EE276 Homework #7

Due on Friday March 8, 5pm

1. Consider a finite alphabet $\mathcal{X}$. Given $D \geq 0$ and a weight function $\rho : \mathcal{X} \to \mathbb{R}_+$, define

$$B_n(\rho, D) := \left\{ x^n : \frac{1}{n} \sum_{i=1}^{n} \rho(x_i) \leq D \right\}.$$

   (a) Show that

$$B_n(\rho, D) = \bigcup_{p \in \mathbb{P}_n : \langle p, \rho \rangle \leq D} T(p)$$

   where

$$\langle p, \rho \rangle := \sum_{x \in \mathcal{X}} p(x)\rho(x).$$

   (b) Show that

$$|B_n(\rho, D)| \doteq 2^{n \max_{p : \langle p, \rho \rangle \leq D} H(p)}.$$

   (Use the expression derived in (a) to obtain lower and upper bounds on $|B_n(\rho, D)|$ that match up to first order in exponent.)

   (c) Specializing the result of (b), show that for $D \in [0, 1/2]$,

$$\left| \left\{ y^n \in \{0,1\}^n : \frac{1}{n} \sum_{i=1}^{n} y_i \leq D \right\} \right| \doteq 2^{n h_2(D)}$$

   where h_2 is the binary entropy function.

2. In what follows, all random variables have finite alphabet, and all pmfs are defined on finite alphabets.

   (a) Let Z be a random variable with pmf p_z and alphabet $\mathcal{Z}$. Let $T_\delta(p_z)$ be the set of strongly δ-typical sequences with respect to p_z.
   Show that for any $g : \mathcal{Z} \to \mathbb{R}$, and $z^n \in T_\delta(p_z)$, we have

$$\left| \frac{1}{n} \sum_{i=1}^{n} g(z_i) - \mathbb{E}[g(Z)] \right| \leq \delta \mathbb{E}[g(Z)].$$

   In what follows, it may be useful to invoke part (a) in your arguments.

   (b) Let (X, Y) be random variables with joint $p_{x,y}$. Let $T_\delta(p_{x,y})$ be the set of jointly δ-typical sequences with respect to $p_{x,y}$. Show that

$$\frac{1}{n} \sum_{i=1}^{n} d(x_i, y_i) \leq (1 + \delta)\mathbb{E}[d(X, Y)]$$

   for any distortion function $d(x, y)$ and $(x^n, y^n) \in T_\delta(p_{x,y})$.

(c) Let $A_\epsilon(p)$ denote the (weakly) typical set with respect to p, and $T_\delta(p)$ be the set of δ-typical sequences with respect to p. Show that

$$T_\delta(p) \subseteq A_\epsilon(p)$$

for $\epsilon = \delta H(p)$.

(d) Let Q and P be pmfs over $\mathcal{X}$, and $T_\delta(P)$ be the set of δ-typical sequences with respect P. Show that

$$Q(T_\delta(P)) \doteq 2^{-n(D(P\|Q)-\alpha(\delta))},$$

where $\alpha(\delta) \geq 0$ and $\alpha(\delta) \to 0$ as $\delta \to 0$.

3. **Modulo 8 Channel and a Binary Source with Erasure Distortion**

Let a, M be arbitrary real numbers, and let $M > 0$. Then we say, $b = a$ mod M, if $b \in [0, M)$ and $b = a - kM$, for some integer k. For example, if $a = 10.2$ and $M = 8$ then $b = 2.2$.

Consider the memoryless channel described by

$$Y = (X + Z) \bmod 8,$$

where the channel input X, output Y, and the noise Z are real valued. The noise is uniformly distributed on $[-B, B]$, i.e. $Z \sim \text{Uniform}[-B, B]$ and is independent of the input X. Assume that $0 < B \leq 4$ is a known channel parameter.

(a) Show that the channel capacity as a function of B is $2 - \log_2 B$.

(b) A communication system is suggested where the permitted channel input values are restricted to the set $\{1, 3, 5, 7\}$. Show that, when $B = 1$, it is still possible to achieve the channel capacity (in fact, to achieve this capacity with zero probability of error and with a very simple scheme).

Consider now a Bernoulli(1/2) source under erasure distortion, i.e., the reconstruction alphabet is $\{0, 1, e\}$ and

$$d(u, v) = \begin{cases} 0 & \text{if } u = v \\ 1 & \text{if } v = e \\ \infty & \text{otherwise} \end{cases}$$

(c) Show that the rate-distortion function of this source is given by $R(D) = 1 - D$ for $0 \leq D \leq 1$.

(d) Can you suggest an explicit scheme (not random coding based) which achieves the optimal rate-distortion performance for the Bernoulli(1/2) source, and distortion D.

(e) Consider now a joint-source-channel-coding setting for communicating the bit source of part (c) through the modulo 8 channel defined in part (a). For every two source symbols we are allotted one channel use, i.e., the encoder translates a block of $2n$ source bits into n channel inputs.
Find the minimal achievable distortion (when $n \to \infty$) as a function of B.

(f) The following communication system is suggested for the setting of the previous part: every two source bits are mapped into a channel input in the following way:

$$\begin{array}{ccc} 00 & \to & 1 \\ 01 & \to & 3 \\ 11 & \to & 5 \\ 10 & \to & 7 \end{array}$$

This mapping is known as a Gray Code. Show that this system (with the corresponding optimal decoder) achieves the minimal distortion in the case where $B = 2$.