創想館　14-201

2008-3

Tuesday　SPRING

13:00－14:30

2008年度

「情報工学」

第3回講義

前回学んだこと

復習

すでに，
一つでも答えられないこと
があったら，大変！

- 覚えていなければいけないこと
- 答えられなければならないこと

✓ 確率過程とは？
✓ マルコフ過程とは？
✓ 生起確率とは？　その一般的性質は？ $\sum_{i=1}^{N} p(i) = 1$
✓ 遷移確率とは？　その一般的性質は？ $\sum_{j=1}^{N} p(j/i) = 1$
✓ シャノン線図とは？　自分で作れるか？
✓ シャノン線図を深く見るとは？　またその意義は？
✓ エルゴード的とは？
✓ シャノンの情報量とは？
✓ エントロピーとは？ $H(X) = -\sum_{i=1}^{N} p(a_i)\log_2 p(a_i)\ [bit]$

ANIMATION

エントロピーの定義

エントロピー：$H(X)$（1通報あたりの平均情報量）

$$H(X) = -\sum_{i=1}^{N} p(a_i)\log_2 p(a_i)\ [bit]$$

復習

・・・(3.2)

（確率事象系 X の平均情報量）

・通報列 $\{a_i\}_{i=1}^{N}$ ： $a_1, a_2, a_3, \cdots a_i, \cdots a_N$

・確率事象系 X ：

$$X = \begin{Bmatrix} a_1 & a_2 & \cdots & a_i & \cdots & a_N \\ p(a_1) & p(a_2) & \cdots & p(a_i) & \cdots & p(a_N) \end{Bmatrix}$$

Example 3.1 : 天気予報のエントロピー

復習

天気予報が提供する情報量を求めてみよう！

ただし、天気予報は、外れないものとする.

天気＼地方	X	Y	Z	W
F	1/2	7/8	1	1/4
C	1/4	0	0	1/4
R	1/8	1/8	0	1/4
S	1/8	0	0	1/4

$$H_X = -\frac{1}{2}log_2\frac{1}{2} - \frac{1}{4}log_2\frac{1}{4} - \left(\frac{1}{8}log_2\frac{1}{8}\right)\times 2 = 1.8 \ \ [bit]$$

Example 3.1 : 天気予報のエントロピー

復習

$$H_X = -\frac{1}{2}log_2\frac{1}{2} - \frac{1}{4}log_2\frac{1}{4} - \left(\frac{1}{8}log_2\frac{1}{8}\right)\times 2 = 1.8 \ \ [bit]$$

$$H_y = -\frac{7}{8}log_2\frac{7}{8} - 0 - \frac{1}{8}log_2\frac{1}{8} - 0 = 0.5 \ \ [bit]$$

$$H_z = -1log_2 1 - 0 - 0 - 0 = 0 \ \ [bit]$$

$$H_w = \left[-\frac{1}{4}log_2\frac{1}{4}\right]\times 4 = 2 \ \ [bit]$$

・・・最大！

memo

生起確率の「ばらつき」大 → **エントロピー 大**

（参考） エントロピー ： 熱力学の用語

- 自然界では、エントロピーは増大方向に向かう
- エントロピー値は、「ばらつき」のパラメータと言える

エントロピー H_X は、確率事象系 X の「不確かさ」の度合

授業後の

質問への答え.1

昨年も，
同じ質問があった・・・

こういう質問は，
是非授業中にして欲しい！

$$H(X) = -\sum_{i=1}^{N} p(a_i)\log_2 p(a_i) \overset{?}{=} \frac{1}{N}\sum_{i=1}^{N} \log_2 p(a_i)$$

とはならないのではないか？

X地方の天気予報の例で計算してみると，合わない．

間違ってはいない！！

平均値とは

（例）100gのりんご８個と200gのりんご2個がある。

りんご1個の重さの平均値は？

解法1. $[100g\text{のりんごの生起確率}]\times 100g + [200g\text{のりんごの生起確率}]\times 200g$

$$= \frac{8}{10}\times 100 + \frac{2}{10}\times 200 = 120g$$

解法2. $\frac{10\text{個のりんごの総量}}{\text{りんごの総個数}} = \frac{100\times 8 + 200\times 2}{10} = 120g$

授業後の

質問への答え.1

昨年も，
同じ質問があった・・・

こういう質問は，
是非授業中にして欲しい！

$$H(X) = -\sum_{i=1}^{N} p(a_i)\log_2 p(a_i) = \frac{1}{N}\sum_{i=1}^{N} \log_2 p(a_i)$$

