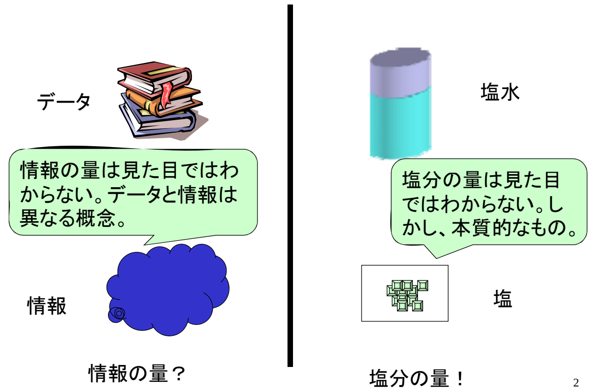
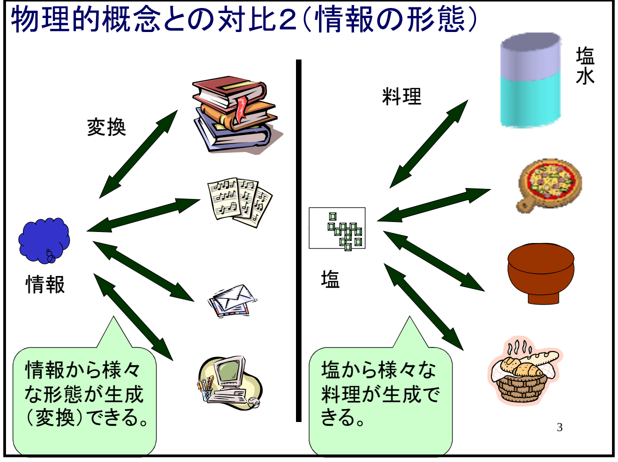
情報量（2章）

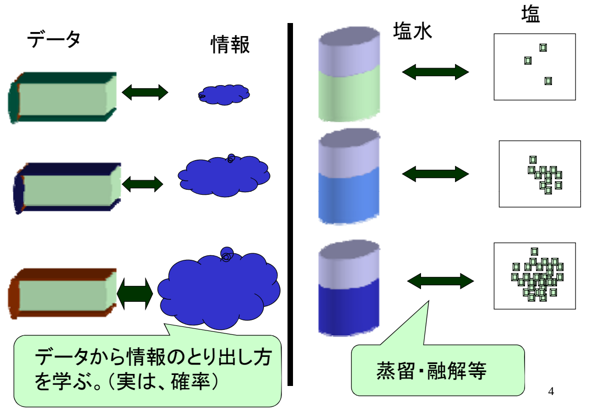
物理的概念との対比１（入れ物と中身）



物理的概念との対比２（情報の形態）



物理的概念との対比３（情報量の計測）



情報の大小（情報量の性質１）

事象（出来事）　　ニュースになる！　　情報量

(a)沖縄に雪が降った。　　$i(a)$

(b)北海道に雪が降った。　　$i(b)$

ニュースになる？

一種の重要度と考える。

どれが妥当か？

A: $i(a) < i(b)$

B: $i(a) = i(b)$

C: $i(a) > i(b)$

練習

次のニュース（事象）の確率と情報量の大小関係を示せ。

（1）

（宝a）買った宝くじが外れた。

（宝b）買った宝くじが１等当たった。

（2）

（事a）今日事故にあった。

（事b）今日事故にあわなかった

（3）

（サa)サイコロを振ったら１が出た。

（サb)偶数がでた。

情報量と確率の関係

事象 a の確率を$P(a)$と表す。

事象	情報量	確率
(a) 沖縄に雪が降った。	$i(a)$	$P(a)$
(b) 北海道に雪が降った。	$i(b)$	$P(b)$
	$i(a) > i(b)$	$P(a) < P(b)$

