

EYP3007 Consultoría Estadística

Caso 2

Diego Aravena

Alonso Campos

Axel González

April 27, 2023

Abstract

En este trabajo se estudió la efectividad de un programa de tutorías implementado por una universidad pública chilena. Se busca determinar un número óptimo de tutorías en el cual pueda verse reflejado si el estudiante logró ser intervenido, entendiéndose como intervenido una vez que existe un impacto en el rendimiento del estudiante, y a su vez al final del primer semestre se busca determinar la mejor decisión que puede tomar un estudiante acerca de seguir o no en el programa el segundo semestre. Se podrá ver que la cantidad óptima de tutorías dependerá de cómo se defina tener un buen rendimiento, aunque en este caso dado la información disponible sólo se puede hablar de notas o más bien números que pueden llegar a distinguir un buen rendimiento. Por otro lado, para decidir si seguir o no en el programa nos veremos enfrentados al gran problema de la inferencia causal y, por ende será necesario añadir supuestos necesarios que podrían cambiar las conclusiones derivadas de cada una.

1 Introducción

Hace ya varios años que las distintas universidades chilenas han tenido que sobrellevar los déficits sociales, culturales, psicológicos, etc y sobre todo de conocimiento, que traen consigo los estudiantes debido al débil programa educacional escolar que llevan a cabo la gran mayoría de colegios en Chile, trayendo consigo fuertes desigualdades al momento de enfrentar la vida universitaria [1]. En resumen una de las soluciones que se ha vuelto famosa este último tiempo y que varias universidades ya la han implementado, trata sobre programas de acompañamiento a los estudiantes de primer año de universidad que carecen de los contenidos mínimos requeridos por cada institución y que no logran ser evaluados en la prueba de selección universitaria.

Es bajo esta premisa que toca verificar la eficacia de un programa de tutorías que se imparte por una universidad chilena, en particular se trata de la Universidad de Santiago de Chile. Este programa de tutorías es ofrecido a estudiantes que recién ingresan a la universidad y se centra principalmente, pero no únicamente, en estudiantes que hayan ingresado a través del acceso inclusivo (VAI). El programa está organizado por semestres y al final del primer semestre cada estudiante puede decidir si continúa con el programa o no; si se liberan los cupos otros estudiantes pueden participar del programa de tutorías.

A continuación se entrega mayor contexto sobre este programa de tutorías y en qué consiste cada una.

1.1 ¿Qué es PAIEP?

PAIEP es el **Departamento de Acceso Inclusivo, Equidad y Permanencia** de la Universidad de Santiago de Chile, cuyo propósito es ofrecer acompañamiento académico a quienes ingresan a primer año en la universidad, tanto por Vías de Acceso Inclusivo en Equidad (VAI), como por ingreso regular (PSU). El acompañamiento aborda un área académica compuesta por: Lectura y Escritura Académica, Matemática, Ciencias y una socio educativa compuesta por el Área Socio educativa.[3]

1.2 ¿Qué es una tutoría?

Las tutorías son instancias de interacción **académica y acompañamiento** entre un tutor par o senior y uno o mas tutorados. En el caso de PAIEP las tutorías tiene como objetivo contribuir y apoyar los aprendizajes y el desarrollo de habilidades de los beneficiarios de la **Beca de Acompañamiento PAIEP** (BAP).[4]

2 Descripción de los datos/variables

Para esta realización se utilizó la base de datos *Exam2023.xls*, el cual contiene información sobre más de 3000 estudiantes que ingresan a primer año en alguna de las carreras que posee la universidad, lo que a su vez no tiene ninguna implicancia en el desarrollo de este trabajo. La tabla 1 nombra y describe cada una de las variables que se le midió a cada estudiante.

Nombre	Descripción
Ptje Ranking	Puntaje correspondiente al percentil obtenido en la enseñanza media
Ptje Psu	Puntaje PSU Lenguaje y Matemática
VAI	Si VAI=1 Estudiante fue admitido a la universidad por medio de programa de acceso inclusivo, VAI=0 Si no.
BAP	Si BAP=1 Estudiante participa de programa de tutorías. BAP=0 Si no.
Tut_20161	Numero de tutorías a la cual asistió el alumno en el primer semestre.
Tut_20162	Numero de tutorías a la cual asistió el alumno en el segundo semestre.
Prom_20161	Promedio de notas del primer semestre.
Prom_20162	Promedio de notas del segundo semestre.

Tabla 1: Descripción de variables de la base de datos

El propósito de este trabajo yace en estudiar la posible existencia de una asociación causal entre la intensidad de un tratamiento y el promedio de notas de cada estudiante (variable respuesta), entendiéndose como tratamiento la cantidad de tutorías a las que estos asisten.

2.1 Análisis exploratorio de datos

2.1.1 Análisis inicial

A continuación se entenderá como estudiante BAP a aquellos cuya variable BAP sea igual a 1. De manera idéntica a la anterior para los estudiantes cuya variable VAI es 1.

La base de datos contiene **3866** registros, es decir, información de 3866 estudiantes de los cuales solo 411 son BAP el primer semestre y 375 son VAI, sin embargo, 353 son estudiantes VAI y BAP. Esto implica que existen 22 estudiantes VAI que no decidieron entrar al programa de tutorías. Toda esta información está recopilada en la tabla 2. Luego, es desde este punto que comienzan a surgir las primeras inquietudes sobre cómo le habría ido en el programa a aquellos estudiantes que no pertenecen a el.

		VAI		
		No	Sí	Total
BAP	No	3433	22	3455
	Sí	58	353	411
	Total	3491	375	3866

Tabla 2: Tabla de contingencia estudiantes BAP y VAI

2.1.2 Análisis de datos faltantes

Se observó la presencia de datos faltantes en la variable respuesta, lo cual vendría siendo un problema dado que es a partir de esta que se desea medir el impacto del programa.

En los promedios del primer semestre existen 161 datos faltantes, de los cuales 17 son de estudiantes BAP. Por otra parte faltan 495 datos en los promedios del segundo semestre, de los cuales a su vez 80 tampoco presentan información en el promedio del primer semestre.

