EE5139/EE6139
Homework 2

Semester I, AY2021/2022
Marco Tomamichel

### Exercise 2.1 Upper bound on entropy [EE5139]

In the lecture notes we show that $H(X) \leq \log|\mathcal{X}|$ for binary random variables. Show this statement for general discrete random variables on any (finite) alphabet $\mathcal{X}$.

### Exercise 2.2 Relative entropy as a parent quantity [all]

Let $X$ and $Y$ be random variables on alphabets $\mathcal{X}$ and $\mathcal{Y}$ with joint pmf $P_{XY}$. Moreover, let $U$ be a uniform random variable on $\mathcal{X}$. Show the following relations:

a.) $H(X) = \log|\mathcal{X}| - D(P_X\|U_X)$.

b.) $H(X|Y) = \log|\mathcal{X}| - D(P_{XY}\|U_X \times P_Y)$.

c.) $I(X:Y) = D(P_{XY}\|P_X \times P_Y)$.

### Exercise 2.3 Example correlations [EE5139]

For each item, find an example of random variables $X$, $Y$ and $Z$ (you can restrict the alphabet size to at most 2 bits) such that the desired relations holds:

a.) $H(X|YZ) = 0$ but $H(X|Y) = H(X|Z) = 1$.

b.) $I(X:Y|Z) = 1$ but $I(X:Y) = 0$.

c.) $I(X:Y) = 1$ but $I(X:Y|Z) = 0$.

d.) $I(X:Y) = I(X:Z) = 1$ but $I(Y:Z) = 0$.

### Exercise 2.4 Information spectrum [EE6139]

Given a random variable $X$ governed by the pmf $P$ or an alternative pmf $Q$, the log-likelihood ratio is defined as the random variable $Z(X) = \log\frac{P(X)}{Q(X)}$.

a.) We have seen that the expectation value of $Z$ (under $P$) is the relative entropy

$$\mathbb{E}[Z] = \sum_{x\in\mathcal{X}} P(x)\log\frac{P(x)}{Q(x)} = D(P\|Q)\,. \tag{1}$$

Give an expression for $\mathrm{Var}[Z]$ (under $P$). This quantity is called the relative entropy variance and denoted by $V(P\|Q)$.

Consider now a sequence of i.i.d. random variables $X^n = (X_1, X_2, \ldots, X_n)$ on $\mathcal{X}^n$ where each $X_i$ is governed by the pmf $P$ or an alternative pmf $Q$. We are interested in pmf of the log-likelihood ratio $Z(X^n)$.

b.) Show that $Z(X^n) = \sum_{i=1}^n Z(X_i)$. What is $\mathbb{E}[Z^n]$ and $\mathrm{Var}[Z^n]$?

c.) Let us now consider the quantity $\Pr[Z(X^n) \leq nR]$ in the limit of large $n$ for different values of $R$. Show that

$$\lim_{n\to\infty} \Pr[Z(X^n) \leq nR] = \begin{cases} 0 & \text{if} \quad R < D(P\|Q) \\ 1 & \text{if} \quad R > D(P\|Q) \end{cases}. \tag{2}$$

**Hint:** Argue using the weak law of large numbers.

d.) Later on in the lecture we will encounter the quantity

$$D_s^\epsilon(P^n\|Q^n) := \sup\{k \in \mathbb{R} : \Pr[Z(X^n) \le k] \le \epsilon\}, \tag{3}$$

which, in words, is asking the largest $k$ such that the tail of the distribution of $Z$ that lies below $k$ has cumulative probability at most $\epsilon$. Show that $D_s^\epsilon(P^n\|Q^n) = nD(P\|Q) + o(n)$, or equivalently,

$$\lim_{n\to\infty} \frac{1}{n} D_s^\epsilon(P^n\|Q^n) = D(P\|Q)\,. \tag{4}$$

**Hint:** Verify that $\frac{1}{n}D_s^\epsilon(P^n\|Q^n) = \sup\{k \in \mathbb{R} : \Pr[\frac{1}{n}Z(X^n) \le k] \le \epsilon\}$.

e.) Optional: Show that in the next order in $n$, we have

$$D_s^\epsilon(P^n\|Q^n) = nD(P\|Q) + \sqrt{nV(P\|Q)}\,\Phi^{-1}(\epsilon) + o\left(\sqrt{n}\right) \tag{5}$$

Can we even say something more about the $o\left(\sqrt{n}\right)$ term?

**Hint:** The statement can be shown using the central limit theorem. A quantitative version of the central limit theorem is the Berry-Esseen theorem. Look it up to make even stronger statements about the remainder term.

### Exercise 2.5 Independence and mutual information [all]

Consider two sequences of random variables $X_1, \dots, X_n$ and $Y_1, \dots, Y_n$. Show that if $X_1, \dots, X_n$ are mutually independent, then

$$I(X_1, \dots, X_n : Y_1, \dots, Y_n) \ge \sum_{i=1}^n I(X_i : Y_i)$$

while if given $Y_i$ the random variable $X_i$ is conditionally independent of all the remaining random variables for all $i = 1, \dots, n$, then

$$I(X_1, \dots, X_n : Y_1, \dots, Y_n) \le \sum_{i=1}^n I(X_i : Y_i)$$