

“Propuesta de un Modelo Para Otorgamiento de Créditos y Estimación de Primas de Seguros en el Sector Agropecuario”

EMMANUEL SORIANO BASILIO*
GERMÁN ANÍBAL NARVÁEZ VÁSQUEZ *
JAVIER CHÁVEZ FERREIRO¹

Resumen:

El presente estudio tiene como objetivo proponer un modelo estadístico para el otorgamiento de créditos en el sector agropecuario que permita generar una adecuada administración del riesgo de crédito. (Caso de los Fideicomisos Instituidos en Relación con la Agricultura, FIRA). La metodología empleada fue la técnica de discriminación, comúnmente conocida como *credit scoring*. Las tablas de puntaje permiten determinar probabilidades de incumplimiento de los créditos garantizados, dichas probabilidades y el saldo contingente son los insumos para estimar el riesgo de crédito a través de tres modelos usualmente utilizados en la práctica: CyRCE, Montecarlo y Credit Risk+. Los resultados obtenidos permitieron concluir que los parámetros determinados a través de un tablero de puntaje con base en una regresión logística y la estimación del riesgo de crédito con el modelo Montecarlo, permiten tener un balance entre los ingresos y egresos.

Modelo - Riesgo Crediticio – FIRA

Abstract:

This study aims to propose a statistical model for lending in the agricultural sector that will generate adequate credit risk management. (Case of the Trust Funds to Agriculture, FIRA). The methodology used was the discrimination technique, commonly known as credit scoring. Score tables for determining default probabilities of secured claims, and these probabilities are contingent balance inputs to estimate credit risk through three models commonly used in practice: CyRCE, Montecarlo and Credit Risk +. The results allowed us to conclude that the parameters determined through a scoreboard based on logistic regression and estimation of credit risk with the Monte Carlo model, let have a balance between revenue and expenditure.

Modelo - Riesgo Crediticio – FIRA

¹ Universidad La Salle Morelia

1. Introducción

Las instituciones financieras en la búsqueda de rendimiento de sus activos cuentan con diferentes productos dirigidos a sus clientes. En particular, el otorgamiento de créditos a personas físicas y morales es uno de sus principales servicios financieros. Las instituciones buscan que los ingresos y los egresos por dicho producto sean tal que los primeros sean iguales o mayores a los segundos; sin embargo, de manera natural al otorgar un crédito las instituciones financieras se encuentran ante un acreditado que pudiera o no cubrir sus obligaciones adquiridas. El presente trabajo tiene como objetivo desarrollar una herramienta integral para la administración del riesgo de crédito, de aplicación general para las instituciones financieras y en particular para los Fideicomisos Instituidos en Relación con la Agricultura (FIRA).

Para lograr dicho objetivo, en primer lugar se presentan los objetivos generales y específicos del trabajo, y las hipótesis relacionadas con dichos objetivos las cuales son aceptadas o rechazadas a través de indicadores de variables citadas en la tabla de congruencia metodológica. En segundo lugar se plantea el marco teórico sobre el cual se sustenta la metodología del presente trabajo. En particular se desarrolla el método de tablas de puntaje, también conocido como *credit scoring*, el cual es una herramienta que permite clasificar a los acreditados en categorías de comportamiento: buenos y malos. Con base en las tablas de puntaje se determinan parámetros que sirven de insumo a los modelos de riesgo de crédito que se plantean los cuales indican un nivel de pérdida para un cierto nivel de confianza. Tanto las estimaciones de las tablas de puntaje como las obtenidas a través de los modelos de riesgo de crédito son la base para la estimación de una prima de seguro basada en un modelo teórico desarrollado de manera interna en FIRA. En tercer lugar, se aplica el marco teórico a la población objetivo del presente trabajo, la cual es una base de datos de créditos garantizados por FIRA entre los años 2006 y 2011. Los resultados obtenidos son contrastados con los efectivamente observados en el mismo periodo. Por último, se muestran de manera conjunta los indicadores de variables obtenidos de forma tal que permitan definir bajo que condiciones se aceptan o rechazan las diferentes hipótesis.

2. Fundamentos De La Investigación

2.1 Antecedentes

Las instituciones financieras a través del otorgamiento de créditos se convierten en uno de los pilares para el crecimiento económico de un país o región. La solidez de la estructura de capital de dichas instituciones contribuye a un sistema financiero estable que ayudará a atenuar impactos económicos y financieros. Para lograr la solidez, los gobiernos deben emitir reglas de supervisión y requerimientos de capital para reducir los riesgos de insolvencia.

Lo anterior llevó a las autoridades de supervisión de Bélgica, Canadá, Francia, Alemania, Italia, Japón, Luxemburgo, Holanda, Suecia, Suiza, Inglaterra y Estados Unidos a formar el Comité de Basilea para la Supervisión Bancaria el cual emitió el documento “Convergencia internacional de medición de capital y estándares de capital”, también denominado “El acuerdo BIS 1988” (después conocido como Basilea I). Este acuerdo, exigía que todos los préstamos a una corporación fueran ponderados por riesgo de crédito en un 100%, sin importar la calidad crediticia de la contraparte. Por lo anterior, y para adecuarse a la operativa de los bancos, en 1999 el comité emitió nuevas reglas conocidas como Basilea II basadas en tres pilares: i) Requerimientos de capital, ii) Revisión del supervisor, iii) Disciplina de mercado.

Según Sandica (2010), la base para una mejor estimación del riesgo de crédito es definir la probabilidad de incumplimiento de la contraparte que tiene una obligación por cumplir ante la institución financiera. Al respecto se han desarrollado varias técnicas, entre ellas el desarrollo de un modelo de puntaje (credit scoring). El pionero de dicho modelo es Durand cuando en 1941 aplicó el análisis de discriminante propuesto por Fisher (1936) para clasificar a los acreditados en buenos y malos con base en las características de los mismos. (Sandica, 2010).

Los modelos de puntaje permiten realizar la aprobación de un crédito de forma tal que la probabilidad de incumplimiento puede ser inferida. Por lo tanto, con un modelo de puntaje el intermediario financiero se encuentra en condiciones de otorgar créditos a un nivel deseado de riesgo.

