

信州大学
大学院総合理工学研究科

修士論文

混合情報源の情報スペクトルの候補情報源による表現

指導教員 西新 幹彦 准教授

工学専攻 電子情報システム工学分野

学籍番号 23W2103K

氏名 山名 岬陽

2025 年 2 月 19 日

目次

1	序論	1
1.1	はじめに	1
2	情報源符号化問題と混合情報源	1
2.1	固定長符号化の符号化レート	2
2.2	混合情報源の固定長符号化レート	3
2.3	強逆情報源の混合とエントロピー・スペクトル	4
3	一般情報源の混合と情報スペクトル	4
3.1	定理の証明 (前半)	7
3.2	定理の証明 (後半)	9
4	まとめ	12
	謝辞	12
	参考文献	12

1 序論

1.1 はじめに

電話やメールなど世の中の通信システムでは、伝えたい情報に対し符号化・復号化という処理を行う。符号化では、送信したい情報を符号化し、符号語を得る。その際、符号化に使用された符号語長を一文字あたりに換算したものを符号化レートと呼ぶ。通信の効率性のためには符号化レートは小さい方がよいが、小さくしすぎると誤り確率が大きくなってしまう。そこで、誤りなく情報を符号化するという条件の下、この符号化レートをどこまで小さくできるかを考える問題を情報源符号化問題という。

情報源符号化問題において、符号化は情報源の特性を基に行われるため、情報源の特性を正確に把握することが重要となる。本研究では情報源の特性を表す指標として情報スペクトルを扱う。情報スペクトルとは情報源の特性を表す確率分布であり、符号化レートの導出に重要な概念である。例えば、一般情報源における固定長符号化では情報スペクトルの上端、可変長符号化では情報スペクトルの期待値が最適な符号化レートとなることが知られている。本研究では、情報スペクトルの上端、下端、期待値ではなく情報スペクトルそのものに着目した。

情報源符号化問題の基本的な問題設定では、情報源の特性は既知とされる。しかし、実際の通信システムでは情報源の特性は不明であることが多い。特性が不明な通信システムの例として、動画ストリーミングシステムが挙げられ、映像データ、音声データ、字幕データなど特性が異なる複数の情報源から符号器への出力が行われる。本論文では特性が不明な情報源を扱う。具体的には複数の候補情報源から確率的に1つが選ばれる情報源を考える、このような情報源を混合情報源と呼ぶ。

混合情報源の情報スペクトルに関する従来研究 [1] では、離散的な混合情報源に対して候補情報源が強逆情報源という条件下で、混合情報源の情報スペクトルの形が候補情報源を使って間接的に表現されている。本論文では、従来研究 [1] に対して情報スペクトルをエントロピー・スペクトルから相互情報量スペクトルへと拡張し、候補情報源を強逆情報源から一般情報源へと拡張した条件下で、混合情報源の情報スペクトルの形を候補情報源を使って直接表現する。

2 情報源符号化問題と混合情報源

情報源アルファベットを $\mathcal{X}$ とする。一般情報源 $\mathbf{X}$ とは $\mathcal{X}^n$ 上の任意の確率分布 P_{X^n} をもつ確率変数 X^n の列である。 $\mathbf{X} = \{X^n\}_{n=1}^{\infty}$ と表す。

本章では、特性が明らかな情報源に対する符号化手法について知られている結果を紹介し、情報源符号化問題における情報スペクトルの重要性を確認する。情報スペクトルの1つにエン

トロピー・スペクトルがあり、エントロピー・スペクトルは $\frac{1}{n} \log \frac{1}{P_{X^n}(X^n)}$ の確率分布のことである。なお、本論文を通じて $\log$ は自然対数を表す。初めに、数学的準備として以下を定義する。

定義 1 実数値確率変数列 $\{Z_n\}_{n=1}^{\infty}$ に対して

$$\text{p-lim sup}_{n \rightarrow \infty} Z_n \triangleq \inf \left\{ \alpha \mid \lim_{n \rightarrow \infty} \Pr\{Z_n > \alpha\} = 0 \right\} \quad (1)$$

