Econometría

Quinta edición

Damodar N. Gujarati

Profesor emérito de Economía
United States Military Academy, West Point

Dawn C. Porter

University of Southern California

Revisión técnica:

Aurora Monroy Alarcón

Instituto Tecnológico Autónomo de México (ITAM)

José Héctor Cortés Fregoso

Centro Universitario de Ciencias Económico-Administrativas (CUCEA)
Universidad de Guadalajara

McGraw Hill

MÉXICO • BOGOTÁ • BUENOS AIRES • CARACAS • GUATEMALA
MADRID • NUEVA YORK • SAN JUAN • SANTIAGO • SÃO PAULO
AUCKLAND • LONDRES • MILÁN • MONTREAL • NUEVA DELHI
SAN FRANCISCO • SINGAPUR • SAN LUIS • SIDNEY • TORONTO

AVISO LEGAL: ESTA INFORMACIÓN PUEDE SER UTILIZADA PARA FINES EDUCATIVOS EXCLUSIVAMENTE.

Capítulo 11

Heteroscedasticidad: ¿qué pasa si la varianza del error no es constante?

USO DIDÁCTICO

Un supuesto importante del modelo clásico de regresión lineal (supuesto 4) es que las perturbaciones u_i que aparecen en la función de regresión poblacional son homoscedásticas; es decir, que todas tienen la misma varianza. En este capítulo examinaremos la validez de este supuesto y también lo que sucede si no se cumple. Como en el capítulo 10, buscamos respuestas a las siguientes preguntas:

1. ¿Cuál es la naturaleza de la heteroscedasticidad?
2. ¿Cuáles son sus consecuencias?
3. ¿Cómo se detecta?
4. ¿Qué remedios existen?

11.1 Naturaleza de la heteroscedasticidad

Como vimos en el capítulo 3, un supuesto importante del modelo clásico de regresión lineal es que la varianza de cada término de perturbación u_i, condicional a los valores seleccionados de las variables explicativas, es algún número constante igual a σ^2. Éste es el supuesto de **homoscedasticidad**, o *igual* (homo) *dispersión* (cedasticidad), es decir, *igual varianza*. Simbólicamente,

$$E\left(u_i^2\right) = \sigma^2 \qquad i = 1, 2, \ldots, n \tag{11.1.1}$$

Gráficamente, la homoscedasticidad en el modelo de regresión con dos variables se observa en la figura 3.4, la cual, por conveniencia, se reproduce como la figura 11.1. Como lo indica esta figura, la varianza condicional de Y_i (la cual es igual a la de u_i), condicional a las X_i dadas, permanece igual sin importar los valores que tome la variable X.

En contraste, observe la figura 11.2, donde la varianza condicional de Y_i aumenta a medida que lo hace X. Aquí, las varianzas de Y_i no son las mismas. Por tanto, hay heteroscedasticidad. Simbólicamente,

$$E\left(u_i^2\right) = \sigma_i^2 \tag{11.1.2}$$

FIGURA 11.1
Perturbaciones homoscedásticas.

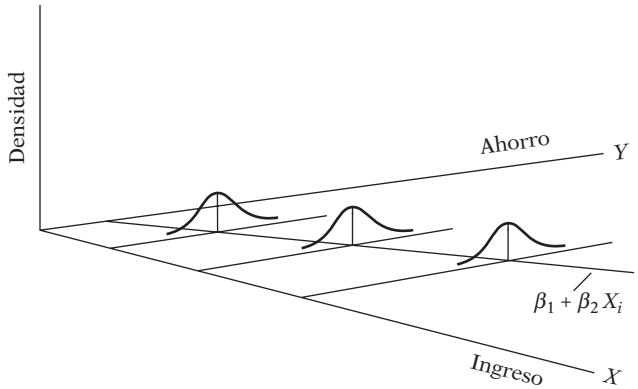
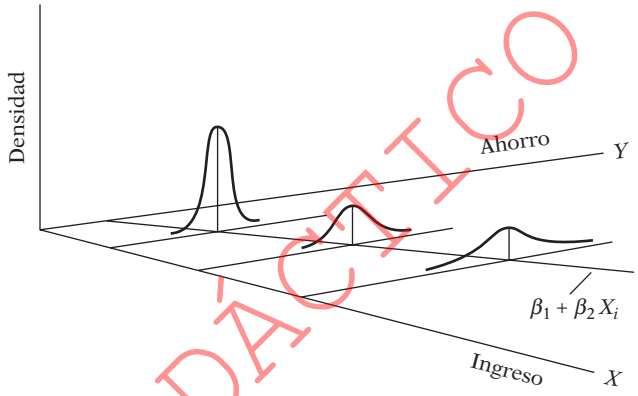


FIGURA 11.2
Perturbaciones heteroscedásticas.



USO DIDÁCTICO

Observe el subíndice de σ^2, que indica que las varianzas condicionales de u_i (= varianzas condicionales de Y_i) ya no son constantes.

Para entender la diferencia entre homoscedasticidad y heteroscedasticidad, suponga que en el modelo con dos variables $Y_i = \beta_1 + \beta_2 X_i + u_i$, Y representa el ahorro y X el ingreso. Las figuras 11.1 y 11.2 indican que, a medida que el ingreso aumenta, el ahorro en promedio también aumenta. Pero, en la figura 11.1, la varianza del ahorro permanece igual en todos los niveles de ingreso, mientras que en la figura 11.2 se incrementa con aumentos del ingreso. Parece que en la figura 11.2, en promedio, las familias de ingresos más altos ahorran más que las de ingresos más bajos, pero también hay más variabilidad en su ahorro.

Hay diversas razones por las cuales las varianzas de u_i pueden ser variables, algunas de las cuales son las siguientes.[1]

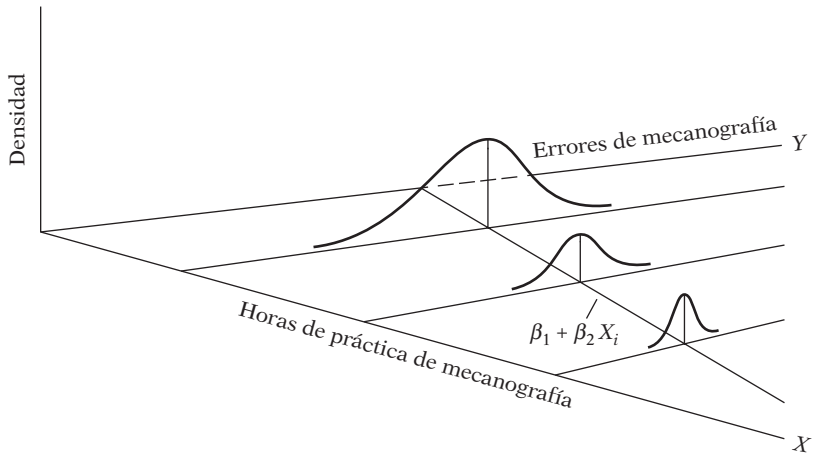
1. Con base en los *modelos de aprendizaje de los errores*, a medida que la gente aprende, disminuyen sus errores de comportamiento con el tiempo. En este caso, esperamos que σ_i^2 se reduzca. Como ejemplo, considere la figura 11.3, que relaciona el número de errores cometidos en una prueba de mecanografía durante un periodo dado. Como indica la figura 11.3, a medida que aumenta el número de horas de esta práctica, se reduce el número promedio de errores de mecanografía, al igual que sus varianzas.

2. A medida que aumentan los ingresos, la gente posee más *ingreso discrecional*[2] y, por tanto, tiene mayores posibilidades de decidir cómo disponer de su ingreso. En consecuencia,

[1] Véase Stefan Valavanis, *Econometrics,* McGraw-Hill, Nueva York, 1959, p. 48.

[2] Como señala Valavanis, "El ingreso aumenta y ahora la gente escasamente distingue los dólares, mientras que antes distinguía las monedas de diez centavos", *ibid.*, p. 48.

FIGURA 11.3
Ilustración de la heteroscedasticidad.



es probable que σ_i^2 aumente con el ingreso. Así, en la regresión del ahorro sobre el ingreso, es probable encontrar que σ_i^2 aumenta con el ingreso (como en la figura 11.2), pues las personas tienen mayores posibilidades de determinar su comportamiento respecto del ahorro. En forma similar, se espera que las compañías con mayores ganancias presenten mayor variabilidad en sus políticas de dividendos que las compañías cuyas ganancias son menores. Además, es probable que las empresas *orientadas al crecimiento* presenten una mayor variabilidad en sus tasas de pago de dividendos que las empresas ya establecidas.

3. A medida que mejoran las técnicas de recolección de datos, es probable que σ_i^2 se reduzca. Así, es probable que los bancos con equipos complejos de procesamiento de información cometan menos errores en los informes mensuales o trimestrales de sus clientes que los bancos que no los posean.

4. La heteroscedasticidad también surge por la presencia de **datos atípicos** o **aberrantes**. Una observación atípica es la que es muy diferente (muy pequeña o muy grande) en relación con las demás observaciones en la muestra. De manera más precisa, un dato atípico es una observación que proviene de una población distinta a la que genera las demás observaciones de la muestra.[3] La inclusión o exclusión de una observación de este tipo, en especial si el tamaño de la muestra es pequeño, puede alterar sustancialmente los resultados del análisis de regresión.

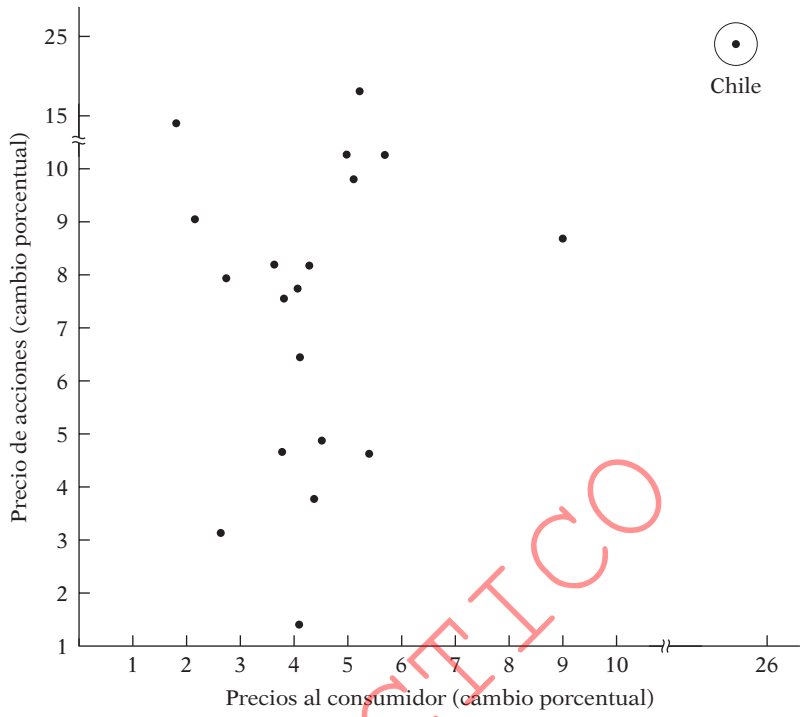
Como ejemplo considere el diagrama de dispersión de la figura 11.4. Con base en los datos de la tabla 11.9 del ejercicio 11.22, en esta figura graficamos la tasa de cambio porcentual de los precios de las acciones (Y) y los precios al consumidor (X) para el periodo posterior a la Segunda Guerra Mundial hasta 1969 en 20 países. En esta figura, la observación sobre Y y X para Chile se considera atípica, pues los valores Y y X son mucho más grandes que para el resto de los países. En estas situaciones es difícil mantener el supuesto de homoscedasticidad. En el ejercicio 11.22 se le pide ver qué sucede con los resultados de la regresión si se retiran del análisis las observaciones de Chile.

5. Otra fuente de heteroscedasticidad surge de la violación del supuesto 9 del MCRL, que establece que el modelo de regresión está correctamente especificado. Aunque analizaremos más a fondo los errores de especificación en el capítulo 13, con mucha frecuencia lo que parece heteroscedasticidad puede deberse a que se omitan del modelo algunas variables importantes. Así, en la función de demanda de un bien, si no se incluyen los precios de los bienes que le son complementarios o con los que compite (sesgo de variable omitida), los residuos de la regresión pueden dar la clara impresión de que la varianza del error no es constante. Pero si se incluyen en el modelo las variables omitidas, esa impresión puede desaparecer.

[3] El autor agradece a Michael McAleer por este señalamiento.

USO DIDÁCTICO

FIGURA 11.4
Relación entre precios de acciones y precios al consumidor.



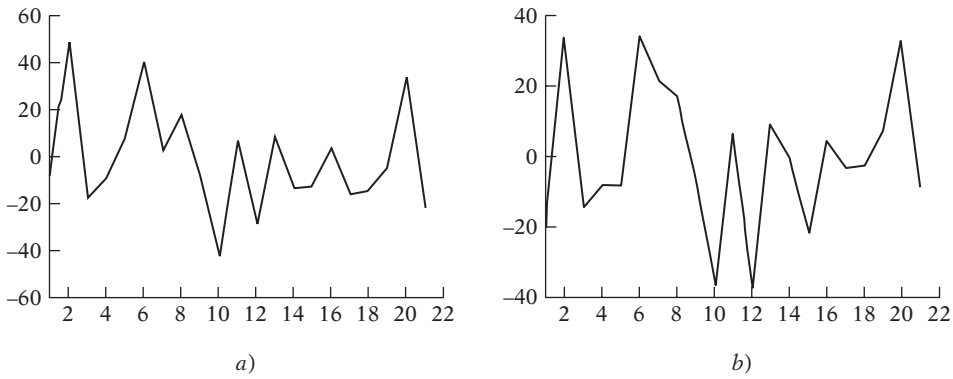
USO DIDÁCTICO

Como un ejemplo concreto, recuerde el estudio de los impactos publicitarios retenidos (Y) en relación con el gasto en publicidad (X). (Consulte el ejercicio 8.32.) Si sólo hace la regresión de Y sobre X y se observan los residuos de esta regresión, se presenta un patrón, pero si se efectúa la regresión de Y sobre X y sobre X^2, se verá otro patrón, lo cual se distingue con claridad en la figura 11.5. Ya comprobamos que X^2 pertenece al modelo. (Consulte el ejercicio 8.32.)

6. Otra fuente de la heteroscedasticidad es la **asimetría** en la distribución de una o más regresoras incluidas en el modelo. Los ejemplos los constituyen variables económicas como ingreso, riqueza y escolaridad. Es bien sabido que la distribución del ingreso y la riqueza en la mayoría de las sociedades es desigual, pues la mayor parte del ingreso o la riqueza le corresponde a unos cuantos individuos pertenecientes a los estratos superiores.

7. Otras fuentes de heteroscedasticidad: como señala David Hendry, la heteroscedasticidad también surge debido a 1) la incorrecta transformación de los datos (por ejemplo, las transformaciones de razón o de primeras diferencias) y 2) una forma funcional incorrecta (por ejemplo, modelos lineales frente a modelos log-lineales).[4]

FIGURA 11.5
Residuos de las regresiones de *a*) impactos publicitarios en el gasto publicitario (GASP) y *b*) impactos publicitarios en GASP y GASP^2.



[4] David F. Hendry, *Dynamic Econometrics,* Oxford University Press, 1995, p. 45.

Observe que el problema de heteroscedasticidad es quizá más común en la información de corte transversal que en la de series de tiempo. En la información de corte transversal, por lo general se trata con miembros de una población en un momento dado, como consumidores individuales o sus familias, empresas, industrias o subdivisiones geográficas, como estados, países, ciudades, etc. Además, estos miembros pueden ser de diferentes tamaños, como empresas pequeñas, medianas o grandes, o ingresos bajos, medios o altos. En las series de tiempo, por el contrario, las variables tienden a ser de órdenes de magnitud similares porque suele recopilarse información sobre el mismo fenómeno o hecho a lo largo de un periodo. Son ejemplos el PIB, el gasto de consumo, el ahorro o el empleo en Estados Unidos, por decir, de 1955 a 2005.

A manera de ilustración sobre la heteroscedasticidad que es posible encontrar en un análisis de corte transversal, considere la tabla 11.1. Esta tabla presenta información sobre salarios por empleado en 10 industrias manufactureras de bienes no duraderos, clasificadas según el número de empleados de la empresa o establecimiento en 1958. En la tabla se presentan además cifras de productividad promedio de nueve clases de empleados.

Aunque las industrias difieran en la composición de su producción, la tabla 11.1 muestra con claridad que, en promedio, las empresas grandes pagan más que las pequeñas. Como ejemplo, las empresas que emplean entre uno y cuatro empleados pagaron, en promedio, sueldos de alrededor de $3 396, mientras que las que emplean entre 1 000 y 2 499 empleados, en promedio, alrededor de $4 843. Pero observe que hay una gran variabilidad en los ingresos entre las diversas clases de empleados, como indican las desviaciones estándar estimadas de los salarios. Esto también se

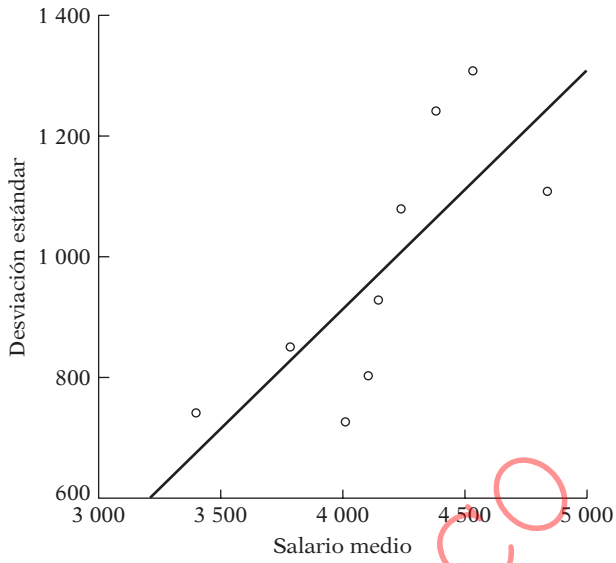
USO DIDÁCTICO

TABLA 11.1 **Salario por empleado (dólares) en la industria de manufactura de bienes no duraderos, de acuerdo con el número de empleados del establecimiento, 1958**

	Tamaño de la plantilla laboral (número promedio de empleados)								
Industria	**1-4**	**5-9**	**10-19**	**20-49**	**50-99**	**100-249**	**250-499**	**500-999**	**1 000-2 499**
Alimentos y similares	2 994	3 295	3 565	3 907	4 189	4 486	4 676	4 968	5 342
Productos del tabaco	1 721	2 057	3 336	3 320	2 980	2 848	3 072	2 969	3 822
Productos textiles	3 600	3 657	3 674	3 437	3 340	3 334	3 225	3 163	3 168
Ropa y productos relacionados	3 494	3 787	3 533	3 215	3 030	2 834	2 750	2 967	3 453
Papel y similares	3 498	3 847	3 913	4 135	4 445	4 885	5 132	5 342	5 326
Impresión y publicación	3 611	4 206	4 695	5 083	5 301	5 269	5 182	5 395	5 552
Productos químicos y similares	3 875	4 660	4 930	5 005	5 114	5 248	5 630	5 870	5 876
Productos petroleros y carboníferos	4 616	5 181	5 317	5 337	5 421	5 710	6 316	6 455	6 347
Productos de caucho y plásticos	3 538	3 984	4 014	4 287	4 221	4 539	4 721	4 905	5 481
Cuero y productos de cuero	3 016	3 196	3 149	3 317	3 414	3 254	3 177	3 346	4 067
Remuneración promedio	3 396	3 787	4 013	4 104	4 146	4 241	4 388	4 538	4 843
Desviación estándar	742.2	851.4	727.8	805.06	929.9	1 080.6	1 241.2	1 307.7	1 110.7
Productividad promedio	9 355	8 584	7 962	8 275	8 389	9 418	9 795	10 281	11 750

Fuente: *The Census of Manufacturers*, U.S. Department of Commerce, 1958 (calculado por el autor).

FIGURA 11.6
Desviación estándar del salario y salario medio.



aprecia en la figura 11.6, que muestra la desviación estándar de los salarios y el salario medio en cada clase de empleados. Como se ve claramente, en promedio, la desviación estándar de los salarios crece con el valor promedio de los salarios.

11.2 Estimación por MCO en presencia de heteroscedasticidad

¿Qué sucede con los estimadores de MCO y sus varianzas si introducimos la heteroscedasticidad permitiendo que $E(u_i^2) = \sigma_i^2$ pero conservamos todos los demás supuestos del modelo clásico? Para responder, recuerde el modelo con dos variables:

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

Al aplicar la fórmula usual, el estimador de MCO de β_2 es

$$\begin{aligned}\hat{\beta}_2 &= \frac{\sum x_i y_i}{\sum x_i^2} \\ &= \frac{n\sum X_i Y_i - \sum X_i \sum Y_i}{n\sum X_i^2 - (\sum X_i)^2}\end{aligned} \tag{11.2.1}$$

pero su varianza está dada ahora por la siguiente expresión (apéndice 11A, sección 11A.1):

$$\operatorname{var}(\hat{\beta}_2) = \frac{\sum x_i^2 \sigma_i^2}{\left(\sum x_i^2\right)^2} \tag{11.2.2}$$

que obviamente difiere de la fórmula usual de varianza obtenida según el supuesto de homoscedasticidad, es decir,

$$\operatorname{var}(\hat{\beta}_2) = \frac{\sigma^2}{\sum x_i^2} \tag{11.2.3}$$

USO DIDÁCTICO

Desde luego, si $\sigma_i^2 = \sigma^2$ para cada i, las dos fórmulas serán idénticas. (¿Por qué?)

