Lec 2: Consistency

Eric Hsienchen Chu*

Spring, 2024

(⊛) Suggested reading: Newey and McFadden (1994), Section 2

1 Consistency

Motivation. We know $\hat{\mathbb{Q}}_n(\theta) \xrightarrow{p} \mathbb{Q}_0(\theta)$ *pointwise* in θ by WLLN. But does it sufficiently imply $\hat{\theta} \xrightarrow{p} \theta_0$? The answer is **NO!** We need $\sup_{\theta\in\Theta}\left|\hat{\mathbb{Q}}_n(\theta) - \mathbb{Q}_0(\theta)\right| \xrightarrow{p} 0$ (uniform consistency in θ) + "regularity conditions". Thus, let's begin by showing **Consistency Theorem**.

> **Theorem 1.1** (Consistency of $\hat{\theta}$)**.** $\hat{\theta} \xrightarrow{p} \theta_0$ if:
>
> (i) θ_0 is **unique** maximizer for $\mathbb{Q}_0$ (*identification*),
>
> (ii) Θ is compact,
>
> (iii) Q_0 is continuous in θ (parameter of interest), and
>
> (iv) $\hat{\mathbb{Q}}_n$ is **uniformly consistent** for $\mathbb{Q}_0$.

Proof. For any $\varepsilon > 0$, we know $\hat{\mathbb{Q}}_n(\hat{\theta}) \geq \hat{\mathbb{Q}}_n(\theta_0) > \hat{\mathbb{Q}}_n(\theta_0) - \varepsilon$, since $\hat{\theta}$ is maximizer of $\hat{\mathbb{Q}}_n$. By (iv), as $n \to \infty$, for any $\theta \in \Theta$, we have $\left|\hat{\mathbb{Q}}_n(\theta) - \mathbb{Q}_0(\theta)\right| < \varepsilon$ with probability 1.

$$\Longrightarrow \begin{cases} \hat{\mathbb{Q}}_n(\hat{\theta}) - \mathbb{Q}_0(\hat{\theta}) < \varepsilon & \Longrightarrow \mathbb{Q}_0(\hat{\theta}) > \hat{\mathbb{Q}}_n(\hat{\theta}) - \varepsilon \\ \mathbb{Q}_0(\theta_0) - \hat{\mathbb{Q}}_n(\theta_0) < \varepsilon & \Longrightarrow \hat{\mathbb{Q}}_n(\theta_0) > \mathbb{Q}_0(\theta_0) - \varepsilon \end{cases} \tag{1.1}$$

Then, as $n \to \infty$, we have:

$$\mathbb{Q}_0(\hat{\theta}) > \hat{\mathbb{Q}}_n(\hat{\theta}) - \varepsilon > (\hat{\mathbb{Q}}_n(\theta_0) - \varepsilon) - \varepsilon \tag{1.2}$$
$$> (\mathbb{Q}_0(\theta_0) - \varepsilon) - 2\varepsilon = \mathbb{Q}_0(\theta_0) - 3\varepsilon, \text{ with probability 1} \tag{1.3}$$

*Department of Economics, University of Wisconsin-Madison. hchu38@wisc.edu. This is lecture notes from the second half of ECON710: Economic Statistics and Econometrics II. Instructor: Prof. Harold Chiang. Materials and sources: Harold's handwritten notes.

Let $\mathcal{N}$ be open set s.t. $\theta_0 \in \mathcal{N} \subseteq \Theta$, then $\mathcal{N}^c := \Theta \cap \mathcal{N}^c$ is compact by (ii) (closed subset of a compact set). Therefore,

$$\exists \theta^* \in \mathcal{N}^c \text{ s.t. } \sup_{\theta \in \Theta} \mathbb{Q}_0(\theta) = \mathbb{Q}_0(\theta^*) \longleftarrow \text{by (iii) continuity} \tag{1.4}$$

$$< \mathbb{Q}_0(\theta_0) \longleftarrow \text{by } \theta_0 = \arg\max \mathbb{Q}_0 \tag{1.5}$$