とし、式 (1) を確率的上極限と呼ぶ。

2.1 固定長符号化の符号化レート

エントロピー・スペクトルは固定長符号化定理を述べる上で重要な役割を担っている。固定長符号化問題は次のように定式化される。 M_n 個の符号語の集合を $\mathcal{M}_n \triangleq \{1, 2, \dots, M_n\}$ とする。そして、符号化を符号器 $\varphi_n: \mathcal{X}^n \rightarrow \mathcal{M}_n$ によって定め、復号化を復号器 $\psi_n: \mathcal{M}_n \rightarrow \mathcal{Y}^n$ によって定める。符号器と復号器の組 (φ_n, ψ_n) を符号と呼ぶ。ここで、 $\frac{1}{n} \log M_n$ は符号化レートと呼ばれる。情報源の出力 $\mathbf{x} \in \mathcal{X}^n$ は符号器 φ_n によって符号化されたあと復号器 ψ_n によって復元される。このような符号化を固定長符号化と呼ぶ。

定義 2 情報源からの長さ n の系列を表す確率変数を $X^n = (X_1, \dots, X_n)$ とおき

$$\varepsilon_n \triangleq \Pr\{X^n \neq \psi_n(\varphi_n(X^n))\} \quad (2)$$

を誤り確率と呼ぶ。

定義 3 固定長符号化においてレート R が達成可能とは、

$$\lim_{n \rightarrow \infty} \varepsilon_n = 0 \quad (3)$$

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log M_n \leq R \quad (4)$$

を満たす符号の列 $\{(\varphi_n, \psi_n)\}_{n=1}^{\infty}$ が存在することである。

定義 4 最小許容固定長符号化レートを

$$R_f(\mathbf{X}) \triangleq \inf\{R \mid R \text{ が達成可能}\} \quad (5)$$

と定義する。

以上のもとで、次の定理が成り立つことが知られている。

定理 1[1] 一般情報源における固定長符号化では、最小許容固定長符号化レート $R_f(\mathbf{X})$ は

$$R_f(\mathbf{X}) = \overline{H}(\mathbf{X}) \quad (6)$$

で与えられる。

ただし, $\overline{H}(\mathbf{X})$ は,

$$\overline{H}(\mathbf{X}) \triangleq \text{p-lim sup}_{n \rightarrow \infty} \frac{1}{n} \log \frac{1}{P_{X^n}(X^n)} \quad (7)$$

と定義する.

式 (7) の右辺より, エントロピー・スペクトルが最小許容固定長符号化レートを表す上で極めて重要であることがわかる.

2.2 混合情報源の固定長符号化レート

定理 1 において, 一般情報源が混合情報源である場合には最小許容固定長符号化レート $R_f(\mathbf{X})$ は別の表現をもつ. まず混合情報源を定義する.

定義 7 可算無限個の任意の候補情報源 $\mathbf{X}_i = \{X_i^n\}_{n=1}^\infty (i = 1, 2, \dots)$ が与えられたとき,

$$P_{X^n}(\mathbf{x}) \triangleq \sum_{i=1}^{\infty} \alpha_i P_{X_i^n}(\mathbf{x}), \mathbf{x} \in \mathcal{X}^n \quad (8)$$

によって定義される情報源 $\mathbf{X} = \{X^n\}_{n=1}^\infty$ を, $\mathbf{X}_i (i = 1, 2, \dots)$ の混合情報源という. ただし, α_i は

$$\sum_{i=1}^{\infty} \alpha_i = 1, (\alpha_i > 0, \forall i = 1, 2, \dots) \quad (9)$$

なる定数である.

混合情報源に対して, 次の定理が成り立つことが知られている.

定理 2[1] 候補情報源 $\mathbf{X}_i (i = 1, 2, \dots)$ で定められる混合情報源 $\mathbf{X}$ に対して

$$R_f(\mathbf{X}) = \sup_i \overline{H}(\mathbf{X}_i) \quad (10)$$

が成立する.

ただし, $\overline{H}(\mathbf{X}_i)$ は,

$$\overline{H}(\mathbf{X}_i) \triangleq \text{p-lim sup}_{n \rightarrow \infty} \frac{1}{n} \log \frac{1}{P_{X_i^n}(X_i^n)} \quad (11)$$

と定義する.