とはならないのではないか？

X地方の天気予報の例で計算してみると，合わない．

間違ってはいない！！

天気予報の情報量では

解法1. $H_X = -\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{4}\log_2\frac{1}{4} - \left(\frac{1}{8}\log_2\frac{1}{8}\right)\times 2 = 1.8\ [bit]$

解法2. $H_X = -\frac{1}{N}\left\{\left(\log_2\frac{1}{2}\right)\times\frac{N}{2} + \left(\log_2\frac{1}{4}\right)\times\frac{N}{4} + \left(\log_2\frac{1}{8}\right)\times\frac{N}{8} + \left(\log_2\frac{1}{8}\right)\times\frac{N}{8}\right\}$

$$= 1.8\ [bit]$$

情報量 $\log_2\frac{1}{2}$ の日数

質問してくれて，
良かった・・・
ありがとう！

授業後の

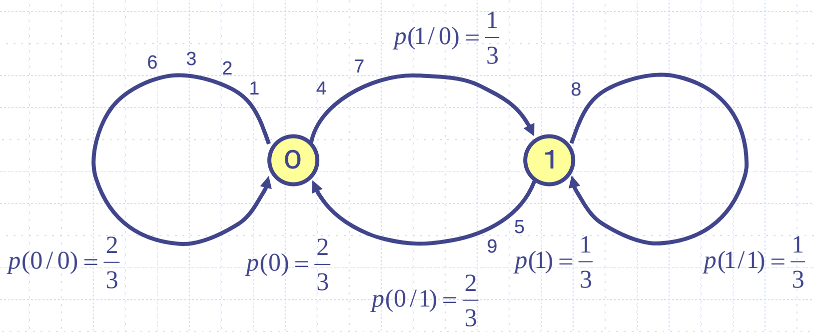
質問への答え.2

今年はなかったが，
昨年あった質問をついでに・・・

こういう質問も，
是非授業中にして欲しい！

A

00001001100100011000100001000100110000110
0100000001100011000 0110001110001010010110
10000010



単純マルコフ過程で見たシャノン線図

授業後の

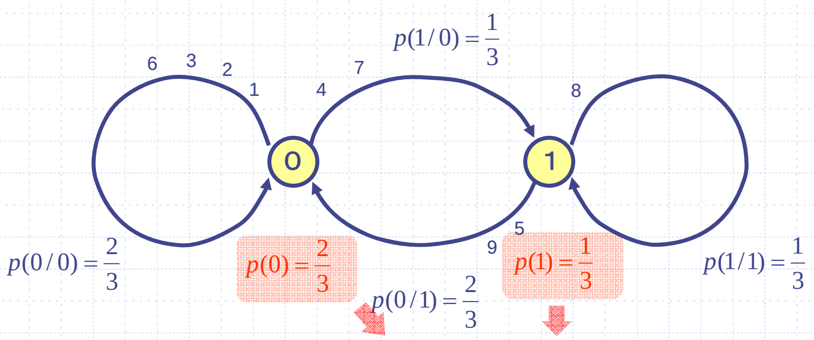
質問への答え.2

今年はなかったが，
昨年あった質問をついでに・・・

こういう質問も，
是非授業中にして欲しい！

A

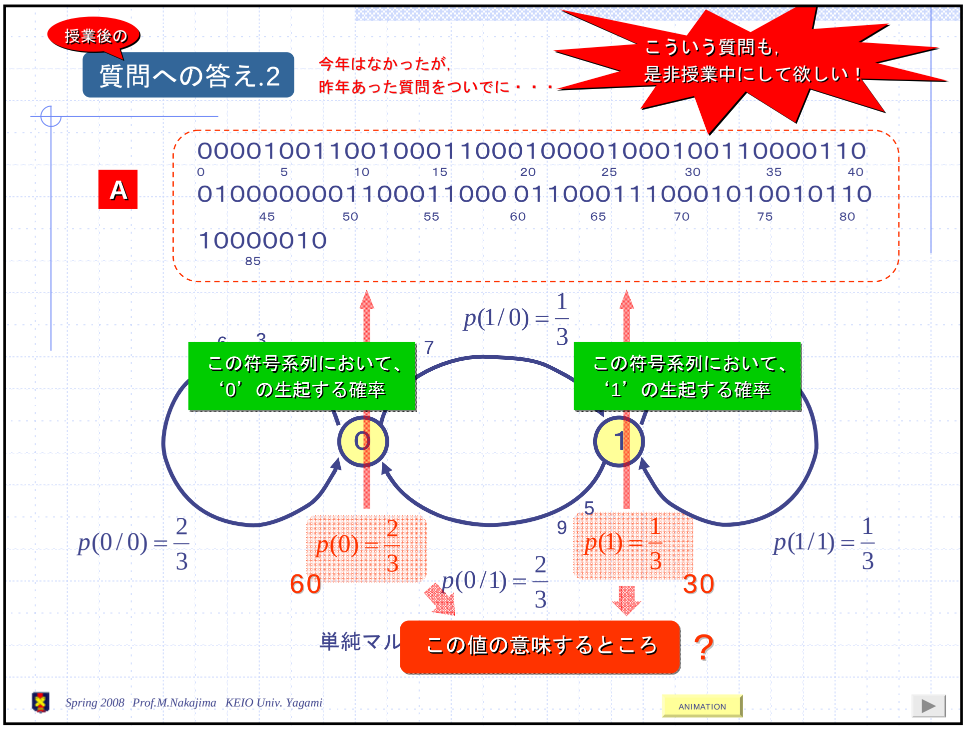
00001001100100011000100001000100110000110
0100000001100011000 0110001110001010010110
10000010

