

離散確率分布における Kullback 情報量の直和分解

関東学院大学 経済学部 布能 英一郎

Eiichiro Funo

School of Economics, Kanto Gakuin University

1. Introduction

本稿は、離散確率分布のみを扱う。確率分布 $P_\theta(x)$ と2つの仮説 $H_1: \theta_i = p_i$, $H_2: \theta_i = q_i$ に対して、Kullback 情報量 $I(1, 2)$ は

$$I(1, 2) = E_{H_1} \log \frac{P(X | H_1)}{P(X | H_2)} = \sum_i p_i \log \frac{p_i}{q_i}$$

で定義される。

2 標本問題における Total Information, Between Information, Within Information

2 標本問題にて、仮説 H_1 を「2つの母集団は異なる」、仮説 H_2 : 「2つの母集団は同じ」に選ぶ。Total Information とは、「仮説 H_1 の下での推定量と、仮説 H_2 の下でのパラメータ $\mathbf{p} = (p_1, \dots, p_k)$ 」間の Kullback 情報量、Between Information とは、「仮説 H_2 の下での推定量と、仮説 H_2 の下でのパラメータ $\mathbf{p} = (p_1, \dots, p_k)$ 」間の Kullback 情報量、Within Information とは、「仮説 H_1 の下での推定量と、仮説 H_2 の下での推定量」間の Kullback 情報量、で定義されるものである。

例：多項分布 2つの独立な多項分布 $\mathbf{X}^{[1]}$, $\mathbf{X}^{[2]}$, すなわち $\mathbf{X}^{[i]} = (X_1^{[i]}, \dots, X_k^{[i]}) \sim M(N^{[i]}, p_1^{[i]}, \dots, p_k^{[i]})$ ($i = 1, 2.$) に対して、 H_1 : 2つの母集団は異なる、 H_2 : 2つの母集団は同じ、は

$$H_1: (p_1^{[1]}, \dots, p_k^{[1]}) = \mathbf{p}^{[1]} \neq \mathbf{p}^{[2]} = (p_1^{[2]}, \dots, p_k^{[2]}),$$

$$H_2: p_i^{[1]} = p_i^{[2]} = p_i, \quad i = 1, \dots, k,$$

と記述できる。よって

$$\begin{aligned} L = \text{Likelihood} &= \frac{P(\mathbf{x}^{[1]}, \mathbf{x}^{[2]} | H_1)}{P(\mathbf{x}^{[1]}, \mathbf{x}^{[2]} | H_2)} = \prod_{i=1}^2 \frac{\prod_{j=1}^k (p_j^{[i]})^{x_j^{[i]}}}{\prod_{j=1}^k (p_j)^{x_j^{[i]}}} = \prod_{i=1}^2 \prod_{j=1}^k \left(\frac{p_j^{[i]}}{p_j} \right)^{x_j^{[i]}} \\ E_{H_1} \log L &= \sum_{i=1}^2 \sum_{j=1}^k E_{H_1}(X_j^{[i]}) \log \frac{p_j^{[i]}}{p_j} = \sum_{i=1}^2 N^{[i]} \sum_{j=1}^k p_j^{[i]} \log \frac{p_j^{[i]}}{p_j} \end{aligned} \quad (1)$$

である。Total Information は、(1) 式の $p_j^{[i]}$ に H_2 の下での (best な) 推定量 $\hat{p}_j^{[i]} = x_j^{[i]} / N^{[i]}$ を代入して

$$\text{Total Information} = \sum_{i=1}^2 \sum_{j=1}^k x_j^{[i]} \log \frac{x_j^{[i]} / N^{[i]}}{p_j} \quad (2)$$

Within Information は、(2) 式の p_j に H_1 の下での (best な) 推定量 $\hat{p}_j = (x_j^{[1]} + x_j^{[2]})/(N^{[1]} + N^{[2]})$ を代入して

$$\text{Within Information} = \sum_{i=1}^2 \sum_{j=1}^k x_j^{[i]} \log \frac{x_j^{[i]}/N^{[i]}}{(x_j^{[1]} + x_j^{[2]})/(N^{[1]} + N^{[2]})} \quad (3)$$

全く同様な考え方で

$$\text{Between Information} = \sum_{j=1}^k (x_j^{[1]} + x_j^{[2]}) \log \frac{(x_j^{[1]} + x_j^{[2]})/(N^{[1]} + N^{[2]})}{p_j} \quad (4)$$

が得られる。

Proposition 1. 多項分布の2標本問題にて、Total Information は Within Information と Between Information の和に等しい。

これは、(2), (3), (4) 式より、容易な計算で得られる。以降、2標本問題にて、Total information = Within information + Between information が成り立つとき、Kullback Information の直和分解が成立する、と略記する。

上記の結果は、「離散分布の場合にも『Sum of Squares Total は Sum of Squares Within と Sum of Squares Between の和 という分散分析の基本結果』に対応したものがある」ことを示している。本稿の筆者はこれまでに、離散モデルの2標本問題において、Kullback Information の直和分解が成立する場合、不成立の場合についていくつかの成果を得てきたが、本稿では新たな結果を加えて報告する。

2. 多項分布の場合

Proposition 1. は多項分布の2標本問題にて考察したものであるが、これを pooling incomplete samples を伴う多項分布 (Asano, 1965) の場合に拡張できる。

Proposition 2.1 $X^{[1]}, X^{[2]}, Y^{[1]}, Y^{[2]}$ は互いに独立で、各 $i = 1, 2$ に対して

$$\left. \begin{aligned} X^{[i]} &= (X_1^{[i]}, \dots, X_m^{[i]}, \dots, X_k^{[i]}) \sim \text{Multinomial}(N_x^{[i]}, p_1^{[i]}, \dots, p_m^{[i]}, \dots, p_k^{[i]}), \\ Y^{[i]} &= (Y_1^{[i]}, \dots, Y_m^{[i]}) \sim \text{Multinomial}(N_y^{[i]}, \frac{p_1^{[i]}}{\sum_{j=1}^m p_j^{[i]}}, \dots, \frac{p_m^{[i]}}{\sum_{j=1}^m p_j^{[i]}}) \end{aligned} \right\} \quad (5)$$

と仮定する。仮説 H_1, H_2 を、

$$H_2 : p_j^{[1]} = p_j^{[2]} = p_j \text{ for all } j = 1, \dots, k, \quad H_1 : \text{not } H_2 \quad (6)$$

