Econometría

Diplomado Banco Central de Honduras

Instituto de Economía

Pontificia Universidad Católica de Chile

Juan Ignacio Urquiza – Junio 2022

Modelo de Regresión Lineal Clásico

- <u>Supuestos</u>:
  - Linealidad en parámetros (RLM.1).
  - Muestreo aleatorio (RLM.2).
  - Colinealidad imperfecta (RLM.3).
  - Media condicional cero (RLM.4).
  - Homocedasticidad (RLM.5).
- En la clase 2 demostramos que:
  - Bajo los supuestos RLM.1 a RLM.4, los estimadores de MCO son insesgados.
  - Bajo los supuestos RLM.1 a RLM.5, los estimadores de MCO son MELI – Teorema de Gauss-Markov.

Heterocedasticidad

- Los errores son heterocedásticos si su varianza condicional en los regresores no es constante.
- ¿Cuáles son las consecuencias?
  - Los estimadores de MCO siguen siendo insesgados pero dejan de ser MELI.
  - Los estimadores de las varianzas $V\left(\widehat{\beta}_j \middle| \boldsymbol{X}\right)$ son sesgados y por ende dejan de ser válidos para hacer inferencia.
  - Las estadísticos t de MCO no tienen distribuciones t, ni los estadísticos F siguen distribuciones F.
  - Estos problemas no se resuelven con tamaños de muestra grandes.

Heterocedasticidad

- Los errores son heterocedásticos si su varianza condicional en los regresores no es constante.
- ¿Cuáles son las consecuencias?
  - Los estimadores de MCO siguen siendo insesgados pero dejan de ser MELI.
- Antes de abandonar MCO, podemos implementar una prueba formal que detecte la presencia de heterocedasticidad.
- Si se detecta heterocedasticidad, hacemos inferencia robusta:
  - Consiste en mantener MCO pero ajustando los errores estándar para que sean válidos en presencia de heterocedasticidad de forma desconocida.

Prueba de Breusch-Pagan

- Considere el MRL múltiple:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + u$$

- La hipótesis nula de homocedasticidad implica que:

$$H_0: V(u|x_1, x_2, \dots, x_k) = \sigma^2$$

- Dado el supuesto de media condicional nula, es equivalente a:

$$H_0: E(u^2|x_1, x_2, \dots, x_k) = \sigma^2$$

- Entonces, la prueba consiste en evaluar si u^2 está relacionado (en valor esperado) con una o más variables explicativas.
- Una forma de hacerlo es suponiendo un modelo sencillo:

$$u^2 = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + \cdots + \delta_k x_k + v$$

$$\rightarrow H_0: \delta_1 = \delta_2 = \cdots = \delta_k = 0$$

Prueba de Breusch-Pagan

- Como los errores son inobservables, se utilizan los residuos:

$$\hat{u}^2 = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + \cdots + \delta_k x_k + error$$

- Luego se calcula el estadístico F para la significancia conjunta de las variables explicativas:

$$F = \frac{R^2_{\hat{u}^2}}{\left(1 - R^2_{\hat{u}^2}\right)} \times \frac{(n-k-1)}{k} \overset{a}{\rightarrow} F(k, n-k-1)$$

donde $R^2_{\hat{u}^2}$ es el R^2 de la regresión auxiliar anterior.

- Si se rechaza H_0, entonces los errores son heterocesdásticos.

```
. reg testscr str el_pct expn_stu

      Source |       SS       df       MS              Number of obs =     420
-------------+------------------------------           F(  3,   416) =  107.45
       Model |  66409.8837     3  22136.6279           Prob > F      =  0.0000
    Residual |  85699.7099   416  206.008918           R-squared     =  0.4366
-------------+------------------------------           Adj R-squared =  0.4325
       Total |  152109.594   419  363.030056           Root MSE      =  14.353

------------------------------------------------------------------------------
     testscr |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         str |  -.2863992   .4805232    -0.60   0.551    -1.230955     .658157
      el_pct |  -.6560227   .0391059   -16.78   0.000    -.7328924   -.5791529
    expn_stu |   .0038679   .0014121     2.74   0.006     .0010921    .0066437
       _cons |   649.5779   15.20572    42.72   0.000     619.6883    679.4676
------------------------------------------------------------------------------

. estat hett, rhs fstat

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
         Ho: Constant variance
         Variables: str el_pct expn_stu

         F(3 , 416)    =     14.84
         Prob > F      =    0.0000
```

