

Sesión 18 (C) — ¿Cubre de verdad el 95 por ciento? La cobertura del intervalo y la frecuencia de detección

Marcos Bujosa

Descripción de la práctica

La lección 12 fue muy cuidadosa con lo que el intervalo de confianza *dice* y lo que *no dice*: dice que el 95 % de las muestras producen un intervalo que contiene a β_2 ; no dice que β_2 esté en *este* intervalo con probabilidad 0,95. Es una afirmación sobre una *frecuencia*, y una frecuencia se comprueba repitiendo. Esta práctica la repite.

Reutilizamos el mundo artificial de la práctica B —el panel disperso de la lección 12: doce valores del regresor de 22,5 a 77,5, $\beta_1 = 10$, $\beta_2 = 0,5$ y perturbaciones normales con $\sigma = 1,1$ — y la mecánica del `loop` que aprendimos en el laboratorio anterior. En cada réplica sorteamos doce perturbaciones nuevas, estimamos, construimos el intervalo y anotamos si contiene al 0,5 verdadero. Después contamos.

Los datos son *simulados* por la razón de siempre: solo conociendo la verdad se puede saber si el intervalo la contiene.

Objetivos

1. Comprobar que el intervalo al 95 % contiene al verdadero β_2 en (aproximadamente) el 95 % de las réplicas, y ver, réplica a réplica, cuáles fallan.
2. Medir el *precio de estimar* σ : qué cobertura se obtiene si se usa 1,96 en lugar del valor crítico de la t_{10} .
3. Contrastar una hipótesis *falsa por poco* con $n = 12$ y con $n = 120$: la identidad $|t| = \sqrt{n-k} \times$ razón de catetos, ejecutada.
4. (Apéndice) Ver qué ocurre con la cobertura cuando las perturbaciones no son normales.

Actividad 1 - ¿Cubre el 95 por ciento?

Construimos el mundo, fijamos el valor crítico de la t con $n - k = 10$ grados de libertad, y repetimos el experimento 4000 veces guardando, en cada réplica, la estimación, su error estándar, los dos extremos del intervalo, un indicador de si contiene al verdadero β_2 y el estadístico t para la hipótesis *cierta* $H_0: \beta_2 = 0,5$.

en línea de comandos:

```
nulldata 12
setobs 1 1 --cross-section
series x = 22.5 + 5*(index - 1)           # el regresor disperso de la leccion 12
scalar b1 = 10
scalar b2 = 0.5
scalar sig = 1.1
scalar R = 4000
scalar tc = critical(t, 10, 0.025)        # valor critico de la t_10 al 5%
scalar zc = 1.96                          # el de la normal, para la Actividad 2

set seed 20261113
loop R --progressive --quiet
  series u = normal(0, sig)
  series y = b1 + b2*x + u
  ols y const x --quiet
```

Licencia: Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0).

```

scalar b2hat = $coeff(x)
scalar ee    = $stderr(x)
scalar cL    = b2hat - tc*ee          # extremos del IC al 95%
scalar cR    = b2hat + tc*ee
scalar dentro = (cL <= b2 && b2 <= cR) # 1 si el IC contiene al verdadero beta_2
scalar tstat  = (b2hat - b2)/ee       # t para la H0 CIERTA beta_2 = 0,5
scalar dentroZ = (b2hat - zc*ee <= b2 && b2 <= b2hat + zc*ee) # con 1,96 en vez de t_c
store "@workdir/replicasIC.gdt" b2hat ee cL cR dentro tstat dentroZ
endloop

```

```

open "@workdir/replicasIC.gdt" --quiet
scalar tc = critical(t, 10, 0.025)      # al abrir otro fichero Gretl olvida los escalares
scalar zc = 1.96

```

```

printf "Replicas" : %d\n", $nobs
printf "Intervalos que contienen al verdadero b2" : %.4f (objetivo: 0,95)\n", mean(dentro)
printf "Rechazos de la H0 cierta al 5%" : %.4f (objetivo: 0,05)\n", mean(abs(tstat) > tc)

```

```

Replicas : 4000
Intervalos que contienen al verdadero b2 : 0,9520 (objetivo: 0,95)
Rechazos de la H0 cierta al 5% : 0,0480 (objetivo: 0,05)

```

Atención al detalle práctico ya conocido del laboratorio anterior: al abrir el fichero de réplicas, Gretl descarta los escalares que teníamos definidos; por eso `tc` y `zc` se vuelven a definir después de abrirlo.

