

# **Modelo Scoring para el otorgamiento de crédito de las pymes**

**Andrea Valencia Echeverri**

*avalen27@eafit.edu.co*

Trabajo de grado para optar al título de Magíster en Administración Financiera

Asesor

Dr. Jimmy A. Saravia Matus

*jsaravia@eafit.edu.co*

**Universidad EAFIT**

**Escuela de Economía y Finanzas**

**Maestría en Administración Financiera**

**Medellín**

**2017**

## Contenido

1. Introducción.....	4
2. Marco conceptual .....	5
2.1 Contexto de las pymes en Colombia .....	5
2.2 Riesgo de crédito y cómo se mide .....	7
2.3 Modelos Scoring, qué son y para qué sirven .....	8
3. Método de solución .....	9
3.1 Metodología y variables seleccionadas .....	13
4. Análisis de resultados .....	19
5. Conclusiones.....	26
Referencias .....	27

## Lista de tablas

Tabla 1. Resumen análisis descriptivo variables .....	18
Tabla 2. Modelo Logit .....	19

## Lista de gráficos

Gráfico 1. Curva probabilidad acierta .....	22
Gráfico 2. Curva probabilidad plazo .....	23
Gráfico 3. Curva probabilidad ventas.....	24
Gráfico 4. Curva probabilidad activo .....	24
Gráfico 5. Curva probabilidad cupo .....	25

## **Resumen**

Los modelos Scoring desde los años 90 se encuentran ampliamente utilizados como métodos de predicción de la probabilidad de pago de personas y empresas, cuando acceden al sistema financiero. Existen varios métodos para el desarrollo de estos modelos, siendo los más utilizados a través del tiempo los econométricos, matemáticos y en la actualidad la inteligencia artificial; lo que ha permitido automatizar en muchas entidades financieras a nivel mundial los estudios de crédito, basándose en las variables más relevantes que determinan un buen comportamiento. Este proyecto se presenta como una forma de entender, a través de la construcción de un modelo Scoring, cuáles son las posibles causas más comunes de que una pyme pague oportunamente o no, la obligación adquirida. Esto permite predecir el comportamiento de pago y a su vez cumplir con las políticas de riesgo definidas en el sistema de administración de riesgo (SARC), el cual se encuentra regulado en Colombia por la Superintendencia Financiera y fundamentado en el marco de referencia de los acuerdos mundiales de Basilea. Todo esto conlleva a un equilibrio entre la búsqueda de maximización de utilidad de una entidad financiera, que en su mayoría es a través de la colocación de recursos, y el nivel de riesgo que se desea asumir.

## **Palabras clave**

Modelo Scoring, Pymes, Riesgo de crédito, Otorgamiento de crédito, Logit.

## **Abstract**

Scoring models since the 1990s are widely used as methods of predicting the probability of payment of individuals and companies when they access the financial system. There are several methods for the development of these models, being the most used over time econometric, mathematical and nowadays artificial intelligence; which has made it possible to automate in many financial institutions worldwide credit studies based on the most relevant variables that determine good behavior. This project is presented as a way of understanding through the construction of a scoring model, which are the most common possible causes for an SME to pay the obligation or not. This allows to predict the behavior of payment and in turn to comply with the risk policies defined in the risk management system (SARC); which is regulated in Colombia by the Financial Superintendence and based on the frame of reference of Basel Accords. All of this leads to a balance between the pursuit of utility maximization of a financial institution, which is mostly through the placement of resources, and the level of risk that one wishes to assume.

## **Key words**

Scoring models, SMEs, Credit risk, Credit granting, Logit.

## **1. Introducción**

Las pymes (pequeñas y medianas empresas) son uno de los motores principales de la economía y, por ende, son entes que acceden constantemente al sector financiero para apalancar su crecimiento. Para las entidades financieras es un segmento apetecido, que en los últimos tiempos ha obtenido un mayor acceso a productos financieros y, por esta razón, deben conocer el comportamiento de su cartera con el fin de controlar su riesgo de crédito y los indicadores de vencimiento. Este proyecto de grado está basado en construir un modelo Scoring (un sistema automatizado de calificación de crédito), a través de una metodología Logit, que permita tener herramientas para la toma de decisiones a la hora de otorgar créditos a este tipo de empresas. Es decir, conocer las principales variables que pueden definir la probabilidad de pago de una pyme al adquirir una obligación financiera. Este modelo se plantea como un complemento a los estudios de crédito actuales, permitiendo tener un segundo filtro en el otorgamiento de créditos. Para el planteamiento del modelo se utiliza la base de datos que fue aportada por una entidad financiera; esta base de datos es la información de los clientes pymes que tienen productos de crédito activos con esta entidad.

El objetivo primordial de este trabajo de grado es desarrollar un modelo Scoring para el segmento pyme, permitiendo de forma ágil y adecuada, el análisis de riesgo crediticio y así otorgar créditos con parámetros definidos. Adicional, a través del planteamiento de este proyecto, también se pretende comprender la dinámica de las pymes en Colombia y el porqué de su importancia en la economía del país y, a partir de este entendimiento, definir los principales aspectos que inciden en el comportamiento de pago de una pyme cuando acceden a créditos, teniendo en cuenta todo lo estipulado por la Superintendencia Financiera en la Administración de Riesgo de Crédito, y así proporcionar una herramienta que pueda ser utilizada como base para aplicar un modelo Scoring en un caso real.

En este documento se presenta en primer lugar el marco conceptual, en el cual se enuncia un panorama de la situación actual de las pymes en Colombia, se contextualiza sobre qué es el riesgo de crédito y qué implica para una entidad financiera, qué son los modelos Scoring

y para qué sirven. Después, se explica el método de solución donde se da una explicación de cómo se obtuvo la información y un detalle de lo que contiene, además de dar una breve ilustración de cuáles son los determinantes de crédito de las pymes y cómo se hace actualmente el estudio de crédito. Posteriormente, se plantea la metodología que será utilizada para el desarrollo del modelo y una explicación ampliada de las variables que fueron seleccionadas y el porqué. A continuación se presentan los resultados y se hace un análisis respectivo de éstos. Por último, teniendo en cuenta los resultados se exponen las conclusiones del trabajo.

## **2. Marco conceptual**

### **2.1 Contexto de las pymes en Colombia**

Las pymes, entendidas como pequeñas y medianas empresas en Colombia, según el artículo 2° de la Ley 905 de 2004, son toda unidad de explotación económica, realizada por persona natural o jurídica, en actividades empresariales, agropecuarias, industriales, comerciales o de servicios, rural o urbana; las cuales se clasifican por sus activos totales y el número de empleados de la siguiente forma:

1. Mediana empresa: a) planta de personal entre cincuenta y uno (51) y doscientos (200) trabajadores, o b) activos totales por valor entre cinco mil uno (5.001) a treinta mil (30.000) salarios mínimos mensuales legales vigentes.
2. Pequeña empresa: a) planta de personal entre once (11) y cincuenta (50) trabajadores, o b) activos totales por valor entre quinientos uno (501) y menos de cinco mil (5.000) salarios mínimos mensuales legales vigentes o,
3. Microempresa: a) planta de personal no superior a los diez (10) trabajadores o, b) activos totales excluida la vivienda por valor inferior a quinientos (500) salarios mínimos mensuales legales vigentes.