情報量の性質1

1. 情報量は、確率の関数
2. 情報量は、確率に関する減少関数
（確率が増加すれば、情報量は減少する。）

7

独立事象の確率と情報量（情報量の性質2）

事象「宝くじが当たって、しかも事故にあった。」の情報量を考えよう。

(宝a) 買った宝くじが外れた。
(宝b) 買った宝くじが1等当たった。

(事a) 今日事故にあった。
(事b) 今日事故にあわなかった。

（互いに無関係な事象を独立な事象と言う。）

”独立”な事象の積事象の確率

$P(\text{宝b}) \times P(\text{事a})$

”独立”な事象の積事象の情報量

$i(\text{宝}b) + i(\text{事a})$

確率統計の復習
積事象の確率が、各事象の確率の和であるとき、独立な事象という。

3. 独立な事象の積事象の情報量は、個々の情報量の和

8

情報量の性質から情報量の数値化へ

事象 x がおきる確率を $P(x)$ と表し、
事象 x がおきたことを知った情報量を $i(x)$ と表す。
このとき、以下を満たす関数$f(x)$ で情報量を定義する。

1. 情報量は、確率の関数である。

$$i(x) = f(P(x))$$

2. 情報量は、確率に対する減少関数である。

$$P(x_1) < P(x_2) \Leftrightarrow i(x_1) > i(x_2)$$

3.独立な積事象を知ったときの情報量は、個々の情報量の和である。

$$\begin{aligned} & i(x_1 \wedge x_2) \\ &= f(P(x_1 \wedge x_2)) = f(P(x_1)) + f(P(x_2)) \\ &= i(x_1) + i(x_2) \end{aligned}$$

9

(自己)情報量と情報量の単位

先のスライドを満たすように、
ある事象 x を知る情報量$i(x)$ は以下の関数で定義される。

情報量は、確率の逆数の対数。

定義（自己情報量）

$$i(x) = -\log_s P(x)$$

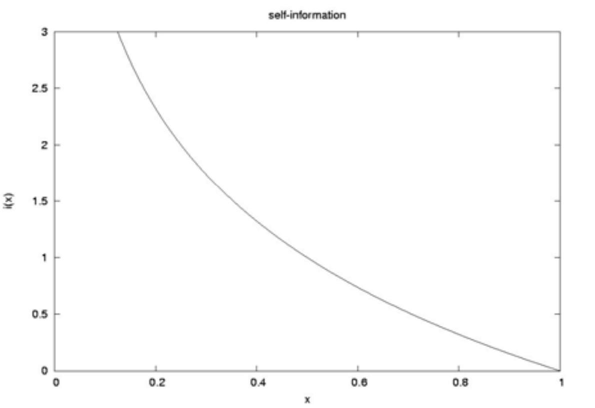
ただし、 $s > 1$

情報量の単位

$(s = 2)$ → [bit]（ビット）
$(s = e)$ → [nat]（ナット）
$(s = 10)$ → [decit]（デシット）

1ビットは確率0.5の事象が起きたことを知る情報量。以後は、ビットだけを扱う。

自己情報量の外形



練習

次の事象の自己情報量を求めよ。

（1）

a ： 2枚のコインを投げて両方とも裏がでる。

（2）

b ： 52枚のトランプから絵札を1枚引く

（3）

c ： アルファベットの書いてある26個の玉から、cの玉を取り出す。

12

事象系

単独の事象ではなくて、事象の集合を考える。
事象の集合とそれが生じる確率を以下のように表す。

$$X=\begin{Bmatrix} x_1 & , & x_2 & , & \cdots & , & x_n \\ P(x_1) & , & P(x_2) & , & \cdots & , & P(x_n) \end{Bmatrix}$$

ここで、

$$0 \le P(x_i) \le 1 \quad (1 \le i \le n)$$

$$\sum_{i=1}^{n} P(x_i) = 1$$

このように、確率が定められた事象の集合を事象系と
言う。

13

事象系例

（1） コイントスの事象系

$$C=\begin{Bmatrix} 表 & , & 裏 \\ \frac{1}{2} & , & \frac{1}{2} \end{Bmatrix}$$

（2） サイコロの事象系

$$D=\begin{Bmatrix} 1 & , & 2 & , & 3 & , & 4 & , & 5 & , & 6 \\ \frac{1}{6} & , & \frac{1}{6} & , & \frac{1}{6} & , & \frac{1}{6} & , & \frac{1}{6} & , & \frac{1}{6} \end{Bmatrix}$$

