

# Jensen's Inequality

The following theorem will be very useful to derive basic properties of entropy.

## Theorem: Jensen's inequality

Let  $f : \mathcal{D} \rightarrow \mathbb{R}$  be a convex function, and let  $n \in \mathbb{N}$ . Then for any  $p_1, \dots, p_n \in \mathbb{R}_{\geq 0}$  such that  $\sum_{i=1}^n p_i = 1$  and for any  $x_1, \dots, x_n \in \mathcal{D}$  it holds that

$$\sum_{i=1}^n p_i f(x_i) \geq f\left(\sum_{i=1}^n p_i x_i\right).$$

If  $f$  is strictly convex and  $p_1, \dots, p_n > 0$ , then equality holds if and only if  $x_1 = \dots = x_n$ . In particular, if  $X$  is a real random variable whose image  $\mathcal{X}$  is contained in  $\mathcal{D}$ , then

$$\mathbb{E}[f(X)] \geq f(\mathbb{E}[X]),$$

and if  $f$  is strictly convex, equality holds if and only if there is a  $c \in \mathcal{X}$  such that  $X = c$  with probability 1.

### Proof

The proof is by induction. The case  $n = 1$  is trivial, and the case  $n = 2$  is identical to the very definition of convexity. Suppose that we have already proved the claim up to  $n - 1 \geq 2$ . Assume, **without loss of generality**, that  $p_n < 1$ . Then:

$$\begin{aligned} \sum_{i=1}^n p_i f(x_i) &= p_n f(x_n) + \sum_{i=1}^{n-1} p_i f(x_i) \\ &= p_n f(x_n) + (1 - p_n) \sum_{i=1}^{n-1} \frac{p_i}{1 - p_n} f(x_i) \\ &\geq p_n f(x_n) + (1 - p_n) f\left(\sum_{i=1}^{n-1} \frac{p_i}{1 - p_n} x_i\right) \quad (\text{induction hypothesis}) \\ &\geq f\left(p_n x_n + (1 - p_n) \sum_{i=1}^{n-1} \frac{p_i}{1 - p_n} x_i\right) \quad (\text{definition of convexity}) \\ &= f\left(p_n x_n + \sum_{i=1}^{n-1} p_i x_i\right) \\ &= f\left(\sum_{i=1}^n p_i x_i\right). \end{aligned}$$

That proves the claim. As for the strictness claim, if  $x_1, \dots, x_n$  are not all identical, then either  $x_1, \dots, x_{n-1}$  are not all identical and the first inequality is strict by induction hypothesis, or  $x_1 = \dots = x_{n-1} \neq x_n$  so that the second inequality is strict by the definition of convexity.