

PONTIFICIA UNIVERSIDAD JAVERIANA

FACULTAD DE CIENCIAS ECONÓMICAS Y  
ADMINISTRATIVAS

MAESTRÍA EN ECONOMÍA

MODELO DE SCORING PARA APROBACIÓN  
DE CRÉDITOS PARA LA CARTERA DE  
CONSUMO, EN UNA COOPERATIVA DE  
APORTE Y CRÉDITO COLOMBIANA

VICTOR MANUEL GUEVARA CASTRO

Y

MANUEL JOSE MORENO NORIEGA

diciembre 12 de 2019

# MODELO DE SCORING PARA APROBACIÓN DE CRÉDITOS PARA LA CARTERA DE CONSUMO, EN UNA COOPERATIVA DE APORTE Y CRÉDITO COLOMBIANA

VICTOR MANUEL GUEVARA CASTRO

Y

MANUEL JOSE MORENO NORIEGA

Trabajo de grado presentado como requisito para optar al título de:

Magister en economía

Director

Ferney Herrera Cruz

PONTIFICIA UNIVERSIDAD JAVERIANA

diciembre 12 de 2019



## Tabla de contenido

1. Resumen.....	7
2. Introducción.....	10
3. Marco Teórico .....	16
4. El scoring de crédito y su avance en Colombia .....	17
5. El modelo logit .....	18
6. El scoring como mecanismo para determinar la probabilidad de Incumplimiento.....	18
7. Metodología .....	19
<b>a) Datos .....</b>	<b>21</b>
<b>b) Variables cualitativas y cuantitativas del Modelo Econométrico .....</b>	<b>22</b>
<b>c) Análisis descriptivo de la muestra.....</b>	<b>25</b>
<b>d) Curva de ROC .....</b>	<b>26</b>
8. Estimación del Modelo.....	30
9. Base de datos y análisis del sesgo .....	30
10. Definición del Default .....	32
11. Presentación de Variables.....	33
12. Análisis descriptivo de Variables .....	33
13. Garantía .....	33
14. Historial Crediticio .....	34
15. Habito de Pago .....	34
16. Ocupación .....	35
17. Cargo.....	36
18. Genero .....	36
19. Estado Civil .....	37
20. Tipo Vivienda .....	37
21. Estrato.....	38
22. Personas a Cargo .....	38
23. Plazo .....	39
24. Rango Ingreso.....	39

25.	Rango Edad .....	40
26.	Rango Acierta .....	40
27.	Monto Crédito.....	41
28.	Antigüedad en la Cooperativa.....	41
29.	Construcción del Modelo .....	43
30.	Metodología .....	43
31.	Salidas del modelo seleccionado .....	43
	Pruebas de bondad de ajuste .....	43
32.	Omnibus Test:.....	43
33.	Test de Wald:.....	44
34.	Tabla de Clasificación:.....	45
35.	Curva Cor O ROC Curve:.....	46
36.	K-S TEST: .....	47
37.	Ecuación.....	48
38.	Score Distribution.....	50
	Conclusiones y recomendaciones.....	51
	Bibliografía.....	53

## Lista de Tablas

Tabla 1: Distribución de Asociados en default y no default .....	22
Tabla 2: Variables disponibles de la muestra .....	22
Tabla 3: Análisis descriptivo de la muestra .....	25
Tabla 4: Resultados de decisión (Matriz de confusión) .....	27
Tabla 5 Garantía .....	34
Tabla 6: Historia de Crédito .....	34
Tabla 7: Hábito de pago .....	35
Tabla 8: Ocupación .....	35
Tabla 9: Cargo .....	36
Tabla 10: Genero .....	37
Tabla 11: Estado Civil .....	37
Tabla 12: Tipo de Vivienda .....	38
Tabla 13: Estrato .....	38
Tabla 14: No personas a cargo .....	39
Tabla 15: Rango de plazo .....	39
Tabla 16: Rango de ingreso .....	40
Tabla 17: Rango de edad .....	40
Tabla 18: Rango acierta .....	41
Tabla 19: Rango monto de crédito .....	41
Tabla 20: Antigüedad en la Cooperativa .....	42
Tabla 21: Prueba ómnibus .....	44
Tabla 22: Variables en la ecuación .....	45
Tabla 23: Tabla de clasificación .....	46
Tabla 24: Estadístico KS .....	48
Tabla 25: Score Distribution .....	50
Tabla 26: Pronostico del modelo .....	51

## Lista de Graficas

Ilustración 1: Ejemplo curva de ROC .....	27
Ilustración 2: Estructura de la Curva ROC .....	28
Ilustración 3: Curva de ROC .....	47

# 1. Resumen

Las Cooperativas dentro del marco financiero colombiano son fundamentales para el otorgamiento de créditos a la población de los estratos socioeconómicos 1, 2 y 3, que tienen acceso restringido a la banca tradicional por ser en su mayoría informales laboralmente, bajo nivel de ingresos o en ocasiones por desconocimiento del manejo financiero. Las Cooperativas brindan la posibilidad de adquirir experiencia crediticia, pero a su vez se genera alto riesgo en el otorgamiento de dichos créditos, por no existir un modelo para el análisis y aprobación de estos, que minimice el default<sup>1</sup> de la cartera. Es por ello por lo que este trabajo se presenta como un aporte, para introducir técnicas estadísticas y econométricas (construcción de un modelo logit) para la toma de decisiones en la aprobación de créditos en una Cooperativa de Aporte y Crédito Colombiana.

En este trabajo se presenta un modelo logit aplicado a datos reales para un scoring de otorgamiento de créditos, como una propuesta alternativa para evaluar supuestos que son utilizados para evaluar el riesgo de crédito en las cooperativas. El modelo logit evalúa el enfoque de la teoría económica de la utilidad aleatoria con respecto a la decisión que enfrenta el Asociado de pagar o no el crédito. Todo lo anterior es a lo que se enfrenta el comité de crédito donde su objetivo gira en maximizar la utilidad del crédito otorgado para incrementar la rentabilidad de la cooperativa. Uno de los objetivos de desarrollar el score de otorgamiento de crédito, es seguir los lineamientos de la superintendencia de la economía solidaria, donde expone en la circular 015 de 2015 la implementación de los sistemas de riesgos en las Cooperativas, dentro de ellos el sistema de administración de riesgo de crédito, el cual debe contribuir a mejorar las decisiones de otorgamiento para minimizar el índice de morosidad<sup>2</sup> de la cooperativa. Es así como este trabajo se expone como

---

<sup>1</sup> Para la Superintendencia Financiera de Colombia financiera el default en la cartera de consumo se define como toda obligación en mora mayor a 90 días.

<sup>2</sup> EL índice de morosidad corresponde al porcentaje de las obligaciones en mora mayor a 30 días con respecto al total de la cartera de la cooperativa (saldo en mora > 30 días / total de la cartera)

una herramienta para minimizar las pérdidas esperadas<sup>3</sup> de la cooperativa, mejorar las herramientas para la toma de decisiones y contribuir desde la teoría económica y del riesgo de crédito para maximizar utilidades de la entidad.

Palabras clave: Default, modelo logit, scoring de crédito, riesgo de crédito, pérdida esperada

## **Abstract**

Cooperatives within the Colombian financial framework are fundamental for granting loans to the population of socioeconomic strata 1, 2 and 3, who have restricted access to traditional banking because they are mostly informal at work, low income or sometimes due to ignorance of financial management. The Cooperatives offer the possibility of acquiring credit experience, but at the same time high risk is generated in the granting of said credits, as there is no model for the analysis and approval of these credits, which minimizes the default of the portfolio. That is why this work is presented as a contribution, to introduce statistical and econometric techniques (construction of a logit model) for decision making in the approval of credits in a Colombian Contribution and Credit Cooperative.

This paper presents a logit model applied to real data for a credit grant scoring, as an alternative proposal to evaluate assumptions that are used to assess credit risk in cooperatives. The logit model evaluates the economic theory approach of random utility with respect to the decision faced by the Associate to pay the credit or not. All of the above is what the credit committee faces where its objective is to maximize the usefulness of the credit granted to increase the profitability of the cooperative. One of the objectives of developing the credit granting score is to follow the guidelines of the superintendence of the solidarity economy, where it sets out in circular 015 of 2015 the implementation of the risk systems in the Cooperatives, within them the system of credit risk management, which should contribute to improving grant decisions to minimize the delinquency rate of the cooperative. This is how this work is presented as a tool to minimize the expected losses of the cooperative, improve the tools for decision-making and contribute from economic theory and credit risk to maximize profits of the entity.

Keywords: Default, logit model, credit scoring, credit risk, expected loss

---

<sup>3</sup> La pérdida crediticia esperada es la estimación (aplicando una probabilidad ponderada, y no sesgada) del valor actual de la falta de pago al momento del vencimiento





## 2. Introducción

La justificación de este trabajo es dada por la función de intermediación que cumplen las cooperativas como alternativa al sistema financiero tradicional, su actividad se encuentra expuesta a un riesgo de crédito, cuyo inadecuado manejo y gestión puede dar origen a profundas crisis con consecuencias muy costosas en términos sociales, en especial para un sector que atiende a aquellos que no son incluidos por el sistema financiero tradicional y, que en el caso Colombiano, constituye la mayoría de la población.

Un ejemplo de ello es la crisis de fines de los años noventa que sufrió el cooperativismo en Colombia. Según los informes de la superintendencia de la economía solidaria (Supersolidaria) y la Confederación de Cooperativas de Colombia (Confecoop), la liberalización financiera y la fuerte entrada de flujos de capital generó que la cartera de créditos creciera a un ritmo acelerado durante la primera mitad de los años noventa. Luego, al tener excesos de recursos, es más difícil diferenciar los buenos proyectos de los malos, por lo que es más factible que se entreguen créditos con mayores riesgos, como consecuencia de las fallas en los mecanismos de control del riesgo de crédito en las entidades del sector solidario.

Así los reportes de la Confederación de Cooperativas de Colombia explican que como consecuencia de la crisis de 1997, alrededor de 50 cooperativas fueron liquidadas con el perjuicio causado a más de 1.000.000 de Asociados.

Se identificaron que los errores en gestión de riesgo fueron identificados como factores de vulnerabilidad e inestabilidad dentro de la actividad financiera y, por ende, desencadenantes de crisis financieras. Por ello analizar el riesgo de crédito e incluir metodologías como las de Basilea, permiten disminuir el riesgo de incumplimiento del pago de las obligaciones.

Por lo anterior la Cooperativas han venido desarrollando un análisis exhaustivo en el otorgamiento de los créditos tanto de consumo como en microcrédito, este último

ha venido ganando un terreno importante como dinamizador de la economía en los sectores sociales más vulnerables y que no cuentan con una estabilidad económica.

