HW Set # 1

Homework due by Tuesday February 3, at 2:40pm, at the *beginning* of class.

1. *Coin flips.* A fair coin is flipped until the first head occurs. Let X denote the number of flips required.

   (a) Find the entropy $H(X)$ in bits.

   (b) A random variable X is drawn according to this distribution. Find an "efficient" sequence of yes-no questions of the form, "Is X contained in the set S?" Compare $H(X)$ to the expected number of questions required to determine X.

2. *Entropy of functions.* Let X be a random variable taking on a finite number of values.

   (i) What is the (general) inequality relationship of $H(X)$ and $H(Y)$ if $Y = 2^X$? If $Y = \cos X$?

   (ii) Show that, in general, the entropy of a function of X is less than or equal to the entropy of X by justifying the following steps:

$$\begin{aligned} H(X, g(X)) &\stackrel{(a)}{=} H(X) + H(g(X) \mid X) \\ &\stackrel{(b)}{=} H(X); \\ H(X, g(X)) &\stackrel{(c)}{=} H(g(X)) + H(X \mid g(X)) \\ &\stackrel{(d)}{\geq} H(g(X)). \end{aligned}$$

   Thus $H(g(X)) \leq H(X)$.

   (iii) When is $H(g(X))$ *equal* to $H(X)$?

3. *Zero conditional entropy.* Show that if $H(Y|X) = 0$, then Y is a function of X, i.e., for all x with $P(x) > 0$, there is only one possible value of y with $P(x, y) > 0$.

4. *Entropy of a disjoint mixture.* Let X_1 and X_2 be discrete random variables drawn according to probability mass functions $P_1(\cdot)$ and $P_2(\cdot)$ over the respective alphabets $\mathcal{X}_1 = \{1, 2, \ldots, m\}$ and $\mathcal{X}_2 = \{m+1, \ldots, n\}$. Let

$$X = \begin{cases} X_1, & \text{with probability } \alpha, \\ X_2, & \text{with probability } 1 - \alpha. \end{cases}$$

   (a) Find $H(X)$ in terms of $H(X_1)$ and $H(X_2)$ and α.

(b) Maximize over α to show that $2^{H(X)} \leq 2^{H(X_1)} + 2^{H(X_2)}$ and interpret using the notion that $2^{H(X)}$ is the effective alphabet size.

5. *Run length coding.* Let $X_1, X_2, \ldots, X_n$ be (possibly dependent) binary random variables. Suppose one calculates the run lengths $R = (R_1, R_2, \ldots)$ of this sequence (in order as they occur). For example, the sequence $X_1^n = 0001100100$ yields run lengths $R = (3, 2, 2, 1, 2)$. Compare $H(X_1, X_2, \ldots, X_n)$, $H(R)$ and $H(X_n, R)$. Show all equalities and inequalities, and bound all the differences.

6. *Markov's inequality for probabilities.* Let $P(x)$ be a probability mass function. Prove, for all $d \geq 0$,

$$\Pr\{P(X) \leq d\} \log\left(\frac{1}{d}\right) \leq H(X).$$

7. *The AEP and source coding.* A memoryless source emits a sequence of independent Bernoulli bits with probabilities $P(1) = 0.005$ and $P(0) = 0.995$. The digits are taken 100 at a time and a binary codeword is provided for every sequence of 100 digits containing three or fewer ones.

   (a) Assuming that all codewords are the same length, find the minimum length required to provide codewords for all sequences with three or fewer ones.

   (b) Calculate the probability of observing a source sequence for which no codeword has been assigned.

   (c) Use Chebyshev's inequality to bound the probability of observing a source sequence for which no codeword has been assigned. Compare this bound with the actual probability computed in part (b).

8. *AEP.* Let $X_1, X_2, \ldots$ be IID random variables drawn according to the probability mass function $P(x), x \in \{1, 2, \ldots, m\}$. Thus $P^n(x_1, x_2, \ldots, x_n) = \prod_{i=1}^n P(x_i)$. We know that $-\frac{1}{n}\log p(X_1, X_2, \ldots, X_n) \to H(X)$ in probability. Let $Q^n(x_1, x_2, \ldots, x_n) = \prod_{i=1}^n Q(x_i)$, where Q is another probability mass function on $\{1, 2, \ldots, m\}$.

   Evaluate $\lim -\frac{1}{n}\log Q^n(X_1, X_2, \ldots, X_n)$, where $X_1, X_2, \ldots$ are IID $\sim P$.

9. *Random box size.* An n-dimensional rectangular box with sides $X_1, X_2, X_3, \ldots, X_n$ is to be constructed. The volume is $V_n = \prod_{i=1}^n X_i$. The edge length l of a n-cube with the same volume as the random box is $l = V_n^{1/n}$. Let $X_1, X_2, \ldots$ be IID uniform random variables over the unit interval $[0, 1]$. Find the value of the limit (in probability or a.s.), $\lim_{n\to\infty} V_n^{1/n}$, and compare it to $(EV_n)^{\frac{1}{n}}$. Clearly the expected edge length does not capture the idea of the volume of the box.