ECON6190 Section 10

Nov. 1, 2024

Yiwei Sun

1. Let $\{X_1 \dots X_n\}$ be a sequence of i.i.d random variables with mean μ and variance σ^2. Let $\hat{\mu} = \frac{1}{n}\sum_{i=1}^n X_i$, and $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \hat{\mu})^2$.

   (a) Suppose $\mathbb{E}X_i^2 < \infty$ → Chebyshev inequality, $i = 1, \dots n$. Show $\hat{\sigma}^2 \xrightarrow{p} \sigma^2$ as $n \to \infty$. and $\hat{\mu} \xrightarrow{P} \mu$.

(a) Need to find $E[\hat{\mu}]$ and $var(\hat{\mu})$

$$E[\hat{\mu}] = E\left[\frac{1}{n}\sum_{i=1}^n X_i\right] \overset{\text{linearity } E[\cdot]}{=} \frac{1}{n}\sum_{i=1}^n \underbrace{E[X_i]}_{iid,=\mu} = \mu$$

$$var(\hat{\mu}) = var\left(\frac{1}{n}\sum_{i=1}^n X_i\right) \overset{iid}{=} \frac{1}{n^2}\sum_{i=1}^n \underbrace{var(X_i)}_{\sigma^2} = \frac{1}{n^2}(n\sigma^2) = \frac{\sigma^2}{n}$$

By Chebyshev inequality / Markov (r=2):

$$P(|\hat{\mu} - \underbrace{E[\hat{\mu}]}_{\mu}| > \delta) \leq \frac{E[(\hat{\mu}-\mu)^2]}{\delta^2} \overset{MSE(\hat{\mu})}{=} \frac{var(\hat{\mu})}{\delta^2} = \frac{\sigma^2}{n\delta^2} \to 0 \text{ as } n \to \infty.$$

$\Rightarrow \hat{\mu} \xrightarrow{P} \mu.$

WTS: $\hat{\sigma}^2 \xrightarrow{P} \sigma^2$

Notice that:
$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n (X_i^2 - 2X_i\hat{\mu} + \hat{\mu}^2)$$
$$= \frac{1}{n}\sum_{i=1}^n X_i^2 \underbrace{- 2\hat{\mu}\underbrace{\frac{1}{n}\sum_{i=1}^n X_i}_{\hat{\mu}}}_{-2\hat{\mu}^2} + \underbrace{\frac{1}{n}\sum_{i=1}^n \hat{\mu}^2}_{\hat{\mu}^2}$$

not indexed

$$= \frac{1}{n}\sum_{i=1}^n X_i^2 - \hat{\mu}^2$$

→ sample version of $var(X) = E[X^2] - (E[X])^2$

Since $E[X_i^2] < \infty$, by Khinchine's WLLN, $\frac{1}{n}\sum_{i=1}^n X_i^2 \xrightarrow{P} E[X^2]$.

We also know $\hat{\mu} \xrightarrow{P} \mu$, then by CMT,

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n X_i^2 - \hat{\mu}^2 \xrightarrow{P} E[X^2] - (E[X])^2 = var(X) = \sigma^2$$

$\Rightarrow \hat{\sigma}^2 \xrightarrow{P} \sigma^2$

(b) Imposing additional assumptions if necessary, find the asymptotic distribution of

$$\sqrt{n}(\hat{\sigma}^2 - \sigma^2)$$

by using delta method. Carefully state your results.

Recall def: $\sigma^2 = E[X^2] - (E[X])^2$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n X_i^2 - (\underbrace{\frac{1}{n}\sum_{i=1}^n X_i}_{\hat{\mu}})^2$$

Consider a function h s.t. $h(a,b) = a - b^2$.

$$\Rightarrow \quad \sigma^2 = h(E[X^2], E[X]) \quad , \quad \hat{\sigma}^2 = h\left(\frac{1}{n}\sum_{i=1}^n X_i^2, \frac{1}{n}\sum_{i=1}^n X_i\right)$$

By first order Taylor expansion,

$$h\left(\frac{1}{n}\sum_{i=1}^n X_i^2, \frac{1}{n}\sum_{i=1}^n X_i\right) = h(E[X^2], E[X]) + \begin{pmatrix} \frac{\partial}{\partial a}h(a,b)\Big|_{\binom{a}{b}=\binom{\tilde{a}}{\tilde{b}}} \\ \frac{\partial}{\partial b}h(a,b)\Big|_{\binom{a}{b}=\binom{\tilde{a}}{\tilde{b}}} \end{pmatrix}^T \begin{pmatrix} \frac{1}{n}\sum_{i=1}^n (X_i^2 - E[X^2]) \\ \frac{1}{n}\sum_{i=1}^n (X_i - E[X]) \end{pmatrix}$$

where $\begin{pmatrix} \tilde{a} \\ \tilde{b} \end{pmatrix}$ is in between $\begin{pmatrix} \frac{1}{n}\sum_{i=1}^n X_i^2 \\ \frac{1}{n}\sum_{i=1}^n X_i \end{pmatrix}$ and $\begin{pmatrix} E[X^2] \\ E[X] \end{pmatrix}$.

