情報理論　#3

第3回講義　5月9日

情報科学研究科　　井上純一

http://chaosweb.complex.eng.hokudai.ac.jp/~j_inoue/

バイアスのかかったコイン投げ

前回の復習

$$\underbrace{A=\{H,T\}}_{\text{1回目の試行}},\ \underbrace{B=\{H,T\}}_{\text{2回目の試行}}$$

$$\boxed{\begin{aligned} P_{B|A}(H\,|\,H) &= P_{B|A}(H\,|\,T) = 1 \\ P_{B|A}(T\,|\,H) &= P_{B|A}(T\,|\,T) = 0 \end{aligned}}$$

2回目には必ず表がでる

$H(A)$:1回目のコインの向きに関する知識のあいまいさ(エントロピー)

$$H(A) = -\sum_{i=\{H,T\}} P_A(i)\log P_A(i) = \log 2 = 1\ (\text{ビット})$$

$H(B|A)$: 1回目の向きを知ったとき、2回目のコインの向きに関する知識のあいまいさ

$$H(B|A) = -\sum_{i=\{H,T\}}\sum_{j=\{H,T\}} P_{B|A}(j\,|\,i)P_A(i)\log P_{B|A}(j\,|\,i) = 0$$

演習問題2より)

$$H(A|B) = -\sum_{i=\{H,T\}}\sum_{j=\{H,T\}} P_{A|B}(i\,|\,j)P_B(j)\log P_{A|B}(i\,|\,j) \longrightarrow$$

どうなるか?

ベイズの公式

これらは既知

$$P_{A|B}(i\,|\,j) = \frac{P_{B|A}(j\,|\,i)P_A(i)}{\displaystyle\sum_{i=\{H,T\}} P_{B|A}(j\,|\,i)P_A(i)}$$

ベイズの公式

証明)

確率に関する積の公式より $P_{A,B}(i,j) = \underline{P_{A|B}(i\,|\,j)P_B(j)} = \underline{P_{B|A}(j\,|\,i)P_A(i)}$

$A \leftrightarrow B, i \leftrightarrow j$ の交換に対して不変である

従って

$$P_{A|B}(i\,|\,j) = \frac{P_{B|A}(j\,|\,i)P_A(i)}{P_B(j)} = \frac{P_{B|A}(j\,|\,i)P_A(i)}{\sum_{i=\{H,T\}} P_{B|A}(j\,|\,i)P_A(i)}$$

このコイン投げの問題に対する具体的な適用例は配布資料#3 p. 16 (51)-(54)式を見よ

相互情報量

1回目の試行に関する知識のあいまいさ

エントロピーの差　：　$H(A) - H(A|B)$

2回目の試行の結果を知ったとき
の1回目の試行に関する知識の
あいまいさ

2回目の試行の結果を知った場合に1回目の試行に
関する知識のあいまいさがどの程度減少したか（知識の増加分）

$$I(A;B) = H(A) - H(A|B)$$

相互情報量

今のコイン投げの例では

$H(A) = 1, H(A|B) = 1$

$I(A;B) = 0$

2回目には必ず表が出るように
仕組まれているので、2回目の
結果を知っても、1回目の試行に
関する知識は増えない

A、Bの交換に対して不変なので

$$I(A;B) = H(B) - H(B|A)$$

とも書ける

例題3

$A = \{a_1, a_2\}$ (a_1 :**晴れ**, a_2: **雨**), $B = \{b_1, b_2\}$ (b_1 :**藻岩山に雲なし**, b_2 :**藻岩山に雲**)

$$P(a_1) = P(a_2) = 1/2, \; P(b_1 \mid a_1) = 1/2, \; P(b_2 \mid a_2) = 1$$

晴れのときには確率1/2で藻岩山に
雲がかかる

雨のときには確率1で必ず
藻岩山には雲がかかる

問題：相互情報量を求めよ

$$I(A;B) = H(A) - \underline{H(A \mid B)}$$

藻岩山に雲がかかっているか、否かを知ったときに
明日の天気に関する知識のあいまいさがどの程度減少
するのか(知識の増加分）

$$H(A \mid B) = -\sum_{i=1,2} \sum_{j=1,2} P(a_i \mid b_j) P(b_j) \log P(a_i \mid b_j)$$

