ECE 6605

HW#2, Assigned Sept 2, Due Sept 15.

1) Network Information Theory. Suppose a Markov chain starts in one of n values, i.e."states", necks down to one of k < n values and then fans back to m > k states. That is

$X \to Y \to Z$ is a Markov Chain in that order and

$$X \in \{1,2,...,n\}, Y \in \{1,2,...,k\}, Z \in \{1,2,...,m\}$$

(a) Show that the dependence of X and Z is limited by the bottleneck by proving that I(X;Z) ≤ log k:

(b) Evaluate I(X;Z) for k = 1, and conclude that no dependence can survive such a bottleneck.

2) *Data Storage Information Theory.* The following sequences of bits are very common in data storage devices.

Let $X = (X_1, X_2, ..., X_n)$ be (possibly dependent) binary random variables. Suppose one calculates the lengths of runs of zeros or ones and let $R = (R_1, R_2, ..., R_m)$ be a sequence of runs. For example, the sequence X = (0001100100) yields run lengths R = (3; 2; 2; 1; 2). As a note, in CD and DVD any run must satisfy $3 \le R_i \le 11$ for all i.

Compare $H(X_1, X_2, ..., X_n), H(R), H(X_n, R)$, namely find as many inequality relations or expressions relating each.

3. Consider a sequence of n binary random variables $(X_1, X_2, ..., X_n)$. Each sequence with an even number of 1's has probability $2^{-(n-1)}$ and each sequence with an odd number of 1's has probability 0. Find the mutual informations ,

$$I(X_1; X_2), I(X_2; X_3 \mid X_1), .., I(X_{n-1}; X_n \mid X_1, ..., X_{n-2}),$$

4. Let each term in $(X_1, X_2, ..., X_n)$ be i.i.d. drawn according to probability mass function p(x). Find

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

5. The problem relates both to the AEP and decision theory where given a set of data we are trying to find which distribution it comes from.

Let $X_1, X_2, X_3, ...,$ be independent identically distributed random variables drawn according to the probability mass function $p(x), x \in \{1,2,...,m\}$ p(x). Thus $p(x_1,...,x_n) = \prod_{i=1}^{n} p(x_i)$. We know that $-\frac{1}{n}\log p(x_1,...,x_n) \to H(X)$. Let $q(x_1,...,x_n) = \prod_{i=1}^{n} q(x_i)$ where q(x) is another probability mass function on $\{1,2,...,m\}$

(a) Evaluate $\lim_{n\to\infty} -\frac{1}{n}\log q(x_1,...,x_n)$ where $X_1, X_2, X_3, ...,$ are i.i.d. each $X_1, X_2, X_3, ...,$ with density p(x)

(b) Now evaluate the limit of the log likelihood ratio

$$\lim_{n\to\infty} \frac{1}{n}\log\frac{q(x_1,...,x_n)}{p(x_1,...,x_n)}$$

when $X_1, X_2, X_3, ...,$ are iid with probability mass function p(x). This result shows that the odds favoring q(x) are exponentially small when p(x) is true.