

Prácticas con Gretl

Marcos Bujosa

28 de septiembre de 2023

Índice

1. TODO	2
2. Precio de casas unifamiliares	2
2.1. Practica Precio Casas Unifamiliares	2
2.1.1. Precio de casas unifamiliares	2
2.1.2. Actividad 1 - mostrar datos	2
2.1.3. Actividad 2 - diagrama de dispersión	3
2.1.4. Actividad 3 - Ajuste por MCO	3
2.1.5. Actividad 4 - Recuperar los valores ajustados	3
2.1.6. Actividad 5 - Otras formas de recuperar el ajuste	4
2.1.7. TODO	4
3. Datos de Anscombe	5
3.1. Practica con los datos de Anscombe	5
3.1.1. La importancia de <i>ver</i> los datos	5
3.1.2. Los datos	5
3.1.3. Los estadísticos de los datos	5
3.1.4. Actividad 1 - Estadísticos descriptivos	5
3.1.5. Actividad 2 - Cuatro regresiones	6
3.1.6. Actividad 3 - Discusión	6
4. Houses	7
4.1. Practica con los datos de las casas añadiendo más características	7
4.1.1. Interpretación de los coeficientes estimados	7
4.1.2. Actividad 2 - Ajuste MCO y resultados	7
4.1.3. Actividad 3 - Previsión con el modelo completo	7
4.1.4. TODO	8

1. TODO

- incluir enlaces a los guiones en github

2. Precio de casas unifamiliares

Guión: [scripts/EjPvivienda.inp](#)

En esta primera práctica con [Gretl](#) reproduciremos el ejemplo visto repetidamente en clase.

Los datos corresponden a los precios de venta y superficie útil de 14 casas unifamiliares en *University City*. San Diego, California. Año 1990. [cite:@Ramanathan:IEA-02]

Veremos como mostrar los datos, generar diagramas de dispersión, realizar un ajuste por MCO, y operar con series de datos y con parámetros estimados. Al final de la práctica aparece el guión completo con el código que evita trabajar con los menús en modo gráfico (que es el peor modo de trabajar con Gretl).

2.1. Practica Precio Casas Unifamiliares

2.1.1. Precio de casas unifamiliares

1. Objetivo

- a) Reproducir el primer ejemplo de regresión visto en clase.
- b) Mostrar datos.
- c) Generar gráficos.
- d) Guardar un modelo para poder consultarlo más tarde.
- e) Recuperar valores ajustados, errores estimados, etc.

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

o bien teclee en linea de comandos:

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

2.1.2. Actividad 1 - mostrar datos

1. Visualice los datos de precios y tamaños de las casas

- En la ventana principal de [Gretl](#), marque con el ratón ambas variables: **price**, **sqft**.
- “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 linea de comandos:

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

2. 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 linea de comandos: help print

2.1.3. Actividad 2 - diagrama de dispersión

1. Scatter plot

- Marque **price** y **sqft** (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

2. 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*.

2.1.4. Actividad 3 - Ajuste por MCO

1. 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
```

2.1.5. Actividad 4 - Recuperar los valores ajustados

1. 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
```

2. 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
```

2.1.6. 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
```

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

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

- phat2: Cálculo correcto: $\hat{\beta}_1 + \hat{\beta}_2 \text{ sqft}$
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
```

2.1.7. TODO

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

3. Datos de Anscombe

Guión: [scripts/anscombe.inp](#)

En esta práctica con [Gretl](#) trabajaremos con los datos diseñados por F.J. Anscombe para ilustrar la importancia de *ver* los diagramas de dispersión entre distintas variables para identificar deficiencias en el planteamiento de los modelos estimados, pues si únicamente se analizan los resultados numéricos dichas deficiencias quedan ocultas.

3.1. Practica con los datos de Anscombe

3.1.1. La importancia de *ver* los datos

1. 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.

3.1.2. 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

3.1.3. 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
$\widehat{y_1} = 3 \cdot 1 + 0.5 \cdot x$	0.67
$\widehat{y_2} = 3 \cdot 1 + 0.5 \cdot x$	0.67
$\widehat{y_3} = 3 \cdot 1 + 0.5 \cdot x$	0.67
$\widehat{y_4} = 3 \cdot 1 + 0.5 \cdot x_4$	0.67

3.1.4. Actividad 1 - Estadísticos descriptivos

1. 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
```

2. 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
```

3. 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
```

3.1.5. Actividad 2 - Cuatro regresiones

Realice las siguientes cuatro regresiones:

- Modelo 1: $Y_1 = a\mathbf{1} + bX + U$
- Modelo 2: $Y_2 = a\mathbf{1} + bX + U$
- Modelo 3: $Y_3 = a\mathbf{1} + bX + U$
- Modelo 4: $Y_4 = a\mathbf{1} + bX_4 + U$

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
```

3.1.6. Actividad 3 - Discusión

1. A la luz de los resultados de las regresiones ¿Qué modelo es mejor?
2. Observe el diagrama de dispersión XY en cada modelo
 - “pinche” en *Ver -->Gráficos -->Gráfico XY (scatter)* Elija la variable para el eje X (regresor) y la variable Y (regresando)

o, por ejemplo, teclee en línea de comandos:

```
gnuplot y1 x
```

podemos pintar varios diagramas juntos con:

```
gnuplot y1 y2 y3 x
```

o varios diagramas separados con:

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

3. De los cuatro modelos... ¿cuáles parecen razonables?

4. Houses

Guión: [scripts/houses.inp](#)

En esta práctica con [Gretl](#) retomamos los datos correspondientes a los precios de venta y superficie útil de 14 casas unifamiliares en *University City*. San Diego, California. Año 1990. [cite:@Ramanathan:IEA-02].

En esta ocasión también incorporamos otras características de las casas: su número de dormitorios y su número de cuartos de baño para tratar de ajustar el precio de venta.

Ello nos permitirá ilustrar cuál es la interpretación correcta de los parámetros estimados en el modelo.

4.1. Practica con los datos de las casas añadiendo más características

4.1.1. Interpretación de los coeficientes estimados

1. 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
```

2. 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

4.1.2. Actividad 2 - Ajuste MCO y resultados

1. 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
```

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

4.1.3. 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
```

1. 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?

4.1.4. TODO

- 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.