



Abril 2016 - ISSN: 1989-4155

REGRESIÓN LOGÍSTICA: UNA APLICACIÓN PARA EL ESTUDIO DE LOS ELEMENTOS DETERMINANTES EN LA SELECCIÓN DE UNA UNIVERSIDAD PARA ALCANZAR UN TÍTULO DE TERCER NIVEL

¹Angel Fredy Castelo Rivas. MBA

²Gelmar García Vidal. Phd

³Reyner Pérez Campdesuñer. Phd

⁴Piedad Alarcón Toabanda. MBA

⁵Walter Castelo Rivas. MBA

¹Universidad Tecnológica Equinoccial (UTE), Santo Domingo - Ecuador, angel.castelo2013@gmail.com

²Universidad Tecnológica Equinoccial (UTE), Santo Domingo – Ecuador, gelmargv@yahoo.com

³Universidad Tecnológica Equinoccial (UTE), Santo Domingo – Ecuador, rperezcampdesuner@gmail.com

⁴Pontificia Universidad Católica del Ecuador (PUCE), Santo Domingo - Ecuador, piaeli@gmail.com

⁵Pontificia Universidad Católica del Ecuador (PUCE), Santo Domingo - Ecuador, castelo_walter@yahoo.es

Para citar este artículo puede utilizar el siguiente formato:

Angel Fredy Castelo Rivas, Gelmar García Vidal, Reyner Pérez Campdesuñer, Piedad Alarcón Toabanda y Walter Castelo Rivas (2016): “Regresión logística: una aplicación para el estudio de los elementos determinantes en la selección de una universidad para alcanzar un título de tercer nivel”, Revista Atlante: Cuadernos de Educación y Desarrollo (abril 2016). En línea: <http://www.eumed.net/rev/atlante/2016/04/regresion.html>

RESUMEN

El artículo hace uso de la regresión logística para determinar qué factores tienen mayor influencia en la elección de una universidad para la obtención de un título de tercer nivel. La variable dependiente asume dos opciones, selección de la Universidad Tecnológica Equinoccial y selección de cualquier otra universidad del territorio de Santo Domingo de los Tsáchilas o fuera de él. Fueron utilizadas cinco variables independientes de las cuales, educación superior que prefiere, el horario de estudio y la categoría del colegio de procedencia fueron aquellas que ejercen mayor influencia en la selección de la universidad.

ABSTRACT

The article makes use of logistic regression to determine which factors have the greatest influence in choosing a university to obtain a bachelor's third level. The dependent variable takes two options, selecting Universidad Tecnológica Equinoccial and selection of any other university in the territory of Santo Domingo de los Tsáchilas or out of it. They were used five independent variables of which prefers higher education, the study schedule and status of the school of origin were those that exert greater influence in the selection of the university.

PALABRAS CLAVES: regresión logística binaria, preferencia de universidad.

INTRODUCCIÓN

La regresión logística suele preferirse como método estadístico cuando se cuentan con varias variable independientes y una variable dependiente no continua y no que sigue una distribución normal (es el caso de una variable binaria), debido a que, en comparación a otras pruebas estadísticas de factura semejante, no se enfrenta a supuestos estrictos que no siempre pueden validarse, incluso es más robusta cuando estos supuestos no se cumplen, haciendo muy apropiada su aplicación en situaciones diversas (Díaz González, Covarrubias Melgar, & Sistachs Vega, 2015; Hair, Anderson, Tatham, & Black, 1999; Heredia Rico, Rodríguez Hernández, & Vilalta Alonso, 2012; Mures Quintana, García Gallego, & Vallejo Pascual, 2005; Salas Velasco, 1996). Una de las ventajas de la regresión logística es que solo se necesita saber si un suceso ocurre o no para entonces utilizar un valor dicotómico como variable dependiente. A partir de este valor dicotómico, el procedimiento predice su estimación de la probabilidad de que el suceso tenga o no lugar; si la predicción de la probabilidad es mayor que 0.50, entonces la predicción es si, y en otro caso no (Balluerka, Gorostiaga, Gómez Benito, & Hidalgo, 2010; Hair, et al., 1999; Menéndez Acuña, Gabriel Argüelles, & Hernández González, 2015). La regresión logística deriva su nombre de la transformación logística utilizada con la variable dependiente (Balluerka, et al., 2010; Fernández Aguerre, 2002; Hosmer & Lemeshow, 2000). Dado que la probabilidad de un evento debe estar entre 0 y 1, no es práctico modelar dichas probabilidades con técnicas de regresión lineal ya que, el modelo de regresión lineal permite a la variable dependiente tomar valores mayores que 1 o menores que 0. El modelo de regresión logística es un tipo de modelo lineal generalizado que va más allá del modelo de regresión lineal mediante la vinculación de la gama de números reales a la gama de 0 - 1. Se debe comenzar

teniendo en cuenta la existencia de una variable continua observada, Z, que se puede considerar como la "propensión hacia" el evento de interés.