Now assume $E\left[\left\|\begin{pmatrix} X^2 \\ X \end{pmatrix}\right\|^2\right] < \infty$, which requires $E[X^4] < \infty$, by mulivariate CLT

$$\sqrt{n}\begin{pmatrix} \frac{1}{n}\sum_{i=1}^n (X_i^2 - E[X^2]) \\ \frac{1}{n}\sum_{i=1}^n (X_i - E[X]) \end{pmatrix} \xrightarrow{d} N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \operatorname{var}\begin{pmatrix} X^2 \\ X \end{pmatrix}\right).$$

Since we assume $E[X^4] < \infty$, $\begin{pmatrix} \frac{1}{n}\sum_{i=1}^n X_i^2 \\ \frac{1}{n}\sum_{i=1}^n X_i \end{pmatrix} \xrightarrow{P} \begin{pmatrix} E[X^2] \\ E[X] \end{pmatrix}$, by WLLN.

and $\begin{pmatrix} \tilde{a} \\ \tilde{b} \end{pmatrix}$ is between $\begin{pmatrix} \frac{1}{n}\sum_{i=1}^n X_i^2 \\ \frac{1}{n}\sum_{i=1}^n X_i \end{pmatrix}$ and $\begin{pmatrix} E[X^2] \\ E[X] \end{pmatrix}$, $\begin{pmatrix} \tilde{a} \\ \tilde{b} \end{pmatrix} \xrightarrow{P} \begin{pmatrix} E[X^2] \\ E[X] \end{pmatrix}$.

Since $\begin{pmatrix} \frac{\partial}{\partial a}h(a,b) \\ \frac{\partial}{\partial b}h(a,b) \end{pmatrix} = \begin{pmatrix} 1 \\ -2b \end{pmatrix}$ is continuous, by CMT,

$$\begin{pmatrix} \frac{\partial}{\partial a} h(a,b) \big|_{\binom{a}{b}=\binom{\tilde a}{\tilde b}} \\ \frac{\partial}{\partial b} h(a,b) \big|_{\binom{a}{b}=\binom{\tilde a}{\tilde b}} \end{pmatrix} \xrightarrow{P} \begin{pmatrix} \frac{\partial}{\partial a} h(a,b) \big|_{\binom{a}{b}=\binom{E[X^2]}{E[X]}} \\ \frac{\partial}{\partial b} h(a,b) \big|_{\binom{a}{b}=\binom{E[X^2]}{E[X]}} \end{pmatrix} = \begin{pmatrix} 1 \\ -2E[X] \end{pmatrix}.$$

$$\sqrt{n}\left(\hat\sigma^2 - \sigma^2\right) = \sqrt{n}\left(h\left(\frac{1}{n}\sum_{i=1}^n X_i^2, \frac{1}{n}\sum_{i=1}^n X_i\right) - h\left(E[X^2] - E[X]\right)\right)$$

$$= \underbrace{\begin{pmatrix} \frac{\partial}{\partial a} h(a,b) \big|_{\binom{a}{b}=\binom{\tilde a}{\tilde b}} \\ \frac{\partial}{\partial b} h(a,b) \big|_{\binom{a}{b}=\binom{\tilde a}{\tilde b}} \end{pmatrix}}_{\xrightarrow{P} \begin{pmatrix} 1 \\ -2E[X] \end{pmatrix}} \underbrace{\sqrt{n} \begin{pmatrix} \frac{1}{n}\sum_{i=1}^n (X_i^2 - E[X^2]) \\ \frac{1}{n}\sum_{i=1}^n (X_i - E[X]) \end{pmatrix}}_{\xrightarrow{d} \mathcal{N}\left(\binom{0}{0}, \operatorname{var}\binom{X^2}{X}\right)}$$

$$\xrightarrow{d} \mathcal{N}\left(\underbrace{\begin{pmatrix} 1 \\ -2E[X] \end{pmatrix}^T}_{1\times 2} \underbrace{\begin{pmatrix} 0 \\ 0 \end{pmatrix}}_{2\times 1}, \underbrace{\begin{pmatrix} 1 \\ -2E[X] \end{pmatrix}^T}_{1\times 2} \underbrace{\operatorname{var}\begin{pmatrix} X^2 \\ X \end{pmatrix}}_{2\times 2} \underbrace{\begin{pmatrix} 1 \\ -2E[X] \end{pmatrix}}_{2\times 1} \right)$$

where
$$\operatorname{var}\begin{pmatrix} X^2 \\ X \end{pmatrix} = \begin{pmatrix} \operatorname{var}(X^2) & \operatorname{cov}(X^2, X) \\ \operatorname{cov}(X^2, X) & \operatorname{var}(X) \end{pmatrix}.$$

