Econometría Básica

Capítulo 08: Heterocedasticidad

José Valderrama & Freddy Rojas
jtvalderrama@gmail.com & frojasca@gmail.com
Universidad Católica Santo Toribio de Mogrovejo

Septiembre de 2021

Contenido

Contenido

¿Qué supuesto no se cumple?

Tenemos el siguiente modelo (o PGD) con una constante y más de un regresos. En el caso de 2 regresares

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + u_i$$

Donde el término error tiene las siguientes propiedades

$$E(u_i|x) = 0$$
$$E(u_i^2|x) = \sigma_i^2$$
$$E(u_i u_j|x) = 0, \quad i \neq j$$

¿Qué supuesto no se cumple?

Debes notar que el estimador es aún insesgado pero no eficiente. Aquí surgen algunas preguntas

- ¿Qué significa eso? Respuesta: Aquí existe otro estimador insesgado, el cual tiene la más baja varianza.
- ¿Por qué es importante? ...

¿Qué es la heterocedasticidad?

- Recordando, el supuesto de homocedasticidad consiste en que condicionado a las variables explicativas, la varianza del error poblacional es constante: $var(u/x) = \sigma^2$
- Si ello no es cierto, esto es si la varianza de u es diferente para valores diferentes de las $x's$, entonces los errores son heterocedásticos
- Un ejemplo típico de lo anterior es lo que ocurre entre los ingresos y los años de educación. Si bien aumentos de la educación aumentan los ingresos, la dispersión de los no observables (u) aumenta con los años de educación (es decir, es posible encontrar gente muy educada ganando muy bien pero también otro grupo ganando muy poco)

USAT Universidad Católica Santo Toribio de Mogrovejo

¿Qué es la heterocedasticidad?

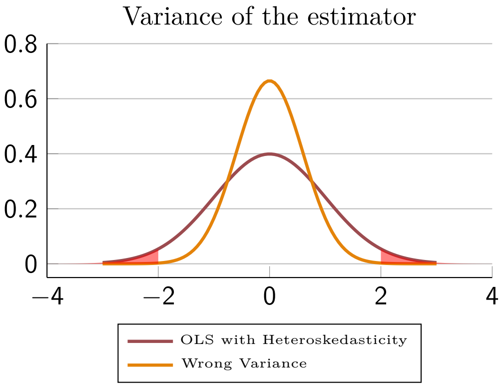
- Recordando, el supuesto de homocedasticidad consiste en que condicionado a las variables explicativas, la varianza del error poblacional es constante: $var(u/x) = \sigma^2$
- Si ello no es cierto, esto es si la varianza de u es diferente para valores diferentes de las $x's$, entonces los errores son heterocedásticos
- Un ejemplo típico de lo anterior es lo que ocurre entre los ingresos y los años de educación. Si bien aumentos de la educación aumentan los ingresos, la dispersión de los no observables (u) aumenta con los años de educación (es decir, es posible encontrar gente muy educada ganando muy bien pero también otro grupo ganando muy poco)

¿Qué es la heterocedasticidad?

- Recordando, el supuesto de homocedasticidad consiste en que condicionado a las variables explicativas, la varianza del error poblacional es constante: $var(u/x) = \sigma^2$
- Si ello no es cierto, esto es si la varianza de u es diferente para valores diferentes de las $x's$, entonces los errores son heterocedásticos
- Un ejemplo típico de lo anterior es lo que ocurre entre los ingresos y los años de educación. Si bien aumentos de la educación aumentan los ingresos, la dispersión de los no observables (u) aumenta con los años de educación (es decir, es posible encontrar gente muy educada ganando muy bien pero también otro grupo ganando muy poco)

Contenido

¿Porqué preocuparnos de la heterocedasticidad?



Si usamos la varianza de homocedasticidad del estimador para hacer inferencia, podemos rechazar la hipótesis nula cuando esa de hecho es verdad (Error Tipo 1).

¿Porqué preocuparnos de la heterocedasticidad?

- Aun con heterocedasticidad los estimadores MCO son insesgados y consistentes.
- Sin embargo, los errores estándar de los estimadores son sesgados
- Al ser sesgados los errores estándar entonces no se pueden usar los estadísticos usuales t y F para hacer inferencia

¿Porqué preocuparnos de la heterocedasticidad?

- Aun con heterocedasticidad los estimadores MCO son insesgados y consistentes.
- Sin embargo, los errores estándar de los estimadores son sesgados
- Al ser sesgados los errores estándar entonces no se pueden usar los estadísticos usuales t y F para hacer inferencia

¿Porqué preocuparnos de la heterocedasticidad?

- Aun con heterocedasticidad los estimadores MCO son insesgados y consistentes.
- Sin embargo, los errores estándar de los estimadores son sesgados
- Al ser sesgados los errores estándar entonces no se pueden usar los estadísticos usuales t y F para hacer inferencia

¿Porqué preocuparnos de la heterocedasticidad?

Entonces tenemos que encontrar un estimador MELI¡.
De hecho, podemos mantener nuestras estimaciones de MCO, debes saber que:

1. MCO es aún insesgado.
2. El problema es que podemos hacer una inferencia incorrecta (surge error tipo 1), solo lo que tenemos que hacer es calcular la verdadera varianza (asumiendo heterocedasticidad) y simplemente hacer la inferencia.
3. Si nos interesa la predicción o el cálculo de intervalos de confianza que necesitamos para encontrar un estimador MELI. Recuerda esto:

$$Var(\beta^{MELI}) < Var(\beta^{het})$$

Un estimador MELI es el estimador de "Mínimos Cuadrados Generalizados" (MCG).

