経時データ解析のための自己回帰線型混合効果モデル

船渡川　伊久子　　データ科学研究系　准教授

【はじめに】

経時データとは複数の対象者それぞれから複数の時点で反応を測定したデータです。例えば、臨床試験では、対象者を無作為に各治療群に割り付け、試験期間中の反応を複数時点観測し、群間で比較します。この時、通常、時点数はそれほど多くありません。反応が連続型の経時データ解析には線型混合効果モデルをよく用います。薬物動態の分野などでは特に、非線型混合効果モデルを用います。私達は経時データ解析のために自己回帰線型混合効果モデルを提案しています[1-5]。図1の左は線型混合効果モデルでしばしば使われる時間の一時式で、各対象者の切片と傾きが変量で、直線的な推移を表しています。図1の右は、自己回帰線型混合効果モデルで、各対象者の開始時とその後の切片が変量で、開始時からある均衡値（漸近値・定常状態での値）に向かう推移を表すことができます。

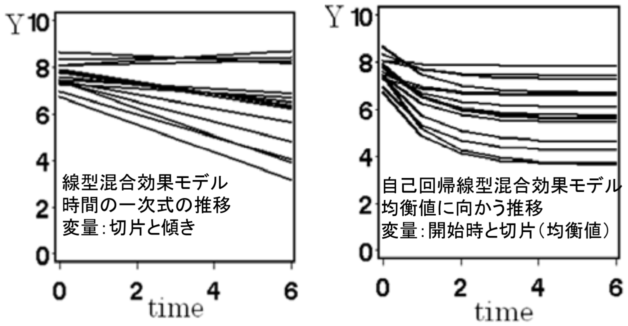

図1.線型混合効果モデルと自己回帰線型混合効果モデルの例

【自己回帰線型混合効果モデル：Autoregressive Linear Mixed Effects Model】

値を一時点前の値に回帰することを自己回帰といいます。反応に関する自己回帰と誤差に関する自己回帰があります。反応に関する自己回帰は医学系の経時データ解析ではほとんど使われていませんでした。自己回帰線型混合効果モデルの例を示します。Y_{it}を対象者i（i=1,…,N）の時点t（t=0,1, …,T_i）の反応とします。

$$\text{Baseline}(t=0): Y_{i,t} = \beta_{\text{base}} + b_{\text{base}\,i} + \varepsilon_{(\text{ME})i,t}$$

$$\text{Autoregressive}(t>0): Y_{i,t} = \rho Y_{i,t-1} + (\beta_{\text{int}} + b_{\text{int}\,i}) + \varepsilon_{(\text{AR})i,t} + \varepsilon_{(\text{ME})i,t} - \rho\varepsilon_{(\text{ME})i,t-1}$$

ここで、βは固定効果、b_iは変量効果、$\varepsilon_{i,t}$は誤差項です。変量効果、誤差項はそれぞれ多変量正規分布に従うと仮定します。反応の変化を表現する式に変形してみます。

$$\text{Change}(t>0): Y_{i,t} - Y_{i,t-1} = (1-\rho)\left\{(1-\rho)^{-1}(\beta_{\text{int}} + b_{\text{int}\,i}) - Y_{i,t-1}\right\} + \varepsilon_{(\text{AR})i,t} + \varepsilon_{(\text{ME})i,t} - \rho\varepsilon_{(\text{ME})i,t-1}$$

変化量は均衡値$(1-\rho)^{-1}(\beta_{\text{int}}-b_{\text{int}\,i})$までの距離に比例定数(1-ρ)で比例しています。これは、連続時間でのmonomolecular（the Mitcherlich）成長曲線に対応します。この例では共変量はありませんが、均衡値は共変量に依存します。右辺に反応変数のない周辺表現に変形します。

$$\text{Marginal}(t>0): Y_{i,t} = \rho^t(\beta_{\text{base}} + b_{\text{base}\,i}) + \sum_{j=1}^{t}\rho^{t-j}\left\{(\beta_{\text{int}} + b_{\text{int}\,i}) + \varepsilon_{(\text{AR})i,j}\right\} + \varepsilon_{(\text{ME})i,t}$$

