情報理論 配布資料 #3

担当：井上 純一 (情報科学研究科棟 8-13)
URL：http://chaosweb.complex.eng.hokudai.ac.jp/~j_inoue/

平成 17 年 5 月 9 日

目 次

演習問題 2 の解答例

1. A 氏, B 氏, C 氏にとってはサイコロの目の出現確率が異なり，これが各自のエントロピー (出るサイコロの目に対する知識のあいまいさ) に反映する．
まず, A 氏にとっては P(目の数) でそれぞれの目の出る確率とするならば, $P(1)=1/3, P(2)=0, P(3)=1/3, P(4)=0, P(5)=1/3, P(6)=0$ であるから, A 氏のエントロピー H_A は

$$H_A = -\sum_{i=1}^{6} P(i)\log P(i) = -\frac{1}{3}\log\frac{1}{3}-\frac{1}{3}\log\frac{1}{3}-\frac{1}{3}\log\frac{1}{3} = \log 3(\text{ビット}) \tag{39}$$

である．これはサイコロの目に関する事前知識が無い場合のエントロピー $\{-(1/6)\log(1/6)\}\times 6=\log 6$ よりも小さいことに注意されたい．「サイコロの目は奇数である」ということを事前知識として持ち合わせていることにより, A 氏のサイコロの目の値に関する知識のあいまいさは $\log 6-\log 3=\log 2=1$(ビット) だけ減少するのである．
B 氏にとって，サイコロの各目が出る確率は $P(1)=0, P(2)=0, P(3)=1/2, P(4)=0, P(5)=1/2, P(6)=0$ であるから, B 氏のエントロピー H_B は A 氏の場合と同様にして

$$H_B = -\frac{1}{2}\log\frac{1}{2}-\frac{1}{2}\log\frac{1}{2} = \log 2 = 1(\text{ビット}) \tag{40}$$

である．従って，「3 か 5 のいずれかの目が出る」という事実を事前知識としてもっている場合には, 何も知らない場合と比べてサイコロの目に関する知識のあいまいさが $\log 6-\log 2=\log 3$(ビット) だけ減少する．
最後に C 氏にとって, サイコロの各目が出る確率は $P(1)=P(2)=P(4)=P(5)=P(6)=0, P(3)=1$ であるから, C 氏のエントロピー H_C は直ちに

$$H_C = 0 \tag{41}$$

であり, エントロピーはゼロであり, 当然, サイコロの目の出方に関する知識のあいまいさはゼロである．従って, エントロピーの大小関係は A 氏, B 氏, C 氏のそれぞれが持つ知識のあいまいさを反映し, $H_A > H_B > H_C = 0$ となる．

2. 1 回目には必ず表が出るコイン投げを考える分けであるから, $\mathsf{H}=$ 表, $\mathsf{T}=$ 裏 のように表記することにし, 事象の集合として 1 回目に出るコインの向き, 2 回目に出るコインの向きをそれぞれ $\mathcal{A}=\{\mathsf{H},\mathsf{T}\}, \mathcal{B}=\{\mathsf{H},\mathsf{T}\}$ と書くことにすれば, 条件付きエントロピーを算出する際に必要な各確率は

$$P_A(\mathsf{H}) = P_A(\mathsf{T}) = \frac{1}{2} \tag{42}$$

$$P_{B|A}(\mathsf{H}|\mathsf{T}) = P_{B|A}(\mathsf{H}|\mathsf{H}) = 1 \tag{43}$$

$$P_{B|A}(\mathsf{T}|\mathsf{H}) = P_{B|A}(\mathsf{T}|\mathsf{T}) = 0 \tag{44}$$

であるから, 求める条件付きエントロピーは

$$\begin{aligned} H(\mathcal{B}|\mathcal{A}) &= -\sum_{\mathcal{A}=\{\mathsf{H},\mathsf{T}\}}\sum_{\mathcal{B}=\{\mathsf{H},\mathsf{T}\}} P_{B|A}(j|i)P_A(i)\log P_{B|A}(j|i) \\ &= -P_{B|A}(\mathsf{H}|\mathsf{H})P_A(\mathsf{H})\log P_{B|A}(\mathsf{H}|\mathsf{H}) - P_{B|A}(\mathsf{H}|\mathsf{T})P_A(\mathsf{T})\log P_{B|A}(\mathsf{H}|\mathsf{T}) \\ &- P_{B|A}(\mathsf{T}|\mathsf{H})P_A(\mathsf{H})\log P_{B|A}(\mathsf{T}|\mathsf{H}) - P_{B|A}(\mathsf{T}|\mathsf{T})P_A(\mathsf{T})\log P_{B|A}(\mathsf{T}|\mathsf{T}) \\ &= -1\times\frac{1}{2}\log 1 - 1\times\frac{1}{2}\log 1 - 0\times\frac{1}{2}\log 0 - 0\times\frac{1}{2}\log 0 = 0 \end{aligned} \tag{45}$$

となり, 求める条件付きエントロピーはゼロである.

