



Análisis causal de los retrasos en la aprobación de materias de primer año en una carrera de Ingeniería mediante técnicas de inferencia

Causal analysis of first-year course approval delays in an engineering major through inference techniques.

PAZ, Hugo Roger¹

Paz, H. R. (2023). Análisis causal de los retrasos en la aprobación de materias de primer año en una carrera de Ingeniería mediante técnicas de inferencia. *RELAPAE*, (19), pp. 172-185.

Resumen

En este estudio se exponen las técnicas de inferencia causal utilizadas para descubrir los factores subyacentes que contribuyen a los retrasos prolongados en la aprobación de cursos de primer año que se observan en la Carrera de Ingeniería Civil de la Facultad de Ciencias Exactas y Tecnología (FACET) de la Universidad Nacional de Tucumán (UNT), Argentina. Hasta el año 2019, el estudiantado tarda un promedio de 5 años en aprobar sus asignaturas del primer año. Determinar las causas de ese retraso fue el motivo de la presente investigación. Nuestro análisis se basó en la utilización del marco de trabajo *DoWhy* (Blöbaum et al., 2022; Sharma y Kiciman, 2020) y Causal Discovery Toolbox (CDT) (Kalainathan et al., 2020), herramientas diseñadas para la inferencia causal. Nuestro objetivo era desentrañar la red de variables que influyen en los tiempos de aprobación de los cursos. Los hallazgos del estudio iluminan un factor causal fundamental que conduce a un aumento estadísticamente significativo en la duración de la aprobación. Específicamente, se reveló que la estructura regulatoria del programa académico y los métodos de evaluación empleados juegan un papel sustancial. En particular, la acumulación de materias regulares que aún no han sido aprobadas mediante un examen final surgió como un determinante clave. Este conocimiento tiene el potencial de guiar intervenciones específicas, mejorando tanto las tasas de éxito del estudiantado como la eficacia general del sistema educativo.

Palabras Clave: inferencia causal, retrasos en la aprobación de cursos, educación en ingeniería, curriculum, ingeniería civil.

Abstract

In this study, we present the causal inference techniques employed to uncover the underlying factors contributing to prolonged delays in the approval of first-year courses observed in the Civil Engineering program at the Faculty of Exact and Technological Sciences (FACET) of the National University of Tucumán (UNT), Argentina. Up until 2019, students took an average of 5 years to pass their first-year subjects. Determining the causes of this delay was the main objective of this investigation. Our analysis was grounded in the utilization of the *DoWhy* framework (Blöbaum et al., 2022; Sharma and Kiciman, 2020) and the Causal Discovery Toolbox (CDT) (Kalainathan et al., 2020), tools designed for causal inference. Our aim was to unravel the network of variables influencing the course approval timelines. The study's findings shed light on a fundamental causal factor that leads to a statistically significant increase in the duration of approval. Specifically, it was revealed that the regulatory structure of the academic program and the assessment methods employed played a substantial role. In particular, the accumulation of regular subjects that had not been approved through a final examination emerged as a key determinant. This knowledge has the potential to guide specific interventions, enhancing both students' success rates and the overall effectiveness of the educational system.

Keywords: causal inference, course approval delays, engineering education, curriculum, civil engineering.

¹ Facultad de Filosofía y Letras- Universidad Nacional de Tucumán, Argentina / hpaz@herrera.unt.edu.ar / <https://orcid.org/0000-0003-1237-7983>

1. Introducción

Este artículo profundiza en las complejidades del retraso en la progresión curricular en la educación superior mediante la investigación de las causas subyacentes en el contexto del plan de estudios actualizado de 2005 del programa de Ingeniería Civil en la Universidad Nacional de Tucumán. El plan de estudios de cinco años y medio, dividido en once semestres, está regido por un Sistema de Correlativas que establece requisitos previos para la inscripción en los cursos. Estos requisitos previos pueden ser "regularizados", indicando la aprobación de evaluaciones parciales o tareas prácticas, o "aprobados", indicando la finalización exitosa de exámenes finales lo que implica que aprobó el curso.

En este estudio, partimos de la hipótesis de que el sistema de evaluaciones existente contribuye a la acumulación de asignaturas "regulares", pero no aprobadas entre el estudiantado. Sostenemos que este fenómeno surge de la falta de tiempo disponible para una adecuada preparación de los exámenes finales, lo que, a su vez, genera retrasos en el proceso de aprobación de los cursos. Nuestro objetivo principal radica en demostrar empíricamente esta hipótesis subyacente, mediante la aplicación de técnicas rigurosas de inferencia causal. A través de este enfoque, buscamos arrojar luz sobre las conexiones causales entre el sistema de evaluaciones, la acumulación de materias pendientes y los retrasos en la progresión curricular.

Como destacó Pearl (2009), la inferencia causal permite la identificación de relaciones causales más allá de las meras correlaciones. El énfasis en la inferencia causal surge de la necesidad de descubrir los verdaderos factores causales que impulsan los retrasos en la progresión curricular. "Los hallazgos de la investigación observacional pueden ser inconsistentes o consistentes pero es poco probable que reflejen verdaderas relaciones de causa y efecto" (Hammerton y Munafò, 2021).

Como concluyen Oloriz et al en un trabajo en donde estudian la relación entre el rendimiento académico de los ingresantes en carreras de ingeniería y el abandono de los estudios universitarios, "existe una importante correlación entre el rendimiento académico en el primer cuatrimestre... y la deserción que se produce durante el segundo, tercero y cuarto cuatrimestre" (Oloriz et al., 2007, p. 1), es por ello que resulta necesario trabajar en la disminución del fracaso académico dado que el mismo conduce, inexorablemente, al abandono de los estudios superiores.

2. Método

2.1. Población y Muestra

Los datos objeto de análisis provienen de las trayectorias académicas de 1615 estudiantes matriculados en la carrera de Ingeniería Civil, abarcando un espectro diverso de avances en la carrera, desde estudiantes de primer año hasta aquellos en etapas cercanas a la graduación, así como individuos que han discontinuado su formación y aquellos que han culminado su proceso educativo. Estos datos se obtuvieron a través del Sistema de Gestión Administrativa y de Estudiantes (SIGEA), que recopila y centraliza información relevante sobre el progreso académico y la gestión del estudiantado en el programa.

