ECS452 2016/2 Part I.4 Dr.Prapun

3 An Introduction to Digital Communication Systems Over Discrete Memoryless Channel (DMC)

In this section, we keep our analysis of the communication system simple by considering purely digital systems. To do this, we assume all non-source-coding parts of the system, including the physical channel, can be combined into an "equivalent channel" which we shall simply refer to in this section as the "channel".

3.1 Discrete Memoryless Channel (DMC) Models

Example 3.1. The **binary symmetric channel (BSC)**, which is the simplest model of a channel with errors, is shown in Figure 4.

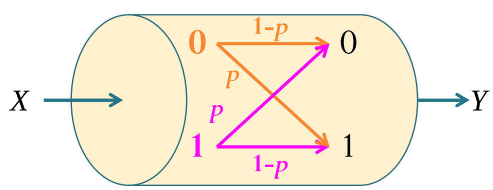


Figure 4: Binary symmetric channel and its channel diagram

- "Binary" means that the there are two possible values for the input and also two possible values for the output. We normally use the symbols 0 and 1 to represent these two values.
- Passing through this channel, the input symbols are complemented with crossover probability p.

switchover probability
bit-flip probability

- It is simple, yet it captures most of the complexity of the general problem.

Example 3.2. Consider a BSC whose samples of input and output are provided below

Time index: 1 2 3 4 5 … 20

Channel input x: 1 0 1 1 1 1 1 1 1 1 1 1 0 1 0 1 1 1 1 1

Channel output y: 1 1 1 1 1 1 1 1 1 0 1 1 0 1 0 1 1 1 1 1

Estimate the following (unconditional and conditional) probabilities by their relative frequencies.

$\underline{p} = [p(0) \;\; p(1)] = [\frac{3}{20} \;\; \frac{17}{20}]$ $\quad \mathcal{X} = \{0, 1\}$

$p(0) = P[X = 0] \approx \frac{3}{20}$ $\qquad p(1) = P[X = 1] \approx 1 - P[X=0] = \frac{17}{20}$

$q(0) = P[Y = 0] \approx \frac{3}{20}$ $\qquad q(1) = P[Y = 1] \approx 1 - \frac{3}{20} = \frac{17}{20}$ $\qquad \underline{q} = [q(0) \;\; q(1)] = [\frac{3}{20} \;\; \frac{17}{20}]$

$Q(0|0) = P[Y = 0|X = 0] \approx \frac{2}{3}$ $\qquad Q(1|0) = P[Y = 1|X = 0] \approx 1 - \frac{2}{3} = \frac{1}{3}$

$Q(0|1) = P[Y = 0|X = 1] \approx \frac{1}{17}$ $\qquad Q(1|1) = P[Y = 1|X = 1] \approx 1 - \frac{1}{17} = \frac{16}{17}$

$$Q = \begin{array}{c|cc} x \backslash y & 0 & 1 \\ \hline 0 & 2/3 & 1/3 \\ 1 & 1/17 & 16/17 \end{array}$$

Definition 3.3. Our model for **discrete memoryless channel (DMC)** is shown in Figure 5.

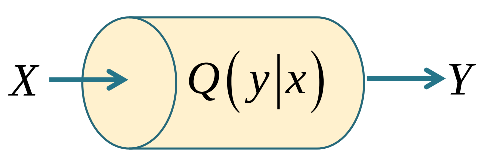


Figure 5: Discrete memoryless channel

- The channel input is denoted by a random variable X.
  - The pmf $p_X(x)$ is usually denoted by simply $p(x)$ and usually expressed in the form of a row vector $\underline{\mathbf{p}}$ or $\underline{\pi}$.
  - The support S_X is often denoted by $\mathcal{X}$.

channel input alphabet

- Similarly, the channel output is denoted by a random variable Y.
  - The pmf $p_Y(y)$ is usually denoted by simply $q(y)$ and usually expressed in the form of a row vector $\underline{\mathbf{q}}$.
  - The support S_Y is often denoted by $\mathcal{Y}$.

