



ESCUELA DE NEGOCIOS

CONSTRUCCIÓN DE UN MODELO DE CREDIT SCORING PARA EL  
OTORGAMIENTO DE CRÉDITO DE CONSUMO DIRIGIDO AL SECTOR  
AUTOMOTRIZ DE UNA INSTITUCIÓN FINANCIERA EN ECUADOR EN  
EL PERIODO 2018-2019

AUTOR

Andrés Sebastián Granja Armendaris

AÑO

2020



ESCUELA DE NEGOCIOS

CONSTRUCCIÓN DE UN MODELO DE *CREDIT SCORING* PARA EL  
OTORGAMIENTO DE CRÉDITO DE CONSUMO DIRIGIDO AL SECTOR  
AUTOMOTRIZ DE UNA INSTITUCIÓN FINANCIERA EN ECUADOR EN EL  
PERIODO 2018-2019

Trabajo de Titulación presentado en conformidad con los requisitos establecidos  
para optar por el título de Máster en Administración de Empresas

Profesor Guía:

Luis Antonio Vaca Hinojosa

Autor:

Andrés Sebastián Granja Armendaris

Año

2020

## DECLARACIÓN DEL PROFESOR GUÍA

Declaro haber dirigido el trabajo, Construcción de un modelo de *credit scoring* para el otorgamiento de crédito de consumo dirigido al sector automotriz de una institución financiera en Ecuador en el período 2018-2019, a través de reuniones periódicas con el estudiante Andrés Sebastián Granja Armendaris, en el semestre 2020-00, orientando sus conocimientos y competencias para un eficiente desarrollo del tema escogido y dando cumplimiento a todas las disposiciones vigentes que regulan los Trabajos de Titulación.

---

Luis Antonio Vaca Hinojosa

C.C. 1712949534

## DECLARACIÓN DEL PROFESOR CORRECTOR

Declaro haber revisado este trabajo, Construcción de un modelo de *credit scoring* para el otorgamiento de crédito de consumo dirigido al sector automotriz de una institución financiera en Ecuador en el período 2018-2019, del estudiante Andrés Sebastián Granja Armendaris, en el semestre 2020-00, dando cumplimiento a todas las disposiciones vigentes que regulan los Trabajos de Titulación.

---

Daniel Andrés Cordero Espinosa

C.C. 070332094-5

## DECLARACIÓN DE AUTORÍA DEL ESTUDIANTE

Declaro que este trabajo es original, de mi autoría, que se han citado las fuentes correspondientes y que en su ejecución se respetaron las disposiciones legales que protegen los derechos de autor vigentes.

---

Andrés Sebastián Granja Armendaris  
C.C. 1714116660

## AGRADECIMIENTOS

Agradezco a todo aquel que ha creído en mí.

## DEDICATORIA

Para Patricia Armendaris, Verónica  
Vallejo y Alberto Granja, son lo más  
lindo de mi vida.

## RESUMEN

El ingreso, la educación, el histórico crediticio y el comportamiento de pago, la facilidad y rapidez para disminuir el ingreso disponible mensual por parte del cliente, la zona de residencia, y el estado civil son algunos de los factores que condicionan la capacidad de pago de las personas en el mundo. Es por esto que un cambio de gran magnitud en alguna de estas variables podría ocasionar el incumplimiento de las obligaciones adquiridas, motivo por el cual una buena administración del riesgo crediticio es eje fundamental para la consecución de las metas de rentabilidad propuestas por las instituciones financieras, en este caso se analiza una entidad dirigida a otorgar créditos de consumo de vehículos livianos. La presente investigación analiza el comportamiento de clientes de una institución crediticia y mediante una regresión de tipo probabilística plantea encontrar las correlaciones y los efectos de distintas variables independientes sobre la variable latente. Ésta mide la probabilidad de caer en mora dividida por estratos ordenados, con el fin de predecir el comportamiento de pago de futuros clientes y mitigar el riesgo crediticio. Finalmente, a través de un método estadístico muy conocido, como lo es el modelo probabilístico ordenado, y la estimación por máxima verosimilitud, se determinó que existe evidencia empírica sobre la correlación entre las variables planteadas y la probabilidad de caer en default clasificado en diferentes categorías.

**Palabras clave:** Crédito, vehículos livianos, riesgo crediticio, variables, default, probabilístico ordenado.



## **ABSTRACT**

Income, education, historical credit, payment behavior, how easy and fast a client can lower his available monthly income, place of residence, and marital status, are among the factors that condition the payment capacity that people have around the world. This is the reason why a long magnitude change in any of the variables previously mentioned could cause a breach of acquired duties. Hence, a good credit risk management is key for reaching the financial institution profitability goals, which in this case is an entity dedicated to grant consumer loans for light vehicles. The present investigation analyzes the behavior of credit institution clients and through a probabilistic regression, proposes to find the correlations and the effects of different independent variables on the latent variable. The regression measures the probability of falling into default divided by ordered strata, in order to predict the payment behavior of future customers and mitigate credit risk. Finally, through a well-known statistical method, such as the ordered probabilistic model, and the maximum likelihood estimation procedure, empirical evidence on the correlation between the proposed variables and the probability of falling into classified default in different categories, was established.

**Keywords:** Credit, light vehicles, credit risk, variables, default, ordinal probit.

# ÍNDICE

|   |    |
|---|----|
| 1. INTRODUCCIÓN.....  | 1  |
| 2. CREDIT SCORING.....  | 3  |
| 2.1. Modelos de Credit Scoring utilizados para otorgar crédito.....                   | 4  |
| 2.1.1. Análisis Discriminante Lineal .....  | 4  |
| 2.1.2. Modelos de Regresión Logística.....  | 5  |
| 2.1.3. Modelos de Regresión Multivariantes Adaptativos .....                          | 6  |
| 2.1.4. Modelos de Programación Matemática .....                                       | 6  |
| 2.1.5. Redes Neuronales .....   | 7  |
| 2.1.6. Árboles Decisionales .....   | 7  |
| 3. METODOLOGÍA.....   | 8  |
| 3.1. Modelo <i>Credit Scoring</i> aplicado a la entidad financiera de investigación.. | 10 |
| 3.2. Muestra .....  | 13 |
| 4. CONTEXTO DE LA PROBLEMÁTICA ANALIZADA.....   | 13 |
| 5. RESULTADOS .....   | 19 |
| 5.1. Validación cruzada del modelo .....  | 21 |
| 6. CONCLUSIONES Y RECOMENDACIONES.....  | 23 |
| REFERENCIAS.....  | 25 |

# 1. INTRODUCCIÓN

El principal y tradicional negocio de las Instituciones Financieras en Ecuador y el mundo es la intermediación financiera. En primera instancia las instituciones realizan la captación de fondos de clientes (pagándoles un interés por el dinero depositado), que, en unión con el capital de la entidad, es colocado o concedido a otros clientes, a los cuales también se les cobra una tasa de interés, que debe ser superior a la mencionada en la captación. Siendo el principal objetivo asegurar la generación de utilidades mediante la recuperación del capital comprometido, una institución financiera debe desarrollar un modelo de *credit scoring*, con el fin de robustecer las herramientas de la administración del riesgo crediticio y reducir al máximo la mora en su cartera. Tomando en consideración la importancia de ésta última afirmación, se presenta a continuación un modelo de *credit scoring* creado para una Institución Financiera en el Ecuador.

De acuerdo al documento investigativo emitido por el Comité de Basilea de Supervisión Bancaria en 2000, “el riesgo de crédito se define como el potencial (probabilidad máxima) que un prestatario bancario no cumpla con sus obligaciones de acuerdo con los términos acordados” (Basel Committee on Banking Supervision, 2000, p. 1). Es por esto, que un manejo óptimo de este tipo de riesgo, es clave para la final consecución de la rentabilidad esperada en el año de operaciones por las instituciones.

