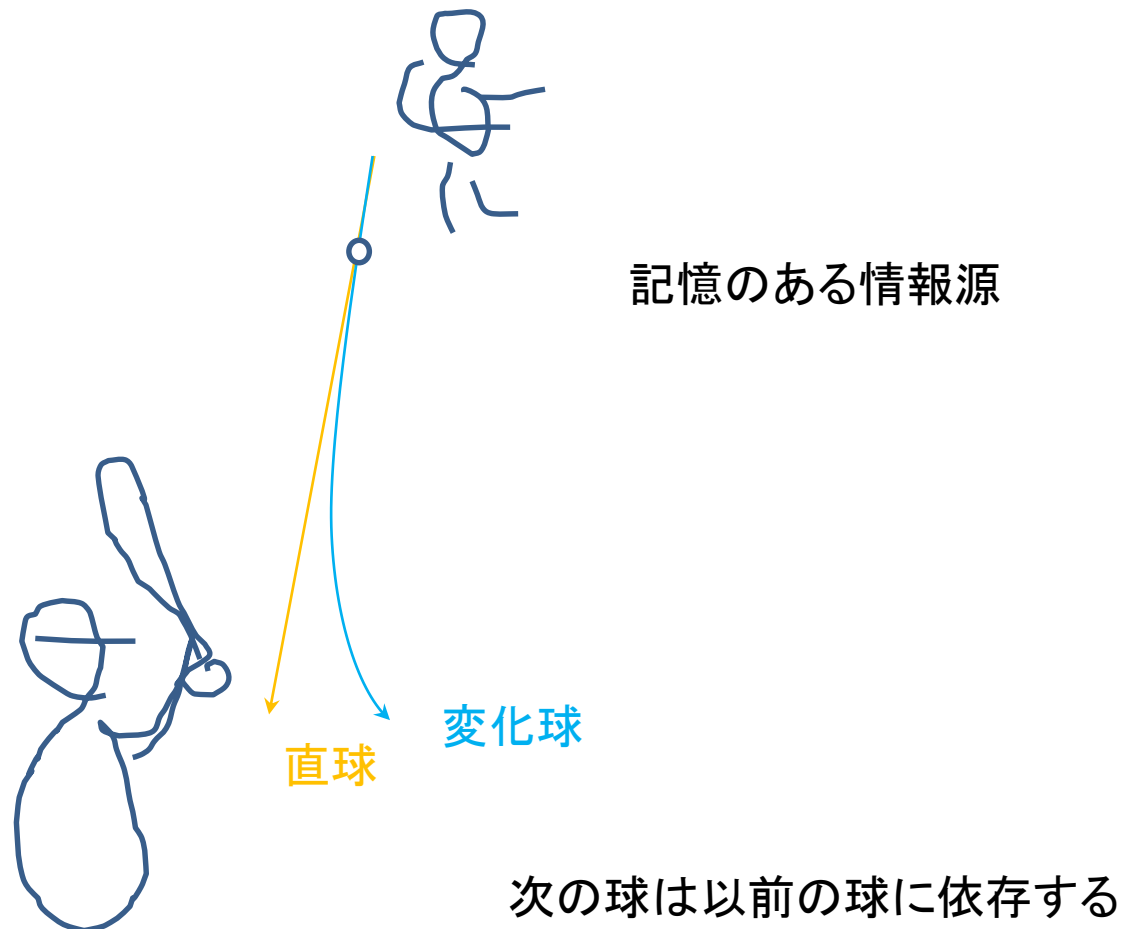


記憶のある情報源

記憶のある情報源とは？



マルコフ情報源

2重マルコフモデルを用いたモデル化

| x_0 | x_1 | x_2 | $P_{X_2 X_0,X_1}(x_2 x_0,x_1)$ |
|-------|-------|-------|--------------------------------|
| 直 | 直 | 直 | 0.6 |
| 直 | 直 | 変 | 0.4 |
| 直 | 変 | 直 | 0.2 |
| 直 | 変 | 変 | 0.8 |
| 変 | 直 | 直 | 0.5 |
| 変 | 直 | 変 | 0.5 |
| 変 | 変 | 直 | 0.6 |
| 変 | 変 | 変 | 0.4 |

x_0, x_1 の次に x_2 が出現する確率

マルコフ情報源

結合確率を用いた記述

| x_0 | x_1 | x_2 | $P_{X_0, X_1, X_2}(x_0, x_1, x_2)$ |
|-------|-------|-------|------------------------------------|
| 直 | 直 | 直 | 0.168 |
| 直 | 直 | 変 | 0.112 |
| 直 | 変 | 直 | 0.024 |
| 直 | 変 | 変 | 0.096 |
| 変 | 直 | 直 | 0.240 |
| 変 | 直 | 変 | 0.240 |
| 変 | 変 | 直 | 0.072 |
| 変 | 変 | 変 | 0.048 |

x_0, x_1, x_2 が順に
出現する確率

マルコフ情報源

結合確率からの条件付き確率の算出

$$\begin{aligned} P_{X_2|X_0,X_1}(\text{変}|\text{直}, \text{変}) &= \frac{P_{X_0,X_1,X_2}(\text{直}, \text{変}, \text{変})}{P_{X_0,X_1}(\text{直}, \text{変})} \\ &= \frac{P_{X_0,X_1,X_2}(\text{直}, \text{変}, \text{変})}{P_{X_0,X_1,X_2}(\text{直}, \text{変}, \text{直}) + P_{X_0,X_1,X_2}(\text{直}, \text{変}, \text{変})} \\ &= \frac{0.096}{0.024 + 0.096} \\ &= 0.8 \end{aligned}$$

