

L13 Operations with $o_p(1)$ and $O_p(1)$

1. Recall

(1) Properties of sequence bounded in probability

X_n is bounded in probability $\iff g(X_n)$ is bounded in probability for all continuous $g(\cdot)$.

X_n is bounded in probability $\iff X_{n_k}$ is bounded in probability for all X_{n_k} .

$X_n = \begin{pmatrix} X_{1n} \\ \vdots \\ X_{kn} \end{pmatrix}$ is bounded in probability $\iff$ In $\begin{pmatrix} X_{in} \\ \vdots \\ X_{kn} \end{pmatrix}$, X_{in} is bounded in probability for all $i = 1, \dots, k$.

(2) Notations

$X_n = O_p(1) \iff X_n$ is bounded in probability

$\iff \forall \epsilon > 0, \exists M_\epsilon$ and $N_\epsilon, \forall n \geq N_\epsilon, P(\|X_n\| \leq M_\epsilon) \geq 1 - \epsilon$.

$X_n = O_p(Y_n) \iff 0 \neq Y_n \in R$ and $\frac{X_n}{Y_n} = O_p(1)$.

$X_n = o_p(1) \iff X_n \xrightarrow{p} 0 \iff \lim_n P(\|X_n\| \leq \delta) = 1, \forall \delta > 0$

$\iff \forall \epsilon > 0, \forall \delta > 0, \exists N_{\epsilon\delta} > 0, \forall n \geq N_{\epsilon\delta}, P(\|X_n\| \leq \delta) \geq 1 - \epsilon$.

$X_n = o_p(y_n) \iff 0 \neq Y_n \in R$ and $\frac{X_n}{Y_n} = o_p(1)$.

2. Operations on $o_p(1)$ and $O_p(1)$

(1) On $O_p(1)$: $O_p(1) + O_p(1) = O_p(1)$, $O_p(1)O_p(1) = O_p(1)$, $AO_p(1) = O_p(1)$.

Proof. Suppose $X_n = O_p(1)$ and $Y_n = O_p(1)$. We show $X_n + Y_n = O_p(1)$, $X_n Y_n = O_p(1)$ and $AX_n = O_p(1)$ assuming that the dimension requirements are met.

X_n and Y_n are bounded in probability. so that $\begin{pmatrix} X_n \\ Y_n \end{pmatrix}$ is bounded in probability.

With continuous $f\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = x + y$, $X_n + Y_n = f\left(\begin{pmatrix} X_n \\ Y_n \end{pmatrix}\right)$ is bounded in probability, i.e., $X_n + Y_n = O_p(1)$;

With continuous $g\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = xy$, $X_n Y_n = g\left(\begin{pmatrix} X_n \\ Y_n \end{pmatrix}\right)$ is bounded in probability, i.e., $X_n Y_n = O_p(1)$;

With continuous $h(X) = AX$, $AX_n = h(X_n)$ is bounded in probability. $\square$

(2) On $o_p(1)$: $o_p(1) + o_p(1) = o_p(1)$, $o_p(1)o_p(1) = o_p(1)$, $AO_p(1) = o_p(1)$.

Sketch of Proof. Suppose $X_n = o_p(1)$ and $Y_n = o_p(1)$, i.e., $X_n \xrightarrow{p} 0$ and $Y_n \xrightarrow{p} 0$.

We can show that $X_n + Y_n \xrightarrow{p} 0$, $X_n Y_n \xrightarrow{p} 0$ and $AX_n \xrightarrow{p} 0$ assuming that the dimension requirements are met. Thus $X_n + Y_n = o_p(1)$, $X_n Y_n = o_p(1)$ and $AX_n = o_p(1)$.

(3) On both $O_p(1)$ and $o_p(1)$: $O_p(1) + o_p(1) = O_p(1)$ and $O_p(1) \cdot o_p(1) = o_p(1)$.

Proof: Treat $o_p(1)$ as $O_p(1)$. Then $O_p(1) + o_p(1) = O_p(1) + O_p(1) = O_p(1)$.

