

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

関東学院大学 経済学部 布能 英一郎

Eiichiro Funo

School of Economics, Kanto Gakuin University

要旨 Kullback (1959 年初版, 1968 年に初版の誤記を修正して出版) は、推測統計の知見を Kullback 情報量によって記述し、更なる展開をも試みた箇所が随所に見られる。この Kullback の試みの成功例の 1 つとして、「多標本問題における分散分析の手法が Kullback 情報量によって記述でき、更に、この記述に従って離散データ解析に適用すると、分散分析と同様な結果を得る」ことが挙げられる。すなわち、Sum of Squares Total は Sum of Squares Within と Sum of Squares Between の和に等しいが、このことを Kullback 情報量の言葉で記述し、更にいくつかの離散確率モデルにおいて Total information が Within information と Between information の和に等しいことを示した。本稿は、多変量離散型分布の 2 標本問題において、Kullback の結果の拡張・発展を考察する。

1. Introduction

本稿では、多変量離散型分布における 2 標本問題を扱う。 $i = 1, 2$, に対して $\mathbf{X}^{[i]} = (X_1^{[i]}, \dots, X_k^{[i]})$ を未知母数 $\theta_1^{[i]}, \dots, \theta_m^{[i]}$, (m と k は、異なっても構わない) を含む多変量離散確率変数とする。 $\mathbf{X}^{[1]}$ と $\mathbf{X}^{[2]}$ は互いに独立と仮定する。仮説 H_1 , H_2 を、 H_2 : 母集団は同じ、 H_1 : 母集団は異なる、に選ぶ。すなわち、 $H_2: \theta_j^{[1]} = \theta_j^{[2]}$ for all $j = 1, \dots, m$, $H_1: \text{not } H_2$ とする。このとき、Kullback 情報量 $I(H_1, H_2)$ は

$$\begin{aligned} I(H_1, H_2) &= E_{H_1} \left(\log \frac{P(\mathbf{X}^{[1]}, \mathbf{X}^{[2]} | H_1)}{P(\mathbf{X}^{[1]}, \mathbf{X}^{[2]} | H_2)} \right) \\ &= E_{H_1} \left(\log \frac{P(\mathbf{X}^{[1]}, \mathbf{X}^{[2]} | \theta_1^{[1]}, \dots, \theta_m^{[1]}, \theta_1^{[2]}, \dots, \theta_m^{[2]})}{P(\mathbf{X}^{[1]}, \mathbf{X}^{[2]} | \theta_1, \dots, \theta_m)} \right) \end{aligned}$$

である。データ $\mathbf{x}^{[i]} = (x_1^{[i]}, \dots, x_k^{[i]})$, $i = 1, 2$, が観測されたとき、 $\theta_j^{[i]}$ の best estimator を $\hat{\theta}_j^{[i]}$ 、 θ_j の best estimator を $\hat{\theta}_j$ と表記する。Kullback (1968) によれば、Total information, Within information, Between information は、以下の式で定義されるものである。

$$\text{Total information} = \hat{E}_{H_1} \left(\log \frac{P(\mathbf{X}^{[1]}, \mathbf{X}^{[2]} | \hat{\theta}_1^{[1]}, \dots, \hat{\theta}_m^{[1]}, \hat{\theta}_1^{[2]}, \dots, \hat{\theta}_m^{[2]})}{P(\mathbf{X}^{[1]}, \mathbf{X}^{[2]} | \theta_1, \dots, \theta_m)} \right),$$

$$\text{Within information} = \hat{E}_{H_1} \left(\log \frac{P(\mathbf{X}^{[1]}, \mathbf{X}^{[2]} | \hat{\theta}_1^{[1]}, \dots, \hat{\theta}_m^{[1]}, \hat{\theta}_1^{[2]}, \dots, \hat{\theta}_m^{[2]})}{P(\mathbf{X}^{[1]}, \mathbf{X}^{[2]} | \hat{\theta}_1, \dots, \hat{\theta}_m)} \right),$$

$$\text{Between information} = \hat{E}_{H_2} \left(\log \frac{P(\mathbf{X}^{[1]}, \mathbf{X}^{[2]} | \hat{\theta}_1, \dots, \hat{\theta}_m)}{P(\mathbf{X}^{[1]}, \mathbf{X}^{[2]} | \theta_1, \dots, \theta_m)} \right).$$

$\mathbf{X}^{[i]} = (X_1^{[i]}, \dots, X_k^{[i]})$, $i = 1, 2$ が多項分布, *i.e.*, $\mathbf{X}^{[i]} \sim \text{multinomial}(N^{[i]}, (\theta_1^{[i]}, \dots, \theta_k^{[i]}))$ の場合

$$\frac{P(\mathbf{X}^{[1]}, \mathbf{X}^{[2]} | \theta_1^{[1]}, \dots, \theta_k^{[1]}, \theta_1^{[2]}, \dots, \theta_k^{[2]})}{P(\mathbf{X}^{[1]}, \mathbf{X}^{[2]} | \theta_1, \dots, \theta_k)} = \prod_{i=1}^2 \frac{\prod_{j=1}^k (\theta_j^{[i]})^{X_j^{[i]}}}{\prod_{j=1}^k \theta_j^{X_j^{[1]} + X_j^{[2]}}} = \prod_{i=1}^2 \prod_{j=1}^k \left(\frac{\theta_j^{[i]}}{\theta_j} \right)^{X_j^{[i]}},$$

$$I(H_1, H_2) = \sum_{i=1}^2 \sum_{j=1}^k E_{H_1}(X_j^{[i]}) \log \frac{\theta_j^{[i]}}{\theta_j} = \sum_{i=1}^2 \sum_{j=1}^k N^{[i]} \theta_j^{[i]} \log \frac{\theta_j^{[i]}}{\theta_j}$$

であり、データ $\mathbf{x}^{[i]} = (x_1^{[i]}, \dots, x_k^{[i]})$ を観測したとき、 $\hat{\theta}_j^{[i]} = x_j^{[i]}/N^{[i]}$, $\hat{\theta}_j = (x_j^{[1]} + x_j^{[2]})/(N^{[1]} + N^{[2]})$ であるから

$$\text{Total information} = \sum_{i=1}^2 \sum_{j=1}^k N^{[i]} \hat{\theta}_j^{[i]} \log \frac{\hat{\theta}_j^{[i]}}{\theta_j} = \sum_{i=1}^2 \sum_{j=1}^k x_j^{[i]} \log \frac{x_j^{[i]}}{N^{[i]} \theta_j}, \quad (1)$$

$$\text{Within information} = \sum_{i=1}^2 \sum_{j=1}^k N^{[i]} \hat{\theta}_j^{[i]} \log \frac{\hat{\theta}_j^{[i]}}{\hat{\theta}_j} = \sum_{i=1}^2 \sum_{j=1}^k x_j^{[i]} \log \frac{x_j^{[i]}/(x_j^{[1]} + x_j^{[2]})}{N^{[i]}/(N^{[1]} + N^{[2]})}, \quad (2)$$

$$\begin{aligned} \text{Between information} &= \hat{E}_{H_2} \left(\log \frac{P(\mathbf{X}^{[1]}, \mathbf{X}^{[2]} | \hat{\theta}_1, \dots, \hat{\theta}_k)}{P(\mathbf{X}^{[1]}, \mathbf{X}^{[2]} | \theta_1, \dots, \theta_k)} \right) \\ &= \sum_{j=1}^k \hat{E}_{H_2}(X_j^{[1]} + X_j^{[2]}) \log \frac{\hat{\theta}_j}{\theta_j} = \sum_{j=1}^k (N^{[1]} + N^{[2]}) \hat{\theta}_j \log \frac{\hat{\theta}_j}{\theta_j} \\ &= \sum_{j=1}^k (x_j^{[1]} + x_j^{[2]}) \log \frac{(x_j^{[1]} + x_j^{[2]})/(N^{[1]} + N^{[2]})}{\theta_j}. \end{aligned} \quad (3)$$

2. Kullback 情報量の直和分解が成立する場合

2 標本問題にて、Between information と Within information の和が Total information に等しいとき、「Kullback 情報量の直和分解が成立する」と略記する。

例 2.1 多項分布の 2 標本問題にて、Kullback 情報量の直和分解が成立する

実際、(1), (2), (3) および

$$\frac{x_j^{[i]}}{N^{[i]} \theta_j} = \frac{x_j^{[i]}/(x_j^{[1]} + x_j^{[2]})}{N^{[i]}/(N^{[1]} + N^{[2]})} \frac{(x_j^{[1]} + x_j^{[2]})/(N^{[1]} + N^{[2]})}{\theta_j}$$

すなわち

$$\log \frac{x_j^{[i]}}{N^{[i]}\theta_j} = \log \frac{x_j^{[i]}/(x_j^{[1]} + x_j^{[2]})}{N^{[i]}/(N^{[1]} + N^{[2]})} + \log \frac{(x_j^{[1]} + x_j^{[2]})/(N^{[1]} + N^{[2]})}{\theta_j}$$

により、容易に示せる。

例 2.2 負の多項分布の 2 標本問題にて、Kullback 情報量の直和分解が成立する。

例 2.3 Pooling incomplete samples を伴う多項分布の 2 標本問題にて、Kullback 情報量の直和分解が成立する。すなわち、 $\mathbf{X}^{[1]}, \mathbf{X}^{[2]}, \mathbf{Y}^{[1]}, \mathbf{Y}^{[2]}$ は互いに独立で、各 $i = 1, 2$ に対して

$$\begin{aligned} \mathbf{X}^{[i]} &= (X_1^{[i]}, \dots, X_m^{[i]}, \dots, X_k^{[i]}) \sim \text{Multinomial}(N_1^{[i]}; p_1^{[i]}, \dots, p_m^{[i]}, \dots, p_k^{[i]}), \\ \mathbf{Y}^{[i]} &= (Y_1^{[i]}, \dots, Y_m^{[i]}) \sim \text{Multinomial}(N_2^{[i]}; \frac{p_1^{[i]}}{\sum_{j=1}^m p_j^{[i]}}, \dots, \frac{p_m^{[i]}}{\sum_{j=1}^m p_j^{[i]}}) \end{aligned}$$

のとき、Kullback 情報量の直和分解が成立する。

例 2.2, 例 2.3 は、布能 (2015) で扱った。なお、例 2.3 については、本稿第 4 章 例 4.1 で簡潔に示す方法を提示する。

3. Kullback 情報量の直和分解が成立しない場合と、その対処

2 標本問題において、Kullback 情報量の直和分解は、常に成り立つとは限らない。というよりも、少し複雑な確率モデルで成立しないどころか、次のようなごく簡単な確率モデルにおいても成り立たない。

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

$$\left. \begin{aligned} \mathbf{X}^{[i]} &= (X_1^{[i]}, \dots, X_m^{[i]}, X_{m+1}^{[i]}, \dots, X_k^{[i]}) \\ &\sim \text{Multinomial}(N_x^{[i]}; \theta_1^{[i]}, \dots, \theta_m^{[i]}, \theta_{m+1}^{[i]}, \dots, \theta_k^{[i]}) \\ \mathbf{Y}^{[i]} &= (Y_1^{[i]}, \dots, Y_m^{[i]}, Y_{m+1}^{[i]}, \dots, Y_k^{[i]}) \\ &\sim \text{Multinomial}(N_y^{[i]}; \theta_1^{[i]}, \dots, \theta_m^{[i]}, \xi_{m+1}^{[i]}, \dots, \xi_k^{[i]}) \end{aligned} \right\} \quad (4)$$

とする。仮説 H_1, H_2 を

$$\left. \begin{aligned} H_2: & \quad \theta_j^{[1]} = \theta_j^{[2]} = \theta_j \quad \text{for } j = 1, \dots, k, \\ & \quad \xi_j^{[1]} = \xi_j^{[2]} = \xi_j \quad \text{for } j = m+1, \dots, k, \\ H_1: & \quad \text{not } H_2. \end{aligned} \right\} \quad (5)$$

に選ぶ。このとき、Total information $\neq$ Within information + Between information.

