例 1.19: 十分統計量 $F_{\overrightarrow{m}}$

$S_{\overrightarrow{m}}$ の定義

- $n \in \mathbf{N}$,
- $\overrightarrow{m} = (m_1, m_2, \ldots, m_n) \in \mathbf{N}^n$
- $m = m_1 + m_2 + \cdots + m_n$

とし，S_m の部分集合 $S_{\overrightarrow{m}}$ を

$$S_{\overrightarrow{m}} = \{p(\cdot;\vec{\eta}) \in S_m \mid \underbrace{\eta_1 = \cdots = \eta_{m_1}, \ldots, \eta_{m_1+\cdots+m_{n-1}+1} = \cdots = \eta_m}_{(\star)},\ \vec{\eta} \in \Xi_m\}$$

で定める．ここで $p(\cdot;\vec{\eta})$ は $p(0;\vec{\eta}) := 1 - \sum_{i=1}^m \eta_i$, $1 \le i \le m$ に対しては $p(i;\vec{\eta}) := \eta_i$ で定められる Ω_m 上の確率関数である．

$S_{\overrightarrow{m}}$ が Ω_m 上の n 次元統計的モデルであることの証明

各 $\vec{\xi} = (\xi_1, \ldots, \xi_n) \in \Xi_n$ に対し，確率関数 $\tilde{p}(\cdot;\vec{\xi})$ を

$$\tilde{p}(\cdot;\vec{\xi}) := p\left(\cdot;\underbrace{\frac{\xi_1}{m_1}, \ldots, \frac{\xi_1}{m_1}}_{m_1\text{個}}, \underbrace{\frac{\xi_2}{m_2}, \ldots, \frac{\xi_2}{m_2}}_{m_2\text{個}}, \ldots, \underbrace{\frac{\xi_n}{m_n}, \ldots, \frac{\xi_n}{m_n}}_{m_n\text{個}}\right) \quad \cdot (\sharp)$$

で定める．$\{\tilde{p}(\cdot;\vec{\xi}) \mid \vec{\xi} \in \Xi_n\}$ は明らかに Ω_m 上の n 次元統計的モデルであるから，

$$S_{\overrightarrow{m}} = \{\tilde{p}(\cdot;\vec{\xi}) \mid \vec{\xi} \in \Xi_n\} \qquad \cdots (*)$$

を示せば $S_{\overrightarrow{m}}$ が Ω_m 上の n 次元統計的モデルであることがいえる．

$(*)$ の $(\subseteq)$ の証明

$S_{\overrightarrow{m}}$ から元 f を勝手にとる．$S_{\overrightarrow{m}}$ の定義より，条件 $(\star)$ をみたす $\vec{\eta} \in \Xi_m$ が存在して $f = p(\cdot;\vec{\eta}) \in S_m$ である．ここで，

- $\xi_1 := \eta_1 + \cdots + \eta_{m_1}$
- $\vdots$
- $\xi_n := \eta_{m_1+\cdots+m_{n-1}+1} + \cdots + \eta_m$

とおく．$\forall j\ \xi_j > 0$，かつ，$\xi_1 + \cdots + \xi_n = \eta_1 + \cdots + \eta_m < 1$ $(\because \vec{\eta} \in \Xi_m)$ なので，$\vec{\xi} := (\xi_1, \ldots, \xi_n) \in \Xi_n$ である．条件 $(\star)$ より

- $\eta_1 = \cdots = \eta_{m_1} \quad \therefore \xi_1 = m_1\eta_1 = \cdots = m_1\eta_{m_1} \quad \therefore \eta_1 = \cdots = \eta_{m_1} = \frac{\xi_1}{m_1}$
- $\vdots$
- $\eta_{m_1+\cdots+m_{n-1}+1} = \cdots = \eta_m \quad \therefore \xi_n = m_n\eta_{m_1+\cdots+m_{n-1}+1} = \cdots = m_n\eta_m \quad \therefore \eta_{m_1+\cdots+m_{n-1}+1} = \cdots = \eta_m = \frac{\xi_n}{m_n}$

