



SUMA DE NEGOCIOS



Artículo de investigación

Modelo para la estimación del deterioro por riesgo de crédito



Iván Mauricio Bermúdez Vera¹, Diego Fernando Manotas Duque² y Javier Olaya Ochoa³

¹ Magíster en Ingeniería: énfasis Ingeniería Industrial. Profesor auxiliar Departamento de Ciencias Naturales, Exactas y Estadística. Universidad Santiago de Cali, Cali, Valle del Cauca, Colombia. (Autor de correspondencia) Correo electrónico: ivan.bermudez00@usc.edu.co. ORCID: 0000-0002-6590-4938.

² Doctor en Ingeniería: énfasis Ingeniería Eléctrica y Electrónica. Profesor titular Escuela de Ingeniería Industrial. Universidad del Valle, Cali, Valle del Cauca, Colombia. Correo electrónico: diego.manotas@correounivalle.edu.co. ORCID: 0000-0003-0148-9840.

³ Doctor en Management Science. Profesor titular Escuela de Estadística. Universidad del Valle, Cali, Valle del Cauca, Colombia. Correo electrónico: javier.olaya@correounivalle.edu.co. ORCID: 0000-0001-7014-2782.

INFORMACIÓN DEL ARTÍCULO

Recibido el 28 de Enero de 2020

Aceptado el 20 de Abril de 2020

Online el 15 de mayo de 2020

Códigos JEL:

C51, C53, G11, G23

Palabras clave:

Riesgo de crédito,
pérdida esperada,
modelo Logit,
método Hold Out,
provisión de cartera.

Keywords:

Credit risk,
expected loss,
Logit model,
Hold Out method,
provision of portfolio.

R E S U M E N

El artículo desarrolla el modelo de estimación de pérdida esperada como soporte al Sistema de Administración del Riesgo de Crédito para una entidad de economía solidaria. Describe la construcción de un modelo de regresión logística para la estimación de la probabilidad de incumplimiento de sus asociados y valora el desempeño del modelo desde el poder de discriminación entre cumplidos e incumplidos mediante el método Hold Out repetido. Definido el modelo y estimada la probabilidad de incumplimiento de los asociados, se determina la tasa de recuperación de acuerdo con el tipo de garantía que presente cada crédito, permitiendo así estimar la pérdida esperada y las provisiones de cartera necesarias para la entidad.

Model for the estimation of impairment due to credit risk

A B S T R A C T

The article develops an estimation model for expected loss as support to the Credit Risk Management System for a corporation of social economy. it describes the setting of a logistic binary regression model to estimate the probability the non-compliance of its affiliates and, it evaluates the model's capacity to discriminate between compliance and non-compliance using the Hold Out recurrent method. Once the model is fixed and the probability is estimated, the recovery rate is determinate according to the guarantee offer by each credit. so, is possible the estimation of expected loss as well as the portfolio provisions required by each fund.

Introducción

De acuerdo con Altman y Saunders (1997), “las fuerzas que han impulsado el desarrollo de la medición del riesgo de crédito son el aumento de quiebras de bancos, la tendencia a la desintermediación y el aumento de la competitividad”. En esencia, Colombia no ha sido ajena a esta dinámica y muestra de ello, la Superintendencia Financiera (SF) ha impulsado normas que tratan de recoger las mejores prácticas en cuanto a la administración del riesgo, particularmente los principios y sugerencias desarrollados por el Comité de Supervisión Bancaria de Basilea (BCBS).

Actualmente, el exceso de demanda de crédito con el aumento en la proliferación de nuevos productos de crédito como la tarjeta de crédito y operaciones de libranzas, expone a los usuarios del sistema financiero a la posibilidad de obtener créditos con más de una entidad, luego en procura de tener una mayor certeza sobre el comportamiento crediticio de un cliente, es necesario que las entidades sean eficientes en identificar y gestionar el riesgo al que están expuestas.

Aunque los sistemas de administración de riesgo de crédito (SARC) han sido definidos como un elemento fundamental en la gestión del riesgo del negocio financiero, la definición de modelos de medición de riesgo de crédito a partir de información estadística no ha sido una práctica ampliamente utilizada en el sector de la economía solidaria. Se registran algunos avances en aspectos de carácter procedimental, sistemas de información y mecanismos de control, pero muchos fondos de empleados aún no cuentan con modelos de medición de riesgo crediticio que puedan servir como base de los modelos internos para el otorgamiento, el seguimiento y el cálculo de la pérdida esperada.

En la gran mayoría de este tipo de entidades existentes en el campo nacional, el mecanismo de pago predominante es el descuento por nómina, lo que reduce significativamente el riesgo crediticio asociado. Sin embargo, la entidad utilizada como muestra en este proyecto tenía la particularidad de que para más del 95% de los asociados, el mecanismo de pago principal son los canales virtuales de las entidades financieras. A pesar del riesgo que podría implicar esta situación, la entidad estudiada como caso presentaba índices de morosidad inferiores al 1%, muy por debajo de los estándares nacionales e internacionales. Esta situación genera un reto adicional en el proceso de modelación, ya que es preciso definir un nuevo umbral de morosidad más exigente que los estándares tradicionales.

Dado lo anterior, resaltamos el aporte de este estudio en la utilización de técnicas probadas en la medición de riesgo de crédito, para configurar un modelo de otorgamiento propio para una entidad del sector de la economía solidaria, considerando las particularidades de la base social de la entidad, que podrían generar potenciales situaciones de riesgo, aun considerando las dificultades para calibrar adecuadamente un modelo de medición en una población con muy bajo nivel de impagos.

Para alcanzar el objetivo, el artículo se desarrolla en cuatro secciones. En primera instancia, se presenta la revisión

de la literatura sobre la evaluación del riesgo de crédito. En la segunda se explica la metodología que se utilizó para la construcción del modelo de pérdida esperada. En la tercera se describen los principales resultados. Finalmente, se concluye el documento.

