

Copyright © 2015 Banco Interamericano de Desarrollo. Esta obra está bajo una licencia Creative Commons IGO 3.0 Reconocimiento-No Comercial-Sin Obra Derivada (CC-IGO BY-NC-ND 3.0 IGO) (<http://creativecommons.org/licenses/by-nc-nd/3.0/igo/legalcode>) y puede ser reproducida para cualquier uso no-comercial otorgando crédito al BID.  No se permiten obras derivadas.

Cualquier disputa relacionada con el uso de las obras del BID que no pueda resolverse amistosamente se someterá a arbitraje de conformidad con las reglas de la CNUDMI. El uso del nombre del BID para cualquier fin que no sea para la atribución y el uso del logotipo del BID, estará sujeta a un acuerdo de licencia por separado y no está autorizado como parte de esta licencia CC-IGO.

Notar que el enlace URL incluye términos y condicionales adicionales de esta licencia.

Las opiniones expresadas en esta publicación son de los autores y no necesariamente reflejan el punto de vista del Banco Interamericano de Desarrollo, de su Directorio Ejecutivo ni de los países que representa.

**Script: Pareamiento**

**Diapositiva 1**

Hola. Me llamo Rosangela Bando y en este video te voy a darte una idea general de cómo se estima el impacto de un programa con el método de pareamiento y de cuáles son los supuestos que se deben cumplir para que esta estimación sea válida.

**Diapositiva 2**

Supongamos que tenemos un programa de tutorías y lo asignamos a un grupo de estudiantes y queremos saber cuál es el efecto de este programa sobre el aprendizaje de los estudiantes medido mediante una prueba estandarizada. Supongamos que observamos un grupo de estudiantes que reciben la tutoría, aquí denotados por tratamiento y observamos que en promedio obtienen 630 puntos. Cuando tomamos la media de un grupo de estudiantes que no reciben la tutoría observamos que el promedio es de 600 puntos. Una estimación ingenua del programa sería simplemente comparar los dos grupos y atribuir la diferencia al programa. Aquí sería un impacto de 630-600 igual a 30 puntos. Sin embargo, una comparación simple integra no sólo diferencias como resultado del programa, sino también diferencias en la composición de los grupos. Supongan por ejemplo que los estudiantes representados aquí gráficamente con gorro blanco tienen entre 10 y 13 años, pero los de la gorra roja son más chiquitos y tienen entre 6 y 9 años. Gráficamente vemos que el grupo de tratamiento tiene más niños grandes mientras que el de control tiene más niños chiquitos. Muy posiblemente la diferencia de 30 puntos se deba también a la diferencia en el desempeño reflejo de las diferencias de edades entre los grupos. El problema de pareamiento propone tomar el grupo de tratamiento y generar un grupo de comparación válido que tenga sus mismas características. Tomemos el grupo de tratamiento.

**Diapositiva 3**

Aquí hemos dividido a los estudiantes de acuerdo no sólo a su estatus de tratamiento sino también de acuerdo a su grupo de edad. Supongamos que el puntaje de los niños entre 6 y 9 años en el grupo de tratamiento es 610 y que el puntaje promedio de los niños de 10 a 13 en el grupo de tratamiento es de 635. ¿Cómo respondemos a la pregunta qué hubiera pasado con el tratamiento si no hubiera habido programa? Una solución es tomar a los niños del grupo de control que tengan las mismas características. Supongamos que los niños de 6 a 9 años en el grupo de comparación tuvieron 600 puntos y los de 10 a 13 tuvieron 600. Estos puntajes pueden modelar lo que hubiera pasado en el grupo de tratamiento sin el programa. Ahora, recordemos que la mayoría de los niños en el grupo de tratamiento son grandes. Supongamos que los niños grandes representan el 80% del grupo de tratamiento. Entonces aquí el efecto sería 610-600=10 para el grupo de niños pequeños, es decir el 20% y de 635-600=35 para el grupo de niños grandes, es decir el 80%. El efecto del programa con pareamiento en el grupo de tratamiento sería de 20 puntos. Noten que este efecto habla sobre el efecto sobre el grupo de tratamiento. Por lo tanto, al estimador se le llama *efecto de tratamiento promedio sobre los tratados*. Para que el efecto de tratamiento promedio sobre los tratados responda a la pregunta de qué hubiera pasado en el grupo de tratamiento en ausencia del programa se deben cumplir dos supuestos.

**Diapositiva 4**

El supuesto número uno que se debe cumplir para que el método de pareamiento proporcione estimaciones de los efectos del programa, es que una vez que se controla por características observables no tiene nada que ver el tratamiento con las características de los estudiantes que influyen en sus calificaciones. Es decir, la selección de tratamiento se basa en observables solamente. Otra forma de decirlo en nuestro ejemplo sería decir que una vez que hemos separado a los niños por grupo de edad, su participación en el programa de tutorías fue equivalente a una lotería, es decir no tuvo nada que ver con la capacidad de los niños de aprender. A este supuesto se le llama independencia condicional en la media. Por ejemplo, supongamos que también hay diferencias por género. Entonces podríamos tener más niñas en el grupo de tratamiento en los niños de edad 10 a 13 que en el de control. La diferencia de 35 se explicaría no sólo por el programa si no por diferencias por género. En este caso, tendríamos que hacer 4 grupos, uno de niños y otro de niñas y uno para niños chiquitos y otro para niños grandes y evaluar en cada grupo. El riesgo más grande es que existan diferencias que no podamos observar que no nos permitan hacer grupos comparables tales como la motivación. Si ignoramos características no observables como la motivación en el cálculo de los efectos, obtendremos una estimación que incorpora una mezcla de los efectos del programa y diferencias entre los grupos de tratamiento y comparación.