For $O_p(1) \cdot o_p(1) = o_p(1)$ assume $X_n \in R$ and $Y_n \in R^k$ so that $\|X_n Y_n\| = |X_n| \cdot \|Y_n\|$. In two cases (i) $X_n = O_p(1)$ and $Y_n = o_p(1)$. (ii) $X_n = o_p(1)$ and $Y_n = O_p(1)$, we show $X_n Y_n = o_p(1)$.

(i) $X_n \in R$, $X_n = O_p(1)$ and $Y_n = o_p(1)$:

$X_n = O_p(1)$: $\forall \epsilon > 0 \exists M_\epsilon > 0$ and $N_\epsilon > 0, \forall n \geq N_\epsilon P(|X_n| \leq M_\epsilon) \geq 1 - \frac{\epsilon}{2}$.

$$Y_n = o_p(1) \implies \forall \delta > 0 \exists N_{\epsilon\delta 1} > 0 \forall n \geq N_{\epsilon\delta 1} P(\|Y_n\| \leq \delta/M_\epsilon) \geq 1 - \epsilon/2.$$

Let $N_{\epsilon, \delta} = \max[N_\epsilon, N_{\epsilon\delta 1}]$. For all $n \geq N_{\epsilon, \delta}$,

With $A = [|X_n| \leq M_\epsilon]$ and $B = [\|Y_n\| \leq \delta/M_\epsilon]$, $A \cap B \subset [\|X_n Y_n\| \leq \delta]$. So

$$\begin{aligned} P(\|X_n Y_n + n\| \leq \delta) &\geq P(A \cap B) = P(A) + P(B) - P(A \cup B) \\ &\geq 1 - \epsilon + [1 - P(A \cup B)] \geq 1 - \epsilon. \end{aligned}$$

Hence $X_n Y_n = o_p(1)$.

(ii) $X_n \in R$, $X = o_p(1)$ and $Y = O_p(1)$

Skipped.

Ex1: $X \xrightarrow{p} X$ and $Y_n \xrightarrow{p} Y$. Show that $aX_n + bY_n + c \xrightarrow{p} aX + bY + c$.
 $(aX_n + bY_n + c) - (aX + bY + c) = a(X_n - X) + b(Y_n - Y)$
 $= O(1)o_p(1) + O(1)o_p(1) = o_p(1).$

So $aX_n + bY_n + c \xrightarrow{p} aX + bY + c$.

Ex2: If $\hat{\theta}_n - \theta = O_p\left(\frac{1}{\sqrt{n}}\right)$, show that $\hat{\theta}_n$ is a consistent estimator for θ .

$\hat{\theta}_n - \theta = O_p\left(\frac{1}{\sqrt{n}}\right)$ implies $\sqrt{n}(\hat{\theta}_n - \theta) = O_p(1)$. So

$$\hat{\theta}_n - \theta = \frac{1}{\sqrt{n}} \left[\sqrt{n}(\hat{\theta}_n - \theta) \right] = o(1) \cdot O_p(1) = o_p(1).$$

Thus $\hat{\theta}_n \xrightarrow{p} \theta$. Hence $\hat{\theta}_n$ is a consistent estimator for θ .

3. Consistent estimators for lower/Upper limits of support

(1) $X_1, \dots, X_n$ is a random sample from X with support (a, θ) .

Let $\hat{\theta}_n = X_{(n)} = \max(X_1, \dots, X_n)$. Then $\hat{\theta}_n$ is a consistent estimator for θ .

Pf: $\forall \epsilon > 0$,

$$\begin{aligned} P(|\hat{\theta}_n - \theta| < \epsilon) &= P(\theta - \epsilon < \hat{\theta}_n < \theta + \epsilon) = P(\hat{\theta}_n > \theta - \epsilon) \\ &= 1 - P(\hat{\theta}_n \leq \theta - \epsilon) = 1 - P(X_1 \leq \theta - \epsilon, \dots, X_n \leq \theta - \epsilon) \\ &= 1 - [P(X \leq \theta - \epsilon)]^n. \end{aligned}$$

Here $0 \leq P(X \leq \theta - \epsilon) < 1$. Thus $\lim_n P(|\hat{\theta}_n - \theta| \leq \epsilon) = 1$ for all ϵ .

Therefore $\hat{\theta}_n$ is a consistent estimator for θ . $\square$

(2) $X_1, \dots, X_n$ is a random sample from X with support (θ, b) .

