Tema 3

Heterocedasticidad

3.1. Concepto de heterocedasticidad. Naturaleza y consecuencias. Ejemplos

Se dice que la varianza del término de perturbación del modelo de regresión lineal es heterocedástica cuando no es constante para todas las observaciones. En este tema nuestro marco de trabajo queda definido como sigue:

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \ldots + \beta_K X_{Ki} + u_i \qquad i = 1, 2, \ldots, N \Longleftrightarrow Y_i = \mathbf{X}_i' \beta + u_i$$

donde los regresores contenidos en la matriz $\mathbf{X}_i' = [1 \quad \mathbf{X}_{2i} \ \ldots \ \mathbf{X}_{Ki}]$ son no estocásticos y las perturbaciones tienen media cero, varianza no constante y covarianzas cero:

$$\begin{array}{l} E(u_i) = 0 \quad \forall i \\ E(u_i^2) = \sigma_i^2 \quad i = 1, 2, \ldots, N \\ E(u_i u_j) \neq 0 \ \forall i, j \quad i \neq j \end{array}$$

En términos matriciales el modelo se escribe:

$$\underset{(N\times 1)}{Y} = \underset{(N\times K)}{X} \ \underset{(K\times 1)}{\beta} + \underset{(N\times 1)}{u}$$

y la estructura de **la matriz de varianzas y covarianzas del vector de perturbaciones** sería la siguiente:

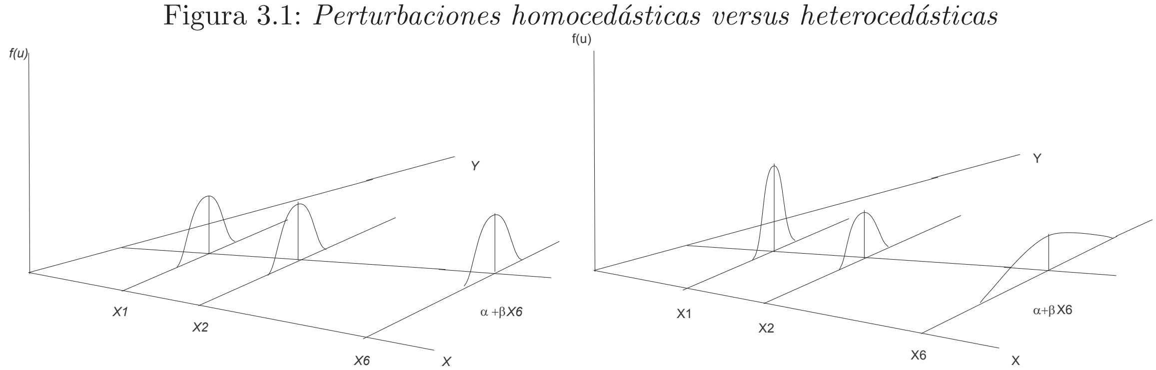
$$\underset{(N\times N)}{E(uu')} = \begin{bmatrix} \sigma_1^2 & 0 & 0 & \ldots & 0 \\ 0 & \sigma_2^2 & 0 & \ldots & 0 \\ 0 & 0 & \sigma_3^2 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & \sigma_N^2 \end{bmatrix} = \sigma^2 \begin{bmatrix} w_{11} & 0 & 0 & \ldots & 0 \\ 0 & w_{22} & 0 & \ldots & 0 \\ 0 & 0 & w_{33} & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & w_{NN} \end{bmatrix}$$

donde $Var(u_i) = \sigma_i^2$ puede cambiar para cada individuo y suponemos que no existe correlación entre perturbaciones de distintos individuos, $E(u_i u_j) = 0 \ \forall i, j \quad i \neq j$. Es decir, sólo consideramos la existencia de heterocedasticidad. Esta matriz es una matriz diagonal de dimensión $(N \times N)$ donde los elementos de la diagonal no son todos iguales. El parámetro σ^2 es un factor de escala común a todos los elementos de la matriz, que perfectamente puede tomar el valor de la unidad.

Para diferenciar entre el concepto de homocedasticidad y el concepto de heterocedasticidad podemos considerar la Figura 3.1. En la izquierda se puede observar que la varianza condicional de Y_i a las X_i permanece igual sin importar los valores que tome la variable X. Hay que recordar que la varianza condicional de Y_i es la misma que la de u_i, por tanto, en el gráfico estamos observando como la varianza de la perturbación permanece constante independientemente del valor que tome el regresor. En la derecha se muestra que la varianza de Y_i aumenta a medida que X aumenta[1]. Hay heterocedasticidad, y lo denotamos:

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

Figura 3.1: *Perturbaciones homocedásticas versus heterocedásticas*



Homocedásticas *Heterocedásticas*

La existencia de heterocedasticidad puede aparecer en numerosas aplicaciones económicas sin embargo, es más habitual en datos de sección cruzada. A continuación veremos algunas situaciones en las cuales las varianzas de u_i pueden no ser constantes.

- **En datos de sección cruzada.**

Ejemplo 3.1 Supongamos que tenemos datos para diferentes comunidades autónomas españolas en el año 2005 sobre gasto sanitario agregado, GS, renta personal disponible, R, el porcentaje de población que supera los 65 años, SEN y población, POP, con los que estimar el siguiente modelo:

$$GS_i = \beta_1 + \beta_2 R_i + \beta_3 SEN_i + \beta_4 POP_i + u_i \qquad i = 1, \ldots, N \tag{3.1}$$

[1]En este caso la varianza de u_i se relaciona con X_i y lo hace de forma creciente, pero podría ser función de otra u otras variables que no sean regresores del modelo. La forma funcional también puede ser diferente.

Las comunidades con más población y/o mayor porcentaje de población con edad superior a 65 años tendrán mayor gasto sanitario que aquellas con menor población o más joven. En esta situación suponer que la dispersión de los gastos sanitarios es la misma para todas las comunidades con distinto nivel de población y composición de la misma no es realista, y se debería proponer que la varianza de la perturbación sea heterocedástica $Var(u_i) = \sigma_i^2$, permitiendo por ejemplo que varíe en función creciente con la población, es decir, $\sigma_i^2 = \sigma^2 POP_i$. Incluso podemos pensar que varíe en función creciente con el porcentaje de población mayor de 65 años, en cuyo caso propondríamos $Var(u_i) = \sigma^2 SEN_i$ o con ambas variables, por lo que la forma funcional pudiera ser $Var(u_i) = \sigma^2(a\, POP_i + b\, SEN_i)$.

Ejemplo 3.2 Un ejemplo recurrente para mostrar la heterocedasticidad es el estudio de la relación entre consumo y renta. Este ejemplo se va a seguir detalladamente a lo largo del tema utilizándolo como ejemplo magistral. Será completamente descrito en la siguiente sección, pero reflexionemos un momento sobre él. Supongamos que tenemos datos sobre renta, R, y gasto en consumo, C, para N familias, con los que estimar el modelo:

$$C_i = \beta_1 + \beta_2 R_i + u_i \qquad i = 1, \dots, N \tag{3.2}$$

Las familias con mayor renta, una vez satisfechas sus necesidades primordiales tienen mayores posibilidades de decidir cuánto ahorrar y cuánto consumir, por lo que es habitual encontrar una mayor variabilidad en el gasto realizado por familias de renta alta que por familias de renta baja. En esta situación suponer que la dispersión de los gastos de consumo es la misma para todas las familias con distinto nivel de renta no es realista y se debería proponer que la varianza de la perturbación sea heterocedástica $Var(u_i) = \sigma_i^2$, permitiendo por ejemplo que varíe en función creciente con la renta de las familias, es decir, $\sigma_i^2 = \sigma^2 R_i$.

Ejemplo 3.3 Un fenómeno parecido ocurre con las empresas que deben decidir qué porcentaje de sus beneficios, B, deben repartir como dividendos, D. Las empresas con mayores beneficios tienen un margen de decisión muy superior al fijar su política de dividendos. Al estimar el modelo:

$$D_i = \beta_1 + \beta_2 B_i + u_i \qquad i = 1, \dots, N \tag{3.3}$$

cabría esperar que la varianza de u_i dependa del nivel de beneficios de la empresa i-ésima y podríamos proponer que por ejemplo, $E(u_i^2) = \sigma_i^2 = \sigma^2 B_i$.

- La heterocedasticidad también puede aparecer **como consecuencia de la agregación de datos**. En este caso la varianza puede depender del número de observaciones del grupo.

Ejemplo 3.4 Supongamos un investigador que desea estimar los coeficientes del siguiente modelo:

$$Y_j = \beta_1 + \beta_2 X_j + u_j \qquad j = 1, \dots, N \tag{3.4}$$

donde $u_j \sim N(0, \sigma^2)$, es decir, la varianza de la perturbación es homocedástica. Supongamos que el número de observaciones N es tal que aconseja agrupar las observaciones en m-grupos de n_i observaciones cada uno. Supongamos que como observación del grupo i-ésimo se toma la media aritmética dentro del grupo. El modelo a estimar sería:

$$\bar{Y}_i = \beta_1 + \beta_2 \bar{X}_i + \bar{u}_i \qquad i = 1, \ldots, m \tag{3.5}$$

y la nueva perturbación $\bar{u}_i$ seguirá teniendo media cero, pero su varianza no será constante ya que dependerá del número de observaciones dentro del grupo,

$$Var(\bar{u}_i) = \frac{\sigma^2}{n_i} \qquad i = 1, \ldots, m.$$

Si el número de observaciones dentro del grupo es el mismo en todos los grupos la varianza de la perturbación $\bar{u}_i$ es homocedástica.

- También encontraremos heterocedasticidad en un **modelo con coeficientes aleatorios** en el cual se permite cierta heterogeneidad entre individuos en los efectos de una variable explicativa.

Ejemplo 3.5 Sea el modelo:

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_{3i} X_{3i} + u_i \qquad i = 1, \ldots, N \tag{3.6}$$

donde:

- las variables exógenas X_{2i} y X_{3i} son no estocásticas.
- $u_i \sim iid(0, \sigma_u^2)$
- $\beta_{3i} = \beta + \epsilon_i$ siendo β una constante a estimar. La variable aleatoria ϵ_i recoge la heterogeneidad tal que, aunque en media el efecto es común a todos los individuos, recogido en β, hay cierta variabilidad recogida por σ_ϵ^2.
- $\epsilon_i \sim iid(0, \sigma_\epsilon^2)$ y $Cov(u_i, \epsilon_i) = 0$

En este caso la ecuación a estimar sería:

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta X_{3i} + v_i \qquad i = 1, \ldots, N \tag{3.7}$$

donde $v_i = u_i + \epsilon_i X_{3i}$ con distribución:

$$v_i \sim (0, \sigma_u^2 + X_{3i}^2 \sigma_\epsilon^2)$$

y vemos que la varianza del término de perturbación del modelo a estimar no es constante ya que depende de la variable exógena X_{3i} que varía con i.

- Otro caso sería la **existencia de un cambio estructural en varianza** recogido por una variable ficticia en la varianza de la perturbación.

Ejemplo 3.6 Supongamos que se desea estudiar la relación entre producción, Y, y mano de obra, X, para un conjunto de 20 trabajadores de los cuales 10 son mujeres y el resto hombres. Si suponemos que la variabilidad de la producción es distinta para los hombres que para las mujeres nuestro modelo a estimar sería:

$$Y_i = \beta_1 + \beta_2 X_i + u_i \qquad i = 1, \ldots, 20 \tag{3.8}$$

donde $u_i \sim (0, \alpha_1 + \alpha_2 D_i)$ siendo D_i una variable ficticia que toma valor la unidad si la observación corresponde a una mujer y cero en el caso contrario. En este caso:

$$\begin{array}{ll} Var(u_i) = \alpha_1 + \alpha_2 & \text{para las observaciones correspondientes a las mujeres} \\ Var(u_i) = \alpha_1 & \text{para las observaciones correspondientes a los hombres} \end{array}$$

Suponiendo que las primeras diez observaciones corresponden a mujeres, la matriz de varianzas y covarianzas del vector de perturbaciones sería la siguiente:

$$E(uu') = \begin{bmatrix} (\alpha_1 + \alpha_2) I_{10} & 0 \\ 0 & \alpha_1 I_{10} \end{bmatrix}$$

Antes de pasar a abordar el problema de hereterocedasticidad más en detalle vamos a recordar cuáles son las consecuencias de la heterocedasticidad sobre el estimador de MCO y su matriz de varianzas y covarianzas:

- Si $u \sim N(0, \sigma^2 \Omega)$ entonces:

  $$\hat{\beta}_{MCO} \sim N(\beta, \sigma^2 (X'X)^{-1} X' \Omega X (X'X)^{-1})$$

  el estimador MCO será lineal e insesgado, pero no será de varianza mínima si $\Omega \neq I$.

- En el caso de un modelo con perturbaciones heterocedásticas la variabilidad de las observaciones de la variable endógena dada la exógena no es la misma para todas las observaciones. El criterio MCO prescinde de esta información ya que concede a todas las observaciones el mismo peso.

- En estas circunstancias parecería más adecuado ponderar más las realizaciones muestrales que sistemáticamente se desvían menos del valor promedio de la variable endógena, que es sobre lo que queremos inferir. Es decir, desearíamos dar mayor peso a aquellas observaciones que surgen de poblaciones con menor variabilidad que a aquellas que provienen de poblaciones con mayor variabilidad.

- Debemos estimar por MCG (cuando Ω es conocida). El criterio de estimación Mínimo Cuadrático Generalizado que estudiábamos en el tema anterior es:

  $$\underset{\hat{\beta}}{Min} \left[(Y - X\hat{\beta})' \, \Omega^{-1} \, (Y - X\hat{\beta}) \right]$$

- El estimador de MCG que se obtiene de la función objetivo anterior se define como:

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

El estimador MCG es lineal, insesgado y de varianza mínima. Su matriz de varianzas y covarianzas es:

$$V(\hat{\beta}_{MCG}) = \sigma^2(X'\Omega^{-1}X)^{-1}$$

Si σ^2 es desconocida, que es lo habitual, la estimamos con la expresión $\hat{\sigma}^2 = \frac{\hat{u}'\Omega^{-1}\hat{u}}{N-K}$ siendo $\hat{u} = Y - X\hat{\beta}_{MCG}$.