En el caso de la entidad “Cooperativa de Aporte y Crédito” se hace necesario el estudio y segmentación de los nichos de mercado, para generar metodologías y análisis estadísticos que permitan minimizar la pérdida esperada, por ello este trabajo se realiza como una solución al problema de decisión de otorgamiento de un crédito.

Todos los negocios requieren de información. Particularmente, el de intermediación de recursos es un negocio fundamentado en información. Ello se debe a la naturaleza propia de estos agentes, es decir, al cumplimiento de su papel económico, en el que transfieren recursos de un sector superavitario a un sector deficitario.

Dentro de un mismo mercado; vale anotar que en Colombia los “Establecimientos de Crédito” son quienes efectúan dicha labor de intermediación de recursos lo cual genera un elemento particular denominado Confianza”

A continuación, explicaremos cómo influye la información en el proceso de intermediación crediticia. Cuando los agentes prestan o colocan recursos deben conocer a la parte beneficiaria del préstamo; este conocer se traduce en confianza. Dicha confianza es la seguridad de que el ahorro del prestamista, manejado por el agente, será devuelto por el prestatario, más un margen de ganancia por la suma de dinero prestada y por un tiempo estipulado

Toda esta labor se traduce en confianza, la cual es obtenida gracias a la información; ello nos genera una visión del futuro que no es incierta, por eso constituye el motivo que determina la voluntad para crear o generar una obligación crediticia. Es justamente esta información la que permitirá al intermediario efectuar un cálculo del riesgo del negocio, para así poder determinar cuánto será el precio del dinero de dicho riesgo.

En todo este proceso o manejo de información, surge un problema que se denomina “Asimetría de Información”. Ésta se presenta cuando los participantes del negocio bancario no cuentan con información o la misma es insuficiente, sea sobre otros participantes o sobre sus intenciones (otros agentes o público en general). Por consiguiente, no se pueden adoptar decisiones adecuadas o acertadas y ello, al circunscribirse al mercado, se refleja en sus costos de transacción. Dicho en otras palabras, un mercado financiero que presente un nivel elevado de asimetrías de información genera costos de transacción que, de una u otra forma, conllevan a que no se presente una correcta asignación de recursos.

Básicamente, las asimetrías de información generan dos efectos nocivos en el mercado bancario y cooperativo: selección adversa y riesgo moral.

La selección adversa se presenta con anterioridad a la transacción. Podemos resumirla diciendo que ocurre cuando los individuos solicitantes de recursos representan un alto riesgo de crédito y por lo tanto, son los más interesados en conseguir uno; por ello actúan con mayor decisión o agresividad para obtener un crédito y, por consiguiente, están dispuestos a pagar mayores tasas de interés. El problema de selección adversa se presenta al existir asimetrías de información en el mercado crediticio; debido a ella, el agente prestatario no contará con suficiente información para saber quién constituye un buen o un mal crédito, por lo cual su decisión no será correcta.

El riesgo moral surge una vez desembolsados o entregado los recursos. En aquel momento hay cambio de incentivos en el prestatario, por ello podrá tomar diferentes actitudes dependiendo de cómo sea su comportamiento con respecto al riesgo; es decir, propenso, adverso o neutro. Si el prestatario es consciente que lo máximo que podrá perder, será el préstamo adquirido u otorgado y si a esto le sumamos una actitud de propensión al riesgo, ello constituirá un incentivo para usar los recursos de una forma temeraria o en proyectos arriesgados, donde las posibilidades de ganancia sea muy rentable pero incierta, como si estuviesen apostando (gambling)

El papel de los agentes crediticios (intermediarios de recursos) es lidiar con estos problemas. Ellos, al estar constituidos como agentes económicos, poseen toda una estructura que se especializa al punto de poder determinar con un bajo costo (es decir, estamos ante la presencia de un negocio con una economía de escala, debido a que para efectuar la operación de análisis para una persona, se requiere de la misma estructura operativa que realizarla para cien personas) quienes son sujetos considerados como buen o mal crédito; saber qué negocios son rentables, seguros o arriesgados, al igual que los riesgos que aquejan éstos; y también poder monitorear los préstamos concedidos para evitar un riesgo moral cuando se concede el crédito. Estas constituyen algunas de las virtudes de los “agentes quizás el papel de los agentes crediticios”.

Es por esta razón que en este trabajo centramos nuestros esfuerzos en tratar de explicar, que es sumamente importante establecer una herramienta para las cooperativas que permita una mejor medición a la hora de otorgar un préstamo ya que la información es limitada, poco precisa y si bien es cierto el mercado al cual atiende las cooperativas es diferente al mercado de los bancos en ambos casos se está expuesto a un deterioro de la cartera.

Las cooperativas se rigen por la ley 79 de 1988, y la ley 454 de 1998<sup>4</sup>. Es cooperativa la empresa asociativa sin ánimo de lucro, en la cual los trabajadores o los usuarios, según el caso, son simultáneamente los aportantes y los gestores de la empresa, creada con el objeto de producir o distribuir conjunta y eficientemente bienes o servicios para satisfacer las necesidades de sus asociados y de la comunidad en general.

El sector solidario contribuye al desarrollo económico y social colombiano. Entre otros aspectos, el sector solidario busca promover la participación de sus asociados

---

<sup>4</sup> La ley 454 del 4 de agosto de 1998 determina el marco conceptual que regula la economía solidaria, transforma el Departamento Administrativo Nacional de Cooperativas en el Departamento Administrativo Nacional de la Economía solidaria y crea la superintendencia de la economía solidaria.

y la ayuda mutua para el establecimiento de productos y servicios en pro de su base social.

Esto ha permitido que el sector de la economía solidaria sea pieza fundamental para el crecimiento de los productos financieros y el acceso a éstos a poblaciones menos favorecidas. Lo anterior contribuye a que los asociados a estas entidades manejen productos financieros con menos restricciones y más conocimiento del sistema financiero.

Es por ello por lo que el sector de la economía solidaria no es ajeno al riesgo de crédito, teniendo en cuenta que dentro de los portafolios de las Cooperativas se encuentra el producto de crédito, como uno de los principales mecanismos para el crecimiento de sus asociados y su entorno económico.

De ahí la Superintendencia de la Economía Solidaria mediante la circular básica contable y financiera, determina que la recuperación de la cartera en las condiciones pactadas inicialmente es uno de los factores principales para mantener la entidad en el tiempo.

Es por lo que la administración del riesgo de crédito cobra importancia para minimizar las pérdidas y generar una mayor participación en el mercado financiero colombiano. El gobierno nacional toma los principios de Basilea para regular la actividad crediticia para las Cooperativas en Colombia a través de la Superintendencia Financiera con modelos de referencia para la cartera de consumo y la Superintendencia de Economía Solidaria con modelos de provisión por deterioro en días de mora; esta última viene ejerciendo control sobre el otorgamiento y seguimiento de los créditos en las Cooperativas de aporte y crédito, a través de la Circular Básica Contable y Financiera, y actualmente trabaja sobre el Sistema de Integral de Administración de Riesgos SIAR con la circular 015 de diciembre 30 de 2015. Sin embargo, en nuestras investigaciones encontramos que las Cooperativas analizan sus créditos y los otorgan con la metodología de expertos y en algunos casos acompañados simplemente de datos estadísticos sencillos del

comportamiento interno con la entidad<sup>5</sup> o capacidad de pago; que en algunos casos no cumplen con referencias como las dadas por el banco de la república en cuanto a carga financiera y nivel de endeudamiento<sup>6</sup>, es así como se requiere mayores indicadores y herramientas a la hora de aprobar un crédito. Por ello este trabajo propone un modelo de scoring para mejorar la tomada de decisiones al momento de la aprobación.

Con este fin se vienen analizando métodos<sup>7</sup> de calificación de cartera y scores de crédito para determinar el punto de Default de las obligaciones y sus determinantes, es así como Edward Altman (1968) fue el primero en proponer una técnica de puntaje para pronosticar el riesgo de quiebra en una empresa con su técnica del Z-score. En el modelo clásico de Altman, un método de puntaje es un enfoque enfatizado en predecir incumplimientos y no explicar por qué ocurren o no. Los elementos básicos son: 1) Un conjunto de “variables” y sus correspondientes atributos que califican y/o cuantifican el peso de la variable en un deudor; 2) Los “puntos” asociados a cada variable según apliquen a los solicitantes, y 3) un “umbral” que separe a los deudores buenos de los malos

En la jerga de las técnicas de **puntaje o score**, las características se refieren a las preguntas que se les hace a los solicitantes de crédito, mientras que los atributos son las posibles respuestas, ya sea que las contesten los solicitantes directamente en una solicitud de crédito o que se obtengan de otras fuentes como las centrales de información (Datacredito o TransUnión “Cifin”)<sup>8</sup>.

Con lo anterior se busca identificar cuáles son las variables que determinan la aprobación de un crédito, minimizando el riesgo de crédito (default del crédito) y de esta manera contribuir a:

---

<sup>5</sup> Observatorio No 14 Confecoop enero 2010, Observatorio No 45 Confecoop marzo 2018

<sup>6</sup> Los informes de estabilidad financiera del banco de la república de Colombia, plantean que el 50% de los colombianos esta endeudado 5.7 veces su ingreso y que este indicador en un 95% de los casos es inferior a 19 veces.

<sup>7</sup> Modelo de referencia de consumo de la Superintendencia Financiera de Colombia SFC

<sup>8</sup> Datacredito y Cifin son centrales de información de carácter privado que se especializan en almacenar y administrar la información concerniente a cómo las personas y empresas han pagado los préstamos o servicios financieros es decir, su comportamiento de pago

- Crear un modelo econométrico (score) que mida el riesgo de crédito y que permita otorgar créditos de consumo de una manera más ágil, con criterios técnicos de acuerdo con la información cualitativa y cuantitativa de los Asociados vinculados a la Cooperativa.
- Minimizar la pérdida esperada de la *Cooperativa*.
- El modelo contribuirá a mejorar los indicadores de índice morosidad (menor porcentaje) y menores provisiones de cartera (mejorando las utilidades o excedentes de la Cooperativa).

Para construir el modelo logit se tuvieron en cuenta medidas de bondad de ajuste como la curva de ROC o ROC CURVE, Test de Wald, Omnibus Test y KS TEST, los cuales son explicados más adelante.

Los datos cualitativos y cuantitativos (tabla 2 variables disponibles) tenidos en cuenta para la construcción del modelo son una muestra de 12.257 asociados de la entidad Cooperativa (entre enero de 2011 y noviembre de 2015).

### **3. Marco Teórico**

Los modelos de riesgo de crédito vienen a ser estudiados desde Beaver (1966) quien bajo un análisis multivariado enfocó su trabajo a determinar cómo podía influir la insolvencia en una entidad, luego Altman (1968) que bajo un análisis discriminante multivariado trabajo el problema de predicción de quiebra corporativa.

