# Práctica 2
# GRETL. Regresión Simple

Estadística-II. INTRODUCCIÓN a la ECONOMETRÍA. UC3M

1. Utilizamos la base de datos ATTEND de Wooldridge para estudiar la relación entre la asistencia a clase, los ejercicios para casa que se han entregado y las notas en los exámenes finales en la Universidad. La asistencia a clase está medida por el tanto por ciento de clases a las que ha asistido el estudiante (*atndrte*); el tanto por ciento de problemas para casa entregados (*hwrte*) y la nota en los exámenes finales estandarizada (*stndfnl*). Utilice el modelo lineal

$$stndfnl = \beta_0 + \beta_1 atndrte + \beta_2 hwrte + \varepsilon$$

para explicar la relación entre la proporción de clases atendidas y las notas en el examen final, donde $\beta_0$ y $\beta_1$ son parámetros desconocidos y $\varepsilon$ es un término de error.

*a*) Estima el modelo por MCO e interpreta los coeficientes estimados, $\hat{\beta}_1$ y $\hat{\beta}_2$.

*b*) Calcula $\widehat{Cov}\,(stndfnl, atndrte)$ y la varianza muestral $\widehat{Var}\,(atndrte)$. Entonces calcula el estimador MCO de la pendiente $\delta_1$ en el modelo "corto"

$$stndfnl = \delta_0 + \delta_1 atndrte + \varepsilon',$$

($\varepsilon'$ es otro error) utilizando la fórmula

$$\hat{\delta}_1 = \frac{\widehat{Cov}\,(stndfnl, atndrte)}{\widehat{Var}\,(atndrte)}.$$

Utiliza ahora gretl para computar $\hat{\delta}_1$.

*c*) Calcula $\widehat{Cov}\,(stndfnl, hwrte)$ y utilizando los resultados de (*a*) y (*b*) y la relación entre los coeficientes del modelo corto y largo que hemos visto en clase,

$$\hat{\delta}_1 = \hat{\beta}_1 + \beta_2 \frac{\widehat{Cov}\,(atndrte, hwrte)}{\widehat{Var}\,(atndrte)},$$

calcula otra vez $\hat{\delta}_1$ y comprueba que es idéntico a los resultados de (*b*).

▶ Comenta sobre los resultados de haber omitido la variable *hwrten* sobre la interpretación del efecto de la atención a clase sobre la nota final.

▶ Indica cómo gretl puede calcular directamente el cociente

$$\frac{\widehat{Cov}\,(atndrte, hwrte)}{\widehat{Var}\,(atndrte)},$$

e interprétalo.

*d*) Compara los errores estándar de los estimadores de los coeficientes de *atndrte* en ambos modelos, çortoz "largo."

*e*) Estima con gretl el siguiente modelo que relaciona las dos variables explicativas en el modelo "largo" para explicar *stndfnl*, es decir, *atndrte* y *hwrte*,

$$atndrte = \gamma_0 + \gamma_1 hwrte + u,$$

y guarda los residuos de MCO, $\hat{u}_i$.

Ajusta ahora el modelo

$$stndfnl = \beta_0 + \beta_1 \hat{u} + v$$

utilizando como variable explicativa los residuos $\hat{u}_i$ del ajuste anterior y llama $\tilde{\beta}_1$ al coeficiente ajustado de $\hat{u}$.

► Compara $\tilde{\beta}_1$ con el valor estimado del coeficiente de *atndrte* en $(a)$, $\hat{\beta}_1$, y en $(b)$, $\hat{\delta}_1$. Comenta el resultado. ¿Cómo interpretas los residuos $\hat{u}$?

2. Utilizamos la base de datos KIELMC de Wooldridge para estudiar el efecto de la instalación de una incineradora de basuras sobre los precios de las casas que están cerca. Tenemos datos de K.A. Kiel and K.T. McClain (1995), "House Prices During Siting Decision Stages: The Case of an Incinerator from Rumor Through Operation," Journal of Environmental Economics and Management 28, 241-255.

   Las variables que utilizamos son:

$$\begin{aligned} price &= \text{precio de venta de la casa} \\ age &= \text{edad de la casa} \\ area &= \text{tamaño de la casa en pies cuadrados} \\ dist &= \text{distancia en pies de la casa al incinerador de basura} \\ baths &= \text{número de baños en la casa} \\ nbh &= \text{barrio en el que se encuentra la casa del 1 a l6 (1 el mejor)} \end{aligned}$$

   y el modelo es

$$\log(price) = \beta_0 + \beta_1 dist + \beta_2 age + \beta_3 age^2 + \beta_4 \log(area) + \beta_5 baths + \varepsilon$$

   *a*) Estima el modelo e interpreta cada uno de los coeficientes.

   *b*) Calcula el estimador la elasticidad esperada ceteris paribus del precio con respecto a la edad de las casas para las casas con la edad crítica (la edad en la que el efecto parcial entre precio y edad cambia).

   *c*) Compara $\hat{\beta}_1$ con el estimador de la pendiente de un modelo con sólo la distancia al incinerador como variable explicativa. ¿Qué consecuencias extraes del análisis?