

Sesión 16 (C) — Dos sorpresas: lo que no hace falta suponer y lo que sí

Marcos Bujosa

Descripción de la práctica

Cerramos la sesión con tres experimentos que responden a dos preguntas incómodas.

La primera nació al final de la práctica A: el histograma de $\hat{\beta}_2$ tenía forma de campana, y dijimos que no debíamos sacar conclusiones de ello porque habíamos *impuesto* esa forma al generar las perturbaciones con una normal. Toca comprobarlo: ¿qué pasa si las generamos con distribuciones muy poco normales? ¿Se desmorona algo?

La segunda es más antigua. En el ejemplo de juguete de la lección 7 —aquel $\mathbf{u} = (1, -2, 0, 2, -1)$ — MCO recuperaba los coeficientes verdaderos *exactamente*, y advertimos que era una casualidad de diseño. La lección 11 explicó por qué: porque $\hat{\beta}_2 = \beta_2 \mathbf{1} + \sum_i W_i U_i$, y aquel $\mathbf{u}$ estaba construido para que el segundo término se anulase. Vamos a fabricar esa casualidad a propósito, miles de veces seguidas, y ver qué le ocurre al histograma.

Objetivos

1. Comprobar que la insesgadez y la fórmula de la varianza se cumplen con perturbaciones uniformes y con perturbaciones muy asimétricas, igual que con perturbaciones normales.
2. Entender por qué eso tenía que ocurrir: repasar la demostración de la lección 11 y localizar en qué paso se usa la distribución de U (en ninguno).
3. Ver que la *forma* de la distribución de $\hat{\beta}_2$, en cambio, sí depende de la distribución de U —y averiguar qué papel juegan, además, el tamaño muestral y el *diseño del regresor*, separando un efecto del otro—.
4. Forzar a la muestra de perturbaciones a cumplir los supuestos de forma *exacta*, no solo en esperanza, y observar el resultado.
5. Dejar planteada, con datos delante, la pregunta que abre la lección 12.

Actividad 1 - Tres perturbaciones muy distintas, la misma media y la misma varianza

Vamos a repetir el experimento de la práctica A tres veces en paralelo, cambiando únicamente cómo se sortea la perturbación. Las tres tendrán media cero y varianza σ^2 ; en todo lo demás son muy diferentes:

- *Normal*: la de la práctica A, simétrica y con colas ligeras.
- *Uniforme*: simétrica también, pero sin colas en absoluto; todos los valores de un intervalo son igual de probables. Para que su varianza sea σ^2 hay que tomar el intervalo $[-\sigma\sqrt{3}, \sigma\sqrt{3}]$.
- *Chi-cuadrado centrada*: marcadamente *asimétrica*. Una χ^2 con 4 grados de libertad tiene media 4 y varianza 8; restándole su media y dividiendo por $\sqrt{8}$ queda con media cero y varianza uno, y basta multiplicar por σ .

o bien teclee en línea de comandos:

```
open hprice2.gdt --quiet
ols price const rooms --quiet
scalar b1_true = $coeff(const)
scalar b2_true = $coeff(rooms)
scalar sigma = sqrt($ess/$df)
scalar Sxx = sum((rooms-mean(rooms))^2)
scalar var_teo = sigma^2 / Sxx
scalar a_unif = sigma*sqrt(3) # semiamplitud de la uniforme

scalar R = 2000
set seed 20261030
loop R --progressive --quiet
  series u_nor = normal(0, sigma)
  series y1 = b1_true + b2_true*rooms + u_nor
  ols y1 const rooms --quiet
  scalar b_nor = $coeff(rooms)

  series u_uni = -a_unif + 2*a_unif*uniform(0,1)
  series y2 = b1_true + b2_true*rooms + u_uni
  ols y2 const rooms --quiet
  scalar b_uni = $coeff(rooms)

  series u_chi = (randgen(X,4) - 4)/sqrt(8) * sigma
  series y3 = b1_true + b2_true*rooms + u_chi
  ols y3 const rooms --quiet
  scalar b_chi = $coeff(rooms)

  store "@workdir/repC.gdt" b_nor b_uni b_chi
endloop
```

y ahora comparamos los tres casos con lo que predice la lección 11:

```
printf "=== LO QUE PREDICE LA LECCION 11 ===\n"
printf "beta_2      = %.4f\n", b2_true
printf "Var[b2|X]   = %.2f   (desv. tipica %.2f)\n\n", var_teo, sqrt(var_teo)

printf "=== LO QUE SALE DE %d REPLICAS, SEGUN COMO SEA LA PERTURBACION ===\n", $nobs
printf "perturbacion normal   : media = %10.2f   varianza = %12.2f\n", mean(b_nor), sum((b_nor-mean(b_nor))^2)/$nobs
printf "perturbacion uniforme : media = %10.2f   varianza = %12.2f\n", mean(b_uni), sum((b_uni-mean(b_uni))^2)/$nobs
printf "perturbacion chi2(4)  : media = %10.2f   varianza = %12.2f\n", mean(b_chi), sum((b_chi-mean(b_chi))^2)/$nobs
```

```
=== LO QUE PREDICE LA LECCION 11 ===
beta_2      = 9119,5484
Var[b2|X]   = 175844,80   (desv. tipica 419,34)
```

```
=== LO QUE SALE DE 2000 REPLICAS, SEGUN COMO SEA LA PERTURBACION ===
perturbacion normal   : media =    9130,35   varianza =    178188,26
perturbacion uniforme : media =    9125,71   varianza =    170057,83
perturbacion chi2(4)  : media =    9118,59   varianza =    178413,11
```

Lo que debe observar

Las tres medias quedan cerca de β_2 y las tres varianzas cerca de σ^2/S_{xx} . Cambiar radicalmente la distribución de la perturbación —de simétrica con colas a simétrica sin colas, y de ahí a francamente asimétrica— no ha estropeado nada.

