Econometría

Tema 5: Heterocedasticidad

2024-2025

1 Concepto, causas y consecuencias **2**
1.1 Concepto 2
1.2 Causas 4
1.3 Consecuencias 7

2 Procedimientos de Detección **8**
2.1 Métodos gráficos 9
2.2 Contrastes 11

3 Estimación bajo heterocedasticidad **20**

1.1. Concepto

$$Y = X\beta + u$$

Hasta ahora suponíamos:

1. $\mathbb{E}[u] = 0$.
2. $\text{var}[u_i] = \sigma^2, \forall i \in 1, 2, \ldots, n$.
3. $\text{Cov}[u_i u_j] = 0, \forall i, j \in 1, 2, \ldots, n$ con $i \neq j$.

Con lo que teníamos:

$$\text{var}[u] = \mathbb{E}[uu^t] = \sigma^2 I_n$$

1.1. CONCEPTO

$$Y = X\beta + u$$

Hasta ahora suponíamos:

1. $\mathbb{E}[u] = 0$.
2. $\text{var}[u_i] = \sigma^2, \forall i \in 1, 2, \ldots, n$.
3. $\text{Cov}[u_i u_j] = 0, \forall i, j \in 1, 2, \ldots, n$ con $i \neq j$.

Con lo que teníamos:

$$\text{var}[u] = \mathbb{E}[uu^t] = \sigma^2 I_n$$

1.1. CONCEPTO

$$Y = X\beta + u$$

Hasta ahora suponíamos:

1. $\mathbb{E}[u] = 0$.
2. $\text{var}[u_i] = \sigma^2, \forall i \in 1, 2, \ldots, n$.
3. $\text{Cov}[u_i u_j] = 0, \forall i, j \in 1, 2, \ldots, n$ con $i \neq j$.

Con lo que teníamos:

$$\text{var}[u] = \mathbb{E}[uu^t] = \sigma^2 I_n$$

¿Y si la matriz de varianzas-covarianzas del vector de perturbaciones pudiera adoptar cualquier forma?

$$\text{VAR}[u] = \sigma^2 \Omega$$

donde Ω es una matriz simétrica y definida positiva.

1.Concepto

Se dice que un modelo de regresión presenta **heteroscedasticidad** cuando la varianza del término de perturbación no permanece constante a lo largo del tiempo.

$$\sigma^2\Omega = \begin{pmatrix} \sigma_1^2 & 0 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & 0 & \cdots & 0 \\ & & \ddots & & \\ 0 & 0 & \cdots & 0 & \sigma_n^2 \end{pmatrix}$$

1.2.Causas de la heteroscedasticidad

- En datos de sección cruzada, donde la escala de la variable dependiente y el poder explicativo de la tendencia del modelo varia a lo largo de las observaciones.
- Cuando en el modelo se trabaja con datos agrupados, es decir, que las observaciones pudieran agruparse en m categorías, cada una de ellas con n_j observaciones, $j = 1, 2, \ldots, m$. Una vez realizada la agrupación, en lugar de trabajar con los datos originales, los cuales son homoscedásticos, utilizamos las medias aritméticas de cada una de las categorías, las cuales resultan ser heteroscedásticas.
- Omisión de una variable relevante en la expresión del modelo de regresión. En esta situación es lógico pensar que el comportamiento de la perturbación dependa de dicha variable omitida, provocando que su varianza sea variable.

1.2. Causas de la heteroscedasticidad: ejemplo

Supongamos que se quiere estudiar el consumo familiar en función de la renta familiar mediante el modelo de regresión

$$C_i = \alpha + \beta R_i + u_i \text{ con } i = 1, 2, \ldots, n$$

donde se verifica que

$$\mathbb{E}[uu^t] = \begin{pmatrix} \sigma_1^2 & 0 & \ldots & 0 \\ 0 & \sigma_2^2 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & \sigma_n^2 \end{pmatrix}$$

El hecho de que la varianza del término de perturbación sea distinta nos da la idea de que la dispersión que existe en los valores del consumo familiar entre los distintos hogares es distinta.

1.2. Causas de la heteroscedasticidad: ejemplo

Supongamos que en lugar de considerar el modelo

$$Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + u_t,$$

se especifica este otro

$$Y_t = \alpha_1 + \alpha_2 X_{2t} + v_t.$$

1.2.Causas de la heteroscedasticidad: ejemplo

Supongamos que en lugar de considerar el modelo

$$Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + u_t,$$

se especifica este otro

$$Y_t = \alpha_1 + \alpha_2 X_{2t} + v_t.$$

Si la variable omitida es realmente relevante, la perturbación aleatoria v_t dependerá de X_{3t} por lo que se puede establecer que $v_t = u_t + \beta_3 X_{3t}$. Y en tal caso:

$$\begin{aligned} Var(v_t) &= \mathbb{E}[v_t^2] = \mathbb{E}[u_t^2 + \beta_3^2 X_{3t}^2 + 2\beta_3 X_{3t} u_t] \\ &= \mathbb{E}[u_t^2] + \beta_3^2 X_{3t}^2 + 2\beta_3 X_{3t} \mathbb{E}[u_t] = \sigma^2 + \beta_3^2 X_{3t}^2 \neq \sigma^2. \end{aligned}$$

Como es evidente, la varianza de la perturbación aleatoria no es constante sino que varía con cada observación.

