

# Unidad 4a.

# Heterocedasticidad

Prof. Dra. Eva Romero  
Ramos



**Universidad  
Europea**

LAUREATE INTERNATIONAL UNIVERSITIES



- 1. Heterocedasticidad: concepto, causas, consecuencias.
- 2. Ejemplo de análisis de heterocedasticidad.
- 3. Medidas de corrección de la heterocedasticidad.

# 1. Heterocedasticidad: concepto, causas, consecuencias.



# Heterocedasticidad

El correcto funcionamiento del modelo de regresión lineal está basado en el cumplimiento de una serie de hipótesis entre las que destacan las relacionadas con el error, que son:

- El término error es completamente aleatorio y **sigue una distribución normal**, de esperanza 0.
- La varianza del error es constante a lo largo de las observaciones del modelo ( $\text{Var}[\varepsilon_i]=\sigma^2$ ).
- El término error no está correlacionado entre los elementos del modelo y es independiente también de la variable explicativa.

La heteroscedasticidad es un problema que surge con el incumplimiento de la segunda de estas hipótesis, la que implica que la varianza del error debe ser constante a lo largo de las observaciones.

# Heterocedasticidad



## Definición.-

La heterocedasticidad se define como uno de los problemas que pueden presentar los modelos de regresión lineal, cuando las varianzas de sus perturbaciones o errores no son constantes para todas las observaciones de la muestra.

Los modelos de regresión lineal presentarán problemas de heterocedasticidad cuando los datos que utilizan provengan de poblaciones heterogéneas con varianzas distintas.

Si la varianza de las perturbaciones o errores es constante, se dice que el modelo es homocedastico, y que hay homogeneidad en la muestra.

## Heterocedasticidad

- La heterocedasticidad *se presenta con mayor frecuencia en muestras de datos de corte transversal*, ya que en ellas los items pueden ser individuos, empresas, países... y no suelen presentar un comportamiento homogéneo.
- Es frecuente también encontrar este problema cuando los *datos se obtienen como agregados o promedios* de datos individuales.
- El ejemplo clásico para ilustrar la presencia de heterocedasticidad en un modelo que explique el comportamiento de variables como el *ahorro* o el *gasto* a partir de la *renta*. En este caso la renta explicará el ahorro por ejemplo con un cierto margen de error. Pero ese error no será igual para los individuos que tienen rentas altas y para los que tienen rentas bajas, puesto que los primeros pueden elegir ahorrar mucho o poco, mientras que los segundos nunca podrán ahorrar mucho porque no tienen acceso a grandes cantidades. Lo mismo pasaría si tomamos como ejemplo el gasto.

## Consecuencias de la heterocedasticidad

- En presencia de heterocedasticidad, **los estimadores de mínimos cuadrados de los parámetros del modelo siguen siendo insesgados y consistentes**. La heterocedasticidad no afecta al  $R^2$  ni al  $\bar{R}^2$ .
- Las principales consecuencias de la heterocedasticidad son:
  - 1. La matriz de varianzas-covarianzas no se calcula correctamente.** Sabemos que esta matriz se obtiene mediante la expresión:

$$\text{Var}(\hat{\beta}) = \sigma_{\epsilon}^2 (X^t X)^{-1}$$

La expresión anterior utiliza la varianza del error asumiéndola constante por lo que si no lo es, la matriz de varianzas-covarianzas de los estimadores no estará bien calculada. Así, en presencia de heterocedasticidad, se hace necesario calcular la estimación de la matriz de varianzas-covarianzas de los estimadores por medio de la expresión:

$$\text{Var}(\hat{\beta}) = \sigma_{\epsilon}^2 (X^t X)^{-1} X^t \Sigma X (X^t X)^{-1}$$

Siendo  $\Sigma$  una matriz que representa el comportamiento de la varianza del error a lo largo de la muestra.

# Consecuencias de la heterocedasticidad



2. **Perdida de eficiencia de los estimadores de mínimos cuadrados ordinarios.** En presencia de heterocedasticidad las varianzas de los parámetros no pueden calcularse con la expresión anterior, ya que se debe tener en cuenta el comportamiento de la varianza del error en el cálculo. Esto incrementa en cualquier caso las varianzas haciendo a las estimaciones menos eficientes.