Contenido

Varianza con heterocedasticidad

- Para ilustrar la idea usemos el modelo de regresión lineal simple en el cual se cumple: $\hat{\beta}_1 = \beta_1 + \frac{\sum(x_i - \overline{x})u_i}{\sum(x_i - \overline{x})^2}$ del cual se deriva:
- $Var(\hat{\beta}_1) = \frac{\sum(x_i - \overline{x})^2 \sigma_i^2}{SCT_x^2}$, donde $SCT_x = \sum(x_i - \overline{x})^2$
- White (1980) demostró empleando teoría asintótica que una estimación válida de lo anterior cuando $\sigma_i^2 \neq \sigma^2$ es $\frac{\sum(x_i - \overline{x})^2 \hat{u}_i^2}{SCT_x^2}$, donde $\hat{u}_i$ son los residuos obtenidos de la regresión.
- La raíz cuadrada de esta última expresión se conoce como error estándar robusto a la heterocedasticidad de β_1
- Finalmente: $t = \frac{\hat{\beta}_1}{\sqrt{Var(\hat{\beta}_1)}}$
- Así como se puede obtener un nuevo 't' bajo la misma filosofía es posible derivar un estadístico F robusto a la heterocedasticidad.

Varianza con heterocedasticidad

- Para ilustrar la idea usemos el modelo de regresión lineal simple en el cual se cumple: $\hat{\beta}_1 = \beta_1 + \frac{\sum(x_i - \overline{x})u_i}{\sum(x_i - \overline{x})^2}$ del cual se deriva:
- $Var(\hat{\beta}_1) = \frac{\sum(x_i - \overline{x})^2 \sigma_i^2}{SCT_x^2}$, donde $SCT_x = \sum(x_i - \overline{x})^2$
- White (1980) demostró empleando teoría asintótica que una estimación válida de lo anterior cuando $\sigma_i^2 \neq \sigma^2$ es $\frac{\sum(x_i - \overline{x})^2 \hat{u}_i^2}{SCT_x^2}$, donde $\hat{u}_i$ son los residuos obtenidos de la regresión.
- La raíz cuadrada de esta última expresión se conoce como error estándar robusto a la heterocedasticidad de β_1
- Finalmente: $t = \frac{\hat{\beta}_1}{\sqrt{Var(\hat{\beta}_1)}}$
- Así como se puede obtener un nuevo 't' bajo la misma filosofía es posible derivar un estadístico F robusto a la heterocedasticidad.

Varianza con heterocedasticidad

- Para ilustrar la idea usemos el modelo de regresión lineal simple en el cual se cumple: $\hat{\beta}_1 = \beta_1 + \frac{\sum(x_i - \overline{x})u_i}{\sum(x_i - \overline{x})^2}$ del cual se deriva:
- $Var(\hat{\beta}_1) = \frac{\sum(x_i - \overline{x})^2 \sigma_i^2}{SCT_x^2}$, donde $SCT_x = \sum(x_i - \overline{x})^2$
- White (1980) demostró empleando teoría asintótica que una estimación válida de lo anterior cuando $\sigma_i^2 \neq \sigma^2$ es $\frac{\sum(x_i - \overline{x})^2 \hat{u}_i^2}{SCT_x^2}$, donde $\hat{u}_i$ son los residuos obtenidos de la regresión.
- La raíz cuadrada de esta última expresión se conoce como error estándar robusto a la heterocedasticidad de β_1
- Finalmente: $t = \frac{\hat{\beta}_1}{\sqrt{Var(\hat{\beta}_1)}}$
- Así como se puede obtener un nuevo 't' bajo la misma filosofía es posible derivar un estadístico F robusto a la heterocedasticidad.

USAT Universidad Católica Santo Toribio de Mogrovejo

Varianza con heterocedasticidad

- Para ilustrar la idea usemos el modelo de regresión lineal simple en el cual se cumple: $\hat{\beta}_1 = \beta_1 + \frac{\sum(x_i - \overline{x})u_i}{\sum(x_i - \overline{x})^2}$ del cual se deriva:
- $Var(\hat{\beta}_1) = \frac{\sum(x_i - \overline{x})^2 \sigma_i^2}{SCT_x^2}$, donde $SCT_x = \sum(x_i - \overline{x})^2$
- White (1980) demostró empleando teoría asintótica que una estimación válida de lo anterior cuando $\sigma_i^2 \neq \sigma^2$ es $\frac{\sum(x_i - \overline{x})^2 \hat{u}_i^2}{SCT_x^2}$, donde $\hat{u}_i$ son los residuos obtenidos de la regresión.
- La raíz cuadrada de esta última expresión se conoce como error estándar robusto a la heterocedasticidad de β_1
- Finalmente: $t = \frac{\hat{\beta}_1}{\sqrt{Var(\hat{\beta}_1)}}$
- Así como se puede obtener un nuevo 't' bajo la misma filosofía es posible derivar un estadístico F robusto a la heterocedasticidad.

Varianza con heterocedasticidad

- Para ilustrar la idea usemos el modelo de regresión lineal simple en el cual se cumple: $\hat{\beta}_1 = \beta_1 + \frac{\sum(x_i - \overline{x})u_i}{\sum(x_i - \overline{x})^2}$ del cual se deriva:
- $Var(\hat{\beta}_1) = \frac{\sum(x_i - \overline{x})^2\sigma_i^2}{SCT_x^2}$, donde $SCT_x = \sum(x_i - \overline{x})^2$
- White (1980) demostró empleando teoría asintótica que una estimación válida de lo anterior cuando $\sigma_i^2 \neq \sigma^2$ es $\frac{\sum(x_i - \overline{x})^2\hat{u}_i^2}{SCT_x^2}$, donde $\hat{u}_i$ son los residuos obtenidos de la regresión.
- La raíz cuadrada de esta última expresión se conoce como error estándar robusto a la heterocedasticidad de β_1
- Finalmente: $t = \frac{\hat{\beta}_1}{\sqrt{Var(\hat{\beta}_1)}}$
- Así como se puede obtener un nuevo 't' bajo la misma filosofía es posible derivar un estadístico F robusto a la heterocedasticidad.

