

UNIVERSIDAD CATOLICA DE LA SANTISIMA CONCEPCION
FACULTAD DE INGENIERIA
INGENIERIA CIVIL INDUSTRIAL



Aplicación de Bilevel-Programming al Problema de Matching en Evaluación de Impacto

Diego Quiroga Alcázar

INFORME DE PROYECTO DE TITULO PARA OPTAR AL TITULO DE
INGENIERO CIVIL INDUSTRIAL

Profesor Guía
Dr. Oscar Cornejo Zúñiga

Concepción, Mayo de 2016

*Dedicado a
mi familia*

Resumen

El presente proyecto de título propone un estimador de matching para ser usado en evaluación de impacto de políticas públicas, mediante modelamiento matemático bilevel (bilevel-programming) y su transformación a un problema de programación lineal. Además, se presenta el pseudocódigo del programa codificado en Matlab, la estructura de las funciones del programa y la interfaz gráfica, así como resultados en términos de dimensionalidad y tiempo CPU de cada problema. El script se usa para generar benchmarks y comparaciones respecto de los estimadores de matching más populares que se encuentran disponibles en la literatura de evaluación de impacto.

El capítulo 1 es una introducción, donde se presenta el objetivo general del proyecto, objetivos específicos, justificación del problema y delimitación del problema.

El capítulo 2 se relaciona con el marco teórico del proyecto. Conceptos relacionados principalmente con la evaluación de impacto como lo son el por qué evaluar, grupos de tratamiento y control, objetivos de la evaluación de impacto, métodos para estimar el impacto de una política pública y principal énfasis a los métodos de matching, los cuales son actualmente los más usados.

El capítulo 3 está relacionado con el estimador BLOPmatching, el cual determina el impacto de un programa de desarrollo mediante la ponderación óptima en el emparejamiento de los distintos individuos que forman parte del experimento, balanceando el valor de las características observadas de cada individuo del grupo tratamiento con su contrafáctico.

El capítulo 4 muestra el desarrollo del software para la implementación de BLOPmatching. Pseudocódigo, funciones utilizadas, interfaz gráfica y los principales errores cometidos al usar el programa son mostrados.

Finalmente, el capítulo 5 pone a prueba a BLOPmatching con respecto al estimador de vecino más cercano basado en Propensity Score y en covariables en términos del impacto generado y su desviación estándar. Se incluye además un análisis en el cambio de la métrica para el estimador y un estudio del tiempo de ejecución del software desarrollado.

Abstract

This thesis project aims to propose a matching estimator for impact evaluation of public policies by solving a bi level optimization problem and its transformation into a linear optimization model. Pseudocode, functional syntax and the graphical user interface are presented in this project which was programmed in a Matlab environment as well as the estimation of CPU usage and benchmarks respect to the Nearest Neighbor Matching Estimator.

The first chapter is an introduction about the general objective, specific objectives, problem's justification and its definition.

The second chapter is about impact evaluation: why do we evaluate?, treatment and control groups, impact evaluation objectives, methods to estimate the impact of a public policy (focused in matching methods).

The third chapter is about the BLOPmatching estimator, which computes a program impact by finding the optimal weighting scheme for each individual, based on a covariates balance criterion.

The fourth chapter is about software development to implementing BLOPmatching. Pseudocode, programmed functions, graphic user interface and error boxes are shown.

Finally, the fifth chapter put on test the BLOPmatching estimator, comparing results with the Nearest Neighbor Matching Estimator, which is the most popular matching estimator currently available.

Agradecimientos

漫长的六年，在晚上学习。

我要感谢我的家人的都支持。我要说谢谢，爸爸Álvaro，妈妈Maritza，哥哥Danilo，哥哥Iván 和哥哥Rodrigo.

另外，我想感谢都朋友们对于你们的支持。谢谢教授对于给这个机会。

谢谢我的朋友们Guardians United CL 很好Destiny 火柴同你们。

不那么重要了，感谢咖啡因。

现在，我可以说: 我是工程师。

Índice general

Resumen	III
Abstract	V
Agradecimientos	VI
Lista de figuras	IX
Lista de tablas	XI
Nomenclatura y Abreviaciones	XIII
1. Introducción	1
1.1. Objetivo General	1
1.2. Objetivos Específicos	1
1.3. Justificación del Problema	2
1.4. Delimitación del Problema	3
2. Evaluación de Impacto	4
2.1. ¿Qué es la Evaluación de Impacto?	4
2.1.1. La Evaluación de Impacto como herramienta en la toma de decisiones.	5
2.1.2. Uso de datos para alimentación de modelos de Impacto	6
2.2. Modelo de inferencia causal de Neymann Rubin	6
2.2.1. Comparando el antes y el después	8
2.2.2. Construyendo el grupo control	9

2.3.	Métodos para estimar el impacto de programas de desarrollo	10
2.3.1.	Regresión discontinua	10
2.3.2.	Diferencias en diferencias	12
2.3.3.	Matching	14
2.3.3.1.	Estimador de vecino más cercano basado en Propensity Score (NN)	17
2.3.3.2.	Estimador de vecino más cercano basado en las covariables	19
2.3.3.3.	Ejemplo de matching	22
2.4.	Implementación de un programa de Evaluación de Impacto	27
3.	El estimador BLOPmatching	29
3.1.	Conceptos formales del estimador	29
3.2.	Linealizando el Estimador	31
3.3.	Estimando el impacto de una política pública mediante BLOPmatching	32
3.4.	Ejemplo	35
4.	Programación, diseño e implementación de BLOPmatching en Matlab	43
4.1.	Pseudocódigo de BLOPmatching	43
4.2.	Script blopmatch.m	50
4.2.1.	Ejemplo en el uso de la función BLOPmatching	50
4.3.	Interfaz de usuario de BLOPmatching	52
4.3.1.	Interfaz principal	52
4.3.2.	Ejemplo	54
4.3.3.	Cuadros de error	58
5.	Resultados y Discusión	61
5.1.	Análisis de tiempo de cálculo	61
5.2.	Benchmark de BLOPmatching respecto de Nearest Neighbor Matching	63
5.3.	Cambios en la métrica de BLOPmatching	68
6.	Conclusiones	70
	Referencias Bibliográficas	73

Índice de figuras

2.1. Ejemplo de regresión discontinua; Regresión entre ingreso per cápita y gasto en servicios básicos para cada familia	11
2.2. Ejemplo de regresión discontinua; Regresión para cada grupo en estudio y estimación de impacto	12
2.3. Ejemplo de match perfecto	15
3.1. Ejemplo BLOPmatching; Distribución de los datos en el plano.	36
4.1. Ejemplo blopmatch.m; Orden de los datos	51
4.2. Ejemplo blopmatch.m; Salidas del Script	52
4.3. Interfaz de usuario BLOPmatching	53
4.4. Interfaz de usuario BLOPmatching; Selección de variables y opciones .	55
4.5. Interfaz de usuario BLOPmatching; Barra de espera	56
4.6. Interfaz de usuario BLOPmatching; Ventana de éxito	56
4.7. Interfaz de usuario BLOPmatching; Resultados del matching en ventana de comandos de Matlab	57
4.8. Interfaz de usuario BLOPmatching; Error por no escoger estimador . .	58
4.9. Interfaz de usuario BLOPmatching; Error, variables de respuesta y tratamiento deben ser distintas.	58
4.10. Interfaz de usuario BLOPmatching; Error, variable de tratamiento debe ser binaria	59
4.11. Interfaz de usuario BLOPmatching; Error, la variable de tratamiento no está permitida como covariable	59

4.12. Interfaz de usuario BLOPmatching; Error, la variable de respuesta no está permitida como covariable	60
5.1. Histograma de la cantidad de vecinos en cada match	65

Índice de tablas

2.1. Ejemplo de Diferencia en Diferencias, estudio del efecto en el aumento del sueldo mínimo en la empleabilidad	14
2.2. Datos ejemplo de matching	22
2.3. Ejemplo de matching; Matriz de distancias entre individuos del grupo control y grupo tratados	23
2.4. Ejemplo de matching; Emparejamientos de cada individuo	24
2.5. Ejemplo de matching; Impacto individual de cada participante	25
2.6. Ejemplo de matching; cantidad de veces que cada individuo es usado en un matching	26
2.7. Ejemplo de matching; resumen estadístico del estimador	27
3.1. Ejemplo BLOPmatching; Grupo de ponderadores para cada individuo del grupo tratamiento	37
3.2. Ejemplo BLOPmatching; Grupo de ponderadores para cada individuo del grupo control	38
3.3. Ejemplo BLOPmatching; Estimación de la respuesta no observable de los individuos del grupo tratamiento	38
3.4. Ejemplo BLOPmatching; Estimación de la respuesta no observable de los individuos del grupo control	38
3.5. Ejemplo BLOPmatching; Impacto individual de cada individuo en el grupo tratamiento	39
3.6. Ejemplo BLOPmatching; Impacto individual de cada individuo en el grupo control	40

3.7. Ejemplo BLOPmatching; Ponderadores intragrupos del grupo en tratamiento	41
3.8. Ejemplo BLOPmatching; Ponderadores intragrupos del grupo control .	41
3.9. Ejemplo BLOPmatching; Varianza condicional de los individuos	42
4.1. Funciones de Matlab utilizadas en el pseudocódigo	48
5.1. Tiempo de cálculo para estimador BLOPmatching en segundos	62
5.2. Tiempo de cálculo para varianza del estimador BLOPmatching en segundos	62
5.3. Resumen estadístico Lalonde (1986)	64
5.4. Resultados de BLOPmatching de Lalonde (1986)	64
5.5. Comparativa entre BLOPmatching y Vecino Más Cercano para ATE .	66
5.6. Comparativa entre BLOPmatching y Vecino Más Cercano para ATT .	66
5.7. Balance post matching para el estimador ATE	67
5.8. Balance post matching para el estimador ATT	67
5.9. Resultados y comparativa de métricas para BLOPmatching	69

Nomenclatura y Abreviaciones

- ATE: Efecto en el tratamiento promedio (Average Treatment Effect).
- ATT: Efecto en el tratamiento promedio del grupo tratados (Average Treatment Effect on the Treated).
- BLOP: Bilevel Optimization Problem.
- BLOPmatching: Bilevel Optimization Matching Estimator.
- DID: Estimador de Diferencia en Diferencias.
- ETC: Equivalente a tiempo completo.
- NN: Nearest Neighbour Matching Estimator.
- NNPS: Nearest Neighbour Matching Estimator based on Propensity Score.
- PS: Propensity Score.

Capítulo 1

Introducción

1.1. Objetivo General

Proponer un estimador de matching (emparejamiento), mediante la solución de un modelo de programación matemática de dos niveles (Bilevel Optimization Model), tal que, se minimice la distancia entre cada individuo de forma separada que es tratado, respecto a la combinación convexa del conjunto de características formada por los individuos del grupo no perteneciente al tratamiento. Paralelamente se minimiza la suma de las distancias entre cada individuo del grupo tratado y su contrafáctico, conocido como grupo control.

1.2. Objetivos Específicos

- Modelar el estimador de matching, bajo un enfoque de programación matemática de dos niveles, siguiendo el trabajo propuesto por Díaz, Rau y Rivera (2015).
- Desarrollar un programa en entorno Matlab para resolver los modelos matemáticos y encontrar solución al estimador.
- Comparar los resultados obtenidos en el inciso anterior respecto del estimador de matching de vecino más cercano (Nearest Neighbour Matching Estimator).

1.3. Justificación del Problema

La evaluación de impacto es una rama de la econometría que permite la medición de la efectividad (en términos cuantitativos) de la implementación de programas de desarrollo o políticas públicas respecto de un grupo de individuos, según Gertler *et al* (2011). Es un campo que lleva poco tiempo de investigación y desarrollo respecto de otras disciplinas, y por ende, es necesario proponer nuevas herramientas que permitan realizar mejores pronósticos y obtener conclusiones fiables respecto a los posibles resultados de cada programa de desarrollo, con el objeto de minimizar el error generado por el sesgo de variables no observadas. Ejemplo en el uso de herramientas de la evaluación de impacto es el caso del programa de maternidad implementado entre los años 2001 y 2005 por el gobierno chino, con el objeto de minimizar la tasa de fallecimiento de madres a la hora de dar a luz, y mejorar la salud en niños recién nacidos en pequeñas localidades rurales del oeste de China, donde las condiciones de los centros de salud no son las óptimas. El caso anterior, fue investigado por Liu, Yan y Wang (2010). Para ello se definieron dos grupos: un grupo de individuos tratados y un grupo de individuos no tratados (grupo control). El resultado de la investigación concluyó que la tasa de mortalidad disminuyó en el periodo de implementación de la política. La estimación en términos cuantitativos del programa desarrollado por el Ministerio de Salud de China fue posible gracias a las herramientas que la evaluación de impacto dispone para los investigadores.

Actualmente, los estimadores de matching han sido ampliamente usados para determinar el impacto de políticas públicas. Estos estimadores en general se basan en Propensity Score, el cual se define como la probabilidad de que un individuo que posee las características para participar del programa en efecto vaya a participar de este, en función de los valores que poseen las variables en estudio para dicho individuo (Rosenbaum y Rubin, 1983). Uno de esos estimadores es el estimador de Vecino Más Cercano (Nearest Neighbour Matching Estimator), que realiza emparejamiento entre aquellos individuos que participan del programa y su grupo contrafactual (individuos no beneficiados), dada una serie de características comunes entre ellos (Abadie e Imbens, 2006). Este estimador divide las observaciones en K subgrupos distintos, donde cada individuo

de cada subgrupo posee características similares respecto de los demás. El resultado del impacto generado está dado por las medias ponderadas del resultado de cada subgrupo, definido por el recíproco de K . El problema es que la cantidad de subgrupos depende enteramente de la decisión arbitraria del investigador y su experiencia. Por lo tanto, no existe una regla que indique o recomiende la cantidad óptima de subgrupos tal que se minimice el efecto del sesgo por variables no observadas.

El estimador propuesto por Díaz, Rau y Rivera (2015) no requiere la determinación de K subgrupos, ya que entrega el esquema óptimo de ponderación para cada individuo en tratamiento respecto del grupo no tratado. El estimador parte de la base de minimizar la distancia entre las características de cada individuo con respecto a la combinación convexa del conjunto de características de cada individuo del grupo no tratado. Además, el estimador busca minimizar la distancia entre las características de los individuos beneficiados con el programa respecto de los no beneficiados. El estimador anterior (y de aquí en adelante) es conocido como BLOPmatching¹, y se basa en la aplicación de modelamiento matemático.

También, el proyecto se justifica en que los investigadores que estén interesados en aplicar BLOPmatching requerirán de herramientas para ello. Es por esto que es necesario un programa para que puedan aplicarlo a distinto tipo de datos e investigaciones de diferente índole.

1.4. Delimitación del Problema

El estimador es diseñado en términos computacionales mediante la aplicación de pseudocódigo, para luego ser traspasado a lenguaje máquina. Se realizarán comparaciones respecto del estimador de Matching de Vecino Más Cercano, el cual es el estimador más popular entre los investigadores de evaluación de impacto. Además, se generarán estudios de tiempo de CPU y dimensionalidad para el estimador. El estimador está programado en Matlab y se usaron subrutinas disponibles en R y STATA para generar los modelos a comparar con BLOPmatching.

¹El nombre proviene de BLOP, cuyas siglas en español significan Problema de Optimización Binivel (Bilevel Optimization Problem).

Capítulo 2

Evaluación de Impacto

El capítulo es una introducción y descripción general de la Evaluación de Impacto que cubre las herramientas que se encuentran disponibles en la literatura y cómo aplicar los métodos de estimación. Este tipo de evaluaciones busca determinar si un programa de desarrollo o un estudio de inferencia causal ha mostrado resultados positivos o no, basado en evidencia. Resultados posibles pueden ser en ejemplo el aumento en los ingresos de un grupo de personas, disminución en el consumo de tabaco, aumento en los resultados de escolares de bajos recursos. La evaluación de impacto busca algo más que sólo mostrar si los objetivos de los programas son alcanzados.

2.1. ¿Qué es la Evaluación de Impacto?

La Evaluación de Impacto es una rama de la Econometría cuyo principal objetivo es el de medir el resultado de una política pública o de un programa de desarrollo, el cual involucra individuos beneficiados e individuos no beneficiados. Lo anterior permite realizar estimaciones del impacto generado sobre esta subpoblación mediante el uso de técnicas estadísticas, en función del valor que toman las variables medidas para cada individuo, las cuales son llamadas Covariables (Gertler *et al.* (2011); Khandker, Koolwal y Samad, 2010).

El principal objetivo de la Evaluación de Impacto es el de entregar respuestas lo

más acertadas posibles de la evolución de un programa de desarrollo sobre individuos en estudio. Para ello hace uso de técnicas de monitoreo sobre la población a ser estudiada. En otras palabras, la Evaluación de Impacto busca responder preguntas de tipo causa y efecto. Imas y Rist (2009) definen tres preguntas que resumen la evaluación de un programa de desarrollo:

- Preguntas descriptivas: Se trata de determinar lo que ocurre en un programa enfocándose en los individuos. Para ello, se van describiendo los diferentes procesos y cambios en el ambiente de los individuos a través del tiempo. Para ello se realizan técnicas de monitoreo y predicción como lo es "Diseño longitudinal".
- Preguntas normativas: Similar a Preguntas descriptivas, sólo que ahora se agregan condiciones estándar para el estudio. Se observa cómo va cambiando el programa con respecto al tiempo. Los resultados se van comparando con respecto a las condiciones establecidas como estándar con el objetivo de ver si se cumplen o no las metas y objetivos del proyecto.
- Preguntas de causa y efecto: Mediante la medición del antes y el después se pretende determinar si un programa de desarrollo es efectivo o no.