3. Cómo las estimaciones de las varianzas de los estimadores se hacen más grandes y *menos fiables* en presencia de heterocedasticidad, *los contrastes de significatividad individual y global* también serán menos fiables, ya que utilizan estas varianzas para calcular sus estadísticos de contraste.

- Por lo demás los estimadores de mínimos cuadrados siguen siendo los mejores estimadores que pueden obtenerse. *Siguen siendo insesgados, pero dejan de ser eficientes.*





## 2. Ejemplo de análisis de heterocedasticidad.

## Ejemplo de análisis de heterocedasticidad



El archivo `gasto_ocio.gdt` contiene datos de una muestra de 50 individuos para las siguientes variables:

GASTOOCIO: Gasto mensual en ocio de cada individuo.

INGRESOS: Ingresos mensuales por individuo.

Se pretende estudiar el gasto en ocio de los individuos a partir de su renta.

# Ejemplo de análisis de heterocedasticidad



La estimación del modelo en Gretl es la siguiente:

```
gretl: modelo 1
Archivo  Editar  Contrastes  Guardar  Gráficos  Análisis  LaTeX
Modelo 1: MCO, usando las observaciones 1-50
Variable dependiente: GASTOOCIO

-----
                Coeficiente   Desv. Típica   Estadístico t   valor p
-----
const           53.9422         19.7912         2.726           0.0089   ***
INGRESOS        0.0812188         0.00329542     24.65           6.73e-029 ***

Media de la vble. dep.  491.4786   D.T. de la vble. dep.  226.2305
Suma de cuad. residuos 183661.7   D.T. de la regresión   61.85697
R-cuadrado          0.926765   R-cuadrado corregido   0.925239
F(1, 48)            607.4223   Valor p (de F)         6.73e-29
Log-verosimilitud   -276.1676   Criterio de Akaike     556.3352
Criterio de Schwarz  560.1593   Crit. de Hannan-Quinn  557.7915

Contraste de normalidad de los residuos -
Hipótesis nula: el error se distribuye normalmente
Estadístico de contraste: Chi-cuadrado(2) = 5.07226
con valor p = 0.079172
```

## Ejemplo de análisis de heterocedasticidad



- Podemos observar que el modelo planteado es significativo y consigue explicar un 92,67% de la variabilidad de la variable Gasto en Ocio.
- No hay evidencias suficientes para rechazar la hipótesis nula de normalidad para los residuos del modelo, a nivel de significación 0,05, por lo que podemos asumir que el error sigue una distribución normal.
- La variable ingreso es significativa y al ser la única variable del modelo no hay problemas de multicolinealidad.
- Analizaremos ahora la posible existencia de problemas de heteroscedasticidad en el modelo.

## Ejemplo de análisis de heterocedasticidad

- Para detectar un posible problema de heterocedasticidad comenzaremos utilizando los *métodos gráficos*, que nos darán una idea del problema, si lo hubiera, y del comportamiento de la heterocedasticidad.
- Posteriormente veremos dos contrastes de hipótesis para detectar la presencia de heterocedasticidad:
  - El *contraste de White*.
  - El *contraste de Goldfeld y Quant*.

# Métodos gráficos para detectar problemas de heterocedasticidad



- Los métodos gráficos para la detección de problemas de heterocedasticidad, consisten en la *observación de los gráficos de dispersión entre el valor absoluto del error, o su cuadrado y las variables explicativas.*
- Estos gráficos nos darán una idea del comportamiento del error del modelo ante variaciones en las variables explicativas.
- *Debemos observar si la dispersión es constante o por el contrario van cambiando a lo largo de las observaciones de la muestra.*
- Observaremos el comportamiento de los residuos como muestra del error del modelo.

# Métodos gráficos para detectar problemas de heterocedasticidad



- Debemos obtener el gráfico de dispersión del valor absoluto de los residuos o su cuadrado frente a las variables explicativas.
- Obtendremos para nuestro ejemplo el gráfico de dispersión del cuadrado de los residuos frente a la variable ingresos. Para ello seguiremos los siguientes pasos:

1) Guardar el cuadrado de los residuos. Desde el menú del modelo:

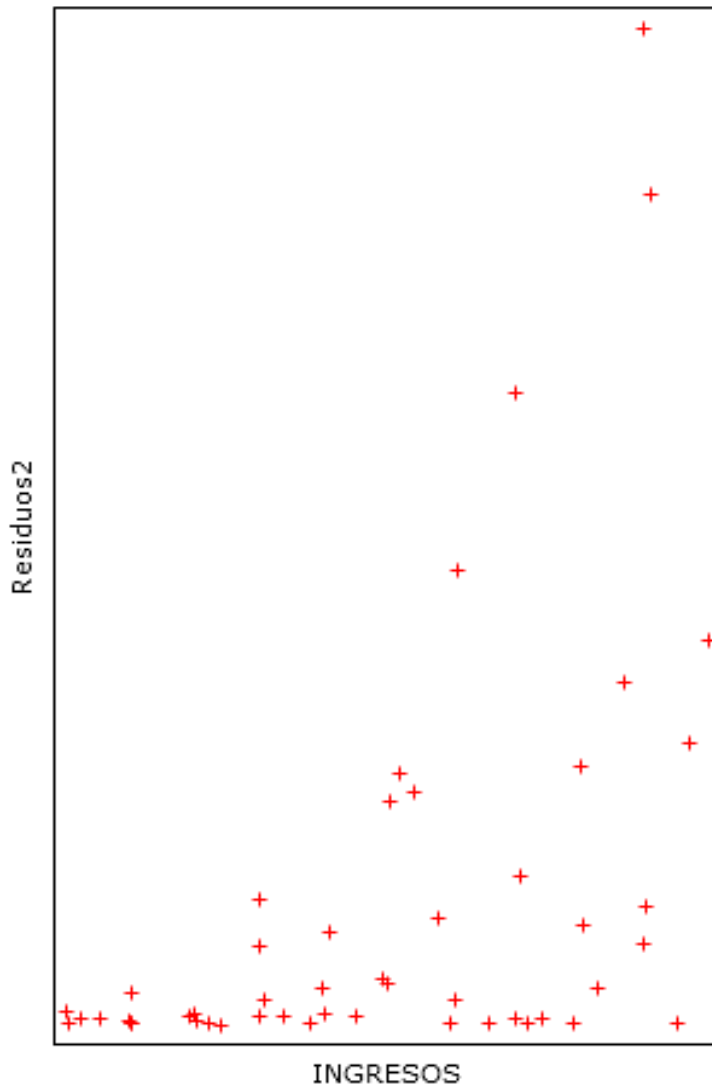
**Guardar -> Residuos al cuadrado**

2) Realizar el gráfico de dispersión. Desde el menú principal:

**Ver -> Gráficos Múltiples -> Gráficos X-Y (Scatters)**

Seleccionando como variable para el eje Y los residuos y para el eje X la variable dependiente correspondiente.

# Métodos gráficos para detectar problemas de heterocedasticidad



- Vemos en el gráfico como la dispersión del cuadrado error aumenta a medida que aumentan los ingresos del individuo.
- Esto es un claro indicio de la presencia de problemas de heterocedasticidad en el modelo.
- En modelos homocedasticos la dispersión del cuadrado del error sería la misma para todos los valores de la variable ingresos.



# Métodos gráficos para detectar problemas de heterocedasticidad



- Si hubiésemos obtenido el gráfico utilizando el valor absoluto de los residuos frente a los ingresos, observaríamos un gráfico muy similar y obtendríamos las mismas conclusiones.
- Para obtener el gráfico usando el valor absoluto en Gretl es necesario guardar los residuos y generar una nueva serie con su valor absoluto antes de obtener el gráfico de dispersión.
- La ventaja de los gráficos frente a los contrastes es que nos permiten observar el comportamiento de la heterocedasticidad, es decir, si crece, decrece o que forma toma.
- En nuestro ejemplo vemos claramente que a medida que aumentan los ingresos, aumenta la dispersión del error.

# Métodos gráficos para detectar problemas de heterocedasticidad



- Vemos en nuestro ejemplo que el gráfico presenta *forma de embudo*, mostrando un crecimiento lineal de la dispersión del cuadrado del error con el aumento de los ingresos.
- Si tuviésemos más de una variable explicativa, deberíamos decidir cuál puede estar causando un problema de heterocedasticidad.
- Si no tenemos indicios sobre que variable puede generar heterocedasticidad en el modelo es recomendable realizar el gráfico con todas las variables explicativas del modelo.
- Otra opción es *utilizar las estimaciones* que hace el modelo de la variable dependiente, ya que son una combinación lineal de todas las variables explicativas incluidas.
- En Gretl se pueden guardar los valores estimados desde el menú del modelo: **Guardar -> Valores estimados**.