Varianza con heterocedasticidad

- Para ilustrar la idea usemos el modelo de regresión lineal simple en el cual se cumple: $\hat{\beta}_1 = \beta_1 + \frac{\sum(x_i - \overline{x})u_i}{\sum(x_i - \overline{x})^2}$ del cual se deriva:
- $Var(\hat{\beta}_1) = \frac{\sum(x_i - \overline{x})^2\sigma_i^2}{SCT_x^2}$, donde $SCT_x = \sum(x_i - \overline{x})^2$
- White (1980) demostró empleando teoría asintótica que una estimación válida de lo anterior cuando $\sigma_i^2 \neq \sigma^2$ es $\frac{\sum(x_i - \overline{x})^2\hat{u}_i^2}{SCT_x^2}$, donde $\hat{u}_i$ son los residuos obtenidos de la regresión.
- La raíz cuadrada de esta última expresión se conoce como error estándar robusto a la heterocedasticidad de β_1
- Finalmente: $t = \frac{\hat{\beta}_1}{\sqrt{Var(\hat{\beta}_1)}}$
- Así como se puede obtener un nuevo 't' bajo la misma filosofía es posible derivar un estadístico F robusto a la heterocedasticidad.

USAT Universidad Católica Santo Toribio de Mogrovejo

T y F robustos a la heterocedasticidad

Stata code

```
reg price rep78 trunk
```

```
Source       |       SS           df       MS      Number of obs   =        69
-------------+----------------------------------   F(2, 66)        =      3.99
Model        |  62200533.8         2  31100266.9   Prob > F        =    0.0232
Residual     |   514596425        66  7796915.53   R-squared       =    0.1078
-------------+----------------------------------   Adj R-squared   =    0.0808
Total        |   576796959        68  8482308.22   Root MSE        =    2792.3

------------------------------------------------------------------------------
price |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
rep78 |   173.0812   346.3682     0.50   0.619    -518.4653    864.6276
trunk |   222.9437   78.94888     2.82   0.006     65.31713    380.5702
_cons |   2451.508   1766.794     1.39   0.170    -1076.009    5979.026
------------------------------------------------------------------------------
```

T y F robustos a la heterocedasticidad

Stata code

```
reg price rep78 trunk, robust
```

```
Linear regression                               Number of obs     =         69
                                                F(2, 66)          =       7.91
                                                Prob > F          =     0.0008
                                                R-squared         =     0.1078
                                                Root MSE          =     2792.3

------------------------------------------------------------------------------
      |               Robust
price |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
rep78 |   173.0812   290.9098     0.59   0.554     -407.739    753.9013
trunk |   222.9437   57.38124     3.89   0.000     108.3783     337.509
_cons |   2451.508   1268.618     1.93   0.058     -81.3697    4984.386
------------------------------------------------------------------------------
```

Contenido

Probando heterocedasticidad

Básicamente, existen 2 pruebas para comprobar la presencia de heterocedasticidad: La prueba de White y la prueba de Breush-Pagan

- La prueba de White considera los efectos lineales y no lineales como fuente de heterocedasticidad.
- El de Breusch Pagan solo considera efectos lineales e incluye otras variables exógenas.

¿Cuál es la hipóteis nula para esas pruebas? *Respuesta*:

$$H_0 : \text{Homocedasticidad}$$
$$H_a : \text{Heterocedasticidad}$$

Probando heterocedasticidad

- Esencialmente lo que se quiere probar es
$H_o : Var(u/x_1, x_2, ..., x_k) = \sigma^2$
- En virtud que se asume $E(u/x) = 0$, $Var(u/x) = E(u^2/x)$, así la hipótesis nula de homocedasticidad es equivalente a:
$H_o : E(u^2/x_1, x_2, ..., x_k) = E(u^2) = \sigma^2$
- Para probar la violación del supuesto de homocedasticidad se busca probar que $E(u^2)$ se relaciona con una o más variables explicativas.
- Así, para $u^2 = \delta_0 + \delta_1 x_1 + ... + \delta_k x_k + v$ ello implica probar $H_o : \delta_1 = ... = \delta_k = 0$

USAT Universidad Católica Santo Toribio de Mogrovejo

Probando heterocedasticidad

- Esencialmente lo que se quiere probar es $H_o : Var(u/x_1, x_2, ..., x_k) = \sigma^2$
- En virtud que se asume $E(u/x) = 0$, $Var(u/x) = E(u^2/x)$, así la hipótesis nula de homocedasticidad es equivalente a: $H_o : E(u^2/x_1, x_2, ..., x_k) = E(u^2) = \sigma^2$
- Para probar la violación del supuesto de homocedasticidad se busca probar que $E(u^2)$ se relaciona con una o más variables explicativas.
- Así, para $u^2 = \delta_0 + \delta_1 x_1 + ... + \delta_k x_k + v$ ello implica probar $H_o : \delta_1 = ... = \delta_k = 0$

Probando heterocedasticidad

- Esencialmente lo que se quiere probar es
  $H_o : Var(u/x_1, x_2, ..., x_k) = \sigma^2$
- En virtud que se asume $E(u/x) = 0$, $Var(u/x) = E(u^2/x)$, así la hipótesis nula de homocedasticidad es equivalente a:
  $H_o : E(u^2/x_1, x_2, ..., x_k) = E(u^2) = \sigma^2$
- Para probar la violación del supuesto de homocedasticidad se busca probar que $E(u^2)$ se relaciona con una o más variables explicativas.
- Así, para $u^2 = \delta_0 + \delta_1 x_1 + ... + \delta_k x_k + v$ ello implica probar H_o : $\delta_1 = ... = \delta_k = 0$

