Assignment 2 Solution

Question 1 Answer

1) Calculating $\sum_{j\in\{a,b\}} \mathbf{P}(X_2 = \mathrm{i}|\ X_1 = \mathrm{j}) = \mathbf{1},\ \textbf{for all}\ i \in \{\mathbf{a}, \mathbf{b}, \boldsymbol{\epsilon}\}$:

$$\mathrm{P}(X_2 = \mathrm{a}|\ X_1 = \mathrm{a}) + \mathrm{P}(X_2 = \mathrm{a}|\ X_1 = \mathrm{b}) = 0.8 + 0.15 = 0.95$$
$$\mathrm{P}(X_2 = \mathrm{b}|\ X_1 = \mathrm{a}) + \mathrm{P}(X_2 = \mathrm{b}|\ X_1 = \mathrm{b}) = 0.2 + 0.8 = 1$$
$$\mathrm{P}(X_2 = \epsilon|\ X_1 = \mathrm{a}) + \mathrm{P}(X_2 = \epsilon|\ X_1 = \mathrm{b}) = 0 + 0.05 = 0.05$$

So, from all above calculations we see that this equation holds only when $\mathrm{P}(X_2 = \mathrm{b}|\ X_1)$.

Calculating $\sum_{i\in\{a,b,\epsilon\}} \mathbf{P}(X_2 = \mathrm{i}|\ X_1 = \mathrm{j}) = \mathbf{1},\ \textbf{for all}\ j \in \{\mathbf{a}, \mathbf{b}\}$:

$$\mathrm{P}(X_2 = \mathrm{a}|\ X_1 = \mathrm{a}) + \mathrm{P}(X_2 = \mathrm{b}|\ X_1 = \mathrm{a}) + \mathrm{P}(X_2 = \epsilon|\ X_1 = \mathrm{a}) = 0.8 + 0.2 + 0 = 1$$
$$\mathrm{P}(X_2 = \mathrm{a}|\ X_1 = \mathrm{b}) + \mathrm{P}(X_2 = \mathrm{b}|\ X_1 = \mathrm{b}) + \mathrm{P}(X_2 = \epsilon|\ X_1 = \mathrm{b}) = 0.15 + 0.8 + 0.05 = 1$$

So, from all above calculations we see that this equation holds. This is a row probability matrix. Channel maps each input to a symbol in output alphabet.

2) $\mathrm{P}(X_1 = \mathrm{a}) = \mathrm{P}(X_1 = \mathrm{b}) = \frac{1}{2}$

So, $\boldsymbol{H(X_1)} = \log_2 2 = \mathbf{1\ bit}$ **(Ans.)**

$\mathrm{P}(X_2 = \mathrm{a}) = \mathrm{P}(X_1 = \mathrm{a})\ \mathrm{P}(X_2 = \mathrm{a}|\ X_1 = \mathrm{a}) + \mathrm{P}(X_1 = \mathrm{b})\ \mathrm{P}(X_2 = \mathrm{a}|\ X_1 = \mathrm{b})$
$= (0.5 \times 0.8) + (0.5 \times 0.15) = 0.475$

$\mathrm{P}(X_2 = \mathrm{b}) = \mathrm{P}(X_1 = \mathrm{a})\ \mathrm{P}(X_2 = \mathrm{b}|\ X_1 = \mathrm{a}) + \mathrm{P}(X_1 = \mathrm{b})\ \mathrm{P}(X_2 = \mathrm{b}|\ X_1 = \mathrm{b})$
$= (0.5 \times 0.2) + (0.5 \times 0.8) = 0.5$

$\mathrm{P}(X_2 = \epsilon) = \mathrm{P}(X_1 = \mathrm{a})\ \mathrm{P}(X_2 = \epsilon|\ X_1 = \mathrm{a}) + \mathrm{P}(X_1 = \mathrm{b})\ \mathrm{P}(X_2 = \epsilon|\ X_1 = \mathrm{b})$
$= (0.5 \times 0) + (0.5 \times 0.05) = 0.025$

So,

$\boldsymbol{H(X_2)} = -[\mathrm{P}(X_2 = \mathrm{a})\log_2 \mathrm{P}(X_2 = \mathrm{a}) + \mathrm{P}(X_2 = \mathrm{b})\log_2 \mathrm{P}(X_2 = \mathrm{b}) + \mathrm{P}(X_2 = \epsilon)\log_2 \mathrm{P}(X_2 = \epsilon)] = \mathbf{1.14\ bits}$ **(Ans.)**

$$\mathrm{P}(X_1 = \mathrm{a}|\ X_2 = \mathrm{a}) = \frac{\mathrm{P}(X_1 = \mathrm{a})\ \mathrm{P}(X_2 = \mathrm{a}|\ X_1 = \mathrm{a})}{\mathrm{P}(X_2 = \mathrm{a})} = \frac{0.5 \times 0.8}{0.475} = 0.84$$
$$\mathrm{P}(X_1 = \mathrm{b}|\ X_2 = \mathrm{a}) = \frac{\mathrm{P}(X_1 = \mathrm{b})\ \mathrm{P}(X_2 = \mathrm{a}|\ X_1 = \mathrm{b})}{\mathrm{P}(X_2 = \mathrm{a})} = \frac{0.5 \times 0.15}{0.475} = 0.16$$
$$\boldsymbol{P_{X_1|X_2=a}} = (\mathbf{0.84, 0.16})$$
$$\mathrm{P}(X_1 = \mathrm{a}|\ X_2 = \mathrm{b}) = \frac{\mathrm{P}(X_1 = \mathrm{a})\ \mathrm{P}(X_2 = \mathrm{b}|\ X_1 = \mathrm{a})}{\mathrm{P}(X_2 = \mathrm{b})} = \frac{0.5 \times 0.2}{0.5} = 0.2$$
$$\mathrm{P}(X_1 = \mathrm{b}|\ X_2 = \mathrm{b}) = \frac{\mathrm{P}(X_1 = \mathrm{b})\ \mathrm{P}(X_2 = \mathrm{b}|\ X_1 = \mathrm{b})}{\mathrm{P}(X_2 = \mathrm{b})} = \frac{0.5 \times 0.8}{0.5} = 0.8$$
$$\boldsymbol{P_{X_1|X_2=b}} = (\mathbf{0.2, 0.8})$$
$$\mathrm{P}(X_1 = \mathrm{a}|\ X_2 = \epsilon) = \frac{\mathrm{P}(X_1 = \mathrm{a})\ \mathrm{P}(X_2 = \epsilon|\ X_1 = \mathrm{a})}{\mathrm{P}(X_2 = \epsilon)} = \frac{0}{0.025} = 0$$
$$\mathrm{P}(X_1 = \mathrm{b}|\ X_2 = \epsilon) = \frac{\mathrm{P}(X_1 = \mathrm{b})\ \mathrm{P}(X_2 = \epsilon|\ X_1 = \mathrm{b})}{\mathrm{P}(X_2 = \epsilon)} = \frac{0.5 \times 0.05}{0.025} = 1$$
$$\boldsymbol{P_{X_1|X_2=\epsilon}} = (\mathbf{0, 1})$$