となり，

$$f = p(\cdot;\vec{\eta}) = p\left(\cdot;\underbrace{\frac{\xi_1}{m_1}, \ldots, \frac{\xi_1}{m_1}}_{m_1\text{個}}, \underbrace{\frac{\xi_2}{m_2}, \ldots, \frac{\xi_2}{m_2}}_{m_2\text{個}}, \ldots, \underbrace{\frac{\xi_n}{m_n}, \ldots, \frac{\xi_n}{m_n}}_{m_n\text{個}}\right) = \tilde{p}(\cdot;\vec{\xi})$$

となる．ある $\vec{\xi} \in \Xi_n$ を用いて $f = \tilde{p}(\cdot;\vec{\xi})$ と表せたので $f \in \{\tilde{p}(\cdot;\vec{\xi}) \mid \vec{\xi} \in \Xi_n\}$ である．

$(*)$ の $(\supseteq)$ の証明

$\{\tilde{p}(\cdot;\vec{\xi}) \mid \vec{\xi} \in \Xi_n\}$ から元 f を勝手にとる．ある $\vec{\xi} \in \Xi_n$ を用いて $f = \tilde{p}(\cdot;\vec{\xi})$ と表されている．まず，定義より

$$f = \tilde{p}(\cdot;\vec{\xi}) = p\left(\cdot;\underbrace{\frac{\xi_1}{m_1},\ldots,\frac{\xi_1}{m_1}}_{m_1\text{個}},\underbrace{\frac{\xi_2}{m_2},\ldots,\frac{\xi_2}{m_2}}_{m_2\text{個}},\ldots,\underbrace{\frac{\xi_n}{m_n},\ldots,\frac{\xi_n}{m_n}}_{m_n\text{個}}\right)$$

である．ここで

- $\eta_1 = \cdots = \eta_{m_1} = \frac{\xi_1}{m_1}$
- $\vdots$
- $\eta_{m_1+\cdots+m_{n-1}+1} = \cdots = \eta_m = \frac{\xi_n}{m_n}$

とおけば

$$f = \tilde{p}(\cdot;\vec{\xi}) = p(\cdot;\vec{\eta})$$

であり，$\forall i\ \eta_i > 0$，かつ，

$$\sum_{i=1}^{m}\eta_i = \underbrace{\frac{\xi_1}{m_1}+\cdots+\frac{\xi_1}{m_1}}_{m_1\text{個}} + \cdots + \underbrace{\frac{\xi_n}{m_n}+\cdots+\frac{\xi_n}{m_n}}_{m_n\text{個}} = \xi_1+\cdots+\xi_n < 1 \quad (\because \vec{\xi} \in \Xi_n)$$

なので $\vec{\eta} \in \Xi_m$ であり，$(\star)$ の条件も満たされる．よって $f \in S_{\vec{m}}$ である．

$F_{\vec{m}}$ の定義

写像 $F_{\vec{m}}\colon \Omega_m \to \Omega_n$ を

- $F_{\vec{m}}(0) := 0$
- $F_{\vec{m}}(i) := j \quad (\text{if } m_1+\cdots+m_{j-1}+1 \leq i \leq m_1+\cdots+m_j)$

で定める (テキストでは $j=1$ を分離しているが，$j=1$ のとき $m_1+\cdots+m_{j-1}=0$ と解釈すれば上二本の式で表せる)．全射性は明らかである．

$F_{\vec{m}}$ が $S_{\vec{m}}$ に関する十分統計量であることの証明

$r(\cdot;\vec{\xi}) := \frac{\tilde{p}(\cdot;\vec{\xi})}{q(F_{\vec{m}}(\cdot);\vec{\xi})}$ が $\vec{\xi}(\in \Xi_n)$ に依存しないことを示せばよい．