Fue Orgler (1970) el que manejo la técnica del análisis de regresión lineal para el análisis de los créditos comerciales, donde sólo revisó los créditos vigentes. Así mismo Orgler en 1971 usó un enfoque de regresión para evaluar los créditos de consumo, con el cual llegó a la conclusión de que la información no incluida en el formulario de solicitud tenía una mejor capacidad de predicción que la información proporcionada inicialmente por el solicitante, al evaluar la calidad de los préstamos futuros.



Chandler y Coffman (1979), hicieron un análisis comparativo entre el scoring de crédito y el juicio de un evaluador de créditos, donde ambos sistemas indicaban, en promedio, cuál de los dos grupos en que se clasificaban los individuos, tenía mejor comportamiento. Con el trabajo de Ohlson (1980), el modelo logit se convirtió en una metodología estándar para pronosticar la probabilidad de quiebra o el estado de Default de los créditos; este modelo (logit) tiene características específicas como son la elección de las variables explicativas, los supuestos relacionados con la forma funcional de las variables, la familia de distribución del término de error, las varianzas de los términos de error y los parámetros (Kukuk y Roennberg, (2013)).

Abdou (2009) y Greene (1998) utilizaron análisis discriminante y regresiones logística y probit, para investigar la eficiencia, en términos de la clasificación correcta de los deudores (buenos y malos), de los modelos tipo scoring para créditos de consumo y comercial, respectivamente. Así mismo el banco de la república en el borrador de economía 466 de 2007, expone los determinantes de la probabilidad de incumplimiento de las empresas colombianas utilizando el modelo logit.

La superintendencia financiera ha definido en términos de provisión para los bancos, un modelo que determina la probabilidad de incumplimiento y la pérdida dado el incumplimiento, de acuerdo a la calificación del crédito y sus garantías (circular básica contable y financiera – circular externa 100 de 1995). Lo anterior nos permite evidenciar que la probabilidad de incumplimiento se puede llegar a determinar a partir de modelos estadísticos, pero que en el sector solidario y propiamente para las cooperativas de aporte y crédito, puede ser diferente por su actividad y nicho de mercado que tienen las entidades con estas características.

#### **4. El scoring de crédito y su avance en Colombia**

Un modelo de scoring o calificación de crédito (puntaje), es una herramienta que construyen las entidades financieras para la toma de decisiones de aprobación o negación de las solicitudes de créditos de los asociados o clientes. Lo anterior

teniendo en cuenta que la entidad que otorga el crédito requiere minimizar la morosidad que pueda presentar la cartera de crédito.

Desde el año 2004 estos modelos generan gran importancia con el acuerdo sobre legislación dado desde el comité de Basilea.

En Colombia el Banco de Republica a través de los borradores de economía, ha hecho análisis en los determinantes de probabilidades de incumplimiento de las empresas (borrador 466-2007) y determinantes del riesgo del crédito comercial en Colombia (reporte de estabilidad financiera en marzo de 2010), informes que han contribuido a generar conciencia que los créditos de consumo de las personas naturales, también se pueden determinar variables para minimizar la probabilidad de incumplimiento.

## **5. El modelo logit**

El modelo logit se genera en 2 fases, la primera genera las estimaciones de la probabilidad de un evento y la segunda fase identifica los ítems que determinan esa probabilidad, así como el peso de cada variable.

Para entender los modelos econométricos de una manera sencilla, se plantea una variable dependiente y la relación que llega a tener con las variables independientes o explicativas. Como función de estimación el logit utiliza la regresión logística (más robusta que la regresión lineal), permitiendo generar las variables más relevantes dentro del modelo y segmentarlas de acuerdo con su nivel de importancia.

## **6. El scoring como mecanismo para determinar la probabilidad de Incumplimiento.**

En el año (1992) Gracia-Diez y serrano definieron la ecuación teórica para la toma de decisión de otorgar un crédito:

Ecuación

$$(1 - P_i) + r V_{S_i} + P_i(-wV_{S_i}) \geq 0$$

$V_{S_i}$ : Valor solicitado por el Asociados i

$P_i$ : Probabilidad de que el asociado i, presente algún tipo de incumplimiento durante la vigencia del crédito.

$r$ : tasa de interés nominal determinada para el tipo de crédito.

$w$ : Fracción del valor solicitado que se pierde por incumplimiento durante la vigencia del crédito.

La ecuación anterior se utiliza para determinar el valor máximo de crédito a otorgar al Asociado o para distribuir de manera óptima el monto de crédito a otorgar. Esto nos permite estimar la probabilidad de incumplimiento del crédito a otorgar y generar un nivel de probabilidades estimadas para mejorar la toma de decisiones al momento de la aprobación.

El objetivo del modelo es poder discriminar los Asociados que ingresan a Default de los que no; de una forma más objetiva a partir de las variables cualitativas y cuantitativas de cada individuo.

## 7. Metodología

El modelo que se utilizará es un modelo logit, el cual modela una estructura de correlación y/o heterocedasticidad entre las variables (cualitativas y cuantitativas de los Asociados de la Cooperativa), determinando de esta manera comportamientos que puedan establecer el default de los créditos.

Se requiere analizar la probabilidad de no pago de un crédito (default del crédito otorgado), a partir de variables cualitativas y cuantitativas con la estimación de un modelo logit, que permitirá definir las probabilidades (regresión) de cada variable.

Actualmente no hay un trabajo de logit aplicado a las Cooperativas de Aporte y Crédito en Colombia, el cual podría influir de manera positiva para el otorgamiento y bancarización de la población colombiana, en las Cooperativas de aporte y crédito.

El modelo de la Superintendencia Financiera calcula la pérdida esperada para realizar las respectivas provisiones individuales, este trabajo pretende anticipar el default de la cartera a partir de variables cualitativas y cuantitativas al momento del otorgamiento ampliando el número de variables.

La variable explicada o variable dependiente, será *default (probabilidad de mora mayor a 90 días)*, que es el número de días en el cual la Cooperativa considera que se entra en mora o en incumplimiento en las obligaciones. Ésta es utilizada para evaluar la salud de la cartera y mediante este análisis administrar mejor el riesgo de probabilidad de incumplimiento. El *default* para la Cooperativa es de 90 días (según metodología<sup>9</sup> aplicada), un *default* bueno será aquel que sea menor o igual a este número de días y uno malo aquel que se encuentre por encima. Además, es de tener en cuenta que el nivel de *default* se determina de acuerdo con las políticas y metodologías de cada entidad y el apetito de riesgo.

El cálculo de la probabilidad de incumplimiento se realiza mediante la siguiente fórmula:

$$Puntaje = \frac{e^z}{1 + e^z}$$

---

<sup>9</sup> La metodología empleada por la cooperativa para calcular el default es el modelo de cadenas de Márkov, el cual mide los rodamientos de la cartera y determina el punto de corte, el cual se haya donde la probabilidad de deterioro es mayor que la probabilidad de recuperación.

Donde, Z varía de acuerdo a las características particulares del crédito-deudor como se muestra a continuación.

$$Z = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 \dots + \beta_n X_n$$

Donde:

La variable dependiente “Z” indica que a medida que las probabilidades de incumplimiento aumentan, los niveles de riesgo de la cartera crecen.

Los parámetros por estimar,  $\beta_{1...}, \beta_n$  corresponden al peso relativo de cada variable en el modelo y  $\beta_0$  que es el promedio del nivel de probabilidad de incumplimiento.

Las variables independientes  $X_{1...}, X_n$  corresponden a cada una de las características cualitativas y cuantitativas de los Asociados – deudores.

En todo el proceso de construcción del modelo se utilizarán dos metodologías:

- Tablas de contingencia: Para realizar el análisis descriptivo y posterior categorización de las variables de entrada y
- Regresión Logística: Con base a las variables analizadas previamente, se dará inicio al proceso de construcción del modelo mediante regresión logística (logit), mediante la técnica de Forward Condicional, que consiste en ingresar una a una cada variable al modelo. dándole prioridad a las que registran mayor predictibilidad.

### **a) Datos**

La base de datos a utilizar es de la *Cooperativa de Aporte y Crédito* que contiene el histórico de seguimiento de créditos de 12.257 asociados de la entidad, (entre enero

de 2011 y noviembre de 2015), donde se identifican variables con relación a cada registro del Asociado de la entidad.

**TABLA 1: DISTRIBUCIÓN DE ASOCIADOS EN DEFAULT Y NO DEFAULT**

<b>DISTRIBUCIÓN DE LOS ASOCIADOS EN DEFAULT Y NO DEFAULT</b>			
	No - default	Default	Total
<i>Número de Créditos</i>	11,666	590	12,257
<i>Porcentaje</i>	95%	5%	100%

**b) Variables cualitativas y cuantitativas del Modelo Econométrico**

- Las variables cualitativas y cuantitativas que son elegidas para calcular el modelo de la Cooperativa son:

**TABLA 2: VARIABLES DISPONIBLES DE LA MUESTRA**

<b>VARIABLES DISPONIBLES EN LA MUESTRA</b>			
Factor	Variable	Nivel de Medición	Fuente
<b>Crediticio</b>	MONTO Valor del Crédito Solicitado	Numérica en pesos	Primaria
	PLAZO Plazo del Crédito Solicitado	Numérica en meses	Primaria
	GARANTIA Tipo de Garantía que respalda el crédito	Cualitativa Nominal Aportes, Codeudor. G. Admisible, Personal.	Primaria
<b>Financiero</b>	INGRESOS Ingresos mensuales del Asociado	Numérica en pesos	Primaria
	ENDEUD. No de veces (Deuda vs salario)	Numérica	Primaria
	CARGA FINANCIERA Porcentaje (cuotas + descuentos vs salario)	Numérica	Primaria

<b>Comportamiento Crediticio</b>	HISTORIAL CREDITICIO Calificación del asociado de acuerdo a historial crediticio con parámetros de la Cooperativa	Cualitativa Nominal Alto, Medio, Bajo, Nulo	Primaria
	HÁBITO DE PAGO Calificación del asociado de acuerdo a historial crediticio con parámetros de la Cooperativa	Cualitativa Nominal Alto Riesgo, Bueno, Aceptable, Regular, Desconocido	Primaria
	PUNTAJE ACIERTA Calificación de la central de información de acuerdo a historial crediticio en el sector financiero	Numérica 150 - 950	Secundaria
<b>Laboral</b>	OCUPACIÓN Ocupación del Asociado	Cualitativa Nominal - Trabajador contrato Indef. - Trabajador contrato Fijo. - Pensionado. - Microempresario. - Prestación de Servicios - Informal - Docente Nombrado - Docente Provisionalidad - Obra o Labor Contratada - Empresario con Empleados	Primaria
	TAMAÑO DE EMPRESA Tamaño de la empresa	Cualitativa Nominal Grande, Mediano, Pequeño, Micro, Informal.	Primaria
	CARGO Tipo de cargo del Asociado	Cualitativa Nominal - Aux o Técnico - Concejal - Director - Docente - Gerente - Informal - Pensionado - Profesional - Vendedor	Primaria
<b>Socio - Demográfica</b>	SEXO Genero del Asociado	Cualitativa Nominal Femenino - Masculino	Primaria
	EDAD Edad del Asociado al momento de solicitar el crédito	Numérica en años	Primaria