2.1.1. La Evaluación de Impacto como herramienta en la toma de decisiones.

La Evaluación de Impacto se considera como una herramienta en la toma de decisiones. El desarrollo de nuevas herramientas permite simular los resultados de futuros programas de desarrollo sobre nuevas poblaciones dado resultados previos de programas anteriores. En términos de toma de decisiones, permite medir relaciones de tipo costo/-beneficio. Los programas de desarrollo y políticas públicas requieren de insumos de tipo monetario y material, para efectuar distintas actividades. Por lo tanto, si una política pública no está generando beneficios para la población y su costo es demasiado alto, entonces esta puede eliminarse. Luego pueden realizarse modificaciones al programa si es que este debe ser implementado sobre otras poblaciones.

2.1.2. Uso de datos para alimentación de modelos de Impacto

El diseño de experimentos para la Evaluación de Impacto requiere de diferentes tipos de datos, los cuales depende del diseño del programa. La determinación de cuáles variables serán medidas es de suma importancia, dado que en este tipo de estudios se generan problemas por el sesgo que generan variables no observadas en estudio. El análisis en Evaluación de Impacto es una combinación de variables cuantitativas y variables cualitativas, las cuales deben ser llevadas a números para poder realizar los análisis estadísticos correspondientes a cada caso. La combinación entre estos tipos de variables se le conoce como "métodos combinados".

La medición de variables es de gran importancia a lo largo de cada programa en estudio. Las variables de tipo cualitativa sirven para generar estimaciones futuras del rumbo que podría tomar en el futuro el programa de desarrollo. Mientras que, las variables cualitativas permiten conocer características no medibles de cada participante y el sentir de cada uno de ellos respecto de la evolución del estudio en el cual están bajo estudio. Lo anterior, puede ser realizado mediante entrevistas a subgrupos que sean estadísticamente representativos de la muestra total.

Otro tipo de variable de suma importancia son las de monitoreo. Estas permiten determinar el avance del programa, cambios en el estado de los individuos y el gasto en recursos en cada actividad concretada.

2.2. Modelo de inferencia causal de Neymann Rubin

La evaluación de impacto en esencia es un problema de estimar datos no observables y el modelo de inferencia causal de Neymann Rubin (Sekhon, 2008) es la base de los métodos y herramientas que se presentarán a lo largo de este proyecto de título. Consideremos que si un individuo se ve beneficiado de un programa de desarrollo entonces no tendremos forma de determinar la respuesta de ese individuo si no se hubiese enrolado en el programa. De la misma forma, para un participante no beneficiado no tendríamos manera de saber cómo hubiese sido el valor de la variable respuesta si hubiese

sido beneficiado. Es por lo tanto, nuestra mejor alternativa comparar los individuos en tratamiento con un grupo contrafáctico de individuos no beneficiados y mediante esto realizar las estimaciones de respuestas no observadas. Esto puede lograrse comparando un individuo del grupo tratamiento con otro del grupo control que posea características similares. Consideremos el siguiente modelo lineal para estimar la respuesta Y_i de un programa de desarrollo para un individuo i :

$$Y_i = \alpha X_i + \beta T_i + \epsilon_i \quad (2.1)$$

Donde T_i es una variable binaria que toma el valor de 1 para aquellos individuos que participan en el programa y 0 para aquellos que no participan. X_i es un grupo de variables y características observables para cada individuo i . Finalmente, ϵ_i es una medida del error reflejado por variables no observables. El problema del modelo presentado es que en general la variable de tratamiento T_i no se encuentra distribuida de forma aleatoria y esto se debe a que los programas de desarrollo son enfocados y diseñados para poblaciones y comunidades con necesidades y características específicas. Además, los programas son afectados por características y variables no observables, lo cual implica que se viola el supuesto principal de mínimos cuadrados para obtener resultados insesgados. En otras palabras la covarianza entre las variables regresoras y el error es distinto de cero. Podemos representar el problema de la siguiente manera: Supongamos un programa para combatir el problema del hambre. Sea Y_i la variable respuesta del programa como peso en kilogramos del individuo i . Para aquellos individuos que participan del programa $T_i = 1$ y el valor respuesta bajo este valor de tratamiento es $Y_i(1)$. Mientras que, aquellos que no participan se tiene que $T_i = 0$ y $Y_i(0)$. El efecto promedio del programa se puede representar como:

$$D = E[Y_i(1)|T_i = 1] - E[Y_i(0)|T_i = 0] \quad (2.2)$$

El problema es que para un individuo i que es beneficiario del programa sólo se conoce $[Y_i(1)|T_i = 1]$, pero no se puede saber el valor de si el individuo no hubiese participado del programa el cual es $[Y_i(0)|T_i = 0]$. En otras palabras, o el individuo participa del programa o no participa. Si se agrega, $[Y_i(0)|T_i = 1]$ a la ecuación anterior,

se tiene lo siguiente:

$$\begin{aligned} D &= E[Y_i(1)|T_i = 1] - E[Y_i(0)|T_i = 0] \\ &+ E[Y_i(0)|T_i = 1] - E[Y_i(0)|T_i = 1] \\ \Rightarrow D &= ATE + E[Y_i(0)|T_i = 1] - E[Y_i(0)|T_i = 0] \\ &\Rightarrow D = ATE + B \end{aligned} \tag{2.3}$$

donde ATE es el efecto en el tratamiento promedio (Average Treatment Effect), el cual representa el beneficio de los individuos en tratamiento con respecto a los no tratados. B es una medida del sesgo por características no observables. Como no se conoce $E[Y_i(0)|T_i = 1]$, no se puede calcular el sesgo. Por otro lado, si se asume que las respuestas Y_i son independientes de la variable tratamiento, entonces el sesgo se puede minimizar. Se conoce como *supuesto de independencia condicional* (Rosenbaum y Rubin, 1983):

$$[Y_i(1), Y_i(0)] \perp T_i | X_i \tag{2.4}$$

Este supuesto es de suma importancia, ya que si uno no lo considera, entonces será imposible tener una aproximación del sesgo B .

2.2.1. Comparando el antes y el después

La inferencia causal es en general un método difícil de utilizar, debido a que se requiere de clones perfectos para cada individuo que está en tratamiento. Esto en general no es posible o de poder lograrse los datos presentarían un sesgo demasiado alto producto de que encontrar clones quitaría la aleatoriedad que las observaciones deben presentar para ser representativas estadísticamente de poblaciones mayores. Dado lo anterior, existe otro método para determinar el impacto de una política pública, el cual

es simplemente comparar el antes y el después en el resultado del programa para cada individuo.¹ Sin embargo, es un método considerado como arriesgado, ya que podrían haber factores ajenos a las variables medidas que podrían afectar el resultado. Usaremos de ejemplo que está disponible en Gertler *et al* (2011) para explicar lo arriesgado de comparar el antes y el después de una política. Considere que un pequeño poblado produjo el año pasado 1000 [kg] de arroz. El gobierno está interesado en implementar un programa de desarrollo para aumentar la productividad. Una vez terminado el programa la producción del pequeño poblado fue de 1100 [kg]. El impacto generado en la política considerando el antes y el después fue de 100 [kg]. Sin embargo, esos 100 [kg] podrían explicarse por el aumento de las lluvias en el período de aplicación del programa, y no por el programa en sí. Por lo tanto, se genera un sesgo muy alto a la hora de estimar el impacto de un programa considerando sólo el antes y el después, causado por factores ajenos y que en general no son controlables.

2.2.2. Construyendo el grupo control

Como se mencionó, el modelo de inferencia causal de Neymann-Rubin es la base de los métodos para evaluar programas de desarrollo. Los métodos que se verán en la siguiente sección requieren de dos grupos: un grupo de tratamiento o de individuos que es beneficiario y otro grupo de control, contrafáctico o simplemente no beneficiario. Un paso clave es que cada individuo del grupo tratamiento tengan su clón, es decir, un individuo del grupo control que posee características similares. Muchas veces esto es muy complicado de lograr, por lo tanto lo que se busca es que la media de las variables y covariables del grupo tratamiento sean lo más cercanas a las del grupo control. Esto puede validarse mediante una prueba de hipótesis para las medias de ambos grupos y separados por cada variable. Es de suma importancia que ambos grupos sean invariantes ante la ausencia del programa. Es decir, que de no existir el programa, las características de los individuos del grupo tratamiento y control no varíen o cambien demasiado con respecto al tiempo, ya que de cumplirse esto, cualquier cambio en los individuos o en las

¹El método no considera un grupo de individuos no tratados, ya que se compara el antes y el después de los individuos que participan del programa.

respuestas será producido enteramente por la aplicación del programa sobre los sujetos de prueba. Ambos grupos no pueden someterse a otras intervenciones durante el período que se introduce en ellos la evaluación de impacto. Por ejemplo, si existe un programa para bajar de peso, los individuos del grupo control no pueden bajar de peso por efecto de otro programa similar. Cumpliéndose esto se asegura que el programa que se quiere aplicar es el que genera los efectos.

2.3. Métodos para estimar el impacto de programas de desarrollo

Actualmente existe diferentes métodos para estimar el impacto de un programa de desarrollo, la mayoría de ellos basados en la Inferencia Causal. A continuación se presentan los métodos clásicos para medir impacto. Sin embargo, pondremos especial énfasis a los métodos de matching, pues el presente proyecto de título propone un nuevo estimador para el método mencionado.

2.3.1. Regresión discontinua

El método de Regresión Discontinua es una herramienta para medir el impacto de políticas públicas o de desarrollo donde se establecen medidas estándar o un umbral fijo para ser beneficiado o no de esta (Lee y Lemieux, 2010). Por ejemplo, un programa para combatir a la pobreza en la región del Bio-bío. Una política de este tipo implicaría que sólo pueden incluirse individuos que posean bajos ingresos respecto de una medida previamente adoptada por los interesados en aplicar la política.

Propongamos un pequeño ejemplo de una situación ficticia para explicar cómo opera la herramienta de Regresión Discontinua. Consideremos que el gobierno está interesado en desarrollar un proyecto para disminuir el gasto en servicios básicos (agua, luz, gas y alimentos) para familias de bajos ingresos. Para ello escoge una subpoblación. Se define una variable que determina quiénes pueden participar del programa y quiénes no. En nuestro caso esa variable será el ingreso per cápita. Mientras que nuestra variable respuesta, la que determina el impacto del programa será el gasto per cápita en servicios

básicos. La figura 2.1 nos muestra el estado del programa antes de la aplicación del programa:

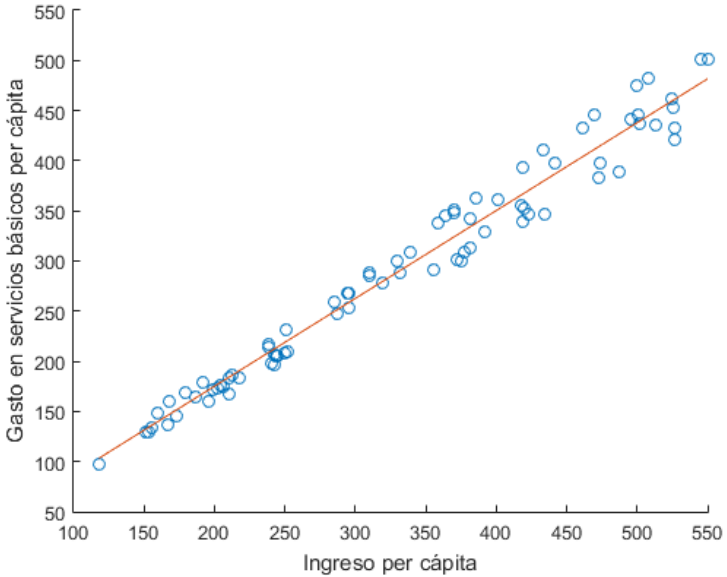


Figura 2.1: Ejemplo de regresión discontinua; Regresión entre ingreso per cápita y gasto en servicios básicos para cada familia

El método de Regresión Discontinua requiere como entrada adicional la norma de selección de los beneficiarios y de los no beneficiarios, el cual dependerá de los requisitos del programa de desarrollo o del criterio de los responsables en aplicar la política. Para este ejemplo, supondremos que el criterio de selección es un ingreso per cápita menor o igual a 400 unidades monetarias. El siguiente paso es monitorear el comportamiento de la variable Gasto en servicios públicos per cápita para cada familia en el grupo de beneficiados. Finalmente, se comparan ambos grupos mediante el uso de técnicas de regresión lineal, mediante el uso de los nuevos datos obtenidos del monitoreo. Se debería esperar que las familias con menos ingresos después del programa disminuyan su gasto per cápita en servicios básicos. Para efectos de comparación, se hace el supuesto de que las familias de mayores recursos no han variado el gasto en servicios básicos en el futuro. La figura 2.2 muestra el estado de ambos grupos después de la aplicación del programa de desarrollo.

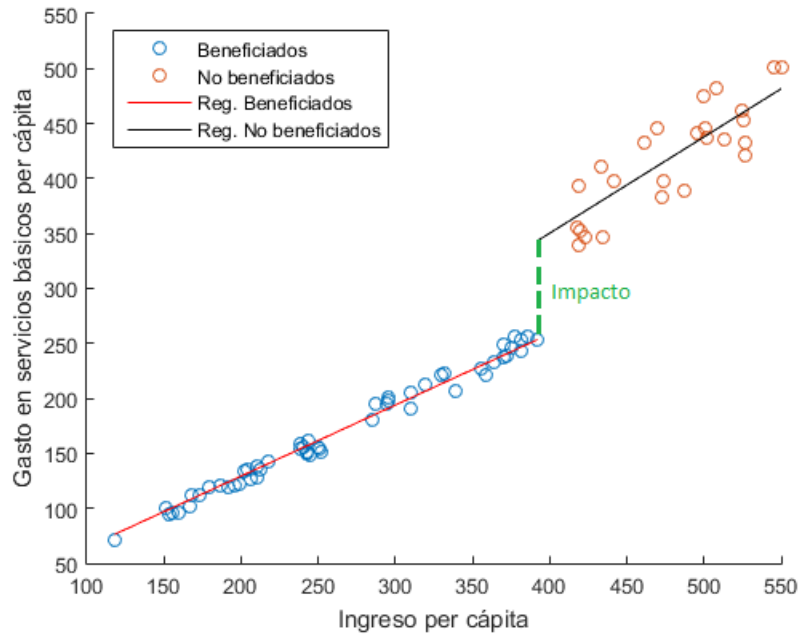


Figura 2.2: Ejemplo de regresión discontinua; Regresión para cada grupo en estudio y estimación de impacto

Como puede apreciarse en la Figura 2.2, las tendencias entre las curvas de regresión para las familias beneficiadas del programa y las no beneficiadas muestran un salto en el punto ingreso per cápita = 400 es decir, en el punto donde se produce el criterio de selección. El impacto es la diferencia entre el punto donde se produce el criterio de selección evaluada en ambas curvas, como puede apreciarse en la figura denotada con una línea discontinua. Para nuestro ejemplo ficticio, se puede observar que hay una disminución el gasto per cápita de servicios básicos para las familias de menores recursos.

2.3.2. Diferencias en diferencias

El método de Diferencias en diferencias (de aquí en adelante llamado DID) es una herramienta para determinar el impacto de una política pública mediante la diferencia entre el antes y el después del grupo de individuos en tratamiento, y la diferencia entre

el antes y el después del grupo de individuos que no están en tratamiento. El método posee la ventaja de que permite hacer un seguimiento completo con respecto al tiempo de la variable respuesta para todos los individuos. Sin embargo, un supuesto fuerte es que la tendencia futura de la respuesta promedio del grupo control (no beneficiados) posee el mismo comportamiento que la tendencia de la respuesta promedio del grupo en tratamiento (Abadie 2005).

El estimador DID es sencillo de aplicar y por ende es bastante popular entre los investigadores de evaluación de impacto. Liu, Yan y Wang (2010) investigan el efecto cuantitativo en la aplicación de un programa para disminuir la tasa de mortandad de madres a la hora de dar a luz en pequeños poblados rurales en el oeste de China. Los autores utilizan DID como herramienta para determinar el impacto del programa. Concluyen que el programa genera impacto sobre la población en estudio, disminuyendo así la tasa de mortandad materna dado el aumento en el cuidado en periodo prenatal.

Como ejemplo utilizaremos el caso estudiado por Card y Krueger (1994). Los autores en su investigación buscaban determinar el impacto que generaba el aumento en el sueldo mínimo en Estados Unidos sobre la empleabilidad en el rubro de la comida rápida. La teoría clásica de la macroeconomía indica que el aumento en el sueldo mínimo provoca aumento en el desempleo de jóvenes y personas con poca formación técnica y académica. Para ello se determinó que el grupo en tratamiento serían cadenas de restaurantes de comida rápida en New Jersey, mientras que el contrafáctico sería cadenas de restaurantes en Pennsylvania. Siguiendo el procedimiento del método, la tendencia futura en el desempleo para los restaurantes de comida rápida en Pennsylvania deben seguir el mismo comportamiento que New Jersey. Los datos fueron tomados entre febrero y diciembre de 1992. El incremento en el sueldo mínimo en New Jersey se inició en abril de 1992. La siguiente tabla muestra el uso de DID para determinar el impacto en el aumento del sueldo mínimo para New Jersey. El empleo fue determinado mediante el uso de una medida utilizada en estudios económicos, conocida como Equivalente a Tiempo Completo (ETC).²

²El Equivalente a Tiempo Completo es una medida usada en economía. Se calcula como la división entre la cantidad de horas semanales que un grupo de trabajadores a tiempo completo o parcial realizan en una semana y la cantidad de horas laborales a tiempo completo de un trabajador (Kimmel *et al*, 2010).

