$X : \Omega \to \mathbb{R}$ is a random variable where $|\Omega| < \infty$;

The Shannon entropy of $X$ is defined to be

$$-\sum_{x \in \mathrm{range}(X)} p_X(x) \log_2 p_X(x), \quad \text{where } 0 \log 0 = 0.$$

We showed symmetry of information: $I(X;Y) = I(Y;X)$, where $I(X;Y) = H(X) - H(X|Y)$.

Two basic tools for dealing with Shannon information are

(1) Jensen's inequality

(2) Fundamental Inequality.

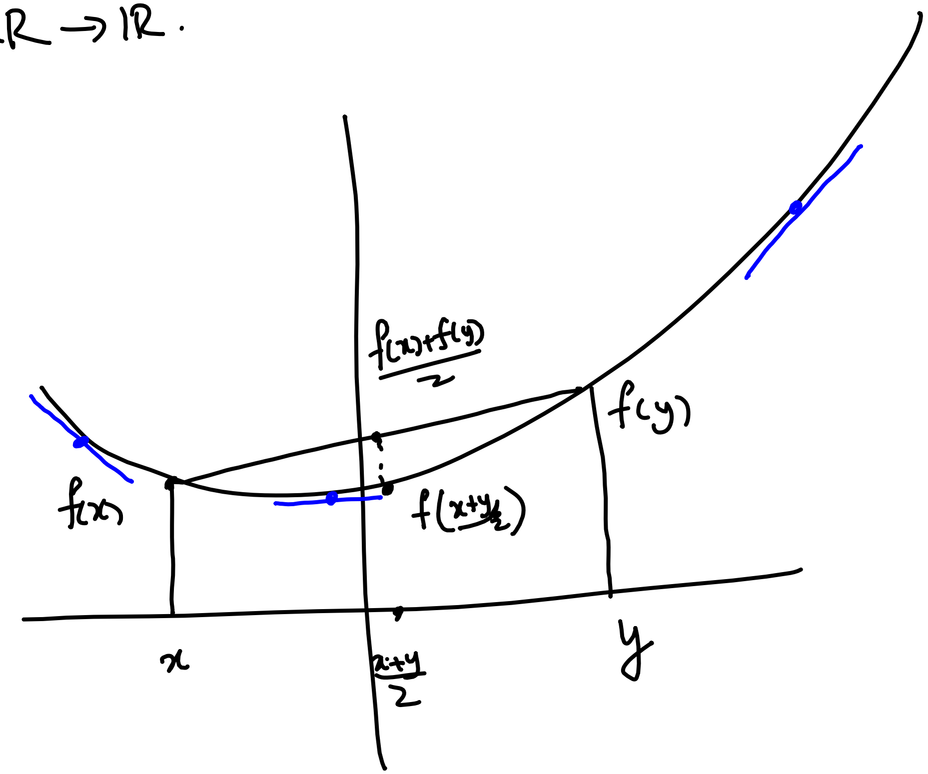
# Jensen's Inequality

Deals with convex functions $f: \mathbb{R} \to \mathbb{R}$.

$f$ is convex if

$\forall \lambda \in [0,1] \quad \forall x, y \in \mathbb{R}$

$$f(\lambda x + (1-\lambda) y) \leq \lambda f(x) + (1-\lambda) f(y)$$




e.g. $\lambda = 1/2$

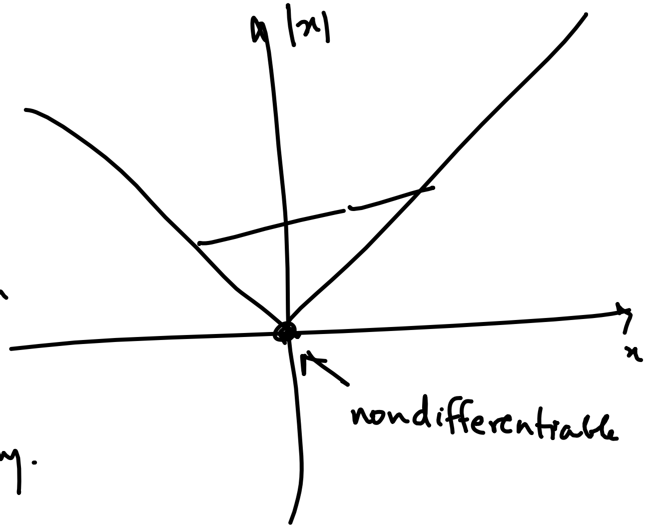
$$f\left(\frac{x+y}{2}\right) \leq \frac{f(x)+f(y)}{2}$$

## Note:

A sufficient condition for convexity of $f: \mathbb{R} \to \mathbb{R}$, if $f$ is twice differentiable is:

$$\frac{d^2 f}{dx^2} > 0.$$

|x| is continuous & convex even though second derivative test does not apply.



