Jensen's Inequality

Intuition and Proof

January 20, 2023

Goals

1. Formally state Jensen's inequality.
2. Provide intuition on the direction of the inequality.
3. Show how to prove it by representing an arbitrary convex function in terms of linear functions.

Theorem (Jensen's Inequality)

Let g be a convex function from I to $\mathbb{R}$, where I is an open interval of reals. Let X be random variable on $(\Omega, \mathcal{F}, P)$, with $X(\omega) \in I$ for all ω. Assume $E[X]$ to be finite. If $\mathcal{H}$ is a sub σ-field of $\mathcal{F}$, then

$$\mathbb{E}[g(X) \mid \mathcal{H}] \geq g(\mathbb{E}[X \mid \mathcal{H}]) \quad a.e. \tag{1}$$

In particular, $\mathbb{E}[g(X)] \geq g(\mathbb{E}[X])$.

Theorem (Jensen's Inequality)

Let g be a convex function from I to $\mathbb{R}$, where I is an open interval of reals. Let X be random variable on $(\Omega, \mathcal{F}, P)$, with $X(\omega) \in I$ for all ω. Assume $E[X]$ to be finite. If $\mathcal{H}$ is a sub σ-field of $\mathcal{F}$, then

$$\mathbb{E}[g(X) \mid \mathcal{H}] \geq g(\mathbb{E}[X \mid \mathcal{H}]) \quad a.e. \tag{1}$$

In particular, $\mathbb{E}[g(X)] \geq g(\mathbb{E}[X])$.

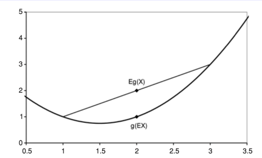
How to recall the direction of the inequality [Durrett, 2010]

Take $P(X = x) = \lambda$, $P(X = y) = 1 - \lambda$.

$$\begin{aligned} \mathbb{E}[g(X)] &= \lambda\, g(x) + (1-\lambda)\, g(y) && \text{LOTUS} \\ &\geq g\big(\lambda x + (1-\lambda) y\big) && \text{def. convexity} \\ &= g\big(\mathbb{E}[X]\big). && \text{def. expectation} \end{aligned}$$

For example:

$P(X = 1) = P(X = 3) = .5$



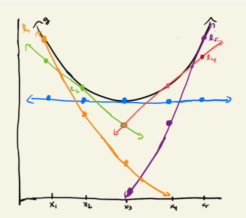
Line of Support Theorem.

Let $g : I \to \mathbb{R}$, where I is an open interval of reals, bounded or unbounded. Assume g is convex, that is,

$$\underbrace{g\left(\alpha x + (1-\alpha)y\right)}_{\text{graph}} \leq \underbrace{\alpha g(x) + (1-\alpha)g(y)}_{\text{chord}}$$

for all $x, y \in I$ and all $\alpha \in [0,1]$. Then there are sequences $\{a_n\}, \{b_n\}$ of real numbers such that for all $x \in I$,

$$g(x) = \sup_n \left(a_n x + b_n\right)$$



Proof (partial)

Here we show that Jensen's inequality (Eq. (1)) holds.

$$
\begin{aligned}
& g(X) = \sup_n \, a_n X + b_n && \text{Line of Support Theorem (Theorem 3)} \\
\implies & g(X) \geq a_n X + b_n && \text{supremum is an upper bound} \\
\implies & \mathbb{E}[g(X) \mid \mathcal{H}] \geq \mathbb{E}[a_n X + b_n \mid \mathcal{H}] \quad \text{a.e.} && \text{monotonicity} \\
\implies & \mathbb{E}[g(X) \mid \mathcal{H}] \geq a_n \, \mathbb{E}[X \mid \mathcal{H}] + b_n \quad \text{a.e.} && \text{linearity} \\
\implies & \mathbb{E}[g(X) \mid \mathcal{H}] \geq \underbrace{\sup_n \{a_n \, \mathbb{E}[X \mid \mathcal{H}] + b_n\}}_{= g(\mathbb{E}[X \mid \mathcal{H}]) \text{ by LoS Thm.}} \quad \text{a.e.} && \text{supremum is least upper bound}
\end{aligned}
$$

The Line of Support Theorem allows us to take an arbitrary convex function and represent it in terms of linear functions (and therefore apply properties that hold under linearity).