情報理論 第3回
– 情報量・エントロピー (2) –
(教科書: 4章・5章)

野崎 隆之

今日の目的と流れ

概要

ある事象から別の事象に関してどれだけ情報をえることができるか？
(ex: 今日の天気の情報から明日の天気の情報をどれだけ得ることができるか？)
$\Rightarrow$ 相互情報量

今日の目標

- 2 個の確率変数のエントロピーを理解し，計算できる
- 以下を計算・説明できる
  - 同時エントロピー
  - 条件付きエントロピー
  - 相互情報量

今日の流れ

1. 確率の復習
2. 同時エントロピー
3. 条件付きエントロピー
4. 相互情報量

1. 確率論の復習 (1: 同時分布)

2 つの確率変数 X, Y に対して，同時確率を以下のように書く

$$P_{X,Y}(x, y)$$

(例) ある地方の天気

天気：晴 or 雨
X: ある日の天気についての確率変数 ($X = 1$ 晴，$X = 2$ 雨)
Y: 次の日の天気についての確率変数 ($Y = 1$ 晴，$Y = 2$ 雨)

		次の日の天気 Y	
		1	2
ある日の天気 X	1	$0.60 = 12/20$	$0.05 = 1/20$
	2	$0.10 = 2/20$	$0.25 = 5/20$

1. 確率論の復習 (2: 周辺化)

(復習) 周辺化

$$P_X(x) = \sum_y P_{X,Y}(x,y), \qquad P_Y(y) = \sum_x P_{X,Y}(x,y),$$

(例・例題) 周辺化

$$P_X(1) = P_{X,Y}(1,1) + P_{X,Y}(1,2) = 0.60 + 0.05 = 0.65$$

$P_X(2) =$ **0.10+0.25**

$P_Y(1) =$ **0.60+0.10**

$P_Y(2) =$ **0.05+0.25**

		次の日の天気 Y	
		1	2
ある日の天気 X	1	$0.60 = 12/20$	$0.05 = 1/20$
	2	$0.10 = 2/20$	$0.25 = 5/20$

1. 確率論の復習 (3: 条件付き確率)

(復習) 条件付き確率

$Y=y$ を知った状況下での $X=x$ になる確率

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

(注意)

- $\sum_x P_{X|Y}(x|y)=1$
- X,Y が独立ならば，$P_{Y|X}(y|x)=P_Y(y), P_{X|Y}(x|y)=P_X(x)$

(例・例題) 条件付き確率

$$P_{Y|X}(1|1)=P_{X,Y}(1,1)/P_X(1)=0.60/0.65=12/13$$

$P_{Y|X}(2|1)=$ **0.05/0.65=1/13**

$P_{X|Y}(2|1)=$ **0.10/0.70**

$P_{X|Y}(2|2)=$ **0.25/0.30**

2. 同時エントロピー (1:定義)

確率変数 (X,Y) が (x,y) であるときの情報量は

$$-\log p_{X,Y}(x,y)$$

同時エントロピー

確率変数 X,Y に対する同時エントロピーは

$$\begin{aligned} H(X,Y) &= \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p_{X,Y}(x,y)(-\log p_{X,Y}(x,y)) \\ &= -\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p_{X,Y}(x,y)\log p_{X,Y}(x,y) \end{aligned}$$

意味： X,Y の曖昧さ

(c.f.) エントロピーの定義は

$$H(X) = -\sum_{x} p_X(x)\log p_X(x)$$

2. 同時エントロピー (1:性質)

(補題 1) 同時エントロピーの上限と下限

$$0 \le H(X,Y) \le \log |\mathcal{X}||\mathcal{Y}|$$

左の不等号の等号成立条件はある $(x,y) \in \mathcal{X} \times \mathcal{Y}$ に対して $p_{X,Y}(x,y) = 1$.
右の不等号の等号成立条件は任意の $(x,y) \in \mathcal{X} \times \mathcal{Y}$ に対して $p_{X,Y}(x,y) = 1/|\mathcal{X}||\mathcal{Y}|$.