Probando heterocedasticidad

- Esencialmente lo que se quiere probar es $H_o : Var(u/x_1, x_2, ..., x_k) = \sigma^2$
- En virtud que se asume $E(u/x) = 0$, $Var(u/x) = E(u^2/x)$, así la hipótesis nula de homocedasticidad es equivalente a: $H_o : E(u^2/x_1, x_2, ..., x_k) = E(u^2) = \sigma^2$
- Para probar la violación del supuesto de homocedasticidad se busca probar que $E(u^2)$ se relaciona con una o más variables explicativas.
- Así, para $u^2 = \delta_0 + \delta_1 x_1 + ... + \delta_k x_k + v$ ello implica probar $H_o : \delta_1 = ... = \delta_k = 0$

Contenido

Diagnóstico: Test de Breusch-Pagan (1979)

- No se observan los $u's$ pero pueden ser estimados a partir de la regresión: $\hat{u}^2$
- Después de regresionar todos los residuos al cuadrado sobre todos los $x's$ se puede usar el estadístico $LM = nR^2$ que tiene una distribución χ_k. En STATA sería:
  1. `reg y x1 x2`
  2. `predict error, resid`
  3. `g error2=error*error`
  4. `reg error2 x1 x2`
  5. display 1-chi2(k,b) -(k es el número de variables y b es igual N°Obs*R2)-
- Si el valor p es lo suficientemente pequeño, es decir, si se halla por debajo del nivel de significancia elegido, entonces rechazamos la hipótesis nula de homocedasticidad.

Diagnóstico: Test de Breusch-Pagan (1979)

- No se observan los $u's$ pero pueden ser estimados a partir de la regresión: $\hat{u}^2$
- Después de regresionar todos los residuos al cuadrado sobre todos los $x's$ se puede usar el estadístico $LM = nR^2$ que tiene una distribución χ_k. En STATA sería:
  1. `reg y x1 x2`
  2. `predict error, resid`
  3. `g error2=error*error`
  4. `reg error2 x1 x2`
  5. display 1-chi2(k,b) -(k es el número de variables y b es igual N°Obs*R2)-
- Si el valor p es lo suficientemente pequeño, es decir, si se halla por debajo del nivel de significancia elegido, entonces rechazamos la hipótesis nula de homocedasticidad.

Diagnóstico: Test de Breusch-Pagan (1979)

- No se observan los $u's$ pero pueden ser estimados a partir de la regresión: $\hat{u}^2$
- Después de regresionar todos los residuos al cuadrado sobre todos los $x's$ se puede usar el estadístico $LM = nR^2$ que tiene una distribución χ_k. En STATA sería:
  1. `reg y x1 x2`
  2. `predict error, resid`
  3. `g error2=error*error`
  4. `reg error2 x1 x2`
  5. display 1-chi2(k,b) -(k es el número de variables y b es igual N°Obs*R2)-
- Si el valor p es lo suficientemente pequeño, es decir, si se halla por debajo del nivel de significancia elegido, entonces rechazamos la hipótesis nula de homocedasticidad.

Diagnóstico: Test de Breusch-Pagan (1979)

- No se observan los $u's$ pero pueden ser estimados a partir de la regresión: $\hat{u}^2$
- Después de regresionar todos los residuos al cuadrado sobre todos los $x's$ se puede usar el estadístico $LM = nR^2$ que tiene una distribución χ_k. En STATA sería:
  1. `reg y x1 x2`
  2. `predict error, resid`
  3. `g error2=error*error`
  4. `reg error2 x1 x2`
  5. display 1-chi2(k,b) -(k es el número de variables y b es igual N°Obs*R2)-

- Si el valor p es lo suficientemente pequeño, es decir, si se halla por debajo del nivel de significancia elegido, entonces rechazamos la hipótesis nula de homocedasticidad.

Resumen: Test de Breusch-Pagan

Pasos:

1. Conseguir los residuos ($\hat{u}$) y elevarlos al cuadrado ($\hat{u}^2$).
2. Ejecutar una regresión de $\hat{u}^2$ en variables sospechosas (puede incluir regresores u otras variables no incluidas en el regresión principal).
$$\hat{u}^2 = \delta_0 + \delta X_i + \gamma Z_i + \epsilon_i$$
3. Realizar una prueba basada en la siguiente hipótesis:
$$H_0 : \delta = 0; \;\; \gamma = 0$$
$$H_a : \text{Al menos uno diferente de cero}$$

Para probar la hipótesis nula, debe calcular el estadístico LM y contraste con valores críticos de χ^2_{k-1} donde k es el número de parámetros a estimar en (2).

Resumen: Test de Breusch-Pagan

El estadístico LM se calcula de la siguiente manera:

$$LM = \frac{SCE_{aux}}{2.\tilde{\sigma}^4}$$

SCE_{aux} es la suma cuadrada de las explicadas de la regresión auxiliar $\tilde{\sigma}^4$ es igual a $\frac{\sum \hat{u}^2}{n}$. Este estadístico es como una distribución χ^2_{k-1} donde k es el número de parámetros a estimar en la regresión auxiliar

Contenido

USAT Universidad Católica Santo Toribio de Mogrovejo

Diagnóstico: Prueba de H. White (1980)

