Lecture 7 Linear Regression and Extensions

Fei Tan

Department of Economics
Chaifetz School of Business
Saint Louis University

E6930 Applied Bayesian Statistics

January 7, 2023

Extending Linear Regressions (LR)

General setup

$$
\begin{aligned}
y_i^* &= x_i'\beta + u_i, \quad u_i|x_i \sim_{i.i.d.} t_\nu(0, \sigma^2) \\
\mathbb{E}(y_i) &= G(x_i'\beta), \qquad i = 1, \ldots, n
\end{aligned}
$$

- Choice of link function $G(\cdot)$
  - standard LR

    $$G(x_i'\beta) = x_i'\beta \quad \Rightarrow \quad y_i = y_i^*$$

  - tobit censored LR ($\nu = \infty$; $\mathbb{N}(0,1)$ p.d.f. ϕ, c.d.f. Φ)

    $$G(x_i'\beta) = x_i'\beta + \frac{\phi(-x_i'\beta/\sigma)}{1 - \Phi(-x_i'\beta/\sigma)}\sigma \quad \Rightarrow \quad y_i = \max\{y_i^*, 0\}$$

  - binary probit LR ($\nu = \infty$; $\sigma^2 = 1$)

    $$G(x_i'\beta) = \Phi(x_i'\beta) \quad \Rightarrow \quad y_i = 1\{y_i^* > 0\}$$

The Road Ahead...

- LR with Gaussian errors
- LR with Student-t errors
- Marginal likelihood
- Tobit censored LR
- Binary probit LR

LR with Gaussian Errors

- Conditionally conjugate prior

$$\beta \sim \mathbb{N}(\beta_0, B_0), \qquad \sigma^2 \sim \mathbb{IG}\text{-}2(\alpha_0/2, \delta_0/2)$$

- Gibbs algorithm
  - step 1: choose $\beta = \beta^{(0)}$, $\sigma^2 = \sigma^{2(0)}$, set $g = 0$
  - step 2: sample recursively

    $$\beta^{(g)} \sim \mathbb{N}(\beta_1^{(g+1)}, B_1^{(g+1)}), \qquad \sigma^{2(g+1)} \sim \mathbb{IG}\text{-}2(\alpha_1/2, \delta_1^{(g+1)}/2)$$

    where

    $$\begin{aligned} B_1^{(g+1)} &= (\sigma^{-2(g)} X'X + B_0^{-1})^{-1} \\ \beta_1^{(g+1)} &= B_1^{(g+1)}(\sigma^{-2(g)} X'y + B_0^{-1}\beta_0) \\ \alpha_1 &= \alpha_0 + n \\ \delta_1^{(g+1)} &= \delta_0 + (y - X\beta^{(g+1)})'(y - X\beta^{(g+1)}) \end{aligned}$$

  - step 3: set $g = g + 1$ and go to step 2

LR with Student-t Errors

- Conditional likelihood

$$f(y_i|\beta,\sigma^2,\lambda_i) = \mathbb{N}(x_i'\beta, \lambda_i^{-1}\sigma^2), \quad \lambda_i \sim \mathbb{G}(\nu/2,\nu/2) \text{ (latent)}$$

- Gibbs sampler for $\pi(\beta,\sigma^2,\lambda|y)$

$$\begin{aligned}
\beta|y,\lambda,\sigma^2 &\sim \mathbb{N}(\beta_1, B_1) \\
\sigma^2|y,\beta,\lambda &\sim \mathbb{IG}\text{-}2(\alpha_1/2, \delta_1/2) \\
\lambda_i|y,\beta,\sigma^2 &\sim \mathbb{G}(\nu_1/2, \nu_{2i}/2), \qquad i=1,\ldots,n
\end{aligned}$$

where $\Lambda = \text{diag}(\lambda_i)$ and

$$\begin{aligned}
B_1 &= (\sigma^{-2}X'\Lambda X + B_0^{-1})^{-1} \\
\beta_1 &= B_1(\sigma^{-2}X'\Lambda y + B_0^{-1}\beta_0) \\
\alpha_1 &= \alpha_0 + n \\
\delta_1 &= \delta_0 + (y - X\beta)'\Lambda(y - X\beta) \\
\nu_1 &= \nu + 1 \\
\nu_{2i} &= \nu + \sigma^{-2}(y_i - x_i'\beta)^2
\end{aligned}$$