Marco teórico

La modelación de riesgo de crédito es un tema de gran trascendencia e importancia para la adecuada gestión del riesgo financiero. Son numerosos los trabajos que se han realizado sobre temas relacionados con modelos de otorgamiento, *scoring*, estimación de pérdida esperada, manejo de provisiones y entorno económico, entre otros. Entre las más recientes publicaciones encontradas en la literatura figuran Bülbül, Hakenes y Lambert (2019), que realizan un completo análisis de los determinantes en las entidades financieras bancarias en Alemania para la adopción de prácticas relacionadas con la gestión de riesgo crediticio. Según los autores, los niveles de sofisticación de las herramientas usadas en los sistemas de administración de riesgo crediticio en los bancos estudiados, son fundamentalmente consecuencia de los niveles de competencia y concentración del mercado de préstamos.

González y González (2019) presentan una aproximación alternativa para la medición de riesgo de créditos en bancos europeos a partir de la información disponible en Google. La aproximación desarrollada se fundamenta en búsquedas lexicográficas, con el propósito de producir un índice que logre capturar el riesgo crediticio considerando elementos de carácter no financiero.

En la actualidad, llama la atención de los investigadores aspectos como la heterogeneidad de los patrones de créditos de los usuarios de créditos de consumo. Esta característica dificulta los procesos de estimación de covarianza entre los diferentes sujetos participantes del portafolio de crédito de las entidades. En este contexto, Madeira (2019) presenta un estudio basado en las estadísticas de créditos hipotecarios en Chile. Entre las conclusiones de dicho estudio, se simulan diferentes condiciones de heterogeneidad asociadas a los clientes hipotecarios en función de lineamientos de orden macroeconómico.

Recientemente se incluyen trabajos como el de Bennouna y Tkouat (2019), que combinan técnicas de análisis discriminantes múltiple y regresión logística para formular un modelo de *scoring* para microcréditos en Marruecos. Chadwick (2018) analiza los impactos de la política monetaria sobre el crecimiento y la volatilidad de los créditos de consumo en Turquía. Su conclusión más relevante señala que la efectividad de la política monetaria se amplifica si se combina con medidas prudenciales en relación con el manejo de las provisiones asociadas a los portafolios de crédito. En este mismo sentido, otro de los trabajos relevantes, relacionados con la formulación de modelos de *scoring* crediticio, es el desarrollado por Fang y Chen (2019), quienes proponen un modelo de evaluación de crédito basado en el indicador Kolmogorov-Smirnov (KS) para medir el desempeño de los criterios de calificación considerados.

Una de las aplicaciones más interesantes en relación con el uso de las técnicas de medición de riesgo de crédito, es la que proponen autores como Walke, Fullerton y Tokle (2018), que utilizan las probabilidades de incumplimiento derivadas de modelos de riesgo de crédito como insumo para la segmentación de bases de datos de usuarios en cooperativas de créditos, con el fin de establecer políticas de tasas de interés basadas en niveles de riesgo crediticio, en lo que se conoce como estrategias de *risk pricing*. Otra aproximación reciente, en el mismo sentido, se presenta en el trabajo desarrollado por Kozodoi, Lessmann, Papakonstantinou, Gatsoulis y Baesens (2019), que proponen segmentos de utilidad a partir de la información brindada por los modelos de riesgo de crédito.

Otros autores como Abdolreza, Farnoosh, Konstantin y Frank (2017), proponen un nuevo método para la predicción de probabilidad, dado el incumplimiento basado en modelos difusos en reglas y métodos de conjunto que incorporan una amplia gama de variables macroeconómicas que proporcionan un aumento significativo en las medidas de rendimiento en comparación con las metodologías propuestas en la literatura. Yu, Ching, Gu y Siu (2017) consideran un modelo de riesgo de crédito en función de la intensidad de forma reducida con un proceso oculto del estado de Markov. Jokivuolle y Peura (2003) optan por procesos estocásticos correlacionados para el valor de la tasa de recuperación y los activos del deudor.

En el ámbito local, Echeverri (2006) en su trabajo de grado compara y evalúa las bondades y desventajas de los modelos de probabilidad lineal, el *logit* y el análisis discriminante. González y Orozco (2009) comparan los modelos estadísticos de alerta temprana frente a los modelos de respuesta binaria en la estimación de la probabilidad de incumplimiento. Ochoa, Galeano y Agudelo (2010) incorporan el uso de variables cualitativas y cuantitativas en la construcción de un modelo de *scoring*, haciendo uso del análisis discriminante. Támara, Aristizábal y Velásquez (2012) describen el uso de modelos econométricos en la estimación de la probabilidad de incumplimiento, obteniendo estimaciones que generan provisiones esperadas de menor cuantía que las descritas por el ente regulador y la entidad financiera. Caicedo, Claramunt y Casanovas (2011) utilizan el enfoque de Merton en la estimación de las probabilidades de incumplimiento y las tasas de recuperación dado el incumplimiento. Salazar (2013), haciendo uso de la teoría de los árboles de decisión, logra determinar las variables que influyen en la capacidad de pago.

Contar con la estimación de la probabilidad de incumplimiento de los asociados solo es un paso, en el camino a cuantificar la pérdida esperada y estimar las provisiones de la entidad. Para ello, es necesario determinar de forma adecuada la demanda solvente del crédito, entendiendo como demanda solvente del crédito, la petición de crédito por parte de aquellos asociados que sean merecedores de ello, mediante la medición de su solvencia (López, 2014). Para que la medición de solvencia no se considere una tarea subjetiva, Basilea II y III establece que cada entidad puede basarse en los puntajes emitidos por las agencias de calificación o definir un sistema de clasificación para determinar dicha solvencia.

Desde el 2006, el Comité de Basilea para la supervisión bancaria, permite a los bancos utilizar sus propios modelos

o enfoques de calificación, con el fin de calcular el capital requerido para la constitución de garantías de crédito, es decir, un modelo o enfoque basado en calificaciones internas de la entidad. Este concepto estipula la idea de la pérdida esperada como producto de tres parámetros: la probabilidad de incumplimiento (PI), la proporción de pérdida dada el incumplimiento (PDI) y la exposición del activo (EA). Siendo importante para la entidad contar con estimaciones confiables de cada componente (PI y PDI), con la finalidad de reducir la pérdida esperada y, por tanto, una reducción de capital de riesgo.

