ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
School of Computer and Communication Sciences

Handout 8 Information Theory and Coding
Homework 4 October 11, 2011

PROBLEM 1. Let X^i denote $X_1, \dots, X_i$.

(a) By the chain rule for entropy,

$$\frac{H(X_1, X_2, \dots, X_n)}{n} = \frac{\sum_{i=1}^{n} H(X_i|X^{i-1})}{n} \tag{1}$$
$$= \frac{H(X_n|X^{n-1}) + \sum_{i=1}^{n-1} H(X_i|X^{i-1})}{n} \tag{2}$$
$$= \frac{H(X_n|X^{n-1}) + H(X_1, X_2, \dots, X_{n-1})}{n}. \tag{3}$$

From stationarity it follows that for all $1 \le i \le n$,

$$H(X_n|X^{n-1}) \le H(X_i|X^{i-1}),$$

which further implies, by summing both sides over $i = 1, \dots, n-1$ and dividing by $n-1$, that,

$$H(X_n|X^{n-1}) \le \frac{\sum_{i=1}^{n-1} H(X_i|X^{i-1})}{n-1} \tag{4}$$
$$= \frac{H(X_1, X_2, \dots, X_{n-1})}{n-1}. \tag{5}$$

Combining (3) and (5) yields,

$$\frac{H(X_1, X_2, \dots, X_n)}{n} \le \frac{1}{n}\left[\frac{H(X_1, X_2, \dots, X_{n-1})}{n-1} + H(X_1, X_2, \dots, X_{n-1})\right] \tag{6}$$
$$= \frac{H(X_1, X_2, \dots, X_{n-1})}{n-1}. \tag{7}$$

(b) By stationarity we have for all $1 \le i \le n$,

$$H(X_n|X^{n-1}) \le H(X_i|X^{i-1}),$$

which implies that,

$$H(X_n|X^{n-1}) = \frac{\sum_{i=1}^{n} H(X_n|X^{n-1})}{n} \tag{8}$$
$$\le \frac{\sum_{i=1}^{n} H(X_i|X^{i-1})}{n} \tag{9}$$
$$= \frac{H(X_1, X_2, \dots, X_n)}{n}. \tag{10}$$

PROBLEM 2. By the chain rule for entropy,

$$H(X_0|X_{-1}, \dots, X_{-n}) = H(X_0, X_{-1}, \dots, X_{-n}) - H(X_{-1}, \dots, X_{-n}) \tag{11}$$
$$= H(X_0, X_1, \dots, X_n) - H(X_1, \dots, X_n) \tag{12}$$
$$= H(X_0|X_1, \dots, X_n), \tag{13}$$

where (12) follows from stationarity.

Problem 3. Let $h_2(p) = -p \log p - (1-p)\log p$ denote the entropy of a binary valued random variable with distribution $p, 1-p$. The entropy per symbol of the source is

$$h_2(p_1) = -p_1 \log p_1 - (1-p_1)\log(1-p_1)$$

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

$$f(p_1) = \frac{h_2(p_1)}{T(p_1)} = \frac{-p_1 \log p_1 - (1-p_1)\log(1-p_1)}{2-p_1}\,.$$

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 $\partial f / \partial p_1 = 0$.

$$\frac{\partial}{\partial p_1}\frac{h_2(p_1)}{T(p_1)} = \frac{T(\partial h_2/\partial p_1) - h_2(\partial T/\partial p_1)}{T^2}$$

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

$$\frac{h_2(p_1)}{T(p_1)} = \frac{-p_1 \log p_1 - p_1^2 \log p_1^2}{2-p_1} = \frac{-(1+p_1^2)\log p_1}{1+p_1^2} = -\log p_1 = 0.69424 \text{ bits.}$$

Problem 4. (a) We can write the following chain of inequalities:

$$\begin{aligned}
Q^n(\mathbf{x}) &\overset{1}{=} \prod_{i=1}^{n} Q(x_i) \overset{2}{=} \prod_{a\in\mathcal{X}} Q(a)^{N(a|\mathbf{x})} \overset{3}{=} \prod_{a\in\mathcal{X}} Q(a)^{nP_{\mathbf{x}}(a)} = \prod_{a\in\mathcal{X}} 2^{nP_{\mathbf{x}}(a)\log Q(a)} & (14)\\
&= \prod_{a\in\mathcal{X}} 2^{n(P_{\mathbf{x}}(a)\log Q(a) - P_{\mathbf{x}}(a)\log P_{\mathbf{x}}(a) + P_{\mathbf{x}}(a)\log P_{\mathbf{x}}(a))} & (15)\\
&= 2^{n\sum_{a\in\mathcal{X}}(-P_{\mathbf{x}}(a)\log\frac{P_{\mathbf{x}}(a)}{Q(a)} + P_{\mathbf{x}}(a)\log P_{\mathbf{x}}(a))} = 2^{n(-D(P_{\mathbf{x}}||Q) + H(P_{\mathbf{x}}))},
\end{aligned}$$

