ECE 563: Information Theory (Fall 2017)

Maxim Raginsky

Homework 1

Assigned August 31, 2017; due September 7, 2017

Reading:

- Polyanskiy and Wu, Sections 1.1–1.3, 1.6, 1.7, 2.1–2.5
- Cover and Thomas, Sections 2.1–2.8

---

1. Let X and Y be two jointly distributed random variables, where X takes values in $\{0,1\}$ and Y takes values in $\mathbb{Z}_+ := \{0,1,2,\ldots\}$ according to the joint distribution

$$P_{XY}(x,y) = \alpha 2^{-x-2y}, \qquad x \in \{0,1\}, y \in \mathbb{Z}_+$$

   where α is a normalization constant. Compute $H(X)$, $H(Y)$, $H(X,Y)$, $H(Y|X)$, and $H(X|Y)$.

2. Let X be an exponential random variable with parameter λ, i.e., with probability density $p_X(x) = \lambda e^{-\lambda x}\mathbf{1}_{\{x \ge 0\}}$. Compute the divergence $D(P_{X+a} \| P_X)$ for an arbitrary $a \in \mathbb{R}$.

3. Recall the definition $d(p\|q) := D(\text{Bern}(p)\|\text{Bern}(q))$ of the binary divergence function:

$$d(p\|q) = p \log \frac{p}{q} + (1-p) \log \frac{1-p}{1-q}.$$

   (a) Use calculus to prove that, for all $p, q \in [0,1]$,

$$d(p\|q) \ge 2(p-q)^2 \log e.$$

   (b) Use the result from part (a) and the data processing inequality to prove the Pinsker–Csiszár inequality

$$\text{TV}(P,Q) \le \sqrt{\frac{1}{2\log e} D(P\|Q)},$$

   where

$$\text{TV}(P,Q) = \sup_E |P(E) - Q(E)|$$

   is the total variation distance between P and Q.

4. One of the most useful inequalities in information theory is

$$\ln x \leq x - 1, \qquad x > 0$$

(it is a consequence of the concavity of the logarithm and the fact that the line $x \mapsto x - 1$ is tangent to the graph of the log function at $x = 1$). In the future, we will refer to it as the *key inequality*.

   (a) Use the key inequality to prove the *log-sum inequality*: for any choice of nonnegative reals $a_1, \ldots, a_n, b_1, \ldots, b_n$,

   $$\sum_{i=1}^{n} a_i \log \frac{a_i}{b_i} \geq a \log \frac{a}{b},$$

   where $a := \sum_{i=1}^{n} a_i$ and $b := \sum_{i=1}^{n} b_i$.

   (b) Use the key inequality to prove that the divergence is nonnegative, i.e., $D(P \| Q) \geq 0$ for any two probability distributions on a common probability space.

5. Let X, Y, Z be three jointly distributed random variables such that $X \to Y \to Z$ is a Markov chain. Assume that Y takes values in a finite set $\mathcal{Y}$. Prove that

$$I(X; Z) \leq \log |\mathcal{Y}|.$$