4 相互情報量

前回から各種エントロピーの定義, 及び, その意味について見てきた. そこで, **演習問題 2** の 2. で取り上げた 2 回続けてのコイン投げの例をとって, これを確認しておくと, $\mathcal{A}=\{\mathsf{H},\mathsf{T}\}, \mathcal{B}=\{\mathsf{H},\mathsf{T}\}$ をそれぞれ, 1 回目, 2 回目に出る目とすると, 1 回目に出る目に関する我々の知識のあいまいさを表すのが, 平均情報量であるエントロピー $H(A)$ であり, これは $P_A(\mathsf{H})=P_A(\mathsf{T})=1/2$ であるから, 定義を思い出すと

$$\begin{aligned} H(A) &= -\sum_{i=\mathcal{A}=\{\mathsf{H},\mathsf{T}\}} P_A(i)\log P_A(i) = -P_A(\mathsf{H})\log P_A(\mathsf{H}) - P_A(\mathsf{T})\log P_A(\mathsf{T}) \\ &= -\frac{1}{2}\log\frac{1}{2} - \frac{1}{2}\log\frac{1}{2} = \log 2 = 1(\text{ビット}) \end{aligned} \tag{46}$$

であった.

一方, 条件付きエントロピー $H(A|B)$ は 2 回目に出るコインの向きを知ったときに, 1 回目に出るコインの向きに関する知識のあいまいさを表しており, 条件付き確率が

$$P_{B|A}(\mathsf{H}|\mathsf{H}) = P_{B|A}(\mathsf{H}|\mathsf{T}) = 1 \tag{47}$$

$$P_{B|A}(\mathsf{T}|\mathsf{H}) = P_{B|A}(\mathsf{T}|\mathsf{T}) = 0 \tag{48}$$

であるが, 前回の **演習問題 2** で求めたのは $H(B|A)$ であって, $H(A|B)$ ではなかったことに注意しなければならない. そこで, $H(A|B)$ を定義に従って書き下してみると

$$H(A|B) = -\sum_{i=\mathcal{A}}\sum_{j=\mathcal{B}} P_{A|B}(i|j)P_B(j)\log P_{A|B}(i|j) \tag{49}$$

を求めるためには, 条件付き確率 $P_{A|B}(i|j)$ を求めなければならないことがわかる. 我々が今知っているのは $P_{B|A}(j|i)$ であるから, この確率を用いて必要な $P_{A|B}(i|j)$ を求めなければならない. このとき, 次の公式が成り立つ.

$$P_{A|B}(i|j) = \frac{P_{B|A}(j|i)P_A(i)}{\sum_{i=\mathcal{A}} P_{B|A}(j|i)P_A(i)} \tag{50}$$

(念のための証明)：確率に関する積の公式より

$$P_{A,B}(i,j) \quad = \quad P_{A|B}(i|j)P_B(j) \tag{51}$$

が成り立つが, $P_{A,B}(i,j)$ は $A \leftrightarrow B$, $i \leftrightarrow j$ の交換に関しても不変であるから

$$P_{A,B}(i,j) \quad = \quad P_{B|A}(j|i)P_A(i) \tag{52}$$

も同時に成り立つ. 従って, (51) を $P_{A|B}(i|j)$ について解き, (52) 式を用いて右辺の分子を書き直すと

$$\begin{aligned} P_{A|B}(i|j) &= \frac{P_{A,B}(i,j)}{P_B(j)} \\ &= \frac{P_{B|A}(j|i)P_A(i)}{P_B(j)} \end{aligned} \tag{53}$$

が得られるが, $P_B(j) = \sum_i P_{B|A}(j|i)P_A(i)$ であるから, 結局, ここで挙げた公式 (**ベイズの公式**と呼ばれる)：

$$P_{A|B}(i|j) \quad = \quad \frac{P_{B|A}(j|i)P_A(i)}{\sum_{i=\mathcal{A}} P_{B|A}(j|i)P_A(i)} \tag{54}$$