En el marco de esta investigación, se ha procedido a examinar el rendimiento académico de estudiantes universitarios a lo largo de un período de catorce años, que comprende desde 2004 hasta 2019. Con el propósito de asegurar la precisión de los resultados, se efectuó una exclusión de estudiantes que habían obtenido la aprobación de asignaturas mediante un sistema de equivalencias, debido a modificaciones en su plan de estudios, trayectoria académica o institución educativa. Dicha medida se implementó considerando que estos alumnos podrían haber ingresado al programa con asignaturas ya aprobadas, lo cual podría impactar en su progreso y desempeño en comparación con sus pares. Para el análisis, se procedió a seleccionar un conjunto de 1343 historias académicas de un total de 1615.

En el presente contexto, se clasifica como "egresados" a estudiantes que han finalizado sus estudios dentro del período bajo evaluación (2005-2019). Por su parte, se considera "permanentes" a estudiantes que han mantenido su inscripción en el programa académico hasta el año 2019 y aún no han concluido su formación.

En última instancia, se define la situación de "deserción" para aquellos alumnos que no registran ninguna inscripción en el año 2020 y no han culminado su proceso educativo. Al adoptar este enfoque, seguimos la propuesta de González Fiegehen (2007) para los alumnos que se matricularon en 2019, representando únicamente el 5% del total analizado (Paz, 2022). Para el resto de estudiantes, se emplea un criterio menos riguroso, incluyendo a aquellos que retoman su inscripción en la carrera y reanudan sus estudios tras un período de inactividad. El cálculo de los lapsos temporales se ha basado en la diferencia entre la última actividad documentada en el sistema SIGEA y la fecha de ingreso al programa académico.

Con relación a la carrera de Ingeniería Civil, el Plan de Estudios se desarrolla en once módulos semestrales con una duración total de cinco años y medio. El total de asignaturas que contiene el Plan de Estudios es de 44. Las asignaturas que comprenden los módulos 1 y modulo 2 son las siguientes:

- 1 Cálculo: Modulo 1
- 2 Física I: Modulo 1
- 3 Sistemas De Representación: Modulo 1
- 4 Álgebra Y Geometría Analítica: Modulo 1
- 5 Cálculo II: Modulo 2
- 6 Física II: Modulo 2
- 7 Fundamentos De Química General: Modulo 2
- 8 Informática: Modulo 2
- 9 Elementos De Álgebra Lineal: Modulo 2

2.2. Técnica de análisis

Para la determinación de los factores subyacentes que contribuyen a los retrasos prolongados en la aprobación de cursos de primer año se utilizaron técnicas de inferencia causal. Este tipo de técnicas busca comprender y establecer relaciones de causa y efecto entre variables.

Como lo destacó Judea Pearl, un pionero en el campo de la inferencia causal, a través de su obra "*Causality: Models, Reasoning, and Inference*" se trata de identificar cómo un cambio en una variable influye directamente en otra, lo que implica la existencia de una relación de causa y efecto (Pearl, 2009). Este enfoque es esencial para responder preguntas críticas en diversas disciplinas, como la medicina, la economía, la epidemiología y las ciencias sociales. A través de técnicas avanzadas, como el uso de experimentos controlados y modelos causales, se pueden descubrir patrones subyacentes y comprender el impacto real de las intervenciones o factores en un sistema. Este enfoque es esencial para la toma de decisiones basadas en evidencia y la formulación de políticas efectivas, lo que contribuye significativamente a la generación de conocimiento sólido en numerosos campos de investigación (Pearl, 2009).

El método se aplicó mediante la utilización de la librería *DoWhy* (Blöbaum et al., 2022; Sharma y Kiciman, 2020) y Causal Discovery Toolbox (CDT) (Kalainathan et al., 2020), herramientas diseñadas para la inferencia causal. La aplicación de estas librerías se realizó mediante la generación de un código en lenguaje Python (Van Rossum & Drake Jr, 1995). Python es un lenguaje de programación ampliamente utilizado en el ámbito científico y de análisis de datos debido a su versatilidad.

La librería *DoWhy* destaca como una herramienta esencial para realizar inferencia causal avanzada en *Python*. *DoWhy* se basa en la teoría del diseño de experimentos y la inferencia causal, y permite identificar relaciones de causa y efecto más allá de simples correlaciones. Esta librería se utiliza para modelar y analizar gráficos causales acíclicos dirigidos (DAG) que representan las relaciones de dependencia causal entre las variables de interés. Además, *DoWhy* ofrece diversos métodos de estimación para calcular el efecto causal promedio, lo que facilita la evaluación de las intervenciones en situaciones del mundo real. La librería *DoWhy* es una herramienta fundamental para llevar a cabo análisis rigurosos de inferencia causal, permitiendo explorar y comprender las relaciones de causa y efecto en una amplia variedad de contextos.

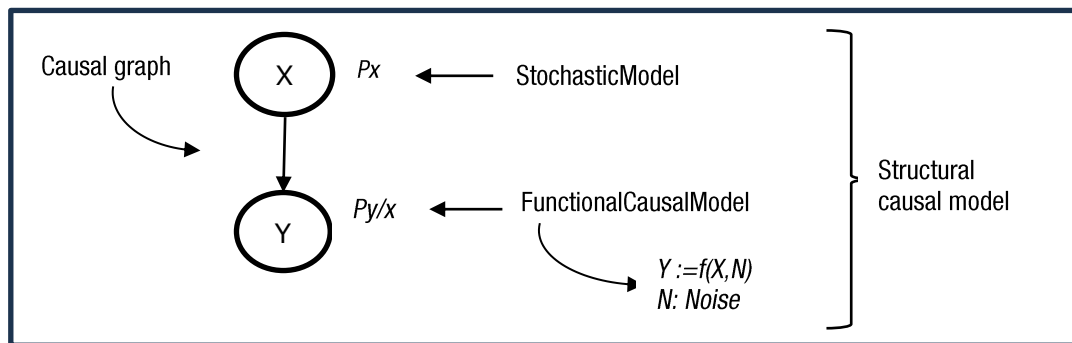
DoWhy se basa en la teoría del diseño de experimentos y la inferencia causal. Su enfoque se fundamenta en el concepto de "intervención", donde se manipulan variables para observar cómo afectan a otras variables en un sistema.

Una vez identificado un efecto causal, se pueden adoptar diferentes métodos de estimación compatible con la estrategia de identificación. Para estimar el efecto causal promedio, *DoWhy* admite los siguientes métodos: *Distance-based matching* (Chen et al., 2007), *Propensity-based Stratification* (Desai et al., 2017), *Propensity Score Matching* (Caliendo & Kopeinig, 2008; Rosenbaum & Rubin, 1983), *Linear Regression, Generalized Linear Models (including logistic regression)* (Chernozhukov et al., 2022), *Binary Instrument/Wald Estimator* (Jiang et al., 2023), *Regression discontinuity* (Imbens & Lemieux, 2008), *Two-stage linear regression* (Angrist & Imbens, 1995).

