

1. (*Basado en Wooldridge (2009), Ch. 8, ej.C8.1*) Considere el modelo siguiente para explicar el comportamiento del 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) Dé un modelo que permita que la varianza de u difiera entre hombres (**male**) y mujeres. La varianza no debe depender de otros factores.
- b) Emplee los datos del archivo **SLEEP75.RAW** para estimar los parámetros de la ecuación de heterocedasticidad. ¿Es la varianza estimada de u mayor para los hombres o para las mujeres?
- c) ¿Es la varianza de u diferente estadísticamente para hombres y para mujeres?
2. (*Basado en Wooldridge (2009), Ch. 8, ej.C8.2*)

- a) Emplee los datos del archivo **HPRICE1.RAW** para obtener errores estándar robustos a la heterocedasticidad para la siguiente ecuación. Analice cualquier diferencia importante con los errores estándar usuales.

$$price = \beta_0 + \beta_1 lotsize + \beta_2 sqft + \beta_3 bdrms + u$$

- b) Repita el apartado a) para la ecuación

$$\log(price) = \beta_0 + \beta_1 \log(lotsize) + \beta_2 \log(sqft) + \beta_3 bdrms + error$$

- c) ¿Qué indican estos ejemplos acerca de la heterocedasticidad y de la transformación usada para la variable dependiente?
3. (*Basado en Wooldridge (2009), Ch. 8, ej.C8.3*) Aplique la prueba completa de White para heterocedasticidad a la siguiente ecuación

$$\log(price) = \beta_0 + \beta_1 \log(lotsize) + \beta_2 \log(sqft) + \beta_3 bdrms + u$$

Usando la forma de la chi-cuadrada del estadístico, obtenga el valor-p. ¿Qué puedes concluir?

4. (*Basado en Wooldridge (2009), Ch. 8, ej.C8.4*) Para este ejercicio emplee el archivo **VOTE1.RAW**.
- a) Estime un modelo en el que **voteA** sea la variable dependiente **ypartystrA**, **democA**, **log(expendA)** y **log(expendB)** sean las variables independientes. Obtenga los residuos MCO, \hat{u}_i , y regrese éstos sobre todas las variables independientes. Explique por qué obtiene $R^2 = 0$.
- b) Ahora calcule la prueba de Breusch-Pagan para heterocedasticidad. Emplee la versión del estadístico F y dé el p-valor.
- c) Calcule el caso especial de la prueba de White para heterocedasticidad, usando de nuevo la forma del estadístico F. ¿Qué tan fuerte es ahora la evidencia de heterocedasticidad?

5. (*Basado en Wooldridge (2009), Ch. 8, ej.C8.5*) Para este ejercicio emplee los datos del archivo **PNTSPRD.RAW**.

- a) La variable **sprdcvr** es una variable binaria que es igual a uno si en un partido de baloncesto universitario se cubrió la diferencia de puntos predicha en Las Vegas. El valor esperado para **sprdcvr**, por ejemplo μ , es la probabilidad de que la diferencia sea cubierta en un partido seleccionado al azar. Pruebe $H_0 : \mu = 0.5$ contra $H_1 : \mu \neq 0$ al nivel de significación del 10% y analice sus hallazgos.
- b) ¿Cuántos de los 553 partidos se jugaron en un campo neutral?
- c) Estime el modelo de probabilidad lineal

$$sprdcvr = \beta_0 + \beta_1 favhome + \beta_2 neutral + \beta_3 fav25 + \beta_4 und25 + u$$

y reporte los resultados en la forma habitual. (Dé los errores estándar usuales de MCO y los errores estándar robustos a la heterocedasticidad.) ¿Qué variable es más significativa tanto estadística como prácticamente?

- d) Explique por qué bajo la hipótesis nula $H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$ no hay heterocedasticidad en el modelo.
- e) Emplee el estadístico F usual para probar la hipótesis del apartado d). ¿Qué concluye usted?
- d) Dado el análisis previo, ¿diría que es posible predecir de manera sistemática si la diferencia de puntos predicha en Las Vegas se logrará usando la información disponible antes del partido?