2.2 Problemática

Los Fideicomisos Instituidos en Relación con la Agricultura (FIRA) se encuentra constituido por cuatro fideicomisos denominados: FEFA, FONDO, FOPESCA y FEGA. Para cumplir su misión, FIRA apoya proyectos productivos relacionados con el sector rural fundamentalmente a través de intermediarios financieros bancarios (Bancos) e intermediarios financieros no bancarios (IFNBs) como SOFOLES, SOFOMES y Uniones de Crédito. A los intermediarios financieros se les ofrecen principalmente dos productos:

- 1) Fondeo a créditos: FIRA entrega recursos a los intermediarios financieros para que los mismos sean otorgados a los acreditados finales, a través de los fideicomisos FEFA, FONDO y FOPESCA.
- 2) Servicio de garantía: FIRA, a través del fideicomiso FEGA, garantiza a los intermediarios financieros el pago de los acreditados finales. Los intermediarios contratan el pago de un porcentaje del saldo insoluto cuando el acreditado incumple sus obligaciones, lo cual se le denomina servicio de garantía. (FEGA cobra una prima de seguro por brindar dicho servicio).

La problemática surge en el procedimiento de pago de garantía, donde se ha podido identificar los siguientes puntos: i) No se evalúa la probabilidad de incumplimiento del acreditado con base en sus

características, ii) No se cuenta con probabilidades de incumplimiento para la estimación del riesgo de crédito (pérdida esperada y no esperada), iii) La prima de garantía se cobra en función del intermediario financiero; sin embargo, el evento que detona el pago es el incumplimiento del acreditado, iv) La falta del conocimiento de la probabilidad de incumplimiento de un acreditado no permite generar políticas de administración de riesgo de crédito que permitan conservar el patrimonio de una institución así como incrementar sus beneficios.

2.3 Pregunta de investigación

¿Un modelo adecuado para el otorgamiento de créditos garantizados en el sector agropecuario puede ser la base para la gestión del riesgo de crédito?. Como preguntas específicas tenemos:

- ¿Cuál es el modelo más adecuado para el otorgamiento de créditos en el sector agropecuario?
- ¿Cómo se puede estimar el riesgo de crédito a través del modelo para el otorgamiento de créditos?
- ¿Cuáles es la prima de riesgo asociada a las características del modelo de otorgamiento?
- ¿Cuál es la relación entre el modelo de otorgamiento de créditos y la gestión del riesgo de crédito?

2.4 Objetivo de la investigación

Determinar un modelo adecuado en el otorgamiento de créditos en el sector agropecuario que sea la base para la gestión del riesgo de crédito. Para llegar a cumplir con el objetivo fue necesario abordar los siguientes objetivos específicos:

- Determinar un modelo adecuado para el otorgamiento de créditos en el sector agropecuario.
- Determinar, con base en el modelo desarrollado de otorgamiento, el riesgo de crédito del portafolio de créditos garantizados.
- Estimar las primas de seguros en el sector agropecuario con base en el modelo de otorgamiento de créditos.
- Identificar la relación entre el modelo de otorgamiento de créditos y la gestión del riesgo de crédito.

2.5 Hipótesis

Un modelo adecuado para el otorgamiento de créditos en el sector agropecuario permitirá generar una adecuada administración del riesgo de crédito.

Variables independientes: Modelo para el otorgamiento de créditos (Modelo de discriminación, riesgo de crédito, primas de seguro y riesgo beneficio)

Variable dependiente: Adecuada administración del riesgo de crédito (Conservación del patrimonio e incremento de los beneficios).

2.5.1 Hipótesis Específicas:

H1: Un modelo con mayor discriminación mejorará el otorgamiento de créditos en el sector agropecuario.

H2: A mayor discriminación del modelo mejor estimación del riesgo de crédito.

H3: Un mejor modelo de discriminación permitirá estimar primas de seguros de equilibrio en el sector agropecuario.

H4: A mayor discriminación del modelo de otorgamiento de créditos, mayor relación riesgo y beneficio.

3. Justificación

El Producto Interno Bruto del sector agropecuario ha incrementado en promedio del primer trimestre de 2008 al segundo trimestre del 2011 el 0.5%. Una de las variables asociadas a la producción del sector es el otorgamiento de créditos ya sea a través de fondeo ó garantías.

Los créditos otorgados por los intermediarios financieros en el sector agropecuario, con ó sin fondeo de FIRA, pueden ser garantizados por FEGA a cambio del pago de una prima de seguro ante el incumplimiento del acreditado. La prima de garantía cobrada por FEGA se encuentra en función de los pagos y recuperaciones que realiza cada intermediario financiero, no se toma como base al acreditado final, quien es el detonante del pago del seguro. Así mismo, el riesgo de crédito no se encuentra asociado a la probabilidad de incumplimiento del acreditado final, además de solo contar con un modelo para la estimación del riesgo. Por lo anterior, se torna necesario el desarrollo de un modelo de puntaje para el otorgamiento de créditos que permita: i) Agilizar el otorgamiento de créditos en el sector agropecuario, ii) Conservar el patrimonio de FEGA, el cual es de carácter gubernamental, iii) Gestionar eficientemente el riesgo de crédito.

4. Marco Conceptual

4.1 Modelo de Puntaje (*Credit Scoring*)

El Modelo de Puntaje, es un conjunto de modelos de decisión y sus técnicas subyacentes que ayudan a los prestamistas en la concesión de créditos al consumidor. Estas técnicas deciden quien obtendrá los créditos, y que estrategias operacionales mejoraran la ganancia de los prestamistas. Las técnicas de los modelos de puntaje (*credit scoring*) evalúan el riesgo en prestar a un consumidor en particular. (Thomas, Edelman, & Crook, 2002).

Los prestamistas deben tomar dos tipos de decisiones: a) conceder el crédito a un nuevo solicitante y b) como manejar a los solicitantes existentes, incluyendo si se les debe incrementar el límite de crédito. En ambos casos, el modelo de puntaje es un predictor del riesgo y se basa en muestras de comportamiento de créditos otorgados en el pasado cuyos consumidores se esperaba sean similares a aquellos que serán evaluados con el modelo de puntaje. El modelo de puntaje para la primera decisión es conocido como Score de Originación y para el segundo es un Score de comportamiento. En particular, para el presente trabajo se desarrolla un Score de Originación.