定理 2 の本質は

$$\overline{H}(\mathbf{X}) = \sup_i \overline{H}(\mathbf{X}_i) \quad (12)$$

であり, 候補情報源と混合情報源のエントロピー・スペクトルの関係について 1 つの側面を表している.

2.3 強逆情報源の混合とエントロピー・スペクトル

式 (12) は情報スペクトルの上端に関する 1 つの性質であるが、候補情報源が強逆性を持つ場合は、情報スペクトルの形についてより詳しい関係が明らかにされている。まず、情報源の強逆性を定義する。

定義 8 固定長情報源符号化において、

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log M_n < R_f(\mathbf{X}) \quad (13)$$

なるどのような (n, M_n, ε_n) 符号に対しても

$$\lim_{n \rightarrow \infty} \varepsilon_n = 1 \quad (14)$$

となるとき、情報源 $\mathbf{X} = \{X^n\}_{n=1}^\infty$ は強逆性を持つという。

定義 7 の式 (8) と定義 8 の式 (14) を用いて、次の定理が成り立つことが知られている。

定理 3[2] 混合情報源に対して、候補情報源が強逆性を持つとき、

$$\sum_{i: \bar{H}(\mathbf{X}_i) > R} \alpha_i \leq \underline{F}(R|\mathbf{X}) \leq \bar{F}(R|\mathbf{X}) \leq \sum_{i: \bar{H}(\mathbf{X}_i) \geq R} \alpha_i \quad (15)$$

となることが知られている。ただし、 $\underline{F}(R|\mathbf{X})$, $\bar{F}(R|\mathbf{X})$ はそれぞれ、

$$\underline{F}(R|\mathbf{X}) \triangleq \liminf_{n \rightarrow \infty} \Pr \left\{ \frac{1}{n} \log \frac{1}{P_{X^n}(X^n)} \geq R \right\} \quad (16)$$

$$\bar{F}(R|\mathbf{X}) \triangleq \limsup_{n \rightarrow \infty} \Pr \left\{ \frac{1}{n} \log \frac{1}{P_{X^n}(X^n)} \geq R \right\} \quad (17)$$

と定義する。

定理 3 は、式 (12) と比べて情報スペクトルの形をより具体的に示しているが、情報スペクトルの形を表す定理としては依然として明示性が不十分である。

3 一般情報源の混合と情報スペクトル

混合情報源と候補情報源の情報スペクトルの関係について式 (15) よりも直接的な表現を導く。これが本論文の主たる成果である。

候補情報源は $\mathbf{X}_i$ でなく $\mathbf{X}_i$ と $\mathbf{Y}_i$ の同時分布で表す一般情報源とし次のように定義する。ここで一般情報源は強逆性を持つ情報源の一般化になっており、強逆性を持たない情報源を含んでいる。

定義 9 可算無限個の候補情報源 $\mathbf{X}_i \mathbf{Y}_i = \{X_i^n Y_i^n\}_{n=1}^\infty (i = 1, 2, \dots)$ が与えられたとき,

$$P_{X^n Y^n}(\mathbf{x}, \mathbf{y}) \triangleq \sum_{i=1}^{\infty} \alpha_i P_{X_i^n Y_i^n}(\mathbf{x}, \mathbf{y}), \mathbf{x} \in \mathcal{X}^n, \mathbf{y} \in \mathcal{Y}^n \quad (18)$$

によって定義される情報源 $\mathbf{X} \mathbf{Y} = \{X^n Y^n\}_{n=1}^\infty$ を, $\mathbf{X}_i \mathbf{Y}_i (i = 1, 2, \dots)$ の混合情報源という. ただし, α_i は

$$\sum_{i=1}^{\infty} \alpha_i = 1, (\alpha_i \geq 0, \forall i = 1, 2, \dots) \quad (19)$$

なる定数である.

この混合情報源に対し,

$$\frac{1}{n} \log \frac{P_{Y^n | X^n}(Y^n | X^n)}{P_{Y^n}(Y^n)} \quad (20)$$

の確率分布を (X^n, Y^n) の相互情報量スペクトルと呼ぶ. ここで, $P_{Y^n}(Y^n) = P_{X^n}(X^n)$ とすれば式 (20) は直ちに

$$\frac{1}{n} \log \frac{1}{P_{X^n}(X^n)} \quad (21)$$

となるため, 相互情報量スペクトルはエントロピー・スペクトルの一般化である. 相互情報量スペクトルは主にレート歪み問題を考える場合に用いられる. 本論文では, エントロピー・スペクトルと相互情報量スペクトルを総称して情報スペクトルと呼ぶ.