Recuerde que $\hat{\beta}_2$ es el mejor estimador lineal e insesgado (MELI) si se mantienen los supuestos del modelo clásico, inclusive el de homoscedasticidad. ¿Seguirá siendo MELI aunque sólo eliminemos el supuesto de homoscedasticidad y lo reemplacemos por el de heteroscedasticidad? Es fácil probar que $\hat{\beta}_2$ es aún lineal e insesgado. En realidad, como se indica en el apéndice 3A, sección 3A.2, para establecer el insesgamiento de $\hat{\beta}_2$ no es necesario que las perturbaciones (u_i) sean homoscedásticas. En realidad, la varianza de u_i, homoscedástica o heteroscedástica, no desempeña papel alguno en la determinación de la propiedad de insesgamiento. Recuerde que en el apéndice 3A, sección 3A.7, vimos que $\hat{\beta}_2$ es un estimador consistente según los supuestos del modelo clásico de regresión lineal. Aunque no lo demostraremos, se puede probar que $\hat{\beta}_2$ es un estimador consistente a pesar de la heteroscedasticidad; es decir, conforme el tamaño de la muestra se incrementa de manera indefinida, el β_2 estimado converge a su valor verdadero. Es más, también puede demostrarse que, en ciertas condiciones (de regulación), $\hat{\beta}_2$ está *distribuida de manera asintótica y normal*. Por supuesto, lo que afirmamos respecto de $\hat{\beta}_2$ también vale para otros parámetros de un modelo de regresión múltiple.

En el supuesto de que $\hat{\beta}_2$ continúe siendo lineal, insesgado y consistente, ¿es "eficiente" o "el mejor"? Es decir, ¿tendrá varianza mínima en la clase de los estimadores lineales e insesgados? ¿Y dicha varianza mínima estará dada por la ecuación (11.2.2)? La respuesta a ambas preguntas es *no*: $\hat{\beta}_2$ deja de ser el mejor y la varianza mínima ya no está dada por (11.2.2). Entonces, ¿cuál estimador es MELI en presencia de heteroscedasticidad? La respuesta se da en la siguiente sección.

USO DIDÁCTICO

11.3 El método de mínimos cuadrados generalizados (MCG)

¿Por qué el estimador usual de MCO de β_2 dado en (11.2.1) no es el mejor, aunque todavía sea insesgado? La razón se intuye en la tabla 11.1. Como se indica ahí, hay una gran variabilidad en los ingresos entre clases de empleados. Si efectuásemos una regresión de salarios por empleado sobre el tamaño de la planta laboral sería recomendable aprovechar el conocimiento de que existe considerable variabilidad entre los salarios de las diferentes clases. Idealmente, querríamos diseñar un esquema de estimación de manera que las observaciones que surgen de poblaciones con mayor variabilidad reciban menos peso que las que provienen de poblaciones con menor variabilidad. Al examinar la tabla 11.1 nos gustaría dar mayor ponderación a las observaciones que provienen de las clases de empleados 10-19 y 20-49 que a las clases de empleados como 5-9 y 250-499, pues las primeras están más concentradas alrededor de sus valores medios que las últimas, lo que permite estimar la FRP en forma más precisa.

Por desgracia, el método de MCO usual no sigue esta estrategia y, por consiguiente, no aprovecha la "información" contenida en la variabilidad desigual de la variable dependiente Y, como sucede con la compensación salarial de los empleados de la tabla 11.1: este método asigna igual peso o importancia a cada observación. Pero existe un método de estimación, conocido como **mínimos cuadrados generalizados** (**MCG**), que toma en cuenta esa información explícitamente y, por consiguiente, es capaz de producir estimadores que son MELI. Para ver cómo se hace, considere el modelo ya familiar con dos variables:

$$Y_i = \beta_1 + \beta_2 X_i + u_i \tag{11.3.1}$$

el cual, para facilitar el reordenamiento algebraico, escribimos como

$$Y_i = \beta_1 X_{0i} + \beta_2 X_i + u_i \tag{11.3.2}$$

donde $X_{0i} = 1$ para cada i. El lector puede ver que estas dos formulaciones son idénticas.

Ahora, suponga que *se conocen* las varianzas heteroscedásticas σ_i^2. Divida ambos lados de (11.3.2) entre σ_i para obtener

$$\frac{Y_i}{\sigma_i} = \beta_1 \left(\frac{X_{0i}}{\sigma_i}\right) + \beta_2 \left(\frac{X_i}{\sigma_i}\right) + \left(\frac{u_i}{\sigma_i}\right) \tag{11.3.3}$$

la cual, para facilidad de exposición, escribimos como

$$Y_i^* = \beta_1^* X_{0i}^* + \beta_2^* X_i^* + u_i^* \tag{11.3.4}$$

en donde las variables con asterisco o transformadas son las variables originales divididas entre σ_i (conocida). Utilizamos la notación β_1^* y β_2^*, los parámetros del modelo transformado, para distinguirlos de los parámetros de MCO usuales β_1 y β_2.

¿Cuál es el propósito de transformar el modelo original? Para ver esto, observe la siguiente característica del término de error transformado u_i^*:

$$\begin{aligned} \operatorname{var}(u_i^*) = E(u_i^*)^2 &= E\left(\frac{u_i}{\sigma_i}\right)^2 && \text{porque } E(u_i^*) = 0 \\ &= \frac{1}{\sigma_i^2} E\left(u_i^2\right) && \text{porque se conoce } \sigma_i^2 \\ &= \frac{1}{\sigma_i^2}\left(\sigma_i^2\right) && \text{porque } E\left(u_i^2\right) = \sigma_i^2 \\ &= 1 \end{aligned} \tag{11.3.5}$$

que es una constante. Es decir, la varianza del término de perturbación transformado u_i^* es ahora homoscedástica. Como aún se conservan los demás supuestos del modelo clásico, el hallazgo de que u^* es homoscedástico sugiere que si se aplican MCO al modelo transformado (11.3.3) se producirán estimadores MELI. En resumen, los β_1^* y β_2^* estimados son ahora MELI y no los estimadores de MCO, $\hat{\beta}_1$ y $\hat{\beta}_2$.

Este procedimiento de transformar las variables originales de forma que las variables transformadas satisfagan los supuestos del modelo clásico y de aplicar luego MCO a ellos se conoce como método de mínimos cuadrados generalizados (MCG). *En resumen, MCG es MCO sobre las variables transformadas que satisfacen los supuestos estándar de mínimos cuadrados.* Los estimadores así obtenidos se conocen como **estimadores de MCG**, y son estos estimadores los que son MELI.

El mecanismo de estimación de β_1^* y β_2^* es el siguiente. Primero escribimos la FRM de (11.3.3)

$$\frac{Y_i}{\sigma_i} = \hat{\beta}_1^* \left(\frac{X_{0i}}{\sigma_i}\right) + \hat{\beta}_2^* \left(\frac{X_i}{\sigma_i}\right) + \left(\frac{\hat{u}_i}{\sigma_i}\right)$$

o

$$Y_i^* = \hat{\beta}_1^* X_{0i}^* + \hat{\beta}_2^* X_i^* + \hat{u}_i^* \tag{11.3.6}$$

Ahora, para obtener los estimadores de MCG, reducimos

$$\sum \hat{u}_i^{2*} = \sum (Y_i^* - \hat{\beta}_1^* X_{0i}^* - \hat{\beta}_2^* X_i^*)^2$$

es decir,

$$\sum \left(\frac{\hat{u}_i}{\sigma_i}\right)^2 = \sum \left[\left(\frac{Y_i}{\sigma_i}\right) - \hat{\beta}_1^* \left(\frac{X_{0i}}{\sigma_i}\right) - \hat{\beta}_2^* \left(\frac{X_i}{\sigma_i}\right)\right]^2 \tag{11.3.7}$$

USO DIDÁCTICO

El mecanismo real para minimizar (11.3.7) sigue las técnicas estándar del cálculo y se presenta en el apéndice 11A, sección 11A.2. Como allí se muestra, el estimador de MCG de β_2^* es

$$\hat{\beta}_2^* = \frac{\left(\sum w_i\right)\left(\sum w_i X_i Y_i\right) - \left(\sum w_i X_i\right)\left(\sum w_i Y_i\right)}{\left(\sum w_i\right)\left(\sum w_i X_i^2\right) - \left(\sum w_i X_i\right)^2} \tag{11.3.8}$$

y su varianza está dada por

$$\operatorname{var}(\hat{\beta}_2^*) = \frac{\sum w_i}{\left(\sum w_i\right)\left(\sum w_i X_i^2\right) - \left(\sum w_i X_i\right)^2} \tag{11.3.9}$$

donde $w_i = 1/\sigma_i^2$.

Diferencia entre MCO y MCG

Recuerde, del capítulo 3, que en MCO reducimos

$$\sum \hat{u}_i^2 = \sum (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)^2 \tag{11.3.10}$$

pero en MCG reducimos la expresión (11.3.7), la cual también se escribe como

$$\sum w_i \hat{u}_i^2 = \sum w_i (Y_i - \hat{\beta}_1^* X_{0i} - \hat{\beta}_2^* X_i)^2 \tag{11.3.11}$$

donde $w_i = 1/\sigma_i^2$ [verifique que (11.3.11) y (11.3.7) sean idénticas].

Por tanto, en MCG se reduce una *suma ponderada de residuos al cuadrado* donde $w_i = 1/\sigma_i^2$ actúan como ponderación, pero en MCO se reduce la SCR sin ponderar o (lo que equivale a lo mismo) con ponderaciones iguales. Como muestra (11.3.7), en MCG, el peso asignado a cada observación es inversamente proporcional a su σ_i, es decir, las observaciones que provienen de una población con una σ_i más grande tendrán una ponderación relativamente menor, y las de una población con un σ_i menor tendrán una ponderación proporcionalmente mayor al reducir la SCR (11.3.11). Para ver claramente la diferencia entre MCO y MCG considere el diagrama hipotético de dispersión de la figura 11.7.

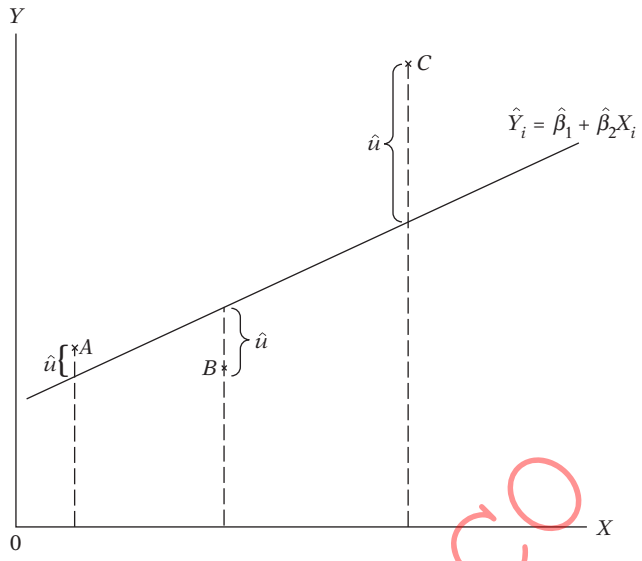
En MCO (sin ponderar), cada $\hat{u}_i^2$ asociada con los puntos A, B y C recibirá el mismo peso al reducir la SCR. Obviamente, en este caso la $\hat{u}_i^2$ asociada al punto C dominará la SCR. Pero en MCG la observación extrema C obtendrá relativamente un peso menor que las otras dos observaciones. Como ya mencionamos, ésta es la estrategia correcta, pues para estimar la función de regresión poblacional (FRP) de una manera más confiable, sería deseable dar más peso a las observaciones agrupadas cerca de su media (poblacional) que a las que están muy dispersas a su alrededor.

Como (11.3.11) reduce una SCR ponderada, esto se conoce apropiadamente como **mínimos cuadrados ponderados** (**MCP**), y los estimadores así obtenidos que aparecen en (11.3.8) y (11.3.9) se conocen como **estimadores de MCP**. Pero MCP es apenas un caso especial de la técnica de estimación más general, MCG. En el contexto de la heteroscedasticidad los términos MCP y MCG se tratan indistintamente. En capítulos posteriores veremos otros casos especiales de MCG.

A propósito, observe que si $w_i = w$, una constante para todas las i, $\hat{\beta}_2^*$ es idéntico a $\hat{\beta}_2$ y var $(\hat{\beta}_2^*)$ es idéntica a la var $(\hat{\beta}_2)$ usual (es decir, homoscedástica) dada en (11.2.3), lo cual no debe sorprender. (¿Por qué?) (Consulte el ejercicio 11.8.)

USO DIDACTICO

FIGURA 11.7
Diagrama de dispersión hipotético.



11.4 Consecuencias de utilizar MCO en presencia de heteroscedasticidad

Como vimos, $\hat{\beta}_2^*$ y $\hat{\beta}_2$ son estimadores (lineales) insesgados: para muestreo repetido, en promedio, $\hat{\beta}_2^*$ y $\hat{\beta}_2$ serán iguales al verdadero β_2, es decir, ambos son estimadores insesgados. Pero sabemos que $\hat{\beta}_2^*$ es el eficiente, es decir, tiene la menor varianza. ¿Qué sucede con los intervalos de confianza, las pruebas de hipótesis y con otros procedimientos si continuamos utilizando el estimador de MCO, $\hat{\beta}_2$? Se distinguen dos situaciones.

Estimación por MCO con heteroscedasticidad

Suponga que utilizamos $\hat{\beta}_2$ y la fórmula de varianza dada en (11.2.2), la cual considera explícitamente la heteroscedasticidad. Con esta varianza y la suposición de que se conocen las σ_i^2, ¿es posible establecer intervalos de confianza y probar hipótesis con las pruebas t y F usuales? La respuesta suele ser no, pues puede demostrarse que var $(\hat{\beta}_2^*) \leq$ var $(\hat{\beta}_2)$,[5] lo cual significa que los intervalos de confianza basados en estos últimos serán innecesariamente grandes. Como resultado, es probable que las pruebas t y F den resultados imprecisos en el sentido de que la var $(\hat{\beta}_2)$ es demasiado grande, y lo que parece un coeficiente estadísticamente no significativo (pues el valor t es más bajo de lo apropiado), de hecho puede resultar significativo si se establecen intervalos de confianza correctos con base en el procedimiento de MCG.

Estimación por MCO sin heteroscedasticidad

La situación se torna muy grave si, además de $\hat{\beta}_2$, también se sigue utilizando la fórmula habitual de varianza (homoscedástica) dada en (11.2.3), aunque exista heteroscedasticidad o se sospeche su existencia: observe que éste es el caso más probable de los dos que aquí se analizan, pues al hacer una regresión estándar por MCO e ignorar (o no conocer) la existencia de la heteroscedasticidad se producirá una varianza de $\hat{\beta}_2$ como la dada en (11.2.3). En primer lugar, la var $(\hat{\beta}_2)$ dada en (11.2.3) es un estimador *sesgado* de var $(\hat{\beta}_2)$ dada en (11.2.2), es decir, en promedio, ésta

[5] Una prueba formal se encuentra en Phoebus J. Dhrymes, *Introductory Econometrics,* Springer-Verlag, Nueva York, 1978, pp. 110-111. A propósito, observe que la pérdida de eficiencia de $\hat{\beta}_2$ [es decir, la medida en que excede var $[\hat{\beta}_2]$ a var $[\hat{\beta}_2^*]$ depende de los valores muestrales de las variables X y del valor de σ_i^2.

USO DIDÁCTICO

sobreestima o subestima la última y *en general* no se puede decir si el sesgo es positivo (sobreestimación) o negativo (subestimación), pues éste depende de la naturaleza de la relación entre σ_i^2 y los valores tomados por la variable explicativa X, como se ve claramente en (11.2.2) (véase el ejercicio 11.9). El sesgo surge de que $\hat{\sigma}^2$, el estimador convencional de σ^2, a saber, $\sum \hat{u}_i^2/(n-2)$, deja de ser un estimador insesgado del último en presencia de heteroscedasticidad (véase el apéndice 11A.3). Como resultado, ya no es posible depender de los intervalos de confianza calculados convencionalmente ni de las pruebas t y F tradicionales.[6] **En resumen, si insistimos en los procedimientos de prueba usuales a pesar de la presencia de heteroscedasticidad, las conclusiones o inferencias que obtengamos pueden ser muy equivocadas**.

Para dar mayor claridad a este tema, veamos un estudio **Monte Carlo** de Davidson y MacKinnon.[7] Consideran el siguiente modelo simple, que en nuestra notación es

$$Y_i = \beta_1 + \beta_2 X_i + u_i \tag{11.4.1}$$

Suponen que $\beta_1 = 1$, $\beta_2 = 1$ y $u_i \sim N(0, X_i^\alpha)$. Como indica la última expresión, suponen que la varianza del error es heteroscedástica y que se relaciona con el valor de la regresora X elevado a la potencia α. Si, por ejemplo, $\alpha = 1$, la varianza del error es proporcional al valor de X; si $\alpha = 2$, la varianza del error es proporcional al cuadrado del valor de X, y así sucesivamente. En la sección 11.6 veremos la lógica de tal procedimiento. Basados en 20 000 repeticiones y con diversos valores para α, obtienen los errores estándar de los dos coeficientes de regresión con MCO [consulte la ecuación (11.2.3)], MCO con heteroscedasticidad [ecuación (11.2.2)] y MCG [ecuación (11.3.9)]. Sus resultados para valores seleccionados de α son:

	Error estándar de $\hat{\beta}_1$			Error estándar de $\hat{\beta}_2$		
Valor de α	**MCO**	**MCO$_{\text{het}}$**	**MCG**	**MCO**	**MCO$_{\text{het}}$**	**MCG**
0.5	0.164	0.134	0.110	0.285	0.277	0.243
1.0	0.142	0.101	0.048	0.246	0.247	0.173
2.0	0.116	0.074	0.0073	0.200	0.220	0.109
3.0	0.100	0.064	0.0013	0.173	0.206	0.056
4.0	0.089	0.059	0.0003	0.154	0.195	0.017

Nota: MCO$_{\text{het}}$ significa MCO con heteroscedasticidad.

La característica más sobresaliente de estos resultados es que los MCO, con o sin corrección por heteroscedasticidad, sobreestiman consistentemente el verdadero error estándar obtenido mediante el procedimiento (correcto) de MCG, en especial para valores grandes de α, con lo cual se establece la superioridad de MCG. Estos resultados también muestran que si no se utilizan MCG y se depende de MCO —permitiendo o no la heteroscedasticidad—, el resultado es una mezcla. Los errores estándar de MCO convencionales son muy grandes (para el intercepto) o *generalmente* muy bajos (para el coeficiente de la pendiente) en relación con los obtenidos por MCO que permite la heteroscedasticidad. El mensaje es claro: ante la presencia de heteroscedasticidad, utilice MCG. Sin embargo, por razones explicadas más adelante en el capítulo, en la práctica no siempre es fácil aplicar MCG. Asimismo, como veremos más adelante, a menos que la heteroscedasticidad sea muy grave, no se abandonarían los MCO en favor de los MCG o de los MCP.

Por el análisis anterior, es claro que la heteroscedasticidad es un problema potencialmente grave y el investigador debe saber si está presente en una situación dada. Si detecta su presen-

[6] De (5.3.6) sabemos que el intervalo de confianza a $100(1-\alpha)\%$ para β_2 es $[\hat{\beta}_2 \pm t_{\alpha/2}\,\text{ee}(\hat{\beta}_2)]$. Pero, si $\text{ee}(\hat{\beta}_2)$ no puede estimarse en forma insesgada, ¿qué confianza se puede tener en el intervalo de confianza calculado convencionalmente?

[7] Russell Davidson y James G. MacKinnon, *Estimation and Inference in Econometrics,* Oxford University Press, Nueva York, 1993, pp. 549-550.

USO DIDÁCTICO

cia, puede tomar acciones correctivas, como una regresión de mínimos cuadrados ponderados o alguna otra técnica. Sin embargo, antes de examinar los procedimientos correctivos, es preciso averiguar si hay heteroscedasticidad o si es probable en un caso dado. Analizamos este tema en la siguiente sección.

Nota técnica

Aunque ya establecimos que, en caso de heteroscedasticidad, son los MCG y no los MCO los que son MELI, existen ejemplos en los que los MCO pueden ser MELI a pesar de la heteroscedasticidad.[8] No obstante, dichos casos son poco frecuentes en la práctica.

11.5 Detección de la heteroscedasticidad

Como con la multicolinealidad, para efectos prácticos la pregunta importante es ¿cómo saber que hay heteroscedasticidad en una situación específica? De nuevo, como en la multicolinealidad, no existen reglas precisas y rápidas para detectar la heteroscedasticidad, sólo algunas reglas prácticas. Pero esta situación es inevitable porque σ_i^2 solamente puede conocerse si se tiene toda la población Y, correspondiente a las X seleccionadas, como la población presentada en la tabla 2.1 o en la tabla 11.1. Pero tal información es una excepción más que la regla en la mayoría de las investigaciones económicas. Al respecto, el econometrista difiere de los científicos en campos como agricultura y biología, donde los investigadores tienen gran parte del control sobre sus temas. En los estudios de economía, es frecuente que sólo haya un valor muestral Y correspondiente a un valor particular de X. Por consiguiente, no hay forma de conocer σ_i^2 a partir de una sola observación Y. Así, en la mayoría de las investigaciones econométricas, la heteroscedasticidad puede ser un asunto de intuición, de conjeturas refinadas, de un trabajo basado en experiencia empírica previa o de pura especulación.