Los intervalos, uno a uno

Un porcentaje es un resumen. Para *ver* qué significa, dibujamos los cien primeros intervalos: sus dos extremos como dos líneas, y el verdadero β_2 como una línea horizontal. Cada vez que la banda inferior sube por encima de 0,5, o la superior baja por debajo, ese intervalo ha fallado.

en línea de comandos:

```

smpl 1 100
series B2 = 0.5
gnuplot cL cR B2 --time-series --with-lines --output=display
printf "Intervalos que fallan entre los 100 primeros: %d\n", 100 - sum(dentro)
smpl full

```

```
Intervalos que fallan entre los 100 primeros: 7
```

Lo que debe observar

La frecuencia empírica de cobertura queda muy cerca de 0,95, y la de rechazos de la hipótesis cierta muy cerca de 0,05: son la misma cosa vista desde dos lados. Con 4 000 réplicas no cabe esperar exactamente 0,9500 —la frecuencia observada es a su vez una variable aleatoria—, pero sí una distancia pequeña.

Y la figura muestra lo que la frecuencia esconde. El intervalo *cambia* de una réplica a otra: sus extremos son variables aleatorias, funciones de la muestra. El β_2 verdadero, en cambio, no se mueve. Esa es la razón de la lectura correcta de la lección 12: lo aleatorio es el intervalo, no el parámetro, y el 95 % es la frecuencia con la que el intervalo aleatorio atrapa al parámetro fijo. Mirando un solo intervalo —cualquiera de los cien— no hay forma de saber si es de los que aciertan.

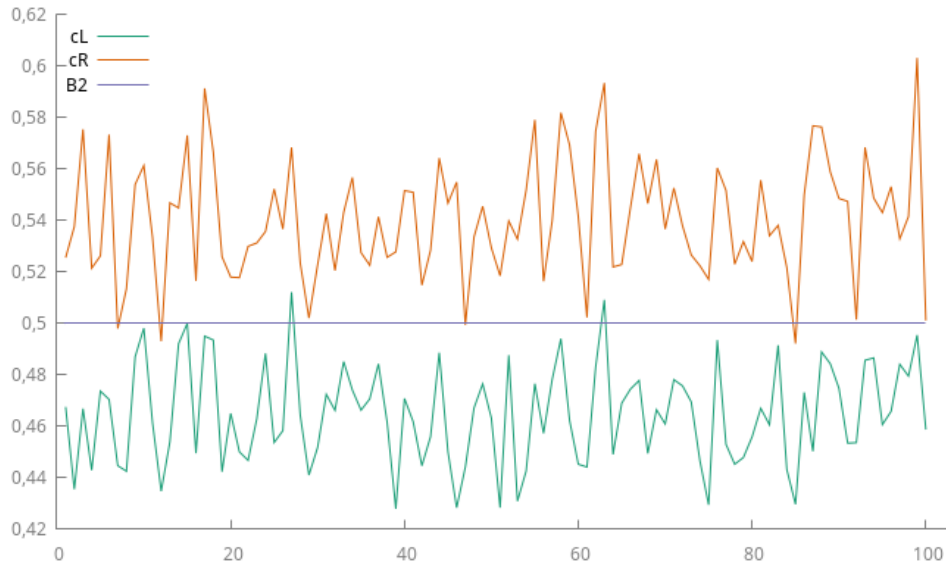


Figura 1: Los cien primeros intervalos de confianza al 95 %: extremo inferior y superior de cada uno, frente al verdadero $\beta_2 = 0,5$ (línea horizontal). Los intervalos que no lo contienen son aquellos en los que una de las dos líneas cruza la horizontal.