Por otro lado, el crédito de consumo destinado para la compra de vehículos en Ecuador, a partir del año 2015 se clasifica como Crédito de Consumo Ordinario. Según la Resolución No. 59-2015-F de la Junta de Regulación Monetaria y Financiera, Art. 2, literal e, se estipula que el “Crédito de Consumo Ordinario.- Es el otorgado a personas naturales destinado a la adquisición o comercialización de vehículos livianos de combustibles fósil” (2015, p. 2), mientras que el resto de préstamos destinados a consumo serán clasificados como Consumo Extraordinario. Una vez dicho esto, según información de la Asociación de Empresas Automotrices del Ecuador (AEADE, 2019, p. 7), en el año 2018, hubo 291 millones de USD en

Créditos de Consumo Ordinario otorgados por las diferentes instituciones financieras y 220 millones de USD de enero a septiembre del 2019, mismos que representaron el 0.93% y el 0.91% de total de créditos concedidos por el sistema financiero respectivamente. Además de ser un rubro bastante considerable en cuanto a montos, es necesario mencionar que según el Banco Central del Ecuador (s.f.) la Tasa Activa Efectiva Referencial para el segmento de Consumo Ordinario se situó en 16.26% anual al 30 de septiembre del 2019 (16.63% anual al 31 de diciembre del 2018), mientras que la Tasa Pasiva Efectiva Referencial Por Plazo a 361 o más días se situó en 7.98% (7.73% anual al cierre del 2018), dando así un diferencial de tasa de 8.28% anual, haciendo muy atractivo para las instituciones financieras el negocio de prestar dinero para la compra de vehículos.

Adicional a lo mencionado en el párrafo precedente, es necesario recalcar que en Ecuador, la venta de vehículos livianos (automóviles, camionetas, SUV) tuvo una tendencia creciente desde el año 2016 hasta el año 2018, pasando de 55.887 unidades vendidas a diciembre de 2017 a 123.457 a finales del 2018 (incremento del 120.90%), mientras que de enero a septiembre del 2019 se vendieron 89.207 unidades de vehículos livianos, valores que van de la mano directamente con el Crédito de Consumo Ordinario y que demuestran la demanda del producto a nivel nacional (AEADE, 2019, p. 7).

Finalmente, el presente documento contiene en el primer apartado la introducción al tema de investigación. En el segundo apartado, se presentan los diferentes tipos de *credit scoring* aplicados a la banca. En el apartado tres, se expone el modelo probabilístico (Probit Ordenado) elegido para determinar la probabilidad de que un cliente caiga en mora, pero ésta definida en diferentes rangos. Luego, en la cuarta parte de este documento se incluye un análisis del contexto de la problemática (mora) relacionada a los Créditos de Consumo Ordinario. En el quinto apartado, se muestran los resultados obtenidos al desarrollar el modelo mencionado y por último el apartado seis, concluye los resultados obtenidos.

## 2. CREDIT SCORING

Los métodos tradicionalmente utilizados por una institución financiera, para conceder un crédito a una persona, normalmente están basados en la opinión subjetiva del responsable a cargo, quien determina la probabilidad que el prestatario no cumpla con el pago de su deuda. Generalmente, se fundamentan en las experiencias de decisiones pasadas, sin embargo, la creciente demanda de créditos en conjunto con la gran competencia que experimentan dichas instituciones y la innovación tecnológica de las últimas décadas, hacen que sea imprescindible la automatización de algunos modelos estadísticos de riesgo de crédito, para complementar el juicio humano. Credit scoring o calificación crediticia es el término utilizado para describir los métodos estadísticos formales para determinar la probabilidad de que los solicitantes incumplan su pago (Hand, D., & Henley, W., 1997).

Un procedimiento general para desarrollar un método de calificación crediticia implica la toma de la base de clientes de la institución financiera en donde constan los “buenos” y “malos” pagadores. Regularmente, se pueden encontrar varias características predictivas (entre 50 y 300), de las propias aplicaciones de crédito (Duffy, 1977) (Capon, 1982, p. 83). Es así, que luego de separar la base total de manera aleatoria en dos partes, una para aplicar el modelo seleccionado y la otra para realizar un proceso de backtesting (ajuste del modelo), se puede aplicar una técnica estadística multivariada. Por ejemplo, se puede realizar una regresión o un análisis discriminante con el fin de obtener la mejor predicción del comportamiento de un cliente, basándose en las características del pasado de los clientes. (Beranek and Taylor 1976; Chatterjee and Barcun 1970; Long 1976; Myers and Forgy 1963) (Capon, 1982, p. 83).

El solicitante del crédito es evaluado en el sistema de calificación desarrollado, mediante la suma de los puntos recibidos en cada una de las diferentes características al momento de llenar la solicitud, para llegar a una puntuación total. El puntaje obtenido se puede tratar de distintas formas, principalmente bajo los

conceptos que rigen a la entidad en cuanto a su apetito al riesgo. De manera que, una vez obtenido este puntaje la institución analiza en qué parte del rango de aprobación se encuentra el cliente, ya sea por un sistema de corte único, donde si este puntaje excede el límite, se otorga crédito, de lo contrario, el solicitante es rechazado o por un sistema más sofisticado, donde el cliente puede recibir un puntaje atado a cada una de las distintas etapas ordenadas de la mora de la cartera (Capon, 1982, p. 83). Es en este último método donde se centra la presente investigación.

## **2.1. Modelos de Credit Scoring utilizados para otorgar crédito**

Según el documento investigativo *“An Overview of Personal Credit Scoring: Techniques and Future Work”*, realizado por Li y Zhong en el 2012 (p. 181), los modelos estadísticos y de inteligencia artificial son dos de los métodos más importantes y más utilizados para la calificación crediticia. Es por esto que resulta útil incluir una breve síntesis con las bases teóricas de vanguardia usadas en la medición del riesgo crediticio, para de esta manera tener una vista general de las posibles soluciones y finalmente elegir el mejor modelo aplicable a la situación de la institución financiera analizada en este documento.

### **2.1.1. Análisis Discriminante Lineal**

El análisis discriminante lineal, fue propuesto por primera vez por Ra Fisher en su documento *“The Use Of Multiple Measurements in Taxonomic Problems”* en este describía que “cuando se miden dos o más poblaciones en varias características  $x_1, \dots, x_8$  se asigna un interés especial a ciertas funciones lineales de las características y mediciones que mejor discriminan a una población” (1936, p. 179). Es decir, el autor propone una técnica multivariante que se centra en el proceder que tiene cierto conjunto de variables independientes, respondiendo al cuestionamiento: ¿Qué función lineal con las características dadas, maximizaría el ratio de diferencia entre las medias y las desviaciones estándar de los grupos de interés? En su estudio, Fisher utiliza esta técnica para analizar diferentes especies de plantas, sin embargo, esta lógica luego fue llevada al campo del riesgo crediticio,

en donde se diferencian las características que definen a un grupo de pagadores “buenos” y “malos” (al momento de pagar un crédito concedido), al igual que las interacciones que se dan entre las mismas. La crítica más común a estos modelos, se da específicamente por su estructura, dado que se necesita relaciones lineales entre variables dependientes e independientes y la asunción de que la distribución de las variables explicativas debe ser Normal.

### **2.1.2. Modelos de Regresión Logística**

Los modelos de regresión logística son una alteración de la regresión lineal. Además, son de los más populares para realizar *credit scoring*, dado que permiten calcular la probabilidad que un cliente caiga o no en *default* (impago) según las variables utilizadas para describir su comportamiento. Wiginton (1980), en uno de los primeros documentos publicados que utilizan estos modelos aplicados al *credit scoring*, realizó una comparación entre modelos de análisis discriminante, llegando a la conclusión de que el modelo de regresión logística dio los mejores resultados de clasificación entre categorías de “buenos” y “malos” riesgos.

