

Diseño de un modelo predictivo para otorgar créditos*

Juan Gabriel Salazar Vergara**

Recibido: 12 de marzo de 2022 • Aprobado: mayo 24 de 2022

<https://doi.org/10.22395/seec.v24n57a15>

RESUMEN

El objetivo de este artículo es presentar evidencias sobre la evaluación de la probabilidad de incumplimiento en el pago de un crédito otorgado por los fondos de empleados de instituciones de educación superior de Medellín y Área Metropolitana en el periodo 2017-2019. Para determinar la probabilidad de pago de los créditos se estimó un modelo econométrico de regresión logística binaria. Las estimaciones permitieron evidenciar que de los 384 casos estudiados doscientos cuarenta (el 62,5 %) tiene probabilidad de pagar y ciento cuarenta y cuatro (el 37,5 %) presentan probabilidad de no pago de sus compromisos.

PALABRAS CLAVE

Regresión logística binaria; riesgo de crédito; asignación de crédito; sistema financiero; Colombia.

CLASIFICACIÓN JEL

G17, G23

CONTENIDO

Introducción; 1. Estado del arte; 2. Marco teórico y conceptual; 3. Metodología; 4. Resultados; 5. Conclusiones; ; Referencias; Anexos.

* Artículo de investigación generado a partir del trabajo de grado titulado "Gestión de riesgos financieros crediticios y operacionales en fondos de empleados universitarios de Medellín y Área Metropolitana", realizado como requisito para optar al título de Doctor en Administración Gerencial, Universidad Benito Juárez García, Puebla, México. El estudio de doctorado se realizó en el período 2017-2020, que contó con la asesoría (tutor) de la Dra. María Esther Alcántara Gutiérrez, Doctora en Dirección y Mercadotecnia, Universidad Benito Juárez, México.

** Profesional en Negocios Internacionales, Universidad de Medellín, Colombia. Especialista en Alta Gerencia, Universidad UNIR, Bogotá, Colombia. Magíster en Finanzas, Institución Universitaria Esumer, Medellín, Colombia. Maestrando en Administración Gerencial (MBA), Universidad Benito Juárez García, Puebla, México. Doctorando en Administración Gerencial, Universidad Benito Juárez García, Puebla, México. Docente posgrados en Ciencias Económicas y Administrativas, Instituto Universitario Veracruzano, Veracruz, México, Correo electrónico: juan.salazar@iuv.edu.mx. Orcid: <https://orcid.org/0000-0002-3006-4262>.

Design of a Predictive Model for Granting Credits

ABSTRACT

The objective of this article is to present evidence on the evaluation of the probability of default in the payment of a loan granted by employee funds of higher education institutions in Medellín and the Metropolitan Area in the period 2017-2019. To determine the probability of loan repayment, a binary logistic regression econometric model was estimated. The estimates showed that of the 384 cases studied, two hundred and forty (62.5 %) are likely to pay and one hundred and forty-four (37.5 %) are likely not to pay their commitments.

KEYWORDS

Binary logistic regression; credit risk; credit allocation; Finance system; Colombia.

JEL CLASSIFICATION

G17, G23.

CONTENT

Introduction; 1. State of the art; 2. Theoretical and conceptual framework; 3. Methodology; 4. Results; 5. Conclusions; Rerefences; Annexes.

Desenho de um modelo preditivo para concessão de créditos

RESUMO

O objectivo deste artigo é apresentar provas sobre a avaliação da probabilidade de incumprimento no pagamento de um empréstimo concedido pelos fundos dos empregados das instituições de ensino superior em Medellín e na Área Metropolitana no período 2017-2019. Para determinar a probabilidade de reembolso do empréstimo, foi estimado um modelo econométrico de regressão logística binária. As estimativas mostraram que dos 384 casos estudados, duzentos e quarenta (62,5 %) são susceptíveis de pagar e cento e quarenta e quatro (37,5 %) são susceptíveis de não pagar os seus compromissos.

PALAVRAS CHAVE

Regressão logística binária; risco de crédito; alocação de crédito; Sistema financeiro; Colômbia.

CLASSIFICAÇÃO JEL

G17, G23.

CONTEÚDO

Introdução; 1. Estado da arte; 2. Enquadramento teórico e conceitual; 3. Metodologia; 4. Resultados; 5. Conclusões; Referências; anexos.

INTRODUCCIÓN

Los rápidos cambios en la economía mundial, unido al crecimiento de la competencia para muchas organizaciones, hacen que la flexibilidad sea protagonista a la hora de otorgar créditos; con ello vienen una serie de riesgos en los cuales incurren muchas organizaciones, entidades solidarias y entidades financieras, por ejemplo, la falta de rigurosidad en la asignación de créditos es uno de los aspectos que puede originar situaciones de riesgo de no pago.

Para el año 2017 en Colombia existían alrededor de 1.484 fondos de empleados reportando información a la Superintendencia de economía solidaria de Colombia, el número de empleados afiliados a los fondos era de 15.492, los excedentes generados por estos fondos ascendieron a COP\$151.027.507.389 (Superintendencia de Economía Solidaria [SES], 2017a), es decir, unos USD\$51.172.821,45¹.

Si bien los fondos de empleados son entidades solidarias, también son colocadoras de créditos. Lo anterior deriva en riesgos muy parecidos a los que incurre cualquier entidad financiera. La falta de controles en los fondos de empleados podría ocasionar pérdidas para dichas entidades, si se tienen en cuenta las estadísticas de la SES (2017b), según la cual, en los últimos años han desaparecido algunos fondos de empleados; esta situación, amerita establecer más control de los riesgos financieros.

Con el fin de evaluar la gestión de riesgos crediticios, por ser la gestión del riesgo crediticio, de acuerdo a De Lara (2013), el más importante riesgo a controlar para cualquier entidad crediticia, en el sector de fondos de empleados universitarios de Medellín y Área Metropolitana en el periodo comprendido entre 2018-2019, se propone la construcción y estimación de un modelo econométrico por medio de la técnica de regresión logística binaria.

La importancia de una adecuada gestión del riesgo en los fondos de empleados y entidades financieras, radica en la detección temprana de probabilidad de incumplimiento de un cliente para su adecuada gestión.

En la actualidad, los fondos de empleados generalmente clasifican a sus deudores por tipo de crédito. Para esto se tienen en cuenta aspectos como: salario², edad y género, entre otras variables.

¹ Valor calculado usando una tasa de cambio (TRM) promedio de \$2951 para 2017, esta TRM promedio se calculó utilizando los valores diarios registrados por el Banco de la República.

² Componente referido a la prestación económica de unos servicios realizados por un trabajador (Llompart, 2007).

El análisis y exploración tiene como punto de partida la búsqueda de evidencias para responder al interrogante: ¿la falta o deficiencia de instrumentos predictivos para la gestión de los créditos de los fondos de empleados de Medellín y el Área Metropolitana, aumenta las probabilidades de riesgos crediticios?

1. ESTADO DEL ARTE

La revisión de los trabajos que se han realizado sobre riesgos financieros se centró en la identificación de la metodología y la información utilizada, así como los principales resultados obtenidos en estos trabajos. A continuación se presenta algunas características de los trabajos identificados sobre el tema:

Cardona (2004) utiliza los árboles de decisión como herramienta para el cálculo de probabilidades de incumplimiento en crédito, mostrando sus ventajas y desventajas.

