Teoría Asintótica

Fundamentos de Econometría

Juan Palomino[1]

[1]Magister en Economía Aplicada con Mención Estudios Regionales
juan.palominoh@pucp.pe

Departamento de Economía

Índice

1. Teoría Asintótica
   - Convergencia de Secuencias Determinísticas
   - Convergencia en Probabilidad
   - Ley de los Grandes Números
   - Convergencia en Distribución
   - Teorema del Límite Central
2. Propiedades Asintóticas de MCO
   - Consistencia
   - Normalidad
   - Consistencia de $\hat{\sigma}^2$
   - Estimación Consistente de la matriz de covarianza

Previamente

- Muestra finita.
- Propiedades del estimador MCO.
- Término de error es normalmente distribuido e independientes de las variables explicativas.

Objetivo

- Obtener resultados aproximados del comportamiento de la distribución de un estimador a medida que la muestra va creciendo.
- Es necesario referirnos a la teoría asintótica:
    - Derivaciones de las propiedades asintóticas o propiedades de grandes muestras de estimadores y test estadísticos.
    - Situación: $n \to \infty$

Ejemplo

- Convergencia de una secuencia de una variable aleatoria $X_n(\omega)$, donde $\omega \in \Omega$, y Ω es el espacio muestral.
- Experimento de lanzar una moneda repetidas veces:

$$X_n(\omega) = \frac{\#\ caras\ después\ de\ n\ lanzadas}{n}$$

- donde $\omega = una\ secuencia\ infinita\ de\ C\ y\ S$. Si:

$$\omega = \{C, C, S, C, S, C, S, C, C, S, S, C\}$$

Ejemplo

- Entonces, las veces que salga cara será:

$$X_1(\omega)=\frac{1}{1}\ \ X_2(\omega)=\frac{2}{2}\ \ X_3(\omega)=\frac{2}{3}\ \ X_4(\omega)=\frac{3}{4}\ \ X_5(\omega)=\frac{3}{5}\ \ X_6(\omega)=\frac{4}{6}$$
$$X_7(\omega)=\frac{4}{7}\ \ X_8(\omega)=\frac{5}{8}\ \ X_9(\omega)=\frac{6}{9}\ \ X_{10}(\omega)=\frac{6}{10}\ \ X_{11}(\omega)=\frac{6}{11}\ \ X_{12}(\omega)=\frac{7}{12}$$

- Observando la secuencia, el sentido común dice que a medida que $n\to\infty$, $x_n(\omega)$ podría converger a 1/2.
- Mayoría de veces denotamos $X_n(\omega)$ simplemente como X_n.

Idea

- La idea es similar en términos de estimadores:

$$\hat{\beta}_1 = (X_1^{'}X_1)^{-1}X_1^{'}y_1$$
$$\hat{\beta}_2 = (X_2^{'}X_2)^{-1}X_2^{'}y_2$$
$$\vdots$$
$$\hat{\beta}_n = (X_n^{'}X_n)^{-1}X_n^{'}y_n$$

- Notar que indexamos el estimador por n indicando que este depende del tamaño de la muestra. Es decir, es una secuencia estocástica.

Índice

1 Teoría Asintótica
- Convergencia de Secuencias Determinísticas
- Convergencia en Probabilidad
- Ley de los Grandes Números
- Convergencia en Distribución
- Teorema del Límite Central

2 Propiedades Asintóticas de MCO
- Consistencia
- Normalidad
- Consistencia de $\hat{\sigma}^2$
- Estimación Consistente de la matriz de covarianza

Convergencia de Secuencias Deterministicas

- Existen dos tipos de sucesión: determinísticas o estocásticas.
- Una sucesión de valores $X_1, X_2, \ldots, X_n$ denotada $\{X_n\}_{n=1}^{\infty}$ o simplemente $\{X_n\}$, es una función de X: $\mathbb{N} \rightarrow \mathbb{R}^n$.

Definición 1: Convergencia Determinística

Una secuencia de valores determinísticos $\{X_n\}$ converge al límite c si para cualquier $\varepsilon > 0$, existe $n^* = n^*(\varepsilon)$ tal que para todo $n > n^*$,

$$|X_n - c| < \varepsilon$$

Es decir, esto ocurre cuando los límites de estas sucesiones existen, $\underset{n\rightarrow\infty}{lim} X_n = c$, donde c es una constante.

