数理解析研究所講究録
第1999巻 2016年 111-130

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 Econmics, 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 $\boldsymbol{X}^{[i]}=(X_1^{[i]},\cdots,X_k^{[i]})$ を未知母数 $\theta_1^{[i]},\cdots,\theta_m^{[i]}$, (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。$\boldsymbol{X}^{[1]}$ と $\boldsymbol{X}^{[2]}$ 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 $j=1,\cdots,m$, H_1 : not H_2 とする。このとき、Kullback 情報量 $I(H_1,H_2)$ は

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

である。データ $\boldsymbol{x}^{[i]}=(x_1^{[i]},\cdots,x_k^{[i]})$, $i=1,2,$ 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。

$$\text{Total information}=\hat E_{H_1}\left(\log\frac{P(\boldsymbol{X}^{[1]},\boldsymbol{X}^{[2]}\mid\hat\theta_1^{[1]},\cdots,\hat\theta_m^{[1]},\hat\theta_1^{[2]},\cdots,\hat\theta_m^{[2]})}{P(\boldsymbol{X}^{[1]},\boldsymbol{X}^{[2]}\mid\theta_1,\cdots,\theta_m)}\right),$$

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

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

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

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

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

であり、データ $\boldsymbol{x}^{[i]}=(x_1^{[i]},\cdots,x_k^{[i]})$ 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

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

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

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

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

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

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

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

とする。仮説 H_1, H_2 を

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

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。なぜであろうか？　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

$$H_2 : \theta_j^{[1]} = \theta_j^{[2]} = \theta_j \quad \text{for } j = 1, \cdots, m, \qquad H_1 : \text{not } H_2 \tag{6}$$

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

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

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

を用いることで、確率モデル (4) は

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

と書くことができ、仮説 (5) は

$$\left.\begin{array}{l} H_2 : u_j^{[1]} = u_j^{[2]} = u_j, \ v_j^{[1]} = v_j^{[2]} = v_j, \ w_j^{[1]} = w_j^{[2]} = w_j, \ s^{[1]} = s^{[2]} = s, \\ H_1 : \text{not } H_2 \end{array}\right\} \tag{8}$$

と書き表せる。このとき、各パラメータの best estimator は

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

であり、

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

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

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

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

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

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

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

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

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

を得る。同様に、

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

が得られる。(10) と (11) の和は (9) に一致しない。

他方、仮説 (6) は

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

と書き直せる。このとき

$$\frac{P(\boldsymbol{x}^{[1]},\boldsymbol{y}^{[1]},\boldsymbol{x}^{[2]},\boldsymbol{y}^{[2]}\mid H_1)}{P(\boldsymbol{x}^{[1]},\boldsymbol{y}^{[1]},\boldsymbol{x}^{[2]},\boldsymbol{y}^{[2]}\mid H_2)}$$
$$=\prod_{i=1}^{2}\frac{\prod_{j=1}^{m}(s^{[i]}u_j^{[i]})^{x_j^{[i]}+y_j^{[i]}}\prod_{j=m+1}^{k}((1-s^{[i]})v_j^{[i]})^{x_j^{[i]}}\prod_{j=m+1}^{k}((1-s^{[i]})w_j^{[i]})^{y_j^{[i]}}}{\prod_{j=1}^{m}(su_j)^{x_j^{[i]}+y_j^{[i]}}\prod_{j=m+1}^{k}((1-s)v_j^{[i]})^{x_j^{[i]}}\prod_{j=m+1}^{k}((1-s)w_j^{[i]})^{y_j^{[i]}}}$$
$$=\prod_{i=1}^{2}\left\{\left(\frac{s^{[i]}}{s}\right)^{\Sigma_{j=1}^{m}(x_j^{[i]}+y_j^{[i]})}\left(\frac{1-s^{[i]}}{1-s}\right)^{\Sigma_{j=m+1}^{k}(x_j^{[i]}+y_j^{[i]})}\prod_{j=1}^{m}\left(\frac{u_j^{[i]}}{u_j}\right)^{x_j^{[i]}+y_j^{[i]}}\right\}$$

すなわち

$$\log\frac{P(\boldsymbol{x}^{[1]},\boldsymbol{y}^{[1]},\boldsymbol{x}^{[2]},\boldsymbol{y}^{[2]}\mid H_1)}{P(\boldsymbol{x}^{[1]},\boldsymbol{y}^{[1]},\boldsymbol{x}^{[2]},\boldsymbol{y}^{[2]}\mid H_2)}=\sum_{i=1}^{2}\Bigg\{(T_m^{[i]}(x)+T_m^{[i]}(y))\log\frac{s^{[i]}}{s}$$
$$+\sum_{j=m+1}^{k}(x_j^{[i]}+y_j^{[i]})\log\frac{1-s^{[i]}}{1-s}+\sum_{j=1}^{m}(x_j^{[i]}+y_j^{[i]})\log\frac{u_j^{[i]}}{u_j}\Bigg\}$$