En el modelo de regresión logística, la relación entre Z y la probabilidad del evento de interés es descrito por esta función de enlace (Alderete, 2006; Kleinbaum, 1994):

$$z_i = \log\left(\frac{\pi_i}{1 - \pi_i}\right)$$

Donde:

π_i : es la probabilidad el caso i-ésimo que experimenta el evento de interés.

z_i : es el valor de la variable continua no observada para el caso i-ésimo.

El modelo asume que Z se relaciona linealmente con sus predictores:

$$z_i = b_0 + b_1x_{i1} + b_2x_{i2} + \dots + b_px_{ip}$$

Donde:

x_{ij} : es el predictor j^{th} para el caso i^{th}

b_j : es el coeficiente j^{th}

p: es el número de predictores

La utilidad de la prueba consiste entonces, en su capacidad para establecer predicciones en las condiciones en que un fenómeno se manifieste o no lo haga condicionado a un grupo de variables planteadas como determinantes (Díaz González, et al., 2015; Hosmer & Lemeshow, 2013; Salas Velasco, 1996). Este método ha sido aplicado en el contexto de la educación para análisis diversos: deserción estudiantil universitaria (Fernández Orrantia & Silva, 2014), estudio del consumo de alcohol en estudiantes universitarios (Restrepo Molina, Agudelo Martínez, Giraldo Torres, & Sánchez Diosa, 2011), estudios de test aplicados en el campo de la educación (Balluerka, et al., 2010), análisis del rendimiento académico (Corengia, Pita, Mesurado, & Centeno, 2013; Chue Gallardo, Barreno Vereau, & Millones Rivalles, 2007; Heredia Rico, et al., 2012; Ibarra & Michalus, 2010), postura corporal en estudiantes universitarios (Agudelo Martínez, 2013), estudio de las relaciones entre el síndrome de burnout y el patrón de conducta tipo A en profesores de educación infantil, primaria y secundaria (Félix Raya, Moriana, & Herruzo, 2010), determinantes sociales e institucionales de la desigualdad educativa (Fernández Aguerre, 2002), asociación entre depresión y bajo rendimiento académico en estudiantes universitarios (Franco Mejía, Gutiérrez Agudelo, & Perea, 2011), detección de estudiantes en riesgo académico (Guillén Sanchez & Chinchilla Brenes, 2005), estimación de la probabilidad de egreso de estudiantes (Nieto Méndez, 2015), relaciones entre ofertas de asignaturas y carreras universitarias (Larrán Jorge &

Andrades Peña, 2015), fracaso escolar (Formichella & Krüger, 2013), dependencia emocional en estudiantes universitarios (Lemos Hoyos, Jaller Jaramillo, González Calle, Díaz León, & de La Ossa, 2012), equidad en el ingreso a la universidad pública (Yusif, Yussof, & Osman, 2013) y estudios de demanda universitaria (Mendoza Requena, Llanos Zavalaga, Rosas Aguirre, Contreras Ríos, & Huayta Zacarías, 2001; Salas Velasco, 1996).

Basado en los estudios antes mencionados el presente trabajo intenta explicar los factores que determinan que los estudiantes del último año de los colegios de Santo Domingo de los Tsáchilas prefieran la Universidad Tecnológica Equinoccial Sede Santo Domingo (UTE) para obtener su título de tercer nivel, todo lo cual permitirá actuar de forma consecuente en aquellos grupos de estudiantes que resulten demanda potencial de dicha universidad. Usando el procedimiento de regresión logística, la UTE puede orientar su labor de orientación vocacional a los estudiantes que tienen más probabilidades de optar por su oferta académica.

