$$\underline{Y}_i = \begin{bmatrix} Y_{i1} \\ \vdots \\ Y_{iT} \end{bmatrix}, \quad \underline{X}_i = \begin{bmatrix} X_{i1}' \\ \vdots \\ X_{iT}' \end{bmatrix}$$

$$Y_i \mid (X_i, \alpha_i) \sim f(y \mid X_i, \alpha_i, \theta)$$

$$\hat{\theta}_{FE} = \underset{\theta}{\operatorname{argmax}} \sum_{i=1}^{n} \ln f(Y_i \mid X_i, \hat{\alpha}_i(\theta), \theta)$$

where $\hat{\alpha}_i(\theta) = \underset{\alpha_i}{\operatorname{argmax}} \ln f(\underline{Y}_i \mid \underline{X}_i, \alpha, \theta)$

Incidental parameters problem:

- only have T observations to estimate α_i, so as $n \to \infty$, T fixed, $\hat{\theta}_{FE} \xrightarrow{p} \theta_T \neq \theta_0$.

Conditional MLE

There is a statistic $S_i = S(Y_i, X_i)$ such that

- $f(Y_i \mid X_i, \alpha_i, S_i, \theta) = f(Y_i \mid X_i, S_i, \theta)$

$$\hat{\theta}_{CMLE} = \underset{\theta}{\operatorname{argmax}} \sum_{i=1}^{n} \ln f(Y_i \mid X_i, S_i, \theta)$$

- need something about minimal sufficiency to get efficiency
- otherwise, it is CAN.

Count Data

- $Y_{it} \in \{0,1,\dots\}$
- $Y_i \mid (\alpha_i, X_i)$ is independent over t and Poisson for each t with conditional mean

$$\exp\{X_{it}'\beta + \alpha_i\} = \lambda_{it}$$

- $$f(y_{it} \mid X_i, \alpha_i, \beta) = \frac{e^{-\lambda_{it}} \lambda_{it}^{y_{it}}}{y_{it}!}$$

Patents: Hausman, Hall, Griliches (EMA 1984)

Take $S_i = \sum_{t=1}^{T} Y_{it}$ ◦ recall: sum of indep. Poisson is Poisson

Then $$f(S_i \mid X_i, \alpha_i, \beta) = \frac{e^{-\sum_{t=1}^{T}\lambda_{it}} \left(\sum_{t=1}^{T}\lambda_{it}\right)^{S_i}}{S_i!}$$

Log of conditional likelihood:

$$\sum_{t=1}^{T}\Big[-\lambda_{it} + y_{it}\ln\lambda_{it} - \overbrace{\ln(y_{it}!)}^{\equiv A_i}\Big]$$

$$-\Big[-\sum_{t=1}^{T}\lambda_{it} + S_i \ln\Big(\sum_{t=1}^{T}\lambda_{it}\Big) - \underbrace{\ln(S_i!)}_{\equiv B_i}\Big]$$

$$= \sum_{t=1}^{T} y_{it}\ln\lambda_{it} - S_i \ln\Big(\sum_{t=1}^{T}\lambda_{it}\Big) + C_i$$

- $\sum_{t=1}^{T} y_{it} \ln \lambda_{it} = \sum_{t=1}^{T} y_{it}(X_{it}'\beta + \alpha_i)$

$$= \sum_{t=1}^{T} y_{it} X_{it}'\beta + \alpha_i S_i$$

- $S_i \ln\left(\sum_{t=1}^{T} \lambda_{it}\right) = S_i \ln\left(\sum_{t=1}^{T} \exp\{X_{it}'\beta + \alpha_i\}\right)$

$$= S_i \ln\left(\exp\{\alpha_i\} \sum_{t=1}^{T} \exp\{X_{it}'\beta\}\right)$$

$$= S_i \alpha_i + S_i \ln\left(\sum_{t=1}^{T} \exp\{X_{it}'\beta\}\right)$$

$$\Rightarrow \ln f(Y_i \mid X_i, S_i, \beta) = \sum_{t=1}^{T} Y_{it} X_{it}'\beta - S_i \ln\left(\sum_{t=1}^{T} \exp\{X_{it}'\beta\}\right) + C_i$$

