## Model

Tobit model is a regression model where some of the dependent variable $y_i^*$ observations $i$ are censored. For censored observations all we can say is that the variable is less(left censoring) or more(right censoring) than some other value $c_i$. Let's indicate that variable is left censored with $l_i = 1, r_i = 0$, right censored with $r_i = 1, l_i = 0$ and is not censored when $r_i = 0 = l_i$. Let's call the latent variable before censoring $y^*$.

$$y_i = \begin{cases} y_i^* & \text{if } r_i = 0 = l_i, \\ c_i & \text{if } l_i = 1, r_i = 0, \\ c_i & \text{if } r_i = 1, l_i = 0 \end{cases} \tag{1}$$

We assume that the underlying generative model is:

$$y_i^* = f(\boldsymbol{x_i}, \boldsymbol{\beta}) + \epsilon \tag{2}$$

Where $\boldsymbol{x_i}$ are independent variables and $\boldsymbol{\beta}$ are model parameters and $\epsilon$ is normally distributed $\epsilon \sim \mathcal{N}(0, \sigma^2)$. Then $y_i^*$ is also normally distributed $y_i^* \sim \mathcal{N}(f(\boldsymbol{x_i}, \boldsymbol{\beta}), \sigma^2)$ After observing real data we try to find parameters $\boldsymbol{\beta}$ that fits the data. We do it using maximum likelihood estimation. We have 3 types of observation that will translate to parts of likelihood. For uncensored data we take probability density function at $y_i$ observed values. Let *pdf* be the probability density function of standard normal distribution

$$\mathcal{L}_{1,i} = pdf(\frac{y_i - f(\boldsymbol{x_i}, \boldsymbol{\beta})}{\sigma})\frac{1}{\sigma} \tag{3}$$

For left censored data we take probability that latent variable is less than observed variable. Let *cdf* be cumulative distribution function of standard normal distribution

$$\mathcal{L}_{2,i} = \mathcal{P}(y_i^* \leq y_i) = cdf(\frac{y_i - f(\boldsymbol{x_i}, \boldsymbol{\beta})}{\sigma}) \tag{4}$$

For right censored data we take probability that latent variable is more than observed variable

$$\mathcal{L}_{3,i} = \mathcal{P}(y_i^* > y_i) = 1 - cdf(\frac{y_i - f(\boldsymbol{x_i}, \boldsymbol{\beta})}{\sigma}) \tag{5}$$

Finally putting everything together

$$\mathcal{L} = \prod_{i=1}^{N} = \mathcal{L}_{1,i}^{(1-l_i)(1-r_i)} * \mathcal{L}_{2,i}^{(l_i)(1-r_i)} * \mathcal{L}_{3,i}^{(1-l_i)(r_i)} \tag{6}$$

Our goal is to find

$$\max_{\boldsymbol{\beta}} \mathcal{L} \tag{7}$$