Toro y Palomo (2014), en su tesis doctoral llamada *Análisis del riesgo financiero en las pymes- estudio de caso aplicado a la ciudad de Manizales*, realizan un estudio integral interesante en cuanto a riesgos financieros, destaca el análisis de diferentes tipos de riesgos financieros.

Por otra parte, Gisbert (2017), en su tesis doctoral llamada *La gestión macro prudencial bancaria: una propuesta de tres modelos de previsión de riesgos*, propone tres interesantes modelos para gestionar de una manera más contundente el riesgo sistémico en entidades financieras internacionales.

Jimber (2015), en su tesis doctoral llamada *Riesgo de crédito. Modelización econométrica*, plantea que:

Un estudio empírico relacionado con la predisposición de las entidades financieras a participar en un proceso de reestructuración bancario, y un estudio estadístico que se centra en el cálculo y posterior análisis de los efectos marginales mediante la estimación de cinco modelos logísticos de variable endógena binaria. (p.4)

Fernández y Pérez (2005) proponen un modelo logístico clasificatorio entre clientes buenos y malos pagadores de una cartera comercial y una cartera de consumo. Para su desarrollo tienen en cuenta los aspectos críticos de las variables utilizadas y la correlación existente entre las variables estudiadas.

Meneses y Macuacé (2011), examinan la metodología de análisis financiero tradicional en Colombia, destacan algunas experiencias de Latinoamérica con nuevos

modelos probabilísticos de análisis financiero y, finalmente, señalan las bondades de dichos modelos a partir de sus ventajas y desventajas.

López (2013) propone un modelo con datos simulados. Al respecto, vale la pena decir que es interesante el modelo, aunque al ser con datos simulados, se da un sesgo con respecto a la realidad, además se ofrece una solución parcial para el manejo de riesgos crediticios.

Pérez (2017) proponen un modelo econométrico para estimar la probabilidad de incumplimiento para muchas Organizaciones de la Economía Social y Solidaria (OESS) del Ecuador; para ello se concentra en el análisis de componentes atribuibles a la demanda de crédito como: registro de impagos, número de años de funcionamiento de la organización, ingresos por ventas, etc.

2. MARCO TEÓRICO Y CONCEPTUAL

2.1. Crédito

El otorgamiento de créditos hace parte fundamental del qué hacer de los fondos de empleados y otras entidades financieras, sumado a los beneficios que reciben los asociados de estas entidades para su crecimiento personal y de la sociedad en general.

El crédito es un préstamo de dinero que una entidad financiera otorga, con el compromiso de que en el futuro sea devuelto en forma gradual.

El crédito es un beneficio que las organizaciones otorgan a sus clientes, conceder a un cliente un producto o un servicio hoy con el compromiso de obtener un pago por ello en el futuro (Dunn, 2006). En el mismo sentido, Miranda (2016) señala que:

El crédito es la cesión provisional de la potestad de disponer de unos fondos con el compromiso de la persona deudora de pagar a la acreedora los intereses devengados y devolver, también, el capital dispuesto en un plazo o dentro de una fecha establecida. (p. 48)

2.2. Riesgo

Otorgar créditos genera unos riesgos, tanto en su otorgamiento como en su seguimiento y control. El riesgo financiero es parte fundamental, inherente a la prestación de servicios de crédito. Se puede definir como la "[p]osibilidad de quebranto o pérdida derivada de la realización de operaciones financieras que pueden afectar a la capitalización bursátil o valor de mercado de la empresa" (Gómez y López, 2002, p. 21)

Los diferentes tipos de riesgos financieros interactúan unos con otros, por ello es difícil delimitar sus fronteras al observar el resultado de un proceso. El riesgo crediticio es la posibilidad de que una organización entre en proceso de pérdidas y pueda, en casos más críticos, caer en *default* o cese de pagos, a causa de que sus deudores falten con sus pagos o paguen parcialmente en las fechas específicas pactadas con anterioridad al vencimiento de los créditos. Vale aclarar que todo tipo de créditos presentan riesgos, en mayor o menor medida (Superintendencia Bancaria de Colombia [SBC], 2002).

Algunas de las causas por las cuales se da el impago de las obligaciones son: desempleo, inflación, corrupción, disminución de recursos. De acuerdo con Alonso y Berggrun (2015):

El riesgo crediticio es aquel asociado a la posibilidad de que un deudor incumpla sus obligaciones, ya sea parcial o completamente. El riesgo de crédito implica no solamente el riesgo que un deudor incumpla su obligación, sino también que pague solamente parcialmente y/o después de la fecha convenida. (p. 4)

De Lara (2013) afirma:

El riesgo de crédito es la pérdida potencial que se registra con motivo del incumplimiento de una contraparte en una transacción financiera (o en algunos de los términos y condiciones de la transacción). También se concibe como un deterioro en la calidad crediticia de la contraparte o en la garantía o colateral pactada originalmente. (p. 163)

2.3. Gestión de riesgos financieros

Los fondos de empleados, por su objeto social y la administración de activos financieros, conllevan unos riesgos de crédito inherentes, por la constitución de dichos elementos, es que se vuelve fundamental el papel de la administración adecuada de riesgos. De acuerdo con Alonso y Berggrun (2015): “[l]a gestión del riesgo es el proceso de proteger los activos e ingresos de una organización empleando una aproximación científica para afrontar el riesgo independientemente de su naturaleza u origen” (p. 8).

El riesgo de incumplimiento se presenta cuando el deudor se encuentra limitado a seguir con el cumplimiento de la deuda. Este aspecto se agudiza cuando variables macroeconómicas como la inflación o el desempleo afectan de manera directa o indirecta al deudor. Dice García (2009) que: “[a]dicionalmente es importante identificar la capacidad de pago, la cual está determinada por la capacidad de generar flujo de caja libre para atender el servicio a la deuda” (p. 236).

La adecuada gestión de los riesgos financieros crediticios es fundamental para una organización, en particular en los fondos de empleados, ya que es importante clasificar a un cliente como bueno o malo para así disminuir pérdidas en cualquier entidad.

Para De Lara (2013), el riesgo crediticio por incumplimiento es el más importante riesgo financiero a tratar por parte de las entidades, sobre todo si su competencia central de negocio es el financiamiento.

Existen diferentes metodologías para estimar probabilidades de incumplimiento; por ejemplo, técnicas estadísticas univariadas, multivariadas, análisis de modelos de clasificación, arboles de decisión, modelos de elección cualitativa (*Probit* y *Logit*) y el análisis de matrices de transición entre otros (Elizondo, 2004).

En la época actual sobresalen modelos probabilísticos mucho más sofisticados para estimar la probabilidad de incumplimiento o *default* de un cliente, y con ello poder otorgarle préstamos que este solicite en un momento dado. Al respecto, se pueden citar algunos de estos modelos: análisis discriminante, análisis de regresión simple, análisis de regresión múltiple, análisis regresión logística binaria y múltiple, redes neuronales³ entre otros.