マルコフ情報源

結合確率からのその他の確率の算出

$$\begin{aligned} P_{X_0}(\text{変}) &= \sum_{x_1} \sum_{x_2} P_{X_0, X_1, X_2}(\text{変}, x_1, x_2) \\ &= P_{X_0, X_1, X_2}(\text{変}, \text{直}, \text{直}) + P_{X_0, X_1, X_2}(\text{変}, \text{直}, \text{変}) \\ &\quad + P_{X_0, X_1, X_2}(\text{変}, \text{変}, \text{直}) + P_{X_0, X_1, X_2}(\text{変}, \text{変}, \text{変}) \\ &= 0.24 + 0.24 + 0.072 + 0.048 = 0.6 \end{aligned}$$

$$\begin{aligned} P_{X_1}(\text{変}) &= \sum_{x_0} \sum_{x_2} P_{X_0, X_1, X_2}(x_0, \text{変}, x_2) \\ &= P_{X_0, X_1, X_2}(\text{直}, \text{変}, \text{直}) + P_{X_0, X_1, X_2}(\text{直}, \text{変}, \text{変}) \\ &\quad + P_{X_0, X_1, X_2}(\text{変}, \text{変}, \text{直}) + P_{X_0, X_1, X_2}(\text{変}, \text{変}, \text{変}) \\ &= 0.024 + 0.096 + 0.072 + 0.048 = 0.24 \end{aligned}$$

マルコフ情報源

結合確率からのその他の確率算出

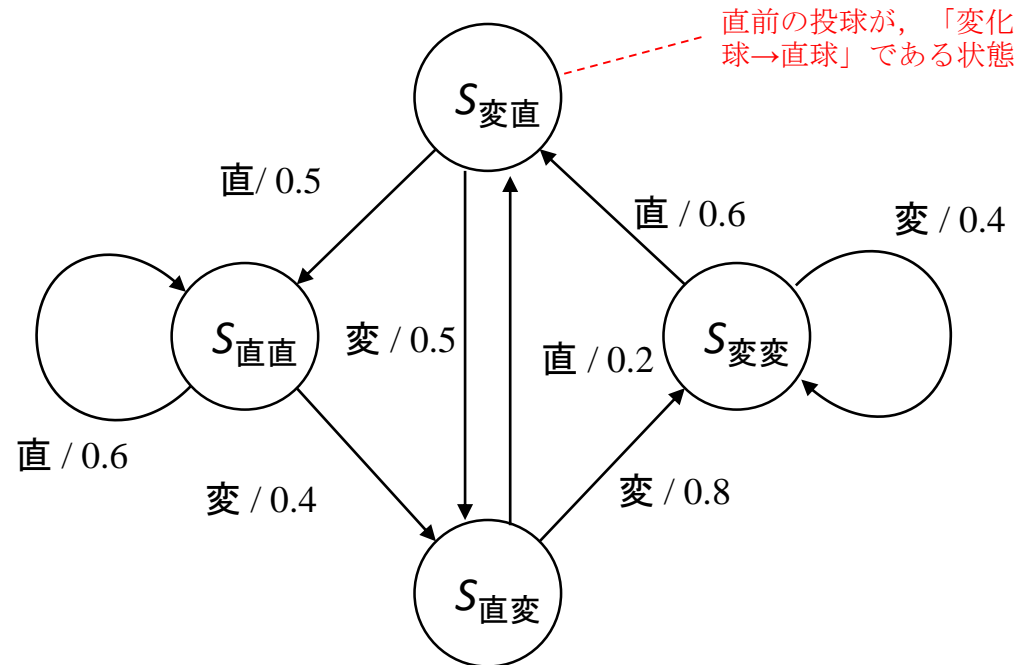
$$\begin{aligned} P_{X_2}(\text{変}) &= \sum_{x_0} \sum_{x_1} P_{X_0, X_1, X_2}(x_0, x_1, \text{変}) \\ &= P_{X_0, X_1, X_2}(\text{直}, \text{直}, \text{変}) + P_{X_0, X_1, X_2}(\text{直}, \text{変}, \text{変}) \\ &\quad + P_{X_0, X_1, X_2}(\text{変}, \text{直}, \text{変}) + P_{X_0, X_1, X_2}(\text{変}, \text{変}, \text{変}) \\ &= 0.112 + 0.096 + 0.24 + 0.048 = 0.496 \end{aligned}$$

マルコフ情報源

条件付確率モデル

状態図

| x_0 | x_1 | x_2 | $P(x_2 x_0, x_1)$ |
|-------|-------|-------|-------------------|
| 直 | 直 | 直 | 0.6 |
| 直 | 直 | 変 | 0.4 |
| 直 | 変 | 直 | 0.2 |
| 直 | 変 | 変 | 0.8 |
| 変 | 直 | 直 | 0.5 |
| 変 | 直 | 変 | 0.5 |
| 変 | 変 | 直 | 0.6 |
| 変 | 変 | 変 | 0.4 |



情報源のモデル化

- 情報源記号 $A = \{a_1, a_2, \dots, a_M\}$ をもつ情報源 S に対して, 時間の経過に伴う S からの出力を X_0, \dots, X_{n-1} と表す (X_i は $\{a_1, a_2, \dots, a_M\}$ のどれかをとり得る確率変数).
- 定常情報源: 時間をずらしても統計的性質は変わらない. つまり, 任意の正整数 n と i , 情報源記号の任意の元 x_0, x_1, \dots, x_{n-1} に対して,

$$P_{X_0, X_1, \dots, X_{n-1}}(x_0, x_1, \dots, x_{n-1}) = P_{X_i, X_{i+1}, \dots, X_{i+n-1}}(x_0, x_1, \dots, x_{n-1})$$

が成立する

記憶のある定常情報源モデル

- 記憶のない情報源: 各時点における情報源記号の発生が他の時点とは独立である. このとき, 次の等式が成立する.

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

- X_0, \dots, X_{n-1} の統計的性質は, あらゆる n に対して結合確率分布:

$$P_{X_0, \dots, X_{n-1}}(x_0, \dots, x_{n-1}) = \left[X_0 = x_0, X_1 = x_1, \dots, X_{n-1} = x_{n-1} \text{ となる確率} \right]$$

を与えれば完全に決まる.

