

講義「情報理論」

第4回 情報源のモデル(前半)

情報理工学部門 情報知識ネットワーク研究室
喜田拓也

二つの確率変数のエントロピー(おさらい)

結合エントロピー:

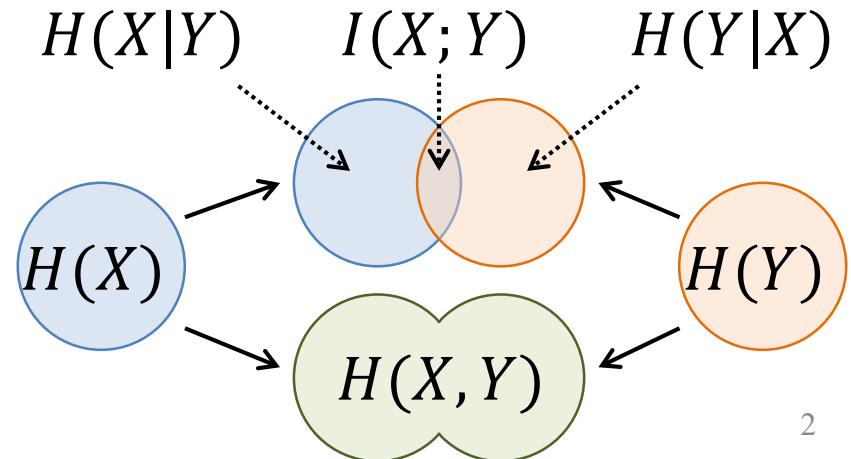
$$H(X, Y) = - \sum_{i=1}^{M_X} \sum_{j=1}^{M_Y} P(x_i, y_j) \log_2 P(x_i, y_j)$$

条件付きエントロピー:

$$H(X|Y) = - \sum_{j=1}^{M_Y} P(y_j) \sum_{i=1}^{M_X} P(x_i | y_j) \log_2 P(x_i | y_j)$$

相互情報量:

$$\begin{aligned} I(X; Y) &= H(X) - H(X|Y) \\ &= H(Y) - H(Y|X) \end{aligned}$$



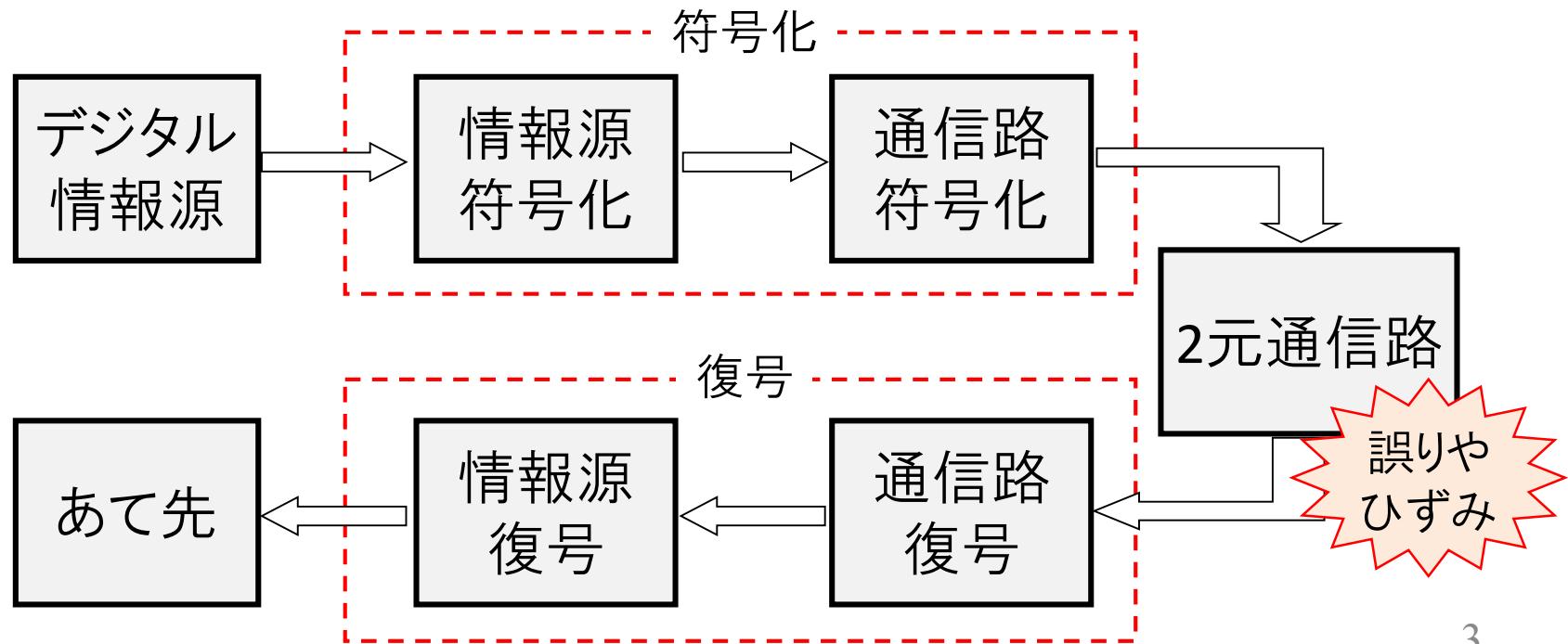
情報理論が取り組む4つの問題(おさらい)

