* Q1 in PS 6 need not be solved
* Final 1 cheat sheet allowed
* Partial credit will be given in the final

Limited Dependent Variable (LIMDEP)

1. Binary Choice Model (Dummy variable appears on LHS)

$Y_i = \alpha + \beta X_i + \varepsilon_i$ where $Y_i = \begin{cases} 1 & 1^{st} \text{ option} \\ 0 & \text{else} \end{cases}$

Suppose $Y_i = 1$ if buy refrigerator, $Y_i = 0$ if not

assume $E[\varepsilon_i] = 0$, X_i nonstochastic

$\Rightarrow \underset{\Pr[Y_i = 1] \equiv P_i}{\underset{\parallel}{E[Y_i]}} = \alpha + \beta X_i$

This gives us $P_i = \alpha + \beta X_i$

But if $X_i > \frac{1-\alpha}{\beta}$, then $P_i > 1$ $\rightarrow\leftarrow$

$\rightarrow$ For X_i sufficiently high or low, $P_i \notin [0,1]$, which does not make sense

This is called the linear probability model.

$Var[Y_i] = P_i(1-P_i)$

$Var[Y_i] = (\alpha + \beta X_i)(1 - \alpha - \beta X_i) \neq \text{constant}$

$\hookrightarrow$ There is heteroskedasticity by assumption.

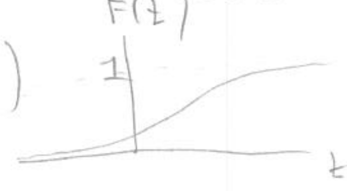
People decided to do something else instead

- Logit
- Probit

$$Y_i = \begin{cases} 1 & \text{1st option} \\ 0 & \text{else} \end{cases}$$

$p_i = \Pr[Y_i = 1]$ characterizes the entire distribution.

$= F(\alpha + \beta X_i)$ st. $F: t \longrightarrow (0,1)$



<u>Logit:</u>

$$F(t) = \frac{\exp\{t\}}{1+\exp\{t\}} \in (0,1) \quad \forall t$$

Clearly $F(t) \uparrow$ in t

$$\lim_{t \to \infty} F(t) = 1$$

$$\lim_{t \to -\infty} F(t) = 0$$

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

<u>Implication:</u>

$$E[Y] = \Pr[Y=1] = \frac{\exp\{\alpha+\beta X\}}{1+\exp\{\alpha+\beta X\}}$$. What is $\frac{\partial E[Y]}{\partial x}$?

$$\frac{\partial E[Y]}{\partial X} = \frac{(1+\exp\{\alpha+\beta X\})\beta\exp\{\alpha+\beta X\} - \exp\{\alpha+\beta X\}\beta\exp\{\alpha+\beta X\}}{(1+\exp\{\alpha+\beta X\})^2}$$

$$= \frac{\beta\exp\{\alpha+\beta X\}}{(1+\exp\{\alpha+\beta X\})^2}$$

Clearly, $\text{sgn}\left[\frac{\partial E[Y]}{\partial X}\right] = \text{sgn}[\beta]$, but $\frac{\partial E[Y]}{\partial X} = g(X)$

That is, the response to a change in X is a function of the current value of X.

Multivariate Generalization

$$E[Y] = Pr[Y=1] = \frac{\exp\{\beta_1 + \beta_2 X_2 + \dots + \beta_K X_K\}}{1 + \exp\{\beta_1 + \beta_2 X_2 + \dots + \beta_K X_K\}}$$

$$\Rightarrow \frac{\partial E[Y]}{\partial X_j} = \frac{\exp\{\beta_1 + \dots + \beta_K X_K\}}{(1 + \exp\{\beta_1 + \dots + \beta_K X_K\})^2} \beta_j$$

Thus $\text{sgn}\left[\frac{\partial E[Y]}{\partial X_j}\right] = \text{sgn}\ \beta_j$

Probit model:

$\Phi(t) \equiv P_r[Z \le t]$ where $Z \sim N(0,1)$

Probit model assumes that $Pr[Y=1] = \Phi(\alpha + \beta X)$

$$\Rightarrow E[Y] = \int_{-\infty}^{\alpha+\beta X} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{t^2}{2}\right\} dt$$

$$\frac{\partial E[Y]}{\partial X} = \varphi(\alpha + \beta X_i) \cdot \beta$$

where $\varphi(t) = \frac{d\Phi(t)}{dt}$ is the pdf of $N(0,1)$

$$= \beta \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{(\alpha+\beta X_i)^2}{2}\right\}$$

