

Pronóstico de la deserción para un programa educativo de ingeniería industrial mediante el uso de MCO, en un campus universitario del centro del país

Drop-out forecast for industrial engineer educative program by OLS, in a university campus at central Mexico

Alicia Sánchez-González *, Héctor D. Molina-Ruíz

Abstract:

The ordinary least squares (OLS) method is a relatively simple and useful statistical tool, which enables decision makers to forecast beyond the data frontier, through a linear adjustment. This method can be applied in different contexts and organizations. Such is the case of higher education institutions, which seek to ensure educational quality, through different strategies and tools. In this document, it is stated statistical contextualization of drop-out's variable and a forecast is made for the following data treated period, with the aim of generating a reference for decision making.

Keywords:

Industrial engineer, Ordinary least squares, Student dropout

Resumen:

El método de mínimos cuadrados ordinarios (MCO), es una herramienta estadística relativamente sencilla y útil, que posibilita a los tomadores de decisiones, efectuar pronósticos más allá de la frontera de los datos, a través de un ajuste lineal. Dicho método, puede ser aplicado en diversos contextos y organizaciones. Tal es el caso de las instituciones de educación superior, que buscan asegurar la calidad educativa, por medio de diferentes estrategias y herramientas. En el presente documento, se plantea la contextualización estadística de la variable "deserción" y se realiza el pronóstico para el periodo consecuente a los datos tratados, con el objetivo de generar un referente para la toma de decisiones.

Palabras Clave:

Deserción estudiantil, Ingeniero industrial, Mínimos cuadrados ordinarios

Introducción

La educación superior se constituye en madre nutricia del conocimiento humano, aportando al individuo los elementos esenciales para su crecimiento y desarrollo personal, intelectual, familiar y social, solo que en ocasiones este accionar se ve frustrado por el fenómeno de la deserción académica, devenido en problema global que afecta la permanencia del claustro estudiantil en las instituciones de educación superior (Romero-Fernández, 2019). El tema de la deserción en México ha sido abordado desde varias décadas atrás, se han realizado diversos estudios y varias propuestas para disminuir el índice de deserción que cuenta el país. Lamentablemente, esta problemática se sigue presentando en los diferentes estados de la república mexicana.

Carballo-Aguilar, Ramírez-León & Álvares-Cisneros (2019), postulan que, la deserción estudiantil en cualquiera de los niveles del sistema educativo, se reconoce como el abandono del proceso educativo formal antes de haber culminado el ciclo educativo al cual aspiraba, para efectos de este estudio. La deserción universitaria es un tema preocupante en diferentes casas de estudio, ya que cada una busca que estos porcentajes disminuyan o que simplemente no exista (Flores-Ferro, E., Bahamondes-Acevedo, V.V., Maureira-Cid, F. & González-Flores, P., 2019).

Para Albarrán-Peña (2019), la deserción, se ve sujeta al surgimiento de dificultades socioeconómicas que provocan en el joven universitario una desmotivación y apatía hacia los estudios de pregrado, pues sienten que la culminación de la carrera no corresponderá con la obtención en un corto plazo de una determinada

^a Autor de Correspondencia, Universidad Autónoma del Estado de Hidalgo, Instituto de Ciencias Básicas e Ingeniería, Av. del Maestro, 41, Tepeji del Río de Ocampo, Hidalgo, México., alicia_sanchez9581@uaeh.edu.mx

^b Autor, Universidad Autónoma del Estado de Hidalgo, Instituto de Ciencias Básicas e Ingeniería, Av. del Maestro, 41, Tepeji del Río de Ocampo, Hidalgo, México.

oportunidad laboral. Como consecuencia de la deserción, los estudiantes, universidades y gobiernos, se ven afectados en términos económicos y sociales institución (Alban & Mauricio, 2019).

La universidad objeto de estudio es una institución pública de educación a nivel media superior, nivel superior y posgrado, se encuentra en la zona centro del país, y al sur de su entidad federativa. la oferta educativa se centra en: biología y química, medicina y ciencias de la salud, humanidades y ciencias de la conducta, ciencias sociales y economía, biotecnología y ciencias agropecuarias, ciencias de la ingeniería y multidisciplinaria. es la institución de enseñanza más antigua de su estado.

Un gran número de estudiantes de ingeniería, abandonar la carrera después del primer año (Pal, 2012). En este documento, se realiza un análisis estadístico del nivel de deserción para un programa educativo de ingeniería industrial, para que, como siguiente paso, se realice la búsqueda de las circunstancias que conllevan a los estudiantes a desertan durante su estancia en la universidad.

