

**APLICACIÓN DE UN MODELO DE RATING PARA MEDIR EL RIESGO DE
CRÉDITO EN LA CARTERA EMPRESARIAL DE UNA ENTIDAD
COOPERATIVA COLOMBIANA**

CAROLINA GÓMEZ MUÑOZ

**UNIVERSIDAD DEL VALLE
FACULTAD DE CIENCIAS SOCIALES Y ECONOMICAS
DEPARTAMENTO DE ECONOMIA
CALI
2009**

**APLICACIÓN DE UN MODELO DE RATING PARA MEDIR EL RIESGO DE
CRÉDITO EN LA CARTERA EMPRESARIAL DE UNA ENTIDAD
COOPERATIVA COLOMBIANA**

CAROLINA GÓMEZ MUÑOZ
Informe de Pasantía para optar el título de economista.

Director
Ramiro Hernán Montaña Sanchez
Economista Especialista en Finanzas

UNIVERSIDAD DEL VALLE
FACULTAD DE CIENCIAS SOCIALES Y ECONOMICAS
DEPARTAMENTO DE ECONOMIA
CALI
2009

**A mi bebé que viene en
Camino, a mi familia y a
Mis compañeros de
labores.**

AGRADECIMIENTOS

La autora expresa sus agradecimientos a:

Dios, por poner en mi camino las personas, los recursos y la fortaleza para culminar con éxito esta etapa tan importante en mi vida.

Mi hermana Ana María quien siempre me alentó a seguir adelante a pesar de las contrariedades.

Rodolfo mi amigo, quien me acompañó desde la distancia y me dio ánimo para continuar.

Mis profesores Juan Byron Correa e Inés María Ulloa por su apoyo en cada uno de los aspectos que dieron origen al presente documento.

Mi amiga Martha Patricia Fajardo y su familia quienes me animaron a continuar.

Mi tutor Ramiro Montaña por elegirme para el empleo que dio origen a este documento y por acompañarme en el proceso con su valiosa retroalimentación.

Mis compañeros de la Gerencia Nacional de Riesgo por su colaboración, paciencia, ánimo y acompañamiento.

Agradecimientos muy especiales para mi compañera y amiga Claudia Johanna Polo Yepes, por su invaluable colaboración y guía. De no ser por ti hoy no estaría escribiendo estas líneas.

TABLA DE CONTENIDO

	Pág.
INTRODUCCION	12
1. PLANTEAMIENTO DEL PROBLEMA.....	13
2. JUSTIFICACION.....	14
3. MARCO TEÓRICO	16
4. ESTADO DEL ARTE.....	18
4.1. METODOLOGÍAS ESTADÍSTICAS	21
4.1.1 Método de Matrices de Transición.....	22
4.1.2 Método de árboles de Decisión.	24
4.1.3 Modelo de Regresión logística.....	24
4.1.4 Método de Análisis discriminante.	26
4.1.4.1. Esquemas de calificación basados en la ponderación de factores de riesgo: El modelo Z-Score de Altman.....	26
4.1.4.2 El modelo Z (1977).....	28
4.2. METODOLOGÍAS DE INFORMACIÓN DE MERCADO	28
5. OBJETIVO GENERAL	29
5.1. OBJETIVOS ESPECIFICOS	29
6. METODOLOGÍA	29

6.1 ENTENDIMIENTO DEL NEGOCIO Y DE LOS DATOS DISPONIBLES	31
6.2 PREPARACION DE DATOS	32
6.2.1 La Base De Datos.....	32
6.2.2 Definición De Variables.....	33
6.2.3 Diagnóstico de Variables Explicativas	37
6.3 MODELOS	39
6.3.1 Selección de Variables	39
6.3.2 Selección de la Muestra.....	40
6.3.3 Metodología Estadística.....	40
6.3.4 Modelo Final	41
6.4 EVALUACIÓN DEL MODELO.....	42
6.4.1 Pruebas de Bondad de Ajuste y Significancia conjunta de variables.....	43
6.4.2 Pruebas de discriminación.....	44
6.4.3 Pruebas de clasificación	45
6.4.4 Pruebas sobre una muestra de validación	46
6.5 DESPLIEGUE DEL MODELO.....	46
7. CONCLUSIONES	46
REFERENCIAS BIBLIOGRÁFICAS.....	48
ANEXO 1: CORRELACIONES ENTRE VARIABLES INDEPENDIENTES	52

LISTA DE TABLAS

	Pág.
Tabla 1: Evolución Histórica de Modelos Predictivos del Fracaso Empresarial.	20
Tabla 2: Tasa de Incumplimiento por Calificación para Bonos Corporativos 1971-1994.....	23
Tabla 3: Matriz de Transición con un Horizonte de Tiempo de un Año Calculada con Bonos Emitidos entre 1971 y 1989 y calificados por S&P	24
Tabla 4: Efecto del Crecimiento de las variables Frente al Incumplimiento.	37
Tabla 5: Valores Válidos de Variables Independientes (<i>Ratios Financieros</i>).....	38
Tabla 6: Distribución del Incumplimiento en la Cartera por Regionales.....	38
Tabla 7: Prueba de Discriminación Kolmogorov-Smirnov	40
Tabla 8: Parámetros estimados para el modelo de regresión logística elegido	41
Tabla 9: Prueba de bondad de ajuste de Hosmer Lemeshow	43
Tabla 10: Prueba de significancia conjunta de las variables.....	43
Tabla 11: Tabla de Clasificación del Modelo.	45

LISTA DE GRÁFICAS

	Pág.
Gráfica 1: Distribución de Probabilidad de las Pérdidas Ocasionadas por Exposición al Riesgo Crediticio.....	17
Gráfica 2: Clasificación de Créditos por la Técnica “Z-Score” de Altman	27
Gráfica 3: Comportamiento cartera empresarial	31
Gráfica 4: Comportamiento Índice de calidad Cartera empresarial.....	32
Gráfica 5: Área bajo la curva ROC.....	44

GLOSARIO

ACTIVO: término contable – financiero con el que se denomina al conjunto de recursos económicos con los que cuenta una persona, sociedad, corporación, entidad, empresa o cualquier organización económica. Se clasifican según su liquidez de mayor a menor en activos corrientes y no corrientes.

AJUSTES POR INFLACION: procedimiento utilizado para reflejar en la contabilidad los efectos en la pérdida del poder adquisitivo de la unidad monetaria y los efectos de la inflación, así como la metodología que se adopta para expresar en moneda del mismo poder adquisitivo la información económica, financiera social de los entes económicos.

AMORTIZACIONES: la eliminación gradual de deuda, a través de pagos periódicos. También se refiere a la eliminación de deuda periódica y al paso de gastos de activos clasificados como intangibles, como por ejemplo las patentes.

CARTERA: designación genérica que comprende los valores o efectos comerciales y documentos a cargo de clientes que forman parte del activo corriente de una empresa comercial, de un banco o una entidad en general.

COSTO DE VENTAS: es el costo en el que se incurre para comercializar un bien o para prestar un servicio. Es el valor en que se ha incurrido para producir o comprar un bien que se vende.

DEPRECIACION: reconocimiento racional y sistemático del costo de los bienes, distribuido durante su vida útil estimada, con el fin de obtener los recursos necesarios para la reposición de los bienes, de manera que se conserve la capacidad operativa o productiva del ente público.

FLUJO DE CAJA: es un cálculo del volumen de ingresos y gastos que ocurrirán en una empresa durante un determinado período. Ello posibilita saber si sobra o falta dinero en un momento dado.

GASTOS FINANCIEROS: corresponden a los gastos que originan los intereses de las deudas financieras en el corto y largo plazo.

LIQUIDEZ: En economía, la liquidez representa la cualidad de los activos para ser convertidos en dinero efectivo de forma inmediata sin pérdida significativa de su valor.

OBLIGACIONES FINANCIERAS: las obligaciones financieras representan el valor de las obligaciones contraídas por las empresas mediante la obtención de recursos provenientes de establecimientos de crédito o de otras instituciones financieras u otros entes distintos de los anteriores, del país o del exterior

PASIVO: Agrupa el conjunto de las cuentas que representan las obligaciones contraídas por el ente económico en desarrollo del giro ordinario de su actividad, pagaderas en dinero, bienes o en servicios

PATRIMONIO: el conjunto de relaciones jurídicas pertenecientes a una persona, que tienen una utilidad económica y por ello son susceptibles de estimación monetaria, y cuya relaciones jurídicas están constituidas por deberes y derechos (activos y pasivos).

POLITICAS DE OTORGAMIENTO: son los lineamientos diseñados por las entidades financieras para administrar la colocación de recursos financieros en el mercado crediticio.

UTILIDAD NETA: Se entiende por utilidad neta, la utilidad resultante después de restar y sumar de la utilidad operacional, los gastos e Ingresos no operacionales respectivamente, los impuestos y la Reserva legal. Es la utilidad que efectivamente se distribuye a los socios

UTILIDAD OPERACIONAL: Es el resultado que se obtiene de restar a las ventas netas los costos de ventas y los gastos de administración y ventas.

RESUMEN

El presente documento es el resultado de la labor realizada por la autora durante el período de pasantía académica en una empresa del sector cooperativo financiero de Colombia, donde se identificó la necesidad de atender a los requerimientos legales de la Superintendencia Financiera de Colombia con respecto a la correcta medición y administración del riesgo crediticio. Este trabajo desarrolla una metodología y un modelo que permite realizar un primer acercamiento al cálculo apropiado de la probabilidad de incumplimiento en el segmento de banca empresarial de la entidad a partir de procedimientos estadísticos como actualmente lo exige la circular externa 022 de 2008 de la Superintendencia Financiera de Colombia.

PALABRAS CLAVES:

- Riesgo de Crédito
- Administración del Riesgo
- Modelos de Rating
- Probabilidad de Incumplimiento
- SARC
- Crédito Empresarial

APLICACIÓN DE UN MODELO DE RATING PARA MEDIR EL RIESGO DE CRÉDITO EN LA CARTERA EMPRESARIAL DE UNA ENTIDAD COOPERATIVA COLOMBIANA

INTRODUCCION

Tradicionalmente, desde que se tiene una crónica de la manera de hacer banca a partir de la edad media y hasta principios del siglo XX, el negocio bancario se mantuvo con pocos cambios. Antiguamente existía el mito del *“olfato de banquero”* que les permitía a estos personajes, conocer perfectamente a sus deudores potenciales, clientes y competidores, de manera que por su pericia sabían a quién le podían prestar y a quien no. Adicionalmente, los banqueros tenían un *“ojo clínico”* para identificar aquellos clientes que implicaban mucho riesgo, pero cuya recompensa era significativamente mas generosa. Así, de los sujetos de crédito que representaban un nivel de riesgo interesante con rendimientos potenciales extraordinarios, los banqueros sabían perfectamente cuánto les podían prestar, cuánto les podían cobrar, qué garantías debían exigir y cómo les iban a pagar. Sin embargo, en la última mitad del siglo XX, y particularmente en las últimas dos décadas, este olfato se ha desvanecido y se ha presentado un cambio radical en la forma de percibir, medir y administrar el riesgo de crédito.

Básicamente, el cambio de visión obedece a un aumento significativo de la complejidad del negocio del crédito. Antiguamente, los banqueros se rodeaban de poderes importantes que les permitían operar con información privilegiada, lo cual representó una fuente importante de negocios y riqueza. De hecho, los banqueros de antaño atendían pequeños grupos de personas poderosas que concentraban su cartera y los mantenían muy bien informados. Hoy en día, aunque los banqueros continúan preocupados por la política, rodeados de las altas esferas del poder y los proyectos grandes son de especial interés, la mayor parte del crédito se atomizó y se pierde en el anonimato de una gran cantidad de deudores para los cuales es imposible hacer una evaluación *“de olfato”* sobre el riesgo que representan en lo individual y en conjunto.

Otra fuente de complejidad, radica en el incremento de instrumentos y operaciones que se registran fuera del balance y que conllevan un riesgo de contraparte; es decir, el riesgo contingente implícito en un contrato, en el que la contraparte del tenedor del título incumple con la obligación contraída (Vg. títulos valores, derivados, letras, entre otros). Debido a la naturaleza del incumplimiento

asociada a este tipo de instrumentos y contratos, se les trata como parte del riesgo de crédito que asumen los agentes que participan en este mercado.

La banca ha atravesado cambios técnicos y estructurales en la composición de su cartera de crédito, la cual presenta cada vez mayores niveles de complejidad de medición de riesgo, a lo cual la banca ha respondido empleando técnicas probabilísticas y estadísticas cada vez más sofisticadas para evaluar el riesgo que representa la mayoría de sus deudores presentes y potenciales en lo individual y su cartera de créditos en conjunto.

1. PLANTEAMIENTO DEL PROBLEMA

Los intermediarios financieros juegan un papel fundamental en el desarrollo de un país. La función de los intermediarios financieros, consiste en redireccionar los recursos de agentes superavitarios a agentes deficitarios (Gurley y Shaw: 1955). Así, dentro de una economía, todos los agentes toman créditos para favorecerse en términos de liquidez y con ello impulsan de manera directa el desarrollo. Los gobiernos, por ejemplo, se endeudan para financiar su déficit y promover la inversión en diversas áreas, las empresas toman créditos para ampliar su capacidad productiva directa o indirectamente y las personas naturales acceden al crédito cuando pagan por bienes o servicios recibidos después de haberlos utilizado, cuando usan sus tarjetas de crédito o cuando concilian con algún tipo de intermediario para financiar su vivienda, educación, viajes, etc. (López: 2005).