H_2 : 同じ母集団からの標本、 H_1 : 異なった母集団からの標本、というフレームワークにおいて、(5) で設定した仮説は、実に自然なものである。ところが、Kullback 情報

量の直和分解は成り立たなかった。これは、我々の直感に反するものである。なぜだろうか？ これに対する回答 (の1つ) として、次のようなことが見い出せた。

Proposition 1. (4) のモデルの下で、仮説を

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

に選ぶと、Kullback 情報量の直和分解が成立する。

例 3.1 および Proposition 1 は、次のようにして示せる：パラメータ変換

$$s^{[i]} = \sum_{l=1}^m \theta_l^{[i]}, \quad u_j^{[i]} = \frac{\theta_j^{[i]}}{\sum_{l=1}^m \theta_l^{[i]}}, \quad v_j^{[i]} = \frac{\theta_j^{[i]}}{\sum_{l=m+1}^k \theta_l^{[i]}}, \quad w_j^{[i]} = \frac{\xi_j^{[i]}}{\sum_{l=m+1}^k \xi_l^{[i]}}$$

すなわち

$$\theta_j^{[i]} = \begin{cases} s^{[i]} u_j^{[i]} & j = 1, \dots, m, \\ (1 - s^{[i]}) v_j^{[i]} & j = m+1, \dots, k, \end{cases}$$

$$\xi_j^{[i]} = (1 - s^{[i]}) w_j^{[i]} \quad j = m+1, \dots, k,$$

を用いることで、確率モデル (4) は

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

と書くことができ、仮説 (5) は

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

と書き表せる。このとき、各パラメータの best estimator は

$$\hat{s}^{[i]} = \frac{\sum_{l=1}^m (x_l^{[i]} + y_l^{[i]})}{N_x^{[i]} + N_y^{[i]}}, \quad \hat{u}_j^{[i]} = \frac{x_j^{[i]} + y_j^{[i]}}{\sum_{l=1}^m (x_l^{[i]} + y_l^{[i]})}, \quad \hat{v}_j^{[i]} = \frac{x_j^{[i]}}{\sum_{l=m+1}^k x_l^{[i]}}, \quad \hat{w}_j^{[i]} = \frac{y_j^{[i]}}{\sum_{l=m+1}^k y_l^{[i]}}$$

であり、

$$\begin{aligned} & \frac{P(\mathbf{x}^{[1]}, \mathbf{y}^{[1]}, \mathbf{x}^{[2]}, \mathbf{y}^{[2]} \mid H_1)}{P(\mathbf{x}^{[1]}, \mathbf{y}^{[1]}, \mathbf{x}^{[2]}, \mathbf{y}^{[2]} \mid H_2)} \\ &= \prod_{i=1}^2 \frac{\prod_{j=1}^m (s^{[i]} u_j^{[i]})^{x_j^{[i]} + y_j^{[i]}} \prod_{j=m+1}^k ((1 - s^{[i]}) v_j^{[i]})^{x_j^{[i]}} \prod_{j=m+1}^k ((1 - s^{[i]}) w_j^{[i]})^{y_j^{[i]}}}{\prod_{j=1}^m (s u_j)^{x_j^{[i]} + y_j^{[i]}} \prod_{j=m+1}^k ((1 - s) v_j)^{x_j^{[i]}} \prod_{j=m+1}^k ((1 - s) w_j)^{y_j^{[i]}}} \end{aligned}$$

$$= \prod_{i=1}^2 \left\{ \left(\frac{s^{[i]}}{s} \right)^{\sum_{j=1}^m (x_j^{[i]} + y_j^{[i]})} \left(\frac{1 - s^{[i]}}{1 - s} \right)^{\sum_{j=m+1}^k (x_j^{[i]} + y_j^{[i]})} \right. \\ \left. \times \prod_{j=1}^m \left(\frac{u_j^{[i]}}{u_j} \right)^{x_j^{[i]} + y_j^{[i]}} \prod_{j=m+1}^k \left(\frac{v_j^{[i]}}{v_j} \right)^{x_j^{[i]}} \prod_{j=m+1}^k \left(\frac{w_j^{[i]}}{w_j} \right)^{y_j^{[i]}} \right\}.$$

$T_m^{[i]}(x) = \sum_{l=1}^m x_l^{[i]}$, $T_m^{[i]}(y) = \sum_{l=1}^m y_l^{[i]}$ という記号を用いると

$$\log \frac{P(\mathbf{x}^{[1]}, \mathbf{y}^{[1]}, \mathbf{x}^{[2]}, \mathbf{y}^{[2]} | H_1)}{P(\mathbf{x}^{[1]}, \mathbf{y}^{[1]}, \mathbf{x}^{[2]}, \mathbf{y}^{[2]} | H_2)} \\ = \sum_{i=1}^2 \left\{ (T_m^{[i]}(x) + T_m^{[i]}(y)) \log \frac{s^{[i]}}{s} + (N_x^{[i]} + N_y^{[i]} - T_m^{[i]}(x) - T_m^{[i]}(y)) \log \frac{1 - s^{[i]}}{1 - s} \right. \\ \left. + \sum_{j=1}^m (x_j^{[i]} + y_j^{[i]}) \log \frac{u_j^{[i]}}{u_j} + \sum_{j=m+1}^k x_j^{[i]} \log \frac{v_j^{[i]}}{v_j} + \sum_{j=m+1}^k y_j^{[i]} \log \frac{w_j^{[i]}}{w_j} \right\}$$

である。そして、 $E_{H_1}(X_j^{[i]}) = N_x^{[i]} s^{[i]} u_j^{[i]}$, ($j \leq m$), $E_{H_1}(X_j^{[i]}) = N_x^{[i]} (1 - s^{[i]}) v_j^{[i]}$, ($j > m$), $E_{H_1}(Y_j^{[i]}) = N_y^{[i]} s^{[i]} u_j^{[i]}$, ($j \leq m$), $E_{H_1}(Y_j^{[i]}) = N_y^{[i]} (1 - s^{[i]}) w_j^{[i]}$, ($j > m$), $E_{H_1}(T_m^{[i]}(X)) = N_x^{[i]} s^{[i]}$, $E_{H_1}(T_m^{[i]}(Y)) = N_y^{[i]} s^{[i]}$ を用いて

$$E_{H_1} \left(\log \frac{P(\mathbf{X}^{[1]}, \mathbf{Y}^{[1]}, \mathbf{X}^{[2]}, \mathbf{Y}^{[2]} | H_1)}{P(\mathbf{X}^{[1]}, \mathbf{Y}^{[1]}, \mathbf{X}^{[2]}, \mathbf{Y}^{[2]} | H_2)} \right) \\ = \sum_{i=1}^2 \left\{ (N_x^{[i]} + N_y^{[i]}) s^{[i]} \log \frac{s^{[i]}}{s} + (N_x^{[i]} + N_y^{[i]}) (1 - s^{[i]}) \log \frac{1 - s^{[i]}}{1 - s} \right. \\ \left. + \sum_{j=1}^m (N_x^{[i]} + N_y^{[i]}) s^{[i]} u_j^{[i]} \log \frac{u_j^{[i]}}{u_j} + \sum_{j=m+1}^k N_x^{[i]} (1 - s^{[i]}) v_j^{[i]} \log \frac{v_j^{[i]}}{v_j} \right. \\ \left. + \sum_{j=m+1}^k N_y^{[i]} (1 - s^{[i]}) w_j^{[i]} \log \frac{w_j^{[i]}}{w_j} \right\},$$

Total information

$$= \sum_{i=1}^2 \left\{ (T_m^{[i]}(x) + T_m^{[i]}(y)) \log \frac{\hat{s}^{[i]}}{s} + (N_x^{[i]} + N_y^{[i]} - T_m^{[i]}(x) - T_m^{[i]}(y)) \log \frac{1 - \hat{s}^{[i]}}{1 - s} \right. \\ \left. + \sum_{j=1}^m (x_j^{[i]} + y_j^{[i]}) \log \frac{\hat{u}_j^{[i]}}{u_j} + \frac{N_x^{[i]}}{N_x^{[i]} + N_y^{[i]}} \frac{\sum_{l=m+1}^k (x_l^{[i]} + y_l^{[i]})}{\sum_{l=m+1}^k x_l^{[i]}} \sum_{j=m+1}^k x_j^{[i]} \log \frac{\hat{v}_j^{[i]}}{v_j} \right. \\ \left. + \frac{N_y^{[i]}}{N_x^{[i]} + N_y^{[i]}} \frac{\sum_{l=m+1}^k (x_l^{[i]} + y_l^{[i]})}{\sum_{l=m+1}^k y_l^{[i]}} \sum_{j=m+1}^k y_j^{[i]} \log \frac{\hat{w}_j^{[i]}}{w_j} \right\}, \quad (9)$$

Within information

$$= \sum_{i=1}^2 \left\{ (T_m^{[i]}(x) + T_m^{[i]}(y)) \log \frac{\hat{s}^{[i]}}{\hat{s}} + (N_x^{[i]} + N_y^{[i]} - T_m^{[i]}(x) - T_m^{[i]}(y)) \log \frac{1 - \hat{s}^{[i]}}{1 - \hat{s}} \right.$$

$$\begin{aligned}
& + \sum_{j=1}^m (x_j^{[i]} + y_j^{[i]}) \log \frac{\hat{u}_j^{[i]}}{\hat{u}_j} + \frac{N_x^{[i]}}{N_x^{[i]} + N_y^{[i]}} \frac{\sum_{l=m+1}^k (x_l^{[i]} + y_l^{[i]})}{\sum_{l=m+1}^k x_l^{[i]}} \sum_{j=m+1}^k x_j^{[i]} \log \frac{\hat{v}_j^{[i]}}{\hat{v}_j} \\
& + \frac{N_y^{[i]}}{N_x^{[i]} + N_y^{[i]}} \frac{\sum_{l=m+1}^k (x_l^{[i]} + y_l^{[i]})}{\sum_{l=m+1}^k y_l^{[i]}} \sum_{j=m+1}^k y_j^{[i]} \log \frac{\hat{w}_j^{[i]}}{\hat{w}_j} \Big\} \quad (10)
\end{aligned}$$

を得る。同様に、

$$\begin{aligned}
\text{Between information} &= \hat{E}_{H_2} \left(\log \frac{\hat{P}(\mathbf{X}^{[1]}, \mathbf{Y}^{[1]}, \mathbf{X}^{[2]}, \mathbf{Y}^{[2]} | H_2)}{P(\mathbf{X}^{[1]}, \mathbf{Y}^{[1]}, \mathbf{X}^{[2]}, \mathbf{Y}^{[2]} | H_2)} \right) \\
&= (T_m^{[1]}(x) + T_m^{[2]}(x) + T_m^{[1]}(y) + T_m^{[2]}(y)) \log \frac{\hat{s}}{s} \\
&+ \sum_{l=m+1}^k (x_l^{[1]} + x_l^{[2]} + y_l^{[1]} + y_l^{[2]}) \log \frac{1 - \hat{s}}{1 - s} + \sum_{j=1}^m (x_j^{[1]} + x_j^{[2]} + y_j^{[1]} + y_j^{[2]}) \log \frac{\hat{u}_j}{u_j} \\
&+ \frac{N_x^{[1]} + N_x^{[2]}}{N_x^{[1]} + N_x^{[2]} + N_y^{[1]} + N_y^{[2]}} \frac{\sum_{l=m+1}^k (x_l^{[1]} + x_l^{[2]} + y_l^{[1]} + y_l^{[2]})}{\sum_{l=m+1}^k (x_l^{[1]} + x_l^{[2]})} \sum_{j=m+1}^k (x_j^{[1]} + x_j^{[2]}) \log \frac{\hat{v}_j}{v_j} \\
&+ \frac{N_y^{[1]} + N_y^{[2]}}{N_x^{[1]} + N_x^{[2]} + N_y^{[1]} + N_y^{[2]}} \frac{\sum_{l=m+1}^k (x_l^{[1]} + x_l^{[2]} + y_l^{[1]} + y_l^{[2]})}{\sum_{l=m+1}^k (y_l^{[1]} + y_l^{[2]})} \sum_{j=m+1}^k (y_j^{[1]} + y_j^{[2]}) \log \frac{\hat{w}_j}{w_j} \quad (11)
\end{aligned}$$