に選ぶ。このとき、Kullback Information の直和分解が成り立つ。

Proposition 2.1 は、pooling incomplete samples が1回行われた場合である。これが階層構造で複数回行われても、同様な結果が成り立つ。たとえば、

Proposition 2.2 $\mathbf{X}^{[i]} = (X_1^{[i]}, \dots, X_{m(1)}^{[i]}, \dots, X_{m(2)}^{[i]}, \dots, X_k^{[i]})$, $\mathbf{Y}^{[i]} = (Y_1^{[i]}, \dots, Y_{m(1)}^{[i]}, \dots, Y_{m(2)}^{[i]})$, $\mathbf{Z}^{[i]} = (Z_1^{[i]}, \dots, Z_{m(1)}^{[i]})$, $(i = 1, 2)$ が互いに独立で

$$\mathbf{X}^{[i]} \sim \text{Multinomial}(N_x^{[i]}; \theta_1^{[i]}, \dots, \theta_{m(1)}^{[i]}, \dots, \theta_{m(2)}^{[i]}, \dots, \theta_k^{[i]}),$$

$$\mathbf{Y}^{[i]} \sim \text{Multinomial}\left(N_y^{[i]}; \frac{\theta_1^{[i]}}{\sum_{a=1}^{m(2)} \theta_a^{[i]}}, \dots, \frac{\theta_{m(1)}^{[i]}}{\sum_{a=1}^{m(2)} \theta_a^{[i]}}, \dots, \frac{\theta_{m(2)}^{[i]}}{\sum_{a=1}^{m(2)} \theta_a^{[i]}}\right),$$

$$\mathbf{Z}^{[i]} \sim \text{Multinomial}\left(N_z^{[i]}; \frac{\theta_1^{[i]}}{\sum_{a=1}^{m(1)} \theta_a^{[i]}}, \dots, \frac{\theta_{m(1)}^{[i]}}{\sum_{a=1}^{m(1)} \theta_a^{[i]}}\right),$$

とする。仮説 H_1, H_2 を $H_2: \theta_j^{[1]} = \theta_j^{[2]} = \theta_j$, $H_1: \text{not } H_2$ に選ぶ。このとき、Kullback Information の直和分解が成り立つ。

注：
$$\frac{\theta_j^{[i]} / \sum_{a=1}^{m(2)} \theta_a^{[i]}}{\sum_{l=1}^{m(1)} (\theta_l^{[i]} / \sum_{a=1}^{m(2)} \theta_a^{[i]})} = \frac{\theta_j^{[i]}}{\sum_{a=1}^{m(1)} \theta_a^{[i]}} \quad j = 1, \dots, m(1).$$

さて、Proposition 2.1 におけるモデルの確率構造および仮説、すなわち、(5), (6) は、 $s = \sum_{l=1}^m p_l$, $u_j = p_j / \sum_{l=1}^m p_l$, $(j \leq m)$, $v_j = p_j / \sum_{l=m+1}^k p_l$ $(j > m)$ とパラメータ変換することで、次のように書き直せる：

$$\left. \begin{aligned} \mathbf{X}^{[i]} &= (X_1^{[i]}, \dots, X_m^{[i]}, X_{m+1}^{[i]}, \dots, X_k^{[i]}) \\ &\sim \text{Multinomial}(N_x^{[i]}; s^{[i]}u_1^{[i]}, \dots, s^{[i]}u_m^{[i]}, (1-s^{[i]})v_{m+1}^{[i]}, \dots, (1-s^{[i]})v_k^{[i]}), \\ \mathbf{Y}^{[i]} &= (Y_1^{[i]}, \dots, Y_m^{[i]}) \sim \text{Multinomial}(N_y^{[i]}; u_1^{[i]}, \dots, u_m^{[i]}) \end{aligned} \right\} \quad (7)$$

$$\left. \begin{aligned} H_2: u_j^{[1]} &= u_j^{[2]} = u_j, \quad v_j^{[1]} = v_j^{[2]} = v_j \text{ for all } j, \text{ and } s^{[1]} = s^{[2]} = s \\ H_1: &\text{not } H_2. \end{aligned} \right\} \quad (8)$$

さて、確率構造に関する仮定 (7) は、次の表 2.1 のように書き表すことができる。

表 2.1

$\mathbf{X}^{[i]}$	$X_1^{[i]}, \dots, X_m^{[i]}$	$X_{m+1}^{[i]}, \dots, X_k^{[i]}$
	$s^{[i]}$	$1 - s^{[i]}$
$\mathbf{Y}^{[i]}$	$Y_1^{[i]}, \dots, Y_m^{[i]}$	
	$u_1^{[i]}, \dots, u_m^{[i]}$	$v_{m+1}^{[i]}, \dots, v_k^{[i]}$

以降、本稿では、混乱の恐れがない限り、多項分布の確率構造に関する仮定を (7) のように記載する代わりに表 2.1 のように略記する。また、仮説 H_1, H_2 は、常に H_2 : パラメータは index $i = 1, 2$ に関して等しい (つまり、 $s^{[1]} = s^{[2]} = s$ 等), $H_1: \text{not } H_2$ に選ぶ。

このような表記により、Proposition 2.2 は、変数変換 $s^{[i]} = \sum_{l=1}^{m(2)} \theta_l^{[i]}$, $t^{[i]} = \frac{\sum_{l=1}^{m(1)} \theta_l^{[i]}}{\sum_{l=1}^{m(2)} \theta_l^{[i]}}$,
 $u_j^{[i]} = \frac{\theta_j^{[i]}}{\sum_{l=1}^{m(1)} \theta_l^{[i]}}$, $v_j^{[i]} = \frac{\theta_j^{[i]}}{\sum_{l=m(1)+1}^{m(2)} \theta_l^{[i]}}$, $w_j^{[i]} = \frac{\theta_j^{[i]}}{\sum_{l=m(2)+1}^k \theta_l^{[i]}}$ によって、

$\mathbf{X}^{[i]}$, $\mathbf{Y}^{[i]}$, $\mathbf{Z}^{[i]}$ ($i = 1, 2$) が互いに独立な多項分布で、その確率構造が表 2.2 で与えられているならば、Kullback Information の直和分解が成り立つ