Ejemplo 3.7 Para el modelo $Y_i = \beta X_i + u_i$ siendo X una variable no estocástica y $u_i \sim NID(0, \sigma_i^2)$. El criterio MCG (o función objetivo) se escribiría:

$$\underset{\hat{\beta}}{Min} \sum_{i=1}^{N} \frac{1}{\sigma_i^2} (Y_i - X_i\hat{\beta})^2$$

de donde

$$\hat{\beta}_{MCG} = \left[\sum_{i=1}^{N} \frac{1}{\sigma_i^2} X_i X_i\right]^{-1} \left[\sum_{i=1}^{N} \frac{1}{\sigma_i^2} X_i Y_i\right] = \frac{\sum_{i=1}^{N} \frac{1}{\sigma_i^2} X_i Y_i}{\sum_{i=1}^{N} \frac{1}{\sigma_i^2} X_i X_i}$$

En la práctica, generalmente no sabemos de antemano si hay o no problemas de heterocedasticidad en las perturbaciones, por lo que se ha desarrollado un gran número de procedimientos para contrastar la hipótesis nula de igualdad de varianzas u homocedasticidad. Esta gran variedad se debe a que la especificación de la hipótesis alternativa de heterocedasticidad no suele ser conocida y puede ser más o menos general. En la siguiente sección vamos a abordar algunos de estos contrastes.

3.2. Contrastes de heterocedasticidad

Nuestro objetivo es claro: **Detectar la existencia de heterocedasticidad en las perturbaciones de un modelo**. La primera aproximación al objetivo es el estudio de los gráficos de residuos y de las variables del modelo.

3.2.1. Detección gráfica.

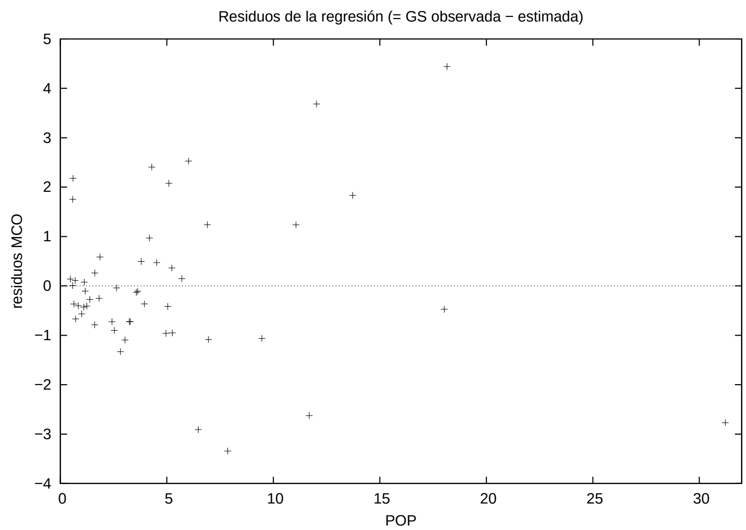
La aplicación del estimador de MCG y algunos contrastes de heterocedasticidad requieren conocer la forma funcional de la varianza de la perturbación. Si suponemos que la varianza de la perturbación depende de uno o más regresores, u otras variables conocidas, un instrumento adecuado para aproximarnos a la misma sería llevar a cabo un análisis de los residuos MCO donde no hemos tenido en cuenta la existencia de heterocedasticidad. Aunque $\hat{u}_{MCO,i}$ no es lo mismo que u_i la detección de patrones sistemáticos en la variabilidad de los residuos MCO nos indicará la posible existencia de heterocedasticidad en las perturbaciones. Además, puede indicarnos una posible forma funcional de la misma.

Consideramos el modelo (3.1) recogido en el Ejemplo 3.1:

$$GS_i = \beta_1 + \beta_2 R_i + \beta_3 SEN_i + \beta_4 POP_i + u_i \qquad i = 1, \ldots, N$$

donde suponemos $E(u_i) = 0 \quad \forall i$ y $E(u_i u_j) = 0 \quad \forall i, j \quad i \neq j$. Si sospechamos que u_i es heterocedástica debido a la variable POP, podemos intentar detectar la existencia de heterocedasticidad en las perturbaciones del modelo ayudándonos del gráfico de los residuos MCO, $(\hat{u}_{MCO,i})$, frente a la variable POP_i.

Figura 3.2: *Residuos MCO versus POP*



Si el gráfico es como el recogido en la Figura 3.2 pensaremos que la variabilidad de los residuos $\hat{u}_{MCO,i}$ se incrementan con POP_i y que el incremento es directamente proporcional. Así, podríamos proponer, por ejemplo:

$$E(u_i^2) = \sigma^2 POP_i \qquad i = 1, 2, \ldots, N$$

Si el gráfico de los residuos MCO frente a POP *hubiera* sido como el recogido en la Figura 3.3 supondríamos que el aumento en la varianza de u_i es inversamente proporcional a POP_i y propondríamos:

$$E(u_i^2) = \sigma^2 POP_i^{-1} \qquad i = 1, 2, \ldots, N$$

También podemos optar por dibujar la serie de los residuos al cuadrados MCO frente a la variable que creemos causa la heterocedasticidad como se muestra en la Figura 3.4. En el gráfico de la izquierda se muestran los pares $(SEN_i, \hat{u}_{MCO,i})$, en el gráfico de la derecha se muestran los pares $(SEN_i, \hat{u}^2_{MCO,i})$. Ambos gráficos muestran la misma información, muestran que la variabilidad de los residuos se incrementa con SEN y podríamos proponer, por ejemplo $Var(u_i) = E(u_i^2) = \sigma^2 SEN_i$.

En general a priori no se conocerá cuál de las variables exógenas genera la heterocedasticidad por lo que resulta aconsejable estudiar los gráficos de los residuos de MCO, contraponiéndolos a cada

Figura 3.3: *Residuos MCO versus POP*

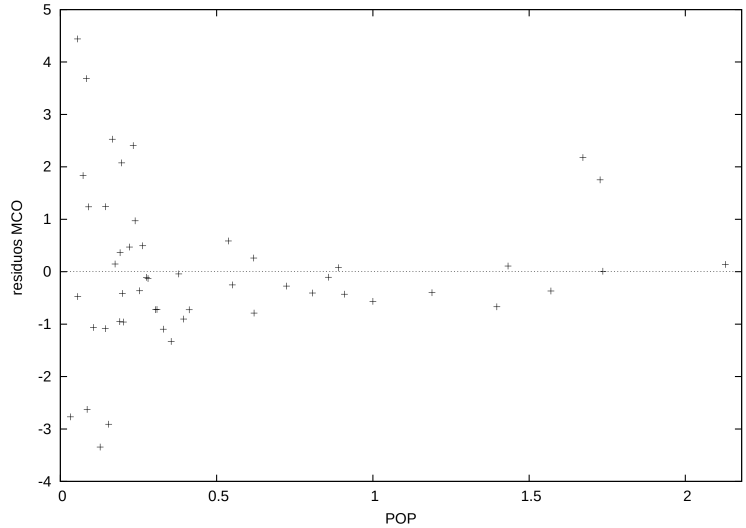
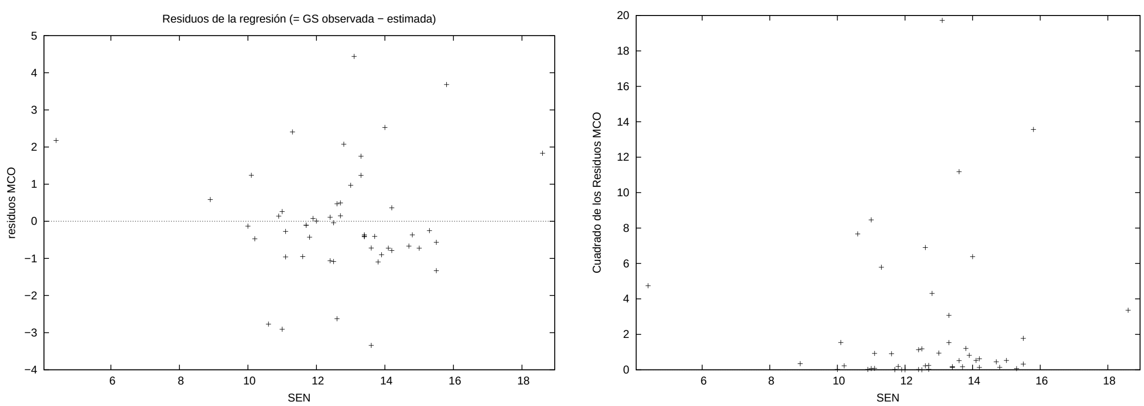


Figura 3.4: *Residuos MCO y sus cuadrados versus SEN*



una de las variables exógenas del modelo, como estamos haciendo al estudiar los residuos frente a POP_i y frente a SEN_i. Notar que ambas variables parecen afectar a la varianza de la perturbación, por ello estaría justificado proponer $Var(u_i) = (a\, POP_i + b\, SEN_i)$, donde a y b son desconocidos y el factor de escala es la unidad, $\sigma^2 = 1$.

Si la gráfica entre $\hat{u}_{MCO,i}$ y POP_i *hubiera* resultado como la de la Figura 3.5, concluiríamos que la varianza de la perturbación no depende de POP_i ya que no se aprecia ningún patrón de comportamiento y parece que hay una distribución aleatoria de los pares $(POP_i, \hat{u}_i)$. En esta situación procede analizar los residuos frente al resto de regresores del modelo.

Las formas anteriores no son las únicas. Si recordamos, en el Ejemplo 3.6 se suponía una situación donde hombres y mujeres en una empresa tenían diferente productividad y se suponía que $Var(u_i) = \alpha_1 + \alpha_2 D_i$ siendo D_i una variable ficticia que toma valor uno si la observación corresponde a una mujer y cero en caso contrario. En esta situación esperaríamos un gráfico como el recogido en la

Figura 3.5: *Perturbaciones homocedásticas*

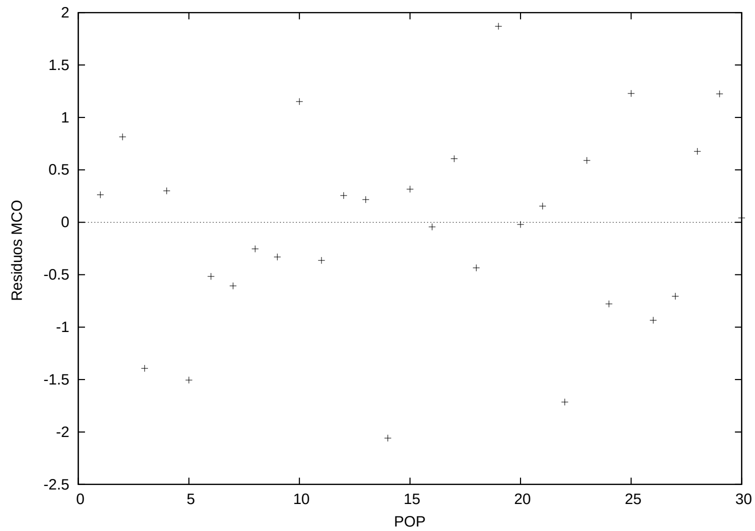


Figura 3.6: *Residuos MCO frente a una variable ficticia*

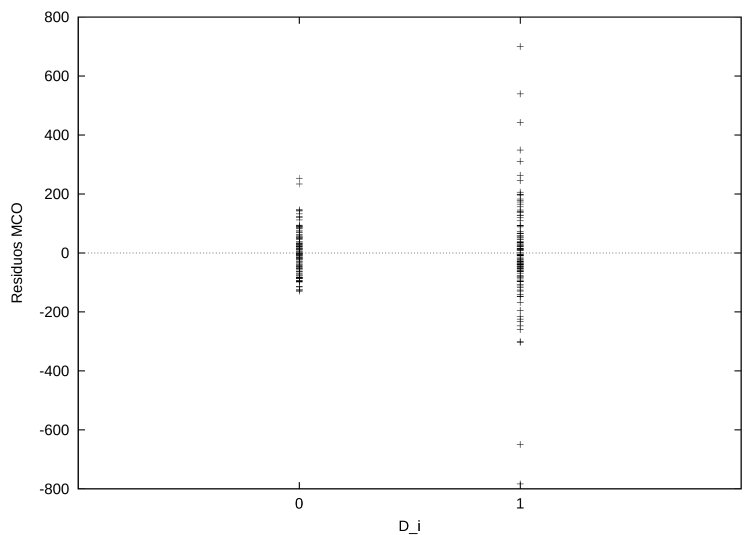


Figura 3.6 donde claramente la dispersión de los residuos para las mujeres es mucho mayor que para los hombres.

Como conclusión diremos que al analizar los gráficos de la relación residuos MCO, o sus cuadrados, con cada uno de los regresores lo que intentaremos detectar visualmente es un crecimiento o decrecimiento en la variabilidad de los residuos con respecto a la variable en cuestión.