が得られる。(10) と (11) の和は (9) に一致しない。

他方、仮説 (6) は

$$H_2 : u_j^{[1]} = u_j^{[2]} = u_j, \quad s^{[1]} = s^{[2]} = s, \quad H_1 : \text{not } H_2 \quad (12)$$

と書き直せる。このとき

$$\begin{aligned}
& \frac{P(\mathbf{x}^{[1]}, \mathbf{y}^{[1]}, \mathbf{x}^{[2]}, \mathbf{y}^{[2]} | H_1)}{P(\mathbf{x}^{[1]}, \mathbf{y}^{[1]}, \mathbf{x}^{[2]}, \mathbf{y}^{[2]} | H_2)} \\
&= \prod_{i=1}^2 \frac{\prod_{j=1}^m (s^{[i]} u_j^{[i]})^{x_j^{[i]} + y_j^{[i]}} \prod_{j=m+1}^k ((1 - s^{[i]}) v_j^{[i]})^{x_j^{[i]}} \prod_{j=m+1}^k ((1 - s^{[i]}) w_j^{[i]})^{y_j^{[i]}}}{\prod_{j=1}^m (s u_j)^{x_j^{[i]} + y_j^{[i]}} \prod_{j=m+1}^k ((1 - s) v_j^{[i]})^{x_j^{[i]}} \prod_{j=m+1}^k ((1 - s) w_j^{[i]})^{y_j^{[i]}}} \\
&= \prod_{i=1}^2 \left\{ \left(\frac{s^{[i]}}{s} \right)^{\sum_{j=1}^m (x_j^{[i]} + y_j^{[i]})} \left(\frac{1 - s^{[i]}}{1 - s} \right)^{\sum_{j=m+1}^k (x_j^{[i]} + y_j^{[i]})} \prod_{j=1}^m \left(\frac{u_j^{[i]}}{u_j} \right)^{x_j^{[i]} + y_j^{[i]}} \right\}
\end{aligned}$$

すなわち

$$\begin{aligned}
\log \frac{P(\mathbf{x}^{[1]}, \mathbf{y}^{[1]}, \mathbf{x}^{[2]}, \mathbf{y}^{[2]} | H_1)}{P(\mathbf{x}^{[1]}, \mathbf{y}^{[1]}, \mathbf{x}^{[2]}, \mathbf{y}^{[2]} | H_2)} &= \sum_{i=1}^2 \left\{ (T_m^{[i]}(x) + T_m^{[i]}(y)) \log \frac{s^{[i]}}{s} \right. \\
&\quad \left. + \sum_{j=m+1}^k (x_j^{[i]} + y_j^{[i]}) \log \frac{1 - s^{[i]}}{1 - s} + \sum_{j=1}^m (x_j^{[i]} + y_j^{[i]}) \log \frac{u_j^{[i]}}{u_j} \right\}
\end{aligned}$$

により

$$\begin{aligned} \text{Total information} = & \sum_{i=1}^2 \left\{ (T_m^{[i]}(x) + T_m^{[i]}(y)) \log \frac{\hat{s}^{[i]}}{s} \right. \\ & \left. + \sum_{l=m+1}^k (x_l^{[1]} + x_l^{[2]} + y_l^{[1]} + y_l^{[2]}) \log \frac{1 - \hat{s}^{[i]}}{1 - s} + \sum_{j=1}^m (x_j^{[i]} + y_j^{[i]}) \log \frac{\hat{u}_j^{[i]}}{u_j} \right\}, \end{aligned} \quad (13)$$

$$\begin{aligned} \text{Within information} = & \sum_{i=1}^2 \left\{ (T_m^{[i]}(x) + T_m^{[i]}(y)) \log \frac{\hat{s}^{[i]}}{\hat{s}} \right. \\ & \left. + \sum_{l=m+1}^k (x_l^{[1]} + x_l^{[2]} + y_l^{[1]} + y_l^{[2]}) \log \frac{1 - \hat{s}^{[i]}}{1 - \hat{s}} + \sum_{j=1}^m (x_j^{[i]} + y_j^{[i]}) \log \frac{\hat{u}_j^{[i]}}{\hat{u}_j} \right\} \end{aligned} \quad (14)$$

を得る。同様に

$$\begin{aligned} \text{Between information} = & \hat{E}_{H_2} \left(\log \frac{\hat{P}(X^{[1]}, Y^{[1]}, X^{[2]}, Y^{[2]} | H_2)}{P(X^{[1]}, Y^{[1]}, X^{[2]}, Y^{[2]} | H_2)} \right) \\ = & (T_m^{[1]}(x) + T_m^{[2]}(x) + T_m^{[1]}(y) + T_m^{[2]}(y)) \log \frac{\hat{s}}{s} \\ & + \sum_{l=m+1}^k (x_l^{[1]} + x_l^{[2]} + y_l^{[1]} + y_l^{[2]}) \log \frac{1 - \hat{s}}{1 - s} + \sum_{j=1}^m (x_j^{[1]} + x_j^{[2]} + y_j^{[1]} + y_j^{[2]}) \log \frac{\hat{u}_j}{u_j} \end{aligned} \quad (15)$$

が得られる。つまり、仮説 (12) の下では、(14) と (15) の和が (13) に等しい。

2 標本問題において、仮説 H_2 : 同じ母集団からの標本、の設定を「注目しているパラメータが2標本で同じであればよい」と設定することに異論もあろう。そこで、確率モデルの設定条件を緩和することで、「すべてのパラメータが同じ」という仮説 H_2 とその対立仮説 H_1 (H_1 は H_2 の否定形) の下で Kullback 情報量の直和分解が成り立つようにすることを考察したところ、次の Proposition を得ることができた。

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

$$\begin{aligned} X^{[i]} &= (X_1^{[i]}, \dots, X_m^{[i]}, X_{m+1}^{[i]}, \dots, X_k^{[i]}) \\ &\sim \text{Multinomial}(N_x^{[i]}; s^{[i]}u_1^{[i]}, \dots, s^{[i]}u_m^{[i]}, (1 - s^{[i]})v_{m+1}^{[i]}, \dots, (1 - s^{[i]})v_k^{[i]}), \\ Y^{[i]} &= (Y_1^{[i]}, \dots, Y_m^{[i]}, Y_{m+1}^{[i]}, \dots, Y_k^{[i]}) \\ &\sim \text{Multinomial}(N_y^{[i]}; t^{[i]}u_1^{[i]}, \dots, t^{[i]}u_m^{[i]}, (1 - t^{[i]})w_{m+1}^{[i]}, \dots, (1 - t^{[i]})w_k^{[i]}) \end{aligned}$$

とする。仮説 H_1, H_2 を

$$\begin{aligned} H_2 : & u_j^{[1]} = u_j^{[2]} = u_j, v_j^{[1]} = v_j^{[2]} = v_j, w_j^{[1]} = w_j^{[2]} = w_j, \\ & s^{[1]} = s^{[2]} = s, t^{[1]} = t^{[2]} = t, \end{aligned}$$

$$H_1 : \text{not } H_2$$

に選ぶ。このとき、Kullback 情報量の直和分解が成立する。

proof $\mathbf{x}^{[1]}, \mathbf{y}^{[1]}, \mathbf{x}^{[2]}, \mathbf{y}^{[2]}$ を観測したとき、 H_1 と H_2 の尤度比は

$$\begin{aligned} & \frac{P(\mathbf{x}^{[1]}, \mathbf{y}^{[1]}, \mathbf{x}^{[2]}, \mathbf{y}^{[2]} \mid H_1)}{P(\mathbf{x}^{[1]}, \mathbf{y}^{[1]}, \mathbf{x}^{[2]}, \mathbf{y}^{[2]} \mid H_2)} \\ &= \prod_{i=1}^2 \left\{ \left(\frac{s^{[i]}}{s} \right)^{\sum_{j=1}^m x_j^{[i]}} \left(\frac{1-s^{[i]}}{1-s} \right)^{\sum_{j=m+1}^k x_j^{[i]}} \left(\frac{t^{[i]}}{t} \right)^{\sum_{j=1}^m y_j^{[i]}} \left(\frac{1-t^{[i]}}{1-t} \right)^{\sum_{j=m+1}^k y_j^{[i]}} \right. \\ & \quad \left. \times \prod_{j=1}^m \left(\frac{u_j^{[i]}}{u_j} \right)^{x_j^{[i]} + y_j^{[i]}} \prod_{j=m+1}^k \left(\frac{v_j^{[i]}}{v_j} \right)^{x_j^{[i]}} \prod_{j=m+1}^k \left(\frac{w_j^{[i]}}{w_j} \right)^{y_j^{[i]}} \right\} \end{aligned}$$

である。これより

$$\begin{aligned} & \log \frac{P(\mathbf{x}^{[1]}, \mathbf{y}^{[1]}, \mathbf{x}^{[2]}, \mathbf{y}^{[2]} \mid H_1)}{P(\mathbf{x}^{[1]}, \mathbf{y}^{[1]}, \mathbf{x}^{[2]}, \mathbf{y}^{[2]} \mid H_2)} \\ &= \sum_{i=1}^2 \left\{ T_m^{[i]}(x) \log \frac{s^{[i]}}{s} + \sum_{j=m+1}^k x_j^{[i]} \log \frac{1-s^{[i]}}{1-s} + T_m^{[i]}(y) \log \frac{t^{[i]}}{t} + \sum_{j=m+1}^k y_j^{[i]} \log \frac{1-t^{[i]}}{1-t} \right. \\ & \quad \left. + \sum_{j=1}^m (x_j^{[i]} + y_j^{[i]}) \log \frac{u_j^{[i]}}{u_j} + \sum_{j=m+1}^k x_j^{[i]} \log \frac{v_j^{[i]}}{v_j} + \sum_{j=m+1}^k y_j^{[i]} \log \frac{w_j^{[i]}}{w_j} \right\} \end{aligned}$$

であり、 $j = 1, \dots, m$ に対して

$$E_{H_1}(X_j^{[i]}) = N_x^{[i]} s^{[i]} u_j^{[i]}, \quad E_{H_1}(Y_j^{[i]}) = N_y^{[i]} t^{[i]} u_j^{[i]},$$

$j = m+1, \dots, k$ に対して

$$E_{H_1}(X_j^{[i]}) = N_x^{[i]} (1-s^{[i]}) v_j^{[i]}, \quad E_{H_1}(Y_j^{[i]}) = N_y^{[i]} (1-t^{[i]}) w_j^{[i]}$$

であることから

$$\begin{aligned} & E_{H_1} \left(\log \frac{P(\mathbf{X}^{[1]}, \mathbf{Y}^{[1]}, \mathbf{X}^{[2]}, \mathbf{Y}^{[2]} \mid H_1)}{P(\mathbf{X}^{[1]}, \mathbf{Y}^{[1]}, \mathbf{X}^{[2]}, \mathbf{Y}^{[2]} \mid H_2)} \right) \\ &= \sum_{i=1}^2 \left\{ N_x^{[i]} s^{[i]} \log \frac{s^{[i]}}{s} + N_x^{[i]} (1-s^{[i]}) \log \frac{1-s^{[i]}}{1-s} + N_y^{[i]} t^{[i]} \log \frac{t^{[i]}}{t} \right. \\ & \quad \left. + N_y^{[i]} (1-t^{[i]}) \log \frac{1-t^{[i]}}{1-t} + \sum_{j=1}^m (N_x^{[i]} s^{[i]} + N_y^{[i]} t^{[i]}) u_j^{[i]} \log \frac{u_j^{[i]}}{u_j} \right. \\ & \quad \left. + \sum_{j=m+1}^k N_x^{[i]} (1-s^{[i]}) v_j^{[i]} \log \frac{v_j^{[i]}}{v_j} + \sum_{j=m+1}^k N_y^{[i]} (1-t^{[i]}) w_j^{[i]} \log \frac{w_j^{[i]}}{w_j} \right\}. \end{aligned}$$