Actividad 2 - El precio de estimar σ

La lección 12 explicó por qué el estadístico sigue una t_{n-k} y no una normal: en el denominador va s , que fluctúa de muestra en muestra, y eso ensancha las colas. Con $n - k = 10$ la diferencia entre $t_c = 2,228$ y $1,96$ es del 14 %. Ya hemos guardado, en cada réplica, si el intervalo construido con $1,96$ contenía al verdadero β_2 ; y tenemos los 4000 estadísticos t bajo la hipótesis cierta, cuya distribución podemos mirar directamente. *en línea de comandos:*

```
printf "Cobertura del intervalo con t_c = %.3f : %.4f\n", tc, mean(dentro)
printf "Cobertura del intervalo con 1,96 : %.4f\n", mean(dentroZ)
printf "\nCuantiles empiricos del estadistico t bajo la H0 cierta:\n"
printf " 2,5%% -> %.3f (la t_10 predice %.3f; la normal, %.3f)\n", quantile(tstat, 0.025), -tc, -zc
printf " 97,5%% -> %.3f (la t_10 predice %.3f; la normal, %.3f)\n", quantile(tstat, 0.975), tc, zc
printf "Fraccion de |t| > 1,96 : %.4f (con 1,96 se rechazaria la H0 cierta mas del 5%)\n",
↪ mean(abs(tstat) > zc)
freq tstat --plot=display
```

```
Cobertura del intervalo con t_c = 2,228 : 0,9520
Cobertura del intervalo con 1,96 : 0,9188
```

```
Cuantiles empiricos del estadistico t bajo la H0 cierta:
```

```
 2,5% -> -2,176 (la t_10 predice -2,228; la normal, -1,960)
 97,5% -> 2,236 (la t_10 predice 2,228; la normal, 1,960)
```

```
Fraccion de |t| > 1,96 : 0,0812 (con 1,96 se rechazaria la H0 cierta mas del 5%)
```

Lo que debe observar

Con $1,96$ el intervalo cubre bastante menos del 95 %, y el contraste rechaza la hipótesis cierta bastante más del 5 %. Los cuantiles empíricos del estadístico lo confirman: se sitúan alrededor de $\pm 2,23$, no de $\pm 1,96$. Esa diferencia es lo que la lección llamó “el precio de estimar σ ”: el s del denominador fluctúa, el cociente se dispersa más que una normal, y la t_{n-k} lo tiene en cuenta. Con diez grados de libertad el precio se paga; con quinientos, como en la práctica A, t_c y $1,96$ son casi lo mismo.

