L-functional and L-estimator

For a function \( J(t) \) on \([0,1]\), define the L-functional as

\[
T(G) = \int xJ(G(x))dG(x), \quad G \in \mathcal{F}.
\]

If \( X_1, \ldots, X_n \) are i.i.d. from \( F \) and \( T(F) \) is the parameter of interest, \( T(F_n) \) is called an L-estimator of \( T(F) \).

\( T(F_n) \) is a linear function of order statistics:

\[
T(F_n) = \int xJ(F_n(x))dF_n(x) = \frac{1}{n} \sum_{i=1}^{n} J\left(\frac{i}{n}\right) X_{(i)},
\]

since \( F_n(X_{(i)}) = i/n, i = 1, \ldots, n \).

Examples

- When \( J(t) \equiv 1 \), \( T(F_n) = \bar{X} \), the sample mean.
- When \( J(t) = (1 - 2\alpha)^{-1} I_{(\alpha,1-\alpha)}(t) \), \( T(F_n) = \bar{X}_\alpha \) is the \( \alpha \)-trimmed sample mean.
Lecture 17: L-estimators and trimmed sample mean

L-functional and L-estimator

For a function $J(t)$ on $[0,1]$, define the L-functional as

$$T(G) = \int x J(G(x)) dG(x), \quad G \in \mathcal{F}.$$  

If $X_1, \ldots, X_n$ are i.i.d. from $F$ and $T(F)$ is the parameter of interest, $T(F_n)$ is called an L-estimator of $T(F)$.

$T(F_n)$ is a linear function of order statistics:

$$T(F_n) = \int x J(F_n(x)) dF_n(x) = \frac{1}{n} \sum_{i=1}^{n} J\left(\frac{i}{n}\right) X_{(i)},$$

since $F_n(X_{(i)}) = i/n$, $i = 1, \ldots, n$.

Examples

- When $J(t) \equiv 1$, $T(F_n) = \bar{X}$, the sample mean.
- When $J(t) = (1 - 2\alpha)^{-1} I_{(\alpha, 1-\alpha)}(t)$, $T(F_n) = \bar{X}_\alpha$ is the $\alpha$-trimmed sample mean.
Although the sample median is also a linear function of order statistics, it is not of the form $T(F_n)$ with an L-functional $T$.

**Asymptotic normality of L-estimators**

To establish the asymptotic normality for L-estimators $T(F_n)$, we follow the following steps.

**Step 1.** For $x \in \mathbb{R}$, calculate

$$
\phi_F(x) = \lim_{t \to 0} \frac{T(F + t(\delta_x - F)) - T(F)}{t}
$$

(if it exists), where $\delta_x$ is the point mass at $x$.

The function $\phi_F$ is called the influence function of $T$ at $F$. The influence function is an important tool in the study of robustness of estimators.

Also, verify that

$$
E[\phi_F(X_1)] = \int \phi_F(x) dF(x) = 0
$$
Although the sample median is also a linear function of order statistics, it is not of the form $T(F_n)$ with an L-functional $T$.

**Asymptotic normality of L-estimators**

To establish the asymptotic normality for L-estimators $T(F_n)$, we follow the following steps.

**Step 1.** For $x \in \mathbb{R}$, calculate

$$
\phi_F(x) = \lim_{t \to 0} \frac{T(F + t(\delta_x - F)) - T(F)}{t}
$$

(if it exists), where $\delta_x$ is the point mass at $x$.

The function $\phi_F$ is called the influence function of $T$ at $F$.

The influence function is an important tool in the study of robustness of estimators.

Also, verify that

$$
E[\phi_F(X_1)] = \int \phi_F(x) dF(x) = 0
$$
Asymptotic normality of L-estimators

Step 2. Verify that $E[\phi_F(X_1)]^2 < \infty$ and obtain

$$\sigma_F^2 = E[\phi_F(X_1)]^2 = \int [\phi_F(x)]^2 dF(x).$$

Step 3. Verify that

$$T(F_n) - T(F) = \frac{1}{n} \sum_{i=1}^{n} \phi_F(X_i) + o_p \left( \frac{1}{\sqrt{n}} \right).$$

This holds when $T$ is differentiable in some sense (§5.2.1). Then

$$\sqrt{n}[T(F_n) - T(F)] \to_d N(0, \sigma_F^2).$$

Step 3 is the most difficult part.

