情報通信システム 小テスト第2回 解答

【問 1】4 個の赤玉と 1 個の白玉が入っているつぼから無作為に連続して 2 個の玉を取り出すという非復元抽出を考える．最初に取り出した玉の色を X_1，2 回目に取り出した玉の色を X_2 とするとき，$H(X_1)$，$H(X_2)$，$H(X_1X_2)$，$H(X_1|X_2)$，$I(X_1;X_2)$ を求めよ．(20 点)

（解答） 赤を R，白を W で表す．確率事象系 X_1, X_2 の確率分布は以下のようになる．

$$
\begin{gathered}
P(X_1\!=\!R) = 4/5,\ \ P(X_1\!=\!W) = 1/5 \\
P(X_2\!=\!R) = 4/5 \cdot 3/4 + 1/5 \cdot 4/4 = 4/5 \\
P(X_2\!=\!W) = 4/5 \cdot 1/4 + 1/5 \cdot 0/4 = 1/5
\end{gathered}
$$

したがって，

$$
H(X_1) = H(X_2) = \mathcal{H}(1/5) = -(1/5)\log(1/5) - (4/5)\log(4/5) = \log 5 - \frac{8}{5}
$$

また結合確率は以下の通りである．

$$
\begin{aligned}
&P(X_1\!=\!R \wedge X_2\!=\!R) = 4/5 \times 3/4 = 3/5, \\
&P(X_1\!=\!R \wedge X_2\!=\!W) = P(X_1\!=\!W \wedge X_2\!=\!R) = 1/5, \\
&P(X_1\!=\!W \wedge X_2\!=\!W) = 0
\end{aligned}
$$

したがって，

$$
\begin{aligned}
H(X_1X_2) &= -(3/5)\log(3/5) - (1/5)\log(1/5) - (1/5)\log(1/5) = \log 5 - \frac{3}{5}\log 3 \\
H(X_1|X_2) &= H(X_1X_2) - H(X_2) = \frac{8}{5} - \frac{3}{5}\log 3 \\
I(X_1;X_2) &= H(X_1) - H(X_1|X_2) = \log 5 - \frac{8}{5} - \left(\frac{8}{5} - \frac{3}{5}\log 3\right) = -\frac{16}{5} + \frac{3}{5}\log 3 + \log 5
\end{aligned}
$$

（別解） 条件付確率は以下の通りである．

$$
\begin{gathered}
P(X_1\!=\!R|X_2\!=\!R) = 3/4,\ \ P(X_1\!=\!R|X_2\!=\!W) = 1, \\
P(X_1\!=\!W|X_2\!=\!R) = 1/4,\ \ P(X_1\!=\!W|X_2\!=\!W) = 0
\end{gathered}
$$

よって，条件付エントロピー $H(X_1|X_2)$ は以下のように計算できる．

$$
H(X_1|X_2) = -(3/5)\log 3/4 - (1/5)\log 1 - (1/5)\log 1/4 = \frac{8}{5} - \frac{3}{5}\log 3
$$

また，相互情報量 $I(X_1;X_2)$ は以下の方法でも計算できる．

$$
\begin{aligned}
I(X_1;X_2) &= H(X_1) + H(X_2) - H(X_1X_2) \\
&= 2 \times (\log 5 - 8/5) - (\log 5 - (3/5)\log 3) \\
&= -\frac{16}{5} + \frac{3}{5}\log 3 + \log 5
\end{aligned}
$$

【問 2】ビット誤り率 $p\,(< 0.5)$ の 2 元対称通信路 BSC とその通信路符号化について，以下の問いに答えよ．(20 点)

(1) 入力 $\{0,1\}$ の生起確率をそれぞれ $w, 1-w$ とするとき，通信路 BSC の伝達情報量を求めよ．(6 点)