Todos estos referentes, dejan en evidencia el auge y constante desarrollo referente al análisis del riesgo crediticio, utilizando diferentes enfoques para la búsqueda de la probabilidad de incumplimiento.

Metodología

Dada la regulación bancaria de Basilea y teniendo en cuenta la reglamentación de la Superintendencia Financiera de Colombia (SFC), el estudio tiene como objetivo describir la construcción del modelo de *scoring* para una entidad de economía solidaria, a partir de la metodología de los modelos de regresión logística binaria, el cual permite calcular la probabilidad de incumplimiento (PI) en el camino a cuantificar la pérdida esperada y estimar las provisiones de una entidad de economía solidaria; teniendo presente que muchas de estas entidades aún no cuentan con modelos de medición de riesgo crediticio, que puedan servir como base de los modelos internos para el otorgamiento, el seguimiento y el cálculo de la pérdida esperada. Es así como la metodología se desarrolla teniendo en cuenta las siguientes tres fases:

- **Fase 1: Construcción del modelo de scoring**

Con el objetivo de determinar los parámetros que permitan estimar la pérdida esperada, de acuerdo con López (2014), se emplea un modelo de puntajes internos básico (conocido generalmente como modelo de *scoring*), en el cual la entidad estima las PI a partir de un modelo interno. Es así que para la construcción del modelo de *scoring*, se usó un modelo de regresión logística, el cual se construye a partir de variables de la historia de crédito del asociado, provenientes de las bases de datos propias de la entidad solidaria y de la Central de Información Financiera (CIFIN). Contando con la información de los créditos realizados por la entidad durante un periodo de 19 meses, y el estado actual del asociado en la central de riesgo CIFIN.

Para la modelación se utilizó el software estadístico R (R Core Team, 2020), usando la función *glm*, la cual permite configurar la modelación y evaluar fácilmente su desempeño. Los modelos de regresión logística permiten predecir el comportamiento de una variable categórica a partir de un conjunto de covariables discretas o continuas, siendo una de sus principales fortalezas. Estos pertenecen al conjunto de Modelos Lineales Generalizados introducidos por Nelder y Wedderburn (1972) y presentados en Dobson y Barnett (2018), parten de una variable de respuesta

dicotómica o binaria (Y), la cual puede tomar dos valores: 0 (cumplido) y 1 (incumplido).

Con la finalidad de expresar la probabilidad de que un asociado sea incumplido con su obligación crediticia como función de otras variables influyentes, en la ecuación (1) se presenta la forma analítica en que esa probabilidad se vincula con las variables explicativas.

$$\pi = \frac{1}{1 + \exp(-Z)} \text{ siendo } Z = \alpha + \beta_1 X_1 + \dots + \beta_k X_k \quad (1)$$

Donde π representa la probabilidad de *default* de un asociado (esto es, $P(Y = 1) = PI$), $\alpha, \beta_1, \dots, \beta_k$ son los parámetros del modelo y \exp denota la función exponencial. La expresión (2) es lo que se conoce como la función logística. Los parámetros de la regresión lineal pueden transformarse con la función logaritmo natural y se interpretan como la relación de odds de la variable dependiente (variable respuesta). Los odds corresponden al cociente entre el número eventos (número de veces que toma el valor de 1) sobre el número de no eventos (número de veces que toma el valor de 0). El modelo lineal generalizado no tiene supuestos teóricos sobre la distribución del error, si no que asume que se dispone de n variables independientes Y_i , tal que todas y cada una sigue la misma distribución, que pertenece a la familia exponencial.

Para determinar qué variables presentan un aporte significativo en la explicación de la probabilidad de incumplimiento para la población de asociados pertenecientes a la entidad, y lograr estimar un modelo adecuado para la entidad, se emplea el procedimiento de selección de variables utilizando el test chi-cuadrado de razón de verosimilitudes (RV), el cual se basa en el estadístico desviación definido en la ecuación (2) para el caso del modelo logístico.

$$D = -2 \sum_{j=1}^p \left[y_j \ln \left(\frac{\hat{\pi}_j}{y_j} \right) + (1 - y_j) \ln \left(\frac{1 - \hat{\pi}_j}{1 - y_j} \right) \right] \quad (2)$$

Donde $\hat{\pi}_j = \hat{\pi}(x_j)$

La desviación compara el logaritmo de la verosimilitud del modelo ajustado con el logaritmo de la verosimilitud del modelo saturado; es decir, el modelo que contiene todas las variables de interés que se quieren evaluar y todas las interacciones posibles, para lograr el contraste de las siguientes hipótesis: H_0 : el modelo ajustado no difiere del modelo saturado, y H_1 : el modelo ajustado difiere del modelo saturado. El modelo saturado es el que contiene la mayor cantidad de parámetros como número de observaciones en el conjunto de datos. El modelo ajustado se refiere al modelo en el que se ha realizado una evaluación de las variables del modelo donde se prueba la significancia de estas en relación con las variables de estudio.

El estadístico D presenta distribución asintótica χ_{n-p}^2 , donde p corresponde al número de parámetros en el modelo, en el que la hipótesis nula será rechazada para un nivel de significancia α cuando $D \geq \chi_{n-p}^2$, equivalente a que el p -valor del contraste sea menor que el nivel de significancia α dado.