次に, 2つの確率分布を比較する手法として以下を定義する.

定義 10[2] 2つの確率変数列 $\{Z_n\}_{n=1}^\infty, \{V_n\}_{n=1}^\infty$ を考える. 任意の $\gamma > 0$, に対して n が十分大きければすべての実数 θ で

$$\Pr\{Z_n \leq \theta\} \geq \Pr\{V_n \leq \theta - \gamma\} - \gamma \quad (22)$$

が成り立つことを本稿では,

$$Z_n \stackrel{\text{d}}{\lesssim} V_n \quad \text{as } n \rightarrow \infty \quad (23)$$

と表す. また,

$$V_n \stackrel{\text{d}}{\lesssim} Z_n \stackrel{\text{d}}{\lesssim} V_n \quad \text{as } n \rightarrow \infty$$

のとき,

$$Z_n \stackrel{\text{d}}{\simeq} V_n \quad \text{as } n \rightarrow \infty \quad (24)$$

と表す。これは、 Z_n と V_n の分布が極限で一致していることを意味する。また、「as $n \rightarrow \infty$ 」は文脈から明らかなときは省略する。

混合情報源の相互情報量スペクトルは、候補情報源を用いて次のように表わされる。

定理 4 混合情報源 $\mathbf{X}$ に対して、その情報スペクトルは候補情報源の情報スペクトルの混合と等しくなる。具体的には、

$$\frac{1}{n} \log \frac{P_{Y^n|X^n}(Y^n|X^n)}{P_{Y^n}(Y^n)} \stackrel{d}{\simeq} Z_n \quad (25)$$

が成り立つ。ただし、 $Z_n, Z_{n,i}$ をそれぞれ、

$$Z_{n,i} \triangleq \frac{1}{n} \log \frac{P_{Y_i^n|X_i^n}(Y_i^n|X_i^n)}{P_{Y_i^n}(Y_i^n)} \quad (26)$$

$$P_{Z_n}(z) \triangleq \sum_{i=1}^{\infty} \alpha_i P_{Z_{n,i}}(z) \quad (27)$$

と定義する。

注意 上記の定理で、 $P_{Y^n}(Y^n) = P_{X^n}(X^n)$ とすれば直ちに、

$$\frac{1}{n} \log \frac{1}{P_{X^n}(X^n)} \stackrel{d}{\simeq} Z'_n \quad (28)$$

が成り立つ。ただし、 $Z'_n, Z'_{n,i}$ をそれぞれ、

$$Z'_{n,i} \triangleq \frac{1}{n} \log \frac{1}{P_{X_i^n}(X_i^n)} \quad (29)$$

$$P_{Z'_n}(z') \triangleq \sum_{i=1}^{\infty} \alpha_i P_{Z'_{n,i}}(z') \quad (30)$$

と定義する。

この定理を証明するために次の補題を用いる。

補題 1[1] 一般に、2つの確率変数 X_1^n, X_2^n と任意の $\gamma > 0$ に対して、

$$\Pr \left\{ \frac{1}{n} \log P_{X_1^n}(X_1^n) - \frac{1}{n} \log P_{X_2^n}(X_1^n) \leq -\gamma \right\} \leq e^{-n\gamma} \quad (31)$$

が成り立つ。

式 (25) の $\stackrel{d}{\simeq}$ を示すには、 $\stackrel{d}{\lesssim}$ と $\stackrel{d}{\gtrsim}$ を示せばよい。以下の節ではこれらをそれぞれ示す。

3.1 定理の証明 (前半)

最初に,

$$Z_n \stackrel{d}{\lesssim} \frac{1}{n} \log \frac{P_{Y^n|X^n}(Y^n|X^n)}{P_{Y^n}(Y^n)} \quad (32)$$

を示す.

