

Prácticas con Gretl

Marcos Bujosa

September 28, 2023

Objetivo

1. Reproducir el primer ejemplo de regresión visto en clase.
2. Mostrar datos.
3. Generar gráficos.
4. Guardar un modelo para poder consultarlo más tarde.
5. Recuperar valores ajustados, errores estimados, etc.

Carga de datos

Archivo --> Abrir datos --> Archivo de muestra y en la pestaña *Ramanathan* seleccione data3-1.

o bien teclee en línea de comandos:

```
open data3-1          # LEEMOS los datos
```

Actividad 1 - mostrar datos

Visualice los datos de precios y tamaños de las casas

- ▶ En la ventana principal de Gretl, marque con el ratón ambas variables: price, sqrt.
- ▶ “Pinche” sobre ellas con el botón derecho del ratón.
- ▶ Seleccione *mostrar valores* del menú desplegable que se ha abierto al pinchar.

o bien teclee en línea de comandos:

```
print -o price sqrt      # MOSTRAMOS los datos
```

Ayuda

Para consultar la documentación sobre cualquier comando, puede emplear el menú desplegable *Ayuda* que aparece arriba, a la derecha de la ventana principal de Gretl.

- ▶ *Ayuda -> Guía de Instrucciones* y “pinche” sobre *print*

o bien teclee en línea de comandos: help print

Actividad 2 - diagrama de dispersión

Scatter plot

- ▶ Marque price y sqrt (pulsando ctrl y pinchando con el botón derecho del ratón sobre ellas). Elija *Gráfico de dos variables XY*
- ▶ Seleccione sqft como variable del eje X

o bien teclee en línea de comandos: gnuplot price sqft

Guardar gráfico como *icono* para editarlo más tarde

- ▶ “Pinche” con el botón derecho sobre la ventana del gráfico.
- ▶ Seleccione *Guardar a sesión como icono*

o bien teclee en línea de comandos:

```
NombreDelGrafico <- gnuplot price sqft
```

(NombreDelGrafico es el nombre con el que se guardará el icono)

En la zona inferior izquierda de la ventana principal puede ver una serie de iconos. Uno de ellos es la *vista de iconos de sesión*.

Actividad 3 - Ajuste por MCO

Ajuste por MCO el modelo de regresión visto en clase

- ▶ Estime el modelo mediante los menús desplegables: *Modelo -> Mínimos Cuadrados Ordinarios*; indique a Gretl el regresando y los regresores y pulse *Aceptar*.

o bien teclee en línea de comandos:

```
ols price 0 sqft
```

(el cero 0 indica el término constante: const)

Actividad 3 - Ajuste por MCO

Ajuste por MCO el modelo de regresión visto en clase

- ▶ Estime el modelo mediante los menús desplegables: *Modelo -> Mínimos Cuadrados Ordinarios*; indique a Gretl el regresando y los regresores y pulse *Aceptar*.

o bien teclee en línea de comandos:

```
ols price 0 sqft
```

(el cero 0 indica el término constante: const)

- ▶ “Pinche” *Archivo* en la ventana del modelo estimado y seleccione *guardar como un icono y cerrar*

o bien teclee en línea de comandos:

```
Regresion <- ols price 0 sqft
```

Actividad 3 - Ajuste por MCO

Ajuste por MCO el modelo de regresión visto en clase

- ▶ Estime el modelo mediante los menús desplegables: *Modelo -> Mínimos Cuadrados Ordinarios*; indique a Gretl el regresando y los regresores y pulse *Aceptar*.

o bien teclee en línea de comandos:

```
ols price 0 sqft
```

(el cero 0 indica el término constante: const)

- ▶ “Pinche” *Archivo* en la ventana del modelo estimado y seleccione *guardar como un icono y cerrar*

o bien teclee en línea de comandos:

```
Regresion <- ols price 0 sqft
```

- ▶ Recupere el modelo “pinchando” sobre su icono

o teclee en línea de comandos el nombre que ha dado al icono seguido de .show, es decir:

```
Regresion.show
```

