5th homework set, Due June 11

1. (2p.) Determine (exactly) the normalized maximum likelihood distribution for binary sequences of length $n = 4$, coming from an i.i.d. process, i.e., the probabilities

$$\frac{P_{ML}(x_1^4)}{\sum_{y_1^4 \in \{0,1\}^4} P_{ML}(y_1^4)}, \; x_1^4 \in \{0,1\}^4, \tag{1}$$

as well as the corresponding (ideal) codelengths.

2. (4p.) Let $\mathcal{P}$ be the class of i.i.d. processes on A^∞ where $A = \{1, \ldots, k\}$, and let Q be the coding process treated in class,

$$Q(x_1^n) = \frac{\prod_{i=1}^k [(n_i - \frac{1}{2})(n_i - \frac{3}{2}) \cdots \frac{1}{2}]}{(n - 1 + \frac{k}{2})(n - 2 + \frac{k}{2}) \cdots \frac{k}{2}}, \tag{2}$$

where n_i is the number of occurrences of symbol i in x_1^n (the product in the numerator is defined to be 1 if $n_i = 0$). Prove that

$$\frac{\prod_{i=1}^k \left(\frac{n_i}{n}\right)^{n_i}}{Q(x_1^n)} \tag{3}$$

is bounded both above and below by a constant (depending on the alphabet size k only) times $n^{\frac{k-1}{2}}$. Finally conclude that it implies that for the class of i.i.d. processes $R_n^* = \frac{k-1}{2} \log n + O(1)$.

Hint: Using that

$$(n - \frac{1}{2})(n - \frac{3}{2}) \cdots \frac{1}{2} = \frac{(2n)!}{2^{2n} n!}, \tag{4}$$

rewrite (2) in terms of factorials (regarding the denominator, the cases $k =$ odd and $k =$ even have to be distinguished), and then apply Stirling's formula. Recall its strong version

$$\sqrt{2\pi} \cdot n^{n+\frac{1}{2}} e^{-n + \frac{1}{12(n+1)}} \leqslant n! \leqslant \sqrt{2\pi} \cdot n^{n+\frac{1}{2}} e^{-n + \frac{1}{12n}}. \tag{5}$$

Remark: It is worthwhile to repeat Exercise 4 of the 4th homework set and to read through Remark 6.1. of the lectures notes.

3. (4p.) Consider the coding process Q defined by (2) in case of $k = 2$. 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.

Hint: You can find the conditional probabilities needed for arithmetic coding on page 481 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.