- The channel corrupts its input X in such a way that when the input is $X = x$, its output Y is randomly selected from the conditional pmf $p_{Y|X}(y|x)$. $= P[Y=y \mid X=x]$

$$Q(y|x) \equiv P[\underbrace{Y=y}_{A} \mid \underbrace{X=x}_{B}] = P(A|B) = \frac{P(A\cap B)}{P(B)} = \frac{P[Y=y, X=x]}{P[X=x]}$$

channel transition probability

  - This conditional pmf $p_{Y|X}(y|x)$ is denoted by $Q(y|x)$ and usually expressed in the form of a probability transition matrix $\mathbf{Q}$:

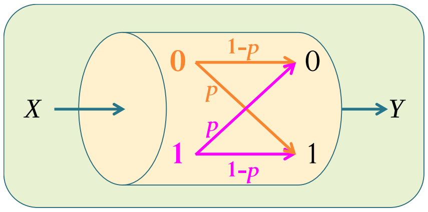
$$\begin{array}{cc} & y \\ x & \begin{bmatrix} \ddots & \vdots & \iddots \\ \cdots & P[Y=y|X=x] & \cdots \\ \iddots & \vdots & \ddots \end{bmatrix} \end{array}$$

  - The channel is called memoryless[9] because its channel output at a given time is a function of the channel input at that time and is not a function of previous channel inputs.
  - Here, the transition probabilities are assumed constant. However, in many commonly encountered situations, the transition probabilities are time varying. An example is the wireless mobile channel in which the transmitter-receiver distance is changing with time.

[9] Mathematically, the condition that the channel is memoryless may be expressed as [12, Eq. 6.5-1 p. 355]

$$p_{X_1^n|Y_1^n}(x_1^n|y_1^n) = \prod_{k=1}^{n} Q(y_k|x_k).$$

Example 3.4. For a binary symmetric channel (BSC) defined in 3.1, we now have three equivalent ways to specify the relevant probabilities:



$$P[Y=0|X=0]=Q(0|0)=1-p$$
$$P[Y=1|X=0]=Q(1|0)=p$$
$$P[Y=0|X=1]=Q(0|1)=p$$
$$P[Y=1|X=1]=Q(1|1)=1-p$$

$$\mathbf{Q}=\begin{array}{c|cc} x \backslash y & 0 & 1 \\ \hline 0 & 1-p & p \\ 1 & p & 1-p \end{array}$$

Example 3.5. Suppose, for a DMC, we have $\mathcal{X} = \{x_1, x_2\}$ and $\mathcal{Y} = \{y_1, y_2, y_3\}$. Then, its probability transition matrix $\mathbf{Q}$ is of the form

$$\mathbf{Q} = \begin{bmatrix} Q(y_1|x_1) & Q(y_2|x_1) & Q(y_3|x_1) \\ Q(y_1|x_2) & Q(y_2|x_2) & Q(y_3|x_2) \end{bmatrix}.$$

You may wonder how this $\mathbf{Q}$ happens in real life. Let's suppose that the input to the channel is binary; hence, $\mathcal{X} = \{0, 1\}$ as in the BSC. However, in this case, after passing through the channel, some bits can be lost[10] (rather than corrupted). In such case, we have three possible outputs of the channel: 0, 1, e where the "e" represents the case in which the bit is erased by the channel.

Example 3.6. Consider a DMC whose samples of input and output are provided below

x: 1 1 1 1 1 1 1 1 1 1 0 0 1 1 1 0 1 1 0 1
y: 1 3 2 2 1 2 1 2 2 3 1 1 1 3 1 3 2 3 1 2

Estimate its input probability vector $\underline{\mathbf{p}}$, output probability vector $\underline{\mathbf{q}}$, and $\mathbf{Q}$ matrix.

$\underbrace{P[X=1, Y=2]}_{A \cap B} = \frac{7}{20}$

$= P[\underbrace{Y=2}_{B} \mid \underbrace{X=1}_{A}]\, P[\underbrace{X=1}_{A}]$

$= \frac{7}{16} \times \frac{16}{20} = \frac{7}{20}$

$\underline{p} = [p(0) \; p(1)] \approx [\frac{4}{20} \;\; \overset{20-4}{\frac{16}{20}}] = [0.2 \;\; 0.8]$

$\underline{q} = [q(1) \; q(2) \; q(3)] \approx [\frac{8}{20} \;\; \frac{7}{20} \;\; \frac{5}{20}] = [0.4 \;\; 0.35 \;\; 0.25]$

$$\mathbf{Q} = \begin{array}{c|ccc} x \backslash y & 1 & 2 & 3 \\ \hline 0 & \frac{3}{4} & \frac{0}{4} & \frac{1}{4} \\ 1 & \frac{5}{16} & \frac{7}{16} & \frac{4}{16} \end{array}$$

$Q(y|x) = P[Y=y \mid X=x]$

[10]The receiver knows which bits have been erased.