De lo anterior, se describen los pasos a realizar en el procedimiento de selección de variables, donde se evalúa el aporte

que presenta cada una de las variables independientes en la explicación de la probabilidad de incumplimiento (PI):

1. Se parte de un modelo que incluye solo una variable predictora. Para determinar cuál es esa primera variable, se ajustan todos los posibles modelos de una sola variable, con cada una de las variables a evaluar.
2. Se elige el modelo que presente menor desviación.
3. Se evalúa la significancia de la variable introducida mediante un test chi-cuadrado de RV.
4. Si la variable resulta significativa, se deja en el modelo y se continúa con la selección de la próxima variable a incorporar, añadiendo por separado cada una de las variables restantes.
5. Se repiten los pasos 2, 3 y 4, hasta que la inclusión de una variable no sea significativa.

La capacidad para predecir la probabilidad o el riesgo de *default* de un individuo debe ser el atributo más útil del modelo. Es decir, la capacidad predictiva de clasificar como “cumplidos” a aquellos que presenten una baja PI y como “incumplidos” a aquellos que presenten una alta PI.

• Fase 2: Validación del modelo

En una segunda fase se valida el modelo valorando su nivel de ajuste y capacidad predictiva. Para ello se estima la curva ROC (*receiver operating characteristic*), la cual indica qué tanto poder de clasificación presenta el modelo (representación gráfica de la proporción de verdaderos positivos respecto a la proporción de falsos positivos). A su vez, la tasa de clasificaciones correctas se entiende como la medida de la bondad del ajuste global y se obtiene como el cociente entre los aciertos y el tamaño de muestra, dependiendo del punto de corte del cual se considera un individuo debe clasificarse como incumplido. Primero se clasifica el estado observado del asociado, si es *incumplido* (1) o *cumplido* (0), frente al estado pronosticado por el modelo ajustado, como se relaciona en la tabla 1.

Tabla 1 - Clasificación del evento observado vs. estimado

Observado	Pronosticado	
	0	1
0	VN	FN
1	FP	VP

Fuente: elaboración propia.

Donde:

VP (*verdaderos positivos*): son incumplidos y el modelo los clasifica como incumplidos.

VN (*verdaderos negativos*): son cumplidos y el modelo los clasifica como cumplidos.

FP (*falsos positivos*): son incumplidos y el modelo los clasifica como cumplidos.

FN (*falsos negativos*): son cumplidos y el modelo los clasifica como incumplidos.

Posteriormente se calcula la sensibilidad (es incumplido y ha sido correctamente clasificado) y especificidad (no es incumplido y ha sido correctamente clasificado) del modelo, utilizando las ecuaciones (3) y (4), respectivamente.

$$\text{Sensibilidad} = \frac{VP}{VP + FP} \quad (3)$$

$$\text{Especificidad} = \frac{VN}{VN + FP} \quad (4)$$

Con el fin de validar la consistencia del comportamiento obtenido con el modelo estimado, se valora el desempeño del modelo evaluando su poder de discriminación de cumplidos e incumplidos utilizando el método Hold Out repetido, en el que se generarán 100 muestras aleatorias de la base de datos original, manteniendo la estructura 80% para la construcción del modelo y 20% para la validación, como se realiza en la estimación del modelo.

• Fase 3: Estimación de pérdida esperada y provisiones

La pérdida esperada es el valor que, por la naturaleza del negocio de crédito, las entidades están dispuestas a perder y se reflejan en la contabilidad como las provisiones. La pérdida se debe a que por diversas razones no todos los deudores pagan sus créditos; además, las entidades no pueden saber de manera anticipada el comportamiento de pago de cada deudor o las dificultades financieras futuras del mismo. Teniendo en cuenta los lineamientos de la SFC (Capítulo II: Reglas relativas a la gestión del riesgo crediticio), la estimación de la pérdida esperada en el marco del sistema de administración de riesgo de crédito (SARC) resulta de la aplicación de la ecuación (5).

$$PE = PI * EA * PDI \quad (5)$$

Donde:

- PE: pérdida esperada o provisión de cartera.
- PI: probabilidad de incumplimiento. Se determina por la calificación que le otorga a cada asociado el modelo construido.
- EA: exposición del activo. Corresponde al valor total neto de dinero que está exponiendo la entidad financiera al otorgar un crédito.
- PDI: pérdida dado el incumplimiento. Depende del tipo de garantía que presente el asociado deudor.

Con el modelo de *scoring* construido y validado en las fases anteriores se estiman las PI para la entidad, mientras que para la estimación de la PDI se tiene en cuenta que un factor importante en la estimación de las pérdidas esperadas en el evento de no pago son las garantías que respaldan las operaciones. Por consiguiente, dado que no hay referentes en la estimación de las PDI por tipo de garantía en el sector solidario, se adoptan los parámetros definidos por la SFC en el capítulo II, como se presenta en la tabla 2.

Finalmente, de acuerdo con la Circular Básica Contable y Financiera 004 de 2008: “todas las organizaciones solidarias sometidas al control y vigilancia de la Superintendencia de la Economía Solidaria, independientemente que califiquen

Tabla 2 - PDI por tipo de garantía

Tipo de garantía	Descripción	PDI
Garantía no admisible	Codeudores y avalistas	55 %
Bienes raíces comerciales y residenciales	Garantías inmobiliarias	40 %
Otros colaterales	Prendas sobre equipo y vehículos	50 %
Sin garantía	Las garantías que no se enuncien en los numerales anteriores	55 %

Fuente: Superintendencia Financiera de Colombia.

y realicen la valoración establecida en la norma, deberán someterse al régimen de provisiones” y deberán constituir como provisión general, “como mínimo una provisión del uno por ciento (1%) sobre el total de la cartera de créditos bruta”. Sin embargo, según las Normas Internacionales de la Información Financiera (NIIF) para las pymes, lo que se llama “provisiones” de acuerdo con el Decreto 2649 de 1993, hoy se denomina “deterioro” y debe ser de forma individual (Cardozo, 2014); esto permite establecer el riesgo crediticio con el fin de registrar la contingencia de pérdida en la cuenta correctora. Es así que dejando a salvo la provisión general, además deberán mantener una provisión individual para la protección de sus créditos.

Resultados

A continuación, se presentan los resultados por cada una de las fases descritas en la metodología.

