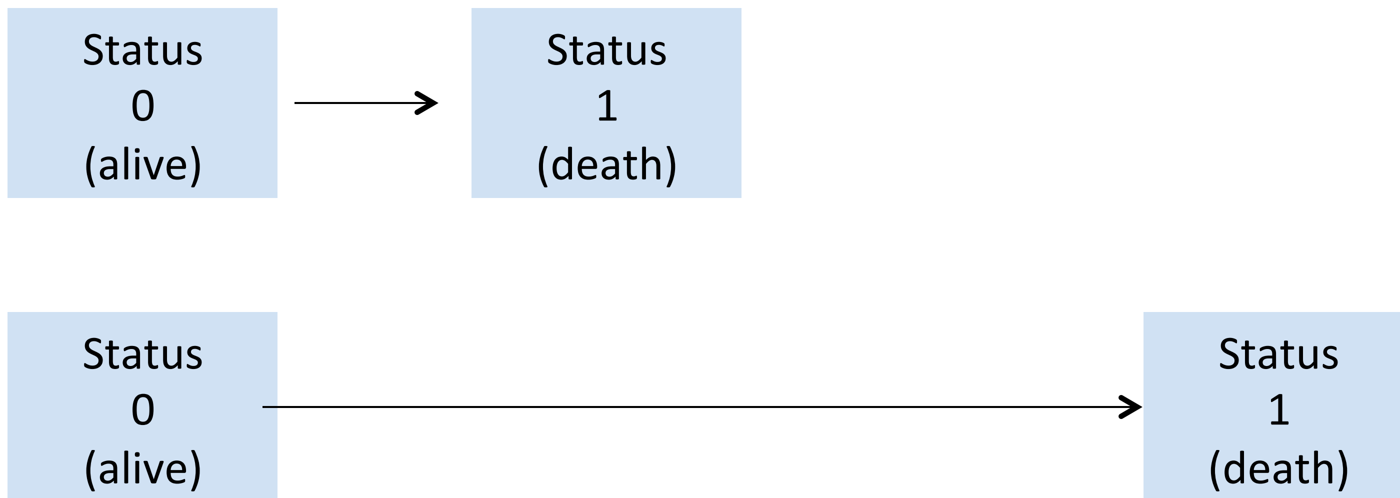
STT 465
Bayesian Multiple Linear Regression:

=> Regression With Censored Data

=> Regression with binary outcomes

Regression with Censored Data

Time to Event Data



Censoring

The event time is larger than the observation time (right censoring).

EXAMPLES:

-The study is closed.

-The subject is lost from follow-up.

The time to event is shorter than the observed time (left censoring).

EXAMPLE:

-When did you started using cigarettes? “I cannot recall”.

-At the beginning of an observational study following up diabetes incidence, some subjects have diabetes at the beginning.

The event is known to fall in an interval, but the exact time is unknown (interval censoring).

EXAMPLE:

-Phenotypes are collected every 5 years.

Censoring

The event time is larger than the observation time (right censoring).

EXAMPLES:

-The study is closed.

-The subject is lost from follow-up.

The time to event is shorter than the observed time (left censoring).

EXAMPLE:

-When did you started using cigarettes? “I cannot recall”.

