

Solución de ejercicios C8.5-C8.9

Constantino Carreto - Luis Valentín Cruz

Laboratorio de Econometría

Octubre, 2018

Ejercicio C8.5

Para este ejercicio emplee los datos del archivo PNTSPRD.RAW.

- (i) 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 $H_0 : \mu = .5$ contra $H_1 : \mu \neq .5$ al nivel de significancia de 10% y analice sus hallazgos. (Sugerencia: esto puede hacerse fácilmente usando una prueba t en la regresión de *sprdcvr* sobre sólo un intercepto.)
- (ii) ¿Cuántos de los 553 partidos se jugaron en un campo neutral?
- (iii) Estime el modelo de probabilidad lineal:

$$\text{sprdcvr} = \beta_0 + \beta_1 \text{favhome} + \beta_2 \text{neutral} + \beta_3 \text{fav25} + \beta_4 \text{und25} + u \quad (1)$$

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?

(iv) Explique por qué bajo la hipótesis nula

$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$, no hay heterocedasticidad en el modelo.

Bajo H_0 la probabilidad de respuesta no depende de ninguna de las variables independientes, lo cual significa que tanto la media como la varianza no dependen de las variables explicativas. Por lo tanto, $Var(u|x)$ no depende de alguna función de las X s.

(v) Emplee el estadístico F usual para probar la hipótesis del inciso iv). ¿Qué concluye usted?

(Aceptamos) No podemos rechazar H_0 : los coeficientes sí son cero

- (vi) 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?

Basados en estas variables, no es posible predecir si el spread será cubierto. El poder explicativo es muy bajo y las variables explicativas de forma conjunta son poco significativas. El coeficiente de *neutral* podría indicar que algo está sucediendo con juegos llevados a cabo en una cancha neutral. Sin embargo, no sería factible hacer un análisis de los juegos (predecirlos) basados únicamente en la variable *neutral* a menos que pudieramos confirmar su significancia con una base de datos individual y más grande.

Ejercicio C8.6

En el ejemplo 7.12 se estimó un modelo de probabilidad lineal para determinar si un joven había sido arrestado durante 1986:

$$\text{arr86} = \beta_0 + \beta_1 \text{pcnv} + \beta_2 \text{avgse} + \beta_3 \text{tottime} + \beta_4 \text{ptime86} + \beta_5 \text{qemp86} + u \quad (2)$$

- narr86: times arrested, 1986
 - arr86: fue arrestado?
- (i) Estime este modelo mediante MCO y verifique que todos los valores ajustados estén estrictamente entre cero y uno.
¿Cuáles son los valores ajustados menor y mayor?
- (ii) Estime la ecuación mediante mínimos cuadrados ponderados, como se analizó en la sección 8.5.

En la sección 8.5 se vió que para cada observación i , $Var(y_i|X_i)$ se estima mediante:

$$\hat{h}_i = \hat{y}_i(1 - \hat{y}_i) \quad (3)$$

donde \hat{y}_i es el valor ajustado de MCO para la observación i .
 \hat{h}_i serán nuestros ponderadores cuando estimemos el modelo por MCP.

- (iii) Use las estimaciones de MCP para determinar si *avgsen* y *totttime* son conjuntamente significativas al nivel de 5%.

Ejercicio C8.7

Para este ejercicio emplee los datos del archivo LOANAPP.RAW.

- (i) Estime la ecuación del inciso iii) del ejercicio para computadora C7.8, calculando los errores estándar robustos a la heterocedasticidad. Compare el intervalo de confianza de 95% para β_{white} con el intervalo de confianza no robusto.

La ecuación del inciso iii) del ejercicio C7.8 es la siguiente:

$$approve = \beta_0 + \beta_1 white + \beta_2 hrat + \beta_3 obrat + \beta_4 loanprc + \beta_5 unem +$$

$$\beta_6 male + \beta_7 married + \beta_8 dep + \beta_9 sch + \beta_{10} cosign + \beta_{11} chist + \beta_{12} pubrec + \beta_{13} mortlatl + \beta_{14} mortlatl^2 + \beta_{15} vr + u \quad (4)$$

donde:

- *approve*: al individuo se le aprueba el préstamo hipotecario?
- *white*: el solicitante es blanco?

