# 2016年度　通信理論

## 第4回　データ圧縮（イライアス符号）

2016. 4.18

植之原　裕行 uenohara.h.aa@m.titech.ac.jp

# 講義スケジュール(前半)

| | 日付 | 教科書 | 内容 |
|---|---|---|---|
| **第1回** | **4/ 7** | 第1章・第2章 | 通信システム概論:<br>通信伝送のモデルと具体例 |
| 第2回 | **4/11** | 第3章 | 情報源符号化<br>(情報量, エントロピー, 情報源符号化定理) |
| 第3回 | **4/14** | 第4章2節 | データ圧縮<br>(情報源符号化, 語頭符号, ハフマン符号) |
| 第4回 | **4/18** | **第4章3節～4節** | **データ圧縮(イライアス符号)** |
| 第5回 | **4/21** | 第4章5節～6節 | データ圧縮(ZL符号, ランレングス符号) |
| 第6回 | **4/25** | 第5章 | 通信路容量<br>(通信路の確率モデル, 相互情報量, 通信路容量) |
| 第7回 | **4/28** | 第6章1節～2節 | 通信路符号化<br>(通信路符号化, 通信路符号化定理, 繰返し符号) |
| 第8回 | **5/ 2** | 第1章～<br>第6章2節 | 通信理論に関する演習 |

# 講義スケジュール(後半)

| | 日付 | 教科書 | 内容 |
|---|---|---|---|
| 第９回 | 5/ 9 | 第６章３、４節 | 誤り訂正符号(線形ブロック符号) |
| 第１０回 | 5/12 | 第６章５節 | 誤り訂正符号(巡回符号, 畳み込み符号) |
| 第１１回 | 5/16 | 第６章６節 | 誤り訂正符号(LDPC符号) |
| 第１２回 | 5/19 | 第7章 | 連続情報と連続通信路 |
| 第１３回 | 5/23 | 第８章１節～２節 | 暗号の基礎理論 |
| 第１４回 | 5/26 | 第８章３節～７節 | 暗号の例(共通鍵暗号、ハッシュ関数) |
| 第１５回 | 5/30 | 第８章３節～７節 | 暗号の例(公開鍵暗号) |
| | 6/ 2 | | 期末試験(予定) |

