

Heteroscedasticidad

Román Salmerón Gómez

Universidad de Granada



Contenidos

Naturaleza del problema

Causas y consecuencias de la heteroscedasticidad

Procedimientos de Detección

Estimación en los modelos con heteroscedasticidad

Naturaleza del problema

Causas y consecuencias de la heteroscedasticidad

Procedimientos de Detección

Estimación en los modelos con heteroscedasticidad

Contenidos

Naturaleza del problema

Heteroscedasticidad

Causas y
consecuencias de la
heteroscedasticidad

Procedimientos de
Detección

Estimación en los
modelos con
heteroscedasticidad

Naturaleza del problema

En el modelo lineal general, $y = X\beta + u$, se supone que la perturbación aleatoria es tal que $E[u] = 0_{n \times 1}$ y $Var(u) = E[u \cdot u^t] = \sigma^2 I_{n \times n}$, lo cual implica que:

- $E[u_t] = 0, \forall t \in \{1, \dots, n\}$.
- $E[u_t^2] = Var(u_t) = \sigma^2, \forall t \in \{1, \dots, n\}$ (varianza constante = homocedasticidad).
- $E[u_i \cdot u_j] = Cov(u_i, u_j) = 0, \forall i \neq j \in \{1, \dots, n\}$ (incorrelación).

Cuando se incumple el supuesto de homocedasticidad, es decir, la varianza no es constante sino que varía con la observación ($E[u_t^2] = \sigma_t^2, \forall t$), se dice que hay heteroscedasticidad.

Consideramos entonces el modelo lineal general $y_{n \times 1} = X_{n \times k} \cdot \beta_{k \times 1} + u_{n \times 1}$, tal que $E[u] = 0_{n \times 1}$ y $Var(u) = \sigma^2 \cdot \Omega_{n \times n}$ donde Ω es una matriz diagonal con diagonal no constante.

Adviértase que en este caso se dice que el modelo tiene una matriz de varianzas-covarianzas no escalar o con perturbaciones no esféricas (ya que la matriz de varianzas-covarianzas de la perturbación aleatoria ya no es igual al producto de una constante por la matriz identidad).

Ejemplo (Novales - Econometría)

Contenidos

Naturaleza del problema

Heteroscedasticidad

Causas y consecuencias de la heteroscedasticidad

Procedimientos de Detección

Estimación en los modelos con heteroscedasticidad

Este problema aparece especialmente cuando se disponen de datos de sección cruzada, es decir, cuando se disponen de observaciones que miden una variable en un momento determinado para distintas entidades (individuos, familias, empresas, etc.).

Así, por ejemplo, en un modelo de consumo en función de los ingresos, $C_t = \alpha_1 + \alpha_2 I_t + u_t$, es esperable que aquellas familias con mayores ingresos tengan mayor varianza que aquellas otras con ingresos inferiores. Esto es debido a que tienen un mayor excedente sobre el que decidir qué parte ahorrar y cuál gastar.

Parecida interpretación se puede realizar sobre un modelo que estudie los dividendos que reparte una empresa a partir sus beneficios, $D_t = \beta_1 + \beta_2 B_t + v_t$. Las empresas con mayores beneficios tendrán mayor margen al fijar su política de dividendos, por lo que es esperable que la varianza de la perturbación aleatoria dependa del nivel de beneficios de cada empresa.

Contenidos

Naturaleza del problema

Causas y
consecuencias de la
heteroscedasticidad

Causas de la
heteroscedasticidad
Consecuencias de la
heteroscedasticidad

Procedimientos de
Detección

Estimación en los
modelos con
heteroscedasticidad

Causas y consecuencias de la heteroscedasticidad

Causas de la heteroscedasticidad

Contenidos

Naturaleza del problema

Causas y consecuencias de la heteroscedasticidad

Causas de la heteroscedasticidad

Consecuencias de la heteroscedasticidad

Procedimientos de Detección

Estimación en los modelos con heteroscedasticidad

Las principales causas de la presencia de heteroscedasticidad en un modelo lineal múltiple son:

- **Naturaleza del fenómeno:** en situaciones en las que se disponen de datos de sección cruzada (como las del ejemplo anterior: consumo de bienes de lujo e ingresos familiares, política de dividendos y ganancias empresariales o política de inversión y ganancias empresariales) es factible la presencia de heteroscedasticidad.
- **Usar datos agregados:** cuando las observaciones de la variable dependiente pueden dividirse en grupos (por ejemplo, una persona que reside en una provincia, un grupo de empresas que pertenecen a un mismo sector, etc.) y se usan como datos los promedios proporcionados por tales grupos es factible la presencia de heteroscedasticidad.
- **Si se omite una variable relevante en el modelo,** es esperable que la perturbación aleatoria dependa de dicha variable omitida, por lo que su varianza difícilmente será constante.