Inferencia robusta

- Recuerde que en el MRL simple:

$$\widehat{\beta_1} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \beta_1 + \frac{\sum_{i=1}^{n}(x_i - \bar{x})u_i}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

- Bajo el supuesto de homocedasticidad, sabemos que:

$$V\left(\widehat{\beta_1}\middle|X\right) = \frac{\sigma^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{\sigma^2}{SCT_x}$$

- Bajo heterocedasticidad, $V(u_i|x_i) = \sigma_i^2$, lo que implica que:

$$V\left(\widehat{\beta_1}\middle|X\right) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2\sigma_i^2}{SCT_x^2}$$

- Cuando $\sigma_i^2 = \sigma^2\ \forall i$ estas fórmulas coinciden, pero si $\sigma_i^2 \neq \sigma^2$ vemos que la fórmula usual ya no es válida.

Inferencia robusta

- Se puede demostrar que un estimador válido de esta varianza se obtiene al reemplazar σ_i^2 (desconocido) por $\hat{u}_i^2$ tal que:

$$\hat{V}\left(\widehat{\beta_1}\middle|X\right) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2 \hat{u}_i^2}{SCT_x^2}$$

- Naturalmente, este análisis se extiende al MRL múltiple.
- La raíz cuadrada de $\hat{V}\left(\widehat{\beta_j}\middle|X\right)$ constituye el estándar robusto.
- Una vez obtenidos los errores estándar robustos, es fácil construir un estadístico t robusto:

$$t = \frac{estimación - valor\ hipotético}{error\ estándar\ robusto}$$

```
. reg testscr str el_pct expn_stu

      Source |       SS           df       MS      Number of obs   =       420
-------------+----------------------------------   F(3, 416)       =    107.45
       Model |  66409.8837         3  22136.6279   Prob > F        =    0.0000
    Residual |  85699.7099       416  206.008918   R-squared       =    0.4366
-------------+----------------------------------   Adj R-squared   =    0.4325
       Total |  152109.594       419  363.030056   Root MSE        =    14.353

------------------------------------------------------------------------------
     testscr |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         str |  -.2863992   .4805232    -0.60   0.551    -1.230955     .658157
      el_pct |  -.6560227   .0391059   -16.78   0.000    -.7328924   -.5791529
    expn_stu |   .0038679   .0014121     2.74   0.006     .0010921    .0066437
       _cons |   649.5779   15.20572    42.72   0.000     619.6883    679.4676
------------------------------------------------------------------------------

. reg testscr str el_pct expn_stu, robust

Linear regression                               Number of obs     =        420
                                                F(3, 416)         =     147.20
                                                Prob > F          =     0.0000
                                                R-squared         =     0.4366
                                                Root MSE          =     14.353

------------------------------------------------------------------------------
             |               Robust
     testscr |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         str |  -.2863992   .4820728    -0.59   0.553    -1.234002     .661203
      el_pct |  -.6560227   .0317844   -20.64   0.000    -.7185008   -.5935446
    expn_stu |   .0038679   .0015807     2.45   0.015     .0007607    .0069751
       _cons |   649.5779   15.45834    42.02   0.000     619.1917    679.9641
------------------------------------------------------------------------------
```

Inferencia robusta

- En este ejemplo se pueden observar varias cosas:
  - Los errores estándar no son muy diferentes. Esto ocurre a menudo, pero no tiene porqué ser así.
  - En este caso particular, las variables que son significativas empleando el estadístico t usual lo siguen siendo cuando se emplea el estadístico t robusto.
  - Los errores estándar robustos pueden ser mayores o menores que los errores estándar usuales. Sin embargo, por lo general son mayores.