Este sector, según cifras recientes en el país, genera 80,8% del empleo del país, de las cuales el 39,9% son sociedades y el restante 60,1% son personas naturales (Revista Dinero, 2016). Por tal razón, se convierten en jugadores indiscutibles en el desarrollo económico del país, en la generación de empleo, en la productividad y en la competitividad.

A pesar del rol primordial que presentan en la economía del país, también se presenta el lado no tan amable, como la falta de apoyo del Estado y de la inexistencia de políticas que permitan su crecimiento, la falta de acceso a fuentes de financiación, el inadecuado uso de tecnologías avanzadas y la alta tasa de mortalidad.

Los beneficios que proporcionan a la sociedad la creación de empresas, se ve afectado por la tasa de mortalidad de las empresas en sus primeros años de vida. Los datos al respecto muestran que la mayoría de las pymes viven por corto tiempo; dentro de los primeros cinco años de vida, cerca de la mitad de las iniciativas empresariales fracasan (Franco y Urbano, 2010).

A pesar de la situación descrita anteriormente, en Colombia las pymes se convirtieron en un atractivo segmento para atender por parte de la mayoría de instituciones financieras, debido a su gran potencial de crecimiento, el gran aporte a la economía, tener mayor participación en los mercados y por la rentabilidad que se puede obtener. Por otro lado, las pymes también deben recurrir al sistema financiero para obtener el apalancamiento que les demanda su negocio, el crecimiento natural o su expansión.

La financiación nace del actuar cotidiano de las empresas, las cuales en su operación demandan recursos para cumplir con sus obligaciones de corto y largo plazo. Estos recursos se pueden obtener de fuentes internas y externas. Las fuentes internas se componen del capital aportado por socios, del flujo de caja generado por el negocio o reinversión de las utilidades. Pero esta situación no se presenta de forma recurrente, ya que las necesidades de caja del negocio solicitan apalancamiento frecuente que no se suplen con el ciclo de caja; es así como se necesita de fuentes externas de financiación que se encuentran en el sector

financiero. Se ha asociado que el mayor acceso al financiamiento externo conlleva a mayores niveles de crecimiento y desarrollo económico (Botello, 2015).

Por todas estas razones, las pymes se convierten en ese sector que debe ser examinado con criterios muy definidos y claros por el riesgo que implican, pero también por lo apetecido que se convierten por su dinámica, por sus necesidades recurrentes y, por ende, por los márgenes importantes que se pueden obtener. No obstante, las entidades financieras antes de tomar alguna decisión de financiación deben estipular y cumplir con sus políticas de riesgo de crédito.

## **2.2 Riesgo de crédito y cómo se mide**

Riesgo de crédito se define como la posibilidad de que una entidad incurra en pérdidas y se disminuya el valor de sus activos, como consecuencia de que un deudor o contraparte incumpla sus obligaciones<sup>1</sup>. El Riesgo de crédito se encuentra regulado en Colombia por la Superintendencia Financiera y está estipulado a través de la Circular Externa 100 de 1995 y la Circular Externa de 2002, donde todas las entidades vigiladas deben cumplir con el establecimiento de un Sistema de Administración de Riesgo, con el fin de evitar poner en peligro la situación de solvencia y liquidez. Por tanto, las entidades financieras deben cumplir con un adecuado sistema de administración de riesgo de crédito y definir unas políticas claras al respecto. Dentro de todo este sistema, se incluyen las diferentes fuentes de financiación, las garantías que se deben exigir, las provisiones que deben tener, el tipo de créditos a financiar, las diferentes variables que van a ser estudiadas por cada segmento que vaya ser atendido (personas naturales, pymes, empresas grandes) y el desarrollo de modelos para el otorgamiento de créditos.

---

<sup>1</sup> Definición de Circular Externa 100 de 1995 de la Superintendencia Financiera de Colombia.

### 2.3 Modelos Scoring, qué son y para qué sirven

Una de las herramientas más utilizadas en los últimos años para la medición y control del riesgo son los modelos Scoring. Los modelos de Credit Scoring, también denominados *scored-cards* o *clasifiers*, son “métodos estadísticos utilizados para clasificar a los solicitantes de crédito, o incluso a quienes ya son clientes de la entidad evaluadora, entre las clases de riesgo bueno o malo” (Hand y Henley, 1997, p. 535).

La utilización de modelos Scoring para la evaluación de riesgo de crédito (probabilidades de default y ordenar los deudores y solicitantes en función de su riesgo de incumplimiento) comenzaron en la década de los años 70's, pero se popularizaron a partir de los 90's, gracias al desarrollo de mejores recursos estadísticos y computacionales y de la creciente necesidad por parte de la industria bancaria de ser más eficaz y eficiente en el estudio de créditos. No obstante, hoy en día continúa siendo utilizado el juicio humano en el origen de créditos; el cual se puede definir como un conjunto de reglas que la entidad aplica de manera sistemática para atender solicitudes de crédito. Actualmente en Colombia se demuestra que ambas metodologías coexisten y se complementan, definiendo sistemas híbridos (Gutiérrez, 2007).

Los modelos Scoring están basados en técnicas estadísticas, matemáticas, econométricas y de inteligencia artificial. Los modelos más usados son: análisis discriminante, regresión lineal, regresión logísticas, modelos Probit, modelos Logit, métodos de programación matemáticas, árboles de decisión y redes neuronales; éstos últimos son los que se han desarrollado y se han comenzado a utilizar a partir de la década de los años 90's.

Muchos de los métodos conducen a una *scorecard*, donde las características de los clientes reciben un puntaje y la suma de los puntajes determina el riesgo del cliente de ser un mal o buen cliente. Otras técnicas no conducen a una *scorecard*, pero en su lugar indican directamente la probabilidad de que el cliente sea bueno y valga entonces la pena aceptar la cuenta.

Dentro de los modelos econométricos, los modelos de probabilidad lineal han dejado de usarse por sus desventajas técnicas. Así, los modelos Probit y Logit son superiores, ya que proveen para cada deudor una probabilidad de default. La ventaja de estos métodos es que se pueden usar las propiedades de los estimadores y las herramientas de los intervalos de confianza y testeo de hipótesis. Se puede conocer la importancia relativa de las diferentes variables, permitiendo así descartar características irrelevantes y asegurar las importantes.