Una vez dada la introducción a la regresión logística en el párrafo precedente, es necesario recalcar que, hay ocasiones en que la escala de un resultado de categoría múltiple no es nominal, sino que tiene un orden específico. Muchas veces este tipo de análisis incluye variables como la cantidad de hijos que un hogar puede tener (0,1,2,3...etc), o la calidad del desempeño de un alumno en una clase (muy mala, mala, pobre, regular, promedio superior al promedio). En esos casos es necesario utilizar un modelo logístico o probabilístico ordenado, que, como su nombre lo indica, toma en cuenta la naturaleza ordinal de los resultados a diferencia de los modelos probabilísticos multinomiales (Bolton, 2009, p. 132).

Finalmente, esta lógica se la puede aplicar al riesgo crediticio. Si una institución financiera quiere determinar si un aplicante es “buen pagador” bajo ciertas características financieras y personales, puede optar por no usar un modelo con variables de respuesta dicotómica (1 o 0, bueno o mala), sino medir la probabilidad

del cliente prospecto en una escala ordenada (excelente, bueno, malo, muy malo) (Bolton, 2009).

### **2.1.3. Modelos de Regresión Multivariantes Adaptativos**

Son modelos no-lineales, no-paramétricos de regresión que en sus inicios fueron propuestos por Jerome Friedman (1991), en su documento investigativo "*Multivariate Adaptive Regression Splines*". El autor se centra en plantear un nuevo modelo que sirva para modelar datos de grandes dimensiones. Es decir, intenta resolver el problema de aproximar adecuadamente una función de varias o muchas variables, dado solo el valor de la función en varios puntos en el espacio variable dependiente. El modelo óptimo de esta técnica es realizado en un proceso de dos etapas. En la primera, un gran número de funciones básicas son construidas para sobre-ajustar a los datos (que pueden ser continuos, categóricos u ordenados). Luego en la segunda, las funciones más básicas son eliminadas según el criterio de validación cruzada generalizada (o sus siglas en inglés GCV). Una medida de la importancia de las variables puede darse al observar el decrecimiento de la VCG cuando se retira una variable del modelo. El proceso continúa hasta que solo queden las funciones básicas que satisfagan los requerimientos dados. Para el análisis crediticio este proceso ayuda a identificar y clasificar las mejores variables posibles en la descripción del cliente (Li y Zhong, 2012, p. 182).

### **2.1.4. Modelos de Programación Matemática**

Estos son métodos que se encuentran dentro del marco de la estadística no paramétrica. Dado un criterio objetivo de optimización, como, por ejemplo, la proporción de aplicantes clasificados correctamente en una entidad financiera, este tipo de modelos resulta muy aplicable. Hand en 1981, describió como minimizar el *perceptron criterion* (una función lineal de la suma de los puntos obtenidos en su rating por los aplicantes mal clasificados) usando programación lineal es decir se logra obtener plantillas o sistemas de calificación adecuados sin dejar de lado la optimización del ajuste de la clasificación (Hand, 1997, p. 532).



### **2.1.5. Redes Neuronales**

Estos modelos normalmente se aplican a la calificación crediticia, viéndolos como modelos estadísticos, con la distinción que involucran combinaciones lineales de secuencias anidadas de transformaciones no lineales de combinaciones lineales de variables. Es decir, estos modelos tratan de emular al sistema nervioso (de ahí su nombre), en donde los llamados “nodos de entrada” (características o variables de la operación de crédito) tienen interconexión con los “nodos de salida” (variable de respuesta es decir probabilidad de caer en default del cliente). El modelo consiste en dar respuesta a una señal de entrada e ir aprendiendo de cada una de estas experiencias (almacenando esta información), o en forma resumida, es un modelo que aprende por sí solo con cada una de estas interacciones de entradas y salidas (Rayo, Lara y Camino, 2010, p. 92) (Davis, Edelman y Gammerman, 1992, p. 46) (Li y Zhong, 2012, p. 183).

### **2.1.6. Árboles Decisionales**

Este modelo decisión también se lo conoce como modelo de partición recursiva. Para este modelo, según un estándar dado, se toma los datos de los clientes y se los divide en subconjuntos limitados, esto con el fin de lograr homogeneidad del riesgo de caer en default, superior a la de los conjuntos originales. Luego se sigue repitiendo este mecanismo, hasta lograr solventar los requerimientos del nodo final de probabilidad. En otras palabras, un árbol de decisión es una forma gráfica de representar todos los eventos (sucesos) que se derivan de la toma de decisiones en cierto momento. Al ser un modelo que no depende de asumir distribuciones de los datos, es un modelo muy didáctico y permite visualizar el problema y sus soluciones medidas por probabilidades, al igual que todos los cálculos que deben realizarse, con el fin de tomar la decisión más acertada (Rayo, Lara y Camino, 2010, p. 93).

Finalmente, los modelos mencionados son algunos de los más comunes para realizar calificación crediticia en la banca. Todos siguen una línea de funcionamiento, en la cual mediante la utilización de variables explicativas del

comportamiento del cliente prospecto, se busca encontrar la manera de predecir un resultado futuro con respecto a su probable manera de pagar la deuda en la que ha incurrido. Además, esta sección del documento permite ver a breves rasgos las técnicas utilizadas para la medición del riesgo crediticio, en las cuales cada vez (a medida que avanza la tecnología y la capacidad de obtener mayor cantidad de datos) siempre se prefiere el modelo que realice la mejor predicción.

### 3. METODOLOGÍA

En este apartado se describe la metodología escogida y se muestra su especificación estadística y matemática. Primero, es necesario conocer el estado de la cartera de los créditos para vehículos otorgada por la entidad de este estudio que se muestra a continuación.

Tabla 1

*Estado de la cartera de Créditos de Consumo Ordinario de la entidad analizada con corte a abril 2019*

|             | Morosidad   |            |             |             |             |              | Total |
|-------------|-------------|------------|-------------|-------------|-------------|--------------|-------|
|             | 0<br>cuotas | 1<br>cuota | 2<br>cuotas | 3<br>cuotas | 4<br>cuotas | >4<br>cuotas |       |
| Casos       | 2611        | 560        | 292         | 123         | 55          | 173          | 3814  |
| % del total | 68.46%      | 14.68%     | 7.66%       | 3.22%       | 1.44%       | 4.54%        | 100%  |

Adaptada de Superintendencia de Bancos, 2019.

Al 30 de abril del 2019, la cartera de créditos de consumo destinados para vehículos estuvo compuesta en un 68.46% por cartera sana que no se encuentra en mora, en un 14,68% por cartera con una cuota vencida (30 días en mora), 7.66% por cartera con dos cuotas vencidas (60 días), 3.22% y 1.44% por cartera con 3 y 4 cuotas vencidas respectivamente y finalmente 4.54% por cartera con más de 4 cuotas vencidas.

Considerando la información presentada en el párrafo precedente y tomando en cuenta como factor preponderante, que, en marzo de 2019 existieron 3.171 deudores que estaban en 0 y 2 cuotas atrasadas y para abril del 2019 tan solo 66 (2.08%) deterioraron su mora, se ha planteado organizar los datos en tres

categorías distintas. La primera, estará compuesta por clientes con mora entre 0 y 2 cuotas vencidas (Riesgo Bajo=1), la segunda representará a los clientes con mora entre 2 y 4 cuotas vencidas (Riesgo Potencial=2) y la tercera será la categoría de clientes con más de 4 cuotas vencidas y (Riesgo Alto=3), las cuales presentan un orden y sirven para modelar la variable de respuesta mediante un modelo de regresión probabilística ordenada.