- Es una generalización del test de Breusch-Pagan y básicamente consiste en considerar no linealiedades en la regresión de los errores al cuadrado.
- Es decir, si antes se tenía: $u^2 = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + v$ ahora se considerará la siguiente estructura $u^2 = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + \delta_3 x_1^2 + \delta_4 x_2^2 + \delta_5 x_1 x_2 + v$
- En stata sería:
  1. `reg y x1 x2`
  2. `predict error,resid`
  3. `g error2=error*error`
  4. `g tem1=x1*x1`
  5. `g tem2=x2*x2`
  6. `g tem3=x1*x2`
  7. `reg error2 x1 x2 tem1 tem2 tem3`
  8. display 1-chi2(k,b) -(k en este caso es igual a 5 y b es igual N°Obs*R2)-

Diagnóstico: Prueba de H. White (1980)

- Es una generalización del test de Breusch-Pagan y básicamente consiste en considerar no linealiedades en la regresión de los errores al cuadrado.
- Es decir, si antes se tenía: $u^2 = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + v$ ahora se considerará la siguiente estructura $u^2 = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + \delta_3 x_1^2 + \delta_4 x_2^2 + \delta_5 x_1 x_2 + v$
- En stata sería:
  1. `reg y x1 x2`
  2. `predict error,resid`
  3. `g error2=error*error`
  4. `g tem1=x1*x1`
  5. `g tem2=x2*x2`
  6. `g tem3=x1*x2`
  7. `reg error2 x1 x2 tem1 tem2 tem3`
  8. display 1-chi2(k,b) -(k en este caso es igual a 5 y b es igual N°Obs*R2)-

Diagnóstico: Prueba de H. White (1980)

- Es una generalización del test de Breusch-Pagan y básicamente consiste en considerar no linealiedades en la regresión de los errores al cuadrado.
- Es decir, si antes se tenía: $u^2 = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + v$ ahora se considerará la siguiente estructura $u^2 = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + \delta_3 x_1^2 + \delta_4 x_2^2 + \delta_5 x_1 x_2 + v$
- En stata sería:
  1. `reg y x1 x2`
  2. `predict error,resid`
  3. `g error2=error*error`
  4. `g tem1=x1*x1`
  5. `g tem2=x2*x2`
  6. `g tem3=x1*x2`
  7. `reg error2 x1 x2 tem1 tem2 tem3`
  8. display 1-chi2(k,b) -(k en este caso es igual a 5 y b es igual N°Obs*R2)-

Resumen: Test de White

Pasos:

1. Conseguir los residuales ($\hat{u}$) y elevarlos al cuadrado ($\hat{u}^2$).
2. Ejecutar una regresión de $\hat{u}^2$ en todos los regresores y en sus productos cuadrados y cruzados (esa regresión es llamada la *regresión auxiliar*. En el caso de 2 regresores;

$$\hat{u}^2 = \delta_0 + \delta_1 X_{1i} + \delta_2 X_{2i} + \delta_3 X_{1i} X_{2i} + \delta_4 X_{1\ i}^2 + \delta_5 X_{2\ i}^2 + \epsilon_i$$

3. Realizar una prueba basada en la siguiente hipótesis

$$H_0 : \delta_1 = 0 : \delta_2 = 0; \ldots ; \delta_{\frac{M(M+3)}{2}} = 0$$

$$H_a : \text{Al menos uno diferente de cero}$$

Para probar la hipótesis nula, se debe calcular el R^2 y contrastar el NR^2 con los valores críticos de una distribución χ^2_{k-1} donde k es el número de parámetros a estimar en el paso 2

Contenido

Forma alternativa del Test de White

- El estadístico tradicional de White tiene el problema de consumir muchos grados de libertad (con dos regresores, la regresión relevante tendría 5 explicativas; con 6 tendría 27!)
- Una sugerencia es usar los valores predichos de $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + ... + \hat{\beta}_k x_k$ ya que si elevamos al cuadrado se obtiene una función que depende de todos los cuadrados y productos cruzados de las variables independientes.
- Es decir, se podría probar heterocedasticidad estimando la ecuación: $\hat{u}^2 = \delta_0 + \delta_1 \hat{y} + \delta_2 \hat{y}^2 + error$. En stata sería:
  1. `reg y x1 x2`
  2. `predict yhat,xb`
  3. `g yhat2=yhat*yhat`
  4. `reg error2 yhat yhat2`
  5. display 1-chi2(2,b) -(nótese que en este caso k siempre es igual a 2 y b es igual N°Obs*R2)-

Forma alternativa del Test de White

- El estadístico tradicional de White tiene el problema de consumir muchos grados de libertad (con dos regresores, la regresión relevante tendría 5 explicativas; con 6 tendría 27!)
- Una sugerencia es usar los valores predichos de $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \ldots + \hat{\beta}_k x_k$ ya que si elevamos al cuadrado se obtiene una función que depende de todos los cuadrados y productos cruzados de las variables independientes.
- Es decir, se podría probar heterocedasticidad estimando la ecuación: $\hat{u}^2 = \delta_0 + \delta_1 \hat{y} + \delta_2 \hat{y}^2 + error$. En stata sería:
  1. `reg y x1 x2`
  2. `predict yhat,xb`
  3. `g yhat2=yhat*yhat`
  4. `reg error2 yhat yhat2`
  5. display 1-chi2(2,b) -(nótese que en este caso k siempre es igual a 2 y b es igual N°Obs*R2)-

Forma alternativa del Test de White

- El estadístico tradicional de White tiene el problema de consumir muchos grados de libertad (con dos regresores, la regresión relevante tendría 5 explicativas; con 6 tendría 27!)
- Una sugerencia es usar los valores predichos de $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + ... + \hat{\beta}_k x_k$ ya que si elevamos al cuadrado se obtiene una función que depende de todos los cuadrados y productos cruzados de las variables independientes.
- Es decir, se podría probar heterocedasticidad estimando la ecuación: $\hat{u}^2 = \delta_0 + \delta_1 \hat{y} + \delta_2 \hat{y}^2 + error$. En stata sería:
  1. `reg y x1 x2`
  2. `predict yhat,xb`
  3. `g yhat2=yhat*yhat`
  4. `reg error2 yhat yhat2`
  5. display 1-chi2(2,b) -(nótese que en este caso k siempre es igual a 2 y b es igual N°Obs*R2)-

