隣接係数の概念

高木英至

文書作成日：2020 年 10 月 13 日(火曜)

この文書は高木の研究報告で用いた「隣接係数」を説明する．内容は文末引用文献の高木(2018)で示した内容と同じである．また，十分な定義をしていなかったが，実質的に同じ概念を次の高木(2003)でも使用している．以下の記述における数式の番号は，高木(2018)の記載のままとした(つまり[3]から始まる)．

複数の集団が空間内で混在する状況で集団間の近さの指標を定義することが今の課題である．もし各意見のセルのクラスタが空間内の１か所にまとまっていれば，クラスタの重心を求めてその重心間の距離を測ればよい．しかし社会的影響のモデルでは多くの場合，各意見のセルのクラスタは複数あって空間内に遍在する．つまり単に意見ごとの重心が近くても，意見クラスタ間の距離は遠いことも十分考えられる．

ここでは，意見間の空間的な近さの指標として，「意見 j に対する意見 i の隣接係数 $c_{i,j}$」を用いる．

最初に，「意見 j に対する意見 i の期待隣接数 $e_{i,j}$」を次のように定義する．

$$
\begin{aligned}
e_{ij} &= n_i m \frac{n_j}{N-1} = m \frac{n_i n_j}{N-1} \quad if \quad i \neq j \\
e_{ij} &= n_i m \frac{n_i - 1}{N-1} \quad\quad\quad\quad\quad if \quad i = j
\end{aligned}
\qquad [3]
$$

ここで n_i，n_j はそれぞれの意見のセル数を指す．セル総数が N である． m は周辺で数えるべきセルの数を指す．隣接を「距離１以内」とすれば，上下左右のセルが該当するから m=4 である．「距離２以内」とすれば m=12 である．

この期待隣接数 $e_{i,j}$ は次の考えに基づく．まず意見 i のあるセルの傍から m 個のセルの意見を調べるとしよう．このとき，セルの意見がランダムに配置されているとすれば，近くの１つのセルが意見 j である確率は $\frac{n_j}{N-1}$ となる．意見 j が i であれば，$\frac{n_i - 1}{N-1}$ である．意見 i のセル数は n_i であり，各セルについて m 個のセルを対象とする，確率的な意見 i 期待隣接数は[3]で表される．

[3]は次の[3’]のように書き換えることができる．[3’]の分数の分母は全体でのセル間の 2 者関係数を，分子は ij 間および i 内部での 2 者関係数を指す(Mayhew & Levinger, 1976)．したがってこの隣接期待数は，隣接可能性が 2 者関係数に比例することを意味している．

$$e_{ij} = Nm\frac{n_i n_j}{N(N-1)} \quad if \quad i \neq j$$
$$e_{ij} = Nm\frac{n_i(n_i-1)}{N-1} \quad if \quad i = j \qquad [3']$$

式[3]から次の関係も導ける．

$$e_{ij} = e_{ji}$$
$$\sum_i \sum_j e_{ij} = mN \qquad [4]$$

隣接期待数が対称的であることは，両集団の成員にとって相手集団の成員が同じ重みを持つことを意味する訳ではない．[3]から 1 人当たりの隣接数をとれば（e_{ij} / n_i），1 人当たりで近くにいる相手集団の人数は，相手集団の人数に比例することが分かる．つまり期待隣接数は同じでも，相対的に小さい集団の成員にとっては，相手集団と接する比率は高いのである（Mayhew & Levinger, 1976）．

同様に，実際に意見 i のセルのある距離内にいる意見 j のセル数（実際の隣接数）を f_{ij} とする．f_{ij} についても，定義上，次の関係があることが分かる．

$$f_{ij} = f_{ji}$$
$$\sum_i \sum_j f_{ij} = mN \qquad [5]$$

ここで，「意見 j に対する意見 i の隣接係数 c_{ij}」を次のように定義する．

$$c_{ij} = \frac{f_{ij}}{e_{ij}} \qquad [6]$$

隣接係数は，意見 i のセルの近くで意見 j のセルが見出せる程度が，偶然からの期待と比べて大きいか小さいかを表している．隣接係数が 1.0 より大きければ，偶然からの期待よりも i と j が近くにいる．1.0 より小さければ，意見 i のセルの周辺では意見 j のセルが少ない．

自己隣接を示す隣接係数 c_{ii} は，意見 i 集団の空間上の凝集性を表すといえる．

なお，隣接係数を c_{ij} ではなく対数変換した値（例えば $Ln(c_{ij})$）とすることも考えられる（高木(2003)では対数変換した値を隣接係数とした）．対数変換すれば，$f_{ij} = c_{ij}$ であればゼロ，$f_{ij} > c_{ij}$ であれば正の値，$f_{ij} > c_{ij}$ であれば負の値となり，その方が分かりやすい．しかし 2020 年の社会心理学会報告の分析では，$f_{ij} = 0$ となることがあり，その場合は $Ln(c_{ij})$ が定義されないため，対数変換しない値を隣接係数として用いている．

引用文献

Mayhew, B.H. & Levinger, R.L. (1976). Size and density of interaction in human aggregates. *American Journal of Sociology*, *82*, 86-110.

高木英至（2003）「エージェントのクラスタ化の相互調整モデル」,『埼玉大学紀要』, 第 39 巻，第 2 号，105-112.

高木英至（2018）「動的社会的影響モデルの拡張可能性」，『埼玉大学紀要（教養学部）』，第 53 巻，第 2 号，207-218.