**Diapositiva 5**

El segundo supuesto es que para todos los individuos en el grupo de tratamiento, vamos a observar individuos en el grupo de comparación que tienen características parecidas. Aquí gráficamente lo hemos introducido como un estudiante con gorro negro.

**Diapositiva 6**

Si utilizamos el método de pareamiento, lo que haremos es clasificar a los estudiantes por grupo de edad. Supongamos que el estudiante con gorro negro tiene 14 años. Cuando hacemos la clasificación, no tendremos un niño de 14 años en el grupo de comparación que nos permita estimar cuál hubiera sido su calificación sin tutor. En este caso decimos que no hay soporte común.

**Dispositiva 7**

Regresemos al caso donde hay soporte común. Recordemos que el supuesto principal es que una vez que hemos controlado por características observables, aquí la edad, no hay características dentro de cada grupo que expliquen diferencias entre los grupos de tratamiento y comparación. El ejemplo que habíamos mencionado era el género. Si hubiera diferencias de género podríamos tener más niñas en el grupo de tratamiento en los niños de edad 10 a 13 que en el de control. En este caso, tendríamos que hacer 4 grupos, uno de niños y otro de niñas y uno para niños chiquitos y otro para niños grandes y evaluar en cada grupo.

**Dispositiva 8**

Pero aparte de la edad, podríamos clasificar a los estudiantes por el nivel educativo de los padres y el ingreso de su hogar. Esto generaría un número enorme de grupos que dadas las combinaciones puede ser nada práctico de manejar. Adicionalmente, las variables continuas se pueden dividir infinitamente. A esto se le llama *problema de dimensionalidad*. La solución al problema de dimensionalidad es calcular la probabilidad de participar en el programa de acuerdo a las características observadas. La probabilidad de participar es un número que va de 0 a 1. Este puntaje refleja la información en todas las variables de forma igualmente buena que generar muchísimos grupos. Una vez que podemos resumir toda la información de muchas variables como el nivel educativo de los padres, el ingreso en el hogar, la edad y el género en una sola dimensión que es el puntaje de propensión, entonces podemos encontrar para cada estudiante una estimación de comparación válida. Aquí voy a representar el puntaje de propensión en esta línea y supongamos que este punto representa un estudiante en el grupo de tratamiento. El control se genera calculando el promedio de los estudiantes en el grupo de comparación con un puntaje de propensión parecido, es decir en un intervalo alrededor de él. Para todos los que caigan en este intervalo se calcula un promedio que se imputa como el valor que hubiera tenido el estudiante si no hubiera tenido el programa. Una vez que todos los estudiantes en el grupo de tratamiento tienen su valor en el escenario sin programa, se hace la diferencia y se calcula el promedio para todos los estudiantes en el grupo de tratamiento. Para calcular de qué tamaño se deben de crear las ventanas y qué tanto cuenta un estudiante del grupo de comparación que está cerca y uno que está lejos existe una riqueza metodológica que puedes consultar en las referencias y que no tenemos tiempo de cubrir aquí.

**Diapositiva 9**

Cuando utilizamos el puntaje de propensión también podríamos tener un problema de soporte común. Aquí hay una gráfica que muestra la densidad de los participantes y no participantes de un programa. Podemos ver que hay zonas en el puntaje para las que hay individuos que no tienen grupo con quién hacer la comparación. A la izquierda tenemos individuos no participantes para los que no tenemos individuos en el grupo de tratamiento. De lado derecho tenemos individuos participantes para los que no contamos con un grupo de comparación. Una opción sería restringir la evaluación al subconjunto de la población en el soporte común. Esto es factible cuando el soporte común es amplio. Sin embargo, en casos donde el soporte común es muy estrecho hay que preguntarse qué tanto va a informar una estimación tan restringida para los propósitos de la evaluación.

**Diapositiva 8**

En resumen vimos que el método de pareamiento estima el impacto de un programa tratando de responder a la pregunta ¿Qué hubiera pasado en el grupo de tratamiento si no hubiera recibido el programa? Para responder a esto, se buscan los estudiantes con características observables parecidas y éstos se usan para estimar un puntaje de lo que hubiera pasado. Para que esta comparación efectivamente refleje el efecto de un programa se deben cumplir dos supuestos. Primero, se debe cumplir independencia condicional en la media, es decir, una vez que tenemos a los individuos clasificados por características observables entonces la asignación al programa no se relaciona con su resultado, por ejemplo el equivalente a una lotería. Noten que en puntajes entonces el supuesto es sobre todos los individuos en un intervalo en el puntaje de propensión. El segundo supuesto es que se cuenta con el soporte común para llevar a cabo la estimación del programa. Recuerden que el método de pareamiento estima el tratamiento promedio sobre los tratados.

**Diapositiva 9**

Si quieres saber más sobre pareamiento te aconsejo consultar el toolkit de Dufló y coautores, la guía de impacto de evaluación en educación y el libro de evaluación de impacto en la práctica.

**Diapositiva 11**

Esto es todo. Espero te sea útil.