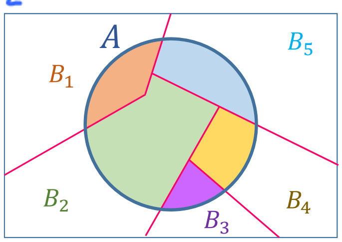
3.7. Knowing the input probabilities $\underline{\mathbf{p}}$ and the channel probability transition matrix $\mathbf{Q}$, we can calculate the output probabilities $\underline{\mathbf{q}}$ from

$$\underline{\mathbf{q}} = \underline{\mathbf{p}}\mathbf{Q}.$$

To see this, recall the **total probability theorem**: If a (finite or infinitely) countable collection of events $\{B_1, B_2, \ldots\}$ is a partition of Ω, then

$$P(A) = \sum_i P(A \cap B_i) = \sum_i P(A|B_i)P(B_i). \tag{5}$$

Ω



$$P(A) = P(A \cap B_1) + P(A \cap B_2) + P(A \cap B_3) + P(A \cap B_4) + P(A \cap B_5)$$

For us, event A is the event $[Y = y]$. Applying this theorem to our variables, we get

$A = [Y=y]$

$B_x = [X=x]$

$$\begin{aligned} q(y) = P[Y = y] &= \sum_x P[X = x, Y = y] \\ &= \sum_x P[Y = y | X = x]\, P[X = x] = \sum_x Q(y|x)p(x). \end{aligned}$$

This is exactly the same as the matrix multiplication calculation performed to find each element of $\underline{\mathbf{q}}$.

Example 3.8. For a binary symmetric channel (BSC) defined in 3.1,

$$\begin{aligned} q(0) &= P[Y = 0] = P[Y = 0, X = 0] + P[Y = 0, X = 1] \\ &= P[Y = 0|\, X = 0]\, P[X = 0] + P[Y = 0|\, X = 1]\, P[X = 1] \\ &= Q(0|0)\, p(0) + Q(0|1)\, p(1) = (1 - p) \times p_0 + p \times p_1 \\ q(1) &= P[Y = 1] = P[Y = 1, X = 0] + P[Y = 1, X = 1] \\ &= P[Y = 1|\, X = 0]\, P[X = 0] + P[Y = 1|\, X = 1]\, P[X = 1] \\ &= Q(1|0)\, p(0) + Q(1|1)\, p(1) = p \times p_0 + (1 - p) \times p_1 \end{aligned}$$

$= [p(0) \;\; p(1)] \begin{bmatrix} Q(0|0) \\ Q(0|1) \end{bmatrix}$ first column of Q matrix

$= [p(0) \;\; p(1)] \begin{bmatrix} Q(1|0) \\ Q(1|1) \end{bmatrix}$ second column of Q matrix

$\underline{q} = [q(0) \;\; q(1)] = \left[\underline{p} \begin{bmatrix} \; \end{bmatrix} \;\; \underline{p} \begin{bmatrix} \; \end{bmatrix} \right] = \underline{p} \overbrace{\left[\begin{bmatrix} \; \end{bmatrix} \begin{bmatrix} \; \end{bmatrix} \right]}^{Q}$

block matrix multiplication (factoring)

$$\underline{q} = \begin{bmatrix} \quad q(y) \quad \end{bmatrix} = \underline{p}\, Q = \underbrace{\begin{bmatrix} \qquad\qquad \end{bmatrix}}_{\underline{p}} \underset{Q}{\begin{array}{c} y \\ \begin{bmatrix} \quad & \boxed{} & \quad \end{bmatrix} \end{array}}$$

