

# Estimación de la tasa de deserción escolar a nivel preparatoria

## Un caso particular de la Universidad Autónoma Chapingo

**Mónica García   Mariana Brizuela   Joel Jaramillo   Miguel Ángel Reyes**

Modelos Lineales Generalizados  
Dr. Luis Enrique Nieto Barajas

Ciudad de México, Mayo 2022



# Motivación

- Cada año la Universidad Autónoma Chapingo (UACH) emite una convocatoria para su **concurso de selección a lo largo del país para estudiantes que terminaron la secundaria o el bachillerato**.
- El ingreso no solamente está determinado por el puntaje en el examen de admisión sino que da **prioridad a estudiantes del medio rural y de recursos económicos escasos** (etnias indígenas, telesecundarias, y habitantes de los municipios más pobres).
- Adicional, **la UACH otorga exención total de pagos por cualquier tipo de servicios** a los estudiantes en condiciones socioeconómicas desfavorables y considera la asignación de servicios asistenciales.

# Deserción Escolar

## Matriz de análisis de factores asociados al fracaso escolar

Dimensión	Factores Exógenos	Factores Endógenos
Material/Estructural	<ul style="list-style-type: none"><li>• Nivel Socioeconómico de la familia</li><li>• Escolaridad de los padres y de adultos del hogar</li><li>• Composición familiar</li><li>• Características de la vivienda</li><li>• Grado de vulnerabilidad social (desempleo, consumo de drogas, delincuencia, etc.)</li><li>• Origen étnico</li><li>• Situación nutricional de los niños</li><li>• Trabajo infantil y de los adolescentes</li></ul>	<ul style="list-style-type: none"><li>• Equipamiento – infraestructura escolar</li><li>• Planta docente</li><li>• Material educativo</li><li>• Programas de alimentación y salud escolar</li><li>• Becas</li></ul>
Cultural	<ul style="list-style-type: none"><li>• Actitud, valoración hacia la educación</li><li>• Pautas de crianza y socialización</li><li>• Consumos culturales</li><li>• Pautas lingüísticas y de comunicación al núcleo familiar</li><li>• Expectativas y Aspiraciones</li><li>• Capital Cultural de las familias</li><li>• Uso del tiempo de los niños y de los jóvenes</li></ul>	<ul style="list-style-type: none"><li>• Capital cultural de los docentes</li><li>• Estilo y prácticas pedagógicas</li><li>• Valoración de expectativas de los docentes y directivos respecto de los alumnos</li><li>• Clima y ambiente escolar</li><li>• Liderazgo y conducción</li></ul>

Fuente: Román C., Marcela. (2013)

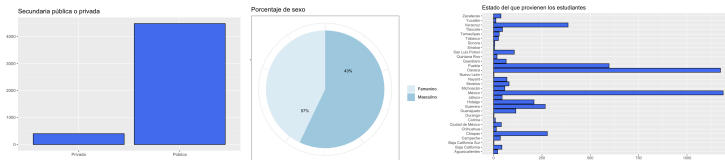
# Objetivo

- Determinar **perfiles de estudiantes con mayor tasa de abandono escolar a nivel preparatoria** a fin de que la universidad destine esfuerzos focalizados para beneficiar a más estudiantes que lo requieren.

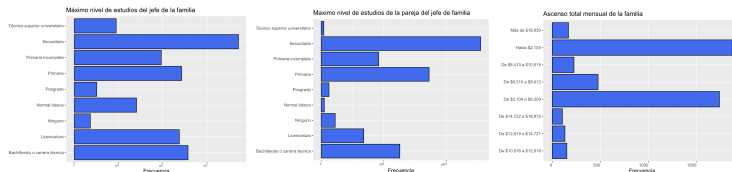
# Análisis Exploración de los Datos - EDA I

Se cuenta con una base con 4,981 registros de estudiantes de preparatoria UACH y 86 variables las cuales están agrupadas en cuatro rubros:

- **Información general del estudiante**

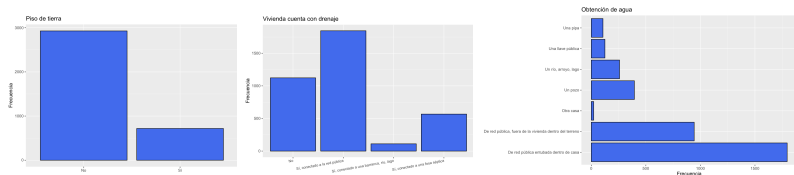


- **Información de la familia**

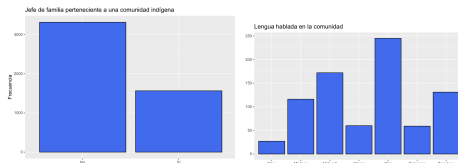


# Análisis Exploración de los Datos - EDA II

## ● Información de la vivienda



## ● Información de la localidad



Se cuenta con la variable categórica **estatus** la cual indica si el alumno sigue vigente o no, así como el registro de sus calificaciones de las materias cursadas.