と言いつけることができる。

表 2.2

$\mathbf{X}^{[i]}$	$X_1^{[i]}, \dots, X_{m(1)}^{[i]}$	$X_{m(1)+1}^{[i]}, \dots, X_{m(2)}^{[i]}$	$X_{m(2)+1}^{[i]}, \dots, X_k^{[i]}$
	$s^{[i]}t^{[i]}$	$s^{[i]}(1-t^{[i]})$	$1-s^{[i]}$
	$u_1^{[i]}, \dots, u_{m(1)}^{[i]}$	$v_{m(1)+1}^{[i]}, \dots, v_{m(2)}^{[i]}$	$w_{m(2)+1}^{[i]}, \dots, w_k^{[i]}$
$\mathbf{Y}^{[i]}$	$Y_1^{[i]}, \dots, Y_{m(1)}^{[i]}$	$Y_{m(1)+1}^{[i]}, \dots, Y_{m(2)}^{[i]}$	
	$t^{[i]}$	$1-t^{[i]}$	
	$u_1^{[i]}, \dots, u_{m(1)}^{[i]}$	$v_{m(1)+1}^{[i]}, \dots, v_{m(2)}^{[i]}$	
$\mathbf{Z}^{[i]}$	$Z_1^{[i]}, \dots, Z_{m(1)}^{[i]}$		
	$u_1^{[i]}, \dots, u_{m(1)}^{[i]}$		

Proposition 2.2 と同様な結果も、直ちに得られる。

Proposition 2.3 $\mathbf{X}^{[i]}$, $\mathbf{Y}^{[i]}$ ($i = 1, 2$) が互いに独立な多項分布で、その確率構造が表 2.3 で与えられているならば、Kullback Information の直和分解が成り立つ

表 2.3

$\mathbf{X}^{[i]}$	$X_1^{[i]}, \dots, X_{m(1)}^{[i]}$	$X_{m(1)+1}^{[i]}, \dots, X_{m(2)}^{[i]}$	$X_{m(2)+1}^{[i]}, \dots, X_k^{[i]}$
	$s^{[i]}t^{[i]}$	$s^{[i]}(1-t^{[i]})$	$1-s^{[i]}$
	$u_1^{[i]}, \dots, u_{m(1)}^{[i]}$	$v_{m(1)+1}^{[i]}, \dots, v_{m(2)}^{[i]}$	$w_{m(2)+1}^{[i]}, \dots, w_k^{[i]}$
$\mathbf{Y}^{[i]}$	$Y_1^{[i]}, \dots, Y_{m(1)}^{[i]}$	$Y_{m(1)+1}^{[i]}, \dots, Y_{m(2)}^{[i]}$	
	$t^{[i]}$	$1-t^{[i]}$	
	$u_1^{[i]}, \dots, u_{m(1)}^{[i]}$	$v_{m(1)+1}^{[i]}, \dots, v_{m(2)}^{[i]}$	

なお パラメータに関しては、各 $i = 1, 2$ に対して $0 \leq s^{[i]}, t^{[i]}, u_j^{[i]}, v_j^{[i]}, w_j^{[i]} \leq 1$, $\sum_{j=1}^{m(1)} u_j^{[i]} = 1$, $\sum_{j=m(1)+1}^{m(2)} v_j^{[i]} = 1$, $\sum_{j=m(2)+1}^k w_j^{[i]} = 1$ が満たされるものとする。(パラメータに関するこの仮定は、以降、暗黙の仮定として、本節で特に断わることなく用いる)

Remark: Proposition 2.3 の確率構造を「 $\mathbf{X}^{[i]}$ と $\mathbf{Y}^{[i]}$ のパラメータが比例関係にある」と述べることができる。

他方、Proposition 2.3 とかなり似通った確率構造を持っていても、Kullback Information の直和分解が成り立たないことがある。

例 2.1 $\mathbf{X}^{[i]}$, $\mathbf{Y}^{[i]}$ ($i = 1, 2$) は互いに独立な多項分布で、その確率構造が表 2.4 で与えら

れている。このとき、Kullback Information の直和分解は不成立。

表 2.4

$\mathbf{X}^{[i]}$	$X_1^{[i]}, \dots, X_{m(1)}^{[i]}$	$X_{m(1)+1}^{[i]}, \dots, X_{m(2)}^{[i]}$	$X_{m(2)+1}^{[i]}, \dots, X_k^{[i]}$
	$s^{[i]}$	$(1 - s^{[i]})t^{[i]}$	$(1 - s^{[i]})(1 - t^{[i]})$
	$u_1^{[i]}, \dots, u_{m(1)}^{[i]}$	$v_{m(1)+1}^{[i]}, \dots, v_{m(2)}^{[i]}$	$w_{m(2)+1}^{[i]}, \dots, w_k^{[i]}$
$\mathbf{Y}^{[i]}$	$Y_1^{[i]}, \dots, Y_{m(1)}^{[i]}$	$Y_{m(1)+1}^{[i]}, \dots, Y_{m(2)}^{[i]}$	
	$s^{[i]}$	$1 - s^{[i]}$	
	$u_1^{[i]}, \dots, u_{m(1)}^{[i]}$	$v_{m(1)+1}^{[i]}, \dots, v_{m(2)}^{[i]}$	

例 2.2 $\mathbf{X}^{[i]}, \mathbf{Y}^{[i]}, \mathbf{Z}^{[i]}$, ($i = 1, 2$) は互いに独立な多項分布で、その確率構造が表 2.5 で与えられている。このとき、Kullback Information の直和分解は不成立。

表 2.5

$\mathbf{X}^{[i]}$	$X_1^{[i]}, \dots, X_{m(1)}^{[i]}$	$X_{m(1)+1}^{[i]}, \dots, X_{m(2)}^{[i]}$	$X_{m(2)+1}^{[i]}, \dots, X_k^{[i]}$
	$s^{[i]}t^{[i]}$	$s^{[i]}(1 - t^{[i]})$	$1 - s^{[i]}$
	$u_1^{[i]}, \dots, u_{m(1)}^{[i]}$	$v_{m(1)+1}^{[i]}, \dots, v_{m(2)}^{[i]}$	$w_{m(2)+1}^{[i]}, \dots, w_k^{[i]}$
$\mathbf{Y}^{[i]}$	$Y_1^{[i]}, \dots, Y_{m(1)}^{[i]}$	$Y_{m(1)+1}^{[i]}, \dots, Y_{m(2)}^{[i]}$	
	$t^{[i]}$	$1 - t^{[i]}$	
	$u_1^{[i]}, \dots, u_{m(1)}^{[i]}$	$v_{m(1)+1}^{[i]}, \dots, v_{m(2)}^{[i]}$	
$\mathbf{Z}^{[i]}$	$Z_1^{[i]}, \dots, Z_{m(1)}^{[i]}$		$Z_{m(2)+1}^{[i]}, \dots, Z_k^{[i]}$
	$s^{[i]}$		$1 - s^{[i]}$
	$u_1^{[i]}, \dots, u_{m(1)}^{[i]}$		$w_{m(2)+1}^{[i]}, \dots, w_k^{[i]}$