1.3. Consecuencias de la heteroscedasticidad

Puesto que en el método de estimación por MCO no influye la matriz de varianzas-covarianzas de la perturbación aleatoria es claro que el estimador por MCO será igualmente:

$$\widehat{\beta}_{MCO} = \left(X^t X\right)^{-1} X^t y.$$

$\widehat{\beta}_{MCO}$ es lineal e insesgado. Sin embargo, ya no se tiene asegurado que la varianza sea mínima:

$$\begin{aligned} Var\left(\widehat{\beta}_{MCO}\right) &= \left(X^t X\right)^{-1} X^t \mathbb{E}[u \cdot u^t] X \left(X^t X\right)^{-1} \\ &= \sigma^2 \left(X^t X\right)^{-1} X^t \Omega X \left(X^t X\right)^{-1}, \end{aligned}$$

distinta a la del modelo con perturbaciones esféricas: $\sigma^2 \left(X^t X\right)^{-1}$.

Por tanto, la consecuencia de la presencia de heteroscedasticidad en un modelo lineal es que los estimadores obtenidos, aunque serán lineales e insesgados, no serán óptimos.

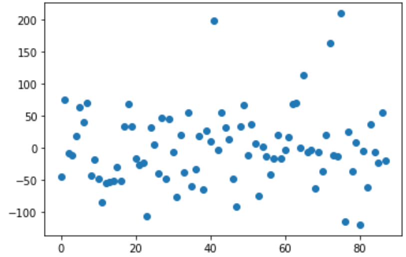
2. Procedimientos de Detección

Para detectar la heteroscedasticidad en un modelo lineal múltiple disponemos de distintos procedimientos.

- **Métodos gráficos** a partir de los cuales intentaremos intuir cuáles son las variables que provocan la existencia de heteroscedasticidad en el modelo. Concretamente, estudiaremos los gráficos de los residuos y de dispersión.
- **Métodos analíticos** para determinar la presencia de heteroscedasticidad en el modelo: Tests de Glesjer, Goldfeld-Quandt, Breusch-Pagan y White. Los dos primeros se deben usar cuando la muestra es pequeña y una variable es la causa de la heteroscedasticidad, mientras que los otros dos cuando la muestra es grande y no se sabe la o las variables que provocan el problema. Además, la hipótesis nula de todos estos contrastes es siempre que el modelo es homocedástico.

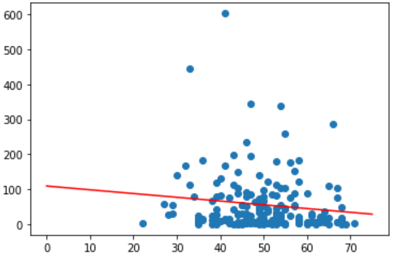
Métodos Gráficos

- Gráfico de los residuos: es un gráfico de dispersión de los residuos o residuos al cuadrado, e_t ó e_t^2, frente a t. Si en dichos gráficos observamos grupos de observaciones con distinta varianza, podemos pensar en la presencia de heteroscedasticidad.



Métodos Gráficos

- Gráficos de dispersión: consiste en el diagrama de dispersión de los residuos o residuos al cuadrado, e_t ó e_t^2, frente a la variable independiente que sospechamos que puede causar la heteroscedasticidad. Si la variabilidad de los residuos aumenta o disminuye conforme aumenta el valor de la variable independiente, entonces podemos pensar que la varianza de la perturbación aleatoria depende de dicha variable y, por tanto, habría presencia de heteroscedasticidad.



CONTRASTE DE GOLDFELD QUANDT

Para muestras pequeñas, y si sospechamos que σ_t^2 está relacionada positivamente con la variable X_i:

1. Ordenar la información muestral, según los valores crecientes de este regresor, de menor a mayor.
2. Omitir los m valores muestrales centrales, (por ejemplo $m = \frac{n}{3}$).
3. Ajustar, mediante MCO, cada uno de los subgrupos obtenidos en el paso 2, una vez eliminados los valores centrales. Cada uno de estos subgrupos estarán formados por $\frac{n-m}{2}$ observaciones. La estimación del modelo del primer subgrupo nos permitirá obtener la SCR_1 y la del subgrupo 2, la SCR_2.
4. Bajo ausencia de heteroscedasticidad ($\sigma_1^2 = \sigma_2^2 = \sigma^2$):