- 確率が明示的に規定されていないものはベイズの原理で求める. 例えば,

$$P(x_2 | x_0, x_1) = \frac{P(x_0, x_1, x_2)}{P(x_0, x_1)} = \frac{P(x_0, x_1, x_2)}{P(x_0, x_1, x_2) + P(x_0, x_1, x_2)}$$

マルコフ情報源

- m 重マルコフ情報源: $m \leq n$ なる全ての n に対して,

$$P_{X_i|X_{i-n}, \dots, X_{i-1}}(x_i | x_{i-n}, \dots, x_{i-1}) = P_{X_i|X_{i-m}, \dots, X_{i-1}}(x_i | x_{i-m}, \dots, x_{i-1})$$

が成り立つ.

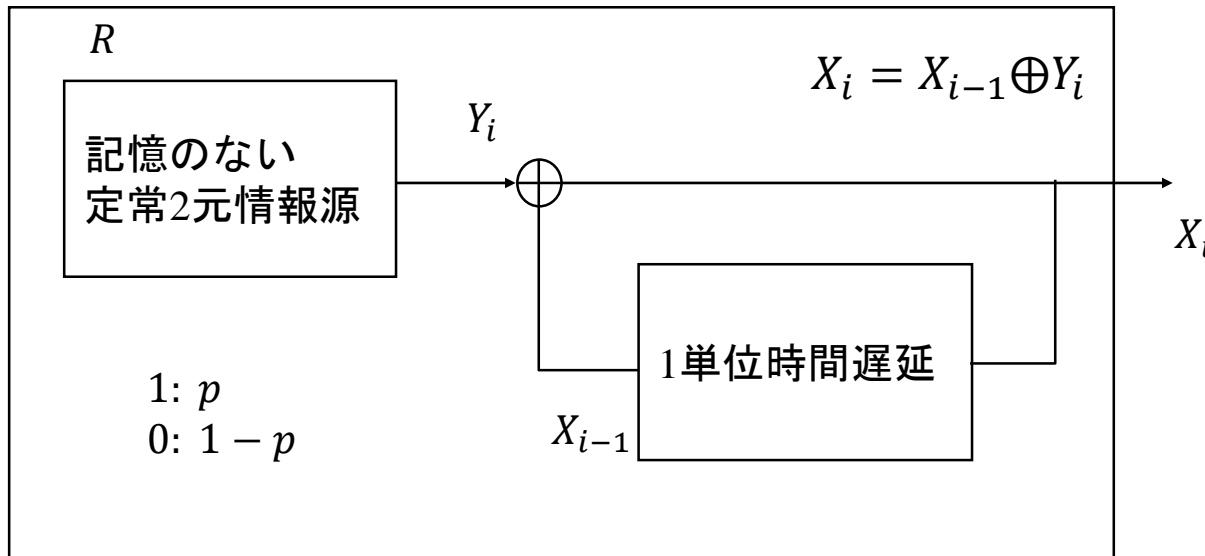
マルコフ情報源

- 一般化されたマルコフ情報源：状態図によって規定される.
- m 重マルコフ情報源は一般化されたマルコフ情報源としてモデル化できる.
- 一般化されたマルコフ情報源は必ずしも m 重マルコフ情報源としてモデル化できない.