# Contrastes de hipótesis para la detección de la heterocedasticidad



- Además de los métodos gráficos existen numerosos contrastes de hipótesis para verificar la presencia de heterocedasticidad en el modelo.
- Gretl tiene implementados el contraste de White, el contraste de Breusch-Pagan y el de Koenker.
- Estudiaremos el contraste de White y el Contraste de Goldfeld y Quandt.

## Contraste de White

El *contrate de White* es el contraste más general para la detección de presencia de heteroscedasticidad en un modelo de regresión lineal múltiple, ya que no requiere una especificación concreta de cómo es la heterocesticidad bajo la hipótesis nula.

Simplemente contrasta:

$H_0: \sigma_i^2 = \sigma^2$  para todo  $i \equiv$  Homocedasticidad

$H_1: \sigma_i^2 \neq \sigma^2$  para todo  $i \equiv$  Heterocesticidad

## Contraste de White

- Este contraste *plantea un modelo auxiliar que pretende explicar el comportamiento del cuadrado de los residuos a partir de las variables explicativas, de sus cuadrados y productos cruzados*, este último si lo solicitamos.
- Si el modelo es homocedástico, este modelo auxiliar no será **significativo** ya que en un modelo homocedastico no se podrá explicar el cuadrado de los residuos a partir de las variables explicativas, sus cuadrados y sus productos cruzados.
- Si por el contrario el modelo presenta un problema de **heterocedasticidad**, este modelo auxiliar **sí será significativo**.
- Cuando el modelo es homocedástico, se cumple que:

$$n \cdot R^2 \sim \chi_{k-1}^2$$

## Contraste de White

- En Gretl tenemos este contraste en las opciones del modelo:

**Contrastes -> Heterocedasticidad -> Contraste de White**

```

gretl: contraste LM(heterocedasticidad)

Contraste de heterocedasticidad de White
MCO, usando las observaciones 1-50
Variable dependiente: uhat^2

      Coeficiente    Desv. Típica    Estadístico t    valor p
-----
const      558.357      3243.35          0.1722          0.8641
INGRESOS   -0.444148         1.37237         -0.3236         0.7476
sq_INGRESOS 0.000152700      0.000125281     1.219           0.2290

R-cuadrado = 0.267028

Estadístico de contraste: TR^2 = 13.351384,
con valor p = P(Chi-cuadrado(2) > 13.351384) = 0.001261
  
```

En este caso vemos que con un p-valor de 0,001261 a nivel de significación 0,05 tenemos evidencias suficientes para rechazar la hipótesis nula de homocedasticidad. *Hay por tanto, un problema de heterocedasticidad en este modelo.*

## Contraste de White



- Este modelo no requiere la especificación de la una hipótesis alternativa que indique la forma que presenta la heterocedasticidad y esto se considera una ventaja, por su flexibilidad y sencillez.
- No obstante, la no especificación de una hipótesis alternativa para la heterocedasticidad es también un inconveniente, ya que si encontramos un problema de heterocedasticidad, no tenemos información sobre el comportamiento de esta.

## Contraste de Goldfeld y Quandt



El *contraste de Goldfeld y Quandt* contraste contrasta la hipótesis nula de homocedasticidad frente a la alternativa, que indica que la heterocedasticidad sigue una función monótona, (es creciente o decreciente), es decir:

Contrasta:

$H_0: \sigma_i^2 = \sigma^2$  para todo  $i \equiv$  Homocedasticidad

$H_1: \sigma_i^2 = h(x_i)$ , para todo  $i \equiv$  Heterocedasticidad que sigue una función monótona  $h(x)$ .



## Contraste de Goldfeld y Quandt

Como el contraste no está implementado en Gretl, para realizarlo debemos seguir los siguientes pasos:

1. *Ordenar la muestra* en orden ascendente de los valores de la variable causante de la heterocedasticidad. Si hubiese más de una variable causante de la heterocedasticidad, ordenaremos según los valores de la estimación de la variable dependiente. Para ello: **Datos -> Ordenar datos**
2. *Haremos dos muestras*, una con los valores más pequeños de la variable para la que hemos ordenado los datos (en nuestro caso INGRESOS) y otra con los más elevados.  
*Eliminaremos las “p” observaciones centrales* de la muestra, de forma que ésta quede dividida en dos submuestras de tamaño  $(N-p)/2$ .