が成り立つ.
従って, この公式を用いることにより

$$P_{A|B}(\mathsf{H}|\mathsf{T}) \quad = \quad \frac{P_{B|A}(\mathsf{T}|\mathsf{H})P_A(\mathsf{H})}{P_{B|A}(\mathsf{T}|\mathsf{H})P_A(\mathsf{H}) + P_{B|A}(\mathsf{T}|\mathsf{T})P_A(\mathsf{T})} = \frac{0 \times \frac{1}{2}}{0 \times \frac{1}{2} + 0 \times \frac{1}{2}} = 0 \tag{55}$$

$$P_{A|B}(\mathsf{H}|\mathsf{H}) \quad = \quad \frac{P_{B|A}(\mathsf{H}|\mathsf{H})P_A(\mathsf{H})}{P_{B|A}(\mathsf{H}|\mathsf{H})P_A(\mathsf{H}) + P_{B|A}(\mathsf{H}|\mathsf{T})P_A(\mathsf{T})} = \frac{1 \times \frac{1}{2}}{1 \times \frac{1}{2} + 1 \times \frac{1}{2}} = \frac{1}{2} \tag{56}$$

$$P_{A|B}(\mathsf{T}|\mathsf{H}) \quad = \quad \frac{P_{B|A}(\mathsf{H}|\mathsf{T})P_A(\mathsf{T})}{P_{B|A}(\mathsf{H}|\mathsf{H})P_A(\mathsf{H}) + P_{B|A}(\mathsf{H}|\mathsf{T})P_A(\mathsf{T})} = \frac{1 \times \frac{1}{2}}{1 \times \frac{1}{2} + 1 \times \frac{1}{2}} = \frac{1}{2} \tag{57}$$

$$P_{A|B}(\mathsf{T}|\mathsf{T}) \quad = \quad \frac{P_{B|A}(\mathsf{T}|\mathsf{T})P_A(\mathsf{T})}{P_{B|A}(\mathsf{T}|\mathsf{H})P_A(\mathsf{H}) + P_{B|A}(\mathsf{T}|\mathsf{T})P_A(\mathsf{T})} = \frac{0 \times \frac{1}{2}}{0 \times \frac{1}{2} + 0 \times \frac{1}{2}} = 0 \tag{58}$$

となる. 従って, この 2 回のコイン投げでは, 2 回目に必ず表が出るので, $P_B(\mathsf{H}) = 1, P_B(\mathsf{T}) = 0$ であることに注意すれば, ここで求めるべき条件付きエントロピーは

$$\begin{aligned} H(A|B) &= -P_{A|B}(\mathsf{H}|\mathsf{H})P_B(\mathsf{H})\log P_{A|B}(\mathsf{H}|\mathsf{H}) - P_{A|B}(\mathsf{T}|\mathsf{H})P_B(\mathsf{H})\log P_{A|B}(\mathsf{T}|\mathsf{H}) \\ &- P_{A|B}(\mathsf{H}|\mathsf{T})P_B(\mathsf{T})\log P_{A|B}(\mathsf{H}|\mathsf{T}) - P_{A|B}(\mathsf{T}|\mathsf{T})P_B(\mathsf{T})\log P_{A|B}(\mathsf{T}|\mathsf{T}) \\ &= -\frac{1}{2} \times 1 \times \log\frac{1}{2} - \frac{1}{2} \times 1 \times \log\frac{1}{2} = \log 2 = 1(\text{ビット}) \end{aligned} \tag{59}$$

と求めることができる. ところで, エントロピー $H(A)$ から $H(A|B)$ を引いたものは, それぞれの意味あいを考えると, **2 回目に投げたコインの裏表を知ったときに, 1 回目に投げられるコインの向きに関する知識のあいまいさがどの程度減少したか**という量を表している. これを**相互情報量**と呼ぶ. この相互情報量は $I(A;B)$ と表記され, 今の場合には

$$I(A;B) \quad = \quad H(A) - H(A|B) = 1 - 1 = 0 \tag{60}$$

であり, ゼロとなる. これは考えてみれば当たり前である. なぜならば, このコイン投げは 2 回目には必ず表が出るように仕組まれているので, 2 回目に出たコインの向きを知っても, いっこうに 1 回目のコインが表か裏かということに関しての知識は増えないのである.
この相互情報量という考え方と, その算出法になれるためにいくつかの例題を見ておこう.

