EE5139/EE6139 Homework 5

Semester I, AY2021/2022
Marco Tomamichel

Exercise 5.1 Min-entropy and Shannon entropy as Rényi entropies [EE5139]

Both the min-entropy and the Shannon entropy are limiting cases of the following family of Rényi entropies:

$$H_\alpha(X) = \frac{1}{1-\alpha} \log \sum_x P(x)^\alpha, \qquad \alpha \in (0,1) \cup (1, +\infty) \,. \tag{1}$$

a.) To verify this, compute the limit of the above quantities for $\alpha \to \{0_+, 1, +\infty\}$. (Here, by saying $\alpha \to 0_+$, we mean α "approaching 0 from right-hand side".)

b.) Plot the Rényi entropy as a function of α for the random variable X distributed as

x	0	1	2
$P(x)$	1/2	1/4	1/4

c.) Show that, for any random variable $X \in \mathcal{X}$ and any pmf $P(x)$, the Rényi entropy is monotonically non-increasing in the parameter α. Argue how this yields an alternative proof of the fact that $H_{\min}(X) \leq H(X) \leq \log |\mathcal{X}|$.

d.) Compute the min-entropy $H_{\min}(X|Y)$ of the joint random variables (X, Y) distributed as

$P(x,y)$		X		
		0	1	2
Y	0	1/6	1/12	1/12
	1	1/12	1/6	1/12
	2	1/12	1/12	1/6

Exercise 5.2 Distributions with a large entropy gap [all]

It is possible to construct distributions that have a large gap between min-entropy and Shannon entropy. This shows that controlling the Shannon entropy or the mutual information is not sufficient for most cryptographic tasks.

a.) Given $\epsilon \in (0,1)$, construct a sequence of random variables $(X_2,\ X_3,\ \ldots,\ X_n,\ \ldots)$ where $X_n \in \{0, 1, \ldots, n-1\}$, such that

$$\left.\begin{aligned} H(X_n) &\geq (1-\epsilon) \log n \\ H_{\min}(X_n) &= C, \end{aligned}\right\} \ \forall n \geq N$$

for some $N \in \mathbb{N}$ and some constant $C > 0$.

b.) Given $\epsilon \in (0,1)$, construct a sequence of random variables $((X_2, Y_2),\ (X_3, Y_3),\ \ldots,\ (X_n, Y_n),\ \ldots)$, where X_n, $Y_n \in \{0, 1, \ldots, n-1\}$, such that

$$H(X_n) = H_{\min}(X_n) = \log n \qquad \forall n$$

$$\left.\begin{aligned} H(X_n|Y_n) &\geq (1-\epsilon) \log n \\ H_{\min}(X_n|Y_n) &= C \end{aligned}\right\} \ \forall n \geq N$$

for some $N \in \mathbb{N}$ and some constant $C > 0$.

Exercise 5.3 Typical sets [all]

Consider a DMS with a two symbol alphabet $\{a, b\}$ where $p_X(a) = 2/3$ and $p_X(b) = 1/3$. Let $X^n = (X_1, \ldots, X_n)$ be a string of symbols emitted by the source with $n = 100,000$. Let $W(X_j)$ be the suprisal for the j-th source output, i.e., $W(X_j) = -\log 2/3$ for $X_j = a$ and $-\log 1/3$ for $X_j = b$. Define $W(X^n) = \sum_{j=1}^{n} W(X_j)$.

a.) Find the variance of $W(X_j)$. For $\epsilon = 0.01$, evaluate a bound on the probability of the typical set $\mathcal{A}_\epsilon^{(n)}$ using Chebyshev's inequality.

b.) Let N_a be the number of a's in the string $X^n = (X_1, \ldots, X_n)$. The random variable (rv) N_a is the sum of n iid rv's. Show what these rv's are.

c.) Express the rv $W(X^n)$ as a function of the rv N_a. Note how this depends on n.

d.) Express the typical set in terms of bounds on N_a. Use Chebyshev's inequality to derive bounds on the probability of the typical set, using properties of N_a instead of $W(X_j)$. **Hint:** You may write $\mathcal{A}_\epsilon^{(n)} = \{x^n : \alpha < N_a < \beta\}$ and calculate α and β.

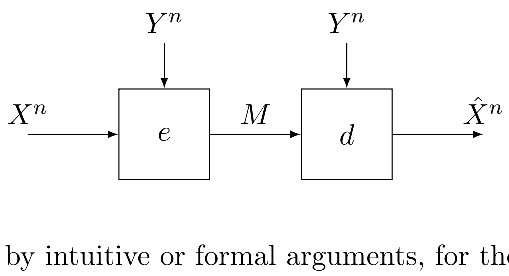
e.) Find $\Pr(N_a = i)$ for $i = 0, 1, 2$. Find the probability of each individual string x^n for those values of i. Find the particular string x^n that has maximum probability over all sample values of X^n. What are the next most probable n-strings. Give a brief discussion of why the most probable n-strings are not regarded as typical strings.