Ejemplo (Novales - Econometría)

Contenidos

Naturaleza del problema

Causas y consecuencias de la heteroscedasticidad

Causas de la heteroscedasticidad

Consecuencias de la heteroscedasticidad

Procedimientos de Detección

Estimación en los modelos con heteroscedasticidad

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:

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

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

Consecuencias de la heteroscedasticidad

Contenidos

Naturaleza del problema

Causas y consecuencias de la heteroscedasticidad

Causas de la heteroscedasticidad

Consecuencias de la heteroscedasticidad

Procedimientos de Detección

Estimación en los modelos con 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 del modelo con perturbaciones no esféricas será:

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

Dicho estimador sigue siendo lineal e insesgado, ya que las condiciones que conducen a verificar dichas propiedades en el modelo con perturbaciones esféricas no se han modificado (las demostraciones son idénticas a tal caso). Sin embargo, ya no se tiene asegurado que la varianza sea mínima, ya que ahora:

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

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

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.

Contenidos

Naturaleza del problema

Causas y
consecuencias de la
heteroscedasticidad

**Procedimientos de
Detección**

Métodos gráficos

Métodos analíticos

Estimación en los
modelos con
heteroscedasticidad

Procedimientos de Detección

Contenidos

Naturaleza del problema

Causas y consecuencias de la heteroscedasticidad

Procedimientos de Detección

Métodos gráficos

Métodos analíticos

Estimación en los modelos con heteroscedasticidad

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

En primer lugar usaremos 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.

Puesto que tomar una decisión a partir de un procedimiento gráfico no es muy adecuado ya que son fácilmente manipulables, recurriremos a métodos analíticos para determinar la presencia de heteroscedasticidad en el modelo. Los test estudiados son el 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.

Dentro de los procedimientos gráficos consideraremos:

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.

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.

Contenidos

Naturaleza del problema

Causas y consecuencias de la heteroscedasticidad

Procedimientos de Detección

Métodos gráficos

Métodos analíticos

Estimación en los modelos con heteroscedasticidad

Supongamos que para el ejemplo anterior en el que deseamos analizar los dividendos de una empresa en función de sus beneficios disponemos de las 20 observaciones de la tabla 1.

Tras realizar la estimación por MCO se obtiene que:

$$\widehat{D}_t = 10'2229 + 0'0638456 \cdot B_t$$

$$(1'35823) \quad (0'011223)$$

con $R^2 = 0'642$ y donde entre paréntesis se especifica la desviación típica estimada de cada coeficiente estimado.

Tabla 1: Dividendos y beneficios de 20 empresas

Empresas	Dividendos	Beneficios
1	13'2	61
2	15	78
3	22'2	158
4	15'2	110
5	16'1	85
6	18'5	150
7	15'5	140
8	15	70
9	20	122
10	15	70
11	21	140
12	16'2	91
13	18'5	105
14	17	115
15	17'5	115
16	22	160
17	18	165
18	23	170
19	17	130
20	17	90

Dada la naturaleza del problema, tal y como se ha indicado, sospechamos la posible presencia de heteroscedasticidad en el modelo. Por tal motivo, en primer lugar usaremos los métodos gráficos para intentar detectarla.