Remark. 例 2.1, 例 2.2 にて、Kullback Information の直和分解が成り立たないのは、「カテゴリーの合併。その後同じ割合での配分」が原因と思える。例 2.1 の場合、確率変数 $X_{m(1)+1}^{[i]}, \dots, X_{m(2)}^{[i]}, X_{m(2)+1}^{[i]}, \dots, X_k^{[i]}$ のカテゴリー $\{m(1)+1, \dots, m(2), m(2)+1, \dots, k\}$ が一旦合併してから 確率変数 $Y_{m(1)+1}^{[i]}, \dots, Y_{m(2)}^{[i]}$ により $\{m(1)+1, \dots, m(2)\} \rightarrow X_{m(1)+1}^{[i]}, \dots, X_{m(2)}^{[i]}$ の生起確率と同じ割合で分岐している。例 2.2 では、確率変数 $X_1^{[i]}, \dots, X_{m(2)}^{[i]}, X_{m(2)+1}^{[i]}, \dots, X_k^{[i]}$ のカテゴリー $\{m(1)+1, \dots, m(1), m(1)+1, \dots, m(2)\}$ が一旦合併してから 確率変数 $Z_1^{[i]}, \dots, Z_{m(1)}^{[i]}$ により $\{1, \dots, m(1)\} \rightarrow X_1^{[i]}, \dots, X_{m(1)}^{[i]}$ の生起確率と同じ割合で分岐している。このことが直和分解不成立の原因と思われるが、詳細説明までには至っていない。

Proposition 2.4 $\mathbf{X}^{[i]}, \mathbf{Y}^{[i]}, \mathbf{Z}^{[i]}$, ($i = 1, 2$) は互いに独立な多項分布で、その確率構造が表 2.6 で与えられている。このとき、Kullback Information の直和分解が成り立つ。

表 2.6

$\mathbf{X}^{[i]}$	$X_1^{[i]}, \dots, X_{m(1)}^{[i]}$	$X_{m(1)+1}^{[i]}, \dots, X_{m(2)}^{[i]}$	$X_{m(2)+1}^{[i]}, \dots, X_k^{[i]}$
	$s_1^{[i]} t^{[i]}$	$s_1^{[i]} (1 - t^{[i]})$	$1 - s_1^{[i]}$
	$u_1^{[i]}, \dots, u_{m(1)}^{[i]}$	$v_{m(1)+1}^{[i]}, \dots, v_{m(2)}^{[i]}$	$w_{m(2)+1}^{[i]}, \dots, w_k^{[i]}$
$\mathbf{Y}^{[i]}$	$Y_1^{[i]}, \dots, Y_{m(1)}^{[i]}$	$Y_{m(1)+1}^{[i]}, \dots, Y_{m(2)}^{[i]}$	
	$t^{[i]}$	$1 - t^{[i]}$	
	$u_1^{[i]}, \dots, u_{m(1)}^{[i]}$	$v_{m(1)+1}^{[i]}, \dots, v_{m(2)}^{[i]}$	
$\mathbf{Z}^{[i]}$	$Z_1^{[i]}, \dots, Z_{m(1)}^{[i]}$	$Z_{m(1)+1}^{[i]}, \dots, Z_{m(2)}^{[i]}$	$Z_{m(2)+1}^{[i]}, \dots, Z_k^{[i]}$
	$s_2^{[i]} t^{[i]}$	$s_2^{[i]} (1 - t^{[i]})$	$1 - s_2^{[i]}$
	$u_1^{[i]}, \dots, u_{m(1)}^{[i]}$	$v_{m(1)+1}^{[i]}, \dots, v_{m(2)}^{[i]}$	$w_{m(2)+1}^{[i]}, \dots, w_k^{[i]}$

Remark: Proposition 2.4 は、Proposition 2.3 の拡張と考えることができる。つまり、 $\mathbf{Y}^{[i]}$ は $\mathbf{X}^{[i]}$ のパラメータと比例関係にあると同時に、 $\mathbf{Z}^{[i]}$ のパラメータとも比例関係にある。このような状況下においても、Kullback Information の直和分解が成り立つ。

例 2.3 独立な多項分布 $\mathbf{X}^{[i]}, \mathbf{Y}^{[i]}, \mathbf{Z}^{[i]}, \mathbf{W}^{[i]}$ ($i = 1, 2$) が表 2.7 の確率構造で書き表せるとき、Kullback 情報量の直和分解は不成立。