MÉTODOS Y MATERIALES

Se utilizó un cuestionario estructurado con ocho preguntas el cual cumple con los criterios de validez aparente y de contenido según criterio de expertos. Se siguió el método de muestreo no probabilístico por cuotas que se basa en seleccionar la muestra después de dividir la población en grupos o estratos que este caso lo constituyeron los colegios. Las cuotas se decidieron a criterio de los investigadores y se eligieron los elementos en cada estrato o grupo por métodos no probabilísticos. Se recopiló información de 556 estudiantes de una población de 14874.

La investigación se realizó en los colegios de Santo Domingo, La Concordia y San Miguel de los Bancos entre los meses de mayo y junio.

Tabla 1: Porcentaje de encuestados por ciudad

	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Santo Domingo	461	82,9	82,9	82,9
San Miguel de los Válidos Bancos	53	9,5	9,5	92,4
La Concordia	42	7,6	7,6	100,0
Total	556	100,0	100,0	

La muestra por colegios y ciudad de ubicación se muestra en la tabla que sigue:

Tabla 2: Colegios estudiados y ciudad de localización

Nombre del Colegio	Ciudad donde se realiza la encuesta			Total
	Santo Domingo	La Concordia	San Miguel de los Bancos	
Villa Florida	32	0	0	32
Víctor Emilio Estrada	15	0	0	15
Santo Domingo	22	0	0	22
San Miguel de los Bancos	0	0	52	52
Providencia	29	0	0	29
Pio XII	38	0	0	38
Nueva Concordia	0	38	0	38
Neumane	2	0	0	2
Mariscal Sucre	33	0	0	33
Madre Laura	26	0	0	26
Latino	22	0	0	22
Kolping	32	0	0	32
Japón	32	0	0	32
Fe y Alegría	48	0	0	48
Dante Alighieri	51	0	0	51
Colegio Adventista del Ecuador	29	0	0	29

Alfredo Pareja	24	0	0	24
Total	435	38	52	525

La base de datos de depuró para realizar el análisis solo con aquellos estudiantes que deseaban continuar estudiando una vez concluidos el bachillerato quedando una total de 525 encuestados para un 94.4% del total como se aprecia en la tabla 3.

Tabla 3: Porcentaje de estudiantes que manifiestan su deseo de seguir con estudios de tercer nivel

	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Estudiar y trabajar	313	56,3	56,3	56,3
Estudiar	212	38,1	38,1	94,4
Válidos Trabajar	29	5,2	5,2	99,6
Otros	2	,4	,4	100,0
Total	556	100,0	100,0	

Con los resultados de la encuesta se pretendió analizar, mediante un modelo de regresión logística, la relación existente entre las características de los estudiantes de colegios encuestados (variables independientes: Naturaleza del colegio del que proviene, Educación superior que prefiere, Modalidad que prefiere para estudiar, Horario para estudiar, Fuente de financiamiento) y la elección de la universidad (variable dependiente dicotómica: Universidad seleccionada en Santo Domingo). El modelo a estimar sería el siguiente:

$$\ln \left[\frac{P(UTE=1)}{P(Otra=0)} \right] = \alpha + \beta_1 \text{Categoría del colegio} + \beta_2 \text{educación preferida} + \beta_3 \text{Modalidad} + \beta_4 \text{Horario} + \beta_5 \text{Fuente de financiamiento}$$

En el modelo propuesto se denomina 1 a la opción de respuesta "Selecciona la UTE" y se define 0 a su alternativa "Cualquier otra universidad en Santo Domingo o fuera de él". Una vez analizados los instrumentos de recopilación de información y tabulados los datos se prepara la prueba estadística utilizando el IBM SPSS 20 (Berlanga Silvente &

Vilà Baños, 2014). La selección de la opción que facilita la corrida de la prueba genera la siguiente sintaxis de comando:

```
LOGISTIC REGRESSION VARIABLES Universidadseleccionada  
  /METHOD=FSSTEP(LR) educaciónpreferida Modalidad Horario  
  Fuentedefinancieamiento Especialidaddelestudiante Categoriadelcolegio  
  /CONTRAST (Especialidaddelestudiante)=Indicator  
  /SAVE=PRED COOK SRESID  
  /CLASSPLOT  
  /PRINT=GOODFIT  
  /CRITERIA=PIN(0.05) POUT(0.10) ITERATE(20) CUT(0.5).
```

- El procedimiento se ajusta a un modelo de regresión logística por defecto.
- Las variables especificadas en el subcomando METHOD se introdujeron en el modelo mediante regresión por pasos hacia adelante sobre la base de la importancia de la relación de probabilidad estadística.
- El subcomando CONTRAST especifica que Especialidaddelestudiante es una variable categórica y debe ser procesada usando indicador de codificación.
- Las subcomando SAVE solicita que las estadísticas de influencia de Cook, y residuos studentizados sean guardados como nuevas variables en el conjunto de datos.
- El subcomando CLASSPLOT solicita una clasificación *plot* de los valores observados y pronosticados.
- El subcomando PRINT solicita la bondad del ajuste estadístico de Hosmer-Lemeshow.
- Todas las demás opciones se establecen en sus valores predeterminados.