En los modelos de regresión, la variable que se va a pronosticar —endógena— se formula como una función matemática de otras variables exógenas. Para Ocampo et al. (2006), las redes neuronales son importantes porque “permiten plantear métodos de pronóstico tipo combinado; entre sus fortalezas se destaca la capacidad de manejar discontinuidades en la información; además, pueden procesar varias series de datos y cruzarlas entre sí” (p. 58).

2.4. Modelos importantes para clasificación de clientes

2.4.1. Modelo Z-Score

Este método se basa en el análisis discriminante o clasificatorio de grupos, un grupo cumple y el otro no. Principalmente es un método estadístico utilizado para clasificar observaciones en grupos clasificados a priori. El modelo Z-Score utiliza indicadores financieros que tienen como objetivo clasificar a las organizaciones en dos grupos: Bancarrotas y No Bancarrotas (Elizondo, 2004).

³ “Las redes neuronales son grafos organizados por capas en cuyos vértices se procesa información, con un tipo de funciones matemáticas determinadas y con la característica de ser densamente interconectadas” (Ocampo et al., 2006, p. 59).

2.4.2. Modelo Zeta

En 1977, Altman, Haldeman y Narayanan mejoran el modelo anterior, ya que este nuevo modelo predice de forma más precisa que el anterior modelo Z-Score original. De acuerdo a Elizondo (2004), el propósito de este modelo conocido como modelo Zeta es clasificar a las empresas en bancarrota incluyendo lo siguiente:

1. Organizaciones medianas y grandes en el análisis.
2. Organizaciones del sector no-manufacturero.
3. Los cambios en los estándares de cálculo de las más importantes razones financieras y nuevas prácticas contables.
4. Técnicas más actuales del análisis estadístico para la estimación del análisis discriminante.

2.4.3. Modelo EMS o *Emerging Market Corporate Bonds Scoring System*

Este modelo muestra un análisis de clasificación del riesgo financiero de una organización y se emplea para calcular el valor relativo entre los diferentes créditos.

2.4.4. Modelo Credit Monitor de KMV (clasificador de empresas)

Para el cálculo de la frecuencia de incumplimiento esperada (EDF), la empresa KMV construyó un modelo de probabilidad de incumplimiento —Credit Monitor (CM)— que permite medir la frecuencia de incumplimiento esperada. Este modelo está construido para convertir la información dada en el precio de una acción en una medida de riesgo de incumplimiento (Sellers et al. 2000, citados por Elizondo, 2004).

2.4.5. Modelos econométricos Probit o Logit

En la actualidad, estos se utilizan de forma amplia para modelos crediticios; en consonancia con De Lara (2013), son modelos de elección cualitativa que consisten en determinar la probabilidad de que un individuo que tiene ciertas características pertenezca a un grupo o a otro, pueden ser malos pagadores en un grupo y en el otro los buenos pagadores. Se denota por “p” a la probabilidad de incumplimiento del acreedor, esta solo puede tomar valores entre 0 y 1.

De acuerdo con Gujarati y Porter (2010), la función de distribución logística se representa como:

$$P_i = \frac{1}{1+e^{-z_i}} = \frac{e^{z_i}}{1+e^{z_i}} \quad [1]$$

Además, argumenta que la razón de probabilidad de ocurrencia de un suceso comparado con la no ocurrencia del mismo se puede representar como:

$$\frac{P_i}{(1-P_i)} = \frac{1+e^{z_i}}{1+e^{-z_i}} = e^{z_i} \quad [2]$$

Al emplear el logaritmo natural se tiene:

$$L_i = \ln\left(\frac{P_i}{1-P_i}\right) \quad [3]$$

$$Z = \beta_0 + \beta_1 x1i + \beta_2 x2i + \dots + \beta_k xki \quad [4]$$

De modo que L corresponde al logaritmo de la razón de las probabilidades o Logit.

Gujarati y Porter (2010), plantean unas características y condiciones para el modelo Logit:

- a) El P toma valores de 0 a 1, o sea, cuando Z varía de $-\infty$ a $+\infty$, el Logit L va de $-\infty$ a $+\infty$. De acuerdo a lo anterior, puede decirse que aunque las probabilidades se encuentran entre 0 y 1, los Logit no están acotados en esta forma.
- b) El modelo permite incluir tantas variables independientes como se requiera.
- c) Si L Logit es positivo, significa que cuando se incrementa el valor de las variables independientes aumentan las posibilidades de que la variable dependiente sea igual a 1. Si el Logit es negativo, las posibilidades de que la endógena sea igual a 1 disminuyen conforme se incrementa el valor de X.
- d) Una interpretación más puntual del modelo: β_2 la pendiente, evalúa el cambio en L ocasionado por el cambio unitario en X. El intercepto β_1 es el valor del logaritmo de las posibilidades en favor de ocurrencia del suceso, si la variable X es igual a 0.
- e) El modelo Logit supone que el logaritmo de la razón de probabilidad está relacionado linealmente con X_i .

De acuerdo con Gujarati y Porter (2010), el modelo se puede reescribir así:

$$L_i = \ln\left(\frac{P_i}{1-P_i}\right) = \beta_1 + \beta_2 X_i + u_i \quad [5]$$

Donde u_i es el término de error estocástico.

Los modelos de regresión simple y múltiple han sido muy utilizados para clasificar clientes potenciales, pero tienen limitantes como: se utilizan con variables dependientes cuantitativas y no permiten utilizar variables de respuesta cualitativa; a diferencia de los modelos logísticos que permiten en su variable respuesta tanto valores cualitativos como cuantitativos, entre este tipo de modelos se encuentran modelos de regresión logística binaria y de regresión logística múltiple.

2.4.6. Modelo de Regresión Logística Binaria

Este tipo de modelos tiene como característica fundamental que su variable respuesta o endógena es una variable dicotómica de tipo mixto, es decir, permite valores tanto cuantitativos como cualitativos.

De acuerdo con De Lara (2013), para formular este tipo de modelos es necesario que la variable endógena o dependiente sea binaria, ya que las dos situaciones posibles es que el cliente sea bueno o malo, es decir, paga (1) o el cliente no paga (0). En este estudio se empleará un modelo de regresión logística binaria con el propósito de evitar las limitaciones que presentan los modelos de regresión lineal.

Se eligió la regresión logística binaria debido a sus ventajas: I. Permite variables categóricas con mayor tolerancia que los modelos lineales; II. Permite pronosticar la probabilidad de impago o *default* del crédito, de acuerdo a los parámetros de las variables independientes; y III. Evalúa el peso de cada variable exógena sobre la variable endógena paga o no paga, de acuerdo al OR (*Odds Ratio* o ventaja).

El OR, en el software SPSS versión 22, se puede establecer como $\exp(\beta)$, donde \exp es la base de los logaritmos neperianos "e" cuyo valor matemático es 2,718, y β es el valor de regresión de la variable independiente en el modelo. Una OR mayor que 1 significa un incremento en la posibilidad del evento de incumplimiento sobre el hecho de pagar cuando la variable exógena se incrementa en una unidad; inversamente, una OR menor que 1 indica lo contrario. El modelo de regresión logística binaria puede expresarse así:

$$\ln\left(\frac{p}{1-p}\right) = b_0 + b_1 \cdot X_1 + b_2 \cdot X_2 + b_3 \cdot X_3 + \dots + b_i \cdot X_i \quad [6]$$

Donde "p" es la probabilidad de ocurrencia del evento, en este caso, impago o *default*.

