政治学方法論 II

第 6 回：階層モデル

矢内　勇生

法学部・法学研究科

2015 年 5 月 30 日

神戸大学

今日の内容

1. 階層モデル入門
   - 階層モデル (hierarchical models)
2. 交換可能性
   - 交換可能性 (exchangeability)
   - 交換可能性について考える
   - 交換可能性とグループ

BDA3: pp.102– の例 (Table 5.1)

実験用のメスネズミ（F344）に腫瘍ができる確率 θ を推定したい。データを集めたところ、14 匹中 4 匹に腫瘍がみつかった。

- データ：$y = 4, n = 14$
- 尤度：$y \sim \mathrm{Bin}(n, \theta)$
- 事前分布：$\theta \sim \mathrm{Beta}(\alpha, \beta)$
- 事後分布：$\theta | y \sim \mathrm{Beta}(y + \alpha, n - y + \beta) = \mathrm{Beta}(\alpha + 4, \beta + 10)$

残された問題は？
α と β の値がわからない

事前分布の母数の設定

推定する母数 θ の母平均と母分散（母標準偏差）がわかるとき

- α, β が特定可能

推定する母数 θ の母平均と母分散（母標準偏差）がわからないとき

- 過去のデータを利用する
- 過去に、70 グループのメスネズミ (F344) のデータを集めた

$$y_i \overset{\text{iid}}{\sim} \text{Bin}(n_i, \theta_i)$$

- 過去のデータ：$i = 1, 2, \ldots, 70$
- 現在のデータ：$i = 71$

過去のデータを用いた事前分布の近似

- 過去のデータ ($i = 1, 2, \ldots, 70$) について

$$\mathrm{mean}\left(\frac{y_i}{n_i}\right) = 0.136, \quad \mathrm{sd}\left(\frac{y_i}{n_i}\right) = 0.103$$

- これを満たすのは、

$$\alpha \approx 1.4, \quad \beta \approx 8.6$$

- したがって、

$$\theta_{71} | y_{71} \sim \mathrm{Beta}(5.4, 18.6),$$

$$\mathrm{E}(\theta_{71} | y_{71}) = 0.225, \quad \mathrm{sd}(\theta_{71} | y_{71}) \approx 0.084$$

過去のデータを生み出した母数の推定

- 依拠している前提：$\theta_i \overset{\text{iid}}{\sim} \text{Beta}(\alpha, \beta)$
- つまり、θ_i の事前分布は同じ
- 同じ事前分布を使って、$\theta_1, \theta_2, \dots, \theta_{70}$ を推定できる？
- 問題：あるデータ y_i を事前分布のパラメタの近似と母数の推定に使う（つまり、二度使う）ことは許されるか？
- $\theta_1, \dots, \theta_{71}$ を別々に推定する？

仮定

- 実験（試行） $j = 1, 2, \ldots, J$
- 実験 j：データ y_j, 母数 θ_j, 尤度 $p(y_j|\theta_j)$
- 母数の一部には重複があってもよい
- 重複する母数の例：$\theta_j = (\mu_j, \sigma^2)$

交換可能性 (exchangeability)

交換可能性

同時確率 $p(\theta_1, \ldots, \theta_J)$ が、インデクス $(1, \ldots, J)$ をどのように入れ替えても変化しないとき、母数 $(\theta_1, \ldots, \theta_J)$ はその同時確率において交換可能である。

- データ y 以外に θ_j を区別する情報がない
- 事前分布においては、どの θ_j も区別せずに扱うべき
- 情報が少ないほど、交換可能性は成り立ちやすい

単純な交換可能性の例

- θ_j：母数 ϕ をもつある事前分布から独立に抽出される
- θ_j の同時事前分布

$$p(\theta_1, \ldots, \theta_J|\phi) = p(\theta|\phi) = \prod_{j=1}^{J} p(\theta_j|\phi)$$

- 通常、ϕ は未知 $\to$ ϕ を消去する

$$p(\theta) = \int p(\theta|\phi)p(\phi)d\phi = \int \left(\prod_{j=1}^{J} p(\theta_j|\phi) \right) p(\phi)d\phi$$

離婚データ：状況 1

- YY が 47 都道府県から 7 つを選び、2014 年の 1000 人あたり離婚数を調べる
- データ：$y_1, \dots, y_7$
- y_7 は？
- y_j を区別する方法がない $\rightarrow$ 交換可能なものとして扱う