• **Ejemplo a seguir en las clases magistrales del resto del tema:**

En lo que sigue del tema resulta muy interesante disponer de un ejercicio sobre el que ir trabajando y viendo resultados, y sobre todo para que podáis replicar los mismos. Por ello os proponemos realizar un estudio sobre la relación entre Consumo, C, y Renta, R, para 40 familias. En la plataforma de apoyo a la docencia, eKASI, tenéis disponible un archivo llamado ejemHETERO.gdt que contiene las observaciones de consumo y renta del ejercicio para que podáis replicarlo por vuestra cuenta. En el Anexo 3.1 que se encuentra al final de este tema podéis encontrar los resultados de gretl que se van mostrando.
Objetivo: Analizar el comportamiento de la varianza de la perturbación en un análisis de la relación entre consumo y renta.
Datos: Observaciones sobre gasto semanal en alimentos y renta semanal para 40 familias. Ambos medidos en dólares. Se muestran en la Tabla 3.1 recogida en el Anexo 3.3[2]. Modelo a estimar:

$$C_i = \beta_1 + \beta_2 R_i + u_i \qquad i = 1, 2, \ldots, 40$$

donde:

$$E(u_i) = 0 \quad \forall i, \qquad E(u_i u_j) = 0 \qquad \forall i, j \quad i \neq j$$

y se cumplen el resto de hipótesis básicas. Nos preguntamos: **¿Es constante la varianza de la perturbación?**

Lo primero que vamos a hacer es estimar el modelo propuesto por MCO y analizar el gráfico de la relación entre consumo y renta, Figura 3.7. Los resultados de la regresión son los siguientes[3]:

Estimaciones MCO utilizando las 40 observaciones 1–40
Variable dependiente: consumo

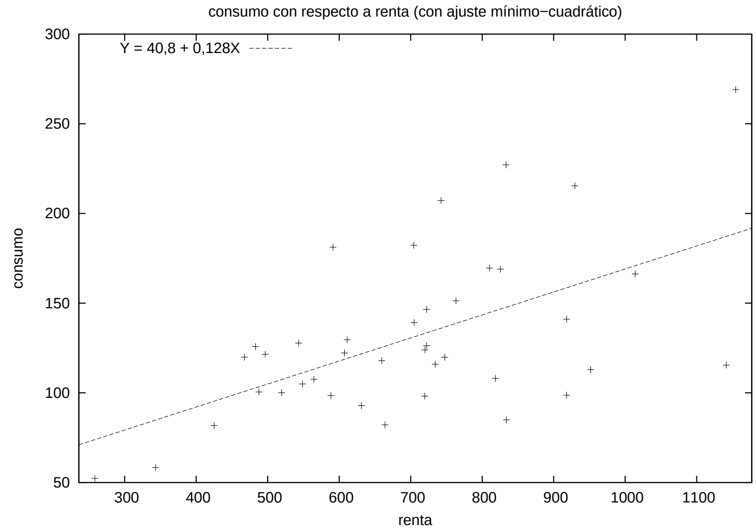
	Coeficiente	Desv. típica	estadístico t	valor p
const	40,7676	22,1387	1,8415	0,0734
renta	0,1282	0,0305	4,2008	0,0002

Suma de cuadrados de los residuos	54311,3
Desviación típica de la regresión ($\hat{\sigma}$)	37,8054
R^2	0,3171
$\bar{R}^2$ corregido	0,2991

En la Figura 3.7 se muestra el gráfico de las observaciones de la variable Consumo, C, sobre la variable explicativa Renta, R, junto con la recta de ajuste mínimo-cuadrático. El gráfico muestra que existe una relación lineal y positiva entre consumo y renta. A mayor renta mayor consumo. Además, se puede observar como no solo que las familias con mayor renta tienen un consumo en media mayor que las de menor renta, sino que además, las familias de rentas altas tienen una mayor variación en su consumo relativamente a las familias de rentas bajas. Lo mismo podemos concluir de los resultados de la estimación. Aunque el ajuste no es muy bueno, ya que el R^2 es bajo, si la

[2] Con la tabla también tenéis suficiente información para introducir vosotros mismos los datos en gretl. El procedimiento aparece descrito en el Anexo 3.2 en el que se muestran instrucciones básicas de gretl.
[3] Ver Anexo 3.1, Resultado 1.

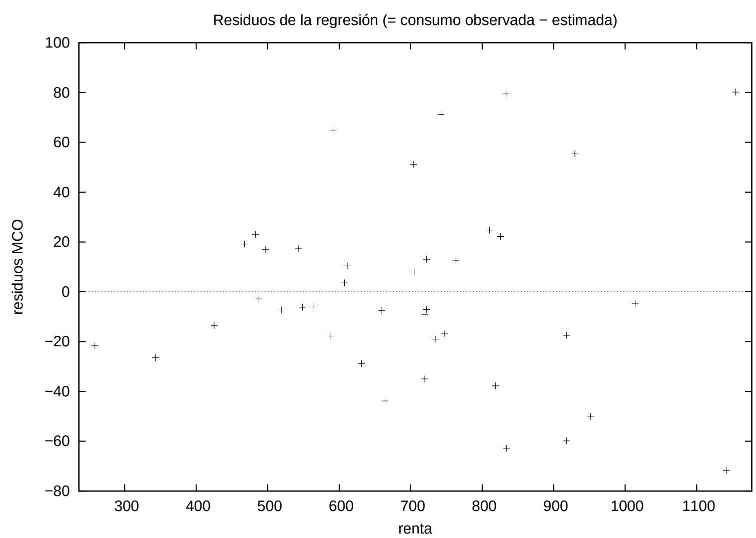
Figura 3.7: *Consumo versus Renta*



perturbación fuese esférica y normal, la renta sería una variable significativa al 5 % para explicar el consumo.

La Figura 3.8 muestra los pares $(R_i, \hat{u}_{i,MCO})$. El gráfico permite dudar sobre el comportamiento homocedástico de la varianza de la perturbación ya que se puede ver claramente como a medida que aumenta la variable renta aumenta la dispersión en los residuos.

Figura 3.8: *Residuos MCO versus Renta*



La apreciación gráfica de una posible existencia de heterocedasticidad debe ser refrendada mediante un contraste.

3.2.2. Test de contraste para heterocedasticidad

A continuación veremos algunos de los test de contraste para heterocedasticidad más importantes. Todos ellos contrastan la existencia de heterocedasticidad suponiendo:

H_0**: ausencia de heterocedasticidad.**
H_a**: existencia de heterocedasticidad.**

Existe una gran variedad de test de contraste de heterocedasticidad que se diferencian entre sí en su generalidad a la hora de modelizar la heterocedasticidad en la hipótesis alternativa y su potencia para detectarla. En este curso aprenderemos dos test de contraste, el test de Goldfeld y Quandt y el test de Breusch y Pagan. El primero de ellos parte del supuesto de que la varianza de la perturbación depende monótonamente de los valores de una única variable y es razonablemente potente si somos capaces de identificar bien esa variable. Por contra, el segundo supone que varianza de la perturbación varía en función de un conjunto de variables independientes. Cuanto mejor seamos capaces de identificar la naturaleza de la heterocedasticidad mejor podremos contrastarla. Pero incluso en el caso en que no seamos capaces de identificar cual es su naturaleza podemos implementar un contraste de tipo general como es el test de White que no requiere hacer supuestos sobre la naturaleza de la misma. El test de White es un test de carácter muy general, y por lo mismo de baja potencia, que en gretl puede ser obtenido directamente como se muestra en el Anexo 3.2.

- **Test de Goldfeld y Quandt (1965):**

Parte del supuesto de que la varianza de la perturbación, σ_i^2 depende monótonamente de los valores de una variable Z_i, que puede ser o no uno de los regresores del modelo. En cualquier caso debe ser una variable observable. Para contrastar la hipótesis nula de ausencia de heterocedasticidad:

$$H_0 : \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_N^2$$

contra la alternativa de existencia de heterocedasticidad:

$$H_a : \sigma_i^2 = \sigma^2 g(Z_i)$$

donde $g(\cdot)$ es una función monótona, supongamos que creciente con Z_i, podemos proceder de la manera siguiente:

1. Ordenar las observaciones de todas las variables del modelo en la muestra según un ordenamiento de los valores de Z_i de menor a mayor.

2. Dividir la muestra en dos bloques de tamaño muestral N_1 y N_2 respectivamente, pudiendo dejar fuera p observaciones centrales para hacer más independientes los dos grupos. El número de observaciones de cada grupo ha de ser similar y mayor que el número de parámetros a estimar[4].

[4] En el caso de tener pocas observaciones, p puede ser cero.

3. Estimar por MCO el modelo de regresión separadamente para cada grupo de observaciones. Guardar la Suma de Cuadrados Residual (SCR) de cada regresión.

4. Construir el siguiente estadístico de contraste, que bajo la hipótesis nula de ausencia de heterocedasticidad y suponiendo que la perturbación sigue una distribución normal, de media cero y no está serialmente correlacionada, sigue una distribución F-Snedecor.

$$GQ = \frac{\hat{\sigma}_2^2}{\hat{\sigma}_1^2} = \frac{\hat{u}_2'\hat{u}_2}{\hat{u}_1'\hat{u}_1}\frac{N_1 - K}{N_2 - K} \overset{H_0}{\sim} F_{(N_2-K,N_1-K)}$$

donde:
$\hat{u}_2'\hat{u}_2$ es la SCR de la regresión de Y sobre X en el segundo grupo de observaciones.
$\hat{u}_1'\hat{u}_1$ es la SCR de la regresión de Y sobre X en el primer grupo de observaciones.

• Interpretación del contraste y regla de decisión:
Si existe homocedasticidad las varianzas han de ser iguales, pero si existe heterocedasticidad del tipo propuesto, con la ordenación de la muestra de menor a mayor, la varianza del término de error será mayor al final de la muestra. Entonces $\hat{u}_2'\hat{u}_2$ debería ser sensiblemente mayor que $\hat{u}_1'\hat{u}_1$.

Cuanto más diverjan las sumas de cuadrados, mayor será el valor del estadístico y mayor será la evidencia contra la H_0, rechazaremos H_0, a un nivel de significación α si:

$$GQ > F_{(N_1-K,N_2-K)|\alpha}$$

Observaciones:

- Si se sospecha que la varianza del término de error depende inversamente de los valores que toma una variable Z_i, entonces se debería ordenar la muestra de acuerdo a un ordenamiento de mayor a menor de los valores decrecientes de dicha variable y proceder del modo descrito anteriormente.

- ¿Cómo elegir p?
  Anteriormente se ha propuesto dividir la muestra en dos partes. Elegir el valor de p es relevante ya que cuanto mayor sea p más grados de libertad se pierden y por tanto, perdemos potencia del contraste. Si p es demasiado pequeño no habrá independencia entre grupos y se prima la homocedasticidad frente a la posibilidad de heterocedasticidad. Harvey y Phillips (1974) sugieren fijar p a $\frac{1}{3}$ de la muestra.

- El contraste se puede utilizar para detectar, en principio, heterocedasticidad de forma general, aunque está pensado para alternativas específicas donde se supone un crecimiento de las varianzas en función de una determinada variable. Si en realidad el problema no es ese, sino que existe otra forma de heterocedasticidad, el estadístico puede no capturarla y no ser significativo.

- Por otro lado si no se rechaza la H_0 también puede deberse a una mala especificación de σ_i^2, que quizá depende de una variable diferente a la supuesta. Por ello puede ser necesario repetir el contraste para otras variables de las que podamos sospechar a priori.

Aplicación del test de Goldfeld y Quandt al ejercicio propuesto. Contrastamos:

$$H_0 : \sigma_1^2 = \sigma_2^2$$
$$H_a : \sigma_1^2 < \sigma_2^2$$

Estadístico de contraste:

$$GQ = \frac{\hat{\sigma}_2^2}{\hat{\sigma}_1^2} = \frac{\hat{u}_2'\hat{u}_2}{\hat{u}_1'\hat{u}_1}\frac{N_1 - K}{N_2 - K} \overset{H_0}{\sim} F_{(N_2-K,N_1-K)}$$

donde:
$\hat{u}_2'\hat{u}_2$ es la SCR de la regresión de C sobre R, incluyendo un término constante, en el segundo grupo de observaciones, $i = 21, \ldots, 40$.
$\hat{u}_1'\hat{u}_1$ es la SCR de la regresión de C sobre R, incluyendo un término constante, en el primer grupo de observaciones, $i = 1, \ldots, 20$.
En el Anexo 3.1, Resultado 2 se recogen los resultados obtenidos de gretl para resolver el contraste y que se muestran a continuación. Los resultados de la estimación en el primer grupo de observaciones son:

$$\begin{array}{llll} \widehat{C}_i & = 12,9388 + 0,1823\, R_i & SCR = 12284,2 & R^2 = 0,4052 \\ (\widehat{des}(\hat{\beta}_{i,MCO})) & \quad (28{,}96) \qquad (0{,}52) & & \end{array}$$

Los resultados de la estimación en el segundo grupo de observaciones son:

$$\begin{array}{llll} \widehat{C}_i & = 48,1767 + 0,1176\, R_i & SCR = 41146,9 & R^2 = 0,1036 \\ (\widehat{des}(\hat{\beta}_{i,MCO})) & \quad (70{,}24) \qquad (0{,}08) & & \end{array}$$

Dada la muestra

$$GQ = \frac{41146,9}{12284,2} = 3{,}34 \qquad F_{(18,18)|0{,}05} = 2{,}19$$

como $3{,}34 > 2{,}19$ rechazo la H_0 para $\alpha = 5\,\%$ y podemos concluir que existe heterocedasticidad explicada por la variable renta. Una forma funcional razonable para la varianza de la perturbación sería:

$$Var(u_i) = \sigma^2 R_i \qquad i = 1, 2, \ldots, 40$$

pero también pudieran ser otras como por ejemplo $Var(u_i) = \sigma^2 R_i^2$ tal que se recoja una relación creciente de la varianza de u_i con la renta.

