1. Let $\mathcal{X}$ be a finite alphabet, assume $\hat{\mathcal{X}} = \mathcal{X}$, and that $d(x, \hat{x})$ satisfies

$$d(x, \hat{x}) = 0 \quad \text{if and only if } x = \hat{x}.$$ 

Define

$$D^* = \min_{\hat{x} \in \hat{\mathcal{X}}} Ed(X, \hat{x}).$$

Assume that $H(X) > 0$ and prove the following:

(a) $R(0) = H(X)$,

(b) $R(D) = 0$ if $D \geq D^*$,

(c) For graduate students: $R(D) > 0$ if $0 \leq D < D^*$.

2. Let $X \sim \text{Bernoulli}(1/2)$ with alphabet $\mathcal{X} = \{0, 1\}$ and let the reproduction alphabet and distortion measure be given by $\hat{\mathcal{X}} = \{0, e, 1\}$ and

$$d(x, \hat{x}) = \begin{cases} 
0 & \text{if } x = \hat{x}, \\
1 & \text{if } \hat{x} = e, \\
\infty & \text{if } x = 0 \text{ and } \hat{x} = 1 \text{, or } x = 1 \text{ and } \hat{x} = 0.
\end{cases}$$

Determine $R(D)$.

Undergraduate students: You can assume that in the definition of $R(D)$ its enough to consider symmetric conditional distributions $p(\hat{x}|x)$ (i.e., such that $p(0|0) = p(1|1)$). Then $I(X; \hat{X}) = H(X) - H(X|\hat{X})$ and $Ed(X, \hat{X})$ are easy to evaluate.

Graduate students: First will have to prove that the symmetry of the problem and the convexity of the mutual information implies that it is enough to consider symmetric conditional distributions $p(\hat{x}|x)$ (i.e., such that $p(0|0) = p(1|1)$).

3. Cover&Thomas 10.8 The left inequality in (10.159) is the Shannon lower bound we proved in class, so only the right inequality needs to be proved and the question answered.

4. Let $X_1, X_2, \ldots, X_n, \ldots$ be a Bernoulli(1/2) source with alphabet $\{0, 1\}$ and let $d(x, \hat{x})$ be the Hamming distortion. The sequence $X^n$ is to be encoded so that its reconstruction satisfies $Ed(X^n, \hat{X}^n) \leq D$ for some $D \in [0, 1]$. Two coding schemes are proposed.

**Scheme 1.** Whenever the source produces a 0, it is flipped to a 1 with probability $\rho$ (it stays as 0 with probability $1 - \rho$), where the value of $\rho$ is chosen in accordance with the maximum allowed expected distortion $D$. (These flips are independent of each other.) Whenever the source produces a 1, the bit remains unchanged. In this manner, the original symmetric source is converted into an asymmetric source. Next, this asymmetric source is losslessly encoded, using $R_1$ bits/source symbol. The reconstruction (or decoding) is performed by reconstructing (losslessly) the asymmetric source. Assume $n$ is large and let $R_1(D)$ denote the minimum rate $R_1$ needed by this scheme to yield an expected distortion at most $D$. 


Scheme 2. Given a source sequence of length $n$, only the first $nR_2$ bits are losslessly encoded. The decoding consists of reconstructing (losslessly) these first $nR_2$ bits and padding them with 0s to obtain a sequence of total length $n$. For large $n$ let $R_2(D)$ denote the minimum number of bits/source symbol needed by this scheme to yield an expected distortion not exceeding $D$.

(a) Determine $R_1(D)$ as a function of $D$ for $0 \leq D \leq 1$.
(b) Determine $R_2(D)$ as a function of $D$ for $0 \leq D \leq 1$.
(c) Compare Schemes 1 and 2 by plotting $R_1(D)$ and $R_2(D)$ in the same figure. Also include the plot of the rate-distortion function $R(D)$ of the given source.