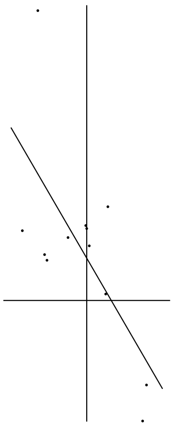
$$F_{exp} = \frac{SCR_2}{SCR_1} > F_{\frac{n-m}{2}-k, \frac{n-m}{2}-k, 1-\alpha}$$

se rechaza la hipótesis de homocedasticidad (hay heteroscedasticidad)

```
sms.het_goldfeldquandt(mco.resid, mco.model.exog)
```

Contraste de GQ

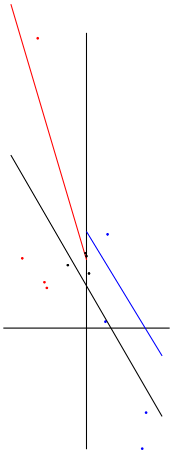
y	x
4.608895	-4.246639
19.1437	-3.225103
3.028715	-2.780972
2.654891	-2.622295
4.155246	-1.234524
4.948342	-0.06467424
4.750159	-0.0002376094
3.608508	0.1671859
0.4248052	1.252264
6.1918	1.394248
-7.961979	3.686417
-5.582933	3.942198



MCO: $y = 2.7941 - 1.7259x$

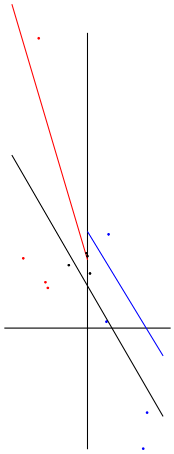
Ejemplo

y	x
4.608895	-4.246639
19.1437	-3.225103
3.028715	-2.780972
2.654891	-2.622295
4.155246	-1.234524
4.948342	-0.06467424
4.750159	-0.0002376094
3.608508	0.1671859
0.4248052	1.252264
6.1918	1.394248
-7.961979	3.686417
-5.582933	3.942198



EJEMPLO

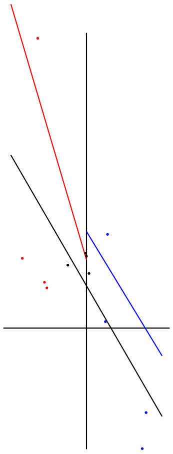
y	x
4.608895	-4.246639
19.1437	-3.225103
3.028715	-2.780972
2.654891	-2.622295
4.155246	-1.234524
4.948342	-0.06467424
4.750159	-0.0002376094
3.608508	0.1671859
0.4248052	1.252264
6.1918	1.394248
-7.961979	3.686417
-5.582933	3.942198



MCO1: $y = 4.5144 - 3.3759x$, $SCR_1 = 0.1918$, $e^t = (-0.3415, 0.1379, 0.2349, -0.0313)$

EJEMPLO

y	x
4.608895	-4.246639
19.1437	-3.225103
3.028715	-2.780972
2.654891	-2.622295
4.155246	-1.234524
4.948342	-0.06467424
4.750159	-0.0002376094
3.608508	0.1671859
0.4248052	1.252264
6.1918	1.394248
-7.961979	3.686417
-5.582933	3.942198

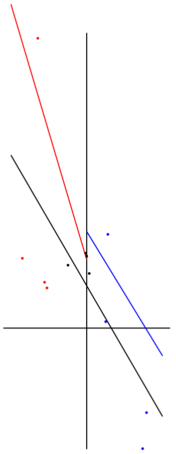


MCO1: $y = 4.5144 - 3.3759x$, $SCR_1 = 0.1918$, $e^t = (-0.3415, 0.1379, 0.2349, -0.0313)$

MCO2: $y = 6.380 - 1.642x$, $SCR_2 = 127.0477$, $e^t = (-1.5382, 7.4685, 2.1003, -8.0306)$

EJEMPLO

y	x
4.608895	-4.246639
19.1437	-3.225103
3.028715	-2.780972
2.654891	-2.622295
4.155246	-1.234524
4.948342	-0.06467424
4.750159	-0.0002376094
3.608508	0.1671859
0.4248052	1.252264
6.1918	1.394248
-7.961979	3.686417
-5.582933	3.942198



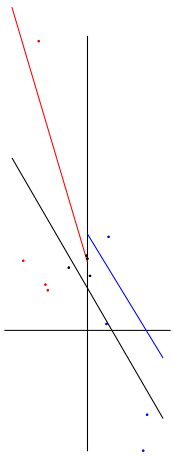
MCO1: $y = 4.5144 - 3.3759x$, $SCR_1 = 0.1918$, $e^t = (-0.3415, 0.1379, 0.2349, -0.0313)$