	ESTRATO Estrato socio económico	Cualitativa Ordinal De 1 a 6	Primaria
	TIPO VIVIENDA Tipo de Vivienda del Asociado	Cualitativa Nominal - Alquilada - Anticresada - Familiar - Propia	Primaria
	ESTADO CIVIL Estado civil del Asociado	Cualitativa Nominal - Casado - Separado - Soltero - Unión Libre - Viudo	Primaria
	NIVEL DE EDUCACIÓN Nivel de estudios del Asociado	Cualitativa Ordinal - Primaria - Bachiller - Técnico - Tecnólogo - Profesional - Especialización - Magister - Doctorado	Primaria
	PERSONAS A CARGO Número de personas a cargo que tiene el Asociado	Cualitativa Ordinal 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10	
	ANTIGÜEDAD EN LA COOP. Antigüedad como Asociado de la cooperativa al momento de realizar el crédito	Numérica en meses 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10	Primaria
<b>ECONOMICO</b>	SECTOR ECONOMICO Sector económico de la actividad que desarrolla el Asociado o la empresa donde labora.	- Servicios - Comercio - Gobierno - Agricultura - Transporte - Textil - Otros	Primaria
	AGENCIA	Cualitativa Nominal	Primaria



Nombre de la agencia donde el Asociado tramito el crédito

- Pampalinda
- Buenaventura
- Floralia
- Pasto
- Tumaco
- Samaniego
- Ipiales
- Popayán
- Pereira

El software utilizado para generar el modelo es “SPSS”

### c) Análisis descriptivo de la muestra

A continuación, se presenta un análisis descriptivo de la muestra de los Asociados a los que se les concedió un crédito

**TABLA 3: ANÁLISIS DESCRIPTIVO DE LA MUESTRA**

## **ANALISIS DESCRIPTIVO DE LA MUESTRA**

	No- Default (11.666 Observaciones)			Default (590 Observaciones)		
	Media	Mediana	Desviación Estándar	Media	Mediana	Desviación Estándar
MONTO	\$ 4,819,105	\$ 3,000,000	\$ 5,644,685	\$ 5,960,488	\$ 4,000,000	\$ 5,737,802
PLAZO	26	22	16	30	29	14
INGRESOS	\$ 1,701,351	\$ 1,246,751	\$ 1,772,433	\$ 1,739,680	\$ 1,220,450	\$ 1,901,276
ENDEUDAMIENTO No VECES	2.63	0.41	4.45	2.39	0.24	4.34
ACIERTA	603	717	269	542	658	275
EDAD AL DESEM	46	44	14	40	38	13
ANTIG COOP	37	19	132	30	1	155

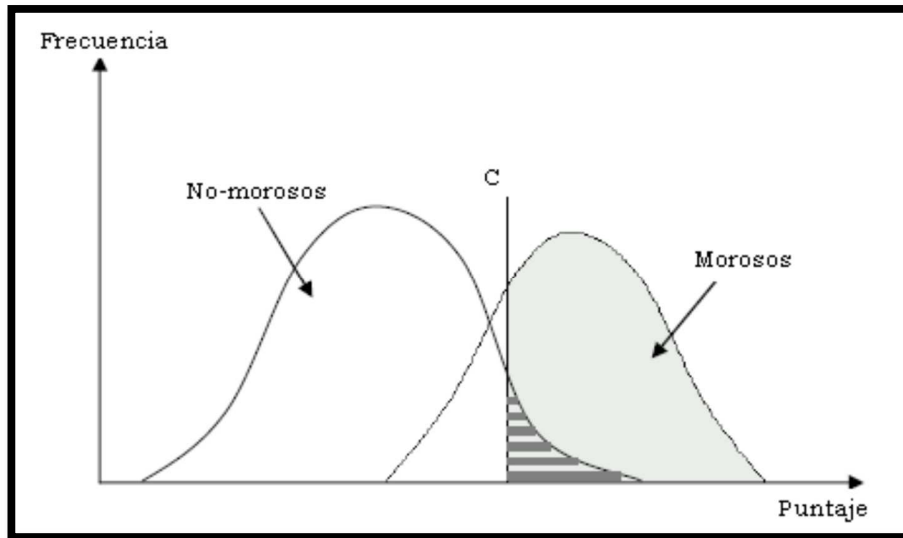
La tabla anterior proporciona el análisis descriptivo de las variables numéricas. En general, se observa que para la variable Endeudamiento, los Asociados que cayeron en Default presentan unos valores menores en media y mediana que los que no llegaron a default, lo cual representa una situación contraria a lo que se

esperaría. En cuanto al Valor del Crédito (en pesos), los Asociados que cayeron en Default tienen un valor promedio de crédito muy similar al de los que no entraron en default, aunque las medianas son un poco diferentes, y los Asociados de No-Default tienen una desviación estándar mayor, lo anterior podría inferir que el aumento de montos sin la debida experiencia del deudor trae consigo un mayor riesgo de crédito. El plazo (en meses), no muestra diferencia significativa en cuanto a la mediana, siendo un poco mayor la media para los que cayeron en Default. El valor de los Ingresos (en pesos), indica que es mayor en promedio para los No-Default, mientras que los de Default tienen una mediana mayor. La Edad (en años), la Antigüedad como Asociado de la Cooperativa (en años) presentan valores menores, tanto en la media como en la mediana, para los que entraron en Default. El acierta que es el puntaje de la central de información y que relaciona el comportamiento de hábito de pago de las obligaciones externas, tiene un menor valor para los que cayeron en Default, sugiriendo que podría haber una influencia sobre la variable dicotómica estudiada.

#### **d) Curva de ROC**

Para representar el poder discriminatorio de un modelo scoring, por lo general se utiliza la curva ROC (acrónimo de Receiver Operating Characteristic), la cual ilustra la posible distribución de los puntajes arrojados por el scoring para los clientes que alcanzan el estado de Default (Morosos) en contraste con los que no llegan a default (No-morosos). En un modelo perfecto de clasificación se podrían distinguir perfectamente los dos tipos de clientes (Morosos y No-morosos), pero como en la realidad esto no es posible, las dos distribuciones se traslapan, como se muestra en la ilustración 1, donde "C" es un valor de corte (o umbral) que establece una regla de decisión para dividir los deudores potenciales en morosos y no-morosos.

## ILUSTRACIÓN 1: EJEMPLO CURVA DE ROC



Considerando el problema de predicción binaria, hay posibles resultados, resumidos en:

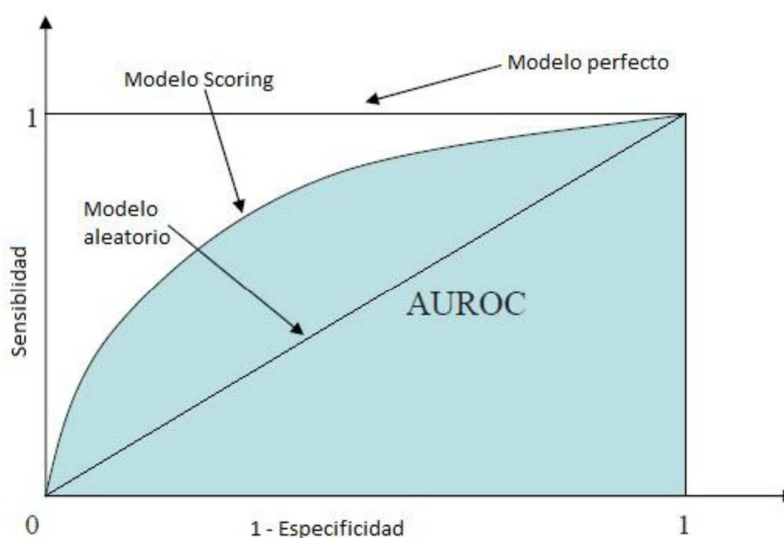
**TABLA 4: RESULTADOS DE DECISIÓN (MATRIZ DE CONFUSIÓN)**

		Observado	
		Defaults	No-defaults
Puntaje	Por encima de "C"	Predicción de los verdaderos positivos (Sensibilidad)	Predicción de los falsos positivos (1-Especificidad)
	Por debajo de "C"	Predicción de los falsos negativos (1-Sensibilidad)	Predicción de los verdaderos negativos (Especificidad)

Si un cliente con un puntaje mayor que "C" llega a default o uno con un puntaje menor que "C", no llega a default, entonces la predicción del scoring es correcta. A la proporción de predicciones correctas de Default se le llama sensibilidad (aciertos) y a la proporción de predicciones correctas de No-default, especificidad (falsas alarmas). Para un punto de corte dado "C", lo ideal es que un scoring tenga una alta

sensibilidad y una alta especificidad. Estadísticamente, la predicción de los falsos positivos (1-especificidad) se conoce como Error Tipo I, el cual se define como el error de rechazar la hipótesis siendo ésta verdadera. La predicción de los falsos negativos (1-sensibilidad) es llamado el Error Tipo II, que se comete cuando se acepta la hipótesis nula siendo falsa.

## ILUSTRACIÓN 2: ESTRUCTURA DE LA CURVA ROC



La curva ROC está definida por la proporción de falsos positivos en el eje X (1-Especificidad) y la proporción de verdaderos positivos (Sensibilidad) en el eje Y, como se muestra en la ilustración 2. Cada resultado de predicción de la tabla 4 (matriz de confusión) representa un punto en el espacio de la curva ROC.

Se denota  $D$  como el conjunto de morosos,  $ND$  el conjunto de no morosos y con  $T$  el conjunto total de clientes. Sea  $G_D$  la distribución de los puntajes de los deudores morosos y  $G_{ND}$  la distribución de los puntajes de los no morosos. Para cualquier punto de corte  $c$ , se tiene que  $F_D(c) = Prob(G_D \geq c)$  es la sensibilidad y  $F_{ND}(c) = Prob(G_{ND} \geq c)$  corresponde a 1- especificidad. Entonces la curva ROC se construye con la variación de  $c$  sobre los posibles puntajes arrojados por el sistema de puntuación.

