6.6 Jensens Lemma

The following theorem will be used to prove some more interesting statements about entropy functions. First, a preliminary definition.

Definition 6.14 *A real function f is said to be concave on an interval I if* $\forall_{x,y \in I}$

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

Theorem 6.15 (Jensens inequality) *If f is continuous and strictly concave on I, and $\sum_{i=1}^{n} a_i = 1,\ a_i \geq 0,\ 1 \leq i \leq n$, then,* $\forall_{x_i \in I}$

$$\sum_{i=1}^{n} a_i f(x_i) \leq f(\sum_{i=1}^{n} a_i x_i)$$

with equality iff $x_1 = x_2 = ... = x_n$

6.6.1 Application

Theorem 6.16 *if X takes values x_1 with probability p_1, x_2 with probability p_2, ..., x_n with probability p_n, then*

$$\mathbf{H}\left[X\right] \leq \lg\ n$$

(with equality when $p_1 = p_2 = ...p_n = 1/n$)

Proof.

$$\begin{aligned} \mathbf{H}\left[X\right] &= -\sum_{i=1}^{n} p_i \lg\ p_i \\ &= \sum_{i=1}^{n} p_i \lg\ 1/p_i \\ &\leq \lg\ \sum_{i=1}^{n} p_i 1/p_i \\ &= \lg\ n \ \ (eq.\ when\ p_1 = p_2 = ... = p_n = 1/n) \end{aligned}$$