（3）トランプを引いたときの数の事象系

$$T=\begin{Bmatrix} 1 & , & 2 & , & \cdots & , & K \\ \frac{1}{13} & , & \frac{1}{13} & , & \cdots & , & \frac{1}{13} \end{Bmatrix}$$

14

練習

次の事象系を形式的に示せ。

（1）トランプを引いたときのカードのマークの事象系

（2）26文字のアルファベットと空白文字が書かれた、27
個の玉が袋に入っている。その袋から1つの玉をとりだ
しす事象系。

15

事象系の平均情報量

定義（エントロピー）

事象系 X のすべての事象に対して、その情報量の平
均をとったものを**平均情報量**または**エントロピー**という。

ある事象系 X が以下のように与えられるとする。

$$X=\begin{Bmatrix} x_1 & , & x_2 & , & \cdots & , & x_n \\ P(x_1) & , & P(x_2) & , & \cdots & , & P(x_n) \end{Bmatrix}$$

このとき、平均情報量 $H(X)$ は自己情報量を元に
以下のように表される。

$$\begin{aligned} H(X) &= \sum_{x \in X} P(x) i(x) \\ &= -\sum_{x \in X} P(x) \log P(x) \end{aligned}$$

慣用的に、エントロ
ピーには記号 H を
用いる。

16

自己情報量と平均情報量

自己情報量は事象に対して定義され、
平均情報量は事象系（事象の集合）に対して定義される。

事象 → **自己情報量**

事象系（情報源） → **平均情報量（エントロピー）**

17

平均情報量の計算例

（1）

$$C=\begin{Bmatrix} 表 & , & 裏 \\ \frac{1}{2} & , & \frac{1}{2} \end{Bmatrix}$$

$$\begin{aligned} H(C) &= -\frac{1}{2}\log_2\frac{1}{2} - -\frac{1}{2}\log_2\frac{1}{2} \\ &= \frac{1}{2}\log_2 2 + \frac{1}{2}\log_2 2 \\ &= 1 \end{aligned}$$

（2）

$$D=\begin{Bmatrix} 1 & , & 2 & , & 3 & , & 4 & , & 5 & , & 6 \\ \frac{1}{6} & , & \frac{1}{6} & , & \frac{1}{6} & , & \frac{1}{6} & , & \frac{1}{6} & , & \frac{1}{6} \end{Bmatrix}$$

$$\begin{aligned} H(D) &= -\frac{1}{6}\log_2\frac{1}{6} - \frac{1}{6}\log_2\frac{1}{6} - \frac{1}{6}\log_2\frac{1}{6} - \frac{1}{6}\log_2\frac{1}{6} - \frac{1}{6}\log_2\frac{1}{6} - \frac{1}{6}\log_2\frac{1}{6} \\ &= 6 \times \frac{1}{6}\log_2 6 \\ &= \log_2 6 \\ &= 2.585\cdots \end{aligned}$$

18

練習1

次の確率事象系(情報源)の平均情報量(エントロピーを求めよ。)

(1)

$$A = \begin{Bmatrix} a_1 & , & a_2 \\ \frac{1}{4} & , & \frac{3}{4} \end{Bmatrix}$$

(2)

$$B = \begin{Bmatrix} b_1 & , & b_2 & , & b_3 \\ \frac{1}{10} & , & \frac{3}{10} & , & \frac{6}{10} \end{Bmatrix}$$

(3)

$$C = \begin{cases} c_1 & \text{さいころを振って１、または４の目} \\ c_2 & \text{さいころを振って１、４以外の目} \end{cases}$$

