ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
School of Computer and Communication Sciences

Handout 7 Information Theory and Coding
Homework 4 October 19, 2007

PROBLEM 1. A discrete memoryless source has alphabet $1, 2$, where symbol 1 has duration 1 and symbol 2 has duration 2. The probabilities of 1 and 2 are p_1 and p_2 respectively. Find the value of p_1 that maximizes the source entropy per unit time, $H(X)/E[l_X]$, where l_x is the duration of the symbol x. What is the maximum value of the entropy per unit time?

PROBLEM 2. A discrete memoryless source emits a sequence of statistically independent binary digits with probabilities $p(1) = 0.005$ and $p(0) = 0.995$. The digits are taken 100 at a time, and a binary codeword is provided for every sequence of 100 digits containing three or fewer 1's.

(a) Assuming that all the codewords are the same length, find the minimum length required to provide codewords to all sequences with three or fewer ones.

(b) Calculate the probability of observing a source sequence for which no codeword has been assigned.

(c) Use Chebyshev's inequality to bound the probability of observing a source sequence for which no codewords has been assigned. Compare this bound to the actual probability computed in part (b).

PROBLEM 3. An n-dimensional rectangular box with sides $X_1, \dots, X_n$ is to be constructed. The volume is $V_n = \prod_{i=1}^n X_n$. The edge length l of an n-dimensional cube of the same volume is $l = V_n^{1/n}$. Let $X_1, \dots, X_n$ be i.i.d. random variables uniformly distributed over the unit interval $[0, 1]$. Find $\lim_{n\to\infty} V_n^{1/n}$ and compare it to $(E[V_n])^{1/n}$.

PROBLEM 4. Let $X_1, X_2, \dots$ be i.i.d. random variables with distribution $p(x)$ taking values in a finite set $\mathcal{X}$. Thus, $p(x_1, \dots, x_n) = \prod_{i=1}^n p(x_i)$. We know that

$$-\frac{1}{n} \log p(X_1, \dots, X_n) \to H(X)$$

in probability. Let $q(x_1, \dots, x_n) = \prod_{i=1}^n q(x_i)$, where $q(x)$ is some other probability distribution on $\mathcal{X}$.[Note: The random variable X_i has distribution $p(x)$]

(b) Evaluate the limit of the log-likelihood-ratio

$$\frac{1}{n} \log \frac{q(X_1, \dots, X_n)}{p(X_1, \dots, X_n)}.$$

PROBLEM 5. In a casino one can bet on the outcome of a random variable X taking values in $\{1, \dots, K\}$. If $X = k$, the casino multiplies the money bet on outcome k by $\frac{1}{q(k)}$, the money bet on other outcomes are lost. The values $q(k)$ represent the casino's belief of the probability of outcome k, and thus $\sum_k q(k) = 1, q(k) > 0$. Suppose that the true probability of $X = k$ is $p(k)$ and these true probabilities are known to the gambler. Let

$$R_n = \frac{1}{n} \log \frac{C_n}{C_0}$$

be the "rate of return" for the gambler where C_0 is the gambler's initial capital and C_n is the capital after playing n successive independent rounds of the game. The gambler's strategy is to allocate his current capital among different bets, betting a fraction f_k of his capital on outcome k.

(a) Use the law of large numbers to find

$$r = \lim_{n \to \infty} R_n$$

in terms of f, p, q.

(b) Find the f that maximizes the "long term rate of return" r, and the value of r.