Observaciones no independientes

- Recién vimos qué pasaba cuando se levantaba el supuesto de homocedasticidad (RLM.5).
- Ahora veremos qué ocurre cuando las observaciones no son independientes. Es decir, cuando se viola el supuesto (RLM.2).
  - En general, hablaremos de dependencia de los errores.
- Formalmente, existe dependencia de los errores cuando:

$$E(u_i u_j) = Cov(u_i, u_j) \neq 0 \text{ para } i \neq j.$$

- Es decir, cuando el término de error de una observación está correlacionado con el término de error de otra.

Dependencia de los errores

- ¿Cuáles son las consecuencias?
  - Los estimadores de MCO siguen siendo insesgados; sin embargo, la dependencia de los errores causa problemas a la hora de estimar la varianza de los estimadores.
  - Por lo tanto, los errores estándar de MCO y los estadísticos de contraste usuales dejan de ser válidos.
- Antes de abandonar MCO, podemos implementar una prueba que detecte la presencia de correlación de los errores.
- Si detectamos correlación, hacemos inferencia robusta:
  - Consiste en mantener MCO pero ajustando los errores estándar para que sean válidos en presencia de correlación serial.

MCO con errores correlacionados

- Considere el MRL simple: $y_t = \beta_0 + \beta_1 x_t + u_t$.
- Por simplicidad, suponga que $\bar{x} = 0$ tal que el estimador de β_1 de MCO puede escribirse como:

$$\widehat{\beta_1} = \beta_1 + SCT_x^{-1} \sum_{t=1}^{T} x_t u_t$$

- Se puede demostrar que la varianza condicional es igual a:

$$V\left(\widehat{\beta_1} \middle| x_t\right) = SCT_x^{-2} \times Var\left(\sum_{t=1}^{T} x_t u_t\right)$$

$$= SCT_x^{-2} \left[\sum_{t=1}^{T} x_t^2 V(u_t) + 2 \sum_{t=1}^{T-1} \sum_{j=1}^{T-t} x_t x_{t+j} E\left(u_t u_{t+j}\right)\right]$$

MCO con errores correlacionados

- El primer término corresponde a la varianza del estimador de MCO cuando no hay correlación de los errores.
- Por lo tanto, si los errores están correlacionados, el estimador usual de la varianza es sesgado.
- En la mayoría de las aplicaciones económicas, tanto las v. explicativas y como los errores suelen presentar correlación serial positiva.
- Entonces, la fórmula usual de la varianza de MCO suele subestimar la verdadera varianza.
- En cualquier caso, esto implica que en presencia de errores correlacionados, los estadísticos de contraste usuales ya no pueden ser utilizados para probar hipótesis.

Prueba de Breusch-Godfrey

- Considere el MRL múltiple con correlación serial de orden 1:

$$y_t = \beta_0 + \beta_1 x_{t,1} + \cdots + \beta_k x_{t,k} + u_t,$$

$$\rightarrow u_t = \rho u_{t-1} + \varepsilon_t$$

- La hipótesis nula de no correlación serial implica que:

$$H_0\colon \rho = 0$$

- Si pudiéramos observar u_t, entonces podríamos estimar ρ a partir de la regresión de u_t sobre u_{t-1}.
- Sin embargo, como los errores son inobservables, usamos los residuos de MCO.
- Entonces, primero se estima la regresión original por MCO y se obtienen los residuos ($\hat{u}_t$) para $t = 1, 2, \ldots, T$.

Prueba de Breusch-Godfrey

- Luego se estima por MCO la regresión auxiliar de los residuos sobre su primer rezago para probar H_0.
- Alternativamente, se estima la regresión auxiliar controlando por todas las v. explicativas, y se utiliza el estadístico ML:

$$\hat{u}_t = \rho\hat{u}_{t-1} + \delta_1 x_{t1} + \cdots + \delta_k x_{tk} + e_t$$

$$ML = (T-1) \times R^2_{\hat{u}} \xrightarrow{a} \chi^2_q$$

donde $(T-1)$ es el tamaño efectivo de la muestra y $R^2_{\hat{u}}$ es el R^2 de la regresión auxiliar.