En la ilustración 2 se observa como la curva ROC empieza en el punto (0,0) y termina en el punto (1,1). Si el valor de corte está en el rango de puntaje de los no morosos, entonces la sensibilidad es menor que 1, la especificidad es igual a 1 y la curva ROC pasa por el punto (0,0). De manera opuesta, si el valor de corte está en el rango de puntaje de los morosos, entonces la sensibilidad es 1, la especificidad es menor que 1 y la curva ROC se incrementa monótonamente hasta el punto (1,1). En un modelo perfecto, los morosos y no morosos serían separados perfectamente, y la predicción se situaría en el punto (0,1) de la curva ROC, representando un 100% de sensibilidad (ningún falso negativo) y un 100% de especificidad (ningún falso positivo). Por el contrario, una clasificación totalmente aleatoria daría un punto a lo largo de la línea diagonal, comprendida desde el punto (0,0) hasta (1,1), e indicaría que el scoring no tiene ningún poder discriminatorio.

El área bajo la curva ROC (AUROC) proporciona una medida del poder discriminatorio del modelo. Para un modelo aleatorio sin poder discriminatorio, el AUROC es de 0.5, y para un modelo perfecto, AUROC es 1. Continuando con la notación anterior, el AUROC se define como:

$$AUROC = \int_0^1 F_D(c) dF_{ND}(c)$$

Para la revisión del score a calcular se empleará la curva ROC (acrónimo de Receiver Operating Characteristic), la cual ilustra la posible distribución de los puntajes arrojados por el scoring, para los Asociados que alcanzan el estado de Default<sup>10</sup> (Morosos) en contraste con los que no llegan a default (No-morosos).

En un modelo perfecto de clasificación se podrían distinguir perfectamente los dos tipos de Asociados (Morosos y No-morosos), pero como en la realidad esto no es

---

<sup>10</sup> Para la superintendencia financiera, el default en los créditos de consumo es todo crédito que alcanza una mora mayor a 90 días en el pago de sus cuotas.

posible, las dos distribuciones se traslapan, como se muestra en ilustración 1, donde “C” es un valor de corte (o umbral) que establece una regla de decisión para dividir los deudores potenciales en morosos y no-morosos.

## **8. Estimación del Modelo**

El objetivo principal de estimar este modelo de scoring, es encontrar una herramienta que le permita a la Cooperativa diferenciar a los asociados que tienen una probabilidad de entrar en default, donde se pretende explicar pagar o no el crédito.

Moreno (2013) explica el modelo logit como el modelo de probabilidad, donde los individuos se enfrentan a varias alternativas que dependen si y solo si de la utilidad que obtiene el individuo al elegir una alternativa.

## **9. Base de datos y análisis del sesgo**

La base para la estimación del modelo fue definida de acuerdo con los siguientes parámetros:

- Desembolsos realizados desde enero 2005
- Operaciones que registraran al menos 12 meses de vigencia, con el fin de poder visualizar un comportamiento de pago que permitiera establecer el hábito de pago

Para efectos de construcción del modelo se seleccionó el 70% de la data mediante la técnica de muestreo aleatorio, mientras que el 30% restante fue utilizado en la fase de evaluación de este.

Hausman y Wise (1977) y Maddala (1983), demostraron que las estimaciones de los parámetros de un modelo de regresión obtenidas a partir de una muestra sesgada son estimaciones inconsistentes de los parámetros poblacionales, debido a que la esperanza de las perturbaciones del modelo no es cero, sino una función de las variables explicativas del mismo.

Gracia-Diez y Serrano (1992) manifiestan que el análisis de la muestra de créditos concedidos, se toma sin tener en cuenta el hecho de que está sesgada. Toda vez que se asume que la entidad concede los créditos a criterio de experto. No obstante en los modelos tipo scoring, el sesgo en la muestra, se trata de corregir asumiendo como se habrían comportado los individuos a los que se les rechazó el crédito si se les hubiera concedido, es decir, prediciendo la probabilidad de default.

Wynand y Bernard (1981) desarrollaron un modelo, que fue aplicado por Boyes et al. (1989) en el análisis de créditos de consumo.

Las ecuaciones estructurales son:

Ecuación para el Default:

$$D = \beta' x_i + \varepsilon_i \quad (\text{Ec. 1})$$

$$D_i = 1 \text{ si y sólo si } D > 0, \text{ y } 0 \text{ en otro caso}$$

Para la aprobación de créditos

$$C = \gamma' v_i + \omega_i \quad (\text{Ec. 2})$$

$$C_i = 1 \text{ si y sólo si } C > 0, \text{ y } 0 \text{ en otro caso}$$

Regla de muestreo

$$D_i \text{ y } x_i \text{ son observadas solamente si } C_i = 1 \quad (\text{Ec. 3})$$

$$C_i \text{ y } \omega_i \text{ son observadas para todos los solicitantes de crédito}$$

Selectividad

$$[\varepsilon_i, \omega_i] \sim N_2 [0,0,1,1, \rho_{\varepsilon\omega}]. \quad (\text{Ec. 4})$$

Esta selectividad es planteada por Green (1992) en un modelo estadístico, para la calificación crediticia.

El vector de atributos,  $v_i$  contiene los factores usados en la decisión de aprobación. La probabilidad de interés es la probabilidad de default dado que la solicitud de crédito fue aprobada, la cual es:

$$Prob [D_i = 1 | C_i = 1] = \frac{\phi_2 [\beta' x_i, \gamma' v_i, \rho]}{\phi [\gamma' v_i]} \quad (\text{Ec. 5})$$

donde  $\phi_2$  es la probabilidad acumulada de la normal bivariada. Si es igual a 0, la selección no tiene ninguna consecuencia y el modelo incondicional descrito anteriormente es apropiado.

En el mismo documento Green (1992) menciona que las contrapartes de los efectos marginales mencionados anteriormente son

$$\frac{\partial \phi_2 (\beta' x_i, \gamma' v_i, \rho) / \phi (\gamma' v_i)}{\partial x_i} = \theta | C_i = 1$$

Este modelo fue desarrollado por Wynand y Bernard (1981) y recientemente aplicado a un análisis de préstamos al consumidor por Boyes, et. Alabama. (1989)

## 10. Definición del Default

De acuerdo con las matrices de transición<sup>11</sup> elaboradas por la Cooperativa, la entidad identificó que el punto de default por mora de la cartera de consumo, corresponde al intervalo de 61-90 días de mora. Es decir, que toda operación de la base de datos que haya registrado mora a corte superior a 90 días se identificó como mala en el momento de realizar el modelo.

---

<sup>11</sup> Matrices de rodamiento o transición metodología recomendada por la Superintendencia Financiera de Colombia SFC



## 11. Presentación de Variables

Para realizar el modelo de otorgamiento de la Cooperativa se consolidó una base de datos más de 30 campos donde se encuentran variables de Identificación, comportamiento en centrales, variables del crédito solicitado, sociodemográficas y Financieras.<sup>12</sup>

**Variables Identificación:** Número de Identificación, Nombre, Dirección, teléfono, correo.

**Variables Sociodemográficas:** Edad, Nivel Académico, Estrato, Tipo de Vivienda, Genero, Personas a Cargo, Ocupación.

**Variables Financieras:** Ingresos, Egresos, Liquidez, Monto de la obligación, Plazo, Tamaño de la Empresa.

**Variables Centrales de Crédito:** Endeudamiento, Carga Financiera, Acierta, Historial Crediticio, Hábito de Pago.

## 12. Análisis descriptivo de Variables

Para cada una de las variables presentes en la base, se realizó análisis descriptivo y de niveles de riesgo con el fin de identificar las de mayor poder predictivo. En el caso de variables continuas se realizó segmentación por percentiles con el objeto de hacer más fácil en análisis la revisión de los niveles de riesgo; mientras que en variables categóricas solo se analizó la probabilidad de Default.

El análisis se realizó teniendo en cuenta todas las variables entregadas en la base de datos, sin embargo, para efectos del presente documento solo se mostrarán los resultados de las más relevantes.

## 13. Garantía

Revisando los niveles de riesgo por tipo de garantía, se puede observar que Codeudor y Garantía Admisible son las de mayores probabilidades con 7.46% y

---

<sup>12</sup> El detalle de los campos utilizados en la base de datos se encuentra en la base de datos en el Excel.

7.79%. En contraste se observan los aportes, con probabilidad de incumplimiento de 0.93%.

**TABLA 5 GARANTÍA**

Garantía	NO DEFAULT	DEFAULT	TOTAL	PD
Aportes	2,131	20	2,152	0.93%
Codeudor	5,038	406	5,444	7.46%
G. Admisible	71	6	77	7.79%
Personal	4,426	158	4,584	3.45%
<b>Total general</b>	<b>11,666</b>	<b>590</b>	<b>12,257</b>	<b>4.81%</b>

## 14. Historial Crediticio

Esta variable define el nivel de historial crediticio reportado en centrales de información, donde aquellos con bajo historial presentan los mayores niveles de riesgo, con probabilidad de incumplimiento de 7,15%. En contraste se logra observar, que aquellos con alto historial registran la probabilidad más baja de la base con 3.60%.

**TABLA 6: HISTORIA DE CRÉDITO**

Historia Crediticio	NO DEFAULT	DEFAULT	TOTAL	PD
Alto	4,310	161	4,472	3.60%
Bajo	1,910	147	2,057	7.15%
Medio	4,784	248	5,032	4.93%
Nulo	662	34	696	4.89%
<b>Total general</b>	<b>11,666</b>	<b>590</b>	<b>12,257</b>	<b>4.81%</b>

## 15. Hábito de Pago

La variable hábito de pago mide la calidad de las obligaciones del asociado y que fueron reportadas a centrales de información. Como era de esperar, el comportamiento aceptable tiene Probabilidad de incumplimiento o default de 1.37%, siendo los de menor riesgo.

Aquellos asociados que no registraron información y los que presentaron un comportamiento regular, fueron los de los mayores niveles de riesgo con 6.99% y 9.84% respectivamente.

**TABLA 7: HÁBITO DE PAGO**

Habito de pago	NO DEFAULT	DEFAULT	TOTAL	PD
Aceptable	287	4	292	1.37%
Alto Riesgo	292	15	307	4.89%
Bueno	9,748	447	10,195	4.38%
Desconocido	652	49	701	6.99%
Regular	687	75	762	9.84%
<b>Total general</b>	<b>11,666</b>	<b>590</b>	<b>12,257</b>	<b>4.81%</b>

## 16. Ocupación

**TABLA 8: OCUPACIÓN**

Ocupacion	NO DEFAULT	DEFAULT	TOTAL	PD
Cooperativo	2		3	0.00%
Docente	3		3	0.00%
Docente En Provisionalidad	155	7	162	4.32%
Docente Nombrado	447	15	462	3.25%
Empresario Con Empleados	14	1	15	6.67%
Informal	813	25	838	2.98%
Microempresario	1,879	95	1,974	4.81%
Obra O Labor Contratada	125	8	133	6.02%
Pensionado	2,217	61	2,278	2.68%
Prestación De Servicios	1,114	55	1,169	4.70%
Trabajador Contrato Fijo	2,161	133	2,294	5.80%
Trabajador Contrato Indefinido	2,736	190	2,926	6.49%
<b>Total general</b>	<b>11,666</b>	<b>590</b>	<b>12,257</b>	<b>4.81%</b>

La ocupación con mayor numero de asociados dentro de la base es Trabajador con contrato indefinido con 2.926, seguida de Trabajador Contrato Fijo con 2.294 y estas representan el 42.5% del total base. Estos consolidan probabilidad de incumplimiento de 6.49% y 5.80% respectivamente.