マルコフ情報源

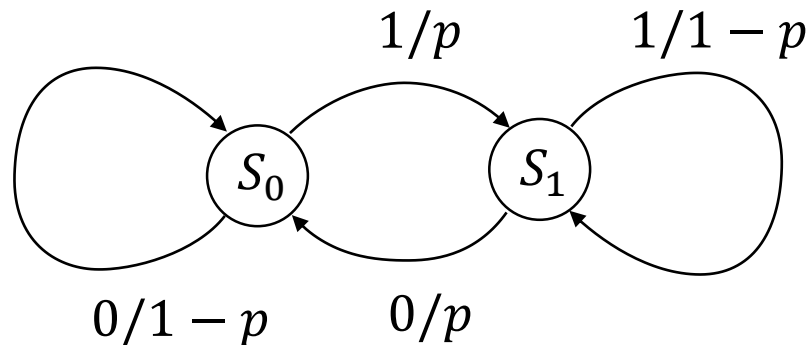
1重マルコフ情報源(単純マルコフ情報源)の例

S



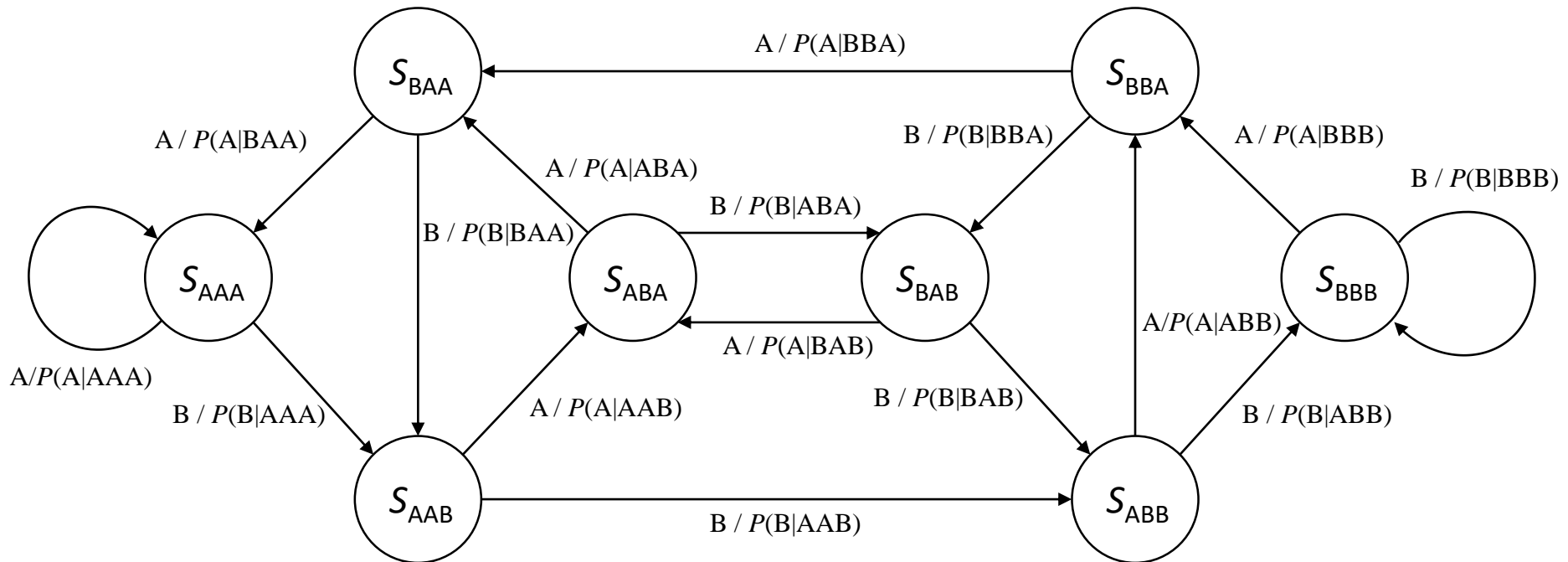
$$\begin{cases} X_i = X_{i-1} \oplus Y_i \\ P_{X_i|X_{i-1}}(0|0) = 1 - p \\ P_{X_i|X_{i-1}}(0|1) = p \\ P_{X_i|X_{i-1}}(1|0) = p \\ P_{X_i|X_{i-1}}(1|1) = 1 - p \end{cases}$$

状態図



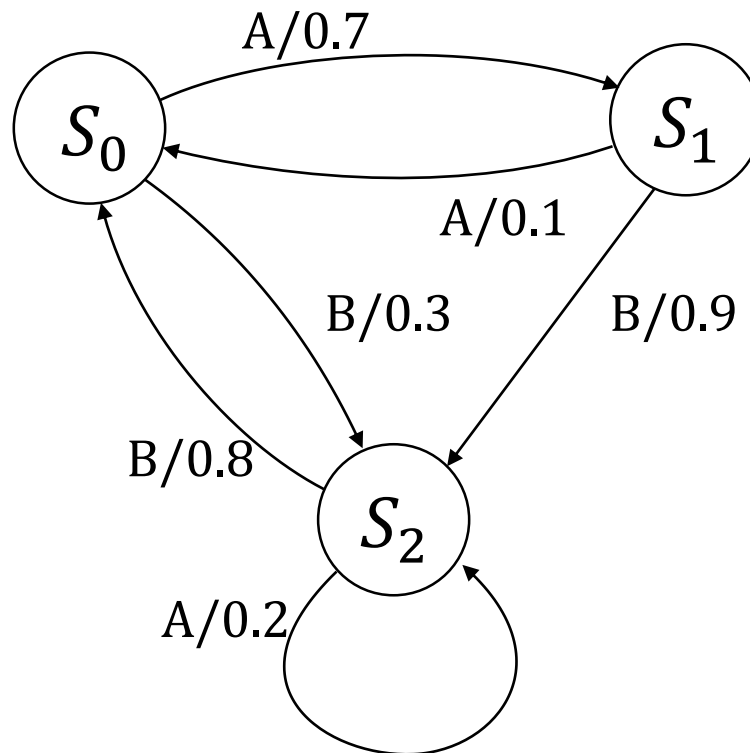
マルコフ情報源

(例) 3重マルコフ情報源の一般化されたマルコフ情報源としてのモデル化



マルコフ情報源

m 重マルコフ情報源としてモデル化できない一般化されたマルコフ情報源の例



エルゴード情報源

エルゴード情報源:

確率変数 X の値の集合平均 $\bar{X} = \sum_{x \in A} x P_X(x)$ と時間平均 $\langle X \rangle = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} x$ が等しい.

$P_X(x)$: ある時刻において $X = x$ になる確率

【エルゴード情報源のイメージ】

同一情報源のコピー

$S=\{A, B\}$

$S=\{A, B\}$

$S=\{A, B\}$

$S=\{A, B\}$

$S=\{A, B\}$

$S=\{A, B\}$

$S=\{A, B\}$

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | B | A | B | A | B | A | B | A | B | A | B | A | B | A | B |
| B | A | B | A | B | A | B | A | B | A | B | A | B | A | B | A |
| B | B | A | A | B | B | A | A | B | B | A | A | B | B | A | A |
| A | A | B | B | A | A | B | B | A | A | B | B | A | A | B | B |
| B | B | B | A | A | A | B | B | B | A | A | A | B | B | B | A |
| A | A | A | B | B | B | A | A | A | B | B | B | A | A | A | B |
| B | B | B | B | A | A | A | A | B | B | B | B | A | A | A | A |

集合平均

時間平均

等しい

t

エルゴード情報源

エルゴード情報源:

確率変数 X の値の集合平均 $\bar{X} = \sum_{x \in A} x P_X(x)$ と時間平均 $\langle X \rangle = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} x$ が等しい.