Forma alternativa del Test de White

- El estadístico tradicional de White tiene el problema de consumir muchos grados de libertad (con dos regresores, la regresión relevante tendría 5 explicativas; con 6 tendría 27!)
- Una sugerencia es usar los valores predichos de $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \ldots + \hat{\beta}_k x_k$ ya que si elevamos al cuadrado se obtiene una función que depende de todos los cuadrados y productos cruzados de las variables independientes.
- Es decir, se podría probar heterocedasticidad estimando la ecuación: $\hat{u}^2 = \delta_0 + \delta_1 \hat{y} + \delta_2 \hat{y}^2 + error$. En stata sería:
  1. `reg y x1 x2`
  2. `predict yhat,xb`
  3. `g yhat2=yhat*yhat`
  4. `reg error2 yhat yhat2`
  5. display 1-chi2(2,b) -(nótese que en este caso k siempre es igual a 2 y b es igual N°Obs*R2)-

Contenido

STATA

Consideremos los siguientes comandos para realizar pruebas de heterocedasticidad

```
use "http://fmwww.bc.edu/ec-p/data/wooldridge/HPRICE1.dta", clear
regress price lotsize sqrft bdrms
//White test
estat imtest, white
//Breusch-Pagan 1
estat hettest, rhs
//Breusch-Pagan 2
estat hettest
```

STATA: Ejemplo 1

```
. *WHITE
. estat imtest, white

White's test for Ho: homoskedasticity
         against Ha: unrestricted heteroskedasticity

         chi2(9)      =     33.73
         Prob > chi2  =    0.0001


. *Breusch-Pagan 1
. estat hettest, rhs

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
         Ho: Constant variance
         Variables: lotsize sqrft bdrms

         chi2(3)      =    30.02
         Prob > chi2  =   0.0000

.
. *Breusch-Pagan 2
. estat hettest

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
         Ho: Constant variance
         Variables: fitted values of price

         chi2(1)      =    20.55
         Prob > chi2  =   0.0000
```

STATA: Ejemplo 2

Consideremos el siguiente conjunto de datos (Fuente: Cameron y Trivelli dataset)

```
use "http://econweb.rutgers.edu/frojas/teaching/
undergraduate/mus03data.dta", clear
regress ltotexp suppins phylim actlim totchr age female income
```

```
. regress ltotexp suppins phylim actlim totchr age female income

      Source |       SS       df       MS              Number of obs =    2955
-------------+------------------------------           F(  7,  2947) =  124.98
       Model |  1264.72124     7  180.674463           Prob > F      =  0.0000
    Residual |  4260.16814  2947  1.44559489           R-squared     =  0.2289
-------------+------------------------------           Adj R-squared =  0.2271
       Total |  5524.88938  2954  1.87030785           Root MSE      =  1.2023

------------------------------------------------------------------------------
     ltotexp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     suppins |   .2556428   .0462264     5.53   0.000     .1650034    .3462821
      phylim |   .3020598   .0569709     5.30   0.000      .190353    .4137666
      actlim |   .3560054   .0621118     5.73   0.000     .2342185    .4777923
      totchr |   .3758201   .0184227    20.40   0.000     .3396974    .4119429
         age |   .0038016   .0036561     1.04   0.299    -.0033672    .0109705
      female |  -.0843275   .0455442    -1.85   0.064    -.1736292    .0049741
      income |   .0025498   .0010194     2.50   0.012      .000551    .0045486
       _cons |   6.703737     .27676    24.22   0.000     6.161075      7.2464
------------------------------------------------------------------------------
```

STATA: Ejemplo 2

```
. *WHITE
. estat imtest, white

White's test for Ho: homoskedasticity
         against Ha: unrestricted heteroskedasticity

         chi2(31)      =    139.90
         Prob > chi2   =    0.0000
.
. *Breusch-Pagan 1
. estat hettest, rhs

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
         Ho: Constant variance
         Variables: suppins phylim actlim totchr age female income

         chi2(7)       =   137.30
         Prob > chi2   =   0.0000

.
. *Breusch-Pagan 2
. estat hettest

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
         Ho: Constant variance
         Variables: fitted values of ltotexp

         chi2(1)       =    48.46
         Prob > chi2   =   0.0000
```

STATA: Ejemplo 3

Consideremos el siguiente conjunto de datos (Fuente: Wooldridge dataset)

```
use "http://fmwww.bc.edu/ec-p/data/
wooldridge/MROZ", clear
regress ltotexp suppins phylim actlim totchr age female income
```

```
.
. regress lwage educ exper expersq

      Source |       SS       df       MS              Number of obs =     428
-------------+------------------------------           F(  3,   424) =   26.29
       Model |  35.0223023     3  11.6741008           Prob > F      =  0.0000
    Residual |  188.305149   424  .444115917           R-squared     =  0.1568
-------------+------------------------------           Adj R-squared =  0.1509
       Total |  223.327451   427  .523015108           Root MSE      =  .66642

------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   .1074896   .0141465     7.60   0.000     .0796837    .1352956
       exper |   .0415665   .0131752     3.15   0.002     .0156697    .0674633
     expersq |  -.0008112   .0003932    -2.06   0.040    -.0015841   -.0000382
       _cons |  -.5220407   .1986321    -2.63   0.009    -.9124668   -.1316145
------------------------------------------------------------------------------
```

STATA: Ejemplo 3

```
.
. *WHITE
. estat imtest, white

White's test for Ho: homoskedasticity
         against Ha: unrestricted heteroskedasticity

         chi2(8)      =     14.11
         Prob > chi2  =    0.0788

.
. *Breusch-Pagan 1
. estat hettest, rhs

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
         Ho: Constant variance
         Variables: educ exper expersq

         chi2(3)      =    34.08
         Prob > chi2  =   0.0000

.
. *Breusch-Pagan 2
. estat hettest

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
         Ho: Constant variance
         Variables: fitted values of lwage

         chi2(1)      =     4.20
         Prob > chi2  =   0.0405
```

Contenido

Tratamiento

Después de la prueba, si detectamos la presencia de heterocedasticidad, tenemos los siguientes pasos para el tratamiento.