4.1.1 Modelo de Discriminación

Si bien los métodos estadísticos fueron los primeros en ser usados para construir las tablas de puntaje, estos han sido modificados. El método estadístico más común actualmente es la regresión logística (modelo logit) que toma como insumo el análisis de discriminante. Otros métodos son la regresión lineal y clasificación de árboles (modelo de árboles) el cual separa el conjunto de solicitantes en un cierto número de subgrupos en función de sus atributos y luego clasifica los subgrupos en satisfactorios o insatisfactorios. (Thomas, Edelman, & Crook, 2002).

En el proceso de otorgamiento de crédito el prestamista debe elegir entre dos opciones: otorgar o rechazar el crédito. El modelo de puntaje intenta ayudar a dicha elección. Por lo anterior, el modelo clasifica al acreditado en uno de los siguientes dos grupos: buenos o malos. Bueno sería un comportamiento aceptable para el prestamista mientras malo significa que el prestamista hubiera deseado rechazarlo.

4.1.2 Separación de dos grupos: Análisis de Discriminante

Los dos grupos a clasificar por el prestamista deben ser buenos y malos. Fisher sugirió que una medida M de sensibilidad de separación de poblaciones con varianza muestral común es:

$$M = \frac{\text{Distancia entre medias muestrales de dos grupos}}{(\text{Varianza muestral de cada grupo})^{1/2}}$$

Si m_G , m_B y S son la media de la muestra del grupo de buenos, la media de los grupos malos y la varianza común respectivamente, la ecuación anterior se redefine como

$$M = \mathbf{w}^T \frac{\mathbf{m}_G - \mathbf{m}_B}{(\mathbf{w}^T \mathbf{S} \mathbf{w})^{1/2}} \quad (1)$$

Diferenciando a la ecuación 1 respecto a \mathbf{w} , el valor de M se maximiza cuando

$$\frac{\mathbf{m}_G - \mathbf{m}_B}{(\mathbf{w}^T \mathbf{S} \mathbf{w})^{1/2}} - \frac{(\mathbf{w}(\mathbf{m}_G - \mathbf{m}_B))^T (\mathbf{S} \mathbf{w}^T)}{(\mathbf{w} \mathbf{S} \mathbf{w}^T)^{3/2}} = 0$$

$$\text{Entonces: } (\mathbf{m}_G - \mathbf{m}_B)(\mathbf{w} \mathbf{S} \mathbf{w}^T) = (\mathbf{S} \mathbf{w}^T)(\mathbf{w}(\mathbf{m}_G - \mathbf{m}_B))^T \quad (2)$$

Intuitivamente la tabla de puntaje busca separar los grupos de forma tal que se busque una ecuación que maximice la separación entre las medias.

4.1.3 Valor de la Información (VI)

El VI permite estimar el valor de predicción de una característica con base en los atributos de esta. (Siddiqi, 2006). La ecuación para estimar esta dada por:

$$VI = \sum_{i=1}^a \ln \left(\frac{f(x_i)_g}{f(x_i)_b} \right) (f(x_i)_g - f(x_i)_b) \quad (3)$$

4.1.4 Regresión Lineal

Los modelos de regresión lineal intentan explicar el valor de una variable dependiente Y en función de valores de una o varias variables explicativas X . Las variables X pueden ser cuantitativas o cualitativas, siendo las primeras aquellas que toman un valor numérico y las cualitativas, también conocidas como variables dummy, son aquellas que implican la existencia o ausencia de una cualidad o atributo, como el ser varón o mujer. El método de cuantificar los atributos consiste en construir variables artificiales que asumen el valor 1 si se tiene la presencia del atributo y 0 en caso contrario. (Gujarati, Principios de econometría, 2006)

La relación entre la variable dependiente y la variable explicativa se encuentra dada en la siguiente ecuación:

$$Y = w_0 + w_{1,1}x_{1,1} + w_{1,2}x_{1,2} + \dots + w_{1,m-1}x_{1,m-1} + \dots + w_{n,1}x_{n,1} + \dots + w_{n,m-1}x_{n,m-1} \quad (4)$$

Si la variable cualitativa (característica) i ($i=1, \dots, n$) tiene m categorías (atributos), sólo hay que introducir $m-1$ variables dummy. Esto se extiende para cada una de las n características que serán utilizadas para pronosticar el valor de Y .

La variable dependiente Y puede tomar dos valores: 1 si el crédito es clasificado como malo y 0 en otro caso. Con la ecuación (4) se busca encontrar los valores de las w que mejor expliquen el comportamiento de los créditos; sin embargo, los resultados que se estiman pueden ser mayores a 1. A pesar de lo anterior, los resultados de la regresión lineal difieren de la regresión logística solo en las colas (Greene, 2011), la cual es la regresión más utilizada para modelos de puntaje. Por lo tanto, los resultados de la regresión lineal son igualmente válidos para obtener un modelo de puntaje. (Thomas, Edelman, & Crook, 2002)

La ecuación (4) puede ser resuelta mediante el método de Mínimos Cuadrados Ordinarios el cual estima los valores de las w 's. La solución se encuentra dada en forma matricial mediante (Gujarati & Porter, Econometría, 2010):

$$W = (X'X)^{-1}X'Y \quad (5)$$

Donde X es una matriz que contiene 0 y 1 indicando las características de cada variable cualitativa en cada uno de los créditos, X' es la matriz traspuesta de X , y Y es un vector que contiene 0 y 1 en función del comportamiento de cada crédito.

4.1.5 Regresión Logística

La regresión logística es una función discriminante que intenta encontrar la mejor combinación de características que expliquen la probabilidad de incumplimiento. Si p_i es la probabilidad de que el solicitante i en la muestra ha incumplido, es decir que Y_i valga 1, es deseable encontrar un vector w tal que (Gujarati & Porter, Econometría, 2010):

ó negativos. A

$$\log\left(\frac{p_i}{1-p_i}\right) = w_0 + w_1x_1 + \dots + w_px_p = wx^T \quad (6)$$

Tomando exponenciales de ambos lados de la ecuación (18) se obtiene:

$$p_i = \frac{e^{w'x}}{1+e^{w'x}} = F(x'_i w) \quad (7)$$

A diferencia de la regresión lineal, la regresión logística permite que la variable p_i solo tome valores entre 0 y 1, el cual es el rango esperado para una variable que estime una probabilidad (en este caso de incumplimiento).