Es importante resaltar que los modelos Scoring en cualquier caso de las metodologías mencionadas anteriormente, emplean información del evaluado, contenidas en la información que provee o fuentes internas y externas. En Colombia, una de las fuentes externas más destacables son las centrales de riesgo. En Colombia existen dos: CIFIN y Datacrédito. Cada una de éstas da una calificación o *score* del cliente según el comportamiento pasado que ha tenido en el sector financiero, es decir, no realizan la evaluación de crédito, simplemente recopilan información que otorgan a las instituciones financieras. Adicional, en la CIFIN se incluyen también el comportamiento que presenta en el sector real.

Después de haber desarrollado el contexto de este trabajo, a continuación se va a presentar cuál fue el modelo escogido para el planteamiento del presente caso, la información utilizada y la base de datos.

### **3. Método de solución**

Teniendo en cuenta la popularización de los modelos Scoring, la inquietud en este trabajo es presentar un modelo que permita entender el comportamiento de pago de las pymes. Se eligen las pymes porque son un segmento que está creciendo en bancarización, pero también por los factores mencionados anteriormente, se convierten en cierta forma en alto riesgo para el sector financiero. Adicional, en algunas entidades ya se encuentran desarrollados estos modelos para la banca personal, pero no incluyen a las pymes. A estas empresas les realizan un estudio de crédito basado en ciertas políticas que debe cumplir

según perfil de riesgo de cada entidad, con el respectivo análisis de sus estados financieros, razones financieras, comportamiento pasado en la entidad, si ya es cliente, y tomando datos de centrales de riesgo. Por tanto, este modelo servirá de herramienta adicional al análisis que se realiza actualmente, para que sirva como un segundo filtro en la toma de decisiones.

Para el desarrollo del modelo se utilizan como fuente primordial una base de datos facilitada por un banco de mediano tamaño, que por razones de seguridad y privacidad no se revelará su nombre, pero para motivos de este trabajo será llamada Entidad Financiera.

Esta base es de los datos actuales de las pymes que tienen o han tenido crédito con el banco. Las pymes en la entidad financiera se clasifican como aquellas empresas que tienen ventas entre 2.000 a 25.000 millones al año. La base de datos está constituida por más de 7.000 registros de clientes que tienen obligaciones activas y canceladas. Para efectos de estudio se toma una muestra de 690 registros de clientes pymes, la cual incluye personas naturales y empresa a quienes ya finalizaron el plazo del crédito acordado inicialmente y así se puede ver qué pasó con el comportamiento de pago durante el tiempo estipulado. La razón por la cual se toman los cancelados es que permite identificar cuáles clientes lograron pagar de forma oportuna al banco o quiénes de esos clientes durante el tiempo establecido no tuvieron un buen comportamiento de pago. Los créditos que presenta la base son de diferentes tipos y modalidades: créditos a largo y corto plazo, algunos son destinados específicamente a capital de trabajo u otros son otorgados para apoyar el crecimiento en infraestructura y tecnología. Para cada registro la base presenta 63 variables, incluyendo cualitativas y cuantitativas (especialmente razones financieras que explican y son indispensables para entender la capacidad de pago de un cliente ante una obligación que vaya a obtener). Más adelante, se explicarán las variables que de esta base fueron tomadas para desarrollar el modelo y el porqué. Antes de eso, quisiera presentar un antecedente de cómo se analiza una operación de crédito a través de juicio humano, con el fin de aplicar dentro del modelo algunos estándares que son importantes al evaluar el riesgo de crédito de una pyme.

Estas evaluaciones están basadas en la experiencia de gerentes de crédito y no sólo toman en cuenta información histórica, sino que tratan de realizar proyecciones de la probable situación futura de los prestatarios y de su capacidad de pago del crédito. La información clave es obtenida a través de la relación con el cliente y de la experiencia de la institución financiera con éste (Dabos, 2007).

Según Botello (2015) existen unos determinantes para la financiación de una pyme, entre los cuales se encuentran:

### **1. Tamaño de la empresa**

Posee una relación directa con la posibilidad de que las empresas puedan acceder a los mercados de financiamiento, vía mayores recursos y garantías. Adicional, muestra qué capacidad de capital disponible posee como respaldo.

### **2. La antigüedad**

La experiencia o el tiempo que la empresa ha operado en los mercados brindan confianza a las instituciones financieras.

### **3. Características de las organizaciones**

Las características propias de la dirección estratégica de la empresa, como la cualificación de la dirección y la concentración accionaria de la empresa, tienen un efecto positivo sobre el acceso al crédito.

### **4. Diferencias sectoriales y localización**

Se encuentran diferencias en el nivel de acceso al crédito según localización geográfica y el sector económico al que pertenece. Se incluyen en este apartado las condiciones económicas que presente cada sector.

## 5. Situación financiera

Ésta se convierte en uno de los determinantes más relevantes, ya que a través de los estados financieros (Balance General y Estado de Resultados), se puede entender la situación financiera de una pyme. Los estados financieros presentados como tal no logran resumir o mostrar la realidad, para esto se les hace su respectivo análisis financiero a través del cálculo de indicadores financieros, análisis vertical y horizontal.

El análisis vertical permite relacionar todas las cuentas de los estados financieros con respecto a una determinada cifra base.

El análisis horizontal permite comparar los resultados de las cuentas de los estados financieros de un período a otro de la misma cifra.

Por último, los indicadores financieros son razones que relacionan dos magnitudes que permiten hacer comparaciones en un momento dado y así analizar la situación financiera de la empresa con una visión global y sintetizada (Romero, Melgarejo, Vera, 2015). Existen tres grupos principales de indicadores: de liquidez, rentabilidad y endeudamiento.

- *Indicadores de Liquidez:* relacionan los activos y pasivos de corto plazo. Con estos indicadores se pretende entender la capacidad de una empresa de cumplir con sus obligaciones de corto plazo mediante la conversión de sus activos corrientes. Para un estudio de crédito se convierte en uno de los indicadores fundamentales, ya que permite entender el ciclo de caja de un negocio. Esto se relaciona también con la capacidad con la que puede asumir las deudas de corto plazo. Los más importantes son: capital de trabajo, razón corriente y prueba ácida.
- *Indicadores de Rentabilidad:* presentan la efectividad de la administración para convertir ventas y la inversión en utilidades, miden el beneficio o productividad de los fondos comprometidos en la empresa. Los más destacados a la hora de un

análisis de crédito son: margen bruto, margen operacional, margen neto, rentabilidad del patrimonio (ROE) y rentabilidad del activo (ROA).

- *Indicadores de Endeudamiento*: mide el grado de financiación con pasivos externos y la participación de los acreedores en la estructura financiera, determinando así el nivel de riesgo. Los más destacados son: endeudamiento total, carga financiera, endeudamiento financiero, cubrimiento de intereses, la deuda analizada sobre el Ebitda (Earnings Before Interest, Taxes, Depreciation and Amortization).