により

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

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

を得る。同様に

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

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

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

とする。仮説 H_1, H_2 を

$$H_2:\ u_j^{[1]}=u_j^{[2]}=u_j,\ v_j^{[1]}=v_j^{[2]}=v_j,\ w_j^{[1]}=w_j^{[2]}=w_j,$$
$$s^{[1]}=s^{[2]}=s,\ t^{[1]}=t^{[2]}=t,$$
$$H_1:\ \text{not } H_2$$

に選ぶ。このとき、 Kullback 情報量の直和分解が成立する。

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

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

である。これより

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

であり、$j = 1, \cdots, m$ に対して

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

$j = m+1, \cdots, k$ 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

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

そして、H_1 の下での best estimator が

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

であることを用いると、

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

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

同様に

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

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

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

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

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

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

$$= \prod_{i=1}^{2} \left\{ \left(\frac{s^{[i]}}{s}\right)^{T_m^{[i]}(x)} \left(\frac{1-s^{[i]}}{1-s}\right)^{N_x^{[i]} - T_m^{[i]}(x)} \prod_{j=1}^{m} \left(\frac{u_j^{[i]}}{u_j}\right)^{x_j^{[i]} + y_j^{[i]}} \right\},$$

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

であるから、これらを用いて

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

$$\begin{aligned}
&\text{Within information} \\
&= \sum_{i=1}^{2} \left\{ T_m^{[i]}(x) \log \frac{\hat{s}^{[i]}}{\hat{s}} + \sum_{j=m+1}^{k} x_j^{[i]} \log \frac{1-\hat{s}^{[i]}}{1-\hat{s}} + \sum_{j=1}^{m} (x_j^{[i]} + y_j^{[i]}) \log \frac{\hat{u}_j^{[i]}}{\hat{u}_j} \right\} \qquad (20)
\end{aligned}$$

を得る。同様に

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

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 $\boldsymbol{X}^{[i]}=(X_1^{[i]},\cdots,X_l^{[i]},\cdots,X_m^{[i]},\cdots,X_k^{[i]})$, $\boldsymbol{Y}^{[i]}=(Y_1^{[i]},\cdots,Y_l^{[i]},\cdots,Y_m^{[i]})$, $\boldsymbol{Z}^{[i]}=(Z_1^{[i]},\cdots,Z_l^{[i]})$ が

$$\boldsymbol{X}^{[i]}\sim \mathrm{Multinomial}(N_x^{[i]};\theta_1^{[i]},\cdots,\theta_l^{[i]},\cdots,\theta_m^{[i]},\cdots,\theta_k^{[i]}),$$

$$\boldsymbol{Y}^{[i]}\sim \mathrm{Multinomial}\Big(N_y^{[i]};\frac{\theta_1^{[i]}}{\sum_{a=1}^m\theta_a^{[i]}},\cdots,\frac{\theta_l^{[i]}}{\sum_{a=1}^m\theta_a^{[i]}},\cdots,\frac{\theta_m^{[i]}}{\sum_{a=1}^m\theta_a^{[i]}}\Big),$$

$$\boldsymbol{Z}^{[i]}\sim \mathrm{Multinomial}\Big(N_z^{[i]};\frac{\theta_1^{[i]}}{\sum_{a=1}^l\theta_a^{[i]}},\cdots,\frac{\theta_l^{[i]}}{\sum_{a=1}^l\theta_a^{[i]}}\Big),$$

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

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

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

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

により、

$$\begin{aligned}\boldsymbol{X}^{[i]}&=(X_1^{[i]},\cdots,X_l^{[i]},X_{l+1}^{[i]},\cdots,X_m^{[i]},X_{m+1}^{[i]},\cdots,X_k^{[i]})\\ &\quad\sim \mathrm{Multinomial}(N_x^{[i]};s^{[i]}t^{[i]}u_1^{[i]},\cdots,s^{[i]}t^{[i]}u_l^{[i]},\ s^{[i]}(1-t^{[i]})v_{l+1}^{[i]},\\ &\qquad\qquad \cdots,s^{[i]}(1-t^{[i]})v_m^{[i]},(1-s^{[i]})w_{m+1}^{[i]},\cdots,(1-s^{[i]})w_k^{[i]}),\\ \boldsymbol{Y}^{[i]}&=(Y_1^{[i]},\cdots,Y_l^{[i]},Y_{l+1}^{[i]},\cdots,Y_m^{[i]})\\ &\quad\sim \mathrm{Multinomial}(N_y^{[i]};t^{[i]}u_1^{[i]},\cdots,t^{[i]}u_l^{[i]},\ (1-t^{[i]})v_{l+1}^{[i]},\cdots,(1-t^{[i]})v_m^{[i]}),\\ \boldsymbol{Z}^{[i]}&=(Z_1^{[i]},\cdots,Z_l^{[i]})\sim \mathrm{Multinomial}(N_z^{[i]};u_1^{[i]},\cdots,u_l^{[i]}),\end{aligned}$$