No debería sorprender: es lo que la lección 11 demostró. Merece la pena volver sobre aquella demostración y comprobar qué se usó en cada paso.

- Para la insesgadez se usó $E[U_i | \mathbf{X}] = 0$ (exogeneidad estricta) y el lema para sacar los W_i fuera de la esperanza condicionada.

- Para la varianza se usó, además, $\text{Var}[U_i | \mathbf{X}] = \sigma^2 \mathbf{1}$ (homocedasticidad) y $\text{Cov}[U_i, U_j | \mathbf{X}] = 0$ para $i \neq j$.

En ningún punto se dijo *qué distribución* tiene U . Solo se usaron su media condicional, su varianza condicional y la ausencia de covarianzas cruzadas. Y esas tres cosas las cumplen igual de bien una normal, una uniforme y una chi-cuadrado centrada. La demostración era, en este sentido, mucho más robusta de lo que a primera vista parece.

Actividad 2 - ¿Y la forma? Aquí sí cambia la historia

Que la media y la varianza no dependan de la distribución de U no significa que *nada* dependa de ella. Veamos los histogramas.

o bien teclee en línea de comandos:

```
freq b_chi --normal --plot=display
```

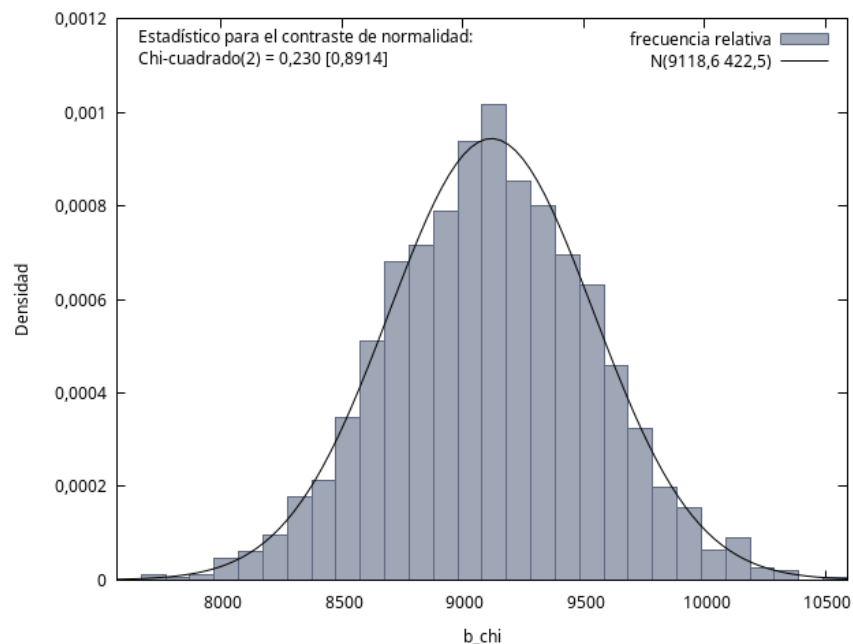


Figura 1: Distribución de $\hat{\beta}_2$ con perturbaciones χ^2 (asimétricas) y $n = 506$.

Probablemente esperaba ver algo torcido, y no lo es: con las 506 observaciones de `hprice2`, el histograma de $\hat{\beta}_2$ sigue pareciendo una campana aunque las perturbaciones sean muy asimétricas.

Antes de sacar conclusiones, probemos con una muestra pequeña y una asimetría aún más agresiva: doce viviendas y perturbaciones χ^2 con *un* grado de libertad, que es de las distribuciones más torcidas que hay a mano.

o bien teclee en línea de comandos:

```
open hprice2.gdt --quiet
ols price const rooms --quiet
scalar b1_true = $coeff(const)
scalar b2_true = $coeff(rooms)
scalar sigma = sqrt($ess/$df)

set seed 160
smpl 12 --random
```

```

set seed 99
loop 4000 --progressive --quiet
  series u_p = (randgen(X,1) - 1)/sqrt(2) * sigma      # chi2 con 1 g.l., muy asimetrica
  series y_p = b1_true + b2_true*rooms + u_p
  ols y_p const rooms --quiet
  scalar b_peq = $coeff(rooms)
  store "@workdir/repCpeq.gdt" b_peq
endloop

```

```

scalar m_p = mean(b_peq)
scalar s_p = sqrt(sum((b_peq-m_p)^2)/$nobs)
printf "n = 12, perturbacion chi2 con 1 g.l.\n"
printf " media      = %.1f\n", m_p
printf " varianza   = %.1f\n", s_p^2
printf " asimetria  = %.3f\n", (sum((b_peq-m_p)^3)/$nobs)/s_p^3

