Properties of Typical Sets

In the coin-flipping example on the previous page, the typical set eventually contained almost all of the probability. The following proposition states that this is a general property of typical sets. The proposition also bounds the size of the typical set.

Proposition

A typical set $A_\epsilon^{(n)}$ satisfies the following:

1. For all $(x_1, \dots, x_n) \in A_\epsilon^{(n)}$,

$$H(X) - \epsilon \leq -\frac{1}{n} \log P_{X^n}(x_1, \dots, x_n) \leq H(X) + \epsilon.$$

2. $P[A_\epsilon^{(n)}] > 1 - \epsilon$ (for large enough n).
3. $|A_\epsilon^{(n)}| \leq 2^{n(H(X)+\epsilon)}$.
4. $|A_\epsilon^{(n)}| \geq (1-\epsilon) 2^{n(H(X)-\epsilon)}$ (for large enough n).

Proof

1. This is immediate from the definition (take the logarithm and divide by $-n$, thereby reversing the inequalities).
2. This follows from the Asymptotic Equipartition Property: for all $\epsilon > 0$,

$$P[| - \frac{1}{n} \log P_{X^n}(X_1, \dots, X_n) - H(X)| > \epsilon] \xrightarrow{n \to \infty} 0,$$

   that is,

$$\forall (\epsilon > 0) \; \forall (\delta > 0) \; \exists n_0 \; \forall (n \geq n_0) \; P[| - \frac{1}{n} \log P_{X^n}(X_1, \dots, X_n) - H(X)| \leq \epsilon] > 1 - \delta.$$

   By choosing $\delta := \epsilon$, the result follows from the first property.
3. First, observe that

$$1 = \sum_{\vec{x} \in \mathcal{X}^n} P_{X^n}(\vec{x}) \geq \sum_{\vec{x} \in A_\epsilon^{(n)}} P_{X^n}(\vec{x}) \geq |A_\epsilon^{(n)}| \cdot 2^{-n(H(X)+\epsilon)},$$

   where the last inequality follows by the definition of typicality. The claim follows by multiplying both sides of the equation by $2^{n(H(X)+\epsilon)}$.
4. By Property 2, we can choose an n large enough so that

$$1 - \epsilon < P[A_\epsilon^{(n)}] = \sum_{\vec{x} \in A_\epsilon^{(n)}} P_{X^n}(\vec{x}) \leq |A_\epsilon^{(n)}| \cdot 2^{-n(H(X)-\epsilon)},$$

where again, the last inequality follows by the definition of typicality.