La Ocupación con los menores niveles de riesgo es Pensionado con 2,68% y estos representan el 18.6% de la base. Otro punto que llama la atención es que los informales registran bajos niveles de riesgo pues su PD es de 2.98%.

## 17. Cargo

El cargo con mayor nivel de riesgo es Auxiliar o Técnico con 6,41%, este comportamiento puede estar influenciado por la corta edad que tienen estas personas y su nivel de ingreso. Estos representan el 35,3% de la base.

**TABLA 9: CARGO**

Cargo	NO DEFAULT	DEFAULT	TOTAL	PD
Aux o Tecnico	4,044	277	4,322	6.41%
Concejal	858	25	883	2.83%
Director	171	6	177	3.39%
Docente	1,030	57	1,087	5.24%
Gerente	96	7	103	6.80%
Informal	1,254	35	1,289	2.72%
Jubilado	2,145	59	2,204	2.68%
pensionado	1		1	0.00%
Profesional	429	35	464	7.54%
Vendedor	1,638	89	1,727	5.15%
<b>Total general</b>	<b>11,666</b>	<b>590</b>	<b>12,257</b>	<b>4.81%</b>

Adicionalmente, se puede observar que los pensionados también tienen una participación importante en la base de datos ya que representan el 18% del total asociados, y su probabilidad es del 2.68%.

Por último, llama la atención que la entidad tiene un número importante de Concejales pues suman 883 asociados.

## 18. Genero

De esta variable se puede observar que hombres y mujeres no registran niveles de riesgo similares ya que la probabilidad de incumplimiento de los hombres es de 5.87%, mientras que la de las mujeres se ubica en el 3.77%. En lo que si se evidencia equidad es en la distribución de asociados por género, ya que ambos representan el 50%.

**TABLA 10: GENERO**

Genero	NO DEFAULT	DEFAULT	TOTAL	PD
F	5,940	233	6,174	3.77%
M	5,726	357	6,083	5.87%
<b>Total general</b>	<b>11,666</b>	<b>590</b>	<b>12,257</b>	<b>4.81%</b>

## 19. Estado Civil

Revisando esta variable se observa que los clasificados como Casados o Viudos tienden a ser más responsables con el pago de sus obligaciones, ya que presentan la PD de menor valor. Estas personas representan el 40% de la base.

En contraste, se observa que aquellos en Unión Libre o Solteros son los que registran los niveles de riesgo más altos.

**TABLA 11: ESTADO CIVIL**

Estado Civil	NO DEFAULT	DEFAULT	TOTAL	PD
Casado (a)	4,047	141	4,189	3.37%
Separado (a)	738	37	775	4.77%
Soltero (a)	3,302	205	3,507	5.85%
Union Libre	2,924	188	3,112	6.04%
Viudo (a)	655	19	674	2.82%
<b>Total general</b>	<b>11,666</b>	<b>590</b>	<b>12,257</b>	<b>4.81%</b>

## 20. Tipo Vivienda

Revisando el resultado del descriptivo para esta variable se logra evidenciar que los asociados con mayor nivel de riesgo (Probabilidad de Default) son aquellos que tienen vivienda Familiar, ya que su PD es 6.80%; siendo la mayor de las diferentes clasificaciones. Estas personas representan el 21,8% de la base de construcción.

Por otra parte, se observa que los asociados con vivienda Propia, registran PD inferior a la del total de la base, permitiendo inferir que aquellas personas con vivienda propia presentan mejor habito de pago. Estos asociados representan el 34% de la base.

En contraste los Asociados con vivienda propia representan un menor riesgos para la Cooperativa, pues su PD es de 3.20%.

**TABLA 12: TIPO DE VIVIENDA**

Tipo Vivienda	NO DEFAULT	DEFAULT	TOTAL	PD
Alquilada	1,957	114	2,072	5.50%
Anticresada	81	3	84	3.57%
Familiar	3,878	283	4,161	6.80%
Propia	5,750	190	5,940	3.20%
<b>Total general</b>	<b>11,666</b>	<b>590</b>	<b>12,257</b>	<b>4.81%</b>

## 21. Estrato

Al revisar el nivel de riesgo de los asociados mediante su estrato socioeconómico no se logran ver diferencias significativas entre sí. Llama la atención que en el caso de estrato 6 la PD es del 11%, sin embargo, la muestra de asociados bajo esta clasificación es muy baja (100) lo que no permite concluir mucho sobre este nivel de riesgo.

Lo que sí se puede observar es que el 93% de la base, se encuentra concentrada en los estratos 1- 3.

**TABLA 13: ESTRATO**

Estrato	NO DEFAULT	DEFAULT	TOTAL	PD
1	3,955	187	4,143	4.51%
2	4,117	207	4,324	4.79%
3	2,893	156	3,049	5.12%
4	524	24	548	4.38%
5	88	5	93	5.38%
6	89	11	100	11.00%
<b>Total general</b>	<b>11,666</b>	<b>590</b>	<b>12,257</b>	<b>4.81%</b>

## 22. Personas a Cargo

Revisando el efecto que tienen el número de personas a cargo contra el nivel de riesgo que registran los asociados se logra observar que a partir de 3 las probabilidades empiezan a registrar tendencia creciente.

**TABLA 14: NO PERSONAS A CARGO**

Personas a Cargo	NO DEFAULT	DEFAULT	TOTAL	PD
0	3,646	189	3,836	4.93%
1	3,137	138	3,275	4.21%
2	2,763	146	2,909	5.02%
3	1,324	66	1,390	4.75%
4	556	33	589	5.60%
5	148	11	159	6.92%
6	64	5	69	7.25%
7	21	1	22	4.55%
8	5	1	6	16.67%
9	1		1	0.00%
10	1		1	0.00%
<b>Total general</b>	<b>11,666</b>	<b>590</b>	<b>12,257</b>	<b>4.81%</b>

### 23. Plazo

Analizando el plazo frente a la probabilidad de incumplimiento de los Asociados, se logra observar que aquellas obligaciones con plazos hasta 12 meses son las que registran el menor nivel de riesgo. Este comportamiento también puede estar relacionado con los montos de crédito, pues generalmente existe una relación directamente proporcional entre el monto y el plazo.

**TABLA 15: RANGO DE PLAZO**

Rango Plazo	NO DEFAULT	DEFAULT	TOTAL	PD
HASTA 12	3,945	81	4,026	2.01%
HASTA 24	3,099	199	3,298	6.03%
>36	4,622	310	4,933	6.28%
<b>Total general</b>	<b>11,666</b>	<b>590</b>	<b>12,257</b>	<b>4.81%</b>

### 24. Rango Ingreso

Los asociados con ingresos entre 801.000 y 1`200.000 registran la mayor PD con 5.42%, sin embargo, esta no resulta ser una diferencia significativa frente a las probabilidades de los demás rangos.

**TABLA 16: RANGO DE INGRESO**

Rango Ingreso	NO DEFAULT	DEFAULT	TOTAL	PD
Hasta 800.000	2,932	138	3,070	4.50%
Hasta 1,2 Millones	2,897	166	3,063	5.42%
Hasta 1,9 Millones	2,914	146	3,060	4.77%
> 1,9 Millones	2,923	140	3,064	4.57%
<b>Total general</b>	<b>11,666</b>	<b>590</b>	<b>12,257</b>	<b>4.81%</b>

## 25. Rango Edad

Otra de las variables que típicamente logra discriminar deudores buenos de los malos es la edad, en este caso se evidencia que los asociados con edad inferior o igual a 44 años son los que registran la PD más altas de la base. Ellos representan el 50% de la base total.

Adicionalmente se logra observar que solo 75 asociados de los 3.016 mayores de 56 años han caído en incumplimiento.

**TABLA 17: RANGO DE EDAD**

Rango Edad	NO DEFAULT	DEFAULT	TOTAL	PD
Hasta 34 años	3,045	245	3,290	7.45%
Hasta 44 años	2,713	149	2,862	5.21%
Hasta 56 años	2,968	121	3,089	3.92%
> 56 años	2,940	75	3,016	2.49%
<b>Total general</b>	<b>11,666</b>	<b>590</b>	<b>12,257</b>	<b>4.81%</b>

## 26. Rango Acierta

Revisando el Acierta frente al nivel de riesgo de los asociados, se logra observar que aquellos con puntajes inferiores a 714 registran niveles de riesgo superiores al consolidado de la cartera (4.81%). De hecho, se logra observar una clara relación inversamente proporcional entre acierta y la probabilidad de incumplimiento.



**TABLA 18: RANGO ACIERTA**

Rango Acierta	NO DEFAULT	DEFAULT	TOTAL	PD
Sin Acierta	1,435	85	1,520	5.59%
Hasta 461	862	73	935	7.81%
Hasta 629	1,146	93	1,239	7.51%
Hasta 681	1,142	84	1,226	6.85%
Hasta 714	1,163	72	1,235	5.83%
Hasta 742	1,187	49	1,236	3.96%
Hasta 764	1,179	37	1,216	3.04%
Hasta 784	1,212	39	1,251	3.12%
Hasta 803	1,153	38	1,191	3.19%
> 803	1,187	20	1,208	1.66%
<b>Total general</b>	<b>11,666</b>	<b>590</b>	<b>12,257</b>	<b>4.81%</b>

## 27. Monto Crédito

Revisando la relación entre el monto y el nivel de riesgo de los asociados se evidencia que los créditos con mayor PD son los de montos superiores a los 6 millones. Sorprendiendo el hecho de que el mayor riesgo se evidencia en las obligaciones con mayor monto.

Esta información indica oportunidades de mejora a la política de crédito dado que se esperaría un comportamiento opuesto al observado.

**TABLA 19: RANGO MONTO DE CRÉDITO**

Rango Monto	NO DEFAULT	DEFAULT	TOTAL	PD
Hasta 1,2 Millones	3,024	50	3,074	1.63%
Hasta 2,58 Millones	3,059	125	3,184	3.93%
Hasta 6 Millones	3,014	238	3,252	7.32%
> 6 Millones	2,569	177	2,747	6.44%
<b>Total general</b>	<b>11,666</b>	<b>590</b>	<b>12,257</b>	<b>4.81%</b>

## 28. Antigüedad en la Cooperativa

La antigüedad permite medir el nivel de fidelización que tiene el asociado con la cooperativa, y se logra observar que aquellos que al momento de solicitar el crédito tienen antigüedad de hasta 1 mes presentan los mayores niveles de riesgo de la base, pues su PD es de 8.93%.

