Second homework set, Due March 27

1. (1p.) Determine $D(P\|Q)$

   (a) when P and Q are binomial distributions

$$P(k) = \binom{n}{k} p^k (1-p)^{n-k},\ Q(k) = \binom{n}{k} q^k (1-q)^{n-k},\ k \in \{0, \dots, n\}. \tag{1}$$

   (b) when P and Q are Poisson distributions

$$P(k) = \frac{\lambda^k}{k!} e^{-\lambda},\ Q(k) = \frac{\mu^k}{k!} e^{-\mu},\ k \in \{0, 1, \dots\}. \tag{2}$$

   Hint: You may use that the expected value $\sum P(k)k$ of the binomial or Poisson distribution is known from probability theory.

In the next three exercises, all alphabets are finite.

2. (3p.) (Multi-information)

   A natural candidate for the mutual information of RVs X, Y and Z taking values in $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$, respectively, is the divergence $\mathrm{D}(\mathbb{P}_{XYZ}||\mathbb{P}_X \times \mathbb{P}_Y \times \mathbb{P}_Z)$, where $\mathbb{P}_{XYZ}$ denotes their joint distribution and $\mathbb{P}_X \times \mathbb{P}_Y \times \mathbb{P}_Z$ is the product distribution on $\mathcal{A} \times \mathcal{B} \times \mathcal{C}$ with first, second and third one dimensional marginal distributions being equal to the marginal distributions of X, Y and Z respectively! Show that it equals $H(X) + H(Y) + H(Z) - H(X, Y, Z) = I(X, Y \wedge Z) + I(X \wedge Y)$. More generally, show that

$$\left(\sum_{i=1}^{n} H(X_i)\right) - H(X_1, X_2, \dots, X_n) = \mathrm{D}(\mathbb{P}_{X_1 X_2 \dots X_k}||\mathbb{P}_{X_1} \times \mathbb{P}_{X_2} \times \cdots \times \mathbb{P}_{X_n}), \tag{3}$$

   and give decompositions of the latter into sum of mutual informations!

3. (3p.) (Optimality of the universal construction)

   Read through the second exercise of the first homework set! Show that the error exponent $e_Q(R)$ is best possible for every Q, i.e., for arbitrary $\tilde{A}_n$ satisfying

$$\frac{1}{n} \log |A_n| \to R \tag{4}$$

   always

$$\liminf_{n\to\infty} \frac{1}{n} \log Q^n(\tilde{A}_n^c) \geq -e_Q(R). \tag{5}$$

   Hint: Check that for any fixed $\varepsilon > 0$, there is n_o such that $\frac{|T_P^n \cap \tilde{A}_n^c|}{|T_P^n|} \geq \frac{1}{2}$ for each n-type P with $H(P) > R + \varepsilon$ if $n \geq n_o$.

4. (2p+1p.)

   (a) Prove the Pinsker inequality (also called Csiszár-Kemperman-Kullback-Pinsker inequality)

$$\mathrm{D}(\mathbb{P}||\mathbb{Q}) \geq \frac{1}{2 \ln 2} \mathrm{d}^2(\mathbb{P}, \mathbb{Q}), \tag{6}$$

   where $d(\mathbb{P}, \mathbb{Q})$ is the variational distance of distributions $\mathbb{P}$ and $\mathbb{Q}$, i.e.,

$$\mathrm{d}(\mathbb{P}, \mathbb{Q}) = \sum_{a \in \mathcal{A}} |\mathbb{P}(a) - \mathbb{Q}(a)|. \tag{7}$$

(b) Show that this bound is tight in the sense that the ratio of $\mathrm{D}(\mathbb{P}||\mathbb{Q})$ and $d^2(\mathbb{P},\mathbb{Q})$ can be arbitrarily close to $\frac{1}{2\ln 2}$.

Hint: With $B \triangleq \{a : \mathbb{P}(a) \geq \mathbb{Q}(a)\}$, $\widetilde{\mathbb{P}} \triangleq (\mathbb{P}(B), \mathbb{P}(A-B))$, $\widetilde{\mathbb{Q}} \triangleq (\mathbb{Q}(B), \mathbb{Q}(A-B))$, we have $\mathrm{D}(\mathbb{P}||\mathbb{Q}) \geq \mathrm{D}(\widetilde{\mathbb{P}}||\widetilde{\mathbb{Q}})$, $\mathrm{d}(\mathbb{P},\mathbb{Q}) = d(\widetilde{\mathbb{P}},\widetilde{\mathbb{Q}})$. Hence it suffices to consider the case $A = \{0,1\}$, i.e., to determine the largest c such that

$$p\log\left(\frac{p}{q}\right) + (1-p)\log\left(\frac{1-p}{1-q}\right) - 4c(p-q)^2 \geq 0, \text{ for every } 0 \leq q \leq p \leq 1.$$

For $q = p$ the equality holds; further, the derivative of the left-hand side with respect to q is negative for $q < p$ if $c \leq \frac{1}{2\ln 2}$ while for $c > \frac{1}{2\ln 2}$ and $p = \frac{1}{2}$ it is positive in the neighborhood of p.

Remark: When checking the statements of the hint above, pay attention to the fact that the base of the logarithm is 2, hence, $(\log x)' = \frac{1}{x\ln 2}$.