Heteroscdasticidad y MCG

Walter Sosa-Escudero

May 16, 2020

Modelo lineal clasico:

1. Linealidad: $Y = X\beta + u$.
2. Exogeneidad: $E(u) = 0$
3. No Multicolinealidad: $\rho(X) = K$.
4. No heteroscedasticidad ni correlacion serial: $V(u) = \sigma^2 I_n$.

Teorema de Gauss/Markov $\hat{\beta} = (X'X)^{-1}X'Y$ es el mejor estimador lineal insesgado.

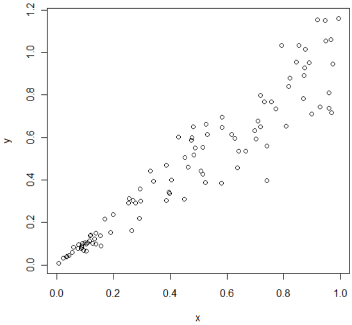
$\hat{V}(\hat{\beta}) = S^2(X'X)^{-1}$ es insesgado para $V(\hat{\beta}) = \sigma^2(X'X)^{-1}$.

El problema

Heterocedasticidad?

- $\hat{\beta}$ lineal e insesgado (por que?).
- Gauss Markov?
- $\hat{V}(\hat{\beta}) = S^2(X'X)^{-1}$ sesgado. Invalida 't' y 'F'.

Intuicion:



- Heterocedasticidad no altera la posicion ‘central’ de la linea de MCO (insesgadez).
- MCO pondera a todas las observaciones por igual. Mas razonable prestar mas atencion a aquellas en donde la varianza es mas chica.

Que hacer con la heterocedasticidad?.

1. Antes de abandonar MCO: tests de heterocedasticidad.
2. *Estrategia 1:* Minimos cuadrados ponderados. Estimador insesgado y eficiente para β bajo heterocedasticidad. Estima insesgadamente la varianza.
3. *Estrategia 2:* Varianza robusta: Mantener MCO (insesgado), reemplazar la varianza por una 'correcta' bajo heterocedasticidad.

Test de Breusch-Pagan/Godfrey/Koenker

Chequea si ciertas variables causan heterocedasticidad.

$$Y = X\beta + u,\ u_i \sim \text{N, con } E(u) = 0 \text{ y}$$

$$V(u_i) = h(\alpha_1 + \alpha_2 Z_{2i} + \alpha_3 Z_{3i} + \ldots + \alpha_p Z_{pi})$$

$h(\ )$ es *cualquier* funcion positiva con dos derivadas.

$\alpha_2 = \ldots = \alpha_p = 0$, $V(u_i) = h(\alpha_1)$, constante!!

Homocedasticidad $\Longleftrightarrow H_0 : \alpha_2 = \alpha_3 = \ldots = \alpha_p = 0$, y
Heterocedasticidad: $H_A : \alpha_2 \neq 0 \vee \alpha_3 \neq 0 \vee \ldots \vee \alpha_p \neq 0$.

El test:

1. Estimar por MCO y guardar residuos al cuadrado, e_i^2.
2. Regresar e_i^2 en las variables Z_{ik}, $k = 2, \dots, p$ obtener (*SCE*). El estadistico es:

$$\frac{1}{2} SCE \sim \chi^2(p-1)$$

bajo H_0, asintoticamente. Rechazar si demasiado grande.

- Intuicion: modelo auxiliar modela la varianza del error. Si su R^2 es alto: se puede explicar la varianza, ergo no es constante.
- Aceptar la H_0no implica que homocedasticidad. Por que?
- Muestras grandes.
- Koenker (1980) propone usar nR_A^2, que ademas es valido si los errores son no-normales.

Test de White

H_0 : homocedasticidad, H_A : heterocedasticidad de alguna forma.

Caso simple, con $K = 3$:

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i \qquad 1, \ldots, n$$

Pasos para implementar el test:

1. Estimar por MCO, guardar los residuos al cuadrado e^2.
2. Regresar e^2 en todas las varaibles, sus cuadrados y todos los posibles productos cruzados no redundantes. En nuestro caso, regresar e^2 en $1, X_2, X_3, X_2^2, X_3^2, X_2 X_3$, obtener el R^2 de esta regresion auxiliar.
3. Bajo H_0, $nR^2 \sim \chi^2(p)$. $p =$ numero de variables explicativas en el modelo auxiliar, menos una.
4. Rechzar H_0 si nR^2 es demasiado grande.

