INTRODUCCIÓN A LA ECONOMETRÍA — GRADO ECO UNED
CORRECCIÓN DEL EXAMEN FEBRERO 2026 S1

Primera parte

1. Un elemento esencial de econometría son los datos. Señale qué tipos de datos usamos en la investigación econométrica:

a) Series temporales, datos de sección cruzada y datos de panel ✓

b) Pueden usarse datos cuantitativos y cualitativos ✓

c) Datos de observación y de experimentación ✓

d) Todas las anteriores son correctas ✓

Las tres alternativas son clasificaciones distintas y compatibles del mismo conjunto de datos: por su estructura (series temporales, sección cruzada y panel), por su naturaleza (cuantitativos y cualitativos, estos últimos incorporados mediante variables ficticias) y por su procedimiento de obtención (datos de observación, los habituales en economía, y datos de experimentación).

---

2. Señale cuál de las siguientes expresiones proporciona el estimador MCO de la pendiente, $\hat{\beta}_1$, en un modelo de regresión simple:

a) $\frac{\sum x_i y_i}{\sum x_i^2}$ ✓

b) $\frac{\sum x_i y_i}{\sum y_i^2}$

c) $\frac{\sum x_i y_i}{(\sum x_i)^2}$

d) Ninguna de las anteriores es correcta

Minimizando $\sum(Y_i - \hat{\alpha} - \hat{\beta}_1 X_i)^2$ se obtiene:

$$\hat{\beta}_1 = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sum(X_i - \bar{X})^2} = \frac{\sum x_i y_i}{\sum x_i^2}$$ de Y sobre X

donde las minúsculas son desviaciones respecto de la media. La opción b) es la pendiente de la regresión inversa, la de X sobre Y. La c) no está definida, porque $\sum x_i = 0$ por construcción.

Usando la notación estándar donde las minúsculas representan desviaciones respecto a la

media ($x_i = X_i - \bar{X}$, $y_i = Y_i - \bar{Y}$):

El enunciado NO especifica si se trata de la regresión de Y sobre X o de X sobre Y:

- Si el modelo es $Y = \alpha + \beta_1 X + \varepsilon \rightarrow \hat{\beta}_1 = \Sigma x_i y_i / \Sigma x_i^2$ (sería opción a)
- Si el modelo es $X = \alpha + \beta_1 Y + \varepsilon \rightarrow \hat{\beta}_1 = \Sigma x_i y_i / \Sigma y_i^2$ (sería opción b)

3. Considere el modelo $Y_i = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{2i}^2 + \varepsilon_i$. Señale la afirmación correcta:

a) El efecto sobre Y de una variación unitaria en X_{2i} es constante en toda la función

b) El efecto sobre Y de una variación unitaria en X_{1i} es constante en toda la función ✓

c) Si $\beta_2 > \beta_1$, el efecto sobre Y de una variación unitaria en X_{2i} siempre será mayor que el efecto de una variación unitaria en X_{1i}

d) Ninguna de las anteriores

Los efectos marginales son:

$$\frac{\partial Y_i}{\partial X_{1i}} = \beta_1 \qquad \frac{\partial Y_i}{\partial X_{2i}} = \beta_2 + 2\beta_3 X_{2i}$$

El de X_1 es constante; el de X_2 depende del nivel de X_2, porque la relación es cuadrática. Eso descarta a) y c); además, comparar coeficientes de variables medidas en unidades distintas carece de sentido.

---

4. En la regresión simple $Y_i = \alpha + \beta X_i + \varepsilon_i$:

a) $R^2 = \overline{R}^2$

b) $R^2 < \overline{R}^2$

c) $R^2 > \overline{R}^2$ ✓

d) $\overline{R}^2 = 0$

El coeficiente corregido penaliza los grados de libertad:

$$\overline{R}^2 = 1 - (1 - R^2)\frac{n-1}{n-k-1}$$

En la regresión simple $k = 1$, de modo que el factor $(n-1)/(n-2)$ es mayor que uno y $\overline{R}^2 < R^2$. Solo coincidirían en el caso límite de ajuste perfecto, $R^2 = 1$.

