution

Exponential Density

Assume expected # of occurrences of an event per unit of time is λ (independently)

- Cars going through intersection
- Number of lightning strikes
- Requests to web server
- Patients admitted to ER
- Rate of radioactive decay

Numbers of occurrences of event: Poisson distribution

$$P(X = i) = e^{-\lambda} \frac{\lambda^i}{i!} \quad (\text{Discrete})$$

How long to wait until next event? Exponential density!

Let's define it and then derive it!

Exponential Density - Warmup

$$X \sim \text{Poi}(\lambda) \Rightarrow P(X = i) = e^{-\lambda} \frac{\lambda^i}{i!}$$

Assume expected # of occurrences of an event per unit of time is λ (independently)

What is the distribution of $Z = \#$ occurrences of event per t units of time?

$$\mathbb{E}[Z] = t\lambda$$

Z is independent over disjoint intervals

$$\text{so } Z \sim \text{Poi}(t\lambda)$$

$$X \sim Poi(\lambda) \Rightarrow P(X = i) = e^{-\lambda} \frac{\lambda^i}{i!}$$

The Exponential PDF/CDF

Assume expected # of occurrences of an event per unit of time is λ (independently)

Numbers of occurrences of event: Poisson distribution

How long to wait until next event? Exponential density!

- The exponential RV has range $[0, \infty]$, unlike Poisson with range $\{0, 1, 2, \dots\}$
- Let $Y \sim Exp(\lambda)$ be the time till the first event. We will compute $F_Y(t)$ and $f_Y(t)$
- Let $Z \sim Poi(t\lambda)$ be the # of events in the first t units of time, for $t \geq 0$.
- $P(Y > t) = P(\text{no event in the first } t \text{ units}) = P(Z = 0) = e^{-t\lambda} \frac{(t\lambda)^0}{0!} = e^{-t\lambda}$
- $F_Y(t) = P(Y \leq t) = 1 - P(Y > t) = 1 - e^{-t\lambda}$
- $f_Y(t) = \frac{d}{dt} F_Y(t) = \lambda e^{-t\lambda}$

$$P(X > t) = e^{-t\lambda}$$

Exponential Distribution

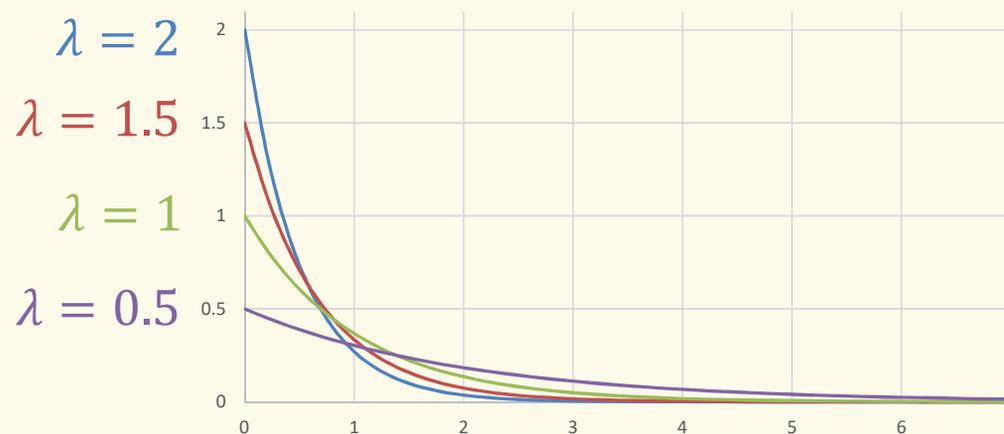
Definition. An **exponential random variable** X with parameter $\lambda \geq 0$ is follows the exponential density

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

We write $X \sim \text{Exp}(\lambda)$ and say X that follows the exponential distribution.

CDF: For $y \geq 0$,

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



Expectation

$$\begin{aligned}\mathbb{E}[X] &= \int_{-\infty}^{+\infty} f_X(x) \cdot x \, dx \\ &= \int_0^{+\infty} \lambda e^{-\lambda x} \cdot x \, dx \\ &= \left(-\left(x + \frac{1}{\lambda}\right) e^{-\lambda x} \right) \Big|_0^{\infty} = \frac{1}{\lambda}\end{aligned}$$

Somewhat complex calculation
use integral by parts

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

$$P(X > t) = e^{-t\lambda}$$

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

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



Memorylessness

Definition. A random variable is **memoryless** if for all $s, t > 0$,

$$P(X > s + t \mid X > s) = P(X > t).$$

Fact. $X \sim \text{Exp}(\lambda)$ is memoryless.

