Department of Electrical and Computer Engineering
ECE 776 - Information Theory

Reading

Cover and Thomas, Chapter 8, 9.

Homework 6

1. **Problem 8.4**
   Let X_i be i.i.d. Bernoulli distribution with $q(x_i = 0) = q(x_i = 1) = 1/2$. Then

$$\begin{aligned}
\max_{p(x_1,\ldots,x_n)} I(X_1,\ldots,X_n;Y_1,\ldots,Y_n) &\geq I(X_1,\ldots,X_n;Y_1,\ldots,Y_n), \qquad \text{where } p(x_1,\ldots,x_n) = q(x_1)\cdots q(x_n) \\
&= H(X_1,\ldots,X_n) - H(X_1,\ldots,X_n|Y_1,\ldots,Y_n) \\
&= H(X_1,\ldots,X_n) - H(Z_1,\ldots,Z_n|Y_1,\ldots,Y_n) \\
&\geq H(X_1,\ldots,X_n) - H(Z_1,\ldots,Z_n) \\
&\geq n - \sum_i H(Z_i) \\
&= n - nH(p, 1-p) = nC.
\end{aligned}$$

2. **Problem 8.11**
   (a) We know that
   $$C = \max_{p(x)} I(X;Y) = \max_{p(x)} H(Y) - H(Y|X).$$
   Further,
   $$H(Y|X) = E[\log \frac{1}{p(Y|X)}] = 1 \text{ bit}.$$
   $H(Y)$ is maximized when $p(x)$ is uniform. Thus $C = \log \frac{5}{2} = 1.32$.

   (b) The concept of zero-error capacity is still not well defined. The Shannon channel capacity is the capacity of channel when small error is allowed and the small error dimishes with large code length. However, when possible, it is more desirable to know the capacity when no error occurs.

   In this case, we can use the code $\{03, 14, 20, 31, 42\}$ to achieve zero-error. The rate of this code is $R = \frac{1}{2}\log 5$, which is the the zero-error capacity of this channel. See reference Lovasz[182].

3. **Problem 8.12**

$$
\begin{aligned}
I(X^n;Y^n) &= H(Y^n) - H(Y^n|X^n) = H(Y^n) - \sum_{i=1}^{n} H(Y_i|X_i) \\
&\leq \sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} H(Y_i|X_i) \\
&\leq n[1 - H(p_i)]
\end{aligned}
$$

The quality is achieved if X_i are i.i.d. with $p(x_i = 0) = p(x_i = 1) = 1/2$. Thus,

$$C = \max_{p(X^n)} I(X^n;Y^n) = n[1 - H(p_i)].$$

4. **Problem 9.1**

   (a) Exponential density.

   $$h(f) = -\int_0^\infty \lambda e^{-\lambda x}[\ln\lambda - \lambda x]dx = \log\frac{e}{\lambda} \text{ bits.}$$

   (b) Laplace density.

   $$h(f) = -\int_{-\infty}^\infty \frac{1}{2}\lambda e^{-\lambda|x|}[\ln\frac{1}{2} + \ln\lambda - \lambda|x|]dx = \log\frac{2e}{\lambda} \text{ bits.}$$

   (c) $X_1 + X_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$. Thus,

   $$h(f) = \frac{1}{2}\log 2\pi e(\sigma_1^2 + \sigma_2^2) \text{ bits.}$$

5. **Problem 9.3**
We have

$$
\begin{aligned}
h(X) &= \frac{1}{2}\log 2\pi e\sigma^2 \\
h(Y) &= \frac{1}{2}\log 2\pi e\sigma^2 \\
h(X,Y) &= \frac{1}{2}\log(2\pi e)^2|K| = \frac{1}{2}\log(2\pi e)^2\sigma^4(1-\rho^2)
\end{aligned}
$$

Thus,

$$I(X;Y) = h(X) + h(Y) - h(X,Y) = -\frac{1}{2}\log(1-\rho^2).$$

For $\rho = 0$, X and Y are independent. Thus, $I(X;Y) = 0$. For $\rho = 1$ (or $\rho = -1$), $X = Y$ (or $X = -Y$). Knowing X, we have perfect knowledge of Y. Thus, $I(X;Y) = \infty$.