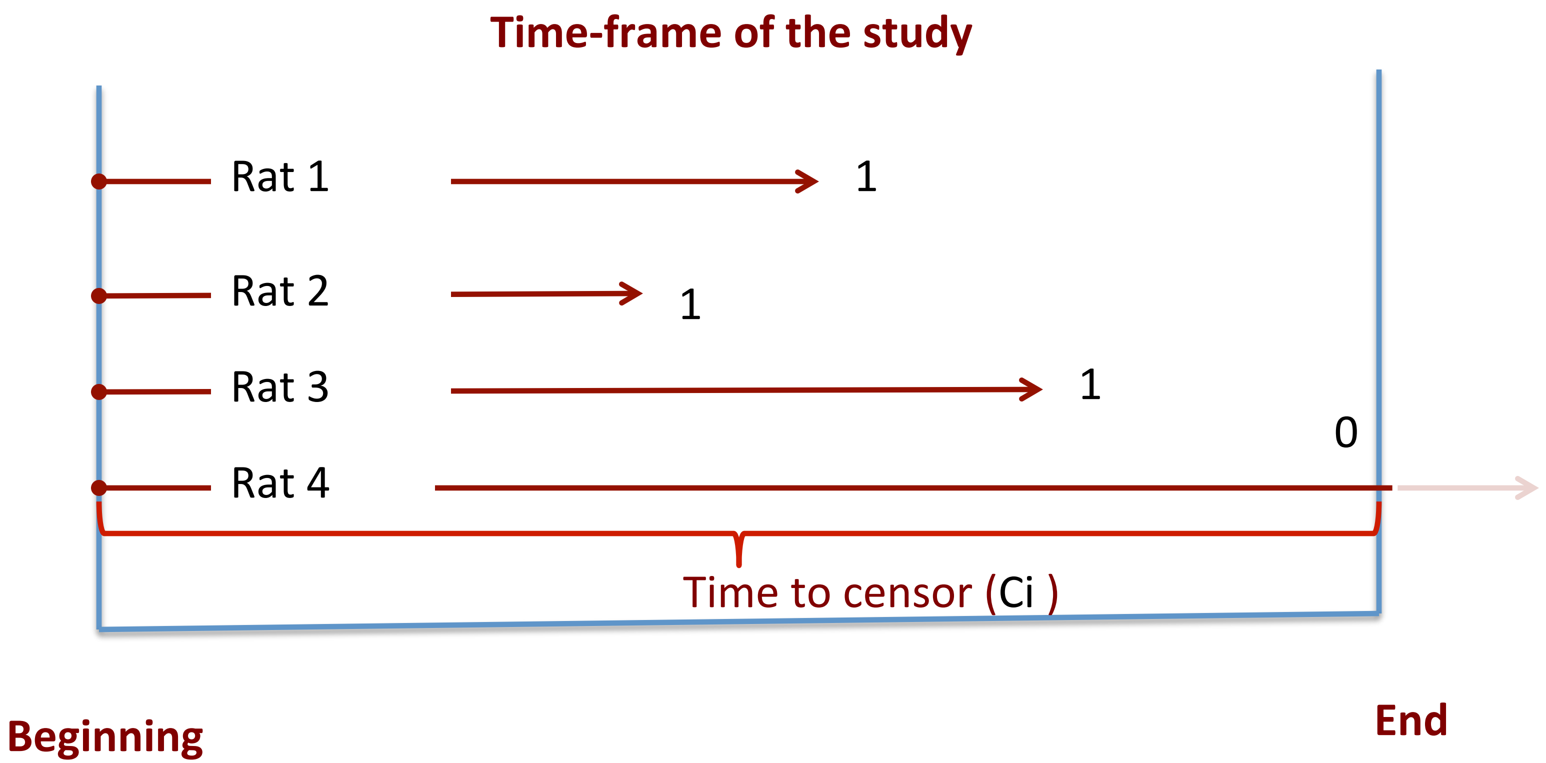
-At the beginning of an observational study following up diabetes incidence, some subjects have diabetes at the beginning.

The event is known to fall in an interval, but the exact time is unknown (interval censoring).

EXAMPLE:

-Phenotypes are collected every 5 years.

Right-Censoring (Y_i>C_i)



Rats 1,2 and 3: time to event is observed.

Rat 4: time to event is unknown; however we know the time to censor, and we know that the event will happens after T_i>Ri (right-censored data)

Dealing with Censoring in Bayesian Regression

Our strategy will be as the same as before:

- Write down the likelihood (for censored data in this case)
- Define the prior
- Using the above, arrive at the joint posterior (it typically does not have a closed form)
- From the joint posterior derive the fully conditional distribution of the censored data points.
- Incorporate the fully conditional on the Gibbs sampler.

Notation:

- Yi: response (e.g., time to event or censoring time)
- Ci: a dummy variable with 1 indicating event and 0 indicating censoring.