• Contraste de Breusch y Pagan (1979):

Breusch y Pagan en 1979 derivan un contraste de heterocedasticidad donde la hipótesis alternativa es bastante general:

$$H_0 : E(u_i^2) = \sigma^2 \quad \forall i \quad u_i \sim NID(0, \sigma^2)$$
$$H_a : \sigma_i^2 = \sigma^2 g(\alpha_0 + \alpha_1 Z_{1i} + \alpha_2 Z_{2i} + \ldots + \alpha_p Z_{pi})$$

las variables Z_{pi} pueden o no ser variables explicativas del modelo de interés, en cualquier caso deben ser observables y $g(\cdot)$ no se especifica. Si todos los coeficientes de la combinación lineal $Z_i'\alpha$

fuesen cero excepto α_0, la varianza sería homocedástica, $\sigma_i^2 = \sigma^2 g(\alpha_0)$. Por tanto un contraste de la hipótesis nula de homocedasticidad vendría dado por la siguiente hipótesis:

$$H_0 : \alpha_1 = \alpha_2 = \ldots = \alpha_p = 0$$

donde se contrastan p restricciones lineales. El proceso de contraste es el siguiente:

1. Estimar por MCO el modelo de interés obteniendo los residuos correspondientes, $\hat{u}_{MCO,i}$.

2. Se obtiene la siguiente serie de residuos normalizados:

$$\hat{e}_i^2 = \frac{\hat{u}_{MCO,i}^2}{\hat{u}'\hat{u}/N} \qquad i = 1, 2, \ldots, N$$

3. Se estima por MCO y se guarda la Suma de Cuadrados Explicada, SCE, de la siguiente regresión auxiliar:

$$\hat{e}_i^2 = \alpha_0 + \alpha_1 Z_{1i} + \alpha_2 Z_{2i} + \ldots + \alpha_p Z_{pi} + \eta_i \quad i = 1, 2, \ldots, N$$

4. Se utiliza como estadístico de contraste el siguiente:

$$BP = \frac{SCE}{2} \xrightarrow{d, H_0} \chi_p^2$$

siendo p los grados de libertad, igual al número de variables Z_{pi}. Rechazamos la H_0 de homocedasticidad a un nivel de significatividad α, si el valor muestral del estadístico excede del valor crítico esto es, $BP > \chi^2_{p|\alpha}$.

• Interpretación del contraste y regla de decisión: si las perturbaciones fuesen homocedásticas, las variables Z_{pi} no deberían tener poder explicativo acerca de los residuos transformados y por tanto la SCE debería ser pequeña. Si $SCE/2$ es grande, rechazaremos la H_0 y existiría heterocedasticidad.

Aplicación del test Breusch y Pagan al ejercicio propuesto. Contrastamos:

$$\begin{array}{l} H_0 : E(u_i^2) = \sigma^2 \quad \forall i \\ H_a : \sigma_i^2 = \sigma^2 g(\alpha_0 + \alpha_1 R_i) \end{array}$$

En este caso la regresión auxiliar es:

$$\hat{e}_i^2 = \alpha_0 + \alpha_1 R_i + \eta_i \qquad i = 1, \ldots, 40$$

En el Anexo 3.1, Resultado 3 se recogen los resultados obtenidos de gretl para resolver el contraste y que se muestran a continuación. Los resultados de la estimación por MCO de la regresión auxiliar son:

$$\begin{array}{rcllll} \widehat{\hat{e}_i^2} & = & -1,6788 \; + & 0,0038 \;\; R_i & \qquad SCR = 52,3940 & \quad R^2 = 0,3010 \\ (\widehat{des}(\hat{\beta}_{i,MCO})) & & (0{,}6876) & (0{,}0009) & & \end{array}$$

El valor muestral del estadístico de contraste es:

$$\frac{SCE}{2} = 11{,}28 \qquad \chi^2_{1|0{,}05} = 3{,}84$$

como 11,28 > 3,84, rechazo la H_0 para $\alpha = 5\,\%$ y podemos concluir que existe heterocedasticidad.

3.3. El estimador MCG bajo heterocedasticidad. Mínimos Cuadrados Ponderados

Una vez detectada estadísticamente la existencia de heterocedasticidad debemos abordar la estimación del modelo teniendo en cuenta esta información. Una posibilidad es obtener el estimador de Mínimos Cuadrados Generalizados, MCG, que tiene en cuenta la estructura existente en la matriz de varianzas y covarianzas. Sin embargo, implica conocer o proponer una forma funcional para la varianza de la perturbación en cada i. Si la estructura de la matriz de varianzas y covarianzas se conoce al menos excepto por un factor de escala, $E(uu') = \sigma^2\Omega$, y Ω es conocida, el vector de coeficientes β se puede estimar por MCG resolviendo el siguiente problema de minimización:

$$\underset{\hat{\beta}}{Min}(Y - X\hat{\beta})'\Omega^{-1}(Y - X\hat{\beta}) = \underset{\hat{\beta}}{Min}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}(Y_i - X_i'\hat{\beta})^2$$

a resultas del cual podemos definir el estimador de MCG como:

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

• **Veamos una aplicación del procedimiento:**
En el ejemplo que estamos desarrollando el modelo a estimar es el siguiente:

$$C_i = \beta_1 + \beta_2 R_i + u_i \qquad i = 1, 2, \ldots, 40 \tag{3.9}$$

donde $E(u_i) = 0 \;\forall\, i,\; E(u_i u_j) = 0 \quad \forall\, i, j \quad i \neq j$ y hemos concluido que un supuesto razonable para el comportamiento de la varianza de la perturbación es:

$$Var(u_i) = \sigma^2 R_i \qquad i = 1, 2, \ldots, 40$$

La estructura de la matriz de varianzas y covarianzas del vector de perturbaciones es la siguiente:

$$E(uu') = \sigma^2 \begin{bmatrix} R_1 & 0 & 0 & \ldots & 0 \\ 0 & R_2 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & 0 & \ldots & R_{40} \end{bmatrix} = \sigma^2\Omega$$

luego Ω es conocida. Si queremos obtener estimadores lineales, insesgados y de varianza mínima estimaremos por MCG. El vector de coeficientes β se puede estimar por MCG resolviendo el siguiente problema de minimización:

$$\underset{\hat{\beta}}{Min}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}(C_i - \hat{\beta}_1 - \hat{\beta}_2 R_i)^2 = \underset{\hat{\beta}}{Min}\sum_{i=1}^{N}\frac{1}{R_i}(C_i - \hat{\beta}_1 - \hat{\beta}_2 R_i)^2 \tag{3.10}$$

Minimizar esta función es equivalente a estimar por MCO el siguiente modelo transformado[5]:

$$\frac{C_i}{\sqrt{R_i}} = \beta_1\frac{1}{\sqrt{R_i}} + \beta_2\sqrt{R_i} + \frac{u_i}{\sqrt{R_i}} \qquad i = 1, 2, \ldots, 40 \tag{3.11}$$

[5] Para el modelo transformado podemos escribir la función a minimizar, equivalente a las anteriores, como: $\underset{\hat{\beta}}{Min}\sum_{i=1}^{N}\left(\frac{C_i}{\sqrt{R_i}} - \hat{\beta}_1\frac{1}{\sqrt{R_i}} - \hat{\beta}_2\sqrt{R_i}\right)^2$. Notar que el criterio y el estimador son invariantes al factor de escala σ^2.

donde la perturbación es homocedástica, no autocorrelada y de media cero:

$$E\left(\frac{u_i}{\sqrt{R_i}}\right) = \frac{E(u_i)}{\sqrt{R_i}} = 0 \quad \forall i$$

$$Var\left(\frac{u_i}{\sqrt{R_i}}\right) = E\left(\frac{u_i}{\sqrt{R_i}} - E\left(\frac{u_i}{\sqrt{R_i}}\right)\right)^2 = E\left(\frac{u_i}{\sqrt{R_i}}\right)^2 = \frac{E(u_i^2)}{R_i} = \frac{\sigma^2 R_i}{R_i} = \sigma^2 \quad \forall i$$

$$Cov\left(\frac{u_i}{\sqrt{R_i}}\frac{u_j}{\sqrt{R_j}}\right) = E\left(\frac{u_i}{\sqrt{R_i}}\frac{u_j}{\sqrt{R_j}}\right) = \frac{E(u_i u_j)}{\sqrt{R_i}\sqrt{R_j}} = 0 \quad \forall i,j \quad i \neq j$$

A la vista de las propiedades de la perturbación del modelo transformado el estimador de MCO de este modelo transformado esto es, el de MCG, es lineal, insesgado y de varianza mínima. Si conocemos la distribución de la perturbación podemos hacer inferencia en muestras finitas con este estimador. Lo veremos más adelante.

Vamos a formalizar matricialmente el modelo transformado (3.11):

$$P^{-1}Y = P^{-1}X\beta + P^{-1}u \Leftrightarrow Y^\star = X^\star\beta + u^\star$$

donde P es la matriz de transformación tal que $PP' = \Omega$ y $(P^{-1})'P^{-1} = \Omega^{-1}$:

$$P = \begin{bmatrix} \sqrt{R_1} & 0 & 0 & \dots & 0 \\ 0 & \sqrt{R_2} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & \sqrt{R_{40}} \end{bmatrix} \qquad P^{-1} = \begin{bmatrix} \frac{1}{\sqrt{R_1}} & 0 & 0 & \dots & 0 \\ 0 & \frac{1}{\sqrt{R_2}} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & \frac{1}{\sqrt{R_{40}}} \end{bmatrix};$$

Las matrices transformadas $X^\star = P^{-1}X$ y $Y^\star = P^{-1}Y$ son las siguientes:

$$X^\star = P^{-1}X = \begin{bmatrix} \frac{1}{\sqrt{R_1}} & \sqrt{R_1} \\ \frac{1}{\sqrt{R_2}} & \sqrt{R_2} \\ \vdots & \vdots \\ \frac{1}{\sqrt{R_{40}}} & \sqrt{R_{40}} \end{bmatrix}; \qquad Y^\star = P^{-1}Y = \begin{bmatrix} \frac{C_1}{\sqrt{R_1}} \\ \frac{C_2}{\sqrt{R_2}} \\ \vdots \\ \frac{C_{40}}{\sqrt{R_{40}}} \end{bmatrix};$$

La expresión matricial del estimador MCG para el modelo sería:

$$\begin{aligned} \hat{\beta}_{MCO} &= (X^{\star\prime}X^\star)^{-1}X^{\star\prime}Y^\star = \\ &= ((P^{-1}X)'(P^{-1}X))^{-1}((P^{-1}X)'(P^{-1}Y)) = \\ &= (X'\Omega^{-1}X)^{-1}(X'\Omega^{-1}Y) = \hat{\beta}_{MCG} \end{aligned}$$

luego:

$$\hat{\beta}_{MCG} = \begin{bmatrix} \sum_1^{40} \frac{1}{R_i} & N \\ N & \sum_1^{40} R_i \end{bmatrix}^{-1} \begin{bmatrix} \sum_1^{40} \frac{C_i}{R_i} \\ \sum_1^{40} C_i \end{bmatrix}$$

La matriz de varianzas y covarianzas del estimador de MCG sería:

$$V(\hat{\beta}_{MCG}) = \sigma^2 (X'\Omega^{-1}X)^{-1} = \sigma^2 \begin{bmatrix} \sum_1^{40} \frac{1}{R_i} & N \\ N & \sum_1^{40} R_i \end{bmatrix}^{-1}$$

Un estimador insesgado de dicha matriz de varianzas y covarianzas sería:

$$\widehat{V}(\hat{\beta}_{MCG}) = \hat{\sigma}^2_{MCG}\, (X'\Omega^{-1}X)^{-1}$$

donde:

$$\hat{\sigma}^2_{MCG} = \frac{\hat{u}'_{MCG}\,\Omega^{-1}\,\hat{u}_{MCG}}{40-K} = \frac{(Y - X\hat{\beta}_{MCG})'\,\Omega^{-1}\,(Y - X\hat{\beta}_{MCG})}{40-K} = \frac{Y'\,\Omega^{-1}\,Y - \hat{\beta}'_{MCG}X'\,\Omega^{-1}\,Y}{40-K}$$

siendo en este caso $Y'\,\Omega^{-1}Y = \sum_1^{40} \frac{C_i^2}{R_i}$.

Resultados de la estimación del modelo[6]:

$$C_i = \beta_1 + \beta_2 R_i + u_i \qquad i = 1, 2, \ldots, 40$$

- por MCO, luego sin tener en cuenta la existencia de heterocedasticidad

$$\begin{array}{cccccc} \widehat{C}_i & = & 40,7676 & + & 0,1282\;\; R_i & R^2 = 0,3171 \\ (\widehat{des}(\hat{\beta}_{i,MCO})) & & (22{,}1387) & & (0{,}0305) & \end{array} \tag{3.12}$$

- por MCG, luego teniendo en cuenta la heterocedasticidad y suponiendo $Var(u_i) = \sigma^2 R_i$:

$$\begin{array}{cccccc} \hat{C}_i & = & 31,9243 & + & 0,1409\;\; R_i & R^2 = 0,4177 \\ (\widehat{des}(\hat{\beta}_{i,MCG})) & & (17{,}9860) & & (0{,}0269) & \end{array} \tag{3.13}$$

- Con respecto a los resultados de la estimación, en (3.12) no podemos decir nada ya que hemos detectado la existencia de heterocedasticidad y por lo tanto, las estimaciones $\hat{\beta}_{i,MCO}$ mostradas no son las mejores posibles ya que han sido obtenidas con un estimador no eficiente y en cuanto a $\widehat{des}(\hat{\beta}_{i,MCO})$, corresponden a un estimador sesgado por lo que no son adecuadas para hacer inferencia.
  Los resultados mostrados en (3.13) tienen en cuenta la existencia de heterocedasticidad. Si la forma funcional propuesta es correcta, el estimador utilizado, $\hat{\beta}_{MCG}$, es un estimador eficiente y las estimaciones $\widehat{des}(\hat{\beta}_{i,MCG})$ corresponden a un estimador insesgado y son válidas para hacer inferencia. Los R^2 no son comparables, ni tampoco las desviaciones estimadas de los coeficientes estimados, $\widehat{des}(\hat{\beta}_i)$.

Al estimador de MCG bajo heterocedasticidad también se le conoce con el nombre de Mínimos Cuadrados Ponderados, MCP, porque pondera las observaciones inversamente al peso de la varianza de la perturbación. En la suma de cuadrados de (3.10) se ponderan más las desviaciones o residuos $[C_i - \widehat{E(C_i)}]$ con menor varianza que las de mayor varianza o dispersión. Esto es muy fácil de entender en el ejemplo desarrollado si observamos la función objetivo, la forma de la matriz P^{-1} y en definitiva, las variables del modelo transformado.

[6]Ver Anexo 3.1, Resultados 1 y 4.

Ejercicio 3.1 Sea el modelo:

$$C_i = \beta_1 + \beta_2 R_i + u_i$$

donde $E(u_i) = 0 \quad \forall i$; $E(u_i^2) = a + bR_i$; $\quad E(u_i u_j) = 0 \quad \forall i, j \quad i \neq j$. Escribe el correspondiente modelo transformado y demuestra las propiedades de la perturbación de dicho modelo suponiendo que a y b son conocidas.