• Fase 1: Construcción del modelo de scoring

En la construcción del modelo de *scoring* (modelo de regresión logística), se inicia con el procedimiento de selección de variables, presentando los resultados de las pruebas chi-cuadrado de RV para cada uno de los modelos resultantes en cada paso realizado, como se observa en la tabla 3. Se encuentra que:

1. Del primer paso, se observa que la variable *compromiso cuota* presenta la mayor reducción de la desviación, con una reducción de 18.59, siendo esta diferencia estadísticamente significativa, con un *p*-valor de 0.000, indicando que debe ser la primera variable en introducirse en el modelo.
2. Para el segundo paso, se parte del modelo con la variable *compromiso cuota* y se ajustan nuevos modelos para cada una de las variables restantes. De los resultados del test de RV sugiere incluir como segunda variable, la *antigüedad en la entidad*, ya que reduce la desviación en 6.72, con un *p*-valor de 0.009535, siendo estadísticamente significativo.
3. Siguiendo el mismo procedimiento, la tercera variable a ingresar sería *monto deuda*, siendo estadísticamente significativa con un *p*-valor de 0.083.
4. Partiendo del modelo con las variables *compromiso cuota*, *antigüedad* y *monto deuda*, se ajustan modelos para cada

una de las restantes variables y se realiza la prueba de RV. A este nivel, la variable *máximo número de cuotas en crédito de consumo*, aunque presenta la más alta diferencia en razón de verosimilitud (0.53), no es estadísticamente significativa con un *p*-valor 0.467. En este paso, la inclusión de cualquiera de las otras variables no mejora el modelo, por tanto, se queda con las variables ingresadas hasta el paso anterior.

- Para finalizar, se probaron las diferentes interacciones entre las variables ingresadas, observando que el aporte del efecto de las interacciones no fue significativo con un *p*-valor de 0.224.

En resumen, de acuerdo con el procedimiento descrito, las variables a ingresar en el modelo como covariables en la explicación de la probabilidad de incumplimiento, son las variables *compromiso cuota*, *antigüedad* y *saldo deuda por crédito de consumo*, sin ninguna interacción de ningún orden, como se presenta en la tabla 3.

Tabla 3 - Bondad ajuste del modelo: test chi-cuadrado de RV

Paso	Variable a ingresar	Desviación	RV	<i>p</i> -valor	Aporte
Nulo	Ninguna	246.6	-	-	-
Paso 1	Compromiso cuota	227.9	18.6	0.000***	Significativo
Paso 2	Antigüedad en la entidad	221.3	6.7	0.010**	Significativo
Paso 3	Monto deuda	218.3	3.1	0.080*	Significativo
Paso 4	Máximo número de cuotas por pagar	217.7	0.5	0.467	No significativo
Paso 5	Interacciones	216.8	1.5	0.224	No significativo

Significancia al *0.1; **0.05; ***0.01.

Fuente: elaboración propia con datos de la entidad.

Reemplazando las estimaciones de los parámetros de la tabla 4 en la ecuación (1), la probabilidad de incumplimiento para la entidad de economía solidaria queda expresada de la siguiente manera:

$$\pi = \frac{1}{1 + \exp(-Z)} \quad (6)$$

donde:

$$Z = -3,416 + (0,041 * \text{CompromisoCuota}) - (0,055 * \text{Antigüedad}) + (0,022 * \text{MontoDeuda})$$

Contiene las variables significativas en la estimación del riesgo crediticio del asociado para la entidad y sus respectivas ponderaciones en el modelo.

Los contrastes sobre los parámetros se realizan a través del test de Wald. Según puede verse en la tabla 4, los *p*-valores asociados a las variables *compromiso cuota* y *antigüedad* son mucho menores de 0.05, mientras que para la variable *monto de la deuda* es menor a 0.1. Por tanto, se obtiene que las tres variables son significativas, es decir, sus coeficientes son significativamente distintos de 0, sus OR (Odds Ratios) son significativamente distintas de 1 y los intervalos de confianza para las mismas no contienen el 1.

Los parámetros del modelo se interpretan en términos de cocientes de ventajas. Se entiende que cuando se habla de una variable en concreto, las conclusiones extraídas asumen que el resto de las variables permanecen fijas.

- *Compromiso cuota*, presenta un OR = 1.042, e indica que por cada unidad (1%) que aumente el *compromiso de la cuota* de un asociado, aumenta algo más de una (1) vez la posibilidad a favor de ser incumplido en el pago de la obligación.

- *Antigüedad en la entidad*, obtuvo un OR = 0.947, e indica que por cada año menos de antigüedad en la entidad, la posibilidad a favor de ser incumplido en el pago de la obligación aumenta en 1.06 veces.

- *Monto deuda por crédito de consumo*, muestra un OR = 1.004, lo que significa que por cada \$100000 que aumenta el saldo de la deuda, aumenta en 1.04 veces la posibilidad a favor de ser incumplido en el pago de la obligación.

De lo anterior, se puede pensar en dos posibles perfiles de asociados en cuanto al riesgo de incumplimiento para la entidad: por un lado, son más propensos al riesgo aquellos que presenten mayores compromisos en cuota, indicando que tendrían comprometido una mayor parte de sus ingresos. En cuanto al tiempo de antigüedad en la entidad, cabe notar que esta variable está directamente relacionada con la edad de los asociados y, por ende, se podría pensar en personas que son mucho más jóvenes que no presentan ingresos muy altos; y finalmente, el monto de la deuda, que entre más grande sea, es sinónimo de un mayor nivel de endeudamiento. Siendo coherente con el sentido de los signos en los parámetros estimados y los resultados obtenidos en la parte descriptiva del estudio. Por otro lado, estarían los menos propensos, como aquellos que presenten un compromiso de cuota más bajo, lleven más tiempo en la entidad y el monto de la deuda sea menor.

