

ECONOMETRÍA

APLICADA UTILIZANDO R.

PAPIME PE302513 LIBRO ELECTRÓNICO Y COMPLEMENTOS DIDÁCTICOS EN MEDIOS COMPUTACIONALES, PARA EL FORTALECIMIENTO DE LA ENSEÑANZA DE LA ECONOMETRÍA

Capítulo 7.

Heterocedasticidad

Jorge Feregrino Feregrino



Objetivo

La detección de la heterocedasticidad juega un papel importante dentro del modelo de regresión lineal, el objetivo de este capítulo es identificar la heterocedasticidad para poder aplicar técnicas de corrección.

.



Introducción

- ❑ El origen de la heterocedasticidad, está asociado a la varianza creciente de las perturbaciones aleatorias del comportamiento, de las variables incluidas en la estimación de un modelo.
- ❑ La detección de la heterocedasticidad en la mayoría de los procedimientos es útil para establecer algún tipo de solución que permite corregir este problema.
- ❑ El análisis de regresión condicional implica, obtener un parámetro estable y útil entre ambas variables, la dispersión entre las variables debe ser aleatoria para evitar problemas de estimación e inferencia econométrica.

Ejercicio aplicado en R

```
#para estimar los coeficientes del modelo de regresión lineal múltiple  
lm(Tiempo ~ Cajas + Distancia ,data=distribución)  
  
#Se le asignara un nombre a los resultados de los coeficientes  
modelo <- lm(Tiempo ~ Cajas + Distancia,data=distribución)  
  
#Para saber que se guardó con ese nombre vamos a poner  
Modelo  
  
#Para obtener las características del modelo se utiliza el siguiente comando  
summary(modelo)  
  
#Con este comando podemos obtener los rango máximos y minimo, así como la media,  
los estimadores, error estándar
```

Los efectos de la heterocedasticidad en los modelos de regresión lineal son los siguientes:

Los estimadores del MCO son lineales insesgados y consistentes, pero en presencia de heterocedasticidad son ineficientes, ya que, la varianza no es la óptima.

Las varianzas del estimador de Mínimos Cuadrados Ordinarios, no pueden calcularse con la expresión usual cuando se ha detectado heterocedasticidad:

$$v(\beta) = \sigma^2(X'X)^{-1}$$

La expresión anterior es un estimador sesgado de la varianza de los parámetros; alternativamente, debe utilizarse la siguiente expresión :

$$covar - var(\hat{\beta}) = \sigma^2[X'X]^{-1}X'\Sigma X[X'X]^{-1}$$

Cuando se realizan estimaciones bajo el supuesto de que las perturbaciones siguen un proceso homocedástico, mediante la aplicación de modelo de regresión con MCO, se cometerá un error de cálculo en la varianza, esto implica, básicamente, que nuestros cálculos sobre la “t” de student, ya no podrán compararse con los valores de referencia correctos, y lo mismo ocurrirá con el resto de cálculos que tienen origen en la varianza estimada.

- El cálculo de la R2 mediante una población, implica que no se utilizarán varianzas condicionales a los valores de las variables explicativas, de modo que el cálculo de la R2 no es afectada por la presencia de heterocedasticidad, de hecho la estimación mediante la siguiente expresión es adecuada en presencia de heterocedasticidad.

$$\tilde{\sigma}^2 = \frac{e'e}{n - k}$$

- La presencia de heteroscedasticidad en los modelos de regresión lineal, rompe con una de las restricciones más importantes en la econometría, cuya hipótesis básica señala que la varianza de los errores aleatorios, condicional a los valores de la variable independiente X, es constante:

$$\text{Var}(v_t|x_t) = \sigma^2$$

Desde el punto de vista técnico, la matriz de varianzas en un modelo de regresión ante la presencia de heterocedastidad se representa así:

$$E(UU') = \begin{bmatrix} E(u_1)^2 & \vdots & \vdots \\ E(u_{12})^2 & E(u_2)^2 & \vdots \\ \vdots & \vdots & \vdots \\ E(u_{1n})^2 & E(u_{2n})^2 & E(u_n)^2 \end{bmatrix} = \begin{bmatrix} E(u_1)^2 & 0 & 0 \\ 0 & E(u_2)^2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & E(u_n)^2 \end{bmatrix} \neq \sigma_i^2 I_n =$$

