COM 5120 Communication Theory

Final Exam Solution

December 27, 2022
15:30 ~ 17:20

Note: There are **6** problems with total 100 points within **2** pages, please write your answer with detail in the answer sheet.

No credit without detail. Closed books. You may use scientific calculator.

1. (12%) Let X be a geometrically distributed random variable, i.e.,

$$P(X = k) = p(1-p)^{k-1}, k = 1, 2, 3, ...$$

(a) Find the entropy of X.
(b) Given that $X > K$, where K is a positive integer, what is the entropy of X?

Solution:
(a)

$$
\begin{aligned}
H(X) &= -\sum_{k=1}^{\infty} p\,(1-p)^{k-1} \log_2\left(p\,(1-p)^{k-1}\right)\\
&= -p\sum_{k=1}^{\infty}(1-p)^{k-1}\log_2 p - p\sum_{k=1}^{\infty}(1-p)^{k-1}\log_2(1-p)^{k-1}\\
&= -p\log_2 p\sum_{k=1}^{\infty}(1-p)^{k-1} - p\log_2(1-p)\sum_{k=1}^{\infty}(k-1)(1-p)^{k-1}\\
&= -p\log_2 p \cdot \frac{1}{p} - p\log_2(1-p)\cdot\frac{1-p}{p^2}\\
&= -\log_2 p - \frac{1-p}{p}\log_2(1-p)
\end{aligned}
$$

[Note] $\displaystyle\sum_{k=1}^{\infty}(k-1)(1-p)^{k-1} = 0 + (1-p)\cdot 1 + (1-p)^2\cdot 2 + \cdots = \sum_{k=0}^{\infty}k\,(1-p)^k$

Let $\displaystyle\sum_{k=0}^{\infty}k\,(1-p)^k = S_n$ and $\begin{cases} S_n = 0 + (1-p)\cdot 1 + (1-p)^2\cdot 2 + \cdots \\ (1-p)S_n = 0 + 0 + (1-p)^2\cdot 1 + \cdots \end{cases}$

$\Rightarrow S_n - (1-p)S_n = 0 + (1-p) + (1-p)^2 + \cdots \quad \Rightarrow [1-(1-p)]\,S_n = \dfrac{1-p}{1-(1-p)}$

$\therefore S_n = \displaystyle\sum_{k=1}^{\infty}(k-1)(1-p)^{k-1} = \frac{1-p}{[1-(1-p)]^2} = \frac{1-p}{p^2}$

and $\displaystyle\sum_{k=1}^{\infty}(1-p)^{k-1} = 1 + (1-p) + (1-p)^2 + \cdots = \frac{1-p}{p} + 1 = \frac{1}{p}$

(b) Clearly $P(X = k|X > K) = 0$ for $k \leq K$. If $k > K$, then

$$P(X = k|X > K) = \frac{P(X = k, X > K)}{P(X > K)} = \frac{p(1 - p)^{k-1}}{P(X > K)}$$

But,

$$
\begin{aligned}
P(X > K) &= \sum_{k=K+1}^{\infty} p (1 - p)^{k-1} \\
&= p \left(\sum_{k=1}^{\infty} (1 - p)^{k-1} - \sum_{k=1}^{K} (1 - p)^{k-1} \right) \\
&= p \left(\frac{1}{1 - (1 - p)} - \frac{1 - (1 - p)^K}{1 - (1 - p)} \right) \\
&= p \cdot \frac{1}{p} - p \cdot \frac{1}{p} - (-p \cdot \frac{(1 - p)^K}{p}) \\
&= (1 - p)^K
\end{aligned}
$$

so that

$$P(X = k|X > K) = \frac{P(X = k, X > K)}{P(X > K)} = \frac{p(1 - p)^{k-1}}{(1 - p)^K}$$