Likelihood Function

Observed Data-points

$$p\left(y_i \mid X, \beta, \sigma_\varepsilon^2\right) = \frac{Exp\left\{-\frac{\left(y_i - x_i'\beta\right)^2}{2\sigma_\varepsilon^2}\right\}}{\sqrt{2\pi\sigma_\varepsilon^2}}$$

Censored Data-points

$$p\left(y_i \mid X, \beta, \sigma_\varepsilon^2, c_i\right) = \frac{Exp\left\{-\frac{\left(y_i - x_i'\beta\right)^2}{2\sigma_\varepsilon^2}\right\}}{\sqrt{2\pi\sigma_\varepsilon^2}} 1\left(y_i > c_i\right)$$

Likelihood Function

$$p\left(y \mid X, \beta, \sigma_\varepsilon^2\right) = \prod_{i:c_i=1} \frac{e^{\left\{-\frac{\left(y_i - x_i'\beta\right)^2}{2\sigma_\varepsilon^2}\right\}}}{\sqrt{2\pi\sigma_\varepsilon^2}} \times \prod_{i:c_i=0} \frac{e^{\left\{-\frac{\left(y_i - x_i'\beta\right)^2}{2\sigma_\varepsilon^2}\right\}}}{\sqrt{2\pi\sigma_\varepsilon^2}} 1\left(y_i > c_i\right)$$

Fully Conditional Distribution

Likelihood Function

$$p\left(y \mid X,\beta,\sigma_\varepsilon^2\right) = \prod_{i:c_i=1} \frac{e^{\left\{-\frac{(y_i - x_i'\beta)^2}{2\sigma_\varepsilon^2}\right\}}}{\sqrt{2\pi\sigma_\varepsilon^2}} \times \prod_{i:c_i=0} \frac{e^{\left\{-\frac{(y_i - x_i'\beta)^2}{2\sigma_\varepsilon^2}\right\}}}{\sqrt{2\pi\sigma_\varepsilon^2}} 1\left(y_i > c_i\right)$$

Prior

$$p\left(\beta,\sigma_\varepsilon^2,\sigma_\beta^2,\ldots\right)$$

Joint Posterior:

$$p\left(\beta,\sigma_\varepsilon^2,\sigma_\beta^2,\ldots \mid y,c\right) \propto \prod_{i:c_i=1} \frac{e^{\left\{-\frac{(y_i - x_i'\beta)^2}{2\sigma_\varepsilon^2}\right\}}}{\sqrt{2\pi\sigma_\varepsilon^2}} \times \prod_{i:c_i=0} \frac{e^{\left\{-\frac{(y_i - x_i'\beta)^2}{2\sigma_\varepsilon^2}\right\}}}{\sqrt{2\pi\sigma_\varepsilon^2}} 1\left(y_i > c_i\right) \times p\left(\beta,\sigma_\varepsilon^2,\sigma_\beta^2,\ldots\right)$$

Fully Conditional (truncated normal):

$$p\left(y_i \mid c_i =\right) \propto \frac{e^{\left\{-\frac{(y_i - x_i'\beta)^2}{2\sigma_\varepsilon^2}\right\}}}{\sqrt{2\pi\sigma_\varepsilon^2}} 1\left(y_i > c_i\right)$$

Regression with Binary Outcomes

Regression with Binary Outcomes

- Binary outcomes follow Bernoulli distributions

$$y_i \in \{0,1\} \qquad\qquad p\left(y_i \middle| \theta\right) = \theta^{y_i}\left(1-\theta\right)^{1-y_i}$$