DoWhy ejecuta una amplia gama de técnicas de inferencia causal, incluyendo el estimado del efecto medio del tratamiento (Average Treatment Effect, ATE) (Crump et al., 2008) y otras medidas de impacto causal. La herramienta también permite la estimación de intervalos de confianza y pruebas de significancia estadística para respaldar los resultados obtenidos. La secuencia de análisis conlleva la ejecución de los siguientes pasos:

A. Definición del Grafo Causal: El primer paso es modelar el grafo causal, que representa las relaciones entre las variables de interés y cómo se influyen mutuamente. Esto ayuda a visualizar el sistema y a comprender las conexiones subyacentes. Un grafo causal (ver Figura 1) modela las relaciones causales, o “relaciones causa-efecto” presentes en un sistema o dominio de problema. En *DoWhy*, se requiere que el gráfico causal sea un gráfico acíclico dirigido (DAG) donde una arista $X \rightarrow Y$ implica que X causa Y . Estadísticamente, un gráfico causal codifica las relaciones de independencia condicional entre variables.

Figura 1. Estructura del Modelo Causal



Fuente: (Blöbaum et al., 2022, p. 2)

Uno de los elementos esenciales en el proceso de análisis causal con *DoWhy* es la definición del grafo causal, que captura las interacciones y relaciones entre las variables de interés. Sin embargo, este paso no es una tarea aislada; más bien, requiere un profundo conocimiento previo sobre el problema en cuestión y las posibles relaciones causales, basado en la experiencia y la intuición.

El proceso de definir un grafo causal dentro de *DoWhy* es como la creación de un mapa que guiará el análisis hacia la comprensión de las relaciones causales subyacentes. Este mapa no puede ser generado únicamente por algoritmos o enfoques automatizados, ya que requiere un entendimiento profundo del dominio específico. Aquí es donde entra en juego el conocimiento previo.

El conocimiento previo es la amalgama de la experiencia, la intuición y las suposiciones informadas. Aquellos que están familiarizados con el problema en cuestión pueden identificar las posibles variables influyentes y las direcciones probables de las relaciones causales. Algunas suposiciones pueden estar basadas en investigaciones anteriores, teorías fundamentales o incluso en las observaciones más simples.

La definición del grafo causal se basa en la idea de que muchas relaciones en el mundo real pueden ser interpretadas como gráficos de causalidad inferida (CI). Un gráfico de CI es una representación visual de las relaciones causales entre variables, donde las flechas indican las influencias directas. En este contexto, modelar el grafo causal en *DoWhy* implica mapear las relaciones de causa y efecto del problema en un formato gráfico.

El proceso de definición del grafo no es meramente técnico; es un arte respaldado por la experiencia. Los expertos en el dominio son quienes pueden aportar ideas valiosas y perspicacia a la tarea. Su capacidad para identificar conexiones no evidentes y captar las sutilezas de las interacciones causales enriquece el análisis y garantiza que el grafo sea una representación fiel de la realidad.

La definición del grafo causal en *DoWhy* es una parte fundamental del proceso de análisis causal. Requiere el aporte del conocimiento previo basado en la experiencia y las suposiciones informadas. Este enfoque colaborativo entre la experiencia humana y las herramientas tecnológicas garantiza que el análisis causal sea un proceso riguroso y orientado a resultados en la búsqueda de conexiones significativas entre las variables.

B. Identificación y Estimación del Efecto Causal: Una vez que el grafo está definido, *DoWhy* utiliza datos observacionales y métodos de inferencia causal para estimar el efecto de las intervenciones en las variables de interés. Esto implica la selección de variables, el cálculo de puntajes de propensión si es necesario y la aplicación de técnicas de ajuste.

Una vez que el grafo causal está definido, se selecciona el tratamiento (variable independiente) y el resultado (variable dependiente) que se investigarán para estimar su efecto causal. *DoWhy* se basa en el concepto de intervención, donde se simula la manipulación del tratamiento para observar su impacto en el resultado.

En muchos casos, los datos observacionales pueden estar sujetos a sesgos debido a la autoselección en los tratamientos. *DoWhy* utiliza el método de "propensity score matching" para estimar la probabilidad de recibir el tratamiento (propensity score) basada en otras variables observadas. Esto ayuda a crear grupos comparables y a ajustar posibles sesgos.

DoWhy ofrece una variedad de métodos de estimación de efectos causales, como el método de "Matching", "Regression Discontinuity", "Instrumental Variables", entre otros. Estos métodos se aplican a los datos observacionales ajustados y controlados, para evaluar cómo cambia el resultado cuando se manipula el tratamiento.

Después de estimar el efecto causal, *DoWhy* permite evaluar la significancia estadística de los resultados. Esto se logra mediante la estimación de intervalos de confianza y la realización de pruebas de hipótesis para determinar si el efecto observado es estadísticamente significativo.

Una vez obtenidos los resultados, *DoWhy* facilita la interpretación de los efectos causales estimados. Los valores obtenidos, junto con los intervalos de confianza y las pruebas de significancia, proporcionan información valiosa sobre cómo el tratamiento afecta al resultado y si esta relación es plausible desde una perspectiva causal.

DoWhy realiza un proceso integral de identificación y estimación del efecto causal, combinando la teoría causal con métodos estadísticos avanzados. A través de la definición del grafo causal, la selección de variables, el uso del propensity score y la aplicación de métodos de estimación, *DoWhy* arroja luz sobre las relaciones causales en los datos observacionales.

C. Evaluación de la Validez del Modelo: *DoWhy* permite evaluar la significancia estadística de los resultados obtenidos, lo que proporciona una medida objetiva de la confiabilidad de los efectos causales estimados.

Se debe tener en cuenta que la parte causal no proviene de datos. Proviene de las suposiciones que conducen a la identificación realizadas en el punto A. Los datos se utilizan simplemente para la estimación estadística. Por lo tanto, se vuelve fundamental verificar si nuestras suposiciones fueron correctas en el primer paso o no. ¿Qué sucede cuando existe otra causa común? ¿Qué sucede cuando el tratamiento en sí fue placebo? Para ello se utilizan tres métodos de refutación de resultados:

Método 1: Causa común aleatoria. Agrega covariables extraídas aleatoriamente a los datos y vuelve a ejecutar el análisis para ver si la estimación causal cambia o no. Si nuestra suposición era originalmente correcta, entonces la estimación causal no debería cambiar mucho.

