Example 12.14. Let \( X = Y = \{0, 1, \ldots, m - 1\} \) with mod-\( m \) addition. We again use the Hamming distortion, \( d(x, y) := 1 \) if \( x \neq y \) and \( d(x, y) = 0 \) if \( x = y \). Let \( P(X = x) = 1/m \) for all \( x \in X \). Find \( \hat{R}(D) \). As in last example, write

\[
I(p \times W) = H(X) - H(X|Y) \\
= H(X) - H(X - Y|Y) \\
\geq H(X) - H(X - Y),
\]

with equality \( \iff \) \( X - Y \) and \( Y \) are independent. Let \( Z := X - Y \). Observe that

\[
E[d(X,Y)] = E[1_{\{X\neq Y\}}] = P(X \neq Y) = P(X - Y \neq 0) = P(Z \neq 0).
\]

Next,

\[
H(X - Y) = H(Z) = P(Z = 0) \log \frac{1}{P(Z = 0)} + \sum_{z=1}^{m-1} P(Z = z) \log \frac{1}{P(Z = z)} \\
\leq P(Z = 0) \log \frac{1}{P(Z = 0)} + P(Z \neq 0) \log \frac{m-1}{P(Z \neq 0)} \\
= h(P(Z \neq 0)) + P(Z \neq 0) \log (m-1),
\]

by the log-sum inequality, with equality \( \iff \)

\[
P(Z = z) = \frac{P(Z \neq 0)}{m-1}, \quad \text{for } z \neq 0.
\]

Assume \( Z \) has this distribution and that \( E[d(X,Y)] = D \). Then

\[
H(X - Y) = H(Z) \leq h(D) + D \log (m-1).
\]

Now suppose also that \( Y \) and \( Z \) are independent with

\[
P(Y = y) = r(y), \quad P(Z = 0) = 1 - D, \quad \text{and} \quad P(Z = z) = \frac{D}{m-1}, \quad z \neq 0.
\]
Can we find a probability mass function $r$ such that if $X := Y + Z$, then $P(X = x) = 1/m$ for all $x \in X$? By a HW Problem, it suffices to take $r$ to be the uniform distribution. With this choice of $r$ and $P(Z = 0) = 1 - D$, we have $E[d(X,Y)] = D$ and

$$I(X \wedge Y) = H(X) - H(X|Y) = \log |X| - [h(D) + D\log(|X| - 1)],$$

which gives,

$$\tilde{R}(D) = \begin{cases} 
\log |X| - [h(D) + D\log(|X| - 1)], & 0 \leq D < \frac{|X| - 1}{|X|}, \\
0, & D \geq \frac{|X| - 1}{|X|}.
\end{cases}$$

The case $D \geq (|X| - 1)/|X|$ can be handled by taking $W(y|x) := 1/|X|$. Then

$$E[d(X,Y)] = \sum_x p(x) \sum_y W(y|x)d(x,y)$$

$$= \sum_x p(x) \sum_{y \neq x} W(y|x)$$

$$= \sum_x p(x) \frac{|X| - 1}{|X|} \leq D.$$

Furthermore, $W(y|x) = 1/|X|$ implies $X$ and $Y$ are independent, which implies that $I(X \wedge Y) = 0$. Since the mutual information cannot be negative, we have shown that the infimum must be zero, that is, $\tilde{R}(D) = 0$ for $D \geq (|X| - 1)/|X|$.

---

12.9. The Shannon Lower Bound

Consider a distortion measure $d(x,y)$ such that for all $y, y' \in Y$, $\{d(\cdot, y)\}$ and $\{d(\cdot, y')\}$ are permutations of each other. Put

$$\eta(D) := \sup_{q: \sum_x q(x)d_c(x) \leq D} H(q)$$

where $d_c(x) := d(x, y_0)$ for some fixed $y_0 \in Y$. Since $H(q) = H(q')$ when $q'$ is a permutation of $q$, we see that $\eta(D)$ does not depend on $y_0$.

**Theorem 12.15** (Shannon Lower Bound). Let $d$ have the permutation property of the preceding paragraph. Assuming finite source and reproduction alphabets and a discrete memoryless source with pmf $p$,

$$R(D) \geq H(p) - \eta(D),$$

where $\eta$ is nondecreasing and concave.
Proof. It is clear that $\eta$ is a nondecreasing function of $D$. We next show that $\eta$ is a concave function of $D$. Suppose $\eta(D_1) = H(q_1)$ and $\eta(D_2) = H(q_2)$, where $\sum_x q_1(x)d_c(x) \leq D_1$ and $\sum_x q_2(x)d_c(x) \leq D_2$. Put $q = \lambda q_1 + (1 - \lambda) q_2$ where $0 \leq \lambda \leq 1$. Then,
\[
\sum_x q(x)d_c(x) = \lambda \sum_x q_1(x)d_c(x) + (1 - \lambda) \sum_x q_2(x)d_c(x)
\leq \lambda D_1 + (1 - \lambda) D_2.
\]
Furthermore,
\[
\eta(\lambda D_1 + (1 - \lambda) D_2) \geq H(q)
\geq \lambda H(q_1) + (1 - \lambda) H(q_2), \quad \text{since } H \text{ is concave},
\]
\[
= \lambda \eta(D_1) + (1 - \lambda) \eta(D_2).
\]
We now establish the lower bound. Suppose that $\mathbb{E}[d(X,Y)] \leq D$. Then
\[
D \geq \sum_x \sum_y P_{XY}(x,y)d(x,y)
= \sum_y P_Y(y) \left[ \sum_x P_{X|Y}(x|y)d(x,y) \right].
\]
With $\hat{d}(y)$ as above,
\[
I(X \wedge Y) = H(X) - H(X|Y)
= H(X) - \sum_y P_Y(y)H(X|Y = y)
= H(X) - \sum_y P_Y(y)H(P_{X|Y}(\cdot|y))
\geq H(X) - \sum_y P_Y(y)\eta(\hat{d}(y)), \quad \text{by def. of } \eta,
\]
\[
\geq H(X) - \eta(\sum_y P_Y(y)\hat{d}(y)), \quad \text{since } \eta \text{ is concave},
\]
\[
= H(X) - \eta(\mathbb{E}[d(X,Y)]),
\]
\[
\geq H(X) - \eta(D), \quad \text{since } \eta \text{ is nondecreasing}. \quad \square
\]

Remark. It can be shown that the Shannon lower bound holds with equality if $d$ has the additional property that for any $x, x' \in X$, $\{d(x, \cdot)\}$ and $\{d(x', \cdot)\}$ are permutations of each other.