Tras la advertencia anterior, podemos examinar algunos métodos informales y formales para detectar la heteroscedasticidad. Como revelará el siguiente análisis, la mayoría de estos métodos se basan en el examen de los residuos $\hat{u}_i$ de MCO, pues son éstos los que se observan y no las perturbaciones u_i. Se espera que $\hat{u}_i$ sean buenas estimaciones de u_i, esperanza que se cumple si el tamaño de la muestra es lo bastante grande.

Métodos informales

Naturaleza del problema

Con mucha frecuencia la naturaleza del problema en consideración sugiere la posibilidad de heteroscedasticidad. Por ejemplo, a partir del trabajo pionero de Prais y Houthakker sobre estudios de presupuesto familiar, en el cual hallaron que la varianza residual correspondiente a la regresión del consumo sobre el ingreso aumentaba con el ingreso, hoy en día generalmente se supone que en encuestas similares se pueden esperar varianzas desiguales entre las perturbaciones.[9] De hecho, en la información de corte transversal que comprende unidades heterogéneas, la heteroscedasticidad puede ser la regla y no la excepción. Así, en el análisis de corte transversal que relaciona el gasto de inversión con las ventas, la tasa de interés, etc., suele esperarse la presencia de heteroscedasticidad si se agrupan empresas pequeñas, medianas y grandes.

USO DIDÁCTICO

[8] La razón reside en que el teorema de Gauss-Markov proporciona la condición suficiente (pero no necesaria) para que los MCO sean eficientes. La condición suficiente y necesaria para que los MCO sean MELI la establece el **teorema de Kruskal**. Pero este tema está más allá del alcance de este libro. El autor está en deuda con Michael McAleer por notar esto. Para mayores detalles, véase Denzil G. Fiebig, Michael McAleer y Robert Bartels, "Properties of Ordinary Least Squares Estimators in Regression Models with Nonspherical Disturbances", *Journal of Econometrics,* vol. 54, núm. 1-3, octubre-diciembre de 1992, pp. 321-334. Para los estudiantes con inclinaciones matemáticas, analizo este tema en más detalle en el **apéndice C** con álgebra matricial.

[9] S.J. Prais y H.S. Houthakker, *The Analysis of Family Budgets,* Cambridge University Press, Nueva York, 1955.

De hecho, ya presentamos ejemplos de lo anterior. En el capítulo 2 analizamos la relación entre los salarios medios o promedios por hora en relación con los años de escolaridad en Estados Unidos. En este capítulo también examinamos la relación entre el gasto en alimentos y el gasto total de 55 familias en India (ejercicio 11.16).

Método gráfico

Si no hay información *a priori* o empírica sobre la naturaleza de la heteroscedasticidad, en la práctica se puede llevar a cabo un análisis de regresión con el supuesto de que no hay heteroscedasticidad y luego hacer un examen *post mortem* de los residuos elevados al cuadrado, $\hat{u}_i^2$, para ver si exhiben algún patrón sistemático. Aunque los $\hat{u}_i^2$ no son lo mismo que los u_i^2, los primeros sirven como representantes de los últimos sobre todo si el tamaño de la muestra es lo bastante grande.[10] Un examen de los $\hat{u}_i^2$ puede revelar patrones como los de la figura 11.8.

En la figura 11.8 se grafican los $\hat{u}_i^2$ frente a los $\hat{Y}_i$, que son los Y_i estimados mediante la línea de regresión, con la idea de averiguar si el valor medio estimado de Y está relacionado sistemáticamente con el residuo al cuadrado. En la figura 11.8*a*) se ve que no hay un patrón sistemático entre las dos variables, lo cual sugiere que tal vez no haya heteroscedasticidad en los datos. Sin embargo, las figuras 11.8*b*) a 11.8*e*) muestran patrones definidos. Por ejemplo, la figura 11.8*c*) sugiere una relación lineal, mientras que las figuras 11.8*d*) y 11.8*e*) indican una relación cuadrática entre $\hat{u}_i^2$ y $\hat{Y}_i$. Con tal conocimiento, aunque informal, es posible transformar los datos de manera que, una vez transformados, no presenten heteroscedasticidad. En la sección 11.6 examinaremos transformaciones de este tipo.

En lugar de graficar los $\hat{u}_i^2$ frente a los $\hat{Y}_i$, se pueden graficar frente a una de las variables explicativas, en especial si el gráfico de $\hat{u}_i^2$ frente a $\hat{Y}_i$ presenta un patrón como el de la figura 11.8*a*). Tal gráfico, que aparece en la figura 11.9, puede revelar patrones similares a los de la figura 11.8. (En el caso del modelo con dos variables, el gráfico de $\hat{u}_i^2$ frente a los $\hat{Y}_i$ equivale a graficar el

USO DIDÁCTICO

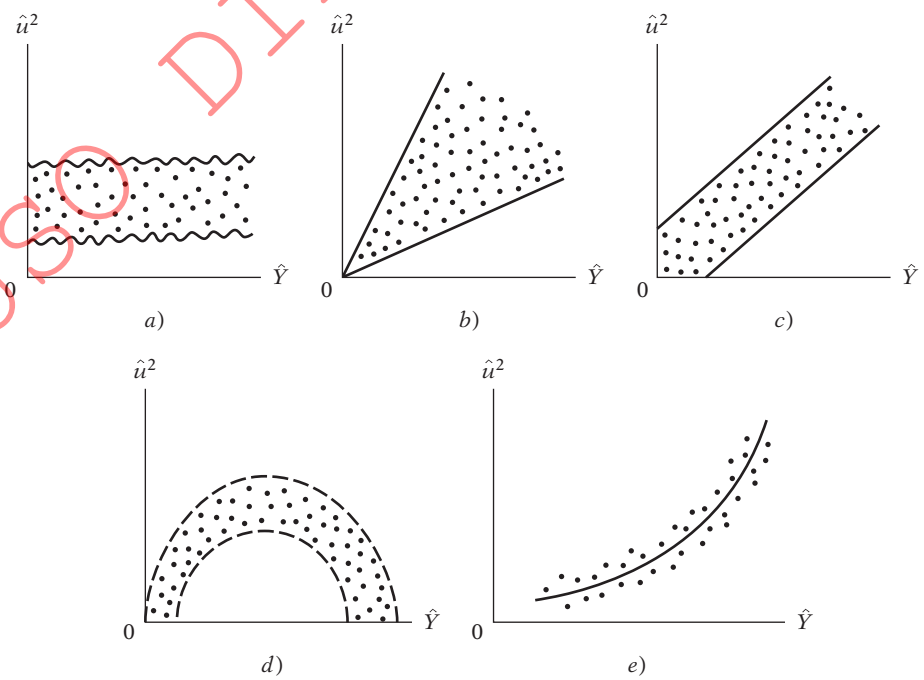
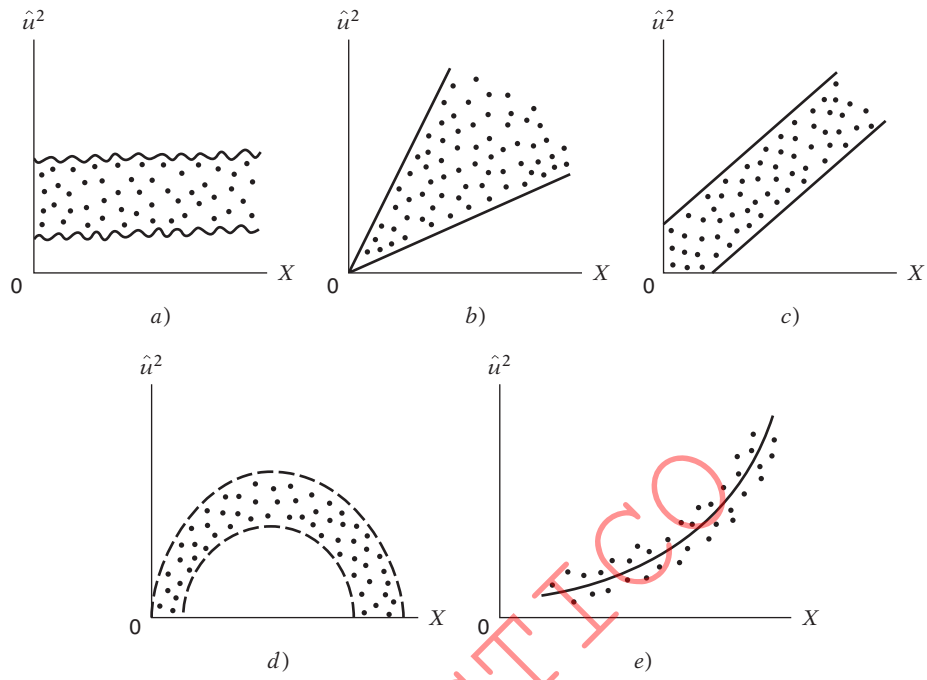


FIGURA 11.8
Patrones hipotéticos de los residuos estimados al cuadrado.

[10] Respecto de la relación entre $\hat{u}_i$ y u_i, véase E. Malinvaud, *Statistical Methods of Econometrics*, North Holland, Amsterdam, 1970, pp. 88-89.

FIGURA 11.9
Diagrama de dispersión de los residuos estimados al cuadrado frente a X.



primero frente a X_i, razón por la cual la figura 11.9 es similar a la 11.8. Pero ésta no es la situación cuando se considera un modelo con dos o más variables X; en este caso, $\hat{u}_i^2$ puede graficarse frente a cualquier variable X incluida en el modelo.)

Un patrón como el de la figura 11.9*c*), por ejemplo, sugiere que la varianza del término de perturbación se relaciona linealmente con la variable X. Así, si en la regresión de ahorro sobre ingreso se encuentra un patrón como el de la figura 11.9*c*), éste sugiere que la varianza heteroscedástica puede ser *proporcional* al valor de la variable ingreso. Saber esto puede contribuir a transformar los datos de manera que, en la regresión sobre los datos transformados, la varianza de las perturbaciones sea homoscedástica. Trataremos este tema en la siguiente sección.

USO DIDÁCTICO

Métodos formales

Prueba de Park[11]

Park formaliza el método gráfico con la sugerencia de que σ_i^2 es algún tipo de función de la variable explicativa X_i. La forma funcional fue

$$\sigma_i^2 = \sigma^2 X_i^{\beta} e^{vi}$$

o

$$\ln \sigma_i^2 = \ln \sigma^2 + \beta \ln X_i + v_i \qquad \textbf{(11.5.1)}$$

donde v_i es el término de perturbación estocástico.

[11] R.E. Park, "Estimation with Heteroscedastic Error Terms", *Econometrica*, vol. 34, núm. 4, octubre de 1966, p. 888. La prueba de Park es un caso especial de la prueba general propuesta por A. C. Harvey en "Estimating Regression Models with Multiplicative Heteroscedasticity", *Econometrica*, vol. 44, núm. 3, 1976, pp. 461-465.

Como σ_i^2 por lo general no se conoce, Park sugiere utilizar $\hat{u}_i^2$ como aproximación y correr la siguiente regresión:

$$\begin{aligned} \ln \hat{u}_i^2 &= \ln \sigma^2 + \beta \ln X_i + v_i \\ &= \alpha + \beta \ln X_i + v_i \end{aligned} \tag{11.5.2}$$

Si β resulta estadísticamente significativo, esto sugerirá heteroscedasticidad en los datos. Si resulta no significativo, podemos aceptar el supuesto de homoscedasticidad. La prueba de Park es, por tanto, un procedimiento de dos etapas. En la primera se efectúa la regresión MCO ignorando el interrogante de la heteroscedasticidad. Se obtiene $\hat{u}_i$ de esta regresión y luego, en la segunda etapa, se efectúa la regresión (11.5.2).

Aunque empíricamente la prueba de Park es atractiva, presenta algunos problemas. Goldfeld y Quandt argumentan que el término de error v_i que entra en (11.5.2) puede no satisfacer los supuestos de MCO y en sí mismo ser heteroscedástico.[12] No obstante, es posible utilizar la prueba de Park como método estrictamente exploratorio.

EJEMPLO 11.1
Relación entre salarios y productividad

Para ilustrar el enfoque de Park, con la información de la tabla 11.1 efectuamos la siguiente regresión:

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

donde Y = salario promedio en miles de dólares, X = productividad promedio en miles de dólares e i = i-ésimo de la planta laboral del establecimiento. Los resultados de la regresión fueron los siguientes:

$$\begin{aligned} \hat{Y}_i &= 1992.3452 + 0.2329X_i \\ \text{ee} &= (936.4791) \quad (0.0998) \\ t &= \quad (2.1275) \quad (2.333) \qquad R^2 = 0.4375 \end{aligned} \tag{11.5.3}$$

Los resultados revelan que el coeficiente de pendiente estimado es significativo en el nivel de 5% con base en una prueba t de una cola. La ecuación muestra que, a medida que aumenta la productividad laboral, por ejemplo, en un dólar, el salario aumenta, en promedio, alrededor de 23 centavos de dólar.

En los residuos obtenidos de la regresión (11.5.3) se hizo la regresión sobre X_i como lo sugiere la ecuación (11.5.2), con los siguientes resultados:

$$\begin{aligned} \widehat{\ln u_i^2} &= 35.817 - 2.8099 \ln X_i \\ \text{ee} &= (38.319) \quad (4.216) \\ t &= \quad (0.934) \quad (-0.667) \qquad R^2 = 0.0595 \end{aligned} \tag{11.5.4}$$

Obvio, no hay una relación estadísticamente significativa entre ambas variables. Según la prueba de Park, se puede concluir que no hay heteroscedasticidad en la varianza del error.[13]

Prueba de Glejser[14]

La prueba de Glejser en esencia es similar a la de Park. Después de obtener los residuos $\hat{u}_i$ de la regresión MCO, Glejser sugiere una regresión sobre los valores absolutos de $\hat{u}_i$ sobre la variable

[12] Stephen M. Goldfeld y Richard E. Quandt, *Nonlinear Methods in Econometrics,* North Holland, Amsterdam, 1972, pp. 93-94.

[13] La forma funcional particular seleccionada por Park es sólo una sugerencia. Una forma funcional diferente revela relaciones significativas. Por ejemplo, se puede utilizar $\hat{u}_i^2$ en lugar de ln $\hat{u}_i^2$ como variable dependiente.

[14] H. Glejser, "A New Test for Heteroscedasticity", *Journal of the American Statistical Association,* vol. 64, 1969, pp. 316-323.

USO DIDÁCTICO

X que se cree muy asociada con σ_i^2. En sus experimentos, Glejser utilizó las siguientes formas funcionales:

$$|\hat{u}_i| = \beta_1 + \beta_2 X_i + v_i$$

$$|\hat{u}_i| = \beta_1 + \beta_2 \sqrt{X_i} + v_i$$

$$|\hat{u}_i| = \beta_1 + \beta_2 \frac{1}{X_i} + v_i$$

$$|\hat{u}_i| = \beta_1 + \beta_2 \frac{1}{\sqrt{X_i}} + v_i$$

$$|\hat{u}_i| = \sqrt{\beta_1 + \beta_2 X_i} + v_i$$

$$|\hat{u}_i| = \sqrt{\beta_1 + \beta_2 X_i^2} + v_i$$

donde v_i es el término de error.

De nuevo, como un asunto empírico o práctico, se puede utilizar el método de Glejser. Sin embargo, Goldfeld y Quandt señalan que el término de error v_i tiene algunos problemas, pues su valor esperado es diferente de cero, está serialmente correlacionado (véase el capítulo 12) e irónicamente es heteroscedástico.[15] Otra dificultad del método Glejser es que los modelos como

$$|\hat{u}_i| = \sqrt{\beta_1 + \beta_2 X_i} + v_i$$

y

$$|\hat{u}_i| = \sqrt{\beta_1 + \beta_2 X_i^2} + v_i$$

no son lineales en los parámetros y, por consiguiente, no pueden estimarse mediante el procedimiento de MCO habitual.

Glejser descubrió que para muestras grandes, los cuatro primeros modelos suelen dar resultados satisfactorios en la detección de la heteroscedasticidad. En la práctica, por consiguiente, la técnica de Glejser es útil para muestras grandes, y en muestras pequeñas sirve estrictamente como herramienta cualitativa para obtener una noción sobre la heteroscedasticidad.

EJEMPLO 11.2
Relación entre el salario y la productividad: prueba de Glejser

Aún con el ejemplo 11.1, se hizo la regresión del valor absoluto de los residuos obtenidos de la regresión (11.5.3) sobre la productividad promedio (X), lo cual dio los siguientes resultados:

$$\begin{aligned} \widehat{|\hat{u}_i|} &= 407.2783 - 0.0203X_i \\ \text{ee} &= (633.1621) \quad (0.0675) \qquad r^2 = 0.0127\text{e} \\ t &= (0.6432) \quad (-0.3012) \end{aligned} \tag{11.5.5}$$

Como se aprecia de esta regresión, no hay relación entre el valor absoluto de los residuos y la regresora, la productividad promedio. Lo anterior refuerza la conclusión basada en la prueba de Park.

Prueba de correlación de orden de Spearman

En el ejercicio 3.8 definimos el coeficiente de correlación de orden de Spearman como

$$r_s = 1 - 6\left[\frac{\sum d_i^2}{n(n^2-1)}\right] \tag{11.5.6}$$

[15] Véanse detalles en Goldfeld y Quandt, *op. cit.*, capítulo 3.

USO DIDÁCTICO

donde d_i = la diferencia en las posiciones o lugares asignados al i-ésimo individuo o fenómeno respecto de dos características y n = número de individuos o fenómenos ordenados. Con el coeficiente de correlación de orden anterior se detecta heteroscedasticidad de la siguiente manera: Suponga que $Y_i = \beta_0 + \beta_1 X_i + u_i$.

Paso 1. Ajuste la regresión a los datos sobre Y y X, y obtenga los residuos $\hat{u}_i$.

Paso 2. Ignore el signo de $\hat{u}_i$, es decir, tome su valor absoluto $|\hat{u}_i|$, y ordene los valores $|\hat{u}_i|$ y X_i (o $\hat{Y}_i$) de acuerdo con un orden ascendente o descendente, y calcule el coeficiente de correlación de orden de Spearman dado antes.

Paso 3. Si supone que el coeficiente poblacional de correlación de orden ρ_s es cero y $n > 8$, la significancia del r_s muestral se prueba mediante la prueba t de la siguiente manera:[16]

$$t = \frac{r_s\sqrt{n-2}}{\sqrt{1-r_s^2}} \tag{11.5.7}$$

con gl $= n - 2$.

Si el valor t calculado excede el valor t crítico, podemos aceptar la hipótesis de heteroscedasticidad; de lo contrario, podemos rechazarla. Si el modelo de regresión considera más de una variable X, r_s se calcula entre $|\hat{u}_i|$ y cada variable X por separado, y la significancia estadística se somete a la prueba t dada en la ecuación (11.5.7).

EJEMPLO 11.3
Ilustración de la prueba de correlación de orden

Para ilustrar la prueba de correlación de orden, considere los datos de la tabla 11.2, que corresponden al rendimiento anual promedio (E_i, %) y la desviación estándar del rendimiento anual (σ_i,%) de 10 fondos de inversión.

TABLA 11.2 Prueba de correlación de orden para heteroscedasticidad

Nombre del fondo mutualista	E_i, rendimiento promedio anual, %	σ_i, desviación estándar del rendimiento anual, %	$\hat{E}_i$†	$\lvert\hat{u}_i\rvert$‡, residuos $\lvert(E_i - \hat{E}_i)\rvert$	Ordenación de $\lvert\hat{u}_i\rvert$	Ordenación de σ_i	d, diferencia entre las dos ordenaciones	d^2
Boston Fund	12.4	12.1	11.37	1.03	9	4	5	25
Delaware Fund	14.4	21.4	15.64	1.24	10	9	1	1
Equity Fund	14.6	18.7	14.40	0.20	4	7	–3	9
Fundamental Investors	16.0	21.7	15.78	0.22	5	10	–5	25
Investors Mutual	11.3	12.5	11.56	0.26	6	5	1	1
Loomis-Sales Mutual Fund	10.0	10.4	10.59	0.59	7	2	5	25
Massachusetts Investors Trust	16.2	20.8	15.37	0.83	8	8	0	0
New England Fund	10.4	10.2	10.50	0.10	3	1	2	4
Putnam Fund of Boston	13.1	16.0	13.16	0.06	2	6	–4	16
Wellington Fund	11.3	12.0	11.33	0.03	1	3	–2	4
Total							0	110

† Obtenido de la regresión: $\hat{E}i = 5.8194 + 0.4590\,\sigma_i$.
‡ Valor absoluto de los residuales.
Nota: La ordenación de valores es ascendente.