Método 2: Refutador de tratamiento con placebo. asigna aleatoriamente cualquier covariable como tratamiento y vuelve a ejecutar el análisis. Si nuestras suposiciones fueran correctas, entonces esta estimación recién descubierta debería llegar a 0.

Método 3: Refutador de subconjunto de datos. crea subconjuntos de datos (similar a la validación cruzada) y verifica si las estimaciones causales varían entre los subconjuntos. Si nuestras suposiciones fueran correctas, no debería haber mucha variación.

Como se ha mencionado anteriormente, la aplicación de las técnicas antes mencionadas se realizó mediante el desarrollo de un código en lenguaje *Python* (Van Rossum & Drake JR, 1995). Se utilizó para el desarrollo del código las librerías específicas *DoWhy* (Sharma & Kiciman, 2020) y *Causal Discovery Volvox* (Kalainathan et al., 2020).

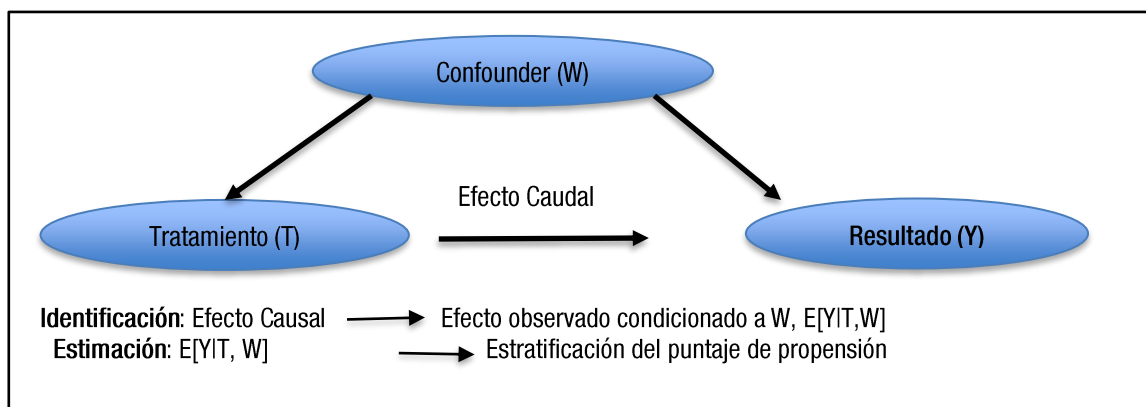
3. Resultados

3.1. Definición de las variables del problema

En este estudio, partimos de la hipótesis de que el sistema de evaluaciones existente contribuye a la acumulación de asignaturas "regulares", pero no aprobadas entre el estudiantado, lo que finalmente produce el efecto causal de generar retrasos en el proceso de aprobación de los cursos. Este análisis se enfocará en este trabajo al retraso que sufre el estudiantado en la aprobación de los cursos del primer año.

Desde este punto de vista, se pueden definir entonces las variables “tratamiento: T” (Acumulación de Asignaturas Regulares) y la variable “resultado: Y” (Tiempo que tardan el estudiantado en aprobar las asignaturas correspondientes al primer año de la carrera). La relación causal T Y involucra otras variables o “confundes” las cuales intervienen en dicha relación causal afectando tanto el tratamiento como el resultado, de acuerdo con la Figura 2:

Figura 2. Grafo de relación causal (DAG)



Con base en esto, las preguntas que se plantean son las siguientes:

¿Cuál es el impacto que tiene la acumulación de asignaturas regulares en el tiempo que tardan el estudiantado en aprobar todas las asignaturas del primer año de la carrera?

Y la pregunta contra fáctica equivalente es:

Si el estudiantado no tiene una cantidad menor a la media de acumulación de asignaturas regulares, ¿Cuál es la probabilidad de que el tiempo de aprobación descienda? O viceversa. En lenguaje formal, nos interesa el efecto promedio del tratamiento en el estudiantado (ATE).

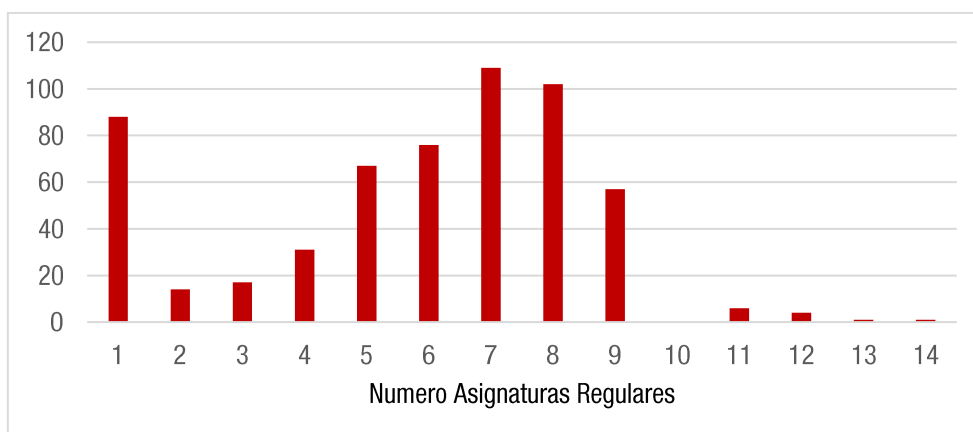
Para el análisis se utilizaron las historias académicas (ver apartado 2.1.), de las cuales se extrajeron los siguientes datos para 1343 estudiantes:

- Cohorte
- Genero
- Edad
- Tiempo Facultad
- Actividades Aprobadas
- Promedio Notas
- Número Total Cursadas
- Número Recursadas
- Número Regulares
- Número Libres
- Total Recursadas Sobre Total Cursadas
- Total Regulares-Total Cursadas
- Número de Exámenes
- Número Promociones
- Número Aprobados
- Número de Reprobados
- Número de Ausentes
- Número Promociones Sobre Total Asignaturas Promocionales
- Número Promociones Sobre Total Asignaturas
- Examen Aprobado Sobre Total Examen
- Examen Promocionado Sobre Total Examen
- Examen Reprobado Sobre Total Examen
- Examen Ausente Sobre Total Examen

- Tiene Promociones
- Tiene Libres Mas 10
- Tiene Reprobados Mas 3
- Tiene Reprobados Menor 3
- Tiempo Aprobación Módulos 1-2 (M1 y 2 = Primer año de la Carrera)
- Tiempo Aprobación M1 2 Mayor a 2 años (M1 y 2 = Primer año de la Carrera)
- Máximo de Asignaturas Regulares Acumuladas (MaxRegAcum)
- Asignaturas Regulares Acumuladas Mayor a 6 (MaxRegAcumMayor6)