まず，$j = 0,1,\ldots,n$ に対して$q(j;\vec{\xi})$ を計算しておく．

- $j=0$ のとき．$q(0;\vec{\xi}) = \sum_{i\in F_{\vec{m}}^{-1}(\{0\})}\tilde{p}(i;\vec{\xi}) = \sum_{i\in\{0\}}\tilde{p}(i;\vec{\xi}) = \tilde{p}(0;\vec{\xi}) = 1-\sum_{i=1}^{n}\xi_i$
- $1 \leq j \leq n$ のとき．$q(j;\vec{\xi}) = \sum_{i\in F_{\vec{m}}^{-1}(\{j\})}\tilde{p}(i;\vec{\xi}) = \sum_{i\in\{\underbrace{m_1+\cdots+m_{j-1}+1,\cdots,m_1+\cdots+m_j}_{m_j\text{個}}\}}\tilde{p}(i;\vec{\xi}) = m_j \times \frac{\xi_j}{m_j} = \xi_j$

ここで $i \in \{m_1+\cdots+m_{j-1}+1,\cdots,m_1+\cdots+m_j\}$ に対し $\tilde{p}(i;\vec{\xi}) = \frac{\xi_j}{m_j}$ であることは $\tilde{p}$ の定義 $(\sharp)$ および p の定義より従う．

以上より，

- $r(0;\vec{\xi}) = \frac{\tilde{p}(0;\vec{\xi})}{q(F_{\vec{m}}(0);\vec{\xi})} = \frac{\tilde{p}(0;\vec{\xi})}{q(0;\vec{\xi})} = \frac{1-\left(\overbrace{\frac{\xi_1}{m_1}+\cdots+\frac{\xi_1}{m_1}}^{m_1\text{個}}+\cdots+\overbrace{\frac{\xi_n}{m_n}+\cdots+\frac{\xi_n}{m_n}}^{m_n\text{個}}\right)}{1-\sum_{j=1}^{n}\xi_j} = \frac{1-\sum_{j=1}^{n}\xi_j}{1-\sum_{j=1}^{n}\xi_j} = 1.$
- $m_1+\cdots+m_{j-1}+1 \leq i \leq m_1+\cdots+m_j$ のとき，$r(i;\vec{\xi}) = \frac{\tilde{p}(i;\vec{\xi})}{q(F_{\vec{m}}(i);\vec{\xi})} = \frac{\tilde{p}(i;\vec{\xi})}{q(j;\vec{\xi})} = \frac{\xi_j/m_j}{\xi_j} = \frac{1}{m_j}.$

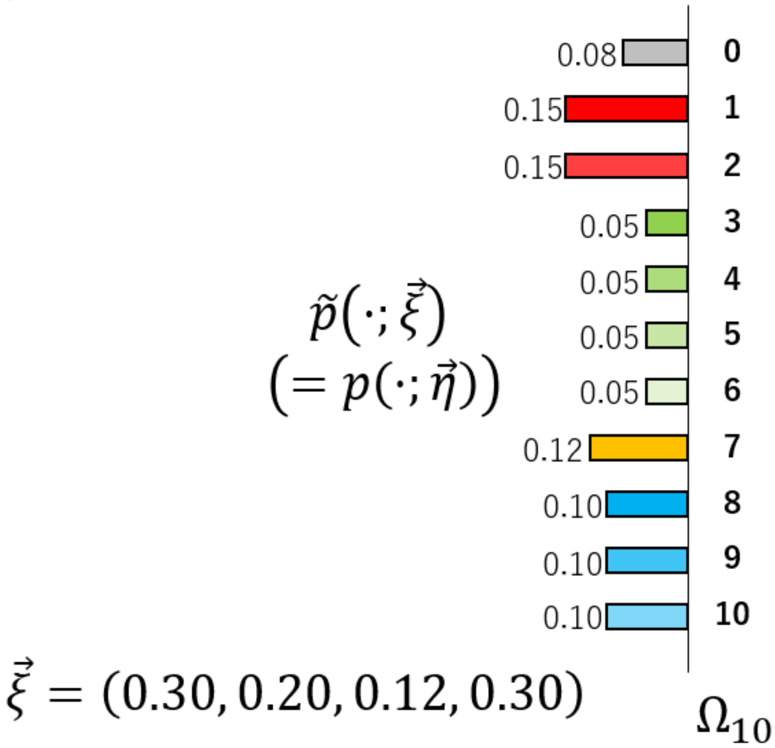
よって $r(\cdot;\vec{\xi})$ は $\vec{\xi}$ に依存せず，$F_{\vec{m}}$ は $S_{\vec{m}}$ に関する十分統計量である．