(*continúa*)

[16] Véase G. Udny Yule y M.G. Kendall, *An Introduction to the Theory of Statistics,* Charles Griffin & Company, Londres, 1953, p. 455.

USO DIDÁCTICO

EJEMPLO 11.3
(*continuación*)

La línea del mercado de capitales (LMC) de la teoría de portafolios postula una relación lineal entre el rendimiento esperado (E_i) y el riesgo (como se mide mediante la desviación estándar, σ) de un portafolio, de la siguiente manera:

$$E_i = \beta_i + \beta_2 \sigma_i$$

Con los datos de la tabla 11.2 se estimó el modelo anterior, a partir del cual se calcularon los residuos. Como los datos se relacionan con 10 fondos mutualistas de distintos tamaños y metas de inversión, se podría esperar *a priori* que hubiera heteroscedasticidad. Para probar esta hipótesis, aplicamos la prueba de correlación de orden. Los cálculos necesarios se proporcionan en la tabla 11.2.

Con la fórmula (11.5.6) obtenemos

$$r_s = 1 - 6\frac{110}{10(100-1)} \qquad \textbf{(11.5.8)}$$
$$= 0.3333$$

Con la prueba t dada en (11.5.7) obtenemos

$$t = \frac{(0.3333)(\sqrt{8})}{\sqrt{1-0.1110}} \qquad \textbf{(11.5.9)}$$
$$= 0.9998$$

Para 8 gl, este valor t no es significativo ni siquiera en el nivel de significancia de 10%; el valor p es 0.17. Por tanto, no hay evidencia de una relación sistemática entre la variable explicativa y los valores absolutos de los residuos, lo cual sugeriría que no hay heteroscedasticidad.

Prueba de Goldfeld-Quandt[17]

Este popular método es aplicable si se supone que la varianza heteroscedástica, σ_i^2, está relacionada positivamente con *una* de las variables explicativas en el modelo de regresión. Por simplicidad, considere el modelo usual con dos variables:

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

Suponga que σ_i^2 está relacionado positivamente con X_i, en la forma

$$\sigma_i^2 = \sigma^2 X_i^2 \qquad \textbf{(11.5.10)}$$

donde σ^2 es una constante.[18]

El supuesto (11.5.10) postula que σ_i^2 es proporcional al cuadrado de la variable X. En su estudio de presupuestos familiares, Prais y Houthakker encontraron muy útil ese supuesto. (Véase la sección 11.5, métodos informales.)

Si (11.5.10) es la relación apropiada, significaría que σ_i^2 sería mayor mientras mayores fueran los valores de X_i. Si éste resulta ser el caso, es muy probable que haya heteroscedasticidad en el modelo. Para probar esto explícitamente, Goldfeld y Quandt sugieren los siguientes pasos:

Paso 1. Ordene las observaciones de acuerdo con los valores de X_i, a partir del valor más bajo de X.

Paso 2. Omita las c observaciones centrales, donde c se especificó *a priori,* y divida las observaciones restantes $(n - c)$ en dos grupos, cada uno de $(n - c)/2$ observaciones.

Paso 3. Ajuste regresiones MCO separadas a las primeras $(n - c)/2$ observaciones y a las últimas $(n - c)/2$ observaciones, y obtenga las respectivas sumas de cuadrados residuales

[17] Goldfeld y Quandt, *op. cit.*, capítulo 3.

[18] Éste es sólo un supuesto plausible. En realidad lo que se requiere es que σ_i^2 esté monotónicamente relacionada con X_i.

USO DIDACTICO

SCR_1 y SCR_2; SCR_1 representa la SCR de la regresión correspondiente a los valores más bajos de X_i (el grupo de varianza pequeña), y SCR_2, a los valores más grandes de X_i (el grupo de varianza grande). Cada SCR tiene

$$\frac{(n-c)}{2} - k \qquad \text{o} \qquad \left(\frac{n-c-2k}{2}\right) \text{gl}$$

donde k es el número de parámetros que deben estimarse, inclusive el intercepto. (¿Por qué?) Sin duda, para el caso de dos variables, k es 2.

Paso 4. Calcule la razón

$$\lambda = \frac{SCR_2/\text{gl}}{SCR_1/\text{gl}} \tag{11.5.11}$$

Si supusimos que las u_i están normalmente distribuidas (lo cual suele hacerse), y *si el supuesto de homoscedasticidad es válido,* entonces se demuestra que λ de (11.5.10) sigue la distribución F con un número de gl en el numerador y uno en el denominador iguales a $(n - c - 2k)/2$.

Si en una aplicación λ $(= F)$ calculada es superior al F crítico en el nivel de significancia seleccionado, podemos rechazar la hipótesis de homoscedasticidad, es decir, podemos afirmar que la heteroscedasticidad es muy probable.

Antes de ilustrar la prueba, conviene explicar la omisión de las observaciones centrales c. Estas observaciones se omiten para agudizar o acentuar la diferencia entre el grupo de varianza pequeña (es decir, SCR_1) y el grupo de varianza grande (es decir, SCR_2). Pero la capacidad de la prueba Goldfeld-Quandt para lograrlo depende de la forma de seleccionar c.[19] Para el modelo con dos variables, los experimentos Monte Carlo realizados por Goldfeld y Quandt sugieren que c sea alrededor de 8 si el tamaño de la muestra es alrededor de 30, y alrededor de 16 si el tamaño de la muestra es alrededor de 60. Sin embargo, Judge *et al.*, encontraron satisfactorios en la práctica los niveles de $c = 4$ si $n = 30$ y $c = 10$ si n es alrededor de 60.[20]

Antes de proseguir, cabe notar que, en caso de que haya más de una variable X en el modelo, el ordenamiento de las observaciones, que es el primer paso en la prueba, puede hacerse de acuerdo con cualquiera de ellas. Por tanto, en el modelo: $Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + u_i$ se pueden ordenar los datos de acuerdo con cualquiera de estas X. Si, *a priori*, no hay seguridad sobre cuál variable X es la adecuada, realice la prueba sobre cada variable X o aplique la prueba de Park, por turnos, sobre cada X.

EJEMPLO 11.4
La prueba de Goldfeld-Quandt

Para ilustrar la prueba de Goldfeld-Quandt presentamos en la tabla 11.3 información sobre el gasto de consumo en relación con el ingreso de una muestra transversal de 30 familias. Suponga que postulamos que el gasto de consumo está relacionado linealmente con el ingreso pero que hay heteroscedasticidad en los datos. Postulamos además que la naturaleza de la heteroscedasticidad es como la de (11.5.10). En la tabla 11.3 presentamos también el reordenamiento necesario de los datos para aplicar la prueba.

Al eliminar las 4 observaciones del medio, las regresiones de MCO basadas en las primeras 13 observaciones y en las últimas 13, y sus sumas de cuadrados residuales asociadas se presentan a continuación (los errores estándar se indican entre paréntesis).

(*continúa*)

[19] Técnicamente, la **potencia** de la prueba depende de la forma de seleccionar c. En estadística, la *potencia* de una *prueba* se mide por la probabilidad de rechazar la hipótesis nula cuando ésta es falsa [es decir, 1 − Prob(error tipo II)]. Aquí la hipótesis nula es que las varianzas de los dos grupos son las mismas, es decir, que hay homoscedasticidad. Para mayor análisis, véase M. M. Ali y C. Giaccotto, "A Study of Several New and Existing Tests for Heteroscedasticity in the General Linear Model", *Journal of Econometrics,* vol. 26, 1984, pp. 355-373.

[20] George G. Judge, R. Carter Hill, William E. Griffiths, Helmut Lütkepohl y Tsoung-Chao Lee, *Introduction to the Theory and Practice of Econometrics,* John Wiley & Sons, Nueva York, 1982, p. 422.

USO DIDACTICO

EJEMPLO 11.4
(*continuación*)

TABLA 11.3 **Datos hipotéticos sobre el gasto de consumo *Y*($) y el ingreso *X*($) para ilustrar la prueba de Goldfeld-Quandt**

		Datos ordenados de acuerdo con los valores *X*		
Y	*X*	*Y*	*X*	
55	80	55	80	
65	100	70	85	
70	85	75	90	
80	110	65	100	
79	120	74	105	
84	115	80	110	
98	130	84	115	
95	140	79	120	
90	125	90	125	
75	90	98	130	
74	105	95	140	
110	160	108	145	
113	150	113	150	
125	165	110	160	4 observaciones del medio
108	145	125	165	
115	180	115	180	
140	225	130	185	
120	200	135	190	
145	240	120	200	
130	185	140	205	
152	220	144	210	
144	210	152	220	
175	245	140	225	
180	260	137	230	
135	190	145	240	
140	205	175	245	
178	265	189	250	
191	270	180	260	
137	230	178	265	
189	250	191	270	

Regresión basada en las primeras 13 observaciones:

$$\hat{Y}_i = 3.4094 + 0.6968X_i$$
$$(8.7049) \quad (0.0744) \qquad r^2 = 0.8887 \qquad \text{SCR}_1 = 377.17 \qquad \text{gl} = 11$$

Regresión basada en las últimas 13 observaciones:

$$\hat{Y}_i = -28.0272 + 0.7941X_i$$
$$(30.6421) \quad (0.1319) \qquad r^2 = 0.7681 \qquad \text{SCR}_2 = 1\,536.8 \qquad \text{gl} = 11$$

De estos resultados obtenemos

$$\lambda = \frac{\text{SCR}_2/\text{gl}}{\text{SCR}_1/\text{gl}} = \frac{1\,536.8/11}{377.17/11}$$
$$\lambda = 4.07$$

El valor *F* crítico para 11 gl en el numerador y 11 gl en el denominador en el nivel de 5% es 2.82. Como el valor $F(=\lambda)$ estimado excede al valor crítico, podemos concluir que hay heteroscedasticidad en la varianza del error. Sin embargo, si el nivel de significancia se fija en 1%, no podemos rechazar el supuesto de homoscedasticidad. (¿Por qué?) Observe que el valor *p* de la λ observada es 0.014.

USO DIDÁCTICO

Prueba Breusch-Pagan-Godfrey[21]

El éxito de la prueba de Goldfeld-Quandt depende no sólo del valor de c (el número de observaciones centrales que se van a omitir), sino también de la identificación de la variable X correcta que servirá de referencia para ordenar las observaciones. Esta limitación de la prueba se evita si consideramos la prueba Breusch-Pagan-Godfrey (BPG).

Para ilustrar esta prueba, considere el modelo de regresión lineal con k variables

$$Y_i = \beta_1 + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + u_i \tag{11.5.12}$$

Suponga que la varianza del error σ_i^2 se describe como

$$\sigma_i^2 = f(\alpha_1 + \alpha_2 Z_{2i} + \cdots + \alpha_m Z_{mi}) \tag{11.5.13}$$

es decir, σ_i^2 es algún tipo de función de las variables Z no estocásticas; alguna de las X o todas ellas pueden servir como Z. Específicamente, suponga que

$$\sigma_i^2 = \alpha_1 + \alpha_2 Z_{2i} + \cdots + \alpha_m Z_{mi} \tag{11.5.14}$$

es decir, σ_i^2 es una función lineal de las Z. Si $\alpha_2 = \alpha_3 = \cdots = \alpha_m = 0$, $\sigma_i^2 = \alpha_1$, que es una constante. Por consiguiente, para probar si σ_i^2 es homoscedástica, se puede probar la hipótesis de que $\alpha_2 = \alpha_3 = \cdots = \alpha_m = 0$. Ésta es la idea básica de la prueba Breusch-Pagan. El procedimiento es el siguiente.

Paso 1. Estime (11.5.12) mediante MCO y obtenga los residuos $\hat{u}_1, \hat{u}_2, \ldots, \hat{u}_n$.

Paso 2. Obtenga $\tilde{\sigma}^2 = \sum \hat{u}_i^2/n$. Recuerde, del capítulo 4, que éste es el estimador de máxima verosimilitud (MV) de σ^2. [*Nota:* El estimador de MCO es $\sum \hat{u}_i^2/[n-k]$.]

Paso 3. Construya las variables p_i definidas como

$$p_i = \hat{u}_i^2/\tilde{\sigma}^2$$

que es simplemente cada residuo elevado al cuadrado dividido entre $\tilde{\sigma}^2$.

Paso 4. Haga la regresión de los p_i así construidos sobre las Z como

$$p_i = \alpha_1 + \alpha_2 Z_{2i} + \cdots + \alpha_m Z_{mi} + v_i \tag{11.5.15}$$

donde v_i es el término de residuo para esta regresión.

Paso 5. Obtenga la SCE (suma de cuadrados explicada) de (11.5.15) y defina

$$\Theta = \frac{1}{2}(\text{SCE}) \tag{11.5.16}$$

Si suponemos que los u_i están normalmente distribuidos, se demuestra que sí hay homoscedasticidad, y si el tamaño n de la muestra aumenta indefinidamente, entonces

$$\Theta \underset{\text{asin}}{\sim} \chi^2_{m-1} \tag{11.5.17}$$

es decir, Θ sigue una distribución ji cuadrada con $(m-1)$ grados de libertad. (*Nota:* asin significa asintóticamente.)

[21] T. Breusch y A. Fagan, "A Simple Test for Heteroscedasticity and Random Coefficient Variation", *Econometrica,* vol. 47, 1979, pp. 1287-1294. Véase también L. Godfrey, "Testing for Multiplicative Heteroscedasticity", *Journal of Econometrics,* vol. 8, 1978, pp. 227-236. Debido a la similitud, estas pruebas se conocen como pruebas de heteroscedasticidad Breusch-Pagan-Godfrey.

USO DIDÁCTICO

Por consiguiente, si en una aplicación el Θ ($= \chi^2$) calculado excede al valor crítico χ^2 en el nivel de significancia seleccionado, se rechaza la hipótesis de homoscedasticidad; de lo contrario, no se rechaza.

El lector puede preguntar por qué la prueba BPG elige $\frac{1}{2}$ SCE como estadístico de prueba. El razonamiento es un poco complicado y se deja para consultar en las referencias.[22]

EJEMPLO 11.5
La prueba Breusch-Pagan-Godfrey (BPG)

A manera de ejemplo, reconsidere la información (tabla 11.3) para ilustrar la prueba de heteroscedasticidad de Goldfeld-Quandt. Al efectuar la regresión de Y sobre X, obtenemos lo siguiente:

Paso 1.

$$\hat{Y}_i = 9.2903 + 0.6378X_i$$
$$\text{ee} = (5.2314) \quad (0.0286) \qquad \text{SCR} = 2\,361.153 \qquad R^2 = 0.9466 \qquad \textbf{(11.5.18)}$$

Paso 2.

$$\tilde{\sigma}^2 = \sum \hat{u}_i^2/30 = 2\,361.153/30 = 78.7051$$

Paso 3. Divida los residuos $\hat{u}_i$ obtenidos de la regresión (11.5.18) entre 78.7051 para construir la variable p_i.

Paso 4. Si supone que las p_i están relacionadas linealmente con X_i ($= Z_i$), como lo establece (11.5.14), obtenemos la regresión

$$\hat{p}_i = -0.7426 + 0.0101X_i$$
$$\text{ee} = (0.7529) \quad (0.0041) \qquad \text{SCE} = 10.4280 \qquad R^2 = 0.18 \qquad \textbf{(11.5.19)}$$

Paso 5.

$$\Theta = \frac{1}{2}(\text{SCE}) = 5.2140 \qquad \textbf{(11.5.20)}$$

Con los supuestos de la prueba BPG, Θ en (11.5.20) sigue asintóticamente la distribución ji cuadrada con 1 gl. [*Nota:* sólo hay una regresora en (11.5.19).] Ahora, de la tabla ji cuadrada vemos que, para 1 gl, el valor crítico de ji cuadrada en 5% es 3.8414, y el valor χ^2 crítico en 1% es 6.6349. Por tanto, el valor observado ji cuadrada de 5.2140 es significativo en el nivel de significancia de 5% pero no en el nivel de 1%. Por consiguiente, llegamos a la misma conclusión obtenida mediante la prueba Goldfeld-Quandt. Pero tenga en mente que, en estricto sentido, la prueba BPG es asintótica o de muestras grandes, y en el ejemplo presente, la muestra de 30 observaciones puede no ser una muestra grande. Debe señalarse también que, en muestras pequeñas, la prueba es sensible al supuesto de que las perturbaciones u_i están normalmente distribuidas. Desde luego, podemos probar el supuesto de normalidad con las pruebas analizadas en el capítulo 5.[23]

USO DIDÁCTICO

Prueba general de heteroscedasticidad de White

A diferencia de la prueba de Goldfeld-Quandt, que requiere reordenar las observaciones respecto de la variable X que supuestamente ocasiona la heteroscedasticidad, o de la prueba BGP, sensible al supuesto de normalidad, la prueba general de heteroscedasticidad propuesta por White no se

[22] Véase Adrian C. Darnell, *A Dictionary of Econometrics,* Edward Elgar, Cheltenham, Reino Unido, 1994, pp. 178-179.

[23] Al respecto, véase R. Koenker, "A Note on Studentizing a Test for Heteroscedasticity", *Journal of Econometrics,* vol. 17, 1981, pp. 1180-1200.

apoya en el supuesto de normalidad y es fácil aplicarla.[24] Como ilustración de la idea básica, considere el siguiente modelo de regresión con tres variables (la generalización al modelo con k variables es sencilla):

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i \tag{11.5.21}$$

Para realizar la prueba de White se procede de la siguiente forma:

Paso 1. Dada la información, estime (11.5.21) y obtenga los residuos $\hat{u}_i$.

Paso 2. Efectúe la siguiente regresión (*auxiliar*):

$$\hat{u}_i^2 = \alpha_1 + \alpha_2 X_{2i} + \alpha_3 X_{3i} + \alpha_4 X_{2i}^2 + \alpha_5 X_{3i}^2 + \alpha_6 X_{2i} X_{3i} + v_i \tag{11.5.22}$$ [25]

Es decir, con el cuadrado de los residuos de la regresión original se hace la regresión sobre las variables o regresoras X originales, sobre sus valores al cuadrado y sobre el (los) producto(s) cruzado(s) de las regresoras. También pueden introducirse potencias más altas de las regresoras. Observe que hay un término constante en esta ecuación, aunque la regresión original puede o no contenerlo. Obtenga R^2 de esta regresión (auxiliar).

Paso 3. Según la hipótesis nula de que no hay heteroscedasticidad, puede demostrarse que el tamaño de la muestra (n) multiplicado por R^2 obtenido de la regresión auxiliar *asintóticamente* sigue la distribución ji cuadrada con gl igual al número de regresoras (sin el término constante) en la regresión auxiliar. Es decir,

$$n \cdot R^2 \underset{\text{asin}}{\sim} \chi^2_{\text{gl}} \tag{11.5.23}$$

donde los gl son iguales a los definidos antes. En el ejemplo, hay 5 gl porque hay 5 regresoras en la regresión auxiliar.

Paso 4. Si el valor ji cuadrada obtenido en (11.5.23) excede al valor ji cuadrada crítico en el nivel de significancia seleccionado, la conclusión es que hay heteroscedasticidad. Si éste no excede el valor ji cuadrada crítico, no hay heteroscedasticidad, lo cual quiere decir que en la regresión auxiliar (11.5.22), $\alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = \alpha_6 = 0$ (véase la nota 25).

USO DIDÁCTICO

EJEMPLO 11.6
Prueba de heteroscedasticidad de White

Con información de corte transversal de 41 países, Stephen Lewis estimó el siguiente modelo de regresión:[26]

$$\ln Y_i = \beta_1 + \beta_2 \ln X_{2i} + \beta_3 \ln X_{3i} + u_i \tag{11.5.24}$$

donde Y = razón entre impuestos arancelarios (impuestos sobre importaciones y exportaciones) y ganancias totales del gobierno, X_2 = razón entre la suma de exportaciones e importaciones y el PNB, X_3 = PNB *per cápita*, y ln representa el logaritmo natural. Sus hipótesis fueron que Y y X_2 estarían relacionadas positivamente (a mayor volumen de comercio exterior, mayor re-

(*continúa*)

[24] H. White, "A Heteroscedasticity Consistent Covariance Matrix Estimator and a Direct Test of Heteroscedasticity", *Econometrica,* vol. 48, 1980, pp. 817-818.

[25] El supuesto de que la varianza del error u_i, σ_i^2, se relaciona funcionalmente con las regresoras, con sus valores al cuadrado y con sus productos cruzados está implícito en este procedimiento. Si todos los coeficientes parciales de pendiente en esta regresión son simultáneamente iguales a cero, entonces la varianza del error es homoscedástica e igual a la constante α_1.

[26] Stephen R. Lewis, "Government Revenue from Foreign Trade", *Manchester School of Economics and Social Studies,* vol. 31, 1963, pp. 39-47.

EJEMPLO 11.6
(*continuación*)

caudo arancelario), y que Y y X_3 estarían negativamente relacionados (a medida que aumenta el ingreso, al gobierno se le facilita recaudar impuestos directos —es decir, el impuesto sobre la renta— que depende de los impuestos sobre el comercio exterior).

