ECON220B Discussion Section 1
Basics of Asymptotic Distribution

Lapo Bini

Intro

- I am a second year interested in Macro and Time Series Econometrics.
- We will meet every Thursday 5/6 pm. The room is booked until 7pm (office hour).
- If you have doubts or need help, I am always available at lbini@ucsd.edu
- ECON220B is challenging: 6 problem sets, midterm + final exam (a lot of material).
- My goal: cover crucial topics for final exam/qual plus tools that might be helpful for your future research.

Roadmap

1. Convergence theorems.
2. Tools to study asymptotic distribution of estimators.
3. Exercise Asymptotic Linear Representation.
4. Superefficient Estimator.

Convergence Theorems

CAN WE RELAX THOSE?

- IND, NOT ID. DIST.
- NOT IND. NOT ID. DIST.
- WEAKER: just uncorrelated

1. WLLN: If $\{x_1, \ldots, x_n\}$ iid where $E[x_i] = \mu < \infty$, then $\forall \varepsilon > 0$ $\lim_{n\to\infty} \mathcal{P}(|\bar{x}_n - \mu| \geq \varepsilon) = 0$, i.e. $\bar{x}_n \xrightarrow{p} \mu$.

2. Small Oh-Pee: Let $\{x_n\}_{n=1}^{\infty}$ be a sequence of RVs, If $x_n \xrightarrow{p} 0$ then we say $x_n = o_p(1)$.

3. CLT: If $\{x_1, \ldots, x_n\}$ iid, $x_i \sim (\mu, \sigma^2)$, where $\mu < \infty, \sigma^2 < \infty$, then $\sqrt{n}(\bar{x} - \mu) \xrightarrow{d} \mathcal{N}(0, \sigma^2)$.

   MUST ALWAYS HOLD

4. Big Oh-Pee: Let $\{x_n\}_{n=1}^{\infty}$ be a sequence of RVs, If $x_n \xrightarrow{d} \mathcal{L}$ then we say $x_n = O_p(1)$.

$$\mathbb{P}(|\bar{x}-\mu| \le \varepsilon) = \mathbb{P}((\bar{x}-\mu)^2 \le \varepsilon^2) \le \frac{\mathbb{E}[|\bar{x}-\mu|^2]}{\varepsilon^2} =$$

$$\mathbb{P}(|\bar{x}-\mu| \le \varepsilon) \le \frac{Var(\bar{x})}{\varepsilon^2} = \frac{1}{\varepsilon^2} Var\left(\frac{1}{m}\Sigma x_i\right) = \frac{1}{\varepsilon^2 m^2} Var(\Sigma x_i)$$ DIFFERENT CASES

1. IF IID $\frac{1}{\varepsilon^2 m^2} Var(\Sigma x_i) = \frac{m}{\varepsilon^2 m^2}\sigma^2 = \frac{\sigma^2}{\varepsilon^2}\cdot\frac{1}{m} \to 0 \quad \therefore \bar{x} \xrightarrow{p} \mu$

2. UNCORRELATED $\frac{1}{\varepsilon^2 m^2} Var(\Sigma x_i) = \frac{1}{\varepsilon^2 m^2}\Sigma Var(x_i) =$

   - IDENTICALLY DISTRIBUTED $\frac{1}{\varepsilon^2 m^2}\Sigma Var(x_i) = \frac{1}{\varepsilon^2 m^2} m\sigma^2 = \frac{\sigma^2}{\varepsilon}\cdot\frac{1}{m} \to 0$
   - NOT IDENTICALLY DISTRIBUTED $\frac{1}{\varepsilon^2 m^2}\Sigma Var(x_i) = \frac{1}{\varepsilon^2 m^2}\sum^m \sigma_i^2$

IF $\frac{\sum^m \sigma_i^2}{\varepsilon^2} = o_p(m^2)$ THEN $\frac{\Sigma\sigma_i^2}{\varepsilon^2 m^2} \xrightarrow{p} 0$