Actividad 4 - Recuperar los valores ajustados

Recuperemos los valores ajustados

- ▶ Desde la ventana del modelo ajustado (recupérese con su icono), “pinche” en *guardar -> valores estimados*. Elija como nombre phat (puede añadir una descripción de la variable). Pulse en *Aceptar*
- ▶ Repita para guardar los residuos con el nombre ehat

o bien teclee en línea de comandos:

```
series phat = $yhat  
series ehat = $uhat
```


Actividad 4 - Recuperar los valores ajustados

Recuperemos los valores ajustados

- ▶ Desde la ventana del modelo ajustado (recupérese con su icono), “pinche” en *guardar -> valores estimados*. Elija como nombre phat (puede añadir una descripción de la variable). Pulse en *Aceptar*
- ▶ Repita para guardar los residuos con el nombre ehat

o bien teclee en línea de comandos:

```
series phat = $yhat  
series ehat = $uhat
```

Mostremos las variables price, sqft, phat y ehat

- ▶ Marque las 4 variables (ctrl y “pinchar” con el botón derecho) y elija *mostrar valores*

o bien teclee en línea de comandos:

```
print -o price sqft phat ehat
```

Actividad 5 - Otras formas de recuperar el ajuste

- ▶ phat2: restar a los precios los errores

Desde la ventana del modelo: *Guardar -> Definir una nueva variable* y teclee: `phat2 = price - ehat`

o bien teclee en línea de comandos:

```
series phat2 = price - ehat
```

Actividad 5 - Otras formas de recuperar el ajuste

Prácticas con Gretl

Marcos Bujosa

Precio de casas
unifamiliares

Datos de
Anscombe

Houses

- ▶ phat2: Cálculo “chapucero”: $52.351 + 0.139 \text{ sqft}$
Guardar -> Definir una nueva variable y teclee:

```
series phat3 = 52.351 + 0.139*sqft
```

Actividad 5 - Otras formas de recuperar el ajuste

- phat2: Cálculo correcto: $\widehat{\beta}_1 + \widehat{\beta}_2 \text{ sqft}$
Guardar -> Definir una nueva variable y teclee:

```
series phat4 = $coeff[1] + $coeff[2]*sqft
```

o bien

```
series phat5 = $coeff(const) + $coeff(sqft)*sqft
```

Visualice los valores ¿Hay diferencias?

```
print -o price phat phat2 phat3 phat4
```

TODO

- ☐ Analizar la varianza de las estimaciones y la sensibilidad de estas a pequeños cambios en la muestra.

Código completo de la práctica: EjPvivienda.inp

```
open data3-1          # LEEMOS los datos
print -o price sqft    # MOSTRAMOS los datos

NombreDelGrafico <- gnuplot price sqft

Regresion <- ols price 0 sqft
Regresion.show

series phat = $yhat
series ehat = $uhat

series phat2 = price - ehat
series phat3 = 52.351 + 0.139*sqft
series phat4 = $coeff[1] + $coeff[2]*sqft
series phat5 = $coeff(const) + $coeff(sqft)*sqft

print -o price phat phat2 phat3 phat4
```

Objetivo

Cuando se ajusta un modelo a los datos, es NECESARIO comenzar observando gráficamente los datos. F.J. Anscombe diseñó este conjunto de datos para ilustrar la importancia de representar gráficamente los datos antes de realizar un análisis empírico.