- Intuicion: misma que BPGK
- Hipotesis nula mas general que el de BPGK.
- Informativo si no se rechaza la hipotesis nula (no heterocedasticidad).
- Si rechaza H_0 : hay heterocedasticidad. No tenemos informacion de que la causa.

Minimos cuadrados ponderados

Caso de dos variables

$$Y_i = \beta_1 + \beta_2 X_i + u_i$$

Bajo heteroscedasticidad:

$$V(u_i) = \sigma_i^2, \quad i = 1, \ldots, n$$

Valen todos los otros supuestos clasicos .

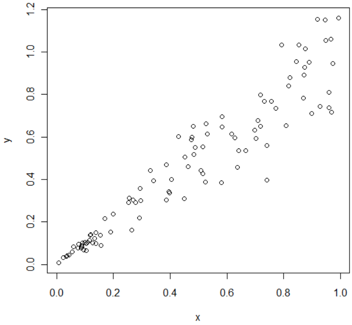
Dividamos cada observacion por σ_i:

$$
\begin{aligned}
\frac{Y_i}{\sigma_i} &= \beta_1 \frac{1}{\sigma_i} + \beta_2 \frac{X_i}{\sigma_i} + \frac{u_i}{\sigma_i} \\
Y_i^* &= \beta_1 X_{1i}^* + \beta_2 X_{2i}^* + u_i^*
\end{aligned}
$$

Notar que $V(u_i^*) = V(u_i/\sigma_i) = 1$: los residuos del modelo transformado son *homocedasticos*

Minimos cuadrados ponderados: el MELI es simplemente el estimador de MCO usando las variables transformadas.

Intuicion:



MCO pondera a todas las observaciones por igual. MCP pondera por la inversa del desvio estandar (menor varianza).

Un modelo implementable

Problema: σ_i^2 desconocido. $Y_i = \beta_1 + \beta_2 X_i + u_i$. Supongamos:

$$V(u_i) = \sigma^2 X_i^2$$

X es la causa de la heterocedasticidad. Dividamos todas las observaciones por X_i

$$\begin{aligned} \frac{Y_i}{X_i} &= \beta_1 \frac{1}{X_i} + \beta_2 + \frac{u_i}{X_i} \\ Y_i^* &= \beta_1 X_{0i}^* + \beta_2 + u_i^* \end{aligned}$$

Notar que $E(u_i^*) = E(u_i/X_i)^2 = \sigma^2 \frac{X_i^2}{X_i^2} = \sigma^2$

Los errores del modelo transformado son homocedaticos: MCO en el modelo transformado!

- No hace falta conocer σ^2. Hemos dividido solo por la parte del error estandar que varia con las obervaciones X_i.
- Cuidado con las interpretaciones. El intercepto de este modelo transformado es la *pendiente* del original y la pendiente del transformado es el intercepto del original.

Otro caso: b) $V(u) = \sigma^2 X_i$

$$
\begin{aligned}
\frac{Y_i}{\sqrt{X_i}} &= \beta_1 \frac{1}{\sqrt{X_i}} + \beta_2 \sqrt{X_i} + \frac{u_i}{\sqrt{X_i}} \\
Y_i^* &= \beta_1 X_{0i}^* + \beta_2 X_{1i}^* + u_i^*
\end{aligned}
$$

Este modelo transformado no tiene intercepto, y el coeficiente de la primera variable explicativa se corresponde con el intercepto y la segudna con la pendiente del modelo original.

Varianza robusta

Problema con MCP: es muy dificil encontrar una forma exacta de heterocedasticidad.
Estrategia alternativa: retener MCO (insesgado pero no eficiente) y buscar un estimador valido para la varianza.

Bajo heterocedasticidad, la varianza de $\hat{\beta}_{MCO}$ es:

$$V(\hat{\beta}_{OLS}) = (X'X)^{-1}X'\Omega X(X'X)^{-1}$$

$\Omega = \text{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2)$.

White (1980): un estimador valido para $X'\Omega X$ es $X'DX$, $D = diag(e_1^2, e_2^2, \ldots, e_n^2)$, e_i's son los residuos de MCO.

Entonces, un estimador *consistente bajo heteroscedasticidad* para la varianza es:

$$\hat{V}(\hat{\beta}_{MCO})_{HC} = (X'X)^{-1}X'DX(X'X)^{-1}$$

Estrategia: usar MCO pero reemplazar $S^2(X'X)^{-1}$ por el estimador de White. Esta estrategia no es eficiente, pero no requiere supuestos acerca de la estructura de la heterocedasticidad.