Cuando un agente accede de alguna manera al crédito, el intermediario queda expuesto a la posibilidad de no recuperar su dinero, esto es lo que se conoce como riesgo crediticio. La correcta medición y administración de este riesgo prevendrá la quiebra de la entidad que actúe como intermediaria y por transitividad protegerá los recursos financieros de los agentes superavitarios de la economía. Como consecuencia de lo anterior, resulta conveniente que los intermediarios financieros adopten un sistema de administración que permita controlar dicho riesgo.

En Colombia, la Superintendencia Financiera de Colombia (en adelante SFC), se encarga de dictar los parámetros fundamentales para alcanzar dicho objetivo. Las reglamentaciones de la SFC exigen a las entidades financieras la implementación de un sistema de administración de riesgos. Entre otros elementos, dicho sistema debe, de una parte, incluir la definición de políticas claras y precisas sobre los diversos procesos de administración de los riesgos y de otro, expresar un alto grado de compromiso y responsabilidad de la junta directiva de las entidades financieras en el conocimiento, seguimiento y control de los procesos correspondientes. (Superintendencia Bancaria de Colombia: 2005)

La normatividad de la SFC incluye todas y cada una de las fases del proceso de intermediación financiera. En este documento se discutirá principalmente lo relacionado con el riesgo crediticio y, específicamente, el riesgo de crédito en el segmento de personas jurídicas¹.

La entidad cooperativa con la que se trabajará, es un Establecimiento de crédito, vigilado por la Superintendencia Financiera de Colombia. Como tal, está obligada a cumplir con los parámetros que ésta imponga para garantizar el correcto funcionamiento del sistema financiero colombiano.

La administración del riesgo al interior de la entidad depende directamente de la Gerencia General y se concentra en la Gerencia Nacional de Riesgo, donde se llevan a cabo diferentes proyectos con el único fin de reducir los riesgos que surgen de la actividad económica que la empresa desarrolla. Uno de ellos es la implementación del Sistema de Administración de Riesgo Crediticio S.A.R.C., que se aplica tanto a personas jurídicas como a personas naturales.

2. JUSTIFICACION

Como consecuencia de la responsabilidad que acarrea la actividad de intermediación financiera con respecto a la protección de los recursos de los agentes superavitarios y la sostenibilidad del negocio, las entidades de crédito pueden ser caracterizadas como aversas al riesgo. Así, cuando una entidad financiera manifiesta balances desfavorables por el deterioro de su cartera, la probabilidad de bancarrota se incrementa. El resultado de esto puede ser el racionamiento del crédito, el alza en las tasas de interés y la reducción de la oferta de fondos prestables, lo que trae consigo devastadoras consecuencias para el desarrollo de los países. (Kiyoutaki y Moore: 1997)

Las causas que dan origen al riesgo de crédito son básicamente la combinación de situaciones desfavorables en el entorno y la imposibilidad de los agentes de superarlas. Dentro de dichas situaciones desfavorables pueden encontrarse por ejemplo: el debilitamiento de un sector económico en particular, la mala administración de las empresas, las crisis y depresiones económicas, entre otras.

Debido a que el riesgo de crédito es inherente a la actividad de intermediación financiera, los intermediarios y los organismos de control se esfuerzan por minimizar el impacto del mismo en el desarrollo de la actividad. Lo anterior se logra a través de la implementación de sistemas de administración de riesgo

¹ El estudio se concentrará en este segmento porque fue en el área específica de crédito empresarial donde se llevó a cabo la práctica que lo motiva.

(SARC) que se construyen y se aplican bajo criterios de evaluación objetivos². Dicho proceso debe tener en cuenta las particularidades del mercado objetivo de la entidad y a la vez debe arrojar resultados estandarizados que permitan alinear la información de todo el sector financiero.

“...La convergencia hacia los nuevos estándares definidos por tales normas³ impone diversos retos en relación con el nivel de desarrollo del mercado financiero local dentro de los que se destacan la mayor profesionalización de los administradores bancarios; el desarrollo y generalización de prácticas de administración integral de los riesgos inherentes al negocio; la identificación de los perfiles de riesgo de las entidades; la incorporación de un enfoque prospectivo en la gestión de este riesgo y en la constitución de las provisiones correspondientes; la consecuente utilización de recursos conceptuales, estadísticos y tecnológicos; la producción y reconstrucción de bases de información de calidad sobre las diversas variables relacionadas con el proceso crediticio; la producción de información pública comparable y oportuna y, por supuesto, la adecuación del proceso supervisor a las demandas de tales estándares.”

Superintendencia Bancaria de Colombia⁴ 2005: *Divulgación del Manual de Supervisión del SARC*. Pág. 2

Dado que existen muchas modalidades de crédito y diferentes tipos de usuarios del mismo, el análisis de la administración del riesgo se hace muy denso. Debido a las facilidades de acceso a la información, el presente documento se enfocará en el segmento de clientes de crédito empresarial de una cooperativa Financiera, que al cierre de Abril de 2009 cuenta con 804 empresas con obligaciones vigentes que constituyen una cartera de \$91.948 millones de pesos.

Se espera que al finalizar el proyecto, la entidad en cuestión, cuente con una herramienta útil para iniciar la construcción de un mecanismo de calificación de clientes en el proceso de otorgamiento de crédito en el segmento de personas jurídicas, que permita la construcción e implementación de modelos internos para el cálculo de pérdidas esperadas y provisiones, de conformidad con la reglamentación de la SFC.

² De acuerdo con la SFC, los criterios objetivos corresponden a aquellos adoptados institucionalmente, previa discusión de las instancias correspondientes, y que son conocidos y comprensibles para los supervisores que los aplicarán. Véase: SFC 2005: *Divulgación del Manual de Supervisión del SARC*.

³ Véase: Circular externa 011 y carta circular 031 de marzo 5 de 2002, Carta Circular 059 de mayo 6 de 2003 y Circulares Externas 014 y 030 de 2003.

⁴ Este documento fue expedido antes de que la Superbancaria y la Supervalores se fusionaran y se creara, en 2006 la Superintendencia Financiera de Colombia.

3. MARCO TEÓRICO

El mercado de crédito es sensible a los cambios en el entorno macroeconómico de los países. De la misma manera, el riesgo en el que se incurre al desarrollar actividades de intermediación en dicho mercado es también variable y depende tanto del comportamiento de la economía como del tipo de agente que solicita el servicio de crédito.

De otro lado, los resultados negativos de las entidades financieras conllevan a la racionalización del crédito. De acuerdo con Stiglitz y Weiss (1981), la tasa de interés que cobran los intermediarios, afecta el riesgo de crédito de dos diferentes maneras: por una parte, puede afectar las acciones de los deudores (riesgo moral) y por otra la dificultad para clasificar a los deudores potenciales conlleva un problema de selección adversa.

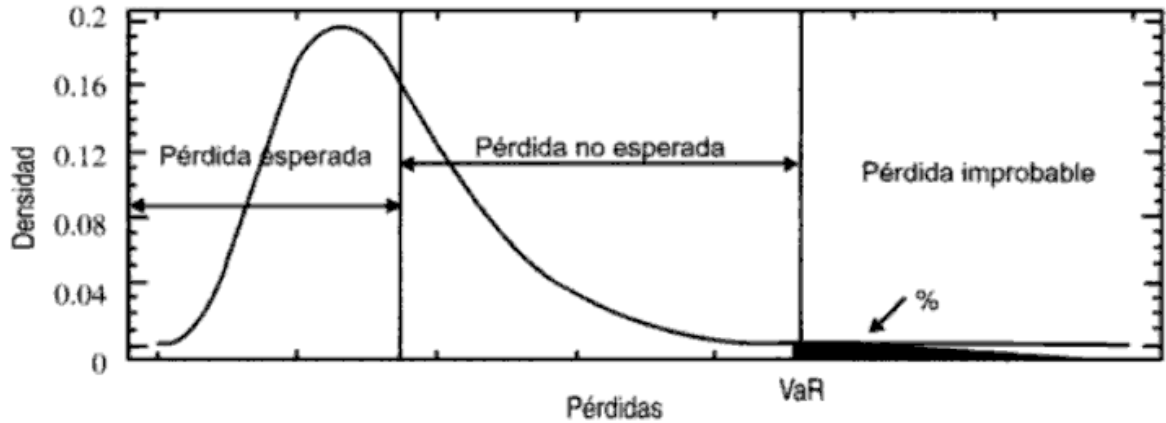
Como en el mercado de crédito se presentan restricciones de información, es decir, existe asimetría en la información, los prestamistas no tienen la certeza necesaria para diferenciar a sus clientes potenciales, es allí cuando el precio del crédito (la tasa de interés) entra a jugar como mecanismo de selección (diferenciación de precios).

Cuando se lleva a cabo el análisis de riesgo de una institución financiera, lo que se espera es encontrar el volumen de pérdidas que la institución puede llegar a tener si sus deudores incumplen con los pagos de sus obligaciones, esto lleva a un deterioro gradual de los activos de la institución. Esto es lo que se conoce como pérdida esperada (De la Fuente: 2003).

Cuando se intenta medir dicha pérdida, es necesario contar con un límite a partir del cual los créditos se consideran como incumplidos. Este límite es también conocido como default y en Colombia se considera que una empresa entra en Default o incumplimiento cuando alcanza una morosidad igual o superior a 150 días. (SFC: 2008)

La pérdida esperada, entonces, es el monto de capital que un intermediario financiero podría llegar a perder como consecuencia de la exposición al riesgo crediticio, para un horizonte de tiempo dado.

Gráfica 1: Distribución de Probabilidad de las Pérdidas Ocasionadas por Exposición al Riesgo Crediticio.



Fuente: DE LA FUENTE, Maria Lourdes (2003). “Modelos de pérdida Esperada”, Capítulo 2. En: *La Medición Integral del Riesgo de Crédito*. Pág.50.

La gráfica 1 muestra la distribución de probabilidad de las pérdidas que se ocasionan por la exposición de una entidad financiera al riesgo crediticio.

La pérdida esperada se representa por la media de la distribución de las pérdidas y se puede identificar en el extremo izquierdo del gráfico. Existe además una pérdida no esperada que surge de las variaciones en el tiempo de la calidad de la cartera ocasionadas por los diferentes escenarios económicos por los que se atraviese en el horizonte temporal, estas pérdidas están representadas por la varianza de los créditos y se pueden identificar en la sección central de la gráfica.

En general, la pérdida esperada de un intermediario financiero se calcula de la siguiente manera⁵:

$$\boxed{\boxed{Pe = Pi \times \text{Saldo Expuesto} \times Pdi}} \quad (1)$$

Donde:

Pe: Pérdida esperada por exposición al riesgo crediticio.

⁵ Esta ecuación es aplicable a cualquier entidad financiera del mundo y constituye el pilar fundamental para la administración del riesgo de crédito en el país. Véase: Capítulo II, Circular Básica 100 de la SFC.

Pi: Probabilidad de Incumplimiento del cliente

PDI: Pérdida de valor de los activos dado el incumplimiento. (Puede expresarse como $(1-T)$ con T = Tasa de recuperación.)

Saldo Expuesto: Constituye el saldo insoluto de la obligación de crédito, la cual se conforma como la suma del capital, los intereses de financiación y demás cuentas por cobrar generadas por el crédito.

Como es evidente, el cálculo de las pérdidas esperadas es sencillo. La dificultad surge cuando se intenta calcular la probabilidad de incumplimiento de los clientes. Esto sucede por múltiples razones, entre las cuales se encuentran principalmente razones de entorno macroeconómico, conocimiento limitado del cliente, debilidades en la información histórica de comportamiento de cartera, entre otras.

Para sortear tal problema, la SFC ha elaborado modelos de referencia para los diferentes tipos de cartera que existen en el sistema financiero colombiano. Como su nombre lo indica, estos modelos pretenden servir de referente para el cálculo de las pérdidas esperadas y en consecuencia de las provisiones que debe realizar cada intermediario. Sin embargo, cada entidad debe procurar la construcción de un modelo que permita realizar un cálculo más preciso que sea consecuente con las particularidades de sus clientes.

Los modelos internos que las entidades financieras construyan, deben cumplir con una serie de condiciones antes de ser aplicados para realizar los cálculos de pérdidas esperadas, por ejemplo: deben ser desarrollados conforme a técnicas de medición estadísticas definidas y precisas; deben estar sustentados con la información histórica del comportamiento de la cartera de la entidad financiera, la cual debe tener por lo menos 7 años de longitud y debe permanecer actualizada; los modelos deben arrojar resultados estandarizados y comparables, etc. Finalmente, los modelos deben ser presentados a la SFC y, después de las pruebas del caso, ésta decidirá si se aprueba la metodología o si la entidad debería seguir usando el modelo de referencia.