# イライアス符号（Elias Code）

- 系列長オーダの計算量で構成可能な符号

ハフマン符号

→ アルファベットサイズ*M*, 符号化系列長*n*に対して
計算量 ∝ $M^n$
メモリサイズ ∝ $M^n \log_2 M^n$

イライアス符号

→ 符号化系列長*n*に比例する程度の実数演算で実行可能

（参考）
・漸近的最良性を有する。
漸近的最良性：十分長いシンボル系列に対して、
平均符号長が情報源エントロピーに漸近する性質。

# イライアス符号の基本的アイデア（1）

二元情報源　$A = \{a_0 = 0, a_1 = 1\}$

発生確率　$p_0 = P(a_0), p_1 = P(a_1)$

情報源系列　${x_1}^n \in A^n$　（*n*次拡大情報源を考える）

$n = 1$　のとき、$0, 1$　の2つ

$n = 2$　のとき、$00, 01, 10, 11$　の4つ

$n = 3$　のとき、$000, 001, \cdots, 111$　の8つ

➡ *n*桁の２進整数

発生確率

$n = 1$　$p_0 = P(a_0)$

$p_1 = P(a_1)$

$\vdots$　$\vdots$

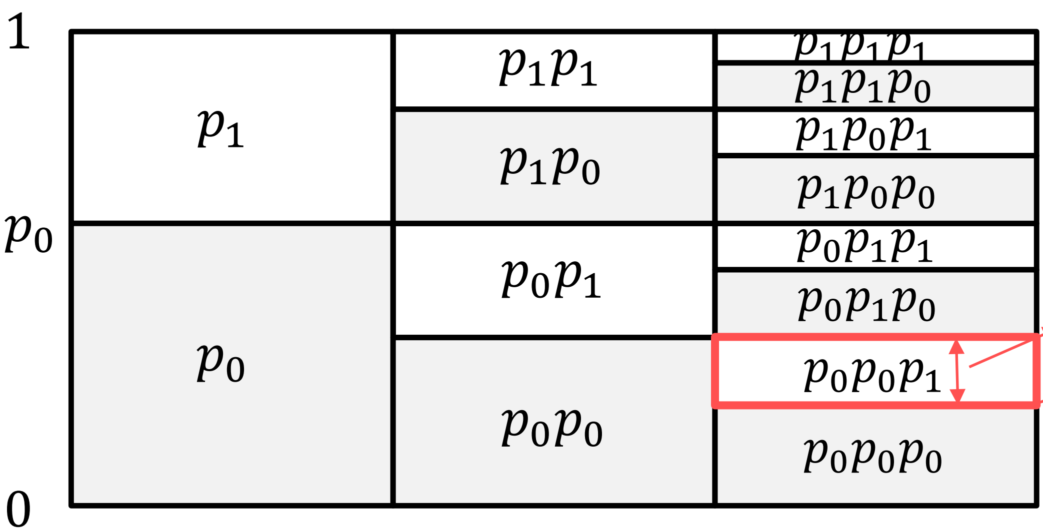
$n = 3$　${p_0}^3 = P(000)$

$\vdots$

➡ $P({x_1}^n) = \prod_{i=1}^{n} P(x_i)$　$, x_1 \in A$

# イライアス符号の基本的アイデア(2)

- n次拡大情報源系列$x_1{}^n$を、その表す整数の短い順に並べ、その発生確率を高さとする積み木を描く。




$p_1p_1p_1 + S_7 = S_8 = (p_0 + p_1)^3$
$p_1p_1p_0 + S_6 = S_7$
$p_1p_0p_1 + S_5 = S_6$
$p_1p_0p_0 + S_4 = S_5$
$p_0p_1p_1 + S_3 = S_4$
$p_0p_1p_0 + S_2 = S_3$
$p_0p_0p_1 + S_1 = S_2$
$p_0p_0p_0 + S_0 = S_1$
$0 = S_0$

(001)は区間$[p_0^3, p_0^2)$内の$\lceil -\log_2 p_0^2 p_1 \rceil$+1桁の2進少数を符号語とする

- 長さ$n$の系列 $x_1{}^n \in A^n$ を $[0,1)$ を覆う交わりのない小区間に一対一に対応付ける。

（例） $x_1{}^3 = 000 \ \rightarrow [0, p_0{}^3)$
$x_1{}^3 = 111 \ \rightarrow [1 - p_1{}^3, 1)$