$\sigma^2 \Sigma$

Ejercicio aplicado en R

```
#para obtener la matriz de covarianzas se aplica el siguiente comando
vcov(modelo)
#para desarrollar, mejor el análisis utilizares la tabla ANOVA, aplicaremos el
siguiente comando
anova(modelo)
# Para obtener el vector de residuos utilizamos el siguiente comando, nos sirven
para revisar las propiedades de los errores
residuales<-modelo$residuals
#Tenemos que revisar los supuestos de normalidad de los residuos, a través de:
rstint<-rstandard(modelo)
#Obtiene los residuales estándares del modelo ajustado
#Vamos a realizar una grafica
win.graph()
par(mfrow=c(1,3))
#para poder observar los gráficos
hist(rstint)
boxplot(rstint)
qqnorm(rstint)
```


- Estimar el modelo inicial, sobre el que se pretende saber si hay o no heterocedasticidad, empleando MCO y determinar los errores.
- Calcular una serie con los errores del modelo anterior al cuadrado estandarizados:

$$\tilde{e}_i^2 = \frac{e_i^2}{\hat{\sigma}^2}$$

$$\hat{\sigma}^2 = \frac{e'e}{n}$$

- 3) Estimar una regresión sobre los determinantes de los errores mediante la incorporación de variables independientes (Z), mediante las cuales se busca establecer si este conjunto de variables explican el proceso de heterocedasticidad de las perturbaciones en el modelo original; la estimación propuesta es la siguiente:

$$\tilde{e}_i^2 = \alpha_0 + \alpha_1 z_{1i} + \alpha_2 z_{2i} + \dots + \alpha_p z_{pi} + \varepsilon_t$$

4) El modelo es ineficiente si la varianza de la variable dependiente estimada y su error estimado es grande. Entonces, podría afirmarse que el poder explicativo del conjunto de variables Z sobre la representación de la varianza de las perturbaciones aleatorias es escaso. Mediante el diseño de un contraste calculado con la sumatoria de los residuales de la estimación planteada en el paso 3, cuando este se encuentre cercano a cero, la probabilidad de que el proceso sea homocedástico es alta. El contraste propuesto sería el siguiente:

$$\frac{\sum \widehat{\tilde{e}}_i^2 * n}{2}$$

•

Breusch y Pagan, mostraron que el contraste se distribuye como una ji-cuadrada, cuando el proceso del modelo es homocedástico, al revisar el contraste tablas, se toman en cuenta las siguientes hipótesis:

H₀: presencia de homcedastacidad

H_a: se acepta la presen de heterocedasticdad

- Cuando la probabilidad de cometer el Error Tipo I, es muy alta no se puede rechazar la hipótesis nula, entonces, la varianza de los errores aleatorios es constante, por lo tanto, homocedásticos.

Ejercicio aplicado en R

```
#para detectar la heterocedasticidad dentro del modelo aplicamos la prueba  
Breusch-Pagan
```

```
bptest (modelo)
```

El Contraste de White para detectar heteocedasticidad

El proceso de estimación es el siguiente:

Estimar el modelo original por MCO, para obtener los errores en la estimación.

Estimar una regresión sobre los determinantes de los errores, con la incorporación de todas las variables incluidas en el estimación del primer modelo, estas elevados al cuadrado y sus combinaciones no repetidas:

$$e_i^2 = \alpha_0 + \alpha_1 x_{1i} + \dots + \alpha_k x_{ki} + \dots + \alpha_p z_{pi} + \varepsilon_t$$



Ejercicio aplicado en R

```
#vamos a aplicar una prueba de nombre White, esto con el fin de devolver las  
matrices de covarianza corregidas para hacer las inferencias  
  
hccm(modelo)  
  
#para obtener los resultados de las pruebas que hicimos aplicamos el siguiente  
comando  
  
coeftest(p,vcov=hccm(modelo))
```