MCO2: $y = 6.380 - 1.642x$, $SCR_2 = 127.0477$, $e^t = (-1.5382, 7.4685, 2.1003, -8.0306)$

$$F_{exp} = \frac{SCR_2}{SCR_1} = 662.2987 > F_{2,2,0.975} = 39$$

EJEMPLO

y	x
4.608895	-4.246639
19.1437	-3.225103
3.028715	-2.780972
2.654891	-2.622295
4.155246	-1.234524
4.948342	-0.06467424
4.750159	-0.0002376094
3.608508	0.1671859
0.4248052	1.252264
6.1918	1.394248
-7.961979	3.686417
-5.582933	3.942198



MCO1: $y = 4.5144 - 3.3759x$, $SCR_1 = 0.1918$, $e^t = (-0.3415, 0.1379, 0.2349, -0.0313)$

MCO2: $y = 6.380 - 1.642x$, $SCR_2 = 127.0477$, $e^t = (-1.5382, 7.4685, 2.1003, -8.0306)$

$F_{exp} = \dfrac{SCR_2}{SCR_1} = 662.2987 > F_{2,2,0.975} = 39$: Se rechaza la hipótesis de homocedasticidad!! (hay heterocedasticidad).

Ejemplo

Utilizando una muestra de 25 observaciones anuales se estima el siguiente modelo de demanda $D_t = \beta_1 + \beta_2 Y_t + \beta_3 PR_t + u_t$.
Utilizando sólo las 10 primeras observaciones se obtiene la siguiente ecuación estimada:

$$\begin{array}{ccccccc} \widehat{D}_t & = & 80.50 & + & 0.93Y_t & - & 0.87PR_t \\ (d.t.estimadas) & & (86.17) & & (1.06) & & (1.9) \end{array}$$

donde se ha obtenido $SCR_1 = 125.7$. Del mismo modo, y utilizando las 10 últimas observaciones, se obtiene la siguiente ecuación estimada

$$\begin{array}{ccccccc} \widehat{D}_t & = & 20.61 & + & 0.53Y_t & - & 0.105PR_t \\ (d.t.estimadas) & & (221.44) & & (0.29) & & (2.41) \end{array}$$

con $SCR_2 = 498.94$. Detectar mediante el test de Goldfeld y Quandt la existencia o no de heteroscedasticidad. Hipótesis nula del contraste es $H_0 : \sigma_1^2 = \sigma_2^2 = \sigma^2$.

EJEMPLO

En esta situación el estadístico de contraste es:

$$F_{exp} = \frac{SCR_2}{SCR_1} = \frac{498.94}{125.7} = 3.9693$$

y se verifica que $F_{exp} \sim F_{\frac{n-m}{2}-k, \frac{n-m}{2}-k} = F_{\frac{25-5}{2}-3, \frac{25-5}{2}-3} = F_{7,7}$ $(m = 5, k = 3)$.

$F_{exp} > 3.79 \Rightarrow$ se rechaza la hipótesis nula a un nivel de significación del 5%.
Por tanto, concluimos que existe heteroscedasticidad.

Contrastes de Breusch-Pagan y White

Para número elevado de observaciones:

1. Estimamos el modelo $y = X\beta + u$ y el vector correspondiente a los residuos de mínimos cuadrados ordinarios, $e = y - \widehat{y}$.

Contrastes de Breusch-Pagan y White

Para número elevado de observaciones:

1. Estimamos el modelo $y = X\beta + u$ y el vector correspondiente a los residuos de mínimos cuadrados ordinarios, $e = y - \widehat{y}$.
2. Breusch-Pagan:
$$e^2 = \delta_0 + \delta_1 x_1 + \ldots + \delta_p X_p + v$$
3. White:
$$e^2 = \delta_0 + \delta_1 x_1 + \ldots + \delta_p X_p + +\delta_{p+1} x_1^2 + \delta_{p+2} x_1 x_2 + \cdots + \delta_{p+p^2} x_p^2 + v$$
4. Se contrasta $H_0 : \delta_1 = \delta_2 = \cdots = 0$ (HOMOCEDASTICIDAD)
5. Si se rechaza H_0: Existe heterocedasticidad!

