

# UNIDAD III. DISEÑO EXPERIMENTAL: PRINCIPIOS Y TÉCNICAS

Joselin Segovia Sarmiento

Correo: [joselin.segovias@ucuenca.edu.ec](mailto:joselin.segovias@ucuenca.edu.ec)

## Contenido

1. El enfoque econométrico al análisis de datos
2. Inferencia basada en experimentos
3. Tipos y usos de los experimentos
4. Validez interna.
  - 4.1 Control indirecto
  - 4.2 Estructura de incentivos
  - 4.3 Control directo
5. Otras consideraciones: muestra, instrucciones, logística, ética.

# Preámbulo / Consideraciones iniciales

¿Qué es un experimento?

- Un experimento es una situación **controlada** en la que muchas características del **entorno** son implementadas por **diseño**, para **observar** las decisiones e interacciones individuales **resultantes**.
- El objetivo de tal observación es inferir las relaciones **causales** entre el **entorno** y el comportamiento de los individuos.

# Preámbulo / Consideraciones iniciales

- Los experimentos son un **proceso de generación de datos controlado**.
  - "Control" significa que la mayoría de los factores que influyen en el comportamiento se mantienen constantes y sólo se varía un factor de interés (el "tratamiento") a la vez.
  - Esta variación controlada de factores es crucial para hacer inferencias causales.
- Esto se puede describir como: un experimento es la elección de un conjunto de **n** inputs, que definen el entorno, asociados con **m** medidas de sus consecuencias con el objetivo de inferir su relación causal, **F**.

- Un experimento se ocupa de las **decisiones de los seres humanos** en un entorno particular del mundo real: laboratorio/campo o sus híbridos.
- Por lo tanto, los resultados medidos en el laboratorio no son solo los resultados de los controles elegidos, sino también la consecuencia de una gama infinita de influencias.
- Por tanto, muchas dimensiones más allá de las **n** entradas que se controlan y eligen realmente, influyen en lo que sucede.
- Ejemplos: el nivel de ruido dentro y fuera del laboratorio, el estado de ánimo de los participantes al llegar al experimento, temperatura del ambiente.

- Dos importantes implicaciones:

1. Un aspecto crítico del diseño experimental radica en la elección del número finito de entradas, **n**, y el número finito de medidas, **m**.
  - La elección depende completamente de la relación y el fenómeno que el experimento busca estudiar/medir: la pregunta de investigación.
2. Los experimentos comparten características comunes con los modelos teóricos y con la vida económica real.
  - Similar a un modelo, se enfoca en un subconjunto de dimensiones relevantes para estudiar el fenómeno de interés.
  - Similar a la realidad, lo que ocurre en el experimento está hecho de infinitas causas y consecuencias que nadie es capaz de controlar u observar.

# Preámbulo / Consideraciones iniciales

Smith (2014) señala que hay algunos aspectos básicos que cualquier individuo “nuevo” en el área debe conocer.

## 1. Teorías de comportamiento:

La economía experimental ha crecido junto con la teoría de juegos y las teorías de toma de decisiones.

Por tanto es importante tener en mente lo siguiente:

- Teoría de la decisión: Esta línea de literatura trata de comprender mejor cómo los individuos toman decisiones bajo riesgo e incertidumbre, qué papel juegan las preferencias en el tiempo y el descuento del futuro y qué lleva a las inconsistencias en la elección.

- Teoría de juegos: Los agentes en una economía interactúan entre sí; su comportamiento está directamente influenciado por las decisiones de otros agentes y, en particular, por las creencias que puedan tener sobre el comportamiento futuro de estos otros agentes. Tales consideraciones conducen a la toma de decisiones estratégica, que es un tema importante en la economía experimental



- Preferencias no estándar: se han desarrollado vistas alternativas del comportamiento que ahora forman parte del bagaje de herramientas de los economistas. Esto incluye:
  - Modelos de decisión no estándar: TP.
  - Preferencias sociales.
- Agregación: La sociedad tiene que tomar decisiones y, por lo tanto, necesita agregar de una u otra forma a las preferencias individuales.
  - Este es el enfoque de la teoría de las subastas, el análisis de mercados y los estudios de toma de decisiones colectivas, como la votación.