Es de relevancia ampliar el concepto de Ebitda y por qué se convierte en un indicador indispensable para el análisis de crédito. Se define como la capacidad de una empresa de generar beneficios considerando únicamente su actividad productiva. Para un análisis de crédito, se convierte en la caja real que produce una empresa durante el tiempo estudiado y se compara con el valor del crédito otorgado. Es decir, el Ebitda debe tener la capacidad de cubrir el endeudamiento que se va a otorgar.

### **3.1 Metodología y variables seleccionadas**

Teniendo en cuenta todas las referencias bibliográficas estudiadas y trabajos anteriores realizados de Credit Scoring, tales como (Gutiérrez, 2007; Moscote y Rincón, 2012; Ochoa, Galeano y Agudelo, 2010; Rayo, Lara y Camino, 2010; Saravia-Matus y Saravia-Matus, 2012); se puede concluir que el modelo Logit es el que mayores ventajas posee, ya que por ser una función logística, desde el punto de vista matemático, es una función flexible y fácil de utilizar y la evidencia empírica ha demostrado que este modelo es adecuado en la mayoría de los casos en los cuales la respuesta es binaria.

Se refiere a respuesta binaria cuando hablamos que la variable dependiente toma sólo dos valores 0 y 1. Para el caso nuestro la variable dependiente será No Pago, que toma valor 1 cuando existe mora de más de 30 días o 0 si la mora es menor de 30 días.

El modelo Logit permite obtener estimaciones de la probabilidad de un suceso e identificar los factores de riesgo que determinan dichas probabilidades.

La modelización Logit es similar a la regresión tradicional, salvo que utiliza como función de estimación la función logística, lo que permite por un lado obtener la estimación de probabilidad, pero al tratarse de un análisis de regresión, permite identificar las variables más relevantes.

En este caso se va a elegir un modelo Logit dicotómico, es decir, cuando el número de alternativas son dos y son excluyentes entre sí. El modelo presenta las siguientes características: la variable dependiente es binaria y toma sólo dos valores y las variables independientes, las cuales pueden ser cualitativas o cuantitativas y al final arroja un vector de parámetros con valores numéricos, que son los coeficientes para cada una de las variables explicativas. De aquí se puede obtener que a cada valor del vector de parámetros le corresponde una variable explicativa, al tenerse en cuenta todas en conjunto y dar valores a cada una de las variables independientes contenidas en el modelo definitivo, se obtiene el valor de la probabilidad de que un individuo posea la característica de interés estudiada en el modelo.

El modelo a desarrollar se puede definir así:

**Ecuación 1. Modelo Logit**

$$\Pr(\text{nopago} = 1 | \mathbf{x}) = \frac{\exp(\mathbf{x}\beta)}{1 + \exp(\mathbf{x}\beta)}$$

Desarrollado el contexto anterior, el modelo de este trabajo y según todas las ventajas descritas será el Modelo Logit, el cual será realizado a través de la herramienta Stata.

El modelo planteado queda así:

## Ecuación 2. Modelo Scoring

$$\begin{aligned} x\beta = & \beta_0 + \beta_1 \text{plazo} + \beta_2 \text{plazo}^2 + \beta_3 \text{cupo} + \beta_4 \text{ventasnetas} + \beta_5 \text{activototal} + \\ & \beta_6 \text{aciertaplus} + \beta_7 \text{missingAP} + \beta_8 \text{pruebaacida} + \beta_9 \text{cubrimientodeinteresesveces} \\ & + \beta_{10} \text{razonendeudamiento} \end{aligned}$$

Más arriba se mencionó que la base posee 63 variables, que incluyen diferentes características a la hora de analizar una evaluación de crédito de la pyme.

Para efectos de este trabajo, se van a elegir las variables que logran ser más relevantes para entender el comportamiento de pago, logrando abordar los determinantes de la financiación de las pymes.

### - Variable dependiente:

Es una variable dicótoma, definida como no pago. Esta variable fue generada en Stata teniendo en cuenta una variable aportada en la base, en la que cada cliente en el momento específico menciona cuántos días de mora llegó a tener. La variable toma valor 1 si los días de no pago es mayor o igual a 30 días y toma valor cero cuando paga menor a 30 días. Se elige mora mayor a 30 días, teniendo en cuenta que cualquier empresa que tengas moras mayor o igual a 30 días será reportado por las instituciones financieras a centrales de riesgo, y adicional comienza a generar alertas del comportamiento de la pyme.

### - Variables independientes:

**Acierta plus:** es la calificación de Datacrédito que tiene cada pyme con el banco al momento que es presentada la base. En este caso puede tomar valores desde puntajes de 400 hasta 900. Con respecto a esta variable se debe realizar una aclaración, ya que no todos los clientes analizados tenían la calificación, pero por la importancia se genera una variable que se denomina MissingAP, que toma los valores de los cuales no tienen datos y éstos se sacan de la variable acierta plus. Para un modelo Scoring esta es una variable relevante, en

tanto esta calificación resume el comportamiento del cliente en el sector financiero, cómo ha pagado o si ha sido reportado por mal hábito de pago. Como más adelante se explicará, para el modelo desarrollado, se convirtió en una de las variables con mayor impacto.

**Plazo:** ésta se convierte en una variable indispensable a la hora de otorgar un crédito, ya que implica el tiempo establecido en el cual el solicitante va a pagar la deuda. Para las pymes es importante entender el plazo versus su información financiera, porque indica según la caja del cliente, el Ebitda (Earnings Before Interests, Taxes, Depreciation, Amortization) y otros, si tiene la capacidad de asumir el plazo. Los plazos pueden ser corto o largo plazo. En la base se presentan plazos desde 12 meses hasta 84 meses. Estos plazos también indican en muchos casos el tipo de operación a tomar, es decir, si es capital de trabajo o créditos para inversiones de crecimiento, infraestructura o demás.

**Cupo:** se toma la variable cupo, ya que indica el monto total que es otorgado y la exposición global de riesgo. El cupo también va relacionado con la capacidad de pago, pues indica si el valor solicitado es capaz de pagarlo o no. Tiene mucho que ver con el tamaño de la empresa y su situación financiera como ventas y caja.

**Activo total:** se elige esta variable porque muestra el tamaño de una empresa. Este es uno de los determinantes fundamentales a la hora de evaluar una pyme, ya que muestra la solidez financiera. Adicional, dentro de los activos se evidencia la capacidad operativa y funcional de una empresa y representan el valor de una compañía.

**Ventas netas:** es el rubro financiero que indica los ingresos que obtiene una empresa proveniente de la actividad que desarrolle. Al igual que los activos, ayuda a entender qué capacidad de servicio a la deuda versus ingresos puede obtener y, por ende, si es capaz de pagar.

**Prueba ácida:** cuando se mencionaron los determinantes, uno de los puntos que hoy más se analiza es la situación financiera. Para entender la situación financiera, una de las herramientas que permite una visión global son las razones financieras y sus diferentes tipos. En este caso se aborda el indicador de liquidez a través de la prueba ácida. La prueba ácida se entiende como la capacidad de la empresa para cancelar sus obligaciones corrientes sin contar con la venta de sus inventarios.