Contenidos

Naturaleza del problema

Causas y consecuencias de la heteroscedasticidad

Procedimientos de Detección

Métodos gráficos

Métodos analíticos

Estimación en los modelos con heteroscedasticidad

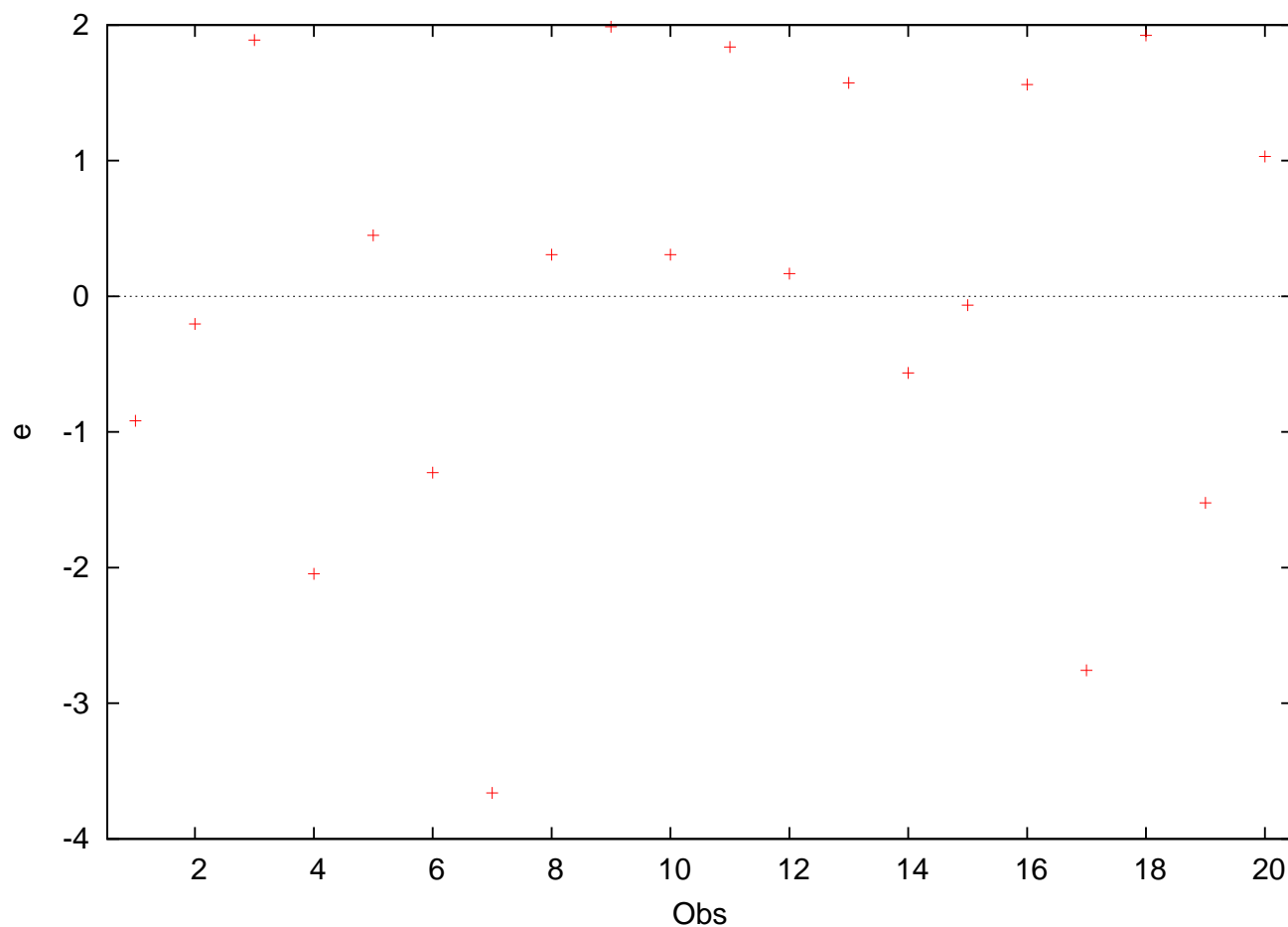


Figura 1: Gráfico de los residuos

Contenidos

Naturaleza del problema

Causas y consecuencias de la heteroscedasticidad

Procedimientos de Detección

Métodos gráficos

Métodos analíticos

Estimación en los modelos con heteroscedasticidad

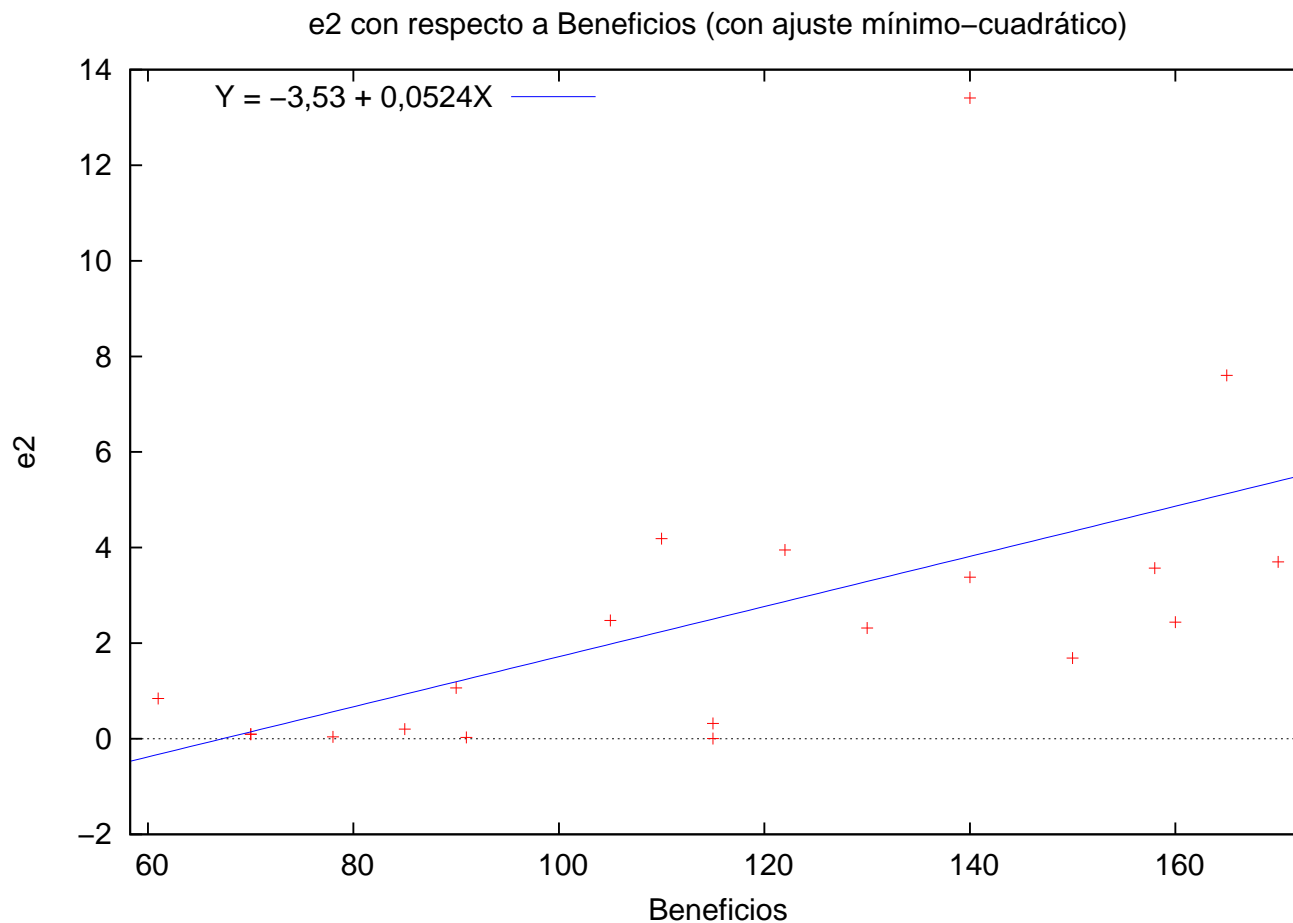


Figura 2: Gráfico de dispersión

Contenidos

Naturaleza del problema

Causas y consecuencias de la heteroscedasticidad

Procedimientos de Detección

Métodos gráficos

Métodos analíticos

Estimación en los modelos con heteroscedasticidad

En el gráfico de los residuos de la figura 1 podemos observar que los grupos de observaciones dados por 1-7, 8-13 y 14-20 tienen distinta varianza. Al mismo tiempo, en el gráfico de dispersión de la figura 2 observamos que la varianza de los residuos al cuadrado aumenta conforme lo hace la variable dividendo.

Todo lo anteriormente expuesto nos hace pensar en la presencia de heteroscedasticidad en el modelo. Si bien, para confirmar este hecho recurriremos a los distintos métodos analíticos que disponemos.

Algunos de dichos procedimientos analíticos presuponen que la heteroscedasticidad está provocada por una variable independiente en concreto. Dicha información se puede obtener a partir de los procedimientos gráficos.

1. Ajustar el modelo original por MCO y obtener los residuos, e_t .