Thus, $\text{sgn}\left[\frac{\partial E[Y]}{\partial X}\right] = \text{sgn}(\beta)$

Multivariate Case:

$$\Pr[Y=1] = \Phi(\beta_1 + \beta_2 X_2 + \dots + \beta_K X_K)$$

$$\Rightarrow \frac{\partial \Pr[Y=1]}{\partial X_j} = \phi(\beta_1 + \beta_2 X_2 + \dots + \beta_K X_K) \cdot \beta_j$$

and $\operatorname{sgn}\left[\frac{\partial \Pr[Y=1]}{\partial X_j}\right] = \operatorname{sgn}(\beta_j)$

Maximum Likelihood Estimation: Let $X \sim N(\theta, 1)$

$$f(x;\theta) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{(x-\theta)^2}{2}\right\}$$

Let $X_1, \dots, X_n \overset{iid}{\sim} N(\theta, 1)$. Then we have that

$$f(x_1, \dots, x_n; \theta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{(x_i-\theta)^2}{2}\right\}$$

$$= \frac{1}{(2\pi)^{n/2}} \exp\left\{-\frac{\sum_{i=1}^{n}(x_i-\theta)^2}{2}\right\}$$

PDF of binary rv w/ $\theta = \Pr[\text{head}]$

$$f(y;\theta) = \Pr[Y=y] = \theta^{y}(1-\theta)^{1-y}, \quad y \in \{0,1\}$$

$$\Rightarrow f(y_1, \dots, y_n; \theta) = \prod_{i=1}^{n} \theta^{y_i}(1-\theta)^{1-y_i} \quad \text{if } Y_1, \dots, Y_n \overset{iid}{\sim} D(\theta)$$

$$L(\theta; y) = \theta^{\sum_{i=1}^{n} y_i}(1-\theta)^{\sum_{i=1}^{n}(1-y_i)}$$

$$\log L(\theta; y) = \left[\sum_{i=1}^{n} y_i\right]\log\theta + \left[\sum_{i=1}^{n}(1-y_i)\right]\log(1-\theta)$$

$$\frac{\partial \log L}{\partial \theta} = 0 \Rightarrow \frac{\sum_{i=1}^{n} y_i}{\theta_{MLE}} = \frac{\sum_{i=1}^{n}(1-y_i)}{1-\theta_{MLE}}$$

$$\Rightarrow \frac{\theta_{MLE}}{1-\theta_{MLE}} = \frac{\Sigma(y_i)}{\Sigma(1-y_i)}$$

θ_{MLE} is consistent and normal

$$\theta_{MLE} = \left(\frac{\Sigma y_i}{\Sigma(1-y_i)}\right) - \theta_{MLE}\left(\frac{\Sigma y_i}{\Sigma(1-y_i)}\right)$$

$$\theta_{MLE}\left[1 + \frac{\Sigma y_i}{\Sigma(1-y_i)}\right] = \frac{\Sigma y_i}{\Sigma(1-y_i)}$$

$$\theta_{MLE}\left(\frac{\Sigma(1-y_i) + \Sigma y_i}{\Sigma(1-y_i)}\right) = \frac{\Sigma y_i}{\Sigma(1-y_i)}$$

$$\theta_{MLE} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

Logit

$$P_i = \frac{\exp\{\alpha + \beta X_i\}}{1 + \exp\{\alpha + \beta X_i\}}$$

$$\max_{\alpha, \beta} \; P_1^{Y_1}(1-P_1)^{1-Y_1} \cdots P_n^{Y_n}(1-P_n)^{1-Y_n}$$

gives us α_{MLE}, β_{MLE}

$$\max_{\alpha, \beta} \prod_{i=1}^{n} \left[\frac{\exp\{\alpha+\beta X_i\}}{1+\exp\{\alpha+\beta X_i\}}\right]^{Y_i} \left[\frac{1}{1+\exp\{\alpha+\beta X_i\}}\right]^{1-Y_i}$$

$$= \max_{\alpha, \beta} \frac{\prod_{i=1}^{n}\left[\exp\{\alpha+\beta X_i\}\right]^{Y_i}}{\prod_{i=1}^{n}\left[1+\exp\{\alpha+\beta X_i\}\right]}$$

(*) In MLE, we can use the t-statistics and other STATA output.