# Modelos y Resultados I

Dado el número de estudiantes dados de baja  $Y_i$  y el número total de estudiantes  $n_i$ , se desea modelar la tasa de deserción  $\theta_i$  en la clase  $i$ . Para ello proponemos 3 modelos:

- 1) **Modelo de efectos intercambiables** con información del número de estudiantes dados de baja clasificados **por el nivel de ingreso del jefe de familia**.

$$Y_i|\theta_i \sim Po(\theta_i n_i); \theta_i|\alpha, \beta \sim Ga(\alpha, \beta); \alpha, \beta \sim Ga(0.01, 0.01)$$

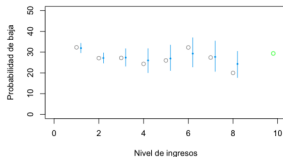
Se corrieron 2 cadenas de Markov con 100 mil iteraciones y un periodo de calentamiento del 10% de las iteraciones.

Parámetro	Media	D. Estándar	2.50%	97.50%
teta 1	31.82027	1.141855	29.67164	34.09736
teta 2	27.19054	1.217528	24.82005	29.59939
teta 3	27.35148	2.194076	23.07083	31.70017
teta 4	26.12628	2.761828	20.71512	31.51186
teta 5	27.13583	3.129808	21.21773	33.59784
teta 6	29.30938	3.61413	22.87393	36.95178
teta 7	27.64807	3.45964	21.03416	34.73572
teta 8	24.59632	3.198173	18.10653	30.41895

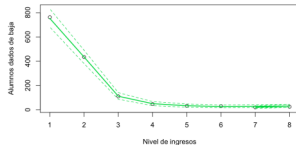
Figure: Estimaciones posteriores de los parámetros

## Modelos y Resultados II

Con este modelo obtenemos una tasa global de deserción del 29.3%, y vemos que un estudiante con un jefe de familia con nivel de ingresos muy bajo es más probable que se deserte de preparatoria, contrario a lo que se observa con estudiantes cuyo jefe de familia percibe altos ingresos.



(a) Tasas estimadas



(b) Bajas estimadas

Con este modelo se hace un buen ajuste a los datos, los intervalos de credibilidad son ligeramente angostos y las predicciones del número de baja de los estudiantes " pasan" por los datos observados.

**DIC: 60.58**



# Modelos y Resultados III

- 2) **Modelo de efectos intercambiables** con información del número de estudiantes dados de baja clasificados **por el sexo, tipo de secundaria, nivel de ingreso del jefe de familia, habla lengua indígena, piso de tierra de la casa, máximo nivel de estudio del jefe de familia e índice de vivienda.**

$$Y_i \sim \text{Poisson}(\mu_i)$$
$$\mu_i \sim \theta_i n_i$$

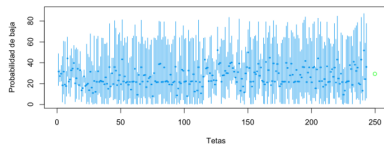
Con las siguientes distribuciones previas no informativas:

$$\theta \sim \mathcal{N}(0, 0.001)$$

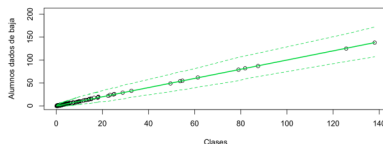
Más Altas		Más bajas	
teta 126 - (mujer, secundaria pública, 1er nivel de ingresos, no habla lengua indígena, no tiene piso de tierra, 1 o 2 carencias en vivienda, nivel de estudios jefe = 1)	52.07179	teta 75 - (hombre, secundaria pública, 8vo nivel de ingresos, no habla lengua indígena, no tiene piso de tierra, 1 o 2 carencias en vivienda, nivel de estudios jefe =1)	7.939968
teta 241 - (mujer, secundaria privada, 8vo nivel de ingresos, no habla lengua indígena, no tiene piso de tierra, 1 o 2 carencias en vivienda, nivel de estudios jefe = 2 )	51.24751	teta 102 - (hombre, secundaria privada, 5to nivel de ingresos, no habla lengua indígena, no tiene piso de tierra, 1 o 2 carencias en vivienda, nivel de estudios jefe = 1)	8.126562
teta 150 - (mujer, secundaria pública, 2do nivel de ingresos, sí habla lengua indígena, no tiene piso de tierra, 1 o 2 carencias en vivienda, nivel de estudios jefe = 3 )	49.97181	teta 94 - (hombre, secundaria privada, 3er nivel de ingresos, no habla lengua indígena, no tiene piso de tierra, 1 o 2 carencias en vivienda, nivel de estudios jefe = 1)	8.554836