We now can pick our $\varepsilon = \frac{1}{3}\left[\mathbb{Q}_0(\theta_0) - \mathbb{Q}_0(\theta^*)\right]$ (> 0) so that, as $n \to \infty$, equation (4) yields $\mathbb{Q}_0(\hat{\theta}) > \mathbb{Q}_0(\theta^*)$ with probability 1, i.e., $\hat{\theta} \notin \mathcal{N}^c$ w.p.1 $\implies \hat{\theta} \in \mathcal{N}$ w.p.1 $\implies \hat{\theta} \xrightarrow{p} \theta_0$. □

Remark. Only Condition (i) *uniqueness of maximizer* θ_0 is required. This makes sure that our estimator $\hat{\theta}$ is centering at the true maximizer θ_0 (and therefore *consistent*), not multiple $\hat{\theta}$ and being inconsistent.

Question. How do we check the Consistency conditions?

Answer. (i) depends case-by-case; (ii) holds normally by assumption; (iii) & (iv) jointly implied by **Uniform LLN (ULLN)**

2 ULLN

Motivation. We rely on **ULLN** to determine Condition (iii) & (iv) in **Consistency Theorem** so that we make sure our estimator $\hat{\theta}$ is consistent for θ_0.

Theorem 2.1 (ULLN). Suppose $(Z_i)_{i=1}^n \overset{iid}{\sim} Z$ and Θ compact. If:

(i) $\theta \longmapsto g(Z;\theta)$ is continuous (a.e.) $\forall \theta \in \Theta$, and

(ii) $\exists$ a function $\zeta \longmapsto h(\zeta)$ s.t. $\begin{cases} \left|g(\zeta;\theta)\right| \leq h(\zeta) \ \forall \theta \in \Theta, & \text{(i.e., } h(\zeta) \textit{ dominating func w/o param)} \\ \mathbb{E}[h(Z)] < \infty \end{cases}$

Then,

① $\theta \longmapsto \mathbb{E}[g(Z;\theta)]$ is continuous in θ ($\longleftarrow$ Condition (iii) ✓)

② $\sup_{\theta \in \Theta} \left| \frac{1}{n} \sum_{i=1}^n g(Z_i;\theta) - \mathbb{E}[g(Z;\theta)] \right| \xrightarrow{p} 0$ ($\longleftarrow$ Condition (iv) ✓)

Proof. (Harold: "Take ECON715") □

Remark. Recall, [**Lec 1**] *uniform consistency* tells us $\sup_{\theta \in \Theta} \left|\hat{\mathbb{Q}}_n \theta) - \mathbb{Q}_0(\theta)\right| \xrightarrow{p} 0 \rightsquigarrow o_p(1)$.

Here, we use $\frac{1}{n} \sum_{i=1}^n g(Z_i;\theta)$ as $\hat{\mathbb{Q}}_n(\theta)$ and use $\mathbb{E}[g(Z;\theta)]$ as true objective function $\mathbb{Q}_0(\theta)$.

Example 2.1 (NLS ✓). Consider $\mathbb{Q}_0(\theta) = -\mathbb{E}[(Y-\mu(x;\theta))^2] =: \mathbb{E}[g(Z;\theta)]$, then ULLN is applicable if μ satisfies ULLN (i) & (ii).

Example 2.2 (MLE ✓). Consider $\mathbb{Q}_0(\theta) = \mathbb{E}[\ell n f(Z;\theta)] =: \mathbb{E}[g(Z;\theta)]$, where $Z \sim^d f(Z;\theta)$ and f is known up to θ (pdf).

Example 2.3 (GMM ×). Goal: $\mathbb{E}[g(Z;\theta_0)] = 0$
$\implies$ Consider $\mathbb{Q}_0(\theta) = -\mathbb{E}[g(Z;\theta)]' \mathbf{W}\mathbb{E}[g(Z;\theta)]$, which is a *quadratic* form of $\mathbb{E}[g(Z;\theta)]$
$\implies$ *Cannot directly apply ULLN!*

Summary. Therefore, we can categorize above discussion into:

- MLE-type (★)
- GMM-type ("minimum distance"): collapse to MLE-type when Just-ID case.