【問題1】できるだけよい情報源符号化法(復号法)を見出すこと

【問題2】情報源符号化の限界を知ること

【問題3】できるだけよい通信路符号化法(復号法)を見出すこと

【問題4】通信路符号化の限界を知ること



今日の内容

3.1 情報源の統計的表現

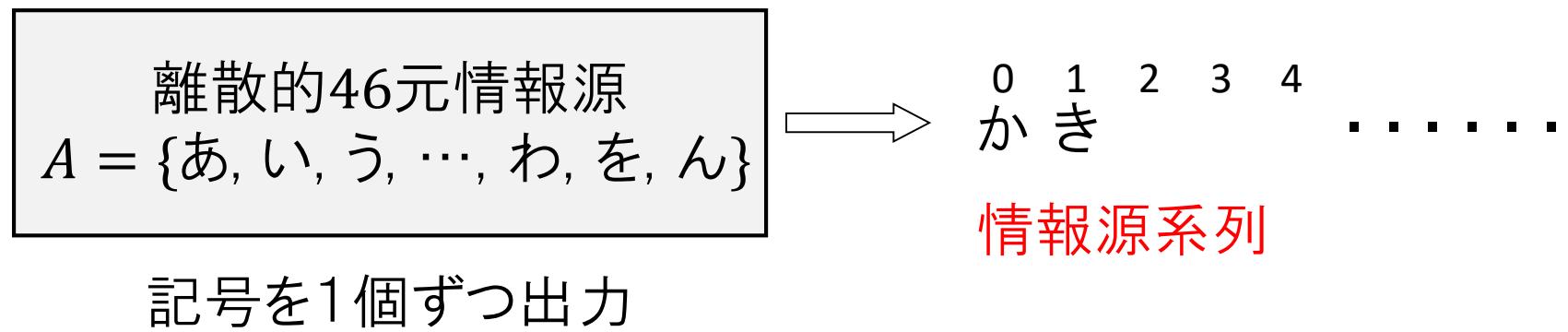
3.2 情報源の基本的なモデル

3.3 マルコフ情報源

離散的M元情報源

M 個の元からなる記号の有限集合 $A = \{a_1, a_2, \dots, a_M\}$ を考える。
これを**情報源アルファベット**とよび、各記号を**情報源記号**という

情報源は、時点0より毎時点(時点は整数値)において、情報源アルファベット A 上の記号をある確率に従って1個ずつ出力する。
このような情報源を、**離散的M元情報源**という



情報源系列の確率分布

時点0から時点 $N - 1$ まで(長さ N)の情報源系列を考える。
その際, 時点 i の情報源の出力を確率変数 X_i で表す

- X_i の取りうる値は, a_1, a_2, \dots, a_M のどれか
- 情報源系列を $X_0X_1X_2 \cdots X_{N-1}$ と書く

このとき, ある情報源系列 $x_0, x_1, x_2, \dots, x_{N-1}$ (ただし, $x_i \in A$)が
出力される確率は, X_0, X_1, \dots, X_{N-1} の結合確率分布で表される

X_0, X_1, \dots, X_{N-1} の結合確率分布

$$P_{X_0X_1\cdots X_{N-1}}(x_0, x_1, \dots, x_{N-1})$$

$$= [X_0 = x_0, X_1 = x_1, \dots, X_{N-1} = x_{N-1} \text{となる確率}]$$

情報源系列の
統計的性質を
決めるもの

部分系列の出現確率を求める(1)

情報源アルファベット $A = \{0, 1\}$

結合確率分布 $P(x_0, x_1, x_2)$ は表3.1のとおり

添え字を省略している
 $P_{X_0 X_1 X_2}(x_0, x_1, x_2)$ の意

結合確率分布から、時点0,1,2における
統計的性質は完全に定まる

例題3.1) $X_0 = 0$ となる確率を求める

$$\begin{aligned} P_{X_0}(0) &= \sum_{x_1=0}^1 \sum_{x_2=0}^1 P(0, x_1, x_2) \\ &= 0.008 + 0.032 \\ &\quad + 0.032 + 0.128 \\ &= 0.2 \end{aligned}$$

表3.1

x_0	x_1	x_2	$P(x_0, x_1, x_2)$
0	0	0	0.008
0	0	1	0.032
0	1	0	0.032
0	1	1	0.128
1	0	0	0.032
1	0	1	0.128
1	1	0	0.128
1	1	1	0.512

部分系列の出現確率を求める(2)

情報源アルファベット $A = \{0, 1\}$

結合確率分布 $P(x_0, x_1, x_2)$ は表3.1のとおり

結合確率分布から、時点0,1,2における
統計的性質は完全に定まる

例) $X_1 = 1, X_2 = 0$ となる確率を求める

$$\begin{aligned} P_{X_1 X_2}(1,0) &= \sum_{x_0=0}^1 P(x_0, 1, 0) \\ &= 0.032 + 0.128 \\ &= 0.16 \end{aligned}$$