4. ESTADO DEL ARTE

A pesar de la importancia que tiene el tema para el funcionamiento del sector financiero, es escasa la literatura acerca de éste que se encuentra en Colombia. La mayor parte de los estudios realizados presentan una caracterización de tipo macroeconómico y se enfocan en la oferta y demanda de dinero, la racionalización del crédito y el desarrollo de los países como consecuencia de la actividad financiera.

Los estudios relacionados con indicadores e índices predictivos de efectividad o eficiencia, datan de la década de los años 30, entre los cuales se puede citar el de Smith y Winakor (1935) quienes partiendo de índices financieros basados en datos contables llegaron a la conclusión de que algunos índices diferían significativamente en empresas en quiebra y no quiebra. Los estudios hasta esa época fueron considerados poco formales. En la década de los 60, se empiezan a usar modelos matemáticos y estadísticos con el objeto de predecir una posible quiebra de empresas. Entre los estudios mas destacados se encuentran: Beaver (1967) y Altman (1968).

Posteriormente, en la década de los 70, los estudios realizados por Edmister (1972), Blum (1974), Deakin (1972), Libby(1975) y Wilcox (1973) entre otros, utilizaron como técnica el Análisis Discriminante Múltiple (ADM), sin embargo, debido a sus características operacionales esta técnica fue criticada por algunos estudiosos que trataban de utilizar metodologías similares a fin de mejorar el ADM.

En las décadas de los 80 y 90, la mayoría de los autores que escribió sobre quiebra, usó otros modelos mas complejos, pero al mismo tiempo mas precisos para determinar la probabilidad de quiebra de las empresas; especialmente los modelos logit, Probit y de redes neuronales. Entre estos se encuentran: Ohlson (1980), Gentry, Newbold & Whitford (1985), Basch y Montenegro (1989), Theodossiou (1991), Narváez (1994), Platt, Platt &Gunnar (1994), Mora (1994), Boritz et al. (1995), Episcopos (1996), Almeida & Siqueira (1997), Gallizo & Serrano (1998), Golinski (1998). En ésta época surge el análisis discriminante de Altman, mediante el cual prueba la correlación entre índices individuales y selecciona aquellos que más contribuyen al valor discriminante. Este enfoque ha tenido amplia aceptación y tal como lo expone Charles Kyd, combinando el análisis discriminante múltiple (ADM), Edward Altman, un economista financiero de la Universidad de New York, se las arregló para contestar una de las preguntas que siempre le era formulada en sus clases y conferencias: ¿cómo puede usted predecir que empresas irán o no a la quiebra?, Indudablemente, desarrollando un modelo para predecir el peligro de que en el futuro una empresa cayese en quiebra.

Regionalmente, se encuentran López & Palacios (2005), quienes en su estudio presentan una aplicación de Data Envelopment Analysis en la clasificación de clientes para la banca empresarial colombiana. El modelo que ellos proponen buscaba discriminar y clasificar correctamente al mayor número de clientes en alguna de dos subpoblaciones: una compuesta por clientes financieramente estables y por ende aptos para recibir crédito, y otra compuesta por clientes que debido a deficiencias en sus condiciones financieras no se les debe prestar dinero porque la probabilidad de un incumplimiento (default) es demasiado alta.

Otras aplicaciones han sido realizadas a empresas financieras y cooperativas colombianas por Aguas y castillo (2005) y Arias (2006).
En la tabla siguiente, se muestra la evolución histórica de los modelos de predicción de quiebras empresariales:

Tabla 1: Evolución Histórica de Modelos Predictivos del Fracaso Empresarial.

Año	Etapas	Nombre del Autor
1932 1935 1942 1965	Etapas Descriptivas Aplicación del método de razones financieras	Fitzpatrick Paul Winaker Arthur y Smith Raymond Mervin Charles Horrigan James
1966 1968	Inicio de la etapa predictiva: Accrual Based Estudios de modelos univariados basados en información contable tradicional: Sistemas de costo histórico	William Beaver
1968,1977 1972,1977 1972 1974 1980 1984 1985	Desarrollo de la etapa predictiva: Accrual Based Estudios de modelos multivariados basados en información contable tradicional: Sistemas de costo histórico	Altman Edward Deakin Edward Blum Mare Ohlson James Rose Peter y Giroux Gary Taffer Richard Zavgrent Christine
1980-1988 1980 1980 1984,1985 1985 1985 1987 1988 1989	Desarrollo de etapa predictiva: Cash Based Estudios de modelos multivariados basados en Cash Flow	Dambolena Ismael y Khory; Dambolena Ismael y Shutman Joel Largay James y Stickney Clyde Casey Cornelius Casey Cornelius y Bartzack Norman Gentry James, Newbold Paul y Whitford David Kurokawua Yukiharu y Takahshi Kichinosuke Gombola Michael, Haskins Mark, Kentz Edward y Williams David Ghalon James y Vigelan Robert Azis Abdul y Lawson Gerald
1973,1975 1981 1983 1985 1987	Desarrollo de la etapa predictiva: Factorial Analysis. Estudios de Modelos Multivariados Basados en el análisis Factorial	Pinches George, Mingo Kent, Cartuthers Kent y Eubank Arthur Cheng Kung y Shinerda Thomas Gombola Michael y Ketz Edward Casey Cornelius y Bartzack Norman Gombola Michael, Haskins Mark, Kentz Edward y Williams David

Fuente: Ibarra Mares Alberto (2001). Pp.25-26

Existe una amplia gama de metodologías que pueden ser aplicadas a la construcción de los modelos de riesgo de crédito, entre ellas se cuenta la econometría, la simulación, la optimización, los modelos financieros e incluso la combinación de todas ellas. (Aguas y Castillo: 2005)

Debido a que cada entidad tiene su propia cultura financiera, los sistemas de medición y calificación internos pueden diferir significativamente entre unas y otras, tanto en su operación como en el uso que cada entidad les da. En la mayoría de los casos estos instrumentos se cuentan como patrimonio intelectual de la entidad y difícilmente se hacen públicos por razones de competitividad. La heterogeneidad en el diseño de los sistemas internos de calificación, hace difícil la comparación de los niveles de riesgo crediticio que presentan los deudores en los diferentes rubros de calificación de diferentes entidades financieras. (Márquez: 2006)

De los enfoques utilizados se identifican dos corrientes: a). los sistemas de clasificación que se basan en una ponderación cualitativa o cuantitativa de los factores que se han identificado como los determinantes en el cumplimiento de las obligaciones, y los que se basan en conceptos de tipo fundamental.

Las técnicas de tipo fundamental parten de la proyección de variables económicas y financieras en el tiempo y del desempeño de la empresa en dichas condiciones, enmarcado claramente en sus estados financieros. El objetivo es determinar las condiciones bajo las cuales un deudor no contaría con la liquidez necesaria para cumplir con los pagos de la deuda o claramente se encontraría en quiebra. De acuerdo con la cantidad y la calidad de la información con que cuente la entidad, se escogerá la metodología.

4.1. METODOLOGÍAS ESTADÍSTICAS

De acuerdo con Aguas y Castillo (2005), si la entidad cuenta con una amplia base de información histórica, el análisis estadístico se hace bastante útil, pero este tipo de metodologías son limitadas porque a pesar de que brindan información estadística importante, el aporte en materia de información económica es limitado; los parámetros no pueden asumirse como constantes en el tiempo, debido a que ninguna teoría subyacente soporta tal condición y por último, el análisis estadístico implica que todas las variables sean medibles.

Aunque la diversidad de modelos es tan amplia como bancos hay en el mundo, los intentos de formalización giran alrededor de dos ideas que son: los que tratan de capturar *el olfato* de los expertos y sistematizarla aprovechando la tecnología, y los que se basan en técnicas econométricas o de análisis discriminante para llegar a un determinante de la calidad del crédito. Los primeros se reflejan, entre otros, en sistemas expertos y redes neuronales⁶. Por lo general este tipo de modelos permite establecer una calificación del cliente, pero su debilidad radica en que no existe un vínculo teórico identificable con la probabilidad de incumplimiento y la severidad de la pérdida. En el mejor de los casos, se puede establecer una

⁶ Véase José Cuenca Bartolomé, et al. (1990)

correspondencia entre las calificaciones arrojadas por el modelo y las probabilidades de quiebra, mediante estudios ex post cuando se cuenta con la información suficiente para tal fin. (Márquez: 2006)

La segunda corriente, se basa en técnicas estadísticas de discriminación. A diferencia del análisis experto mencionado anteriormente, la estimación de los ponderadores de la función discriminante se hace con datos reales de casos de empresas quebradas y sanas. El principal exponente de este tipo de modelo es Edward Altman con su trabajo pionero de Z-Score (1968) que será expuesto en detalle posteriormente.

En este sentido, existen varios métodos y modelos estadísticos empleados para el cálculo de la probabilidad de incumplimiento entre los cuales se encuentran: Matrices de transición, árboles de decisión, Regresión Logística, análisis discriminante, entre otros. (Arias: 2006)

4.1.1 Método de Matrices de Transición.

Este es el método que usa la SFC en los modelos de referencia. Permite calcular la probabilidad de pasar de un estado i en un período t a un estado j en un período $t + 1$.

En 1989, Altman desarrolló una metodología para estimar la probabilidad de incumplimiento a través de estimaciones de la tasa de incumplimiento; es decir: la razón entre los deudores que incumplieron con sus pagos, con respecto a los que estaban sirviendo puntualmente a su deuda.

Para este fin, Altman seleccionó una muestra histórica de bonos corporativos calificados por Standar&Poor's, para los cuales calculó la tasa de incumplimiento anual y acumulada, dada la calificación original. Los resultados obtenidos en ese estudio se muestran en la tabla a continuación

Tabla 2: Tasa de Incumplimiento por Calificación para Bonos Corporativos 1971-1994

Calificación original	Año	1	2	3	4	5	6	7	8	9	10
AAA	Mortalidad anual	0.00%	0.00%	0.00%	0.00%	0.60%	0.00%	0.00%	0.00%	0.00%	0.00%
	Mortalidad acumulada	0.00%	0.00%	0.00%	0.00%	0.60%	0.06%	0.06%	0.06%	0.06%	0.06%
AA	Mortalidad anual	0.00%	0.00%	0.47%	0.27%	0.60%	0.00%	0.01%	0.00%	0.04%	0.04%
	Mortalidad acumulada	0.00%	0.00%	0.47%	0.74%	7.40%	0.74%	0.74%	0.74%	0.78%	0.82%
A	Mortalidad anual	0.00%	0.00%	0.05%	0.15%	0.08%	0.16%	0.06%	0.17%	0.12%	0.00%
	Mortalidad acumulada	0.00%	0.00%	0.05%	0.19%	0.27%	0.43%	0.50%	0.67%	0.79%	0.79%
BBB	Mortalidad anual	0.03%	0.39%	0.41%	0.67%	0.40%	0.54%	0.21%	0.10%	0.10%	0.46%
	Mortalidad acumulada	0.03%	0.42%	0.82%	1.49%	1.88%	2.41%	2.62%	2.72%	2.81%	3.27%
BB	Mortalidad anual	0.44%	0.98%	3.41%	1.78%	2.80%	1.33%	2.75%	0.29%	1.69%	4.22%
	Mortalidad acumulada	0.44%	1.41%	4.77%	6.47%	9.09%	10.30%	12.76%	13.01%	14.49%	18.09%
B	Mortalidad anual	1.41%	4.31%	7.27%	6.93%	7.06%	6.24%	3.76%	1.96%	1.26%	1.64%
	Mortalidad acumulada	1.41%	5.65%	12.51%	18.58%	24.33%	29.25%	31.72%	33.06%	33.90%	34.99%
CCC	Mortalidad anual	2.46%	16.57%	17.67%	12.17%	4.50%	12.98%	1.63%	5.71%	0.00%	4.41%
	Mortalidad acumulada	2.46%	18.62%	33.02%	41.17%	43.82%	51.11%	51.91%	54.65%	54.65%	56.65%

Fuente: GUTIERREZ, Rodolfo y MARQUEZ, Javier (2003). “Los Métodos de Calificación de Cartera y su Importancia para los Paradigmas de Medición del Riesgo de Crédito”. Capítulo 1 en *La Medición Integral del Riesgo de Crédito*. Pág. 24. –Altman y Kishore (1997) –

En 1991, el mismo Altman utilizó datos de Standar&Poor’s para determinar la manera en que los bonos corporativos migran de un nivel de calidad a otro. Este trabajo dió origen al uso de cadenas de Markov, finitas y discretas, para modelar los procesos Aleatorios que caracterizan los cambios en la calidad de los créditos en el tiempo mediante matrices de transición. Dicha matriz contiene las probabilidades de incumplimiento de deudores en un período, partiendo de los diferentes niveles de calidad o calificaciones de cada crédito. Por último, las potencias de la matriz “k”, contienen las probabilidades de transición de un estado a otro en los próximos “k” períodos y, por lo tanto, las probabilidades de incumplimiento en esos mismos períodos.