Los datos

y_1	y_2	y_3	y_4	x	x_4
8.04	9.14	7.46	6.58	10	8
6.95	8.14	6.77	5.76	8	8
7.58	8.74	12.74	7.71	13	8
8.81	8.77	7.11	8.84	9	8
8.33	9.26	7.81	8.47	11	8
9.96	8.10	8.84	7.04	14	8
7.24	6.13	6.08	5.25	6	8
4.26	3.10	5.39	12.50	4	19
10.84	9.13	8.15	5.56	12	8
4.82	7.26	6.42	7.91	7	8
5.68	4.74	5.73	6.89	5	8

Los estadísticos de los datos

Estadísticos de los datos	
Media de cada una de las variables x :	9.0
Varianza de cada una de las variables x :	11.0
Media de cada una de las variables y :	7.5
Varianza de cada una de las variables y :	4.12
Correlación entre las variables x e y de cada regresión:	0.816

Rectas de regresión	R^2
$\widehat{y}_1 = 3 \cdot \mathbf{1} + 0.5 \cdot x$	0.67
$\widehat{y}_2 = 3 \cdot \mathbf{1} + 0.5 \cdot x$	0.67
$\widehat{y}_3 = 3 \cdot \mathbf{1} + 0.5 \cdot x$	0.67
$\widehat{y}_4 = 3 \cdot \mathbf{1} + 0.5 \cdot x_4$	0.67

Actividad 1 - Estadísticos descriptivos

Carga de datos

Archivo --> Abrir datos --> Archivo de muestra y en la pestaña *Gretl* seleccione *anscombe*.

o bien teclee en línea de comandos:

```
open anscombe
```

Visualice los estadísticos descriptivos de las datos

- ▶ Marque una variable (o varias) y “Pinche” con el botón derecho. Elija *Estadísticos principales*

o bien teclee en línea de comandos `summary` seguido de las series. Por ejemplo:

```
summary --simple y1 y2 y3 y4 x x4
```

Observar la correlación entre variables

- ▶ *Ver --> Matriz de correlación* y elija variables

o, por ejemplo, teclee en línea de comandos:

```
corr y1 y2 y3 y4 x x4
```

Actividad 2 - Cuatro regresiones

Realice las siguientes cuatro regresiones:

► Modelo 1: $Y_1 = a1 + bX + U$

► Modelo 2: $Y_2 = a1 + bX + U$

► Modelo 3: $Y_3 = a1 + bX + U$

► Modelo 4: $Y_4 = a1 + bX_4 + U$

y compare los resultados.

► Para cada modelo:

Modelo -> Mínimos Cuadrados Ordinarios; indique regresando y regresores. Pulse *Aceptar* (guarde el modelo como icono).

o bien teclee en línea de comandos:

```
Modelo1 <- ols y1 0 x
```

```
Modelo2 <- ols y2 0 x
```

```
Modelo3 <- ols y3 0 x
```

```
Modelo4 <- ols y4 0 x4
```


Código completo de la práctica: `anscombe.inp`

```
open anscombe

summary --simple y1 y2 y3 y4 x x4

corr y1 y2 y3 y4 x x4

Modelo1 <- ols y1 0 x
Modelo2 <- ols y2 0 x
Modelo3 <- ols y3 0 x
Modelo4 <- ols y4 0 x4

scatters y1 y2 y3 ; x
scatters y4 ; x4
```

Interpretación de los coeficientes estimados

Los datos

Son los del primer ejemplo de clase, junto a dos características adicionales de las casas: número de dormitorios (*bedrms*) y cuartos de baño (*baths*).

Archivo --> Abrir datos --> Archivo de muestra y en la pestaña *Ramanathan* seleccione data4-1.

o bien teclee en línea de comandos:

```
open data4-1
```

Actividad 1

Piense cuáles son los signos esperados de los parámetros del modelo

$$PRICE = \beta_1 1 + \beta_2 SQFT + \beta_3 BEDRMS + \beta_4 BATHS + U$$

Este siempre debe ser el primer paso en un análisis empírico

Actividad 2 - Ajuste MCO y resultados

Ajuste el modelo por MCO y guárdelo como un icono

- *Modelo -> Mínimos Cuadrados Ordinarios*; e indique regresando y regresores. Pulse *Aceptar*.

o bien teclee en línea de comandos:

```
Modelo1 <- ols price const sqft bedrms baths
```

Actividad 2 - Ajuste MCO y resultados

Ajuste el modelo por MCO y guárdelo como un icono

- ¿Confirman los resultados su previsión sobre los signos en β ?