添え字を省略している
 $P_{X_0 X_1 X_2}(x_0, x_1, x_2)$ の意

表3.1

x_0	x_1	x_2	$P(x_0, x_1, x_2)$
0	0	0	0.008
0	0	1	0.032
0	1	0	0.032
0	1	1	0.128
1	0	0	0.032
1	0	1	0.128
1	1	0	0.128
1	1	1	0.512

部分系列の出現確率を求める(3)

条件付確率分布: $P_{X_1|X_0}(x_1|x_0) = P_{X_0X_1}(x_0, x_1) / P_{X_0}(x_0)$

$X_0 = x_0$ という条件の下での $X_1 = x_1$ となる確率(割合)

$$\begin{aligned}P_{X_1|X_0}(0|0) &= \frac{P_{X_0X_1}(0,0)}{P_{X_0}(0)} \\&= \frac{0.008 + 0.032}{0.2} \\&= 0.2.\end{aligned}$$

X_0 と X_1 で条件を付けた X_2 の条件付確率分布

$$P_{X_2|X_0X_1}(x_2|x_0, x_1) = \frac{P(x_0, x_1, x_2)}{P_{X_0X_1}(x_0, x_1)}$$

表3.1

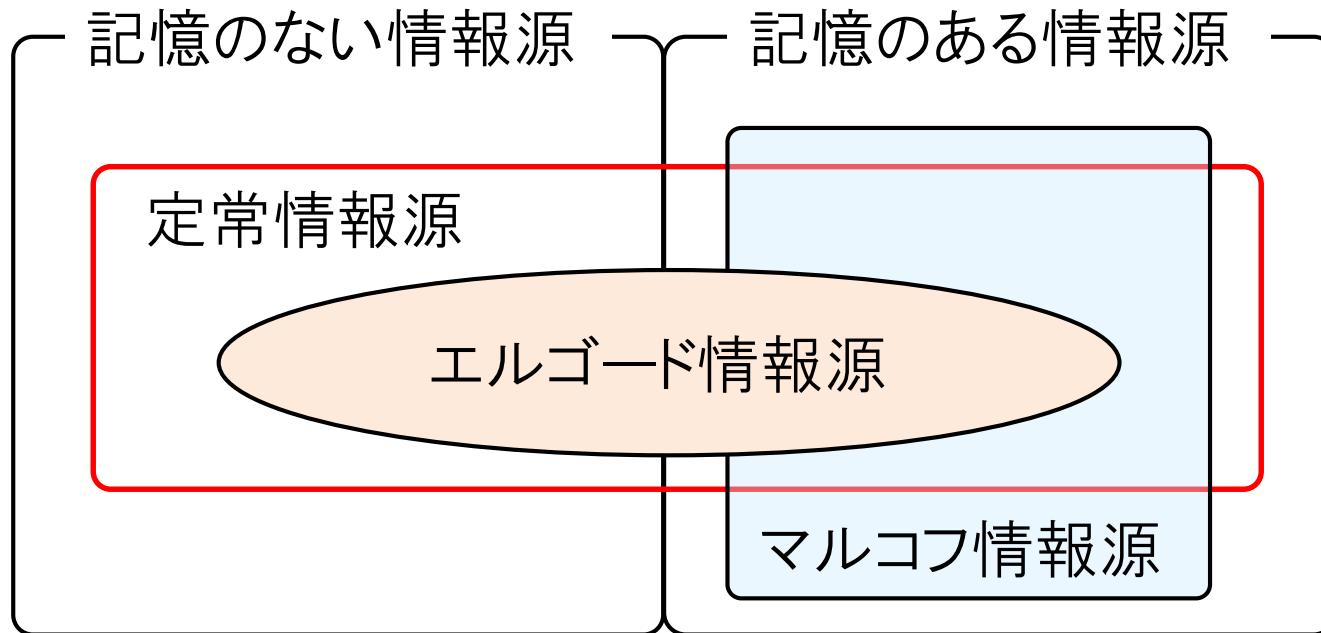
x_0	x_1	x_2	$P(x_0, x_1, x_2)$
0	0	0	0.008
0	0	1	0.032
0	1	0	0.032
0	1	1	0.128
1	0	0	0.032
1	0	1	0.128
1	1	0	0.128
1	1	1	0.512

Try 練習問題3.2

情報源のモデル

どんな大きい N についても, X_0, X_1, \dots, X_{N-1} の結合確率分布を与えることができれば, この情報源の統計的性質は完全に記述できる. しかし, **一般には困難**.

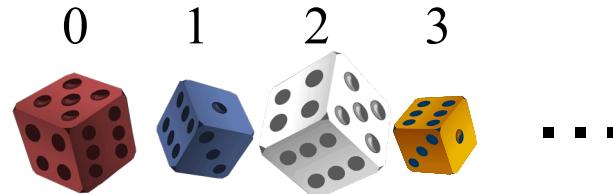
議論を進めるためには, 情報源に何らかの**扱いやすい性質**を仮定する必要がある.



