Perplexity

Like entropy, **perplexity** provides a measure of the amount of uncertainty of a random variable. In fact, perplexity is simply a monotonic function of entropy:

$$\text{Perp}(X) := 2^{H(X)}$$

That is, the perplexity of a random variable X is simply two raised to the entropy of X.

Definition 1 *Given a random variable X, its* ***perplexity*** *is*

$$Perp(X) := 2^{H(X)}$$

where $H(X)$ is the entropy of X.

Intuition

Perplexity is often more interpretable than the entropy of a random variable and is therefore often used when we want uncertainty to be expressed in more "human" terms. The interpretation of perplexity arises from the fact that the perplexity of a uniform, discrete random variable with K outcomes is K (Theorem 1). For example, the perplexity of a fair coin is two and the perplexity of a fair six-sided die is six. This provide a frame of reference when interpreting a perplexity value. That is, if the perplexity of some random variable X is $\mathcal{P}$, our uncertainty towards the outcome of X is equal to the uncertainty we would have towards a $\mathcal{P}$-sided die.

Theorem 1 *Given a discrete uniform random variable*

$$X \sim Cat(p_1, p_2, \ldots, p_K)$$

where $\forall i, j \in [K]\ p_i = p_j$, the perplexity of X is K.

Proof:

$$\begin{aligned}
\text{Perp}(X) &= 2^{H(X)} \\
&= 2^{-\frac{1}{K}\log_2 \frac{1}{K}} \\
&= 2^{-\log_2 \frac{1}{K}} \\
&= \frac{1}{2^{\log_2 \frac{1}{K}}} \\
&= \frac{1}{1/K} \\
&= K
\end{aligned}$$

□