---

5. En un modelo de regresión con tres variables explicativas más la constante, la expresión $\sigma_\varepsilon^2(\mathbf{X}'\mathbf{X})^{-1}$,

a) Es una matriz de dimensiones (n x 4)

b) Es una matriz de dimensiones (4 x 4) ✓

c) Es una matriz de dimensiones (3 x 3)

d) Es un escalar

$Var(\hat{\beta}) = \hat{\sigma}_\varepsilon^2 (X'X)^{-1} = \begin{pmatrix} Var(\hat{\beta}_0) & & \\ & \ddots & \\ & & Var(\hat{\beta}_3) \end{pmatrix}$ 4x4

Es la matriz de varianzas y covarianzas de $\widehat{\boldsymbol{\beta}}$. Con tres variables más la constante hay 4 parámetros, luego $\mathbf{X}$ es $(n \times 4)$ y $\mathbf{X}'\mathbf{X}$ y su inversa son (4×4). Multiplicar por el escalar σ_ε^2 no altera la dimensión. En su diagonal figuran las varianzas de los cuatro estimadores, cuyas raíces cuadradas son los errores estándar.

---

6. La razón por la que el método de *Mínimos Cuadrados Ponderados* no suele emplearse en la práctica es:

a) El proceso de cálculo es muy complejo

b) Exige que las variables explicativas no estén autocorrelacionadas

c) No se conoce la varianza del término de error ✓

d) Ninguna de las anteriores

$\sigma_i^2 = f(x)$

MCP corrige la heterocedasticidad dividiendo toda la ecuación por σ_i, de forma que el nuevo error tenga varianza constante. El problema es que σ_i^2 es un parámetro poblacional desconocido, por lo que en la práctica se recurre a los MCG factibles o, más habitualmente, a los errores estándar robustos de White. El cálculo, en cambio, es sencillo, y la autocorrelación afecta al error, no a las explicativas.

---

7. Con datos de 88 países se ha estimado (errores estándar entre paréntesis):

$$\log(Y_i) = 0.54 - 0.66X_{1i} + 1.82\log(X_{2i}) - 0.12D_i$$
$$(0.1) \quad (0.23) \quad (0.71) \quad (0.04)$$

$H_0: \beta_D = 0 \quad t = \frac{-0.12}{0.04} = -3$ RH0

donde Y es el número de muertos por millón de habitantes en accidentes de tráfico, X_1 el impuesto sobre bebidas alcohólicas, X_2 la renta per cápita y D una dummy que toma el valor 1 si el país tiene una legislación que prohíbe la venta de bebidas alcohólicas a los menores de 20 años. Según la estimación anterior, el efecto de esta legislación:

a) Disminuye en 12 los muertos por millón en accidente de tráfico

b) No tiene ningún efecto porque el coeficiente no es estadísticamente distinto de cero F

c) Disminuye en un 12% el número de muertos por millón de habitantes ✓

d) Ninguna de las anteriores

tiene efecto

Di => log-lin

El estadístico es $t = -0.12/0.04 = -3$, superior en valor absoluto a 1.99 (valor crítico con 84 grados de libertad), luego el coeficiente sí es significativo y se descarta la opción b).

Al estar la dependiente en logaritmos, el coeficiente mide un cambio relativo y no un cambio en unidades de Y: los países con esa legislación tienen en torno a un 12% menos de muertos, lo que descarta también la opción a).

Nota. El efecto exacto es $100(e^{-0.12} - 1) = -11.31\%$; el 12% es la aproximación habitual y, entre las alternativas ofrecidas, c) sigue siendo la mejor.

---

8. Observe los modelos:

$$\log Y_i = \beta_0 + \beta_1 \log X_{1i} + \beta_2 \log X_{2i} + \beta_3 \log X_{3i} + \varepsilon_i$$

$$(2) \quad \log\left(\frac{Y_i}{X_{1i}}\right) = \beta_0 + \beta_2 \log X_{2i} + \beta_3 \log X_{3i} + \varepsilon_i$$

$H_0: \beta_1 = 1 \rightarrow MR\ (2)$

y señale la respuesta correcta

a) El segundo modelo está anidado en el primero ✓

b) El primer modelo está anidado en el segundo

c) Los dos modelos son exactamente iguales

d) Ninguna de las anteriores

Como $\log(Y_i/X_{1i}) = \log Y_i - \log X_{1i}$, el segundo modelo equivale a:

$$\log Y_i = \beta_0 + 1 \cdot \log X_{1i} + \beta_2 \log X_{2i} + \beta_3 \log X_{3i} + \varepsilon_i$$

Es decir, es el primero con la restricción $\beta_1 = 1$ impuesta: el modelo restringido está anidado en el general. La validez de la restricción puede contrastarse con un test F de una restricción o con un t sobre $\beta_1 = 1$.