A medida que su antigüedad se incrementa, las probabilidades de incumplimiento disminuyen, permitiendo ver la relación inversamente proporcional que existe entre estas dos variables.

**TABLA 20: ANTIGÜEDAD EN LA COOPERATIVA**

<b>Rango Antigüedad</b>	<b>NO DEFAULT</b>	<b>DEFAULT</b>	<b>TOTAL</b>	<b>PD</b>
Hasta 1 Mes	2,947	289	3,236	8.93%
Hasta 24 Meses	2,894	172	3,066	5.61%
Hasta 43 Meses	2,822	78	2,900	2.69%
> 43 Meses	3,003	51	3,055	1.67%
<b>Total general</b>	<b>11,666</b>	<b>590</b>	<b>12,257</b>	<b>4.81%</b>

Una vez realizado el análisis descriptivo de las diferentes variables presentes en la base de datos, se procede a diseñar un número de modelos que competirán entre ellos para definir cuál es el de mayor poder predictivo y mejor ajuste.

## 29. Construcción del Modelo

### 30. Metodología

Se realizaron análisis exploratorios con la construcción de al menos 30 modelos<sup>13</sup>, para finalmente definir con la ayuda de las pruebas de bondad de ajuste<sup>14</sup>, el de mejor rendimiento y mayor robustez.

### 31. Salidas del modelo seleccionado

#### Pruebas de bondad de ajuste

Para la selección del modelo se tuvo en cuenta las siguientes pruebas con un nivel de significancia del 5%.

### 32. Omnibus Test:

Las pruebas ómnibus son una especie de prueba estadística, que prueban si la varianza explicada en un conjunto de datos es significativamente mayor que la varianza no explicada. Como prueba estadística se implementa en una hipótesis general que tiende a encontrar una significación general entre la varianza de los parámetros, mientras se examinan parámetros del mismo tipo. Su objetivo es probar más de dos parámetros del mismo tipo y su función es encontrar la importancia general de al menos uno de los parámetros involucrados.

Con base en el estadístico x-Cuadrado se establece si globalmente el modelo es o no significativo; es decir, si al menos una de las variables explica el default.

$$H_0: \beta_1 = \beta_2 = \beta_3 \dots \dots = \beta_n = 0$$

---

<sup>13</sup> Ver archivo en Excel "Salidas Modelo"

<sup>14</sup> La bondad de ajuste de un modelo estadístico describe lo bien que se ajusta un conjunto de observaciones. Las medidas de bondad en general resumen la discrepancia entre los valores observados y los valores esperados en el modelo de estudio

*Ha: Al menos alguno de los coeficientes de  $\beta$  es diferente de cero.*

## **TABLA 21: PRUEBA ÓMNIBUS**

**Pruebas ómnibus de coeficientes de modelo**

	Chi-cuadrado	gl	Sig.
Paso	8,462	1	0,004
Bloque	483,159	7	0,000
Modelo	483,159	7	0,000

De acuerdo con el nivel de significancia de la prueba, se rechaza la hipótesis nula, con lo que se puede afirmar que al menos una de las variables explica el default. Lo anterior quiere decir que al menos unas de las variables analizadas, explica porque los créditos entran en mora mayor a 90 días. Esto quiere decir que la Cooperativa dentro de la información solicitada al Asociado, tiene datos que determinan el incumplimiento del asociado frente al pago del crédito otorgado. Se observa que el nivel de significancia es bueno 0.004, para lo cual podemos concluir que los datos del modelo pueden determinar el default de los créditos.

### **33. Test de Wald:**

El estadístico de Wald se utiliza para poner a prueba el verdadero valor del parámetro basado en la estimación de la muestra, esto quiere decir que se estima el valor de máxima verosimilitud del parámetro para compararlo con el valor propuesto.

Con base en el estadístico de Wald, se establece si marginalmente una a una de las variables incluidas en el modelo son significativas. (Cabe aclarar que para las variables categóricas se debe mirar la significancia de manera global y no de manera marginal cada categoría).

Tabla 21

**TABLA 22: VARIABLES EN LA ECUACIÓN**

Variables en la ecuación						
	B	Error estándar	Wald	gl	Sig.	Exp(B)
CATEGORIA PLAZO(1)	-0,687	0,132	26,893	1	0,000	0,503
CATEGORIA GARANTIA(1)	-0,557	0,103	29,245	1	0,000	0,573
CATEGORIA ACIERTA(1)	0,704	0,093	57,042	1	0,000	2,022
CATEGORIA GENERO(1)	-0,480	0,089	29,308	1	0,000	0,619
CATEGORIA TIP VIVIENDA(1)	-0,401	0,097	17,228	1	0,000	0,669
ANTIG COOP	-0,021	0,002	79,588	1	0,000	0,979
EDAD AL DESEM	-0,010	0,004	8,299	1	0,004	0,990
Constante	-1,770	0,162	118,817	1	0,000	0,170

### 34. Tabla de Clasificación:

A pesar de no constituirse como una prueba estadística, la tabla de clasificación ofrece un análisis de la bondad de ajuste del modelo, permitiendo identificar la exactitud que tiene el modelo para clasificar correctamente a los Asociados NO DEFAULT y DEFAULT. Dicho de otra manera, el porcentaje de Asociados buenos "NO DEFAULT" clasificados como buenos y el porcentaje de Asociados malos "DEFAULT" clasificados como malos.

Para fijar el punto de corte a partir del cual se clasifican los Asociados cumplidos y los incumplidos, se toma la probabilidad de incumplimiento que presentaba la base analizada. Esto no es más que la proporción de asociados malos que presentaba la base.

## TABLA 23: TABLA DE CLASIFICACIÓN

Tabla de clasificación

Observado	Pronosticado		Porcentaje correcto
	BUENOS	MALOS	
DEFAULT BUENOS	7704	3962	66,0
MALOS	155	435	73,7
<b>Porcentaje global</b>			<b>66,4</b>

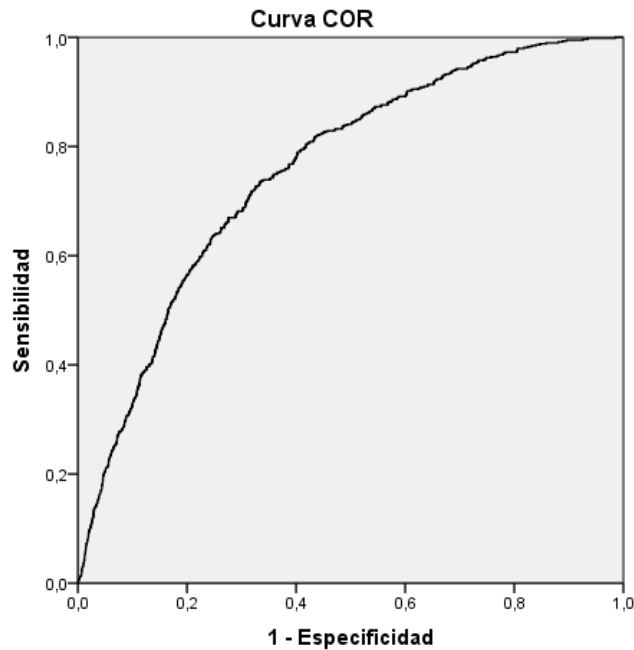
Se observa que a partir del modelo se clasifican adecuadamente el 66% de los asociados buenos y el 73.3% de los asociados malos; para una tasa de aciertos global que alcanza el 66.4%.

### 35. Curva Cor O ROC Curve:

La curva de ROC es una representación gráfica de la sensibilidad, frente a la especificidad para un sistema clasificador binario según se varía el umbral de discriminación. Este gráfico es la representación de la razón o ratio de verdaderos positivos frente a la razón o ratio de falsos positivos, éste también puede variar según el umbral de discriminación (valor a partir del cual decidimos que un caso es un positivo).

Esta prueba de bondad de ajuste para las regresiones logísticas es la representación gráfica de la sensibilidad y el inverso de la especificidad.

### ILUSTRACIÓN 3: CURVA DE ROC



Los segmentos de diagonal se generan mediante empates.

Área
0,756

El área bajo la curva se puede interpretar como la probabilidad de que un asociado cualquiera sea bien clasificado por el modelo ajustado. La literatura recomienda para modelos de otorgamiento, el área bajo la curva sea superior o igual a 65%; con lo que este modelo estaría presentando muy buen ajuste (75.6%).

### 36. K-S TEST:

La prueba de Kolmogórov-Smirnov (también prueba K-S), es una prueba no paramétrica que determina la bondad de ajuste de dos distribuciones de probabilidad entre sí. Esta prueba es más sensible a los valores cercanos a la mediana, que a los extremos de la distribución. La interpretación es bastante

sencilla dado que su valor varía entre 0 y 1; donde un modelo perfecto estaría cercano a 1 y un modelo deficiente en 0.

En este caso el test se ubicó en 9.518, por lo que se puede afirmar que el modelo tiene buen ajuste.

Otra interpretación que se le puede dar al modelo, es que la máxima diferencia entre el acumulado de buenos y el acumulado de malos que se encontrarán en el modelo es del 40.2%. Esta información se puede validar con el Score Distribution que está calculado más adelante.

**TABLA 24: ESTADÍSTICO KS**

Estadísticos de prueba		Probabilidad pronosticada
Máximas diferencia s extremas	Absoluta	0,402
	Positivo	0,402
	Negativo	-0,001
Z de Kolmogorov-Smirnov		9,518
Sig. asintótica (bilateral)		0,000

a. Variable de agrupación: DEFAULT

### 37. Ecuación

Para estimar el modelo logit se utilizó el paquete estadístico SPSS. Se realizó un proceso paso a paso, con datos disponibles en la tabla No 2 (variables disponibles de la muestra). Las variables que resultaron significativas fueron plazo, garantía, genero, tipo de vivienda, antigüedad en la Cooperativa y Edad. Aunque esto es coherente con el resultado de la exploración por factores realizado anteriormente, el objetivo del módelo es hallar al Asociado que entra en Default, el cual se define a partir del criterio de quiebra establecido en el desarrollo de los modelos tipo scoring. De acuerdo con el Capítulo II de la Circular Externa 100 de la Superintendencia Financiera de Colombia (2011), el Default o evento de incumplimiento se define como el estado en que entra un deudor, cuando tiene una mora mayor a 90 días en cualquiera de las obligaciones crediticias que tenga en la entidad financiera (créditos de consumo). Esta definición viene desde el Acuerdo



Capital de Basilea II (2004), donde El Comité de Supervisión Bancaria de Basilea dio una definición del evento de Default que sirve de referencia regulatoria para las entidades financieras que desarrollan sus modelos internos de puntuación.

