

Errores de Medicion

Walter Sosa-Escudero

May 10, 2009

Un caso particular en que los regresores son necesariamente aleatorios es cuando hay **errores de medicion**.

Consideremos el siguiente caso simple

$$C_i = \beta_1 + \beta_2 X_i^* + u_i$$

Por ejemplo, C puede ser consumo y X_i^* el 'ingreso permanente'.

Supongamos que solo observamos una version 'ruidosa' de X_i^* :

$$X_i = X_i^* + \omega_i$$

ω_i es un **error de medicion**.

Por ejemplo, X_i podria ser el ingreso corriente.

Supondremos $E(\omega_i) = 0$, $V(\omega_i) = \sigma_\omega^2$ y ω no correlacionado con X_i^* y u_i .

Reemplazando $X_i^* = X_i - \omega_i$, el modelo puede ser escrito como

$$C_i = \beta_1 + \beta_2 X_i + \nu_i$$

con $\nu = -\beta_2 \omega_i + u_i$

Pero

$$\begin{aligned} Cov(X_i, \nu_i) = E(X_i, \nu_i) &= E[(X_i^* + \omega_i)(-\beta_2\omega_i + u_i)] \\ &= -\beta_2\sigma_\omega^2 \neq 0 \end{aligned}$$

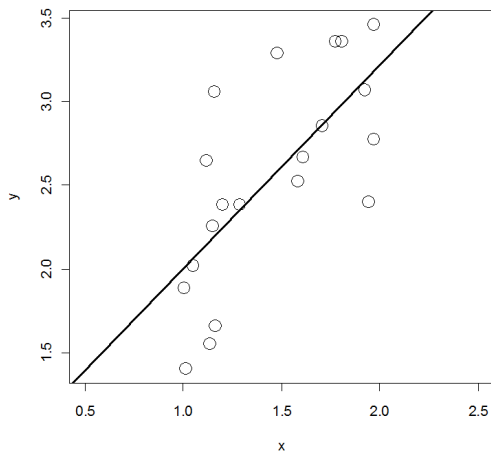
Entonces, el estimador de MCO que regresa Y_i en X_i es **sesgado**, ya que X_i esta correlacionado con el termino de error ν_i .

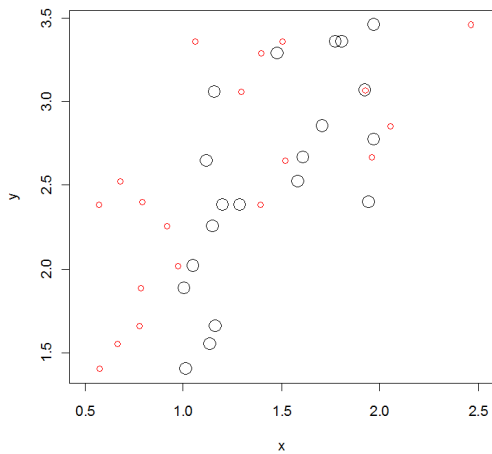
Es posible mostrar que cuando n es relativamente grande:

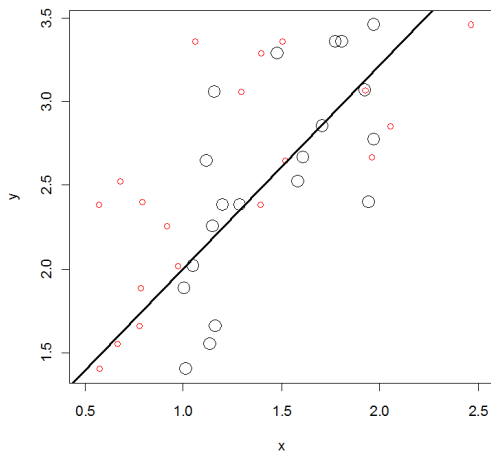
$$\hat{\beta} \simeq \beta \left[\frac{\sigma_{X^*}^2}{\sigma_{\omega}^2 + \sigma_{X^*}^2} \right]$$