**Tabla 3: Matriz de Transición con un Horizonte de Tiempo de un Año
Calculada con Bonos Emitidos entre 1971 y 1989 y calificados por S&P**

Calificación	AAA	AA	A	BBB	BB	B	CCC	Impago
AAA	94.30%	5.50%	0.10%	0.00%	0.00%	0.00%	0.00%	0.00%
AA	0.70%	92.60%	6.40%	0.20%	0.10%	0.10%	0.00%	0.00%
A	0.00%	2.60%	92.10%	4.70%	0.30%	0.20%	0.00%	0.00%
BBB	0.00%	0.00%	5.50%	90.00%	2.80%	1.00%	0.10%	0.30%
BB	0.00%	0.00%	0.00%	6.80%	86.10%	6.30%	0.90%	0.00%
B	0.00%	0.00%	0.20%	1.60%	1.70%	93.70%	1.70%	1.10%
CCC	0.00%	0.00%	0.00%	0.00%	9.00%	2.80%	92.50%	4.60%

Fuente: GUTIERREZ, Rodolfo y MARQUEZ, Javier (2003). “Los Métodos de Calificación de Cartera y su Importancia para los Paradigmas de Medición del Riesgo de Crédito”. Capítulo 1, En: *La Medición Integral del Riesgo de Crédito*. Pág. 25. –Altman y Kao (1992) –

En la tabla 3 se muestra la matriz obtenida por Altman. Los elementos sobre la diagonal principal de la matriz representan las probabilidades de que la calidad de un crédito no cambie en el período de tiempo especificado, por otra parte, en la última columna se obtienen las probabilidades de incumplimiento dada la calificación del crédito al inicio del periodo. Análogamente los números que se encuentran por fuera de la diagonal principal son las probabilidades de transición o migración de un nivel de calidad a otro.

4.1.2 Método de árboles de Decisión.

Con este método se pueden asignar probabilidades mediante la asignación de particiones recursivas, hasta llegar a una clasificación final. Para emplear esta metodología todas las variables que la componen deben ser categóricas. Las variables continuas deben, por lo tanto, convertirse en categóricas a través del uso de rangos de valores.

4.1.3 Modelo de Regresión logística.

Es un caso especial del modelo de regresión lineal. El modelo busca una ecuación con el fin de estimar o predecir una variable de respuesta de tipo binario o dicotómico que indica la presencia o ausencia de determinada condición en los individuos. Para aplicar la técnica de MCO se introduce una transformación

utilizando una función exponencial, la cual asume que la probabilidad (P) de éxito (Y=1) satisface la ecuación:

$$\ln\left(\frac{p}{1-p}\right) = \alpha + \beta' X \quad (2)$$

Donde β es un vector de parámetros y α representa el intercepto.

Para encontrar la probabilidad de éxito asociada ($\Pr [Y=1]$), se despeja p de la fórmula anterior, obteniendo

$$\Pr(Y = 1) = \frac{\exp^{(\alpha + \beta' X)}}{1 + \exp^{(\alpha + \beta' X)}} = \frac{1}{1 + \exp^{-(\alpha + \beta' X)}} \quad (3)$$

La probabilidad de no éxito ($\Pr [Y=0]$) corresponde al complemento de la probabilidad de éxito.

Características de la regresión logística:

1. no requiere supuestos de normalidad y homogeneidad de varianza para la variable dependiente.
2. acepta variables explicativas de escala o intervalo así como variables nominales u ordinales.
3. requiere que las variables explicativas sean no correlacionadas
4. Para que los resultados de las estimaciones del modelo de regresión logística sean válidos, “la presunta causa de cierto efecto necesariamente ha de preceder a este último en el tiempo” (SILVA: 1995).

Estimación del modelo de regresión logística

Para encontrar las estimaciones de los coeficientes del modelo logístico, se asume que el conjunto de valores objetivo $\{y_i, i = 1, \dots, n\}$ sigue una distribución binomial, tal que $y_i = 1$ representa la ocurrencia del suceso y $y_i = 0$ representa una falla o no ocurrencia, de esta forma, se deben encontrar los estimados de máxima verosimilitud de los coeficientes del modelo, utilizando un procedimiento numérico iterativo (algoritmo de Newton-Ralphson) con un programa de computadora.

Una vez calculados los coeficientes del modelo, se calcula la probabilidad de ocurrencia del suceso para todos los individuos de la muestra analizada.

Al igual que en la regresión lineal, se calculan pruebas para evaluar la significancia de los coeficientes del modelo y se realizan pruebas de bondad de ajuste.

4.1.4 Método de Análisis discriminante.

Esta técnica es utilizada para clasificar observaciones en grupos clasificados *a priori*. La clasificación se lleva a cabo a partir de la observación de un conjunto de variables que caracterizan a los individuos u objetos que se desea clasificar.

El análisis discriminante busca generar una combinación lineal de las características que mejor discriminen entre los grupos, es decir que maximice la varianza entre los grupos y minimice la varianza al interior de cada grupo.

Dentro de este tipo de análisis se reconocen los modelos Z-score, Z, Z1 y Z2 de Altman, entre otros.

4.1.4.1. Esquemas de calificación basados en la ponderación de factores de riesgo: El modelo Z-Score de Altman

Fue desarrollado originalmente para predecir las quiebras de las empresas. Se analizó la información financiera de una muestra inicial compuesta por 2 grupos de 33 empresas del sector manufacturero. Las empresas en bancarrota, agrupaba las empresas que habían entrado en bancarrota en el periodo comprendido entre 1946 y 1965. El grupo de las que no presentaron bancarrota fue seleccionado de una muestra de industrias manufactureras, elegidas y estratificadas aleatoriamente.

Después de que los grupos fueran definidos y las empresas seleccionadas, se recolectó la información de los estados financieros (balance y estado de resultados). Se seleccionó una lista de 22 variables (indicadores financieros o *ratios*), que fueron clasificadas en cinco categorías: Liquidez, rentabilidad, apalancamiento, solvencia y actividad (Heine: 2000).

De la lista de 22 variables, cinco se escogieron como las que mejor trabajo hacían como predictoras de quiebra en las empresas. La función discriminante finalmente obtenida fue:

$$Z = 0.012X_1 + 0.014X_2 + 0.033X_3 + 0.006X_4 + 0.999X_5 \quad (4)$$

Donde:

X_1 : Capital de trabajo/Activo total

X_2 : Utilidades retenidas/Activo total

X_3 : Utilidades antes de intereses e impuestos/Activo Total

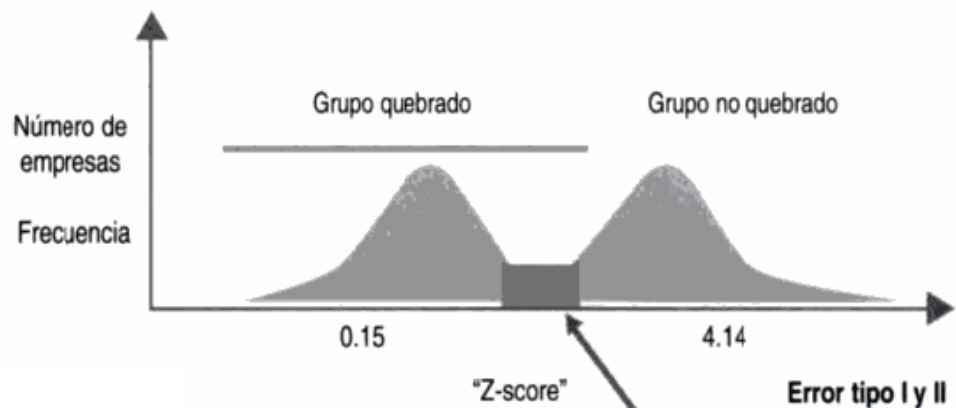
X_4 : Capital a valor de mercado/Pasivos Totales

X_5 : Ventas/Activos totales

Z: Discriminante

Del estudio empírico, se obtuvo que la clasificación “Z” estaba distribuida normalmente, y el promedio de las empresas sanas era de 4.14, mientras que la de las empresas quebradas era de 0.15. Así, obteniendo la calificación “Z” para una empresa deudora, se pueden hacer pruebas de hipótesis para saber a qué grupo pertenece.

Gráfica 2: Clasificación de Créditos por la Técnica “Z-Score” de Altman



Fuente: Márquez Díez-Canedo, Javier (2006). *“La Nueva Visión del Riesgo de Crédito”*. Pág. 23

El principal argumento usado en contra del modelo Z-score es que usa razones financieras y éstas tienen un efecto de subestimación en el tamaño de las estadísticas, es decir, no se pueden identificar datos relevantes en el otorgamiento de un crédito, como por ejemplo el tamaño de las empresas.

Otra de las falencias de este tipo de modelos es que su aplicación tanto a grandes como a pequeñas empresas es cuestionable, ya que para el desarrollo del mismo solo se emplearon las cifras de empresas grandes con problemas de solvencia. Adicionalmente, dado el alto grado de correlación que se presenta entre las variables explicativas del modelo, se devaluó el poder analítico del mismo. En consecuencia, los sistemas de evaluación basados en la decisión experta de los ejecutivos de crédito de las entidades financiera se vieron favorecidos (De la Fuente: 2003).

4.1.4.2 El modelo Z (1977).

Es, en esencia, el modelo Z-score con algunas modificaciones introducidas por Altman, Haldeman y Narayanan. El objetivo sigue siendo clasificar las empresas en bancarota.

Básicamente lo que cambia es que se incluyen empresas medianas y pequeñas en el análisis; se amplía el campo de actividad económica al incluir empresas por fuera del sector manufacturero; se incluyen los cambios en los estándares para el cálculo de las principales razones financieras y nuevas prácticas contables y, se cuenta con mejores y más recientes técnicas estadísticas para la estimación.

El resultado de la estimación de este modelo resulta ser mejor que el de z-score, pues la inclusión de empresas no manufactureras no arroja resultados inconsistentes y el modelo permite predecir quiebras con cinco años de anticipación con una confianza del 70% y con un año de anticipación con una confiabilidad del 96%.

Las debilidades fundamentales de este modelo es que no considera los diferentes ciclos económicos y por lo tanto presenta sesgos de ciclo económico, y adicionalmente, no cuenta con información sectorial que permita hacer análisis y consideraciones precisas sobre determinadas empresas.

Los modelos Z1 y Z2 son variaciones del modelo Z de Altman y modifican básicamente la composición de algunos vectores de la matriz X de información.

4.2. METODOLOGÍAS DE INFORMACIÓN DE MERCADO

Dado el caso en el que no se cuenta con bases de información robusta que permitan emplear las técnicas estadísticas, y si la información de los clientes es pública, entonces esta metodología podría resultar apropiada. Sin embargo, de

acuerdo con Aguas y Castillo (2005), esta metodología cuenta con serias limitaciones que incluyen:

- Limitación en la información de algunos clientes en el mercado, de acuerdo con la diversidad de la cartera con la que cuente el intermediario.
- Altos costos de acceso a la información de clientes que no sea pública.
- Ineficiencia para clasificar clientes en masa.

Adicionalmente, estas metodologías suponen neutralidad de riesgo y eficiencia del mercado, por lo cual no siempre resulta conveniente aplicar dichas técnicas en mercados como el colombiano.

5. OBJETIVO GENERAL

Aplicar un modelo de rating para medir el riesgo de crédito de la cartera empresarial de una entidad cooperativa, de conformidad con lo establecido en el capítulo II de la circular externa 022 de la SFC (2008).

5.1. OBJETIVOS ESPECIFICOS

- Identificar las razones o *ratios* financieros que determinan el incumplimiento en el pago de las obligaciones de la cartera empresarial en la entidad.
- Seleccionar la mejor metodología de rating que permita predecir el riesgo de crédito, para el segmento objetivo de acuerdo con la información disponible en la entidad.

6. METODOLOGÍA

Debido que el proyecto que origina el presente documento hace parte de la construcción de modelos de otorgamiento de crédito en la cooperativa financiera, deberá emplearse para tal fin la metodología aprobada por el consejo de administración de la entidad.

La metodología que se debe aplicar se conoce como CRISP-DM (CHAPMAN, 1999) por sus iniciales en inglés *Cross-Industry Standard Process for Data Mining*. Como su nombre lo indica, esta metodología puede ser utilizada en gran variedad de sectores y problemas de negocio.

El proceso general de la metodología CRISP-DM incluye seis fases que dirigen los puntos primordiales del proceso de minería de datos⁷. Las seis fases se ajustan y conjugan en un proceso cíclico.