$\varepsilon_{(\text{AR})}$ は1次の自己回帰、$\varepsilon_{(\text{ME})}$は時点間で独立な誤差であることがわかります。時間依存性共変量がある場合には時間がたつにつれ過去の共変量の影響が減衰することがわかります。図2に行列表現による自己回帰線型混合効果モデル、変化量への変形、周辺表現への変形、-2対数尤度を示します。行列表現により見通しが良くなります。対数尤度は周辺表現に対応しているため途中欠測があっても計算できます。

$$\mathbf{Y}_i = \rho\,\mathbf{F}_i\mathbf{Y}_i + \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i + \boldsymbol{\varepsilon}_i$$

$$Var(\mathbf{b}_i) = \mathbf{G},\quad Var(\boldsymbol{\varepsilon}_i) = \mathbf{R}_i \qquad \mathbf{F}_i = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$\mathbf{V}_i \equiv Var(\mathbf{Y}_i) = \mathbf{Z}_i\mathbf{G}\mathbf{Z}_i^T + \mathbf{R}_i$$

$$\mathbf{Y}_i - \mathbf{F}_i\mathbf{Y}_i = (1-\rho)\left\{(1-\rho)^{-1}\left(\mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i\right) - \mathbf{F}_i\mathbf{Y}_i\right\} + \boldsymbol{\varepsilon}_i$$

$$\mathbf{Y}_i = (\mathbf{I}_i - \rho\mathbf{F}_i)^{-1}\left(\mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i + \boldsymbol{\varepsilon}_i\right)$$

$$\boldsymbol{\Sigma}_i \equiv Var(\mathbf{Y}_i) = (\mathbf{I}_i - \rho\mathbf{F}_i)^{-1}\mathbf{V}_i\left\{(\mathbf{I}_i - \rho\mathbf{F}_i)^{-1}\right\}$$

$$-2ll_{ML} = K\log(2\pi) + \sum_{i=1}^{N}\log|\Sigma_i|$$

$$+\left\{\sum_{i=1}^{N}\left(\mathbf{Y}_i - (\mathbf{I}_i - \rho\mathbf{F}_i)^{-1}\mathbf{X}_i\boldsymbol{\beta}\right)^T\Sigma_i^{-1}\left(\mathbf{Y}_i - (\mathbf{I}_i - \rho\mathbf{F}_i)^{-1}\mathbf{X}_i\boldsymbol{\beta}\right)\right\}$$

図2. 自己回帰線型混合効果モデルの行列表現と対数尤度

【状態空間表現：State Space Representation】

Jones [6]は線型混合効果モデルの尤度を計算するために状態空間表現を用いました。状態空間表現を使うと計算に必要な行列のサイズは時点数に依らないため、小さくすみます。ただし、時点数が少ない時には通常の計算方法で問題ないでしょう。Jonesに倣い、自己回帰線型混合効果モデルの尤度を求める方法を提案しました[4]。状態空間表現は状態方程式と観測方程式からなります。詳細は省略しますが、前述の式を図3の状態空間で表します。この表現からは周辺尤度が計算されるので、途中欠測があっても尤度が計算できます。また、測定時間が人によって異なる場合でも、時間単位を十分短くし、欠測が沢山あるデータと捉えることで尤度が計算できます。式中のμ_{it}は潜在的な真の反応と解釈できます。

$$\text{State Eq.}: \mathbf{s}_{i(t)} = \boldsymbol{\Phi}_{i(t;t-1)}\mathbf{s}_{i(t-1)} + \mathbf{f}_{i,t}^{(s)} + \boldsymbol{\upsilon}$$