19

練習2

次の事象系のエントロピーを求めよ。

(1) トランプを引いたときのカードのマークの事象系

(2)26文字のアルファベットと空白文字が書かれた、27個の玉が袋に入っている。その袋から1つの玉をとりだす事象系。

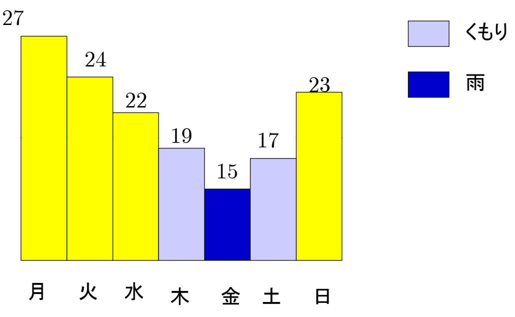
20

補足:平均(期待値)の話

21

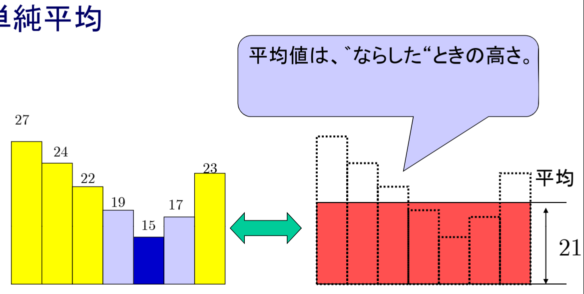
期待値の式

1週間の間の平均気温を求めたい。



22

単純平均



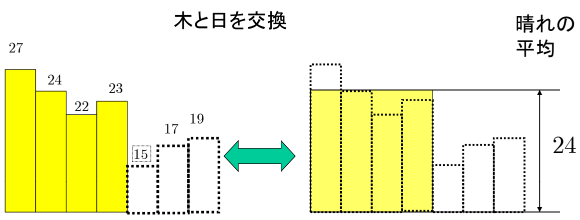
面積が等しい

$$\overline{T} = \frac{27+24+22+19+15+17+23}{7}$$

$$= 21$$

23

晴れの平均気温



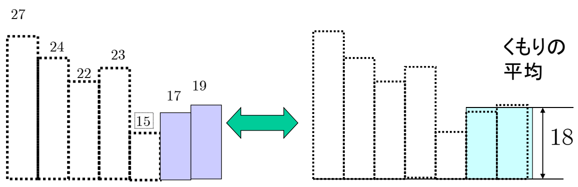
面積が等しい

$$\overline{T_f} = \frac{27+24+22+23}{4}$$

$$= 24$$

24

曇りの平均気温

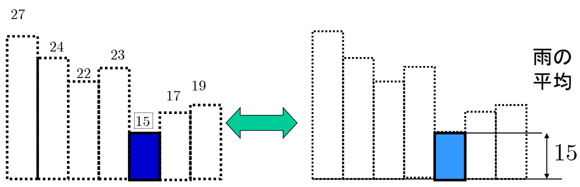


面積が等しい

$$\overline{T_c} = \frac{17+19}{2}$$
$$= 18$$

25

雨の平均気温



面積が等しい

$$\overline{T_r} = \frac{15}{1}$$
$$= 15$$

26

よくある間違い

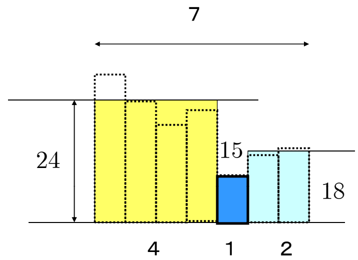
1週間のすべての日は、「晴れ」か「くもり」か「雨」であった。
「晴れ」の日の平均気温は $\overline{T_f} = 24$ 度 、
「くもり」の日の平均気温は $\overline{T_c} = 18$ 度、
「雨」の日の平均気温は $\overline{T_r} = 15$
であった。
一週間の平均気温を求めよ。

$$\overline{T_{wrong}} = \frac{\overline{T_f} + \overline{T_c} + \overline{T_r}}{3}$$
$$= \frac{24+18+15}{3}$$
$$= 19$$

この計算は間違い。
（平均の平均を求めるときには、注意が必要。）
このような計算ができるためには、どのような条件が必要か？

27

確率



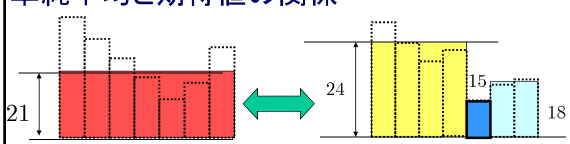
晴れの確率：$P_f = \frac{4}{7}$

雨の確率： $P_r = \frac{1}{7}$

曇りの確率：$P_c = \frac{2}{7}$

28

単純平均と期待値の関係



面積が等しい

$$\overline{T} = P_f\overline{T_f} + P_r\overline{T_r} + P_c\overline{T_c}$$
$$= \frac{4}{7} \cdot \underbrace{\frac{27+24+22+23}{4}}_{24} + \frac{1}{7} \cdot \underbrace{\frac{15}{1}}_{15} + \frac{2}{7} \cdot \underbrace{\frac{19+17}{2}}_{18}$$
$$= \frac{27+24+22+19+15+17+23}{7}$$
$$= 21$$