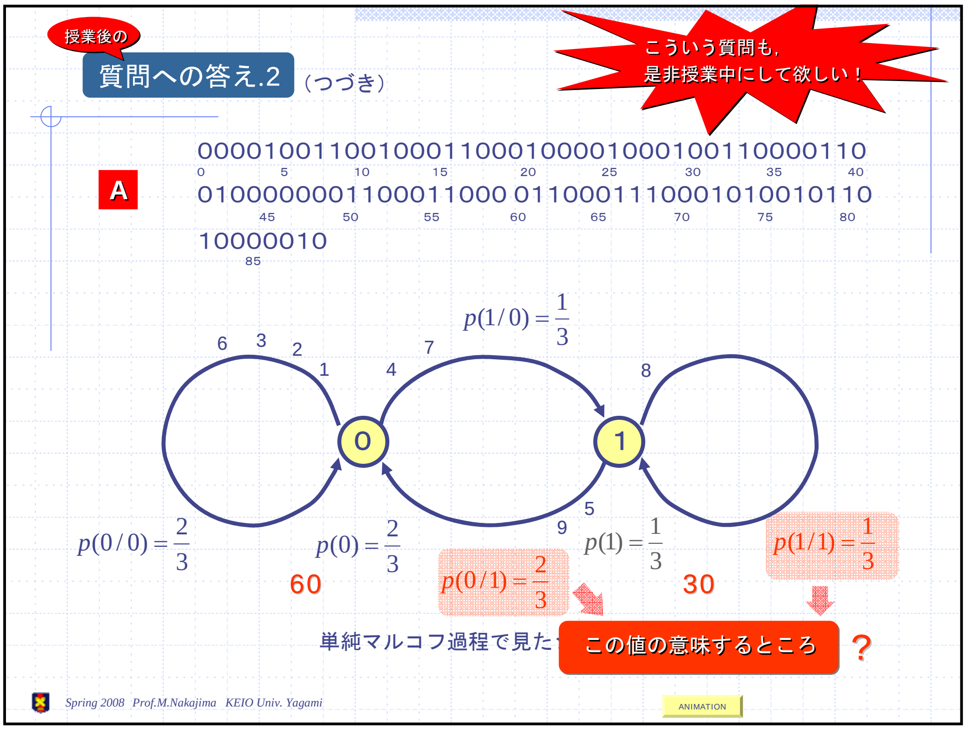


単純マル

この値の意味するところ ？

ANIMATION


授業後の
質問への答え.2
今年はなかったが，
昨年あった質問をついでに・・・
こういう質問も，
是非授業中にして欲しい！
A
00001001100100011000100001000100110000110
01000000011000110000110001110001010010110
10000010
この符号系列において、
‘0’の生起する確率
この符号系列において、
‘1’の生起する確率
$p(1/0)=\frac{1}{3}$
$p(0/0)=\frac{2}{3}$
$p(0)=\frac{2}{3}$
60
$p(0/1)=\frac{2}{3}$
$p(1)=\frac{1}{3}$
30
$p(1/1)=\frac{1}{3}$
単純マル
この値の意味するところ
？
Spring 2008 Prof.M.Nakajima KEIO Univ. Yagami
ANIMATION


授業後の
質問への答え.2（つづき）
こういう質問も，
是非授業中にして欲しい！
A
00001001100100011000100001000100110000110
01000000011000110000110001110001010010110
10000010
$p(1/0)=\frac{1}{3}$
$p(0/0)=\frac{2}{3}$
$p(0)=\frac{2}{3}$
60
$p(0/1)=\frac{2}{3}$
$p(1)=\frac{1}{3}$
30
$p(1/1)=\frac{1}{3}$
単純マルコフ過程で見た
この値の意味するところ
？
Spring 2008 Prof.M.Nakajima KEIO Univ. Yagami
ANIMATION