2. Ajustar por MCO la regresión auxiliar que tiene como variable dependiente el valor absoluto de los residuos anteriores y como independiente la variable que se supone provoca la heteroscedasticidad elevada a h . Esto es:

$$|e_t| = \alpha + \beta X_t^h + v_t.$$

Los valores más comunes para h son ± 2 , ± 1 y $\pm 1/2$.

3. Observar los contrastes de significación individual de la pendiente de la regresión auxiliar. Si rechazamos $H_0 : \beta = 0$, indicará presencia de heteroscedasticidad.

Al rechazar $H_0 : \beta = 0$ entonces los residuos dependerán de X_t^h , por lo que habrá heteroscedasticidad en el modelo y podemos considerar que $E[u_t^2] = \sigma^2 X_t^h$.

Contenidos

Naturaleza del problema

Causas y consecuencias de la heteroscedasticidad

Procedimientos de Detección

Métodos gráficos

Métodos analíticos

Estimación en los modelos con heteroscedasticidad

Tras realizar la estimación por MCO del modelo original (ejemplo anterior) hemos obtenido los errores o residuos de la tabla 2. A partir de los cuales hemos ajustado las siguientes regresiones auxiliares:

$$|\hat{e}_t| = 2'15499 - 8633'51 \cdot B_t^{-2} \quad R^2 = 0'3617$$

(0'3191) (2703'4)

$$|\hat{e}_t| = 3'1728 - 197'755 \cdot B_t^{-1} \quad R^2 = 0'4163$$

(0'5479) (55'1884)

$$|\hat{e}_t| = 5'212 - 40'6989 \cdot B_t^{-1/2} \quad R^2 = 0'4374$$

(1'0576) (10'8788)

$$|\hat{e}_t| = -2'93457 + 0'39747 \cdot B_t^{1/2} \quad R^2 = 0'4619$$

(1'09025) (0'1011)

$$|\hat{e}_t| = -0'891826 + 0'01888 \cdot B_t \quad R^2 = 0'4646$$

(0'5783) (0'004778)

$$|\hat{e}_t| = 0'138792 + 0'000079 \cdot B_t^2 \quad R^2 = 0'4536$$

(0'3424) (0'0000205)