Teniendo en cuenta la estratificación mencionada en el párrafo anterior, el análisis partirá de una función probabilística con un desarrollo ordenado. Este tipo de modelos analizan la reacción en probabilidad de una variable dependiente frente a un conjunto de variables independientes, además permite modelar variables cualitativas, que, en este caso específico se ajusta muy bien a las características recogidas de los clientes en la base de datos.

Según Rodríguez y Cáceres (2007), la variable latente  $Y_i^*$ , va a depender de un conjunto de variables independientes (explicativas) que determinan la elección de cada individuo  $i$ . Se representa de la siguiente manera:  $Y_i^* = x_i\beta' + \varepsilon_i$

Donde  $\mu_i$  representa al error aleatorio. El individuo se sitúa dentro de las alternativas de rangos  $\mu_n$   $n= 1, \dots, N$  que son un reflejo del valor de  $Y_i^*$ , como se muestra a continuación (Hung, Cheng, Chen y Huang. 2013, p. 95-96) (Rodríguez y Cáceres, 2007):

$$Y_i = \begin{cases} 0, & \text{si } Y_i^* \leq 0 \\ 1, & \text{si } 0 < Y_i^* \leq \mu_1 \\ 2, & \text{si } \mu_1 < Y_i^* \leq \mu_2 \\ \vdots & \\ N, & \text{si } \mu_{N-1} < Y_i^* \end{cases}$$

asumiendo que  $\mu_0 = 0$  y se mantiene la perspectiva de variable latente el modelo adopta la siguiente forma:

$$P(Y_i = 0) = P(Y_i^* \leq 0) = P(\varepsilon_i \leq \beta'x_i),$$

$$P(Y_i = n) = P(\mu_{n-1} < Y_i^* \leq \mu_n) = P(\mu_{n-1}\beta'x_i < \varepsilon_i \leq \mu_n - \beta'x_i), n = 1, \dots, N - 1,$$

$$P(Y_i = N) = P(Y_i^* > \mu_{N-1}) = P(\varepsilon_i > \mu_{N-1} - \beta'x_i),$$

Si se reemplaza a la distribución de la variable aleatoria  $\varepsilon_i$  por F, entonces

$$\begin{aligned} P(Y_i = 0) &= F(-\beta'x_i), \\ P(Y_i = n) &= P(\mu_n - \beta'x_i) - F(\mu_{n-1} - \beta'x_i), \quad n = 1, \dots, N - 1, \\ P(Y_i = N) &= 1 - F(\mu_{N-1} - \beta'x_i) \end{aligned}$$

De acuerdo, al trabajo de Barron (2018), muchas veces se utiliza la función de distribución normal acumulativa (CDF) que suele llamarse también función probit. “Esto es virtualmente idéntico a la función logística, y como la interpretación de resultados usando la función logística es generalmente más sencillo (..)” (p. 7).

En el caso del modelo probabilístico ordenado para la regresión se utiliza una función de enlace normal inversa (Newsom, 2019, p.2).

$$Y^* = \Phi^{-1}[\hat{p}] = \beta_0 + \beta_1 x$$

En donde  $\Phi$  representa:

$$\Phi = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha + \beta x} \exp\left(-\frac{1}{2}Z^2\right) dZ$$

La estimación del modelo se lleva a cabo por máxima verosimilitud recurriendo a algoritmos iterativos de optimización. Los umbrales se estiman simultáneamente con los parámetros del vector  $\beta$ , común para todas las alternativas (Spanos, 2003, p.659-668).

### 3.1. Modelo *Credit Scoring* aplicado a la entidad financiera de investigación

En el caso particular del presente estudio el modelo probabilístico ordenado se presenta de la siguiente forma:

$$Y_i = \begin{cases} 1, & \text{si } Y_i^* \leq a \\ 2, & \text{si } a \leq Y_i^* \leq b \\ 3, & \text{si } Y_i^* > b \end{cases}$$

donde las categorías 1, 2 y 3 representan Riesgo Bajo, Riesgo Potencial y Riesgo Alto respectivamente,  $Y_i^*$  es la variable latente ordenada que adopta las categorías

de riesgo,  $a$  es el primer corte de la estimación y  $b$  es el segundo corte. Las probabilidades que  $Y_i^*$  adopte un determinado valor ordenado de categoría se calcula de la siguiente forma:

Probabilidad que  $Y_i^* = \text{Riesgo Bajo}$

$$P(Y_i^* \leq a) = P(x_i\beta = \mu_i \leq a)$$

$$P(Y_i^* \leq a) = P(\mu_i \leq a - x_i\beta)$$

$$P(Y_i^* \leq a) = \Phi(a - x_i\beta)$$

Probabilidad que  $Y_i^* = \text{Riesgo Potencial}$

$$P(a \leq Y_i^* \leq b) = P(a \leq x_i\beta + a_i \leq b)$$

$$P(a \leq Y_i^* \leq b) = P(x_i\beta + \mu_i \leq b) - P(x_i\beta + \mu_i \leq a)$$

$$P(a \leq Y_i^* \leq b) = P(\mu_i \leq b - x_i\beta) - P(\mu_i \leq a - x_i\beta)$$

$$P(a \leq Y_i^* \leq b) = \Phi(b - x_i\beta) - \Phi(a - x_i\beta)$$

Probabilidad que  $Y_i^* = \text{Riesgo Alto}$

$$P(Y_i^* > b) = 1 - P(Y_i^* \leq b)$$

$$P(Y_i^* > b) = 1 - \Phi(b - x_i\beta)$$

Especificación del modelo

$Y_i$  (Prob. de mora ordenada)

$$= \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i} + \beta_6 X_{6i} + \beta_7 X_{7i} + \beta_8 X_{8i} + \varepsilon_i$$

Donde:

$Y_i$  : *estratoMora*. Es la variable dependiente que adopta el número de la categoría de riesgo, Riesgo Bajo=1, Riesgo Potencial=2, Riesgo Alto=3, según la clasificación

mencionada antes, que se ajusta a la cantidad de meses caídos en mora por el cliente.

$X_{1i}$ : *score\_actual*. Indica el valor medido en puntos sobre 999 del score de riesgo crediticio dado por el buró de crédito cuando realizó la solicitud inicial del crédito.

$X_{2i}$ : *historico\_cliente*. Indica el comportamiento histórico del cliente con relación a créditos previos concedidos por la institución, tomando el valor de 1 si el cliente fue un mal pagador en el pasado y 0 en el caso contrario, cuando realizó la solicitud del crédito.

$X_{3i}$ : *ingreso*. Es una variable que recoge todos los ingresos del cliente al momento de aplicar para el crédito, unificando ingresos en relación de dependencia con ingresos por negocios propios (con sus debidos respaldos de declaraciones de impuestos), cuando realizó la solicitud del crédito.

$X_{4i}$ : *educ\_primarianula*. Es una variable dicótoma que toma el valor de 1 si el cliente tuvo un máximo de educación primaria o no tuvo educación y 0 si fue otra categoría de educación cuando realizó la solicitud del crédito.

$X_{5i}$ : *soltero\_new*. Es una variable dicótoma que adopta el valor de 1 si el cliente tenía estado civil soltero o divorciado y 0 si tenía el resto de estados civiles cuando realizó la solicitud del crédito.

$X_{6i}$ : *UIO*. Es una variable dicótoma que adopta el valor de 1 si el cliente vivía en la ciudad de Quito y 0 si residía en cualquier otra ciudad del país cuando realizó la solicitud del crédito.

$X_{7i}$ : *GYE*. Es una variable dicótoma que adopta el valor de 1 si el cliente vivía en la ciudad de Guayaquil y 0 si residía en cualquier otra ciudad del país cuando realizó la solicitud del crédito.