- Psicología del comportamiento: para estudiar la toma de decisiones individuales, se requiere pedir prestado mucho de la psicología.
  - Una gran parte de esta literatura está dedicada a explorar las **desviaciones sistemáticas de la toma racional de decisiones**, asociadas con varios sesgos bien documentados, como el anclaje y el sesgo de status quo, el efecto de dotación, el sesgo de confirmación, la falacia de la conjunción, el efecto de marco, la ilusión de control, aversión a la pérdida.
- Neuroeconomía: El análisis del comportamiento individual también se ha basado recientemente en la teoría de decisiones en ciencias médicas, lo que llevó al desarrollo de la neuroeconomía, que utiliza medidas fisiológicas para relacionarlas con el comportamiento.

## 2. Juegos y entornos de toma de decisiones.

La implementación de los marcos teóricos descritos en un laboratorio por lo general requieren el uso de entornos, procedimientos y reglas particulares.

Son herramientas diseñadas para estudiar diferentes aspectos del comportamiento individual y son un aspecto básico para todo sujeto en el campo.

- Formas de elicitación:

Mecanismos que fuerzan a los agentes a revelar algo sobre ellos mismos, como el riesgo, las preferencias intertemporales, o las creencias sobre lo que otros harán.

- Los juegos experimentales:

Son juegos estructurados con propiedades teóricas específicas que se utilizan y estudian ampliamente en la EE.

Incluyen el dilema del prisionero, el juego de confianza, el juego de la caza del ciervo, el juego del dictador, el juego de adivinanzas, y muchos otros.

- Los cuestionarios psicológicos:

Se pueden utilizar para recopilar datos sobre cómo las personas reflexionan sobre sus decisiones y cómo evalúan las diferentes situaciones.

Por ejemplo: cuestionarios psicométricos que incluyen medidas de habilidades cognitivas y no cognitivas, rasgos de personalidad o emociones.

### 3. Campos y aplicaciones.

Economía laboral: que se centra en los efectos de las políticas de mercado laboral, el intercambio entre el consumo y el ocio, la función de producción de la educación, etc.

La economía personal: se centra en la forma en que las personas se comportan en las empresas, abordando cuestiones como la forma en que las personas eligen los empleos y sus razones, la cantidad de personas que trabajan y cómo responden a los incentivos monetarios y no monetarios.

- La organización industrial: se centra en cómo las empresas interactúan entre sí según variables de decisión como volúmenes, precios o niveles de publicidad. Además de las diferentes estructuras de mercado: subastas, oligopolios, perfecta competencia.
- La economía Ambiental: estudio de las políticas para regular comportamiento en detrimento del medio ambiente. Son críticos aquí el estudio de la toma de decisiones colectiva, y comportamiento polizante / free-riding.
- Economía de la salud: la salud preventiva es un asunto conductual clave. Se estudia el Mercado de servicios de salud, diseños de incentivos, esquemas de pagos.

- Economía y derecho: estudio de cómo los individuos reaccionan a diferentes leyes, reglas. Se enfoca a identificar si la gente cumple la ley, y cómo esta cambia normas sociales y los equilibrios.



#### 4. Aspectos metodológicos.

- Perspectiva del participante.
- Cómo EE se relaciona como otros métodos empíricos de la economía, la relación con la teoría.
- Cómo diseñar e implementar un experimento en el laboratorio
- Métodos estadísticos para el análisis de datos.

# Términos básicos

- Sujeto/participante: persona que está dispuesta (o es seleccionada) para participar en un experimento.
- Sesión experimental: instancia particular de un experimento en tiempo y espacio.
- Ronda/período: fases, etapas, repeticiones o rondas en las cuales se divide el experimento.
- Tratamiento línea base: versión de línea base, sirve como referencia para comparar con tratamientos (grupo de control).
- Tratamiento experimental: diferentes versiones de contexto experimental. Se caracteriza por cambiar algún elemento de interés, relacionado con la pregunta de investigación

# 1. El enfoque econométrico al análisis de datos

- Los experimentos son controlados porque:
  - El ambiente/contexto experimental es construido intencionalmente de acuerdo a los objetivos de la pregunta de investigación.
  - Al decidir el diseño, las reglas del “juego”, la información provista, el experimentador decide lo que el econometrista llama “proceso de generación de datos” (PGD).
  - En econometría, el reto principal/básico: encontrar conjunto de supuestos que mejor se adapten al PGD (desconocido).
  - Los experimentos invierten este reto.
  - Vamos a centrarnos en esta ESPECIAL característica de los experimentos.