小区間に属する

$$l_i = \left\lceil -\log_r \prod_{i=1}^{n} P(x_i) \right\rceil + 1$$

桁の$r$進小数を符号語とする。

（⌈ ⌉:⌈ ⌉内の数以上の最小整数）

# イライアス符号の符号化アルゴリズム（1）

各シンボルに対応する区間を求めるための初期化

① 情報源シンボル発生確率 $\{P(a_i)\}_{i=0}^{M-1}$ に対して

$$F(a_j) = \sum_{s=0}^{j-1} P(a_s), j = 0, 1, \cdots, M$$

符号に応じた区間の開始点を決める

とおき、

$$Q_0 = 0$$

符号を考える区間の開始の値（初期値）

$$w_0 = 1$$

区間の大きさ（初期値）

とする。

ただし、

$F(a_0) = 0, F(a_M) = 1$ の階段関数。

（2元情報源の例）

$$F(a_0) = 0, F(a_1) = P(a_0), F(a_2) = P(a_0) + P(a_1) = 1$$

# イライアス符号の符号化アルゴリズム(2)

各シンボルに対応する区間を求める

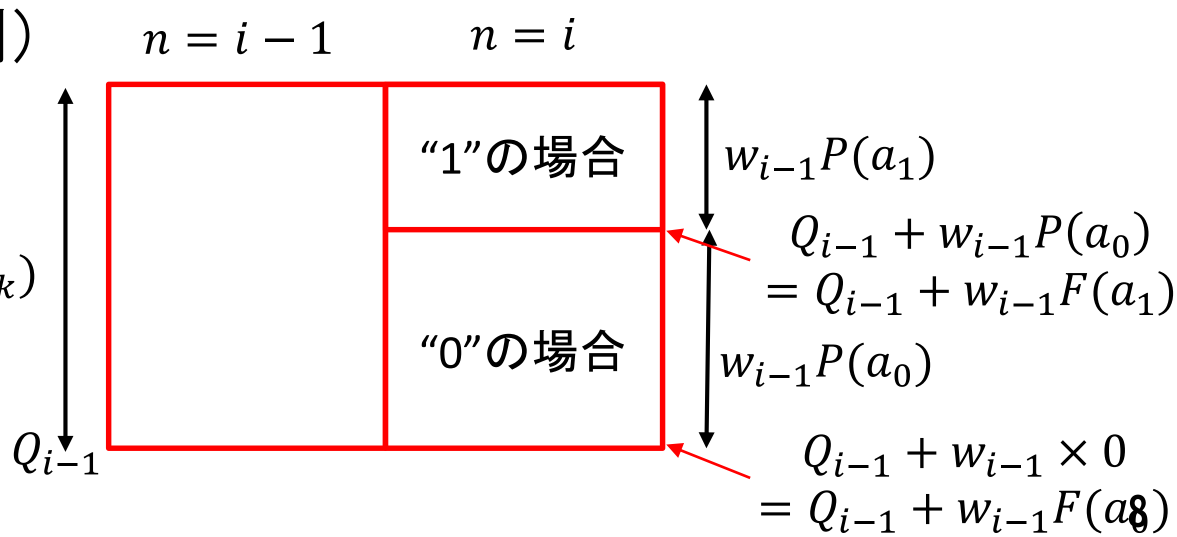
② 長さ$n$の系列 $x_1{}^n = x_1 x_2 \cdots x_n \in A^n$ に対して
$Q_i, w_i$ を以下の規則で$n$まで順に求める。

$$Q_i = Q_{i-1} + \underline{w_{i-1}}\,\underline{F(x_i)}$$ 区間の開始の値

$$w_i = w_{i-1} P(x_i)$$ 区間の大きさ

（2元情報源の例）

$$w_{i-1} = \prod_{k=1}^{i-1} P(x_k)$$



# イライアス符号の符号化アルゴリズム（３）

区間内の指定桁数の符号語を求める

③ 区間　$[Q_n, Q_n + w_n)$ にて、$r$進表現において小数点以下

$$l = \lceil -\log_r w_n \rceil + 1 = \left\lceil -\log_r \prod_{i=1}^{n} P(x_i) \right\rceil + 1$$

桁（ $l$ +1桁以下は0）で表される小数

$$c = [0.c_1 c_2 \cdots c_l]_r$$

が存在し、 $[c, c + r^{-l}) \subset [Q_n, Q_n + w_n)$　が成立。

$c_l = c_1 c_2 \cdots c_l$　を符号語とする。

# イライアス符号の復号化アルゴリズム（1）

符号語の存在する区間を特定すれば、その区間に対応する元の符号が求まる。

- 長さ$n$の情報源系列に対するイライアス符号の最大符号語長は

$$l_n = \left\lceil -n\log_r P^{\min} \right\rceil + 1$$

$$P_{\min} = \min_{0 \le s \le M-1} P(a_s)$$

で与えられる。

長さ$l_0$の符号系列$c_1 c_2 \cdots c_l c_{l+1} \cdots c_{l_0} \left(\in B^{l_0}\right)$を受け取って $z_0 := \emptyset$（空シンボル）として以下の手順により復号系列 $z_1 z_2 \cdots z_n \in A^n$を定める。