具体例

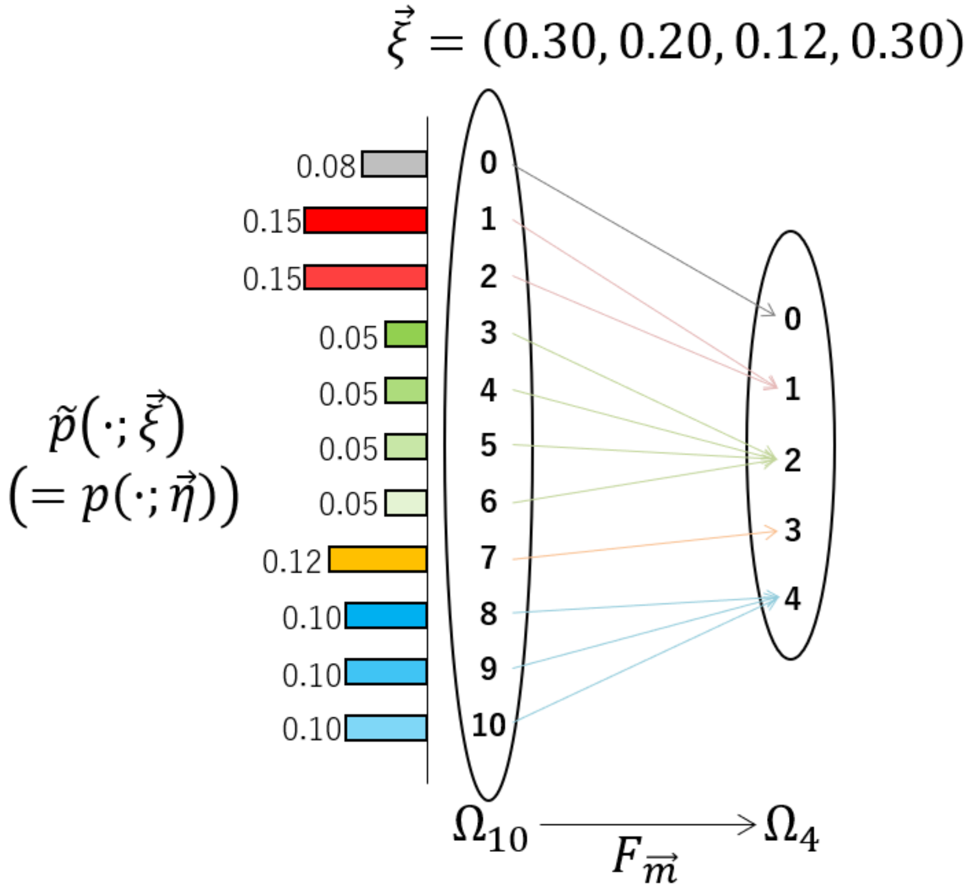
- $n = 4$
- $\vec{m} = (2, 4, 1, 3)$
    - $m_1 = 2$
    - $m_2 = 4$
    - $m_3 = 1$
    - $m_4 = 3$
    - $m = 2 + 4 + 1 + 3 = 10$