# El enfoque econométrico al análisis de datos

Primero, ¿Cuál es el objetivo de los trabajos empíricos en economía?

- Tienen por objetivo extraer lecciones generales sobre la evidencia disponible en los datos. Ejm:
  - ¿Qué nos dice la sensibilidad de los consumidores a los precios acerca de la forma de su función de utilidad?
  - ¿Qué nos dicen las diferencias salariales de género acerca de la discriminación?
  - ¿Qué nos dice el acceso al agua sobre la desigualdad?

- La econometría se ha desarrollado, para responder a preguntas de este tipo, denominados aspectos inferenciales.
- Resolver estas preguntas depende de la consistencia entre dos conjuntos de supuestos acerca del proceso de generación de datos:
  1. Relacionado con los mecanismos que producen lo observado.
  2. Relacionado con cuán informativas son las estadísticas generadas a partir de lo observado (datos).
- Esto se operacionaliza tradicionalmente mediante supuestos de identificación.
- Reto por lo gral: encontrar el conjunto de supuestos que mejor se adapta a proceso de generación de datos.

- Experimentos invierten este proceso y permiten elegir el proceso de generación de datos que mejor se adapte a la pregunta de investigación.
- Entonces, los supuestos de identificación proveen una guía de cómo diseñar los experimentos.
- Fuente importante para guiarnos = cuasi experimentos / experimentos naturales
  - Característica de estos: usar cambios espontáneos en las reglas institucionales como un (cuasi) experimento.

- Ejemplo de identificación de una relación causal: Efectos de esquemas de remuneración sobre el desempeño laboral.
- Dos problemas del análisis de datos desde la econometría
  - Cómo está la muestra relacionada con la población = problema de **identificación**
  - Dado que trabajamos con realizaciones de una variable aleatoria lo que observamos siempre diferirá poco/mucho de lo que queremos medir, a pesar de las propiedades de identificación que se consideren = problema de **inferencia estadística**.

- Dado que el verdadero PGD no se conoce, se supone/asume uno.
- Esto implica un conjunto de supuestos acerca de los datos observados: la forma funcional del modelo, la estrategia de muestreo.
- De asumir el PGD, se deriva la identificación e inferencia estadística del estimador de interés.
  - Por ejm: el estimador de la media muestral converge a la media de la poblacional si las observaciones son aleatorias.



# Estimación de efectos causales en tratamientos.

- Identificación = la pregunta principal que se debe abordar al diseñar cualquier investigación empírica.
  - Sin identificación, inferencia no tiene sentido.
- En el contexto de experimentos, la identificación se puede basar en el marco de evaluación causal.
- En donde se mide el efecto causal de un cambio en el ambiente de decisiones: efecto causal de un tratamiento.
- Se considera dos estados de la naturaleza:  $T = 0$  situación base, y  $T = 1$  cambia una dimension, donde el *tratamiento* genera ese cambio que queremos medir

- El resultado se denota  $y_{i0}$  en el estado sin tratamiento y  $y_{i1}$  en el estado con el tratamiento.
- Problema: observar la misma unidad muestral en ambos estados (individuo/firma/país) simultáneamente.
- Por tanto, no se puede medir  $y_{i1} - y_{i0} = \text{efecto causal}$
- El problema de la evaluación consiste entonces en encontrar **contrafactuales**: observaciones empíricas que miden/aproximan de manera convincente lo que el investigador no observa.
- Dos dimensiones críticas: (i) cuánto refleja el contrafactual lo que queremos observar, (ii) el tipo del efecto causal.

- Para ello, tomar en cuenta:

Datos de experimentos naturales proveen información de corte transversal de dos tipos de individuos: (i)  $T=1$  (ii)  $T=0$

Tenemos algo así, donde vectores para resultados son incompletos:

$i$	$\mathcal{T}_i$	$y_i(0)$	$y_i(1)$
1	1	–	10
2	0	2	–
3	1	–	3
...	...	...	...
$n$	0	5	–

## **Parámetros del efecto del tratamiento.**

Dadas estas circunstancias, la forma de estimar el efecto causal es a través del cambio promedio inducido por el tratamiento para cualquier individuo en la población = **efecto promedio del tratamiento** en evaluación de política.