例題 3

A さんは札幌在住のアマチュア気象予報士であり, 彼独自の方法により明日の天気を予測する. その方法によると, 藻岩山に雲がかかることと明日の天気には相関があるそうで, $\mathcal{A}=\{a_1,a_2\}$, (a_1 : 晴れ, a_2 : 雨), $\mathcal{B}=\{b_1,b_2\}$, (b_1 : 藻岩山に雲なし, b_2 : 藻岩山に雲) とすると, 過去 10 年間にわたる統計により

$$p(a_1)=p(a_2)=\frac{1}{2},\ \ p(b_1|a_1)=\frac{1}{2},\ \ p(b_2|a_2)=1$$

であることがわかった. このとき以下の問いに答えよ.

(1) 次の関係式：

$$\sum_{j=1,2} p(b_j|a_i) \quad = \quad 1$$

が成り立つことを示し, これを用いて $p(b_2|a_1), p(b_1|a_2)$ を求めよ.

(2) 「藻岩山にかかる雲」を確認することにより, 明日の天気に関してもたらされる相互情報量：

$$I(\mathcal{A};\mathcal{B}) \quad = \quad H(\mathcal{A})-H(\mathcal{A}|\mathcal{B})$$

を求めよ. ただし, 必要であれば次の公式：

$$P(a_i|b_j) \quad = \quad \frac{p(b_j|a_i)p(a_i)}{\sum_{k=1,2} p(b_j|a_k)p(a_k)}=\frac{p(b_j|a_i)p(a_i)}{p(b_j|a_1)p(a_1)+p(b_j|a_2)p(a_2)}$$

を用いても良い.

(解答例)

問題文の誘導に従って相互情報量を求める.

(1) まず

$$\begin{aligned}\sum_{j=1,2} p(b_j|a_i) &= \sum_{j=1,2}\frac{p(b_j,a_i)}{p(a_i)} \\ &= \frac{\sum_{j=1,2} p(a_i|b_j)p(b_j)}{p(a_i)}=\frac{p(a_i)}{p(a_i)}=1 \end{aligned} \tag{61}$$

となり, 与えられた関係式の成立が示される. 従って, これを用いれば

$$p(b_2|a_1) \quad = \quad 1-p(b_1|a_1)=\frac{1}{2} \tag{62}$$

$$p(b_1|a_2) \quad = \quad 1-p(b_2|a_2)=0 \tag{63}$$

が直ちに得られる.

(2) エントロピー $H(\mathcal{A})$, 条件付きエントロピー $H(\mathcal{A}|\mathcal{B})$ を計算しよう.

$$H(\mathcal{A}) \quad = \quad -\sum_{j=1,2} p(a_j)\log p(a_j) = -\frac{1}{2}\log\frac{1}{2}-\frac{1}{2}\log\frac{1}{2}=1 \tag{64}$$

$$H(\mathcal{A}|\mathcal{B}) \quad = \quad -\sum_{i=1,2}\sum_{j=1,2} p(a_i,b_j)\log p(a_i|b_j)$$

$$
\begin{aligned}
&= -\sum_{i=1,2}\sum_{j=1,2} p(a_i|b_j)p(b_j)\log p(a_i|b_j) \\
&= -p(a_1|b_1)p(b_1)\log p(a_1|b_1) - p(a_2|b_1)p(b_1)\log p(a_2|b_1) \\
&- p(a_1|b_2)p(b_2)\log p(a_1|b_2) - p(a_2|b_2)p(b_2)\log p(a_2|b_2) \qquad (65)
\end{aligned}
$$

さて, ここから先の計算を進めるためには, 確率 $p(a_1|b_2), p(a_1|b_2)$ 等が必要であるが, これは問題文に与えたヒントの公式より

$$
\begin{aligned}
p(a_1|b_1) &= \frac{p(b_1|a_1)p(a_1)}{\sum_{j=1,2} p(b_1|a_j)p(a_j)} = \frac{p(b_1|a_1)p(a_1)}{p(b_1|a_1 p(a_1) + p(b_1|a_2)p(a_2)} = 1 \qquad (66) \\
p(a_2|b_1) &= \frac{p(b_1|a_2)p(a_2)}{\sum_{j=1,2} p(b_1|a_j)p(a_j)} = 0 \qquad (67) \\
p(a_1|b_2) &= \frac{p(b_2|a_1)p(a_1)}{\sum_{j=1,2} p(b_2|a_j)p(a_j)} = \frac{p(b_2|a_1)p(a_1)}{p(b_2|a_1)p(a_1) + p(b_2|a_2)p(a_2)} = \frac{1}{3} \qquad (68) \\
p(a_2|b_2) &= \frac{p(b_2|a_2)p(a_2)}{\sum_{j=1,2} p(b_2|a_j)p(a_j)} = \frac{p(b_2|a_2)p(a_2)}{p(b_2|a_1)p(a_1) + p(b_2|a_2)p(a_2)} = \frac{2}{3} \qquad (69)
\end{aligned}
$$

