CS3236: Homework 11

Pierre Senellart ⟨pierre.senellart@nus.edu.sg⟩

Due Monday, November 3, 2pm

Submission: Submission can be made in person before the beginning of the lecture, or by upload to the IVLE Workbin. You may submit handwritten answers (possibly scanned), but they must be **clearly readable**. Answers that are not readable may receive only partial points. Late answers are still accepted on Tuesday by 2pm (on the IVLE Workbin only), but receive only 50% of the possible points. Any later submissions receive no points. Assignments are marked out of 20 points.

1. **Properties of the min-entropy.** (10 points) During the lecture we talked about the min-entropy as a means to quantify information. In this homework, you will prove some useful properties of the min-entropy. Show that for all distributions $\Pr_{XYE}$ and $\Pr_{XE}$ the following hold. Remember, that you may use both the standard definition of the min-entropy as well as its more operational interpretation as the guessing entropy.
   a) (1 points) $H_{\min}(X \mid E) \geq 0$.
   b) (2 points) $H_{\min}(X \mid E) \leq \log|X|$, where $|X|$ is the support of the distribution $\Pr_X$. (The support of a random variable X over $\mathcal{X}$ is given by $|X| = |\{x \in \mathcal{X} \mid \Pr_X(x) > 0\}|$.)
   c) (1 point) $H_{\min}(XY \mid E) \geq H_{\min}(X \mid E)$.
   d) (4 points) $H_{\min}(X \mid f(E)) \geq H_{\min}(X \mid E)$ for any function f applied to $E = e$.
   e) (2 points) $H_{\min}(X \mid E) \geq H_{\min}(X) - \log|E|$.
2. **Properties of the statistical distance.** (10 points)
   a) Remember our die guessing game. Let's investigate in more detail just how well we can distinguish two different dice. Suppose that I have two dice in my possession with possible symbols $\mathcal{R} = \{1, 2, 3, 4, 5, 6\}$. Die 0 is a fair die, i.e., the probability distribution $\Pr_X$ over the possible outcomes is uniform. Die 1 however is not quite fair and favors the outcome 6 with probability distribution $\Pr_Y$ given by $\rho_Y = (\frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{3}{8})$. With probability $\frac{1}{2}$ I will roll die 0 and with probability $\frac{1}{2}$ I will roll die 1. You cannot see which die I used, but I tell you the outcome of my die roll. Can you guess whether I used die 0 or die 1?
      i. (1 point) Describe a strategy to determine which die I used. What is the probability that your strategy is correct (averaged over the choice of die)?

ii. (1 point) Compute $\Delta(\Pr_X, \Pr_Y)$.

iii. (1 point) Using the total variation distance, show that any strategy has success probability at most $\frac{29}{48}$.

iv. (1 point) How does this compare to your strategy? Can you think of a strategy that attains this bound?

b) (3 points) Let X and Y be random variables distributed over a set R. Show that we can also write the statistical distance as

$$\Delta(\Pr_X, \Pr_Y) = \max_{S \subseteq R} (\Pr_X(S) - \Pr_Y(S)), \tag{1}$$

where the maximization is taken over subsets S of R.

c) (3 points) Consider a function $f : R \to [0, 1]$, and suppose that we apply f to X. Let $\mathbb{E}(f(X))$ denote the expected value of $f(X)$. Show that we can also write the statistical distance as

$$\Delta(\Pr_X, \Pr_Y) = \max_{f} |\mathbb{E}(f(X)) - \mathbb{E}(f(Y))|, \tag{2}$$

where the maximization is taken over functions $f : R \to [0, 1]$. *(Hint: Think about what functions might maximize* $\mathbb{E}(f(X)) - \mathbb{E}(f(Y))$*.)*

d) (Bonus) Show that for any randomized function F

$$\Delta(\Pr_{F(X)}, \Pr_{F(Y)}) \leq \Delta(\Pr_X, \Pr_Y). \tag{3}$$

(Hint: show your result first for a deterministic function. Then use the fact that a randomized function can be seen as choosing a deterministic function f *with probability* $\Pr_F(f)$*.)*

e) (Bonus) Show that equality holds if the (deterministic) function F is one-to-one.