- $O_p(\cdot)$: IF $X_m \xrightarrow{d} \mathcal{L}$ $\quad Z_m := m^3 X_m \quad Z_m = O_p(?)$

$$Z_m = O_p(m^3) \to \frac{Z_m}{m^3} = X \xrightarrow{d} \mathcal{L}$$

Tools Asymptotic Distribution (1/3)

- Slutsky's theorem: let X_n,Y_n be a sequence of scalar/vector/matrix random elements, If $X_n \xrightarrow{p} x$ and $Y_n \xrightarrow{d} Y$ then:

(i) $X_n Y_n \xrightarrow{d} xY$

(ii) $X_n + Y_n \xrightarrow{d} x + Y$

(iii) $\frac{Y_n}{X_n} \xrightarrow{d} \frac{Y}{x}$ assuming $x \neq 0$

- Relationship between $o_p(1)$ and $O_p(1)$:

$o_p(1) \cdot O_p(1) = o_p(1)$

$O_p(1) + O_p(1) = O_p(1)$

$o_p(1) + o_p(1) = o_p(1)$

$o_p(1) + O_p(1) = O_p(1)$

- Continuous mapping theorem let X_n be a sequence of scalar/vector/matrix random elements and $g(\cdot)$ be a continuous function, If $X_n \xrightarrow{p} x$ then $g(X_n) \xrightarrow{p} g(x)$

Tools Asymptotic Distribution (2/3)

- Taylor's Mean Value Theorem: Let $g : \mathbb{R}^m \to \mathbb{R}$, if g continuous in $[\theta, \hat{\theta}]$, differentiable in $(\theta, \hat{\theta})$, then:

$$g(\hat{\theta}) = g(\theta) + \underline{\nabla g(\tilde{\theta})'\left(\hat{\theta} - \theta\right)} \quad \text{with:} \quad \tilde{\theta} \in [\theta, \hat{\theta}]$$

REMINDER: $g(x') = \sum_{\ell=0}^{m} \frac{g^{(\ell)}(x)}{\ell!}(x'-x)^{\ell} + \frac{g^{(m+1)}(\tilde{x})}{(m+1)!}(x'-x)^{m+1}$

$m = 0 \to 1$

- Delta Method: just trivial manipulation:

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} \mathcal{N}(0, \Sigma)$$

$$\sqrt{n}(g(\hat{\theta}) - g(\theta)) = \underline{\nabla g(\tilde{\theta})'\sqrt{n}(\hat{\theta} - \theta)}$$

$$\sqrt{n}(g(\hat{\theta}) - g(\theta)) \xrightarrow{d} \mathcal{N}(0, \nabla g(\theta)' \Sigma \nabla g(\theta))$$

$$\sqrt{n}\begin{bmatrix} \bar{x} - E[x_i] \\ \bar{y} - E[y_i] \end{bmatrix} \xrightarrow{d} N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{bmatrix}\right)$$

REMEMBER: SCALAR CASE

$x \sim N(\mu, \sigma^2)$

$ax \sim N(a\mu, a^2\sigma^2)$

$ax + c \sim N(a\mu + c, a^2\sigma^2)$

A, Σ (K×K) MATRIX

Y, Z, ν (K×1) VECTOR

$AY \sim N(A\nu, A\Sigma A^T)$

$\underset{(1\times K)(K\times K)}{Z^T Y} \sim N(\underset{(1\times K)(K\times 1)}{Z^T \nu}, \underset{(1\times K)(K\times K)(K\times 1)}{Z^T \Sigma Z})$

Tools Asymptotic Distribution (3/3)

THIS IS UNIVARIATE NORMAL

- Asymptotic linear representation: fantastic tool to study asymptotic distribution of estimators. Based on the so-called "Levy-Lindeberg" CLT:

$$\sqrt{n}(\bar{x}-\mu) = \sqrt{n}\left[\frac{1}{n}\sum_{i=1}^{n} x_i - \frac{n}{n}\mu\right] = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}(x_i-\mu) \xrightarrow{d} \mathcal{N}(0,\sigma^2)$$

- Why is it useful? Suppose $\hat{\theta} = \left(\sum_{i=1}^{n} x_i\right)/\left(\sum_{i=1}^{n} y_i\right)$, can we apply CLT?