$$\begin{pmatrix} \mu_{i,t} \\ b_{\text{base}\,i} \\ b_{\text{int}\,i} \end{pmatrix} = \begin{pmatrix} \rho & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} \mu_{i,t-1} \\ b_{\text{base}\,i} \\ b_{\text{int}\,i} \end{pmatrix} + \begin{pmatrix} \mathbf{X}_{i,t}\boldsymbol{\beta} \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} \varepsilon_{(\text{AR})i,t} \\ 0 \\ 0 \end{pmatrix}$$

$$\text{Observation Eq.}: \mathbf{Y}_{i,t} = \mathbf{H}_{i,t}\mathbf{s}_{i(t)} + \mathbf{f}_{i,t}^{(o)} + \xi_{i,t}$$

$$Y_{i,t} = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} \mu_{i,t} \\ b_{\text{base}\,i} \\ b_{\text{int}\,i} \end{pmatrix} + \varepsilon_{(\text{ME})i,t}$$

図3. 自己回帰線型混合効果モデルの状態空間表現

【関連する話題】

自己回帰線型混合効果モデルと関連していくつかの研究を行っています。

- 多くの臨床研究では、複数の反応変数の推移を観測しています。2変量自己回帰線型混合効果モデルを提案しています[2]。相互に影響を与えながらある均衡値に向かう推移を表します。
- 臨床試験では、反応の値に応じて、しばしば対象者が途中脱落します。反応の値が悪い人ほど脱落すると、単純平均は実際よりも良くなるでしょう。欠測が観測された反応に依存していても、観測されていない反応に依存していなく、モデルが正しければ尤度に基づく方法は妥当であることが分かっています。実際のデータをよく表すモデルの開発が望まれます。自己回帰線型混合効果モデルは線型モデルでは用いられていなかった特徴的な分散共分散構造を与え、均衡値に向かう推移を表します[3]。
- 臨床研究では、反応に応じて投与量を変更する場合があります。投与量は時間依存性共変量です。前述のように自己回帰モデルは投与量の履歴を考慮した解析となっています（図4参照）。反応の値が悪い人ほど高用量を投与すると、高用量の単純平均は実際よりも悪くなるでしょう。欠測の問題に倣い、投与量が観測された反応に依存していても、観測されていない反応に依存していなく、モデルが正しければ尤度に基づく解析が妥当となります[5]。

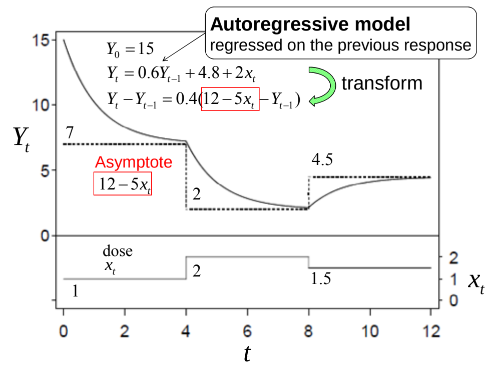

図4. 時間依存性共変量がある場合の自己回帰モデル

【参考文献】

[1] Funatogawa I et al. (2007) An autoregressive linear mixed effects model for the analysis of longitudinal data which show profiles approaching asymptotes. *Stat Med* **26**:2113-30.

[2] Funatogawa I et al. (2008) A bivariate autoregressive linear mixed effects model for the analysis of longitudinal data. *Stat Med* **27**:6367-78.

[3] Funatogawa T et al. (2008) An autoregressive linear mixed effects model for the analysis of longitudinal data which include dropouts and show profiles approaching asymptotes. *Stat Med* **27**:6351-66.

[4] Funatogawa I and Funatogawa T (2012) An autoregressive linear mixed effects model for the analysis of unequally spaced longitudinal data with dose-modification. *Stat Med* **31**, 589–599.

[5] Funatogawa I and Funatogawa T (2012) Dose-response relationship from longitudinal data with response-dependent dose-modification using likelihood methods. *Biometrical J* **54**, 494–506.

[6] Jones RH (1993) *Longitudinal Data with Serial Correlation: A State-space Approach.* Chapman & Hall.