が得られるので, これらを (65) に代入して

$$
H(\mathcal{A}|\mathcal{B}) = -\frac{1}{3}\times\frac{3}{4}\log\frac{1}{3} - \frac{2}{3}\times\frac{3}{4}\log\frac{2}{3} = \frac{3}{4}\log 3 - \frac{1}{2} \qquad (70)
$$

となる. 従って, 相互情報量 $I(\mathcal{A};\mathcal{B})$ は

$$
I(\mathcal{A};\mathcal{B}) = 1 - \left\{\frac{3}{4}\log 3 - \frac{1}{2}\right\} = \frac{3}{2} - \frac{3}{4}\log 3 \simeq 0.31 \ (\text{ビット}) \qquad (71)
$$

となる.

例題 4

例題 1 で考えた 2 元対称消失通信路について以下の問いに答えよ.

(1) 入力値 $\mathcal{A} = \{0,1\}$, 及び出力値 $\mathcal{B} = \{0,1,\mathrm{x}\}$ を確率変数とするとき, $\mathcal{A},\mathcal{B}$ の**エントロピー**：$H(\mathcal{A}), H(\mathcal{B})$, **結合エントロピー**：$H(\mathcal{A},\mathcal{B})$, 及び, **条件付きエントロピー**：$H(\mathcal{A}|\mathcal{B}), H(\mathcal{B}|\mathcal{A})$ を求めよ. ただし, $p_{\mathcal{A}}(1) = p_{\mathcal{A}}(0) = 1/2$ とする.

(2) (1) で求めたエントロピー, 結合エントロピー, 条件付きエントロピーに対して関係式：

$$
H(\mathcal{A},\mathcal{B}) = H(\mathcal{A}) + H(\mathcal{B}|\mathcal{A}) = H(\mathcal{B}) + H(\mathcal{A}|\mathcal{B})
$$

が成立することを確かめよ.

(3) 確率変数 $\mathcal{A},\mathcal{B}$ の**相互情報量**：$I(\mathcal{A};\mathcal{B})$ を求めよ. また, $q = 0$ のとき相互情報量 $I(\mathcal{A};\mathcal{B})$ を最小とするような p の値を求めよ.

(解答例)

(1) エントロピー, 結合エントロピー, 条件付きエントロピーをそれぞれ求めていこう.

$$
H(\mathcal{A}) = -\sum_{j=0,1} p_{\mathcal{A}}(j)\log p_{\mathcal{A}}(j) = -\frac{1}{2}\log\frac{1}{2} - \frac{1}{2}\log\frac{1}{2} = 1 \qquad (72)
$$

出力値の確率が

$$
\begin{aligned}
p_{\mathcal{B}}(0) &= \sum_{k=0,1} p_{\mathcal{B}|\mathcal{A}}(0|k)p_{\mathcal{A}}(k) = p_{\mathcal{B}|\mathcal{A}}(0|0)p_{\mathcal{A}}(0) + p_{\mathcal{B}|\mathcal{A}}(0|1)p_{\mathcal{A}}(1) = \frac{1}{2}(1-q) \quad (73)\\
p_{\mathcal{B}}(1) &= \sum_{k=0,1} p_{\mathcal{B}|\mathcal{A}}(1|k)p_{\mathcal{A}}(k) = p_{\mathcal{B}|\mathcal{A}}(1|0)p_{\mathcal{A}}(0) + p_{\mathcal{B}|\mathcal{A}}(1|1)p_{\mathcal{A}}(1) = \frac{1}{2}(1-q) \quad (74)\\
p_{\mathcal{B}}(\times) &= \sum_{k=0,1} p_{\mathcal{B}|\mathcal{A}}(\times|k)p_{\mathcal{A}}(k) = p_{\mathcal{B}|\mathcal{A}}(\times|0)p_{\mathcal{A}}(0) + p_{\mathcal{B}|\mathcal{A}}(\times|1)p_{\mathcal{A}}(1) = q \quad (75)
\end{aligned}
$$

であることから

$$
H(\mathcal{B}) = -\sum_{j=0,1,\times} p_{\mathcal{B}}(j)\log p_{\mathcal{B}}(j) = 1-q-(1-q)\log(1-q) - q\log q \quad (76)
$$