3.4. Estimador de Mínimos Cuadrados Generalizados Factibles. Especificación de un modelo para la heterocedasticidad

En la sección anterior la matriz de varianzas y covarianzas de la perturbación es conocida excepto por un factor de escala σ^2. Cuando los elementos de $E(uu')$ no son conocidos, $E(uu') = \sigma^2 \Omega = \Sigma$, no es posible estimar N varianzas más K coeficientes de regresión con N observaciones[7]. Una forma de abordar el problema es modelar la varianza de la perturbación en función de un conjunto de variables observables, Z_i, que pueden ser o no regresores del modelo, y de un vector de parámetros desconocido θ, cuya dimensión es estimable y no crece con el tamaño muestral N:

$$\sigma_i^2 = g(Z_i, \theta), \quad \forall i \quad \text{de forma que} \quad \Sigma = \Sigma(\theta)$$

Una vez obtenido un estimador de θ, $\hat{\theta}$, se puede definir un estimador de $\Sigma, \widehat{\Sigma} = \Sigma(\hat{\theta})$, y estimar el vector de coeficientes β por el método de Mínimos Cuadrados Generalizados Factibles, MCGF. Bajo ciertas condiciones de regularidad, si el estimador $\widehat{\Sigma}$ es consistente, el estimador de MCGF tiene buenas propiedades asintóticas.

El estimador de MCGF es un estimador en dos etapas. En la primera etapa utilizamos los residuos de MCO para estimar θ. Dado que el residuo es una aproximación a la perturbación,

$$\hat{u}_i = Y_i - X_i'\hat{\beta}_{MCO} = Y_i - X_i'\beta - X_i'(\hat{\beta}_{MCO} - \beta) = u_i + \text{error}$$

y que hemos supuesto $E(u_i^2) = \sigma_i^2 = g(Z_i, \theta)$, entonces podemos definir la ecuación $u_i^2 = g(Z_i, \theta) +$ error. Si $g(Z_i, \theta)$ es lineal en θ, por ejemplo $\theta' Z_i$, se puede considerar la siguiente regresión auxiliar para estimar los parámetros θ:

$$\hat{u}_i^2 = \theta' Z_i + \zeta_i \qquad i = 1, \ldots, N$$

Bajo ciertas condiciones de regularidad, se puede demostrar que el estimador MCO de θ es consistente. Una vez obtenido el estimador consistente de θ, y por lo tanto, de $\hat{\sigma}_i^2 = g(Z_i, \hat{\theta})$, en la segunda etapa, se sustituye en la función suma de cuadrados ponderada y se minimiza con respecto a β:

$$\underset{\hat{\beta}}{Min}\ (Y - X\hat{\beta})'\widehat{\Sigma}^{-1}(Y - X\hat{\beta}) = \underset{\hat{\beta}}{Min} \sum_{i=1}^{N} \frac{1}{\hat{\sigma}_i^2}(Y_i - X_i'\hat{\beta})^2$$

[7]Notar que cuando los elementos de Ω no son conocidos difícilmente será conocida σ^2, lo lógico es pensar que nada en las varianzas es conocido, ni tan siquiera que exista una parte común en todas ellas, así es habitual denotar a $E(uu')$ como Σ.

obteniendo el estimador de MCGF definido como:

$$\hat{\beta}_{MCGF} = (X'\widehat{\Sigma}^{-1}X)^{-1}X'\widehat{\Sigma}^{-1}Y$$

• **Veamos una aplicación de este procedimiento utilizando el ejercicio magistral:**
El modelo a estimar es:

$$C_i = \beta_1 + \beta_2 R_i + u_i \qquad i = 1, 2, \ldots, 40 \tag{3.14}$$

donde $E(u_i) = 0 \quad \forall i$ y $E(u_i u_j) = 0 \quad \forall i, j \quad i \neq j$. Suponemos que $Var(u_i) = a + bR_i$ siendo a y b constantes desconocidas. La forma funcional que hemos supuesto para la varianza de la perturbación es razonable con respecto a la información disponible, existencia de heterocedasticidad y dependencia creciente con respecto a R_i y depende de un número de parámetros pequeño que son estimables de forma consistente como veremos a continuación.

La estructura de la matriz de varianzas y covarianzas del vector de perturbaciones, que evidencia que Σ es desconocida, es la siguiente:

$$E(uu') = \begin{bmatrix} a + bR_1 & 0 & 0 & \ldots & 0 \\ 0 & a + bR_2 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & 0 & \ldots & a + bR_{40} \end{bmatrix} = \Sigma$$

¿Cómo estimamos los parámetros desconocidos del modelo anterior?
Nuestro problema ahora es cómo estimar los parámetros desconocidos β_1, β_2, a y b del modelo. El modelo transformado para la forma funcional propuesta para la varianza de u_i es:

$$\frac{C_i}{\sqrt{a + bR_i}} = \beta_1 \frac{1}{\sqrt{a + bR_i}} + \beta_2 \frac{R_i}{\sqrt{a + bR_i}} + \frac{u_i}{\sqrt{a + bR_i}} \qquad i = 1, \ldots, 40 \tag{3.15}$$

En este caso en el modelo transformado tenemos dos constantes desconocidas a y b que deben ser previamente estimadas para poder aplicar el estimador de MCO al modelo (3.15). Para obtener estimadores consistentes de a y b podemos proceder de la forma siguiente:

1. Estimamos por MCO el modelo (3.9) y guardamos los residuos de mínimos cuadrados ordinarios.

2. Estimamos por MCO la siguiente regresión auxiliar:

   $$\hat{u}^2_{MCO,i} = a + bR_i + \epsilon_i \qquad i = 1, 2, \ldots, 40$$

   De esta regresión obtenemos $\hat{a}_{MCO}$ y $\hat{b}_{MCO}$ estimados consistentemente.

Una vez estimados consistentemente $\hat{a}_{MCO}$ y $\hat{b}_{MCO}$ podemos estimar por MCO, o lo que es lo mismo MCGF, el siguiente modelo transformado:

$$\frac{C_i}{\sqrt{\hat{a} + \hat{b}R_i}} = \beta_1 \frac{1}{\sqrt{\hat{a} + \hat{b}R_i}} + \beta_2 \frac{R_i}{\sqrt{\hat{a} + \hat{b}R_i}} + \frac{u_i}{\sqrt{\hat{a} + \hat{b}R_i}} \qquad i = 1, \ldots, 40 \tag{3.16}$$

Los estimadores $(\hat{\beta}_1, \hat{\beta}_2)_{MCGF}$ así obtenidos son consistentes dado que los estimadores de a y b a su vez lo son. Además, serán asintóticamente eficientes. En muestras pequeñas no se conocen sus propiedades porque MCGF es un estimador no lineal. No obstante, al ser un estimador consistente y tener distribución asintótica conocida podemos utilizarlo para hacer inferencia asintótica válida, como veremos más adelante.

A la hora de implementar el estimador de MCGF es fundamental tener en cuenta que Σ es una matriz de varianzas y covarianzas, luego la variable de ponderación es $\hat{\sigma}_i^2$ y ésta debe ser positiva $\forall i$. Es necesario comprobarlo, ya que si a la hora de estimar σ_i^2 no se impone esta restricción no tiene porqué satisfacerse.

En el caso del ejemplo $\hat{\sigma}_i^2 = \hat{a} + \hat{b}R_i$ no es positiva para las 40 familias de la muestra, por lo que la forma funcional elegida para la varianza $Var(u_i) = a + bR_i$ a pesar de ser razonable para la muestra, no es implementable. Debemos probar otras, que también sean razonables claro está, por ejemplo $Var(u_i) = (a + bR_i)^2$. Para esta forma funcional la regresión auxiliar es:

$$\hat{u}_{MCO,i}^2 = a + bR_i + cR_i^2 + \epsilon_i \qquad i = 1, 2, \ldots, 40$$

El correspondiente modelo transformado una vez estimados consistentemente a, b y c es:

$$\frac{C_i}{\sqrt{\hat{a} + \hat{b}R_i + \hat{c}R_i^2}} = \beta_1 \frac{1}{\sqrt{\hat{a} + \hat{b}R_i + \hat{c}R_i^2}} + \beta_2 \frac{R_i}{\sqrt{\hat{a} + \hat{b}R_i + \hat{c}R_i^2}} + \frac{u_i}{\sqrt{\hat{a} + \hat{b}R_i + \hat{c}R_i^2}}$$

$$i = 1, \ldots, 40$$

La estimación del modelo por MCO, o lo que igual MCGF para esta última forma funcional de la varianza de la perturbación proporciona los siguientes resultados[8]:

- Modelo estimado por MCGF:

$$\begin{array}{rcccc} \hat{C}_i & = & 34,2386 & + & 0,1410\ R_i \\ (\widehat{des}(\hat{\beta}_{i,MCGF})) & & (17{,}6042) & & (0{,}0289) \end{array}$$

  Si observamos los resultados de la estimación vemos como cuantitativamente las realizaciones muestrales de los estimadores MCG y MCGF son muy similares. Sin embargo, es importante recalcar que las varianzas de los estimadores MCG y MCGF no son comparables entre sí. El marco de trabajo es distinto. Con el estimador MCG estamos trabajando en términos de propiedades en muestras finitas de un estimador y bajo un supuesto concreto en la varianza de la perturbación y en base a su certeza se demuestran las propiedades del estimador. Con el estimador de MCGF al ser no lineal buscamos propiedades asintóticas bajo el supuesto de que la forma funcional propuesta para la varianza de la perturbación es adecuada y además, se ha estimado consistentemente. Pero si el supuesto sobre $Var(u_i)$ es adecuado con MCGF, éste es asintóticamente equivalente a MCG, en el sentido de que ambos estimadores tienen la misma distribución asintótica.

[8]Ver Anexo 3.1, Resultado 5.

3.5. MCO: Estimador de $V(\hat{\beta}_{MCO})$ robusto a heterocedasticidad

En presencia de heterocedasticidad el estimador Mínimo Cuadrático Ordinario es lineal, insesgado y consistente, pero no es de varianza mínima. Su matriz de varianzas y covarianzas se define $V(\hat{\beta}_{MCO}) = \sigma^2(X'X)^{-1}X'\Omega X(X'X)^{-1}$. Un estimador de la matriz de varianzas y covarianzas del estimador MCO es $\widehat{V}(\hat{\beta}_{MCO}) = \hat{\sigma}^2(X'X)^{-1}$. Utilizar este estimador para hacer inferencia cuando hay heterocedasticidad no es adecuado. Los estadísticos t y F habituales para hacer inferencia sobre β definidos en base a este estimador de la matriz de varianzas y covarianzas de $\hat{\beta}_{MCO}$ son inapropiados ya que es un estimador sesgado e inconsistente.

La dificultad que entraña el conocimiento de Ω, o Σ en su caso, hace interesante el poder contar con un estimador de $V(\hat{\beta}_{MCO})$ consistente y robusto a la posible existencia de heterocedasticidad y de esta forma derivar estadísticos válidos, al menos asintóticamente, para contrastar hipótesis sobre el vector de coeficientes β utilizando $\hat{\beta}_{MCO}$.

White (1980) demuestra que un estimador consistente de la matriz de varianzas y covarianzas asintótica de $\hat{\beta}_{MCO}$ en presencia de heterocedasticidad es:

$$\widehat{V}(\hat{\beta}_{MCO})_{White} = (X'X)^{-1}(X'SX)(X'X)^{-1}$$

donde $S = diag(\hat{u}_1^2, \hat{u}_2^2, \ldots, \hat{u}_N^2)$. Esta matriz de varianzas y covarianzas consistente asintóticamente puede ser utilizada para hacer inferencia válida al menos asintóticamente utilizando $\hat{\beta}_{MCO}$ sin tener que especificar a priori la estructura de heterocedasticidad.

- Aplicación del estimador de White de la matriz de varianzas y covarianzas del estimador MCO al ejercicio propuesto[9]:

$$\begin{array}{ccccccc} \widehat{C}_i & = & 40,7676 & + & 0,1282 & R_i \\ \widehat{des}(\hat{\beta}_{i,MCO}) & & (22,1387) & & (0,0305) & \\ \widehat{des}(\hat{\beta}_{i,MCO})_{White} & & (27,3213) & & (0,0439) & \end{array} \quad (3.17)$$

  Notar que los coeficientes han sido estimados por MCO, estimador lineal, insesgado, consistente y no de varianza mínima dado que $E(u_i^2) = \sigma_i^2$. El estimador, $\widehat{des}(\hat{\beta}_{i,MCO})$ se obtiene de $\hat{\sigma}^2(X'X)^{-1}$ estimador sesgado e inconsistente bajo heterocedasticidad. Sin embargo, $\widehat{des}(\hat{\beta}_{i,MCO})_{White}$ es un estimador consistente de $des(\hat{\beta}_{i,MCO})$ y ha sido obtenido con la expresión $(X'X)^{-1}(X'SX)(X'X)^{-1}$. En este ejercicio en concreto se observa que con el estimador $\widehat{des}(\hat{\beta}_{i,MCO}) = \hat{\sigma}^2(X'X)^{-1}$, sesgado e inconsistente se estaba subestimando la desviación típica poblacional de $\hat{\beta}_{MCO}$, que es la raíz cuadrada del elemento i-ésimo de la diagonal principal de $\hat{\sigma}^2(X'X)^{-1}$.

- El problema de estimación de los parámetros de interés como una cuestión de objetivos e información disponible:

Si conocemos la forma funcional de la varianza de la perturbación, es decir, si conocemos Ω y si nuestro objetivo es obtener estimadores **lineales, insesgados y de varianza mínima** estimaremos

[9] Anexo 3.1, Resultado 6.

por MCG. Si no conocemos Ω habremos de estimarla y una vez estimada podremos sustituirla en la expresión del estimador. En este caso nuestro estimador es MCGF, definido como:

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

Bajo ciertas condiciones suficientes el estimador es consistente y podremos obtener una distribución asintótica conocida que nos permitirá realizar inferencia asintótica.