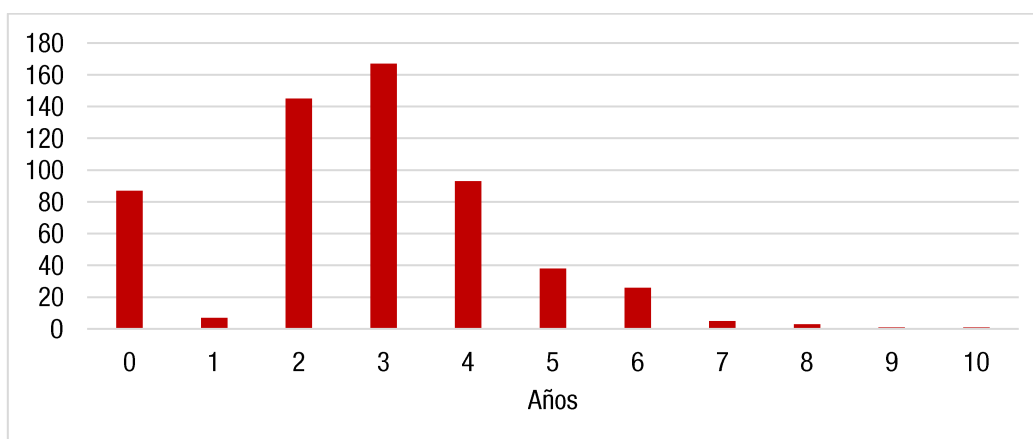
Con relación a las variables tratamiento (Asignaturas Regulares Acumuladas) y resultado (Tiempo Aprobación Módulos 1 y 2), se realizó una distribución binaria de manera de permitir la ejecución del análisis. Dicha distribución (Asignaturas Regulares Acumuladas Mayor a 6 para el tratamiento y Tiempo Aprobación M1 2 Mayor a 2 años para el resultado) se realizó tomando como base el valor medio de dichas variables para el total de la muestra. De esta manera se busca que los casos para ambas muestras estén balanceados, ya que, de otro modo le implementación del método arroja errores de inferencia (Ver Figuras 3 y 4).

Figura 3. Histograma Máximas Regularidades Acumuladas



Fuente: Elaboración propia (Base Datos SIGEA)

Figura 4. Histograma Tiempo Aprobación Asignaturas Año 1



Fuente: Elaboración propia (Base Datos SIGEA)

3.2. Definición del Grafo Causal

La generación del grafo causal en este estudio ha sido guiada por una serie de suposiciones fundamentales sobre la

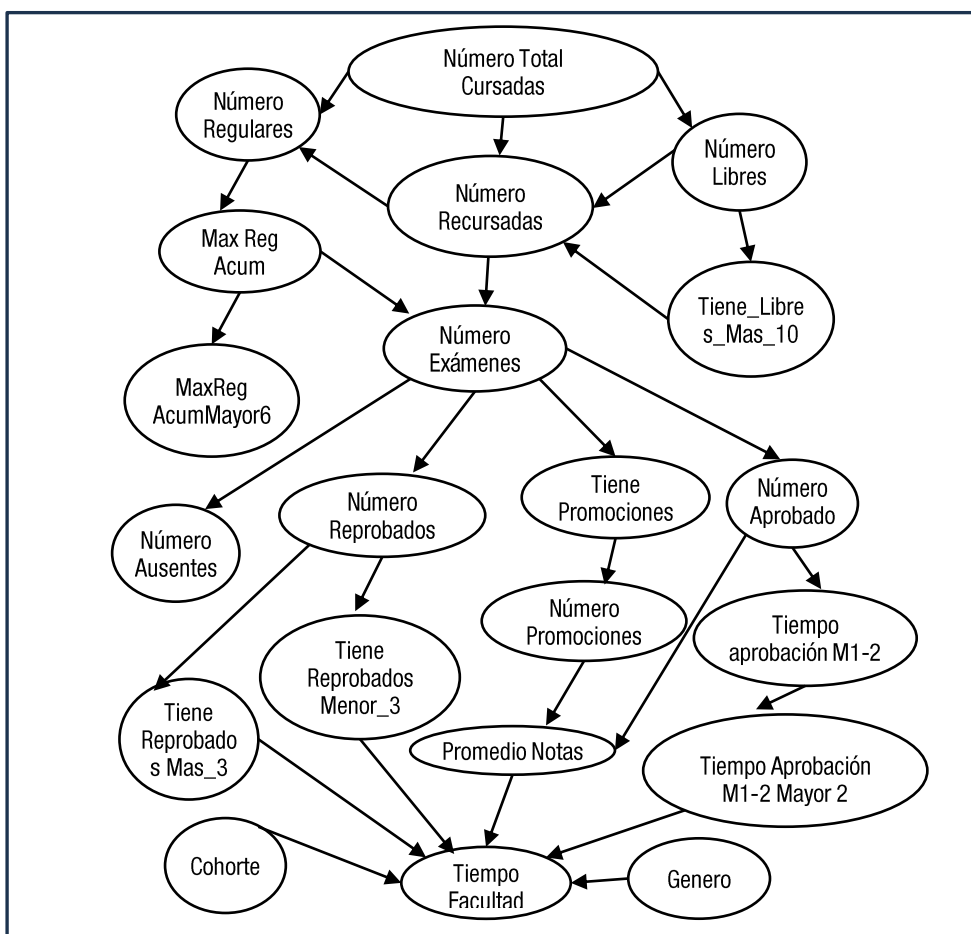
estructura del plan de estudios y las interacciones entre las actividades de evaluación, como la regularización y aprobación de asignaturas. En este proceso, se han incorporado variables de tratamiento y resultado siguiendo la distribución binaria que se detalló en la sección anterior. Estas variables esenciales fueron incorporadas para modelar las intervenciones y los efectos observados en el contexto de estudio.

Cabe destacar que el grafo causal presentado aquí representa una primera aproximación al análisis de causalidad. Las relaciones y conexiones entre las variables han sido delineadas de acuerdo con las suposiciones disponibles y el conocimiento experto. Sin embargo, para confirmar la validez y robustez de estas relaciones causales, se realizarán análisis de refutación posteriores que permitirán someter las hipótesis causales a pruebas rigurosas.