が得られる. また, 結合エントロピーは

$$
\begin{aligned}
H(\mathcal{A},\mathcal{B}) &= -\sum_i\sum_j p_{\mathcal{A},\mathcal{B}}(i,j)\log p_{\mathcal{A},\mathcal{B}}(i,j)\\
&= -\sum_i\sum_j p_{\mathcal{B}|\mathcal{A}}(i|j)p_{\mathcal{A}}(j)\log p_{\mathcal{B}|\mathcal{A}}(i|j)p_{\mathcal{A}}(j)\\
&= -p_{\mathcal{B}|\mathcal{A}}(0|0)p_{\mathcal{A}}(0)\log p_{\mathcal{B}|\mathcal{A}}(0|0)p_{\mathcal{A}}(0) - p_{\mathcal{B}|\mathcal{A}}(1|0)p_{\mathcal{A}}(0)\log p_{\mathcal{B}|\mathcal{A}}(1|0)p_{\mathcal{A}}(0)\\
&- -p_{\mathcal{A}|\mathcal{B}}(\times|0)p_{\mathcal{A}}(0)\log p_{\mathcal{B}|\mathcal{A}}(\times|0)p_{\mathcal{A}}(0)\\
&- p_{\mathcal{B}|\mathcal{A}}(0|1)p_{\mathcal{A}}(1)\log p_{\mathcal{B}|\mathcal{A}}(0|1)p_{\mathcal{A}}(1) - p_{\mathcal{B}|\mathcal{A}}(1|1)p_{\mathcal{A}}(1)\log p_{\mathcal{B}|\mathcal{A}}(1|1)p_{\mathcal{A}}(1)\\
&- p_{\mathcal{B}|\mathcal{A}}(\times|1)p_{\mathcal{A}}(1)\log p_{\mathcal{B}|\mathcal{A}}(\times|1)p_{\mathcal{A}}(1)\\
&= 1-(1-p-q)\log(1-p-q) - p\log p - q\log q \quad (77)
\end{aligned}
$$

となる. 条件付きエントロピーは

$$
\begin{aligned}
H(\mathcal{B}|\mathcal{A}) &= -\sum_i\sum_j p_{\mathcal{B}|\mathcal{A}}(i|j)p_{\mathcal{A}}(j)\log p_{\mathcal{B}|\mathcal{A}}(i|j)\\
&= -p_{\mathcal{B}|\mathcal{A}}(0|0)p_{\mathcal{A}}(0)\log p_{\mathcal{B}|\mathcal{A}}(0|0) - p_{\mathcal{B}|\mathcal{A}}(1|0)p_{\mathcal{A}}(0)\log p_{\mathcal{B}|\mathcal{A}}(1|0)\\
&- p_{\mathcal{A}|\mathcal{B}}(\times|0)p_{\mathcal{A}}(0)\log p_{\mathcal{B}|\mathcal{A}}(\times|0)\\
&- p_{\mathcal{B}|\mathcal{A}}(0|1)p_{\mathcal{A}}(1)\log p_{\mathcal{B}|\mathcal{A}}(0|1) - p_{\mathcal{B}|\mathcal{A}}(1|1)p_{\mathcal{A}}(1)\log p_{\mathcal{B}|\mathcal{A}}(1|1)\\
&- p_{\mathcal{B}|\mathcal{A}}(\times|1)p_{\mathcal{A}}(1)\log p_{\mathcal{B}|\mathcal{A}}(\times|1)\\
&= -(1-p-q)\log(1-p-q) - p\log p - q\log q \quad (78)
\end{aligned}
$$

である. 従って, これらの結果から

$$
H(\mathcal{A}) + H(\mathcal{B}|\mathcal{A}) = 1-(1-p-q)\log(1-p-q) - p\log p - q\log q = H(\mathcal{A},\mathcal{B}) \quad (79)
$$