Tabla 4 - Estimación de parámetros para la probabilidad de incumplimiento

Variable explicativa	Coefficiente	Error estándar	Wald	<i>p</i> -valor	OR	I.C. 95% para EXP(B)	
						inferior	superior
Intercepto	-3.416	0.433	62.3	0.000***	0.03	-	-
Compromiso cuota	0.041	0.013	9.7	0.002***	1.04	1.015	1.069
Antigüedad en la entidad	-0.055	0.020	7.6	0.006***	0.95	0.911	0.984
Monto deuda	0.002	0.001	3.1	0.078*	1.00	1.001	1.004

Significancia al *0.1; **0.05; ***0.01.

Fuente: elaboración propia con datos de la entidad.

• Fase 2: Validación del modelo

Una vez estimado el modelo, se valora su nivel de ajuste y capacidad predictiva. La figura 1 presenta la curva característica de operación para el modelo ajustado, presentando un área de 76.2%, lo que permite considerar que el modelo estimado es adecuado en su capacidad de discriminación, ya que, según Hosmer et al. (2004), un modelo se considera preciso y tiene alta capacidad de discriminación cuando el área es al menos de 70%. En cuando a la bondad de ajuste del modelo, la figura 2 presenta las sensibilidades y especificidades obtenidas para diferentes puntos de corte entre 0.05 y 0.9, tanto para el 80% de los datos con que se construyó el modelo, como para el 20% que se dejó para la validación. Para ambos casos se obtiene que las curvas convergen al mismo punto de corte del 0.05, donde se maximiza la sensibilidad y especificidad del modelo.

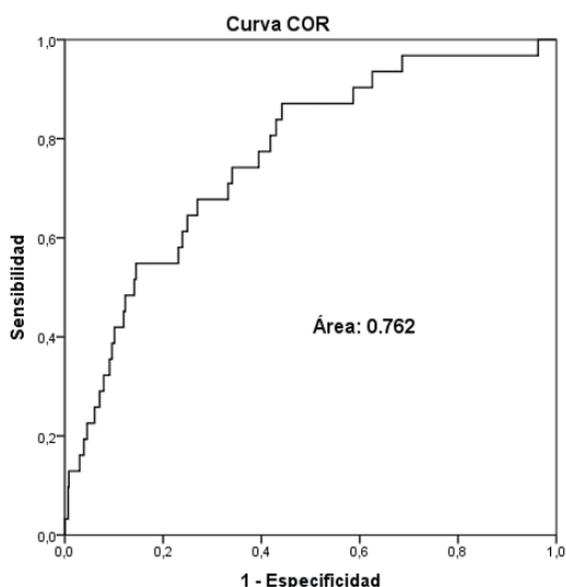


Figura 1 - Curva ROC del modelo estimado
Fuente: elaboración propia con datos de la entidad.

La figura 3 presenta las sensibilidades y especificidades para cada uno de los 100 modelos estimados, tanto para los datos con que se construyó el modelo, como para el 20% que se dejó para la validación, observando un comportamiento consistente con el obtenido en el modelo estimado para la entidad en el que la máxima sensibilidad y especificidad del modelo se obtiene para un punto de corte del 0.05. Este punto de corte indica el punto a partir del cual debe clasificarse a un asociado como posible *default* para la entidad.

De lo anterior se obtiene que para la población de asociados a la entidad el *score* de clasificación entre *cumplido* e *incumplido* es del 5%. Sin embargo, tomando como referencia que en la mayoría de las entidades financieras el *score* de clasificación está entre 50% y 65%, se sugiere que la utilización del modelo esté alineada como un generador de alarmas, en cuyos casos de asociados identificados como posibles *default* sean objeto de un estudio más detallado, complementario al proceso de otorgamiento de créditos que realiza la entidad.

• Fase 3: Estimación de pérdida esperada y provisiones

Una vez estimado y validado el modelo, para efectos de los registros en los estados financieros y reportes a las centrales de información, las entidades deben homologar las calificaciones de incumplimiento de sus modelos internos. Por tanto, el modelo construido, además de discriminar entre *cumplidos* e *incumplidos*, asigna una calificación a los clientes necesaria para determinar las provisiones de cartera. Con este propósito y siguiendo las buenas prácticas enunciadas por la SFC, se basó en la regla de calificación expuesta en el modelo de referencia de consumo, encontrando así los puntos de corte y escalas de calificación, como se muestra en la tabla 5. Una vez recalificados los asociados de acuerdo con su probabilidad de incumplimiento, la entidad deberá mantener una provisión individual mínima en los porcentajes que se relacionan en la tabla 6.

Finalmente, tomando como base la cartera de un mes específico, se tiene un saldo expuesto de \$10868203851, de los cuales se estima una pérdida esperada de \$974455726,

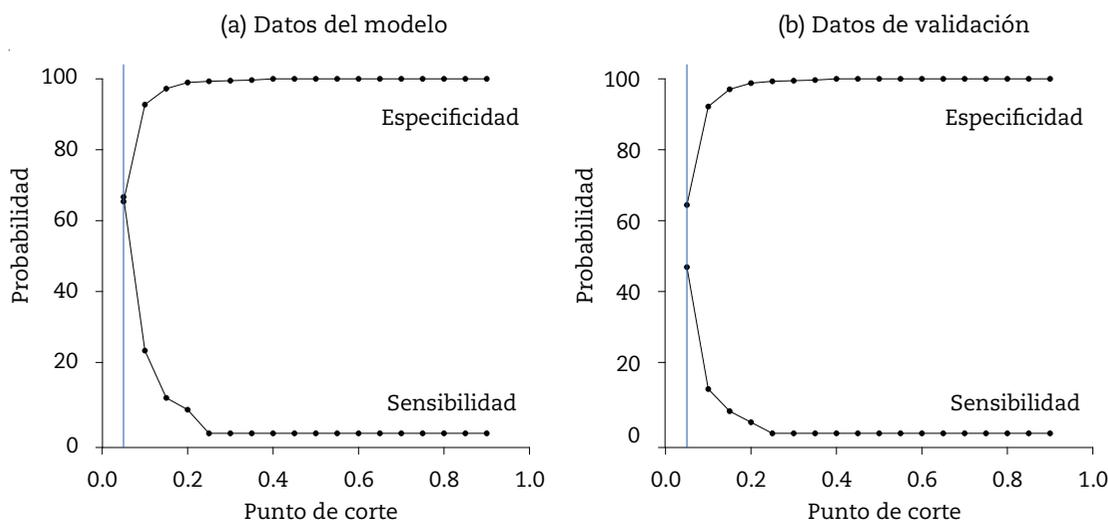


Figura 2 - Sensibilidad y especificidad: modelo estimado
Fuente: elaboración propia con datos de la entidad.