Tabla 2.1: Ejemplo de Diferencia en Diferencias, estudio del efecto en el aumento del sueldo mínimo en la empleabilidad

	New Jersey	Pennsylvania
ETC antes del programa	20,44	23,33
ETC después del programa	21,03	21,17
Diferencia	0,59	-2,16
DID	0,59 -(-2,16) = 2,75	

El resultado del estudio concluye que aumento en el sueldo mínimo genera 2,75 ETC y eso impacta de forma positiva el empleo. Por lo tanto no hay evidencia que muestren que aumento en el sueldo mínimo genere aumento en la tasa de desempleo.

Se puede apreciar que el procedimiento es bastante sencillo, sin embargo, no es un método robusto para determinar el impacto de estudios en efectos causales, debido a que deja de lado factores ajenos al estudio que podrían influir en el resultado final. Además, suponer que el grupo control tendrá el mismo comportamiento futuro que el grupo en tratamiento es un supuesto demasiado fuerte.

2.3.3. Matching

Matching o emparejamiento, es un método para la estimación del impacto de políticas públicas, programas de desarrollo y estudios de causa efecto que puede ser aplicado en la gran mayoría de los casos de evaluación de impacto. Las entradas para realizar matching son una variable respuesta, una variable binaria que determina si un individuo pertenece al grupo de tratamiento o al grupo control y una serie de variables medibles de interés para cada estudio. El método se basa en comparar cada individuo que será tratado con aquel individuo del grupo control que posea características similares. A esto se le conoce como 'matching' o 'emparejamiento'. La siguiente figura muestra un ejemplo de matching entre un grupo de individuos que pertenecen a dos subgrupo; grupo control y tratamiento. Se puede apreciar que al comparar las mismas características sólo dos individuos pueden ser emparejados.

Grupo control				Grupo en tratamiento			
Individuo	Cigarros consumidos al día	Edad	Peso (kg)	Individuo	Cigarros consumidos al día	Edad	Peso (kg)
1	23	25	87	6	15	18	72
2	21	15	67	7	14	22	67
3	15	47	91	8	30	35	83
4	18	51	77	9	21	15	67
5	30	33	73	10	22	66	108

Figura 2.3: Ejemplo de match perfecto

La teoría es sencilla, pero en la práctica es complicado emparejar a cada individuo en tratamiento con un individuo del grupo control. Esto debido a que el emparejamiento depende del valor de las variables que se tomen. Si se determinan pocas variables y si son de características (edad, nacionalidad, etc.), es probable que la mayoría de los individuos puedan ser emparejados, sin embargo esto produce problemas de sesgo estadístico por variables no observables. Mientras que a mayor cantidad de variables, la probabilidad de poder emparejar un individuo con otro disminuyen pues se determina una menor cantidad de similitudes conforme aumenta la dimensionalidad del estudio. El no poder generar matching dado los datos disponibles implica pérdida de observaciones; un ejemplo de esto se puede ver en Chapin (1947). En su estudio concluyó que de 671 individuos en el grupo de tratamiento y 523 en el grupo control sólo se pudieron realizar 23 matching para un problema de 6 variables. Sin embargo, Rosenbaum y Rubin (1983) desarrollaron un método conocido como 'Propensity Score' o 'Puntaje a la propensión', el cuál significó un importante avance para los métodos de matching usados en la actualidad y que resuelve el problema de pérdida de datos al no encontrar matching perfectos. El Propensity Score es la probabilidad de que un individuo cualquiera participe del programa de desarrollo en aplicación. La probabilidad o puntaje a la propensión (Propensity Score) es un número entre 0 y 1 que se determina mediante el valor observado de las variables para cada individuo de forma independiente. La estimación del Propensity Score puede ser mediante métodos paramétricos como lo son los modelos de regresión de tipo logit y probit. Sin embargo, se deben asumir supuestos fuertes con respecto a la distribución de probabilidad de los datos o bien se deben 'nor-

malizar' para que se cumplan los supuestos. En la literatura se encuentran disponibles métodos semi paramétricos y no paramétricos para la estimación del Propensity Score. Entre los métodos semi paramétricos Todd (1995) propone el uso de regresión lineal para la estimación de la probabilidad de participación de los individuos y Hahn (1998) propone estimar el Propensity Score mediante el uso de regresión no paramétrica.

Luego, para cada individuo del grupo tratado se determina mediante una medida métrica (como lo puede ser la distancia Euclidea) la distancia de la probabilidad de participar del individuo en tratamiento respecto de todos los individuos del grupo control. Finalmente, se escoge aquel individuo que posea una probabilidad similar o cercana. Los matching que se generan son uno a uno, es decir, cada individuo del grupo tratamiento es emparejado con un individuo del grupo control. El resultado final o el impacto de la política aplicada, es el promedio en la diferencia del resultado observado con respecto al resultado estimado producto del matching. Esto debido a que se desea tener una medida que sea representativa de todo el programa de desarrollo y además porque el impacto es el resultado más importante y por ende el que representa mayor interés para los investigadores (Heckman, 1991).

Los métodos de matching deben lidiar con el problema de sesgo debido a variables no observables y a individuos que no pueden ser emparejados. Rosenbaum y Rubin en su trabajo de 1985 muestran que al combinar una distancia métrica como la distancia de Mahalanobis (la cual considera la covarianza entre las variables) y el puntaje a la propensión se puede reducir el sesgo de variables no observables. Agregando ciertas condiciones propias del programa (por ejemplo, que sólo se puedan emparejar individuos del mismo sexo) se puede disminuir aún más el sesgo. Una vez realizado el matching, el cual es uno a uno, los individuos emparejados no pueden volver a ser usados en el estudio. Esto se conoce como matching sin reemplazo y posee la desventaja de que al sacar individuos ya emparejados del estudio, aumenta la posibilidad de que debamos emparejar individuos que tienen una probabilidad a la propensión muy distinta del resto de individuos que queden en estudio. Esto incrementa el sesgo y además los resultados podrían estar sobrestimados o subestimados ya que el match dependerá del orden en que se emparejen los individuos (Rosenbaum, 1995). Mientras que el uso de reemplazo, es decir, poder un individuo ser emparejado más de una vez nos permite reducir el

sesgo y además el investigador no debe preocuparse del orden en que cada individuo será pareado (Dehejia y Wahba, 1999). A continuación se presentarán dos métodos de matching de vecino más cercano: uno de ellos basado en Propensity Score y otro basado en el valor directo de las covariables y el uso de distancias métricas.

2.3.3.1. Estimador de vecino más cercano basado en Propensity Score (NN)

El estimador de Matching de Vecino Más cercano o 'NN' por sus siglas en inglés (Nearest Neighbour) es un estimador basado en Propensity Score que determina los matching de cada individuo tratado con respecto de un sub grupo 'm' de individuos del grupo no tratado (Becker e Ichino, 2002). Una primera versión de este método puede apreciarse en el trabajo de Rubin (1973), en el cual propone realizar los matching para cada individuo en tratamiento con respecto a aquel individuo del grupo control con características similares.

El estimador trabaja buscando aquellos 'm' individuos más cercanos al individuo con quien buscamos realizar el match. Para ello se determina la distancia entre cada individuo del grupo en tratamiento (en términos de Propensity Score) y los individuos del grupo control. Luego se escogen los 'm' vecinos más cercanos y se procede a realizar el matching. Consideremos i como el índice de cada individuo del grupo tratado y j como los individuos del grupo control. Sea p_i la probabilidad de un individuo i de participar en el programa y p_j la probabilidad de participación de un individuo j del grupo control y a $M(i)$ el conjunto de los 'm' vecinos más cercanos de i . Encontrar los 'm' vecinos más cercanos de i es equivalente a:

$$M(i) = \underset{j}{\text{mín}} \|p_i - p_j\| \quad (2.5)$$

La distancia puede obtenerse en términos de la diferencia en valor absoluto o inducido por una distancia métrica como la de Mahalanobis (Rosenbaum y Rubin, 1985). Luego se estima el 'Efecto en el Tratamiento de los Individuos Tratados' ATT (Average Treatment on Treated), el cual determina el impacto del programa en estudio. ATT se calcula de la siguiente forma:

$$ATT = \frac{1}{N_1} \left(\sum_{i \in N_1} Y_i - \sum_{i \in N_1} \sum_{j \in M(i)} w_{ij} Y_j^M \right) \quad (2.6)$$

donde:

- N_1 : Cantidad de individuos en el grupo de tratamiento.
- Y_i : Resultado del programa para el individuo i .
- Y_j^M : Resultado del programa para el individuo j perteneciente al grupo no tratado del vecindario M .
- w_{ij} : Ponderación de cada individuo j al impacto de i . Se calcula como el recíproco del total de individuos pertenecientes a M .

Para que ATT sea un estimador válido se deben satisfacer las siguientes hipótesis:

Lema 1 *Si $p(x)$ es el puntaje a la propensión, entonces, dado $p(x)$ la variable de tratamiento T es estadísticamente independiente (usualmente denotado como \perp) de la matriz de covariables x :*

$$T \perp x \mid p(x)$$

Lema 2 *Supongamos que la variable respuesta es independiente de la variable de tratamiento dado los valores de la matriz de covariables:*

$$Y(1), Y(0) \perp T \mid x$$

Entonces dada la probabilidad de participación de un individuo al programa, se tiene que la variable respuesta es independiente de la variable de tratamiento:

$$Y(1), Y(0) \perp T \mid p(x)$$

El cumplimiento de estas hipótesis implica que, individuos con valores similares de Propensity Score tendrán la misma distribución de características observables y no observables, independiente de la variable de tratamiento.³

También se define la varianza del estimador ATT como una medida del sesgo por variables no medidas. La varianza se calcula como sigue:

$$Var(ATT) = \frac{1}{N_1} Var(Y_i) + \frac{1}{(N_1)^2} \sum_{j \in M} (w_j)^2 Var(Y_j^M) \quad (2.7)$$

con $w_j = \sum_i w_{ij}$.

La ventaja de éste método de matching respecto de Regresión Discontinua y Diferencia en Diferencias es que permite la disminución del sesgo estadístico provocado por variables no observables. Lo anterior se logra debido a que un individuo del grupo no tratado puede usarse más de una vez para realizar el matching. Esto se conoce como 'matching con reemplazo'. Sin embargo, el método presenta la desventaja de que el número de individuos para realizar cada match depende del criterio del investigador, pues no existe ninguna regla o norma que indique si más individuos o menos individuos generan más o menos sesgo.

2.3.3.2. Estimador de vecino más cercano basado en las covariables

Otra forma en que opera el estimador de vecino más cercano es mediante la aplicación de medidas métricas directamente sobre las covariables, sin aplicar Propensity Score. Basándonos en la notación de Abadie *et al.* (2004), consideraremos un conjunto de individuos $i, i = 1, \dots, N$. Sea $Y_i(1)$ como la respuesta del programa para el individuo i si forma parte del programa y a $Y_i(0)$ como la respuesta del programa si el individuo no hubiese formado parte del programa. Retomando ideas del inciso 2.2. relacionado con inferencia causal, para un individuo i que se somete al programa sólo conocemos $Y_i(1)$, por lo tanto debemos estimar $Y_i(0)$. De forma similar esto aplica a los individuos del grupo control, es decir, no conocemos $Y_i(1)$ para aquellos individuos. Para realizar la estimación se deben escoger los 'm' individuos más cercanos de i . Consideremos a

³Demostración de estos lemas pueden encontrarse en Imbens, (2000).

$M(i)$ el conjunto de los 'm' vecinos más cercanos de i y a j el subíndice de los individuos del grupo control. Encontrar los 'm' vecinos más cercanos es equivalente a:

$$M(i) = \min_j \|x_i - x_j\| = \min_j [(x_i - x_j)'S^{-1}(x_i - x_j)]^{1/2} \quad (2.8)$$

donde S puede ser inducida por la distancia euclidiana, distancia de Mahalanobis o la diagonal de la matriz de varianzas y covarianzas de la matriz de datos x de la siguiente manera:

$$S = \begin{cases} \frac{(x - \bar{x}'1_{n \times 1})'W(x - \bar{x}'1_{n \times 1})}{\sum_{i=1}^n w_i - 1} & ; \text{ si la métrica es Mahalanobis} \\ \text{diag} \left(\frac{(x - \bar{x}'1_{n \times 1})'W(x - \bar{x}'1_{n \times 1})}{\sum_{i=1}^n w_i - 1} \right) & ; \text{ si la métrica es diag(cov)} \\ I_p & ; \text{ si la métrica es euclidiana} \end{cases}$$

donde:

- $1_{n \times 1}$: Vector $n \times 1$ de unos.
- I_p : Matriz identidad de orden p .
- $\bar{x} = (\sum_{i=1}^n w_i x_i) / (\sum_{i=1}^n w_i)$.
- W : Matriz diagonal de dimensión $n \times n$ que contiene los ponderadores w_i ⁴

Luego, se estima para cada individuo del grupo tratamiento la respuesta al programa si no se hubiesen tratado $\hat{Y}_i(0)$ como una media ponderada de los 'm' vecinos más cercanos de i (de forma recíproca, se estima $\hat{Y}_i(1)$ para los individuos del grupo control):

⁴Los ponderadores dependen de la cantidad de vecinos 'm' que el investigador desea usar. Por ejemplo, si se usan 2 vecinos para cada individuo i , w_i será el recíproco de esos dos vecinos, vale decir $1/2 = 0,5$.

$$\hat{Y}_i(T_i) = \begin{cases} Y_i & \text{Si } T_i = T, \text{ para } T \in \{0, 1\} \\ \sum_{j \in M(i)} w_j Y_j & \text{En otro caso} \end{cases}$$

Luego, estimadores para ATE y ATT respectivamente son:

$$\widehat{ATE} = \frac{1}{N}(\hat{Y}_i(1) - \hat{Y}_i(0)) \quad (2.9)$$

$$\widehat{ATT} = \frac{1}{N_1}(\hat{Y}_i(1) - \hat{Y}_i(0)) \quad (2.10)$$

Siguiendo a Abadie *et al* (2004), definiremos a $K_M(i)$ como la cantidad de veces que un individuo i es usado para emparejar a un individuo del grupo contrario. Definido esto, se tiene que la varianza para ATE se puede estimar de la siguiente forma:

$$\widehat{Var}_{ATE} = \frac{1}{N^2} \sum_{i=1}^N (1 + K_M(i))^2 \hat{\sigma}_{ATE} \quad (2.11)$$

Mientras que para ATT, se tiene:

$$\widehat{Var}_{ATT} = \frac{1}{N_1^2} \sum_{i=1}^N (T_i - (1 - T_i)K_M(i))^2 \hat{\sigma}_{ATT} \quad (2.12)$$

con $\hat{\sigma}_{ATE}$ y $\hat{\sigma}_{ATT}$ las varianzas condicionales entre las covariables X_i y la variable de tratamiento T_i .

Para estimar la varianza condicional, se debe asumir que el impacto o efecto individual, $Y_i(1) - Y_i(0)$ es constante y que la varianza condicional entre la respuesta Y_i y la variable de tratamiento T_i condicionado a X_i no varía. Lo anterior nos lleva a que nuestro modelo de inferencia causal presenta homocedasticidad. En base a estos supuestos, podemos calcular la varianza de ATE como:

$$\hat{\sigma}_{ATE} = \frac{1}{2N} \sum_{i=1}^N \left(\frac{1}{M(i)} \sum_{m \in M(i)} [T_i(Y_i - Y_m - \widehat{ATE}) + (1 - T_i)(Y_m - Y_i - \widehat{ATE})]^2 \right) \quad (2.13)$$

Mientras que para ATT se tiene:

$$\hat{\sigma}_{ATT} = \frac{1}{2N_1} \sum_{i:T_i=1} \left(\frac{1}{M(i)} \sum_{m \in M(i)} (Y_i - Y_m - \widehat{ATT})^2 \right) \quad (2.14)$$

Finalmente, la varianza de los estimadores ATE y ATT se pueden calcular reemplazando (2.10) y (2.11) en (2.8) y (2.9) respectivamente.

