HW6 - Introduction to Information Theory
due December 22

1. Consider the probability distribution $p(x, y)$ for $x, y \in \{0, 1\}$:

$$p(x,y) = \begin{Bmatrix} 1/6, & x=0, y=0, \\ 1/6, & x=0, y=1, \\ 0, & x=1, y=0, \\ 2/3, & x=1, y=1 \end{Bmatrix} \tag{1}$$

   Find

   (a) $H(X)$, $H(Y)$,

   (b) $H(X|Y)$, $H(Y|X)$,

   (c) $H(X, Y)$,

   (d) $I(X; Y)$.

2. Show that the following expressions hold. X, Y, Z are jointly distributed random variables.

   (a) $H(X, Y|Z) \geq H(X|Z)$,

   (b) $I(X, Y; Z) \geq I(X; Z)$,

   (c) $H(X, Y, Z) - H(X, Y) \leq H(X, Z) - H(X)$,

   (d) $I(X; Z|Y) \geq I(Z; Y|X) - I(Z; Y) + I(X; Z)$.

3. Assume that we have four symbols, namely a, b, c, d. These symbols are mapped to codewords of $10, 00, 11, 110$, respectively.

   (a) Does the given code have the prefix property?

   (b) Is this a uniquely decodable code? Prove it.

4. The probability distributions $p(x)$ and $q(x)$ are defined on the same random variable X with 5 outcomes, $\{1, 2, 3, 4, 5\}$. $C_1(x)$ and $C_2(x)$ are two different codes on X:

Symbol	$p(x)$	$q(x)$	$C_1(x)$	$C_2(x)$
1	1/2	1/2	0	0
2	1/4	1/8	10	100
3	1/8	1/8	110	101
4	1/16	1/8	1110	110
5	1/16	1/8	1111	111

(a) Find $H(p), H(q), D(p||q)$ and $D(q||p)$.

(b) Prove the optimality of C_1 and C_2 on X under p and q, respectively.

(c) We decide to use C_2 when the distribution is p. Find the average codeword length. Find its difference to the entropy of p.

(d) We decide to use C_1 when the distribution is q. Find the average codeword length. How much do we lose due to miscoding?

5. For each of the following codes, find a probability assignment that results in Huffman coding. If there is no such a probability assignment, explain why.

(a) {0,10,11}

(b) {00,01,10,110}

(c) {01,10}.

6. Consider the probability distribution on X with 8 outcomes:

$p(X = 1) = p(X = 2) = 2p(X = 3) = 2p(X = 4) = 4p(X = 5) = 4(X = 6) = 4p(X = 7) = 4p(X = 8)$.

(a) Give the Huffman binary code.

(b) Find the expected code length of a) and $H(X)$.

(c) Find the Huffman 3-ary code.

7. For each of the following, find the differential entropy, $h(X) = -\int f \ln f$.

(a) $f(x) = \lambda e^{-\lambda x}$ for nonnegative x.

(b) $f(x) = 1/2\lambda e^{-\lambda|x|}$ for real x.

(c) $f(x)$ where, $X = X_1 + X_2$ where X_1 and X_2 are independent with $f(x_i) = N(\mu_i, \sigma_i^2)$, $i = 1, 2$.

8. Consider the channel transition matrix, Q:

$$Q = \begin{bmatrix} 1 & 0 \\ 1/2 & 1/2 \end{bmatrix} \tag{2}$$

where $Q(i, j)$ gives the probability of receiving y_j if x_i is transmitted. Assume $x_1 = y_1 = 0$ and $x_2 = y_2 = 1$.

(a) Draw the channel model with the transmitted bits on the left side, the received bits on the right side and the transition probabilities connecting them. What is the name of this channel?

(b) Find the capacity of this channel.

(c) Find the input distribution that achieves the capacity.

9. Find the capacity of following channels using the given probability transition matrices.

   (a) Outcomes: $\{0, 1, 2\}$ for both the input and output distributions,

$$p(y|x) = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix} \tag{3}$$

   (b) Outcomes: $\{0, 1, 2\}$ for both the input and output distributions,

$$p(y|x) = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \end{bmatrix} \tag{4}$$

   (c) Outcomes: $\{0, 1, 2, 3\}$ for both the input and output distributions,

$$p(y|x) = \begin{bmatrix} p & 1-p & 0 & 0 \\ 1-p & p & 0 & 0 \\ 0 & 0 & q & 1-q \\ 0 & 0 & 1-q & q \end{bmatrix} \tag{5}$$

   (d) Outcomes: $\{0, 1, 2\}$ for both the input and output distributions,

$$p(y|x) = \begin{bmatrix} 1-p & p & 0 \\ p & 1-p & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{6}$$

10. Consider a cascaded BSC ($x \to y \to z$) with two identical BSCs with the following probability transition matrices:

$$p(z|y) = p(y|x) = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix} \tag{7}$$

    Find the channel capacity of the cascaded BSC.

Report and Submission

You can prepare your report with any text editor or in handwriting. Upload a single pdf (not a .zip file) on MOODLE. Late submission is penalized with 10% per day.