O también, de acuerdo con (Rayo et al., 2010), se puede calcular la función logística de la siguiente forma:

$$f(p) = \frac{e^{(z)}}{1+e^{(z)}} = \frac{1}{1+e^{-(z)}} \quad [7]$$

Donde Z, sería:

$$Z = b_0 + b_1 \cdot X_1 + b_2 \cdot X_2 + b_3 \cdot X_3 + \dots + b_i \cdot X_i \quad [8]$$

En los modelos de regresión logística en general, y la binaria en este caso, se utilizan unos indicadores que sirven para evaluar el ajuste del modelo, la predictibilidad etc. El coeficiente R cuadrado de Cox y Snell es un coeficiente de determinación que se usa para calcular la proporción de varianza de la variable endógena explicada por las variables exógenas (Cox y Snell, 1989).

El R cuadrado de Nagelkerke, de acuerdo a Menard (2002), es semejante al coeficiente R2 de determinación en las regresiones lineales, es una versión corregida de la R cuadrado de Cox y Snell (1989). Muestra que los coeficientes de las variables del modelo son significativos y distintos de cero, además que el modelo es representativo.

Para evaluar la bondad del ajuste de un modelo se usa el estadístico Hosmer-Lemeshow. De acuerdo con Hosmer y Lemeshow (2000), este calcula la bondad de ajuste del modelo, es decir, un valor alto de probabilidad predicha se relaciona con un valor 1 de la variable endógena dicotómica, se trata de comparar las probabilidades estimadas o predichas con las observadas en una tabla de contingencia para determinar el ajuste del modelo.

Un instrumento importante a tener en cuenta en modelos logísticos para clasificación de individuos, es la llamada curva ROC; según Hosmer y Lemeshow (2000), es un mecanismo de carácter estadístico usado en el análisis de la clasificación de individuos en un grupo específico o en otro de acuerdo a sus características. Es una prueba para clasificar a los individuos de una población en dos grupos: un grupo que muestre un suceso específico y el otro que tenga ausencia de este.

2.5. Marco legal

2.5.1. Circular Externa 088 de 2000 de la Superintendencia Bancaria

Esta circular aclaró las pautas mínimas de gestión de riesgos que deberán cumplir las organizaciones vigiladas por la Superintendencia Financiera para la realización de sus operaciones de tesorería.

2.5.2. Comité de Basilea

Es un comité conformado por bancos centrales y supervisores o reguladores bancarios de los principales países industrializados.

De acuerdo a Muñoz y Soler (2018), los acuerdos importantes de Basilea III son: a). La instauración de un piso en el cálculo de los activos ponderados por riesgo (APR). b). El consenso sobre un enfoque normalizado y examinado relativo al riesgo de crédito optimizará la estabilidad y la sensibilidad al riesgo. c). Restricción del empleo de modelos avanzados para algunos tipos de activos, para las que es complejo instaurar un modelo consistente.

En Colombia se trata de adoptar estos derroteros mediante un marco llamado el sistema de administración de riesgo crediticio (SARC).

3. CARACTERIZACIÓN Y DEFINICIÓN DE LAS VARIABLES

La caracterización de variables es importante porque facilita desagregar las variables que conforman el problema de investigación, los objetivos, las hipótesis etc.; para su medición, dichas variables se pueden segmentar en dimensiones, áreas, indicadores, índices, etc. La tabla 1 muestra las variables del presente estudio.

La variable endógena o dependiente será el *default* o incumplimiento, se refiere a la posibilidad de que el deudor efectúe o no pagos a la deuda, la cancele completamente o no. Esta variable tomará valores de cero cuando de acuerdo a las variables independientes se da una posibilidad de que el deudor no pague la deuda y uno cuando se dé la posibilidad de pagos.

Las variables independientes, las cuales explican la variable dependiente y de las cuales se escogerán las más significativas de acuerdo a su contribución en la predicción de la variable dependiente son: tipo de crédito, educación, plazo, tasa de colocación, calificaciones de riesgo, edad, género, localización, jurídico, empleo, ingresos, disponible para embargo, mora, saldo a capital.

Tabla 1. Caracterización de las variables

<i>Variables potenciales explicativas de la variable dependiente para la construcción del modelo econométrico</i>				
<i>Objetivo</i>	<i>Variable</i>	<i>Tipo de variable</i>	<i>Definición</i>	<i>Dimensión</i>
Proponer un modelo que prediga la morosidad de riesgo financiero crediticio sujeto a unas restricciones.	Tipo de crédito	Categoría-dicotómica	Informa sobre los tipos de crédito que maneja cada entidad.	Crédito de consumo, vivienda, comercial y microcrédito
	Educación	Categoría-dicotómica	Informa sobre el nivel de instrucción de cada cliente.	No bachiller, bachiller, pregrado, posgrado
	Plazo	Numérica-discreta	Informa sobre el tiempo establecido para cubrir cada crédito.	Plazo menor de treinta días. Plazo de treinta a sesenta días. Plazo de sesenta a noventa días Plazo mayor a noventa días
	Tasa de colocación	Numérica-continua	Informa sobre la tasa de interés a la cual se prestarán los recursos.	
	Calificaciones de riesgo	Categoría-dicotómica	Informa sobre el historial en centrales de riesgo de los clientes.	Calificaciones en centrales de riesgo. No calificaciones en centrales de riesgo
	Edad	Numérica-discreta		
	Genero	Categoría-dicotómica	Informa sobre el género de cada cliente. En el trimestre móvil febrero – abril 2018, "la población desocupada estuvo compuesta por un 44,4 % hombres y un 55,6 % mujeres" (Departamento Nacional de Estadística [DANE], 2018).	Femenino. Masculino
	Disponible embargo	Numérica-continua	Informa sobre el disponible para embargo en caso de incumplimiento de un cliente.	Nivel de disponible para embargo
	Mora	Numérica-discreta	Informa si el cliente está o no en mora y los días de mora que presenta.	Mora Sin mora

Cont. Tabla 1

<i>Variables potenciales explicativas de la variable dependiente para la construcción del modelo econométrico</i>				
<i>Objetivo</i>	<i>Variable</i>	<i>Tipo de variable</i>	<i>Definición</i>	<i>Dimensión</i>
Proponer un modelo que prediga la morosidad de riesgo financiero crediticio sujeto a unas restricciones.	Variable dependiente: Incumplimiento o <i>default</i> .	Categórica, dicotómica	Informa sobre el incumplimiento total o parcial del crédito por parte del cliente.	Cliente bueno-Cliente malo.
	Localización	Categórica, dicotómica	Indica si se tiene la dirección de residencia o trabajo del cliente, para lograr su ubicación (si el cliente está localizado, el riesgo de impago disminuye, pero si no está localizado el riesgo aumenta).	Cliente localizado Cliente no localizado
	Jurídico	Categórica, dicotómica	Indica, si el cliente presenta o no, procesos jurídicos en la entidad.	Proceso jurídico No proceso jurídico
	Empleo	Categórica, dicotómica	Indica si el cliente presenta o no una labor determinada (si el cliente está empleado el riesgo de impago disminuye, pero si está desempleado el riesgo aumenta).	Cliente empleado Cliente desempleado
	Ingresos	Numérica, Escala continua	Monto que recibe el cliente, derivado de actividades diversas (si el cliente presenta ingresos el riesgo de impago disminuye, pero si no presenta el riesgo aumenta).	Nivel de ingresos.
	Saldo a capital	Numérica, Escala continua	Informa sobre el capital adeudado, en una determinada fecha (entre más alto sea el saldo a capital adeudado, mayor será el riesgo y viceversa).	Nivel de saldo a capital.