1. Si nos interesa solo la inferencia, basta con calcular la varianza del estimador excluyendo el supuesto de homocedasticidad [Sugerencia de STATA: ayuda a la regresión]. Un estimador consistente es el estimador de varianza robusta de White. (sugerimos tener una muestra grande para este tratamiento).
2. Si estamos interesados en la predicción en una muestra pequeña, debe utilizar el estimador de mínimos cuadrados generalizados (MCG o GLS).

Tratamiento #1: Varianza Robusta de White

Recuerde que la varianza del estimador es (en el caso de solo 1 regresor)

$$
\begin{aligned}
E[\widehat{\beta}_1 - \beta_1]^2 &= E\left[\frac{\sum_{i=1}^{n}(x-\overline{x})u}{\sum_{i=1}^{n}(x-\overline{x})^2}\right]^2 \\
&= \frac{\sum_{i=1}^{n}(x-\overline{x})^2 E(u^2|x)}{\left[\sum_{i=1}^{n}(x-\overline{x})^2\right]^2}
\end{aligned}
$$

Como sabe, $E(u^2)$ ya no es una constante entre las observaciones (heterocedasticidad). White H. (1980) sugiere construir la varianza del estimador asumiendo $E(u^2|x) = u^2$

Varianza Robusta de White

Entonces, tenemos

$$
\begin{aligned}
E[\widehat{\beta}_1 - \beta_1]^2 &= \frac{\sum_{i=1}^{n}(x-\overline{x})^2 E(u^2|x)}{\left[\sum_{i=1}^{n}(x-\overline{x})^2\right]^2} \\
&= \frac{\sum_{i=1}^{n}(x-\overline{x})^2 \widehat{u}^2}{\left[\sum_{i=1}^{n}(x-\overline{x})^2\right]^2}
\end{aligned}
$$

Finalmente, la expresión anterior es la varianza rebusta de White.

Necesitamos algunas correciones para los grados de libertad (pequeño ajuste de muestra)

$$Var(\widehat{\beta}_1)^R = \frac{1}{n(n-k)} \frac{\sum_{i=1}^{n}(x-\overline{x})^2\widehat{u}^2}{\left[\frac{\sum_{i=1}^{n}(x-\overline{x})^2}{n}\right]^2}$$

STATA calcula la varianza robusta como en la expresión anterior.

Varianza Robusta de White, Ejemplo 1

Tengamos en cuenta el conjunto de datos en el ejemplo #1. Para considerar la varianza robusta de White, deberíamos escribir -después de declarar la regresión- lo siguiente “ `, robust` ”.

```
.
. *Robust VARIANCE
. regress price lotsize sqrft bdrms, robust

Linear regression                                      Number of obs =      88
                                                       F(  3,    84) =   23.72
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.6724
                                                       Root MSE      =  59.833

------------------------------------------------------------------------------
             |               Robust
       price |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     lotsize |   .0020677   .0012514     1.65   0.102    -.0004209    .0045563
       sqrft |   .1227782   .0177253     6.93   0.000     .0875294     .158027
       bdrms |   13.85252   8.478625     1.63   0.106    -3.008154     30.7132
       _cons |  -21.77031   37.13821    -0.59   0.559    -95.62371     52.0831
------------------------------------------------------------------------------
```

Varianza Robusta de White, Ejemplo 2

Tengamos en cuenta el conjunto de datos del ejemplo #2.

```
.
. *Robust VARIANCE
. regress ltotexp suppins phylim actlim totchr age female income, vce(robust)

Linear regression                                      Number of obs =    2955
                                                       F(  7,  2947) =  126.97
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.2289
                                                       Root MSE      =  1.2023

------------------------------------------------------------------------------
             |               Robust
     ltotexp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     suppins |   .2556428   .0465982     5.49   0.000     .1642744    .3470112
      phylim |   .3020598    .057705     5.23   0.000     .1889136     .415206
      actlim |   .3560054   .0634066     5.61   0.000     .2316797    .4803311
      totchr |   .3758201   .0187185    20.08   0.000     .3391175    .4125228
         age |   .0038016   .0037028     1.03   0.305    -.0034587     .011062
      female |  -.0843275    .045654    -1.85   0.065    -.1738444    .0051894
      income |   .0025498   .0010468     2.44   0.015     .0004973    .0046023
       _cons |   6.703737   .2825751    23.72   0.000     6.149673    7.257802
------------------------------------------------------------------------------
```

Varianza Robusta de White, Ejemplo 3

Tengamos en cuenta el conjunto de datos del ejemplo #3.

```
.
. regress lwage educ exper expersq, vce(robust)

Linear regression                               Number of obs =     428
                                                F(  3,   424) =   27.30
                                                Prob > F      =  0.0000
                                                R-squared     =  0.1568
                                                Root MSE      =  .66642

------------------------------------------------------------------------------
             |               Robust
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   .1074896    .013219     8.13   0.000     .0815068    .1334725
       exper |   .0415665    .015273     2.72   0.007     .0115462    .0715868
     expersq |  -.0008112   .0004201    -1.93   0.054    -.0016369    .0000145
       _cons |  -.5220407   .2016505    -2.59   0.010    -.9183997   -.1256816
------------------------------------------------------------------------------
```

Tratamiento #2: Mínimos cuadrados ponderados

- Siempre es posible estimar errores estándar robustos para MCO.
- Pero si se conoce la forma funcional de la heterocedasticidad, se pueden obtener estimadores más eficientes que MCO.
- La idea básica es transformar el modelo para conseguir errores homocedásticos. Este procedimiento se conoce como Mínimos Cuadrados Ponderados (MCP).