$X_{8i}$ : *dbito*. Es una variable dicótoma que adopta el valor de 1 si el cliente tenía una cuenta de la cual se le hacían débitos directos de otros créditos obtenidos o pagos de servicios y 0 si no la tenían cuando realizó la solicitud del crédito.

### 3.2. Muestra

La muestra elegida es de corte transversal y tiene como base la cartera de clientes que han obtenido un crédito de consumo ordinario por parte de una entidad financiera, ubicada en la ciudad de Quito. Esta base cuenta con un total de 3814 clientes al 30 de abril del 2019. Para poder realizar la metodología mencionada en el apartado anterior, se ha dividido a la base en dos partes (de manera aleatoria), la primera para realizar las estimaciones (2670 observaciones) y la segunda para realizar *backtesting* y poder comprobar el ajuste del modelo (1144 observaciones).

Para la variable dependiente o latente se tuvo que transformar las cuotas caídas en mora de los clientes a cada rango mencionado; para la variable *ingresos* se tuvo que unificar los ingresos en relación de dependencia de los clientes con los ingresos de negocios propios respaldados por las debidas declaraciones de impuestos; para la variable *score\_actual* se tomó el valor reportado por el buró de crédito (de 0 a 999 puntos); para las variables *histórico\_cliente*, *educ\_primarianula*, *soltero\_new*, *UIO*, *GYE* y *dbito* se tuvo que transformar en unos a las características buscadas por cada una y en ceros si era el caso contrario o el resto de categorías.

Finalmente, cabe mencionar la composición de la base de datos utilizada para estimar el modelo, donde del total de 2670, la cartera con Riesgo Bajo representó el 92.58% (2472), la de Riesgo Potencial presentó un 4.11% (110) y la de Riesgo Alto exhibió un 3.30% (88).

## 4. CONTEXTO DE LA PROBLEMÁTICA ANALIZADA

El ente regulador del sistema financiero ecuatoriano es la Superintendencia de Bancos (2019). En su página web se puede encontrar toda la información necesaria para describir al sector, es así que como datos iniciales se menciona que para el 30

de septiembre del 2019 existieron 24 bancos privados y 5 instituciones financieras públicas (dentro de las instituciones financieras públicas se encuentran BIESS, Fondo Nacional de Garantías, BanEcuador B.P., Banco de Desarrollo y CFN B.P.).

En cuanto a la actividad del sistema financiero la Superintendencia de Bancos (2019) mencionó que al 30 de septiembre que los activos totales de las distintas entidades que comprenden el Sector Financiero Ecuatoriano, alcanzaron un valor de USD 50.864.290, principalmente gracias al crecimiento de la cartera de créditos, misma que para septiembre del 2019 totalizó USD 34.191.590, con una participación del 85.88% por parte de la Banca Privada y un 14.12% por parte de la Banca Pública.

Tabla 2

*Estructura de la cartera de créditos en el Ecuador (en miles de dólares)*

| Período   | Bancos Privados | Sociedades Financieras | Banca Pública | Total      |
|-----------|-----------------|------------------------|---------------|------------|
| 2016      | 20.375.061      | 1.339.567              | 3.891.175     | 25.605.803 |
| 2017      | 24.600.644      |                        | 4.203.580     | 28.804.225 |
| Sept-2018 | 26.883.705      |                        | 4.549.605     | 31.433.310 |
| 2018      | 27.325.175      |                        | 4.667.572     | 31.992.748 |
| Sept-2019 | 29.364.456      |                        | 4.827.133     | 34.191.590 |

Adaptada de Superintendencia de Bancos, 2019.

Por otro lado, en cuanto a la morosidad de la cartera total del Sistema Financiero, se observó que para septiembre de 2019 se situó en 3.88%, valor superior al del año anterior a septiembre (3.61%), esto por el comportamiento de la mora en la Banca Pública.

Tabla 3

*Índices de morosidad de la cartera total*

| Período   | Bancos Privados | Sociedades Financieras | Banca Pública | Total |
|-----------|-----------------|------------------------|---------------|-------|
| 2016      | 3,54%           | 5,68%                  | 5,28%         | 3,92% |
| 2017      | 2,96%           |                        | 6,52%         | 3,48% |
| Sept-2018 | 3,11%           |                        | 6,60%         | 3,61% |
| 2018      | 2,62%           |                        | 4,86%         | 2,95% |
| Sept-2019 | 3,00%           |                        | 7,93%         | 3,88% |

Adaptada de Superintendencia de Bancos, 2019.



Dentro de la problemática analizada para la presente investigación es necesario observar cómo se ha comportado el margen de intermediación financiera del Sistema Financiero, mismo que se puede notar ha decrecido en un 12,30%.

Tabla 4

*Margen de Intermediación del sistema financiero (en miles de dólares)*

| Período                  | 2016    | 2017    | Sept-2018 | 2018    | Sept-2019 |
|--------------------------|---------|---------|-----------|---------|-----------|
| Margen de Intermediación | 183.020 | 402.409 | 489.473   | 686.115 | 429.292   |

Adaptada de Superintendencia de Bancos, 2019.

Además, se observó que la rentabilidad del Sistema Financiero se ha deteriorado, ya que para septiembre de 2018 el ROA del total del sistema se ubicó en 2.18%, mientras que para septiembre del 2019 se ubicó en 1.33%. De la misma manera la rentabilidad para los accionistas sufrió un decrecimiento, ya que el ROE pasó de 11.21% en septiembre del 2018 a 8.55% en septiembre del 2019, esto debido a la situación de la Banca Pública, que disminuyó los indicadores globales en gran manera, aun cuando los Bancos Privados, por el contrario, aumentaron su rentabilidad.

Tabla 5

*Índices de Rentabilidad sobre los activos ROA*

| Período   | Bancos Privados | Sociedades Financieras | Banca Pública | Total |
|-----------|-----------------|------------------------|---------------|-------|
| 2016      | 0,62%           | 2,15%                  | 2,27%         | 1,68% |
| 2017      | 1,02%           |                        | 1,29%         | 1,15% |
| Sept-2018 | 1,40%           |                        | 2,97%         | 2,18% |
| 2018      | 1,35%           |                        | 2,89%         | 2,12% |
| Sept-2019 | 1,49%           |                        | 1,17%         | 1,33% |

Adaptada de Superintendencia de Bancos, 2019.

Tabla 6

*Índices de Rentabilidad sobre el patrimonio ROE*

| Período   | Bancos Privados | Sociedades Financieras | Banca Pública | Total  |
|-----------|-----------------|------------------------|---------------|--------|
| 2016      | 6,72%           | 12,34%                 | 7,58%         | 8,88%  |
| 2017      | 10,36%          |                        | 4,22%         | 7,29%  |
| Sept-2018 | 13,28%          |                        | 9,13%         | 11,21% |
| 2018      | 13,65%          |                        | 9,00%         | 11,32% |
| Sept-2019 | 13,71%          |                        | 3,39%         | 8,55%  |

Adaptada de Superintendencia de Bancos, 2019.

También es necesario analizar la composición de la cartera de los Bancos Privados en Ecuador con el fin de poder delimitar un poco más el contexto en el que se desenvuelve la entidad de la presente investigación. En la tabla subsecuente, se puede ver que la participación de los créditos de consumo ordinario ha aumentado en su sobre el total de la cartera, pasando de USD 324.333.000 que representaron un 1.59% del total de la cartera en el 2016, a USD 618.649.000 que representaron un 2.26% a finales del 2018 (2.28% a finales de septiembre de 2019).