Los componentes del vector w son encontrados a través del método de máxima verosimilitud, mediante encontrar el vector que maximice el logaritmo de la función de verosimilitud la cual es (Novales, 1993).

$$L = \prod_{Y_i=1} F(x'_i w) \prod_{Y_i=0} [1 - F(x'_i w)] \quad (8)$$

4.1.6 Peso de la evidencia

Tanto la regresión lineal como el modelo de regresión logística pueden realizarse sobre variables dummies o con base en el Peso de la Evidencia o WOE por sus siglas en inglés (*Weight of Evidence*). El $WOE_{i,k}$ del atributo i que pertenece a la característica k se estima a través de:

$$WOE_{i,k} = \sum_{i=1}^a \ln\left(\frac{f(x_i)_g}{f(x_i)_b}\right) * 100 \quad (9)$$

La práctica sugiere la utilización del WOE en lugar de las variables dummies.

4.1.7 Tablero de puntaje

En general la relación entre odds (división de buenos a malos) y el puntaje puede ser presentado como una transformación lineal: (Siddiqi, 2006)

$$\text{Score} = \text{offset} + \text{factor} * \ln(\text{odds}) \quad (10)$$

El tablero de puntaje se desarrolla utilizando odds especificados a un puntaje y “puntos para doblar odds” (pdo) también especificados. El factor y el offset pueden ser calculados mediante resolver las siguientes ecuaciones:

$$\text{Score} = \text{offset} + \text{factor} * \ln(\text{odds})$$

$$\text{Score} + \text{pdo} = \text{offset} + \text{factor} * \ln(2 * \text{odds})$$

Resolviendo para las diferentes variables se obtiene:

$$\text{Pdo} = \text{factor} * \ln(2) \quad (11)$$

$$\text{Factor} = \text{pdo} / \ln(2) \quad (12)$$

$$\text{Offset} = \text{score} - \text{factor} * \ln(\text{odds}) \quad (13)$$

Por ejemplo, si se quiere una relación de 50:1 en 600 puntos y doblar puntaje (pdo) cada 20 puntos, entonces el factor y el offset son 28.8539 y 487.123 respectivamente. Cada puntaje correspondiente a cada conjunto de odds (o cada atributo) puede ser calculado como:

$$\text{Score} = 487.123 + 28.8539 \ln(\text{odds}) \quad (14)$$

Dado que el tablero de puntaje se realiza utilizando el WOE como insumo, la ecuación (10) puede ser modificada como:

$$\begin{aligned} \text{score} &= \ln(\text{odds}) * \text{factor} + \text{offset} = -\left(\sum_{j,i=1}^{k,n} (\text{woe}_j * w_i) + a\right) * \text{factor} + \text{offset} = \\ &= -\left(\sum_{j,i=1}^{k,n} \left(\text{woe}_j * w_i + \frac{a}{n}\right)\right) * \text{factor} + \text{offset} = \sum_{j,i=1}^{k,n} \left(-\left(\text{woe}_j * w_i + \frac{a}{n}\right) * \text{factor} + \right. \\ &\quad \left. \text{offset}/n\right) \quad (15) \end{aligned}$$

Con la ecuación (15) se encuentra el puntaje de cada atributo. La suma de los puntajes de cada atributo del crédito da como resultado el puntaje total otorgado al crédito.

4.1.8 Validación del Tablero de Puntaje

El tablero de puntaje debe demostrar que efectivamente separa de una manera adecuada a los créditos clasificados como malos de aquellos clasificados como buenos, otorgando un menor puntaje a los primeros. Para esto, se han desarrollado una serie de pruebas estadísticas, siendo las comúnmente utilizadas para llevar a cabo la validación de los modelos de puntaje (Thomas, Edelman, & Crook, 2002):

- Índice de Ginni,
- Índice de Divergencia,
- K-S (Kolmogorov – Smirnov),
- Valor de la Información,
- Índice de Estabilidad de la Población,
- Prueba T.

4.1.8.1 Índice de Ginni

El coeficiente de Ginni mide la eficacia que tiene el modelo de puntaje de separar los créditos buenos de los malos. El máximo valor que puede alcanzar el coeficiente de Ginni es de 0.5. El área contenida entre la línea de la identidad y la curva es denominada coeficiente de Ginni IG y es estimado con la siguiente fórmula:

$$IG = \frac{1}{2} - \frac{1}{2} \sum_{i=1}^n \frac{b_i - b_{i-1}}{b_i} (g_i - g_{i-1}) + \sum_{i=1}^n \frac{b_i - b_{i-1}}{b_i} g_{i-1} \quad (16)$$

4.1.8.2 Índice de divergencia

Su objetivo es medir la eficacia de un modelo de puntaje a través de estimar la separación entre la media de las distribuciones de los créditos de buen \bar{k} y mal \bar{h} comportamiento. Su fórmula es la siguiente:

$$ID = 2 \frac{(\bar{k} - \bar{h})^2}{Var_k + Var_h} \quad (17)$$

4.1.8.3 K-S (Kolmogorov – Smirnov)

Otra prueba que se utiliza para medir la eficacia del modelo de separar las cuentas buenas de las malas es la denominada K-S (Kolmogorov-Smirnov), la cual mide la distancia máxima que existe entre la distribución acumulada de los créditos con comportamiento malo b_i y bueno g_i .

$$K - S = \max(b_i - g_i) \quad \forall i$$

4.1.8.4 Valor de la Información

La prueba del Valor de la Información (VI) es utilizada para determinar si las variables y el modelo sirven para predecir el comportamiento de los acreditados. Valores cercanos o superiores a 0.3 indican que el modelo contiene una gran cantidad de información.

$$VI = \sum_{i=1}^n \ln \left(\frac{f(x_i)_g}{f(x_i)_b} \right) (f(x_i)_g - f(x_i)_b) \quad (18)$$

4.1.8.5 Índice de la Estabilidad de la Población (IEP)

Este índice es utilizado para determinar la estabilidad de un modelo, variable o población en el tiempo, así como para establecer si la muestra de desarrollo de un modelo es representativa de la población total. El IEP se calcula con la siguiente fórmula:

$$IEP = \sum_{i=1}^n (O_i - E_i) * \ln \frac{O_i}{E_i} \quad (19)$$

4.1.8.6 Prueba T

El método de mínimos cuadrados se emplea cuando la relación entre x , en este caso el puntaje, y la media de y es lineal o bastante cercano a una línea recta, de tal manera que se produzcan predicciones a través de $y_i = \alpha + \beta x_i + \varepsilon_i$.