Fuente: elaboración propia.

4. METODOLOGÍA

El tipo de investigación será aplicada con un enfoque mixto, cuyo nivel de investigación será explicativo correlacional o causal y donde se cuantifican relaciones entre conceptos y variables (Hernández et al., 2014). Por cuanto busca gestionar de una forma adecuada los riesgos financieros crediticios en fondos de empleados de Medellín y el Área Metropolitana. Para Carrasco (2005) los estudios de tipo correlacional pretenden visualizar cómo se relacionan o vinculan diversos fenómenos entre sí. El procedimiento para realizar el modelo econométrico para mitigación del riesgo crediticio es el siguiente:

- I. Se limpia la base de datos de clientes suministrada por un fondo de empleados universitario de la ciudad de Medellín, es decir, se excluyen aquellos datos que no cumplen con los estándares que exige una regresión logística, esto es, que los clientes tengan datos completos, de lo contrario se excluyen. En este caso se trata de un fondo de empleados universitario de cuatrocientos registros de los cuales se tomarán todos los datos por ser una base de datos considerada pequeña, claro está, restando los registros que no cumplen los requisitos exigidos, la base de datos final se compone de trescientos ochenta y cuatro registros.
- II. Se llevan los datos al programa estadístico SPSS versión 22 y se codifican las variables endógenas y variables independientes, donde se toma como variable endógena, *default* o incumplimiento; dicha variable es de tipo dicotómica para tomar dos posibles valores: $Y=1$ si el cliente resulta ser bueno o cumplido, $Y=0$ si el cliente resulta ser malo o incumplido.
- III. En la opción de regresión logística binaria del programa SPSS se clasifican dichas variables entre cualitativas y cuantitativas, se corre el modelo varias veces con las variables que se consideran valiosas para el modelo mediante el método de pasos hacia adelante de Wall. Esto, con el fin de ir eliminando las variables que no son significativas para el modelo.
- IV. Obteniendo el modelo final después de haber corrido modelos menos significativos, se procede a evaluar sus indicadores con la prueba ómnibus que arroja el SPSS, se evalúan la restricción de que los coeficientes del modelo no sean iguales a cero, observando que el p-valor sea menor del 5 %.
- V. Se evalúa la capacidad de predicción del modelo, es decir, las variaciones de la variable *default* dependen de las variables independientes incluidas en el modelo con el R cuadrado de Nagelkerke y la R cuadrado de Cox y Snell (1989).

- VI. Se observa la justeza del modelo predicho, comparado con el observado, mediante el estadístico de Hosmer y Lemeshow y su tabla de contingencia.
- VII. Se realiza una curva llamada ROC en SPSS para observar el área bajo la curva, este también es un importante indicador de nivel de predicción del modelo.
- VIII. Se obtiene la ecuación del modelo, para cada variable de la ecuación se obtiene: coeficiente (b), error típico de (b), estadístico de Wald (R), razón de ventajas estimada (exp (b)), intervalo de confianza para la exp (b), log de verosimilitud si el término se ha eliminado del modelo.
- IX. Una vez estimado el modelo, es posible hacer predicciones sobre la probabilidad de que un crédito nuevo sea malo o incumplido, conociendo las variables características del cliente.

Para la validación de este modelo, se toma una muestra aleatoria de aproximadamente 20 % del total de la base de datos final estipulada para el análisis.

4.1. Recolección de información

Las variables potenciales para el modelo fueron recopiladas con base en entrevistas a expertos, la técnica de investigación empleada fue la entrevista semiestructurada. De acuerdo con Hernández et al. (2014), esta técnica se fundamenta en un diálogo revelador con el entrevistado, también a través de informes previos de la Superintendencia Solidaria, gremios del sector solidario, bases de datos académicas, artículos académicos etc. La base de datos para la construcción del modelo fue suministrada por un fondo de empleados.

4.2. Restricciones del Modelo Econométrico

4.2.1. Prueba de significancia

Para definir las variables que son significativas estadísticamente se debe validar el cumplimiento de las siguientes condiciones:

1. El coeficiente de las variables sea distinto de cero.
2. P valor < 0.05, esto rechaza la hipótesis de que el coeficiente sea 0.
3. La presentación de no correlaciones altas entre las variables independientes.

5. RESULTADOS

5.1. Hipótesis particulares del modelo econométrico

5.1.1. Hipótesis nula

La predictibilidad de la variable dependiente “incumplimiento” de un crédito de un fondo de empleados no depende de las variables explicativas o independientes propias del modelo (cliente).

5.1.2. Hipótesis de trabajo

La predictibilidad de la variable dependiente “incumplimiento” de un crédito depende de las variables explicativas o independientes propias del modelo (cliente).

5.2. Desarrollo Modelo Econométrico

El método de selección de variables fue por pasos hacia delante de Wald, el cual comienza con el modelo que incluye solo el término constante y se van añadiendo al mismo modelo las variables independientes, según su grado de relación con la variable dependiente y su significación estadística para llegar a escoger el mejor modelo. Comprobación de supuestos de no colinealidad del modelo mediante el programa SPSS

La tabla 2 muestra que los factores de inflación de la varianza (FIV) son iguales a 1; por lo tanto, se cumple el supuesto de la no colinealidad entre las variables independientes del modelo.

Tabla 2. Estadísticas de colinealidad

Modelo	Estadísticas de colinealidad	
	Tolerancia	VIF
(Constante)		
ubicación o no del cliente	0,974	1,026
1 procesos jurídicos	0,991	1,009
empleado o desempleado	0,977	1,023
ingresos totales	0,983	1,017
Saldo a capital	0,985	1,015

Fuente: elaboración propia.

La tabla 3 muestra el número de casos, donde no se presentan casos perdidos.

Tabla 3. Resumen de procesamiento de casos

<i>Casos sin ponderar^a</i>		N	Porcentaje
Casos seleccionados	Incluido en el análisis	384	100,0
	Casos perdidos	0	0,0
	Total	384	100,0
Casos no seleccionados		0	0,0
Total		384	100,0

Fuente: elaboración propia.

La tabla 4 revela que la variable dependiente es dicotómica, asigna el valor de 0 a clientes que tienen más probabilidades de no pagar que de pagar y de 1 a clientes que tienen más probabilidades de efectuar el pago.

Tabla 4. Codificación de variable dependiente

<i>Valor original</i>	<i>Valor interno</i>
no paga	0
paga	1

Fuente: elaboración propia.