が成り立っていることが確かめられる.
一方, 条件付きエントロピー $H(\mathcal{A}|\mathcal{B})$ は

$$
\begin{aligned}
H(\mathcal{A}|\mathcal{B}) &= -\sum_i\sum_j p_{\mathcal{B}|\mathcal{A}}(i|j)p_{\mathcal{A}}(j)\log p_{\mathcal{A}|\mathcal{B}}(j|i)\\
&= -p_{\mathcal{B}|\mathcal{A}}(0|0)p_{\mathcal{A}}(0)\log p_{\mathcal{A}|\mathcal{B}}(0|0) - p_{\mathcal{B}|\mathcal{A}}(1|0)p_{\mathcal{A}}(0)\log p_{\mathcal{A}|\mathcal{B}}(1|0)\\
&- p_{\mathcal{A}|\mathcal{B}}(\times|0)p_{\mathcal{A}}(0)\log p_{\mathcal{A}|\mathcal{B}}(\times|0)\\
&- p_{\mathcal{B}|\mathcal{A}}(0|1)p_{\mathcal{A}}(1)\log p_{\mathcal{A}|\mathcal{B}}(0|1) - p_{\mathcal{B}|\mathcal{A}}(1|1)p_{\mathcal{A}}(1)\log p_{\mathcal{A}|\mathcal{B}}(1|1)\\
&- p_{\mathcal{B}|\mathcal{A}}(\times|1)p_{\mathcal{A}}(1)\log p_{\mathcal{A}|\mathcal{B}}(\times|1) \quad (80)
\end{aligned}
$$

となるが,

$$p_{\mathcal{A}|\mathcal{B}}(0|0) = \frac{p_{\mathcal{B}|\mathcal{A}}(0|0)p_{\mathcal{A}}(0)}{p_{\mathcal{B}|\mathcal{A}}(0|0)p_{\mathcal{A}}(0)+p_{\mathcal{B}|\mathcal{A}}(0|1)p_{\mathcal{A}}(1)} = \frac{1-p-q}{1-q} \tag{81}$$

$$p_{\mathcal{A}|\mathcal{B}}(0|1) = \frac{p_{\mathcal{B}|\mathcal{A}}(1|0)p_{\mathcal{A}}(0)}{p_{\mathcal{B}|\mathcal{A}}(1|0)p_{\mathcal{A}}(0)+p_{\mathcal{B}|\mathcal{A}}(1|1)p_{\mathcal{A}}(1)} = \frac{p}{1-q} \tag{82}$$

$$p_{\mathcal{A}|\mathcal{B}}(0|\times) = \frac{p_{\mathcal{B}|\mathcal{A}}(\times|0)p_{\mathcal{A}}(0)}{p_{\mathcal{B}|\mathcal{A}}(\times|0)p_{\mathcal{A}}(0)+p_{\mathcal{B}|\mathcal{A}}(\times|1)p_{\mathcal{A}}(1)} = \frac{1}{2} \tag{83}$$

$$p_{\mathcal{A}|\mathcal{B}}(1|0) = \frac{p_{\mathcal{B}|\mathcal{A}}(0|1)p_{\mathcal{A}}(1)}{p_{\mathcal{B}|\mathcal{A}}(0|0)p_{\mathcal{A}}(0)+p_{\mathcal{B}|\mathcal{A}}(0|1)p_{\mathcal{A}}(1)} = \frac{p}{1-q} \tag{84}$$

$$p_{\mathcal{A}|\mathcal{B}}(1|1) = \frac{p_{\mathcal{B}|\mathcal{A}}(1|1)p_{\mathcal{A}}(1)}{p_{\mathcal{B}|\mathcal{A}}(1|0)p_{\mathcal{A}}(0)+p_{\mathcal{B}|\mathcal{A}}(1|1)p_{\mathcal{A}}(1)} = \frac{1-p-q}{1-q} \tag{85}$$

$$p_{\mathcal{A}|\mathcal{B}}(1|\times) = \frac{p_{\mathcal{B}|\mathcal{A}}(\times|1)p_{\mathcal{A}}(1)}{p_{\mathcal{B}|\mathcal{A}}(\times|0)p_{\mathcal{A}}(0)+p_{\mathcal{B}|\mathcal{A}}(\times|1)p_{\mathcal{A}}(1)} = \frac{1}{2} \tag{86}$$

であることに注意すると

$$H(\mathcal{A}|\mathcal{B}) = q-(1-p-q)\log(1-p-q)-p\log p+(1-q)\log(1-q) \tag{87}$$

が得られ, 関係式

$$H(\mathcal{B})+H(\mathcal{A}|\mathcal{B}) = 1-(1-p-q)\log(1-p-q)-p\log p-q\log q = H(\mathcal{A},\mathcal{B}) \tag{88}$$

が成り立っていることが確かめられる.