授業後の

質問への答え.2 （つづき）

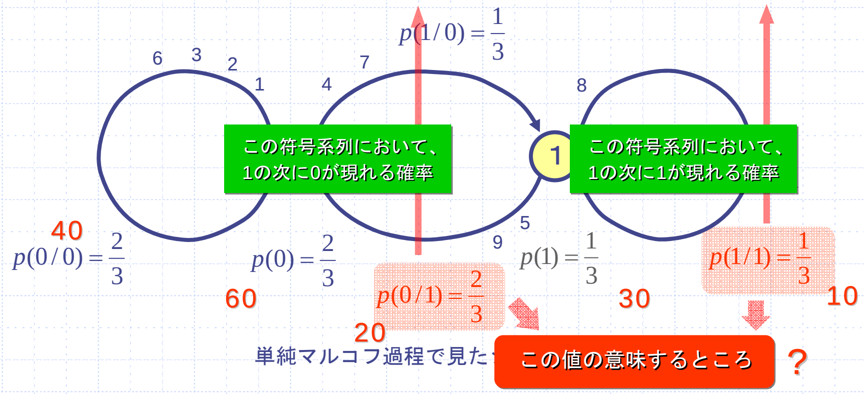
こういう質問も，
是非授業中にして欲しい！

A

00001001100100011000100001000100110000110

010000000110001100 0 0110001110001010010110

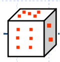
10000010



Example 3.2

復習の最後に簡単なクイズをだしてみよう．

（１）サイコロを振って，その目が何であったかを知らされた時に受け取るエントロピーは，いかほどになるか？



（略解）

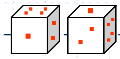
$$X=\left(\begin{array}{cccccc} ① & ② & ③ & ④ & ⑤ & ⑥ \\ \hline 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{array}\right)$$

$$H(X)=-\sum_{i=1}^{6}\frac{1}{6}log_2\frac{1}{6}=6\times\frac{1}{6}log_2\frac{1}{6}=2.5\ [bit]$$

Example 3.3

もうひとつ・・・

（２）２個のサイコロを振って，その目の和が何であったかを知らされた時に受け取るエントロピーは，いかほどになるか？



（略解）

$$Y = \begin{pmatrix} 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\ 1/36 & 2/36 & 3/36 & 4/36 & 5/36 & 6/36 & 5/36 & 4/36 & 3/36 & 2/36 & 1/36 \end{pmatrix}$$

$$H(Y) = -\frac{1}{36}\log_2\frac{1}{36} - \frac{2}{36}\log_2\frac{2}{36} - \frac{3}{36}\log_2\frac{3}{36} \cdots\cdots$$

かっこ良く書くと・・・

$$H(Y) = -2\sum_{i=1}^{6}\frac{i-1}{36}\log_2\frac{i-1}{36} - \frac{6}{36}\log_2\frac{6}{36}$$

$$= \cdot\cdot\cdot\cdot\cdot \quad = 3.3\ [bit]$$

ここからが，新しい話し

まず，情報量の式で，
対数の底が「２」になっている理由について

・事象が「0」と「1」の通報列（バイナリー符号列）を考えてみよう．

・前例題（天気予報の話）において，そのエントロピーが一番大きくなるのは，各天候の生起確率が‘共に1/4’の時であった．
したがって，確率事象系のエントロピーが最大となるのは，全事象の生起確率が，全て‘等確率’となる時であることが推測される．

・そうであるとすれば，「0」，「1」の２事象の場合，両者の生起確率が共に1/2の時にエントロピーは最大となるものと考えられ，その値は，

$$H = -\frac{1}{2}log_2\frac{1}{2} - \frac{1}{2}log_2\frac{1}{2} = 1\ [bit]$$

となる．

ここからが，新しい話し

まず，情報量の式で，
対数の底が「２」になっている理由について

・事

・前
各

地方 / 天気	X	Y	Z	W
F	1/2	7/8	1	1/4
C	1/4	0	0	1/4
R	1/8	1/8	0	1/4
S	1/8	0	0	1/4

$H_X = 1.8 \ [bit]$

$H_y = 0.5 \ [bit]$

$H_z = 0 \ [bit]$

$H_w = 2 \ [bit]$

最大！？

ANIMATION

ここからが，新しい話し

まず，情報量の式で，
対数の底が「２」になっている理由について

・事象が「0」と「1」の通報列(バイナリー符号列)を考えてみよう．

・前例題(天気予報の話)において，そのエントロピーが最も大きくなるのは，各天候の生起確率が'共に1/4'の時であった．
したがって，確率事象系のエントロピーが最大となるのは，全事象の生起確率が，全て'等確率'となる時であることが推測される．

・そうであるとすれば，「0」「1」の２事象の場合，両者の生起確率が共に1/2の時にエントロピーは最大となるものと考えられ，その値は，

$$H = -\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{2}\log_2\frac{1}{2} = 1 \ [bit]$$

となる．

・これが，理由である !!