3 Consistency of MLE

Theorem 3.1 (Consistency; MLE). Suppose $(Z_i)_{i=1}^n \overset{iid}{\sim} f(\zeta;\theta)$, where f : known pdf given $\theta \in \Theta$, then $\hat{\theta} \xrightarrow{p} \theta_0$ if:

① $\theta \neq \theta_0 \implies f(Z;\theta) \neq f(Z;\theta_0)$ (i.e., different density) ,

② Θ is compact,

③ $\theta \longmapsto \ell n f(Z;\theta)$ is continuous (a.e.) $\forall \theta \in \Theta$ and Z_i, and

④ $\mathbb{E}\left[\sup_{\theta\in\Theta}\left|\ell n f(Z;\theta)\right|\right] < \infty$.

Proof. Since Theorem 1.1 (Consistency) condition (ii) is checked by ②, we now need to verify condition (i): unique maximizer **[Spring 2023 Final Q1]**.
Recall that $\mathbb{Q}_0 := \mathbb{E}[\ell n f(Z;\theta)]$ for MLE.
$\implies$ **WTS.** $\mathbb{Q}_0(\theta_0) > \mathbb{Q}_0(\theta)\ \forall \theta \neq \theta_0$ (⊛ intuition: $\theta_0 = \arg\max_\theta \mathbb{Q}_0$)

$$\begin{aligned}
\mathbb{Q}_0(\theta) - \mathbb{Q}_0(\theta_0) &= \mathbb{E}[\ell n f(Z;\theta) - \ell n f(Z;\theta_0)] = \mathbb{E}\left[\ell n\left(\frac{f(Z;\theta)}{f(Z;\theta_0)}\right)\right] &(3.1)\\
&< \ell n \mathbb{E}\left[\frac{f(Z;\theta)}{f(Z;\theta_0)}\right] \longleftarrow \text{"<" holds by Jensen's Ineq \& ① diff density } \forall\theta\neq\theta_0 &(3.2)\\
&= \ell n \int \frac{f(\zeta;\theta)}{\cancel{f(\zeta;\theta_0)}} \cdot \underbrace{\cancel{f(\zeta;\theta_0)}}_{\text{true pdf}} d\zeta = \ell n \int f(\zeta;\theta) d\zeta = \ell n 1 = 0 &(3.3)
\end{aligned}$$

Thus, we verify $\mathbb{Q}_0(\theta_0) > \mathbb{Q}_0(\theta)\ \forall\theta \neq \theta_0$, i.e., θ_0 is unique maximizer & condition (i) (✓). We now use ③ & ④ to check if ULLN is applicable so that Consistency condition (iii) & (iv) will be jointly satisfied.

- Define $g(\zeta;\theta) := \ell n f(\zeta;\theta)$, then by ③ we note $g(Z;\theta)$ is conti. $\forall\theta\in\Theta$ & Z w.p.1.
- Let $h(\zeta) = \sup_{\theta\in\Theta}\Big|\ell n f(\zeta;\theta)\Big|$, then $\begin{cases}\Big|g(\zeta;\theta)\Big| = \Big|\ell n f(\zeta;\theta)\Big| \le \sup_{\theta\in\Theta}\Big|\ell n f(\zeta;\theta)\Big| = h(\zeta;\theta)\ \forall\theta\in\Theta \\ \mathbb{E}[h(Z)] = \mathbb{E}\left[\sup_{\theta\in\Theta}\Big|\ell n f(Z;\theta)\Big|\right] < \infty \longleftarrow \text{by ④}\end{cases}$

So, ULLN is satisfied, and by ULLN we know Consistency condition (iii) & (iv) (✓). Ultimately, by Theorem 1.1, we conclude MLE is consistent. □

4 Exercise from DIS SEC

Exercise 4.1 (Spring24 TA Handout7 Ex3)**.** Consider the simple linear model $Y_i = \beta_0 X_i + e_i$ where $\mathbb{E}[e_i|X_i] = 0$. Here Y_i and X_i are scalars with $\mathbb{E}[Y_i^4] < \infty$ and $\mathbb{E}[X_i^4] < \infty$. Take the parameter space $\Theta = [-1,1]$ and assume $\beta_0 \in int(\Theta)$. Then, we define an M–estimator for β as follows:

$$\hat{\beta} = \operatorname*{argmin}_{\beta\in[-1,1]} S_n(\beta) \text{ where } S_n(\beta) = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \beta X_i)^2 \tag{4.1}$$

(a) Show that $\sup_{\beta\in[-1,1]}\Big|S_n(\beta) - S(\beta)\Big| \xrightarrow{p} 0$ where $S(\beta) = \mathbb{E}[(Y_i - \beta X_i)^2]$.