Además, como parte del proceso de modelado, se han introducido todas las variables que se consideraron pertinentes como posibles “*confounders*”. La inclusión de estas variables adicionales tiene como objetivo mejorar la precisión de la inferencia causal al controlar posibles fuentes de sesgo y confusión. Esta estrategia de ajuste busca garantizar que los efectos estimados sean atribuibles de manera más precisa al tratamiento bajo estudio.

En resumen, el proceso de generación del grafo causal ha implicado la realización de asunciones fundamentales sobre la estructura del plan de estudios y las interacciones entre las actividades de evaluación. A medida que avance el análisis, se aplicarán métodos de refutación para confirmar y fortalecer las relaciones causales propuestas en este grafo inicial. La inclusión de variables de tratamiento, resultado y “*confounders*” busca ofrecer una representación sólida y detallada de las relaciones causales subyacentes en este contexto educativo. El grafo resultante se puede observar en la Figura 5.

Figura 5. Grafo Causal Directo Acíclico (DAG)



Fuente: elaboración propia.

2. Identificación y estimación del efecto causal

La implementación de la identificación y estimación del efecto causal se llevó a cabo mediante la adopción de enfoques metodológicos específicos, en consonancia con las premisas de rigor científico y validez causal. En este estudio, se emplearon dos estimadores, a saber, "*backdoor.propensity_score_matching*" y "*backdoor.distance_matching*", con el objetivo de discernir y cuantificar los efectos causales bajo análisis. Estos estimadores se fundamentan en las nociones de "*backdoor*" y "*matching*" para abordar las posibles vías alternativas de confusión y sesgo en los datos observacionales (Caliendo & Kopeinig, 2008; Rosenbaum & Rubin, 1983).

En el marco de esta metodología, el concepto de "*backdoor*" se refiere a variables que abren vías adicionales entre el tratamiento y el resultado, lo que podría generar confusión en la estimación del efecto causal (Luo et al., 2023). Para contrarrestar este efecto, se aplicaron los estimadores de "*propensity score matching*" y "*distance matching*". El primero, basado en el puntaje de propensión, busca lograr equilibrio entre los grupos de tratamiento y control en términos de las características observables relevantes. El segundo, utilizando la distancia entre las covariables, realiza un emparejamiento más preciso y calibrado para lograr la comparabilidad de los grupos (Caliendo & Kopeinig, 2008; Rosenbaum & Rubin, 1983).

La elección de estos estimadores se justifica por su capacidad para mitigar el potencial sesgo y las posibles variables de confusión en los datos observacionales. La teoría que subyace en estos estimadores se sustenta en la idea de controlar adecuadamente las variables que podrían intervenir en la relación causal entre el tratamiento y el resultado, de manera que el efecto causal pueda ser aislado y estimado de manera más precisa.

La implementación de la Identificación y Estimación del Efecto Causal en este estudio se basa en la selección de estimadores específicos, "*backdoor.propensity_score_matching*" y "*backdoor.distance_matching*", que son respaldados por la teoría causal y metodológica. Estos estimadores abordan las posibles rutas de confusión mediante el control riguroso de variables mediante la técnica de emparejamiento. La aplicación de estos enfoques metodológicos avanza hacia una comprensión más sólida y fundamentada de los efectos causales en el contexto estudiado. Los resultados se pueden observar en la Figura 6.

Figura 6. Resultados de Identificación y Estimación del Efecto Causal

```
PROPENSITY_SCORE_MATCHING

*** Causal Estimate ***
## Identified estimand
Estimand type: nonparametric-ate
### Estimand : 1
Estimand name: backdoor
Estimand expression:
d
-----
(E[Tiempo_Aprobacion_M1_2_Mayor2|MaxRegAcum])
d[MaxRegAcumMayor6]
Estimand assumption 1, Unconfoundedness: If  $U \rightarrow \{MaxRegAcumMayor6\}$  and  $U \rightarrow Tiempo\_Aprobacion\_M1\_2\_Mayor2$ 
then  $P(Tiempo\_Aprobacion\_M1\_2\_Mayor2|MaxRegAcumMayor6,MaxRegAcum,U) = P(Tiempo\_Aprobacion\_M1\_2\_Mayor2|MaxRegAcumMayor6,MaxRegAcum)$ 

## Realized estimand
b: Tiempo_Aprobacion_M1_2_Mayor2~MaxRegAcumMayor6+MaxRegAcum
Target units: ate
## Estimate
Mean value: 0.6963350785340314
Causal Estimate is: 0.6963350785340314
Textual Interpreter
-----
Increasing the treatment variable(s) [MaxRegAcumMayor6] from 0 to 1 causes an increase of
0.6963350785340314 in the expected value of the outcome [Tiempo_Aprobacion_M1_2_Mayor2], over the data
```

distribution/population represented by the dataset.

DISTANCE_MATCHING

*** Causal Estimate ***

Identified estimand

Estimand type: nonparametric-ate

Estimand : 1

Estimand name: backdoor

Estimand expression:

$$d \frac{E[\text{Tiempo_Aprobacion_M1_2_Mayor2} | \text{MaxRegAcum}]}{d[\text{MaxRegAcumMayor6}]}$$

Estimand assumption 1, Unconfoundedness: If $U \rightarrow \{\text{MaxRegAcumMayor6}\}$ and $U \rightarrow \text{Tiempo_Aprobacion_M1_2_Mayor2}$ then $P(\text{Tiempo_Aprobacion_M1_2_Mayor2} | \text{MaxRegAcumMayor6}, \text{MaxRegAcum}, U) = P(\text{Tiempo_Aprobacion_M1_2_Mayor2} | \text{MaxRegAcumMayor6}, \text{MaxRegAcum})$

Realized estimand

b: $\text{Tiempo_Aprobacion_M1_2_Mayor2} \sim \text{MaxRegAcumMayor6} + \text{MaxRegAcum}$

Target units: ate

Estimate

Mean value: 0.6963350785340314

Causal Estimate is: 0.6963350785340314

Textual Interpreter

Increasing the treatment variable(s) [MaxRegAcumMayor6] from 0 to 1 causes an increase of 0.6963350785340314 in the expected value of the outcome [Tiempo_Aprobacion_M1_2_Mayor2], over the data distribution/population represented by the dataset.

Fuente: Elaboración propia (Salida código Python).

3.3. Evaluación de la Validez del Modelo

La evaluación rigurosa de la validez del modelo causal desplegado constituye una etapa esencial en el proceso de análisis, permitiendo la confirmación de las relaciones causales propuestas y fortaleciendo la credibilidad de los resultados obtenidos. Para lograr esto, en este estudio se han empleado refutadores basados en el enfoque contrafactual y métodos de *resampling*.