3.9. Recall, from ECS315, that there is another matrix called the **joint probability matrix P**. This is the matrix whose elements give the joint probabilities $P_{X,Y}(x,y) = P[X = x, Y = y]$:

$$\mathbf{P} = \begin{array}{c} \\ x \\ \\ \end{array} \begin{array}{c} y \\ \begin{bmatrix} \ddots & \vdots & \iddots \\ \cdots & P[X = x, Y = y] & \cdots \\ \iddots & \vdots & \ddots \end{bmatrix} \end{array}.$$

Recall also that we can get $p(x)$ by adding the elements of $\mathbf{P}$ in the row corresponding to x. Similarly, we can get $q(y)$ by adding the elements of $\mathbf{P}$ in the column corresponding to y.

By definition, the relationship between the conditional probability $Q(y|x)$ and the joint probability $P_{X,Y}(x,y)$ is

$$Q(y|x) = \frac{P_{X,Y}(x,y)}{p(x)}.$$

$P(A|B) = \frac{P(A \cap B)}{P(B)}$ $\quad A = [Y = y]$, $B = [X = x]$

$P[Y = y | X = x] = \frac{P[X = x, Y = y]}{P[X = x]}$

Equivalently,

$$P_{X,Y}(x,y) = p(x)Q(y|x).$$

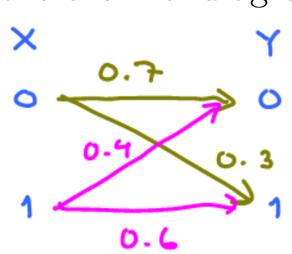
Therefore, to get the matrix $\mathbf{P}$ from matrix $\mathbf{Q}$, we need to multiply each row of $\mathbf{Q}$ by the corresponding $p(x)$. This could be done easily in MATLAB by first constructing a diagonal matrix from the elements in $\underline{\mathbf{p}}$ and then multiply this to the matrix $\mathbf{Q}$:

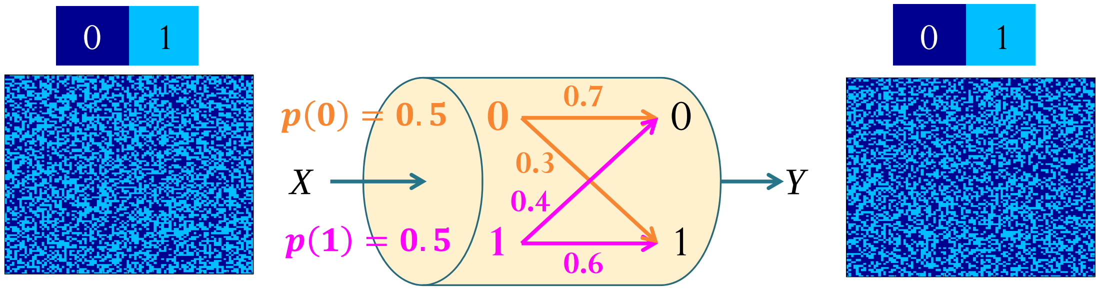
$$\mathbf{P} = (\operatorname{diag}(\underline{\mathbf{p}}))\,\mathbf{Q}.$$

Example 3.10. Binary Asymmetric Channel (BAC): Consider a binary input-output channel whose matrix of transition probabilities is

$$\mathbf{Q} = \begin{array}{c} x\backslash y \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix} \end{array}.$$

(a) Draw the channel diagram.





$\underline{p} = \left[\tfrac{1}{2} \quad \tfrac{1}{2}\right]$

(b) If the two inputs are equally likely, find the corresponding output probability vector $\underline{\mathbf{q}}$ and the joint probability matrix $\mathbf{P}$ for this channel.

$$\underline{q} = \underline{p}\,Q = \left[\tfrac{1}{2} \quad \tfrac{1}{2}\right]\begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix} = [0.55 \quad 0.45]$$

