

離散数理工学 第 12 回

離散確率論：エントロピー

岡本 吉央

okamotoy@uec.ac.jp

電気通信大学

2025 年 1 月 28 日

最終更新：2025 年 1 月 19 日 21:20

今日の目標

- ▶ エントロピーの基本的な性質を導出できる
 - ▶ 条件つきエントロピーと結合エントロピー
 - ▶ 劣加法性
 - ▶ シアラーの不等式
- ▶ エントロピーを使って、離散数学の基礎事項を証明できる
 - ▶ 二項係数の和の評価
 - ▶ 体積と射影の面積の関係

格言

エントロピーを使って、離散数学の事項を証明できる

目次

- ① エントロピー
- ② エントロピーの劣加法性とシアラーの不等式
- ③ エントロピーの応用：二項係数の和の評価
- ④ エントロピーの応用：体積と射影の面積
- ⑤ 今日のまとめ

エントロピー

整数値確率変数 X

定義：エントロピー

 X の **エントロピー** とは次の量のこと

$$H(X) = \sum_{i \in \mathbb{Z}} \Pr(X = i) \log_2 \frac{1}{\Pr(X = i)}$$

ただし、 $\Pr(X = i) = 0$ のとき、 $\Pr(X = i) \log_2 \frac{1}{\Pr(X = i)} = 0$ とする

例：次の確率変数 X を考える

$$\Pr(X = 1) = \frac{1}{2}, \quad \Pr(X = 2) = \frac{1}{4}, \quad \Pr(X = 3) = \frac{1}{8}, \quad \Pr(X = 4) = \frac{1}{8}$$

$$\text{このとき、} H(X) = \frac{1}{2} \log_2 2 + \frac{1}{4} \log_2 4 + \frac{1}{8} \log_2 8 + \frac{1}{8} \log_2 8 = \frac{9}{4}$$

エントロピーの考え方

エントロピーの直感

$H(X) \approx X$ のもつ不確実性

$H(X)$ が大きい $\leftrightarrow X$ のもつ不確実性が大きい

例

X が $\{0, 1, 2, 3\}$ の値を取る確率変数であるとき

▶	i	0	1	2	3	$\rightsquigarrow H(X) = 0$
	$\Pr(X = i)$	1	0	0	0	

▶	i	0	1	2	3	$\rightsquigarrow H(X) = 1$
	$\Pr(X = i)$	1/2	1/2	0	0	

▶	i	0	1	2	3	$\rightsquigarrow H(X) = 2$
	$\Pr(X = i)$	1/4	1/4	1/4	1/4	

二値エントロピー関数

X の取る値が 2 つ (例えば, 0 と 1) の場合

$\Pr(X = 0) = p$, $\Pr(X = 1) = 1 - p$ だとすると ($0 \leq p \leq 1$)

$$H(X) = p \log_2 \frac{1}{p} + (1 - p) \log_2 \frac{1}{1 - p}$$

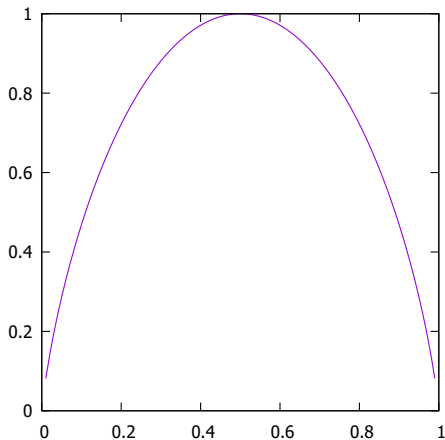
定義：二値エントロピー関数

次の関数 $H: [0, 1] \rightarrow \mathbb{R}$ を**二値エントロピー関数** と呼ぶ

$$H(p) := p \log_2 \frac{1}{p} + (1 - p) \log_2 \frac{1}{1 - p}$$

二値エントロピー関数：プロット

$$H(p) := p \log_2 \frac{1}{p} + (1-p) \log_2 \frac{1}{1-p}$$



二値エントロピー関数の性質

$$H(p) := p \log_2 \frac{1}{p} + (1-p) \log_2 \frac{1}{1-p}$$

性質：二値エントロピー関数の導関数

(演習問題)