2.3.3.3. Ejemplo de matching

Consideremos un programa para la reducción del consumo de tabaco de una población. Tengamos en cuenta como variables de interés el peso de cada persona (kg), la edad. Como variable respuesta consideraremos la cantidad de cigarrillos consumidos al día y la variable tratamiento, la cual es binaria y toma el valor de 1 si el individuo i participa en el programa y 0 en otro caso. Los datos se muestra a continuación:

Tabla 2.2: Datos ejemplo de matching

Individuo	Peso	Edad	Consumo de cigarrillos	Tratamiento
1	85	33	23	0
2	99	27	27	0
3	71	62	21	0
4	87	49	30	0
5	67	52	29	0
6	102	39	18	0
7	78	19	19	0
8	84	22	22	0
9	101	57	25	0
10	85	60	26	0
11	74	31	19	1
12	65	30	17	1
13	104	70	15	1
14	83	28	16	1
15	94	56	20	1

El método a usar será el de vecino más cercano basado en covariables y la métrica a usar será la de euclidiana. Escogeremos los dos vecinos más cercanos de cada individuo, en otras palabras $m = 2$. Para cada individuo del grupo tratamiento (Tratamiento = 1) calculamos la distancia euclidiana respecto de los demás individuos del grupo control, usando sólo las variables observables (Peso y Edad). El resultado puede apreciarse en la siguiente tabla:

Tabla 2.3: Ejemplo de matching; Matriz de distancias entre individuos del grupo control y grupo tratados

	Individuo 11	Individuo 12	Individuo 13	Individuo 14	Individuo 15
Individuo 1	11,1803	20,2237	41,5933	5,3852	24,6982
Individuo 2	25,3180	34,1321	43,2897	16,0312	29,4279
Individuo 3	31,1448	32,5576	33,9559	36,0555	23,7697
Individuo 4	22,2036	29,0689	27,0185	21,3776	9,8995
Individuo 5	22,1359	22,0907	41,1461	28,8444	27,2947
Individuo 6	29,1204	38,0789	31,0644	21,9545	18,7883
Individuo 7	12,6491	17,0294	57,2451	10,2956	40,3113
Individuo 8	13,4536	20,6155	52,0000	6,0828	35,4401
Individuo 9	37,4833	45,0000	13,3417	34,1321	7,0711
Individuo 10	31,0161	36,0555	21,4709	32,0624	9,8489

Comenzamos con los emparejamientos. Como la cantidad de vecinos escogidos para el ejemplo es 2, entonces para cada individuo del grupo tratamiento debemos escoger los 2 individuos del grupo control que posean la menor distancia respecto del sujeto a emparejar. De forma similar, debemos hacer lo mismo para cada individuo perteneciente al grupo control. La siguiente tabla resume los emparejamientos a realizar:

Tabla 2.4: Ejemplo de matching: Emparejamientos de cada individuo

Individuo	Grupo	Emparejamiento
1	Control	11, 14
2	Control	11, 14
3	Control	11, 15
4	Control	14, 15
5	Control	11, 12
6	Control	14, 15
7	Control	11, 14
8	Control	11, 14
9	Control	13, 15
10	Control	13, 15
11	Tratamiento	1, 7
12	Tratamiento	1, 7
13	Tratamiento	9, 10
14	Tratamiento	1, 8
15	Tratamiento	9, 10

El siguiente paso es estimar la respuesta no observada para cada individuo y el impacto individual. Recordar que, los individuos del grupo control requieren la estimación de $Y_i(1)$, mientras que los individuos del grupo tratamiento $Y_i(0)$. La respuesta de cada individuo se calcula como el promedio ponderado de las respuestas observadas de los individuos que serán emparejados. Por ejemplo, para el individuo 1 que pertenece al grupo control (Tratamiento = 0), estimaremos $Y_1(1)$ como el promedio de $Y_{11}(1)$ y $Y_{14}(1)$. Mientras que para el individuo 13 que pertenece al grupo de tratamiento (Tratamiento = 1), estimaremos $Y_{13}(0)$ como el promedio de $Y_9(0)$ y $Y_{10}(0)$. La siguiente tabla muestra en resumen todos los cálculos y el impacto individual de todos los individuos:

Tabla 2.5: Ejemplo de matching; Impacto individual de cada participante

Individuo	Emparejamiento	$Y_i(1)$	$Y_i(0)$	Impacto
		A	B	A - B
1	11, 14	$0,5(19+16)=17,5$	23	-5,5
2	11, 14	$0,5(19+16)=17,5$	27	-9,5
3	11, 15	$0,5(19+20)=19,5$	21	-1,5
4	14, 15	$0,5(16+20)=18,0$	30	-12
5	11, 12	$0,5(19+17)=18,0$	29	-11
6	14, 15	$0,5(16+20)=18,0$	18	0
7	11, 14	$0,5(19+16)=17,5$	19	-1,5
8	11, 14	$0,5(19+16)=17,5$	22	-4,5
9	13, 15	$0,5(15+20)=17,5$	25	-7,5
10	13, 15	$0,5(15+20)=17,5$	26	-8,5
11	1, 7	19	$0,5(23+19)=21,0$	-2
12	1, 7	17	$0,5(23+19)=21,0$	-4
13	9, 10	15	$0,5(25+26)=25,5$	-10,5
14	1, 8	16	$0,5(23+22)=22,5$	-6,5
15	9, 10	20	$0,5(25+26)=25,5$	-5,5

Luego, el impacto para todo el programa se calcula como el promedio de los impactos individuales. Esto se conoce como efecto de tratamiento promedio (ATE por sus siglas en inglés). El ATE calculado para nuestro ejemplo nos da como resultado -6 cigarros. Esto quiere decir, que los individuos que recibieron el tratamiento consumen en promedio 6 cigarros menos que los individuos que no recibieron el tratamiento. También es de interés el efecto de tratamiento promedio sobre los individuos tratados. El ATT calculado se obtiene como el promedio del impacto de los individuos tratados, en nuestro caso ATT tiene un valor de -5,7, lo cual significa que los individuos que formaron parte del programa disminuyeron en promedio su consumo de cigarrillos en 5,7 unidades versus si ellos no hubiesen participado del programa.

Otro estadístico de interés es la varianza de los estimadores ATE y ATT. Para ello debemos determinar la cantidad de veces que los individuos fueron usados en los match:

Tabla 2.6: Ejemplo de matching; cantidad de veces que cada individuo es usado en un matching

Individuo	Cantidad de veces usado en matching $K_m(i)$
1	3
2	0
3	0
4	0
5	0
6	0
7	2
8	1
9	2
10	2
11	6
12	1
13	2
14	6
15	5

El primer paso para calcular la varianza mediante (2.8) y (2.9) es determinar la varianza condicional (2.10) y (2.11). Para los datos de nuestro ejemplo, la varianza condicional para ATE y ATT son 12,08 y 6,05 respectivamente. A partir de las varianzas condicionales determinamos que la varianza de cada estimador es de 10,68 y 6,53 para ATE y ATT respectivamente. A partir de los estimadores ATE y ATT, los cuales son un promedio de el efecto del tratamiento sobre la muestra y el grupo de tratamiento, podemos estimar intervalos de confianza para la media de estos. Para ello considerare-

mos un nivel de confianza del 95 %.⁵ El resultado final del match puede apreciarse en la siguiente tabla:

Tabla 2.7: Ejemplo de matching; resumen estadístico del estimador

Estimador	Efecto medio	Error estándar (raíz de la varianza)	Intervalo de confianza del efecto medio
ATE	-6,00	3,27	[-7,65 ; -4,35]
ATT	-5,70	2,55	[-7,94 ; -3,46]

2.4. Implementación de un programa de Evaluación de Impacto

En las secciones anteriores se trataron temas referentes a lo que es la Evaluación de Impacto, qué es lo que busca como objetivo principal y los métodos existentes en la literatura para poder estimar el impacto de un estudio de inferencial causal. La implementación va más allá de todo esto; el poder demostrar que el programa a aplicar es beneficioso o poder inferir sobre el efecto causal de un experimento binario dependen de cómo se haya elaborado el experimento. Por lo tanto, en esta sección mostraremos los pasos a seguir en la implementación que Gertler *et al* (2011) consideran para la exitosa aplicación de un programa de desarrollo:

1. Preparativos:

- Considerar el motivo y que es lo que se evaluará.
- Objetivo general y específicos.
- Hipótesis.
- Seleccionar indicadores.

2. Operación

⁵Para la estimación de un intervalo de confianza para la media se requiere de un tamaño muestral grande, por lo tanto considérese este intervalo simplemente como ejemplo.

- Elegir el diseño experimental.
- Confirmar si el diseño es ético.
- Seleccionar a los involucrados en el experimento.
- Cronograma de actividades.
- Presupuesto para las actividades.

3. Muestreo y análisis de datos

- Determinar el tamaño de muestra y cómo realizar el muestreo.
- Determinar las variables a medir. De suma importancia, determinar la variable respuesta.
- En el caso de programas de políticas públicas o de desarrollo, generar un cuestionario para determinar los individuos que participarán del experimento.
- Procesar los datos, es decir, determinar existencia de datos atípicos y balancear los grupos de control y tratamiento en término de las observaciones.
- Analizar los datos obtenidos.
- Redactar un informe y mostrar los resultados.

Capítulo 3

El estimador BLOPmatching

Este capítulo trata sobre cómo fue modelado el estimador BLOPmatching propuesto por Díaz, Rau y Rivera (2015). Se presentará el modelo en sí y una forma alternativa de este, supuestos a considerar, propiedades y el método para resolver el problema.

3.1. Conceptos formales del estimador

Consideremos que se desea implementar una política pública sobre un grupo de personas, bajo un conjunto de variables medibles (X, Y, T) , con X el conjunto de variables y covariables medidos para cada individuo, Y el resultado para cada individuo una vez finalizado el programa y con T una variable binaria que toma el valor de 1 si el individuo forma parte del grupo de beneficiados y el valor de 0 cuando el individuo no forma parte del grupo beneficiado. Consideraremos K como la dimensión de la matriz de variables y covariables X .

Algunos supuestos a considerar, y que son estándar de los métodos de matching¹son los siguientes:

Supuesto 1. X es un subconjunto convexo y compacto de \mathbb{R}^K .

Supuesto 2. Existe $0 < c < 1$ tal que $0 < P(T = 1 | X) < 1 - c$.

¹Ver Rosenbaum y Rubin (1983) para mayor detalle.

Retomando ideas, consideraremos que cada individuo que no pertenece al programa es indexado por $1, \dots, N_0$, mientras que los que participan del programa son indexados como $N_0 + 1, \dots, N_1$. Luego, $N_0 + N_1 = N$, siendo fácil ver que N es el total de individuos en estudio.

Actualmente, los estimadores consideran un valor M , que representa a los primeros m -vecinos más cercanos respecto del individuo en tratamiento con el cual queremos realizar el matching. Por otro lado, el estimador BLOPmatching presentado por Díaz *et al* (2015) considera que el matching de cada individuo en tratamiento debe ser realizado con cada individuo que no está en tratamiento. Esto implica que existe un grupo de ponderadores para cada individuo en tratamiento. Lo anterior se resuelve mediante el siguiente problema de optimización:

$$\begin{aligned} \min_{\lambda = \lambda_j} \quad & \sum_{j=1}^{N_0} \lambda_j \|X_i - X_j\| \\ \text{s.a} \quad & \\ & \lambda \in \text{argmin}(P_i) \end{aligned} \tag{3.1}$$

donde $\text{argmin}(P_i)$ es el conjunto solución del siguiente problema:

$$\min_{\lambda_j \in \Delta_{N_0}} \left\| X_i - \sum_{j=1}^{N_0} \lambda_j X_j \right\| \tag{3.2}$$

denotando el simplex de dimensión N_0 a Δ_{N_0} , que contiene a los ponderadores λ_j .

Lo anterior se conoce como Bilevel Optimization Model, pues contiene un modelo líder y otro seguidor (Colson *et al*, 2007). El problema (3.1) busca ponderaciones λ_j altas para individuos i del grupo tratados que tengan la menor distancia respecto de individuos j del grupo control. En otras palabras, en (3.1) se busca ponderar alto aquellos individuos que posean características similares, mientras que ponderar bajo aquellos individuos que posean características disímiles a la hora de realizar el matching. Mientras que en (3.2) se busca minimizar la distancia entre un individuo en tratamiento

respecto de la combinación convexa de las características de los individuos del grupo de control. Vale decir, se busca un individuo dentro del grupo control con características óptimas y que sea un buen representante del individuo tratado en el grupo de los individuos no tratados.

3.2. Linealizando el Estimador

Usando como base el problema (3.2), se tiene que encontrar la menor distancia entre X_i y la combinación convexa entre los X_j , con $j \in 1, \dots, N_0$, lo cual es equivalente a determinar el vector proyección de X_i sobre la envoltura convexa formada por las X_j covariables.² De lo anterior se concluye que:

$$\|X_i - \text{proj}(X_i)\| = \min_{\lambda_j \in \Delta_{N_0}} \left\| X_i - \sum_{j=1}^{N_0} \lambda_j X_j \right\| \quad (3.3)$$

siendo $\text{proj}(X_i)$ la proyección del vector X_i sobre la envoltura convexa formada por las características de los individuos X_j del grupo opuesto.

De la relación obtenida en (3.3) aplicada sobre el problema (3.1) se obtiene el siguiente problema de programación lineal:

$$\begin{aligned} & \min_{\lambda = \lambda_j} \sum_{j=1}^{N_0} \lambda_j \|X_i - X_j\| & (3.4) \\ & \text{s.a} \\ & \sum_{j=1}^{N_0} \lambda_j X_j = \text{proj}(X_i) \\ & \sum_{j=1}^{N_0} \lambda_j = 1 \\ & \lambda_j \geq 0 \quad \forall j = 1, \dots, N_0 \end{aligned}$$

El principal problema es determinar la proyección de X_i sobre la envoltura convexa

²La envoltura convexa de X_j es el poliedro convexo más pequeño que puede ser formado con los X_j covariables (Berg *et al*, 2008).

de las covariables X_j , lo cual puede ser resuelto aplicando métodos disponibles en la literatura de optimización (Botkin y Stoer, 2005).

La solución del problema anterior viene dado por el conjunto:

$$\lambda(i) = (\lambda_1(i), \dots, \lambda_{N_1-T_i}(i)) \quad (3.5)$$

El impacto individual estimado para cada individuo se determina como sigue:

$$\hat{Y}_i = T_i \sum_{j=1}^{N_0} \lambda_j(i) Y_j + (1 - T_i) \sum_{j=1}^{N_1} \lambda_j(i) Y_{N_0+j} \quad (3.6)$$

Una vez que cada individuo del grupo ha sido emparejado por separado con la totalidad de los individuos del grupo control es que se debe de realizar lo opuesto. Vale decir, se debe hacer el match entre cada individuo del grupo control por separado con la totalidad de los individuos del grupo beneficiado. Lo anterior implica que se deben resolver $N_0 + N_1 = N$ problemas de optimización.

3.3. Estimando el impacto de una política pública mediante BLOPmatching

Como base se tomarán ideas de la sección 2.2.2., para explicar el cómo BLOPmatching estima el efecto en el tratamiento promedio y el efecto en el tratamiento de los individuos tratados. Ambos, ATE y ATT respectivamente son estimadores del impacto de un programa. ATE mide el efecto promedio considerando a todos los individuos en estudio, vale decir, los que pertenecen al grupo en tratamiento y a los del grupo control. Mientras que ATT mide el efecto del programa sólo considerando a los individuos que pertenecen al grupo de tratamiento.

Dado que se quiere minimizar el sesgo generado por variables no observables, cambios en variables con respecto al tiempo o por tendencias en las variables propias del comportamiento de cada subpoblación en estudio es que para los individuos que son tratados se debe estimar el resultado Y correspondientes a si ellos no hubiesen sido tratados. Mientras que a los individuos del grupo control se debe estimar el resultado

de si hubiesen recibido el tratamiento. Considerando lo anterior, se tiene que:

$$\hat{Y}_i(0) = \begin{cases} Y_i & \text{si } T_i = 0 \\ \hat{Y}_i & \text{si } T_i = 1 \end{cases}$$

$$\hat{Y}_i(1) = \begin{cases} \hat{Y}_i & \text{si } T_i = 0 \\ Y_i & \text{si } T_i = 1 \end{cases}$$

Así, el estimador de matching para ATE y ATT son respectivamente:

$$ATE = \frac{1}{N} \sum_{i=1}^N (\hat{Y}_i(1) - \hat{Y}_i(0)) \quad (3.7)$$

$$ATT = \frac{1}{N_1} \sum_{i=1}^N (\hat{Y}_i(1) - \hat{Y}_i(0)) \quad (3.8)$$

Para estimar la varianza, siguiendo el trabajo de Abadie e Imbens (2006), se tiene que son estimadores robustos de la varianza para ATE y ATT las siguientes relaciones:

$$\hat{V}(ATE) = \frac{1}{N} \sum_{i=1}^N \left((\hat{Y}_i(1) - \hat{Y}_i(0) - \widehat{ATE})^2 + ((1 + c_i^{[1]})^2 - (1 + c_i^{[2]})) \sigma_i^2 \right) \quad (3.9)$$

$$\hat{V}(ATT) = \frac{1}{N_1} \sum_{i=1}^N \left(W_i (\hat{Y}_i(1) - \hat{Y}_i(0) - \widehat{ATT})^2 + (1 - W_i) ((c_i^{[1]})^2 - c_i^{[2]}) \sigma_i^2 \right) \quad (3.10)$$

con:

- $c_i^{[a]} = \sum_{j=1}^{N_0} (\lambda_{i-N_0}(j))^a$, si i pertenece al grupo control.
- $c_i^{[a]} = \sum_{j=1}^{N_1} (\lambda_i(N_0 + j))^a$ si i pertenece al grupo tratado.
- $\sigma_i^2 = \sigma^2(X_i, W_i)$, es la varianza condicional.

Estimar la varianza condicional σ_i^2 implica resolver el problema (3.4) para cada individuo i del grupo control, emparejando a aquel individuo con el resto de individuos de su mismo grupo. De forma similar si el individuo pertenece al grupo de tratamiento. Lo anterior se resume de la siguiente manera:

$$\begin{aligned}
 & \min_{\phi=\phi_j} \sum_{j=1}^{N_0} \phi_j \|X_i - X_j\| & (3.11) \\
 & \text{s.a} \\
 & \sum_{j=1}^{N_0} \phi_j X_j = \text{proj}(X_i) \\
 & \sum_{j=1}^{N_0} \phi_j = 1 \\
 & \phi_j \geq 0 \quad \forall j = 1, \dots, N_0
 \end{aligned}$$

La solución al nuevo problema (3.11) es el vector $\phi(i) = \{\phi_1(i), \dots, \phi_{i-1}(i), \phi_{i+1}(i), \dots, \phi_{n_0}(i)\}$, que es el conjunto de ponderadores del proceso de matching intragrupos. Con lo cual es fácil ver que $\phi_i(i) = 0$. Mientras que si i es un individuo del grupo en tratamiento, el vector solución está compuesto por $\phi(i) = \{\phi_{n_0+1}(i), \dots, \phi_{i-1}(i), \phi_{i+1}(i), \dots, \phi_{n_0+n_1}(i)\}$. Dado eso, se define:

$$\rho(i) = (1 - T_i) \sum_{j=1}^{N_0} [\phi_j(i)]^2 + T_i \sum_{j=1}^{N_1} [\phi_{N_0+j}(i)]^2 \quad (3.12)$$

La relación anterior nos servirá para estimar la varianza condicional. De lo anterior, se tiene:

$$\hat{\sigma}^2(i) = \frac{1 - T_i}{1 + \rho(i)} \left(Y_i - \sum_{j=1}^{N_0} \phi_j(i) Y_j \right)^2 + \frac{T_i}{1 + \rho(i)} \left(Y_i - \sum_{j=1}^{N_1} \phi_j(i) Y_{N_0+j} \right)^2 \quad (3.13)$$

Finalmente, reemplazando (3.13) en (3.9) y (3.10) se tiene un estimador para la varianza de ATE y ATT respectivamente.

