



USAID
FROM THE AMERICAN PEOPLE

MÓDULO EVALUACIÓN DE PLANES, PROGRAMAS Y PROYECTOS DE DESARROLLO

Experimentos Aleatorios

NOTACIÓN

$$D_i = \begin{cases} 1 & \text{individuo } i \text{ recibe tratamiento} \\ 0 & \text{de lo contrario} \end{cases}$$

$$Y_i(D_i) \begin{cases} \text{Variable de resultado} \\ \text{para cada individuo } i \\ \text{dado su estado } D_i \end{cases}$$

$$\tau_i = Y_i(1) - Y_i(0)$$

NOTACIÓN

$E[Y(1) | D = 1]$ → Media de la variable de resultado en el grupo de participantes.

$E[Y(0) | D = 1]$ → Media que hubieran obtenido los participantes si el programa no hubiera existido.

$E[Y(1) | D = 0]$ → Media de la variable de resultado entre los no participantes del programa, si hubieran podido participar.

$E[Y(0) | D = 0]$ → Media de la variable de resultado de quienes no participaron en el programa.

¿PARA QUÉ HACER EXPERIMENTOS ALEATORIOS?

Para evitar el sesgo de selección y, por ende, asegurar que:

$$E[Y(0) | D = 1] - E[Y(0) | D = 0] = 0$$



$$E[u_i | D_i] = 0$$

Cuando la asignación al programa se hace de manera aleatoria.

Experimento social controlado



¿CÓMO HACER EXPERIMENTOS ALEATORIOS?

- Idealmente: { Se toman dos individuos IDÉNTICOS y se trata a uno de ellos.



IMPOSIBLE

- { Se asignan aleatoriamente los individuos del grupo de estudio al grupo de tratamiento o al grupo de control.
- Alternativa:

VENTAJAS Y LIMITACIONES

Ventaja

- Es la manera más transparente para determinar el efecto de políticas gubernamentales. La estrategia de identificación es muy clara y no requiere supuestos.

Limitaciones

- Consideraciones éticas.
- Costo de la aleatorización → monetario y político.
- Sólo se puede evaluar el programa tal y como está diseñado.
- Sólo evalúa los efectos de corto plazo del programa.

INDIVIDUOS VS. CONGLOMERADOS

Individuos

- Se aleatorizan individuos a partir de un listado de los individuos elegibles.

Conglomerados (clústeres)

- Se aleatoriza no a individuos, sino barrios o comunidades.
- Se aprovechan restricciones logísticas o presupuestales, de capacidad operativa del programa. Se aleatoriza el orden de entrada al programa; se pospone la entrada de algunos grupos, de tal manera que actúen como grupo de control.
- No existe una clara ventaja de alguno de los diseños básicos sobre el otro.

MODELO EN DIFERENCIAS

Si la aleatorización es exitosa, es decir, no hay sesgo de selección, entonces el efecto del tratamiento se puede estimar como:

$$[\bar{Y} | D = 1] - [\bar{Y} | D = 0]$$

○

$$Y_i = \beta_0 + \beta_1 D_i + u_i$$



Estimador de diferencias.

Insesgado y consistente pero no necesariamente eficiente.

EJEMPLO: CANASTA

RESULTADOS

Educación de la madre	% niños bien nutridos, tratamiento	% niños bien nutridos, control	% niños bien nutridos, contrafactual	Educación materna, tratamiento	Educación materna, control
Baja	50	30	40	10	20
Media	70	40	50	70	60
Alta	90	70	80	20	20

$$\begin{array}{ccc}
 \downarrow & & \downarrow \\
 E[Y(1) | D = 1] & & E[Y(0) | D = 0] \\
 & & \downarrow \\
 & & E[Y(0) | D = 1]
 \end{array}$$