表 2.7

$\mathbf{X}^{[i]}$	$X_1^{[i]}, \dots, X_{m(1)}^{[i]}$	$X_{m(1)+1}^{[i]}, \dots, X_{m(2)}^{[i]}$	$X_{m(2)+1}^{[i]}, \dots, X_k^{[i]}$
	$s^{[i]} t_1^{[i]}$	$s^{[i]} (1 - t_1^{[i]})$	$1 - s^{[i]}$
	$\mu_1^{[i]}, \dots, \mu_{m(1)}^{[i]}$	$\xi_{m(1)+1}^{[i]}, \dots, \xi_{m(2)}^{[i]}$	$\eta_{m(2)+1}^{[i]}, \dots, \eta_k^{[i]}$
$\mathbf{Y}^{[i]}$	$Y_1^{[i]}, \dots, Y_{m(1)}^{[i]}$	$Y_{m(1)+1}^{[i]}, \dots, Y_{m(2)}^{[i]}$	
	$t_1^{[i]}$	$1 - t_1^{[i]}$	
	$\mu_1^{[i]}, \dots, \mu_{m(1)}^{[i]}$	$\xi_{m(1)+1}^{[i]}, \dots, \xi_{m(2)}^{[i]}$	
$\mathbf{Z}^{[i]}$	$Z_1^{[i]}, \dots, Z_{m(1)}^{[i]}$	$Z_{m(1)+1}^{[i]}, \dots, Z_{m(2)}^{[i]}$	$Z_{m(2)+1}^{[i]}, \dots, Z_k^{[i]}$
	$s^{[i]} t_2^{[i]}$	$s^{[i]} (1 - t_2^{[i]})$	$1 - s^{[i]}$
	$\mu_1^{[i]}, \dots, \mu_{m(1)}^{[i]}$	$\xi_{m(1)+1}^{[i]}, \dots, \xi_{m(2)}^{[i]}$	$\eta_{m(2)+1}^{[i]}, \dots, \eta_k^{[i]}$
$\mathbf{W}^{[i]}$	$W_1^{[i]}, \dots, W_{m(1)}^{[i]}$	$W_{m(1)+1}^{[i]}, \dots, W_{m(2)}^{[i]}$	
	$t_2^{[i]}$	$1 - t_2^{[i]}$	
	$\mu_1^{[i]}, \dots, \mu_{m(1)}^{[i]}$	$\xi_{m(1)+1}^{[i]}, \dots, \xi_{m(2)}^{[i]}$	

なお、例 2.3 におけるパラメータ $0 \leq \mu_j^{[i]} \leq 1$, $0 \leq \xi_j^{[i]} \leq 1$, $0 \leq \eta_j^{[i]} \leq 1$ は、

$$\sum_{j=1}^{m(1)} \mu_j^{[i]} = 1, \quad \sum_{j=m(1)+1}^{m(2)} \xi_j^{[i]} = 1, \quad \sum_{j=m(2)+1}^k \eta_j^{[i]} = 1 \text{ を満たすものである。}$$

例 2.4 独立な多項分布 $\mathbf{X}^{[i]}, \mathbf{Y}^{[i]}, \mathbf{Z}^{[i]}$ ($i = 1, 2$) が下記のような確率構造で書き表せると

き、Kullback 情報量の直和分解は不成立。

表 2.8

$\mathbf{X}^{[i]}$	$X_1^{[i]}, \dots, X_{m(1)}^{[i]}$	$X_{m(1)+1}^{[i]}, \dots, X_{m(2)}^{[i]}$	$X_{m(2)+1}^{[i]}, \dots, X_k^{[i]}$
	$s^{[i]}t_1^{[i]}$	$s^{[i]}(1 - t_1^{[i]})$	$1 - s$
	$u_1^{[i]}, \dots, u_{m(1)}^{[i]}$	$v_{m(1)+1}^{[i]}, \dots, v_{m(2)}^{[i]}$	$w_{m(2)+1}^{[i]}, \dots, w_k^{[i]}$
$\mathbf{Y}^{[i]}$	$Y_1^{[i]}, \dots, Y_{m(1)}^{[i]}$	$Y_{m(1)+1}^{[i]}, \dots, Y_{m(2)}^{[i]}$	
	$t_1^{[i]}$	$1 - t_1^{[i]}$	
	$u_1^{[i]}, \dots, u_{m(1)}^{[i]}$	$v_{m(1)+1}^{[i]}, \dots, v_{m(2)}^{[i]}$	
$\mathbf{Z}^{[i]}$	$Z_1^{[i]}, \dots, Z_{m(1)}^{[i]}$	$Z_{m(1)+1}^{[i]}, \dots, Z_{m(2)}^{[i]}$	$Z_{m(2)+1}^{[i]}, \dots, Z_k^{[i]}$
	$s^{[i]}t_2^{[i]}$	$s^{[i]}(1 - t_2^{[i]})$	$1 - s^{[i]}$
	$u_1^{[i]}, \dots, u_{m(1)}^{[i]}$	$v_{m(1)+1}^{[i]}, \dots, v_{m(2)}^{[i]}$	$w_{m(2)+1}^{[i]}, \dots, w_k^{[i]}$

Remark: 例 2.4 は「Proposition 2.4 におけるパラメータ $s_1^{[i]}, s_2^{[i]}$ を $s_1^{[i]} = s_2^{[i]} = s^{[i]}$ とし、パラメータ $t^{[i]}$ を別のものにした」確率モデルである。Proposition 2.4 では直和分解が成り立つのに例 2.4 では Kullback 情報量の直和分解が成り立たない。この原因が何であるか。著者には全くわからず、困っている。

例 2.5 独立な多項分布 $\mathbf{X}^{[i]}, \mathbf{Y}^{[i]}$ ($i = 1, 2$) が

表 2.9

$\mathbf{X}^{[i]}$	$X_1^{[i]}, \dots, X_{m(1)}^{[i]}$	$X_{m(1)+1}^{[i]}, \dots, X_{m(2)}^{[i]}$	$X_{m(2)+1}^{[i]}, \dots, X_{m(3)}^{[i]}$	$X_{m(3)+1}^{[i]}, \dots, X_k^{[i]}$
	$s^{[i]}t^{[i]}r^{[i]}$	$s_1^{[i]}t^{[i]}(1 - r^{[i]})$	$s^{[i]}(1 - t^{[i]})$	$1 - s^{[i]}$
	$\mu_1^{[i]}, \dots, \mu_{m(1)}^{[i]}$	$\xi_{m(1)+1}^{[i]}, \dots, \xi_{m(2)}^{[i]}$	$\zeta_{m(2)+1}^{[i]}, \dots, \zeta_{m(3)}^{[i]}$	$\eta_{m(3)+1}^{[i]}, \dots, \eta_k^{[i]}$
$\mathbf{Y}^{[i]}$	$Y_1^{[i]}, \dots, Y_{m(1)}^{[i]}$	$Y_{m(1)+1}^{[i]}, \dots, Y_{m(2)}^{[i]}$		$Y_{m(3)+1}^{[i]}, \dots, Y_k^{[i]}$
	$s^{[i]}r^{[i]}$	$s^{[i]}(1 - r^{[i]})$		$1 - s^{[i]}$
	$\mu_1^{[i]}, \dots, \mu_{m(1)}^{[i]}$	$\xi_{m(1)+1}^{[i]}, \dots, \xi_{m(2)}^{[i]}$		$\eta_{m(3)+1}^{[i]}, \dots, \eta_k^{[i]}$