En el marco de esta evaluación, se aplicaron dos refutadores específicos: el "*random common cause*" y el "*Bootstrap Sample Dataset*". Estos refutadores buscan desafiar y poner a prueba la robustez de las conclusiones causales derivadas del modelo.

El refutador "*random common cause*" se basa en la introducción de una variable aleatoria como posible causa común, con el fin de evaluar si esta variable puede alterar los efectos estimados. Esta metodología busca verificar si el modelo es sensible a variables no contempladas en el análisis original y, por lo tanto, evalúa la posible influencia de factores no considerados en la inferencia causal (Farnè & Montanari, 2022).

Por otro lado, el refutador "*Bootstrap Sample Dataset*" opera mediante la generación de muestras de datos repetidas a partir de la muestra original (Farnè & Montanari, 2022). Esta técnica de remuestreo permite evaluar la estabilidad de los efectos causales estimados al considerar variaciones en la muestra observacional. Mediante la comparación de los resultados obtenidos en diferentes muestras generadas, se evalúa la consistencia y confiabilidad de las conclusiones causales. Los resultados de la aplicación de los métodos antes mencionados se pueden observar en la Figura 7.

Figura 7. Resultados de Evaluación de la Validez del Modelo

```
REFUTATION 1
*** Class Name ***
backdoor.propensity_score_matching
Refute: Add a random common cause
Refute: Add a random common cause
Estimated effect:0.5000000000000001
New effect:0.5000000000000001
p value:1.0
REFUTATION 2
*** Class Name ***
backdoor.propensity_score_matching
Refute: Bootstrap Sample Dataset
Refute: Bootstrap Sample Dataset
Estimated effect:0.5000000000000001
New effect:0.5054347826086957
p value:0.55
REFUTATION 3
*** Class Name ***
backdoor.distance_matching
Refute: Add a random common cause
Refute: Add a random common cause
Estimated effect:0.5000000000000001
New effect:0.5000000000000001
p value:1.0
REFUTATION 4
*** Class Name ***
backdoor.distance_matching
Refute: Bootstrap Sample Dataset
Refute: Bootstrap Sample Dataset
Estimated effect:0.5000000000000001
New effect:0.49663043478260865
p value:0.43999999999999995
```

Fuente: Elaboración propia (Salida código Python).

Los resultados de las evaluaciones de refutación muestran la robustez y validez del modelo causal propuesto. En la Refutación 1 y 3 (*Refutation 1* y *Refutation 3* de la Figura 7), donde se introduce una causa común aleatoria, se observa que los efectos estimados permanecen consistentes y no se ven alterados significativamente. Esto sugiere que el modelo es resistente a la inclusión de nuevas variables y refuerza la coherencia de las relaciones causales establecidas.

En la Refutación 2 y 4 (*Refutation 2 y Refutation 4* de la Figura 7), en las cuales se aplicó el uso del refutador "*Bootstrap Sample Dataset*", se observa una ligera variación en los efectos causales estimados. Sin embargo, esta variación se encuentra dentro de márgenes aceptables y no compromete la solidez de las conclusiones. Estos resultados refuerzan la fiabilidad y estabilidad de los efectos causales identificados en el análisis.

En conjunto, la evaluación de la validez del modelo a través de refutadores y métodos de *resampling* refuerza la confianza en las relaciones causales propuestas. Estos procedimientos proporcionan una rigurosa confirmación de la robustez de los resultados y subrayan la solidez de las conclusiones derivadas del análisis causal.

5. Discusión y Conclusiones

Los resultados derivados de la aplicación de los estimadores de *propensity score matching* y *distance matching* proporcionan valiosa información sobre el impacto del incremento de la variable de tratamiento: "más de 6 asignaturas regulares acumuladas sin examen final aprobado" (MaxRegAcumMayor6) en el valor esperado del resultado: "tiempo de aprobación de las asignaturas de los módulos 1 y 2 mayor a 2 años" (Tiempo_Aprobacion_M1_2_Mayor2). Estos resultados revelan patrones significativos que pueden tener repercusiones profundas en el contexto de la Carrera de Ingeniería Civil y la toma de decisiones educativas. En este análisis, se ha explorado cómo el acumulado de más de 6 asignaturas regulares no aprobadas se relaciona con el tiempo de aprobación de las asignaturas del primer año.

Los hallazgos indican que el aumento de la variable de tratamiento de 0 a 1 está asociado con un incremento de 0.696 en el valor esperado del tiempo de aprobación. En otras palabras, el estudiantado que acumula más de 6 asignaturas regulares sin aprobar tiene una probabilidad del 70% de que el tiempo de aprobación de las asignaturas del primer año de Ingeniería Civil sea mayor a 2 años. Este resultado pone de manifiesto una conexión preocupante entre el acumulado de asignaturas no aprobadas y la prolongación en el tiempo necesario para completar el primer año de la carrera.

Este impacto tiene implicaciones significativas en el retraso en la progresión de la carrera y en la tasa de abandono estudiantil. La alta probabilidad de que el estudiantado en esta situación experimente un retraso sustancial en el logro de los objetivos académicos puede tener efectos negativos en su motivación, compromiso y perspectiva hacia la educación superior. Además, el aumento de la duración del primer año puede influir en las decisiones de continuar o abandonar la carrera, lo que potencialmente contribuye a la tasa de abandono.

Es importante señalar que este análisis se centró en el primer año de la carrera, pero sus implicaciones se pueden extrapolar al cursado del ciclo superior. Futuras líneas de investigación podrían profundizar en el análisis particularizado por año y por asignatura. Examinar cómo el patrón de acumulación de asignaturas no aprobadas se relaciona con la duración de cada año y con las asignaturas específicas podría brindar información más detallada y personalizada para la toma de decisiones educativas.

Los resultados obtenidos subrayan la importancia de abordar el retraso en la aprobación de asignaturas en el contexto de la Carrera de Ingeniería Civil. La relación identificada entre el acumulado de asignaturas no aprobadas y el tiempo de aprobación del primer año plantea cuestiones significativas sobre el diseño de estrategias de apoyo al estudiantado y la optimización de la estructura curricular. Estos resultados sirven como punto de partida para futuras investigaciones que puedan tener un impacto positivo en la retención estudiantil y la mejora de la experiencia académica.

Por otra parte, los resultados obtenidos en este estudio respaldan una preocupación que ha sido ampliamente discutida en la literatura académica en décadas anteriores. Trabajos emblemáticos como el realizado por Tinto (1975) sobre la deserción estudiantil en la educación superior y la investigación de (Bean & Metzner, 1985) sobre la persistencia estudiantil, ya habían identificado la relevancia crítica de los retrasos académicos en los primeros años de la carrera como un precursor significativo del abandono.