Desarrollo

Para el presente estudio se realizó la búsqueda de datos en los históricos del área de planeación, para la universidad objeto de estudio, previo sustento (teórico) de la temática de deserción escolar. Una vez obtenidos los datos, de efectuó la caracterización básica de los datos, cálculo de las medidas de tendencia central, medidas de dispersión, coeficiente de correlación y regresión lineal (MCO), con el objetivo de caracterizar los datos, y aportar los resultados a los tomadores de decisiones, para proveerles de un referente que les permita ejercer acciones con miras a la reducción del índice de deserción, específicamente para el programa educativo de ingeniería industrial, impartido en la institución objeto de estudio.

El presente, es un estudio no experimental, existiendo observación del fenómeno tal y como ocurre, en la institución objeto de estudio, por parte de los investigadores, sin intervención directa en el desarrollo de los hechos (deserciones). Además, se considera un estudio cualitativo, longitudinal de evolución de cohortes, dado que se utilizan los históricos con los cuales cuenta la institución al para las diferentes cohortes, con base en dichos datos, se efectúa el desarrollo de los cálculos de medidas de tendencia central, dispersión, coeficiente de correlación y regresión lineal.

Resultados

La institución objeto de estudio y, particularmente, el programa académico bajo análisis, cuenta con un seguimiento a nivel central, por parte de la dirección general de planeación, con lo cual, se tienen datos de las

diferentes cohortes, al respecto de los ingresos, egresos, deserciones, entre otros datos históricos.

A continuación, se muestran los datos obtenidos de la deserción de estudiantes en el campus bajo análisis, para el programa objeto de estudio, a partir de su fecha de creación (Tabla 1).

Tabla 1: Estadísticas Institucionales de la universidad objeto de estudio. Fuente: Elaboración propia con base en UAEH (2019)

Cohorte	ID	Ingresos	Egresos	Desertados
JULIO-DICIEMBRE 2002	0	41	8	33
ENERO - JUNIO 2003	1	18	7	11
JULIO-DICIEMBRE 2003	2	44	16	28
ENERO - JUNIO 2004	3	16	3	13
JULIO-DICIEMBRE 2004	4	31	12	19
ENERO - JUNIO 2005	5	7	3	4
JULIO-DICIEMBRE 2005	6	16	12	4
JULIO-DICIEMBRE 2006	7	25	8	17
ENERO - JUNIO 2007	8	4	2	2
JULIO-DICIEMBRE 2007	9	29	16	13
ENERO - JUNIO 2008	10	15	6	9
JULIO-DICIEMBRE 2008	11	41	21	20
ENERO - JUNIO 2009	12	14	4	10
JULIO-DICIEMBRE 2009	13	30	14	16
ENERO - JUNIO 2010	14	15	7	8
JULIO-DICIEMBRE 2010	15	24	11	13
ENERO - JUNIO 2011	16	15	5	10
JULIO-DICIEMBRE 2011	17	32	24	8
ENERO - JUNIO 2012	18	24	9	15
JULIO-DICIEMBRE 2012	19	39	9	30
ENERO - JUNIO 2013	20	9	5	4
JULIO-DICIEMBRE 2013	21	35	11	24
ENERO - JUNIO 2014	22	15	6	9
JULIO-DICIEMBRE 2014	23	32	11	21

Es preciso mencionar, que la generación enero-junio 2015, aún no alcanza el egreso, por ello, se excluyen los datos de dicha generación en adelante, para el presente análisis.

3.1 Caracterización básica de datos

A continuación, se muestra caracterización básica de los datos de correspondientes al ingreso por cohorte generacional para las clases determinadas por el criterio del investigador (Tabla 2).

Tabla 2: Caracterización de frecuencias, tallo-hoja con límites inferiores y superiores para las diferentes clases, correspondientes

al ingreso de estudiantes. Elaboración propia con base en UAEH (2019)

# Clase	Límite Inferior	Límite Superior	Hoja	Frecuencia absoluta	Frecuencia acumulada absoluta	Frecuencia relativa	Frecuencia relativa acumulada
1	4.00	11.0	III	3	3	0.125	0.125
2	11.0	18.0	HHH II	7	10	0.292	0.417
3	18.0	25.0	III	3	13	0.125	0.542
4	25.0	32.0	IIII	4	17	0.167	0.708
5	32.0	39.0	III	3	20	0.125	0.833
6	39.0	46.0	IIII	4	24	0.167	1.000
				Σ=24		Σ=1.000	