$P_X(x)$: ある時刻において $X = x$ になる確率

【非エルゴード情報源のイメージ】

同一情報源のコピー

$S=\{A, B\}$

$S=\{A, B\}$

$S=\{A, B\}$

$S=\{A, B\}$

$S=\{A, B\}$

$S=\{A, B\}$

$S=\{A, B\}$

A A A A A A A A A A A A A A A A A

B B B B B B B B B B B B B B B B B

A A A A A A A A A A A A A A A A A

A A A A A A A A A A A A A A A A A

B B B B B B B B B B B B B B B B B

B B B B B B B B B B B B B B B B B

A A A A A A A A A A A A A A A A A

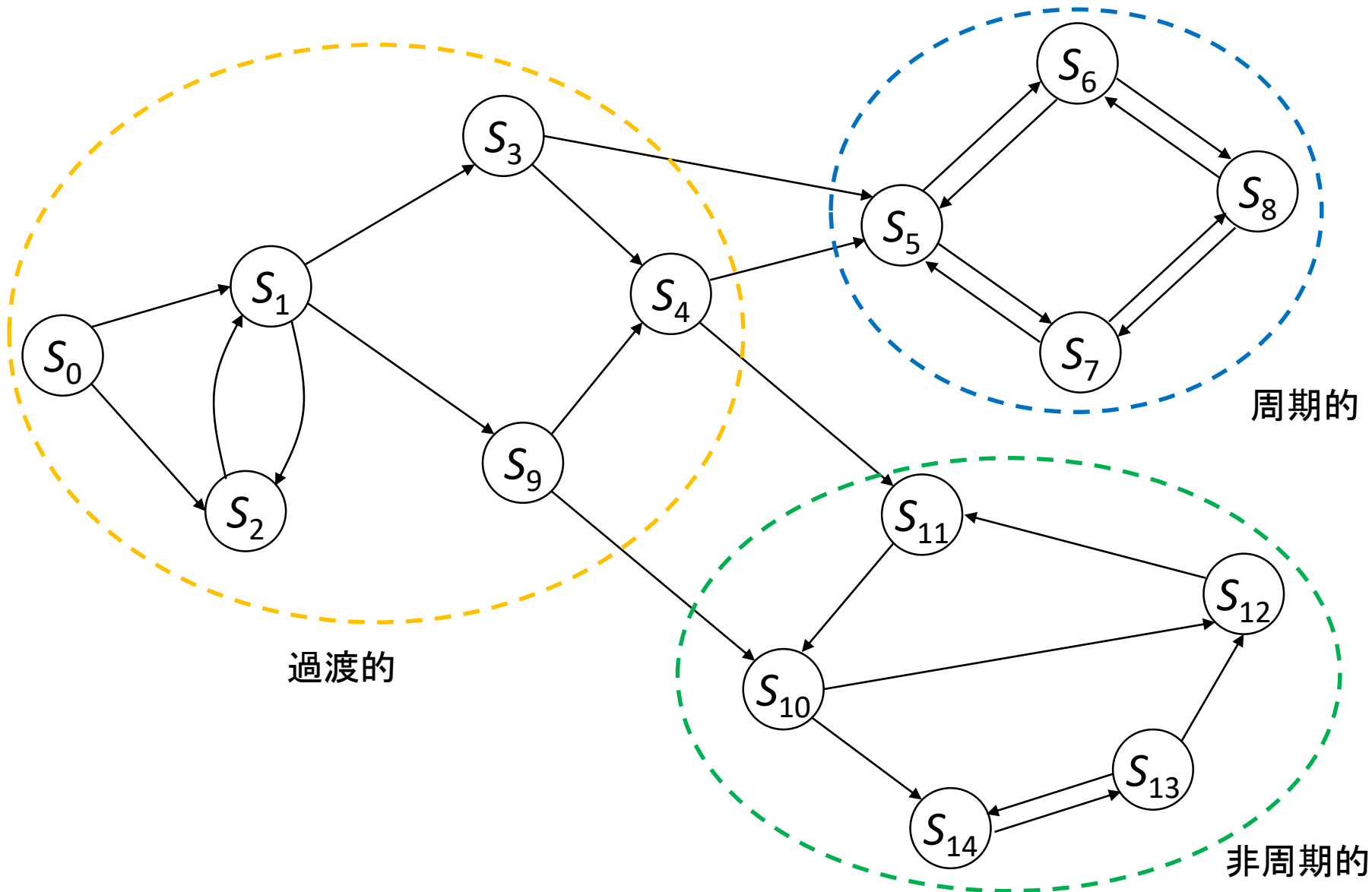
集合平均

異なる

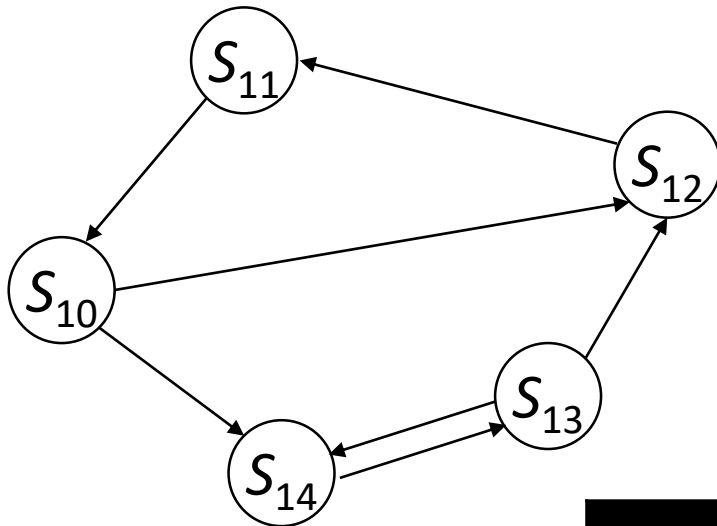
時間平均

t

マルコフ情報源



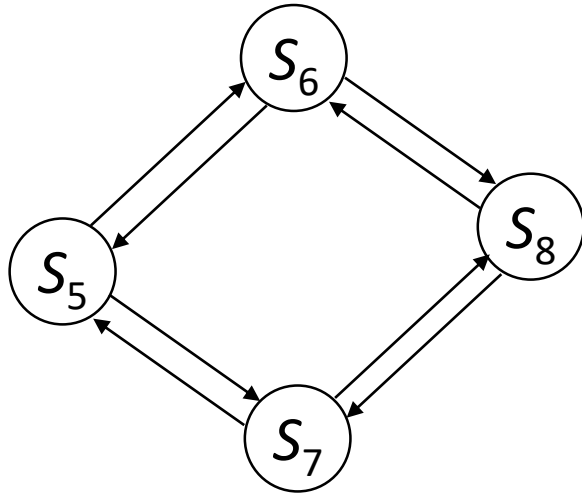
マルコフ情報源



非周期的: はじめにどのような状態であっても, ある時間が経過したあとの任意の時点において, どの状態になることも可能(確率は0ではない)