La tabla 5 presenta la prueba significancia chi-cuadrado, la cual indica si el modelo es significativo o no. En este caso la significancia del modelo es menor a 0,05; es decir, el modelo es significativo.

Tabla 5. Pruebas de ómnibus de coeficientes del modelo

		<i>Chi-cuadrado</i>	<i>gl</i>	<i>Sig.</i>
Paso 1	Escalón	358,290	5	0,000
	Bloque	358,290	5	0,000
	Modelo	358,290	5	0,000

Fuente: elaboración propia.

La tabla 6 presenta el R cuadrado de Nagelkerke, este es una versión corregida de la R cuadrado de Cox y Snell, la R cuadrado de Cox y Snell tiene un valor máximo inferior a 1, incluso para un modelo "perfecto". La R cuadrado de Nagelkerke corrige la escala del estadístico para cubrir el rango completo de 0 a 1.

Tabla 6. Resumen del modelo

Escalón	Logaritmo de la verosimilitud -2	R cuadrado de Cox y Snell	R cuadrado de Nagelkerke
1	146,626 ^a	0,607	0,829

Fuente: elaboración propia.

La tabla 7 muestra la diferencia entre las frecuencias observadas y esperadas muy próximas, el modelo presenta un ajuste adecuado.

Tabla 7. Tabla de contingencia de Hosmer y Lemeshow para frecuencias observadas y esperadas

Observado	<i>paga o no paga = no paga</i>		<i>paga o no paga = paga</i>		Total
	Esperado	Observado	Esperado	Observado	
1	37	37,945	1	0,055	38
2	36	37,274	2	0,726	38
3	37	32,540	1	5,460	38
4	23	22,345	15	15,655	38
5	7	8,320	31	29,680	38
6	0	2,102	38	35,898	38
7	1	0,389	37	37,611	38
8	0	0,068	38	37,932	38
9	0	0,015	38	37,985	38
10	0	0,002	42	41,998	42

Fuente: elaboración propia.

La tabla 8 de clasificación indica el porcentaje de clientes que ha sido correctamente clasificado. En la base de datos de clientes para generar el modelo de regresión logística, se obtiene un 93,5 % de clasificación correcta.

Tabla 8. Tabla de clasificación

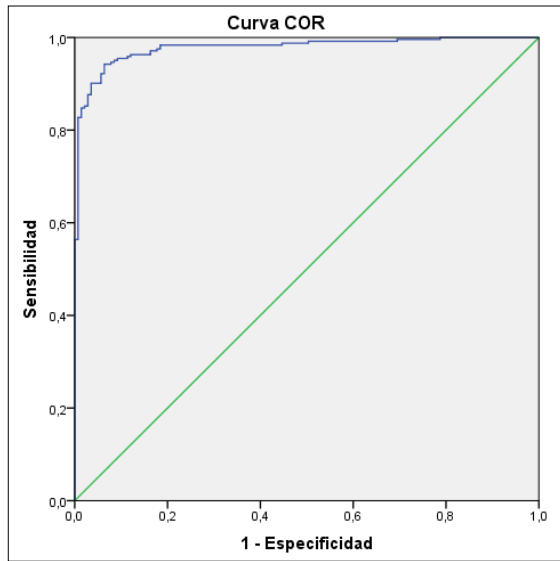
	Observado		Pronosticado		
			<i>paga o no paga</i>		Corrección de porcentaje
			<i>no paga</i>	<i>paga</i>	
Paso 1	<i>paga o no paga</i>	<i>no paga</i>	130	11	92,2
		<i>paga</i>	14	229	94,2
Porcentaje global					93,5

Fuente: elaboración propia.

5.3. Curva COR o ROC en inglés

La curva COR indica la capacidad de discriminación del modelo, analizando el área bajo la curva. El área bajo la curva para este modelo (grafico 1) es del 97,8 %; esta es la probabilidad de que al seleccionar al azar un asociado o cliente este sea clasificado correctamente como bueno o malo según el caso.

Gráfico 1. Curva COR



Fuente: elaboración propia.

En la tabla 9 se puede observar que el modelo se ajusta adecuadamente a los datos, el porcentaje de error es muy bajo 0,7 % y la precisión de discriminación del modelo es excelente.

Tabla 9. Datos curva COR

Área	Error estándar	Significación asintótica	95 % de intervalo de confianza asintótico	
			Límite inferior	Límite superior
0,978	0,007	0,000	0,965	0,991

Fuente: elaboración propia.

La tabla 10 permite formar el modelo logístico con las variables representativas como: Localización, jurídico, empleo, ingresos, saldo a capital. Este incluye en la segunda columna, la estimación de los coeficientes de las variables del modelo

y el estadístico de prueba Wald en la columna cuatro, el cual es un indicador para evidenciar la significancia estadística de cada variable en la ecuación.

Tabla 10. Variables en la ecuación

		Variables en la ecuación						95 % C.I. para EXP(B)	
		B	Error estándar	Wald	gl	Sig.	Exp(B)	Inferior	Superior
Paso 1ª	Localización(1)	-5,324	0,760	49,088	1	0,000	0,005	0,001	0,022
	jurídico(1)	-1,363	0,688	3,928	1	0,047	0,256	0,066	0,985
	Empleo(1)	-4,909	0,757	42,039	1	0,000	0,007	0,002	0,033
	ingresos	0,000	0,000	64,152	1	0,000	1,000	1,000	1,000
	Saldo capital	0,000	0,000	13,425	1	0,000	1,000	1,000	1,000
	Constante	-1,139	0,805	2,003	1	0,157	0,320		

Fuente: elaboración propia.

En el resumen del modelo, se puede evidenciar que los coeficientes de determinación son muy adecuados (tabla 6), por ejemplo, entre un 60,7 % el R cuadrado de Cox y Snell y un 82,9 %, esto indica que las variaciones de la variable *default* o incumplimiento dependen de las variables independientes incluidas en el modelo.

De esta forma se rechaza la hipótesis nula: la predictibilidad de la variable dependiente "incumplimiento" de un crédito de un fondo de empleados, no depende de las variables explicativas o independientes propias del modelo (cliente).

Y se acepta la hipótesis de trabajo: la predictibilidad de la variable dependiente "incumplimiento" de un crédito, depende de las variables explicativas o independientes propias del modelo (cliente).

El poder de predicción de un modelo es fundamental para clasificar bien los deudores de cualquier entidad colocadora de créditos. El modelo propuesto predice adecuadamente, en la tabla 8 se advierte que el 7,8 % resultado del complemento del 92,2 % de clientes malos han sido clasificados como buenos; esto puede generar una pérdida para el fondo al aumentar el riesgo de no recuperar el monto prestado. De igual forma, el 5,8 % resultado del complemento del 94,2 % de los clientes buenos han sido clasificados como malos, esto puede conducir a la pérdida de un cliente potencial que podría haber sido objeto de crédito para el fondo de empleados (costo de oportunidad).

$$f(p) = \frac{e^{(z)}}{1 + e^{(z)}} = \frac{1}{1 + e^{-(Z = -1,139 - 5,324 * Localización(1) - 1,363 * jurídico(1) - 4,909 * Empleo(1) - 0 * ingresos - 0 * Saldo.capital)}} \quad [8]$$

Además, la tabla 10 también permite observar en la tercera columna que los errores de las variables son bajos, menores a 1; por lo tanto, las variables son adecuadas para el modelo. De igual forma, se advierte en la sexta columna que las variables significativas fueron todas debido a que las variables en mención tienen una significancia menor a 5 % ($p < 0,05$), porcentaje utilizado para este estudio. Hay que resaltar que las variables ingresos y saldo a capital presentan "0" en sus coeficientes, pero al eliminarlas del modelo, este presenta un desempeño más pobre de predictibilidad; por lo tanto, se dejan en el modelo dichas variables.