Los resultados empíricos apoyaron las hipótesis. Para el propósito, el punto importante es averiguar si hay heteroscedasticidad en los datos. Como los datos son de corte transversal e implican una heterogeneidad de países, podemos esperar *a priori* heteroscedasticidad en la varianza del error. Con la prueba de heteroscedasticidad de White en los residuos obtenidos de la regresión (11.5.24) se obtuvieron los siguientes resultados:[27]

$$\begin{aligned}\widehat{\hat{u}_i^2} = &-5.8417 + 2.5629 \ln \text{Comercio}_i + 0.6918 \ln \text{PNB}_i \\ &-0.4081(\ln \text{Comercio}_i)^2 - 0.0491(\ln \text{PNB}_i)^2 \\ &+0.0015(\ln \text{Comercio}_i)(\ln \text{PNB}_i) \qquad R^2 = 0.1148\end{aligned} \tag{11.5.25}$$

Nota: Los errores estándar no están dados porque no son pertinentes para nuestro propósito.

Ahora, $n \cdot R^2 = 41(0.1148) = 4.7068$, que tiene, asintóticamente, una distribución ji cuadrada con 5 gl (¿por qué?). El valor ji cuadrada crítico en 5% para 5 gl es 11.0705, el valor crítico en 10% es 9.2363, y el valor crítico en 25% es 6.62568. Para todos los fines prácticos, podemos concluir, con base en la prueba de White, que no hay heteroscedasticidad.

Conviene hacer un comentario relacionado con la prueba de White. Si un modelo tiene muchas regresoras, la introducción de todas las regresoras, de sus términos elevados al cuadrado (o a potencias más elevadas) y de sus productos cruzados pueden consumir grados de libertad rápidamente. Por consiguiente, se debe tener cautela con esta prueba.[28]

En los casos en que el estadístico de prueba de White es significativo estadísticamente, la heteroscedasticidad puede no necesariamente ser la causa, sino los errores de especificación, los cuales veremos en mayor detalle en el capítulo 13 (recuerde el punto 5 de la sección 11.1). En otras palabras, **la prueba de White puede ser una prueba de heteroscedasticidad (pura), de error de especificación o de ambos**. Se argumenta que, si no están presentes términos con productos cruzados en el procedimiento de prueba de White, esto constituye una prueba de heteroscedasticidad pura. Si existen tales términos, es una prueba de heteroscedasticidad y de sesgo de especificación.[29]

Otras pruebas de heteroscedasticidad

Hay muchas otras pruebas de heteroscedasticidad, cada una con supuestos determinados. El lector interesado puede consultar las referencias.[30] Mencionamos sólo una de estas pruebas debido a su simplicidad. Es la **prueba de Koenker-Basset (KB)**. Al igual que las pruebas Park, Breusch-Pagan-Godfrey y la de White, la prueba KB se basa en los residuos al cuadrado, $\hat{u}_i^2$, pero en vez de hacer la regresión sobre una o más regresoras, se efectúa la regresión de los residuos al cuadrado sobre los valores estimados de la regresora al cuadrado. De manera específica, si el modelo original es:

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \cdots + \beta_k X_{ki} + u_i \tag{11.5.26}$$

[27] Estos resultados, con cambio en la notación, se reproducen de William F. Lott y Subhash C. Ray, *Applied Econometrics: Problems with Data Sets, Instructor's Manual*, capítulo 22, pp. 137-140.

[28] A veces la prueba se puede modificar para conservar los grados de libertad. Consulte el ejercicio 11.18.

[29] Véase Richard Harris, *Using Cointegration Analysis in Econometrics Modelling,* Prentice Hall & Harvester Wheatsheaf, Reino Unido, 1995, p. 68.

[30] Véase M.J. Harrison y B.E. McCabe, "A Test for Heteroscedasticity Based on Ordinary Least Squares Residuals", *Journal of the American Statistical Association,* vol. 74, 1979, pp. 494-499; J. Szroeter, "A Class of Parametric Tests for Heteroscedasticity in Linear Econometric Models", *Econometrica,* vol. 46, 1978, pp. 1311-1327; M.A. Evans y M.L. King, "A Further Class of Tests for Heteroscedasticity", *Journal of Econometrics,* vol. 37, 1988, pp. 265-276; y R. Koenker y G. Bassett, "Robust Tests for Heteroscedasticity Based on Regression Quantiles", *Econometrica,* vol. 50, 1982, pp. 43-61.

USO DIDÁCTICO

se estima este modelo, se obtiene $\hat{u}_i$ de dicho modelo y luego se calcula

$$\hat{u}_i^2 = \alpha_1 + \alpha_2(\hat{Y}_i)^2 + v_i \tag{11.5.27}$$

donde $\hat{Y}_i$ son los valores estimados del modelo (11.5.26). La hipótesis nula es que $\alpha_2 = 0$. Si no se rechaza, se puede concluir que no existe heteroscedasticidad. La hipótesis nula se prueba con las pruebas t o F usuales. (Observe que $F_{1,k} = t_k^2$.) Si el modelo (11.5.26) es doble logaritmo, se lleva a cabo la regresión de los residuos al cuadrado sobre $(\log \hat{Y}_i)^2$. Otra ventaja de la prueba KB es que es aplicable aunque el término de error en el modelo original (11.5.26) no esté normalmente distribuido. Si aplica la prueba KB al ejemplo 11.1, descubrirá que el coeficiente de la pendiente en la regresión de los residuos cuadrados obtenida de (11.5.3) sobre el $\hat{Y}_i^2$ estimado a partir de (11.5.3) no es estadísticamente distinto de cero, por lo que se refuerza la prueba de Park. Este resultado no debe sorprender, pues en estos momentos sólo se tiene una sola regresora. No obstante, la prueba KB es aplicable si hay una o muchas regresoras.

Nota sobre las pruebas de heteroscedasticidad

Ya analizamos varias pruebas de heteroscedasticidad en esta sección. Pero, ¿cómo decidir cuál es la mejor? No es una pregunta fácil, pues estas pruebas se basan en supuestos diversos. Al comparar las pruebas, es necesario prestar atención al tamaño (o nivel de significancia), potencia (la probabilidad de rechazar una hipótesis falsa) y sensibilidad a los valores atípicos.

Ya señalamos algunas limitaciones de la prueba de heteroscedasticidad de White, que es popular y fácil de aplicar. Como resultado de estas limitaciones, tal vez tenga poca potencia en relación con las opciones. Además, la prueba no sirve para identificar los factores o variables que causan heteroscedasticidad.

Asimismo, la prueba de Breusch-Pagan-Godfrey es sensible al supuesto de normalidad. En contraste, la prueba de Koenker-Bassett no se basa en el supuesto de normalidad y, en consecuencia, puede ser más potente.[31] En la prueba de Goldfeld-Quandt, si se omiten muchas observaciones, puede disminuir la potencia de la prueba.

Está fuera del ámbito de este texto proporcionar un análisis comparativo de las diferentes pruebas de heteroscedasticidad. Sin embargo, el lector interesado puede consultar el artículo de John Lyon y Chin-Ling Tsai para darse una idea de los puntos fuertes y débiles de las diversas pruebas de heteroscedasticidad.[32]

USO DIDÁCTICO

11.6 Medidas correctivas

Como vimos, la heteroscedasticidad no destruye las propiedades de insesgamiento y consistencia de los estimadores de MCO; sin embargo, éstos ya no son eficientes, ni siquiera asintóticamente (es decir, en muestras grandes). Esta falta de eficiencia resta credibilidad a los procedimientos habituales de pruebas de hipótesis. Por consiguiente, es necesario introducir medidas correctivas. Existen dos enfoques para remediar el problema de heteroscedasticidad: cuando se conoce σ_i^2 y cuando no se conoce σ_i^2.

Cuando se conoce σ_i^2: método de los mínimos cuadrados ponderados

Como vimos en la sección 11.3, si se conoce σ_i^2, el método más directo de corregir la heteroscedasticidad es con los mínimos cuadrados ponderados, pues los estimadores obtenidos mediante este método son MELI.

[31] Para detalles, véase William H. Green, *Econometric Analysis,* 6a. ed., Pearson/Prentice-Hall, Nueva Jersey, 2008, pp. 165-167.

[32] Véase su artículo "A Comparison of Tests of Heteroscedasticity", *The Statistician,* vol. 45, núm. 3, 1996, pp. 337-349.

EJEMPLO 11.7
Ilustración del método de los mínimos cuadrados ponderados

Para ilustrar el método, suponga que deseamos estudiar la relación entre la remuneración salarial y el tamaño de la planta laboral de los datos de la tabla 11.1. Por simplicidad, se mide el tamaño de la planta laboral con las siguientes categorías: 1 (1-4 empleados), 2 (5-9 empleados), . . . , 9 (1 000-2 499 empleados), aunque también se puede medir con el punto medio de las diversas clases de empleados en la tabla.

Ahora, sea Y la remuneración salarial promedio por empleado ($) y X el tamaño de empleados. Efectuamos la siguiente regresión [véase la ecuación (11.3.6)]:

$$Y_i/\sigma_i = \hat{\beta}_1^*(1/\sigma_i) + \hat{\beta}_2^*(X_i/\sigma_i) + (\hat{u}_i/\sigma_i) \qquad \textbf{(11.6.1)}$$

donde σ_i son las desviaciones estándar de los salarios como aparecen en la tabla 11.1. Los datos simples necesarios para efectuar esta regresión están en la tabla 11.4.

TABLA 11.4
Ilustración de una regresión de mínimos cuadrados ponderados

Remuneración, Y	Tamaño de la planta laboral, X	σ_i	Y_i/σ_i	X_i/σ_i
3 396	1	742.2	4.5664	0.0013
3 787	2	851.4	4.4480	0.0023
4 013	3	727.8	5.5139	0.0041
4 104	4	805.06	5.0978	0.0050
4 146	5	929.9	4.4585	0.0054
4 241	6	1 080.6	3.9247	0.0055
4 387	7	1 241.2	3.5288	0.0056
4 538	8	1 307.7	3.4702	0.0061
4 843	9	1 110.7	4.3532	0.0081

Fuente: La información sobre Y y σ_i (desviación estándar de la remuneración salarial) corresponde a la tabla 11.1. El tamaño de la planta laboral: 1 = 1-4 empleados, 2 = 5-9 empleados, etc. Los últimos datos son también de la tabla 11.1.

Nota: En la regresión (11.6.2), la variable dependiente es (Y_i/σ_i), y las independientes, $(1/\sigma_i)$ y (X_i/σ_i).

Antes de proseguir con el análisis de los resultados de la regresión, observe que (11.6.1) no tiene término de intercepto (¿por qué?). Por consiguiente, debemos utilizar el modelo de regresión a través del origen para estimar β_1^* y β_2^*, tema analizado en el capítulo 6. Pero, hoy en día, la mayoría de los paquetes de computación dan la opción de suprimir el término de intercepto (*Minitab* o *EViews*, por ejemplo). Observe también otra característica interesante de (11.6.1): tiene dos variables explicativas, $(1/\sigma_i)$ y (X_i/σ_i), mientras que si utilizáramos MCO, la regresión del salario sobre el tamaño de la planta laboral tendría una sola variable explicativa, X_i. (¿Por qué?)

Los resultados de la regresión de MCP son los siguientes:

$$\begin{aligned} \widehat{(Y_i/\sigma_i)} &= 3\,406.639(1/\sigma_i) + 154.153(X_i/\sigma_i) \\ &\quad\;\; (80.983) \qquad\qquad (16.959) \\ t = &\quad\;\; (42.066) \qquad\qquad (9.090) \\ &\quad\;\; R^2 = 0.9993^{33} \end{aligned} \qquad \textbf{(11.6.2)}$$

Para su comparación, presentamos a continuación los resultados de la regresión de MCO usual o no ponderada:

$$\begin{aligned} \hat{Y}_i &= 3\,417.833 + 148.767\,X_i \\ &\quad\;\; (81.136) \quad (14.418) \\ t = &\quad\;\; (42.125) \quad (10.318) \qquad R^2 = 0.9383 \end{aligned} \qquad \textbf{(11.6.3)}$$

En el ejercicio 11.7 se le pide comparar estas dos regresiones.

[33] Como se señaló en la nota 3 del capítulo 6, R^2 de la regresión a través del origen no es directamente comparable con R^2 del modelo con presencia del intercepto. El valor R^2 de 0.9993 obtenido tiene en cuenta esta diferencia. (Véanse los diversos paquetes de cómputo para mayores detalles sobre la forma de corregir R^2 para tener en cuenta la ausencia del término del intercepto. Véase también el apéndice 6A, sección 6A1.)

USO DIDÁCTICO

Cuando no se conoce σ_i^2

Como ya vimos, si se conocen las verdaderas σ_i^2, podemos utilizar el método de MCP para obtener estimadores MELI. Como pocas veces se conocen las verdaderas σ_i^2, ¿existe alguna forma de obtener estimaciones *consistentes* (en el sentido estadístico) de las varianzas y covarianzas de los estimadores de MCO aunque haya heteroscedasticidad? La respuesta es sí.

Varianzas y errores estándar consistentes con heteroscedasticidad de White

White demostró que esta estimación puede realizarse de forma que las inferencias estadísticas sean *asintóticamente* válidas (es decir, para muestras grandes) sobre los verdaderos valores de los parámetros.[34] No presentaremos aquí los detalles matemáticos porque no están al alcance de este libro. En el apéndice 11A.4 se detalla el procedimiento de White. Sin embargo, en la actualidad hay diversos paquetes de computación que presentan varianzas y errores estándar con la corrección de heteroscedasticidad de White en forma simultánea con las varianzas y los errores estándar de MCO usuales.[35] A propósito, los errores estándar de White corregidos mediante heteroscedasticidad también se conocen como **errores estándar robustos**.

EJEMPLO 11.8
Ilustración del procedimiento de White

Como ejemplo, citamos los siguientes resultados obtenidos por Greene:[36]

$$\begin{aligned}
\hat{Y}_i &= 832.91 - 1\,834.2\,(\text{Ingreso}) + 1\,587.04\,(\text{Ingreso})^2 \\
\text{ee de MCO} &= (327.3) \quad (829.0) \quad (519.1) \\
t &= (2.54) \quad (2.21) \quad (3.06) \qquad \textbf{(11.6.4)} \\
\text{ee de White} &= (460.9) \quad (1\,243.0) \quad (830.0) \\
t &= (1.81) \quad (-1.48) \quad (1.91)
\end{aligned}$$

donde Y = gasto *per cápita* en escuelas públicas por estado en 1979 e Ingreso = ingreso *per cápita* por estado en 1979. La muestra consistió en 50 estados más Washington, D.C.

Como demuestran los resultados anteriores, los errores estándar corregidos por heteroscedasticidad (de White) resultan considerablemente más grandes que los errores estándar de MCO, y, por consiguiente, los valores t estimados son mucho menores que los obtenidos por MCO. Con base en estos últimos, ambas regresoras son estadísticamente significativas en el nivel de 5%, mientras que con base en los estimadores de White, no lo son. Sin embargo, cabe señalar que los errores estándar corregidos por heteroscedasticidad de White pueden ser más grandes o más pequeños que los errores estándar sin corregir.

Como los estimadores de las varianzas consistentes con heteroscedasticidad de White están disponibles ahora en paquetes de computación para regresión, se recomienda que el lector los reporte. Como recomiendan Wallace y Silver:

> En términos generales, quizá sea buena idea utilizar la opción WHITE [disponible en los programas de regresión] sistemáticamente, tal vez comparar estos resultados con los resultados de MCO regulares como forma de verificar si la heteroscedasticidad es un problema grave en un conjunto particular de datos.[37]

Supuestos razonables sobre el patrón de heteroscedasticidad

Una desventaja del procedimiento de White, además de ser de muestras grandes, es que los estimadores obtenidos por este medio pueden no ser tan eficientes como los obtenidos por métodos

[34] Véase H. White, *op. cit.*

[35] Más técnicamente, se conocen como **estimadores de la matriz de covarianzas consistentes con heteroscedasticidad**: para abreviar, **EMCCH**.

[36] William H. Greene, *Econometric Analysis,* 2a. ed., Macmillan, Nueva York, 1993, p. 385.

[37] T. Dudley Wallace y J. Lew Silver, *Econometrics: An Introduction,* Addison-Wesley, Reading, Massachusetts, 1988, p. 265.

USO DIDÁCTICO

que transforman la información para reflejar tipos específicos de heteroscedasticidad. Para ilustrar esto, recordemos el modelo de regresión con dos variables:

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

Consideraremos ahora diversos supuestos sobre el patrón de heteroscedasticidad.

SUPUESTO 1 La varianza del error es proporcional a X_i^2:

$$E\left(u_i^2\right) = \sigma^2 X_i^2 \qquad \textbf{(11.6.5)}^{38}$$

Si, por razones de "especulación", por los métodos gráficos, o por los métodos Park y Glejser, se cree que la varianza de u_i es proporcional al cuadrado de la variable explicativa X (figura 11.10), se puede transformar el modelo original de la siguiente manera. Divida el modelo original entre X_i:

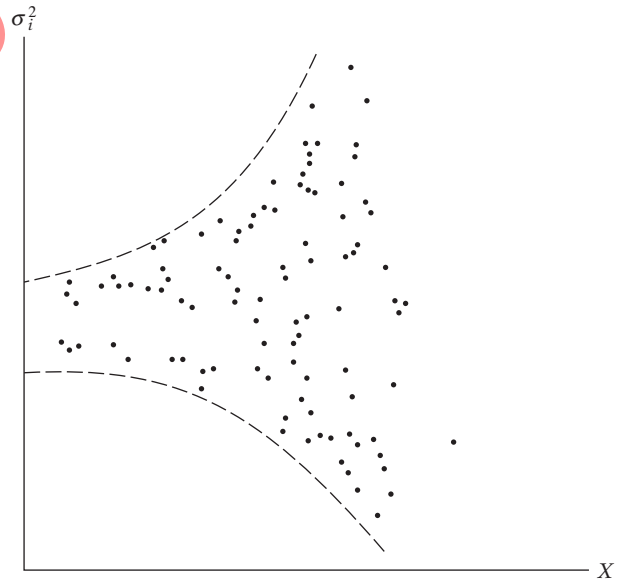
$$\begin{aligned} \frac{Y_i}{X_i} &= \frac{\beta_1}{X_i} + \beta_2 + \frac{u_i}{X_i} \\ &= \beta_1 \frac{1}{X_i} + \beta_2 + v_i \end{aligned} \qquad \textbf{(11.6.6)}$$

donde v_i es el término de perturbación transformado, igual a u_i/X_i. Ahora, es fácil verificar que

$$\begin{aligned} E\left(v_i^2\right) &= E\left(\frac{u_i}{X_i}\right)^2 = \frac{1}{X_i^2} E\left(u_i^2\right) \\ &= \sigma^2 \qquad \text{utilizando (11.6.5)} \end{aligned}$$

Por tanto, la varianza de v_i es ahora homoscedástica y podemos aplicar MCO a la ecuación transformada (11.6.6), con la regresión Y_i/X_i sobre $1/X_i$.

FIGURA 11.10
Varianza del error proporcional a X^2.



[38] Recuerde que ya tratamos este supuesto en el análisis de la prueba Goldfeld-Quandt.

USO DIDACTICO

Observe que, en la regresión transformada, el término del intercepto β_2 es el coeficiente de pendiente en la ecuación original, y el coeficiente de la pendiente β_1 es el término del intercepto en el modelo original. Por consiguiente, para retornar al modelo original tenemos que multiplicar la estimación (11.6.6) por X_i. Una aplicación de esta transformación está en el ejercicio 11.20.

SUPUESTO 2 La varianza del error es proporcional a X_i. La **transformación de raíz cuadrada**:

$$E\left(u_i^2\right) = \sigma^2 X_i \tag{11.6.7}$$

Si se cree que la varianza de u_i, en lugar de ser proporcional al cuadrado X_i, es proporcional a la misma X_i, entonces el modelo original se transforma de la siguiente manera (figura 11.11):

$$\begin{aligned}\frac{Y_i}{\sqrt{X_i}} &= \frac{\beta_1}{\sqrt{X_i}} + \beta_2\sqrt{X_i} + \frac{u_i}{\sqrt{X_i}} \\ &= \beta_1\frac{1}{\sqrt{X_i}} + \beta_2\sqrt{X_i} + v_i\end{aligned} \tag{11.6.8}$$

donde $v_i = u_i/\sqrt{X_i}$ y donde $X_i > 0$.

Con el supuesto 2 se verifica fácilmente que $E(v_i^2) = \sigma^2$, una situación homoscedástica. Por consiguiente, se puede aplicar MCO a (11.6.8), con la regresión de $Y_i/\sqrt{X_i}$ sobre $1/\sqrt{X_i}$ y $\sqrt{X_i}$.

Observe una característica importante del modelo transformado: no tiene término de intercepto. Por consiguiente, será necesario el modelo de regresión a través del origen para estimar β_1 y β_2. Tras efectuar la regresión (11.6.8), retornamos al modelo original con tan sólo multiplicar (11.6.8) por $\sqrt{X_i}$.