Los estimadores de α y β están dados respectivamente por

$$a = \bar{y} - b \bar{x}$$

$$b = S_{xy} / S_{xx}$$

donde

$$s_e^2 = \frac{S_{xx}S_{yy} - (S_{xy})^2}{n(n-2)S_{xx}}$$

$$S_{xx} = n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2$$

$$S_{yy} = n \sum_{i=1}^n y_i^2 - \left(\sum_{i=1}^n y_i \right)^2$$

$$S_{xy} = n \sum_{i=1}^n x_i y_i - \frac{\sum_{i=1}^n x_i}{n} \frac{\sum_{i=1}^n y_i}{n}$$

Siendo \bar{x}, \bar{y} la media de x y y respectivamente, n el número de observaciones y s_e el error estándar de la estimación.

4.2 Modelo de Riesgo de Crédito

El riesgo de crédito lo definen las “Disposiciones de carácter general aplicables a las instituciones de crédito” emitidas por la Comisión Nacional Bancaria y de Valores (CNBV) como la pérdida potencial por la falta de pago de un acreditado o contraparte. El riesgo de crédito deriva de la posibilidad de que el acreditado pueda incumplir. (Hull, 2007) En particular para el presente trabajo se definirá el riesgo de crédito como la posibilidad de que FEGA pague una garantía al intermediario financiero por el incumplimiento de las obligaciones de un acreditado.

4.2.1 Pérdida Esperada

Como se trata de una pérdida esperada, las instituciones financieras deben tomar como dado la disminución de sus utilidades conforme a dicha pérdida. A través de la generación de reservas para riesgos crediticios se reconoce dicha pérdida. Conforme se detalla más adelante, esta pérdida es la base para el cobro de la prima de seguro. La pérdida no esperada, la cual es la diferencia entre el riesgo de crédito y la pérdida esperada, es cubierta con el capital de la Institución y también debe ser considerada como componente de la prima de seguro.

4.2.2 Modelo CyRCE

El resultado de otorgar un crédito se puede clasificar en dos escenarios: bueno (cumplió) o malo (incumplió). El resultado de que el crédito i se le clasifique como malo se encuentra en función de una probabilidad de incumplimiento p_i para un horizonte de tiempo ht . Dicha probabilidad esta asociada a las características del acreditado y se estima a partir de los modelos de puntaje anteriormente descritos. Por simplicidad, se supondrá independencia entre acreditados, es decir, el

que un acreditado incumpla no implica que otro también lo hará. Una vez descrito lo anterior, se procede a encontrar la distribución de pérdida de la cartera donde cada crédito tiene un saldo de incumplimiento S_i y una probabilidad de incumplimiento p_i . La distribución se puede estimar por medio de una distribución de probabilidad normal, la cual requiere obtener una probabilidad de incumplimiento común p_π . Dicha probabilidad se estima a través de

$$p_\pi = \frac{\sum_{i=1}^n p_i S_i}{S} \quad (20)$$

$$S = \sum_{i=1}^n S_i \quad (21)$$

Dado que cada crédito presenta un comportamiento binomial, es decir, pago o no pago, el riesgo de mercado se estima a través de

$$\text{VaR}_\alpha = p_\pi S + z_\alpha \sqrt{V} = p_\pi S + z_\alpha \sqrt{p_\pi(1-p_\pi) \sum_{i=1}^n S_i^2} \quad (22)$$

4.2.3 Modelo CREDIT RISK +

A diferencia del modelo CyRCE, el modelo Credit Risk + modela dos procesos aleatorios que presenta cualquier cartera de crédito: el proceso de número de incumplimiento y el proceso del monto de las pérdidas. La distribución de Poisson modela eventos discretos en un espacio o tiempo continuo. Es decir, los eventos de incumplimiento (discreto) en un dado lapso de tiempo ht . En la cartera se desea modelar el número de incumplimiento de los cuales es imposible pronosticar el número exacto ni el momento preciso. Sin embargo, es posible asociar una probabilidad al número de incumplimientos en un dado lapso de tiempo.

Sea $p_n = Pr[N = n]$ la probabilidad de que se presenten exactamente n incumplimientos. Utilizando la distribución de Poisson, p_n se encuentra dado por (Gutiérrez & Elizondo, 2002)

$$p_n = \frac{e^{-\mu} \mu^n}{n!} \quad (23)$$

Donde μ representa el número esperado de incumplimientos de la cartera en el periodo.

4.2.4 Modelo Montecarlo

Se tiene una cartera compuesta por n acreditados, cada uno con un saldo s_i , independientes entre ellos con una probabilidad de incumplimiento p_i para un horizonte de tiempo ht . Dado que el incumplimiento es una variable aleatoria, el resultado de cada crédito se puede modelar a través de una función indicadora I_i que toma los siguientes valores (Márquez, 2009):

$$I_i = \begin{cases} 1 & \text{si } u_i \leq p_i \\ 0 & \text{eoc} \end{cases} \quad (24)$$

Donde I_i toma el valor 1 si el acreditado i incumple y 0 en otro caso (*eoc*), y u_i es una variable aleatoria distribuida de manera uniforme en el rango $0 - 1$ $U \sim [0,1]$ para $i = 1, 2, \dots, n$.