Convergencia de Secuencias Deterministicas

Convergencia Determinística

Si $X_n = 2 + \frac{3}{n}$, entonces el límite es 2 ya que $|X_n - c| = \left|2 + \frac{3}{n} - 2\right| = \left|\frac{3}{n}\right| < \varepsilon$ para todo $n > n^* = \frac{3}{\varepsilon}$

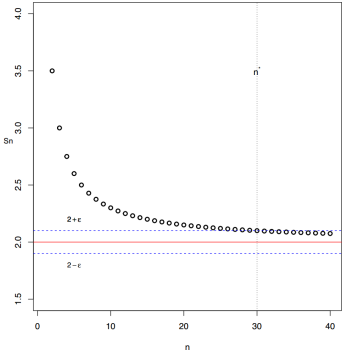

Figura: Convergencia de secuencia $2 + \frac{3}{n}$

Índice

1 Teoría Asintótica
- Convergencia de Secuencias Determinísticas
- Convergencia en Probabilidad
- Ley de los Grandes Números
- Convergencia en Distribución
- Teorema del Límite Central

2 Propiedades Asintóticas de MCO
- Consistencia
- Normalidad
- Consistencia de $\hat{\sigma}^2$
- Estimación Consistente de la matriz de covarianza

Convergencia en Probabilidad

- Anteriormente hemos visto convergencia de números, pero ¿qué hay de las secuencias de variables aleatorias?
- Para una secuencia de variables aleatorias no podemos estar seguros de que $|X_n - c| < \varepsilon$, incluso para n grande, debido a la aleatoriedad.

Definición 2: Convergencia en Probabilidad

Una secuencia de variables aleatorias $\{X_n\}$ converge en probabilidad a una constante (no aleatoria) α si, para cualquier $\varepsilon > 0$,

$$\lim_{n \to \infty} Pr(|X_n - \alpha| > \varepsilon) = 0$$

La constante α se conoce como el límite de probabilidad de X_n, y es escrito como $plim X_n = \alpha$ o $X_n \xrightarrow{p} \alpha$.

$$X_n \xrightarrow{p} \alpha \quad \textit{es lo mismo que} \quad X_n - \alpha \xrightarrow{p} 0$$

Teorema del Mapeo Continuo

Teorema 1: Teorema del Mapeo Continuo

Dada una función continua $g(X)$, si

$$X_n \xrightarrow{p} X \Longrightarrow g(X_n) \xrightarrow{p} g(X)$$

a medida que $n \longrightarrow \infty$, o equivalente,

$$plim[g(X_n)] = g[plim(X_n)]$$

Estimador Consistente

Definición 3: Estimador Consistente

Un estimador $\hat{\theta}_n$ de un parámetro θ es un estimador consistente si y solo si:

$$plim\hat{\theta}_n = \theta$$

- La convergencia en la probabilidad también se conoce como **consistencia débil**, y dado que este ha sido el concepto de convergencia estocástica más familiar en econometría, la palabra "débil" a menudo simplemente se descarta.

Convergencia Fuerte en Probabilidad

Definición 4: Convergencia Fuerte en Probabilidad

Una secuencia de variables aleatorias $\{X_n\}$ converge en probabilidad fuerte o casi segura a una constante (no aleatoria) α si para cualquier $\varepsilon > 0$,

$$Pr(\underset{n\to\infty}{lim} X_n = \alpha) = 1$$

Se escribe como $X_n \xrightarrow{c.s} \alpha$, a medida que $n \to \infty$. Una condición equivalente para la convergencia casi segura es:

$$limPr(|X_m \to X| < \varepsilon)\ \forall m \geq n) = 1$$

- Este concepto es más fuerte que convergencia en probabilidad, es decir, si una secuencia converge casi segura, este converge en probabilidad:

$$\xrightarrow{c.s} \Longrightarrow \xrightarrow{p}$$

Convergencia en Media Cuadrática

Teorema 2: Convergencia en Media Cuadrática

Si X_n tiene media μ_n y varianza σ_n^2 tal que los límites ordinarios de μ_n y σ_n^2 son c y 0, respectivamente, entonces X_n converge en media cuadrática a c,

$$X_n \xrightarrow{m.c} c$$

y

$$plim X_n = c$$

- Este teorema implica que $X_n \xrightarrow{m.c} c \Longrightarrow X_n \xrightarrow{p} c$.