Distribuciones educativas de las madres

EJEMPLO: CANASTA ATT

$$\tau_{ATT} = E[\tau \mid D = 1]$$
$$E[Y(1) \mid D = 1] - E[Y(0) \mid D = 1]$$

$$\hat{E}[Y(1) \mid D = 1] = (50 * 0.1) + (70 * 0.7) + (90 * 0.2)$$
$$= 72$$

$$\hat{E}[Y(0) \mid D = 1] = (40 * 0.1) + (50 * 0.7) + (80 * 0.2)$$
$$= 55$$

$$\hat{\tau}_{ATT} = 72 - 55 = 17$$

EJEMPLO: CANASTA ATT

Dado que $E[Y(0) | D = 1]$ no es observable
(hipotético),

τ_{ATT} calculamos $E[\tau | D = 1]$

$$E[Y(1) | D = 1] - E[Y(0) | D = 0]$$

$$\begin{aligned}\hat{E}(Y(1) | D = 1) &= (50 * 0.1) + (70 * 0.7) + (90 * 0.2) \\ &= 72\end{aligned}$$

$$\begin{aligned}\hat{E}(Y(0) | D = 0) &= (30 * 0.2) + (40 * 0.6) + (70 * 0.2) \\ &= 44\end{aligned}$$

$$\hat{\tau}_{ATT} = 72 - 44 = 28$$

ESTIMADOR DE DIFERENCIAS CON REGRESORES ADICIONALES

La eficiencia del estimador se puede mejorar dependiendo de la disponibilidad de datos.

El estimador de diferencias puede no ser eficiente, porque:

- Otras variables, distintas al tratamiento, pueden también determinar la variable de resultado aun cuando la aleatorización es exitosa.

ESTIMADOR DE DIFERENCIAS CON REGRESORES ADICIONALES

Si la información que afecta la variable de interés está contenida en la base de datos, se puede incluir en la regresión como:

$$Y_i = \beta_0 + \beta_1 D_i + \beta_2 X_i + u_i$$

Estimador de diferencias con regresores adicionales.

Características del individuo que explican también la variable de resultado y preceden al tratamiento.

EJEMPLO: CANASTA CON ASIGNACIÓN ALEATORIA

Suponga que se debe medir la talla para la edad de los niños, donde la asignación al programa está hecha de manera aleatoria.

Estimación:

$$Y_i = \beta_0 + \beta_1 D_i + u_i$$



Estimador de diferencias que mide el efecto del programa *Canasta*.

EJEMPLO: CANASTA CON REGRESORES ADICIONALES

- Suponga que la población de interés está compuesta por dos razas: altos y bajitos.
 - La variable raza afecta la estatura de los niños y no depende de si la familia es beneficiaria o no.

$$Y_i = \beta_0 + \beta_1 D_i + \beta_2 X_i + u_i$$

↑
Estimador de diferencias
con regresores
adicionales.

↑
Raza = 1 para los altos
y 0 para los bajitos.

ESTIMADOR DE DIFERENCIAS CON EFECTOS HETEROGÉNEOS

Estima el efecto de la intervención entre distintos subgrupos de la población, dado que:



El efecto puede diferir entre individuos dependiendo de sus características.

$$Y_i = \beta_0 + \beta_1 D_i + \beta_2 X_i + \beta_3 D_i X_i + u_i$$

Cuando X_i es binaria el coeficiente β_3 mide el impacto adicional en la variable objetivo para un subgrupo de la población.

EJEMPLO: CANASTA CON EFECTOS HETEROGÉNEOS

Si la mejora en nutrición es más eficiente en altos que en bajitos, por su predisposición genética, el impacto dependerá de la variable raza:

$$Y_i = \beta_0 + \beta_1 D_i + \beta_2 X_i + \beta_3 D_i X_i + u_i$$

↑
Impacto adicional en estatura para los altos con respecto a los bajitos.


ESTIMADOR DE DIFERENCIAS EN EL TIEMPO

Calcula el impacto de estar expuesto a un programa utilizando datos del grupo de tratamiento únicamente.

Donde:

$$T_i = \begin{cases} i=0 & \text{período anterior a la aplicación del tratamiento} \\ i=1 & \text{período posterior a la aplicación del tratamiento} \end{cases}$$

Estimador: $\tau = [\bar{Y}_1 | D = 1] - [\bar{Y}_0 | D = 1]$



Indica el período de la observación

ESTIMADOR DE DIFERENCIAS EN EL TIEMPO

$[\bar{Y}_1 | D = 1]$ \longrightarrow Media de la variable de resultado en el grupo de los tratados en el período posterior a la intervención.

