Semana 3. Mínimos Cuadrados Ordinarios

Equipo Econometría Avanzada

Universidad de los Andes

26 de agosto de 2022

Contenido

Contexto

Bryan et al. (2021) utilizan información de personas pertenecientes a comunidades evangélicas pentecostales vulnerables de Filipinas.

- Aleatoriamente se seleccionaron 320 comunidades religiosas para ser invitadas a un programa sobre valores cristianos.
- El programa consistió en 15 sesiones semanales.
- El currículo enseñó la “bondad del mundo material” e incentiva el ahorro, el no malgastar dinero, y el no apostar ni beber.

Pregunta de investigación

¿Cuál es el efecto de la asignación al programa de valores cristianos (D) sobre la religiosidad (Y^1) y el salario (Y^2)?

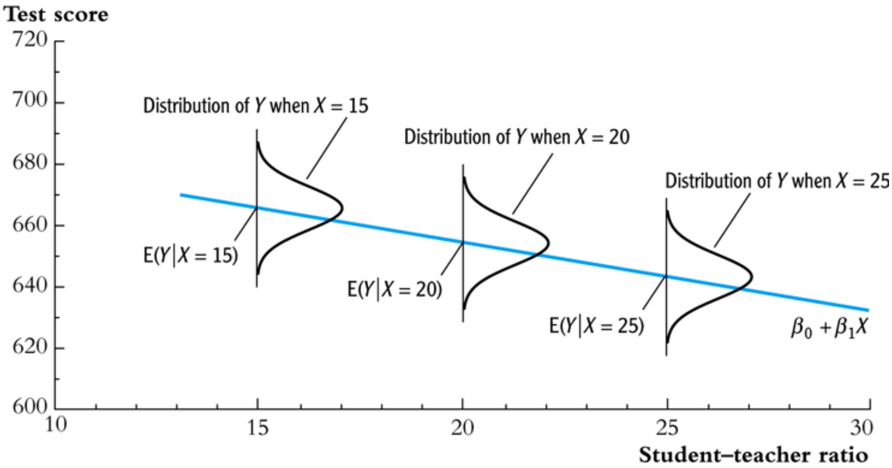
Considere el siguiente modelo de regresión:

$$Y_i^j = \alpha_0^j + \alpha_1^j D_i + u_i^j,$$

con $j = \{1, 2\}$

- ¿Qué contienen los términos u_i^1 y u_i^2?

El modelo de regresión lineal



Mínimos Cuadrados Ordinarios (MCO)

Estimación

Para un modelo generalizado, $Y = X\alpha + U$, el estimador de MCO de α es:

$$\widehat{\alpha}_{MCO} = \arg\min_{\widehat{\alpha}} \widehat{U}' \widehat{U}$$

$$= \arg\min_{\widehat{\alpha}} \sum_{i=1}^{N} \widehat{U}_i^2$$

$$\boxed{\widehat{\alpha}_{MCO} = \left(X'X\right)^{-1} X'Y}$$

MCO

Valores predichos

- **Outcome predicho:**

$$\widehat{Y} = X\widehat{\alpha}$$

- **Término de error predicho:**

$$\widehat{U} = Y - \widehat{Y}$$

Pilas: Modelo $\neq$ Método de estimación

MCO

Propiedades del estimador de MCO

En general, para un modelo

$$Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \epsilon_i$$

El estimador del vector de coeficientes β por MCO, $\widehat{\beta}_{MCO}$, tiene las siguientes propiedades:

- **(Insesgamiento):** Es insesgado siempre que $\mathbb{E}[\epsilon_i|X] = 0$. (Exogeneidad estricta)
- **(Consistencia):** Es consistente siempre que $Cov(x_i, \epsilon_i) = \mathbb{E}[x_i \epsilon_i] = 0$ (No correlación)
- **(Eficiencia):** Es el estimador MELI si:
  - $\mathbb{E}[\epsilon_i|X] = 0$. (Exogeneidad estricta)
  - $Var(\epsilon_i) = \sigma^2 \; \forall i$ (Homocedasticidad)
  - $Cov(\epsilon_i, \epsilon_j) = 0 \; \forall i \neq j$ (No autocorrelación)

- ¿Qué supuesto es más débil entre exogeneidad estricta o no correlación?
- ¿Es razonable pensar que $\widehat{\alpha}^j_{1,MCO}$ es insesgado y consistente?
- ¿Es razonable pensar que es el estimador MELI?

MCO

Estimadores particulares

El estimador de MCO toma formas particulares en función de la cantidad y naturaleza de los regresores. Algunas son:

1. Cuando X es de tamaño $(N \times 1)$ y solo contiene una columna de 1s:
$$Y = \alpha_0 + U$$
$$\widehat{\alpha}_{0,MCO} = \bar{Y}$$
2. Cuando X es de tamaño $(N \times 2)$ y contiene una columna de 1s y una columna con D (dummy):
$$Y = \alpha_0 + \alpha_1 D + U$$
$$\widehat{\alpha}_{1,MCO} = \bar{Y}_1 - \bar{Y}_0$$

MCO

Estimadores particulares

3. Cuando X es de tamaño $(N \times 2)$ y contiene una columna de 1s y una columna con x_1 (continua):

$$Y = \alpha_0 + \alpha_1 x_1 + U$$

$$\widehat{\alpha}_{1,MCO} = \frac{\widehat{Cov}(Y, x_1)}{\widehat{Var}(x_1)}$$

Nota: Bajo el supuesto de que $\mathbb{E}[U|x_1] = 0$, entonces

$$\alpha_1 = \frac{\partial \mathbb{E}[Y|x_1]}{\partial x_1}$$

esto es, es el efecto marginal de incrementar/disminuir x_1 **en el promedio** de Y.