### **1. Efecto promedio del tratamiento (ATE)**

Este parámetro mide el impacto del tratamiento en cualquier individuo de la población dotada de ciertas características individuales  $X$

## **2. Tratamiento promedio en los tratados (ATT)**

Este parámetro se restringe a la sub-población que recibe el tratamiento (condicionado en el grupo de individuos tratados).

Mide el cambio para aquellos individuos involucrados en el cambio.

- Estos dos parámetros pueden ser iguales o no, dependiendo del mecanismo detrás del cambio en el resultado.
- Si la respuesta individual al tratamiento es homogénea en la población, entonces los dos parámetros serán los mismos.

- Sin embargo, son diferentes si la población es heterogénea en términos de su respuesta al tratamiento, y la asignación al tratamiento está relacionada con dicha heterogeneidad, lo que induce una diferencia sistemática en el efecto esperado del tratamiento entre individuos de los dos grupos.
- Dicha diferencia se debe a la existencia de una relación entre el beneficio del tratamiento (su efecto esperado en el resultado) y la participación en el tratamiento.
- Dos tipos de implementación típicamente dan lugar a este tipo de mecanismo.

1. Si la participación en el tratamiento es gratuita. Es muy probable que aquellos individuos con mayor probabilidad de beneficio derivado del tratamiento, ingresen a él.
2. Si el tratamiento se dirige a una subpoblación en particular (trabajadores con salarios bajos, por ejemplo) y está diseñado a propósito para cambiar su resultado.

- El ATE y el ATT son parámetros verdaderos de la distribución del cambio causal inducido por el tratamiento en estudio.
- Cuando difieren, la pregunta es ¿cuál queremos conocer y / o cuál informará mejor sobre las consecuencias del tratamiento.
- Depende de la pregunta de investigación.
- Por lo tanto, la estrategia de estimación debe adaptarse al verdadero parámetro de interés.
- Los requisitos de observación de los dos parámetros son bastante diferentes.



- El ATT se basa en los resultados de la submuestra de individuos tratados:
- $ATT = E[y_i(1) | \mathbf{X}, T_i = 1] - E[y_i(0) | \mathbf{X}, T_i = 1]$
- Se tiene  $E[y_i(1) | \mathbf{X}, T_i = 1]$ , ¿cómo hallar  $E[y_i(0) | \mathbf{X}, T_i = 1]$ ?
- Problema contrafactual: medir/hallar datos sobre los resultados de los individuos tratados si el tratamiento no se hubiera implementado

- El requerimiento contrafactual para medir el ATE es aún más fuerte
- Denotando  $p = \Pr [T = 1]$  la probabilidad de ser tratado, el ATE se define como:

$$ATE = p E [ y_i(1) - y_i(0) | \mathbf{X}, T_i = 1 ] + (1 - p) E [ y_i(1) - y_i(0) | \mathbf{X}, T_i = 0 ]$$

Además del comportamiento observado de los individuos no tratados, también se necesita encontrar un contrafactual para el resultado de los individuos no tratados si hubieran sido tratados.

# Identificación con datos observacionales.

- El estimador de corte transversal compara los resultados promedio de los individuos tratados y no tratados en un solo período de tiempo.
- Resultados  $y(1) \mid X, T = 1$  para quienes recibieron el tratamiento
- y  $y(0) \mid X, T = 0$  para los que no lo hicieron
- El estimador de corte transversal busca medir el efecto del tratamiento basándose en:

$$\hat{\Delta}^{Cross} = \frac{1}{n} \sum_{i=1}^n y_i(1) - \frac{1}{n} \sum_{i=1}^n y_i(0)$$

- Esta expresión deja claro que el estimador utiliza el resultado en el grupo de control como un contrafactual del resultado del grupo tratado.
- Es decir: el estimador estaría midiendo un parámetro relevante solo si  $y(0) | X, T = 0$  es una medida correcta del resultado que el individuo tratado habría obtenido sin tratamiento.

Sin embargo, esto no está garantizado.