NO

- We want to derive the influence function $\psi(x_i, y_i)$ such that:

$$\hat{\theta} = \frac{\frac{1}{m}\sum x_i}{\frac{1}{m}\sum y_i} = \frac{1}{m}\sum x_i/\bar{y}$$

$$\sqrt{n}(\hat{\theta}-\theta) = \frac{1}{\sqrt{n}}\underbrace{\sum_{i=1}^{n}\psi(x_i,y_i)}_{\{\psi_i,\dots,\psi_m\}\ \text{IID} \implies \text{CLT applies.}} + o_p(1)$$

$(x_i, \bar{y}), \dots, (x_m, \bar{y}_m)$

Exercise on Asymptotic Linear approximation

- Let $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ be independently and identically distributed, where both x_i and y_i are univariate but they may not be independent. We define the following parameter of interest and the estimators:

$$\theta = \frac{E[x_i]}{E[y_i]} = \frac{\mu}{\nu}, \quad \hat{\theta} = \frac{\bar{x}}{\bar{y}}, \quad \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

(1) Derive the asymptotic linear representation of $\hat{\theta}$.

(2) Derive the asymptotic distribution of $\hat{\theta}$.

(3) Propose an estimator for the asymptotic variance.

(4) Use the delta method to derive the asymptotic distribution, and compare with your previous answers.

(1) $\theta = \frac{\mu}{\nu}$ $\quad \hat\theta = \frac{\bar x}{\bar y}$

$$\sqrt{n}(\hat\theta - \theta) = \sqrt{n}\left[\frac{\bar x}{\bar y} - \frac{\mu}{\nu}\right] = \sqrt{n}\left[\frac{1}{\bar y \nu}(\bar x \nu - \mu \bar y)\right]$$

$$= \sqrt{n}\left[\frac{1}{\bar y \nu}\left(\underline{\bar x \nu} - \underline{\mu \bar y} + \underline{\mu\nu} - \underline{\mu\nu}\right)\right]$$

$$= \sqrt{n}\left[\frac{1}{\bar y \nu}\left(\underline{\nu(\bar x - \mu)} - \mu\underline{(\bar y - \nu)}\right)\right]$$

$$= \frac{1}{\bar y}\overbrace{\left(\sqrt{n}(\bar x - \mu) - \frac{\mu}{\nu}\sqrt{n}(\bar y - \nu)\right)}^{:= A}$$

$$= \frac{1}{\bar y} A + \frac{1}{\nu} A - \frac{1}{\nu} A = \frac{1}{\nu} A - \overset{O_p(1)}{A}\left(\frac{\bar y - \nu}{\bar y \nu}\right) \xrightarrow{P} 0$$

$\bar y - \nu \xrightarrow{P} 0$

$\bar y \xrightarrow{P} \nu$

$\bar y \nu \longrightarrow \nu^2$ by corollary plim

$\frac{\bar y - \nu}{\bar y \nu} \xrightarrow{P} 0$

$$= \frac{1}{\nu}\left[\underbrace{\sqrt{n}(\bar x - \mu)}_{\xrightarrow{d} N(0,\sigma_x^2)} - \frac{\mu}{\nu}\underbrace{\sqrt{n}(\bar y - \nu)}_{\xrightarrow{d} N(0,\sigma_y^2)}\right] + o_p(1)$$

$$= \sqrt{n}\left(\frac{1}{n}\sum \frac{1}{\nu}(x_i - \mu) - \frac{1}{n}\sum \frac{\mu}{\nu^2}(y_i - \nu)\right) + o_p(1)$$

$$= \frac{1}{\sqrt{n}}\sum\left(\underbrace{\frac{1}{\nu}(x_i - \mu)}_{:= \Psi(x_i)} - \underbrace{\frac{\mu}{\nu^2}(y_i - \nu)}_{:= \emptyset(y_i)}\right) + o_p(1)$$

$O_p(1)\, o_p(1) = o_p(1)$ by slutsky.