## Contraste de Goldfeld y Quandt



Elegiremos el valor de  $p$  de forma arbitraria, teniendo en cuenta que cada submuestra debe tener un tamaño superior al de los parámetros a estimar en el modelo.

Es habitual, siempre que sea posible eliminar aproximadamente la cuarta parte de la muestra disponible.

Con 50 observaciones, es razonable eliminar los 12 valores centrales, es decir  $p=12$ , dejando una primera muestra que tendrá las observaciones de la primera a la número 19 y la segunda de la 32 a la última.

- 3. Estimar dos modelos de regresión múltiple*, uno para cada submuestra y *guardar los valores de la desviación típica de los residuos de la regresión.*

## Contraste de Goldfeld y Quandt



Para estimar modelos con parte de la muestra debemos usar la opción: **muestra -> Establecer** rango desde el menú principal.

Para guardar las desviaciones típicas de los residuos de la regresión haremos: **Guardar -> Desviación típica de la regresión** desde el menú del modelo.

4. Con los valores de las desviaciones típicas de ambos modelos, generaremos nuestro estadístico de contraste, desde la consola de Gretl:

```
genr EstCon=(DTGrande/DTPeques)^2
```

## Contraste de Goldfeld y Quandt



Bajo la hipótesis nula, este estadístico sigue una distribución F de Fisher-Snedecor con  $(N-p)/2-(K+1)$  grados de libertad (en nuestro ejemplo 17).

5. Para obtener el p-valor haremos:

*genr* pvalor=`pvalue(F,17,17,EstCon)`

En nuestro ejemplo se obtiene un valor de  $5,036E-7$ , por lo que debemos *rechazar la hipótesis nula de homocedasticidad*.

## Contraste de Goldfeld y Quandt

Una vez detectada la heterocedasticidad, podemos comparar directamente los valores obtenidos para las desviaciones típicas de los residuos de la regresión en ambas muestras, de modo que:

- Si  $DT_{Grande} > DT_{Peques}$ , observamos que cuando aumentan los valores de la variable INGRESOS, la varianza del error aumenta, por lo que es creciente.
- Si  $DT_{Grande} < DT_{Peques}$  vemos que cuando los valores de la variable INGRESOS aumenta, la varianza del error disminuye, por lo que será decreciente.

En nuestro caso vemos que es creciente, como se observaba en los gráficos.



# 3. Medidas de corrección de la heterocedasticidad.

## Medidas de corrección de la heterocedasticidad.



Analizaremos dos posibles soluciones a la heterocedasticidad:

- Transformar las variables, por ejemplo tomado Logaritmos.
- Estimar por mínimos cuadrados ponderados.

# Medidas de corrección de la heterocedasticidad - Logaritmos



- Para resolver posibles problemas de heterocedasticidad en el modelo, en ocasiones se recurre a realizar transformaciones en las variables, que puedan eliminar estos problemas.
- La *transformación logarítmica* es de las más utilizadas, ya que al tomar logaritmos las series se suavizan y los problemas de heterocedasticidad tienden a desaparecer.
- Esta transformación consiste simplemente en tomar logaritmos en una o varias variables del modelo.
- Es importante tener en cuenta que con este tipo de transformaciones *la interpretación de los parámetros del modelo cambia, pasando a ser en términos porcentuales.*



# Medidas de corrección de la heterocedasticidad - Logaritmos



Interpretación de los parámetros en presencia de logaritmos:

Modelo	Variable dependiente	Variable independiente	Interpretación del parámetro
Nivel-nivel	y	x	$\Delta y = \beta_i \Delta x$
Nivel-log	y	Log(x)	$\Delta y = (\beta_i / 100) \% \Delta x$
Log-nivel	Log(y)	x	$\Delta \% y = (100 \beta_i) \Delta x$
Log-log	Log(y)	Log(x)	$\Delta \% y = \beta_i \% \Delta x$

(Wooldridge, J. Introducción a la econometría: un enfoque moderno. 2006. Ed. Thomson. Pg. 49)

# Medidas de corrección de la heterocedasticidad - Logaritmos



Interpretación de los parámetros en presencia de logaritmos:

- $\Delta y = \beta_i \Delta x$  : Cuando X aumenta en una unidad, Y aumenta en  $\beta$  unidades.
- $\Delta y = (\beta_i/100)\% \Delta x$  : Cuando X aumenta en un 1%, Y aumenta en  $\beta/100$  unidades.
- $\Delta \%y = (100\beta_i)\Delta x$  : Cuando X aumenta en una unidad Y aumenta en un  $100 \cdot \beta$  %.
- $\Delta \%y = \beta_i \% \Delta x$  : Cuando X aumenta en un 1%, Y aumenta en un  $\beta$  %.