Dado lo anterior, la pérdida estimada LGD_w del portafolio para el escenario w viene dada por

$$LGD_w = \sum_{i=1}^n I_i * s_i \quad (25)$$

Se simulan W escenarios que permiten recrear la distribución de pérdidas LGD

$$LGD = \begin{matrix} \hat{e} & LGD_1 & \hat{u} \\ \hat{e} & LGD_2 & \hat{u} \\ \hat{e} & \cdot & \hat{u} \\ \hat{e} & \cdot & \hat{u} \\ \hat{e} & LGD_w & \hat{u} \end{matrix} \quad (26)$$

Los W escenarios son ordenados de la mayor a la menor pérdida, siendo el VaR a un α nivel de confianza la observación Y que cumple con (Hull, 2007):

$$Y = W * (1 - \alpha) \quad (27)$$

4.2.5 Prueba Retrospectiva (*backtesting*)

El objetivo de la prueba retrospectiva (*backtesting*) es determinar la capacidad predictiva de los modelos de riesgo. Lo anterior se hace contabilizando en un cierto horizonte de tiempo el número de veces que las pérdidas por riesgo de crédito exceden el VaR (*rebases*). La metodología sugerida es la desarrollada por Kupiec, la cual establece el número de rebases máximos y mínimos conforme a la siguiente ecuación:

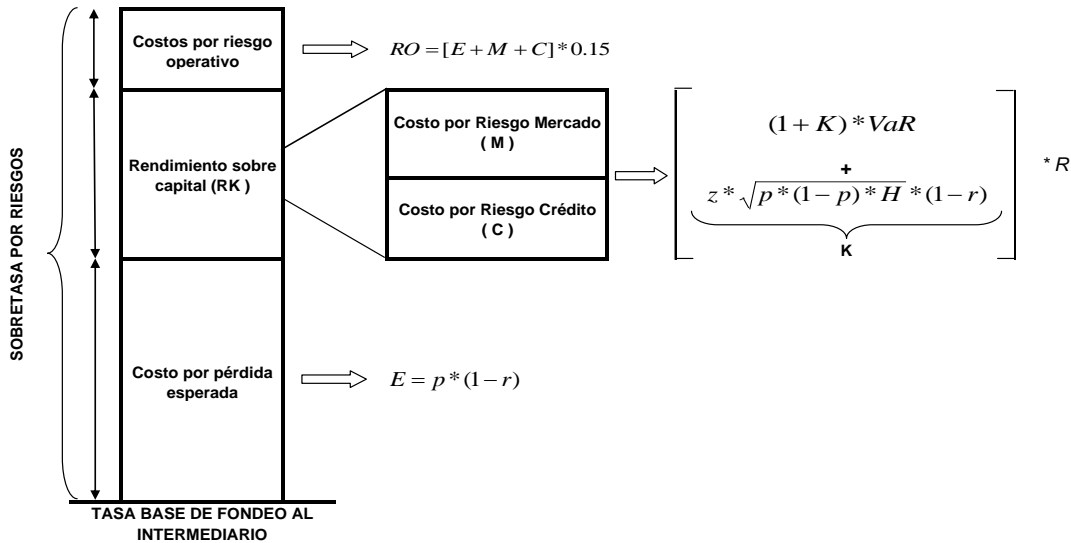
$$\underline{-2 \ln \left[(1-p)^{m-e} p^e \right] + 2 \ln \left[\left(1 - \frac{e}{m} \right)^{m-e} \left(\frac{e}{m} \right)^e \right]} \quad (28)$$

4.3 Modelo para la Estimación de Primas de Seguro

El contrato de seguro es el acuerdo por el cual una de las partes, el asegurador –FEGA-, se obliga de resarcir un daño a través de una suma asegurada de dinero a la otra parte, tomador – Intermediario Financiero-, a cambio del pago de un precio denominado prima.

En la figura 1 se esquematiza el modelo general de estimación de prima por riesgos. El concepto general se encuentra construido sobre: i) La base de una tasa de fondeo a los acreditados, ii) Un primer componente que integra el costo por pérdidas esperadas de riesgo crédito, iii) Un segundo componente integrado por el rendimiento sobre el capital que soporta la operación, iv) Un componente final que incluye la consideración estándar sobre los posibles costos operativos de la operación.

figura 1. Componentes de la sobretasa por riesgos. (También denominada prima de riesgos).



Fuente: Modelo para sobretasa de riesgo desarrollado por FIRA

5. Metodología

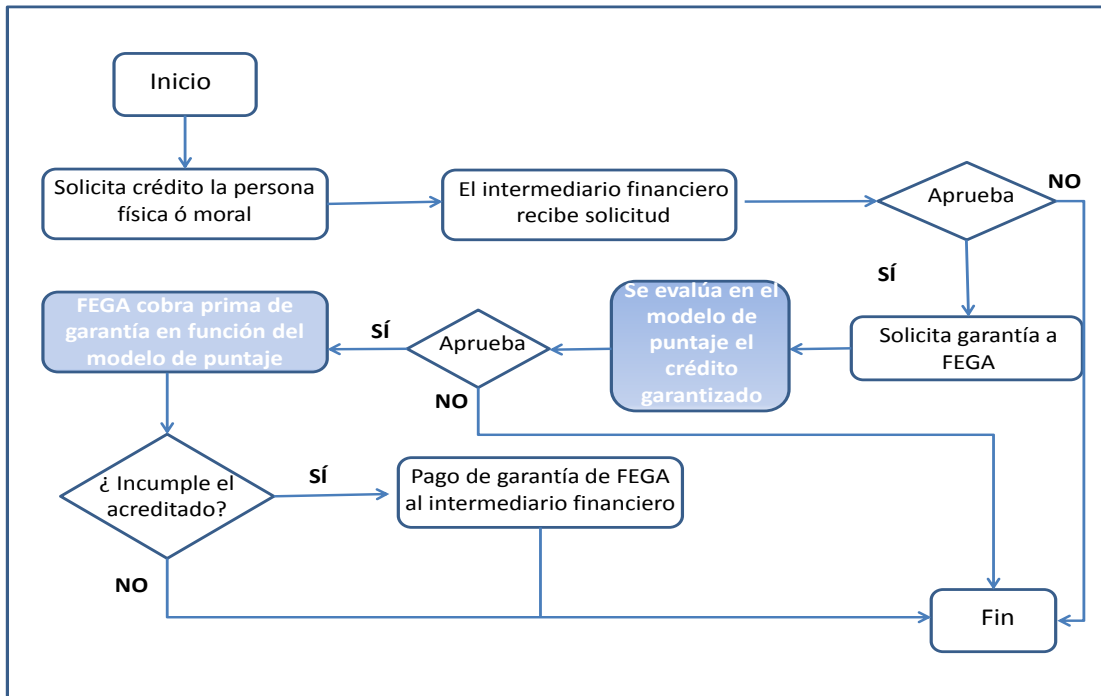
La presente investigación tiene un diseño descriptivo correlacional, para determinar el modelo más adecuado en el otorgamiento de créditos en el sector agropecuario que permita ser la base para la gestión del riesgo de crédito. Para realizar el presente estudio, se tomó una muestra de 110,278 créditos garantizados de personas físicas por FEGA entre enero 2006 y mayo 2011. De cada crédito se cuenta con 69 variables ligadas al perfil del acreditado. La misma muestra sirvió para realizar las pruebas retrospectivas de la estimación del valor en riesgo del portafolio garantizado.