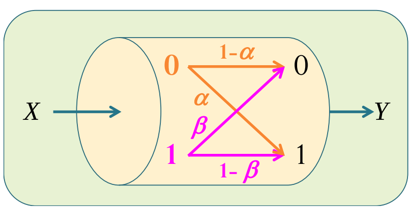
$$Q = \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix} \xrightarrow[\times p(1)]{\times p(0)} \begin{array}{c|cc} x \backslash y & 0 & 1 \\ \hline 0 & 0.35 & 0.15 \\ 1 & 0.2 & 0.3 \end{array} = P$$

$$\downarrow \Sigma \qquad \downarrow \Sigma$$

$$[0.55 \quad 0.45] = \underline{q}$$

[17, Ex. 11.3]

Example 3.11. Similar to Example 3.4 where we have three equivalent ways to specify BSC. We also have three different ways to describe BAC:



$P[Y=0|X=0] = Q(0|0) = 1-\alpha$

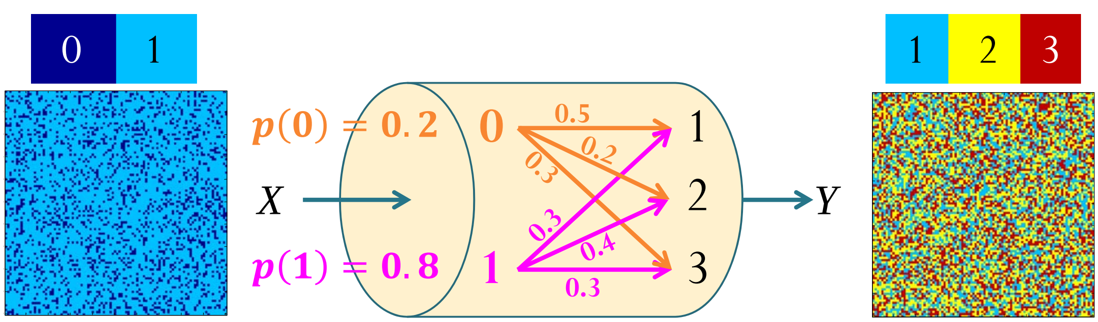
$P[Y=1|X=0] = Q(1|0) = \alpha$

$P[Y=0|X=1] = Q(0|1) = \beta$

$P[Y=1|X=1] = Q(1|1) = 1-\beta$

$$\mathbf{Q} = \begin{array}{c} x \backslash y \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & \quad 1 \end{array} \\ \begin{bmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{bmatrix} \end{array}$$

Example 3.12. Find the output probability vector $\underline{\mathbf{q}}$ and the joint probability matrix $\mathbf{P}$ for the following DMC:



$$Q = \begin{array}{c|ccc} x \backslash y & 1 & 2 & 3 \\ \hline 0 & 0.5 & 0.2 & 0.3 \\ 1 & 0.3 & 0.4 & 0.3 \end{array} \xrightarrow[\times 0.8]{\times 0.2} \begin{array}{c|ccc} x \backslash y & 1 & 2 & 3 \\ \hline 0 & 0.10 & 0.04 & 0.06 \\ 1 & 0.24 & 0.32 & 0.24 \end{array} = P$$

$$\downarrow \Sigma \quad \downarrow \Sigma \quad \downarrow \Sigma$$

$$\underline{q} = [0.34 \quad 0.36 \quad 0.30]$$

Summary:

DMC: X: Channel Input
Y: Channel Output

Notation used in digital commu. (and info. theory) class is different from probability class

Channel input alphabet $\rightarrow$ $S_X \to \mathcal{X}$

$S_Y \to \mathcal{Y}$ $\leftarrow$ Channel output alphabet

$P[X=x] \to p_X(x) \to p(x) \to \underline{p}$

$P[Y=y] \to p_Y(y) \to q(y) \to \underline{q}$

$P[Y=y|X=x] \to p_{Y|X}(y|x) \to Q(y|x) \to$ **Q** matrix

$P[X=x \text{ "AND" } Y=y] \to p_{X,Y}(x,y) \to p(x,y) \to$ **P** matrix

Put together all values at correrspending positions in the matrix

When the alphabets are lists of integers,

we usually write $p(x)$ and $q(y)$ as p_x and q_y respectively.