そして、 H_1 の下での best estimator が

$$\begin{aligned} \hat{s}^{[i]} &= \frac{T_m^{[i]}(x)}{N_x^{[i]}}, \quad \hat{t}^{[i]} = \frac{T_m^{[i]}(y)}{N_y^{[i]}}, \quad \hat{u}_j^{[i]} = \frac{x_j^{[i]} + y_j^{[i]}}{T_m^{[i]}(x) + T_m^{[i]}(y)}, \\ \hat{v}_j^{[i]} &= \frac{x_j^{[i]}}{N_x^{[i]} - T_m^{[i]}(x)} = \frac{x_j^{[i]}}{\sum_{l=m+1}^k x_l^{[i]}}, \quad \hat{w}_j^{[i]} = \frac{y_j^{[i]}}{N_y^{[i]} - T_m^{[i]}(y)} = \frac{y_j^{[i]}}{\sum_{l=m+1}^k y_l^{[i]}} \end{aligned}$$

であることを用いると、

Total information

$$\begin{aligned}
 &= \sum_{i=1}^2 \left\{ T_m^{[i]}(x) \log \frac{\hat{s}^{[i]}}{s} + (N_x^{[i]} - T_m^{[i]}(x)) \log \frac{1 - \hat{s}^{[i]}}{1 - s} + T_m^{[i]}(y) \log \frac{\hat{t}^{[i]}}{t} \right. \\
 &\quad + (N_y^{[i]} - T_m^{[i]}(y)) \log \frac{1 - \hat{t}^{[i]}}{1 - t} + \sum_{j=1}^m (x_j^{[i]} + y_j^{[i]}) \log \frac{\hat{u}_j^{[i]}}{u_j} + \sum_{j=m+1}^k x_j^{[i]} \log \frac{\hat{v}_j^{[i]}}{v_j} \\
 &\quad \left. + \sum_{j=m+1}^k y_j^{[i]} \log \frac{\hat{w}_j^{[i]}}{w_j} \right\}, \tag{16}
 \end{aligned}$$

Within information

$$\begin{aligned}
 &= \sum_{i=1}^2 \left\{ T_m^{[i]}(x) \log \frac{\hat{s}^{[i]}}{\hat{s}} + (N_x^{[i]} - T_m^{[i]}(x)) \log \frac{1 - \hat{s}^{[i]}}{1 - \hat{s}} + T_m^{[i]}(y) \log \frac{\hat{t}^{[i]}}{\hat{t}} \right. \\
 &\quad + (N_y^{[i]} - T_m^{[i]}(y)) \log \frac{1 - \hat{t}^{[i]}}{1 - \hat{t}} + \sum_{j=1}^m (x_j^{[i]} + y_j^{[i]}) \log \frac{\hat{u}_j^{[i]}}{\hat{u}_j} + \sum_{j=m+1}^k x_j^{[i]} \log \frac{\hat{v}_j^{[i]}}{\hat{v}_j} \\
 &\quad \left. + \sum_{j=m+1}^k y_j^{[i]} \log \frac{\hat{w}_j^{[i]}}{\hat{w}_j} \right\}. \tag{17}
 \end{aligned}$$

同様に

Between information

$$\begin{aligned}
 &= (T_m^{[1]}(x) + T_m^{[2]}(x)) \log \frac{\hat{s}}{s} + (N_x^{[1]} + N_x^{[2]} - T_m^{[1]}(x) - T_m^{[2]}(x)) \log \frac{1 - \hat{s}}{1 - s} \\
 &\quad + (T_m^{[1]}(y) + T_m^{[2]}(y)) \log \frac{\hat{t}}{t} + (N_y^{[1]} + N_y^{[2]} - T_m^{[1]}(y) - T_m^{[2]}(y)) \log \frac{1 - \hat{t}}{1 - t} \\
 &\quad + \sum_{j=1}^m (x_j^{[1]} + x_j^{[2]} + y_j^{[1]} + y_j^{[2]}) \log \frac{\hat{u}_j}{u_j} \\
 &\quad + \sum_{j=m+1}^k (x_j^{[1]} + x_j^{[2]}) \log \frac{\hat{v}_j}{v_j} + \sum_{j=m+1}^k (y_j^{[1]} + y_j^{[2]}) \log \frac{\hat{w}_j}{w_j} \tag{18}
 \end{aligned}$$

が得られる。あとは単純計算で (17) と (18) の和が (16) であることが示せる。

4. パラメータ変換のメリット

前節では、パラメータ変換を上手に用いて Kullback 情報量の直和分解を示したが、パラメータ変換の有用性について更に調べてみた。

例 4.1 例 2.3 (Pooling incomplete samples の場合の 2 標本問題) は、パラメータ変換を行って Kullback 情報量を計算すると、以下のように簡潔に示せる：パラメータ変換

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

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

$$\text{と書き下せる。} \quad \hat{s}^{[i]} = \frac{T_m^{[i]}(x)}{N_x^{[i]}}, \quad \hat{u}_j^{[i]} = \frac{x_j^{[i]} + y_j^{[i]}}{T_m^{[i]}(x) + N_y^{[i]}}, \quad \hat{v}_j^{[i]} = \frac{x_j^{[i]}}{\sum_{l=m+1}^k x_l^{[i]}} \quad \text{であり}$$

$$\begin{aligned} \frac{P(\mathbf{x}^{[1]}, \mathbf{y}^{[1]}, \mathbf{x}^{[2]}, \mathbf{y}^{[2]} \mid H_1)}{P(\mathbf{x}^{[1]}, \mathbf{y}^{[1]}, \mathbf{x}^{[2]}, \mathbf{y}^{[2]} \mid H_2)} &= \prod_{i=1}^2 \frac{\prod_{j=1}^m (s^{[i]}u_j^{[i]})^{x_j^{[i]}} \prod_{j=m+1}^k ((1-s^{[i]})v_j^{[i]})^{x_j^{[i]}} \prod_{j=1}^m (u_j^{[i]})^{y_j^{[i]}}}{\prod_{j=1}^m (su_j)^{x_j^{[i]}} \prod_{j=m+1}^k ((1-s)v_j)^{x_j^{[i]}} \prod_{j=1}^m (u_j)^{y_j^{[i]}}} \\ &= \prod_{i=1}^2 \left\{ \left(\frac{s^{[i]}}{s} \right)^{T_m^{[i]}(x)} \left(\frac{1-s^{[i]}}{1-s} \right)^{N_x^{[i]} - T_m^{[i]}(x)} \prod_{j=1}^m \left(\frac{u_j^{[i]}}{u_j} \right)^{x_j^{[i]} + y_j^{[i]}} \right\}, \end{aligned}$$

$$\begin{aligned} E_{H_1} \left(\log \frac{P(\mathbf{X}^{[1]}, \mathbf{Y}^{[1]}, \mathbf{X}^{[2]}, \mathbf{Y}^{[2]} \mid H_1)}{P(\mathbf{X}^{[1]}, \mathbf{Y}^{[1]}, \mathbf{X}^{[2]}, \mathbf{Y}^{[2]} \mid H_2)} \right) &= \sum_{i=1}^2 \left\{ E_{H_1} T_m^{[i]}(X) \log \frac{s^{[i]}}{s} \right. \\ &\quad \left. + E_{H_1} (N_x^{[i]} - T_m^{[i]}(X)) \log \frac{1-s^{[i]}}{1-s} + \sum_{j=1}^m E_{H_1} (X_j^{[i]} + Y_j^{[i]}) \log \frac{u_j^{[i]}}{u_j} \right\} \\ &= \sum_{i=1}^2 \left\{ N_x^{[i]} s^{[i]} \log \frac{s^{[i]}}{s} + N_x^{[i]} (1-s^{[i]}) \log \frac{1-s^{[i]}}{1-s} + \sum_{j=1}^m (N_x^{[i]} s^{[i]} + N_y^{[i]}) u_j^{[i]} \log \frac{u_j^{[i]}}{u_j} \right\} \end{aligned}$$

であるから、これらを用いて

$$\begin{aligned} \text{Total information} &= \hat{E}_{H_1} \left(\log \frac{\hat{P}(\mathbf{X}^{[1]}, \mathbf{Y}^{[1]}, \mathbf{X}^{[2]}, \mathbf{Y}^{[2]} \mid H_1)}{P(\mathbf{X}^{[1]}, \mathbf{Y}^{[1]}, \mathbf{X}^{[2]}, \mathbf{Y}^{[2]} \mid H_2)} \right) \\ &= \sum_{i=1}^2 \left\{ T_m^{[i]}(x) \log \frac{\hat{s}^{[i]}}{s} + \sum_{j=m+1}^k x_j^{[i]} \log \frac{1-\hat{s}^{[i]}}{1-s} + \sum_{j=1}^m (x_j^{[i]} + y_j^{[i]}) \log \frac{\hat{u}_j^{[i]}}{u_j} \right\}, \end{aligned} \quad (19)$$

Within information

$$= \sum_{i=1}^2 \left\{ T_m^{[i]}(x) \log \frac{\hat{s}^{[i]}}{\hat{s}} + \sum_{j=m+1}^k x_j^{[i]} \log \frac{1-\hat{s}^{[i]}}{1-\hat{s}} + \sum_{j=1}^m (x_j^{[i]} + y_j^{[i]}) \log \frac{\hat{u}_j^{[i]}}{\hat{u}_j} \right\} \quad (20)$$

を得る。同様に

$$\begin{aligned} \text{Between information} &= \hat{E}_{H_2} \left(\log \frac{\hat{P}(\mathbf{X}^{[1]}, \mathbf{Y}^{[1]}, \mathbf{X}^{[2]}, \mathbf{Y}^{[2]} \mid H_2)}{P(\mathbf{X}^{[1]}, \mathbf{Y}^{[1]}, \mathbf{X}^{[2]}, \mathbf{Y}^{[2]} \mid H_2)} \right) \\ &= \sum_{i=1}^2 T_m^{[i]}(x) \log \frac{\hat{s}}{s} + \sum_{j=m+1}^k \sum_{i=1}^2 x_j^{[i]} \log \frac{1-\hat{s}}{1-s} + \sum_{j=1}^m \sum_{i=1}^2 (x_j^{[i]} + y_j^{[i]}) \log \frac{\hat{u}_j}{u_j} \end{aligned} \quad (21)$$

が得られる。よって、(20) と (21) の和は (19) に等しい。

同様に、Pooling incomplete samples が nested の状態で複数回行われる場合の 2 標本問題においても、Kullback 情報量の直和分解を、パラメータ変換を上手に用いることで簡単に示すことができる。

Proposition 3. 各 $i = 1, 2$ に対して $\mathbf{X}^{[i]} = (X_1^{[i]}, \dots, X_l^{[i]}, \dots, X_m^{[i]}, \dots, X_k^{[i]})$, $\mathbf{Y}^{[i]} = (Y_1^{[i]}, \dots, Y_l^{[i]}, \dots, Y_m^{[i]})$, $\mathbf{Z}^{[i]} = (Z_1^{[i]}, \dots, Z_l^{[i]})$ が

$$\mathbf{X}^{[i]} \sim \text{Multinomial}(N_x^{[i]}; \theta_1^{[i]}, \dots, \theta_l^{[i]}, \dots, \theta_m^{[i]}, \dots, \theta_k^{[i]}),$$