Contrastes de Breusch-Pagan y White

Para número elevado de observaciones:

1. Estimamos el modelo $y = X\beta + u$ y el vector correspondiente a los residuos de mínimos cuadrados ordinarios, $e = y - \widehat{y}$.
2. Breusch-Pagan:
$$e^2 = \delta_0 + \delta_1 x_1 + \ldots + \delta_p X_p + v$$
3. White:
$$e^2 = \delta_0 + \delta_1 x_1 + \ldots + \delta_p X_p + +\delta_{p+1} x_1^2 + \delta_{p+2} x_1 x_2 + \cdots + \delta_{p+p^2} x_p^2 + v$$
4. Se contrasta $H_0 : \delta_1 = \delta_2 = \cdots = 0$ (HOMOCEDASTICIDAD)
5. Si se rechaza H_0: Existe heterocedasticidad!

```
sms.het_breuschpagan(mco.resid, mco.model.exog)
sms.het_white(mco.resid, mco.model.exog)
```

Contraste de Glejser

Si sospechamos de que la heteroscedasticidad está ligada a X_i:

Contraste de Glejser

Si sospechamos de que la heteroscedasticidad está ligada a X_i:

1. Estimamos el modelo $y = X\beta + u$ y $e = y - \widehat{y}$.

Contraste de Glejser

Si sospechamos de que la heteroscedasticidad está ligada a X_i:

1. Estimamos el modelo $y = X\beta + u$ y $e = y - \widehat{y}$.
2. Se ajusta por MCO el modelo en el que la variable endógena es $|e_t|$ y la variable exógena es la variable X_i, la cual pensamos que es la que provoca la heteroscedasticidad, que será denotada por z_t, es decir

$$|e_t| = \delta_0 + \delta_1 z_t^h + v_t$$

donde v_t es ruido blanco. La regresión se realiza para distintos valores de $h = \pm 1, \pm 2, \pm \frac{1}{2} \ldots$.

Contraste de Glejser

Si sospechamos de que la heteroscedasticidad está ligada a X_i:

1. Estimamos el modelo $y = X\beta + u$ y $e = y - \widehat{y}$.
2. Se ajusta por MCO el modelo en el que la variable endógena es $|e_t|$ y la variable exógena es la variable X_i, la cual pensamos que es la que provoca la heteroscedasticidad, que será denotada por z_t, es decir

$$|e_t| = \delta_0 + \delta_1 z_t^h + v_t$$

   donde v_t es ruido blanco. La regresión se realiza para distintos valores de $h = \pm 1, \pm 2, \pm \frac{1}{2} \ldots$.
3. Para cada h, contrastar $H_0 : \delta_1 = 0$ (t-test). Si se rechaza H_0, entonces existe heteroscedasticidad en el modelo.

Ejemplo

y_i	55	65	70	80	79	84	...	178	191	137	189
X_i	80	100	85	110	120	115	...	265	270	230	250
$\lvert e_i \rvert$	5.31	8.07	6.5	0.55	6.82	1.36	...	0.3	9.51	18.98	20.26

Para $h = \pm 1, \pm 2, \pm\frac{1}{2}$ se estima el modelo de regresión en el cual la variable endógena es el valor absoluto de los errores de estimación, $|e_t|$, y la variable exógena es la variable que pensamos que puede ser la causante del problema de heteroscedasticidad, X_i.

$$
\begin{array}{lll}
h = 1 & |\widehat{e}_t| = \underset{(3.6348)}{1.1126} + \underset{(0.01988)}{0.0331} \cdot z_t & R^2 = 0.1188 \\
h = -1 & |\widehat{e}_t| = \underset{(1.9876)}{11.0619} - \underset{(282.044)}{639.379} \cdot z_t^{-1} & R^2 = 0.0814 \\
h = 2 & |\widehat{e}_t| = \underset{(1.3159)}{3.6566} + \underset{(3.34\cdot 10^{-5})}{9.56\cdot 10^{-5}} \cdot z_t^{2} & R^2 = \mathbf{0.1233} \\
h = -2 & |\widehat{e}_t| = \underset{(1.1709)}{8.5926} - \underset{(18621.3)}{35036.3} \cdot z_t^{-2} & R^2 = 0.0575 \\
h = 1/2 & |\widehat{e}_t| = \underset{(4.0212)}{-3.8895} + \underset{(0.3056)}{0.8289} \cdot z_t^{1/2} & R^2 = 0.1126 \\
h = -1/2 & |\widehat{e}_t| = \underset{(3.8293)}{16.0441} - \underset{(47.1948)}{115.249} \cdot z_t^{-1/2} & R^2 = 0.0932
\end{array}
$$

Ejemplo

Utilizando $|t_{exp}| = |\frac{\widehat{\delta_1}}{\sqrt{\mathrm{var}[\widehat{\delta_1}]}}|$ o el p-valor se estudia la significatividad del parámetro δ_1 de la regresión $|e_t| = \delta_0 + \delta_1 z_t^h + v_t$. Si se rechaza la hipótesis nula, existirá heteroscedasticidad.