If we let $k = K + t$ with $t = 1, 2, 3, \dots$ then

$$P(X = k|X > K) = \frac{p(1 - p)^K (1 - p)^{t-1}}{(1 - p)^K} = p(1 - p)^{t-1}$$

that is $P(X = k|X > K)$ is the geometrically distributed. Hence, using the results of the first part we obtain

$$
\begin{aligned}
H(X|X > K) &= -\sum_{t=1}^{\infty} p(1 - p)^{t-1} \log_2 \left(p(1 - p)^{t-1} \right) \\
&= -\log_2 p - \frac{1 - p}{p} \log_2 (1 - p)
\end{aligned}
$$

$\blacksquare$

2. (12%) For the channel shown in Figure 1 and given that $P(A) = 1 - p$, $P(B) = P(C) = \frac{p}{2}$, find the **channel capacity** and the **input distribution** that achieves capacity.
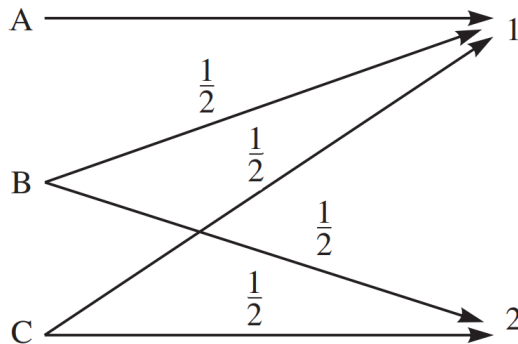

Figure 1: channel model

Solution:

By symmetry of the inputs **B** and **C**, their probabilities have to be equal and we have

$$\begin{cases} P(A) = 1 - p \\ P(B) = P(C) = p/2 \end{cases}$$

Then we have

$$\Rightarrow \begin{cases} P(Y = 1) = 1 - p + p/2 = 1 - p/2 \\ P(Y = 2) = p/2 \end{cases}$$

$$\Rightarrow \begin{cases} H_b(p) = -p \log_2 p - (1 - p) \log_2(1 - p) \\ H_b(\frac{p}{2}) = -\frac{p}{2} \log_2 \frac{p}{2} - (1 - \frac{p}{2}) \log_2(1 - \frac{p}{2}) \end{cases}$$

thus $H(Y) = H_b(p/2)$. We also note that

$$\begin{cases} H(Y|X = A) = 0 \\ H(Y|X = B) = H(Y|X = C) = 1 \end{cases}$$

hence

$$H(Y) = -\frac{p}{2} \log_2 \frac{p}{2} - (1 - \frac{p}{2}) \log_2(1 - \frac{p}{2})$$
$$H(Y|X) = (1 - p) \cdot H(Y|X = A) + p/2 \cdot H(Y|X = B) + p/2 \cdot H(Y|X = C) = p$$

Therefore,

$$C = \max_p I(X; Y)$$
$$= \max_p H(Y) - H(Y|X)$$
$$= \max_p H_b(p/2) - p$$
$$= \max_p \left(-\frac{p}{2} \log_2 \frac{p}{2} - (1 - \frac{p}{2}) \log_2(1 - \frac{p}{2}) \right) - p$$

straightforward differentiation results in

$$\frac{dC}{dp} = \frac{d}{dp}\left(-\frac{p}{2} \log_2 \frac{p}{2} - (1 - \frac{p}{2}) \log_2(1 - \frac{p}{2}) \right) - \frac{d}{dp}p = 0$$

$$\Rightarrow -\left[\frac{1}{2} \log_2(\frac{p}{2}) + \frac{p}{2} \cdot \left(\frac{\frac{1}{2}}{\frac{p}{2} \ln 2} \right) - \frac{1}{2} \log_2(1 - \frac{p}{2}) + (1 - \frac{p}{2}) \cdot \left(\frac{-\frac{1}{2}}{(1 - \frac{p}{2}) \ln 2} \right) \right] - 1 = 0$$

$$\Rightarrow -\frac{1}{2} \log_2(\frac{\frac{p}{2}}{1 - \frac{p}{2}}) - 1 = 0 \quad \left(\because \log_a b = \frac{\log_e b}{\log_e a} \right)$$

