EE276 Homework #6

Due on March 1, 5pm

1. **Rate distortion for uniform source with Hamming distortion.**
Consider a source X uniformly distributed on the set $\{1, 2, ..., m\}$. Find the rate distortion function for this source with Hamming distortion, i.e.,

$$d(x, \hat{x}) = \begin{cases} 0, & x = \hat{x} \\ 1, & x \neq \hat{x} \end{cases}$$

via the following steps:

   (a) Argue that $R(D) = 0$ when $D \geq 1 - \frac{1}{m}$.

   (b) Show that for $D \leq 1 - \frac{1}{m}$, $I(X; \hat{X}) \geq \log_2 m - h_2(D) - D \log_2(m-1)$ for any joint distribution $(X, \hat{X})$ satisfying the distortion constraint D.
   Hint: Fano's inequality.

   (c) Find distribution $p(\hat{x}|x)$ that achieves the above lower bound when $0 \leq D \leq 1 - \frac{1}{m}$.

   (d) Use the above parts to write down the rate-distortion function $R(D)$ for $D \geq 0$.

2. **Convexity of rate distortion function.**
Assume $(X, Y) \sim p(x, y) = p(x)p(y|x)$. In this problem, you will show that for fixed $p(x)$, $I(X; Y)$ is a convex function of $p(y|x)$.

   (a) The log sum inequality states that for n positive numbers $a_1, a_2, \cdots, a_n$, and $b_1, b_2, \cdots, b_n$, we have

   $$\sum_{i=1}^{n} a_i \log \frac{a_i}{b_i} \geq \left(\sum_{i=1}^{n} a_i \right) \log \left(\frac{\sum_{i=1}^{n} a_i}{\sum_{i=1}^{n} b_i} \right)$$

   with equality if and only if $\frac{a_i}{b_i}$ =constant. Using this inequality (you don't have to prove this inequality), show that $D(p||q)$ is convex in (p, q), i.e.,

   $$\lambda D(p_1||q_1) + (1-\lambda) D(p_2||q_2) \geq D(\lambda p_1 + (1-\lambda)p_2 || \lambda q_1 + (1-\lambda) q_2)$$

   (b) Let $p_1(y|x)$ and $p_2(y|x)$ be two different conditional distributions. For $i \in \{1, 2\}$, let $p_i(x, y) = p_i(y|x)p(x)$, i.e., their corresponding joint distributions. For $0 \leq \lambda \leq 1$, let $p_\lambda(y|x) \triangleq \lambda p_1(y|x) + (1-\lambda) p_2(y|x)$. Show that

   $$p_\lambda(y) = \lambda p_1(y) + (1-\lambda) p_2(y)$$

   (c) The mutual information between random variables X and Y can be alternatively written as

   $$I(X; Y) = D(p(x, y) || p(x)p(y))$$

   Using this in addition to the results of the previous parts show that for fixed $p(x)$, $I(X; Y)$ is convex in $p(y|x)$.

(d) Using the previous part, show that the rate distortion function $R^{(I)}(D)$ is convex in the distortion parameter D.

3. **Shannon lower bound.**
Let X be a continuous random variable with mean zero and variance σ^2. $R(D)$ is the corresponding rate-distortion function for mean-squared distortion.

(a) Show the lower bound:

$$h(X) - \frac{1}{2}\log(2\pi e D) \leq R(D).$$

(b) Using the joint distribution shown in Figure 1, show the upper bound on $R(D)$:

$$R(D) \leq \frac{1}{2}\log\frac{\sigma^2}{D} \tag{1}$$

Are Gaussian random variables harder or easier to describe than other random variables with the same variance?

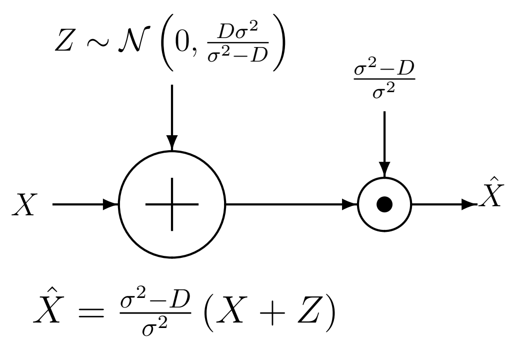

Figure 1: Joint distribution for upper bound on rate distortion function. The circle with the dot represents multiplication.

4. **Rate distortion for two independent sources.** Let $\{X_i\}$ be iid $\sim p(x)$ with distortion $d(x, \hat{x})$ and rate distortion function $R_X(D)$. Similarly, let $\{Y_i\}$ be iid $\sim p(y)$ with distortion $d(y, \hat{y})$ and rate distortion function $R_Y(D)$.

Suppose we now wish to describe the process $\{(X_i, Y_i)\}$ subject to distortions $\mathbb{E}[d(X, \hat{X})] \leq D_1$ and $\mathbb{E}[d(Y, \hat{Y})] \leq D_2$. Thus a rate $R_{X,Y}(D_1, D_2)$ is sufficient, where

$$R_{X,Y}(D_1, D_2) = \min_{p(\hat{x},\hat{y}|x,y):\mathbb{E}[d(X,\hat{X})]\leq D_1, \mathbb{E}[d(Y,\hat{Y})]\leq D_2} I(X, Y; \hat{X}, \hat{Y})$$

Suppose the $\{X_i\}$ process and the $\{Y_i\}$ process are independent of each other.

Express $R_{X,Y}(D_1, D_2)$ in terms of $R_X(D_1)$ and $R_Y(D_2)$. Can one simultaneously compress two independent sources better than by compressing the sources individually?

5. **Distortion-rate function.** Let

$$D(R) = \min_{p(\hat{x}|x):I(X;\hat{X})\leq R} \mathbb{E}[d(X,\hat{X})] \tag{2}$$

be the distortion rate function.

(a) Is $D(R)$ increasing or decreasing in R?

(b) Is $D(R)$ convex or concave in R?

(c) Let $X_1, X_2, \ldots, X_n$ be i.i.d. $\sim p(x)$. Suppose one is given a code $(X^n, \hat{X}^n)$ with

$$\frac{1}{n} I(X^n; \widehat{X}^n) \leq R$$

and resulting distortion $D = \mathbb{E}[d(X^n, \hat{X}^n)]$. We want to show that $D \geq D(R)$. Give reasons for the following steps in the proof:

$$\begin{aligned}
D &= \mathbb{E}[d(X^n, \hat{X}^n(i(X^n)))] && (3)\\
&\overset{(a)}{=} \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n} d(X_i, \hat{X}_i)\right] && (4)\\
&\overset{(b)}{=} \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[d(X_i, \hat{X}_i)] && (5)\\
&\overset{(c)}{\geq} \frac{1}{n}\sum_{i=1}^{n} D\left(I(X_i; \hat{X}_i)\right) && (6)\\
&\overset{(d)}{\geq} D\left(\frac{1}{n}\sum_{i=1}^{n} I(X_i; \hat{X}_i)\right) && (7)\\
&\overset{(e)}{\geq} D\left(\frac{1}{n} I(X^n; \hat{X}^n)\right) && (8)\\
&\overset{(f)}{\geq} D(R) && (9)
\end{aligned}$$