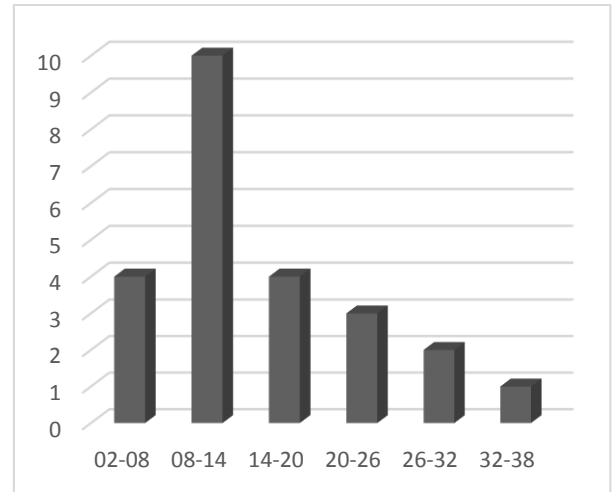


Figura 1: Histograma de deserción por cohorte generacional. Elaboración propia con base en UAEH (2019)

En la siguiente tabla (Tabla 3), se muestra caracterización básica de los datos de correspondientes a la deserción por cohorte generacional para las clases determinadas por el criterio del investigador y que, para el presente estudio, funge como variable de interés.

Tabla 3: Caracterización de frecuencias, tallo-hoja con límites inferiores y superiores para las diferentes clases, correspondientes a la deserción de estudiantes. Elaboración propia con base en UAEH (2019)

# Clase	Límite Inferior	Límite Superior	Hoja	Frecuencia absoluta	Frecuencia acumulada absoluta	Frecuencia relativa	Frecuencia relativa acumulada
1	2	8		4	4	0.16667	0.167
2	8	14	HHH HHH	10	14	0.41667	0.583
3	14	20		4	18	0.16667	0.750
4	20	26		3	21	0.125	0.875
5	26	32		2	23	0.08333	0.958
6	32	38		1	24	0.04167	1.000
				Σ=24		Σ=1.000	

En la figura (Figura 1), que se muestra a continuación, se presenta el histograma correspondiente a la caracterización de los datos de deserción por cohorte generacional.

3.2 Medidas de tendencia central

Para el caso de estudio, se definen las variables aleatorias X (deserción de estudiantes por cohorte generacional) y Y (Ingreso de estudiantes por cohorte generacional), como sigue:

Para el caso de la institución objeto de estudio tiene un estimador para la media igual a:

$$\bar{x} = 14.20833333 \text{ estudiantes que han desertado}$$

Para lo cual se tiene un valor esperado de 14.20833333 estudiantes desertados, por cohorte generacional.

Al realizar el ordenamiento de los datos para la variable X, se tiene que $x_{\frac{n}{2}}$ es igual a 13 y $x_{\frac{n}{2}+1}$ es igual a 13, por lo tanto, la mediana es:

Aunado a esto, con la información de variable X se puede ver que, la moda está dada por:

$$\text{moda} = \check{x} = \{x_4, x_{10}, x_{16}\} = 13 \text{ estudiantes desertados}$$

Dados los resultados anteriores, dos de las medias de tendencia central consideradas, indican que la propensión tiende a 13 estudiantes desertados, por cohorte generacional, por otro lado, la media, indica que se esperan 14.20833333 estudiantes desertados, por cohorte generacional. De forma general, se puede ver que las medidas de tendencia central, en efecto tienden al valor esperado para el estimador de la media.

3.3 Medidas de dispersión

Para el caso de la institución objeto de estudio, y con base en los registros para la variable X, tiene un estimador para la desviación estándar:

$$S_x = 8.459515695 \text{ estudiantes desertados}$$

En cuanto a la varianza, se tiene el cálculo para el estadístico, como sigue:

$$S_x^2 = 71.5634058$$

El rango de los datos vinculados a los estudiantes desertados es de (ecuación 10):

$$\text{rango} = 33 - 2 = 31 \text{ estudiantes desertados}$$

Para los datos obtenidos de la observación del proceso se tiene que el coeficiente de asimetría es de:

$$\text{asimetría} = 0.691184806$$

La curtosis de los datos vinculados a las acciones vencidas es:

$$\text{curtosis} = -0.431714406$$

De acuerdo con los cálculos anteriores de tiene un rango de datos igual a 31 estudiantes desertados. Por otro lado, la desviación estándar es igual a 8.459515695 estudiantes desertados, lo cual implica que el 65% de los datos se encuentra a una desviación estándar (hacia la izquierda y hacia la derecha) de la media, el 95% de los datos se distribuyen a dos desviaciones estándar y 99% de los datos se encuentran a tres desviaciones estándar.