- ただし $c_{l+1} \cdots c_{l_0}$ は $c_l$ が符号系列の終わりの場合空シンボル、そうでなければ後に続く符号語の先頭部分になる。

# イライアス符号の復号化アルゴリズム（2）

① $$F(a_j) = \sum_{s=0}^{j-1} P(a_s), j = 0, 1, \cdots, M$$

に定め、

$$\widetilde{Q_0} = 0, \widetilde{w_0} = 1, P(z_0) = 1, F(z_0) = 0$$

とする。

② $i = 1, 2, \cdots, n$ に対して、以下を行う（$Q_i, w_i, z_i$が求まる）。

$$\widetilde{Q_i} = \widetilde{Q_{i-1}} + \widetilde{w_{i-1}} F(z_{i-1})$$

$$\widetilde{w_i} = \widetilde{w_{i-1}} P(z_{i-1})$$

とおき、

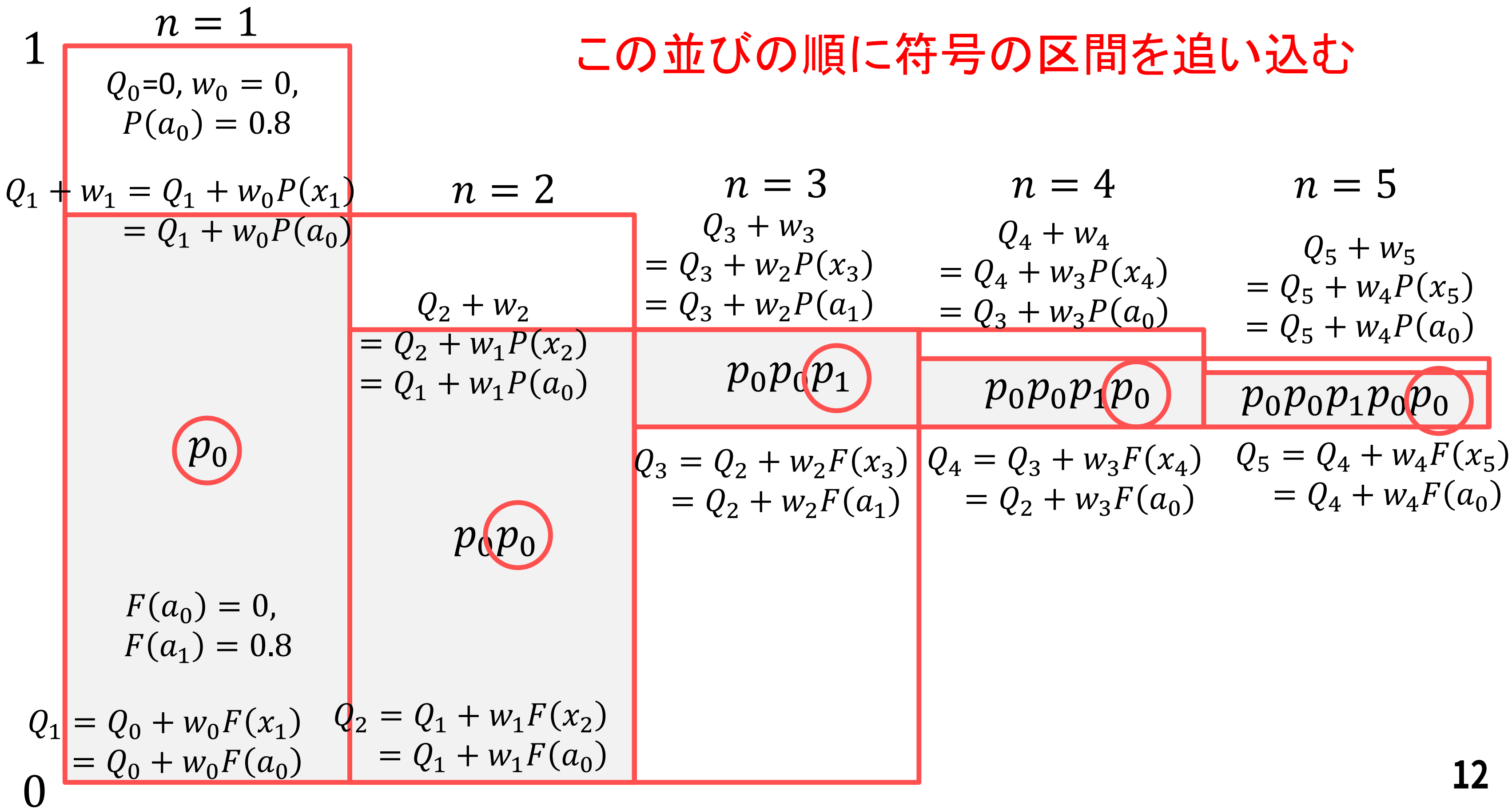
$\tilde{c} = \left[0.c_1 c_2 \cdots c_l \cdots c_{l_0}\right]_r$ に対して

