EL606: Homework 2 – Solutions

Due date: Sept. 28, 2004

1. *Problem 10, Ch. 2, Cover&Thomas*

   (a) The last corollary to Theorem 2.8.1 in the text states that if $X \to Y \to Z$, then $I(X;Y) \geq I(X;Y|Z)$. Equality holds iff $I(X;Z) = 0$ *i.e.,* X and Z are independent. A simple example of random variables satisfying the inequality conditions above is the following: X is a fair binary random variable and $Y = X$ and $Z = Y$. In this case,

   $$I(X;Y) = H(X) - H(X|Y) = H(X) = 1$$

   and

   $$I(X;Y|Z) = H(X|Z) - H(X|Y,Z) = 0.$$

   Therefore, $I(X;Y) > I(X;Y|Z)$.

   (b) This example is also given in the text. Let X, Y be independent fair binary random variables and let $Z = X + Y$. In this case we have that

   $$I(X;Y) = 0 \text{ and } I(X;Y|Z) = H(X|Z) = 1/2.$$

   Therefore, $I(X;Y) < I(X;Y|Z)$.

2. *Problem 21, Ch. 2, Cover&Thomas*
   By the chain rule for mutual information,

   $$I(X_1; X_2, \ldots, X_n) = I(X_1; X_2) + I(X_1; X_3|X_2) + \cdots + I(X_1; X_n|X_2, \ldots, X_{n-2}).$$

   By the Markov property, the past and the future are conditionally independent given the present and hence all terms except the first are zero. Therefore

   $$I(X_1; X_2, \ldots, X_n) = I(X_1; X_2).$$

3. *Problem 2, Ch. 3, Cover&Thomas*
   Notice that

   $$p(X_1, X_2, \ldots, X_n)^{\frac{1}{n}} = \exp\Big[\frac{1}{n}\log[p(X_1, X_2, \ldots, X_n)\Big].$$

   It follows from Theorem 3.1.1 that

   $$\lim(p(X_1, X_2, \ldots, X_n))^{\frac{1}{n}} = e^{-H(X)} \text{ in probability.}$$

4. *Problem 3, Ch. 3, Cover&Thomas*

   (a) The number of 100-bit binary sequences with three or fewer ones is

   $$\binom{100}{0} + \binom{100}{1} + \binom{100}{2} + \binom{100}{3} = 1 + 100 + 4950 + 161700 = 166751\,.$$

   The required codeword length is $\lceil \log_2 166751 \rceil = 18$. (Note that $H(0.005) = 0.0454$, so 18 is quite a bit larger than the 4.5 bits of entropy.)

(b) The probability that a 100-bit sequence has three or fewer ones is

$$\sum_{i=0}^{3} \binom{100}{i} (0.005)^i (0.995)^{100-i} = 0.60577 + 0.30441 + 0.7572 + 0.01243 = 0.99833$$

Thus the probability that the sequence that is generated cannot be encoded is $1 - 0.99833 = 0.00167$.

(c) In the case of a random variable S_n that is the sum of n i.i.d. random variables $X_1, X_2, \ldots, X_n$, Chebyshev's inequality states that

$$\Pr(|S_n - n\mu| \geq \epsilon) \leq \frac{n\sigma^2}{\epsilon^2},$$

where μ and σ^2 are the mean and variance of X_i. In this problem, $n = 100$, $\mu = 0.005$, and $\sigma^2 = 0.005 \cdot 0.995$. Note that $S_{100} \geq 4$ if and only if $|S_{100} - 100(0.005)| \geq 3.5$, so we should choose $\epsilon = 3.5$. Then

$$\Pr(S_{100} \geq 4) \leq \frac{100(0.005)(0.995)}{(3.5)^2} \approx 0.04061\,.$$

This bound is much larger than the actual probability 0.00167.

5. *Problem 4, Ch. 3, Cover&Thomas*
Let

$$Y_n = (X_1 X_2 \ldots X_n)^{\frac{1}{n}}\,.$$

We have $Y_n \to 2^{E \log X}$ in probability (similarly as in Problem. 2 of this homework). We calculate $E \log X = \frac{1}{2}\log 1 + \frac{1}{4}\log 2 + \frac{1}{4}\log 3 = \frac{1}{4}\log 6$, and therefore

$$Y_n \to 2^{\frac{1}{4}\log 6} = 1.565.$$

6. *Problem 5, Ch. 3, Cover&Thomas*

(a) Since the $X_1, X_2, \ldots, X_n$ are i.i.d., so are $q(X_1), q(X_2), \ldots, q(X_n)$, and hence we can apply the law of large numbers to obtain

$$\begin{aligned}
\lim -\frac{1}{n} \log q(X_1, X_2, \ldots, X_n) &= \lim -\frac{1}{n} \sum \log q(X_i) \\
&= -E(\log q(X)) \\
&= -\sum p(x) \log q(x) \\
&= \sum p(x) \log \frac{p(x)}{q(x)} - \sum p(x) \log p(x) \\
&= D(\boldsymbol{p}||\boldsymbol{q}) + H(\boldsymbol{p}).
\end{aligned}$$

(b) Again, by the law of large numbers,

$$\begin{aligned}
\lim -\frac{1}{n}\log\frac{q(X_1,X_2,\ldots,X_n)}{p(X_1,X_2,\ldots,X_n)} &= \lim -\frac{1}{n}\sum\log\frac{q(X_i)}{p(X_i)} \\
&= -E\Big[\log\frac{q(X)}{p(X)}\Big] \\
&= -\sum p(x)\log\frac{q(x)}{p(x)} \\
&= \sum p(x)\log\frac{p(x)}{q(x)} \\
&= D(\boldsymbol{p}||\boldsymbol{q}).
\end{aligned}$$

7. *Problem 6, Ch. 3, Cover&Thomas*
By Problem 5 of this homework, we have

$$\lim_{n\to\infty} V_n^{\frac{1}{n}} = e^{E\ln X}.$$

Now,

$$E\ln X = \int_0^1 \ln x dx = -1,$$

and thus the "effective" edge length of the box is

$$\lim_{n\to\infty} V_n^{\frac{1}{n}} = e^{E\ln X} = e^{-1}.$$

To compute the expected volume of the box, note that the X_i's are independent. Therefore,

$$E(V_n) = \prod E(X_i) = \frac{1}{2^n} > \frac{1}{e^n}.$$