Después de construir un modelo, es necesario determinar si este se aproxima razonablemente al comportamiento de los datos procesados. Para esto es necesario analizar la Prueba de ajuste del modelo que el procedimiento de regresión logística binaria informa la bondad de ajuste de Hosmer-Lemeshow. Además es necesario analizar los Plot residuales.

RESULTADOS

Una vez procesados los datos en el paquete estadístico utilizando la opción Regresión Logística Binaria se presentan los hallazgos encontrados. El estadístico de Hosmer-

Lemeshow muestra que el modelo obtenido se ajusta adecuadamente los datos con una significación de 0.218 teniendo en cuenta el criterio de que un valor de significación es inferior a 0,05 indica un mal ajuste. Aquí, La tabla que sigue muestras estadísticas más fiables del ajuste del modelo de regresión logística binaria.

Tabla 4: Prueba de Hosmer y Lemeshow

Paso	Chi cuadrado	gl	Sig.
1	,000	0	.
2	24,302	5	,000
3	50,434	8	,000
4	10,721	8	,218

El método utilizado comienza con un modelo que no incluye ninguno de los predictores utilizados.

Tabla 5: Paso 0, variables que no están en la ecuación

		Puntuación	gl	Sig.
		n		
Paso 0	educaciónpreferida	168,055	2	,000
	educaciónpreferida(1)	139,810	1	,000
	educaciónpreferida(2)	167,974	1	,000
	Modalidad	12,095	1	,001
	Horario	61,385	1	,000
	Fuente de financiamiento	1,251	1	,263
	Categoría del colegio	2,415	1	,120
	Estadísticos globales	205,441	6	,000

En cada paso, el predictor con mayor puntuación estadística cuyo significado valor es menor que un valor especificado (por defecto 0.05) se añade al modelo.

Las variables elegidas siguiendo el método paso a paso deben tener cambios significativos en $-2 \log\text{-likelihood}$.

Tabla 6: Resumen del modelo

Paso	-2 log de la verosimilitud	R cuadrado de Cox y Snell	R cuadrado de Nagelkerke
1	465,697 ^a	,334	,460
2	415,227 ^a	,395	,544
3	405,418 ^a	,406	,560

a. La estimación ha finalizado en el número de iteración 7 porque las estimaciones de los parámetros han cambiado en menos de ,001.

En el modelo de regresión lineal, el coeficiente de determinación, R^2 , resume la proporción de la varianza en la variable dependiente asociada con el predictor de variables (independientes), con valores más grandes R^2 se indica que se explica por el modelo un mayor porcentaje de la variación, a un máximo de 1. Para los modelos de regresión con una variable dependiente categórica, no es posible calcular un único R^2 statistic como en caso el modelo de regresión lineal, por lo que estas aproximaciones se calculan en su lugar. El en caso que se estudia se obtiene un R^2 de Cox y Snell de 0,406 y un R^2 de Nagelkerke a 0,560. Estos resultados, aunque podrían ser superiores, conducen a pensar que la capacidad explicativa del modelo es adecuada (Morrison, 1972; Salas Velasco, 1996).

Tabla 7: Tabla de clasificación ^a

Observado			Pronosticado		
			Universidad que está pensando o le gustaría matricularse en Santo Domingo		Porcentaje correcto
			Otra universidad del territorio o fuera de él	Universidad Tecnológica Equinoccial (UTE)	
Paso 1	Universidad que está pensando o le gustaría matricularse en Santo Domingo	Otra universidad del territorio o fuera de él	204	138	59,6
		Universidad Tecnológica Equinoccial (UTE)	3	180	98,4
	Porcentaje global				73,1
Paso 2	Universidad que está pensando o le gustaría matricularse en Santo Domingo	Otra universidad del territorio o fuera de él	258	84	75,4
		Universidad Tecnológica Equinoccial (UTE)	23	160	87,4
	Porcentaje global				79,6
Paso 3	Universidad que está pensando o le gustaría matricularse en Santo Domingo	Otra universidad del territorio o fuera de él	250	92	73,1
		Universidad Tecnológica Equinoccial (UTE)	22	161	88,0
	Porcentaje global				78,3