$$\widetilde{Q_i} + \widetilde{w_i} F(a_j) \le \tilde{c} < \widetilde{Q_i} + \widetilde{w_i} F(a_{j+1})$$

を満足する$j$（$0 \le j \le M-1$）を求めて復号系列$i$番目の文字とする。

# イライアス符号の符号化例(1)

- $A = \{a_0, a_1\}, P(a_0) = 0.8, P(a_1) = 0.2$
  で与えられる無記憶情報源を考え、$n = 5$の系列
  $x_1{}^5 = x_1 x_2 x_3 x_4 x_5 = a_0 a_0 a_1 a_0 a_0$ の符号化を考える。




この並びの順に符号の区間を追い込む

# イライアス符号の符号化例（2）

① 初期化

$$F(a_0) = 0, F(a_1) = P(a_1) = 0.8, F(a_2) = 1$$

$$Q_0 = 0, w_0 = 1$$

② 符号化②に基づき$i = 1$とおくと、対応する文字が$a_0$なので、

$Q_1 = Q_0 + w_0F(x_1) = 0 + 1 \times \boxed{F(a_0)} = 0$ <- $Q_0$のまま

$$w_1 = w_0P(x_1) = 1 \times \boxed{P(a_0)} = 0.8$$

$$\rightarrow [Q_1, Q_1 + w_1) = [0, 0.8)$$

# イライアス符号の符号化例(3)

$i = 2$の場合、対応する文字が$a_0$

$$Q_2 = Q_1 + w_1 F(x_2) = 0 + 0.8 \times \boxed{F(a_0)} = 0$$

$$w_2 = w_1 P(x_2) = 0.8 \times \boxed{P(a_0)} = 0.64$$

$$\rightarrow [Q_2, Q_2 + w_2) = [0, 0.64)$$

$i = 3$の場合、対応する文字が$a_1$

$$Q_3 = Q_2 + w_2 F(x_3) = 0 + 0.64 \times \boxed{F(a_1)} = 0.512$$

$$w_3 = w_2 P(x_3) = 0.64 \times \boxed{P(a_1)} = 0.128$$

$$\rightarrow [Q_3, Q_3 + w_3) = [0.512, 0.64)$$

# イライアス符号の符号化例(4)

$i = 4$の場合、対応する文字が$a_0$

$$Q_4 = Q_3 + w_3 F(x_4) = 0.512 + 0.128 \times \boxed{F(a_0)} = 0.512$$

$$w_4 = w_3 P(x_4) = 0.128 \times \boxed{P(a_0)} = 0.1024$$

$$\rightarrow [Q_4, Q_4 + w_4) = [0.512, 0.6144)$$

$i = 5$の場合、対応する文字が$a_0$

$$Q_5 = Q_4 + w_4 F(x_5) = 0.512 + 0.1024 \times \boxed{F(a_0)} = 0.512$$

$$w_5 = w_4 P(x_5) = 0.1024 \times \boxed{P(a_0)} = 0.08192$$

$$\rightarrow [Q_5, Q_5 + w_5) = [0.512, 0.59392)$$

# イライアス符号の符号化例(５)

③ $[Q_5, Q_5 + w_5) = [0.512, 0.59392)$
$= \left[[0.1000001 \cdots]_2, [0.1001100 \cdots]_2\right)$