$$
\begin{array}{lll}
h = -2 & |t_{exp}| = 1.8815 \not> t_{29,0.975} = 2.045 & \text{Se mantiene } H_0 \\
h = -1 & |t_{exp}| = 2.2669 > t_{29,0.975} = 2.045 & \text{Se rechaza } H_0 \\
h = -\frac{1}{2} & |t_{exp}| = 2.4419 > t_{29,0.975} = 2.045 & \text{Se rechaza } H_0 \\
h = 1 & |t_{exp}| = 1.6649 \not> t_{29,0.975} = 2.045 & \text{Se mantiene } H_0 \\
h = 2 & |t_{exp}| = 2.8713 > t_{29,0.975} = 2.045 & \text{Se rechaza } H_0 \\
h = \frac{1}{2} & |t_{exp}| = 2.7124 > t_{29,0.975} = 2.045 & \text{Se rechaza } H_0
\end{array}
$$

Ejemplo

Utilizando $|t_{exp}| = |\frac{\widehat{\delta}_1}{\sqrt{\text{var}[\widehat{\delta}_1]}}|$ o el p-valor se estudia la significatividad del parámetro δ_1 de la regresión $|e_t| = \delta_0 + \delta_1 z_t^h + v_t$. Si se rechaza la hipótesis nula, existirá heteroscedasticidad.

$$\begin{array}{lll}
h = -2 & |t_{exp}| = 1.8815 \not> t_{29,0.975} = 2.045 & \text{Se mantiene } H_0 \\
h = -1 & |t_{exp}| = 2.2669 > t_{29,0.975} = 2.045 & \text{Se rechaza } H_0 \\
h = -\frac{1}{2} & |t_{exp}| = 2.4419 > t_{29,0.975} = 2.045 & \text{Se rechaza } H_0 \\
h = 1 & |t_{exp}| = 1.6649 \not> t_{29,0.975} = 2.045 & \text{Se mantiene } H_0 \\
h = 2 & |t_{exp}| = 2.8713 > t_{29,0.975} = 2.045 & \text{Se rechaza } H_0 \\
h = \frac{1}{2} & |t_{exp}| = 2.7124 > t_{29,0.975} = 2.045 & \text{Se rechaza } H_0
\end{array}$$

Para $h = -1; 2; \frac{1}{2}; -\frac{1}{2}$, existe heteroscedasticidad, y el modelo $|\widehat{e}_t| = 3.6566 + 9.56 \cdot 10^{-5} \cdot z_t^2$ el mejor (mayor R^2). Podemos pensar que $\mathbb{E}[u_t^2] = \sigma^2 z_t^2$.

3. Estimación bajo heterocedasticidad

Supongamos que queremos ajustar un modelo $y = X\beta + u$, donde sabemos que existe heterocedasticidad, es decir:

$$\mathbb{E}[uu^t] = \begin{pmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \sigma_n^2 \end{pmatrix}$$

3. Estimación bajo heterocedasticidad

Supongamos que queremos ajustar un modelo $y = X\beta + u$, donde sabemos que existe heterocedasticidad, es decir:

$$\mathbb{E}[uu^t] = \begin{pmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \sigma_n^2 \end{pmatrix} \overset{\sigma^2=\sigma_1^2}{\Longrightarrow} \mathbb{E}[uu^t] = \sigma^2 \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \frac{\sigma_2^2}{\sigma_1^2} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \frac{\sigma_n^2}{\sigma_1^2} \end{pmatrix} = \sigma^2\Omega$$

$\Omega = \mathrm{diag}(w_1, \dots, w_n)$, y $w_i = \frac{\sigma_i^2}{\sigma_1^2}$

3. Estimación bajo heterocedasticidad

Supongamos que queremos ajustar un modelo $y = X\beta + u$, donde sabemos que existe heterocedasticidad, es decir:

$$\mathbb{E}[uu^t] = \begin{pmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \sigma_n^2 \end{pmatrix} \overset{\sigma^2=\sigma_1^2}{\Longrightarrow} \mathbb{E}[uu^t] = \sigma^2 \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \frac{\sigma_2^2}{\sigma_1^2} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \frac{\sigma_n^2}{\sigma_1^2} \end{pmatrix} = \sigma^2\Omega$$

$\Omega = \operatorname{diag}(w_1, \dots, w_n)$, y $w_i = \frac{\sigma_i^2}{\sigma_1^2} \Rightarrow Var(u_i) = \mathbb{E}[u_i^2] = \sigma^2 w_i, \forall i.$