〔証明〕 まず混合情報源の定義より,

$$\Pr \left\{ \frac{1}{n} \log \frac{P_{Y^n|X^n}(Y^n|X^n)}{P_{Y^n}(Y^n)} \leq R \right\} \quad (33)$$

$$= \sum_{i=1}^{\infty} \alpha_i \Pr \left\{ \frac{1}{n} \log \frac{P_{Y^n|X^n}(Y_i^n|X_i^n)}{P_{Y^n}(Y_i^n)} \leq R \right\} \quad (34)$$

が得られる. ここで, 式 (34) の $\Pr\{\dots\}$ を $P(A)$ と表せば,

$$P(A) = \Pr \left\{ \frac{1}{n} \log \frac{P_{Y^n|X^n}(Y_i^n|X_i^n)}{P_{Y^n}(Y_i^n)} \leq R \right\} \quad (35)$$

$$= \Pr \left\{ \frac{1}{n} \log \frac{P_{X^n Y^n}(X_i^n, Y_i^n)}{P_{X^n}(X_i^n) P_{Y^n}(Y_i^n)} \leq R \right\} \quad (36)$$

$$= \Pr \left\{ \frac{1}{n} \log P_{X^n Y^n}(X_i^n, Y_i^n) + \frac{1}{n} \log \frac{1}{P_{X^n}(X_i^n)} + \frac{1}{n} \log \frac{1}{P_{Y^n}(Y_i^n)} \leq R \right\} \quad (37)$$

となる. この式の $\frac{1}{n} \log \frac{1}{P_{X^n}(X_i^n)}$, $\frac{1}{n} \log \frac{1}{P_{Y^n}(Y_i^n)}$ について, 補題 1 から,

$$\Pr \left\{ \frac{1}{n} \log \frac{1}{P_{X^n}(X_i^n)} \leq \frac{1}{n} \log \frac{1}{P_{X_i^n}(X_i^n)} - \gamma \right\} \leq e^{-n\gamma}, \quad (38)$$

$$\Pr \left\{ \frac{1}{n} \log \frac{1}{P_{Y^n}(Y_i^n)} \leq \frac{1}{n} \log \frac{1}{P_{Y_i^n}(Y_i^n)} - \gamma \right\} \leq e^{-n\gamma} \quad (39)$$

が得られる. ここで,

$$P(B) = \Pr \left\{ \frac{1}{n} \log \frac{1}{P_{X^n}(X_i^n)} > \frac{1}{n} \log \frac{1}{P_{X_i^n}(X_i^n)} - \gamma \right\} \quad (40)$$

$$P(C) = \Pr \left\{ \frac{1}{n} \log \frac{1}{P_{Y^n}(Y_i^n)} > \frac{1}{n} \log \frac{1}{P_{Y_i^n}(Y_i^n)} - \gamma \right\} \quad (41)$$

とおくと、式 (38),(39) より、

$$P(A) = P(A \cap B \cap C) + P(A \cap (B \cap C)^c) \quad (42)$$

$$\leq P(A \cap B \cap C) + P((B \cap C)^c) \quad (43)$$

$$= P(A \cap B \cap C) + P((B^c \cup C^c)) \quad (44)$$

$$\leq P(A \cap B \cap C) + P(B^c) + P(C^c) \quad (45)$$

$$\leq P(A \cap B \cap C) + 2e^{-n\gamma} \quad (46)$$

$$= \Pr \left\{ \frac{1}{n} \log P_{X^n Y^n}(X_i^n, Y_i^n) + \frac{1}{n} \log \frac{1}{P_{X_i^n}(X_i^n)} \right. \\ \left. + \frac{1}{n} \log \frac{1}{P_{Y_i^n}(Y_i^n)} \leq R + 2\gamma \right\} + 2e^{-n\gamma} \quad (47)$$

が成り立つ．さらにこの式の $\frac{1}{n} \log P_{X^n Y^n}(X_i^n, Y_i^n)$ に着目すると、定義 9 より、

$$P_{X^n Y^n}(\mathbf{x}, \mathbf{y}) \geq \alpha_i P_{X_i^n Y_i^n}(\mathbf{x}, \mathbf{y}) \quad (48)$$