Un caso interesante es el modelo de intercepto cero, es decir, $Y_i = \beta_2 X_i + u_i$. En este caso, la ecuación (11.6.8) se convierte en:

$$\frac{Y_i}{\sqrt{X_i}} = \beta_2\sqrt{X_i} + \frac{u_i}{\sqrt{X_i}} \tag{11.6.8a}$$

USO DIDÁCTICO

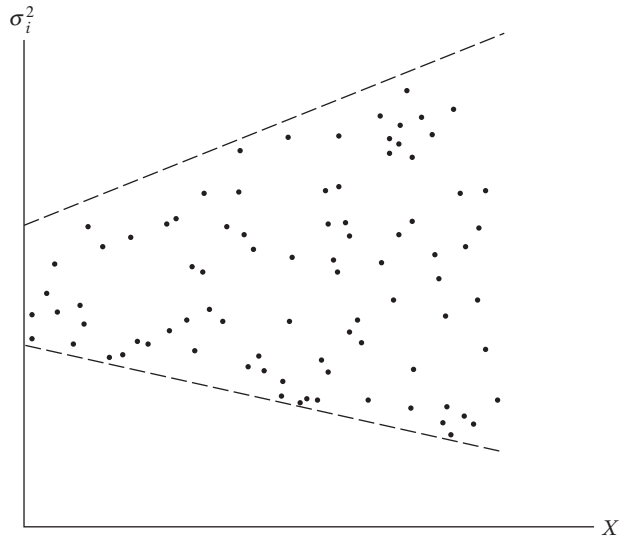


FIGURA 11.11
Varianza del error proporcional a X.

Y se puede demostrar que

$$\hat{\beta}_2 = \frac{\bar{Y}}{\bar{X}} \tag{11.6.8b}$$

Es decir, el estimador de mínimos cuadrados ponderados es simplemente la razón de las medias de las variables dependiente y explicativa. (Para probar la ecuación [11.6.8b], se aplica la fórmula de regresión a través del origen dada en la ecuación [6.1.6].)

SUPUESTO 3 La varianza del error es proporcional al cuadrado del valor medio de Y.

$$E\left(u_i^2\right) = \sigma^2[E(Y_i)]^2 \tag{11.6.9}$$

La ecuación (11.6.9) postula que la varianza de u_i es proporcional al cuadrado del valor esperado de Y (figura 11.8*e*). Ahora,

$$E(Y_i) = \beta_1 + \beta_2 X_i$$

Por consiguiente, si transformamos la ecuación original de la siguiente manera,

$$\begin{aligned} \frac{Y_i}{E(Y_i)} &= \frac{\beta_1}{E(Y_i)} + \beta_2 \frac{X_i}{E(Y_i)} + \frac{u_i}{E(Y_i)} \\ &= \beta_1 \left(\frac{1}{E(Y_i)}\right) + \beta_2 \frac{X_i}{E(Y_i)} + v_i \end{aligned} \tag{11.6.10}$$

donde $v_i = u_i/E(Y_i)$, se ve que $E(v_i^2) = \sigma^2$; es decir, las perturbaciones v_i son homoscedásticas. Por tanto, es la regresión (11.6.10) la que satisfará el supuesto de homoscedasticidad del modelo clásico de regresión lineal.

La transformación (11.6.10), sin embargo, no funciona, porque $E(Y_i)$ depende de β_1 y β_2, los cuales no se conocen. Por supuesto, se conoce $\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i$, que es un estimador de $E(Y_i)$. Por consiguiente, podemos proceder en dos etapas: primero, efectuamos la regresión de MCO usual sin considerar el problema de heteroscedasticidad y obtenemos $\hat{Y}_i$. Luego, con el $\hat{Y}_i$ estimado, transformamos el modelo de la siguiente manera:

$$\frac{Y_i}{\hat{Y}_i} = \beta_1 \left(\frac{1}{\hat{Y}_i}\right) + \beta_2 \left(\frac{X_i}{\hat{Y}_i}\right) + v_i \tag{11.6.11}$$

donde $v_i = (u_i/\hat{Y}_i)$. En el paso 2 efectuamos la regresión (11.6.11). Aunque $\hat{Y}_i$ no es exactamente $E(Y_i)$, *estos estimadores son consistentes*; es decir, a medida que el tamaño de la muestra aumenta indefinidamente, éstos convergen hacia el verdadero $E(Y_i)$. Por tanto, la transformación (11.6.11) tendrá un desempeño adecuado en la práctica si el tamaño de la muestra es razonablemente grande.

SUPUESTO 4 Una transformación logarítmica como

$$\ln Y_i = \beta_1 + \beta_2 \ln X_i + u_i \tag{11.6.12}$$

con gran frecuencia reduce la heteroscedasticidad cuando se compara con la regresión $Y_i = \beta_1 + \beta_2 X_i + u_i$.

USO DIDACTICO

Este resultado surge porque la transformación logarítmica comprime las escalas en las cuales se miden las variables, y por tanto reduce una diferencia entre dos valores de diez veces a una diferencia de dos veces. Así, el número 80 es diez veces el número 8, pero el ln 80 ($= 4.3280$) es cerca de dos veces tan grande como ln 8 ($= 2.0794$).

Otra ventaja de la transformación logarítmica es que el coeficiente de pendiente β_2 mide la elasticidad de Y respecto de X, es decir, el cambio porcentual en Y ante un cambio porcentual en X. Por ejemplo, si Y es el consumo y X el ingreso, β_2 en (11.6.12) mide la elasticidad del ingreso, mientras que en el modelo original, β_2 sólo mide la tasa de cambio del consumo medio por cambio unitario del ingreso. Ésta es una razón por la cual los modelos logarítmicos son muy populares en la econometría empírica. (En el ejercicio 11.4 se aprecian algunos problemas asociados a la transformación logarítmica.)

Para concluir la exposición sobre medidas correctivas, de nuevo destacamos que todas las transformaciones analizadas antes son *ad hoc*; en esencia, especulamos sobre la naturaleza de σ_i^2. El que las transformaciones estudiadas en verdad funcionen depende de la naturaleza del problema y de la gravedad de la heteroscedasticidad. Hay otros problemas con las transformaciones que deben tomarse en cuenta:

1. Cuando se va más allá del modelo con dos variables tal vez no se conozca *a priori* la variable X que debe seleccionarse para transformar los datos.[39]

2. La transformación logarítmica como se analiza en el supuesto 4 no es aplicable si algún valor de Y o de X es cero o negativo.[40]

3. Además, hay un problema de **correlación espuria**. Este término, atribuido a Karl Pearson, se refiere a la situación de correlación entre las razones de variables, aunque las variables originales no estén correlacionadas o sean aleatorias.[41] Así, en el modelo $Y_i = \beta_1 + \beta_2 X_i + u_i$, Y y X pueden no estar correlacionados, pero en el modelo transformado $Y_i/X_i = \beta_1(1/X_i) + \beta_2$, a menudo se encuentra que Y_i/X_i y $1/X_i$ sí lo están.

4. Cuando las σ_i^2 no se conocen directamente y se estiman a partir de una o más de las transformaciones ya analizadas, todos los procedimientos de prueba t, F, etc., son, *en estricto sentido, válidos sólo para muestras grandes*. Por consiguiente, se debe tener cuidado al interpretar resultados fundados en las diversas transformaciones cuando las muestras son pequeñas o finitas.[42]

11.7 Ejemplos para concluir

Para concluir el análisis de heteroscedasticidad presentamos tres ejemplos que ilustran los principales temas de este capítulo.

USO DIDACTICO

[39] Sin embargo, como un asunto práctico, se puede graficar $\hat{u}_i^2$ contra cada una de las variables y decidir cuál variable X puede utilizarse para transformar la información (figura 11.9).

[40] Algunas veces podemos utilizar ln $(Y_i + k)$ o ln $(X_i + k)$, donde k es un número positivo seleccionado de forma que todos los valores de Y y X se tornan positivos.

[41] Por ejemplo, si X_1, X_2 y X_3 están mutuamente no correlacionadas, $r_{12} = r_{13} = r_{23} = 0$ y se encuentra que las razones (sus valores) X_1/X_3 y X_2/X_3 están correlacionadas, entonces hay correlación espuria. "En términos más generales, la correlación se describe como espuria si es inducida por el método de manejo de datos y no está presente en la información original." M. G. Kendall y W. R. Buckland, *A Dictionary of Statistical Terms*, Hafner, Nueva York, 1972, p. 143.

[42] Para mayores detalles, véase George G. Judge *et al., op. cit.*, sección 14.4, pp. 415-420.

EJEMPLO 11.9
De nuevo la mortalidad infantil

Regresemos al ejemplo de mortalidad infantil que hemos analizado en diversas ocasiones. De los datos de 64 países obtuvimos los resultados de la regresión mostrados en la ecuación (8.1.4). En vista de que los datos son transversales e implican diferentes países con distintas experiencias sobre mortalidad infantil, parece muy probable que haya heteroscedasticidad. Para descubrirlo, primero tomaremos en cuenta los residuos obtenidos de la ecuación (8.1.4). Tales residuos se grafican en la figura 11.12. De acuerdo con dicha gráfica, al parecer los residuos no muestran ningún patrón específico que sugiera heteroscedasticidad. No obstante, las apariencias engañan, así que aplicaremos las pruebas de Park, Glejser y White para descubrir alguna evidencia de heteroscedasticidad.

Prueba de Park. Como hay dos regresoras, PIB y TAM, hacemos la regresión de los residuos al cuadrado a partir de la regresión (8.1.4) sobre cualquiera de esas variables. O podemos efectuar la regresión de ellos sobre los valores estimados de MI ($= \widehat{\text{MI}}$) a partir de la regresión (8.1.4). Con la regresión anterior obtenemos los siguientes resultados:

$$\begin{aligned} \widehat{\hat{u}_i^2} &= 854.4006 + 5.7016\,\widehat{\text{MI}}_i \\ t &= \quad (1.2010) \quad (1.2428) \qquad r^2 = 0.024 \end{aligned} \tag{11.7.1}$$

Nota: $\hat{u}_i$ son los residuos obtenidos de la regresión (8.1.4) y $\widehat{\text{MI}}$ son los valores estimados de MI a partir de la regresión (8.1.4).

Como muestra esta regresión, no existe ninguna relación sistemática entre el cuadrado de los residuos y los valores estimados de MI (¿por qué?), lo cual sugiere que el supuesto sobre la homoscedasticidad puede ser válido. A propósito, si hacemos la regresión del logaritmo de los residuos al cuadrado sobre el logaritmo de $\widehat{\text{MI}}$, no cambia la conclusión.

Prueba de Glejser. Los valores absolutos de los residuos obtenidos a partir de (8.1.4), cuando se hizo la regresión de los valores estimados de MI a partir de esa misma regresión, dieron estos resultados:

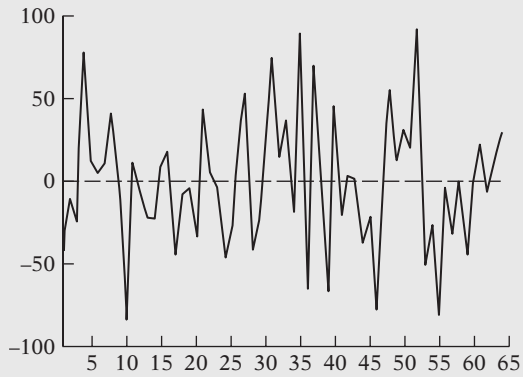
$$\begin{aligned} \widehat{|\hat{u}_i|} &= 22.3127 + 0.0646\,\widehat{\text{MI}}_i \\ t &= (2.8086) \quad (1.2622) \qquad r^2 = 0.0250 \end{aligned} \tag{11.7.2}$$

De nuevo, no existe una gran relación sistemática entre los valores absolutos de los residuos y los valores estimados de MI, pues el valor t del coeficiente de la pendiente no es estadísticamente significativo.

Prueba de White. Al aplicar la prueba de White para la heteroscedasticidad con y sin los términos de los productos cruzados no hallamos ninguna evidencia de heteroscedasticidad. También volvimos a estimar (8.1.4) para obtener los errores estándar consistentes con heteroscedasticidad de White y los valores t, pero los resultados fueron muy parecidos a los de la ecuación (8.1.4), lo cual no sorprende en vista de lo arrojado por las diversas pruebas de heteroscedasticidad anteriores.

En resumen, parece que la regresión de mortalidad infantil (8.1.4) no experimenta heteroscedasticidad alguna.

FIGURA 11.12
Residuos de la regresión (8.1.4).



USO DIDÁCTICO

EJEMPLO 11.10

Gastos en investigación y desarrollo (IyD), ventas y utilidades de 14 sectores industriales de Estados Unidos, 2005

En la tabla 11.5 se presentan datos sobre el gasto en investigación y desarrollo (IyD), ventas y utilidades de 14 sectores industriales en Estados Unidos (todas las cifras se expresan en millones de dólares). Como los datos transversales presentados en la tabla 11.5 son muy heterogéneos, en una regresión de IyD sobre las ventas es probable la presencia de heteroscedasticidad. Los resultados obtenidos de la regresión fueron los siguientes:

$$\begin{aligned} \widehat{\text{IyD}}_i &= 1\,338 + 0.0437\ \text{Ventas}_i \\ \text{ee} &= (5\,015) \quad (0.0277) \\ t &= (0.27) \quad (1.58) \qquad r^2 = 0.172 \end{aligned} \tag{11.7.3}$$

No sorprende una relación positiva entre IyD y ventas, aunque no sea estadísticamente significativa en los niveles tradicionales.

TABLA 11.5

Ventas y empleo para empresas que realizan IyD industrial en Estados Unidos, por industria, 2005 (los valores se expresan en millones de dólares)

Fuente: National Science Foundation, Division of Science Resources Statistics, Survey of Industrial Research and Development, 2005, y U.S. Census Bureau Annual Survey of Manufacturers, 2005.

Industria	Ventas	IyD	Utilidades
1 Alimentos	374 342	2 716	234 662
2 Textiles, ropa y cuero	51 639	816	53 510
3 Productos químicos básicos	109 899	2 277	75 168
4 Resinas, caucho sintético, fibras y filamentos	132 934	2 294	34 645
5 Productos farmacéuticos y medicamentos	273 377	34 839	127 639
6 Productos de plástico y caucho	90 176	1 760	96 162
7 Productos metálicos fabricados	174 165	1 375	155 801
8 Maquinaria	230 941	8 531	143 472
9 Computadoras y equipo periférico	91 010	4 955	34 004
10 Semiconductores y otros componentes electrónicos	176 054	18 724	81 317
11 Instrumentos de navegación, medición, electromédicos y de control	118 648	15 204	73 258
12 Equipo eléctrico, aparatos electrodomésticos y componentes	101 398	2 424	54 742
13 Productos y partes aeroespaciales	227 271	15 005	72 090
14 Equipo y suministros médicos	56 661	4 374	52 443

USO DIDÁCTICO

Para verificar si la regresión (11.7.3) experimenta heteroscedasticidad, obtuvimos los residuos $\hat{u}_i$ y los residuos al cuadrado $\hat{u}_i^2$ del modelo, y se graficaron respecto de las ventas, como se muestra en la figura 11.13. Según esta figura, parece que existe un patrón sistemático entre los residuos y los cuadrados de los residuos y las ventas, lo cual indica heteroscedasticidad. Para probar esto de manera formal, utilizamos las pruebas de Park, Glejser y White, las cuales dieron los siguientes resultados:

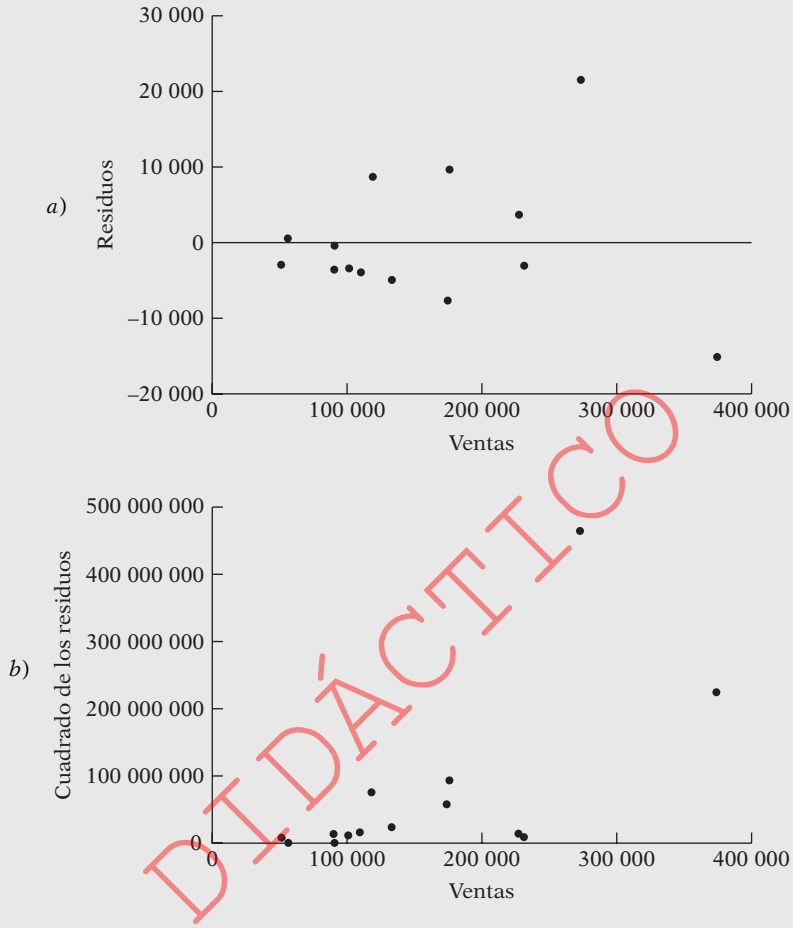
Prueba de Park

$$\begin{aligned} \widehat{\hat{u}_i^2} &= -72\,493\,719 + 916.1\ \text{Ventas}_i \\ \text{ee} &= (54\,940\,238) \quad (303.9) \\ t &= (-1.32) \quad (3.01) \qquad r^2 = 0.431 \end{aligned} \tag{11.7.4}$$

La prueba de Park indica una relación positiva estadísticamente significativa entre el cuadrado de los residuos y las ventas.

EJEMPLO 11.10 (*continuación*)

FIGURA 11.13 Residuos *a*) y cuadrado de los residuos *b*) sobre ventas.



USO DIDÁCTICO

Prueba de Glejser

$$\widehat{|\hat{u}_i|} = -1\,003 + 0.04639\ \text{Ventas}_i$$
$$\text{ee} = (2\,316) \quad (0.0128) \qquad \textbf{(11.7.5)}$$
$$t = (-0.43) \quad (3.62) \qquad r^2 = 0.522$$

La prueba de Glejser también indica una relación sistemática entre los valores absolutos de los residuos y las ventas, con lo que aumenta la posibilidad de que la regresión (11.7.3) presente heteroscedasticidad.

Prueba de White

$$\widehat{\hat{u}_i^2} = -46\,746\,325 + 578\ \text{Ventas}_i + 0.000846\ \text{Ventas}_i^2$$
$$\text{ee} = (112\,224\,348) \quad (1\,308) \quad (0.003171) \qquad \textbf{(11.7.6)}$$
$$t = (-0.42) \quad (0.44) \quad (0.27)$$
$$R^2 = 0.435$$

Con el valor R^2 y $n = 14$, se obtiene $nR^2 = 6.090$. Con la hipótesis nula de inexistencia de heteroscedasticidad, esto debe seguir una distribución ji cuadrada con 2 gl [porque hay dos regresoras en (11.7.6)]. El valor *p*, que resulta de obtener un valor ji cuadrado de 6.090 o mayor, es de casi 0.0476. Como se trata de un valor bajo, la prueba de White también indica heteroscedasticidad.

En resumen, con base en las gráficas de los residuos y las pruebas de Park, Glejser y White, parece que en la regresión de lyD (11.7.3) existe heteroscedasticidad. Como la verdadera varianza del error se desconoce, no podemos utilizar el método de los mínimos cuadrados ponderados para obtener los errores estándares corregidos por heteroscedasticidad y los valores *t*. Por consiguiente, con base en los datos disponibles, tendríamos que hacer conjeturas respecto de la naturaleza de la varianza de error.

Para concluir el ejemplo, a continuación presentamos los errores estándar de White consistentes con heteroscedasticidad, como se analizaron en la sección 11.6:

$$\begin{aligned} \widehat{\text{lyD}}_i &= 1\,337.87 + 0.0437\ \text{Ventas}_i \\ \text{ee} &= (4\,892.447) \quad (0.0411) \\ t &= \quad (0.27) \quad (1.06) \qquad r^2 = 0.172 \end{aligned} \tag{11.7.7}$$

Al comparar la ecuación (11.7.7) con la ecuación (11.7.3) (la segunda sin corrección por heteroscedasticidad), observamos que, a pesar de que no cambiaron los parámetros estimados (como era de esperarse), el error estándar del coeficiente del intercepto disminuyó un poco y el error estándar del coeficiente de la pendiente se incrementó un poco. Pero recuerde que el procedimiento de White es estrictamente para muestras grandes, en tanto que aquí sólo tenemos 14 observaciones.

USO DIDÁCTICO

EJEMPLO 11.11

La tabla 11.6 del sitio Web del libro contiene datos sobre salarios y otros aspectos relacionados correspondientes a 94 distritos escolares en el noroeste de Ohio. En principio, se estimó la siguiente regresión con base en estos datos:

$$\ln(\text{Salario})_i = \beta_1 + \beta_2 \ln(\text{Ingresofam}) + \beta_3 \ln(\text{Valorinm}) + u_i$$

Donde Salario = salario promedio de los maestros ($), Ingresofam = ingreso familiar promedio en el distrito ($) y Valorinm = valor promedio de los inmuebles en el distrito ($).