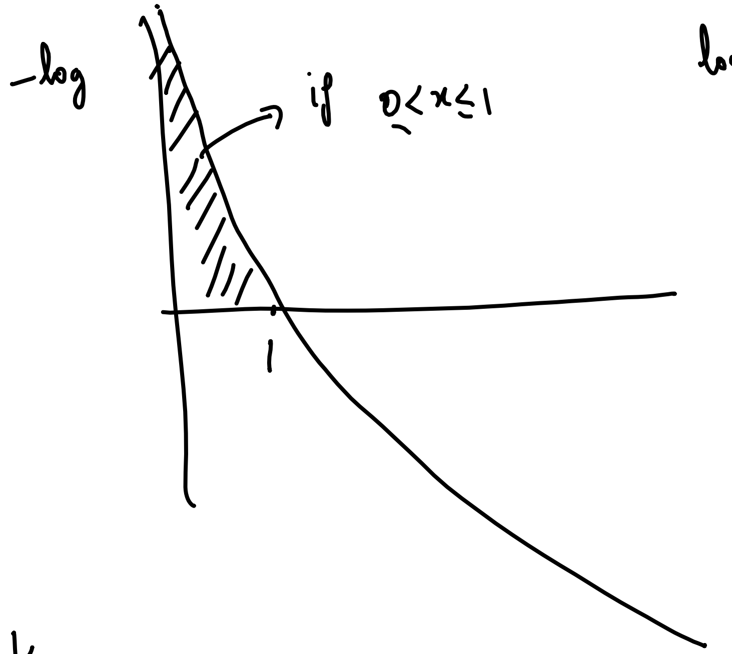

## Jensen's inequality

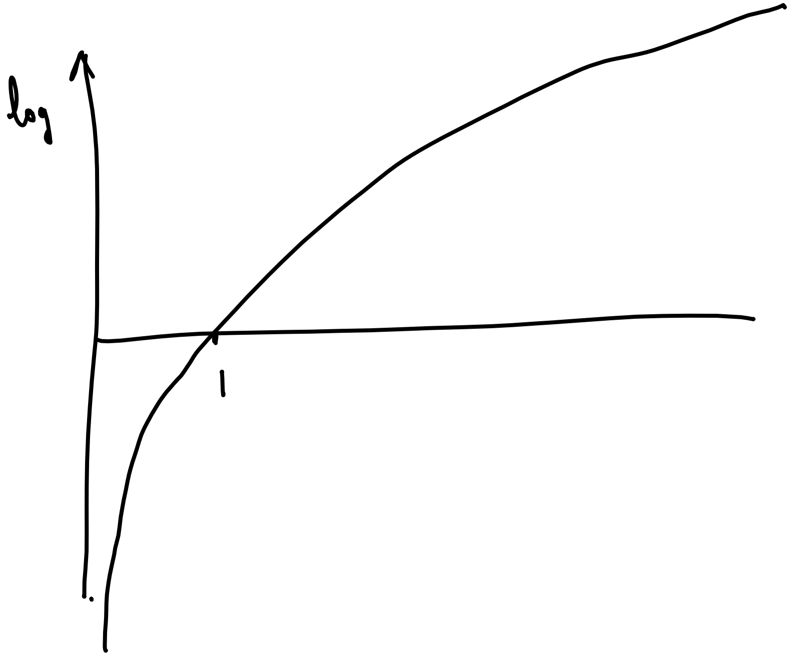
For every random variable $X: \Omega \to \mathbb{R}$, $|\Omega| < \infty$, and every continuous, convex function $f: \mathbb{R} \to \mathbb{R}$, we have

$$f(\mathbb{E} X) \leq \mathbb{E}(f(X)).$$

## Application

Consider $f: (0,\infty) \to \mathbb{R}$ defined by $f(x) = -\ln(x)$.




$$\frac{d}{dx}(-\ln x) = -1/x$$

$$\frac{d^2}{dx}(-\ln x) = \frac{d}{dx}(-1/x) = +1/x^2 \quad \& \quad \forall x \in \mathbb{R},$$

$1/x^2 > 0$. $\Rightarrow -\ln x$ is convex.

Similarly $-\log_2(x)$ is convex.

