Sample problems for the 12th lecture (May 31)

1. Let

$$P_{X^3}(x^3) = \begin{cases} \frac{1}{20}, & x^3 = 000 \\ \frac{1}{20}, & x^3 = 001 \\ \frac{1}{10}, & x^3 = 010 \\ \frac{1}{10}, & x^3 = 011 \\ \frac{3}{20}, & x^3 = 100 \\ \frac{3}{20}, & x^3 = 101 \\ \frac{1}{5}, & x^3 = 110 \\ \frac{1}{5}, & x^3 = 111 \end{cases}$$

Define $h_{X^3}(x^3) \triangleq \log \frac{1}{P_{X^3}(x^3)}$.

(a) List the distribution of the entropy density $h_{X^3}(X^3)$.

(b) Is it true that $h_{X^3}(u^3) > h_{X^3}(v^3)$ iff $P_{X^3}(u^3) < P_{X^3}(v^3)$? Justify your answer.

(c) Define $\mathcal{D}(R) \triangleq \{x^3 \in \{0,1\}^3 : \frac{1}{3} h_{X^3}(x^3) \leq R\}$. Find the range of R such that $\Pr[X^3 \in \mathcal{D}(R)] = 0.7$.

Solution.

(a)

$$h_{X^n}(X^n) = \begin{cases} \log(20) & \text{with probability } \frac{1}{10} = 0.1 \\ \log(10) & \text{with probability } \frac{1}{5} = 0.2 \\ \log(20/3) & \text{with probability } \frac{3}{10} = 0.3 \\ \log(5) & \text{with probability } \frac{2}{5} = 0.4 \end{cases}$$

(b) Yes since

$$\begin{aligned} & h_{X^3}(u^3) > h_{X^3}(v^3) \\ \Leftrightarrow \quad & \log \frac{1}{P_{X^n}(u^3)} > \log \frac{1}{P_{X^n}(v^3)} \\ \Leftrightarrow \quad & P_{X^n}(u^3) < P_{X^n}(v^3) \end{aligned}$$

(c) $\log(20/3) \leq R < \log(10)$

2. (a) Prove that the number of x^n's satisfying $P_{X^n}(x^n) \geq \frac{1}{N}$ is at most N.

(b) Let $\mathcal{C}_n^*$ be the set that maximizes $\Pr[X^n \in \mathcal{C}_n^*]$ among all sets of the same size M_n. Prove that

$$\Pr[X^n \notin \mathcal{C}_n^*] \leq \Pr\left[\frac{1}{n} h_{X^n}(X^n) > \frac{1}{n}\log M_n\right].$$

Hint: Use (a).

(c) Let $\mathcal{C}_n$ be a subset of $\mathcal{X}^n$, satisfying that $|\mathcal{C}_n| = M_n$. Prove that for every $\gamma > 0$,

$$\Pr[X^n \notin \mathcal{C}_n] \geq \Pr\left[\frac{1}{n} h_{X^n}(X^n) > \frac{1}{n}\log M_n + \gamma\right] - \exp\{-n\gamma\},$$

where $h_{X^n}(X^n)$ is defined in Problem 1.

Solution.

(a) This can be proved by contradiction. Suppose there are $N+1$ x^n's satisfying $P_{X^n}(x^n) \geq \frac{1}{N}$. Then, $1 = \sum_{x^n \in \mathcal{X}^n} P_{X^n}(x^n) \geq \frac{N+1}{N}$, which is a contradiction. Hence, the number of x^n's satisfying $P_{X^n}(x^n) \geq \frac{1}{N}$ is at most N.

(b) From (a), the number of x^n's satisfying $P_{X^n}(x^n) \geq \frac{1}{M_n}$ is at most M_n. Since $\mathcal{C}_n^*$ should consist of M_n words with larger probabilities, we have

$$\begin{aligned}
\Pr[X^n \in \mathcal{C}_n^*] &\geq \Pr\left[P_{X^n}(X^n) \geq \frac{1}{M_n}\right] \\
&= \Pr\left[\frac{1}{n}\log\frac{1}{P_{X^n}(X^n)} \leq \frac{1}{n}\log M_n\right] \\
&= \Pr\left[\frac{1}{n} h_{X^n}(X^n) \leq \frac{1}{n}\log M_n\right],
\end{aligned}$$

which implies

$$\Pr[X^n \notin \mathcal{C}_n^*] \leq \Pr\left[\frac{1}{n} h_{X^n}(X^n) > \frac{1}{n}\log M_n\right].$$

(c) It suffices to prove that

$$\begin{aligned}
\Pr\left[X^n \in \mathcal{C}_n\right] &\leq \Pr\left[\frac{1}{n} h_{X^n}(X^n) \leq \frac{1}{n}\log M_n + \gamma\right] + \exp\{-n\gamma\} \\
&= \Pr\left[\frac{1}{n}\log\frac{1}{P_{X^n}(X^n)} \leq \frac{1}{n}\log M_n + \gamma\right] + \exp\{-n\gamma\} \\
&= \Pr\left[P_{X^n}(X^n) \geq \frac{1}{M_n} e^{-n\gamma}\right] + \exp\{-n\gamma\}.
\end{aligned}$$

