

## Hoja de Ejercicios 7

### Heterocedasticidad

**Nota:** En aquellos ejercicios en los que se incluyen estimaciones y referencia al archivo de datos utilizado, el estudiante debería comprobar los resultados obtenidos en Gretl.

1. Suponga que  $Y_1$  e  $Y_2$  son variables aleatorias independientes, con  $E[Y_1] = E[Y_2] = \mu$  y  $Var[Y_1] = \sigma^2, Var[Y_2] = 4\sigma^2$ . Dadas 2 observaciones  $(y_1, y_2)$ , consideramos los dos estimadores siguientes del parámetro desconocido  $\mu$  :

$$m = \frac{1}{2}y_1 + \frac{1}{2}y_2$$

$$m' = \frac{3}{5}y_1 + \frac{2}{5}y_2.$$

- (a) Compruebe que tanto  $m$  como  $m'$  son estimadores insesgados y calcule sus varianzas. ¿Qué estimador escogería?
- (b) Obtenga el estimador insesgado  $m^* = cy_1 + (1 - c)y_2$  con varianza mínima. ¿Qué relación guarda el cociente de las ponderaciones de cada observación,  $c/(1 - c)$ , con las varianzas respectivas?
- (c) Suponga ahora que conoce las medias muestrales  $m_1$  y  $m_2$  obtenidas a partir de dos muestras aleatorias de una misma población de tamaños 100 y 25, respectivamente. ¿Cuál es el mejor estimador que puede obtener? (Pista: recuerde que bajo muestreo aleatorio la varianza de una media muestral es inversamente proporcional al tamaño muestral, y que las observaciones de una y otra muestra son independientes.)
2. [Basado en el ejercicio 8.6 del libro de Wooldridge] Utilice los datos de SLEEP75 para estimar la siguiente ecuación sobre las horas de sueño.

$$sleep = \beta_0 + \beta_1 totwrk + \beta_2 educ + \beta_3 age + \beta_4 age^2 + \beta_5 yngkid + \beta_6 male + u.$$

- (a) Contraste mediante el contraste de White en GRETL (menu *contrastes*) la presencia de heteroscedasticidad en el modelo con un nivel de significación del 10%.
- (b) Escriba un modelo que permita que la varianza de  $u$  difiera entre hombres y mujeres. La varianza no debería depender de otros factores.
- (c) Estime los parámetros de ese modelo para la heterocedasticidad. (Se tiene que estimar la ecuación *sleep* por MCO, primero, para obtener los residuos MCO.) ¿Es la varianza estimada de  $u$  mayor para los hombres o para las mujeres?
- (d) ¿Difiere de manera estadísticamente significativa la varianza de  $u$  para hombres y para mujeres?
3. En un modelo de regresión lineal en el que hay heterocedasticidad, pero los demás supuestos habituales se cumplen, indique razonadamente si cada una de las siguientes afirmaciones es verdadera o falsa:
- (a) El estimador MCO es consistente.

- (b) La forma normal de calcular la varianza del estimador MCO es incorrecta.
  - (c) El cumplimiento o no del supuesto de homocedasticidad no afecta a las propiedades del estimador MCO y este sigue siendo el Estimador Lineal e Insesgado de Mínima Varianza.
  - (d) Las respuestas 3a y 3b son ambas correctas.
4. Tener heterocedasticidad es un problema porque (indique la afirmación de entre las siguientes que es correcta):
- (a) En su forma habitual MCO supone homocedasticidad y calcula las estimaciones de los parámetros del modelo en base a ese supuesto erróneo.
  - (b) En su forma habitual MCO supone homocedasticidad y calcula las estimaciones de la varianza del estimador en base a ese supuesto erróneo.
  - (c) Sesga el estimador MCO de los parámetros.
  - (d) No sabemos cuál es la varianza del estimador MCO.
5. Si hay heterocedasticidad pero es ignorada completamente y se usa el comando para obtener el estimador de MCO habitual (sin ninguna opción especial), entonces (indique la afirmación de entre las siguientes que es correcta):
- (a) obtendremos estimaciones de los parámetros inconsistentes.
  - (b) obtendremos estimaciones de la varianza del estimador de los parámetros que no son consistentes.
  - (c) obtendremos siempre estimaciones de la varianza sesgadas al alza.
  - (d) Ninguna de las otras respuestas es correcta.
6. Si para estimar un modelo de regresión lineal en el que se cumple el supuesto de homocedasticidad usamos MCO, pero empleamos la fórmula que nos da estimaciones robustas a heterocedasticidad de los errores estándar, entonces (indique la afirmación de entre las siguientes que es correcta):
- (a) No podremos hacer inferencia usando esas estimaciones porque la estimación usando la fórmula que nos da errores robustos a heterocedasticidad es inconsistente bajo homocedasticidad.
  - (b) Dado que hay homocedasticidad, estaríamos mejor usando la fórmula habitual de los errores estándar, en lugar de calcular los errores robustos.
  - (c) Tanto bajo homocedasticidad como bajo heterocedasticidad la estimación de los errores estándar que hemos hecho es consistente.
  - (d) Las respuestas 6b y 6c son ambas correctas.
7. Si para estimar un modelo de regresión lineal bajo heterocedasticidad usamos MCO, pero teniendo cuidado de emplear la fórmula que nos da estimaciones robustas a heterocedasticidad de los errores estándar, entonces (indique la afirmación de entre las siguientes que es correcta):
- (a) Tendremos el mejor estimador de entre los lineales e insesgados.
  - (b) Ninguna de las otras respuestas es correcta.

- (c) Tendremos estimaciones de los parámetros que son insesgadas, pero errores estandar sesgados.
- (d) El  $R^2$  calculado es una medida de bondad de ajuste válida puesto que hemos tenido en cuenta que hay heterocedasticidad a la hora de calcular los errores estandar.