であるため、

$$P(A) \leq \Pr \left\{ \frac{1}{n} \log P_{X_i^n Y_i^n}(X_i^n, Y_i^n) + \frac{1}{n} \log \frac{1}{P_{X_i^n}(X_i^n)} \right. \\ \left. + \frac{1}{n} \log \frac{1}{P_{Y_i^n}(Y_i^n)} \leq R + \frac{1}{n} \log \frac{1}{\alpha_i} + 2\gamma \right\} + 2e^{-n\gamma} \quad (49)$$

$$= \Pr \left\{ \frac{1}{n} \log \frac{P_{Y_i^n | X_i^n}(Y_i^n | X_i^n)}{P_{Y_i^n}(Y_i^n)} \leq R + \frac{1}{n} \log \frac{1}{\alpha_i} + 2\gamma \right\} + 2e^{-n\gamma} \quad (50)$$

が得られる．

式 (50) を式 (34) に代入したもとの、式 (34) 全体を評価する．そのために、

$$\sum_{i=k+1}^{\infty} \alpha_i \leq \gamma \quad (51)$$

となるように k を取り、式 (34) の和を $\sum_{i=1}^k + \sum_{i=k+1}^{\infty}$ のように分ける．前半部分 $\sum_{i=1}^k$ は

i が有限個であるため、十分大きなすべての n で $\frac{1}{n} \log \frac{1}{\alpha_i} < \gamma$ となる。したがって、

$$\sum_{i=1}^k \alpha_i \Pr \left\{ \frac{1}{n} \log \frac{P_{Y_i^n|X_i^n}(Y_i^n|X_i^n)}{P_{Y_i^n}(Y_i^n)} \leq R + \frac{1}{n} \log \frac{1}{\alpha_i} + 2\gamma \right\} \quad (52)$$

$$\leq \sum_{i=1}^k \alpha_i \Pr \left\{ \frac{1}{n} \log \frac{P_{Y_i^n|X_i^n}(Y_i^n|X_i^n)}{P_{Y_i^n}(Y_i^n)} \leq R + 3\gamma \right\} \quad (53)$$

$$\leq \sum_{i=1}^{\infty} \alpha_i \Pr \left\{ \frac{1}{n} \log \frac{P_{Y_i^n|X_i^n}(Y_i^n|X_i^n)}{P_{Y_i^n}(Y_i^n)} \leq R + 3\gamma \right\} \quad (54)$$

$$= \sum_{i=1}^{\infty} \alpha_i \Pr \{Z_{n,i} \leq R + 3\gamma\} \quad (55)$$

$$= \Pr \{Z_n \leq R + 3\gamma\} \quad (56)$$

と書ける。ただし式 (54)~(56) の式変形では式 (26),(27) を用いた。後半部分 $\sum_{i=k+1}^{\infty}$ は、式 (51) より、

$$\sum_{i=k+1}^{\infty} \alpha_i \Pr \left\{ \frac{1}{n} \log \frac{P_{Y_i^n|X_i^n}(Y_i^n|X_i^n)}{P_{Y_i^n}(Y_i^n)} \leq R + \frac{1}{n} \log \frac{1}{\alpha_i} + 2\gamma \right\} \leq \gamma \quad (57)$$

となる。以上を総合すると、十分大きなすべての n で、

$$\Pr \left\{ \frac{1}{n} \log \frac{P_{Y^n|X^n}(Y^n|X^n)}{P_{Y^n}(Y^n)} \leq R \right\} \leq \Pr \{Z_n \leq R + 3\gamma\} + 2e^{-n\gamma} + \gamma \quad (58)$$

が成り立ち、定義 10 から、

$$Z_n \stackrel{d}{\lesssim} \frac{1}{n} \log \frac{P_{Y^n|X^n}(Y^n|X^n)}{P_{Y^n}(Y^n)} \quad (59)$$

が得られる。

3.2 定理の証明 (後半)

次に、

$$\frac{1}{n} \log \frac{P_{Y^n|X^n}(Y^n|X^n)}{P_{Y^n}(Y^n)} \stackrel{d}{\lesssim} Z_n \quad (60)$$

