

- Resumen
- 1. Aleatorización
- 2. Control experimental (homogeneidad inducida al nivel de la unidad)
- 3. Experimentos naturales (o “como si fueran” aleatorizados )
- 4. Comparaciones de antes/después
- 5. Controlando ex post I: Regresión
- 6. Controlando ex post: II: Emparejar y Ponderar
- 7. Variables instrumentales (Instrumental Variables, IV)
- 8. Diseños de regresión discontinua (Regression discontinuity design, RDD)
- 9. Rastreo de procesos
- 10. Estrategias de entrada (Front-door strategies) (Argumento de mecanismos)

## Resumen

Los experimentos son una forma de averiguar si una cosa causa algo más. La idea básica es: pruébelo y descúbralo. Lo complicado es averiguar cómo probarlo de una manera que genere confianza en nuestras creencias sobre los efectos causales. Una estrategia comúnmente utilizada por los investigadores es la intervención aleatoria. Esta es la estrategia que está en el centro de la mayoría de las investigaciones experimentales realizadas por los miembros de EGAP. Pero hay otras estrategias que a veces son más apropiadas. Aquí describimos las diez estrategias más destacadas para determinar efectos causales.<sup>1</sup>

## 1. Aleatorización

La estrategia utilizada en pruebas de control aleatorias (o intervenciones aleatorias, experimentos aleatorios) consiste en utilizar algún tipo de sorteo para determinar quién, dentro de algún grupo, tendrá o no acceso a un tratamiento o programa (o quizás quién lo recibirá primero y quién después, o quién recibirá una versión y quién otra). El atractivo de este enfoque es que utiliza la aleatorización para averiguar cuáles son los efectos sistemáticos de un programa. La aleatorización reduce la posibilidad de que una relación observada entre el tratamiento y las variables de resultado se deba a “posibles factores que puedan generar confusión” (confounding factors), es decir, a otras cosas que son diferentes entre los grupos (por ejemplo, podría preocuparnos que las cosas parezcan mejores en áreas tratadas precisamente porque los programas son implementados en áreas donde las cosas funcionan bien, pero al saber que la selección fue aleatoria esta preocupación se elimina por completo). Esta es estrategia es poderosa porque garantiza que no hay una relación sistemática entre el tratamiento y todas las demás características que pueden afectar a las variables de resultado, seamos consciente de ellas o no. Por esta razón, a menudo se considera el mejor modelo a seguir. Sin embargo la aleatorización no puede utilizarse siempre y en todas partes, tanto por razones éticas como prácticas. Pero puede utilizarse en muchas más situaciones de lo que se tiende a creer. Léase Humphreys y Weinstein ([http://www.columbia.edu/~mh2245/papers1/HW\\_ARPS09.pdf](http://www.columbia.edu/~mh2245/papers1/HW_ARPS09.pdf)) para una discusión de las ventajas y las limitaciones de este método para la investigación en la economía política de desarrollo.

## 2. Control experimental (homogeneidad inducida al nivel de la unidad)

Una segunda estrategia que se usa más en la física y espacios de laboratorio es utilizar el control experimental para asegurar que dos unidades son idénticas entre sí en todos los aspectos relevantes excepto en el tratamiento. Por ejemplo, para ver si una pelota pesada cae más rápido que una más ligera, nos podemos asegurar de que tengan la misma forma y tamaño y dejarlas caer al mismo tiempo, en las mismas condiciones meteorológicas, etc. Luego podemos atribuir cualquier diferencia en los resultados a las características que no se han mantenido constante entre las dos unidades. Esta estrategia es fundamentalmente diferente a la utilizada en los ensayos aleatorios. En dichos ensayos, normalmente se renuncia a la idea de mantener todo fijo y se intenta asegurar que la variación natural sobre variables que se pueden o no observar, no produzca un sesgo en las estimaciones; adicionalmente se busca evaluar los efectos promedio en una gama de factores de contexto en lugar de un conjunto fijo de factores de contexto. Las ventajas de hacer este tipo control dependen de la confianza que tengamos en que se pueden controlar todos los factores relevantes; si no se puede, el enfoque aleatorio puede ser más apropiado.