Let $\hat{\theta}_n = X_{(n)} = \min(X_1, \dots, X_n)$. Then $\hat{\theta}_n$ is a consistent estimator for θ .

Proof. Skipped.

L14 Consistent estimator for population median

1. Population median and sample median

(1) Population median

Suppose X is a continuous random variable.

$$\begin{aligned} \theta \text{ is the median for } X &\stackrel{\text{def}}{\iff} P(X < \theta) = P(X > \theta) \\ &\iff P(X < \theta) = 0.5 \iff P(X > \theta) = 0.5 \end{aligned}$$

(2) Sample median

Let $X_1, \dots, X_n$ be a random sample from X .

Rearrange the sample to $X_{(1)} \leq \dots \leq X_{(n)}$. Then the sample median is

$$\hat{\theta}_n = \begin{cases} \frac{X_{(\frac{n}{2})} + X_{(\frac{n}{2}+1)}}{2} & n \text{ is even} \\ X_{(\frac{n+1}{2})} & n \text{ is odd} \end{cases}$$

Ex1: With $X_{(1)} \leq X_{(2)} \leq X_{(3)}$, $X_{(2)}$ is the sample median.

With $X_{(1)} \leq X_{(2)} \leq X_{(3)} \leq X_{(4)}$, $\frac{X_{(2)} + X_{(3)}}{2}$ is the sample median.

2. Analysis on how to prove the consistency

(1) Sufficient condition I

$\hat{\theta}_n$ is consistent estimator for $\theta \iff \sqrt{n}(\hat{\theta}_n - \theta) = O_p(1)$ by Ex2 in L13

$$\iff \forall \epsilon > 0, \exists M_\epsilon > 0 \text{ and } N_\epsilon > 0, \forall n \geq N_\epsilon, P(|\sqrt{n}(\hat{\theta}_n - \theta)| \leq M_\epsilon) > 1 - \epsilon$$

$$\iff \forall \epsilon > 0, \exists M_\epsilon > 0 \text{ and } N_\epsilon > 0, \forall n \geq N_\epsilon, P\left(\theta - \frac{M_\epsilon}{\sqrt{n}} \leq \hat{\theta}_n \leq \theta + \frac{M_\epsilon}{\sqrt{n}}\right) > 1 - \epsilon$$

$$\iff \forall \epsilon > 0, \exists M_\epsilon > 0 \text{ and } N_\epsilon > 0, \forall n \geq N_\epsilon, P\left(\hat{\theta}_n \leq \theta - \frac{M_\epsilon}{\sqrt{n}} \cup \hat{\theta}_n \geq \theta + \frac{M_\epsilon}{\sqrt{n}}\right) \leq \epsilon.$$

(2) Sufficient condition II

$$\forall \epsilon > 0, \exists M_{1\epsilon} > 0 \text{ and } N_{1\epsilon} > 0, \forall n \geq N_{1\epsilon} P\left(\hat{\theta}_n \leq \theta - \frac{M_{1\epsilon}}{\sqrt{n}}\right) \leq \frac{\epsilon}{2}$$

$$\exists M_{2\epsilon} > 0 \text{ and } N_{2\epsilon} > 0, \forall n \geq N_{2\epsilon} P\left(\hat{\theta}_n \geq \theta + \frac{M_{2\epsilon}}{\sqrt{n}}\right) \leq \frac{\epsilon}{2}$$

$$\implies \forall \epsilon > 0, \exists M_\epsilon = \max[M_{1\epsilon}, M_{2\epsilon}] > 0 \text{ and } N_\epsilon = \max[N_{1\epsilon}, N_{2\epsilon}] > 0, \forall n \geq N_\epsilon$$

$$P\left(\hat{\theta}_n \leq \theta - \frac{M_\epsilon}{\sqrt{n}}\right) < \frac{\epsilon}{2} \text{ and } P\left(\hat{\theta}_n \geq \theta + \frac{M_\epsilon}{\sqrt{n}}\right) \leq \frac{\epsilon}{2}$$