Mínimos Cuadrados Generalizados

Como $\Omega \succ 0$ y simétrica, por Descomposición de Cholesky, existirá una matriz $P (= Q^{-}1)$, no singular, tal que,

$$\Omega^{-1} = P^t P$$

Mínimos Cuadrados Generalizados

Como $\Omega \succ 0$ y simétrica, por Descomposición de Cholesky, existirá una matriz $P(=Q^{-}1)$, no singular, tal que,

$$\Omega^{-1} = P^t P$$

De manera que se podría transformar el modelo original para que la matriz Ω se convierta en la matriz identidad y se pudiera aplicar MCO sobre el modelo transformado aplicando estas expresiones:

$$\begin{aligned} \widehat{\beta}_{MCG} &= \left[(X^*)^t X^*\right]^{-1}\left[(X^*)^t y^*\right] = \left[X^t P^t P X\right]^{-1}\left[X^t P^t P y\right] = \\ &= \left[X^t \Omega^{-1} X\right]^{-1}\left[X^t \Omega^{-1} y\right] \end{aligned}$$

$$\text{var}\left[\widehat{\beta}_{MCG}\right] = \sigma^2\left[(X^*)^t X^*\right]^{-1} = \sigma^2\left[X^t P^t P X\right]^{-1} = \sigma^2\left[X^t \Omega^{-1} X\right]^{-1}$$

Mínimos Cuadrados Generalizados

Como $\Omega \succ 0$ y simétrica, por Descomposición de Cholesky, existirá una matriz $P(= Q^-1)$, no singular, tal que,

$$\Omega^{-1} = P^t P$$

De manera que se podría transformar el modelo original para que la matriz Ω se convierta en la matriz identidad y se pudiera aplicar MCO sobre el modelo transformado aplicando estas expresiones:

$$\begin{aligned}
\widehat{\beta}_{MCG} &= \left[(X^*)^t X^*\right]^{-1} \left[(X^*)^t y^*\right] = \left[X^t P^t P X\right]^{-1} \left[X^t P^t P y\right] = \\
&= \left[X^t \Omega^{-1} X\right]^{-1} \left[X^t \Omega^{-1} y\right]
\end{aligned}$$

$$\operatorname{var}\left[\widehat{\beta}_{MCG}\right] = \sigma^2 \left[(X^*)^t X^*\right]^{-1} = \sigma^2 \left[X^t P^t P X\right]^{-1} = \sigma^2 \left[X^t \Omega^{-1} X\right]^{-1}$$

Estimador Mínimos Cuadrados Ponderados

En el caso particular de la existencia de heterocedastidad, ¿Cómo se descompone $\Omega^{-1} = P^t P$?.

$$\Omega = \begin{pmatrix} w_1 & 0 & \dots & 0 \\ 0 & w_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & w_n \end{pmatrix}$$

Estimador Mínimos Cuadrados Ponderados

En el caso particular de la existencia de heterocedastidad, ¿Cómo se descompone $\Omega^{-1} = P^t P$?.

$$\Omega = \begin{pmatrix} w_1 & 0 & \dots & 0 \\ 0 & w_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & w_n \end{pmatrix} \Rightarrow \Omega^{-1} = \begin{pmatrix} \frac{1}{w_1} & 0 & \dots & 0 \\ 0 & \frac{1}{w_2} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \frac{1}{w_n} \end{pmatrix}$$

Estimador Mínimos Cuadrados Ponderados

En el caso particular de la existencia de heterocedastidad, ¿Cómo se descompone $\Omega^{-1} = P^t P$?.

$$\Omega = \begin{pmatrix} w_1 & 0 & \dots & 0 \\ 0 & w_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & w_n \end{pmatrix} \Rightarrow \Omega^{-1} = \begin{pmatrix} \frac{1}{w_1} & 0 & \dots & 0 \\ 0 & \frac{1}{w_2} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \frac{1}{w_n} \end{pmatrix}$$

Considerando:

$$P = \begin{pmatrix} \frac{1}{\sqrt{w_1}} & 0 & \dots & 0 \\ 0 & \frac{1}{\sqrt{w_2}} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \frac{1}{\sqrt{w_n}} \end{pmatrix}$$

Mínimos Cuadrados Ponderados

$$y^* = Py \quad X^* = PX \quad u^*Pu$$

Mínimos Cuadrados Ponderados

$$y^* = Py \quad X^* = PX \quad u^* Pu$$

$$y^* = \begin{pmatrix} \frac{y_1}{\sqrt{w_1}} \\ \vdots \\ \frac{y_n}{\sqrt{w_n}} \end{pmatrix}, \; X^* = \begin{pmatrix} \frac{1}{\sqrt{w_1}} & \frac{x_{21}}{\sqrt{w_1}} & \cdots & \frac{x_{k1}}{\sqrt{w_1}} \\ \frac{1}{\sqrt{w_2}} & \frac{x_{22}}{\sqrt{w_2}} & \cdots & \frac{x_{k2}}{\sqrt{w_2}} \\ & \ddots & \ddots & \\ \frac{1}{\sqrt{w_n}} & \frac{x_{2n}}{\sqrt{w_n}} & \cdots & \frac{x_{kn}}{\sqrt{w_n}} \end{pmatrix}, u^* = \begin{pmatrix} \frac{u_1}{\sqrt{w_1}} \\ \vdots \\ \frac{u_n}{\sqrt{w_n}} \end{pmatrix}.$$