```

```

n = 12, perturbacion chi2 con 1 g.l.
media      = 9208,6
varianza   = 12853823,5
asimetria  = -0,082

```

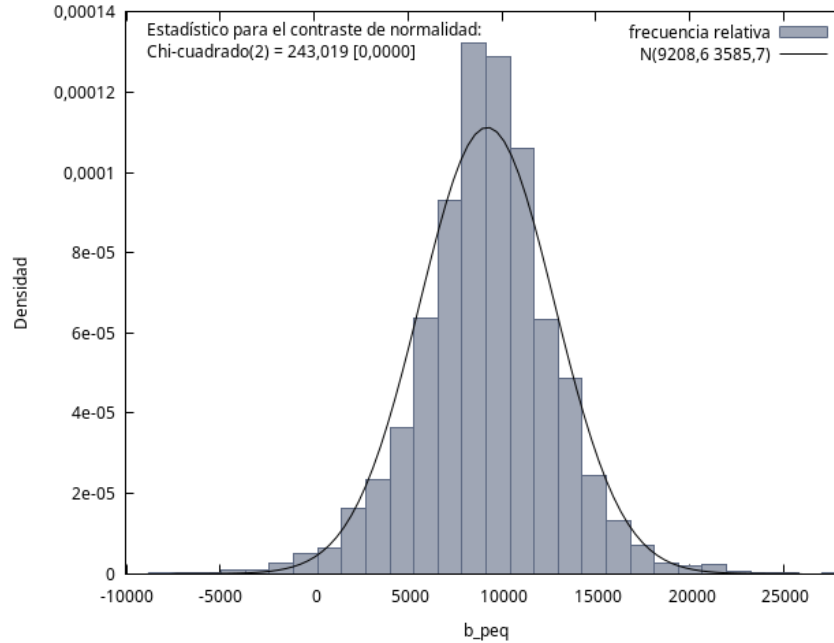


Figura 2: Distribución de $\hat{\beta}_2$ con perturbaciones χ^2 de un grado de libertad y solo $n = 12$. Sigue pareciendo simétrica.

Una simetría que no esperábamos

El resultado desconcierta: aunque la perturbación es de las más asimétricas imaginables y la muestra es diminuta, el histograma de $\hat{\beta}_2$ *sigue pareciendo simétrico*, y el coeficiente de asimetría que acabamos de calcular sale prácticamente nulo. (El contraste de normalidad de `freq --normal` sí rechaza, pero lo hace por las colas —la curtosis—, no por la asimetría.)

La explicación está, una vez más, en la fórmula de la lección 11:

$$\hat{\beta}_2 = \beta_2 + \frac{1}{n} \sum_i W_i U_i, \quad W_i \text{ proporcional a } (X_i - \bar{X}).$$

Los pesos W_i son proporcionales a las *desviaciones* del regresor respecto a su media, y esas desviaciones son negativas para aproximadamente la mitad de las observaciones. Así que $\sum_i W_i U_i$ suma perturbaciones torcidas hacia la derecha con peso positivo y perturbaciones torcidas hacia la derecha con peso *negativo* — que, al cambiarles el signo, quedan torcidas hacia la izquierda—. Las dos asimetrías se compensan en buena medida, y el resultado sale casi simétrico.

¿De quién es la culpa: de la perturbación o del regresor?

Llegados aquí conviene detenerse, porque hay dos sospechosos y todavía no hemos hecho nada para distinguirlos. Cuando el histograma salga torcido, ¿será por la distribución de la *perturbación* o por la del *regresor*?

Para separarlos hace falta cambiar una cosa cada vez. Montamos los cuatro experimentos que resultan de combinar dos perturbaciones (normal y χ^2 con un grado de libertad) con dos regresores: uno *equilibrado* —las doce viviendas reales— y otro con una *observación dominante*, en el que once viviendas comparten el mismo valor y una se sale mucho del grupo.

o bien teclee en línea de comandos:

```
open hprice2.gdt --quiet
ols price const rooms --quiet
scalar b1_true = $coeff(const)
scalar b2_true = $coeff(rooms)
scalar sigma = sqrt($ess/$df)

smpl 1 12
series x_eq = rooms          # regresor equilibrado: las doce viviendas
series x_pal = 6             # regresor con una observacion dominante:
x_pal[1] = 14                # once iguales y una que se sale
```

```
set seed 99
loop 8000 --progressive --quiet
  series u_nor = normal(0, sigma)          # perturbacion normal
  series u_chi = (randgen(X,1) - 1)/sqrt(2) * sigma # perturbacion chi2(1)

  series y1 = b1_true + b2_true*x_eq + u_nor
  ols y1 const x_eq --quiet
  scalar eq_nor = $coeff(x_eq)

  series y2 = b1_true + b2_true*x_pal + u_nor
  ols y2 const x_pal --quiet
  scalar pal_nor = $coeff(x_pal)

  series y3 = b1_true + b2_true*x_eq + u_chi
  ols y3 const x_eq --quiet
  scalar eq_chi = $coeff(x_eq)

  series y4 = b1_true + b2_true*x_pal + u_chi
  ols y4 const x_pal --quiet
  scalar pal_chi = $coeff(x_pal)

  store "@workdir/repDD.gdt" eq_nor pal_nor eq_chi pal_chi
endloop
```

```
scalar m1 = mean(eq_nor)
scalar s1 = sqrt(sum((eq_nor-m1)^2)/$nobs)
scalar m2 = mean(pal_nor)
scalar s2 = sqrt(sum((pal_nor-m2)^2)/$nobs)
scalar m3 = mean(eq_chi)
scalar s3 = sqrt(sum((eq_chi-m3)^2)/$nobs)
scalar m4 = mean(pal_chi)
scalar s4 = sqrt(sum((pal_chi-m4)^2)/$nobs)

printf "ASIMETRIA DE b2 ESTIMADO      regresor equilibrado      regresor con palanca\n"
printf "perturbacion NORMAL              %8.3f              %8.3f\n", (sum((eq_nor-m1)^3)/$nobs)/s1^3,
↵ (sum((pal_nor-m2)^3)/$nobs)/s2^3
printf "perturbacion CHI2(1)            %8.3f              %8.3f\n", (sum((eq_chi-m3)^3)/$nobs)/s3^3,
↵ (sum((pal_chi-m4)^3)/$nobs)/s4^3
```