記憶のない情報源 [定義3.1]

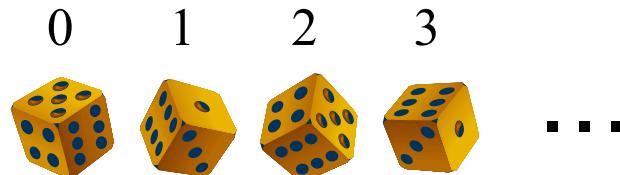
記憶のない情報源 (memoryless information source)

各時点における情報源記号の発生が、他の時点と独立である
情報源



記憶のない定常情報源 (i.i.d.(independently and identically distributed))

記憶のない情報源において、各時点における情報源記号の発生が同一の確率分布 $P_X(x)$ に従う情報源



時点 i によらないので
 X の添え字がない

記憶のない定常情報源の結合確率分布

記憶のない定常情報源における長さ n の系列の結合確率分布

$$P_{X_0 \dots X_{n-1}}(x_0, \dots, x_{n-1}) = \prod_{i=0}^{n-1} P_X(x_i)$$

例題3.3) さいころの目の出方を調べる. 目が $6 \rightarrow a$, それ以外 b を出力する記憶のない定常情報源 S が系列 $aaba$ を出力する確率はいくらか?

⇒ $A = \{a, b\}$ の 2 元情報源で, かつ a, b の発生確率が $1/6$ と $5/6$ となる記憶のない定常情報源になる. よって,

$$P(a, a, b, a) = \frac{1}{6} \times \frac{1}{6} \times \frac{5}{6} \times \frac{1}{6} = \frac{5}{1296}.$$

Try 練習問題3.3

(一般の)定常情報源の定義

定常情報源(stationary information source)

時間をずらしても、統計的性質の変わらない情報源

任意の正整数 n と i および情報源アルファベットの任意の元 x_0, x_1, \dots, x_{N-1} に対し、

$$P_{X_0 X_1 \dots X_{N-1}}(x_0, x_1, \dots, x_{N-1})$$

$$= P_{X_i X_{i+1} \dots X_{i+N-1}}(x_0, x_1, \dots, x_{N-1})$$

が成立するとき、この情報源を定常情報源という。

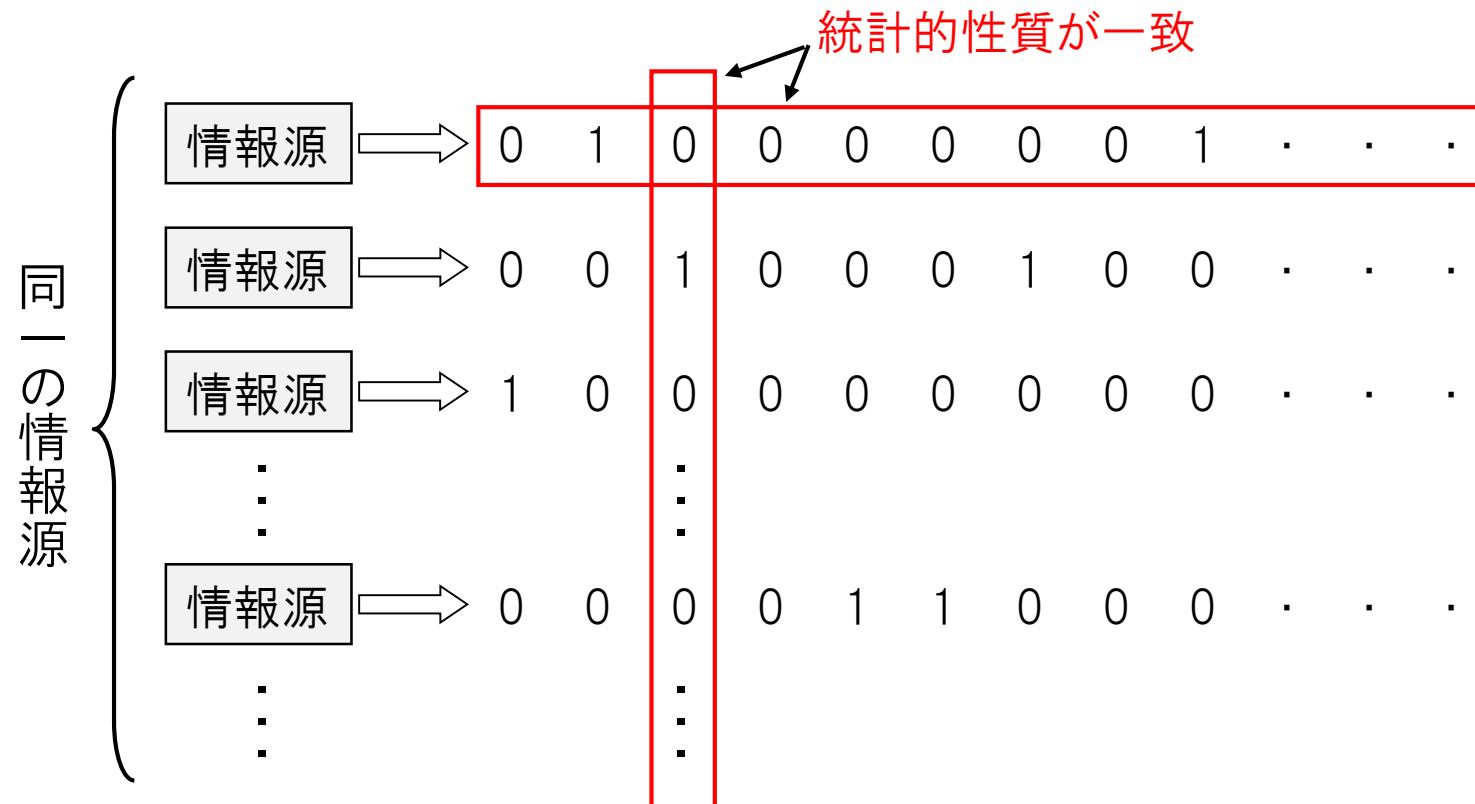
定常情報源の出力は、各時点において、同一の確率分布に従う。この確率分布を定常分布と呼ぶ