This approach can also be applied to other functionals (§5.2).

We now apply this approach to show the asymptotic normality of the trimmed sample mean.
Asymptotic normality of L-estimators

Step 2. Verify that $E[\phi_F(X_1)]^2 < \infty$ and obtain

$$
\sigma_F^2 = E[\phi_F(X_1)]^2 = \int [\phi_F(x)]^2 dF(x).
$$

Step 3. Verify that

$$
T(F_n) - T(F) = \frac{1}{n} \sum_{i=1}^{n} \phi_F(X_i) + o_p \left( \frac{1}{\sqrt{n}} \right).
$$

This holds when $T$ is differentiable in some sense (§5.2.1). Then

$$
\sqrt{n}[T(F_n) - T(F)] \to_d N(0, \sigma_F^2).
$$

Step 3 is the most difficult part.

This approach can also be applied to other functionals (§5.2).

We now apply this approach to show the asymptotic normality of the trimmed sample mean.
Asymptotic normality of L-estimators

Step 2. Verify that \( E[\phi_F(X_1)]^2 < \infty \) and obtain
\[
\sigma_F^2 = E[\phi_F(X_1)]^2 = \int [\phi_F(x)]^2 dF(x).
\]

Step 3. Verify that
\[
T(F_n) - T(F) = \frac{1}{n} \sum_{i=1}^{n} \phi_F(X_i) + o_p \left( \frac{1}{\sqrt{n}} \right).
\]
This holds when \( T \) is differentiable in some sense (§5.2.1). Then
\[
\sqrt{n}[T(F_n) - T(F)] \to_d N(0, \sigma_F^2).
\]
Step 3 is the most difficult part.

This approach can also be applied to other functionals (§5.2).

We now apply this approach to show the asymptotic normality of the trimmed sample mean.
Step 1: Derivation of the influence function $\phi_F$

\[ T(G) = \int xJ(G(x))dG(x), \quad G \in \mathcal{F} \]

For $F$ and $G$ in $\mathcal{F}$,

\[
T(G) - T(F) = \int xJ(G(x))dG(x) - \int xJ(F(x))dF(x)
\]

\[
= \int_0^1 [G^{-1}(t) - F^{-1}(t)]J(t)dt
\]

\[
= \int_0^1 \int_{F^{-1}(t)}^{G^{-1}(t)} dxJ(t)dt
\]

\[
= \int_{-\infty}^{\infty} \int_{G(x)}^{F(x)} J(t)dtdx
\]

\[
= \int_{-\infty}^{\infty} [F(x) - G(x)]J(F(x))dx
\]

\[
- \int_{-\infty}^{\infty} U_G(x)[G(x) - F(x)]J(F(x))dx,
\]
Step 1: Derivation of the influence function $\phi_F$

where

$$U_G(x) = \begin{cases} \frac{\int_F G(x) J(t) dt}{[G(x) - F(x)] J(F(x))} - 1 & G(x) \neq F(x), J(F(x)) \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

and the fourth equality follows from Fubini’s theorem and the fact that the region in $\mathbb{R}^2$ between curves $F(x)$ and $G(x)$ is the same as the region in $\mathbb{R}^2$ between curves $G^{-1}(t)$ and $F^{-1}(t)$.

Let $G = F + t(\delta_x - F)$, where $\delta_x$ is the degenerated distribution at $x$.