a. El valor de corte es ,500

La tabla de clasificación muestra los resultados prácticos de la utilización del modelo de regresión logística. Para cada caso, la respuesta pronosticada es Sí selecciona a la UTE la probabilidad predicha por el modelo en cada caso es mayor que el valor de corte (*cutoff*) especificado en los cuadros de diálogo (en este caso, el valor por defecto de 0,5). De los casos utilizados para crear el modelo, 161 de las 183 son clasificados correctamente para la selección de la Universidad Tecnológica Equinoccial (UTE) y 250 de 342 son clasificados de forma correcta para aquellos que seleccionan otra universidad del territorio o fuera de él. De forma general, el 78,3% de todos los casos fueron clasificados correctamente correspondiéndose con el criterio de que un buen modelo debe identificar correctamente el mayor porcentaje de los casos.

La tabla de estimaciones de los parámetros resume el efecto de cada predictor. Si el nivel de significación de la estadística de Wald es pequeño (inferior a 0,05), entonces el parámetro es útil, para el modelo generado se han seleccionado los predictores *Educación preferida*, el *Horario de preferencia para estudiar* y la *Categoría del colegio de procedencia*.

Tabla 8: Variables en la ecuación

		B	E.T.	Wald	gl	Sig.	Exp(B)
Paso 1 ^a	educaciónpreferida			54,306	2	,000	
	educaciónpreferida(1)	-1,322	1,244	1,129	1	,288	,267
	educaciónpreferida(2)	3,444	1,027	11,247	1	,001	31,304
	Constante	-3,178	1,021	9,696	1	,002	,042
Paso 2 ^b	educaciónpreferida			52,289	2	,000	
	educaciónpreferida(1)	-1,487	1,255	1,405	1	,236	,226
	educaciónpreferida(2)	3,272	1,040	9,909	1	,002	26,377
	Horario	-,824	,134	37,945	1	,000	,438
Paso 3 ^c	Constante	-1,202	1,069	1,263	1	,261	,301
	educaciónpreferida			54,802	2	,000	
	educaciónpreferida(1)	-1,486	1,255	1,402	1	,236	,226
	educaciónpreferida(2)	3,422	1,042	10,790	1	,001	30,637
	Horario	-,818	,132	38,200	1	,000	,442
	Categoríadelcolegio	-,784	,254	9,494	1	,002	,457
	Constante	-,091	1,128	,007	1	,936	,913

a. Variable(s) introducida(s) en el paso 1: educaciónpreferida.

b. Variable(s) introducida(s) en el paso 2: Horario.

c. Variable(s) introducida(s) en el paso 3: Categoríadelcolegio.

Los predictores y valores de los coeficientes mostrados en el paso 3 se utilizarán por el procedimiento para hacer predicciones. El significado de un coeficiente de regresión logística no es tan sencillo como el de un coeficiente de regresión lineal. Mientras que B es adecuado para probar la utilidad de predictores, Exp (B) es más fácil de interpretar. Exp (B) encontrado en la última columna de la tabla representa la relación entre el cambio en la probabilidad del evento de interés, la selección de la UTE como universidad para estudiar, para un cambio de una unidad en el predictor.

Por ejemplo, Exp (B) para la variable Categoría del colegio es igual a ,457, lo que significa que las probabilidades de seleccionar a la UTE para una persona que provenga de un colegio privado es de ,457 veces las probabilidades de seleccionar la UTE que para una persona que provenga de un colegio público, todos en igualdad de condiciones. Las probabilidades de seleccionar a la UTE para una persona que prefiere a una Universidad pública son;

$$P(\text{seleccionar UTE}) = \frac{\text{Odds}(\text{seleccionar UTE})}{1 + \text{Odds}(\text{seleccionar UTE})} = \frac{0,457}{1 + 0,457} = 0,3136$$

por lo que la probabilidad correspondiente de seleccionar la UTE por parte de una persona que prefiera una universidad pública se reduce a 0.3136.

