$$\frac{1}{n}\sum_{i=1}^{n} \underbrace{h(X_i)}_{Y_i} \to E[h(X_i)]$$

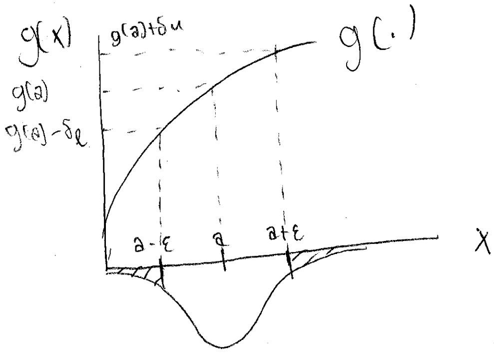
what about $g\left(\sum_{i=1}^{n} h(X_i)\right)$

<u>Thm 5.5.4</u> If $g(\cdot)$ is continuous everywhere and $X_1, X_2, \dots$ is such that $\operatorname{plim} X_n = a$, then $g(X_1), g(X_2), \dots$ is such that $\operatorname{plim} g(X_n) = g(a)$.

<u>Remarks:</u>

$\boxed{1}$ Given $\lim_{n\to\infty} \Pr[|X_n - a| > \varepsilon] = 0 \quad \forall \varepsilon > 0$

need $\lim_{n\to\infty} \Pr[|g(X_n) - g(a)| > \delta] = 0 \quad \forall \delta > 0$



e.g. $\operatorname{plim} X_n = a$, $\quad \Pr[X_n > 0] = 1 \quad \forall n$

Let $Y_n = g(X_n) = \sqrt{X_n}$

$$\Pr[|\sqrt{X_n} - \sqrt{a}| < \delta] = \Pr[|(\sqrt{X_n} - \sqrt{a})(\sqrt{X_n} + \sqrt{a})| < \delta|\sqrt{X_n} + \sqrt{a}|]$$

$$= \Pr[|X_n - a| < \delta|\sqrt{X_n} + \sqrt{a}|]$$

$$\geq \Pr[|X_n - a| < \underbrace{\delta\sqrt{a}}_{=\varepsilon}]$$

$$\to 1$$

2] a.s. convergence. It is easier to establish these results wrt a.s. convergence. (sets of measure zero.)

3] $\text{plim } X_n = a$, $\text{plim } Y_n = b$

a) $\text{plim}(X_n + Y_n) = a + b$

b) $\text{plim}(X_n Y_n) = ab$

c) $\text{plim}\left(\frac{X_n}{Y_n}\right) = \frac{a}{b}$ if $b \neq 0$

Thm 5.5.17 (Slutsky)

If $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{P} a$, then

a) $Y_n X_n \xrightarrow{d} aX$

b) $Y_n + X_n \xrightarrow{d} a + X$

e.g. $\sqrt{n}\left(\frac{\bar{X}_n - \mu}{\sigma}\right) \xrightarrow{\sigma} N(0,1)$

Let $s_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2 \xrightarrow{P} \sigma^2$

$$t = \frac{\sqrt{n}(\bar{X}_n - \mu)}{s_n} = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \underbrace{\frac{\sigma}{s_n}}_{\xrightarrow{P} 1} \xrightarrow{d} N(0,1)$$

Recall: approximation

n-1	t(0.025)	Z(0.025)
15	-2.13	-1.96
30	-2.04	-1.96
120	-1.98	-1.96

Thm (Continuous mapping thm) If $X_n \xrightarrow{d} X$ and $g(\cdot)$ is continuous on support of X, then $g(X_n) \xrightarrow{d} g(X)$

e.g. Suppose $Z_n = \sqrt{n}\left(\frac{\bar{X}_n - \mu}{\sigma}\right) \xrightarrow{d} N(0,1)$

$g(Z_n) = (Z_n)^2$, so $Z_n^2 \xrightarrow{d} \chi^2(1)$

Thm 5.5.4 (Delta method) If X_n is s.t. $\sqrt{n}(X_n - a) \xrightarrow{d} N(0, \sigma^2)$. If $g(\cdot)$ is differentiable at a, if $\frac{\partial g(a)}{\partial a} \neq 0$, then $\sqrt{n}(g(X_n) - g(a))$
$\xrightarrow{d} N\left(0, \left(\frac{\partial g}{\partial a}(a)\right)^2 \sigma^2\right)$.