Más adelante se podrá ver que el impacto del programa quedará sucinto al comportamiento entre los promedios del primer y segundo semestre, por lo tanto, al no poseer información del primer semestre se habría perdido el progreso que hubiese tenido el estudiante, dado que se entenderá como buen rendimiento al suceder una de las siguientes situaciones: el promedio del segundo semestre es superior al del primero ó el promedio del primer semestre es superior o igual a una nota X sujeta a algún criterio. Por este motivo es que se ignorarán a los estudiantes que no posean información de su promedio en el primer semestre, eliminando a su vez 80 estudiantes del segundo semestre que tampoco poseen información, quedando así 415 estudiantes con datos faltantes el segundo semestre.

Finalmente se reemplazó los datos faltantes por ceros en las variables asociadas al promedio de notas del segundo semestre y cantidad de tutorías asistidas en el primer y segundo semestre.

2.1.3 Selección de variables

Dado el propósito del estudio, las variables a considerar en principio son:

Prom_20161	Tut_20161	BAP
Prom_20162	Tut_20162	VAI

Para el caso de las variables asociadas a los puntajes PSU de matemáticas, lenguaje y ranking, en la tabla 3 se puede observar que la correlación entre el promedio del primer semestre y el puntaje ranking es menos de un 1%, por otro lado, tanto el puntaje PSU de matemáticas como de lenguaje están bordeando el 25% de correlación con el promedio del primer semestre. Todo lo anterior no logra apoyar la idea de incluir estas variables al análisis dada la poca evidencia presente

Correlación	Puntaje PSU Lenguaje	Puntaje PSU Matemática	Puntaje Ranking	Promedio primer semestre
Puntaje PSU Lenguaje	1.000	0.451	-0.039	0.235
Puntaje PSU Matemática	0.451	1.000	-0.115	0.270
Puntaje Ranking	-0.039	-0.115	1.000	0.006
Promedio primer semestre	0.235	0.270	0.006	1.000

Tabla 3: Matriz de correlación

Desde otro punto de vista, en la figura 1 se muestra el gráfico de dispersión entre el promedio PSU de lenguaje y matemáticas diferenciando por los estudiantes que logran igualar o mejorar su promedio. En principio no se logra observar ninguna distinción entre estos estudiantes según sus puntajes PSU. Además observando el histograma de la figura 2 se ve que existen estudiantes con bajos puntajes PSU pero con una mejora en el promedio, mientras que hay estudiantes con puntajes PSU altos pero que no mejoran, es por estos motivos que no logra ser viable votar a favor sobre incluir estas variables en el análisis.

3 Problemáticas

Este trabajo busca realizar una evaluación de impacto del programa de tutorías tal que se pueda tomar una decisión que implique llevarlo a cabo o no. Para ello, el enfoque estará en estudiar los siguientes

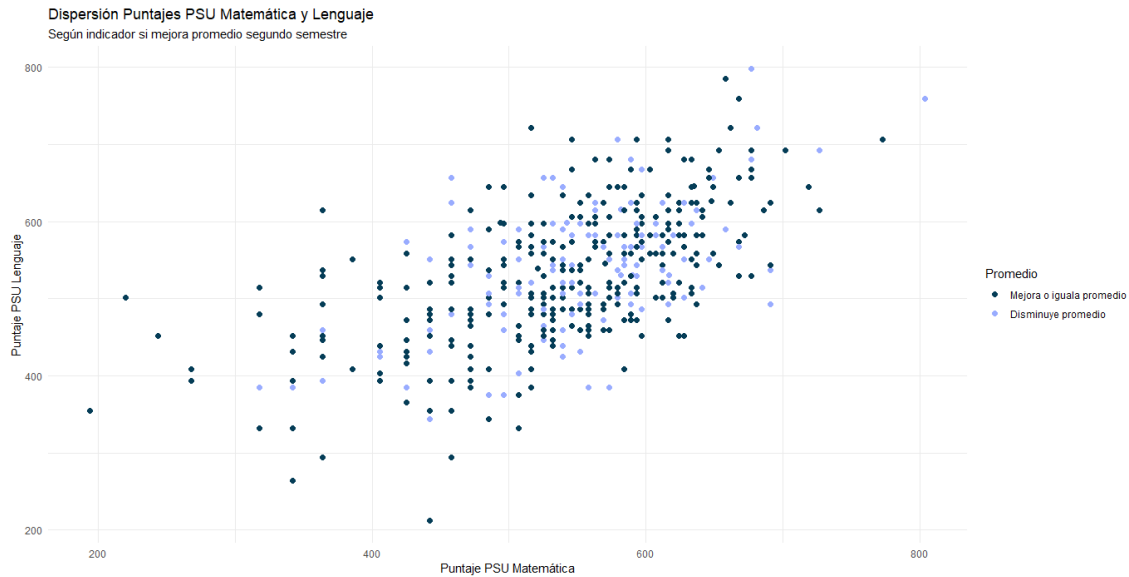


Figura 1: Gráfico de dispersión entre PSU de Matemáticas y Lenguaje distinguiendo a los estudiantes por nota 4