$$\Rightarrow -\frac{1}{2} \cdot \frac{\log_e(\frac{\frac{p}{2}}{1 - \frac{p}{2}})}{\log_e 2} - 1 = 0 \quad \left(\because \log_a b = \frac{1}{\log_b a} \right)$$

$$\Rightarrow -\frac{1}{2} \log_2 e \cdot \ln(\frac{\frac{p}{2}}{1 - \frac{p}{2}}) - 1 = 0$$

$$\Rightarrow \ln(\frac{\frac{p}{2}}{1 - \frac{p}{2}}) = \frac{2}{\log_2 e} \quad \Rightarrow \quad \frac{\frac{p}{2}}{1 - \frac{p}{2}} = \exp(\frac{2}{\log_2 e}) = \frac{1}{4}$$

$$\Rightarrow \frac{p}{2} = \frac{1}{4} - \frac{p}{8} \quad \Rightarrow \quad \frac{5p}{8} = \frac{2}{8}$$

resulting in

$$p = \frac{2}{5} = 0.4$$

$$C = H_b(\frac{p}{2}) - p = \left(-\frac{p}{2}\log_2\frac{p}{2} - (1 - \frac{p}{2})\log_2(1 - \frac{p}{2})\right) - p$$

$$= H_b(0.2) - 0.4 = (-0.2\log_2 0.2 - (1 - 0.2)\log_2(1 - 0.2)) - 0.4$$

$$\approx 0.7219 - 0.4 = 0.3219$$

∎

3. (12%) A voice-band telephone channel passes the frequencies in the band from 300 to 3300 Hz. It is desired to design a modem that transmits at a symbol rate of 2400 symbols/s, with the objective of achieving 9600 bits/s. **Select** an appropriate **QAM signal constellation**, **carrier frequency**, and the **roll-off factor** of a pulse with a raised cosine spectrum that utilizes the entire frequency band. **Sketch the spectrum** of the transmitted signal pulse and **indicate the important frequencies**.

Solution:
The bandwidth of the bandpass channel is:

$$W = 3300 - 300 = 3000 \text{ Hz}$$

In order to transmit 9600 bps with a symbol rate

$$R = \frac{1}{T} = 2400 \text{ symbols/sec}$$

the number of information bits per symbol should be

$$k = \frac{9600}{2400} = 4 \text{ bits}$$

Hence, a $2^4 = 16$ QAM signal constellation is needed.

The carrier frequency f_c is:

$$f_c = \frac{300 + 3300}{2} = 1800 \text{ Hz}$$

which is the mid-frequency of the frequency band that the bandpass channel occupies. If a pulse with raised cosine spectrum and roll-off factor β is used for spectral shaping, then for the bandpass signal with bandwidth W:

$$\frac{1}{2T}(1 + \beta) = \frac{W}{2} = 1500$$

$$\Rightarrow \frac{1}{2} \cdot 2400 \cdot (1 + \beta) = 1500$$

$$\Rightarrow 1 + \beta = \frac{1500}{1200} = 1.25$$

$$\therefore \beta = 0.25$$

A sketch of the spectrum of the transmitted signal pulse is shown in Figure 2. (The horizontal coordinate, say 3000, is computed as follows: $f_c + \frac{1}{2T} = 1800 + 2400/2 = 3000$)
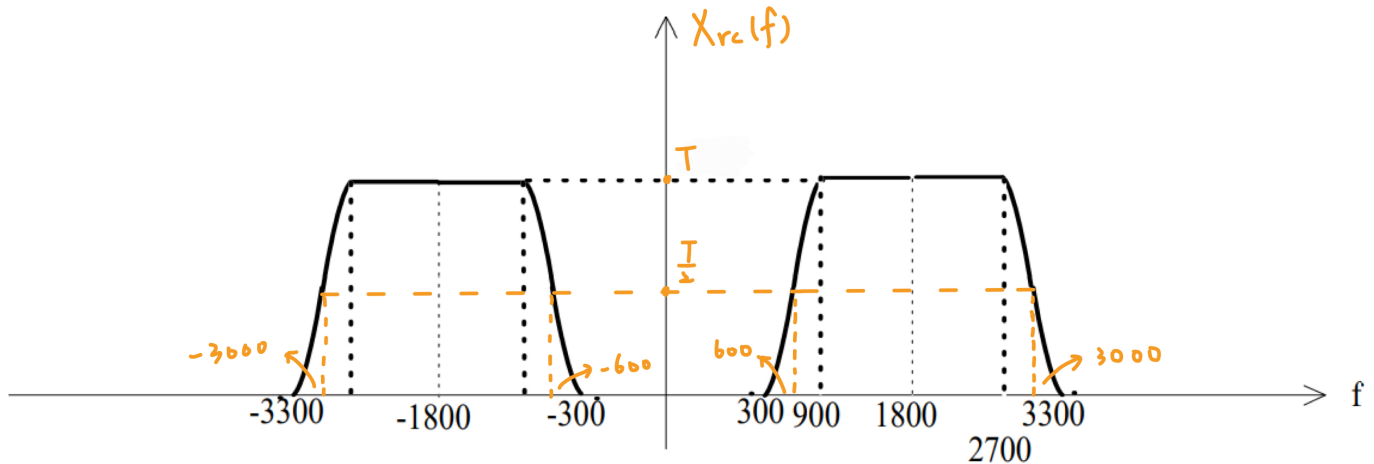

Figure 2: transmitter block diagram

4. (24%) Three equiprobable messages m_1, m_2 and m_3 are to be transmitted over an AWGN channel with noise power spectral density $\frac{1}{2}N_0$. The messages are

$$s_1(t) = \begin{cases} 1 & 0 \le t \le T \\ 0 & \text{otherwise} \end{cases} \qquad s_2(t) = -s_3(t) \begin{cases} 1 & 0 \le t \le \frac{1}{2}T \\ -1 & \frac{1}{2}T < t \le T \\ 0 & \text{otherwise} \end{cases}$$

(a) What is the **dimensionality** of the signal space?
(b) Find an **appropriate basis** for the signal space.
(c) Please **Draw the signal constellation** of the messages m_1, m_2, m_3 and **the optimal decision regions** R_1, R_2, R_3 for this problem.
(d) Which of the three messages is most vulnerable to errors and why? In other words, which of $P(\text{error} \mid m_i \text{ transmitted})$, $i = 1, 2, 3$, is largest?

Solution:
(a) Since $m_2(t) = -m_3(t)$ the dimensionality of the signal space is two.
(b) As a basis of the signal space we consider the functions

$$f_1(t) = \begin{cases} \dfrac{1}{\sqrt{T}} & 0 \le t \le T \\ 0 & \text{otherwise} \end{cases} \qquad f_2(t) = \begin{cases} \dfrac{1}{\sqrt{T}} & 0 \le t \le \frac{1}{2}T \\ -\dfrac{1}{\sqrt{T}} & \frac{1}{2}T < t \le T \\ 0 & \text{otherwise} \end{cases}$$

The vector representation of the signals is

$$\mathbf{m_1} = [\sqrt{T}, \ 0], \quad \mathbf{m_2} = [0, \ \sqrt{T}], \quad \mathbf{m_3} = [0, \ -\sqrt{T}]$$