3.4. Ejemplo

Consideremos los datos del ejemplo de matching de la sección 2.3.3.3. En resumen, los datos constaban de 15 individuos, de los cuales 5 recibieron tratamiento y 10 no. Las variables independientes son 'Peso' y 'Edad', la variable respuesta es 'Consumo de cigarros' y finalmente la variable de tratamiento es 'Tratamiento'. El estimador BLOP-matching busca la ponderación óptima de cada individuo a emparejar con respecto de los demás, minimizando la distancia entre ellos y mediante las restricciones de igualdad buscando individuos con valores de covariables similares. Para ello, se resuelve el problema (3.4) para cada individuo del grupo tratamiento con respecto a todos los individuos del grupo control. De similar forma, se resuelve (3.4) para los individuos del grupo control. Como el problema posee dos dimensiones, podemos graficar los datos para tener una idea de cuales podría ser posibles emparejamientos en términos de la distancia entre ellos.

Así, a simple vista deberíamos esperar por ejemplo para el individuo 12 encontrar emparejamientos con los individuos 7 y 5, ya que estos son los más cercanos. Mientras que para el individuo 15 podríamos encontrar emparejamientos con los individuos 10, 4 y 9 (ver figura 3.1).

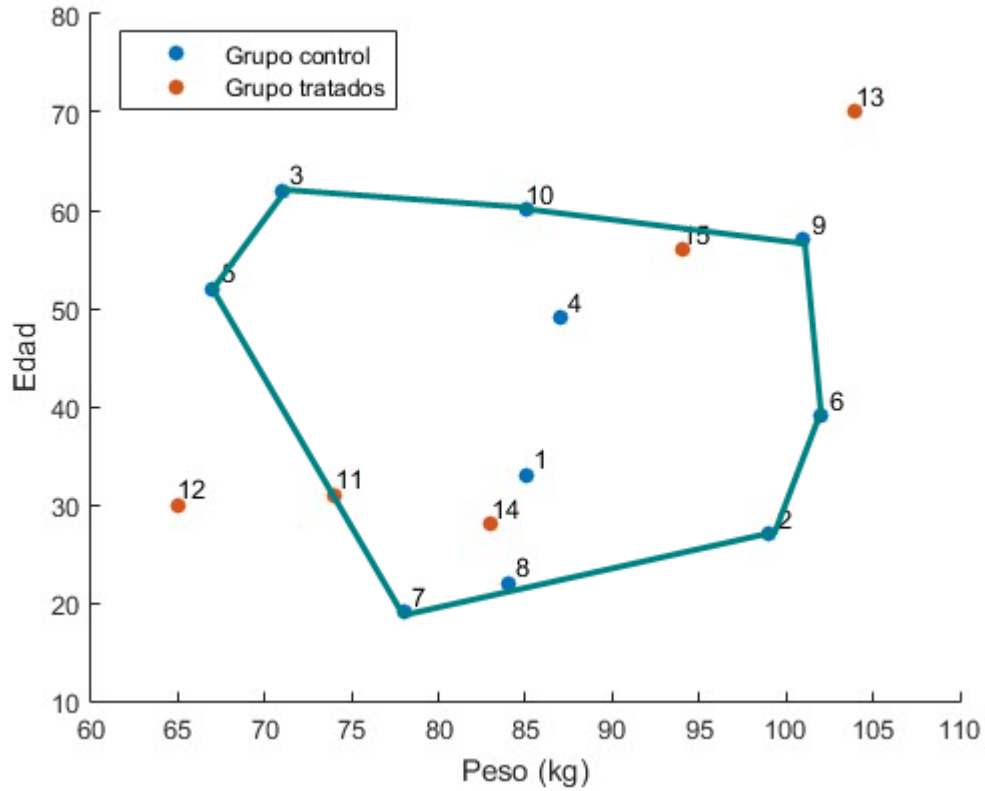


Figura 3.1: Ejemplo BLOPmatching; Distribución de los datos en el plano.

Haciendo uso del script `blopmatch.m` (De elaboración propia. Ver capítulo 4.2. para ver su desarrollo y modelado) procederemos a resolver el problema 3.4 para cada individuo del grupo en tratamiento. Los resultados son los siguientes:

Tabla 3.1: Ejemplo BLOPmatching; Grupo de ponderadores para cada individuo del grupo tratamiento

	Individuos i del grupo de tratamiento				
	11	12	13	14	15
$\lambda_1(i)$	0,0000	0,0000	0,0000	0,6190	0,0000
$\lambda_2(i)$	0,0000	0,0000	0,0000	0,0000	0,0000
$\lambda_3(i)$	0,0000	0,0000	0,0000	0,0000	0,0000
$\lambda_4(i)$	0,0000	0,0000	0,0000	0,0000	0,2176
$\lambda_5(i)$	0,3636	0,3333	0,0000	0,0000	0,0000
$\lambda_6(i)$	0,0000	0,0000	0,0000	0,0000	0,0000
$\lambda_7(i)$	0,6364	0,6667	0,0000	0,2698	0,0000
$\lambda_8(i)$	0,0000	0,0000	0,0000	0,1111	0,0000
$\lambda_9(i)$	0,0000	0,0000	1,0000	0,0000	0,5353
$\lambda_{10}(i)$	0,0000	0,0000	0,0000	0,0000	0,2471

Así, el individuo 11 es emparejado en un 36,36 % con el individuo 5 y en un 63,64 % con el individuo 7. El individuo 12 es emparejado en un 33,33 % con el individuo 5 y en un 66,67 % con el individuo 7. Se genera un matching uno a uno entre los individuos 13 y 9. El individuo 14 es emparejado con los individuos 1, 7 y 8 en un 61,90 %, 26,98 % y 11,11 % respectivamente. Finalmente, para el individuo 15, éste es emparejado con los individuos 4, 9 y 10 en un 21,76 %, 53,53 % y 24,71 % respectivamente. Los resultados obtenidos coinciden con lo observado en la figura 3.1; los individuos son emparejados con aquellos más cercanos y que posean características similares balanceando el matching. Del mismo modo, se deben realizar los matching para los individuos del grupo de control. Los resultados se pueden ver a continuación:

Tabla 3.2: Ejemplo BLOPmatching; Grupo de ponderadores para cada individuo del grupo control

	Individuos i del grupo control									
	1	2	3	4	5	6	7	8	9	10
$\lambda_{11}(i)$	0,0000	0,0000	0,0000	0,1024	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
$\lambda_{12}(i)$	0,0000	0,0000	0,2000	0,1707	0,4500	0,0000	0,2778	0,0000	0,0000	0,2500
$\lambda_{13}(i)$	0,0079	0,0000	0,8000	0,0000	0,5500	0,2619	0,0000	0,0000	0,6905	0,7500
$\lambda_{14}(i)$	0,8254	1,0000	0,0000	0,0000	0,0000	0,7381	0,7222	1,0000	0,3095	0,0000
$\lambda_{15}(i)$	0,1667	0,0000	0,0000	0,7268	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000

A partir de los ponderadores que se obtuvieron y pueden observarse en la tabla 3.1 y 3.2 procedemos a calcular la respuesta no observada para cada individuo. Para los individuos del grupo tratamiento estimamos como respuesta no observable $Y_i(0)$, mientras que para el grupo de control $Y_i(1)$. La respuesta no observada se calcula como una media ponderada de los $\lambda_j(i)$ obtenidos. Así, por ejemplo, para el individuo 1 que pertenece al grupo control, se tiene que la respuesta no observable es $\hat{Y}_1(1) = 0 \cdot 19 + 0 \cdot 17 + 0,0070 \cdot 15 + 0,8254 \cdot 16 + 0,1667 \cdot 20 = 16,65$. Siguiendo el procedimiento mostrado, se tienen las respuestas para los individuos del grupo tratamiento en la siguiente tabla:

Tabla 3.3: Ejemplo BLOPmatching; Estimación de la respuesta no observable de los individuos del grupo tratamiento

Individuo i	$\hat{Y}_i(0)$
11	22,63
12	22,33
13	25,00
14	21,80
15	26,33

Mientras que para los individuos del grupo control se tiene el siguiente resultado:

Tabla 3.4: Ejemplo BLOPmatching; Estimación de la respuesta no observable de los individuos del grupo control

Individuo i	$\widehat{Y}_i(1)$
1	16,65
2	15,99
3	15,40
4	19,38
5	15,90
6	15,73
7	16,27
8	16,00
9	15,30
10	15,50

Ahora procedemos a estimar el impacto de cada individuo en el programa. Retomando ideas de la sección de matching, el impacto se estima como la diferencia en la respuesta de un individuo cuando recibe el programa y la respuesta del mismo individuo cuando no recibe el programa. Para los individuos del grupo tratamiento, esto se resume como $Y_i(1) - \widehat{Y}_i(0)$. Mientras que para los individuos del grupo control se tiene $\widehat{Y}_i(1) - Y_i(0)$. Los resultados pueden apreciarse en las siguientes dos tablas:

Tabla 3.5: Ejemplo BLOPmatching; Impacto individual de cada individuo en el grupo tratamiento

Individuo i	$Y_i(1) - \widehat{Y}_i(0)$
11	-3,6364
12	-5,3333
13	-10,0000
14	-5,8095
15	-6,3353

Tabla 3.6: Ejemplo BLOPmatching; Impacto individual de cada individuo en el grupo control

Individuo i	$\widehat{Y}_i(1) - Y_i(0)$
1	-6,3413
2	-11,0000
3	-5,6000
4	-10,6146
5	-13,1000
6	-2,2619
7	-2,7222
8	-6,0000
9	-9,6905
10	-10,5000

Con los impactos individuales calculados, podemos obtener una estimación para ATE y ATT. Para ATE se tiene que puede ser estimado como la media simple entre los impactos de todos los individuos en estudio, esto quiere decir que consideramos grupo control y de tratamiento. El resultado de esto da $\widehat{ATE} = -7,263$. Por lo tanto, los individuos que fueron sometidos al tratamiento disminuyeron su consumo de cigarrillos en promedio un 7,263 cigarros respecto de aquellos que no se sometieron al programa. ATT se calcula como el promedio simple entre el impacto individual de los individuos que se sometieron al programa; de lo anterior obtenemos como resultado $\widehat{ATT} = -6,2229$. Esto quiere decir que aquellos individuos que se sometieron al tratamiento disminuyeron su consumo de cigarros en un promedio de 6,2229 cigarros respecto a que si estos individuos no hubiesen participado en el programa. En palabras simples, el programa da resultados.

Cada estimador presentado posee una medida de la desviación o error estándar. Para estimar cada uno de estos estadísticos debemos resolver el problema (3.11) para cada individuo en estudio, emparejándolos con individuos de su mismo grupo. Por ejemplo, para el individuo 3 que pertenece al grupo control, debemos emparejarlo con

los individuos 1, 2, 4, 5, 6, 7, 8, 9 y 10, los cuales también pertenece al grupo no tratado. Los resultados de resolver (3.11) para cada grupo se pueden apreciar en las siguientes tablas:

Tabla 3.7: Ejemplo BLOPmatching; Ponderadores intragrupos del grupo en tratamiento

	Individuos i del grupo tratamiento				
	11	12	13	14	15
$\phi_{11}(i)$	-	1,0000	0,0000	1,0000	0,2857
$\phi_{12}(i)$	0,5418	-	0,0000	0,0000	0,0000
$\phi_{13}(i)$	0,0000	0,0000	-	0,0000	0,6463
$\phi_{14}(i)$	0,3897	0,0000	0,0000	-	0,0680
$\phi_{15}(i)$	0,0684	0,0000	1,0000	0,0000	-

Tabla 3.8: Ejemplo BLOPmatching; Ponderadores intragrupos del grupo control

	Individuos i del grupo control									
	1	2	3	4	5	6	7	8	9	10
$\phi_1(i)$	-	0,0000	0,0000	0,3159	0,0000	0,0000	0,0000	0,0630	0,0000	0,0000
$\phi_2(i)$	0,0000	-	0,0000	0,0000	0,0000	0,6000	0,0000	0,2647	0,0000	0,0000
$\phi_3(i)$	0,0000	0,0000	-	0,0000	0,7674	0,0000	0,0000	0,0000	0,0000	0,5333
$\phi_4(i)$	0,3953	0,0000	0,0000	-	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
$\phi_5(i)$	0,0109	0,0000	0,7778	0,0000	-	0,0000	0,0000	0,0000	0,0000	0,0000
$\phi_6(i)$	0,0000	0,8333	0,0000	0,1176	0,0000	-	0,0000	0,0000	0,1429	0,0000
$\phi_7(i)$	0,0000	0,0000	0,0000	0,0000	0,2326	0,0000	-	0,6723	0,0000	0,0000
$\phi_8(i)$	0,5938	0,1667	0,0000	0,0000	0,0000	0,0000	1,0000	-	0,0000	0,0000
$\phi_9(i)$	0,0000	0,0000	0,0000	0,0000	0,0000	0,4000	0,0000	0,0000	-	0,4667
$\phi_{10}(i)$	0,0000	0,0000	0,2222	0,5664	0,0000	0,0000	0,0000	0,0000	0,8571	-

Con los ponderadores $\phi_j(i)$ reemplazamos en (3.12) para calcular $\rho(i)$. Esto nos permitirá estimar la varianza condicional $\hat{\sigma}(i)^2$ para cada individuo en estudio, la cual es fundamental para la estimación de la varianza de ATE y ATT. Los resultados para $\rho(i)$ y la varianza condicional $\hat{\sigma}(i)^2$ se muestran en la siguiente tabla:

Tabla 3.9: Ejemplo BLOPmatching; Varianza condicional de los individuos

Individuo i	$\rho(i)$	Varianza condicional $\widehat{\sigma}(i)^2$
1	0,5090	3,3211
2	0,7222	40,3226
3	0,6543	32,5075
4	0,4345	24,1750
5	0,6431	43,6129
6	0,5200	44,2368
7	1,0000	4,5000
8	0,5260	0,2603
9	0,7551	0,0116
10	0,5022	6,5355
11	0,4502	3,2905
12	1,0000	2,0000
13	1,0000	12,5000
14	1,0000	4,5000
15	0,5039	9,5467

Con la varianza condicional de cada individuo podemos estimar la varianza para ATE y ATT. Reemplazando los valores de $\widehat{\sigma}(i)^2$ en las ecuaciones 3.9 y 3.10 obtenemos como resultado que la varianza de ATE es de 3,074, mientras que para ATT se tiene una varianza de 1,342. Si comparamos con el ejemplo realizado en 2.3.3.3, se puede apreciar que el encontrar la ponderación óptima de cada individuo en el proceso de emparejamiento impacta de forma directa y positiva sobre la estimación de la varianza, la cual con BLOPmatching es menor.

Capítulo 4

Programación, diseño e implementación de BLOPmatching en Matlab

Este capítulo trata sobre la programación y desarrollo de scripts y archivos necesarios para el funcionamiento de BLOPmatching, en entorno Matlab. Se tratan temas de la sintaxis de programación del lenguaje, uso de pseudocódigo para el diseño previo a la programación, funcionamiento de las distintas funciones del código programado, interfaz de usuario para BLOPmatching y cómo se debe usar el programa para evaluar estudios de causa efecto o programas de desarrollo.

4.1. Pseudocódigo de BLOPmatching

BLOPmatching es un estimador que requiere de cálculos computacionales complejos, lo cual justifica el uso de recursos computacionales para estimar el impacto de un programa de desarrollo. A continuación, se presenta el pseudocódigo de BLOPmatching, el cual realiza los siguientes pasos principales:

- Construir los grupos de tratamiento y control en función de la variable binaria que determina si un individuo recibe tratamiento o no.