---

9. Para contrastar si el modelo está correctamente especificado se usa,

a) El test de Davidson y McKinnon

b) El Reset test de Ramsey ✓

c) El test de Mizon y Richard

d) Todos ellos

El RESET test de Ramsey (Regression Equation Specification Error Test) es el test específico para contrastar si la forma funcional del modelo está correctamente especificada.

El test añade potencias de los valores ajustados ($\hat{Y}^2$, $\hat{Y}^3$, ...) como regresores adicionales y contrasta su significatividad conjunta.

H_o: El modelo está correctamente especificado (los términos adicionales no son significativos)

H_1: El modelo tiene errores de especificación

El test de Davidson y McKinnon (J-test) se utiliza para comparar modelos NO anidados.

El test de Mizon y Richard es un test de especificación más general pero menos utilizado en la práctica.

RESET: $Y_i = \beta_o + \beta_1 X_{1i} + ... + \beta_k X_{ki} + \delta_1 \hat{Y}_i^2 + \delta_2 \hat{Y}_i^3 + v_i$

H_o: $\delta_1 = \delta_2 = 0$ (especificación correcta)

10. En general el sesgo debido a variables medidas con error se produce cuando:

a) Alguna variable explicativa está medida con error ✓

b) La variable dependiente está medida con error

c) Solo si la variable dependiente y alguna de las independiente están medidas con error

d) Siempre está presente dado que en economía dado que no hay variables económicas medidas sin error

Si el error afecta a la dependiente, se absorbe en la perturbación y MCO sigue siendo insesgado: solo pierde precisión. Si afecta a una explicativa, $X_i^* = X_i + u_i$, el regresor pasa a estar correlacionado con el término de error y el estimador resulta sesgado e inconsistente:

$$\text{plim}\hat{\beta} = \beta \cdot \frac{\sigma_X^2}{\sigma_X^2 + \sigma_u^2}$$

El factor es menor que uno: es el sesgo de atenuación, que se corrige con variables instrumentales. La opción d) confunde la existencia de errores de medida con la existencia de sesgo.

PLANTILLA DE SOLUCIONES – TEST

1	2	3	4	5	6	7	8	9	10
d	a	b	c	b	c	c	a	b	a

Parte práctica

El Ministerio de Educación de un país hipotético está interesado en conocer cómo afecta el número de estudiantes por profesor a los resultados académicos. Con este objetivo recaba datos de 420 centros en todo el territorio y estima la ecuación:

$$\hat{Y}_i = 70 - 2.3X_{1i}, \quad n = 420, \quad R^2 = 0.051$$

donde la variable dependiente es el resultado de un test estandarizado que mide las competencias de los estudiantes en lengua y matemáticas (varía entre 0 y 100 puntos) y X_{1i} es el número de alumnos por professor

a) ¿Es el signo del coeficiente de X acorde con lo esperado? ¿Qué efecto tendría una reducción de 5 unidades en el número de estudiantes por profesor?

El signo es el esperado: una ratio alumnos/profesor más alta supone menos atención por estudiante y, por tanto, peores resultados. El coeficiente estimado es negativo, $\hat{\beta}_1 = -2.3$.

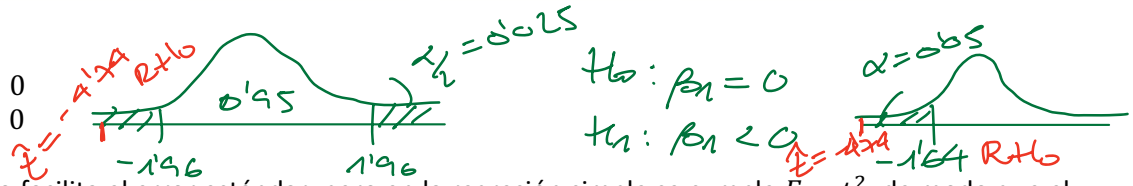
El efecto de reducir la ratio en 5 unidades es:

$$\Delta\hat{Y} = (-2.3) \times (-5) = +11.5 \text{ puntos}$$

La puntuación media prevista aumentaría en 11.5 puntos, manteniendo constante todo lo demás.

