階層一般化線形モデルの尤度方程式のラプラス近似について

広島大・理　上野　哲矢　 若木　宏文

社会心理学や医学などの分野において, データが階層的な構造を持つことが多い. 階層的な構造とは, 例えば, 被験者がある病院に入院している, といったように, 個人データが上位の集団に包含されている状況を指す. 階層一般化線形モデルは, このような階層的な構造を持つデータを解析することを目的としており, 各観測対象について, 目的変量と説明変量の観測値に対して一般化線形モデルを当てはめる. 本研究においては線形予測子の切片項をランダム係数にすることで, 集団間の相違を考慮したモデルと解釈することができる. 切片項にランダム係数を導入することで, 例えば, 複数個の病院の中から無作為に病院を抽出し, その病院に属する被験者が抱える病気の治癒するまでの期間の変動を, 各病院ごとに表現することが可能となる. 本発表では, 以上のような場面を想定し, 目的変量に生存時間解析でよく用いられる指数分布を仮定した場合について発表する.

ランダム係数 $\alpha_1, \cdots, \alpha_m$ は互いに独立な確率変数で $\alpha_i \sim N(0, 1/\tau)$ とし, 目的変数ベクトル $\boldsymbol{y}_i = (y_{i1}, \cdots, y_{in})'$, 説明変数行列 $\boldsymbol{X}_i = (\boldsymbol{x}_{i1}, \cdots, \boldsymbol{x}_{in})'$ の組 $(\boldsymbol{y}_i, \boldsymbol{X}_i)$ が与えられているとする. ここでは, 簡単のために τ は既知とする. また, $i = 1, \cdots, m,\ j = 1, \cdots, n$ とし, $\boldsymbol{x}_{ij}$ は p 次元説明変数ベクトルとする. $\alpha_i, \boldsymbol{X}_i$ を与えたもとで $y_{i1}, \cdots, y_{in}$ は互いに独立な確率変数で, $y_{ij} \mid \alpha_i, \boldsymbol{x}_{ij} \sim Ex[\exp^{-1}(\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \alpha_i)]$ とする. p 次元回帰係数ベクトル $\boldsymbol{\beta}$ の推定に関心があり, $\boldsymbol{\beta}$ の推定量を得るために最尤法を適用することが多い. したがって, 尤度関数を導出する必要があるが, $\boldsymbol{y} = (\boldsymbol{y}'_1, \cdots, \boldsymbol{y}'_m)'$ の尤度関数 $L(\boldsymbol{y}; \boldsymbol{\beta}, \tau \mid \boldsymbol{X}_1 \cdots, \boldsymbol{X}_m)$ は以下の式を扱う必要がある.

$$L(\boldsymbol{y}_i; \boldsymbol{\beta}, \tau \mid \boldsymbol{X}_i) = \frac{\tau^{m/2}}{(2\pi)^{m/2}} \prod_{i=1}^{m} \int \exp\Big[\sum_{j=1}^{n} \big\{\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \alpha_i - y_{ij}\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \alpha_i)\big\} - \frac{\tau\alpha_i^2}{2}\Big] d\alpha_i.$$

上の式から分かるように, ランダム係数に連続型分布を仮定した場合は積分を計算する必要があるが, 多くの場合, 解析的に求めることができない. 最尤推定量は, 反復法で求めることが多い. しかし, 尤度関数に積分を含むため, 繰り返しのたびに数値計算が必要となってしまう. 数値計算のための計算コストを減らすために, ガウスエルミート法などが用いられるが, 計算式が複雑であるため, 推定量の漸近的な性質を考察できない. これらの課題を改善する方法として, ラプラス近似を用いる手法が提案されている. $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_m)'$ とすると, ラプラス近似を用いた近似対数尤度関数 $l(\boldsymbol{\beta}, \tau; \boldsymbol{y}, \hat{\boldsymbol{\alpha}} \mid \boldsymbol{X}_1, \cdots, \boldsymbol{X}_m)$ は次のように表される.

$$l(\boldsymbol{\beta}, \tau; \boldsymbol{y}, \hat{\boldsymbol{\alpha}} \mid \boldsymbol{X}_1, \cdots, \boldsymbol{X}_m) = \sum_{i=1}^{m} \left\{ \lambda(\boldsymbol{y}_i, \hat{\alpha}_i; \boldsymbol{\beta}, \tau) + \frac{1}{2}\log\tau - \frac{1}{2}\log\left(- \left.\frac{\partial^2 \lambda(\boldsymbol{y}_i, \alpha_i; \boldsymbol{\beta}, \tau)}{\partial \alpha_i^2}\right|_{\alpha_i = \hat{\alpha}_i} \right) \right\},$$

$$\lambda(\boldsymbol{y}_i, \alpha_i; \boldsymbol{\beta}, \tau) = \sum_{j=1}^{n} \big\{\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \alpha_i - y_{ij}\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \alpha_i)\big\} - \frac{\tau\alpha_i^2}{2}.$$

ただし, $\partial\lambda/\partial\alpha_i = 0$ の解を $\hat{\alpha}_i$ とする. 本研究では n が十分大きいとき, ラプラス近似を用いて得られる近似対数尤度方程式の解として得られる推定量の漸近性質等について報告する.

参考文献

[1] Y. Lee and J. A. Nelder. (1996). Hierarchical Generalized Linear Models, *J. R. Statist. Soc. B,* **58**, 91-108