Ex. For BSC, $\mathcal{X} = \{0,1\}$
$p_0 = p(0) = P[X=0]$
$p_1 = p(1) = P[X=1]$

Alternatively, when the members of the alphabet(s) are explicitly indexed,

we often define $p_i \equiv p(x_i)$ and $q_j \equiv q(y_j)$

DMC is defined by its Q matrix:

When $\mathcal{X} = \{x_1, x_2, x_3, \dots, x_m\}$ and $\mathcal{Y} = \{y_1, y_2, y_3, \dots, y_n\}$,

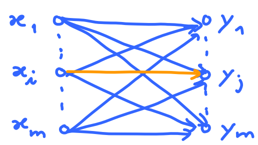
$$Q = \begin{array}{c|c} x \backslash y & y_1 \cdots y_j \cdots y_n \\ \hline x_1 & \\ \vdots & \\ x_i & Q(y_j|x_i) \\ \vdots & \\ x_m & \end{array} \xrightarrow[\times p(x_m)]{\times p(x_1) \;\cdots\; \times p(x_i)} \begin{array}{c|c} x \backslash y & y_1 \cdots y_j \cdots y_n \\ \hline x_1 & \\ \vdots & \\ x_i & p_{X,Y}(x_i, y_j) \\ \vdots & \\ x_m & \end{array} = P$$

The P matrix can be derived from the Q matrix by scaling each row of the Q matrix by the corresponding p(x)

$$\left[\ \Sigma\downarrow \quad \Sigma\downarrow \quad \Sigma\downarrow\ \right] = \underline{q}$$

Alternatively, $\underline{q} = \underline{p}\, Q$

Equivalently, we may define a DMC via its channel diagram (of transition probabilities):



Each arrow should be labeled with $Q(y_j|x_i)$.

3.2 Decoder and Symbol Error Probability

3.13. Knowing the characteristics of the channel and its input, on the receiver side, we can use this information to build a "good" receiver.

We now consider a part of the receiver called the **(channel) decoder**. Its job is to guess the value of the channel input[11] X from the value of the received channel output Y. We denote this guessed value by $\hat{X}$.

3.14. A "good" receiver is the one that (often) guesses correctly. So, our goal here is to

- maximize the probability of correct guessing
- minimize " " " wrong guessing

Quantitatively, to measure the performance of a decoder, we define a quantity called the (symbol) error probability.

Definition 3.15. The **(symbol) error probability**, denoted by $P(\mathcal{E})$, can be calculated from

$$P(\mathcal{E}) = P\left[\hat{X} \neq X\right].$$

3.16. A "good" detector should guess based on all the information it has obtained. Here, the only information it can observe is the value of Y. So, a detector is a function of Y, say, $g(Y)$. Therefore, $\hat{X} = g(Y)$.

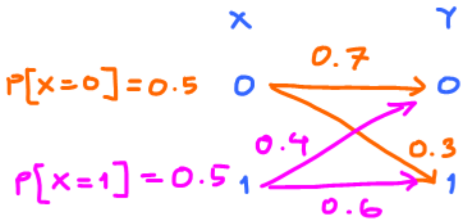
We will write $\hat{X}$ as $\hat{x}(Y)$ to emphasize that the decoded value $\hat{X}$ depends on the observed value Y and that the detector is a deterministic function of the channel output Y; the randomness in the decoded value $\hat{X}$ comes from the randomness in Y.

Definition 3.17. A "naive" decoder is a decoder that simply sets $\hat{X} = Y$.

"pass-through"
"do-nothing"

[11]To simplify the analysis, we still haven't considered the channel encoder. (It may be there but is included in the equivalent channel or it may not be in the system at all.)

Example 3.18. Consider the BAC channel and input probabilities specified in Example 3.10. Find $P(\mathcal{E})$ when $\hat{X} = Y$.



$$P(\mathcal{E}) \equiv P[\hat{X} \neq X] = P[Y \neq X]$$
$$= 0.2 + 0.15 = 0.35$$

$$Q = \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix} \xrightarrow[\times p(1)]{\times p(0)} \begin{array}{c|cc} x \backslash y & 0 & 1 \\ \hline 0 & \underline{0.35} & \underline{0.15} \\ 1 & \underline{0.2} & 0.3 \end{array} = P$$

