Purchase Cameron and Trivedi's <u>Microeconometrics</u> and Wooldridge
└ good with intuition

Newey, McFadden, H of E, Volume 4?, Chapter 36?

1 Treatment Effect
- Causal analysis
- Angrist and Levy (smaller classes good)

2 Structural

Duflo and Hanna (Treatment effects) - Read abstract
- Experiments with teacher attendance and cameras.

Duflo, Hanna, Ryan (Structural)
- Structural approach - nonlinearity in incentive system. Try to estimate elasticities of labor supply.

Read section 3 of Wooldridge
- Ch 12: M estimation
- Ch 13: MLE
- Ch 14: GMM/MDE

Section 4: Applications
- Ch 15: Discrete response
- Ch 16: Censored
- Ch 17: Sample selection
- Ch 20: Duration analysis

Alternatively, read ch. 5 in Cameron and Trivedi
- Section 4 is similar to section 4 in Wooldridge

Random utility

$$U_i = x_i'\beta + \varepsilon_i$$

agent works iff $U_i > 0$

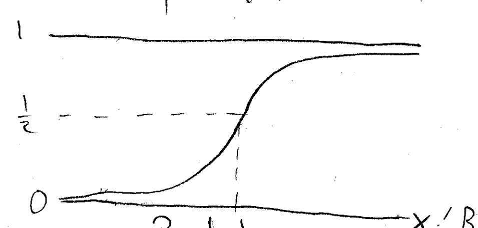
iff $\varepsilon_i > -x_i'\beta$

all we see is $y_i = 1_{\{\varepsilon_i > -x_i'\beta\}}$

$$\Pr[y_i = 1 \mid x_i] = \Pr[\varepsilon_i > -x_i'\beta \mid x_i] = \frac{\exp\{x_i'\beta\}}{1+\exp\{x_i'\beta\}} \equiv \Lambda(x_i'\beta)$$

if $\varepsilon_i \sim$ Type I extreme value

$$\Rightarrow \Pr[y_i = 0 \mid x_i] = 1 - \Lambda(x_i'\beta)$$



Borderline case
∘ small Δ in observables leads to big Δ in behavior

$$f_{Y_i}(y_i \mid x_i) = (\Lambda(x_i'\beta))^{y_i}(1-\Lambda(x_i'\beta))^{1-y_i}$$

$$\Rightarrow \log f(y_i \mid x_i) = y_i x_i'\beta - \log(1+\exp\{x_i'\beta\})$$

$$\Rightarrow L_i(\beta \mid y_i, x_i) \Rightarrow l_i = \log L_i$$

$$\Rightarrow \mathcal{L} = \sum l_i \quad \text{if iid data}$$

$$\hat{\beta}_{ML} = \underset{\beta \in B}{\operatorname{argmax}}\ \mathcal{L}(\beta \mid y, X)$$

$$= \underset{\beta \in B}{\operatorname{argmax}}\ \frac{1}{N}\sum_{i=1}^{N} l_i(\beta)$$

$$= \underset{\beta \in B}{\operatorname{argmax}}\ Q_N(\beta)$$

$\hat{\beta}$ is s.t. $\frac{\partial}{\partial \beta} Q_N(\hat{\beta}) = 0$

$$\Leftrightarrow \frac{1}{N} \sum_{i=1}^{N} \frac{\partial}{\partial \beta} l_i(\hat{\beta}) = 0$$

$$\Leftrightarrow \frac{1}{N} \sum_{i=1}^{N} (y_i - \Lambda(x_i'\hat{\beta}))x_i = 0$$

- Must do numerical compotation to characterize $\hat{\beta}$.

Define $\overbrace{s_i(\beta)}^{k \times 1} = \frac{\partial}{\partial \beta} l_i(\beta) = (y_i - \Lambda(x_i'\beta))x_i$.

$= [y_i - E[y_i|x_i]]x_i$

and $S(\beta) = \frac{1}{N} \sum_{i=1}^{N} s_i(\beta)$

$$= \frac{1}{N} \sum_{i=1}^{N} (y_i - E[y_i|x_i])x_i = 0$$

(cf normal equations)

Define $H_i(\beta) = \frac{\partial^2}{\partial \beta \partial \beta'} l_i(\beta)$

$$= - \frac{\Lambda(x_i'\beta)}{1 + \exp(x_i'\beta)} x_i x_i'$$

$$= - \Lambda(x_i'\beta)(1 - \Lambda(x_i'\beta)) x_i x_i'$$

Need $H(\beta) = -\frac{1}{N} \sum_{i=1}^{N} \Lambda(x_i'\beta)(1 - \Lambda(x_i'\beta)) x_i x_i'$ is negative semi-definite.

We have:

I Consistency: $\hat{\beta} \xrightarrow{p} \beta_0$

2] Asymptotic Normality:

$$\sqrt{N}(\hat{\beta}-\beta_0) \xrightarrow{d} N(0, A_0^{-1} B_0 A^{-1})$$

where $A_0 \equiv E[H_i(\beta_0)]$, $B_0 \equiv E[s_i(\beta_0)(s_i(\beta_0))']$

Can estimate covariance matrix by $\hat{A}^{-1}\hat{B}\hat{A}^{-1}$.

Under MLE correct specification,

1] Consistency: $\hat{\beta} \xrightarrow{p} \beta_0$

2] $\sqrt{N}(\hat{\beta}-\beta_0) \xrightarrow{d} N(0, A_0^{-1})$, since $A_0 = -B_0$

Can estimate A_0 with $\hat{A} = \frac{1}{N}\sum_{i=1}^{N} H_i(\hat{\beta})$

- Robust MLE involves using $\hat{A}^{-1}\hat{B}\hat{A}^{-1}$

To compute $\hat{\beta}$, we

1] Guess $\hat{\beta}^0$

2] Update

3] Iterate to convergence $\|\hat{\beta}^{j+1}-\hat{\beta}^j\|$ small

2] updating: 1^{st} derivative tells which direction to go 2^{nd} derivative tells how much. (larger 2^{nd} derivative implies more curvature $\Rightarrow$ take smaller step)

eg $\hat{\beta}^{(t+1)} = \hat{\beta}^{(t)} + C^{(t)} g^{(t)}$, where $C^{(t)}$ is $K \times K$, $g^{(t)}$ is $K \times 1$ gradient