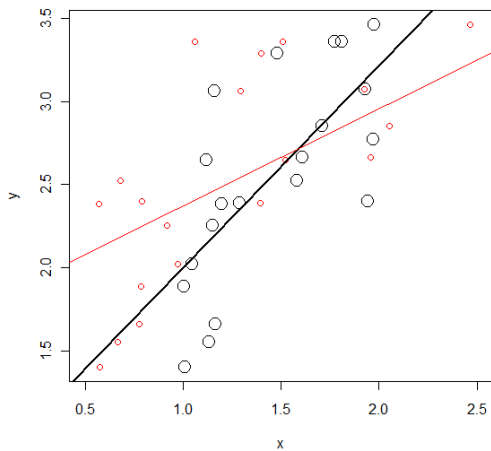
Este es un resultado muy importante

- $\sigma_{X^*}^2 / (\sigma_{\omega}^2 + \sigma_{X^*}^2)$ es un numero entre 0 y 1.
- Entonces, $|\hat{\beta}|$ tiende a ser menor que β : el error de medicion en la variable explicativa implica un **sesgo de atenuacion**.
- Los errors de medicion contraen las estimaciones hacia cero.
- Cuidado con echarle la culpa a los errores de medicion por las no-significatividades!









Errores de medicion en la variable explicada

$$Y^* = \beta_1 + \beta_2 X_i + u_i$$

$$Y_i = Y^* + \omega_i$$

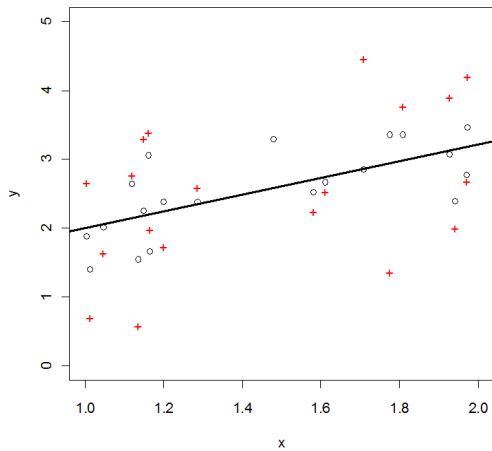
ω_i no correlacionada con u_i , $V(\omega_i) = \sigma_\omega^2$

Reemplazando

$$\begin{aligned} Y_i &= \beta_1 + \beta_2 X_i + \omega_i + u_i \\ &= \beta_1 + \beta_2 X_i + \nu_i \end{aligned}$$

Error de medicion en la variable explicada

- Notar que bajo nuestros supuestos, $\nu_i \equiv \omega_i + u_i$ no esta correlacinada con X_i , entonces, la presencia de errores de mediicion en la variable explicada **no sesga al estimador de MCO**.
- $V(\nu_i) = \sigma_\omega^2 + \sigma^2 > \sigma^2$, la varianza del 'nuevo' termino de error es mayor que la que habria en un modelo sin errores de medicion: los errores de medicion en la variable dependiente hacen que el estimador OLS sea **mas ineficiente**.



Ejemplo: ahorro micro y macro

Modelo de ahorro para personas:

$$s_i = \alpha + \beta y_i^* + u_i$$

donde: s_i = ahorro; y_i^* = ingreso permanente (no observable)

Estimamos:

$$s_i = \alpha + \beta y_i + \nu_i$$

y_i = ingreso corriente

Recordar que la version 'macro' (en base a las cuentas nacionales) tiene un R^2 muy alto. . Consideremos la version 'micro' con datos individuales.

Dependent Variable: Y
 Method: Least Squares
 Date: 05/16/00 Time: 19:48
 Sample: 1 50
 Included observations: 50

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.342769	0.856116	5.072643	0.0000
X	-0.005185	0.011163	-0.464539	0.6444
R-squared	0.004476	Mean dependent var	3.950660	
Adjusted R-squared	-0.016264	S.D. dependent var	1.003267	
S.E. of regression	1.011393	Akaike info criterion	2.899713	
Sum squared resid	49.09998	Schwarz criterion	2.976193	
Log likelihood	-70.49281	F-statistic	0.215796	
Durbin-Watson stat	1.856169	Prob(F-statistic)	0.644362	

El ingreso no es relevante: sesgo por atenuacion.

Que es lo que explica las diferencias entre las versiones micro y macro?

En la versin 'macro' Y_t es el ingreso corriente, una suerte de promedio

$$\begin{aligned}y_t &= \frac{1}{n} \sum_{i=1}^n y_{it} \\&= \frac{1}{n} \sum_{i=1}^n y_{it}^* + \frac{1}{n} \sum_{i=1}^n \omega_{it}^*\end{aligned}$$

Notar que

$$V \left[\frac{1}{n} \sum_{i=1}^n \omega_{it}^* \right] = \frac{1}{n} \sigma_{\omega}^2$$

Entonces, la 'importancia' del error de medicion tiende a desaparecer: el ingreso corriente agregado hace un trabajo relativamente buen en aproximar el ingreso permanente, pero no a nivel individual.