Tabla 7

*Composición de la cartera bruta de bancos privados*

| Período      | 2016    | 2017    | Sep-18  | 2018    | Sep-19  |
|--------------|---------|---------|---------|---------|---------|
| Comercial    | 45,30%  | 40,10%  | 37,92%  | 37,33%  | 34,16%  |
| Prioritario  |         |         |         |         |         |
| Consumo      | 29,73%  | 32,88%  | 34,11%  | 34,28%  | 36,40%  |
| Prioritario  |         |         |         |         |         |
| Inmobiliario | 9,61%   | 8,61%   | 8,18%   | 8,25%   | 7,91%   |
| Microempresa | 7,29%   | 6,45%   | 6,20%   | 6,35%   | 6,57%   |
| Productivo   | 3,50%   | 6,72%   | 8,45%   | 8,74%   | 9,96%   |
| Comercial    | 0,77%   | 1,01%   | 1,09%   | 1,10%   | 1,06%   |
| Ordinario    |         |         |         |         |         |
| Consumo      | 1,59%   | 2,09%   | 2,29%   | 2,26%   | 2,28%   |
| Ordinario    |         |         |         |         |         |
| Vivienda de  |         |         |         |         |         |
| Interés      | 0,11%   | 0,36%   | 0,13%   | 0,23%   | 0,25%   |
| Público      |         |         |         |         |         |
| Educativo    | 2,11%   | 1,76%   | 1,61%   | 1,47%   | 1,41%   |
| Total        | 100,00% | 100,00% | 100,00% | 100,00% | 100,00% |

Adaptada de Superintendencia de Bancos, 2019.

De igual manera se observó la morosidad dividida por cartera, en donde se muestra una marcada tendencia creciente en los períodos anuales en los créditos de consumo ordinario, al igual que un ligero incremento de septiembre del 2018 a septiembre del 2019.

Tabla 8  
*Composición de la cartera bruta de bancos privados*

| Período                     | 2016  | 2017  | Sep-18 | 2018  | Sep-19 |
|-----------------------------|-------|-------|--------|-------|--------|
| Comercial Prioritario       | 1,26% | 1,15% | 1,35%  | 0,90% | 1,23%  |
| Consumo Prioritario         | 6,91% | 5,23% | 5,00%  | 4,59% | 4,84%  |
| Inmobiliario                | 2,99% | 2,89% | 3,33%  | 2,80% | 3,27%  |
| Microempresa                | 6,57% | 4,99% | 5,03%  | 4,93% | 5,09%  |
| Productivo                  | 0,10% | 0,27% | 0,30%  | 0,21% | 0,43%  |
| Comercial Ordinario         | 0,06% | 0,15% | 0,28%  | 0,20% | 0,35%  |
| Consumo Ordinario           | 2,92% | 4,91% | 6,90%  | 5,78% | 6,92%  |
| Vivienda de Interés Público | 0,00% | 0,17% | 0,97%  | 2,11% | 2,33%  |
| Educativo                   | 4,68% | 4,79% | 7,15%  | 0,71% | 1,29%  |
| Total                       | 3,54% | 2,96% | 3,11%  | 2,62% | 3,00%  |

Adaptada de Superintendencia de Bancos, 2019.

Para finalizar, se presenta una figura en la cual se describe la interacción entre los Créditos de Consumo Ordinario concedidos por el sistema financiero y las unidades vendidas de vehículos en una serie de tiempo mensual. Aquí se puede notar claramente como el crédito es responsable de guiar la compra de vehículos en el Ecuador, motivo por el cual se puede observar una relación directa entre las gráficas de barras y líneas.

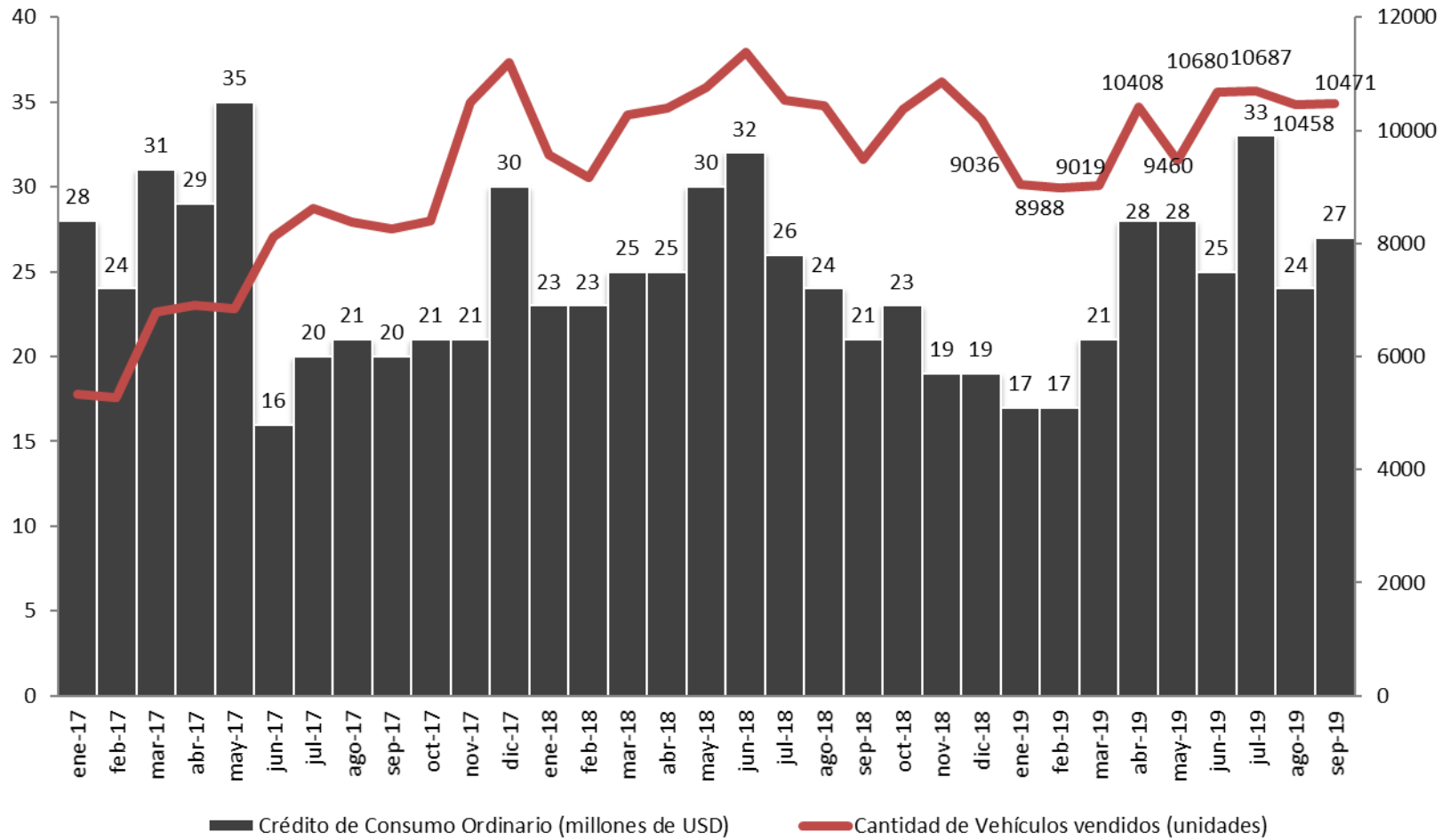


Figura1. Crédito de Consumo Ordinario vs. ventas de vehículos livianos. Adaptada de AEADE,2019.

