Definition: Binary Entropy

For a **binary random variable** X with image $\mathcal{X} = \{x_0, x_1\}$ and probabilities $P_X(x_0) = p$ and $P_X(x_1) = 1 - p$, we can write $H(X) = h(p)$, where h denotes the binary entropy function:

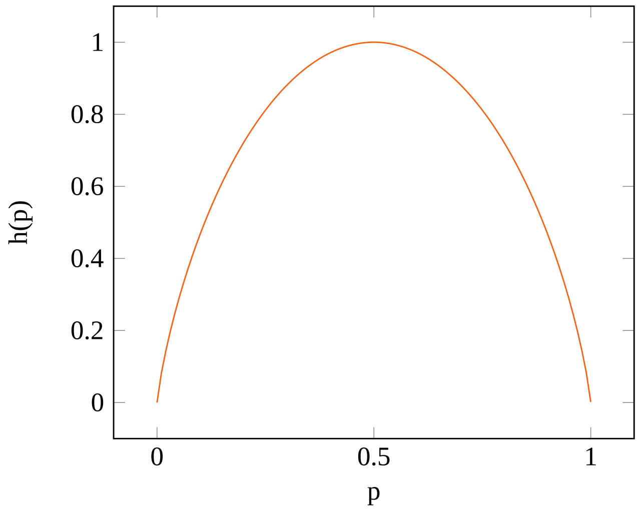
Definition: Binary entropy function h

The binary entropy function is defined for $0 < p < 1$ as

$$h(p) := p \log \frac{1}{p} + (1-p) \log \frac{1}{1-p},$$

and is defined as $h(p) = 0$ for $p = 0$ or $p = 1$.

The **graph of h** on the interval $[0, 1]$, as a function of p, looks as follows:



If we think of X as the random variable describing the outcome of a coin flip, we see that a relatively fair coin ($p \approx \frac{1}{2}$) yields a higher expected surprisal value than a very biased coin (where p is closer to 0 or 1). If the coin is completely fair ($p = \frac{1}{2}$), the entropy is exactly 1 bit.