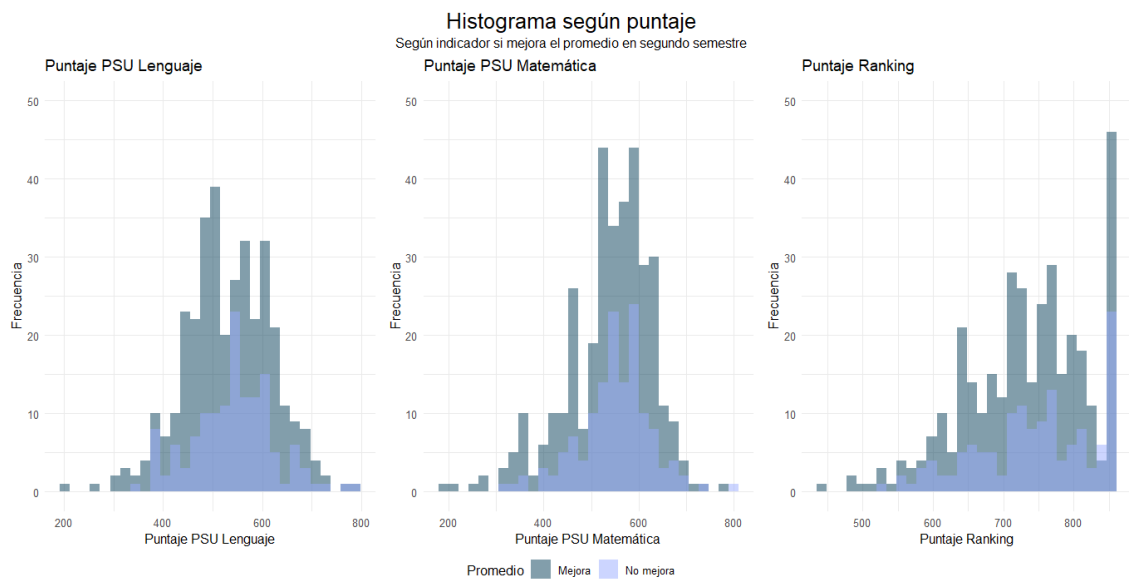


Figura 2: Histogramas PSU Lenguaje, Matemáticas y puntaje Ranking respectivamente, distinguiendo a los estudiantes que mejoran o no su promedio el segundo semestre

aspectos:

- Determinar cuál es la cantidad de tutorías que genera un impacto positivo en el rendimiento académico de los estudiantes, es decir, cual es la asistencia necesaria a clases de tutorías con tal de que los estudiantes logren a final de semestre un promedio igual o superior a cierta nota.
- Al final de cada semestre los estudiantes que hayan sido parte de las tutorías deciden si continuar en el programa o no. Dicho esto se busca determinar cuál es la mejor decisión en pos de obtener un cambio positivo en el rendimiento académico.

3.1 Problemática 1. Determinar el número de tutorías necesarias para generar un impacto en el rendimiento académico

Para medir el impacto de las tutorías se consideran sólo los estudiantes BAP, ya que, sólo ellos tendrán información sobre la cantidad de tutorías asistidas. En base a esta se dividirá a los estudiantes en dos grupos, los que fueron al menos a k tutorías, que en adelante se mencionarán como intervenidos, y lo que asisten a menos de k tutorías. De esta manera se define como

$$Z_k = \begin{cases} 1, & \text{si el estudiante asiste al menos a } k \text{ tutorías} \\ 0, & \text{si no} \end{cases}$$

Por otro lado, se determina si la intervención de las tutorías ha tenido un impacto positivo en el promedio del estudiante. Para ello se tendrá como referencia una nota τ , de modo que se define la siguiente variable

$$Y = \begin{cases} 1, & \text{si el promedio del 1° semestre es mayor a } \tau \\ 0, & \text{si no} \end{cases}$$

Por consiguiente mediante el teorema de probabilidad total es posible determinar la probabilidad de que los estudiantes obtengan un promedio igual o mayor a τ particionando por Z_k , es decir,

$$P(Y = 1) = P(Y = 1|Z_k = 1)P(Z_k = 1) + P(Y = 1|Z_k = 0)P(Z_k = 0) \quad (1)$$

En particular se hará un seguimiento de la probabilidad condicional $P(Y = 1|Z_k = 1)$ de la ecuación (1), que corresponde a la probabilidad de obtener un promedio igual o superior a τ dado que se asiste al menos a k tutorías, puesto que interesa ver si la cantidad de tutorías tiene algún efecto sobre el promedio. Por ende, bajo este contexto se puede entender esta probabilidad condicional como la probabilidad de éxito.

La figura 3 muestra para distintos valores de k cómo fluctúa la probabilidad de éxito, distinguiendo para $\tau = (4.24, 5.04, 5.48)$, notas correspondientes a los cuartiles 1, 2 y 3 de los promedios de estudiantes que no son BAP. Se escogen estos valores meramente para tener algunas referencias entre los estudiantes BAP y no BAP, luego se podrá deducir que con cualquier otra nota se pueden obtener las mismas conclusiones.

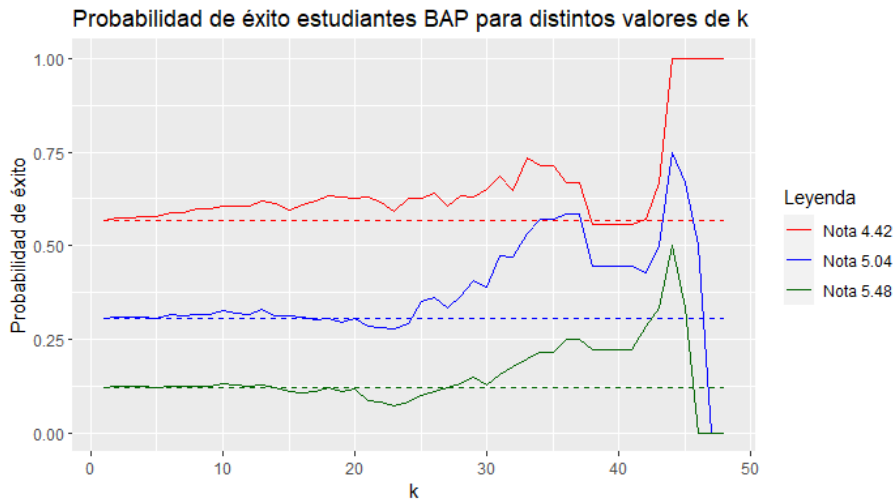


Figura 3: Probabilidad de éxito según la cantidad de tutorías asistidas en el primer semestre tomando como referencia tres notas distintas