を示す。

〔証明〕 Z_n の分布は式 (26),(27) より,

$$\Pr \{Z_n \leq R\} = \sum_{i=1}^{\infty} \alpha_i \Pr \{Z_{n,i} \leq R\} \quad (61)$$

$$= \sum_{i=1}^{\infty} \alpha_i \Pr \left\{ \frac{1}{n} \log \frac{P_{Y_i^n | X_i^n}(Y_i^n | X_i^n)}{P_{Y_i^n}(Y_i^n)} \leq R \right\} \quad (62)$$

$$= \sum_{i=1}^{\infty} \alpha_i \Pr \left\{ \frac{1}{n} \log P_{X_i^n Y_i^n}(X_i^n, Y_i^n) + \frac{1}{n} \log \frac{1}{P_{X_i^n}(X_i^n)} + \frac{1}{n} \log \frac{1}{P_{Y_i^n}(Y_i^n)} \leq R \right\} \quad (63)$$

と書ける. ここで式 (63) の $\Pr\{\dots\}$ を,

$$P(C) = \Pr \left\{ \frac{1}{n} \log P_{X_i^n Y_i^n}(X_i^n, Y_i^n) + \frac{1}{n} \log \frac{1}{P_{X_i^n}(X_i^n)} + \frac{1}{n} \log \frac{1}{P_{Y_i^n}(Y_i^n)} \leq R \right\} \quad (64)$$

とおく. この式の $\frac{1}{n} \log P_{X_i^n Y_i^n}(X_i^n, Y_i^n)$ について, 補題 1 から,

$$\Pr \left\{ \frac{1}{n} \log P_{X_i^n Y_i^n}(X_i^n, Y_i^n) \leq \frac{1}{n} \log P_{X^n Y^n}(X_i^n, Y_i^n) - \gamma \right\} \leq e^{-n\gamma} \quad (65)$$

が得られる. ここで,

$$P(D) = \Pr \left\{ \frac{1}{n} \log P_{X_i^n Y_i^n}(X_i^n, Y_i^n) > \frac{1}{n} \log P_{X^n Y^n}(X_i^n, Y_i^n) - \gamma \right\} \quad (66)$$

とおくと, 式 (65) より,

$$P(C) = P(C \cap D) + P(C \cap D^c) \quad (67)$$

$$\leq P(C \cap D) + P(D^c) \quad (68)$$

$$\leq P(C \cap D) + e^{-n\gamma} \quad (69)$$

$$= \Pr \left\{ \frac{1}{n} \log P_{X^n Y^n}(X_i^n, Y_i^n) + \frac{1}{n} \log \frac{1}{P_{X_i^n}(X_i^n)} + \frac{1}{n} \log \frac{1}{P_{Y_i^n}(Y_i^n)} \leq R + \gamma \right\} + e^{-n\gamma} \quad (70)$$

が成り立つ. 次に式 (70) 内の $\frac{1}{n} \log \frac{1}{P_{X_i^n}(X_i^n)}, \frac{1}{n} \log \frac{1}{P_{Y_i^n}(Y_i^n)}$ に着目する. 定義 9 から,

$$P_{X^n}(\mathbf{x}) = \sum_{i=1}^{\infty} \alpha_i P_{X_i^n}(\mathbf{x}), P_{Y^n}(\mathbf{y}) = \sum_{i=1}^{\infty} \alpha_i P_{Y_i^n}(\mathbf{y}) \quad (71)$$

が導き出せるので,

$$\frac{\alpha_i}{P_{X^n}(\mathbf{x})} < \frac{1}{P_{X_i^n}(\mathbf{x})}, \frac{\alpha_i}{P_{Y^n}(\mathbf{y})} < \frac{1}{P_{Y_i^n}(\mathbf{y})} \quad (72)$$

を用いることで,

$$P(C) \leq \Pr \left\{ \frac{1}{n} \log P_{X^n Y^n}(X_i^n, Y_i^n) + \frac{1}{n} \log \frac{1}{P_{X^n}(X_i^n)} + \frac{1}{n} \log \frac{1}{P_{Y^n}(Y_i^n)} \right. \\ \left. \leq R + \frac{2}{n} \log \frac{1}{\alpha_i} + \gamma \right\} + e^{-n\gamma} \quad (73)$$

$$= \Pr \left\{ \frac{1}{n} \log \frac{P_{Y^n|X^n}(Y_i^n|X_i^n)}{P_{Y^n}(Y_i^n)} \leq R + \frac{2}{n} \log \frac{1}{\alpha_i} + \gamma \right\} + e^{-n\gamma} \quad (74)$$

が得られる.