Tabla 2: Residuos del modelo original

Empresas	Beneficios	$ e_t $
1	61	0'917530
2	78	0'202905
3	158	1'889445
4	110	2'045965
5	85	0'450176
6	150	1'299790
7	140	3'661333
8	70	0'307860
9	122	1'987888
10	70	0'307860
11	140	1'838667
12	91	0'167102
13	105	1'573263
14	115	0'565193
15	115	0'065193
16	160	1'561754
17	165	2'757474
18	170	1'923298
19	130	1'522877
20	90	1'030948

En todos los casos rechazamos H_0 , por lo que hay heteroscedasticidad. Además, como el mejor modelo es el quinto (mayor coeficiente de determinación), pensamos que $E[u_t^2] = \sigma^2 B_t$.

Contenidos

Naturaleza del problema

Causas y consecuencias de la heteroscedasticidad

Procedimientos de Detección

Métodos gráficos

Métodos analíticos

Estimación en los modelos con heteroscedasticidad

1. Ordenar de menor a mayor las observaciones respecto de la variable que consideramos que produce la heteroscedasticidad.
2. Eliminar m observaciones centrales (normalmente un tercio de la muestra).
3. Ajustar por MCO los dos grupos de observaciones restantes y calcular sus suma de cuadrados de los residuos (SCR_1 para el primer grupo y SCR_2 para el segundo).
4. Calcular el estadístico de contraste $F_{exp} = \frac{SCR_2}{SCR_1}$, que sigue una distribución F de Snedecor con $\frac{n-m}{2}$ grados de libertad en el numerador y en el denominador.
5. Si $F_{exp} > F_{\frac{n-m}{2}, \frac{n-m}{2}}(1 - \alpha)$ rechazamos la hipótesis nula de homocedasticidad.

Adviértase que algunos textos consideran que los grados de libertad de la F de Snedecor son $\frac{n-m-2k}{2}$. Nosotros usaremos cualquiera de las dos versiones.

Contenidos

Naturaleza del problema

Causas y consecuencias de la heteroscedasticidad

Procedimientos de Detección

Métodos gráficos

Métodos analíticos

Estimación en los modelos con heteroscedasticidad

Tras reordenar de forma creciente los datos con respecto a la variable que provoca la heteroscedasticidad (beneficios), ajustamos por MCO las regresiones formadas por las 7 primeras observaciones y las 7 últimas, ya que suprimimos las 6 observaciones centrales:

$$\hat{D}_t = 7'5343 + 0'100477 \cdot B_t \quad R^2 = 0'8739$$

$$(1'3407) \quad (0'017) \quad SCR_1 = 1'134109$$

$$\hat{D}_t = 1'15735 + 0'121975 \cdot B_t \quad R^2 = 0'2735$$

$$(13'7885) \quad (0'0889) \quad SCR_2 = 32'934$$

Tabla 3: Observaciones reordenadas

Empresa	Dividendos	Beneficios
1	13'2	61
8	15	70
10	15	70
2	15	78
5	16'1	85
20	17	90
12	16'2	91
13	18'5	105
4	15'2	110
15	17'5	115
14	17	115
9	20	122
19	17	130
7	15'5	140
11	21	140
6	18'5	150
3	22'2	158
16	22	160
17	18	165
18	23	170

Como $F_{exp} = \frac{SCR_2}{SCR_1} = \frac{32'934}{1'134109} = 29'03954 > 3'17889 = F_{7,7}(0'95)$, entonces rechazamos la hipótesis nula de homocedasticidad. Esto es, hay heteroscedasticidad en el modelo considerado.

1. Ajustar el modelo original por MCO y obtener los residuos, e_t .
2. Ajustar por MCO la regresión auxiliar en la que la variable dependiente son los residuos al cuadrado y las independientes todas las variables del modelo original, sus cuadrados y todos los posibles productos cruzados (omitiendo las repeticiones).
3. Calcular el estadístico de contraste $\chi_{exp} = nR_{aux}^2$, donde R_{aux}^2 es el coeficiente de determinación de la regresión auxiliar, que sigue una distribución chi-cuadrado con $p - 1$ grados de libertad donde p es el número de regresores de la regresión auxiliar.
4. Si $\chi_{exp} > \chi_{p-1}^2(1 - \alpha)$, entonces rechazamos la hipótesis nula de homocedasticidad.

El test de Breusch-Pagan coincide con este test con la única diferencia de que la regresión auxiliar normalmente sólo se consideran como independientes todas las variables del modelo.

Al aplicar el test de Glesjer ya hemos ajustado el modelo original por MCO y obtenido los residuos, por lo que procedemos a ajustar directamente la regresión auxiliar:

$$\widehat{e}_t^2 = -5'17545 + 0'0833 \cdot B_t - 0'0001328 \cdot B_t^2 \quad R^2 = 0'3115679$$

(8'2864) (0'1505) (0'00064)

Puesto que $\chi_{exp} = n \cdot R_{aux}^2 = 20 \cdot 0'315679 = 6'313575 > 5'99146 = \chi_2^2$, entonces rechazamos la hipótesis nula de homocedasticidad y, por tanto, en el modelo considerado hay presencia de heteroscedasticidad.