29

2事象の事象系の平均情報量（重要）

ある事象系 X が以下のように与えられるとする。

$$X = \begin{Bmatrix} x & , & \overline{x} \\ p & , & 1-p \end{Bmatrix}$$

引数が、事象系。
記号も注意する。

このとき、エントロピーは次式となる。

$$H(X) = -p\log p - (1-p)\log(1-p)$$

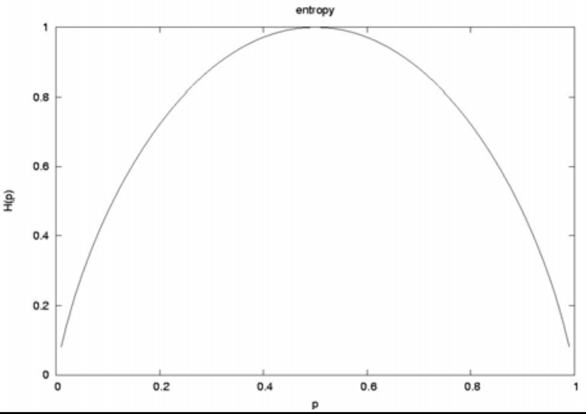
この形の事象系は非常によく用いられ、この右辺の形の関数をエントロピー関数といい以下のように表す。

$$\mathcal{H}(p) \equiv -p\log p - (1-p)\log(1-p)$$

引数が、数。
記号も注意する。

30

エントロピー関数の外形 $\mathcal{H}(p)$



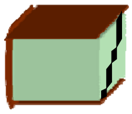
具体例1：文書の情報量（1記号あたりの平均情報量の意味）

アルファベット $\mathcal{A} = \{a, b, c, \cdots, z\}$ に対して、
個々の記号の出現確率を考えよう。

記号 $\alpha \in \mathcal{A}$ の出現する確率を $P(\alpha)$ と書く。

例えば、 $P(a)$: a の現れる確率
$P(b)$: b の現れる確率

辞書



このとき、一般の英文において、各記号の出現確率は異なる。すなわち、以下である。

$$P(a) \neq P(b) \neq \cdots \neq P(z) \neq \frac{1}{26}$$

32

まず、アルファベット $\mathcal{A} = \{a, b, c, \cdots, z\}$ 中の文字を用いた文書の情報量を考える。

今、$\mathcal{A}$ の n 個の文字からなる英文

$$S = s_1 s_2 \cdots s_n$$

出現確率に基づく事象とみなせることに注意する。

の情報量を $I(S)$ とする。

$I(S)$ は各記号が S に出現する自己情報量の総和である。

$$I(S) = \sum_{i=1}^{n} i(s_i)$$

事象に対する
確率の逆数の対数

33

今度は逆に、情報量 $I(S)$ を持つ事象を文書 S で表すことを考える。

まず、アルファベットの（1記号あたりの）平均情報量（エントロピー）が以下の式で表される。

$$H(\mathcal{A}) = \sum_{\alpha=a}^{z} \{P(\alpha) \times i(\alpha)\}$$

$$= P(a)i(a) + P(b)i(b) + \cdots + P(z)i(z)$$

$$= -P(a)\log P(a) - P(b)\log P(b) - \cdots - P(z)\log P(z)$$

$$= -\sum_{\alpha=a}^{z} P(\alpha)\log P(\alpha)$$ [bit/記号]

