ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
School of Computer and Communication Sciences

Handout 8 Information Theory and Coding
Solutions to homework 4 October 26, 2007

PROBLEM 1. Let $\mathcal{H}(p) = -p\log p - (1-p)\log p$ denote the entropy of a binary valued random variable with distribution $p, 1-p$. The entropy per symbol of the source is

$$\mathcal{H}(p_1) = -p_1 \log p_1 - (1-p_1)\log(1-p_1)$$

and the average symbol duration (or time per symbol) is

$$T(p_1) = 1 \cdot p_1 + 2 \cdot p_2 = p_1 + 2(1-p_1) = 2 - p_1 = 1 + p_2\,.$$

Therefore the source entropy per unit time is

$$f(p_1) = \frac{\mathcal{H}(p_1)}{T(p_1)} = \frac{-p_1\log p_1 - (1-p_1)\log(1-p_1)}{2-p_1}\,.$$

Since $f(0) = f(1) = 0$, the maximum value of $f(p_1)$ must occur for some point p_1 such that $0 < p_1 < 1$ and $\partial f/\partial p_1 = 0$.

$$\frac{\partial}{\partial p_1}\frac{\mathcal{H}(p_1)}{T(p_1)} = \frac{T(\partial\mathcal{H}/\partial p_1) - \mathcal{H}(\partial T/\partial p_1)}{T^2}$$

After some calculus, we find that the numerator of the above expression (assuming natural logarithms) is

$$T(\partial H/\partial p_1) - H(\partial T/\partial p_1) = \ln(1-p_1) - 2\ln p_1\,,$$

which is zero when $1 - p_1 = p_1^2 = p_2$, that is, $p_1 = \frac{1}{2}(\sqrt{5}-1) = 0.61803$, the reciprocal of the golden ratio, $\frac{1}{2}(\sqrt{5}+1) = 1.61803$. The corresponding entropy per unit time is

$$\frac{\mathcal{H}(p_1)}{T(p_1)} = \frac{-p_1\log p_1 - p_1^2\log p_1^2}{2-p_1} = \frac{-(1+p_1^2)\log p_1}{1+p_1^2} = -\log p_1 = 0.69424 \text{ bits.}$$

PROBLEM 2.

(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 the entropy of the source is $-0.005\log_2(0.005) - 0.995\log_2(0.995) = 0.0454$ bits, so 18 is quite a bit larger than the 4.5 bits of entropy per 100 source letters.)

(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

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

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

PROBLEM 3. The volume $V_n = \prod_{i=1}^n X_i$ is a random variable, since the X_i are random variables uniformly distributed on $[0,1]$. V_n tends to 0 as $n \to \infty$. However

$$\log_e V_n^{1/n} = \frac{1}{n}\log_e V_n = \frac{1}{n}\sum \log_e X_i \to E[\log_e(X)] \text{ a.e.}$$

by the Strong Law of Large Numbers, since X_i and $\log_e(X_i)$ are i.i.d. and $E[\log_e(X)] < \infty$. Now

$$E[\log_e(X_i)] = \int_0^1 \log_e(x)\,dx = -1$$

Hence, since e^x is a continuous function,

$$\lim_{n\to\infty} V_n^{1/n} = \exp\Big[\lim_{n\to\infty}\frac{1}{n}\log_e V_n\Big] = \frac{1}{e} < \frac{1}{2}.$$

Thus the "effective" edge length of this solid is e^{-1}. Note that since the X_i's are independent, $E(V_n) = \prod E(X_i) = (1/2)^n$.

PROBLEM 4.

By the strong law of large numbers,

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

PROBLEM 5. Let the random variable $X(i)$ represent the outcome of the i^{th} toss. The $X(i)$'s are i.i.d. (distribution p) random variables taking values in $\{1 \ldots K\}$. The capital C_n after the n^{th} toss is related to the capital C_{n-1} after the $(n-1)^{\text{th}}$ toss as

$$C_n = C_{n-1}\frac{f(X(n))}{q(X(n))}$$

Using the above relation recursively, C_n can be expressed in terms of C_0 as

$$C_n = C_0\prod_{i=1}^{n}\frac{f(X(i))}{q(X(i))}$$

(a) The "long term" rate of return $r = \lim_{n\to\infty} R_n$ is given by

$$
\begin{aligned}
\lim_{n\to\infty} R_n &= \lim_{n\to\infty} \frac{1}{n} \log \frac{C_n}{C_0} \\
&= \lim_{n\to\infty} \frac{1}{n} \log \prod_{i=1}^{n} \frac{f(X(i))}{q(X(i))} \\
&= E \log \frac{f(X)}{q(X)} \qquad \text{a.s} \\
&= \sum_{i=1}^{K} p(x) \log \frac{f(x)}{q(x)} \\
&= \sum_{i=1}^{K} p(x) (\log \frac{p(x)}{q(x)} + \log \frac{f(x)}{p(x)}) \\
&= D(p\|q) - D(p\|f)
\end{aligned}
$$

where X is a random variable corresponding to the outcome of a toss. The third equality follows from the strong law of large numbers.

(b) Note that only the second divergence term depends on f and is minimum when $f = p$. Therefore the gambler maximizes r by choosing $f = p$ and this maximal $r = D(p\|q)$. Note that the maximal long term return is positive as long as the casinos odds are different from the odds implied by the true distribution.