- Por esta razón, el grupo de control no proporciona un contrafactual preciso del resultado del grupo de tratamiento.
- Sesgo = dado por medida de la diferencia en resultados entre tratados y no tratados si no hubieran sido tratados (diferencia de línea base).
- Dado que esta diferencia es dependiente de si el tratamiento beneficia a los individuos o no = **efecto selección**.
- Si existe tal selección, entonces el estimador de corte transversal resulta en una estimación sesgada

- Este efecto de selección equivale a una violación de un supuesto de identificación porque implica que la heterogeneidad no observada se correlaciona con el tratamiento

= La distribución es diferente en la subpoblación de individuos que posteriormente recibirán el tratamiento y aquellos que no lo reciben.

- Esto resulta de selección endógena. Ejm: Selección de individuos con ciertas características para piloto de un programa.

## 2. Inferencia basada en experimentos controlados.

- Inferencia de experimentos controlados será diferente a lo anteriormente revisado sobre experimentos naturales.
- Principal razón: decidimos el proceso de generación de datos.
  1. La variación de interés se implementa intencionalmente de acuerdo con la pregunta de investigación.
  2. La participación es un elemento más del experimento.

- Así, supuestos de identificación de econometría = directrices de cómo diseñar para observar efecto de tratamiento de interés.
- De lo discutido respecto a los parámetros de identificación de efectos, dos mecanismos principales de identificación emergen:
  1. La eliminación de correlación entre elementos no observados y el cambio en la variable explicativa de interés.
  2. Medir y así eliminar el efecto de los factores confusores.



# Identificación mediante aleatorización.

- Condición crucial para identificación: exogeneidad de las variables explicativas = no correlación entre el componente no observado del resultado y las variables de interés.
- Forma más simple de evitar correlación es aleatorización.
- ¿a qué supuesto de OLS nos referimos?

$$\mathbb{E}(\varepsilon_i|\mathbf{X}) = 0, \forall i \Rightarrow \mathbb{E}(\hat{\theta}_{OLS}|\mathbf{X}) = \theta$$

- En el caso extremo, esto obviamente se obtendría si uno pudiera decidir de manera puramente aleatoria el valor de la variable explicativa para cada individuo.

- Ejemplo: estudio para investigar razones conductuales de diferencias de género en comportamiento.
- Se busca ver si hay una diferencia de género en el desempeño (ambientes competitivos).
- Se consideran grupos cuyos miembros: todos mujeres/hombres/mixto. Realizan una tarea simple.
- Esto es conocimiento general.
- Resultados muestran que composición de género del grupo tiene un efecto sobre desempeño. Mujeres tienen mucho peores resultados al competir en grupo con miembros únicamente hombres.

- Aquí, a pesar de que el género no se puede aleatorizar. Se puede aleatorizar la pertenencia de un miembro a un grupo.
- Regla: asignación aleatoria al tratamiento.
- Por diseño/construcción entonces adoptamos el supuesto de identificación del estimador de corte transversal

$$\mathbb{E}[\varepsilon_i(0) | \mathbf{X}, T_i = 1] = \mathbb{E}[\varepsilon_i(0) | \mathbf{X}, T_i = 0]$$

- Entonces, para medir ATT simplemente comparamos entre tratados y no tratados.

- Pero, ¿qué es población en el contexto experimental?
- La identificación y el efecto del tratamiento es específico para la población experimental.
- Fuerte crítica a experimentos.
- Clave: definición de población sobre la cual hacer inferencia.
- En términos de identificar efecto, sin embargo, se sigue obteniendo estimación insesgada del efecto.

# Identificación mediante variables de control.

- Con datos observacionales: mejoran la identificación, en lugar de porque su efecto realmente pertenece a la pregunta de investigación.
- Con datos experimentales: son parte misma de la pregunta de investigación.
- Variables explicativas ingresan directamente en el estudio como condicionantes. Siendo exógenas, se identifica su efecto.
- Pero también indirectamente en el componente no observado.
- Por construcción, cualquier componente de variable explicada se elimina de residuo tan pronto como se incluye en X.

- Esto, permite otro método de identificación.
- Agregar medidas al modelo reducirá el ruido y, por lo tanto, reducirá efecto de posibles factores de confusión en el término de error.
- Mejora la identificación cuando dimensiones no observables que se correlacionan con los observables de interés, se vuelven medibles.