### 3. Experimentos naturales (o “como si fueran” aleatorizados )

A veces los investigadores no pueden aleatorizar, pero la inferencia causal sigue siendo posible porque la naturaleza ha hecho la aleatorización por ellos. La característica clave del enfoque del “experimento natural” es que tenemos razones para creer que la variación en algún tratamiento natural es “como si fuera aleatoria”. Por ejemplo, digamos que los cupos en un colegio se asignan por sorteo. Entonces podríamos analizar los efectos de la asistencia a la escuela como si se tratara de un ensayo de control aleatorio. Una investigación ingeniosa sobre los efectos del conflicto en los niños realizada por Annan y Blattman (<http://www.chrisblattman.com/documents/research/2010.Consequences.RESTAT.pdf>) utilizó el hecho de que el Ejército de Resistencia del Señor (Lord’s Resistance Army, LRA) en Uganda secuestró a niños de forma casi aleatoria. Otro estudio ingenioso sobre los programas de Desarme, Desmovilización y Reintegración (DDR) realizado por Gilligan, Mvukiyehe y Samii ([http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1911968](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1911968)) tomó provecho del hecho de que los trabajos de una ONG se interrumpieron debido a una disputa contractual, lo que dio lugar a un grupo de control “natural” de excombatientes que no recibieron programas de desmovilización. Puede ver el libro de Dunning (<http://www.cambridge.org/us/academic/subjects/politics-international-relations/research-methods-politics/natural-experiments-social-sciences-design-based-approach>) para una guía sobre cómo encontrar y analizar experimentos naturales.

### 4. Comparaciones de antes/después

A menudo, lo primero que se busca para determinar efectos causales es la comparación de las unidades antes y después del control. Aquí se utiliza el pasado como control del presente. La idea básica es muy intuitiva: se apaga el interruptor de la luz y se ve que la luz se apaga; atribuir el cambio de luz a la acción parece fácil incluso en ausencia de aleatorización o control. Pero este enfoque no es tan fiable para muchas intervenciones sociales, especialmente en entornos cambiantes. El problema es que las cosas mejoran o empeoran por muchas razones, no necesariamente relacionadas con los tratamientos o programas que nos interesan. De hecho, es posible que debido a todos los demás cambios, las cosas puedan empeorar en un área del programa, incluso si esos programas tuvieron un efecto positivo (las cosas empeoraron pero no están tan mal como lo habrían estado sin el programa!). Una aproximación más sofisticada que la simple comparación antes/después se denomina “diferencia de diferencias” (difference in difference): básicamente

se compara la diferencia antes/después de las áreas de tratamiento con las de las áreas de control. Este es un buen enfoque, pero hay que asegurarse de que se dispone de buenos grupos de control y, en particular, de que no es probable que los grupos de control y de tratamiento cambien por razones distintas al tratamiento.

## 5. Controlando ex post I: Regresión

Quizás la aproximación más común para la identificación causal en el análisis estadístico aplicado es el uso de la regresión múltiple para controlar los posibles factores que pueden generar confusión. La idea es tratar de utilizar cualquier información que se tenga sobre por qué las áreas tratadas y no tratadas no son fácilmente comparables y ajustar estas diferencias estadísticamente. Este método funciona bien en la medida en que se puedan averiguar y medir los factores de confusión y su relación con el tratamiento, pero no funciona si no se sabe cuáles son los factores de confusión. En general, no conocemos cuáles son todos los posibles factores de confusión, y eso expone este enfoque a todo tipo de sesgos (de hecho, si se controlan las variables equivocadas, es posible *introducir* sesgos donde antes no existían).

## 6. Controlando ex post: II: Emparejar y Ponderar

Una variedad de enfoques alternativos trata de tener en cuenta las variables de confusión al emparejar de manera cuidadosa las unidades de tratamiento con una o varias unidades de control. Emparejar tiene algunas ventajas sobre la regresión (por ejemplo, las estimaciones pueden ser menos sensibles a las elecciones de la forma funcional), pero la idea básica es, sin embargo, similar, y de hecho los métodos de emparejamiento pueden implementarse en un marco de regresión utilizando ponderaciones adecuadas. Al igual que la regresión, en el fondo esta estrategia depende de la convicción de que no hay variables de confusión importantes que el investigador desconozca o no pueda medir. Algunos métodos específicos incluyen:

- emparejamiento óptimo total o por pares (<http://dept.stat.lsa.umich.edu/~bbh/hansen2004.pdf>) y revisar paquete `optmatch` (<https://github.com/markmfredrickson/optmatch>)
- emparejamiento óptimo por pares con buen balance mediante programación mixta de enteros (<https://projecteuclid.org/euclid.aoas/1396966284>) Véase también el paquete `designmatch` (<https://cran.rstudio.com/web/packages/designmatch> (<https://cran.rstudio.com/web/packages/designmatch>)) y el artículo de comparación de enfoques (<http://jrzubizarreta.com/evaluation.pdf>)
- emparejamiento óptimo de multinivel (para diseños con colegios y alumnos) (<https://projecteuclid.org/euclid.aoas/1536652962>)
- emparejamiento óptimo disperso (<https://www.stat.berkeley.edu/~spi/software.html>)
- emparejamiento completo generalizado (<https://arxiv.org/abs/1703.03882>)
- emparejamiento exacto con baja granularidad (<http://gking.harvard.edu/cem>)
- combinaciones genéticas (<http://sekhon.berkeley.edu/papers/GenMatch.pdf>)
- balance de entropía (<http://web.stanford.edu/~jhain/Paper/PA2012.pdf>)
- ponderación de propensidad inversa (<http://pan.oxfordjournals.org/content/18/1/36.short>)
- pesos de balance estables (<http://www.columbia.edu/~jz2313/sbw.pdf>), y el uso de
- controles sintéticos (<http://web.stanford.edu/~jhain/Paper/AJPS2015a.pdf>).

## 7. Variables instrumentales (Instrumental Variables, IV)

Se puede utilizar un método muy diferente para estimar los efectos causales si los investigadores pueden encontrar alguna característica que explique por qué un grupo determinado recibió un tratamiento, pero que no esté relacionada con la variable de resultado de interés. Esta característica se denomina instrumento. Por ejemplo, supongamos que estamos interesados en el efecto de un programa de medios de sustento sobre el empleo y resulta que la mayoría de las personas que accedieron a dicho programa lo hicieron porque eran parientes de un funcionario del programa. Entonces, si no hay otras formas de relacionar el hecho de ser pariente de esta persona con las perspectivas de algún empleo, se puede calcular el efecto del programa al trabajar el efecto de ser pariente de este individuo en las perspectivas de empleo. Este ha sido un enfoque bastante popular, pero parte del entusiasmo por él ha decaído un poco, básicamente porque es difícil encontrar un buen instrumento. Una aplicación ingeniosa para estudiar los efectos de la pobreza en el conflicto utilizó las precipitaciones de lluvia en África como instrumento para ingresos/crecimiento. Aunque existe la preocupación de que la correlación entre conflicto y pobreza se deba a que el conflicto pueda causar la pobreza, ¡no suena posible que el conflicto cause las precipitaciones! Por lo tanto, el uso de las precipitaciones de lluvia como instrumento dio mucha más confianza en que realmente existe una relación causal, y no sólo correlacional, entre la pobreza y el conflicto (<http://emiguel.econ.berkeley.edu/research/economic-shocks-and-civil-conflict-an-instrumental-variables-approach>).