**Prueba ácida = (Activo Corriente - Inventario) / Pasivo Corriente**

**Razón endeudamiento:** esta es una de las razones más fundamentales a la hora de asumir el riesgo con una pyme, puesto que muestra cuántos de sus activos están financiados con deuda. Un alto endeudamiento por lo general implica más riesgo. Hay variaciones del endeudamiento como observar el endeudamiento financiero, el cubrimiento de interés con el Ebitda. En este caso se va a tomar como **Pasivo Total / Activo Total**.

**Cubrimiento intereses:** esta es una razón que indica el número de veces que una empresa puede pagar los intereses de la deuda con el Ebitda. Esta se incluye como índice adicional para entender los indicadores de endeudamiento y ser más específicos en la capacidad de asumir el pago de la carga financiera que asume.

A continuación se anexa la tabla con el resumen estadístico de las variables seleccionadas:

**Tabla 1.** Resumen análisis descriptivo variables

<b>Variable</b>	<b>Unidad de Medida</b>	<b>Número de observaciones</b>	<b>Promedio</b>	<b>Desviación estándar</b>	<b>Mínimo</b>	<b>Máximo</b>
No Pago	N/A	689	0,08127	0,273	0	1
Plazo	Meses	689	52,6865	43,05594	1	240
Cupo	Millones	689	360,8946	605,1986	0.00034	7000
Ventas netas	Millones	662	7614,27	13224,6	0	120593,6
Activos total	Millones	662	10836,7	31691,05	97,2	351239,1
Prueba ácida	Veces	662	2,733012	15,83692	0	201,7329
Cubrimiento de Intereses	Veces	662	3,387464	13,98954	-132,037	150
Razón Endeudamiento	Veces	662	0,62641	0,59037	0	9,9057
Acierta Plus	Puntaje	168	749,77	151,78	150	904

Fuente: Elaboración propia, 2017.

La tabla 1 resume el análisis descriptivo de todas las variables dependientes e independientes. Adicional, se menciona por variable el número de observaciones que se tuvieron en cuenta y la unidad de medida. Importante aclarar que la variable Acierta Plus toma como referencia 168 observaciones, debido a que no todas tenían incluido este indicador. No obstante, se elige esta variable, como fue explicado más arriba, por la importancia que arrojó el modelo. Con estos valores de la media, y los máximos y mínimos, se va a realizar y mostrar más adelante el comportamiento de las probabilidades para valores de las variables.

Es importante aclarar que para efectos del desarrollo de este trabajo se eligieron éstas variables como las más relevantes. No obstante, en la práctica se analizan de forma detallada otras. Es por esta razón que esta herramienta se propone como un apoyo a un área de crédito y se puede utilizar como el segundo filtro de la primera decisión revisada a través de juicio humano.

Existen otras variables en la base que eran importantes, pero cuando se realizó el análisis en Stata arrojó como resultado variables no significativas que no explicaban el no pago. Se analizaron otras como: edad, la situación con ésta es que la base tiene combinación de personas naturales que se les asignaba edad, pero también hay empresas constituidas las cuales no tenían edad; no obstante, se trató de generar pero el resultado final era no significativa. La misma situación sucedió con el sector que se agrupó en los macrosectores, pero en realidad no logró la relevancia para explicar el comportamiento de pago.

#### 4. Análisis de resultados

Después de explicar el porqué y el cómo del desarrollo del modelo, se mostrará de forma detallada cómo se realizó en la herramienta y las debidas interpretaciones. Como ya se ha explicado, se realizó a través de un modelo Logit, como se muestra en la Tabla 2. Se tomaron en cuenta 662 observaciones.

**Tabla 2. Modelo Logit**

<b>Logistic Regression</b>	Number of obs = 662 LR Chi2 (10) = 54.47 Prob > chi2 = 0 Pseudo R2 = 0.1495			
Log likelihood = -154.95462				
<b>No Pago</b>	<b>Coef</b>	<b>Std Err</b>	<b>z</b>	<b>P &gt; z</b>
<b>Plazo</b>	0.35145	0.15450	2.27	0.023
<b>Plazo2</b>	0.00021	0.00011	1.79	0.073
<b>Cupo</b>	0.00005	0.00033	0.17	0.868
<b>Ventas netas</b>	0.00031	0.00001	2.06	0.040
<b>Activo Total</b>	0.00006	0.00002	2.39	0.017
<b>Acierta Plus</b>	0.01130	0.00036	3.09	0.002
<b>Missing AP</b>	558344	169162	3.30	0.001
<b>Prueba acida</b>	0.02306	0.04030	0.57	0.567
<b>Cubrimiento de intereses</b>	0.00926	0.01160	0.80	.0425
<b>Razon Endeudamiento</b>	0.17667	0.32171	0.55	0.583
<b>Constante</b>	2.71358	1.79753	1.51	0.131

Fuente: Elaboración propia, 2017.

Como se muestra en la Tabla 2, el modelo arroja un coeficiente para cada variable independiente, que en conjunto podrá explicar la probabilidad de pago del sujeto que se analice. El signo del coeficiente se entiende como la relación que tiene cada variable independiente con no pago; es decir, si es positivo significa que tienen una relación directamente proporcional y si es negativo tiene una relación inversamente proporcional. Adicional, el modelo Logit por ser estadístico, permite obtener la significancia estadística de cada variable con la dependiente. La significancia estadística es dada por  $P > |z|$  y se interpreta que si la probabilidad de cometer un error de tipo I (aceptar un falso positivo) es menor a 5% es una variable estadísticamente significativa; si la probabilidad de cometer un error de tipo I es mayor 5% es una variable no significativa. No obstante, en este modelo a pesar que algunas variables no obtuvieron el nivel de significancia, se dejan dentro del modelo teniendo en cuenta que en la realidad son muy importantes para evaluar una solicitud de crédito, y además suele pasar que necesitan una muestra de mayor número de datos para lograr ser explicativas. Más adelante serán utilizados los coeficientes para dar la probabilidad total de no pago.