ASIMETRIA DE b2 ESTIMADO	regresor equilibrado	regresor con palanca
perturbacion NORMAL	-0,044	0,032
perturbacion CHI2(1)	0,221	2,460

Y los dos histogramas de la columna de la derecha —los dos con el regresor de observación dominante, que es donde cabría esperar el estropicio—:

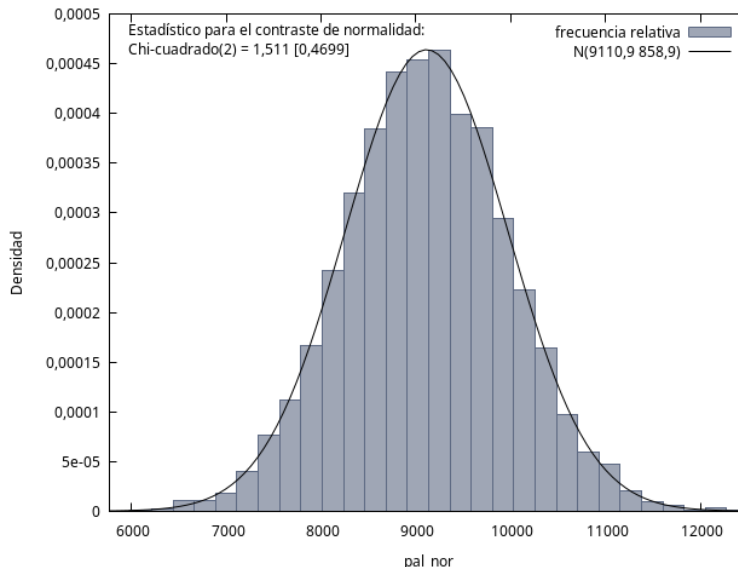


Figura 3: Regresor con observación dominante y perturbación **normal**: $\hat{\beta}_2$ sale simétrico.

Lo que debe observar

Lea la tabla por filas, que es donde está la respuesta:

- *Fila de arriba (perturbación normal)*: la asimetría es prácticamente nula en las DOS columnas. Por muy extraño que sea el regresor, con perturbaciones normales $\hat{\beta}_2$ sale simétrico.
- *Fila de abajo (perturbación χ^2)*: la asimetría es pequeña con el regresor equilibrado y grande con el de observación dominante.

De donde se sigue la respuesta a la pregunta del epígrafe, y conviene enunciarla sin ambigüedad:

La asimetría procede de la distribución de la perturbación. El regresor no la crea: solo decide si sobrevive.

Que el regresor no pueda crearla no es casualidad ni un accidente de estos números, y se ve directamente en la fórmula de la lección 11. En estos cuatro experimentos el regresor lo hemos *fijado nosotros* — x_{eq} y x_{pal} se construyen una vez y no vuelven a tocarse dentro del bucle—, de modo que los pesos son números: son los w_i de la lección 7, proporcionales a $(x_i - \mu_x)$. Y entonces

$$\hat{\beta}_2 = \beta_2 1 + \sum_i w_i U_i$$

es, ahora sí, una auténtica *combinación lineal* de las perturbaciones, con coeficientes que son números.¹ Si esas perturbaciones son normales, cualquier combinación lineal suya vuelve a ser normal —sean cuales sean los pesos—. Por eso la fila de arriba es plana: el diseño del regresor no tiene por dónde introducir asimetría.

¹ Conviene no generalizar el atajo. En la lección 11 el regresor es aleatorio, los pesos W_i son variables aleatorias, y NO se convierten en números por el hecho de condicionar en $\mathbf{X}$: lo que autoriza allí a sacarlos fuera de la esperanza es el Lema, no ninguna transmutación de funciones en números. Aquí el atajo es legítimo por un motivo distinto y mucho más humilde: el regresor está fijado de antemano, así que los w_i son números desde el principio.