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

$$\begin{aligned}H_2:&\ \ u_j^{[1]}=u_j^{[2]}=u_j,\ v_j^{[1]}=v_j^{[2]}=v_j,\ w_j^{[1]}=w_j^{[2]}=w_j,\\ &\ \ s^{[1]}=s^{[2]}=s,\ t^{[1]}=t^{[2]}=t,\\ H_1:&\ \ \text{not }H_2\end{aligned}$$

となる。

$$\frac{P(\boldsymbol{x}^{[1]},\boldsymbol{y}^{[1]},\boldsymbol{z}^{[1]},\boldsymbol{x}^{[2]},\boldsymbol{y}^{[2]},\boldsymbol{z}^{[2]}\mid H_1)}{P(\boldsymbol{x}^{[1]},\boldsymbol{y}^{[1]},\boldsymbol{z}^{[1]},\boldsymbol{x}^{[2]},\boldsymbol{y}^{[2]},\boldsymbol{z}^{[2]}\mid H_2)}$$

$$=\prod_{i=1}^{2}\left\{\left(\frac{s^{[i]}}{s}\right)^{\Sigma_{j=1}^{m}x_j^{[i]}}\left(\frac{1-s^{[i]}}{1-s}\right)^{\Sigma_{j=m+1}^{k}x_j^{[i]}}\left(\frac{t^{[i]}}{t}\right)^{\Sigma_{j=1}^{l}(x_j^{[i]}+y_j^{[i]})}\left(\frac{1-t^{[i]}}{1-t}\right)^{\Sigma_{j=l+1}^{m}(x_j^{[i]}+y_j^{[i]})}\right.$$

$$\left.\times\prod_{j=1}^{l}\left(\frac{u_j^{[i]}}{u_j}\right)^{x_j^{[i]}+y_j^{[i]}+z_j^{[i]}}\prod_{j=l+1}^{m}\left(\frac{v_j^{[i]}}{v_j}\right)^{x_j^{[i]}+y_j^{[i]}}\prod_{j=m+1}^{k}\left(\frac{w_j^{[i]}}{w_j}\right)^{x_j^{[i]}}\right\},$$

$$E_{H_1}\left(\log\frac{P(\boldsymbol{X}^{[1]},\boldsymbol{Y}^{[1]},\boldsymbol{Z}^{[1]},\boldsymbol{X}^{[2]},\boldsymbol{Y}^{[2]},\boldsymbol{Z}^{[2]}\mid H_1)}{P(\boldsymbol{X}^{[1]},\boldsymbol{Y}^{[1]},\boldsymbol{Z}^{[1]},\boldsymbol{X}^{[2]},\boldsymbol{Y}^{[2]},\boldsymbol{Z}^{[2]}\mid H_2)}\right)$$

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

$$+(N_x^{[i]}s^{[i]}+N_y^{[i]})(1-t^{[i]})\log\frac{1-t^{[i]}}{1-t}+(N_x^{[i]}s^{[i]}t^{[i]}+N_y^{[i]}t^{[i]}+N_z^{[i]})\sum_{j=1}^{l}u_j^{[i]}\log\frac{u_j^{[i]}}{u_j}$$

$$\left.+(N_x^{[i]}s^{[i]}+N_y^{[i]})(1-t^{[i]})\sum_{j=l+1}^{m}v_j^{[i]}\log\frac{v_j^{[i]}}{v_j}+N_x^{[i]}\sum_{j=m+1}^{k}w_j^{[i]}\log\frac{w_j^{[i]}}{w_j}\right\},$$

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

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

により

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

$$+(T_l^{[i]}(x)+T_l^{[i]}(y))\log\frac{\hat{t}^{[i]}}{t}+\sum_{j=l+1}^{m}(x_j^{[i]}+y_j^{[i]})\log\frac{1-\hat{t}^{[i]}}{1-t}$$

$$\left.+\sum_{j=1}^{l}(x_j^{[i]}+y_j^{[i]}+z_j^{[i]})\log\frac{\hat{u}_j^{[i]}}{u_j}+\sum_{j=l+1}^{m}(x_j^{[i]}+y_j^{[i]})\log\frac{\hat{v}_j^{[i]}}{v_j}+\sum_{j=m+1}^{k}x_j^{[i]}\log\frac{\hat{w}_j^{[i]}}{w_j}\right\}$$

$$\text{Within information}=\sum_{i=1}^{2}\left\{T_m^{[i]}(x)\log\frac{\hat{s}^{[i]}}{\hat{s}}+\sum_{j=m+1}^{k}x_j^{[i]}\log\frac{1-\hat{s}^{[i]}}{1-\hat{s}}\right.$$