を具体的に計算する

$$I(A;B) = 0.31 \text{ (ビット)}$$

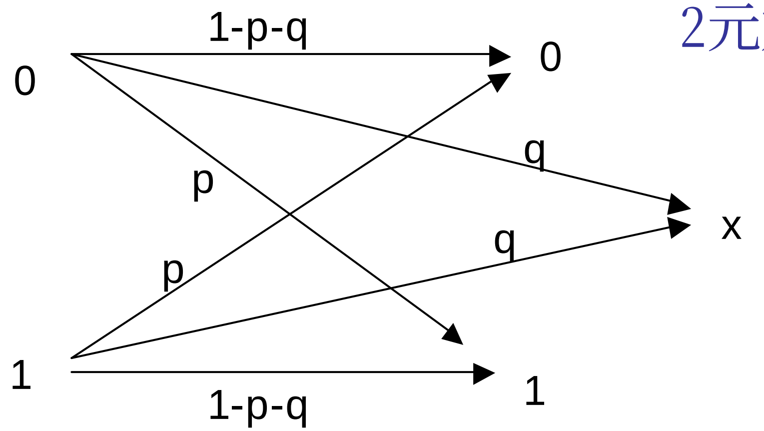
具体的な計算は配布資p.17-18

例題4　#1

2元対称消失通信路

$$p_A(0) = p_A(1) = 1/2$$

(ランダム入力)



出力値の確率:

$$p_B(0),\ p_B(1),\ p_B(x)$$

$$p_B(0) = \sum_{k=0,1} p_{B|A}(0\,|\,k)\,p_A(k) = p_{B|A}(0\,|\,0)\,p_A(0) + p_{B|A}(0\,|\,1)\,p_A(1) = \frac{1}{2}(1-q)$$

同様に

$$p_B(1) = \frac{1}{2}(1-q),\ p_B(x) = q$$

エントロピーは

$$H(A) = -\sum_{j=0,1} p_A(j)\log p_A(j) = 1$$

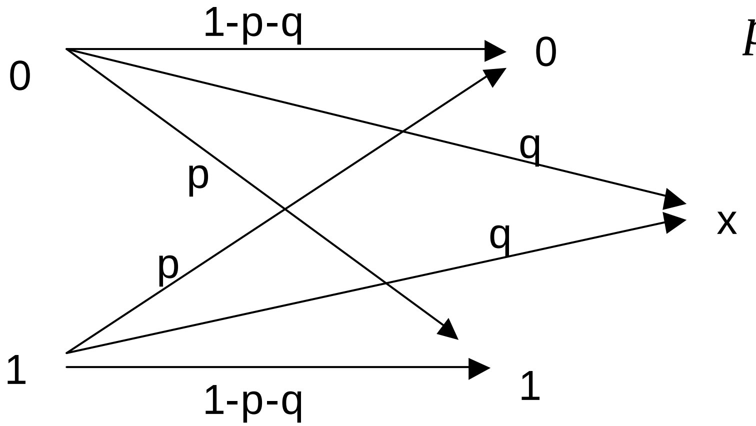
$$H(B) = -\sum_{j=0,1,x} p_B(j)\log p_B(j) = 1 - q - (1-q)\log(1-q) - q\log q$$

例題4　#2

2元対称消失通信路

$$p_A(0) = p_A(1) = \tfrac{1}{2}$$

(ランダム入力)



$$p_B(0) = \frac{1}{2}(1-q)$$

$$p_B(x) = q$$

$$p_B(1) = \frac{1}{2}(1-q)$$

結合エントロピー

$$H(A,B) = -\sum_{i=0,1,x}\sum_{j=0,1} p_{B|A}(i\,|\,j)\,p_A(j)\log p_{B|A}(i\,|\,j)\,p_A(j)$$

