EE2340/EE5847: Information Sciences/Information Theory **2020**

Homework 3: 3rd Feb 2020

Instructor: Shashank Vatedka

Instructions: *You are encouraged to discuss and collaborate with your classmates. However, you* must explicitly *mention at the top of your submission who you collaborated with. Copying is NOT permitted, and solutions must be written independently and in your own words.*

Homeworks must be submitted on Google classroom. Please scan a copy of your handwritten assignment and upload as pdf with filename <your ID>_HW<homework no>.pdf. Example: EEB19BTECH00000_HW1.pdf.

For programming questions, submit as separate files. Please use the naming convention <your ID>_HW<homework no>_problem<problem no>.. Example: EEB19BTECH00000_HW1_problem1.c*

Exercise 3.1. Compute the KL divergence $D(p\|q)$ for the following pairs of distributions:

1. $p = \mathcal{N}(\mu_1, \sigma^2)$ and $q = \mathcal{N}(\mu_2, \sigma^2)$
2. $p =$Poisson(λ_1) and $q =$Poisson(λ_2)
3. $p =$Binomial(p) and $q =$Binomial(q)
4. $p =$Exponential(α) and $q =$Exponential(β)

Exercise 3.2. Give two separate examples of triplets of probability mass functions p, q, r on $\{0, 1\}$ that satisfy

1. $D(p\|q) + D(q\|r) < D(p\|r)$
2. $D(p\|q) + D(q\|r) > D(p\|r)$

This confirms that the KL divergence between distributions does not satisfy the triangle inequality, unlike the Euclidean distance.

Exercise 3.3. Prove the following version of the second law of thermodynamics: Suppose that a system evolves according to a discrete-time (first order) Markov chain. Then, conditioned on the initial state of the chain, the entropy of the system cannot decrease with time.

Exercise 3.4. Prove that for any joint distribution p_{XY} on $\mathcal{X} \times \mathcal{Y}$,

$$I(X;Y) = \min_{q_X, q_Y} D(p_{XY} \| q_X q_Y)$$

where q_X, q_Y are minimized over distributions on $\mathcal{X}$ and $\mathcal{Y}$ respectively.

Exercise 3.5 (Nonuniform distributions can achieve capacity)**.** Find the capacity, and the capacity-achieving input distribution for the following asymmetric channels:

1. Erasure channel: $\mathcal{X} = \{0, 1\}$, $\mathcal{Y} = \{0, 1, e\}$ and

$$p_{Y|X}(y|x) = \begin{cases} 0.7 & \text{if } (x, y) = (0, 0) \\ 0.3 & \text{if } (x, y) = (0, e) \\ 0.9 & \text{if } (x, y) = (1, 1) \\ 0.1 & \text{if } (x, y) = (1, e) \end{cases}$$

2. Binary-input binary-output channel: $\mathcal{X} = \{0, 1\}$, $\mathcal{Y} = \{0, 1\}$ and

$$p_{Y|X}(y|x) = \begin{cases} 0.7 & \text{if } (x, y) = (0, 0) \\ 0.3 & \text{if } (x, y) = (0, 1) \\ 0.9 & \text{if } (x, y) = (1, 1) \\ 0.1 & \text{if } (x, y) = (1, 0) \end{cases}$$

Exercise 3.6 (Capacity-achieving input distribution need not be unique)**.** Consider the following channel with $\mathcal{X} = \{0, 1, 2\}$ and $\mathcal{Y} = \{0, 1, 2\}$:

$$p_{Y|X}(y|x) = \begin{cases} 0.5 & \text{if } (x, y) = (0, 0) \text{ or } (0, 1) \text{ or } (1, 0) \text{ or } (1, 1) \\ 1 & \text{if } (x, y) = (2, 2) \\ 0 & \text{otherwise} \end{cases}$$

Give two input distributions p_X that maximize $I(X; Y)$.

Exercise 3.7. Let $\mathcal{P}$ be the set of all probability mass functions on $\{0, 1, 2\}$ that satisfy $p(0) + p(1) \leqslant 0.3$. Compute

$$\max_{p \in \mathcal{P}} H(X)$$

Exercise 3.8. For $\alpha \geqslant 0$ and $\alpha \neq 1$, the Rényi entropy of order α for a pmf p_X is defined as

$$R_\alpha(X) = \frac{1}{1 - \alpha} \log_2 \left(\sum_{x \in \mathcal{X}} (p_X(x))^\alpha \right).$$

Prove the following

1. $$\lim_{\alpha \to 1} R_\alpha(X) = H(X).$$

2. $$R_2(X) = -\log_2 \Pr[X = Y]$$

   where X, Y are independent and identically distributed random variables having pmf p_X.

3. Assuming that $\arg\max_x p_X(x)$ is unique,

   $$\lim_{\alpha \to \infty} R_\alpha(X) = -\log_2 \left(\max_x p_X(x) \right)$$