Las seis fases se relacionan a continuación con la respectiva descripción de las actividades a realizarse en cada una de ellas:

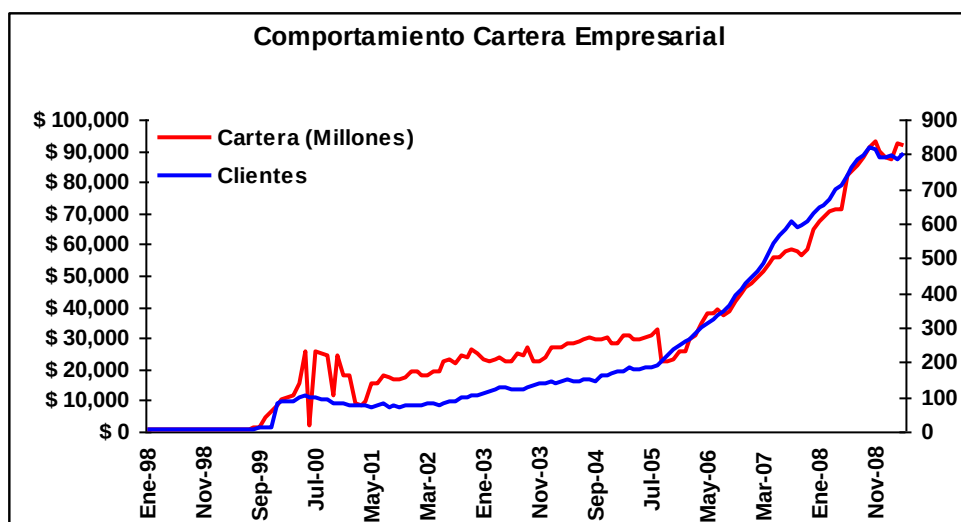
1. **Entendimiento del Negocio:** Esta fase llevó a cabo durante un período de práctica empresarial, y constituye la parte más importante del proyecto. Como su nombre lo indica, implica conocer el funcionamiento de la empresa, la normatividad vigente y determinar las necesidades que tiene la entidad para poder cumplir con sus objetivos estratégicos.
2. **Entendimiento de los datos:** De la misma manera que con la primera fase, ésta también se llevó a cabo durante el período de práctica empresarial. Consta de reconocer el tipo de información de la que se dispone y la manera como ésta puede llegar a ser usada para alcanzar el objetivo del proyecto. En este caso, se identificó información cuantitativa y cualitativa, representada en los archivos de análisis de crédito del segmento de banca empresarial, comprendido entre los periodos 2005 y 2008.
3. **Preparación de los datos:** Después de identificar las fuentes de información y el tipo de información con que se cuenta, deben iniciarse procesos de depuración, reconstrucción, integración y formateo de datos, con el fin de contar con datos de alta calidad para la etapa de modelación. Esta etapa consiste en tomar cada uno de los registros de la base de datos y verificar que cumplan con las características indispensables para iniciar la fase de modelación. Dichas características incluyen: calidad, completitud, oportunidad y fidelidad.
4. **Modelación:** Esta fase incluye la selección de la técnica de modelación, generación del diseño de las pruebas, construcción y valoración de modelos.
5. **Evaluación:** Una vez se haya elegido el modelo, se procederá a evaluar el aporte que los resultados del mismo hagan a la consecución del objetivo inicial, es decir a calificar de manera acertada los clientes pertenecientes a la cartera empresarial en el momento del otorgamiento del crédito para predecir la probabilidad de incumplimiento con un margen mínimo de error.
6. **Despliegue:** Esta fase se concentra en integrar los conocimientos y resultados de las fases anteriores en el día a día del negocio. Esta fase incluye la elaboración del plan de despliegue, monitoreo y mantenimiento. En resumen, lo que se busca es generar una herramienta que, al ser integrada dentro del proceso de otorgamiento, permita incrementar la asertividad de las mesas de crédito a la hora de aprobar o negar una solicitud de crédito del segmento empresarial.

⁷ La minería de datos es un proceso interactivo e iterativo, que permite integrar la tecnología con el conocimiento del negocio para identificar relaciones subyacentes y características en los datos. (SPSS Inc.: 2003, Pág. 3)

6.1 ENTENDIMIENTO DEL NEGOCIO Y DE LOS DATOS DISPONIBLES

A abril de 2009, la entidad cuenta con 804 clientes empresariales con una cartera que asciende a \$91.948 millones. El crecimiento sostenido de la cartera se ha observado a partir del año 2005 donde se mejoraron las herramientas de análisis crediticio y se dedicó personal exclusivo con experiencia en este segmento de clientes.

Gráfica 3: Comportamiento cartera empresarial

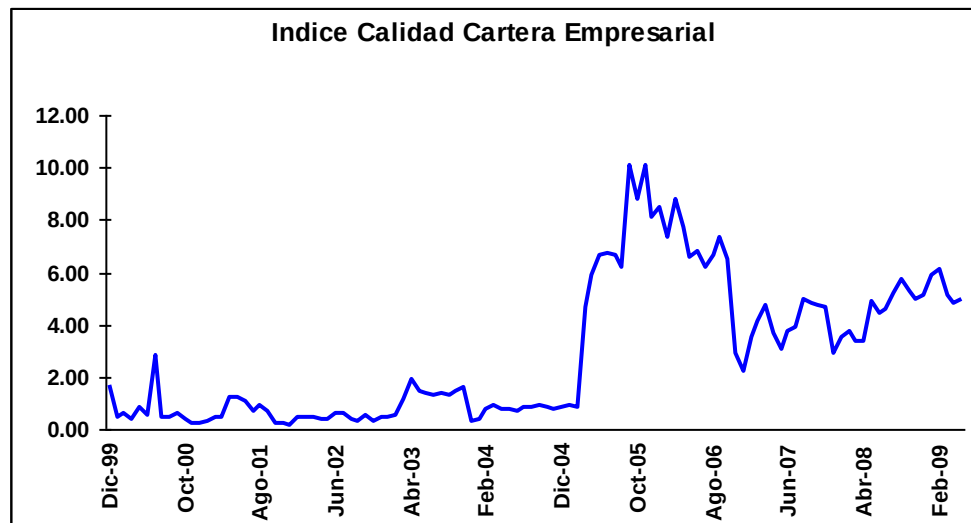


Fuente: Bodega SARC- Cálculos Propios.

En la historia de la Cooperativa, se tiene un total de 1.309 clientes que han solicitado crédito al menos una vez; de estos clientes, no se contaba con la información de los estados financieros en el aplicativo de la entidad, razón por la cual a partir del año 2007, se inició un proceso de recolección de esta información. A la fecha, se ha logrado recolectar información de 444 clientes.

En cuanto a calidad de la cartera empresarial (mora mayor a 30 días), se observa que la misma ha tenido un buen comportamiento a partir de mediados del año 2006 con una tendencia en los últimos dos años al crecimiento de la cartera vencida.

Gráfica 4: Comportamiento Índice de calidad Cartera empresarial



Fuente: Bodega SARC- Cálculos Propios.

La línea representativa de la cartera empresarial es la COE (Cartera Ordinaria Empresarial) que representa al cierre de abril el 66% de la cartera de este segmento (\$60.871 millones)

La información a utilizar está contenida en diversas fuentes: 1) En la bodega SARC, de donde se extrae lo referente a la colocación, la cartera y su posterior comportamiento de pago, así como la información cualitativa y cuantitativa no financiera de los clientes y, 2) Estados financieros en Excel que se obtienen durante el proceso de análisis crediticio.

6.2 PREPARACION DE DATOS

6.2.1 La Base De Datos

Se cuenta con información de 444 empresas que han tenido obligaciones de crédito con la entidad desde 2003 hasta septiembre de 2008. Cuenta con información cuantitativa (cifras de estados financieros, indicadores, etc.), cualitativa no financiera (antigüedad de constitución de la empresa) y cualitativa (nombre de la empresa, CIIU, entre otros)

Dicha base de datos se sometió a un proceso de depuración de datos inconsistentes e incompletos.

Las variables independientes correspondientes a los indicadores financieros fueron normalizadas para eliminar el efecto de la escala de medición de las diferentes variables con las que se cuenta.

6.2.2 Definición De Variables

Incumplimiento: Para efectos de estudio, se toma como referencia una altura de mora de 90 días, es decir: si una empresa que tiene un crédito vigente alcanza una altura de mora de 90 días o más entonces se considerará que la empresa ha caído en incumplimiento. Si la empresa alcanza una altura de mora de 89 días o menos se considerará como No incumplida.

Esta es entonces, una variable latente que se define de la siguiente manera:

Y= incumplimiento

$$Y = \begin{cases} Incumple_si_Y = 1 \\ No_incumple_si_Y = 0 \end{cases}$$

Variables explicativas: las principales variables son los indicadores o *ratios* financieros, los cuales proporcionan información importante sobre el estado de la estructura financiera de las empresas. Entre ellos se encuentran:

EBITDA (*“Earnings before taxes, depreciations and amortizations”*): Corresponde a las ganancias que tiene la empresa antes de realizar los descuentos por depreciaciones, amortizaciones e impuestos. Su cálculo se realiza de la siguiente manera:

EBITDA= Utilidad Operacional + Depreciaciones + Otros Ingresos operacionales – Otros Egresos Operacionales+ Amortizaciones + Ajustes por Inflación.

Es un indicador de capacidad de pago. Su crecimiento indica mayor capacidad de la empresa para atender a sus necesidades de financiación.

EBITDA/GASTOS FINANCIEROS. Es una razón que indica cuántas veces la empresa puede cubrir los intereses de sus obligaciones financieras. En general se desea que esta razón se encuentre por encima de 1.5 veces.

OBLIGACIONES FINANCIERAS/EBITDA: Esta relación muestra en qué magnitud puede responder la empresa a sus obligaciones financieras de corto y largo plazo. A medida que el EBITDA crece, la relación se hace más pequeña. Así, si el endeudamiento de la empresa crece en mayor proporción que el EBITDA, la relación se hará más grande y la capacidad de pago de la empresa se verá afectada negativamente. Lo deseable es que esta razón se encuentre por debajo de 3 veces.

OBLIGACIONES FINANCIERAS/VENTAS: Corresponde al cociente entre el total de las obligaciones financieras (corto y largo plazo) y las ventas. Indica la proporción de las ventas que es acaparada por el endeudamiento financiero de la empresa. Entre más pequeño sea este porcentaje, mejor será la situación financiera de la empresa. Lo ideal es que este indicador no sobrepase el 50%

PASIVO TOTAL/EBITDA: Esta razón muestra cuántas veces estaría la empresa en capacidad de asumir el total de sus pasivos con el EBITDA. Entre más pequeña sea esta relación menor será la probabilidad de incumplimiento de la empresa. Lo ideal es que este indicador permanezca por debajo de 5 (Domash: 2002)

RAZON CORRIENTE: Es un indicador que determina cómo está respaldado el pasivo de la empresa. La razón corriente se calcula como el cociente entre los Activos Corrientes y los Pasivos Corrientes de la empresa. Normalmente, se interpreta como el número de pesos de activo que la empresa tiene por cada peso de pasivo. De acuerdo con los expertos, una razón corriente “sana” se encuentra alrededor de 2. Si es muy superior puede que existan fondos ociosos que indiquen una mala gestión administrativa; si es muy pequeña, la empresa puede presentar serios problemas de liquidez.

PRUEBA ACIDA: Básicamente tiene el mismo propósito que la razón corriente, pero ésta castiga de alguna manera la menor liquidez de los inventarios así: $\text{Prueba ácida} = (\text{Activo corriente} - \text{inventarios}) / \text{Pasivo corriente}$. Esta razón normalmente, debería encontrarse alrededor de 1.5.

CAPITAL DE TRABAJO: Es la diferencia entre el activo corriente y el pasivo corriente de la empresa. Muestra el monto de capital que puede invertirse en la actividad económica de la empresa. Entre mayor sea esta relación, mejor será la situación financiera de la empresa.

ROTACIÓN DE CARTERA: Muestra cuántos días le toma a una empresa recuperar su cartera. Se calcula de la siguiente manera.

Rotación de Cartera= (cartera total/Ventas)*360 días

Este indicador no tiene valores de referencia por ser discrepante entre sectores macroeconómicos, pero lo ideal es que se mantenga en niveles acordes con el promedio del sector (p.e. 30 días para comercio)

ROTACION DE INVENTARIOS: Análogamente a la rotación de cartera, este indicador muestra cuántos días le toma a la empresa vender y renovar todo su inventario. Su cálculo se realiza de la siguiente manera:

Rotación de inventario= (inventario/costo de ventas)*360 días

ROTACIÓN DE ACTIVO: Este índice está dado por la relación entre las ventas netas y el activo total de la empresa. Muestra en qué proporción crecen las ventas por cada peso invertido en el activo. La disminución de un año a otro en la rotación de activo, puede indicar que la empresa emplea ahora un nivel proporcionalmente mayor de activos para generar ventas.

ROTACIÓN DE PROVEEDORES: Este indicador muestra cuántos días le toma a la empresa pagar sus cuentas a los proveedores. Este, también depende del sector en el que la empresa desarrolla su actividad económica. Se calcula de la siguiente manera:

Rotación de proveedores= (cuentas con proveedores/costo de ventas)*360 días.

Es conveniente en el análisis ver cómo es la relación entre éste indicador y el de rotación de cartera para conocer cómo se afecta el flujo de caja.⁸

NIVEL DE ENDEUDAMIENTO: Como su nombre lo indica, mide el endeudamiento total de la empresa mediante la relación de los pasivos y los activos; en general, este indicador no debería sobrepasar el 70%, sin embargo ello depende de las políticas de otorgamiento definidas por cada entidad.

⁸ Si la empresa debe pagar más a sus proveedores en un tiempo menor al que se demora en recuperar la cartera, entonces se puede afectar negativamente el flujo de caja.