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

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

であり、 $\mathbf{X}^{[1]}, \mathbf{Y}^{[1]}, \mathbf{Z}^{[1]}, \mathbf{X}^{[2]}, \mathbf{Y}^{[2]}, \mathbf{Z}^{[2]}$ は互いに独立とする。仮説 H_1, H_2 を

$$H_1 : (\theta_1^{[1]}, \dots, \theta_k^{[1]}) \neq (\theta_1^{[2]}, \dots, \theta_k^{[2]}), \quad H_2 : \theta_i^{[1]} = \theta_i^{[2]} = \theta_i$$

に選ぶ。このとき、Kullback 情報量の直和分解が成り立つ。

proof 上記 Proposition は、パラメータ変換

$$\theta_j^{[i]} = \begin{cases} s^{[i]} t^{[i]} u_j^{[i]} & j = 1, \dots, l, \\ s^{[i]} (1 - t^{[i]}) v_j^{[i]} & j = l + 1, \dots, m, \\ (1 - s^{[i]}) w_j^{[i]} & j = m + 1, \dots, k \end{cases}$$

により、

$$\begin{aligned} \mathbf{X}^{[i]} &= (X_1^{[i]}, \dots, X_l^{[i]}, X_{l+1}^{[i]}, \dots, X_m^{[i]}, X_{m+1}^{[i]}, \dots, X_k^{[i]}) \\ &\sim \text{Multinomial}(N_x^{[i]}; s^{[i]} t^{[i]} u_1^{[i]}, \dots, s^{[i]} t^{[i]} u_l^{[i]}, s^{[i]} (1 - t^{[i]}) v_{l+1}^{[i]}, \\ &\quad \dots, s^{[i]} (1 - t^{[i]}) v_m^{[i]}, (1 - s^{[i]}) w_{m+1}^{[i]}, \dots, (1 - s^{[i]}) w_k^{[i]}), \end{aligned}$$

$$\begin{aligned} \mathbf{Y}^{[i]} &= (Y_1^{[i]}, \dots, Y_l^{[i]}, Y_{l+1}^{[i]}, \dots, Y_m^{[i]}) \\ &\sim \text{Multinomial}(N_y^{[i]}; t^{[i]} u_1^{[i]}, \dots, t^{[i]} u_l^{[i]}, (1 - t^{[i]}) v_{l+1}^{[i]}, \dots, (1 - t^{[i]}) v_m^{[i]}), \end{aligned}$$

$$\mathbf{Z}^{[i]} = (Z_1^{[i]}, \dots, Z_l^{[i]}) \sim \text{Multinomial}(N_z^{[i]}; u_1^{[i]}, \dots, u_l^{[i]}),$$

と書き換えることができる。このような書き換えにより、仮説 H_1, H_2 は

$$\begin{aligned} H_2 : & \quad u_j^{[1]} = u_j^{[2]} = u_j, \quad v_j^{[1]} = v_j^{[2]} = v_j, \quad w_j^{[1]} = w_j^{[2]} = w_j, \\ & \quad s^{[1]} = s^{[2]} = s, \quad t^{[1]} = t^{[2]} = t, \end{aligned}$$

$$H_1 : \text{not } H_2$$

となる。

$$\begin{aligned} & \frac{P(\mathbf{x}^{[1]}, \mathbf{y}^{[1]}, \mathbf{z}^{[1]}, \mathbf{x}^{[2]}, \mathbf{y}^{[2]}, \mathbf{z}^{[2]} \mid H_1)}{P(\mathbf{x}^{[1]}, \mathbf{y}^{[1]}, \mathbf{z}^{[1]}, \mathbf{x}^{[2]}, \mathbf{y}^{[2]}, \mathbf{z}^{[2]} \mid H_2)} \\ &= \prod_{i=1}^2 \left\{ \left(\frac{s^{[i]}}{s} \right)^{\sum_{j=1}^m x_j^{[i]}} \left(\frac{1-s^{[i]}}{1-s} \right)^{\sum_{j=m+1}^k x_j^{[i]}} \left(\frac{t^{[i]}}{t} \right)^{\sum_{j=1}^l (x_j^{[i]} + y_j^{[i]})} \left(\frac{1-t^{[i]}}{1-t} \right)^{\sum_{j=l+1}^m (x_j^{[i]} + y_j^{[i]})} \right. \\ & \quad \times \prod_{j=1}^l \left(\frac{u_j^{[i]}}{u_j} \right)^{x_j^{[i]} + y_j^{[i]} + z_j^{[i]}} \prod_{j=l+1}^m \left(\frac{v_j^{[i]}}{v_j} \right)^{x_j^{[i]} + y_j^{[i]}} \prod_{j=m+1}^k \left(\frac{w_j^{[i]}}{w_j} \right)^{x_j^{[i]}} \left. \right\}, \end{aligned}$$

$$\begin{aligned} & E_{H_1} \left(\log \frac{P(\mathbf{X}^{[1]}, \mathbf{Y}^{[1]}, \mathbf{Z}^{[1]}, \mathbf{X}^{[2]}, \mathbf{Y}^{[2]}, \mathbf{Z}^{[2]} \mid H_1)}{P(\mathbf{X}^{[1]}, \mathbf{Y}^{[1]}, \mathbf{Z}^{[1]}, \mathbf{X}^{[2]}, \mathbf{Y}^{[2]}, \mathbf{Z}^{[2]} \mid H_2)} \right) \\ &= \sum_{i=1}^2 \left\{ N_x^{[i]} s^{[i]} \log \frac{s^{[i]}}{s} + N_x^{[i]} (1-s^{[i]}) \log \frac{1-s^{[i]}}{1-s} + (N_x^{[i]} s^{[i]} + N_y^{[i]} t^{[i]}) \log \frac{t^{[i]}}{t} \right. \\ & \quad + (N_x^{[i]} s^{[i]} + N_y^{[i]})(1-t^{[i]}) \log \frac{1-t^{[i]}}{1-t} + (N_x^{[i]} s^{[i]} t^{[i]} + N_y^{[i]} t^{[i]} + N_z^{[i]}) \sum_{j=1}^l u_j^{[i]} \log \frac{u_j^{[i]}}{u_j} \\ & \quad \left. + (N_x^{[i]} s^{[i]} + N_y^{[i]})(1-t^{[i]}) \sum_{j=l+1}^m v_j^{[i]} \log \frac{v_j^{[i]}}{v_j} + N_x^{[i]} \sum_{j=m+1}^k w_j^{[i]} \log \frac{w_j^{[i]}}{w_j} \right\}, \end{aligned}$$

$$\hat{s}^{[i]} = \frac{T_m^{[i]}(x)}{N_x^{[i]}}, \quad \hat{t}^{[i]} = \frac{T_l^{[i]}(x) + T_l^{[i]}(y)}{T_m^{[i]}(x) + N_y^{[i]}}, \quad \hat{u}_j^{[i]} = \frac{x_j^{[i]} + y_j^{[i]} + z_j^{[i]}}{T_l^{[i]}(x) + T_l^{[i]}(y) + N_z^{[i]}}, \quad j \leq l,$$

$$\hat{v}_j^{[i]} = \frac{x_j^{[i]} + y_j^{[i]}}{\sum_{j=l+1}^m (x_j^{[i]} + y_j^{[i]})}, \quad l < j \leq m, \quad \hat{w}_j^{[i]} = \frac{x_j^{[i]}}{\sum_{j=m+1}^k x_j^{[i]}}, \quad m < j \leq k$$

により

$$\begin{aligned} \text{Total information} &= \sum_{i=1}^2 \left\{ T_m^{[i]}(x) \log \frac{\hat{s}^{[i]}}{s} + \sum_{j=m+1}^k x_j^{[i]} \log \frac{1-\hat{s}^{[i]}}{1-s} \right. \\ & \quad + (T_l^{[i]}(x) + T_l^{[i]}(y)) \log \frac{\hat{t}^{[i]}}{t} + \sum_{j=l+1}^m (x_j^{[i]} + y_j^{[i]}) \log \frac{1-\hat{t}^{[i]}}{1-t} \\ & \quad \left. + \sum_{j=1}^l (x_j^{[i]} + y_j^{[i]} + z_j^{[i]}) \log \frac{\hat{u}_j^{[i]}}{u_j} + \sum_{j=l+1}^m (x_j^{[i]} + y_j^{[i]}) \log \frac{\hat{v}_j^{[i]}}{v_j} + \sum_{j=m+1}^k x_j^{[i]} \log \frac{\hat{w}_j^{[i]}}{w_j} \right\} \end{aligned}$$

$$\begin{aligned} \text{Within information} &= \sum_{i=1}^2 \left\{ T_m^{[i]}(x) \log \frac{\hat{s}^{[i]}}{\hat{s}} + \sum_{j=m+1}^k x_j^{[i]} \log \frac{1-\hat{s}^{[i]}}{1-\hat{s}} \right. \\ & \quad + (T_l^{[i]}(x) + T_l^{[i]}(y)) \log \frac{\hat{t}^{[i]}}{\hat{t}} + \sum_{j=l+1}^m (x_j^{[i]} + y_j^{[i]}) \log \frac{1-\hat{t}^{[i]}}{1-\hat{t}} \\ & \quad \left. + \sum_{j=1}^l (x_j^{[i]} + y_j^{[i]} + z_j^{[i]}) \log \frac{\hat{u}_j^{[i]}}{\hat{u}_j} + \sum_{j=l+1}^m (x_j^{[i]} + y_j^{[i]}) \log \frac{\hat{v}_j^{[i]}}{\hat{v}_j} + \sum_{j=m+1}^k x_j^{[i]} \log \frac{\hat{w}_j^{[i]}}{\hat{w}_j} \right\}. \end{aligned}$$

同様に

$$\begin{aligned}
 \text{Between information} &= \hat{E}_{H_2} \left(\log \frac{\hat{P}(\mathbf{X}^{[1]}, \mathbf{Y}^{[1]}, \mathbf{Z}^{[1]}, \mathbf{X}^{[2]}, \mathbf{Y}^{[2]}, \mathbf{Z}^{[2]} | H_2)}{P(\mathbf{X}^{[1]}, \mathbf{Y}^{[1]}, \mathbf{Z}^{[1]}, \mathbf{X}^{[2]}, \mathbf{Y}^{[2]}, \mathbf{Z}^{[2]} | H_2)} \right) \\
 &= \sum_{i=1}^2 T_m^{[i]}(x) \log \frac{\hat{s}}{s} + \sum_{j=m+1}^k \sum_{i=1}^2 x_j^{[i]} \log \frac{1-\hat{s}}{1-s} + \sum_{i=1}^2 (T_l^{[i]}(x) + T_l^{[i]}(y)) \log \frac{\hat{t}}{t} \\
 &\quad + \sum_{j=l+1}^m \sum_{i=1}^2 (x_j^{[i]} + y_j^{[i]}) \log \frac{1-\hat{t}}{1-t} + \sum_{j=1}^l \sum_{i=1}^2 (x_j^{[i]} + y_j^{[i]} + z_j^{[i]}) \log \frac{\hat{u}_j}{u_j} \\
 &\quad + \sum_{j=l+1}^m \sum_{i=1}^2 (x_j^{[i]} + y_j^{[i]}) \log \frac{\hat{v}_j}{v_j} + \sum_{j=m+1}^k \sum_{i=1}^2 x_j^{[i]} \log \frac{\hat{w}_j}{w_j}.
 \end{aligned}$$