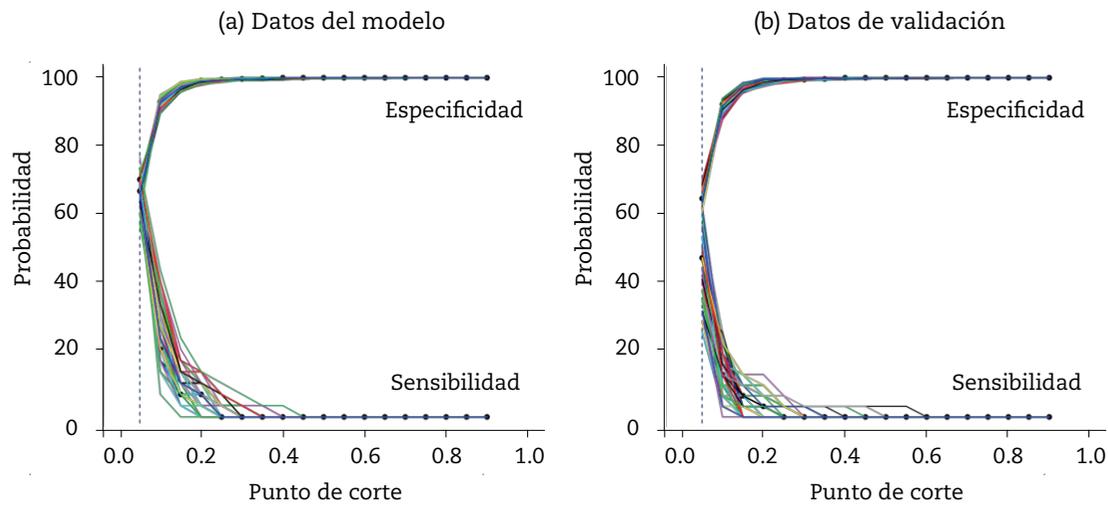


Figura 3 - Sensibilidad y especificidad: remuestreo (100 modelos)
Fuente: elaboración propia con datos de la entidad.

Tabla 5 - Recalificación de categorías por probabilidad de incumplimiento

Categoría	Probabilidad de incumplimiento (%)		Categorías agregadas
	Mínimo	Máximo	
AA	0	3	A
A	>3	5	
BB	>5	28	B
B	>28	40	
CC	>40	53	C
C	>53	70	
D	>70	83	D
E	>82	100	

Fuente: elaboración propia.

Tabla 6 - Provisión individual

Calificación	Provisión individual
A	0%
B	1%
C	10%
D	20%
E	100%

Fuente: Superintendencia de la Economía Solidaria.

calculada a partir de la ecuación (1) definida en la primera etapa de la metodología, donde la PI es estimada por el modelo construido y la PDI es estimada a partir del referente definido en la tabla 1 dependiendo el tipo de garantía. De la pérdida esperada, se estima la provisión general e individual para la entidad solidaria, siendo de \$108 millones y \$418 millones, respectivamente, como se muestra en la tabla 7. Sin embargo, “la decisión de constituir una provisión general superior al mínimo exigido corresponderá a una

política adoptada por el consejo de administración, junta directiva o quien haga sus veces” (Circular Básica Contable y Financiera 004 de 2008).

Tabla 7 - Provisión entidad

Evento	Monto
Saldo expuesto	\$ 10868203851
Pérdida esperada	\$ 974455726
Provisión general del 1% (norma)	\$ 108682039
Provisión individual	\$ 418165742
Provisión total	\$ 526847781

Fuente: elaboración propia.

Conclusiones

Usando la información propia de la entidad, en condiciones de bajo *default*, se logra construir en este trabajo un modelo para determinar el riesgo de crédito y para limitar y tener conocimiento sobre el nivel de riesgo que puede afectar el capital de la entidad, logrando una adecuada cobertura acorde con la realidad de su portafolio.

Los resultados obtenidos permiten valorar el uso de la regresión logística, ya que por ser una herramienta flexible logra modelar de modo sencillo y elegante la probabilidad de incumplimiento de los asociados. El modelo admite el uso de covariables de cualquier naturaleza, sean categóricas, discretas o continuas y limita los supuestos a la independencia de variables aleatorias independientes provenientes de una única distribución de la familia exponencial. Estos elementos teóricos, unidos a su riqueza interpretativa, hacen de esta herramienta un instrumento versátil para perfilar a los asociados y, a partir de allí, proponer la implementación de modelos acordes con el entorno en que se desenvuelve el sector solidario.

Otro elemento destacable a partir de los resultados obtenidos es la necesidad de seguir utilizando métodos clásicos, como el de regresión logística para la evaluación de riesgo de crédito; ya que en algunos casos, como el de la entidad utilizada en el artículo, el tamaño de la base de usuarios de los productos crediticios y sus hábitos de pago dificultan la utilización de métodos basados en el manejo de grandes cantidades de datos. De hecho, los bajos niveles de morosidad que presenta la entidad analizada, obligaron a una redefinición de los umbrales de impago, para el modelo de riesgo, considerando como clientes riesgosos aquellos que han obtenido en el algún momento de su vida crediticia calificación B. Esto presenta una diferencia sustancial con los modelos de riesgo crediticio que normalmente se usan en el sector financiero tradicional.

Financiación

El artículo es resultado posterior del trabajo de grado de maestría de Iván Mauricio Bermúdez con la asesoría de los profesores Diego Manotas y Javier Olaya. No contó con financiación.