5

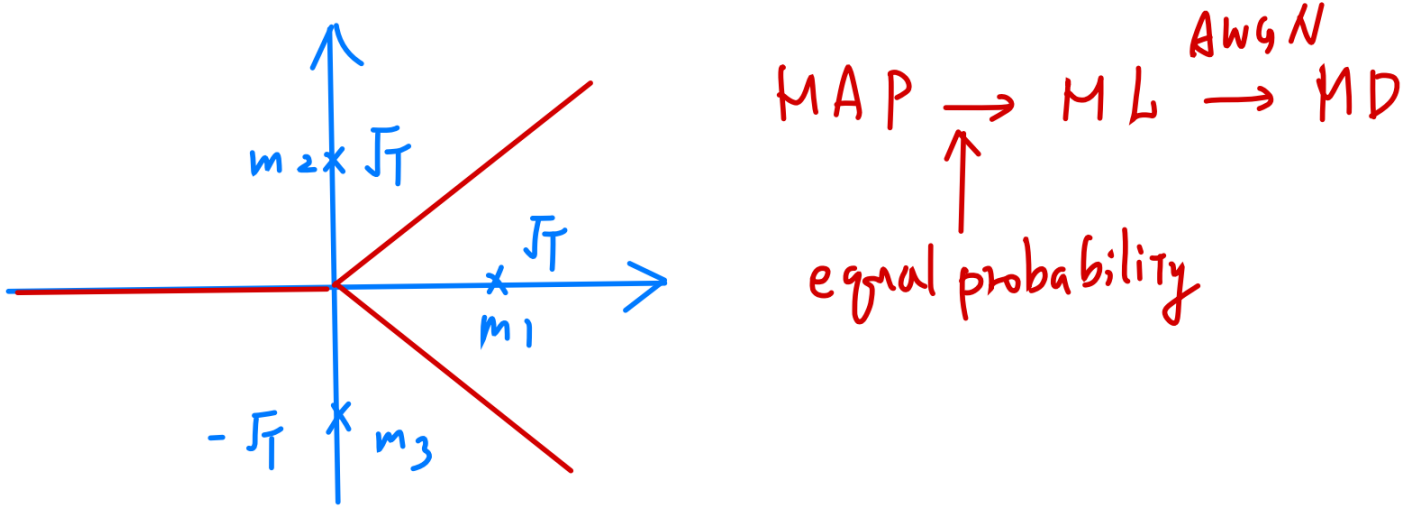**(c)** The signal constellation and the decision regions is depicted in Figure 3.

Figure 3: signal constellation and decision regions

(d) If the signals are equiprobable then:

$$P(e|\mathbf{m_1}) = \mathbf{P}(|\mathbf{r} - \mathbf{m_1}|^2 > |\mathbf{r} - \mathbf{m_2}|^2|\mathbf{m_1}) + \mathbf{P}(|\mathbf{r} - \mathbf{m_1}|^2 > |\mathbf{r} - \mathbf{m_3}|^2|\mathbf{m_1})$$

When $\mathbf{m_1}$ is transmitted then $\mathbf{r} = \left[\sqrt{\mathbf{T}} + \mathbf{n_1}, \mathbf{n_2}\right]$ and therefore, $P(e|\mathbf{m_1})$ is written as:

$$P(e|\mathbf{m_1}) = \mathbf{P}(\mathbf{n_2} - \mathbf{n_1} > \sqrt{\mathbf{T}}) + \mathbf{P}(\mathbf{n_1} + \mathbf{n_2} < \sqrt{\mathbf{T}})$$

Since n_1, n_2 are zero-mean statistically independent Gaussian random variables, each with variance $\frac{N_0}{2}$, the random variables $x = n_1 - n_2$ and $y = n_1 + n_2$ are zero-mean Gaussian with variance N_0. When $\mathbf{m_2}$ is transmitted then $\mathbf{r} = \left[\mathbf{n_1}, \mathbf{n_2} + \sqrt{\mathbf{T}}\right]$ and therefore, $P(e|\mathbf{m_2})$ is written as:

$$P(e|\mathbf{m_2}) = \mathbf{P}(\mathbf{n_1} - \mathbf{n_2} > \sqrt{\mathbf{T}}) + \mathbf{P}(\mathbf{n_2} < \sqrt{\mathbf{T}})$$