とする．

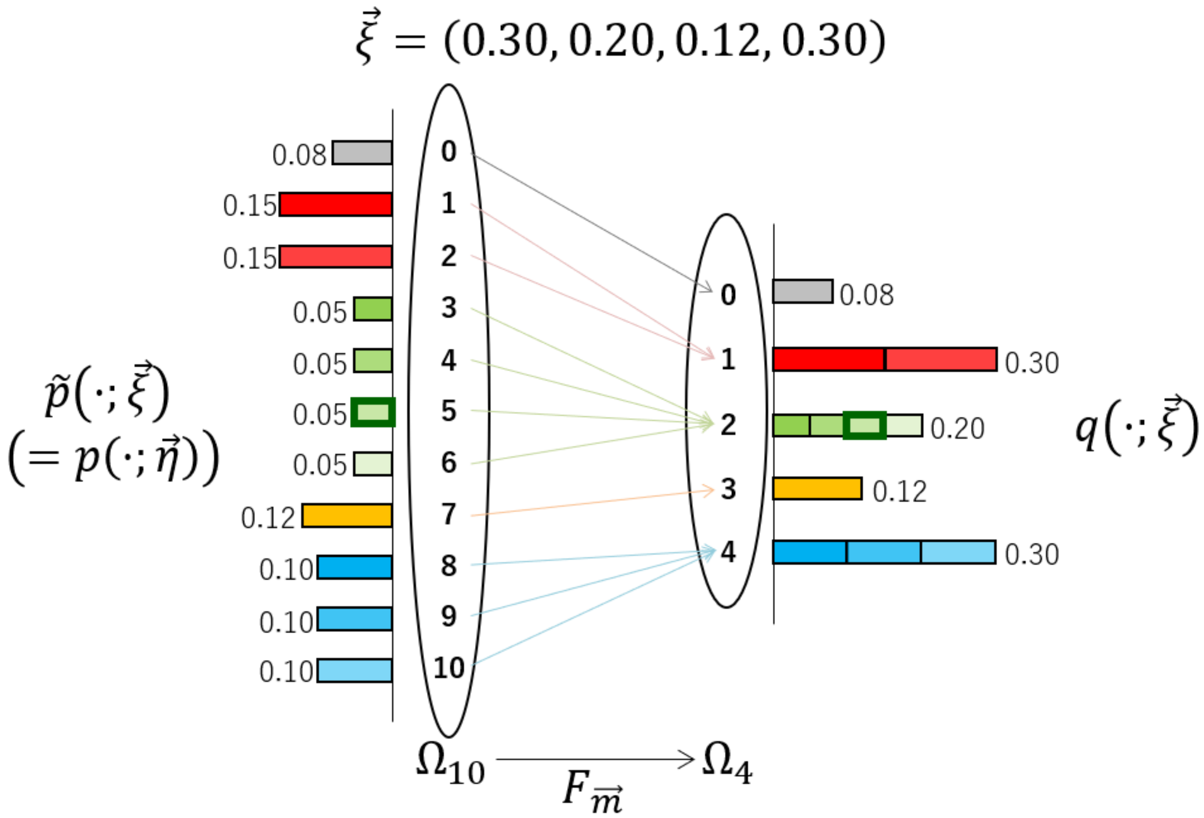
- $\vec{\eta} = (0.15, 0.15,\ \ 0.05, 0.05, 0.05, 0.05,\ \ 0.12,\ \ 0.10, 0.10, 0.10)$ とおく．
    - $\vec{\eta}$ の成分はすべて正で総和は $0.92\ (< 1)$ なので $\vec{\eta} \in \Xi_{10}$
    - さらに，
        - $\eta_1 = \eta_2$
        - $\eta_3 = \eta_4 = \eta_5 = \eta_6$
        - (η_7 に関しては条件なし)
        - $\eta_8 = \eta_9 = \eta_{10}$
    - よって $p(\cdot; \vec{\eta}) \in S_{\vec{m}}$ である．
- $\vec{\xi} = (0.30,\ 0.20,\ 0.12,\ 0.30)$ とおく．
    - $\vec{\xi}$ の成分はすべて正で総和は $0.92\ (< 1)$ なので $\vec{\xi} \in \Xi_4$
- $\tilde{p}$ の定めかた (式(1.132)) より
    - $\tilde{p}(\cdot; \vec{\xi}) = p(\cdot; \vec{\eta}) \in S_{\vec{m}}$
    - もっと明示的に書けば $\tilde{p}(\cdot; (0.30, 0.20, 0.12, 0.30)) = p(\cdot; (0.15, 0.15, 0.05, 0.05, 0.05, 0.05, 0.12, 0.10, 0.10, 0.10))$
- $\tilde{p}(\cdot; \vec{\xi})$ は次のような姿をした Ω_{10} 上の確率関数である．



- 式(1.134)(1.135) にしたがって写像 $F_{\vec{m}} \colon \Omega_{10} \to \Omega_4$ を求めると次の通りとなる．要するに，$F_{\vec{m}}$ は “同じグループ” をまとめるはたらきをもつ：