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|----------|---|---|---|---|---|---|---|---|
| S_{10} | ● | | | ● | | ● | ● | ● |
| S_{11} | | | ● | | ● | ● | ● | ● |
| S_{12} | | ● | | ● | ● | ● | ● | ● |
| S_{13} | | | ● | | ● | ● | ● | ● |
| S_{14} | | ● | | ● | ● | ● | ● | ● |

マルコフ情報源



周期的: どれだけ時間が経過しても, 任意の時点において, すべての状態に到達することはできない.

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|-------|---|---|---|---|---|---|---|---|
| S_5 | ● | | ● | | ● | | ● | |
| S_6 | | ● | | ● | | ● | | ● |
| S_7 | | ● | | ● | | ● | | ● |
| S_8 | | | ● | | ● | | ● | |

マルコフ情報源

- 正規マルコフ情報源では、はじめにどんな状態分布が与えられても、十分な時間が経過すれば、状態遷移は定常的になり、出力も定常的確率分布に従う。十分な時間の後には、エルゴード情報源とみなせる。
- 一般のマルコフ情報源では、定常分布は少なくとも一つ存在。初期分布として定常分布を与えると定常情報源となる。
- 規約情報源では、周期的であっても定常分布は一意的に存在する。それを初期分布とすればエルゴード情報源になる。
- 周期的規約情報源では、定常分布以外の初期分布を与えると、十分な時間が経過しても定常情報源にはならない。

マルコフ情報源

- 遷移確率行列:

$$\Pi = \begin{bmatrix} p_{0,0} & \cdots & p_{0,N-1} \\ \vdots & \ddots & \vdots \\ p_{N-1,0} & \cdots & p_{N-1,N-1} \end{bmatrix}$$

ここで, $p_{ij} = P(s_j | s_i)$: 状態 s_i の次に状態 s_j に遷移する確率.

- 遷移確率行列について, 次の性質が成立する.

$$\Pi^{(t)} = \begin{bmatrix} p_{0,0}^{(t)} & \cdots & p_{0,N-1}^{(t)} \\ \vdots & \ddots & \vdots \\ p_{N-1,0}^{(t)} & \cdots & p_{N-1,N-1}^{(t)} \end{bmatrix} = \Pi^t$$

ここで, $p_{ij}^{(t)}$ は, 状態 s_i から状態 s_j に t 時点後に遷移する確率.

マルコフ情報源

- 正規マルコフ情報源: 十分な時間が経過した後は, 任意の状態間の遷移が生じ得る情報源.
- 正規マルコフ情報源では, $t \rightarrow \infty$ のとき, $p_{ij}^{(t)}$ は i に無関係な値に収束する. つまり,

$$\lim_{t \rightarrow \infty} p_{ij}^{(t)} = u_j$$

- 定常分布 $w = (w_0, w_1, \dots, w_{N-1})$ の計算: 次の連立方程式を解く.

$$\begin{cases} w_0 + w_1 + \dots + w_{N-1} = 1 \\ w\Pi = w \end{cases}$$

マルコフ情報源

正規マルコフ情報源では, $t \rightarrow \infty$ のとき, $p_{ij}^{(t)}$ は i に無関係な値に収束する. すなわち,

$$\lim_{t \rightarrow \infty} p_{ij}^{(t)} = u_j$$

つまり,

$$\lim_{t \rightarrow \infty} \Pi^t = U = \begin{bmatrix} u_0 & \cdots & u_{N-1} \\ \vdots & \ddots & \vdots \\ u_0 & \cdots & u_{N-1} \end{bmatrix}$$

例

$$\Pi = \begin{bmatrix} 0.7 & 0.3 \\ 0.2 & 0.8 \end{bmatrix}$$

$$\Pi^2 = \begin{bmatrix} 0.55 & 0.45 \\ 0.3 & 0.7 \end{bmatrix}$$

$$\Pi^3 = \begin{bmatrix} 0.475 & 0.525 \\ 0.35 & 0.65 \end{bmatrix}$$

$$\Pi^4 = \begin{bmatrix} 0.4375 & 0.5625 \\ 0.375 & 0.625 \end{bmatrix}$$

$$\Pi^5 = \begin{bmatrix} 0.41875 & 0.58125 \\ 0.3875 & 0.6125 \end{bmatrix}$$

$$\Pi^6 = \begin{bmatrix} 0.409375 & 0.590625 \\ 0.39375 & 0.60625 \end{bmatrix}$$

$$\Pi^7 = \begin{bmatrix} 0.404688 & 0.595313 \\ 0.396875 & 0.603125 \end{bmatrix}$$

$$\Pi^8 = \begin{bmatrix} 0.402344 & 0.597656 \\ 0.38438 & 0.601563 \end{bmatrix}$$

$$\Pi^9 = \begin{bmatrix} 0.401172 & 0.598828 \\ 0.399219 & 0.600781 \end{bmatrix}$$

$$\Pi^{10} = \begin{bmatrix} 0.400586 & 0.599414 \\ 0.399609 & 0.600391 \end{bmatrix}$$

マルコフ情報源

定常分布 $w = (w_0, w_1, \dots, w_{N-1})$ の計算: 次の連立方程式を解く.

