PS in TA's mailbox

Censored Regression Model

$$Y_i^* = \alpha + \beta X_i + \varepsilon_i$$

"toy" $Y_i = \begin{cases} Y_i^* & \text{if } Y_i^* \geq 0 \\ 0 & \text{otherwise} \end{cases}$ if you see zero, all you know is that $Y_i^* < 0$

Y_i^* is "left-censored at zero."

Practical Example

Top Coding $Y_i = \begin{cases} \alpha & Y_i^* \geq \alpha \\ Y_i^* & 0 < Y_i^* < \alpha \end{cases}$ "censoring from the right"

It is important to select this α in such a way that the dataset is informative and privacy is not violated.

Tobin - Demand for refrigerator.

$$Y_i^* = \alpha + \beta X_i + \varepsilon_i$$

$$Y_i = \begin{cases} Y_i^* & Y_i^* < C \\ C & \text{else} \end{cases}$$

$$\Rightarrow C - Y_i^* = (C - \alpha) - \beta X_i - \varepsilon_i$$

Observe: $\begin{cases} C - Y_i^* & \text{if } C - Y_i^* > 0 \iff Y_i^* \leq C \\ 0 & \text{if } C - Y_i^* \leq 0 \iff Y_i^* > C \end{cases}$

$$y_i^* = \gamma + \delta X_i + u_i$$

$$y_i = \begin{cases} y_i^* & y_i^* \geq 0 \\ 0 & \text{else} \end{cases}$$

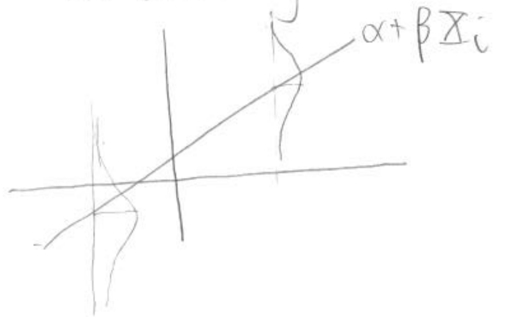
This is the same as the model above "toy".

(Tobit model) - Left-censored model

<u>This</u> is solved by Heckman (Heckit)

No censoring

Story (I)

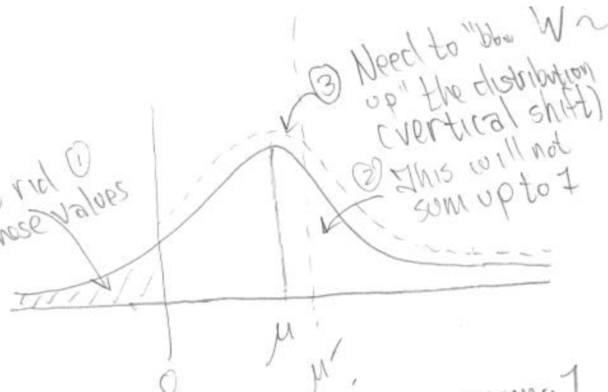


$$Y_i = \alpha + \beta X_i + \varepsilon_i$$

$$\Rightarrow E[Y_i] = \alpha + \beta X_i$$

$W \sim N(\mu, \sigma^2)$

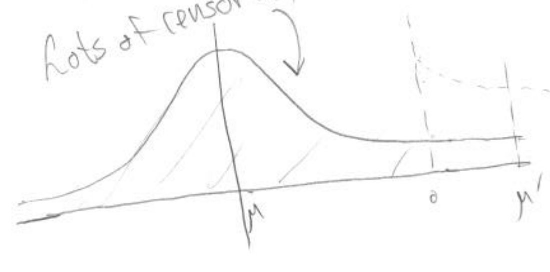
Story (II)



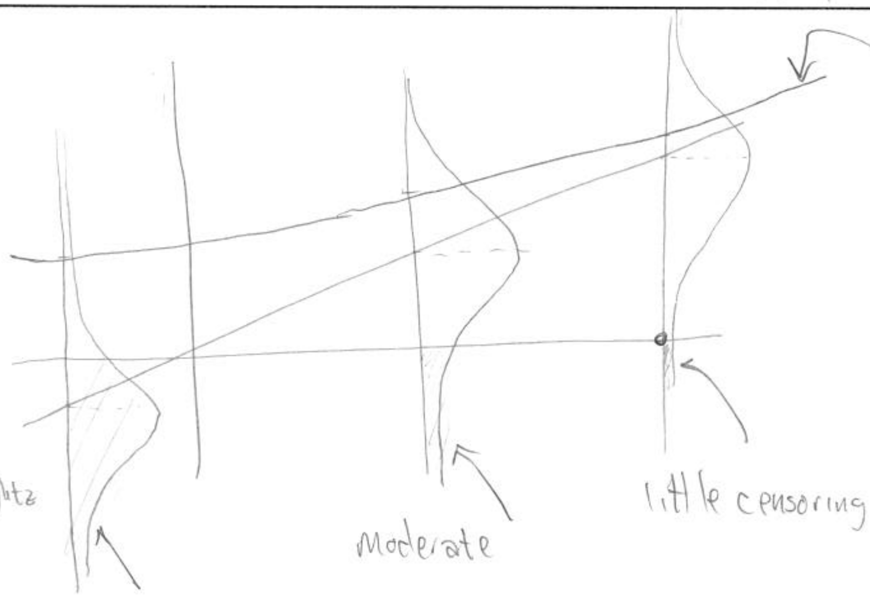
Condition this distribution on the event that $W_i \geq 0$

ignoring 1 $\Rightarrow$

$$\Pr[\{2\} \mid \{1\}^c] = \frac{\Pr[\{2\} \cap \{1\}^c]}{\Pr[\{1\}^c]} = \frac{\Pr[\{2\}]}{\Pr[\{1\}^c]} = \frac{\frac{1}{6}}{\frac{5}{6}} = \frac{1}{5}$$