上式において、 $N = 1$ とおくと、 $P_{X_0}(x_0) = P_{X_i}(x_0)$ となる

エルゴード情報源

エルゴード情報源(ergodic information source)

それが発生する十分長い任意の系列に、その情報源の統計的性質が完全に現れている定常情報源



エルゴード情報源の性質

エルゴード情報源では、**集合平均と時間平均が一致する**

集合平均を求めるのは難しい。時間平均を観測するのは容易

定常分布が $P_X(x)$ である定常情報源の出力 X を変数とする任意の関数 $f(X)$ を考える。ただし $f(X)$ は実数値をとるものとする。

$f(X)$ の**集合平均**

$$\overline{f(X)} = \sum_{x \in A} f(x)P_X(x)$$

$f(X)$ の**時間平均**

$$\langle f(X) \rangle = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} f(X)$$

エルゴード情報源においては次が成り立つ

$$\overline{f(X)} = \langle f(X) \rangle$$

記憶のない定常情報源は
エルゴード情報源

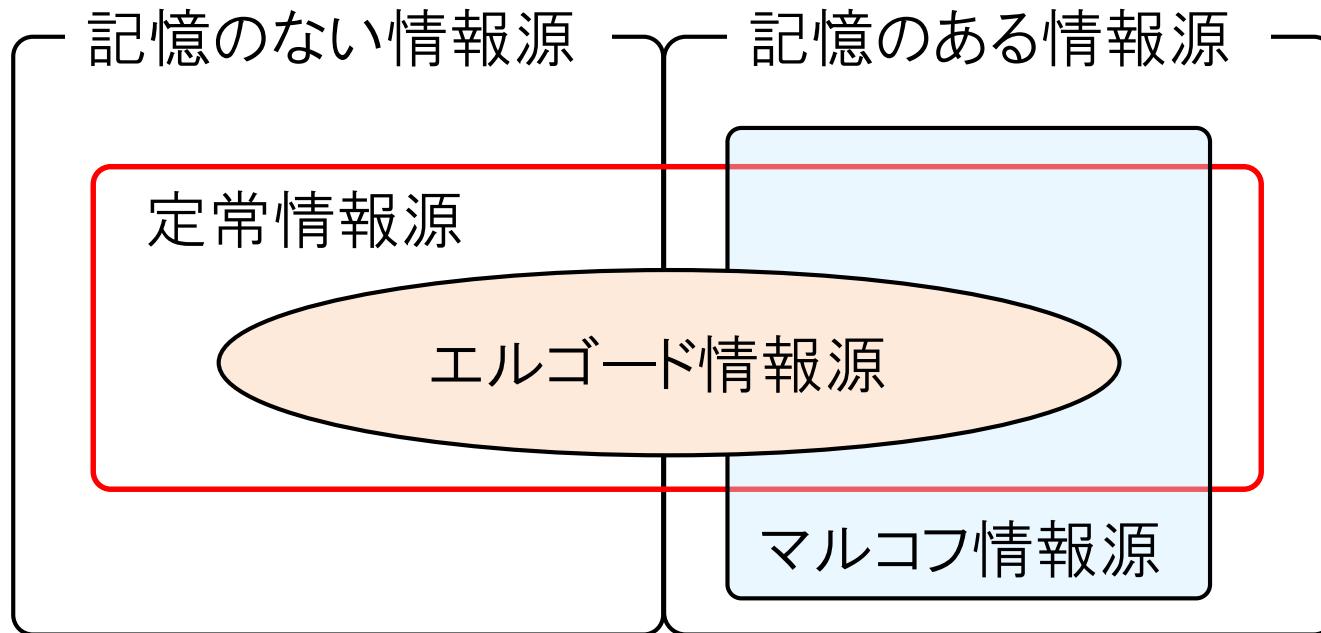
※ 定常情報源であってもエルゴード的でないものはいくらでもある。

例) 確率 $1/2$ で 0 だけからなる系列か 1 だけからなる系列を発生する情報源

情報源のモデル（再掲）

どんな大きい N についても, X_0, X_1, \dots, X_{N-1} の結合確率分布を与えることができれば, この情報源の統計的性質は完全に記述できる. しかし, **一般には困難**.