証明は 1 つの確率変数の場合と同様

3. 条件付きエントロピー (1: 導入)

- 今日の天気が "晴" という情報を受け取ったときに，明日の天気にはどれだけ曖昧さが残るのか？

(定義) $X = x$ を取る条件下の Y の条件付きエントロピー

$$H(Y|X=x) := -\sum_{y} P_{Y|X}(y|x) \log P_{Y|X}(y|x) \quad \text{[bit]}$$

(例・例題)

さっきの天気の問題

$$\begin{aligned} H(Y|X=1) &= -P_{Y|X}(1|1)\log P_{Y|X}(1|1) \\ &\quad - P_{Y|X}(2|1)\log P_{Y|X}(2|1) \\ &= -\tfrac{12}{13}\log\tfrac{12}{13} - \tfrac{1}{13}\log\tfrac{1}{13} \end{aligned}$$

$H(Y|X=2) =$ **-PYIX(1I2)log PYIX(1I2)**
-PYIX(2I2)logPYIX(2I2)

$H(X|Y=1) =$

$H(X|Y=2) =$

2.1 条件付きエントロピー (2: 定義)

- 天気予報 X を知ったときに，
  実際の天気 Y にはどれだけ曖昧さが残るのか？

(定義) 条件付きエントロピー

$$
\begin{aligned}
H(Y|X) &:= \sum_x P_X(x) H(Y|X=x) \\
&= -\sum_x \sum_y P_X(x) P_{Y|X}(y|x) \log P_{Y|X}(y|x) \\
&= -\sum_x \sum_y P_{X,Y}(x,y) \log P_{Y|X}(y|x) \qquad \text{[bit]}
\end{aligned}
$$

最悪これだけでも覚える

$$
H(Y|X=x) = -\sum_y P_{Y|X}(y|x) \log P_{Y|X}(y|x)
$$

$$
H(Y|X) = -\sum_x \sum_y P_{X,Y}(x,y) \log P_{Y|X}(y|x)
$$

2.1 条件付きエントロピー (3: 例題)

(例・例題) 条件付きエントロピー

$$
\begin{aligned}
H(Y|X) =& P_X(1)H(Y|X=1) + P_X(2)H(Y|X=2) \\
=& \frac{13}{20}\left[-\frac{12}{13}\log\frac{12}{13} - \frac{1}{13}\log\frac{1}{13}\right] + \frac{7}{20}\left[-\frac{2}{7}\log\frac{2}{7} - \frac{5}{7}\log\frac{5}{7}\right] \\
=& -\frac{12}{20}\log\frac{12}{13} - \frac{1}{20}\log\frac{1}{13} - \frac{2}{20}\log\frac{2}{7} - \frac{5}{20}\log\frac{5}{7} \\
H(X|Y) =&
\end{aligned}
$$

注意: $H(X|Y) \neq H(Y|X)$

2.1 条件付きエントロピー (4: 性質)

(補題 2) エントロピーのチェイン則

$$H(X,Y) = H(X|Y) + H(Y)$$

証明の概略： 定義にしたがって右辺を展開し，まとめれば良い

$$\begin{aligned} H(Y) &= -\sum_{y} P_Y(y) \log P_Y(y) \\ &= -\sum_{x}\sum_{y} P_{X,Y}(x,y) \log P_Y(y) \qquad \text{(周辺化の逆)} \end{aligned}$$

2.1 条件付きエントロピー (5: 性質 2)

(補題 2) 条件付きエントロピーの上限と下限

$$0 \le H(X|Y) \le H(X) \tag{1}$$

左の不等号の等号成立条件は $\forall y \exists x$ s.t. $P_{X|Y}(x|y) = 1$
右の不等号の等号成立条件は $P_{X|Y}(x|y) = P_X(x)$ (X, Y が独立)

- (左側) Y から X を一意に知ること ($\forall y \exists x$ s.t. $P_{X|Y}(x|y) = 1$) ができれば，X に曖昧さは残らない
- (右側) X と Y が独立であれば，Y を知っても X の曖昧さはそのまま

証明の概略： 第 2 回補題 1 の証明法とほぼ同様

(c.f.) エントロピーの上限と下限 (第 2 回補題 1)

$$0 \le H(X) \le \log k \tag{2}$$

左の不等号の等号成立条件はある i に対して $p_i = 1$.
右の不等号の等号成立条件は $p_1 = p_2 = \cdots = p_k = 1/k$.