$\boldsymbol{H(X_1|X_2=a)} = -[0.84\ \log_2 0.84 + 0.16\ \log_2 0.16] = \mathbf{0.63\ bit}$ **(Ans.)**
$\boldsymbol{H(X_1|X_2=b)} = -[0.2\ \log_2 0.2 + 0.8\ \log_2 0.8] = \mathbf{0.72\ bit}$ **(Ans.)**
$\boldsymbol{H(X_1|X_2=\epsilon)} = -[0 + 1\ \log_2 1] = \mathbf{0\ bit}$ **(Ans.)**

$\boldsymbol{H(X_1|X_2)} =$ $\mathrm{P}(X_2 = \mathrm{a})\ H(X_1|X_2=a) + \mathrm{P}(X_2 = \mathrm{b})\ H(X_1|X_2=b) + \mathrm{P}(X_2 = \epsilon)\ H(X_1|X_2=\epsilon)$
= (0.475x0.63)+(0.5x0.72)+(0.025x0) = **0.659 bit** **(Ans.)**

$H(X_2|X_1=a) = -[0.8\ \log_2 0.8 + 0.2\ \log_2 0.2 +\ 0\ ] = 0.72$ bit
$H(X_2|X_1=b) = -[0.15\ \log_2 0.15 + 0.8\ \log_2 0.8 +\ 0.05\ \log_2 0.05]\ = 0.88$ bit
$\boldsymbol{H(X_2|X_1)} = \mathrm{P}(X_1 = \mathrm{a})\ H(X_2|X_1=a) + \mathrm{P}(X_1 = \mathrm{b})\ H(X_2|X_1=b) = \mathbf{0.8\ bit}$ **(Ans.)**

3) $\boldsymbol{H(X_1)} - \boldsymbol{H(X_1|X_2=a)} = 1 - 0.63 = 0.37$ bit
$\boldsymbol{H(X_1)} - \boldsymbol{H(X_1|X_2=b)} = 1 - 0.72 = 0.28$ bit
$\boldsymbol{H(X_1)} - \boldsymbol{H(X_1|X_2=\epsilon)} = 1 - 0 = 1$ bit
So, $\boldsymbol{X_2} = \epsilon$ gives most amount of information about input source.

4) $P(X_1 = a, X_2 = a) = \mathrm{P}(X_1 = \mathrm{a})\ \mathrm{P}(X_2 = \mathrm{a}|\ X_1 = \mathrm{a}) = 0.5\mathrm{x}0.8 = 0.4$
$P(X_1 = a, X_2 = b) = \mathrm{P}(X_1 = \mathrm{a})\ \mathrm{P}(X_2 = \mathrm{b}|\ X_1 = \mathrm{a}) = 0.5\mathrm{x}0.2 = 0.1$
$P(X_1 = a, X_2 = \epsilon) = \mathrm{P}(X_1 = \mathrm{a})\ \mathrm{P}(X_2 = \epsilon|\ X_1 = \mathrm{a}) = 0.5\mathrm{x}0 = 0$
$P(X_1 = b, X_2 = a) = \mathrm{P}(X_1 = \mathrm{b})\ \mathrm{P}(X_2 = \mathrm{a}|\ X_1 = \mathrm{b}) = 0.5\mathrm{x}0.15 = 0.075$
$P(X_1 = b, X_2 = b) = \mathrm{P}(X_1 = \mathrm{b})\ \mathrm{P}(X_2 = \mathrm{b}|\ X_1 = \mathrm{b}) = 0.5\mathrm{x}0.8 = 0.4$
$P(X_1 = b, X_2 = \epsilon) = \mathrm{P}(X_1 = \mathrm{b})\ \mathrm{P}(X_2 = \epsilon|\ X_1 = \mathrm{b}) = 0.5\mathrm{x}0.05 = 0.025$
$\boldsymbol{H(X_1, X_2)} = -[0.4\ \log_2 0.4 + 0.1\ \log_2 0.1 +\ 0\ + 0.075\ \log_2 0.075 +\ 0.4\ \log_2 0.4 +$
$0.025\ \log_2 0.025]$ = **1.803 bits**

$\boldsymbol{H(X_1)} = 1$ bit
$\boldsymbol{H(X_2)} = 1.14$ bits
$\boldsymbol{H(X_1)} + \boldsymbol{H(X_2)}\ = 2.14$ bits

So, from above calculations we see that,

$$\boldsymbol{H(X_1, X_2)} < \boldsymbol{H(X_1)} + \boldsymbol{H(X_2)}$$

The result is expected because:

$$\boldsymbol{H(X_1, X_2)} = \boldsymbol{H(X_1)} + \boldsymbol{H(X_2)} - \boldsymbol{I(X_1;\ X_2)}$$