$l = \lceil -\log_2 w_5 \rceil + 1 = 5$

区間$[Q_5, Q_5 + w_5)$に対して小数点以下5桁の小数$c$として

$c = [0.10001]_2, [0.10010]_2)$の２つが求まる。

$c = [0.10001]_2$を採用し、符号語$c_l = 10001$とする。

# イライアス符号の復号化例（１）

- 符号化例の$c_l = 10001$を復号化する。

$$n = 5, l_0 = \lceil -5\log_2 0.2 \rceil + 1 = 13$$

復号対象となる符号シンボル系列は、

$$\tilde{c} = [0.10001c_6 \cdots c_{13}]_2$$

$$c = [0.10001]_2 = 0.53125 \leq \tilde{c} < [0.10010]_2 = 0.5625$$

① 初期化

$$\widetilde{Q_0} = 0, \widetilde{w_0} = 1, F(a_0) = 0, F(a_1) = 0.8, F(a_2) = 1,$$
$$P(a_0) = 0.8, P(a_1) = 0.2, P(z_0) = 1, F(z_0) = 0$$

# イライアス符号の復号化例(2)

対象とする符号が先頭から順にどの文字の区間に存在するかを求める

② $i$ =1に対して復号化②の手順を行う。($P(a_0)$側か$P(a_1)$側か)

$\widetilde{Q_1} = \widetilde{Q_0} + \widetilde{w_0}F(z_0) = 0 + 1 \times 0 = 0$　対象とする範囲の最小値

$\widetilde{w_1} = \widetilde{w_0}P(z_0) = 1 \times 1 = 1$　対象とする範囲

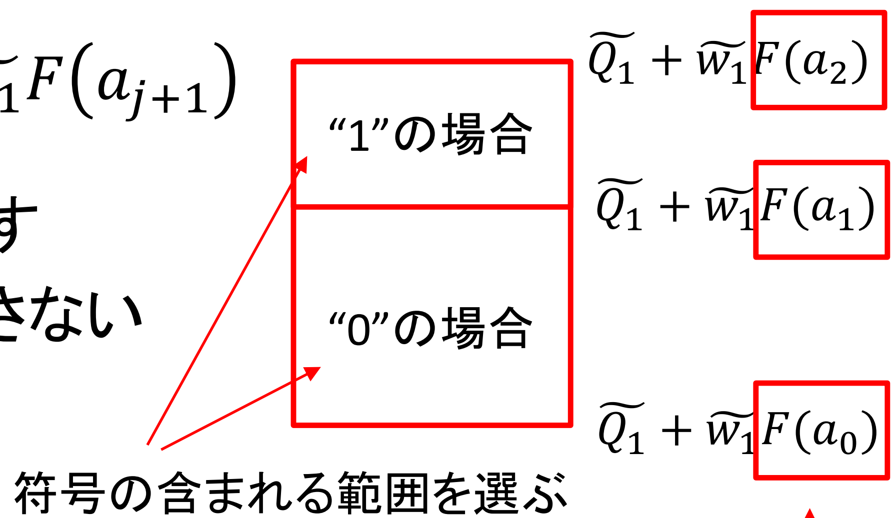
$\widetilde{Q_1} + \widetilde{w_1}F(a_j) \leq \tilde{c} < \widetilde{Q_1} + \widetilde{w_1}F(a_{j+1})$