Para el test de Breusch-Pagan la regresión auxiliar sería:

$$\widehat{e}_t^2 = -3'52638 + 0'0524 \cdot B_t \quad R^2 = 0'313955$$

(2'2116) (0'0182)

En tal caso, $\chi_{exp} = n \cdot R_{aux}^2 = 20 \cdot 0'313955 = 6'2791 > 3'84146 = \chi_1^2$ y entonces rechazamos la hipótesis nula de homocedasticidad.

Contenidos

Naturaleza del problema

Causas y
consecuencias de la
heteroscedasticidad

Procedimientos de
Detección

Estimación en los
modelos con
heteroscedasticidad

Mínimos Cuadrados
Ponderados

Ejemplo

Estimación en los modelos con heteroscedasticidad

Mínimos Cuadrados Ponderados

Contenidos

Naturaleza del problema

Causas y consecuencias de la heteroscedasticidad

Procedimientos de Detección

Estimación en los modelos con heteroscedasticidad

Mínimos Cuadrados Ponderados

Ejemplo

Hemos comprobado que en un modelo con perturbaciones no esféricas el estimador por mínimos cuadrados ordinarios es lineal e insesgado pero no tenemos asegurado que sea óptimo (en el sentido de varianza mínima).

Para resolver este problema surgen los mínimos cuadrados generalizados. Dicho método consiste en transformar un modelo con perturbaciones no esféricas en otro con perturbaciones esféricas, de forma que al aplicarle a este último el método de mínimos cuadrados ordinarios se obtenga un estimador lineal, insesgado y óptimo.

En dicha transformación es fundamental el teorema de Aitken, el cual afirma que al ser Ω una matriz simétrica definida positiva entonces existe una matriz regular, P , tal que $P^t P = \Omega^{-1}$ (en tal caso se verifica que $\Omega = P^{-1} (P^{-1})^t$, de donde se deduce que $P \Omega P^t = I_{n \times n}$).

Esto es, premultiplicando el modelo con perturbaciones no esféricas por una matriz no estocástica, P , se obtiene que $y_* = X_* \beta + u_*$, donde $y_* = Py$, $X_* = PX$ y $u_* = Pu$. Entonces:

$$E[u_*] = E[Pu] = PE[u] = 0_{n \times 1},$$

$$E[u_* \cdot u_*^t] = E[Pu u^t P^t] = PE[uu^t]P^t = \sigma^2 P \Omega P^t = \sigma^2 I_{n \times n}.$$

Mínimos Cuadrados Ponderados

Contenidos

Naturaleza del problema

Causas y consecuencias de la heteroscedasticidad

Procedimientos de Detección

Estimación en los modelos con heteroscedasticidad

Mínimos Cuadrados Ponderados

Ejemplo

Consideremos el modelo lineal general, $y = X\beta + u$, tal que

$$E[u] = 0_{n \times 1}, \quad Var(u) = E[u \cdot u^t] = \begin{pmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_n^2 \end{pmatrix},$$

es decir, la perturbación aleatoria es incorrelada, $E[u_i \cdot u_j] = 0, \forall i \neq j \in \{1, \dots, n\}$, y heterocedástica, $Var(u_t) = E[u_t^2] = \sigma_t^2, t = 1, \dots, n$.

Por cuestión de notación consideraremos:

$$E[u \cdot u^t] = \sigma_1^2 \cdot \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \frac{\sigma_2^2}{\sigma_1^2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{\sigma_n^2}{\sigma_1^2} \end{pmatrix} = \sigma^2 \cdot \Omega,$$

donde $\sigma^2 = \sigma_1^2$ y $\Omega = diag(w_1, \dots, w_n)$ con $w_t = \frac{\sigma_t^2}{\sigma_1^2}, t = 1, \dots, n$.

Mínimos Cuadrados Ponderados

Contenidos

Naturaleza del problema

Causas y consecuencias de la heteroscedasticidad

Procedimientos de Detección

Estimación en los modelos con heteroscedasticidad

Mínimos Cuadrados Ponderados

Ejemplo

Por tanto, se ha considerado que la varianza de la perturbación aleatoria tiene la siguiente estructura:

$$\text{Var}(u_t) = E[u_t^2] = \sigma^2 \cdot w_t, \quad \forall t.$$

Esta suposición final es fundamental para que el modelo pueda ser estimado. En principio, si suponemos que todas las varianzas son distintas entre sí, tendríamos que estimar $k + n$ parámetros: k correspondientes a los coeficientes de las variables independientes y n correspondientes a la varianza de la perturbación aleatoria. Ante esta situación, habría más parámetros a estimar que observaciones disponibles, n . Estimación absolutamente imposible.