5) **(i)** $\boldsymbol{I(X_1;\ X_2)} = H(X_1) - H(X_1|X_2) = 1 - 0.659 = \mathbf{0.341\ bit}$
(Ans.)
(ii) $\boldsymbol{D(P(X_1, X_2)\ ) \parallel P(X_1)\ P(X_2))}$
$= P(a,a)\ \log_2 \frac{P(a,a)}{P(X_1=a)\ P(X_2=a)} +\ P(a,b)\ \log_2 \frac{P(a,b)}{P(X_1=a)\ P(X_2=b)} +\ P(a,\epsilon)\ \log_2 \frac{P(a,\epsilon)}{P(X_1=a)\ P(X_2=\epsilon)} +$
$P(b,a)\ \log_2 \frac{P(b,a)}{P(X_1=b)\ P(X_2=a)} +\ P(b,b)\ \log_2 \frac{P(b,b)}{P(X_1=b)\ P(X_2=b)} +\ P(b,\epsilon)\ \log_2 \frac{P(b,\epsilon)}{P(X_1=b)\ P(X_2=\epsilon)}$

= 0.34

So, $\boldsymbol{I(X_1;\, X_2) = D(P(X_1, X_2)) \parallel P(X_1)\, P(X_1))}$ **(Ans.)**

6) Given, $\mathrm{P}(X_1 = \mathrm{a}) = 0.4\ and\ \mathrm{P}(X_1 = \mathrm{b}) = 0.6$
$\boldsymbol{H(X_1)} = -[0.4 \log_2 0.4 + 0.6 \log_2 0.6] = \mathbf{0.97\ bit}$

$\mathrm{P}(X_2 = \mathrm{a}) = \mathrm{P}(X_1 = \mathrm{a})\, \mathrm{P}(X_2 = \mathrm{a}|\, X_1 = \mathrm{a}) + \mathrm{P}(X_1 = \mathrm{b})\, \mathrm{P}(X_2 = \mathrm{a}|\, X_1 = \mathrm{b})$
$= (0.4 \text{ x } 0.8) + (0.6 \text{ x } 0.15) = 0.41$
$\mathrm{P}(X_2 = \mathrm{b}) = \mathrm{P}(X_1 = \mathrm{a})\, \mathrm{P}(X_2 = \mathrm{b}|\, X_1 = \mathrm{a}) + \mathrm{P}(X_1 = \mathrm{b})\, \mathrm{P}(X_2 = \mathrm{b}|\, X_1 = \mathrm{b})$
$= (0.4 \text{ x } 0.2) + (0.6 \text{ x } 0.8) = 0.56$
$\mathrm{P}(X_2 = \epsilon) = \mathrm{P}(X_1 = \mathrm{a})\, \mathrm{P}(X_2 = \epsilon|\, X_1 = \mathrm{a}) + \mathrm{P}(X_1 = \mathrm{b})\, \mathrm{P}(X_2 = \epsilon|\, X_1 = \mathrm{b})$
$= 0 + (0.6 \text{ x } 0.05) = 0.03$
So, $\boldsymbol{H(X_2)} = -[0.41 \log_2 0.41 + 0.56 \log_2 0.56 + 0.03 \log_2 0.03] = \mathbf{1.15\ bits}$

$\mathrm{P}(X_1 = \mathrm{a}|\, X_2 = \mathrm{a}) = \dfrac{0.4 \text{ x } 0.8}{0.41} = 0.78$
$\mathrm{P}(X_1 = \mathrm{b}|\, X_2 = \mathrm{a}) = \dfrac{0.6 \text{ x } 0.15}{0.41} = 0.22$
$\boldsymbol{P_{X_1|X_2=a} = (0.78, 0.22)}$
$\mathrm{P}(X_1 = \mathrm{a}|\, X_2 = \mathrm{b}) = \dfrac{0.4 \text{ x } 0.2}{0.56} = 0.14$
$\mathrm{P}(X_1 = \mathrm{b}|\, X_2 = \mathrm{b}) = \dfrac{0.6 \text{ x } 0.8}{0.56} = 0.86$
$\boldsymbol{P_{X_1|X_2=b} = (0.14, 0.86)}$
$\mathrm{P}(X_1 = \mathrm{a}|\, X_2 = \epsilon) = \dfrac{0}{0.03} = 0$
$\mathrm{P}(X_1 = \mathrm{b}|\, X_2 = \epsilon) = \dfrac{0.6 \text{ x } 0.05}{0.03} = 1$
$\boldsymbol{P_{X_1|X_2=\epsilon} = (0, 1)}$
$\boldsymbol{H(X_1|X_2 = a)} = -[0.78 \log_2 0.78 + 0.22 \log_2 0.22] = \mathbf{0.76\ bit}$
$\boldsymbol{H(X_1|X_2 = b)} = -[0.14 \log_2 0.14 + 0.86 \log_2 0.86] = \mathbf{0.58\ bit}$
$\boldsymbol{H(X_1|X_2 = \epsilon)} = -[0 + 1 \log_2 1] = \mathbf{0\ bit}$
$\boldsymbol{H(X_1|X_2)} = (0.41\ x\ 0.76) + (0.56\ x\ 0.58) + (0.03\ x\ 0) = \mathbf{0.64\ bit}$

$\boldsymbol{I(X_1;\, X_2)} = H(X_1) - H(X_1|X_2) = 0.97 - 0.64 = \mathbf{0.33\ bit}$
Yes, the amount of information that passes through the channel has reduced for the distribution of the source $\boldsymbol{S'}$.