En ocasiones modelizar Ω, o estimarla si es desconocida no resulta fácil. En estas ocasiones puede resultar preferible quedarnos con el estimador de los coeficientes β por MCO y ocuparnos de cómo realizar inferencia válida con este estimador en estas circunstancias. En este caso una solución podría ser obtener un estimador de la matriz de varianzas y covarianzas de los estimadores MCO con adecuadas propiedades, en concreto consistente y robusto a la posible existencia de heterocedasticidad. La utilización del estimador MCO, consistente en presencia de heterocedasticidad y la matriz de varianzas y covarianzas propuesta nos permitiría realizar inferencia asintótica sin tener que establecer la forma funcional de la heterocedasticidad.

Hay que notar que el estimador de White permite a los investigadores no arriesgarse a manipular los datos a la búsqueda de una forma funcional de la heterocedasticidad cuando esta no está muy clara y realizar inferencia asintótica. Sin embargo, no soluciona la cuestión de la ineficiencia del estimador.

3.6. Contraste de restricciones lineales

Una vez hemos estimado adecuadamente los coeficientes del modelo de interés es probable que nuestro objetivo sea realizar contraste de hipótesis. En el tema anterior ya se vio con detalle cómo contrastar restricciones lineales[10]. Sean las hipótesis nula y alternativa para el contraste de q restricciones lineales:

$$H_0 : R\beta = r$$
$$H_a : R\beta \neq r$$

donde R es una matriz $(q \times K)$ y r es un vector de dimensión $(q \times 1)$, siendo q el número de restricciones lineales a contrastar.

• **Estadísticos basados en el estimador de β por MCG o MCP.**
En este caso

$$Y = X\beta + u \qquad u \sim N(0, \sigma^2\Omega) \quad \text{y} \quad \Omega \quad \text{es conocida}$$

el estimador MCG es lineal en u, insesgado y de varianza mínima, su distribución en muestras finitas es:

$$\hat{\beta}_{MCG} \sim N(\beta, \sigma^2(X'\Omega^{-1}X)^{-1})$$

Estadísticos de contraste y distribución asociada:

[10] La notación utilizada para el tamaño muestral en ese tema era T en lugar de N.

Para $q = 1$

$$\frac{R\hat{\beta}_{MCG} - r}{\sqrt{R\,\hat{\sigma}^2(X'\Omega^{-1}X)^{-1}\,R'}} \overset{H_0}{\sim} t_{(N-K)}$$

para $q > 1$

$$\frac{(R\hat{\beta}_{MCG} - r)'\,[R\,(X'\Omega^{-1}X)^{-1}\,R'\,]^{-1}\,(R\hat{\beta}_{MCG} - r)/q}{\hat{\sigma}^2} \overset{H_0}{\sim} F_{(q,N-K)}$$

siendo un estimador insesgado de σ^2, $\hat{\sigma}^2_{MCG} = \dfrac{(Y - X\hat{\beta}_{MCG})'\,\Omega^{-1}(Y - X\hat{\beta}_{MCG})}{N-K}$.

Las reglas de decisión son las habituales.

Ejemplo 3.8 En el modelo $C_i = \beta_1 + \beta_2 R_i + u_i \quad u_i \sim N(0, \sigma^2 R_i)$ contrastamos la significatividad de la variable R_i con el estadístico y distribución siguientes:

$$H_0 : \beta_2 = 0 \quad \text{versus} \quad H_a : \beta_2 \neq 0 \qquad \frac{\hat{\beta}_{2,MCG}}{\widehat{des}(\hat{\beta}_{2,MCG})} \overset{H_0}{\sim} t_{(40-2)}$$

Rechazamos la hipótesis nula de no significatividad de la variable renta para $\alpha = 5\,\%$ si el valor muestral del estadístico t es tal que t $> t_{(40-2)|0,025}$.

Ejercicio 3.2 Sea el modelo:

$$Y = X\beta + u \qquad u \sim N(0, \Sigma) \qquad \Sigma \text{ conocida}$$

Deriva la distribución del estimador MCG de β y el estadístico de contraste y su distribución asociada para contrastar $H_0 : R\beta = r$.

Ejercicio 3.3 Sea el modelo

$$Y = X\beta + u \qquad u \sim N(0, \sigma^2\Omega) \quad \text{y} \quad \Omega \quad \text{es conocida}$$

En el modelo transformado correspondiente, deriva la distribución del estimador MCO de β y el estadístico de contraste y su distribución asociada para contrastar $H_0 : R\beta = r$.

• Estadísticos basados en el estimador de β por MCGF

En este caso

$$Y = X\beta + u \qquad u \sim N(0, \Sigma) \qquad \Sigma \text{ desconocida, pero estimable consistentemente}$$

El estimador de MCGF, $\hat{\beta}_{MCGF} = (X'\widehat{\Sigma}^{-1}X)^{-1}(X'\widehat{\Sigma}^{-1}Y)$, es un estimador no lineal y en general sesgado, cuya distribución en muestras finitas no es conocida. En muestras grandes es un estimador consistente si $\widehat{\Sigma}$ es un estimador consistente de Σ. Su distribución asintótica es:

$$\sqrt{N}(\hat{\beta}_{MCGF} - \beta) \xrightarrow{d} N(0, G^{-1})$$

Los estadísticos de contraste y distribución asociada utilizando como estimador de $G^{-1} = \text{plim}\left[\frac{1}{N} X'\Sigma X\right]^{-1}$ a $N\widehat{V}(\hat{\beta}_{MCGF}) = N(X'\widehat{\Sigma}^{-1}X)^{-1}$ son:

para $q = 1$

$$\frac{R\hat{\beta}_{MCGF} - r}{\sqrt{R\ (X'\widehat{\Sigma}^{-1}X)^{-1}\ R'}} \xrightarrow{d,H_0} N(0,1)$$

para $q > 1$

$$(R\hat{\beta}_{MCGF} - r)'[\ R\ (X'\widehat{\Sigma}^{-1}X)^{-1}\ R'\ ]^{-1}(R\hat{\beta}_{MCGF} - r) \xrightarrow{d,H_0} \chi^2_q$$

Las reglas de decisión son las habituales.

Ejemplo 3.9 En el modelo $C_i = \beta_1 + \beta_2 R_i + u_i \quad u_i \sim N(0, a + bR_i)$ contrastamos la significatividad de la variable R_i con el estadístico y distribución siguientes:

$$H_0 : \beta_2 = 0 \quad \text{versus} \quad H_a : \beta_2 \neq 0$$

$$\frac{\hat{\beta}_{2,MCGF}}{\widehat{des}(\hat{\beta}_{2,MCGF})} \xrightarrow{d,H_0} N(0,1)$$

Rechazamos la hipótesis nula de no significatividad de la variable renta para $\alpha = 5\,\%$ si el valor muestral del estadístico t es tal que t $> N(0,1)_{|\frac{0,05}{2}}$.

• **Estadísticos basados en el estimador de β por MCO y un estimador robusto a heterocedasticidad de $V(\hat{\beta}_{MCO})$**

En este caso

$$Y = X\beta + u \qquad u \sim N(0, \Sigma) \qquad \Sigma \text{ desconocida y difícil de modelar y estimar consistentemente}$$

Dado que el estimador MCO es consistente y $\widehat{V}(\hat{\beta}_{MCO})_{White}$ también lo es bajo heterocedasticidad podemos utilizarlos conjuntamente para hacer inferencia asintótica válida. En este caso el estadístico válido para realizar contraste de hipótesis de la forma $H_0 : R\beta = r$ y su distribución asintótica asociada son:

$$(R\hat{\beta}_{MCO} - r)'(R\,\widehat{V}(\hat{\beta}_{MCO})_{White}\ R')^{-1}(R\hat{\beta}_{MCO} - r) \xrightarrow{d,H_0} \chi^2_q$$

Siendo q el número de restricciones de contraste. Para $q = 1$ podemos escribir el estadístico anterior como:

$$\frac{R\hat{\beta}_{MCO} - r}{\sqrt{R\,\widehat{V}(\hat{\beta}_{MCO})_{White}\ R'}} \xrightarrow{d,H_0} N(0,1)$$

Ejemplo 3.10 En el modelo $C_i = \beta_1 + \beta_2 R_i + u_i \quad u_i \sim N(0, \sigma_i^2)$, siendo σ_i^2 desconocida, contrastamos la significatividad de la variable R_i utilizando el estimador MCO de β_2 con el estadístico y distribución siguientes:

$$H_0 : \beta_2 = 0 \quad \text{versus} \quad H_a : \beta_2 \neq 0$$

$$\frac{\hat{\beta}_{2,MCO}}{\widehat{des}(\hat{\beta}_{2,MCO})_{White}} \xrightarrow{d,H_0} N(0,1)$$

Rechazamos la hipótesis nula de no significatividad de la variable renta para $\alpha = 5\,\%$ si el estadístico calculado t es tal que t $> N(0,1)_{|\frac{0,05}{2}}$.

3.7. Resumen de los resultados obtenidos en el ejercicio magistral

En esta sección vamos a resumir los resultados obtenidos en este tema. Para ello vamos utilizar el ejercicio magistral. En el ejemplo nos proponíamos estudiar la relación entre consumo y renta para lo cual se disponía de 40 observaciones de ambas variables.

$$C_i = \beta_1 + \beta_2 R_i + u_i \qquad i = 1, \ldots, 40$$

Los resultados de la estimación MCO son:

$$\begin{array}{ccccc} \widehat{C}_i & = & 40,7676 & + & 0,1282 \;\; R_i \\ (\widehat{des}(\hat{\beta}_{i,MCO})) & & (22{,}1387) & & (0{,}0305) \end{array}$$
$$R^2 = 0,317118 \qquad GQ = 3,34 \qquad BP = 11,28$$

Dados lo valores muestrales de los estadísticos GQ y BP, para ambos rechazamos la hipótesis nula de homocedasticidad. La perturbación del modelo es heterocedástica y el estimador MCO utilizado para estimar, aunque es un estimador lineal, insesgado y consistente no es de varianza mínima. Las desviaciones típicas estimadas para los parámetros se han calculado en base a la expresión $\widehat{V}(\hat{\beta}_{MCO}) = \hat{\sigma}^2(X'X)^{-1}$, estimador sesgado e inconsistente de $V(\hat{\beta}_{MCO})$ por lo que no deben ser utilizadas para hacer inferencia. Los estadísticos t y F habituales para realizar inferencia no siguen las distribuciones t-Student y F-Snedecor que les corresponden.

A partir de la Figura 3.8 y de los resultados de los contrastes de heterocedasticidad se ha supuesto que $Var(u_i) = \sigma_i^2 = \sigma^2 R_i$ y se ha reestimado el modelo por MCP con los siguientes resultados:

$$\begin{array}{ccccc} \hat{C}_i & = & 31,9243 & + & 0,1409 \;\; R_i \qquad R^2 = 0,417757 \\ (\widehat{des}(\hat{\beta}_{i,MCG})) & & (17{,}9860) & & (0{,}0269) \end{array}$$

Si la forma funcional propuesta para σ_i^2 es correcta, el estimador MCP es un estimador lineal, insesgado y de varianza mínima. Si la perturbación es normal, el estimador MCP sigue una distribución normal en muestras finitas, $\hat{\beta}_{MCG} \sim N(\beta, \sigma^2(X'\Omega^{-1}X)^{-1})$ válida para hacer inferencia. El estadístico calculado para la $H_0 : \beta_2 = 0$ frente a $H_a : \beta_2 \neq 0$ es $\text{t} = 4,20 > 2,02 = t_{(40-2)|0,025}$, luego para un nivel de significatividad $\alpha = 5\,\%$ la variable es significativa para explicar el consumo de los individuos.

Respecto a la bondad del ajuste no podemos decir nada. Bajo heterocedasticidad el coeficiente R^2 calculado en el modelo original no tiene la interpretación que conocemos. Además, hay que tener mucho cuidado con su interpretación cuando trabajamos en el modelo transformado ya que en muchos casos tras la transformación, el modelo no tiene término independiente.

Una alternativa a la forma funcional propuesta para $Var(u_i)$ y coherente tanto con la Figura 3.8 como con los resultados de los estadísticos de contraste utilizados para contrastar la existencia de heterocedasticidad es $Var(u_i) = (a + bR_i)^2$. En este caso hay dos coeficientes desconocidos en la forma funcional propuesta por lo que el estimador MCG (o MCP) no es aplicable. El modelo debe ser estimado por MCGF y los resultados de dicha estimación son:

$$\begin{array}{ccccc} \hat{C}_i & = & 34,2386 & + & 0,1410 \;\; R_i \\ (\widehat{des}(\hat{\beta}_{i,MCGF})) & & (17{,}6042) & & (0{,}0289) \end{array}$$

Si la forma funcional propuesta es correcta el estimador MCGF es un estimador no lineal y consistente si $\hat{\sigma}_i^2$ a su vez es consistente. Además, es eficiente asintóticamente y tiene distribución asintótica conocida y válida para hacer inferencia asintótica. En este caso el estadístico calculado para la $H_0 : \beta_2 = 0$ frente a $H_a : \beta_2 \neq 0$ es t $= 4,87 > 1,96 = N(0,1)_{|0{,}025}$, luego para un nivel de significatividad $\alpha = 5\,\%$ la variable renta es significativa asintóticamente para explicar el consumo de los individuos.