以上より、Kullback 情報量の直和分解が成り立つ。

5. パラメータを分離して、Kullback 情報量の直和分解が成立するモデルを構築する

例 5.1 $\mathbf{X}^{[1]}, \mathbf{X}^{[2]}, \mathbf{Y}^{[1]}, \mathbf{Y}^{[2]}$ は互いに独立で、各 $i = 1, 2$ に対し

$$\begin{aligned}
 \mathbf{X}^{[i]} &= (X_1^{[i]}, \dots, X_m^{[i]}, \dots, X_k^{[i]}) \sim \text{Negative Multinomial}(r_x^{[i]}, \theta_1^{[i]}, \dots, \theta_m^{[i]}, \dots, \theta_k^{[i]}), \\
 \mathbf{Y}^{[i]} &= (Y_1^{[i]}, \dots, Y_m^{[i]}) \sim \text{Negative Multinomial}(r_y^{[i]}, \frac{\theta_1^{[i]}}{\sum_{j=0}^m \theta_j^{[i]}}, \dots, \frac{\theta_m^{[i]}}{\sum_{j=0}^m \theta_j^{[i]}})
 \end{aligned}$$

なる確率モデルでは、Total information $\neq$ Within information + Between information であった。そこで、上記に「似たモデル」で Kullback 情報量の直和分解が成り立つモデルを探してみたところ、次のようなものが見つかった。

例 5.2 $\mathbf{X}^{[1]}, \mathbf{X}^{[2]}, \mathbf{Y}^{[1]}, \mathbf{Y}^{[2]}$ は互いに独立で、各 $i = 1, 2$ に対し

$$\begin{aligned}
 \mathbf{X}^{[i]} &= (X_1^{[i]}, \dots, X_m^{[i]}, \dots, X_k^{[i]}) \sim \text{Negative Multinomial}(r_x^{[i]}, p_1^{[i]}, \dots, p_m^{[i]}, \dots, p_k^{[i]}), \\
 \mathbf{Y}^{[i]} &= (Y_1^{[i]}, \dots, Y_m^{[i]}) \sim \text{Negative Multinomial}(r_y^{[i]}, q_1^{[i]}, \dots, q_m^{[i]}),
 \end{aligned}$$

但し、

$$p_j^{[i]} = \begin{cases} 1 - s^{[i]} & j = 0, \\ s^{[i]} t^{[i]} u_j^{[i]} & j = 1, \dots, m, \\ s^{[i]} (1 - t^{[i]}) v_j^{[i]} & j = m+1, \dots, k, \end{cases} \quad q_j^{[i]} = \begin{cases} 1 - w^{[i]} & j = 0, \\ w^{[i]} u_j^{[i]} & j = 1, \dots, m \end{cases}$$

とする。仮説 H_1, H_2 を $H_2 : u_j^{[1]} = u_j^{[2]} = u_j, v_j^{[1]} = v_j^{[2]} = v_j, s^{[1]} = s^{[2]} = s, H_1 : \text{not } H_2$ に選ぶ。このとき、Kullback 情報量の直和分解が成立する。

実際、

$$\begin{aligned} & \frac{P(\mathbf{x}^{[1]}, \mathbf{y}^{[1]}, \mathbf{x}^{[2]}, \mathbf{y}^{[2]} | H_1)}{P(\mathbf{x}^{[1]}, \mathbf{y}^{[1]}, \mathbf{x}^{[2]}, \mathbf{y}^{[2]} | H_2)} \\ &= \prod_{i=1}^2 \left\{ \left(\frac{1-s^{[i]}}{1-s} \right)^{r_x^{[i]}} \left(\frac{s^{[i]}}{s} \right)^{T_k^{[i]}(x)} \left(\frac{t^{[i]}}{t} \right)^{T_m^{[i]}(x)} \left(\frac{1-t^{[i]}}{1-t} \right)^{\sum_{j=m+1}^k x_j^{[i]}} \right. \\ & \quad \left. \times \left(\frac{1-w^{[i]}}{1-w} \right)^{r_y^{[i]}} \left(\frac{w^{[i]}}{w} \right)^{T_m^{[i]}(y)} \prod_{j=1}^m \left(\frac{u_j^{[i]}}{u_j} \right)^{x_j^{[i]}+y_j^{[i]}} \prod_{j=m+1}^k \left(\frac{v_j^{[i]}}{v_j} \right)^{x_j^{[i]}} \right\} \end{aligned}$$

であるから

$$\begin{aligned} \text{Total information} &= \hat{E}_{H_1} \left(\log \frac{\hat{P}(\mathbf{X}^{[1]}, \mathbf{Y}^{[1]}, \mathbf{X}^{[2]}, \mathbf{Y}^{[2]} | H_1)}{P(\mathbf{X}^{[1]}, \mathbf{Y}^{[1]}, \mathbf{X}^{[2]}, \mathbf{Y}^{[2]} | H_2)} \right) \\ &= \sum_{i=1}^2 \left\{ r_x^{[i]} \log \frac{1-\hat{s}^{[i]}}{1-s} + \hat{E}_{H_1} T_k^{[i]}(X) \log \frac{\hat{s}^{[i]}}{s} + \hat{E}_{H_1} T_m^{[i]}(X) \log \frac{\hat{t}^{[i]}}{t} \right. \\ & \quad + \hat{E}_{H_1} (T_k^{[i]}(X) - T_m^{[i]}(X)) \log \frac{1-\hat{t}^{[i]}}{1-t} + r_y^{[i]} \log \frac{1-\hat{w}^{[i]}}{1-w} + \hat{E}_{H_1} T_m^{[i]}(Y) \log \frac{\hat{w}^{[i]}}{w} \\ & \quad \left. + \sum_{j=1}^m \hat{E}_{H_1} (X_j^{[i]} + Y_j^{[i]}) \log \frac{\hat{u}_j^{[i]}}{u_j} + \sum_{j=m+1}^k \hat{E}_{H_1} (X_j^{[i]}) \log \frac{\hat{v}_j^{[i]}}{v_j} \right\}. \end{aligned}$$

$j = 1, \dots, m$ に対して $E_{H_1}(X_j^{[i]}) = r_x^{[i]} \frac{s^{[i]} t^{[i]} u_j^{[i]}}{1-s^{[i]}}$, $E_{H_1}(Y_j^{[i]}) = r_y^{[i]} \frac{w^{[i]} u_j^{[i]}}{1-w^{[i]}}$ であり、
 $j = m+1, \dots, k$ に対して $E_{H_1}(X_j^{[i]}) = r_x^{[i]} \frac{s^{[i]} (1-t^{[i]}) v_j^{[i]}}{1-s^{[i]}}$ である。そして、各パラメータの best estimator は

$$\begin{aligned} \hat{s}^{[i]} &= \frac{T_k^{[i]}(x)}{r_x^{[i]} + T_k^{[i]}(x)}, \quad \hat{w}^{[i]} = \frac{T_m^{[i]}(y)}{r_y^{[i]} + T_m^{[i]}(y)}, \quad \hat{t}^{[i]} = \frac{T_m^{[i]}(x)}{T_k^{[i]}(x)}, \\ \hat{u}_j^{[i]} &= \frac{x_j^{[i]} + y_j^{[i]}}{T_m^{[i]}(x) + T_m^{[i]}(y)}, \quad \hat{v}_j^{[i]} = \frac{x_j^{[i]}}{T_k^{[i]}(x) - T_m^{[i]}(x)} \end{aligned}$$

であるから、

$$\hat{E}_{H_1}(X_j^{[i]}) = \frac{T_m^{[i]}(x)(x_j^{[i]} + y_j^{[i]})}{T_m^{[i]}(x) + T_m^{[i]}(y)}, \quad \hat{E}_{H_1}(Y_j^{[i]}) = \frac{T_m^{[i]}(y)(x_j^{[i]} + y_j^{[i]})}{T_m^{[i]}(x) + T_m^{[i]}(y)} \quad \text{for } j = 1, \dots, m,$$

$$\hat{E}_{H_1}(X_j^{[i]}) = x_j^{[i]} \quad \text{for } j = m+1, \dots, k$$

を得る。これらを代入して

$$\text{Total information} = \sum_{i=1}^2 \left\{ r_x^{[i]} \log \frac{1-\hat{s}^{[i]}}{1-s} + T_k^{[i]}(x) \log \frac{\hat{s}^{[i]}}{s} + T_m^{[i]}(x) \log \frac{\hat{t}^{[i]}}{t} \right.$$

$$\begin{aligned}
& + \sum_{j=m+1}^k x_j^{[i]} \log \frac{1 - \hat{t}^{[i]}}{1 - \hat{t}} + r_y^{[i]} \log \frac{1 - \hat{w}^{[i]}}{1 - \hat{w}} + T_m^{[i]}(y) \log \frac{\hat{w}^{[i]}}{\hat{w}} \\
& + \sum_{j=1}^m (x_j^{[i]} + y_j^{[i]}) \log \frac{\hat{u}_j^{[i]}}{\hat{u}_j} + \sum_{j=m+1}^k x_j^{[i]} \log \frac{\hat{v}_j^{[i]}}{\hat{v}_j} \Big\}. \quad (22)
\end{aligned}$$

これより直ちに

$$\begin{aligned}
\text{Within information} = & \sum_{i=1}^2 \left\{ r_x^{[i]} \log \frac{1 - \hat{s}^{[i]}}{1 - \hat{s}} + T_k^{[i]}(x) \log \frac{\hat{s}^{[i]}}{\hat{s}} + T_m^{[i]}(x) \log \frac{\hat{t}^{[i]}}{\hat{t}} \right. \\
& + \sum_{j=m+1}^k x_j^{[i]} \log \frac{1 - \hat{t}^{[i]}}{1 - \hat{t}} + r_y^{[i]} \log \frac{1 - \hat{w}^{[i]}}{1 - \hat{w}} + T_m^{[i]}(y) \log \frac{\hat{w}^{[i]}}{\hat{w}} \\
& \left. + \sum_{j=1}^m (x_j^{[i]} + y_j^{[i]}) \log \frac{\hat{u}_j^{[i]}}{\hat{u}_j} + \sum_{j=m+1}^k x_j^{[i]} \log \frac{\hat{v}_j^{[i]}}{\hat{v}_j} \right\} \quad (23)
\end{aligned}$$

を得る。同様な計算により

$$\begin{aligned}
\text{Between information} = & \hat{E}_{H_2} \left(\log \frac{\hat{P}(\mathbf{X}^{[1]}, \mathbf{Y}^{[1]}, \mathbf{X}^{[2]}, \mathbf{Y}^{[2]} | H_2)}{P(\mathbf{X}^{[1]}, \mathbf{Y}^{[1]}, \mathbf{X}^{[2]}, \mathbf{Y}^{[2]} | H_2)} \right) \\
= & (r_x^{[1]} + r_x^{[2]}) \log \frac{1 - \hat{s}}{1 - s} + (r_y^{[1]} + r_y^{[2]}) \log \frac{1 - \hat{w}}{1 - w} + \sum_{i=1}^2 \left\{ T_k^{[i]}(x) \log \frac{\hat{s}}{s} + T_m^{[i]}(x) \log \frac{\hat{t}}{t} \right. \\
& \left. + \sum_{j=m+1}^k x_j^{[i]} \log \frac{1 - \hat{t}}{1 - t} + T_m^{[i]}(y) \log \frac{\hat{w}}{w} + \sum_{j=1}^m (x_j^{[i]} + y_j^{[i]}) \log \frac{\hat{u}_j}{u_j} + \sum_{j=m+1}^k x_j^{[i]} \log \frac{\hat{v}_j}{v_j} \right\} \quad (24)
\end{aligned}$$

が得られる。以上より、(23) と (24) の和が (22) に等しい。

例 5.3 整数 $m(1), m(2), k$ を $2 \leq m(1) < m(2) < k$ に選ぶ。確率変数 $\mathbf{X}, \mathbf{Y}$ は互いに独立で

$$\begin{aligned}
\mathbf{X} &= (X_1, \dots, X_k) \sim \text{Multinomial}(N_x; \theta_1, \theta_2, \dots, \theta_k) \\
\mathbf{Y} &= (Y_1, \dots, Y_{m(1)}, Y_{m(2)+1}, \dots, Y_k) \\
&\sim \text{Multinomial}(N_y; \eta_1, \dots, \eta_{m(1)}, \eta_{m(2)+1}, \dots, \eta_k)
\end{aligned}$$

$$\text{where } \eta_j = \frac{\sum_{l=1}^{m(2)} \theta_l}{\sum_{l=1}^{m(1)} \theta_l} \theta_j, \quad j \leq m(1), \quad \eta_j = \theta_j \quad j \geq m(2) + 1$$

とする。このモデルの下での2標本問題にて、Total information $\neq$ Within information + Between information. そこで

$$\theta_j = \begin{cases} sru_j & j \leq m(1), \\ s(1-r)v_j & m(1)+1 \leq j \leq m(2), \\ (1-s)w_j & m(2)+1 \leq j, \end{cases} \quad \eta_j = \begin{cases} tu_j & j \leq m(1), \\ (1-t)w_j & m(2)+1 \leq j \end{cases}$$

と変更すると、Kullback 情報量の直和分解が成り立つ。他方、 $s = t$ とすると、Total information $\neq$ Within information + Between information.