7) **(i)** $\boldsymbol{H_\infty(X_1)} = -\log_2 0.5 = \mathbf{1\ bit}$ **(Ans.)**
(ii) $\boldsymbol{H_\infty(X_1|X_2 = a)} = -\log_2 0.84 = \mathbf{0.25\ bit}$
$\boldsymbol{H_\infty(X_1|X_2 = b)} = -\log_2 0.8 = \mathbf{0.32\ bit}$
$\boldsymbol{H_\infty(X_1|X_2 = \epsilon)} = -\log_2 1 = \mathbf{0\ bit}$
So, $\boldsymbol{X_2} = \epsilon$ has best success chance. **(Ans.)**

8)
Let probability distribution of $\boldsymbol{X_1}$: $\{a = \alpha,\ b = 1 - \alpha\}$

Then we have the probability distribution for $\boldsymbol{X_2}$
$\{a = 0.65\alpha + 0.15, b = 0.8 - 0.6\alpha, \epsilon = 0.05 - 0.05\alpha\}$

$\boldsymbol{H}(\boldsymbol{X_2})$ = -(0.65α + 0.15)$\log_2$(0.65α + 0.15) − (0.8 − 0.6α)$\log_2$(0.8 − 0.6α) − (0.05 − 0.05α)$\log_2$(0.05 − 0.05α)

$\boldsymbol{H}(\boldsymbol{X_1}|\boldsymbol{X_2}) = \alpha \times 0.7219 + (1-\alpha) \times 0.8841 = 0.8841 - 0.1622\alpha$

$\boldsymbol{I}(\boldsymbol{X_1};\ \boldsymbol{X_2})$ = -$(0.65\alpha + 0.15)log_2(0.65\alpha + 0.15) - (0.8 - 0.6\alpha)log_2(0.8 - 0.6\alpha) - (0.05 - 0.05\alpha)log_2(0.05 - 0.05\alpha) - (0.8841 - 0.1622\alpha)$

Finding max value of $\boldsymbol{I}(\boldsymbol{X_1};\ \boldsymbol{X_2})\ while\ \boldsymbol{\alpha} \in [\boldsymbol{0}, \boldsymbol{1}]$ (using **MATLAB fplot**)

C = MAX $\boldsymbol{I}(\boldsymbol{X_1};\ \boldsymbol{X_2})$ exists when $\boldsymbol{\alpha = 0.49}$, the maximum channel capacity approximates to 0.34.

Question 2 Answer

For sequence 1: $n_1 = 16$

$\boldsymbol{S_{n_1}} = 1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1 = \mathbf{16}$

Test statistic, $\boldsymbol{S_{obs}} = \frac{|16|}{\sqrt{16}} = 4$

$\boldsymbol{P}$ value = $erfc\left(\frac{S_{obs}}{\sqrt{2}}\right) = erfc\left(\frac{4}{\sqrt{2}}\right) = 0.00006275$

As, $\mathbf{0.00006275 < 0.01}$, this sequence appears to be **non-random** or this sequence is not generated by a good random number generator.

For sequence 2: $n_2 = 26$

$\boldsymbol{S_{n_2}} = 1 + (-1) + (-1) + 1 + (-1) + 1 + (-1) + 1 + (-1) + 1 + (-1) + (-1) + (-1) + (-1) + (-1) + (-1) + (-1) + (-1) + (-1) + (-1) + (-1) + (-1) + (-1) + (-1) + (-1) + (-1) = \mathbf{-16}$

Test statistic, $\boldsymbol{S_{obs}} = \frac{|-16|}{\sqrt{26}} = 3.14$

$\boldsymbol{P}$ value = $erfc\left(\frac{S_{obs}}{\sqrt{2}}\right) = erfc\left(\frac{3.14}{\sqrt{2}}\right) = 0.001692052$

As, $\mathbf{0.001692052 < 0.01}$, this sequence appears to be **non-random** or this sequence is not generated by a good random number generator.

For sequence 3: $n_3 = 26$

$\boldsymbol{S_{n_3}} = 1 + (-1) + (-1) + 1 + (-1) + 1 + (-1) + 1 + (-1) + 1 + 1 + (-1) + (-1) + (-1) + 1 + (-1) + 1 + (-1) + 1 + 1 + 1 + (-1) + (-1) + 1 + (-1) + 1 = \mathbf{0}$

Test statistic, $\boldsymbol{S_{obs}} = \frac{|0|}{\sqrt{26}} = 0$

$\boldsymbol{P}$ value = $erfc\left(\frac{S_{obs}}{\sqrt{2}}\right) = erfc\left(\frac{0}{\sqrt{2}}\right) = 1$

As, $\mathbf{1 \geq 0.01}$, this sequence is **random,** or this sequence is generated by a good random number generator.

Question 3 Answer

Given the distribution of $\boldsymbol{X}$ and $\boldsymbol{K}$ are uniform.

So, $\mathrm{P}(X = 1) = \mathrm{P}(X = 2) = \mathrm{P}(X = 3) = \frac{1}{3}$

And, $\mathrm{P}(K = k_1) = \mathrm{P}(K = k_2) = \mathrm{P}(K = k_3) = \mathrm{P}(K = k_4) = \frac{1}{4}$

$\mathbf{P}(\boldsymbol{Y} = \mathbf{1}) = \mathrm{P}(X = 2)\,\mathrm{P}(K = k_1) + \mathrm{P}(X = 2)\,\mathrm{P}(K = k_3) = (\frac{1}{3} \times \frac{1}{4}) + (\frac{1}{3} \times \frac{1}{4}) = \frac{\mathbf{1}}{\mathbf{6}}$
$\mathbf{P}(\boldsymbol{Y} = \mathbf{2}) = \mathrm{P}(X = 1)\,\mathrm{P}(K = k_1) + \mathrm{P}(X = 3)\,\mathrm{P}(K = k_2) + \mathrm{P}(X = 1)\,\mathrm{P}(K = k_3) + \mathrm{P}(X = 2)\,\mathrm{P}(K = k_4) = (\frac{1}{12} + \frac{1}{12} + \frac{1}{12} + \frac{1}{12}) = \frac{\mathbf{1}}{\mathbf{3}}$
$\mathbf{P}(\boldsymbol{Y} = \mathbf{3}) = \mathrm{P}(X = 2)\,\mathrm{P}(K = k_2) + \mathrm{P}(X = 3)\,\mathrm{P}(K = k_3) + \mathrm{P}(X = 1)\,\mathrm{P}(K = k_4) = (\frac{1}{12} + \frac{1}{12} + \frac{1}{12}) = \frac{\mathbf{1}}{\mathbf{4}}$
$\mathbf{P}(\boldsymbol{Y} = \mathbf{4}) = \mathrm{P}(X = 3)\,\mathrm{P}(K = k_1) + \mathrm{P}(X = 1)\,\mathrm{P}(K = k_2) + \mathrm{P}(X = 3)\,\mathrm{P}(K = k_4) = (\frac{1}{12} + \frac{1}{12} + \frac{1}{12}) = \frac{\mathbf{1}}{\mathbf{4}}$

1) $$\mathbf{P}(\boldsymbol{X} = \mathbf{1} | \boldsymbol{Y} = \mathbf{1}) = \frac{\mathrm{P}(X = 1)\,\mathrm{P}(Y = 1 | X = 1)}{\mathrm{P}(Y = 1)} = \frac{\mathrm{P}(X = 1) \times 0}{\mathrm{P}(Y = 1)} = 0 \neq \mathbf{P}(\boldsymbol{X} = \mathbf{1})$$

So, this system doesn't provide perfect secrecy.

