Problem 1: Properties of mutual information. Let X, Y_1, Y_2 be three RVs.
(a) Given $I(X;Y_1) = I(X;Y_2) = 0$, does it follow that $I(X;Y_1,Y_2) = 0$?
(b) Given $I(X;Y_1) = I(X;Y_2) = 0$, does it follow that $I(Y_1;Y_2) = 0$?

Problem 1

(a) No. Let X, Y_1, Y_2 be three pairwise independent Ber($\frac{1}{2}$) RV,
$I(X;Y_1) = I(X;Y_2) = I(Y_1;Y_2) = 0$ (See Sept. 14 Recitation)

Let $X = Y_1 \oplus Y_2$ then

$I(X;Y_1,Y_2) = H(X) - H(X|Y_1,Y_2) = H(X) \neq 0$ (with $H(X|Y_1,Y_2) = 0$)

(b) No. Let $Y_1 = Y_2$ where Y_1 is a RV s.t. $I(X;Y_1) = 0$. This implies $I(X;Y_2) = 0$. However, $I(Y_1;Y_2) = H(Y_1) - H(Y_1|Y_2) = H(Y_1) \neq 0$ (with $H(Y_1|Y_2) = 0$)

Problem 2: Data Processing Inequality. Let the RVs X, U, Z form a Markov chain, $X \to Y \to Z$.
(a) Show that $H(X|Y) = H(X|Y,Z)$.
(b) Show that $H(X|Y) \leq H(X|Z)$.
(c) Show that $I(X;Y) \geq I(X;Z)$.
(d) Show that $I(X;Z|Y) = 0$.

Problem 2

(a) $H(X,Z|Y) = H(Z|Y) + H(X|Y,Z)$
$= H(X|Y) + H(Z|X,Y)$

Since $X \to Y \to Z$ is a Markov Chain we have

$p(z|x,y) = p(z|y)$

$\Rightarrow H(Z|Y) = H(Z|X,Y)$

$\Rightarrow H(X|Y,Z) = H(X|Y)$

(b) $H(X|Y) \overset{\text{from (a)}}{=} H(X|Y,Z) \leq H(X|Z)$ (conditioning decreases entropy)

(c) $I(X;Y) - I(X;Z) = [H(X) - H(X|Y)] - [H(X) - H(X|Z)]$
$= H(X|Z) - H(X|Y) \geq 0$ (from b)

(d) $I(X;Z|Y) = H(X|Y) - H(X|Z,Y) = 0$ (from a)

6) **An AEP-like limit.** Let $X_1, X_2, \ldots$ be i.i.d. drawn according to probability mass function $p(x)$. Find

$$\lim_{n\to\infty} [p(X_1, X_2, \ldots, X_n)]^{\frac{1}{n}}.$$

6) *An AEP-like limit.* $X_1, X_2, \ldots,$ i.i.d. $\sim p(x)$. Hence $\log(X_i)$ are also i.i.d. and

$$\begin{aligned}
\lim(p(X_1, X_2, \ldots, X_n))^{\frac{1}{n}} &= \lim 2^{\log(p(X_1,X_2,\ldots,X_n))^{\frac{1}{n}}} \\
&= 2^{\lim \frac{1}{n}\sum \log p(X_i)} \text{ a.e.} \\
&= 2^{E(\log(p(X)))} \text{ a.e.} \\
&= 2^{-H(X)} \text{ a.e.}
\end{aligned}$$

by the strong law of large numbers (assuming of course that H(X) exists).

10) **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 i.i.d. uniform random variables over the unit interval $[0,1]$. Find $\lim_{n\to\infty} V_n^{1/n}$, and compare to $(EV_n)^{\frac{1}{n}}$. Clearly the expected edge length does not capture the idea of the volume of the box. The geometric mean, rather than the arithmetic mean, characterizes the behavior of products.