Story (III)



OLS will try to trace this curve
└ The fitted line will be flatter than the original line
└ There will be a bias

05 Aumann, Schelling
04 Kydland, Prescott
03 Engel, Granger
02 Kahneman, Smith
01 Spence, Akerlof, Stiglitz
00 Heckman?

OLS will not work ⇒ Use Heckman's two-step method
└ Not in the final.

Binary Choice Model:

$Pr[Y_i = 1] = F(\alpha + \beta X_i)$

$U_{i1}^* = \alpha_1 + \beta_1 X_i + \varepsilon_{i1}$ indirect utility fcn, if purchase refrigerator

X_i: eg price of refrigerator

$U_{i0}^* = \alpha_0 + \varepsilon_{i0}$ indirect utility fcn if not purchase refrigerator

Choose $\max\{U_{i1}^*, U_{i0}^*\}$

$Y_i = 1$ iff $U_{i1}^* \geq U_{i0}^*$

$\iff \alpha_1 + \beta_1 X_i + \varepsilon_{i1} \geq \alpha_0 + \varepsilon_{i0}$

$\iff \varepsilon_{i0} - \varepsilon_{i1} \leq (\alpha_1 - \alpha_0) + \beta_1 X_i$

$\iff \varepsilon_i \leq \alpha + \beta X_i$

$\Rightarrow Pr[Y_i = 1] = Pr[\varepsilon_i \leq \alpha + \beta X_i]$

assume $\varepsilon_i \sim N(0,1)$

Recall $\Phi(\cdot)$ is the cdf of $N(0,1)$.

$$\Rightarrow \Pr[Y_i = 1] = \Pr[\varepsilon_i \leq \alpha + \beta X_i] = \Phi(\alpha + \beta X_i)$$

assume that ε_i has cdf $\frac{e^t}{1+e^t}$

Then, $\Pr[Y_i = 1] = \frac{e^{\alpha+\beta X_i}}{1+e^{\alpha+\beta X_i}}$

In general, if ε_i has cdf F, then

$$\Pr[Y_i = 1] = F(\alpha + \beta X_i)$$

McFadden gave this utility-maximization interpretation and won the Nobel prize.

Likelihood of $Y_1, \ldots, Y_n$

$$\max_{\alpha,\beta} \; [F(\alpha+\beta X_1)]^{Y_1}[1-F(\alpha+\beta X_1)]^{1-Y_1} \cdots [F(\alpha+\beta X_n)]^{Y_n}[1-F(\alpha+\beta X_n)]^{1-Y_n}$$

This will give us $\hat{\alpha}_{MLE}, \hat{\beta}_{MLE}$

$$\hat{\alpha}_{MLE} \xrightarrow{d} N(\alpha, SE_\alpha^2)$$

$$\hat{\beta}_{MLE} \xrightarrow{d} N(\beta, SE_\beta^2)$$

Probit model of voting for McGovern

$$Y_i = \begin{cases} 1 & \text{if } i \text{ voted for McGovern} \\ 0 & \text{else} \end{cases}$$

$$Pr[Y_i = 1] = \Phi\left(-.713 - \underset{(.082)}{.375}F_i - \underset{(.066)}{.257}T_i - \underset{(.092)}{.593}V_i - \underset{(.058)}{.075}M_i - \ldots\right)$$

$$Pr[Y_i = 1] = \Phi(\beta_1 + \beta_2 X_{1i} + \ldots + \beta_k X_{ki})$$

$$\frac{\partial Pr[Y_i = 1]}{\partial X_2} = \underbrace{\varphi(\beta_1 + \ldots + \beta_K X_{Ki})}_{\geq 0} \cdot \beta_2 \geq 0 \text{ iff } \beta_2 \geq 0$$

Suppose H_0: X_2 does not matter

$$\Leftrightarrow H_0: \frac{\partial Pr(Y_i = 1)}{\partial X_2} = 0$$

$$\Leftrightarrow H_0: \beta_2 = 0$$

Does F_i matter? $\left|\frac{-.375 - 0}{.082}\right| > 4 > 1.96$

$H_0: \beta_2 = 0$, $H_1: \beta_2 \neq 0$

$\Rightarrow$ Yes, F_i matters

Note $H_0: \beta_2 = c \neq 0$ does not have any meaning.

Final Exam) The mathematical form of binary choice is important

Marginal impact is important

Determine whether or not X_j matters.