Mean: 91

Extremum Estimators (consistency/normality)

Know the necessary assumptions!

a good source is the Newey / McFadden chapter of the Handbook of Econometrics (~94)

◦ may be posted on Juergen's website.

Example: $X_i \sim \text{Poisson}(\theta_0)$ iid $\Rightarrow f_{X_i}(x_i, \theta) = \dfrac{e^{-\theta}\theta^{x_i}}{x_i!}$, $x_i = 0, 1, 2, \ldots$

What is $\hat{\theta}_n^{MLE}$ and what are its asymptotic properties?

$$L(\theta, x_1, \ldots, x_n) = \prod_{i=1}^{n} \frac{e^{-\theta}\theta^{x_i}}{x_i!} = \frac{e^{-n\theta}\theta^{\Sigma x_i}}{\prod x_i!}$$

$$\Rightarrow \ln L = -n\theta + \left(\Sigma x_i\right)\ln\theta - \ln \prod_{i=1}^{n} x_i!$$

FOC: $$\frac{\partial \ln L}{\partial \theta} = -n + \frac{\sum_{i=1}^{n} x_i}{\hat{\theta}_n^{MLE}} = 0 \Rightarrow \hat{\theta}_n^{MLE} = \frac{1}{n}\sum_{i=1}^{n} x_i \equiv \bar{X}_n$$

Now, using the extremum estimator framework:

$$\hat{\theta}_n \equiv \underset{\theta \in \Theta}{\arg\max}\ Q_n(\theta)$$

Here, $$Q_n(\theta) = \ln L(\theta, x_1, \ldots, x_n) = \frac{1}{n}\sum_{i=1}^{n} \ln\left[\frac{e^{-\theta}\theta^{x_i}}{x_i!}\right]$$

$$\left.\begin{aligned} &= \frac{1}{n}\sum_{i=1}^{n} m(w_i, \theta) \\ &= \frac{1}{n}\sum_{i=1}^{n} \ln f_{X_i}(x_i, \theta) \end{aligned}\right\} \text{notation.}$$

Consistency (Thm 1.2): Check the assumptions.

(i) Θ compact $\Leftrightarrow$ Θ is closed and bounded

◦ In practice, we never really know whether or not this is satisfied.

◦ permitive condition - just assume it

(ii) $Q_n(w_i\ \theta)$ is continuous in θ $\forall w_i$.

Eg: $Q_n(w_i, \theta) = \frac{1}{n}(-n\theta + (\sum_{i=1}^{n} x_i)\ln(\theta) - \sum_{i=1}^{n}\ln(x_i!))$

└ Yes, this is continuous

(iii) $Q_n(w_i, \theta)$ is measurable in w_i $\forall\theta$. (ignore this)

(iv) This is tricky. Two parts:

(a) Identification

(b) Uniform convergence in probability

(a) Want to show: $\theta \neq \theta_0 \Rightarrow f_{X_i}(x_i, \theta) \neq f_{X_i}(x_i, \theta_0)$

Let $\theta \neq \theta_0 \Rightarrow \frac{f_{X_i}(x_i,\theta)}{f_{X_i}(x_i,\theta_0)} = \frac{(e^{-\theta}\theta^{x_i})/x_i!}{(e^{-\theta_0}\theta_0^{x_i})/x_i!} = \underbrace{\underbrace{e^{-(\theta-\theta_0)}}_{\neq 1 \text{ if } \theta\neq\theta_0}\underbrace{\left[\frac{\theta}{\theta_0}\right]^{x_i}}_{\neq 1 \text{ if } \theta\neq\theta_0}}_{\neq 1 \text{ if } \theta\neq\theta_0} \neq 1$

$\Rightarrow f_{X_i}(x_i, \theta) \neq f_{X_i}(x_i, \theta_0)$.

└ this is not trivial, though

(b) uniform convergence: (Proposition 1.3)

Need:
- X_i iid — cannot check this
- Θ compact — already satisfied
- $\ln f(x_i, \theta)$ continuous in θ — already satisfied
- $E\left[\sup_{\theta\in\Theta} |\ln f_{X_i}(w_i,\theta)|\right] < +\infty$

dominated convergence theorem $\Rightarrow \sup_{\theta\in\Theta}\left|\frac{1}{n}\sum_{i=1}^{n}\ln f_{X_i}(w_i,\theta) - E[\ln f(X_i,\theta)]\right| \xrightarrow{P} 0$

Then, $\hat{\theta}_n^{MLE} \xrightarrow{P} \theta$. □

Normality (Thm 1.5):

Recall: $Q_n(\theta) = \frac{1}{n}\sum_{i=1}^{n}\ln f(x_i,\theta)$

$= \frac{1}{n}\sum \ln\left(\frac{e^{-\theta}\theta^{x_i}}{x_i!}\right) = \frac{1}{n}$

$m(x_i, \theta) \equiv \ln f(x_i, \theta) = -\theta + x_i\ln\theta - \ln x_i!$.

