AEP for Ergodic Stationary Processes

We just saw that Shannon's source coding theorem generalizes from i.i.d. sources to stationary stochastic sources. Under some natural assumptions, many statements and interpretations of the Shannon entropy we have seen in the context of the asymptotic equipartition property (AEP) can be generalized to the entropy rate $H(\{X_i\})$ of stochastic processes. We do require that the process behaves naturally in the following sense:

Definition: Ergodic Random Processes

A stochastic process $\{X_i\}$ is **ergodic** if its statistical properties can be deduced from a single, sufficiently long, random sample of the process.

There are many equivalent ways of giving precise mathematical definitions of this notion, but this goes beyond the scope of this course. Instead, we consider the following examples.

Example

Suppose we repeatedly pick a letter at random and print it three times:
`LLL EEE HHH QQQ MMM QQQ OOO TTT EEE YYY XXX GGG...`

This random process is ergodic (as we see all letters eventually) but not stationary.

On the other hand, if we pick a letter at random and print it forever:
`GGGGGGGGGGGGGGGGGGGG...`

This process is stationary, but not ergodic (from one sample of the process, we can only see a single letter).

Shannon–McMillan–Breiman theorem: AEP

If $H(\{X_i\})$ is the entropy rate of a finite-valued stationary ergodic process $\{X_i\}$, then

$$-\frac{1}{n}\log P_{X_1,\ldots X_n}(X_1,\ldots X_n) \to H(\{X_i\}) \qquad \text{with probability 1.}$$

Typesetting math: 100%

created: 2019-10-21

In other words, the asymptotic equipartition property holds for such processes: we can define typical sets and all the results about data compression we have seen in the previous module not only hold for iid processes, but more generally for stationary ergodic processes.

created: 2019-10-21