- Si $ML > \chi^2_q$, se rechaza H_0 y decimos que hay evidencia en favor de la presencia de correlación serial (o autocorrelación).

```
. reg i3 inf def

      Source |       SS           df       MS      Number of obs   =        56
-------------+----------------------------------   F(2, 53)        =     40.09
       Model |  272.420338         2  136.210169   Prob > F        =    0.0000
    Residual |  180.054275        53  3.39725047   R-squared       =    0.6021
-------------+----------------------------------   Adj R-squared   =    0.5871
       Total |  452.474612        55  8.22681113   Root MSE        =    1.8432

------------------------------------------------------------------------------
          i3 | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
         inf |   .6058659   .0821348     7.38   0.000     .4411243    .7706074
         def |   .5130579   .1183841     4.33   0.000     .2756095    .7505062
       _cons |   1.733266    .431967     4.01   0.000     .8668497    2.599682
------------------------------------------------------------------------------

. estat bgodfrey, lag(1) nomiss0

Breusch-Godfrey LM test for autocorrelation
---------------------------------------------------------------------------
    lags(p)  |          chi2               df                 Prob > chi2
-------------+-------------------------------------------------------------
       1     |         21.553               1                   0.0000
---------------------------------------------------------------------------
                        H0: no serial correlation
```

```
. predict u, r

. reg u L.u inf def
```

Source	SS	df	MS			
				Number of obs	=	55
				F(3, 51)	=	10.95
Model	66.3679315	3	22.1226438	Prob > F	=	0.0000
Residual	102.993175	51	2.01947402	R-squared	=	0.3919
				Adj R-squared	=	0.3561
Total	169.361106	54	3.13631678	Root MSE	=	1.4211

u	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
u						
L1.	.62339	.1111842	5.61	0.000	.4001784	.8466015
inf	.0066283	.0657049	0.10	0.920	-.1252798	.1385364
def	-.1105103	.1009978	-1.09	0.279	-.3132719	.0922513
_cons	.1795108	.3352319	0.54	0.595	-.4934953	.8525169

$$\rightarrow LM = (T-1) \times R^2_{\hat{u}} = (56-1) \times 0.3919 \cong 21.55$$

TABLE C: Chi-Square distributions

cum probability	0.025	0.80	0.90	0.95	0.975	0.99	0.995	0.999	0.9995
right tail	0.975	0.2	0.1	0.05	0.025	0.01	0.005	0.001	0.0005
df									
1	0.00098	1.64	2.71	3.84	5.02	6.63	7.88	10.83	12.12
2	0.051	3.22	4.61	5.99	7.38	9.21	10.60	13.82	15.20
3	0.216	4.64	6.25	7.81	9.35	11.34	12.84	16.27	17.73
4	0.48	5.99	7.78	9.49	11.14	13.28	14.86	18.47	20.00
5	0.83	7.29	9.24	11.07	12.83	15.09	16.75	20.51	22.11
6	1.24	8.56	10.64	12.59	14.45	16.81	18.55	22.46	24.10
7	1.69	9.80	12.02	14.07	16.01	18.48	20.28	24.32	26.02
8	2.18	11.03	13.36	15.51	17.53	20.09	21.95	26.12	27.87
9	2.70	12.24	14.68	16.92	19.02	21.67	23.59	27.88	29.67
10	3.25	13.44	15.99	18.31	20.48	23.21	25.19	29.59	31.42
11	3.82	14.63	17.28	19.68	21.92	24.73	26.76	31.26	33.14
12	4.40	15.81	18.55	21.03	23.34	26.22	28.30	32.91	34.82
13	5.01	16.98	19.81	22.36	24.74	27.69	29.82	34.53	36.48
14	5.63	18.15	21.06	23.68	26.12	29.14	31.32	36.12	38.11
15	6.26	19.31	22.31	25.00	27.49	30.58	32.80	37.70	39.72
16	6.91	20.47	23.54	26.30	28.85	32.00	34.27	39.25	41.31
17	7.56	21.61	24.77	27.59	30.19	33.41	35.72	40.79	42.88
18	8.23	22.76	25.99	28.87	31.53	34.81	37.16	42.31	44.43
19	8.91	23.90	27.20	30.14	32.85	36.19	38.58	43.82	45.97
20	9.59	25.04	28.41	31.41	34.17	37.57	40.00	45.31	47.50
21	10.28	26.17	29.62	32.67	35.48	38.93	41.40	46.80	49.01
22	10.98	27.30	30.81	33.92	36.78	40.29	42.80	48.27	50.51
23	11.69	28.43	32.01	35.17	38.08	41.64	44.18	49.73	52.00
24	12.40	29.55	33.20	36.42	39.36	42.98	45.56	51.18	53.48
25	13.12	30.68	34.38	37.65	40.65	44.31	46.93	52.62	54.95
30	16.79	36.25	40.26	43.77	46.98	50.89	53.67	59.70	62.16
40	24.43	47.27	51.81	55.76	59.34	63.69	66.77	73.40	76.10
50	32.36	58.16	63.17	67.50	71.42	76.15	79.49	86.66	89.56
60	40.48	68.97	74.40	79.08	83.30	88.38	91.95	99.61	102.7
80	57.15	90.41	96.58	101.9	106.6	112.3	116.3	124.8	128.3
100	74.22	111.7	118.5	124.3	129.6	135.8	140.2	149.4	153.2