実際、

$$\begin{aligned}
& \frac{P(\mathbf{x}^{[1]}, \mathbf{y}^{[1]}, \mathbf{x}^{[2]}, \mathbf{y}^{[2]} \mid H_1)}{P(\mathbf{x}^{[1]}, \mathbf{y}^{[1]}, \mathbf{x}^{[2]}, \mathbf{y}^{[2]} \mid H_2)} \\
&= \prod_{i=1}^2 \frac{\prod_{j=1}^{m(1)} (s^{[i]} r^{[i]} u_j^{[i]})^{x_j^{[i]}} \prod_{j=m(1)+1}^{m(2)} (s^{[i]} (1-r^{[i]}) v_j^{[i]})^{x_j^{[i]}} \prod_{j=m(2)+1}^k ((1-s^{[i]}) w_j^{[i]})^{x_j^{[i]}}}{\prod_{j=1}^{m(1)} (s r u_j)^{x_j^{[i]}} \prod_{j=m(1)+1}^{m(2)} (s(1-r) v_j)^{x_j^{[i]}} \prod_{j=m(2)+1}^k ((1-s) w_j)^{x_j^{[i]}}} \\
&\quad \times \frac{\prod_{j=1}^{m(1)} (t^{[i]} u_j^{[i]})^{y_j^{[i]}} \prod_{j=m(2)+1}^k ((1-t^{[i]}) w_j^{[i]})^{y_j^{[i]}}}{\prod_{j=1}^{m(1)} (t u_j)^{y_j^{[i]}} \prod_{j=m(2)+1}^k ((1-t) w_j)^{y_j^{[i]}}} \\
&= \prod_{i=1}^2 \left\{ \left(\frac{s^{[i]}}{s} \right)^{T_{m(2)}^{[i]}(x)} \left(\frac{1-s^{[i]}}{1-s} \right)^{N_x^{[i]} - T_{m(2)}^{[i]}(x)} \left(\frac{r^{[i]}}{r} \right)^{T_{m(1)}^{[i]}(x)} \left(\frac{1-r^{[i]}}{1-r} \right)^{T_{m(2)}^{[i]}(x) - T_{m(1)}^{[i]}(x)} \right. \\
&\quad \times \left(\frac{t^{[i]}}{t} \right)^{T_{m(1)}^{[i]}(y)} \left(\frac{1-t^{[i]}}{1-t} \right)^{N_y^{[i]} - T_{m(1)}^{[i]}(y)} \prod_{j=1}^{m(1)} \left(\frac{u_j^{[i]}}{u_j} \right)^{x_j^{[i]} + y_j^{[i]}} \\
&\quad \times \left. \prod_{j=m(1)+1}^{m(2)} \left(\frac{v_j^{[i]}}{v_j} \right)^{x_j^{[i]}} \prod_{j=m(2)+1}^k \left(\frac{w_j^{[i]}}{w_j} \right)^{x_j^{[i]} + y_j^{[i]}} \right\}
\end{aligned}$$

より

$$\begin{aligned}
& E_{H_1} \left(\log \frac{P(\mathbf{X}^{[1]}, \mathbf{Y}^{[1]}, \mathbf{X}^{[2]}, \mathbf{Y}^{[2]} \mid H_1)}{P(\mathbf{X}^{[1]}, \mathbf{Y}^{[1]}, \mathbf{X}^{[2]}, \mathbf{Y}^{[2]} \mid H_2)} \right) \\
&= \sum_{i=1}^2 \left\{ E_{H_1} T_{m(2)}^{[i]}(X) \log \frac{s^{[i]}}{s} + E_{H_1} \left(\sum_{l=m(2)+1}^k X_l^{[i]} \right) \log \frac{1-s^{[i]}}{1-s} + E_{H_1} T_{m(1)}^{[i]}(X) \log \frac{r^{[i]}}{r} \right. \\
&\quad + E_{H_1} \left(\sum_{l=m(1)+1}^{m(2)} X_l^{[i]} \right) \log \frac{1-r^{[i]}}{1-r} + E_{H_1} T_{m(1)}^{[i]}(Y) \log \frac{t^{[i]}}{t} \\
&\quad + E_{H_1} \left(\sum_{l=m(2)+1}^k Y_l^{[i]} \right) \log \frac{1-t^{[i]}}{1-t} + \sum_{j=1}^{m(1)} E_{H_1} (X_j^{[i]} + Y_j^{[i]}) \log \frac{u_j^{[i]}}{u_j} \\
&\quad + \sum_{j=m(1)+1}^{m(2)} E_{H_1} X_j^{[i]} \log \frac{v_j^{[i]}}{v_j} + \sum_{j=m(2)+1}^k E_{H_1} (X_j^{[i]} + Y_j^{[i]}) \log \frac{w_j^{[i]}}{w_j} \Big\} \\
&= \sum_{i=1}^2 \left\{ N_x^{[i]} s^{[i]} \log \frac{s^{[i]}}{s} + N_x^{[i]} (1-s^{[i]}) \log \frac{1-s^{[i]}}{1-s} + N_x^{[i]} s^{[i]} r^{[i]} \log \frac{r^{[i]}}{r} \right. \\
&\quad + N_x^{[i]} s^{[i]} (1-r^{[i]}) \log \frac{1-r^{[i]}}{1-r} + N_y^{[i]} t^{[i]} \log \frac{t^{[i]}}{t} + N_y^{[i]} (1-t^{[i]}) \log \frac{1-t^{[i]}}{1-t}
\end{aligned}$$

$$\begin{aligned}
& + \sum_{j=1}^{m(1)} (N_x^{[i]} s^{[i]} r^{[i]} + N_y^{[i]} t^{[i]}) u_j^{[i]} \log \frac{u_j^{[i]}}{u_j} + \sum_{j=m(1)+1}^{m(2)} N_x^{[i]} s^{[i]} (1 - r^{[i]}) v_j^{[i]} \log \frac{v_j^{[i]}}{v_j} \\
& + \sum_{j=m(2)+1}^k (N_x^{[i]} (1 - s^{[i]}) + N_y^{[i]} (1 - t^{[i]})) w_j^{[i]} \log \frac{w_j^{[i]}}{w_j} \Big\}.
\end{aligned}$$

そして、各パラメータの best estimator は

$$\hat{s}^{[i]} = \frac{T_{m(2)}^{[i]}(x)}{N_x^{[i]}}, \quad \hat{r}^{[i]} = \frac{T_{m(1)}^{[i]}(x)}{T_{m(2)}^{[i]}(x)}, \quad \hat{t}^{[i]} = \frac{T_{m(1)}^{[i]}(y)}{N_y^{[i]}}, \quad \hat{u}_j^{[i]} = \frac{x_j^{[i]} + y_j^{[i]}}{\sum_{l=1}^{m(1)} (x_l^{[i]} + y_l^{[i]})}$$

$$\hat{v}_j^{[i]} = \frac{x_j^{[i]}}{\sum_{l=m(1)+1}^{m(2)} x_l^{[i]}}, \quad \hat{w}_j^{[i]} = \frac{x_j^{[i]} + y_j^{[i]}}{\sum_{l=m(2)+1}^k (x_l^{[i]} + y_l^{[i]})}$$

であったから

$$\begin{aligned}
\text{Total information} &= \hat{E}_{H_1} \left(\log \frac{\hat{P}(\mathbf{X}^{[1]}, \mathbf{Y}^{[1]}, \mathbf{X}^{[2]}, \mathbf{Y}^{[2]} | H_1)}{P(\mathbf{X}^{[1]}, \mathbf{Y}^{[1]}, \mathbf{X}^{[2]}, \mathbf{Y}^{[2]} | H_2)} \right) \\
&= \sum_{i=1}^2 \left\{ T_{m(2)}^{[i]}(x) \log \frac{\hat{s}^{[i]}}{s} + \left(\sum_{l=m(2)+1}^k x_l^{[i]} \right) \log \frac{1 - \hat{s}^{[i]}}{1 - s} + T_{m(1)}^{[i]}(x) \log \frac{\hat{r}^{[i]}}{r} \right. \\
&\quad + \left(\sum_{l=m(1)+1}^{m(2)} x_l^{[i]} \right) \log \frac{1 - \hat{r}^{[i]}}{1 - r} + T_{m(1)}^{[i]}(y) \log \frac{\hat{t}^{[i]}}{t} + \left(\sum_{l=m(2)+1}^k y_l^{[i]} \right) \log \frac{1 - \hat{t}^{[i]}}{1 - t} \\
&\quad \left. + \sum_{j=1}^{m(1)} (x_j^{[i]} + y_j^{[i]}) \log \frac{\hat{u}_j^{[i]}}{u_j} + \sum_{j=m(1)+1}^{m(2)} x_j^{[i]} \log \frac{\hat{v}_j^{[i]}}{v_j} + \sum_{j=m(2)+1}^k (x_j^{[i]} + y_j^{[i]}) \log \frac{\hat{w}_j^{[i]}}{w_j} \right\}.
\end{aligned}$$

これより直ちに

Within information

$$\begin{aligned}
&= \sum_{i=1}^2 \left\{ T_{m(2)}^{[i]}(x) \log \frac{\hat{s}^{[i]}}{\hat{s}} + \left(\sum_{l=m(2)+1}^k x_l^{[i]} \right) \log \frac{1 - \hat{s}^{[i]}}{1 - \hat{s}} + T_{m(1)}^{[i]}(x) \log \frac{\hat{r}^{[i]}}{\hat{r}} \right. \\
&\quad + \left(\sum_{l=m(1)+1}^{m(2)} x_l^{[i]} \right) \log \frac{1 - \hat{r}^{[i]}}{1 - \hat{r}} + T_{m(1)}^{[i]}(y) \log \frac{\hat{t}^{[i]}}{\hat{t}} + \left(\sum_{l=m(2)+1}^k y_l^{[i]} \right) \log \frac{1 - \hat{t}^{[i]}}{1 - \hat{t}} \\
&\quad \left. + \sum_{j=1}^{m(1)} (x_j^{[i]} + y_j^{[i]}) \log \frac{\hat{u}_j^{[i]}}{\hat{u}_j} + \sum_{j=m(1)+1}^{m(2)} x_j^{[i]} \log \frac{\hat{v}_j^{[i]}}{\hat{v}_j} + \sum_{j=m(2)+1}^k (x_j^{[i]} + y_j^{[i]}) \log \frac{\hat{w}_j^{[i]}}{\hat{w}_j} \right\}
\end{aligned}$$

