Source Coding and the Entropy Typical Set

Submit a PDF of your answers to Canvas

1. This problem will explore four properties of the *entropy typical set*. The properties imply the acheivability of the source coding theorem.

Consider n i.i.d. realizations of a discrete random variable $X \in \mathscr{X}$ with $|\mathscr{X}| < \infty$. The entropy typical set is defined as

$$A_n^{\varepsilon} = \left\{ \boldsymbol{x} : \left| \frac{1}{n} \log_2 \left(\frac{1}{p(\boldsymbol{x})} \right) - H(X) \right| \leq \varepsilon \right\},$$

where $\boldsymbol{x} \in \mathscr{X}^n$ is a vector of the n realizations.

from def

$-\varepsilon \leq \frac{1}{n} \log_2 \left(\frac{1}{p(x)}\right) - H(x) \leq \varepsilon$

$-n(H(x)+\varepsilon) \leq \log_2 p(x) \leq -n(H(x)-\varepsilon)$

$2^{-n(H(x)+\varepsilon)} \leq p(x) \leq 2^{-n(H(x)-\varepsilon)}$

a) Show that for any $\boldsymbol{x} \in A_n^{\varepsilon}$, $2^{-n(H(X)+\varepsilon)} \leq p(\boldsymbol{x}) \leq 2^{-n(H(X)-\varepsilon)}$.

b) Show that for any fixed $\varepsilon > 0$, by lln → $\frac{1}{n} \log_2 \left(\frac{1}{p(x)}\right)$ converges to $H(x)$ A.S. (prob=1)

$$\mathbb{P}(\boldsymbol{x} \in A_n^{\varepsilon}) \geq 1 - \varepsilon.$$

By ε-δ → large enough n can lead to more than 1-ε of the set converging

You may find results from a previous activity useful. Specify a value of n_0 for which $\mathbb{P}(\boldsymbol{x} \in A_n^{\varepsilon}) \geq 1 - \varepsilon$ holds for $n \geq n_0$.

c) Next show that $|A_{\varepsilon}^n| \leq 2^{n(H(X)+\varepsilon)}$ for sufficiently large n by justifying each line below:

$$\begin{aligned}
1 &\overset{(1)}{=} \sum_{\boldsymbol{x} \in \mathscr{X}} p(\boldsymbol{x}) \\
&\overset{(2)}{\geq} \sum_{\boldsymbol{x} \in A_{\varepsilon}^n} p(\boldsymbol{x}) && \text{is a subset, thus } \Sigma \text{ prob is } \geq \\
&\overset{(3)}{\geq} \sum_{\boldsymbol{x} \in A_{\varepsilon}^n} 2^{-n(H(X)+\varepsilon)} && \text{from (a) we derived} \\
&\overset{(4)}{=} 2^{-n(H(X)+\varepsilon)} |A_{\varepsilon}^n|. && \text{after summation}
\end{aligned}$$

d) Show that $|A_{\varepsilon}^n| \geq (1-\varepsilon) 2^{n(H(X)-\varepsilon)}$ for sufficiently large n by justifying each step:

$$\begin{aligned}
1 - \varepsilon &\overset{(1)}{<} \mathbb{P}(\boldsymbol{x} \in A_n^{\varepsilon}) \text{ for sufficiently large } n && \text{from (b)} \\
&\overset{(2)}{\leq} \sum_{\boldsymbol{x} \in A_{\varepsilon}^n} 2^{-n(H(X)-\varepsilon)} && \text{from (a), upper bound} \\
&\overset{(3)}{=} 2^{-n(H(X)-\varepsilon)} |A_{\varepsilon}^n|. && \text{after summation}
\end{aligned}$$

2. Write a short paragraph to explain how problem (1) shows that n i.i.d. random variables can be compressed into $nH(X)$ bits with a negligible risk of error as n grows. Note that we have shown that the entropy typical set contains a relatively small number of the possible realizations of $\boldsymbol{x}$, but most of the probability. While we have not shown the converse, there is not a significantly smaller set (in an asymptotic sense, see [Cover and Thomas, Theorem 3.3.1.]).

A_n^ε denotes the set of $x \in \mathcal{X}$ where, once compressed to $n\,H(x)$ bits, the error remains within ε. Since $P(x \in A_n^\varepsilon)$ converges to 1 with smaller ε and sufficiently large n, we can pick an n large enough the error is negligible (within ε)