Ahora bien, teniendo en cuenta que w_t estará relacionado con las variables independientes (tal y como se ha puesto de manifiesto en los métodos analíticos), simplemente habría que estimar σ^2 . Por tanto, hemos pasado a estimar un sólo parámetro en lugar de n . Es decir, hay que estimar en el modelo $k + 1$ parámetros. Cuestión que ahora si es asumible.

Mínimos Cuadrados Ponderados

Contenidos

Naturaleza del problema

Causas y consecuencias de la heteroscedasticidad

Procedimientos de Detección

Estimación en los modelos con heteroscedasticidad

Mínimos Cuadrados Ponderados

Ejemplo

Para resolver el problema de heteroscedasticidad, como hemos indicado, hay que encontrar una matriz P no estocástica tal que $P^t P = \Omega^{-1}$. En este caso, puesto que

$$\Omega^{-1} = \begin{pmatrix} \frac{1}{w_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{w_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{w_n} \end{pmatrix},$$

es claro que

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

Mínimos Cuadrados Ponderados

Contenidos

Naturaleza del problema

Causas y consecuencias de la heteroscedasticidad

Procedimientos de Detección

Estimación en los modelos con heteroscedasticidad

Mínimos Cuadrados Ponderados

Ejemplo

Comprobemos que el modelo transformado a partir de la matriz P anterior, $y_* = X_*\beta + u_*$, donde

$$y_* = Py = \begin{pmatrix} \frac{Y_1}{\sqrt{w_1}} \\ \frac{Y_2}{\sqrt{w_2}} \\ \vdots \\ \frac{Y_n}{\sqrt{w_n}} \end{pmatrix}, \quad u_* = Pu = \begin{pmatrix} \frac{u_1}{\sqrt{w_1}} \\ \frac{u_2}{\sqrt{w_2}} \\ \vdots \\ \frac{u_n}{\sqrt{w_n}} \end{pmatrix},$$

$$X_* = PX = \begin{pmatrix} \frac{1}{\sqrt{w_1}} & \frac{X_{21}}{\sqrt{w_1}} & \cdots & \frac{X_{k1}}{\sqrt{w_1}} \\ \frac{1}{\sqrt{w_2}} & \frac{X_{22}}{\sqrt{w_2}} & \cdots & \frac{X_{k2}}{\sqrt{w_2}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{\sqrt{w_n}} & \frac{X_{2n}}{\sqrt{w_n}} & \cdots & \frac{X_{kn}}{\sqrt{w_n}} \end{pmatrix},$$

es un modelo con perturbaciones esféricas. En efecto, es claro que:

$$Y_t^* = \frac{Y_t}{\sqrt{w_t}}, \quad u_t^* = \frac{u_t}{\sqrt{w_t}}, \quad X_{it}^* = \frac{X_{it}}{\sqrt{w_t}}, \quad i = 1, \dots, k, \quad t = 1, \dots, n.$$

Mínimos Cuadrados Ponderados

Contenidos

Naturaleza del problema

Causas y consecuencias de la heteroscedasticidad

Procedimientos de Detección

Estimación en los modelos con heteroscedasticidad

Mínimos Cuadrados Ponderados

Ejemplo

Y entonces:

$$E[u_t^*] = E\left[\frac{u_t}{\sqrt{w_t}}\right] = \frac{1}{\sqrt{w_t}} \cdot E[u_t] = 0, \forall t$$

$$Var(u_t^*) = Var\left(\frac{u_t}{\sqrt{w_t}}\right) = \frac{1}{w_t} \cdot Var(u_t) = \frac{\sigma^2 \cdot w_t}{w_t} = \sigma^2, \forall t$$

$$Cov(u_i^*, u_j^*) = E[u_i^* \cdot u_j^*] = E\left[\frac{u_i}{\sqrt{w_i}} \cdot \frac{u_j}{\sqrt{w_j}}\right] = \frac{E[u_i \cdot u_j]}{\sqrt{w_i w_j}} = 0, \forall i \neq j$$

Es evidente que el modelo transformado es un modelo con perturbaciones esféricas, por lo que las estimaciones obtenidas a partir del mismo serán lineales, insesgadas y óptimas.

Adviértase que en el modelo transformado se ponderan los datos inversamente al valor de la desviación típica de las perturbaciones. Por tal motivo, el método de mínimos cuadrados generalizados (MCG) recibe en este caso el nombre de mínimos cuadrados ponderados (MCP).

En el modelo considerado sobre el análisis de los dividendos de las empresas en función de sus beneficios hemos constatado, a partir de 20 observaciones, la presencia de heteroscedasticidad en dicho modelo. Por tanto, las estimaciones obtenidas

$$\hat{D}_t = 10'2229 + 0'0638456 \cdot B_t \quad R^2 = 0'642$$

$$(1'35823) \quad (0'011223)$$

no son óptimas.