$\hat{\beta}_{CMLE}$ satisfies:

$$0 = \sum_{i=1}^{n} \left[\sum_{t=1}^{T} Y_{it} X_{it} - S_i \frac{\sum_{t=1}^{T} X_{it} \exp\{X_{it}'\hat{\beta}\}}{\sum_{t=1}^{T} \exp\{X_{it}'\hat{\beta}\}} \right]$$

- FOCs are linear in Y_{it}

Consistency of CMLE only depends on the conditional mean assumption: $E[Y_{it} \mid X_i, \alpha_i] = \exp\{X_{it}'\beta + \alpha_i\}$ (Wooldridge)

e.g.
- $E[\exp\{X_{it-1}'\beta_0\} y_{it} - y_{it-1}\exp\{X_{it}'\beta_0\} \mid X_i, \alpha_i] = 0$
- $E[\exp\{-X_{it}'\beta_0\} y_{it} - \exp\{-X_{it-1}'\beta_0\} y_{it-1} \mid X_i, \alpha_i] = 0$

$$= \exp\{-X_{it}'\beta_0\}\exp\{X_{it}'\beta_0+\alpha_i\}$$
$$-\exp\{-X_{it-1}'\beta_0\}\exp\{X_{it-1}'\beta_0+\alpha_i\}$$
$$= e^{\alpha_i} - e^{\alpha_i} = 0$$

- ie multiplicative fixed effects are okay.

Correlated Random Effects

- model for $\alpha_i | X_i$: $g(\alpha | X, \gamma)$ • this is how individual effect is determined by X.
- $\underbrace{f(y|X,\beta,\gamma)}_{\text{identified from data}} = \int \underbrace{f(y|X,\alpha,\beta)}_{\text{specify this}} \underbrace{g(\alpha|X,\gamma)}_{\text{this is the unknown}} d\alpha$
  - This is sensitive to the specification of $g(\alpha|X,\gamma)$
- Not Logit, bounded X_{it}'s, full time dummies, $T=2$, β_0 is not identified.

Large T bias correction:

$Pr[Y_{it}=1|X_i,\alpha_i] = \Phi(X_{it}'\beta_0+\alpha_i)$

$$\frac{\partial}{\partial X}\int \Phi(X'\beta+\alpha)\underbrace{f(\alpha)}_{\text{marginal distribution}}d\alpha$$ • marginal effect

$T=3$ or 4, get small biases in FE estimator of this

$$\frac{\partial}{\partial X}\frac{1}{n}\sum_{i=1}^{n}\Phi(X'\hat{\beta}_{FE}+\hat{\alpha}_i)$$

Recall, $\hat{\theta}_{FE} \xrightarrow{p} \theta_T \neq \theta_0$

$$\theta_T = \theta_0 + \frac{B}{T} + \frac{D}{T^2} + o\left(\frac{1}{T^2}\right)$$

Bias corrections for Large T

- allow T to grow with n.

$$\sqrt{nT}\,(\hat{\theta}_{FE} - \theta_T) \xrightarrow{d} N(0, \Omega)$$

$$\hat{\theta}_{FE} = \underset{\theta}{\operatorname{argmax}} \sum_{i=1}^{n} \ln f(Y_i \mid X_i, \hat{\alpha}_i(\theta), \theta)$$

$$\sqrt{nT}\,(\hat{\theta}_{FE} - \theta_0) = \sqrt{nT}\,(\hat{\theta}_{FE} - \theta_T) + \sqrt{nT}\,(\theta_T - \theta_0)$$

$$= \sqrt{nT}\,(\hat{\theta}_{FE} - \theta_T) + \sqrt{nT}\left(\frac{B}{T} + \frac{D}{T^2} + o\left(\frac{1}{T^2}\right)\right)$$

$$\sqrt{nT}\,(\theta_T - \theta_0) = \sqrt{\frac{n}{T}}\,B + \sqrt{\frac{n}{T^3}}\,D + o\left(\sqrt{\frac{n}{T^3}}\right)$$

If $\sqrt{\frac{n}{T}} \to \rho$, $\sqrt{nT}\,(\theta_T - \theta_0) = \rho B$

$$\Rightarrow \sqrt{nT}\,(\hat{\theta}_{FE} - \theta_0) \xrightarrow{d} N(\rho B, \Omega)$$

Take $\hat{\hat{\theta}} = \hat{\theta}_{FE} - \frac{\hat{B}}{T}$

Next time, we will discuss how to estimate B.