

Econometría I

Introducción

Ramiro de Elejalde

Facultad de Economía y Finanzas
Universidad Alberto Hurtado

Preliminares

Cuestiones domésticas

- **Profesor:** Ramiro de Elejalde
Email: ramirode@gmail.com
- **Horario de oficina:** No hay un horario fijo, enviar un correo para fijar una cita, o interactuar por el chat de Microsoft Teams.
- **Ayudante:** Carla Gómez Solano
Email: carlagomezs.11@outlook.com
- **Horario de la Ayudantía:** Lunes de 16:00 a 17:20.
- **Evaluación:** Tareas 15 %, 2 pruebas 40 % (20 % cada una), trabajo grupal 15 % y examen 30 %.

- Bibliografía

- Wooldridge, J. M. (2010), *Econometric Analysis of Cross Section and Panel Data*. MIT Press.
- Angrist, J. D. and J-S Pischke (2009), *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.
- Stock, J. H. and M.W. Watson (2006), *Introduction to econometrics, 2nd Edition*. Addison-Wesley.

1. Introducción
2. Esperanza condicional y modelo lineal de regresión
3. Teoría asintótica
4. Modelo de regresión lineal: Estimación por MCO
5. Modelo de regresión lineal: Inferencia por MCO
6. Temas adicionales de MCO
7. Variables instrumentales

Introducción

Efecto causal y experimentos ideales

Predicción versus causalidad

Estructura de datos

Enfoque del curso

Trabajo grupal

Referencias: Cap. 1 de Wooldridge, y cap. 1, 2 y 3 de Stock and Watson.

Introducción

Introducción

Efecto causal y experimentos ideales

La teoría economía propone relaciones relevantes, muchas veces con implicaciones de política, pero no cuantifica las magnitudes de los efectos causales.

- **Ejemplos:**

- ¿Cuál es el efecto de reducir el tamaño de clase en el rendimiento de los estudiantes?
- ¿Cuál es el efecto de tener hijos sobre la oferta laboral de las madres?
- ¿Cuál es el efecto de un aumento del precio sobre la cantidad demandada?

- **Predicción versus causalidad**

- ¿Cuál será la tasa de inflación en el mes siguiente en Chile?
- Ejemplo: Correlación no implica causalidad

<https://www.latercera.com/tendencias/noticia/>

[ninos-bajan-tres-decimas-notas-hora-uso-del-celular/123002/](https://www.latercera.com/tendencias/noticia/ninos-bajan-tres-decimas-notas-hora-uso-del-celular/123002/)

Experimento ideal

Experimentos aleatorios controlados (Randomized controlled trials o RCT)

- Grupo de tratamiento
- Grupo de control
- **Asignación aleatoria** al grupo de control o tratamiento
- Ejemplos: Medicina, agricultura y, en forma más reciente, en ciencias sociales (PROGRESA, Rand Health Insurance Experiment, Tennessee class size experiment).

Experimento ideal

- ¿Cómo diseñaría el experimento ideal para medir el efecto del tamaño de clase sobre el rendimiento educativo?

Datos observacionales versus experimentales

Datos experimentales

- Proviene de experimentos diseñados para evaluar el efecto causal de un tratamiento o política.

Datos observacionales (no experimentales)

- Obtenidos a través de censos, encuestas o datos administrativos.
- Fuentes de datos para Chile: Censo de población y vivienda (censos), CASEN, ENIA (encuestas), y SII, Aduana, DEMRE, Ministerio de Educación, Superintendencia de Salud, y Superintendencia de Entidades Financieras (datos administrativos).

Datos observacionales versus experimentales

Dificultades para medir un efecto causal con datos observacionales

- **Omisión de variable relevante**

Hay una variable no observada que afecta tanto al tratamiento como al resultado.

- **Simultaneidad o causalidad reversa**

El resultado afecta al tratamiento.

- **Error de medida** en alguna de las variables observadas

- **Selección en la muestra**

Disponibilidad de datos depende de un proceso de selección relacionado con la variable dependiente.

Predicción versus causalidad

Para realizar una predicción no es necesario medir una relación causal.

- **Predicción**

¿Cuál será la tasa de crecimiento del PIB en el próximo en Chile?

- **Causalidad**

¿Cuál es el efecto de la reforma tributaria en la tasa de crecimiento del PIB en el 2015 en Chile?

Introducción

Estructura de datos

Datos de corte transversal (cross-section)

- Datos para distintas unidades (personas, hogares, firmas, unidades gubernamentales, etc.) para un sólo período de tiempo.
- El orden de los datos es arbitrario.

Datos de corte transversal (cross-section)

- **Ejemplo:** Datos de 420 distritos educacionales en California, EEUU para 1998

observat	dist_code	enrl_tot	teachers	testscr
1	75119	195	10.9	17.9
2	61499	240	11.15	21.5
3	61549	1550	82.9	18.7
.
.
.
420	72751	1778	93.8	19.0

Datos de series de tiempo

- Son observaciones de una variable o varias variables a lo largo de varios periodos de tiempo (años, trimestres, meses, etc.).
- Ejemplos: precio de activos financieros y reales, series macroeconómicas (tipos de interés, tipos de cambio, tasa de inflación, PIB, etc.)
- Las observaciones no son independientes: la evolución temporal puede explotarse con fines predictivos.
- La frecuencia y orden con la que se observan los datos es muy importante.