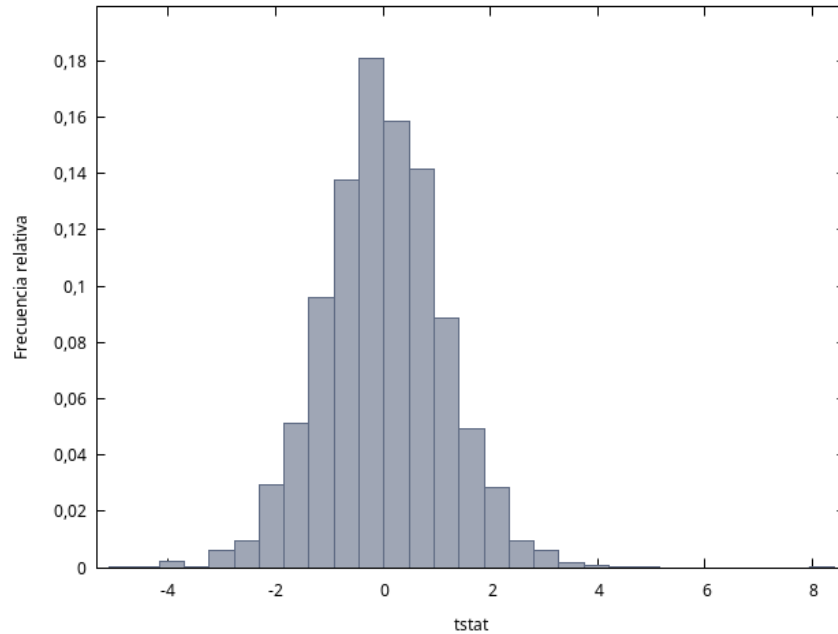


Figura 2: Histograma de los 4000 estadísticos $t = (\hat{\beta}_2 - 0,5)/ee(\hat{\beta}_2)$, calculados bajo la hipótesis cierta. Sus cuantiles del 2,5 % y del 97,5 % caen donde los sitúa la t_{10} , no donde los sitúa la normal.

Actividad 3 - Una hipótesis falsa por poco, con $n = 12$ y con $n = 120$

Hasta aquí hemos contrastado una hipótesis cierta. Ahora contrastamos una *falsa por poco*: $H_0: \beta_2 = 0,47$, cuando el verdadero es 0,5. ¿Con qué frecuencia la detecta el contraste? Lo hacemos dos veces: con las doce observaciones de siempre, y con ciento veinte —el mismo patrón de regresor repetido diez veces, el mismo σ , la misma hipótesis—. Lo único que cambia es la cantidad de evidencia.

en línea de comandos:

```

nulldata 120 --preserve
setobs 1 1 --cross-section
series x = 22.5 + 5*((index - 1) % 12)      # el mismo patron, diez veces
scalar b1 = 10                             # los parametros del mundo, de nuevo
scalar b2 = 0.5
scalar sig = 1.1
scalar R = 4000
scalar c_falsa = 0.47

```

```

smpl 1 12                                # una sola vuelta del patron: n = 12
scalar tc = critical(t, $nobs - 2, 0.025)
set seed 20261113
scalar rech = 0
scalar cov = 0
loop R --quiet
    series u = normal(0, sig)
    series y = b1 + b2*x + u
    ols y const x --quiet
    rech += (abs(($coeff(x) - c_falsa)/$stderr(x)) > tc)
    cov += ($coeff(x) - tc*$stderr(x) <= b2 && b2 <= $coeff(x) + tc*$stderr(x))
endloop
printf "n = %3d:  t_c = %.3f  rechazos de la H0 falsa (0,47) = %.3f  cobertura del IC = %.3f\n", $nobs, tc, rech/R, cov/R

```

```

smpl full                                # las diez vueltas: n = 120
scalar tc = critical(t, $nobs - 2, 0.025)
set seed 20261113
scalar rech = 0

```

```

scalar cov = 0
loop R --quiet
  series u = normal(0, sig)
  series y = b1 + b2*x + u
  ols y const x --quiet
  rech += (abs(($coeff(x) - c_falsa)/$stderr(x)) > tc)
  cov += ($coeff(x) - tc*$stderr(x) <= b2 && b2 <= $coeff(x) + tc*$stderr(x))
endloop
printf "n = %3d:  t_c = %.3f  rechazos de la H0 falsa (0,47) = %.3f  cobertura del IC = %.3f\n", $nobs, tc, rech/R, cov/R

```

```

n = 12:  t_c = 2,228  rechazos de la H0 falsa (0,47) = 0,319  cobertura del IC = 0,952
n = 120:  t_c = 1,980  rechazos de la H0 falsa (0,47) = 0,999  cobertura del IC = 0,946

```

Lo que debe observar

La cobertura del intervalo es la misma en los dos casos —en torno al 95%—, como debe: el intervalo cumple su garantía con doce datos y con ciento veinte. Lo que cambia es la *detección*: con $n = 12$ la hipótesis falsa se rechaza menos de una vez de cada tres; con $n = 120$, prácticamente siempre.