REFERENCIAS

- Abdolreza, N., Farnoosh, F., Konstantin H., & Frank, J. (2017). Fuzzy decision fusion approach for loss-given-default modeling. *European Journal of Operational Research*, (262), 780-791. <https://doi.org/10.1016/j.ejor.2017.04.008>.
- Altman, E. I., & Saunders, A. (1997). Credit risk measurement: Developments over the last 20 years. *Journal of Banking & Finance*, 21(11-12), 1721-1742. [https://doi.org/10.1016/S0378-4266\(97\)00036-8](https://doi.org/10.1016/S0378-4266(97)00036-8).
- Bennouna, G., & Tkiouat, M. (2019). Scoring in microfinance: Credit risk management tool –Case of Morocco–. *Procedia Computer Science*, 148, 522-531. <https://doi.org/10.1016/j.procs.2019.01.025>.
- Bülbül, D., Hakenes, H., & Lambert, C. (2019). What influences banks' choice of credit risk management practices? Theory and evidence. *Journal of Financial Stability*, 40, 1-14. <https://doi.org/10.1016/j.jfs.2018.11.002>.
- Caicedo, E., Claramunt, M., & Casanovas, M. (2011). Medición del riesgo de crédito mediante modelos estructurales: una aplicación al mercado colombiano. *Cuadernos de Administración*, 24(42), 73-100. http://www.scielo.org.co/scielo.php?script=sci_abstract&pid=S0120-35922011000100004.
- Cardozo, H. (2014). *IFRS/NIIF para el sector solidario. Aplicadas a las PYME: modelo catálogo de cuentas*. Bogotá: Ecoe Ediciones.
- Chadwick, M. G. (2018). Effectiveness of monetary and macroprudential shocks on consumer credit growth and volatility in Turkey. *Central Bank Review*, 18(2), 69-83. <https://doi.org/10.1016/j.cbrev.2018.03.001>.
- Circulares externas 003 de 2013 y 006 de 2014. Superintendencia de la Economía Solidaria.
- Circular Básica Contable y Financiera 004 de 2008. Superintendencia de la Economía Solidaria.
- Dobson, A. J., & Barnett, A. G. (2018). *An introduction to generalized linear models*. Florida: Chapman & Hall/CRC.
- Echeverri, F. (2006). *Evaluación de modelos para la medición de riesgo de incumplimiento en créditos para una entidad financiera del Eje Cafetero* (tesis de maestría). Manizales, Colombia: Universidad Nacional de Manizales. <http://www.bdigital.unal.edu.co/1083/1/faneryecheverri-valdes.2006.pdf>.
- Fang, F., & Chen, Y. (2019). A new approach for credit scoring by directly maximizing the Kolmogorov–Smirnov statistic. *Computational Statistics & Data Analysis*, 133, 180-194. <https://doi.org/10.1016/j.csda.2018.10.004>.
- Gómez, J. E., & Orozco, I. P. (2009). Un modelo de alerta temprana para el sistema financiero colombiano. *Borradores de Economía*, 565. <http://www.banrep.gov.co/docum/ftp/borra565.pdf>.
- González, M., & González, C. (2019). An alternative approach to predicting bank credit risk in Europe with Google data. *Finance Research Letters*. <https://doi.org/10.1016/j.frl.2019.08.029>.
- Hosmer, D. W., & Lemeshow, S. (2004). *Applied logistic regression*. Nueva York: John Wiley and Sons.
- Jokivuolle, E., & Peura, S. (2003). Incorporating collateral value uncertainty in loss given default estimates and loan-to-value ratios. *European Financial Management*, 9(3), 299-314. <https://doi.org/10.1111/1468-036X.00222>.
- Kozodoi, N., Lessmann, S., Papakonstantinou, K., Gatsoulis, Y., & Baesens, B. (2019). A multi-objective approach for profit-driven feature selection in credit scoring. *Decision Support Systems*, 120, 106-117. <https://doi.org/10.1016/j.dss.2019.03.011>.
- López, F. (2014). Pero ¿existe demanda solvente de crédito? *Estrategia Financiera*, 315, 76-77.
- Madeira, C. (2019). Measuring the covariance risk of consumer debt portfolios. *Journal of Economic Dynamics and Control*, 104, 21-38. <https://doi.org/10.1016/j.jedc.2019.05.005>.
- Nelder, J. A., & Wedderburn, R. W. (1972). Generalized linear models. *Source Journal of the Royal Statistical Society. Series A (General)*, 135(3), 370-384. <https://doi.org/10.2307/2344614>.
- Ochoa, J., Galeano, W., & Agudelo, L. (2010). Construcción de un modelo de scoring para el otorgamiento de crédito en una entidad financiera. *Perfil de Coyuntura Económica*, 16, 191-222. <https://www.redalyc.org/articulo.oa?id=86120022009>.
- R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Salazar, F. E. (2013). Cuantificación del riesgo de incumplimiento en créditos de libre inversión: Universidad del Valle. *Estudios Gerenciales*, 29, 416-427. <https://doi.org/10.1016/j.estger.2013.11.007>.
- Támara, A., Aristizábal, R., & Velásquez, E. (2012). Matrices de transición en el análisis del riesgo crediticio como elemento fundamental en el cálculo de la pérdida esperada en una institución financiera colombiana. *Revista Ingenierías Universidad de Medellín*, 11(20), 105-114. <https://revistas.udem.edu.co/index.php/ingenierias/article/view/669>.
- Walke, A. G., Fullerton, T. M., & Tokle, R. J. (2018). Risk-based loan pricing consequences for credit unions. *Journal of Empirical Finance*, 47, 105-119. <https://doi.org/10.1016/j.jempfin.2018.02.006>.
- Yu, F., Ching, W., Gu, J., & Siu, T. (2017). Interacting default intensity with hidden Markov process. *Quantitative Finance*, 7(5), 781-794. <https://doi.org/10.1080/14697688.2016.1237036>.