議論を進めるためには, 情報源に何らかの**扱いやすい性質**を仮定する必要がある.



ちょっと休憩

マルコフ情報源 [定義3.3]

記憶のある情報源の基本的なモデル

***m*重マルコフ情報源**(*m*-th order Markov information source):

N を m 以上の任意の整数とする

任意の時点 i について, その直前の N 個の出力で条件を付けた X_i の条件付確率分布が, 直前の m 個の出力だけで条件を付けた X_i の条件付確率分布と一致する, すなわち

$$\begin{aligned} & P_{X_i|X_{i-1}\dots X_{i-N}}(x_i|x_{i-1}, \dots, x_{i-N}) \\ &= P_{X_i|X_{i-1}\dots X_{i-m}}(x_i|x_{i-1}, \dots, x_{i-m}) \end{aligned}$$

となる.

系列中の過去 m 文字の並びにしたがって,
次の文字の生起確率が決まるモデル



Andrey Andreyevich Markov
(1856 – 1922, aged 66)

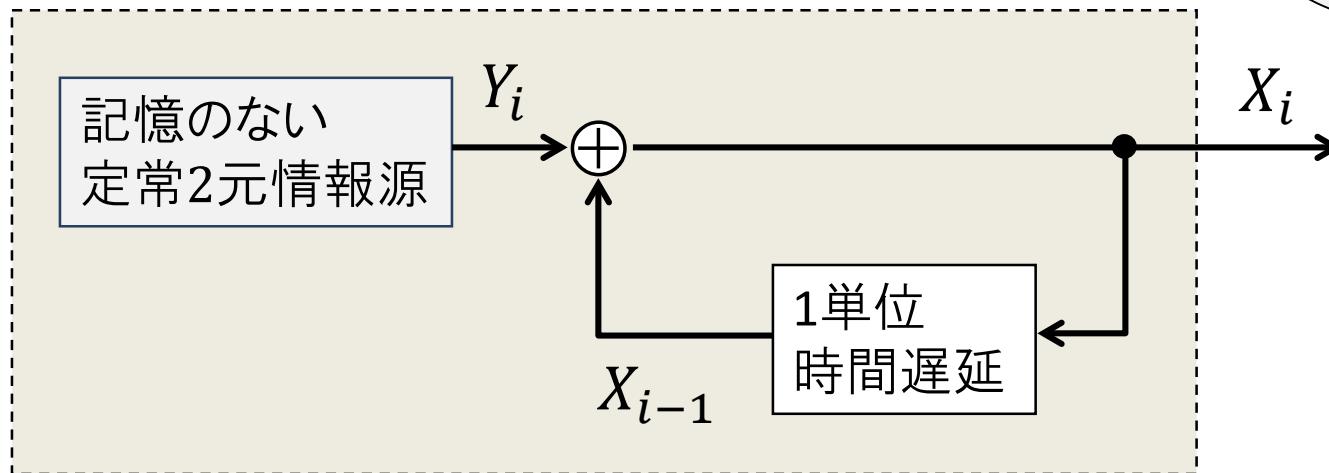
1重マルコフ情報源(単純マルコフ情報源)の例

例題3.4) 1, 0をそれぞれ確率0.2, 0.8で発生する記憶のない2元情報源の出力 Y_i と、1時点前のこの情報源の出力 X_{i-1} とから

$$X_i = X_{i-1} \oplus Y_i$$

により、現時点の出力 X_i が定まる2元情報源

⊕は、
排他的論理和



$$\begin{aligned} P_{X_i|X_{i-2}X_{i-1}}(0|00) &= P_{X_i|X_{i-2}X_{i-1}}(0|10) \\ &= P_{X_i|X_{i-1}}(0|0) = 0.8 \end{aligned}$$