La línea roja continua corresponde a la probabilidad de éxito de que los estudiantes tengan una nota igual o superior a 4.42 dado que asisten al menos a k tutorías, mientras que la recta roja interlineada corresponde a la probabilidad de que todos los estudiantes BAP tengan un promedio igual o superior a la misma nota 4.42, es decir, $P(Y = 1)$ que es cercano a 0.6. Se observa que para $0 \leq k \leq 36$ la probabilidad de éxito siempre es mayor o igual que $P(Y = 1)$ y cuya diferencia se incrementa a medida que k aumenta, lo cual nos da a entender que con asistir al menos a 5 tutorías aproximadamente la probabilidad de éxito tiende a aumentar en mayor medida llegando hasta un 15% sobre $P(Y = 1)$, visto de otra manera, al asistir a las tutorías se estaría superando el cuartil 1 de los promedios de estudiantes no BAP. Como observación la cantidad de estudiantes con promedio igual o superior a 4.42 y con una asistencia de al menos 40 tutorías es de tan solo 5 estudiantes por lo que se estaría perdiendo precisión. Lo que vemos hace sentido, ya que, la probabilidad de superar aquella nota es menos costoso que superar una nota 5.04 o 5.48, que son mayores, pues es posible observar que para estas notas se requiere cada vez más de una mayor cantidad de asistencia y además las probabilidades son más bajas respectivamente.

De la recta azul continua se puede ver que la probabilidad de éxito empieza a aumentar a partir de un k igual 25, mientras que la probabilidad de éxito de la recta verde continua aumenta a partir de un k cercano a 30.

Por lo tanto, dependiendo de la nota por la cual se quiera condicionar, el k óptimo será diferente. La siguiente problemática a analizar consiste en determinar la mejor decisión sobre seguir o no en el programa, y esta dependerá del k escogido. Para efectos de este estudio y con el ánimo de continuar con el mismo es que se escogerá de manera arbitraria un k igual a 20. Dicha cifra considera un número importante de individuos que permanecen en las tutorías, pues al aumentar k , disminuye el número de estudiantes que cumplen con la condición. Además de ser cercano al valor donde se aprecian mejoras en la probabilidad de tener un promedio superior a x nota, que corresponde a $k = 25$, sin embargo, esta cifra posee un menor número de estudiantes.

3.2 Problemática 2

Tal como se mencionó anteriormente, ahora es de interés medir el impacto que tienen las tutorías sobre los promedios tanto del primer y segundo semestre. Para ello se crea la variable Y_1 que está asociada a la variable respuesta cuando los estudiantes van al menos a k tutorías y la variable Y_0 asociada a la variable respuesta cuando el estudiante asiste a menos de k tutorías. En este caso la variable respuesta estará asociada al comportamiento entre ambos promedios, determinando como respuesta exitosa cuando el promedio del estudiante en el segundo semestre es mayor o igual que el primero. Es decir,

$$Y = \begin{cases} 1, & \text{si el estudiante mejora su promedio el 2° semestre} \\ 0, & \text{si no} \end{cases}$$

Como se mencionó anteriormente, en este caso el valor de k corresponde a 20, por lo que se considerará la variable

$$Z = \begin{cases} 1, & \text{si el estudiante asiste al menos a 20 tutorías el 2° semestre} \\ 0, & \text{si no} \end{cases}$$

Ahora bien, utilizando el Teorema de probabilidades totales para **determinar** la probabilidad de éxito de las variables Y_1 e Y_0 , se tiene lo siguiente

$$P(Y_1 = 1) = P(Y_1 = 1|Z = 1) \cdot P(Z = 1) + P(Y_1 = 1|Z = 0) \cdot P(Z = 0) \quad (2)$$

$$P(Y_0 = 1) = P(Y_1 = 0|Z = 1) \cdot P(Z = 1) + P(Y_0 = 1|Z = 0) \cdot P(Z = 0) \quad (3)$$

Como se puede apreciar las probabilidades de interés dependen de factores que no se pueden conocer, estos son, $P(Y_1 = 1|Z = 0)$ y $P(Y_1 = 0|Z = 1)$. Estos eventos se conocen como contrafactuales y se pueden interpretar respectivamente como la probabilidad de que un estudiante mejore su promedio cuando asiste a las 20 tutorías dado que este en verdad no asistió, y como la probabilidad de que un estudiante mejore su promedio cuando asiste a menos de 20 tutorías dado que este en verdad si asistió. Estas probabilidades corresponden a situaciones hipotéticas, dado que los estudiantes sólo pueden estar categorizados como intervenidos o no ($Z=1$ o $Z=0$), por ende, se deben realizar supuestos sobre estas y en base a ello obtener los posibles resultados.

En particular, para estudiar el impacto del programa, en las siguientes secciones se estudiarán algunos casos que pueden estar ocurriendo con las probabilidades que desconocemos.

3.2.1 Estudio de casos

Caso 1: Probabilidades Límite

De lo único que se puede estar seguro es que ambas probabilidades $P(Y_1 = 1|Z = 0)$ y $P(Y_0 = 1|Z = 1)$ están acotadas en el intervalo $[0, 1]$. Por este motivo es que se reemplazará cada una de estas por sus cotas, logrando así obtener un rango en donde puede estar 'moviéndose' la probabilidad de interés $P(Y_1 = 1)$ y $P(Y_0 = 1)$. Por tanto, reemplazando estos valores en la ecuación (2) y (3) se obtiene los siguientes intervalos.

$$\begin{aligned} P(Y = 1, Z = 1) &\leq P(Y_1 = 1) \leq P(Y = 1, Z = 1) + P(Z = 0) \\ P(Y = 1, Z = 0) &\leq P(Y_0 = 1) \leq P(Y = 1, Z = 0) + P(Z = 1) \end{aligned}$$

En este caso los intervalos descritos anteriormente sólo dependen de valores conocidos, por lo que sí es posible determinar dichos límites.

Luego, utilizando la información proporcionada, en la Figura 4 se muestran los límites para cada caso.