## 8. Diseños de regresión discontinua (Regression discontinuity design, RDD)

El método de regresión discontinua no es muy utilizado, pero tiene un gran potencial. Funciona de la siguiente manera: digamos que se va a poner un programa a disposición de un conjunto de beneficiarios potenciales. Estos beneficiarios potenciales se clasifican en función de un conjunto de criterios relevantes, como pueden ser niveles de educación previos, la situación laboral, etc. Estos criterios pueden ser cuantitativos, pero también pueden incluir información cualitativa, como una valoración a partir de una entrevista. Estos criterios individuales se agregan en una única puntuación y se identifica un umbral. Los candidatos con una puntuación superior a este umbral son admitidos en el programa, mientras que los que están por debajo no son admitidos. Los grupos de “Proyecto” y “Comparación” son identificados, al seleccionar a los candidatos que están cerca de este umbral por ambos lados. Al usar este método podemos estar seguros de que las unidades tratadas y no tratadas son similares, al menos cerca al umbral. Incluso tenemos una medida directa de la característica principal en la que difieren (su puntuación en los criterios de selección). Esta información proporciona la clave para estimar un efecto del programa a partir de la comparación de las variables de resultados entre estos dos grupos. La ventaja de este enfoque es que todo lo que se necesita es que la agencia que lo implementa utilice un conjunto claro de criterios (que pueden convertirse en una puntuación), sobre los que se toman las decisiones de asignación de tratamiento. La desventaja es que sólo se pueden hacer estimaciones realmente fiables del impacto para las unidades situadas cerca al umbral. Para una visión general de la RDD, léase Skovron y Titiunik (<http://www-personal.umich.edu/~titiunik/papers/SkovronTitiunik2015.pdf>) y Lee y Lemieux (<http://econ.sites.olt.ubc.ca/files/2014/02/Lee-Lemieux-rev.pdf>); para dos aplicaciones interesantes, léase

Manacorda et al. sobre Uruguay (<http://emiguel.econ.berkeley.edu/research/government-transfers-and-political-support>) y Samii sobre Burundi (<http://journals.cambridge.org/action/displayAbstract?fromPage=online&aid=8963107>).

## 9. Rastreo de procesos

En gran parte del trabajo cualitativo, los investigadores tratan de establecer la causalidad, no sólo observando si la participación en un programa se asocia con mejores resultados, sino también (a) buscando los pasos del proceso a lo largo del camino que dirían si un programa tuvo los efectos que se cree que tuvo y (b) buscando pruebas de otras variables de resultado que deberían verse si (o tal vez: si y sólo si) el programa fue eficaz. Por ejemplo, no sólo si las personas de un programa de medios de sustento consiguieron un trabajo, sino si además recibieron formación en algo útil, si recibieron ayuda de personas del programa para encontrar un empleador en esa zona, etc. Si todos estos pasos se dan, se puede confiar en que la relación es causal y no espuria. Si se implementó un programa pero nadie participó realmente en él, esto podría dar pie a sospechar que cualquier correlación entre el tratamiento y las variables de resultados es espuria. La dificultad con este enfoque es que puede ser difícil saber si alguna prueba dentro del caso tiene valor probatorio. Por ejemplo, un programa puede tener efectos positivos (o negativos) a través de muchos procesos de los que no se sabe nada, y los procesos que se consideran importantes, pueden no serlo. Léase Humphreys y Jacobs (<http://www.columbia.edu/%7Emh2245/papers1/BIQQ.pdf>) para una descripción de la lógica Bayesiana subyacente al rastreo de procesos e ilustraciones de cómo combinarlo con otros enfoques estadísticos.

## 10. Estrategias de entrada (Front-door strategies) (Argumento de mecanismos)

Un último método, conceptualmente cercano al rastreo de procesos, consiste en en la argumentación usando mecanismos. Digamos que sabemos que sólo  $A$  puede causar  $C$  sólo a través de  $B$ . Digamos, además, que sabemos que ninguna tercera cosa puede causar tanto  $B$  como  $C$  (salvo, quizás, a través de  $A$ ) y que ninguna tercera cosa puede causar tanto  $A$  como  $B$ . Entonces se puede utilizar la covariación entre  $A$  y  $B$  y entre  $B$  y  $C$  para evaluar el efecto de  $A$  en  $C$ . La ventaja es que la causalidad puede establecerse incluso en presencia de factores de confusión, por ejemplo, incluso si variables no observadas causan tanto  $A$  como  $C$ . La dificultad, sin embargo, es que la estrategia requiere mucha confianza en sus creencias sobre la estructura de las relaciones causales. Para más información, léase Pearl (2000) ([https://books.google.com/books?id=wnGU\\_TsW3BQC&hl=en](https://books.google.com/books?id=wnGU_TsW3BQC&hl=en)).

- 
1. Autor de origen: Macartan Humphreys. Revisiones menores: Winston Lin, 30 de agosto de 2016. La guía es un documento activo y está sujeto a actualización por parte de los miembros de EGAP en cualquier momento; los colaboradores enumerados no son responsables de las ediciones posteriores. ↩