- **Ejemplo:** Datos de inflación, empleo, desempleo y crecimiento para un país

year	inflation	unemployment	growth
1975	3.8	5.8	3.6
1976	5.4	6.4	2.8
1977	5.3	8.9	2.9
.	.	.	.
.	.	.	.
.	.	.	.
2013	1.3	6.3	1.5
2014	1.1	6.9	1.2

Datos de panel o longitudinales

- Consisten en una serie temporal por cada unidad de corte transversal. La longitud de las series temporales (T) suele ser mucho más corta que el número de unidades de corte transversal (N).
- Generalmente $N > T$
- Datos de panel \neq cortes transversales repetidos
- Aportan más información y permite responder a preguntas que los cortes transversales no pueden
- Permiten incluir una estructura temporal en el razonamiento económico

Datos de panel o longitudinales

- **Ejemplo:** Panel de firmas

firm	year	profit	employment	stock
1	1995	200	150	0
1	1996	3000	135	0
:	:	:	:	:
:	:	:	:	:
1	2000	4566	356	1
2	1995	200	150	0
2	1996	3000	135	0
:	:	:	:	:
525	1995	200	150	0
525	1996	3000	135	0
:	:	:	:	:
:	:	:	:	:
525	2000	4566	356	1

Introducción

Enfoque del curso

Muestreo aleatorio simple

- Un modelo poblacional
- Se eligen N individuos en forma aleatoria de una población de interés. Cada individuo tiene la misma probabilidad de ser elegido.
- Denotamos la muestra aleatoria como $\{w_1, w_2, \dots, w_N\}$.
- Dado el muestreo aleatorio, las w_i 's son variables aleatorias **independiente e idénticamente distribuidas (i.i.d.)**
- Supuesto razonable para datos de corte transversal y datos de panel.

- Regresores estocásticos

Se asume que los regresores (las x 's) son estocásticos. NO se asume que son fijos en muestras repetidas.

Si los regresores son fijos no existe correlación entre los inobservados y los regresores. No es realista para datos observacionales.

- Teoría asintótica

Distribución muestral y propiedades de los estimadores cuando el tamaño de la muestra converge a infinito.

No es necesario hacer supuestos distribucionales como normalidad.

No suponemos homoscedasticidad porque no es realista y sus consecuencias poco importantes.

Introducción

Trabajo grupal

Trabajo grupal: Retornos a la educación

- **Pregunta:** ¿Cuál es el efecto de un año adicional de educación en los salarios?
- **Población:** Personas entre 30 y 65 años que residen y trabajan en Chile. Año: 2017.
- **Fuente de datos:** Casen 2017.
- **Variables**
 - *ytrabajacor*: Ingreso del trabajo corregido
 - *esc*: Años de escolaridad.

Estadísticas descriptivas

```
. summ esc ytrabajacor, d
```

Escolaridad

```
-----  
Percentiles      Smallest  
1%               1             0  
5%               4             0  
10%              6             0      Obs           67,481  
25%              9             0      Sum of wgt.    67,481  
  
50%              12  
Largest          Std. dev.     3.937814  
75%              15             22  
90%              17             22      Variance       15.50638  
95%              17             22      Skewness       -.3869452  
99%              19             22      Kurtosis       3.097094
```

Estadísticas descriptivas

Ingreso del trabajo

Percentiles		Smallest		
1%	35000	250		
5%	100000	417		
10%	154167	500	Obs	66,789
25%	270000	833	Sum of wgt.	66,789
50%	370000		Mean	582796
Largest		Std. dev.	886911.8	
75%	608334	3.40e+07		
90%	1100000	4.00e+07	Variance	7.87e+11
95%	1680000	4.88e+07	Skewness	21.85263
99%	3530000	8.00e+07	Kurtosis	1298.178

- Ecuación empírica

$$\log y_{trabajador_i} = \beta_0 + \beta_1 esc_i + u_i$$

Estimaciòn

```
. regress lwage esc, r
```

```
Linear regression               Number of obs   =   66,366
F(1, 66364)                    =   17519.91
Prob > F                       =   0.0000
R-squared                      =   0.2493
Root MSE                       =   .73898
```

```
-----+-----
                Robust
lwage | Coefficient  std. err.      t    P>|t|    [95% conf. interval]
-----+-----
   esc |   .1082153   .0008176   132.36   0.000   .1066129   .1098177
 _cons |   11.62096   .009928   1170.53   0.000   11.6015   11.64042
-----+-----
```

Trabajo grupal: Retornos a la educación

- Interpretación
- ¿Usted cree que estamos midiendo un efecto causal?
 - Causas
 - Sesgo
 - Soluciones
- Forma funcional
- Heterogeneidad

Trabajo grupal: Algunas ideas aleatorias

- Función de producción/productividad: ENIA (Encuesta Nacional Industrial Anual)
- Educación: Datos de Ministerio de Educación (notas, asistencia, rendimiento en el colegio, universidad)
- Salud, Educación, Empleo: Casen
- Innovación: Encuesta de Innovación
- Etc