なる確率構造で書き表せるとき、Kullback 情報量の直和分解不成立。パラメータ s を $\mathbf{X}^{[i]}$ と $\mathbf{X}^{[i]}$ では別々のものにとすると、すなわち、確率構造が表 2.10 で与えられるとき、Kullback 情報量の直和分解は成立。

表 2.10

$\mathbf{X}^{[i]}$	$X_1^{[i]}, \dots, X_{m(1)}^{[i]}$	$X_{m(1)+1}^{[i]}, \dots, X_{m(2)}^{[i]}$	$X_{m(2)+1}^{[i]}, \dots, X_{m(3)}^{[i]}$	$X_{m(3)+1}^{[i]}, \dots, X_k^{[i]}$
	$s_1^{[i]}t^{[i]}r^{[i]}$	$s_1^{[i]}t^{[i]}(1 - r^{[i]})$	$s_1^{[i]}(1 - t^{[i]})$	$1 - s_1^{[i]}$
	$\mu_1^{[i]}, \dots, \mu_{m(1)}^{[i]}$	$\xi_{m(1)+1}^{[i]}, \dots, \xi_{m(2)}^{[i]}$	$\zeta_{m(2)+1}^{[i]}, \dots, \zeta_{m(3)}^{[i]}$	$\eta_{m(3)+1}^{[i]}, \dots, \eta_k^{[i]}$
$\mathbf{Y}^{[i]}$	$Y_1^{[i]}, \dots, Y_{m(1)}^{[i]}$	$Y_{m(1)+1}^{[i]}, \dots, Y_{m(2)}^{[i]}$		$Y_{m(3)+1}^{[i]}, \dots, Y_k^{[i]}$
	$s_2^{[i]}r^{[i]}$	$s_2^{[i]}(1 - r^{[i]})$		$1 - s_2^{[i]}$
	$\mu_1^{[i]}, \dots, \mu_{m(1)}^{[i]}$	$\xi_{m(1)+1}^{[i]}, \dots, \xi_{m(2)}^{[i]}$		$\eta_{m(3)+1}^{[i]}, \dots, \eta_k^{[i]}$

3. 負の多項分布 (NM) の場合

例 3.1 $X^{[1]}, X^{[2]}, Y^{[1]}, Y^{[2]}$ は互いに独立で、各 $i = 1, 2$ に対し

$$\begin{aligned} \mathbf{X}^{[i]} &= (X_1^{[i]}, \dots, X_m^{[i]}, \dots, X_k^{[i]}) \sim \text{NM}(r_x^{[i]}, \theta_1^{[i]}, \dots, \theta_m^{[i]}, \dots, \theta_k^{[i]}), \\ \mathbf{Y}^{[i]} &= (Y_1^{[i]}, \dots, Y_m^{[i]}) \sim \text{NM}(r_y^{[i]}, \sum_{j=1}^k \theta_l^{[i]} \frac{\theta_1^{[i]}}{\sum_{l=1}^m \theta_l^{[i]}}, \dots, \sum_{j=1}^k \theta_l^{[i]} \frac{\theta_m^{[i]}}{\sum_{l=1}^m \theta_l^{[i]}}) \end{aligned}$$

なる確率モデルでは、Kullback 情報量の直和分解不成立。なお、上記の確率モデルにて、パラメータ $\theta_j^{[i]}$ の制約条件は $\sum_{j=1}^m \theta_j^{[i]} \leq 1$ であって、 $\sum_{j=1}^m \theta_j^{[i]} = 1$ ではない。

他方、次の場合には、Kullback 情報量の直和分解が成立する。

Proposition 3.1 上記例 3.1 の確率変数 $\mathbf{Y}^{[i]}$ に関する仮定を

$$\mathbf{Y}^{[i]} = (Y_1^{[i]}, \dots, Y_m^{[i]}) \sim \text{Multinomial}(N_y^{[i]}, \frac{\theta_1^{[i]}}{\sum_{j=1}^m \theta_j^{[i]}}, \dots, \frac{\theta_m^{[i]}}{\sum_{j=1}^m \theta_j^{[i]}})$$

に変更すると、Kullback 情報量の直和分解が成立する

Proposition 3.1 は、「ベースとなるのが負の多項分布。これに Pooling incomplete sample として付随するのが多項分布。このような確率モデルにおける 2 標本問題」と言える。しかし、このような確率モデルの状況設定は、自然なものとは言い難い。

さて、

$$s^{[i]} = \sum_{l=1}^k \theta_l^{[i]}, \quad t^{[i]} = \frac{\sum_{l=1}^m \theta_l^{[i]}}{\sum_{l=1}^k \theta_l^{[i]}}, \quad u_j^{[i]} = \frac{\theta_j^{[i]}}{\sum_{l=1}^m \theta_l^{[i]}} \quad (j \leq m), \quad v_j^{[i]} = \frac{\theta_j^{[i]}}{\sum_{l=m+1}^k \theta_l^{[i]}} \quad (j > m)$$

とパラメータ変換することで、例 3.1, Proposition 3.1 の確率構造は、それぞれ表 3.1 および表 3.2 で書き表せる。

表 3.1

$\mathbf{X}^{[i]}$	$r_x^{[i]}$	$X_1^{[i]}, \dots, X_m^{[i]}$	$X_{m+1}^{[i]}, \dots, X_k^{[i]}$
	$1 - s^{[i]}$	$s^{[i]} t^{[i]}$	$s^{[i]} (1 - t^{[i]})$
		$u_1^{[i]}, \dots, u_m^{[i]}$	$v_{m+1}^{[i]}, \dots, v_k^{[i]}$
$\mathbf{Y}^{[i]}$	$r_y^{[i]}$	$Y_1^{[i]}, \dots, Y_m^{[i]}$	
	$1 - s^{[i]}$	$s^{[i]}$	
		$u_1^{[i]}, \dots, u_m^{[i]}$	

表 3.2

$\mathbf{X}^{[i]}$	$r_x^{[i]}$	$X_1^{[i]}, \dots, X_m^{[i]}$	$X_{m+1}^{[i]}, \dots, X_k^{[i]}$
	$1 - s^{[i]}$	$s^{[i]} t^{[i]}$	$s^{[i]} (1 - t^{[i]})$
		$u_1^{[i]}, \dots, u_m^{[i]}$	$v_{m+1}^{[i]}, \dots, v_k^{[i]}$
$\mathbf{Y}^{[i]}$		$Y_1^{[i]}, \dots, Y_m^{[i]}$	
		$u_1^{[i]}, \dots, u_m^{[i]}$	