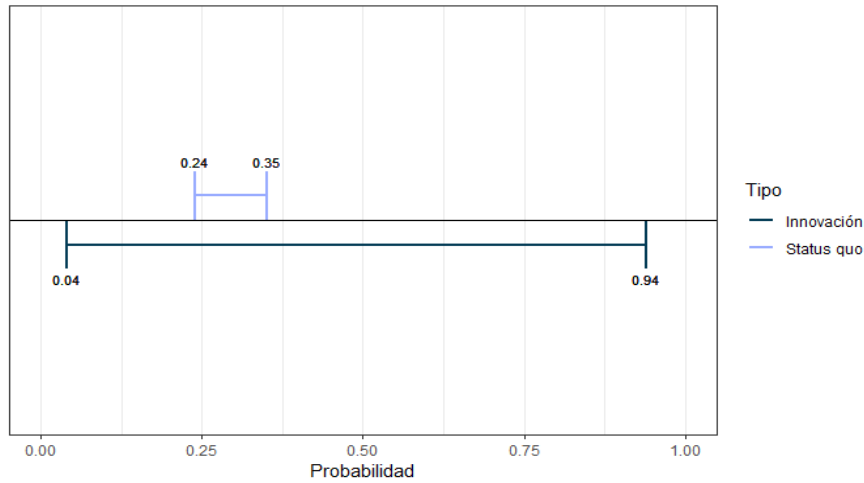


Figura 4: Intervalo de identificación Parcial caso 1, donde innovación corresponde a $P(Y_1 = 1)$ y status quo como $P(Y_0 = 1)$

Como se aprecia en la figura 4 si se reemplaza las probabilidades por los valores extremos, entonces, la probabilidad $P(Y_1 = 1)$ presenta una mayor variabilidad dado el ancho del intervalo, en comparación a $P(Y_0 = 1)$. Luego se puede concluir que la probabilidad de que todos los estudiantes mejoren su promedio cuando asisten al menos a k tutorías, en el mejor de los escenarios el 94% lo lograría. En caso contrario,

en una situación adversa, se espera que tan solo el 4% de ellos mejore su promedio. Por otro lado, la probabilidad de que todos los estudiantes mejoren su promedio cuando asisten a menos de k tutorías, es por lo mucho un 35% y por lo bajo sería de un 24%.

Caso 2: Solución Optimista

Otra supuesto que se utilizara es creer que la innovación es siempre mejor que el status quo. En este caso, significa que asistir a las tutorías aumentará la probabilidad de obtener una mejora en el promedio en el segundo semestre. En otras palabras, suponiendo que si un estudiante que asiste a las tutorías no lo hiciese, disminuiría su probabilidad de éxito; en caso contrario, si un estudiante que no asiste lo hiciese entonces mejoraría su rendimiento. Lo anterior se puede expresar en los siguientes supuestos

$$\begin{aligned} P(Y_1 = 1|Z = 0) &\geq P(Y_0 = 1|Z = 0) \\ P(Y_0 = 1|Z = 1) &\leq P(Y_1 = 1|Z = 1) \end{aligned}$$

Posteriormente, reemplazando en la ecuación (2) y (3) respectivamente se obtienen los siguientes intervalos de identificación parcial

$$\begin{aligned} P(Y = 1) \leq P(Y_1 = 1) \leq P(Y = 0, Z = 0) + P(Y = 1) \\ P(Y = 1, Z = 0) \leq P(Y_0 = 1) \leq P(Y = 1) \end{aligned}$$

Luego, utilizando la información proporcionada, en la Figura 5 se muestran los límites obtenidos para cada caso.

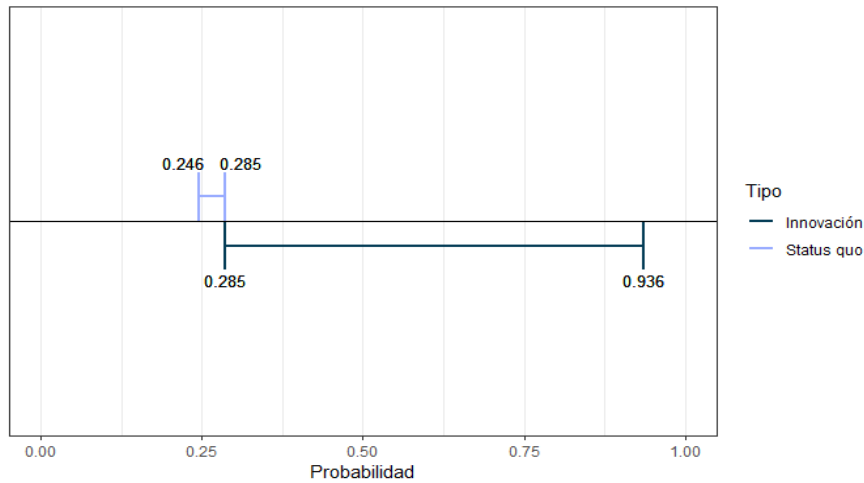


Figura 5: Intervalo de Identificación Parcial Caso 2

Analizando la figura 5 es posible apreciar que al suponer que la innovación es mejor que el status quo, la probabilidad de que si todos los estudiantes participaran en al menos k tutorías la probabilidad de mejorar su rendimiento sería al menos 29% y a lo más 94%. Ahora bien, en el caso contrario, si todos los estudiantes no participasen de las tutorías, entonces la probabilidad de mejorar será de al menos 25% y como máximo 29%.

