ECONOMETRÍA I

Ayudantía 10

Profesor: *Víctor Macías E.*
Ayudante: *Juan Felipe Ly*
09 de Diciembre del 2021

Comentes

1. En el contexto de un Modelo de Regresión Lineal (MRL), la recta de regresión pasa por los valores promedios de la variable dependiente. Comente.

R: Este se cumple siempre que el modelo de regresión incluya un término constante, es la constante justamente quien cumple el rol de asegurar que el promedio muestral del valor estimado de Y se iguale con el promedio muestral de los valores observados de Y . Si el modelo incluye término constante por condición de primer orden se cumple que $\sum u_i = 0$, y que el valor estimado de la constante (β) es $\hat{\beta}_1 = \overline{Y} - \hat{\beta}_2 \overline{X}$, de lo que deduce inmediatamente que la recta de regresión pasa por los valores promedio: $\overline{Y} = \hat{\beta}_1 + \hat{\beta}_2 \overline{X}$. Por lo tanto, esta afirmación es verdadera sólo si estamos hablando de un modelo que incluye término constante.

2. La estimación por variables instrumentales siempre es mejor en términos de consistencia y sesgo que la estimación por mínimos cuadrados ordinarios. Comente.

R: Falso. Si existe problema de endogeneidad y se tiene un instrumento exógeno y relevante el estimador de VI es consistente mientras que MCO es sesgado e inconsistente. Pero si se tiene el problema de endogeneidad y se utiliza un instrumento débil (que no sea exógeno o que sea irrelevante) la estimación VI puede ser más sesgada e inconsistente que el estimador MCO.

3. La significancia global del modelo depende de la significancia individual de las variables independientes incluidas.

R: Falso. Es posible que, aun cuando los coeficientes estimados no sean individualmente significativos, sí lo sean en conjunto. Lo anterior podría darse en el caso que la covarianza de los regresores sea capaz de "explicar" la variabilidad de la variable dependiente. Como existe un alto grado de correlación entre las variables independientes del modelo (es decir, un alto grado de multicolinealidad), entonces, al menos una de estas variables tiene una influencia significativa sobre la variable dependiente, pero no se puede establecer cual. Por otro lado, al existir un alto grado de multicolinedlidad, es muy probable que los coeficientes, a nivel individual, no sean estadísticamente significativos. Recuerde que:

$$var(\hat{\beta}) = \hat{\sigma}^2(X'X)^{-1}$$

Si existe un alto grado de colinealidad entre las variables independientes del modelo, el determinante de $(X'X)$ tiende a cero, por lo que se obtienen varianzas "gigantes".

4. La omisión de una variable relevante siempre sesga la estimación MCO de β. Comente.

R: Falso, la omisión de una variable relevante siempre produce sesgo en los parámetros a menos que la correlación entre la variable omitida y las explicativas incluidas sea cero, el signo del sesgo depende de dos cosas: la correlación entre la variable omitida y las variables explicativas incluidas y el valor del parámetro que tendría asociado la variable omitida(valor poblacional). Para un modelo sencillo con una variable explicativa (x_1) y una variable omitida (x_2) se mostró en clases que el sesgo es: $\frac{cov(x_1,x_2)}{V(x_1)}\beta_2$.

5. Siempre es posible obtener los estimadores de MCO en el modelo múltiple utilizando la expresión $\hat{\beta} = (X'X)^{-1}X'Y$.

R: Si la matriz $(X'X)$ no es invertible, no será posible computar la fórmula señalada. Ahora bien, para que la matriz $(X'X)$ sea invertible, es necesario que el rango de la matriz X sea completo, es decir, que sea igual al número de regresores incluidos en el modelo múltiple $(rango(X) = k)$.

6. La significancia estadística de una variable x_i está por completo determinada por la magnitud de $\hat{\beta}_i$, mientras que la significancia económica de esa variable está relacionada con la magnitud (y signo) de $t_{\hat{\beta}_i}$.

R: Es al revés, la significancia estadística de una variable x_i está por completo determinada por la magnitud de $t_{\hat{\beta}_i}$, mientras que la significancia económica de esa variable está relacionada con la magnitud (y signo) de $\hat{\beta}_i$.

7. Un síntoma de la existencia de multicolinealidad es la presencia de un R^2 ajustado alto.

R: El síntoma corresponde no simplemente a la observación de un R^2 alto, sino que además acompañado por estadísticos t que nos hacen concluir que los parámetros no son significativos (t bajos). Esto porque en presencia de multicolinealidades, las varianzas de los estimadores son altas, pues $(X'X)$ es cercano a cero. Lo que resulta en estadísticos t pequeños.

