6th homework set, Due June 11

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.

2. (2p.) Find the (ideal) codelength for encoding the sequence 2221111222 using (i) the coding process assigned to the class of i.i.d. processes (see page 481 of the lecture notes) and (ii) the coding process assigned to the class of (first order) Markov sources (see page 485 of the lecture notes). Interpret the result. (Actual codes are not required, only the codelengths)

3. (4p.) (Universal source coding with fixed length codes of rate R)

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

   $$\frac{1}{n} \log |A_n| \to R \tag{3}$$

   and that

   $$Q^n(A_n) \to 1 \tag{4}$$

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

   $$Q^n(A_n^c) = 2^{-n \cdot e_Q(R) + o(n)},\ e_Q(R) = \min_{P : H(P) \geqslant R} D(P||Q). \tag{5}$$

   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 $\{C_n, n = 1, 2, \dots\}$ 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 $n \to \infty$.

   (b) Show that the error exponent $e_Q(R)$ is best possible for every Q, i.e., for arbitrary $\tilde{A}_n$ satisfying (3) above, always

   $$\liminf_{n \to \infty} \frac{1}{n} \log Q^n(\tilde{A}_n^c) \geqslant -e_Q(R). \tag{6}$$

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

4. (3p.) (Application of exercise 2b of the third homework set)

   For binary valued i.i.d. $X_1, \dots, X_n$ with common distribution $Q = (Q(0), Q(1)) = (1-q, q)$. Let $p \leqslant q$. Show that

   $$Pr\left(\sum_{i=1}^{n} X_i \leqslant np\right) \leqslant 2^{-nD(p||q)} \tag{7}$$

where

$$D(p||q) = p \log \frac{p}{q} + (1-p) \log \frac{1-p}{1-q}. \tag{8}$$

How is this related to Sanov's theorem?

Hint: First prove that

$$Pr\left(X_i = 1 \Big| \sum_{i=1}^{n} X_i \leqslant np\right) \leqslant p \tag{9}$$

via determining

$$Pr\left(X_i = 1 \Big| \sum_{i=1}^{n} X_i = k\right),\ 0 \leqslant k \leqslant np. \tag{10}$$