Resumen

Resumen

- Heterocedasicidad: MCO ineficiente, invalida el estimador estandar de la varianza, y los procesos inferenciales estandar (t tests, F tests, etc.).
- MCP es eficiente e insesgado, pero su implementacion depende de conocer la estructura de varianza.
- En la practica es mas comun mantener MCO y reemplazar el estimador de la varianza por el de White, que no requiere supuestos.

Minimos cuadrados generalizados

En general, valen todos los supuestos clasicos valen, pero ahora

- $V(u) = \sigma^2 \Omega$ en donde Ω es cualquier matriz simetrica $n \times n$ y positiva definida.

Permiteo heterocedasticidad y/o correlacion serial, de cualquier forma.

Plan

1. Consecuencias de relajar $V(u) = \sigma^2 I_n$.
2. Estimador optimo (MCG).

Consecuencias de relajar $V(u|X) = \sigma^2 I_n$

- $\hat{\beta}$ lineal e insesgado, Gauss Markov no vale.
- $V(\hat{\beta})$ es ahora $\sigma^2(X'X)^{-1}\Omega(X'X)^{-1}$.
- $S^2(X'X)^{-1}$ es sesgado para $V(\hat{\beta})$.
- Tests t ni F funcionan.

Entonces, ignorar heterocedasticidad o correlacion serial (usar $\hat{\beta}$ y $\hat{V}(\hat{\beta}|X) = S^2(X'X)^{-1}$), preserva insesadez, inferencia estandar deja de ser valida. Eficiencia?

Resultado simple: si Ω es $n \times n$ simetrica y pd, entonces existe una matriz $n \times n$, C, no singular tal que:

$$\Omega^{-1} = C'C$$

Que significa esto, intuitivamente?

Consideremos el siguiente modelo transformado

$$Y^* = X^*\beta + u^*$$

con $Y^* = CY$, $X^* = CX$ y $u^* = Cu$.

Es facil verificar:

1. $Y^* = X^*\beta + u^*$, el modelo transformado es trivialmente lineal.
2. $E(u^*) = CE(u) = 0$
3. $\rho(X^*) = \rho(CX) = K$,
4. $V(u^*) =$

$$
\begin{aligned}
V(Cu|X) = E(CuuC'|X) &= CE(uu'|X)C' \\
&= C\sigma^2\Omega C' \\
&= \sigma^2 C[\Omega^{-1}]^{-1}C' \\
&= \sigma^2 C\left[(C'C)^{-1}\right]^{-1} C \\
&= \sigma^2 I_n
\end{aligned}
$$

Entonces...?

Vale el Teorema de Gauss-Markov Theorem *para el modelo transformado*, entonces, el MELI es:

$$\hat{\beta}_{gls} = (X^{*\prime}X^*)^{-1}X^{*\prime}Y^*$$

Este es el estimador de minimos cuadrados generalizado

- MCG es un estimador MCO para un modelo transformado.
- MELI bajo heterocedasticidad o correlacion serial.
- Ahoa es claro en que sentido $\hat{\beta}$ es ineficiente.
- Importante: las propiedades estadisticas dependen de la estructura subyacente (no son propiedades del estimador *per-se*).

Notar que

$$\begin{aligned}\hat{\beta}_{gls} &= (X^{**}X^{*})^{-1}X^{**}Y^{*} \\ &= (X'C'CX)^{-1}X'C'CY \\ &= (X'\Omega^{-1}X)X'\Omega^{-1}Y\end{aligned}$$

- Cuando $\Omega = I_n$, $\hat{\beta}_{gls} = \hat{\beta}$.
- La implementacion practica de $\hat{\beta}_{gls}$ require conocer Ω (si bien no σ^2.)
- Es facil chequear que $V(\hat{\beta}_{gls}) = \sigma^2(X^{*\prime}X^{*})^{-1}$.
- Minimos cuadrados ponderados: $\Omega = \text{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2)$, entonces, $C = \text{diag}(\sigma_1^{-1}, \sigma_2^{-1}, \ldots, \sigma_n^{-1})$

MGC Factible

Supongamos que existe un estimador para Ω, que llamaremos $\hat{\Omega}$. Entonces, reemplazando Ω por $\hat{\Omega}$:

$$\hat{\beta}_{fgls} = (X'\hat{\Omega}^{-1}X)X'\hat{\Omega}^{-1}Y$$

Este es el estimador de MCG factible.

Es lineal e insesgado? Eficiente?