例 3.1 の確率構造において、パラメータ $s^{[i]}$ を $\mathbf{X}^{[i]}$, $\mathbf{Y}^{[i]}$ で別々のものにとすると、Kullback 情報量の直和分解が成立する。

Proposition 3.2 負の多項分布 $\mathbf{X}^{[1]}$, $\mathbf{X}^{[2]}$, $\mathbf{Y}^{[1]}$, $\mathbf{Y}^{[2]}$ は互いに独立で、その確率構造が表 3.3 で与えられるならば、Kullback 情報量の直和分解が成立する。

表 3.3

$\mathbf{X}^{[i]}$	$r_x^{[i]}$	$X_1^{[i]}, \dots, X_m^{[i]}$	$X_{m+1}^{[i]}, \dots, X_k^{[i]}$
	$1 - s_1^{[i]}$	$s_1^{[i]} t^{[i]}$	$s_1^{[i]} (1 - t^{[i]})$
		$u_1^{[i]}, \dots, u_m^{[i]}$	$v_{m+1}^{[i]}, \dots, v_k^{[i]}$
$\mathbf{Y}^{[i]}$	$r_y^{[i]}$	$Y_1^{[i]}, \dots, Y_m^{[i]}$	
	$1 - s_2^{[i]}$	$s_2^{[i]}$	
		$u_1^{[i]}, \dots, u_m^{[i]}$	

Remark 例 3.1 すなわち、 $\mathbf{X}^{[i]}$, $\mathbf{Y}^{[i]}$ が負の多項分布で、確率構造が表 3.1 の時は、直和分解不成立。このことは、 $\mathbf{X}^{[i]}$, $\mathbf{Y}^{[i]}$ が多項分布で確率構造が表 3.4 で与えられている時に直和分解不成立だったことから類推できる

表 3.4

$\mathbf{X}^{[i]}$	$X_1^{[i]}, \dots, X_{m(1)}^{[i]}$	$X_{m(1)+1}^{[i]}, \dots, X_{m(2)}^{[i]}$	$X_{m(2)+1}^{[i]}, \dots, X_k^{[i]}$
	$1 - s^{[i]}$	$s^{[i]} t^{[i]}$	$s^{[i]} (1 - t^{[i]})$
	$u_1^{[i]}, \dots, u_{m(1)}^{[i]}$	$v_{m(1)+1}^{[i]}, \dots, v_{m(2)}^{[i]}$	$w_{m(2)+1}^{[i]}, \dots, w_k^{[i]}$
$\mathbf{Y}^{[i]}$	$Y_1^{[i]}, \dots, Y_{m(1)}^{[i]}$	$Y_{m(1)+1}^{[i]}, \dots, Y_{m(2)}^{[i]}$	
	$1 - s^{[i]}$	$s^{[i]}$	
	$u_1^{[i]}, \dots, u_{m(1)}^{[i]}$	$v_{m(1)+1}^{[i]}, \dots, v_{m(2)}^{[i]}$	

Remark 2 $\mathbf{X}^{[i]}$, $\mathbf{Y}^{[i]}$ が負の多項分布で、確率構造が表 3.5 の時は、直和分解が成立

表 3.5

$\mathbf{X}^{[i]}$	$r_x^{[i]}$	$X_1^{[i]}, \dots, X_m^{[i]}$	$X_{m+1}^{[i]}, \dots, X_k^{[i]}$
	$1 - s_1^{[i]}$	$s_1^{[i]} t^{[i]}$	$s_1^{[i]} (1 - t^{[i]})$
		$u_1^{[i]}, \dots, u_m^{[i]}$	$v_{m+1}^{[i]}, \dots, v_k^{[i]}$
$\mathbf{Y}^{[i]}$	$r_y^{[i]}$	$Y_1^{[i]}, \dots, Y_m^{[i]}$	
	$1 - s_2^{[i]}$	$s_2^{[i]}$	
		$u_1^{[i]}, \dots, u_m^{[i]}$	

これは、 $\mathbf{X}^{[i]}$, $\mathbf{Y}^{[i]}$:多項分布で、確率構造が表 3.6 の時は、直和分解が成立することから、負の多項分布の場合でも直和分解成立が類推できる

表 3.6

$\mathbf{X}^{[i]}$	$X_1^{[i]}, \dots, X_{m(1)}^{[i]}$	$X_{m(1)+1}^{[i]}, \dots, X_{m(2)}^{[i]}$	$X_{m(2)+1}^{[i]}, \dots, X_k^{[i]}$
	$1 - s_1^{[i]}$	$s_1^{[i]} t^{[i]}$	$s_1^{[i]} (1 - t^{[i]})$
	$u_1^{[i]}, \dots, u_{m(1)}^{[i]}$	$v_{m(1)+1}^{[i]}, \dots, v_{m(2)}^{[i]}$	$w_{m(2)+1}^{[i]}, \dots, w_k^{[i]}$
$\mathbf{Y}^{[i]}$	$Y_1^{[i]}, \dots, Y_{m(1)}^{[i]}$	$Y_{m(1)+1}^{[i]}, \dots, Y_{m(2)}^{[i]}$	
	$1 - s_2^{[i]}$	$s_2^{[i]}$	
	$u_1^{[i]}, \dots, u_{m(1)}^{[i]}$	$v_{m(1)+1}^{[i]}, \dots, v_{m(2)}^{[i]}$	

Remark 3 負の多項分布は、「負の 2 項分布 × 多項分布」と分解できるので、「多項分布の部分」で Pooling incomplete samples が行われていれば、直和分解が成立する、と言える。