El coeficiente de asimetría (*asimetría* = 0.691184806), indica que existe asimetría positiva, por lo cual los datos tienden a agruparse a la izquierda de la media. En el caso del coeficiente de curtosis (*curtosis* = -0.431714406), la concentración de datos se distribuye de platocúrtica, por lo cual muestra una forma achatada.

3.4 Índice de correlación

Para el caso de la institución objeto de estudio, se utilizan como variables para el cálculo del índice de correlación, los estudiantes que han desertado y los estudiantes inscritos, por cohorte generacional. Como ya se hizo mención la el promedio para la variable X es 14.208 ($\bar{x} = 14.20833333$ estudiantes que han desertado de nivel universitario), en tanto que para la variable Y se tiene:

$$\bar{y} = 23.79166667 \text{ estudiantes de nuevo ingreso}$$

La desviación estándar para la variable aleatoria X esta dada por 16.599 ($S_x = 16.59933369$ acciones vencidas), mientras que la desviación estándar para la variable Y es:

$$S_y = 27.31785408 \text{ estudiantes de nuevo ingreso}$$

La covarianza entre las variables X y Y , esta dada por:

$$S_{XY} = \text{cov}(X, Y) = 379.0763889$$

Por lo tanto, el índice de correlación se expresa por:

$$\rho_{XY} = 0.851368777$$

Lo cual implica que existe un 83.13% de relación positiva entre las variables X y Y , con base en ello se puede expresar que un cambio en la alguna de ellas, afecta directamente a la otra.

Al considerar las desviaciones estandar poblacionales, se tiene:

$$\rho_{XY} = \frac{83.12673611}{(8.28140068)(11.29889068)} = 0.88838481$$

Que de hecho, representa el resultado que arroja la formula COEF.DE.CORREL(matriz1;matriz2), de Microsoft® Excel® y la función de correlación (menú "estadística", submenú "estadísticas básicas", función "correlación") del programa Minitab®.

Lo cual indica una correlación de 88.83%. Sin embargo, dado que los datos históricos usados en el presente estudio, corresponden a la muestra para diferentes cohortes, se recomienda el uso de las ecuaciones (8) y (14) para cálculo de desviaciones estándar, aunadas a la ecuación (15) para el cálculo de la correlación estadística entre las variables y el consecuente cálculo del índice de correlación (ecuación 16)

3.5 Regresión lineal (MCO)

La regresión lineal, también conocida como mínimos cuadrados ordinarios, es una herramienta útil para el cálculo del pronóstico del periodo $t + 1$, con base en la información histórica. Cabe hacer mención que, para esta sección, las variables serán tratadas de una forma que se adecue a la necesidad del pronóstico. Para nuestro caso, se define como (refiérase a la Tabla 1):

$$X = \{\text{número de estudiantes que ingresan, por cohorte generacional}\} \\ = \{41, 18, \dots, 15, 32\}$$

$$Y = \{\text{deserción de estudiantes por cohorte generacional}\} \\ = \{33, 11, \dots, 9, 21\}$$

Usando las diferentes formulaciones matemáticas, a saber: \bar{y} , \bar{x} , S_x^2 , S_{XY} , b, a , \hat{Y} , VE , VNE , VT , y R^2 , se tiene:

$$b = 0.651132114$$

$$a = -1.283184878$$

$$\hat{Y} = -1.28318487794928 + 0.651132113959339 * X$$

Es decir, el pronóstico para el periodo 24 (enero – junio 2015), con $x_i = 14$, es:

$$\begin{aligned}\hat{y}_{t+1} &= \hat{y}_{25} = \hat{y}_{x=14} \\ &= -1.28318487794928 \\ &+ 0.651132113959339 * (24) \\ &= 7.832664717\end{aligned}$$

Lo cual indica que para el periodo $t + 1$, se espera que se presenten alrededor de 8 alumnos desertados.

Las variaciones para el método de mínimos cuadrados son:

$$VE = 1299.035698$$

$$VNE = 346.9226355$$

$$VT = 1645.958333$$

Con base en los cálculos anteriores, el coeficiente de determinación es:

$$R^2 = 1 - \frac{346.9226355}{1645.958333} = 1 - 0.210772429 = 0.789227571$$

o en su caso:

$$R^2 = \frac{1299.035698}{1645.958333} = 0.789227571$$

Que, en el caso de nuestro estudio, el ajuste presenta un 78.92% de calidad de ajuste para la línea de regresión.

