Propiedades de Muestras Grandes y Simulación

Microeconomía Cuantitativa

R. Mora

Departmento of Economía
Universidad Carlos III de Madrid

Esquema

Las propiedades de muestras grandes

- las propiedades de muestras grandes describen cómo un estimador $\hat{\theta}$ de un parámetro θ "se comporta" conforme aumenta el tamaño muestral:
    1. si $\hat{\theta}$ es muy diferente del verdadero valor θ conforme $n \to \infty$
    2. cómo varía la distribución de $\hat{\theta}$ conforme $n \to \infty$?
- en consonancia con estas dos cuestiones, hay dos conceptos de "comportamiento":
    1. convergencia en probabilidad: consistencia
    2. convergencia en distribución

Convergencia en probabilidad: consistencia

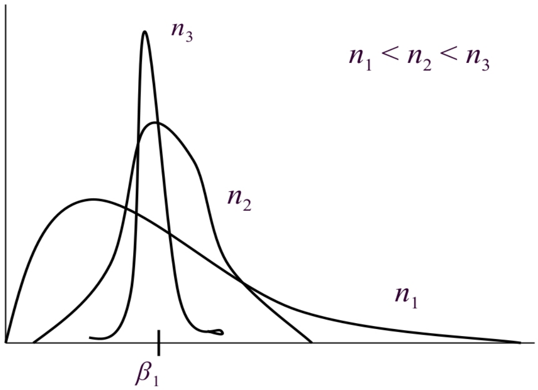
Definición

- Conforme aumenta la muestra, cualquier diferencia, no importa lo pequeña que sea, entre $\hat{\theta}$ y θ será arbitrariamente improbable

- Técnicamente: $\Pr\left(\left|\hat{\theta}-\theta\right|>\varepsilon\right)\to 0$ as $n\to\infty$

- θ es la probabilidad en el límite de $\hat{\theta}$
- $\hat{\theta}$ converge en probabilidad a θ
- $plim(\hat{\theta})=\theta$

Una interpretación gráfica de la consistencia



Source: Wooldrigde (2003)

La ley de grandes números: intro

- en su versión más simple fue formulada por Bernoulli en 1713: le costó 20 años obtener la prueba
- esencialmente nos dice que la media converge en probabilidad a la esperanza
- la LGN es importante porque "garantiza" resultados estables a largo plazo para fenómenos aleatorios

Una ley de grandes números

Para una variable aleatoria y con esperanza μ tal que la media de una muestra de tamaño n es $\overline{y}_n$. Entonces

$$plim(\overline{y}_n) = \mu$$

Example 1

$$plim(\hat{cov}_n(y, x)) = cov(y, x)$$

Example 2

$$plim(\hat{var}_n(x)) = var(x)$$

Propiedades del operador *plim*

Teorema de Mann-Wald

Para cualquier función contínua $g(\cdot)$ y variable aleatoria x:

$$plim(g(x)) = g(plim(x))$$

Example 1 $plim(x+y) = plim(x) + plim(y)$

Example 2 $plim\left(\frac{x}{y}\right) = \frac{plim(x)}{plim(y)}$ if $plim(y) \neq 0$

Ejemplo 1: El Teorema Fundamental de Estadística

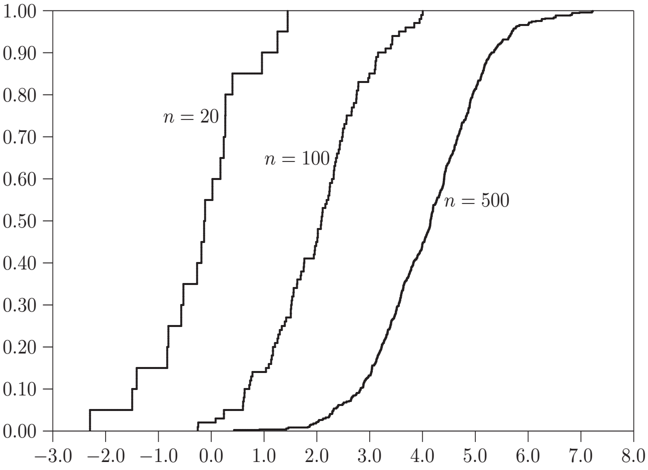
- Supón que X es una variable aleatoria con fda $F(X)$ y que obtenemos una muestra aleatoria de tamaño n donde el elemento típico x_i es una realización independiente de X
- La **distribución empírica** la distribución discreta que asigna el peso $\frac{1}{n}$ a cada observación x_i , $i = 1, ..., n$
- La **EDF** es la función de distribución de la distribución empírica:

$$\hat{F}(x) \equiv \frac{1}{n} \sum_{i=1}^{n} I(x_i \leq x)$$

donde $I(\cdot)$ es la función indicador

El Teorema Fundamental de la Estadística

$$plim\, \hat{F}(x) = F(x)$$



EDFs para tres muestras de 20, 100 y 500 observaciones extraídas de tres distribuciones normales , cada una con varianza 1 y medias 0, 2 y 4,

Ejemplo 2: MCO bajo los supuestos clásicos

Supuestos Gauss-Markov

- A1: Linealidad: $y = \beta + \beta_1 x_1 + \ldots + \beta_k x_k + v$
- A2: Muestreo aleatorio
- A3: Independencia en media condicionada: $E[y\,|\mathbf{x}] = \beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k$
- A4: Invertibilidad de la matriz de varianzas-covarianzas
- A5: Homoscedasticidad: $Var[v\,|\mathbf{x}] = \sigma^2$

Normalidad

- A6: Normalidad: $y\,|\mathbf{x} \sim N(\beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k, \sigma^2)$

Consistencia MCO

Teorema

Bajo Gauss-Markov A1-A4, MCO es consistente

Ejemplo: $wages = \beta_0 + \beta_1 educ + u$ con $cov(educ, u) = 0$

- $\hat{\beta}_1 = \beta_1 + \frac{c\hat{o}v(educ_i, u_i)}{v\hat{a}r(educ_i)}$
- $plim\left(\hat{\beta}_1\right) = plim(\beta_1) + \frac{plim(c\hat{o}v(educ_i, u_i))}{plim(v\hat{a}r(educ_i))} = \beta_1 + \frac{cov(educ, u)}{var(educ)}$
- Puesto que $cov(educ, u) = 0 \Rightarrow plim\left(\hat{\beta}_1\right) = \beta_1$

Un ejemplo de inconsistencia

Modelo Verdadero: $wages = \beta_0 + \beta_1 educ + \beta_2 IQ + v$

- $cov(educ, v) = cov(IQ, v) = 0$
- $cov(educ, IQ) \neq 0$, $\beta_2 \neq 0$

- Ecuación estimada por MCO: $wages = \hat{\gamma}_0 + \hat{\gamma}_1 educ + \hat{u}_{educ}$

$$\hat{\gamma}_1 = \hat{\beta}_1 + \hat{\beta}_2 \frac{\hat{cov}(educ, IQ)}{\hat{var}(educ)} \Rightarrow plim(\hat{\gamma}_1) = \beta_1 + \beta_2 \frac{cov(educ, IQ)}{var(educ)}$$

- $plim(\hat{\gamma}_1) \neq \beta_1$ if
  - la inteligencia es relevante: $\beta_2 \neq 0$
  - la educación está correlacionada con la inteligencia: $cov(educ, IQ) \neq 0$

Normalidad asintótica

Definición

Conforme aumenta la muestra, la distribución de $\hat{\beta}_j$ se acerca arbitrariamente a la normal

- Técnicamente: $\Pr(\hat{\beta}_j \leq z) \rightarrow \Phi(z)$ as $n \rightarrow \infty$

El Teorema Central del Límite: algo de historia

- probablemente una de las leyes matemáticas más interesantes, la prueba es sorprendentemente sencilla (pero reconozco que el resultado me es intuitivamente inexplicable)
- el primer matemático que la postula es el francés de Moivre, en 1733
- Pierre Simon Laplace, otro francés, demostró la versión más simple en 1812
- el ruso Aleksander Liapunov probó el caso general en 1901