- To see how this works,

1. Taylor series expansion:

$$g(X_n) = g(a) + \frac{\partial g(a)}{\partial a}(X_n - a) + \underbrace{\frac{1}{2}\frac{\partial^2 g(a)}{\partial^2 a}(X_n - a)^2 + \cdots}_{\text{ignore this}}$$

$$\Rightarrow \sqrt{n}(g(X_n) - g(a)) = \frac{\partial g(a)}{\partial a}\sqrt{n}(X_n - a)$$

$$\xrightarrow{d} N\left(0, \left(\frac{\partial g(a)}{\partial a}\right)^2 \sigma^2\right)$$

$$\sqrt{n}(X_n - a) = O_p(1)$$

$$(\sqrt{n}(X_n - a))^2 = O_p(1)O_p(1) = O_p(1)$$

$$\Rightarrow n(X_n - a)^2 = O_p(1) \quad \Rightarrow \sqrt{n}(X_n - a)^2 = o_p(1)$$

This is "why" we can ignore higher-order terms.

2] Mean value expansion:

$$g(\bar{X}_n) \cong g(a) + \frac{\partial g(\bar{X}_n^*)}{\partial a}(\bar{X}_n - a)$$

where $\underbrace{\bar{X}_n^* \in \text{conv}\{\bar{X}_n, a\}}$

convex hull of $\bar{X}_n$ and a

"on the line segment connecting $\bar{X}_n$ and a"

$$\Rightarrow \sqrt{n}\left(g(\bar{X}_n) - g(a)\right) \cong \frac{\partial g(\bar{X}_n^*)}{\partial a} \underbrace{\sqrt{n}(\bar{X}_n - a)}$$

$$\rightarrow \frac{\partial g}{\partial a}(a) \qquad \xrightarrow{d} N(0, \sigma^2)$$

if we assume g is C^1,

e.g. 5.5.23

$$\sqrt{n}\left(\frac{\bar{X}_n - a}{\sigma}\right) \xrightarrow{d} N(0, 1) \qquad g(\bar{X}_n) = \frac{1}{\bar{X}_n}$$

$$g'(a) = -\frac{1}{a^2}$$

$$\Rightarrow \sqrt{n}\left(\frac{1}{\bar{X}_n} - \frac{1}{a}\right) \xrightarrow{d} N\left(0, \frac{\sigma^2}{a^4}\right)$$

Suppose $g(\bar{X}_n) = \bar{X}_n^2 \qquad g'(a) = 2a$

$$\Rightarrow \sqrt{n}\left(\bar{X}_n^2 - a^2\right) \xrightarrow{d} N(0, 4a^2\sigma^2)$$

Multivariate delta method: Spse $\sqrt{n}(X_n - a) \xrightarrow{d} N(0, \Sigma)$ and $g: \mathbb{R}^k \to \mathbb{R}^l$ $l \le k$. Spse g is differentiable at a with Jacobian $\underset{l \times k}{A}$. Then

$$\sqrt{n}(g(X_n) - g(a)) \xrightarrow{d} N(0, A \Sigma A').$$

$T(X) = T(X_1, \ldots, X_n)$

e.g. $T(X) = \underbrace{g\Big(\underbrace{\sum_{i=1}^{n} \underbrace{f(X_i)}_{Y_i}}_{\text{use CLTs and LLNs}}\Big)}_{\text{use } \delta\text{-method, continuous mapping thm.}}$

Davidson: Stochastic Limit theory

- Great reference for econometric theory.

$T(X)$

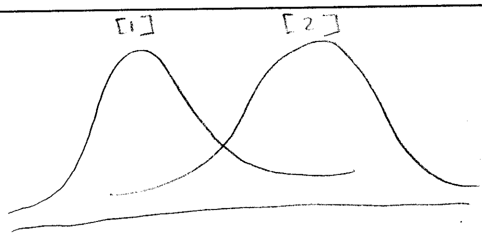
Exact

1. Lots of work
2. Depends critically on assumptions
3. Results are exact

asymptotic

1. much easier
2. general assumptions on X_i's
3. Only an approximation

Myths: ① Parsiomony of assumptions is a great reason to do asymptotics over exact theory.



[1] assumption set 1
[2] assumption set 2

- There is no single distribution that can approximate both of these distributions.

The real reason we do asymptotics is that it is easier.

<u>Statistics:</u> What does $T(\underline{X})$ look like?

- We will construct useful $T(\underline{X})$ functions.

Statistical Inference

- point estimation
- hypothesis testing
- confidence intervals

Point estimation

<u>Defn:</u> a <u>point estimator</u> is a function $W(\underline{X}_1, \ldots, \underline{X}_n)$ of a sample. (ie any statistic is a potential estimator.)

Methods of deriving estimators