8. El estimador de White me permite solucionar el problema de ineficiencia en el estimador MCO cuando los errores son Heterocedásticos. Comente.

R: Falso, el estimador de White sólo me permite estimar en forma consistente la matriz de varianzas y covarianzas de los parámetros en presencia de heterocedasticidad estimados por MCO. El estimador MCO siempre será ineficiente;el estimador eficiente en este contexto es el de Mínimos Cuadrados Generalizados. Por lo tanto, White me permite estimar consistentemente la matriz de varianzas y covarianzas del estimador MCO (que es ineficiente) para realizar la inferencia en forma correcta, utilizando esta matriz.

Problemas

1. Considere el modelo:

$$y_i = \beta_0 + \beta_1 x_i + u_i \qquad (1)$$

donde $u_i \sim N(0, \sigma_u^2)$. Se sabe que la variable x_i está medida con error, de modo que sólo se observa:

$$x_i{}^* = x_i + v_i$$

donde $v_i \sim N(0, \sigma_v^2)$ y es independiente de u_i, de x_i y de y_i.

a) Reescriba el modelo en la ecuación (1) en función sólo de las variables observables y de los términos de error.

R:

$$\begin{aligned}
y_i &= \beta_0 + \beta_1 x_i + u_i \\
y_i &= \beta_0 + \beta_1 (x_i{}^* - v_i) + u_i \\
y_i &= \beta_0 + \beta_1 x_i{}^* - \beta_1 v_i + u_i \\
y_i &= \beta_0 + \beta_1 x_i{}^* - \beta_1 v_i + u_i \\
y_i &= \beta_0 + \beta_1 x_i{}^* + (u_i - \beta_1 v_i) \\
y_i &= \beta_0 + \beta_1 x_i{}^* + \varepsilon_i
\end{aligned}$$

b) Muestre que en el modelo en (a) el error está correlacionado con la variable observable $x_i{}^*$. Explique las implicancias de este problema.

R:

$$\begin{aligned}
Cov(\varepsilon_i, x_i{}^*) &= Cov(u_i - \beta_1 v_i \ , \ x_i + v_i) \\
Cov(\varepsilon_i, x_i{}^*) &= Cov(u_i, x_i) + Cov(u_i, v_i) - \beta_1 \ Cov(v_i, x_i) - \beta_1 \ Cov(v_i, v_i) \\
Cov(\varepsilon_i, x_i{}^*) &= 0 + 0 - \beta_1 \cdot 0 - \beta_1 \cdot \sigma_v{}^2 \\
Cov(\varepsilon_i, x_i{}^*) &= -\beta_1 \cdot \sigma_v{}^2
\end{aligned}$$

Se observa que el error ϵ_i está correlacionado con la variable observable $x_i{}^*$. Esto implica que la estimación por MCO generará estimadores sesgados de los parámetros β.

c) Explique cómo se podría solucionar el problema en (b).

R: Una forma de correción es la utilización de variables instrumentales. La idea es utilizar un instrumento en vez de la variable medida con error. Para ello, se requiere que el instrumento no esté correlacionado con el término de error y que el instrumento sí esté correlacionado con la variable explicativa para la cual actúa como instrumento (la variable medida con error).

d) En el modelo en la ecuación (1) considere ahora que sólo la variable dependiente y_i está medida con error $w_i \sim N(0, \sigma_w^2)$, el que es independiente de x_i y de u_i. ¿Cuáles son las implicancias de este problema de medición? Explique y justifique algebraicamente.

R: Si sólo la variable dependiente está medida con error, se tiene que sólo se observa:

$$y_i{}^* = y_i + w_i$$

De este modo, el modelo observable es

$$y_i{}^* = \beta_0 + \beta_1\ x_i + u_i$$

$$y_i + w_i = \beta_0 + \beta_1\ x_i + u_i$$

$$y_i = \beta_0 + \beta_1\ x_i + u_i - w_i$$

$$y_i = \beta_0 + \beta_1\ x_i + (u_i - w_i)$$

$$y_i = \beta_0 + \beta_1\ x_i + \epsilon_i$$

Se tendrá entonces que la covarianza entre el término de error y las variables explicativas es igual a cero:

$$Cov(\epsilon_i, x_i) = Cov(u_i - w_i,\ x_i)$$

$$Cov(\epsilon_i, x_i) = Cov(u_i, x_i) + Cov(w_i, x_i)$$

$$Cov(\epsilon_i, x_i) = 0$$

Por lo tanto, no habrá problema de sesgo en la estimación de los parámetros.