We then derive

$$\begin{aligned}
\Pr\left[X^n \in \mathcal{C}_n\right] &= \Pr\left[X^n \in \mathcal{C}_n \text{ and } P_{X^n}(X^n) \geq \frac{1}{M_n}\, e^{-n\gamma}\right] \\
&\quad + \Pr\left[X^n \in \mathcal{C}_n \text{ and } P_{X^n}(X^n) < \frac{1}{M_n}\, e^{-n\gamma}\right] \\
&\leq \Pr\left[P_{X^n}(X^n) \geq \frac{1}{M_n}\, e^{-n\gamma}\right] \\
&\quad + \Pr\left[X^n \in \mathcal{C}_n \text{ and } P_{X^n}(X^n) < \frac{1}{M_n}\, e^{-n\gamma}\right] \\
&= \Pr\left[P_{X^n}(X^n) \geq \frac{1}{M_n}\, e^{-n\gamma}\right] \\
&\quad + \sum_{x^n \in \mathcal{C}_n} P_{X^n}(x^n) \cdot \mathbf{1}\left\{P_{X^n}(x^n) < \frac{1}{M_n}\, e^{-n\gamma}\right\} \\
&< \Pr\left[P_{X^n}(X^n) \geq \frac{1}{M_n}\, e^{-n\gamma}\right] \\
&\quad + \sum_{x^n \in \mathcal{C}_n} \frac{1}{M_n}\, e^{-n\gamma} \cdot \mathbf{1}\left\{P_{X^n}(x^n) < \frac{1}{M_n}\, e^{-n\gamma}\right\} \\
&= \Pr\left[P_{X^n}(X^n) \geq \frac{1}{M_n}\, e^{-n\gamma}\right] + |\mathcal{C}_n| \frac{1}{M_n} e^{-n\gamma} \\
&= \Pr\left[P_{X^n}(X^n) \geq \frac{1}{M_n}\, e^{-n\gamma}\right] + e^{-n\gamma}.
\end{aligned}$$

3. Let $\{\mathcal{C}_n^*\}_{n=1}^{\infty}$ be a sequence of the optimal codes that satisfy i) $\mathcal{C}_n^* \subset \mathcal{X}^n$; ii) $\Pr[X^n \in \mathcal{C}_n^*]$ is maximized among codes of the same size; and iii) $\lim_{n\to\infty} \frac{1}{n} \log |\mathcal{C}_n^*| = R$. Show that

$$\lim_{n\to\infty} \Pr[X^n \notin \mathcal{C}_n^*] = \Lambda(R),$$

provided that

$$\Lambda(x) \triangleq \lim_{n\to\infty} \Pr\left[\frac{1}{n} h_{X^n}(X^n) > x\right]$$

is continuous at $x = R$.

Hint: Use Problems 2(b) and 2(c).

Solution. $\lim_{n\to\infty} \frac{1}{n} \log |\mathcal{C}_n^*| = R$ implies that for every $\delta > 0$,

$$R - \delta < \frac{1}{n} \log |\mathcal{C}_n^*| < R + \delta \text{ for sufficiently large } n.$$

Hence, from Problem 2(b), we have

$$\begin{aligned}\Pr[X^n \notin \mathcal{C}_n^*] &\le \Pr\left[\frac{1}{n}h_{X^n}(X^n) > \frac{1}{n}\log|\mathcal{C}_n^*|\right]\\ &\le \Pr\left[\frac{1}{n}h_{X^n}(X^n) > R-\delta\right].\end{aligned}$$

Problem 2(c) additionally gives

$$\begin{aligned}\Pr[X^n \notin \mathcal{C}_n^*] &\ge \Pr\left[\frac{1}{n}h_{X^n}(X^n) > \frac{1}{n}\log|\mathcal{C}_n^*|+\gamma\right] - e^{-n\gamma}\\ &\ge \Pr\left[\frac{1}{n}h_{X^n}(X^n) > R+\delta+\gamma\right] - e^{-n\gamma}.\end{aligned}$$

These two inequalities indicate

$$\begin{aligned}\Lambda(R+\delta+\gamma) &= \liminf_{n\to\infty}\Pr\left[\frac{1}{n}h_{X^n}(X^n) > R+\delta+\gamma\right]\\ &\le \liminf_{n\to\infty}\Pr[X^n\notin\mathcal{C}_n^*] \le \limsup_{n\to\infty}\Pr[X^n\notin\mathcal{C}_n^*]\\ &\le \limsup_{n\to\infty}\Pr\left[\frac{1}{n}h_{X^n}(X^n) > R-\delta\right] = \Lambda(R-\delta).\end{aligned}$$

Since $\Lambda(x)$ is continuous at $x = R$, we can take $\delta \downarrow 0$ and $\gamma \downarrow 0$ to obtain

$$\lim_{n\to\infty}\Pr[X^n\notin\mathcal{C}_n^*] = \Lambda(R).$$

4. Continue from Problem 3.

(a) If

$$\Lambda(R) = \begin{cases}1, & 0 \le R < 0.1;\\ 2-10R, & 0.1 \le R < 0.2;\\ 0, & R \ge 0.2,\end{cases}$$

what is the minimum (asymptotic) data compression rate (in nats per source letter) for a sequence of codes with decompression error no larger than 0.1?

(b) If $X_1, X_2, X_3, \ldots$ is an i.i.d. sequence of random variables, what is $\Lambda(R)$ for this sequence? Is this $\Lambda(R)$ continuous for $R > H(X)$? Is this $\Lambda(R)$ continuous for $R < H(X)$?

Solution.

(a)

$$\begin{aligned}\min\{R : \Lambda(R) \le 0.1\} &\Leftrightarrow \min\{R : 2 - 10R \le 0.1\}\\ &\Leftrightarrow \min\{R : 10R \ge 1.9\} = 0.19\end{aligned}$$

(b) Since

$$\frac{\log\frac{1}{P_X(X_1)} + \cdots + \log\frac{1}{P_X(X_n)}}{n}$$

converges to $E[\log\frac{1}{P_X(X)}] = H(X)$ in probability, we have

$$\Lambda(R) = \begin{cases} 1, & R < H(X); \\ 0, & R > H(X) \end{cases}.$$

Apparently, $\Lambda(R)$ is only discontinuous at $R = H(X)$.