Esto es la identidad de la lección 12, $|t| = \sqrt{n-k} \times$ razón de catetos, ejecutada en vez de enunciada. La relación entre x e y es la misma en los dos mundos, el ruido es el mismo, la hipótesis es igual de falsa; lo único que ha crecido es $\sqrt{n-k}$, y con ello la frecuencia con la que un efecto real se hace visible. (Los manuales llaman *potencia* del contraste a esa frecuencia de detección.)

Y tiene una lectura inmediata para la práctica A, donde con doce viviendas el intervalo iba de 1 900 a 20 000: con doce observaciones, dos de cada tres veces *no se rechaza* una hipótesis que es falsa. “No rechazar” no era “aceptar”: era, la mayoría de las veces, no tener evidencia suficiente para decidir.

Apéndice - ¿Y si las perturbaciones no son normales?

Material complementario, con el mismo estatus que el apéndice matricial del laboratorio anterior: va más allá de lo que da tiempo en clase y queda aquí para quien quiera completarlo.

La lección 12 fue explícita: la normalidad es el único supuesto nuevo, y su cometido es dar *forma* a la distribución de $\hat{\beta}_2$. El laboratorio anterior mostró que insesgadez y varianza no la necesitan; la pregunta natural es si la cobertura del intervalo —que sí depende de la forma— la necesita. Repetimos la Actividad 1 con perturbaciones χ^2 de un grado de libertad, centradas y reescaladas a la misma σ : una distribución muy asimétrica, la misma con la que el laboratorio anterior obtuvo un histograma inconfundiblemente torcido con $n = 12$.

en línea de comandos:

```

smpl 1 12
scalar tc = critical(t, $nobs - 2, 0.025)
set seed 20261113
scalar cov = 0
scalar rech0 = 0
loop R --quiet
  series u = (randgen(X, 1) - 1)/sqrt(2) * sig  # chi2(1) centrada, con desviacion tipica sig
  series y = b1 + b2*x + u
  ols y const x --quiet
  cov += ($coeff(x) - tc*$stderr(x) <= b2 && b2 <= $coeff(x) + tc*$stderr(x))
  rech0 += (abs(($coeff(x) - b2)/$stderr(x)) > tc)
endloop
printf "chi2(1) centrada, n = %3d:  cobertura del IC = %.4f  rechazos de la H0 cierta = %.4f\n", $nobs, cov/R, rech0/R

```

```

smpl full
scalar tc = critical(t, $nobs - 2, 0.025)
set seed 20261113
scalar cov = 0
scalar rech0 = 0
loop R --quiet
  series u = (randgen(X, 1) - 1)/sqrt(2) * sig

```

```

series y = b1 + b2*x + u
ols y const x --quiet
cov += ($coeff(x) - tc*$stderr(x) <= b2 && b2 <= $coeff(x) + tc*$stderr(x))
rech0 += (abs(($coeff(x) - b2)/$stderr(x)) > tc)
endloop
printf "chi2(1) centrada, n = %3d:  cobertura del IC = %.4f  rechazos de la H0 cierta = %.4f\n", $nobs, cov/R, rech0/R

```

```

chi2(1) centrada, n = 12:  cobertura del IC = 0,9620  rechazos de la H0 cierta = 0,0380
chi2(1) centrada, n = 120:  cobertura del IC = 0,9478  rechazos de la H0 cierta = 0,0523

```

Lo que debe observar

Con $n = 12$ la cobertura y la tasa de rechazo se apartan de sus valores nominales, y lo hacen en la dirección que menos se espera: el intervalo cubre *más* del 95 % y el contraste rechaza *menos* del 5 %. El aparato de la lección 12, construido bajo normalidad, se vuelve conservador con estas perturbaciones y esta muestra. No es un desastre —nadie saca conclusiones falsas por rechazar de menos—, pero el 5 % que imprime el programa ya no es el 5 % real.

