Asignación Óptima de Horas de Tutoría Basada en Riesgo Académico

Cindy Liliana Vargas Duque

Luis Angel Vargas Narvaez

Jesús Ariel González Bonilla

# 1. Introducción

Este informe aborda la **asignación óptima de horas de tutoría** como una herramienta para mitigar el riesgo académico en programas universitarios. La relevancia del problema radica en que las instituciones disponen de recursos limitados — horas adicionales de acompañamiento — que deben distribuirse de manera estratégica para maximizar la permanencia estudiantil y mejorar indicadores de retención. La masificación de la educación superior y la presión por resultados cuantificables exigen mecanismos reproducibles basados en datos. Bajo esta perspectiva se plantea un marco formal que prioriza intervenciones a nivel de programa, área disciplinar y nivel de riesgo, con el objetivo general de **minimizar el riesgo académico agregado** posterior a la intervención sujeto a restricciones de presupuesto y capacidad. Los objetivos específicos incluyen: i) construir métricas de riesgo comparables; ii) parametrizar un modelo lineal interpretable; iii) realizar un análisis de sensibilidad presupuestal; y iv) derivar precios sombra para apoyar la toma de decisiones estratégicas.

# 2. Descripción del escenario

El contexto corresponde a la **Corporación Universitaria del Huila (Corhuila)**, que gestiona diversos programas académicos cuyos estudiantes presentan desempeños heterogéneos en cuatro áreas: *lectura crítica*, *competencias ciudadanas*, *razonamiento cuantitativo* e *inglés*. A partir de bases de datos institucionales se construyen indicadores de riesgo inversamente proporcionales al puntaje obtenido en cada área. La institución dispone de un presupuesto finito de horas de tutoría, y cada área presenta una **capacidad máxima** que puede absorber durante el semestre debido a disponibilidad de docentes u otras restricciones. Adicionalmente, se supone que una hora de tutoría reduce el riesgo en una proporción constante específica del área.

La Figura 1 muestra la distribución de puntajes observada en las cuatro áreas. Áreas con mayor densidad de puntajes bajos evidencian mayor riesgo académico y son candidatas a recibir más horas de tutoría.