$$= 1-(1-p-q)\log(1-p-q) - p\log p - q\log q$$

条件付きエントロピー

$$H(B\,|\,A) = -\sum_{i=0,1,x}\sum_{j=0,1} p_{B|A}(i\,|\,j)\,p_A(j)\log p_{B|A}(i\,|\,j) = -(1-p-q)\log(1-p-q) - p\log p - q\log q$$

$$\boxed{H(A,B) - H(B\,|\,A) = H(A)}$$

エントロピーのチェイン則

例題4　#3

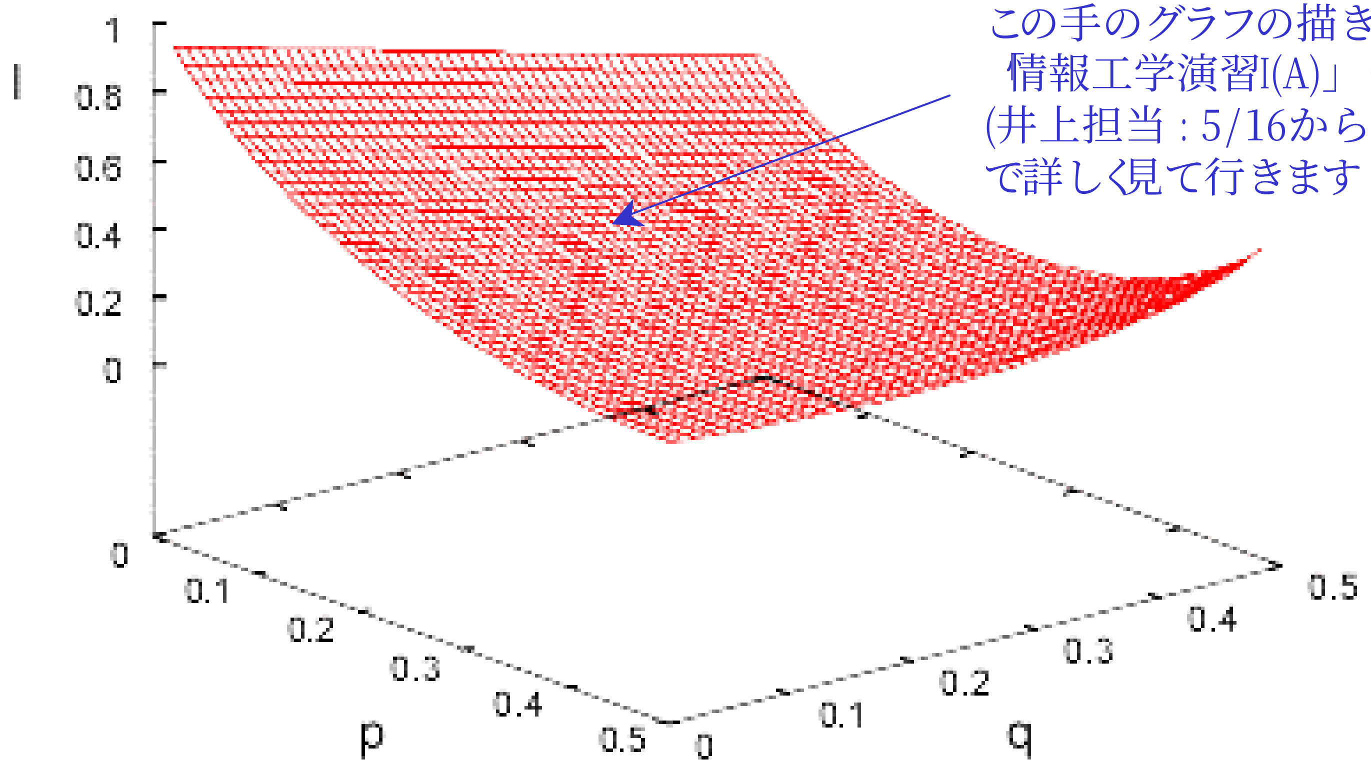
エントロピーのチェイン則　$H(A,B)-H(B|A)=H(A)$　は一般的に示せる