NIVEL DE ENDEUDAMIENTO FINANCIERO. Este indicador muestra qué parte del endeudamiento de la empresa corresponde a obligaciones con el sector financiero, es decir, qué parte del activo está apalancado con endeudamiento financiero. Se calcula como sigue:

Endeudamiento financiero= Obligaciones Financieras (C.P+L.P)/Activo Total, donde: C.P representa el corto plazo y L.P el largo plazo.

El nivel máximo deseable de este indicador se ubica alrededor del 50%, sin embargo el límite de tolerancia lo imponen las políticas de otorgamiento de las entidades financieras.

CONCENTRACION DEL PASIVO EN EL CORTO PLAZO: Como su nombre lo indica, muestra cuál es la proporción del pasivo exigible por los acreedores en un período de tiempo inferior a un año. El efecto Neto del crecimiento de esta variable sobre la probabilidad de incumplimiento, depende de otros aspectos financieros de la empresa como la liquidez y las características del sector, entre otros.

OBLIGACIONES FINANCIERAS/PATRIMONIO: Mide la proporción del patrimonio que se compromete con las obligaciones financieras de la empresa. Los niveles deseables de este indicador dependen igualmente del contexto en el que se ubique la empresa, sin embargo, por lo general se desea que se encuentre en niveles bajos.

ROA (RENTABILIDAD DEL ACTIVO): Es la relación existente entre la utilidad Neta y el activo total. Como su nombre lo indica, muestra la rentabilidad de los activos de operación de la empresa. Se espera que una empresa rentable desde el punto de vista financiero, presente resultados en este indicador por encima del IPC.

ROE (RENTABILIDAD DEL PATRIMONIO): De la misma manera que con el ROA, este indicador relaciona la utilidad con el patrimonio total como un indicador de rentabilidad financiera de la empresa. Se espera que en empresas rentables financieramente, éste se encuentre por encima del valor del IPC.

Estos *ratios* influyen positiva o negativamente en el crecimiento de la probabilidad de que un sujeto de crédito sea cumplido o no. En la tabla 4, se presentan los diferentes efectos que el crecimiento de cada variable ejerce sobre el incremento de la probabilidad de incumplimiento. Por ejemplo, la variable I_EDAGTOSF a

medida que crece, se reduce la probabilidad de que la empresa incurra en incumplimiento ya que el crecimiento de dicho indicador implica que la empresa tiene mayor capacidad de atender a sus obligaciones crediticias.

Para efectos de modelación, se presentan los indicadores con la letra “I” antecediendo a una abreviación de su nombre.

Tabla 4: Efecto del Crecimiento de las variables Frente al Incumplimiento.

VARIABLE	EFFECTO DEL CRECIMIENTO
I_EDAGTOSF	(-)
I_EDAVTAS	(-)
I_OBFEDA	(+)
I_OBLFVTAS	(+)
I_PASTEDA	(+)
I_RZCTE	(-)*
I_PACIDA	(-)*
I_CPTTRAB	(-)
I_ROTACART	(+)
I_ROTAINVE	(+)
I_ROTAACT	(+)
I_ROTAPROVE	(-)
I_NIVENDEU	(+)
I_NIVENDF	(+)
I_CONCENPCP	(+)**
I_OBFPATR	(+)
I_ROA	(-)
I_ROE	(-)

* Estos *ratios* preferiblemente debe encontrarse alrededor de 2

** El efecto neto del crecimiento de esta variable depende de otros aspectos financieros de la empresa como la liquidez y las características del sector, entre otros.

Fuente: Bodega SARC, cálculos propios.

6.2.3 Diagnóstico de Variables Explicativas

En la siguiente tabla, se puede observar que las variables en general no tienen valores inválidos, con excepción de la variable I_EDAGTOSF, la cual presenta un

27% de valores nulos o inválidos, por dicha razón se suprime dicha variable, a pesar de que la misma tiene sentido de negocio.

Tabla 5: Valores Válidos de Variables Independientes (*Ratios Financieros*).

INDICADOR	Válidos	% Faltantes	Promedios	
			Cumplidos	No cumplidos
EDAD	444	0%	33.335	8.011
I_EDAGTOSF	326	27%	917.824	7.257
I_EDAVTAS	444	0%	0.733	0.113
I_OBFEDA	444	0%	1.301	0.477
I_OBLFVTAS	444	0%	0.128	0.433
I_PASTEDA	444	0%	8.281	8.204
I_RZCTE	444	0%	8.626	2.151
I_PACIDA	444	0%	8.235	1.850
I_CPTTRAB	444	0%	453.085	4.378
I_ROTACART	444	0%	98.982	154.720
I_ROTAINVE	444	0%	63.056	59.539
I_ROTAACT	444	0%	1.909	1.783
I_ROTAPROVE	444	0%	49.344	113.709
I_NIVENDEU	444	0%	0.503	0.539
I_NIVENDF	444	0%	0.124	0.138
I_CONCENPCP	444	0%	0.783	0.774
I_OBFPATR	444	0%	0.424	0.487
I_ROA	444	0%	0.096	0.106
I_ROE	444	0%	0.221	0.233

Fuente: Bodega de datos SARC, cálculos propios

Numero de empresas incumplidas y no incumplidas

Como puede verse en la tabla 6, la cartera de la entidad en cuestión es considerablemente sana, teniendo en cuenta que el 81% de las empresas no sobrepasan la altura de mora de los 89 días.

Tabla 6: Distribución del Incumplimiento en la Cartera por Regionales

REGIONAL	NO INCUMPLE		INCUMPLE		TOTAL
	#	% FILA	#	% FILA	
Bogota	173	80.84%	41	19.16%	214
Cali	76	76.77%	23	23.23%	99
Caribe	45	91.84%	4	8.16%	49
Eje Cafetero	22	66.67%	11	33.33%	33
Medellin	40	85.11%	7	14.89%	47
Palmira	2	100.00%	0	0.00%	2
TOTAL	358	80.63%	86	19.37%	444

Fuente: Bodega de Datos SARC, cálculos propios.

6.3 MODELOS

6.3.1 Selección de Variables

Se realizó un análisis de correlación entre estas variables con el fin de evitar el problema de variables redundantes. Los resultados se presentan en el Anexo 1.

En este, se puede apreciar que existen relaciones fuertes entre las variables I_EDAGTOSF e I_EDAVTAS, la relación es muy cercana; esto se debe a que ambas variables tienen un indicador común que es el EBITDA, debido a ello y adicionalmente por la cantidad de valores atípicos de la primera anteriormente mencionados (ver tabla 5), se tomó la decisión de excluirla de las variables de trabajo. Este tipo de relaciones se presentan a lo largo de la base de datos y la mayoría de estas se presentan en condiciones similares a la mencionada anteriormente, por ello, se tomó la decisión de trabajar con aquellas variables que pesan más (desde el punto de vista de los expertos) en el comportamiento de pago de las empresas. Cabe anotar que la exclusión de las demás variables no implica que estas no sean importantes en el otorgamiento de crédito pero para efectos de modelación ellas resultarían redundantes y esto podría traducirse en problemas de especificación en el modelo.

Se probó el nivel de discriminación de cada variable cuantitativa a través de la prueba Kolmogorov-Smirnov para hacer una selección inicial de aquellas variables con más alto poder explicativo.

Con un nivel de confianza del 95% se observa en la muestra que las variables I_CPTTRAB, EDAD y I_PACIDA son las únicas que está presentando discriminación a nivel univariado entre el grupo de clientes cumplidos e incumplidos. Los resultados de la prueba se muestran en la Tabla 7 a continuación:

Tabla 7: Prueba de Discriminación Kolmogorov-Smirnov

Variables Independientes	Estadísticos de contraste ^a			Z de Kolmogorov-Smirnov	Sig. asíntot. (bilateral)
	Absoluta	Positiva	Negativa		
EDAD	.306	.000	-.306	2.547	.000
L_EDAVTAS	.077	.077	-.031	.640	.807
L_OBLFVTAS	.107	.107	-.017	.890	.406
L_PASTEDA	.087	.087	-.057	.722	.675
L_PACIDA	.181	.037	-.181	1.505	.022
L_CPTTRAB	.199	.006	-.199	1.661	.008
L_ROTAINVE	.110	.110	-.035	.920	.366
L_ROTAACT	.072	.037	-.072	.598	.867
L_NIVENDF	.102	.102	-.020	.848	.469
L_CONCENPCP	.065	.065	-.060	.542	.931
L_OBFPATR	.091	.091	-.022	.756	.617
L_ROE	.096	.096	-.076	.802	.541

a Variable de Agrupación: Incumple

Fuente: Bodega de Datos SARC, cálculos Propios.

Utilizando la información de CIU, se creó la variable sector económico, la cual se agrupó de acuerdo al porcentaje de clientes cumplidos e incumplidos en cada categoría, para dar lugar a una variable indicadora: los clientes pertenecientes a los sectores de Otras Actividades de Servicios Comunitarios, Sociales y Personales, Intermediación Financiera, Servicios Sociales y de Salud, Educación, Administración Pública y Defensa; Seguridad Social de Afiliación Obligatoria y Explotación de Minas y Canteras se agruparon en el grupo de clientes con mejor comportamiento y el resto se agrupó en otra categoría.

6.3.2 Selección de la Muestra

Dado que generalmente la proporción de clientes incumplidos dentro de la muestra de análisis es muy pequeña⁹, se usan métodos que permitan disminuir la base de clientes cumplidos. Para esto, se aplicó un muestreo sistemático sobre el grupo de clientes cumplidos y aleatoriamente se escogieron 90 registros de este grupo de clientes. La muestra final estuvo conformada por 50 clientes cumplidos (51%) y 86 clientes incumplidos (49%) de la muestra. Dado que se tiene una muestra muy pequeña de clientes, el total de la muestra se utilizará para desarrollar el modelo.

6.3.3 Metodología Estadística

Para la elaboración de los modelos se seleccionó el método de regresión logística, dado que es una metodología sencilla que no requiere de supuestos

⁹ En este caso en particular se tienen únicamente 76 casos (17% de la muestra)

distribucionales y que sus resultados se pueden interpretar directamente como probabilidades de incumplimiento sin necesidad de realizar transformaciones adicionales como lo requieren algunas técnicas como redes neuronales y árboles de decisión.

Con las diferentes variables disponibles y teniendo en cuenta que el modelo logístico es sensible a correlaciones altas entre las variables independientes, se llevó a cabo un proceso de pruebas de diferentes conjuntos de variables, hasta llegar a un modelo que cumpliera las pruebas establecidas.

6.3.4 Modelo Final

Se aplicó el modelo de regresión logística sobre la muestra seleccionada usando el paquete estadístico STATA. Las variables consideradas fueron:

Variable Dependiente: Incumple

Variables Independientes: i_nivendf_std, i_edavtas_std, i_cpctrab_std, i_rotaact_std, i_obfpatr_std, sector, edad_r, inter1

Los resultados del modelo final, elegido para el presente ejercicio, se muestran en la tabla 8 a continuación:

Tabla 8: Parámetros estimados para el modelo de regresión logística elegido

incumple	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]
i_nivendf_~d	.502974	.3001736	1.68	0.094	-.0853556 1,091303
i_edavtas_~d	-1750631	1,306951	-1.34	0.180	-4.312.207 8,109459
i_cpctrab_~d	-2808321	1,037099	-2.71	0.007	-4.840.998 -.7756439
i_rotaact_~d	-.3306086	.1849333	-1.79	0.074	-.6930712 .031854
i_obfpatr_~d	.0703841	.6383478	0.11	0.912	-1.180.755 1,321523
sector	1,270802	.3897574	3.26	0.001	.5068919 2,034713
edad_r	1,491664	.4389816	3.40	0.001	.6312755 2,352052
inter1	-.3843648	.1939267	-1.98	0.047	-.7644542 -.0042755
_cons	-2281621	.7550101	-3.02	0.003	-3.761.414 -.8018284

Fuente: Bodega de Datos SARC, cálculos propios.

A nivel general, haciendo una validación cualitativa¹⁰ del modelo, se observa que los signos de los coeficientes son coherentes y tienen sentido de negocio.

¹⁰ Es decir que tenga sentido de negocio por las variables que el modelo usó.

$\beta_{I_NIVENDF}$: es positivo, esto implica que a medida que el valor del indicador crece, la probabilidad de incumplimiento asociada también se incrementa.

$\beta_{I_EDAVTAS}$: es negativo. Esto implica que a medida que este indicador crece, la probabilidad de que la empresa incurra en incumplimiento se reduce.

$\beta_{I_CPTTRAB}$: es negativo. Esto implica que cuando el valor de la variable crece, la probabilidad de que el cliente incurra en incumplimiento se reduce.

$\beta_{I_ROTAECT}$: este coeficiente, muestra que entre mayor sea el valor de la variable, menor será la probabilidad de que ésta incurra en incumplimiento.

$\beta_{I_OBFPATR}$: este coeficiente es positivo, por lo tanto la probabilidad de incumplimiento crece.