# Medidas de corrección de la heterocedasticidad - Logaritmos



- Probaremos en nuestro ejemplo a resolver la heteroscedasticidad tomando logaritmos en ambas variables:

```
gretl: modelo 4
-----
Archivo  Editar  Contrastes  Guardar  Gráficos  Análisis  LaTeX
-----
Modelo 4: MCO, usando las observaciones 1-50
Variable dependiente: LOGGASTOOCIO

-----
                Coeficiente  Desv. Típica  Estadístico t  valor p
-----
const          -0.956043    0.198149      -4.825        1.46e-05 ***
LOGINGRESOS    0.833879    0.0234578     35.55        3.88e-036 ***

Media de la vble. dep.  6.067135    D.T. de la vble. dep.  0.554473
Suma de cuad. residuos  0.551283    D.T. de la regresión  0.107168
R-cuadrado      0.963405    R-cuadrado corregido  0.962643
F(1, 48)       1263.668    Valor p (de F)        3.88e-36
Log-verosimilitud  41.74133    Criterio de Akaike    -79.48266
Criterio de Schwarz -75.65861    Crit. de Hannan-Quinn -78.02644

Log-verosimilitud para GASTOOCIO = -261.615

Contraste de normalidad de los residuos -
Hipótesis nula: el error se distribuye normalmente
Estadístico de contraste: Chi-cuadrado(2) = 0.337144
con valor p = 0.84487

Contraste de heterocedasticidad de White -
Hipótesis nula: No hay heterocedasticidad
Estadístico de contraste: LM = 1.09845
con valor p = P(Chi-cuadrado(2) > 1.09845) = 0.577397
```

# Medidas de corrección de la heterocedasticidad - Logaritmos



- Como vemos en el modelo anterior *se ha resuelto el problema de heterocedasticidad*, ya que el p-valor obtenido para el contraste de White es de 0,5773, por lo que a nivel de significación 0,05 no tenemos evidencias suficientes para rechazar la hipótesis nula de homocedasticidad en el modelo.
- Interpretaríamos el parámetro beta estimado como la variación porcentual que se produce en la variable dependiente ante cambios porcentuales en la variable independiente, es decir: Cuando los ingresos aumentan en un 1% el ocio aumenta en un 0,83%.

# Medidas de corrección de la heterocedasticidad - Logaritmos



- Veamos como se interpretaría el modelo si hubiésemos decidido tomar logaritmo únicamente en una variable:

1. Tomando logaritmos en la variable explicativa también se resolvería el problema de heterocedasticidad a nivel de significación 0,05.

gretl: modelo 6

Archivo Editar Contrastes Guardar Gráficos Análisis LaTeX

Modelo 6: MCO, usando las observaciones 1-50  
Variable dependiente: GASTOOCIO

	Coefficiente	Desv. Típica	Estadístico t	valor p
const	-2203.52	162.504	-13.56	4.91e-018 ***
LOGINGRESOS	319.983	19.2379	16.63	1.47e-021 ***

Media de la vble. dep. 491.4786 D.T. de la vble. dep. 226.2305  
Suma de cuad. residuos 370781.9 D.T. de la regresión 87.88983  
R-cuadrado 0.852150 R-cuadrado corregido 0.849070  
F(1, 48) 276.6544 Valor p (de F) 1.47e-21  
Log-verosimilitud -293.7306 Criterio de Akaike 591.4612  
Criterio de Schwarz 595.2852 Crit. de Hannan-Quinn 592.9174

Contraste de heterocedasticidad de White -  
Hipótesis nula: No hay heterocedasticidad  
Estadístico de contraste: LM = 5.91124  
con valor p = P(Chi-cuadrado(2) > 5.91124) = 0.0520464

Diríamos ahora que, por cada 1% que aumentan los ingresos, el gasto en ocio aumentará en 3,19 €.

