The Asymptotic Equipartition Property (AEP)

The weak law of large numbers has an entropy variant, which follows almost directly:

Theorem: Asymptotic Equipartition Property (AEP)

Let $X_1, X_2, X_3, \ldots$ be i.i.d. random variables with distribution P_X. Then

$$-\frac{1}{n} \log P_{X_1 \cdots X_n}(X_1, \ldots, X_n) \xrightarrow{p} H(X).$$

(Note that $P_{X_1 \cdots X_n}(X_1, \ldots, X_n)$ is itself a random variable, and $H(X)$ can be regarded as a constant random variable.)

Proof

Since the variables X_i are independent, so are the random variables $\log P_X(X_i)$. Then

$$\begin{aligned} -\frac{1}{n} \log P_{X_1 \cdots X_n}(X_1, \ldots, X_n) &= -\frac{1}{n} \sum_{i=1}^{n} \log P_X(X_i) \\ &\xrightarrow{p} -\mathbb{E}[\log P_X(X_i)] = H(X) \end{aligned}$$

by the weak law of large numbers.

In terms of surprisal values, the AEP states that your surprisal value, averaged over all samples, will eventually approach the entropy $H(X)$ of a single sample.