First homework set, Due March 13, 16:00

1. (4p.) Determine the I-divergences $\mathrm{D}(\mathbb{P}||\mathbb{Q})$ and $\mathrm{D}(\mathbb{Q}||\mathbb{P})$ if

   (a) $\mathbb{Q}$ is an arbitrary distribution over a finite set A and $\mathbb{P}(a) = \frac{\mathbb{Q}(a)}{\mathbb{Q}(B)}$ if $a \in B$ and 0 otherwise, where $B \subset A$ and $\mathbb{Q}(B) = \sum_{a \in B} Q(a) > 0$.

   (b) $\mathbb{P}$ and $\mathbb{Q}$ are defined as follows. A hat contains k_1 slips of paper marked with 1, k_2 slips of paper marked with 2 ... and k_r slips of paper marked with r. Let n be equal to $k_1 + \cdots + k_r$. We draw n times from the hat (i) without replacement (ii) with replacement. For a given sequence $\mathbf{x} = x_1 \ldots x_n \in \{1, \ldots, r\}^n$ let $\mathbb{P}(\mathbf{x})$ and $\mathbb{Q}(\mathbf{x})$ denote the probability of drawing the given sequence in cases (i) and (ii), respectively. Is there any connection with part (a)?

   (c) Calculate also the entropies of $\mathbb{P}$ and $\mathbb{Q}$ defined in part (b)! Which entropy is the larger?

2. (3p.) (Universal source coding with fixed length codes of rate R)

   Let $A_n \subset \mathcal{X}^n$ be the union of all type classes T_P^n with $H(P) \leqslant R$. Show that

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

   and that

   $$Q^n(A_n) \to 1 \tag{2}$$

   for every distribution Q on $\mathcal{X}$ with $H(Q) < R$. More exactly, prove that

   $$Q^n(A_n^c) = 2^{-n \cdot e_Q(R) + o(n)}, \; e_Q(R) = \min_{P: H(P) \geqslant R} D(P||Q). \tag{3}$$

   Hint: Use the results of Section 2 of the lecture notes.

   Remark: For each n take a **fixed length** n-code C_n as follows. Let $l_n = \lceil \log |A_n| \rceil$ and choose an injective mapping $f_n : A_n \to \{0\} \times \{0,1\}^{l_n}$. Then let $C_n : \mathcal{X}^n \to \{0,1\}^{l_n+1}$ be the code which encodes the element of A_n using f_n and encode all the other sequences of $\mathcal{X}^n$ with the all 1's sequence of length $l+1$. According to this exercise the code sequence $\{C_n, n = 1, 2, \ldots\}$ uses asymptotically R bits per source symbol, moroever, it is true simultaneously for each Q with $H(Q) < R$ that the Q^n probability that the source sequence can not be recovered from its code tends to 0 with exponent $e_Q(R)$ as $n \to \infty$.

3. (3p.) Let $Q = \{Q_n\}_{n=1}^{\infty}$ be the iid process on the alphabet $\{1, 2\}$ with marginal distribution $Q_1(1) = \frac{9}{10}$ and $Q_1(2) = \frac{1}{10}$. Determine the codeword assigned to $x_1^7 = 1211112$ by arithmetic coding determined by coding process Q, for both versions of arithmetic coding found on pages 427-428 of the lecture notes.

   Supplement: When you divide the interval $J(x_1^n)$ into two parts, let $J(x_1^n 1)$ be the left and $J(x_1^n 2)$ be the right subinterval.