要するに、シャノン先生は
0，1のバイナリー符号系列に対する
エントロピー最大値を
「１」にしたかったと思われる！

エントロピーの最大値

前の話から大方の予想はついたであろうが，以下‘エントロピーの最大値’について厳密に考えてみたいと思う．

答えから言ってしまうと，

$$\text{エントロピーの最大値} = \log_2(\text{事象数})$$

となる．

- N 個の事象 $\{a_i\}_{i=1}^{N}$ でできた確率事象系において不等式 $H \le \log_2 N$ の成り立つことが証明できれば良いことになる．
- 等号は、全ての事象が等確率で生起する時に成立する．

ANIMATION

証明について

方針：$H(x)$ の式の最大化

（１）微分によって極値を求める

（２）式の変形

・常識的
・しかし難しい
・不完全

・数学的
・式を知らなければできない！
・完全！

（２）の方法から先にやる！

（１）の方法については，事象数３の場合について説明する．一般的な証明は出来ない！

私には，

ANIMATION

証明(2)

$$H \leq log_2 N \qquad \underline{log_2 N - H \geq 0}$$

(1.1) ⇨ $\underline{\sum p_i = 1}$

(3.2) ⇨ $\underline{H = -\sum p_i log_2 p_i}$

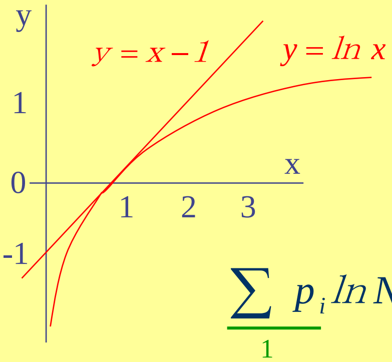
$$\underline{log_2 N - H} = log_2 N \underline{\sum p_i} + \underline{\sum p_i log_2 p_i}$$
$$= \sum p_i log_2 N + \sum p_i log_2 pi$$
$$= \sum p_i log_2 N p_i = *$$

ここで

$$ln x = log_e x = \frac{log_2 x}{log_2 e} \Rightarrow log_2 x = log_e x log_2 e$$

証明(2)

数学的準備： $x - 1 \geq ln x$ … 等式は、$x = 1$ の時



よって、$ln \frac{1}{x} \geq 1 - x$

$x = \frac{1}{Np_i}$ とすれば

$$\underbrace{\sum p_i ln Np_i}_{1} \geq \underbrace{\sum p_i}_{1} \left(1 - \frac{1}{Np_i}\right)$$

証明(2)

$$\log_2 N - H = \log_2 N \sum p_i + \sum p_i \log_2 p_i$$

$$= \sum p_i \log_2 N + \sum p_i \log_2 p_i$$

$$= \sum p_i \log_2 N p_i$$

$$\log_2 x = \log_e x \log_2 e$$

$$= \sum p_i \log_e N p_i \log_2 e$$

$$\sum p_i \ln N p_i \geq \sum p_i \left(1 - \frac{1}{N p_i}\right)$$

$$\geq \log_2 e \sum p_i \left(1 - \frac{1}{N p_i}\right)$$

$$= \log_2 e \left(\sum p_i - \frac{1}{N} \sum \frac{p_i}{p_i} \right) = 0$$

↑ 正　↑ 1　↑ 1

よって, $\log_2 N \geq H$

ANIMATION

クイズ.1

学籍番号末尾の数字が‘偶数’の人のみ、
そのまま教室に残って下さい.
偶数の人は、速やかに退室して下さい.

教室に残った人は、隣りの人との間に
一席空けて座って下さい.

回答時間は、10分間です.
終わった人は、回答用紙を裏返してそのまま
机の上に残し、静かに退出して下さい.

ANIMATION

クイズ.1

学籍番号末尾の数字が‘偶数’の人のみ、

（1）1個のコインを振って、表が出るか裏がでるかを見る時に得られる情報量の平均値を求めなさい。

（2）次に、同様のことを3個のコインで行った場合に得られる情報量の平均値を求めなさい。

2008年度

「情報工学」

第3回講義　おわり