Probit Regression:

$$y_i \sim \text{Bern}(\theta_i), \quad \Phi^{-1}(\theta_i) = \underline{x}_i'\underline{\beta}$$

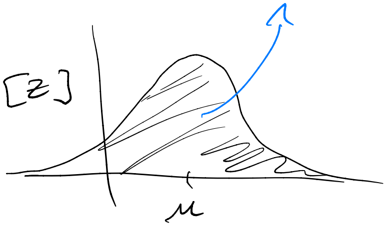
$$\Rightarrow \theta_i = \Phi(\underline{x}_i'\underline{\beta})$$

↑ Std. Norm. CDF (in R 'pnorm').

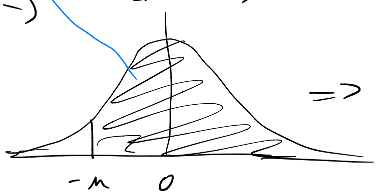
Notes on Probit Link function:

If $z \sim N(\mu, 1)$, then $z-\mu \sim N(0,1)$

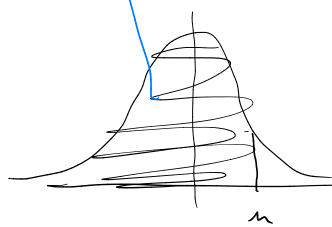
$$\Rightarrow P(z > 0) = P(z > \mu - \mu)$$
$$= P(z - \mu > -\mu)$$
$$= P(z - \mu < \mu)$$
$$= \Phi(\mu)$$



$\Rightarrow$



$\Rightarrow$



Thus, $y \sim \text{Bern}(\theta)$, with $\Phi^{-1}(\theta) = \mu$,

is equivalent to:
$$y = \begin{cases} 0, & z \le 0 \\ 1, & z > 0 \end{cases}$$
$$z \sim N(\mu, 1)$$
↑ auxiliary ("latent") variable

because:

$$P(y=1) = P(z>0) = \Phi(\mu).$$

Thus, for probit regression, there are 2 options:

1.) Use M-H w/ binary regression using probit link.

2.) Use latent variable approach:

$$y_i \sim \begin{cases} 0 & , z_i \le 0 \\ 1 & , z_i > 0 \end{cases} \quad , i = 1, \ldots, n$$

$$z_i \sim N(\underline{x}_i'\underline{\beta}, 1) \implies y_i = 1 \text{ w.p. } \Phi(\underline{x}_i'\underline{\beta})$$

$$\underline{\beta} \sim N(\mu_\beta, \Sigma_\beta)$$

Posterior Distn:

$$[\underline{\beta}, \underline{z} \mid \underline{y}] \propto \Big(\prod_{i=1}^{n} \underbrace{[y_i \mid z_i]}_{\text{constraints}} [z_i \mid \underline{\beta}]\Big)[\underline{\beta}]$$

$$\propto \Big(\prod_{i=1}^{n} \big(1_{\{z_i > 0\}}^{y_i} + 1_{\{z_i \le 0\}}^{1-y_i} \big) [z_i \mid \underline{\beta}] \Big)[\underline{\beta}]$$

Full-conditional distns:

$$[\underline{\beta} \mid \cdot] \propto \prod_{i=1}^{n} [z_i \mid \underline{\beta}][\underline{\beta}]$$

$$\propto \exp\left\{-\tfrac{1}{2}(\underline{z} - X\underline{\beta})' I (\underline{z} - X\underline{\beta})\right\} \exp\left\{-\tfrac{1}{2}(\underline{\beta} - \mu_\beta)' \Sigma_\beta^{-1} (\underline{\beta} - \mu_\beta)\right\}$$

$$\propto \exp\Big\{-\tfrac{1}{2}\Big(-2\underbrace{\big(\underline{z}' I X + \mu_\beta' \Sigma_\beta^{-1}\big)}_{\underline{b}'}\underline{\beta} + \underline{\beta}' \underbrace{\big(X' I X + \Sigma_\beta^{-1}\big)}_{A} \underline{\beta}\Big)\Big\},$$

$$= N(A^{-1}\underline{b}, A^{-1})$$

Conjugate!

$$[z_i | \cdot] \propto \left(1_{\{z_i > 0\}}^{y_i} + 1_{\{z_i \leq 0\}}^{1-y_i}\right)[z_i | \boldsymbol{\beta}]$$

$$\propto \begin{cases} N(\mathbf{x}_i'\boldsymbol{\beta}, 1)\, 1_{\{z_i > 0\}} & , y_i = 1 \\ N(\mathbf{x}_i'\boldsymbol{\beta}, 1)\, 1_{\{z_i \leq 0\}} & , y_i = 0 \end{cases}$$

$$= \begin{cases} TN(\mathbf{x}_i'\boldsymbol{\beta}, 1)\big|_0^{\infty} & , y_i = 1 \\ TN(\mathbf{x}_i'\boldsymbol{\beta}, 1)\big|_{-\infty}^{0} & , y_i = 0 \end{cases}$$

MCMC Algorithm:

1.) Set $\boldsymbol{\beta}^{(0)}$ at init., $k=0$
2.) $k = k+1$
3.) sample $z_i^{(k)} \sim [z_i | \boldsymbol{\beta}^{(k-1)}, y_i]$ for $i = 1, \ldots, n$
4.) Sample $\boldsymbol{\beta}^{(k)} \sim [\boldsymbol{\beta} | \mathbf{z}^{(k)}, \cdot]$
5.) Goto 2 until $k = K$.

This latent variable representation can be extended to many settings:

1.) multinomial data (Albert and Chib, 1993)

2.) hierarchical models (Albert and Chib, 1993)

3.) hier. spatial models (Hooten et al., 2003)

4.) zero-inflated models (Johnson et al., 2013)