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

No hay valores ajustados menores que cero, pero hay 213 mayores a 1. A menos que hagamos algo con esos valores ajustados, no podemos aplicar directamente MCP.

Ejercicio C8.8

Para este ejercicio emplee los datos del archivo GPA1.RAW.

- (i) Emplee MCO para estimar un modelo que relacione colGPA con hsGPA , ACT , skipped y PC . Obtenga los residuales de MCO.
- (ii) Calcule el caso especial de la prueba de White para heterocedasticidad. En la regresión de \hat{u}_i^2 sobre $\text{col}\hat{\text{GPA}}_i$, $\text{col}\hat{\text{GPA}}_i^2$, obtenga los valores ajustados, es decir \hat{h}_i .

El estadístico F para el test de White es 3.58, con 138 df y un p-value de 0.031. Entonces, a un nivel de 5%, concluimos que hay evidencia de heterocedasticidad.

- (iii) Verifique que los valores ajustados del inciso ii) sean todos estrictamente positivos. Después obtenga las estimaciones de mínimos cuadrados ponderados usando como ponderadores $\frac{1}{\hat{h}_i}$. 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?

El valor ajustado más pequeño de la regresión del inciso anterior es 0.27, mientras el más grande es 0.165. Los estimadores de MCO y MCP para las variables skipped y PC son muy similares y ambos estimadores tienen la misma significancia estadística.

- (iv) En la estimación por MCP del inciso iii), obtenga los errores estándar robustos a la heterocedasticidad. En otras palabras, considere la posibilidad de que la función de varianza estimada en el inciso ii) pueda estar mal especificada. (Vea la pregunta 8.4.) ¿Varían mucho los errores estándar de los del inciso iii)?

Los errores estándar robustos obtenidos son iguales a tres dígitos de los obtenidos en iii).

Ejercicio C8.9

En el ejemplo 8.7 se calcularon las estimaciones de MCO y un conjunto de estimaciones de MCP en una ecuación sobre la demanda de cigarros.

- (i) Obtenga las estimaciones de MCO en la ecuación (8.35).
- (ii) Obtenga los \hat{h}_i usados en la estimación de MCP de la ecuación (8.36) y reproduzca la ecuación (8.36). A partir de esta ecuación, obtenga los residuales no ponderados y los valores ajustados; llámelos \hat{u}_i y \hat{y}_i , respectivamente.
- (iii) Sean $\bar{u}_i = \frac{\hat{u}_i}{\sqrt{\hat{h}_i}}$ y $\bar{y}_i = \frac{\hat{y}_i}{\sqrt{\hat{h}_i}}$ las cantidades ponderadas. Realice el caso especial de la prueba de White para heterocedasticidad regresando \bar{u}_i^2 sobre \bar{y}_i y \bar{y}_i^2 , teniendo cuidado de incluir un intercepto, como siempre. ¿Encuentra heterocedasticidad en los residuales ponderados?

La R-cuadrada de la regresión \bar{u}_i^2 sobre \bar{y}_i y \bar{y}_i^2 es 0.027 y un p-value alrededor de cero. Rechazar H_0 , hay heterocedasticidad.

(iv) ¿Qué implican los hallazgos del inciso iii) respecto a la forma de heterocedasticidad usada para obtener (8.36)?

La heterocedasticidad encontrada en iii) muestra que es posible que con MCG, no se elimine la heterocedasticidad. Entonces, los errores estándar y las pruebas de hipótesis no son válidas.

- (v) Obtenga errores estándar válidos para las estimaciones de MCP que permitan que la función de la varianza esté mal especificada.

Las diferencias sustanciales en los errores estándar con la ecuación 8.36 es otro indicador de que el propósito de eliminar la heterocedasticidad no funcionó. Con excepción de restaurn, todos los errores estándar son mayores. Todas las variables que eran significativas con errores estándar no robustos permanecen significativas, pero los intervalos de confianza son mucho mayores.