(b) Show that $\hat{\beta} \xrightarrow{p} \beta_0$.

Solution (a)**.** Essentially, we want to invoke **ULLN**. For ULLN (i): *continuous in parameter*, we see that $(Y_i - \beta X_i)^2$ is continuous in β (✓). To check ULLN (ii) *dominating function w/o parameter*, we define $g(Y_i, X_i;\beta) = (Y_i - \beta X_i)^2$ and find that:

$$\begin{aligned}\Big|g(Y_i,X_i;\beta)\Big| = \Big|(Y_i-\beta X_i)^2\Big| &= (Y_i - \beta X_i)^2 && (4.2)\\ &\le 2(Y_i^2 + \beta^2 X_i^2) \longleftarrow \text{by CR Ineq.} && (4.3)\\ &\le 2(Y_i^2 + 1^2 X_i^2) \longleftarrow \text{by } \beta\in[-1,1] \Rightarrow \beta^2\in[0,1]. && (4.4)\end{aligned}$$

So, we can let $h(Y_i, X_i) = 2(Y_i^2 + X_i^2)$ and that ULLN (ii) is satisfied by $\Big|g(Y_i,X_i;\beta)\Big| \le h(Y_i,X_i)$, with $\mathbb{E}[h(Y_i,X_i)] = 2(\mathbb{E}[Y_i^2] + \mathbb{E}[X_i^2]) < \infty$ (since 4^{th} moments exist). By Theorem 2.1 (ULLN), the statement is true.

Solution (b)**.** Since $\Theta = [-1,1] \subset R$ is closed and bounded, by Heine-Borel Theorem we know Θ is compact. We now only need to check Theorem 1.1 (Consistency) (i): β_0 *being* *unique minimizer for* $S(\beta)$[1]. First note that $S(\beta_0) = \mathbb{E}[(Y_i - \beta_0 X_i)^2] = \mathbb{E}[e_i^2]$. Then, as we take any $\tilde{\beta} \neq \beta_0$, we find that:

$$\begin{aligned}
S(\tilde{\beta}) &= \mathbb{E}[(Y_i - \tilde{\beta} X_i)^2] && (4.5)\\
&= \mathbb{E}[(\beta_0 X_i + e_i - \tilde{\beta} X_i)^2] \longleftarrow \text{plugging in } Y_i. && (4.6)\\
&= \mathbb{E}[((\beta_0 - \tilde{\beta}) X_i + e_i)^2] && (4.7)\\
&= (\beta_0 - \tilde{\beta})^2 \mathbb{E}[X_i^2] + \mathbb{E}[e_i^2] + \underbrace{2(\beta_0 - \tilde{\beta})\mathbb{E}[X_i e_i]}_{= 0 \text{ by } \mathbb{E}[e_i | X_i] = 0} && (4.8)\\
&= (\beta_0 - \tilde{\beta})^2 \mathbb{E}[X_i^2] + \mathbb{E}[e_i^2] && (4.9)\\
&> \mathbb{E}[e_i^2] = S(\beta_0) && (4.10)
\end{aligned}$$

So, we verify that β_0 is unique minimizer of $S(\beta)$. By Theorem 1.1 (Consistency), $\hat{\beta} \xrightarrow{p} \beta_0$.

References

Newey, W. K., & McFadden, D. (1994). Chapter 36 large sample estimation and hypothesis testing. Elsevier. https://doi.org/10.1016/S1573-4412(05)80005-4

[1]Here, we construct $\hat{\beta} = \arg\min S_n(\beta)$ rather than $\arg\max$, so we need to verify unique MINIMIZER.