（解答） $0, 1$ が出力される確率をそれぞれ q_1, q_2 とすると，$q_1 = w(1-p)+(1-w)p$，$q_2 = wp+(1-w)(1-p)$ となる．これらから，

$$
\begin{aligned}
H(B) &= -q_1\log q_1 - q_2\log q_2 \;=\; \mathcal{H}(wp+(1-w)(1-p)) \\
H(B|A) &= -w(p_{11}\log p_{11}+p_{12}\log p_{12})-(1-w)(p_{21}\log p_{21}+p_{22}\log p_{22}) \\
&= -w(1-p)\log(1-p)-wp\log p-(1-w)p\log p \\
&\quad -(1-w)(1-p)\log(1-p) \\
&= \mathcal{H}(p)
\end{aligned}
$$

よって，通信路 BSC の伝達情報量 $I(A;B)$ は以下のように求められる．

$$
I(A;B) = H(B)-H(B|A) = \mathcal{H}(wp+(1-w)(1-p))-\mathcal{H}(p)
$$

(2) (1) の結果を用いて，通信路 BSC の通信路容量 C を求めよ．(4 点)

（解答） w について $I(A;B)$ が最大になるのは，$\mathcal{H}(wp+(1-w)(1-p))$ が最大のときである．$\mathcal{H}(wp+(1-w)(1-p))$ は，$wp+(1-w)(1-p)=\frac{1}{2}$ となる w が存在すれば，そのときに最大値 $\mathcal{H}(1/2)=1$ を取る．$wp+(1-w)(1-p)=\frac{1}{2}$ を整理すると $(w-\frac{1}{2})(2p-1)=0$ となり，$w=1/2$ のときに $wp+(1-w)(1-p)=\frac{1}{2}$ が成立する．したがって，$C = 1-\mathcal{H}(p)$ となる．

(3) $p=0.20$ のとき，$(n,2)$ 符号（2 個の情報記号に $n-2$ 個の検査記号を付加して符号長を n にした組織符号）を用いていくらでも誤り確率が小さくなるような符号化を行いたい．そのような符号化が存在する最小の整数値 n を求めよ．必要ならば，$\log_2 3 = 1.58$，$\log_2 5 = 2.32$ を用いよ．(4 点)

（解答） 通信路符号化定理より，$R<C$ のときにはいくらでも復号誤り確率が小さくなるような符号化が存在する．$p=0.20$ のとき，この通信路の通信路容量 $C = 1-\mathcal{H}(\frac{1}{5}) = 1+\frac{1}{5}\log_2\frac{1}{5}+\frac{4}{5}\log_2\frac{4}{5} = \frac{13}{5}-\log_2 5 = 0.28$ である．一方，$(n,2)$ 符号の情報速度は $R=\frac{2}{n}$ である．よって，$R=\frac{2}{n}<0.28(=C)$ を満たす最小の整数 n を求めればよい．そのような整数は $n=8$ である．

(4) 通信路 BSC において，以下の符号 $\boldsymbol{C}$ を用いたときの復号誤り確率 P_e を求めよ．ただし，符号語は等確率で生起するものとし，最尤復号を実行するものとする．(6 点)