Ampliando información de cada variable se hacen las siguientes explicaciones:

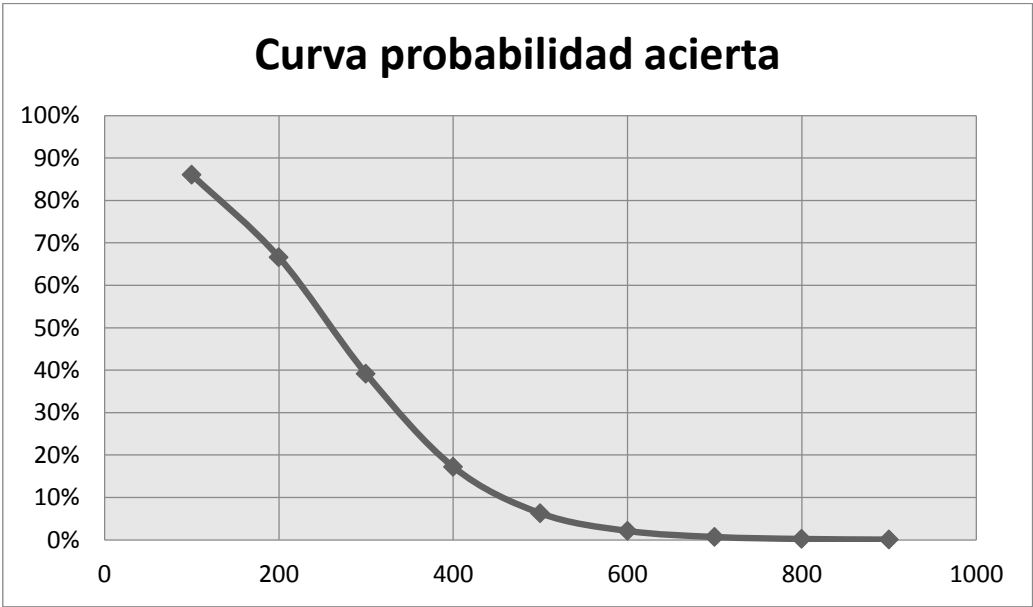
- Se aclara que además de la variable plazo, se genera y se incluye una llamada  $\text{plazo}^2$ , (plazo al cuadrado). El coeficiente de plazo es positivo, lo cual indica que a mayor número de meses que se otorgue un crédito aumenta la probabilidad de no pago y viceversa. Esto es muy razonable, ya que se asume mayor riesgo por parte de una entidad financiera porque se demora más en obtener el capital otorgado. El  $\text{plazo}^2$  tiene comportamiento negativo, lo que indica que el efecto de esta variable en la probabilidad de no pago tiende a mitigarse a medida que aumenta el plazo.
- El cupo, a pesar que no es una variable estadísticamente significativa en el modelo, se incluye porque para cualquier evaluación la exposición de crédito, comparando con su situación financiera, define si puede o no pagar. El coeficiente es negativo, lo que indica que a mayor cupo otorgado la probabilidad de no pago disminuye y viceversa.

- Ventas netas: en el modelo resulta ser una variable significativa, ya que su probabilidad de cometer un error de tipo I es menor a 5%. En realidad, es uno de los temas siempre a tener en cuenta, en tanto son los ingresos que genera una pyme producto del desarrollo de su negocio productivo. El signo es positivo lo cual indica que, manteniendo fijo el tamaño de la empresa (ver variable total de activos abajo), a medida que una empresa incrementa sus ventas, la probabilidad de no pago aumenta. Esto probablemente porque la empresa tiene un mayor volumen de activos por cobrar, lo cual limita su liquidez y consecuentemente su capacidad de pago.
- Activos totales: resultó ser también una variable significativa que va muy relacionada con el tamaño y la robustez de una pyme. El coeficiente es negativo, lo que indica que a mayor tamaño de la empresa menor probabilidad de no pago.
- Acierta plus: el modelo nos arroja como una de las variables más importantes. Es significativa y con un coeficiente negativo, es decir, que a un mayor *score* de Datacrédito (lo que significa un mejor comportamiento financiero) menor es la probabilidad de no pago. Es un indicador de suma importancia en una evaluación de crédito, ya que permite entender con un comportamiento antiguo, la posibilidad que sea buen pagador.
- Prueba ácida, cubrimiento de intereses y razón de endeudamiento: estas tres razones financieras resultaron en el modelo no significativas. Se van a dejar dentro de éste, puesto que permiten de forma global entender cómo está la situación de una empresa. Tienen coeficiente negativo. Se aclara que el signo negativo de la razón de endeudamiento no es un comportamiento a lo esperado, pero puede suceder por cuestiones de esta muestra, pero no se espera que sea igualmente negativo en otras muestras.

A continuación se anexan unos gráficos del comportamiento de la probabilidad de las principales variables. En cada variable se hace una comparación teniendo en cuenta la variable acierta y cuando no se tiene en cuenta. Esto con el fin de demostrar que en el modelo planteado la variable acierta, que corresponde al *score* o puntaje de Datacrédito, es

la más significativa y con mayor peso en la probabilidad de pago o no de una pyme de un crédito otorgado.

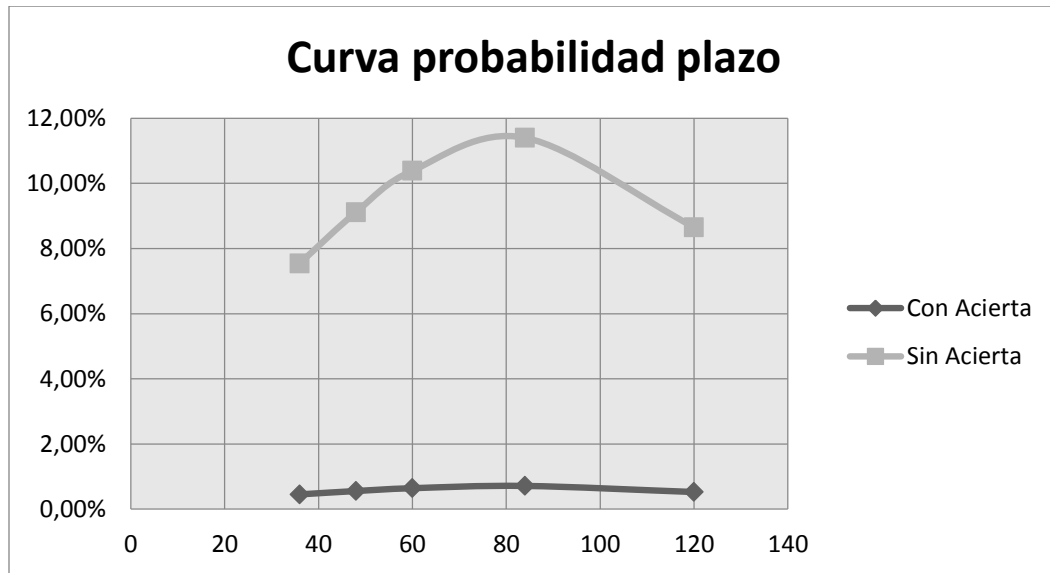
**Gráfico 1.** Curva probabilidad acierta



Fuente: Elaboración propia, 2017.

La curva de probabilidad acierta muestra el comportamiento de esta variable, tomando diferentes *scores* o puntajes, teniendo todas las demás variables de la ecuación constantes. Importante aclarar que cuando se habla de constante es que no se modifican y los valores que se le asigna corresponde a la media de cada una. Este gráfico nos muestra la relevancia que toma acierta en el modelo, pues se convierte en un indicador de riesgo indispensable, a la hora de entender la forma como se comporta una pyme en el pago de una obligación adquirida. A medida que el *score* es muy bajo (a partir de 400 hacia abajo), se evidencia un aumento muy drástico en la probabilidad de no pago, llegando a niveles hasta del 80% para *score* de 100. Es decir, que el comportamiento pasado de un cliente en el sector financiero se convierte determinante para el otorgamiento de créditos a pyme.