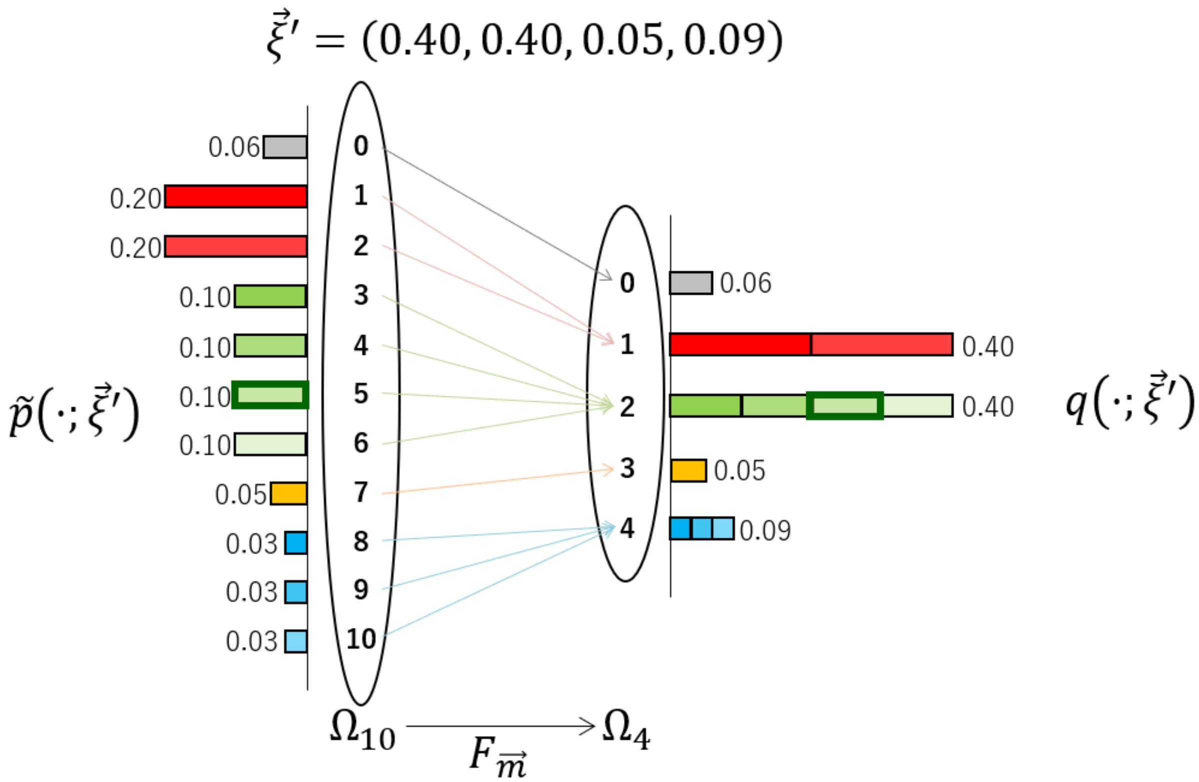
- $q(\cdot;\vec{\xi})$ は次のような姿となる：



- $r(5;\vec{\xi})$ を考えよう．図より $r(5;\vec{\xi}) = 1/4$ である．
- このようにして $r(\cdot;\vec{\xi})$ を計算すると，図から明らかなように次のようになる：

i	$r(i;\vec{\xi})$
0	1
1	$1/2$
2	$1/2$
3	$1/4$
4	$1/4$
5	$1/4$
6	$1/4$
7	1
8	$1/3$
9	$1/3$
10	$1/3$

- 別の $\vec{\xi}' \in \Xi_4$ を取っても $r(\cdot;\vec{\xi}')$ を計算すると必ず上の表と同等になるから，$r(\cdot;\vec{\xi})$ は $\vec{\xi}$ に依存しないことがわかる．したがって，$F_{\vec{m}}$ は $S_{\vec{m}}$ に関する十分統計量である．



- $F_{\vec{m}}$ は Ω_{10} に対する「良いグルーピング」「良い同一視」を与えている！