情報系列 $\boldsymbol{i}$	符号 $\boldsymbol{C}$
00	000
01	011
10	101
11	110

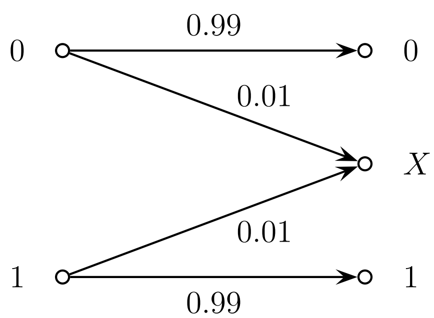
（解答） 各符号語に対して，以下が成立する．

- 誤りが 0 個のときには，必ず正しく復号されるので，$P_e = 0$.
- 誤りが 1 個のときには，必ず誤りが生じたことが検出される．ここですべての符号語（000, 011, 101, 110）から受信語 $\boldsymbol{y}$ が生じる確率を計算する．すると，それらのうち 3 個の符号語からは 1 個の誤りが生じることで $\boldsymbol{y}$ になり，残りの 1 個の符号語からは 3 個の誤りが生じることで $\boldsymbol{y}$ になる．1 個の誤りが生じる確率 $p(1-p)^2$ と 3 個の誤りが生じる確率 p^3 を比較すると，$p < 1/2$ より $p(1-p)^2 > p^3$ となる．よって，最尤復号法では 3 個の符号語のいずれかに確率 1/3 で復号することになる．符号語が等確率（$= 1/4$）に生起することを考えあわせると，以上から $(2/3) \times 3p(1-p)^2$ の確率で復号誤りが生じる．したがって，$P_e = 2p(1-p)^2$.
- 誤りが 2 個のときには，必ず誤った復号になる（誤りが検出できない）．誤りが 2 個だけ生じる確率は $3p^2(1-p)$ なので，$P_e = 3p^2(1-p)$.
- 誤りが 3 個のときには，必ず誤りが生じたことが検出され，誤りが 1 個のときの説明から最尤復号法では必ず復号誤りとなる．したがって，$P_e = (1-p)^3$.

符号語が等確率に生起するので，以上から復号誤り確率 P_e は以下のように求まる．

$$P_e = 2p(1-p)^2 + 3p^2(1-p) + p^3 = p(2-p)$$

【問 3】以下の通信路線図で示される通信路の通信路容量を求めよ．(10 点)



（解答） 一様性を利用しない場合：

入力 $\{0,1\}$ の生起確率をそれぞれ $w, 1-w$ として，まず伝達情報量を求める．0, X, 1 が出力される確率をそれぞれ q_1, q_2, q_3 とすると，$q_1 = 0.99w$，$q_2 = 0.01$，$q_3 = 0.99(1-w)$ となる．これらから，

$$\begin{aligned}
H(B) &= -q_1 \log q_1 - q_2 \log q_2 - q_3 \log q_3 \\
&= -0.99w \log 0.99w - 0.01 \log 0.01 - 0.99(1-w) \log 0.99(1-w) \\
&= -0.99w \log w - 0.99(1-w) \log(1-w) - 0.99w \log 0.99 - 0.99(1-w) \log 0.99 \\
&\quad -0.01 \log 0.01 \\
&= 0.99\mathcal{H}(w) + \mathcal{H}(0.99)
\end{aligned}$$

$$
\begin{aligned}
H(B|A) &= -p_1(p_{11}\log p_{11} + p_{12}\log p_{12} + p_{13}\log p_{13}) \\
&\quad -p_2(p_{21}\log p_{21} + p_{22}\log p_{22} + p_{23}\log p_{23}) \\
&= -w\{0.99\log 0.99 + 0.01\log 0.01\} - (1-w)\{0.99\log 0.99 + 0.01\log 0.01\} \\
&= \mathcal{H}(0.99)
\end{aligned}
$$

よって，伝達情報量は $I(A;B) = H(B) - H(B|A) = 0.99\mathcal{H}(w)$ となる．w に関して $I(A;B)$ が最大となるのは，$w = 1/2$ のときである．このとき $\mathcal{H}(w) = 1$ より，$C = \max_w I(A;B) = 0.99$ となる．

（別解） 一様性を利用する場合：
この通信路は入力に対して一様であるので，この通信路容量は以下の式で与えられる．

$$
\begin{aligned}
C &= \max_w\{H(B)\} + \sum_{j=1}^{3} p_{1j}\log p_{1j} \\
&= \max_w\{-q_1\log q_1 - q_2\log q_2 - q_3\log q_3\} + 0.99\log 0.99 + 0.01\log 0.01 \\
&= \max_w\{0.99\mathcal{H}(w) + \mathcal{H}(0.99)\} - \mathcal{H}(0.99)
\end{aligned}
$$

w に関して $H(B)$ が最大となるのは $w = 1/2$ のときであり，このとき $H(B) = 0.99 + \mathcal{H}(0.99)$ となる．したがって，$C = 0.99 + \mathcal{H}(0.99) - \mathcal{H}(0.99) = 0.99$ となる．