Inferencia robusta

- Recuerde que en el MRL simple, la varianza condicional de $\widehat{\beta_1}$ viene dada por:

$$V\left(\widehat{\beta_1} \middle| x_t\right) = SCT_x^{-2}\left[\sum_{t=1}^{T} x_t^2 V(u_t) + 2\sum_{t=1}^{T-1}\sum_{j=1}^{T-t} x_t x_{t+j} E\left(u_t u_{t+j}\right)\right]$$

- Por lo tanto, si los errores están serialmente correlacionados, es necesario tomar en cuenta la correlación entre u_t y u_{t+j}.
- En la práctica, es común suponer que una vez que los términos de error están alejados por más de unos cuantos períodos, dicha correlación es fundamentalmente 0.
- Para ello, se debe elegir un número entero $J > 0$ que controle cuánta correlación serial se estará permitiendo en el cálculo.

Inferencia robusta

- Entonces, el error estándar robusto viene dado por la raíz de:

$$\hat{V}\left(\widehat{\beta_1}\middle|x_t\right) = SCT_x^{-2}\left[\sum_{t=1}^{T} x_t^2 \hat{u}_t^2 + 2\sum_{j=1}^{J}\sum_{t=j+1}^{T} w_j x_t x_{t-j} \hat{u}_t \hat{u}_{t-j}\right]$$

donde $w_j = \left(1 - \frac{j}{J+1}\right)$.

- Estos errores estándar de Newey-West son robustos tanto a la heterocedasticidad como a la correlación serial (HAC).
- El número de rezagos se elige según la frecuencia de los datos: anual (1 a 3), trimestral (4 a 12), mensual (12 a 36).
- En Stata, simplemente utilizamos el comando "newey" (después de "tsset"), especificando el número de rezagos.

```
. reg i3 inf def
```

Source	SS	df	MS
Model	272.420338	2	136.210169
Residual	180.054275	53	3.39725047
Total	452.474612	55	8.22681113

Number of obs	=	56
F(2, 53)	=	40.09
Prob > F	=	0.0000
R-squared	=	0.6021
Adj R-squared	=	0.5871
Root MSE	=	1.8432

i3	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
inf	.6058659	.0821348	7.38	0.000	.4411243	.7706074
def	.5130579	.1183841	4.33	0.000	.2756095	.7505062
_cons	1.733266	.431967	4.01	0.000	.8668497	2.599682

```
. newey i3 inf def, lag(3)
```

Regression with Newey-West standard errors
Maximum lag = 3

Number of obs	=	56
F(2, 53)	=	18.32
Prob > F	=	0.0000

i3	Coefficient	Newey-West std. err.	t	P>|t|	[95% conf.	interval]
inf	.6058659	.1075138	5.64	0.000	.3902205	.8215112
def	.5130579	.2196756	2.34	0.023	.0724444	.9536713
_cons	1.733266	.5438208	3.19	0.002	.6424995	2.824032