Recall: $\hat{\theta}_T = \arg\min \bar{E}\left[\sum_{t=1}^{T} \ln f(z_{it} \mid \theta, \hat{\alpha}_i(\theta))\right]$

$\hat{\theta}_T \to \theta_T$, which in general is not θ_0

- Need bias correction

$\theta_T = \theta_0 + \frac{B}{T} + O\left(\frac{1}{T^2}\right)$, where B is a bias correction term.

Fixed effects estimator:

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

- $\hat{\theta} = \arg\max \hat{Q}(\theta)$

  $\hat{Q}(\theta) \xrightarrow{p} Q(\theta) \Rightarrow \hat{\theta} \xrightarrow{p} \arg\max Q(\theta)$

  - may not be θ_0, but it is at least converging!

- $\bar{\theta} = \arg\max Q(\theta)$

  $\frac{\partial \hat{Q}}{\partial \theta} = 0 = \frac{\partial \hat{Q}}{\partial \theta}(\bar{\theta}) + \frac{\partial^2 \hat{Q}}{\partial\theta\partial\theta'}(\hat{\theta} - \bar{\theta})$

  $\Rightarrow \sqrt{nT}\,(\hat{\theta} - \bar{\theta}) = \left[-\frac{\partial^2 \hat{Q}}{\partial\theta\partial\theta'}\right]^{-1} \sqrt{nT}\,\frac{\partial \hat{Q}}{\partial \theta}(\bar{\theta})$

For $n/T \to \rho$,

$$(nT)^{1/2}(\hat{\theta}-\theta_0) = \underbrace{(nT)^{1/2}(\hat{\theta}-\theta_T)}_{\xrightarrow{d} N(0,\Omega)} + (nT)^{1/2}(\theta_T-\theta_0)$$

and $(nT)^{1/2}\left(\frac{B}{T} + O\left(\frac{1}{T^2}\right)\right)$

$$= \left(\frac{n}{T}\right)^{1/2} B + O\left((nT)^{1/2}/T^2\right)$$

$$\to \rho^{1/2} B \qquad \to 0$$

$$\Rightarrow (nT)^{1/2}(\hat{\theta}-\theta_0) \xrightarrow{d} N(\rho^{1/2}B, \Omega)$$

confidence intervals:

$$\left[\hat{\theta} \pm \bar{q}\sqrt{\frac{\hat{\Omega}}{nT}}\right] = I$$

$\Pr[\theta_0 \in I] \to$ something $\neq \tau \Pr\{[N(0,1)] \geq \bar{q}\}$

<u>analytical bias correction</u>

Construct estimator $\hat{B}$. Bias corrected estimator is $\hat{\theta}_1 = \hat{\theta} - \frac{\hat{B}}{T}$

assume $(nT)^{1/2}(\hat{B}-B)/T \xrightarrow{p} 0$

assume $n/T^3 \to 0$. This kills off the remainder term.

Plugging in,

$$(nT)^{1/2}(\hat{\theta}_1 - \theta_0) = (nT)^{1/2}(\hat{\theta} - \theta_T) + (nT)^{1/2}\left(\theta_T - \theta_0 - \frac{\hat{B}}{T}\right)$$

$$= (nT)^{1/2}(\hat{\theta} - \theta_T) + \underbrace{(nT)^{1/2}(B - \hat{B})\frac{1}{T}}_{\to 0} + \underbrace{O\left((nT)^{1/2} \cdot \frac{1}{T^2}\right)}_{\to 0}$$

$$(nT)^{1/2}(\hat{\theta} - \theta_T) \xrightarrow{d} N(0, \Omega)$$

$$\Rightarrow (nT)^{1/2}(\hat{\theta}_1 - \theta_0) \xrightarrow{d} N(0, \Omega).$$

We can iterate this, since bias formula usually depends on θ: $\hat{B} = \tilde{B}(\hat{\theta})$

- $\hat{\theta}_j = \hat{\theta} - \tilde{B}(\hat{\theta}_{j-1}) \cdot \frac{1}{T}$
- $\hat{\theta}_\infty = \hat{\theta} - \tilde{B}(\hat{\theta}_\infty) \cdot \frac{1}{T}$

Jackknife Bias Correction:

Use how $\hat{\theta}$ changes as T changes to get an implicit bias correction.

