Práctica 3
Regresión Múltiple. Inferencia

Estadística-II. INTRODUCCIÓN a la ECONOMETRÍA. UC3M

1. Utilizamos la base de datos HPRICE1 de Wooldridge para estimar este modelo

$$\log(price) = \beta_0 + \beta_1 \log(assess) + \beta_2 \log(lotsize) + \beta_3 \log(sqrft) + \beta_4 bdrms + u \quad (1)$$

donde

$$\begin{aligned} price &= \text{precio de una casa} \\ assess &= \text{tasación de la casa antes de la venta} \\ lotsize &= \text{tamaño del solar} \\ sqrft &= \text{superficie en pies cuadrados de la casa} \\ bdrms &= \text{número de dormitorios.} \end{aligned}$$

Se quiere contrastar si la tasación del precio de la casa es una valoración racional. Si fuese así, entonces un cambio de un 1 % en *assess* debería ir asociado con un cambio de un 1 % en *price*, es decir $\beta_1 = 1$. Además *lotsize*, *sqrft* y *bdrms* no deberían ayudar a predecir $\log(price)$, una vez que hemos controlado por *assess*.

Esta hipótesis se puede establecer como

$$H_0 : \beta_1 = 1, \beta_2 = 0, \beta_3 = 0, \beta_4 = 0.$$

Son 4 restricciones, tres de exclusión, pero $\beta_1 = 1$ no lo es.

a) Estima el modelo sin restringir y explora si cada una de las hipótesis individuales contenidas en H_0 pueden ser rechazadas o no. Anota SCE_{nr} y R^2_{nr}.

b) Estima el modelo restringido bajo H_0,

$$\log(price) - \log(assess) = \beta_0 + u$$

y obtén su SCE_r y R^2_r. ¿A qué sería igual SCE_r en términos de las variables $\log(price)$ y $\log(assess)$? ¿Cuántas hipótesis (q) hemos impuesto para ajustar este modelo?

c) Calcula el estadístico F

$$F = \frac{SCE_r - SCE_{nr}}{SCE_{nr}} \frac{n-k-1}{q}$$

y compáralo con el valor crítico de una $F_{n-k-1,q}$ al 5 %.

► ¿Se podría utilizar el estadístico F basado en R^2_r y R^2_{nr}? ¿Por qué?

d) Ahora nos planteamos contrastar la hipótesis

$$H_0 : \beta_1 = 0, \beta_2 = 0, \beta_3 = 0, \beta_4 = 0,$$

sobre el modelo (1). Interpreta H_0. ¿Cuál es el modelo restringido en este caso? Realiza el contraste correspondiente de la F.

► ¿Se podría utilizar el estadístico F basado en R^2_r y R^2_{nr}? ¿Por qué?

2. Utilizando el modelo

$$\log(price) = \delta_0 + \delta_1 \log(lotsize) + \delta_2 \log(sqrft) + \delta_3 bdrms + v \tag{2}$$

se quiere contrastar la hipótesis nula de que añadir una habitación y disminuir el 5 % la superficie de la vivienda (sin cambiar *lotsize*) no va a tener efectos sobre el precio de la vivienda.

a) Estima el modelo (1) y comprueba la significación individual y conjunta de las variables explicativas.

b) Para establecer dicha hipótesis en términos de los parámetros del modelo hay que comparar las predicciones sobre el aumento promedio en *price* de $\Delta bdrms = 1$ y de $\Delta sqrft = 5\,\%$:

- $\Delta E\left[price|\Delta bdrms = 1\right] \approx 100 \cdot \delta_3\,\%$.
- $\Delta E\left[price|\Delta sqrft = 5\,\%\right] \approx 5 \cdot \delta_2\,\%$.

▶ Define la hipótesis nula en términos de $(\delta_0, \delta_1, \delta_2, \delta_3)$ y estima el correspondiente modelo restringido:

$$\log(price) = \delta_0 + \delta_1 \log(lotsize) + \delta_3 \left\{bdrms + 20 \log(sqrft)\right\} + v$$

Realiza el correspondiente test de la F para contrastar dicha hipótesis.

c) Alternativamente, se puede contrastar la hipótesis anterior definiendo un nuevo parámetro

$$\theta = 20 \cdot \delta_3 - \delta_2.$$

Sustituye la ecuación $\delta_2 = 20\delta_3 - \theta$ en el modelo (2) y estima el modelo resultante en función de los parámetros $(\delta_0, \delta_1, \theta, \delta_3)$.

▶ ¿Cuál es la relación entre las estimaciones MCO de este modelo reparametrizado y las del modelo original (2)?

▶ Realiza el contraste de la hipótesis de interés

$$H_0 : \theta = 0$$

con un test de la t.

▶ ¿Cuál es la relación entre los valores de la t y de la F obtenida en (b)? ¿Qué ventajas tiene el contraste de la t con respecto al de la F?