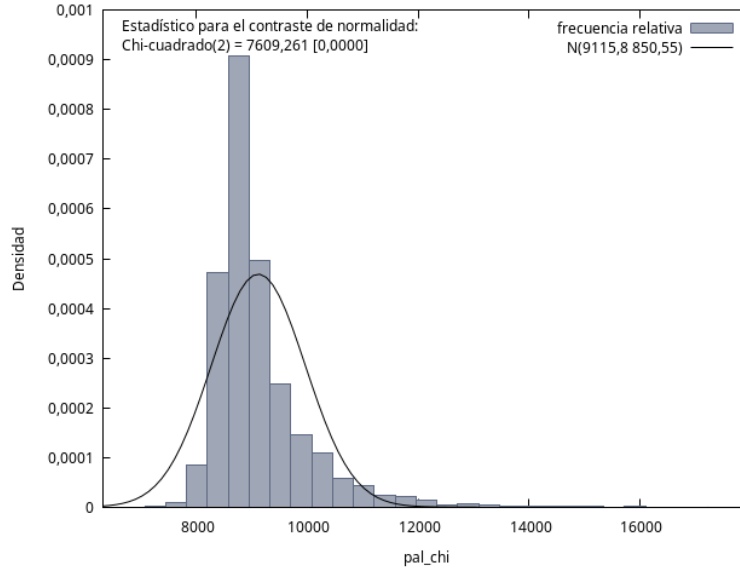


Figura 4: El **mismo** regresor, con perturbación χ^2 de un grado de libertad: ahora $\hat{\beta}_2$ hereda la asimetría.

Lo que el regresor sí controla es *cuánta* de la asimetría de U llega al resultado. Con pesos repartidos, la suma mezcla muchas perturbaciones y sus asimetrías se compensan; con un peso que domina a los demás, $\hat{\beta}_2 \approx \beta_2 + W_1 U_1$ y el estimador reproduce casi tal cual la forma de una sola perturbación.

Recapitulando las dos preguntas

Conviene no confundir lo que hemos establecido con lo que solo hemos observado:

Propiedad de $\hat{\beta}_2$	¿Depende de la distribución de U ?	¿Depende del regresor?
Media (insesgadez)	No	No
Varianza	No	Sí (a través de S_{xx})
Forma	Sí	Solo cuánto se nota

Las dos primeras filas están *demostradas* en la lección 11 y verificadas en la Actividad 1. La tercera es lo que acabamos de explorar experimentalmente, y es también la razón por la que la lección 11 se detuvo donde se detuvo: demostró exactamente lo que se podía demostrar sin suponer nada sobre la distribución de U . Todo lo que dependa de la forma —y los intervalos de confianza y los contrastes de hipótesis dependen de la forma— exige un supuesto adicional. Ese supuesto, y lo que se puede hacer con él, es el asunto de la lección 12.

Una observación. Habrá notado una regularidad: la asimetría se diluye cuando hay muchas observaciones, o cuando ninguna domina. Hay un resultado general detrás de ese comportamiento, y explica por qué la econometría aplicada puede permitirse a menudo no preocuparse demasiado por la distribución de U . No lo vamos a abordar aquí: anótelos como pregunta pendiente para la próxima lección.

Actividad 3 - El experimento imposible: forzar los supuestos exactamente

Tanto los supuestos del modelo como el corolario de la exogeneidad se enuncian *en esperanza*: $E[U_i | \mathbf{X}] = 0$ dice que la perturbación vale cero *en promedio*, no que valga cero en ninguna muestra concreta. En cualquier muestra real, el vector de perturbaciones tendrá una media muestral que no será cero y no será exactamente ortogonal al regresor.

¿Y si lo forzáramos? Podemos construir un vector de perturbaciones a medida: se sortea al azar, se le quita su proyección sobre los regresores —con lo que queda *exactamente* ortogonal a $\mathbf{1}$ y a $\mathbf{x}$ — y se reescala para que su varianza sea exactamente σ^2 . El truco es el mismo que usamos en el ejemplo de juguete de la lección 7, ahora automatizado.

o bien teclee en línea de comandos:

```
open hprice2.gdt --quiet
ols price const rooms --quiet
scalar b1_true = $coeff(const)
scalar b2_true = $coeff(rooms)
scalar sigma = sqrt($ess/$df)

set seed 20261030
loop 2000 --progressive --quiet
  series u0 = normal(0,1)
  ols u0 const rooms --quiet
  series uex = $uhat
  series uex = uex / sqrt(sum(uex^2)/$nobs) * sigma # ortogonal EXACTAMENTE a 1 y a rooms
  series y_e = b1_true + b2_true*rooms + uex # y con varianza exactamente sigma^2
  ols y_e const rooms --quiet
  scalar b_exa = $coeff(rooms)
  store "@workdir/repExact.gdt" b_exa
endloop

printf "beta_2 verdadero = %.6f\n", b2_true

printf "media de las 2000 estimaciones = %.6f\n", mean(b_exa)
printf "varianza de las estimaciones = %.10f\n", sum((b_exa-mean(b_exa))^2)/$nobs
printf "minimo = %.6f maximo = %.6f\n", min(b_exa), max(b_exa)
```

```
beta_2 verdadero = 9119,548390
media de las 2000 estimaciones = 9119,548390
varianza de las estimaciones = 0,0000000000
minimo = 9119,548390 maximo = 9119,548390
```

Lo que debe observar

Las dos mil estimaciones son *el mismo número*, y ese número es β_2 . La varianza es cero. El histograma, si lo dibujara, sería una única barra.