Convergencia en Media Cuadrática

- **Demostración:** Decimos que $\{X_n\}$ una sucesión de variables aleatorias i.i.d. converge en media cuadrática al límite fijo c si:

$$limE[(X_n - c)^2] = 0$$

- Partiendo de $(X_n - c)^2$ y tomando el valor esperado tenemos:

$$E[(X_n - c)^2] = Var(X_n - c) + (E(X_n) - c)^2$$

- Tomando límites:

$$\lim_{n\to\infty} E[(X_n - c)^2] = \lim_{n\to\infty} Var(X_n) + \lim_{n\to\infty} (E(X_n) - c)^2$$

- Si $\lim_{n\to\infty} Var(X_n) = 0$ y $\lim_{n\to\infty} E(X_n) = c$ se cumple la definición de convergencia media cuadrática.

Consistencia de la media muestral

Teorema 3: Consistencia de la media muestral

La media de una muestra aleatoria de cualquier población con media finita μ y varianza finita σ^2 es un estimador consistente de μ.

- **Demostración:** ya que $E[\bar{X}_n] = \mu$ y $Var(\bar{X}_n) = \frac{\sigma^2}{n}$. Asimismo, usando el teorema de la convergencia en media cuadratica:

$$\bar{X}_n \xrightarrow{m.c} \mu \Longrightarrow \bar{X}_n \xrightarrow{p} \mu$$

Consistencia del Estimador Insesgado

Teorema 4: Consistencia del Estimador Insesgado

Un estimador insesgado $\hat{\theta}_n$ es consistente si

$$lim_{n\to\infty} Var(\hat{\theta}(X_1,\ldots,X_n)) = 0$$

Reglas para límites de probabilidad

- Si X_n e Y_n son variables aleatorias con $X_n \xrightarrow{p} c$ y $Y_n \xrightarrow{p} d$:
  1. Regla de suma:
  $$X_n + Y_n \xrightarrow{p} c + d$$
  2. Regla de producto:
  $$X_n Y_n \xrightarrow{p} cd$$
  3. Regla del ratio
  $$X_n / Y_n \xrightarrow{p} c/d \;si\; d \neq 0$$
  4. Regla de matriz inversa: Si W_n es una matriz cuyo elementos son variables aleatorias y si $W_n \xrightarrow{p} \Omega$, entonces:
  $$W_n^{-1} \xrightarrow{p} \Omega^{-1}$$
  5. Regla de producto de matrices: Si X_n e Y_n son matrices aleatorias con $X_n \xrightarrow{p} A$ y $Y_n \xrightarrow{p} B$, entonces:
  $$X_n Y_n \xrightarrow{p} AB$$

Índice

1. Teoría Asintótica
   - Convergencia de Secuencias Determinísticas
   - Convergencia en Probabilidad
   - Ley de los Grandes Números
   - Convergencia en Distribución
   - Teorema del Límite Central
2. Propiedades Asintóticas de MCO
   - Consistencia
   - Normalidad
   - Consistencia de $\hat{\sigma}^2$
   - Estimación Consistente de la matriz de covarianza

Ley de los Grandes Números

- Es un teorema para convergencia en probabilidad en el caso especial donde la secuencia $\{X_n\}$ es un promedio muestral, es decir, $X_n = \overline{X}_n$, donde:

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Ley de los Grandes Números

- Al respecto, suponer una muestra aleatoria simple de una variable aleatoria X, siendo esta muestra de tamaño n.
- La muestra es la colección de $X_1, X_2, ..., X_n$.
- Supongamos que cada X_i tiene como media a μ_x y como varianza a σ_x^2. El promedio de esta muestra es:

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Ley de los Grandes Números

- El valor esperado de la media muestral es:

$$E[\bar{X}_n] = E[\frac{1}{n}\sum_{i=1}^{n} X_i] = \frac{1}{n}\sum_{i=1}^{n} E[X_i] = \frac{1}{n} n\mu_x = \mu_x$$