Figure: Clases más/menos probables

# Modelos y Resultados IV



(a) Tasas estimadas



(b) Bajas estimadas

**DIC: 546**

- 3) **Modelo de regresión** Hasta ahora hemos modelado el número de éxitos con una distribución *Poisson*, la cual utilizamos para conteos no acotados, dado que tiene un soporte positivo usando  $\lambda$  como tasa. Para hacerlo interpretable, podemos pasarlo a los reales y modelarlo con un predictor lineal. Para ello ocupamos una liga logística. A continuación el modelo: 1.0

$$\begin{aligned} Y_i &\sim \text{Poisson}(\mu_i) \\ \mu_i &\sim \lambda_i n_i \\ \log(\lambda_i) &= \alpha + \beta x_1 + \delta x_2 + \gamma x_3 \end{aligned}$$

# Modelos y Resultados V

Al igual que el modelo 1, se corrieron 2 cadenas de Markov con 100 mil iteraciones y un periodo de calentamiento de 10 mil de las iteraciones.

Parámetro	Media	D. Estándar	2.50%	97.50%
$\alpha$	-1.27	0.05632	-1.27	-1.162
$\beta_1$	-0.1368	0.02718	-0.1899	-0.08439
$\beta_2$	0.1368	0.02718	0.08439	0.1899
$\delta_1$	0.09744	0.05354	0.05988	0.2035
$\delta_2$	-0.09744	0.05354	-0.2035	-0.05988
$\gamma_1$	0.1473	0.02875	0.9134	0.2043
$\gamma_2$	-0.1473	0.02875	0.2043	-0.09134

Figure: Estimaciones posteriores de los coeficientes

Podemos apreciar que todos los coeficientes son significativos. La tasa global de deserción es del 28% ( $e^\alpha = e^{-1.27} = 0.28$ ), muy parecida al modelo 1!

## Modelos y Resultados VI

$$\frac{\lambda_i}{\lambda_j} = \frac{e^{\alpha+\beta_1}}{e^{\alpha}} = e^{(-0.1368)} = 0.87$$

$$\frac{\lambda_i}{\lambda_j} = \frac{e^{\alpha+\beta_1}}{e^{\alpha+\beta_2}} = e^{\beta_1-\beta_2} = e^{-0.1368-(0.1368))} = 0.76$$

Figure: Tasas de deserción por sexo

$$\frac{\lambda_i}{\lambda_j} = \frac{e^{\alpha+\delta_1}}{e^{\alpha}} = e^{(0.09744)} = 1.10$$

$$\frac{\lambda_i}{\lambda_j} = \frac{e^{\alpha+\delta_1}}{e^{\alpha+\delta_2}} = e^{\delta_1-\delta_2} = e^{0.09744-(-0.09744))} = 1.21$$

Figure: Tasas de deserción por tipo de secundaria

## Modelos y Resultados VII

$$\frac{\lambda_i}{\lambda_j} = \frac{e^{\alpha+\gamma_1}}{e^{\alpha}} = e^{(0.1473)} = 1.1587$$

$$\frac{\lambda_i}{\lambda_j} = \frac{e^{\alpha+\gamma_1}}{e^{\alpha+\gamma_2}} = e^{\gamma_1-\gamma_2} = e^{0.1473-(-0.1473)} = 1.34$$

Figure: Tasas de deserción por lengua indígena

# Conclusiones I

- Hay una constante en las estimaciones, el ingreso y las características sociodemográficas juegan un papel importante en el desempeño de los alumnos.
- Además de haber probado la hipótesis que nos planteamos al inicio, también encontramos que existe una relación entre el sexo del alumno, puesto que en el modelo 2 las probabilidades de deserción son más altas para las mujeres que para los hombres.
- En otros modelos que evaluamos, la variable Estados no proporcionó información que contribuya a la determinación del abandono escolar.
- "Trade-off" entre eficiencia e inclusión

# ¡Gracias!