Vale la pena entender por qué, porque es inmediato a partir de la lección 11. Aquella lección estableció

$$\hat{\beta}_2 = \beta_2 \mathbf{1} + \sum_i W_i U_i,$$

y los pesos W_i son proporcionales a $(X_i - \bar{X})$. Ahora bien, una vez fijada la muestra concreta con la que trabajamos, esos pesos son números —los w_i de la lección 7—. Si hemos construido $\mathbf{u}$ exactamente ortogonal al regresor centrado, entonces $\sum_i w_i u_i = 0$ *en esa muestra*, no solo en promedio. El segundo sumando desaparece y lo que queda es β_2 exactamente. Cada vez. Por construcción.

Y aquí está la moraleja, que conviene no malinterpretar:

La dispersión de $\hat{\beta}_2$ no procede de ningún defecto de MCO. Procede de que, en una muestra real, la perturbación no es ortogonal al regresor —solo lo es en promedio.

Dicho de otro modo: si viviéramos en un mundo en el que cada muestra cumpliera los supuestos de forma exacta, MCO sería un instrumento perfecto y no haría falta ninguna inferencia estadística. Toda la teoría que viene a continuación —errores estándar, intervalos, contrastes— existe porque ese mundo no es el nuestro. Este experimento lo hace visible fabricándolo artificialmente.

Conviene además cerrar el círculo con la lección 7: aquel $\mathbf{u} = (1, -2, 0, 2, -1)$ del ejemplo de juguete, que parecía una curiosidad, era exactamente esto. Advertimos entonces que recuperar $\hat{\beta}_2 = 3$ exactamente “no era lo habitual”. Ahora sabemos decir con precisión qué lo hacía posible, y podemos reproducirlo a voluntad.

Preguntas de interpretación para la clase

1. Un compañero afirma: “para que MCO sea insesgado hace falta que las perturbaciones sean normales”. A la vista de la Actividad 1, ¿qué le respondería? ¿En qué paso concreto de la demostración de la lección 11 se apoyaría?
2. En la Actividad 2, con $n = 506$ el histograma parecía normal aunque U no lo fuera. ¿Diría entonces que la normalidad de U es irrelevante para todo? Distinga entre lo que la práctica demuestra y lo que solo sugiere.
3. ¿Por qué la varianza de $\hat{\beta}_2$ es cero en la Actividad 3? ¿Contradice eso la fórmula de la lección 11?
4. En la Actividad 3 hemos “hecho trampa” construyendo perturbaciones ortogonales al regresor. ¿Podría un investigador hacer esa trampa con datos reales? ¿Por qué?
5. Suponga que en una muestra real la perturbación resultara, por pura casualidad, casi ortogonal al regresor. ¿Sería su estimación especialmente buena? ¿Podría saberlo mirando los datos?

Para profundizar:

- Lección 11 ([lección 11](#)), sección “Sustituyendo el modelo poblacional”, donde se obtiene $\hat{\beta}_2 = \beta_2 1 + \sum_i W_i U_i$.
- Lección 7 ([lección 7](#)), ejemplo numérico con $\mathbf{u}$ diseñado ortogonal a los regresores.
- Wooldridge, J. M. (2020). *Introductory Econometrics*, cap. 2, sección 2.5 y cap. 4, sección 4.1 (donde aparece, por primera vez, el supuesto de normalidad).