Caso 3: Es mejor lo observado que lo contrafactual

Otro supuesto es creer que es mejor lo observado que lo contrafactual. Este contexto se justifica con el ejemplo de que si un estudiante que asiste al menos a k tutorías no lo hiciera, su probabilidad de éxito disminuiría; por el contrario, el resultado sería peor para las personas que no participan si lo hicieran. Esto lo podemos denotar como

$$\begin{aligned} P(Y_1 = 1|Z = 0) &\leq P(Y_0 = 1|Z = 0) \\ P(Y_1 = 1|Z = 1) &\geq P(Y_0 = 1|Z = 1) \end{aligned}$$

Reemplazando en las ecuaciones (2) y (3) se derivan los siguientes intervalos

$$\begin{aligned} P(Y = 1, Z = 1) &\leq P(Y_1 = 1) \leq P(Y = 1) \\ P(Y = 1, Z = 0) &\leq P(Y_0 = 1) \leq P(Y = 1) \end{aligned}$$

Luego, utilizando la información proporcionada, en la Figura 6 se muestran los límites obtenidos para cada caso. Si se supone que el rendimiento observado es mejor que si el estudiante estuviera en su situación contrafactual, al haber ido como mínimo a k tutorías, su probabilidad de éxito sería de al menos 4% y en el mejor escenario un 29%. Ahora bien, si todos los estudiantes que no asistieran al número de tutorías lo hicieran, entonces la probabilidad de mejorar su rendimiento está acotada por 25% y 29%.

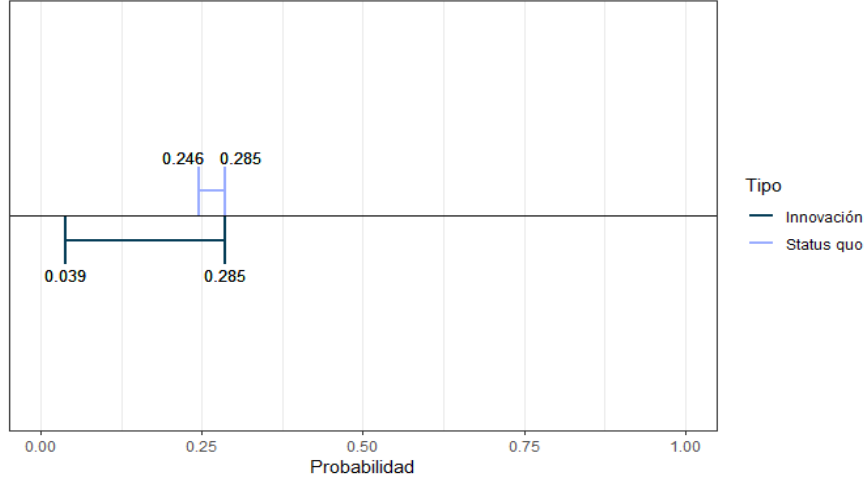


Figura 6: Intervalo de Identificación Parcial Caso 3

Caso 4: Lo contrafactual es mejor que lo observado

Finalmente, también se supondrá que lo contrafactual es mejor que lo observado. En el contexto del problema se interpreta por ejemplo como el caso en que si un estudiante que asiste al menos a k tutorías no lo hiciera, su rendimiento en el segundo semestre mejoraría; caso contrario, si un estudiante que no participa, decidiría optar por asistir a las tutorías, su rendimiento mejoraría. De esta forma, lo anterior lo denotamos como

$$\begin{aligned} P(Y = 1) &\leq P(Y_1 = 1) \leq P(Y = 1) + P(Y = 0, Z = 0) \\ P(Y = 1) &\leq P(Y_0 = 1) \leq P(Y = 1, Z = 0) + P(Z = 1) \end{aligned}$$

Reemplazando en las ecuaciones (2) y (3), se derivan los siguientes intervalos

$$P(Y = 1) \leq P(Y_1 = 1) \leq P(Y = 1) + P(Y = 0, Z = 0)$$

$$P(Y = 1) \leq P(Y_0 = 1) \leq P(Y = 1, Z = 0) + P(Z = 1)$$

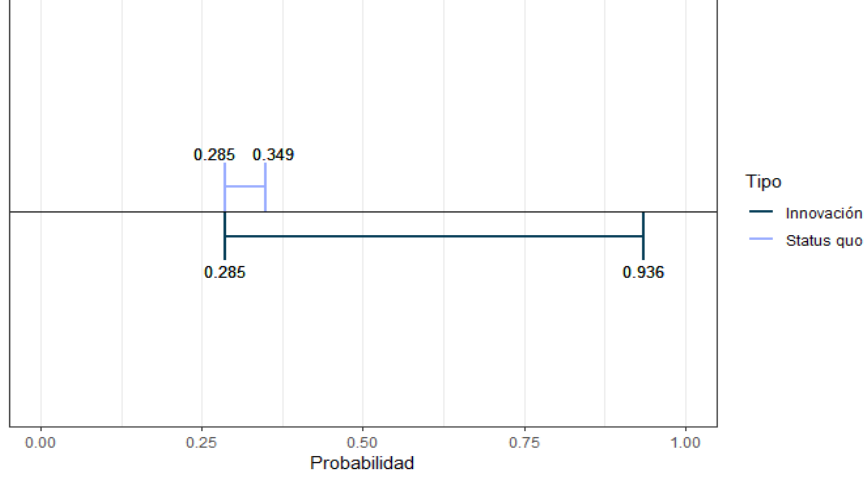


Figura 7: Intervalo de Identificación Parcial Caso 4

La figura 7 muestra los límites de los intervalos de las probabilidades $P(Y_1 = 1)$ y $P(Y_0 = 1)$, en el caso de suponer que el rendimiento de los estudiantes sería mejor si los estudiantes se les asignase su respectivo contrafactual. De esta manera se concluye que si todos los alumnos fuesen intervenidos asistiendo a las tutorías, entonces la probabilidad de superar su promedio en el primer semestre será de, a lo más 94% y en el peor escenario un 29%. Para los estudiantes que pertenecen al status quo su probabilidad de mejorar es al menos un 29% y en el caso más favorable será de 35%.

3.3 Estudio de sensibilidad de la elección de k

Tal como se observa en el gráfico 3, se aprecia que al aumentar en una tutoría no se mejora significativamente la probabilidad de éxito que puedan llegar a tener los estudiantes. Es por este motivo que se realiza un estudio de sensibilidad de la elección de k para determinar cuánto puede llegar a afectar las conclusiones y estimaciones de las probabilidades.