Estas contribuciones tempranas sentaron las bases para la comprensión de cómo los obstáculos en la culminación de los cursos en los primeros años pueden tener un impacto adverso en la continuidad de la educación superior. Tal como argumentaron Terenzini y Pascarella (1991), la prolongación del período académico inicial puede erosionar la motivación del estudiantado y presentar desafíos para su compromiso y éxito académico.

Nuestros resultados, que sugieren un aumento substancial en el tiempo de aprobación del primer año para estudiantes con un historial de asignaturas no aprobadas, están en consonancia con estas conclusiones previas. La persistencia de este patrón en el contexto de la Carrera de Ingeniería Civil pone en relieve la importancia continua de abordar esta problemática académica.

La extrapolación de nuestros hallazgos a los años subsiguientes, si bien demanda investigaciones futuras, encuentra respaldo en la perspectiva propuesta por Tinto (1989) acerca de la "teoría de la integración". Según esta teoría, los desafíos iniciales en la integración académica pueden influir en el éxito posterior en la educación superior. En este sentido, nuestras conclusiones proporcionan una base para investigaciones ulteriores que consideren un análisis detallado por año y asignatura, expandiendo así la comprensión de cómo el atraso académico puede mantener una influencia prolongada en la trayectoria estudiantil.

Finalmente, creemos que este estudio contribuye a una corriente investigativa en constante evolución que enfatiza la importancia fundamental de abordar los retrasos en los primeros años de la carrera en relación con el abandono estudiantil. Al edificar sobre investigaciones previas que han establecido los vínculos entre retraso y deserción, este análisis aporta un contexto empírico específico y subraya la relevancia de políticas educativas destinadas a mitigar el retraso y promover la persistencia del estudiantado en el ámbito de la educación superior.

Referencias bibliográficas

- Angrist, J. D., & Imbens, G. W. (1995). Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity. *Journal of the American Statistical Association*, 90(430), 431-442. <https://doi.org/10.1080/01621459.1995.10476535>
- Bean, J. P., & Metzner, B. S. (1985). A Conceptual Model of Nontraditional Undergraduate Student Attrition. *Review of Educational Research*, 55(4), Article 4. <https://doi.org/10.3102/00346543055004485>
- Blöbaum, P., Götz, P., Budhathoki, K., Mastakouri, A. A., & Janzing, D. (2022). DoWhy-GCM: An extension of DoWhy for causal inference in graphical causal models. *arXiv preprint arXiv:2206.06821*.
- Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of economic surveys*, 22(1), 31-72.
- Chen, Y. M., Gong, J. Y., & Shi, W. Z. (2007). A distance-based matching algorithm for multi-scale road networks. *Acta Geodaetica et Cartographica Sinica*, 36(1), 84-90.
- Chernozhukov, V., Cinelli, C., Newey, W., Sharma, A., & Syrgkanis, V. (2022). *Long Story Short: Omitted Variable Bias in Causal Machine Learning*.
- Crump, R. K., Hotz, V. J., Imbens, G. W., & Mitnik, O. A. (2008). Nonparametric tests for treatment effect heterogeneity. *The Review of Economics and Statistics*, 90(3), 389-405.
- Desai, R. J., Rothman, K. J., Bateman, B. T., Hernandez-Diaz, S., & Huybrechts, K. F. (2017). A Propensity-score-based Fine Stratification Approach for Confounding Adjustment When Exposure Is Infrequent. *Epidemiology (Cambridge, Mass.)*, 28(2), 249-257. <https://doi.org/10.1097/EDE.0000000000000595>
- Farnè, M., & Montanari, A. (2022). A bootstrap method to test granger-causality in the frequency domain. *Computational Economics*, 59(3), 935-966.
- González Fiegehen, L. E. (2007). Repitencia y deserción universitaria en América Latina. En *Informe sobre la Educación Superior en America Latina y el Caribe 200-2005. La Metamorfosis de la educación superior*. (pp. 156-170). IESALC.
- Hammerton, G., & Munafò, M. R. (2021). Causal inference with observational data: The need for triangulation of evidence. *Psychological Medicine*, 51(4), 563-578. <https://doi.org/10.1017/S0033291720005127>
- Imbens, G. W., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, 142(2), 615-635. <https://doi.org/10.1016/j.jeconom.2007.05.001>
- Jiang, Z., Chen, S., & Ding, P. (2023). *An instrumental variable method for point processes: Generalised Wald estimation based on deconvolution*.
- Kalainathan, D., Goudet, O., & Dutta, R. (2020). Causal discovery toolbox: Uncovering causal relationships in python. *The Journal of Machine Learning Research*, 21(1), 1406-1410.
- Luo, D., Zou, L., Ai, Q., Chen, Z., Li, C., Yin, D., & Davison, B. D. (2023). *Unconfounded Propensity Estimation for Unbiased Ranking*.

Oloriz, M., Lucchini, M. L., & Ferrero, E. (2007). *Relación entre el Rendimiento Académico de los Ingresantes en Carreras de Ingeniería y el Abandono de los Estudios Universitarios*. VII Coloquio Internacional sobre Gestión Universitaria en América del Sur., Mar del Plata, Argentina.

Paz, H. R. (2022). Hacia un nuevo paradigma en el diseño curricular. *Revista Argentina de Investigación Educativa*, 11(4), 139-160.

Pearl, J. (2009). *Causality*. Cambridge university press.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. En *Biometrika* (Vol. 70, Número 1, pp. 41-55). <https://doi.org/10.1093/biomet/70.1.41>

Sharma, A., & Kiciman, E. (2020). *DoWhy*. An End-to-End Library for Causal Inference. *arXiv preprint arXiv:2011.04216*.

Terenzini, P. T., & Pascarella, E. T. (1991). Twenty years of research on college students: Lessons for future research. *Research in Higher Education*, 32, 83-92.

Tinto, V. (1975). Dropout from Higher Education: A Theoretical Synthesis of Recent Research. *Review of Educational Research*, 45(1), Article 1. <https://doi.org/10.3102/00346543045001089>

Tinto, V. (1989). DEFINIR LA DESERCIÓN: UNA CUESTIÓN DE PERSPECTIVA. *Revista de Educación Superior*, 18(3), Article 3.

Van Rossum, G., & Drake Jr, F. L. (1995). *Python reference manual*. Centrum voor Wiskunde en Informatica Amsterdam.

Fecha de recepción: 26-8-2023

Fecha de aceptación: 2-11-2023