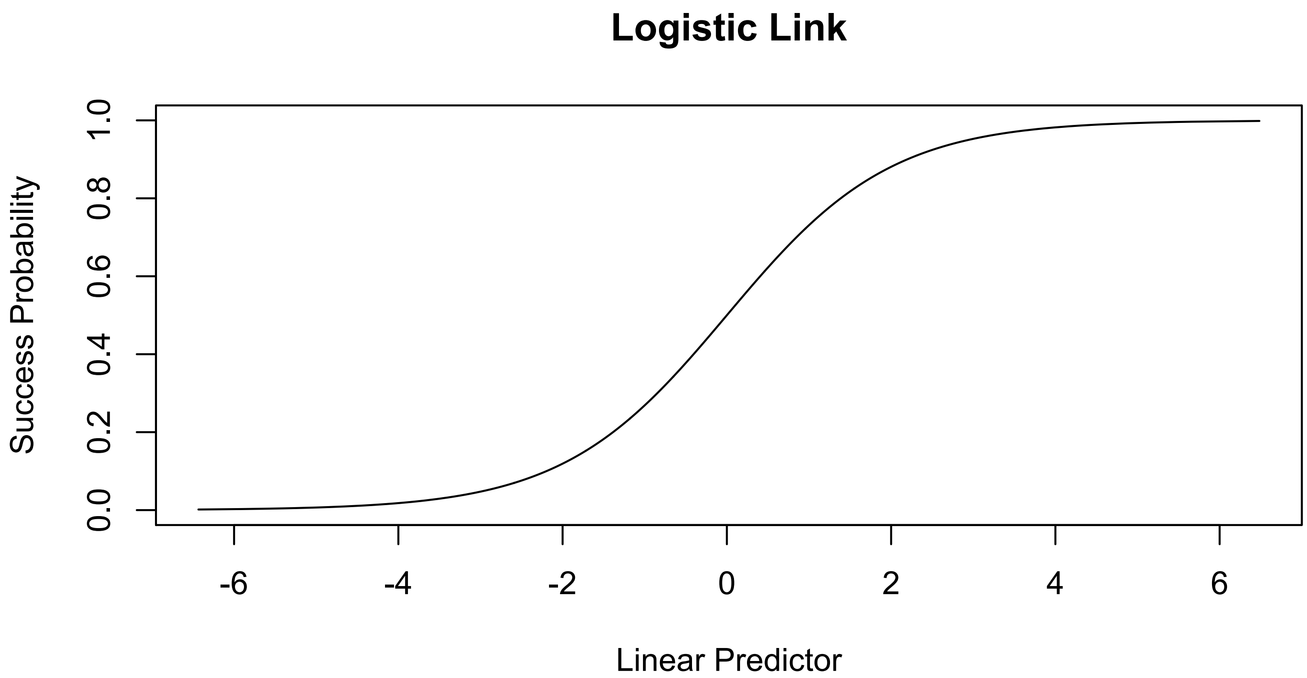
- Regression with Binary outcomes: we want to make θ a function of one or more predictors.
- Problem:
  - θ lives in the 0-1 interval,
  - while a regression function $\sum_{j=1}^{p} x_{ij}\beta_j$ lives in the real line.

- To get around this we need to introduce a link function that maps from the real line to the 0-1 interval.

- The most commonly used links are the logit and probit link.

Logistic Regression

$$p\left(y_i = 1\right) = \theta \qquad \log\left(\frac{\theta}{1-\theta}\right) = \sum_{j=1}^{p} x_{ij}\beta_j \qquad \Rightarrow \theta = \frac{e^{\sum_{j=1}^{p} x_{ij}\beta_j}}{1+e^{\sum_{j=1}^{p} x_{ij}\beta_j}}$$



Logistic Regression

- Likelihood Function

$$p(y) = \prod_{i=1}^{n} \theta^{y_i} (1-\theta)^{1-y_i}$$

$$\log\left(\frac{\theta}{1-\theta}\right) = \eta_i = \sum_{j=1}^{p} x_{ij}\beta_j \qquad \Rightarrow \theta_i = \frac{e^{\sum_{j=1}^{p} x_{ij}\beta_j}}{1+e^{\sum_{j=1}^{p} x_{ij}\beta_j}}$$

$$p(y) = \prod_{i=1}^{n} \left[\frac{e^{\eta_i}}{1+e^{\eta_i}}\right]^{y_i} \left[1 - \frac{e^{\eta_i}}{1+e^{\eta_i}}\right]^{1-y_i}$$

Threshold Model (Probit Link)

$$l_i = \sum_{j=1}^{p} x_{ij}\beta_j + \varepsilon_i \qquad\qquad y_i = \begin{cases} 0 & y_i < 0 \\ 1 & \text{Otherwise} \end{cases}$$

Logistic Regression

- Likelihood Function

$$p(y) = \prod_{i=1}^{n} \theta^{y_i} (1-\theta)^{1-y_i}$$

$$\log\left(\frac{\theta}{1-\theta}\right) = \sum_{j=1}^{p} x_{ij}\beta_j \Rightarrow \theta = \frac{e^{\sum_{j=1}^{p} x_{ij}\beta_j}}{1+e^{\sum_{j=1}^{p} x_{ij}\beta_j}}$$

$$p(y) = \prod_{i=1}^{n} \left[\frac{e^{\sum_{j=1}^{p} x_{ij}\beta_j}}{1+e^{\sum_{j=1}^{p} x_{ij}\beta_j}}\right]^{y_i} \left[1-\frac{e^{\sum_{j=1}^{p} x_{ij}\beta_j}}{1+e^{\sum_{j=1}^{p} x_{ij}\beta_j}}\right]^{1-y_i}$$