- Cálculo de la distancia euclidiana de cada individuo del grupo tratamiento con respecto a los individuos del grupo control. El resultado es un vector con los coeficientes de la función objetivo del modelo (3.4).
- Resolver el problema (3.4) para cada individuo del grupo en tratamiento con respecto a todos los individuos del grupo control. Lo mismo para cada individuo del grupo control con respecto a todos los individuos del grupo de tratamiento.
- Estimar el resultado de cada individuo y calcular la diferencia entre resultado estimado y resultado observado.
- Calcular ATE y/o ATT.
- Resolver el problema (3.11) para cada individuo del grupo en tratamiento con respecto a los demás individuos de su mismo grupo. Lo mismo para cada individuo del grupo control con respecto al resto de individuos de su mismo grupo.
- Estimar la varianza condicional de cada individuo mediante la ponderación óptima obtenida de (3.11).
- Estimar la varianza de ATE y/o ATT a partir de los resultados de varianza condicional de cada individuo.

```

1 % Pseudocodigo de BLOPmatching
2
3 % Entradas
4 %X: Matriz de variables y covariables
5 %Y: Variable respuesta o dependiente
6 %T: Variable binaria de tratamiento
7
8 % Salidas
9 %Lambda0: Matriz con las ponderaciones de cada match del grupo control
.
10 %Lambda1: Matriz con las ponderaciones de cada match del grupo tratado
.
11 %ATE: Estimador de matching para ATE

```

```

12 %ATT: Estimador de matching para ATT
13 %VAR(ATE): Varianza del estimador ATE
14 %VAR(ATT): Varianza del estimador ATT
15
16 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
17
18 function [ATE, ATT] = blopmatching(X, Y, T)
19
20 %Determinar los grupos de tratamiento y control:
21
22 [n,k] = size(X)
23
24 n0 = 0;
25 n1 = 0;
26
27 for i = 1 : n
28     if Y(i) == 0
29         n0 = n0 + 1;
30     else
31         n1 = n1 + 1;
32     end
33 end
34
35 %Ordenamos los datos segun el valor de la variable de tratamiento
36
37 X = sort(X, T)
38
39 %Construyendo cada grupo
40
41 X_n0 = X(1:n0)
42 X_n1 = X(n0 + 1: n1)
43
44 %Calculamos las distancias euclidianas entre cada individuo respecto del
45     otro
46
47 distancias = euc_distance(X_n0, X_n1)

```

```

48 Resolvemos el problema 3.4 para los individuos del grupo control
49
50 for i = 1 : n0
51     funcion_objetivo = distancias(i,:);
52     Aeq = [transpuesta(X_n1), ones(n1)];
53     rhs = [X_n0(i) , 1];
54     lambda0(i) = linprog(funcion_objetivo , Aeq, rhs)
55 end
56
57 % Resolvemos el problema 3.4 para los individuos del grupo tratamiento
58
59 for i = 1 : n1
60     funcion_objetivo = distancias(:,i);
61     Aeq = [X_01' , ones(n0)];
62     rhs = [X_n1(i) , 1];
63     lambda1(i) = linprog(funcion_objetivo , Aeq, rhs)
64 end
65
66 % Calculamos la respuesta estimada de cada individuo
67
68 for i = 1 : n0
69     resp_est0(i) = lambda0(i)*Y(i)';
70 end
71
72 for i = 1 : n1
73     resp_est1(i) = lambda1(i)*Y(n0 + i)';
74 end
75
76 % Con la respuesta estimada, se calcula el impacto individual de cada
    participante
77
78 for i = 1 : n0
79     Y_est0(i) = Y(i) - resp_est0(i);
80 end
81
82 for i = 1 : n1
83     Y_est1(i) = Y(n0 + i) - resp_est1(i);

```

```

84 end
85
86 Y_est = [Y_est0; Y_est1];
87
88 % Estimando ATT y ATE
89
90 ATT = (1/. n1)*Y_est;
91 ATE = (1/. n)*Y_est;
92
93 % Estimando el poder de alfa
94
95 pow_x0_1 = sum(lambda0);
96 pow_x0_2 = sum(lambda0).^2;
97 pow_x1_1 = sum(lambda1);
98 pow_x1_2 = sum(lambda1).^2;
99
100 pow1 = [pow_x0_1; pow_x1_1];
101 pow2 = [pow_x0_2; pow_x1_2];
102
103 % Resolviendo 3.11 para cada individuo del grupo control
104
105 distancias = euc_distance(X_n0, X_n0);
106
107 for i = 1 : n0
108     funcion_objetivo = distancias(i,:);
109     Aeq = [[X_n1(1:i-1) X_n1(i+1:n0)], ones(n1)];
110     rhs = [X_n0(i) , 1];
111     phi0(i) = linprog(funcion_objetivo , Aeq, rhs)
112 end
113
114 % Resolviendo 3.11 para cada individuo del grupo tratamiento
115
116 distancias = euc_distance(X_n1, X_n1);
117
118 for i = 1 : n1
119     funcion_objetivo = distancias(i,:);
120     Aeq = [[X_n0(1:i-1) X_n0(i+1:n1)], ones(n0)];

```

```

121   rhs = [X_n1(i) , 1];
122   phi1(i) = linprog(funcion_objetivo , Aeq, rhs)
123 end
124
125 rho0 = sum(phi0.^2);
126 rho1 = sum(phi1.^2);
127
128 % Varianza condicional ...
129
130 for i = 1 : n0
131   sigma0 = (1/(1 + rho0(i, :))) * ((y1(i, 1) - [y1(1:i-1) y1(i+1:n1)] * phi0(:, i)
132 end
133
134 for i = 1 : n1
135   sigma1 = (1/(1 + rho1(i, :))) * ((y0(i, 1) - [y0(1:i-1) y0(i+1:n0)] * phi1(:, i)
136 end
137
138 sigma = [sigma0; sigma1];
139
140 % Calcular la varianza de ATE y ATT
141
142 for i = 1: n0+n1
143   ATE_var(i, 1) = (1+pow1(i, 1)).^2 * sigma(i, 1) + (y_est(i, 1) - ATE).^2 - (1+pow2(
       i, 1)) * sigma(i, 1);
144 end
145
146 for i = 1: n1
147   ATT_var(i, 1) = (y_est(i+n0, 1) - ATT).^2 + ((power1(i, 1).^2 - power2(i, 1)) *
       sigma(n0+i, 1));
148 end
149
150 Varianza_ATE = sum(ATE_var) / ((n0+n1).^2);
151 Varianza_ATT = sum(ATT_var) / ((n1).^2);

```

A continuación se presentan observaciones sobre el uso de funciones propias de Matlab en el pseudocódigo presentado:

Tabla 4.1: Funciones de Matlab utilizadas en el pseudocódigo

Función de Matlab	Línea(s)	Uso de la función
size(x)	22	Función que obtiene la cantidad de filas y columnas de un arreglo; en nuestro caso la cantidad de observaciones y variables.
sort(x,t)	37	Ordena los valores de la matriz 'x' en función de los valores de la variable de tratamiento 't'.
euc_distance(Xn0, Xn1)	46, 205 y 116	Función de elaboración propia la cual calcula la distancia euclidiana entre cada individuo Xn0 del grupo control con respecto a cada individuo Xn1 del grupo tratamiento.
Apóstrofe	52, 61, 69 y 73	Obtiene la inversa de un arreglo matricial.
ones(n,k)	52 y 61	Función que genera un arreglo de n x m, cuyos elementos son solo unos. Representa los coeficientes de lambda para el problema (3.4) y el lado derecho de la restricción asociada a que la suma de los ponderadores debe ser igual a 1.
linprog(fobj, Aeq, rhs)	54, 63, 111 y 122	Solver que viene por defecto en el Optimization Toolbox para Matlab. Permite resolver modelos de programación lineal.

4.2. Script `blopmatch.m`

`blopmatch.m` es una función escrita en Matlab del estimador BLOPmatching presentado en esta memoria.

La función contiene la siguiente estructura:

$$[\text{results}] = \text{blopmatch}(x, y, t, \text{weights})$$

Donde las entradas son:

- **x:** Matriz de variables del conjunto de individuos; `blopmatch` admite como entrada arreglos matriciales, tablas o arreglos de celdas.
- **y:** Vector con la variable respuesta de los individuos. Se admite como entrada arreglos vectoriales, la columna de una tabla o la columna de un arreglo de celdas.
- **t:** Vector de tratamiento de los individuos. Se admite como entrada arreglos vectoriales, la columna de una tabla o la columna de un arreglo de celdas.
- **weights:** Si el parámetro toma el valor de 1 entonces se guardan los ponderadores del matching de los individuos del grupo control y tratamiento en un archivo con extensión `.mat`.

La salida es `'results'`, el cual es un objeto de Matlab de tabla formateada que contiene la estimación de ATE y ATT, como también el error estándar de cada estimador, el estadístico z y su valor p.

4.2.1. Ejemplo en el uso de la función BLOPmatching

Consideremos los datos de los ejemplos de matching y BLOPmatching. Los datos se encuentran en un archivo llamado `'fumadores.xls'`. Las primeras dos columnas contienen las variables `'Peso'` y `'Edad'`, la tercera columna contiene la variable de respuesta `'Consumo de cigarrillos'`, mientras que la cuarta columna contiene la variable de tratamiento.

	A	B	C	D
1	Peso	Edad	consumo	Tratamiento
2	85	33	23	0
3	99	27	27	0
4	71	62	21	0
5	87	49	30	0
6	67	52	29	0
7	102	39	18	0
8	78	19	19	0
9	84	22	22	0
10	101	57	25	0
11	85	60	26	0
12	74	31	19	1
13	65	30	17	1
14	104	70	15	1
15	83	28	16	1
16	94	56	20	1

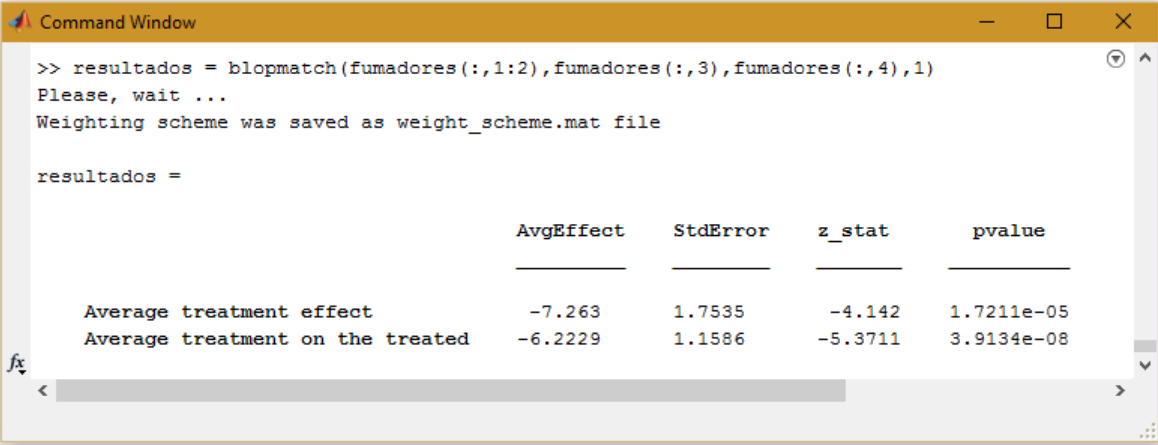
Figura 4.1: Ejemplo blopmatch.m; Orden de los datos

Como la función fue diseñada para funcionar con arreglos matriciales, tablas y arreglos de celdas, entonces no debemos preocuparnos de cómo se importarán los datos. La única consideración es que si los datos son importados como vectores, entonces para la entrada x del script, los vectores a ingresar deben estar concatenados en un arreglo matricial. En nuestro caso, los datos son importados como tabla y son guardados en el workspace de Matlab con el nombre del archivo de Excel. Luego, las entradas del programa quedan de la siguiente forma:

```
resultados = blopmatch(fumadores(:,1:2),fumadores(:,3),fumadores(:,4), 1)
```

Así, 'resultados' es una variable que guardará las salidas de blopmatch, fumadores(:, 1 : 2) indica que de la tabla 'fumadores' se usen todas las filas y las columnas 1 y 2 que son las variables 'Peso' y 'Edad'; fumadores(:, 3) indica que se use la variable 'consumo' como respuesta y finalmente fumadores(:, 4) nos indica que en la columna 4 del arreglo fumadores se encuentra la variable de tratamiento. Además, como la entrada que pertenece a 'weights' contiene un 1 le estamos indicando a blopmatch.m que guarde los ponderadores en un archivo de Matlab, el cual se llama por defecto

'weight_scheme.mat'. Una vez finalizado todos los cálculos, los resultados se mostrarán en la ventana de comandos:



```
>> resultados = blopmatch(fumadores(:,1:2),fumadores(:,3),fumadores(:,4),1)
Please, wait ...
Weighting scheme was saved as weight_scheme.mat file

resultados =
```

	AvgEffect	StdError	z_stat	pvalue
Average treatment effect	-7.263	1.7535	-4.142	1.7211e-05
Average treatment on the treated	-6.2229	1.1586	-5.3711	3.9134e-08

Figura 4.2: Ejemplo blopmatch.m; Salidas del Script

4.3. Interfaz de usuario de BLOPmatching

Adicional al programa BLOPmatching.m se presenta una versión con interfaz de usuario del script, con el objetivo de que la utilización del método BLOPmatching sea más sencilla. La interfaz fue programada en Matlab versión 2015b.

4.3.1. Interfaz principal

El programa contiene una interfaz principal la cual puede apreciarse en la siguiente figura:

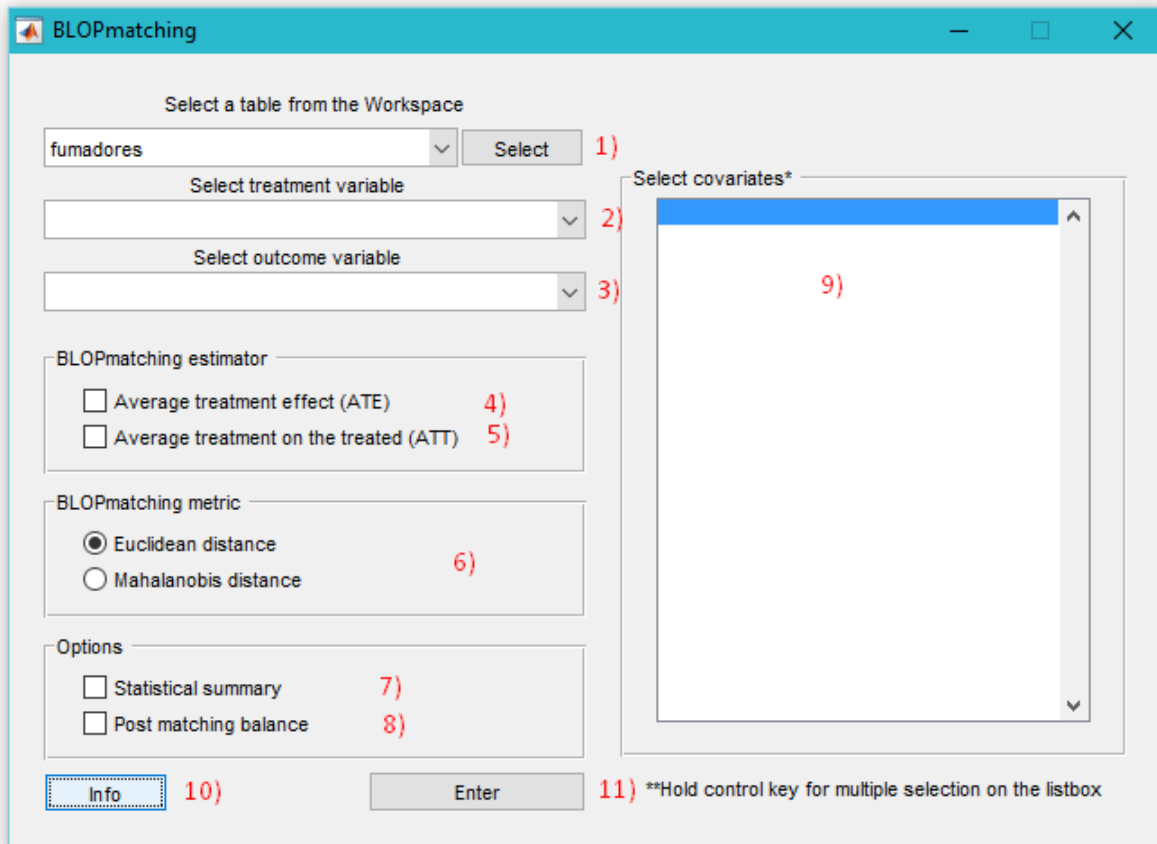


Figura 4.3: Interfaz de usuario BLOPmatching

La interfaz contiene los siguientes elementos:

1. **Workspace pop-up:** Menu para seleccionar los datos desde el workspace de Matlab. Los datos deben ser importados como un arreglo tipo tabla hacia el workspace para luego ser cargados.
2. **Select treatment variable pop-up:** Botón de tipo menu para seleccionar la variable de tratamiento dentro de los datos cargados.
3. **Select outcome variable pop-up:** Botón de tipo menu para seleccionar la variable respuesta dentro de los datos cargados.

4. **Average treatment effect checkbox:** Botón de tipo caja que indica al programa que calcule el estimador ATE.
5. **Average treatment on the treated checkbox:** Botón de tipo caja que indica al programa que calcule el estimador ATT.
6. **BLOPmatching metric button group:** Botones circulares para seleccionar la métrica del estimador.
7. **Statistical summary checkbox:** Botón tipo caja que entrega un resumen estadístico de los datos.
8. **Post matching balance checkbox:** Botón tipo caja que entrega un balance post matching.
9. **Select covariates listbox:** Cuadro de lista que permite seleccionar las variables que serán usadas para emparejar. Para selección de múltiples variables se debe mantener presionada la tecla ctrl.
10. **Info button:** Botón que muestra un cuadro de mensaje con información relacionada al programa.
11. **Enter button:** Botón que inicia el proceso de matching para las opciones y variables seleccionadas.

4.3.2. Ejemplo

Consideremos el conjunto de datos del ejemplo en 2.3.3.3. Las variables de interés son 'Peso' y 'Edad'. Además, una variable respuesta 'Consumo de cigarrillos' y una variable binaria llamada 'tratamiento'. El primer paso es correr el archivo BLOPmatching.m, el cual abrirá la interfaz de usuario. Luego se deben importar los datos hacia el workspace de Matlab como un arreglo de tipo tabla. Luego se debe seleccionar el nombre del arreglo mediante el pop-up 1) mostrado en la sección anterior.

El siguiente paso es seleccionar la variable tratamiento del conjunto de datos, la variable de respuesta y seleccionar alguno o ambos estimadores. En el listbox aparecen

las variables del conjunto de datos. Seleccionamos las variables 'Peso' y 'Edad'. También seleccionaremos las opciones para desplegar un resumen estadístico y un balance postmatching en la ventana de comandos. Mientras que en el pop-up de la variable de respuesta seleccionaremos 'Consumo de cigarrros' y finalmente en el pop-up de la variable de tratamiento seleccionaremos la variable 'tratamiento'. La siguiente figura muestra la configuración deseada.