## 5. RESULTADOS

El modelo de calificación de riesgo crediticio presentado fue estimado mediante la utilización del programa Stata en su versión 16. El mismo que a través del comando `oprobit` permitió obtener los coeficientes de cada una de las variables, junto con los errores estándar, los valores z, los valores p, los intervalos de confianza y los cortes necesarios para poder realizar la estimación de las probabilidades para cada una de las categorías. A continuación, se muestran los resultados obtenidos:

```
oprobit estratomora score_actual historico_cliente ingreso educ_primarianula
soltero_new Quito GYE dbito
```

```
Iteration 0 log likelihood = -841.59851
Iteration 1 log likelihood = -582.67919
Iteration 2 log likelihood = -546.23916
Iteration 3 log likelihood = -544.38962
Iteration 4 log likelihood = -544.37707
Iteration 5 log likelihood = -544.37707
```

```
Ordered probit regression           Number of obs       = 2670
                                   LR chi2(8)           = 594.44
                                   Prob > chi2           = 0.0000
Log likelihood          -54.437.707  Pseudo R2           = 0.3532
```

Tabla 9

*Estimación del modelo probabilístico de riesgo de crédito*

| <i>estratomora</i>       | <i>Coef.</i>     | <i>Std. Err.</i> | <i>z</i>     | <i>P&gt;z</i> | <i>[95% Conf. Interval]</i> |                  |
|--------------------------|------------------|------------------|--------------|---------------|-----------------------------|------------------|
| <i>score_actual</i>      | <i>-.0010664</i> | <i>.0001491</i>  | <i>-7.15</i> | <i>0.000</i>  | <i>-.0013585</i>            | <i>-.0007742</i> |
| <i>historico_cliente</i> | <i>2.011098</i>  | <i>.1387213</i>  | <i>14.50</i> | <i>0.000</i>  | <i>1.739209</i>             | <i>2.282986</i>  |
| <i>ingreso</i>           | <i>-.0000845</i> | <i>.0000383</i>  | <i>-2.21</i> | <i>0.027</i>  | <i>-.0001596</i>            | <i>-9.49e-06</i> |
| <i>educ_primarianula</i> | <i>.2885069</i>  | <i>.121284</i>   | <i>2.38</i>  | <i>0.017</i>  | <i>.0507946</i>             | <i>.5262192</i>  |
| <i>soltero_new</i>       | <i>.185682</i>   | <i>.097838</i>   | <i>1.90</i>  | <i>0.058</i>  | <i>-.006077</i>             | <i>.3774411</i>  |
| <i>UIO</i>               | <i>-.082802</i>  | <i>.133406</i>   | <i>-0.62</i> | <i>0.535</i>  | <i>-.3442729</i>            | <i>.1786689</i>  |
| <i>GYE</i>               | <i>-.1069768</i> | <i>.1169488</i>  | <i>-0.91</i> | <i>0.360</i>  | <i>-.3361922</i>            | <i>.1222387</i>  |
| <i>dbito</i>             | <i>.035727</i>   | <i>.1099068</i>  | <i>0.33</i>  | <i>0.745</i>  | <i>-.1796863</i>            | <i>.2511403</i>  |
| <i>/cut1</i>             | <i>2.01229</i>   | <i>.2088903</i>  |              |               | <i>1.602872</i>             | <i>2.421707</i>  |
| <i>/cut2</i>             | <i>2.627649</i>  | <i>.2139043</i>  |              |               | <i>2.208404</i>             | <i>3.046894</i>  |

En la tabla precedente se puede observar la dirección de los coeficientes sobre las probabilidades. Se denota que, el coeficiente atado a la variable `score_actual` tiene signo negativo, es decir una relación inversa en el efecto del puntaje que tiene en el buró de crédito y la probabilidad de mora de los clientes. Luego la variable que recoge el comportamiento histórico de los clientes con la entidad (`histórico_cliente`), refleja un coeficiente de gran magnitud y positivo, determinando una relación directa en el efecto del cliente “mal” pagador con la mora. En tercer lugar, la variable ingreso refleja una relación inversa en el efecto del dinero que recibe mensualmente una persona y la mora en la que puede incurrir. La educación primaria o nula (`educ_primarianula`), demuestra según la estimación, una relación directa sobre el efecto de la cantidad de años de educación de un cliente y la mora incurrida, de la misma manera se comporta la variable `soltero_new` al recoger el sentido del efecto entre los clientes que son solteros o divorciados y la mora incurrida. Por otro lado, las variables de ubicación geográfica del cliente muestran una relación inversa en el efecto de vivir en la ciudad de Quito o Guayaquil (UIO,GYE) sobre la mora incurrida. Y finalmente la variable `dbito` que toma el valor de 1 si el cliente tiene una cuenta de donde se le realicen débitos directos de pagos de otros créditos o servicios y 0 si no la tiene, presenta una directa con la mora incurrida.

Por otro lado, en cuanto a los valores que permiten inferir significancia estadística de las distintas variables, se puede decir que 5 de las 8 variables tiene valores p que reflejan significancia estadística, sin embargo, en este tipo de modelos no se puede concluir significancia simplemente por los valores p observados, sino que se suele realizar un análisis de los efectos marginales de cada una de las variables para ver los rangos en los que se mueve la significancia de cada una. Es por esto que se han conservado todas las variables originales del modelo.

### 5.1. Validación cruzada del modelo

Para realizar la validación cruzada del modelo se utilizó un 70% de la base de datos total para realizar la estimación (2670), el restante 30% de la base (1144) se utilizaron para realizar pruebas cruzadas de la validación del modelo.

El proceso comienza con la estimación de los coeficientes de las variables dadas que luego permitirán realizar las predicciones de las probabilidades. Luego se procedió analizar cada una de las categorías planteadas para la variable latente como si fuera una variable discreta, es decir categoría1 o Bajo Riesgo = 1, todo lo demás cero. Y se contrastó si la probabilidad predicha por el modelo se asemeja a la realidad de la categoría, mediante un punto de corte de referencia para las probabilidades predichas. Este punto de corte es dado a través del apetito al riesgo de la institución. Para la segunda categoría o Riesgo Potencial se hace lo mismo y para tercera categoría o Riesgo Alto de igual manera. El núcleo de este análisis, es poder analizar las predicciones del modelo frente a la realidad. Cabe mencionar que los puntos de corte para cada una de las categorías fueron, 70% para Riesgo Bajo, y 17% para Riesgo Potencial y 17% para riesgo Alto.

Tabla 10

*Validación cruzada por categoría de Riesgo Bajo*

|          |   | Predicción |        |         |
|----------|---|------------|--------|---------|
|          |   | 0          | 1      | Total   |
| Realidad | 0 | 75,16%     | 24,84% | 100,00% |
|          | 1 | 12,51%     | 87,49% | 100,00% |

En la tabla superior se hace referencia a la validación cruzada del Riesgo Bajo, en donde se puede notar que las predicciones de no caer en el valor 1 (es decir caer en estratos distintos de 1) fueron correctas en un 75,16% al contrastar con la realidad. Mientras que el 87,49% fue el valor de las veces que el modelo predijo acertadamente la probabilidad de caer en la categoría de mora 1.

Tabla 11  
*Validación cruzada por categoría de Riesgo Potencial*

|          |   | Predicción |        |         |
|----------|---|------------|--------|---------|
|          |   | 0          | 1      | Total   |
| Realidad | 0 | 81,23%     | 18,77% | 100,00% |
|          | 1 | 39,71%     | 60,29% | 100,00% |

En la tabla precedente se hace referencia a la validación cruzada del Riesgo Potencial, en donde se puede notar que las predicciones de no caer en el valor 1 (es decir caer en estratos distintos de 1) fueron correctas en un 81,23% al contrastar con la realidad. Mientras que el 60,29% fue el valor de las veces que el modelo predijo acertadamente la probabilidad de caer en 1 que hace referencia a la categoría de mora 2.