10) *Random box size.* The volume $V_n = \prod_{i=1}^n X_i$ is a random variable, since the X_i are random variables uniformly distributed on $[0,1]$. V_n tends to 0 as $n \to \infty$. However

$$\log_e V_n^{\frac{1}{n}} = \frac{1}{n}\log_e V_n = \frac{1}{n}\sum \log_e X_i \to E(\log_e(X)) \text{ a.e.}$$

by the Strong Law of Large Numbers, since X_i and $\log_e(X_i)$ are i.i.d. and $E(\log_e(X)) < \infty$. Now

$$E(\log_e(X_i)) = \int_0^1 \log_e(x)\,dx = -1$$

Hence, since e^x is a continuous function,

$$\lim_{n\to\infty} V_n^{\frac{1}{n}} = e^{\lim_{n\to\infty}\frac{1}{n}\log_e V_n} = \frac{1}{e} < \frac{1}{2}.$$

Thus the "effective" edge length of this solid is e^{-1}. Note that since the X_i's are independent, $E(V_n) = \prod E(X_i) = (\frac{1}{2})^n$. Also $\frac{1}{2}$ is the arithmetic mean of the random variable, and $\frac{1}{e}$ is the geometric mean.

2) **How many fingers has a Martian?** Let

$$S = \begin{pmatrix} S_1, \ldots, S_m \\ p_1, \ldots, p_m \end{pmatrix}.$$

The S_i's are encoded into strings from a D-symbol output alphabet in a uniquely decodable manner. If $m = 6$ and the codeword lengths are $(l_1, l_2, \ldots, l_6) = (1, 1, 2, 3, 2, 3)$, find a good lower bound on D. You may wish to explain the title of the problem.

2) *How many fingers has a Martian?*

Uniquely decodable codes satisfy Kraft's inequality. Therefore

$$f(D) = D^{-1} + D^{-1} + D^{-2} + D^{-3} + D^{-2} + D^{-3} \le 1. \quad (379)$$

We have $f(2) = 7/4 > 1$, hence $D > 2$. We have $f(3) = 26/27 < 1$. So a possible value of D is 3. Our counting system is base 10, probably because we have 10 fingers. Perhaps the Martians were using a base 3 representation because they have 3 fingers. (Maybe they are like Maine lobsters ?)

3) **Slackness in the Kraft inequality.** An instantaneous code has word lengths $l_1, l_2, \ldots, l_m$ which satisfy the strict inequality

$$\sum_{i=1}^{m} D^{-l_i} < 1.$$

The code alphabet is $\mathcal{D} = \{0, 1, 2, \ldots, D-1\}$. Show that there exist arbitrarily long sequences of code symbols in $\mathcal{D}^*$ which cannot be decoded into sequences of codewords.

3) *Slackness in the Kraft inequality.* Instantaneous codes are prefix free codes, i.e., no codeword is a prefix of any other codeword. Let $n_{max} = \max\{n_1, n_2, ..., n_q\}$. There are $D^{n_{max}}$ sequences of length n_{max}. Of these sequences, $D^{n_{max}-n_i}$ start with the i-th codeword. Because of the prefix condition no two sequences can start with the same codeword. Hence the total number of sequences which start with some codeword is $\sum_{i=1}^{q} D^{n_{max}-n_i} = D^{n_{max}} \sum_{i=1}^{q} D^{-n_i} < D^{n_{max}}$. Hence there are sequences which do not start with any codeword. These and all longer sequences with these length n_{max} sequences as prefixes cannot be decoded. (This situation can be visualized with the aid of a tree.)

Alternatively, we can map codewords onto dyadic intervals on the real line corresponding to real numbers whose decimal expansions start with that codeword. Since the length of the interval for a codeword of length n_i is D^{-n_i}, and $\sum D^{-n_i} < 1$, there exists some interval(s) not used by any codeword. The binary sequences in these intervals do not begin with any codeword and hence cannot be decoded.