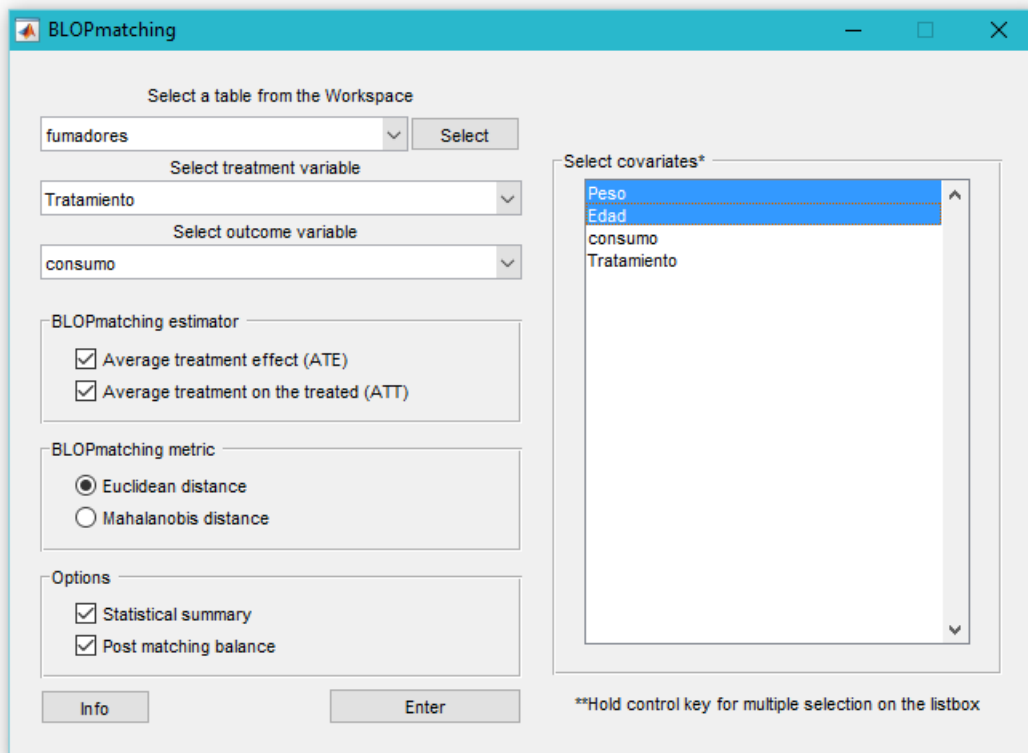


Figura 4.4: Interfaz de usuario BLOPmatching; Selección de variables y opciones

Seleccionado todas las variables y opciones se debe de hacer click en el botón 'Enter' para dar inicio al programa. Si la estructura de los datos es la correcta, el programa entregará los resultados en la ventana de comandos de Matlab. De producirse algún error, el programa lo detectará y pedirá mediante un cuadro de error corregir el problema según lo recomendado. Mientras el programa realiza los diferentes cálculos, una barra de espera nos indica el tiempo de espera de cada actividad, las cuales son calcular los

pesos del grupo tratamiento y control, y calcular los pesos intragrupos para estimar la varianza de cada estimador.

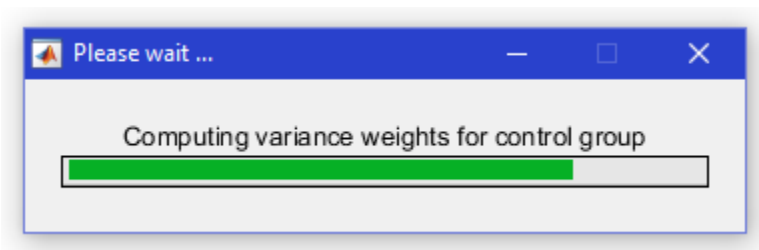


Figura 4.5: Interfaz de usuario BLOPmatching; Barra de espera

Una vez el programa haya hecho los cálculos correspondientes se desplegará una ventana indicando que todos los procesos han finalizado con éxito.

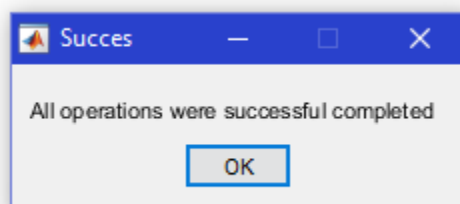


Figura 4.6: Interfaz de usuario BLOPmatching; Ventana de éxito

Por último, se despliegan los resultados del proceso de matching en la pantalla de comandos de Matlab, en el orden de las opciones escogidas por el usuario. En nuestro caso, el resumen estadístico con la media y la desviación estandar divididas por grupo control y grupo de tratamiento, el resultado del matching, indicando los coeficientes de los estimadores, error estandar, estadístico z y su valor p y finalmente el balance post matching.¹

¹La hipótesis bajo la aplicación del estadístico z es determinar si el coeficiente de cada estimador es distinto o no de cero. Las hipótesis son $H_0 : ATE = 0$ versus $H_1 : ATE \neq 0$. Si no puede rechazarse la hipótesis nula, entonces el programa no genera impacto o beneficio sobre la población. Mientras que el balance post matching es una comparativa entre la media observale del grupo tratamiento y la estimación de $Y_i(1)$ del grupo control. El objetivo es determinar que tan buenas son las estimaciones para cada variable.

```

Command Window

Number of observations: Total = 15
                      Control group = 10
                      Treatment group = 5

Statistical Summary

      Ctrl_Mean   Treat_Mean   Ctrl_Std   Treat_Std
      -----   -
Peso      85.9      84          12.05     15.508
Edad      42        43          16.131    18.947

Matching Summary

                        Coef   Std_Error   z   p_Value
                        -----
Average treatment effect   -7.263   1.7535   -4.142   1.7211e-05
Average treatment on the treated -6.2229  1.1586   -5.3711  3.9134e-08

Optimal number of matches: 2.0667

Post Matching Balance

      Treat_Mean   Ctrl_Mean   t_stat   p_value
      -----   -
Peso      84        87.89     -0.53898  0.32045
Edad      43        43.656    -0.068698 0.38168

fx >>

```

Figura 4.7: Interfaz de usuario BLOPmatching; Resultados del matching en ventana de comandos de Matlab

4.3.3. Cuadros de error

Los siguientes cuadros de mensaje pueden aparecer si no se cumplen requisitos para el uso del programa o si bien las opciones seleccionadas no son las correctas. Se debe tener principal cuidado en la estructura de los datos a importar al programa, respetando que no hayan espacios vacíos en las celdas del archivo excel y que los nombres de las variables estén en la primera fila.

Error 1: El primer cuadro de error se da al no escoger ningún estimador:

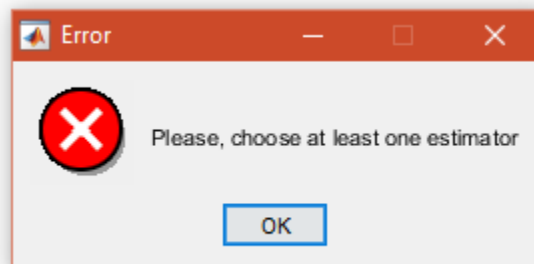


Figura 4.8: Interfaz de usuario BLOPmatching; Error por no escoger estimador

Error 2: El segundo posible cuadro de error se da cuando las variables 'respuesta' y de 'tratamiento' no son distintas:

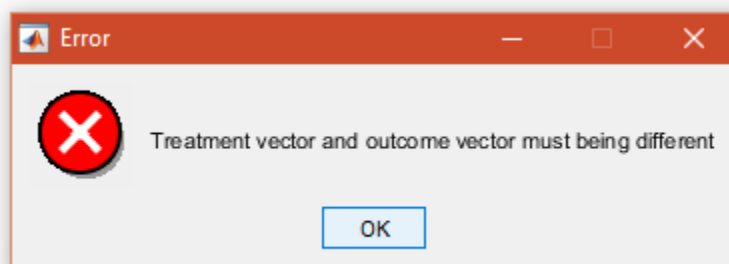


Figura 4.9: Interfaz de usuario BLOPmatching; Error, variables de respuesta y tratamiento deben ser distintas.

Error 3: La variable de tratamiento debe ser binaria, de no cumplirse el programa arrojará el siguiente mensaje de error:

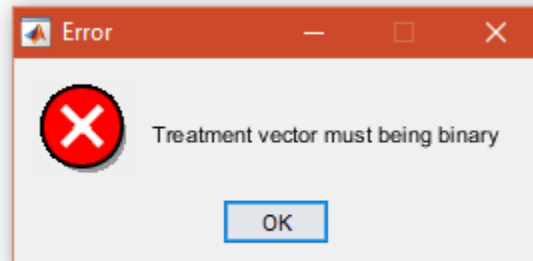


Figura 4.10: Interfaz de usuario BLOPmatching; Error, variable de tratamiento debe ser binaria

Error 4: La variable de tratamiento (o la de respuesta) no puede ser usada en el grupo de covariables como variable independiente:

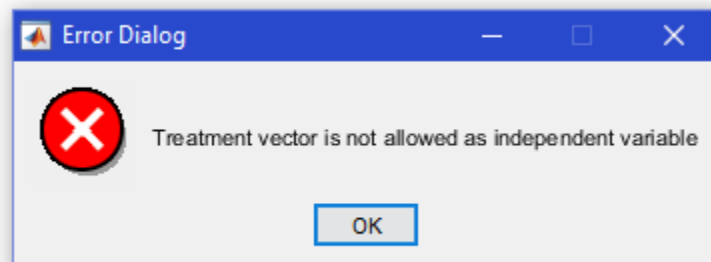


Figura 4.11: Interfaz de usuario BLOPmatching; Error, la variable de tratamiento no está permitida como covariable

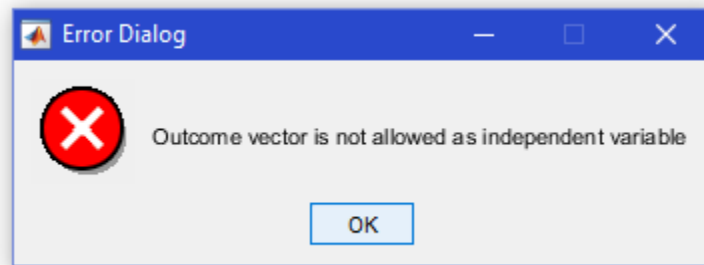


Figura 4.12: Interfaz de usuario BLOPmatching; Error, la variable de respuesta no está permitida como covariable

Capítulo 5

Resultados y Discusión

En este capítulo se toma en discusión el tiempo CPU que toma al estimador BLOP-matching en encontrar el esquema de ponderación óptimo de cada individuo para realizar la estimación del impacto generado por un estudio de inferencia causal o un programa de desarrollo. También se hará una comparación entre BLOPmatching y los estimadores de vecino más cercano basados en Propensity Score y en las covariables, en términos de la estimación de ATE y ATT.

5.1. Análisis de tiempo de cálculo

El estimador fué programado en una computadora personal ASUS X555L, con procesador Intel i7-4510U, velocidad de reloj de 3,1Ghz, 8gb en memoria RAM, bajo sistema operativo Windows 10. El software es compilado en Matlab en su versión 2015b. El estimador BLOPmatching resuelve N problemas de optimización, donde N es la cantidad de individuos en estudio (grupo control mas grupo tratado). El solver a utilizar es Linprog, el cual viene por defecto en el Optimization Toolbox de Matlab. El algoritmo a usar es el de punto interior.

Consideraremos diferentes grupos de individuos de distintos tamaños muestrales N . Los dividiremos en un 25 % grupo tratados y un 75 % grupo control. Mientras que K será la cantidad de variables de la matriz de covariables.

La siguiente tabla muestra el tiempo CPU en segundos que toma al programa realizar los cálculos para la media del estimador BLOPmatching. Los datos fueron generados de forma aleatoria y distribuidos normalmente en el intervalo $[-1,1]$.

Tabla 5.1: Tiempo de cálculo para estimador BLOPmatching en segundos

Individuos (N)	Número de variables (K)			
	2	4	8	16
100	4,995	5,253	7,131	9,577
200	11,500	11,624	15,875	19,876
400	25,330	28,771	35,507	42,753
800	65,091	64,391	82,750	105,089
1600	153,377	173,343	201,053	288,194
3200	331,056	408,306	445,446	660,592

Como puede apreciarse en la tabla 5.1, la tendencia del programa es a aumentar el tiempo de cálculo a medida que aumenta la cantidad de individuos a una rapidez mucho mayor que al aumentar el número de variables. Lo anterior se explica debido a que los problemas a resolver contienen la misma cantidad de variables que cantidad de individuos. Mientras que la cantidad de restricciones del problema es igual a la cantidad de variables de los datos más uno.

Para la varianza se tienen los siguientes resultados:

Tabla 5.2: Tiempo de cálculo para varianza del estimador BLOPmatching en segundos

Individuos (N)	Número de variables (K)			
	2	4	8	16
100	5,472	5,735	7,999	9,435
200	10,834	13,300	16,514	20,154
400	35,707	34,918	42,873	49,723
800	87,129	89,247	108,541	130,422
1600	217,974	243,006	275,177	354,175
3200	337,658	425,659	532,229	808,592

5.2. Benchmark de BLOPmatching respecto de Nearest Neighbor Matching

Para realizar la comparativa utilizaremos un conjunto de datos usados en Lalonde (1986) para determinar el efecto de un programa de entrenamiento para ayudar a trabajadores desempleados a mejorar sus habilidades básicas y luego introducirlas al mercado laboral, dándoles experiencia en un ambiente protegido. El programa fue desarrollado a mediados de los años 70 y es conocido como The National Supported Work Demonstration (NSW). El impacto del programa es medido en términos de cuál fue el ingreso monetario de los individuos en el año 1978, tanto como para aquellos que fueron beneficiados como para los que no. Los datos pueden ser fácilmente obtenidos, ya que se encuentran dentro del paquete 'Matching' de R.

Realizaremos un primer análisis de este conjunto de datos. Se muestran las variables en estudio y una breve descripción de ellas:

- **Age:** Edad en años de cada individuo.
- **Education:** Años de educación.
- **Black:** Variable dummy que toma el valor de 1 si el individuo es de raza negra y 0 en otro caso.
- **Hispanic:** Variable dummy que toma el valor de 1 si el individuo es hispano y 0 en otro caso.
- **Married:** Variable dummy que toma el valor de 1 si el individuo es casado y 0 en otro caso.
- **Earnings 1974:** Ingresos monetarios del año 1974.
- **Earnings 1975:** Ingresos monetarios del año 1975.
- **Earnings 1978:** Ingresos monetarios del año 1978. La variable será usada como respuesta.

- **Unemployed 1974:** Variable dummy que toma el valor de 1 si el individuo estuvo desempleado el año 1974 y 0 en otro caso.
- **Unemployed 1975:** Variable dummy que toma el valor de 1 si el individuo estuvo desempleado el año 1975 y 0 en otro caso.
- **Treatment:** Variable dummy que toma el valor de 1 si el individuo se somete al programa y 0 si no se somete. Será usada como variable de tratamiento.

Los datos cuentan con 185 individuos en el grupo de tratamiento y 260 en el grupo control. La siguiente tabla muestra estadísticos de tendencia central (media, varianza y prueba de diferencia de medias) para los grupos de tratamiento y control.

Tabla 5.3: Resumen estadístico Lalonde (1986)

Variable	Grupo tratados		Grupo control		p-value
	Media	DE	Media	DE	
Age	25,816	7,155	25,056	7,056	0,265
Education	10,346	2,010	10,089	1,614	0,149
Black	0,843	0,364	0,827	0,379	0,647
Hispanic	0,059	0,237	0,108	0,311	0,063
Married	0,189	0,392	0,154	0,362	0,334
Earnings 1974	2095,6	4886,6	2107,0	5687,9	0,982
Earnings 1975	1532,1	3219,3	1266,9	3103	0,385
Unemployed 1974	0,708	0,455	0,750	0,434	0,330
Unemployed 1975	0,600	0,491	0,685	0,466	0,067

Las medias de los grupos de tratamiento y control son similares en términos de sus coeficientes. Mayor evidencia de esto nos lo dan sus p-value, que nos indican que no se puede rechazar la hipótesis de que las medias del grupo tratado sea igual que la del grupo control.¹ Por lo tanto, las observaciones están bien balanceadas en ambos grupos. En términos de la estimación de ATE y ATT, los resultados para el conjunto de datos son los siguientes:

¹Considerese un nivel de significancia del %5 de aquí en adelante.

Tabla 5.4: Resultados de BLOPmatching de Lalonde (1986)

	Coefficiente	E. Estandar	valor-p
E. en el tratamiento promedio (ATE)	1707,8	607,34	0,0049
E. en el tratamiento sobre los tratados (ATT)	1791,8	700,06	0,010

Esto nos dice que aquellas personas que fueron beneficiadas con el programa percibieron 1707,8 USD más que aquellos que no fueron beneficiados. ATT nos dice que los individuos que se enrolaron en el programa percibieron 1791,8 USD más que si ellos no se hubiesen beneficiado del programa. Los valores p de cada estimador indican que se rechaza la hipótesis de que el efecto promedio del programa es 0. Por lo tanto estos estimadores son significativos a un valor de $\alpha = 0,05$. En resumen, el programa genera un impacto positivo sobre los participantes.