Gracias al test de Glesjer hemos supuesto que $E[u_t^2] = \sigma^2 \cdot B_t$, es decir, la perturbación aleatoria depende directamente de los beneficios. Por tanto, lo que en teoría se denotaba como w_t corresponde a la variable beneficios, B_t . Luego para transformar los datos habrá que dividir por la raíz cuadrada de dicha variable:

$$D_t^* = \frac{D_t}{\sqrt{B_t}}, \quad cte_t^* = \frac{1}{\sqrt{B_t}}, \quad B_t^* = \frac{B_t}{\sqrt{B_t}} = \sqrt{B_t},$$

con $t = 1, \dots, 20$.

A partir de los datos transformados (ver la tabla 4) se obtiene la siguiente estimación por MCO del modelo transformado:

$$\widehat{D}_t^* = 10'2147 \cdot cte_t^* + 0'0639 \cdot B_t^*$$

(1'1196) (0'011)

con $R^2 = 0'9931$ (adviértase que los valores numéricos de las nuevas estimaciones no difieren mucho de las originales).

Para estudiar si se ha resuelto el problema de heteroscedasticidad, realizaremos el contraste de White para su detección:

$$\widehat{e}_t^2 = 101'913 \cdot cte_t^* + 1'42678 \cdot B_t^* - 199'428 \cdot cte_t^{*2} - 0'03957 \cdot B_t^{*2} - 18'4964 \cdot cte_t^* \cdot B_t^*$$

(501'18) (4'907) (1248'59) (74'7421) (0'1197)

con $R^2 = 0'301086$. Como $n \cdot R_{aux}^2 = 20 \cdot 0'301086 = 6'021711 \not\geq 9'48773 = \chi_4^2(0'95)$ no se rechaza la hipótesis nula de homocedasticidad. Por tanto, se ha resuelto el problema.

Tabla 4: Datos del modelo transformado

Empresa	B_t^*	D_t^*	cte_t^*
1	7'81025	1'690087	0'1280369
2	8'83176	1'698416	0'1132277
3	12'56981	1'766137	0'07956
4	10'48809	1'449263	0'09535
5	9'21954	1'746290	0'1084652
6	12'24745	1'510519	0'08165
7	11'83216	1'309989	0'08452
8	8'36660	1'792843	0'1195229
9	11'04536	1'810715	0'09054
10	8'36660	1'792843	0'1195229
11	11'83216	1'774824	0'08452
12	9'53939	1'698221	0'1048285
13	10'24695	1'805415	0'09759
14	10'72381	1'585258	0'09325
15	10'72381	1'631883	0'09325
16	12'64911	1'739253	0'07906
17	12'84523	1'401298	0'07785
18	13'03840	1'764019	0'07670
19	11'40175	1'490999	0'08771
20	9'48683	1'791957	0'1054093

Contenidos

Naturaleza del problema

Causas y consecuencias de la heteroscedasticidad

Procedimientos de Detección

Estimación en los modelos con heteroscedasticidad

Mínimos Cuadrados Ponderados

Ejemplo

- [1] Salmerón, R. y Tamayo, J. (2010). Técnicas cuantitativas aplicadas al análisis de la flexibilidad en la producción, la explotación y la exploración en las empresas. Revista Estadística Española, Volumen 52, Número 175, Páginas 529-567.
- [2] Gómez, S. y Salmerón, R. (2011). Influencia del entorno institucional en el desarrollo del emprendimiento español. Un análisis empírico. Revista Venezolana de Gerencia, Volumen 16, Número 54, Páginas 191-208.
- [3] Salmerón, R., Gutiérrez, R., López, M.M. y García, C. (2012). Evaluando la evaluación continua. III Jornadas sobre Innovación Docente y Adaptación al EEES en las Titulaciones Técnicas. Granada 20-21 septiembre 2012.

Puedes encontrarlas en la dirección web:

<http://www.ugr.es/local/romansg/material/WebEco/index.html>

Contenidos

Naturaleza del problema

Causas y consecuencias de la heteroscedasticidad

Procedimientos de Detección

Estimación en los modelos con heteroscedasticidad

Mínimos Cuadrados Ponderados

Ejemplo

- [1] Gujarati, D. (1997). *Econometría*. Ed. McGraw Hill. Capítulo 9.
- [2] Johnston, J. (1989). *Métodos de Econometría*. Ed. Vicens-Vives. Capítulo 8.
- [3] Novales, A. (1993). *Econometría*. McGraw Hill. Capítulo 6.
- [4] Uriel, E., Contreras, D., Moltó, M.L. y Peiró, A. (1990). *Econometría. El Modelo Lineal*. Editorial AC. Capítulo 10.
- [5] Wooldridge, J.M. (2005). *Introducción a la Econometría: Un enfoque moderno*. Thomson. Capítulo 8.