が得られる。同様な計算で

$$\begin{aligned}
\text{Between information} &= \hat{E}_{H_2} \log \frac{\hat{P}(\mathbf{X}^{[1]}, \mathbf{Y}^{[1]}, \mathbf{X}^{[2]}, \mathbf{Y}^{[2]} | H_2)}{P(\mathbf{X}^{[1]}, \mathbf{Y}^{[1]}, \mathbf{X}^{[2]}, \mathbf{Y}^{[2]} | H_2)} \\
&= (T_{m(2)}^{[1]}(x) + T_{m(2)}^{[2]}(x)) \log \frac{\hat{s}}{s} + \sum_{l=m(2)+1}^k (x_l^{[1]} + x_l^{[2]}) \log \frac{1 - \hat{s}}{1 - s}
\end{aligned}$$

$$\begin{aligned}
& + (T_{m(1)}^{[1]}(y) + T_{m(1)}^{[2]}(y)) \log \frac{\hat{r}}{r} + \sum_{l=m(1)+1}^{m(2)} (x_j^{[1]} + x_j^{[2]}) \log \frac{1 - \hat{r}}{1 - r} \\
& + (T_{m(1)}^{[1]}(y) + T_{m(1)}^{[2]}(y)) \log \frac{\hat{t}}{t} + \sum_{l=m(2)+1}^k (y_l^{[1]} + y_l^{[2]}) \log \frac{1 - \hat{t}}{1 - t} \\
& + \sum_{j=1}^{m(1)} (x_j^{[1]} + x_j^{[2]} + y_j^{[1]} + y_j^{[2]}) \log \frac{\hat{u}_j}{u_j} + \sum_{j=m(1)+1}^{m(2)} (x_j^{[1]} + x_j^{[2]}) \log \frac{\hat{v}_j}{v_j} \\
& + \sum_{j=m(2)+1}^k (x_j^{[1]} + x_j^{[2]} + y_j^{[1]} + y_j^{[2]}) \log \frac{\hat{w}_j}{w_j}.
\end{aligned}$$

よって、Total = Between + Within が成り立つ。

他方、 $s = t$ の場合、

$$\begin{aligned}
\hat{s}^{[i]} &= \frac{T_{m(2)}^{[i]}(x) + T_{m(1)}^{[i]}(y)}{N_x^{[i]} + N_y^{[i]}}, \quad \hat{r}^{[i]} = \frac{T_{m(1)}^{[i]}(x)}{T_{m(2)}^{[i]}(x)}, \quad \hat{u}_j^{[i]} = \frac{x_j^{[i]} + y_j^{[i]}}{\sum_{l=1}^{m(1)} (x_l^{[i]} + y_l^{[i]})}, \\
\hat{v}_j^{[i]} &= \frac{x_j^{[i]}}{\sum_{l=m(1)+1}^{m(2)} x_l^{[i]}}, \quad \hat{w}_j^{[i]} = \frac{x_j^{[i]} + y_j^{[i]}}{\sum_{l=m(2)+1}^k (x_l^{[i]} + y_l^{[i]})}
\end{aligned}$$

により

Total information

$$\begin{aligned}
&= \sum_{i=1}^2 \left\{ (T_{m(2)}^{[i]}(x) + T_{m(1)}^{[i]}(y)) \log \frac{\hat{s}^{[i]}}{s} + \sum_{l=m(2)+1}^k (x_l^{[i]} + y_l^{[i]}) \log \frac{1 - \hat{s}^{[i]}}{1 - s} \right. \\
&+ \sum_{j=m(2)+1}^k (x_j^{[i]} + y_j^{[i]}) \log \frac{\hat{w}_j^{[i]}}{w_j} + N_x^{[i]} \frac{T_{m(2)}^{[i]}(x) + T_{m(1)}^{[i]}(y)}{N_x^{[i]} + N_y^{[i]}} \frac{T_{m(1)}^{[i]}(x)}{T_{m(2)}^{[i]}(x)} \log \frac{\hat{r}^{[i]}}{r} \\
&+ N_x^{[i]} \frac{T_{m(2)}^{[i]}(x) + T_{m(1)}^{[i]}(y)}{N_x^{[i]} + N_y^{[i]}} \frac{T_{m(2)}^{[i]}(x) - T_{m(1)}^{[i]}(x)}{T_{m(2)}^{[i]}(x)} \log \frac{1 - \hat{r}^{[i]}}{1 - r} \\
&+ \left(\frac{N_x^{[i]} T_{m(1)}^{[i]}(x)}{N_x^{[i]} + N_y^{[i]} T_{m(2)}^{[i]}(x)} + \frac{N_y^{[i]}}{N_x^{[i]} + N_y^{[i]}} \right) \frac{T_{m(2)}^{[i]}(x) + T_{m(1)}^{[i]}(y)}{T_{m(1)}^{[i]}(x) + T_{m(1)}^{[i]}(y)} \sum_{j=1}^{m(1)} (x_j^{[i]} + y_j^{[i]}) \log \frac{\hat{u}_j^{[i]}}{u_j} \\
&\left. + \frac{N_x^{[i]} T_{m(2)}^{[i]}(x) + T_{m(1)}^{[i]}(y)}{N_x^{[i]} + N_y^{[i]} T_{m(2)}^{[i]}(x)} \sum_{j=m(1)+1}^{m(2)} x_j^{[i]} \log \frac{\hat{v}_j^{[i]}}{v_j} \right\}, \tag{25}
\end{aligned}$$

Within information

$$= \sum_{i=1}^2 \left\{ (T_{m(2)}^{[i]}(x) + T_{m(1)}^{[i]}(y)) \log \frac{\hat{s}^{[i]}}{\hat{s}} + \sum_{l=m(2)+1}^k (x_l^{[i]} + y_l^{[i]}) \log \frac{1 - \hat{s}^{[i]}}{1 - \hat{s}} \right.$$

$$\begin{aligned}
& + \sum_{j=m(2)+1}^k (x_j^{[i]} + y_j^{[i]}) \log \frac{\hat{w}_j^{[i]}}{\hat{w}_j} + N_x^{[i]} \frac{T_{m(2)}^{[i]}(x) + T_{m(1)}^{[i]}(y)}{N_x^{[i]} + N_y^{[i]}} \frac{T_{m(1)}^{[i]}(x)}{T_{m(2)}^{[i]}(x)} \log \frac{\hat{r}^{[i]}}{\hat{r}} \\
& + N_x^{[i]} \frac{T_{m(2)}^{[i]}(x) + T_{m(1)}^{[i]}(y)}{N_x^{[i]} + N_y^{[i]}} \frac{T_{m(2)}^{[i]}(x) - T_{m(1)}^{[i]}(x)}{T_{m(2)}^{[i]}(x)} \log \frac{1 - \hat{r}^{[i]}}{1 - \hat{r}} \\
& + \left(\frac{N_x^{[i]} T_{m(1)}^{[i]}(x)}{N_x^{[i]} + N_y^{[i]} T_{m(2)}^{[i]}(x)} + \frac{N_y^{[i]} T_{m(2)}^{[i]}(x) + T_{m(1)}^{[i]}(y)}{N_x^{[i]} + N_y^{[i]} T_{m(1)}^{[i]}(y)} \right) \frac{m(1)}{\sum_{j=1}^{m(1)} (x_j^{[i]} + y_j^{[i]})} \log \frac{\hat{u}_j^{[i]}}{\hat{u}_j} \\
& + \frac{N_x^{[i]} T_{m(2)}^{[i]}(x) + T_{m(1)}^{[i]}(y)}{N_x^{[i]} + N_y^{[i]} T_{m(2)}^{[i]}(x)} \sum_{j=m(1)+1}^{m(2)} x_j^{[i]} \log \frac{\hat{v}_j^{[i]}}{\hat{v}_j} \} \quad (26)
\end{aligned}$$

を得る。また、同様な計算により

Between information

$$\begin{aligned}
& = \sum_{i=1}^2 (T_{m(2)}^{[i]}(x) + T_{m(1)}^{[i]}(y)) \log \frac{\hat{s}}{s} + \sum_{j=m(2)+1}^k (x_j^{[1]} + x_j^{[2]} + y_j^{[1]} + y_j^{[2]}) \log \frac{1 - \hat{s}}{1 - s} \\
& + \sum_{j=m(2)+1}^k (x_j^{[1]} + x_j^{[2]} + y_j^{[1]} + y_j^{[2]}) \log \frac{\hat{w}_j}{w_j} \\
& + \left(\sum_{i=1}^2 N_x^{[i]} \frac{\sum_{i=1}^2 (T_{m(2)}^{[i]}(x) + T_{m(1)}^{[i]}(y))}{\sum_{i=1}^2 (N_x^{[i]} + N_y^{[i]})} \frac{\sum_{i=1}^2 T_{m(1)}^{[i]}(x)}{\sum_{i=1}^2 T_{m(2)}^{[i]}(x)} \log \frac{\hat{r}}{r} \right. \\
& + \left(\sum_{i=1}^2 N_x^{[i]} \frac{\sum_{i=1}^2 (T_{m(2)}^{[i]}(x) + T_{m(1)}^{[i]}(y))}{\sum_{i=1}^2 (N_x^{[i]} + N_y^{[i]})} \frac{\sum_{i=1}^2 (T_{m(2)}^{[i]}(x) - T_{m(1)}^{[i]}(x))}{\sum_{i=1}^2 T_{m(2)}^{[i]}(x)} \log \frac{1 - \hat{r}}{1 - r} \right. \\
& + \left(\frac{\sum_{i=1}^2 N_x^{[i]} \sum_{i=1}^2 T_{m(1)}^{[i]}(x)}{\sum_{i=1}^2 (N_x^{[i]} + N_y^{[i]}) \sum_{i=1}^2 T_{m(2)}^{[i]}(x)} + \frac{\sum_{i=1}^2 N_y^{[i]}}{\sum_{i=1}^2 (N_x^{[i]} + N_y^{[i]})} \right) \\
& \quad \times \frac{\sum_{i=1}^2 (T_{m(2)}^{[i]}(x) + T_{m(1)}^{[i]}(y))}{\sum_{i=1}^2 (T_{m(1)}^{[i]}(x) + T_{m(1)}^{[i]}(y))} \frac{m(1)}{\sum_{j=1}^{m(1)} (x_j^{[1]} + x_j^{[2]} + y_j^{[1]} + y_j^{[2]})} \log \frac{\hat{u}_j}{u_j} \\
& + \left. \frac{\sum_{i=1}^2 N_x^{[i]} \sum_{i=1}^2 (T_{m(2)}^{[i]}(x) + T_{m(1)}^{[i]}(y))}{\sum_{i=1}^2 (N_x^{[i]} + N_y^{[i]}) \sum_{i=1}^2 T_{m(2)}^{[i]}(x)} \sum_{j=m(1)+1}^{m(2)} (x_j^{[1]} + x_j^{[2]}) \log \frac{\hat{v}_j}{v_j} \right) \quad (27)
\end{aligned}$$

を得る。(26) と (27) の和は、(25) に一致しない。

結論・今後の研究課題

本稿は、いくつかの多変量離散確率モデルにおける 2 標本問題において、Kullback 情報量の直和分解、すなわち、Total information が Within information と Between information の和になるか否かについて調べた成果の一部である。

かなり単純で、Kullback 情報量の直和分解が成り立ちような確率モデルにおいても成立しなかったり、複数回の pooling incomplete samples を伴う確率モデルで Kullback

情報量の直和分解が成立するなど、直和分解の可否はさまざまであることがわかった。直和分解が成り立つ場合、これを示すのに、パラメータを分離すると見通しよく割と簡単な計算で示すことができた。また、パラメータ分離によって、直和分解が成立しない確率モデルでも、確率モデルのパラメータの一部を別々なパラメータに置き換えることで、直和分解が成立する確率モデルを構築できるようないくつかの例を見い出した。

直和分解が成立するような確率モデルの条件を求めること、および、このような現象の背後に何らかの確率構造が潜んでいるのなら、これを探求することが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] 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 京都大学数理解析研究所講究録 No. 1954, 90-103, 2015 年 6 月.
- [3] Johnson, N.L., Kotz, S. and Balakrishnan, N. (1997). *Discrete Multivariate Distributions*. Wiley.
- [4] Kullback, S. (1959). *Information Theory and Statistics*. Wiley.
- [5] Kullback, S. (1968). *Information Theory and Statistics*, Revised edition. Dover.