離婚データ：状況 2

- 7 つのうちから 6 つをランダムに選ぶ
- 選ばれたデータ：$1.83, 1.67, 1.65, 1.78, 1.82, 1.79$
- y_7 は？
- 6 つの観測値に基づく y_7 の事後予測：平均 1.76, おおよそ 1.6 〜 2 程度
- インデクスを付け替えても、予測は変わらない
- 各観測値は都道府県で、基礎となる離婚率は同じ分布から生じていると考えられる
  - y_j は交換可能
  - y_j は独立ではない

離婚データ：状況 3

- あらかじめ、7 つの都道府県は関西とその近郊県であることを教える
- 兵庫、大阪、京都、滋賀、奈良、岡山、三重
- ただし、順番はランダムで、どのインデクスがどの県に対応するかは不明
- 観測値を手に入れる前：y_j は交換可能
- ただし、事前分布は変わる
  - 大阪は大都市なので離婚率が他より高そう
  - 奈良は伝統を守るので、離婚率が低そう
- 事前分布の分散を大きくする必要
- 選ばれたデータ：$1.83, 1.67, 1.65, 1.78, 1.82, 1.79$
- 特に他とかけ離れた値はない：y_7 は大阪または奈良？
- y_7 の事後予測分布：データの分布よりも大きいまたは小さい値を予測

離婚データ：状況 4

- y_7 は大阪であることを教える
- y_7 を他の観測値と区別できるので、y_j 全体は交換可能ではない
- $y_7 > 1.83$ となる確率が高いと予測する
- 実際の $y_8 = 2.08$

部分的な交換可能性 (1)

グループ分けによる階層モデル

- 観測値をグループに分けることが可能
- 各グループに別の確率モデルを適用
- グループごとの特性が不明
- グループの特性を交換可能なものとして扱い、グループの特性に対して共通の事前分布を置く

例

- 2 つの異なる研究室からのメスネズミのデータ
- どちらの研究室から得られたがわかれば、データを区別できる
- 研究室の特性についての情報がないので、研究室ごとにグループ分けし、共通の事前分布を使う

部分的な交換可能性 (2)

追加情報がある場合の交換可能性

- y_i が追加情報 x_i を伴っている
- y_i は交換可能ではない
- (y_i, x_i) は交換可能
- (y_i, x_i) の同時分布 または条件付きモデル $y_i|x_i$ を考える

例

- 追加情報として、2013 年の離婚データ x_j が手元にある
- y_j は交換可能ではない：x_j の値によって、y_j を区別できる
- (y_j, x_j) は交換可能：x_j の値が全く同じ 2 県は区別できない
- 条件付き交換可能性を利用：x_j を共変量として利用する

条件付き交換可能性

- 条件付き交換可能性を使うのが普通
- 説明変数 x_j で条件付けを行う

$$p(\theta_1,\ldots,\theta_J|x_1,\ldots,x_J) = \int \left[\prod_{j=1}^{J} p(\theta_j|\phi,x_j)\right] p(\phi|x)d(\phi)$$

階層ベイズモデル

- 母数 (parameter) θ を推定したい
- θ の事前分布は母数 (hyperparameter) ϕ を持つ
- ϕ も未知
- θ と ϕ の同時事後分布を推定する！
- θ と ϕ の同時事前分布

$$p(\phi, \theta) = p(\phi)p(\theta|\phi)$$

- θ と ϕ の同時事後分布

$$\begin{aligned} p(\phi, \theta|y) &\propto p(\phi, \theta)p(y|\phi, \theta) \\ &= p(\phi, \theta)p(y|\theta) \\ &= p(\phi)p(\theta|\phi)p(y|\theta) \end{aligned}$$

2 種類の事後予測分布

既存の θ_j に対応する将来の観測値 $\tilde{y}_j$

- 現在の実験で、メスネズミを新たに 1 匹観測したときの観測値の予測
- 得られた θ_j の事後分布から、無作為に y を抽出

将来の母数 $\tilde{\theta}$ に対応する $\tilde{y}$

- 将来新たに実験を行ったときに得られる観測値の予測
- まず、ϕ の事後分布 を利用し、$\tilde{\theta}$ を抽出
- 抽出された $\tilde{\theta}$ を利用し、$\tilde{y}$ を抽出