- ▶ $\frac{d}{dp} H(p) = \log_2 \frac{1-p}{p}$
- ▶ $\frac{d^2}{dp^2} H(p) = -\frac{1}{p(1-p) \ln 2}$

$0 \leq p \leq 1$ なので、 $H(p)$ は $p = 1/2$ で最大値 1 を取る

p	0		1/2		1
$H(p)$	0	↗	1	↘	0
$H(p)'$	+	+	0	-	-
$H(p)''$	-	-	-	-	-

エントロピーの最大値

性質：エントロピーの最大値

確率変数 X が n 個の異なる値しか正の確率で取らないとき

$$H(X) \leq \log_2 n$$

注： X が異なる n 個の値をそれぞれ確率 $\frac{1}{n}$ でとると

$$H(X) = n \cdot \frac{1}{n} \log_2 n = \log_2 n$$

(一様分布がエントロピーを最大にする)

エントロピーの最大値：補題 (ギブスの不等式)

補題：ギブスの不等式

正実数 x_1, x_2, \dots, x_n と y_1, y_2, \dots, y_n が $\sum_{i=1}^n x_i \geq \sum_{i=1}^n y_i$ を満たすとき

$$\sum_{i=1}^n x_i \log_2 x_i \geq \sum_{i=1}^n x_i \log_2 y_i$$

証明： \log_2 ではなく \ln で証明すれば十分 (なぜ?)

$$\begin{aligned} \sum x_i \ln y_i - \sum x_i \ln x_i &= \sum x_i \ln \frac{y_i}{x_i} \\ &\leq \sum x_i \left(\frac{y_i}{x_i} - 1 \right) = \sum y_i - \sum x_i \leq 0 \quad \square \end{aligned}$$

エントロピーの最大値：補題 (ギブスの不等式)

補題：ギブスの不等式

正実数 x_1, x_2, \dots, x_n と y_1, y_2, \dots, y_n が $\sum_{i=1}^n x_i \geq \sum_{i=1}^n y_i$ を満たすとき

$$\sum_{i=1}^n x_i \log_2 x_i \geq \sum_{i=1}^n x_i \log_2 y_i$$

証明： \log_2 ではなく \ln で証明すれば十分 (なぜ?)

$$\begin{aligned} \sum x_i \ln y_i - \sum x_i \ln x_i &= \sum x_i \ln \frac{y_i}{x_i} \\ &\leq \sum x_i \left(\frac{y_i}{x_i} - 1 \right) = \sum y_i - \sum x_i \leq 0 \quad \square \end{aligned}$$

事実：有用な不等式

(第1回講義の復習)

任意の実数 x に対して $1 + x \leq e^x$

エントロピーの最大値：証明

- ▶ 確率変数 X が取る値は $1, 2, \dots, n$ であるとして,
 $\Pr(X = i) = p_i > 0$ とする
- ▶ ギブスの不等式を $x_i = p_i, y_i = \frac{1}{n}$ として用いると

$$\begin{aligned} H(X) &= \sum_{i=1}^n p_i \log_2 \frac{1}{p_i} \\ &= - \sum_{i=1}^n p_i \log_2 p_i \\ &\leq - \sum_{i=1}^n p_i \log_2 \frac{1}{n} = \log_2 n \end{aligned}$$

□

条件つきエントロピー：条件が事象

整数値確率変数 X , 事象 E

定義：条件つきエントロピー

E のもとでの X の **条件つきエントロピー** は

$$H(X | E) = \sum_{i \in \mathbb{Z}} \Pr(X = i | E) \log_2 \frac{1}{\Pr(X = i | E)}$$

