Information Theory

Information Theory

In information theory, the analog of the law of large numbers is the Asymptotic Equipartition Property (AEP).

The Law Of Large Numbers states that for Indenpendent, Identically Distributed (*i.i.d.*) random variables states:

$\frac{1}{N}\Sigma_{k=1}^{N} X_k \to E\left(N\right)$ for sufficiently large N

The Asymptotic Equipartition Property (AEP) states:

$\frac{1}{N}\log_2 \frac{1}{p(X_1, X_2, \cdots, X_N)} \to H\left(X\right)$ where H is the entropy, X_k's are *i.i.d.* and $p\left(X_1, X_2, \cdots, X_N\right)$ is the probability of observing the sequence $X_1, X_2, \cdots, X_N$.

This enables us to divide the set of all sequences into two sets, the *typical set*, where the sample entropy is close to the true entropy, and the nontypical set, which contains the other sequences.

Entropy

The entropy is a measure of the average uncertainty in a random variable.

The entropy of a random variable X with a probability mass function $p(x)$ is defined by

$$H\left(X\right) = -\sum_{x} p\left(x\right) \log_2 p\left(x\right) \tag{1}$$

Asymptotic Equipartition Property (AEP)

If $X_1, X_2, \cdots, X_N$ are $i.i.d$ ~ p(x), then

$$\frac{1}{N}\log_2 p\left(X_1, X_2, \cdots, X_N\right) \to H\left(X\right) \text{ in probability} \tag{2}$$

Proof:

$$\begin{aligned}\frac{1}{N}\log_2 p\left(X_1, X_2, \cdots, X_N\right) &= -\frac{1}{N}\sum_k \log_2 p\left(X_k\right) \\ &\to -E\log_2 p\left(X\right) \\ &\to H\left(X\right)\end{aligned}$$

Typical Set $A_\epsilon^{(n)}$

A *typical set* $A_\epsilon^{(n)}$ with respect to p(x) is the set of sequences $(x_1, x_2, \cdots, x_N) \in \mathcal{X}^N$ with the property

$$2^{-N(H(X)+\epsilon)} \leq p\left(x_1, x_2, \cdots, x_n\right) \leq 2^{-N(H(X)-\epsilon)}$$