Generalized Linear Model for Binary Data using the Probit Link

Brian M. Brost

10 August 2015

Description

A regression model for binary data using the probit link.

Implementation

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

Model statement

Let y_t, for $t = 1, \ldots, T$, 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.

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

Full conditional distributions

Observation model auxiliary variable (u_t):

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

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

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

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