# 政治学方法論 II

## 第 14 回：回帰モデル

矢内　勇生

法学部・法学研究科

2015 年 7 月 22 日

神戸大学

# 今日の内容

1. 回帰モデル
   - 回帰モデルの基礎
   - 古典的回帰モデルのベイズ分析

# 回帰モデルを作る意味

- ある変数 $y$ が他の変数 $x$ の関数としてどのように変化するか？
- これまで対象としてきたデータの分布：$p(y|\theta)$
- 回帰モデルが対象とするデータの分布：$p(y|\theta,x)$
- 仮定：$n$ 個の観測値 $(x,y)_i$ は交換可能

# 記号

- $y$：応答変数または結果変数
- $x = (x_1, \ldots, x_k)$：説明変数
- 観測単位 $i = 1, \ldots, n$
- $y$ は $n$ 次の列ベクトル
- $X$ は $n \times k$ 行列（計画行列 [design matrix]）：各列が 1 つの説明変数

## Normal linear モデル

- $X$ が所与のときの $y$ の分布が正規分布 (**normal**) で、正規分布の平均が $X$ の線形 (**linear**) 関数：

$$\mathrm{E}(y_i|\beta,X) = \beta_1 x_{i1} + \cdots + \beta_k x_{ik}$$

for $i = 1,\ldots,n$.
- 多くの場合、$x_{i1} = 1$ for all $i$
- 推定する母数：$\theta = (\beta_1,\ldots,\beta_k,\sigma)$
- 単純化するため、当面以下を仮定する
  - $\mathrm{var}(y_i|\theta,X) = \sigma^2$ for all $i$
  - $y_i$ は $\theta,X$ が与えられた条件の下で独立

# Normal linear モデルによる分析の概要

- 変数 $x$ と $y$ を定義する（必要に応じて変数変換する）
- 推定する母数に対する事前分布を設定する
  - 手元にあるデータを用いて推定ができる程度に強い分布
  - 事前情報がデータを支配してしまわない程度に弱い分布
- $p(\theta|y,X)$ を求める！

# 条件付きモデルの正当化

- $X$ もデータの一部：$X$ の尤度 $p(X|\psi)$
- 同時尤度 $p(X,y|\psi,\theta)$ と同時事前分布 $p(\psi,\theta)$ を考える必要
- 通常の回帰分析：$p(\psi,\theta)=p(\psi)p(\theta)$ を仮定
- 同時事後分布を分解すると

$$p(\psi,\theta|X,y)=p(\psi|X)p(\theta|X,y)$$

- 後の因数に注目すると、

$$p(\theta|X,y)\propto p(\theta)p(y|X,\theta)$$

- $k+1$ 個の変数の同時分布ではなく、$k$ 個の変数（説明変数）を所与として 1 つの変数（応答変数）の条件付き分布を分析できる

# 基本モデルと無情報事前分布

- 基本モデル：

$$y|\beta,\sigma,X \sim \mathrm{N}(X\beta,\sigma^2 I)$$

  - 観測上の誤差が独立
  - 分散は一定
  - $I$ は $n$ 次の単位行列
- 無情報事前分布：$(\beta,\log\sigma)$ 上の一様分布を仮定

$$p(\beta,\sigma^2|X) \propto \sigma^{-2}$$

  - 推定する母数の数に対して観測単位の数 $n$ が十分大きい：無情報事前分布を使って推定可能
  - 推定する母数が多い or $n$ が小さい：情報を有する事前分布を使う or 階層モデルを利用する

# 事後分布

- 求める事後分布：$p(\beta, \sigma^2|y, X)$
- 分解すると

$$p(\beta, \sigma^2|y, X) = p(\beta|\sigma^2, y, X)p(\sigma^2|y, X)$$

- 記法を簡略化するため、$X$ を省略して書くと、

$$p(\beta, \sigma^2|y) = p(\beta|\sigma^2, y)p(\sigma^2|y)$$

# $\beta$ の条件付き事後分布と $\sigma^2$ の周辺事後分布

- $\beta$ の条件付き事後分布は正規分布：

$$\beta|\sigma, y \sim \mathrm{N}(\hat{\beta}, V_\beta \sigma^2)$$

$$\hat{\beta} = (X^T X)^{-1} X^T y$$

$$V_\beta = (X^T X)^{-1}$$

- $\sigma^2$ の周辺事後分布：

$$p(\sigma^2|y) = \frac{p(\beta, \sigma^2|y)}{p(\beta|\sigma^2, y)}$$

$$\sigma^2|y \sim \text{Inv-}\chi^2(n-k, s^2)$$

$$s^2 = \frac{1}{n-k}(y - X\hat{\beta})^T (y - X\hat{\beta})$$