En la tabla 10 también se evidencia que los coeficientes negativos de las variables enmarcan un OR menor a 1 como señalador de disminución del riesgo a medida que se incrementa la variable independiente; un coeficiente mayor que cero no influirá en la estimación de la probabilidad, mientras que un valor negativo originará una probabilidad cercana a cero. Lo anterior, debido al cambio de signo en la función logística. Con los coeficientes de las variables de la tabla 10, columna B, y los valores de las variables, se puede estimar la probabilidad de incumplimiento de cada deudor, mediante el modelo construido de regresión logística binaria. Finalmente, de la tabla 10 se pueden interpretar los OR de la siguiente forma:

Localización (1) Exp (B)= si el cliente está localizado adecuadamente en la dirección actual reportada en el fondo de empleados, se reduce el riesgo de impago en 95,5 %; por el contrario, se aumenta el riesgo de impago en doscientas veces con un cliente que no está localizado.

Jurídico (1) Exp (B)= si el cliente presenta procesos jurídicos por parte del fondo de empleados, se reduce el riesgo de impago en 74,5 %; por el contrario, se aumenta el riesgo de impago cuatro veces al no adelantarle al cliente procesos jurídicos.

Empleo (1) Exp (B)= si el cliente posee empleo se reduce el riesgo de impago en 99,30 %; por el contrario, se aumenta el riesgo de impago en ciento cuarenta y tres veces, cuando el cliente está desempleado o se queda sin empleo.

CONCLUSIONES

La adopción de un modelo econométrico como el propuesto en este estudio se convierte en una herramienta eficiente frente a los riesgos crediticios, ya que el modelo adquiere relevancia en la medida que ataca el riesgo crediticio de incumplimiento al clasificar los clientes entre buenos y malos. De acuerdo con el estudio de riesgos

en fondos de empleados, el riesgo de probabilidad de incumplimiento presenta alta incidencia, por lo tanto, se hace imprescindible mitigarlo con un modelo de estas características.

Al evaluar la situación de riesgos financieros crediticios de los fondos de empleados universitarios de Medellín y Área Metropolitana, se puede inferir la ineludible necesidad de controlar los riesgos y causas asociadas con el incumplimiento, que al no ser controlados, por ejemplo, con un modelo econométrico de predicción, inicialmente en la etapa de asignación del crédito, posteriormente se debería controlar y hacer un adecuado seguimiento mientras el crédito este activo; de lo contrario se podría incurrir en pérdidas considerables.

La financiación para los clientes mediante fondos de empleados se convierte en una opción muy interesante frente a la financiación tradicional, ya que generalmente los fondos presentan unas bajas tasas de interés, debido principalmente a que los excedentes se traducen en beneficios para los asociados, por ejemplo, en bajos costos por el capital recibido.

Los resultados obtenidos en este estudio muestran la importancia que representa el modelo de regresión logística binaria como modelo de predicción, como instrumento de apoyo para el análisis del riesgo de incumplimiento de compromisos de carácter crediticio, de los clientes hacia los fondos de empleados universitarios de Medellín y Área Metropolitana y para cualquier entidad que otorgue créditos.

En lo que respecta a la base de datos objeto del estudio, de los trescientos ochenta y cuatro casos estudiados solo doscientos cuarenta presentan una probabilidad de pagar sus compromisos, es decir, el 62,5 % de los deudores tiene la posibilidad de realizar pagos. Es recomendable que se haga un buen seguimiento de estos sujetos cumplidores para que en el futuro no incurran en impagos. De los trescientos ochenta y cuatro casos estudiados, ciento cuarenta y cuatro presentan una probabilidad de impago de sus compromisos, es decir, el 37,5 % de los deudores tiene la posibilidad de realizar pagos parciales o de no realizarlos. Es recomendable que se realice un buen seguimiento, se revisen los casos para establecer, por ejemplo, acuerdos personalizados más blandos de crédito, con el fin de evitar posibles pérdidas.

El modelo propuesto presenta un poder de predictibilidad correspondiente a un 93,5 %, por ello se puede considerar un modelo adecuado para predecir la probabilidad de incumplimiento de un sujeto de crédito.

Es fundamental mencionar que el modelo econométrico no reemplaza a los tomadores de decisiones crediticias. En general los modelos son instrumentos para ayudar en la toma de decisiones sin pretender reemplazar a nadie.

Finalmente, hay que mencionar que un modelo econométrico mediante regresión logística binaria es importante para predecir la morosidad o incumplimiento; sin embargo, los riesgos de una entidad deben igualmente gestionarse de forma adecuada, tomarse de manera integral y no parcializar los riesgos, de lo contrario la tarea se realiza de forma incompleta.

REFERENCIAS

- Alonso C., J. C. y Berggrun P., L. (2015). *Introducción al análisis de Riesgo Financiero* (3.^a ed.). ECOE Ediciones.
- Cardona Hernández, P. A. (2004). Aplicación de árboles de decisión en modelos de riesgo crediticio. *Revista Colombiana de Estadística*, 27(2), 139–151. <https://revistas.unal.edu.co/index.php/estad/article/view/28808>.
- Carrasco, S. (2005). *Metodología de la Investigación Científica*(1.^a ed.). Editorial San Marcos.
- Cox, D. y Snell, E. J. (1989). *Analysis of Binary Data* (2.^a ed.). Ed. Chapman and Hall/CRC.
- De Lara, A. (2013). *Medición y Control de Riesgos Financieros* (4.^a ed.). Limusa.
- Departamento Nacional de Estadística (DANE). (2018). *Mercado Laboral por Sexo*. https://www.dane.gov.co/files/investigaciones/boletines/ech/ech_genero/bol_ejesexo_feb18_abr18.pdf.
- Dunn, M. A. (2006). *Ultimate Credit and Collections Handbook* (1st First Edition). *Entrepreneur magazine* s.
- Elizondo, A. (Coord.) (2004). *Medición integral del riesgo de crédito*. Editorial Limusa.
- Fernández, H. y Pérez, F. O. (2005). El modelo logístico: una herramienta estadística para evaluar el riesgo de crédito. *Revista Ingenierías Universidad de Medellín*, 4(6), 55-75. <http://www.redalyc.org/articulo.oa?id=75040605>.
- García S., O. L. (2009). *Administración Financiera: Fundamentos y Aplicaciones*(4.^a ed.): Editorial Prensa Moderna Impresores S.A.
- Gisbert, A. (2017). *La gestión macro prudencial bancaria: una propuesta de tres modelos de previsión de riesgos* [tesis de doctorado, Universitat Abat Oliva CEU]. Repositorio Institucional Tesis Doctorals en Xarxa (TDX).
- Gómez, D. y López, M. (2002). *Riesgos financieros y operaciones internacionales*. ESIC Editorial.
- Gujarati, D. y Porter, D. (2010). *Econometría* (5.^a ed.). McGraw-Hill.
- Hernández, R., Fernández, C. y Baptista, P. (2014). *Metodología de la Investigación*. McGraw-Hill.
- Hosmer, D. y Lemeshow, S. (2000). *Applied Logistic Regression* (2nd ed.). John Wiley & Sons, Inc.