$$\begin{cases} w_0 + w_1 + \dots + w_{N-1} = 1 \\ w\Pi = w \end{cases}$$

例 次の例については,

$$\begin{cases} w_{\text{直直}} + w_{\text{直変}} + w_{\text{変直}} + w_{\text{変変}} = 1 \\ 0.6w_{\text{直直}} + 0.5w_{\text{変直}} = w_{\text{直直}} \\ 0.4w_{\text{直直}} + 0.5w_{\text{変直}} = w_{\text{直変}} \\ 0.2w_{\text{直変}} + 0.6w_{\text{変変}} = w_{\text{変直}} \\ 0.8w_{\text{直変}} + 0.4w_{\text{変変}} = w_{\text{変変}} \end{cases}$$

ないしは,

$$\begin{cases} w_{\text{直直}} + w_{\text{直変}} + w_{\text{変直}} + w_{\text{変変}} = 1 \\ \begin{pmatrix} w_{\text{直直}}, w_{\text{直変}}, w_{\text{変直}}, w_{\text{変変}} \end{pmatrix} \begin{bmatrix} 0.6 & 0.4 & 0 & 0 \\ 0 & 0 & 0.2 & 0.8 \\ 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0.6 & 0.4 \end{bmatrix} = \begin{pmatrix} w_{\text{直直}}, w_{\text{直変}}, w_{\text{変直}}, w_{\text{変変}} \end{pmatrix} \end{cases}$$

を解いて, $w_{\text{直直}} \approx 0.273, w_{\text{直変}} \approx 0.218, w_{\text{変直}} \approx 0.218, w_{\text{変変}} \approx 0.291$ を得る (上で, 1個の等式は冗長).

マルコフ情報源のエントロピー

一般マルコフ情報源のエントロピー計算法:

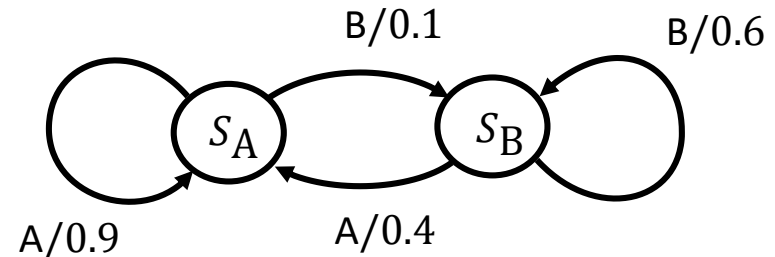
- 情報源記号: $\{a_1, \dots, a_M\}$
- 状態: $\{S_0, \dots, S_{N-1}\}$
- その定常確率分布: $\{w_0, \dots, w_{N-1}\}$
- 状態 S_i での情報源記号 a_i の発生確率: $P(a_i|S_i)$ とすれば, マルコフ情報源のエントロピーは,

$$H(S) = \sum_{i=0}^{N-1} w_i \left(- \sum_{j=1}^M P(a_j|S_i) \log_2 P(a_j|S_i) \right)$$

マルコフ情報源のエントロピー

(例) マルコフ情報源

$$\begin{cases} P_{X_i|X_{i-1}}(A|A) = 0.9 \\ P_{X_i|X_{i-1}}(B|A) = 0.1 \\ P_{X_i|X_{i-1}}(A|B) = 0.4 \\ P_{X_i|X_{i-1}}(B|B) = 0.6 \end{cases}$$



定常的な状態分布を (w_A, w_B) , 状態が S_A である確率を w_A , S_B である確率を w_B とすると

$$\begin{cases} w_A + w_B = 1 \\ (w_A, w_B) \begin{bmatrix} 0.9 & 0.1 \\ 0.4 & 0.6 \end{bmatrix} = (w_A, w_B) \end{cases}$$

を解いて, $w_A = 0.8, w_B = 0.2$ を得る.

状態が S_A であるときのエントロピーは, $H_{S_A}(S) = \mathcal{H}(0.1) \approx 0.469$ である.

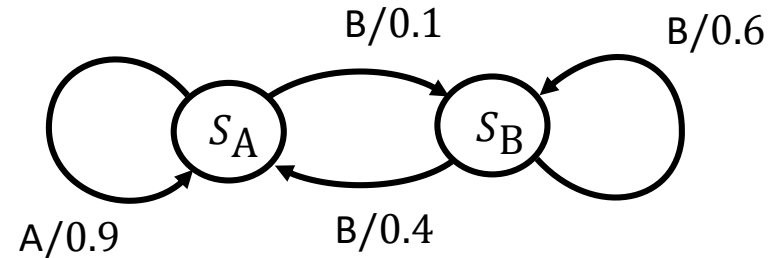
状態が S_B であるときのエントロピーは, $H_{S_B}(S) = \mathcal{H}(0.4) \approx 0.971$ である.

それぞれの状態の確率に応じてブレンドすると, $H(S) = 0.8H_{S_A}(S) + 0.2H_{S_B}(S) \approx 0.5694$

マルコフ情報源のエントロピー

(例) マルコフ情報源

$$\begin{cases} P_{X_i|X_{i-1}}(A|A) = 0.9 \\ P_{X_i|X_{i-1}}(B|A) = 0.1 \\ P_{X_i|X_{i-1}}(A|B) = 0.4 \\ P_{X_i|X_{i-1}}(B|B) = 0.6 \end{cases}$$



