Chapter 4: Entropy rate of a stochastic process

Chapter 4 outline

- Entropy rate
- Stochastic processes and Markov chains
Stochastic processes

- Know entropy of a sequence of iid RVs. What about the notion of entropy of a more general random process?

**Definition:** A stochastic process \( \{X_i\} \) is an indexed sequence of random variables.

**Definition:** A discrete-time stochastic process \( \{X_i\}_{i \in \mathbb{Z}} \) is one for which we associate the discrete index set \( I = \{1, 2, \ldots\} \) with time.

**Entropy:** \( H(\{X_i\}) = H(X_1) + H(X_2|X_1) + \cdots = \infty \) (often)

- Should probably normalize by \( n \) somehow!

Entropy rate

- **Entropy Rate:** The entropy rate of a stochastic process \( \{X_i\} \) is defined by
  \[
  H(X) = \lim_{n \to \infty} \frac{1}{n} H(X_1, X_2, \ldots, X_n)
  \]
  when the limit exists. We can also define an alternative notion:
  \[
  H'(X) = \lim_{n \to \infty} H(X_n|X_{n-1}, X_{n-2}, \ldots, X_1).
  \]

- Entropy rate estimates the additional entropy per new sample.
- Gives lower bound on number of code bits per sample.
- If the \( X_i \) are not i.i.d. the entropy rate limit may not exist.
- \( X_i \) i.i.d. random variables: \( H(X) = H(X_i) \)
Stationary processes

Definition: A discrete-time stochastic process is said to be stationary if the joint distribution of any subset of the sequence of random variables is invariant with respect to shifts in the time index; that is,

$$\Pr\{X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n\} = \Pr\{X_{1+l} = x_1, X_{2+l} = x_2, \ldots, X_{n+l} = x_n\}$$

for every $n$ and every shift $l$ and for all $x_1, x_2, \ldots, x_n \in \mathcal{X}$.

Lemma: For a stationary stochastic process, $H(X_n|X_{n-1}, X_{n-2}, \ldots, X_1)$ is nonincreasing in $n$ and has a limit $H'(\mathcal{X})$.

Lemma: Cesáro mean If $a_n \to a$ and $b_n = \frac{1}{n} \sum_{i=1}^{n} a_i$, then $b_n \to a$.

Theorem: For a stationary stochastic process, $H(\mathcal{X})$ and $H'(\mathcal{X})$ exist and are equal:

$$H(\mathcal{X}) = H'(\mathcal{X}).$$

Markov chains

- Book has nice results on the entropy rate of Markov chains in Chapter 4. We’re skipping it for now, but if you’re interested, have a look!