(2) $\sqrt{n}(\hat\theta - \theta) = \frac{1}{\sqrt{n}}\sum \underline{\Psi(x_i) - \emptyset(y_i)} + o_p(1) \longrightarrow N(0, V)$

$(\Psi_1, \emptyset_1) \dots (\Psi_n, \emptyset_n)$ are now IID BETWEEN COUPLES BUT WITHIN! THEY COULD BE CORRELATED

$$\left.\begin{aligned} \frac{1}{\sqrt{n}}\sum \Psi(x_i) &\xrightarrow{d} N(0; V_1) \\ \frac{1}{\sqrt{n}}\sum \emptyset(y_i) &\xrightarrow{d} N(0; V_2) \end{aligned}\right\}$$

REMEMBER IF $X \sim N(\mu_x, \sigma_x)$ $Y \sim N(\mu_Y, \sigma_Y^2)$

THEN $X + Y \sim N(\mu_x + \mu_Y;\ \sigma_x^2 + \sigma_Y^2 + 2\underline{\sigma_{xY}})$ COVARIANCE

$$\therefore \frac{1}{\sqrt{n}}\sum \Psi(x_i) - \frac{1}{\sqrt{n}}\sum \emptyset(y_i) = \frac{1}{\sqrt{n}}\sum \Psi(x_i) - \emptyset(y_i) \xrightarrow{d} N(0; V)$$

(3) ASYMPTOTIC VARIANCE

$(x_i, y_i) \ldots (x_m, y_m)$ INDENDENT, ALTHOUG $x_i\ y_i$ MAY BE CORRELATED

$$V = Var\left(\frac{1}{\sqrt{m}}\sum \Psi(x_i) - \phi(y_i)\right) = \frac{1}{m} Var\left(\sum \Psi(x_i) - \phi(y_i)\right) = \frac{1}{m}\sum Var\left(\Psi(x_i) - \phi(y_i)\right)$$

$(x_1, y_1) \ldots (x_m, y_m)$ IDENTICALLY DISTR.

$$= \frac{m}{m} Var\left(\Psi(x_i) - \phi(y_i)\right) = Var\left(\Psi_i(x_i)\right) + Var\left(\phi(y_i)\right) + 2 Cov\left(\Psi(x_i), \phi(y_i)\right)$$

$$= \frac{1}{\nu^2}\sigma_x^2 + \frac{\mu^2}{\nu^4}\sigma_y^2 - 2\frac{\mu}{\nu^3}\sigma_{xy} \longrightarrow$$

PUT HAT ON EVERYTHING FOR ESTIMATOR. COROLLARY PLIM GIVES YOU CONSISTENCY!

(4) ASYMPTOTIC DISTRIBUTION VIA DELTA METHOD

GIVEN $\alpha = \begin{vmatrix} \mu \\ \nu \end{vmatrix}$ $\hat{\alpha} = \begin{vmatrix} \bar{x} \\ \bar{y} \end{vmatrix}$ by CLT $\sqrt{m}\left(\begin{vmatrix} \bar{x} \\ \bar{y} \end{vmatrix} - \begin{vmatrix} \mu \\ \nu \end{vmatrix}\right) \xrightarrow{d} N\left(\begin{vmatrix} 0 \\ 0 \end{vmatrix}, \begin{vmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{vmatrix}\right)$

DELTA METHOD:

$$g(\hat{\alpha}) = g(\alpha) + \nabla g(\tilde{\alpha})(\hat{\alpha} - \alpha) \implies (g(\hat{\alpha}) - g(\alpha)) = \nabla g(\tilde{\alpha})^T(\hat{\alpha} - \alpha)$$

$$\implies \sqrt{m}(g(\hat{\alpha}) - g(\alpha)) = \nabla g(\tilde{\alpha})^T \sqrt{m}(\hat{\alpha} - \alpha) \longrightarrow N(0, \nabla g(\alpha)^T \Sigma \nabla g(\alpha))$$

CLT FOR RANDOM VECTOR

WHY IT BECOMES α?