Utilizando el procedimiento de regresión logística, que ha construido un modelo para predecir la probabilidad de que un cliente seleccione a la UTE como universidad para realizar sus estudios de tercer nivel. Este modelo incluye solo tres variables que indiquen en la probabilidad de seleccionar la UTE como universidad para continuar estudios de tercer nivel, el resto de las variables no explicaron esa variable dependiente quedando el modelo como sigue:

$$\text{Ln} \left[\frac{P(UTE=1)}{P(Otra=0)} \right] = \alpha + \beta_1 \text{Categoríadelcolegio} + \beta_2 \text{educaciónpreferida} + \beta_3 \text{Horario}$$

$$\begin{aligned} \text{Ln} \left[\frac{P(UTE = 1)}{P(Otra = 0)} \right] \\ = ,936 + 0,226 \text{educaciónpreferida} + 0,442 \text{Horario} \\ + 0,457 \text{Categoríadelcolegio} \end{aligned}$$

DISCUSIÓN Y CONCLUSIONES

El estudio realizado se apoya en la lógica seguida con los realizados por Lemos Hoyos, et. al. (2012), Yusif, et. al (2013), Mendoza Requena, et. al (2001) y Salas Velasco (1996).

No se encontraron estudios dirigidos específicamente a determinar las variables que influyen en la elección de los estudiantes respecto a la selección de la Universidad Tecnológica Equinoccial como opción para continuar estudios de tercer nivel.

De acuerdo con el modelo obtenido las variables que explican la selección de la Universidad Tecnológica Equinoccial como opción para continuar estudios de tercer nivel son el tipo de educación superior que prefiere, el horario de estudio y la categoría del colegio de procedencia.

De igual forma se determinó que las variables excluidas no poseen suficiente influencia para explicar la selección de la Universidad Tecnológica Equinoccial como opción para continuar estudios de tercer nivel.

RECOMENDACIONES

En función a lo anterior la Universidad Tecnológica Equinoccial debe identificar a los estudiantes de colegios privados que prefieran una educación superior privada en días de semana en cualquiera de sus horarios orientar su actividad para orientar la formación vocacional y promoción de sus servicios.

REFERENCIAS BIBLIOGRÁFICAS

- Agudelo Martínez, A. (2013). Factores asociados a la postura corporal en estudiantes universitarios. *Revista CES Movimiento y Salud*, 1(1), 11-18.
- Alderete, A. M. (2006). Fundamentos del Análisis de Regresión Logística en la Investigación Psicológica. *Evaluar*, 6, 52- 67.
- Balluerka, N., Gorostiaga, A., Gómez Benito, J., & Hidalgo, M. D. (2010). Use of multilevel logistic regression to identify the causes of differential item functioning. *Psicothema*, 22(4), 1018-1025.
- Berlanga Silvente, V., & Vilà Baños, R. (2014). Cómo obtener un Modelo de Regresión Logística Binaria con SPSS. *REIRE, Revista d'Innovació i Recerca en Educació*, 7(2), 105-118.
- Corengia, A., Pita, M., Mesurado, B., & Centeno, A. (2013). Predicting Academic Performance and Attrition in Undergraduate Students. *Liberabit*, 19(1), 101-112.
- Chue Gallardo, J., Barreno Vereau, E., & Millones Rivalles, R. (2007). Sistema para el análisis estadístico de Técnicas multivariadas del rendimiento académico de los estudiantes de una institución de enseñanza superior. *INTERFASES. Revista digital de la Facultad de Ingeniería de Sistemas*(2), 51-82.
- Díaz González, L., Covarrubias Melgar, D., & Sistachs Vega, V. (2015). Enfoque bayesiano del modelo de regresión logística usando Cadenas de Markov Monte Carlo. *Revista Investigación Operacional*, 36(2), 178-185.
- Félix Raya, A., Moriana, J. A., & Herruzo, J. (2010). Relación entre el síndrome de burnout y el patrón de conducta tipo a en profesores. *Ansiedad y Estrés*, 16(1), 61-70.
- Fernández Aguerre, T. (2002). Determinantes sociales e institucionales de la desigualdad educativa en sexto año de educación primaria de Argentina y