By Jensen's inequality for log

$$H(X) = \sum_{x \in \text{range}(X)} p(x) \cdot \log_2 \frac{1}{p(x)} \leq \log\left(\sum_{x \in \text{range}(X)} p(x) \cdot \frac{1}{p(x)}\right) = \log\left(\sum_{x \in \text{range}(X)} 1.\right)$$

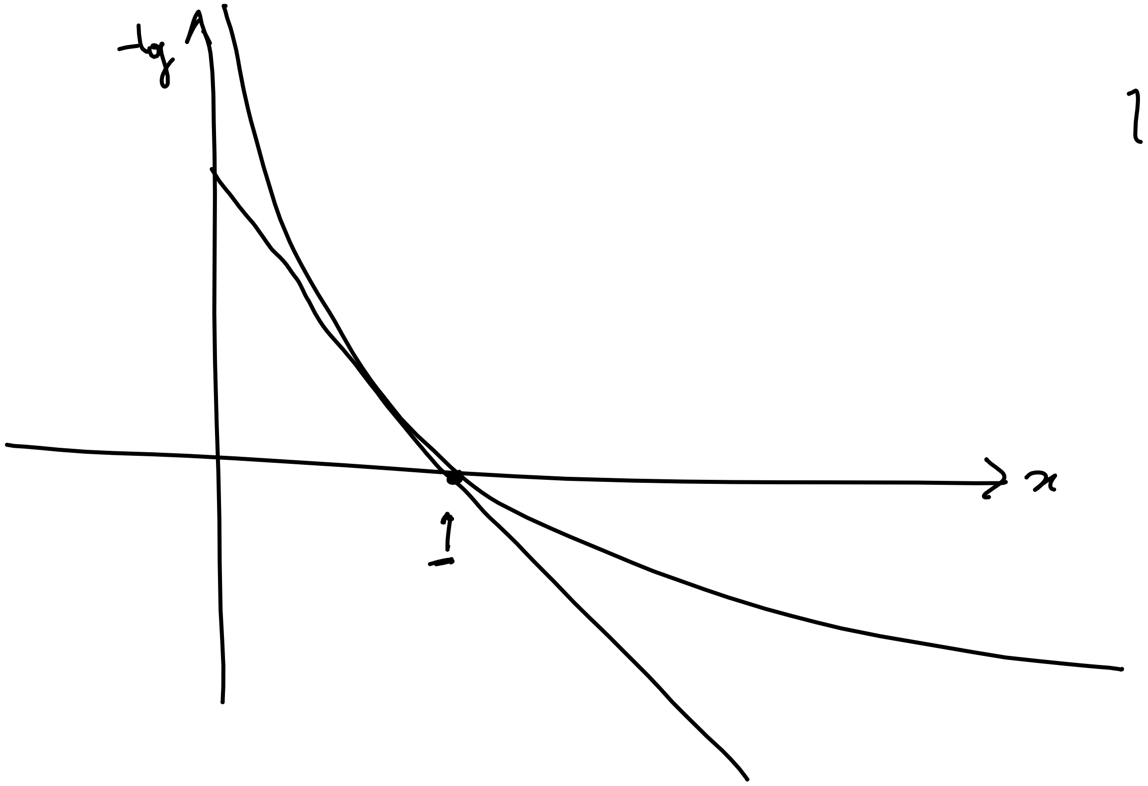
$$= \log(|\Omega|).$$

For example, if $|\Omega| = n$, then $H(X) \leq \log_2(n)$.

The uniform distribution, i.e. $p(x) = 1/n$ for all $x \in \Omega$, $|\Omega| = n$, has:

$$H(X) = \sum_{x \in \text{range}(X)} p(x) \cdot \log_2\left(\frac{1}{p(x)}\right) = \sum_{x \in \text{range}(X)} 1/n \cdot \log_2(n) = \log(n) \times n \times 1/n = \log(n).$$

## Fundamental Inequality (only for $-\log$)




"Tangent is below a convex curve"

In particular, for $\log_2(x)$,

the tangent at $1$ is below the curve.

$$(1-x) \leq -\log_2 x.$$

## Application of fundamental inequality

$$H(p) = \sum_{x \in range(X)} p(x) \log \left( \frac{1}{p(x)} \right)$$

$$\geq \sum_{i=0}^{n-1} p_i \left[ 1 - p_i \right]$$

Both $p_i$ & $1 - p_i$ are $\geq 0$. Hence their sum $\geq 0$.

$$\geq 0.$$