# Convex Functions and Jensen's Inequality

Andrew Nobel

September, 2024

# Convex Functions

**Definition:** Let $I$ be an interval in $\mathbb{R}$. A function $f : I \to \mathbb{R}$ is *convex* if for every $x, y \in I$ and every $\alpha \in [0, 1]$,

$$f(\alpha x + (1-\alpha)y) \;\leq\; \alpha f(x) + (1-\alpha) f(y)$$

If the inequality is strict when $\alpha \neq 0, 1$ then $f$ is called strictly convex

**Note:** The set of points $\{\alpha x + (1-\alpha)y : \alpha \in [0,1]\}$ is a line segment (interval) between $x$ and $y$. If $x, y$ belong to $I$ then so do all the points $\alpha x + (1-\alpha)y$

**Picture**

- Graph of $f$ is curved upwards or linear
- For each $x, y \in I$ the line connecting $(x, f(x))$ and $(y, f(y))$ lies *above* the graph of $f$

# Concave Functions

**Definition:** Let $I$ be an interval in $\mathbb{R}$. A function $f : I \to \mathbb{R}$ is *concave* if for every $x, y \in I$ and every $\alpha \in [0, 1]$,

$$f(\alpha x + (1 - \alpha)y) \;\geq\; \alpha f(x) + (1 - \alpha)f(y)$$

If the inequality is strict when $\alpha \neq 0, 1$ then $f$ is called strictly concave

**Picture**

- Graph of $f$ is curved downwards or linear
- For each $x, y \in I$ the line connecting $(x, f(x))$ and $(y, f(y))$ lies *below* the graph of $f$

**Fact:** A function $f$ is concave if and only if $-f$ is convex

# Establishing Convexity and Concavity

1. **Check the definition:** In many cases it is possible to check the definition directly

2. **Second derivative condition**

   - A function $f : I \to \mathbb{R}$ is convex if $f'' \geq 0$ and strictly convex if $f'' > 0$
   - A function $f : I \to \mathbb{R}$ is concave if $f'' \leq 0$ and strictly convex if $f'' < 0$

# Examples of Convex/Concave Functions

1. $|x|$ is convex, but *not* strictly convex on $I = \mathbb{R}$

2. $x^2$, $e^x$, and $e^{-x}$ are strictly convex on $I = \mathbb{R}$

3. $x^{-1}$ and $x \log x$ are strictly convex on $I = (0, \infty)$

4. $\log x$ and $\sqrt{x}$ are strictly concave on $I = (0, \infty)$

5. Function $f(x) = e^{-x^2/2}$ is
   - strictly concave on $I = (-1, 1)$
   - strictly convex on $I = (-\infty, 1)$ and $I = (1, \infty)$
   - neither convex nor concave on $I = \mathbb{R}$

6. Linear function $f(x) = ax + b$ is convex and concave

# Basic Properties of Convex Functions

**Fact:** Let $I \subseteq \mathbb{R}$ be an interval

1. If $f_1, \ldots, f_m$ are convex on $I$ then $f(x) = \max_{1 \leq j \leq m} f_j(x)$ is convex
2. If $f_1, \ldots, f_m$ are convex on $I$ and $a_1, \ldots, a_m \geq 0$ then $f = \sum_{i=1}^{m} a_i f_i$ is convex
3. If $f$ is convex and $g : \mathbb{R} \to \mathbb{R}$ is convex and non-decreasing, then $g \circ f$ is convex
4. If $f$ is convex on $I$ it is continuous except possibly at the endpoints of $I$

**Note:** Corresponding properties for concave functions follow from the fact that $f$ is concave if and only if $-f$ is convex

# Jensen’s Inequality

# Jensen's Inequality

**Jensen's Inequality:** If a random variable $X \in I$ then $\mathbb{E}X \in I$ and the following inequalities hold

(a) If $f : I \to \mathbb{R}$ is convex then $f(\mathbb{E}X) \leq \mathbb{E}f(X)$.

(b) If $f : I \to \mathbb{R}$ is concave then $f(\mathbb{E}X) \geq \mathbb{E}f(X)$.

**Idea:** In the special case where $X$ is defined by

$$X = \begin{cases} x & \text{with probability } \alpha \\ y & \text{with probability } 1 - \alpha \end{cases}$$

for two points $x, y \in I$, inequality (a) coincides with the definition of convexity. Jensen's inequality extends this to general random variables $X$ with values in $I$

# Applications of Jensen's Inequality

**Immediate Applications**

- $\mathbb{E}X^2 \geq (\mathbb{E}X)^2$
- $\mathbb{E}e^X \geq e^{\mathbb{E}X}$ and $\mathbb{E}e^{-X} \geq e^{-\mathbb{E}X}$
- $\mathbb{E}(X \log X) \geq (\mathbb{E}X) \log(\mathbb{E}X)$ if $X > 0$
- $\mathbb{E} \log X \leq \log \mathbb{E}X$ if $X > 0$
- $\mathbb{E}\sqrt{X} \leq \sqrt{\mathbb{E}X}$ if $X \geq 0$

**Fact:** Arithmetic-Geometric mean inequality. If $a_1, \ldots, a_n$ are positive then

$$(\prod_{i=1}^{n} a_i)^{1/n} \leq n^{-1} \sum_{i=1}^{n} a_i$$