2.2 相互情報量 (1: 定義)

- 天気予報 X を知った時に実際の天気 Y の曖昧さはどれだけ減少したのか?
- 受信結果 Y を知った時に送信情報 X の曖昧さはどれだけ減少したのか?

(定義) 相互情報量

$$I(X;Y) := H(X) - H(X|Y) \tag{3}$$

$$= [X \text{ のもつ曖昧さ}] - [Y \text{ を知った時の } X \text{ のもつ曖昧さ}] \tag{4}$$

意味：

- Y を知った時の X の持つ曖昧さの減少量
- Y から X の情報をどれだけ得られるか (漏れ出るか)？

2.2 相互情報量 (2: 性質)

(補題 3) 相互情報量の別表現

$$I(X;Y) = \sum_x \sum_y P_{X,Y}(x,y) \log \frac{P_{X,Y}(x,y)}{P_X(x) P_Y(y)} \tag{5}$$

証明:

$$\begin{aligned}
I(X;Y) &= H(X) - H(X|Y) && (6)\\
&= -\sum_x \sum_y P_X(x,y) \log P_X(x) + \sum_x \sum_y P_X(x,y) \log P_{X|Y}(x|y) && (7)\\
&= \sum_x \sum_y P_X(x,y) \log \frac{P_{X|Y}(x|y)}{P_X(x)} && (8)\\
&= \sum_x \sum_y P_X(x,y) \log \frac{P_{X,Y}(x,y)}{P_X(x) P_Y(y)} \qquad \text{(条件付き確率の定義)} && (9)
\end{aligned}$$

2.2 相互情報量 (3: 性質 2)

(補題 4) 相互情報量の対称性

$$I(X;Y) = I(Y;X) \tag{10}$$

証明：

$$\begin{aligned}
I(Y;X) &= H(Y) - H(Y|X) \\
&= H(Y) - [H(X,Y) - H(X)] \quad \text{(補題 1)} \\
&= H(X) - [H(X,Y) - H(Y)] \\
&= H(X) - H(X|Y) \quad \text{(補題 1)} \\
&= I(X;Y)
\end{aligned}$$

2.2 相互情報量 (4: 性質 3)

(補題 5) 相互情報量の上限と下限

$$0 \leq I(X;Y) \leq \min\{H(X), H(Y)\} \tag{11}$$

左の不等号の等号成立条件は $P_{X|Y}(x|y) = P_X(x)$ (X, Y が独立)
右の不等号の等号成立条件は $\forall y \exists x$ s.t. $P_{X|Y}(x|y) = 1$

- (左側) どんな情報でも X を知る上で無駄にはならない

証明：
(左側) 補題 2 ($0 \leq H(X|Y) \leq H(X)$) を利用する

$$I(X;Y) = H(X) - H(X|Y) \geq H(X) - H(X) = 0 \tag{12}$$

(右側) 補題 2 と補題 4 ($I(X;Y) = I(Y;X)$) を利用する

$$I(X;Y) = H(X) - H(X|Y) \leq H(X) - 0 = H(X) \tag{13}$$

$$I(X;Y) = I(Y;X) = H(Y) - H(Y|X) \leq H(Y) - 0 = H(Y) \tag{14}$$

以上から $I(X;Y) \leq \min\{H(X), H(Y)\}$

今日のまとめ (1)

(定義) 同時エントロピー

$$H(X,Y) = -\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p_{X,Y}(x,y)\log p_{X,Y}(x,y)$$

(定義) 条件付きエントロピー

$$H(Y|X) = -\sum_{x}\sum_{y} P_{X,Y}(x,y)\log P_{Y|X}(y|x) \qquad \text{[bit]}$$

(定義) 相互情報量

$$I(X;Y) := H(X) - H(X|Y) = \sum_{x}\sum_{y} P_{X,Y}(x,y)\log\frac{P_{X,Y}(x,y)}{P_X(x)P_Y(y)}$$

今日のまとめ (2)

$$
\begin{aligned}
&H(X, Y) = H(X|Y) + H(Y) \\
&0 \leq H(X|Y) \leq H(X) \\
&I(X;Y) = I(Y;X) \\
&0 \leq I(X;Y) \leq \min\{H(X), H(Y)\}
\end{aligned}
$$