

PROBIT REGRESSION MODEL FOR BINARY DATA

Brian M. Brost

10 August 2015

Description

A regression model for binary data using the probit link.

Implementation

The file `probit.reg.sim.R` simulates data according to the model statement presented below, and `probit.reg.mcmc.R` contains the MCMC algorithm for model fitting.

Model statement

Let y_t , for $t = 1, \dots, T$, be observed data that take on the values $\{0,1\}$. Also let \mathbf{X} be a design matrix containing covariates for which inference is desired, and the vector $\boldsymbol{\beta}$ be the corresponding coefficients.

$$\begin{aligned} y_t &\sim \begin{cases} 0, & u_t \leq 0 \\ 1, & u_t > 0 \end{cases} \\ u_t &\sim \mathcal{N}(\mathbf{x}'_t \boldsymbol{\beta}, 1) \\ \boldsymbol{\beta} &\sim \mathcal{N}(\mathbf{0}, \sigma_{\boldsymbol{\beta}}^2 \mathbf{I}) \end{aligned}$$

Full conditional distributions

Observation model auxiliary variable (u_t):

$$\begin{aligned} [u_t | \cdot] &\propto [y_t | u_t][u_t] \\ &\propto (1_{\{y_t=0\}} 1_{\{u_t \leq 0\}} + 1_{\{y_t=1\}} 1_{\{u_t > 0\}}) \times \mathcal{N}(u_t | \mathbf{x}'_t \boldsymbol{\beta}, 1) \\ &= \begin{cases} \mathcal{TN}(\mathbf{x}'_t \boldsymbol{\beta}, 1)_{-\infty}^0, & y_t = 0 \\ \mathcal{TN}(\mathbf{x}'_t \boldsymbol{\beta}, 1)_0^{\infty}, & y_t = 1 \end{cases} \end{aligned}$$

Regression coefficients ($\boldsymbol{\beta}$):

$$\begin{aligned} [\boldsymbol{\beta} | \cdot] &\propto [\mathbf{u} | \boldsymbol{\beta}, \sigma^2][\boldsymbol{\beta}] \\ &\propto \mathcal{N}(\mathbf{u} | \mathbf{X}\boldsymbol{\beta}, 1) \mathcal{N}(\boldsymbol{\beta} | \mathbf{0}, \sigma_{\boldsymbol{\beta}}^2 \mathbf{I}) \\ &\propto \exp \left\{ -\frac{1}{2} (\mathbf{u} - \mathbf{X}\boldsymbol{\beta})' (\mathbf{u} - \mathbf{X}\boldsymbol{\beta}) \right\} \exp \left\{ -\frac{1}{2} (\boldsymbol{\beta} - \mathbf{0})' (\sigma_{\boldsymbol{\beta}}^2 \mathbf{I})^{-1} (\boldsymbol{\beta} - \mathbf{0}) \right\} \\ &\propto \exp \left\{ -\frac{1}{2} (\mathbf{u} - \mathbf{X}\boldsymbol{\beta})' (\mathbf{u} - \mathbf{X}\boldsymbol{\beta}) \right\} \exp \left\{ -\frac{1}{2} (\boldsymbol{\beta} - \mathbf{0})' (\sigma_{\boldsymbol{\beta}}^2 \mathbf{I})^{-1} (\boldsymbol{\beta} - \mathbf{0}) \right\} \\ &\propto \exp \left\{ -\frac{1}{2} (-2\mathbf{u}'\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}) \right\} \exp \left\{ -\frac{1}{2} (\boldsymbol{\beta}' (\sigma_{\boldsymbol{\beta}}^2 \mathbf{I})^{-1} \boldsymbol{\beta}) \right\} \\ &\propto \exp \left\{ -\frac{1}{2} \left(-2(\mathbf{u}'\mathbf{X})\boldsymbol{\beta} + \boldsymbol{\beta}' (\mathbf{X}'\mathbf{X} + (\sigma_{\boldsymbol{\beta}}^2 \mathbf{I})^{-1}) \boldsymbol{\beta} \right) \right\} \\ &= \mathcal{N}(\mathbf{A}^{-1}\mathbf{b}, \mathbf{A}^{-1}) \end{aligned}$$

where $\mathbf{A} = \mathbf{X}'\mathbf{X} + (\sigma_{\boldsymbol{\beta}}^2 \mathbf{I})^{-1}$ and $\mathbf{b}' = \mathbf{u}'\mathbf{X}$.