EEE 551 Information Theory (Spring 2022)

Homework 2

Due: February 16, 2022

---

1. *The Binary Typical Set.* Let $X \sim \text{Bern}(p)$ (i.e. it is binary with $p_X(1) = p,\ p_X(0) = 1 - p$). Given a sequence $x^n \in \{0,1\}^n$, let n_1 be the number of 1s that appear in the sequence (e.g. if $x^n = (1,1,0,1)$, then $n_1 = 3$).

   (a) Write $p(x^n)$ in terms of only p, n, and n_1.

   (b) Find a condition (again in terms of only p, n, and n_1) for the sequence x^n being in the typical set $A_\epsilon^{(n)}$.

   (c) If $\epsilon = 0$, is the typical set empty?

   (d) Suppose $p = 1/3$, $n = 20$, and $\epsilon = 0.05$. Calculate $\Pr\{A_\epsilon^{(n)}\}$. *Hint:* If $X^n \overset{\text{iid}}{\sim} \text{Bern}(p)$, then n_1 is a binomial random variable.

2. Problem 3.9 from Cover-Thomas: *AEP.* Let $X_1, X_2, \ldots$ be independent, identically distributed random variables drawn according to the probability mass function $p(x), x \in \{1, 2, \ldots, m\}$. Thus $p(x_1, x_2, \ldots, x_n) = \prod_{i=1}^n p(x_i)$. We know that $-\frac{1}{n} \log p(X_1, X_2, \ldots, X_n) \to H(X)$ in probability. Let $q(x_1, x_2, \ldots, x_n) = \prod_{i=1}^n q(x_i)$, where q is another probability mass function on $\{1, 2, \ldots, m\}$.

   (a) Evaluate $\lim -\frac{1}{n} \log q(X_1, X_2, \ldots, X_n)$, where $X_1, X_2, \ldots$ are i.i.d. $\sim p(x)$.

   (b) Now evaluate the limit of the log likelihood ratio $\frac{1}{n} \log \frac{q(X_1,\ldots,X_n)}{p(X_1,\ldots,X_n)}$ when $X_1, X_2, \ldots$ are i.i.d. $\sim p(x)$. Thus, the odds favoring q are exponentially small when p is true.

3. *Optimal Fixed-to-Fixed Source Code.* Consider the following distribution:

x	a	b	c
$p(x)$	0.5	0.3	0.2

   This problem requires you to find the fixed-to-fixed source code that minimizes the probability of error for a given number of compressed bits. This optimal code will not necessarily be the code based on the typical set. A fixed-to-fixed code compresses a sequences of length n to a fixed number ℓ bits. That is, there is an encoding function
   $$f : \mathcal{X}^n \to \{0,1\}^\ell$$
   and a decoding function
   $$g : \{0,1\}^\ell \to \mathcal{X}^n.$$
   For each of the following values of n and ℓ, find the code that minimizes the probability of error, and the associated probability of error.

   (a) $n = 1, \ell = 1$

   (b) $n = 2, \ell = 1$

(c) $n = 2, \ell = 2$

(d) $n = 2, \ell = 3$

4. *Coding for Two Different Distributions.* Consider two distributions $p(x)$ and $q(x)$ on the same alphabet $\mathcal{X}$. We want to design a source code that works with *either* distribution. Let $R > \max\{H_p(X), H_q(X)\}$, where H_p and H_q refer to the entropies under $p(x)$ and $q(x)$ respectively. Find a fixed-to-fixed source code with rate R such that the probability of error is less than ϵ if the underlying distribution is either $p(x)$ or $q(x)$.

5. Problem 5.9 from Cover-Thomas: *Optimal code lengths that require one bit above entropy.* The source coding theorem shows that the optimal code for a random variable X has an expected length less than $H(X) + 1$. Give an example of a random variable for which the expected length of the optimal code is close to $H(X) + 1$ [i.e., for any $\epsilon > 0$, construct a distribution for which the optimal code has $L(C) > H(X) + 1 - \epsilon$]. *Hint:* Try a Bernoulli distribution.

6. Problem 5.12 from Cover-Thomas: *Shannon codes and Huffman codes.* Consider a random variable X that takes on four values with probabilities $(\frac{1}{3}, \frac{1}{3}, \frac{1}{4}, \frac{1}{12})$.

   (a) Construct a Huffman code for this random variable.

   (b) Show that there exist two different sets of optimal lengths for the codewords; namely, show that codeword length assignments $(1, 2, 3, 3)$ and $(2, 2, 2, 2)$ are both optimal.

   (c) Conclude that there are optimal codes with codeword lengths for some symbols that exceed the Shannon code length $\lceil \log \frac{1}{p(x)} \rceil$.

7. *Huffman Codes on Five Letters.* Consider a random variable with alphabet $\mathcal{X} = \{\mathsf{a}, \mathsf{b}, \mathsf{c}, \mathsf{d}, \mathsf{e}\}$, where

$$p_{\mathsf{a}} \geq p_{\mathsf{b}} \geq p_{\mathsf{c}} \geq p_{\mathsf{d}} \geq p_{\mathsf{e}}.$$

   Let $(l_{\mathsf{a}}, l_{\mathsf{b}}, l_{\mathsf{c}}, l_{\mathsf{d}}, l_{\mathsf{e}})$ be the codeword lengths for the Huffman code for this distribution. Determine the Huffman codeword lengths as a function of the distribution; i.e. give all possible Huffman word lengths that could occur, and for each one, say what conditions on the distribution need to hold in order for this to be the Huffman code.

8. *Coding for the Wrong Distribution.* Let $p(x)$ and $q(x)$ be two different distributions on the same alphabet $\mathcal{X}$. Suppose you construct a code for distribution $q(x)$, but the true distribution is $p(x)$. Consider the fixed-to-variable code with Shannon code length for $q(x)$, i.e., the length is $\ell(x) = \lceil \log \frac{1}{q(x)} \rceil$. Prove that the expected length (under distribution $p(x)$) satisfies

$$H(X) + D(p\|q) \leq L(C) < H(X) + D(p\|q) + 1.$$