Since $\lim_{t \to 0} U_{F+t(\delta_x-F)}(y) = 0$, by the dominated convergence theorem,

$$\lim_{t \to 0} \int_{-\infty}^{\infty} U_{F+t(\delta_x-F)}(y)[\delta_x(y) - F(y)] J(F(y)) dy = 0.$$  

Hence

$$\lim_{t \to 0} \frac{T(F + t(\delta_x - F)) - T(F)}{t} = -\int_{-\infty}^{\infty} [\delta_x(y) - F(y)] J(F(y)) dy,$$

which is $\phi_F(x)$, the influence function of $T$. 
Step 1: Derivation of the influence function $\phi_F$

By Fubini’s theorem and the fact that $\int \delta_x(y) dF(x) = F(y)$,

$$
\int \phi_F(x) dF(x) = - \int_{-\infty}^{\infty} \left[ \int (\delta_x - F(y)) dF(x) \right] J(F(y)) dy = 0,
$$

Consider now $J(t) = (\beta - \alpha)^{-1} l_{(\alpha, \beta)}(t)$,

$$
\phi_F(x) = - \frac{1}{\beta - \alpha} \int_{F^{-1}(\alpha)}^{F^{-1}(\beta)} [\delta_x(y) - F(y)] dy.
$$

Assume that $F$ is continuous at $F^{-1}(\alpha)$ and $F^{-1}(\beta)$. $F(F^{-1}(\alpha)) = \alpha$ and $F(F^{-1}(\beta)) = \beta$.

When $x < F^{-1}(\alpha)$,

$$
\phi_F(x) = - \frac{1}{\beta - \alpha} \int_{F^{-1}(\alpha)}^{F^{-1}(\beta)} [1 - F(y)] dy
$$

$$
= - \left. \frac{y[1 - F(y)]}{\beta - \alpha} \right|_{F^{-1}(\alpha)}^{F^{-1}(\beta)} - \frac{1}{\beta - \alpha} \int_{F^{-1}(\alpha)}^{F^{-1}(\beta)} y dF(y)
$$

$$
= \frac{F^{-1}(\alpha)(1 - \alpha) - F^{-1}(\beta)(1 - \beta)}{\beta - \alpha} - T(F)
$$
Step 1: Derivation of the influence function $\phi_F$

Similarly, when $x > F^{-1}(\beta)$,

$$
\phi_F(x) = \frac{1}{\beta - \alpha} \int_{F^{-1}(\alpha)}^{F^{-1}(\beta)} F(y) dy
= \frac{F^{-1}(\beta)\beta - F^{-1}(\alpha)\alpha}{\beta - \alpha} - T(F).
$$

Finally, when $F^{-1}(\alpha) \leq x \leq F^{-1}(\beta)$,

$$
\phi_F(x) = \frac{1}{\beta - \alpha} \int_{F^{-1}(\alpha)}^{x} F(y) dy - \frac{1}{\beta - \alpha} \int_{x}^{F^{-1}(\beta)} [1 - F(y)] dy
= \frac{x - F^{-1}(\alpha)\alpha - F^{-1}(\beta)(1 - \beta)}{\beta - \alpha} - T(F).
$$
Step 1: Derivation of the influence function $\phi_F$

Hence,

$$
\phi_F(x) = \begin{cases} 
    \frac{F^{-1}(\alpha)(1-\alpha)-F^{-1}(\beta)(1-\beta)}{\beta-\alpha} - T(F) & x < F^{-1}(\alpha) \\
    \frac{x-F^{-1}(\alpha)(1-\alpha)-F^{-1}(\beta)(1-\beta)}{\beta-\alpha} - T(F) & F^{-1}(\alpha) \leq x \leq F^{-1}(\beta) \\
    \frac{F^{-1}(\beta)(1-\alpha)-F^{-1}(\alpha)(1-\beta)}{\beta-\alpha} - T(F) & x > F^{-1}(\beta). 
\end{cases}
$$

If $F$ is symmetric about $\theta$, $J$ is symmetric about $\frac{1}{2}$ ($J(t) = J(1 - t)$), and $\int_0^1 J(t)dt = 1$, then $F(x) = F_0(x - \theta)$, where $F_0$ is a c.d.f. that is symmetric about 0, i.e., $F_0(x) = 1 - F_0(-x)$, and

$$
\int xJ(F_0(x))dF_0(x) = \int xJ(1 - F_0(-x))dF_0(x) = \int xJ(F_0(-x))dF_0(x) = -\int yJ(F_0(y))dF_0(y),
$$

i.e., $\int xJ(F_0(x))dF_0(x) = 0$. 
Step 1: Derivation of the influence function $\phi_F$