Como se trata de un modelo de doble logaritmo, todos los coeficientes de las pendientes son elasticidades. Con base en las diversas pruebas de heteroscedasticidad analizadas en el texto, se concluyó que el modelo anterior tenía heteroscedasticidad. Por tanto, obtuvimos los errores estándar robustos (de White). La siguiente tabla presenta los resultados de la regresión precedente con y sin errores estándar robustos.

Variable	Coeficiente	ee MCO	ee robusto
Intercepto	7.0198	0.8053	0.7721
		(8.7171)	(9.0908)
ln(ingresofam)	0.2575	0.0799	0.1009
		(3.2230)	(2.5516)
ln(valorinm)	0.0704	0.0207	0.0460
		(3.3976)	(1.5311)
R^2	0.2198		

Nota: Las cifras entre paréntesis son razones *t* estimadas.

Aunque los valores de los coeficientes y R^2 son iguales con el método de MCO o con el de White, los errores estándar cambiaron; el cambio más radical es el del error estándar del coeficiente ln(valorinm). El método habitual de MCO indicaría que el coeficiente estimado de esta variable es muy significativo estadísticamente, mientras que el error estándar robusto de White indica que este coeficiente no es significativo ni siquiera en el nivel de 10%. Lo importante de este ejemplo es que, si existe heteroscedasticidad, debe tomarse en cuenta para estimar el modelo.

11.8 Advertencia respecto de una reacción exagerada ante la heteroscedasticidad

En el ejemplo sobre IyD analizado en la sección anterior vimos que, cuando utilizamos la transformación raíz cuadrada para corregir la heteroscedasticidad en el modelo original (11.7.3), el error estándar del coeficiente de la pendiente disminuyó y su valor t aumentó. ¿Este cambio es tan importante como para causar preocupación en la práctica? En otras palabras, ¿cuándo debe preocupar el problema de la heteroscedasticidad? Como sostiene un autor: "la heteroscedasticidad jamás ha sido una razón para desechar un modelo que de otra forma sería adecuado".[43]

Aquí vale la pena recordar la advertencia de John Fox:

> . . .vale la pena corregir una varianza del error desigual sólo cuando el problema es grave.
>
> El impacto de una varianza del error no constante sobre la eficacia de un estimador de mínimos cuadrados ordinarios y sobre la validez de la inferencia de mínimos cuadrados depende de diversos factores, como el tamaño de la muestra, el grado de variación de σ_i^2, la configuración de los valores X (es decir, la regresora) y de la relación entre la varianza del error y las X. Por consiguiente, no es posible llegar a conclusiones generales respecto del daño producido por la heteroscedasticidad.[44]

De regreso al modelo (11.3.1), vimos ya que la varianza del estimador de pendiente, var $(\hat{\beta}_2)$, está dada por la fórmula usual mostrada en (11.2.3). Según MCG, la varianza del estimador de pendiente, var $(\hat{\beta}_2^*)$, está dada por (11.3.9). Sabemos que esta última es más eficaz que la primera. Pero, ¿qué tan grande debe ser la primer varianza (es decir, la de MCO) en relación con la varianza de MCG antes de que cause preocupación? Como regla práctica, Fox sugiere que el problema empiece a preocupar "...cuando la varianza del error más grande sea mayor que 10 veces la más pequeña".[45] Por consiguiente, al revisar los resultados de las simulaciones Monte Carlo de Davidson y MacKinnon presentadas antes, considere el valor de $\alpha = 2$. La varianza del β_2 estimado es 0.04 con MCO, y 0.012 con MCG, por lo que la razón de la primera respecto de la segunda es de casi 3.33.[46] De acuerdo con la regla de Fox, en este caso la gravedad de la heteroscedasticidad quizá no sea lo bastante grande para provocar preocupación.

También recuerde que, a pesar de la heteroscedasticidad, los estimadores de MCO (en condiciones generales) son lineales e insesgados, y están asintótica y normalmente distribuidos (es decir, en muestras grandes).

Como veremos cuando analicemos otra violación a los supuestos del modelo clásico de regresión lineal, la advertencia de esta sección resulta apropiada como regla general. Si hace caso omiso de lo anterior, puede cometer errores.

USO DIDÁCTICO

Resumen y conclusiones

1. Un supuesto importante del modelo clásico de regresión lineal es que todas las perturbaciones u_i tienen la misma varianza σ^2. Si este supuesto no se satisface, hay heteroscedasticidad.
2. La heteroscedasticidad no destruye las propiedades de insesgamiento y consistencia de los estimadores de MCO.
3. Sin embargo, estos estimadores dejan de tener varianza mínima, es decir, de ser eficientes. Por consiguiente, no son MELI.
4. Los estimadores MELI son proporcionados por el método de mínimos cuadrados ponderados, siempre que se conozcan las varianzas heteroscedásticas de error, σ_i^2.

[43] N. Gregory Mankiw, "A Quick Refresher Course in Macroeconomics", *Journal of Economic Literature,* vol. XXVIII, diciembre de 1990, p. 1648.

[44] John Fox, *Applied Regression Analysis, Linear Models, and Related Methods,* Sage, California, 1997, p. 306.

[45] *Ibid.*, p. 307.

[46] Observe que los errores estándar se elevaron al cuadrado para obtener las varianzas.

5. En presencia de heteroscedasticidad, las varianzas de los estimadores de MCO no se obtienen con las fórmulas usuales de MCO. Sin embargo, si insiste en utilizar las fórmulas habituales de MCO, las pruebas t y F basadas en ellas pueden conducir a grandes desatinos que darán por resultado conclusiones erróneas.
6. Es más fácil documentar las consecuencias de la heteroscedasticidad que detectarlas. Existen diversas pruebas de diagnóstico disponibles, pero no se puede decir con seguridad cuál funcionará en una situación dada.
7. Aunque se sospeche y se detecta la heteroscedasticidad, no es fácil corregir el problema. Si la muestra es grande, se pueden obtener los errores estándar de los estimadores de MCO corregidos por el método de heteroscedasticidad de White y realizar inferencia estadística basada en estos errores estándar.
8. De lo contrario, con base en los residuos de MCO, se pueden hacer conjeturas con ciertos fundamentos acerca del patrón probable de heteroscedasticidad y transformar la información original de manera que en la información transformada no haya heteroscedasticidad.

USO DIDÁCTICO

EJERCICIOS

Preguntas

11.1. Establezca si las siguientes afirmaciones son verdaderas, falsas o inciertas y *comente sus razones brevemente*:

a) En presencia de heteroscedasticidad, los estimadores de MCO son sesgados e ineficientes.

b) Si hay heteroscedasticidad, las pruebas convencionales t y F son inválidas.

c) En presencia de heteroscedasticidad, el método de MCO habitual siempre sobreestima los errores estándar de los estimadores.

d) Si los residuales estimados mediante una regresión por MCO exhiben un patrón sistemático, significa que hay heteroscedasticidad en los datos.

e) No hay una prueba general de heteroscedasticidad que no esté basada en algún supuesto acerca de cuál variable está correlacionada con el término de error.

f) Si el modelo de regresión está mal especificado (por ejemplo, si se omitió una variable importante), los residuos de MCO mostrarán un patrón claramente distinguible.

g) Si una regresora con varianza no constante se omite (incorrectamente) de un modelo, los residuos (MCO) serán heteroscedásticos.

11.2. En una regresión de salarios promedio (W, \$) sobre el número de empleados (N) de una muestra aleatoria de 30 empresas se obtuvieron los siguientes resultados:*

$$\widehat{W} = 7.5 + 0.009N \qquad (1)$$
$$t = \text{n.a.} \quad (16.10) \qquad R^2 = 0.90$$
$$\widehat{W}/N = 0.008 + 7.8(1/N) \qquad (2)$$
$$t = (14.43) \quad (76.58) \qquad R^2 = 0.99$$

a) ¿Cómo interpreta las dos regresiones?

b) ¿Qué supone el autor al pasar de la ecuación (1) a la (2)? ¿Le preocupaba la heteroscedasticidad? ¿Cómo sabe?

c) ¿Puede relacionar las pendientes y los interceptos de los dos modelos?

d) ¿Puede comparar los valores de R^2 de los dos modelos? ¿Por qué?

* Véase Dominick Salvatore, *Managerial Economics*, McGraw-Hill, Nueva York, 1989, p. 157.

11.3. *a*) ¿Puede estimar los parámetros de los modelos

$$|\hat{u}_i| = \sqrt{\beta_1 + \beta_2 X_i} + v_i$$

$$|\hat{u}_i| = \sqrt{\beta_1 + \beta_2 X_i^2} + v_i$$

mediante el método de mínimos cuadrados ordinarios? ¿Por qué?

b) Si la respuesta es negativa, ¿puede sugerir un método informal o formal de estimación de los parámetros de tales modelos? (Véase el capítulo 14.)

11.4. Aunque los modelos logarítmicos como el de la ecuación (11.6.12) a menudo reducen la heteroscedasticidad, se debe prestar cuidadosa atención a las propiedades del término de perturbación de estos modelos. Por ejemplo, el modelo

$$Y_i = \beta_1 X_i^{\beta_2} u_i \tag{1}$$

puede escribirse como

$$\ln Y_i = \ln \beta_1 + \beta_2 \ln X_i + \ln u_i \tag{2}$$

a) Si $\ln u_i$ tiene valor esperado cero, ¿cuál debe ser la distribución de u_i?

b) Si $E(u_i) = 1$, ¿será $E(\ln u_i) = 0$? ¿Por qué?

c) Si $E(\ln u_i)$ es diferente de cero, ¿qué puede hacer para volverlo cero?

11.5. Muestre que β_2^* de (11.3.8) también se expresa como

$$\beta_2^* = \frac{\sum w_i y_i^* x_i^*}{\sum w_i x_i^{2*}}$$

y var (β_2^*) dada en (11.3.9) también se expresa como

$$\text{var}\,(\beta_2^*) = \frac{1}{\sum w_i x_i^{2*}}$$

donde $y_i^* = Y_i - \bar{Y}^*$ y $x_i^* = X_i - X^*$ representan las desviaciones en relación con las medias ponderadas $\bar{Y}^*$ y $\bar{X}^*$ definidas como

$$\bar{Y}^* = \sum w_i Y_i \Big/ \sum w_i$$

$$\bar{X}^* = \sum w_i X_i \Big/ \sum w_i$$

11.6. Con propósitos pedagógicos, Hanushek y Jackson estiman el siguiente modelo:

$$C_t = \beta_1 + \beta_2 \text{PNB}_t + \beta_3 D_t + u_i \tag{1}$$

donde C_t = gasto agregado de consumo privado en el año t, PNB_t = producto nacional bruto en el año t y D_t = gastos de defensa nacional en el año t, con el objetivo de estudiar el efecto de los gastos de defensa sobre otros gastos en la economía.

Los autores postulan que $\sigma_t^2 = \sigma^2(\text{PNB}_t)^2$, luego transforman (1) y estiman

$$C_t/\text{PIB}_t = \beta_1\,(1/\text{PIB}_t) + \beta_2 + \beta_3\,(D_t/\text{PIB}_t) + u_t/\text{PIB}_t \tag{2}$$

USO DIDÁCTICO

Los resultados empíricos basados en la información de 1946 a 1975 fueron los siguientes (errores estándar entre paréntesis):*

$$\hat{C}_t = 26.19 \quad + 0.6248\ \text{PNB}_t - \ 0.4398\ D_t$$
$$(2.73) \qquad (0.0060) \qquad (0.0736) \qquad R^2 = 0.999$$
$$\widehat{C_t/\text{PNB}_t} = 25.92\,(1/\text{PNB}_t) + 0.6246 \quad - \ 0.4315\,(D_t/\text{PNB}_t)$$
$$(2.22) \qquad (0.0068) \qquad (0.0597) \qquad R^2 = 0.875$$

a) ¿Qué supuesto hacen los autores sobre la naturaleza de la heteroscedasticidad? ¿Puede justificarlo?

b) Compare los resultados de las dos regresiones. ¿La transformación del modelo original mejora los resultados, es decir, reduce los errores estándar estimados? ¿Por qué?

c) ¿Puede comparar los dos valores de R^2? ¿Por qué? (*Sugerencia:* Examine las variables dependientes.)

11.7. Consulte las regresiones estimadas (11.6.2) y (11.6.3). Los resultados de la regresión son muy similares. ¿A qué se debe esta conclusión?

11.8. Pruebe que si $w_i = w$, una constante, para cada i, β_2^* y $\hat{\beta}_2$ son idénticos, así como sus varianzas.

11.9. Consulte las fórmulas (11.2.2) y (11.2.3), y suponga que

$$\sigma_i^2 = \sigma^2 k_i$$

donde σ^2 es una constante y k_i son ponderaciones *conocidas*, no necesariamente todas iguales.

Con este supuesto, muestre que la varianza dada en (11.2.2) se expresa como

$$\text{var}\,(\hat{\beta}_2) = \frac{\sigma^2}{\sum x_i^2} \cdot \frac{\sum x_i^2 k_i}{\sum x_i^2}$$

El primer término del miembro derecho es la fórmula de la varianza dada en (11.2.3), es decir, var (β_2^*) con homoscedasticidad. ¿Qué puede decir sobre la naturaleza de la relación entre var $(\hat{\beta}_2)$ con heteroscedasticidad y con homoscedasticidad? (*Sugerencia:* Examine, en la fórmula anterior, el segundo término del miembro derecho.) ¿Puede derivar alguna conclusión general sobre las relaciones entre (11.2.2) y (11.2.3)?

11.10. En el modelo

$$Y_i = \beta_2 X_i + u_i \quad (\textit{Nota:}\ \text{No hay intercepto})$$

se le informa que $\text{var}\,(u_i) = \sigma^2 X_i^2$. Demuestre que

$$\text{var}\,(\hat{\beta}_2) = \frac{\sigma^2 \sum X_i^4}{\left(\sum X_i^2\right)^2}$$

USO DIDÁCTICO

* Eric A. Hanushek y John E. Jackson, *Statistical Methods for Social Scientists*, Academic, Nueva York, 1977, p. 160.

Ejercicios empíricos

11.11. Con la información de la tabla 11.1, efectúe la regresión de la remuneración salarial promedio Y sobre la productividad promedio X, y considere el tamaño de la planta laboral como unidad de observación. Interprete sus resultados y vea si están de acuerdo con los presentados en (11.5.3).

a) De la regresión anterior, obtenga los residuos $\hat{u}_i$.

b) Según la prueba de Park, efectúe la regresión de $\ln \hat{u}_i^2$ sobre $\ln X_i$ y verifique la regresión (11.5.4).

c) Según el método de Glejser, efectúe la regresión de $|\hat{u}_i|$ sobre X_i y luego la regresión de $|\hat{u}_i|$ sobre $\sqrt{X_i}$. Comente sus resultados.

d) Encuentre la correlación de orden entre $|\hat{u}_i|$ y X_i, y comente sobre la naturaleza de la heteroscedasticidad presente en los datos, si existe.

11.12. La tabla 11.6 presenta información sobre la razón ventas/efectivo en las industrias manufactureras de Estados Unidos, clasificadas por tamaño de activos del establecimiento de 1971-I a 1973-IV. (Información trimestral.) La razón ventas/efectivo puede considerarse una medida de la velocidad del ingreso en el sector empresarial, es decir, el número de veces que circula un dólar.

a) Por cada tamaño de activos, calcule la media y la desviación estándar de la razón ventas/efectivo.

b) Grafique el valor de la media frente a la desviación estándar obtenida en *a*), con el tamaño de activos como unidad de observación.

c) Con un modelo de regresión apropiado, determine si la desviación estándar de la razón se incrementa con el valor de la media. De no ser así, ¿cómo interpreta el resultado?

d) Si hay una relación estadísticamente significativa entre los dos, ¿cómo transformaría la información de manera que no haya heteroscedasticidad?

11.13. *Prueba de homogeneidad de varianza de Bartlett.** Suponga que hay k varianzas muestrales independientes $s_1^2, s_2^2, \ldots, s_k^2$ con $f_1, f_2, \ldots, f_k$ gl, cada una proveniente de poblaciones normalmente distribuidas con media μ y varianza σ_i^2. Suponga además que deseamos probar la hipótesis nula H_0: $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_k^2 = \sigma^2$; es decir, cada varianza muestral es una estimación de la misma varianza poblacional σ^2.

Si la hipótesis nula es verdadera, entonces

$$s^2 = \frac{\sum_{i=1}^{k} f_i s_i^2}{\sum f_i} = \frac{\sum f_i s_i^2}{f}$$

TABLA 11.6
Tamaño de activos (millones de dólares)

Año y trimestre	1-10	10-25	25-50	50-100	100-250	250-1 000	1 000 +
1971-I	6.696	6.929	6.858	6.966	7.819	7.557	7.860
-II	6.826	7.311	7.299	7.081	7.907	7.685	7.351
-III	6.338	7.035	7.082	7.145	7.691	7.309	7.088
-IV	6.272	6.265	6.874	6.485	6.778	7.120	6.765
1972-I	6.692	6.236	7.101	7.060	7.104	7.584	6.717
-II	6.818	7.010	7.719	7.009	8.064	7.457	7.280
-III	6.783	6.934	7.182	6.923	7.784	7.142	6.619
-IV	6.779	6.988	6.531	7.146	7.279	6.928	6.919
1973-I	7.291	7.428	7.272	7.571	7.583	7.053	6.630
-II	7.766	9.071	7.818	8.692	8.608	7.571	6.805
-III	7.733	8.357	8.090	8.357	7.680	7.654	6.772
-IV	8.316	7.621	7.766	7.867	7.666	7.380	7.072

Fuente: *Quarterly Financial Report for Manufacturing Corporations*, Federal Trade Commission y Securities and Exchange Commission, gobierno de Estados Unidos, variables sobre diversos temas (calculadas).

* Véase "Properties of Sufficiency and Statistical Tests", *Proceedings of the Royal Society of London A*, vol. 160, 1937, p. 268.

USO DIDACTICO

constituye una estimación de la estimación común (agrupada) de la varianza poblacional σ^2, donde $f_i = (n_i - 1)$, con n_i como el número de observaciones en el i-ésimo grupo y donde $f = \sum_{i=1}^{k} f_i$.

Bartlett demostró que la hipótesis nula se prueba por la razón A/B, distribuida aproximadamente como la distribución χ^2 con $k - 1$ gl, donde

$$A = f \ln s^2 - \sum \left(f_i \ln s_i^2\right)$$

y

$$B = 1 + \frac{1}{3(k-1)}\left[\sum\left(\frac{1}{f_i}\right) - \frac{1}{f}\right]$$

Aplique la prueba de Bartlett a los datos de la tabla 11.1 y verifique que no se puede rechazar la hipótesis de que las varianzas poblacionales de la remuneración salarial son las mismas para cada tamaño de la planta laboral del establecimiento, en el nivel de significancia de 5%.

Nota: f_i, los gl para cada varianza muestral, es 9, pues n_i para cada muestra (es decir, clase de empleados) es 10.

11.14. Considere el siguiente modelo de regresión a través del origen:

$$Y_i = \beta X_i + u_i, \quad \text{para } i = 1, 2$$

Se tiene que $u_1 \sim N(0, \sigma^2)$ y $u_2 \sim N(0, 2\sigma^2)$, y que son estadísticamente independientes. Si $X_1 = +1$ y $X_2 = -1$, obtenga la estimación por mínimos cuadrados *ponderados* (MCP) de β y su varianza. Si en esta situación supuso de manera incorrecta que la dos varianzas de los errores son iguales (por ejemplo, iguales a σ^2), ¿cuál sería el estimador de MCO de β?, ¿y su varianza? Compare estas estimaciones con las obtenidas por el método de MCP. ¿Qué conclusión general deduce?*

11.15. La tabla 11.7 proporciona datos sobre 81 automóviles respecto de su MPG (millas promedio por galón), CF (caballos de fuerza de su motor), VOL (pies cúbicos de su cabina), VM (velocidad máxima en millas por hora) y su PS (peso del vehículo en cientos de lb).

a) Considere el siguiente modelo:

$$\text{MPG}_i = \beta_1 + \beta_2 \text{VM}_i + \beta_3 \text{CF}_i + \beta_4 \text{PS}_i + u_i$$

Estime los parámetros de este modelo e interprete los resultados. Desde el punto de vista económico, ¿tiene sentido?

b) ¿Esperaría que la varianza del error en el modelo anterior sea heteroscedástica? ¿Por qué?

c) Con la prueba de White determine si la varianza de error es heteroscedástica.

d) Obtenga los errores estándar de White consistentes con la heteroscedasticidad, así como los valores t, y compare los resultados con los obtenidos mediante MCO.

e) Si se establece heteroscedasticidad, ¿cómo puede transformar los datos de manera que en los datos transformados la varianza del error sea homoscedástica? Muestre los cálculos necesarios.

11.16. *Gasto alimentario en India.* En la tabla 2.8 se proporcionaron datos sobre el gasto en alimentos y el gasto total de 55 familias de India.

a) Haga la regresión del gasto alimentario sobre el gasto total y examine los residuos obtenidos en dicha regresión.

b) Grafique los residuos obtenidos en el inciso *a*) contra el gasto total y verifique si existe algún patrón sistemático.

* Adaptado de F.A.F. Seber, *Linear Regression Analysis*, John Wiley & Sons, Nueva York, 1977, p. 64.