$$\implies \forall \epsilon > 0, \exists M_\epsilon > 0 \text{ and } N_\epsilon > 0, \forall n \geq N_\epsilon, P\left(\hat{\theta}_n \leq \theta - \frac{M_\epsilon}{\sqrt{n}} \cup \hat{\theta}_n \geq \theta + \frac{M_\epsilon}{\sqrt{n}}\right) \leq \epsilon.$$

$$\iff \hat{\theta}_n \text{ is consistent estimator for } \theta$$

(3) What we show

In the two statements in sufficient condition II we only show one:

$$\forall \epsilon > 0, \exists M_\epsilon > 0 \text{ and } N_\epsilon > 0, \forall n \geq N_\epsilon, P\left(\hat{\theta}_n \geq \theta + \frac{M_\epsilon}{\sqrt{n}}\right) \leq \frac{\epsilon}{2}.$$

3. Proof

(1) Further analysis

Let B be the number of $X_1, \dots, X_n$ that fall in $\left[\theta + \frac{M_\epsilon}{\sqrt{n}}, \infty\right)$.

Then $B \sim B(n, p) \sim (np, np(1-p))$ where $p = P\left(X > \theta + \frac{M_\epsilon}{\sqrt{n}}\right)$. Note that

$$\begin{aligned}
\left[\hat{\theta}_n \geq \theta + \frac{M_\epsilon}{\sqrt{n}} \right] &= \left[\begin{cases} X_{(\frac{n+1}{2})} & n \text{ is odd} \\ \frac{X_{(\frac{n}{2})} + X_{(\frac{n}{2}+1)}}{2} & n \text{ is even} \end{cases} \geq \theta + \frac{M_\epsilon}{\sqrt{n}} \right] \\
&\subset \left[\begin{cases} X_{(\frac{n+1}{2})} & n \text{ is odd} \\ X_{(\frac{n}{2}+1)} & n \text{ is even} \end{cases} \geq \theta + \frac{M_\epsilon}{\sqrt{n}} \right] \\
&\subset \left[B \geq \begin{cases} n - \frac{n+1}{2} + 1 & n \text{ is odd} \\ n - \frac{n}{2} & n \text{ is even} \end{cases} \right] = [B \geq \frac{n}{2}] \\
&= [B - np \geq \frac{n}{2} - np]
\end{aligned}$$

By Chebyshev's inequality and $p(1-p) \leq \frac{1}{4}$,

$$P\left(\hat{\theta}_n \geq \theta + \frac{M_\epsilon}{\sqrt{n}}\right) \leq P\left(B - np \geq \frac{n}{2} - np\right) \leq \frac{np(1-p)}{n^2(\frac{1}{2}-p)^2} \leq \frac{1}{4n(\frac{1}{2}-p)^2}$$

where $\frac{1}{2} - p = P(X > \theta) - P(X > \theta + \frac{M_\epsilon}{\sqrt{n}}) = P\left(\theta \leq X < \theta + \frac{M_\epsilon}{\sqrt{n}}\right)$. Thus it suffices to show that $\forall \epsilon > 0$, $\exists M_\epsilon > 0$ and $N_\epsilon > 0$, $\forall n \geq N_\epsilon$, $\frac{1}{4n[P(\theta \leq X < \theta + M_\epsilon/\sqrt{n})]^2} \leq \frac{\epsilon}{2}$.

(2) Proof

For all $\epsilon > 0$,

suppose $[\theta - d, \theta + d]$ is in the support of X and contains all $X_1, \dots, X_n$.

With $f_m = \min[f(x) : \theta - d \leq x \leq \theta + d]$ where $f(x)$ is the pdf of X ,

let $M_\epsilon = \frac{1}{f_m \sqrt{2\epsilon}}$.

$$\text{Then } \left[P\left(\theta \leq X < \theta + \frac{M_\epsilon}{\sqrt{n}}\right) \right]^2 \geq \left(f_m \frac{M_\epsilon}{\sqrt{n}} \right)^2 = \frac{1}{2n\epsilon}.$$

Thus $\frac{1}{4n[P(\theta \leq X < \theta + M_\epsilon/\sqrt{n})]^2} \leq \frac{1}{4n \frac{1}{2n\epsilon}} = \frac{\epsilon}{2}$. $\square$