- La varianza del promedio muestral es:

$$\begin{aligned} Var[\bar{X}_n] &= Var[\frac{1}{n}\sum_{i=1}^{n} X_i] = \frac{1}{n^2} Var[\sum_{i=1}^{n} X_i] \\ &= \frac{1}{n^2}[\sum_{i=1}^{n} Var[X_i] + 2\sum_{i<j}\sum Cov(X_i, X_j)] \\ &= \frac{1}{n^2}[\sum \sigma_x^2] = \frac{1}{n^2} n\sigma_x^2 = \frac{\sigma_x^2}{n} \end{aligned}$$

Ley de los Grandes Números

Teorema 5: Ley Débil de los Grandes Números (Ley de Khinchine)

Sea $\{X_n\}$ una muestra aleatoria i.i.d con $E(X_i) = \mu$ y sea $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. Entonces:

$$\lim_{n\to\infty} Pr[|X_n - \mu| > \varepsilon] = 0$$

O equivalente

$$\lim_{n\to\infty} Pr[|X_n - \mu| \leq \varepsilon] = 1$$

En otras palabras,

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{p} E(X_i)$$

o $plim \bar{X}_n = \mu$. Lo que evidencia que el estimador $\hat{\mu} = \bar{X}_n$ converge en probabilidad a la verdadera media de la población μ.

Ejemplo 1

Ejemplo de media de una normal

Considere n muestras con pdf $N(1,2)$. Por ejemplo X_1 es la primera muestra con solo una observación que viene de una $N(1,2)$, X_2 es la segunda muestra con dos observaciones (X_1, X_2) que también proviene de una $N(1,2)$; X_3 con tres observaciones (X_1, X_2, X_3) y así. Notar que cada muestra (o secuencia) es una muestra aleatoria i.i.d con $E(X_i) = \mu$. La muestra para cada secuencia es también una secuencia:

$$\bar{X}_1 = g(X_1) = X_1$$

$$\bar{X}_2 = g(X_1, X_2) = \frac{1}{2}\sum_{i=1}^{2} X_i$$

$$\bar{X}_3 = g(X_1, X_2, X_3) = \frac{1}{3}\sum_{i=1}^{3} X_i$$

$$\vdots$$

$$\bar{X}_N = g(X_1, X_2, ..., X_n) = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Ejemplo 1

```
# Setup
set.seed(123)  # Set the seed
N <- 10000
n <- 1:N       # vector: n = 1, 2, ..., N

n_dat <- rnorm(n = n, mean = 1, sd = 0.5) # Sample from N(1, 0.5^2)
xbar <- cumsum(n_dat) /  n                # Cumulated mean
plot(n, xbar, type = "l")
abline(h = 1.01, col = "blue", lty = 2)
abline(h = 0.99, col = "blue", lty = 2)
```

Figura: Simulación 1: R-Script de la media de una normal

Ejemplo 1

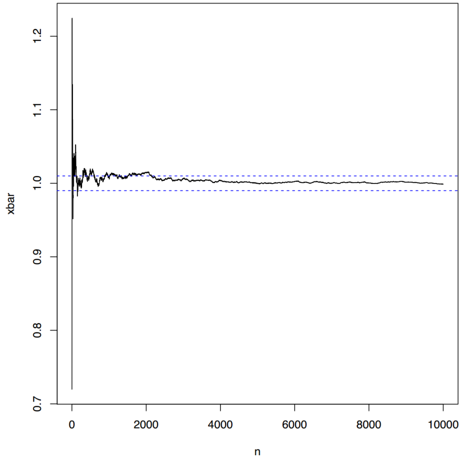

Figura: Convergencia de la media de una distribución normal ($\varepsilon = 0{,}01$)

Ejemplo 2

Lanzando una Moneda

Ahora simulamos $N = 10000$ lanzamientos de monedas. Después de cada lanzamiento simulado, gráficamos la proporción X_n de caras obtenidas contra el número n de lanzamientos total. La LGN dice que deberíamos ver un rastro que se acerca mucho a 1/2 a medida que aumenta n.