USO DIDACTICO

TABLA 11.7 Datos de millajes de automóviles de pasajeros

Observación	MPG	VM	CF	VOL	PS	Observación	MPG	VM	CF	VOL	PS
1	65.4	96	49	89	17.5	42	32.2	106	95	106	30.0
2	56.0	97	55	92	20.0	43	32.2	109	102	92	30.0
3	55.9	97	55	92	20.0	44	32.2	106	95	88	30.0
4	49.0	105	70	92	20.0	45	31.5	105	93	102	30.0
5	46.5	96	53	92	20.0	46	31.5	108	100	99	30.0
6	46.2	105	70	89	20.0	47	31.4	108	100	111	30.0
7	45.4	97	55	92	20.0	48	31.4	107	98	103	30.0
8	59.2	98	62	50	22.5	49	31.2	120	130	86	30.0
9	53.3	98	62	50	22.5	50	33.7	109	115	101	35.0
10	43.4	107	80	94	22.5	51	32.6	109	115	101	35.0
11	41.1	103	73	89	22.5	52	31.3	109	115	101	35.0
12	40.9	113	92	50	22.5	53	31.3	109	115	124	35.0
13	40.9	113	92	99	22.5	54	30.4	133	180	113	35.0
14	40.4	103	73	89	22.5	55	28.9	125	160	113	35.0
15	39.6	100	66	89	22.5	56	28.0	115	130	124	35.0
16	39.3	103	73	89	22.5	57	28.0	102	96	92	35.0
17	38.9	106	78	91	22.5	58	28.0	109	115	101	35.0
18	38.8	113	92	50	22.5	59	28.0	104	100	94	35.0
19	38.2	106	78	91	22.5	60	28.0	105	100	115	35.0
20	42.2	109	90	103	25.0	61	27.7	120	145	111	35.0
21	40.9	110	92	99	25.0	62	25.6	107	120	116	40.0
22	40.7	101	74	107	25.0	63	25.3	114	140	131	40.0
23	40.0	111	95	101	25.0	64	23.9	114	140	123	40.0
24	39.3	105	81	96	25.0	65	23.6	117	150	121	40.0
25	38.8	111	95	89	25.0	66	23.6	122	165	50	40.0
26	38.4	110	92	50	25.0	67	23.6	122	165	114	40.0
27	38.4	110	92	117	25.0	68	23.6	122	165	127	40.0
28	38.4	110	92	99	25.0	69	23.6	122	165	123	40.0
29	46.9	90	52	104	27.5	70	23.5	148	245	112	40.0
30	36.3	112	103	107	27.5	71	23.4	160	280	50	40.0
31	36.1	103	84	114	27.5	72	23.4	121	162	135	40.0
32	36.1	103	84	101	27.5	73	23.1	121	162	132	40.0
33	35.4	111	102	97	27.5	74	22.9	110	140	160	45.0
34	35.3	111	102	113	27.5	75	22.9	110	140	129	45.0
35	35.1	102	81	101	27.5	76	19.5	121	175	129	45.0
36	35.1	106	90	98	27.5	77	18.1	165	322	50	45.0
37	35.0	106	90	88	27.5	78	17.2	140	238	115	45.0
38	33.2	109	102	86	30.0	79	17.0	147	263	50	45.0
39	32.9	109	102	86	30.0	80	16.7	157	295	119	45.0
40	32.3	120	130	92	30.0	81	13.2	130	236	107	55.0
41	32.2	106	95	113	30.0						

Notas:

VOL = pies cúbicos del espacio de cabina
CF = caballos de fuerza del motor
MPG = millas promedio por galón
VM = velocidad máxima, millas por hora
PS = peso del vehículo, cientos de libras
Observación = número de observaciones del automóvil (el nombre de los vehículos no se proporciona).

Fuente: U.S. Environmental Protection Agency, 1991, Informe EPA/AA/CTAB/91-02.

USO DIDÁCTICO

c) Si la gráfica del inciso *b*) sugiere heteroscedasticidad, aplique las pruebas de Park, Glejser y White para determinar si la sensación respecto de la heteroscedasticidad observada en *b*) se sustenta con estas pruebas.

d) Obtenga los errores estándar de White consistentes con la heteroscedasticidad y compárelos con los errores estándar de MCO. Decida si vale la pena corregir este ejemplo a causa de la heteroscedasticidad.

11.17. Repita el ejercicio 11.16, pero en esta ocasión efectúe la regresión del logaritmo del gasto alimentario sobre el logaritmo del gasto total. Si observa heteroscedasticidad en el modelo lineal del ejercicio 11.16 pero no en el modelo log-lineal, ¿a qué conclusión llega? Muestre todos los cálculos necesarios.

11.18. *Un atajo de la prueba de White.* Como mencionamos en el texto, la prueba de White consume grados de libertad si existen varias regresoras y se introducen todas las regresoras, sus términos cuadrados y sus productos cruzados. Por consiguiente, en vez de estimar las regresiones como la (11.5.22), ¿por qué no simplemente efectúa la siguiente regresión?

$$\hat{u}_i^2 = \alpha_1 + \alpha_2 \hat{Y}_i + \alpha_2 \hat{Y}_i^2 + \nu_i$$

donde $\hat{Y}_i$ son los valores estimados Y (es decir, la regresada) de cualquier modelo que se calcule. Después de todo, $\hat{Y}_i$ es tan sólo el promedio ponderado de las regresoras, donde los coeficientes estimados de la regresión sirven como ponderaciones.

Obtenga el valor R^2 de la regresión anterior y utilice (11.5.22) para probar la hipótesis de que no existe heteroscedasticidad.

Aplique la prueba anterior al ejemplo de gasto alimentario del ejercicio 11.6.

11.19. Reconsidere el ejemplo sobre IyD de la sección 11.7. Repita ese ejemplo con las ganancias como la regresora. *A priori,* ¿esperaría que los resultados fuesen diferentes de los que utilizan las ventas como regresoras?, ¿por qué?

11.20. La tabla 11.8 proporciona datos sobre la mediana de los salarios de catedráticos en estadística que laboraron en centros universitarios de investigación de Estados Unidos durante el año académico 2007.

a) Grafique la mediana de los salarios respecto de los rangos de años (como medida de los años de experiencia). Para propósitos de la gráfica, suponga que la mediana de los salarios está referida al punto medio del rango de años correpondiente. Por consiguiente, el salario de $124 578 del rango 4-5 está referido a 4.5 años del rango correspondiente, y así sucesivamente. Para el último grupo, suponga que el rango es 31-33.

b) Considere los siguientes modelos de regresión:

$$Y_i = \alpha_1 + \alpha_2 X_i + u_i \tag{1}$$

$$Y_i = \beta_1 + \beta_2 X_i + \beta_3 X_i^2 + \nu_i \tag{2}$$

TABLA 11.8
Mediana de los salarios de los catedráticos en estadística, 2007

Fuente: American Statistical Association, "2007 Salary Report".

Rango de años	Cuenta	Mediana
0 a 1	40	$101 478
2 a 3	24	102 400
4 a 5	35	124 578
6 a 7	34	122 850
8 a 9	33	116 900
10 a 14	73	119 465
15 a 19	69	114 900
20 a 24	54	129 072
25 a 30	44	131 704
31 o más	25	143 000

USO DIDÁCTICO

donde Y = mediana del salario, X = año en el rango (medido como el punto medio del intervalo), y u y v son los términos de error. ¿Puede justificar por qué el modelo (2) sería preferible al modelo (1)? A partir de estos datos, estime los modelos.

c) Si observa heteroscedasticidad en el modelo (1) pero no en el modelo (2), ¿a qué conclusiones llega? Muestre los cálculos necesarios.

d) Si observa heteroscedasticidad en el modelo (2), ¿cómo puede transformar los datos de manera que en el modelo transformado no existiera heteroscedasticidad?

11.21. Tiene la siguiente información:

$$SCR_1 \text{ basado en las primeras 30 observaciones} = 55, \text{ gl} = 25$$
$$SCR_2 \text{ basado en las últimas 30 observaciones} = 140, \text{ gl} = 25$$

Realice la prueba de heteroscedasticidad de Goldfeld-Quandt en el nivel de significancia de 5%.

11.22. La tabla 11.9 presenta información acerca de los precios de acciones (Y) y los precios al consumidor (X) expresados en cambios porcentuales anuales para un corte transversal de 20 países.

a) Grafique los datos en un diagrama de dispersión.

b) Efectúe la regresión de Y sobre X y examine los residuos de esta regresión. ¿Qué observa?

c) Como los datos de Chile parecen atípicos, repita la regresión en *b*) sin la información sobre Chile. Ahora examine los residuos de esta regresión. ¿Qué observa?

d) Si, con base en los resultados de *b*), concluye que hubo heteroscedasticidad en la varianza del error, pero con base en los resultados de *c*) modifica este resultado, ¿qué conclusiones generales obtiene?

TABLA 11.9
Precios de acciones y precios al consumidor, periodo posterior a la Segunda Guerra Mundial (hasta 1969)

Fuente: Phillip Cagan, *Common Stock Values and Inflation: The Historical Record of Many Countries*, National Bureau of Economic Research, suplemento, marzo de 1974, tabla 1, p. 4.

	Tasa de crecimiento, % anual	
País	Precio de acciones, Y	Precio al consumidor, X
1. Australia	5.0	4.3
2. Austria	11.1	4.6
3. Bélgica	3.2	2.4
4. Canadá	7.9	2.4
5. Chile	25.5	26.4
6. Dinamarca	3.8	4.2
7. Finlandia	11.1	5.5
8. Francia	9.9	4.7
9. Alemania	13.3	2.2
10. India	1.5	4.0
11. Irlanda	6.4	4.0
12. Israel	8.9	8.4
13. Italia	8.1	3.3
14. Japón	13.5	4.7
15. México	4.7	5.2
16. Países Bajos	7.5	3.6
17. Nueva Zelanda	4.7	3.6
18. Suecia	8.0	4.0
19. Reino Unido	7.5	3.9
20. Estados Unidos	9.0	2.1

USO DIDÁCTICO

11.23 La tabla 11.10 del sitio Web presenta datos sobre salarios y otros aspectos relacionados de 447 ejecutivos de 500 empresas de *Fortune*. Los datos incluyen salario = salario y gratificaciones de 1999; remtot = remuneración total del director general en 1999; antigüedad = número de años como director general (0 si tiene menos de 6 meses en el cargo); edad = edad del director general; ventas = total de los ingresos por ventas de la empresa en 1998; utilidades = utilidades de la empresa en 1998; y activos = total de activos de la empresa en 1998.

a) Estime la siguiente regresión con base en estos datos y obtenga el estadístico de Breusch-Pagan-Godfrey para ver si hay heteroscedasticidad:

$$\text{salario}_i = \beta_1 + \beta_2\text{antigüedad}_i + \beta_3\text{edad}_i + \beta_4\text{ventas}_i + \beta_5\text{utilidades}_i + \beta_6\text{activos}_i + u_i$$

¿Parece existir un problema de heteroscedasticidad?

b) Ahora cree un segundo modelo con ln(salario) como variable dependiente. ¿Observa alguna mejora en la heteroscedasticidad?

c) Cree diagramas de dispersión del salario sobre cada variable independiente. ¿Puede discernir qué variables contribuyen al problema? ¿Qué propondría ahora para resolverlo? ¿Cuál es el modelo final?

USO DIDACTICO

Apéndice 11A

11A.1 Prueba de la ecuación (11.2.2)

Del apéndice 3A, sección 3A.3, tenemos

$$\begin{aligned}\text{var}(\hat{\beta}_2) &= E\left(k_1^2u_1^2 + k_2^2u_2^2 + \cdots + k_n^2u_n^2 + 2 \text{ términos de productos cruzados}\right)\\ &= E\left(k_1^2u_1^2 + k_2^2u_2^2 + \cdots + k_n^2u_n^2\right)\end{aligned}$$

pues los valores esperados de los términos de productos cruzados son cero debido al supuesto de no correlación serial,

$$\text{var}(\hat{\beta}_2) = k_1^2E\left(u_1^2\right) + k_2^2E\left(u_2^2\right) + \cdots + k_n^2E\left(u_n^2\right)$$

porque se conocen las k_i. (¿Por qué?)

$$\text{var}(\hat{\beta}_2) = k_1^2\sigma_1^2 + k_2^2\sigma_2^2 + \cdots + k_n^2\sigma_n^2$$

pues $E(u_i^2) = \sigma_i^2$.

$$\begin{aligned}\text{var}(\hat{\beta}_2) &= \sum k_i^2\sigma_i^2\\ &= \sum\left[\left(\frac{x_i}{\sum x_i^2}\right)^2\sigma_i^2\right] \qquad \text{pues } k_i = \frac{x_i}{\sum x_i^2} \qquad (11.2.2)\\ &= \frac{\sum x_i^2\sigma_i^2}{\left(\sum x_i^2\right)^2}\end{aligned}$$

11A.2 Método de mínimos cuadrados ponderados

Para ilustrar el método utilizamos el modelo de dos variables $Y_i = \beta_1 + \beta_2X_i + u_i$. El método de mínimos cuadrados no ponderado reduce

$$\sum \hat{u}_i^2 = \sum(Y_i - \hat{\beta}_1 - \hat{\beta}_2X_i)^2 \qquad (1)$$

para obtener las estimaciones, mientras que el método de mínimos cuadrados ponderados reduce la suma de cuadrados ponderada de los residuales:

$$\sum w_i \hat{u}_i^2 = \sum w_i (Y_i - \hat{\beta}_1^* - \hat{\beta}_2^* X_i)^2 \tag{2}$$

donde β_1^* y β_2^* son los estimadores de mínimos cuadrados ponderados y las ponderaciones w_i son tales que

$$w_i = \frac{1}{\sigma_i^2} \tag{3}$$

es decir, las ponderaciones son inversamente proporcionales a la varianza de u_i o Y_i condicional a las X_i dadas, y queda entendido que $\text{var}(u_i \mid X_i) = \text{var}(Y_i \mid X_i) = \sigma_i^2$.

Al diferenciar (2) respecto de $\hat{\beta}_1^*$ y $\hat{\beta}_2^*$, obtenemos

$$\frac{\partial \sum w_i \hat{u}_i^2}{\partial \beta_1^*} = 2 \sum w_i (Y_i - \hat{\beta}_1^* - \hat{\beta}_2^* X_i)(-1)$$

$$\frac{\partial \sum w_i \hat{u}_i^2}{\partial \beta_2^*} = 2 \sum w_i (Y_i - \hat{\beta}_1^* - \hat{\beta}_2^* X_i)(-X_i)$$

Igualamos a cero las expresiones anteriores y obtenemos las dos siguientes ecuaciones normales:

$$\sum w_i Y_i = \hat{\beta}_1^* \sum w_i + \hat{\beta}_2^* \sum w_i X_i \tag{4}$$

$$\sum w_i X_i Y_i = \hat{\beta}_1^* \sum w_i X_i + \hat{\beta}_2^* \sum w_i X_i^2 \tag{5}$$

Observe la semejanza entre estas ecuaciones normales y las ecuaciones normales de los mínimos cuadrados no ponderados.

Resolvemos estas ecuaciones simultáneamente y obtenemos

$$\hat{\beta}_1^* = \bar{Y}^* - \hat{\beta}_2^* \bar{X}^* \tag{6}$$

y

$$\hat{\beta}_2^* = \frac{\left(\sum w_i\right)\left(\sum w_i X_i Y_i\right) - \left(\sum w_i X_i\right)\left(\sum w_i Y_i\right)}{\left(\sum w_i\right)\left(\sum w_i X_i^2\right) - \left(\sum w_i X_i\right)^2} \qquad (11.3.8) = (7)$$

La varianza de $\hat{\beta}_2^*$ que aparece en (11.3.9) se obtiene en la forma de la varianza de $\hat{\beta}_2$ que aparece en el apéndice 3A, sección 3A.3.

Nota: $\bar{Y}^* = \sum w_i Y_i / \sum w_i$ y $\bar{X}^* = \sum w_i X_i / \sum w_i$. Como se verifica fácilmente, estas medias ponderadas coinciden con las medias usuales o sin ponderar $\bar{Y}$ y $\bar{X}$ cuando $w_i = w$, una constante, para todo i.

11A.3 Prueba de que $E(\hat{\sigma}^2) \neq \sigma^2$ en presencia de heteroscedasticidad

Considere el modelo de dos variables:

$$Y_i = \beta_1 + \beta_2 X_i + u_i \tag{1}$$

donde $\text{var}(u_i) = \sigma_i^2$.

Ahora

$$\begin{aligned} \hat{\sigma}^2 = \frac{\sum \hat{u}_i^2}{n-2} &= \frac{\sum (Y_i - \hat{Y}_i)^2}{n-2} = \frac{\sum [\beta_1 + \beta_2 X_i + u_i - \hat{\beta}_1 - \hat{\beta}_2 X_i]^2}{n-2} \\ &= \frac{\sum [-(\hat{\beta}_1 - \beta_1) - (\hat{\beta}_2 - \beta_2) X_i + u_i]^2}{n-2} \end{aligned} \tag{2}$$

USO DIDÁCTICO

Observe que $(\hat{\beta}_1 - \beta_1) = -(\hat{\beta}_2 - \beta_2)\bar{X} + \bar{u}$, y sustituya lo anterior en (2); al tomar las esperanzas en ambos miembros, tenemos:

$$\begin{aligned} E(\hat{\sigma}^2) &= \frac{1}{n-2}\left\{-\sum x_i^2 \operatorname{var}(\hat{\beta}_2) + E\left[\sum (u_i - \bar{u})^2\right]\right\} \\ &= \frac{1}{n-2}\left[-\frac{\sum x_i^2 \sigma_i^2}{\sum x_i^2} + \frac{(n-1)\sum \sigma_i^2}{n}\right] \end{aligned} \tag{3}$$

donde se utiliza (11.2.2).

Como puede observar de (3), si existe homoscedasticidad, es decir, $\sigma_i^2 = \sigma^2$ para cada i, $E(\hat{\sigma}^2) = \sigma^2$. En consecuencia, el valor esperado de $\hat{\sigma}^2 = \sum \hat{u}^2/(n-2)$ calculado de manera convencional no será igual a la verdadera σ^2 en presencia de heteroscedasticidad.[1]

11A.4 Errores estándar robustos de White

Para dar alguna idea respecto de los errores estándar de White corregidos por heteroscedasticidad, considere el siguiente modelo de regresión de dos variables:

$$Y_i = \beta_1 + \beta_2 X_i + u_i \qquad \operatorname{var}(u_i) = \sigma_i^2 \tag{1}$$

Como se muestra en (11.2.2),

$$\operatorname{var}(\hat{\beta}_2) = \frac{\sum x_i^2 \sigma_i^2}{\left(\sum x_i^2\right)^2} \tag{2}$$

Como σ_i^2 no son directamente observables, White sugiere utilizar $\hat{u}_i^2$, el residuo al cuadrado para cada i, en lugar de σ_i^2, y calcular var $(\hat{\beta}_2)$ de la siguiente forma:

$$\operatorname{var}(\hat{\beta}_2) = \frac{\sum x_i^2 \hat{u}_i^2}{\left(\sum x_i^2\right)^2} \tag{3}$$

White demostró que (3) es un estimador consistente de (2); es decir, conforme el tamaño de la muestra crece de manera indefinida, (3) converge a (2).[2]

A propósito, observe que si el paquete de software utilizado no contiene el procedimiento de errores estándar robustos de White, se puede efectuar como se muestra en (3): primero la regresión usual por MCO, después se obtienen los residuos de esta regresión y luego se utiliza la fórmula (3).

El procedimiento de White se generaliza al modelo de regresión de k variables:

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \cdots + \beta_k X_{ki} + u_i \tag{4}$$

La varianza de todo coeficiente de regresión parcial, por ejemplo $\hat{\beta}_j$, se obtiene así:

$$\operatorname{var}(\hat{\beta}_j) = \frac{\sum \hat{w}_{ji}^2 \hat{u}_i^2}{\left(\sum \hat{w}_{ji}^2\right)^2} \tag{5}$$

donde $\hat{u}_i$ son los residuos obtenidos de la regresión original (4), y $\hat{w}_j$ son los residuos proporcionados por la regresión (auxiliar) de la regresora X_j sobre las regresoras restantes en (4).

Obvio, es un procedimiento tardado, pues se debe estimar (5) para cada variable X. Por supuesto, todo este trabajo se evita si se cuenta con un software estadístico que lo haga de manera rutinaria. En la actualidad, programas como PC-GIVE, *Eviews*, MICROFIT, SHAZAM, STATA y LIMDEP calculan de manera muy sencilla los errores estándar robustos de White con heteroscedasticidad.

USO DIDÁCTICO

[1] Se pueden obtener mayores detalles en Jan Kmenta, *Elements of Econometrics*, 2a. ed., Macmillan, Nueva York, 1986, pp. 276-278.

[2] Para mayor precisión, n veces (3) converge en probabilidad a $E[(X_i - \mu_X)^2 u_i^2]/(\sigma_X^2)^2$, que es la probabilidad límite de n veces (2), donde n es el tamaño de la muestra, μ_x es el valor esperado de X y σ_X^2 es la varianza (poblacional) de X. Para mayores detalles, véase Jeffrey M. Wooldridge, *Introductory Econometrics: A Modern Approach*, South-Western, 2000, p. 250.