$[\bar{Y}_0 | D = 1]$ \longrightarrow Media de la variable de resultado en el grupo de los tratados en el período anterior a la intervención.

ESTIMADOR DE DIFERENCIAS EN EL TIEMPO

- Regresión:

$$Y_i = \beta_0 + \beta_1 T_i + u_i$$



Estimador de diferencias en el tiempo.

Importante:


Algunas variables de resultado pueden tener tendencia natural en el tiempo.

$\hat{\beta}_1$ → Es insesgado, si **no** existe tendencia temporal en la variable de resultado.

CONTROLANDO POR LA TENDENCIA

Se puede controlar por la tendencia temporal natural siempre que se tengan más de dos observaciones para cada individuo.

$$Y_i = \beta_0 + \beta_1 t_i + \beta_2 T_i + u_i$$



Tendencia de tiempo que afecta la variable de resultado.

Diferencia entre el antes y el después de la variable de interés para el grupo de tratamiento, neto del cambio que se debe al paso del tiempo.

Supuesto fuerte: la tendencia de tiempo es lineal.

Útil para la evaluación de programas universales o en curso.

EJEMPLO: CANASTA CON DIFERENCIAS EN EL TIEMPO

Podemos estimar el impacto del programa del tipo antes-después, usando únicamente el grupo de

$$Y_i = \beta_0 + \beta_1 t_i + \beta_2 T_i + u_i$$

Donde:

$$T = \begin{cases} 0 & \text{antes de iniciar el programa} \\ 1 & \text{después de la intervención} \end{cases}$$

$$t_i = \begin{cases} \text{Tendencia de tiempo que captura el hecho de que} \\ \text{los niños entre 0 y 5 años están en proceso de} \\ \text{crecimiento} \end{cases}$$

POTENCIALES PROBLEMAS DE ALEATORIZACIÓN

- Efecto experimental o de Hawthorne.
- Mala aleatorización.
- Pérdida de muestra o atrición.
- Debido al costo, es posible que el tamaño de la muestra sea pequeño, hecho que genera que los efectos no sean suficientemente precisos.

POTENCIALES PROBLEMAS DE ALEATORIZACIÓN

- Problemas de validez externa:
 - El experimento no es suficientemente similar al programa.
 - Efectos de equilibrio general o contaminación del grupo de control debido a externalidades.
 - Autoselección posterior a la aleatorización.



Si la aleatorización no es exitosa:



El estimador es sesgado y se requieren otros métodos econométricos para encontrar el efecto del programa.

PARA VERIFICAR LA ALEATORIZACIÓN

- Si el tratamiento está asignado de manera aleatoria, no debe estar correlacionado con las características individuales, entonces:

$$D_i = \beta_0 + \beta_1 X_i + u_i$$

Matriz de características del individuo.

Cuando X_i no está correlacionada con el tratamiento D_i los coeficientes que acompañan a X_i deben ser cero.

EJEMPLO: CANASTA PARA VERIFICAR LA ALEATORIZACIÓN

- Para determinar si la aleatorización de *Canasta* fue exitosa estimamos:

$$D_i = \beta_0 + \beta_1 X_i + u_i$$



Tratamiento.



Características de los niños y de los hogares.

Para que los resultados sean confiables, debemos rechazar la hipótesis sobre que el conjunto de la variables observables explican el tratamiento.

LA CAJA NEGRA

INS → ACTIVIDADES → PROD → RES → IMP



LA CAJA NEGRA

No se ha explicitado y no se sabe como las actividades operan los cambios, por medio de qué mecanismos o dinámicas, y cómo la población responde a la actuación del proyecto

Con la Teoría de cambio se busca identificar y explicitar esta dinámica para una mejor comprensión de la actuación y resultados del proyecto.



USAID
FROM THE AMERICAN PEOPLE