Exercise 5.4 Source coding with side information [EE5139]

Consider a memoryless source $(\boldsymbol{X}, \boldsymbol{Y})$ that produces in each iteration two random variables, X_i and Y_i, where X_i is private information and Y_i is public information. The pairs (X_i, Y_i) follow a joint distribution P_{XY} and are i.i.d.. We are looking for a fixed-length block code that compresses the private information $X^n = (X_1, X_2, \ldots, X_n)$ using the public information $Y^n = (Y_1, Y_2, \ldots, Y_n)$ such that the code can be decoded asymptotically error-free with help of the public information.
An $(n, 2^L)$-code for such a source is given by an encoder, $e : (X^n, Y^n) \to M$, and decoder, $d : (M, Y^n) \to \hat{X}^n$, as illustrated in the figure below. The codeword $M \in \{0, 1\}^L$ is a binary string of length L. We define $R^*(\boldsymbol{X}|\boldsymbol{Y})$ as the infimum over all rates R such that there exists a sequence of $(n, 2^{nR})$-codes satisfying

$$\lim_{n\to\infty} \Pr\left[X^n \neq \hat{X}^n\right] = 0, \quad \text{where} \quad \hat{X}^n = d_n(e_n(X^n, Y^n), Y^n) \tag{2}$$

is a function of both X^n and Y^n. We want to establish that $R^*(\boldsymbol{X}|\boldsymbol{Y}) = H(X|Y)$.



a.) Determine $R^*(\boldsymbol{X}|\boldsymbol{Y})$, by intuitive or formal arguments, for the simple cases where

  i.) X and Y are independent,

  ii.) $X = Y$,

b.) By explicitly constructing a code for the source (X, Y) using codes for the sources Y and X (with side information Y), show that $R^*(\boldsymbol{X}, \boldsymbol{Y}) \leq R^*(\boldsymbol{X}|\boldsymbol{Y}) + R^*(\boldsymbol{Y})$.

c.) Show that the converse, $R^*(\boldsymbol{X}|\boldsymbol{Y}) \geq H(X|Y)$ using Fano's inequality. **Hint:** You will also need the following sequence of inequalities, which needs to be verified.

$$\begin{aligned} H(X^n|\hat{X}^n) &\geq H(X^n|Y^nM) && (3)\\ &= H(X^nM|Y^n) - H(M|Y^n) && (4)\\ &\geq H(X^nM|Y^n) - L && (5)\\ &\geq H(X^n|Y^n) - L\,. && (6) \end{aligned}$$

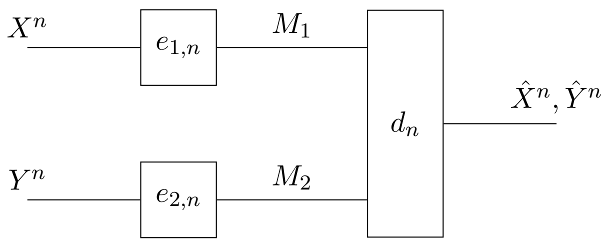
d.) Give a formal proof or a sketch of a proof that $R^*(\boldsymbol{X}|\boldsymbol{Y}) \leq H(X|Y)$. **Hint:** Consider the typical set

$$\mathcal{A}_\epsilon^{(n)}(\boldsymbol{X}|\boldsymbol{Y}) := \left\{ (x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n : \left| \frac{1}{n} \log \frac{1}{P_{X^n|Y^n}(x^n|y^n)} - H(X|Y) \right| \leq \epsilon \right\}. \quad (7)$$

Exercise 5.5 Achievability for the Slepian–Wolf coding problem [EE6139]

We return to the Exercise 4.4 from the last homework. Let X and Y be a pair of jointly distributed random variables. (X is distributed on finite set $\mathcal{X}$, and Y is distributed on finite set $\mathcal{Y}$.) An $(n, 2^{nL_1}, 2^{nL_2})$-separately-encoded-jointly-decoded source code consists of a pair of encoders e_1, e_2, and a decoder d, where

- $e_1 : \mathcal{X}^n \to \{0,1\}^{nL_1}$,
- $e_2 : \mathcal{Y}^n \to \{0,1\}^{nL_2}$, and
- $d : \{0,1\}^{nL_1} \times \{0,1\}^{nL_2} \to \mathcal{X}^n \times \mathcal{Y}^n$.