MCO

Regresión múltiple y partialling-out

Cuando X es de tamaño $(N \times k)$:

$$Y = \alpha_0 + \alpha_1 D + x\beta + U \quad (1)$$

¿Cuál es la interpretación de $\widehat{\alpha}_1$ en este contexto?

La manera más adecuada para interpretar este coeficiente se deriva del Teorema de Waugh-Frisch-Lovell, conocido como el método de "partialling-out".

"Partialling-out" nos permite entender qué hacen los controles y por qué son importantes.

MCO

Regresión múltiple y partialling-out

Teorema de Waugh-Frisch-Lovell

Considere el modelo (1)

$$Y = \alpha_0 + \alpha_1 D + x\beta + U \quad (1)$$

junto con los siguientes modelos

$$D = \gamma_0 + x\gamma_1 + \eta \quad (2)$$

$$Y = \mu_0 + x\mu_1 + \xi \quad (3)$$

$$\hat{\xi} = \zeta_0 + \zeta_1\hat{\eta} + \varepsilon \quad (4)$$

Entonces $\widehat{\alpha}_{1,MCO} = \widehat{\zeta}_{1,MCO}$.

Ver simulación

Otros supuestos: matriz de varianza y covarianza (VarCov)

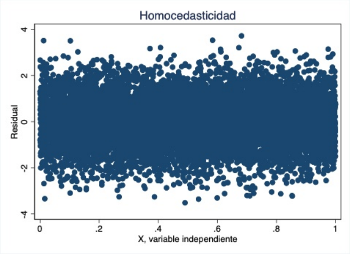
$$\Sigma = \text{Var}(\epsilon|\mathbf{X}) = \mathbb{E}[\epsilon\epsilon'|\mathbf{X}] = \begin{bmatrix} \sigma_{11}^2 & \sigma_{12} & \dots & \sigma_{1n-1} & \sigma_{1n} \\ \sigma_{21} & \sigma_{22}^2 & \dots & \sigma_{2n-1} & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \sigma_{n-11} & \sigma_{n-12} & \dots & \sigma_{n-1n-1}^2 & \sigma_{n-1n} \\ \sigma_{n1} & \sigma_{n2} & \dots & \sigma_{nn-1}^2 & \sigma_{nn}^2 \end{bmatrix}_{N\times N}$$

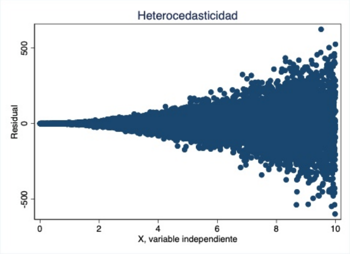
Matriz cuadrada (NxN), contiene las varianzas y covarianzas de los términos de error de N observaciones

- La diagonal principal contiene las varianzas de los errores.
- Fuera de la diagonal están las covarianzas entre los errores de las distintas observaciones.
- ¿Cómo se ve la VarCov cuando se cumple el supuesto de homoscedasticidad y no autocorrelación?
- ¿Y cómo se ve cuando se cumple el supuesto de no autocorrelación pero no el de homoscedasticidad?¿Y cuando se cumple el supuesto de homoscedasticidad pero no el de no autocorrelación?

Inference Estadística

Heterocedasticidad





Heterocedasticidad: $Var(\epsilon_i|X) \neq constante$

- **Problema:** Podemos sobre o subestimar los errores. No hay una regla general, pues depende de la relación entre el error y la variable independiente.
  - Esto invalida la inferencia estadística.
- **Solución:** Usar errores estándar robustos de White (en muestras grandes).
  - Esto valida la inferencia estadística en presencia de heteroscedasticidad. **No la elimina.** MCO sigue siendo ineficiente.

Auto-correlación: $cov(\epsilon_i \epsilon_j | X) \neq 0$ dado que $i \neq j$

- **Problema:** Subestima los errores estándar.
  - Esto invalida la inferencia estadística.
- **Solución:** Errores estándar robustos a correlación a nivel de clúster (se sugieren 50 o más clusters).
  - Esto valida la inferencia estadística en presencia de autocorrelación. **No la elimina.** MCO sigue siendo ineficiente.
- ¿En qué situaciones pueden comúnmente surgir problemas de autocorrelación?

En general, el no corregir la estimación de los errores en presencia de heteroscedasticidad y/o autocorrelación resulta en:

- Estimadores sesgados de la varianza de los estimadores.
- Intervalos de confianza más grandes o pequeños de los reales.

Por ende, **podemos llegar a conclusiones incorrectas**. Entonces, en muchos casos, hace falta usar errores estándar robustos para realizar inferencia estadística correctamente.

- ¿Entonces nunca se deben usar los errores estándar clásicos?

Corrección de errores

1. Heteroscedasticidad:
   - Mínimos Cuadrados Ponderados (MCP) o Mínimos Cuadrados Generalizados (MCG).
     - Funciona en caso de conocer la forma funcional de la heteroscedasticidad.
     - Esto sí corrige el problema de heteroscedasticidad.
   - Errores estándar de White en el resto de situaciones.
2. Autocorrelación:
   - Mínimos Cuadrados Generalizados (MCG) o Mínimos Cuadrados Factibles (MCF).
     - Funciona en caso de conocer la forma funcional de la autocorrelación.
     - Esto sí corrige el problema de autocorrelación.
   - Errores estándar robustos a correlación a nivel de clúster.
   - ¿Cuándo corregir por errores estándar cluster?
     - Si el proceso de muestreo fue clusterizado.
     - La asignación al tratamiento fue clusterizada.

¿Hay alguna otra alternativa? Sí, el cálculo de errores estándar mediante la metodología de *Bootstrap*.

¡Gracias!