where 1 follows because the sequence is i.i.d., grouping symbols gives 2, and 3 is the definition of type.

(b) Upper bound: We know that

$$\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = 1.$$

Consider one term and set $p = k/n$. Then,

$$1 \ge \binom{n}{k}\left(\frac{k}{n}\right)^k \left(1 - \frac{k}{n}\right)^{n-k} = \binom{n}{k} 2^{n\left(\frac{k}{n}\log\frac{k}{n} + \frac{n-k}{n}\log\frac{n-k}{n}\right)} = \binom{n}{k} 2^{-nh_2\left(\frac{k}{n}\right)}$$

Lower bound: Define $S_j = \binom{n}{j} p^j (1-p)^{n-j}$. We can compute

$$\frac{S_{j+1}}{S_j} = \frac{n-j}{j+1} \frac{p}{1-p}.$$

One can see that this ratio is a decreasing function in j. It equals 1, if $j = np + p - 1$, so $\frac{S_{j+1}}{S_j} < 1$ for $j = \lfloor np + p \rfloor$ and $\frac{S_{j+1}}{S_j} \geq 1$ for any smaller j. Hence, S_j takes its maximum value at $j = \lfloor np + p \rfloor$, which equals k in our case. From this we have that

$$\begin{aligned} 1 = \sum_{j=0}^{n} \binom{n}{j} p^j (1-p)^{n-j} &\leq (n+1) \max_j \binom{n}{j} p^j (1-p)^j \\ &\leq (n+1) \binom{n}{k} \left(\frac{k}{n}\right)^k \left(1 - \frac{k}{n}\right)^{n-k} = \binom{n}{k} 2^{-nh_2(\frac{k}{n})}. \end{aligned} \tag{16}$$

The last equality comes from the derivation we had when proving the upper bound.

PROBLEM 5. (a)

$$E[F_n] = E[F_0 X_0 X_1 \dots X_n] = F_0 (E[X_1])^n = F_0 (9/8)^n$$

We exploited the i.i.d. property of the sequence. One can see that $E[F_n] \to \infty$ with $n \to \infty$.

(b)

$$\begin{aligned} l_n = E[\log_2 F_n] = E[\log_2 (F_0 X_0 X_1 \dots X_n)] = E\left[\log_2 F_0 + \sum_{i=1}^{n} \log_2 X_i\right] = \\ = E[\log_2 F_0] + nE[\log_2 X_1] = \log_2 F_0 - \frac{n}{2}. \end{aligned} \tag{17}$$

(c) It concentrates around 2^{l_n}. F_n in itself is not a sum of i.i.d. variables. Taking its logarithm results such a sum, so the law of large numbers applies.

$$\log_2 F_n = \log_2 F_0 + \sum_{i=1}^{n} \log_2 X_i \to \log_2 F_0 + nE[\log X_1] = \log_2 F_0 - \frac{n}{2}.$$

(d) From the previous result it follows that although it seems appealing that the expected value of our fortune goes to infinity, it actually converges to 0 (very rapidly).

(e) If we keep a proportion r of the money in reserve at each play, we change the evolution of the game as follows: starting from a unit fortune, our fortune become $r + (1-r)X$. Thus, this is equivalent to replacing X with $r + (1-r)X$. Consequently, the quantity to maximize is

$$E[\log_2(r+(1-r)X)] = \frac{1}{2}\log(r+(1-r)2) + \frac{1}{2}\log(r+(1-r)/4) = \frac{1}{2}[\log(2-r) + \log(1+3r) - \log 4].$$

By differentiation, the maximum appears at the value of r that solves

$$\frac{1}{2-r} = \frac{3}{1+3r},$$

which is $\frac{5}{6}$; with a growth rate $\frac{1}{2}\log\frac{49}{48}$. Note that now, the fortune does grow to infinity almost surely.