6. (*Basado en Wooldridge (2009), Ch. 8, ej.C8.6*) Estime el siguiente modelo de probabilidad lineal para determinar si un joven había sido arrestado durante 1986:

$$arr86 = \beta_0 + \beta_1 pcnv + \beta_2 avgsen + \beta_3 tottime + \beta_4 ptime86 + \beta_5 qemp86 + u.$$

- a) Verifique que todos los valores ajustados estén estrictamente entre cero y uno. ¿Cuáles son los valores ajustados menor y mayor?
 - b) Estime la ecuación mediante mínimos cuadrados ponderados.
 - c) Use las estimaciones de MCP para determinar si **avgsen** y **tottime** son conjuntamente significativas al nivel de 5%.
7. (*Basado en Wooldridge (2009), Ch. 8, ej.C8.7*) Para este ejercicio emplee los datos del archivo **LOANAPP.RAW**.

- a) Estime la siguiente ecuación calculando los errores estándar robustos a la heterocedasticidad

$$approve = \beta_0 + \beta_1 white + other$$

Compare el intervalo de confianza de 95% para con el intervalo de confianza no robusto.

- b) Obtenga los valores ajustados con la regresión del inciso a). ¿Son algunos de ellos menores a cero? ¿Qué significa esto acerca de la aplicación de mínimos cuadrados ponderados?

8. (*Basado en Wooldridge (2009), Ch. 8, ej.C8.8*) Para este ejercicio emplee los datos del archivo **GPAL.RAW**.

- Emplee MCO para estimar un modelo que relacione **colGPA** con **hsGPA**, **ACT**, **skipped** y **PC**. Obtenga los residuos MCO.
- Calcule el caso especial de la prueba de White para heterocedasticidad. En la regresión de \hat{u}_i^2 sobre \widehat{colGPA}_i y \widehat{colGPA}_i^2 , obtenga los valores ajustados, es decir \hat{h}_i .
- Verifique que los valores ajustados del apartado b) sean todos estrictamente positivos. Después obtenga las estimaciones de mínimos cuadrados ponderados usando como ponderadores $1/\hat{h}_i$. Compare las estimaciones de mínimos cuadrados ponderados con las correspondientes estimaciones de MCO para el efecto del número de clases perdidas por semana (**skipped**) y el efecto de poseer una computadora (**PC**). ¿Qué puede decir de su significancia estadística?
- En la estimación por MCP del apartado c), obtenga los errores estándar robustos a la heterocedasticidad. En otras palabras, considere la posibilidad de que la función de varianza estimada en el apartado b) pueda estar mal especificada ¿Varían mucho los errores estándar de los del apartado c)?

9. (*Basado en Wooldridge (2009), Ch. 8, ej.C8.9*) Utilice la base de datos **SMOKE.RAW** para analizar la demanda de cigarrillos a partir de la siguiente ecuación

$$cigs = \beta_0 + \beta_1 \log(income) + \beta_2 \log(cigprice) + \beta_3 educ + \beta_4 age + \beta_4 age^2 + \beta_5 restaurn + u$$

- Obtenga las estimaciones de MCO.
 - Obtenga los \hat{h}_i usados en la estimación de MCP de esta ecuación. A partir de esta ecuación, obtenga los residuos no ponderados y los valores ajustados; (\hat{u}_i, \hat{y}_i , respectivamente).
 - Sean $\check{u}_i = \hat{u}_i/\sqrt{\hat{h}_i}$ y $\check{y}_i = \hat{y}_i/\sqrt{\hat{h}_i}$ las cantidades ponderadas. Realice el caso especial de la prueba de White para heterocedasticidad regresando \check{u}_i^2 sobre $\check{y}_i, \check{u}_i^2$, teniendo cuidado de incluir un intercepto. ¿Encuentra heterocedasticidad en los residuos ponderados?
 - ¿Qué implican los hallazgos del apartado c) respecto a la forma de heterocedasticidad usada para obtener los resultados estimadores de a)?
 - Obtenga errores estándar válidos para las estimaciones de MCP que permitan que la función de la varianza esté mal especificada.
10. (*Basado en Wooldridge (2009), Ch. 8, ej.C8.10*) Para este ejercicio emplee la base de datos del archivo **401KSUBS.RAW**.
- Empleando MCO estime un modelo de probabilidad lineal para **e401k**, usando como variables explicativas **inc**, **inc²**, **age**, **age²** y **male**. Obtenga tanto los errores estándar usuales de MCO como sus versiones robustas a la heterocedasticidad. ¿Existen diferencias importantes?
 - En el caso especial de la prueba de White para heterocedasticidad en la que se regresan los residuales cuadrados de MCO sobre un cuadrado de los valores ajustados por MCO, \hat{u}_i^2 sobre $\hat{y}_i, \hat{y}_i^2, i = 1, \dots, n$, argumente que el límite de probabilidad del coeficiente de \hat{y}_i debe de ser uno, el límite de probabilidad del coeficiente de \hat{y}_i^2 debe de ser -1 , y que el límite de probabilidad del intercepto debe de ser cero.