Las ratas y el TCL

- ve al sistema de alcantarillas de Madrid
- captura 100 ratas, mide sus colas, estandariza las medidas y calcula la media por la raiz cuadrada de 100
- ahora captura 500 ratas, y vuelve a hacer lo mismo, esta vez multiplicando por la raíz cuadrada de 500
  - si la distribución de la segunda media se parece más a la normal estándar, éste es intuitivamente el resultado
  - si no, en vez de 500 ratas prueba con 1000
- la idea es que si sigues aumentando la muestra de las ratas, SEGURO que te acercarás tanto como quieras a la distribución normal

Las estrellas y el TCL

- mide el brillo de 100 estrellas
- ¡sí, eso es! SEGURO que aumentando el tamaño muestral puedes acercarte a la distribución normal tanto como quieras simplemente calculando la media estandarizada
- lo sorprendente del TCL es que la distribución aleatoria original es irrelevante: el esplendor de las estrellas no tiene nada que ver con el tamaño de la cola de las ratas
- el resultado es el producto de la operación matemática

Un TCL

Para cualquier variable aleatoria y con esperanza μ y varianza σ^2 cuya media muestral es $\overline{y}_n$,

$$n^{1/2}\frac{\overline{y}_n-\mu}{\sigma}\rightarrow N(0,1)\,\text{conforme}\; n\rightarrow\infty$$

Bajo Gauss-Markov A.1-A.5

$$n^{1/2}\frac{\hat{\beta}_j-\beta_j}{\sigma/a_j}\rightarrow N(0,1)\,\text{conforme}\; n\rightarrow\infty\;\text{donde}\; a_j{}^2 = plim\left(\frac{1}{n}\sum_i r\hat{e}s_{ji}{}^2\right)$$

Se puede utilizar junto con LGN: $plim\left(\hat{\sigma}^2\right) = plim\left(\frac{SSR}{n-k-1}\right) = \sigma^2$

Normalidad asintótica de MCO

- conforme aumenta la muestra n, los estimadores MCO—convenientemente reescalados—se aproximan arbitrariamente a una normal

$$n^{1/2}\hat{\beta}_j \approx N\left(\beta_j, \sigma^2 {a_j}^2\right)$$

- cuanto mayor muestra, más precisas serán las estimaciones
- nota importante: esta expresión de la varianza asintótica es correcta solo si el supuesto A5 (homoscedasticidad) es correcto

El test t

- del TCL (y la aplicación de una LGN) se puede demostrar que

$$t = \frac{\hat{\beta}_j - \beta_j}{se\left(\hat{\beta}_j\right)} \rightarrow N\left(0,1\right) \text{ conforme } n \rightarrow \infty$$

- este resultado puede usarse para constrastar si un coeficiente es significativo

Propiedades de muestras pequeñas

- la idea es entender el comportamiento de un estimador cuando el tamaño muestral n está "fijo"

Bajo Gauss-Markov A.1:A.5 Y Normalidad A.6

$$\frac{\hat{\beta}_j - \beta_j}{se(\hat{\beta}_j)} \backsim t_{n-k-1}$$

- el problema es que el supuesto de normalidad es muy restrictivo

Experimentos Monte Carlo: definición y motivación

Definición

- Los experimentos Monte Carlo son un tipo de algoritmos computacionales que utilizan muestreo aleatorio repetido

- los métodos de Monte Carlo se utilizan para simular sistemas matemáticos
- se tienden a utilizar cuando calcular resultados exactos es imposible o muy costoso
- en este curso vamos a utilizarlos para generar datos observacionales y estudiar el comportamiento de las técnicas econométricas cuando las muestras no son infinitas

¿Por qué se llaman Monte Carlo?