Python Pseudo Code

```
To be added...
```

Marginal Likelihood

Chib method

$$m(y) = \frac{\prod_{i=1}^{n} t_\nu(x_i'\beta^*, \sigma^{2*})\pi(\beta^*)\pi(\sigma^{2*})}{\pi(\beta^*, \sigma^{2*}|y)}, \quad \forall\theta^* = (\beta^*, \sigma^{2*}) \in \Theta$$

- Compute $\pi(\beta^*, \sigma^{2*}|y)$ (not involving λ) at high-density point θ^* from Gibbs output

$$\pi(\beta^*, \sigma^{2*}|y) = \pi(\beta^*|y)\pi(\sigma^{2*}|\beta^*, y)$$

  - full run: $\hat{\pi}(\beta^*|y) = \frac{1}{G}\sum_{g=1}^{G} \pi(\beta^*|\sigma^{2(g)}, \lambda^{(g)}, y)$, where $(\theta^{(g)}, \lambda^{(g)}) \sim \pi(\theta, \lambda|y)$
  - reduced run: $\hat{\pi}(\sigma^{2*}|\beta^*, y) = \frac{1}{G}\sum_{g=1}^{G} \pi(\sigma^{2*}|\beta^*, \lambda^{(g)}, y)$, where $(\sigma^{2(g)}, \lambda^{(g)}) \sim \pi(\sigma^2, \lambda|\beta^*, y)$

Tobit Censored LR

Model

$$
\begin{aligned}
y_i^* &= x_i'\beta + u_i, \qquad u_i|x_i \sim_{i.i.d.} \mathbb{N}(0,\sigma^2) \\
y_i &= \max\{y_i^*, 0\}, \qquad i = 1,\ldots,n
\end{aligned}
$$

- Chib (1992) introduces latent variables z for censored observations and Gibbs sampler for $\pi(\beta,\sigma^2,z|y)$
  - conditionally conjugate prior for (β,σ^2) as before
  - sample $\beta|y_z,\sigma^2 \sim \mathbb{N}(\beta_1,B_1)$, where y_z replaces $y_i = 0$ by $z_i < 0$
  - sample $\sigma^2|y_z,\beta \sim \mathbb{IG}\text{-}2(\alpha_1/2,\delta_1/2)$
  - sample $z_i|y,\beta,\sigma^2 \sim \mathbb{TN}_{(-\infty,0)}(x_i'\beta,\sigma^2)$ (truncated normal)
  - exercise: Student-t version
- Data augmentation technique [Tanner & Wong (1987)]

Python Pseudo Code

```
To be added...
```

Binary Probit LR

Model

$$\begin{aligned} y_i^* &= x_i'\beta + u_i, \qquad u_i|x_i \sim_{i.i.d.} \mathbb{N}(0,1) \\ y_i &= 1\{y_i^* > 0\}, \qquad i = 1, \ldots, n \end{aligned}$$

- Albert and Chib (1993) introduces latent variables $z = y^*$ and Gibbs sampler for $\pi(\beta, z|y)$
  - $\beta \sim \mathbb{N}(\beta_0, B_0)$ as before; $\sigma^2 = 1$ for identification
  - sample $\beta|z \sim \mathbb{N}(\beta_1, B_1)$ (B_1 not updated)
  - sample $z_i|y, \beta \sim \mathbb{TN}_{(-\infty,0]}(x_i'\beta, 1)$ if $y_i = 0$ or $\mathbb{TN}_{(0,\infty)}(x_i'\beta, 1)$ if $y_i = 1$
  - exercise: Student-t version
- Binary logit LR: $u_i|x_i \sim_{i.i.d.} \mathbb{L}(0,1)$ (logistic distribution)

Python Pseudo Code

```
To be added...
```

Readings

- Albert & Chib (1993), "Bayesian Analysis of Binary and Polychotomous Response Data," *Journal of the American Statistical Association*
- Chib (1992), "Bayes Inference in the Tobit Censored Regression Model," *Journal of Econometrics*
- Tanner & Wong (1987), "The Calculation of Posterior Distributions by Data Augmentation (with Discussion)," *Journal of the American Statistical Association*