- c) Para el modelo estimado en el apartado a), obtenga la prueba de White y vea si las estimaciones de los coeficientes corresponden aproximadamente a los valores teóricos descritos en el apartado b).
- d) Después de verificar que todos los valores ajustados del inciso a) se encuentran entre cero y uno, obtenga las estimaciones de mínimos cuadrados ponderados del modelo de probabilidad lineal. ¿Difieren éstas, de manera importante, de las estimaciones de MCO?
11. (Basado en Wooldridge (2009), Ch. 8, ej.C8.11) Utilice los datos del archivo **401KSUBS.RAW**, restringiendo la muestra a $fsize = 1$, para estimar el siguiente modelo

$$nettfa = \beta_0 + \beta_1 inc + \beta_2 (age - 25)^2 + \beta_3 male + \beta_4 e401k + \beta_5 e401k * inc + u$$

- a) Estime esta ecuación por MCO y obtenga los errores estándar usuales y robustos. ¿Qué concluye acerca de la significancia estadística del término de interacción?
- b) Ahora estime el modelo por MCP empleando los ponderadores, $1/inc_i$. Calcule los errores estándar usuales y robustos para los estimadores de MCP. Usando los errores estándar robustos, ¿es el término de interacción estadísticamente significativo?
- c) Analice el coeficiente de MCP de **e401k**. ¿Es este coeficiente de mucho interés en sí mismo? Explique.
- d) Vuelva a estimar el modelo por MCP, pero use el término de interacción $e401k \cdot (inc - 30)$; el ingreso promedio en la muestra es aproximadamente 29.44. Interprete ahora el coeficiente de **e401k**.
12. (Basado en Wooldridge (2009), Ch. 8, ej.C8.12) Para responder esta pregunta emplee los datos del archivo **MEAPOO01.RAW**.

- a) Estime el modelo mediante MCO y obtenga los errores estándar usuales y los errores estándar completamente robustos. ¿Qué observa al compararlos?

$$math4 = \beta_0 + \beta_1 lunch + \beta_2 \log(enroll) + \beta_3 \log(exppp) + u.$$

- b) Aplique el caso especial de la prueba de White para heterocedasticidad. ¿Cuál es el valor de la prueba F? ¿Qué concluye usted?
- c) Obtenga los como los valores ajustados de la regresión de $\log(\hat{u}_i^2)$ sobre $\widehat{math4}_i$, $\widehat{math4}_i^2$, donde $\widehat{math4}_i$ son los valores ajustados de MCO y las \hat{u}_i son los residuos de MCO. Sea $\hat{h}_i = \exp(\hat{g}_i)$. Use las \hat{h}_i para obtener las estimaciones de MCP. ¿Hay grandes diferencias con los coeficientes de MCO?
- d) Obtenga los errores estándar de MCP que permiten una mala especificación de la función de la varianza. ¿Difieren mucho estos errores estándar de los errores estándar usuales de MCP?
- e) Para estimar el efecto del gasto (**exppp**) sobre **math4**, ¿parece ser más preciso MCO o MCP?
13. (Basado en Wooldridge (2009), Ch. 8, ej.8.1) Considere un modelo lineal para explicar el consumo mensual de cerveza (**beer**):

$$beer = \beta_0 + \beta_1 inc + \beta_2 price + \beta_3 educ + \beta_4 female + u$$

$$E(u|inc, price, educ, female) = 0$$

$$Var(u|inc, price, educ, female) = \sigma^2 inc^2.$$

Escriba la ecuación transformada que tiene un término de error homocedástico.