UC Berkeley
Department of Electrical Engineering and Computer Sciences

EECS 126: PROBABILITY AND RANDOM PROCESSES

Problem Set 7
Fall 2019

1. **r^{th} Moment Convergence**

   Let X_n converge in the r^{th} mean to X. Prove that it converges in probability.

2. **Mutual Information and Channel Coding**

   The **mutual information** of X and Y is defined as

   $$I(X;Y) := H(X) - H(X \mid Y)$$

   Here, $H(X \mid Y)$ denotes the **conditional entropy** of X given Y, which is defined as:

   $$\begin{aligned} H(X \mid Y) &= \sum_{y \in \mathcal{Y}} p_Y(y) H(X \mid Y = y) \\ &= \sum_{y \in \mathcal{Y}} p_Y(y) \sum_{x \in \mathcal{X}} p_{X|Y}(x \mid y) \log_2 \frac{1}{p_{X|Y}(x \mid y)} \end{aligned}$$

   The interpretation of conditional entropy is the average amount of uncertainty remaining in the random variable X after observing Y. The interpretation of mutual information is therefore the amount of information about X gained by observing Y.

   The channel coding theorem says that if X is passed into the channel and Y is received, then the capacity of the channel is

   $$C = \max_{p_X} I(X;Y) = \max_{p_X} H(X) - H(X \mid Y)$$

   (a) Let X be the roll of a fair die and $Y = \mathbf{1}\{X \geq 5\}$. What is $H(X \mid Y)$?

   (b) Suppose the channel is a noiseless binary channel, i.e. $X \in 0, 1$ and $Y = X$. Use the theorem to find C.

(c) Consider a binary erasure channel with probability of erasure p. Use the theorem to find C.

Hint: To find the optimal p_X, it is helpful to let $p_X(1) = P(X = 1) = \alpha$.

3. Huffman Questions

Consider a set of n objects. Let $X_i = 1$ or 0 accordingly as the i-th object is good or defective. Let $X_1, X_2, \ldots, X_n$ be independent with $\mathbb{P}(X_i = 1) = p_i$; and $p_1 > p_2 > \cdots > p_n > 1/2$. We are asked to determine the set of all defective objects. Any yes-no question you can think of is admissible.

(a) Propose an algorithm based on Huffman coding in order to identify all defective objects.

(b) If the longest sequence of questions is required by nature's answers to our questions which are based on Huffman coding, then what (in words) is the last question we should ask? And what two sets are we distinguishing with this question?

Note: This problem is related to the 'Entropy and Information Content' section of Lab 4.