Conclusión

Por su parte Flores-Ferro *et al.* (2019), puntualiza varias causas de la deserción escolar, para el nivel universitario. Para Pal (2012), las razones de importancia relativa para la deserción, se pueden agrupar en: razones personales y razones vinculadas al programa (curso). Albarrán-Peña (2019), lista algunas causales para la deserción escolar como lo zona la carencia de dinero para sufragar los gastos de vivienda, alimentación, vestido, calzado, transporte, libros, fotocopiado de textos, adquisición de materiales y equipos para prácticas de laboratorio, entre otros aspectos, genera incertidumbre e incapacidad para cubrir estas necesidades, obligando al joven a desincorporarse de sus estudios. Para Carballo-Aguilar, Ramírez-León & Álvarez-Cisneros (2019), la principal causa de deserción, consiste en el alto índice de reprobación.

El fenómeno de deserción se vuelve crítico cuando los encargados de las entidades universitarias no procesan las herramientas necesarias para identificar a los estudiantes que están en riesgo de dejar la institución (Alban & Mauricio, 2019). Con el análisis de los datos históricos se crea la oportunidad, para que los tomadores de decisiones, generen acciones de mejora, con objeto de reducir, los niveles de deserción en el programa educativos bajo análisis. El siguiente paso, consiste en determinar las causas raíz que enfrentan los estudiantes

para abandonar sus estudios antes de concluir la carrera, en el campus objeto de estudio.

Referencias

- Alban, M., & Mauricio, D. (2019). Predicting University Dropout through Data Mining: A Systematic Literature. *Indian Journal of Science and Technology*, 12, 4, consultado de: [https://www.researchgate.net/profile/Mayra_Taipe3/publication/330981790_Predicting_University_Dropout_through_Data_Mining_A_Systematic_Literature/links/5c5e54a9a6fdccb608b287cb/Predicting-University-Dropout-through-Data-Mining-A-Systematic-Literature.pdf].
- Albarrán-Peña, J. M. (2019). La deserción estudiantil en la Universidad de Los Andes (Venezuela). *Educación y Humanismo*, 21(36), 69-100, consultado de: [http://revistas.unisimon.edu.co/index.php/educacion/article/download/2806/4023].
- Carballo-Aguilar, O.A., Ramírez-León, R.H. & Álvarez-Cisneros, W. (2019). Tipo de baja y motivo con mayor frecuencia en la deserción de la licenciatura en estadística y sistemas de información/Type of withdrawal and reason most frequently in the desertion of the degree in statistics and information systems. *Revista Iberoamericana de Producción Académica y Gestión Educativa*, 6(11), consultado de: [https://www.pag.org.mx/index.php/PAG/article/download/790/1143].
- Flores-Ferro, E., Bahamondes-Acevedo, V.V., Maureira-Cid, F. & González-Flores, P. (2019). Motivos de deserción universitaria de estudiantes de educación física de Chile. *EmásF: revista digital de educación física*, (57), 14-23, consultado de: [https://dialnet.unirioja.es/descarga/articulo/6860150.pdf].
- Pal, S. (2012). Mining educational data to reduce dropout rates of engineering students. *International Journal of Information Engineering and Electronic Business*, 4(2), 1, consultado de: [http://www.mecspress.org/ijieeb/ijieeb-v4-n2/IJIEEB-V4-N2-1.pdf].
- Romero-Fernández, A., Robayo-Andrade, E., Padilla-Buñay, A., Álvarez-Cortés, L. & Pailiacho Yucta., H.R. (2019). Análisis de la deserción estudiantil en carreras de ciencias técnicas de la Universidad Regional Autónoma de los Andes (UNIANDÉS), *Revista Dilemas Contemporáneos: Educación, Política y Valores*, VI(2), consultado de: [https://dilemascontemporaneoseducacionpoliticayvalores.com/_files/200004231-b38aeb48e3/19.01.17%20An%C3%A1lisis%20de%20la%20deserción%20de%20estudiantil%20en%20carreras%20de%20ciencias.....pdf].
- Serratos, F., & Sanfeliu, A. (2006). Signatures versus histograms: Definitions, distances and algorithms. *Pattern recognition*, 39(5), 921-934, consultado de: [https://www.sciencedirect.com/science/article/pii/S0031320305004619].
- Strauss, O. (2009). Quasi-continuous histograms. *Fuzzy Sets and Systems*, 160(17), 2442-2465, consultado de: [http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.154.698&rep=rep1&type=pdf].
- UAEH (2019). Escuela Superior Tepeji, Concentrado de la Eficiencia Terminal, Información para el estudio de Trayectorias Escolares, Dirección General de Planeación, Universidad Autónoma del Estado de Hidalgo, consultado de: [http://intranet.uaeh.edu.mx/DGP/trayectorias.htm].