$$+(T_l^{[i]}(x)+T_l^{[i]}(y))\log\frac{\hat{t}^{[i]}}{\hat{t}}+\sum_{j=l+1}^{m}(x_j^{[i]}+y_j^{[i]})\log\frac{1-\hat{t}^{[i]}}{1-\hat{t}}$$

$$\left.+\sum_{j=1}^{l}(x_j^{[i]}+y_j^{[i]}+z_j^{[i]})\log\frac{\hat{u}_j^{[i]}}{\hat{u}_j}+\sum_{j=l+1}^{m}(x_j^{[i]}+y_j^{[i]})\log\frac{\hat{v}_j^{[i]}}{\hat{v}_j}+\sum_{j=m+1}^{k}x_j^{[i]}\log\frac{\hat{w}_j^{[i]}}{\hat{w}_j}\right\}.$$

同様に

$$\text{Between information} = \hat{E}_{H_2}\left(\log\frac{\hat{P}(\boldsymbol{X}^{[1]},\boldsymbol{Y}^{[1]},\boldsymbol{Z}^{[1]},\boldsymbol{X}^{[2]},\boldsymbol{Y}^{[2]},\boldsymbol{Z}^{[2]}\mid H_2)}{P(\boldsymbol{X}^{[1]},\boldsymbol{Y}^{[1]},\boldsymbol{Z}^{[1]},\boldsymbol{X}^{[2]},\boldsymbol{Y}^{[2]},\boldsymbol{Z}^{[2]}\mid H_2)}\right)$$

$$= \sum_{i=1}^{2} T_m^{[i]}(x)\log\frac{\hat{s}}{s} + \sum_{j=m+1}^{k}\sum_{i=1}^{2} x_j^{[i]}\log\frac{1-\hat{s}}{1-s} + \sum_{i=1}^{2}(T_l^{[i]}(x)+T_l^{[i]}(y))\log\frac{\hat{t}}{t}$$

$$+ \sum_{j=l+1}^{m}\sum_{i=1}^{2}(x_j^{[i]}+y_j^{[i]})\log\frac{1-\hat{t}}{1-t} + \sum_{j=1}^{l}\sum_{i=1}^{2}(x_j^{[i]}+y_j^{[i]}+z_j^{[i]})\log\frac{\hat{u}_j}{u_j}$$

$$+ \sum_{j=l+1}^{m}\sum_{i=1}^{2}(x_j^{[i]}+y_j^{[i]})\log\frac{\hat{v}_j}{v_j} + \sum_{j=m+1}^{k}\sum_{i=1}^{2} x_j^{[i]}\log\frac{\hat{w}_j}{w_j}.$$

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

$$\mathbf{X}^{[i]} = (X_1^{[i]},\cdots,X_m^{[i]},\cdots,X_k^{[i]}) \sim \text{Negative Multinomial}(r_x^{[i]},\ \theta_1^{[i]},\cdots,\theta_m^{[i]},\cdots,\theta_k^{[i]}),$$

$$\mathbf{Y}^{[i]} = (Y_1^{[i]},\cdots,Y_m^{[i]}) \sim \text{Negative Multinomial}(r_y^{[i]},\ \frac{\theta_1^{[i]}}{\sum_{j=0}^{m}\theta_j^{[i]}},\cdots,\frac{\theta_m^{[i]}}{\sum_{j=0}^{m}\theta_j^{[i]}})$$

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

$$\mathbf{X}^{[i]} = (X_1^{[i]},\cdots,X_m^{[i]},\cdots,X_k^{[i]}) \sim \text{Negative Multinomial}(r_x^{[i]},\ p_1^{[i]},\cdots,p_m^{[i]},\cdots,p_k^{[i]}),$$

$$\mathbf{Y}^{[i]} = (Y_1^{[i]},\cdots,Y_m^{[i]}) \sim \text{Negative Multinomial}(r_y^{[i]},\ q_1^{[i]},\cdots,q_m^{[i]}),$$

但し、

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

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、

$$\frac{P(\boldsymbol{x}^{[1]},\boldsymbol{y}^{[1]},\boldsymbol{x}^{[2]},\boldsymbol{y}^{[2]}\mid H_1)}{P(\boldsymbol{x}^{[1]},\boldsymbol{y}^{[1]},\boldsymbol{x}^{[2]},\boldsymbol{y}^{[2]}\mid H_2)} = \prod_{i=1}^{2}\left\{\Big(\frac{1-s^{[i]}}{1-s}\Big)^{r_x^{[i]}}\Big(\frac{s^{[i]}}{s}\Big)^{T_k^{[i]}(x)}\Big(\frac{t^{[i]}}{t}\Big)^{T_m^{[i]}(x)}\Big(\frac{1-t^{[i]}}{1-t}\Big)^{\Sigma_{j=m+1}^{k}x_j^{[i]}}\right.$$
$$\left.\times\Big(\frac{1-w^{[i]}}{1-w}\Big)^{r_y^{[i]}}\Big(\frac{w^{[i]}}{w}\Big)^{T_m^{[i]}(y)}\prod_{j=1}^{m}\Big(\frac{u_j^{[i]}}{u_j}\Big)^{x_j^{[i]}+y_j^{[i]}}\prod_{j=m+1}^{k}\Big(\frac{v_j^{[i]}}{v_j}\Big)^{x_j^{[i]}}\right\}$$