$$H_1(S) = \mathcal{H}(0.8) \approx 0.722$$

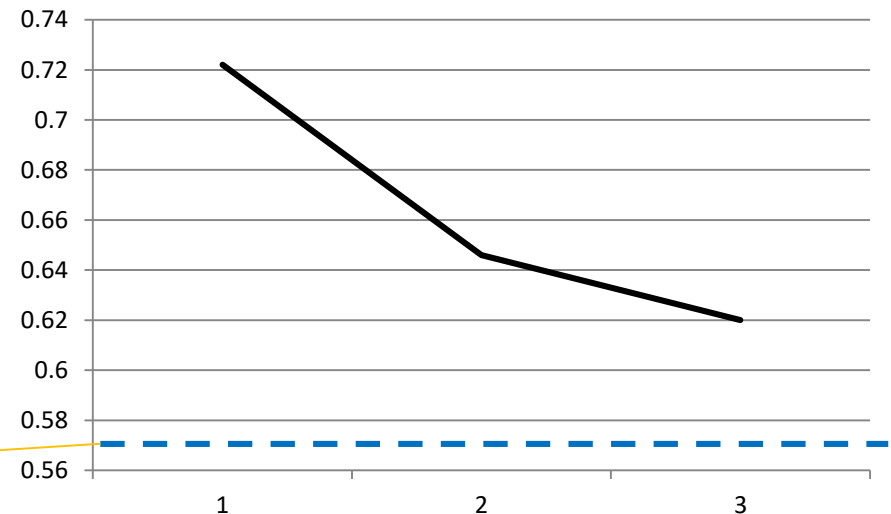
$$H_2(S) = \frac{H_1(S^2)}{2} \approx 0.646$$

$$H_3(S) = \frac{H_1(S^3)}{3} \approx 0.620$$



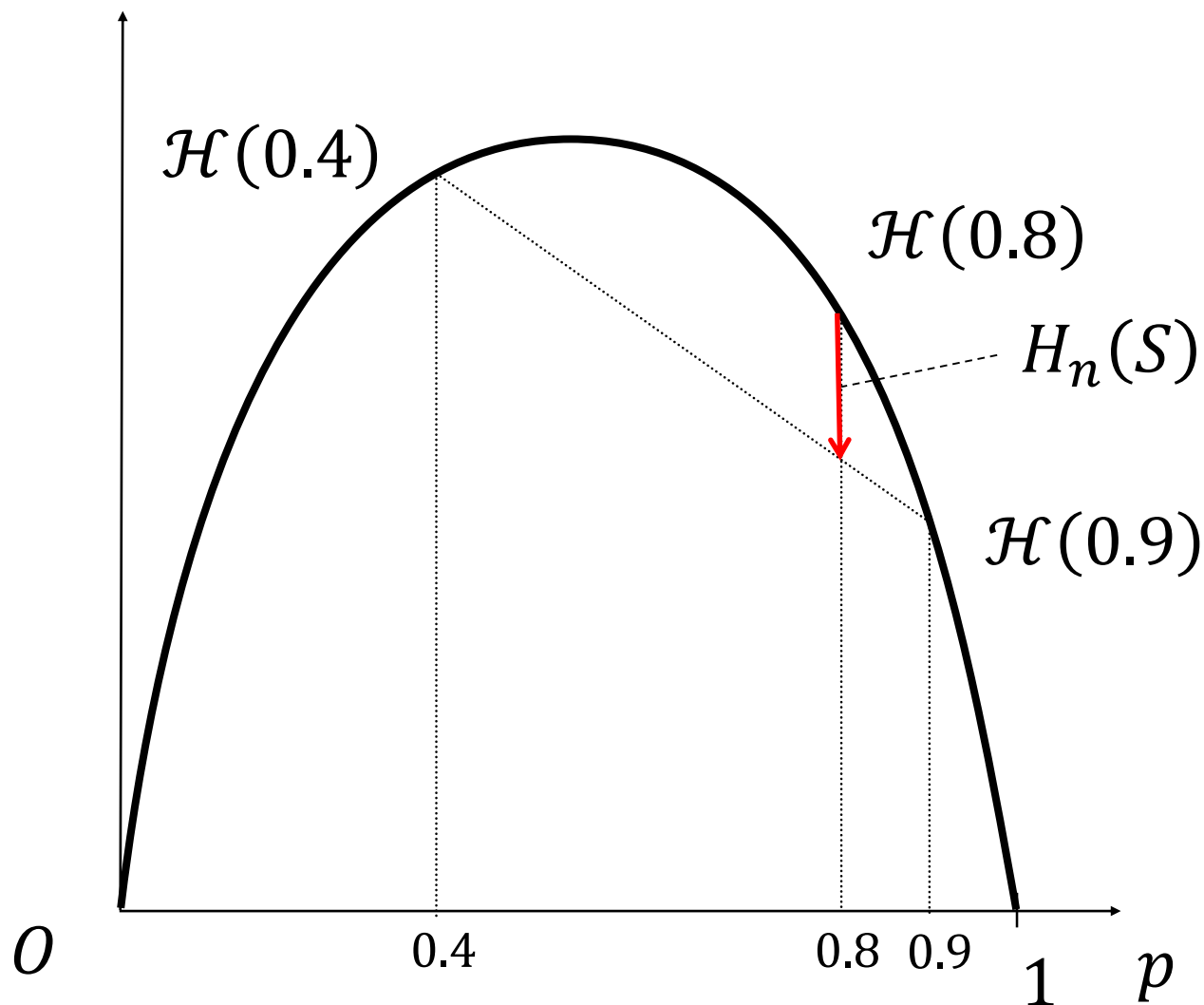
$$H(S) \approx 0.5694$$

$H_n(S)$



n

n 次エントロピーの背後にある構造



マルコフ情報源のエントロピー

- 正規マルコフ情報源のエントロピーと n 次エントロピーの間に次の関係がある.

$$H(S) \leq H_n(S)$$

$$\lim_{n \rightarrow \infty} H_n(S) = H(S)$$

まとめ

- 記憶を持つ情報源とは？
- 条件付確率を使ったモデル化
- マルコフ情報源モデル
- エルゴード性
- 状態の分類, マルコフ情報源モデルの分類
- マルコフ情報源の遷移確率行列表現
- 正規マルコフ情報源モデルの漸近的性質
- 正規マルコフ情報源の定常分布の計算法
- 一般マルコフ情報源のエントロピー計算法
- マルコフ情報源の n 次エントロピー