# Inferencia y control.

- Control: descartar variaciones confusas en el error, bloqueando su valor a un nivel específico.
- Si hay una variable no observada,  $z$ , pertenece a error.
- En evaluación de impacto, por tanto, genera endogeneidad al generar variaciones en  $T$  y en resultado,  $y$ .
- Aleatorización rompe la correlación entre las causas desconocidas del resultado y la variación debido a  $T$ .
- Consideremos ahora, un diseño en el que  $z$  se mantenga constante entre individuos.

- Si se mantiene constante, su efecto no observado sobre el resultado, y, ya no es diferente entre submuestras ( $T=1$ ,  $T=0$ ).
- Mientras más dimensiones no observadas se puedan mantener constantes, mejor estimación.
- Por tanto, control mejora también inferencia.
- Conclusión: los valores muestrales derivados de datos experimentales son más informativos a medida que más dimensiones inobservables se mantienen constantes entre observaciones.



### 3. Tipos de experimentos.

- Diversos. Los más comunes:
  1. Experimentos de campo.
  2. Experimentos de laboratorio.
- Diferencia principal: entorno natural vs. entorno artificial.

- Experimento de campo produce observaciones:
  - (i) De una asignación aleatoria de individuos al tratamiento
  - (ii) En un entorno "natural" o "real"
- Sujetos toman decisiones en el mismo contexto social que lo harían en circunstancias normales.
- Ensayos de control aleatorio (RCT / ECA).
- A pesar de ello, varias dimensiones que hacen que decisiones observadas estén cerca o lejos de sus circunstancias naturales.

- Harrison and List (2004) caracterizan experimentos de acuerdo a:
  1. La naturaleza de la muestra de sujetos.
  2. La naturaleza de la información que los sujetos aportan.
  3. La naturaleza de la mercancía.
  4. La naturaleza de la tarea o las normas comerciales aplicadas.
  5. La naturaleza de lo que está en juego.
  6. La naturaleza del entorno en el que opera el sujeto.

- Obvias categorías: cercano – lejano a realidad.
- Sin embargo, autores definen  $6*6 = 36$  diferentes clases de experimentos.
- De acuerdo a Gerber y Green (2012), un experimento es de campo según lo siguiente:
  1. Autenticidad de los tratamientos.
  2. Identidad de los participantes.
  3. La naturaleza del contexto.
  4. Las medidas de resultado.

## **Guía para elegir tipo de diseño.**

- Ninguno tiene superioridad metodológica absoluta.
- Depende de pregunta de investigación.
- La elección implica un trade-off en diversas dimensiones:
  1. Aspectos prácticos y facilidad de implementación. Costo es lo más evidente: mayor mientras más real. Además, complejidad, duración.
    - Acuerdos, hábitos locales, planificación, pilotos largos.
  2. Aspectos inferenciales. Más real, mejor representación = validez externa. Pero, también implica más variables y factores difíciles de controlar = reduce capacidad de evidencia causal pura.

### 3. Habilidad de aleatorizar en la práctica.

- Abandono. Entonces:
  - Tamaño muestra disminuye si abandono es aleatorio.
  - Endogeneidad: si abandono es específico de individuos con determinadas características relacionadas con la variable de resultado. Entonces: problema identificación.
- Incumplimiento. En campo es más difícil restringir individuos a seguir reglas. Entonces:
  - Algunas personas en el grupo de tratamiento en realidad no lo reciben.
  - Reciben un tratamiento diferente al estudiado.
- Efecto derrame. Efecto de tratamiento puede tener impacto en no tratados. Intensifica y sesga el efecto de tratamiento.

# Usos de los experimentos.

1. Probar teoría, es decir, evaluar la relevancia empírica de los modelos teóricos.

Pregunta clave: ¿es el comportamiento observado inducido por el ambiente de decisión que replica un modelo teórico? = validez interna.

2. Buscar hechos, en cuyo caso utilizan la realidad para informar la teoría; los experimentos se usan luego como "exhibiciones" en lugar de "pruebas" (Sugden, 2005).

Pregunta clave: ¿se observará el mismo comportamiento fuera del entorno experimental? = validez externa.

3. Basándose en ambos tipos de contribución, los experimentos pueden usarse para informar las decisiones de los responsables políticos.

Fuerte base teórica + evidencia bien documentada = validez interna + validez externa.

Estudiar por ejemplo: mecanismos de mercado, cambios de política o nuevas estructuras organizativas.