# Medidas de corrección de la heterocedasticidad - Logaritmos



2. Y si el logaritmo lo tomamos en la variable dependiente el problema de heterocedasticidad no se resuelve.

```
gret: modelo 7
Archivo  Editar  Contrastes  Guardar  Gráficos  Análisis  LaTeX
Modelo 7: MCO, usando las observaciones 1-50
Variable dependiente: LOGGASTOOCIO

-----
                Coeficiente   Desv. Típica   Estadístico t   valor p
-----
const          5.01160        0.0572758      87.50           1.35e-054 ***
INGRESOS       0.000195936        9.53699e-06    20.54           1.99e-025 ***

Media de la vble. dep.  6.067135   D.T. de la vble. dep.  0.554473
Suma de cuad. residuos  1.538221   D.T. de la regresión  0.179015
R-cuadrado       0.897892   R-cuadrado corregido  0.895764
F(1, 48)         422.0888   Valor p (de F)        1.99e-25
Log-verosimilitud  16.08799   Criterio de Akaike     -28.17598
Criterio de Schwarz -24.35194   Crit. de Hannan-Quinn -26.71976

Log-verosimilitud para GASTOOCIO = -287.269

Contraste de heterocedasticidad de White -
Hipótesis nula: No hay heterocedasticidad
Estadístico de contraste: LM = 9.39363
con valor p = P(Chi-cuadrado(2) > 9.39363) = 0.00912429
```

No obstante,  
diremos que  
cuando los  
ingresos  
aumentan en 1€,  
el gasto en ocio  
aumentará un  
0,019%.

## Medidas de corrección de la heterocedasticidad – Mínimos cuadrados ponderados

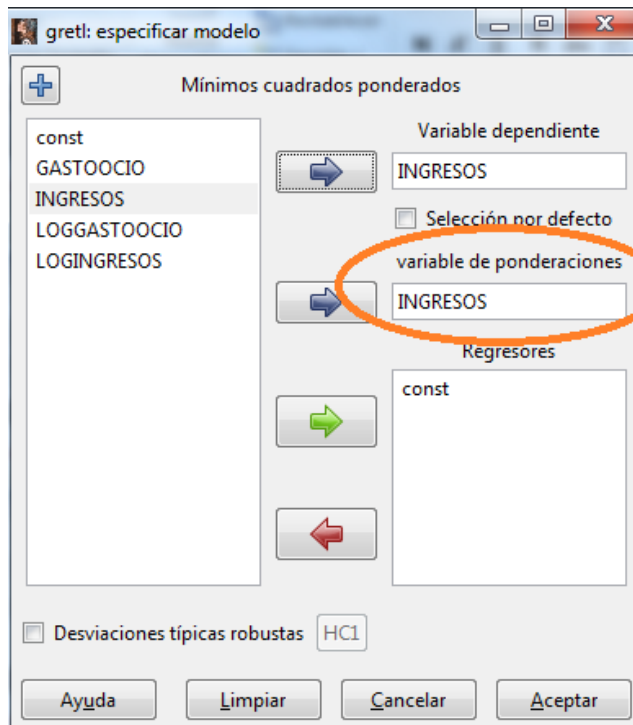
- Otra solución a la heterocedasticidad es estimar el modelo mediante el método de mínimos cuadrados ponderados.
- El método de mínimos cuadrados ponderados nos dará una mejor estimación de los parámetros en caso de presencia de heterocedasticidad en el modelo, ya que los estimará teniendo en cuenta que la varianza del error no es constante.
- Para obtenerlos es necesario tener una idea de la forma que presenta la heterocedasticidad y la variable que la está causando.
- En este caso no eliminamos el problema de nuestro modelo, sino que realizamos la estimación asumiéndolo.