Con $n = 120$ la desviación prácticamente desaparece. Es el cierre de la lección 12: en muestras grandes, el teorema central del límite devuelve a $\hat{\beta}_2$ una distribución aproximadamente normal aunque U no lo sea, y los contrastes e intervalos vuelven a ser válidos sin el supuesto. En muestras pequeñas, la normalidad es el supuesto que sostiene el resultado exacto —y cuando falla, lo que falla es la calibración del nivel.

Preguntas de interpretación para la clase

1. En la Actividad 1, la frecuencia de cobertura no ha salido exactamente 0,95. ¿Qué esperaría que ocurriese con la diferencia si el experimento se repitiera con 400 000 réplicas en lugar de 4 000? ¿Y si se repitiera con otras 4 000, cambiando la semilla?
2. Escoja uno de los intervalos de la figura que *no* contiene a 0,5. Si ese fuera el único que tuviera —como ocurre siempre con datos reales—, ¿qué diría de él? ¿Habría alguna forma de saber que es de los que fallan?
3. En la Actividad 2, ¿por qué usar 1,96 con $n - k = 10$ hace que el contraste rechace la hipótesis cierta *más* del 5 % de las veces, y no menos?
4. En la Actividad 3, la cobertura es la misma con doce y con ciento veinte observaciones, pero la detección no. ¿Qué propiedad del intervalo se conserva al cambiar n , y cuál cambia? Relaciónelo con la anchura de los intervalos de la práctica A.
5. Tras el Apéndice: si en un estudio con muestra pequeña las perturbaciones fueran claramente asimétricas, ¿en qué sentido habría que tomar con cautela un valor p de 0,04? ¿Y uno de 0,0001?

Para profundizar

- Lección 12 ([lección 12](#)), secciones “El intervalo de confianza” y “Por qué el valor crítico depende de $n - k$ ”.
- Wooldridge, J. M. (2020). *Introductory Econometrics*, cap. 4, sección 4.3 (intervalos de confianza) y cap. 5, sección 5.2 (validez aproximada de los contrastes en muestras grandes sin normalidad).
- Manual de Gretl, capítulo sobre *loop constructs*; funciones `quantile` y `randgen`.

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

nulldata 12
setobs 1 1 --cross-section
series x = 22.5 + 5*(index - 1)           # el regresor disperso de la lección 12
scalar b1 = 10
scalar b2 = 0.5
scalar sig = 1.1
scalar R = 4000
scalar tc = critical(t, 10, 0.025)        # valor crítico de la t_10 al 5%
scalar zc = 1.96                         # el de la normal, para la Actividad 2

set seed 20261113
loop R --progressive --quiet
  series u = normal(0, sig)
  series y = b1 + b2*x + u
  ols y const x --quiet
  scalar b2hat = $coeff(x)
  scalar ee = $stderr(x)
  scalar cL = b2hat - tc*ee                # extremos del IC al 95%
  scalar cR = b2hat + tc*ee
  scalar dentro = (cL <= b2 && b2 <= cR)    # 1 si el IC contiene al verdadero beta_2
  scalar tstat = (b2hat - b2)/ee           # t para la H0 CIERTA beta_2 = 0,5
  scalar dentroZ = (b2hat - zc*ee <= b2 && b2 <= b2hat + zc*ee) # con 1,96 en vez de t_c
  store "@workdir/replicasIC.gdt" b2hat ee cL cR dentro tstat dentroZ
endloop
open "@workdir/replicasIC.gdt" --quiet
scalar tc = critical(t, 10, 0.025)
scalar zc = 1.96
outfile --quiet cobertura.txt
  printf "Replicas" : %d\n", $nobs
  printf "Intervalos que contienen al verdadero b2" : %.4f (objetivo: 0,95)\n", mean(dentro)
  printf "Rechazos de la H0 cierta al 5%" : %.4f (objetivo: 0,05)\n", mean(abs(tstat) > tc)
end outfile