b) Valore la significatividad estadística de la variable explicativa indicando las hipótesis nula y alternativa que somete a contraste

H_0: $\beta_1 = 0$
H_1: $\beta_1 \neq 0$



El enunciado no facilita el error estándar, pero en la regresión simple se cumple $F = t^2$, de modo que el estadístico puede obtenerse a partir del R^2:

$$F = \frac{R^2/k}{(1-R^2)/(n-k-1)} = \frac{0.051}{0.949/418} = 22.46 \qquad |t| = \sqrt{22.46} = 4.74$$

Con 418 grados de libertad el valor crítico al 5% es 1.96. Como $4.74 > 1.96$, **se rechaza** H_0: la ratio alumnos/profesor es estadísticamente significativa.

c) Explique qué problemas plantea el resultado anterior y si considera que es apropiado para tomar una decisión sobre la reducción de la ratio alumnos/professor

No es apropiado, por tres razones:

El $R^2 = 0.051$ indica que la ratio explica apenas el 5.1% de la variabilidad de los resultados.

Se omiten variables relevantes (renta familiar, idioma, recursos del centro) que están correlacionadas con la ratio, lo que sesga el estimador.

Los datos son de observación, no experimentales, de modo que la significatividad no permite hablar de causalidad.

La segunda estimación lo confirma: al incluir controles, el coeficiente pasa de -2.3 a -1.05 y el R^2 sube a 0.78.

Considere ahora la estimación:

$$\hat{Y}_i = \underset{\substack{(4.7)\\[7.4]}}{66} - \underset{\substack{(0.24)\\[0.33]}}{1.05X_{1i}} + \underset{\substack{(0.002)\\[0.016]}}{0.003X_{2i}} - \underset{\substack{(0.03)\\[0.04]}}{0.14X_{3i}} - \underset{\substack{(0.02)\\[0.04]}}{0.54X_{4i}}, \quad n = 420, \quad R^2 = 0.78, \quad \textbf{White test} = 40.2$$

donde las nuevas variables son X_2 = dummy = 1 si el estudiante dispone de ordenador personal,

X_3 = porcentaje de estudiantes en el centro que tienen problemas con el idioma por ser emigrantes, X_4 = porcentaje de estudiantes que tienen beca de comedor.

Entre paréntesis, los errores estándar y entre corchetes los errores estándar robustos.

d) Valore si los signos de los coeficientes son razonables y contraste la significatividad de las variables. Indique cuál de las versiones de los errores estándar ha de emplearse y porqué.

Primero hay que decidir qué errores estándar emplear, con el contraste de White:

H_0: homocedasticidad
H_1: heterocedasticidad

El estadístico, 40.2, supera el valor crítico de la χ^2 al 5% (entre 15.5 y 23.7 según se incluyan o no los productos cruzados), de modo que **se rechaza** H_0: hay heterocedasticidad. MCO sigue siendo insesgado, pero los errores estándar convencionales no son válidos, así que deben emplearse los **robustos**, los que figuran entre corchetes.

Regresión auxiliar del test de White:

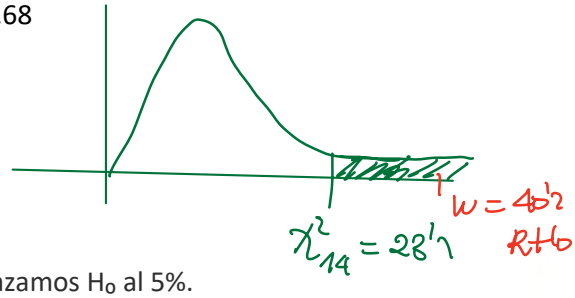
Se regresa el cuadrado de los residuos sobre las variables originales, sus cuadrados y sus interacciones 2 a 2:

$\hat{u}^2 = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3 + \alpha_4 X_4 + \alpha_5 X_1^2 + \alpha_6 X_2^2 + \alpha_7 X_3^2 + \alpha_8 X_4^2 + \alpha_9 X_1 X_2 + \alpha_{10} X_1 X_3 + \alpha_{11} X_1 X_4 + \alpha_{12} X_2 X_3 + \alpha_{13} X_2 X_4 + \alpha_{14} X_3 X_4 + v$

• Total: 4 + 4 + 6 = 14 grados de libertad

Contraste: Estadístico White = 40.2

Valor crítico $\chi^2_{0.05,14} \approx 23.68$



Como 40.2 > 23.68, rechazamos H_0 al 5%.