$$
\begin{aligned}
H(A,B) &= -\sum_{i=0,1,x}\sum_{j=0,1} p_{B|A}(i|j)p_A(j)\log p_{B|A}(i|j)p_A(j) \\
&= H(B|A) - \sum_{i=0,1,x}\sum_{j=0,1} p_{B|A}(i|j)p_A(j)\log p_A(j) \\
&= H(B|A) - \sum_{i=0,1,x}\sum_{j=0,1} p_{A|B}(j|i)p_B(i)\log p_A(j) \\
&= H(B|A) - \sum_{j=0,1}\left\{\sum_{i=0,1,x} p_{A|B}(j|i)p_B(i)\right\}\log p_A(j) \\
&= H(B|A) + H(A)
\end{aligned}
$$

例題4 #4

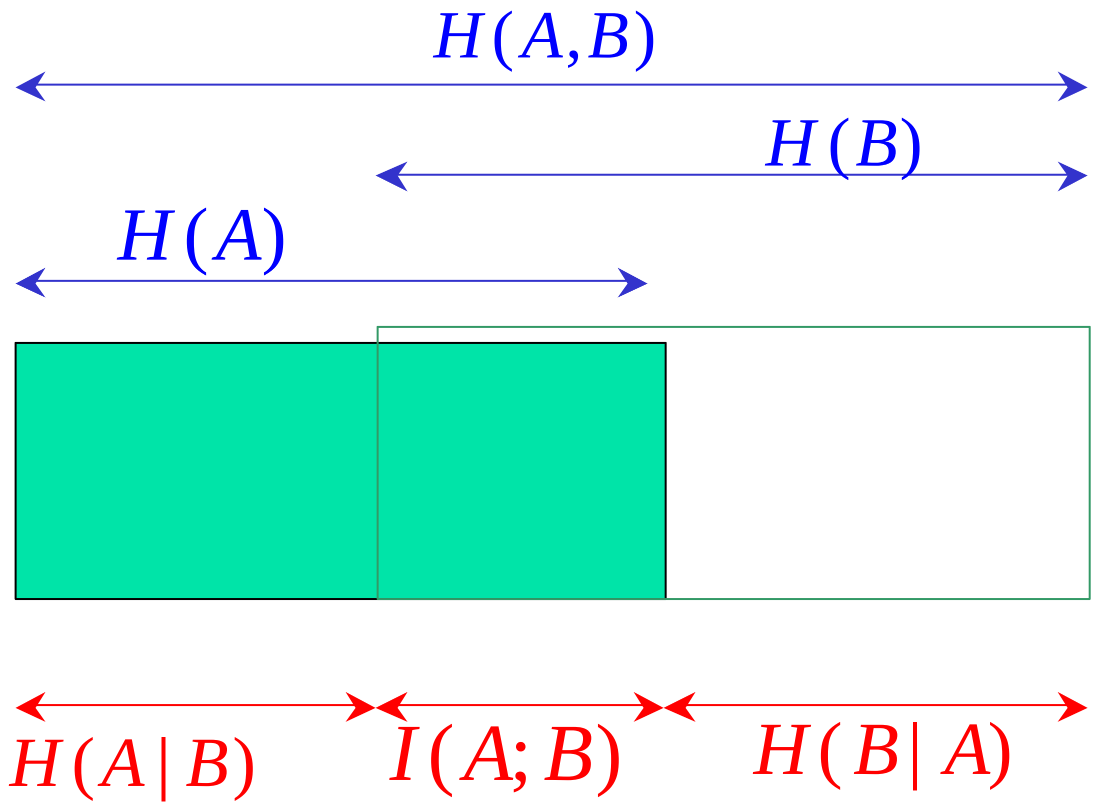
相互情報量

$$I(A;B) = H(A) - H(A|B) = 1 - q + (1 - p - q)\log(1 - p - q) + p\log p - (1-q)\log(1-q)$$



この手のグラフの描き方は
情報工学演習I(A)」(ファイル操作)
(井上担当 : 5/16から3週間)
で詳しく見て行きます

今週のまとめ



各エントロピー、相互情報量の間の関係