Mínimos Cuadrados Ponderados

- $\mathbb{E}[u_i^*] = \mathbb{E}[\frac{u_i}{\sqrt{w_i}}] = \frac{1}{\sqrt{w_i}}\mathbb{E}[u_i] = 0.$

Mínimos Cuadrados Ponderados

- $\mathbb{E}[u_i^*] = \mathbb{E}[\frac{u_i}{\sqrt{w_i}}] = \frac{1}{\sqrt{w_i}}\,\mathbb{E}[u_i] = 0.$
- $Var(u_i^*) = Var(\frac{u_i}{\sqrt{w_i}}) = \frac{1}{w_i} Var(u_i) = \frac{1}{w_i}\,\sigma^2\,w_i = \sigma^2$: HOMOCEDASTICIDAD!!!

Mínimos Cuadrados Ponderados

- $\mathbb{E}[u_i^*] = \mathbb{E}[\frac{u_i}{\sqrt{w_i}}] = \frac{1}{\sqrt{w_i}}\,\mathbb{E}[u_i] = 0.$
- $Var(u_i^*) = Var(\frac{u_i}{\sqrt{w_i}}) = \frac{1}{w_i} Var(u_i) = \frac{1}{w_i}\,\sigma^2\,w_i = \sigma^2$: HOMOCEDASTICIDAD!!!
- $\mathbb{E}[u_i^*\,u_j^*] = \mathbb{E}[\frac{u_i}{\sqrt{w_i}}\,\frac{u_j}{\sqrt{w_j}}] = \frac{1}{\sqrt{w_i\,w_j}}\mathbb{E}[u_i\,u_j] = 0$: INCORRELACIÓN!!

Mínimos Cuadrados Ponderados

- $\mathbb{E}[u_i^*] = \mathbb{E}[\frac{u_i}{\sqrt{w_i}}] = \frac{1}{\sqrt{w_i}}\,\mathbb{E}[u_i] = 0.$
- $Var(u_i^*) = Var(\frac{u_i}{\sqrt{w_i}}) = \frac{1}{w_i} Var(u_i) = \frac{1}{w_i}\,\sigma^2\,w_i = \sigma^2$: HOMOCEDASTICIDAD!!!
- $\mathbb{E}[u_i^*\,u_j^*] = \mathbb{E}[\frac{u_i}{\sqrt{w_i}}\,\frac{u_j}{\sqrt{w_j}}] = \frac{1}{\sqrt{w_i\,w_j}}\mathbb{E}[u_i\,u_j] = 0$: INCORRELACIÓN!!

Podemos aplicar MCO para estimar $y^* = X^*\beta + u^*$.

EJEMPLO

En el Ejemplo del Test de Glejser se tenía que para el modelo $y = \beta_1 + \beta_2 x$ (solo una variable independiente):

$$\mathbb{E}[u_i^2] = \sigma^2 x_i^2$$

Ejemplo

En el Ejemplo del Test de Glejser se tenía que para el modelo $y = \beta_1 + \beta_2 x$ (solo una variable independiente):

$$\mathbb{E}[u_i^2] = \sigma^2 x_i^2$$

En tal caso, la transformación que debemos hacer de nuestros datos es tomando $w_i = x_i^2$, esto es:

$$y^* = \begin{pmatrix} \frac{y_1}{x_1} \\ \vdots \\ \frac{y_n}{x_n} \end{pmatrix}, \; x^* = \begin{pmatrix} \frac{1}{x_1} & \frac{x_1}{x_1} \\ \frac{1}{x_2} & \frac{x_2}{x_2} \\ \vdots & \vdots \\ \frac{1}{x_n} & \frac{x_n}{x_n} \end{pmatrix} = \begin{pmatrix} \frac{1}{x_1} & 1 \\ \frac{1}{x_2} & 1 \\ \vdots & \vdots \\ \frac{1}{x_n} & 1 \end{pmatrix}.$$

EJEMPLO

Los estimadores se obtendrían como:

$$\begin{aligned}\widehat{\beta}_{MCG} &= \left[(X^*)^t X^*\right]^{-1}\left[(X^*)^t y^*\right] = \left[X^t P^t P X\right]^{-1}\left[X^t P^t P y\right] = \\ &= \left[X^t \Omega^{-1} X\right]^{-1}\left[X^t \Omega^{-1} y\right]\end{aligned}$$

donde $\Omega^{-1} = \begin{pmatrix} \frac{1}{x_1^2} & 0 & \dots & 0 \\ 0 & \frac{1}{x_2^2} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \frac{1}{x_n^2} \end{pmatrix}$

A partir de un patrón de comportamiento de $\mathbb{E}[u_i u_j]$ en base a otra variable, realizamos la transformación de los datos y aplicamos MCO.