smpl 1 100
series B2 = 0.5
gnuplot cL cR B2 --time-series --with-lines --output="@workdir/baile.png"
outfile --quiet baile.txt
  printf "Intervalos que fallan entre los 100 primeros: %d\n", 100 - sum(dentro)
end outfile
smpl full

outfile --quiet precioSigma.txt
```

```

printf "Cobertura del intervalo con t_c = %.3f : %.4f\n", tc, mean(dentro)
printf "Cobertura del intervalo con 1,96      : %.4f\n", mean(dentroZ)
printf "\nCuantiles empiricos del estadistico t bajo la H0 cierta:\n"
printf "  2,5%% -> %.3f      (la t_10 predice %.3f; la normal, %.3f)\n", quantile(tstat, 0.025), -tc,
↳ -zc
printf "  97,5%% -> %.3f      (la t_10 predice %.3f; la normal, %.3f)\n", quantile(tstat, 0.975),
↳ tc, zc
printf "Fraccion de |t| > 1,96      : %.4f      (con 1,96 se rechazaria la H0 cierta mas del
↳ 5%%)\n", mean(abs(tstat) > zc)
end outfile
freq tstat --plot="@workdir/histT.png"

nulldata 120 --preserve
setobs 1 1 --cross-section
series x = 22.5 + 5*((index - 1) % 12)      # el mismo patron, diez veces
scalar b1 = 10                             # los parametros del mundo, de nuevo
scalar b2 = 0.5
scalar sig = 1.1
scalar R = 4000
scalar c_falsa = 0.47
outfile --quiet deteccion.txt
  smpl 1 12                                # una sola vuelta del patron: n = 12
  scalar tc = critical(t, $nobs - 2, 0.025)
  set seed 20261113
  scalar rech = 0
  scalar cov = 0
  loop R --quiet
    series u = normal(0, sig)
    series y = b1 + b2*x + u
    ols y const x --quiet
    rech += (abs(($coeff(x) - c_falsa)/$stderr(x)) > tc)
    cov += ($coeff(x) - tc*$stderr(x) <= b2 && b2 <= $coeff(x) + tc*$stderr(x))
  endloop
  printf "n = %3d:  t_c = %.3f  rechazos de la H0 falsa (0,47) = %.3f  cobertura del IC = %.3f\n",
↳ $nobs, tc, rech/R, cov/R
  smpl full                                # las diez vueltas: n = 120
  scalar tc = critical(t, $nobs - 2, 0.025)
  set seed 20261113
  scalar rech = 0
  scalar cov = 0
  loop R --quiet
    series u = normal(0, sig)
    series y = b1 + b2*x + u
    ols y const x --quiet
    rech += (abs(($coeff(x) - c_falsa)/$stderr(x)) > tc)
    cov += ($coeff(x) - tc*$stderr(x) <= b2 && b2 <= $coeff(x) + tc*$stderr(x))
  endloop
  printf "n = %3d:  t_c = %.3f  rechazos de la H0 falsa (0,47) = %.3f  cobertura del IC = %.3f\n",
↳ $nobs, tc, rech/R, cov/R
end outfile

outfile --quiet apendice.txt
  smpl 1 12
  scalar tc = critical(t, $nobs - 2, 0.025)
  set seed 20261113
  scalar cov = 0
  scalar rech0 = 0
  loop R --quiet
    series u = (randgen(X, 1) - 1)/sqrt(2) * sig      # chi2(1) centrada, con desviacion tipica sig
    series y = b1 + b2*x + u
    ols y const x --quiet
    cov += ($coeff(x) - tc*$stderr(x) <= b2 && b2 <= $coeff(x) + tc*$stderr(x))
    rech0 += (abs(($coeff(x) - b2)/$stderr(x)) > tc)
  endloop
  printf "chi2(1) centrada, n = %3d:  cobertura del IC = %.4f  rechazos de la H0 cierta = %.4f\n",
↳ $nobs, cov/R, rech0/R
  smpl full

```