Can simplify (not required) to $\sqrt{n}(\hat\sigma^2 - \sigma^2) \xrightarrow{d} \mathcal{N}\left(0, \operatorname{var}\left((X - E[X])^2\right)\right)$

Let's find $\sqrt{n}(\hat\sigma^2 - \sigma^2)$ w/o using delta method.

$$\hat\sigma^2 - \sigma^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \hat\mu)^2 - \sigma^2$$

$$= \frac{1}{n}\sum_{i=1}^n (X_i - \mu + \mu - \hat\mu)^2 - \sigma^2$$

$$= \frac{1}{n}\sum_{i=1}^n (X_i - \mu)^2 + \underbrace{2(\mu - \hat\mu) \underbrace{\frac{1}{n}\sum_{i=1}^n (X_i - \mu)}_{= \hat\mu - \mu}}_{-2(\mu - \hat\mu)^2} + (\mu - \hat\mu)^2 - \sigma^2$$

$$= \frac{1}{n}\sum_{i=1}^n (X_i - \mu)^2 - \sigma^2 - (\mu - \hat\mu)^2 \quad \cdots\cdots (1)$$

$$\sqrt{n}\left(\hat\sigma^2-\sigma^2\right) \overset{①}{=} \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n (x_i-\mu)^2-\sigma^2\right) - \underbrace{\sqrt{n}\left(\mu-\hat\mu\right)^2}_{=\underbrace{\sqrt{n}(\hat\mu-\mu)}_{O_p(1)}\underbrace{(\hat\mu-\mu)}_{o_p(1)} = o_p(1)}$$

$O_p(1)$: b/c by CLT, $\xrightarrow{d} \mathcal{N}(0,\sigma^2)$

$o_p(1)$ b/c $\hat\mu \xrightarrow{P} \mu$

"sample average of $\tilde{x}_i$" $\Rightarrow$ apply CLT

$$= \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n \underbrace{(x_i-\mu)^2}_{} - \sigma^2\right) + o_p(1)$$

view this as my new RV $\tilde{x}_i$

$$E[\tilde{x}_i] = E[(x_i-\mu)^2] = \sigma^2$$

$$\Rightarrow \frac{1}{n}\sum_{i=1}^n \tilde{x}_i - E[\tilde{x}_i]$$

$\Rightarrow$ by CLT, $\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n (x_i-\mu)^2-\sigma^2\right) \xrightarrow{d} \mathcal{N}\left(0,\ \mathrm{var}\left((x-\mu)^2\right)\right)$

Q: What's the stochastic order of $\hat\mu-\mu$ and $\hat\sigma^2-\sigma^2$?

$\hat\mu \xrightarrow{P} \mu \quad \Rightarrow \quad \hat\mu-\mu = o_p(1)$

What's the rate of $\hat\mu-\mu$ converges to 0?

$$\hat\mu-\mu = O_p\left(\sqrt{MSE(\hat\mu)}\right) = O_p\left(\sqrt{\frac{\sigma^2}{n}}\right) = O_p\left(\frac{1}{\sqrt{n}}\right).$$

Alternatively, by CLT, $\sqrt{n}(\hat\mu-\mu) \xrightarrow{d} \mathcal{N}(0,\sigma^2) \Rightarrow \sqrt{n}(\hat\mu-\mu) = O_p(1)$

$\Rightarrow \hat\mu-\mu = O_p\left(\frac{1}{\sqrt{n}}\right)$

$\left|\frac{\sqrt{n}(\hat\sigma^2-\sigma^2)}{1}\right|$ bounded in prob asymptotically

Since $\sqrt{n}\left(\hat\sigma^2-\sigma^2\right) \xrightarrow{d} \mathcal{N}\left(0,\ \mathrm{var}\left((x-E[x])^2\right)\right) \Rightarrow \sqrt{n}(\hat\sigma^2-\sigma^2) = O_p(1)$

$\Rightarrow \hat\sigma^2-\sigma^2 = O_p\left(\frac{1}{\sqrt{n}}\right)$