Como BLOPmatching busca la cantidad de vecinos y ponderaciones óptimas, se tiene que para nuestro caso el promedio de vecinos más cercanos es de 3,03 emparejamientos, con un emparejamiento mínimo de un vecino y como máximo 8 vecinos. El siguiente gráfico muestra la distribución de la cantidad de vecinos:

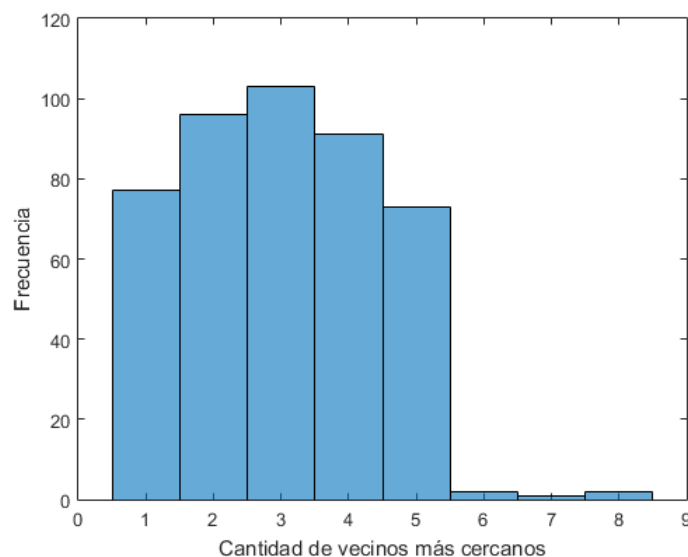


Figura 5.1: Histograma de la cantidad de vecinos en cada match

En orden de demostrar el rendimiento del estimador BLOPmatching es que comparamos este con los estimadores de vecino más cercano en Propensity Score y en covariables. Las siguientes tablas muestran los resultados para ATE y ATT:²

Tabla 5.5: Comparativa entre BLOPmatching y Vecino Más Cercano para ATE

Estimador	Cantidad de vecinos más cercanos (m)						
	Óptimo	1	2	4	16	64	Todos
BLOPmatching	1707,8 (607,34)	-	-	-	-	-	-
NN Propensity Score	-	1427 (789,72)	1564,9 (728,6)	1508,8 (676,13)	1648,3 (654,99)	1539,3 (648,78)	1896,5 (651,78)
NN Covariables	-	2057,2 (809,30)	1627,9 (748,87)	1903,7 (747,78)	1413,9 (666,23)	1346,3 (659,4)	1539,5 (654,08)

Tabla 5.6: Comparativa entre BLOPmatching y Vecino Más Cercano para ATT

Estimador	Cantidad de vecinos más cercanos (m)						
	Óptimo	1	2	4	16	64	Todos
BLOPmatching	1791,8 (700,06)	-	-	-	-	-	-
NN Propensity Score	-	1608,2 (824,14)	1898,2 (775,56)	1970,9 (751,93)	1863,8 (731,81)	1847,1 (729,91)	1794,3 (712,61)
NN Covariables	-	1554,7 (780,39)	1704 (733,62)	1911,3 (713,71)	1793,9 (725,30)	2348,1 (721,46)	2441,65 (706,18)

El estimador BLOPmatching es menor en términos de error estandar del tratamiento medio, tanto para ATE como para ATT. Esto se da debido a que BLOPmatching busca los ponderadores óptimos para cada individuo, en cambio los estimadores de vecino más cercano ponderan de la misma forma a cada observación, lo cual no asegura un balance entre los valores de las covariables de los individuos para formar cada match. Por lo tanto se concluye que el performance de BLOPmatching es bueno en contraste

²La estimación del Propensity Score para NN se calculó mediante un modelo de regresión binomial con factores lineales. La métrica para NN en covariables es Mahalanobis.

con los otros estimadores. A continuación presentamos los balances post matching de cada estimador, que vienen a asegurar nuestra aseveración:

Tabla 5.7: Balance post matching para el estimador ATE

Variable	Media grupo		valor-p
	tratamiento	control	
Age	25,781	24,987	0,391
Education	10,259	10,188	0,727
Black	0,870	0,883	0,719
Hispanic	0,061	0,054	0,742
Married	0,142	0,164	0,622
Earnings 1974	2095,6	2092,1	0,797
Earnings 1975	1495,8	1320,6	0,665
Unemployed 1974	0,742	0,785	0,433
Unemployed 1975	0,632	0,715	0,137

Tabla 5.8: Balance post matching para el estimador ATT

Variable	Media grupo		valor-p
	tratamiento	control	
Age	25,816	24,987	0,370
Education	10,346	10,188	0,537
Black	0,843	0,883	0,371
Hispanic	0,059	0,0535	0,765
Married	0,189	0,164	0,611
Earnings 1974	2095,6	2092,1	0,797
Earnings 1975	1532,1	1320,6	0,626
Unemployed 1974	0,708	0,785	0,139
Unemployed 1975	0,600	0,715	0,030

El balance post matching del estimador ATE es la comparativa entre la media del grupo de tratamiento estimado mediante las ponderaciones del grupo control y la media

del grupo control estimado con los ponderadores del grupo tratamiento. Mientras que para ATT es una comparativa entre la media observable del grupo tratamiento y la media de las variables del grupo control estimadas con la ponderación obtenida para los individuos del grupo en tratamiento. Nos indica que tan bueno es el esquema de ponderación por cada variable. En nuestro caso se puede ver que las medias observables del grupo tratamiento y las medias estimadas del grupo control son similares. Mayor evidencia de esto nos lo da el valor-p, las cuales al ser mayor que $\alpha = 0,05$ nos indica que no podemos rechazar la hipótesis de que las medias sean iguales. Esto confirma que nuestro esquema de ponderación generó un buen resultado en las estimaciones. Por otro lado, se tiene que para la variable de desempleo en el año 1975 (Unemployed 1975) se debe rechazar la hipótesis mencionada. Por lo tanto, se recomienda probar un modelo de inferencia causal sin esa variable y comparar los resultados.

5.3. Cambios en la métrica de BLOPmatching

En los métodos de matching basados en covariables, las métricas más usadas son las de distancia euclidiana y distancia de Mahalanobis. Estudiaremos en cómo cambia la solución al usar ambas métricas en el estimador. Consideraremos datos reales usados en Almond *et al* (2005). Los datos pueden ser fácilmente obtenidos en STATA. El estudio en análisis es un problema de inferencia causal, que busca determinar el impacto que genera el consumo de tabaco de madres en el período de embarazo sobre el peso de sus hijos recién nacidos. Esto debido a que aquellos infantes que poseen un peso menor que el resto pueden padecer problemas de salud a futuro. Las variables consideradas en el modelo son:

- Mother smokes: Variable dummy que toma el valor de 1 si la madre fuma durante el periodo de embarazo y 0 en otro caso. Será usado como variable de tratamiento.
- Baby weight: Peso del bebé nacido. Será usado como variable de respuesta.
- First baby: Variable dummy que toma el valor de 1 si el niño es el primero hijo, 0 en otro caso.

- Married: Variable dummy que toma el valor de 1 si el individuo está casado y 0 en otro caso.
- Mother's age: Edad de la madre.
- Mother's educational level: Nivel educacional de la madre.

Tabla 5.9: Resultados y comparativa de métricas para BLOPmatching

Estimador	D. Euclidiana			Mahalanobis		
	Coef.	E. Estandar	valor-p	Coef.	E. Estandar	valor-p
ATE	-206	16,669	0,000	-205,62	16,697	0,000
ATT	-244,94	25,578	0,000	-239,56	25,618	0,000

Como puede apreciarse en las estimaciones de la tabla anterior, ambos estimadores ATE y ATT poseen valores cercanos al ser estimados usando como métrica la distancia euclidiana y la distancia de Mahalanobis. Entonces la pregunta es cuándo conviene usar una métrica como la euclidiana y cuándo una como la de Mahalanobis. La respuesta la define el resultado del matching y del balance post matching. Esto quiere decir, que si para una métrica los estimadores ATE y ATT no son significativos, entonces se recomienda usar otra distancia. También es favorable cambiar de métrica cuando para alguna variable o más de una no hay un buen balance en términos de las respuestas no observables estimadas, en otras palabras, si no puede rechazarse la hipótesis de que las estimaciones de la media de cada variable para el grupo control y tratamiento en términos de nuestros ponderadores sean distintas de cero. Finalmente, si nos quedáramos con el modelo con métrica de Mahalanobis, la interpretación quedaría para ATE, que las mujeres que fuman en periodo de embarazo dan a luz a niños con un peso menor en 205,62 gramos que aquellas mujeres que no fumaron durante ese periodo. Mientras que para ATT, las mujeres fumadoras dieron a luz a niños 239,56 gramos menos que en el caso de que ellas no hubiese fumado. Por lo tanto, basado en nuestros estimadores, podemos afirmar que las mujeres que fuman durante el periodo de embarazo darán a luz a niños con menor peso, lo cual podría generar a futuro problemas de salud en los menores.

Capítulo 6

Conclusiones

La evaluación de impacto es un campo que aún requiere de bastante desarrollo. Nuevas metodologías se han implementado a través del tiempo para formar las bases para que esta ciencia siga creciendo. La magnitud de aplicaciones que posee es bastante amplia y existen diversos métodos para estimar el impacto de un programa de desarrollo o política pública, así también como estimar el efecto de un estudio de inferencia causal. Entre estos métodos se encuentra Regresión Discontinua, el cual estima el impacto de un programa midiendo el antes y el después de una política y desarrollando modelos de regresión lineal sobre los resultados observados. Luego el impacto se genera mediante la evaluación del parámetro con el cual se enrolan a los individuos en cada programa. El segundo método discutido es el de diferencias en diferencias, el cual es una herramienta bastante sencilla de utilizar, pero que requiere de la aceptación de supuestos bastante fuertes acerca de la tendencia de las respuestas, lo cual implica que el método no sea robusto. Finalmente, están los métodos de matching, los cuales utilizan técnicas de estadística avanzada y grandes bases de datos para estimar las respuestas no observables sobre un grupo de individuos en estudios.

Actualmente, los métodos de matching son la forma de estimar el impacto de un programa más populares. Pueden ser aplicados a diferentes modelos y situaciones distintas, aunque requieren de grandes bases de datos. El estimador introducido por Díaz, Rau y Rivera (2015) se aleja de los principios básicos de los estimadores de vecino más cercano, los cuales son comparar aquellos vecinos con menor diferencia en términos de

alguna distancia métrica. En cambio, BLOPmatching busca la ponderación óptima bajo un criterio de balance de variables y covariables, minimizando la distancia total del individuo que debe ser emparejado respecto del resto. Esto se logra resolviendo múltiples problemas de optimización que cumplen con los requisitos expuestos.

Uno de los objetivos es generar herramientas informáticas para la implementación del estimador BLOPmatching. Se cuenta con una aplicación desarrollada para Matlab, con una interfaz gráfica sólida que permite realizar estimaciones de la forma más intuitiva y fácil posible, permitiendo al usuario seleccionar las variables de tratamiento, respuesta y el conjunto de variables independientes, como así también los estimadores a calcular y las diferentes opciones disponibles asociadas a resúmenes estadísticos divididos para cada grupo y el balance post matching, el cual es una herramienta para validar nuestro modelo y la calidad de los ponderadores calculados. La interfaz gráfica cubre todos los posibles errores que pueden darse por problemas en la estructura de los datos o entradas erróneas en el modelo. Además, se presenta una función llamada `blopmatch.m` que permite la utilización de BLOPmatching mediante la aplicación directa de comandos propios de Matlab, así el usuario puede usar la función para los propósitos que estime conveniente en combinación con otras aplicaciones, con la principal ventaja de que el usuario no debe preocuparse del formato de entrada de los datos (`blopmatch.m` admite arreglos matriciales, tablas y arreglos de celdas).

El presente proyecto de título ha demostrado que BLOPmatching se desenvuelve bastante bien en términos de error estándar y efecto promedio sobre la estimación de inferencia causal en un grupo de individuos. Mediante la aplicación de la metodología de BLOPmatching se mostró que el estimador es una herramienta competente, generando estimaciones dentro de los parámetros esperados minimizando así el error estándar. Además, el balance post matching da mayor evidencia del rendimiento del estimador al asegurar que el esquema de ponderación es significativo para el modelo planteado. Una gran ventaja que presenta BLOPmatching es que el usuario no debe preocuparse de la cantidad de vecinos a usar en el proceso (acción dejada al arbitrio del investigador) ni tampoco de la estimación de un modelo de Propensity Score, el cual requiere de la implementación de modelos de regresión logístico y con ello cumplir los supuestos propios de este tipo de curvas. La desventaja de contar con una solución óptima al problema

de matching es que se requiere el hardware apropiado para que las estimaciones sean realizadas en el menor tiempo posible, ya que el programa requiere de resolver modelos de optimización por cada individuo para estimar así la respuesta no observable y la varianza condicional. Se puede concluir que los tiempos de cálculos son bastante sensibles a cambios en el tamaño muestral, no siendo tan sensible a cambios en la cantidad de variables. Se puede apreciar que cambios en la métrica de los estimadores no generan diferencias tan grandes en término de los coeficientes calculados para ATE y ATT, sin embargo es recomendable cambiar el tipo de distancia cuando haya problemas para determinar si los estimadores son significativos o no.

Basados en el trabajo de Díaz, Rau y Rivera (2015), este proyecto de título aporta con herramientas informáticas diseñadas y programadas para el uso abierto de la comunidad científica, priorizando que sean herramientas intuitivas y sencillas de utilizar. Además, contribuye a la literatura de Evaluación de Impacto como introducción a ésta disciplina. considerando tópicos desde cómo realizar un experimento de inferencia causal y métodos disponibles para estimar el impacto.

Se espera como desafío para el futuro que los métodos que se han desarrollado puedan permitir la evaluación del impacto de programas de desarrollo en términos multivariantes. Por ejemplo, si se desea determinar si un programa para controlar la obesidad genera disminuye el peso de los participantes; no solo medir ese factor, si no que además determinar si se genera un impacto para combatir la diabetes u otras enfermedades relacionadas con la obesidad. También se busca que estos métodos puedan ser la base para la replicación de los programas de desarrollo sobre otras poblaciones basados en los resultados de estudios anteriores y a partir de la evidencia disponible poder ampliar los programas, quitar los que tengan una relación beneficio costo bajo y mejorar las políticas implementadas.

Referencias Bibliográficas

1. Abadie, A. (2005). Semiparametric Difference-in-Differences Estimator, *Review of Economics Studies* 72: 1-19.
2. Abadie, A. Drukker, D. Leber, J. e Imbens, G. (2004). Implementing matching estimators for average treatment effects in Stata, *The Stata Journal* 4(3): 290-311.
3. Abadie, A. e Imbens, G. W. (2006), Large sample properties of matching estimator for average treatment effect, *Econometrica* 74: 235–267.
4. Almond, D. Chay, K. y Lee, D. (2005). The costs of low birth weight. *The Quarterly Journal of Economics* 120(3): 1031–1083.
5. Berg, M. Cheong, O. Kreveld, M. y Overmars, M. (2008). *Computational geometry: Algorithms and applications*. Berlín: Springer.
6. Becker, S, e Ichino, A. (2002). Estimation of average treatment effects based on propensity scores, *The Stata Journal*, 2(4): 290-311.
7. Botkin, N. y Stoer, J. (2005). Minimization of convex functions on the convex hull of a point set. *Mathematical Methods of Operations Research*, 62: 167-185.

8. Card, D, y Krueger, A. (1994). Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania. *American Economic Review* 84(4): 772–793.
9. Chapin, F. (1947). *Experimental designs in sociological research*. New York: Harper.
10. Colson, B., Marcotte, P., y Savard, G. (2007). An overview of bilevel optimization. *Annals of Operations Research*, 153: 235-256.
11. Dehejia, R. y Wahba, S. (1999). Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs. *Journal of the American Statistical Association* 94(448): 1053-1062.
12. Díaz, J. Rau, T. y Rivera, J. (2015), A matching estimator based on a bi-level optimization problem, *The Review of Economics and Statistics*.
13. Gertler, P. Martinez, S. Premand, P. Rawlings, L. y Vermeersch, C. (2011). *Impact evaluation in practice*. Washington D.C: The World Bank.
14. Hahn, J. (1998). On the role of the propensity score in efficient semiparametric estimation of average treatment effects. *Econometrica*: 66(2): 315-331.
15. Heckman, J. (1991). Randomization and social policy evaluation. Cambridge: National Bureau of Economic Research.
16. Imas, L. y Rist, R. (2009). *Road to Results: Designing and Conducting Effective Development Evaluations*. Washington D.C: The World Bank.
17. Imbens, G. (2000). The role of propensity score in estimating dose—response functions. *Biometrika* 87(3): 706-710.
18. Khandker, S. Koolwal, G. y Samad, H. (2010). *Handbook of impact evaluation quantitative methods and practices*. Washington D.C: The World Bank.

19. Kimmel, P. Weygandt, J. y Kieso, D. (2010). *Accounting: Tools for Business Decision Makers*. New York: Wiley.
20. Lalonde, R. (1986). Evaluating the econometric evaluations of training programs with experimental data. *The American Economic Review* 76(4): 604-620.
21. Lee, D. y Lemieux, T. (2010). Regression Discontinuity Designs in Economics. *Journal of Economic Literature*: 281-355.
22. Liu, X. Yan, H. y Wang, D. (2010). The evaluation of “Safe Motherhood” program on maternal care utilization in rural western China: a difference in difference approach. *BMC Public Health* 10:566.
23. Rosenbaum, P. (1995). *Observational studies*. New York: Springer Verlag.
24. Rosenbaum, P. y Rubin, D. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika* 70(1): 41-55.
25. Rosenbaum, P. y Rubin, D. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity. *American Statistician* 39(1): 33-38.
26. Rubin, D. (1973). Matching to remove bias in observational studies. *Biometrics* 29(1): 159-183.
27. Sekhon, J. (2008). The Neyman-Rubin model of causal inference and estimation via matching methods. *Oxford Handbooks Online*.
28. Todd, P. (1995). Matching and local linear regression approaches to solving the evaluation problem with a semi-parametric propensity score. University of Chicago.