El cálculo de la probabilidad de incumplimiento “Default” para el modelo se realiza mediante la siguiente fórmula:

$$PI = \frac{e^z}{1 + e^z}$$

Donde, se obtuvo el siguiente resultado:

Z: **-1,770** + (si Plazo<=12; toma el valor de **-0,687**; en caso contrario 0) + (si Garantía es Codeudor o Garantía Admisible; toma el valor de **0**; en caso contrario **0,557**) + (si Acierta <=714; toma el valor de **0,704**; en caso contrario **0**) + (si Genero es Femenino; toma el valor de **-0,480**; en caso contrario 0) + (si Tipo de Vivienda es Alquilada o Familiar; toma valor de 0; en caso contrario **-0,401**) + (Antigüedad Cooperativa\***-0,021087**) + (Edad\***-0,01015**)

Lo anterior nos describe los pesos de cada una de las variables, en la ecuación. De ahí la importancia que la Cooperativa pueda solicitar esta información, en la solicitud de crédito de cada Asociado.

## 38. Score Distribution

**TABLA 25: SCORE DISTRIBUTION**

SCORE DISTRIBUTION														
SCORE ORIGINACION (agrupado)	BUENOS	%BUENOS	%ACUM	MALOS	%MALOS	%ACUM	Total	%TOTAL	%ACUM	ODDS	ODDS ACUM	PD	PD ACUM	K-S
<= ,00724	1.221	10%	10%	4	1%	1%	1.225	10%	10%	305	20	0,33%	4,81%	9,79%
,00725 - ,01227	1.215	10%	21%	12	2%	3%	1.227	10%	20%	101	19	0,98%	5,05%	18,17%
,01228 - ,01794	1.204	10%	31%	20	3%	6%	1.224	10%	30%	60	18	1,63%	5,28%	25,10%
,01795 - ,02410	1.197	10%	41%	30	5%	11%	1.227	10%	40%	40	17	2,44%	5,49%	30,28%
,02411 - ,03259	1.192	10%	52%	33	6%	17%	1.225	10%	50%	36	17	2,69%	5,68%	34,90%
,03260 - ,04278	1.186	10%	62%	43	7%	24%	1.229	10%	60%	28	16	3,50%	5,88%	37,78%
<b>,04279 - ,05853</b>	<b>1.169</b>	<b>10%</b>	<b>72%</b>	<b>53</b>	<b>9%</b>	<b>33%</b>	<b>1.222</b>	<b>10%</b>	<b>70%</b>	<b>22</b>	<b>16</b>	<b>4,34%</b>	<b>6,05%</b>	<b>38,82%</b>
,05854 - ,07709	1.145	10%	82%	80	14%	47%	1.225	10%	80%	14	15	6,53%	6,18%	35,07%
,07710 - ,11656	1.096	9%	91%	134	23%	69%	1.230	10%	90%	8	15	10,89%	6,15%	21,75%
,11657+	1.041	9%	100%	181	31%	100%	1.222	10%	100%	6	16	14,81%	5,72%	0,00%
<b>Total</b>	<b>11.666</b>			<b>590</b>	<b>100%</b>		<b>12.256</b>			<b>20</b>		<b>4,81%</b>	<b>4,81%</b>	

En el score distribution se puede observar que se cumple el supuesto de **monotonicidad** el cual indica que a medida que los puntajes disminuyen, los niveles de riesgo de la cartera crecen. Es por este motivo que a medida que a medida que disminuye el rango del Score arrojado por el modelo, las probabilidades reales de la cartera presentan comportamiento inversamente proporcional.

También se observa que el máximo K-S del modelo se ubica en el rango de **4.27% – 5.85%**; siendo el punto en que mayor diferencia acumulada existe entre la cantidad de asociados buenos y malos. Es por este motivo que la academia recomienda que se haga como punto de corte para aprobación de créditos en este rango.

Sin embargo, es importante recordar que el hecho de establecer el punto de corte para aprobar créditos, depende del apetito de riesgo que desea la entidad y el que puede asumir bajo las condiciones financieras de la misma.

## Conclusiones y recomendaciones

Se puede observar que el modelo logit tiene un buen poder discriminatorio, reflejado en las tasas de aciertos, sobre todo para los Asociados que ingresan a Default.

En el pronóstico, el modelo determina buenos como buenos en un 66% y agrupa malos como malos en un 74%. Lo cual nos permite señalar que el modelo tiene un porcentaje alto de predicción, al separar los dos grupos (default y no default)

**TABLA 26: PRONOSTICO DEL MODELO**

Observado	Pronosticado		Porcentaje correcto
	BUENOS	MALOS	
DEFAULT BUENOS	7704	3962	66.0
MALOS	155	435	73.7
<b>Porcentaje global</b>			<b>66.4</b>

Así mismo se pudo constatar que las variables empleadas en el modelo (según ecuación) determinan si un asociado ingresa a Default, demostrando ser significativas y explicando el modelo requerido. Por otro lado, se descartaron variables que no fueron significativas y se recomienda a la Cooperativa verificarlas para saber si se requieren a los Asociados o no.

Lo anterior teniendo en cuenta que al Asociado se le solicitan variables cualitativas y cuantitativas, que no explican el "default" y solo generan más tramites de documentos.

Para la Cooperativa resulta primordial contar con una herramienta estadística como esta, ya que mejora la rentabilidad de los asociados y disminuye la probabilidad de incumplimiento "Default", situación que es recomendada por la superintendencia de

economía solidaria a las Cooperativas, donde recomienda evaluar el riesgo de crédito para mitigar la probabilidad de incumplimiento “Default” de los Asociados.

Se recomienda calibrar el modelo cada 6 meses, para analizar los comportamientos de las variables analizadas y los nichos de mercado nuevos a los cuales pueda llegar la entidad en su plan de crecimiento.

Se hace necesario para futuros análisis, que la entidad disponga de una base de datos de los créditos aplazados y negados con el objetivo de incluirlos en el modelo.

El aporte microeconómico del modelo a la entidad tiene que ver con el tema de las preferencias del Asociado al momento de solicitar un crédito y poderlo pagar o no en el tiempo, ello conlleva a un análisis más estadístico mejorando lo que hoy se vienen desarrollando para la aprobación de créditos (modelo experto). El score mejorara la segmentación de los asociados para determinar el default de éstos.

## Bibliografía

Abdou, H. 2009. An evaluation of alternative scoring models in private banking. *Journal of Risk Finance* 10 (1): 38-53.

Abdou, H. and Pointon, J. (2011). Credit scoring, statistical techniques and evaluation criteria: a review of the literature. *Intelligent Systems in Accounting, Finance & Management*, 18 (2-3), pp. 59-88.

Aguas, D. y Castillo, M. (2002). "Modelo de administración del riesgo crediticio para la cartera comercial de una entidad financiera colombiana", *Apuntes de Banca y Finanzas*, No. 6, pp.1-8, 2002.

Altman, E. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*.

Amaya, C. A. (2005). Evaluación del riesgo de crédito en el sistema financiero colombiano. Reporte de estabilidad financiera, Banco de la Republica

Arango, J. P., N. Zamudio, and I. Orozco (2005). Riesgo de crédito: Un análisis desde las firmas. Reporte de estabilidad financiera, Banco de la Republica.

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In *Proceedings of the Second International Symposium on Information Theory*, B. N. Petrov and F. Csaki (eds). Budapest: Akademiai Kiado, 267-281.

Boyes, W., D. Hoffman, and S. Low (1989). An Econometric Analysis of the Bank Credit Scoring Problem. *Journal of Econometrics*, 40, pp. 3-14.

Chandler, G. G., Coffman, J. Y. (1979). A comparative analysis of empirical vs. judgemental credit evaluation. *The Journal of Retail Banking* 1 (2): 15-26. Chatterjee, S., A. S. Hadi and B. Price. (2000). *Regression analysis by example*. John Wiley and Sons, New York, New York. USA.

Fernández, H. y Pérez, F. (2005). "El modelo logístico: una herramienta estadística para evaluar el riesgo de crédito", *Revista de Ingenierías Universidad de Medellín*, Vol. 4, No. 006, pp. 55-75.

Gracia-Diez, M. y Serrano, G. R. (1992). Algunos aspectos sobre el análisis empírico de credit scoring. *Estadística Española*. 34 (130): 261 – 283.

Greene, W. (1992). A statistical model for credit scoring. Department of Economics Stern School of Business. New York University.

Gutierrez, M. (2007). "Modelos de credit scoring: qué, cómo, cuándo y para qué", *MPRA Paper*, No. 16377, pp.1-30.

Hauck, W. W. & Donner, A. (1977), 'Wald's test as applied to hipotesis in logit analysis', *JASA* 72, 851–853.

Hosmer, D. and Lemeshow, S. (2000). Applied Logistic Regression. Wiley Series in Probability and Statistics. Second Edition. JOHN WILEY & SONS, INC. United State of America. Pp. 91 -128

Hosmer, D. & Lemeshow, S. (1989), Applied Logistic Regression, Wiley & Sons, New York.

Jiménez, G. and J. Saurina (2003). Collateral, type of lender and relationship banking as determinants of credit risk. *Journal of Banking & Finance*.

Kukuk, M. and Roennberg M. (2013). Corporate Credit Default Models: A Mixed Logit Approach. *Review of Quantitative Finance and Accounting*. 40 (3): 467-483

Maddala, G. S. (1983). Limited dependent and qualitative variables in Econometrics. New York: Cambridge University Press, pp. 165 – 178.

McFadden, D. (1976). A comment on discriminant analysis versus logit analysis. *Annals of Economic and Social Measurement*, vol. 5, pp. 511 – 523.

Moreno V. Sandra, (2013). El Modelo Logit Mixto para la construcción de un Scoring de Crédito.

Orgler, Y. E. (1970). A credit scoring model for commercial loans. *Journal of Money, Credit and Banking* II (4): 435-445.

Ohlson, J. A., (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy, *Journal of Accounting Research*, Vol. 18, No. 1, pp. 109–131.

Porath, D., (2004). Estimating Probabilities of Default for German Savings Banks and Credit Cooperatives, Discussion Paper Series 2: Banking and Financial Supervision No 06/2004.

Superintendencia de Economía Solidaria. (2012). Circular básica contable y financiera, Colombia. [www.supersolidaria.gov.co](http://www.supersolidaria.gov.co)

Superintendencia Financiera de Colombia. (1995). Circular básica contable y financiera 100 de 1995, Colombia.

Torres, G. (2005). "El acuerdo de Basilea: Estado del arte del SARC en Colombia", Ad-minister, No. 6, pp.114-135.

Wu, X. (2008). Credit Scoring Model Validation. Universiteit van Amsterdam. Faculty of Science. Korteweg-de Vries Institute for Mathematics. Máster Thesis.