based on a single observation. $s(x_i,\theta) = \frac{\partial m}{\partial \theta}(x_i,\theta) = -1 + \frac{x_i}{\theta}$

Mean value theorem around θ_0:

$$0 = \underbrace{\frac{1}{n}\sum_{i=1}^{n} s(X_i,\hat{\theta}_n^{MLE})}_{=0 \text{ by FOC}} = \frac{1}{n}\sum_{i=1}^{n} s(X_i,\theta_0) + \frac{1}{n}\sum_{i=1}^{n} \underbrace{\frac{\partial s(X_i,\hat{\theta})}{\partial \theta'}}_{\equiv H(x_i,\hat{\theta})} (\hat{\theta}_n^{MLE} - \theta_0)$$

where $\hat{\theta} \in |\hat{\theta}_n^{MLE}, \theta_0|$

Notation for Hessians

$g(\tau_1,\tau_2) = 3\tau_1^2\tau_2 \iff g(\tau),\ \tau = \begin{bmatrix}\tau_1\\ \tau_2\end{bmatrix}$

$$\frac{\partial g(\tau)}{\partial \tau} = \begin{bmatrix}\frac{\partial g(\tau)}{\partial \tau_1}\\ \frac{\partial g(\tau)}{\partial \tau_2}\end{bmatrix} = \begin{bmatrix}6\tau_1\tau_2\\ 3\tau_1^2\end{bmatrix}$$

$$\frac{\partial^2 g(\tau)}{\partial \tau \partial \tau'} = \frac{\partial \begin{bmatrix}6\tau_1\tau_2\\ 3\tau_1^2\end{bmatrix}}{\partial \tau'} = \begin{bmatrix}\frac{\partial 6\tau_1\tau_2}{\partial \tau_1} & \frac{\partial 6\tau_1\tau_2}{\partial \tau_2}\\ \frac{\partial 3\tau_1^2}{\partial \tau_1} & \frac{\partial 3\tau_1^2}{\partial \tau_2}\end{bmatrix} = \begin{bmatrix}6\tau_2 & 6\tau_1\\ 6\tau_1 & 0\end{bmatrix}$$

$$\Rightarrow \sqrt{n}(\hat{\theta}_n^{MLE} - \theta_0) = -\underbrace{\left[\frac{1}{n}\sum_{i=1}^{n} H(X_i,\hat{\theta})\right]^{-1}}_{(1)} \left[\underbrace{\sqrt{n}\left[\frac{1}{n}\sum_{i=1}^{n} s(X_i,\theta_0)\right]}_{(2)}\right]$$

$$\left\{ \text{cf } \left(\frac{1}{n}\sum_{i=1}^{n} X_i' X_i\right)\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} X_i'\varepsilon_i\right) \right\}$$

Since $\hat{\theta}_n^{MLE} \xrightarrow{P} \theta_0$, and $\hat{\theta} \in |\hat{\theta}_n^{MLE}, \theta_0|$, $\hat{\theta} \xrightarrow{P} \theta_0$

(1) $\frac{1}{n}\sum_{i=1}^{n} H(X_i,\hat{\theta}) \xrightarrow{P} E[H(X_i,\theta_0)]$ if $E\left[\sup_{\theta\in\Theta}\|H(X_i,\theta)\|\right] < +\infty$

ULLN

$\ln f(X_i,\theta) \in C^2$

By the continuity theorem $\left[\frac{1}{n}\sum_{i=1}^{n} H(X_i,\hat{\theta})\right]^{-1} \xrightarrow{P} [\underbrace{E[H(X_i,\theta_0)]}_{=H_0}]^{-1}$

(2) : $\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} s(X_i, \theta_0)\right) \xrightarrow{d} N(0, \Omega_0)$ by CLT

where $\Omega_0 \equiv Var[s(X_i, \theta_0)]$

$= E[s(X_i, \theta_0) s(X_i, \theta_0)']$

$= I_0 = -H_0$

CLT holds since we have iid data and if we assume $\Omega_0 < +\infty$

Then, by Slutsky's theorem,

$\sqrt{n}(\hat{\theta}_n^{MLE} - \theta_0) \xrightarrow{d} N(0, (-H_0)^{-1}\Omega_0(-H_0)^{-1})$

Here, since $\Omega_0 = -H_0 = I_0$,

$\sqrt{n}(\hat{\theta}_n^{MLE} - \theta_0) \xrightarrow{d} N(0, (I_0)^{-1} I_0 (I_0^{-1}))$

$= N(0, (I_0)^{-1})$

Recall: $\ln f(X_i, \theta_0) = -\theta + x_i \ln \theta - \ln x_i!$

$s(X_i, \theta_0) = \frac{x_i}{\theta} - 1$

$H_0 \equiv E\left[-\frac{x_i}{\theta_0^2}\right] = -\frac{1}{\theta_0}$

$\Rightarrow \sqrt{n}(\hat{\theta}_n^{MLE} - \theta_0) \xrightarrow{d} N(0, \theta_0)$ for our case.