Assuming an exponential distribution, if you've waited s minutes, The probability of waiting t more is exactly same as when $s = 0$.

Memorylessness of Exponential

$$P(X > t) = e^{-\lambda t}$$

Proof that assuming exp distr, if you've waited s minutes, prob of waiting t more is exactly same as when $s = 0$

Fact. $X \sim \text{Exp}(\lambda)$ is memoryless.

Proof.

$$\begin{aligned} P(X > s + t \mid X > s) &= \frac{P(\{X > s + t\} \cap \{X > s\})}{P(X > s)} \\ &= \frac{P(X > s + t)}{P(X > s)} \\ &= \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = e^{-\lambda t} = P(X > t) \end{aligned}$$

The only memoryless RVs are the geometric RV (discrete) and Exp RV (continuous)

Example

- Time it takes to check someone out at a grocery store is exponential with an expected value of 10 mins.
- Independent for different customers
- If you are the second person in line, what is the probability that you will have to wait between 10 and 20 mins?

$$T \sim \text{Exp}\left(\frac{1}{10}\right)$$

$$P(10 \leq T \leq 20) = \int_{10}^{20} \frac{1}{10} e^{-\frac{x}{10}} dx$$

$$y = \frac{x}{10} \text{ so } dy = \frac{dx}{10}$$

$$P(10 \leq T \leq 20) = \int_1^2 e^{-y} dy = -e^{-y} \Big|_1^2 = e^{-1} - e^{-2}$$

Example

- Time it takes to check someone out at a grocery store is exponential with an expected value of 10 mins.
- Independent for different customers
- If you are the second person in line, what is the probability that you will have to wait between 10 and 20 mins?

$$T \sim \text{Exp}\left(\frac{1}{10}\right)$$

$$\text{so } F_T(t) = 1 - e^{-\frac{t}{10}}$$

$$\begin{aligned} P(10 \leq T \leq 20) &= F_T(20) - F_T(10) \\ &= 1 - e^{-\frac{20}{10}} - \left(1 - e^{-\frac{10}{10}}\right) = e^{-1} - e^{-2} \end{aligned}$$

Agenda

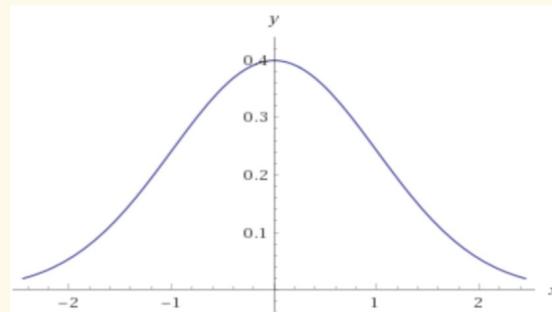
- Uniform Distribution
- Exponential Distribution
- Normal Distribution ◀

The Normal Distribution

Definition. A **Gaussian (or normal)** random variable with parameters $\mu \in \mathbb{R}$ and $\sigma \geq 0$ has density

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

We say that X follows the Normal Distribution, and write $X \sim \mathcal{N}(\mu, \sigma^2)$.



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



Carl Friedrich
Gauss

The Normal Distribution

Definition. A **Gaussian (or normal) random variable** with parameters $\mu \in \mathbb{R}$ and $\sigma \geq 0$ has density

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

We say that X follows the Normal Distribution, and write $X \sim \mathcal{N}(\mu, \sigma^2)$.

Fact. If $X \sim \mathcal{N}(\mu, \sigma^2)$, then $\mathbb{E}[X] = \mu$, and $\text{Var}(X) = \sigma^2$

Proof of expectation is easy because density curve is symmetric around μ ,

$f_X(\mu - x) = f_X(\mu + x)$, but proof for variance requires integration of $e^{-x^2/2}$

We will see next time why the normal distribution is (in some sense) the most important distribution.



Carl Friedrich
Gauss

The Normal Distribution

Aka a “Bell Curve” (imprecise name)