Reason: An encryption system is perfectly secure if for all the P(X|Y) observations on X, we have **P(X|Y) = P(X)**. In the above observation, this condition doesn't hold. So it is not perfectly secure.

2)

$$\mathbf{P}(\boldsymbol{X} = \mathbf{1} | \boldsymbol{Y} = \mathbf{1}) = \frac{\mathrm{P}(X = 1)\,\mathrm{P}(Y = 1 | X = 1)}{\mathrm{P}(Y = 1)} = \frac{\frac{1}{3} \times 0}{\frac{1}{6}} = 0$$

$$\mathbf{P}(\boldsymbol{X} = \mathbf{2} | \boldsymbol{Y} = \mathbf{1}) = \frac{\mathrm{P}(X = 2)\,\mathrm{P}(Y = 1 | X = 2)}{\mathrm{P}(Y = 1)} = \frac{\frac{1}{3} \times \frac{1}{2}}{\frac{1}{6}} = 1$$

$$\mathbf{P}(\boldsymbol{X} = \mathbf{3} | \boldsymbol{Y} = \mathbf{1}) = \frac{\mathrm{P}(X = 3)\,\mathrm{P}(Y = 1 | X = 3)}{\mathrm{P}(Y = 1)} = \frac{\frac{1}{3} \times 0}{\frac{1}{6}} = 0$$

$$\boldsymbol{P_{X|Y=1}} = (\mathbf{0}, \mathbf{1}, \mathbf{0})$$

$$\mathbf{P}(\boldsymbol{X} = \mathbf{1} | \boldsymbol{Y} = \mathbf{2}) = \frac{\mathrm{P}(X = 1)\,\mathrm{P}(Y = 2 | X = 1)}{\mathrm{P}(Y = 2)} = \frac{\frac{1}{3} \times \frac{1}{2}}{\frac{1}{3}} = \frac{1}{2}$$

$$\mathbf{P}(\boldsymbol{X} = \mathbf{2} | \boldsymbol{Y} = \mathbf{2}) = \frac{\mathrm{P}(X = 2)\,\mathrm{P}(Y = 2 | X = 2)}{\mathrm{P}(Y = 2)} = \frac{\frac{1}{3} \times \frac{1}{4}}{\frac{1}{3}} = \frac{1}{4}$$

$$\mathbf{P}(\boldsymbol{X} = \mathbf{3} | \boldsymbol{Y} = \mathbf{2}) = \frac{\mathrm{P}(X = 3)\,\mathrm{P}(Y = 2 | X = 3)}{\mathrm{P}(Y = 2)} = \frac{\frac{1}{3} \times \frac{1}{4}}{\frac{1}{3}} = \frac{1}{4}$$

$$\boldsymbol{P_{X|Y=2}} = (\mathbf{1/2}, \mathbf{1/4}, \mathbf{1/4})$$

$$\mathbf{P}(\boldsymbol{X} = \mathbf{1} | \boldsymbol{Y} = \mathbf{3}) = \frac{\mathrm{P}(X = 1)\,\mathrm{P}(Y = 3 | X = 1)}{\mathrm{P}(Y = 3)} = \frac{\frac{1}{3} \times \frac{1}{4}}{\frac{1}{4}} = \frac{1}{3}$$

$$\mathbf{P}(\boldsymbol{X} = \mathbf{2} | \boldsymbol{Y} = \mathbf{3}) = \frac{\mathrm{P}(X = 2)\,\mathrm{P}(Y = 3 | X = 2)}{\mathrm{P}(Y = 3)} = \frac{\frac{1}{3} \times \frac{1}{4}}{\frac{1}{4}} = \frac{1}{3}$$

$$\mathbf{P}(\boldsymbol{X} = \mathbf{3} | \boldsymbol{Y} = \mathbf{3}) = \frac{\mathrm{P}(X = 3)\,\mathrm{P}(Y = 3 | X = 3)}{\mathrm{P}(Y = 3)} = \frac{\frac{1}{3} \times \frac{1}{4}}{\frac{1}{4}} = \frac{1}{3}$$

$$\boldsymbol{P_{X|Y=3}} = (\mathbf{1/3}, \mathbf{1/3}, \mathbf{1/3})$$

$$\mathbf{P}(\boldsymbol{X}=\mathbf{1} \mid \boldsymbol{Y}=\mathbf{4}) = \frac{\mathrm{P}(X=1)\ \mathrm{P}(Y=4 \mid X=1)}{\mathrm{P}(Y=4)} = \frac{\frac{1}{3} \text{ x } \frac{1}{4}}{\frac{1}{4}} = \frac{1}{3}$$

$$\mathbf{P}(\boldsymbol{X}=\mathbf{2} \mid \boldsymbol{Y}=\mathbf{4}) = \frac{\mathrm{P}(X=2)\ \mathrm{P}(Y=4 \mid X=2)}{\mathrm{P}(Y=4)} = \frac{\frac{1}{3} \text{ x } 0}{\frac{1}{4}} = 0$$

$$\mathbf{P}(\boldsymbol{X}=\mathbf{3} \mid \boldsymbol{Y}=\mathbf{4}) = \frac{\mathrm{P}(X=3)\ \mathrm{P}(Y=4 \mid X=3)}{\mathrm{P}(Y=4)} = \frac{\frac{1}{3} \text{ x } \frac{1}{2}}{\frac{1}{4}} = \frac{2}{3}$$

$$\boldsymbol{P}_{\boldsymbol{X}|\boldsymbol{Y}=\mathbf{4}} = (\mathbf{1/3}, \mathbf{0}, \mathbf{2/3})$$