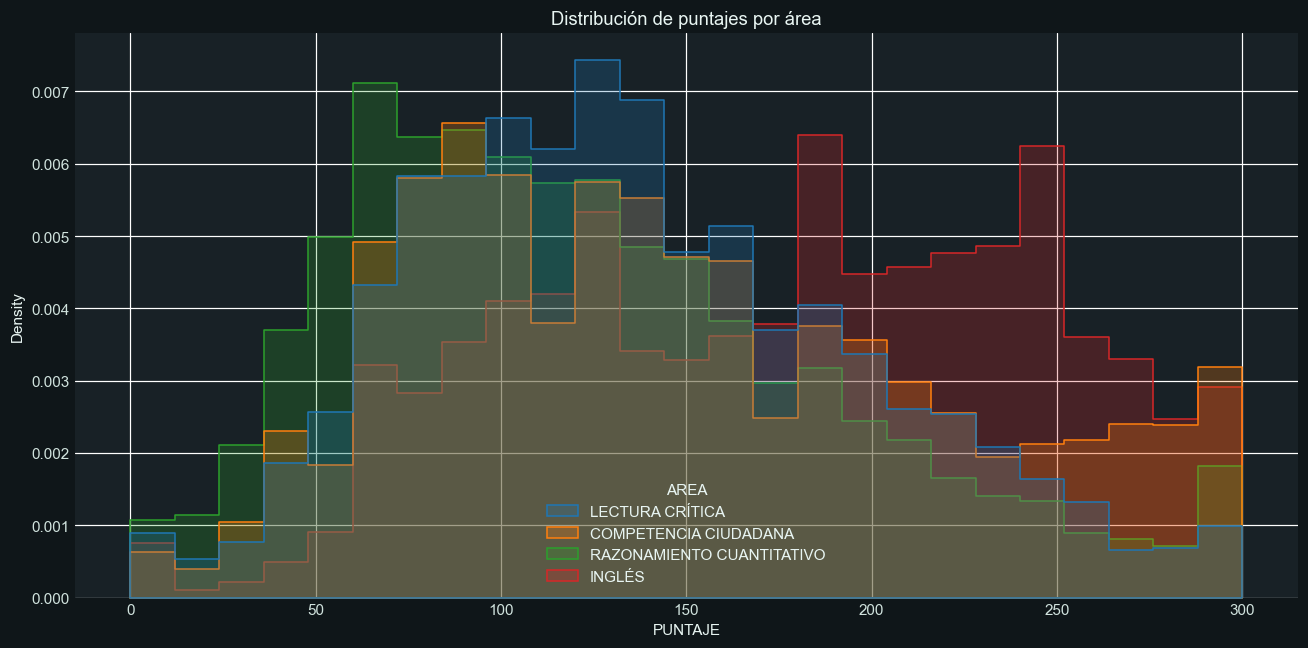


Figura 1. Distribución de puntajes por área

La figura muestra la distribución de puntajes en lectura crítica (azul), competencias ciudadanas (naranja), razonamiento cuantitativo (verde) e inglés (rojo). Las áreas con picos alrededor de 60–100 concentran muchos puntajes bajos.

# 3. Formulación matemática del problema

Para formalizar el problema se define el conjunto de cohortes y el conjunto de áreas . Se introduce la variable de decisión continua que representa las **horas de tutoría** asignadas a la cohorte en el área . Los parámetros incluyen:

* : presupuesto global de horas disponibles.
* : capacidad máxima de horas que el área puede absorber.
* : riesgo base de la cohorte en el área , inversamente proporcional al puntaje.
* : coeficiente de efectividad del área , que mide la reducción del riesgo por hora de tutoría.
* : capacidad máxima de horas que la cohorte puede recibir (opcional).

## 3.1 Función objetivo

Se busca **maximizar la reducción agregada del riesgo** (equivalentemente minimizar el riesgo residual) mediante la siguiente función objetivo:

## 3.2 Restricciones

1. **Presupuesto global**:
2. **Capacidad por área**:
3. **Límite por cohorte (opcional)**:
4. **No negatividad**:

## 3.3 Clasificación del problema

La instancia es **determinista** porque los parámetros se consideran conocidos y fijos; **lineal**, ya que la función objetivo y las restricciones son combinaciones lineales de las variables de decisión; posee **restricciones** de presupuesto y capacidad; y utiliza **variables continuas**. Por lo tanto, se trata de un problema de programación lineal continua que puede resolverse mediante el método simplex.

# 4. Selección del método de resolución

Debido a la estructura lineal del modelo y a la necesidad de satisfacer exactamente las restricciones de presupuesto y capacidad, se selecciona la **programación lineal** como técnica de resolución. El método **simplex** explota la geometría de la región factible y garantiza encontrar la solución óptima en un número finito de iteraciones. Además, su interpretación dual proporciona precios sombra útiles para evaluar el beneficio marginal de relajar cada restricción. Se descarta el gradiente descendente, diseñado para optimizar funciones continuas no necesariamente sujetas a restricciones estrictas, porque no aprovecha la naturaleza lineal ni ofrece información dual valiosa. Para la implementación se utiliza la biblioteca *PuLP* con el solver *HiGHS*, que combina eficiencia y acceso a los multiplicadores duales.

# 5. Implementación en Python

La resolución computacional del modelo se desarrolló en **Python 3.11** combinando varias bibliotecas:

* **pandas** para cargar y preparar los datos de programas, cohortes y puntajes.
* **PuLP** para modelar las variables de decisión y codificar la función objetivo y las restricciones.
* **HiGHS** como solver para resolver el problema y obtener precios sombra.
* **matplotlib** y **seaborn** para generar visualizaciones, incluida la distribución de puntajes y la curva de sensibilidad presupuestal.

Un fragmento representativo del código utilizado para construir y resolver el problema se muestra a continuación:

import pulp  
  
import pulp

# Crear el modelo de maximización

model = pulp.LpProblem('Asignacion\_Horas', pulp.LpMaximize)

# Variables de decisión x[(c,a)] para cada cohorte y área

x = {(c, a): pulp.LpVariable(name=f'x\_{c}\_{a}', lowBound=0)

     for c in cohortes for a in areas}

# Función objetivo: maximizar la reducción de riesgo

model += pulp.lpSum(alpha[a] \* x[(c, a)] for c in cohortes for a in areas)

# Restricciones de capacidad por área

for a in areas:

    model += pulp.lpSum(x[(c, a)] for c in cohortes) <= H\_area[a]