Ejemplo 2

```
# Set up
set.seed(123)
N <- 10000                          # total number of tosses
n <- 1:N                            # vector: n = 1, 2, ...., N; Toss number

# Simulate and plot
h <- rbinom(n = n, size = 1, prob = 1/2) # vector: H = 0 or 1 each with p = 1./2
x <- cumsum(h) / n                  # vector: proportion of heads
plot(n, x, type = "l", ylim = c(0, 1))
abline(h = 0.52, col = "blue", lty = 2)
abline(h = 0.48, col = "blue", lty = 2)
```

Figura: Simulación 2: R-Script de distribución binomial

Ejemplo 2

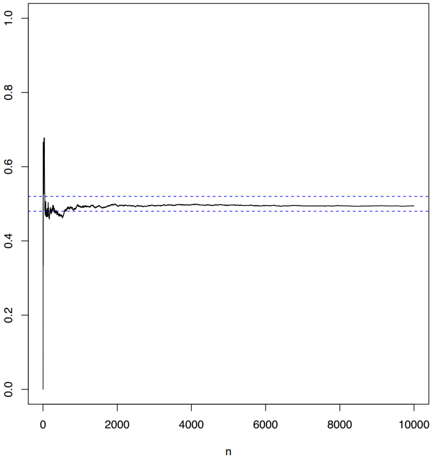

Figura: Convergencia de la media de una distribución binomial ($\varepsilon = 0{,}02$)

Índice

1. Teoría Asintótica
   - Convergencia de Secuencias Determinísticas
   - Convergencia en Probabilidad
   - Ley de los Grandes Números
   - Convergencia en Distribución
   - Teorema del Límite Central
2. Propiedades Asintóticas de MCO
   - Consistencia
   - Normalidad
   - Consistencia de $\hat{\sigma}^2$
   - Estimación Consistente de la matriz de covarianza

Convergencia en Distribución

Definición: Convergencia en Distribución

Si la función de distribución acumulada F_{X_n} de la secuencia de variables aleatorias $\{X_n\}$ converge a la función de distribución acumula F_X a medida que $n \longrightarrow \infty$ en todos los puntos z donde $F_X(z)$ es continuo, entonces $\{X_n\}$ converge en distribución a X. Se denota como:

$$X_n \xrightarrow{d} X$$

o

$$\lim_{n \to \infty} |F_{X_n} - F_X| = 0$$

Convergencia en Distribución

Teorema: Convergencia en probabilidad implica convergencia en distribución

Si la secuencia de variables aleatorias $\{X_n\}$ converge en probabilidad a la variable aleatoria X, la secuencia también converge en distribución a X.

$$X_n \xrightarrow{p} X \Rightarrow X_n \xrightarrow{d} X$$

Índice

1 Teoría Asintótica
- Convergencia de Secuencias Determinísticas
- Convergencia en Probabilidad
- Ley de los Grandes Números
- Convergencia en Distribución
- Teorema del Límite Central

2 Propiedades Asintóticas de MCO
- Consistencia
- Normalidad
- Consistencia de $\hat{\sigma}^2$
- Estimación Consistente de la matriz de covarianza

Teorema del Límite Central

Teorema del Límite Central

Sea $\{X_n\}$ una secuencia de variables aleatorias i.i.d.. Si $E(X_n) = \mu$ y la varianza es estrictamente positiva y finita, $0 < \sigma^2 < \infty$. Se define $\bar{X}_n = n^{-1}\sum_{i=1}^{n} X_i$. Entonces la distribución de:

$$\begin{aligned} Z_n &= \frac{\bar{X}_n - E(\bar{X}_n)}{\sqrt{Var(\bar{X}_n)}} \\ &= \frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} \\ &= \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{d} N(0,1) \end{aligned}$$

- A medida que n se aproxima al infinito. Esto es lo mismo que:

$$\sqrt{(\bar{X}_n - \mu)} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (X_i - \mu) \xrightarrow{d} N(0, \sigma^2)$$

Índice

1. Teoría Asintótica
   - Convergencia de Secuencias Determinísticas
   - Convergencia en Probabilidad
   - Ley de los Grandes Números
   - Convergencia en Distribución
   - Teorema del Límite Central
2. Propiedades Asintóticas de MCO
   - Consistencia
   - Normalidad
   - Consistencia de $\hat{\sigma}^2$
   - Estimación Consistente de la matriz de covarianza