```

scalar tc = critical(t, $nobs - 2, 0.025)
set seed 20261113
scalar cov = 0
scalar rech0 = 0
loop R --quiet
    series u = (randgen(X, 1) - 1)/sqrt(2) * sig
    series y = b1 + b2*x + u
    ols y const x --quiet
    cov += ($coeff(x) - tc*$stderr(x) <= b2 && b2 <= $coeff(x) + tc*$stderr(x))
    rech0 += (abs(($coeff(x) - b2)/$stderr(x)) > tc)
endloop
printf "chi2(1) centrada, n = %3d:  cobertura del IC = %.4f  rechazos de la H0 cierta = %.4f\n",
↪ $nobs, cov/R, rech0/R
end outfile

```

Enlace al guión: [S18-Prct-C-cobertura.inp](#)

Respuestas

1. *Más réplicas, otra semilla.* La frecuencia de cobertura observada es la media de R indicadores 0/1 independientes, cada uno con probabilidad 0,95 de valer 1; su desviación típica es $\sqrt{0,95 \cdot 0,05/R}$, unos 0,0034 con $R = 4000$ y diez veces menos con $R = 400\,000$. Cabe esperar, por tanto, que la distancia a 0,95 se reduzca a la décima parte. Con otra semilla y las mismas 4000 réplicas, la frecuencia sería otra, distinta de la obtenida en el mismo orden de magnitud (unas milésimas): la propia frecuencia empírica tiene su dispersión.
2. *Un intervalo que falla, visto en solitario.* Nada lo distingue de los que aciertan: tiene el mismo aspecto, la misma anchura aproximada, se ha construido con la misma receta. Con datos reales no hay forma de saber si el intervalo que tenemos delante es uno de los del 5 %. Eso es lo que significa que la garantía sea sobre el procedimiento y no sobre el intervalo concreto: sabemos cuántas veces acierta el método, no si ha acertado esta vez.
3. *Por qué 1,96 rechaza de más.* Porque $1,96 < 2,228$: la región crítica “ $|t| > 1,96$ ” es más grande que “ $|t| > 2,228$ ”, y el estadístico —que sigue una t_{10} , con colas más gruesas que la normal— cae en ella con más frecuencia que el 5 %. Usar el valor crítico de la normal es actuar como si σ fuese conocida; como no lo es y se estima, el estadístico se dispersa más de lo que la normal supone, y el contraste, calibrado para una normal, se vuelve demasiado laxo.
4. *Lo que se conserva y lo que cambia con n .* Se conserva la cobertura: sea cual sea n , el intervalo al 95 % contiene al parámetro el 95 % de las veces, porque su valor crítico se ajusta a los grados de libertad. Cambia la *anchura*: con más observaciones el error estándar es menor (el regresor en desviaciones es más largo) y el intervalo, más estrecho; por eso un valor falso por poco queda fuera de él con más frecuencia. Es lo que se veía en la práctica A al pasar de doce viviendas a quinientas: la misma garantía del 95 %, con un intervalo diez veces más estrecho y, en consecuencia, mucho más informativo.
5. *Valores p con muestra pequeña y perturbaciones asimétricas.* El Apéndice muestra que, en ese caso, el nivel real del contraste no coincide con el nominal: aquí el contraste resultó conservador, pero con otras distribuciones podría ser lo contrario, y en cualquier caso el 0,05 del programa deja de ser una cifra fiable. Un valor p de 0,04 está tan cerca del umbral que esa descalibración puede cambiar el veredicto: hay que tomarlo como “en el límite”, no como “significativo”. Uno de 0,0001 está tan lejos del umbral que ninguna descalibración razonable lo convierte en no significativo: la conclusión es robusta aunque la cifra exacta no lo sea. La regla práctica es que la cautela hace falta cerca del umbral, y que con muestras grandes el problema se desvanece.