Entropy, Relative Entropy and Mutual Information Exercises

all exercises except the altered 2.8 are by Cover and Thomas

Exercise 2.1:
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. The following expressions may be useful:

$$\sum_{n=1}^{\infty} r^n = \frac{r}{1-r}, \qquad \sum_{n=1}^{\infty} nr^n = \frac{r}{(1-r)^2}$$

(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.

Exercise 2.3:
Minimum entropy. What is the minimum value of $H(p_1, \ldots, p_n) = H(\mathbf{p})$ as $\mathbf{p}$ ranges over the set of n-dimensional proabability vectors? Find all $\mathbf{p}$'s which achieve this minimum.

Exercise 2.8 (adjusted):
Wimbledon Men's Final. The Men's Final is a five-set series that terminates as soon as either player wins three sets. Let X be the random variable that represents the outcome of a Men's Final between players Andre Agassi (A) and Björn Borg (B); some possible values for X are AAA, BABAB, BBAAA. Let Y be the number of sets played, which ranges from 3 to 5. Assuming that A and B are equally matched and that the results of the sets are independent, calculate $H(X)$, $H(Y)$, $H(Y|X)$, $H(X|Y)$, $H(X,Y)$, and $I(X;Y)$.

Exercise 2.11:
Average entropy. Let $H(p) = -p\log_2 p - (1-p)\log_2(1-p)$ be the binary entropy function.

(a) Evaluate $H(1/4)$.

(b) Calculate the average entropy $H(p)$ when the probability p is chosen uniformly in the range $0 \le p \le 1$.

It may help to know that:

$$\int u^n \ln u \, du = \frac{u^{n+1}}{n+1}\left(\ln u - \frac{1}{n+1}\right) + C$$

Exercise 2.16:

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

X \ Y	0	1
0	1/3	1/3
1	0	1/3

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 quantites in (a) through (e).

Exercise 2.18:
Entropy of a sum. Let X and Y be random variables that take on values $x_1, x_2, \ldots, x_r$ and $y_1, y_2, \ldots, y_s$. Let $Z = X + Y$.

(a) Show that $H(Z|X) = H(Y|X)$. Argue that if X, Y are independent then $H(Y) \leq H(Z)$ and $H(X) \leq H(Z)$. Thus the addition of independent random variables adds uncertainty.

(b) Give an example (of necessarily dependent random variables) in which $H(X) > H(Z)$ and $H(Y) > H(Z)$.

(c) Under what conditions does $H(Z) = H(X) + H(Y)$?

Exercise 2.21:
Data processing. Let $X_1 \to X_2 \to X_3 \to \cdots \to X_n$ form a Markov chain in this order; i.e., let

$$p(x_1, x_2, \ldots, x_n) = p(x_1)p(x_2|x_1)\cdots p(x_n|x_{n-1}) \tag{1}$$

Reduce $I(X_1; X_2, \ldots, X_n)$ to its simplest form.