PHYS 273, Fall 2016, Homework 1

Due date: Tuesday, October 4th, 2016

1. A source produces a character x that is chosen randomly from the alphabet $A = \{0, 1, \ldots, 9, a, b, \ldots, z\}$; with probability $1/3$, x is a numeral $(0, ..., 9)$; with probability $1/3$, x is a vowel (a, e, i, o, u); and with probability $1/3$ it is one of the 21 consonants. All numerals are equiprobable, and the same goes for vowels and consonants. Estimate the entropy of X regrouping first the events in the three groups of numerals, vowels and consonants and then calculating the entropy within the groups.

2. A random variable $x \in \{0, 1, 2, 3\}$ is selected by flipping a bent coin with bias f to determine whether the outcome is in $\{0, 1\}$ or $\{2, 3\}$; then either flipping a second bent coin with bias g or a third bent coin with bias h, respectively. Write down the probability distribution of X. Use also the regrouping property of the Shannon entropy to find the entropy of X.

3. An unbiased coin is flipped until one head is thrown. What is the entropy of the random variable $x \in \{1, 2, 3, \ldots\}$, the number of flips? Repeat the calculation for the case of a biased coin with probability f of coming up heads. Solve the problem both directly and by using the regrouping property of the Shannon entropy.

4. The frequency p_n of the n-th most frequent word in English is roughly approximated by $p_n = 0.1/n$ for $n \leq 12367$ and zero otherwise. This remarkable $1/n$ law is known as Zipfs law, and approximately applies to the word frequencies of many languages (Zipf, 1949). If we assume that English is generated by picking words at random according to this distribution, what is the entropy of English (per word)? [This calculation can be found in Prediction and entropy of printed English, C.E. Shannon, Bell Syst. Tech. J. 30, pp.5064 (1950), though some numerical errors are present so copying is not recommended!]

5. Problem 4.1, pg.66 from the book by MacKay, "Information Theory, Inference, and Learning Algorithms". The PDF can be downloaded from `http://www.inference.phy.cam.ac.uk/itprnn/book.pdf`.