Let $\hat{\theta}_{(t)}$ denote f.e. estimator not using t^{th} time period:

$$\tilde{\hat{\theta}}^{JK} \equiv T\hat{\theta} - (T-1)\sum_{t=1}^{T}\frac{\hat{\theta}_{(t)}}{T}$$

To see why we have bias correction:

$$\theta_T = \theta_0 + \frac{B}{T} + \frac{D}{T^2} + O\left(\frac{1}{T^3}\right)$$

For T fixed:

this expression wipes out B term

$$\tilde{\hat{\theta}} \xrightarrow{P} T\theta_T - (T-1)\theta_{T-1}$$

$$= \theta_0 + \left(\frac{1}{T} - \frac{1}{T-1}\right)D + O\left(\frac{1}{T^2}\right)$$

$$= \theta_0 + O\left(\frac{1}{T^2}\right)$$

$$y_{it}^* = X_{it}'\beta + \alpha_i - \varepsilon_{it}, \quad \varepsilon_{it} \sim N(0,1)$$

$$y_{it} = 1\{y_{it}^* > 0\}$$

$$\Pr[y_{it}=1 \mid X_{it}, \alpha_i, \beta] = \Phi(X_{it}'\beta_0 + \alpha_i)$$

- $\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n}\frac{\partial \Phi}{\partial X_{it}}(X_{it}'\hat{\beta} + \hat{\alpha}_i)$

- averaging over individuals the effect of X on the probabilities

- using just those fixed effects, we find that there is very little bias for these marginal effects.

- Bias - corrected estimator:
  - $\hat{\theta}_1 = \hat{\theta} - \frac{\hat{B}}{T}$

Bias - corrected objective fcn:
  - $\hat{L}(\theta) = \frac{1}{n} \sum_{i=1}^{n} \sum_{t=1}^{T} \ln f(z_{it} | \theta, \hat{\alpha}_i(\theta)) \cdot \frac{1}{T}$
  - $\hat{L}(\theta) - \frac{\hat{D}(\theta)}{T}$

Let $\alpha_i(\theta) = \underset{\alpha_i}{\operatorname{argmax}} E\left[\frac{1}{T} \sum_{t=1}^{T} \ln f(z_{it} | \theta, \alpha)\right]$

- Expand $\hat{L}$ around same thing with $\alpha_i(\theta)$ instead of $\hat{\alpha}_i(\theta)$

$$\hat{L} = \frac{1}{Tn} \sum_{i=1}^{n} \sum_{t=1}^{T} \ln f(z_{it} | \theta, \alpha_i) + \frac{1}{2n} \sum_{i=1}^{n} \left[\frac{1}{T} \sum_{t=1}^{T} \frac{\partial^2 \ln f(z_{it} | \theta, \tilde{\alpha}_i)}{\partial \alpha_i^2}\right] (\hat{\alpha}_i - \alpha_i)^2$$

The first-order condition for $\hat{\alpha}_i$ is

$$\frac{\partial}{\partial \alpha_i} \frac{1}{T} \sum_{t=1}^{T} \ln f(z_{it} | \theta, \hat{\alpha}_i) = 0$$

$$\Rightarrow \quad \hat{\alpha}_i - \alpha_i = \left[-\frac{\partial^2}{\partial \alpha_i^2} \frac{1}{T} \sum_{t=1}^{T} \ln f(z_{it} | \theta, \tilde{\alpha}_i)\right]^{-1} \frac{1}{T} \sum_{t=1}^{T} \frac{\partial \ln f(z_{it}, \theta_{\alpha_i})}{\partial \alpha_i}$$

- There is a simple expression for the variance of $\hat{\alpha}_i$

- $\hat{D}(\theta) = \frac{1}{2n}\sum_{i=1}^{n}\left\{\left[\frac{1}{T}\sum_{t=1}^{T}\frac{\partial^2 \ln f(z_{it},\theta,\hat{\alpha}_i(\theta))}{\partial \alpha_i^2}\right]^{-1}\frac{1}{T}\sum_{t=1}^{T}\left[\frac{\partial \ln f(z_{it},\theta,\hat{\alpha}_i(\theta))}{\partial \alpha_i}\right]^2\right\}$

$$\Rightarrow \underset{\substack{\uparrow \\ \text{bias-corrected}}}{\hat{\theta}^{BC}} = \operatorname{argmax}\left(\hat{L}(\theta) - \frac{\hat{D}(\theta)}{T}\right)$$

You can jackknife correct the objective function as well

- Jackknife estimator: $T\,\hat{\theta}_T - (T-1)\frac{1}{T}\sum_{t=1}^{T}\hat{\theta}_{-t}$
- Jackknife objective fcn: $T\hat{L}(\theta) - (T-1)\frac{1}{T}\sum_{t=1}^{T}\hat{L}_{-t}(\theta)$,

where $\hat{L}_{-t}(\theta) = \sum_{i=1}^{n}\max_{\alpha_i}\frac{1}{T-1}\sum_{s\neq t}\ln f(z_{is},\theta,\alpha_i)\cdot\frac{1}{n}$

Can do a simpler version of the jackknife that preserves time series structure (cannot just drop t^{th} observation.)