Actividad 2 - Ajuste MCO y resultados

Ajuste el modelo por MCO y guárdelo como un icono

- La interpretación de los coeficientes es *ceteris paribus*, es decir, manteniendo el resto de variables sin cambios.
Así, β_3 es el efecto esperado de añadir una habitación *manteniendo el mismo tamaño de casa!*

Actividad 3 - Previsión con el modelo completo

Para una casa de 1600 *sqft* con 3 habitaciones y 2 cuartos de baño, el modelo estimado *prevé un precio* (en miles de dólares) *de*

$$129.062 + 0.154800 \cdot (1600) - 21.5875 \cdot (3) - 12.1928 \cdot (2) = 287.593$$

```
yhat1 = $coeff(const)+$coeff(sqft)*1600+$coeff(bedrms)*3+$coeff(baths)*2
```

Actividad 3 - Previsión con el modelo completo

Para una casa de 1600 *sqft* con 3 habitaciones y 2 cuartos de baño, el modelo estimado *prevé un precio* (en miles de dólares) *de*

$$129.062 + 0.154800 \cdot (1600) - 21.5875 \cdot (3) - 12.1928 \cdot (2) = 287.593$$

```
yhat1 = $coeff(const)+$coeff(sqft)*1600+$coeff(bedrms)*3+$coeff(baths)*2
```

Previsión al *ampliar* una casa con una habitación más

Sabiendo que cada pie cuadrado son 0.092 m², el modelo prevé que *ampliar la casa con una habitación adicional de 20 m² (unos 220 pies²) aumenta su precio en...* ¿cuanto?

```
yhat2 = $coeff(const)+$coeff(sqft)*(1600+220)+$coeff(bedrms)*(3+1)+$coeff(baths)*2
```

Actividad 3 - Previsión con el modelo completo

Para una casa de 1600 *sqft* con 3 habitaciones y 2 cuartos de baño, el modelo estimado *prevé un precio* (en miles de dólares) *de*

$$129.062 + 0.154800 \cdot (1600) - 21.5875 \cdot (3) - 12.1928 \cdot (2) = 287.593$$

```
yhat1 = $coeff(const)+$coeff(sqft)*1600+$coeff(bedrms)*3+$coeff(baths)*2
```

Previsión al *ampliar* una casa con una habitación más

Sabiendo que cada pie cuadrado son 0.092 m², el modelo prevé que *ampliar la casa con una habitación adicional de 20 m² (unos 220 pies²) aumenta su precio en...* ¿cuanto?

```
yhat2 = $coeff(const)+$coeff(sqft)*(1600+220)+$coeff(bedrms)*(3+1)+$coeff(baths)*2
```

► ...unos 300 mil dólares. ¿Contradice esto su idea inicial?

- ▶ Analizar la varianza de las estimaciones y la sensibilidad de estas a pequeños cambios en la muestra.
- ▶ Comparar el R^2 del ajuste con y sin estos datos adicionales.
- ▶ Comparar el R^2 ajustado con y sin estos datos adicionales.
- ▶ Comparar previsiones entre el modelo completo y el modelo sin datos sobre dormitorios y cuartos de baño.

Código completo de la práctica: houses.inp

```
open data4-1
```

```
Modelo1 <- ols price const sqft bedrms baths
```

```
yhat1 = $coeff(const)+$coeff(sqft)*1600+$coeff(bedrms)*3+$coeff(baths)*2
```

```
yhat2 = $coeff(const)+$coeff(sqft)*(1600+220)+$coeff(bedrms)*(3+1)+$coeff(baths)*2
```