式 (74) を式 (63) に代入したもとので, 式 (63) 全体を評価する. そのために先ほどと同様に,

$$\sum_{i=k+1}^{\infty} \alpha_i \leq \gamma \quad (75)$$

となるように k を取り, 式 (63) の和を $\sum_{i=1}^k + \sum_{i=k+1}^{\infty}$ のように分ける. 前半部分 $\sum_{i=1}^k$ は i が有限個であるため, 十分大きなすべての n で $\frac{2}{n} \log \frac{1}{\alpha_i} < \gamma$ となる. したがって,

$$\sum_{i=1}^k \alpha_i \Pr \left\{ \frac{1}{n} \log \frac{P_{Y^n|X^n}(Y_i^n|X_i^n)}{P_{Y^n}(Y_i^n)} \leq R + \frac{2}{n} \log \frac{1}{\alpha_i} + \gamma \right\} \quad (76)$$

$$\leq \sum_{i=1}^k \alpha_i \Pr \left\{ \frac{1}{n} \log \frac{P_{Y^n|X^n}(Y_i^n|X_i^n)}{P_{Y^n}(Y_i^n)} \leq R + 2\gamma \right\} \quad (77)$$

$$\leq \sum_{i=1}^{\infty} \alpha_i \Pr \left\{ \frac{1}{n} \log \frac{P_{Y^n|X^n}(Y_i^n|X_i^n)}{P_{Y^n}(Y_i^n)} \leq R + 2\gamma \right\} \quad (78)$$

$$= \Pr \left\{ \frac{1}{n} \log \frac{P_{Y^n|X^n}(Y^n|X^n)}{P_{Y^n}(Y^n)} \leq R + 2\gamma \right\} \quad (79)$$

と示せる. 後半部分 $\sum_{i=k+1}^{\infty}$ は式 (75) より,

$$\sum_{i=k+1}^{\infty} \alpha_i \Pr \left\{ \frac{1}{n} \log \frac{P_{Y^n|X^n}(Y_i^n|X_i^n)}{P_{Y^n}(Y_i^n)} \leq R + \frac{2}{n} \log \frac{1}{\alpha_i} + \gamma \right\} \leq \gamma \quad (80)$$

が成り立つ. 以上を総合すると, 十分大きなすべての n で,

$$\Pr \{Z_n \leq R\} \leq \Pr \left\{ \frac{1}{n} \log \frac{P_{Y^n|X^n}(Y^n|X^n)}{P_{Y^n}(Y^n)} \leq R + 2\gamma \right\} + e^{-n\gamma} + \gamma \quad (81)$$

となり, 定義 10 から,

$$\frac{1}{n} \log \frac{P_{Y^n|X^n}(Y^n|X^n)}{P_{Y^n}(Y^n)} \stackrel{d}{\lesssim} Z_n \quad (82)$$

が得られた. $\square$

4 まとめ

本研究では混合情報源の情報スペクトルを候補情報源を使って直接的に表現することができた．具体的には $\stackrel{d}{\simeq}$ を用いることで，混合情報源の情報スペクトルと候補情報源の情報スペクトルの混合が一致することが示せた．従来結果では候補情報源には強逆性が課されていたが，本研究では一般情報源まで拡張させることができた．さらに表現する情報スペクトルはエントロピー・スペクトル，相互情報量スペクトルどちらでもよい．この結果より，特性が不明な情報源は候補情報源を用いることで情報源の特性を正確に把握することができる．一方で，本論文では離散的な混合について議論したが，混合の仕方として離散的な混合よりも一般的である連続的な混合があるため，連続的な混合情報源でも同様のことが成り立つかどうかは今後の問題となる．しかし，連続的な混合情報源における定理の証明で3章の証明手法を用いる際には，式(48)，式(72)を使用できなくなるため，代用できる補題を導入することが必要となる．

謝辞

本研究を行うにあたり，丁寧なご指導を賜りました指導教員の西新幹彦准教授に感謝の意を表する．

参考文献

- [1] 韓 太舜，情報理論における情報スペクトル的方法，培風館，1998.
- [2] 西新幹彦，「情報スペクトル指向の情報源可変長符号化定理について」，第45回情報理論とその応用シンポジウム (SITA2022) 予稿集，pp.442–447，Dec. 2022.