The rate pair (R_1, R_2) is said to be achievable for DMS (X, Y) if there exists a sequence of $(n, 2^{nL_1}, 2^{nL_2})$-codes with encoders $e_{1,n}$, $e_{2,n}$ and decoder d_n such that

$$\lim_{n\to\infty} P\{(\hat{X}^n, \hat{Y}^n) \neq (X^n, Y^n)\} = 0$$

where

$$(\hat{X}^n, \hat{Y}^n) = d_n(M_1, M_2),\ M_1 = e_{1,n}(X^n),\ \text{and } M_2 = e_{2,n}(Y^n)$$

are the reconstructed source and codewords respectively.
This time, we are interested in the achievability of the problem.

a.) **An alternative for typical sequences** Given $n \in \mathbb{N}$ and $\epsilon \in (0,1)$, we define the set of Y-sequences as

$$\mathcal{T}_\epsilon^{(n)}(Y) \triangleq \left\{ \boldsymbol{y} \in \mathcal{Y}^n : \left| \frac{\sum_{i=1}^n \delta_{y,y_i}}{n} - p_Y(y) \right| < \left\lceil \sqrt{\frac{|\mathcal{Y}|}{\epsilon}} \right\rceil \sqrt{\frac{p_Y(y)(1-p_Y(y))}{n}} \quad \forall y \in \mathcal{Y} \right\}.$$

Show that

i.) $P\left[Y^n \in \mathcal{T}_\epsilon^{(n)}(Y)\right] \geq 1 - \epsilon$

ii.) There exists some $A > 0$ independent from n and ϵ such that

$$2^{-nH(Y)-A\sqrt{n/\epsilon}} < p_{Y^n}(\boldsymbol{y}) < 2^{-nH(Y)+A\sqrt{n/\epsilon}}$$

for all $\boldsymbol{y} \in \mathcal{T}_\epsilon^{(n)}(Y)$.

iii.) $\lim_{n\to\infty} \frac{1}{n} \log_2 \left|\mathcal{T}_\epsilon^{(n)}(Y)\right| = H(Y)$.

b.) **Position-based coding** Given positive integer n and $\epsilon > 0$, let $M_X \triangleq \left\lfloor 2^{n(H(Y|X)+\epsilon)} \right\rfloor$, and let M be another positive integer. Let $\{\boldsymbol{X}_{i,j}\}_{i,j}$ be a set of i.i.d. random variables on $\mathcal{X}^n$, where $i \in \{1, \ldots, M_X\}$, $j \in \{1, \ldots, M\}$, and

$$p_{\boldsymbol{X}_{i,j}}(\boldsymbol{x}) = \prod_{k=1}^{n} p_X(x_k)$$

for each (i, j).

i.) Suppose $I(X, Y) > \frac{1}{2}\epsilon$, and let $M = \left\lfloor 2^{n(I(X,Y)-\frac{1}{2}\epsilon)} \right\rfloor$. Prove that, for n large enough,

$$P\left[X^n \neq \boldsymbol{X}_{i,j}\ \forall (i, j)\right] < 2\epsilon.$$

ii.) Let A and B be a pair of random variable denoting the "smallest" indices a, b such that that $X^n = \boldsymbol{X}_{a,b}$. Namely,

$$p_{A,B|X^n,\{\boldsymbol{X}_{i,j}\}}(a, b|\boldsymbol{x}, \{\boldsymbol{x}_{i,j}\}) = \begin{cases} & \boldsymbol{x} = \boldsymbol{x}_{a,b} \\ 1 \quad \text{if} & \boldsymbol{x} \neq \boldsymbol{x}_{i,j} \quad \forall i < a \\ & \boldsymbol{x} \neq \boldsymbol{x}_{a,j} \quad \forall j < b \\ 0 \quad \text{otherwise} & \end{cases}.$$

We take the convention that $(A, B) = (\infty, \infty)$ if $X^n \neq \boldsymbol{X}_{i,j}$ for all i, j. Prove that

$$P\left[A < \infty,\, B < \infty,\, p_{Y^n|X^n}(Y^n|\boldsymbol{X}_{A,j}) \geq p_{Y^n|X^n}(Y^n|X^n)\ \exists j \neq B\right] < \epsilon$$

for n large enough.

c.) Based on the arguments in a.) and b.), show that, for any δ_1, $\delta_2 > 0$, the following rates are achievable

$$R_1 = H(X|Y) + \delta_1, \tag{8}$$
$$R_2 = H(Y) + \delta_2. \tag{9}$$

d.) Show that any (R_1, R_2) satisfying the following inequalities are achievable

$$R_1 > H(X|Y), \tag{10}$$
$$R_2 > H(Y|X), \tag{11}$$
$$R_1 + R_2 > H(X, Y). \tag{12}$$