EEE 551 Information Theory (Spring 2022)

Homework 1

Due: February 2, 2022

---

1. *Example of joint entropy.* Let $p(x, y)$ be given by

$$
\begin{array}{cc|ccc}
 & & & X & \\
 & & 0 & 1 & 2 \\
\hline
\multirow{2}{*}{Y} & 0 & \frac{1}{3} & \frac{1}{6} & 0 \\
 & 1 & 0 & \frac{1}{6} & \frac{1}{3} \\
\hline
\end{array}
$$

   Find:

   (a) $H(X)$, $H(Y)$.

   (b) $H(X|Y)$, $H(Y|X)$.

   (c) $H(X, Y)$.

   (d) $H(Y) - H(Y|X)$.

   (e) $I(X; Y)$.

   (f) Draw a Venn diagram for the quantities in parts (a) through (e).

2. Problem 2.4 from Cover-Thomas: *Entropy of functions of a random variable.* Let X be a discrete random variable. Show that the entropy of a function of X is less than or equal to the entropy of X by justifying the following steps:

$$
\begin{aligned}
H(X, g(X)) &\overset{\textbf{(a)}}{=} H(X) + H(g(X)|X) \\
&\overset{\textbf{(b)}}{=} H(X), \\
H(X, g(X)) &\overset{\textbf{(c)}}{=} H(g(X)) + H(X|g(X)) \\
&\overset{\textbf{(d)}}{\geq} H(g(X)).
\end{aligned}
$$

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

3. *Yes/No question.* Let X be a random variable with PMF given by

x	a	b	c	d	e	f	g
$p(x)$	1/4	1/4	1/8	1/8	1/8	1/16	1/16

   Let Y be a random variable with PMF given by

y	a	b	c	d
$p(y)$	0.4	0.3	0.2	0.1

(a) Devise a strategy to determine X by a series of Yes/No questions such that the expected number of questions is exactly equal to $H(X)$. Each question may depend on the outcome of the previous question.

(b) For Y, find a strategy of Yes/No questions to minimize the expected number of questions. What is the smallest possible expected number of questions? In this case, how does it compare to $H(Y)$?

4. *Mutual information vs. conditional mutual information.* In general, conditioning neither decreases nor increases mutual information; i.e. either the mutual information $I(X;Y)$ or the conditional mutual information $I(X;Y|Z)$ can be larger. This problem shows that under certain conditions, one can conclude that one of these quantities is larger than the other. Prove the following two statements:

   (a) If $X \to Y \to Z$, then $I(X;Y|Z) \le I(X;Y)$.

   (b) If X and Z are independent, then $I(X;Y) \le I(X;Y|Z)$.

   Hint: Expand $I(X;Y,Z)$ using the chain rule in two different ways.

5. Problem 2.27 from Cover-Thomas: *Grouping rule for entropy.* Let $\mathbf{p} = (p_1, p_2, \ldots, p_m)$ be a probability distribution on m elements (i.e., $p_i \ge 0$ and $\sum_{i=1}^m p_i = 1$). Define a new distribution $\mathbf{q}$ on $m-1$ elements as $q_1 = p_1, q_2 = p_2, \ldots, q_{m-2} = p_{m-2}$, and $q_{m-1} = p_{m-1} + p_m$ [i.e., the distribution $\mathbf{q}$ is the same as $\mathbf{p}$ on $\{1, 2, \ldots, m-2\}$, and the probability of the last element in $\mathbf{q}$ is the sum of the last two probabilities of $\mathbf{p}$]. Show that

$$H(\mathbf{p}) = H(\mathbf{q}) + (p_{m-1} + p_m) H\left(\frac{p_{m-1}}{p_{m-1} + p_m}, \frac{p_m}{p_{m-1} + p_m}\right).$$

6. *Rényi Entropy.* Rényi entropy is a different way of defining an "entropy" that generalizes Shannon's measure. Rényi entropy has a parameter α, which can be any positive number except 1. Rényi entropy is defined as

$$H_\alpha(X) = \frac{1}{1-\alpha} \log \left[\sum_{x \in \mathcal{X}} p(x)^\alpha\right].$$

   (a) Find $H_\alpha(X)$ if X is a uniform random variable with an alphabet of size m. Compare your answer to the corresponding value for the standard entropy.

   (b) Plot $H_\alpha(X)$ as a function of α where X is Bern(0.1).

   (c) Prove that $\lim_{\alpha \to 1} H_\alpha(X) = H(X)$, where $H(X)$ is the standard entropy. For this reason, the standard entropy is sometimes written $H_1(X)$. *Hint:* Use L'Hôpital's rule.

   (d) If X and Y are independent, prove that $H_\alpha(X,Y) = H_\alpha(X) + H_\alpha(Y)$.

   (e) Show that Rényi entropy does *not* satisfy the grouping rule, as defined in Problem 5. To do this, find a distribution $\mathbf{p}$ where the formula in Problem 5 is violated for Rényi entropy.

7. Problem 2.32 from Cover-Thomas: *Fano.* We are given the following joint distribution on (X, Y):

		Y		
		a	b	c
X	1	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{12}$
	2	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{12}$
	3	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{6}$

   Let $\hat{X}(Y)$ be an estimator for X (based on Y) and let $P_e = \Pr\{\hat{X}(Y) \ne X\}$.

(a) Find the minimum probability of error estimator $\hat{X}(Y)$ and the associated probability P_e.

(b) Evaluate Fano's inequality for this problem and compare.

8. Problem 2.37 from Cover-Thomas: *Relative entropy.* Let X, Y, Z be three random variables with a joint probability mass function $p(x, y, z)$. The relative entropy between joint distribution and the product of the marginals is

$$D(p(x, y, z) \| p(x)p(y)p(z)) = E\left[\log \frac{p(x, y, z)}{p(x)p(y)p(z)}\right].$$

Expand this in terms of entropies. When is this quantity zero?

9. Problem 2.42 from Cover-Thomas: *Inequalities.* Which of the following inequalities are generally $\geq, =, \leq$? Label each with $\geq$, $=$, or $\leq$.

(a) $H(5X)$ vs. $H(X)$

(b) $I(g(X); Y)$ vs. $I(X; Y)$

(c) $H(X_0|X_{-1})$ vs. $H(X_0|X_{-1}, X_1)$

(d) $H(X, Y)/(H(X) + H(Y))$ vs. 1