(i) $\boldsymbol{H}(\boldsymbol{X}) = \log_2 3 = \mathbf{1.58}$ **bits**

$\boldsymbol{H}(\boldsymbol{X}|\boldsymbol{Y}=\mathbf{1}) = \mathbf{0}$ **bit**

$\boldsymbol{H}(\boldsymbol{X}|\boldsymbol{Y}=\mathbf{2}) = -\left[\frac{1}{2}\log_2\frac{1}{2} + \frac{1}{4}\log_2\frac{1}{4} + \frac{1}{4}\log_2\frac{1}{4}\right] = \mathbf{1.5}$ **bits**

$\boldsymbol{H}(\boldsymbol{X}|\boldsymbol{Y}=\mathbf{3}) = \log_2 3 = \mathbf{1.58}$ **bits**

$\boldsymbol{H}(\boldsymbol{X}|\boldsymbol{Y}=\mathbf{4}) = -\left[\frac{1}{3}\log_2\frac{1}{3} + 0 + \frac{2}{3}\log_2\frac{2}{3}\right] = \mathbf{0.92}$ **bit**

$\boldsymbol{H}(\boldsymbol{X}) - \boldsymbol{H}(\boldsymbol{X}|\boldsymbol{Y}=\mathbf{1}) = \mathbf{1.58}$ **bits**

$\boldsymbol{H}(\boldsymbol{X}) - \boldsymbol{H}(\boldsymbol{X}|\boldsymbol{Y}=\mathbf{2}) = \mathbf{0.08}$ **bits**

$\boldsymbol{H}(\boldsymbol{X}) - \boldsymbol{H}(\boldsymbol{X}|\boldsymbol{Y}=\mathbf{3}) = \mathbf{0}$ **bit**

$\boldsymbol{H}(\boldsymbol{X}) - \boldsymbol{H}(\boldsymbol{X}|\boldsymbol{Y}=\mathbf{4}) = \mathbf{0.66}$ **bits**

So, ϵ = 1.58 bits. Therefore, ***Y=1*** leaks most amount of information.

(ii) $\boldsymbol{SD}(\boldsymbol{P}_{\boldsymbol{X}}, \boldsymbol{P}_{\boldsymbol{X}|\boldsymbol{Y}=\mathbf{1}}) = \frac{1}{2}\left(\left|\frac{1}{3} - 0\right| + \left|\frac{1}{3} - 1\right| + \left|\frac{1}{3} - 0\right|\right) = \mathbf{2/3}$

$\boldsymbol{SD}(\boldsymbol{P}_{\boldsymbol{X}}, \boldsymbol{P}_{\boldsymbol{X}|\boldsymbol{Y}=\mathbf{2}}) = \frac{1}{2}\left(\left|\frac{1}{3} - \frac{1}{2}\right| + \left|\frac{1}{3} - \frac{1}{4}\right| + \left|\frac{1}{3} - \frac{1}{4}\right|\right) = \mathbf{1/6}$

$\boldsymbol{SD}(\boldsymbol{P}_{\boldsymbol{X}}, \boldsymbol{P}_{\boldsymbol{X}|\boldsymbol{Y}=\mathbf{3}}) = \mathbf{0}$

$\boldsymbol{SD}(\boldsymbol{P}_{\boldsymbol{X}}, \boldsymbol{P}_{\boldsymbol{X}|\boldsymbol{Y}=\mathbf{4}}) = \frac{1}{2}\left(\left|\frac{1}{3} - \frac{1}{3}\right| + \left|\frac{1}{3} - 0\right| + \left|\frac{1}{3} - \frac{2}{3}\right|\right) = \mathbf{1/3}$

So, ϵ = **2/3** bits. Therefore, ***Y=1*** leaks most amount of information.
Yes, both measures point to the same ciphertext.

3)

$$\mathbf{P}(\boldsymbol{K}=\boldsymbol{k}_\mathbf{1} \mid \boldsymbol{Y}=\mathbf{1}) = \frac{\mathrm{P}(K=k_1)\ \mathrm{P}(Y=1 \mid K=k_1)}{\mathrm{P}(Y=1)} = \frac{\frac{1}{4} \text{ x } \frac{1}{3}}{\frac{1}{6}} = \frac{1}{2}$$

$$\mathbf{P}(\boldsymbol{K}=\boldsymbol{k}_\mathbf{2} \mid \boldsymbol{Y}=\mathbf{1}) = \frac{\mathrm{P}(K=k_2)\ \mathrm{P}(Y=1 \mid K=k_2)}{\mathrm{P}(Y=1)} = \frac{\frac{1}{4} \text{ x } 0}{\frac{1}{6}} = 0$$

$$\mathbf{P}(\boldsymbol{K}=\boldsymbol{k}_\mathbf{3} \mid \boldsymbol{Y}=\mathbf{1}) = \frac{\mathrm{P}(K=k_3)\ \mathrm{P}(Y=1 \mid K=k_3)}{\mathrm{P}(Y=1)} = \frac{\frac{1}{4} \text{ x } \frac{1}{3}}{\frac{1}{6}} = \frac{1}{2}$$

$$\mathbf{P}(\boldsymbol{K}=\boldsymbol{k}_\mathbf{4} \mid \boldsymbol{Y}=\mathbf{1}) = \frac{\mathrm{P}(K=k_4)\ \mathrm{P}(Y=1 \mid K=k_4)}{\mathrm{P}(Y=1)} = \frac{\frac{1}{4} \text{ x } 0}{\frac{1}{6}} = 0$$

$$\boldsymbol{P}_{\boldsymbol{K}|\boldsymbol{Y}=\mathbf{1}} = (\mathbf{1/2}, \mathbf{0}, \mathbf{1/2}, \mathbf{0})$$