であるから

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

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

$$\hat{s}^{[i]} = \frac{T_k^{[i]}(x)}{r_x^{[i]}+T_k^{[i]}(x)},\quad \hat{w}^{[i]} = \frac{T_m^{[i]}(y)}{r_y^{[i]}+T_m^{[i]}(y)},\quad \hat{t}^{[i]} = \frac{T_m^{[i]}(x)}{T_k^{[i]}(x)},$$

$$\hat{u}_j^{[i]} = \frac{x_j^{[i]}+y_j^{[i]}}{T_m^{[i]}(x)+T_m^{[i]}(y)},\quad \hat{v}_j^{[i]} = \frac{x_j^{[i]}}{T_k^{[i]}(x)-T_m^{[i]}(x)}$$

であるから、

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

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

を得る。これらを代入して

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

$$+\sum_{j=m+1}^{k} x_j^{[i]} \log \frac{1-\hat{t}^{[i]}}{1-t} + r_y^{[i]} \log \frac{1-\hat{w}^{[i]}}{1-w} + T_m^{[i]}(y) \log \frac{\hat{w}^{[i]}}{w}$$
$$+\sum_{j=1}^{m} (x_j^{[i]} + y_j^{[i]}) \log \frac{\hat{u}_j^{[i]}}{u_j} + \sum_{j=m+1}^{k} x_j^{[i]} \log \frac{\hat{v}_j^{[i]}}{v_j} \Bigg\}. \tag{22}$$

これより直ちに

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

を得る。同様な計算により

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

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 $2 \le m(1) < m(2) < k$ に選ぶ。確率変数 $\boldsymbol{X}$, $\boldsymbol{Y}$ は互いに独立で

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

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

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

$$\theta_j = \begin{cases} s r u_j & j \le m(1), \\ s(1-r) v_j & m(1)+1 \le j \le m(2), \\ (1-s) w_j & m(2)+1 \le j, \end{cases} \qquad \eta_j = \begin{cases} t u_j & j \le m(1), \\ (1-t) w_j & m(2)+1 \le j \end{cases}$$

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、

$$\frac{P(\boldsymbol{x}^{[1]},\boldsymbol{y}^{[1]},\boldsymbol{x}^{[2]},\boldsymbol{y}^{[2]}\mid H_1)}{P(\boldsymbol{x}^{[1]},\boldsymbol{y}^{[1]},\boldsymbol{x}^{[2]},\boldsymbol{y}^{[2]}\mid H_2)}$$

$$=\prod_{i=1}^{2}\frac{\prod_{j=1}^{m(1)}(s^{[i]}r^{[i]}u_j^{[i]})^{x_j^{[i]}}\prod_{j=m(1)+1}^{m(2)}(s^{[i]}(1-r^{[i]})v_j^{[i]})^{x_j^{[i]}}\prod_{j=m(2)+1}^{k}((1-s^{[i]})w_j^{[i]})^{x_j^{[i]}}}{\prod_{j=1}^{m(1)}(sru_j)^{x_j^{[i]}}\prod_{j=m(1)+1}^{m(2)}(s(1-r)v_j)^{x_j^{[i]}}\prod_{j=m(2)+1}^{k}((1-s)w_j)^{x_j^{[i]}}}$$

$$\times\frac{\prod_{j=1}^{m(1)}(t^{[i]}u_j^{[i]})^{y_j^{[i]}}\prod_{j=m(2)+1}^{k}((1-t^{[i]})w_j^{[i]})^{y_j^{[i]}}}{\prod_{j=1}^{m(1)}(tu_j)^{y_j^{[i]}}\prod_{j=m(2)+1}^{k}((1-t)w_j)^{y_j^{[i]}}}$$

$$=\prod_{i=1}^{2}\left\{\left(\frac{s^{[i]}}{s}\right)^{T_{m(2)}^{[i]}(x)}\left(\frac{1-s^{[i]}}{1-s}\right)^{N_x^{[i]}-T_{m(2)}^{[i]}(x)}\left(\frac{r^{[i]}}{r}\right)^{T_{m(1)}^{[i]}(x)}\left(\frac{1-r^{[i]}}{1-r}\right)^{T_{m(2)}^{[i]}(x)-T_{m(1)}^{[i]}(x)}\right.$$

$$\times\left(\frac{t^{[i]}}{t}\right)^{T_{m(1)}^{[i]}(y)}\left(\frac{1-t^{[i]}}{1-t}\right)^{N_y^{[i]}-T_{m(1)}^{[i]}(y)}\prod_{j=1}^{m(1)}\left(\frac{u_j^{[i]}}{u_j}\right)^{x_j^{[i]}+y_j^{[i]}}$$

