正方分割表における対称性の尺度と累積確率の対称性

山田　章史　(東京理科大学大学院・理工学研究科)
山本　紘司　(東京理科大学大学院・理工学研究科)
田畑　耕治　(東京理科大学・理工学部)
富澤　貞男　(東京理科大学・理工学部)

第 1 部：正方分割表における対称性の尺度

行と列が同じ分類からなる正方分割表において，Tomizawa, Seo and Yamamoto (1998) は対称モデル (Bowker, 1948) からの隔たりを測る尺度 $\phi^{(\lambda)}$ を提案した．尺度 $\phi^{(\lambda)}$ は，0 以上 1 以下の値をとり，(i) 任意の対称的な 2 つのセル確率のどちらか一方が 0 であるときに，対称モデルからの隔たりは最大（すなわち $\phi^{(\lambda)}=1$）であり，(ii) 対称モデルが成り立つときに最小（すなわち $\phi^{(\lambda)}=0$）になるという性質を持つ．しかし，実際のデータ解析では，多くの場合に各セル確率は 0 ではないため，尺度 $\phi^{(\lambda)}$ は 1 に到達出来ない．そこで，本報告では各セル確率が 0 でないときでも尺度の値が 1 に到達出来るような尺度を提案した．

正方 $R \times R$ 分割表において，(i, j) セル確率を p_{ij} とする $(i = 1, \ldots, R; j = 1, \ldots, R)$．対称モデルは次のように定義される (Bowker, 1948):

$$p_{ij} = p_{ji} \quad (i \neq j).$$

すべての $i \neq j$ に対して，$1 - d \leq p^c_{ij} \leq d$ を満たす定数 d $(0.5 < d \leq 1)$ を与えて，対称モデルからの隔たりを測る尺度を次のように導入した:

$$\Phi^{(\lambda)}(d) = \frac{1}{C^{(\lambda)}(d)}\phi^{(\lambda)} \qquad (\lambda > -1),$$

ただし，

$$C^{(\lambda)}(d) = 1 - \frac{\lambda 2^{\lambda}}{2^{\lambda} - 1}H^{(\lambda)}(d), \quad H^{(\lambda)}(d) = \frac{1}{\lambda}\left[1 - d^{\lambda+1} - (1-d)^{\lambda+1}\right].$$

ここに $\Phi^{(0)}(d) = \lim_{\lambda \to 0} \Phi^{(\lambda)}(d)$．尺度 $\Phi^{(\lambda)}(d)$ は $0 \leq \Phi^{(\lambda)}(d) \leq 1$ を満たし，任意の $\lambda > -1$ に対して，(i) 対称モデルが成り立つための必要十分条件は $\Phi^{(\lambda)}(d) = 0$ であり， (ii) 対称モデルからの隔たりが最大 [$i < j$ に対して，$p^c_{ij} = 1 - d$（このとき $p^c_{ji} = d$）または $p^c_{ji} = 1 - d$（このとき $p^c_{ij} = d$）のときと定義] であるための必要十分条件は $\Phi^{(\lambda)}(d) = 1$ である．

さらに，尺度の近似信頼区間や実データへの適用例を示し，提案した尺度の有用性を示した．

第 2 部：正方分割表における累積確率に基づく対称性に関するモデル

行と列が順序のある同じ分類からなる $R \times R$ 正方分割表において，(i, j) セル確率を p_{ij} とする $(i = 1, \ldots, R; j = 1, \ldots, R)$．累積確率 $\{G_{ij}\}$，$i \neq j$，を

$$G_{ij} = \sum_{s=1}^{i} \sum_{t=j}^{R} p_{st} \quad (i < j), \qquad G_{ij} = \sum_{s=i}^{R} \sum_{t=1}^{j} p_{st} \quad (i > j),$$

のように定義する．本報告では，対称的な累積確率の対数オッズが，分割表の主対角線からの距離に関する多項式で表されるモデル（C モデル）を次のように考えた：$0 \leq k_1 < k_2 < \cdots < k_m \leq R - 2$ $(1 \leq m \leq R - 1)$ に対して，

$$\log G_{ij} = \begin{cases} \Delta_{j-i} + \Psi_{ij} & (i < j), \\ \Psi_{ij} & (i > j), \end{cases} \qquad \log p_{ii} = \Psi_{ii},$$

ただし，

$$\Delta_{j-i} = \sum_{t=1}^{m} (j - i)^{k_t} \Theta_{k_t}, \quad \Psi_{ij} = \Psi_{ji}.$$

(Tomizawa, Miyamoto and Yamamoto, 2006; Yamamoto and Tomizawa, 2008)．

さらに，累積確率の和をとったものを定義し，これに基づいた新たなモデルを導入することで，C モデルに関する分解定理を得た．加えて，新たに提案したモデルやモデルの分解定理の有用性を，実データを用いた応用例とともに示した．

参考文献

Bowker, A. H. (1948). A test for symmetry in contingency tables. *Journal of the American Statistical Association*, **43**, 572-574.

Tomizawa, S., Miyamoto, N. and Yamamoto, K. (2006). Decomposition for polynomial cumulative symmetry model in square contingency tables with ordered categories. *Metron: International Journal of Statistics*, **64**, 303-314.

Tomizawa, S., Seo, T. and Yamamoto, H. (1998). Power-divergence-type measure of departure from symmetry for square contingency tables that have nominal categories. *Journal of Applied Statistics*, **25**, 387-398.

Yamamoto, K. and Tomizawa, S. (2008). A generalization of cumulative symmetry models and its decomposition for square contingency tables with ordered categories. *Far East Journal of Theoretical Statistics*, **24**, 201-211.