Consistencia

Teorema: Consistencia de estimadores MCO

Bajo el supuesto de linealidad, muestra aleatoria, ortogonalidad, condición de rango, los estimadores MCO son consistentes. Es decir:

$$\hat{\beta}_n \xrightarrow{p} \beta_0$$

Consistencia

- Prueba: dado los supuestos de linealidad y condición de rango, el error muestral es dado por:

$$\hat{\beta}_n - \beta_0 = (X^{'}X)^{-1}X^{'}\varepsilon$$
$$= (\frac{1}{n}X^{'}X)^{-1}(\frac{1}{n}X^{'}\varepsilon)$$
$$= (\frac{1}{n}\sum_{i=1}^{n} x_i x_i^{'})^{-1}(\frac{1}{n}\sum_{i=1}^{n} x_i \varepsilon_i)$$

- Ya que el error muestral está en términos de promedios, aplicamos el Teorema del Límite Central:

$$plim(\hat{\beta}_n - \beta_0) = plim(\frac{1}{n}\sum_{i=1}^{n} x_i x_i^{'})^{-1} plim\left(\frac{1}{n}\sum_{i=1}^{n} x_i \varepsilon_i\right)$$
$$= \left(plim\frac{1}{n}\sum_{i=1}^{n} x_i x_i^{'}\right)^{-1} plim\left(\frac{1}{n}\sum_{i=1}^{n} x_i \varepsilon_i\right)$$

Consistencia

- Ya que $\{x_i\}$ es una muestra aleatoria i.i.d, entonces $\{x_i x_i^{'}\}$ son también secuencias i.i.d. Asimismo, aplicamos la Ley Débil de Grandes Números:

$$\frac{1}{n}\sum_{i=1}^{n} x_i x_i^{'} \xrightarrow{p} E[x_i x_i^{'}] = Q_{xx}$$

- Este existe por el supuesto de condición rango. Ya que Q_{xx} es invertible:

$$\left(\frac{1}{n}\sum_{i=1}^{n} x_i x_i^{'}\right)^{-1} \xrightarrow{p} Q_{xx}^{-1}$$

- El producto $\{x_i \varepsilon_i\}$ es i.i.d. (ya que las observaciones son i.i.d). Asimismo, aplicamos la Ley Débil de Grandes Números:

$$\frac{1}{n}\sum_{i=1}^{n} x_i \varepsilon_i \xrightarrow{p} E[x_i \varepsilon_i] = 0$$

- por supuesto de condición de ortonalidad. Por lo tanto:

$$\hat{\beta}_n - \beta_0 \xrightarrow{p} Q_{xx}^{-1} 0 = 0$$

Índice

1. Teoría Asintótica
   - Convergencia de Secuencias Determinísticas
   - Convergencia en Probabilidad
   - Ley de los Grandes Números
   - Convergencia en Distribución
   - Teorema del Límite Central
2. Propiedades Asintóticas de MCO
   - Consistencia
   - Normalidad
   - Consistencia de $\hat{\sigma}^2$
   - Estimación Consistente de la matriz de covarianza

Normalidad

Teorema: Normalidad Asintótica del Estimador MCO

Bajo los supuestos de linealidad, muestra aleatoria, ortogonalidad, condición de rango, momentos finitos, a medida que $n \to \infty$:

$$\sqrt{n}(\hat{\beta}_n - \beta_0) \xrightarrow{d} N(0, V_\beta)$$

donde $V_\beta = Q_{xx}^{-1} \Omega Q_{xx}^{-1}$, $Q_{xx} = E[x_i x_i^{'}]$, y $\Omega = E[\varepsilon_i^2 x_i x_i^{'}]$.