条件つきエントロピー：条件が確率変数

整数値確率変数 X, Y

定義：条件つきエントロピー

Y のもとでの X の **条件つきエントロピー** は

$$H(X | Y) = \sum_{j \in \mathbb{Z}} \Pr(Y = j) H(X | Y = j)$$

条件つきエントロピー：公式の導出

$$\begin{aligned} H(X | Y) &= \sum_{j \in \mathbb{Z}} \Pr(Y = j) H(X | Y = j) \\ &= \sum_{j \in \mathbb{Z}} \Pr(Y = j) \sum_{i \in \mathbb{Z}} \Pr(X = i | Y = j) \log_2 \frac{1}{\Pr(X = i | Y = j)} \\ &= \sum_{i \in \mathbb{Z}} \sum_{j \in \mathbb{Z}} \Pr(Y = j) \Pr(X = i | Y = j) \log_2 \frac{1}{\Pr(X = i | Y = j)} \\ &= \sum_{i \in \mathbb{Z}} \sum_{j \in \mathbb{Z}} \Pr(X = i, Y = j) \log_2 \frac{1}{\Pr(X = i | Y = j)} \end{aligned}$$

(この式を条件つきエントロピーの定義とすることもある)

条件つきエントロピーの性質：条件づけ

有限個の値をとる確率変数 X, Y

性質：条件づけ

$$H(X | Y) \leq H(X)$$

証明：

$$\begin{aligned} H(X) &= \sum_i \Pr(X = i) \log_2 \frac{1}{\Pr(X = i)} \\ &= \sum_i \sum_j \Pr(X = i, Y = j) \log_2 \frac{1}{\Pr(X = i)} \end{aligned}$$

なので、

(次のページに続く)

条件つきエントロピーの性質：条件づけ (証明の続き)

$$\begin{aligned} & H(X) - H(X | Y) \\ &= \sum_i \sum_j \Pr(X = i, Y = j) \left(\log_2 \frac{1}{\Pr(X = i)} - \log_2 \frac{1}{\Pr(X = i | Y = j)} \right) \\ &= \sum_i \sum_j \Pr(X = i, Y = j) \log_2 \frac{\Pr(X = i | Y = j)}{\Pr(X = i)} \\ &= \sum_i \sum_j \Pr(X = i, Y = j) \log_2 \frac{\Pr(X = i, Y = j)}{\Pr(X = i) \Pr(Y = j)} \\ &\geq 0 \quad (\text{ギブスの不等式}) \end{aligned}$$

したがって、 $H(X) \geq H(X | Y)$

□

目次

- ① エントロピー
- ② エントロピーの劣加法性とシアラーの不等式
- ③ エントロピーの応用：二項係数の和の評価
- ④ エントロピーの応用：体積と射影の面積
- ⑤ 今日のまとめ

結合エントロピー

整数値確率変数 X, Y

定義：結合エントロピー (同時エントロピー)

 X, Y の **結合エントロピー** とは、次の量のこと

$$H(X, Y) = \sum_{i \in \mathbb{Z}} \sum_{j \in \mathbb{Z}} \Pr(X = i, Y = j) \log_2 \frac{1}{\Pr(X = i, Y = j)}$$

3つ以上の確率変数の結合エントロピーも同様に定義する

注1 : X, Y は独立であるとは限らない注2 : $H(X, Y) = H(Y, X)$

結合エントロピーと条件つきエントロピー

整数値確率変数 X, Y

性質：結合エントロピーと条件つきエントロピー

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

証明：演習問題

エントロピーの劣加法性

有限個の異なる値を取る確率変数 X, Y

性質：エントロピーの劣加法性

$$H(X, Y) \leq H(X) + H(Y)$$

証明：

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

□

エントロピーの劣加法性 (3 つ以上の場合)

劣加法性は 3 つ以上の確率変数 X_1, X_2, \dots, X_n に対しても成り立つ (ただし, 取る値は有限とする)

性質: エントロピーの劣加法性

(演習問題)

$$H(X_1, X_2, \dots, X_n) \leq \sum_{i=1}^n H(X_i)$$

証明は, 2 変数の場合と同様にできる (n に関する数学的帰納法)