Similarly from the symmetry of the problem, we obtain:

$$P(e|\mathbf{m_3}) = \mathbf{P}(e|\mathbf{m_2})$$

Hence:

$$P(e|\mathbf{m_1}) = \frac{1}{\sqrt{2\pi N_0}} \int_{\sqrt{T}}^{\infty} e^{-\frac{x^2}{2N_0}} dx + \frac{1}{\sqrt{2\pi N_0}} \int_{-\infty}^{-\sqrt{T}} e^{-\frac{y^2}{2N_0}} dy$$

$$= Q(\sqrt{\frac{2T}{2N_0}}) + Q(\sqrt{\frac{2T}{2N_0}}) = 2Q(\sqrt{\frac{T}{N_0}})$$

$$P(e|\mathbf{m_2}) = Q(\sqrt{\frac{2T}{2N_0}}) + Q(\sqrt{\frac{4T}{2N_0}}) = Q(\sqrt{\frac{T}{N_0}}) + Q(\sqrt{\frac{2T}{N_0}}) = P(e|\mathbf{m_3})$$

Since $Q(\cdot)$ is momononically decreasing, we obtain:

$$Q(\sqrt{\frac{2T}{N_0}}) < Q(\sqrt{\frac{T}{N_0}})$$

and therefore, the probability of error $P(e|\mathbf{m_1})$ is larger than $P(e|\mathbf{m_2})$ and $P(e|\mathbf{m_3})$ Hence, the message $\mathbf{m_1}$ is more vulnerable to errors. The reason for that is that it has both threshold lines close to it, while the other two signals have one of the their threshold lines further away.

∵ The probability of error $P(e|\mathbf{m_1})$ is larger than $P(e|\mathbf{m_2})$ and $P(e|\mathbf{m_3})$.

∴ The message $\mathbf{m_1}$ is the most vulnerable to errors.

■

5. (24%) Two discrete memoryless information sources X and Y each have an alphabet with six symbols, $\mathcal{X} = \mathcal{Y} = \{1, 2, 3, 4, 5, 6\}$. The probabilities of the letters for X are $\{\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{32}, \frac{1}{32}\}$. The source Y has a uniform distribution.

(a) Find the entropy of both X and Y.

(b) Design Huffman codes for each source. Which Huffman code is more efficient? (**Hint**: Efficiency of a Huffman code is defined as the ratio of the source entropy $H(X)$ to the average codeword length $\bar{R}$.)

(c) If Huffman codes were designed for the second extension of these sources (i.e., two letters at a time), **for which source (X or Y) would you expect a performance improvement compared to the single-letter Huffman code and why?**

Solution:

(a)

The entropy of X:
$$H(X) = -\sum_{i=1}^{5} (\frac{1}{2})^i \log_2 (\frac{1}{2})^i - \frac{1}{32} \log_2 \frac{1}{32}$$
$$= -\frac{1}{2} \log_2 \frac{1}{2} - \frac{1}{4} \log_2 \frac{1}{4} - \frac{1}{8} \log_2 \frac{1}{8} - \frac{1}{16} \log_2 \frac{1}{16} - \frac{2}{32} \log_2 \frac{1}{32}$$
$$= \frac{1}{2} \log_2 2 + \frac{1}{4} \log_2 4 + \frac{1}{8} \log_2 8 + \frac{1}{16} \log_2 16 + \frac{1}{16} \log_2 32$$
$$= \frac{1}{2} \cdot 1 + \frac{1}{4} \cdot 2 + \frac{1}{8} \cdot 3 + \frac{1}{16} \cdot 4 + \frac{1}{16} \cdot 5$$
$$= \frac{8}{16} + \frac{8}{16} + \frac{6}{16} + \frac{4}{16} + \frac{5}{16} = \frac{31}{16} = 1.9375$$