3.19. For general DMC, the error probability of the naive decoder is

$$\begin{aligned} P(\mathcal{E}) &= P\left[\hat{X} \neq X\right] = P[Y \neq X] = 1 - P[Y = X] \\ &= 1 - \sum_x P[Y = x, X = x] = 1 - \sum_x P[Y = x \,| X = x] P[X = x] \\ &= 1 - \sum_x Q(x \,| x)\, p(x) \end{aligned}$$

Example 3.20. With the derived formula, let's revisit Example 3.18.

$$P(\mathcal{E}) = 1 - (Q(0\,|0)\,p(0) + Q(1\,|1)\,p(1)) = 1 - \left(0.7 \times \frac{1}{2} + 0.6 \times \frac{1}{2}\right) = 0.35.$$

Example 3.21. Find the error probability $P(\mathcal{E})$ when a naive decoder is used with a DMC channel in which $\mathcal{X} = \{0, 1\}$, $\mathcal{Y} = \{1, 2, 3\}$, $\mathbf{Q} = \begin{bmatrix} 0.5 & 0.2 & 0.3 \\ 0.3 & 0.4 & 0.3 \end{bmatrix}$ and $\underline{\mathbf{p}} = [0.2, 0.8]$.

From Example 3.12,

$$Q = \begin{array}{c|ccc} x \backslash y & 1 & 2 & 3 \\ \hline 0 & 0.5 & 0.2 & 0.3 \\ 1 & 0.3 & 0.4 & 0.3 \end{array} \xrightarrow[\times 0.8]{\times 0.2} \begin{array}{c|ccc} x \backslash y & 1 & 2 & 3 \\ \hline 0 & \underline{0.10} & \underline{0.04} & \underline{0.06} \\ 1 & 0.24 & \underline{0.32} & \underline{0.24} \end{array} = P$$

$$P(\mathcal{E}) \equiv P[\hat{X} \neq X]$$
$$= P[Y \neq X]$$
$$= 1 - 0.24 = 0.76$$

$$P(C) = P[Y = X]$$
$$= 0.24$$

Example 3.22. DIY Decoder: Consider a different decoder specified in the decoding table below. Find the error probability $P(\mathcal{E})$ when such decoder is used in Example 3.21.

Decoding table →

y	$\hat{x}(y)$
1	0
2	1
3	0

$$\begin{array}{c|ccc} \hat{x} & 0 & 1 & 0 \\ x \backslash y & 1 & 2 & 3 \\ \hline \begin{array}{c} 0 \\ 1 \end{array} & \multicolumn{3}{c}{\left[\begin{array}{ccc} 0.10 & 0.04 & 0.06 \\ 0.24 & 0.32 & 0.24 \end{array}\right]} \end{array} = P$$

$P(\mathcal{C}) = P[\hat{X} = X] = 0.1 + 0.32 + 0.06 = 0.48$

$P(\mathcal{E}) = P[\hat{X} \neq X] = 1 - 0.48 = 0.52$

Example 3.23. Repeat Example 3.22 but use the following decoder

y	$\hat{x}(y)$
1	1
2	1
3	0

$$\begin{array}{c|ccc} \hat{x} & 1 & 1 & 0 \\ x \backslash y & 1 & 2 & 3 \\ \hline \begin{array}{c} 0 \\ 1 \end{array} & \multicolumn{3}{c}{\left[\begin{array}{ccc} 0.10 & 0.04 & 0.06 \\ 0.24 & 0.32 & 0.24 \end{array}\right]} \end{array} = P$$

$P(\mathcal{C}) = 0.24 + 0.32 + 0.06 = 0.62$

$P(\mathcal{E}) = 1 - P(\mathcal{C}) = 1 - 0.62 = 0.38$

Observation: For each column of the **P** matrix, we circle the probability corresponding to the row of x that has the same value as $\hat{x}(y)$.

3.24. A recipe for finding $P(\mathcal{E})$ of any (DIY) decoder:

(a) Find the **P** matrix by scaling each row of the **Q** matrix by its corresponding $p(x)$.

(b) Write $\hat{x}(y)$ values on top of the y values for the **P** matrix.

(c) For each y column in the **P** matrix, circle the element whose corresponding x value is the same as $\hat{x}(y)$.

(d) $P(\mathcal{C})$ = the sum of the circled probabilities. $P(\mathcal{E}) = 1 - P(\mathcal{C})$.