シアラーの不等式 (簡単な場合)

有限個の値をとる確率変数 X, Y, Z

性質：シアラーの不等式 (簡単な場合)

$$2H(X, Y, Z) \leq H(X, Y) + H(X, Z) + H(Y, Z)$$

証明：

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

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

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

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

シアラーの不等式 (簡単な場合)

有限個の値をとる確率変数 X, Y, Z

性質：シアラーの不等式 (簡単な場合)

$$2H(X, Y, Z) \leq H(X, Y) + H(X, Z) + H(Y, Z)$$

証明：

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

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

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

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

シアラーの不等式 (簡単な場合)

有限個の値をとる確率変数 X, Y, Z

性質：シアラーの不等式 (簡単な場合)

$$2H(X, Y, Z) \leq H(X, Y) + H(X, Z) + H(Y, Z)$$

証明：

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

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

$$H(X, Z) = H(X) + H(Z | X) \geq H(X) + H(Z | X, Y)$$

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

シアラーの不等式 (簡単な場合)

有限個の値をとる確率変数 X, Y, Z

性質：シアラーの不等式 (簡単な場合)

$$2H(X, Y, Z) \leq H(X, Y) + H(X, Z) + H(Y, Z)$$

証明：

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

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

$$H(X, Z) = H(X) + H(Z | X) \geq H(X) + H(Z | X, Y)$$

$$H(Y, Z) = H(Y) + H(Z | Y) \geq H(Y) + H(Z | X, Y)$$

シアラーの不等式 (簡単な場合)

有限個の値をとる確率変数 X, Y, Z

性質：シアラーの不等式 (簡単な場合)

$$2H(X, Y, Z) \leq H(X, Y) + H(X, Z) + H(Y, Z)$$

証明：

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

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

$$H(X, Z) = H(X) + H(Z | X) \geq H(X) + H(Z | X, Y)$$

$$H(Y, Z) = H(Y) + H(Z | Y) \geq H(Y) + H(Z | X, Y)$$

したがって、

$$\begin{aligned} 2H(X, Y, Z) &= 2H(X) + 2H(Y | X) + 2H(Z | X, Y) \\ &\leq 2H(X) + H(Y | X) + H(Y) + 2H(Z | X, Y) \\ &\leq H(X, Y) + H(X, Z) + H(Y, Z) \end{aligned}$$

□

シアラーの不等式 (一般的な場合)

設定

有限個の値をとる確率変数 X_1, X_2, \dots, X_n

- ▶ 添字の集合 $I \subseteq \{1, 2, \dots, n\}$ に対して,
 X_I と書いたら, X_i ($i \in I$) を並べた確率変数とする
- ▶ 添字の集合 $I_1, I_2, \dots, I_m \subseteq \{1, 2, \dots, n\}$ に対して,
 各 $i \in \{1, 2, \dots, n\}$ がこの中の k 個以上に含まれるとする

性質：シアラーの不等式 (一般的な場合)

(演習問題)

上の設定のもとで, 次が成り立つ

$$k \cdot H(X_1, X_2, \dots, X_n) \leq \sum_{\ell=1}^m H(X_{I_\ell})$$

先ほどの「簡単な場合」は

$n = 3, m = 3, I_1 = \{1, 2\}, I_2 = \{1, 3\}, I_3 = \{2, 3\}, k = 2$ の場合に対応する

シアラーの不等式：例

- ▶ $2H(X_1, X_2, X_3) \leq H(X_1, X_2) + H(X_1, X_3) + H(X_2, X_3)$
- ▶ $2H(X_1, X_2, X_3, X_4) \leq H(X_1, X_2, X_3) + H(X_1, X_4) + H(X_2, X_4) + H(X_3, X_4)$
- ▶ $(n - 1)H(X_1, X_2, \dots, X_n) \leq \sum_{i=1}^{n-1} \sum_{j=i+1}^n H(X_i, X_j)$

設定 (再掲)