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

より

$$E_{H_1}\left(\log\frac{P(\boldsymbol{X}^{[1]},\boldsymbol{Y}^{[1]},\boldsymbol{X}^{[2]},\boldsymbol{Y}^{[2]}\mid H_1)}{P(\boldsymbol{X}^{[1]},\boldsymbol{Y}^{[1]},\boldsymbol{X}^{[2]},\boldsymbol{Y}^{[2]}\mid H_2)}\right)$$

$$=\sum_{i=1}^{2}\left\{E_{H_1}T_{m(2)}^{[i]}(X)\log\frac{s^{[i]}}{s}+E_{H_1}\Big(\sum_{l=m(2)+1}^{k}X_l^{[i]}\Big)\log\frac{1-s^{[i]}}{1-s}+E_{H_1}T_{m(1)}^{[i]}(X)\log\frac{r^{[i]}}{r}\right.$$

$$+E_{H_1}\Big(\sum_{l=m(1)+1}^{m(2)}X_l^{[i]}\Big)\log\frac{1-r^{[i]}}{1-r}+E_{H_1}T_{m(1)}^{[i]}(Y)\log\frac{t^{[i]}}{t}$$

$$+E_{H_1}\Big(\sum_{l=m(2)+1}^{k}Y_l^{[i]}\Big)\log\frac{1-t^{[i]}}{1-t}+\sum_{j=1}^{m(1)}E_{H_1}(X_j^{[i]}+Y_j^{[i]})\log\frac{u_j^{[i]}}{u_j}$$

$$\left.+\sum_{j=m(1)+1}^{m(2)}E_{H_1}X_j^{[i]}\log\frac{v_j^{[i]}}{v_j}+\sum_{j=m(2)+1}^{k}E_{H_1}(X_j^{[i]}+Y_j^{[i]})\log\frac{w_j^{[i]}}{w_j}\right\}$$

$$=\sum_{i=1}^{2}\left\{N_x^{[i]}s^{[i]}\log\frac{s^{[i]}}{s}+N_x^{[i]}(1-s^{[i]})\log\frac{1-s^{[i]}}{1-s}+N_x^{[i]}s^{[i]}r^{[i]}\log\frac{r^{[i]}}{r}\right.$$

$$+N_x^{[i]}s^{[i]}(1-r^{[i]})\log\frac{1-r^{[i]}}{1-r}+N_y^{[i]}t^{[i]}\log\frac{t^{[i]}}{t}+N_y^{[i]}(1-t^{[i]})\log\frac{1-t^{[i]}}{1-t}$$

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

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

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

これより直ちに

$$\text{Within information}$$
$$=\sum_{i=1}^{2}\Bigg\{T_{m(2)}^{[i]}(x)\log\frac{\hat{s}^{[i]}}{\hat{s}}+\Big(\sum_{l=m(2)+1}^{k}x_l^{[i]}\Big)\log\frac{1-\hat{s}^{[i]}}{1-\hat{s}}+T_{m(1)}^{[i]}(x)\log\frac{\hat{r}^{[i]}}{\hat{r}}$$
$$+\Big(\sum_{l=m(1)+1}^{m(2)}x_l^{[i]}\Big)\log\frac{1-\hat{r}^{[i]}}{1-\hat{r}}+T_{m(1)}^{[i]}(y)\log\frac{\hat{t}^{[i]}}{\hat{t}}+\Big(\sum_{l=m(2)+1}^{k}y_l^{[i]}\Big)\log\frac{1-\hat{t}^{[i]}}{1-\hat{t}}$$
$$+\sum_{j=1}^{m(1)}(x_j^{[i]}+y_j^{[i]})\log\frac{\hat{u}_j^{[i]}}{\hat{u}_j}+\sum_{j=m(1)+1}^{m(2)}x_j^{[i]}\log\frac{\hat{v}_j^{[i]}}{\hat{v}_j}+\sum_{j=m(2)+1}^{k}(x_j^{[i]}+y_j^{[i]})\log\frac{\hat{w}_j^{[i]}}{\hat{w}_j}\Bigg\}$$