Código completo de la práctica

```
# -----
# Copyright (C) 2026 Marcos Bujosa
# Licencia: GNU General Public License v3.0 o posterior
# Este código es software libre y puede ser redistribuido y/o modificado bajo los términos de la GPL.
# Ver el archivo LICENSE del repositorio para más detalles.
# -----

# Los dos primeros comandos son necesarios para que Gretl guarde los resultados de la práctica en el
↪ directorio de trabajo
# al ejecutar lo siguiente desde un terminal (use los nombres y ruta que correspondan)
#
#   DIRECTORIO="Nombre_Directorio_trabajo" gretlcli -b ruta/nombre_fichero_de_la_practica.inp
#
# Si esto no le funciona en su sistema, comente las siguientes dos líneas y sitúese en el directorio de
↪ trabajo de gretl
# que corresponda (configure dicho directorio de trabajo desde la ventana principal de Gretl).

string directory = getenv("DIRECTORIO")
set workdir "@directory"

set verbose off

open hprice2.gdt --quiet
ols price const rooms --quiet
scalar b1_true = $coeff(const)
scalar b2_true = $coeff(rooms)
scalar sigma   = sqrt($ess/$df)
scalar Sxx     = sum((rooms-mean(rooms))^2)
scalar var_teo = sigma^2 / Sxx
scalar a_unif  = sigma*sqrt(3)           # semiamplitud de la uniforme

scalar R = 2000
set seed 20261030
loop R --progressive --quiet
  series u_nor = normal(0, sigma)
  series y1 = b1_true + b2_true*rooms + u_nor
  ols y1 const rooms --quiet
  scalar b_nor = $coeff(rooms)

  series u_uni = -a_unif + 2*a_unif*uniform(0,1)
  series y2 = b1_true + b2_true*rooms + u_uni
  ols y2 const rooms --quiet
  scalar b_uni = $coeff(rooms)

  series u_chi = (randgen(X,4) - 4)/sqrt(8) * sigma
  series y3 = b1_true + b2_true*rooms + u_chi
  ols y3 const rooms --quiet
  scalar b_chi = $coeff(rooms)

  store "@workdir/repC.gdt" b_nor b_uni b_chi
endloop

outfile --quiet comparacionTres.txt
printf "=== LO QUE PREDICE LA LECCION 11 ===\n"
printf "beta_2      = %.4f\n", b2_true
printf "Var[b2|X]    = %.2f   (desv. típica %.2f)\n\n", var_teo, sqrt(var_teo)
end outfile
open "@workdir/repC.gdt" --quiet
outfile --quiet --append comparacionTres.txt
printf "=== LO QUE SALE DE %d REPLICAS, SEGUN COMO SEA LA PERTURBACION ===\n", $nobs
printf "perturbacion normal      : media = %10.2f   varianza = %12.2f\n", mean(b_nor),
↪ sum((b_nor-mean(b_nor))^2)/$nobs
printf "perturbacion uniforme    : media = %10.2f   varianza = %12.2f\n", mean(b_uni),
↪ sum((b_uni-mean(b_uni))^2)/$nobs
```

```

    printf "perturbacion chi2(4)    : media = %10.2f    varianza = %12.2f\n", mean(b_chi),
    ↪ sum((b_chi-mean(b_chi))^2)/$nobs
end outfile

freq b_chi --normal --plot="@workdir/hist_chi_grande.png"

open hprice2.gdt --quiet
ols price const rooms --quiet
scalar b1_true = $coeff(const)
scalar b2_true = $coeff(rooms)
scalar sigma   = sqrt($ess/$df)

set seed 160
smpl 12 --random
set seed 99
loop 4000 --progressive --quiet
    series u_p = (randgen(X,1) - 1)/sqrt(2) * sigma    # chi2 con 1 g.l., muy asimetrica
    series y_p = b1_true + b2_true*rooms + u_p
    ols y_p const rooms --quiet
    scalar b_peq = $coeff(rooms)
    store "@workdir/repCpeq.gdt" b_peq
endloop
open "@workdir/repCpeq.gdt" --quiet
outfile --quiet chiPequeno.txt
    scalar m_p = mean(b_peq)
    scalar s_p = sqrt(sum((b_peq-m_p)^2)/$nobs)
    printf "n = 12, perturbacion chi2 con 1 g.l.\n"
    printf " media      = %.1f\n", m_p
    printf " varianza    = %.1f\n", s_p^2
    printf " asimetria    = %.3f\n", (sum((b_peq-m_p)^3)/$nobs)/s_p^3
end outfile
freq b_peq --normal --plot="@workdir/hist_chi_peq.png"

open hprice2.gdt --quiet
ols price const rooms --quiet
scalar b1_true = $coeff(const)
scalar b2_true = $coeff(rooms)
scalar sigma   = sqrt($ess/$df)

smpl 1 12
series x_eq = rooms    # regresor equilibrado: las doce viviendas
series x_pal = 6        # regresor con una observacion dominante:
x_pal[1] = 14          # once iguales y una que se sale
set seed 99
loop 8000 --progressive --quiet
    series u_nor = normal(0, sigma)    # perturbacion normal
    series u_chi = (randgen(X,1) - 1)/sqrt(2) * sigma    # perturbacion chi2(1)

    series y1 = b1_true + b2_true*x_eq + u_nor
    ols y1 const x_eq --quiet
    scalar eq_nor = $coeff(x_eq)

    series y2 = b1_true + b2_true*x_pal + u_nor
    ols y2 const x_pal --quiet
    scalar pal_nor = $coeff(x_pal)

    series y3 = b1_true + b2_true*x_eq + u_chi
    ols y3 const x_eq --quiet
    scalar eq_chi = $coeff(x_eq)

    series y4 = b1_true + b2_true*x_pal + u_chi
    ols y4 const x_pal --quiet
    scalar pal_chi = $coeff(x_pal)

    store "@workdir/repDD.gdt" eq_nor pal_nor eq_chi pal_chi
endloop
open "@workdir/repDD.gdt" --quiet