The entropy of Y:
$$H(Y) = 6 \cdot (-\frac{1}{6} \log_2 \frac{1}{6}) = \log_2 6 = 2.5850$$
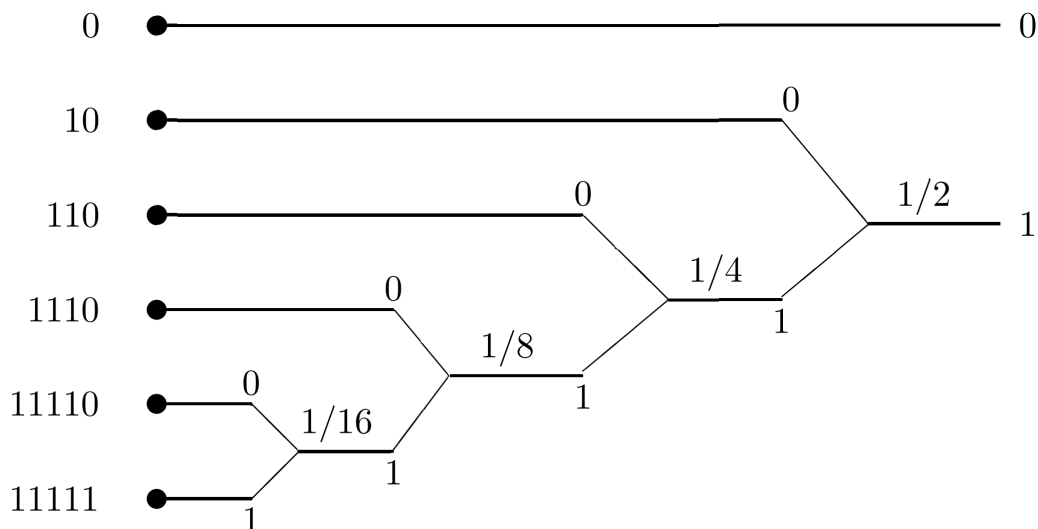
(b) For X the Huffman tree is

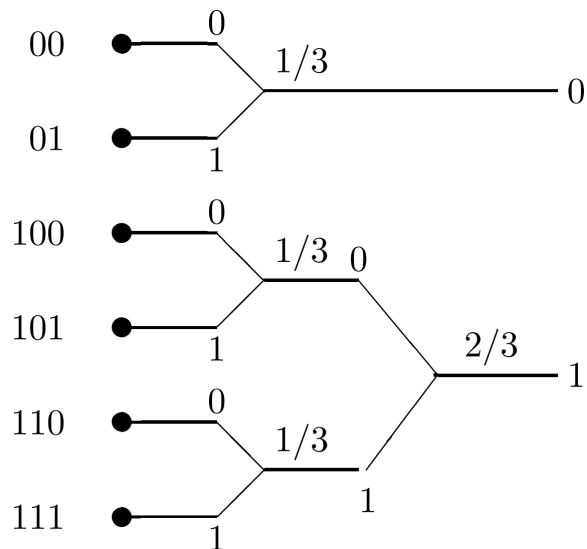Figure 4: the Huffman tree of X

and for Y the Huffman tree is

Figure 5: the Huffman tree of Y

For X the average codeword length is

$$\bar{R}_X = 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{8} + 4 \cdot \frac{1}{16} + 5 \cdot \frac{2}{32}$$
$$= \frac{8}{16} + \frac{8}{16} + \frac{6}{16} + \frac{4}{16} + \frac{5}{16} = \frac{31}{16} = 1.9375$$

hence the efficiency of X is

$$\eta_X = \frac{H(X)}{\bar{R}_X} = \frac{1.9375}{1.9375} = 1 = 100\%$$

For Y the average codeword length is

$$\bar{R}_Y = 2 \cdot \frac{1}{3} + 3 \cdot \frac{2}{3} = \frac{8}{3} = 2.666$$

hence the efficiency of Y is

$$\eta_Y = \frac{H(Y)}{\bar{R}_Y} = \frac{2.5850}{2.666} \approx 0.9696 = 96.96\%$$

Therefore X is more efficient.