$\because \tilde{\alpha} \in [\alpha, \hat{\alpha}]$ BUT $\hat{\alpha} \xrightarrow{P} \alpha$

$\therefore \nabla g(\tilde{\alpha})^T \xrightarrow{P} \nabla g(\alpha)$

REMEMBER: $\nabla g(\alpha)^T$ is $(K \times 1)$ VECTOR $\therefore \nabla g(\tilde{\alpha})^T \sqrt{m}(\hat{\alpha} - \alpha)$ is (1×1), IT IS UNIVARIATE NORMAL!

$g: \mathbb{R}^2 \to \mathbb{R}$ s.t. $g(\alpha) = g\left(\begin{vmatrix} \mu \\ \nu \end{vmatrix}\right) = \frac{\mu}{\nu}$ TO COMPUTE GRADIENT TAKE PARTIAL DERIVATIVES: $\nabla g(\alpha) = \begin{vmatrix} \frac{1}{\nu} \\ -\frac{\mu}{\nu^2} \end{vmatrix}$. THEN ASYMPTOTIC VARIANCE IS

$$\nabla g(\alpha)^T \Sigma \nabla g(\alpha) = \begin{vmatrix} 1/\nu & -\mu/\nu^2 \end{vmatrix} \begin{vmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{vmatrix} \begin{vmatrix} 1/\nu \\ -\mu/\nu^2 \end{vmatrix}$$

$$\nabla g(\alpha)^T \Sigma \nabla g(\alpha) = \begin{vmatrix} \frac{1}{\nu}\sigma_x^2 - \frac{\mu}{\nu^2}\sigma_{xy} & + \frac{1}{\nu}\sigma_{xy} - \frac{\mu}{\nu^2}\sigma_y^2 \end{vmatrix} \begin{vmatrix} \frac{1}{\nu} \\ -\frac{\mu}{\nu^2} \end{vmatrix}$$

$$= \frac{1}{\nu^2}\sigma_x^2 - \frac{\mu}{\nu^3}\sigma_{xy} - \frac{\mu}{\nu^3}\sigma_{xy} + \frac{\mu^2}{\nu^4}\sigma_y^2 = \frac{1}{\nu^2}\sigma_x^2 + \frac{\mu^2}{\nu^4}\sigma_y^2 - 2\frac{\mu}{\nu^3}\sigma_{xy} = V \checkmark$$

BONUS : PS1 Q1.3 (ASKED DURING OH)

$(\hat{\theta}-\theta) = [(\bar{x}-\mu)(\bar{x}-\mu) + \cancel{2\mu(\bar{x}-\mu)}]$ WHAT RATE DO WE NEED?

$=0$ since $\mu=0$

$\therefore \; n(\hat{\theta}-\theta) = n(\bar{x}-\mu)(\bar{x}-\mu) + o_p(1)$

$$n(\hat{\theta}-\theta) = \sqrt{n}(\bar{x}-\mu)\sqrt{n}(\bar{x}-\mu) + o_p(1) \xrightarrow{d} N(0,\sigma^2)^2 = \sigma^2\chi^2(1)$$

WHAT IF SAME ISSUE IN Q2?

$\Rightarrow \begin{cases} E[x_i] = \mu = 0 \\ E[z_i] = c = 0 \end{cases}$ || IN THE QUESTION IT IS y_i.

$(\hat{\theta}-\theta) = (\bar{x}-\mu)(\bar{z}-c) + \mu(\bar{z}-c) + c(\bar{x}-\mu)$

$n(\hat{\theta}-\theta) = \sqrt{n}(\bar{x}-\mu)\cdot\sqrt{n}(\bar{z}-c) \xrightarrow{d} N(0,\sigma^2_x)N(0,\sigma^2_z) = \mathcal{L}$ $\quad E[x_i] = \mu = 0$

$n(\hat{\theta}-\theta)$ converges at FASTER RATE n since $(\bar{x}-\mu) = O_p\left(\frac{1}{\sqrt{n}}\right)$ $\quad (\bar{z}-c) = O_p\left(\frac{1}{\sqrt{n}}\right)$

$\therefore \; (\bar{x}-\mu)(\bar{z}-c) = O_p\left(\frac{1}{\sqrt{n}}\frac{1}{\sqrt{n}}\right) = O_p\left(\frac{1}{n}\right)$. Now, WHAT DISTRIBUTION IS $\mathcal{L}$?