$\rightarrow F(a_0) \leq \tilde{c} < F(a_1)$を満たす

$F(a_1) \leq \tilde{c} < F(a_2)$を満たさない

よって、$j = 0$

$\therefore z_1 = a_0$




$F(a_0) \leftrightarrow F(a_1)$か
$F(a_1) \leftrightarrow F(a_2)$かを選ぶ

# イライアス符号の復号化例（３）

$i = 2$の場合、

$$\widetilde{Q_2} = \widetilde{Q_1} + \widetilde{w_1}F(z_1) = 0 + 1 \times F(a_0) = 0$$
$$\widetilde{w_2} = \widetilde{w_1}P(z_1) = 1 \times P(a_0) = 0.8$$

$$\widetilde{Q_2} + \widetilde{w_2}F\left(a_j\right) \leq \tilde{c} < \widetilde{Q_2} + \widetilde{w_2}F\left(a_{j+1}\right)$$
$$0.8F\left(a_j\right) \leq \tilde{c} < 0.8F\left(a_{j+1}\right)$$

$\rightarrow j = 0$

$\therefore z_2 = a_0$

# イライアス符号の復号化例（4）

$i = 3$の場合、

$$\widetilde{Q_3} = \widetilde{Q_2} + \widetilde{w_2}F(z_2) = 0 + 0.8 \times F(a_0) = 0$$

$$\widetilde{w_3} = \widetilde{w_2}P(z_2) = 0.8 \times P(a_0) = 0.64$$

$$\widetilde{Q_3} + \widetilde{w_3}F\left(a_j\right) \leq \tilde{c} < \widetilde{Q_3} + \widetilde{w_3}F\left(a_{j+1}\right)$$

$$0.64F\left(a_j\right) \leq \tilde{c} < 0.64F\left(a_{j+1}\right)$$

$\rightarrow j = 1$

$\therefore z_3 = a_1$

# イライアス符号の復号化例（５）

$i = 4$の場合、

$$\widetilde{Q_4} = \widetilde{Q_3} + \widetilde{w_3}F(z_3) = 0 + 0.64 \times F(a_1) = 0.512$$

$$\widetilde{w_4} = \widetilde{w_3}P(z_3) = 0.64 \times P(a_1) = 0.128$$

$$\widetilde{Q_4} + \widetilde{w_4}F(a_j) \leq \tilde{c} < \widetilde{Q_4} + \widetilde{w_4}F(a_{j+1})$$

$$\boxed{0.512} + 0.128F(a_j) \leq \tilde{c} < \boxed{0.512} + 0.128F(a_{j+1})$$

$$\rightarrow j = 0$$

$$\therefore z_4 = a_0$$

$z_3 = a_1$だったので下限が$i = 3$までから変化

# イライアス符号の復号化例(6)

$i = 5$の場合、

$$\widetilde{Q_5} = \widetilde{Q_4} + \widetilde{w_4}F(z_4) = 0.512 + 0.128 \times F(a_0) = 0.512$$

$$\widetilde{w_5} = \widetilde{w_4}P(z_4) = 0.128 \times P(a_0) = 0.1024$$

$$\widetilde{Q_5} + \widetilde{w_5}F\left(a_j\right) \leq \tilde{c} < \widetilde{Q_5} + \widetilde{w_5}F\left(a_{j+1}\right)$$

$$0.512 + 0.1024F\left(a_j\right) \leq \tilde{c} < 0.512 + 0.1024F\left(a_{j+1}\right)$$

$\rightarrow j = 0$

$\therefore z_5 = a_0$

よって、

$$z_1^5 = a_0 a_0 a_1 a_0 a_0$$

# まとめ

## イライアス(Elias)符号

- 情報源系列を、長さの短い順に並べ、その発生確率を高さとする積み木を描く。
- 長さ$n$の系列 $x_1{}^n \in A^n$ を $[0,1)$ を覆う交わりのない小区間に一対一に対応付ける。