34

よって、文書 S 中に含まれる文字数の期待値 $\overline{n}$ は次式で求められる。

$$\overline{n}[\text{記号}] = \frac{I(S)[bit]}{H(\mathcal{A})[bit/\text{記号}]}$$

したがって、文書の長さは情報源 $\mathcal{A}$ の1記号あたりの平均情報量（エントロピー） $H(\mathcal{A})$ を"単位"とした情報量で表される。
また、情報源 $\mathcal{A}$ の長さ n の文書（事象）の平均情報量 $\overline{I}(S)$ に関して次式が成り立つ。

$$\overline{I}(S)[bit] = n[\text{記号}] \times H(\mathcal{A})[bit/\text{記号}]$$

35

練習1

次の事象系のエントロピー（平均情報量）を求めよ。

（1）

$$A = \begin{Bmatrix} a_1 & , & a_2 & , & a_3 \\ \frac{1}{5} & , & \frac{2}{5} & , & \frac{2}{5} \end{Bmatrix}$$

（2）

$$B = \begin{Bmatrix} b_1 & , & b_2 & , & b_3 & , & b_4 \\ \frac{1}{9} & , & \frac{2}{9} & , & \frac{2}{9} & , & \frac{5}{9} \end{Bmatrix}$$

36

練習2

情報源(事象系)から生成された文字列の自己情報量を求めよ。

（1）

$$A = \begin{Bmatrix} a_1 & , & a_2 & , & a_3 \\ \frac{1}{5} & , & \frac{2}{5} & , & \frac{2}{5} \end{Bmatrix}$$

から生成された、文字列 $s_1 = a_3 a_2 a_1 a_3$ の自己情報量 $I(s_1)$

（2）

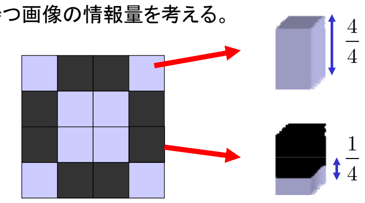
$$B = \begin{Bmatrix} b_1 & , & b_2 & , & b_3 & , & b_4 \\ \frac{1}{9} & , & \frac{2}{9} & , & \frac{2}{9} & , & \frac{5}{9} \end{Bmatrix}$$

から生成された、文字列 $s_2 = b_3 b_4 b_4 b_1 b_2$ の自己情報量 $I(s_2)$

37

具体例2：画像の情報量

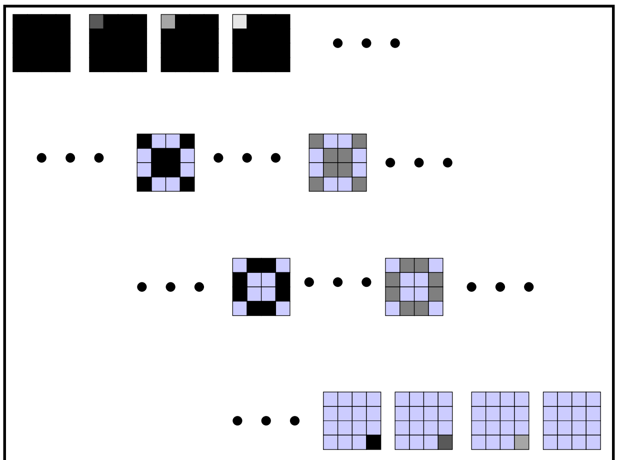
４×４の画素数を持ち、個々の画素が4階調の輝度値を持つ画像の情報量を考える。



各画素がある輝度値を持つ確率を、$\frac{1}{4}$とする。このとき、この画像が持つ情報量I(画)は次式で表される。

$$I(画)=\sum_{j=1}^{16} i(画素_j) = 16 \times \log 4 = 32[bit] = 4[byte]$$

38



前のスライドでは、輝度値が全ての画素で確率が均等であると仮定してあった。
実は、風景画、人物画等では、画素における輝度値の出現確率は一様ではない。
このような場合、より少ない情報量しか持たない。
すなわち、確率が一様でない場合には、１画素あたりの平均情報量は、一様な場合より小さくなる。

40

練習

下の画像の自己情報量を求めよ。ただし、各画素の諧調値はすべて均等に現れるのものとする。

（1）

256×256 の画素で、各画素が 16 階調の白黒画像の自己情報量を求めよ。

（1）

256×256 の画素のカラー画像の自己情報量を求めよ。ただし、1画素は赤、緑、青（RGB）の各色で表され、各色はそれぞれ16階調で表されるものとする。

41