REMEMBER : Product of Two densities is different from the density of The product of Two random variables

a. PRODUCT TWO DENSITIES

IF $X \overset{d}{\equiv} N(a,b)$ $\quad Z \overset{d}{\equiv} N(c,d) \longrightarrow f_x \cdot f_z$ COULD BE PDF OF NORMAL DISTRIBUTION, That is why in Bayesian Inference The product of NORMAL PRIOR × NORMAL LIKELIHOOD gives us a NORMAL POSTERIOR

b. PRODUCT RANDOM VARIABLES $\qquad$ TAUGHT BY GRAHAM !

In This case we are doing a MULTIVARIATE TRANSFORMATION and we have $Y = XZ$ where $X \overset{d}{\equiv} N(a,b)$ $\quad Z \overset{d}{\equiv} N(c,d)$

DISTRIBUTION of This multivariate Transformation : Let

$U = g_1(X,Z)$ $\quad Y = g_2(X,Z) \longrightarrow X = h_1(U,Y)$ $\quad Z = h_2(U,Y)$, and $J = \begin{vmatrix} \partial h_1/\partial u & \partial h_1/\partial y \\ \partial h_2/\partial u & \partial h_2/\partial y \end{vmatrix}$ is The Jacobian, Then PDF of Y would be :

$$f_Y(y) = \int f_{UY}(u,y)\,du = \int f_{XZ}(h_1(u,y), h_2(u,y)) \cdot \det(J)\,du$$

TRICK TO SEMPLYFY THE CALCULATION

$XZ = \frac{1}{4}(X+Z)^2 - \frac{1}{4}(X-Z)^2$ $\quad$ where $\begin{matrix} X+Z \sim N(a+c,\; b^2+d^2+2\sigma_{xz}) \\ X-Z \sim N(a-c,\; b^2+d^2-2\sigma_{xz}) \end{matrix}$

$\therefore \; (X+Z)^2 \sim N(\mu_{x+z}, \sigma^2_{x+z})^2 \implies$ IF $\mu_{x+z} = 0$ THEN $(X+Z)^2 \sim \sigma^2_{x+z}\chi^2(1)$

and $(X-Z)^2 \sim N(\mu_{x-z}, \sigma^2_{x-z})^2 \implies$ IF $\mu_{x-z} = 0$ THEN $(X-Z)^2 \sim \sigma^2_{x-z}\chi^2(1)$

IN OUR CASE IT HOLDS

FINALLY : $Y = XZ \sim \frac{\sigma^2_{x+z}}{4}\chi^2(1) - \frac{\sigma^2_{x-z}}{4}\chi^2(1)$

LINEAR COMBINATION OF 2 $\chi^2(1)$ distributions

Hodge's Estimator

- Assume we have an iid sample, $\{x_1, x_2, \ldots, x_n\}$, from a distribution with mean μ and variance σ^2. To estimate μ, we will consider two estimators. The first one is just the sample mean $\hat{\mu} \equiv \bar{x}$, while the second one is called superefficient estimator:

$$\tilde{\mu} = \begin{cases} \hat{\mu} & \text{if } |\hat{\mu}| > n^{-1/4} \\ 0 & \text{if } |\hat{\mu}| \leq n^{-1/4} \end{cases}$$

(1) Derive the asymptotic distribution of $\hat{\mu}$ and the asymptotic MSE.

(2) Show that $\mu \neq 0 \Rightarrow \mathbb{P}(\tilde{\mu} = \hat{\mu}) \to 1$ and $\therefore \sqrt{n}(\tilde{\mu} - \mu) \xrightarrow{d} \mathcal{N}(0, \sigma^2)$

(3) Show that $\mu = 0 \Rightarrow \mathbb{P}(\tilde{\mu} = \hat{\mu}) \to 0$ and $\therefore \sqrt{n}(\tilde{\mu} - \mu) \xrightarrow{p} 0$.