Distribución chi-cuadrado (X^2)
(niveles de confianza)

G.L.	,995	,99	,975	,95	,90
1	7,88	6,63	5,02	3,84	2,71
2	10,6	9,21	7,38	5,99	4,61
3	12,8	11,3	9,35	7,81	6,25
4	14,9	13,3	11,1	9,49	7,78
5	16,7	15,1	12,8	11,1	9,24
6	18,5	16,8	14,4	12,6	10,6
7	20,3	18,5	16,0	14,1	12,0
8	22,0	20,1	17,5	15,5	13,4
9	23,6	21,7	19	16,9	14,7
10	25,2	23,2	20,5	18,3	16,0
11	26,8	24,7	21,9	19,7	17,3
12	28,3	26,2	23,3	21,0	18,5
13	29,8	27,7	24,7	22,4	19,8
14	31,3	29,1	26,1	23,7	21,1
15	32,8	30,6	27,5	25,0	22,3
16	34,3	32	28,8	26,3	23,5
17	35,7	33,4	30,2	27,6	24,8
18	37,2	34,8	31,5	28,9	26,0

Con 415 grados de libertad, el valor crítico al 5% es 1.96. Los signos y los contrastes son:

- X_1(alumnos por profesor): -1.05, signo negativo esperado.
  $t = -1.05/0.33 = -3.18$ → significativa. RHo
- X_2(ordenador personal): 0.003, signo positivo esperado pero magnitud despreciable.
  $t = 0.003/0.016 = 0.19$ → no significativa. No RHo
- X_3(problemas con el idioma): -0.14, signo negativo esperado.
  $t = -0.14/0.04 = -3.50$ → significativa. RHo
- X_4(beca de comedor, indicador de renta baja): -0.54, signo negativo esperado.
  $t = -0.54/0.04 = -13.50$ → significativa. RHo

Los cuatro signos son razonables y solo el ordenador personal deja de ser significativo una vez controlado el nivel socioeconómico del centro.

e) Contraste la significatividad global de la regression

H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$ ⟶ $q = 4 = k$
H_1: al menos un $\beta_j \neq 0$

$$F = \frac{R^2/k}{(1-R^2)/(n-k-1)} = \frac{0.78/4}{0.22/415} = 367.8$$

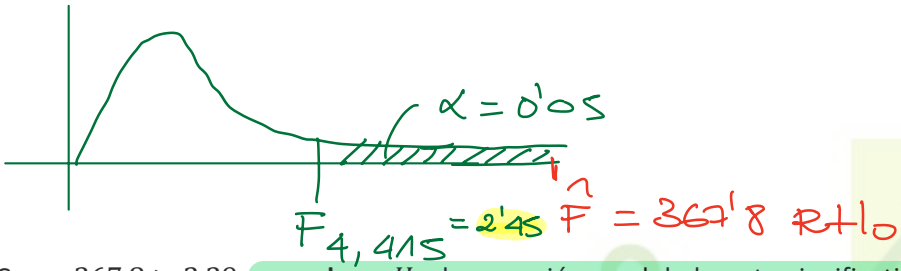
420 - 4 - 1

exacto

Distribución F de Snedecor (5%)

g.l.	Grados de libertad en el numerador									
	1	2	3	4	5	6	7	8	9	10
1	161,4	199,5	215,7	224,6	230,2	234,0	236,8	238,9	240,5	241,9
2	18,51	19,00	19,16	19,25	19,3	19,33	19,35	19,37	19,38	19,4
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,20
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19
29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22	2,18
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99
120	3,92	3,07	2,68	2,45	2,29	2,17	2,09	2,02	1,96	1,91

El valor crítico de la F con 4 y 415 grados de libertad al 5% es aproximadamente 2.39.



Como $367.8 > 2.39$, **se rechaza** H_0: la regresión es globalmente significativa, lo que resulta coherente con un R^2 de 0.78.