(3) (2) で得られた結果から相互情報量 $I(\mathcal{A};\mathcal{B})$ は

$$\begin{aligned} I(\mathcal{A};\mathcal{B}) &= H(\mathcal{A})-H(\mathcal{A}|\mathcal{B}) \\ &= 1-q+(1-p-q)\log(1-p-q)+p\log p-(1-q)\log(1-q) \end{aligned} \tag{89}$$

となるが, $q=0$ とおけば (これは教科書の 2 元対称通信路の場合)

$$I(\mathcal{A};\mathcal{B}) = 1+(1-p)\log(1-p)+p\log p \tag{90}$$

となり, p に関する極値条件 $(\partial I/\partial p)=0$ から直ちに

$$p = \frac{1}{2} \tag{91}$$

が得られるが, $(\partial^2 I/\partial p^2)|_{p=1/2}=4>0$ であるから, この $p=1/2$ で相互情報量は最小となる. この通信路の相互情報量は「出力値を観測することによって減少する入力値に関する『あいまいさ』の量」であるから, 2 元対称通信路の場合, 確率 1/2 で入力信号が反転してしまえば, 出力値を知ることにより入力値に関してもたらされる情報量は最小 (ゼロ) となる. これは直観と照らし合わせて, もっともらしい. なお, この相互情報量は $p=0,1$ のとき, すなわち, 「全くノイズが無い」「全て反転する」場合につき最大値をとることも容易に見て取れるであろう. 図 9 にここでの相互情報量 I を p,q の関数としてプロットしておこう.

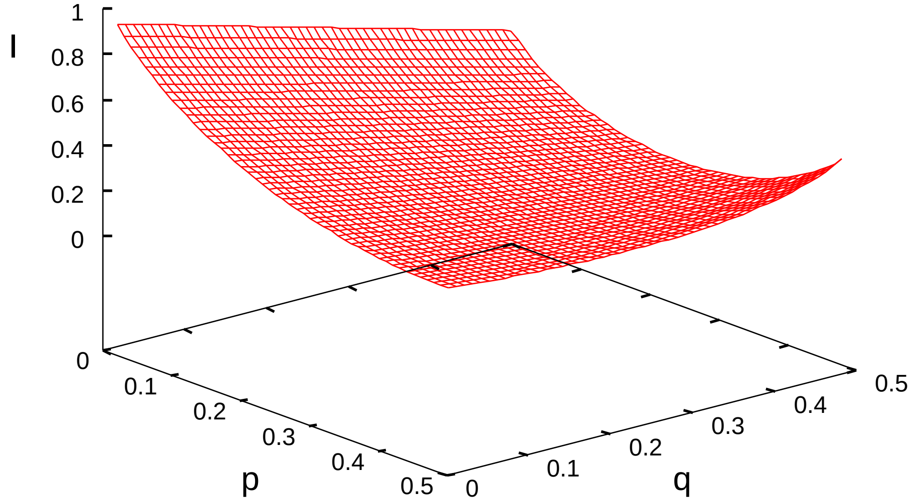

図 9: この問題で扱う通信路の相互情報量 I. p, q はそれぞれ $[0, 0.5]$ の間でプロットしてある. $q=0$ のとき, I は $p=1/2$ で最小値をとる (I は $p=1/2$ の軸に関して対称であることに注意).

演習問題 3

講義で取りあげた 2 回のコイン投げに関し,「2 回目の試行 (コイン投げ) では, 1 回目に現れたコインの向きと反対向きが確率 q で, 同じ向きが $1-q$ の確率で出る」ものとする. このとき以下の問いに答えよ.

(1) このコイン投げは 1 回目の試行でのコインの向き: $A = \{\mathsf{H}, \mathsf{T}\}$ を「入力」, 2 回目の試行でのコインの向き: $B = \{\mathsf{H}, \mathsf{T}\}$ を「出力」とみなしたとき, 既に学んだ **2 元対称通信路** (**配布資料** #**2** 参照) と同様の状態遷移図で表すことができることを実際に状態遷移図を描いて示し, また, 条件付き確率: $P_{B|A}(\mathsf{H}|\mathsf{H}), P_{B|A}(\mathsf{H}|\mathsf{T}), P_{B|A}(\mathsf{T}|\mathsf{H}), P_{B|A}(\mathsf{T}|\mathsf{T})$ をそれぞれ求めよ.

(2) ベイズの公式 (50) を用いることにより, $P_{A|B}(\mathsf{H}|\mathsf{H}), P_{A|B}(\mathsf{H}|\mathsf{T}), P_{A|B}(\mathsf{T}|\mathsf{H}), P_{A|B}(\mathsf{T}|\mathsf{T})$ をそれぞれ求めよ.

(3) $H(A)$, $H(A|B)$ を計算し, 相互情報量: $I(A;B) = H(A) - H(A|B)$ を q の関数として図示せよ. また, $I(A;B)$ を最大/最小にする q をそれぞれ求め, それらの結果が意味するところを簡潔に述べよ.