Homework problem set 3

Your homework must be handed in **electronically via Moodle before Friday November 25th, 20:00h**. This deadline is strict and late submissions are graded with a 0. At the end of the course, the lowest of your 6 weekly homework grades will be dropped. You are strongly encouraged to work together on the exercises, including the homework. However, after this discussion phase, you have to write down and submit your own individual solution. Numbers alone are never sufficient, always motivate your answers.

Problem 1: Shannon code

(CT, exercise 5.25) Consider the following method for generating a code for a random variable X which takes on m values $\{1, 2, ..., m\}$. Assume that the probabilities are ordered su that $P_X(1) \geq P_X(2) \geq ... \geq P_X(m)$. Define

$$F_i := \sum k = 1^{i-1} P_X(k),$$

the sum of the probabilities of all symbols less than i. Then the Shannon code is defined by assigning the (binary representation of the) number $F_i \in [0, 1]$ as the codeword for i, where F_i is rounded off to $\lceil \log \frac{1}{P_X(i)} \rceil$ bits.

(a) Construct the code for the probability distribution $P_X(1) = \frac{1}{2}$, $P_X(2) = \frac{1}{4}$, $P_X(3) = P_X(4) = \frac{1}{8}$

(b) Construct the code for the probability distribution $P_Y(1) = P_Y(2) = P_Y(3) = \frac{1}{3}$.

(c) Show that the Shannon code is prefix-free.

(d) Show that the average length ℓ_S of the Shannon code satisfies

$$H(X) \leq \ell_S(P_X) < H(X) + 1.$$

   define shannon code, construct the code for some example distribution, and then show that the length is optimal. (and prefix-free)

Problem 2: Calculation of the typical set

To clarify the notion of a typical set $A_\varepsilon^{(n)}$ and the smallest set of high probability $B_\delta^{(n)}$, we will calculate these sets for a simple example. Consider a sequence of i.i.d. binary random variables $X_1, X_2, ...X_n$, where the probability that $P_X(1) = 0.6$ and $P_X(0) = 0.4$.

(a) **(1pt)** Calculate $H(X)$.

(b) **(3pt)** With $n = 25$ and $\varepsilon = 0.1$, which sequences fall in the typical set $A_\varepsilon^{(n)}$? What is the probability of the typical set? How many elements are there in the typical set? (This involves computation of a table of probabilities for sequences with k 1's, $0 \leq k \leq 25$, and finding those sequences that are in the typical set.)
Hint: Here is the table: `http://goo.gl/sQCPM0`

(c) **(2pt)** How many elements are there in the smallest set that has probability 0.9? In other words, what is $|B_\delta^{(n)}|$ for $n = 25$ and $\delta = 0.1$?

(d) **(2pt)** How many elements are there in the intersection $|A_\varepsilon^{(n)} \cap B_\delta^{(n)}|$ of the sets computed in parts (b) and (c)? What is the probability of this intersection?

Problem 3: 3

more to come