(c) Source X has achieved 100% efficiency. Therefore we expect an improvement for source Y.

∎

6. (16%) The transmission of a signal pulse with a raised cosine spectrum through a channel results in the following (noise-free) sampled output from the demodulator:

$$x_k = \begin{cases} -0.5 & m = -2 \\ 0.1 & m = -1 \\ 1 & m = 0 \\ -0.2 & m = 1 \\ 0.05 & m = 2 \\ 0 & \text{otherwise} \end{cases}$$

(a) Design a three-tap zero-forcing linear equalizer so that the output is

$$q_m = \begin{cases} 1 & m = 0 \\ 0 & m = \pm 1 \end{cases}$$

(b) Determine q_m for $m = \pm 2, \pm 3$ by convolving the impulse response of the equalizer with the channel response.

Solution:
(a) The output of the zero-forcing linear equalizer is:

$$q_m = \sum_{n=-1}^{1} c_n x_{m-n}$$

$$\Rightarrow \begin{cases} x_0 c_{-1} + x_{-1} c_0 + x_{-2} c_1 = q_{-1} \\ x_1 c_{-1} + x_0 c_0 + x_{-1} c_1 = q_0 \\ x_2 c_{-1} + x_1 c_0 + x_0 c_1 = q_1 \end{cases} \Rightarrow \begin{pmatrix} x_0 & x_{-1} & x_{-2} \\ x_1 & x_0 & x_{-1} \\ x_2 & x_1 & x_0 \end{pmatrix} \begin{pmatrix} c_{-1} \\ c_0 \\ c_1 \end{pmatrix} = \begin{pmatrix} q_{-1} \\ q_0 \\ q_1 \end{pmatrix}$$

With $q_0 = 1$ and $q_1 = 0$ for $m \neq 0$, we obtain the system:

$$\begin{pmatrix} 1 & 0.1 & -0.5 \\ -0.2 & 1 & 0.1 \\ 0.05 & -0.2 & 1 \end{pmatrix} \begin{pmatrix} c_{-1} \\ c_0 \\ c_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

Solving the previous system in terms of the equalizer's coefficients, we obtain:

$$\begin{pmatrix} c_{-1} \\ c_0 \\ c_1 \end{pmatrix} = \begin{pmatrix} 1 & 0.1 & -0.5 \\ -0.2 & 1 & 0.1 \\ 0.05 & -0.2 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$$\approx \begin{pmatrix} 0.9756 & 0 & 0.4878 \\ 0.1961 & 0.9804 & 0 \\ 0.0096 & 0.1961 & 0.9756 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0.9804 \\ 0.1961 \end{pmatrix}$$

(b) The output values of the equalizer q_m for $m = \pm 2, \pm 3$ are given by

$$q_{-3} = \sum_{n=-1}^{1} c_n x_{-3-n} = c_{-1} x_{-2} + c_0 x_{-3} + c_1 x_{-4} = 0 + 0 + 0 = 0$$

$$q_{-2} = \sum_{n=-1}^{1} c_n x_{-2-n} = c_{-1} x_{-1} + c_0 x_{-2} + c_1 x_{-3} = 0 + 0.9804 \cdot (-0.5) + 0 = -0.490196$$

$$q_2 = \sum_{n=-1}^{1} c_n x_{2-n} = c_{-1} x_3 + c_0 x_2 + c_1 x_1 = 0 + 0.9804 \cdot 0.05 + 0.1961 \cdot (-0.2) = 0.0098$$

$$q_3 = \sum_{n=-1}^{1} c_n x_{3-n} = c_{-1} x_4 + c_0 x_3 + c_1 x_2 = 0 + 0 + 0.1961 \cdot 0.05 = 0.009805$$

Therefore

$$q_m = \begin{cases} 0 & m = -3 \\ -0.490196 & m = -2 \\ 0.0098 & m = 2 \\ 0.009805 & m = 3 \end{cases}$$

$\blacksquare$