Tal como se puede observar en la figura 8, existen diferencias en las estimaciones de las probabilidades utilizando distintos valores de k , es decir, el ancho de los intervalos aumenta o disminuye. Por ende, al modificar el número de tutorías que al menos debe asistir un alumno también se altera el impacto en la probabilidad de éxito. Si bien, las probabilidades asociadas a cada intervalo cambian, se siguen manteniendo "las mismas estructuras". Por ende las conclusiones serán similares pero tendrán distintos rangos de variabilidad. Esto evidencia que las conclusiones no solo dependen de los datos, sino que además de los supuestos. [2]

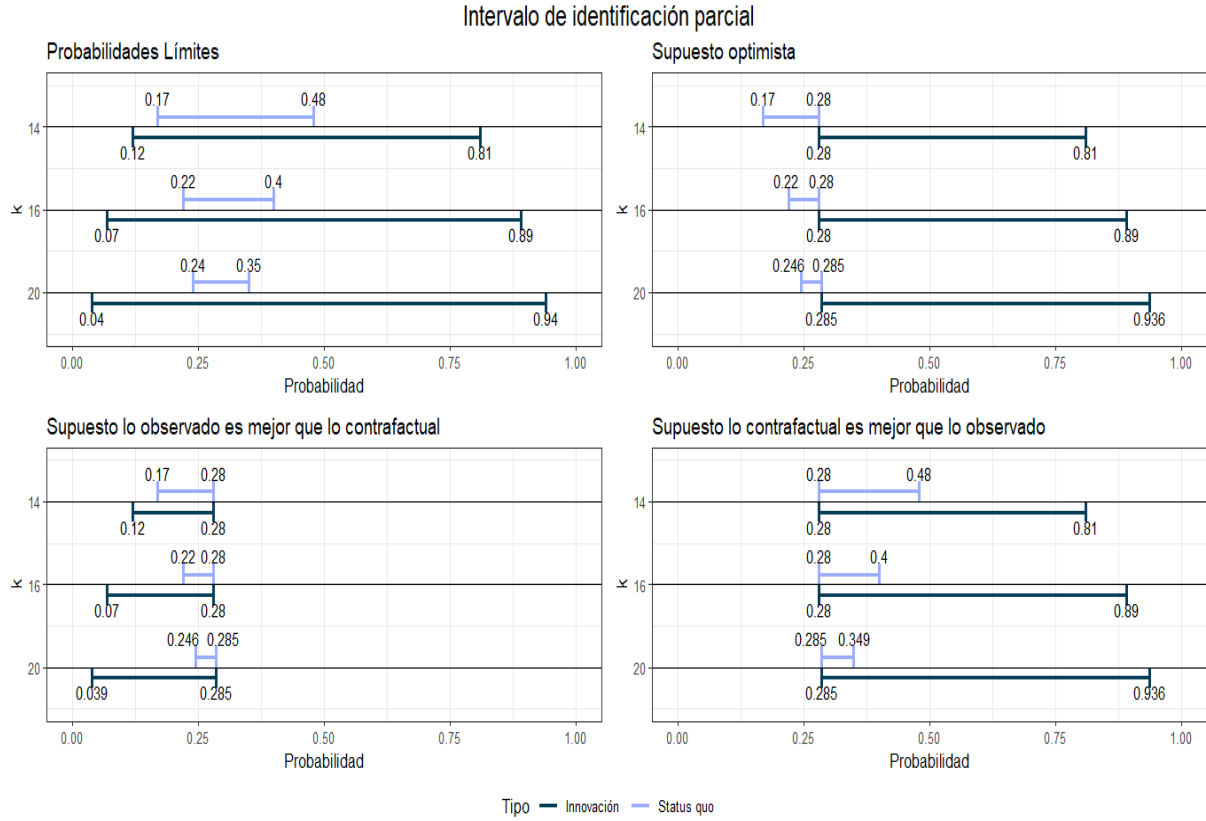


Figura 8: Intervalos de identificación parcial según supuesto y elección de k

4 Recomendaciones

4.1 Problemática 2

Podemos resumir la información de los intervalos obtenida en la sección 3.2 en la Tabla 4

Caso	$P(Y_1 = 1 X)$		$P(Y_0 = 1 X)$	
	Límite Inferior	Límite Superior	Límite Inferior	Límite Superior
Probabilidad Límite	0.04	0.94	0.24	0.35
Optimista	0.29	0.94	0.25	0.29
Factual > contrafactual	0.04	0.29	0.25	0.29
Factual < contrafactual	0.29	0.94	0.29	0.35

Tabla 4: Resumen intervalos de identificación parcial según cada supuesto

Caso 1: Probabilidades Límite

Como primera aproximación para determinar los intervalos de $P(Y_1 = 1)$ y $P(Y_0 = 1)$, se utilizó los límites de las probabilidades relacionadas con los eventos contrafactuales, es decir, que las probabilidades están acotadas por 0 y 1. De esta manera, se considera todos los posibles valores de $P(Y_1 = 1)$ y $P(Y_0 = 1)$. Sin embargo, dichos intervalos no entregan mayor información para elegir una decisión sobre

si los estudiantes deben asistir a las tutorías o no, pues si la $P(Y_1 = 1) > 0.35$ es conveniente que todos los estudiantes asistan a las tutorías, por el contrario, si $P(Y_1 = 1) < 0.24$, entonces no es recomendable que los estudiantes participen el programa de tutorías el segundo semestre.

Por ende, al utilizar las probabilidades límites se obtienen todos los posibles valores, además no se realizan supuestos. La información obtenida no entrega una evidencia significativa para escoger una opción.

Caso 2: Solución Optimista

Como se mencionó anteriormente, la solución optimista hace referencia que para obtener mejores resultados el estudiante debe asistir a las tutorías, y si no asistiese su resultado es peor. Por el contrario, si un alumno que no es intervenido fuese inscrito en el programa, la probabilidad de mejorar su rendimiento aumentaría. Esto se puede justificar, ya que, en el caso más desfavorable en que las tutorías no aporten en el aprendizaje de los estudiantes, su rendimiento debiese ser al menos igual como si no participara del programa.