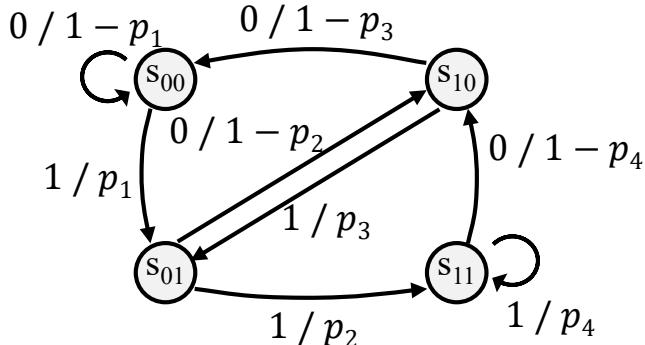
2つ前の出力には
左右されない

m 重マルコフ情報源の遷移図

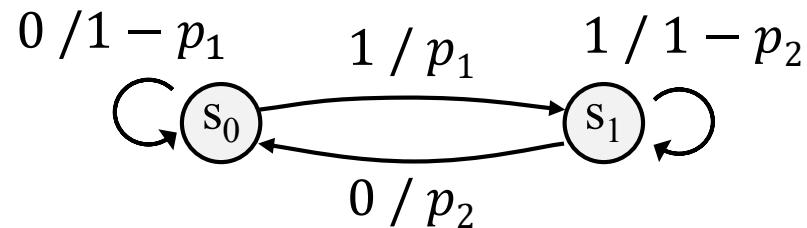
離散的 q 元情報源を考える

これが m 重マルコフ情報源だとする

- q^m 個の状態を持ち、各状態において記号の出力確率が定まっているとみなせる
- 出力を一つ発生するたびに q^m 個の状態の中を遷移していく

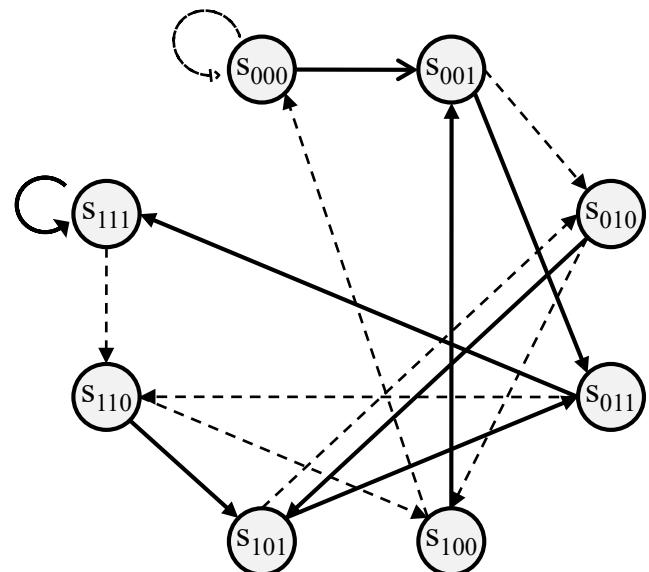


2重マルコフ情報源



1重マルコフ情報源

→ 1の遷移
---> 0の遷移
※確率は省略



3重マルコフ情報源

一般化されたマルコフ情報源

「直前の m 個の出力に対応する」という条件をなくし、より抽象的に（自由な状態図で）情報源を定義できる

図3.6はもはや m 重マルコフ情報源ではない

このような情報源を**一般化されたマルコフ情報源**という

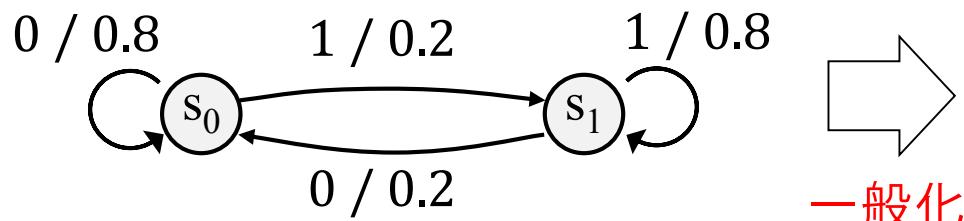


図3.5 例題3.4の状態図

マルコフ連鎖(Markov chain)
モデルとも言う

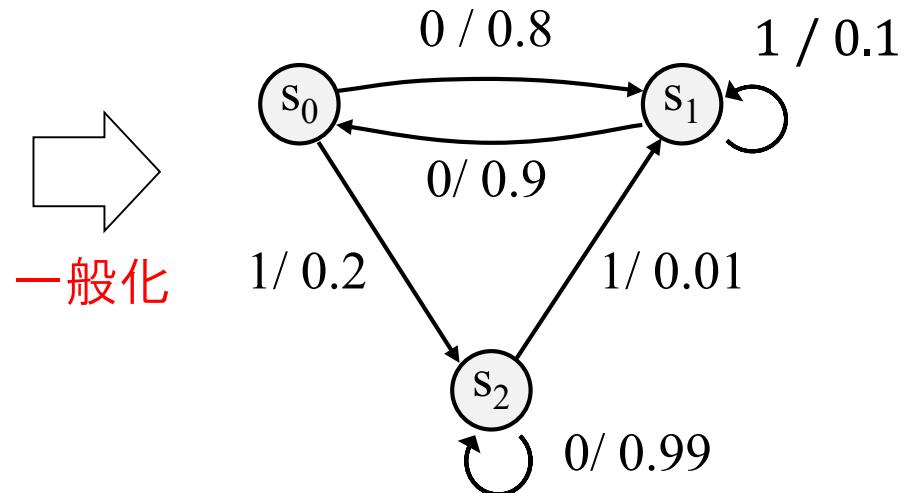
