

二元分割表における周辺点対称性に関する方向付き尺度

東京理科大学・理工 生亀 清貴

順序カテゴリ二元 $R \times C$ 分割表において, (i, j) セル確率を p_{ij} とする ($i = 1, \dots, R; j = 1, \dots, C$). 周辺点対称モデルは次のように定義される (Tomizawa, 1985):

$$p_{i\cdot} = p_{i^*}, \quad (i = 1, \dots, R), \quad p_{\cdot j} = p_{\cdot j^{**}} \quad (j = 1, \dots, C),$$

ただし $p_{i\cdot} = \sum_{t=1}^C p_{it}$, $p_{\cdot j} = \sum_{s=1}^R p_{sj}$, $i^* = R+1-i$, $j^{**} = C+1-j$. このモデルは行周辺確率と列周辺確率が, 中心カテゴリまたは中心点に対してそれぞれ対称的であることを示している. 二元分割表において周辺点対称モデルが成り立たないとき, その隔たりがどの程度かを測ることに興味がある. Yamamoto et al. (2011) は Power-divergence (Cressie and Read, 1984) を用いた尺度を提案した. しかしながら, Yamamoto et al. (2011) の尺度では, $i = 1, \dots, \lfloor \frac{R}{2} \rfloor$ に対して行周辺確率 $p_{i\cdot}$ と p_{i^*} のどちらの方が大きいかを判別することができない (列周辺確率に関しても同様).

本講演では中心カテゴリ (中心点) に関する二種類の完全非対称を区別することができる, 周辺点対称性からの隔たりを測る尺度 φ_{MPS} を提案する. この尺度は, 行周辺確率が中心カテゴリ (中心点) に対して対称的かどうかを測る部分尺度と, 列周辺確率に関する部分尺度との重み付き和として表される. 提案尺度は $-1 \leq \varphi_{MPS} \leq 1$ を満たす. 周辺比例対称モデルを

$$p_{i\cdot} = \Delta_1 p_{i^*}, \quad \left(i = 1, \dots, \left\lfloor \frac{R}{2} \right\rfloor\right), \quad p_{\cdot j} = \Delta_2 p_{\cdot j^{**}} \quad \left(j = 1, \dots, \left\lfloor \frac{C}{2} \right\rfloor\right),$$

と定義すると, このモデルが成り立つ下で $\varphi_{MPS} = 0$ ならば, 周辺点対称モデルが成り立つ. $\varphi_{MPS} = 1$ となるための必要十分条件は, $i = 1, \dots, \lfloor \frac{R}{2} \rfloor$ に対して $p_{i\cdot} = 0$ ($p_{i^*} > 0$), かつ $j = 1, \dots, \lfloor \frac{C}{2} \rfloor$ に対して $p_{\cdot j} = 0$ ($p_{\cdot j^{**}} > 0$) である. また, $\varphi_{MPS} = -1$ となるための必要十分条件は, $i = 1, \dots, \lfloor \frac{R}{2} \rfloor$ に対して $p_{i^*} = 0$ ($p_{i\cdot} > 0$), かつ $j = 1, \dots, \lfloor \frac{C}{2} \rfloor$ に対して $p_{\cdot j^{**}} = 0$ ($p_{\cdot j} > 0$) である.

二元 $R \times C$ 分割表の (i, j) セル観測度数を n_{ij} とする ($i = 1, \dots, R; j = 1, \dots, C$). また $\{n_{ij}\}$ は多項分布に従うとし, 全標本数を n ($= \sum_{i=1}^R \sum_{j=1}^C n_{ij}$) とする. (i, j) セル確率 p_{ij} の推定値として $\hat{p}_{ij}$ ($= n_{ij}/n$) を用いた推定尺度を $\hat{\varphi}_{MPS}$ とおくと, デルタ法により $\sqrt{n}(\hat{\varphi}_{MPS} - \varphi_{MPS})$ は大標本の下で平均 0, 分散 $\sigma^2[\varphi_{MPS}]$ の正規分布に従う. 実際のデータ解析においては, $\sigma^2[\varphi_{MPS}]$ の $\{p_{ij}\}$ を $\{\hat{p}_{ij}\}$ に置き換えることにより得られる分散の推定値を用いて, φ_{MPS} の近似の信頼区間を構成する.

参考文献

- Cressie, N. and Read, T. R. C. (1984). *Journal of the Royal Statistical Society, Series B*, **46**, 440-464.
Tomizawa, S. (1985). *Biometrical Journal*, **27**, 895-905.
Yamamoto, K., Tahata, K., Suzuki, M. and Tomizawa, S. (2011). *Statistica*, **71**, 367-380.