4. Poisson 分布の場合

例 4.1 $m < k$ とする。 $X_j^{[i]}$ ($i = 1, 2, j = 1, 2, \dots, k$) および $Y_j^{[i]}$ ($i = 1, 2, j = 1, 2, \dots, m$) はすべて独立で

$$X_j^{[i]} \sim \text{Poisson}(\lambda_j^{[i]}), \quad j = 1, \dots, k,$$

$$Y_j^{[i]} \sim \text{Poisson}((\lambda_1^{[i]} + \dots + \lambda_m^{[i]} + \dots + \lambda_k^{[i]}) \frac{\lambda_j^{[i]}}{\lambda_1^{[i]} + \dots + \lambda_m^{[i]}}), \quad j = 1, \dots, m.$$

$$H_1 : (\lambda_1^{[1]}, \dots, \lambda_k^{[1]}) \neq (\lambda_1^{[2]}, \dots, \lambda_k^{[2]}), \quad H_2 : (\lambda_1^{[1]}, \dots, \lambda_k^{[1]}) = (\lambda_1^{[2]}, \dots, \lambda_k^{[2]}).$$

この場合、Kullback 情報量の直和分解は成り立たない。

Proposition 4.1 $m \leq k$ とする。 $X_j^{[i]}$ ($i = 1, 2, j = 1, 2, \dots, k$) および $\mathbf{Y}^{[i]}$ ($i = 1, 2$) は互いに独立で

$$X_j^{[i]} \sim \text{Poisson}(\lambda_j^{[i]}), \quad i = 1, \dots, k,$$

$$\mathbf{Y}^{[i]} = (Y_1^{[i]}, \dots, Y_m^{[i]}) \sim \text{Multinomial}(N_y^{[i]}, \frac{\lambda_1^{[i]}}{\sum_{a=1}^m \lambda_a^{[i]}}, \dots, \frac{\lambda_1^{[i]}}{\sum_{a=1}^m \lambda_a^{[i]}}).$$

$$H_1 : (\lambda_1^{[1]}, \dots, \lambda_k^{[1]}) \neq (\lambda_1^{[2]}, \dots, \lambda_k^{[2]}), \quad H_2 : (\lambda_1^{[1]}, \dots, \lambda_k^{[1]}) = (\lambda_1^{[2]}, \dots, \lambda_k^{[2]}).$$

このとき、Kullback 情報量の直和分解は成立する。**Remark** この現象に対する 1 つの解釈 Poisson 分布の積 = 1 次元 Poisson × 多項分布 と分解される。よって、Pooling incomplete samples が「多項分布」の部分なら Kullback 情報量の直和分解が成り立つ。

Proposition 4.2. $m \leq k$ とする。 $X_j^{[i]}$ ($i = 1, 2, j = 1, 2, \dots, k$) および $X_j^{[i]}$ ($i = 1, 2, j = 1, 2, \dots, m$) は互いに独立。パラメータ $\lambda_x^{[i]}, \lambda_y^{[i]}, \theta_j^{[i]}$ ($i = 1, 2, j = 1, 2, \dots, k$) は、

$0 \leq \lambda_x^{[i]}, 0 \leq \lambda_y^{[i]}, 0 \leq \theta_j^{[i]} \leq 1, \sum_{j=1}^k \theta_j^{[i]} = 1$ を満たすとする。このとき、

$$X_j^{[i]} \sim \text{Poisson}(\lambda_x^{[i]} \theta_j^{[i]}), \quad j = 1, \dots, k,$$

$$Y_j^{[i]} \sim \text{Poisson}(\lambda_y^{[i]} \frac{\theta_j^{[i]}}{\sum_{l=1}^m \theta_l^{[i]}}), \quad j = 1, \dots, m.$$

ならば、Kullback 情報量の直和分解が成立する。

Remark $t^{[i]} = \sum_{a=1}^m \theta_a^{[i]}$, $u_j = \theta_j / \sum_{a=1}^m \theta_a$, ($j \leq m$), $v_j = \theta_j / \sum_{a=m+1}^k \theta_a$, ($j > m$) と変数変換することで表 4.1, 更に $\lambda^{[i]} = \lambda_x^{[i]} + \lambda_y^{[i]}$, $s^{[i]} = \lambda_x^{[i]} / (\lambda_x^{[i]} + \lambda_y^{[i]})$ とすることで表 4.2 のように確率構造を表記できる。

表 4.1

$X^{[i]}$	$X_1^{[i]}, \dots, X_m^{[i]}$	$X_{m+1}^{[i]}, \dots, X_k^{[i]}$
	$\lambda_x^{[i]} t^{[i]}$	$\lambda_x^{[i]} (1 - t^{[i]})$
	$u_1^{[i]}, \dots, u_m^{[i]}$	$v_{m+1}^{[i]}, \dots, v_k^{[i]}$
$Y^{[i]}$	$Y_1^{[i]}, \dots, Y_m^{[i]}$	
	$\lambda_y^{[i]}$	
	$u_1^{[i]}, \dots, u_m^{[i]}$	

表 4.2

$X^{[i]}$	$X_1^{[i]}, \dots, X_m^{[i]}$	$X_{m+1}^{[i]}, \dots, X_k^{[i]}$
	$\lambda^{[i]} s^{[i]} t^{[i]}$	$\lambda^{[i]} s^{[i]} (1 - t^{[i]})$
	$u_1^{[i]}, \dots, u_m^{[i]}$	$v_{m+1}^{[i]}, \dots, v_k^{[i]}$
$Y^{[i]}$	$Y_1^{[i]}, \dots, Y_m^{[i]}$	
	$\lambda^{[i]} (1 - s^{[i]})$	
	$u_1^{[i]}, \dots, u_m^{[i]}$	

今後の課題 本稿で扱った内容を、情報理論の立場、あるいは情報幾何の立場で説明できるように思える。今後の研究課題である。

References

- [1] Asano, C. (1965). On estimating multinomial probabilities by pooling incomplete samples. *Annals of the Institute of Statistical Mathematics* **17**, 1-13.
- [2] Johnson, N.L., Kotz, S. and Balakrishnan, N. (1997). *Discrete Multivariate Distributions*. Wiley.
- [3] Kullback, S. (1959). *Information Theory and Statistics*. Wiley.
- [4] Kullback, S. (1968). *Information Theory and Statistics*, Revised edition. Dover.
- [5] 布能英一郎 (2015). 負の多項分布における Kullback 情報量の直和分解—— Pooling incomplete samples の場合を含めた考察—— 京都大学数理解析研究所講究録 No 1954, 90-103.
- [6] 布能英一郎 (2016). 多変量離散型分布の 2 標本問題における Kullback 情報量の直和分解 京都大学数理解析研究所講究録 No 1999, 111-130.
- [7] 稲垣宣生 (2003). 数理統計学 (改訂版), 裳華房.