El sustento teórico se ubica, en su primer parte, en los modelos de discriminación haciendo un análisis de los diferentes modelos de puntaje, entre los más importantes se encuentran el modelo de regresión lineal y Logit. Después de hacer un comparativo entre los modelos de discriminación, se aplicaron los resultados a los modelos de estimación del riesgo de crédito CyRCE, Credit Risk + y Montecarlo. Posteriormente, se aplicaron los resultados del modelo de discriminación a un modelo paramétrico para estimar las primas de riesgo en el sector agropecuario. En términos generales la presente investigación desea dar a conocer la relación entre las variables independientes (modelo de discriminación, riesgo de crédito, primas de seguro y riesgo beneficio) con la variable dependiente (conservación del patrimonio e incremento de los beneficios).

6. Análisis E Interpretación De La Información

En la siguiente figura se demuestra esquemáticamente el modelo que propone generar un sistema de administración de riesgo de crédito que solucione los problemas detectados en el procedimiento de pago de garantía que tenga como base un modelo de puntaje para el otorgamiento de créditos.

Figura 2. Procedimiento para el pago de garantía considerando un modelo de puntaje



Fuente: Elaboración propia

6.1 Generación de la muestra de Desarrollo

La presente investigación, se respaldó con una base de datos inicial de 110,278 créditos garantizados, de los cuales se eliminaron 43 por falta de información, quedando 110,235 créditos para análisis. Los créditos garantizados presentan 69 distintas características, aplicándose a cada una la prueba del valor de la información, conforme a lo descrito supra para determinar su grado de utilidad para el desarrollo del modelo de puntaje. El periodo de garantía de los créditos es del 2006 al 2011 conforme se muestra en la siguiente tabla.

Tabla 1. Distribución de crédito por año de apertura

Fecha de apertura de crédito	
Año	Número de créditos
2006	5,104
2007	5,504
2008	13,975
2009	28,794
2010	41,612
2011	15,246
Total	110,235

Fuente: Elaboración propia

El objetivo de un modelo de puntaje es descubrir las características de aquellos acreditados que la institución financiera no hubiera deseado otorgarles un crédito ya que está los clasificaría como malos. Con base en lo anterior, se vuelve menester definir la clasificación de los créditos:

- Malo: se define como un crédito malo aquel que presentó pago de garantía y a los 90 días no se recuperó el total.
- Bueno: es aquel crédito al que no se le otorgó pago de garantía.
- Indeterminado: son aquellos tuvieron pago de garantía pero a los 90 días ya se había recuperado el monto total.

Tabla 2. Clasificación de créditos por muestra total

Muestra total		
Clasificación	Número de créditos	Porcentaje
Bueno	101,751	92.30%
Malo	666	0.60%
Indeterminados	7,818	7.09%
Total	110,235	100.00%

Fuente: Elaboración propia.

Para desarrollar la tabla de puntaje, se sugiere (Siddiqi, 2006) que la muestra inicial se divida de la siguiente forma:

- 80% del total de los créditos utilizados como muestra de desarrollo.
- 20% del total de los créditos utilizados como muestra de prueba.

Para el desarrollo del tablero de puntaje (*scorecard*) se eliminaron los créditos indeterminados, quedando únicamente créditos definidos como buenos o malos, con dichos cambios la muestra total se redujo a 102,417 créditos, esto con la finalidad de darle mayor potencia al modelo. De éstos créditos se eliminaron 2 por presentar información incompleta, con lo cual la muestra final quedó en 102,415 créditos. La muestra de desarrollo es una selección aleatoria a partir de la base de datos total, quedando de la siguiente forma:

Tabla 3. Clasificación de créditos por muestra de desarrollo

MUESTRA DE DESARROLLO		
Clasificación	Número de créditos	Porcentaje
Bueno	80,651	99.31%
Malo	562	0.69%
TOTAL	81,213	100.00%

Fuente: Elaboración propia

La muestra de prueba quedó como se aprecia en la tabla siguiente:

Tabla 4. Clasificación de créditos por muestra de prueba

MUESTRA DE PRUEBA		
Clasificación	Número de créditos	Porcentaje
Bueno	21,098	99.51%
Malo	104	0.49%
TOTAL	21,202	100.00%

Fuente: Elaboración propia

La muestra de prueba guarda consistencia con la de desarrolla en el sentido de mantener la proporción de buenos y de malos.

6.2 Aplicación y contrastación de los modelos

En la siguiente Tabla 5 se muestran los resultados de aplicar varios modelos tendientes a comprobar la hipótesis originalmente planteada. Esta información nos permitirá realizar algunas conclusiones.

