EE276: Problem Session #3 Problems

1. **AEP.**
   Let X_i be iid $\sim p(x),\ x \in \{1, 2, \ldots, m\}$. Let $\mu = EX$, and $H = -\sum p(x)\log p(x)$. Let $A^n = \{x^n \in \mathcal{X}^n : |-\frac{1}{n}\log p(x^n) - H| \leq \epsilon\}$. Let $B^n = \{x^n \in \mathcal{X}^n : |\frac{1}{n}\sum_{i=1}^n X_i - \mu| \leq \epsilon\}$.

   (a) Does $\Pr\{X^n \in A^n\} \longrightarrow 1$ as $n \to \infty$?

   (b) Does $\Pr\{X^n \in A^n \cap B^n\} \longrightarrow 1$ as $n \to \infty$?

   (c) Show $|A^n \cap B^n| \leq 2^{n(H+\epsilon)}$, for all n.

   (d) Show $|A^n \cap B^n| \geq (\frac{1}{2})2^{n(H-\epsilon)}$, for n sufficiently large.

2. **Uniquely decodable and instantaneous codes.** Let $L = \sum_{i=1}^m p_i l_i^{100}$ be the expected value of the 100th power of the word lengths associated with an encoding of the random variable X. Let $L_1 = \min L$ over all instantaneous codes; and let $L_2 = \min L$ over all uniquely decodable codes. What inequality relationship exists between L_1 and L_2?

3. **Classes of codes.** Consider the code $\{0, 01\}$

   (a) Is it instantaneous?

   (b) Is it uniquely decodable?

   (c) Is it nonsingular?

4. **Data compression.** Find an optimal set of binary codeword lengths $l_1, l_2, \ldots$ (minimizing $\sum p_i l_i$) for an instantaneous code for each of the following probability mass functions:

   (a) $\mathbf{p} = (\frac{10}{41}, \frac{9}{41}, \frac{8}{41}, \frac{7}{41}, \frac{7}{41})$

   (b) $\mathbf{p} = (\frac{9}{10}, (\frac{9}{10})(\frac{1}{10}), (\frac{9}{10})(\frac{1}{10})^2, (\frac{9}{10})(\frac{1}{10})^3, \ldots)$

5. **Slackness in the Kraft inequality.**

   (a) In the lecture, we have seen Kraft inequality for prefix codes over the alphabet $\{0, 1\}$ with alphabet size $D = 2$. The proof used a binary tree, where the branches represent the symbols of the codeword and each codeword is represented by a leaf on the tree. For prefix codes over an alphabet of size D, what would be the degree of the tree?

   (b) In the lecture, we have observed that when $D = 2$, an optimal code satisfies Kraft inequality with equality. Is this still true when $D > 2$?

   (c) A prefix code has word lengths $l_1, l_2, \ldots, l_m$ which satisfy the strict inequality

   $$\sum_{i=1}^m D^{-l_i} < 1.$$

   The code alphabet is $\mathcal{D} = \{0, 1, 2, \ldots, D-1\}$. Show that there exist arbitrarily long sequences of code symbols in $\mathcal{D}^*$ which cannot be decoded into sequences of codewords.