が得られる。同様な計算で

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

$$+(T^{[1]}_{m(1)}(y)+T^{[2]}_{m(1)}(y))\log\frac{\hat{r}}{r}+\sum_{l=m(1)+1}^{m(2)}(x^{[1]}_j+x^{[2]}_j)\log\frac{1-\hat{r}}{1-r}$$
$$+(T^{[1]}_{m(1)}(y)+T^{[2]}_{m(1)}(y))\log\frac{\hat{t}}{t}+\sum_{l=m(2)+1}^{k}(y^{[1]}_l+y^{[2]}_l)\log\frac{1-\hat{t}}{1-t}$$
$$+\sum_{j=1}^{m(1)}(x^{[1]}_j+x^{[2]}_j+y^{[1]}_j+y^{[2]}_j)\log\frac{\hat{u}_j}{u_j}+\sum_{j=m(1)+1}^{m(2)}(x^{[1]}_j+x^{[2]}_j)\log\frac{\hat{v}_j}{v_j}$$
$$+\sum_{j=m(2)+1}^{k}(x^{[1]}_j+x^{[2]}_j+y^{[1]}_j+y^{[2]}_j)\log\frac{\hat{w}_j}{w_j}.$$

よって、Total = Between + Within が成り立つ。

他方、$s=t$ の場合、

$$\hat{s}^{[i]}=\frac{T^{[i]}_{m(2)}(x)+T^{[i]}_{m(1)}(y)}{N^{[i]}_x+N^{[i]}_y},\quad \hat{r}^{[i]}=\frac{T^{[i]}_{m(1)}(x)}{T^{[i]}_{m(2)}(x)},\quad \hat{u}^{[i]}_j=\frac{x^{[i]}_j+y^{[i]}_j}{\sum_{l=1}^{m(1)}(x^{[i]}_l+y^{[i]}_l)},$$

$$\hat{v}^{[i]}_j=\frac{x^{[i]}_j}{\sum_{l=m(1)+1}^{m(2)}x^{[i]}_l},\quad \hat{w}^{[i]}_j=\frac{x^{[i]}_j+y^{[i]}_j}{\sum_{l=m(2)+1}^{k}(x^{[i]}_l+y^{[i]}_l)}$$

により

Total information

$$=\sum_{i=1}^{2}\left\{(T^{[i]}_{m(2)}(x)+T^{[i]}_{m(1)}(y))\log\frac{\hat{s}^{[i]}}{s}+\sum_{l=m(2)+1}^{k}(x^{[i]}_l+y^{[i]}_l)\log\frac{1-\hat{s}^{[i]}}{1-s}\right.$$
$$+\sum_{j=m(2)+1}^{k}(x^{[i]}_j+y^{[i]}_j)\log\frac{\hat{w}^{[i]}_j}{w_j}+N^{[i]}_x\frac{T^{[i]}_{m(2)}(x)+T^{[i]}_{m(1)}(y)}{N^{[i]}_x+N^{[i]}_y}\frac{T^{[i]}_{m(1)}(x)}{T^{[i]}_{m(2)}(x)}\log\frac{\hat{r}^{[i]}}{r}$$
$$+N^{[i]}_x\frac{T^{[i]}_{m(2)}(x)+T^{[i]}_{m(1)}(y)}{N^{[i]}_x+N^{[i]}_y}\frac{T^{[i]}_{m(2)}(x)-T^{[i]}_{m(1)}(x)}{T^{[i]}_{m(2)}(x)}\log\frac{1-\hat{r}^{[i]}}{1-r}$$
$$+\left(\frac{N^{[i]}_x}{N^{[i]}_x+N^{[i]}_y}\frac{T^{[i]}_{m(1)}(x)}{T^{[i]}_{m(2)}(x)}+\frac{N^{[i]}_y}{N^{[i]}_x+N^{[i]}_y}\right)\frac{T^{[i]}_{m(2)}(x)+T^{[i]}_{m(1)}(y)}{T^{[i]}_{m(1)}(x)+T^{[i]}_{m(1)}(y)}\sum_{j=1}^{m(1)}(x^{[i]}_j+y^{[i]}_j)\log\frac{\hat{u}^{[i]}_j}{u_j}$$
$$\left.+\frac{N^{[i]}_x}{N^{[i]}_x+N^{[i]}_y}\frac{T^{[i]}_{m(2)}(x)+T^{[i]}_{m(1)}(y)}{T^{[i]}_{m(2)}(x)}\sum_{j=m(1)+1}^{m(2)}x^{[i]}_j\log\frac{\hat{v}^{[i]}_j}{v_j}\right\},\tag{25}$$