(4) Assume $\mu = cn^{-1/3}$, what is the asymptotic MSE of $\tilde{\mu}$?

(1) $\hat{\mu}$ is just standard sample mean $\bar{x} \sim N\left(\mu ; \frac{\sigma^2}{m}\right) \therefore \sqrt{m}(\bar{x}-\mu) \sim N(0,\sigma^2)$

$\bar{x}$ IS UNBIASED $\therefore$ MSE

(2) SHOW THAT $\mu \neq 0$ $\quad P(\tilde{\mu} = \hat{\mu}) \xrightarrow{P} 1$

$-m^{1/2 - 1/4} = m^{1/4}$

$\underline{P(\tilde{\mu}=\hat{\mu}) = P(|\hat{\mu}| > m^{-1/4})} = P(\hat{\mu} < \underline{-m^{-1/4}} \text{ or } \hat{\mu} > \underline{m^{-1/4}})$

SUBTRACT μ AND MULTIPLY BY $\sqrt{m}$ ON EACH SIDE

$\tilde{\mu} = \hat{\mu}$ IFF $|\hat{\mu}| \geq m^{-1/4}$

$= P(\sqrt{m}(\hat{\mu}-\mu) < -m^{1/4} - \sqrt{m}\mu \text{ or } \sqrt{m}(\hat{\mu}-\mu) > m^{1/4} - \sqrt{m}\mu)$

$\geq P(\underbrace{\sqrt{m}(\hat{\mu}-\mu)}_{\xrightarrow{d} N(\cdot)} > \underbrace{m^{1/4} - \sqrt{m}\mu}_{\text{EXPLODE}})$

EXPLODE $\lim_{m\to\infty} m^{1/4} - \sqrt{m}\mu = \lim_{m\to\infty} m^{1/2}\left(\cancelto{0}{\frac{1}{m^{1/2}}} - \mu\right) = \pm\infty$

THEN 2 CASES

(i) $\mu > 0$ $\quad P(\sqrt{m}(\hat{\mu}-\mu) > m^{1/4} - \sqrt{m}\mu) \xrightarrow{A.S.} 1$

$\lim_{m\to\infty} \sqrt{m}\left(-\mu - \cancel{\frac{m^{1/4}}{m^{1/2}}}\right) = -\mu\sqrt{m}$ $\quad = 0 \because m^{1/2} \gg m^{1/4}$

(ii) $\mu < 0$ $\quad$ I consider $\quad P(\sqrt{m}(\hat{\mu}-\mu) < -m^{1/4} - \sqrt{m}\mu) \overset{m\to\infty}{=} P(\sqrt{m}(\hat{\mu}-\mu) < -\mu\sqrt{m} = +\infty) \longrightarrow 1$

$\therefore$ IF $\mu \neq 0$ THEN $\tilde{\mu}$ BEHAVES ALMOST SURELY as $\hat{\mu}$: $\sqrt{m}(\tilde{\mu}-\mu) \xrightarrow{d} N(0,\sigma^2)$

Remember that converges in probability implies converges in distribution.

(3) SHOW THAT $\mu = 0 \implies P(\tilde{\mu}=\hat{\mu}) \xrightarrow{P} 0$

$P(\tilde{\mu}=\hat{\mu}) = P(|\hat{\mu}| > m^{-1/4}) = P(\hat{\mu} < -m^{-1/4} \text{ or } \hat{\mu} > m^{-1/4})$

$= P(\sqrt{m}\hat{\mu} < -m^{1/4} \text{ or } \sqrt{m}\hat{\mu} > m^{1/4}) \xrightarrow{P} P(N(0,\sigma^2) < -\infty \text{ or } N(0,\sigma^2) > +\infty) = 0$