```

```

outfile --quiet dosPorDos.txt
    scalar m1 = mean(eq_nor)
    scalar s1 = sqrt(sum((eq_nor-m1)^2)/$nobs)
    scalar m2 = mean(pal_nor)
    scalar s2 = sqrt(sum((pal_nor-m2)^2)/$nobs)
    scalar m3 = mean(eq_chi)
    scalar s3 = sqrt(sum((eq_chi-m3)^2)/$nobs)
    scalar m4 = mean(pal_chi)
    scalar s4 = sqrt(sum((pal_chi-m4)^2)/$nobs)

    printf "ASIMETRIA DE b2 ESTIMADO      regresor equilibrado   regresor con palanca\n"
    printf "perturbacion NORMAL              %8.3f              %8.3f\n",
    ↪ (sum((eq_nor-m1)^3)/$nobs)/s1^3, (sum((pal_nor-m2)^3)/$nobs)/s2^3
    printf "perturbacion CHI2(1)            %8.3f              %8.3f\n",
    ↪ (sum((eq_chi-m3)^3)/$nobs)/s3^3, (sum((pal_chi-m4)^3)/$nobs)/s4^3
end outfile
freq pal_nor --normal --plot="@workdir/hist_pal_nor.png"
freq pal_chi --normal --plot="@workdir/hist_pal_chi.png"

open hprice2.gdt --quiet
ols price const rooms --quiet
scalar b1_true = $coeff(const)
scalar b2_true = $coeff(rooms)
scalar sigma = sqrt($ess/$df)

set seed 20261030
loop 2000 --progressive --quiet
    series u0 = normal(0,1)
    ols u0 const rooms --quiet
    series uex = $uhat
    series uex = uex / sqrt(sum(uex^2)/$nobs) * sigma # ortogonal EXACTAMENTE a 1 y a rooms
    series y_e = b1_true + b2_true*rooms + uex # y con varianza exactamente sigma^2
    ols y_e const rooms --quiet
    scalar b_exa = $coeff(rooms)
    store "@workdir/repExacta.gdt" b_exa
endloop

outfile --quiet uExacta.txt
    printf "beta_2 verdadero = %.6f\n", b2_true
end outfile
open "@workdir/repExacta.gdt" --quiet
outfile --quiet --append uExacta.txt
    printf "media de las 2000 estimaciones = %.6f\n", mean(b_exa)
    printf "varianza de las estimaciones = %.10f\n", sum((b_exa-mean(b_exa))^2)/$nobs
    printf "minimo = %.6f    maximo = %.6f\n", min(b_exa), max(b_exa)
end outfile

```

Enlace al guión: [S16-PrcT-C-sorpresas.inp](#)

Respuestas

1. *¿Hace falta normalidad para la insesgadez?* No, y el experimento lo confirma: con perturbaciones uniformes y con perturbaciones muy asimétricas la media de las estimaciones sigue quedando en β_2 . La razón está en la demostración de la lección 11: la insesgadez sale de tomar esperanza condicional en $\hat{\beta}_2 = \beta_2 + \sum_i W_i U_i$, usar el lema para sacar los W_i fuera, y aplicar $E[U_i | \mathbf{X}] = 0$. Ese supuesto habla de la *media condicional* de U , no de su distribución. Una uniforme y una chi-cuadrado centrada tienen media condicional cero igual que una normal, y eso es cuanto se necesita.
2. *¿Es irrelevante la normalidad?* Hay que separar dos cosas. Lo que la práctica *demuestra* es que la insesgadez y la fórmula de la varianza no requieren normalidad: se cumplen con las tres distribuciones ensayadas. Lo que la práctica solo *sugiere*, y no prueba, es algo sobre la forma: con muchas observaciones el histograma de $\hat{\beta}_2$ tiende a parecer una campana aunque U no sea normal, mientras que con pocas observaciones y una perturbación muy asimétrica deja de parecerlo. De ahí no se puede concluir que la normalidad sobre. Puede concluirse, como mucho, que el papel de ese supuesto es más sutil de lo que parecía, y que hace falta un resultado teórico —que aquí no tenemos— para saber cuándo cabe prescindir de él.
3. *Varianza cero.* Porque hemos construido cada perturbación de modo que $\sum_i w_i u_i = 0$ exactamente en cada muestra, de forma que $\hat{\beta}_2 = \beta_2$ siempre, sin variabilidad alguna. No contradice la fórmula de la lección 11, sino que ilustra su alcance: aquella fórmula calcula la varianza de $\hat{\beta}_2$ *bajo los supuestos del modelo*, entre los cuales está que $\mathbf{U}$ sea aleatoria con $E[U_i | \mathbf{X}] = 0$. Nuestras perturbaciones fabricadas no cumplen ese modelo: no son un sorteo libre, sino un sorteo al que después le hemos impuesto una restricción que depende de $\mathbf{X}$. Al condicionar el sorteo a esa restricción hemos salido del marco en el que la fórmula era válida.
4. *¿Podría hacerse con datos reales?* No. Para restar a $\mathbf{u}$ su proyección sobre los regresores hay que *conocer* $\mathbf{u}$, y la perturbación es justamente lo que no se observa: sabemos $\mathbf{y}$ y sabemos $\mathbf{x}$, pero $\mathbf{u}$ solo existe en el modelo. Lo que sí observamos son los residuos $\hat{\mathbf{e}}$, que por construcción *siempre* son exactamente ortogonales a los regresores —esa es una identidad algebraica de MCO, no un supuesto—. Precisamente por eso los residuos no sirven para detectar si la perturbación era o no ortogonal: lo son siempre, se cumpla o no el supuesto.
5. *Una muestra afortunada.* Sí sería una estimación especialmente buena: si $\sum_i w_i u_i$ resultara casi nulo, $\hat{\beta}_2$ quedaría casi en β_2 . Pero no hay forma de saberlo mirando los datos, por lo dicho en la respuesta anterior: la cantidad que habría que mirar, $\sum_i w_i u_i$, depende de $\mathbf{u}$, que es inobservable. Este es, en el fondo, el motivo de ser de toda la inferencia estadística: como nunca sabremos si nuestra muestra concreta fue afortunada o desafortunada, lo único que podemos hacer es cuantificar cuánta suerte cabe esperar —y eso es lo que mide un error estándar.