Within information

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

$$+\sum_{j=m(2)+1}^{k}(x_j^{[i]}+y_j^{[i]})\log\frac{\hat{w}_j^{[i]}}{\hat{w}_j}+N_x^{[i]}\frac{T_{m(2)}^{[i]}(x)+T_{m(1)}^{[i]}(y)}{N_x^{[i]}+N_y^{[i]}}\frac{T_{m(1)}^{[i]}(x)}{T_{m(2)}^{[i]}(x)}\log\frac{\hat{r}^{[i]}}{\hat{r}}$$

$$+N_x^{[i]}\frac{T_{m(2)}^{[i]}(x)+T_{m(1)}^{[i]}(y)}{N_x^{[i]}+N_y^{[i]}}\frac{T_{m(2)}^{[i]}(x)-T_{m(1)}^{[i]}(x)}{T_{m(2)}^{[i]}(x)}\log\frac{1-\hat{r}^{[i]}}{1-\hat{r}}$$

$$+\left(\frac{N_x^{[i]}}{N_x^{[i]}+N_y^{[i]}}\frac{T_{m(1)}^{[i]}(x)}{T_{m(2)}^{[i]}(x)}+\frac{N_y^{[i]}}{N_x^{[i]}+N_y^{[i]}}\right)\frac{T_{m(2)}^{[i]}(x)+T_{m(1)}^{[i]}(y)}{T_{m(1)}^{[i]}(x)+T_{m(1)}^{[i]}(y)}\sum_{j=1}^{m(1)}(x_j^{[i]}+y_j^{[i]})\log\frac{\hat{u}_j^{[i]}}{\hat{u}_j}$$

$$+\frac{N_x^{[i]}}{N_x^{[i]}+N_y^{[i]}}\frac{T_{m(2)}^{[i]}(x)+T_{m(1)}^{[i]}(y)}{T_{m(2)}^{[i]}(x)}\sum_{j=m(1)+1}^{m(2)}x_j^{[i]}\log\frac{\hat{v}_j^{[i]}}{\hat{v}_j}\Bigg\} \tag{26}$$

を得る。また、同様な計算により

Between information

$$=\sum_{i=1}^{2}(T_{m(2)}^{[i]}(x)+T_{m(1)}^{[i]}(y))\log\frac{\hat{s}}{s}+\sum_{j=m(2)+1}^{k}(x_j^{[1]}+x_j^{[2]}+y_j^{[1]}+y_j^{[2]})\log\frac{1-\hat{s}}{1-s}$$

$$+\sum_{j=m(2)+1}^{k}(x_j^{[1]}+x_j^{[2]}+y_j^{[1]}+y_j^{[2]})\log\frac{\hat{w}_j}{w_j}$$

$$+\Big(\sum_{i=1}^{2}N_x^{[i]}\Big)\frac{\sum_{i=1}^{2}(T_{m(2)}^{[i]}(x)+T_{m(1)}^{[i]}(y))}{\sum_{i=1}^{2}(N_x^{[i]}+N_y^{[i]})}\frac{\sum_{i=1}^{2}T_{m(1)}^{[i]}(x)}{\sum_{i=1}^{2}T_{m(2)}^{[i]}(x)}\log\frac{\hat{r}}{r}$$

$$+\Big(\sum_{i=1}^{2}N_x^{[i]}\Big)\frac{\sum_{i=1}^{2}(T_{m(2)}^{[i]}(x)+T_{m(1)}^{[i]}(y))}{\sum_{i=1}^{2}(N_x^{[i]}+N_y^{[i]})}\frac{\sum_{i=1}^{2}(T_{m(2)}^{[i]}(x)-T_{m(1)}^{[i]}(x))}{\sum_{i=1}^{2}T_{m(2)}^{[i]}(x)}\log\frac{1-\hat{r}}{1-r}$$

$$+\left(\frac{\sum_{i=1}^{2}N_x^{[i]}}{\sum_{i=1}^{2}(N_x^{[i]}+N_y^{[i]})}\frac{\sum_{i=1}^{2}T_{m(1)}^{[i]}(x)}{\sum_{i=1}^{2}T_{m(2)}^{[i]}(x)}+\frac{\sum_{i=1}^{2}N_y^{[i]}}{\sum_{i=1}^{2}(N_x^{[i]}+N_y^{[i]})}\right)$$

$$\times\frac{\sum_{i=1}^{2}(T_{m(2)}^{[i]}(x)+T_{m(1)}^{[i]}(y))}{\sum_{i=1}^{2}(T_{m(1)}^{[i]}(x)+T_{m(1)}^{[i]}(y))}\sum_{j=1}^{m(1)}(x_j^{[1]}+x_j^{[2]}+y_j^{[1]}+y_j^{[2]})\log\frac{\hat{u}_j}{u_j}$$

$$+\frac{\sum_{i=1}^{2}N_x^{[i]}}{\sum_{i=1}^{2}(N_x^{[i]}+N_y^{[i]})}\frac{\sum_{i=1}^{2}(T_{m(2)}^{[i]}(x)+T_{m(1)}^{[i]}(y))}{\sum_{i=1}^{2}T_{m(2)}^{[i]}(x)}\sum_{j=m(1)+1}^{m(2)}(x_j^{[1]}+x_j^{[2]})\log\frac{\hat{v}_j}{v_j} \tag{27}$$

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そ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せた。

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.