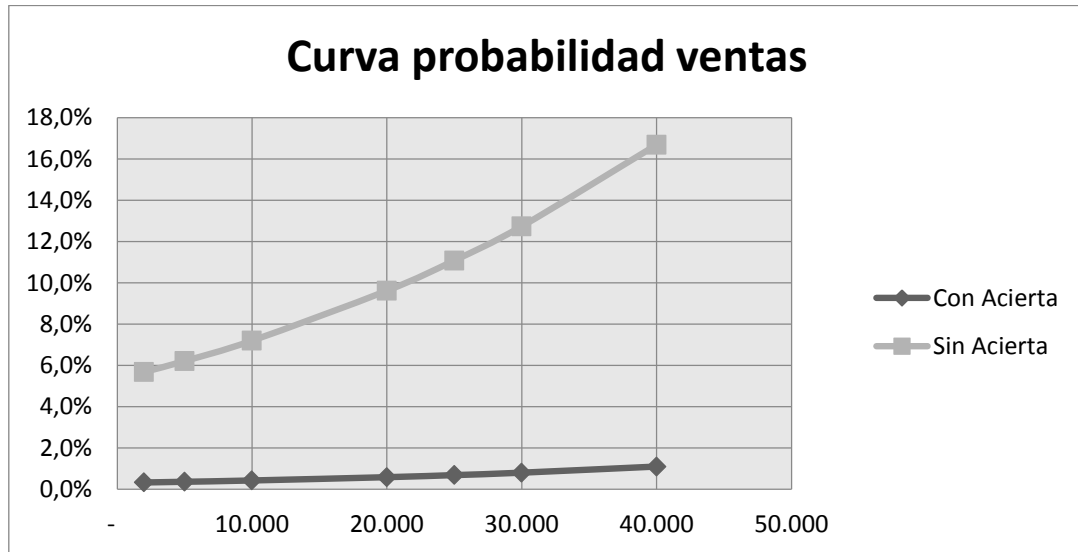
**Gráfico 2.** Curva probabilidad plazo



Fuente: Elaboración propia, 2017.

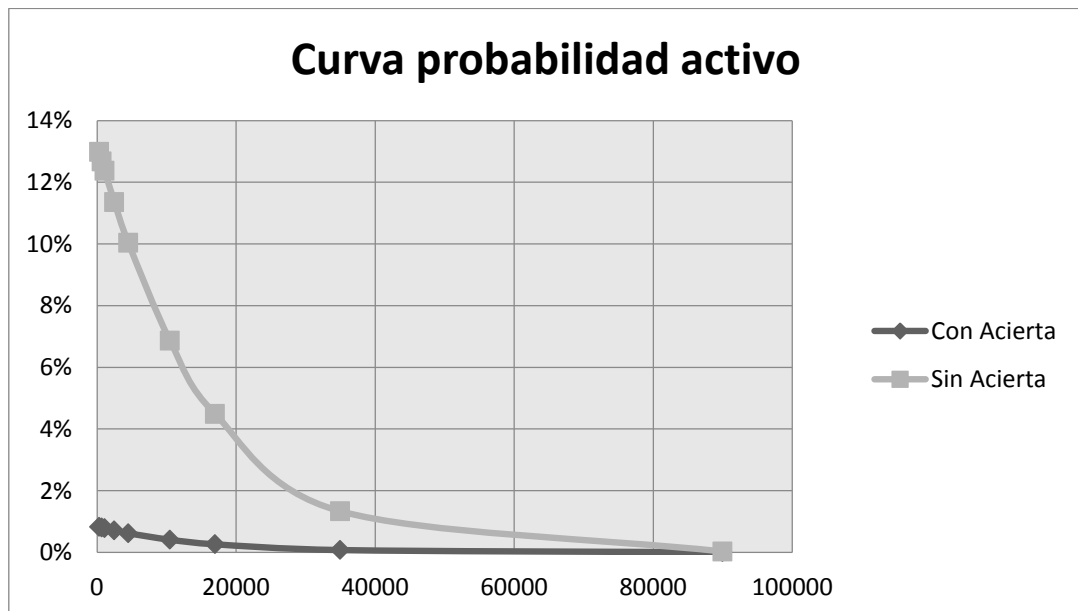
La curva de probabilidad de la variable plazo, se analiza de dos formas: 1). Teniendo en cuenta la variable acierta, a la cual se le asigna su valor promedio de 749 puntos. En este caso se obtienen para los diferentes plazos dados, unas probabilidades bajas que no superan el 1%. 2). Se analiza asignando el valor cero a la variable acierta plus y asignando a la variable missingAP el valor de 1 (es decir, los valores para estas variables en el caso de las empresas pymes para las cuales no se cuenta con un *score* de Acierta Plus) y las demás variables se ingresan con sus valores promedio; aquí podemos observar cómo cambia de forma inmediata la curva y las probabilidades se aumenta mucho obteniendo valores hasta de 11%. Se concluye la importancia que toma acierta dentro del modelo establecido y se evidencia cómo en ausencia de tener esta información, aumenta el riesgo y asume la probabilidad de no pago. Además, se evidencia esta gráfica en particular que presenta una función logística ya que como se mencionó al inicio esta variable fue se le creo una al cuadrado.

**Gráfico 3. Curva probabilidad ventas**



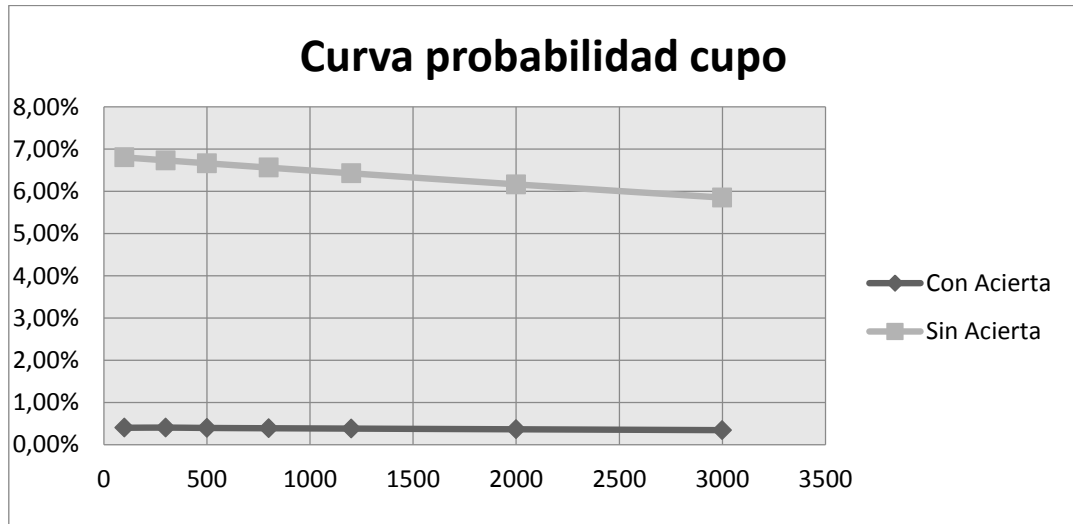
Fuente: Elaboración propia, 2017.