有限個の値をとる確率変数 X_1, X_2, \dots, X_n

- ▶ 添字の集合 $I \subseteq \{1, 2, \dots, n\}$ に対して,
 X_i と書いたら, $X_i (i \in I)$ を並べた確率変数とする
- ▶ 添字の集合 $I_1, I_2, \dots, I_m \subseteq \{1, 2, \dots, n\}$ に対して,
 各 $i \in \{1, 2, \dots, n\}$ がこの中の k 個以上に含まれるとする

目次

- ① エントロピー
- ② エントロピーの劣加法性とシアラーの不等式
- ③ エントロピーの応用：二項係数の和の評価
- ④ エントロピーの応用：体積と射影の面積
- ⑤ 今日のまとめ

極値組合せ論のある定理

設定

- ▶ $S_1, S_2, \dots, S_m \subseteq \{1, 2, \dots, n\}$
- ▶ 各 $i \in \{1, 2, \dots, n\}$ に対して, $p_i = \frac{1}{m} |\{j \in \{1, 2, \dots, m\} \mid i \in S_j\}|$
(p_i は S_1, \dots, S_m の中で i を含むものの割合)

定理 (Kleitman, Shearer, Sturtevant '81)

上の設定のもとで

$$\log_2 m \leq \sum_{i=1}^n H(p_i)$$

極値組合せ論のある定理：証明 (1)

証明： $S_1, \dots, S_m \subseteq \{1, 2, \dots, n\}$ から次のような確率変数を作る

- ▶ S_1, \dots, S_m の 1 つを等確率 $1/m$ で選ぶ (選んだものを S_j とする)
- ▶ 各 $i \in \{1, 2, \dots, n\}$ に対して、次のように X_i を作る

$$X_i = \begin{cases} 1 & (i \in S_j), \\ 0 & (i \notin S_j) \end{cases}$$

このとき、次が成り立つ

- ▶ $H(X_1, X_2, \dots, X_n) = m \cdot \frac{1}{m} \log_2 m = \log_2 m$
- ▶ 各 $i \in \{1, 2, \dots, n\}$ に対して、 $H(X_i) = H(p_i)$

極値組合せ論のある定理：証明 (2)

証明 (続き)：したがって、エントロピーの劣加法性より

$$\log_2 m = H(X_1, X_2, \dots, X_n) \leq \sum_{i=1}^n H(X_i) = \sum_{i=1}^n H(p_i)$$



応用：二項係数の和の評価

定理：二項係数の和の評価

任意の正整数 n と実数 $p \in (0, 1/2]$ に対して、次が成り立つ

$$\sum_{i=0}^{\lfloor np \rfloor} \binom{n}{i} \leq 2^{nH(p)}$$

証明：Kleitman–Shearer–Sturtevant の定理を使う

- ▶ $\{1, 2, \dots, n\}$ の部分集合で要素数が $0, 1, \dots, \lfloor np \rfloor$ であるものを全部考えて、 S_1, S_2, \dots, S_m とする

- ▶ このとき、 $m = \sum_{i=0}^{\lfloor np \rfloor} \binom{n}{i}$

応用：二項係数の和の評価：続き

定理：二項係数の和の評価

任意の正整数 n と実数 $p \in (0, 1/2]$ に対して、次が成り立つ

$$\sum_{i=0}^{\lfloor np \rfloor} \binom{n}{i} \leq 2^{nH(p)}$$

証明 (続き)：Kleitman–Shearer–Sturtevant の定理を使う

- ▶ p_1, p_2, \dots, p_n は次を満たす
 - ▶ $p_1 = p_2 = \dots = p_n$
 - ▶ $p_1 + p_2 + \dots + p_n \leq np$ (なぜ?)
- ▶ $\therefore p_1 = p_2 = \dots = p_n \leq p$
- ▶ したがって、Kleitman–Shearer–Sturtevant の定理より