Tabla 12  
*Validación cruzada por categoría de Riesgo Alto*

|          |   | Predicción |        |         |
|----------|---|------------|--------|---------|
|          |   | 0          | 1      | Total   |
| Realidad | 0 | 90,46%     | 9,54%  | 100,00% |
|          | 1 | 31,76%     | 68,24% | 100,00% |

En la tabla presentada se hace referencia a la validación cruzada del Riesgo Alto, en donde se puede notar que las predicciones de no caer en el valor 1 (es decir caer en estratos distintos de 1) fueron correctas en un 90,46% al contrastar con la realidad. Mientras que el 68,24% fue el valor de las veces que el modelo predijo acertadamente la probabilidad de caer en 1 que hace referencia a la categoría de mora 3.

Concluyendo, se puede comprobar a través de los resultados obtenidos tanto en la estimación como en las tablas de validación cruzada, que el modelo se ajusta bien a la realidad de la cartera de la entidad financiera y que dados los puntos de corte apegados a su posición frente al riesgo se alcanza un nivel óptimo de predicción del comportamiento de sus clientes.



## 6. CONCLUSIONES Y RECOMENDACIONES

A lo largo del documento se ha descrito la idea principal del negocio de las entidades financieras y como la administración del riesgo crediticio es fundamental para la consecución de las metas de rentabilidad planteada al inicio del año de operaciones. A la par, queda claro que la implementación de un modelo de *credit scoring* representa una ventaja al ser utilizado para determinar objetivamente la probabilidad que una persona caiga en mora o no y por cuanto tiempo.

La investigación presentada comprueba la existencia de correlación entre las variables establecidas y la probabilidad de incurrir en alguna de las categorías de mora. Es decir, que mediante un análisis práctico se pudo encontrar la relación probabilística entre caer en diferentes escalas de default y variables de comportamiento de los clientes. Siendo la variable más preponderante comportamiento histórico del cliente con la institución, seguido de la educación del mismo.

Enmarcado en el tema, se puede concluir que es clara la evidencia empírica que argumenta la hipótesis de esta investigación, aun cuando los valores de significancia estadística para tres de las ocho variables están por encima del límite convencional en la estimación inicial, éstas no fueron removidas dado que el análisis de los efectos marginales es en donde se determinarían la significancia por rangos. Es importante mantener todas las variables dentro del modelo ya que en el panorama global todas aportan con información relevante.

En el análisis realizado de predictibilidad (validación cruzada) el modelo se acerca y representa la realidad del negocio para la base de datos obtenida de la entidad estudiada. Cabe recalcar que los datos y resultados conseguidos se aplican solo para la entidad en mención y se necesitarían tomar datos de diferentes instituciones para poder ampliar el espectro de aplicación del modelo presentado.

Para concluir, al ser un modelo que trabaja con información provista por los clientes existe un margen de error en las variables que dependen de la veracidad del cliente. Es por esto que se recomienda mitigar dicho riesgo, mediante el cruce de información entre bases de datos de distintos organismos, con la previa autorización del cliente.

## REFERENCIAS

- AEDE. (2019). Sector automotor en cifras No. 39. Recuperado de: <http://www.aeade.net/wp-content/uploads/2019/12/Boletin-sector-automotor-resumido-N-%C2%B0-39.pdf>
- Banco Central del Ecuador. (s.f.). Tasas de interés efectivas vigentes datos históricos. Recuperado de: <https://contenido.bce.fin.ec/documentos/Estadisticas/SectorMonFin/TasasInteres/TasasHistorico.htm>
- Barron, David. (2018). Oxford University. Recuperado de: <http://users.ox.ac.uk/~jesu0073/Lecture%203/LogisticRegression.pdf>
- Basel Committee on Banking Supervision. (2000). Principles for the Management of Credit Risk. Recuperado de: <https://www.bis.org/publ/bcbs75.pdf>
- Bolton, Christine. (2009). Logistic regression and its application in credit scoring. University of Pretoria. Recuperado de: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.1024.2660&rep=rep1&type=pdf>
- Capon, Noel. (1982). Credit Scoring Systems: A Critical Analysis. Columbia University. 46(2), 82-91. Recuperado de: <https://doi.org/10.1177/002224298204600209>
- Davis, R. H., Edelman, D. B., & Gamberman, A. J. (1992). Machine-Learning Algorithms for Credit-Card Applications. *Journal of Management Mathematics*, 4(1), 43-51.
- Fisher, Ra. (1936). The Use of Multiple Measurements in Taxonomic Problems. *Annals of Human Genetics*. Recuperado de: <https://doi.org/10.1111/j.1469-1809.1936.tb02137.x>
- Friedman, J. H. (1991). Multivariate Adaptive Regression Splines. *Annals of Statistics*, 19(1), pp. 1-141. doi:10.1214/aos/1176347963
- Hand, D.J. and Henley, W.E. (1997). Statistical Classification Methods in Consumer Credit Scoring: A Review. *Journal of Royal Statistical Society*, 160, 523-541. Recuperado de: <https://doi.org/10.1111/j.1467-985X.1997.00078.x>
- Hung, Chang, Hui Wen Cheng, Shih-shen Chen y Ying-Chen Huang. (2013). Factors That Affect Credit Rating: An Application Of Ordered Probit Models. Recuperado de: [http://www.ipe.ro/rjef/rjef4\\_13/rjef4\\_2013p94-108.pdf](http://www.ipe.ro/rjef/rjef4_13/rjef4_2013p94-108.pdf)

- Junta de Regulación Monetaria y Financiera. (2015). Resolución No. 59-2015-F. Recuperado de: <https://contenido.bce.fin.ec/documentos/Estadisticas/SectorMonFin/TasasInteres/RegTasas059.pdf>
- Li, Xiao-Lin and Yu Zhong. (2012). An Overview of Personal Credit Scoring: Techniques and Future Work. *International Journal of Intelligence Science*, 2, 181-189. Recuperado de: <http://dx.doi.org/10.4236/ijis.2012.224024>
- Newson, Jason. (2019). Link Functions and the Generalized Linear Model. *Multiple Regression and Multivariate Quantitative Methods*. Recuperado de: [http://web.pdx.edu/~newsomj/mvclass/ho\\_link.pdf](http://web.pdx.edu/~newsomj/mvclass/ho_link.pdf)
- Rayo Cantón, Salvador, Juan Lara Rubio y David Camino Blasco. (2010). Un Modelo de Credit scoring para instituciones de microfinanzas en el marco de Basilea II. Recuperado de: <http://www.scielo.org.pe/pdf/jefas/v15n28/a05v15n28.pdf>
- Rodriguez, C y Cáceres, J. (2007). Modelos de elección discreta y especificaciones ordenadas: una reflexión metodológica. *Estadística Española* 49 (166), 451-471. Recuperado de: [https://www.researchgate.net/publication/239931007\\_Modelos\\_de\\_eleccion\\_discreta\\_y\\_especificaciones\\_ordenadas\\_una\\_reflexion\\_metodologica](https://www.researchgate.net/publication/239931007_Modelos_de_eleccion_discreta_y_especificaciones_ordenadas_una_reflexion_metodologica)
- Spanos, A (2003). *Probability Theory and Statistical Inference*. (1.a ed.). Cambridge: Cambridge University Press (Virtual Publishing) doi: ISBN 0 511 01097 4
- Superintendencia de Bancos. (2019). Boletín Financiero septiembre 2019. Recuperado de: [http://estadisticas.superbancos.gob.ec/portalestadistico/portalestudios/?page\\_id=415](http://estadisticas.superbancos.gob.ec/portalestadistico/portalestudios/?page_id=415)
- Superintendencia de Bancos. (2019). Volumen de Crédito. Recuperado de: [http://estadisticas.superbancos.gob.ec/portalestadistico/portalestudios/?page\\_id=327](http://estadisticas.superbancos.gob.ec/portalestadistico/portalestudios/?page_id=327)
- Wiginton, J.C. (1980). A note on the comparison of logit and discriminant models of consumer credit behaviour. *J. Finan. Quant. Anal.*, 15(3), 757-770. doi: 10.2307/2330408