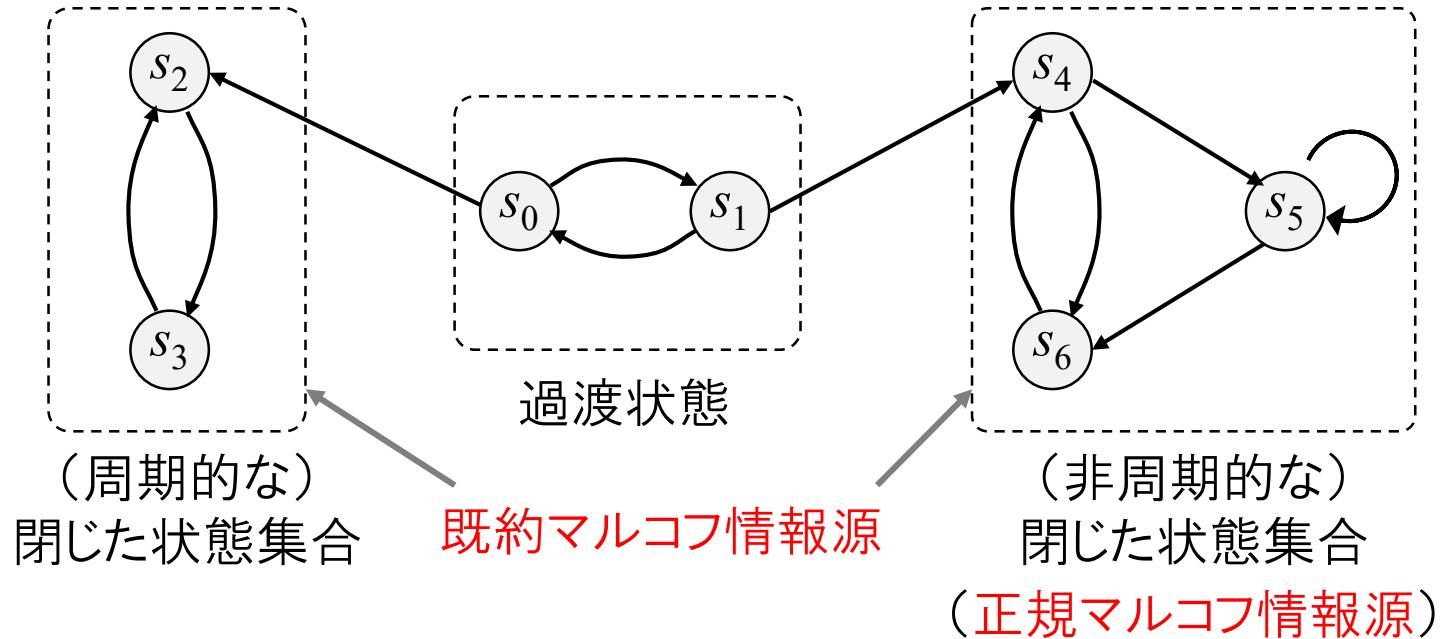


図3.6 一般化されたマルコフ情報源の
状態図

状態の分類



過渡状態: 十分時間が経過すれば抜け出てしまい、戻ることのない状態

閉じた状態集合: 任意の状態から任意の状態へ矢印をたどっていくことできる状態の集合

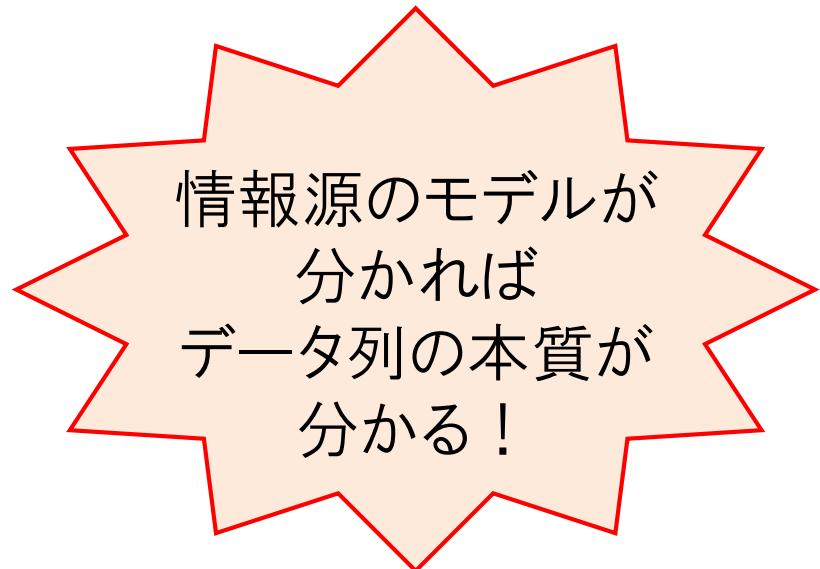
非周期的状態集合: 閉じた状態集合で、ある時間が経過した後の任意の時点において、どの状態にある確率も0でないようなもの

周期的状態集合: 閉じた状態集合がいくつかの部分集合に分割され、そのおのが周期的な時点においてのみ現れるもの

こんなモデルがどう役に立つのか？

次の3元情報源の系列を考える。これをうまく符号化したい。

う ん た ん う ん た ん う ん た ん た ん



今日のまとめ

情報源のモデル

情報源の統計的性質は結合確率分布によって決定される

一般的には結合確率分布を知ることは難しい

情報源の各種モデルについて

- 記憶のない情報源
- 定常情報源
- エルゴード情報源
- マルコフ情報源(m 重マルコフ, 一般化マルコフ, 正規マルコフ)

次回：情報源のモデル続き