Normalidad

- Demostración: Multiplicando el error muestral por $\sqrt{n}$, tenemos:

$$\sqrt{n}(\hat{\beta}_n - \beta_0) = \underbrace{(\frac{1}{n}\sum_{i=1}^{n} x_i x_i^{'})^{-1}}_{\xrightarrow{p} Q_{xx}^{-1}} \underbrace{(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} x_i \varepsilon_i)}_{\xrightarrow{d} N(0, Var(x_i\varepsilon_i))}$$

- El producto $\{x_i\varepsilon_i\}$ es i.i.d y por supuesto de ortogonalidad, $E(x_i\varepsilon_i) = 0$. Entonces,

$$\begin{aligned} Var(x_i\varepsilon_i) &= E[(x_i\varepsilon)(x_i\varepsilon)^{'}] - [E(x_i\varepsilon_i]^2 \\ &= E[(x_i\varepsilon)(x_i\varepsilon)^{'}] \\ &= E[x_i\varepsilon_i^2 x_i^{'}] \\ &= E[\varepsilon_i^2 x_i x_i^{'}] \\ &= \Omega \end{aligned}$$

Normalidad

- Tenemos:

$$\sqrt{n}(\hat{\beta}_n - \beta_0) = \underbrace{(\frac{1}{n}\sum_{i=1}^{n} x_i x_i^{'})^{-1}}_{\xrightarrow{p} Q_{xx}^{-1}} \underbrace{(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} x_i \varepsilon_i)}_{\xrightarrow{d} N(0,\Omega)}$$

- a medida que $n \to \infty$, donde $\Omega = E[\varepsilon_i^2 x_i x_i^{'}]$. Por lo tanto,

$$\sqrt{n}(\hat{\beta}_n - \beta_0) \xrightarrow{d} Q_{xx}^{-1} N(0,\Omega) = N(0, Q_{xx}^{-1}\Omega Q_{xx}^{-1})$$

- a medida que $n \to \infty$.

Normalidad

- Ya que $\sqrt{n}(\hat{\beta}_n - \beta_0) \xrightarrow{d} N(0, V_{\hat{\beta}})$, donde $V_{\hat{\beta}}$ es una matriz semidefinida positiva $K \times K$. Entonces, $\hat{\beta}_n$ es $\sqrt{n}$-distribuido normalmente asintótico y $V_{\hat{\beta}}$ es la varianza asintótica de $\sqrt{n}(\hat{\beta}_n - \beta_0)$, denotado como $Avar\sqrt{n}(\hat{\beta}_n - \beta_0) = V_{\hat{\beta}}$. Por lo tanto:

$$\hat{\beta}_n \overset{a}{\sim} \frac{V_{\hat{\beta}}}{n}$$

- donde $\overset{a}{\sim}$ se lee como "distribuido asintóticamente como".
- Encontrar una estimación de $\hat{V}_{\hat{\beta}}$ de $V_{\hat{\beta}}$. El estimador de $Avar(\hat{\beta}_n)$ es $\hat{V}_n/n$, y podemos escribirlo:

$$\widehat{Avar}(\hat{\beta}_n) = \hat{V}_n/n$$

Normalidad

Teorema: Estimador Consistente de Varianza Asintótica V

Suponer que el estimador consistente de $\Omega(K \times K)$ es $\hat{\Omega}$. Entonces, bajo el supuesto de muestra aleatoria, la matriz de varianza-covarianza asintótica $V_{\hat{\beta}}$ es estimador consistentemente por:

$$\hat{V}_{\hat{\beta}} = S_{xx}^{-1} \hat{\Omega} S_{xx}^{-1}$$

donde S_{xx} es el promedio muestral de $x_i x_i^{'}$.

$$S_{xx} \equiv \frac{1}{n} \sum_{i=1}^{n} x_i x_i^{'} = \frac{1}{n} X^{'} X$$

Normalidad

- Demostración: Ya que $\hat{\Omega}$ es estimador consistente o Ω, sabemos que $\hat{\Omega} \xrightarrow{p} \Omega$. Además:

$$S_{xx} \xrightarrow{p} Q_{xx}$$

- Asimismo,

$$\hat{V}_{\hat{\beta}} \xrightarrow{p} Q_{xx}^{-1} \Omega Q_{xx}^{-1} \equiv V_{\hat{\beta}}$$

Índice

1 Teoría Asintótica
- Convergencia de Secuencias Determinísticas
- Convergencia en Probabilidad
- Ley de los Grandes Números
- Convergencia en Distribución
- Teorema del Límite Central