β_{SECTOR} : el coeficiente es positivo. Este resultado indica que dependiendo del sector macroeconómico al que pertenezca la empresa, la probabilidad de incumplimiento crece. Anteriormente se mencionaron las categorías de referencia, por lo que las empresas que no pertenecen a esos sectores tienen una P.I. mayor.

β_{EDAD_R} : el coeficiente es positivo, por lo que se explica que si una empresa está recién constituida, entonces la probabilidad de incumplimiento es más alta.

Al analizar la significancia individual de cada variable, se observa en la tabla 8 que las variables significativas a un nivel de confianza del 95% son: $i_cpttrab_std$, $sector$, $edad_r$ e $inter1$ (esta última corresponde a la interacción entre las variables $i_obfpatr_std$ y $i_nivendf_std$), por este motivo en el modelo se dejan las variables individuales a pesar que no sean significativas.

El cálculo de la probabilidad de incumplimiento a partir de este modelo para cada uno de los clientes se haría con el siguiente modelo:

$$P(Y = 1) = \frac{1}{1 + e^{\left(0.502974x_{i_nivendf_std} - 1.750631x_{i_edavtas_std} - 2.808631x_{i_cpttrab_std} - 0.3306086x_{i_rotaect_std} + 0.703841x_{i_obfpatr_std} + 1.270802x_{sector} + 1.491664x_{edad_r} - 0.3843648x_{inter1} - 2.281621\right)}}$$

6.4 EVALUACIÓN DEL MODELO

El modelo debe ser evaluado desde diferentes perspectivas, la primera de ellas es que supere las pruebas establecidas para un modelo de regresión logística sobre la muestra usada para su ajuste, estas son: la prueba chi-cuadrado de significancia conjunta de las variables y la prueba de Hosmer Lemeshow. De igual

forma se deben aplicar pruebas de discriminación sobre las probabilidades estimadas por ese modelo, la prueba usada aquí es la del Área bajo la curva ROC.

6.4.1 Pruebas de Bondad de Ajuste y Significancia conjunta de variables

Los resultados de la prueba de bondad de ajuste de Hosmer Lemeshow se presentan en la tabla 9.

Tabla 9: Prueba de bondad de ajuste de Hosmer Lemeshow

Group	Prob	Obs_1	Exp_1	Obs_0	Exp_0	Total
1	0.1677	2	1.4	16	16.6	18
2	0.2532	2	3.9	16	14.1	18
3	0.3369	4	5.0	13	12.0	17
4	0.3948	7	6.7	11	11.3	18
5	0.4653	8	7.4	9	9.6	17
6	0.5573	11	9.3	7	8.7	18
7	0.6672	12	11.2	6	6.8	18
8	0.7506	11	12.1	6	4.9	17
9	0.8244	15	14.1	3	3.9	18
10	0.9878	14	14.9	3	2.1	17

number of observations	=	176
number of groups	=	10
Hosmer-Lemeshow chi2(8)	=	3.64
Prob > chi2	=	0.8877

Fuente: Bodega de datos SARC, Cálculos propios.

De esta tabla, se observa que el modelo se ajusta a los datos, por cuanto a un nivel de confianza del 95% no existe suficiente evidencia para rechazar la hipótesis nula (“El modelo se ajusta a los datos”).

Los resultados de la prueba de significancia conjunta de las variables se presentan en la tabla 10:

Tabla 10: Prueba de significancia conjunta de las variables

Number of obs	=	176
Wald chi2(8)	=	31.34
Prob > chi2	=	0.0001
Pseudo R2	=	0.1997

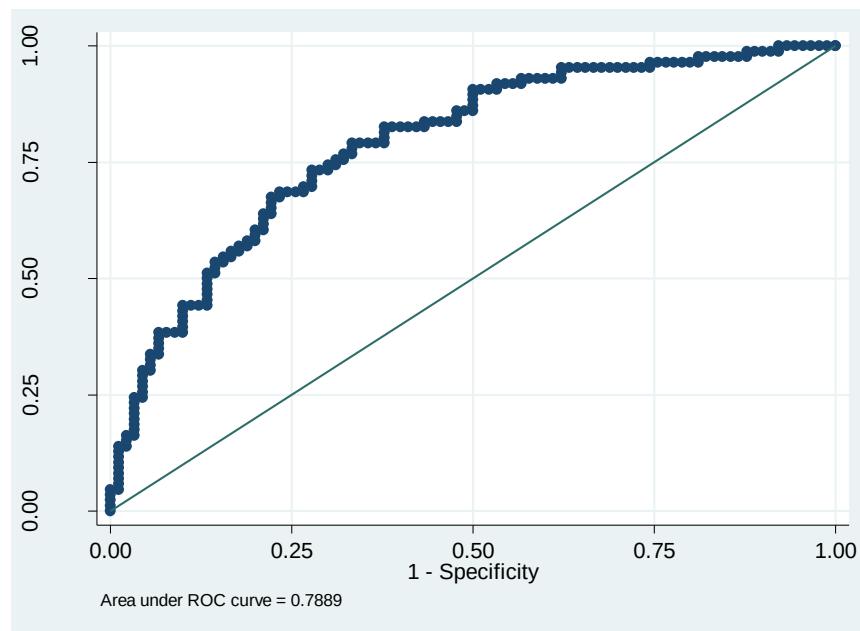
Esta prueba con un valor $p=0.0001$, muestra que hay suficiente evidencia estadística a un nivel del 95% de confiabilidad para concluir que en la muestra analizada al menos una de las variables contribuye de forma significativa a explicar el modelo.

6.4.2 Pruebas de discriminación

Al aplicar la prueba de discriminación del área bajo la curva ROC, se obtiene un valor de 0.7889, lo cual indica una buena capacidad de discriminación. En general, se espera que los valores de este estadístico se encuentren por encima de 0.5.

Si se calcula el Coeficiente de Gini, este da un valor de 0.5778.¹¹, el cual se espera también que esté por encima de 0.5.

Gráfica 5: Área bajo la curva ROC



El área bajo la curva ROC, puede interpretarse como la probabilidad de que ante un par de individuos, uno moroso y otro no moroso, el modelo los clasifique correctamente. (ARIAS: 2006)

Dado el resultado presentado en la gráfica 5, se puede concluir que el modelo presenta una buena capacidad de discriminación con un área aproximada de 78,89%.

¹¹ Se calcula a partir del área bajo la curva ROC como: $2 \times \text{Área curva ROC} - 1$

6.4.3 Pruebas de clasificación

Tabla 11: Tabla de Clasificación del Modelo.

estat classification, all
Logistic model for incumple

Classified	D	~D	Total
+	60	24	84
-	26	66	92
Total	86	90	176

Classified + if predicted $\Pr(D) \geq .5$
True D defined as incumple != 0

Sensitivity	$\Pr(+ D)$	69.77%
Specificity	$\Pr(- \sim D)$	73.33%
Positive predictive value	$\Pr(D +)$	71.43%
Negative predictive value	$\Pr(\sim D -)$	71.74%

False + rate for true ~D	$\Pr(+ \sim D)$	26.67%
False - rate for true D	$\Pr(- D)$	30.23%
False + rate for classified	$\Pr(\sim D +)$	28.57%
False - rate for classified	$\Pr(D -)$	28.26%
Correctly classified		71.59%

Fuente: Bodega de datos SARC, Cálculos propios.

Para el cálculo de la tabla de clasificación es necesario establecer un punto de corte para diferenciar las observaciones en sucesos probables y sucesos improbables. Este nivel se fija generalmente en 0.5 que es el punto medio de la distribución logarítmica. En este caso, puede apreciarse que el modelo clasifica correctamente el 71.59% de las observaciones, lo que indica que, en principio, el modelo tiene un buen nivel de exactitud. Es importante anotar que este modelo como todos los usados en el sector financiero se expone a calibraciones permanentes que permitan elevar gradualmente el nivel de clasificaciones correctas del mismo. En cuanto a los clientes incumplidos efectivamente clasificados como tal, se tiene que el modelo presenta una sensibilidad del 69,77% y para los clientes cumplidos clasificados correctamente, se tiene una especificidad del modelo de 73,33%.

6.4.4 Pruebas sobre una muestra de validación

Dado que se contaba con una muestra muy pequeña, no se tuvo información suficiente para probar el modelo sobre una muestra de validación tomada por fuera del rango de tiempo de la muestra de desarrollo del modelo. Esta actividad se deja para un proyecto futuro cuando la entidad crediticia cuente con información suficiente.

6.5 DESPLIEGUE DEL MODELO

La documentación del modelo desarrollado, será entregada la entidad financiera con el fin de que la misma pueda llevar a cabo la implementación del modelo en su aplicativo actual. Este modelo se constituye en una herramienta valiosa para establecer procesos de otorgamiento que cuantifiquen el riesgo de pérdidas para la entidad y de esta forma se seleccionen los clientes menos riesgosos.

7. CONCLUSIONES

- Durante el proceso de construcción del modelo, se evidenció que los ratios financieros no son las únicas variables determinantes en el comportamiento de pago de los deudores, y que existen variables de tipo no financiero que marcan notablemente dicho comportamiento tales como el sector macroeconómico, la edad de constitución, entre otras.
- Se evidencia que para la cartera empresarial de la entidad, las principales variables financieras que marcan el comportamiento de pago de los clientes corresponden a aquellas relacionadas con el nivel de endeudamiento financiero y capital de trabajo.
- Dado que este es el primer modelo que se genera para el segmento de Banca Empresarial, es necesario que en el proceso de clasificación de clientes nuevos se empleen adicionalmente procedimientos de tipo experto, que complementen el análisis hasta calibrar completamente el modelo.
- Se comprobó que la técnica de regresión logística dado el tamaño de muestra pequeña ofrece buenos resultados para predecir el comportamiento de la cartera de crédito empresarial en la fase de otorgamiento.
- Se recomienda que la entidad haga un esfuerzo por reconstruir la información histórica faltante de sus clientes y automatizar el proceso de análisis y almacenamiento de la información de solicitudes de crédito de

sus clientes Banca Empresarial, dado que la principal dificultad encontrada para probar el modelo desarrollado fue la falta de información.

- Se recomienda que la entidad adelante todas las actividades necesarias para poner en marcha la implementación del nuevo modelo de rating, como herramienta para seleccionar los sujetos de crédito en el otorgamiento.

REFERENCIAS BIBLIOGRÁFICAS

AGUAS, D. y CASTILLO, M. 2005. *Modelo de Administración de Riesgo Crediticio Para la Cartera Comercial de una Entidad Financiera Colombiana*. Universidad de los Andes. Bogotá D.C.

ALMEIDA, F., SIQUEIRA, J. 1997. "Comparação Entre Regressão Logística e Redes Neurais na Previsão de Falência de Bancos Brasileiros" *Terceiro Congresso Brasileiro de Redes Neurais*, 4. Florianópolis pp. 1-6.

ALTMAN, Edward I. 1968. "Financial Ratios, Discriminant Analysis and Prediction of Corporate Bankruptcy", *Journal of Finance* Vol. 23, No. 4 pp. 589-609

ARIAS, Jéimar 2006. *Modelo de Riesgo de Crédito en el Sector Cooperativo*. Trabajo de grado presentado como requisito para optar el título de Especialista estadística. Universidad Nacional de Colombia, Bogotá D.C.

BASCH, Michael, MONTENEGRO, C. 1989. "Aplicación de Modelos Multivariados a la Predicción de Quiebra, Evidencia Empírica" *Paradigmas en Administración*. No. 14, Primer Semestre pp. 37-35.

BEAVER, W.H. 1966. "Accounting ratios and the Prediction of Failure: Some Behavioral Evidence" *Journal of Accounting Research*. No.1 pp. 1-25.

BLUM, M.P. 1974. "Failing Company Discriminant Analysis". *Journal of Accounting Research*. Vol. 4 pp. 71-111.

BONTZ, J.E., KENNEDY, D.B., DE MIRANDA, E., ALBURQUENQUE, A. 1995. "Predicting Corporate Failure Using a Neural Network Approach. Intelligent Systems" En: *Accounting, Finance and Management School of Accountancy*. John Wiley & Sons Ltda. Vol. 4 pp.95-111

CHAPMAN, Pete, CLINTON, Julian, KERBER, Randy, KHABAZA, Thomas, REINARTZ, Thomas, SHEARER, Colin y WIRTH, Rüdiger. 1999. "Step-by-Step Data Mining Guide" *CRISP-DM 1.0*. SPSS Inc. Chicago, IL.

CUENA BARTOLOMÉ, José, GARCÍA-SERRANO, Ana y MARQUEZ DIEZ-CANEDO, Javier. (1990), "CYRCE: A Pattern of Knowledge Representation and Knowledge Acquisition for Decision on Personal Loans". *Expert Systems With Applications*, Vol. 1, pp. 155-160, Pergamon Press Pic.

DEAKIN, E.B. 1972. "A Discriminant Analysis of Business Failure." *Journal of Accounting Research*. No. 1 pp. 167-179.

DE LA FUENTE, Maria de Lourdes 2003. "Modelos de Pérdida Esperada". Capítulo 2 En: *La Medición Integral del Riesgo de Crédito*. Editorial Limusa S.A. de C.V., México, D.F. 2003.