**Gráfico 4. Curva probabilidad activo**



Fuente: Elaboración propia, 2017.

**Gráfico 5.** Curva probabilidad cupo



Fuente: Elaboración propia, 2017.

Los gráficos de ventas, activos y cupo se construyeron, al igual que plazo, teniendo en cuenta en un caso la variable acierta y en otro sin ésta. En resumen, se repite el comportamiento enunciado anteriormente, en el cual la probabilidad aumenta bastante cuando se calculan sin acierta. Los resultados de las ventas no superan un indicador de 2% de probabilidad de pago y, en el caso de los activos, se presenta las menores probabilidades.

Por tanto, se deduce que las variables ventas, activos, plazo y acierta son significativas en el modelo y permiten entender el comportamiento de una pyme a la hora del pago de un crédito. Los indicadores financieros y el cupo se convierten en herramientas claves a la hora de realizar un estudio de crédito, para conocer la situación financiera de la empresa y contar como apoyo para la toma de decisiones.

## 5. Conclusiones

La construcción de un modelo Scoring para pyme surgió de la inquietud de entender, a partir de herramientas estadísticas y econométricas, cómo puede ser el comportamiento de pago de este tipo de empresas; teniendo en cuenta que es un segmento de alto crecimiento y desarrollo en la economía colombiana.

El desarrollo del proyecto permitió definir, después de un exhaustivo estudio y teniendo en cuenta el conocimiento previo en la aplicación real, los determinantes para el acceso de las pymes a crédito en Colombia y, a partir de allí, plantear las variables del modelo, siempre regidos por el marco regulatorio de la administración del riesgo que deben tener las entidades financieras en Colombia.

A partir del desarrollo del modelo se puede concluir que el tamaño de una pyme medida a través de las ventas y el activo resultan ser variables indispensables y con una relación directa al comportamiento de pago de una pyme. Esto es evidente en la práctica, ya que a este segmento en algunos casos, por su informalidad, se le tiene restricciones al acceso a crédito. Es decir, a medida que se van consolidando en su crecimiento, en mayores ingresos el sector financiero va asumiendo mayores riesgos. Hoy en Colombia se puede evidenciar que algunas entidades dentro de sus políticas tienen un tope mínimo de ventas y pequeñas empresas que están empezando no tienen oportunidad de acceso al mercado financiero. Además, el microcrédito en Colombia se encuentra en proceso de desarrollo y de penetración en el mercado. Otras variables como cupo o plazo, que son de manejo interno de cada entidad otorgante también son relevantes, puesto que, son los indicadores de riesgo y nivel de exposición que se desea asumir cuando se otorga un crédito. Se encuentran dentro del modelo los indicadores financieros, que ayudan a establecer un panorama de una empresa y entender la capacidad que tienen para el pago de una obligación que se va a otorgar; aunque en el modelo son no significativos (muy probablemente por el tamaño de la muestra), en la realidad son *input* de suma importancia a la hora de una evaluación de riesgo de crédito. Por último, se encuentra cierta, que en Colombia es el puntaje de una empresa

o persona natural en las centrales de riesgo; ésta resume el comportamiento histórico de un sujeto de crédito en el sector financiero. En el modelo planteado se convirtió en la variable de mayor peso para definir la probabilidad de pago o no; donde aumentaba drásticamente las probabilidades de no pago cuando se realizaban los cálculos sin tenerla disponible.

De este modelo se obtiene un resultado con respecto al nivel de riesgo que ayuda a la toma de decisiones y es un complemento para los estudios de crédito que están basados en juicio humano y revisión de cifras financieras.

## **Referencias**

- Adriazola Román, N. L. (2015). Construcción de un modelo rating de admisión para la clasificación de riesgo de crédito. *Tesis*. Santiago de Chile.
- Botello, H. (Mayo de 2015). Determinantes del acceso al crédito las PYMES en Colombia. *Ensayos de Economía*, 1-22.
- Coloma, P., Weber, R., Guajardo, J., & Miranda, J. (2006). Modelos analíticos para el manejo del riesgo de crédito. *Trend Managemet*, 45-51.
- Dabos (2007). *Credit Scoring*. Buenos Aires: Universidad de Belgrano.
- Espín, O., & Rodríguez, C. (2013). Metodología para un scoring de clientes sin referencias crediticias. *Cuadernos de Economía*, 1-22.
- Franco, M. y Urbano, D. E. (2010). El éxito de las pymes en Colombia: un estudio de casos en el sector salud. *Estudios Gerenciales*, 77-97.

- Gutiérrez, M. (2007). Modelos de credit scoring; qué, cómo, cuándo y para qué. *MPRA Paper*, 1-30.
- Hand, D. J. and Henley, W. E. (1997). "Statistical Classification Methods in Consumer CreditScoring: a Review". *Royal Statistical Society*, 523-541.
- Moscote, O; Rincón, W. (2012). Modelo Logit y Probit: un caso de aplicación. *Comunicaciones en Estadística*, 5 (2).
- Ochoa, J., Galeano, W., & Agudelo, L. (17 de enero de 2010). Construcción de un modelo scoring para el otorgamiento de crédito en una entidad financiera. *Perfil de Conyuntura Económica*, 191-222.
- Rayo, S; Lara, J; Camino, D. (2010). Modelo de Credit Scoring para instituciones de microfinanzas en el marco de Basilea II. *Journal of Economics, Finance and Administrative Science*, 90-124.
- Revista Dinero (2016). Pymes contribuyen con más del 80% del empleo en Colombia. Consultado el 18 de octubre de 2016 en <http://www.dinero.com/edicion-impresa/caratula/articulo/porcentaje-y-contribucion-de-las-pymes-en-colombia/231854>
- Romero, F; Melgarejo, Z; Vera, M. (2015). Fracaso empresarial de las pequeñas y medianas empresas (pymes) en Colombia. *Suma de Negocios*, 29-41.
- Saravia-Matus, S; Saravia-Matus, J, (2012). Gender Issues in Microfinance and Repayment Performance: The Case of Nicaraguan Microfinance Institution. *Encuentro No 91*, 7-31.
- Torres, G. (2005). El acuerdo de Basilea: Estado del arte del SARC en Colombia. *Ad-Minister*, 114-135.
- Superfinanciera (1995). Circular Externa 100 de 1995. Bogotá, Colombia.