ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
School of Computer and Communication Sciences

Handout 26 Information Theory and Coding
Graded Homework, due Monday, Dec. 18, 2023 Dec. 4, 2023

You are allowed (even encouraged) to discuss the problems on the homework with your colleagues. However, your solutions should be in your own words. If you collaborated on your solution, write down the name of your collaborators and your sources; no points will be deducted. But similarities in solutions beyond the listed collaborations will be considered as malpractice.

PROBLEM 1. Suppose (U_1, V) is a pair of random variables with distribution p_{UV}, and suppose $U_2, \dots, U_m$ are i.i.d. random variables with distribution p_U, independent of (U_1, V).

Let $\text{score}(u, v) := p_{V|U}(v|u)$, and let $S_i = \text{score}(U_i, V)$. For $i = 2, \dots, m$, let $B_i = \mathbb{1}\{S_i \geq S_1\}$, and let $L = \sum_{i=2}^m B_i$. Note that the event $\{L \geq 1\}$ includes the event $\{S_1 \text{ is not the highest score}\}$.

(a) Show that for any $r \geq 0$ and $i \geq 2$,

$$\mathbb{E}[B_i \mid U_1 = u_1, V = v] \leq \sum_u p_U(u) \left[\frac{p_{V|U}(v|u)}{p_{V|U}(v|u_1)} \right]^r.$$

Hint: For non-negative a, b, r, the inequality $\mathbb{1}\{a \geq b\} \leq (a/b)^r$ holds.

(b) For $i \geq 2$, show that $\mathbb{E}[B_i] \leq \sum_v \left[\sum_u p_U(u) \sqrt{p_{V|U}(v|u)} \right]^2$.
Hint: Use (a) with a careful choice of r.

(c) Show that

$$\Pr(S_1 \text{ is not the highest score}) \leq (m-1) \sum_v \left[\sum_u p_U(u) \sqrt{p_{V|U}(v|u)} \right]^2.$$

Hint: $\Pr(L \geq 1) \leq \mathbb{E}[L]$.

Define $R_1(p_U, p_{V|U}) := -\log \sum_v \left[\sum_u p_U(u) \sqrt{p_{V|U}(v|u)} \right]^2$.

(d) With $p_{X^n}(x^n) = \prod_{i=1}^n p_X(x_i)$, and $p_{Y^n|X^n}(y^n|x^n) = \prod_{i=1}^n p_{Y|X}(y_i|x_i)$, show that $R_1(p_{X^n}, p_{Y^n|X^n}) = nR_1(p_X, p_{Y|X})$.

(e) Given a channel $p_{Y|X}$ and input distribution p_X, show that for every $0 \leq R < R_1(p_X, p_{Y|X}) =: R_1$, and positive integer n, there is a code with $m = \lceil 2^{nR} \rceil$ codewords and with average probability of error $\bar{p}_e \leq 2^{-n(R_1 - R)}$.
Hint: Choose m codewords $X^n(1), \dots, X^n(m)$, i.i.d. from distribution p_{X^n}. Make use of what you already showed in (d) and (c).

(f) With $p_{Y|X}$ being the Binary Erasure Channel and for p_X the uniform distribution on the input alphabet, compute and sketch R_1 (defined above) and C (the channel capacity) as a function of the erasure probability. Comment on the plots obtained.

(g) Continuing with the notation of (a)–(c), for any $r \geq 0$, and for any $0 \leq \rho \leq 1$, show that

$$\mathbb{E}[L^\rho \mid U_1 = u_1, V = v] \leq (m-1)^\rho \bigg(\sum_u p_U(u) \left[\frac{p_{V|U}(v|u)}{p_{V|U}(v|u_1)} \right]^r \bigg)^\rho.$$

Hint: Use of the bound you found in (a) to upper bound $\mathbb{E}[L \mid U_1 = u_1, V = v]$; note that $z \in [0, \infty) \mapsto z^\rho$ is concave, so, $\mathbb{E}[Z^\rho] \leq \mathbb{E}[Z]^\rho$.

(h) For any $0 \leq \rho \leq 1$, and $r \geq 0$, show that

$$\mathbb{E}[L^\rho] \leq (m-1)^\rho \sum_v \Big[\sum_{u'} p_U(u') p_{V|U}(v|u')^{1-r\rho} \Big] \Big[\sum_u p_U(u) p_{V|U}(v|u)^r \Big]^\rho.$$

Hint: Use (g).

(i) For any $0 \leq \rho \leq 1$, show that

$$\mathbb{E}[L^\rho] \leq (m-1)^\rho \sum_v \Big[\sum_u p_U(u) p_{V|U}(v|u)^{1/(1+\rho)} \Big]^{1+\rho}.$$

Hint: Examine (h) for the choice $r = 1/(1+\rho)$.

For $0 < \rho \leq 1$, define $R_\rho(p_U, p_{V|U}) := -\rho^{-1} \log \sum_v \Big[\sum_u p_U(u) p_{V|U}(v|u)^{1/(1+\rho)} \Big]^{1+\rho}$. (Observe that setting $\rho = 1$ recovers R_1.)

(j) Given a channel $p_{Y|X}$ and input distribution p_X, show that for every $0 \leq R < R_\rho(p_X, p_{Y|X}) =: R_\rho$, positive integer n, there is a code with $m = \lceil 2^{nR} \rceil$ codewords and with average probability of error $\bar{p}_e \leq 2^{-n\rho(R_\rho - R)}$.
Hint: Observe that $\Pr(L \geq 1) \leq \mathbb{E}[L^\rho]$ and follow the reasoning in (d) and (e).

(k) Show that $\lim_{\rho \to 0^+} R_\rho(p_U, p_{V|U}) = I(U; V)$. Conclude from this and (j) that for any channel $p_{Y|X}$ and $R < C(p_{Y|X})$ there is a number $\beta > 0$ such that, for every positive integer n there is a code for the channel with $m = \lceil 2^{nR} \rceil$ codewords and with error probability $\bar{p}_e \leq 2^{-n\beta}$.