- Jimber, J. A. (2015). *Riesgo de crédito. Modelización econométrica* [tesis de doctorado, Universidad de Cordova]. Helvia: Repositorio Institucional de la Universidad de Córdoba.
- Llompart, M. (2007). *El salario: concepto, estructura y cuantía (1.ª ed.)*. Editorial La Ley.
- López, A. (2013). Modelo predictivo de riesgo de morosidad para créditos bancarios usando datos simulados. *Epistemus*, (14), 22-30.
- Menard, S. (2002). *Applied Logistic Regression Analysis (2nd ed.)*. Sage university papers series: Quantitative applications in the social sciences. Sage publicacions Ltd.
- Meneses, L. Á. y Macuacé, R. A. (2011). Valoración y riesgo crediticio en Colombia. *Revista Finanzas y Política Económica*, 3(2), 65-82. <https://www.redalyc.org/pdf/3235/323527256006.pdf>.
- Miranda, M. (2016). *Análisis de productos y servicios de financiación (1.ª ed.)*. IC Editorial.
- Muñoz, S. y Soler, P. (2018). Acuerdo sobre Basilea III: mayor claridad y estabilidad regulatoria para los bancos europeos. BBVA. <https://www.bbva.com/es/acuerdo-basilea-iii-mayor-claridad-estabilidad-regulatoria-bancos-europeos/>.
- Ocampo, E., Cabrera, A. y Ruiz, A. (2006). Pronóstico de bolsa de valores empleando técnicas inteligentes. *Tecnura*, 9(18), 57-66. <http://www.redalyc.org/articulo.oa?id=257021033007>.
- Pérez, J. (2017). La regresión logística como modelo de predicción del riesgo crediticio en las organizaciones de la economía social y solidaria. *Revista Ciencia Administrativa*, (2), 232-243. <https://www.uv.mx/iiesca/files/2018/03/23CA201702.pdf>.
- Rayo, S., Lara, J. y Camino, D. (2010). Un Modelo de Credit Scoring para instituciones de microfinanzas en el marco de Basilea II. *Journal of Economics, Finance and Administrative Science*, 15(28), 89-124. <http://www.scielo.org.pe/pdf/jefas/v15n28/a05v15n28.pdf>.
- Superintendencia Bancaria de Colombia [SBC] (2002). Circular externa 011 que modifica el Capítulo II de la Circular Externa No. 100 de 1995, referente a la Gestión del Riesgo de Crédito <https://www.superfinanciera.gov.co/descargas?com=institucional&name=pubFile16971&downloadname=ance011.doc>.
- Superintendencia de la Economía Solidaria [SES] (2017a). *Estadísticas de entidades del sector solidario*. <http://www.supersolidaria.gov.co/es/entidad/estadisticas>.
- Superintendencia de la Economía Solidaria [SES] (2017b). *Principales cuentas reportadas por entidades del sector solidario por categoría de entidad años 2002–2017*. www.supersolidaria.gov.co/sites/default/files/public/entidades/estadisticas_130718.xls.
- Toro, J. y Palomo, R. (2014). Análisis del riesgo financiero en las pymes – estudio de caso aplicado a la ciudad de Manizales. *Lasallista de investigación*, 11(2), 78-88.

ANEXOS

Tabla A1. Validación del Modelo. Resumen de procesamiento de casos

Casos sin ponderar ^a		N	Porcentaje
Casos seleccionados	Incluido en el análisis	73	100,0
	Casos perdidos	0	0,0
	Total	73	100,0
Casos no seleccionados		0	0,0
Total		73	100,0

Fuente: elaboración propia.

No se observan casos perdidos en el modelo para validación

Tabla A2. Pruebas ómnibus de coeficientes de modelo

		Chi-cuadrado	gl	Sig.
Paso 1	Escalón	58,275	5	0,000
	Bloque	58,275	5	0,000
	Modelo	58,275	5	0,000

Fuente: elaboración propia.

Se observa en la prueba de ómnibus del bloque 1 que el modelo en su validación sigue ajustando bien, al igual que el modelo original.

Tabla A3. Resumen del modelo

Escalón	Logaritmo de la verosimilitud -2	R cuadrado de Cox y Snell	R cuadrado de Nagelkerke
1	27,452 ^a	0,550	0,796

Fuente: elaboración propia.

El R cuadrado de Nagelkerke de la tabla del modelo para validación, muestra que los coeficientes de determinación que son explicados por este modelo, son muy adecuados, al igual que en el modelo original.

Tabla A4. Contingencia para la prueba de Hosmer y Lemeshow

Observado	<i>paga o no paga = no paga</i>		<i>paga o no paga = paga</i>		Total	
	Esperado	Observado	Esperado			
Paso 1	1	7	6,930	0	0,070	7
	2	6	6,347	1	0,653	7
	3	5	4,181	2	2,819	7
	4	1	1,684	6	5,316	7
	5	1	0,686	6	6,314	7
	6	0	0,129	7	6,871	7
	7	0	0,028	7	6,972	7
	8	0	0,011	7	6,989	7
	9	0	0,003	7	6,997	7
	10	0	0,001	10	9,999	10

Fuente: elaboración propia.

Tabla A5. Tabla de clasificación

Observado	Pronosticado				
	<i>paga o no paga</i>		Corrección de porcentaje		
	<i>no paga</i>	<i>paga</i>			
Paso 1	paga o no paga	no paga	17	3	85,0
		paga	2	51	96,2
	Porcentaje global				93,2

a. El valor de corte es 0,500

Fuente: elaboración propia.

Se observa en la tabla de clasificación, que el modelo en su validación sigue pronosticando adecuadamente.

Tabla A6. Variables en la ecuación

	B	Error estándar	Wald	gl	Sig.	Exp(B) Inferior	95 % C.I. para EXP(B) Superior		
Paso 1 ^a	Localización(1)	-4,473	1,777	6,332	1	0,012	0,011	0,000	0,372
	jurídico(1)	-2,695	1,911	1,990	1	0,158	0,068	0,002	2,856
	Empleo(1)	-4,212	1,689	6,219	1	0,013	0,015	0,001	0,406
	ingresos	0,000	0,000	10,941	1	0,001	1,000	1,000	1,000
	Saldo capital	0,000	0,000	4,824	1	0,028	1,000	1,000	1,000
	Constante	1,389	1,971	0,497	1	0,481	4,011		

a. Variables especificadas en el paso 1: Localización, jurídico, Empleo, ingresos, Saldo capital.

Fuente: elaboración propia.

Se observa que las variables utilizadas para el modelo en su validación siguen siendo representativas en la ecuación final, tal como en el modelo original.