Introduction to Information Theory, Fall 2020

Practice problems for exercise class #3

You do **not** have to hand in these exercises, they are for your practice only.

0. **Exercises from MacKay:** 5.8, 5.19, 5.20, 5.31

1. **Probability theory refresher:**

   (a) Show that if X has a Bernoulli distribution with parameter p, i.e. $X \sim \text{Bernoulli}(p)$, then

   $$\mathbb{E}X = p,$$
   $$\text{Var}(X) = p(1-p).$$

   (b) Use part (a) to compute the expectation value and variance of a random variable with a $\text{Binomial}(n, p)$ distribution.

   (c) In class we discussed two ways of defining the variance of random variable X:

   $$\text{Var}(X) = \mathbb{E}\left[(X - \mathbb{E}X)^2\right] = \mathbb{E}[X^2] - (\mathbb{E}X)^2.$$

   Show that the two formulas are equivalent.

2. **Computing entropies:**

   (a) Compute the entropy $H(X)$ of a random variable X that has three possible outcomes that occur with probabilities $1/2$, $1/4$, and $1/4$. Compute $H(X)$.

   (b) Consider two random variables Y and Z with the following joint distribution:

$Z \setminus Y$	sun	moon
morning	$1/2$	0
evening	$1/4$	$1/4$

   Compute $H(Y, Z)$, $H(Y)$, and $H(Z)$. Are Y and Z independent?

   (c) Let $X_1, \ldots, X_n$ be n independent random variables. We abbreviate $X^n = (X_1, \ldots, X_n)$. Show that

   $$H(X^n) = \sum_{i=1}^{n} H(X_i).$$

   In particular, what is $H(X^n)$ if $X_1, \ldots, X_n \overset{\text{IID}}{\sim} P$?

3. **Convexity (mathematics challenge):** In class, we said that a (twice differentiable) function $f \colon \mathbb{R} \to \mathbb{R}$ is *convex* if $f'' \geqslant 0$.

   (a) Show that if f is convex, then for any $x, z \in \mathbb{R}$

   $$f(z) \leqslant f(x) + f'(z)(z - x).$$

   Hint: Use the fundamental theorem of calculus.

(b) Show that if f is convex, then for any $x, y \in \mathbb{R}$ and $p \in [0, 1]$

$$f(px + (1-p)y) \leqslant pf(x) + (1-p)f(y). \tag{1}$$

In mathematics, this inequality is often taken as the definition of convexity, since it also works for functions that are not differentiable.
Hint: Apply part (a) with $z = px + (1-p)y$.

(c) Interpret Eq. (1) graphically. (We discussed this briefly in class!)

(d) Explain the relation of Eq. (1) to Jensen's inequality.

(e) **Optional:** Prove Jensen's inequality. *Hint: Use induction.*