$$\log_2 m \leq \sum_{i=1}^n H(p_i) \leq nH(p)$$

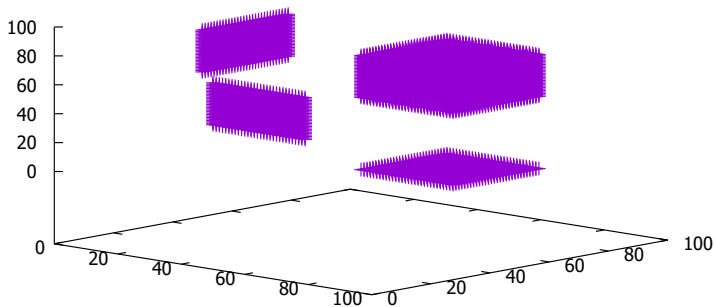
$0 \leq p_i \leq p \leq 1/2$ より、 $H(p_i) \leq H(p)$ であることに注意

□

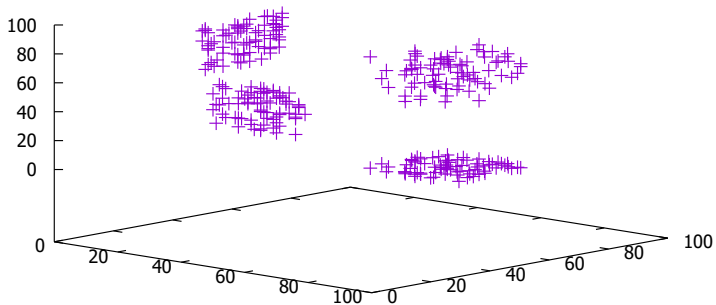
目次

- ① エントロピー
- ② エントロピーの劣加法性とシアラーの不等式
- ③ エントロピーの応用：二項係数の和の評価
- ④ エントロピーの応用：体積と射影の面積
- ⑤ 今日のまとめ

応用：体積と射影の面積 (例 1)



応用：体積と射影の面積 (例 2)



応用：体積と射影の面積

設定

- ▶ $A \subseteq \{(x, y, z) \mid x, y, z \in \{1, 2, \dots, n\}\}$
- ▶ $A_x = \{(0, y, z) \mid (x, y, z) \in A\},$
 $A_y = \{(x, 0, z) \mid (x, y, z) \in A\},$
 $A_z = \{(x, y, 0) \mid (x, y, z) \in A\},$

性質

$$|A|^2 \leq |A_x| \cdot |A_y| \cdot |A_z|$$

体積と射影の面積：証明

証明：次のように確率変数 X, Y, Z を構成する

- ▶ A の要素を 1 つ一様分布に従って選ぶ
- ▶ 選んだ要素を (X, Y, Z) とする

シアラーの不等式より

$$2H(X, Y, Z) \leq H(X, Y) + H(X, Z) + H(Y, Z)$$

ここで,

- ▶ $H(X, Y, Z) = \log_2 |A|$
- ▶ $H(X, Y) \leq \log_2 |A_z|$, $H(X, Z) \leq \log_2 |A_y|$, $H(Y, Z) \leq \log_2 |A_x|$

したがって, $\log_2 |A|^2 \leq \log_2 |A_z| |A_y| |A_x|$

□

目次

- ① エントロピー
- ② エントロピーの劣加法性とシアラーの不等式
- ③ エントロピーの応用：二項係数の和の評価
- ④ エントロピーの応用：体積と射影の面積
- ⑤ 今日のまとめ

今日の目標

今日の目標

- ▶ エントロピーの基本的な性質を導出できる
 - ▶ 条件つきエントロピーと結合エントロピー
 - ▶ 劣加法性
 - ▶ シアラーの不等式
- ▶ エントロピーを使って、離散数学の基礎事項を証明できる
 - ▶ 二項係数の和の評価
 - ▶ 体積と射影の面積の関係

格言

エントロピーを使って、離散数学の事項を証明できる

目次

- ① エントロピー
- ② エントロピーの劣加法性とシアラーの不等式
- ③ エントロピーの応用：二項係数の和の評価
- ④ エントロピーの応用：体積と射影の面積
- ⑤ 今日のまとめ