# Restricción de presupuesto global

model += pulp.lpSum(x[(c, a)] for c in cohortes for a in areas) <= H\_total

# Resolver el modelo

model.solve(pulp.PULP\_CBC\_CMD(msg=False))

# Recuperar soluciones

soluciones = {(c, a): x[(c, a)].value() for c in cohortes for a in areas}

# 6. Resultados y análisis

Al resolver el modelo se obtuvieron las horas óptimas por cohorte y área, así como las reducciones de riesgo asociadas. En general, las cohortes con puntajes base más bajos recibieron más horas de tutoría, especialmente en **razonamiento cuantitativo** y **competencias ciudadanas**. La suma del riesgo residual disminuyó de forma notable respecto al escenario sin tutoría, destacándose una mayor reducción en las áreas con mayor pendiente de mitigación. Los **precios sombra** de las restricciones permitieron identificar cuellos de botella: valores positivos en las restricciones de capacidad de razonamiento cuantitativo e inglés indicaron que ampliar su capacidad generaría mejoras importantes.

La Figura 2 presenta la **curva de sensibilidad** al presupuesto de horas. Se observa que la reducción total crece rápidamente hasta alrededor de horas, punto a partir del cual los rendimientos decrecientes se hacen evidentes. Esta información orienta sobre la saturación del recurso y dónde invertir horas adicionales.

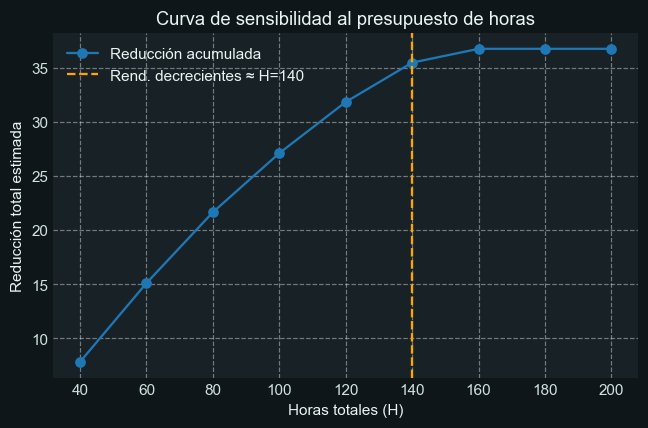


Figura 2. Curva de sensibilidad al presupuesto de horas

La curva muestra la reducción total estimada en función del número de horas disponibles. A medida que se incrementa el presupuesto, la reducción aumenta hasta alcanzar un umbral () a partir del cual los rendimientos se vuelven decrecientes.

El análisis de sensibilidad reveló un **punto de rendimientos decrecientes** en torno a horas. Por debajo de este valor, cada hora adicional genera reducciones significativas; por encima, la ganancia marginal disminuye. Los multiplicadores duales asociados al presupuesto global y a las capacidades de áreas actuaron como precios sombra, sugiriendo que la institución debería priorizar inversiones en aquellas restricciones con mayor valor dual.

# 7. Conclusiones

* El modelo de programación lineal propuesto proporciona un marco transparente y eficiente para asignar horas de tutoría en función del riesgo académico, garantizando trazabilidad en las decisiones y facilidad de interpretación para los actores institucionales.
* Su aplicación en Corhuila demostró que la distribución óptima focaliza recursos en las cohortes y áreas de mayor necesidad, y que los precios sombra derivados del modelo ofrecen señales accionables para decidir ampliaciones de capacidades o ajustes presupuestales.
* Las principales limitaciones se concentran en la suposición de linealidad entre horas e impacto y en la estimación manual de los coeficientes de efectividad, factores que pueden introducir sesgos y reducir la validez externa si no se calibran con evidencia.
* Futuras líneas de trabajo incluyen calibrar empíricamente el parámetro α mediante experimentos controlados, incorporar variables enteras y costos diferenciados a través de programación lineal mixta, extender el enfoque a formulaciones estocásticas o robustas que contemplen la incertidumbre, e integrar el modelo con sistemas de analítica institucional y aprendizaje automático para cerrar el ciclo de retroalimentación y fortalecer la toma de decisiones a largo plazo.

# Referencias

PuLP Community (2023). *PuLP: Linear Programming in Python*.

Python Software Foundation (2024). *Python Language Reference*.

Hunter, J. D. (2007). *Matplotlib: A 2D graphics environment*. Computing in Science & Engineering.