$$\mathbf{P}(\boldsymbol{K}=\boldsymbol{k_1}|\,\boldsymbol{Y}=\mathbf{2}) = \frac{\mathrm{P}(K=k_1)\ \mathrm{P}(Y=2|\ K=k_1)}{\mathrm{P}(Y=2)} = \frac{\frac{1}{4}\ \mathrm{x}\ \frac{1}{3}}{\frac{1}{3}} = \frac{1}{4}$$

$$\mathbf{P}(\boldsymbol{K}=\boldsymbol{k_2}|\,\boldsymbol{Y}=\mathbf{2}) = \frac{\mathrm{P}(K=k_2)\ \mathrm{P}(Y=2|\ K=k_2)}{\mathrm{P}(Y=2)} = \frac{\frac{1}{4}\ \mathrm{x}\ \frac{1}{3}}{\frac{1}{3}} = \frac{1}{4}$$

$$\mathbf{P}(\boldsymbol{K}=\boldsymbol{k_3}|\,\boldsymbol{Y}=\mathbf{2}) = \frac{\mathrm{P}(K=k_3)\ \mathrm{P}(Y=2|\ K=k_3)}{\mathrm{P}(Y=2)} = \frac{\frac{1}{4}\ \mathrm{x}\ \frac{1}{3}}{\frac{1}{3}} = \frac{1}{4}$$

$$\mathbf{P}(\boldsymbol{K}=\boldsymbol{k_4}|\,\boldsymbol{Y}=\mathbf{2}) = \frac{\mathrm{P}(K=k_4)\ \mathrm{P}(Y=2|\ K=k_4)}{\mathrm{P}(Y=2)} = \frac{\frac{1}{4}\ \mathrm{x}\ \frac{1}{3}}{\frac{1}{3}} = \frac{1}{4}$$

$$\boldsymbol{P_{K|Y=2}} = (\mathbf{1/4}, \mathbf{1/4}, \mathbf{1/4}, \mathbf{1/4})$$

$$\mathbf{P}(\boldsymbol{K}=\boldsymbol{k_1}|\,\boldsymbol{Y}=\mathbf{3}) = \frac{\mathrm{P}(K=k_1)\ \mathrm{P}(Y=3|\ K=k_1)}{\mathrm{P}(Y=3)} = \frac{\frac{1}{4}\ \mathrm{x}\ 0}{\frac{1}{4}} = 0$$

$$\mathbf{P}(\boldsymbol{K}=\boldsymbol{k_2}|\,\boldsymbol{Y}=\mathbf{3}) = \frac{\mathrm{P}(K=k_2)\ \mathrm{P}(Y=3|\ K=k_2)}{\mathrm{P}(Y=3)} = \frac{\frac{1}{4}\ \mathrm{x}\ \frac{1}{3}}{\frac{1}{4}} = \frac{1}{3}$$

$$\mathbf{P}(\boldsymbol{K}=\boldsymbol{k_3}|\,\boldsymbol{Y}=\mathbf{3}) = \frac{\mathrm{P}(K=k_3)\ \mathrm{P}(Y=3|\ K=k_3)}{\mathrm{P}(Y=3)} = \frac{\frac{1}{4}\ \mathrm{x}\ \frac{1}{3}}{\frac{1}{4}} = \frac{1}{3}$$

$$\mathbf{P}(\boldsymbol{K}=\boldsymbol{k_4}|\,\boldsymbol{Y}=\mathbf{3}) = \frac{\mathrm{P}(K=k_4)\ \mathrm{P}(Y=3|\ K=k_4)}{\mathrm{P}(Y=3)} = \frac{\frac{1}{4}\ \mathrm{x}\ \frac{1}{3}}{\frac{1}{4}} = \frac{1}{3}$$

$$\boldsymbol{P_{K|Y=3}} = (\mathbf{0}, \mathbf{1/3}, \mathbf{1/3}, \mathbf{1/3})$$

$$\mathbf{P}(\boldsymbol{K}=\boldsymbol{k_1}|\,\boldsymbol{Y}=\mathbf{4}) = \frac{\mathrm{P}(K=k_1)\ \mathrm{P}(Y=4|\ K=k_1)}{\mathrm{P}(Y=4)} = \frac{\frac{1}{4}\ \mathrm{x}\ \frac{1}{3}}{\frac{1}{4}} = \frac{1}{3}$$

$$\mathbf{P}(\boldsymbol{K}=\boldsymbol{k_2}|\,\boldsymbol{Y}=\mathbf{4}) = \frac{\mathrm{P}(K=k_2)\ \mathrm{P}(Y=4|\ K=k_2)}{\mathrm{P}(Y=4)} = \frac{\frac{1}{4}\ \mathrm{x}\ \frac{1}{3}}{\frac{1}{4}} = \frac{1}{3}$$

$$\mathbf{P}(\boldsymbol{K}=\boldsymbol{k_3}|\,\boldsymbol{Y}=\mathbf{4}) = \frac{\mathrm{P}(K=k_3)\ \mathrm{P}(Y=4|\ K=k_3)}{\mathrm{P}(Y=4)} = \frac{\frac{1}{4}\ \mathrm{x}\ 0}{\frac{1}{4}} = 0$$

$$\mathbf{P}(\boldsymbol{K}=\boldsymbol{k_4}|\,\boldsymbol{Y}=\mathbf{4}) = \frac{\mathrm{P}(K=k_4)\ \mathrm{P}(Y=4|\ K=k_4)}{\mathrm{P}(Y=4)} = \frac{\frac{1}{4}\ \mathrm{x}\ \frac{1}{3}}{\frac{1}{4}} = \frac{1}{3}$$

$$\boldsymbol{P_{K|Y=3}} = (\mathbf{1/3}, \mathbf{1/3}, \mathbf{0}, \mathbf{1/3})$$

$\boldsymbol{H}(\boldsymbol{K}) = \log_2 4 = \mathbf{2\ bits}$

$\boldsymbol{H}(\boldsymbol{K}|\boldsymbol{Y}=\mathbf{1}) = \mathbf{1\ bit}$

$\boldsymbol{H}(\boldsymbol{K}|\boldsymbol{Y}=\mathbf{2}) = \log_2 4 = \mathbf{2\ bits}$

$\boldsymbol{H}(\boldsymbol{K}|\boldsymbol{Y}=\mathbf{3}) = -\left[0 + \left(\frac{1}{3}\log_2\frac{1}{3}\right).3\ \right] = \mathbf{1.58\ bits}$

$$H(K|Y=4) = -\left[\left(\frac{1}{3}\log_2\frac{1}{3}\right).3+0\right] = \mathbf{1.58\ bit}$$

$\boldsymbol{H(K) - H(K|Y=1) = 1}$ **bit**
$\boldsymbol{H(K) - H(K|Y=2) = 0}$ **bit**
$\boldsymbol{H(K) - H(K|Y=3) = 0.42}$ **bit**
$\boldsymbol{H(K) - H(K|Y=4) = 0.42}$ **bits**

So, ϵ = 1 bit. Therefore, ***Y=1*** leaks most amount of information about the key.