- Uruguay, 1999.: Una aproximación mediante un modelo de regresión logística. *Revista Mexicana de Investigación Educativa*, 7(16), 501-536.
- Fernández Orrantia, X., & Silva, E. (2014). Deserción estudiantil universitaria en el primer semestre. El caso de una institución de educación superior ecuatoriana. *Cuadernos del Contrato Social por la Educación*(10), 34-48.
- Formichella, M. M., & Krüger, N. (2013). El fracaso escolar en el nivel medio argentino: ¿es menos frecuente en las escuelas de gestión privada debido a su administración? *Regional and Sectoral Economic Studies*, 13(3), 127-144.
- Franco Mejía, C., Gutiérrez Agudelo, S., & Perea, E. (2011). Asociación entre depresión y bajo rendimiento académico en estudiantes universitarios. *Psicogente*, 14(25), 67-75.
- Guillén Sanchez, E., & Chinchilla Brenes, S. (2005). Detección de estudiantes en riesgo académico en el Instituto Tecnológico de Costa Rica. *Educación*, 29(2), 123-138.
- Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1999). *Análisis multivariante* (Quinta ed.). Madrid: Prentice Hall Iberia, SRL.
- Heredia Rico, J. J., Rodríguez Hernández, A. G., & Vilalta Alonso, J. A. (2012). Empleo de la regresión logística ordinal para la predicción del rendimiento académico. *Revista Investigación Operacional*, 33(3), 252-267.
- Hosmer, D. W., & Lemeshow, S. (2000). *Applied Logistic Regression* (2nd ed.). New York: John Wiley and Sons.
- Hosmer, D. W., & Lemeshow, S. (2013). *Applied logistic regression*. New York: John Wiley & Sons.
- Ibarra, M. d. C., & Michalus, J. C. (2010). Análisis del rendimiento académico mediante un modelo logit. *Ingeniería Industrial*, 9(2), 47-56.
- Kleinbaum, D. G. (1994). *Logistic Regression: A Self-Learning Text*. New York: Springer-Verlag.
- Larrán Jorge, M., & Andrades Peña, F. J. (2015). La oferta de asignaturas de responsabilidad social corporativa y ética empresarial en las titulaciones de finanzas y contabilidad: análisis comparativo con el ámbito de la gestión de organizaciones. *Revista de Contabilidad – Spanish Accounting Review*, 18(1), 1-10.

- Lemos Hoyos, M., Jaller Jaramillo, C., González Calle, A. M., Díaz León, Z. T., & de La Ossa, D. (2012). Perfil cognitivo de la dependencia emocional en estudiantes universitarios en Medellín, Colombia. *Universitas Psychologica*, 11(2), 395-404.
- Mendoza Requena, D., Llanos Zavalaga, F., Rosas Aguirre, Á., Contreras Ríos, C., & Huayta Zacarías, E. (2001). Factores asociados al ingreso a medicina en la Universidad Peruana Cayetano Heredia y sus tendencias durante el periodo 1994-1999. *Revista Médica Herediana*, 12(1), 8-15.
- Menéndez Acuña, E. P., Gabriel Argüelles, E., & Hernández González, S. (2015). Nonparametric Regression: An Alternative to the Scatter Diagram. *Revista Investigación Operacional*, 36(2), 156-160.
- Morrison, D. G. (1972). Upper Bounds for Correlations Between Binary Outcomes and Probabilistic Predictions. *Journal of the American Statistical Association*, 7(68-70).
- Mures Quintana, M. J., García Gallego, A., & Vallejo Pascual, M. E. (2005). Aplicación del Análisis Discriminante y Regresión Logística en el estudio de la morosidad en las entidades financieras. Comparación de resultados. *Pecvnia*, 1, 175-199.
- Nieto Méndez, A. L. (2015). *Estimación de la probabilidad de egreso de estudiantes de licenciatura en ciencias de la BUAP usando Regresión Logística*. Unpublished Tesis que para obtener el título de Licenciada en Matemáticas Aplicadas, Benemérita Universidad Autónoma de Puebla, Puebla.
- Restrepo Molina, L., Agudelo Martínez, A., Giraldo Torres, M. T., & Sánchez Diosa, L. (2011). Factores de contexto asociados al consumo de alcohol en estudiantes universitarios. *Revista CES Salud Pública*, 2(2), 136-147.
- Salas Velasco, M. (1996). La regresión logística . Una aplicación a la demanda de estudios universitarios. *Estadística Española*, 38(141), 193-217.
- Yusif, H., Yussof, I., & Osman, Z. (2013). Public university entry in Ghana: Is it equitable? *International Review of Education*, 59, 7-27.