3F1: Signals and Systems

INFORMATION THEORY

Examples Paper

1. The output of a discrete memoryless source consists of the possible letters $X_1, X_2, \cdots, X_n$, which occur with probabilities $P_1, P_2, \cdots, P_n$, respectively. Prove that the entropy $H(X)$ of the source is at most $\log_2(n)$.

2. A discrete memoryless source has an alphabet of eight letters, $x_i, i = 1, 2, \cdots, 8$ with probabilities 0.25, 0.20, 0.15, 0.12, 0.10, 0.08, 0.05 and 0.05.

   (a) Use the Huffman encoding to determine a binary code for the source output.
   (b) Determine the average codeword length $L$.
   (c) Determine the entropy of the source and hence its efficiency.

3. Show that for statistically independent events

   $$H(X_1, X_2, \cdots, X_n) = \sum_{i=1}^{n} H(X_i)$$

4. A five-level non-uniform quantizer for a zero-mean signal results in the 5 levels $-b, -a, 0, a, b$ with corresponding probabilities of occurrence $p_{-b} = p_b = 0.05$, $p_{-a} = p_a = 0.1$ and $p_0 = 0.7$.

   (a) Design a Huffman code that encodes one signal sample at a time and determine the average bit rate per sample.
   (b) Design a Huffman code that encodes two output samples at a time and determine the average bit rate per sample.
   (c) What are the efficiencies of these two codes?

5. Given two random variables $X$ and $Y$, $I(X; Y)$ is defined as:

   $$I(X; Y) = \sum_{x \in X, y \in Y} P(x, y) \log_2 \left( \frac{P(x|y)}{P(x)} \right)$$

   Show that $I(X; Y) = I(Y; X)$
6. What is the entropy of the following continuous probability density functions?

(a) \[ P(x) = \begin{cases} 
0 & x < -2 \\
0.25 & -2 < x < 2 \\
0 & x > 2 
\end{cases} \]

(b) \[ P(x) = \frac{1}{2} e^{-\lambda|x|} \]

(c) \[ P(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2} \]

7. Continuous variables \( X \) and \( Y \) are independent and normally distributed with standard deviation \( \sigma = 1 \).

\[ P(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \quad P(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2} \]

A variable \( Z \) is defined by \( z = x + y \). What is the mutual information of \( X \) and \( Z \)?

8. A symmetric binary communications channel operates with signalling levels of \( \pm 2 \) volts at the detector in the receiver, and the rms noise level at the detector is 0.5 volts. The binary symbol rate is 100 kbit/s.

(a) Determine the probability of error on this channel and hence, based on mutual information, calculate the theoretical capacity of this channel for error-free communication.

(b) If the binary signalling were replaced by symbols drawn from a continuous process with a Gaussian (normal) pdf with zero mean and the same mean power at the detector, determine the theoretical capacity of this new channel, assuming the symbol rate remains at 100 ksym/s and the noise level is unchanged.

**Numerical Answers**

2. b) 2.83 bits; c) 2.798 bits, 98.9%

4. a) 1.6 bit / sample; b) 1.465 bit / sample; c) 91.05%, 99.44%

6. a) \( \log_2(4) = 2 \); b) \( \log_2(2e/\lambda) \); c) \( \log_2(\sigma\sqrt{2\pi e}) \)

7. 0.5 bit

8. a) \( p_e = 3.17 \times 10^{-5}, \quad 99.948 \text{ kbit/s.} \); b) 204.37 kbit/s.

Nick Kingsbury, November 12, 2012.