2 Propiedades Asintóticas de MCO
- Consistencia
- Normalidad
- Consistencia de $\hat{\sigma}^2$
- Estimación Consistente de la matriz de covarianza

Consistencia de $\hat{\sigma}^2$

Teorema: Estimación consistente de la varianza del error

Sea $\hat{\varepsilon}_i \equiv y_i - x_i^{'}\hat{\beta}$ el residuo MCO para la observación i. Bajo el supuesto de linealidad, muestra aleatoria, ortogonalidad, condición de rango:

$$\hat{\sigma}^2 \equiv \frac{1}{n}\sum_{i=1}^{n}\hat{\varepsilon}_i^2 \xrightarrow{p} E(\varepsilon_i^2)$$

siempre que $E(\varepsilon_i^2)$ existe y es finito.

Consistencia de $\hat{\sigma}^2$

- Demostración: tenemos que:

$$\begin{aligned}\hat{\varepsilon}_i &= y_i - x_i^{'}\hat{\beta} \\ &= x_i^{'}\beta_0 + \varepsilon_i - x_i^{'}\hat{\beta} \\ &= \varepsilon_i - x_i^{'}(\hat{\beta} - \beta_0)\end{aligned}$$

- Por lo tanto, los residuos al cuadrado es igual errores al cuadrado más una desviación:

$$\hat{\varepsilon}_i^2 = \varepsilon_i^2 - 2\varepsilon_i x_i^{'}(\hat{\beta}_n - \beta_0) + (\hat{\beta}_n - \beta_0)^{'} x_i x_i^{'}(\hat{\beta}_n - \beta_0)$$

- Tomando el promedio de residuos cuadrados:

$$\sigma_i^2 = \frac{1}{n}\sum_{i=1}^{n}\varepsilon_i^2 - 2\left(\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i x_i^{'}\right)(\hat{\beta}_n - \beta_0) + (\hat{\beta}_n - \beta_0)^{'}\left(\frac{1}{n}\sum_{i=1}^{n}x_i x_i^{'}\right)(\hat{\beta}_n - \beta_0)$$

Consistencia de $\hat{\sigma}^2$

- Por el Teorema Débil de la Ley de los Grandes Números mostramos que:

$$\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i^2 \xrightarrow{p} \sigma_0^2$$

$$\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i x_i^{'} \xrightarrow{p} E(\varepsilon_i x_i^{'}) = 0$$

$$\frac{1}{n}\sum_{i=1}^{n}x_i x_i^{'} \xrightarrow{p} E(x_i x_i^{'}) = Q_{xx}$$

- Por el teorema de consistencia de estimadores MCO se muestra que $\hat{\beta}_n \xrightarrow{p} \beta_0$. Por lo tanto, $\frac{1}{n}\sum_{i=1}^{n}\hat{\varepsilon}_i^2 \xrightarrow{p} E(\varepsilon_i^2)$.

Índice

1. Teoría Asintótica
   - Convergencia de Secuencias Determinísticas
   - Convergencia en Probabilidad
   - Ley de los Grandes Números
   - Convergencia en Distribución
   - Teorema del Límite Central

2. Propiedades Asintóticas de MCO
   - Consistencia
   - Normalidad
   - Consistencia de $\hat{\sigma}^2$
   - Estimación Consistente de la matriz de covarianza

Estimación consistente de la matriz de covarianza

Teorema: Estimación consistente de la matriz de covarianza

Suponer que el coeficiente estimado $\hat{\beta}$ usado para calcular los residuos $\hat{\varepsilon}_i$ para $\hat{V}$ es consistente, y suponer que $\Omega = E(\varepsilon_i^2 x_i x_i^{'})$ existe y es finito. Bajo el supuesto de linealidad, muestra aleatoria:

$$\hat{\Omega} = \frac{1}{n} \sum_{i=1}^{n} \hat{\varepsilon}_i{}^2 x_i x_i^{'} \xrightarrow{p} \Omega$$

donde $\hat{\varepsilon}_i = y_i - x_i^{'} \hat{\beta}$.

Referencias

- **Capítulo 2.1 y 2.3.** Hayashi, F. (2011). Econometrics. Princeton University Press.