DOMASH, Harry 2002. *"Fire Your Stock Analyst! Analyzing Stocks on Your Own"* Prentice Hall. 2002.

EDMINSTER, R.O. 1972. "An Empirical Test of Financial Ratio Analysis for Small Business Failure Prediction." *Journal of Financial and Quantitative Analysis*. No.2 pp. 1477-1493.

EPISCOPOS, A. 1997. "Artificial Neural Networks", Dean Paxson and Douglas Wood (eds) *Blackwell Encyclopedic Dictionary of Finance*, pp. 3-4, Oxford, Blackwell.

GENTRY J.A., WHITFORD D.T., NEWBOLD, P. 1968. "Predicting Industrial Band Ratings With a Probit Model And Funds Flow Components." *Financial Review*. Vol. 23 pp. 269-286.

GOLINSKY, G. 1998. *"Bankruptcy Prediction Using Neural Networks"* MBA Paper.

GURLEY, J.G y E.S. Shaw 1955. "Financial Aspects of Economic Development" *American Economic Review*, No.45, pp. 515-538

GUTIERREZ, Rodolfo y MARQUEZ, Javier (2003). "Los Métodos de Calificación de Cartera y su Importancia para los Paradigmas de Medición del Riesgo de Crédito". Capítulo 1 en *La Medición Integral del Riesgo de Crédito*. Editorial Limusa S.A. de C.V., México, D.F. 2003.

HEINE, M 2000. *"Predicting Financial Distress of Companies: Revisiting the Z-score and ZETA® models"* New York, Julio de 2000.

KIYOUTAKI, Nobuhiro y MOORE John 1997. "Credit Cycles". *The Journal of Political Economy*, Vol. 105 No.2 (Apr., 1997), pp.211-248.

KYD, Charles W. 1988. *"Aplicaciones de Lotus 123. Modelos Financieros"*. McGraw-Hill, España.

LIBBY, R. 1975. "Accounting Ratios and the Prediction of Failure.: Some Behavioral Evidence". *Journal of Accounting Research*. No.1 pp. 150-161

LÓPEZ GARCIA, Mauricio José, PALACIOS, Fernando 2005. *Data Envelopment Analysis en la Clasificación de Clientes Para la Banca Empresarial*. Universidad de los Andes, Bogotá D.C.

MARQUEZ DIEZ-CANEDO, Javier 2006. *“Una Nueva Visión del Riesgo de Crédito”*. Editorial Limusa S.A. de C.V., México, D.F. 2006.

MORA, A. 1994. “Los modelos de Predicción del Fracaso Empresarial: una Aplicación Empírica del Logit.” *Artículos Doctrinales Revista Española de Financiación y Contabilidad*. Vol.23 No. 78.

NARVAEZ, L.A. 1994. *“Predicción de la Quiebra de Empresas”*. Apuntes: Universidad Austral de Chile. Master en Administración de Empresas. Valdivia, Junio (1994).

OHLSON, J.1980. “Financial Ratios and the Probabilistic Prediction of Bankruptcy”. *Journal of Accounting Research*, Spring. Pp. 109-131

PLATT, PLATT y GUNNAR. Junio de 1994. “Bankruptcy Discrimination With Real Variables”. *Journal of Business Finance and Accounting*. 24(4).

SERRANO, C., GALLIZO, J.L, 1998. *“Las Redes Neuronales Artificiales en el Tratamiento de la Información Financiera.”* Departamento de Contabilidad y Finanzas, Facultad de Ciencias Económicas y Empresariales. Universidad de Zaragoza.

SILVA, Aycaguer L.C.: *“Excursión a la Regresión Logística en Ciencias de la Salud”* Díaz de Santos, Madrid. 1995.

SPSS Inc. 2003. “A Strategy for Data Mining: The CRISP-DM Process Methodology” En: *Introduction to Clementine*. Capitulo 1, pp. 1-4.

STIGLITZ, Joseph E. y WEISS, Andrew 1981. “Credit Rationing in Markets With Imperfect Information”. *The American Economic Review*, Vol. 71, No. 3 (Jun., 1981), pp. 393-410.

TAMARI, M. 1966. “Financial Ratios as Means of Forecasting Bankruptcy”. En: *Management International Review*. N0.4 pp. 15-21

THEODOSSIOU, P. 1991. “Alternative Models for Assessing the Financial Condition of Business in Greece”. *Journal of Business & Accounting*. Vol.18. pp. 697-723. Sep. 1991.

WILCOX, A. 1971 “A Simple Theory of Financial Ratios as Predictors of Failure.” *Journal of Accounting Research*. pp. 389-395

WINAKOR, A., SMITH, R. 1935. *“Changes in the Financial Structure of Unsuccessful Industrial Corporations.”* Bulletin No.51. University of Illinois, Bureau of Business Research, Urbana, Illinois.

Superintendencia Bancaria de Colombia 2005. Documento de Divulgación del Manual de Supervisión del SARC.

<http://www.superfinanciera.gov.co/ComunicadosyPublicaciones/supervisionsarc0805.pdf>

Superintendencia Financiera de Colombia 2008. Circular Externa 022 de 2008.

ANEXO 1: CORRELACIONES ENTRE VARIABLES INDEPENDIENTES

		Correlaciones																			
		EDAD	L_EDAGTOSF	L_EDAVTAS	L_OBFEDA	L_OBLFVTAS	L_PASTEDA	L_RZCTE	L_PACIDA	L_CPTTRAB	L_ROTACART	L_ROTAINVE	L_ROTAACT	L_ROTAPROVE	L_NIVENDEU	L_NIVENDF	L_CONCENPCP	L_OBFPATR	L_ROA	L_ROE	
EDAD	Correlación de Pearson	1	-0.006	-0.004	0.015	0.011	-0.005	-0.008	-0.008	-0.002	-0.012	0.068	-0.038	0.000	0.016	0.035	0.050	0.007	-0.012	-0.005	
	Sig. (bilateral)		.910	.932	.751	.825	.918	.865	.866	.967	.797	.154	.421	.996	.731	.456	.291	.886	.803	.911	
L_EDAGTOSF	Correlación de Pearson	-0.006	1	1.000	-0.006	-0.037	-0.013	-0.004	-0.004	-0.014	-0.033	-0.016	-0.016	-0.016	-0.109	-0.051	0.048	-0.021	.132	.035	
	Sig. (bilateral)	.910		.000	.916	.501	.811	.940	.941	.806	.558	.774	.776	.776	.049	.357	.385	.701	.017	.525	
L_EDAVTAS	Correlación de Pearson	-0.004	1.000	1	-0.004	-0.009	-0.014	-0.004	-0.004	-0.009	-0.023	-0.014	-0.017	-0.010	-0.090	-0.037	0.037	-0.012	.108	.026	
	Sig. (bilateral)	.932	.000		.940	.845	.775	.933	.935	.853	.636	.773	.718	.836	.058	.436	.434	.793	.023	.591	
L_OBFEDA	Correlación de Pearson	0.015	-0.006	-0.004	1	-.261	.700	-0.005	-0.005	-0.009	-.165	.005	-.039	-.231	-.040	.001	.020	.026	-.033	-.044	
	Sig. (bilateral)	.751	.916	.940		.000	.000	.920	.918	.849	.000	.921	.418	.000	.401	.986	.669	.581	.492	.350	
L_OBLFVTAS	Correlación de Pearson	0.011	-0.037	-0.009	-.261	1	-0.002	-0.009	-0.009	-0.010	.724	.030	-.094	.851	.091	.220	-.158	.123	-.083	-.055	
	Sig. (bilateral)	.825	.501	.845	.000		.963	.843	.845	.827	.000	.531	.047	.000	.056	.000	.001	.010	.080	.247	
L_PASTEDA	Correlación de Pearson	-0.005	-0.013	-0.014	.700	-0.002	1	-.015	-.015	-0.008	.033	-0.002	-0.059	.018	.093	-0.008	.000	.075	-.109	-.083	
	Sig. (bilateral)	.918	.811	.775	.000	.963		.758	.758	.863	.483	.974	.218	.708	.051	.868	.993	.114	.021	.082	
L_RZCTE	Correlación de Pearson	-0.008	-0.004	-0.004	-0.005	-0.009	-0.015	1	1.000	.014	.052	-0.014	-0.029	-0.016	-.116	-.047	.026	-.013	-.027	-.031	
	Sig. (bilateral)	.865	.940	.933	.920	.843	.758		.000	.774	.277	.774	.547	.734	.014	.319	.583	.782	.566	.521	
L_PACIDA	Correlación de Pearson	-0.008	-0.004	-0.004	-0.005	-0.009	-0.015	1.000	1	.007	.052	-0.016	-0.028	-0.016	-.116	-.047	.026	-0.014	-.027	-.030	
	Sig. (bilateral)	.866	.941	.935	.918	.845	.758	.000		.879	.271	.734	.556	.734	.015	.321	.579	.771	.573	.531	
L_CPTTRAB	Correlación de Pearson	-0.002	-0.014	-0.009	-0.009	-0.010	-0.008	.014	.007	1	.005	.017	-0.070	-0.024	-.082	-0.017	-0.030	.027	-.023	-.039	
	Sig. (bilateral)	.967	.806	.853	.849	.827	.863	.774	.879		.911	.725	.138	.610	.083	.723	.533	.573	.636	.416	
L_ROTACART	Correlación de Pearson	-0.012	-0.033	-0.023	-.165	.724	.033	.052	.052	.005	1	-0.005	-.198	.646	-.001	.059	-.054	.036	-.119	-.082	
	Sig. (bilateral)	.797	.558	.636	.000	.000	.483	.277	.271	.911		.914	.000	.000	.985	.216	.253	.450	.012	.084	
L_ROTAINVE	Correlación de Pearson	0.068	-0.016	-0.014	.005	.030	-0.002	-0.014	-0.016	.017	-0.005	1	-.111	.157	-.011	.021	.030	.032	-.092	-.055	
	Sig. (bilateral)	.154	.774	.773	.921	.531	.974	.774	.734	.725	.914		.019	.001	.810	.662	.532	.505	.052	.244	
L_ROTAACT	Correlación de Pearson	-0.038	-0.016	-0.017	-.039	-.094	-.059	-.029	-.028	-.070	-.198	-.111	1	-.106	.041	-.069	.154	-.076	.280	.194	
	Sig. (bilateral)	.421	.776	.718	.418	.047	.218	.547	.556	.138	.000	.019		.025	.387	.144	.001	.108	.000	.000	
L_ROTAPROVE	Correlación de Pearson	.000	-0.016	-0.010	-.231	.851	.018	-0.016	-0.016	-.024	.646	.157	-.106	1	.123	.076	-.033	.062	-.054	-.040	
	Sig. (bilateral)	.996	.776	.836	.000	.000	.708	.734	.734	.610	.000	.001	.025		.009	.109	.482	.191	.256	.396	
L_NIVENDEU	Correlación de Pearson	0.016	-.109	-0.090	-0.040	.091	.093	-.116	-.116	-.082	-.001	-0.011	.041	.123	1	.414	-.144	.213	-.309	.120	
	Sig. (bilateral)	.731	.049	.058	.401	.056	.051	.014	.015	.083	.985	.810	.387	.009		.000	.002	.000	.000	.012	
L_NIVENDF	Correlación de Pearson	0.035	-0.051	-0.037	.001	.220	-0.008	-.047	-.047	-0.017	.059	.021	-.069	.076	.414	1	-.205	.236	-.246	.077	
	Sig. (bilateral)	.456	.357	.436	.986	.000	.868	.319	.321	.723	.216	.662	.144	.109	.000		.000	.000	.000	.106	
L_CONCENPCP	Correlación de Pearson	.050	.048	.037	.020	-.158	.000	.026	.026	-0.030	-.054	.030	.154	-.033	-.144	-.205	1	-.219	.156	.143	
	Sig. (bilateral)	.291	.385	.434	.669	.001	.993	.583	.579	.533	.253	.532	.001	.482	.002	.000		.000	.001	.003	
L_OBFPATR	Correlación de Pearson	.007	-0.021	-0.012	.026	.123	.075	-.013	-.014	.027	.036	.032	-.076	.062	.213	.236	-.219	1	-.044	-.504	
	Sig. (bilateral)	.886	.701	.793	.581	.010	.114	.782	.771	.573	.450	.505	.108	.191	.000	.000	.000		.354	.000	
L_ROA	Correlación de Pearson	-0.012	.132	.108	-.033	-.083	-.109	-.027	-.027	-.023	-.119	-.092	.280	-.054	-.309	-.246	.156	-.044	1	.476	
	Sig. (bilateral)	.803	.017	.023	.492	.080	.021	.566	.573	.636	.012	.052	.000	.256	.000	.000	.001	.354		.000	
L_ROE	Correlación de Pearson	-0.005	.035	.026	-.044	-.055	-.083	-.031	-.030	-.039	-.082	-.055	.194	-.040	.120	.077	.143	-.504	.476	1	
	Sig. (bilateral)	.911	.525	.591	.350	.247	.082	.521	.531	.416	.084	.244	.000	.396	.012	.106	.003	.000	.000		

** La correlación es significativa al nivel 0,01 (bilateral).

* La correlación es significante al nivel 0,05 (bilateral).