4) **(a)** Shannon entropy of the key after observing ciphertext $\boldsymbol{Y_1=3}$ is:

$$\boldsymbol{H(K|Y_1=3) = 1.58}\ \textbf{bits}$$

Shannon entropy of the plaintext after observing ciphertext $\boldsymbol{Y_1=3}$ is:

$$\boldsymbol{H(X|Y_1=3) = 1.58}\ \textbf{bits}$$

(b) $\underline{\boldsymbol{Y_2=4}}$ seen:

$\boldsymbol{P(K=k_1|\,Y_1=3, Y_2=4) = 0,}$ as k_1 does not produce ciphertext 3.
$\boldsymbol{P(K=k_3|\,Y_1=3, Y_2=4) = 0,}$ as k_3 does not produce ciphertext 4.
$\boldsymbol{P(K=k_2|\,Y_1=3, Y_2=4) = P(K=k_4|\,Y_1=3, Y_2=4) = 1/2,}$ as k_2 , k_4 each produce ciphertext 3 and 4 with equal probability.

$$\boldsymbol{H(K|Y_1=3, Y_2=4)} = -\left[0+\left(\frac{1}{2}\log_2\frac{1}{2}\right).2\right] = \mathbf{1\ bit}$$

$\boldsymbol{H(K) - H(K|Y_1=3, Y_2=4)} = 2-1 = \mathbf{1\ bit}$

From above calculations we have seen that, $\boldsymbol{H(K|Y_1=3) = 1.58}$ **bits** and $\boldsymbol{H(K|Y_1=3, Y_2=4) = 1}$ **bit.** That means observing the second ciphertext has reduced the uncertainty of the key.

5) (**Bonus**)

For the given table, reasons for going ϵ value high are: (i) plaintexts **X=1** and **X=3** are never mapped to ciphertext 1 and (ii) plaintext **X=2** is never mapped to ciphertext 4.

Considering the above facts, we re-design the encryption table using same size ciphertext and key space as follows:

		X		
		1	**2**	**3**
K	$\boldsymbol{k_1}$	2	1	4
	$\boldsymbol{k_2}$	1	3	2
	$\boldsymbol{k_3}$	4	2	3
	$\boldsymbol{k_4}$	3	4	1

Table: Encryption/Decryption Table

Assuming the probability distributions on ***X*** and ***K*** uniform, we calculate the new probability distribution over ***Y***.

$\boldsymbol{P(Y=1)} = (\frac{1}{12}+\frac{1}{12}+\frac{1}{12}) = \boldsymbol{\frac{1}{4}}$
$\boldsymbol{P(Y=2)} = (\frac{1}{12}+\frac{1}{12}+\frac{1}{12}) = \boldsymbol{\frac{1}{4}}$
$\boldsymbol{P(Y=3)} = (\frac{1}{12}+\frac{1}{12}+\frac{1}{12}) = \boldsymbol{\frac{1}{4}}$

$\mathbf{P}(\boldsymbol{Y}=\mathbf{4})=(\frac{1}{12}+\frac{1}{12}+\frac{1}{12})=\frac{\mathbf{1}}{\mathbf{4}}$

By using the following formula, we construct the table with posterior probabilities $P(X = x|Y = y)$:

$$P(X=x|Y=y)=\frac{P(X=x)\,P(Y=y|X=x)}{P(Y=y)}$$

P(X|Y)		***X*** 1	2	3
Y	***1***	1/3	1/3	1/3
	2	1/3	1/3	1/3
	3	1/3	1/3	1/3
	4	1/3	1/3	1/3

Now we calculate statistical distance to measure the leakage of this new design.

$\boldsymbol{SD}(\boldsymbol{P_X},\boldsymbol{P_{X|Y=1}})=\frac{1}{2}(\left|\frac{1}{3}-\frac{1}{3}\right|+\left|\frac{1}{3}-\frac{1}{3}\right|+\left|\frac{1}{3}-\frac{1}{3}\right|)=0$

$\boldsymbol{SD}(\boldsymbol{P_X},\boldsymbol{P_{X|Y=2}})=\frac{1}{2}(\left|\frac{1}{3}-\frac{1}{3}\right|+\left|\frac{1}{3}-\frac{1}{3}\right|+\left|\frac{1}{3}-\frac{1}{3}\right|)=\mathbf{0}$

$\boldsymbol{SD}(\boldsymbol{P_X},\boldsymbol{P_{X|Y=3}})=\frac{1}{2}(\left|\frac{1}{3}-\frac{1}{3}\right|+\left|\frac{1}{3}-\frac{1}{3}\right|+\left|\frac{1}{3}-\frac{1}{3}\right|)=\mathbf{0}$

$\boldsymbol{SD}(\boldsymbol{P_X},\boldsymbol{P_{X|Y=4}})=\frac{1}{2}(\left|\frac{1}{3}-\frac{1}{3}\right|+\left|\frac{1}{3}-\frac{1}{3}\right|+\left|\frac{1}{3}-\frac{1}{3}\right|)=\mathbf{0}$

Therefore, ϵ = **0** bit.

The ϵ value is reduced by **2/3 – 0 = 2/3** bit. So, in our new design the leakage about plaintext is minimized.