Los estimadores MCG y MCGF propuestos son correctos y adecuados para las formas funcionales propuestas para la varianza de la perturbación, cada uno en su contexto. Sin embargo, si proponemos $Var(u_i) = \sigma^2 R_i$ y la forma funcional correcta es otra distinta, por ejemplo $Var(u_i) = (a + bR_i)^2$ el estimador MCG no posee las propiedades indicadas. Y viceversa. Es por ello, que ante esta dificultad, en muchas ocasiones sea más adecuado no realizar ningún supuesto sobre $Var(u_i)$, y utilizar el estimador de MCO combinado con un estimador consistente de su matriz de varianzas y covarianzas bajo heterocedasticidad y realizar inferencia con ellos. Para el caso de heterocedasticidad el estimador robusto a heterocedasticidad de $V(\hat{\beta}_{MCO})$ es el estimador de White. Si optamos por esta solución para estimar y realizar inferencia, los resultados obtenidos son:

$$\begin{array}{ccccc} \widehat{C}_i & = & 40,7676 & + & 0,1282 \;\; R_i \\ (\widehat{des}(\hat{\beta}_{i,MCO})_{White}) & & (27{,}3213) & & (0{,}0439) \end{array}$$

El estimador MCO utilizado para estimar los coeficientes del modelo es lineal, insesgado y no es de varianza mínima. Sin embargo, es consistente, al igual que lo son las desviaciones típicas estimadas con el estimador de White. Por ello podemos utilizar el estadístico t-habitual con distribución asintótica $N(0,1)$ para contrastar la significatividad de la variable renta, $H_0 : \beta_2 = 0$ frente a $H_a : \beta_2 \neq 0$. En este caso t $= 2,91 > 1,96 = N(0,1)_{|0{,}025}$, luego para un nivel de significatividad $\alpha = 5\,\%$ la variable es significativa para explicar el consumo de los individuos.

3.8. Anexo 3.1: Resultados de gretl utilizados en las clases magistrales

• **Resultado 1:** Estimación por MCO:

Estimaciones MCO utilizando las 40 observaciones 1–40
Variable dependiente: consumo

	Coeficiente	Desv. típica	estadístico t	valor p
const	40,7676	22,1387	1,8415	0,0734
renta	0,1282	0,0305	4,2008	0,0002

Media de la var. dependiente	130,3130
D.T. de la variable dependiente	45,1586
Suma de cuadrados de los residuos	54311,30
Desviación típica de la regresión ($\hat{\sigma}$)	37,8054
R^2	0,3171
$\bar{R}^2$ corregido	0,2991
Grados de libertad	38

• **Resultado 2:** Regresiones parciales para computar el estadístico de Goldfeld y Quandt:

Estimaciones MCO utilizando las 20 observaciones 1–20
Variable dependiente: consumo

	Coeficiente	Desv. típica	estadístico t	valor p
const	12,9388	28,9666	0,4467	0,6604
renta	0,1823	0,0520	3,5023	0,0025

Media de la var. dependiente	112,3040
D.T. de la variable dependiente	32,9714
Suma de cuadrados de los residuos	12284,2
Desviación típica de la regresión ($\hat{\sigma}$)	26,1238
R^2	0,4052
$\bar{R}^2$ corregido	0,3722
Grados de libertad	18

Estimaciones MCO utilizando las 20 observaciones 21–40
Variable dependiente: consumo

	Coeficiente	Desv. típica	estadístico t	valor p
const	48,1767	70,2419	0,6859	0,5015
renta	0,1176	0,0815	1,4425	0,1663

Media de la var. dependiente	148,322
D.T. de la variable dependiente	49,1527
Suma de cuadrados de los residuos	41146,9
Desviación típica de la regresión ($\hat{\sigma}$)	47,8115
R^2	0,1036
$\bar{R}^2$ corregido	0,0538
Grados de libertad	18

• **Resultado 3:** Regresión auxiliar para computar el estadístico de Breusch Pagan:

Estimaciones MCO utilizando las 40 observaciones 1–40
Variable dependiente: enorm

	Coeficiente	Desv. típica	estadístico t	valor p
const	−1,6788	0,6876	−2,4415	0,0194
renta	0,0038	0,0009	4,0461	0,0002

Media de la var. dependiente	1,0000
D.T. de la variable dependiente	1,3864
Suma de cuadrados de los residuos	52,3940
Desviación típica de la regresión ($\hat{\sigma}$)	1,1742
R^2	0,3010
$\bar{R}^2$ corregido	0,2827
Grados de libertad	38

• **Resultado 4:** Estimación por MCP:

Estimaciones MC.Ponderados utilizando las 40 observaciones 1–40
Variable dependiente: consumo
Variable utilizada como ponderación: 1/renta

	Coeficiente	Desv. típica	estadístico t	valor p
const	31,9244	17,9861	1,7749	0,0839
renta	0,1409	0,0269	5,2216	0,0000

Estadísticos basados en los datos ponderados:

Suma de cuadrados de los residuos	68,7020
Desviación típica de la regresión ($\hat{\sigma}$)	1,3446
R^2	0,4177
$\bar{R}^2$ corregido	0,4024
Grados de libertad	38

Estadísticos basados en los datos originales:

Media de la var. dependiente	130,313
D.T. de la variable dependiente	45,1586
Suma de cuadrados de los residuos	54557,3
Desviación típica de la regresión ($\hat{\sigma}$)	37,8909

• **Resultado 5:** Regresión auxiliar para el estimador de MCGF:

Estimaciones MCO utilizando las 40 observaciones 1–40
Variable dependiente: $\hat{u}^2_{i,MCO}$

	Coeficiente	Desv. típica	estadístico t	valor p
const	1923,6000	2365,4800	0,8132	0,4213
renta	−7,4202	6,6882	−1,1094	0,2744
sq_renta	0,0087	0,0045	1,9221	0,0623

Media de la var. dependiente	1357,78
D.T. de la variable dependiente	1882,48
Suma de cuadrados de los residuos	8,78226e+07
Desviación típica de la regresión ($\hat{\sigma}$)	1540,64
R^2	0,3645
$\bar{R}^2$ corregido	0,3302
$F(2, 37)$	10,6133
valor p para $F()$	0,0002

• Estimación por MCGF:

Estimaciones MC.Ponderados utilizando las 40 observaciones 1–40
Variable dependiente: consumo
Variable utilizada como ponderación: $\hat{\sigma}^2_i$

	Coeficiente	Desv. típica	estadístico t	valor p
const	34,2386	17,6042	1,9449	0,0592
renta	0,1410	0,0289	4,8757	0,0000

Estadísticos basados en los datos ponderados:

Suma de cuadrados de los residuos	39,3094
Desviación típica de la regresión ($\hat{\sigma}$)	1,0170
R^2	0,3848
$\bar{R}^2$ corregido	0,3686
Grados de libertad	38

Estadísticos basados en los datos originales:

Media de la var. dependiente	130,3130
D.T. de la variable dependiente	45,1586
Suma de cuadrados de los residuos	54793,9
Desviación típica de la regresión ($\hat{\sigma}$)	37,9730

• **Resultado 6:** Estimación de White de $V(\hat{\beta}_{MCO})$:

Estimaciones MCO utilizando las 40 observaciones 1–40
Variable dependiente: consumo
Desviaciones típicas robustas ante heterocedasticidad, variante HC3

	Coeficiente	Desv. típica	estadístico t	valor p
const	40,7676	27,3213	1,4922	0,1439
renta	0,1282	0,0439	2,9168	0,0059

Media de la var. dependiente	130,3130
D.T. de la variable dependiente	45,1586
Suma de cuadrados de los residuos	54311,3
Desviación típica de la regresión ($\hat{\sigma}$)	37,8054
R^2	0,3171
$\bar{R}^2$ corregido	0,2991
Grados de libertad	38

3.9. Anexo 3.2: Instrucciones básicas de gretl para heterocedasticidad

En las clases del Centro de Cálculo o Laboratorio Informático nuestro objetivo es aprender el manejo de un software libre especialmente indicado y creado para el aprendizaje de la Econometría. Es un software muy sencillo en el que podéis ser autodidactas. Sin embargo, vamos a comenzar con un recordatorio de las nociones básicas que ya se han visto en Introducción a la Econometría. A continuación se pasará a mostrar las instrucciones específicas más usuales para el caso de heterocedasticidad.

Comienzo de la sesión. Conexión y lectura de datos:

- Encender el terminal → Hacer click cuando lo pide → Introducir login y password donde lo pide. Introducir el pendrive o diskette.
- Si pensáis guardar los resultados en un documento Word. Pulsar:

  Inicio → Todos los programas → Microsoft Office → Microsoft Office Word.

  Así estamos abriendo un documento en Word para ir guardando los resultados[11].

  Minimizaremos la ventana del documento .doc (ó .tex) si es el caso para usarlo cuando haya resultados que guardar.
- Pulsar *Inicio → Todos los programas → gretl*

  Ya estamos dentro de gretl y veremos una ventana con diferentes opciones que podemos utilizar.
- En el Centro de Cálculo de la Facultad los resultados van a una carpeta compartida que está previamente creada, pero en nuestro PC necesitaremos guardar los resultados en una carpeta donde previamente hemos abierto el documento Word, Tex, etc. Por lo tanto, lo primero que haremos será predeterminar el destino. Pulsar:
  Archivo → Preferencias → General

  En la ventana *Directorio gretl de usuario* buscaremos la situación de la carpeta citada → Pulsar *Aceptar*
- Para leer los datos de la tarea. Pulsar:
  Archivo ⟶ Abrir datos ⟶ Archivo de muestra ⟶ Nombre del "fichero de datos" por ejemplo → *Ramanathan* → *data7-24.gdt*

  Aparecerán las variables de la muestra y en la barra superior diferentes etiquetas, por ejemplo en *Datos* podremos ver las observaciones y sus características, en *Modelo* podremos realizar estimaciones.

[11]Si no dominamos Word podemos guardar los resultados en formato texto para pegar en el block de notas o "Notepad". En este caso no crearemos el documento .doc. gretl también permite guardar resultados en formato Latex previo crear el documento .tex de la misma forma que el documento Word con la opción adecuada.

Algunas etiquetas de la pantalla principal:

- La etiqueta *Datos*:
  Algunas de las opciones que contiene la etiqueta Datos son las siguientes:

  Mostrar valores
  Editar valores
  Leer información
  Ver descripción
  Estructura del conjunto de datos

  Para obtener lo que necesitamos, sólo tenemos que seleccionar la etiqueta correspondiente y la variable o variables a estudiar. Por ejemplo, para ver la estructura del conjunto de datos pulsamos en la etiqueta *Estructura del conjunto de datos* y obtendremos una pantalla en la que aparecerá seleccionado el tipo de datos con el que estamos trabajando, en este caso *Sección Cruzada* seleccionamos *Adelante* y nos confirma que la muestra es una sección cruzada junto con su tamaño, *1 a 224 observaciones*.

  Si la muestra fuese de serie temporal hubiera indicado *Serie temporal*, seleccionaríamos aceptar y veríamos la frecuencia, *mensual*, y el inicio y final de la muestra *1968:1 a 1998:12*, por ejemplo. La etiqueta estructura del conjunto de datos es muy útil cuando necesitamos cambiar alguno de ellos por ejemplo si añadimos nuevas observaciones.

  La misma información contenida en la estructura del conjunto de datos podemos encontrarla en la etiqueta: *Ver descripción*, que describe el conjunto de datos junto con cada una de las variables que lo componen.

- La etiqueta *Ver:*
  Se obtienen gráficos de las variables y sus estadísticos principales entre otros. Para obtener los estadísticos principales de las variables de la muestra podemos hacerlo pulsando en:

  Ver $\rightarrow$ *Estadísticos principales*

  La ventana de output mostrará la media, moda, valor máximo y mínimo de la serie, desviación típica, coeficiente de variación, curtosis y asimetría, para una única serie o para el conjunto de ellas seleccionándolo previamente.

- La etiqueta *Variable*:
  Sirve para trabajar con una única serie de la muestra. Algunas de las opciones que incluye esta etiqueta son:

  Buscar
  Mostrar valores
  Estadísticos principales
  Distribución de frecuencias
  Gráfico de frecuencias (simple, contra la normal, contra la gamma)
  Gráfico de series temporales
  Editar atributos
  etc

- La etiqueta *Añadir*:
  Con esta etiqueta podemos añadir variables o transformaciones de las existentes al conjunto de datos original, para ello tras pulsar en *Añadir* → tenemos las siguientes posibilidades:

  - Logaritmos de las variables seleccionadas
  - Cuadrados de las variables seleccionadas
  - Retardos de las variables seleccionadas
  - Primeras diferencias de las variables seleccionadas
  - Diferencias del logaritmo las variables seleccionadas
  - Diferencias estacionales de las variables seleccionadas
  - Variable índice:
    $$\text{index } i = 1, \ldots, N, \;\; \text{index} \equiv i$$
    $$\text{time } t = 1, \ldots, T, \;\; \text{time} \equiv t$$
  - Tendencia temporal
  - Variable aleatoria (uniforme, normal, chi cuadrado y t-Student) Por ejemplo para crear una variable normal de media 0 y desviación 1 haremos *nombre de la variable 0 1*
  - Variables ficticias, etc.
  - Definir una nueva variable. Esta opción podemos utilizarla para crear combinaciones de variables por ejemplo $Z_t = 4 + \epsilon_t \qquad \epsilon_t \sim N(0, 1)$. Permite los operadores,

    `+,  -,  *,  /, ^`

    (suma, resta, producto, potencia) entre otros.
  - Para crear un conjunto de datos:

Datos no incluidos en gretl:
En ocasiones debemos trabajar con un fichero de datos que no está incluido en gretl. Podemos importarlo y trabajar con él con solo que esté en alguno de los formatos compatibles con gretl. Para ello pulsamos en *Archivo* → *Abrir datos* → *importar* y seleccionamos el formato adecuado siguiendo la secuencia de órdenes que se nos pida. Esta secuencia va dirigida a definir la muestra, tipo de datos, longitud, etc.

En otras ocasiones debemos crear directamente la muestra en gretl. Para ello debemos crear un nuevo conjunto de datos como sigue:

Archivo → *Nuevo conjunto de datos*

y completar la información que pide sobre:
número de observaciones
estructura del conjunto de datos (serie temporal o sección cruzada)
frecuencia
observación inicial

A la pregunta *¿Desea empezar a introducir los valores de los datos usando la hoja de cálculo de gretl?* contestar *Sí*

- Introducir el nombre de la variable. El máximo de caracteres que acepta es 15, no usar acentos ni la letra *ñ*. Pulsar *Aceptar*
- En la hoja de cálculo situarnos en la primera celda y teclear la observación correspondiente, a continuación pulsar *intro*. Si nos saltamos alguna observación podemos insertar una fila en el lugar correspondiente con solo situarnos en la celda posterior e ir a *observación → insertar obs.* Una vez introducidas todas las variables pulsar *Aplicar*.
- Para guardar los datos: en menú *Archivo → Guardar datos*. Dar nombre al conjunto de datos, por ejemplo *Azar* y se grabará automáticamente con la extensión *gdt*.