$\therefore$ when $\mu = 0$ THEN $P(\tilde{\mu} = 0) \longrightarrow 1$ since $\tilde{\mu}$ only takes two possible values

and $P(\tilde{\mu} = 0) = 1 - P(\tilde{\mu} = \hat{\mu}) = 1$

we also have that $P(\sqrt{m}\tilde{\mu} = \sqrt{m}\,0) \longrightarrow 1$ $\quad$ (in probability theory $\lim_{m\to\infty} 0 \cdot m = 0$)

This means $\sqrt{m}\tilde{\mu} \xrightarrow{P} 0$

$\sqrt{m}(\tilde{\mu}-\mu) \xrightarrow{P} 0$ $\quad$ ACTUALLY $m^{\alpha}(\tilde{\mu}-\mu) \xrightarrow{P} 0$ $\quad \forall \alpha \in \mathbb{R}$ since $m^{\alpha}\, 0 \xrightarrow{m\to 0} 0$ $\quad \forall \alpha \in \mathbb{R}$

In conclusion, we have
$$\begin{cases} m^{\alpha}(\tilde{\mu}-\mu) \xrightarrow{P} 0, \alpha \in \mathbb{R} & \text{IF } \mu = 0 \\ \sqrt{m}(\tilde{\mu}-\mu) \xrightarrow{d} N(0,\sigma^2) & \text{IF } \mu \neq 0 \end{cases}$$

(4) Is $\tilde{\mu}$ better than $\hat{\mu}$? NO, here there is why:

ASSUME $\mu = cm^{-1/3} \longrightarrow 0$ $\quad$ and we know $E[\hat{\mu}] = cm^{-1/3}$

NOTE THAT $\quad cm^{-1/3} < m^{-1/4}$ (under some assumption on c.) $\therefore \tilde{\mu} = 0$

$|\hat{\mu}| \leq m^{-1/4} \therefore \tilde{\mu} = 0$

$\therefore \sqrt{m}(\tilde{\mu}-\mu) = \cancelto{=0}{\sqrt{m}\tilde{\mu}} - cm^{-1/3} \cdot m^{1/2} = \sqrt{m}\tilde{\mu} - \overset{\frown}{c m^{1/6}} = 0 - cm^{1/6} \longrightarrow \pm\infty$

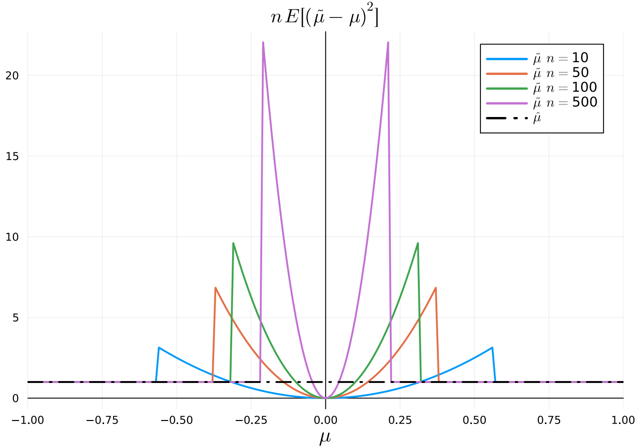
$\therefore$ MSE explodes, we can do systematically worse than $\hat{\mu}$ IN FINITE SAMPLE since

$MSE[\tilde{\mu}] = E[(\sqrt{m}(\tilde{\mu}-\mu))^2] = E[c^2 m^{2/6}] \longrightarrow +\infty$

HODGE'S ESTIMATOR is consistent for μ, it's asymptotic distribution is the same as $\hat{\mu}$ except for $\mu = 0$ where the RATE OF CONVERGENCE BECOMES ARBITRARILY FAST, BUT AMSE POTENTIALLY UNBOUNDED!

Deceiving Asymptotics

Let's assume $\{x_1, x_2, \ldots, x_n\}$ iid with $x_i \sim \mathcal{N}(\mu, 1)$. This is the asymptotic MSE (scaled by the rate n) of the Hodge's estimator for different values of μ.



$n \uparrow$

AMSE $\uparrow \infty$