Information Theory: Exercise II

1) Let $E_1, \ldots, E_m$ be such that $E_m > \cdots > E_1 > 0$. We say that a random variable X has the Boltzmann distribution with parameter $\beta > 0$, if for every i,

$$Pr[X = E_i] = \frac{e^{-\beta E_i}}{\sum_j e^{-\beta E_j}}$$

Let E be such that $E_1 < E < \sum_i E_i/m$. Show that among all the distributions $p_1, \ldots, p_m$ such that $\sum_i p_i E_i = E$, the distribution with maximal entropy is a Boltzmann distribution.
What happens when $E > \sum_i E_i/m$?
What happens when $E = \sum_i E_i/m$?

2) Let X be a random variable with distribution $(p_1, \ldots, p_m)$, such that for every i the probability p_i is a power of 2 (i.e., it is 2^{-k} for some k). Prove that the average code length obtained by Huffman's code in this case is exactly $H(X)$.

3) Prove that the average code length obtained by Huffman's code on any random variable X is $\geq H(X)$.
Hint: use Question 1 from Exercise I.

4) Prove that the average code length obtained by Huffman's code on any random variable X is $\leq H(X) + 1$.
Hint: use Question 2 and the optimality of Huffman's code.

5) Assume that you want to generate a random variable X with distribution $(p_1, \ldots, p_m)$ by coin-flips as follows:

In each step, based on all previous coin-flips, you can either stop or flip another coin with distribution $(p, 1-p)$ of your choice (where the bias $(p, 1-p)$ is not necessarily the same in all steps). After you stop, you have to output a value in $\{1, \ldots, m\}$ (based on all coin-flips). The requirement is that for every i the probability to output i is p_i.

Show that in every procedure as above the expectation of the number of coin flips is at least $H(X)$ and that there exists such a procedure where the expectation of the number of coin flips is at most $H(X) + 1$.