Hence,

$$T(F) = \int xJ(F(x))dF(x)$$

$$= \theta \int J(F(x))dF(x) + \int (x - \theta)J(F_0(x - \theta))dF_0(x - \theta)$$

$$= \theta \int_0^1 J(t)dt + \int yJ(F_0(y))dF_0(y)$$

$$= \theta.$$

Assume that $F$ is continuous at $F^{-1}(\alpha)$ and $F^{-1}(1 - \alpha)$. When $\beta = 1 - \alpha$, $J$ is symmetric about $\frac{1}{2}$ and

$$\phi_F(x) = \begin{cases} 
\frac{F^{-1}_0(\alpha)}{1 - 2\alpha} & x < F^{-1}(\alpha) \\
\frac{x - \theta}{1 - 2\alpha} & F^{-1}(\alpha) \leq x \leq F^{-1}(1 - \alpha) \\
\frac{F^{-1}_0(1 - \alpha)}{1 - 2\alpha} & x > F^{-1}(1 - \alpha),
\end{cases}$$

where $F^{-1}(\alpha) + F^{-1}(1 - \alpha) = 2\theta$, $F^{-1}_0(\alpha) = F^{-1}(\alpha) - \theta$ and $F^{-1}_0(1 - \alpha) = F^{-1}(1 - \alpha) - \theta$. 
Step 2: Calculation of $\sigma_F^2 = E[\phi_F(X_1)]^2$

Because $F_0^{-1}(\alpha) = -F_0^{-1}(1 - \alpha)$, we obtain that

$$\int [\phi_F(x)]^2 dF(x) = \frac{[F_0^{-1}(\alpha)]^2}{(1 - 2\alpha)^2} \alpha + \frac{[F_0^{-1}(1 - \alpha)]^2}{(1 - 2\alpha)^2} \alpha$$

$$+ \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} \frac{(x - \theta)^2}{(1 - 2\alpha)^2} dF(x)$$

$$= 2\alpha \frac{[F_0^{-1}(1 - \alpha)]^2}{(1 - 2\alpha)^2} + \int_{F_0^{-1}(\alpha)}^{F_0^{-1}(1-\alpha)} \frac{x^2}{(1 - 2\alpha)^2} dF_0(x)$$

$$= \sigma_{\alpha}^2.$$

Step 3: Asymptotic normality of the trimmed sample mean

It can be shown that the L-functional $T(G)$ is differentiable in some sense (see the textbook).

Hence, for the $\alpha$-trimmed sample mean $\bar{X}_\alpha$,

$$\sqrt{n}(\bar{X}_\alpha - \theta) \rightarrow_d N(0, \sigma_{\alpha}^2).$$
Step 2: Calculation of $\sigma^2_F = E[\phi_F(X_1)]^2$

Because $F_0^{-1}(\alpha) = -F_0^{-1}(1 - \alpha)$, we obtain that

$$
\int [\phi_F(x)]^2 dF(x) = \frac{[F_0^{-1}(\alpha)]^2}{(1 - 2\alpha)^2} \alpha + \frac{[F_0^{-1}(1 - \alpha)]^2}{(1 - 2\alpha)^2} \alpha
\]

$$

$$
+ \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} \frac{(x - \theta)^2}{(1 - 2\alpha)^2} dF(x)
\]

$$

$$
= 2\alpha [F_0^{-1}(1 - \alpha)]^2 \frac{1}{(1 - 2\alpha)^2} + \int_{F_0^{-1}(\alpha)}^{F_0^{-1}(1-\alpha)} \frac{x^2}{(1 - 2\alpha)^2} dF_0(x)
\]

$$

$$
= \sigma^2_\alpha.
\]

Step 3: Asymptotic normality of the trimmed sample mean

It can be shown that the L-functional $T(G)$ is differentiable in some sense (see the textbook).

Hence, for the $\alpha$-trimmed sample mean $\bar{X}_\alpha$,

$$
\sqrt{n}(\bar{X}_\alpha - \theta) \to_d N(0, \sigma^2_\alpha).
\]