Caso: Varianza proporcional a un valor constante

- En este caso la heterocedastidas puede ser modelado como:

$$Var(\mu|x) = \sigma^2 h(x)$$

- Donde se puede representar $h(x) = h_i$.
- $E(\mu_i|\sqrt{h_i(x)}) = 0$, debido h_i es una función de x, y $Var(\mu_i|\sqrt{h_i|x}) = \sigma^2$, debido que se sabe $Var(\mu|(x)) = \sigma^2 h_i$.
- Así, si se divide la ecuación de regresión por $\sqrt{h_i}$ se tendrá un modelo donde los errores son homocedásticos.

USAT Universidad Católica Santo Toribio de Mogrovejo

Mínimos Cuadrados Generalizados

- Estimar la ecuación transformada por MCO es un ejemplo de Mínimos Cuadrados Generalizados (MCG).
- MCG es MELI en este caso.
- MCG es un procedimiento de Mínimos Cuadrados Ponderado (MCP) donde cada residuo es ponderado por el recíproco de $Var(\mu|x)$

Mínimos Cuadrados Ponderados

- Mientras que es intuitivo ver la razón por la que se realiza la transformación en MCO para conseguir homocedasticidad, dicha transformación podría ser problemática.
- MCP es un camino para alcanzar el mismo objetivo sin necesidad de hacer la transformación.
- La idea es minimizar la suma ponderada de errores al cuadrado (ponderada por $1/h_i$):

$$\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_{i1} - \ldots - \beta_k x_{ik})^2 / h_i$$

$$\sum_{i=1}^{n}(y_i^* - \beta_0 x_{i0}^* - \beta_1 x_{i1}^* - \ldots - \beta_k x_{ik}^*)^2$$

Mínimos Cuadrados Ponderados

- MCP es el camino adecuado si se conoce como es $Var(\mu|x)$.
- Aunque en la mayoría de los casos no conocemos la forma funcional de la heterocedasticidad.
- Un ejemplo donde si es conocido el patrón de heterocedasticidad es cuando los datos están agregados.
- En ese caso se puede ponderar cada observación agregada por el recíproco de la cantidad de individuos en cada grupo.

USAT Universidad Católica Santo Toribio de Mogrovejo

MCG factibles

- Lo usual es el caso donde no sepamos la forma funcional de la heterocedasticidad.
- En este caso, necesitamos estimar $h(x_i)$-
- Tipicamente se asume un forma funcional flexible:

$$Var(\mu|x) = \sigma^2 exp(\alpha_0 + \alpha_1 x_1 + ... + \alpha_k x_k)$$

- El objetivo entonces es estimar los α.

USAT Universidad Católica Santo Toribio de Mogrovejo

MCG factibles

- El supuesto implica que:

$$\mu^2 = \sigma^2 exp(\alpha_0 + \alpha_1 x_1 + ... + \alpha_k x_k) v$$

- Donde $E(v|x) = 1$, que implica $E(v) = 1$
- $ln(\mu^2) = \alpha_0 + \alpha_1 x_1 + ... + \alpha_k x_k + \epsilon$
- Donde $E(\epsilon) = 1$ y ϵ es independiente de x.
- Ahora, sabiendo que $\hat{\mu}$ es un estimado de μ, podemos estimar esto por MCO.

MCG factibles

- Así, un estimado de h es obtenido como $\hat{h} = exp(\hat{g})$, y el recíproco de esto es nuestro ponderador.
- Resumiendo:
- Estimar el modelo original MCO, guardar los residuos $\hat{\mu}$, elevarlos al cuadrado, y tomar el logaritmo. .
- Regresionar $ln(\hat{\mu}^2)$ sobre todas las variables independientes y obtener los valores estimados $\hat{g}$.
- Estimar MCG usando $1/exp(\hat{g})$ como ponderador.

Consideraciones

1. MCO sigue siendo insesgado y coherente.
2. Después de comprobar que existe evidencia de heterocedasticidad en nuestro modelo, necesitamos tratamiento.
3. Hay 2 tratamientos para el problema (reducción) de heterocedasticidad: GLS o calcular la opción robusta de White usando STATA. Nota: Heterocedasticidad no deseparece, se asume.
4. GLS debe ser factible; necesitamos conocer el verdadero patrón de heterocedasticidad.
5. El uso de una opción robusta en STATA es factible; podemos calcular la varianza de los estimadores de MCO asumiendo heterocedasticidad.
6. Los errores estándar deben corregirse asumiendo heterocedasticidad, entonces estamos listos para hacer inferencias.

Contenido

Referencias

- Referencia 1
- Referencia 2. *colocar alguna referencia*
- Referencia 3. *colocar alguna referencia*

Referencia 1
Referencia 2. *colocar alguna referencia*
Referencia 3. *colocar alguna referencia*

- Agregar alguna nota

Econometría Básica

Capítulo 08: Heterocedasticidad

José Valderrama & Freddy Rojas
jtvalderrama@gmail.com & frojasca@gmail.com
Universidad Católica Santo Toribio de Mogrovejo

Septiembre de 2021