Tabla 5. Resultado de las variables

Variable independiente	Definición Real	Dimensión	Indicadores	Regresión Lineal	Regresión Logística
Modelo de discriminación	Modelo de combinación de variables que mejor separa dos grupos en función de sus características.	Modelo de regresión lineal	Índice de Ginni	0.37	NA
			Prueba K-S	60%	NA
			Índice de divergencia	2.91	NA
			Valor de la información	2.68	NA
		Modelo logit	Índice de Ginni	NA	0.40
			Prueba K-S	NA	62%
			Índice de divergencia	NA	2.73
			Valor de la información	NA	2.49
Estimación del riesgo de crédito	El riesgo de crédito es la pérdida potencial del producto del incumplimiento de la contraparte en una operación	Modelo CyRCE	Valor en riesgo	1.50%	1.65%
			Índice de Kupiec	1	0
		Credit Risk +	Valor en riesgo	1.35%	1.49%
			Índice de Kupiec	1	1
		Modelo Montecarlo	Valor en riesgo	1.56%	1.69%
			Índice de Kupiec	1	0

	que incluye un compromiso de pago.				
Estimación de prima de seguro	Cantidad de dinero que abona el asegurado a la compañía de seguros para poder contar con la cobertura de seguro contratada.	Modelo paramétrico	Primas asociadas al crédito	0.61%	0.67%
			Relación cobro estimado y valor en riesgo	0.95	1.05
			Prueba retrospectiva	1	0
Relación riesgo beneficio	Valorización de evaluación que relaciona las utilidades en el capital invertido o el valor de la producción con los recursos empleados y el beneficio generado	Riesgo beneficio de FEGA	Montos de pago de garantía	1.56%	
			Relación riesgo beneficio	0.95	1.05

Fuente: Elaboración propia sobre la base de los objetivos de la investigación

Los modelos de puntaje, tanto el de regresión lineal como el de regresión logística, presenta indicadores muy parecidos. Lo anterior es consistente con lo descrito supra, (resultados similares en el centro y con diferencia en los extremos); sin embargo, las probabilidades de incumplimiento estimadas con el modelo de regresión logística en los extremos (mayor y menor puntaje) son ligeramente mayores. Esto se refleja en los modelos de estimación de riesgo de crédito. Bajo los tres diferentes enfoques, -CyRCE, Credit Risk + y Montecarlo- el riesgo de crédito con base en el modelo de regresión logística es superior al modelo de regresión lineal.

En el caso del riesgo de crédito estimado con la regresión logística, dicho riesgo siempre es superior a las pérdidas observadas por el otorgamiento de garantías; sin embargo, con el de regresión lineal el riesgo de crédito estimado es menor a la pérdida observada en los modelos CyRCE y Credit Risk +, lo cual se indica con un 1 en el indicador de intervalo de Kupiec en la tabla anterior.

La prima de seguro por riesgos estimada con los parámetros de la regresión logística es superior a la estimada con la regresión lineal. En ambos casos la prima es cercana al equilibrio, pero solo bajo el

modelo de regresión logística y modelo Montecarlo se tiene una prima por riesgo que cubre las pérdidas efectivamente observadas.

El modelo de regresión logística es, para la población objetivo en estudio, el modelo más adecuado para el otorgamiento de créditos en el sector agropecuario que permite una adecuada administración del riesgo de crédito mediante utilizar el modelo Montecarlo y tener una prima de riesgo de equilibrio conforme a los modelos anteriormente propuestos.

7. Conclusiones

La falta del conocimiento de la probabilidad de incumplimiento de un acreditado no permite generar políticas de administración de riesgo de crédito que permitan conservar el patrimonio de una institución así como incrementar sus beneficios, debido principalmente a que: i) No se evalúa la probabilidad de incumplimiento del acreditado con base en sus características, ii) No se cuenta con probabilidades de incumplimiento para la estimación del riesgo de crédito (pérdida esperada y no esperada), iii) La prima de garantía se cobra en función del intermediario financiero; sin embargo, el evento que detona el pago es el incumplimiento del acreditado.

Con este antecedente, se ha propuesto generar un sistema de administración de riesgo de crédito que solucione los problemas detectados en el procedimiento de pago de garantía que tenga como base un modelo de puntaje para el otorgamiento de créditos, cuyos resultados sirvan para estimar el riesgo de crédito del portafolio de garantías y las primas de riesgo de equilibrio que deben cobrarse en el sector agropecuario. Bajo estas circunstancias, la combinación del modelo de regresión logística, modelo Montecarlo y el modelo para la estimación de primas de riesgo permiten cumplir las expectativas planteadas en las hipótesis general y específicas.

Referencias

- Greene, W. (2011). *Econometric Analysis*. Estados Unidos de América.
- Gujarati, D. (2006). *Principios de econometría*. España: McGraw Hill.
- Gujarati, D., & Porter, D. (2010). *Econometría*. México: McGraw Hill.
- Gutiérrez, J., & Elizondo, A. (2002). *Riesgo de Crédito: el enfoque actuarial*. Ciudad de México.
- Hull, J. C. (2007). *Risk management and financial institutions*. United States of America: Pearson Prentice Hall.
- Márquez, J. (2009). *Una nueva visión del riesgo de crédito*. Distrito Federal, México: Limusa.
- Novales, A. (1993). *Econometría*. Madrid: Mc Graw Hill.

- Sandica, A. M. (2010). Credit Scoring Modelling: A micro-macro approach. *Academy of economic studies. Doctoral school of finance-dofin* , 110.
- Siddiqi, N. (2006). *Credit Risk Scorecards: Developing and implementing intelligent Credit Scoring*. New Jersey: John Wiley & Sons, Inc.
- Thomas, L., Edelman, D., & Crook, J. (2002). *Credit Scoring and its applications*. United States of America: SIAM.

8. Anexos

8.1 Anexo 1.

La definición de las variables utilizadas en los modelos de puntaje son:

- Monto de crédito: Valor monetario original de otorgamiento del crédito.
- Porcentaje de cobertura efectiva. Porcentaje del crédito efectivamente cubierto por FEGA, una vez descontada la garantía líquida.
- Cadena productiva: conjunto de actividades que tienen por objeto la producción, transformación, distribución y consumo de bienes y servicios generados para satisfacer las necesidades sociales y de intercambio de una región o País.
- Número de amortizaciones: número de pagos de capital.
- Municipio: Región en la que se desarrollará la actividad para la cual fue otorgada el crédito.
- Tipo de garantía: En función del programa que otorga la garantía.
- Esquema de cobro: Periodicidad de cobro de los intereses.
- Tipo de cobertura: Variable dicotómica referida a la tasa de interés que se le cobro al acreditado, siendo las posibilidades fija y variable.
- Garantía sin fondeo: Créditos garantizados cuyo origen de los recursos no provienen de FIRA.
- Edad: edad del acreditado a la fecha de apertura del crédito.