# Medidas de corrección de la heterocedasticidad – Mínimos cuadrados ponderados



Para realizar una estimación por mínimos cuadrados ponderados en Gretl, utilizaremos la secuencia del menú principal:

Modelo -> Otros modelos lineales -> Mínimos cuadrados ponderados

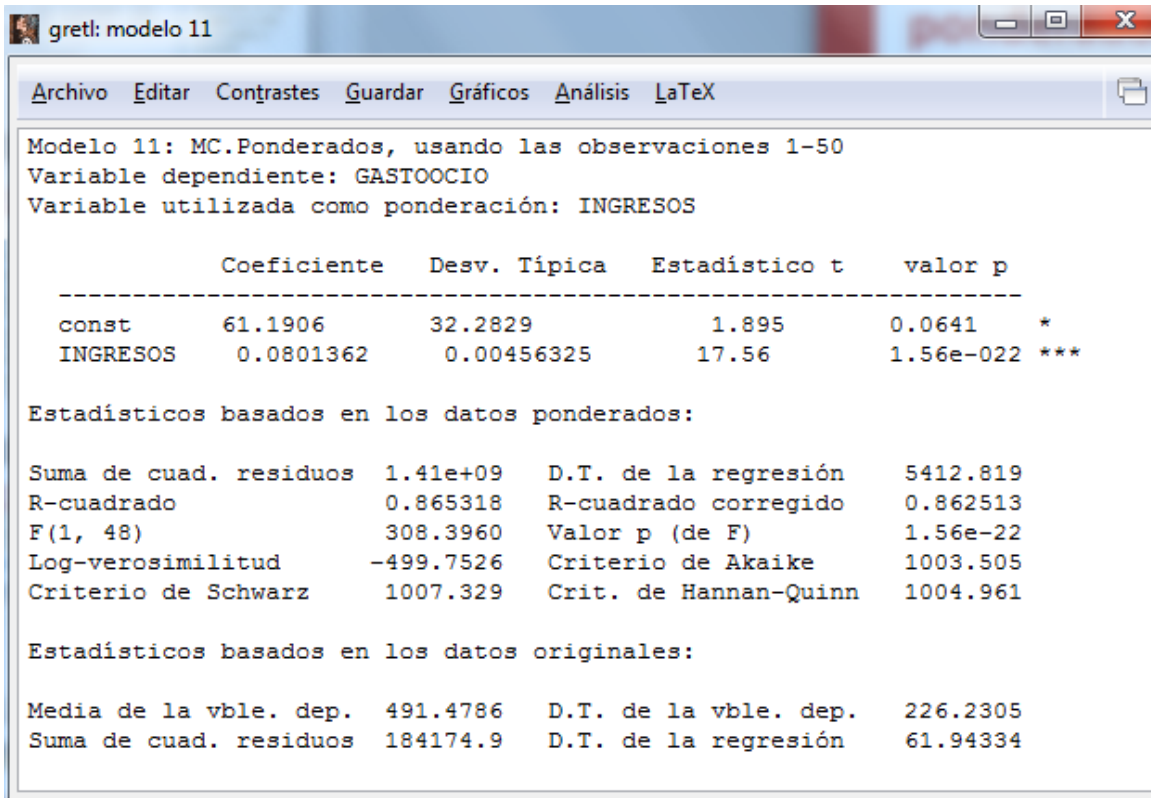


Es importante indicar la variable que utilizaremos para la ponderación, que será la causante del problema.



# Medidas de corrección de la heterocedasticidad – Mínimos cuadrados ponderados

Para nuestro modelo obtendremos la siguiente estimación:



gretl: modelo 11

Archivo Editar Contrastes Guardar Gráficos Análisis LaTeX

Modelo 11: MC.Ponderados, usando las observaciones 1-50  
Variable dependiente: GASTOOICIO  
Variable utilizada como ponderación: INGRESOS

	Coefficiente	Desv. Típica	Estadístico t	valor p	
const	61.1906	32.2829	1.895	0.0641	*
INGRESOS	0.0801362	0.00456325	17.56	1.56e-022	***

Estadísticos basados en los datos ponderados:

Suma de cuad. residuos	1.41e+09	D.T. de la regresión	5412.819
R-cuadrado	0.865318	R-cuadrado corregido	0.862513
F(1, 48)	308.3960	Valor p (de F)	1.56e-22
Log-verosimilitud	-499.7526	Criterio de Akaike	1003.505
Criterio de Schwarz	1007.329	Crit. de Hannan-Quinn	1004.961

Estadísticos basados en los datos originales:

Media de la vble. dep.	491.4786	D.T. de la vble. dep.	226.2305
Suma de cuad. residuos	184174.9	D.T. de la regresión	61.94334

Esta nueva estimación del modelo presenta mejores propiedades que la estimación por Mínimos cuadrados ordinarios.