Si en otro momento queremos usar este conjunto de datos sólo habrá que pulsar el botón izquierdo del ratón dos veces para que se active.

Un repaso a lo más básico:

• **Estimación MCO:** *Modelo ⟶ Mínimos Cuadrados Ordinarios*

Seleccionar la variable endógena y exógenas mediante el siguiente proceso:

1. Variable endógena, pulsar *nombre de la variable dependiente → Elegir*
2. Elegir los regresores, pulsar *Añadir* con cada una. Por defecto tendremos predeterminada una constante que se puede eliminar si es necesario. Para realizar la regresión pulsar *Aceptar*.

Se muestran los resultados de la estimación y diferentes estadísticos. Las desviaciones típicas son calculadas con la expresión $\hat{\sigma}^2(X'X)^{-1}$.

Notar que en la ventana abierta por MCO, abajo a la izquierda aparece una casilla con la leyenda estimaciones típicas robustas. En principio no debe estar activada. Corresponden a la estimación consistente de la matriz de varianzas y covarianzas de los estimadores MCO como veremos más adelante. En la ventana de resultados de la estimación MCO tenemos diferentes opciones, podemos hacer contrastes, gráficos etc.

• **Para guardar los resultados en formato word**: *Editar → Elegir formato RTF(Ms Word).*
Abrir el documento Word creado anteriormente. Seleccionar:

Edición → Pegar → Guardar → Minimizar ventana[12]

• **Gráficos de residuos:**
Tenemos varias posibilidades para dibujar los residuos, podemos dibujar su evolución en la muestra o contra alguna de las variables exógenas de la muestra, dependiendo de lo que nos interese. Partimos de la ventana de resultados MCO:

1. Evolución de los residuos:

    a) si la muestra es de serie temporal:
    Gráficos ⟶ Gráfico de residuos ⟶ contra "el tiempo"

[12] Para guardarlo en Latex, .tex, o modo texto, .txt, proceder igual con la opción adecuada.

b) si la muestra es de sección cruzada:
Gráficos ⟶ *Gráfico de residuos* ⟶ *"por observación"*

2. Gráfico de residuos frente a alguno de los regresores:
   Gráficos ⟶ *Gráfico de residuos* ⟶ *contra "nombre del regresor"*

3. Para dibujar los residuos frente a una variable incluida en el fichero de datos, pero que no sea uno de los regresores debemos situarnos en la pantalla inicial de gretl y seguir la secuencia:
   Datos ⟶ *Gráficos* ⟶ *Gráfico X-Y scatter* ⟶ *seleccionar X e Y adecuadamente*

• **Para guardar gráficos**: pulsar con el botón derecho del ratón en cualquier parte del gráfico y elegir la opción en que queremos que nos lo guarde, por ejemplo postcript (.eps) o cualquier otra que nos convenga. En la ventana que aparece indicarle donde queremos que nos lo guarde.

• **Para ver el gráfico variable ajustada-observada:**
Gráficos ⟶ *Gráfico de ajustada-observada* ⟶ *elegir por número de observación* (si es una muestra de sección cruzada) o *frente al tiempo* (si la muestra es de serie temporal)

• **Para ver el gráfico variable ajustada-observada frente a alguno de los regresores:**
Gráficos ⟶ *Gráfico de ajustada-observada* ⟶ *contra "nombre del regresor"*

• Podemos guardar los datos de la variable endógena estimada, los residuos y los residuos al cuadrado que posiblemente, necesitemos después. En la pantalla de resultados de la estimación:
Guardar ⟶ *valores ajustados*

Guardar ⟶ *residuos*

Guardar ⟶ *residuos al cuadrado*

entre otros. gretl los va a añadir al conjunto de datos con el que trabajamos y los denota respectivamente por yhat1, uhat1 e usq1 respectivamente, donde 1 indica que corresponde al modelo 1, así que si lo buscáis para el modelo que habéis estimado en tercer lugar los llamara uhat3. Además, añade una leyenda explicativa de la variable. Como veis en la pestaña hay otros estadísticos que se pueden guardar de la misma forma por ejemplo la suma residual de cuadrados por ejemplo, ess1, etc.
También podemos hacer nuevas estimaciones o añadir variables explicativas a la anterior repitiendo los pasos anteriores.

Instrucciones de gretl específicas para heterocedasticidad

• **Contrastes de heterocedasticidad**:

gretl tiene implementados diferentes contrastes de heterocedasticidad. Por ejemplo el contraste de White o el contraste de Breusch-Pagan. Ambos se encuentran en la pantalla de resultados de la estimación MCO. Pulsar:

Contrastes $\rightarrow$ *Heterocedasticidad* $\rightarrow$ *seleccionar el contraste deseado*

Por supuesto ambos los podéis realizar siguiendo explícitamente todos los pasos aprendidos en clase. De la misma forma se puede llevar a cabo el contraste de Goldfeld y Quandt. Veamos como realizar el contraste de Breusch-Pagan.

• Contraste de **Breusch-Pagan**:
Para computar la regresión auxiliar para el contraste de Breusch-Pagan hemos de guardar en primer lugar, los cuadrados de los residuos MCO y la SCR del modelo de interés. A continuación hemos de definir la variable endógena de la regresión auxiliar, los residuos normalizados, que hemos llamado $e_i^2 = \hat{u}_{i,MCO}^2/\hat{\sigma}^2$, para ello en:

Variable $\longrightarrow$ *Definir una nueva variable*

introducir la fórmula $e2 = usq1/(ess1/N)$ con los valores que previamente habéis guardados o anotado de la suma de cuadrados residual y el tamaño muestral[13]. A continuación estimáis la regresión auxiliar por MCO de la manera habitual y anotar los estadísticos necesarios para computar el test.

• Contraste de **Goldfeld y Quandt**:
En el contraste de Goldfeld y Quandt se supone que la varianza de la perturbación es monótonamente creciente (o decreciente) con una variable cuyos valores son conocidos. Es necesario ordenar la muestra conforme al crecimiento (decrecimiento) de esa variable. Para ello podemos seguir la siguiente secuencia desde la pantalla inicial:

Datos	$\longrightarrow$	*ordenar*
	$\longrightarrow$	*seleccionar la variable por cuyo crecimiento o decrecimiento queremos ordenar la muestra*
	$\longrightarrow$	*seleccionar el criterio, creciente o decreciente*

Además, es necesario que la muestra se divida en dos submuestras dejando un número de observaciones centrales que de independencia a ambas submuestras. A continuación se realiza la regresión MCO en ambas submuestras y se aplica el estadístico de contraste. La única dificultad de este test es dividir la muestra. Para ello debemos seguir la siguiente secuencia de órdenes partiendo de la pantalla principal de gretl. Seleccionar, según el criterio adecuado:

Muestra	$\rightarrow$	*seleccionar rango*
	$\rightarrow$	*restringir a partir de un criterio*
	$\rightarrow$	*definir a partir de una variable ficticia*

con ello hemos restringido la muestra y podemos realizar la regresión en dicha submuestra. Tomar nota de los estadísticos necesarios. A continuación hay que recuperar el rango completo antes de volver a restringir la muestra para crear la segunda submuestra. Pulsar:

Muestra $\rightarrow$ *recuperar rango completo*

volver a restringir la muestra para realizar la segunda regresión y tomar nota de los datos necesarios para realizar el contraste. Antes de seguir trabajando recordar recuperar el rango completo de la muestra.

[13] Se ha denotado $\hat{u}_{i,MCO}^2$ =usq1 y $\hat{\sigma}^2$ =ess1/N.

• Los contraste de Goldfeld- Quandt y Breusch-Pagan no son los únicos posibles. De hecho gretl ofrece la posibilidad de contrastar heterocedasticidad usando el test de White y una variante del mismo como habéis podido ver al desplegar la pestaña *Contrastes* $\longrightarrow$ *Heterocedasticidad* sin embargo, en los contenidos del tema no se suele incluir este contraste ya que el tiempo disponible es limitado.

• **Estimador de White**:
El estimador de White es el estimador consistente de la matriz de varianzas y covarianzas de los parámetros estimados por MCO, $\widehat{V}(\hat{\beta}_{MCO})$. Lógicamente ha de encontrarse dentro de la estimación MCO y ya lo hemos abordado antes, estaba en la ventana abierta por MCO. Abajo a la izquierda aparece una casilla con la leyenda estimaciones típicas robustas. Corresponden a la estimación consistente de la matriz de varianzas y covarianzas de los estimadores MCO a la White. Si deseamos esta estimación haremos *click* en la casilla. Pulsar en *configurar* y elegir *HC3a* que es la correspondiente al estimador visto en clase.

• **Mínimos Cuadrados Generalizados o Ponderados (MCP)**
Bajo MCP gretl implementa el estimador de MCG vía MCO en el modelo transformado.

Lo primero que va a necesitar gretl es conocer la ponderación por ello, si la ponderación no es una variable previamente definida, debéis definirla de la forma habitual:

Variable $\longrightarrow$ *Definir una nueva variable* $\longrightarrow$ *introducir la fórmula*

Una vez que tenéis construida la ponderación para estimar por MCP seleccionamos:
Modelo $\longrightarrow$ *otros modelos lineales* $\longrightarrow$ *mínimos cuadrados ponderados*

Se eligen adecuadamente la variable dependiente, las independientes y la ponderación seleccionando y añadiendo de la forma habitual.

Con la definición de la variable de ponderación hay que tener un poco de cuidado ya que es necesario adecuarse a la forma que utiliza el programa para construir las variables en el modelo transformado. Así para diferentes versiones hace diferentes cosas y necesitamos saber a ciencia cierta cómo construye el modelo transformado a partir de la definición de la variable de ponderación. Lo sabemos en la pestaña Ayuda, dentro de Mínimos Cuadrados Ponderados dependiendo de la versión. Por ejemplo para la versión 1.7.1, incluye la siguiente leyenda en inglés[14]:

Weighted Least Squares

If “wtvar”is a dummy variable, WLS estimation is equivalent to eliminating all observations with value zero for wtvar.

Let “wtvar”denote the variable selected in the ”Weight variable”box. An OLS regression is run, where the dependent variable is the product of the positive square root of wtvar and the selected dependent variable, and the independent variables are also multiplied by the square root of wtvar. Statistics such as R-squared are based on the weighted data. If wtvar is a dummy variable, weighted least squares estimation is equivalent to eliminating all observations with value zero for wtvar.

[14]Las etiquetas de ayuda están en inglés ya que no se traducen como el resto del programa.

Donde "wtvar" denota la variable de ponderación y el modelo transformado se construye multiplicando la raíz cuadrada de "wtvar" por cada variable del modelo original incluida la constante. Por ejemplo si suponemos que $Var(u_i) = \sigma^2 X_{2i}^2$ tomaremos como ponderación a $\frac{1}{X_{2i}^2}$ y si $Var(u_i) = \sigma^2 X_{2i}$ tomaremos como ponderación a $\frac{1}{X_{2i}}$.

- **Mínimos Cuadrados Generalizados Factibles**

1. Primero hemos de pensar en la forma funcional de la varianza. Supongamos que la forma funcional de la varianza que vamos a proponer es:

   $$Var(u_i) = \sigma_i^2 = \alpha_1 + \alpha_2 W_{1i} + \alpha_3 W_{2i}$$

   A continuación hemos de crear la regresión auxiliar y estimarla por MCO. Si la variable endógena son los residuos MCO al cuadrado del modelo de interés y los hemos guardado podemos utilizarlos directamente y si no lo hemos hecho elevaremos al cuadrado los residuos MCO guardados en la estimación MCO original. Estimamos la regresión auxiliar y guardamos las estimaciones de la variable endógena de esta regresión. Con esta serie construiremos la ponderación, es decir, vamos a obtener $\hat{\sigma}_i^2$ por lo que debemos comprobar que es siempre positiva.

2. Construimos la ponderación $1/\hat{\sigma}_i^2$ y la añadimos al conjunto de datos usando como habitualmente

   Variable $\longrightarrow$ *Definir una nueva variable* $\longrightarrow$ *introducir la fórmula*

3. A continuación seleccionamos:

   Modelo $\longrightarrow$ *Mínimos Cuadrados Ponderados*

   y la variable dependiente, las independientes y la ponderación seleccionando y añadiendo de la forma habitual. Este modo de proceder proporciona el estimador de MCGF.

3.10. Anexo 3.3: Tablas de datos

- Datos sobre Consumo y Renta. Ejemplo magistral.

Tabla 3.1: Observaciones de Consumo y Renta

Observación	Consumo	Renta	Observación	Consumo	Renta
i= 1	52,25	258,3	i= 21	98,14	719,8
i= 2	58,32	343,1	i= 22	123,94	720,0
i= 3	81,79	425,0	i= 23	126,31	722,3
i= 4	119,90	467,5	i= 24	146,47	722,3
i= 5	125,80	482,9	i= 25	115,98	734,4
i= 6	100,46	487,7	i= 26	207,23	742,5
i= 7	121,51	496,5	i= 27	119,80	747,7
i= 8	100,08	519,4	i= 28	151,33	763,3
i= 9	127,75	543,3	i= 29	169,51	810,2
i= 10	104,94	548,7	i= 30	108,03	818,5
i= 11	107,48	564,6	i= 31	168,90	825,6
i= 12	98,48	588,3	i= 32	227,11	833,3
i= 13	181,21	591,3	i= 33	84,94	834,0
i= 14	122,23	607,3	i= 34	98,70	918,1
i= 15	129,57	611,2	i= 35	141,06	918,1
i= 16	92,84	631	i= 36	215,40	929,6
i= 17	117,92	659,6	i= 37	112,89	951,7
i= 18	82,13	664,0	i= 38	166,25	1014,0
i= 19	182,28	704,2	i= 39	115,43	1141,3
i= 20	139,13	704,8	i= 40	269,03	1154,6