

Jensen's Inequality

Jensen's inequality is one of the fundamental tools very frequently used in information theory.

A function $f(x)$ is said to be *convex* over an interval (a,b) , if for every $x_1, x_2 \in (a,b)$ and $0.0 \leq \lambda \leq 1.0$,

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2). \quad (1.1)$$

It is said *strictly convex* if the equality holds only if $\lambda = 1.0$ or $\lambda = 0.0$.

A function f is concave if $-f$ is convex.

Examples of convex functions include x^2 and $|x|$.

Examples of concave functions include $\log(x), \sqrt{x}$ over $x \geq 0$.