NUS National University of Singapore

EE5139/EE6139
Homework 6

Semester I, AY2021/2022
Marco Tomamichel

Exercise 6.1 Data processing inequality for the total variation distance (EE5139)

Let P_X, Q_X be two distributions over some finite set $\mathcal{X}$. Let $\{T(y|x)\}_{y\in\mathcal{Y},x\in\mathcal{X}}$ be some conditional pmf, where $\mathcal{Y}$ is some finite set. Suppose $P_Y(y) = \sum_{x\in\mathcal{X}} T(y|x) \cdot P_X(x)$ and $Q_Y(y) = \sum_{x\in\mathcal{X}} T(y|x) \cdot Q_X(x)$ for each $y \in \mathcal{Y}$. Prove that

$$\delta_{\text{tvd}}(P_Y, Q_Y) \le \delta_{\text{tvd}}(P_X, Q_X).$$

Exercise 6.2 Empirical typical set (all)

Let P_X be a probability distribution over some finite set $\mathcal{X}$. Recall that the empirical typical set (of length n and tolerance ϵ) w.r.t. P_X is defined as

$$\mathcal{A}^{(n)}_{\text{emp},\epsilon}(P_X) := \{\boldsymbol{x}^n \in \mathcal{X}^n : \delta_{\text{tvd}}(f_{\boldsymbol{x}^n}, P_X) \le \epsilon\}$$

where $f_{\boldsymbol{x}^n}$ is the empirical distribution on $\mathcal{X}$ induced by the sequence $\boldsymbol{x}^n$, *i.e.*,

$$f_{\boldsymbol{x}^n}(x) := \frac{1}{n} \left|\{i \in \{1, \ldots, n\} : x_i = x\}\right|.$$

Prove that, for any $\epsilon \in (0, 1]$,

$$\lim_{n\to\infty} \mathrm{P}\left[X^n \in \mathcal{A}^{(n)}_{\text{emp},\epsilon}(P_X)\right] = 1.$$

Exercise 6.3 Hypothesis testing (all)

Consider the three pmfs P and Q, and S given by the probability vectors $p = (\frac{1}{2}, \frac{1}{4}, \frac{1}{4})$, $q = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ and $s = (0, \frac{1}{2}, \frac{1}{2})$, respectively.

a.) Compute the symmetric error probabilities $\epsilon^*_{\text{sym},1}$ for hypothesis tests for all three pairs.

b.) Consider the problem of ternary hypothesis testing between the three distributions, when all three have equal priors. Can you find the minimal error probability and optimal test?

c.) Compute $D(P\|Q)$ and $D(P\|S)$.

d.) Consider asymmetric hypothesis testing for P and S. For each $\epsilon > 0$, find $N_0(\epsilon) \in \mathbb{N}$ such that $\beta_n(\epsilon) = 0$ for all $n \ge N_0(\epsilon)$. What does that say about the limit $\lim_{n\to\infty} -\frac{1}{n}\log\beta_n(\epsilon)$? Interpret your result in item c) in this light.

Exercise 6.4 A simple parity check code (EE5139)

In the first lecture we encountered a code that stores $k = 4$ bits in $n = 8$ bits by computing the parities $x_1 \oplus x_2$, $x_3 \oplus x_4$, $x_1 \oplus x_3$ and $x_2 \oplus x_4$.

a.) Give the codewords for this code and compute the minimal distance. How many errors can it detect and correct?

b.) Is it a linear code? If so, compute matrices G and H.

c.) Use the Hamming bound to determine if this code is perfect or not.

d.) Construct the dual code for this code.

Exercise 6.5 Asymptotic property of the smooth min-entropy (EE6139)

Given a random variable X on $\mathcal{X}$ distributed according to P_X, we recall that the min-entropy of X is defined as

$$H_{\min}(X) := -\log \max_{x\in\mathcal{X}} P_X(x) = \min_{x\in\mathcal{X}} -\log P_X(x).$$

For each $\epsilon \in [0,1)$, the ϵ-smooth min-entropy of X is defined as

$$H^{\epsilon}_{\min}(X) := \max_{\delta_{\text{tvd}}(P_X,\tilde{P}_X)\leq\epsilon} H_{\min}(X)_{\tilde{P}_X} = \max_{\delta_{\text{tvd}}(P_X,\tilde{P}_X)\leq\epsilon} \left\{-\log \max_{x\in\mathcal{X}} \tilde{P}_X(x)\right\}.$$

We are interested in showing $\frac{1}{n}H^{\epsilon}_{\min}(X^n) \to H(X)$ as $n\to\infty$, $\epsilon\to 0$ (in that order), where X^n consists of n i.i.d. copies of X, namely $P_{X^n} = \prod_{i=1}^n P_{X_i}$.

a.) For each $\epsilon > 0$, let $\mathcal{A}^{(n)}_{\epsilon}(X)$ denote the ϵ-typical set of X (see eq. (2.52) from the lecture notes). For each positive integer n, define a distribution on $\mathcal{X}^n$ as

$$\hat{P}_{X^n}(\boldsymbol{x}^n) := \begin{cases} P_{X^n}(\boldsymbol{x}^n) & \boldsymbol{x}^n \in \mathcal{A}^{(n)}_{\epsilon}(X) \\ \frac{1-P_{X^n}(\mathcal{A}^{(n)}_{\epsilon}(X))}{|\mathcal{X}|^n - \left|\mathcal{A}^{(n)}_{\epsilon}(X)\right|} & \boldsymbol{x}^n \notin \mathcal{A}^{(n)}_{\epsilon}(X) \end{cases}$$

Prove that for n large enough, $\delta_{\text{tvd}}(P_{X^n}, \hat{P}_{X^n}) \leq \epsilon$.

b.) Using $\hat{P}_{X^n}$ from the previous step, show that

$$\frac{1}{n}H^{\epsilon}_{\min}(X^n) \geq H(X) - \epsilon$$

for n large enough.

c.) **Continuity of entropy.** For any two distributions P_X, $\tilde{P}_X$ (on the same set) such that $\delta_{\text{tvd}}(P_X, \tilde{P}_X) \leq \epsilon$, prove that for all $\epsilon < 1/2$

$$\left|H(X)_{P_X} - H(X)_{\tilde{P}_X}\right| \leq h(\epsilon) + \epsilon \cdot \log|\mathcal{X}|.$$

where $h : t \mapsto -t\log t - (1-t)\log(1-t)$. **Hint:** Construct a pair of joint random variables (X_1, X_2) with marginals $P_{X_1} = P_X$, $P_{X_2} = \tilde{P}_X$, and $\mathrm{P}[X_1 = X_2]$ large. Use Fano's inequality.

d.) Using the previous step, show that

$$\frac{1}{n}H^{\epsilon}_{\min}(X^n) \leq H(X) + \frac{1}{n}\cdot h(\epsilon) + \epsilon\cdot\log|\mathcal{X}|$$

for $0 < \epsilon < 1/2$.