- este nombre se usó por físicos nucleares americanos en los años 40
- querían resoveer un problema de difusión de radiación, pero no sabían hacerlo analíticamente, así que decidieron resolverlo utilizando la aleatoriedad
- al proyecto lo llamaron "Monte Carlo" en referencia al casino en Monaco donde el tío de uno de ellos se gastaba el dinero
- desde entonces se ha utilizado en todos los campos de la ciencia
- para los economistas, las técnicas de Monte Carlo son importantes para resolver ciertos problemas matemáticos que surgen en muchos modelos económicos (integración, optimización), en juegos y en estadística aplicada

Monte Carlo en Estadística Aplicada

- las técnicas de Monte Carlo se utilizan frecuentemente en dos contextos en la Estadística Aplicada
  - para comparar y contrastar estadísticas para muestras pequeñas
  - para estudiar cómo el comportamiento en muestras pequeñas se transforma en comportamiento en muestras grandes
- las técnicas de Monte Carlo son útiles
  - frecuentemente sus resultados son más precisos que los resultados asintóticos
  - pero no son tan difíciles de obtener como los tests exactos

Monte Carlo vs. Simulación

- una simulación es una representación ficticia de la realidad mediante ordenador
- un estudio de Monte Carlo es una técnica que puede ser usada para resolver un problema matemático/estadístico
- una simulación Monte Carlo utiliza el muestreo repetido de datos simulados para determinar las propiedades de algún fenómeno

Un sencillo ejercicio de simulación

1. con la ayuda de un ordenador, extrae un valor pseudo-aleatorio de la uniforme [0,1]
2. si el valor es menor o igual a 0.50 designa el resultado del experimento como “cara”
3. si el valor es mayor que 0.50 designa el resultado como “cruz”

- ésta es una simulación de lanzar una moneda al aire

Un sencillo estudio de Monte Carlo

el area de una figura irregular inscrita en el cuadrado unitario

1. estrae dos valores pseudo-aleatorios de la uniforme [0,1]
2. si el punto identificado por esos valores se encuentra dentro de la figura, designa el resultado del experimento como "éxito", si no, "fracaso"
3. repite pasos 1 y 2 muchas veces
4. la proporción de éxitos nos da el área de la figura

- ésto es usar técnicas de Monte Carlo para computar una integral compleja

Un sencillo experimento de Monte Carlo

1. extrae un valor pseudo-aleatorio de la uniforme [0,1]
2. si el valor es menor o igual a 0.50 designa el resultado como "cara", si no, "cruz"
3. repite pasos 1 y 2 muchas veces
4. la proporción de caras es una replicación Monte Carlo del estimador "media" de la probabilidad de que salga cara al lanzar una moneda al aire

Usando Monte Carlo en Estadística Aplicada

- suponemos un modelo estadístico/econométrico y lo simulamos muchas veces
- de cada población simulada extraemos una muestra de tamaño n y computamos un estimador
- observando las propiedades estadísticas del estimador en todas las simulaciones, "estimamos" las propiedades del estimador cuando la muestra está fija en n
- incrementando n y realizando todo de nuevo, "estimamos" cómo el estimador se comporta al aumentar la muestra

El algoritmo básico de un experimento Monte Carlo

Experimento Monte Carlo para una muestra de tamaño fijo N

1. supón valores para las variables exógenas del modelo o extráelas de sus distribuciones
2. extrae una muestra aleatoria de tamaño N para el error
3. calcula las variables endógenas del modelo usando el modelo
4. calcula el estimador
5. replica pasos 1 a 4 R veces
6. examina la distribución empírica de los R valores obtenidos del estimador

Un sencillo experimento Monte Carlo

$log(wages) = 10 + 0.05 * D + u, \; u \sim N(0,1), \; D = 1$ con prob. 0.3

1. extrae N realizacones de D
2. extrae N realizaciones de u
3. computa $log(wages)$
4. MCO $log(wages)$ sobre D y guarda $\hat{\beta}_1^r$
5. replica pasos 1 a 4 R veces
6. examina la distribución empírica de $\hat{\beta}_1^r$

Resumen

- las propiedades de muestras grandes nos dicen cómo se comporta un estimador conforme la muestra se hace arbitrariamente grande
- las propiedades exactas de los estimadores para un tamaño muestral fijo son difíciles de obtener y cuando se obtienen es frecuentemente con supuestos muy restrictivos
- las simulaciones de Monte Carlo son útiles en la Estadística Aplicada. Con ellas podemos estudiar:
    - las propiedades en muestras pequeñas de los estimadores
    - cómo estas propiedades se convierten en propiedades de muestras grandes