政治学方法論 II

第 9 回：階層モデル（続）

矢内　勇生

法学部・法学研究科

2015 年 6 月 10 日

神戸大学

今日の内容

1. 8 つの学校での並列実験
   - イントロダクション
   - 階層モデルを使わない推定
   - 階層モデルを使った推定
2. Stan によるベイズ推定
   - Stan と RStan
   - 8 校の並列実験、メタ分析、その他

8 つの学校での並列実験

実験の目的

特別指導が SAT-V のスコアに与える影響を調べたい

- 8 つの高校 $(j = 1, \ldots, 8)$
- 各校の生徒を統制群（特別指導なし）と処置群（特別指導あり）に無作為に分ける
- 処置前に模試の点数を測る（PSAT-V と PSAT-M）
- 学校ごとに、SAT-V の点数を PSAT-V と PSAT-M に回帰し、特別指導の平均処置効果 (y_j) とその標準誤差 (σ_j) を求める

問題の設定

- 8 つの独立した高校; $J = 8$
- 高校 j における平均処置効果 y_j から、θ_j を推定する
- y_j の標準誤差 σ_j は既知（OLS により計算済み）

学校	A	B	C	D	E	F	G	H
y_j	28	8	-3	7	-1	1	18	12
σ_j	15	10	16	11	9	11	10	18

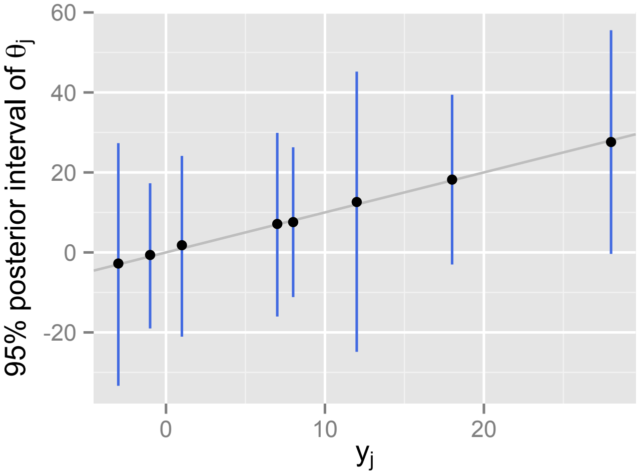
- θ_j の尤度（y_j のサンプリングモデル）：

$$y_j|\theta_j \sim \mathrm{N}(\theta_j, \sigma_j^2)$$

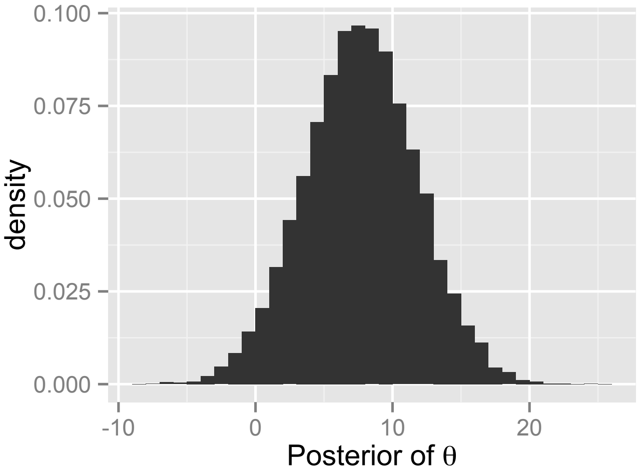
- 母数分布（事前分布）

$$\theta_j \overset{\text{iid}}{\sim} \mathrm{N}(\mu, \tau^2)$$

Separating (no pooling)



Complete pooling



階層モデル

- 尤度（サンプリングモデル）：

$$y_j|\theta_j \sim \mathrm{N}(\theta_j, \sigma_j^2)$$

- 事前分布：

$$\theta_j|\mu, \tau \sim \mathrm{N}(\mu, \tau^2)$$

- 無情報超事前分布

$$p(\mu|\tau) \propto 1$$

かつ

$$p(\tau) \propto 1$$

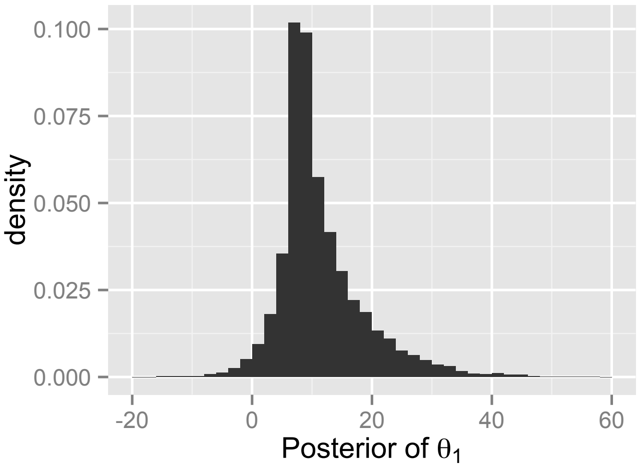
事後分布

- 同時事後分布

$$p(\theta,\mu,\tau|y) = p(\tau|y)p(\mu|\tau,y)(\theta|\mu,\tau,y)$$

- したがって、以下の順番でシミュレーションを行う
  1. $\tau|y$ を抽出する
  2. $\mu|\tau,y$ を抽出する
  3. $\theta_j|\mu,\tau,y$ $(j=1,\dots,8)$ を抽出する
  4. 1 から 3 を L 回繰り返し、事後分布を得る

θ_1（A 校の平均処置効果）の事後分布



Stan とは？

Stan

- ベイズ推定を実行するプログラム
- MCMC によるシミュレーション（サンプリング）：NUTS, HMC
- C++：Linux, Mac, Windows 上で動く
- Free: オープンソース、無料
- RStan：Stan の R 用インターフェース

RStan のインストールと使用

- インストールについてはココ を参照
- 使い方：read the manual

RStan を使った分析例

授業のウェブサイトを参照

階層モデルと Stan によるベイズ推定