Luego, si se supone que la innovación es siempre mejor, entonces la decisión que se debiese tomar es que todos los estudiantes asistan a las tutorías, pues en el caso más desfavorable el resultado sería igual que no asistir. Gráficamente se puede apreciar en la figura 6, pues los intervalos son disjuntos y las probabilidades de mejorar el rendimiento en el segundo semestre son siempre mayores o iguales a la innovación.

Caso 3: Es mejor lo observado que lo contrafactual

En este caso se supone que es mejor lo factual que lo contrafactual. En otras palabras, si un estudiante que participa del programa no lo hiciese, su probabilidad de mejorar en el segundo semestre disminuiría; del mismo modo, si el estudiante que no asiste a las tutorías decidiera participar, su rendimiento empeoraría.

Lo anterior es posible justificarlo en el caso de existir un mecanismo de selección al momento de asignar tutorías a los estudiantes, es decir, que si un estudiante asiste al programa es porque lo requiere, por lo tanto, mejoraría su promedio en el segundo semestre, de lo contrario, no asistir sería perjudicial. Por otra parte, los alumnos que no son asignados no necesitan del programa, y en el caso de ser asignados disminuiría su rendimiento, pues, por ejemplo estudiando individualmente obtendrían mejores resultados.

A diferencia del Caso 2, al decidir si todos los estudiantes debiesen asistir o no las tutorías, un grupo será perjudicado. Lo anterior se puede reflejar en la intersección de los intervalos de identificación parcial de la figura 6. Por lo tanto, la decisión se debe centrar en perjudicar la menor cantidad de alumnos posible. Por consiguiente, al suponer que lo factual es mejor que lo contrafactual se recomienda que los estudiantes no asistan a las tutorías, debido a que, si todos asistieran a las tutorías es más probable que perjudique a un mayor número de alumnos, puesto que, podría suceder que las probabilidades de mejorar en el segundo semestre sea de 4%. Por el contrario, en el caso más desfavorable en caso de no asistir, la probabilidad de éxito es de un 25%.

Caso 4: Es mejor lo contrafactual que lo observado

Finalmente, se supone que lo contrafactual es mejor que lo observado. En otras palabras, es suponer que los estudiantes que participan del programa mejorarían su rendimiento si no asistiesen a las tutorías. En el caso contrario, si no asisten a las tutorías mejorarían su promedio en el segundo semestre si participaran del programa. Luego, es posible justificar este supuesto si el mecanismo de selección mencionado en el caso anterior no funcione adecuadamente y sugiera incorrectamente la asignación de tutorías, es decir, las personas que no asisten debiesen participar, y viceversa.

Al igual que en el caso anterior, al decidir por el status quo o la innovación perjudicará un grupo, puesto que, como se muestra en la figura 7, existe intersección entre los intervalos, por lo que nuevamente la

decisión se debe basar en perjudicar la menor cantidad de alumnos posible. De este modo, se debe decidir por la innovación, pues en el mejor escenario ya que la probabilidad de éxito si todos los estudiantes no asistiesen a las tutorías será de 0.35, mientras que si participaran del programa será de 0.94.

5 Conclusión

El objetivo de este informe es realizar una evaluación de impacto para las tutorías realizadas a los beneficiarios de la Beca de Acompañamiento PAIEP, con la finalidad de sugerir si los estudiantes deben participar en el programa de tutorías de la universidad o no.

Para ello, en primer lugar, se determina el número de tutorías que generan un mayor impacto en la probabilidad de superar diferentes valores para el promedio del primer semestre. De este modo, se busca establecer el número de tutorías para definir que un estudiante es intervenido.

Luego, se realiza la evaluación de impacto. En este informe se utilizaron 4 casos, el primero con las probabilidades límite, el supuesto de optimista, suponer que lo factual es mejor que lo contrafactual y viceversa. En esta sección se concluye que los supuestos junto con los datos disponibles orientan la toma de decisiones, ya que, con supuestos diferentes se pueden concluir decisiones diferentes. Además, se realizó un análisis de sensibilidad para el número de tutorías, donde se concluye que cambian los valores de los intervalos, no obstante, no cambian las conclusiones.

Posteriormente, se realizan recomendaciones según el supuesto. Pues dependiendo de éste, se concluirá si se recomienda asistir a las tutorías o no. De modo general, sin considerar la disposición de fondos por parte de la universidad y asumiendo que los estudiantes irán al menos a 20 tutorías, se recomienda asistir al programa de tutorías, pues es razonable suponer que en el escenario más desfavorable la probabilidad de mejorar en el segundo semestre sea al menos como si no participara del programa. Además, como se menciona anteriormente, los beneficiarios de esta beca no solamente reciben apoyo académico sino que también los ayudan en su proceso de transición a la universidad.

References

- [1] ASAMBLEA DE ACADÉMICOS/AS/ES AUTOCONVOCADOS DE LA UNIVERSIDAD DE CHILE. (2020, FEBRUARY 24).. *Otro sistema de ingreso a la Educación Superior para un Chile que despertó..* CIPER Chile; Fundación CIPER. <https://www.ciperchile.cl/2020/02/24/otro-sistema-de-ingreso-a-la-educacion-superior-para-un-chile-que-desperto/>
- [2] SAN MARTÍN E, GONZÁLEZ-LARRONDO T. *How to Broker the Evaluation of Public Policies? A Strategy to Reveal Inherent Uncertainties in Policy-Decision Making. 2020.*
- [3] Quiénes Somos. PAIEP — Universidad de Santiago de Chile [Internet]. 2017 [citado 27 de Abril de 2023]. Disponible desde: <https://www.paiep.usach.cl/somos>
- [4] Programa de Acceso Inclusivo, Equidad y Permanencia (PAIEP) [Internet]. Facultad de Ciencia. 2021 [citado 27 de Abril de 2023]. Disponible desde: <https://fciencia.usach.cl/programa-de-acceso-inclusivo-equidad-y-permanencia-paiep>